



EDCARLOS MIRANDA DE SOUZA

**TEORIA DA REGRESSÃO ISOTÔNICA E
DELINEAMENTOS SEQUENCIAIS
APLICADOS A ENSAIOS DE
DOSE-RESPOSTA**

LAVRAS-MG

2011

EDCARLOS MIRANDA DE SOUZA

**TEORIA DA REGRESSÃO ISOTÔNICA E DELINEAMENTOS
SEQUENCIAIS APLICADOS A ENSAIOS DE DOSE-RESPOSTA**

Tese apresentada à Universidade Federal de Lavras,
como parte das exigências do Programa de Pós-
Graduação em Estatística e Experimentação Agro-
pecuária, área de concentração em Estatística e
Experimentação Agropecuária, para a obtenção do
título de Doutor.

Orientador

Dr. Joel Augusto Muniz

Co-orientador

Dr. Lucas Monteiro Chaves

LAVRAS-MG

2011

**Ficha Catalográfica Preparada pela Divisão de Processos Técnicos da
Biblioteca da UFLA**

Souza, Edcarlos Miranda de.

Teoria da regressão isotônica e delineamentos sequenciais aplicados a ensaios de dose-resposta / Edcarlos Miranda de Souza. – Lavras : UFLA, 2011.

411 p. : il.

Tese (doutorado) – Universidade Federal de Lavras, 2011.

Orientador: Joel Augusto Muniz.

Bibliografia.

1. Método probit. 2. Método logit. 3. Método Up and Down. 4. Delineamento bayesiano isotônico sequencial. 5. Ensaios clínicos. I. Universidade Federal de Lavras. II. Título.

CDD – 519.536

EDCARLOS MIRANDA DE SOUZA

**TEORIA DA REGRESSÃO ISOTÔNICA E DELINEAMENTOS
SEQUENCIAIS APLICADOS A ENSAIOS DE DOSE-RESPOSTA**

Tese apresentada à Universidade Federal de Lavras,
como parte das exigências do Programa de Pós-
Graduação em Estatística e Experimentação Agro-
pecuária, área de concentração em Estatística e
Experimentação Agropecuária, para a obtenção do
título de Doutor.

APROVADA em 21 de Março de 2011.

Dr. Daniel Furtado Ferreira UFLA

Dr. Fredy Walther Castellares Caceres UFMG

Dr. Lucas Monteiro Chaves UFLA

Dr. Renato Martins Assunção UFMG

Dr. Joel Augusto Muniz
Orientador

LAVRAS-MG

2011

*Dedico este trabalho à minha família:
minha mãe, Cosma, minha esposa,
Mara, meus irmãos, Eliana, Gleide,
Emanuel, Ângelo, Luciano (in
memoriam) e aos meus filhos, Luis
Gustavo e Carlos Eduardo.*

AGRADECIMENTOS

À Deus, por estar sempre presente na minha vida, iluminando meus caminhos e tornando tudo possível.

À Universidade Federal de Lavras e ao Departamento de Ciências Exatas, juntamente com seus docentes, pela oportunidade de crescimento, aprendizado e realização profissional.

À Universidade Federal do Acre e ao Centro de Ciências Exatas e Tecnológicas, juntamente com seus docentes, pela oportunidade de crescimento, aprendizado e realização profissional, da qual me propiciaram e incentivaram.

À Coordenação de Aperfeiçoamento de Pessoal de Nível Superior - CAPES, pelo suporte financeiro concedido para a realização deste trabalho.

Ao apoio, confiança, ensinamentos, amizade e acompanhamento do orientador Dr. Joel Augusto Muniz, desde o mestrado sempre atencioso e preocupado com minhas atividades profissionais. Espero realizarmos muitos trabalhos ainda juntos. Muito obrigado, as palavras são poucas para agradecer.

Ao Professor, Amigo e co-orientador, Dr. Lucas Monteiro Chaves (São Lucas), por toda a sua disposição em colaborar, pelo apoio em todas as horas, um verdadeiro amigo, ao qual tenho muita satisfação e orgulho de termos trabalhado juntos e do qual ainda espero realizar muitos trabalhos, a curiosidade e animação sempre estão presentes em sua vida, além de um sincero agradecimento, quero deixar meu eterno abraço.

Aos professores, Renato, Daniel e Freddy, pelas valiosas sugestões.

Aos funcionários do DEX - UFLA, pela boa vontade e eficiência com que sempre me atenderam. Especialmente a Josi, Selminha, Edila e Maria.

Aos meus irmãos e amigos: Eliana, Gleide, Ângelo, Emanuel e Luciano, pessoas do qual eu amo muito, e que sempre me incentivaram. Embora Luciano

não esteja mais presente, sinto sempre a sua felicidade na realização deste sonho.

À minha amada mãe, sempre me apoiando, me amando incondicionalmente, uma verdadeira guerreira, que sempre soube superar as dificuldades com perseverança e fé em Deus. Mãe, você é o meu maior exemplo de vida, se nascer de novo fosse possível, não desejaria outra mãe que não fosse você. Um grande beijo.

Aos meus lindos filhos, meu agradecimento e desculpas pela ausência em uma fase de suas vidas que não voltará mais, espero compensar de alguma forma daqui pra frente o tempo que fiquei longe (fisicamente) de vocês. Amo muito vocês.

À minha amada, mais amada das amadas (minha Mara) , meus agradecimentos, pelo apoio, amor incondicional, carinho e felicidade que me propiciou durante toda esta jornada, amo você.

A todos os meus amigos da graduação e especialização, especialmente, Itamar, Jairo, Osvaldo, André, Aledias, Edileudo, José Roberto (Empresário), Wilson, Tony, Rodiney, Alessandro e Marquinhos, tudo começou ali, com muitas “horas alegres” de estudo.

A todos os meus amigos das vilas Santa Cecília, Albert Sampaio e Dom Moacir e a todos os membros da minha família, sobrinhos, primos e tios.

Ao meu grande amigo Mário Luiz, pelo apoio, incentivo, torcida e principalmente, pela nossa bela e sincera amizade.

À “dona” Maria, que me motivou a continuar na carreira docente com sua alegria, respeito e amizade.

À minha família em Lavras: Janice, Paulo, André, Fabrício, Ricardo, Altemir, Lucy e Gabriel. A convivência com vocês foi ótima, vocês são pessoas maravilhosas.

À galera da Bahia: Gláucia, Tânia, Ângela, Cleide, Jaílson, Jaime, Asly, Wasquez, Edmary, Regilson, Isabel, Otaviano, Márcio, Nelson, Norma, Walter, Gabriel, Wal, Vitória, Cleíton, Laine e Elma. Valeu pela torcida e alegria natural em todos vocês.

Ao meu amigo Gérson, do qual tenho muito respeito e admiração, você merece todas as bençãos que Deus tem lhe proporcionado, gostaria de agradecer pelos bons momentos de estudo e aprendizado que tive com você, um abraço irmão.

Aos meus amigos do grupo de estudo do mestrado: Patrícia, Augusto, Ana Paula, Tânia, Ricardo, Altemir, Richardson e Isabel, pelo companheirismo, amizade e inúmeros momentos de alegria.

Aos meus amigos do grupo de estudo do doutorado: Gérson, Ademária, Wanderley, Naje, Ana Lúcia, Moysés, Augusto, Tânia, Ana Paula, Adriano, Renata e Devanil, foi ótimo aprender com vocês um pouco de estatística e outros valores que não aparecem em livros.

Ao casal de amigos, Eustáquio e Francisca (Franzinha), por todos os momentos alegres, estímulo, sugestões e a bela amizade.

Ao meu amigo Altemir, companheiro de todas as batalhas, sempre presente, torcendo, participando e ajudando, valeu mesmo por tudo, principalmente por nossa amizade.

Aos amigos dos quais não tive a honra de cursar alguma disciplina, mas que torceram e acreditaram muito no meu trabalho: Diogo e Gi, Talita, Carol, Luzia, Tiago, Ênio, Elson, Danilo, Leandro, Maíra, Andressa, Juliano, Hernani, Lucas, Jair, Púpico e Dani.

A todos que, de alguma forma, contribuíram para a realização deste trabalho, meu muito obrigado!!!

*"Insanity is doing the same things over and over
again and expecting the different results."*

Albert Einstein.

RESUMO

Diversos métodos para análise de dados de sensibilidade, obtidos em ensaios de dose-resposta, tem como objetivo estimar uma dose alvo que produzirá determinado efeito em Γ por cento de uma população. Inicialmente, os métodos tinham como alvo estudar a dose para o qual $\Gamma = 50\%$, comumente chamada de DL_{50} (Dose Letal a 50%). Métodos como o *probit*, *logit* e *up and down*, estimam com eficiência tal dose, sendo que, o método *up and down*, baseado na Teoria dos Delineamentos Sequenciais, tem a vantagem de reduzir o tamanho da amostra. Delineamentos alternativos derivados do *up and down*, tais como o Delineamento da Moeda Viciada e o Delineamento de Narayana Modificado, são eficazes para a estimação de qualquer quantil da curva de Tolerâncias que seja de interesse prático em Ensaios Clínicos de Fase I ou Ensaios de Toxicidade Aguda. Como a curva de dose-resposta é geralmente crescente, um método estatístico adequado para o processo de estimação é a regressão isotônica. Um dos objetivos deste trabalho é propor novos estimadores obtidos a partir da teoria da regressão isotônica para serem aplicados a delineamentos sequenciais, bem como novos estimadores aplicados à teoria de *probit* e *logit*. Também é proposto um delineamento bayesiano isotônico sequencial. Para avaliar o desempenho dos estimadores em relação aos vários delineamentos e em relação ao delineamento proposto foi utilizada simulação Monte Carlo, calculando-se o EQM e viés destes estimadores para diversos tamanhos amostrais. Os resultados obtidos indicam a viabilidade do uso prático tanto dos novos estimadores como do delineamento proposto.

Palavras-chave: Regressão Isotônica. Delineamentos Sequenciais. Método *Probit*. Método *Logit*. Método *Up and Down*. Delineamento Bayesiano Isotônico Sequencial.

ABSTRACT

The goal of many analysis methods for sensibility data, obtained in dose-response designs, is to estimate the dose-target that will produce a predefined result on Γ percent of the population. Initially the usual was $\Gamma = 0.5$ corresponding to the so called 50% lethal dose, LD_{50} . Methods like *probit*, *logit* and *Up and Down* estimate efficiently such dose and, among them, the *Up and Down* method, based on the theory of sequential designs, has the advantage of reduced sample size. When dealing with estimation of quantile of the tolerance curve, as in Phase I clinical essays or acute toxicity tests, some *Up and Down* based design, like the *biased coin design* and the *modified Narayana design*, has shown to be efficient. As the dose-response is usually a increasing curve, the appropriate statistical method for estimation is the *isotonic regression*. One of the objectives of this work is to propose new estimators, based on the theory of isotonic regression, to be used when dealing with sequential designs as well as estimators applied to *probit* and *logit* theory. It is also proposed a sequential isotonic Bayesian design. To evaluate the accomplishment of the estimator we used Monte Carlo Method to evaluate MSE and bias for different sample sizes. Results has shown feasibility for practical use of the new estimators and the proposed design.

Key-words: Isotonic Regression. Sequential Design. *Probit* Method. *Logit* Method. *Up and Down* Method . Sequential Isotonic Bayesian Design.

SUMÁRIO

1	INTRODUÇÃO	14
2	REFERENCIAL TEÓRICO	16
2.1	Regressão isotônica	16
2.1.1	A história da regressão isotônica	16
2.1.2	Regressão linear, não linear e regressão isotônica	17
2.1.3	Visão algébrica da regressão isotônica	20
2.1.4	Aspectos geométricos da regressão isotônica	38
2.1.5	Análise convexa	50
2.1.5.1	Estimação por máxima verossimilhança para uma família de distribuições binomiais com parâmetros ordenados	58
2.1.5.2	Estimação por máxima verossimilhança para uma família de distribuições geométricas com parâmetros ordenados	62
2.1.5.3	Estimação por máxima verossimilhança para uma família de distribuições poisson com parâmetros ordenados	66
2.1.5.4	Estimação por máxima verossimilhança para uma família de distribuições gama com parâmetros ordenados	69
2.1.6	Estimação por máxima verossimilhança de parâmetros ordenados na família exponencial	72
2.1.7	O problema das estimativas de componentes de variâncias negativas	81
2.1.7.1	Estimação de variâncias ordenadas de distribuições normais com médias conhecidas	81
2.1.7.2	Estimativas de componentes de variâncias negativas	83
2.1.7.3	O grafo de um delineamento	97
2.1.8	Estimação bayesiana e regressão isotônica	100
2.1.9	O cálculo da regressão isotônica para ordem simples utilizando o algoritmo PAVA	105
2.1.9.1	O diagrama de somas acumuladas	105
2.1.9.2	O maior minorante convexo	109
2.1.9.3	Um algoritmo para o cálculo da regressão isotônica	112
2.2	Análise de dados de sensibilidade e delineamentos sequenciais .	115
2.2.1	Análise de dados de sensibilidade	115
2.2.2	O método <i>probit</i>	118
2.2.2.1	O caso dos 0% e 100% de sucessos em análise <i>probit</i>	129
2.2.3	O método <i>logit</i>	131

2.2.3.1	O caso dos 0% e 100% de sucessos em análise <i>logit</i>	138
2.2.4	Delineamentos sequenciais	141
2.2.4.1	O método <i>up and down</i>	141
2.2.4.2	O delineamento <i>up and down</i> da moeda viciada	156
2.2.5	Ensaio clínico de fase I no tratamento de câncer	160
2.2.5.1	O delineamento tradicional ou 3+3	162
2.2.5.2	O delineamento <i>K-in-a-row</i>	164
2.2.5.3	O delineamento <i>up and down</i> de média móvel	165
2.2.5.4	O delineamento <i>up and down</i> de Narayana	165
2.2.5.5	O delineamento <i>up and down</i> de Narayana modificado	166
2.2.5.6	O delineamento isotônico de Leung e Wang	166
2.2.6	Delineamentos <i>up and down</i> generalizados	171
2.2.7	Processo de estimação	175
2.2.7.1	Estimadores paramétricos	177
2.2.7.2	Estimadores não paramétricos	177
3	METODOLOGIA	180
3.1	Estimadores propostos	181
3.1.1	Estimadores paramétricos	181
3.1.1.1	Estimador <i>probit</i> isotonizado	182
3.1.1.2	Estimadores <i>logit</i> isotonizado	182
3.1.2	Estimadores não paramétricos	183
3.1.2.1	Estimador moda da frequência empírica isotonizada	185
3.1.2.2	Estimador por interpolação da regressão unimodal	185
3.2	Delineamento proposto	187
3.3	Estudo de simulação	190
3.3.1	Comparação dos estimadores	191
3.3.1.1	Os delineamentos	195
3.3.1.2	Os estimadores	196
3.3.1.3	A avaliação dos estimadores	197
3.3.2	Avaliação do delineamento proposto	197
3.3.3	Software	198
4	RESULTADOS E DISCUSSÃO	202
4.1	Estimadores	202
4.1.1	Cenário paramétrico - distribuição normal	202
4.1.2	Cenário paramétrico - distribuição logística	212
4.1.3	Estimadores não paramétricos	222
4.1.4	Delineamento bayesiano isotônico sequencial - DBIS	233
4.1.4.1	Primeiro conjunto das doses	233
4.1.4.2	Segundo conjunto das doses	237

5	CONCLUSÃO	241
	REFERÊNCIAS	243
	APÊNDICE	247

1 INTRODUÇÃO

Existem diversas situações experimentais em que se quer medir em uma população a eficácia de uma determinada droga. Um procedimento comum em tais experimentos é a aplicação da droga em vários níveis, observando a resposta à mesma por contagem do número de indivíduos que forneceram resposta positiva. Esses ensaios são chamados de dose-resposta. Assim, um ensaio do tipo dose-resposta é aquele em que uma determinada droga é administrada em k diferentes doses (níveis), d_1, d_2, \dots, d_k em, respectivamente, m_1, m_2, \dots, m_k indivíduos, obtendo-se como resposta, após um período especificado, y_1, y_2, \dots, y_k indivíduos que mudam de estado (ocorrência de um sucesso, por exemplo, morte).

Nos ensaios de dose-resposta, a tolerância é comumente modelada pela distribuição normal ou pela distribuição logística e estes métodos são denominados *probit* e *logit*. Em geral, nos métodos *probit* e *logit* é necessário que muitos indivíduos sejam expostos a doses elevadas ou muito baixas, o que gera problemas práticos e éticos. Métodos alternativos, que em média expõem um número menor de indivíduos a doses elevadas ou muito baixas, são derivados da teoria de passeios aleatórios, denominados métodos *up and down*.

Os métodos *probit*, *logit* e *up and down* estimam com eficiência a dose letal de um indivíduo a 50% (DL_{50}). Entretanto, para proporções próximas das caudas das distribuições tais como DL_{10} ou DL_{33} (comuns em ensaios oncológicos), estes métodos se tornam menos eficientes. Diversos delineamentos sequenciais derivados do método *up and down* tem sido propostos na literatura, visando a uma melhor estimação para qualquer quantil da curva de tolerâncias.

Os tipos de situações que se utilizam dos ensaios de dose-resposta são bem diversificado. Dentre elas, destacam-se os experimentos com inseticidas,

ensaios oncológicos ou de toxicidade aguda e até mesmo experimentos com explosivos. Em tais situações, espera-se naturalmente que a medida em que a dose de uma determinada droga é aumentada, a probabilidade de resposta à mesma também aumenta, e que um procedimento de estimação deve contemplar este fato. Dada uma curva, um procedimento matemático que obtém a curva crescente mais próxima possível no sentido dos quadrados mínimos é a regressão isotônica. Esta teoria é uma ferramenta natural para se estudar ensaios de dose-resposta, no entanto, a literatura apresenta poucas aplicações. Este trabalho visa utilizar de forma essencial os resultados da teoria da regressão isotônica na teoria dos ensaios de dose-resposta.

Assim, este trabalho foi realizado com o objetivo de desenvolver estimadores para quantis que sejam utilizados em ensaios de dose-resposta, fazendo o uso da teoria da regressão isotônica, bem como dos delineamentos sequenciais, que sejam úteis tanto em ensaios em que as probabilidades de toxicidade sejam modeladas por uma distribuição normal, Logística ou outra qualquer pertencente à família exponencial. Objetivou-se também a proposição de um delineamento sequencial para ensaios de dose-resposta que fizesse uso da metodologia bayesiana, bem como da teoria da regressão isotônica. O desempenho dos procedimentos propostos foi avaliado utilizando-se do Erro Quadrático Médio e Viés dos estimadores, por meio de simulação Monte Carlo.

O referencial teórico apresenta inicialmente os principais resultados sobre regressão isotônica. Como a literatura disponível é bastante restrita nesta área, as passagens matemáticas tanto nos resultados teóricos, quanto nas aplicações, foram desenvolvidas em detalhes, podendo-se afirmar que este trabalho é uma contribuição didática em língua portuguesa para a teoria da regressão isotônica.

2 REFERENCIAL TEÓRICO

2.1 Regressão isotônica

O problema estatístico, objetivo central neste estudo, que se relaciona com a teoria da regressão isotônica é a estimação em espaços paramétricos restritos. Antes de iniciar as definições formais e principais teoremas e propriedades ligados a esta teoria, bem como sua importância em problemas práticos, convém fazer de forma breve uma revisão histórica de tal tema, desde sua origem por volta dos anos 50 até os dias atuais.

2.1.1 A história da regressão isotônica

Segundo Eeden (1996), a teoria surgiu nos anos 50 com o problema em que duas probabilidades θ_1 e θ_2 , com $\theta_1 \leq \theta_2$, precisavam ser estimadas. A estimação por Máxima Verossimilhança foi utilizada para este propósito. Os estimadores de Máxima Verossimilhança irrestrito mostraram-se ser inadmissíveis pelo critério de erro quadrático. Isto conduziu a estimação por Máxima Verossimilhança em espaços paramétricos restritos.

Diversos trabalhos surgiram de forma independente na mesma década, dentre os quais destaca-se o trabalho de Constance van Eeden em sua dissertação defendida na Universidade de Amsterdam sob o título *Testing and Estimating Ordered Parameters of Probability Distributions*, e posteriormente em quatro artigos publicados em 1956 – 1957 em *Indagationes Mathematicae* (LEEUEW; HORNIK; MAIR, 2009).

Paralelamente, como em muitas vezes acontece em ciência, outros trabalhos com o objetivo de solucionar os mesmos problemas foram surgindo ao mesmo tempo, dando origem a diversos artigos científicos com o foco em estimação de

parâmetros ordenados, dentre eles destacam-se os trabalhos de Ayer et al. (1955), Bartholomew (1959a, 1959b) e Brunk (1955).

Um compêndio com toda esta teoria interligada aparece no excelente livro de Barlow et al. (1972), também conhecido como o livro dos quatro *B's*. A próxima grande contribuição foi o livro de Robertson, Wright e Dykstra (1988), que apresentou os resultados fornecidos em Barlow et al. (1972) acrescidos de algumas contribuições posteriores.

O problema da regressão isotônica geral é também um problema de programação quadrática, motivando o surgimento de diversos trabalhos nas áreas ligadas à informática, desde a apresentação de algoritmos mais eficientes até a apresentação de programas em softwares estatísticos convencionais.

Recentemente, por volta de 2002, esta teoria passa a ter também conexão com experimentos de dose-resposta que se utilizam de delineamentos sequenciais, com trabalhos voltados para a área de ensaios clínicos, em que o tamanho da amostra geralmente é pequeno. Tais trabalhos aparecem do ponto de vista prático, utilizando a teoria apenas como ferramenta.

O tema será apresentado com um certo rigor matemático e com uma preocupação de ser de certa forma autosuficiente, uma vez que a principal referência nessa área, o livro dos quatro *B's*, apresenta a maioria dos resultados de forma muito sucinta e portanto de difícil leitura.

2.1.2 Regressão linear, não linear e regressão isotônica

A análise de regressão consiste na realização de um estudo estatístico com o objetivo de verificar a existência de relação funcional significativa entre variáveis. Essa teoria está ligada diretamente com os conceitos de esperança condicional e mínimos quadrados. A função de regressão de uma variável aleatória

Y em outra X é a esperança condicional $\mu(x) = E[Y|X = x]$ e deve fornecer o melhor ajuste para a distribuição de Y como uma função de X no sentido dos mínimos quadrados. A variável X é denominada variável independente e Y de variável dependente.

A análise de regressão pode ser realizada usando modelos estatísticos paramétricos ou não paramétricos. Os modelos paramétricos consistem basicamente na obtenção de uma equação, com parâmetros a serem estimados, que explica a variação da variável dependente pela variação dos níveis da(s) variável(is) independente(s). Os modelos não paramétricos são aqueles onde nenhuma pressuposição é feita sobre a distribuição conjunta de X e Y .

Os modelos paramétricos podem ser classificados com relação aos parâmetros em duas categorias: lineares e não lineares. Segundo Draper e Smith (1998), um modelo é linear nos parâmetros se qualquer uma das derivadas parciais desse modelo em relação aos parâmetros não dependem especificamente de nenhum deles. A definição de um modelo não linear, apresentada pela maioria dos autores, dentre eles Bates e Watts (1988), é que pelo menos uma derivada parcial da variável dependente, com relação a algum parâmetro presente no modelo, depende de algum parâmetro.

De modo geral, a análise de regressão em qualquer situação está direcionada para as inferências sobre $\mu(x)$, estimado-a, ou testando hipóteses sobre a mesma. Comumente, em Estatística, o termo x é um número real ou um vetor de números reais, no entanto, este termo terá aqui um sentido mais amplo, podendo representar um determinado atributo, por exemplo, em Estatística Experimental x poderá representar um efeito de um determinado tratamento, e as esperanças condicionais a serem comparadas nesta situação, serão do tipo $\mu(1) = E[Y|\text{tratamento 1}]$ e $\mu(2) = E[Y|\text{tratamento 2}]$. Em geral, Y será uma variável

aleatória cujo valores serão números reais, mas X poderá ser ou não aleatório, seus valores poderão ser ou não números reais ou vetores de números reais.

O interesse principal, neste trabalho, são os modelos de regressão que requerem, por pressupostos teóricos, uma ordenação parcial ou total dos números correspondentes $\mu(x)$. Tal fato ocorre em diversas situações experimentais, como por exemplo, em experimentos de dose-resposta utilizados em bioensaios, a variável independente x pode ser a dose de um determinado inseticida e a probabilidade de uma resposta (por exemplo, a morte do inseto) é $\mu(x)$. Naturalmente é de se esperar que se a dose x aumenta, a probabilidade de morte dos insetos $\mu(x)$ não diminuirá, ou seja, $\mu(x)$ deve ser uma função não decrescente de x .

Situações experimentais em que a ordenação na variável regressora implica em uma ordenação na variável resposta são das mais diversas possíveis. Um outro exemplo, bem diferente dos bioensaios, pode ser encontrado no trabalho de McArthur et al. (2003), que investigaram a influência das algas marinhas sobre a quantidade de espécies marinhas disponíveis para a pesca no Sul da Austrália. Existem diversos pressupostos biológicos para acreditar que o declínio das algas marinhas teriam consequências negativas sobre a abundância e composição dessas espécies. Em seu estudo eles consideraram duas variáveis, a *captura* que era definida como o peso vivo de espécies pescadas em um local espacialmente definido e o *esforço* definido como sendo o número de dias de barco, indiferentemente do número de pescadores em cada barco, gasto para se obter a referida captura. Já era sabido por experiência dos pescadores que a *captura* era uma função crescente do *esforço*, e de posse destas informações eles estudaram a influência das algas marinhas nas áreas de pesca considerando a informação da ordenação existente.

Na análise de regressão usual não se usa uma informação a priori sobre uma possível relação de ordem na variável resposta. Na teoria da regressão isotô-

nica essa informação é utilizada de forma essencial, conseqüentemente é de se esperar que a eficiência da análise seja aumentada quando faz-se o uso da mesma.

A Regressão Isotônica faz o uso desta informação em um contexto bastante geral, considerando até mesmo situações em que a ordenação entre os elementos é parcial, e sua teoria apresenta uma geometria interessante do ponto de vista matemático.

Parte deste trabalho está baseado no livro clássico *Statistical Inference under Order Restrictions: The Theory and Application of Isotonic Regression*, de Barlow et al. (1972). Entretanto, as demonstrações e a forma de exposição que serão dadas a seguir, no que se refere ao assunto regressão isotônica, podem ser entendidas como uma contribuição didática à teoria. As demonstrações foram detalhadas e ampliadas, com o objetivo de elucidar as passagens mais complicadas do livro. Há uma opção de ordenação do conteúdo ligeiramente diferente do livro por motivos de ordem didática. Espera-se, com isso, que este trabalho seja uma fonte onde o leitor possa entender com propriedade os principais conceitos ligados à regressão isotônica, bem como a interpretação e utilização dos mesmos em áreas práticas do conhecimento científico.

As citações serão usualmente omitidas, uma vez que estes resultados podem ser encontrados nos livros de Barlow et al. (1972) e Robertson, Wright e Dykstra (1988). Entretanto, como já mencionado, a sequência da apresentação, os detalhes das explicações, os gráficos e figuras, são parte inerente deste trabalho e de certa forma, uma contribuição à área.

2.1.3 Visão algébrica da regressão isotônica

Um problema comum em regressão está relacionado a um conjunto de valores pré-fixados $X = \{x_1, x_2, \dots, x_k\}$, onde $x_1 < x_2 < \dots < x_k$. Para

$i = 1, 2, \dots, k$ os valores $y_j(x_i)$, $j = 1, 2, \dots, m_i$ são medidas de alguma quantidade de interesse (variável resposta). Isso é, para $x \in X$, $y_1(x), \dots, y_{m_i}(x)$ são observações de alguma distribuição (No caso do exemplo de bioensaios, os valores em X seriam as doses de inseticida, os valores $y_j(x_i)$ a proporção de insetos mortos para a dose x_i , m_i o número de repetições para cada dose, x_i).

Seja $\mu = \mu(x) = E[Y \mid X = x]$, a média da distribuição da variável Y dado $x \in X$. Se μ é modelada como uma função linear nos parâmetros, é razoável estimar μ por regressão linear. Nesse caso, a solução do problema é ajustar uma função linear nos parâmetros aos dados no sentido dos mínimos quadrados, isto é, minimizar

$$\sum_{x \in X} \sum_{j=1}^{m(x)} [y_j(x) - f(x)]^2 \quad (2.1)$$

na classe das funções f , em que f é uma função linear nos parâmetros. Essa teoria é bem discutida em Draper e Smith (1998). Para o caso de um modelo de regressão não linear nos parâmetros ser adotado para $f(x)$, uma referência é Bates e Watts (1988).

No contexto não paramétrico, se quer minimizar

$$\sum_{x \in X} \sum_{j=1}^{m(x)} [y_j(x) - f(x)]^2 \quad (2.2)$$

em que f é uma função qualquer definida em X . A solução deste problema é dada por $\hat{\mu}(x) = \bar{y}(x)$.

Em certas situações é razoável supor que μ seja uma função não decrescente de x . Uma estimativa de μ poderia ser obtida minimizando a soma de quadrados (2.1) na classe das funções não decrescente f em X , ou seja, na classe

das funções f tais que $x_i < x_j$ implica em $f(x_i) \leq f(x_j)$.

Barlow et al. (1972) usam as seguintes definições:

Definição 2.1 *Seja X um conjunto de números $\{x_1, \dots, x_k\}$. Uma função de valor real f em X é isotônica (ou não decrescente) se para $x, y \in X$, $x < y$ implicar em $f(x) \leq f(y)$.*

Definição 2.2 *Sejam X como na Definição 2.1, g uma função em X e w uma dada função positiva em X . Uma função isotônica g^* em X é uma regressão isotônica de g com pesos w , se ela minimiza, na classe das funções isotônicas f em X , a soma*

$$\sum_{x \in X} [g(x) - f(x)]^2 w(x). \quad (2.3)$$

De posse da Definição 2.2 surgem algumas perguntas: a primeira delas é se a regressão isotônica existe e uma outra é, existindo, será ela única? A resposta a estas perguntas é sim, e tal fato será demonstrado em teoremas posteriores. O problema de se encontrar a regressão isotônica é então o seguinte: encontrar uma função não decrescente que minimize a soma de quadrados dos erros 2.3, ou seja, dentre todas as funções não decrescentes possíveis, busca-se uma que torna a soma de quadrados dos erros a menor possível.

Nas definições 2.1 e 2.2, foi utilizado o conceito de ordem dos números reais, chamada de ordem simples, sem que este termo fosse explicitamente definido, isso ocorreu por ser esta ordem a natural que é trabalhada comumente nos problemas de regressão. No entanto, os conceitos e resultados teóricos que são apresentados com a regressão isotônica são bem mais gerais e requerem as seguintes definições apresentadas em Lipschutz (1972):

Definição 2.3 Uma relação binária “ \preceq ” em X estabelece uma ordem simples em X se:

1. Ela é reflexiva: $x \preceq x$ para todo $x \in X$;
2. Ela é transitiva: $x, y, z \in X, x \preceq y, y \preceq z \Rightarrow x \preceq z$;
3. Ela é anti-simétrica: $x, y \in X, x \preceq y, y \preceq x \Rightarrow x = y$;
4. Quaisquer dois elementos são comparáveis: $x, y \in X \Rightarrow$ ou $x \preceq y$ ou $y \preceq x$.

Definição 2.4 Uma relação binária “ \preceq ” em X estabelece uma ordem parcial em X se ela satisfaz as condições 1, 2 e 3 da Definição 2.3, ou seja, se ela é reflexiva, transitiva e anti-simétrica.

Definição 2.5 Uma relação binária “ \preceq ” em X estabelece uma quase ordem em X se ela satisfaz as condições 1 e 2 da Definição 2.3, ou seja, se ela é reflexiva e transitiva.

Como pode ser observado, toda ordem simples é uma ordem parcial e toda ordem parcial é uma quase ordem. Assim, se a teoria for feita sobre uma quase ordem, ela será mais geral do que se fosse feita sobre uma ordem simples, pois englobará todas elas. Para elucidar as Definições 2.3, 2.4 e 2.5, considere o Exemplo 2.1:

Exemplo 2.1 Considere o conjunto $X = \{2,3,4\}$ e a relação \preceq : $x \preceq y$ se x divide y . Observe que esta relação é reflexiva e transitiva, pois x divide x , e se x divide y e y divide z então x divide z , para quaisquer $x, y, z \in X$. Ela é anti-simétrica, pois se x divide y e y divide x então $x = y$. No entanto, existem elementos não comparáveis, 2 não é comparável com 3, pois 2 não divide 3 e 3 não divide 2. Portanto, a relação \preceq é uma ordem parcial.

Considerando o conjunto X e a relação de ordem \preceq como no Exemplo 2.1, pode-se definir uma função f de X em \mathbb{R} , definida por $f(x) = 2x$, observe que para os elementos comparáveis tem-se uma relação de ordem mantida nas imagens de f , por exemplo, $2 \preceq 4$ e $f(2) = 4 \leq 8 = f(4)$, e isso vale para quaisquer dois elementos comparáveis, ou seja, nos elementos que são comparáveis vale uma certa monotonicidade, o que motiva uma definição mais geral de função isotônica.

Definição 2.6 *Uma função de valor real $f : X \rightarrow \mathbb{R}$ é isotônica com respeito a uma quase-ordem \preceq em X se $x, y \in X$, $x \preceq y$ implicar em $f(x) \leq f(y)$.*

Naturalmente, de posse da Definição 2.6, pode-se pensar no problema da regressão isotônica geral sobre um conjunto quase-ordenado, e a definição para esta situação é similar à apresentada pela Definição 2.2.

Definição 2.7 *Se g é uma função em um conjunto quase-ordenado X e w uma dada função positiva sobre X . Uma função isotônica g^* sobre X é uma regressão isotônica de g com pesos w se ela minimiza, na classe das funções isotônicas em X a soma*

$$\sum_{x \in X} [g(x) - f(x)]^2 w(x). \quad (2.4)$$

Estes conceitos podem ser aplicados também para o caso onde f é uma função antitônica (Definição 2.8).

Definição 2.8 *Seja X um conjunto finito $\{x_1, \dots, x_k\}$ com a ordem simples $x_1 \prec x_2 \prec \dots \prec x_k$. Uma função de valor real f sobre X é antitônica se para $x, y \in X$, $x \prec y$ implicar em $f(x) \geq f(y)$.*

Note que, com as mudanças naturais, todas as definições e consequentemente todos os teoremas para a regressão isotônica também serão válidos para a

regressão antitônica. O termo “regressão monótona” foi utilizado inicialmente, no entanto, Barlow et al. (1972) sugeriram o termo “isotônico” que significa “preservando a ordem” e é preferido à monotonicidade que pode ser interpretado como “preservando a ordem” ou “ordem reversa”.

Alguns resultados algébricos serão utilizados na demonstração de alguns teoremas posteriores, um deles está na seguinte notação: se a e b são dois números reais, $a \wedge b$ e $a \vee b$ denotam respectivamente o menor e o maior entre eles, ou seja:

$$a \wedge b = \min(a, b)$$

$$a \vee b = \max(a, b)$$

Se f e g são duas funções sobre X , as funções $f \wedge g$ e $f \vee g$ são definidas por:

$$(f \wedge g)(x) = f(x) \wedge g(x), x \in X$$

$$(f \vee g)(x) = f(x) \vee g(x), x \in X.$$

A Figura 1 ilustra esta situação para um caso onde f e g não tem nenhuma restrição, a Figura 2 para o caso onde f e g são funções isotônicas. Observe que quando as funções não são isotônicas, as funções $f \vee g$ e $f \wedge g$ também podem não ser isotônicas (Figura 1), no entanto, no caso onde f e g são isotônicas, as funções $f \vee g$ e $f \wedge g$ também são (Figura 2). Este resultado será utilizado na demonstração de um teorema que implicará na existência da regressão isotônica. Este fato está exposto formalmente no teorema 2.1.

Teorema 2.1 *Sejam f e g duas funções de valores reais sobre um conjunto X . Então, se f e g são isotônicas sobre X , as funções $(f \wedge g)(x)$ e $(f \vee g)(x)$ também serão.*

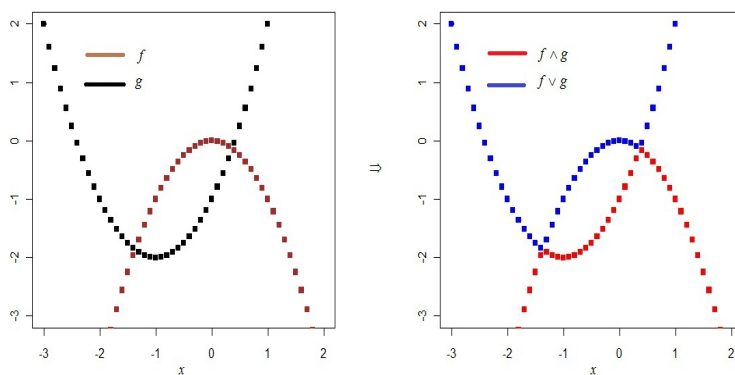


Figura 1 Representação para $f \vee g$ e $f \wedge g$, com f e g funções quaisquer

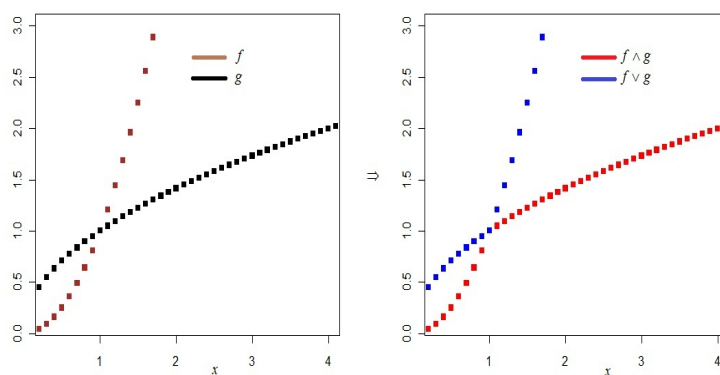


Figura 2 Representação para $f \vee g$ e $f \wedge g$, com f e g funções isotônicas

Demonstração: Sejam $x_1, x_2 \in X$, com $x_1 \preceq x_2$, então, por hipótese (f e g são isotônicas) tem-se $f(x_1) \leq f(x_2)$ e $g(x_1) \leq g(x_2)$. Daí,

$$(f \wedge g)(x_1) = \min\{f(x_1), g(x_1)\} \leq f(x_1) \leq f(x_2)$$

e

$$(f \wedge g)(x_1) = \min\{f(x_1), g(x_1)\} \leq g(x_1) \leq g(x_2).$$

Portanto,

$$(f \wedge g)(x_1) = \min\{f(x_1), g(x_1)\} \leq \min\{f(x_2), g(x_2)\} = (f \wedge g)(x_2),$$

o que implica que $f \wedge g$ é isotônica. A demonstração para $f \vee g$ é análoga.

■

O Teorema 2.2 afirma que se a regressão isotônica g^* de g existe, e se g está limitada acima e abaixo por funções isotônicas, então g^* tem os mesmos limites. Este teorema é muito parecido com o *Teorema do Sanduíche* ou *do Confronto* utilizado no cálculo diferencial e integral.

Teorema 2.2 Se g_1 e g_2 são funções isotônicas sobre X tais que $g_1(x) \leq g(x) \leq g_2(x)$, para todo $x \in X$, e se g^* é uma regressão isotônica de g , então também $g_1(x) \leq g^*(x) \leq g_2(x)$ para $x \in X$. Em particular, se “ a ” e “ b ” são constantes tais que $a \leq g(x) \leq b$ para $x \in X$, então $a \leq g^*(x) \leq b$ para $x \in X$.

Demonstração: Sejam f uma função isotônica qualquer e $h = f \vee g_1$. Pelo teorema 2.1, h é também isotônica. Se $f(x) \geq g_1(x)$ para um particular $x \in X$, então $h(x) = f(x)$ de modo que $g(x) - f(x) = g(x) - h(x)$. Se $f(x) < g_1(x)$ então $0 \leq g(x) - h(x) = g(x) - g_1(x) < g(x) - f(x)$. Assim,

para todo $x \in X$,

$$[g(x) - h(x)]^2 \leq [g(x) - f(x)]^2.$$

Então,

$$\sum_{x \in X} [g(x) - h(x)]^2 w(x) \leq \sum_{x \in X} [g(x) - f(x)]^2 w(x),$$

com a desigualdade estrita se $f(x) < g_1(x)$ para algum $x \in X$. De modo análogo, pode-se definir $l = f \wedge g_2$ e concluir que $l(x) \leq g_2(x)$ (Figura 3):

$$\sum_{x \in X} [g(x) - l(x)]^2 w(x) \leq \sum_{x \in X} [g(x) - f(x)]^2 w(x),$$

ou seja, existem funções isotônicas $h(x)$ e $l(x)$ com $h(x) \geq g_1(x)$ e uma $l(x) \leq g_2(x)$ que reduzem a soma de quadrados

$$\sum_{x \in X} [g(x) - f(x)]^2 w(x).$$

Portanto, como g^* minimiza a soma de quadrados, g^* satisfaz $g^*(x) \geq g_1(x)$ e $g^*(x) \leq g_2(x)$ para todo $x \in X$.

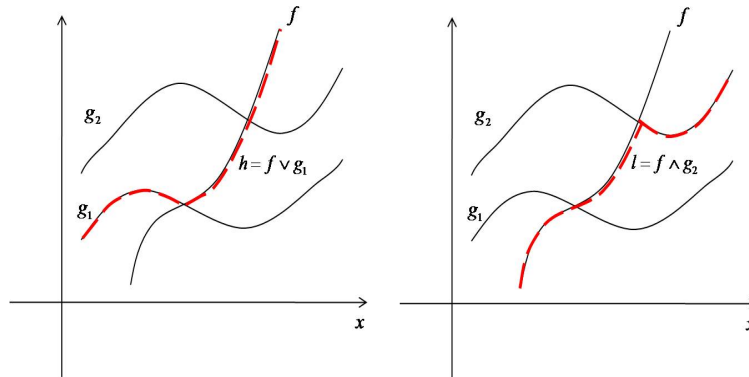


Figura 3 Representação das funções h e l para a demonstração do Teorema 2.2



Os resultados demonstrados até aqui utilizaram-se do fato de uma possível existência da regressão isotônica g^* , o Teorema 2.3 afirmará que uma regressão isotônica g^* existe, embora não deixará claro ainda como encontrá-la.

Teorema 2.3 *Uma regressão isotônica de g como na Definição 2.7 existe.*

Demonstração: *Sejam “ a ” e “ b ” os limites de g , ou seja, $a \leq g(x) \leq b$ para $x \in X$. Da prova do Teorema 2.2 segue que, se f é qualquer função isotônica sobre X , existe uma função isotônica h compreendida entre “ a ” e “ b ” tais que:*

$$\sum_{x \in X} [g(x) - h(x)]^2 w(x) \leq \sum_{x \in X} [g(x) - f(x)]^2 w(x).$$

Assim, é suficiente mostrar que existe função isotônica minimizando a soma:

$$S_X = \sum_{x \in X} [g(x) - f(x)]^2 w(x)$$

sujeito a restrição $a \leq f(x) \leq b$ para todo $x \in X$.

Seja k o número de elementos do conjunto X , em \mathbb{R}^k defina o subconjunto $\mathfrak{B} = \{(y_1, y_2, \dots, y_k) \mid a \leq y_i \leq b \text{ e } y_1 \leq y_2 \leq \dots \leq y_k\}$, \mathfrak{B} é um subconjunto fechado e limitado de \mathbb{R}^k . A soma S_X pode ser considerada como uma função contínua de k argumentos reais $f(x)$, ou seja, S_X é uma função definida de \mathbb{R}^k em \mathbb{R} da forma:

$$\begin{aligned} S_X(y_1, \dots, y_k) &= \sum_{x \in X} [g(x) - f(x)]^2 w(x) \\ &= [g(x_1) - f(x_1)]^2 w(x_1) + \dots + [g(x_k) - f(x_k)]^2 w(x_k), \end{aligned}$$

em que $f(x_i) = y_i$. Esta função tem um mínimo em \mathfrak{B} , pois o conjunto \mathfrak{B} é

limitado e fechado (LIMA, 1981). Um ponto de mínimo (y_1^*, \dots, y_k^*) define uma regressão isotônica $g^*(x_i) = y_i^*$

■

Os próximos teoremas que serão enunciados e demonstrados servirão, dentre outras coisas, para demonstrar a unicidade da regressão isotônica. Além disso, o Teorema 2.4 mostrará uma relação entre as funções isotônicas e as funções convexas, essa relação será discutida com mais detalhes na próxima seção, onde será abordado o aspecto geométrico da regressão isotônica.

Teorema 2.4 *Se f_1 e f_2 são funções isotônicas sobre X e α é um número real tal que $0 \leq \alpha \leq 1$, então $\alpha f_1 + (1 - \alpha)f_2$ é isotônica.*

Demonstração: *Sejam $x_i \preceq x_j$ elementos em X , então, por hipótese:*

$$f_1(x_i) \leq f_1(x_j) \Rightarrow \alpha f_1(x_i) \leq \alpha f_1(x_j)$$

$$f_2(x_i) \leq f_2(x_j) \Rightarrow (1 - \alpha)f_2(x_i) \leq (1 - \alpha)f_2(x_j).$$

Somando as desigualdades acima, tem-se:

$$\alpha f_1(x_i) + (1 - \alpha)f_2(x_i) \leq \alpha f_1(x_j) + (1 - \alpha)f_2(x_j)$$

$$\Rightarrow (\alpha f_1 + (1 - \alpha)f_2)(x_i) \leq (\alpha f_1 + (1 - \alpha)f_2)(x_j).$$

Portanto, $\alpha f_1 + (1 - \alpha)f_2$ é uma função isotônica sobre X .

■

Teorema 2.5 *Uma regressão isotônica g^* de g com pesos w sobre X munido de*

uma quase ordem \preceq satisfaz:

$$\sum_{x \in X} [g(x) - g^*(x)] [g^*(x) - f(x)] w(x) \geq 0 \quad (2.5)$$

para toda função isotônica f em X .

Demonstração: De acordo com o Teorema 2.4, se α é um número real tal que $0 \leq \alpha \leq 1$, então

$$(1 - \alpha)g^* + \alpha f$$

é também isotônica. Então, a expressão

$$S = \sum_{x \in X} \{g(x) - [(1 - \alpha)g^*(x) + \alpha f(x)]\}^2 w(x)$$

assume seu mínimo em $\alpha = 0$, o valor de α para o qual $(1 - \alpha)g^* + \alpha f = g^*$. Por outro lado, esta soma é uma função quadrática de α com algumas características a serem observadas. Primeiro, é que tal função está definida somente no intervalo $[0,1]$, segundo, tal função é o gráfico de uma parábola côncava para cima se

$f \neq g^*$, como pode ser observado no desenvolvimento da expressão abaixo:

$$\begin{aligned}
S &= S(\alpha) \\
&= \sum_{x \in X} \{g(x) - [(1 - \alpha)g^*(x) + \alpha f(x)]\}^2 w(x) \\
&= \sum_{x \in X} \{g(x) - [g^*(x) - \alpha g^*(x) + \alpha f(x)]\}^2 w(x) \\
&= \sum_{x \in X} \{g(x) - g^*(x) + \alpha g^*(x) - \alpha f(x)\}^2 w(x) \\
&= \sum_{x \in X} \{\alpha [g^*(x) - f(x)] + [g(x) - g^*(x)]\}^2 w(x) \\
&= \sum_{x \in X} \{\alpha^2 [g^*(x) - f(x)]^2 + 2\alpha [g^*(x) - f(x)][g(x) - g^*(x)] \\
&\quad + [g(x) - g^*(x)]^2\} w(x) \\
&= \alpha^2 \sum_{x \in X} [g^*(x) - f(x)]^2 w(x) + 2\alpha \sum_{x \in X} [g^*(x) - f(x)][g(x) - g^*(x)] w(x) \\
&\quad + \sum_{x \in X} [g(x) - g^*(x)]^2 w(x).
\end{aligned}$$

Podem ocorrer apenas duas situações com relação à derivada de S no ponto $\alpha = 0$: como $\alpha = 0$ é o mínimo de S para $0 \leq \alpha < 1$, ou $\frac{dS}{d\alpha} = 0$, significando que $\alpha = 0$ é um mínimo global ou $\alpha = 0$ é um mínimo restrito aos valores de α em $[0,1]$ e neste caso $\frac{dS}{d\alpha} > 0$ (Figura 4). Então, de modo geral pode-se concluir que:

$$\left. \frac{dS}{d\alpha} \right|_{\alpha=0} \geq 0.$$

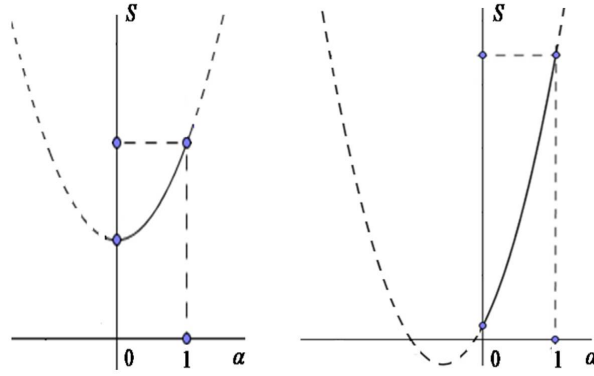


Figura 4 Representação da função $S = S(\alpha)$

Mas,

$$\begin{aligned} \frac{dS}{d\alpha} &= \sum_{x \in X} \left\{ 2\alpha [g^*(x) - f(x)]^2 + 2[g(x) - g^*(x)][g^*(x) - f(x)] \right\} w(x) \\ &= 2 \sum_{x \in X} \left\{ \alpha [g^*(x) - f(x)]^2 + [g(x) - g^*(x)][g^*(x) - f(x)] \right\} w(x). \end{aligned}$$

Agora, em $\alpha = 0$ tem-se:

$$\left. \frac{dS}{d\alpha} \right|_{\alpha=0} = 2 \sum_{x \in X} [g(x) - g^*(x)][g^*(x) - f(x)]w(x) \geq 0.$$

Portanto,

$$\sum_{x \in X} [g(x) - g^*(x)][g^*(x) - f(x)]w(x) \geq 0.$$

■

Teorema 2.6 Uma regressão isotônica g^* de g com pesos w em relação a uma

quase ordem \preceq , satisfaz:

$$\begin{aligned} \sum_{x \in X} [g(x) - f(x)]^2 w(x) &\geq \sum_{x \in X} [g(x) - g^*(x)]^2 w(x) \\ &\quad + \sum_{x \in X} [g^*(x) - f(x)]^2 w(x) \end{aligned} \quad (2.6)$$

para toda função f isotônica sobre X .

Demonstração: Note que:

$$\begin{aligned} [g(x) - f(x)]^2 w(x) &= [g(x) - g^*(x) + g^*(x) - f(x)]^2 w(x) \\ &= \{[g(x) - g^*(x)] + [g^*(x) - f(x)]\}^2 w(x) \\ &= \{[g(x) - g^*(x)]^2 + 2[g(x) - g^*(x)][g^*(x) - f(x)] \\ &\quad + [g^*(x) - f(x)]^2\} w(x) \\ &= [g(x) - g^*(x)]^2 w(x) + [g^*(x) - f(x)]^2 w(x) \\ &\quad + 2[g(x) - g^*(x)][g^*(x) - f(x)] w(x). \end{aligned}$$

Daí,

$$\begin{aligned} \sum_{x \in X} [g(x) - f(x)]^2 w(x) &= \sum_{x \in X} [g(x) - g^*(x)]^2 w(x) + \sum_{x \in X} [g^*(x) - f(x)]^2 w(x) \\ &\quad + 2 \sum_{x \in X} [g(x) - g^*(x)][g^*(x) - f(x)] w(x). \end{aligned}$$

Mas, pelo Teorema 2.5,

$$\sum_{x \in X} [g(x) - g^*(x)][g^*(x) - f(x)] w(x) \geq 0.$$

Consequentemente,

$$\sum_{x \in X} [g(x) - f(x)]^2 w(x) \geq \sum_{x \in X} [g(x) - g^*(x)]^2 w(x) + \sum_{x \in X} [g^*(x) - f(x)]^2 w(x).$$

■

Teorema 2.7 *Se uma função isotônica u satisfaz*

$$\sum_{x \in X} [g(x) - u(x)] [u(x) - f(x)] w(x) \geq 0$$

para toda função isotônica f sobre X , então u é uma regressão isotônica de g com pesos w .

Demonstração: Note que:

$$\begin{aligned} [g(x) - f(x)]^2 w(x) &= [g(x) - u(x) + u(x) - f(x)]^2 w(x) \\ &= \{[g(x) - u(x)] + [u(x) - f(x)]\}^2 w(x) \\ &= [g(x) - u(x)]^2 w(x) + [u(x) - f(x)]^2 w(x) \\ &\quad + 2 [g(x) - u(x)] [u(x) - f(x)] w(x). \end{aligned}$$

Logo,

$$\begin{aligned} \sum_{x \in X} [g(x) - f(x)]^2 w(x) &= \sum_{x \in X} [g(x) - u(x)]^2 w(x) + \sum_{x \in X} [u(x) - f(x)]^2 w(x) \\ &\quad + 2 \sum_{x \in X} [g(x) - u(x)] [u(x) - f(x)] w(x). \end{aligned}$$

Mas, por hipótese,

$$\sum_{x \in X} [g(x) - u(x)] [u(x) - f(x)] w(x) \geq 0,$$

assim, tem-se que:

$$\sum_{x \in X} [g(x) - f(x)]^2 w(x) \geq \sum_{x \in X} [g(x) - u(x)]^2 w(x) + \sum_{x \in X} [u(x) - f(x)]^2 w(x)$$

para toda f isotônica, e como

$$\sum_{x \in X} [u(x) - f(x)]^2 w(x) \geq 0,$$

vale que

$$\sum_{x \in X} [g(x) - f(x)]^2 w(x) \geq \sum_{x \in X} [g(x) - u(x)]^2 w(x)$$

para toda f isotônica. Logo, u minimiza a soma

$$\sum_{x \in X} [g(x) - f(x)]^2 w(x)$$

na classe das funções f isotônicas, ou seja, u é uma regressão isotônica de g .

■

De posse destes resultados, será demonstrado no Teorema 2.8 que a regressão isotônica é única.

Teorema 2.8 *A regressão isotônica de uma função g com pesos w com respeito a uma quase ordem \preceq sobre um conjunto finito X é única.*

Demonstração: *Suponha que para uma dada função g existam duas regressões isotônicas g_1^* e g_2^* associadas à mesma. Então ambas devem satisfazer a*

Equação 2.5, ou seja,

$$\begin{aligned} \sum_{x \in X} [g(x) - g_1^*(x)] [g_1^*(x) - f(x)] w(x) &\geq 0 \\ \sum_{x \in X} [g(x) - g_2^*(x)] [g_2^*(x) - f(x)] w(x) &\geq 0. \end{aligned}$$

Estas desigualdades são verdadeiras para qualquer função isotônica f sobre X , em particular para as funções g_1^* e g_2^* . Dessa maneira, substituindo f por g_2^* na primeira desigualdade e f por g_1^* na segunda, tem-se:

$$\begin{aligned} \sum_{x \in X} [g(x) - g_1^*(x)] [g_1^*(x) - g_2^*(x)] w(x) &\geq 0 \\ \sum_{x \in X} [g(x) - g_2^*(x)] [g_2^*(x) - g_1^*(x)] w(x) &\geq 0. \end{aligned}$$

Somando-se ambos os membros destas desigualdades:

$$\begin{aligned} &\sum_{x \in X} \{ [g(x) - g_1^*(x)] [g_1^*(x) - g_2^*(x)] + [g(x) - g_2^*(x)] [g_2^*(x) - g_1^*(x)] \} w(x) \geq 0 \\ &\Rightarrow \sum_{x \in X} \{ [g(x) - g_1^*(x)] [g_1^*(x) - g_2^*(x)] - [g(x) - g_2^*(x)] [g_1^*(x) - g_2^*(x)] \} w(x) \geq 0 \\ &\Rightarrow \sum_{x \in X} \{ [g_1^*(x) - g_2^*(x)] [g(x) - g_1^*(x) - g(x) + g_2^*(x)] \} w(x) \geq 0 \\ &\Rightarrow - \sum_{x \in X} [g_1^*(x) - g_2^*(x)]^2 w(x) \geq 0 \\ &\Rightarrow \sum_{x \in X} [g_1^*(x) - g_2^*(x)]^2 w(x) \leq 0 \\ &\Rightarrow \sum_{x \in X} [g_1^*(x) - g_2^*(x)]^2 w(x) = 0 \\ &\Rightarrow g_1^*(x) = g_2^*(x), \quad \forall x \in X. \end{aligned}$$

■

Os Teoremas 2.3 e 2.8 afirmam que a regressão isotônica existe e é única, no entanto, eles não mostram como encontrar a mesma. Isso será feito na seção que trata dos algoritmos utilizados para o cômputo da regressão isotônica. Alguns teoremas e resultados da regressão isotônica estão diretamente ligados com o aspecto geométrico do espaço das funções isotônicas, que será discutido na próxima seção.

2.1.4 Aspectos geométricos da regressão isotônica

Uma função $f : X \mapsto \mathbb{R}$, em que X é um conjunto finito com k elementos, pode ser interpretada como um ponto no Espaço Euclidiano k - dimensional $\mathbb{R}^k = \{(y_1, \dots, y_k), y_i \in \mathbb{R}\}$. Por exemplo, se $X = \{2, 3\}$, e $f : X \mapsto \mathbb{R}$ tal que $f(x) = x^2$, então, f pode ser identificada ao vetor

$$\begin{bmatrix} f(2) \\ f(3) \end{bmatrix} = \begin{bmatrix} 4 \\ 9 \end{bmatrix}.$$

Isto é, o conjunto de todas as funções $f : X \mapsto \mathbb{R}$, $\mathcal{A} = \{f : X = \{x_1, x_2\} \mapsto \mathbb{R}\}$, pode ser identificado com o espaço \mathbb{R}^2 (Figura 5). De modo semelhante, para $X = \{x_1, x_2, x_3\}$ (Figura 6).

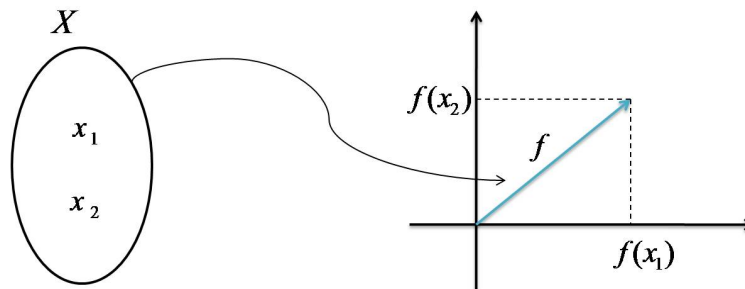


Figura 5 Identificação do conjunto $\mathcal{A} = \{f : X = \{x_1, x_2\} \mapsto \mathbb{R}\}$ e o \mathbb{R}^2

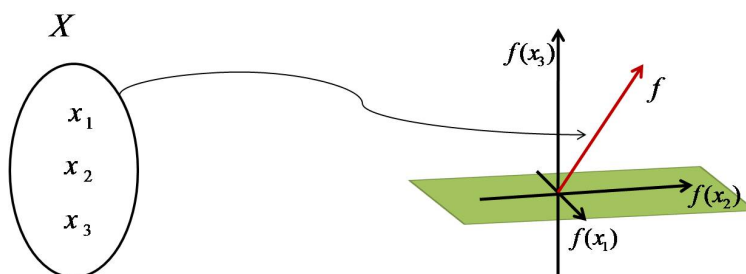


Figura 6 Identificação do conjunto $\mathcal{A} = \{f : X = \{x_1, x_2, x_3\} \mapsto \mathbb{R}\}$ e o \mathbb{R}^3

O conjunto das funções $f : X \mapsto \mathbb{R}$ com alguma restrição, pode não ser necessariamente todo o espaço \mathbb{R}^k , mas um subconjunto do mesmo. No caso das funções isotônicas, o *Espaço das funções isotônicas* não será todo o \mathbb{R}^k , mas um subconjunto com propriedades interessantes.

Para explicitar este subconjunto, considere uma função f isotônica de $X = \{x_1, x_2\}$ em \mathbb{R} , ou seja, para $x_1 \preceq x_2$ tem-se $f(x_1) \leq f(x_2)$, assim, o espaço das isotônicas é o conjunto $\Upsilon = \{(f(x_1), f(x_2)) \in \mathbb{R}^2 : f(x_1) \leq f(x_2)\}$ que é a parte hachurada na Figura 7. Dada uma função g não isotônica, a regressão isotônica g^* é um ponto no conjunto Υ o mais próximo possível do ponto g . Neste caso, g^* é a projeção ortogonal de g em Υ .

O espaço das funções isotônicas está relacionado aos conjuntos convexos. As Definições 2.9 e 2.10 serão necessárias para a compreensão desta relação.

Definição 2.9 Um conjunto C é convexo em \mathbb{R}^k se para quaisquer par de pontos p_1 e p_2 em C , e $\alpha \in [0, 1]$, a combinação linear convexa $\alpha p_1 + (1 - \alpha)p_2 \in C$.

O Teorema 2.4 mostra que o espaço das funções isotônicas é um conjunto convexo. De fato, este conjunto convexo é ainda mais específico.

Definição 2.10 Um conjunto C é um cone se $\forall p \in C$ e $0 \leq \alpha \in \mathbb{R}$, tem-se $\alpha p \in C$.

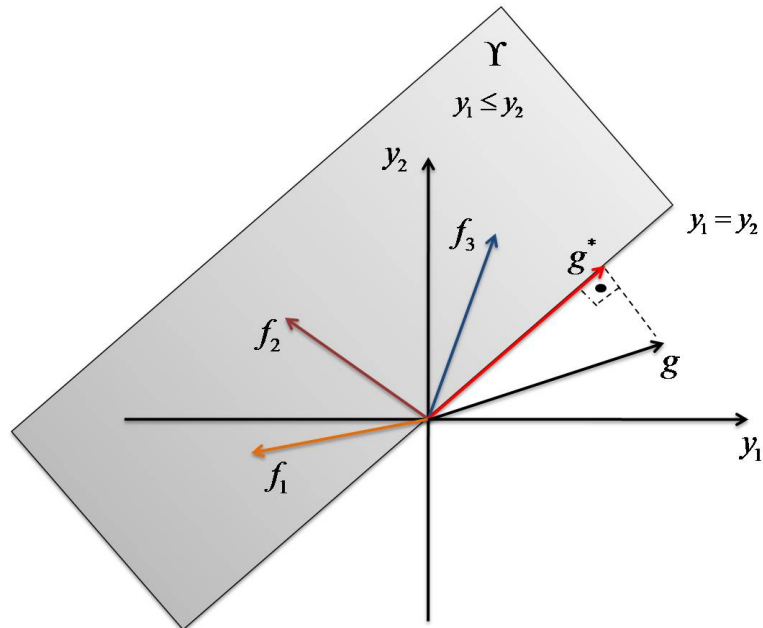


Figura 7 Espaço das isotônicas para $f : \{x_1, x_2\} \rightarrow \mathbb{R}$

Teorema 2.9 *Se f é isotônica e c é uma constante não negativa, então cf é isotônica.*

Demonstração: *Dados dois elementos do domínio de f , x_1 e x_2 tais que $x_1 \preceq x_2$, então:*

$$f(x_1) \leq f(x_2) \Rightarrow cf(x_1) \leq cf(x_2) \Rightarrow (cf)(x_1) \leq (cf)(x_2).$$

Portanto, cf é uma função isotônica.

■

Desta maneira, está provado que o espaço das funções isotônicas é um *Cone Convexo*. Agora, será demonstrado que a regressão isotônica de g é na verdade a projeção ortogonal de g neste cone convexo. Este é o conteúdo do

Teorema 2.10.

Antes de enunciar o Teorema 2.10 convém explicitar o uso de alguns conceitos e notações. Se $w = (w_1, w_2, \dots, w_k)$ é um vetor cujos componentes são números reais, com $w_i > 0$, então, pode-se definir um novo produto interno em \mathbb{R}^k , da forma: $\langle \vec{x}, \vec{y} \rangle_w = \sum_{i=1}^k x_i y_i w_i$. Com este novo produto interno tem-se todos os conceitos geométricos usuais, tais como: projeção ortogonal, norma, etc..

Exemplo 2.2 Considere os vetores em \mathbb{R}^3 : $\vec{x} = (0, 1, 2)$, $\vec{y} = (4, -1, 2)$ e $w = (1, 4, 1)$, então \vec{x} e \vec{y} são ortogonais com relação ao produto interno $\langle \cdot, \cdot \rangle_w$, pois:

$$\langle \vec{x}, \vec{y} \rangle_w = 0 \times 4 \times 1 + 1 \times (-1) \times 4 + 2 \times 2 \times 1 = 0 - 4 + 4 = 0$$

e

$$\|\vec{x}\|_w = \sqrt{\langle \vec{x}, \vec{x} \rangle} = \sqrt{0^2 \times 1 + 1^2 \times 4 + 2^2 \times 1} = \sqrt{8},$$

em que $\|\vec{x}\|_w$ denota a norma do vetor \vec{x} com relação aos pesos w .

Este é o produto interno que o espaço das funções isotônicas com peso w estará munido, e será utilizada a seguinte notação:

$$\langle f, g \rangle_w = \sum_{x \in X} f(x)g(x)w(x).$$

Teorema 2.10 Uma função isotônica u sobre X é a regressão isotônica de g com peso w se e somente se

$$\langle g - u, u \rangle_w = \sum_{x \in X} [g(x) - u(x)] u(x) w(x) = 0 \quad (2.7)$$

e

$$\langle g - u, f \rangle_w = \sum_{x \in X} [g(x) - u(x)] f(x) w(x) \leq 0, \quad (2.8)$$

para todo $x \in X$ e qualquer f isotônica em X .

Demonstração: Suponha que u satisfaz 2.7 e 2.8 acima, então,

$$\begin{aligned} & \sum_{x \in X} [g(x) - u(x)][u(x) - f(x)]w(x) \\ &= \underbrace{\sum_{x \in X} [g(x) - u(x)]u(x)w(x)}_{=0 \text{ (Equação 2.7)}} - \underbrace{\sum_{x \in X} [g(x) - u(x)]f(x)w(x)}_{\leq 0 \text{ (Desigualdade 2.8)}} \geq 0, \end{aligned}$$

ou seja, u satisfaz o Teorema 2.7, e portanto u é a regressão isotônica de g . Agora, suponha que u é a regressão isotônica de g , ou seja, $u = g^*$. Considere c um número real qualquer, então, a função $f(x) = cg^*$ é isotônica e satisfaz a Desigualdade 2.5, ou seja,

$$\begin{aligned} & \sum_{x \in X} [g(x) - g^*(x)][g^*(x) - cg^*(x)]w(x) \geq 0 \\ & \Rightarrow \sum_{x \in X} [g(x) - g^*(x)]g^*(x)[1 - c]w(x) \geq 0 \\ & \Rightarrow [1 - c] \sum_{x \in X} [g(x) - g^*(x)]g^*(x)w(x) \geq 0. \end{aligned}$$

Para $c > 1$, tem-se

$$\sum_{x \in X} [g(x) - g^*(x)]g^*(x)w(x) \leq 0,$$

e para $c < 1$, tem-se

$$\sum_{x \in X} [g(x) - g^*(x)]g^*(x)w(x) \geq 0,$$

logo:

$$\sum_{x \in X} [g(x) - g^*(x)]g^*(x)w(x) = 0.$$

Portanto, g^* satisfaz 2.7. Agora, falta provar que g^* também satisfaz 2.8. Novamente, por hipótese, g^* satisfaz 2.5:

$$\begin{aligned}
 & \sum_{x \in X} [g(x) - g^*(x)][g^*(x) - f(x)]w(x) \geq 0 \\
 \Rightarrow & \underbrace{\sum_{x \in X} [g(x) - g^*(x)]g^*(x)w(x)}_{=0 \text{ (Equação 2.7)}} - \sum_{x \in X} [g(x) - g^*(x)]f(x)w(x) \geq 0 \\
 \Rightarrow & \sum_{x \in X} [g(x) - g^*(x)]f(x)w(x) \leq 0,
 \end{aligned}$$

para toda f isotônica sobre X , ou seja, g^* satisfaz 2.8. ■

Os resultados apresentados até aqui sugerem importantes aspectos geométricos da regressão isotônica. Primeiro, o que está expresso na Equação 2.7 é que o produto interno $\langle g - g^*, g \rangle_w = 0$, ou seja, geometricamente g^* é a projeção ortogonal de g com pesos w no cone das isotônicas. A Desigualdade 2.8 implica que o produto interno $\langle g - g^*, f \rangle_w \leq 0$, onde f é uma função isotônica (Figura 8). O Teorema 2.7 indica que $\langle g - g^*, g^* - f \rangle_w \geq 0$, que está ilustrado na Figura 9.

O próximo resultado mostra uma propriedade interessante da regressão isotônica, tal fato será demonstrando usando os resultados do Teorema 2.10. Mas antes, note que toda função constante $f(x) = c$ é isotônica, em particular, $f(x) = 1$ e $f(x) = -1$. Estas funções estão relacionadas aos vetores de constantes em \mathbb{R}^k , $\vec{1} = (1, \dots, 1)$ e $-\vec{1} = (-1, \dots, -1)$. A Figura 10 ilustra esta situação para o caso onde $X = \{x_1, x_2\}$.

Teorema 2.11 *A regressão isotônica g^* de g satisfaz a seguinte igualdade:*

$$\left\langle g, \vec{1} \right\rangle_w = \left\langle g^*, \vec{1} \right\rangle_w \quad \text{isto é,} \quad \sum_{x \in X} g(x)w(x) = \sum_{x \in X} g^*(x)w(x) \quad (2.9)$$

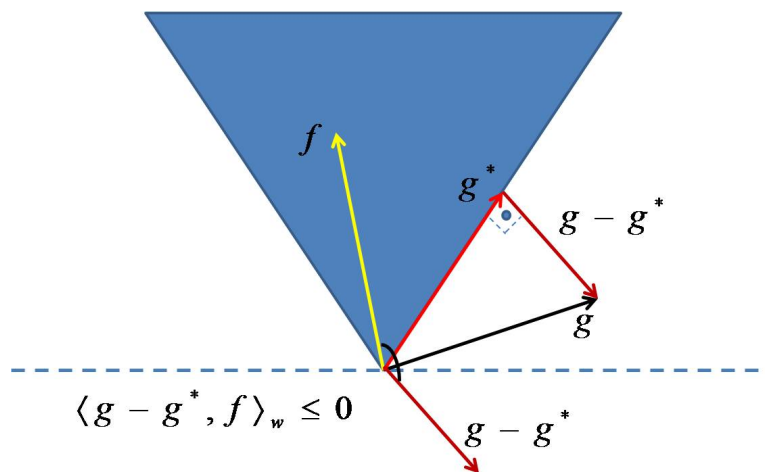


Figura 8 Representação da Desigualdade 2.8

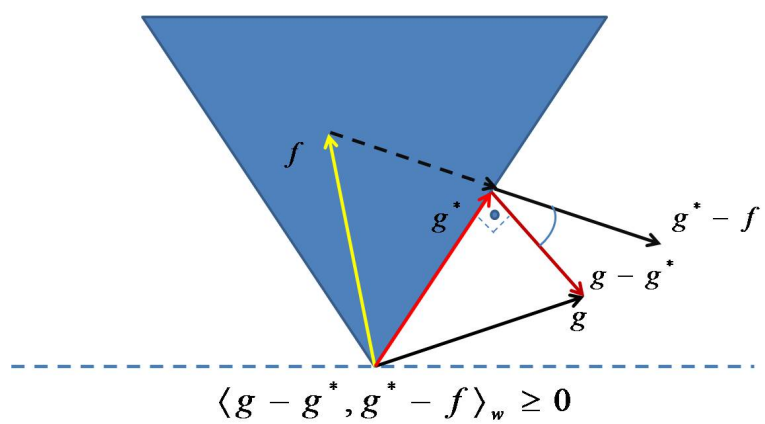


Figura 9 Representação do Teorema 2.7

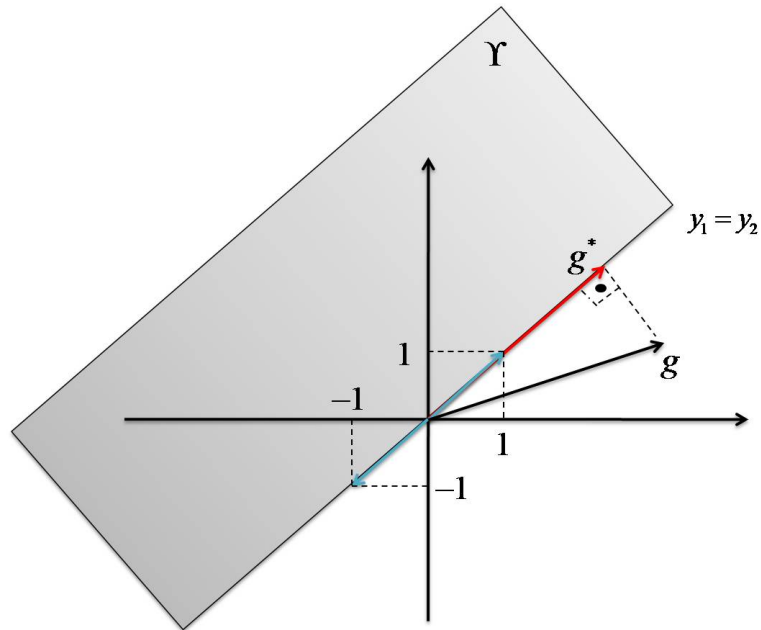


Figura 10 Representação das funções isotônicas constantes $f(x) = 1$ e $f(x) = -1$

Demonstração: A regressão isotônica satisfaz a Desigualdade 2.8 para toda f isotônica, ou seja,

$$\sum_{x \in X} [g(x) - g^*(x)] f(x) w(x) \leq 0,$$

qualquer que seja f isotônica. Dessa maneira, tomando as funções constantes $f(x) = 1$ e $f(x) = -1$ tem-se os seguintes resultados:

$$\sum_{x \in X} [g(x) - g^*(x)] w(x) \leq 0$$

e

$$\sum_{x \in X} [g(x) - g^*(x)] w(x) \geq 0.$$

Portanto,

$$\begin{aligned}
 \sum_{x \in X} [g(x) - g^*(x)] w(x) &= 0 \\
 \Rightarrow \sum_{x \in X} g(x) w(x) - \sum_{x \in X} g^*(x) w(x) &= 0 \\
 \Rightarrow \sum_{x \in X} g(x) w(x) &= \sum_{x \in X} g^*(x) w(x).
 \end{aligned}$$

■

Uma generalização das Equações 2.7 e 2.9 será apresentada agora como uma propriedade adicional da regressão isotônica. Uma definição e um lema preparam o método para a prova desta generalização.

Definição 2.11 *Sejam X munido da quase ordem \preceq , $Z \neq \emptyset$ um subconjunto de X e $g : Z \rightarrow \mathbb{R}$. A média de g sobre Z com pesos w é definida por:*

$$Av_g(Z) = \frac{\sum_{x \in Z} g(x) w(x)}{\sum_{x \in Z} w(x)} \quad (2.10)$$

O Lema 2.1 afirma que nos pontos onde a regressão isotônica é constante, essa constante será a média ponderada de g nos pontos deste conjunto.

Lema 2.1 *Seja c um número real, se o subconjunto $[g^* = c] = \{x \in X, g^*(x) = c\}$ é não vazio, então*

$$c = Av_g([g^* = c])$$

Demonstração:

$$\begin{aligned}
 & \sum_{x \in X} [g(x) - g^*(x)]^2 w(x) = \\
 & = \sum_{[g^* \neq c]} [g(x) - g^*(x)]^2 w(x) + \sum_{[g^* = c]} [g(x) - g^*(x)]^2 w(x) \\
 & = \sum_{[g^* \neq c]} [g(x) - g^*(x)]^2 w(x) + \sum_{[g^* = c]} [g(x) - c]^2 w(x).
 \end{aligned}$$

Agora, considere a função

$$S(t) = \sum_{[g^* = c]} [g(x) - t]^2 w(x),$$

$S(t)$ é uma função quadrática em t ,

$$\begin{aligned}
 \frac{dS}{dt} &= 0 \\
 \Rightarrow -2 \sum_{[g^* = c]} [g(x) - t] w(x) &= 0 \\
 \Rightarrow \sum_{[g^* = c]} g(x) w(x) - t \sum_{[g^* = c]} w(x) &= 0 \\
 \Rightarrow t &= \frac{\sum_{[g^* = c]} g(x) w(x)}{\sum_{[g^* = c]} w(x)} \\
 \Rightarrow t &= Av_g([g^* = c]),
 \end{aligned}$$

ou seja, S alcança seu mínimo em $t = Av_g([g^* = c])$, assim, se c não fosse igual a $Av_g([g^* = c])$, uma função isotônica \tilde{g}^* poderia ser encontrada coincidindo com g^* no conjunto $[g^* \neq c]$ e ligeiramente diferente de g^* em $[g^* = c]$ tendo uma soma de quadrados ponderada menor. Por exemplo, para $c > Av_g([g^* = c])$,

suponha:

$$\tilde{g}^*(x) = \begin{cases} g^*(x) & \text{para } x \in [g^* \neq c] \\ k & \text{para } x \in [g^* = c] \text{ com } Av_g[g^* = c] < k < c. \end{cases}$$

Então,

$$\sum_{[g^*=c]} [g(x) - k]^2 w(x)$$

é menor do que (Figura 11):

$$\sum_{[g^*=c]} [g(x) - c]^2 w(x),$$

o que contradiz o fato de g^* ser a regressão isotônica de g . Raciocínio análogo vale para o caso onde $c < Av_g([g^* = c])$. Portanto,

$$c = Av_g([g^* = c]).$$

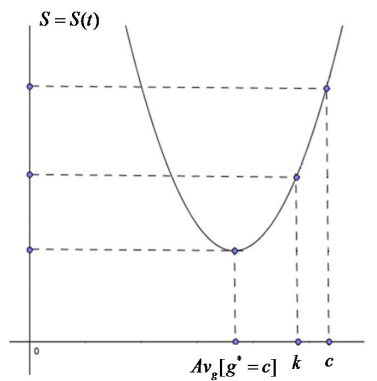


Figura 11 Representação da função $S = S(t)$

■

O Teorema 2.12 é uma generalização da Equação 2.7.

Teorema 2.12 Para $\Psi : \mathbb{R} \rightarrow \mathbb{R}$ qualquer ,

$$\sum_{x \in X} [g(x) - g^*(x)] \Psi(g^*(x)) w(x) = 0 \quad (2.11)$$

Demonstração: Considere que todos os valores que a regressão isotônica g^* assume estão no conjunto $C = \{c_1, c_2, \dots, c_k\}$ com $c_i \neq c_j, \forall i \neq j$, isto é, $g^*(X) = C$. Então,

$$\begin{aligned} & \sum_{x \in X} [g(x) - g^*(x)] \Psi(g^*(x)) w(x) = \\ &= \sum_{[g^*=c_1]} [g(x) - c_1] \Psi(c_1) w(x) + \sum_{[g^*=c_2]} [g(x) - c_2] \Psi(c_2) w(x) \\ &+ \dots + \sum_{[g^*=c_k]} [g(x) - c_k] \Psi(c_k) w(x) \\ &= \Psi(c_1) \sum_{[g^*=c_1]} [g(x) - c_1] w(x) + \Psi(c_2) \sum_{[g^*=c_2]} [g(x) - c_2] w(x) \\ &+ \dots + \Psi(c_k) \sum_{[g^*=c_k]} [g(x) - c_k] w(x) \\ &= \sum_{i=1}^k \left\{ \Psi(c_i) \sum_{[g^*=c_i]} [g(x) - c_i] w(x) \right\}. \end{aligned}$$

Mas, para cada c_i , pelo Lema 2.1, tem-se sobre o conjunto $[g^* = c_i]$:

$$\begin{aligned} \sum_{[g^*=c_i]} [g(x) - c_i] w(x) &= \sum_{[g^*=c_i]} g(x) w(x) - c_i \sum_{[g^*=c_i]} w(x) \\ &= \sum_{[g^*=c_i]} g(x) w(x) - \frac{\sum_{[g^*=c_i]} g(x) w(x)}{\sum_{[g^*=c_i]} w(x)} \sum_{[g^*=c_i]} w(x) \\ &= \sum_{[g^*=c_i]} g(x) w(x) - \sum_{[g^*=c_i]} g(x) w(x) \\ &= 0. \end{aligned}$$

Portanto,

$$\sum_{i=1}^k \left\{ \Psi(c_i) \sum_{[g^*=c_i]} [g(x) - c_i] w(x) \right\} = \sum_{i=1}^k \{ \Psi(c_i) \times 0 \} = 0.$$

■

O Teorema 2.12 afirma que $\langle g - g^*, \Psi(g^*) \rangle_w = 0$, isto significa que $g - g^*$ é ortogonal a qualquer função real Ψ composta com g^* .

2.1.5 Análise convexa

Existe uma relação entre determinadas funções convexas e os Estimadores de Máxima Verossimilhança para algumas distribuições importantes. Antes de expor o principal teorema que relaciona estes fatos, convém explicitar o uso de algumas notações. A seguinte definição de função convexa adotada por Lima (1976) será utilizada:

Definição 2.12 *Uma função $\Phi : I \mapsto \mathbb{R}$, definida em um intervalo, chama-se convexa quando $a, x, b \in I$ e $a < x < b$, o ponto $(x, \Phi(x))$ do gráfico de Φ está situado abaixo da secante (segmento de reta) que liga os pontos $(a, \Phi(a))$ e $(b, \Phi(b))$ (Figura 12).*

Observe que, com esta definição, Φ não é necessariamente contínua. Seja Φ uma função convexa definida da seguinte forma: Φ é finita para os pontos do intervalo I e infinita para pontos não pertencentes à I . Considere φ uma determinação arbitrária de suas derivadas, e caso Φ não seja derivável em determinado ponto, φ denotará qualquer número entre as derivadas direita e esquerda deste ponto, com φ definida e finita sobre I . Então φ é não decrescente (pois Φ é convexa). Para

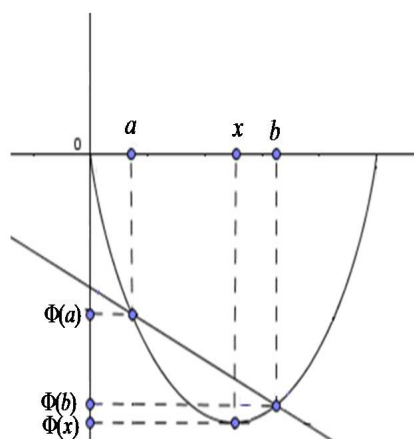


Figura 12 Representação de uma função convexa Φ

números u e v , defina:

$$\Delta_{\Phi}(u,v) = \Phi(u) - \Phi(v) - (u-v)\varphi(v)$$

se u e $v \in I$,

$$\Delta_{\Phi}(u,v) = \infty \quad (2.12)$$

se u e $v \notin I$.

O gráfico da Figura 13 ilustra o exposto na Equação 2.12. Observe que Δ_{Φ} é não negativo, e pode ser interpretado como o excesso do ponto pertencente ao gráfico de Φ entre u e v sobre o ponto pertencente à reta tangente ao gráfico em v . O fato de Δ_{Φ} ser não negativo decorre do pressuposto que Φ é uma função convexa, e desta maneira:

$$\Phi(u) \geq \Phi(v) + (u-v)\varphi(v) \quad (2.13)$$

para $u, v \in I$, com a desigualdade estrita se $u \neq v$ e se Φ é estritamente convexa. Em outras palavras, como Φ é convexa, o gráfico de Φ estará acima da reta tangente. Claramente, a adição de uma função linear em Φ não afeta Δ_Φ .

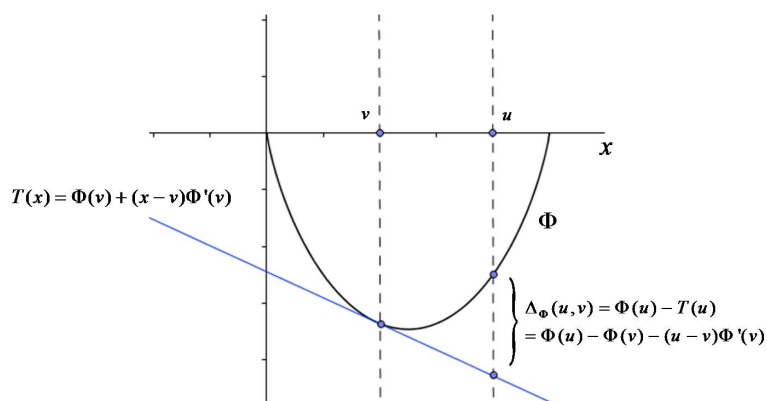


Figura 13 Representação da Equação 2.12

Da definição 2.12, tem-se os seguintes resultados:

$$\Delta_\Phi(r, s) = \Phi(r) - \Phi(s) - (r - s)\varphi(s)$$

$$\Delta_\Phi(s, t) = \Phi(s) - \Phi(t) - (s - t)\varphi(t).$$

Somando-se ambos os membros das igualdades acima, obtém-se:

$$\Delta_\Phi(r, s) + \Delta_\Phi(s, t) = \Phi(r) - \Phi(t) - (r - s)\varphi(s) - (s - t)\varphi(t).$$

Mas,

$$\Delta_\Phi(r, t) = \Phi(r) - \Phi(t) - (r - t)\varphi(t) \Rightarrow \Phi(r) - \Phi(t) = \Delta_\Phi(r, t) + (r - t)\varphi(t).$$

Substituindo este resultado na igualdade anterior:

$$\begin{aligned} \Delta_{\Phi}(r,s) + \Delta_{\Phi}(s,t) &= \Delta_{\Phi}(r,t) + (r-t)\varphi(t) - (r-s)\varphi(s) - (s-t)\varphi(t) \\ \Rightarrow \Delta_{\Phi}(r,s) + \Delta_{\Phi}(s,t) &= \Delta_{\Phi}(r,t) + [r-t-s+t]\varphi(t) - (r-s)\varphi(s) \\ \Rightarrow \Delta_{\Phi}(r,s) + \Delta_{\Phi}(s,t) &= \Delta_{\Phi}(r,t) - (r-s)[\varphi(s) - \varphi(t)] \\ \Rightarrow \Delta_{\Phi}(r,t) &= \Delta_{\Phi}(r,s) + \Delta_{\Phi}(s,t) + (r-s)[\varphi(s) - \varphi(t)], \quad (2.14) \\ &\text{se } s, t \in I. \end{aligned}$$

A Figura 14 ilustra geometricamente a Equação 2.14.

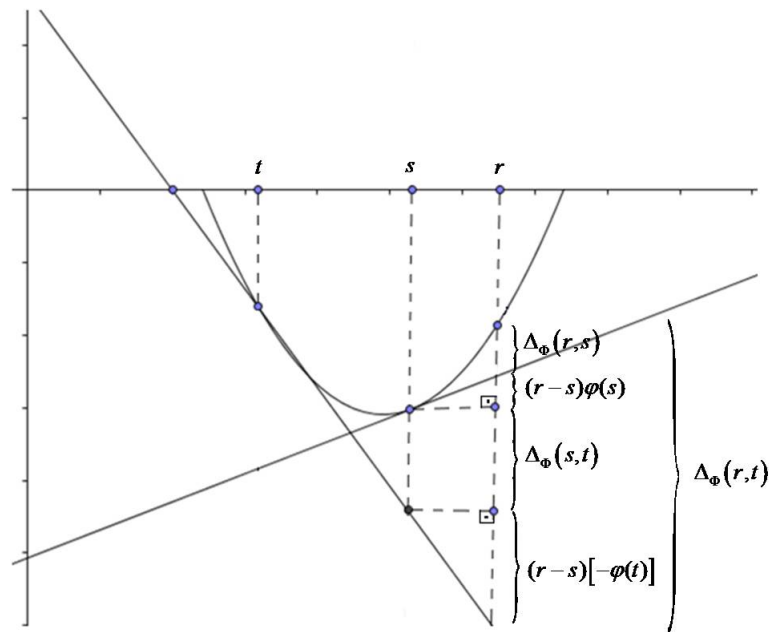


Figura 14 Representação da Equação 2.14

Estimação por Máxima Verossimilhança é maximizar a função de verossimilhança ou a log-verossimilhança com relação aos parâmetros, encontrando

assim o Estimador de Máxima Verossimilhança. Este processo também pode ser feito minimizando o negativo da log-verossimilhança e o problema passa a ser de minimização. No caso de restrições de ordem no espaço paramétrico, existe uma relação entre este processo de estimação e o de encontrar o mínimo da expressão

$$\sum_{x \in X} \Delta_{\Phi}[f(x), g(x)]w(x).$$

A ligação destas teorias está apresentada no Teorema 2.13.

Teorema 2.13 *Se f é isotônica em X e a imagem de f está contida em I e Φ é convexa, então:*

$$\sum_{x \in X} \Delta_{\Phi}[g(x), f(x)]w(x) \geq \sum_{x \in X} \Delta[g(x), g^*(x)]w(x) + \sum_{x \in X} \Delta[g^*(x), f(x)]w(x). \quad (2.15)$$

Consequentemente, g^* minimiza

$$\sum_{x \in X} \Delta_{\Phi}[g(x), f(x)]w(x) \quad (2.16)$$

e maximiza

$$\sum_{x \in X} \{\Phi(f(x)) + [g(x) - f(x)]\varphi(f(x))\} w(x). \quad (2.17)$$

na classe das funções isotônicas f com imagem em I . A função que minimiza (ou maximiza) é única se Φ é estritamente convexa.

Demonstração: Usando 2.14 com $r = g(x)$, $t = f(x)$ e $s = g^*(x)$,

tem-se:

$$\begin{aligned}
\Delta_{\Phi}(r,t) &= \Delta_{\Phi}(r,s) + \Delta_{\Phi}(s,t) + (r-s)[\varphi(s) - \varphi(t)] \\
\Rightarrow \Delta_{\Phi}[g(x),f(x)] &= \Delta_{\Phi}[g(x),g^*(x)] + \Delta_{\Phi}[g^*(x),f(x)] \\
&+ [g(x) - g^*(x)]\{\varphi(g^*(x)) - \varphi(f(x))\} \\
\Rightarrow \sum_{x \in X} \Delta_{\Phi}[g(x),f(x)]w(x) &= \sum_{x \in X} \Delta_{\Phi}[g(x),g^*(x)]w(x) \\
&+ \sum_{x \in X} \Delta_{\Phi}[g^*(x),f(x)]w(x) + \sum_{x \in X} [g(x) - g^*(x)]\{\varphi(g^*(x)) - \varphi(f(x))\}w(x).
\end{aligned}$$

Agora,

$$\begin{aligned}
&[g(x) - g^*(x)]\{\varphi(g^*(x)) - \varphi(f(x))\} \\
&= [g(x) - g^*(x)]\varphi(g^*(x)) - [g(x) - g^*(x)]\varphi(f(x)),
\end{aligned}$$

e dado que φ é não decrescente (Φ é uma função convexa), $\varphi(f(x))$ é isotônica se f também for. Logo, aplicando a Desigualdade 2.8, e a Equação 2.11, obtém-se:

$$\begin{aligned}
&\sum_{x \in X} [g(x) - g^*(x)]\{\varphi(g^*(x)) - \varphi(f(x))\}w(x) = \\
&= \underbrace{\sum_{x \in X} [g(x) - g^*(x)]\varphi(g^*(x))w(x)}_{=0 \text{ (Equação 2.11)}} - \underbrace{\sum_{x \in X} [g(x) - g^*(x)]\varphi(f(x))w(x)}_{\leq 0 \text{ (Desigualdade 2.8)}} \geq 0.
\end{aligned}$$

Portanto, a igualdade

$$\begin{aligned}
\sum_{x \in X} \Delta_{\Phi}[g(x),f(x)]w(x) &= \sum_{x \in X} \Delta_{\Phi}[g(x),g^*(x)]w(x) \\
&+ \sum_{x \in X} \Delta_{\Phi}[g^*(x),f(x)]w(x) + \underbrace{\sum_{x \in X} [g(x) - g^*(x)]\{\varphi(g^*(x)) - \varphi(f(x))\}w(x)}_{\geq 0}
\end{aligned}$$

implica em

$$\sum_{x \in X} \Delta_{\Phi}[g(x), f(x)]w(x) \geq \sum_{x \in X} \Delta[g(x), g^*(x)]w(x) + \sum_{x \in X} \Delta[g^*(x), f(x)]w(x).$$

Dessa maneira g^* minimiza $\sum_{x \in X} \Delta_{\Phi}[g(x), f(x)]w(x)$, isto é, maximiza

$$- \sum_{x \in X} \Delta_{\Phi}[g(x), f(x)]w(x).$$

Como em 2.12:

$$\Delta_{\Phi}[g(x), f(x)] = \Phi(g(x)) - \Phi(f(x)) - [g(x) - f(x)]\varphi(f(x)),$$

portanto g^* maximiza

$$- \sum_{x \in X} \Phi(g(x)) + \sum_{x \in X} \{\Phi(f(x)) + [g(x) - f(x)]\varphi(f(x))\},$$

como o termo

$$- \sum_{x \in X} \Phi(g(x))$$

não depende de f , g^* maximiza

$$\sum_{x \in X} \{\Phi(f(x)) + [g(x) - f(x)]\varphi(f(x))\},$$

na classe das funções isotônicas f com imagem em I .

Agora, se Φ é estritamente convexa, a conclusão da unicidade decorre do fato de que o termo do lado direito de 2.15 é estritamente positivo a menos que $f = g^*$.

■

Corolário 2.1 *Sejam $\Psi_1, \Psi_2, \dots, \Psi_p$ funções de valores reais arbitrárias e h_1, h_2, \dots, h_m funções isotônicas sobre X . Então g^* minimiza:*

$$\sum_{x \in X} \Delta_{\Phi}[g(x), f(x)]w(x)$$

na classe das funções isotônicas f com imagem em I satisfazendo qualquer uma ou todas das seguintes condições:

$$\sum_{x \in X} [g(x) - f(x)]\Psi_j(f(x))w(x) = 0, \quad j = 1, 2, \dots, p$$

$$\sum_{x \in X} f(x)h_j(x)w(x) \geq \sum_{x \in X} g(x)h_j(x)w(x), \quad j = 1, 2, \dots, m.$$

Demonstração: *Pelo Teorema 2.12, é imediato que:*

$$\sum_{x \in X} [g(x) - g^*(x)]\Psi_j[g^*(x)]w(x) = 0,$$

e pela Equação 2.8 do Teorema 2.10, tem-se que:

$$\begin{aligned} \sum_{x \in X} [g(x) - g^*(x)]h_j(x)w(x) &\leq 0 \\ \Rightarrow \sum_{x \in X} g(x)h_j(x)w(x) - \sum_{x \in X} g^*(x)h_j(x)w(x) &\leq 0 . \\ \Rightarrow \sum_{x \in X} g^*(x)h_j(x)w(x) &\geq \sum_{x \in X} g(x)h_j(x)w(x) \end{aligned}$$

■

Os resultados dessa seção serão utilizados a seguir no processo de Estimativa por Máxima Verossimilhança de parâmetros ordenados de famílias de distribuições binomiais, geométricas, poisson e gama.

2.1.5.1 Estimação por máxima verossimilhança para uma família de distribuições binomiais com parâmetros ordenados

Existem diversas situações onde a variável resposta é modelada por uma distribuição binomial, dentre elas, pode-se destacar problemas ligados a bioensaios, como por exemplo, experimentos com inseticidas. Considere x o valor da dose de um determinado inseticida variando em um conjunto de doses $X = \{x_1, x_2, \dots, x_n\}$, em que $x_1 < x_2 < \dots < x_n$.

Suponha que, após a aplicação de uma dose x , a probabilidade de uma resposta (por exemplo, morte do inseto) seja $p(x)$. O ensaio pode ser modelado da seguinte forma: considere que para cada $x \in X$, $m(x)$ insetos sejam expostos à dose x , com $a(x)$ mortes ocorrendo, e $b(x) = m(x) - a(x)$ o número de sobreviventes. Dessa maneira, a verossimilhança para uma amostra aleatória para todas as doses em X é dada por:

$$L \propto \prod_{x \in X} [p(x)]^{a(x)} [1 - p(x)]^{b(x)}.$$

Considerando

$$\bar{y}(x) = \frac{a(x)}{m(x)}$$

o número médio de respostas por dose, tem-se

$$a(x) = \bar{y}(x)m(x)$$

e

$$b(x) = m(x) - a(x) = m(x) - \bar{y}(x)m(x) = [1 - \bar{y}(x)]m(x).$$

Assim, a log-verossimilhança é dada por

$$\begin{aligned}
 l &\propto \log L \\
 &\propto \sum_{x \in X} \log \left\{ [p(x)]^{a(x)} [1 - p(x)]^{b(x)} \right\} \\
 &\propto \sum_{x \in X} \{ a(x) \log p(x) + b(x) \log [1 - p(x)] \} \\
 &\propto \sum_{x \in X} \{ \bar{y}(x) m(x) \log p(x) + [1 - \bar{y}(x)] m(x) \log [1 - p(x)] \} \\
 &\propto \sum_{x \in X} \{ \bar{y}(x) \log p(x) + [1 - \bar{y}(x)] \log [1 - p(x)] \} m(x).
 \end{aligned}$$

É conhecido que na ausência de restrições o estimador de máxima verossimilhança para $p(x)$ é $\bar{y}(x)$ (MOOD; GRAYBILL; BOES, 1974). É natural esperar que a medida que a dose aumenta, a probabilidade de resposta (morte do inseto) $p(x)$ também aumente, ou seja, é natural esperar que $p(x)$ seja uma função crescente de x . Entretanto, como a resposta de um inseto para cada dose é uma variável aleatória Bernoulli (1 - morte, 0 - sobrevivência) com parâmetro $p(x)$, pode acontecer que mesmo com o aumento da dose, ocorra uma diminuição na resposta, ou seja, o número de mortos na dose x_{i+1} seja menor do que o número de mortos na dose x_i . Isso não é raro e nem mesmo algo difícil de conceber dada a aleatoriedade do processo se $m(x)$ é pequeno.

Quando $m(x)$ é grande para cada x , a razão $\bar{y}(x)$ é geralmente crescente em função de x , como realmente era de se esperar. Mas, se por algum motivo as médias amostrais $\bar{y}(x)$ estão na ordem errada, seria interessante utilizar, no processo de estimação, a informação da restrição pertinente ao problema, maximizando a verossimilhança em um espaço restrito definido por $p(x_1) \leq p(x_2) \leq \dots \leq p(x_n)$.

Maximizar l é equivalente a minimizar $-l$, ou seja, minimizar a expressão

$$\sum_{x \in X} \{-\bar{y}(x) \log p(x) - [1 - \bar{y}(x)] \log[1 - p(x)]\} m(x). \quad (2.18)$$

Assim, a solução do problema da estimação por máxima verossimilhança restrita de $p(x)$ é obter a função que minimiza 2.18 na classe das funções isotônicas sobre X .

Para se colocar a expressão 2.18 no contexto do Teorema 2.13, considere a função estritamente convexa (Figura 15):

$$\begin{aligned} \Phi(u) &= u \log u + (1 - u) \log(1 - u), \quad 0 < u < 1 \\ \Phi(0) &= 0, \\ \Phi(1) &= 0, \end{aligned} \quad (2.19)$$

e usando a notação dada pela Equação 2.12, a determinação arbitrária da derivada de Φ :

$$\begin{aligned} \varphi(u) &= \log u - \log(1 - u), \quad 0 < u < 1 \\ \varphi(0) &= -\infty, \\ \varphi(1) &= \infty. \end{aligned} \quad (2.20)$$

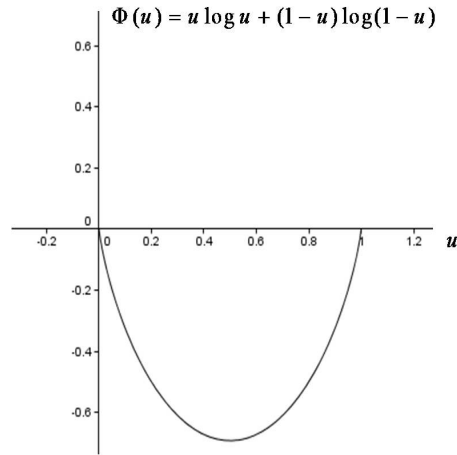


Figura 15 Representação da função $\Phi(u)$

Deste modo, de 2.12, 2.19 e 2.20, para $0 < u < 1$:

$$\begin{aligned}
 \Delta_{\Phi}[\bar{y}(x), p(x)] &= \Phi(\bar{y}(x)) - \Phi(p(x)) - [\bar{y}(x) - p(x)]\varphi(p(x)) \\
 &= \bar{y}(x) \log \bar{y}(x) + [1 - \bar{y}(x)] \log[1 - \bar{y}(x)] - p(x) \log p(x) \\
 &\quad - [1 - p(x)] \log[1 - p(x)] - [\bar{y}(x) - p(x)][\log p(x) - \log[1 - p(x)]] \\
 &= \bar{y}(x) \log \bar{y}(x) + [1 - \bar{y}(x)] \log[1 - \bar{y}(x)] - p(x) \log p(x) \\
 &\quad - \log[1 - p(x)] + p(x) \log[1 - p(x)] - \bar{y}(x) \log p(x) + \bar{y}(x) \log[1 - p(x)] \\
 &\quad + p(x) \log p(x) - p(x) \log[1 - p(x)] \\
 &= \bar{y}(x) \log \bar{y}(x) + [1 - \bar{y}(x)] \log[1 - \bar{y}(x)] - \bar{y}(x) \log p(x) \\
 &\quad - [1 - \bar{y}(x)] \log[1 - p(x)].
 \end{aligned}$$

Portanto,

$$\begin{aligned}
 \sum_{x \in X} \Delta_{\Phi}[\bar{y}(x), p(x)] m(x) &= \\
 &= \sum_{x \in X} \{ \bar{y}(x) \log \bar{y}(x) + [1 - \bar{y}(x)] \log[1 - \bar{y}(x)] - \bar{y}(x) \log p(x) \\
 &\quad - [1 - \bar{y}(x)] \log[1 - p(x)] \} m(x)
 \end{aligned}$$

e como

$$\bar{y}(x) \log \bar{y}(x) + [1 - \bar{y}(x)] \log[1 - \bar{y}(x)]$$

é constante, isto é, não depende de $p(x)$, minimizar

$$\sum_{x \in X} \Delta_{\Phi}[\bar{y}(x), p(x)] m(x)$$

em $p(x)$, é equivalente a minimizar

$$\sum_{x \in X} \{-\bar{y}(x) \log p(x) - [1 - \bar{y}(x)] \log[1 - p(x)]\} m(x),$$

que é a -log-verossimilhança 2.18. Dessa maneira, o Estimador de Máxima Verossimilhança restrito é a função isotônica que minimiza

$$\sum_{x \in X} \Delta_{\Phi}[\bar{y}(x), p(x)] m(x)$$

na classe das funções isotônicas p sobre X , e pelo Teorema 2.13 esta função é a regressão isotônica $\bar{y}^*(x)$ da função $\bar{y}(x)$ com pesos $m(x)$. Observe que, para esta conclusão, foi fundamental o uso da análise convexa. Além disso, todos estes resultados são válidos para uma quase ordem em X .

2.1.5.2 Estimação por máxima verossimilhança para uma família de distribuições geométricas com parâmetros ordenados

Considere as variáveis aleatórias Y_i , $i = 1, 2, \dots, n$ com densidade geométrica, ou seja,

$$f_{Y_i}(y_i; p_i) = p_i q_i^{y_i} \quad \text{para } y_i = 0, 1, 2, \dots$$

onde $p_i + q_i = 1$, com $0 < p_i \leq 1$. Dadas as observações independentes y_1, y_2, \dots, y_n , a verossimilhança para tal amostra é:

$$L = L(p_1, p_2, \dots, p_n; y_1, y_2, \dots, y_n) = \prod_{i=1}^n p_i q_i^{y_i}. \quad (2.21)$$

A esperança de Y_i é dada por $\mu_i = \mu(i) = \frac{q_i}{p_i}$ (MOOD; GRAYBILL; BOES, 1974). Desta maneira,

$$\mu_i = \frac{q_i}{p_i} \Rightarrow q_i = p_i \mu_i \Rightarrow q_i = (1 - q_i) \mu_i \Rightarrow q_i + q_i \mu_i = \mu_i \Rightarrow q_i = \frac{\mu_i}{1 + \mu_i}$$

e

$$p_i = 1 - q_i \Rightarrow p_i = 1 - \frac{\mu_i}{1 + \mu_i} \Rightarrow p_i = \frac{1}{1 + \mu_i}.$$

Daí, fazendo uma mudança de parâmetros em 2.21,

$$\begin{aligned} L &= \prod_{i=1}^n p_i q_i^{y_i} \\ &= \prod_{i=1}^n \left[\left(\frac{1}{1 + \mu_i} \right) \left(\frac{\mu_i}{1 + \mu_i} \right)^{y_i} \right] \\ &= \prod_{i=1}^n \left[(1 + \mu_i)^{-1} \mu_i^{y_i} (1 + \mu_i)^{-y_i} \right] \\ &= \prod_{i=1}^n (1 + \mu_i)^{-(1+y_i)} \mu_i^{y_i}. \end{aligned}$$

Considere o problema de estimação por máxima verossimilhança restrita:

$$\text{Maximizar}_{\mu=(\mu_1, \dots, \mu_n)} \left\{ \prod_{i=1}^n (1 + \mu_i)^{-(1+y_i)} \mu_i^{y_i} \right\}$$

sujeito à restrição $\mu_i \leq \mu_j$, $i < j$. Maximizar a verossimilhança é equivalente a

minimizar a -log-verossimilhança, ou seja, o problema em questão é equivalente

a:

$$\text{Minimizar}_{\mu=(\mu_1, \dots, \mu_n)} \left\{ \sum_{i=1}^n [-y_i \log \mu_i + (1 + y_i) \log(1 + \mu_i)] \right\}$$

sujeito à restrição $\mu_i \leq \mu_j$ quando $i < j$. Agora, considerando a função estritamente convexa (Figura 16):

$$\Phi(u) = u \log u - (1 + u) \log(1 + u), \quad u > 0,$$

tem-se,

$$\varphi(u) = \log u - \log(1 + u), \quad u > 0.$$

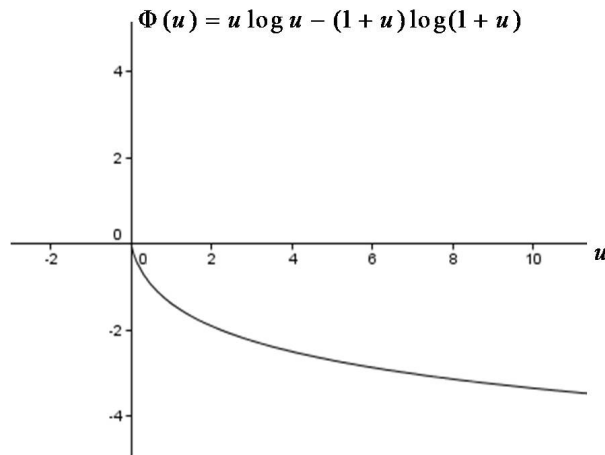


Figura 16 Representação da função $\Phi(u)$

Portanto, por 2.12:

$$\begin{aligned}
 \Delta_{\Phi}[y_i, \mu_i] &= \Phi(y_i) - \Phi(\mu_i) - (y_i - \mu_i)\varphi(\mu_i) \\
 &= y_i \log y_i - (1 + y_i) \log(1 + y_i) - \mu_i \log \mu_i + (1 + \mu_i) \log(1 + \mu_i) \\
 &\quad - (y_i - \mu_i)[\log \mu_i - \log(1 + \mu_i)] \\
 &= y_i \log y_i - (1 + y_i) \log(1 + y_i) - \mu_i \log \mu_i + \log(1 + \mu_i) \\
 &\quad + \mu_i \log(1 + \mu_i) - y_i \log \mu_i + y_i \log(1 + \mu_i) + \mu_i \log \mu_i - \mu_i \log(1 + \mu_i) \\
 &= y_i \log y_i - (1 + y_i) \log(1 + y_i) - y_i \log \mu_i + (1 + y_i) \log(1 + \mu_i).
 \end{aligned}$$

Desta maneira,

$$\begin{aligned}
 \sum_{i=1}^n \Delta_{\Phi}[y_i, \mu_i] &= \\
 &= \sum_{i=1}^n \{y_i \log y_i - (1 + y_i) \log(1 + y_i) - y_i \log \mu_i + (1 + y_i) \log(1 + \mu_i)\},
 \end{aligned}$$

e como

$$y_i \log y_i - (1 + y_i) \log(1 + y_i)$$

é constante em relação a μ_i , minimizar

$$\sum_{i=1}^n \Delta_{\Phi}[y_i, \mu_i]$$

em μ_i é equivalente a minimizar

$$\sum_{i=1}^n \{-y_i \log \mu_i + (1 + y_i) \log(1 + \mu_i)\},$$

ou seja, o mínimo para a função -log-verossimilhança. Assim, o Estimador de

Máxima Verossimilhança procurado é a função isotônica que minimiza

$$\sum_{i=1}^n \Delta_{\Phi}[y_i, \mu_i]$$

na classe das funções isotônicas sobre o conjunto dos índices $X = \{1, 2, \dots, n\}$, com pesos $w(x) = 1$ e pelo Teorema 2.13 essa função é a regressão isotônica y_i^* de y_i (em que se considerou a função $y(i) = y_i$). Novamente, a análise convexa foi fundamental para a comprovação deste fato. Note que todos estes resultados valem para uma quase ordem \preceq no conjunto dos índices X .

2.1.5.3 Estimação por máxima verossimilhança para uma família de distribuições poisson com parâmetros ordenados

Suponha que um sistema seja posto em operação e n_1 falhas ou erros ocorrem durante o primeiro período de tempo $[0, t_1]$. O dispositivo pode então ser modificado para melhorar a performance. Subsequentemente, n_2 falhas podem ocorrer no segundo intervalo de tempo de tamanho t_2 e assim para k períodos de tempo. É assumido que o número de falhas, Y_i , observado durante o i -ésimo período de tempo, tem uma distribuição de Poisson com média $\lambda_i t_i$, com $\lambda_i > 0$, isto é,

$$f_{Y_i}(y_i; \lambda_i) = \frac{(\lambda_i t_i)^{y_i} e^{-\lambda_i t_i}}{y_i!}, \quad y_i = 0, 1, \dots$$

Considere uma amostra de falhas independentes resultantes em n períodos de tempo y_1, y_2, \dots, y_n . A verossimilhança para tal amostra é:

$$L = L(\lambda_1, \lambda_2, \dots, \lambda_n; y_1, y_2, \dots, y_n) = \prod_{i=1}^n \left[\frac{(\lambda_i t_i)^{y_i} e^{-\lambda_i t_i}}{y_i!} \right]. \quad (2.22)$$

Com a hipótese de que o sistema é melhorado a cada teste, deseja-se maximizar a verossimilhança 2.22 sujeito à restrição $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_n$. A log-verossimilhança é dada por:

$$\begin{aligned}
 l &= \log L \\
 &= \sum_{i=1}^n \log \left\{ \frac{(\lambda_i t_i)^{y_i} e^{-\lambda_i t_i}}{y_i!} \right\} \\
 &= \sum_{i=1}^n \{y_i \log(\lambda_i t_i) - \lambda_i t_i - \log y_i!\} \\
 &= \sum_{i=1}^n \{y_i \log(\lambda_i) + y_i \log(t_i) - \lambda_i t_i - \log y_i!\} \\
 &= \sum_{i=1}^n \{y_i \log(\lambda_i) - \lambda_i t_i\} + \sum_{i=1}^n \{y_i \log(t_i) - \log y_i!\} \\
 &= \sum_{i=1}^n \left\{ \frac{y_i}{t_i} \log(\lambda_i) - \lambda_i \right\} t_i + \sum_{i=1}^n \{y_i \log(t_i) - \log y_i!\},
 \end{aligned}$$

e como,

$$\sum_{i=1}^n \{y_i \log(t_i) - \log y_i!\}$$

é constante em relação aos λ_i , maximizar a log-verossimilhança restrita é um problema equivalente a

$$\text{Maximizar}_{\lambda=(\lambda_1, \dots, \lambda_n)} \sum_{i=1}^n \left\{ \frac{y_i}{t_i} \log(\lambda_i) - \lambda_i \right\} t_i \quad (2.23)$$

sujeito a $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_n$.

Agora, considere a função estritamente convexa (Figura 17):

$$\Phi(u) = u \log u, \quad u > 0,$$

então

$$\varphi(u) = 1 + \log u, \quad u > 0.$$

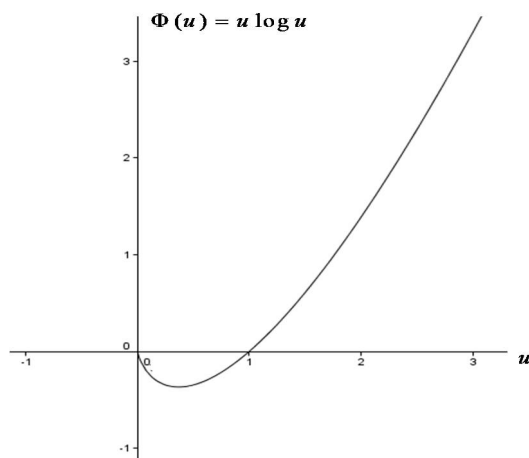


Figura 17 Representação da função $\Phi(u)$

Considere $X = \{1, 2, \dots, n\}$ e as funções $g(i) = \frac{y_i}{t_i}$, $f(i) = \lambda_i$ e $w(i) = t_i$, com $i = 1, 2, \dots, n$. Observe que a função f é antitônica com relação aos índices i , e segue do Teorema 2.13 que a regressão antitônica g^* de g maximiza a expressão:

$$\begin{aligned} & \sum_{i=1}^n \{ \Phi[f(i)] + [g(i) - f(i)]\varphi[f(i)] \} w(i) \\ &= \sum_{i=1}^n \{ f(i) \log[f(i)] + [g(i) - f(i)][\log[f(i) + 1]] \} w(i) \\ &= \sum_{i=1}^n \{ f(i) \log[f(i)] + g(i) \log[f(i)] + g(i) - f(i) \log[f(i)] - f(i) \} w(i) \\ &= \sum_{i=1}^n \{ g(i) \log[f(i)] + g(i) - f(i) \} w(i) \\ &= \sum_{i=1}^n \left\{ \frac{y_i}{t_i} \log(\lambda_i) + \frac{y_i}{t_i} - \lambda_i \right\} t_i, \end{aligned}$$

que é equivalente a maximizar a expressão:

$$\sum_{i=1}^n \left\{ \frac{y_i}{t_i} \log(\lambda_i) - \lambda_i \right\} t_i$$

na classe das funções antitônicas f em X . Mas, isso é exatamente a solução do problema da Equação 2.23, ou seja, a regressão antitônica $g^*(i)$ de $g(i) = \frac{y_i}{t_i}$ com pesos $w(i) = t_i$ é o Estimador de Máxima Verossimilhança para $\lambda_1, \lambda_2, \dots, \lambda_n$ sujeito à restrição $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_n$.

2.1.5.4 Estimação por máxima verossimilhança para uma família de distribuições gama com parâmetros ordenados

Sejam α_i e λ_i números reais positivos, e Y_i uma variável aleatória com densidade gama, para $i = 1, 2, \dots, n$:

$$f_{Y_i}(y_i; \lambda_i, \alpha_i) = \frac{\lambda_i^{-\alpha_i}}{\Gamma(\alpha_i)} y_i^{\alpha_i-1} e^{-\frac{y_i}{\lambda_i}}. \quad (2.24)$$

Considere y_1, y_2, \dots, y_n como observações independentes das variáveis aleatórias Y_1, Y_2, \dots, Y_n com densidades $f(y_1; \lambda_1, \alpha_1), \dots, f(y_n; \lambda_n, \alpha_n)$ respectivamente, em que α_i é conhecido, e o problema da estimação por Máxima Verossimilhança de $\lambda_1, \lambda_2, \dots, \lambda_n$ com restrições de ordem $\lambda_1 \leq \lambda_2 \leq \dots \leq \lambda_n$. Considere $\lambda = (\lambda_1, \lambda_2, \dots, \lambda_n)$, $\alpha = (\alpha_1, \alpha_2, \dots, \alpha_n)$, $\mathbf{y} = (y_1, y_2, \dots, y_n)$. A verossimilhança para esta amostra é dada por:

$$L = L(\lambda; \alpha, \mathbf{y}) = \prod_{i=1}^n \left[\frac{\lambda_i^{-\alpha_i}}{\Gamma(\alpha_i)} y_i^{\alpha_i-1} e^{-\frac{y_i}{\lambda_i}} \right],$$

e a log-verossimilhança:

$$l = \log L = \sum_{i=1}^n \left\{ -\alpha_i \log \lambda_i + (\alpha_i - 1) \log y_i - \frac{x_i}{\lambda_i} - \log[\Gamma(\alpha_i)] \right\}.$$

Observe que a parte

$$(\alpha_i - 1) \log y_i - \log[\Gamma(\alpha_i)]$$

é constante em relação aos λ_i , logo, o problema de maximização de l sujeito às restrições de ordem é equivalente a:

$$\text{Maximizar}_{\lambda=(\lambda_1, \dots, \lambda_n)} \sum_{i=1}^n \left[-\log \lambda_i - \frac{x_i}{\alpha_i \lambda_i} \right] \alpha_i \quad (2.25)$$

sujeito às restrições $\lambda_1 \leq \lambda_2 \leq \dots \leq \lambda_n$.

Considere a função estritamente convexa (Figura 18):

$$\Phi(u) = -\log u, \quad u > 0,$$

então,

$$\varphi(u) = -\frac{1}{u}, \quad u > 0.$$

Considerando $X = \{1, 2, \dots, n\}$, $g(i) = \frac{y_i}{\alpha_i}$, $f(i) = \lambda_i$ e $w(i) = \alpha_i$, com

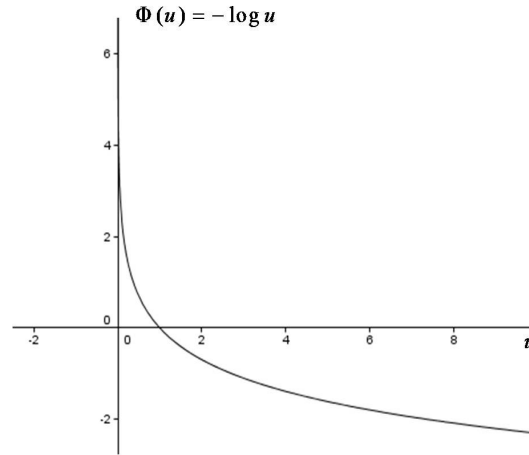


Figura 18 Representação da função $\Phi(u)$

$i = 1, 2, \dots, n$, segue do Teorema 2.13 que a regressão isotônica g^* de g , maximiza

$$\begin{aligned}
 & \sum_{i=1}^n \{ \Phi[f(i)] + [g(i) - f(i)] \varphi[f(i)] \} w(i) \\
 &= \sum_{i=1}^n \left\{ -\log[f(i)] + [g(i) - f(i)] \left[-\frac{1}{f(i)} \right] \right\} w(i) \\
 &= \sum_{i=1}^n \left\{ -\log[f(i)] - \frac{g(i)}{f(i)} + 1 \right\} w(i) \\
 &= \sum_{i=1}^n \left\{ -\log \lambda_i - \frac{y_i}{\alpha_i \lambda_i} + 1 \right\} \alpha_i,
 \end{aligned}$$

que é equivalente a maximizar

$$\sum_{i=1}^n \left\{ -\log \lambda_i - \frac{y_i}{\alpha_i \lambda_i} \right\} \alpha_i$$

na classe das funções isotônicas $\lambda(\cdot)$. Mas isso é exatamente a solução do problema da Equação 2.25, ou seja, a regressão isotônica $g^*(i)$ de $g(i) = \frac{y_i}{\alpha_i}$ com pesos $w(i) = \alpha_i$ é o Estimador de Máxima Verossimilhança sujeito à restrição $\lambda_1 \leq \lambda_2 \leq \dots \leq \lambda_n$. Mais adiante será considerado o caso mais geral, em que os parâmetros α_i podem ser desconhecidos.

2.1.6 Estimação por máxima verossimilhança de parâmetros ordenados na família exponencial

Os resultados anteriores podem ser colocados em um contexto mais geral. Considere uma família de densidades com dois parâmetros θ e h com respeito à alguma medida $v_h(dy)$ (que não precisa ser especificada) que pode depender do parâmetro h mas não do parâmetro θ , descrita da seguinte maneira:

$$f(y; \theta, h) = \exp\{\Phi(\theta) + (y - \theta)\varphi(\theta)\}h \quad (2.26)$$

em que Φ é estritamente convexa e φ é uma medida de determinação de suas derivadas, digamos suas derivadas à direita, θ assume valores em um intervalo de números reais, e h positivo. Então,

$$\int f(y; \theta, h)v_h(dy) = 1.$$

Afirmção 2.1 θ é a média da distribuição com a densidade em (2.26).

Demonstração: *Por hipótese,*

$$\begin{aligned} \int f(y; \theta, h)v_h(dy) &= 1 \\ \Rightarrow \frac{\partial}{\partial \theta} \int f(y; \theta, h)v_h(dy) &= 0 \\ \Rightarrow \frac{\partial}{\partial \theta} \int e^{[\Phi(\theta) + (y - \theta)\varphi(\theta)]h} v_h(dy) &= 0 \\ \Rightarrow \int \frac{\partial}{\partial \theta} \{e^{\Phi(\theta)h} \times e^{(y - \theta)\varphi(\theta)h} v_h(dy)\} &= 0 \\ \Rightarrow \int \{h\varphi(\theta)e^{\Phi(\theta)h} \times e^{(y - \theta)\varphi(\theta)h}\} v_h(dy) + \\ \int \{e^{\Phi(\theta)h} [-\varphi(\theta)h + (y - \theta)\varphi'(\theta)h] e^{(y - \theta)\varphi(\theta)h}\} v_h(dy) &= 0 \\ \Rightarrow \{h\varphi(\theta) \int e^{[\Phi(\theta) + (y - \theta)\varphi(\theta)]h} v_h(dy)\} - \{h\varphi(\theta) \int e^{[\Phi(\theta) + (y - \theta)\varphi(\theta)]h} v_h(dy)\} + \\ \{\varphi'(\theta)h \int (y - \theta)e^{[\Phi(\theta) + (y - \theta)\varphi(\theta)]h} v_h(dy)\} &= 0 \\ \Rightarrow \varphi'(\theta)h \int (y - \theta)e^{[\Phi(\theta) + (y - \theta)\varphi(\theta)]h} v_h(dy) &= 0. \end{aligned}$$

Agora, como por hipótese $h > 0$ e Φ estritamente convexa, tem-se que $\varphi'(\theta) = \Phi''(\theta) > 0$, ou seja, $\varphi'(\theta)h > 0$. Portanto,

$$\int (y - \theta)e^{[\Phi(\theta)+(y-\theta)\varphi(\theta)]h}v_h(dy) = 0$$

ou

$$\int (y - \theta)f(y; \theta, h)v_h(dy) = 0, \quad (2.27)$$

o que implica que θ é a média da densidade dada na Equação (2.26). ■

Afirmção 2.2 A variância da distribuição com a densidade dada na Equação (2.26) é:

$$\frac{1}{h\varphi'(\theta)}.$$

Demonstração: Considerando a Equação ,

$$\begin{aligned} & \int (y - \theta)f(y; \theta, h)v_h(dy) = 0 \\ \Rightarrow & \frac{\partial}{\partial \theta} \left\{ \int (y - \theta)e^{[\Phi(\theta)+(y-\theta)\varphi(\theta)]h}v_h(dy) \right\} = 0 \\ \Rightarrow & \int \frac{\partial}{\partial \theta} \left\{ (y - \theta)e^{[\Phi(\theta)+(y-\theta)\varphi(\theta)]h}v_h(dy) \right\} = 0 \\ \Rightarrow & \int \left\{ [(-1)e^{\Phi(\theta)h} + (y - \theta)\varphi(\theta)he^{\Phi(\theta)h}] e^{(y-\theta)\varphi(\theta)h} + \right. \\ & \left. (y - \theta)e^{\Phi(\theta)h} [(-\varphi(\theta)h + (y - \theta)\varphi'(\theta)he^{(y-\theta)\varphi(\theta)h})] \right\} v_h(dy) = 0 \\ \Rightarrow & \int \left\{ -e^{\Phi(\theta)h+(y-\theta)\varphi(\theta)h} + (y - \theta)\varphi(\theta)he^{\Phi(\theta)h+(y-\theta)\varphi(\theta)h} \right. \\ & \left. - (y - \theta)\varphi(\theta)he^{\Phi(\theta)h+(y-\theta)\varphi(\theta)h} + (y - \theta)^2\varphi'(\theta)he^{\Phi(\theta)h+(y-\theta)\varphi(\theta)h} \right\} v_h(dy) = 0 \\ \Rightarrow & - \underbrace{\int e^{[\Phi(\theta)+(y-\theta)\varphi(\theta)]h}v_h(dy)}_{=1} + \varphi'(\theta)h \int (y - \theta)^2 e^{[\Phi(\theta)+(y-\theta)\varphi(\theta)]h}v_h(dy) = 0 \\ \Rightarrow & -1 + \varphi'(\theta)h \int (y - \theta)^2 e^{[\Phi(\theta)+(y-\theta)\varphi(\theta)]h}v_h(dy) = 0 \\ \Rightarrow & \int (y - \theta)^2 e^{[\Phi(\theta)+(y-\theta)\varphi(\theta)]h}v_h(dy) = \frac{1}{\varphi'(\theta)h} \end{aligned}$$

ou

$$\int (y - \theta)^2 f(y; \theta, h) v_h(dy) = \frac{1}{h\varphi'(\theta)}, \quad (2.28)$$

o que implica que a variância da distribuição com a densidade dada na Equação (2.26) é:

$$\frac{1}{h\varphi'(\theta)}.$$

■

As Afirmações 2.3 e 2.4 aparecem em Barlow et al. (1972). Entretanto, os detalhes nas demonstrações são partes inerente a este trabalho, bem como a apresentação dos Exemplos 2.3 e 2.4.

Afirmção 2.3 *A distribuição Normal com parâmetros μ e σ^2 pertence à família de distribuições em (2.26):*

Demonstração:

$$\theta = \mu, \Phi(\theta) = \frac{\theta^2}{2}, \varphi(\theta) = \theta, h = \frac{1}{\sigma^2} \text{ e } v_h(dy) = \left(\frac{h}{2\pi}\right)^{\frac{1}{2}} \exp\left(-\frac{hy^2}{2}\right) dy.$$

Daí,

$$\begin{aligned} f(y; \theta, h)v_h(dy) &= \exp\{[\Phi(\theta) + (y - \theta)\varphi(\theta)]h\} v_h(dy) \\ &= \exp\left\{\left[\frac{\theta^2}{2} + (y - \theta)\theta\right]h\right\} v_h(dy) \\ &= \exp\left\{\left[\frac{\theta^2}{2} + y\theta - \theta^2\right]h\right\} v_h(dy) \\ &= \exp\left\{\left[y\theta - \frac{\theta^2}{2}\right]h\right\} \left(\frac{h}{2\pi}\right)^{\frac{1}{2}} \exp\left(-\frac{hy^2}{2}\right) dy \\ &= \left(\frac{h}{2\pi}\right)^{\frac{1}{2}} \exp\left\{\left[y\theta - \frac{\theta^2}{2} - \frac{y^2}{2}\right]h\right\} dy \\ &= \left(\frac{h}{2\pi}\right)^{\frac{1}{2}} \exp\left\{-\frac{h}{2}(y - \theta)^2\right\} dy \\ &= \left(\frac{1}{2\pi\sigma^2}\right)^{\frac{1}{2}} \exp\left\{-\frac{1}{2\sigma^2}(y - \mu)^2\right\} dy. \end{aligned}$$



Afirmção 2.4 A distribuição Bernoulli com parâmetro p pertence à família de distribuições em (2.26):

Demonstração:

$$\theta = p, \Phi(\theta) = \theta \log \theta + (1 - \theta) \log(1 - \theta), \varphi(\theta) = \log \left(\frac{\theta}{1 - \theta} \right), h = 1 \text{ e } v_h(dy) = dy.$$

Daí,

$$\begin{aligned} f(y; \theta, h) v_h(dy) &= \exp \{ [\Phi(\theta) + (y - \theta) \varphi(\theta)] h \} v_h(dy) \\ &= \exp \left\{ \theta \log \theta + (1 - \theta) \log(1 - \theta) + (y - \theta) \log \left(\frac{\theta}{1 - \theta} \right) \right\} dy \\ &= \exp \{ \theta \log \theta + (1 - \theta) \log(1 - \theta) + (y - \theta) [\log \theta - \log(1 - \theta)] \} dy \\ &= \exp \{ \theta \log \theta + \log(1 - \theta) - \theta \log(1 - \theta) + y \log \theta - y \log(1 - \theta) \\ &\quad - \theta \log \theta + \theta \log(1 - \theta) \} dy \\ &= \exp \{ \log(1 - \theta) + y \log \theta - y \log(1 - \theta) \} dy \\ &= \exp \{ y \log \theta + (1 - y) \log(1 - \theta) \} dy \\ &= \exp \left\{ \log \theta^y + \log (1 - \theta)^{1 - y} \right\} dy \\ &= \exp \{ \log \theta^y \} \cdot \exp \left\{ \log (1 - \theta)^{1 - y} \right\} dy \\ &= \theta^y (1 - \theta)^{1 - y} dy \\ &= p^y (1 - p)^{1 - y} dy \end{aligned}$$



Afirmção 2.5 A distribuição Poisson com parâmetro λ pertence à família de distribuições em (2.26):

Demonstração:

$$\theta = \lambda, \Phi(\theta) = \theta \log \theta, \varphi(\theta) = 1 + \log \theta, h = 1 \text{ e } v_h(dy) = \frac{e^{-y}}{y!} dy.$$

Daí,

$$\begin{aligned}
f(y; \theta, h)v_h(dy) &= \exp \{[\Phi(\theta) + (y - \theta)\varphi(\theta)] h\} v_h(dy) \\
&= \exp \{ \theta \log \theta + (y - \theta)(1 + \log \theta) \} \frac{e^{-y}}{y!} dy \\
&= \exp \{ \theta \log \theta + y + y \log \theta - \theta - \theta \log \theta \} \frac{e^{-y}}{y!} dy \\
&= \exp \{ y + y \log \theta - \theta \} \frac{e^{-y}}{y!} dy \\
&= \frac{e^{-\theta} \theta^y}{y!} dy \\
&= \frac{e^{-\lambda} \lambda^y}{y!} dy.
\end{aligned}$$

■

Afirmação 2.6 A distribuição Exponencial com parâmetro λ pertence à família de distribuições em (2.26).

Demonstração:

$$\theta = \frac{1}{\lambda}, \Phi(\theta) = -1 + \log\left(\frac{1}{\theta}\right), \varphi(\theta) = -\frac{1}{\theta}, h = 1, v_h(dy) = dy.$$

$$\begin{aligned}
f(y; \theta, h)v_h(dy) &= \exp \{[\Phi(\theta) + (y - \theta)\varphi(\theta)] h\} v_h(dy) \\
&= \exp \left\{ -1 + \log\left(\frac{1}{\theta}\right) + (y - \theta)\left(-\frac{1}{\theta}\right) \right\} dy \\
&= \exp \left\{ -1 + \log\left(\frac{1}{\theta}\right) - \frac{y}{\theta} + 1 \right\} dy \\
&= \exp \left\{ \log\left(\frac{1}{\theta}\right) - \frac{y}{\theta} \right\} dy \\
&= \frac{1}{\theta} e^{-\frac{1}{\theta}y} dy \\
&= \lambda e^{-\lambda y} dy.
\end{aligned}$$

■

Afirmção 2.7 A distribuição Gama com parâmetros α e λ pertence à família de distribuições em (2.26).

Demonstração:

$$\theta = \frac{\alpha}{\lambda}, \Phi(\theta) = \theta - \log\theta - 1, \varphi(\theta) = 1 - \frac{1}{\theta}, h = \alpha, v_h(dy) = \frac{h}{\Gamma(h)}(hy)^{h-1}e^{-hy}dy.$$

Daí,

$$\begin{aligned} f(y; \theta, h)v_h(dy) &= \exp\{[\Phi(\theta) + (y - \theta)\varphi(\theta)]h\}v_h(dy) \\ &= \exp\left\{\left[\theta - \log\theta - 1 + (y - \theta)\left(1 - \frac{1}{\theta}\right)\right]h\right\}v_h(dy) \\ &= \exp\left\{\left[\theta - \log\theta - 1 + y - \frac{y}{\theta} - \theta + 1\right]h\right\}v_h(dy) \\ &= \exp\left\{-h\log\theta + hy - \frac{hy}{\theta}\right\}\frac{h}{\Gamma(h)}(hy)^{h-1}\exp\{-hy\}dy \\ &= \exp\left\{-h\log\theta + hy - \frac{hy}{\theta} - hy\right\}\frac{h}{\Gamma(h)}(hy)^{h-1}dy \\ &= \exp\left\{\log\theta^{-h}\right\}\exp\left\{-\frac{hy}{\theta}\right\}\frac{h}{\Gamma(h)}(hy)^{h-1}dy \\ &= \theta^{-h}\frac{h}{\Gamma(h)}(hy)^{h-1}\exp\left\{-\frac{hy}{\theta}\right\}dy \\ &= \left(\frac{\alpha}{\lambda}\right)^{-\alpha}\frac{\alpha}{\Gamma(\alpha)}\alpha^{\alpha-1}y^{\alpha-1}\exp\left\{-\frac{\lambda}{\alpha}\alpha y\right\}dy \\ &= \frac{\alpha^{-\alpha}}{\lambda^{-\alpha}}\frac{\alpha}{\Gamma(\alpha)}\alpha^{\alpha-1}y^{\alpha-1}e^{-\lambda y}dy \\ &= \frac{\lambda^\alpha}{\Gamma(\alpha)}y^{\alpha-1}e^{-\lambda y}dy. \end{aligned}$$

■

O Teorema 2.14 fornece um importante resultado para situações práticas, em que o processo de modelagem considera uma família de distribuições pertencentes à família exponencial.

Teorema 2.14 *Seja X um conjunto finito, quase ordenado por \preceq . Para cada $x \in X$, considere a distribuição condicional de Y dado x tendo densidade dada em (2.26), ou seja,*

$$f_Y(y; \theta, h) = \exp\{[\Phi(\theta) + (y - \theta)\varphi(\theta)]h\},$$

com $\theta = \mu(x) = E[Y|x]$ e $h = a\nu(x)$, em que $\nu(x)$ é um número positivo conhecido para cada x , mas o número positivo a pode ser desconhecido. Se amostras aleatórias são retiradas destas distribuições condicionais, com tamanho $m(x) > 0$, $x \in X$, então, a Estimativa de Máxima Verossimilhança de $\mu(\cdot)$ dado que $\mu(\cdot)$ é isotônica, é fornecida pela regressão isotônica das médias amostrais, com pesos $w(x) = \nu(x)m(x)$, $x \in X$.

Demonstração: Considere os valores amostrais representados por $y_j(x)$, $j = 1, 2, \dots, m(x)$, $x \in X$, e $\bar{y}(x)$ a média amostral:

$$\bar{y}(x) = \frac{\sum_{j=1}^{m(x)} y_j(x)}{m(x)}.$$

Então, a densidade nos valores amostrais para cada $x \in X$ é:

$$f(y_j(x); \theta, h) = \exp\{[\Phi(\mu(x)) + (y_j(x) - \mu(x))\varphi(\mu(x))]a\nu(x)\}.$$

A função de verossimilhança dado os valores amostrais $y_j(x)$ é:

$$\begin{aligned} L &= \prod_{x \in X} \prod_{j=1}^{m(x)} \exp\{[\Phi(\mu(x)) + (y_j(x) - \mu(x))\varphi(\mu(x))]a\nu(x)\} \\ &= \exp\left\{\sum_{x \in X} \sum_{j=1}^{m(x)} \{[\Phi(\mu(x)) + (y_j(x) - \mu(x))\varphi(\mu(x))]a\nu(x)\}\right\}. \end{aligned}$$

Dessa maneira, a função de log-verossimilhança é:

$$\begin{aligned}
l &= \log L \\
&= \sum_{x \in X} \sum_{j=1}^{m(x)} \{[\Phi(\mu(x)) + (y_j(x) - \mu(x)) \varphi(\mu(x))] a \nu(x)\} \\
&= \sum_{x \in X} \left\{ \sum_{j=1}^{m(x)} \{[\Phi(\mu(x)) + (y_j(x) - \mu(x)) \varphi(\mu(x))] a \nu(x)\} \right\} \\
&= a \sum_{x \in X} \left\{ \nu(x) \sum_{j=1}^{m(x)} \{[\Phi(\mu(x)) + (y_j(x) - \mu(x)) \varphi(\mu(x))]\} \right\} \\
&= a \sum_{x \in X} \{ \nu(x) [m(x) \Phi(\mu(x)) + m(x) (\bar{y}(x) - \mu(x)) \varphi(\mu(x))] \} \\
&= a \sum_{x \in X} \{ [\Phi(\mu(x)) + (\bar{y}(x) - \mu(x)) \varphi(\mu(x))] \nu(x) m(x) \}.
\end{aligned}$$

Como a é positivo (h e $\nu(x)$ são positivos), maximizar l sujeito a $\mu(\cdot)$ isotônica é equivalente a maximizar a expressão:

$$\sum_{x \in X} \{ [\Phi(\mu(x)) + (\bar{y}(x) - \mu(x)) \varphi(\mu(x))] \nu(x) m(x) \},$$

sujeito a $\mu(\cdot)$ ser isotônica. Mas, pelo Teorema 2.13, esta expressão é unicamente maximizada na classe das funções isotônicas $\mu(\cdot)$ pela regressão isotônica $\bar{y}^*(x)$ de $\bar{y}(x)$ com pesos $w(x) = \nu(x) m(x)$, $x \in X$.

■

Exemplo 2.3 Considere o caso usual em Estatística Experimental em que as respostas $y_j(x)$ provém de uma distribuição normal com médias $\mu(x)$ e variâncias constantes iguais a σ^2 , ou seja, $y_j(x) \sim N(\mu(x), \sigma^2)$, com $\mu(\cdot)$ conhecida ser isotônica em $x \in X$. Levando em conta os resultados obtidos na demonstração da Afirmação 2.3, tem-se: $\theta = \mu(x)$ e $h = \frac{1}{\sigma^2}$ e, portanto, considerando $a = \frac{1}{\sigma^2}$,

$\nu(x) = 1$ e o Teorema 2.14, garante-se que a estimativa de $\mu(\cdot)$ sujeita à restrição de $\mu(\cdot)$ ser isotônica é simplesmente a regressão isotônica de $\bar{y}(x)$ com os pesos conhecidos $w(x) = m(x)$, $x \in X$. Assim, pode-se estimar $\mu(\cdot)$ sujeito às restrições de ordem mesmo desconhecendo a variância, supondo apenas que a mesma é constante para cada $x \in X$.

Exemplo 2.4 Considerando o caso normal e variâncias proporcionais, $y_j(x) \sim N(\mu(x), k(x)\sigma^2)$ com a constante de proporção $k(x)$ conhecida, tem-se, pela Afirmação 2.3: $\theta = \mu(x)$ e $h = \frac{1}{k(x)\sigma^2}$, assim, considerando $a = \frac{1}{\sigma^2}$, $\nu(x) = \frac{1}{k(x)}$ e o Teorema 2.14, garante-se que a Estimativa de Máxima Verossimilhança sob a restrição de isotonicidade para $\mu(x)$ é a regressão isotônica $\bar{y}^*(x)$ de $\bar{y}(x)$ com os pesos conhecidos $w(x) = \frac{m(x)}{k(x)}$, $x \in X$.

Exemplo 2.5 Um resultado mais geral poder ser considerado para uma amostra semelhante a dos exemplos 2.3 e 2.4, proveniente de uma distribuição Gama com parâmetros α e λ , em que ambos os parâmetros são desconhecidos, mas com α constante para cada x . Suponha $y_j(x) \sim \text{Gama}(\alpha, \lambda(x))$, ou seja, para cada x o parâmetro α é desconhecido mas é constante, enquanto que os parâmetros $\lambda(x)$ são desconhecidos. Deseja-se estimar, sob restrição de isotonicidade, as médias $\mu(x)$ das variáveis aleatórias $Y(x)$. Para isso, conforme Afirmação 2.7, basta considerar $h = \alpha$, $\nu(x) = 1$ e o Teorema 2.14. A Estimativa de Máxima Verossimilhança sob restrição de isotonicidade para $\mu(x)$ é a regressão isotônica $\bar{y}^*(x)$ de $\bar{y}(x)$ com os pesos conhecidos $w(x) = m(x)$. Observe que, neste caso, como α é constante e $\mu(x) = \frac{\alpha}{\lambda(x)}$, os resultados da restrição de ordem se estendem para os parâmetros $\frac{1}{\lambda(x)}$.

Exemplo 2.6 Um resultado mais geral pode ser obtido para o Exemplo 2.5, considerando $y_j(x) \sim \text{Gama}(\nu(x)\alpha, \lambda(x))$, com $\nu(x)$ conhecido. Neste caso basta

considerar $h = \alpha\nu(x)$ e o Teorema 2.14. A estimativa de Máxima Verossimilhança sob restrição de isotonicidade para $\mu(x)$ é a regressão isotônica $\bar{y}^*(x)$ de $\bar{y}(x)$ com os pesos conhecidos $w(x) = m(x)\nu(x)$.

2.1.7 O problema das estimativas de componentes de variâncias negativas

Antes de expor uma situação mais geral, relacionada à análise de variância, convém destacar alguns resultados inerentes à estimação de variâncias ordenadas de distribuições normais, para o caso em que se conhece a média. Este resultado será utilizado posteriormente em uma situação mais geral. Alguns resultados de estatística básica relacionando a normal com a gama serão expostos.

2.1.7.1 Estimação de variâncias ordenadas de distribuições normais com médias conhecidas

Suponha $X \sim N(0, \sigma^2)$, deseja-se encontrar a distribuição de $Y = X^2$.

Por transformações de variáveis (MOOD; GRAYBILL; BOES, 1974):

$$\begin{aligned}
 f_Y(y) &= \left[\frac{1}{2} \frac{1}{\sqrt{y}} f_X(-\sqrt{y}) + \frac{1}{2} \frac{1}{\sqrt{y}} f_X(\sqrt{y}) \right] I_{(0, \infty)} \\
 &= \left[\frac{1}{2} \frac{1}{\sqrt{y}} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2\sigma^2}y} + \frac{1}{2} \frac{1}{\sqrt{y}} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2\sigma^2}y} \right] I_{(0, \infty)} \\
 &= \frac{1}{\sqrt{y}} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2\sigma^2}y} I_{(0, \infty)} \\
 &= \left(\frac{1}{2\sigma^2} \right)^{\frac{1}{2}} \frac{1}{\sqrt{\pi}} y^{-\frac{1}{2}} e^{-\frac{1}{2\sigma^2}y} I_{(0, \infty)} \\
 &= \frac{\left(\frac{1}{2\sigma^2}\right)^{\frac{1}{2}}}{\sqrt{\pi}} y^{-\frac{1}{2}} e^{-\frac{1}{2\sigma^2}y} I_{(0, \infty)} \\
 &= \frac{\lambda^\alpha}{\Gamma(\alpha)} y^{\alpha-1} e^{-\lambda y} I_{(0, \infty)},
 \end{aligned}$$

com $\lambda = \frac{1}{2\sigma^2}$ e $\alpha = \frac{1}{2}$. Portanto, se $X \sim N(0, \sigma^2)$, a variável $Y = X^2 \sim \text{Gama}(\alpha = \frac{1}{2}, \lambda = \frac{1}{2\sigma^2})$, que tem média σ^2 e variância $2\sigma^4$.

Suponha observações amostrais independentes $y_j(x)$, provindas de uma variável aleatória $Y_j(x) \sim N(\mu(x), \sigma^2(x))$, $j = 1, 2, \dots, m(x)$, $x \in X$. Se os $\mu(x)$ são conhecidos, pode-se subtraí-los obtendo distribuições com média zero, isto é, pode-se, sem perda de generalidade, considerar as observações $\tilde{Y}_j(x) \sim N(0, \sigma^2(x))$. Deseja-se estimar $\sigma^2(x)$ sob a hipótese de isotonicidade, isto é, $x_1 < x_2 \Rightarrow \sigma^2(x_1) < \sigma^2(x_2)$.

Barlow et al. (1972) discutem este problema, relacionando o mesmo com a teoria da regressão isotônica. Observe que, considerando o exposto anteriormente, a estatística

$$\tilde{t}_j(x) = \tilde{y}_j^2(x)$$

tem distribuição Gama com parâmetros $\alpha = \frac{1}{2}$ e $\lambda = \frac{1}{2\sigma^2(x)}$. A média desta distribuição é $\sigma^2(x)$ (a média da Gama é $\frac{\alpha}{\lambda}$) e a variância $2\sigma^4(x)$ (a variância da Gama é $\frac{\alpha}{\lambda^2}$). Desta maneira, o problema passa a ser encontrar estimativas de médias ordenadas de uma família de distribuições Gama com parâmetro α conhecido. Mas isto já foi solucionado, como resultado do Teorema 2.14 no Exemplo 2.5. A Estimativa de Máxima Verossimilhança de $\sigma^2(x)$ sob a hipótese de isotonicidade com respeito a uma quase ordem sobre X é regressão isotônica \tilde{t}^* de \tilde{t} com pesos $m(x)$, em que:

$$\tilde{t}_j(x) = \frac{\sum_{j=1}^{m(x)} \tilde{y}_j^2(x)}{m(x)}, \quad x \in X.$$

Desses resultados, segue a Afirmação 2.8.

Afirmação 2.8 *Suponha observações $y_j(x)$, provindas das variáveis aleatórias*

$Y_j(x) \sim N(\mu(x), \sigma^2(x))$, $j = 1, 2, \dots, m(x)$, para cada $x \in X$. Se os $\mu(x)$ são conhecidos, as Estimativas de Máxima Verossimilhança de $\sigma^2(x)$, com a restrição de $\sigma^2(x)$ ser isotônica com relação a uma quase ordem sobre X é a regressão isotônica \tilde{t}^* de \tilde{t} com pesos $m(x)$, em que

$$\tilde{t}_j(x) = \frac{\sum_{j=1}^{m(x)} (y_j(x) - \mu(x))^2}{m(x)}.$$

2.1.7.2 Estimativas de componentes de variâncias negativas

Em análise de variância é comum decompor a variância em componentes que modelam a variabilidade dos vários aspectos aleatórios envolvidos na análise de um fenômeno. Frequentemente, o experimentador se depara com a ocorrência de estimativas negativas de tais componentes que são positivos. Uma das técnicas usuais na estimação dos componentes das variâncias é o método dos momentos, este método consiste basicamente em igualar os quadrados médios amostrais com suas esperanças (BARBIN, 1998). Tal procedimento pode apresentar, em diversas situações, estimativas negativas para os componentes da variância.

Nelder (1954) sugeriu que estimativas negativas ocorrem em razão do fato de que o modelo assumido está incorreto. Mas, supondo que o modelo estatístico assumido está correto, uma explicação para tal ocorrência seria que o ruído estatístico poderia ter afetado drasticamente a inferência.

Thompson (1962) e Thompson e Moore (1963) fazem uma discussão sobre tal problema, assumindo que o modelo estatístico está correto e que uma mudança no processo de estimação pode, de forma satisfatória, contornar este problema. Algo semelhante já havia sido feito por Herbach (1959) que usou o princípio da máxima verossimilhança para estimar os componentes da variância.

A teoria da regressão isotônica pode contornar este problema, mas antes de expor tal exemplo, convém apresentar alguns resultados discutidos em Rao (2001). Considere um vetor de observações:

$$\mathbf{X} = \begin{bmatrix} x_{11} \\ x_{12} \\ \vdots \\ x_{1J} \\ x_{21} \\ x_{22} \\ \vdots \\ x_{2J} \\ \vdots \\ x_{I1} \\ x_{I2} \\ \vdots \\ x_{IJ} \end{bmatrix}_{IJ \times 1},$$

e a transformação ortogonal

$$\mathbf{B} = \begin{bmatrix} \frac{1}{\sqrt{n}} & \frac{1}{\sqrt{n}} & \cdots & \frac{1}{\sqrt{n}} \\ b_{21} & b_{22} & \cdots & b_{2n} \\ b_{31} & b_{32} & \cdots & b_{3n} \\ \vdots & \vdots & \cdots & \vdots \\ b_{n1} & b_{n2} & \cdots & b_{nn} \end{bmatrix}.$$

O produto interno entre quaisquer duas linhas de \mathbf{B} é zero, isso implica que, exceto a primeira linha, todas as outras devem ter soma dos seus elementos

iguais a zero. Por exemplo, considerando o produto entre a segunda e primeira linha, tem-se:

$$\begin{aligned} \frac{1}{\sqrt{n}} \times b_{21} + \frac{1}{\sqrt{n}} \times b_{22} + \cdots + \frac{1}{\sqrt{n}} \times b_{2n} &= 0 \\ \Rightarrow \frac{1}{\sqrt{n}} (b_{21} + b_{22} + \cdots + b_{2n}) &= 0 \\ \Rightarrow b_{21} + b_{22} + \cdots + b_{2n} &= 0. \end{aligned}$$

Então,

$$\mathbf{W} = \mathbf{B}\mathbf{X} = \begin{bmatrix} \frac{1}{\sqrt{n}} & \frac{1}{\sqrt{n}} & \cdots & \frac{1}{\sqrt{n}} \\ b_{21} & b_{22} & \cdots & b_{2n} \\ b_{31} & b_{32} & \cdots & b_{3n} \\ \vdots & \vdots & \cdots & \vdots \\ b_{n1} & b_{n2} & \cdots & b_{nn} \end{bmatrix} \begin{bmatrix} x_{11} \\ x_{12} \\ \vdots \\ x_{1J} \\ x_{21} \\ x_{22} \\ \vdots \\ x_{2J} \\ \vdots \\ x_{I1} \\ x_{I2} \\ \vdots \\ x_{IJ} \end{bmatrix} = \begin{bmatrix} \sqrt{n}\bar{x} \\ w_{12} \\ \vdots \\ w_{1J} \\ w_{21} \\ w_{22} \\ \vdots \\ w_{2J} \\ \vdots \\ w_{I1} \\ w_{I2} \\ \vdots \\ w_{IJ} \end{bmatrix}.$$

Observe que a primeira coordenada de \mathbf{W} é a média aritmética do vetor \mathbf{X} . Por

outro lado, considerando um vetor de constantes $\mu = [\mu \ \mu \ \dots \ \mu]'$, tem-se:

$$\mathbf{B}\mu = \begin{bmatrix} \frac{1}{\sqrt{n}} & \frac{1}{\sqrt{n}} & \dots & \frac{1}{\sqrt{n}} \\ b_{21} & b_{22} & \dots & b_{2n} \\ b_{31} & b_{32} & \dots & b_{3n} \\ \vdots & \vdots & \dots & \vdots \\ b_{n1} & b_{n2} & \dots & b_{nn} \end{bmatrix} \begin{bmatrix} \mu \\ \mu \\ \mu \\ \vdots \\ \mu \end{bmatrix} = \begin{bmatrix} \sqrt{n}\mu \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix}.$$

Dessa maneira,

$$\mathbf{B}(\mathbf{X} - \mu) = \mathbf{B}\mathbf{X} - \mathbf{B}\mu = \mathbf{W} - \mathbf{B}\mu = \begin{bmatrix} \sqrt{n}\bar{x} - \sqrt{n}\mu \\ w_{12} \\ \vdots \\ w_{1J} \\ w_{21} \\ w_{22} \\ \vdots \\ w_{2J} \\ \vdots \\ w_{I1} \\ w_{I2} \\ \vdots \\ w_{IJ} \end{bmatrix}.$$

A norma quadrática do vetor $\mathbf{X} - \mu$ é:

$$\|\mathbf{X} - \mu\|^2 = \sum_j \sum_i (x_{ij} - \mu)^2,$$

e a norma quadrática do vetor $\mathbf{B}(\mathbf{X} - \mu)$ é:

$$\|\mathbf{B}(\mathbf{X} - \mu)\|^2 = (\sqrt{n}\bar{x} - \sqrt{n}\mu)^2 + \sum_{(i,j) \neq (1,1)} \sum w_{ij}^2.$$

Mas, uma transformação ortogonal preserva a norma de vetores, ou seja:

$$\sum_j \sum_i (x_{ij} - \mu)^2 = (\sqrt{n}\bar{x} - \sqrt{n}\mu)^2 + \sum_{(i,j) \neq (1,1)} \sum w_{ij}^2. \quad (2.29)$$

Agora, do ponto de vista Estatístico, pode-se considerar o vetor \mathbf{X} como um vetor onde cada componente é uma variável aleatória seguindo o seguinte modelo:

$$x_{ij} = \mu + a_i + \varepsilon_{ij} \quad i = 1, 2, \dots, I \quad e \quad j = 1, 2, \dots, J, \quad (2.30)$$

em que μ é uma constante, $\{a_i\}$ e $\{\varepsilon_{ij}\}$ são variáveis aleatórias normalmente e independentemente distribuídas com $E[a_i] = E[\varepsilon_{ij}] = 0$, $Var[a_i] = \sigma_a^2$ e $Var[\varepsilon_{ij}] = \sigma^2$, ou seja, $a_i \sim N(0, \sigma_a^2)$ e $\varepsilon_{ij} \sim N(0, \sigma^2)$.

A distribuição conjunta dos x_{ij} é uma normal multivariada com matriz de variância-covariância:

$$\Sigma = \begin{bmatrix} \Sigma_1 & 0 & \cdots & 0 \\ 0 & \Sigma_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \Sigma_I \end{bmatrix}.$$

Como são os elementos da matriz Σ_i ? Para ver isto são necessários os

seguintes cálculos:

$$\begin{aligned}
 \text{cov}[x_{ij}, x_{is}] &= E[x_{ij}x_{is}] - E[x_{ij}]E[x_{is}] \\
 &= E[(\mu + a_i + \varepsilon_{ij})(\mu + a_i + \varepsilon_{is})] - \mu^2 \\
 &= E[(\mu + a_i + \varepsilon_{ij})(\mu + a_i + \varepsilon_{is})] - \mu^2 \\
 &= E[\mu^2 + 2\mu a_i + \mu\varepsilon_{is} + a_i^2 + a_i\varepsilon_{is} + \varepsilon_{ij}\mu + \varepsilon_{ij}a_i + \varepsilon_{ij}\varepsilon_{is}] - \mu^2 \\
 &= \mu^2 + E[a_i^2] - \mu^2 \\
 &= \sigma_a^2, \quad (j \neq s),
 \end{aligned}$$

e

$$\begin{aligned}
 \text{cov}[x_{ij}, x_{ij}] &= \text{var}[x_{ij}] \\
 &= \text{var}[\mu + a_i + \varepsilon_{ij}] \\
 &= \text{var}[a_i] + \text{var}[\varepsilon_{ij}] \\
 &= \sigma_a^2 + \sigma^2.
 \end{aligned}$$

Logo, para todo i , tem-se:

$$\Sigma_i = \begin{bmatrix} \sigma_a^2 + \sigma^2 & \sigma_a^2 & \cdots & \sigma_a^2 \\ \sigma_a^2 & \sigma_a^2 + \sigma^2 & \cdots & \sigma_a^2 \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_a^2 & \sigma_a^2 & \cdots & \sigma_a^2 + \sigma^2 \end{bmatrix}$$

isto é,

$$\Sigma_i = \sigma_a^2 \mathbf{M} + \sigma^2 \mathbf{I},$$

em que \mathbf{M} é uma matriz com todos os elementos iguais a 1 de tamanho $J \times J$.

Considere agora, uma matriz ortogonal \mathbf{A} , tal que:

$$\mathbf{A} = \begin{bmatrix} \mathbf{A}_1 & 0 & \cdots & 0 \\ 0 & \mathbf{A}_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \mathbf{A}_I \end{bmatrix}_{n \times n},$$

em que:

$$\mathbf{A}_i = \begin{bmatrix} \frac{1}{\sqrt{J}} & \frac{1}{\sqrt{J}} & \cdots & \frac{1}{\sqrt{J}} \\ a_{21} & a_{22} & \cdots & a_{2J} \\ a_{31} & a_{32} & \cdots & a_{3J} \\ \vdots & \vdots & \cdots & \vdots \\ a_{J1} & a_{J2} & \cdots & a_{JJ} \end{bmatrix}$$

é ortogonal. Seja então $\mathbf{Z} = \mathbf{A}\mathbf{X}$, a matriz de variância-covariância de \mathbf{Z} é:

$$\text{cov}[\mathbf{Z}] = \text{cov}[\mathbf{A}\mathbf{X}] = \mathbf{A} \text{cov}[\mathbf{X}] \mathbf{A}' = \mathbf{A}\Sigma\mathbf{A}'.$$

Mas,

$$\begin{aligned}
\mathbf{A}_i \boldsymbol{\Sigma}_i \mathbf{A}'_i &= \mathbf{A}_i [\sigma_a^2 \mathbf{M} + \sigma^2 \mathbf{I}] \mathbf{A}'_i \\
&= \mathbf{A}_i \sigma_a^2 \mathbf{M} \mathbf{A}'_i + \mathbf{A}_i \sigma^2 \mathbf{I} \mathbf{A}'_i \\
&= \sigma_a^2 \mathbf{A}_i \mathbf{M} \mathbf{A}'_i + \sigma^2 \mathbf{A}_i \mathbf{I} \mathbf{A}'_i \\
&= \sigma_a^2 \mathbf{A}_i \mathbf{M} \mathbf{A}'_i + \sigma^2 \mathbf{I} \\
&= \sigma_a^2 \begin{bmatrix} \frac{1}{\sqrt{J}} & \frac{1}{\sqrt{J}} & \cdots & \frac{1}{\sqrt{J}} \\ a_{21} & a_{22} & \cdots & a_{2J} \\ a_{31} & a_{32} & \cdots & a_{3J} \\ \vdots & \vdots & \cdots & \vdots \\ a_{J1} & a_{J2} & \cdots & a_{JJ} \end{bmatrix} \begin{bmatrix} 1 & 1 & \cdots & 1 \\ 1 & 1 & \cdots & 1 \\ 1 & 1 & \cdots & 1 \\ \vdots & \vdots & \cdots & \vdots \\ 1 & 1 & \cdots & 1 \end{bmatrix} \mathbf{A}'_i + \sigma^2 \mathbf{I} \\
&= \sigma_a^2 \begin{bmatrix} \sqrt{J} & \sqrt{J} & \cdots & \sqrt{J} \\ 0 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & 0 \end{bmatrix} \begin{bmatrix} \frac{1}{\sqrt{J}} & a_{21} & a_{31} & \cdots & a_{J1} \\ \frac{1}{\sqrt{J}} & a_{22} & a_{32} & \cdots & a_{J2} \\ \vdots & \vdots & \vdots & \vdots & \vdots \\ \frac{1}{\sqrt{J}} & a_{2J} & a_{3J} & \cdots & a_{JJ} \end{bmatrix} + \sigma^2 \mathbf{I} \\
&= \sigma_a^2 \begin{bmatrix} J & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & 0 \end{bmatrix} + \sigma^2 \mathbf{I} \\
&= \begin{bmatrix} \sigma_a^2 J + \sigma^2 & 0 & \cdots & 0 \\ 0 & \sigma^2 & \cdots & 0 \\ \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & \sigma^2 \end{bmatrix}_{J \times J} .
\end{aligned}$$

Assim,

$$\Sigma^* = \text{cov} [\mathbf{Z}] = \begin{bmatrix} \Sigma_1^* & 0 & \cdots & 0 \\ 0 & \Sigma_2^* & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \Sigma_I^* \end{bmatrix}_{n \times n},$$

em que:

$$\Sigma_i^* = \begin{bmatrix} \sigma_a^2 J + \sigma^2 & 0 & \cdots & 0 \\ 0 & \sigma^2 & \cdots & 0 \\ \vdots & \vdots & \cdots & \vdots \\ 0 & 0 & \cdots & \sigma^2 \end{bmatrix}_{J \times J}.$$

Agora, observe que:

$$\begin{aligned} E[z_{i1}] &= E \left[\frac{1}{\sqrt{J}} x_{i1} + \cdots + \frac{1}{\sqrt{J}} x_{iJ} \right] \\ &= \frac{1}{\sqrt{J}} E[x_{i1}] + \cdots + \frac{1}{\sqrt{J}} E[x_{iJ}] \\ &= \frac{1}{\sqrt{J}} \mu + \cdots + \frac{1}{\sqrt{J}} \mu \\ &= \frac{J\mu}{\sqrt{J}} \\ &= \sqrt{J}\mu. \end{aligned}$$

Para $j > 1$:

$$\begin{aligned} E[z_{ij}] &= E[a_{s,l} x_{i1} + \cdots + a_{s,l+J} x_{iJ}] \\ &= \left(\sum_{l+J} a_{s,l+J} \right) \mu \\ &= 0. \end{aligned}$$

Note que a matriz de variância-covariância de \mathbf{Z} é diagonal, indicando que as componentes deste vetor são não-correlacionadas. Mas toda transformação linear invertível preserva a normalidade (FERREIRA, 2008), em particular uma transformação ortogonal. Para o caso normal, a não correlação implica em independência (MOOD; GRAYBILL; BOES, 1974). Assim, pode-se concluir que $z_{i1} \sim N(\sqrt{J}\mu, \sigma_a^2 J + \sigma^2)$ e $z_{ij} \sim N(0, \sigma^2)$ com $j > 1$, e além disso, z_{i1} e z_{ij} são independentes, para $j > 1$.

Considere agora o vetor \mathbf{Z}^* como sendo um vetor com as mesmas componentes do vetor \mathbf{Z} só que com os elementos que tem mesma distribuição agrupados da seguinte forma:

$$\mathbf{Z}^* = \begin{bmatrix} \mathbf{Z}_1^* \\ \mathbf{Z}_2^* \end{bmatrix},$$

em que $\mathbf{Z}_1^* = [z_{11} \ z_{21} \ \cdots \ z_{I1}]'$ e \mathbf{Z}_2^* tem os demais elementos. Assim, a matriz de covariância de \mathbf{Z}^* é:

$$\text{cov}[\mathbf{Z}^*] = \begin{bmatrix} \mathbf{D}_1 & \mathbf{0} \\ \mathbf{0} & \mathbf{D}_2 \end{bmatrix},$$

em que $\mathbf{D}_1 = (J\sigma_a^2 + \sigma^2) \mathbf{I}_I$ e $\mathbf{D}_2 = \sigma^2 \mathbf{I}_{I(J-1)}$. Pode-se realizar uma nova transformação ortogonal $\tilde{\mathbf{Z}} = \mathbf{C}\mathbf{Z}^*$, com:

$$\mathbf{C} = \begin{bmatrix} \mathbf{H}_I & \mathbf{0} \\ \mathbf{0} & \mathbf{I} \end{bmatrix}_{n \times n}$$

em que:

$$\mathbf{H}_I = \begin{bmatrix} \frac{1}{\sqrt{I}} & \frac{1}{\sqrt{I}} & \cdots & \frac{1}{\sqrt{I}} \\ h_{21} & h_{22} & \cdots & h_{2I} \\ \vdots & \vdots & \cdots & \vdots \\ h_{I1} & h_{I2} & \cdots & h_{II} \end{bmatrix}_{I \times I} .$$

A matriz de variância-covariância de $\tilde{\mathbf{Z}}$ é:

$$\begin{aligned} \text{cov} [\tilde{\mathbf{Z}}] &= \mathbf{C} \text{cov} [\mathbf{Z}^*] \mathbf{C}' = \begin{bmatrix} \mathbf{H}_I & \mathbf{0} \\ \mathbf{0} & \mathbf{I} \end{bmatrix} \begin{bmatrix} \mathbf{D}_1 & \mathbf{0} \\ \mathbf{0} & \mathbf{D}_2 \end{bmatrix} \begin{bmatrix} \mathbf{H}_I' & \mathbf{0} \\ \mathbf{0} & \mathbf{I} \end{bmatrix} \\ &= \begin{bmatrix} \mathbf{H}_I \mathbf{D}_1 & \mathbf{0} \\ \mathbf{0} & \mathbf{D}_2 \end{bmatrix} \begin{bmatrix} \mathbf{H}_I' & \mathbf{0} \\ \mathbf{0} & \mathbf{I} \end{bmatrix} = \begin{bmatrix} \mathbf{H}_I \mathbf{D}_1 \mathbf{H}_I' & \mathbf{0} \\ \mathbf{0} & \mathbf{D}_2 \end{bmatrix} \\ &= \begin{bmatrix} \mathbf{H}_I [(J\sigma_a^2 + \sigma^2) \mathbf{I}] \mathbf{H}_I' & \mathbf{0} \\ \mathbf{0} & \mathbf{D}_2 \end{bmatrix} \\ &= \begin{bmatrix} J\sigma_a^2 \mathbf{H}_I \mathbf{I} \mathbf{H}_I' + \sigma^2 \mathbf{H}_I \mathbf{I} \mathbf{H}_I' & \mathbf{0} \\ \mathbf{0} & \mathbf{D}_2 \end{bmatrix} \\ &= \begin{bmatrix} J\sigma_a^2 \mathbf{I} + \sigma^2 \mathbf{I} & \mathbf{0} \\ \mathbf{0} & \mathbf{D}_2 \end{bmatrix} = \begin{bmatrix} \mathbf{D}_1 & \mathbf{0} \\ \mathbf{0} & \mathbf{D}_2 \end{bmatrix} \\ &= \text{cov} [\mathbf{Z}^*], \end{aligned}$$

ou seja, a transformação ortogonal \mathbf{C} preservou as variâncias de todos os componentes do vetor \mathbf{Z}^* e consequentemente do vetor \mathbf{Z} . Agora, as esperanças dos componentes do vetor $\tilde{\mathbf{Z}}$ são:

$$\begin{aligned}
E[\tilde{z}_{11}] &= E\left[\frac{1}{\sqrt{I}}z_{11} + \frac{1}{\sqrt{I}}z_{21} + \cdots + \frac{1}{\sqrt{I}}z_{I1}\right] \\
&= \frac{1}{\sqrt{I}}(E[z_{11}] + E[z_{21}] + \cdots + E[z_{I1}]) \\
&= \frac{1}{\sqrt{I}}(\sqrt{J}\mu + \sqrt{J}\mu + \cdots + \sqrt{J}\mu) \\
&= \frac{1}{\sqrt{I}}I\sqrt{J}\mu \\
&= \sqrt{I}\sqrt{J}\mu \\
&= \sqrt{n}\mu.
\end{aligned}$$

Para $i > 1$, tem-se:

$$\begin{aligned}
E[\tilde{z}_{i1}] &= E[h_{i1}z_{11} + h_{i2}z_{21} + \cdots + h_{iI}z_{I1}] \\
&= h_{i1}E[z_{11}] + h_{i2}E[z_{21}] + \cdots + h_{iI}E[z_{I1}] \\
&= h_{i1}\sqrt{J}\mu + h_{i2}\sqrt{J}\mu + \cdots + h_{iI}\sqrt{J}\mu \\
&= \sqrt{J}\mu \underbrace{\left(\sum_{s=1}^I h_{is}\right)}_{=0} \\
&= 0.
\end{aligned}$$

Como a transformação \mathbf{C} não altera os outros elementos do vetor \mathbf{Z}^* , tem-se que as componentes do vetor $\tilde{\mathbf{Z}}$ são as variáveis aleatórias:

$$\begin{aligned}
\tilde{z}_{11} &\sim N(\sqrt{n}\mu, J\sigma_a^2 + \sigma^2) \\
\tilde{z}_{i1} &\sim N(0, J\sigma_a^2 + \sigma^2), \text{ para } i > 1 \\
\tilde{z}_{ij} &\sim N(0, \sigma^2), \text{ para } j = 2, 3, \dots, J \text{ e } i = 1, 2, \dots, I.
\end{aligned}$$

Pela matriz de variância-covariância de $\tilde{\mathbf{Z}}$ e a normalidade das suas componentes, pode-se concluir que tais variáveis aleatórias são independentes. Observe que todo este procedimento implica em uma fatoração na verossimilhança que facilitará o cálculo dos estimadores de máxima verossimilhança para as variâncias.

Considerando $\lambda_1 = J\sigma_a^2 + \sigma^2$ e $\lambda_2 = \sigma^2$, a verossimilhança para μ , λ_1 e λ_2 é:

$$\begin{aligned}
 L &= L(\mu, \lambda_1, \lambda_2) \\
 &= \frac{1}{\sqrt{2\pi}} \lambda_1^{-\frac{1}{2}} e^{-\frac{1}{2\lambda_1} (\tilde{z}_{11} - \sqrt{n}\mu)^2} \times \prod_{i=2}^I \left\{ \frac{1}{\sqrt{2\pi}} \lambda_1^{-\frac{1}{2}} e^{-\frac{1}{2\lambda_1} \tilde{z}_{i1}^2} \right\} \\
 &\times \prod_{i=1}^I \prod_{j=2}^J \left\{ \frac{1}{\sqrt{2\pi}} \lambda_2^{-\frac{1}{2}} e^{-\frac{1}{2\lambda_2} \tilde{z}_{ij}^2} \right\} \\
 &= \left(\frac{1}{\sqrt{2\pi}} \right)^n \lambda_1^{-\frac{1}{2}} e^{-\frac{1}{2\lambda_1} (\tilde{z}_{11} - \sqrt{n}\mu)^2} \\
 &\times \lambda_1^{-\frac{I-1}{2}} e^{-\frac{1}{2\lambda_1} \sum_{i=2}^I \tilde{z}_{i1}^2} \times \lambda_2^{-\frac{I(J-1)}{2}} e^{-\frac{1}{2\lambda_2} \sum_{i=1}^I \sum_{j=2}^J \tilde{z}_{ij}^2} \\
 &= \left(\frac{1}{\sqrt{2\pi}} \right)^n \lambda_1^{-\frac{I}{2}} \times e^{-\frac{1}{2\lambda_1} \left\{ (\tilde{z}_{11} - \sqrt{n}\mu)^2 + \sum_{i=2}^I \tilde{z}_{i1}^2 \right\}} \times \lambda_2^{-\frac{I(J-1)}{2}} e^{-\frac{1}{2\lambda_2} \sum_{i=1}^I \sum_{j=2}^J \tilde{z}_{ij}^2}.
 \end{aligned}$$

Observe que é natural querer estimar μ , λ_1 e λ_2 por máxima verossimilhança sujeito às restrições $\lambda_1 \geq \lambda_2 \geq 0$. A fatoração efetuada permite estimar facilmente μ :

$$\frac{\partial L}{\partial \mu} = 0 \Rightarrow 2(\tilde{z}_{11} - \sqrt{n}\hat{\mu})\sqrt{n} = 0 \Rightarrow \hat{\mu} = \frac{\tilde{z}_{11}}{\sqrt{n}}.$$

Substituindo este resultado na verossimilhança, e lembrando que $n = IJ = I + I(J-1)$, tem-se:

$$L = \left(\frac{1}{\sqrt{2\pi}} \right)^{\frac{I}{2}} \lambda_1^{-\frac{I}{2}} e^{-\frac{1}{2\lambda_1} \sum_{i=1}^I \zeta_i^2} \left(\frac{1}{\sqrt{2\pi}} \right)^{\frac{I(J-1)}{2}} \lambda_2^{-\frac{I(J-1)}{2}} e^{-\frac{1}{2\lambda_2} \sum_{i=1}^I \sum_{j=2}^J \tilde{z}_{ij}^2}, \quad (2.31)$$

em que, $\zeta_i = \tilde{z}_{i1}$ para $i = 2, 3, \dots, I$ e $\zeta_1 = 0$.

A verossimilhança em 2.31 é precisamente aquela que seria obtida, se amostras independentes de tamanhos I e $I(J-1)$ fossem tomadas de duas distribuições normais com médias 0 e variâncias λ_1 e λ_2 respectivamente, com valores observados $\zeta_1, \zeta_2, \dots, \zeta_I$ da primeira e valores \tilde{z}_{ij} , $i = 1, 2, \dots, I$, $j = 2, 3, \dots, J$ para a segunda. Sejam $X = \{1, 2\}$, $m(1) = I$, $m(2) = I(J-1)$, $\lambda_1 = \sigma^2(1)$, $\lambda_2 = \sigma^2(2)$ e a ordem simples $1 \succeq 2$ a quase ordem sobre X . Pela Afirmação 2.8, a estimativa de máxima verossimilhança de $\lambda_x = \sigma^2(x)$, $x = 1, 2$, sujeito a $\lambda_1 \geq \lambda_2$ para $1 \succeq 2$, é a regressão isotônica $t^*(\cdot)$ de $\tilde{t}(\cdot)$ com pesos $m(\cdot)$ em que:

$$\tilde{t}(1) = \frac{\sum_{i=1}^I \zeta_i^2}{I}$$

e

$$\tilde{t}(2) = \frac{\sum_{j=2}^J \sum_{i=1}^I \tilde{z}_{ij}^2}{I(J-1)}.$$

As estimativas obtidas por este processo jamais serão negativas, e satisfazem a condição $\lambda_1 \geq \lambda_2 > 0$.

A condição $1 \succeq 2$ pode ser entendida como uma comparação genérica, 1 e 2 não são necessariamente números reais expostos em uma reta, aqui, estes valores podem representar, por exemplo, dois tratamentos qualitativos.

Estes resultados podem ser generalizados para modelos mais complexos da análise de componentes de variância em que se tem uma ordenação parcial.

2.1.7.3 O grafo de um delineamento

Os grafos serão usados para indicar a ordenação parcial em um conjunto de pontos. Um *grafo* é um diagrama de pontos consistindo de um conjunto finito de pontos e linhas, as linhas descrevem a ordem entre os pontos (Figura 19).

Uma *árvore enraizada* é um grafo com um ponto especial, chamado raiz, tal que, para cada ponto do grafo existe um único caminho, consistindo alternativamente de pontos e linhas, conectando o ponto dado e a raiz. A Figura 19 é um exemplo de árvore enraizada, já a Figura 20 não, pois para alguns pontos existem mais de um caminho até a raiz.

O *predecessor* de um ponto P é o ponto mais próximo a P que está no caminho de P e a raiz. Na Figura 19 tem-se uma representação de um grafo com 9 pontos e 8 linhas. A raiz deste grafo é o ponto 1 que é o predecessor dos pontos 2 e 3, os pontos 4, 5 e 6 tem como predecessor o ponto 2, os pontos 7 e 8 tem como predecessor o ponto 3, o ponto 9 tem como predecessor o ponto 8.

Thompson e Moore (1963), no contexto da análise de variância em estatística experimental, consideram os pontos de um grafo como Quadrados Médios (ou Quadrados Médios Esperados). Estes autores representam três delineamentos por grafos, tal representação se encontra na Figura 20. Detalhes destes delineamentos podem ser encontrados em Searle (1987).

A ordenação considerada na Figura 20 está representada pelo fato de que um Quadrado Médio Esperado é maior ou igual do que outro se o caminho conectando ele e a raiz passa através do outro Quadrado Médio. Por exemplo, em (2) na Figura 20 a raiz é $E(QM_E)$, com $E(QM_A) \leq E(QM_{AB}) \leq E(QM_E)$ e $E(QM_B) \leq E(QM_{AB}) \leq E(QM_E)$, a ordenação é parcial pois $E(QM_A)$ e $E(QM_B)$ são não comparáveis.

Thompson e Moore (1963) resolveram o problema de estimar $E(QM's)$

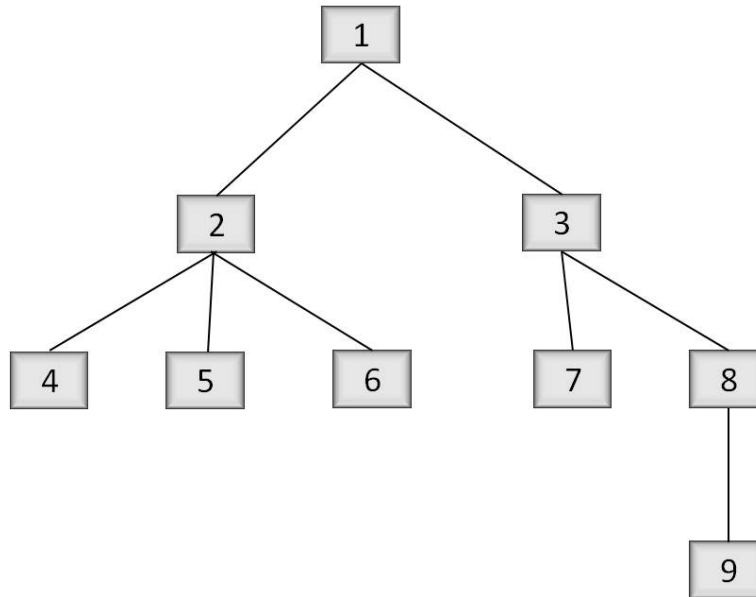


Figura 19 Representação de um grafo

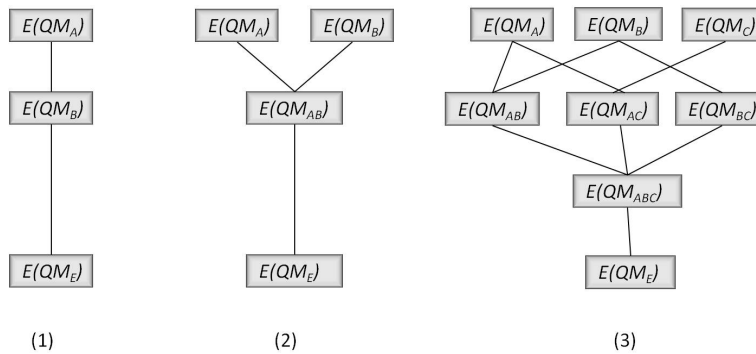


Figura 20 Grafos Lineares para: (1) Classificação Hierárquica *two-fold*, (2) Classificação *two-way* com $n > 1$ observações por célula e (3) classificação *three-way* com $n > 1$ observações por células. A notação $E(QM_X)$ significa: Esperança do Quadrado Médio do fator X e $E(QM_{XY})$, Esperança do Quadrado Médio da Interação entre os fatores X e Y , $E(QM_E)$, Esperança do Quadrado Médio do Erro

seguindo a ordenação parcial em cada delineamento apenas para o caso de uma ordenação cujo grafo segue uma árvore enraizada (Figura 20, parte (1) e (2), por exemplo).

O problema geral, para qualquer situação de ordenação parcial possível, pode ser pensando como um problema de regressão isotônica, ou seja, deseja-se estimar os $E(QM's)$ a partir de alguma ordenação parcial particular a cada delineamento. O grande entrave é que não existe um algoritmo eficiente para resolver este problema para qualquer situação, entretanto, avanços tem sido obtidos nesta área.

Leeuw, Hornik e Mair (2009) têm mostrado soluções para alguns tipos de ordenação que se encaixam uma boa parte dos problemas de análise de variância. A Figura 21 ilustra quatro tipos de ordenações consideradas por estes autores. Um pacote em linguagem R (software estatístico R) é desenvolvido, com implementação de algoritmos que visam resolver situações mais gerais como as apresentadas aqui.

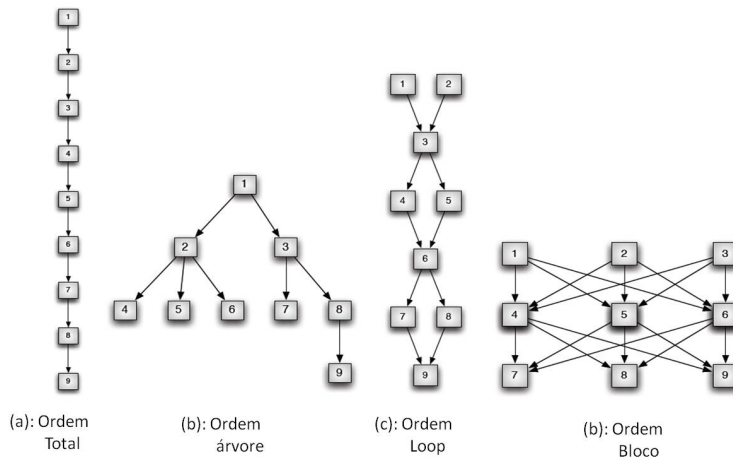


Figura 21 Grafos para diversos tipos de ordenação

Naturalmente, as estimativa obtidas via regressão isotônica seguirão a or-

denação esperada nos parâmetros (Quadrados Médios Esperados), resolvendo, assim, o problema de estimativas de componentes de variância negativos. Além disso, os valores encontrados serão as estimativas de máxima verossimilhança sob a restrição imposta pelo grafo considerado.

2.1.8 Estimação bayesiana e regressão isotônica

Um dos objetivos deste trabalho é a utilização de conceitos bayesianos em ensaios de dose-resposta por meio de delineamentos sequenciais.

Suponha que para cada $x \in X$ o observador considera $m(x)$ observações em uma distribuição com densidade da forma 2.26 com $\theta = \mu(x)$, $h = \nu(x)$, em que $\mu(\cdot)$ é desconhecida e $\nu(\cdot)$ conhecida, ou seja:

$$f(y; \mu(x), \nu(x)) = \exp \{ (\Phi(\mu(x)) + [y - \mu(x)]\varphi(\mu(x))) \nu(x) \}. \quad (2.32)$$

Suponha que é conhecido, além disso, que $\mu(\cdot)$ é isotônica com respeito a uma quase ordem sobre X . O pesquisador poderá selecionar uma distribuição a priori para os parâmetros $\mu(x)$, para cada $x \in X$ de acordo com algum conhecimento prévio, como é comum em experimentos bayesianos (PAULINO; TURKMAN; MURTEIRA, 2003).

Considere uma densidade a priori, como uma função de um argumento θ com respeito a alguma medida $d_\rho(\theta)$, com parâmetros θ_0 e $0 < h_0$, da forma:

$$\pi(\theta) \propto \exp \{ [\Phi(\theta) + (\theta_0 - \theta)\varphi(\theta)] h_0 \}. \quad (2.33)$$

Afirmção 2.9 A função $J(\theta) = \Phi(\theta) + (\theta_0 - \theta)\varphi(\theta)$ assume seu máximo em $\theta = \theta_0$.

Demonstração:

$$\begin{aligned}\frac{dJ}{d\theta} &= \varphi(\theta) + \theta_0\varphi'(\theta) - \varphi(\theta) - \theta\varphi'(\theta) \\ &= \varphi'(\theta)(\theta_0 - \theta).\end{aligned}$$

Como, para qualquer θ , tem-se $\varphi'(\theta) > 0$ (φ' é a derivada segunda de Φ e Φ é estritamente convexa):

$$\varphi'(\theta)(\theta_0 - \theta) = 0 \Rightarrow \theta = \theta_0.$$

Da Afirmação 2.9 é evidente que θ_0 é a moda da densidade em 2.33. Se o parâmetro h_0 é grande, a densidade tem uma forma mais leptocúrtica sobre θ_0 ; $h_0 = 0$ fornece uma distribuição a priori uniforme (não informativa). O parâmetro h_0 será chamado *parâmetro de precisão*.

Suponha que para cada $x \in X$ o pesquisador considere $\mu_0(x)$ um valor provável de $\mu(x)$. São escolhidos para cada x um parâmetro de precisão $h_0(x)$ e uma densidade a priori com respeito a alguma medida d_p :

$$\exp\{(\Phi(\theta) + [\mu_0(x) - \theta]\varphi(\theta))h_0(x)\}.$$

Para a obtenção de uma priori para o vetor $\mu = (\mu(x_1), \mu(x_2), \dots, \mu(x_k))$, considera-se os $\mu(x)$ independentes. Dessa forma, a distribuição a priori para μ é:

$$\begin{aligned}
\pi[\mu] &= \pi[(\mu(x_1), \mu(x_2), \dots, \mu(x_k))] \\
&\propto \prod_{x \in X} \exp\{(\Phi(\mu(x)) + [\mu_0(x) - \mu(x)]\varphi(\mu(x)))h_0(x)\} \\
&\propto \exp\left\{\sum_{x \in X} ((\Phi(\mu(x)) + [\mu_0(x) - \mu(x)]\varphi(\mu(x)))h_0(x))\right\}.
\end{aligned}$$

Se $\mu(x)$ é suposta ser isotônica, o pesquisador deve selecionar uma função modal $\mu_0(x)$ que seja isotônica. Uma forma de incorporar a restrição de ordem na priori para o vetor μ é definir esta priori da forma:

1. Se $\mu(\cdot)$ for isotônica sobre X :

$$\pi[\mu] \propto \prod_{x \in X} \exp\{(\Phi(\mu(x)) + [\mu_0(x) - \mu(x)]\varphi(\mu(x)))h_0(x)\} \quad (2.34)$$

2. Se $\mu(\cdot)$ não for isotônica sobre X :

$$\pi[\mu] = 0.$$

Observe que $\pi[\mu]$ é proporcional ao núcleo de uma densidade, ou seja, $\pi[\mu]$ determina a probabilidade do vetor μ pertencer a uma determinada região. Como é de conhecimento prévio do experimento que $\mu(x)$ é isotônica sobre x , a probabilidade de μ não pertencer ao espaço das isotônicas é zero.

A moda da densidade em 2.34 é $\mu_0 = (\mu_0(x_1), \mu_0(x_2), \dots, \mu_0(x_k))$, e a função $\mathbf{h}_0 = (h_0(x_1), h_0(x_2), \dots, h_0(x_k))$ determina a concentração desta densidade ao redor da moda. Para $\mathbf{h}_0 = \mathbf{0}$ tem-se uma priori uniforme ou plana (note que se está falando de uma priori em um espaço k -dimensional).

Agora, suponha que para cada $x \in X$ o pesquisador tenha tomado $m(x)$ observações sob a densidade em 2.32 e a amostra observada representada pelo vetor $\mathbf{y} = (y_1(x_1), \dots, y_{m(x_1)}(x_1); \dots; y_1(x_k), \dots, y_{m(x_k)}(x_k))$ em que $y_j(x)$ representa a j -ésima observação tomada em x . Considere ainda o vetor precisão conhecido $\nu = (\nu(x_1), \nu(x_2), \dots, \nu(x_k))$. A verossimilhança para esta amostra é:

$$\begin{aligned}
L &= L(\mu; \mathbf{y}, \nu) \\
&\propto \prod_{x \in X} \prod_{j=1}^{m(x)} \exp \{ [\Phi(\mu(x)) + [y_j(x) - \mu(x)]\varphi(\mu(x))] \nu(x) \} \\
&\propto \exp \left(\sum_{x \in X} \sum_{j=1}^{m(x)} \{ [\Phi(\mu(x)) + [y_j(x) - \mu(x)]\varphi(\mu(x))] \nu(x) \} \right) \\
&\propto \exp \left(\sum_{x \in X} \{ [m(x)\Phi(\mu(x)) + m(x)\bar{y}(x)\varphi(\mu(x)) - m(x)\mu(x)\varphi(\mu(x))] \nu(x) \} \right) \\
&\propto \exp \left(\sum_{x \in X} \{ [\Phi(\mu(x)) + [\bar{y}(x) - \mu(x)]\varphi(\mu(x))] \nu(x)m(x) \} \right).
\end{aligned}$$

Utilizando-se do Teorema de Bayes, tem-se que a distribuição a posteriori para o vetor μ após a amostragem realizada é:

$$\begin{aligned}
\pi[\mu|\mathbf{y}] &\propto L(\mu; \mathbf{y}, \nu)\pi[\mu] \\
&\propto \exp \left(\sum_{x \in X} \{ [\Phi(\mu(x)) + [\bar{y}(x) - \mu(x)]\varphi(\mu(x))] \nu(x)m(x) \} \right) \\
&\times \exp \left(\sum_{x \in X} \{ [\Phi(\mu(x)) + [\mu_0(x) - \mu(x)]\varphi(\mu(x))] h_0(x) \} \right) \\
&\propto \exp \left(\sum_{x \in X} \{ [\Phi(\mu(x))(\nu(x)m(x) + h_0(x)) + (\bar{y}(x)\nu(x)m(x) \right. \\
&\quad \left. + \mu_0(x)h_0(x))\varphi(\mu(x)) - \mu(x)\varphi(\mu(x))(\nu(x)m(x) + h_0(x))] \} \right) \\
&\propto \exp \left\{ \sum_{x \in X} \left\{ \left[\Phi(\mu(x)) + \left(\frac{\bar{y}(x)\nu(x)m(x) + \mu_0(x)h_0(x)}{\nu(x)m(x) + h_0(x)} - \mu(x) \right) \varphi(\mu(x)) \right] \right. \right. \\
&\quad \left. \left. \times (\nu(x)m(x) + h_0(x)) \right\} \right\}
\end{aligned}$$

ou melhor, a distribuição a posteriori do vetor μ é:

1. Se $\mu(\cdot)$ for isotônica sobre X :

$$\pi[\mu | \mathbf{y}] \propto \exp \left\{ \sum_{x \in X} \{ [\Phi(\mu(x)) + (\tilde{\mu}(x) - \mu(x)) \varphi(\mu(x))] w(x) \} \right\} \quad (2.35)$$

2. Se $\mu(\cdot)$ não for isotônica sobre X :

$$\pi[\mu | \mathbf{y}] = 0,$$

em que,

$$\tilde{\mu}(x) = \frac{\bar{y}(x)\nu(x)m(x) + \mu_0(x)h_0(x)}{\nu(x)m(x) + h_0(x)}$$

e,

$$w(x) = \nu(x)m(x) + h_0(x).$$

A moda a posteriori em 2.35 maximiza a soma aparecendo no expoente na classe das funções isotônicas $\mu(x)$. Mas isso é precisamente o problema de valor extremo já solucionado no Teorema 2.13 (Equação 2.17) e a solução para o mesmo é a regressão isotônica de $\tilde{\mu}(x)$, denotada por $\tilde{\mu}^*(x)$ com pesos $w(x) = \nu(x)m(x) + h_0(x)$.

Os resultados discutidos nesta seção servem como demonstração para o Teorema 2.15:

Teorema 2.15 *Seja $X = \{x_1, x_2, \dots, x_k\}$ um conjunto finito, quase ordenado por \preceq . Para cada $x \in X$ associe uma distribuição tendo densidade como em 2.26 com $\theta = \mu(x)$ desconhecido, e com $h = \nu(x)$, conhecido. Considere uma amostra de variáveis aleatórias independentes tomadas destas distribuições com tamanhos $m(x) > 0$, $x \in X$. Considere o vetor de componentes aleatórios k -dimensional $\mu = (\mu(x_1), \mu(x_2), \dots, \mu(x_k))$ tendo distribuição a priori sobre*

a classe das funções isotônicas $\mu(x)$ com densidade como em 2.34, com moda isotônica $\mu_0 = (\mu_0(x_1), \mu_0(x_2), \dots, \mu_0(x_k))$ (especificada) e função de precisão $\mathbf{h}_0 = (h_0(x_1), h_0(x_2), \dots, h_0(x_k))$ (especificada). Então, a distribuição a posteriori de μ tem densidade na mesma família que a priori com moda dada por $\tilde{\mu}^* = (\tilde{\mu}^*(x_1), \tilde{\mu}^*(x_2), \dots, \tilde{\mu}^*(x_k))$, em que, cada $\tilde{\mu}^*(x)$ é a regressão isotônica com pesos $w(x) = \nu(x)m(x) + h_0(x)$, da média ponderada de $\bar{y}(x)$ e $\mu_0(x)$:

$$\tilde{\mu}(x) = \frac{\bar{y}(x)\nu(x)m(x) + \mu_0(x)h_0(x)}{\nu(x)m(x) + h_0(x)}, \quad x \in X.$$

Em particular, se $\mathbf{h}_0 = \mathbf{0}$ (priori plana) a moda a posteriori coincide com as Estimativas de Máxima Verossimilhança.

2.1.9 O cálculo da regressão isotônica para ordem simples utilizando o algoritmo PAVA

Até agora muitos resultados foram colocados sobre a regressão isotônica e a relação da mesma com os estimadores de máxima verossimilhança e parâmetros com restrição de ordem. Entretanto, resta a questão de como obtê-la explicitamente. Será discutido um algoritmo para o caso de ordenação simples no conjunto X , este é o caso mais comum. O problema de se calcular a regressão isotônica em relação a uma quase ordem qualquer ainda é um problema em aberto.

2.1.9.1 O diagrama de somas acumuladas

Uma interpretação gráfica da regressão isotônica é obtida com a seguinte construção: dada uma ordenação simples $x_1 \prec x_2 \prec \dots \prec x_n$ e uma função real

qualquer g em X , plotando as somas acumuladas

$$G_j = \sum_{i=1}^j g(x_i)w(x_i), \quad j = 1, 2, \dots, n$$

contra as somas acumuladas

$$W_j = \sum_{i=1}^j w(x_i), \quad j = 1, 2, \dots, n,$$

isso é, o gráfico no plano cartesiano dos pontos $P_j = (W_j, G_j)$, $j = 0, 1, 2, \dots, n$. Por definição $P_0 = (0, 0)$. Estes pontos constituem o diagrama de somas acumuladas (*cumulative sum diagram - CSD*) da função dada g com pesos w . A inclinação do segmento que une os pontos P_{j-1} e P_j é justamente $g(x_j)$, $j = 1, 2, \dots, n$:

$$tg\alpha = \frac{G_j - G_{j-1}}{W_j - W_{j-1}} = \frac{\sum_{i=1}^j g(x_i)w(x_i) - \sum_{i=1}^{j-1} g(x_i)w(x_i)}{\sum_{i=1}^j w(x_i) - \sum_{i=1}^{j-1} w(x_i)} = \frac{g(x_j)w(x_j)}{w(x_j)} = g(x_j).$$

Como pode ser observado na Figura 22.

Desta maneira, cada valor que $g(x_j)$ assume pode ser interpretado como a inclinação de um segmento de reta que une dois pontos no CSD. Para exemplificar, considere a Figura 23, nela o diagrama apresenta quatro segmentos de reta, conseqüentemente quatro inclinações a serem analisados e quatro valores para $g(x_i)$, ou seja, o CSD mostra geometricamente o comportamento de $g(x_1)$, $g(x_2)$, $g(x_3)$ e $g(x_4)$, de modo que é fácil comparar dois elementos consecutivos:

$$tg\alpha_1 > tg\alpha_2 \Rightarrow g(x_1) > g(x_2)$$

$$tg\alpha_2 > tg\alpha_3 \Rightarrow g(x_2) > g(x_3)$$

$$tg\alpha_3 < tg\alpha_4 \Rightarrow g(x_3) < g(x_4)$$

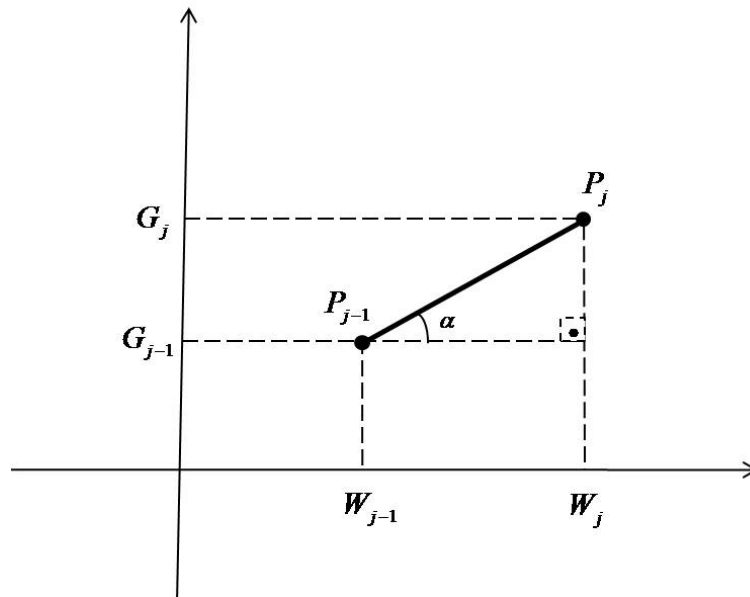


Figura 22 Analisando dois pontos consecutivos no Diagrama de Somas Acumuladas

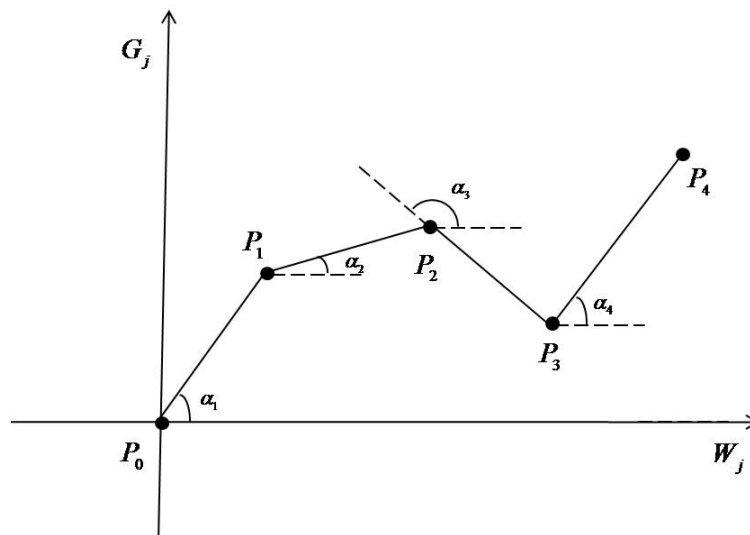


Figura 23 Interpretando os declives no Diagrama de Somas Acumuladas

Para uma função g isotônica, tem-se:

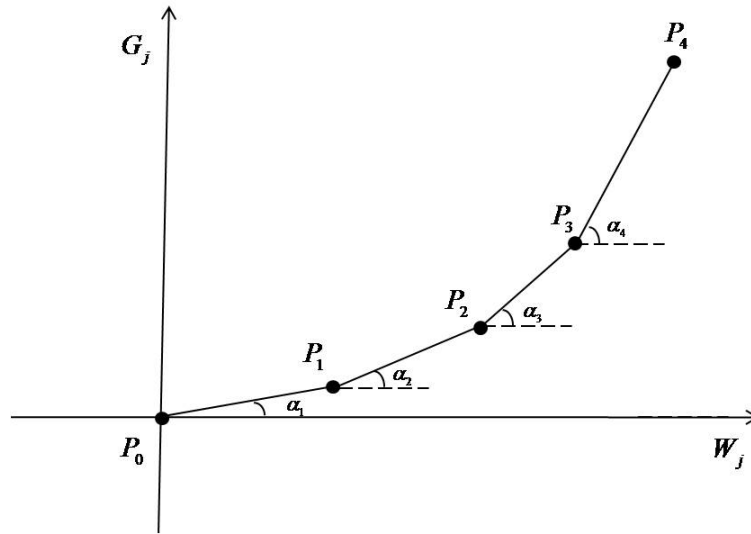


Figura 24 Diagrama de Somas Acumuladas para uma função isotônica

$$\operatorname{tg} \alpha_1 < \operatorname{tg} \alpha_2 \Leftrightarrow g(x_1) < g(x_2)$$

$$\operatorname{tg} \alpha_2 < \operatorname{tg} \alpha_3 \Leftrightarrow g(x_2) < g(x_3)$$

$$\operatorname{tg} \alpha_3 < \operatorname{tg} \alpha_4 \Leftrightarrow g(x_3) < g(x_4).$$

Neste caso, a reta definida por P_{i-1} e P_i está abaixo de todos os P_j , $j > 0$, e portanto, o CSD é uma função convexa (Figura 24).

Outra propriedade é que a inclinação da corda que une P_{i-1} a P_j com ($i \leq j$) é a média ponderada de g em relação aos pesos dados (Figura 25).

$$\operatorname{tg}(\alpha) = \frac{\sum_{r=1}^j g(x_r)w(x_r) - \sum_{r=1}^{i-1} g(x_r)w(x_r)}{\sum_{r=1}^j w(x_r) - \sum_{r=1}^{i-1} w(x_r)} = \frac{\sum_{r=i}^j g(x_r)w(x_r)}{\sum_{r=i}^j w(x_r)}.$$

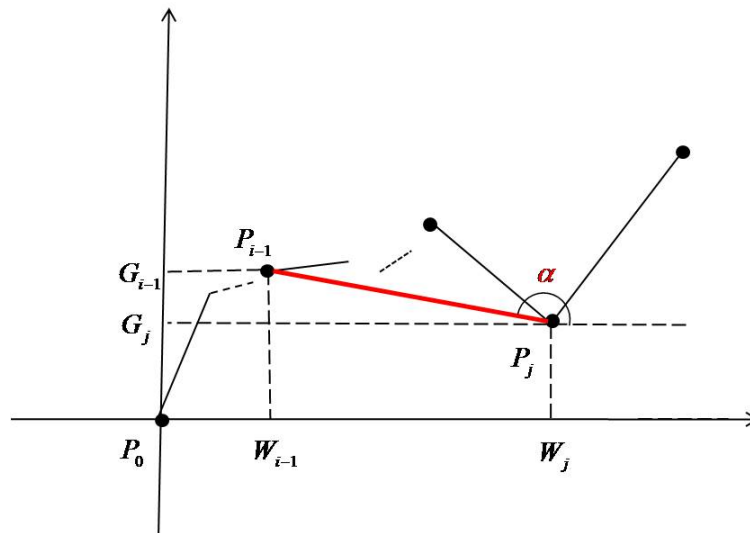


Figura 25 Interpretando o declive entre dois pontos quaisquer no Diagrama de Somas Acumuladas

2.1.9.2 O maior minorante convexo

Existem infinitos polígonos convexos com extremidade em P_0 e P_n que estão completamente abaixo do CSD. A figura 26 ilustra alguns deles. Cada um destes polígonos é uma função convexa e será denominada de minorante convexo. Será visto que a regressão isotônica, para o caso de uma ordenação simples, estará relacionada com o gráfico de uma destas funções.

O gráfico do supremo de todas as funções convexas cujos gráficos estão abaixo do CSD é denominado o maior minorante convexo (*greatest convex minorant - GCM*). Considere uma reta que intersecta o CSD tal que todo o CSD está sobre ou acima dela. Se ela intersecta o CSD em mais de um ponto, então o segmento unindo suas intersecções à esquerda e à direita é uma parte do gráfico do GCM. Tal fato permite a seguinte interpretação geométrica: Fixe um elástico no ponto P_0 e esticando este elástico por baixo do CSD fixe a outra extremidade no

ponto P_n . O polígono obtido pela sequência de retas é o GCM (Figura 26).

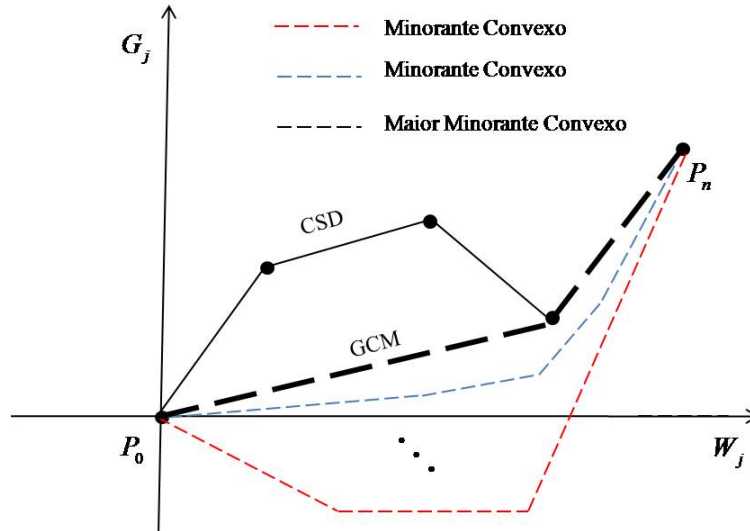


Figura 26 Representação do Maior Minorante Convexo - GCM

Uma outra relação entre o CSD e o GCM é que o CSD e GCM coincidem em P_n , isto é:

$$G_n^* = G_n. \quad (2.36)$$

Seja $\tilde{g}(x_j)$ a função definida como a inclinação do GCM no ponto P_j^* com abscissa:

$$W_j = \sum_{i=1}^j w(x_i).$$

$$\text{Declive em } P_j \text{ do CSD: } \frac{G_j - G_{j-1}}{W_j - W_{j-1}} = g(x_j)$$

$$\text{Declive em } P_j^* \text{ do GCM: } \frac{G_j^* - G_{j-1}^*}{W_j - W_{j-1}} = \tilde{g}(x_j), \quad j = 1, 2, 3, 4.$$

Considere o exemplo simples, utilizando os dados que estão descritos na Tabela 1, para ilustrar algebricamente os Diagrama de Somas Acumuladas e a

interpretação dos resultados para o maior minorante convexo.

Tabela 1 Exemplo do diagrama de somas acumuladas - CSD e do maior minorante convexo - GCM.

j	w_j	W_j	$g(x_j)$	G_j	G_j^*	$\tilde{g}(x_j)$
1	1	1	-2	-2	-2	-2
2	2	3	$\frac{5}{2}$	3	$-\frac{8}{5}$	$\frac{1}{5}$
3	3	6	$-\frac{4}{3}$	-1	-1	$\frac{1}{5}$
4	2	8	1	1	1	1

Se para algum i , P_{i-1}^* está estritamente abaixo de P_{i-1} , então as inclinações do GCM entre P_{i-1}^* para a esquerda e para a direita são as mesmas, isto é, a poligonal não tem vértices em P_{i-1}^* :

$$G_{i-1}^* < G_{i-1} \Rightarrow \tilde{g}_i^* = \tilde{g}_{i-1}^*, \quad i = 1, 2, \dots, n. \quad (2.37)$$

Finalmente, se $P_r = P_r^*$, $P_s = P_s^*$ se $P_r P_s = P_r^* P_s^*$ é um lado do GCM e $P_s P_t = P_j^* P_t^*$ é o lado direito adjacente, e se $r < j < s$, então a inclinação de $P_j P_s$ é menor do que a inclinação de $P_j^* P_s^*$. Se $s < j < t$, então a inclinação de $P_s P_j$ é maior do que a inclinação de $P_s^* P_j^*$. Isto é, se $\tilde{g}(x)$ tem um valor constante a para $x_r < x \leq x_s$ e o valor constante $b > a$ para $x_s < x < x_t$, então:

$$\frac{\sum_{i=j+1}^s g(x_i)w(x_i)}{\sum_{i=j+1}^s w(x_i)} = \frac{G(x_s) - G(x_j)}{W(x_s) - W(x_j)} \leq a \quad \text{para } j < s, \quad (2.38)$$

e

$$\frac{\sum_{i=s+1}^j g(x_i)w(x_i)}{\sum_{i=s+1}^j w(x_i)} = \frac{G(x_j) - G(x_s)}{W(x_j) - W(x_s)} \geq b \quad \text{para } s < j \leq t. \quad (2.39)$$

Em particular,

$$g(x_s) \leq \tilde{g}(x_s) = a < b = \tilde{g}(x_{s+1}) < g(x_{s+1}).$$

Afirmação 2.10 A inclinação à esquerda do GCM no ponto (W_i, P_i^*) , $\tilde{g}(x_i)$ é de fato a regressão isotônica $g^*(x_i)$.

A demonstração deste fato segue das propriedades descritas (BARLOW et al., 1972).

2.1.9.3 Um algoritmo para o cálculo da regressão isotônica

Dado g uma função qualquer em X munido de uma ordem simples, se g é isotônica sua regressão isotônica é a própria função. Caso contrário, para algum i , $g(x_{i-1}) > g(x_i)$, neste caso, denomina-se x_i como um violador. Neste caso altera-se o CSD ligando-se os pontos P_{i-2}^* e P_i^* por um segmento de reta. A observação fundamental é que o GCM deste novo CSD é igual ao anterior, isto é, a reta que liga P_{i-2}^* e P_i^* está abaixo deste segmento de reta unindo P_{i-2}^* e P_i^* . Este método de aproximações sucessivas do GCM pode ser descrito algebricamente como o algoritmo *Pool Adjacent Violators - PAVA*.

O algoritmo começa fazendo uma partição de $X = \{x_1, x_2, \dots, x_n\}$ em blocos unitários: $\{x_1\}, \{x_2\}, \dots, \{x_n\}$, e por um processo iterativo vai unindo estes blocos até que uma partição final seja alcançada. Se $g(x_1) \leq g(x_2) \leq \dots \leq g(x_n)$, então g é isotônica e nada se tem a fazer, ou seja, a partição inicial é também

a partição final, e $g^*(x_i) = g(x_i)$, $i = 1, 2, \dots, n$. Se não, seleciona-se qualquer um dos violadores, isto é, seleciona-se um x_{i+1} tal que $g(x_i) > g(x_{i+1})$. Faz-se então um *pool* (junção de coisas diferentes) entre os blocos que contém x_i e x_{i+1} em um novo bloco $\{x_i, x_{i+1}\}$ ordenado entre os blocos $\{x_{i-1}\}$ e $\{x_{i+2}\}$. O valor da função g nos pontos x_i e x_{i+1} é alterado para o valor da média:

$$\frac{g(x_i)w(x_i) + g(x_{i+1})w(x_{i+1})}{w(x_i) + w(x_{i+1})}$$

com peso associado aos novos valores de $g(x_i)$ e $g(x_{i+1})$ igual a $w(x_i) + w(x_{i+1})$.

Depois deste passo do algoritmo, os valores das médias associadas com os blocos são examinados para ver se eles estão ou não na ordem requerida. Se estão, a partição final foi encontrada e os valores de g^* para cada ponto de um bloco são os valores dos *pooled* associados a cada bloco. Se não, um novo par de blocos violadores adjacentes é selecionado e eles são colocados juntos para formar um novo bloco, com os pesos associados sendo a soma dos pesos deles e o valor da média associada sendo a média ponderada dos valores médios deles, completando outro passo do algoritmo. Isto é repetido até a partição final ser encontrada. A Figura 27 ilustra esta situação, nela foram necessários dois *pooling* até que o Maior Minorante Convexo fosse alcançado. A regressão isotônica $g^*(x_j)$ para cada ponto é o declive do segmento de reta, estendido à esquerda do ponto P_j^* do GCM.

Da Figura 27 pode-se concluir que $g^*(x_1) = g^*(x_2) = g^*(x_3)$, pois o declive do segmento à esquerda dos pontos P_1^* , P_2^* e P_3^* é o mesmo. Para o cômputo da regressão isotônica com uma ordenação simples, pode ser usado um software estatístico, como por exemplo o R Development Core Team (2009), que já apresenta rotina pronta para este tipo de análise.

Leeuw, Hornik e Mair (2009) criaram um novo pacote para o cálculo da regressão isotônica, o pacote *isotone*. Este pacote tem implementado o PAVA, bem

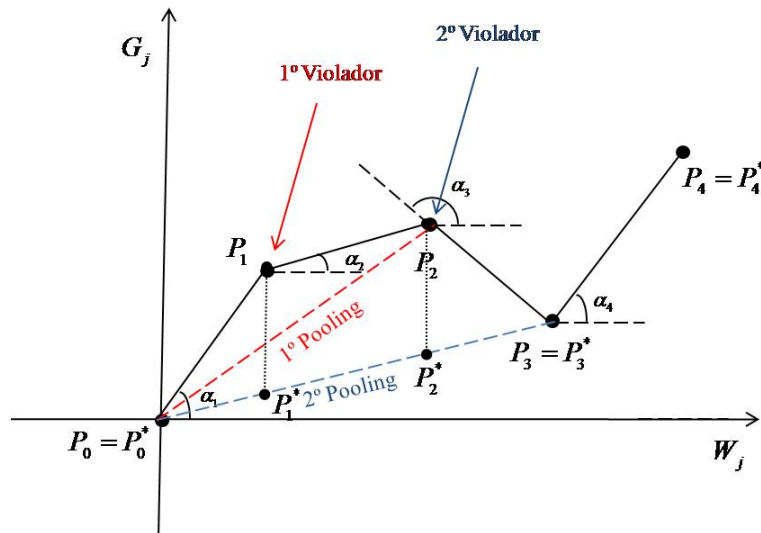


Figura 27 Representação dos violadores adjacentes e os *pooling* necessários até alcançar o Maior Minorante Convexo

como utiliza-se de outros algoritmos mais sofisticados para o caso de alguns tipos de ordem parcial. A função principal para o uso do PAVA é *gpava()*. O pacote está disponível na página <http://CRAN.R-project.org/package=isotone>.

Várias outras implementações do PAVA estão configuradas em R, a Tabela 2 fornece algumas delas. Cada uma destas funções pode resolver problemas específicos, por exemplo, a função *pava()* do pacote *Cir* resolve problemas de regressão isotônica para o caso de uma ordem simples, entretanto, para o caso de uma ordem parcial é necessária a utilização de uma outra função, como por exemplo a *gpava()* citada anteriormente. A consulta destas funções pode ser feita na página www.r-project.org/.

Os resultados discutidos nas últimas secções valem também para o caso de uma ordem reversa com algumas modificações. Um par violador será aquele para o qual $g(x_i) \leq g(x_{i+1})$, e algumas alterações devem ser feitas, mas sem perda de generalidade. No caso do software R, em algumas funções são necessários alguns

Tabela 2 Diversos pacotes e funções com implementações do PAVA em R.

Pacote	Função	Linguagem
Stats	<i>isoreg()</i>	C
monreg	<i>monreg()</i>	C
fdrtool	<i>monoreg()</i>	C
Cir	<i>pava()</i> , <i>cir.pava()</i> , <i>pava()</i>	R
Iso	<i>pava()</i>	Fortran
clue	<i>pava()</i>	R
logcondens	<i>isoMean()</i>	R
sandwich	<i>pava.blocks</i>	R
intcox	<i>intcox.pavaC()</i>	C

ajustes, por exemplo, na função *pava()* do pacote *Cir*, basta trocar o *default* dentro da função, onde aparece *dec=FALSE* substituir por *dec=TRUE* e a ordem a ser considerada passará a ser a reversa.

2.2 Análise de dados de sensibilidade e delineamentos sequenciais

Os dados de sensibilidade provenientes de experimentos de dose-resposta são, em geral, analisados sob procedimentos estatísticos específicos. A análise destes dados, bem como a relação entre os delineamentos sequenciais com os experimentos de dose-resposta são discutidos e definidos a seguir.

2.2.1 Análise de dados de sensibilidade

Existem diversas situações experimentais em que a variável resposta não pode ser medida na prática, isto é, tem-se variáveis aleatórias não observáveis. Por exemplo, em experimentos com explosivos, deixa-se cair um peso padrão sobre uma amostra de explosivo, de várias alturas, no sentido de se obter a menor altura que faz com que a amostra exploda. O que se pode conhecer é se determinada altura é maior ou menor do que a altura crítica desejada, ou seja, se ocorre a

explosão para uma determinada altura, pode-se concluir que aquela altura é maior do que a altura crítica, caso não ocorra a explosão conclui-se que esta altura é menor do que a altura crítica (DIXON; MOOD, 1948).

Uma outra situação em que aparecem experimentos com característica semelhantes a dos explosivos são os experimentos de dose-resposta a um determinado medicamento ou inseticida. Considerando a situação dos inseticidas, pode-se estar interessado na dose crítica que está associado com cada inseto, ou seja, se o inseto receber uma dose maior do que a dose crítica, ele morre, caso contrário, ele continua vivendo. Neste caso também não é possível obter esta dose experimentalmente, o que se consegue é saber se determinada dose é maior ou menor do que a dose crítica, para cada inseto testado.

Em ensaios de oncologia ou toxicidade aguda um dos objetivos é determinar a dose μ de uma droga que produzirá toxicidade (supõe-se uma resposta binária, toxicidade ou não toxicidade) com determinada probabilidade em uma população alvo, por exemplo, em um experimento com determinada droga que serve para combater um certo tipo de câncer, pode-se estar interessado em se conhecer a dose, para a qual 30% da população alvo (pessoas com o mesmo tipo de cancer) serão afetadas pela toxicidade do medicamento. Situação semelhante ocorre em pesquisas farmacêuticas de maneira geral, com pesquisas sobre germicidas, anestésicos e qualquer outro tipo de droga em diversos campos de pesquisas biológicas ou médicas.

O nome *análise de sensibilidade* parece bem sugestivo para todas as situações citadas anteriormente pois, em cada uma delas, o que se sabe é se a amostra a ser utilizada será *sensível* ou não ao tratamento aplicado. No caso dos explosivos o tratamento seria a altura, no caso dos inseticidas, a dose do inseticida e no caso dos ensaios oncológicos a dose de droga utilizada, ou seja, o que se tem é se a

unidade experimental respondeu ou não ao tratamento aplicado.

Nos experimentos de sensibilidade não é possível fazer mais do que uma observação em uma mesma unidade experimental, pois, uma vez que um teste foi realizado com a mesma, ela passa a ser alterada pelo experimento. No caso dos explosivos, o explosivo é danificado; no caso dos inseticidas, o inseto é morto; nos ensaios oncológicos, o paciente fica afetado pela medicação.

O procedimento utilizado por Bliss (1935) e Fisher (1935) em experimentos com as características citadas anteriormente é dividir a amostra em grupos (podendo ter ou não o mesmo tamanho amostral) e testar cada grupo para um nível escolhido, um primeiro grupo é testado em um nível de tratamento, um segundo grupo em um segundo nível e assim por diante. A análise realizada pelos autores para este delineamento foi denominada *probit*.

A partir de 1943 em trabalhos com experimentos que envolviam explosivos no Laboratório de Pesquisas sobre Explosivos em Bruceton, Pennsylvania, Dixon e Mood (1948) desenvolveram um método alternativo ao *probit*, ao qual denominaram de método *up and down*, que será posto em detalhes na próxima seção. Este método trabalha essencialmente com um delineamento do tipo sequencial e apresenta uma vantagem de concentrar as estimativas em torno da média com grande rapidez.

Segundo Stylianou e Flournoy (2002), em pesquisas médicas do tipo oncológicas, o tamanho da amostra, por questões éticas, não deve ser grande. Deseja-se estimar com precisão e acurácia o parâmetro de interesse, por exemplo, a dose que produz determinada toxicidade, com o menor número de pacientes possíveis.

Na utilização do método proposto por Bliss (1935) e Fisher (1935) os pacientes deveriam ser divididos em grupos, de modo que em determinado grupo os pacientes receberiam doses muito baixas que seriam certamente ineficazes e

outros pacientes receberiam doses muito elevadas que certamente seriam tóxicas, e isso ocorreria estando o pesquisador ciente desses efeitos, o que leva ao problema de questões éticas.

Segundo Ghosh, Mukhopadhyay e Sen (1997) e Stylianou e Flournoy (2002), a estimação sequencial em geral diminui o tamanho da amostra, entretanto, a análise se utiliza da teoria assintótica, sendo necessária, a busca por delineamentos ótimos, que mantenham um certo nível de confiança no processo de estimação, mesmo quando o tamanho da amostra é pequeno.

2.2.2 O método *probit*

O *probit* é um método para análise de dados de sensibilidade, que foi discutido inicialmente por Bliss (1934a, 1934b). O método passa a ser discutido com mais detalhes em Bliss (1935) onde o autor está interessado em calcular a curva de dose-mortalidade.

Um ensaio do tipo dose-resposta é aquele em que uma determinada droga é administrada em k diferentes doses (níveis), d_1, d_2, \dots, d_k em, respectivamente, m_1, m_2, \dots, m_k indivíduos, obtendo-se como resposta, após um período especificado, y_1, y_2, \dots, y_k indivíduos que mudaram de estado (ocorrência de um sucesso, por exemplo, morte). A resposta ao tratamento para cada unidade experimental é uma variável aleatória com distribuição Bernnoui com parâmetro p_i , que é a probabilidade de ocorrência de sucesso.

Portanto, neste tipo de ensaio uma amostra com indivíduos de uma mesma espécie é selecionada, ou seja, os elementos amostrais devem ter características as mais similares possíveis. Tal amostra é dividida em k grupos, cada um com m_i indivíduos, $i = 1, 2, \dots, k$ (em geral o número de indivíduos em cada grupo é o mesmo). Cada dose é aplicada a cada grupo e o número de sucessos por grupo

(morte do indivíduo, por exemplo) é contado, dessa maneira, a resposta para cada grupo é uma variável aleatória com distribuição binomial com parâmetros m_i e p_i , ou seja, $Y_i \sim \text{Bin}(m_i, p_i)$.

A Figura 28 ilustra esta situação. Nela, tem-se $m_1 = 4$, $m_2 = 5$, $m_3 = 8$ e $m_4 = 6$, ou seja, $Y_1 \sim \text{Bin}(4, p_1)$, $Y_2 \sim \text{Bin}(5, p_2)$, $Y_3 \sim \text{Bin}(8, p_3)$ e $Y_4 \sim \text{Bin}(6, p_4)$. As realizações destas variáveis aleatórias: $y_1 = 1$, $y_2 = 3$, $y_3 = 6$ e $y_4 = 6$ com proporção de mortos na amostra iguais a 0,25; 0,60; 0,75 e 1,00, respectivamente para cada grupo.

É natural que para doses pequenas ocorram pouco sucessos e que a medida que a dose aumenta ocorra um aumento neste número. Em outras palavras, para doses muito pequenas a probabilidade de morte dos insetos (por exemplo) é baixa e em doses muito altas essa probabilidade se aproxima de 1.

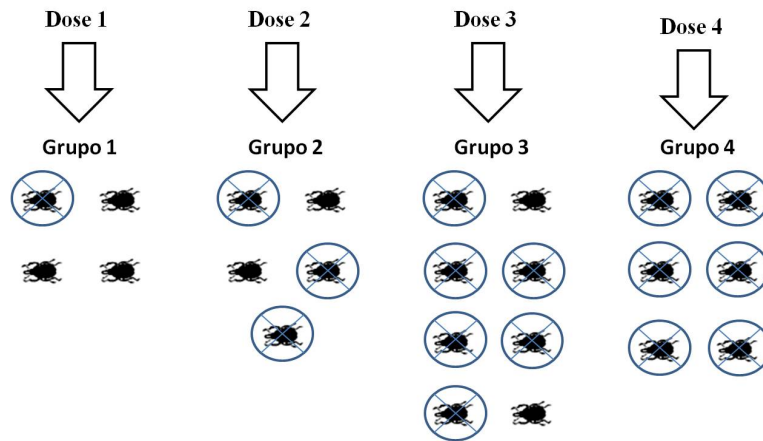


Figura 28 Representação do delineamento experimental para o *probit*

Segundo Bliss (1934a), o resultado de uma investigação da ação de um agente tóxico na mortalidade de um organismo é usualmente expresso como uma curva assimétrica em formato da letra S , ou seja, uma sigmóide, em que a porcentagem de mortalidade dos indivíduos está relacionada às doses a que eles teriam

seido expostos, conforme Figura 29.

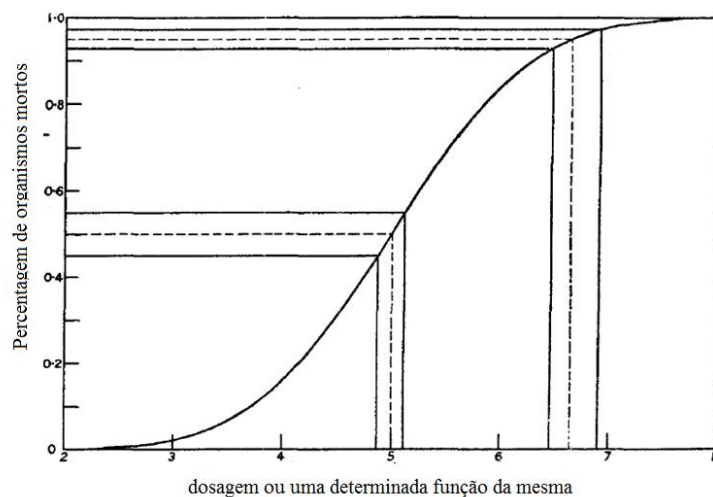


Figura 29 Curva de dose-mortalidade

Bliss (1935) chama a curva da Figura 29 de *Curva de Dose-Mortalidade* e argumenta que tal representação possui algumas desvantagens, tais como, pouca informação nas caudas com respeito à variação da porcentagem de mortos por variação nas dosagens. Observe que a mudança na porcentagem de mortos por variação da abscissa (doses) é pequena próximo às mortalidades de 0 e 100 por cento, e alta perto de 50 por cento.

A curva de *dose mortalidade* (Figura 29) não fornece diretamente para o pesquisador uma informação que é de interesse prático, que é a *susceptibilidade* ou *tolerância* da unidade experimental ao tratamento que lhe foi submetido. Esta tolerância pode ser entendida como o valor crítico que o tratamento assume no limiar entre sucesso ou fracasso do ensaio. Por exemplo, nos experimentos de dose-resposta, ela pode ser entendida como a menor dose possível para matar um determinado indivíduo, ou seja, a dose letal do indivíduo. Já nos experimentos com explosivos, a mesma pode ser entendida como a menor altura com que o peso

provocará a explosão de uma determinada amostra.

A tolerância é uma variável aleatória contínua que não pode ser medida diretamente, como já discutido, e o que o pesquisador tem em mãos são as proporções de sucessos (ou fracassos) para cada grupo do qual foi submetido o tratamento. A questão é, de posse destes dados, é possível modelar a susceptibilidade de cada indivíduo na população? A resposta a essa pergunta é sim, e isso pode ser feito de várias maneiras, uma delas denominada de *método probit*.

Suponha dois indivíduos I_1 e I_2 com susceptibilidades d_1 e d_2 respectivamente, ou seja, quando o indivíduo I_1 é submetido à dose d_1 ele morre, e não morre para uma dose menor do que d_1 . O mesmo ocorre para o indivíduo I_2 , com a dose d_2 . Note que, se o indivíduo I_1 é submetido a um dose maior do que a sua dose crítica, ele também morrerá, e isso também ocorre para o indivíduo I_2 . Dessa maneira, se $d_1 < d_2$, e os dois indivíduos forem submetidos à dose d_2 ocorrerá a morte dos dois. Em outras palavras, um indivíduo escolhido ao acaso responderá a uma determinada dose se a dose a qual ele foi submetido for maior ou igual que a sua dose crítica.

Assim, como no caso de qualquer outra característica biológica, a tolerância varia de indivíduo para indivíduo de uma população. No método *probit*, tal variável será considerada no processo de modelagem como uma variável aleatória com distribuição normal com média μ e variância σ^2 , ou seja, se U é a variável aleatória que representa a tolerância de cada indivíduo da população, tem-se que $U \sim N(\mu, \sigma^2)$, conforme Figura 30.

A modelagem pela distribuição normal não é uma pressuposição tão difícil de ser aceita, dado que a resposta ou não de cada indivíduo depende de vários fatores, é uma média de respostas de várias fontes, tais como, genéticas, imunológicas, ambientais, etc.

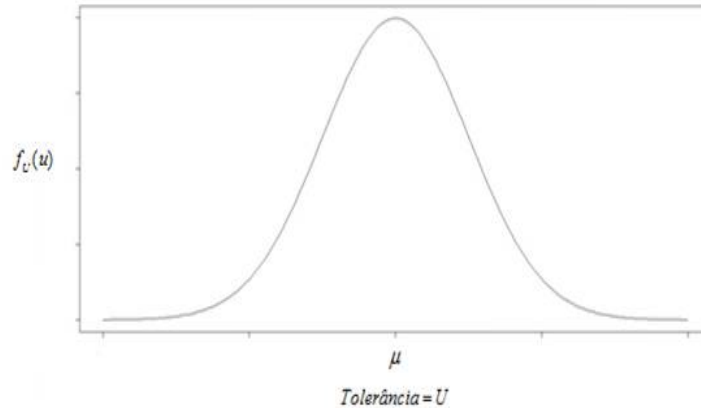


Figura 30 Representação da distribuição de tolerâncias

Se a dose d_i é aplicada à população toda e $f_U(u)$ é a função de densidade de probabilidade para a distribuição de tolerâncias, todo indivíduo cuja tolerância é menor do que d_i responderá à droga (tratamento), e a probabilidade de que um indivíduo escolhido ao acaso responda à dose é dada por

$$p_i = P[U \leq d_i] = F(d_i) = \int_{-\infty}^{d_i} f_U(u) du. \quad (2.40)$$

O problema, então, está em encontrar uma curva sigmóide que se ajuste bem aos dados e a partir dela obter as doses letais (ou críticas) desejadas (Figura 31). O que ocorre, porém, é que o modelo exposto na Equação 2.40 é não linear nos parâmetros e como já discutido, as curvas sigmóides apresentam pouca informação nas caldas a respeito da variação no percentual de mortos com relação a uma determinada variação das doses. A ideia é fazer uma transformação tal que essa curva sigmóide passe a ser vista como uma reta e assim procedimentos comuns de regressão possam ser usados para estimação dos parâmetros.

A ideia do *probit* é fazer uma simples manipulação na Equação 2.40, como

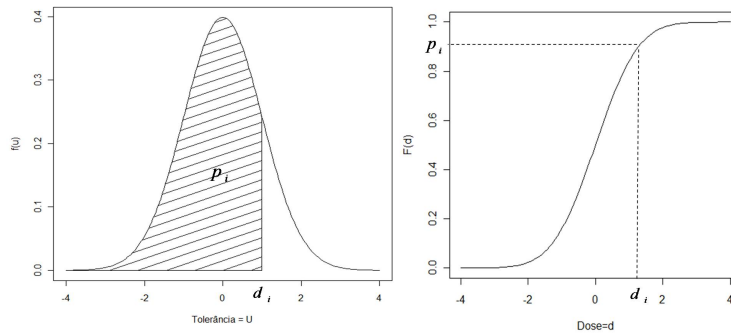


Figura 31 Distribuição densidade de probabilidade e função de distribuição acumulada das tolerâncias

segue:

$$\begin{aligned}
 p_i &= P[U \leq d_i] = P\left[\frac{(U - \mu)}{\sigma} \leq \frac{(d_i - \mu)}{\sigma}\right] \\
 &= P\left[Z \leq -\frac{\mu}{\sigma} + \frac{1}{\sigma}d_i\right] = P[Z \leq \beta_1 + \beta_2 d_i],
 \end{aligned}$$

com $\beta_1 = -\frac{\mu}{\sigma}$ e $\beta_2 = \frac{1}{\sigma}$. Considerando $\Phi(x)$ a distribuição acumulada da normal padrão no ponto x , tem-se:

$$p_i = \Phi(\beta_1 + \beta_2 d_i) \quad (2.41)$$

em que $\Phi(\cdot)$ é uma função não linear aplicada em uma relação linear nos parâmetros. Aplicando a inversa de $\Phi(\cdot)$ em ambos os membros da Equação 2.41:

$$\Phi^{-1}(p_i) = \beta_1 + \beta_2 d_i.$$

O *probit* de p_i é definido como sendo $\Phi^{-1}(p_i)$, ou seja, o *probit* nada mais é do que uma transformação dos valores que estão no eixo das ordenadas da sigmóide da Figura 31:

$$\text{probit}(p_i) = \Phi^{-1}(p_i) = \beta_1 + \beta_2 d_i. \quad (2.42)$$

Considerando $\hat{p}_i = \frac{y_i}{m_i}$ a proporção de sucessos (mortos) observados para a i -ésima dose, pode-se estimar β_1 e β_2 pelo método dos mínimos quadrados ponderados. Tal método consiste em encontrar estimativas $\hat{\beta}_1$ e $\hat{\beta}_2$ que minimizam a expressão:

$$\sum_{i=1}^n [\hat{p}_i - p_i]^2 w_i = \sum_{i=1}^n [\hat{p}_i - \Phi(\beta_1 + \beta_2 d_i)]^2 w_i \quad (2.43)$$

em que os pesos w_i são usualmente tomados como o inverso das variâncias das binomiais correspondentes, ou seja,

$$\begin{aligned} w_i &= \frac{1}{\text{Var}[\hat{P}_i]} = \frac{1}{\text{Var}[Y_i/m_i]} = \frac{1}{\text{Var}[Y_i]/m_i^2} \\ &= \frac{1}{(m_i p_i (1 - p_i))/m_i^2} = \frac{m_i}{p_i (1 - p_i)}. \end{aligned}$$

Quando a transformação *probit* é aplicada, uma nova relação de mínimos quadrados deve ser estipulada, e uma pergunta que surge é: Que tipo de minimização deve ser requerida para as variáveis transformadas, para que se tenha uma equivalência entre esta nova relação e à Equação 2.43? Para responder a esta pergunta deve-se observar que a transformação *probit* pode ser pensada como uma função g de $A \subset \mathbb{R}^2$ em $B \subset \mathbb{R}^2$, definida por $g(x, y) = (x, f(y))$, conforme Figura 32, em que, $f(y) = \Phi^{-1}(y)$.

Observe que, do cálculo diferencial, para cada dose d_i fixada, vale a seguinte relação:

$$f'(y) \cong \frac{f(y + \Delta y) - f(y)}{\Delta y},$$

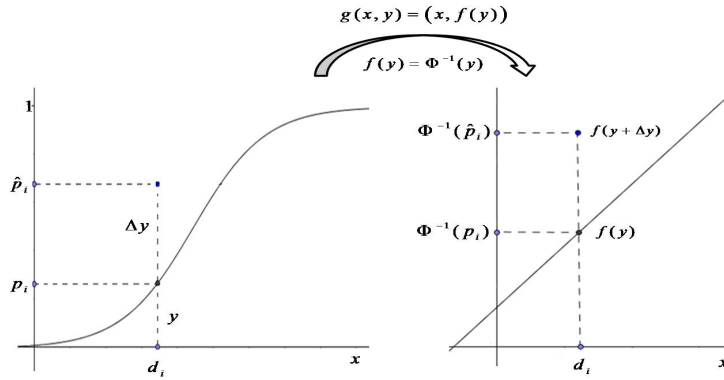


Figura 32 O probit como uma transformação $g(x, y) = (x, f(y))$

ou

$$\Delta y \cong \frac{1}{f'(y)} [f(y + \Delta y) - f(y)]. \quad (2.44)$$

Para $f(y) = \Phi^{-1}(y)$, $y = p_i$ e $\Delta y = \hat{p}_i - p_i$, conforme Figura 32, tem-se os seguintes resultados:

$$\frac{df(y)}{dy} = \frac{d\Phi^{-1}(y)}{dy} = \frac{1}{\frac{d\Phi}{dy}(\Phi^{-1}(y))} = \frac{1}{\phi(\Phi^{-1}(y))},$$

ou seja,

$$f'(p_i) = \frac{1}{\phi(\Phi^{-1}(p_i))} = \frac{1}{z_i}.$$

Além disso,

$$f(y + \Delta y) = f(\hat{p}_i) = \Phi^{-1}(\hat{p}_i)$$

e

$$f(y) = f(p_i) = \Phi^{-1}(p_i) = \beta_1 + \beta_2 d_i,$$

em que z_i é a ordenada da distribuição normal no ponto de abscissa $\Phi^{-1}(p_i)$.

Substituindo estes resultados na Equação 2.44:

$$\begin{aligned}
(\hat{p}_i - p_i) &\cong z_i[\Phi^{-1}(\hat{p}_i) - \Phi^{-1}(p_i)] & (2.45) \\
\Rightarrow (\hat{p}_i - p_i)^2 &\cong z_i^2[\Phi^{-1}(\hat{p}_i) - \Phi^{-1}(p_i)]^2 \\
\Rightarrow (\hat{p}_i - p_i)^2 \frac{m_i}{p_i(1-p_i)} &\cong [\Phi^{-1}(\hat{p}_i) - \Phi^{-1}(p_i)]^2 \frac{z_i^2 m_i}{p_i(1-p_i)} \\
&\Rightarrow (\hat{p}_i - p_i)^2 w_i \cong [\Phi^{-1}(\hat{p}_i) - \Phi^{-1}(p_i)]^2 w_i^* \\
\Rightarrow \sum_{i=1}^n (\hat{p}_i - p_i)^2 w_i &\cong \sum_{i=1}^n [\Phi^{-1}(\hat{p}_i) - \Phi^{-1}(p_i)]^2 w_i^*.
\end{aligned}$$

Dessa maneira, o problema de estimação por mínimos quadrados ponderados passa a ser encontrar $\hat{\beta}_1$ e $\hat{\beta}_2$ que minimizam a expressão:

$$\sum_{i=1}^n [\Phi^{-1}(\hat{p}_i) - \Phi^{-1}(p_i)]^2 w_i^*,$$

com os novos pesos $w_i^* = (z_i^2 m_i)/(p_i(1-p_i))$. Porém, esses pesos contém grandezas desconhecidas. Uma alternativa é substituir estas grandezas por suas estimativas \hat{p}_i e \hat{q}_i , e o problema de estimação por mínimos quadrados ponderados passa a ser encontrar $\hat{\beta}_1$ e $\hat{\beta}_2$ que minimizam:

$$\sum_{i=1}^n [\Phi^{-1}(\hat{p}_i) - \Phi^{-1}(p_i)]^2 \frac{z_i^2 m_i}{\hat{p}_i \hat{q}_i}$$

ou

$$\sum_{i=1}^n [\text{probit}(\hat{p}_i) - (\beta_1 + \beta_2 d_i)]^2 \frac{z_i^2 m_i}{\hat{p}_i \hat{q}_i}, \quad (2.46)$$

em que z_i é a ordenada da normal padrão no ponto $\Phi^{-1}(\hat{p}_i) = \text{probit}(\hat{p}_i)$.

A forma como este resultado foi obtido é contribuição deste trabalho. Fisher (1935) obteve este mesmo resultado pelo processo de estimação por máxima verossimilhança, o que era de se esperar, dado a suposição de normalidade para a

distribuição de tolerâncias.

A Figura 33 ilustra uma situação com dados simulados, consistindo das proporções antes e depois da transformação *probit*. Pode-se observar o formato sigmóide nos dados antes da transformação e a oscilação dos mesmos em torno de uma reta após a realização do *probit*.

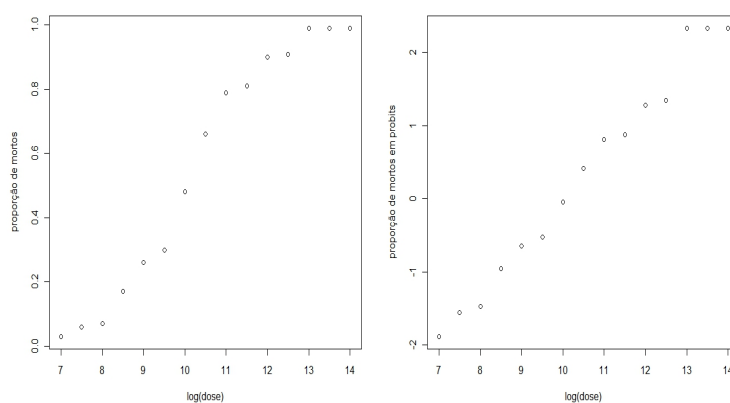


Figura 33 Amostra simulada antes e depois da transformação *probit*

Observe na Figura 33 que, na escala *probit*, alguns valores aparecem com sinal negativo. Isto não ocasiona problema algum nas estimativas finais de μ e σ . Entretanto, pode causar confusão na mente de pesquisadores com pouca familiaridade em estatística, pois os mesmos podem pensar que estão trabalhando com proporções negativas. Pensando nisso, Bliss (1935) sugere usar uma normal centrada em 5 ao invés de uma centrada em 0 e apresenta as suas tabelas de *probits* com valores que tem esta característica. Este fato não é de muita importância, pois o objetivo final são as estimativas de μ e σ que não são alteradas por esta translação.

O uso de tabelas também é algo ultrapassado, tendo em vista o avanço computacional nas últimas décadas. Dessa forma, para calcular o *probit* de um

valor, pode-se usar um software estatístico, como por exemplo o R. Para calcular o *probit* usando este software, basta utilizar a função *qnorm* do mesmo.

A utilização do método *probit* requer que a tolerância seja normalmente distribuída. A experiência mostra que em experimentos de dose mortalidade e em experimentos com explosivos o logaritmo da dose ou da altura é que tem distribuição normal. Então, todas as análises não seriam feitas diretamente com a dose ou altura, mas com a $\log(\text{dose})$ ou $\log(\text{altura})$. Para as conclusões a respeito da variável de interesse deve-se aplicar a função inversa do logaritmo nos resultados obtidos ou fazer adequações necessárias para as conclusões finais (DIXON; MOOD, 1948).

É muito comum encontrar resultados de gráficos: $\log(\text{dose})$ versus *probit* das proporções. Em outras áreas outras transformações podem ser requeridas e o pesquisador deve atentar a este fato.

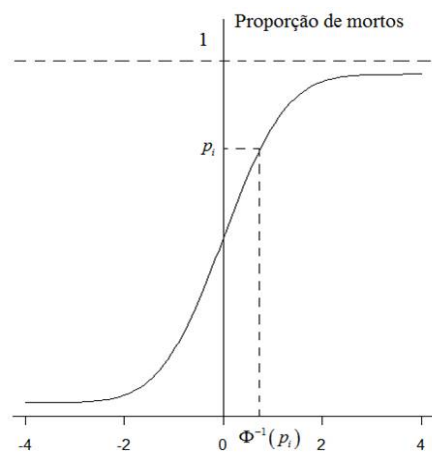


Figura 34 O *probit* na curva de distribuição acumulada

O método *probit* apresenta um problema quando ocorrem 100% ou 0% de sucessos para um determinado nível, pois a função inversa da acumulada da normal não está definida para 0 e 1, não sendo possível o cálculo do *probit* destes

valores.

A medida que as proporções se aproximam de zero, o *probit* tende a $-\infty$ e quando as proporções se aproximam de 1 ele tende a $+\infty$ (Figura 34). Além disso, tal fato pode influenciar bastante nas estimativas de β_1 e β_2 na Equação 2.42. A solução deste problema foi apresentada por Fisher (1935) em apêndice do artigo de Bliss (1935) e será apresentada a seguir, de forma ligeiramente diferente da apresentação original.

2.2.2.1 O caso dos 0% e 100% de sucessos em análise *probit*

Da Equação 2.45, tem-se que:

$$(\hat{p}_i - p_i) \cong z_i[\text{probit}(\hat{p}_i) - \text{probit}(p_i)]. \quad (2.47)$$

Entretanto, quando p_i é muito pequeno, o número de mortos observados pode frequentemente ser igual a zero ou muito pequeno, e a aproximação em 2.47 não será satisfatória, pois o valor $\text{probit}(\hat{p}_i)$ não estará definido, ou tenderá a menos infinito (para valores de \hat{p}_i próximo de zero).

Deve ser observado que se o experimento fosse repetido várias vezes, em algumas vezes o número de mortos observados poderia não ser igual a zero e conseqüentemente a indefinição em $\text{probit}(\hat{p}_i)$ não ocorreria. A solução apresentada por Fisher (1935) consiste em substituir o valor $\text{probit}(\hat{p}_i)$ por um valor fictício $\text{probit}^*(\hat{p}_i)$. Dessa maneira, supondo que para a dose d_j o número de mortos foi

igual a zero, a Equação 2.47 pode ser substituída por:

$$\begin{aligned}
 (\hat{p}_j - p_j) &\cong z_j[\text{probit}^*(\hat{p}_j) - \text{probit}(p_j)] \\
 \Rightarrow (0 - p_j) &\cong z_j[\text{probit}^*(\hat{p}_j) - \text{probit}(p_j)] \\
 \Rightarrow -\frac{p_j}{z_j} &\cong \text{probit}^*(\hat{p}_j) - \text{probit}(p_j) \\
 \Rightarrow \text{probit}^*(\hat{p}_j) &\cong \text{probit}(p_j) - \frac{p_j}{z_j} \\
 \Rightarrow \text{probit}^*(\hat{p}_j) &\cong (\beta_1 + \beta_2 d_j) - \frac{p_j}{z_j}.
 \end{aligned}$$

Agora, suponha que para a dose d_j o número de mortos é 100%, então:

$$\begin{aligned}
 (\hat{p}_j - p_j) &\cong z_j[\text{probit}^*(\hat{p}_j) - \text{probit}(p_j)] \\
 \Rightarrow (1 - p_j) &\cong z_j[\text{probit}^*(\hat{p}_j) - \text{probit}(p_j)] \\
 \Rightarrow \frac{1 - p_j}{z_j} &\cong \text{probit}^*(\hat{p}_j) - \text{probit}(p_j) \\
 \Rightarrow \text{probit}^*(\hat{p}_j) &\cong \text{probit}(p_j) + \frac{1 - p_j}{z_j} \\
 \Rightarrow \text{probit}^*(\hat{p}_j) &\cong (\beta_1 + \beta_2 d_j) + \frac{q_j}{z_j}.
 \end{aligned}$$

Dessa maneira, o procedimento para a análise *probit* deve ser o seguinte:

Toma-se todos os pares (d_i, \hat{p}_i) em que $\hat{p}_i \neq 0$ e $\hat{p}_i \neq 1$ e faz-se uma primeira estimativa de β_1 e β_2 por mínimos quadrados ponderados com pesos $w_i = \frac{z_i^2 m_i}{\hat{p}_i q_i}$, conforme exposto anteriormente.

Encontradas estas estimativas e utilizando a Equação 2.41, considera-se o

probit para a dose d_j em que ocorreu $\hat{p}_j = 0$ por:

$$probit^*(\hat{p}_j) = \left(\hat{\beta}_1 + \hat{\beta}_2 d_j \right) - \frac{\Phi \left(\hat{\beta}_1 + \hat{\beta}_2 d_j \right)}{\phi \left(\hat{\beta}_1 + \hat{\beta}_2 d_j \right)},$$

e para o caso onde ocorreu $\hat{p}_j = 1$:

$$probit^*(\hat{p}_j) = \left(\hat{\beta}_1 + \hat{\beta}_2 d_j \right) + \frac{1 - \Phi \left(\hat{\beta}_1 + \hat{\beta}_2 d_j \right)}{\phi \left(\hat{\beta}_1 + \hat{\beta}_2 d_j \right)},$$

em que $\Phi(\cdot)$ e $\phi(\cdot)$ representam a função de distribuição acumulada e função densidade de probabilidade da normal padrão respectivamente.

De posse destes valores, considera-se todos os valores iniciais juntamente com os anteriores e a regressão linear é refeita, calculando-se novas estimativas para β_1 e β_2 , que serão as estimativas finais do processo. Fisher (1935) argumenta que esta passagem é importante, pois a simples exclusão dos dados pode tornar as estimativas viciadas.

2.2.3 O método *logit*

O método *probit* pressupõe que a distribuição da tolerância dos indivíduos seja uma normal. Berkson (1944) argumenta que a densidade da distribuição logística está muito próxima da curva normal e é aplicada largamente em estudos de fenômenos físico-químicos, podendo ter, assim, em determinadas situações, um melhor apelo teórico.

Como este método consiste basicamente em uma readequação do *probit*, substituindo a curva normal pela curva logística, Berkson (1944) denominou o mesmo de *método logit* (*logistic unit*). Mantendo as mesmas notações utilizadas na seção anterior, assume-se que U tem distribuição logística com parâmetros μ e

τ , que é muito similar à curva de distribuição normal na forma (Figura 35), com caudas um pouco mais longas e tem função densidade de probabilidade dada por (MOOD; GRAYBILL; BOES, 1974):

$$f_U(u; \mu, \tau) = \frac{1}{\tau} \frac{\exp\left(\frac{u-\mu}{\tau}\right)}{\left[1 + \exp\left(\frac{u-\mu}{\tau}\right)\right]^2}, \quad \mu, \tau \in \mathbb{R} \text{ e } \tau > 0, \quad (2.48)$$

com esperança $E(U) = \mu$ e variância $\sigma^2 = Var(U) = \frac{\pi^2\tau^2}{3}$.

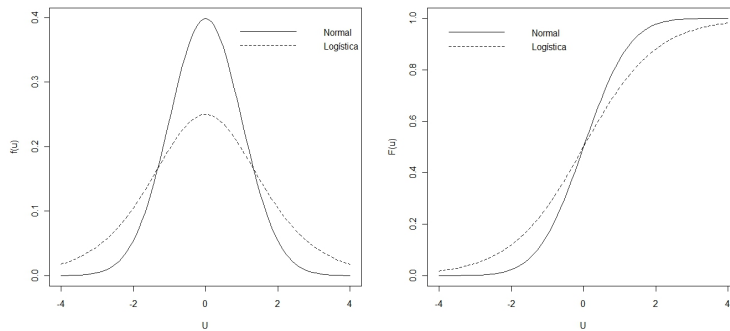


Figura 35 Funções densidade de probabilidade e funções de distribuição acumulada da Normal e Logística com $\mu = 0$ e $\sigma^2 = 1$

Fazendo uma mudança de parâmetros em 2.48, com $\beta_1 = -\frac{\mu}{\tau}$ e $\beta_2 = \frac{1}{\tau}$, tem-se:

$$f_U(u; \beta_1, \beta_2) = \frac{\beta_2 e^{\beta_1 + \beta_2 u}}{(1 + e^{\beta_1 + \beta_2 u})^2}.$$

Logo,

$$p_i = P[U \leq d_i] = F(d_i) = \frac{e^{\beta_1 + \beta_2 d_i}}{1 + e^{\beta_1 + \beta_2 d_i}}.$$

Observe que,

$$1 - p_i = 1 - \frac{e^{\beta_1 + \beta_2 d_i}}{1 + e^{\beta_1 + \beta_2 d_i}} = \frac{1}{1 + e^{\beta_1 + \beta_2 d_i}},$$

e que,

$$\frac{p_i}{1 - p_i} = \frac{e^{\beta_1 + \beta_2 d_i}}{1 + e^{\beta_1 + \beta_2 d_i}} \times (1 + e^{\beta_1 + \beta_2 d_i}) = e^{\beta_1 + \beta_2 d_i},$$

dessa forma,

$$\log \left(\frac{p_i}{1 - p_i} \right) = \beta_1 + \beta_2 d_i.$$

A transformação feita nas unidades $\frac{p_i}{1 - p_i}$ pelo logaritmo neperiano é denominada por Berkson (1944) *logit* de p_i . Assim,

$$\text{logit}(p_i) = \log \left(\frac{p_i}{1 - p_i} \right) = \beta_1 + \beta_2 d_i. \quad (2.49)$$

Para estimação dos parâmetros β_1 e β_2 são usados métodos convencionais como os mínimos quadrados e máxima verossimilhança.

Se a taxa de mortalidade em d_i é p_i e a proporção de mortos observados para a dose d_i em m_i ensaios é $\hat{p}_i = \frac{y_i}{m_i}$, em que y_i é o número de mortos observados, tem-se que:

$$P(\hat{P}_i = \hat{p}_i) = \frac{m_i!}{m_i!(m_i - y_i)!} p_i^{y_i} (1 - p_i)^{m_i - y_i} = C_{m_i}^{y_i} p_i^{y_i} (1 - p_i)^{m_i - y_i}.$$

Considerando todos os valores dos \hat{p}_i para as doses d_1, d_2, \dots, d_k , tem-se a verossimilhança:

$$L = L(\beta_1, \beta_2 | \hat{p}_1, \hat{p}_2, \dots, \hat{p}_k) = \prod_{i=1}^k C_{m_i}^{y_i} p_i^{y_i} (1 - p_i)^{m_i - y_i}.$$

Tomando o logaritmo natural em L :

$$l = \log L = \sum_{i=1}^k \log C_{m_i}^{y_i} + \sum_{i=1}^k y_i \log p_i + \sum_{i=1}^k (m_i - y_i) \log(1 - p_i).$$

As seguintes derivadas parciais serão utilizadas no cômputo das equações de máxima verossimilhança:

$$\begin{aligned}
 \frac{\partial p_i}{\partial \beta_1} &= \frac{e^{\beta_1 + \beta_2 d_i} (1 + e^{\beta_1 + \beta_2 d_i}) - e^{\beta_1 + \beta_2 d_i} e^{\beta_1 + \beta_2 d_i}}{(1 + e^{\beta_1 + \beta_2 d_i})^2} \\
 &= \frac{e^{\beta_1 + \beta_2 d_i} (1 + e^{\beta_1 + \beta_2 d_i} - e^{\beta_1 + \beta_2 d_i})}{(1 + e^{\beta_1 + \beta_2 d_i})^2} \\
 &= \frac{e^{\beta_1 + \beta_2 d_i}}{1 + e^{\beta_1 + \beta_2 d_i}} \times \frac{1}{1 + e^{\beta_1 + \beta_2 d_i}} \\
 &= p_i(1 - p_i)
 \end{aligned}$$

e

$$\begin{aligned}
 \frac{\partial p_i}{\partial \beta_2} &= \frac{d_i e^{\beta_1 + \beta_2 d_i} (1 + e^{\beta_1 + \beta_2 d_i}) - e^{\beta_1 + \beta_2 d_i} d_i e^{\beta_1 + \beta_2 d_i}}{(1 + e^{\beta_1 + \beta_2 d_i})^2} \\
 &= \frac{d_i e^{\beta_1 + \beta_2 d_i} (1 + e^{\beta_1 + \beta_2 d_i} - e^{\beta_1 + \beta_2 d_i})}{(1 + e^{\beta_1 + \beta_2 d_i})^2} \\
 &= d_i \frac{e^{\beta_1 + \beta_2 d_i}}{1 + e^{\beta_1 + \beta_2 d_i}} \times \frac{1}{1 + e^{\beta_1 + \beta_2 d_i}} \\
 &= d_i p_i(1 - p_i).
 \end{aligned}$$

Assim,

$$\begin{aligned}
\frac{\partial l}{\partial \beta_1} &= \sum_{i=1}^k \left(\frac{\partial p_i}{\partial \beta_1} \times y_i \times \frac{1}{p_i} \right) + \sum_{i=1}^k \left(-\frac{\partial p_i}{\partial \beta_1} \times (m_i - y_i) \times \frac{1}{1 - p_i} \right) \\
&= \sum_{i=1}^k \frac{\partial p_i}{\partial \beta_1} \left(\frac{y_i}{p_i} + \frac{y_i - m_i}{1 - p_i} \right) = \sum_{i=1}^k \left(\frac{\partial p_i}{\partial \beta_1} \times \frac{y_i - p_i m_i}{p_i(1 - p_i)} \right) \\
&= \sum_{i=1}^k \left(p_i(1 - p_i) \times \frac{y_i - p_i m_i}{p_i(1 - p_i)} \right) \\
&= \sum_{i=1}^k (y_i - p_i m_i) = \sum_{i=1}^k \left(\frac{y_i}{m_i} - p_i \right) m_i \\
&= \sum_{i=1}^k (\hat{p}_i - p_i) m_i \\
&= \sum_{i=1}^k \left(\hat{p}_i - \frac{e^{\beta_1 + \beta_2 d_i}}{1 + e^{\beta_1 + \beta_2 d_i}} \right) m_i
\end{aligned}$$

e

$$\begin{aligned}
\frac{\partial l}{\partial \beta_2} &= \sum_{i=1}^k \left(\frac{\partial p_i}{\partial \beta_2} \times y_i \times \frac{1}{p_i} \right) + \sum_{i=1}^k \left(-\frac{\partial p_i}{\partial \beta_2} \times (m_i - y_i) \times \frac{1}{1 - p_i} \right) \\
&= \sum_{i=1}^k \frac{\partial p_i}{\partial \beta_2} \left(\frac{y_i}{p_i} + \frac{y_i - m_i}{1 - p_i} \right) = \sum_{i=1}^k \left(\frac{\partial p_i}{\partial \beta_2} \times \frac{y_i - p_i m_i}{p_i(1 - p_i)} \right) \\
&= \sum_{i=1}^k \left(d_i p_i(1 - p_i) \times \frac{y_i - p_i m_i}{p_i(1 - p_i)} \right) \\
&= \sum_{i=1}^k d_i (y_i - p_i m_i) = \sum_{i=1}^k \left(\frac{y_i}{m_i} - p_i \right) d_i m_i \\
&= \sum_{i=1}^k (\hat{p}_i - p_i) d_i m_i \\
&= \sum_{i=1}^k \left(\hat{p}_i - \frac{e^{\beta_1 + \beta_2 d_i}}{1 + e^{\beta_1 + \beta_2 d_i}} \right) d_i m_i.
\end{aligned}$$

Portanto, os estimadores de máxima verossimilhança $\hat{\beta}_1$ e $\hat{\beta}_2$ devem satisfazer as equações:

$$\sum_{i=1}^k \left(\hat{p}_i - \frac{e^{\hat{\beta}_1 + \hat{\beta}_2 d_i}}{1 + e^{\hat{\beta}_1 + \hat{\beta}_2 d_i}} \right) m_i = 0$$

e

$$\sum_{i=1}^k \left(\hat{p}_i - \frac{e^{\hat{\beta}_1 + \hat{\beta}_2 d_i}}{1 + e^{\hat{\beta}_1 + \hat{\beta}_2 d_i}} \right) d_i m_i = 0,$$

ou ainda,

$$\sum_{i=1}^k \left(m_i \times \frac{e^{\hat{\beta}_1 + \hat{\beta}_2 d_i}}{1 + e^{\hat{\beta}_1 + \hat{\beta}_2 d_i}} \right) = \sum_{i=1}^k \hat{p}_i m_i \quad (2.50)$$

$$\sum_{i=1}^k \left(d_i m_i \times \frac{e^{\hat{\beta}_1 + \hat{\beta}_2 d_i}}{1 + e^{\hat{\beta}_1 + \hat{\beta}_2 d_i}} \right) = \sum_{i=1}^k \hat{p}_i d_i m_i. \quad (2.51)$$

As soluções para as Equações 2.50 e 2.51 só podem ser encontradas por métodos numéricos.

Uma outra metodologia que pode ser utilizada na estimação de β_1 e β_2 é o método dos mínimos quadrados ponderados, similarmente ao que foi exposto no método *probit*. Agora, deseja-se encontrar estimativas $\hat{\beta}_1$ e $\hat{\beta}_2$ que minimizam a expressão:

$$\sum_{i=1}^k [\hat{p}_i - p_i]^2 w_i = \sum_{i=1}^k \left[\hat{p}_i - \frac{e^{\beta_1 + \beta_2 d_i}}{1 + e^{\beta_1 + \beta_2 d_i}} \right]^2 w_i, \quad (2.52)$$

em que os pesos w_i são tomados como o inverso das variâncias das proporções consideradas, ou seja,

$$w_i = \frac{1}{\text{Var}[\hat{P}_i]} = \frac{m_i}{p_i(1 - p_i)}.$$

Do mesmo modo do que foi colocado no *probit*, considere a transformação *logit* nas proporções como uma função g de $A \subset \mathbb{R}^2$ em $B \subset \mathbb{R}^2$, definida por $g(x, y) = (x, f(y))$, conforme Figura 36, em que, $f(y) = \log(y/(1 - y))$.

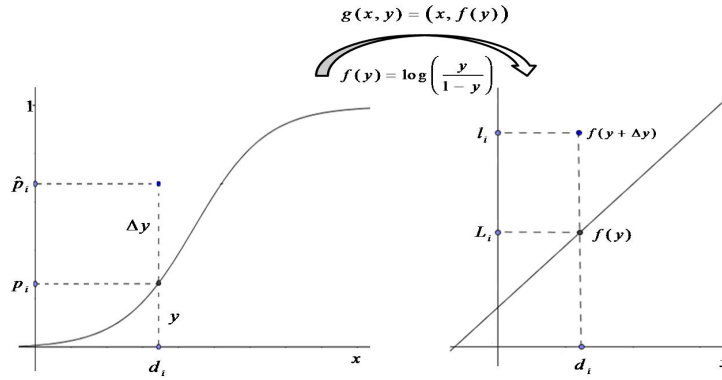


Figura 36 Visualização do *logit* como uma transformação $g(x, y) = (x, f(y))$

Na Figura 36,

$$l_i = \text{logit}(\hat{p}_i) = \log(\hat{p}_i / (1 - \hat{p}_i))$$

e

$$L_i = \text{logit}(p_i) = \log(p_i / (1 - p_i)) = \beta_1 + \beta_2 d_i.$$

Novamente, para cada dose d_i fixada:

$$\Delta y \cong \frac{1}{f'(y)} [f(y + \Delta y) - f(y)]. \quad (2.53)$$

Para $f(y) = \log(y/(1-y))$, $y = p_i$ e $\Delta y = \hat{p}_i - p_i$, conforme Figura 36, tem-se os seguintes resultados:

$$f'(y) = \frac{1}{y(1-y)} \Rightarrow f'(p_i) = \frac{1}{p_i(1-p_i)}$$

$$f(y + \Delta y) = f(\hat{p}_i) = \log(\hat{p}_i / (1 - \hat{p}_i)) = l_i$$

$$f(y) = f(p_i) = \log(p_i / (1 - p_i)) = \beta_1 + \beta_2 d_i = L_i.$$

Substituindo estes resultados na Equação 2.53:

$$\begin{aligned}
(\hat{p}_i - p_i) &\cong p_i(1 - p_i)[l_i - L_i] \\
\Rightarrow (\hat{p}_i - p_i)^2 &\cong [p_i(1 - p_i)]^2 [l_i - L_i]^2 \\
\Rightarrow m_i(\hat{p}_i - p_i)^2 &\cong m_i [p_i(1 - p_i)]^2 [l_i - L_i]^2 \\
\Rightarrow (\hat{p}_i - p_i)^2 \frac{m_i}{p_i(1-p_i)} &\cong [l_i - L_i]^2 m_i p_i (1 - p_i) \\
\Rightarrow (\hat{p}_i - p_i)^2 w_i &\cong [l_i - L_i]^2 w_i^* \Rightarrow \sum_{i=1}^k (\hat{p}_i - p_i)^2 w_i \cong \sum_{i=1}^k [l_i - L_i]^2 w_i^*.
\end{aligned}$$

Dessa maneira, o problema de estimação por mínimos quadrados ponderados passa a ser encontrar $\hat{\beta}_1$ e $\hat{\beta}_2$ que minimizam a expressão:

$$\sum_{i=1}^k [l_i - L_i]^2 w_i^*,$$

com os novos pesos $w_i^* = m_i p_i (1 - p_i)$. Porém, estes pesos contém grandezas desconhecidas, e assim como no *probit*, uma alternativa é substituir estas grandezas por suas estimativas \hat{p}_i e \hat{q}_i , e o problema de estimação por mínimos quadrados passa a ser encontrar $\hat{\beta}_1$ e $\hat{\beta}_2$ que minimizam:

$$\sum_{i=1}^k [l_i - L_i]^2 m_i \hat{p}_i \hat{q}_i$$

ou

$$\sum_{i=1}^k [\text{logit}(\hat{p}_i) - (\beta_1 + \beta_2 d_i)]^2 m_i \hat{p}_i \hat{q}_i. \quad (2.54)$$

2.2.3.1 O caso dos 0% e 100% de sucessos em análise *logit*

Berkson (1944) não apresenta em detalhes uma discussão para esta situação, embora presente em notas de rodapé os resultados finais. Aqui será apresentado suas conclusões com os cálculos detalhados. Da Equação 2.53, tem-se

que:

$$(\hat{p}_i - p_i) \cong p_i(1 - p_i)[\text{logit}(\hat{p}_i) - \text{logit}(p_i)]. \quad (2.55)$$

Novamente, quando p_i é muito pequeno, o número de mortos observados pode frequentemente ser igual a zero ou muito pequeno, e a aproximação em 2.55 não será satisfatória, pois então, o valor $\text{logit}(\hat{p}_i)$ não estará definido, ou tenderá a menos infinito (para valores de \hat{p}_i próximo de zero).

De novo, deve ser observado que se o experimento fosse repetido várias vezes, em algumas vezes o número de mortos observados poderia não ser igual a zero e conseqüentemente a indefinição em $\text{logit}(\hat{p}_i)$ não ocorreria. A solução consiste em substituir o valor $\text{logit}(\hat{p}_i)$ por um valor fictício $\text{logit}^*(\hat{p}_i)$.

Se para a dose d_j o número de sucessos foi igual a zero, a Equação 2.55 pode ser substituída por:

$$\begin{aligned} (\hat{p}_j - p_j) &\cong p_j(1 - p_j)[\text{logit}^*(\hat{p}_j) - \text{logit}(p_j)] \\ \Rightarrow (0 - p_j) &\cong p_j(1 - p_j)[\text{logit}^*(\hat{p}_j) - \text{logit}(p_j)] \\ \Rightarrow -\frac{p_j}{p_j(1 - p_j)} &\cong \text{logit}^*(\hat{p}_j) - \text{logit}(p_j) \\ \Rightarrow \text{logit}^*(\hat{p}_j) &\cong \text{logit}(p_j) - \frac{p_j}{p_j(1 - p_j)} \\ \Rightarrow \text{logit}^*(\hat{p}_j) &\cong (\beta_1 + \beta_2 d_j) - \frac{1}{q_j}, \end{aligned}$$

em que,

$$q_j = \frac{1}{1 + e^{\beta_1 + \beta_2 d_j}}.$$

Agora, suponha que para a dose d_j o número de mortos é 100%, então:

$$\begin{aligned}
 (\hat{p}_j - p_j) &\cong p_j(1 - p_j)[\text{logit}^*(\hat{p}_j) - \text{logit}(p_j)] \\
 \Rightarrow (1 - p_j) &\cong p_j(1 - p_j)[\text{logit}^*(\hat{p}_j) - \text{logit}(p_j)] \\
 \Rightarrow \frac{1 - p_j}{p_j(1 - p_j)} &\cong \text{logit}^*(\hat{p}_j) - \text{logit}(p_j) \\
 \Rightarrow \text{logit}^*(\hat{p}_j) &\cong \text{logit}(p_j) + \frac{1}{p_j} \\
 \Rightarrow \text{logit}^*(\hat{p}_j) &\cong (\beta_1 + \beta_2 d_j) + \frac{1}{p_j},
 \end{aligned}$$

em que,

$$p_j = \frac{e^{\beta_1 + \beta_2 d_j}}{1 + e^{\beta_1 + \beta_2 d_j}}.$$

Assim, o procedimento para a análise *logit* deve ser o seguinte: Toma-se todos os pares (d_i, \hat{p}_i) onde $\hat{p}_i \neq 0$ e $\hat{p}_i \neq 1$ e faz-se uma primeira estimativa de β_1 e β_2 por mínimos quadrados ponderados com pesos $w_i = m_i \hat{p}_i \hat{q}_i$, conforme exposto anteriormente. Encontradas estas estimativas, considera-se o *logit* para a dose d_j em que ocorreu $\hat{p}_j = 0$ por:

$$\text{logit}^*(\hat{p}_j) \cong (\hat{\beta}_1 + \hat{\beta}_2 d_j) - \left(1 + e^{\hat{\beta}_1 + \hat{\beta}_2 d_j}\right),$$

e para o caso onde ocorreu $\hat{p}_j = 1$:

$$\text{logit}^*(\hat{p}_j) \cong (\hat{\beta}_1 + \hat{\beta}_2 d_j) + \frac{1 + e^{\hat{\beta}_1 + \hat{\beta}_2 d_j}}{e^{\hat{\beta}_1 + \hat{\beta}_2 d_j}}.$$

De posse destes valores, considera-se todos os valores iniciais juntamente com os valores anteriores e calcula-se novas estimativas para β_1 e β_2 , que serão as estimativas finais do processo.

2.2.4 Delineamentos sequenciais

Os delineamentos utilizados na análise *probit* e *logit* utilizam-se de muitos valores amostrais, além do que existem experimentos em que submeter as unidades experimentais a doses baixas ou altas pode não ser desejável. Os delineamentos sequenciais amenizam este problema e diversos deles serão apresentados agora, sendo que a grande maioria teve origem com o clássico método *up and down* apresentado por Dixon e Mood (1948).

2.2.4.1 O método *up and down*

Este método pode ser utilizado em qualquer experimento de análise de sensibilidade, mas pode ser proveitoso em termos didáticos apresentá-lo em função de um experimento específico. Isso será feito utilizando experimentos com os explosivos discutidos em Dixon e Mood (1948), e ao mesmo tempo com analogia aos experimentos em oncologia apresentados em Stylianou e Flournoy (2002).

No caso dos explosivos, a técnica consiste em escolher alguma altura inicial h_0 , e uma sucessão de alturas h_1, h_2, h_3, \dots acima de h_0 juntamente com uma sucessão de alturas $h_{-1}, h_{-2}, h_{-3}, \dots$ abaixo de h_0 . A primeira amostra de explosivo é testada derrubando o peso para uma altura h_0 . Se a primeira amostra explode, a segunda será testada na altura h_{-1} , caso contrário ela será testada na altura h_1 , em outras palavras, escolhe-se uma altura inicial para o processo, se nessa altura ocorre uma explosão, o próximo explosivo deverá ser testado em uma altura imediatamente inferior, caso isso não ocorra, deve-se subir para altura imediatamente superior. O processo justifica o seu nome, *método up and down*.

Para experimentos com toxicidade, o processo poderia ocorrer de maneira similar, escolhe-se uma dose inicial para começar o processo, se o paciente responde a dose, ou seja, se ele apresenta toxicidade, deve-se baixar a dose para o

próximo paciente, caso contrário deve-se aumentar a dose.

O nome *delineamento sequencial* é bem aplicado a experimentos deste tipo, pois, os tratamentos (altura ou dose de inseticida) são aplicados nas unidades experimentais (explosivo ou paciente com câncer) de modo sequencial, ou seja, a maneira como os tratamentos são aplicados nas parcelas segue um *passeio aleatório* nos valores das alturas ou doses. Assim, o pesquisador não sabe a priori qual parcela receberá determinado tratamento.

Dixon e Mood (1948) argumentam que este método estabiliza muito rapidamente em torno da média, aumentando a acurácia com que a média é estimada, ou seja, para uma dada acurácia, o método *up and down* requer menos testes do que o método ordinário de dividir as parcelas em grupos. Dessa maneira, o tamanho amostral requerido para uma estabilização em torno da média é menor. Obviamente, existem situações em que o tamanho da amostra não é um limitante importante, por exemplo, em experimentos com determinadas espécie de insetos, pode-se ter o tamanho da amostra grande sem muito esforço experimental. Entretanto, isso não ocorre em pesquisas médicas, onde a unidade experimental é o ser humano. Dixon e Mood (1948) argumentam que a economia no número de observações pode ser na ordem de 30 a 40 por cento.

O método foi desenvolvido sob pressuposições, que serão expostas a seguir. Dentre elas, está a normalidade dos dados, que dificilmente é observada em experimentos com explosivos ou dose-resposta. É comum o uso de transformações do tipo logaritmo para alcançar tal feito. O desenvolvimento será realizado pressupondo que essas variáveis já estão normalizadas por algum tipo de transformação.

Seja Y uma variável aleatória normalmente distribuída com média μ e variância σ^2 , e procedimentos serão feitos sobre os $y_{i's}$ observados, com o objetivo de se estimar μ e σ^2 . A variável Y pode representar por exemplo o logaritmo

natural da altura com que os pesos são derrubados. Além disso, será suposto que a probabilidade de ocorrer um sucesso (uma explosão) para um valor (altura) igual ou menor que y_i é igual a $\Phi_{\mu, \sigma^2}(y_i)$, em que Φ_{μ, σ^2} é a acumulada da distribuição normal com média μ e variância σ^2 , e tal probabilidade será denotada por p_i . Assim,

$$p_i = \Phi_{\mu, \sigma^2}(y_i) = \int_{-\infty}^{y_i} \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2\sigma^2}(t-\mu)^2} dt = 1 - q_i. \quad (2.56)$$

Observe que para cada altura selecionada, tem-se uma probabilidade de sucesso (explosão) e fracasso (não explosão) relacionadas. A título de ilustração considere o Exemplo 2.7 adaptado de Dixon e Mood (1948).

Exemplo 2.7 *Considere que em um experimento com explosivos foram selecionadas as seguintes alturas: $y_0 = 2,0m$, $y_1 = 1,7m$, $y_2 = 1,4m$, $y_3 = 1,1m$ e $y_4 = 0,8m$ e que um peso foi derrubado a partir destas alturas sobre o explosivo que explodia ou não dependendo de cada altura e da aleatoriedade do processo. Se ocorria uma explosão em uma determinada altura (sucesso) então a próxima altura inferior era selecionada, caso contrário, a próxima altura superior era selecionada. A sequência na Tabela 3 representa um resultado possível no experimento, a altura inicial é $2,0m$, cada explosão é representada pela letra S de sucesso, e cada não explosão pela letra F de fracasso.*

Tabela 3 Exemplo do passeio aleatório nas alturas para o método up and down.

	y_0	y_1	y_2	y_1	y_2	y_3	y_2	y_1	y_1
Alturas	2,0	1,7	1,4	1,7	1,4	1,1	1,4	1,7	1,4
Passeio	S	S	F	S	S	F	F	S	F

No Exemplo 2.7 a probabilidade de uma explosão na altura $1,4m$ é $p_2 =$

$\Phi_{\mu,\sigma^2}(y_2) = \Phi_{\mu,\sigma^2}(1,4)$ e tal probabilidade é um parâmetro populacional desconhecido, que deverá ser estimado. Entretanto, embora tal parâmetro seja desconhecido, sabe-se que ele depende exclusivamente dos parâmetros μ e σ^2 da distribuição normal.

O número de vezes que o passeio visitou a altura i e ocorreu um sucesso será denotado por n_i , ou seja, n_i é o número de vezes que o passeio visitou a altura i e ocorreu uma explosão. No Exemplo 2.7, o passeio visitou 3 vezes a altura $y_2 = 1,4m$, no entanto, só ocorreu explosão em 1 caso, logo $n_2 = 1$, e para as outras alturas selecionadas: $n_0 = 1$, $n_1 = 3$, e $n_3 = 0$. De maneira análoga, m_i representará o número de vezes que o passeio visitou a altura i e ocorreu uma falha, ou seja, não ocorreu uma explosão. No caso: $m_0 = 0$, $m_1 = 0$, $m_2 = 3$, e $m_3 = 1$.

Defina também as variáveis aleatórias $N = \sum n_i$ e $M = \sum m_i$, ou seja, o número total de sucessos observados durante um ensaio será denotado por N e o número total de falhas por M . Obviamente, se k representa o número total de elementos amostrados a contar do ponto inicial, tem-se que: $k = M + N$. No exemplo 2.7 tem-se os seguintes resultados: $N = 5$, $M = 4$ e $k = 9$.

O processo discutido aqui é essencialmente sequencial, não se sabe à priori qual será o tamanho da amostra, isto é, o número de experimentos de se deixar cair um peso de uma determinada altura sobre o explosivo para que ocorra uma estabilização do sistema.

Observe que se o número k de ensaios é dado, tem-se uma sequência de sucessos S e fracassos F . Note ainda que dada a sequência de S e F e o ponto inicial y_0 , é possível determinar todos os n_i e m_i , como no Exemplo 2.7. Já o contrário não é verdadeiro, dado os n_i e m_i existem várias sequências de S e F possíveis, por exemplo, para $k = 9$, como no Exemplo 2.7, pode-se ter 512 sequências possíveis,

alguns exemplos são: $SSSSFFFS$ ou $SFFSFFFS$.

Considerando um passeio aleatório sobre os pontos

$$\{\dots, y_{-3}, y_{-2}, y_{-1}, y_0, y_1, y_2, y_3, \dots\},$$

tem-se que, associado a este passeio, está uma sequência de sucessos e fracassos, como exposto anteriormente. Além disso, se

$$(n, m) = ((\dots, n_{-2}, n_{-1}, n_0, n_1, n_2, \dots), (\dots, m_{-2}, m_{-1}, m_0, m_1, m_2, \dots)),$$

em que as sequências m e n possuem apenas um número finito de termos diferentes de zero. A distribuição de probabilidade dos n_i e m_i é dada por:

$$P(n, m | y_0) = K \prod_{i=-\infty}^{\infty} p_i^{n_i} q_i^{m_i}, \quad (2.57)$$

em que K é o número de sequências de S e F com (n, m) dado.

Existe uma relação entre n_i e m_{i-1} . Para ocorrer um novo ensaio em y_i tem-se duas possibilidades: ocorreu um sucesso (explosão) do ensaio em y_{i+1} e neste caso volta-se a y_i descendo, ou ocorreu um fracasso (não explodiu) do ensaio em y_{i-1} e neste caso volta-se a y_i subindo (Figura 37). Portanto, quando se tem um sucesso em y_i , o passeio retorna a y_i apenas se ocorrer um fracasso em y_{i-1} nos ensaios subsequentes. Logo, $|n_i - m_{i-1}| = 0$ ou 1 . Dessa maneira, pode-se substituir m_{i-1} por n_i na Equação 2.57, fornecendo uma expressão mais simplificada:

$$P(n, m | y_0) \cong K \prod_{i=-\infty}^{\infty} (p_i q_{i-1})^{n_i}. \quad (2.58)$$

Observe que $M - N$ pode ser grande (em módulo) se ocorrerem sequen-

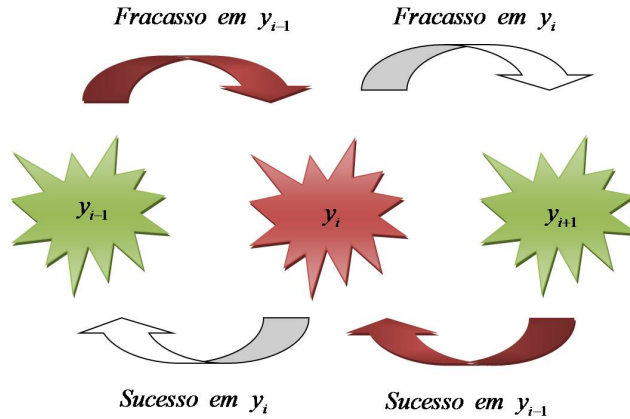


Figura 37 Ilustração do passeio aleatório

cias longas de fracassos (não explosões) ou de sucessos (explosão). Por exemplo, nas seqüências $SSSSSSSSSSSSSSSSSSSSFFFSFS$ ou $FFFFFFFFFFFFFFFF FFFSSFSSF$, tem-se 18 sucessos e 3 fracassos na primeira, com $M - N = 18 - 3 = 15$, tal fato poderia ocorrer se uma escolha do valor da altura inicial y_0 fosse bem acima da altura crítica real, o que provocaria um grande número de explosões até que o passeio começasse a oscilar. Da mesma maneira, uma situação como na segunda seqüência ocorreria, se o valor da altura inicial fosse muito abaixo da altura crítica. No caso da segunda seqüência, tem-se $M - N = 4 - 18 = -14$.

De modo geral, supondo $N \leq M$, pode-se escrever a Equação 2.58 por:

$$P = P(n, m | y_0, M - N) \cong K \prod_{i=-\infty}^{\infty} (p_i q_{i-1})^{n_i}, \quad (2.59)$$

e esta expressão é que será considerada no processo de maximização. Se $M - N$ é pequeno pouca informação estará sendo negligenciada, isso ocorre quando o valor inicial não está muito longe do verdadeiro valor do parâmetro (média). Obviamente, quando $M - N$ é grande haverá pouca contribuição dos valores iniciais do ensaio.

Note então, que toda informação fica definida pelas sequências $\{n_i\}$ ou $\{m_i\}$, mas isso é verdade apenas para este tipo de passeio aleatório e para uma boa escolha do ponto inicial y_0 .

A equação de log-verossimilhança resultante da expressão 2.59 é dada por 2.60:

$$\begin{aligned}\log P &= \log K' + \sum_{-\infty}^{\infty} \log (p_i q_{i-1})^{n_i} \\ &= \log K' + \sum_{-\infty}^{\infty} n_i [\log p_i + \log q_{i-1}].\end{aligned}\quad (2.60)$$

Calculando as derivadas parciais com relação a μ e σ e igualando a zero, tem-se as seguintes equações:

$$\begin{aligned}\frac{\partial \log P}{\partial \mu} &= \sum_{-\infty}^{\infty} n_i \left(\frac{1}{p_i} \frac{\partial p_i}{\partial \mu} + \frac{1}{q_{i-1}} \frac{\partial q_{i-1}}{\partial \mu} \right) = 0 \\ \frac{\partial \log P}{\partial \sigma} &= \sum_{-\infty}^{\infty} n_i \left(\frac{1}{p_i} \frac{\partial p_i}{\partial \sigma} + \frac{1}{q_{i-1}} \frac{\partial q_{i-1}}{\partial \sigma} \right) = 0.\end{aligned}$$

Estas equações dependem das seguintes derivadas parciais

$$\frac{\partial p_i}{\partial \mu}, \frac{\partial q_{i-1}}{\partial \mu}, \frac{\partial p_i}{\partial \sigma}, \frac{\partial q_{i-1}}{\partial \sigma},$$

que serão calculadas a seguir:

$$\begin{aligned}\frac{\partial p_i}{\partial \mu} &= \frac{\partial}{\partial \mu} \int_{-\infty}^{y_i} \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2}\frac{(t-\mu)^2}{\sigma^2}} dt \\ &= \int_{-\infty}^{y_i} \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2}\frac{(t-\mu)^2}{\sigma^2}} \left(-\frac{1}{\sigma^2} \right) 2(t-\mu)(-1) dt \\ &= \int_{-\infty}^{y_i} \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2}\frac{(t-\mu)^2}{\sigma^2}} \left(\frac{1}{\sigma^2} \right) (t-\mu) dt.\end{aligned}$$

Fazendo

$$w = -\frac{(t - \mu)^2}{2\sigma^2}$$

tem-se:

$$\frac{dw}{dt} = -\frac{(t - \mu)}{\sigma^2}$$

ou

$$dt = -\frac{\sigma^2}{(t - \mu)} dw$$

e, para $t \rightarrow -\infty$ tem-se $w \rightarrow -\infty$. Do mesmo modo, para $t = y_i$ tem-se

$$w = -\frac{(t - \mu)^2}{2\sigma^2}.$$

Logo,

$$\begin{aligned} \frac{\partial p_i}{\partial \mu} &= \int_{-\infty}^{-\frac{(y_i - \mu)^2}{2\sigma^2}} \frac{1}{\sqrt{2\pi}\sigma} e^w \left(\frac{t - \mu}{\sigma^2} \right) \left(-\frac{\sigma^2}{t - \mu} \right) dw \\ &= \int_{-\infty}^{-\frac{(y_i - \mu)^2}{2\sigma^2}} \frac{1}{\sqrt{2\pi}\sigma} e^w dw \\ &= \left[-\frac{1}{\sqrt{2\pi}\sigma} e^w \right]_{-\infty}^{-\frac{(y_i - \mu)^2}{2\sigma^2}} \\ &= -\frac{1}{\sqrt{2\pi}\sigma} \left[e^{-\frac{(y_i - \mu)^2}{2\sigma^2}} - \lim_{w \rightarrow -\infty} e^w \right] \\ &= -\frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(y_i - \mu)^2}{2\sigma^2}} \\ &= -\phi(y_i) \\ &= -z_i. \end{aligned}$$

Como $p_{i-1} = 1 - q_{i-1}$ ou $q_{i-1} = 1 - p_{i-1}$, vale o seguinte resultado:

$$\frac{\partial q_{i-1}}{\partial \mu} = \frac{\partial}{\partial \mu} (1 - p_{i-1}) = -\frac{\partial p_{i-1}}{\partial \mu} = z_{i-1}.$$

Portanto, a primeira equação de verossimilhança é dada em 2.61:

$$\frac{\partial \log P}{\partial \mu} = \sum_{-\infty}^{\infty} n_i \left(\frac{z_{i-1}}{q_{i-1}} - \frac{z_i}{p_i} \right) = 0. \quad (2.61)$$

Agora, será encontrada a segunda equação de verossimilhança que depende das derivadas parciais

$$\frac{\partial p_i}{\partial \sigma} \text{ e } \frac{\partial q_{i-1}}{\partial \sigma}.$$

$$\begin{aligned} \frac{\partial p_i}{\partial \sigma} &= \frac{\partial}{\partial \sigma} \int_{-\infty}^{y_i} \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(t-\mu)^2}{2\sigma^2}} dt \\ &= \int_{-\infty}^{y_i} \left(-\frac{1}{\sqrt{2\pi}\sigma^2} \right) e^{-\frac{(t-\mu)^2}{2\sigma^2}} \left[-\frac{1}{2}(t-\mu)^2 \right] \left(-\frac{2}{\sigma^3} \right) dt \\ &= \int_{-\infty}^{y_i} \left(-\frac{1}{\sqrt{2\pi}\sigma^3} \right) e^{-\frac{1}{2}\left(\frac{t-\mu}{\sigma}\right)^2} \left(\frac{t-\mu}{\sigma} \right)^2 dt. \end{aligned}$$

Fazendo

$$w = \frac{t-\mu}{\sigma} \Rightarrow dt = \sigma dw$$

$$t \rightarrow -\infty \Rightarrow w \rightarrow -\infty$$

$$t = y_i \Rightarrow w = \frac{y_i - \mu}{\sigma} = x_i.$$

Assim,

$$\begin{aligned}\frac{\partial p_i}{\partial \sigma} &= \int_{-\infty}^{x_i} \left(-\frac{1}{\sqrt{2\pi}\sigma^3} \right) e^{-\frac{1}{2}w^2} w^2 \sigma dw \\ &= -\frac{1}{\sqrt{2\pi}\sigma^2} \int_{-\infty}^{x_i} e^{-\frac{1}{2}w^2} w^2 dw.\end{aligned}$$

Usando integrações por partes pode-se observar os seguintes resultados:

$$\int e^{-\frac{1}{2}w^2} w^2 dw = \int \underbrace{w}_u \underbrace{we^{-\frac{1}{2}w^2} dw}_{dv},$$

assim,

$$u = w \Rightarrow du = dw$$

$$dv = we^{-\frac{1}{2}w^2} dw \Rightarrow v = -e^{-\frac{1}{2}w^2}.$$

Logo,

$$\int e^{-\frac{1}{2}w^2} w^2 dw = -we^{-\frac{1}{2}w^2} + \int e^{-\frac{1}{2}w^2} dw,$$

O que fornece,

$$\begin{aligned}
\frac{\partial p_i}{\partial \sigma} &= \left[\frac{1}{\sqrt{2\pi\sigma^2}} w e^{-\frac{1}{2}w^2} - \frac{1}{\sqrt{2\pi\sigma^2}} \int e^{-\frac{1}{2}w^2} dw \right]_{-\infty}^{x_i} \\
&= \frac{1}{\sqrt{2\pi\sigma^2}} x_i e^{-\frac{1}{2}x_i^2} - \underbrace{\lim_{w \rightarrow \infty} \left[\frac{1}{\sqrt{2\pi\sigma^2}} w e^{-\frac{1}{2}w^2} \right]}_{=0} - \frac{1}{\sqrt{2\pi\sigma^2}} \int_{-\infty}^{x_i} e^{-\frac{1}{2}w^2} dw \\
&= \frac{1}{\sqrt{2\pi\sigma^2}} x_i e^{-\frac{1}{2}x_i^2} - \frac{1}{\sigma^2} \int_{-\infty}^{x_i} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}w^2} dw \\
&= \frac{1}{\sigma} \left[x_i \underbrace{\frac{1}{\sqrt{2\pi\sigma}} e^{-\frac{1}{2}\left(\frac{y_i-\mu}{\sigma}\right)^2}}_{z_i} - \frac{1}{\sigma} \int_{-\infty}^{y_i} \underbrace{\frac{1}{\sqrt{2\pi\sigma}} e^{-\frac{1}{2}\left(\frac{t-\mu}{\sigma}\right)^2} dt}_{p_i} \right] \\
&= \frac{1}{\sigma} \left[x_i z_i - \frac{p_i}{\sigma} \right],
\end{aligned}$$

onde z_i representa a ordenada da distribuição de Y em y_i e $x_i = (y_i - \mu)/\sigma$.

Agora, $q_{i-1} = 1 - p_{i-1}$, então:

$$\frac{\partial q_{i-1}}{\partial \sigma} = \frac{\partial}{\partial \sigma} (1 - p_{i-1}) = -\frac{1}{\sigma} \left[x_{i-1} z_{i-1} - \frac{p_{i-1}}{\sigma} \right].$$

Logo,

$$\begin{aligned}
\frac{\partial P}{\partial \sigma} &= 0 \\
\Rightarrow \sum_{i=-\infty}^{\infty} n_i \left[\frac{1}{p_i} \frac{\partial p_i}{\partial \sigma} + \frac{1}{q_{i-1}} \frac{\partial q_{i-1}}{\partial \sigma} \right] &= 0 \\
\Rightarrow \sum_{i=-\infty}^{\infty} n_i \left\{ \frac{1}{p_i} \frac{1}{\sigma} \left[x_i z_i - \frac{p_i}{\sigma} \right] + \frac{1}{q_{i-1}} \left(-\frac{1}{\sigma} \right) \left[x_{i-1} z_{i-1} - \frac{p_{i-1}}{\sigma} \right] \right\} &= 0 \\
\Rightarrow \sum_{i=-\infty}^{\infty} n_i \left\{ \frac{x_i z_i}{p_i} - \frac{1}{\sigma} - \frac{x_{i-1} z_{i-1}}{q_{i-1}} + \frac{1}{\sigma} \right\} &= 0 \\
\Rightarrow \sum_{i=-\infty}^{\infty} n_i \left\{ \frac{x_i z_i}{p_i} - \frac{x_{i-1} z_{i-1}}{q_{i-1}} \right\} &= 0.
\end{aligned}$$

Multiplicando a última equação acima por (-1) , tem-se a segunda equação de máxima verossimilhança dada em 2.62:

$$\frac{\partial \log P}{\partial \sigma} = \sum_{i=-\infty}^{\infty} n_i \left\{ \frac{x_{i-1} z_{i-1}}{q_{i-1}} - \frac{x_i z_i}{p_i} \right\} = 0. \quad (2.62)$$

As estimativas de máxima verossimilhança de μ e σ são as raízes ditas $\hat{\mu}$ e $\hat{\sigma}$ das equações 2.61 e 2.62.

Agora, observe que o número de sucessos em y_{i+1} é n_{i+1} , assim, em n_{i+1} vezes o ensaio resulta em sucesso em y_{i+1} e conseqüentemente retorna a y_i , logo, em média o passeio retorna a y_i em $E[n_{i+1}]$ vezes. A probabilidade de ocorrer um sucesso em y_i é p_i , portanto, ocorrem $p_i E[n_{i+1}]$ sucessos em y_i quando o passeio chega em y_i vindo de y_{i+1} . Do mesmo modo, se o passeio está em y_{i-1} , pode-se concluir que o número de sucessos em y_i quando o passeio chega em y_i vindo de y_{i-1} é $p_i E[m_{i-1}]$, ou seja, em média o passeio retorna a y_i vindo de y_{i-1} , $E[m_{i-1}]$ vezes (Figura 38).

$$E[n_i] = p_i E[m_{i-1}] + p_i E[n_{i+1}].$$

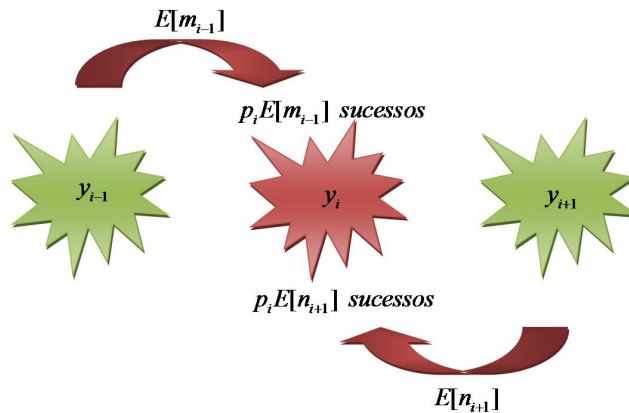


Figura 38 Ilustração do passeio aleatório nas doses relacionando $E[n_i]$ com $E[n_{i+1}]$

Considerando a aproximação $m_{i-1} \simeq n_i$, pode-se dizer que o número de sucessos (explosões) em y_i quando o passeio vem de y_{i-1} é $p_i E[n_i]$. Assim,

$$\begin{aligned}
 E[n_i] &= p_i E[n_i] + p_i E[n_{i+1}] \\
 \Rightarrow (1 - p_i) E[n_i] &= p_i E[n_{i+1}] \\
 \Rightarrow q_i E[n_i] &= p_i E[n_{i+1}] \\
 \Rightarrow \frac{E[n_{i+1}]}{q_i} &= \frac{E[n_i]}{p_i}. \tag{2.63}
 \end{aligned}$$

Da recorrência em 2.63, considerando $i > 0$ tem-se a seguinte relação:

$$\begin{aligned}
 E[n_i] &= \frac{q_{i-1}}{p_{i-1}} E[n_{i-1}] \\
 &= \frac{q_{i-1}}{p_{i-1}} \frac{q_{i-2}}{p_{i-2}} E[n_{i-2}] \\
 &\vdots \\
 &= \frac{q_{i-1}}{p_{i-1}} \frac{q_{i-2}}{p_{i-2}} \dots \frac{q_2}{p_2} \frac{q_1}{p_1} \frac{q_0}{p_0} E[n_0] \\
 &= \prod_{j=0}^{i-1} \frac{q_j}{p_j} E[n_0].
 \end{aligned}$$

Do mesmo modo, para $i < 0$ tem-se:

$$\begin{aligned}
 E[n_i] &= \frac{p_i}{q_i} E[n_{i+1}] \\
 &= \frac{p_i p_{i+1}}{q_i q_{i+1}} E[n_{i+2}] \\
 &\vdots \\
 &= \frac{p_i p_{i+1} p_{i+2} \dots p_{-2} p_{-1}}{q_i q_{i+1} q_{i+2} \dots q_{-2} q_{-1}} E[n_0] \\
 &= \prod_{j=-1}^i \frac{p_j}{q_j} E[n_0].
 \end{aligned}$$

Pode-se simplificar a expressão, da forma:

$$\begin{aligned}
 w_0 &= 1 \\
 w_i &= \prod_{j=0}^{i-1} \frac{q_j}{p_j}, \quad i > 0 \\
 &= \prod_{j=-1}^i \frac{p_j}{q_j}, \quad i < 0.
 \end{aligned}$$

Consequentemente,

$$E[n_i] = w_i E[n_0].$$

Mas,

$$\begin{aligned}
 N &= \sum_{-\infty}^{\infty} E[n_i] \\
 \Rightarrow N &= \sum_{-\infty}^{\infty} w_i E[n_0] \\
 \Rightarrow N &= E[n_0] \sum_{-\infty}^{\infty} w_i
 \end{aligned}$$

$$\Rightarrow E[n_0] = \frac{N}{\sum_{-\infty}^{\infty} w_i}.$$

Portanto,

$$E[n_i] = \frac{Nw_i}{\sum_{-\infty}^{\infty} w_i}. \quad (2.64)$$

Os estimadores de máxima verossimilhança de μ e σ são as raízes das Equações 2.61 e 2.62, e não existe uma forma fechada simples para estas raízes. Dixon e Mood (1948) obtiveram fórmulas aproximadas para estes estimadores.

Para a média, o estimador de máxima verossimilhança é a média amostral ponderada. Tal fato é razoável, pois no método *up and down* o passeio aleatório rapidamente se concentra em torno dos pontos onde as probabilidades p_i não são nem grandes e nem pequenas, isto é, para as doses onde p_i é próximo de $\frac{1}{2}$.

O estimador de máxima verossimilhança aproximado para a média μ é simplesmente:

$$\hat{\mu} = \frac{1}{N} \sum n_i y_i, \quad (2.65)$$

que é a média ponderada das doses pelo número de vezes em que cada uma delas foi visitada.

Uma expressão alternativa para esta fórmula é (DIXON; MOOD, 1948):

$$\hat{\mu} = y' + d \left(\frac{A}{Q} \pm \frac{1}{2} \right), \quad (2.66)$$

em que $A = \sum i n_i$, $Q = \min\{M, N\}$. A análise é feita somente com sucessos ou somente com fracassos, ela será feita com sucessos se $Q = N$ e com fracassos se $Q = M$, assim, n_i representa: ou a frequência com que ocorreu um sucesso na dose y_i ou a frequência com que ocorreu um fracasso na dose y_i . Além disso, só

são considerados na análise doses que foram visitadas pelo passeio.

Na fórmula 2.66, y' é a menor dose visitada com menor frequência, e o fator 0,5 segue de correção de continuidade, sendo que o sinal é “+” se $Q = M$, e o sinal “-” se $Q = N$. O estimador para o desvio padrão é:

$$s = 1,620d \left(\frac{QB - A^2}{Q^2} + 0,029 \right), \quad (2.67)$$

em que $B = \sum i^2 n_i$.

2.2.4.2 O delineamento *up and down* da moeda viciada

Um outro tipo de delineamento sequencial considerado para análise de dados de sensibilidade é o delineamento da moeda viciada (*biased coin design - BCD*) introduzido na literatura por Derman (1957), discutido também em Stylianou e Flournoy (2002). O processo é muito semelhante ao que ocorre no método *up and down* de Dixon e Mood (1948) discutido na seção anterior, com a diferença que, neste passeio o valor quantitativo do tratamento (por exemplo, dose de medicamento) pode aumentar, diminuir ou permanecer o mesmo.

As unidades experimentais continuam com as mesmas características das apresentadas na seção anterior, o que muda é a forma como os tratamentos são alocados.

Considere a situação de um ensaio clínico, por exemplo, testes de determinado medicamento em oncologia, em que um dos objetivos é determinar a dose θ que produzirá toxicidade com probabilidade Γ na população alvo. Suponha que, para a dose $D = d$, $p(d) = P(\text{toxicidade} | D = d)$, onde $p(d)$ é uma função crescente em d , quer se determinar a dose θ , tal que $p(\theta) = \Gamma$, com Γ previamente definido.

Sejam $p_i = p(d_i)$, $i = 1, 2, \dots, k$ e $D = \{d_1, d_2, \dots, d_k\}$ o conjunto das

doses a serem testadas. O processo é iniciado escolhendo-se uma dose inicial. Esta dose inicial depende da experiência do pesquisador. Uma boa sugestão na presença de conhecimento prévio é começar com uma dose próxima da dose desconhecida θ . Tem-se duas situações a serem consideradas no BCD: $\Gamma \leq 0,5$ e $\Gamma > 0,5$.

Para $\Gamma \leq 0,5$, o procedimento começa aplicando-se uma dose d_i em um paciente, se é observado toxicidade, a dose a ser aplicada no próximo paciente é d_{i-1} . Se não é observado toxicidade, o procedimento é aleatorizado da seguinte forma: com probabilidade $b = \Gamma/(1 - \Gamma)$ aplica-se a dose d_{i+1} e com probabilidade $1 - b$ a mesma dose d_i .

Para $\Gamma > 0,5$, o procedimento começa aplicando-se uma dose d_i em um paciente, se não for observado toxicidade, a dose a ser aplicada no próximo paciente é d_{i+1} , caso seja observado toxicidade, o procedimento é aleatorizado da seguinte forma: com probabilidade $b = (1 - \Gamma)/\Gamma$ aplica-se a dose d_{i-1} e com probabilidade $1 - b$ a mesma dose d_i . A Figura 39 apresenta um resumo do BCD.

Para uma justificativa deste delineamento, considere $\Gamma = 0,2$. Neste caso, o interesse é estimar a dose θ que produzirá toxicidade em 20% dos indivíduos da população. Observe que a porcentagem de toxicidade que se quer estimar é pequena, logo, seria interessante que o passeio ficasse estabilizado em torno de doses que atingem um percentual próximo deste valor. Por exemplo, se o passeio começar com uma dose muito alta, é de interesse que essa dose caia rapidamente, pois doses altas produzirão toxicidade com probabilidade alta. Assim, quando a dose produz toxicidade, deve-se baixar um nível da mesma para o próximo paciente. Agora, quando não for constatado toxicidade, deve-se ser mais conservador em aumentar a dose, uma vez que se está interessado em doses baixas. Este procedimento é obtido pelo processo de aleatorização que coloca uma maior probabilidade (0,75) em não aumentar a dose.

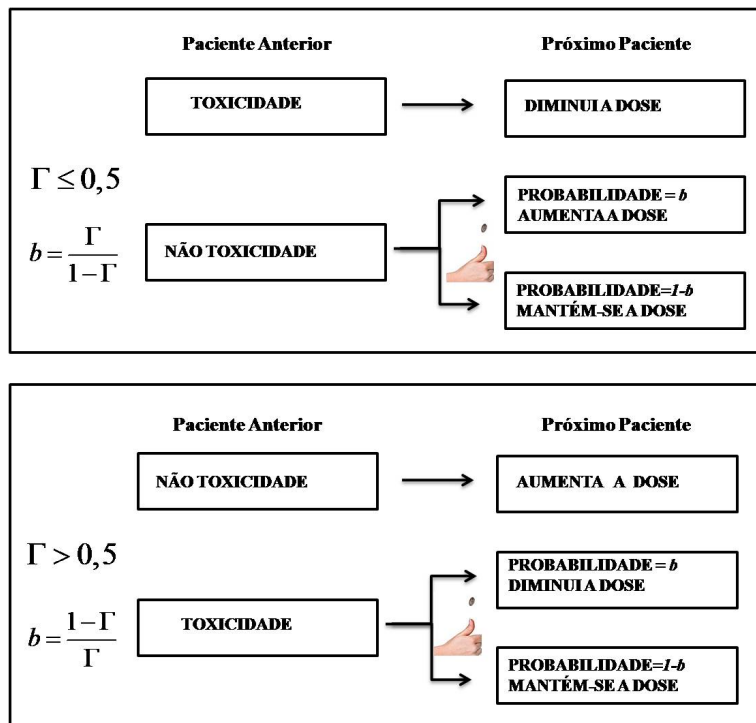


Figura 39 Resumo do Delineamento da Moeda Viciada

Para $\Gamma > 0,5$, o interesse é o inverso, ou seja, quer-se estimar uma dose que produz uma probabilidade de toxicidade alta na população. Por exemplo, para $\Gamma = 0,8$, quer-se estimar uma dose que produz toxicidade em 80% dos indivíduos da população, e neste caso é interessante que o passeio caminhe rapidamente para doses relativamente altas. Isto é obtido da seguinte maneira: quando não é observado toxicidade, a dose é aumentada. Quando a toxicidade é observada, deve-se ser conservador em abaixá-la. Isso é obtido pela aleatorização, em que a probabilidade de se abaixar a dose é $b = 0,25$.

A origem do termo *delineamento da moeda viciada* é devido ao fato de que o processo de aleatorização é realizado de modo semelhante a jogar uma moeda viciada, para então se tomar a decisão.

Dentre as classes dos delineamentos *up and down*, o BCD tem apresentado características ótimas, no sentido de ter uma distribuição das doses mais concentrada em torno de θ (GIOVAGNOLI; PINTACUDA, 1998).

No processo BCD devem ser feitos alguns ajustes para garantir que as doses mais altas ou mais baixas sempre permaneçam dentro da amplitude das doses especificadas pelo experimentador.

É comum, em muitos experimentos clínicos, que a amostra não esteja disponível de imediato, devido ao fato dos pacientes se apresentarem de acordo com a incidência da doença. Além disso, existem em muitos casos, períodos de avaliação pré-determinados por agências reguladoras, que estipulam prazos para a resposta final do paciente ser computada. Por exemplo, se um período de avaliação durar 70 dias e um paciente responder a droga em 40 dias, o pesquisador só poderá fazer uso desta informação após os 70 dias (término do período de avaliação).

Em tais situações o uso do BCD prolongará muito o resultado final do experimento. Diante destas circunstâncias específicas, Stylianou e Follmann (2004)

aprimoraram o BCD, criando o delineamento da moeda viciada acelerado (*accelerated biased coin desing - ABCD*), permitindo a avaliação de múltiplos pacientes no tempo e reduzindo o tempo de espera.

2.2.5 Ensaios clínicos de fase I no tratamento de câncer

Os estatísticos tem contribuído com diversos artigos científicos, com importantes metodologias em experimentos toxicológicos. Este envolvimento por parte destes cientistas tem conduzido a muitas inovações estatísticas. A maior parte destes trabalhos foram motivadas por avaliações farmacológicas, como já destacado nos resultados de Berkson (1944), Bliss (1935) e Dixon e Mood (1948), citados anteriormente.

Segundo Ryan (2000), a toxicologia é o estudo da natureza e mecanismos de efeitos tóxicos de substâncias sobre organismos vivos e outros sistemas biológicos. A complexidade e variabilidade inerente em populações humanas torna difícil a avaliação de efeitos ambientais adversos. Por esta razão, o campo da toxicologia tem tradicionalmente trabalhado com grande ênfase em estudos controlados em laboratório envolvendo animais.

No contexto da avaliação de risco em pacientes com câncer, estratégias alternativas baseadas em modelos de dose-resposta tem amplamente sido aceitas durante décadas em ensaios clínicos oncológicos de Fase I.

Os ensaios clínicos oncológicos de Fase I são estudos planejados, que representam a primeira aplicação de uma nova droga ou combinação de drogas em humanos, e como tal, são fundamentais para o sucesso clínico do processo de desenvolvimento de drogas (STORER, 1989; TOURNEAU; LEE; SIU, 2009).

O objetivo principal de cada ensaio é estabelecer a dose recomendada e/ou programa de uma droga experimental ou combinação de drogas para a realização

de testes eficazes em ensaios de fase II. Segundo Storer (1989), o estudo é planejado para estimar a então chamada máxima dose tolerada (MDT) de uma nova droga. A MDT pode ser considerada como a medida de algum percentual de uma distribuição de tolerâncias com respeito a alguma definição objetiva de toxicidade clínica.

Durante vários anos, foram propostos delineamentos específicos para ensaios de fase I, tais delineamentos tem muitos componentes a serem analisados (TOURNEAU; LEE; SIU, 2009): Dose inicial, incremento de dose, método de escalar as doses, número de pacientes por nível, especificação da toxicidade alvo, definição da MDT e a dose recomendada para ensaios de fase II.

Storer (1989) afirma que as características nestes tipos de ensaios não são adequadas às dos bioensaios, pois estas incluem, por exemplo, um número relativamente pequeno de pacientes disponíveis para o experimento, em que o tempo de avaliação para cada um deles é relativamente longo. Além disso, existem requerimentos éticos para estimar a MDT de modo conservador, sem contar a subjetividade relativa da resposta (tóxica/não tóxica no lugar de vivo/morto), a heterogeneidade da população de pacientes, e a possível dificuldade em classificar a resposta devido à desistência prematura do tratamento, por razões sem conexão com a toxicidade.

Paul, Rosenberger e Flournoy (2004) têm trabalhado em situações onde a resposta em cada dose não é necessariamente binária, podendo ter mais do que dois resultados possíveis. Eles argumentam que existe uma normatização da Organização Mundial da Saúde em que o grau de 0 a 4 é assinalado para refletir o grau de toxicidade para uma lista de doenças baseadas em alguns fatores. O que é comumente usado é uma dicotomização de uma variável aleatória que tem resposta ordinal do tipo: 0 – 2 (nenhuma toxicidade) e 3 – 4 (toxicidade). Estes autores

desenvolvem metodologias sem fazer uso desta dicotomização, argumentando que a mesma tem como consequência a perda de informação.

Segundo Gerke e Siedentop (2007), durante boa parte das últimas décadas a MDT tem sido usualmente estimada sob o enfoque do Delineamento Tradicional, que surgiu antes de 1973. Esta regra também é conhecida como Delineamento $3 + 3$ e é considerado conservador com respeito ao número de toxicidades que ocorrem, porque a subida das doses é executada com cautela. Ela pode conduzir a um número grande de pacientes submetidos a uma mesma dose necessários para estimar a MDT. Tal Delineamento é exposto a seguir.

2.2.5.1 O delineamento tradicional ou 3+3

Resumos deste delineamento podem ser encontrados em Gerke e Siedentop (2007), Storer (1989) e Tourneau, Lee e Siu (2009), dentre outros. Uma versão mais geral do mesmo é apresentada em Leung e Wang (2001) e será apresentada aqui.

Suponha os k níveis de doses sendo $d_1 < \dots < d_k$ e as probabilidades de toxicidade sendo $0 \leq p_1 < \dots < p_k \leq 1$. Grupos de M pacientes são usados e o ensaio começa com a menor dose. A partir daí:

1. Avalia-se M pacientes em d_i :
 - (a) Se ocorre 0 toxicidade, ir para d_{i+1} .
 - (b) Se ocorrer toxicidade e o número de eventos tóxicos for menor ou igual a m_1 , ir para o passo 2, caso contrário parar e recomendar a dose d_{i-1} como a MDT.
2. Avalie um adicional de M pacientes em d_i :

- (a) Se ocorrerem menos do que $m_1 + m_2$ eventos tóxicos no total, deve-se subir a dose para d_{i+1} .
- (b) Caso contrário parar e recomendar a dose d_{i-1} como a MDT.

A estratégia usa um máximo de $N = 2k$ grupos de M pacientes cada para encontrar a MDT. Em ensaios oncológicos é comum utilizar $M = 3$, $m_1 = 1$ e $m_2 = 0$, em que a MDT é a maior dose com uma probabilidade de toxicidade não excedendo um valor crítico, dito $p^* = 0,33$.

Segundo Tourneau, Lee e Siu (2009) este delineamento ainda é muito utilizado, pois é de fácil implementação, não requerendo nenhuma modelagem da curva de toxicidade, e pressupondo apenas que a toxicidade cresce com a dose.

Em geral os grupos tem 3 pacientes. O primeiro grupo é tratado em uma dose considerada segura. Se nenhum dos três pacientes no grupo apresenta toxicidade, outros três pacientes serão tratados no próximo nível de dose mais alto. Entretanto, se um dos primeiros três pacientes apresenta toxicidade, mais três pacientes serão tratados na mesma dose.

A subida continua até que ao menos dois pacientes entre um grupo de três ou seis experimenta uma dose de toxicidade limite (isto é mais do que 33% dos pacientes com toxicidade limite naquele nível de dose). A dose recomendada para o ensaio de fase II é convencionalmente definida pelo nível de dose logo abaixo deste nível de dose tóxica (isto difere entre Estados Unidos e Europa).

Existem propostas de variações deste método, tais como o delineamento $2 + 4$, $3 + 3 + 3$ e $3 + 1 + 1$, que não serão discutidas aqui, mas que são muitos similares ao Delineamento $3 + 3$.

A grande vantagem do delineamento tradicional é que ele é simples e seguro. Entretanto, uma desvantagem do mesmo é que ele envolve um grande número de pacientes que são tratados em doses baixas (isto é, potencialmente sub-

terapêuticas), enquanto poucos pacientes receberão doses realmente próximas da dose recomendada para ensaios de fase II.

Simon et al. (1997) propõem um delineamento alternativo para ensaios de fase I, em que menos pacientes são tratados em níveis de doses subterapêuticos, com redução na duração do ensaio. Entretanto, o delineamento sugerido por eles tem como regra o tratamento feito em grupos de pacientes, não permitindo que somente um paciente seja tratado em cada passo do ensaio.

2.2.5.2 O delineamento *K-in-a-row*

Segundo Ivanova et al. (2003), este delineamento foi introduzido na literatura em 1996, e tal método utiliza mais informação do que o BCD, além de evitar a aleatorização. Estes autores fazem um resumo do mesmo: suponha que o n -ésimo paciente foi alocado no nível de dose d_j , $j = 1, 2, \dots, k$. A regra consiste em alocar o próximo paciente no:

1. nível de dose d_{j-1} se $Y(n) = 1$, isto é, se a última resposta foi tóxica.
2. nível de dose d_{j+1} se todos os K sujeitos mais recentes receberam a dose d_j e não houve toxicidade observada, isto é, $Y(n) = Y(n-1) = \dots = Y(n-K+1) = 0$;
3. nível de dose d_j , caso contrário.

Como pode ser observado, este delineamento utiliza as K últimas informações do ensaio. Para $K = 1, 2, 3, \dots$, tal regra pode ter qualquer alvo $\Gamma = 1 - (0,5)^{\frac{1}{K}}$.

2.2.5.3 O delineamento *up and down* de média móvel

Tal delineamento é bem similar ao *K-in-a-row* e também pode ser encontrado em Ivanova et al. (2003): Suponha que o n -ésimo paciente foi alocado no nível d_j , $j = 1, 2, \dots, k$. Alocar o próximo paciente no:

1. nível de dose d_{j-1} , se existe ao menos uma toxicidade entre as k respostas mais recentes no atual nível de dose.
2. nível de dose d_{j+1} , se não existe nenhuma toxicidade observada entre as k respostas mais recentes no atual nível de dose.

Observe que esta regra não permite que o ensaio repita a mesma dose para o próximo paciente, ela permite apenas subir ou descer a dose para o mesmo, se assemelhando ao método *up and down* de Dixon e Mood (1948), exceto que ela faz uso de mais informação do que apenas o uso da última dose. Para $k = 1, 2, 3, \dots$, a regra *up and down* de média móvel tem alvo $\Gamma = 1 - (0,5)^{\frac{1}{k}}$.

2.2.5.4 O delineamento *up and down* de Narayana

Este delineamento surgiu por volta de 1953, com o objetivo de melhorar o delineamento de Dixon e Mood (1948) para uma probabilidade alvo $\Gamma = 0,5$, ou seja, com o objetivo de se estimar a dose θ_{50} , que produz toxicidade em 50% da população. Neste delineamento a distribuição das doses assinaladas convergem para dois ou três pontos próximo desta dose.

Considere $R_j(n)$ e $N_j(n)$ o número de respostas tóxicas e o número de pacientes assinalados na dose d_j respectivamente, incluindo o n -ésimo paciente. A razão $\frac{R_j(n)}{N_j(n)}$ fornece uma estimativa para a probabilidade de toxicidade na dose d_j . Considere que a razão $\frac{R_j(n)}{N_j(n)}$ é não nula. A regra de Narayana para estimar a dose θ_{50} é dada por: Alocar o próximo paciente para o nível de dose

1. d_{j-1} , se $\frac{R_j(n)}{N_j(n)} > 0,5$ e $Y(n) = 1$;
2. d_{j+1} , se $\frac{R_j(n)}{N_j(n)} < 0,5$ e $Y(n) = 0$;
3. d_j , caso contrário.

2.2.5.5 O delineamento *up and down* de Narayana modificado

Ivanova et al. (2003) fizeram uma modificação no delineamento original de Narayana para situações em que $\Gamma = 1 - (0,5)^{\frac{1}{k}}$. Isso foi feito adicionando uma condição sobre $\frac{R_j(n)}{N_j(n)}$ pela regra de decisão de média móvel *up and down*. A regra é definida da seguinte maneira: Suponha que o n -ésimo paciente foi alocado no nível d_j , $j = 1, 2, \dots, k$. Assinale o próximo paciente no nível de dose:

1. d_{j-1} , se $\frac{R_j(n)}{N_j(n)} > \Gamma$ e se existe ao menos uma toxicidade entre as k respostas mais recentes no atual nível de dose;
2. d_{j+1} , se $\frac{R_j(n)}{N_j(n)} < \Gamma$ e não existe toxicidade observada entre as k respostas mais recentes no atual nível de dose;
3. d_j , caso contrário.

2.2.5.6 O delineamento isotônico de Leung e Wang

Leung e Wang (2001) introduzem um novo método que retém a simplicidade do método tradicional, utilizando apenas a pressuposição de que a toxicidade é não decrescente com a dose. O método consiste em re-escalar as estimativas das toxicidades via regressão isotônica.

A apresentação deste método no artigo original de Leung e Wang (2001) é um pouco confusa, de modo que o texto apresentado a seguir será bem mais

explícito que o artigo original. Os resultados serão dados de modo genérico e será desenvolvido um exemplo para maior clareza.

O método tradicional para estimar a dose que produz determinada toxicidade leva em conta apenas resultados relativos à dose que está sendo experimentada no atual momento, desconsiderando a informação relativa às outras doses. Tal método não faz uso da pressuposição de uma monotonicidade na relação dose-toxicidade. Um método então mais eficiente pode ser considerado.

Considere o vetor com as doses ordenadas $\mathbf{d} = (d_1, d_2, \dots, d_k)$. Para o j -ésimo passo do ensaio considere o vetor de toxicidades observadas até o presente passo $\mathbf{t}^j = (t_1^j, t_2^j, \dots, t_k^j)$ e o vetor com o número de testes efetuados em cada dose $\mathbf{n}^j = (n_1^j, n_2^j, \dots, n_k^j)$, onde t_i^j e n_i^j representam respectivamente o número de casos de toxicidade e o número de testes no nível de dose d_i no passo j .

Em cada passo para as doses que são alcançadas pelo ensaio, considere a estimativa da probabilidade de toxicidade na dose d_i por:

$$\hat{p}_i^j = \frac{t_i^j}{n_i^j},$$

e considere $\hat{p}_i^j = 0$ para as doses que não foram alcançadas pelo ensaio, ou seja, aquelas doses que não foram experimentadas em nenhum paciente. De posse desta definição, considere o vetor de estimativas das probabilidades de toxicidades para cada dose em cada passo como:

$$\hat{\mathbf{p}}^j = (\hat{p}_1^j, \hat{p}_2^j, \dots, \hat{p}_k^j).$$

Em cada passo do ensaio calcule a regressão isotônica do vetor $\hat{\mathbf{p}}^j$ com pesos \mathbf{n}^j usando o PAVA, obtendo o vetor de estimativas isotônicas:

$$\mathbf{p}^{*j} = (p_1^{*j}, p_2^{*j}, \dots, p_k^{*j}).$$

Cada passo do ensaio consiste em atualizar este valor a partir da nova dose indicada seguindo o seguinte procedimento:

1. Considere um grupo de M pacientes em um nível de dose d_i .
2. Encontre o vetor \mathbf{p}^{*j} conforme mostrado anteriormente, escolhendo a dose em que p_i^{*j} está próxima da toxicidade alvo Γ , em que i é a última dose usada. No $(j + 1)$ -ésimo passo:
 - (a) Se $p_i^{*j} < \Gamma$,
 - i. subir um nível de dose se $\Gamma - p_i^{*j} \geq p_{i+1}^{*j} - \Gamma$, $i < k$;
 - ii. continuar no mesmo nível de dose, caso contrário.
 - (b) Se $p_i^{*j} \geq \Gamma$,
 - i. descer um nível de dose se $\Gamma - p_{i-1}^{*j} < p_i^{*j} - \Gamma$, $i > 1$;
 - ii. continuar no mesmo nível de dose, caso contrário.
3. Iteração entre os passos (1) e (2) e parar para algum critério especificado.

Na prática, para $i = 1$, considere $p_0^{*j} = p_1^{*j}$ e para $i = k$, considere $p_{k+1}^{*j} = p_k^{*j}$. A regressão isotônica ordena a toxicidade em um modo monotônico de modo que as decisões de subida ou descida são óbvias pelas comparações sugeridas no passo (2) do algoritmo. O seguinte exemplo é uma adaptação de um exemplo apresentado em Leung e Wang (2001).

Exemplo 2.8 *Assumindo seis níveis de dose e considerando que a verdadeira probabilidade de toxicidade em cada nível é $p_1 = 0,15$, $p_2 = 0,2$, $p_3 = 0,3$,*

$p_4 = 0,65$, $p_5 = 0,96$ e $p_6 = 1$. A toxicidade alvo é $\Gamma = 0,33$. Resultados foram simulados por Leung e Wang (2001), considerando cada grupo experimentado com $M = 3$ pacientes. Os resultados em cada passo estão expostos na Tabela 4. A dose d_1 foi aplicada em um grupo de três pacientes, obtendo uma resposta positiva (toxicidade), então, o vetor com a estimativas isotônicas foi calculado, ou seja, utilizando-se o PAVA, foi calculado a regressão isotônica do vetor $\hat{\mathbf{p}} = (\frac{1}{3}, 0, 0, 0, 0, 0)$ com pesos $\mathbf{n} = (3, 0, 0, 0, 0, 0)$, obtendo-se o vetor $\mathbf{p}^* = (0,33; 0,33; 0,33; 0,33; 0,33; 0,33)$. Os resultados deste vetor são utilizados pelo algoritmo para a tomada de decisão em subir, descer ou permanecer na mesma dose. Neste primeiro caso, $p_1^* = 0,33$, $\Gamma = 0,33$ e $p_0^* = 0,33$, assim: $p_1^* \geq \Gamma$ e $\Gamma - p_0^* \geq p_1^* - \Gamma$, portanto, deve-se continuar na mesma dose. Desta maneira, um novo grupo de 3 pacientes é selecionado recebendo a dose d_1 , deste grupo, apenas 1 apresentou toxicidade, de modo que, considerando os dois grupos na dose d_1 , tem-se uma amostra de 6 pacientes com 2 toxicidades observadas, e o processo é repetido (Passo 2 na Tabela 4), obtendo mais uma vez a dose d_1 para o passo seguinte. Mais uma vez, um grupo de 3 pacientes é submetido à dose d_1 e dessa vez nenhuma toxicidade é observada, então, na dose d_1 foram testados uma amostra com 9 pacientes com 2 toxicidades observadas. O processo é então repetido, até que algum critério estabelecido seja cumprido. A Tabela 4 apresenta resultados para 9 passos, ou seja, uma amostra total de 27 pacientes foi utilizada.

Ivanova e Flournoy (2009) tem comparado este delineamento com outros por meio de simulação, tendo observado que tal delineamento é conservador, além disso, o mesmo tende a administrar doses baixas para muitos pacientes. Estes autores também testaram este delineamento sob diversos tamanhos amostrais por grupo, concluindo que é aceitável fixar pacientes em grupos de não necessaria-

Tabela 4 Ensaio de Fase I simulado para utilização do delineamento isotônico.

Passo		Histórico dos Tratamentos						Dose Atual
		d_1	d_2	d_3	d_4	d_5	d_6	
1	n_i^1	3	0	0	0	0	0	d_1
	t_i^1	1	0	0	0	0	0	
	\hat{p}_i^1	$\frac{1}{3}$	0	0	0	0	0	
	p^{*1}_i	0,33	0,33	0,33	0,33	0,33	0,33	
2	n_i^2	6	0	0	0	0	0	d_1
	t_i^2	2	0	0	0	0	0	
	\hat{p}_i^2	$\frac{2}{6}$	0	0	0	0	0	
	p^{*2}_i	0,33	0,33	0,33	0,33	0,33	0,33	
3	n_i^3	9	0	0	0	0	0	d_1
	t_i^3	2	0	0	0	0	0	
	\hat{p}_i^3	$\frac{2}{9}$	0	0	0	0	0	
	p^{*3}_i	0,22	0,22	0,22	0,22	0,22	0,22	
4	n_i^4	9	3	0	0	0	0	d_2
	t_i^4	2	0	0	0	0	0	
	\hat{p}_i^4	$\frac{2}{9}$	$\frac{0}{3}$	0	0	0	0	
	p^{*4}_i	0,17	0,17	0,17	0,17	0,17	0,17	
5	n_i^5	9	3	3	0	0	0	d_3
	t_i^5	2	0	0	0	0	0	
	\hat{p}_i^5	$\frac{2}{9}$	$\frac{0}{3}$	$\frac{0}{3}$	0	0	0	
	p^{*5}_i	0,13	0,13	0,13	0,13	0,13	0,13	
6	n_i^6	9	3	3	3	0	0	d_4
	t_i^6	2	0	0	2	0	0	
	\hat{p}_i^6	$\frac{2}{9}$	$\frac{0}{3}$	$\frac{0}{3}$	$\frac{2}{3}$	0	0	
	p^{*6}_i	0,13	0,13	0,13	0,67	0,67	0,67	
7	n_i^7	9	3	6	3	0	0	d_3
	t_i^7	2	0	1	2	0	0	
	\hat{p}_i^7	$\frac{2}{9}$	$\frac{0}{3}$	$\frac{1}{6}$	$\frac{2}{3}$	0	0	
	p^{*7}_i	0,17	0,17	0,17	0,67	0,67	0,67	
8	n_i^8	9	3	9	3	0	0	d_3
	t_i^8	2	0	2	2	0	0	
	\hat{p}_i^8	$\frac{2}{9}$	$\frac{0}{3}$	$\frac{2}{9}$	$\frac{2}{3}$	0	0	
	p^{*8}_i	0,17	0,17	0,22	0,67	0,67	0,67	
9	n_i^9	9	3	12	3	0	0	d_3
	t_i^9	2	0	4	2	0	0	
	\hat{p}_i^9	$\frac{2}{9}$	$\frac{0}{3}$	$\frac{4}{12}$	$\frac{2}{3}$	0	0	
	p^{*9}_i	0,17	0,17	0,33	0,67	0,67	0,67	

mente tamanho igual para encurtar a duração do ensaio.

2.2.6 Delineamentos *up and down* generalizados

Segundo Durham, Flournoy e Rosenberger (1997), as estratégias para delineamentos de dose-resposta sequenciais incluem: métodos de aproximação estocástica, métodos de máxima verossimilhança, métodos bayesianos e métodos *up and down*.

Como anteriormente citado, Dixon e Mood (1948) obtiveram resultados elegantes com foco nos estimadores da mediana. Aparentemente, Derman (1957) foi o primeiro a demonstrar que os níveis de dose poderiam ser encontrados ao redor de qualquer quantil pela aleatorização dos delineamentos *up and down* usando uma moeda viciada. Stylianou e Flournoy (2002) voltam a utilizar estas ideias, fazendo o uso de resultados de simulação para acompanhamento do desempenho do mesmo. Adequações são propostas para aumentar o desempenho.

Nesta seção será proposta uma generalização dos métodos *up and down*, que incluirá todos os métodos propostos anteriormente. De acordo com Pace e Stylianou (2007), em ensaios clínicos existem sempre muitas restrições quanto ao número de pacientes envolvidos, desde o custo econômico do desenvolvimento de uma nova droga, até mesmo as questões éticas ligadas ao experimento, de modo que os delineamentos sequenciais do tipo *up and down* passam a ganhar cada vez mais espaço nestas áreas.

Pensando de forma genérica, Giovagnoli e Pintacuda (1998) discutem uma generalização de um problema em inferência não paramétrica, que é a busca por quantis de uma distribuição acumulada F (no caso, a curva de dose-resposta), tipicamente a mediana $LD_{50} = F^{-1}(0,50)$ (*Lethal Dose* - LD, ou Dose Letal a

50%), ou de modo mais geral, $LD_p = F^{-1}(p)$, com $0 < p < 1$, em que:

$$F^{-1}(p) = \inf\{x : F(x) \geq p\}.$$

Considere que níveis crescentes de x são medidas em um conjunto discreto $D = \{d_0, d_1, \dots, d_M, \dots\}$. Um ensaio neste espaço discreto é uma observação aleatória que é tomada para cada d_i . Seja X_n a observação pontual no n -ésimo ensaio, ou seja, X_n é uma variável aleatória que tem como resposta um valor d_i . Considere ainda Y_n a resposta em X_n , tomada como binária: $Y_n = 0$ ou $Y_n = 1$.

Uma definição geral de um Método *up and down* para dados binários com uma curva de resposta quantal $F(x)$ crescente é:

1. Inicie em um ponto x_0 escolhido entre os $d_{k's}$.
2. O ponto para a próxima observação é escolhido de acordo com as seguintes probabilidades de transição condicionais sobre a resposta:

(a) $k \geq 1$:

- $P(X_{n+1} = d_{k-1} \mid X_n = d_k, Y_n = 1) = \gamma_k > 0$,
- $P(X_{n+1} = d_{k-1} \mid X_n = d_k, Y_n = 0) = \gamma'_k \geq 0$,

(b) $k \geq 0$:

- $P(X_{n+1} = d_{k+1} \mid X_n = d_k, Y_n = 0) = \alpha_k > 0$,
- $P(X_{n+1} = d_{k+1} \mid X_n = d_k, Y_n = 1) = \alpha'_k \geq 0$,
- $P(X_{n+1} = d_k \mid X_n = d_k, Y_n = 1) = 1 - \gamma_k - \alpha'_k \geq 0$
- $P(X_{n+1} = d_k \mid X_n = d_k, Y_n = 0) = 1 - \gamma'_k - \alpha_k \geq 0$.

3. $\gamma'_0 = \gamma_0 = 0$ e, se existe um nível máximo d_M (como acontece em muitas situações práticas), é entendido que $\alpha_M = \alpha'_M = 0$. Isto assegura que a probabilidade de ter um nível menor que d_0 ou maior que d_M é nula.

O processo de escolher um valor em D é basicamente uma cadeia de Markov com espaço $\{0,1,\dots,M,\dots\}$, que é condicional à resposta obtida em cada valor d_i em D , com as probabilidades de transição acima citadas. O passeio aleatório considerado é não homogêneo.

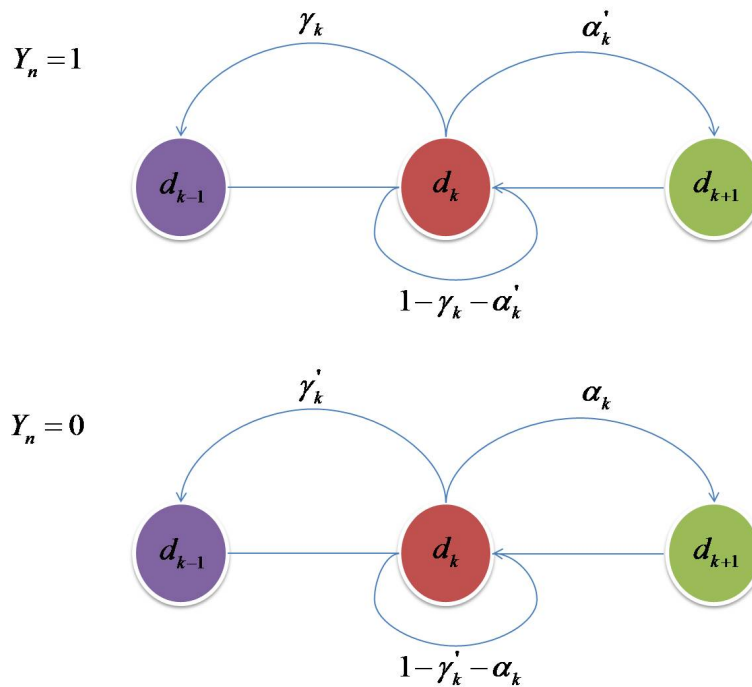


Figura 40 Ilustração de um Delineamento *up and down* Geral

Fazendo analogia aos experimentos de dose-resposta, observa-se que esta definição é mais geral do que as apresentadas anteriormente, pois, independentemente da resposta, sempre existe a possibilidade de descer, subir ou ficar na mesma dose.

Note que, quando a resposta é positiva, existe sempre uma probabilidade de descer ($\lambda_k > 0$), já no caso de uma resposta negativa a probabilidade de descer pode ser zero ou não ($\lambda'_k \geq 0$). Resultado análogo ocorre quando a resposta é

negativa (Figura 40).

O delineamento de Dixon e Mood (1948) e o BCD (moeda viciada) de Derman (1957) e Stylianou e Flournoy (2002) são casos particulares deste delineamento mais geral.

A sequência $\{X_n\}$ é um passeio aleatório não homogêneo no espaço das doses D , com:

$$\begin{aligned} p_k &= \alpha_k(1 - F_k) + \alpha'_k F_k, \\ q_k &= \lambda_k F_k + \lambda'_k(1 - F_k), \end{aligned}$$

em que, $F(d_k) = F_k$, para todo $0 \leq k \leq M$.

Segundo Giovagnoli e Pintacuda (1998), as condições experimentais mais comuns não são muito severas, assegurando a existência da distribuição invariante da cadeia, para a qual o delineamento converge quase certamente pela lei dos grandes números. Estes autores demonstraram que se o alvo do algoritmo é encontrar o quantil $F^{-1}(p)$, sempre que $\{p_k\}$ e $\{q_k\}$ forem sequências crescentes, o algoritmo tem uma única distribuição invariante que é unimodal e sua moda d_m é tal que:

$$d_{m-1} < F^{-1}(p) \leq d_{m+1},$$

sempre que as seguintes condições são satisfeitas,

1. $\alpha_k \geq \alpha'_k$: a probabilidade de subir para uma resposta negativa ($Y_n = 0$) é maior do que para uma resposta positiva ($Y_n = 1$);
2. $\lambda_k \geq \lambda'_k$: a probabilidade de descer para uma resposta positiva ($Y_n = 1$) é maior do que para uma resposta negativa ($Y_n = 0$);

3. $\alpha_k \geq \lambda'_k$: a probabilidade de subir para uma resposta negativa é maior do que a probabilidade de descer para a mesma resposta;
4. $\lambda_k \geq \alpha'_k$: a probabilidade de descer para uma resposta positiva é maior do que a probabilidade de subir para a mesma resposta,

para $k = 1, \dots, M - 1$.

2.2.7 Processo de estimação

Considere que um estudo de dose-resposta utiliza k níveis de doses em ordem crescente, $D = \{d_1 \leq d_2 \leq \dots \leq d_k\}$, e o objetivo de um experimento é encontrar a dose θ com probabilidade de toxicidade Γ (estipulado pelo pesquisador), ou seja, encontrar θ , tal que $P(\theta) = \Gamma$. Para todos os estimadores que serão apresentados a seguir, considere:

- n : o número total de pacientes envolvidos no experimento (tamanho da amostra);
- N_i : o número de pacientes que experimentaram a dose d_i ao final do ensaio;
- R_i : o número de eventos tóxicos no nível de dose d_i ao final do ensaio;
- $X(n)$: o nível de dose do n -ésimo paciente;
- $Y(n)$: uma resposta binária indicando ou não toxicidade no n -ésimo paciente, com $Y(n) = 1$ se o mesmo apresentar toxicidade e $Y(n) = 0$ caso contrário;
- $p(d_i) = p_i$: a probabilidade de toxicidade no nível de dose d_i ;
- $\hat{p}_i = \frac{R_i}{N_i}$: a proporção de sucessos (toxicidades) para cada nível de dose, que é a estimativa natural de p_i ;

- $\{p_i^*\}$: a regressão isotônica de $\{\hat{p}_i\}$ com pesos $\{N_i\}$, que é a estimativa isotônica de p_i .

Vários estimadores vem sendo propostos para ensaios de dose-resposta, dentre os quais destacam-se duas classes: estimadores paramétricos e estimadores não paramétricos. Os estimadores paramétricos fazem uso da pressuposição que a probabilidade de resposta positiva (a presença de toxicidade, por exemplo), são tomadas por meio de uma distribuição, em que, $p_i = F(d_i)$ e F é a função de densidade acumulada de alguma variável aleatória. Os casos mais trabalhados são a distribuição normal ou a distribuição logística.

O'Quigley, Pepe e Fisher (1990) trabalham de forma paramétrica com o uso da estatística bayesiana, estes autores argumentam que na maioria dos ensaios, em particular em ensaios de quimioterapia em pacientes com câncer, é raro que não se tenha nenhuma ideia sobre a curva de dose-resposta. Para eles, tal informação está implicitamente usada na escolha dos níveis de dose disponíveis para o uso do ensaio. Eles propõem atualizar a relação dose-resposta à medida que as observações sobre toxicidade severa se tornam disponíveis.

Embora a metodologia de O'Quigley, Pepe e Fisher (1990) tenha apresentado bons resultados, ela não é tão utilizada na prática, devido à necessidade de um estatístico para a realização das análises e da elaboração do experimento. Segundo, pelo fato de existir muita discussão com relação à distribuição a priori que será considerada, pois qualquer erro nesta hipótese influirá diretamente na vida dos pacientes, podendo expor os mesmos a dose excessivamente altas (TOURNEAU; LEE; SIU, 2009). Tal metodologia não será exposta aqui.

Para os casos não paramétricos não existe nenhuma pressuposição sobre os p_i 's, além daquela de que as probabilidades p_i crescem a medida que os níveis de dose d_i aumentam.

2.2.7.1 Estimadores paramétricos

Stylianou e Flournoy (2002) avaliaram via simulação o desempenho de dois estimadores paramétricos propostos com base no método *logit*, utilizando-se dos métodos dos mínimos quadrados ponderados e máxima verossimilhança:

$$\text{logit}(\Gamma) = \beta_1 + \beta_2\theta.$$

Resolvendo para θ , tem-se os estimadores:

$$\hat{\theta}_\Gamma^7 = \frac{\log\left(\frac{\Gamma}{1-\Gamma}\right) - \hat{\beta}_1^{LS}}{\hat{\beta}_2^{LS}} \quad (2.68)$$

e

$$\hat{\theta}_\Gamma^8 = \frac{\log\left(\frac{\Gamma}{1-\Gamma}\right) - \hat{\beta}_1^{ML}}{\hat{\beta}_2^{ML}}. \quad (2.69)$$

em que $\hat{\beta}_1^{LS}$ e $\hat{\beta}_2^{LS}$ são os Estimadores de Mínimos Quadrados Ponderados e $\hat{\beta}_1^{ML}$ e $\hat{\beta}_2^{ML}$ são os estimadores de máxima verossimilhança. Os autores não avaliaram estimadores, considerando o procedimento *probit*, que pode ser obtido de modo similar.

O estimador $\hat{\theta}_\Gamma^8$ em 2.69 também é avaliado nos trabalhos de Durham, Flournoy e Rosenberger (1997), Gerke e Siedentop (2007), Ivanova et al. (2003), Pace e Stylianou (2007) e Paul, Rosenberger e Flournoy (2004).

2.2.7.2 Estimadores não paramétricos

Estimadores não paramétricos são também interessantes, pois, em muitos ensaios de dose-resposta, não se tem um modelo paramétrico pré-definido. Um dos métodos, *Estimador Empírico da Média*, consiste na média simples truncada de algumas observações (PACE; STYLIANOU, 2007; STYLIANOU; FLOURNOY,

2002):

$$\hat{\theta}_{\Gamma}^9 = \frac{1}{M} \sum_{i=s}^{N+1} y_i \quad (2.70)$$

em que $M = N - s + 2$ e $s = \max\{i: \text{os primeiros } i \text{ pacientes tendo a mesma resposta}\}$.

O truncamento consiste em descartar os primeiros valores idênticos, que surgem em geral por um valor inicial afastado da dose de interesse. Eles argumentam que este estimador tem se mostrado superior ao estimador simples da média, especialmente quando a dose inicial está longe de θ .

Um problema que surge com o uso da proporção de sucessos observado $\hat{p}_i = \frac{y_i}{m_i}$ para estimar p_i é que as estimativas podem não ser crescentes com a dosagem. Nestes casos podem ser usados conceitos de estimação isotônica. Assim, dois estimadores isotônicos são apresentados e avaliados pela primeira vez na literatura por meio de estudos de simulação em ensaios de dose-resposta, utilizando-se de delineamentos do tipo *up and down*, foram apresentados em Stylianou e Flournoy (2002).

O primeiro, denominado *Estimador Isotônico*:

$$\hat{\theta}_{\Gamma}^{10} = \max_{d_i \in D} \{d_i : p_i^* \leq \Gamma\}, \quad (2.71)$$

em que p_i^* é a regressão isotônica de \hat{p}_i . Em outras palavras, $\hat{\theta}_{\Gamma}^{10}$ é a maior dose com uma toxicidade estimada não excedendo um nível pré-especificado Γ . Este estimador também é discutido posteriormente em Pace e Stylianou (2007) e em Stylianou, Proschan e Flournoy (2003). Paul, Rosenberger e Flournoy (2004) fazem uma adaptação do mesmo para o caso de uma resposta não binária.

Observe que $\hat{\theta}_{\Gamma}^{10}$ só pode assumir valores no conjunto das doses D , ou seja,

se está utilizando um estimador com valores discretos para se estimar um parâmetro com valores reais. Desta maneira, pode ser preferível um estimador isotônico com alguns ajustes, denominado *Estimador Isotônico Modificado*. Considerando $\hat{\theta}_\Gamma^{10}$ como na Equação 2.71, este estimador é definido por:

$$\hat{\theta}_\Gamma^{11} = \frac{\Gamma - p_i^*}{p_{i+1}^* - p_i^*} (d_{i+1} - \hat{\theta}_\Gamma^{10}) + \hat{\theta}_\Gamma^{10}. \quad (2.72)$$

O estimador $\hat{\theta}_\Gamma^{11}$ também é analisado em trabalhos de Pace e Stylianou (2007) e Stylianou, Proschan e Flournoy (2003).

3 METODOLOGIA

O objetivo principal deste trabalho é fornecer ferramentas para estimação de quantis de uma distribuição acumulada F , que representa a curva de dose-resposta em diversas situações experimentais, utilizando amostras de pequeno tamanho, fato que geralmente ocorre em ensaios em que imperativos éticos limitam o número de repetições, que ocorrem, por exemplo, em ensaios na área de oncologia.

A proposta então é fornecer métodos estatísticos para estimar uma dose θ , que produzirá toxicidade com probabilidade Γ (especificado pelo pesquisador) em uma população alvo. Tal dose será denotada por θ_Γ . Em todas as análises será considerado que esta toxicidade é crescente com a dose. Desta maneira, será suposto que, para a dose $D = d$, $p(d) = P(\text{toxicidade}|D = d)$ é uma função crescente em d e $F(d) = P(\text{toxicidade}|D \leq d)$. Deseja-se determinar a dose θ , tal que $p(\theta) = \Gamma$, com Γ conhecido a priori, ou seja, para $0 < \Gamma < 1$,

$$\theta_\Gamma = F^{-1}(\Gamma) = \inf \{ \theta : F(\theta) \geq \Gamma \}. \quad (3.1)$$

Considere $D = \{d_1, d_2, \dots, d_k\}$ o conjunto das doses a serem testadas para todas as situações e as probabilidades de toxicidade na dose d_i por $p_i = p(d_i)$, $i = 1, 2, \dots, k$. Serão considerados dois cenários com relação ao processo de estimação: paramétrico e não paramétrico.

Estimadores propostos serão apresentados e comparados com estimadores já apresentados na literatura para diversos tipos de delineamentos. Além disso, está sendo proposto um novo tipo de delineamento, que se utiliza da metodologia bayesiana. Os resultados dos desempenhos das novas metodologias serão comparados com aqueles apresentados na literatura, por meio de simulação estocástica. Os detalhes dos métodos e dos estudos de simulação são dados a seguir.

3.1 Estimadores propostos

Todos os estimadores propostos fazem o uso da pressuposição de que a toxicidade é crescente com a dose. Serão consideradas duas situações nas proposições: estimação paramétrica e estimação não paramétrica.

3.1.1 Estimadores paramétricos

Serão consideradas duas situações para este caso: O caso normal e o logístico. Em ambas as situações, serão propostos métodos por mínimos quadrados ponderados e por máxima verossimilhança. Para o caso normal estes estimadores coincidem. Entretanto, para o caso logístico isso não acontece.

Estimadores com ênfase no métodos *probit* e *logit* tem sido utilizados principalmente para estimar a dose mediana $\theta_{0,50}$, onde tradicionalmente todas as doses têm o mesmo número de repetições, ou seja, para cada dose são testados o mesmo número de pacientes. Isso não ocorre em ensaios oncológicos, principalmente devido às questões éticas ou de expor um número excessivo de pacientes a níveis de dose baixos ou altos.

A maioria dos delineamentos utilizados em ensaios clínicos são do tipo sequenciais, isto é, delineamentos onde o número de pacientes e as doses a serem testadas não são previamente determinadas. Conjectura-se, neste trabalho, que em relação aos delineamentos onde todas as doses são igualmente testadas, o uso de delineamentos sequenciais não modificará severamente as estimativas obtidas pelo *probit* ou *logit*, tendo em vista que as doses mais visitadas por esses ensaios serão aquelas próximo ao quantil alvo. Além disso, os pesos para as doses mais visitadas terão maior influência no processo de estimação.

3.1.1.1 Estimador *probit* isotonzado

Considere os pares de doses (d_i, \hat{p}_i) , em que \hat{p}_i e a proporção de sucessos observados em N_i repetições na dose d_i . Encontre a regressão isotônica $\{p_i^*\}$ de $\{\hat{p}_i\}$ com pesos $\{N_i\}$ obtendo novos pares (d_i, p_i^*) . Com estes novos pares, obtenha os estimadores pelo método *probit* dos coeficientes de regressão β_1 e β_2 . Denote estes estimadores por $\hat{\beta}_1^*$ e $\hat{\beta}_2^*$. O *estimador probit isotonzado - EPI*, para a dose alvo θ_Γ é definido por:

$$\hat{\theta}_\Gamma^1 = \frac{\text{probit}(\Gamma) - \hat{\beta}_1^*}{\hat{\beta}_2^*}. \quad (3.2)$$

Isso segue do fato de que, na população,

$$\text{probit}(\Gamma) = \beta_1 + \beta_2\theta.$$

3.1.1.2 Estimadores *logit* isotonzado

Considerando que a distribuição de tolerâncias é uma logística, serão propostos dois estimadores baseados em regressão isotônica, diferindo apenas pelo método de estimação para a obtenção dos coeficientes de regressão. Tais estimadores para a dose alvo θ_Γ serão chamados de *estimadores logit isotonzado - ELI*, e são definidos da seguinte maneira:

$$\hat{\theta}_\Gamma^2 = \frac{\log\left(\frac{\Gamma}{1-\Gamma}\right) - \hat{\beta}_1^{*LS}}{\hat{\beta}_2^{*LS}} \quad (3.3)$$

e

$$\hat{\theta}_\Gamma^3 = \frac{\log\left(\frac{\Gamma}{1-\Gamma}\right) - \hat{\beta}_1^{*ML}}{\hat{\beta}_2^{*ML}}, \quad (3.4)$$

em que, $\hat{\beta}_1^{*LS}$ e $\hat{\beta}_2^{*LS}$ são os estimadores de mínimos quadrados ponderados e $\hat{\beta}_1^{*ML}$ e $\hat{\beta}_2^{*ML}$ são os estimadores de máxima verossimilhança de β_1 e β_2 respectivamente, obtidos após a isotonização das proporções observadas.

As estimativas de máxima verossimilhança são obtidas por meio da teoria dos modelos lineares generalizados, que se utiliza de métodos numéricos, como o Newton-Rhapson. Existe possibilidade de tais estimativas não serem encontradas ou por conta de valores iniciais impróprios ou tamanho amostral pequeno.

Deseja-se aqui também comparar estes dois estimadores, sob a conjectura de que o estimador de mínimos quadrados ponderados apresentará resultados similares ao de máxima verossimilhança.

As relações apresentadas nas Equações 3.3 e 3.4 podem ser justificadas pela seguinte relação existente na população:

$$\text{logit}(\Gamma) = \beta_1 + \beta_2\theta,$$

em que

$$\text{logit}(\Gamma) = \log\left(\frac{\Gamma}{1-\Gamma}\right).$$

3.1.2 Estimadores não paramétricos

Quando não se tem um modelo paramétrico pré-definido para a curva de dose-resposta, pode-se trabalhar com métodos não paramétricos. Tais métodos não fazem nenhuma pressuposição sobre os p'_i s, exceto que tais probabilidades crescem com a dosagem. Os estimadores não paramétricos discutidos na literatura utilizam basicamente as proporções de sucessos observadas, os mesmos não fazem o uso da frequência com que uma dose é repetida.

Diversos trabalhos tem apontado na direção de que a frequência com que o passeio aleatório nas doses pelos delineamentos anteriormente descritos é tal que

a frequência de visita às doses é unimodal em torno do quantil de interesse θ_Γ , se o tamanho da amostra for suficientemente grande. Como em ensaios clínicos, o tamanho da amostra é geralmente pequeno, tal condição deve ser severamente violada e provavelmente a ausência de estimadores que se utilizem deste fato aconteça por tal motivo.

A Figura 41 (parte (a)) exibe o diagrama de frequências das visitas às doses para realização de um ensaio. Se é conhecido que a distribuição de frequência deste passeio é unimodal, isto é, $f_1 \leq f_2 \leq \dots \leq f_l \leq f_\Gamma$ se $d_1 < d_2 < \dots < d_l < \theta_\Gamma$ e $f_\Gamma \geq f_{l+1} \geq \dots \geq f_k$ se $\theta_\Gamma > d_{l+1} > \dots > d_k$, com f_Γ representando a frequência na dose alvo θ_Γ . Um método de estimação que pode ser utilizado é o da regressão isotônica unimodal. Esta teoria de certa forma segue da teoria da regressão isotônica usual, apesar de apresentar algumas dificuldades. A regressão unimodal está representada na Figura 41 (parte (b)). O cálculo da regressão isotônica unimodal pode ser realizado pelo pacote *Iso* e a função *ufit*.

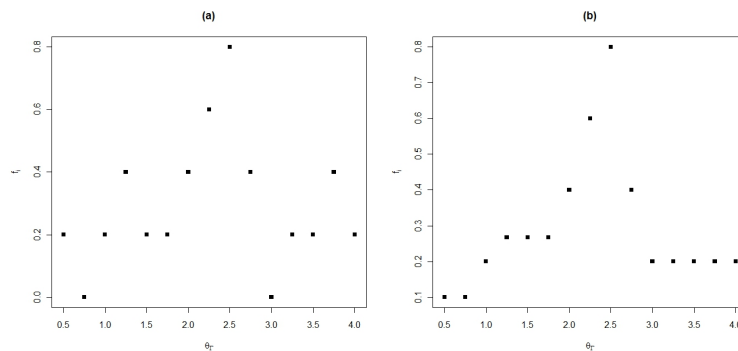


Figura 41 Doses *versus* frequências: (a) frequências resultantes de um ensaio hipotético, (b) frequências isotonizadas por meio de regressão isotônica unimodal

Mediante tais resultados são propostos neste trabalhos dois estimadores não paramétricos para a estimação de um quantil alvo em ensaios de dose resposta,

ambos baseados na Teoria da Regressão Isotônica Unimodal. O primeiro estimador será um estimador “ingênuo”, que será simplesmente a moda da distribuição de frequências após o processo de isotonização. Um segundo estimador irá considerar uma interpolação de tais resultados.

3.1.2.1 Estimador moda da frequência empírica isotonizada

Considere um ensaio de dose resposta em que o par (d_i, f_i) representa a dose utilizada pelo pesquisador e a frequência com que o passeio visitou tal dose. Para a sequência de frequências $\{f_i\}$ considere o Processo de Regressão Unimodal fornecendo como resultados $\{f_i^*\}$. O *estimador moda da frequência empírica isotonizada* - *EMFEI* é definido simplesmente por:

$$\hat{\theta}_\Gamma^A = (d_i : \text{A frequência } f_i^* \text{ é máxima}). \quad (3.5)$$

Este estimador fornece como estimativas doses que foram testadas, fato que ocorre em muitas situações práticas, pois alguns pesquisadores só prescrevem doses que foram testadas.

3.1.2.2 Estimador por interpolação da regressão unimodal

Para explicar a origem do próximo estimador, considere que o processo de regressão isotônica unimodal já foi realizado. Anote a frequência máxima obtida para este processo por f_{\max}^* e a dose relacionada à esta frequência por d_{\max} . Considere ainda que as frequências relacionadas com as doses imediatamente anterior (d_{ant}) e posterior (d_{pos}) são f_{ant}^* e f_{pos}^* respectivamente.

Supõe-se aqui, que a dose alvo θ_Γ será um valor compreendido entre d_{ant} e d_{pos} . O processo a ser utilizado é semelhante àquele realizado para a obtenção da moda pela fórmula de Czuber (para estimação da moda com dados disposto em

uma tabela com distribuição de frequências por intervalos de classe). Note que aqui não se está necessariamente trabalhando com intervalos de classe.

Para ilustrar tal resultado considere a Figura 42. O *estimador por regressão unimodal interpolado - EIRU* $\hat{\theta}_\Gamma^5$ para a dose alvo θ_Γ , será a abscissa do ponto B (Encontro dos segmentos de reta \overline{AD} e \overline{EC}).

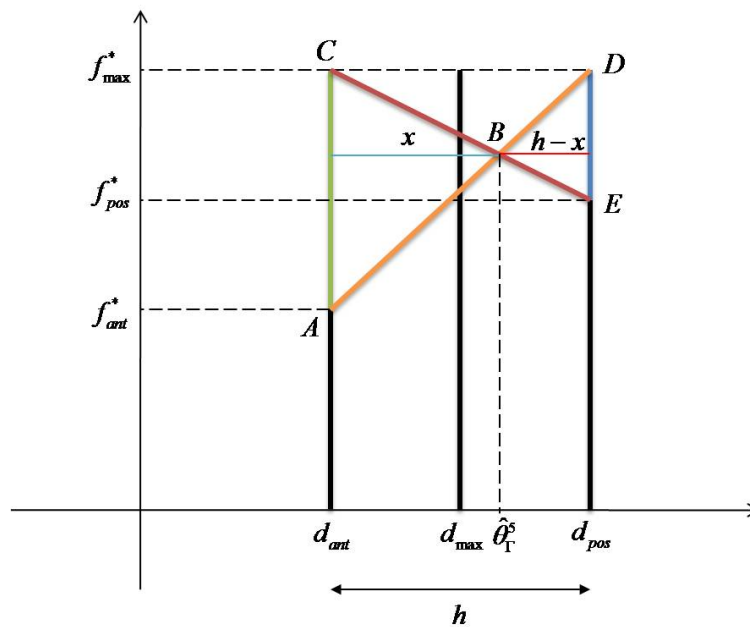


Figura 42 Ilustração para a dedução do estimador por interpolação da regressão unimodal

Observe que $\hat{\theta}_\Gamma^5$ é influenciado diretamente pelas frequências anterior e posterior, de modo que o mesmo pode oscilar no intervalo (d_{ant}, d_{pos}) de acordo com tais frequências. Se $\overline{DE} < \overline{AC}$, tem-se $f_{pos}^* > f_{ant}^*$ e o estimador $\hat{\theta}_\Gamma^5$ estará mais próximo de d_{pos} do que de d_{ant} . Já o contrário, se $\overline{AC} < \overline{DE}$, tem-se $f_{ant}^* > f_{pos}^*$ e $\hat{\theta}_\Gamma^5$ estará mais próximo de d_{ant} .

Na Figura 42, os triângulos ABC e BDE são semelhantes, logo:

$$\begin{aligned}
\frac{x}{h-x} &= \frac{\overline{AC}}{\overline{DE}} \\
\Rightarrow x\overline{DE} &= (h-x)\overline{AC} \\
\Rightarrow x(\overline{DE} + \overline{AC}) &= h\overline{AC} \\
\Rightarrow x &= \left(\frac{\overline{AC}}{\overline{DE} + \overline{AC}} \right) h \\
\Rightarrow x &= \frac{f_{\max}^* - f_{ant}^*}{(f_{\max}^* - f_{pos}^*) + (f_{\max}^* - f_{ant}^*)} (d_{pos} - d_{ant}) \\
\Rightarrow x &= \left(\frac{f_{\max}^* - f_{ant}^*}{2f_{\max}^* - (f_{pos}^* + f_{ant}^*)} \right) (d_{pos} - d_{ant}).
\end{aligned}$$

Portanto,

$$\hat{\theta}_{\Gamma}^{\tilde{5}} = d_{ant} + \left(\frac{f_{\max}^* - f_{ant}^*}{2f_{\max}^* - (f_{pos}^* + f_{ant}^*)} \right) (d_{pos} - d_{ant}). \quad (3.6)$$

3.2 Delineamento proposto

Considere o conjunto das doses $D = \{d_1, d_2, \dots, d_k\}$, com $d_1 < d_2 < \dots < d_k$. O objetivo do ensaio é estimar a dose θ_{Γ} que produz toxicidade Γ na população, ou seja, $p(\theta_{\Gamma}) = \Gamma$. Considere que as probabilidades $p(d_i)$ são variáveis aleatórias tendo alguma distribuição especificada.

Para cada dose d_i associe uma distribuição condicional, tendo densidade na família exponencial 2.26, com $\theta = p(d_i)$ desconhecido e $h = v(d_i)$ conhecido, ou seja,

$$f(y | p(d_i), v(d_i)) = \exp \{ [\Phi(p(d_i)) + (y - p(d_i)) \varphi(p(d_i))] v(d_i) \} \quad (3.7)$$

Note que para cada dose d_i , está associada uma variável aleatória Bernoulli Y_i , com $E[Y_i | p(d_i), d_i] = p(d_i)$. A suposição natural aqui, é que $d_1 < d_2 < \dots < d_k$ implica em $p(d_1) \leq p(d_2) \leq \dots \leq p(d_k)$. O parâmetro de precisão $h = v(d_i)$ deve ser conhecido, tal parâmetro determina a variabilidade dos dados em torno de $p(d_i)$. Foi demonstrado na Afirmação 2.4 que para a distribuição Bernoulli tem-se $h = 1$.

Para cada dose d_i , considere amostras independentes tomadas com tamanho N_i . O pesquisador seleciona um vetor de probabilidades a priori $\mathbf{p}_0 = (p_0(d_1), p_0(d_2), \dots, p_0(d_k))$, em que cada $p_0(d_i)$ é o valor que o pesquisador acha ser plausível, para representar a moda da distribuição da variável aleatória $p(d_i)$. O vetor de componentes aleatórios $\mathbf{p} = (p(d_1), p(d_2), \dots, p(d_k))$ terá distribuição a priori como em 2.34:

$$\pi[\mathbf{p}] \propto \prod_{i=1}^k \exp\{[\Phi(p(d_i)) + (y - p_0(d_i))\varphi(p(d_i))]h_0(d_i)\},$$

se $p(d_i)$ for isotônica sobre D , ou

$$\pi[\mathbf{p}] = 0,$$

caso contrário.

O Teorema 2.15 afirma que a distribuição a posteriori de \mathbf{p} tem densidade na mesma família que a priori com moda $\tilde{\mathbf{p}}^* = (\tilde{p}^*(d_1), \tilde{p}^*(d_2), \dots, \tilde{p}^*(d_k))$, em que cada $\tilde{p}^*(d_i)$ é a regressão isotônica com pesos $w(d_i) = v(d_i)N(d_i) + h_0(d_i)$, da média ponderada de $\bar{y}(d_i)$ e $p_0(d_i)$:

$$\tilde{p}^*(d_i) = \frac{\bar{y}(d_i)v(d_i)N(d_i) + p_0(d_i)h_0(d_i)}{v(d_i)N(d_i) + h_0(d_i)}.$$

Em um ensaio de dose resposta, é natural o pesquisador considerar a variável aleatória $Y_i \mid p(d_i), n \sim Ber(p(d_i))$. Desta maneira, $\bar{y}(d_i)$ será a proporção de respostas positivas para a dose d_i , ou seja,

$$\bar{y}(d_i) = \hat{p}(d_i) = \frac{R(d_i)}{N(d_i)},$$

em que $R(d_i)$ representa o número de resposta positivas para a dose d_i . Assim, como $v(d_i) = 1$ para a Bernoulli,

$$\tilde{p}(d_i) = \frac{\hat{p}(d_i) N(d_i) + p_0(d_i) h_0(d_i)}{N(d_i) + h_0(d_i)}.$$

O delineamento proposto consiste dos quatro seguintes passos:

1. O pesquisador deverá especificar previamente, o vetor com as modas a priori $\mathbf{p}_0 = (p_0(d_1), p_0(d_2), \dots, p_0(d_k))$ e o vetor de precisão a priori $\mathbf{h}_0 = (h_0(d_1), h_0(d_2), \dots, h_0(d_k))$. O vetor com as modas a priori deve obedecer a relação de ordem $p_0(d_1) \leq p_0(d_2) \leq \dots \leq p_0(d_k)$.
2. O pesquisador inicia o ensaio testando um grupo de M pacientes em um nível de dose d_i , e observa a proporção de sucessos $\hat{p}(d_i)$. De posse destes valores (considerando $\hat{p}(d_l) = p_0(d_l)$ e $N(d_l) = 0$, para as doses d_l 's que não foram visitadas pelo ensaio), calcula-se a regressão isotônica $\{\tilde{p}^*(d_i)\}$, com pesos $w(d_i) = N(d_i) + h_0(d_i)$ de

$$\tilde{p}(d_i) = \frac{\hat{p}(d_i) N(d_i) + p_0(d_i) h_0(d_i)}{N(d_i) + h_0(d_i)}.$$

3. Assinale o próximo paciente no nível de dose:

(a) d_{i-1} , se $\tilde{p}^*(d_i) > \Gamma$ e se existe ao menos uma toxicidade entre as k

respostas mais recentes;

(b) d_{i+1} , se $\tilde{p}^*(d_i) < \Gamma$ e não existe toxicidade observada entre as k respostas mais recentes;

(c) d_i , caso contrário.

4. Iteração entre os passos (2) e (3) e parar quando algum critério especificado é conhecido.

As condições do passeio aleatório nas doses, isto é, o comportamento do passeio para d_1 e d_k podem ser convenientemente ajustados. Este delineamento permite ao pesquisador fazer uso de informação conhecida previamente sobre o experimento. Tanto as distribuições a priori como posteriori pertencem a uma família bastante flexível, pois os resultados foram deduzidos considerando a família exponencial.

Observe que não é necessário a especificação detalhada da família, sendo necessária apenas a especificação da moda e do parâmetro de precisão, assim, os cálculos que devem ser efetuados são razoavelmente simples, não sendo necessário o uso de computação intensiva. Este método será chamado no decorrer do trabalho de *delineamento bayesiano isotônico sequencial - DBIS*.

3.3 Estudo de simulação

Foram realizados diversos estudos de Simulação Monte Carlo para a avaliação dos estimadores propostos, comparando o desempenho dos mesmos com estimadores apresentados na literatura sobre diversos tipos de delineamentos, tanto para situações paramétricas como não paramétricas. O delineamento bayesiano isotônico proposto foi simulado, e para o mesmo foram utilizados estimadores não paramétricos. Diversos cenários foram analisados e serão detalhados a seguir.

3.3.1 Comparação dos estimadores

No estudo de simulação, foram avaliados 11 estimadores, dentre eles, 5 propostos. Estes estimadores foram usados em seis delineamentos apresentados na literatura. Diversos tamanhos de amostra foram utilizados, com $n = 3, \dots, 60$, sendo que, para aqueles delineamentos que utilizavam metodologias semelhantes a do Delineamento Tradicional, foram tomadas amostras múltiplas de 3, ou seja, $n = 3, 6, 9, \dots, 60$.

Para todos os cenários foram considerados 8 níveis de doses, ou seja, o vetor de doses $\mathbf{d} = (d_1, d_2, d_3, d_4, d_5, d_6, d_7, d_8)$ com probabilidades associadas $\mathbf{p} = (p_1, p_2, p_3, p_4, p_5, p_6, p_7, p_8)$. Foram consideradas doses igualmente espaçadas e três probabilidades alvo de toxicidade, $\Gamma = 0,10$, $\Gamma = 0,33$ e $\Gamma = 0,50$.

Para os modelos paramétricos foram considerados dados gerados de uma normal e de uma logística, com três variações das médias. A primeira variação consiste em tomar a média como um ponto central entre as doses, neste cenário considera-se que o pesquisador tem uma boa noção da dose média. Na segunda situação, a maior parte das doses ficará acima da média da distribuição e na terceira a maior parte das doses ficará abaixo, simulando situações em que o pesquisador tem maior desconhecimento da média da distribuição. A Tabela 5 apresenta todos os valores que foram utilizados.

O verdadeiro valor do parâmetro θ_Γ nas seis curvas consideradas na Tabela 5 foram:

1. Normal($\mu = 2,25$): $\theta_{0,10} = 0,97$, $\theta_{0,33} = 1,81$ e $\theta_{0,50} = 2,25$;
2. Normal($\mu = 1,25$): $\theta_{0,10} = -0,03$, $\theta_{0,33} = 0,81$ e $\theta_{0,50} = 1,25$;
3. Normal($\mu = 3,25$): $\theta_{0,10} = 1,97$, $\theta_{0,33} = 2,81$ e $\theta_{0,50} = 3,25$;
4. Logística($\mu = 2,25$): $\theta_{0,10} = 1,04$, $\theta_{0,33} = 1,86$ e $\theta_{0,50} = 2,25$;

Tabela 5 Distribuições e probabilidades de toxicidade associadas aos níveis de dose $\mathbf{d} = (0,5; 1,0; 1,5; 2,0; 2,5; 3,0; 3,5; 4,0)$, utilizados no estudo de simulação para a modelagem paramétrica.

Distribuição	μ	σ^2	\mathbf{p}
Normal	2,25	1	(0,04; 0,11; 0,23; 0,40; 0,60; 0,77; 0,89; 0,96)
Normal	1,25	1	(0,23; 0,40; 0,60; 0,77; 0,89; 0,96; 0,99; 1,00)
Normal	3,25	1	(0,00; 0,01; 0,04; 0,11; 0,23; 0,40; 0,60; 0,77)
Logística	2,25	1	(0,04; 0,09; 0,20; 0,39; 0,61; 0,80; 0,91; 0,96)
Logística	1,25	1	(0,20; 0,39; 0,61; 0,80; 0,91; 0,96; 0,98; 0,99)
Logística	3,25	1	(0,01; 0,02; 0,04; 0,09; 0,20; 0,39; 0,61; 0,80)

5. Logística($\mu = 1,25$): $\theta_{0,10} = 0,39$, $\theta_{0,33} = 0,86$ e $\theta_{0,50} = 1,25$;

6. Logística($\mu = 3,25$): $\theta_{0,10} = 2,04$, $\theta_{0,33} = 2,86$ e $\theta_{0,50} = 3,25$;

Cada linha na Tabela 5 representa uma curva de dose-resposta específica. Nas Figuras 43, 44 e 45 estas curvas estão representadas conjuntamente com as probabilidades alvo, $\Gamma = 0,10$, $\Gamma = 0,33$ e $\Gamma = 0,50$, permitindo uma melhor compreensão do cenário. Vale lembrar que para o caso dos modelos paramétricos, é suposto que $p(d_i) = F(d_i)$, em que F (distribuição de tolerâncias) é a função de distribuição acumulada da variável aleatória considerada na modelagem.

Os valores considerados para análise não paramétrica estão expostos na Tabela 6, sua representação conjunta às toxicidades alvo estão representadas na Figura 46.

Tabela 6 Distribuições e probabilidades de toxicidade associadas aos níveis de dose $\mathbf{d} = (0,5; 1,0; 1,5; 2,0; 2,5; 3,0; 3,5; 4,0)$, utilizados no estudo de simulação para a modelagem não paramétrica.

Curvas	Distribuição	\mathbf{p}
Curva 1	exp(0,3)	(0,14; 0,26; 0,36; 0,45; 0,53; 0,59; 0,65; 0,70)
Curva 2	Gama(3;0,5)	(0,08; 0,32; 0,58; 0,76; 0,88; 0,94; 0,97; 0,99)
Curva 3	Gama(2;1,25)	(0,06; 0,19; 0,34; 0,48; 0,59; 0,69; 0,77; 0,83)

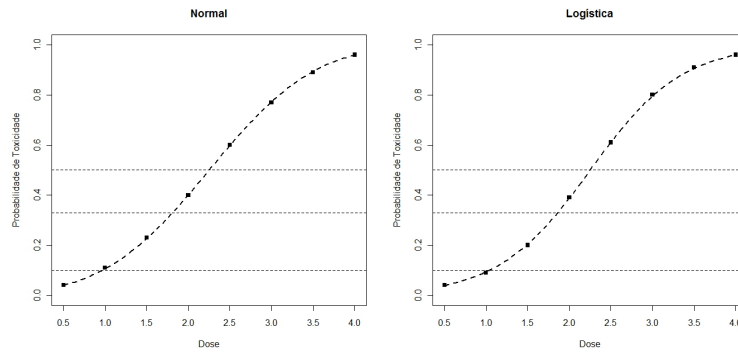


Figura 43 Curva de dose-resposta considerada para o estudo de simulação para dois modelos paramétricos, normal e logístico, ambos com média $\mu = 2,25$ e $\sigma^2 = 1$. As linhas tracejadas representam as probabilidades de toxicidade alvo, $\Gamma = 0,10$, $\Gamma = 0,33$ e $\Gamma = 0,50$

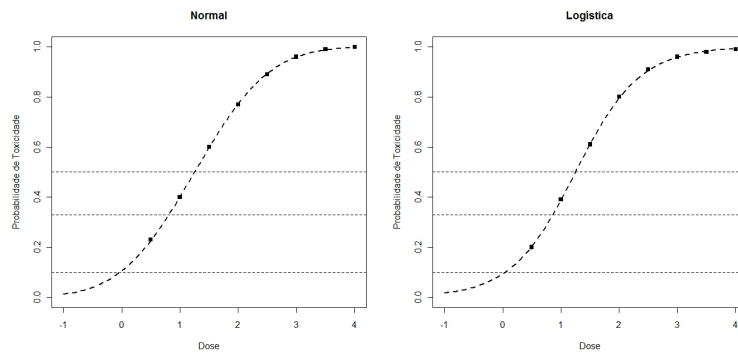


Figura 44 Curva de dose-resposta considerada para o estudo de simulação para dois modelos paramétricos, normal e logístico, ambos com média $\mu = 1,25$ e $\sigma^2 = 1$. As linhas tracejadas representam as probabilidades de toxicidade alvo, $\Gamma = 0,10$, $\Gamma = 0,33$ e $\Gamma = 0,50$

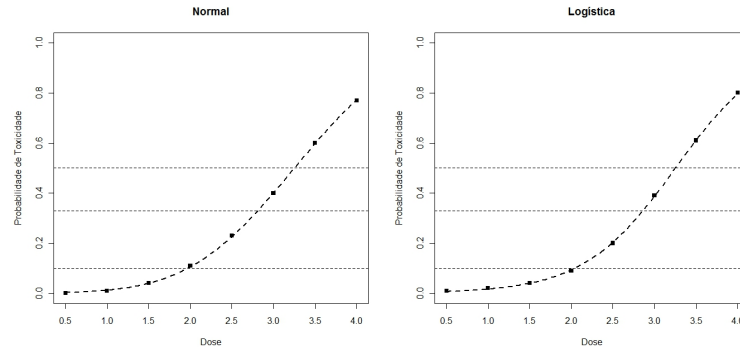


Figura 45 Curva de dose-resposta considerada para o estudo de simulação para dois modelos paramétricos, normal e logístico, ambos com média $\mu = 3,25$ e $\sigma^2 = 1$. As linhas tracejadas representam as probabilidades de toxicidade alvo, $\Gamma = 0,10$, $\Gamma = 0,33$ e $\Gamma = 0,50$

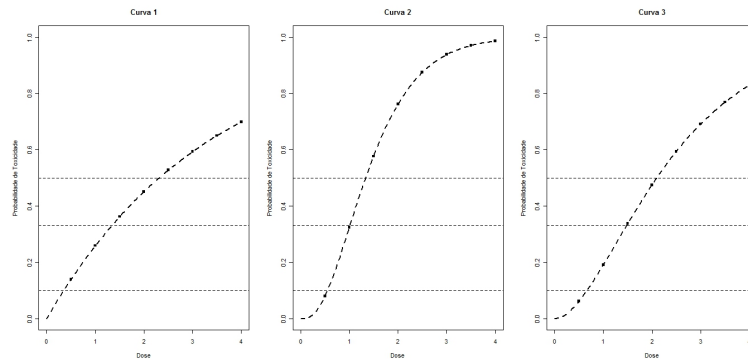


Figura 46 Curva de dose-resposta considerada para o estudo de simulação para os modelos não paramétricos. As linha tracejadas representam as probabilidades de toxicidade alvo, $\Gamma = 0,10$, $\Gamma = 0,33$ e $\Gamma = 0,50$

Assim, no estudo de simulação, sabe-se que o verdadeiro valor do parâmetro θ_{Γ} nas três curvas apresentadas na Tabela 6 são:

1. Curva 1: $\theta_{0,10} = 0,35$, $\theta_{0,33} = 1,33$ e $\theta_{0,50} = 2,31$;
2. Curva 2: $\theta_{0,10} = 0,55$, $\theta_{0,33} = 1,01$ e $\theta_{0,50} = 1,34$;
3. Curva 3: $\theta_{0,10} = 0,66$, $\theta_{0,33} = 1,47$ e $\theta_{0,50} = 2,10$;

Os estimadores não paramétricos também foram testados sobre as curvas apresentadas na Tabela 5 (análise paramétrica). Observe que as probabilidades que foram utilizadas na Tabela 6 são de distribuições conhecidas na simulação. Entretanto, os estimadores não paramétricos não utilizam tal informação no processo de estimação.

3.3.1.1 Os delineamentos

Os delineamentos usados para a comparação dos estimadores foram:

1. O Delineamento da Moeda Viciada - BCD;
2. O Delineamento de Narayana Modificado - DNM;
3. O Delineamento Isotônico de Leung e Wang - DILW.

Os delineamentos BCD e DNM mostram-se ser semelhantes quanto à distribuição das doses aos pacientes. Espera-se que o DNM apresente alguma vantagem quanto ao BCD, dado que o mesmo utiliza informação de outras doses além daquela que está sendo testada. O Delineamento DILW tende a concentrar muitos pacientes em uma única dose, enquanto que o BCD e DNM concentraram a distribuição das doses em torno da dose alvo. Os estudos de simulação de Ivanova e Flournoy (2009) mostraram que o DILW pode não apresentar um desempenho

tão bom, conjecturando que os dados coletados em doses baixas não influenciam as estimativas isotônicas em doses altas. Os próprios Leung e Wang (2001) consideram o delineamento conservador, no sentido de evitar doses altas, potencialmente tóxicas.

3.3.1.2 Os estimadores

Foram utilizados estimadores considerando duas situações possíveis: estimação paramétrica e estimação não paramétrica. Tais estimadores já foram apresentados anteriormente e estão resumidos na Tabela 7. Os estimadores $\hat{\theta}_\Gamma^1, \hat{\theta}_\Gamma^2, \hat{\theta}_\Gamma^3, \hat{\theta}_\Gamma^4$ e $\hat{\theta}_\Gamma^5$ são aqueles que estão sendo propostos neste trabalho.

Tabela 7 Resumo de todos os estimadores analisados no estudo de simulação, tanto para a modelagem paramétrica como não paramétrica.

Estimador	Procedimento
$\hat{\theta}_\Gamma^1 = \frac{probit(\Gamma) - \hat{\beta}_1^*}{\hat{\beta}_2^*}$	Paramétrico - Normal
$\hat{\theta}_\Gamma^2 = \frac{\log\left(\frac{\Gamma}{1-\Gamma}\right) - \hat{\beta}_1^{*LS}}{\hat{\beta}_2^{*LS}}$	Paramétrico - Logística
$\hat{\theta}_\Gamma^3 = \frac{\log\left(\frac{\Gamma}{1-\Gamma}\right) - \hat{\beta}_1^{*ML}}{\hat{\beta}_2^{*ML}}$	Paramétrico - Logística
$\hat{\theta}_\Gamma^4 = (d_i : \text{A frequência } f_i^* \text{ é máxima})$	Não Paramétrico
$\hat{\theta}_\Gamma^5 = d_{ant} + \left(\frac{f_{\max}^* - f_{ant}^*}{2f_{\max}^* - (f_{pos}^* + f_{ant}^*)} \right) (d_{pos} - d_{ant})$	Não Paramétrico
$\hat{\theta}_\Gamma^6 = \frac{probit(\Gamma) - \hat{\beta}_1}{\hat{\beta}_2}$	Paramétrico - Normal
$\hat{\theta}_\Gamma^7 = \frac{\log\left(\frac{\Gamma}{1-\Gamma}\right) - \hat{\beta}_1^{LS}}{\hat{\beta}_2^{LS}}$	Paramétrico - Logística
$\hat{\theta}_\Gamma^8 = \frac{\log\left(\frac{\Gamma}{1-\Gamma}\right) - \hat{\beta}_1^{ML}}{\hat{\beta}_2^{ML}}$	Paramétrico - Logística
$\hat{\theta}_\Gamma^9 = \frac{1}{M} \sum_{i=s}^{N+1} y_i$	Não Paramétrico
$\hat{\theta}_\Gamma^{10} = \max_{d_i \in D} \{d_i : p_i^* \leq \Gamma\}$	Não Paramétrico
$\hat{\theta}_\Gamma^{11} = \frac{\Gamma - p_i^*}{p_{i+1}^* - p_i^*} (d_{i+1} - \hat{\theta}_\Gamma^{10}) + \hat{\theta}_\Gamma^{10}$	Não Paramétrico

3.3.1.3 A avaliação dos estimadores

A avaliação dos estimadores se dará por meio do Erro Quadrático Médio (EQM) e do viés:

$$EQM = \frac{1}{m} \sum_{i=1}^m \left(\hat{\theta}_{\Gamma,i}^j - \theta_{\Gamma} \right)^2 \quad (3.8)$$

e

$$\text{viés} = \frac{1}{m} \sum_{i=1}^m \left(\hat{\theta}_{\Gamma,i}^j - \theta_{\Gamma} \right), \quad (3.9)$$

em que $\hat{\theta}_{\Gamma,i}^j$ é a estimativa de θ_{Γ} obtida pelo estimador $\hat{\theta}_{\Gamma}^j$ na i -ésima amostra e m é o número total de amostras, uma vez que: $EQM = \text{Variância} + \text{viés}^2$. Em todos os cenários considerou-se $m = 2000$.

3.3.2 Avaliação do delineamento proposto

Para avaliação do delineamento bayesiano isotônico sequencial - DBIS, foi utilizado o estimador $\hat{\theta}_{\Gamma}^0$. O desempenho deste estimador sobre o DBIS foi comparado com o desempenho sobre o DNM.

Dois conjuntos de doses foram utilizados: $D_1 = \{0,5; 1,0; 1,5; 2,0; 2,5; 3,0; 3,5; 4,0\}$ e $D_2 = \{0,5; 0,75; 1,0; 1,25; 1,5; 1,75; 2,0; 2,25; 2,5; 2,75; 3,0; 3,25; 3,5; 3,75; 4,0\}$. Espera-se que, com um número maior de doses e consequentemente uma maior variação entre elas, os dados apresentarão um maior número de violadores e o processo de isotonização terá um maior efeito. Neste caso, é de se esperar que o DBIS seja mais eficiente que o DNM.

Foram analisados 21 cenários sob o conjunto D_1 e 18 cenários sob o conjunto D_2 . Para o conjunto D_1 sob o DBIS foram utilizados dois cenários de informação à priori, com parâmetros populacionais diferentes para diversos cenários (Tabela 8). Para o conjunto de doses D_2 sob o DBIS foi considerado um único vetor de modas à priori com variação nos parâmetros conforme Tabela 9.

Tabela 8 Vetores de moda (p_{01} e p_{02}) e precisão (h_{01} e h_{02}) à priori utilizados nos cenários sob o conjunto de dose D_1 , os vetores p_1 , p_2 e p_3 representam os vetores de probabilidades populacionais e h_1 , h_2 e h_3 os de precisão populacionais considerados no estudo de simulação.

Vetores
$p_{01} = (0,02; 0,07; 0,15; 0,25; 0,40; 0,60; 0,70; 0,80)$
$p_{02} = (0,10; 0,15; 0,30; 0,60; 0,80; 0,90; 0,95; 0,99)$
$p_1 = (0,04; 0,11; 0,23; 0,40; 0,60; 0,77; 0,89; 0,96)$
$p_2 = (0,14; 0,26; 0,36; 0,45; 0,53; 0,59; 0,65; 0,70)$
$p_3 = (0,08; 0,32; 0,58; 0,76; 0,88; 0,94; 0,97; 0,99)$
$h_{01} = (2,00; 2,00; 2,00; 2,00; 2,00; 2,00; 2,00; 2,00)$
$h_{02} = (3,00; 3,00; 3,00; 3,00; 3,00; 3,00; 3,00; 3,00)$

Os valores iniciais para o ensaio considerando o DBIS dependiam da informação a priori considerada, sendo escolhida a maior dose menor que a proporção alvo dada no vetor de modas a priori, por exemplo, se a dose alvo era aquela em que 33% da população alvo seria atingida, a dose inicial considerando p_{01} seria $d_4 = 2,0$, e para p_{02} seria $d_3 = 1,5$ (Tabela 8). As tabelas 10 e 11 resumem todos os cenários.

3.3.3 Software

Foram elaboradas rotinas simulando cada delineamento em linguagem R. As análises que se utilizam dos conceitos de regressão isotônica simples foram realizadas através do pacote *isotone* e função *gpava*. Para o cálculo da regressão isotônica unimodal foi considerado o pacote *Iso* e a função *ufit*. As rotinas simulando os delineamentos e os estimadores encontram-se nos APÊNDICES de AA até AP.

Tabela 9 Vetores de moda (p_{03}) e precisão (h_{03} e h_{04}) à priori utilizados nos cenários sob o conjunto de dose D_2 , os vetores p_4 , p_5 e p_6 representam os vetores de probabilidades populacionais e h_4 , h_5 e h_6 os de precisão populacionais considerados no estudo de simulação.

Vetores	
p_{03}	$(0,02; 0,04; 0,11; 0,18; 0,25; 0,30; 0,35; 0,42; 0,49; 0,57; 0,70; 0,84; 0,87; 0,92; 0,94)$
p_4	$(0,04; 0,07; 0,11; 0,16; 0,23; 0,31; 0,40; 0,50; 0,60; 0,69; 0,77; 0,84; 0,89; 0,93; 0,96)$
p_5	$(0,14; 0,20; 0,26; 0,31; 0,36; 0,41; 0,45; 0,49; 0,53; 0,56; 0,59; 0,62; 0,65; 0,68; 0,70)$
p_6	$(0,08; 0,19; 0,32; 0,46; 0,58; 0,68; 0,76; 0,83; 0,88; 0,91; 0,94; 0,96; 0,97; 0,98; 0,99)$
h_{03}	$(2,00; 2,00; 2,00; 2,00; 2,00; 2,00; 2,00; 2,00; 2,00; 2,00; 2,00; 2,00; 2,00; 2,00; 2,00)$
h_{04}	$(20,0; 20,0; 20,0; 20,0; 20,0; 20,0; 20,0; 20,0; 20,0; 20,0; 20,0; 20,0; 20,0; 20,0; 20,0)$

Tabela 10 Cenários utilizados no estudo de simulação para avaliação do DBIS sob o conjunto de doses D_1 . **PP** significa: Parâmetros Populacionais.

Cenário	Delineamento	Distribuição	Γ	p_0	h_0	j	K	PP
(a)	DBIS	$N(2,25; 1)$	0,10	p_{01}	h_{01}	2	7	p_1
(b)	DBIS	$N(2,25; 1)$	0,10	p_{02}	h_{02}	1	7	p_1
(c)	DNM	$N(2,25; 1)$	0,10	—	—	1	7	p_1
(d)	DBIS	$N(2,25; 1)$	0,33	p_{01}	h_{01}	4	2	p_1
(e)	DBIS	$N(2,25; 1)$	0,33	p_{02}	h_{01}	3	2	p_1
(f)	DNM	$N(2,25; 1)$	0,33	—	—	1	2	p_1
(g)	DBIS	$N(2,25; 1)$	0,50	p_{01}	h_{01}	5	1	p_1
(h)	DBIS	$N(2,25; 1)$	0,50	p_{02}	h_{02}	4	1	p_1
(i)	DNM	$N(2,25; 1)$	0,50	—	—	1	1	p_1
(j)	DBIS	$Exp(0,3)$	0,10	p_{01}	h_{01}	2	7	p_2
(m)	DNM	$Exp(0,3)$	0,10	—	—	1	7	p_2
(n)	DBIS	$Exp(0,3)$	0,33	p_{01}	h_{01}	4	2	p_2
(o)	DNM	$Exp(0,3)$	0,33	—	—	1	2	p_2
(p)	DBIS	$Exp(0,3)$	0,50	p_{01}	h_{01}	5	1	p_2
(q)	DNM	$Exp(0,3)$	0,50	—	—	1	1	p_2
(r)	DBIS	$Gama(3; 0,5)$	0,10	p_{01}	h_{01}	2	7	p_3
(s)	DNM	$Gama(3; 0,5)$	0,10	—	—	1	7	p_3
(t)	DBIS	$Gama(3; 0,5)$	0,33	p_{01}	h_{01}	4	2	p_3
(u)	DNM	$Gama(3; 0,5)$	0,33	—	—	1	2	p_3
(v)	DBIS	$Gama(3; 0,5)$	0,50	p_{01}	h_{01}	5	1	p_3
(x)	DNM	$Gama(3; 0,5)$	0,50	—	—	1	1	p_3

Tabela 11 Cenários utilizados no estudo de simulação para avaliação do DBIS sob o conjunto de doses D_2 . **PP** significa: Parâmetros Populacionais.

Cenário	Delineamento	Distribuição	Γ	p_0	h_0	j	K	PP
(A)	DBIS	$N(2,25; 1)$	0,10	p_{03}	h_{03}	2	7	p_4
(B)	DNM	$N(2,25; 1)$	0,10	—	—	1	7	p_4
(C)	DBIS	$N(2,25; 1)$	0,33	p_{03}	h_{03}	6	2	p_4
(D)	DNM	$N(2,25; 1)$	0,33	—	—	1	2	p_4
(E)	DBIS	$N(2,25; 1)$	0,50	p_{03}	h_{03}	9	1	p_4
(F)	DNM	$N(2,25; 1)$	0,50	—	—	1	1	p_4
(G)	DBIS	$Exp(0,3)$	0,10	p_{03}	h_{04}	2	7	p_5
(H)	DNM	$Exp(0,3)$	0,10	—	—	1	7	p_5
(I)	DBIS	$Exp(0,3)$	0,33	p_{03}	h_{04}	6	2	p_5
(J)	DNM	$Exp(0,3)$	0,33	—	—	1	2	p_5
(M)	DBIS	$Exp(0,3)$	0,50	p_{03}	h_{04}	9	1	p_5
(N)	DNM	$Exp(0,3)$	0,50	—	—	1	1	p_5
(O)	DBIS	$Gama(3; 0,5)$	0,10	p_{03}	h_{03}	2	7	p_6
(P)	DNM	$Gama(3; 0,5)$	0,10	—	—	1	7	p_6
(Q)	DBIS	$Gama(3; 0,5)$	0,33	p_{03}	h_{03}	6	2	p_6
(R)	DNM	$Gama(3; 0,5)$	0,33	—	—	1	2	p_6
(S)	DBIS	$Gama(3; 0,5)$	0,50	p_{03}	h_{03}	9	1	p_6
(T)	DNM	$Gama(3; 0,5)$	0,50	—	—	1	1	p_6

4 RESULTADOS E DISCUSSÃO

4.1 Estimadores

Como citado anteriormente, os delineamentos usados para a comparação dos estimadores foram: delineamento da moeda viciada (BCD), delineamento de Narayana modificado (DNM) e o delineamento isotônico de Leung e Wang (DILW) e o conjunto da doses $D = \{0,5; 1,0; 1,5; 2,0; 2,5; 3,0; 3,5; 4,0\}$. São muitos os cenários avaliados. Como os resultados são similares, a apresentação em detalhes se dará apenas para o delineamento DNM. Para os demais delineamentos, apenas em casos particulares serão feitos alguns comentários. Seus resultados completos encontram-se nos APÊNDICES de A até T.

4.1.1 Cenário paramétrico - distribuição normal

Para este cenário foram comparados dois estimadores, $\hat{\theta}_T^1$ (proposto) e $\hat{\theta}_T^6$. O desempenho do estimador proposto $\hat{\theta}_T^1$ com relação ao EQM foi bem superior ao $\hat{\theta}_T^6$ para os cenários que se utilizaram da distribuição normal, como pode ser observado nas Figuras 47, 48,49, 50, 51, 52, 53, 54 e 55. Estes estimadores para o delineamento BCD apresentaram desempenho semelhantes, exceto para o caso em que $\mu = 1,25$ e $\Gamma = 0,10$, uma vez que a dose alvo se encontra fora da amplitude das doses utilizadas (APÊNDICE A).

Para um tamanho de amostra a partir de 16, o EQM do estimador $\hat{\theta}_T^1$ começa a se estabilizar em torno de zero praticamente para todos os cenários. Para o caso $\mu = 3,25$ e $\Gamma = 0,10$ a estabilização do EQM em torno do zero foi mais lenta (Figura 53).

Na literatura consultada foi observado que no caso de delineamentos sequenciais o estimador mais utilizado é o *logit*.

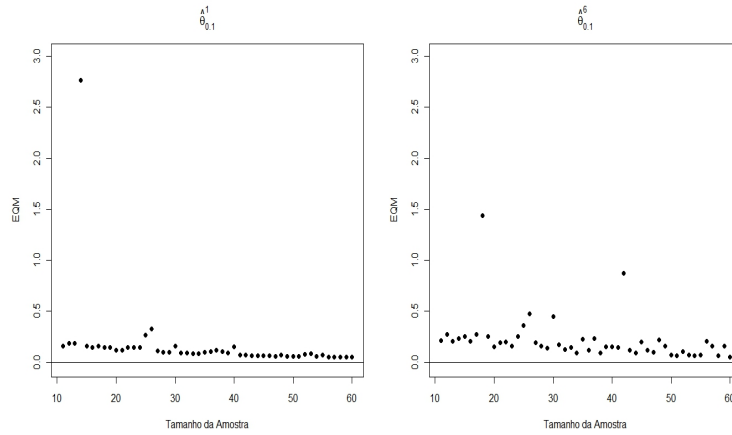


Figura 47 EQM para cada tamanho amostral e os estimadores $\hat{\theta}_{\Gamma}^1$ e $\hat{\theta}_{\Gamma}^6$, amostras geradas considerando uma **Distribuição Normal** com $\mu = 2,25$ e $\sigma^2 = 1$ para $\Gamma = 0,10$ e o **DNM**

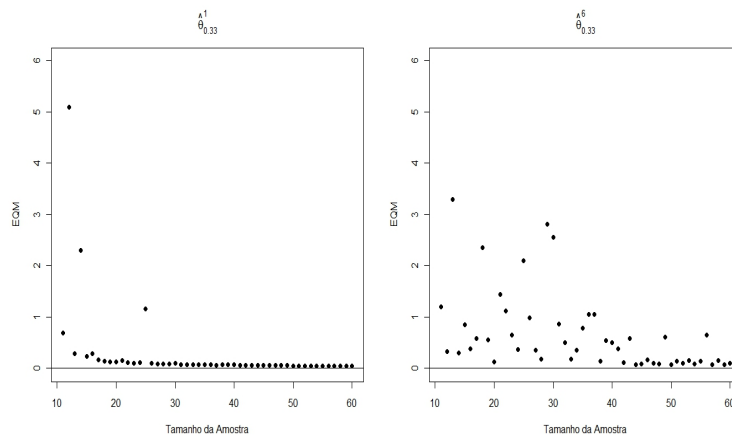


Figura 48 EQM para cada tamanho amostral e os estimadores $\hat{\theta}_{\Gamma}^1$ e $\hat{\theta}_{\Gamma}^6$, amostras geradas considerando uma **Distribuição Normal** com $\mu = 2,25$ e $\sigma^2 = 1$ para $\Gamma = 0,33$ e o **DNM**

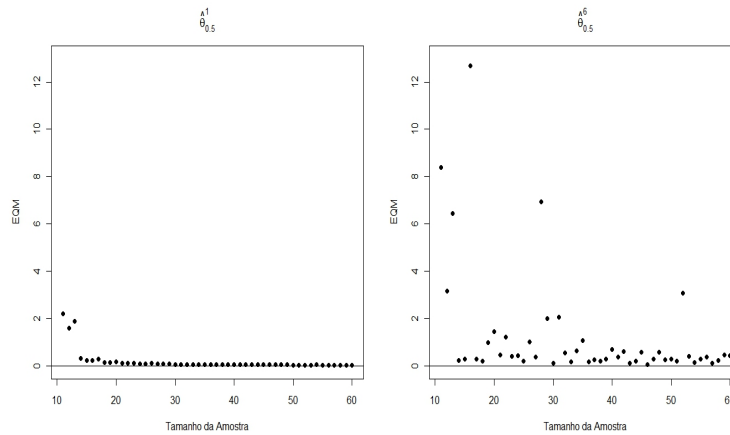


Figura 49 EQM para cada tamanho amostral e os estimadores $\hat{\theta}_{\Gamma}^1$ e $\hat{\theta}_{\Gamma}^6$, amostras geradas considerando uma **Distribuição Normal** com $\mu = 2,25$ e $\sigma^2 = 1$ para $\Gamma = 0,50$ e o **DNM**

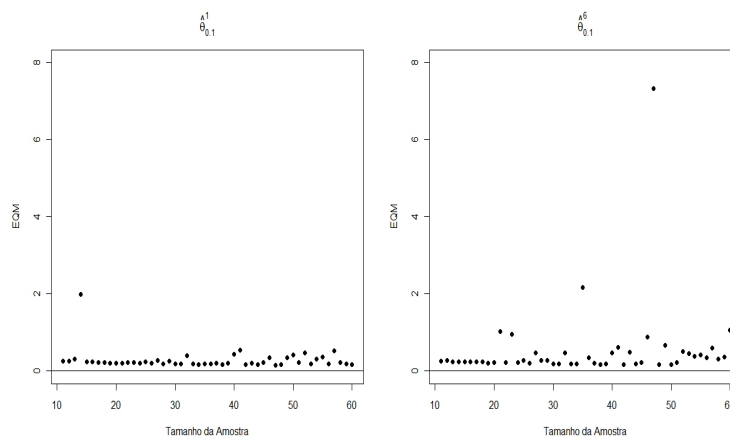


Figura 50 EQM para cada tamanho amostral e os estimadores $\hat{\theta}_{\Gamma}^1$ e $\hat{\theta}_{\Gamma}^6$, amostras geradas considerando uma **Distribuição Normal** com $\mu = 1,25$ e $\sigma^2 = 1$ para $\Gamma = 0,10$ e o **DNM**

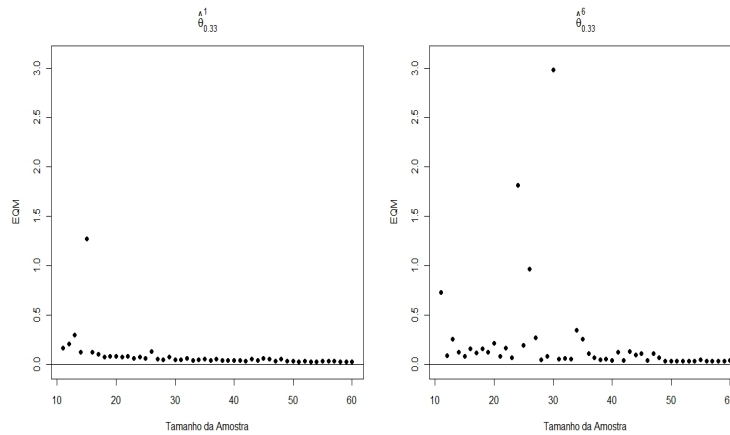


Figura 51 EQM para cada tamanho amostral e os estimadores $\hat{\theta}_{\Gamma}^1$ e $\hat{\theta}_{\Gamma}^6$, amostras geradas considerando uma **Distribuição Normal** com $\mu = 1,25$ e $\sigma^2 = 1$ para $\Gamma = 0,33$ e o **DNM**

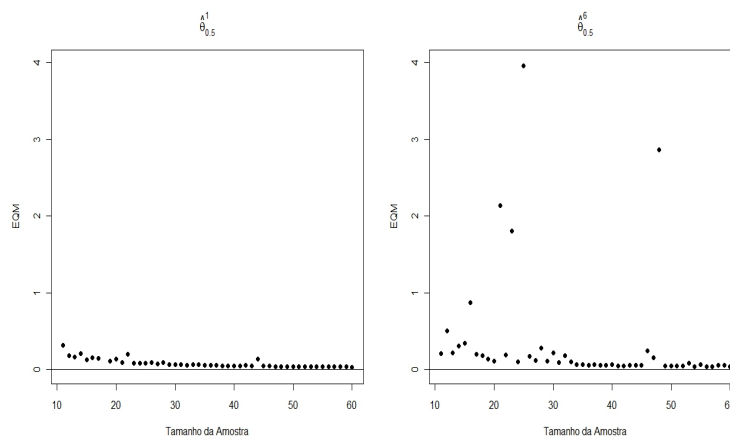


Figura 52 EQM para cada tamanho amostral e os estimadores $\hat{\theta}_{\Gamma}^1$ e $\hat{\theta}_{\Gamma}^6$, amostras geradas considerando uma **Distribuição Normal** com $\mu = 1,25$ e $\sigma^2 = 1$ para $\Gamma = 0,50$ e o **DNM**

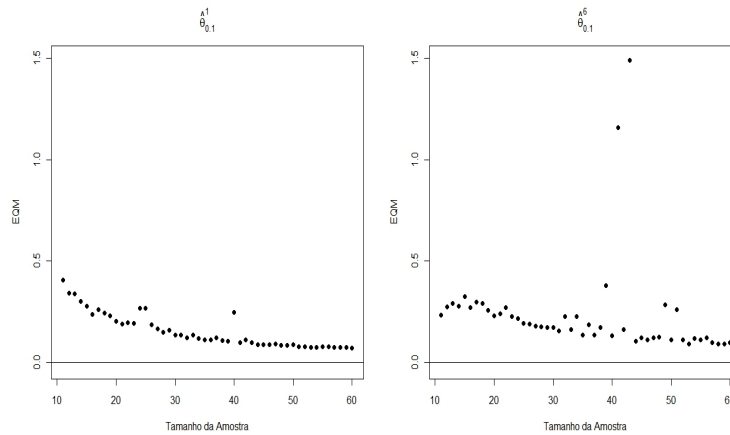


Figura 53 EQM para cada tamanho amostral e os estimadores $\hat{\theta}_{\Gamma}^1$ e $\hat{\theta}_{\Gamma}^6$, amostras geradas considerando uma **Distribuição Normal** com $\mu = 3,25$ e $\sigma^2 = 1$ para $\Gamma = 0,10$ e o **DNM**

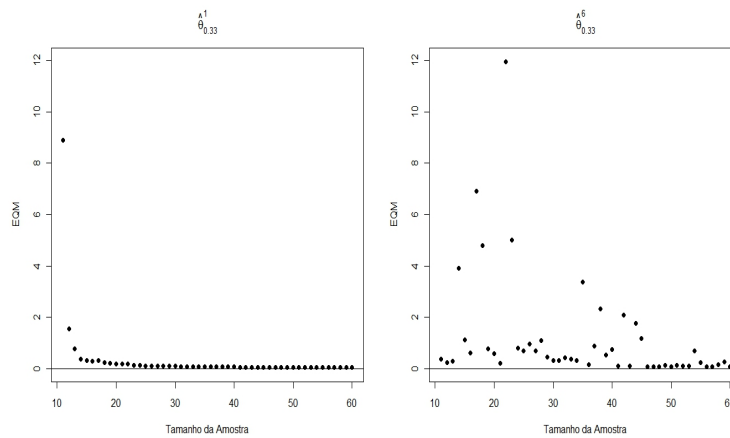


Figura 54 EQM para cada tamanho amostral e os estimadores $\hat{\theta}_{\Gamma}^1$ e $\hat{\theta}_{\Gamma}^6$, amostras geradas considerando uma **Distribuição Normal** com $\mu = 3,25$ e $\sigma^2 = 1$ para $\Gamma = 0,33$ e o **DNM**

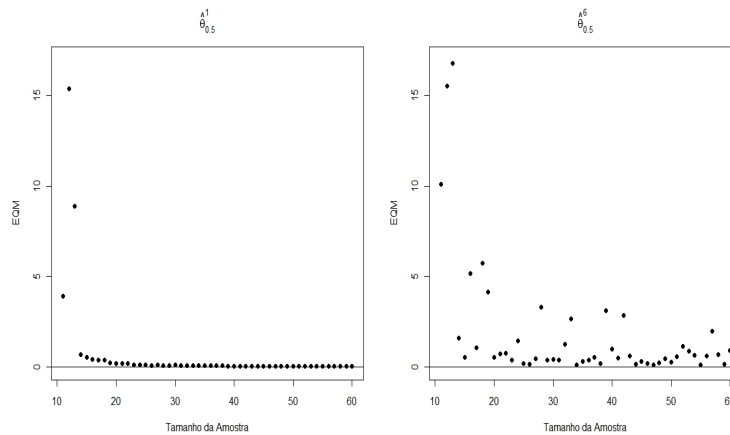


Figura 55 EQM para cada tamanho amostral e os estimadores $\hat{\theta}_{\Gamma}^1$ e $\hat{\theta}_{\Gamma}^6$, amostras geradas considerando uma **Distribuição Normal** com $\mu = 3,25$ e $\sigma^2 = 1$ para $\Gamma = 0,50$ e o **DNM**

Com relação ao viés, houve um desempenho mais modesto do estimador $\hat{\theta}_{\Gamma}^1$ comparado ao estimador $\hat{\theta}_{\Gamma}^6$, mesmo assim, o viés de $\hat{\theta}_{\Gamma}^1$ se mostrou mais estável do que o viés de $\hat{\theta}_{\Gamma}^6$ para os cenários que utilizaram a distribuição normal, como pode ser observado nas figuras 56, 57, 58, 59, 60, 61, 62, 63 e 64.

No estudo de simulação foram excluídas situações em que o método não poderia ser aplicado, por exemplo, houve casos em que em determinado ensaio nenhuma proporção de sucessos foi observada, desta maneira, o *probit* não poderia ser aplicado. Tal fato ocorreu, em geral, para amostras de tamanho pequeno, mas também ocorreu para amostras grandes com certa raridade. Ocorreram também situações em que as estimativas de β_2 eram muito próximas de zero, fornecendo assim estimativas inviáveis do ponto de vista prático. Um outro caso que impossibilitou a aplicação do método foi aquele em que todas as proporções observadas eram iguais a zero ou um. Neste caso o estimador $\hat{\theta}_{\Gamma}^6$ não poderia ser aplicado. Este estimador também não pode ser aplicado em casos onde ao descartar as

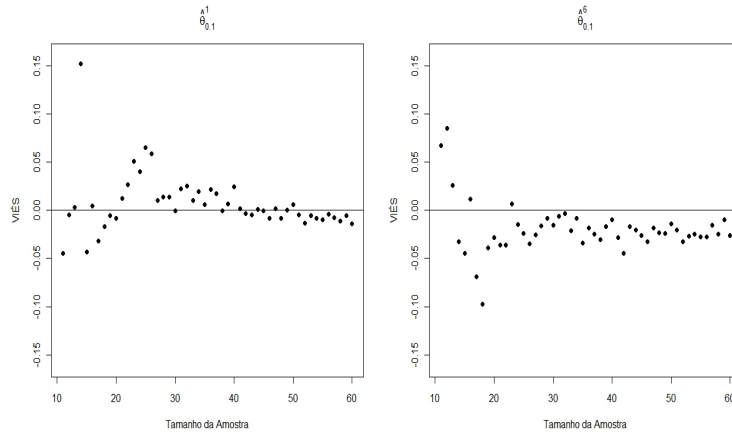


Figura 56 Viés para cada tamanho amostral e os estimadores $\hat{\theta}_{\Gamma}^1$ e $\hat{\theta}_{\Gamma}^6$, amostras geradas considerando uma **Distribuição Normal** com $\mu = 2,25$ e $\sigma^2 = 1$ para $\Gamma = 0,10$ e o **DNM**

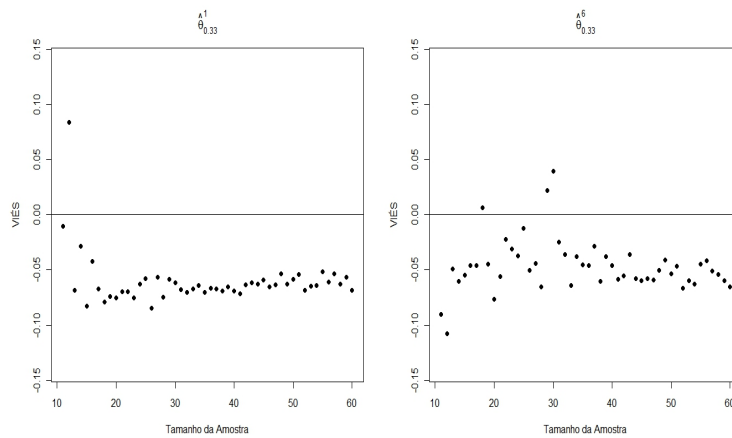


Figura 57 Viés para cada tamanho amostral e os estimadores $\hat{\theta}_{\Gamma}^1$ e $\hat{\theta}_{\Gamma}^6$, amostras geradas considerando uma **Distribuição Normal** com $\mu = 2,25$ e $\sigma^2 = 1$ para $\Gamma = 0,33$ e o **DNM**

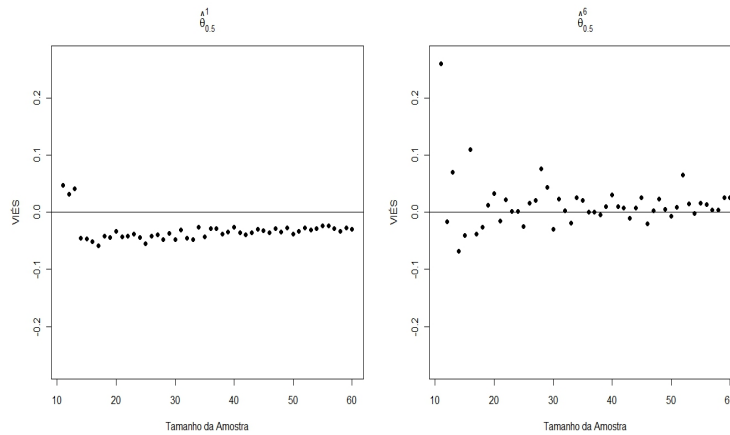


Figura 58 Viés para cada tamanho amostral e os estimadores $\hat{\theta}_{\Gamma}^1$ e $\hat{\theta}_{\Gamma}^6$, amostras geradas considerando uma **Distribuição Normal** com $\mu = 2,25$ e $\sigma^2 = 1$ para $\Gamma = 0,50$ e o **DNM**

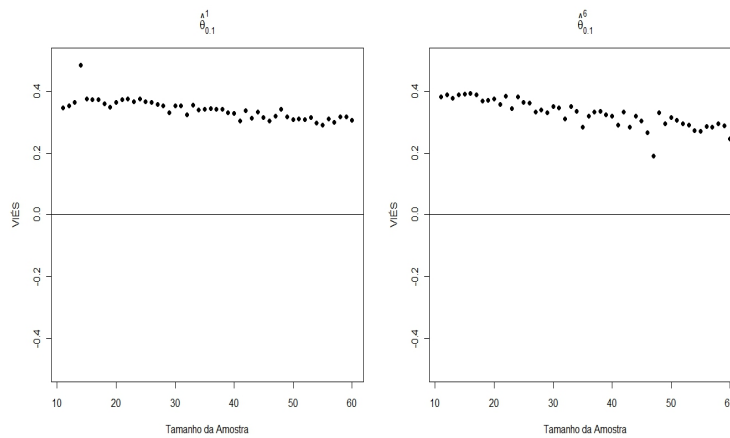


Figura 59 Viés para cada tamanho amostral e os estimadores $\hat{\theta}_{\Gamma}^1$ e $\hat{\theta}_{\Gamma}^6$, amostras geradas considerando uma **Distribuição Normal** com $\mu = 1,25$ e $\sigma^2 = 1$ para $\Gamma = 0,10$ e o **DNM**

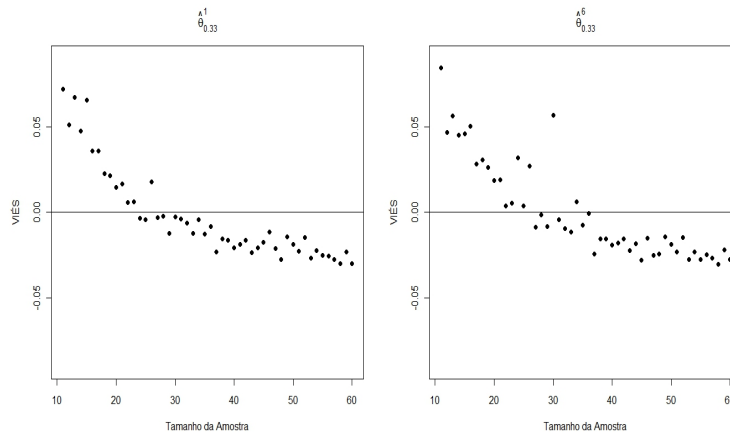


Figura 60 Viés para cada tamanho amostral e os estimadores $\hat{\theta}_{\Gamma}^1$ e $\hat{\theta}_{\Gamma}^6$, amostras geradas considerando uma **Distribuição Normal** com $\mu = 1,25$ e $\sigma^2 = 1$ para $\Gamma = 0,33$ e o **DNM**

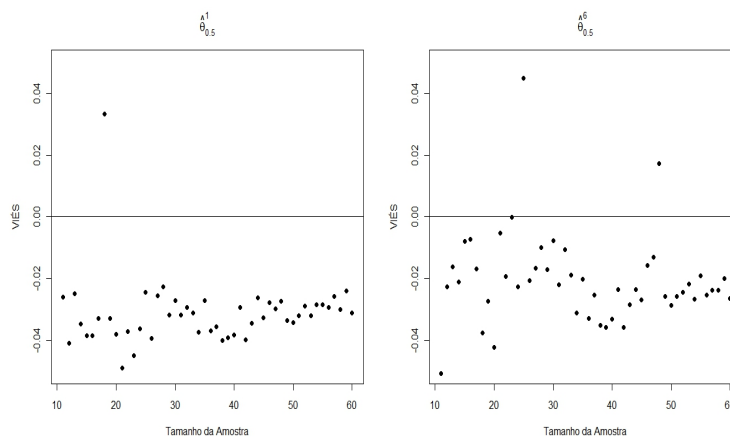


Figura 61 Viés para cada tamanho amostral e os estimadores $\hat{\theta}_{\Gamma}^1$ e $\hat{\theta}_{\Gamma}^6$, amostras geradas considerando uma **Distribuição Normal** com $\mu = 1,25$ e $\sigma^2 = 1$ para $\Gamma = 0,50$ e o **DNM**

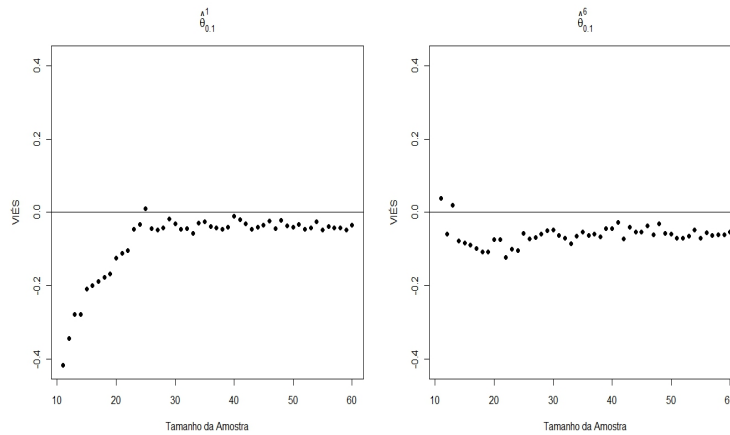


Figura 62 Viés para cada tamanho amostral e os estimadores $\hat{\theta}_{\Gamma}^1$ e $\hat{\theta}_{\Gamma}^6$, amostras geradas considerando uma **Distribuição Normal** com $\mu = 3,25$ e $\sigma^2 = 1$ para $\Gamma = 0,10$ e o **DNM**

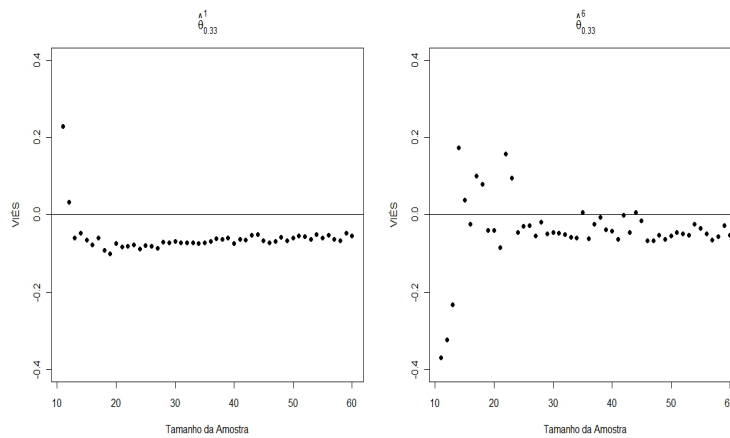


Figura 63 Viés para cada tamanho amostral e os estimadores $\hat{\theta}_{\Gamma}^1$ e $\hat{\theta}_{\Gamma}^6$, amostras geradas considerando uma **Distribuição Normal** com $\mu = 3,25$ e $\sigma^2 = 1$ para $\Gamma = 0,33$ e o **DNM**

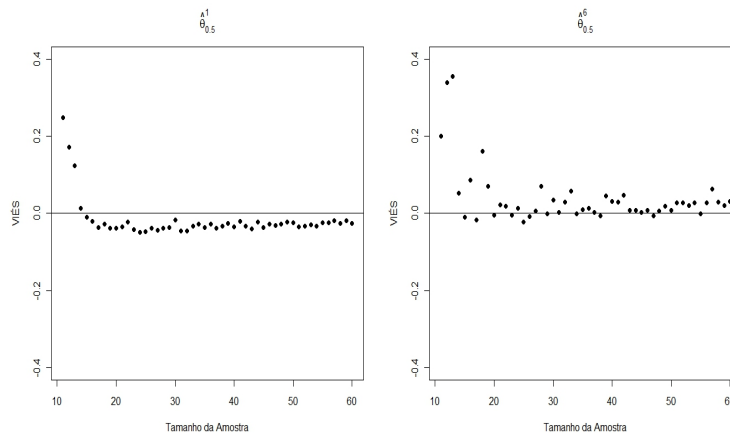


Figura 64 Viés para cada tamanho amostral e os estimadores $\hat{\theta}_{\Gamma}^1$ e $\hat{\theta}_{\Gamma}^6$, amostras geradas considerando uma **Distribuição Normal** com $\mu = 3,25$ e $\sigma^2 = 1$ para $\Gamma = 0,50$ e o **DNM**

proporções zero e um resulte apenas um ponto para o estudo da regressão linear. A isotonização das proporções feita pelo estimador $\hat{\theta}_{\Gamma}^1$, antes da aplicação do *probit* amenizou tais problemas. Sendo que, para tamanho de amostras menores o método ainda pode ser aplicado com maior eficiência do que o estimador $\hat{\theta}_{\Gamma}^6$.

Ensaio reais que forneçam resultados estranhos devem ser analisados com cuidado, por exemplo, um ensaio em que a proporção de sucessos observados é zero para todas as doses pode ser obtido em uma amostra (rara) que seja resistente à droga, tal amostra pode ser objeto de estudo, entretanto, as estimativas obtidas para tal ensaio não são confiáveis. Uma outra explicação para resultados deste tipo pode ser uma escolha errada dos níveis de doses utilizados.

4.1.2 Cenário paramétrico - distribuição logística

Para este cenário foram estudados quatro estimadores paramétricos: dois propostos, $\hat{\theta}_{\Gamma}^2$ e $\hat{\theta}_{\Gamma}^3$ e dois da literatura, $\hat{\theta}_{\Gamma}^7$ e $\hat{\theta}_{\Gamma}^8$.

Com relação ao EQM, o desempenho dos estimadores propostos foi superior aos outros dois, apesar de o estimador $\hat{\theta}_{\Gamma}^8$ também apresentar bons resultados, como pode ser visto nas Figuras 65, 66, 67, 68, 69, 70, 71, 72 e 73.

O tamanho amostral para o qual houve estabilização do EQM do estimadores $\hat{\theta}_{\Gamma}^2$ e $\hat{\theta}_{\Gamma}^3$ em torno de zero é a partir de 10. Em relação ao estimador $\hat{\theta}_{\Gamma}^8$ os resultados obtidos estão de acordo com os obtidos por Ivanova et al. (2003).

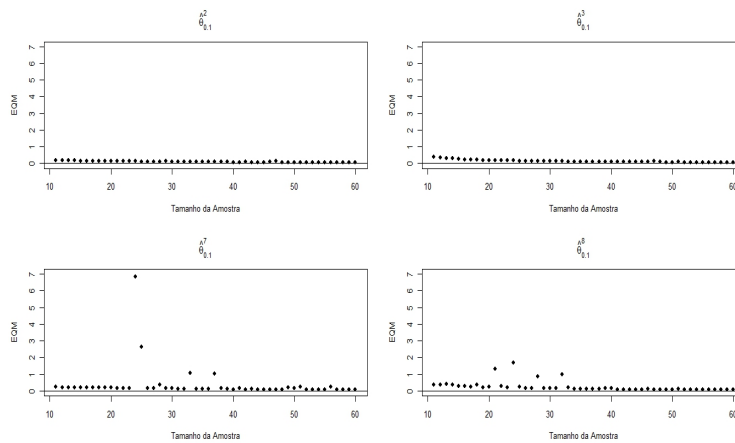


Figura 65 EQM para cada tamanho amostral e os estimadores $\hat{\theta}_{\Gamma}^2$, $\hat{\theta}_{\Gamma}^3$, $\hat{\theta}_{\Gamma}^7$ e $\hat{\theta}_{\Gamma}^8$, amostras geradas considerando uma **Distribuição Logística** com $\mu = 2,25$ e $\sigma^2 = 1$ para $\Gamma = 0,10$ e o **DNM**

Dentre os quatro estimadores, $\hat{\theta}_{\Gamma}^3$ foi aquele que apresentou ligeiramente melhores resultados, tanto quanto ao EQM, quanto ao viés (Figuras 74, 75, 76, 77, 78, 79, 80, 81 e 82.). Mesmo para o cenário em que a dose alvo se encontrava fora da amplitude das doses, este estimador apresentou bons resultados (Figuras 68 e 77).

Os estimadores $\hat{\theta}_{\Gamma}^3$ e $\hat{\theta}_{\Gamma}^8$ no processo de simulação, que foram calculados pela rotina *glm* do software R, foram mais estáveis que os outros dois estimadores. Em geral, para amostras de tamanho maior ou igual 10 o método pode ser aplicado

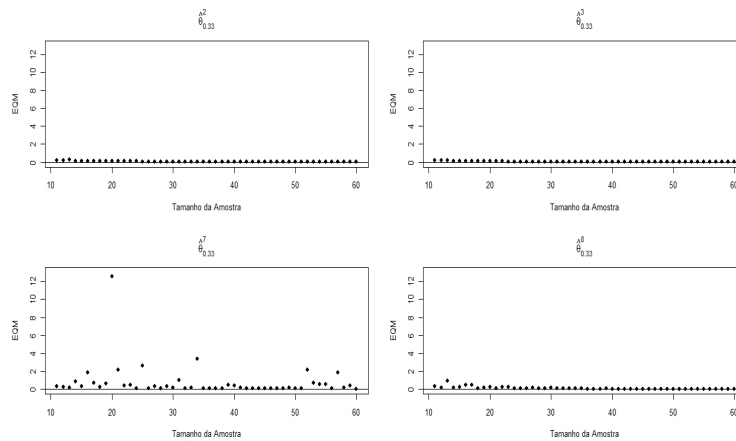


Figura 66 EQM para cada tamanho amostral e os estimadores $\hat{\theta}_{\Gamma}^2$, $\hat{\theta}_{\Gamma}^3$, $\hat{\theta}_{\Gamma}^7$ e $\hat{\theta}_{\Gamma}^8$, amostras geradas considerando uma **Distribuição Logística** com $\mu = 2,25$ e $\sigma^2 = 1$ para $\Gamma = 0,33$ e o **DNM**

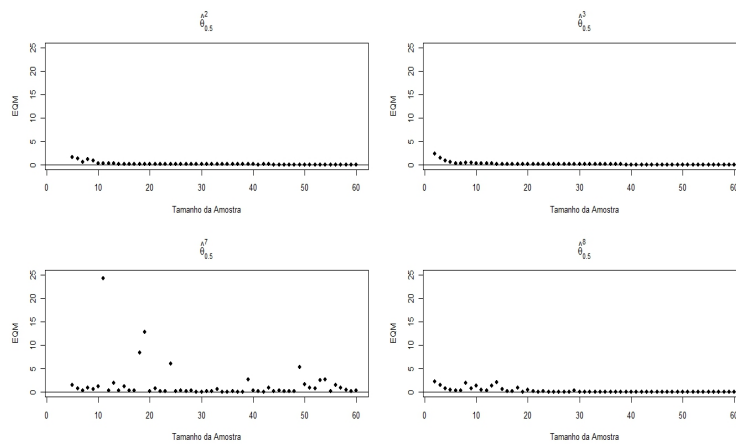


Figura 67 EQM para cada tamanho amostral e os estimadores $\hat{\theta}_{\Gamma}^2$, $\hat{\theta}_{\Gamma}^3$, $\hat{\theta}_{\Gamma}^7$ e $\hat{\theta}_{\Gamma}^8$, amostras geradas considerando uma **Distribuição Logística** com $\mu = 2,25$ e $\sigma^2 = 1$ para $\Gamma = 0,50$ e o **DNM**

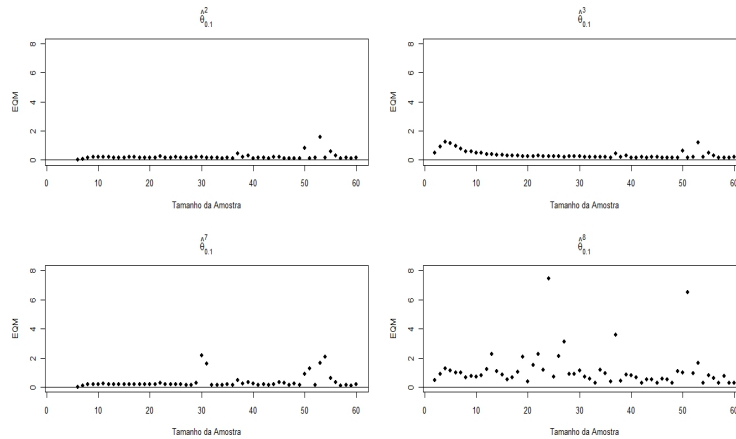


Figura 68 EQM para cada tamanho amostral e os estimadores $\hat{\theta}_{\Gamma}^2$, $\hat{\theta}_{\Gamma}^3$, $\hat{\theta}_{\Gamma}^7$ e $\hat{\theta}_{\Gamma}^8$, amostras geradas considerando uma **Distribuição Logística** com $\mu = 1,25$ e $\sigma^2 = 1$ para $\Gamma = 0,10$ e o **DNM**

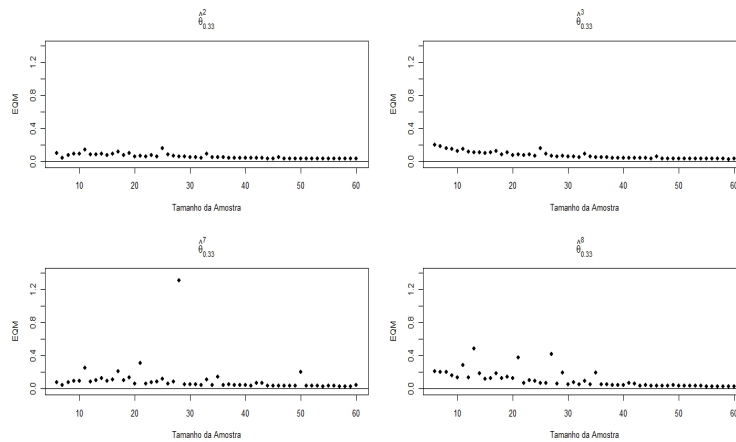


Figura 69 EQM para cada tamanho amostral e os estimadores $\hat{\theta}_{\Gamma}^2$, $\hat{\theta}_{\Gamma}^3$, $\hat{\theta}_{\Gamma}^7$ e $\hat{\theta}_{\Gamma}^8$, amostras geradas considerando uma **Distribuição Logística** com $\mu = 1,25$ e $\sigma^2 = 1$ para $\Gamma = 0,33$ e o **DNM**

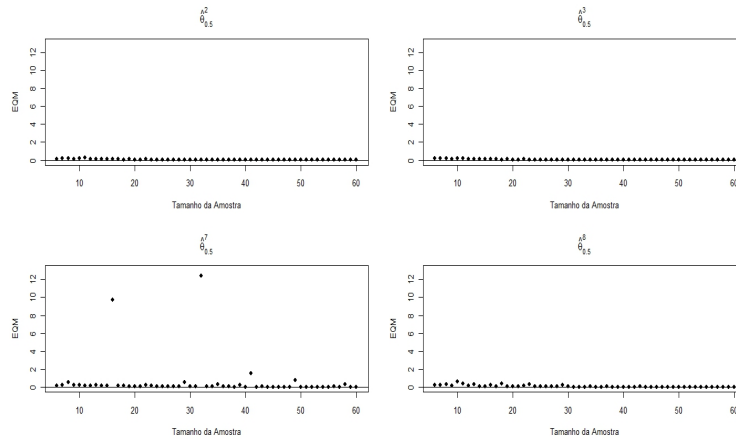


Figura 70 EQM para cada tamanho amostral e os estimadores $\hat{\theta}_{\Gamma}^2$, $\hat{\theta}_{\Gamma}^3$, $\hat{\theta}_{\Gamma}^7$ e $\hat{\theta}_{\Gamma}^8$, amostras geradas considerando uma **Distribuição Logística** com $\mu = 1,25$ e $\sigma^2 = 1$ para $\Gamma = 0,50$ e o **DNM**

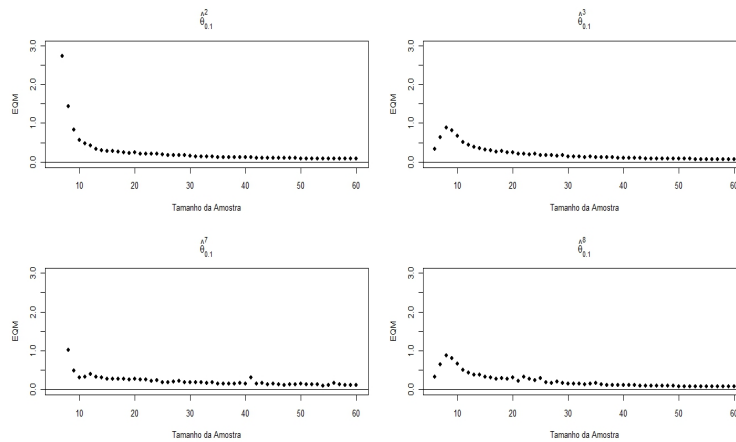


Figura 71 EQM para cada tamanho amostral e os estimadores $\hat{\theta}_{\Gamma}^2$, $\hat{\theta}_{\Gamma}^3$, $\hat{\theta}_{\Gamma}^7$ e $\hat{\theta}_{\Gamma}^8$, amostras geradas considerando uma **Distribuição Logística** com $\mu = 3,25$ e $\sigma^2 = 1$ para $\Gamma = 0,10$ e o **DNM**

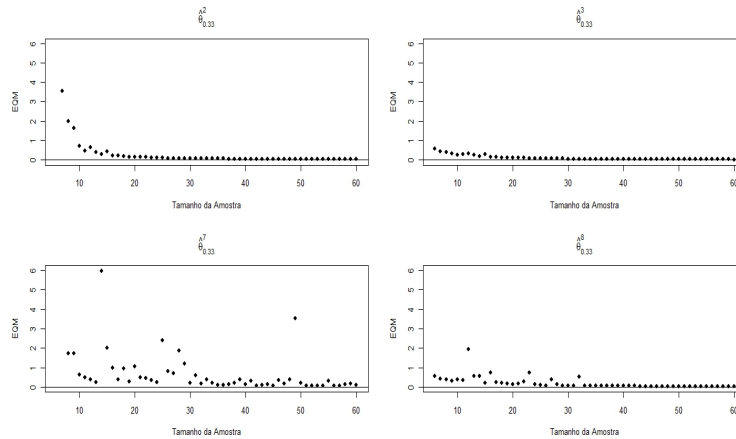


Figura 72 EQM para cada tamanho amostral e os estimadores $\hat{\theta}_{\Gamma}^2$, $\hat{\theta}_{\Gamma}^3$, $\hat{\theta}_{\Gamma}^7$ e $\hat{\theta}_{\Gamma}^8$, amostras geradas considerando uma **Distribuição Logística** com $\mu = 3,25$ e $\sigma^2 = 1$ para $\Gamma = 0,33$ e o **DNM**

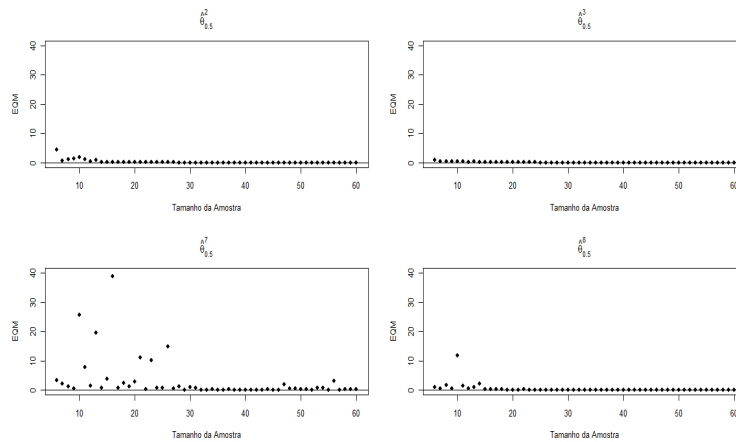


Figura 73 EQM para cada tamanho amostral e os estimadores $\hat{\theta}_{\Gamma}^2$, $\hat{\theta}_{\Gamma}^3$, $\hat{\theta}_{\Gamma}^7$ e $\hat{\theta}_{\Gamma}^8$, amostras geradas considerando uma **Distribuição Logística** com $\mu = 3,25$ e $\sigma^2 = 1$ para $\Gamma = 0,50$ e o **DNM**

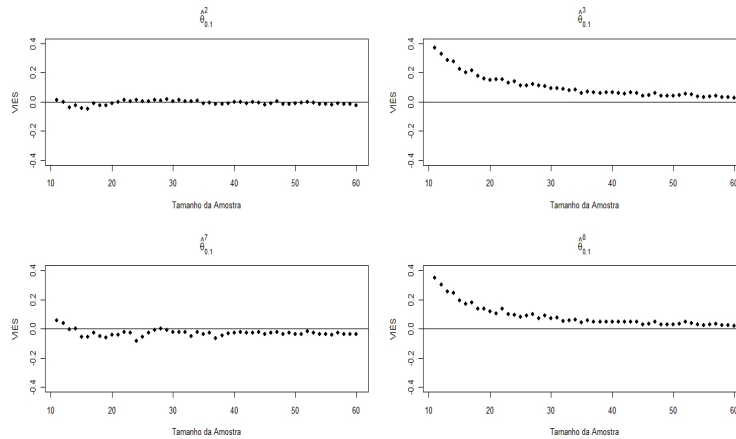


Figura 74 Viés para cada tamanho amostral e os estimadores $\hat{\theta}_{\Gamma}^2$, $\hat{\theta}_{\Gamma}^3$, $\hat{\theta}_{\Gamma}^7$ e $\hat{\theta}_{\Gamma}^8$, amostras geradas considerando uma **Distribuição Logística** com $\mu = 2,25$ e $\sigma^2 = 1$ para $\Gamma = 0,10$ e o **DNM**

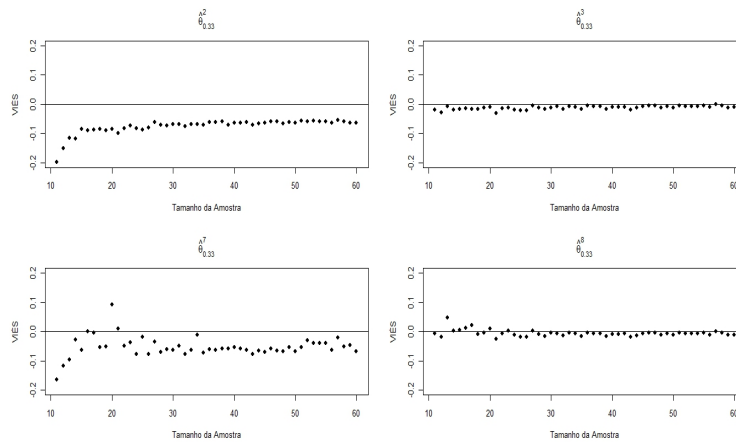


Figura 75 Viés para cada tamanho amostral e os estimadores $\hat{\theta}_{\Gamma}^2$, $\hat{\theta}_{\Gamma}^3$, $\hat{\theta}_{\Gamma}^7$ e $\hat{\theta}_{\Gamma}^8$, amostras geradas considerando uma **Distribuição Logística** com $\mu = 2,25$ e $\sigma^2 = 1$ para $\Gamma = 0,33$ e o **DNM**

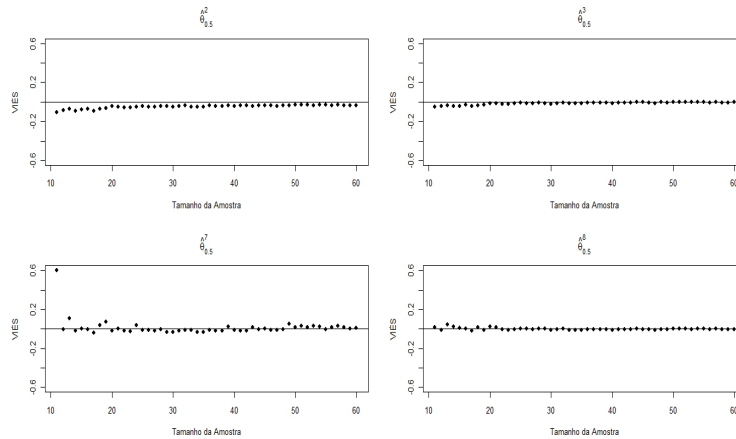


Figura 76 Viés para cada tamanho amostral e os estimadores $\hat{\theta}_{\Gamma}^2$, $\hat{\theta}_{\Gamma}^3$, $\hat{\theta}_{\Gamma}^7$ e $\hat{\theta}_{\Gamma}^8$, amostras geradas considerando uma **Distribuição Logística** com $\mu = 2,25$ e $\sigma^2 = 1$ para $\Gamma = 0,50$ e o **DNM**

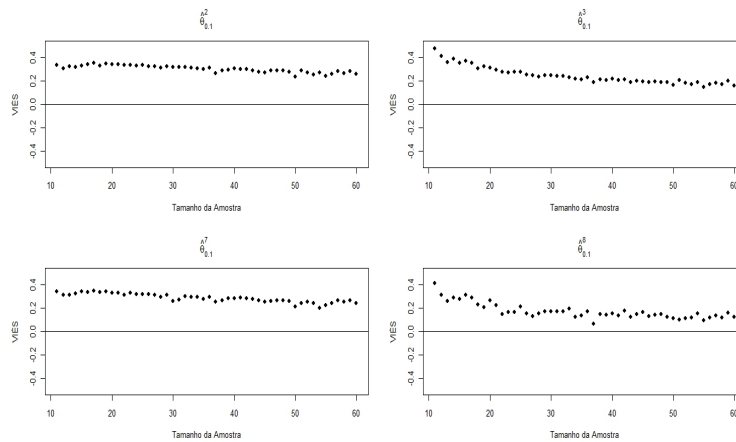


Figura 77 Viés para cada tamanho amostral e os estimadores $\hat{\theta}_{\Gamma}^2$, $\hat{\theta}_{\Gamma}^3$, $\hat{\theta}_{\Gamma}^7$ e $\hat{\theta}_{\Gamma}^8$, amostras geradas considerando uma **Distribuição Logística** com $\mu = 1,25$ e $\sigma^2 = 1$ para $\Gamma = 0,10$ e o **DNM**

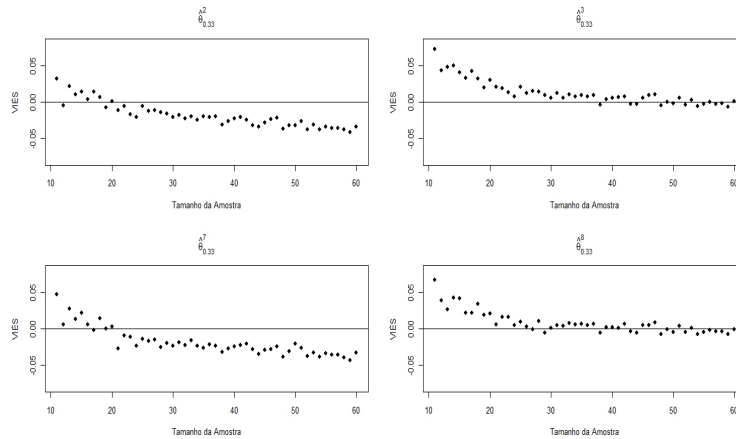


Figura 78 Viés para cada tamanho amostral e os estimadores $\hat{\theta}_{\Gamma}^2$, $\hat{\theta}_{\Gamma}^3$, $\hat{\theta}_{\Gamma}^7$ e $\hat{\theta}_{\Gamma}^8$, amostras geradas considerando uma **Distribuição Logística** com $\mu = 1,25$ e $\sigma^2 = 1$ para $\Gamma = 0,33$ e o **DNM**

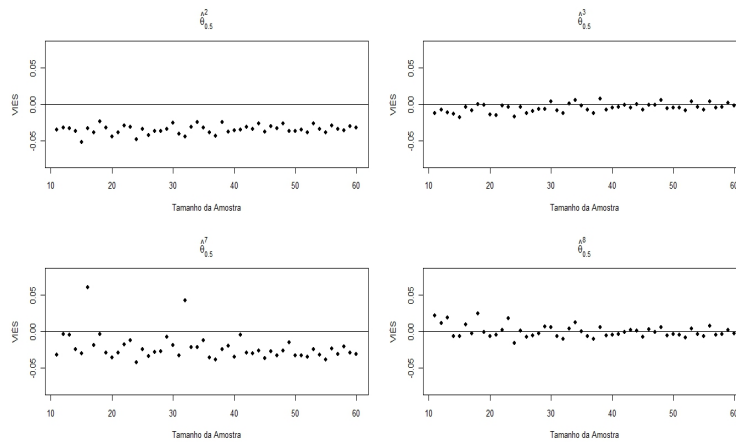


Figura 79 Viés para cada tamanho amostral e os estimadores $\hat{\theta}_{\Gamma}^2$, $\hat{\theta}_{\Gamma}^3$, $\hat{\theta}_{\Gamma}^7$ e $\hat{\theta}_{\Gamma}^8$, amostras geradas considerando uma **Distribuição Logística** com $\mu = 1,25$ e $\sigma^2 = 1$ para $\Gamma = 0,50$ e o **DNM**

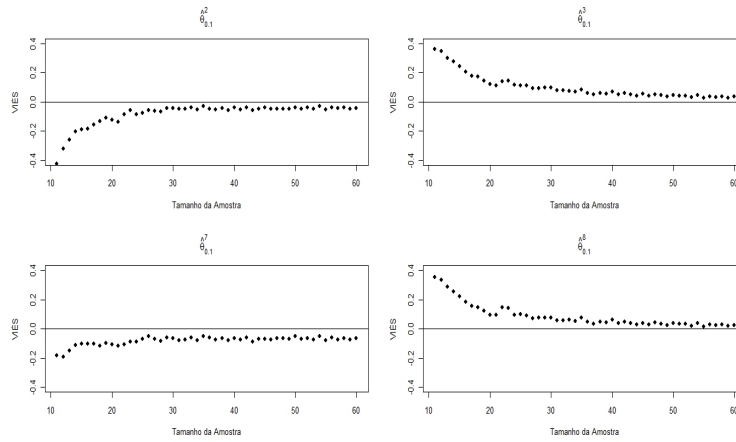


Figura 80 Viés para cada tamanho amostral e os estimadores $\hat{\theta}_{\Gamma}^2$, $\hat{\theta}_{\Gamma}^3$, $\hat{\theta}_{\Gamma}^7$ e $\hat{\theta}_{\Gamma}^8$, amostras geradas considerando uma **Distribuição Logística** com $\mu = 3,25$ e $\sigma^2 = 1$ para $\Gamma = 0,10$ e o **DNM**

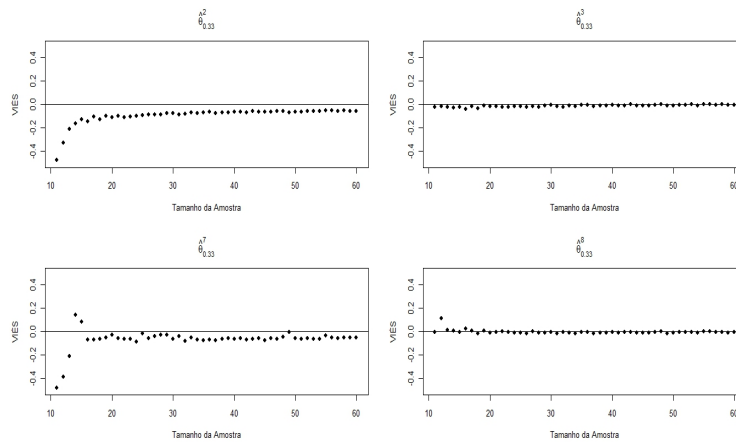


Figura 81 Viés para cada tamanho amostral e os estimadores $\hat{\theta}_{\Gamma}^2$, $\hat{\theta}_{\Gamma}^3$, $\hat{\theta}_{\Gamma}^7$ e $\hat{\theta}_{\Gamma}^8$, amostras geradas considerando uma **Distribuição Logística** com $\mu = 3,25$ e $\sigma^2 = 1$ para $\Gamma = 0,33$ e o **DNM**

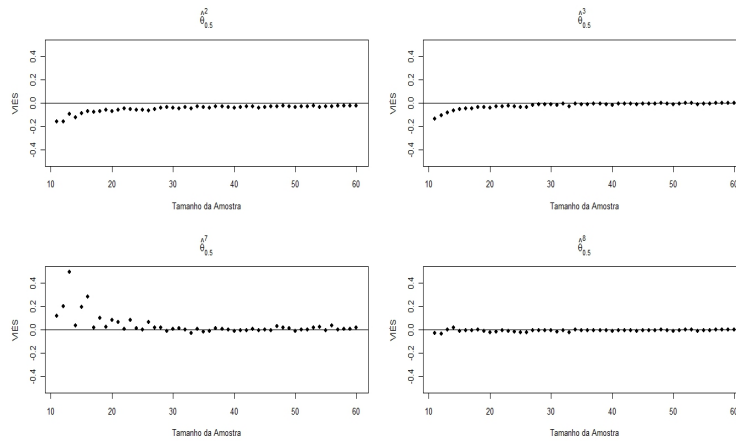


Figura 82 Viés para cada tamanho amostral e os estimadores $\hat{\theta}_{\Gamma}^2$, $\hat{\theta}_{\Gamma}^3$, $\hat{\theta}_{\Gamma}^7$ e $\hat{\theta}_{\Gamma}^8$, amostras geradas considerando uma **Distribuição Logística** com $\mu = 3,25$ e $\sigma^2 = 1$ para $\Gamma = 0,50$ e o **DNM**

praticamente em 100% das vezes. No entanto, para os outros dois estimadores o tamanho amostral deve ser maior do que 20 para que a aplicação do método funcione aproximadamente em 80% das vezes.

Stylianou e Follmann (2004) comentam que estimadores como $\hat{\theta}_{\Gamma}^3$ e $\hat{\theta}_{\Gamma}^8$ poderiam apresentar problemas para amostras reduzidas, devido à instabilidade dos estimadores de máxima verossimilhança para amostras pequenas, e portanto sugeriram o uso dos estimadores de mínimos quadrados para tais situações. Neste trabalho, para amostras relativamente pequenas, com tamanho maior do que 10, o desempenho dos estimadores de máxima verossimilhança é superior aos de mínimos quadrados.

4.1.3 Estimadores não paramétricos

Foram estudados cinco estimadores não paramétricos, dois propostos, $\hat{\theta}_{\Gamma}^4$ e $\hat{\theta}_{\Gamma}^5$ e três apresentados na literatura, $\hat{\theta}_{\Gamma}^9$, $\hat{\theta}_{\Gamma}^{10}$ e $\hat{\theta}_{\Gamma}^{11}$. Estes estimadores foram

estudados para o DNM em 27 cenários, sob o enfoque do EQM e viés para cada tamanho amostral. As conclusões são similares para quase todos os cenários, assim, serão expostos alguns casos, os demais encontram-se nos APÊNDICES B, D, E, F, G, H, I, J, L, M, O, Q, R, S e T.

O estimador $\hat{\theta}_T^4$ pode ser aplicado em qualquer situação, sendo que o mesmo escolhe uma das doses estipuladas pelo pesquisador como a estimativa da dose alvo. Os estimadores $\hat{\theta}_T^5$ e $\hat{\theta}_T^9$ puderam ser aplicados em quase 100% dos casos para amostras maiores ou iguais a 10. Os estimadores $\hat{\theta}_T^{10}$ e $\hat{\theta}_T^{11}$ apresentaram problemas na simulação para os casos em que a dose alvo era menor do que a menor dose testada, sendo que, quanto maior a amostra, menos o estimador podia ser aplicado. Uma justificativa para este mal comportamento é que para amostras grandes todos os delineamentos têm uma tendência em concentrar as respostas em única dose, no caso a menor delas.

Com relação ao EQM todos os estimadores apresentaram resultados bem similares quando comparados dentro dos mesmos cenários, como pode ser observado, por exemplo na Figura 83.

Para um tamanho amostral a partir de 20 o EQM é praticamente zero para qualquer um dos estimadores, resultados semelhantes aos apresentados por Stylianou, Proscha e Flournoy (2003) para os estimadores $\hat{\theta}_T^{10}$ e $\hat{\theta}_T^{11}$.

O estimador $\hat{\theta}_T^{10}$ também sempre toma uma das doses testadas como uma estimativa. O desempenho de $\hat{\theta}_T^4$ e $\hat{\theta}_T^{10}$ foi semelhante (Figura 84), entretanto, $\hat{\theta}_T^4$ pode ser usado em qualquer situação, e $\hat{\theta}_T^{10}$ apresentou muitos problemas quanto a sua aplicabilidade quando a dose alvo estava fora da amplitude das doses. Neste sentido, pode-se afirmar que o estimador proposto apresenta um desempenho superior ao estimador $\hat{\theta}_T^{10}$ utilizado na literatura.

Dentre os estimadores $\hat{\theta}_T^9$, $\hat{\theta}_T^{10}$ e $\hat{\theta}_T^{11}$ já apresentados na literatura, aquele

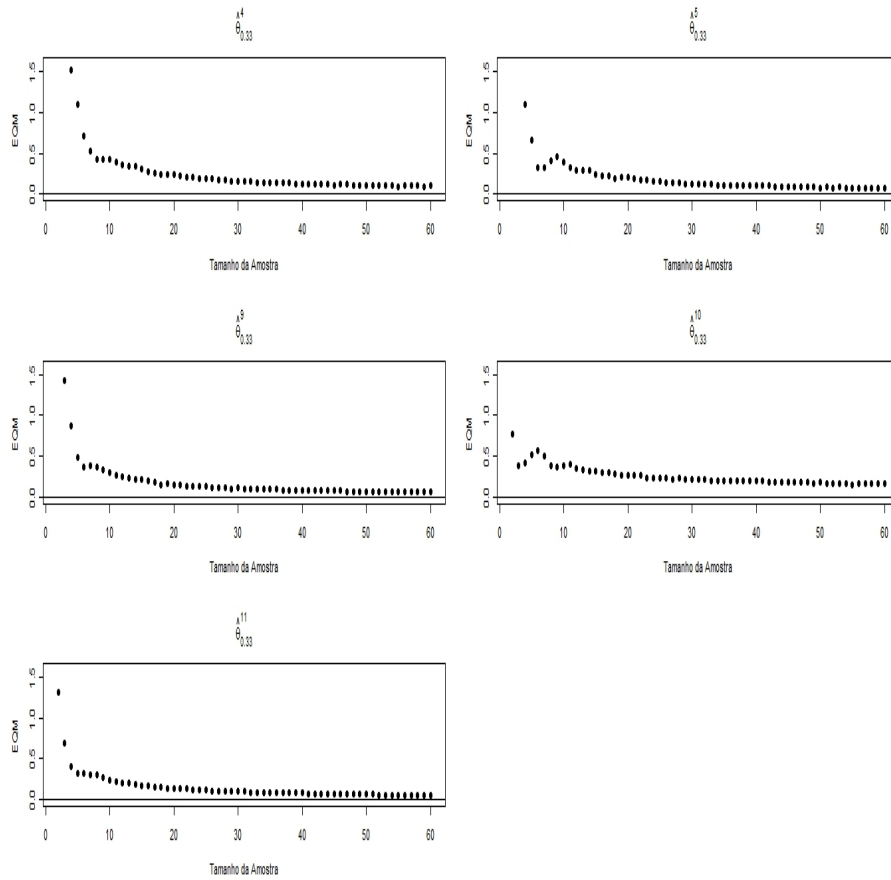


Figura 83 EQM para cada tamanho amostral e os estimadores $\hat{\theta}_{\Gamma}^4$, $\hat{\theta}_{\Gamma}^5$, $\hat{\theta}_{\Gamma}^9$, $\hat{\theta}_{\Gamma}^{10}$ e $\hat{\theta}_{\Gamma}^{11}$, amostras geradas considerando uma **Distribuição Normal** com $\mu = 2,25$ e $\sigma^2 = 1$ para $\Gamma = 0,33$ e o **DNM**

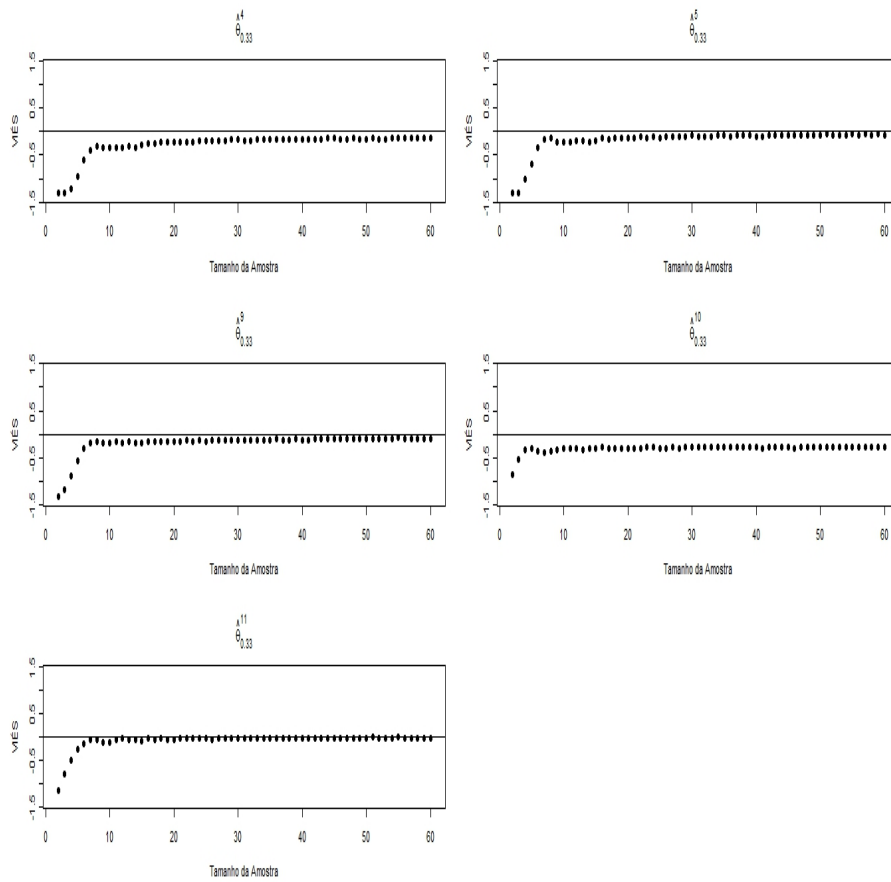


Figura 84 Viés para cada tamanho amostral e os estimadores $\hat{\theta}_{\Gamma}^4$, $\hat{\theta}_{\Gamma}^5$, $\hat{\theta}_{\Gamma}^9$, $\hat{\theta}_{\Gamma}^{10}$ e $\hat{\theta}_{\Gamma}^{11}$, amostras geradas considerando uma **Distribuição Normal** com $\mu = 2,25$ e $\sigma^2 = 1$ para $\Gamma = 0,33$ e o **DNM**

que apresentou melhor desempenho foi $\hat{\theta}_\Gamma^9$. Este estimador se mostrou muito bom também quando comparado aos estimadores propostos $\hat{\theta}_\Gamma^4$ e $\hat{\theta}_\Gamma^5$, obtendo na grande maioria dos casos um desempenho ligeiramente melhor com relação ao EQM. Aparentemente, os estimadores $\hat{\theta}_\Gamma^4$ e $\hat{\theta}_\Gamma^5$ o superaram ligeiramente somente no caso em que as probabilidades foram modeladas por distribuições não simétricas e a dose alvo estava fora da amplitude das doses, como nas Figuras 85, 86, 87, 88, 89 e 90.

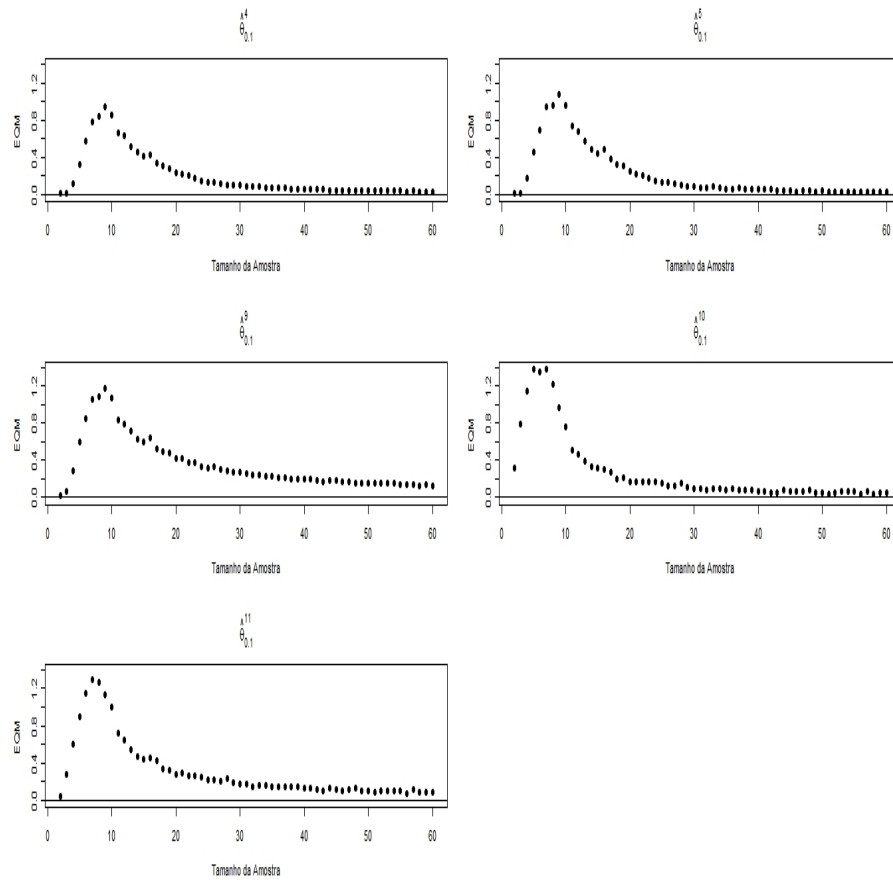


Figura 85 EQM para cada tamanho amostral e os estimadores $\hat{\theta}_{\Gamma}^4$, $\hat{\theta}_{\Gamma}^5$, $\hat{\theta}_{\Gamma}^9$, $\hat{\theta}_{\Gamma}^{10}$ e $\hat{\theta}_{\Gamma}^{11}$, amostras geradas considerando uma **Distribuição Exponencial** com parâmetro $\lambda = 0,30$, para $\Gamma = 0,10$ e o **DNM**

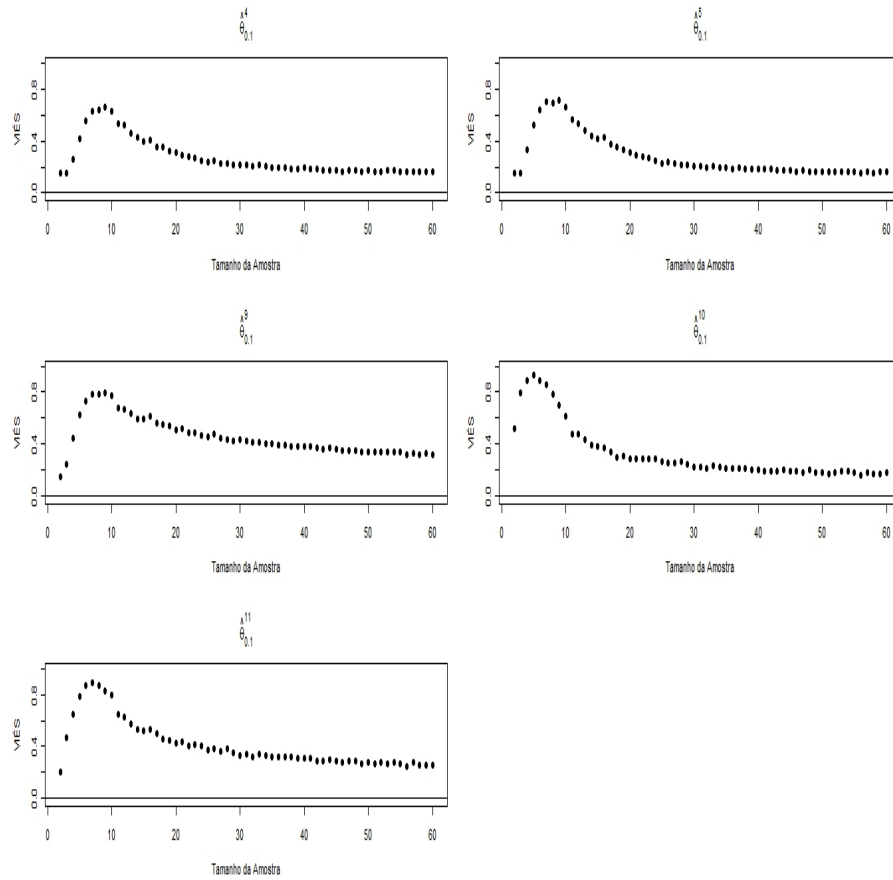


Figura 86 Viés para cada tamanho amostral e os estimadores $\hat{\theta}_{\Gamma}^4$, $\hat{\theta}_{\Gamma}^5$, $\hat{\theta}_{\Gamma}^9$, $\hat{\theta}_{\Gamma}^{10}$ e $\hat{\theta}_{\Gamma}^{11}$, amostras geradas considerando uma **Distribuição Exponencial** com parâmetro $\lambda = 0,30$, para $\Gamma = 0,10$ e o **DNM**

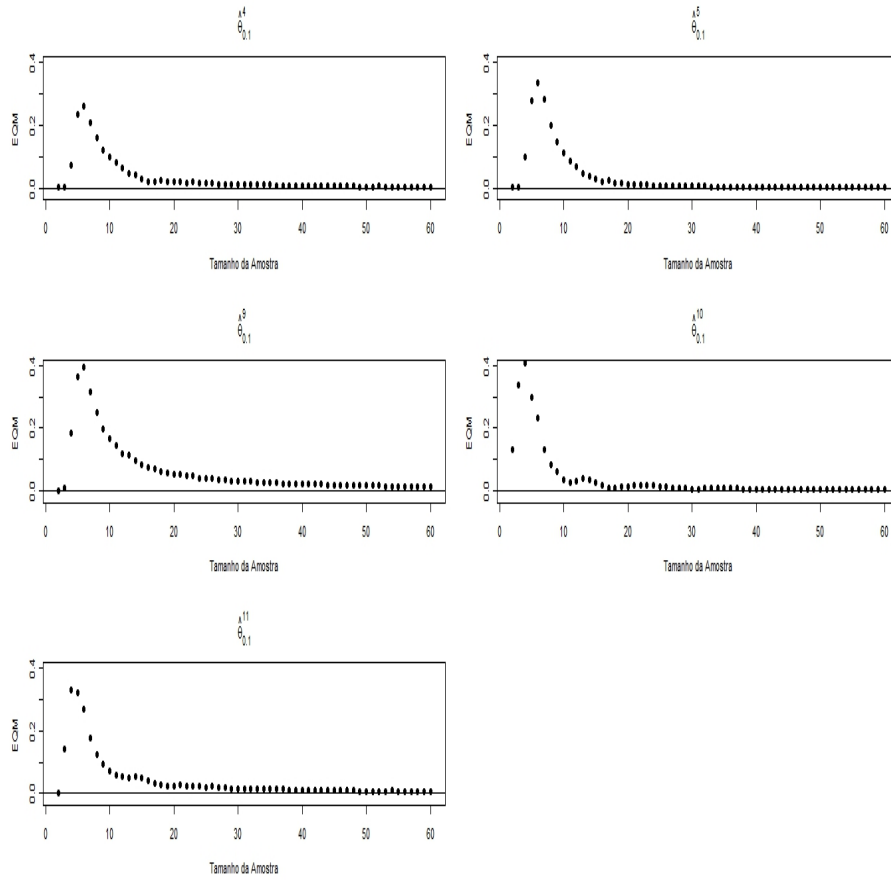


Figura 87 EQM para cada tamanho amostral e os estimadores $\hat{\theta}_{\Gamma}^4$, $\hat{\theta}_{\Gamma}^5$, $\hat{\theta}_{\Gamma}^9$, $\hat{\theta}_{\Gamma}^{10}$ e $\hat{\theta}_{\Gamma}^{11}$, amostras geradas considerando uma **Distribuição Gama** com parâmetros $\alpha = 3$ e $\lambda = 0,50$, para $\Gamma = 0,10$ e o **DNM**

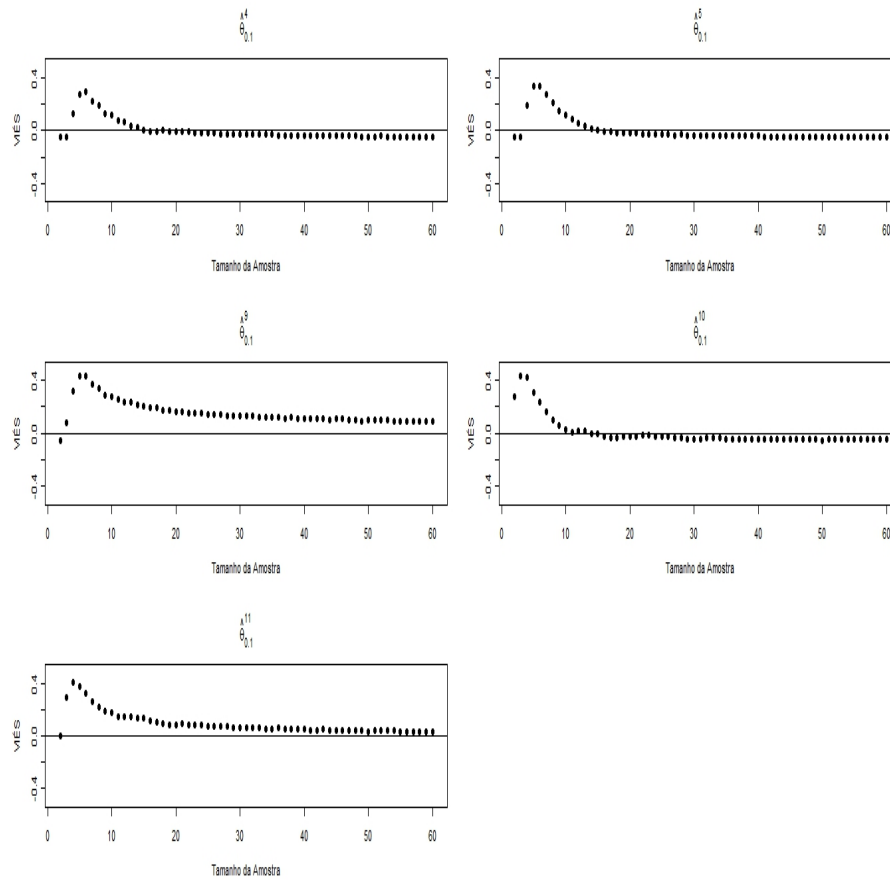


Figura 88 Viés para cada tamanho amostral e os estimadores $\hat{\theta}_{\Gamma}^4$, $\hat{\theta}_{\Gamma}^5$, $\hat{\theta}_{\Gamma}^9$, $\hat{\theta}_{\Gamma}^{10}$ e $\hat{\theta}_{\Gamma}^{11}$, amostras geradas considerando uma **Distribuição Gama** com parâmetros $\alpha = 3$ e $\lambda = 0,50$, para $\Gamma = 0,10$ e o **DNM**

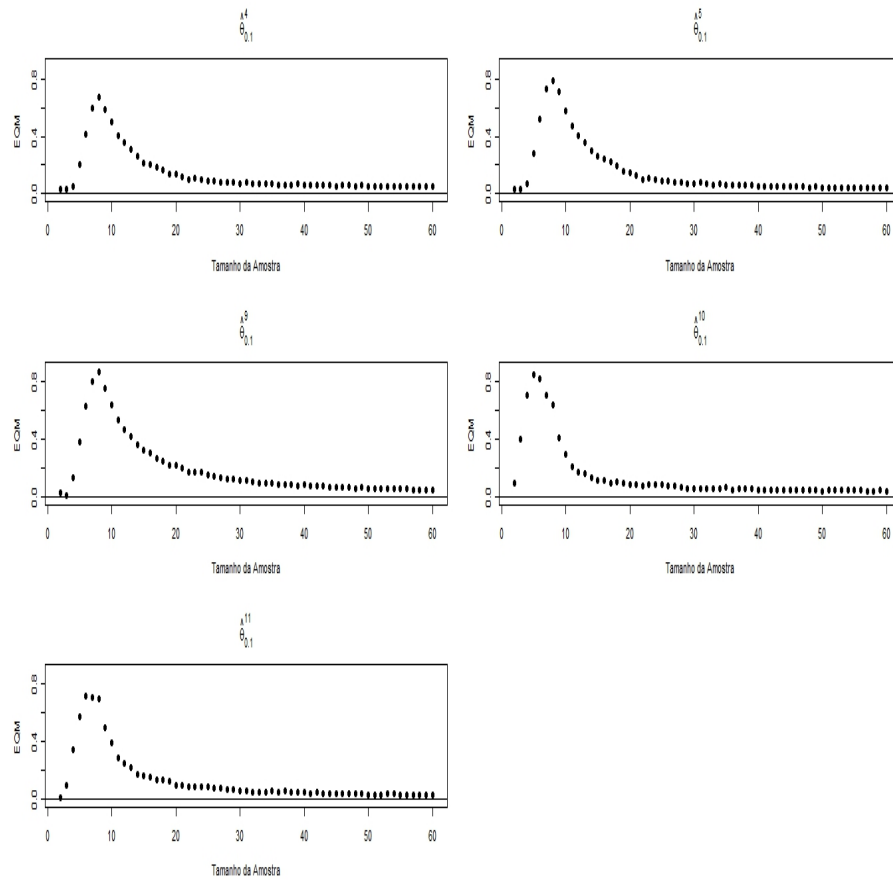


Figura 89 EQM para cada tamanho amostral e os estimadores $\hat{\theta}_{\Gamma}^4$, $\hat{\theta}_{\Gamma}^5$, $\hat{\theta}_{\Gamma}^9$, $\hat{\theta}_{\Gamma}^{10}$ e $\hat{\theta}_{\Gamma}^{11}$, amostras geradas considerando uma **Distribuição Gama** com parâmetros $\alpha = 2$ e $\lambda = 1,25$, para $\Gamma = 0,10$ e o **DNM**

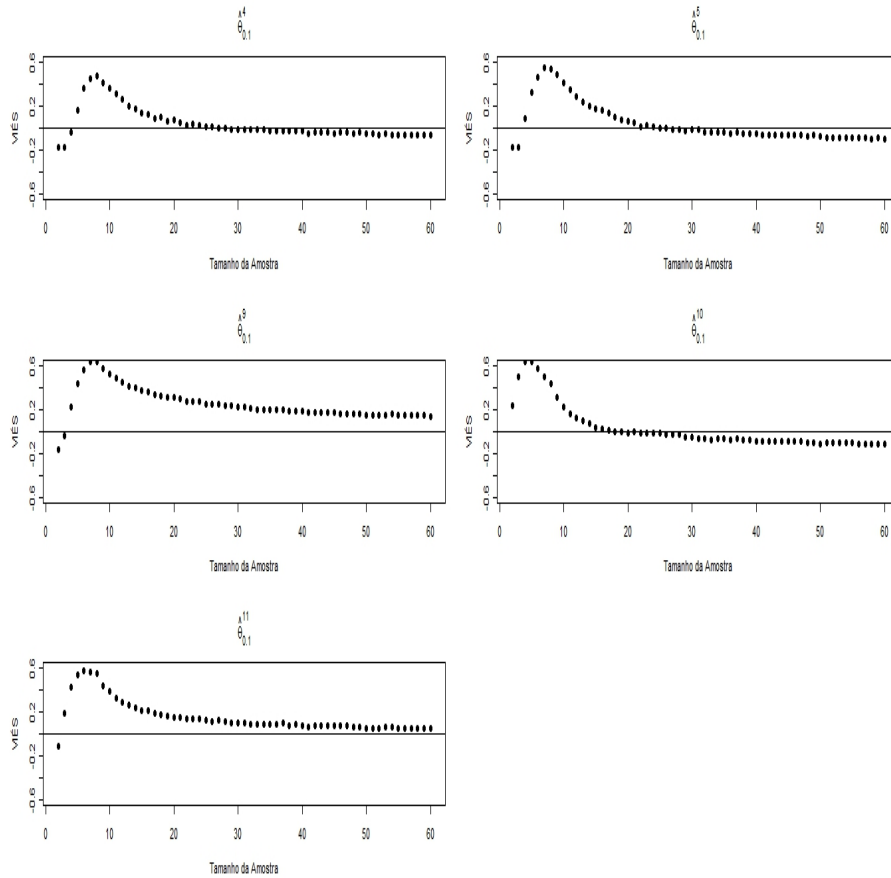


Figura 90 Viés para cada tamanho amostral e os estimadores $\hat{\theta}_{\Gamma}^4$, $\hat{\theta}_{\Gamma}^5$, $\hat{\theta}_{\Gamma}^9$, $\hat{\theta}_{\Gamma}^{10}$ e $\hat{\theta}_{\Gamma}^{11}$, amostras geradas considerando uma **Distribuição Gama** com parâmetros $\alpha = 2$ e $\lambda = 1,25$, para $\Gamma = 0,10$ e o **DNM**

4.1.4 Delineamento bayesiano isotônico sequencial - DBIS

O delineamento proposto DBIS foi avaliado para dois conjuntos de doses, com amplitudes iguais e dois níveis de espaçamento diferentes. Todos os cenários foram avaliados considerando um mesmo estimador, calculando-se seu EQM e viés. Os resultados para os vários cenários foram semelhantes e por simplicidade serão apresentados e discutidos apenas alguns casos específicos, sendo que os demais estão apresentados nos APÊNDICES U, V, X e Z.

4.1.4.1 Primeiro conjunto das doses

O estudo de simulação para a avaliação do DBIS para as doses em $D_1 = \{0,5; 1,0; 1,5; 2,0; 2,5; 3,0; 3,5; 4,0\}$ e os diversos cenários expostos na Tabela 10 apresentaram resultados semelhantes, e serão mostrados ou discutidos os resultados de apenas um caso em que as verdadeiras probabilidades de transição entre as doses foram obtidas de uma Distribuição Normal com média $\mu = 2,25$ e variância $\sigma^2 = 1,0$. Os demais resultados gráficos para as distribuições Exponencial e Gama encontram-se nos APÊNDICES U e V respectivamente.

Para a estimação dos três quantis considerados, $\theta_{0,10}$, $\theta_{0,33}$ e $\theta_{0,50}$ tanto o EQM quanto o viés tendem a se aproximar de zero mais rapidamente pelo DBIS do que pelo DNM (Figuras 91, 92 e 93).

Para o quantil mais extremo ($\theta_{0,10}$) percebe-se uma estabilização do EQM e do viés em torno de zero para tamanhos de amostra em torno de 20 para o DBIS, enquanto que resultado semelhante ocorre pelo DNM para um tamanho amostral em torno de 30 (Figura 91). Para os outros dois quantis ($\theta_{0,33}$ e $\theta_{0,50}$) a estabilização ocorreu mais rápido para ambos os delineamentos, entretanto, o DBIS apresenta melhores resultados para tamanho amostral em torno de 10 (Figura 92 e 93).

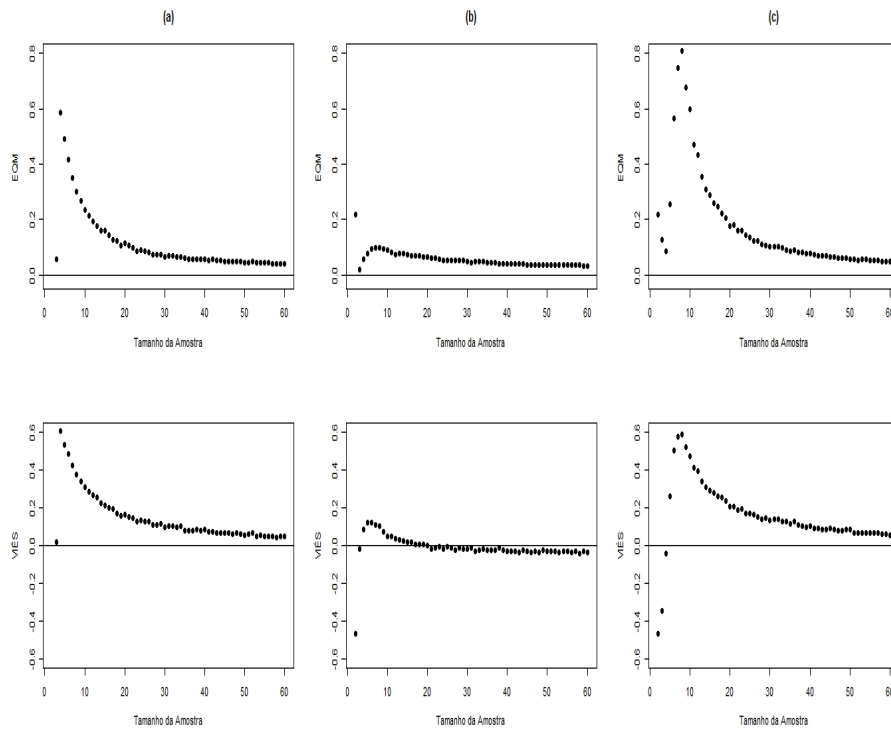


Figura 91 EQM e Viés para cada tamanho amostral e o estimador $\hat{\theta}_{\Gamma}^g$, amostras geradas considerando uma **Distribuição Normal** com parâmetros $\mu = 2,25$ e $\sigma^2 = 1$, para $\Gamma = 0,10$. Cenários (a), (b) e (c) apresentados na Tabela 10, os cenários (a) e (b) referentes ao **DBIS** com priors diferentes e o cenário (c) referente ao **DNM**

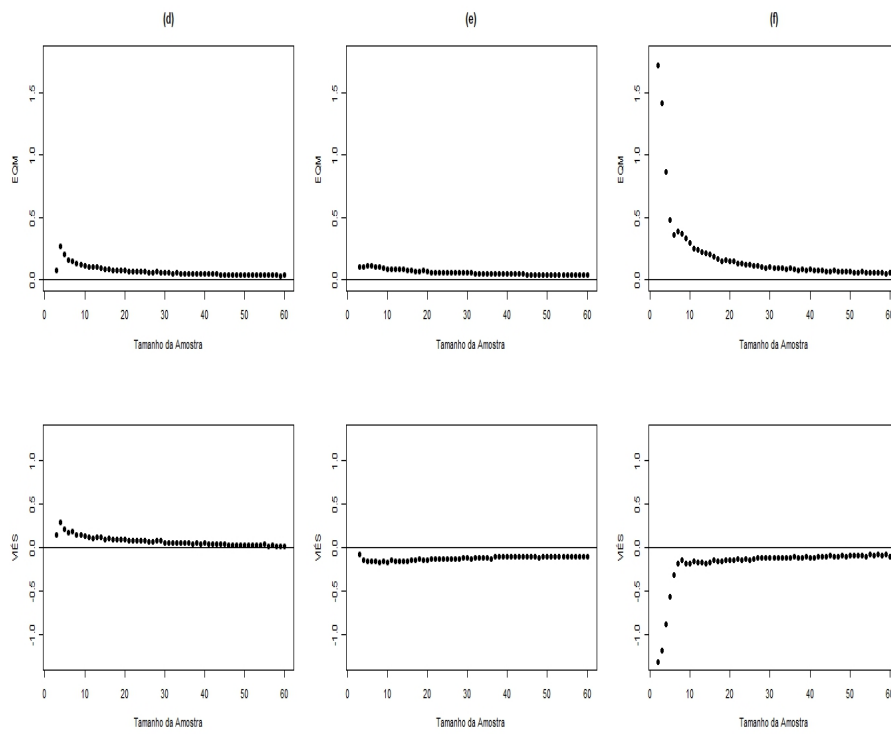


Figura 92 EQM e Viés para cada tamanho amostral e o estimador $\hat{\theta}_{\Gamma}^g$, amostras geradas considerando uma **Distribuição Normal** com parâmetros $\mu = 2,25$ e $\sigma^2 = 1$, para $\Gamma = 0,33$. Cenários (d), (e) e (f) apresentados na Tabela 10, os cenários (d) e (e) referentes ao **DBIS** com priors diferentes e o cenário (f) referente ao **DNM**

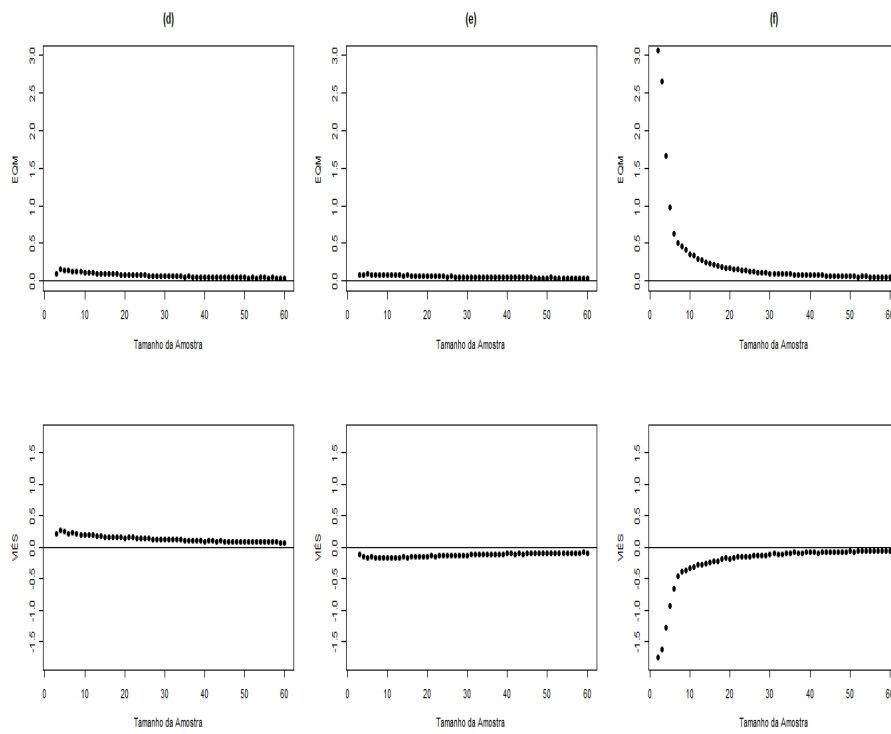


Figura 93 EQM e Viés para cada tamanho amostral e o estimador $\hat{\theta}_{\Gamma}^9$, amostras geradas considerando uma **Distribuição Normal** com parâmetros $\mu = 2,25$ e $\sigma^2 = 1$, para $\Gamma = 0,50$. Cenários (d), (e) e (f) apresentados na Tabela 10, os cenários (d) e (e) referentes ao **DBIS** com priors diferentes e o cenário (f) referente ao **DNM**

Quando o verdadeiro valor das probabilidades de transição entre as doses foram obtidos de uma distribuição Exponencial, o DBIS e o DNM apresentaram resultados semelhantes àqueles já mencionados para o caso Normal (APÊNDICE U). Uma ligeira vantagem foi apresentada pelo DNM em relação ao DBIS para o caso Gama (APÊNDICE V). Em relação às prioris utilizadas o desempenho do DBIS mostrou-se robusto, fornecendo melhores estimativas do que pelo DNM.

4.1.4.2 Segundo conjunto das doses

O estudo de simulação para a avaliação do DBIS para as doses em $D_2 = \{0,5; 0,75; 1,0; 1,25; 1,5; 1,75; 2,0; 2,25; 2,5; 2,75; 3,0; 3,25; 3,5; 3,75; 4,0\}$, e os diversos cenários expostos na Tabela 11 apresentaram resultados semelhantes, e serão apresentados os resultados de apenas um caso em que as verdadeira probabilidades de transição entre as doses, foram obtidos de uma Distribuição Normal com média $\mu = 2,25$ e variância $\sigma^2 = 1,0$. Os demais resultados gráficos para as distribuições Exponencial e Gama encontram-se nos APÊNDICES X e Z respectivamente.

Os resultados mostram que, diminuindo o espaçamento das doses, o desempenho do DBIS em relação ao DNM melhora. Observa-se que tanto para o caso Normal (Figuras 94, 95 e 96), como para os outros casos (APÊNDICES X e Z), o DBIS estabilizou mais rapidamente o EQM e o viés do que o DNM. O desempenho do DNM não foi melhor para nenhum dos cenários para este conjunto de doses.

Com o menor espaçamento entre as doses, o desempenho do DBIS se mostrou ainda mais robusto em relação a priori escolhida.

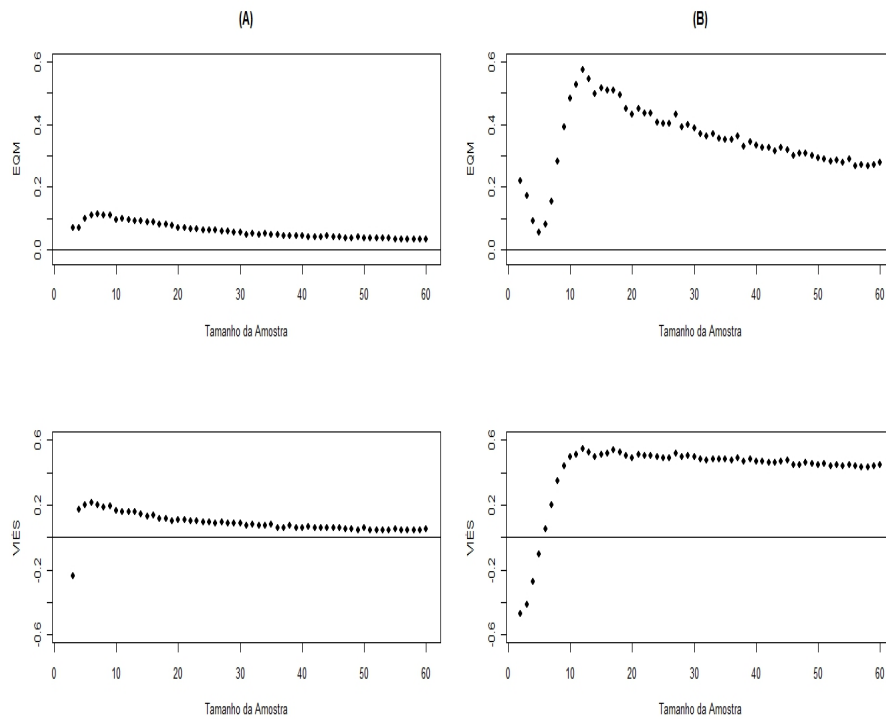


Figura 94 EQM e Viés para cada tamanho amostral e o estimador $\hat{\theta}_{\Gamma}^g$, amostras geradas considerando uma **Distribuição Normal** com parâmetros $\mu = 2,25$ e $\sigma^2 = 1$, para $\Gamma = 0,10$. Cenários (A) e (B) apresentados na Tabela 11, referentes ao **DBIS** e **DNM** respectivamente

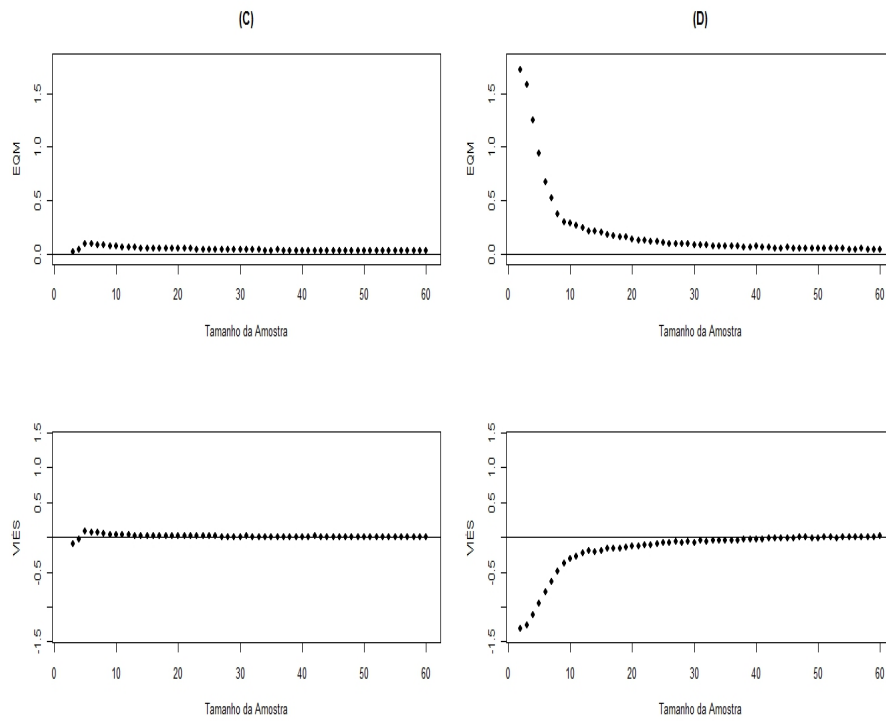


Figura 95 EQM e Viés para cada tamanho amostral e o estimador $\hat{\theta}_{\Gamma}^9$, amostras geradas considerando uma **Distribuição Normal** com parâmetros $\mu = 2,25$ e $\sigma^2 = 1$, para $\Gamma = 0,33$. Cenários (C) e (D) apresentados na Tabela 11, referentes ao **DBIS** e **DNM** respectivamente

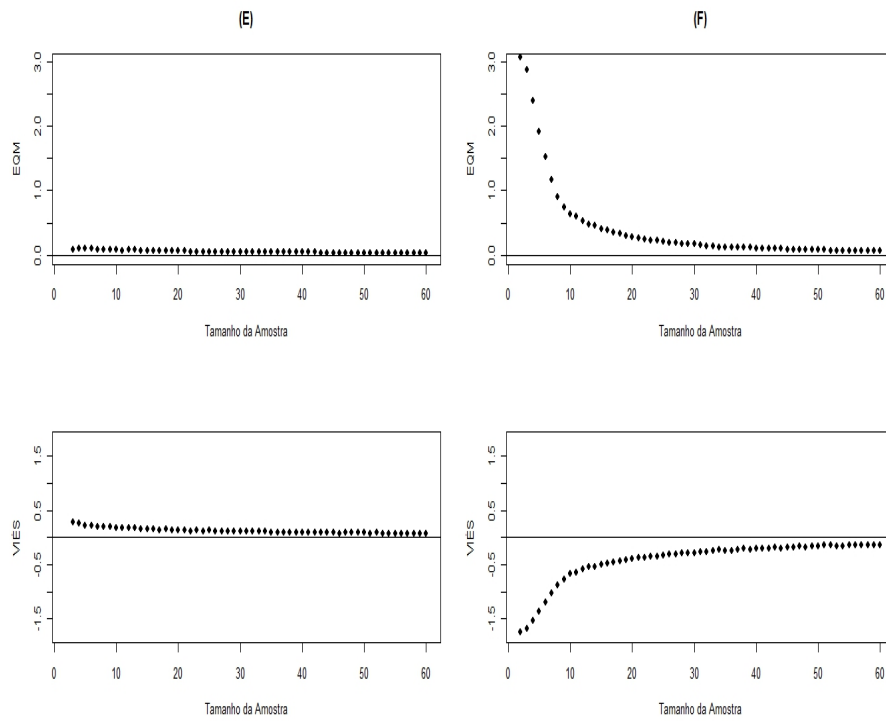


Figura 96 EQM e Viés para cada tamanho amostral e o estimador $\hat{\theta}_{\Gamma}^g$, amostras geradas considerando uma **Distribuição Normal** com parâmetros $\mu = 2,25$ e $\sigma^2 = 1$, para $\Gamma = 0,50$. Cenários (E) e (F) apresentados na Tabela 11, referentes ao **DBIS** e **DNM** respectivamente

5 CONCLUSÃO

- Para os estimadores de curvas de dose-resposta normais, o estimador proposto, θ_{Γ}^1 (*Probit Isotonizado*), apresentou melhores resultados com relação ao EQM e viés do que o estimador usual apresentado na literatura para os três delineamentos, BCB, DNM e DILW;
- Dentre os estimadores de curvas de dose-resposta logística, os estimadores propostos, θ_{Γ}^2 (*Logit Isotonizado* baseado no método dos mínimos quadrados) e o θ_{Γ}^3 (*Logit Isotonizado* baseado no método da máxima verossimilhança), apresentaram melhores resultados com relação ao EQM e viés do que os demais apresentados na literatura para os três delineamentos, BCB, DNM e DILW. Dentre os dois propostos, o estimador de máxima verossimilhança apresentou melhor resultado do que o de mínimos quadrados ponderados para todos os cenários, além disso, mostrou-se ser mais estável para amostras pequenas;
- Os dois Estimadores não paramétricos propostos, θ_{Γ}^4 (*estimador moda da frequência empírica isotonizada*) e θ_{Γ}^5 (*estimador por regressão unimodal interpolado*), apresentaram resultados tão bom quanto àqueles já apresentados na literatura, sendo que, entre aqueles que estimam a dose alvo por uma das doses testadas, o estimador $\hat{\theta}_{\Gamma}^4$ apresentou resultado ligeiramente melhor do que $\hat{\theta}_{\Gamma}^{10}$, com a vantagem de poder ser utilizado em amostras pequenas.
- O delineamento bayesiano isotônico sequencial - DBIS apresentou melhores resultados para os cenários estudados em relação ao delineamento DNM, com redução do EQM e viés, quando se utiliza θ_{Γ}^9 (Média Ponderada) como estimador .

- Os resultados das simulações recomendam o uso do DBIS para ensaios de dose-resposta.

REFERÊNCIAS

- AYER, M. et al. An empirical distribution function for sampling with incomplete information. **The Annals of Mathematical Statistics**, Beachwood, v. 26, n. 4, p. 641-647, Dec. 1955.
- BARBIN, D. **Componentes de variância: teoria e aplicações**. 2. ed. Piracicaba: FEALQ, 1998. 123 p.
- BARLOW, R. E. et al. **Statistical inference under order restrictions: the theory and application of isotonic regression**. New York: J. Wiley, 1972. 388 p.
- BARTHOLOMEW, D. J. A test of homogeneity for ordered alternatives. **Biometrika**, Oxford, v. 46, n. 1/2, p. 36-48, June 1959a.
- _____. A test of homogeneity for ordered alternatives II. **Biometrika**, Oxford, v. 46, n. 3/4, p. 328-335, Dec. 1959b.
- BATES, D. M.; WATTS, D. G. **Nonlinear regression analysis and its applications**. New York: J. Wiley, 1988. 365 p.
- BERKSON, J. Application to the logistic function to bio-assay. **Journal of the American Statistical Association**, Washington, v. 39, n. 227, p. 357-365, Sept. 1944.
- BLISS, C. I. The calculator of the dosage-mortality curve. **Annals of Applied Biology**, Warwick, v. 22, n. 1, p. 134-167, Feb. 1935.
- _____. The method of probits. **Science**, New York, v. 79, n. 2037, p. 38-39, Jan. 1934a.
- _____. The method of probits: a correction. **Science**, New York, v. 79, n. 2053, p. 409-410, May 1934b.
- BRUNK, H. D. Maximum likelihood estimates of monotone parameters. **The Annals of Mathematical Statistics**, Beachwood, v. 26, n. 4, p. 607-616, Dec. 1955.
- DERMAN, C. Nonparametric up and down experimentation. **The Annals of Mathematical Statistics**, Beachwood, v. 28, n. 3, p. 795-798, Sept. 1957.

DIXON, W. J.; MOOD, A. M. A method for obtaining and analyzing sensitivity data. **Journal of the American Statistical Association**, Washington, v. 43, n. 241, p. 109-126, Mar. 1948.

DRAPER, N. R.; SMITH, H. **Applied regression analysis**. New York: J. Wiley, 1998. 706 p.

DURHAM, S. D.; FLOURNOY, N.; ROSENBERGER, W. F. A random walk rule for phase 1 clinical trials. **Biometrics**, Arlington, v. 53, n. 2, p. 745-760, June 1997.

EEDEN, C. V. Estimation in restricted parameter spaces: some history and some recent developments. **Quarterly**, Amsterdam, v. 9, n. 1, p. 69-76, Mar. 1996.

FERREIRA, D. F. **Estatística multivariada**. Lavras: UFLA, 2008. 662 p.

FISHER, R. A. Appendix to Bliss (1935): the case of zero survivors. **Annals of Applied Biology**, Warwick, v. 22, n. 1, p. 164-165, Feb. 1935.

GERKE, O.; SIEDENTOP, H. Optimal phase 1 dose-escalation trial designs in oncology: a simulation study. **Statistics in Medicine**, Malden, v. 27, n. 28, p. 5329-5344, Sept. 2007.

GHOSH, M.; MUKHOPADHYAY, N.; SEN, P. K. **Sequential estimation**. New York: J. Wiley, 1997. 480 p.

GIOVAGNOLI, A.; PINTACUDA, N. Properties of frequency distribution induced by general "up-and-down" methods for estimating quantiles. **Journal of Statistical Planning and Inference**, Cupertino, v. 74, n. 1, p. 51-63, Jan. 1998.

HERBACH, L. H. Properties of model II: type analysis of variance tests, a optimum nature of the f-test for model II, in the balanced case. **The Annals of Mathematical Statistics**, Beachwood, v. 30, n. 4, p. 939-959, Dec. 1959.

IVANOVA, A. et al. Improved up-and-down designs for phase 1 trials. **Statistics in Medicine**, Malden, v. 22, n. 1, p. 69-82, Jan. 2003.

IVANOVA, A.; FLOURNOY, N. Comparison of isotonic designs for dose-finding. **Statistics in Biopharmaceutical Research**, Washington, v. 1, n. 1, p. 101-112, Feb. 2009.

- LEEUW, J.; HORNIK, K.; MAIR, P. Isotone optimization in R: pool, adjacent, violators algorithm (pava) and active set methods. **Journal of Statistical Software**, Los Angeles, v. 32, n. 5, p. 1-24, Oct. 2009.
- LEUNG, D. H. Y.; WANG, Y. G. Isotonic designs for phase 1 trials. **Controlled Clinical Trials**, New York, v. 22, n. 2, p. 126-138, Apr. 2001.
- LIMA, E. L. **Curso de análise 1**. Rio de Janeiro: LTC, 1976. 344 p.
- _____. **Curso de análise 2**. Rio de Janeiro: LTC, 1981. 547 p.
- LIPSCHUTZ, S. **Teoria dos conjuntos**. São Paulo: Makron Books do Brasil, 1972. 336 p.
- MCARTHUR, L. C. et al. Investigation of the influence of seagrass on fisheries catch level using isotonic regression. **Environmental Modeling and Assessment**, Heidelberg, v. 8, n. 4, p. 285-290, Dec. 2003.
- MOOD, A. M.; GRAYBILL, F.; BOES, D. C. **Introduction to the theory of statistics**. Singapore: McGraw-Hill Book, 1974. 564 p.
- NELDER, J. A. The interpretation of negative components of variance. **Biometrika**, Oxford, v. 41, n. 3/4, p. 544-548, Dec. 1954.
- O'QUIGLEY, J.; PEPE, M.; FISHER, L. Continual reassessment method: a practical design for phase 1 clinical trials in cancer. **Biometrics**, Arlington, v. 46, n. 1, p. 33-48, Mar. 1990.
- PACE, N. L.; STYLIANOU, M. P. Advances in and limitations of up-and-down methodology. **Anesthesiology**, Bethesda, v. 107, n. 1, p. 144-152, July 2007.
- PAUL, R. K.; ROSENBERGER, W. F.; FLOURNOY, N. Quantile estimation following non-parametric phase 1 clinical trials with ordinal response. **Statistics in Medicine**, Malden, v. 23, n. 16, p. 2483-2495, Aug. 2004.
- PAULINO, C. D.; TURKMAN, M. A. A.; MURTEIRA, B. **Estatística bayesiana**. Lisboa: Calouste Gulbenkian, 2003. 400 p.
- R DEVELOPMENT CORE TEAM. **R: a language and environment for statistical computing**. Vienna: R Foundation for Statistical Computing, 2009. Disponível em: <<http://www.R-project.org>>. Acesso em: 15 jul. 2010.

- RAO, C. R. **Linear statistical inference and its applications**. New York: J. Wiley, 2001. 656 p.
- ROBERTSON, T.; WRIGHT, F. T.; DYKSTRA, R. L. **Order restricted statistical inference**. New York: J. Wiley, 1988. 521 p.
- RYAN, L. M. Statistical issues in toxicology. **Journal of the American Statistical Association**, Washington, v. 95, n. 449, p. 304-308, Mar. 2000.
- SEARLE, S. R. **Linear models for unbalanced data**. New York: J. Wiley, 1987. 536 p.
- SIMON, R. et al. Accelerated titration designs for phase 1 clinical trials in oncology. **Journal of the National Cancer Institute**, Bethesda, v. 89, n. 15, p. 1138-1147, Aug. 1997.
- STORER, B. E. Design and analysis of phase 1 clinical trials. **Biometrics**, Arlington, v. 45, n. 3, p. 925-937, Sept. 1989.
- STYLIANOU, M.; FLOURNOY, N. Dose finding using the biased coin up-and-down design and isotonic regression. **Biometrics**, Arlington, v. 58, n. 1, p. 171-177, Mar. 2002.
- STYLIANOU, M.; FOLLMANN, D. A. The accelerated biased coin up-and-down design in phase 1 trials. **Journal of Biopharmaceutical Statistics**, New York, v. 14, n. 1, p. 249-260, Dec. 2004.
- STYLIANOU, M.; PROSCHAN, M.; FLOURNOY, N. Estimating the probability of toxicity at the target dose following an up-and-down design. **Statistics in Medicine**, Malden, v. 22, n. 4, p. 535-543, Feb. 2003.
- THOMPSON, J. W. A. The problem of negative estimates of variance components. **The Annals of Mathematical Statistics**, Beachwood, v. 33, n. 1, p. 273-289, Mar. 1962.
- THOMPSON, J. W. A.; MOORE, J. R. Non-negative estimates of variance components. **Technometrics**, Washington, v. 5, n. 4, p. 441-449, Nov. 1963.
- TOURNEAU, C. L.; LEE, J. J.; SIU, L. L. Dose escalation methods in phase 1 cancer clinical trials. **Journal of the National Cancer Institute**, Bethesda, v. 101, n. 10, p. 708-720, May 2009.

APÊNDICE

APÊNDICE	Páginas
APÊNDICE A Resultados gráficos para o EQM e Viés dos estimadores $\hat{\theta}_\Gamma^1$ e $\hat{\theta}_\Gamma^6$, confrontados para tamanhos de amostras $n = 2, 3, \dots, 60$, amostras geradas considerando uma Distribuição Normal e o Delineamento da Moeda Viciada (BCD).	252
APÊNDICE B Resultados gráficos para o EQM e Viés dos estimadores $\hat{\theta}_\Gamma^4$, $\hat{\theta}_\Gamma^5$, $\hat{\theta}_\Gamma^9$, $\hat{\theta}_\Gamma^{10}$ e $\hat{\theta}_\Gamma^{11}$, confrontados para tamanhos de amostras $n = 2, 3, \dots, 60$, amostras geradas considerando uma Distribuição Normal e o Delineamento da Moeda Viciada (BCD).	261
APÊNDICE C Resultados gráficos para o EQM e Viés dos estimadores $\hat{\theta}_\Gamma^2$, $\hat{\theta}_\Gamma^3$, $\hat{\theta}_\Gamma^7$ e $\hat{\theta}_\Gamma^8$, confrontados para tamanhos de amostras $n = 2, 3, \dots, 60$, amostras geradas considerando uma Distribuição Logística e o Delineamento da Moeda Viciada (BCD).	270
APÊNDICE D Resultados gráficos para o EQM e Viés dos estimadores $\hat{\theta}_\Gamma^4$, $\hat{\theta}_\Gamma^5$, $\hat{\theta}_\Gamma^9$, $\hat{\theta}_\Gamma^{10}$ e $\hat{\theta}_\Gamma^{11}$, confrontados para tamanhos de amostras $n = 2, 3, \dots, 60$, amostras geradas considerando uma Distribuição Logística e o Delineamento da Moeda Viciada (BCD).	279
APÊNDICE E Resultados gráficos para o EQM e Viés dos estimadores $\hat{\theta}_\Gamma^4$, $\hat{\theta}_\Gamma^5$, $\hat{\theta}_\Gamma^9$, $\hat{\theta}_\Gamma^{10}$ e $\hat{\theta}_\Gamma^{11}$, confrontados para tamanhos de amostras $n = 2, 3, \dots, 60$, amostras geradas considerando uma Distribuição Exponencial e o Delineamento da Moeda Viciada (BCD).	288
APÊNDICE F Resultados gráficos para o EQM e Viés dos estimadores $\hat{\theta}_\Gamma^4$, $\hat{\theta}_\Gamma^5$, $\hat{\theta}_\Gamma^9$, $\hat{\theta}_\Gamma^{10}$ e $\hat{\theta}_\Gamma^{11}$, confrontados para tamanhos de amostras $n = 2, 3, \dots, 60$, amostras geradas considerando uma Distribuição Gama e o Delineamento da Moeda Viciada (BCD).	291
APÊNDICE G Resultados gráficos para o EQM e Viés dos estimadores $\hat{\theta}_\Gamma^4$, $\hat{\theta}_\Gamma^5$, $\hat{\theta}_\Gamma^9$, $\hat{\theta}_\Gamma^{10}$ e $\hat{\theta}_\Gamma^{11}$, confrontados para tamanhos de amostras $n = 2, 3, \dots, 60$, amostras geradas considerando uma Distribuição Gama e o Delineamento da Moeda Viciada (BCD).	294

- APÊNDICE H** Resultados gráficos para o EQM e Viés dos estimadores $\hat{\theta}_\Gamma^4$, $\hat{\theta}_\Gamma^5$, $\hat{\theta}_\Gamma^9$, $\hat{\theta}_\Gamma^{10}$ e $\hat{\theta}_\Gamma^{11}$, confrontados para tamanhos de amostras $n = 2, 3, \dots, 60$, amostras geradas considerando uma Distribuição Normal e o Delineamento de Narayana Modificado (DNM). . 297
- APÊNDICE I** Resultados gráficos para o EQM e Viés dos estimadores $\hat{\theta}_\Gamma^4$, $\hat{\theta}_\Gamma^5$, $\hat{\theta}_\Gamma^9$, $\hat{\theta}_\Gamma^{10}$ e $\hat{\theta}_\Gamma^{11}$, confrontados para tamanhos de amostras $n = 2, 3, \dots, 60$, amostras geradas considerando uma Distribuição Logística e o Delineamento de Narayana Modificado (DNM). 305
- APÊNDICE J** Resultados gráficos para o EQM e Viés dos estimadores $\hat{\theta}_\Gamma^4$, $\hat{\theta}_\Gamma^5$, $\hat{\theta}_\Gamma^9$, $\hat{\theta}_\Gamma^{10}$ e $\hat{\theta}_\Gamma^{11}$, confrontados para tamanhos de amostras $n = 2, 3, \dots, 60$, amostras geradas considerando uma Distribuição Exponencial e o Delineamento de Narayana Modificado (DNM). 314
- APÊNDICE L** Resultados gráficos para o EQM e Viés dos estimadores $\hat{\theta}_\Gamma^4$, $\hat{\theta}_\Gamma^5$, $\hat{\theta}_\Gamma^9$, $\hat{\theta}_\Gamma^{10}$ e $\hat{\theta}_\Gamma^{11}$, confrontados para tamanhos de amostras $n = 2, 3, \dots, 60$, amostras geradas considerando uma Distribuição Gama e o Delineamento de Narayana Modificado (DNM). . . 316
- APÊNDICE M** Resultados gráficos para o EQM e Viés dos estimadores $\hat{\theta}_\Gamma^4$, $\hat{\theta}_\Gamma^5$, $\hat{\theta}_\Gamma^9$, $\hat{\theta}_\Gamma^{10}$ e $\hat{\theta}_\Gamma^{11}$, confrontados para tamanhos de amostras $n = 2, 3, \dots, 60$, amostras geradas considerando uma Distribuição Gama e o Delineamento de Narayana Modificado (DNM). . . 318
- APÊNDICE N** Resultados gráficos para o EQM e Viés dos estimadores $\hat{\theta}_\Gamma^1$ e $\hat{\theta}_\Gamma^6$, confrontados para tamanhos de amostras $n = 2, 3, \dots, 60$, amostras geradas considerando uma Distribuição Normal e o Delineamento Isotônico de Leung e Wang (DILW). 320
- APÊNDICE O** Resultados gráficos para o EQM e Viés dos estimadores $\hat{\theta}_\Gamma^4$, $\hat{\theta}_\Gamma^5$, $\hat{\theta}_\Gamma^9$, $\hat{\theta}_\Gamma^{10}$ e $\hat{\theta}_\Gamma^{11}$, confrontados para tamanhos de amostras $n = 2, 3, \dots, 60$, amostras geradas considerando uma Distribuição Normal e o Delineamento Isotônico de Leung e Wang (DILW). 329
- APÊNDICE P** Resultados gráficos para o EQM e Viés dos estimadores $\hat{\theta}_\Gamma^2$, $\hat{\theta}_\Gamma^3$, $\hat{\theta}_\Gamma^7$ e $\hat{\theta}_\Gamma^8$, confrontados para tamanhos de amostras $n = 2, 3, \dots, 60$, amostras geradas considerando uma Distribuição Logística e o Delineamento Isotônico de Leung e Wang (DILW). 338

- APÊNDICE Q** Resultados gráficos para o EQM e Viés dos estimadores $\hat{\theta}_\Gamma^4$, $\hat{\theta}_\Gamma^5$, $\hat{\theta}_\Gamma^9$, $\hat{\theta}_\Gamma^{10}$ e $\hat{\theta}_\Gamma^{11}$, confrontados para tamanhos de amostras $n = 2, 3, \dots, 60$, amostras geradas considerando uma Distribuição Logística e o Delineamento Isotônico de Leung e Wang (DILW). 347
- APÊNDICE R** Resultados gráficos para o EQM e Viés dos estimadores $\hat{\theta}_\Gamma^4$, $\hat{\theta}_\Gamma^5$, $\hat{\theta}_\Gamma^9$, $\hat{\theta}_\Gamma^{10}$ e $\hat{\theta}_\Gamma^{11}$, confrontados para tamanhos de amostras $n = 2, 3, \dots, 60$, amostras geradas considerando uma Distribuição Exponencial e o Delineamento Isotônico de Leung e Wang (DILW).....356
- APÊNDICE S** Resultados gráficos para o EQM e Viés dos estimadores $\hat{\theta}_\Gamma^4$, $\hat{\theta}_\Gamma^5$, $\hat{\theta}_\Gamma^9$, $\hat{\theta}_\Gamma^{10}$ e $\hat{\theta}_\Gamma^{11}$, confrontados para tamanhos de amostras $n = 2, 3, \dots, 60$, amostras geradas considerando uma Distribuição Gama e o Delineamento Isotônico de Leung e Wang (DILW).359
- APÊNDICE T** Resultados gráficos para o EQM e Viés dos estimadores $\hat{\theta}_\Gamma^4$, $\hat{\theta}_\Gamma^5$, $\hat{\theta}_\Gamma^9$, $\hat{\theta}_\Gamma^{10}$ e $\hat{\theta}_\Gamma^{11}$, confrontados para tamanhos de amostras $n = 2, 3, \dots, 60$, amostras geradas considerando uma Distribuição Gama e o Delineamento Isotônico de Leung e Wang (DILW).362
- APÊNDICE U** Resultados gráficos para o EQM e Viés do estimador $\hat{\theta}_\Gamma^9$ para comparação do DBIS e DNM, confrontados para tamanhos de amostras $n = 2, 3, \dots, 60$, amostras geradas considerando uma Distribuição Exponencial e conjunto de doses $D_1 = \{0,5; 1,0; \dots; 4,0\}$ 365
- APÊNDICE V** Resultados gráficos para o EQM e Viés do estimador $\hat{\theta}_\Gamma^9$ para comparação do DBIS e DNM, confrontados para tamanhos de amostras $n = 2, 3, \dots, 60$, amostras geradas considerando uma Distribuição Gama e conjunto de doses $D_1 = \{0,5; 1,0; 1,5 \dots; 4,0\}$ 368
- APÊNDICE X** Resultados gráficos para o EQM e Viés do estimador $\hat{\theta}_\Gamma^9$ para comparação do DBIS e DNM, confrontados para tamanhos de amostras $n = 2, 3, \dots, 60$, amostras geradas considerando uma Distribuição Exponencial e conjunto de doses $D_2 = \{0,5; 0,75; 1,0; 1,25; \dots; 4,0\}$ 371

APÊNDICE Z	Resultados gráficos para o EQM e Viés do estimador $\hat{\theta}_T^9$ para comparação do DBIS e DNM, confrontados para tamanhos de amostras $n = 2, 3, \dots, 60$, amostras geradas considerando uma Distribuição Gama e conjunto de doses $D_2 = \{0,5; 0,75; 1,0; 1,25; \dots; 4,0\}$	374
APÊNDICE AA	Rotina para simular uma amostra obtida pelo Delineamento da Moeda Viciada - BCD.	377
APÊNDICE AB	Rotina para simular uma amostra obtida pelo Delineamento de Narayana Modificado - DNM.	379
APÊNDICE AC	Rotina para simular uma amostra obtida pelo Delineamento de Isotônico de Leung e Wang - DILW.	381
APÊNDICE AD	Rotina para simular uma amostra obtida pelo Delineamento Bayesiano Isotônico Sequencial - DBIS.	383
APÊNDICE AE	Rotina para calcular estimativas pelo estimador $\hat{\theta}_T^1$	385
APÊNDICE AF	Rotina para calcular estimativas pelo estimador $\hat{\theta}_T^2$	389
APÊNDICE AG	Rotina para calcular estimativas pelo estimador $\hat{\theta}_T^3$	393
APÊNDICE AH	Rotina para calcular estimativas pelo estimador $\hat{\theta}_T^4$	395
APÊNDICE AI	Rotina para calcular estimativas pelo estimador $\hat{\theta}_T^5$	396
APÊNDICE AJ	Rotina para calcular estimativas pelo estimador $\hat{\theta}_T^6$	398
APÊNDICE AL	Rotina para calcular estimativas pelo estimador $\hat{\theta}_T^7$	402
APÊNDICE AM	Rotina para calcular estimativas pelo estimador $\hat{\theta}_T^8$	406
APÊNDICE AN	Rotina para calcular estimativas pelo estimador $\hat{\theta}_T^9$	408

APÊNDICE AO Rotina para calcular estimativas pelo estimador $\hat{\theta}_{\Gamma}^{10}$ 409

APÊNDICE AP Rotina para calcular estimativas pelo estimador $\hat{\theta}_{\Gamma}^{11}$ 410

APÊNDICE A

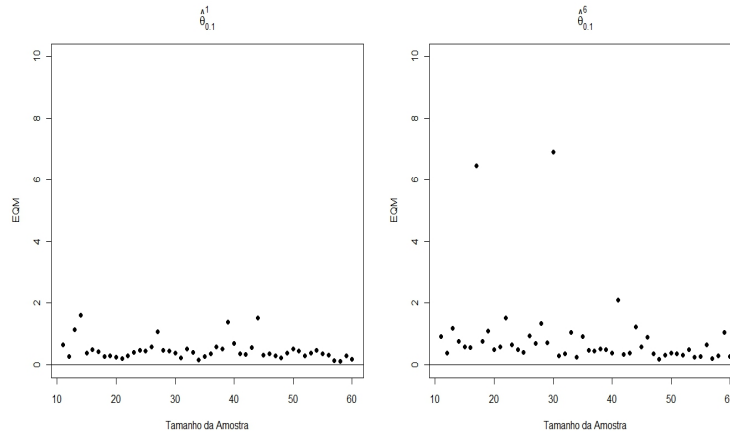


Figura 97 EQM para cada tamanho amostral e os estimadores $\hat{\theta}_{\Gamma}^1$ e $\hat{\theta}_{\Gamma}^6$, amostras geradas considerando uma **Distribuição Normal** com $\mu = 2,25$ e $\sigma^2 = 1$ para $\Gamma = 0,10$ e o **BCD**

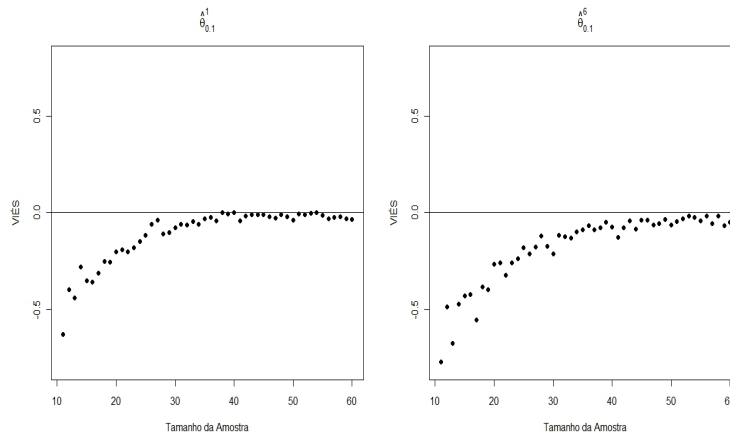


Figura 98 Viés para cada tamanho amostral e os estimadores $\hat{\theta}_{\Gamma}^1$ e $\hat{\theta}_{\Gamma}^6$, amostras geradas considerando uma **Distribuição Normal** com $\mu = 2,25$ e $\sigma^2 = 1$ para $\Gamma = 0,10$ e o **BCD**

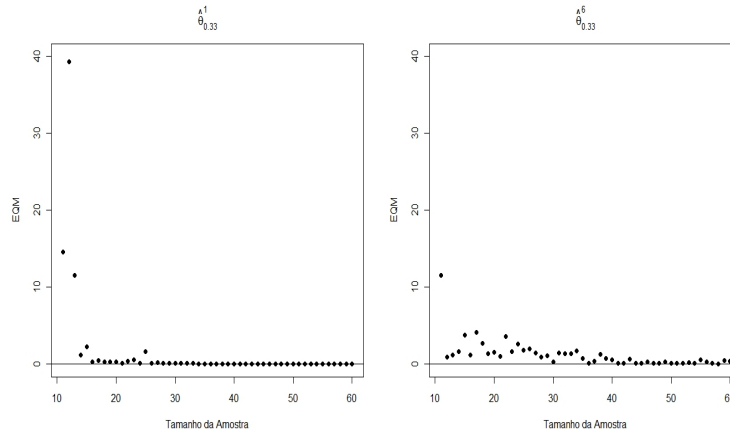


Figura 99 EQM para cada tamanho amostral e os estimadores $\hat{\theta}_{\Gamma}^1$ e $\hat{\theta}_{\Gamma}^6$, amostras geradas considerando uma **Distribuição Normal** com $\mu = 2,25$ e $\sigma^2 = 1$ para $\Gamma = 0,33$ e o **BCD**

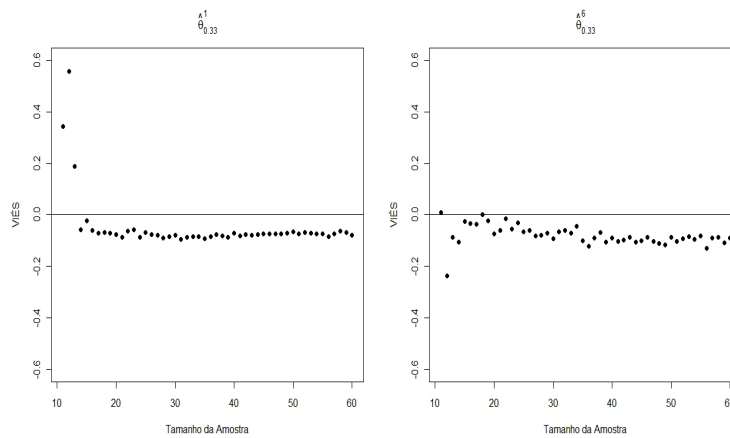


Figura 100 Viés para cada tamanho amostral e os estimadores $\hat{\theta}_{\Gamma}^1$ e $\hat{\theta}_{\Gamma}^6$, amostras geradas considerando uma **Distribuição Normal** com $\mu = 2,25$ e $\sigma^2 = 1$ para $\Gamma = 0,33$ e o **BCD**

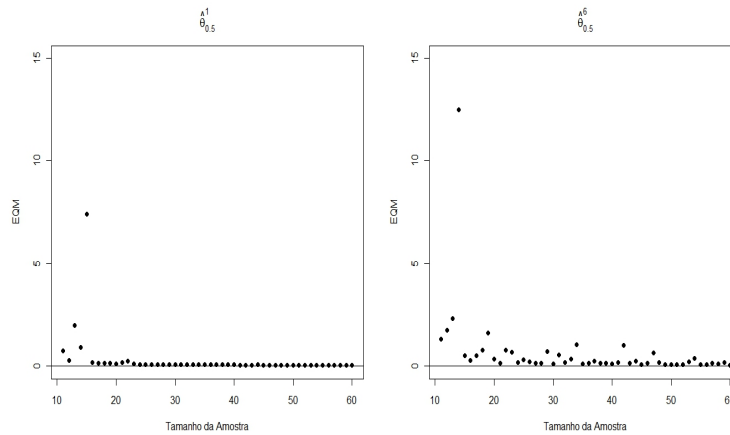


Figura 101 EQM para cada tamanho amostral e os estimadores $\hat{\theta}_{\Gamma}^1$ e $\hat{\theta}_{\Gamma}^6$, amostras geradas considerando uma **Distribuição Normal** com $\mu = 2,25$ e $\sigma^2 = 1$ para $\Gamma = 0,50$ e o **BCD**

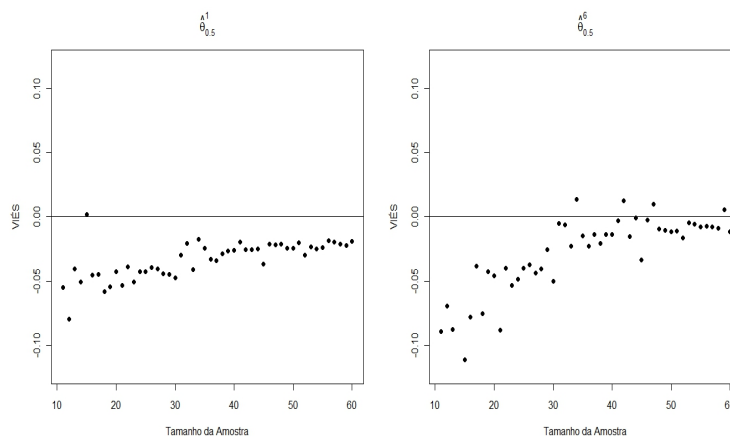


Figura 102 Viés para cada tamanho amostral e os estimadores $\hat{\theta}_{\Gamma}^1$ e $\hat{\theta}_{\Gamma}^6$, amostras geradas considerando uma **Distribuição Normal** com $\mu = 2,25$ e $\sigma^2 = 1$ para $\Gamma = 0,50$ e o **BCD**

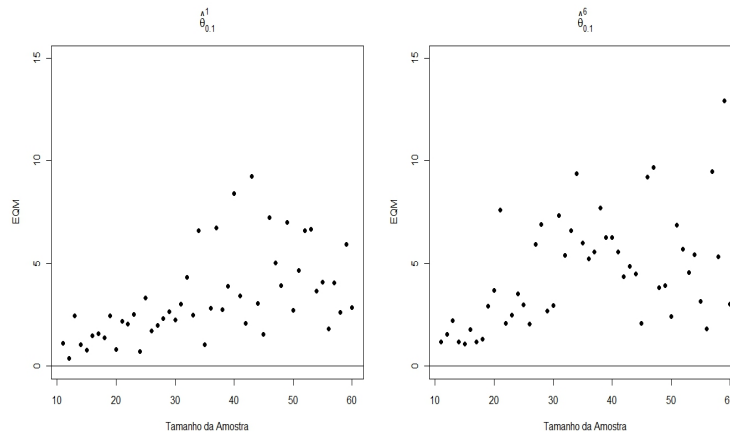


Figura 103 EQM para cada tamanho amostral e os estimadores $\hat{\theta}_{\Gamma}^1$ e $\hat{\theta}_{\Gamma}^6$, amostras geradas considerando uma **Distribuição Normal** com $\mu = 1,25$ e $\sigma^2 = 1$ para $\Gamma = 0,10$ e o **BCD**

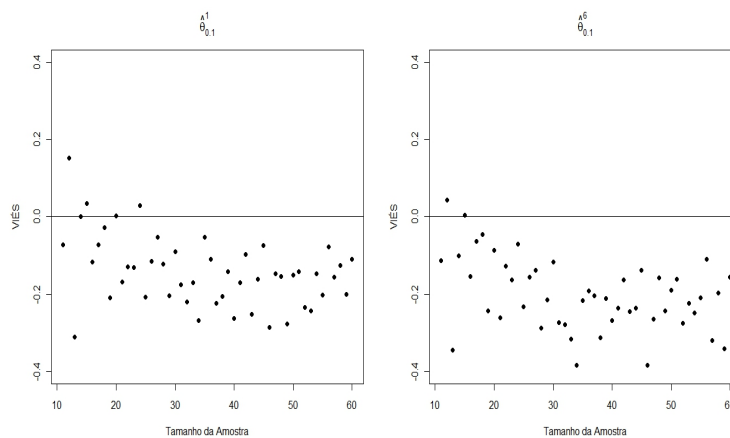


Figura 104 Viés para cada tamanho amostral e os estimadores $\hat{\theta}_{\Gamma}^1$ e $\hat{\theta}_{\Gamma}^6$, amostras geradas considerando uma **Distribuição Normal** com $\mu = 1,25$ e $\sigma^2 = 1$ para $\Gamma = 0,10$ e o **BCD**

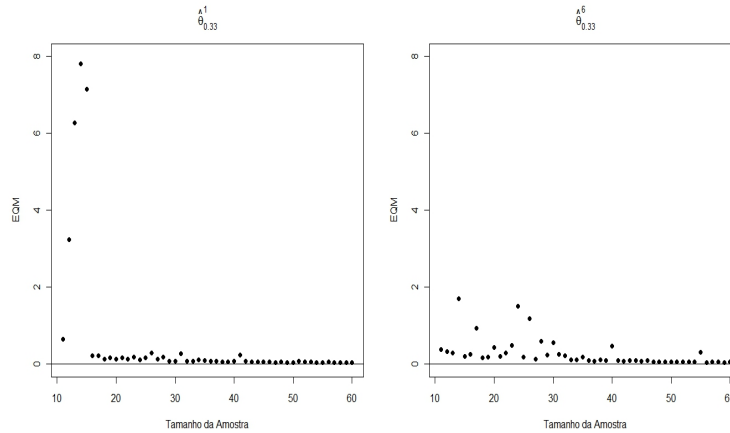


Figura 105 EQM para cada tamanho amostral e os estimadores $\hat{\theta}_{\Gamma}^1$ e $\hat{\theta}_{\Gamma}^6$, amostras geradas considerando uma **Distribuição Normal** com $\mu = 1,25$ e $\sigma^2 = 1$ para $\Gamma = 0,33$ e o **BCD**

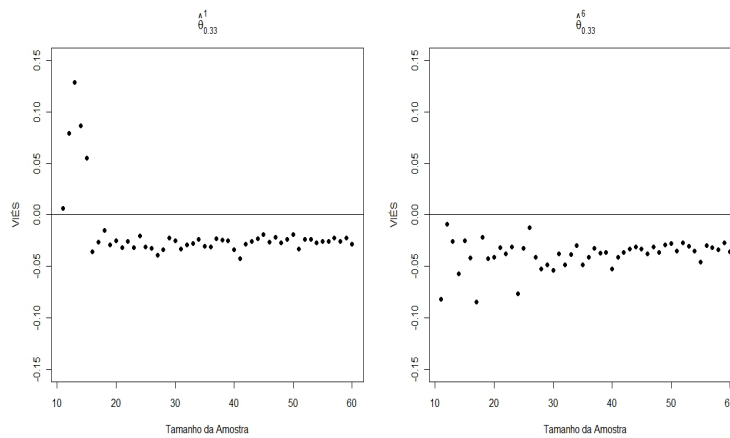


Figura 106 Viés para cada tamanho amostral e os estimadores $\hat{\theta}_{\Gamma}^1$ e $\hat{\theta}_{\Gamma}^6$, amostras geradas considerando uma **Distribuição Normal** com $\mu = 1,25$ e $\sigma^2 = 1$ para $\Gamma = 0,33$ e o **BCD**

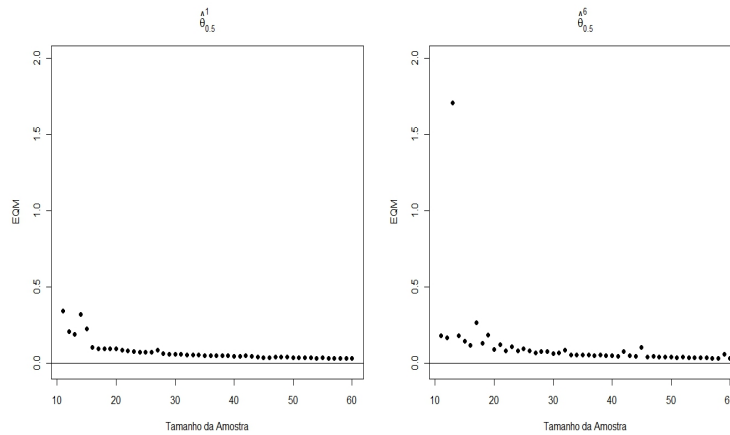


Figura 107 EQM para cada tamanho amostral e os estimadores $\hat{\theta}_{\Gamma}^1$ e $\hat{\theta}_{\Gamma}^6$, amostras geradas considerando uma **Distribuição Normal** com $\mu = 1,25$ e $\sigma^2 = 1$ para $\Gamma = 0,50$ e o **BCD**

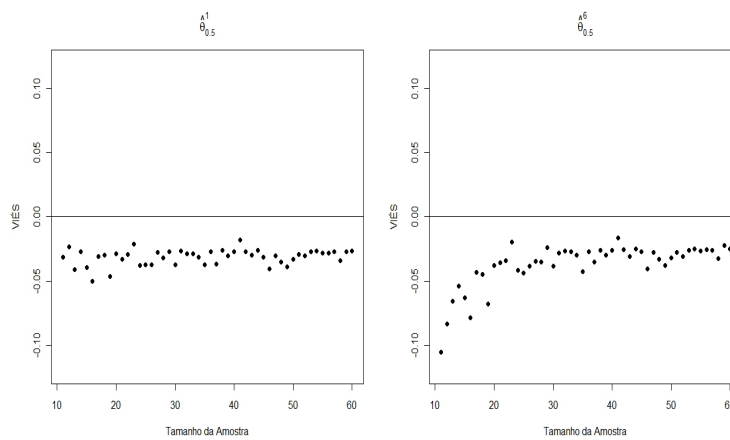


Figura 108 Viés para cada tamanho amostral e os estimadores $\hat{\theta}_{\Gamma}^1$ e $\hat{\theta}_{\Gamma}^6$, amostras geradas considerando uma **Distribuição Normal** com $\mu = 1,25$ e $\sigma^2 = 1$ para $\Gamma = 0,50$ e o **BCD**

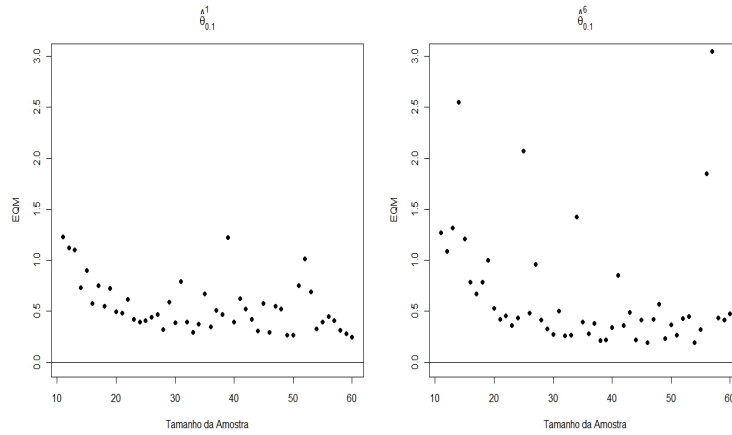


Figura 109 EQM para cada tamanho amostral e os estimadores $\hat{\theta}_{\Gamma}^1$ e $\hat{\theta}_{\Gamma}^6$, amostras geradas considerando uma **Distribuição Normal** com $\mu = 3,25$ e $\sigma^2 = 1$ para $\Gamma = 0,10$ e o **BCD**

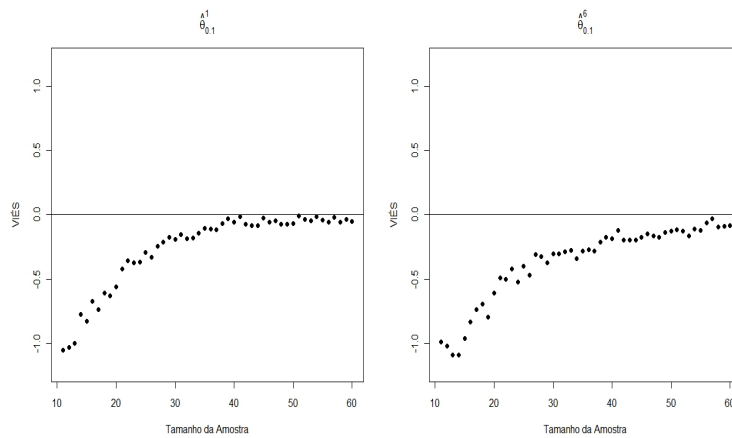


Figura 110 Viés para cada tamanho amostral e os estimadores $\hat{\theta}_{\Gamma}^1$ e $\hat{\theta}_{\Gamma}^6$, amostras geradas considerando uma **Distribuição Normal** com $\mu = 3,25$ e $\sigma^2 = 1$ para $\Gamma = 0,10$ e o **BCD**

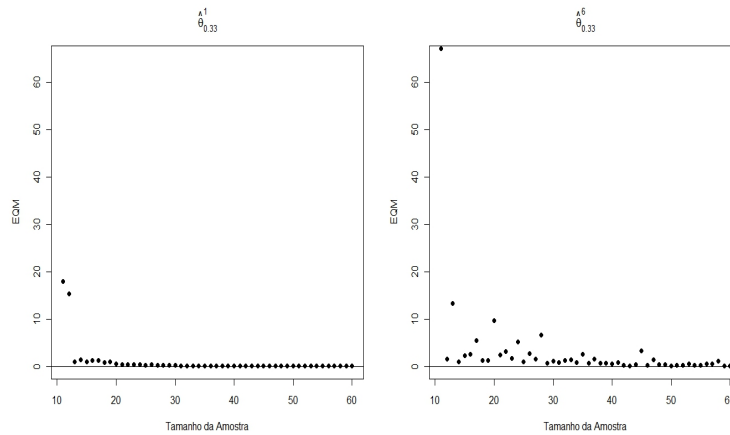


Figura 111 EQM para cada tamanho amostral e os estimadores $\hat{\theta}_{\Gamma}^1$ e $\hat{\theta}_{\Gamma}^6$, amostras geradas considerando uma **Distribuição Normal** com $\mu = 3,25$ e $\sigma^2 = 1$ para $\Gamma = 0,33$ e o **BCD**

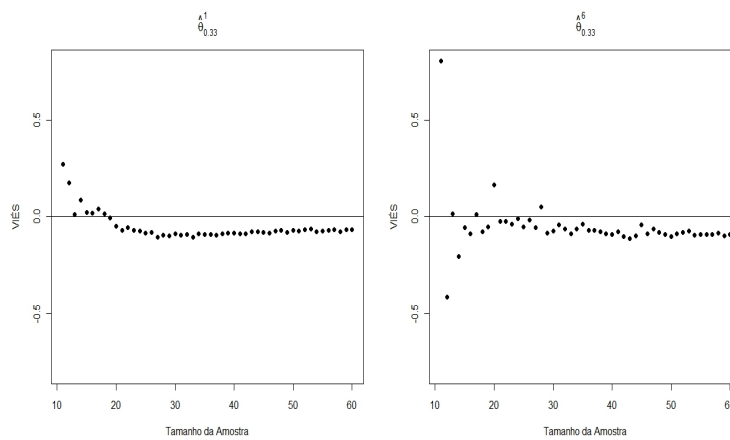


Figura 112 Viés para cada tamanho amostral e os estimadores $\hat{\theta}_{\Gamma}^1$ e $\hat{\theta}_{\Gamma}^6$, amostras geradas considerando uma **Distribuição Normal** com $\mu = 3,25$ e $\sigma^2 = 1$ para $\Gamma = 0,33$ e o **BCD**

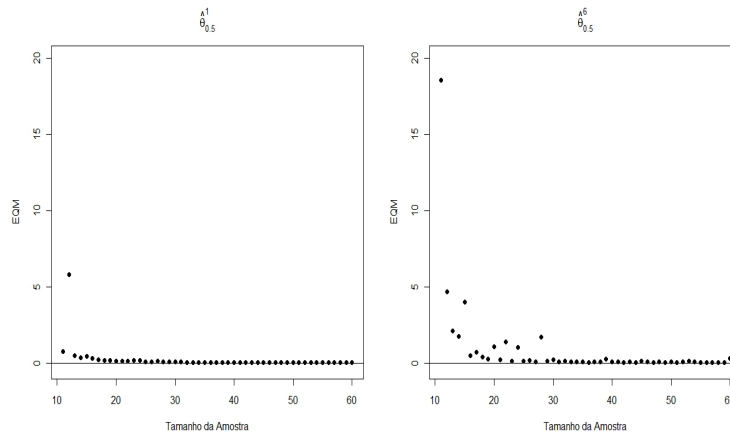


Figura 113 EQM para cada tamanho amostral e os estimadores $\hat{\theta}_{\Gamma}^1$ e $\hat{\theta}_{\Gamma}^6$, amostras geradas considerando uma **Distribuição Normal** com $\mu = 3,25$ e $\sigma^2 = 1$ para $\Gamma = 0,50$ e o **BCD**

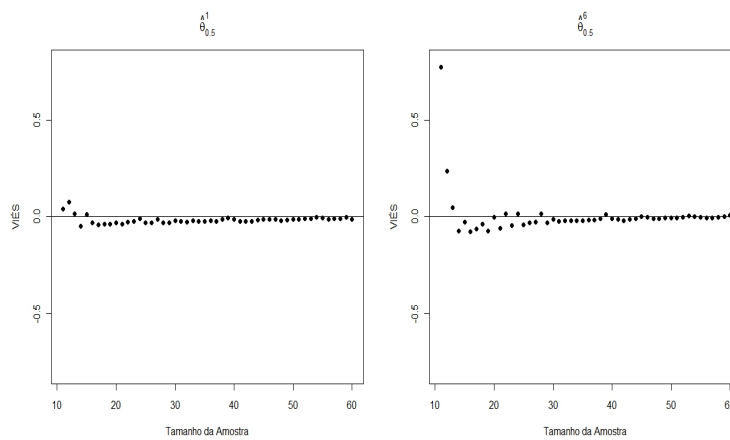


Figura 114 Viés para cada tamanho amostral e os estimadores $\hat{\theta}_{\Gamma}^1$ e $\hat{\theta}_{\Gamma}^6$, amostras geradas considerando uma **Distribuição Normal** com $\mu = 3,25$ e $\sigma^2 = 1$ para $\Gamma = 0,50$ e o **BCD**

APÊNDICE B

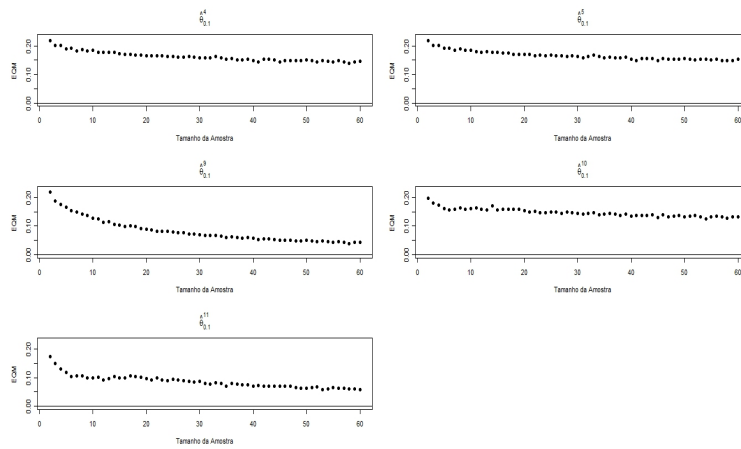


Figura 115 EQM para cada tamanho amostral e os estimadores $\hat{\theta}_{\Gamma}^4$, $\hat{\theta}_{\Gamma}^5$, $\hat{\theta}_{\Gamma}^9$, $\hat{\theta}_{\Gamma}^{10}$ e $\hat{\theta}_{\Gamma}^{11}$, amostras geradas considerando uma **Distribuição Normal** com $\mu = 2,25$ e $\sigma^2 = 1$ para $\Gamma = 0,10$ e o **BCD**

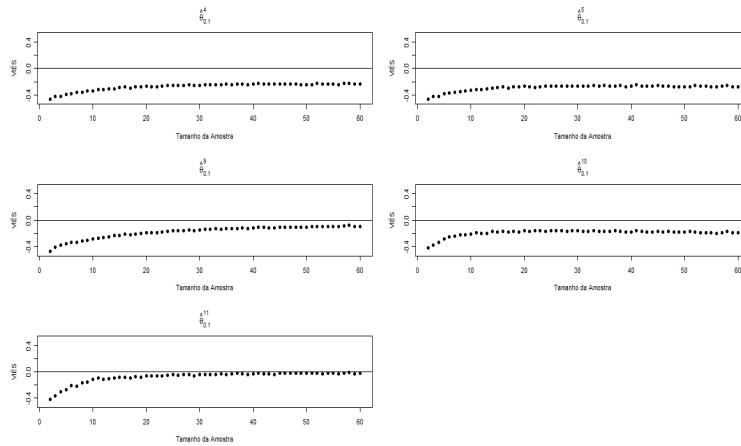


Figura 116 Viés para cada tamanho amostral e os estimadores $\hat{\theta}_{\Gamma}^4$, $\hat{\theta}_{\Gamma}^5$, $\hat{\theta}_{\Gamma}^9$, $\hat{\theta}_{\Gamma}^{10}$ e $\hat{\theta}_{\Gamma}^{11}$, amostras geradas considerando uma **Distribuição Normal** com $\mu = 2,25$ e $\sigma^2 = 1$ para $\Gamma = 0,10$ e o **BCD**

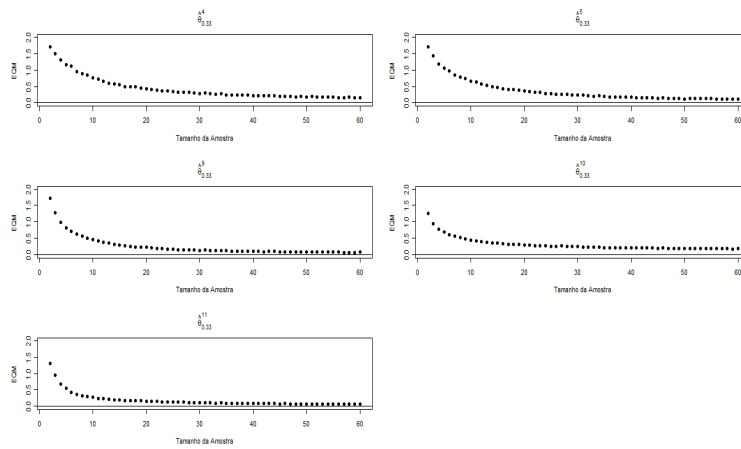


Figura 117 EQM para cada tamanho amostral e os estimadores $\hat{\theta}_{\Gamma}^4$, $\hat{\theta}_{\Gamma}^5$, $\hat{\theta}_{\Gamma}^9$, $\hat{\theta}_{\Gamma}^{10}$ e $\hat{\theta}_{\Gamma}^{11}$, amostras geradas considerando uma **Distribuição Normal** com $\mu = 2,25$ e $\sigma^2 = 1$ para $\Gamma = 0,33$ e o **BCD**

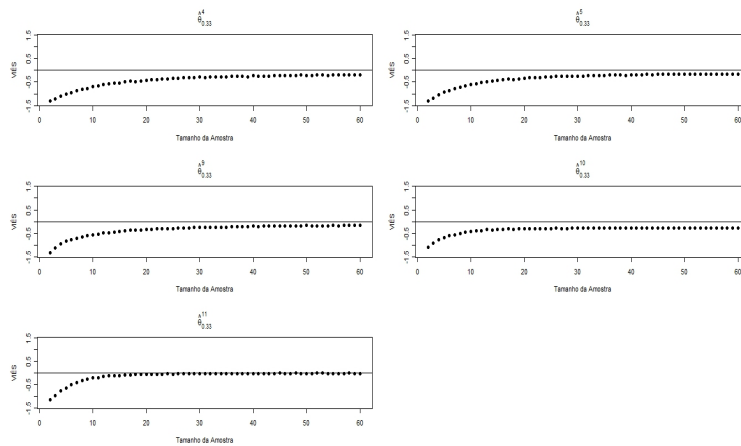


Figura 118 Viés para cada tamanho amostral e os estimadores $\hat{\theta}_{\Gamma}^4$, $\hat{\theta}_{\Gamma}^5$, $\hat{\theta}_{\Gamma}^9$, $\hat{\theta}_{\Gamma}^{10}$ e $\hat{\theta}_{\Gamma}^{11}$, amostras geradas considerando uma **Distribuição Normal** com $\mu = 2,25$ e $\sigma^2 = 1$ para $\Gamma = 0,33$ e o **BCD**

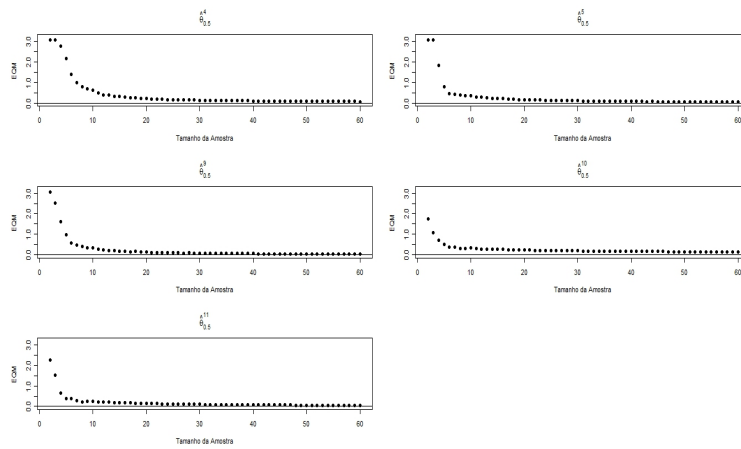


Figura 119 EQM para cada tamanho amostral e os estimadores $\hat{\theta}_{\Gamma}^4$, $\hat{\theta}_{\Gamma}^5$, $\hat{\theta}_{\Gamma}^9$, $\hat{\theta}_{\Gamma}^{10}$ e $\hat{\theta}_{\Gamma}^{11}$, amostras geradas considerando uma **Distribuição Normal** com $\mu = 2,25$ e $\sigma^2 = 1$ para $\Gamma = 0,50$ e o **BCD**

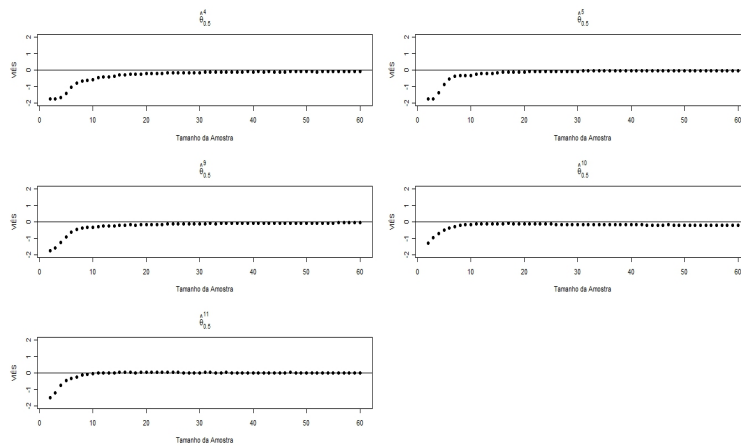


Figura 120 Viés para cada tamanho amostral e os estimadores $\hat{\theta}_{\Gamma}^4$, $\hat{\theta}_{\Gamma}^5$, $\hat{\theta}_{\Gamma}^9$, $\hat{\theta}_{\Gamma}^{10}$ e $\hat{\theta}_{\Gamma}^{11}$, amostras geradas considerando uma **Distribuição Normal** com $\mu = 2,25$ e $\sigma^2 = 1$ para $\Gamma = 0,50$ e o **BCD**

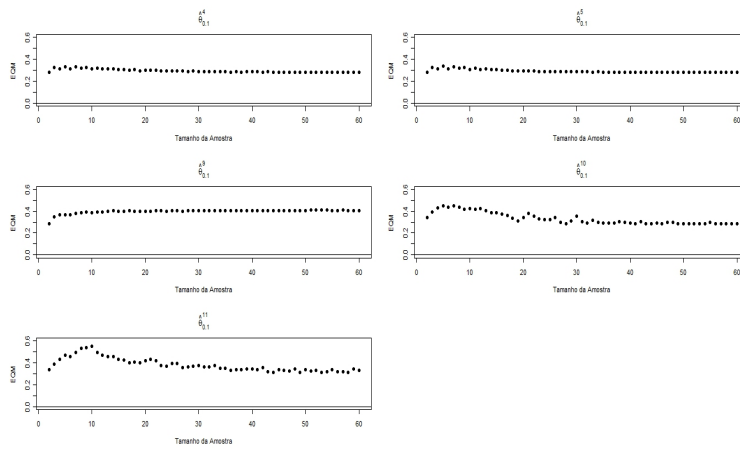


Figura 121 EQM para cada tamanho amostral e os estimadores $\hat{\theta}_{\Gamma}^4$, $\hat{\theta}_{\Gamma}^5$, $\hat{\theta}_{\Gamma}^9$, $\hat{\theta}_{\Gamma}^{10}$ e $\hat{\theta}_{\Gamma}^{11}$, amostras geradas considerando uma **Distribuição Normal** com $\mu = 1,25$ e $\sigma^2 = 1$ para $\Gamma = 0,10$ e o **BCD**

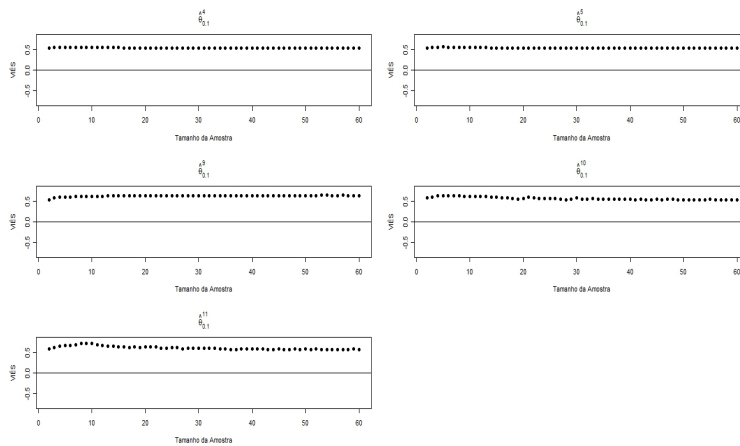


Figura 122 Viés para cada tamanho amostral e os estimadores $\hat{\theta}_{\Gamma}^4$, $\hat{\theta}_{\Gamma}^5$, $\hat{\theta}_{\Gamma}^9$, $\hat{\theta}_{\Gamma}^{10}$ e $\hat{\theta}_{\Gamma}^{11}$, amostras geradas considerando uma **Distribuição Normal** com $\mu = 1,25$ e $\sigma^2 = 1$ para $\Gamma = 0,10$ e o **BCD**

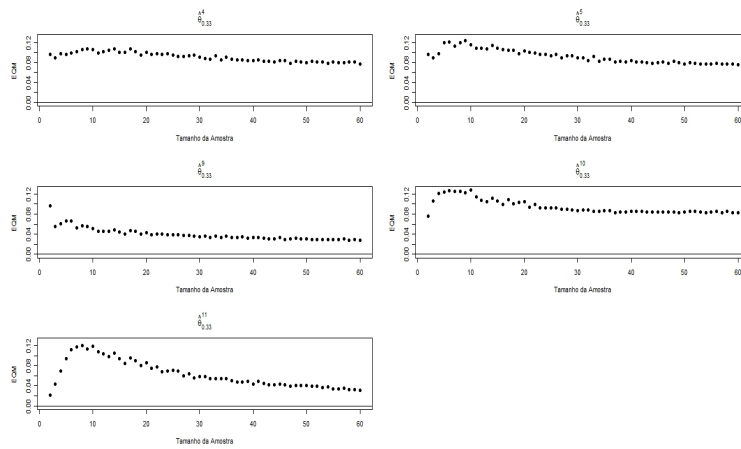


Figura 123 EQM para cada tamanho amostral e os estimadores $\hat{\theta}_{\Gamma}^4$, $\hat{\theta}_{\Gamma}^5$, $\hat{\theta}_{\Gamma}^9$, $\hat{\theta}_{\Gamma}^{10}$ e $\hat{\theta}_{\Gamma}^{11}$, amostras geradas considerando uma **Distribuição Normal** com $\mu = 1,25$ e $\sigma^2 = 1$ para $\Gamma = 0,33$ e o **BCD**

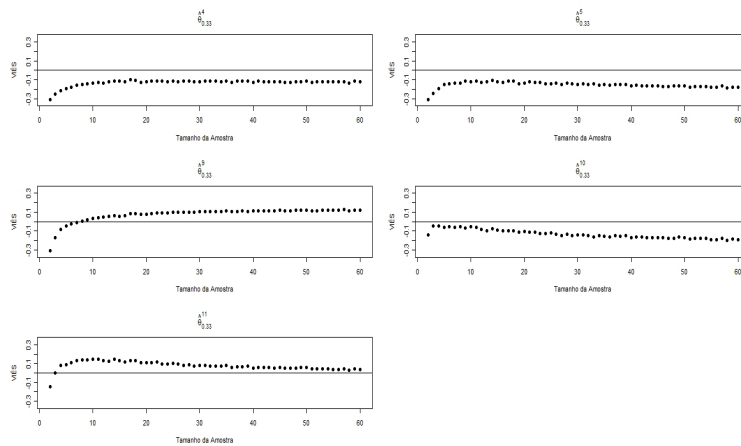


Figura 124 Viés para cada tamanho amostral e os estimadores $\hat{\theta}_{\Gamma}^4$, $\hat{\theta}_{\Gamma}^5$, $\hat{\theta}_{\Gamma}^9$, $\hat{\theta}_{\Gamma}^{10}$ e $\hat{\theta}_{\Gamma}^{11}$, amostras geradas considerando uma **Distribuição Normal** com $\mu = 1,25$ e $\sigma^2 = 1$ para $\Gamma = 0,33$ e o **BCD**

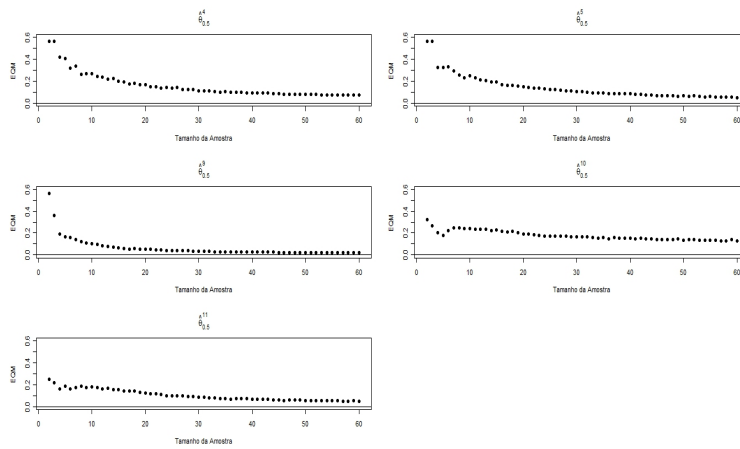


Figura 125 EQM para cada tamanho amostral e os estimadores $\hat{\theta}_{\Gamma}^4$, $\hat{\theta}_{\Gamma}^5$, $\hat{\theta}_{\Gamma}^9$, $\hat{\theta}_{\Gamma}^{10}$ e $\hat{\theta}_{\Gamma}^{11}$, amostras geradas considerando uma **Distribuição Normal** com $\mu = 1,25$ e $\sigma^2 = 1$ para $\Gamma = 0,50$ e o **BCD**

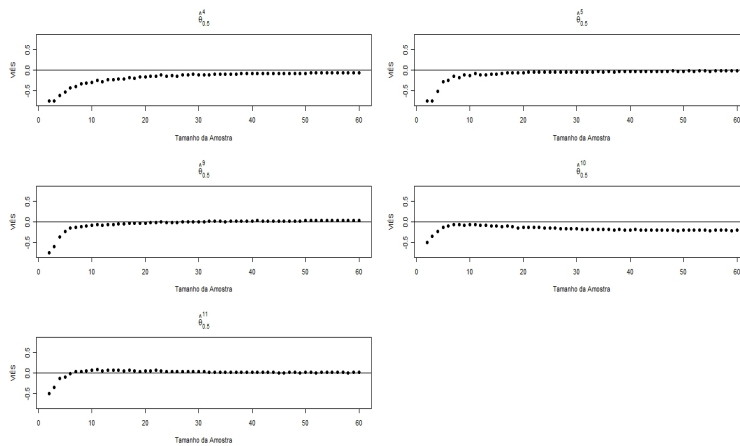


Figura 126 Viés para cada tamanho amostral e os estimadores $\hat{\theta}_{\Gamma}^4$, $\hat{\theta}_{\Gamma}^5$, $\hat{\theta}_{\Gamma}^9$, $\hat{\theta}_{\Gamma}^{10}$ e $\hat{\theta}_{\Gamma}^{11}$, amostras geradas considerando uma **Distribuição Normal** com $\mu = 1,25$ e $\sigma^2 = 1$ para $\Gamma = 0,50$ e o **BCD**

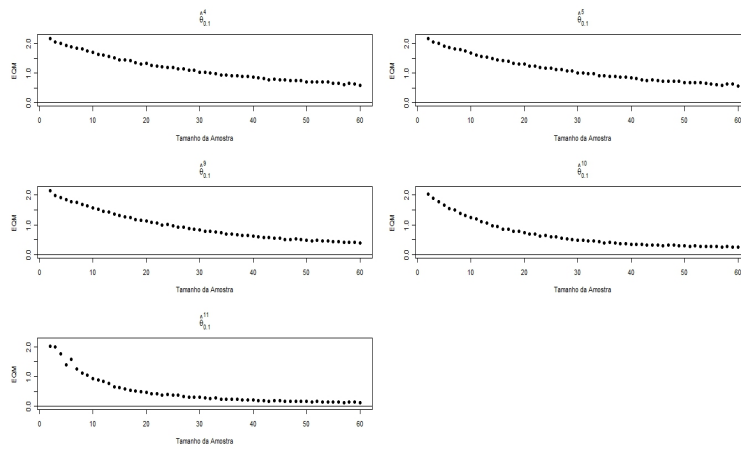


Figura 127 EQM para cada tamanho amostral e os estimadores $\hat{\theta}_{\Gamma}^4$, $\hat{\theta}_{\Gamma}^5$, $\hat{\theta}_{\Gamma}^9$, $\hat{\theta}_{\Gamma}^{10}$ e $\hat{\theta}_{\Gamma}^{11}$, amostras geradas considerando uma **Distribuição Normal** com $\mu = 3,25$ e $\sigma^2 = 1$ para $\Gamma = 0,10$ e o **BCD**

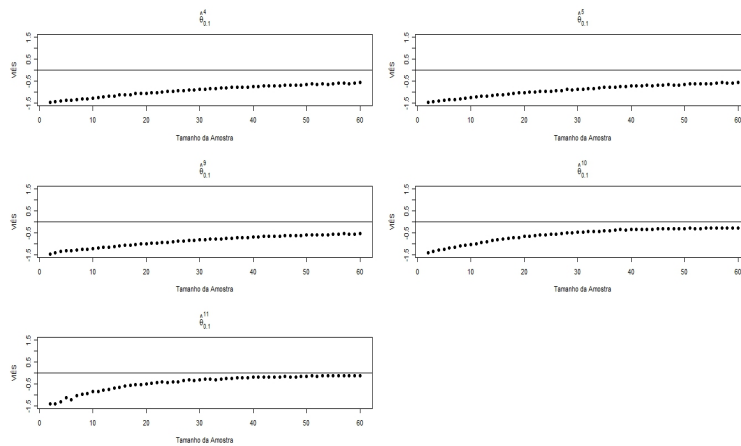


Figura 128 Viés para cada tamanho amostral e os estimadores $\hat{\theta}_{\Gamma}^4$, $\hat{\theta}_{\Gamma}^5$, $\hat{\theta}_{\Gamma}^9$, $\hat{\theta}_{\Gamma}^{10}$ e $\hat{\theta}_{\Gamma}^{11}$, amostras geradas considerando uma **Distribuição Normal** com $\mu = 3,25$ e $\sigma^2 = 1$ para $\Gamma = 0,10$ e o **BCD**

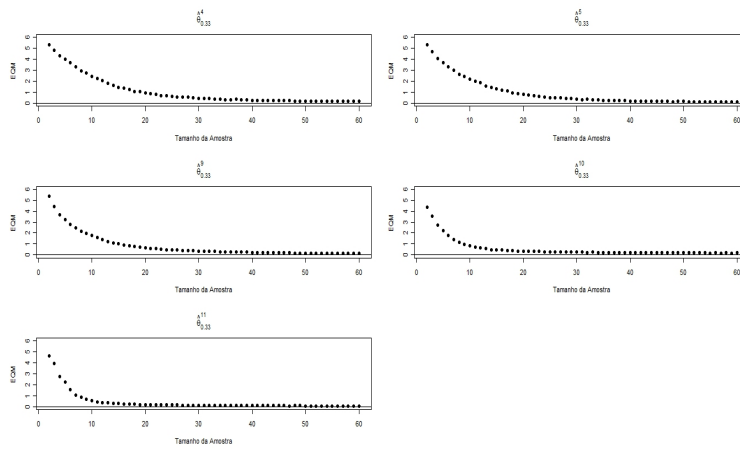


Figura 129 EQM para cada tamanho amostral e os estimadores $\hat{\theta}_{\Gamma}^4$, $\hat{\theta}_{\Gamma}^5$, $\hat{\theta}_{\Gamma}^9$, $\hat{\theta}_{\Gamma}^{10}$ e $\hat{\theta}_{\Gamma}^{11}$, amostras geradas considerando uma **Distribuição Normal** com $\mu = 3,25$ e $\sigma^2 = 1$ para $\Gamma = 0,33$ e o **BCD**

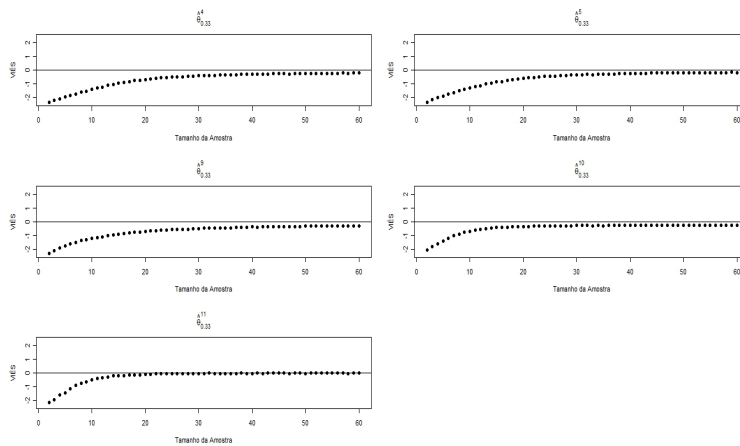


Figura 130 Viés para cada tamanho amostral e os estimadores $\hat{\theta}_{\Gamma}^4$, $\hat{\theta}_{\Gamma}^5$, $\hat{\theta}_{\Gamma}^9$, $\hat{\theta}_{\Gamma}^{10}$ e $\hat{\theta}_{\Gamma}^{11}$, amostras geradas considerando uma **Distribuição Normal** com $\mu = 3,25$ e $\sigma^2 = 1$ para $\Gamma = 0,33$ e o **BCD**

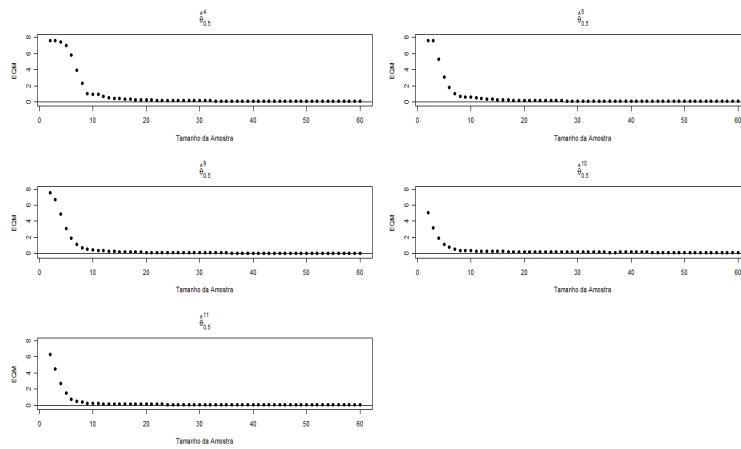


Figura 131 EQM para cada tamanho amostral e os estimadores $\hat{\theta}_{\Gamma}^4$, $\hat{\theta}_{\Gamma}^5$, $\hat{\theta}_{\Gamma}^9$, $\hat{\theta}_{\Gamma}^{10}$ e $\hat{\theta}_{\Gamma}^{11}$, amostras geradas considerando uma **Distribuição Normal** com $\mu = 3,25$ e $\sigma^2 = 1$ para $\Gamma = 0,50$ e o **BCD**

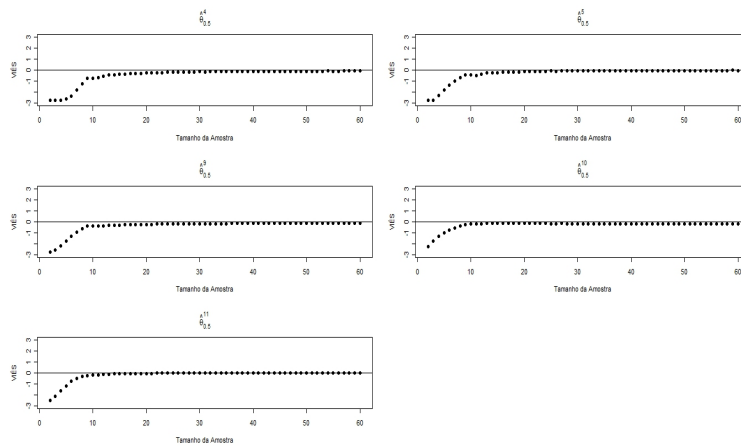


Figura 132 Viés para cada tamanho amostral e os estimadores $\hat{\theta}_{\Gamma}^4$, $\hat{\theta}_{\Gamma}^5$, $\hat{\theta}_{\Gamma}^9$, $\hat{\theta}_{\Gamma}^{10}$ e $\hat{\theta}_{\Gamma}^{11}$, amostras geradas considerando uma **Distribuição Normal** com $\mu = 3,25$ e $\sigma^2 = 1$ para $\Gamma = 0,50$ e o **BCD**

APÊNDICE C

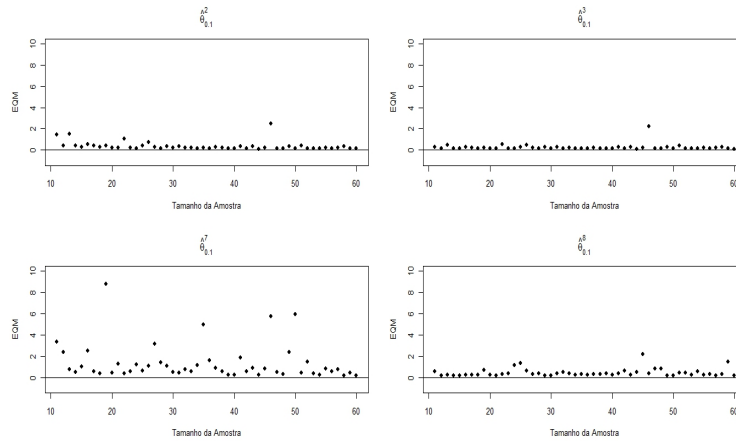


Figura 133 EQM para cada tamanho amostral e os estimadores $\hat{\theta}_{\Gamma}^2$, $\hat{\theta}_{\Gamma}^3$, $\hat{\theta}_{\Gamma}^7$ e $\hat{\theta}_{\Gamma}^8$, amostras geradas considerando uma **Distribuição Logística** com $\mu = 2,25$ e $\sigma^2 = 1$ para $\Gamma = 0,10$ e o **BCD**

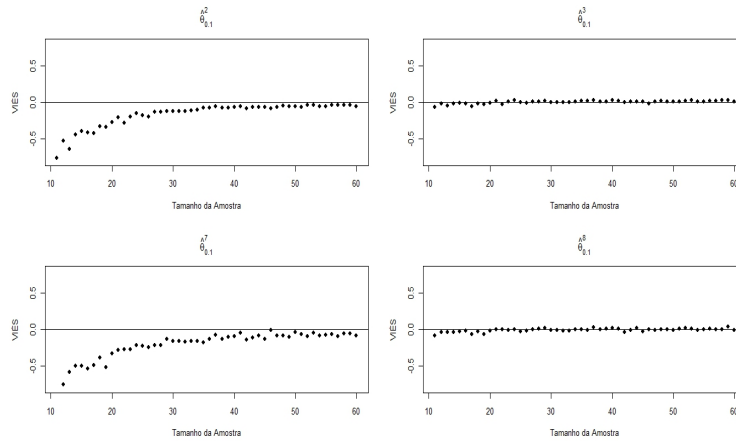


Figura 134 Viés para cada tamanho amostral e os estimadores $\hat{\theta}_{\Gamma}^2$, $\hat{\theta}_{\Gamma}^3$, $\hat{\theta}_{\Gamma}^7$ e $\hat{\theta}_{\Gamma}^8$, amostras geradas considerando uma **Distribuição Logística** com $\mu = 2,25$ e $\sigma^2 = 1$ para $\Gamma = 0,10$ e o **BCD**

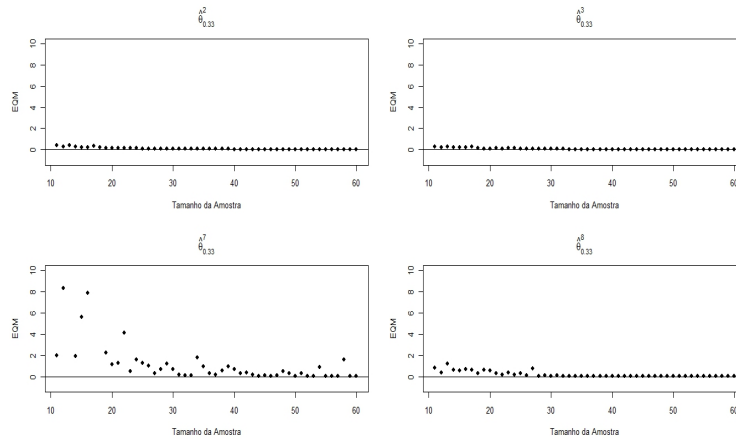


Figura 135 EQM para cada tamanho amostral e os estimadores $\hat{\theta}_{\Gamma}^2$, $\hat{\theta}_{\Gamma}^3$, $\hat{\theta}_{\Gamma}^7$ e $\hat{\theta}_{\Gamma}^8$, amostras geradas considerando uma **Distribuição Logística** com $\mu = 2,25$ e $\sigma^2 = 1$ para $\Gamma = 0,33$ e o **BCD**

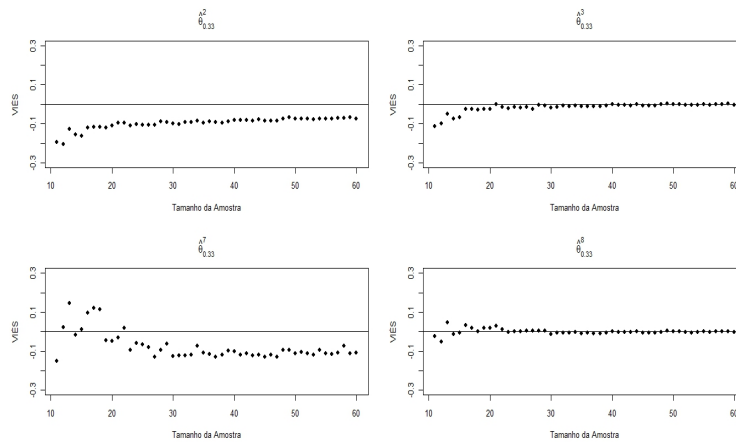


Figura 136 Viés para cada tamanho amostral e os estimadores $\hat{\theta}_{\Gamma}^2$, $\hat{\theta}_{\Gamma}^3$, $\hat{\theta}_{\Gamma}^7$ e $\hat{\theta}_{\Gamma}^8$, amostras geradas considerando uma **Distribuição Logística** com $\mu = 2,25$ e $\sigma^2 = 1$ para $\Gamma = 0,33$ e o **BCD**

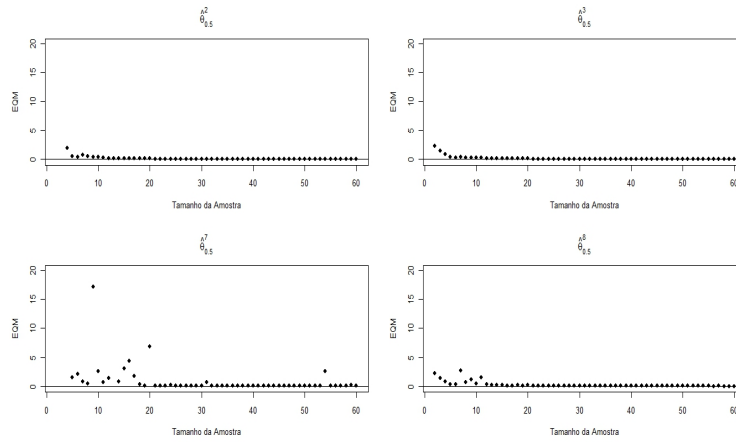


Figura 137 EQM para cada tamanho amostral e os estimadores $\hat{\theta}_{\Gamma}^2$, $\hat{\theta}_{\Gamma}^3$, $\hat{\theta}_{\Gamma}^7$ e $\hat{\theta}_{\Gamma}^8$, amostras geradas considerando uma **Distribuição Logística** com $\mu = 2,25$ e $\sigma^2 = 1$ para $\Gamma = 0,50$ e o **BCD**

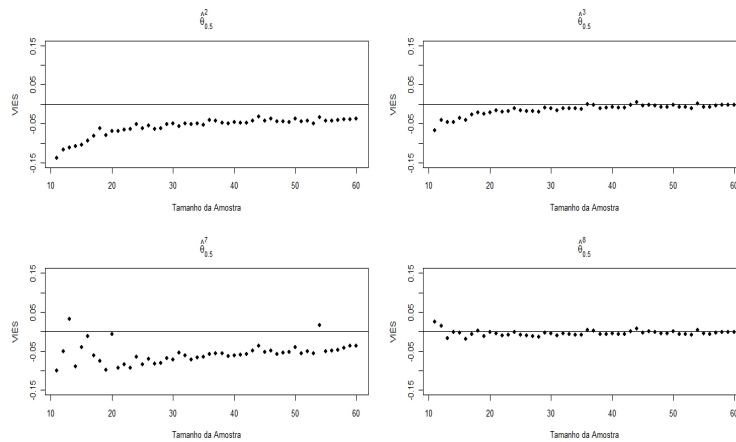


Figura 138 Viés para cada tamanho amostral e os estimadores $\hat{\theta}_{\Gamma}^2$, $\hat{\theta}_{\Gamma}^3$, $\hat{\theta}_{\Gamma}^7$ e $\hat{\theta}_{\Gamma}^8$, amostras geradas considerando uma **Distribuição Logística** com $\mu = 2,25$ e $\sigma^2 = 1$ para $\Gamma = 0,50$ e o **BCD**

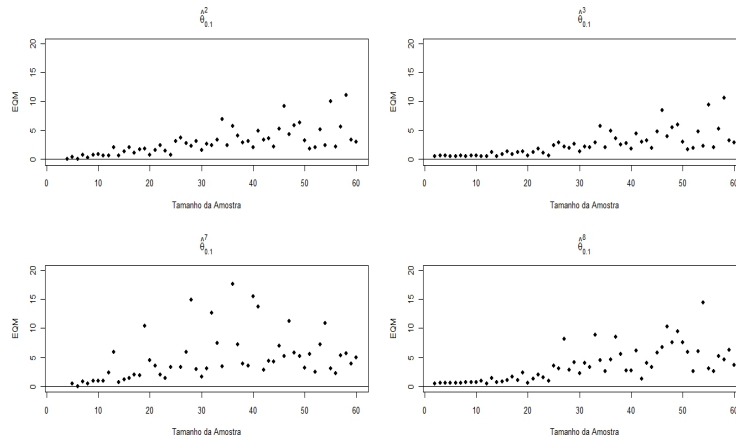


Figura 139 EQM para cada tamanho amostral e os estimadores $\hat{\theta}_{\Gamma}^2$, $\hat{\theta}_{\Gamma}^3$, $\hat{\theta}_{\Gamma}^7$ e $\hat{\theta}_{\Gamma}^8$, amostras geradas considerando uma **Distribuição Logística** com $\mu = 1,25$ e $\sigma^2 = 1$ para $\Gamma = 0,10$ e o **BCD**

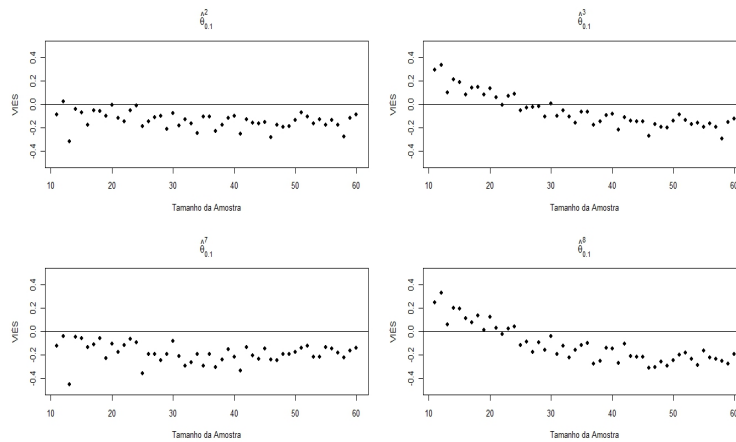


Figura 140 Viés para cada tamanho amostral e os estimadores $\hat{\theta}_{\Gamma}^2$, $\hat{\theta}_{\Gamma}^3$, $\hat{\theta}_{\Gamma}^7$ e $\hat{\theta}_{\Gamma}^8$, amostras geradas considerando uma **Distribuição Logística** com $\mu = 1,25$ e $\sigma^2 = 1$ para $\Gamma = 0,10$ e o **BCD**

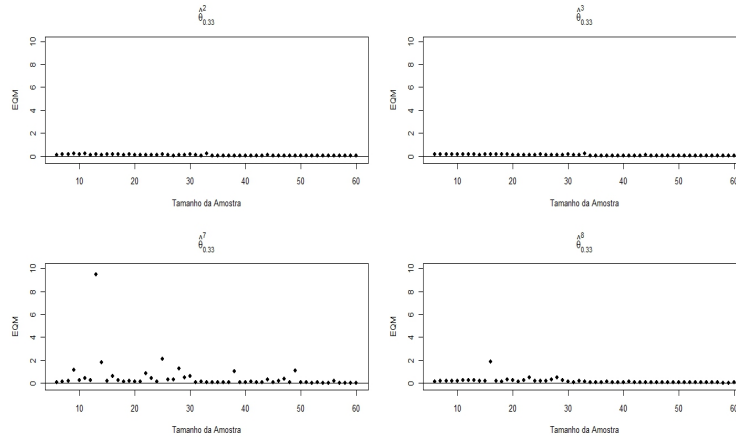


Figura 141 EQM para cada tamanho amostral e os estimadores $\hat{\theta}_{\Gamma}^2$, $\hat{\theta}_{\Gamma}^3$, $\hat{\theta}_{\Gamma}^7$ e $\hat{\theta}_{\Gamma}^8$, amostras geradas considerando uma **Distribuição Logística** com $\mu = 1,25$ e $\sigma^2 = 1$ para $\Gamma = 0,33$ e o **BCD**

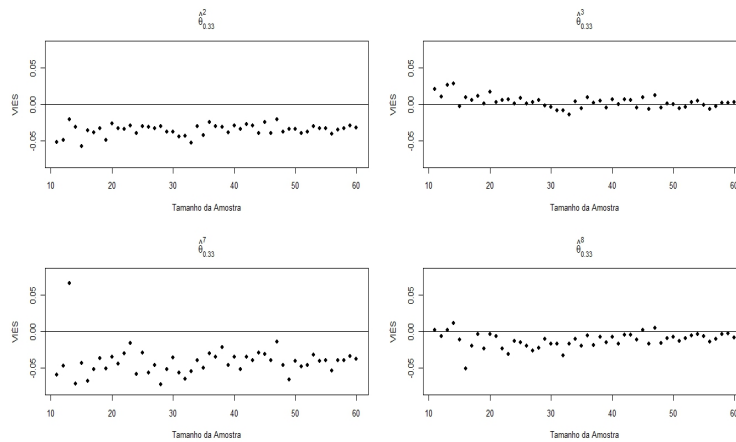


Figura 142 Viés para cada tamanho amostral e os estimadores $\hat{\theta}_{\Gamma}^2$, $\hat{\theta}_{\Gamma}^3$, $\hat{\theta}_{\Gamma}^7$ e $\hat{\theta}_{\Gamma}^8$, amostras geradas considerando uma **Distribuição Logística** com $\mu = 1,25$ e $\sigma^2 = 1$ para $\Gamma = 0,33$ e o **BCD**

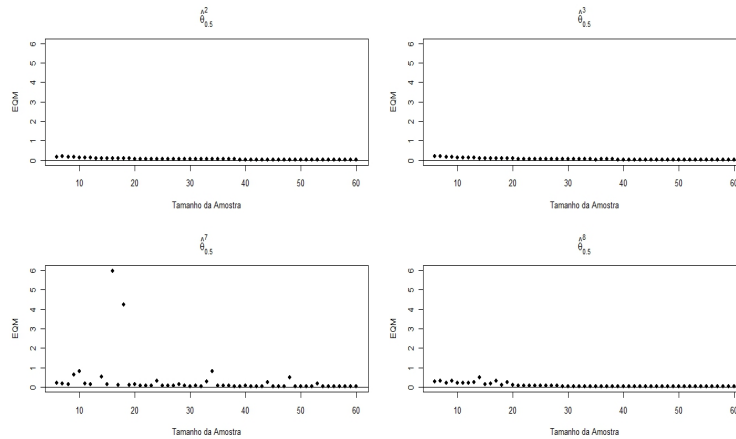


Figura 143 EQM para cada tamanho amostral e os estimadores $\hat{\theta}_{\Gamma}^2$, $\hat{\theta}_{\Gamma}^3$, $\hat{\theta}_{\Gamma}^7$ e $\hat{\theta}_{\Gamma}^8$, amostras geradas considerando uma **Distribuição Logística** com $\mu = 1,25$ e $\sigma^2 = 1$ para $\Gamma = 0,50$ e o **BCD**

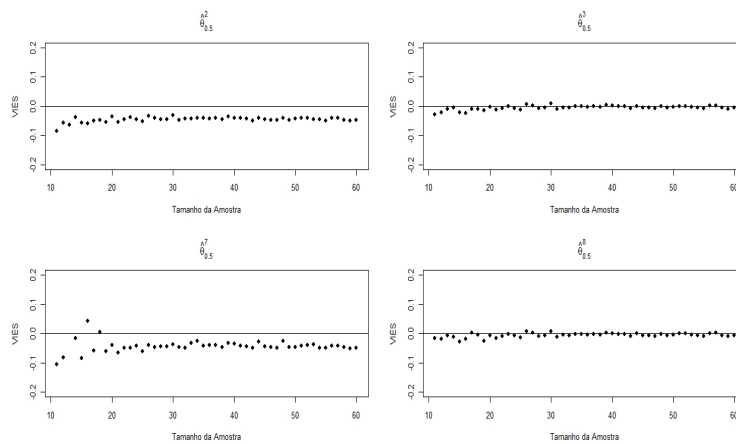


Figura 144 Viés para cada tamanho amostral e os estimadores $\hat{\theta}_{\Gamma}^2$, $\hat{\theta}_{\Gamma}^3$, $\hat{\theta}_{\Gamma}^7$ e $\hat{\theta}_{\Gamma}^8$, amostras geradas considerando uma **Distribuição Logística** com $\mu = 1,25$ e $\sigma^2 = 1$ para $\Gamma = 0,50$ e o **BCD**

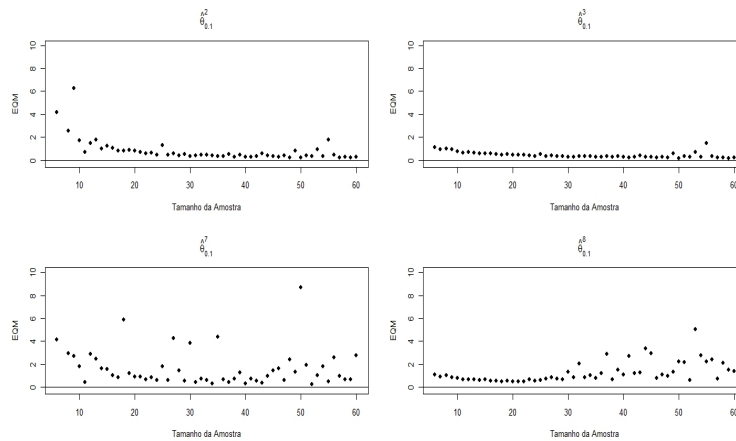


Figura 145 EQM para cada tamanho amostral e os estimadores $\hat{\theta}_{\Gamma}^2$, $\hat{\theta}_{\Gamma}^3$, $\hat{\theta}_{\Gamma}^7$ e $\hat{\theta}_{\Gamma}^8$, amostras geradas considerando uma **Distribuição Logística** com $\mu = 3,25$ e $\sigma^2 = 1$ para $\Gamma = 0,10$ e o **BCD**

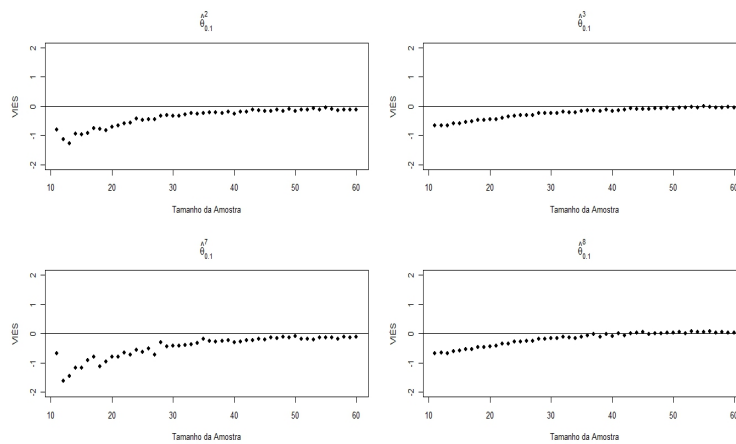


Figura 146 Viés para cada tamanho amostral e os estimadores $\hat{\theta}_{\Gamma}^2$, $\hat{\theta}_{\Gamma}^3$, $\hat{\theta}_{\Gamma}^7$ e $\hat{\theta}_{\Gamma}^8$, amostras geradas considerando uma **Distribuição Logística** com $\mu = 3,25$ e $\sigma^2 = 1$ para $\Gamma = 0,10$ e o **BCD**

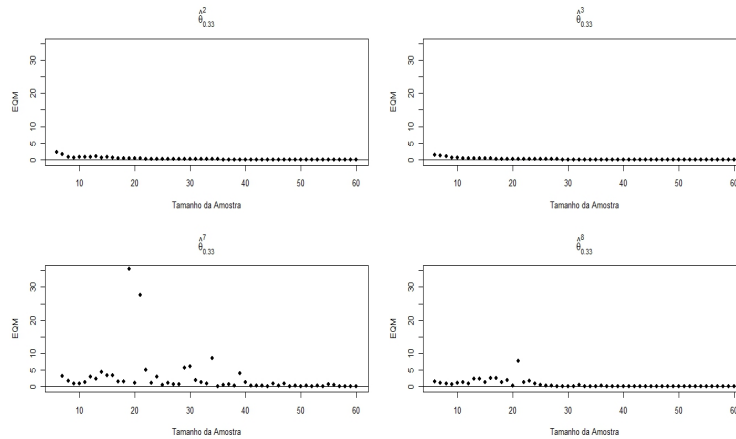


Figura 147 EQM para cada tamanho amostral e os estimadores $\hat{\theta}_{\Gamma}^2$, $\hat{\theta}_{\Gamma}^3$, $\hat{\theta}_{\Gamma}^7$ e $\hat{\theta}_{\Gamma}^8$, amostras geradas considerando uma **Distribuição Logística** com $\mu = 3,25$ e $\sigma^2 = 1$ para $\Gamma = 0,33$ e o **BCD**

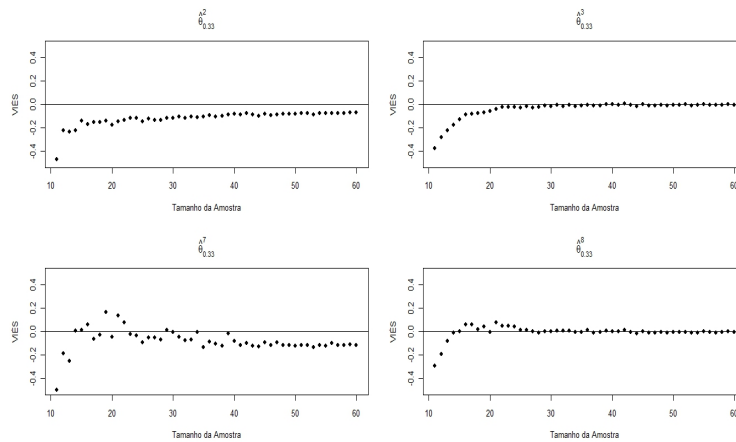


Figura 148 Viés para cada tamanho amostral e os estimadores $\hat{\theta}_{\Gamma}^2$, $\hat{\theta}_{\Gamma}^3$, $\hat{\theta}_{\Gamma}^7$ e $\hat{\theta}_{\Gamma}^8$, amostras geradas considerando uma **Distribuição Logística** com $\mu = 3,25$ e $\sigma^2 = 1$ para $\Gamma = 0,33$ e o **BCD**

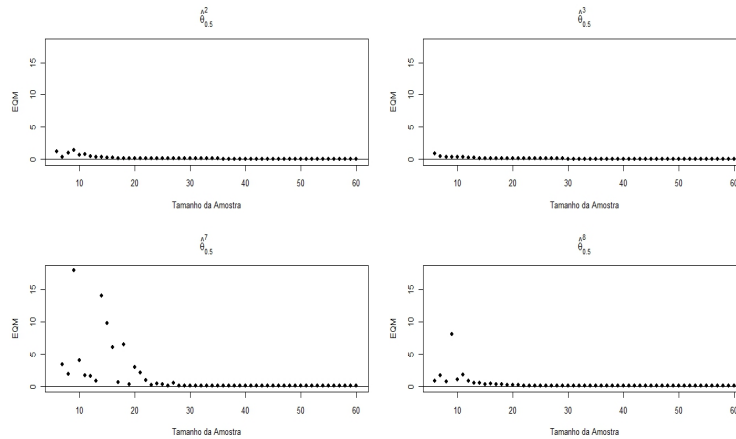


Figura 149 EQM para cada tamanho amostral e os estimadores $\hat{\theta}_{\Gamma}^2$, $\hat{\theta}_{\Gamma}^3$, $\hat{\theta}_{\Gamma}^7$ e $\hat{\theta}_{\Gamma}^8$, amostras geradas considerando uma **Distribuição Logística** com $\mu = 3,25$ e $\sigma^2 = 1$ para $\Gamma = 0,50$ e o **BCD**

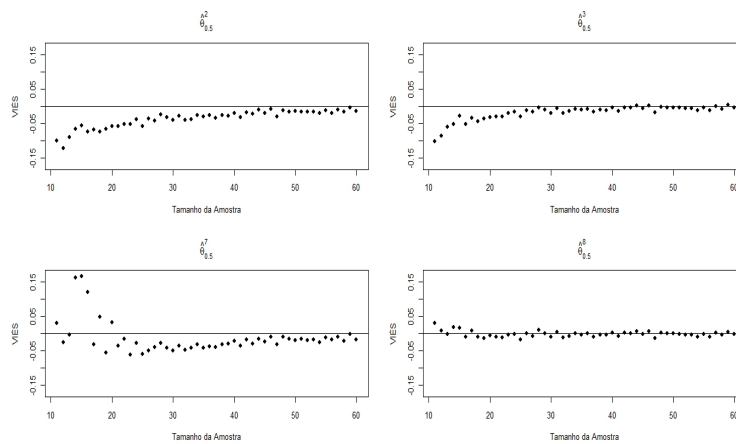


Figura 150 Viés para cada tamanho amostral e os estimadores $\hat{\theta}_{\Gamma}^2$, $\hat{\theta}_{\Gamma}^3$, $\hat{\theta}_{\Gamma}^7$ e $\hat{\theta}_{\Gamma}^8$, amostras geradas considerando uma **Distribuição Logística** com $\mu = 3,25$ e $\sigma^2 = 1$ para $\Gamma = 0,50$ e o **BCD**

APÊNDICE D

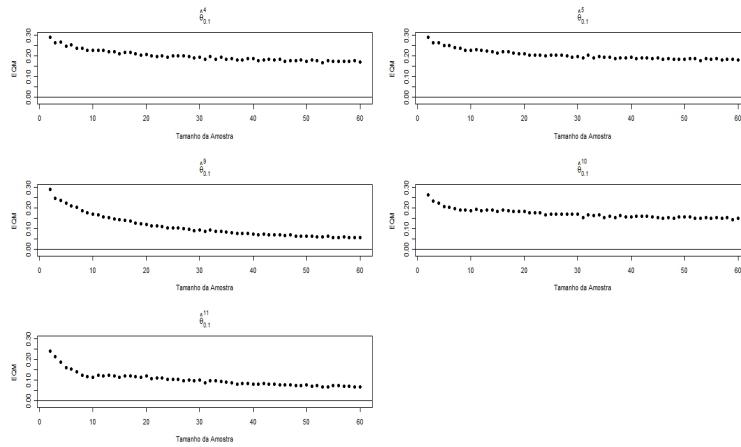


Figura 151 EQM para cada tamanho amostral e os estimadores $\hat{\theta}_{\Gamma}^4$, $\hat{\theta}_{\Gamma}^5$, $\hat{\theta}_{\Gamma}^9$, $\hat{\theta}_{\Gamma}^{10}$ e $\hat{\theta}_{\Gamma}^{11}$, amostras geradas considerando uma **Distribuição Logística** com $\mu = 2,25$ e $\sigma^2 = 1$ para $\Gamma = 0,10$ e o **BCD**

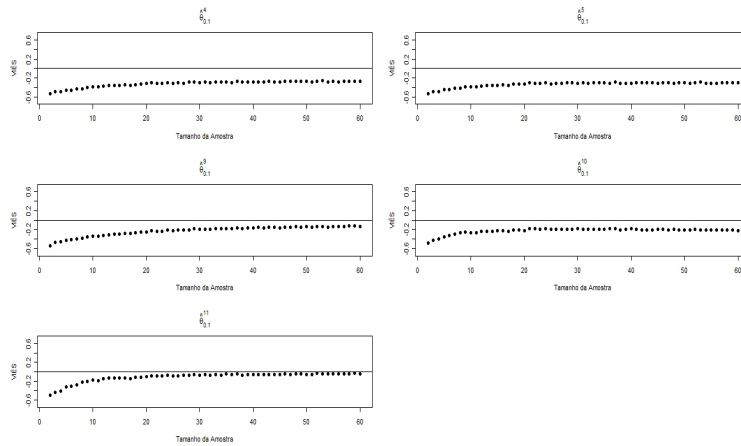


Figura 152 Viés para cada tamanho amostral e os estimadores $\hat{\theta}_{\Gamma}^4$, $\hat{\theta}_{\Gamma}^5$, $\hat{\theta}_{\Gamma}^9$, $\hat{\theta}_{\Gamma}^{10}$ e $\hat{\theta}_{\Gamma}^{11}$, amostras geradas considerando uma **Distribuição Logística** com $\mu = 2,25$ e $\sigma^2 = 1$ para $\Gamma = 0,10$ e o **BCD**

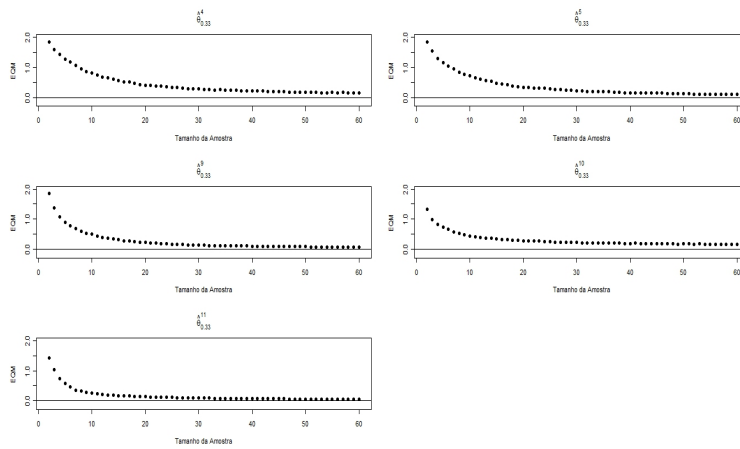


Figura 153 EQM para cada tamanho amostral e os estimadores $\hat{\theta}_{\Gamma}^4$, $\hat{\theta}_{\Gamma}^5$, $\hat{\theta}_{\Gamma}^9$, $\hat{\theta}_{\Gamma}^{10}$ e $\hat{\theta}_{\Gamma}^{11}$, amostras geradas considerando uma **Distribuição Logística** com $\mu = 2,25$ e $\sigma^2 = 1$ para $\Gamma = 0,33$ e o **BCD**

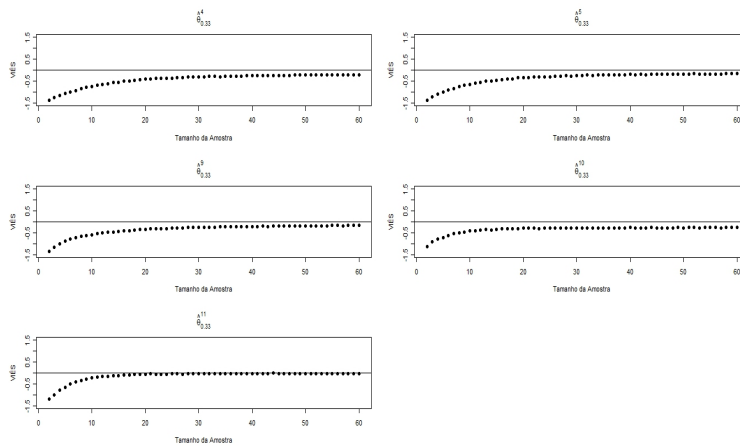


Figura 154 Viés para cada tamanho amostral e os estimadores $\hat{\theta}_{\Gamma}^4$, $\hat{\theta}_{\Gamma}^5$, $\hat{\theta}_{\Gamma}^9$, $\hat{\theta}_{\Gamma}^{10}$ e $\hat{\theta}_{\Gamma}^{11}$, amostras geradas considerando uma **Distribuição Logística** com $\mu = 2,25$ e $\sigma^2 = 1$ para $\Gamma = 0,33$ e o **BCD**

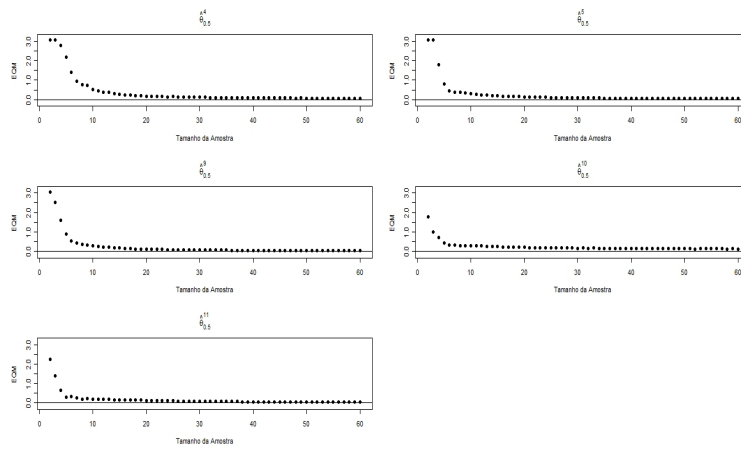


Figura 155 EQM para cada tamanho amostral e os estimadores $\hat{\theta}_{\Gamma}^4$, $\hat{\theta}_{\Gamma}^5$, $\hat{\theta}_{\Gamma}^9$, $\hat{\theta}_{\Gamma}^{10}$ e $\hat{\theta}_{\Gamma}^{11}$, amostras geradas considerando uma **Distribuição Logística** com $\mu = 2,25$ e $\sigma^2 = 1$ para $\Gamma = 0,50$ e o **BCD**

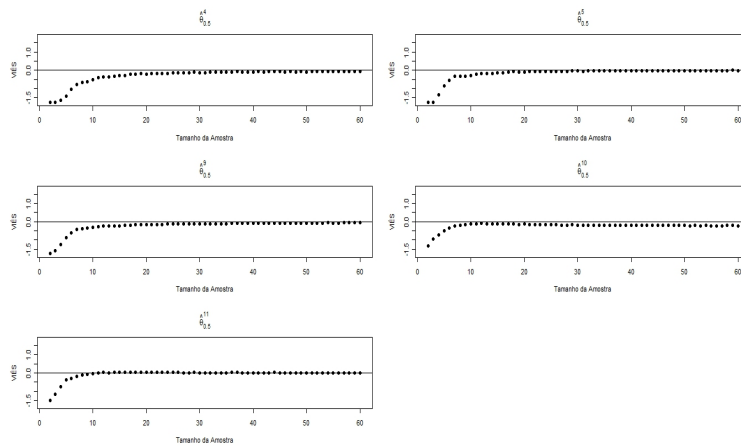


Figura 156 Viés para cada tamanho amostral e os estimadores $\hat{\theta}_{\Gamma}^4$, $\hat{\theta}_{\Gamma}^5$, $\hat{\theta}_{\Gamma}^9$, $\hat{\theta}_{\Gamma}^{10}$ e $\hat{\theta}_{\Gamma}^{11}$, amostras geradas considerando uma **Distribuição Logística** com $\mu = 2,25$ e $\sigma^2 = 1$ para $\Gamma = 0,50$ e o **BCD**

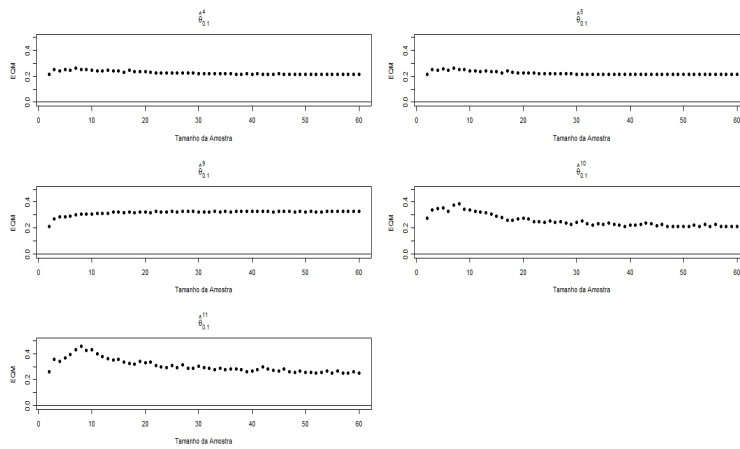


Figura 157 EQM para cada tamanho amostral e os estimadores $\hat{\theta}_{\Gamma}^4$, $\hat{\theta}_{\Gamma}^5$, $\hat{\theta}_{\Gamma}^9$, $\hat{\theta}_{\Gamma}^{10}$ e $\hat{\theta}_{\Gamma}^{11}$, amostras geradas considerando uma **Distribuição Logística** com $\mu = 1,25$ e $\sigma^2 = 1$ para $\Gamma = 0,10$ e o **BCD**

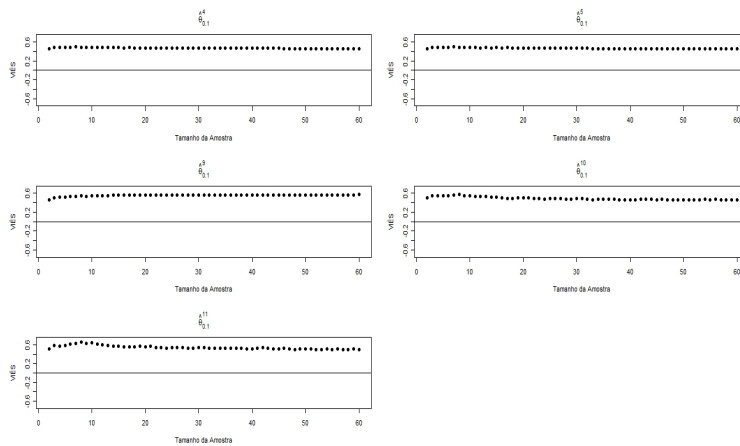


Figura 158 Viés para cada tamanho amostral e os estimadores $\hat{\theta}_{\Gamma}^4$, $\hat{\theta}_{\Gamma}^5$, $\hat{\theta}_{\Gamma}^9$, $\hat{\theta}_{\Gamma}^{10}$ e $\hat{\theta}_{\Gamma}^{11}$, amostras geradas considerando uma **Distribuição Logística** com $\mu = 1,25$ e $\sigma^2 = 1$ para $\Gamma = 0,10$ e o **BCD**

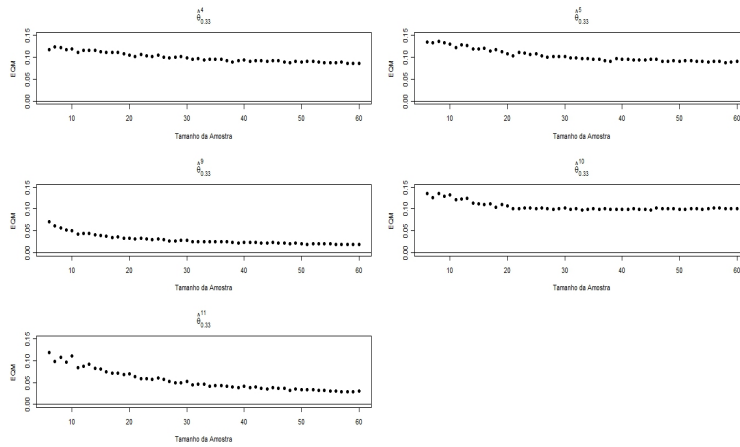


Figura 159 EQM para cada tamanho amostral e os estimadores $\hat{\theta}_{\Gamma}^4$, $\hat{\theta}_{\Gamma}^5$, $\hat{\theta}_{\Gamma}^9$, $\hat{\theta}_{\Gamma}^{10}$ e $\hat{\theta}_{\Gamma}^{11}$, amostras geradas considerando uma **Distribuição Logística** com $\mu = 1,25$ e $\sigma^2 = 1$ para $\Gamma = 0,33$ e o **BCD**

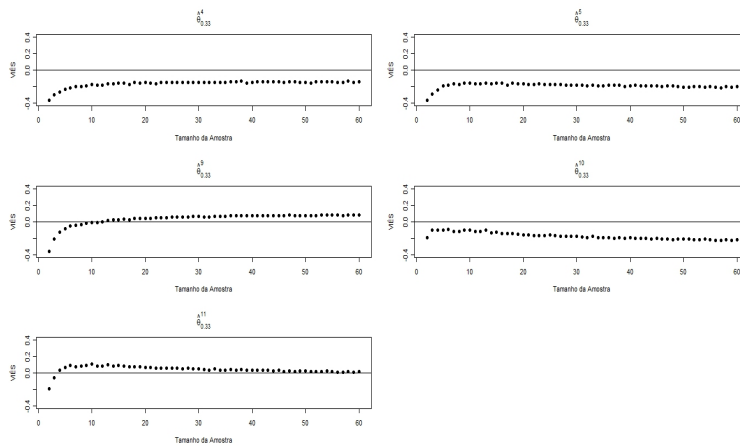


Figura 160 Viés para cada tamanho amostral e os estimadores $\hat{\theta}_{\Gamma}^4$, $\hat{\theta}_{\Gamma}^5$, $\hat{\theta}_{\Gamma}^9$, $\hat{\theta}_{\Gamma}^{10}$ e $\hat{\theta}_{\Gamma}^{11}$, amostras geradas considerando uma **Distribuição Logística** com $\mu = 1,25$ e $\sigma^2 = 1$ para $\Gamma = 0,33$ e o **BCD**

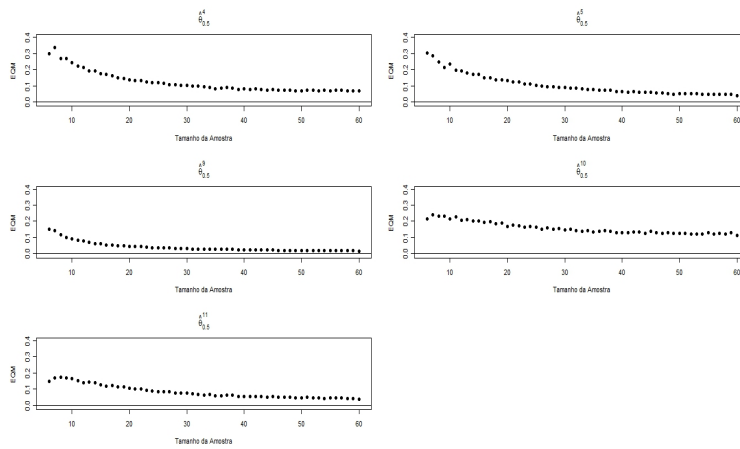


Figura 161 EQM para cada tamanho amostral e os estimadores $\hat{\theta}_{\Gamma}^4$, $\hat{\theta}_{\Gamma}^5$, $\hat{\theta}_{\Gamma}^9$, $\hat{\theta}_{\Gamma}^{10}$ e $\hat{\theta}_{\Gamma}^{11}$, amostras geradas considerando uma **Distribuição Logística** com $\mu = 1,25$ e $\sigma^2 = 1$ para $\Gamma = 0,50$ e o **BCD**

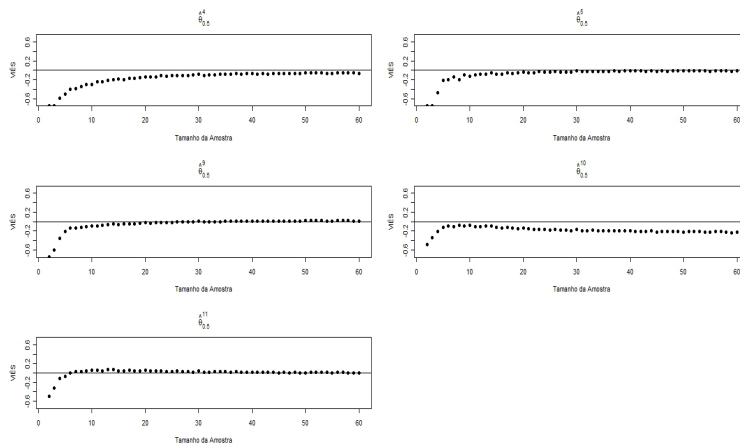


Figura 162 Viés para cada tamanho amostral e os estimadores $\hat{\theta}_{\Gamma}^4$, $\hat{\theta}_{\Gamma}^5$, $\hat{\theta}_{\Gamma}^9$, $\hat{\theta}_{\Gamma}^{10}$ e $\hat{\theta}_{\Gamma}^{11}$, amostras geradas considerando uma **Distribuição Logística** com $\mu = 1,25$ e $\sigma^2 = 1$ para $\Gamma = 0,50$ e o **BCD**

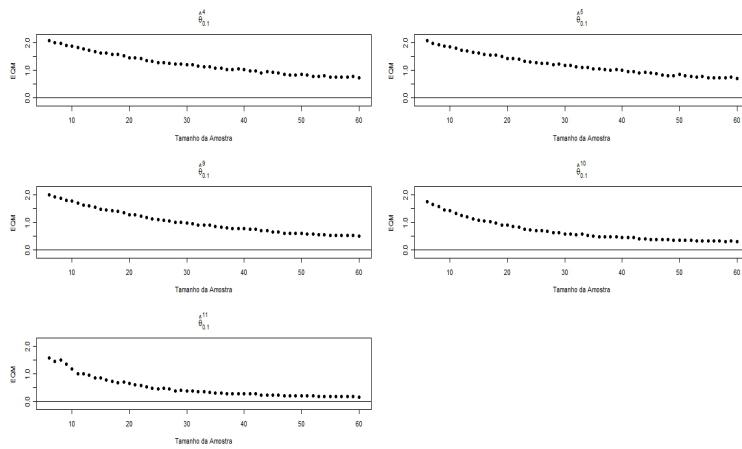


Figura 163 EQM para cada tamanho amostral e os estimadores $\hat{\theta}_\Gamma^4$, $\hat{\theta}_\Gamma^5$, $\hat{\theta}_\Gamma^9$, $\hat{\theta}_\Gamma^{10}$ e $\hat{\theta}_\Gamma^{11}$, amostras geradas considerando uma **Distribuição Logística** com $\mu = 3,25$ e $\sigma^2 = 1$ para $\Gamma = 0,10$ e o **BCD**

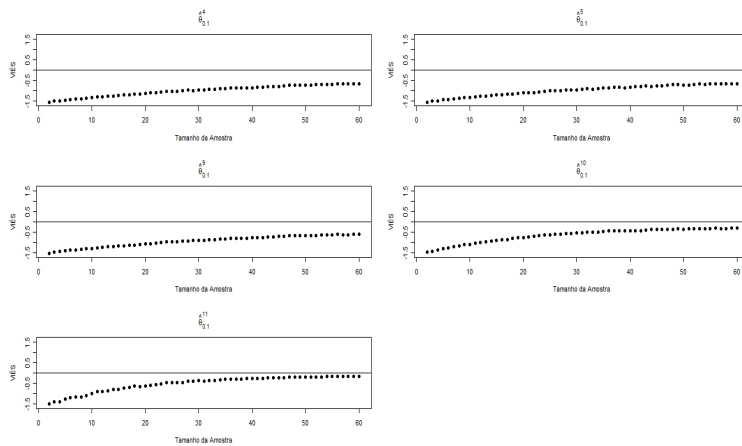


Figura 164 Viés para cada tamanho amostral e os estimadores $\hat{\theta}_\Gamma^4$, $\hat{\theta}_\Gamma^5$, $\hat{\theta}_\Gamma^9$, $\hat{\theta}_\Gamma^{10}$ e $\hat{\theta}_\Gamma^{11}$, amostras geradas considerando uma **Distribuição Logística** com $\mu = 3,25$ e $\sigma^2 = 1$ para $\Gamma = 0,10$ e o **BCD**

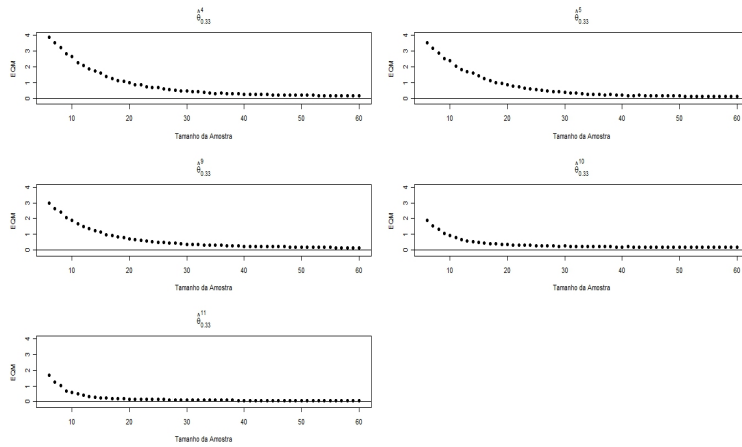


Figura 165 EQM para cada tamanho amostral e os estimadores $\hat{\theta}_{\Gamma}^4$, $\hat{\theta}_{\Gamma}^5$, $\hat{\theta}_{\Gamma}^9$, $\hat{\theta}_{\Gamma}^{10}$ e $\hat{\theta}_{\Gamma}^{11}$, amostras geradas considerando uma **Distribuição Logística** com $\mu = 3,25$ e $\sigma^2 = 1$ para $\Gamma = 0,33$ e o **BCD**

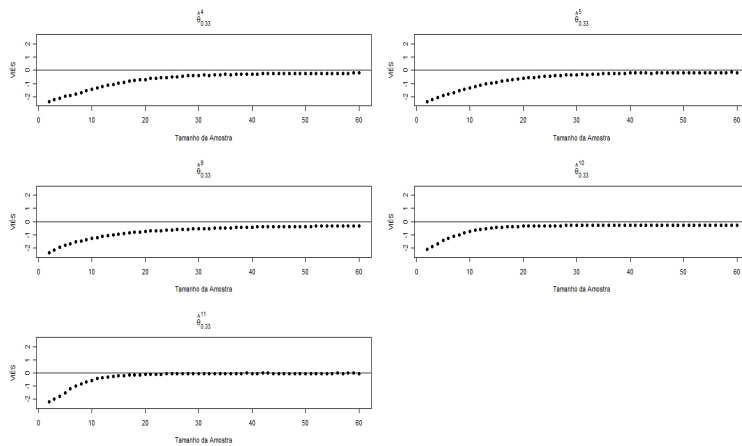


Figura 166 Viés para cada tamanho amostral e os estimadores $\hat{\theta}_{\Gamma}^4$, $\hat{\theta}_{\Gamma}^5$, $\hat{\theta}_{\Gamma}^9$, $\hat{\theta}_{\Gamma}^{10}$ e $\hat{\theta}_{\Gamma}^{11}$, amostras geradas considerando uma **Distribuição Logística** com $\mu = 3,25$ e $\sigma^2 = 1$ para $\Gamma = 0,33$ e o **BCD**

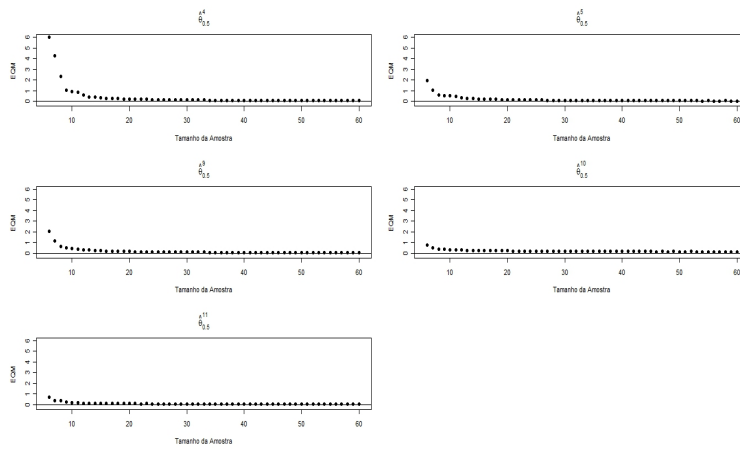


Figura 167 EQM para cada tamanho amostral e os estimadores $\hat{\theta}_{\Gamma}^4$, $\hat{\theta}_{\Gamma}^5$, $\hat{\theta}_{\Gamma}^9$, $\hat{\theta}_{\Gamma}^{10}$ e $\hat{\theta}_{\Gamma}^{11}$, amostras geradas considerando uma **Distribuição Logística** com $\mu = 3,25$ e $\sigma^2 = 1$ para $\Gamma = 0,50$ e o **BCD**

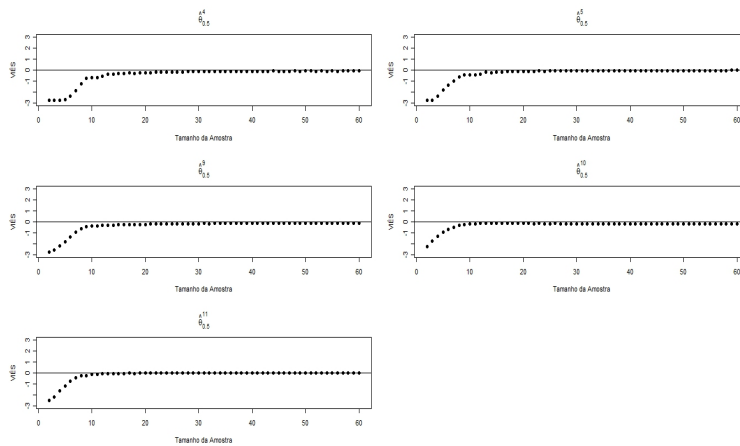


Figura 168 Viés para cada tamanho amostral e os estimadores $\hat{\theta}_{\Gamma}^4$, $\hat{\theta}_{\Gamma}^5$, $\hat{\theta}_{\Gamma}^9$, $\hat{\theta}_{\Gamma}^{10}$ e $\hat{\theta}_{\Gamma}^{11}$, amostras geradas considerando uma **Distribuição Logística** com $\mu = 3,25$ e $\sigma^2 = 1$ para $\Gamma = 0,50$ e o **BCD**

APÊNDICE E

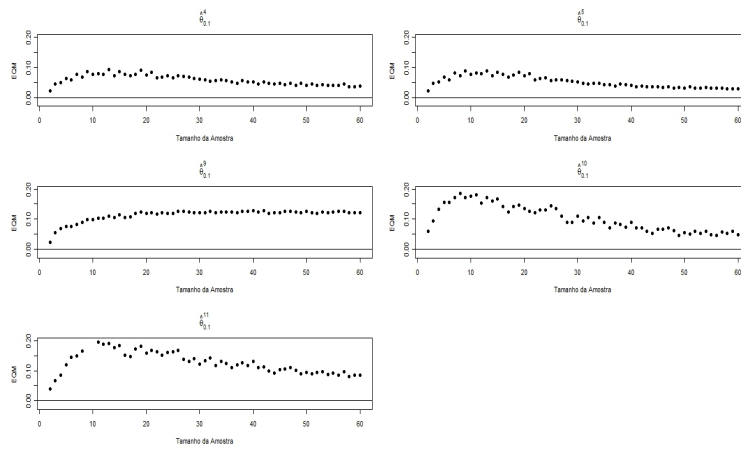


Figura 169 EQM para cada tamanho amostral e os estimadores $\hat{\theta}_{\Gamma}^4$, $\hat{\theta}_{\Gamma}^5$, $\hat{\theta}_{\Gamma}^9$, $\hat{\theta}_{\Gamma}^{10}$ e $\hat{\theta}_{\Gamma}^{11}$, amostras geradas considerando uma **Distribuição Exponencial** com parâmetro $\lambda = 0,30$, para $\Gamma = 0,10$ e o **BCD**

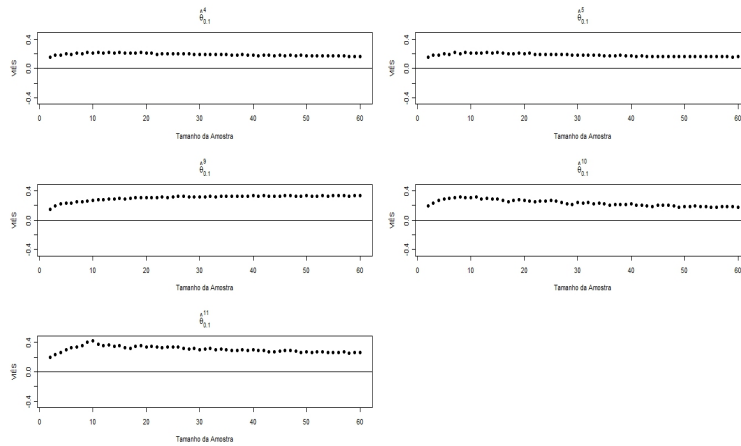


Figura 170 Viés para cada tamanho amostral e os estimadores $\hat{\theta}_{\Gamma}^4$, $\hat{\theta}_{\Gamma}^5$, $\hat{\theta}_{\Gamma}^9$, $\hat{\theta}_{\Gamma}^{10}$ e $\hat{\theta}_{\Gamma}^{11}$, amostras geradas considerando uma **Distribuição Exponencial** com parâmetro $\lambda = 0,30$, para $\Gamma = 0,10$ e o **BCD**

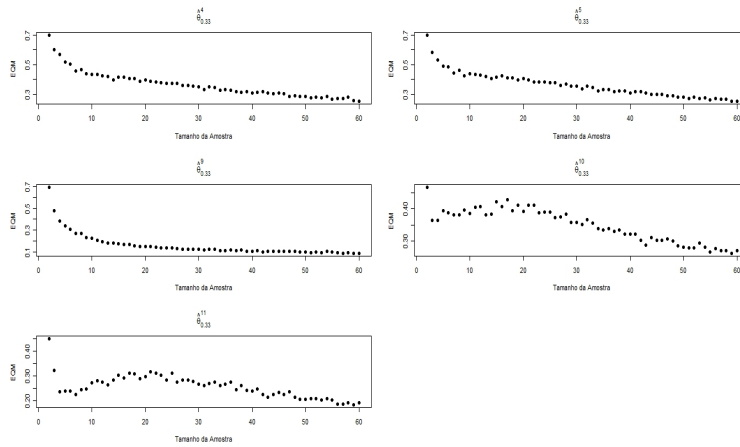


Figura 171 EQM para cada tamanho amostral e os estimadores $\hat{\theta}_{\Gamma}^4$, $\hat{\theta}_{\Gamma}^5$, $\hat{\theta}_{\Gamma}^9$, $\hat{\theta}_{\Gamma}^{10}$ e $\hat{\theta}_{\Gamma}^{11}$, amostras geradas considerando uma **Distribuição Exponencial** com parâmetro $\lambda = 0,30$, para $\Gamma = 0,33$ e o **BCD**

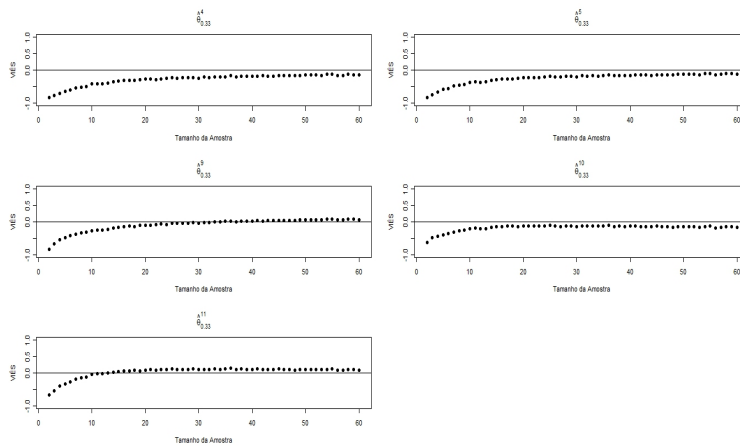


Figura 172 Viés para cada tamanho amostral e os estimadores $\hat{\theta}_{\Gamma}^4$, $\hat{\theta}_{\Gamma}^5$, $\hat{\theta}_{\Gamma}^9$, $\hat{\theta}_{\Gamma}^{10}$ e $\hat{\theta}_{\Gamma}^{11}$, amostras geradas considerando uma **Distribuição Exponencial** com parâmetro $\lambda = 0,30$, para $\Gamma = 0,33$ e o **BCD**

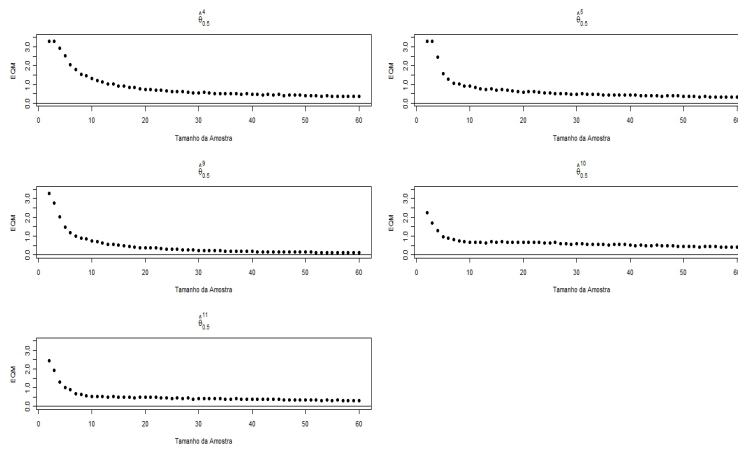


Figura 173 EQM para cada tamanho amostral e os estimadores $\hat{\theta}_{\Gamma}^4$, $\hat{\theta}_{\Gamma}^5$, $\hat{\theta}_{\Gamma}^9$, $\hat{\theta}_{\Gamma}^{10}$ e $\hat{\theta}_{\Gamma}^{11}$, amostras geradas considerando uma **Distribuição Exponencial** com parâmetro $\lambda = 0,30$, para $\Gamma = 0,50$ e o **BCD**

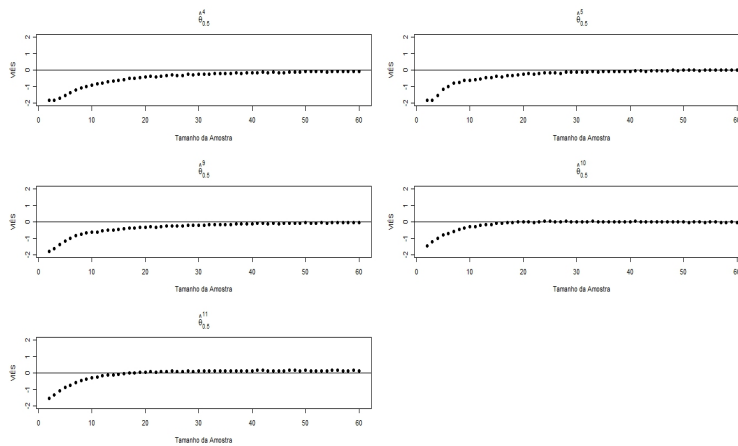


Figura 174 Viés para cada tamanho amostral e os estimadores $\hat{\theta}_{\Gamma}^4$, $\hat{\theta}_{\Gamma}^5$, $\hat{\theta}_{\Gamma}^9$, $\hat{\theta}_{\Gamma}^{10}$ e $\hat{\theta}_{\Gamma}^{11}$, amostras geradas considerando uma **Distribuição Exponencial** com parâmetro $\lambda = 0,30$, para $\Gamma = 0,50$ e o **BCD**

APÊNDICE F

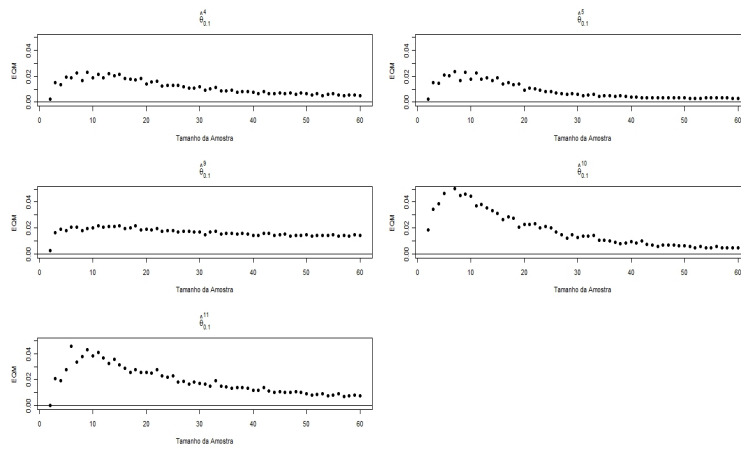


Figura 175 EQM para cada tamanho amostral e os estimadores $\hat{\theta}_{\Gamma}^4$, $\hat{\theta}_{\Gamma}^5$, $\hat{\theta}_{\Gamma}^9$, $\hat{\theta}_{\Gamma}^{10}$ e $\hat{\theta}_{\Gamma}^{11}$, amostras geradas considerando uma **Distribuição Gama** com parâmetros $\alpha = 3$ e $\lambda = 0,50$, para $\Gamma = 0,10$ e o **BCD**

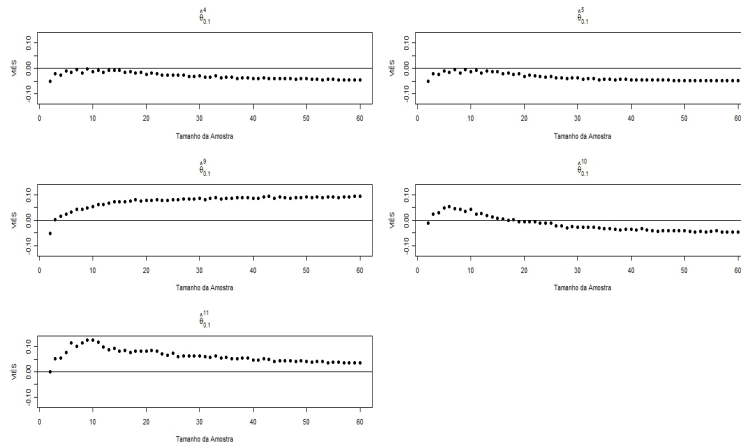


Figura 176 Viés para cada tamanho amostral e os estimadores $\hat{\theta}_{\Gamma}^4$, $\hat{\theta}_{\Gamma}^5$, $\hat{\theta}_{\Gamma}^9$, $\hat{\theta}_{\Gamma}^{10}$ e $\hat{\theta}_{\Gamma}^{11}$, amostras geradas considerando uma **Distribuição Gama** com parâmetros $\alpha = 3$ e $\lambda = 0,50$, para $\Gamma = 0,10$ e o **BCD**

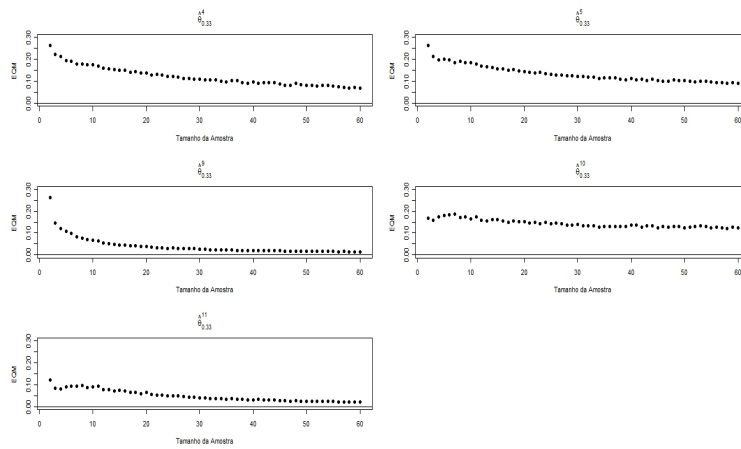


Figura 177 EQM para cada tamanho amostral e os estimadores $\hat{\theta}_{\Gamma}^4$, $\hat{\theta}_{\Gamma}^5$, $\hat{\theta}_{\Gamma}^9$, $\hat{\theta}_{\Gamma}^{10}$ e $\hat{\theta}_{\Gamma}^{11}$, amostras geradas considerando uma **Distribuição Gama** com parâmetros $\alpha = 3$ e $\lambda = 0,50$, para $\Gamma = 0,33$ e o **BCD**

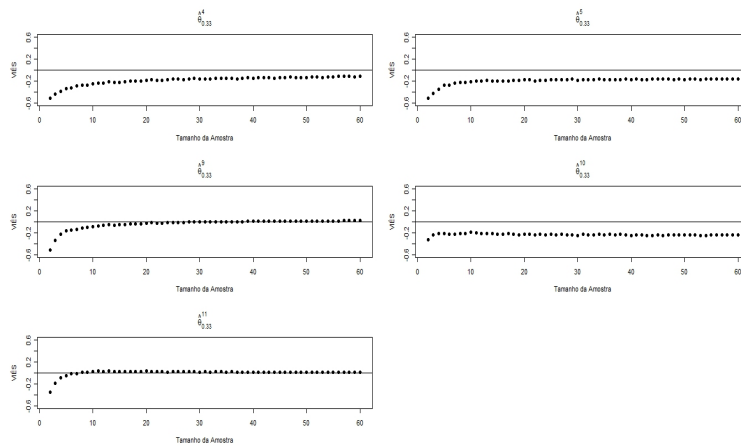


Figura 178 Viés para cada tamanho amostral e os estimadores $\hat{\theta}_{\Gamma}^4$, $\hat{\theta}_{\Gamma}^5$, $\hat{\theta}_{\Gamma}^9$, $\hat{\theta}_{\Gamma}^{10}$ e $\hat{\theta}_{\Gamma}^{11}$, amostras geradas considerando uma **Distribuição Gama** com parâmetros $\alpha = 3$ e $\lambda = 0,50$, para $\Gamma = 0,33$ e o **BCD**

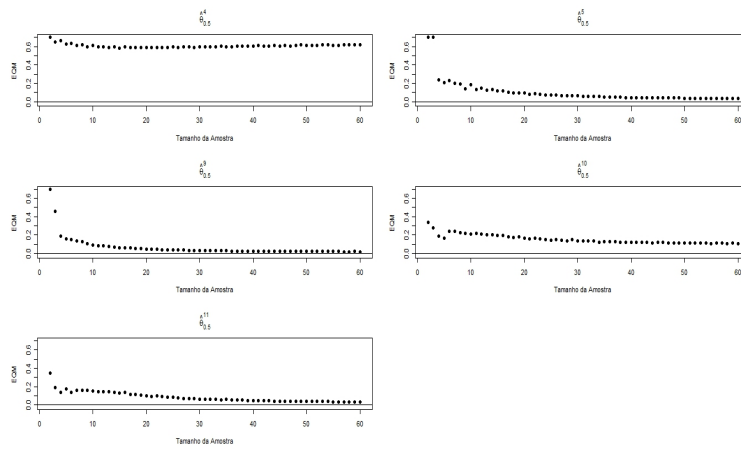


Figura 179 EQM para cada tamanho amostral e os estimadores $\hat{\theta}_{\Gamma}^4$, $\hat{\theta}_{\Gamma}^5$, $\hat{\theta}_{\Gamma}^9$, $\hat{\theta}_{\Gamma}^{10}$ e $\hat{\theta}_{\Gamma}^{11}$, amostras geradas considerando uma **Distribuição Gama** com parâmetros $\alpha = 3$ e $\lambda = 0,50$, para $\Gamma = 0,50$ e o **BCD**

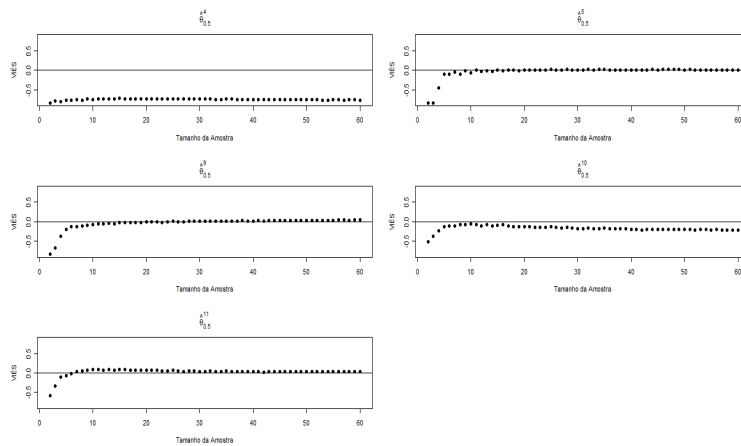


Figura 180 Viés para cada tamanho amostral e os estimadores $\hat{\theta}_{\Gamma}^4$, $\hat{\theta}_{\Gamma}^5$, $\hat{\theta}_{\Gamma}^9$, $\hat{\theta}_{\Gamma}^{10}$ e $\hat{\theta}_{\Gamma}^{11}$, amostras geradas considerando uma **Distribuição Gama** com parâmetros $\alpha = 3$ e $\lambda = 0,50$, para $\Gamma = 0,50$ e o **BCD**

APÊNDICE G

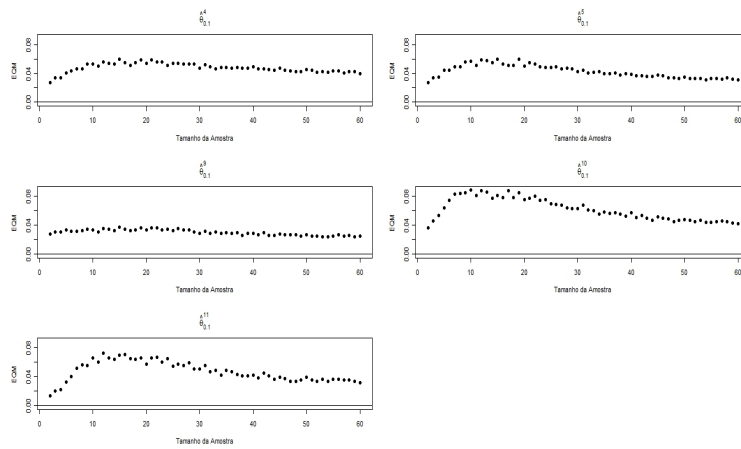


Figura 181 EQM para cada tamanho amostral e os estimadores $\hat{\theta}_{\Gamma}^4$, $\hat{\theta}_{\Gamma}^5$, $\hat{\theta}_{\Gamma}^9$, $\hat{\theta}_{\Gamma}^{10}$ e $\hat{\theta}_{\Gamma}^{11}$, amostras geradas considerando uma **Distribuição Gama** com parâmetros $\alpha = 2$ e $\lambda = 1,25$, para $\Gamma = 0,10$ e o **BCD**

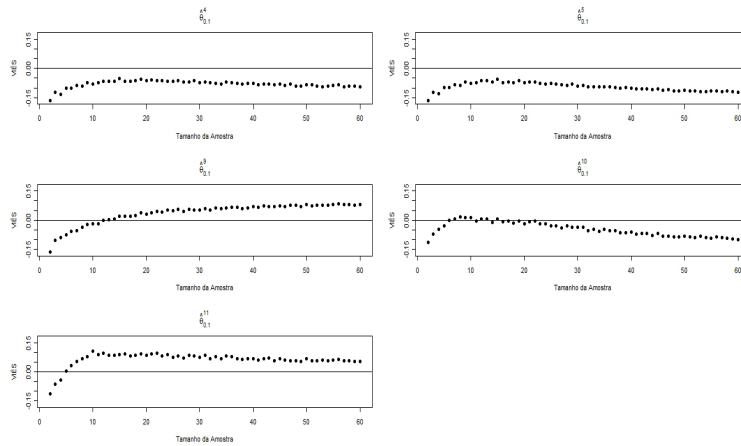


Figura 182 Viés para cada tamanho amostral e os estimadores $\hat{\theta}_{\Gamma}^4$, $\hat{\theta}_{\Gamma}^5$, $\hat{\theta}_{\Gamma}^9$, $\hat{\theta}_{\Gamma}^{10}$ e $\hat{\theta}_{\Gamma}^{11}$, amostras geradas considerando uma **Distribuição Gama** com parâmetros $\alpha = 2$ e $\lambda = 1,25$, para $\Gamma = 0,10$ e o **BCD**

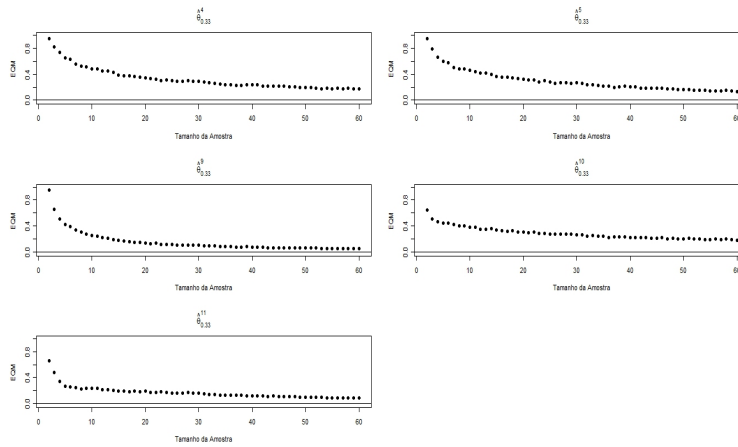


Figura 183 EQM para cada tamanho amostral e os estimadores $\hat{\theta}_{\Gamma}^4$, $\hat{\theta}_{\Gamma}^5$, $\hat{\theta}_{\Gamma}^9$, $\hat{\theta}_{\Gamma}^{10}$ e $\hat{\theta}_{\Gamma}^{11}$, amostras geradas considerando uma **Distribuição Gama** com parâmetros $\alpha = 2$ e $\lambda = 1,25$, para $\Gamma = 0,33$ e o **BCD**

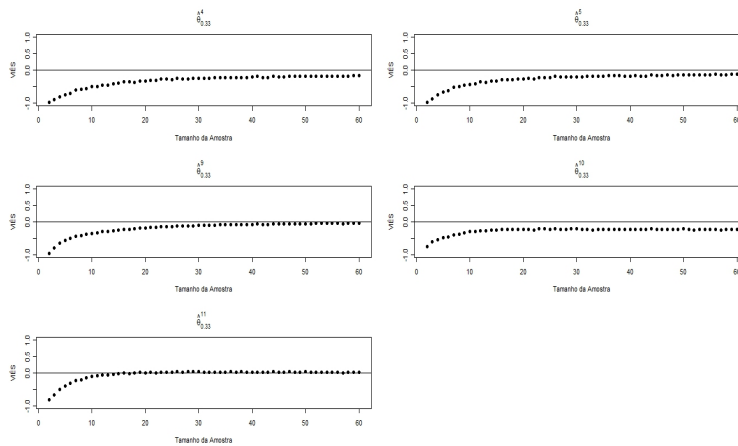


Figura 184 Viés para cada tamanho amostral e os estimadores $\hat{\theta}_{\Gamma}^4$, $\hat{\theta}_{\Gamma}^5$, $\hat{\theta}_{\Gamma}^9$, $\hat{\theta}_{\Gamma}^{10}$ e $\hat{\theta}_{\Gamma}^{11}$, amostras geradas considerando uma **Distribuição Gama** com parâmetros $\alpha = 2$ e $\lambda = 1,25$, para $\Gamma = 0,33$ e o **BCD**

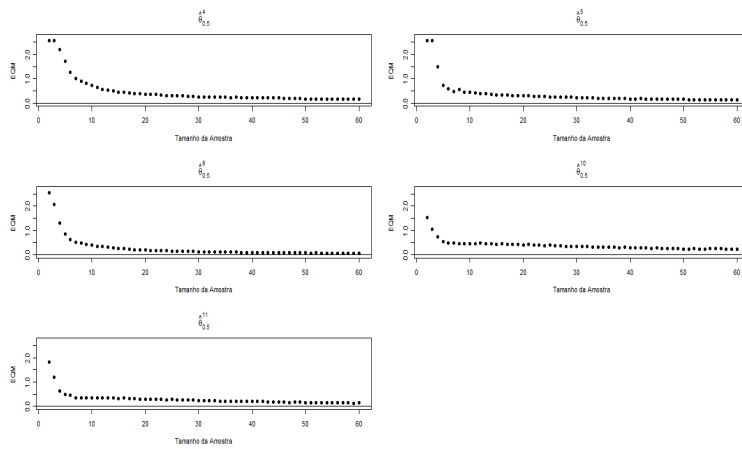


Figura 185 EQM para cada tamanho amostral e os estimadores $\hat{\theta}_{\Gamma}^4$, $\hat{\theta}_{\Gamma}^5$, $\hat{\theta}_{\Gamma}^9$, $\hat{\theta}_{\Gamma}^{10}$ e $\hat{\theta}_{\Gamma}^{11}$, amostras geradas considerando uma **Distribuição Gama** com parâmetros $\alpha = 2$ e $\lambda = 1,25$, para $\Gamma = 0,50$ e o **BCD**

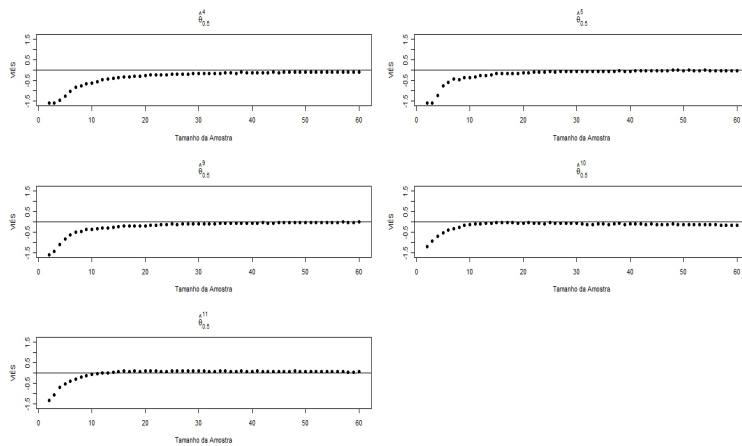


Figura 186 Viés para cada tamanho amostral e os estimadores $\hat{\theta}_{\Gamma}^4$, $\hat{\theta}_{\Gamma}^5$, $\hat{\theta}_{\Gamma}^9$, $\hat{\theta}_{\Gamma}^{10}$ e $\hat{\theta}_{\Gamma}^{11}$, amostras geradas considerando uma **Distribuição Gama** com parâmetros $\alpha = 2$ e $\lambda = 1,25$, para $\Gamma = 0,50$ e o **BCD**

APÊNDICE H

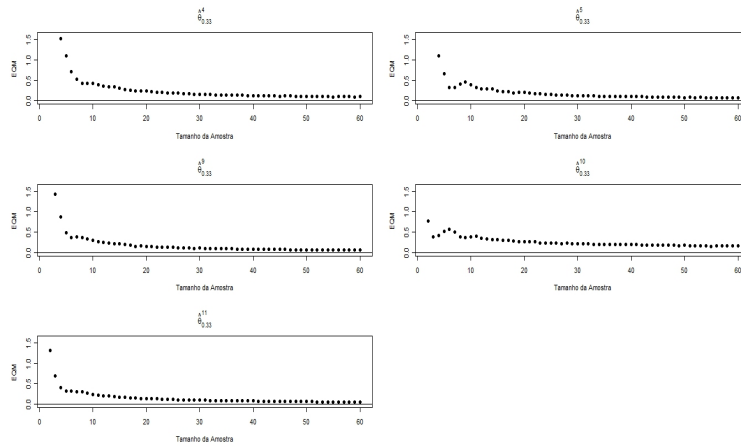


Figura 187 EQM para cada tamanho amostral e os estimadores $\hat{\theta}_{\Gamma}^4$, $\hat{\theta}_{\Gamma}^5$, $\hat{\theta}_{\Gamma}^9$, $\hat{\theta}_{\Gamma}^{10}$ e $\hat{\theta}_{\Gamma}^{11}$, amostras geradas considerando uma **Distribuição Normal** com $\mu = 2,25$ e $\sigma^2 = 1$ para $\Gamma = 0,33$ e o **DNM**

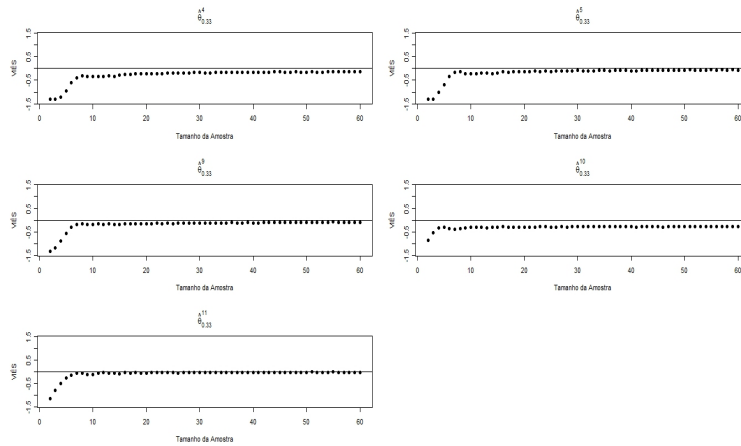


Figura 188 Viés para cada tamanho amostral e os estimadores $\hat{\theta}_{\Gamma}^4$, $\hat{\theta}_{\Gamma}^5$, $\hat{\theta}_{\Gamma}^9$, $\hat{\theta}_{\Gamma}^{10}$ e $\hat{\theta}_{\Gamma}^{11}$, amostras geradas considerando uma **Distribuição Normal** com $\mu = 2,25$ e $\sigma^2 = 1$ para $\Gamma = 0,33$ e o **DNM**

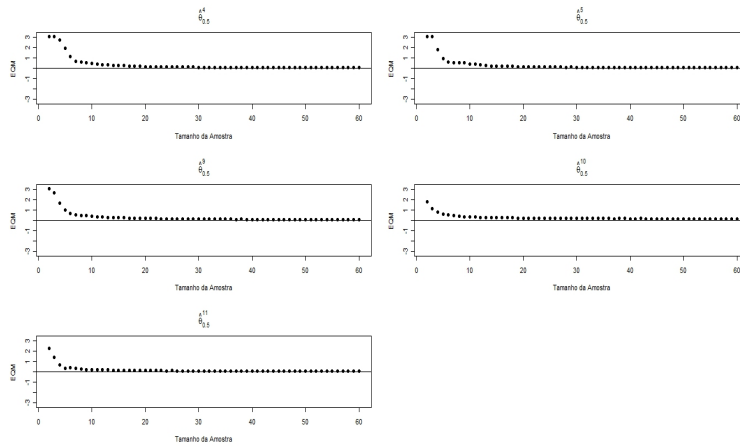


Figura 189 EQM para cada tamanho amostral e os estimadores $\hat{\theta}_{\Gamma}^4$, $\hat{\theta}_{\Gamma}^5$, $\hat{\theta}_{\Gamma}^9$, $\hat{\theta}_{\Gamma}^{10}$ e $\hat{\theta}_{\Gamma}^{11}$, amostras geradas considerando uma **Distribuição Normal** com $\mu = 2,25$ e $\sigma^2 = 1$ para $\Gamma = 0,50$ e o **DNM**

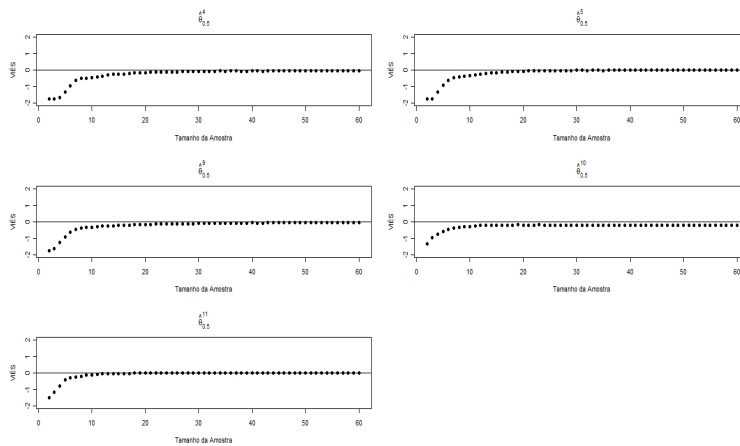


Figura 190 Viés para cada tamanho amostral e os estimadores $\hat{\theta}_{\Gamma}^4$, $\hat{\theta}_{\Gamma}^5$, $\hat{\theta}_{\Gamma}^9$, $\hat{\theta}_{\Gamma}^{10}$ e $\hat{\theta}_{\Gamma}^{11}$, amostras geradas considerando uma **Distribuição Normal** com $\mu = 2,25$ e $\sigma^2 = 1$ para $\Gamma = 0,50$ e o **DNM**

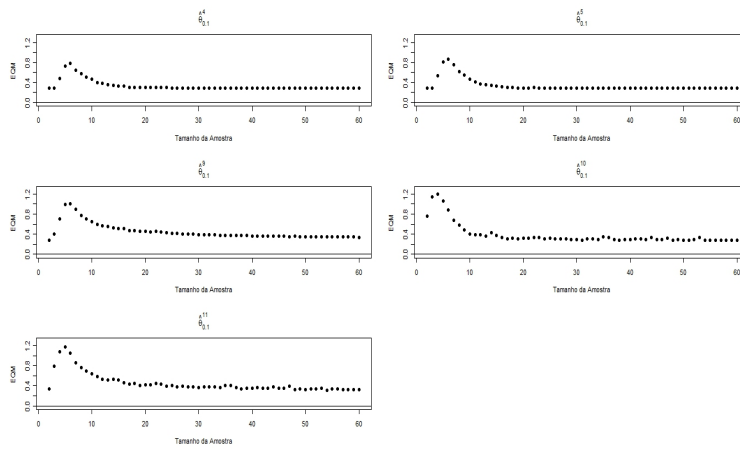


Figura 191 EQM para cada tamanho amostral e os estimadores $\hat{\theta}_{\Gamma}^4$, $\hat{\theta}_{\Gamma}^5$, $\hat{\theta}_{\Gamma}^9$, $\hat{\theta}_{\Gamma}^{10}$ e $\hat{\theta}_{\Gamma}^{11}$, amostras geradas considerando uma **Distribuição Normal** com $\mu = 1,25$ e $\sigma^2 = 1$ para $\Gamma = 0,10$ e o **DNM**

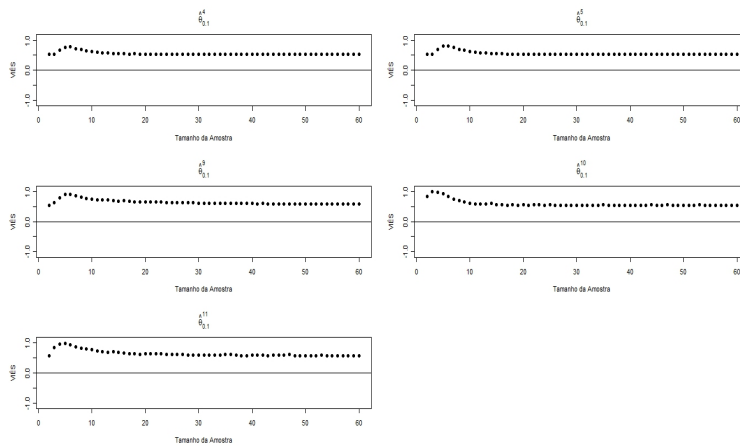


Figura 192 Viés para cada tamanho amostral e os estimadores $\hat{\theta}_{\Gamma}^4$, $\hat{\theta}_{\Gamma}^5$, $\hat{\theta}_{\Gamma}^9$, $\hat{\theta}_{\Gamma}^{10}$ e $\hat{\theta}_{\Gamma}^{11}$, amostras geradas considerando uma **Distribuição Normal** com $\mu = 1,25$ e $\sigma^2 = 1$ para $\Gamma = 0,10$ e o **DNM**

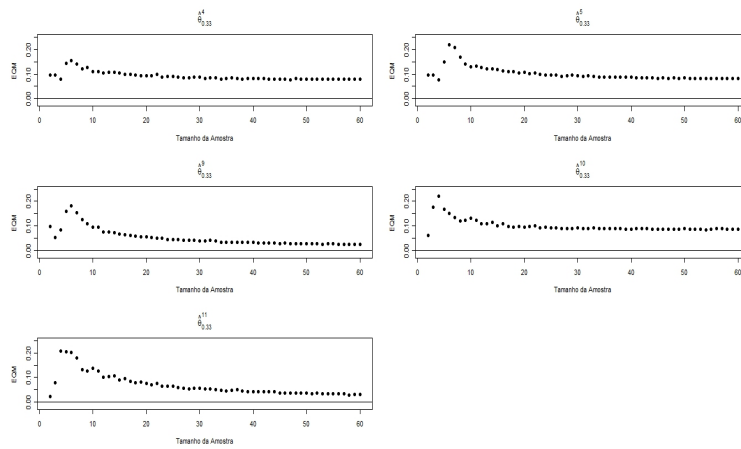


Figura 193 EQM para cada tamanho amostral e os estimadores $\hat{\theta}_{\Gamma}^4$, $\hat{\theta}_{\Gamma}^5$, $\hat{\theta}_{\Gamma}^9$, $\hat{\theta}_{\Gamma}^{10}$ e $\hat{\theta}_{\Gamma}^{11}$, amostras geradas considerando uma **Distribuição Normal** com $\mu = 1,25$ e $\sigma^2 = 1$ para $\Gamma = 0,33$ e o **DNM**

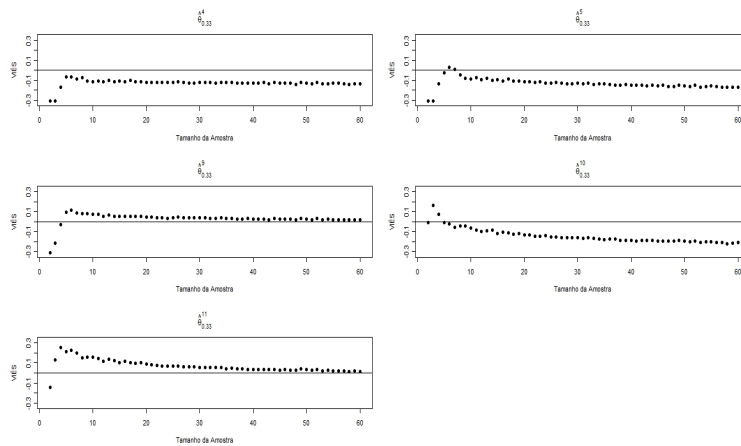


Figura 194 Viés para cada tamanho amostral e os estimadores $\hat{\theta}_{\Gamma}^4$, $\hat{\theta}_{\Gamma}^5$, $\hat{\theta}_{\Gamma}^9$, $\hat{\theta}_{\Gamma}^{10}$ e $\hat{\theta}_{\Gamma}^{11}$, amostras geradas considerando uma **Distribuição Normal** com $\mu = 1,25$ e $\sigma^2 = 1$ para $\Gamma = 0,33$ e o **DNM**

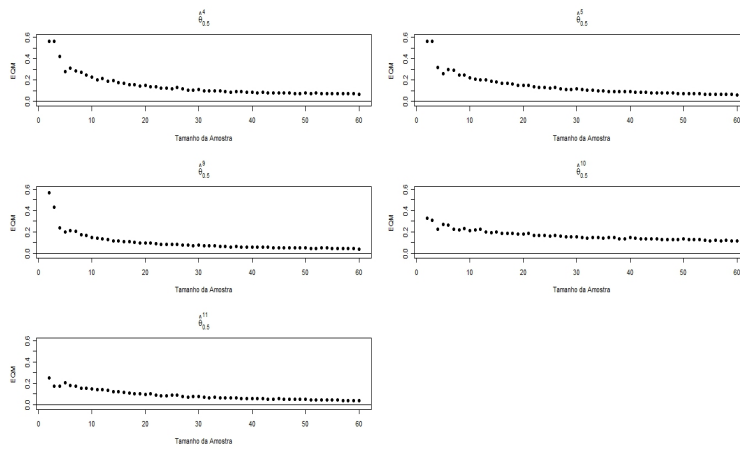


Figura 195 EQM para cada tamanho amostral e os estimadores $\hat{\theta}_{\Gamma}^4$, $\hat{\theta}_{\Gamma}^5$, $\hat{\theta}_{\Gamma}^9$, $\hat{\theta}_{\Gamma}^{10}$ e $\hat{\theta}_{\Gamma}^{11}$, amostras geradas considerando uma **Distribuição Normal** com $\mu = 1,25$ e $\sigma^2 = 1$ para $\Gamma = 0,50$ e o **DNM**

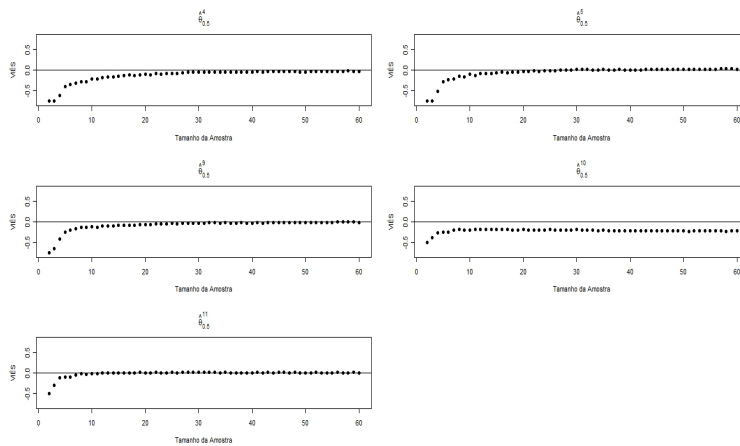


Figura 196 Viés para cada tamanho amostral e os estimadores $\hat{\theta}_{\Gamma}^4$, $\hat{\theta}_{\Gamma}^5$, $\hat{\theta}_{\Gamma}^9$, $\hat{\theta}_{\Gamma}^{10}$ e $\hat{\theta}_{\Gamma}^{11}$, amostras geradas considerando uma **Distribuição Normal** com $\mu = 1,25$ e $\sigma^2 = 1$ para $\Gamma = 0,50$ e o **DNM**

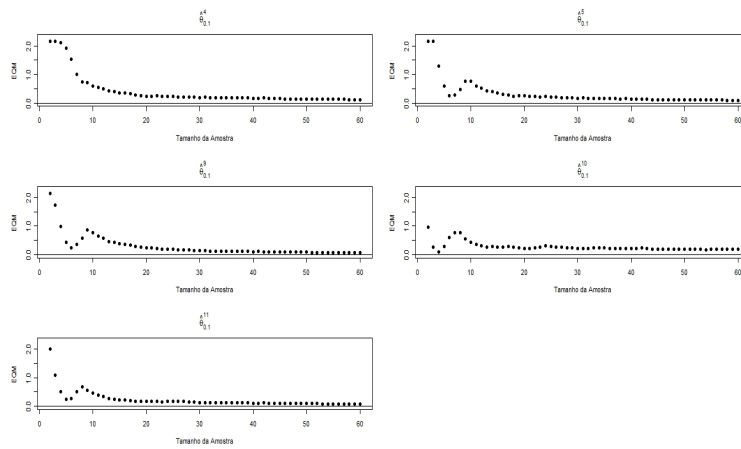


Figura 197 EQM para cada tamanho amostral e os estimadores $\hat{\theta}_{\Gamma}^4$, $\hat{\theta}_{\Gamma}^5$, $\hat{\theta}_{\Gamma}^9$, $\hat{\theta}_{\Gamma}^{10}$ e $\hat{\theta}_{\Gamma}^{11}$, amostras geradas considerando uma **Distribuição Normal** com $\mu = 3,25$ e $\sigma^2 = 1$ para $\Gamma = 0,10$ e o **DNM**

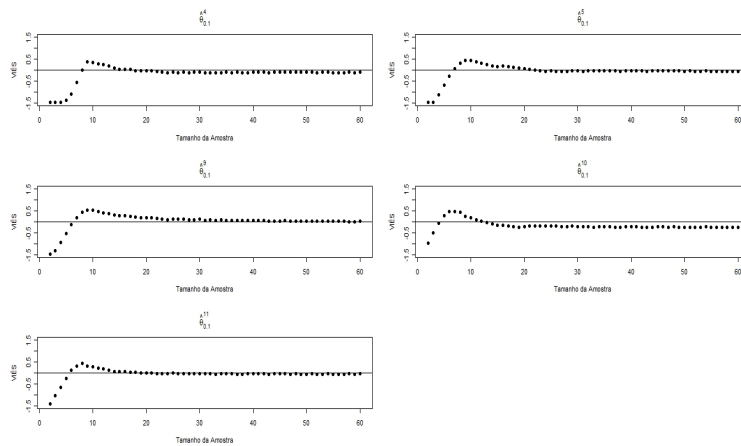


Figura 198 Viés para cada tamanho amostral e os estimadores $\hat{\theta}_{\Gamma}^4$, $\hat{\theta}_{\Gamma}^5$, $\hat{\theta}_{\Gamma}^9$, $\hat{\theta}_{\Gamma}^{10}$ e $\hat{\theta}_{\Gamma}^{11}$, amostras geradas considerando uma **Distribuição Normal** com $\mu = 3,25$ e $\sigma^2 = 1$ para $\Gamma = 0,10$ e o **DNM**

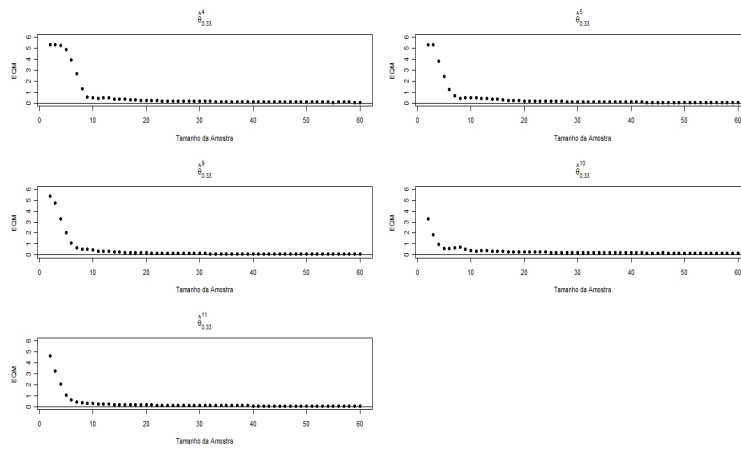


Figura 199 EQM para cada tamanho amostral e os estimadores $\hat{\theta}_{\Gamma}^4$, $\hat{\theta}_{\Gamma}^5$, $\hat{\theta}_{\Gamma}^9$, $\hat{\theta}_{\Gamma}^{10}$ e $\hat{\theta}_{\Gamma}^{11}$, amostras geradas considerando uma **Distribuição Normal** com $\mu = 3,25$ e $\sigma^2 = 1$ para $\Gamma = 0,33$ e o **DNM**

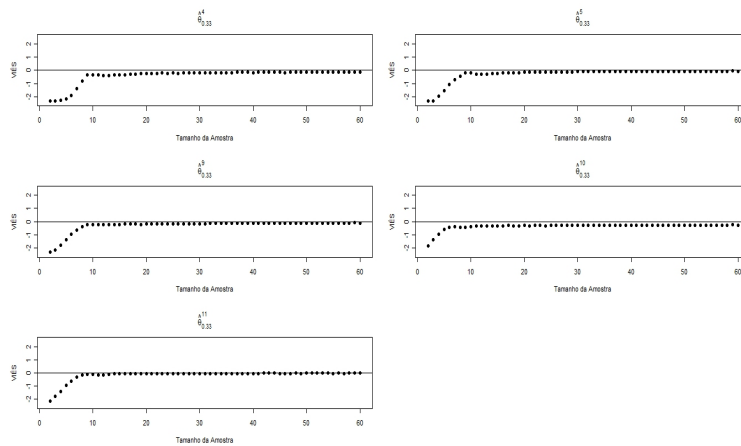


Figura 200 Viés para cada tamanho amostral e os estimadores $\hat{\theta}_{\Gamma}^4$, $\hat{\theta}_{\Gamma}^5$, $\hat{\theta}_{\Gamma}^9$, $\hat{\theta}_{\Gamma}^{10}$ e $\hat{\theta}_{\Gamma}^{11}$, amostras geradas considerando uma **Distribuição Normal** com $\mu = 3,25$ e $\sigma^2 = 1$ para $\Gamma = 0,33$ e o **DNM**

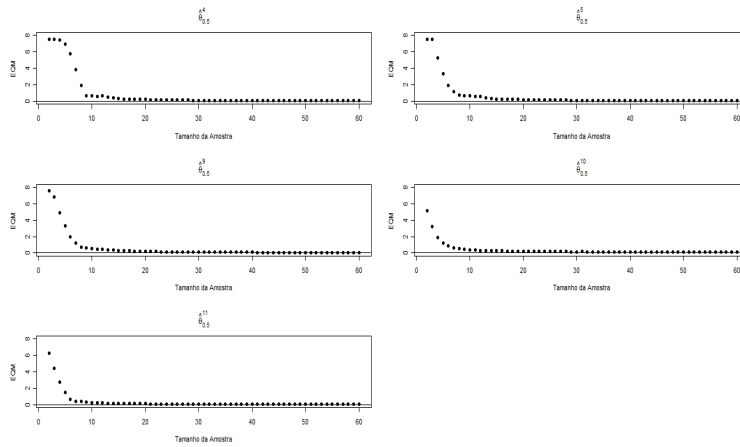


Figura 201 EQM para cada tamanho amostral e os estimadores $\hat{\theta}_{\Gamma}^4$, $\hat{\theta}_{\Gamma}^5$, $\hat{\theta}_{\Gamma}^9$, $\hat{\theta}_{\Gamma}^{10}$ e $\hat{\theta}_{\Gamma}^{11}$, amostras geradas considerando uma **Distribuição Normal** com $\mu = 3,25$ e $\sigma^2 = 1$ para $\Gamma = 0,50$ e o **DNM**

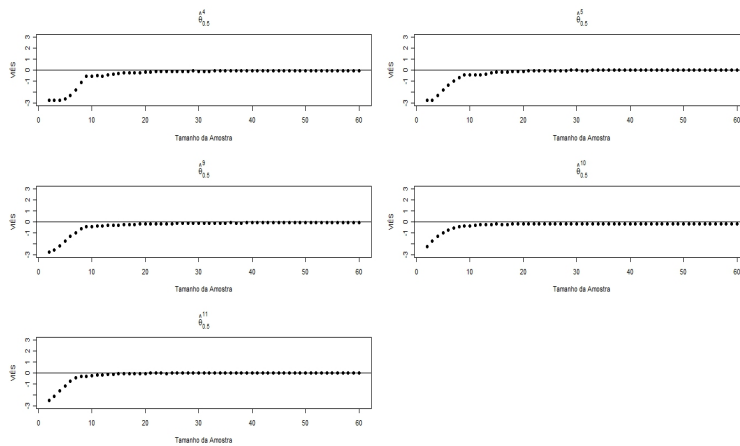


Figura 202 Viés para cada tamanho amostral e os estimadores $\hat{\theta}_{\Gamma}^4$, $\hat{\theta}_{\Gamma}^5$, $\hat{\theta}_{\Gamma}^9$, $\hat{\theta}_{\Gamma}^{10}$ e $\hat{\theta}_{\Gamma}^{11}$, amostras geradas considerando uma **Distribuição Normal** com $\mu = 3,25$ e $\sigma^2 = 1$ para $\Gamma = 0,50$ e o **DNM**

APÊNDICE I

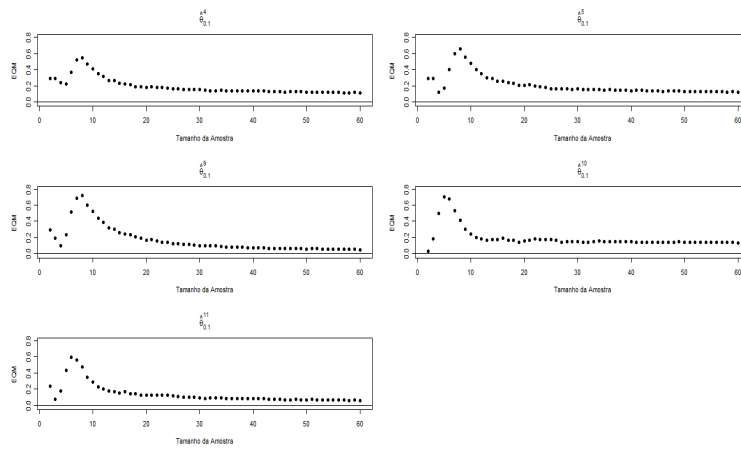


Figura 203 EQM para cada tamanho amostral e os estimadores $\hat{\theta}_{\Gamma}^4$, $\hat{\theta}_{\Gamma}^5$, $\hat{\theta}_{\Gamma}^9$, $\hat{\theta}_{\Gamma}^{10}$ e $\hat{\theta}_{\Gamma}^{11}$, amostras geradas considerando uma **Distribuição Logística** com $\mu = 2,25$ e $\sigma^2 = 1$ para $\Gamma = 0,10$ e o **DNM**

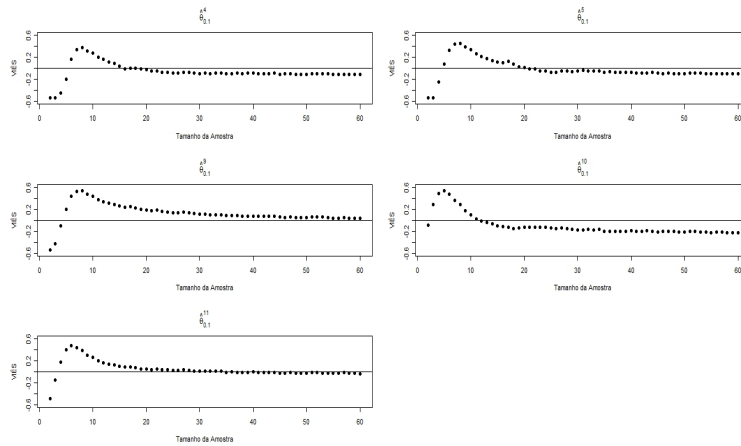


Figura 204 Viés para cada tamanho amostral e os estimadores $\hat{\theta}_{\Gamma}^4$, $\hat{\theta}_{\Gamma}^5$, $\hat{\theta}_{\Gamma}^9$, $\hat{\theta}_{\Gamma}^{10}$ e $\hat{\theta}_{\Gamma}^{11}$, amostras geradas considerando uma **Distribuição Logística** com $\mu = 2,25$ e $\sigma^2 = 1$ para $\Gamma = 0,10$ e o **DNM**

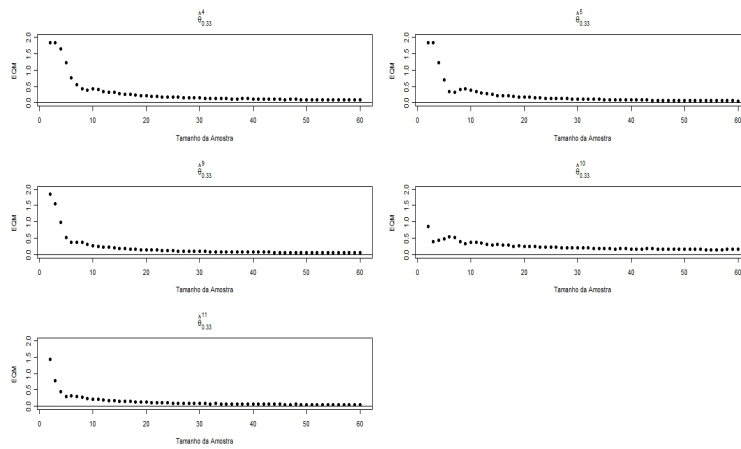


Figura 205 EQM para cada tamanho amostral e os estimadores $\hat{\theta}_{\Gamma}^4$, $\hat{\theta}_{\Gamma}^5$, $\hat{\theta}_{\Gamma}^9$, $\hat{\theta}_{\Gamma}^{10}$ e $\hat{\theta}_{\Gamma}^{11}$, amostras geradas considerando uma **Distribuição Logística** com $\mu = 2,25$ e $\sigma^2 = 1$ para $\Gamma = 0,33$ e o **DNM**

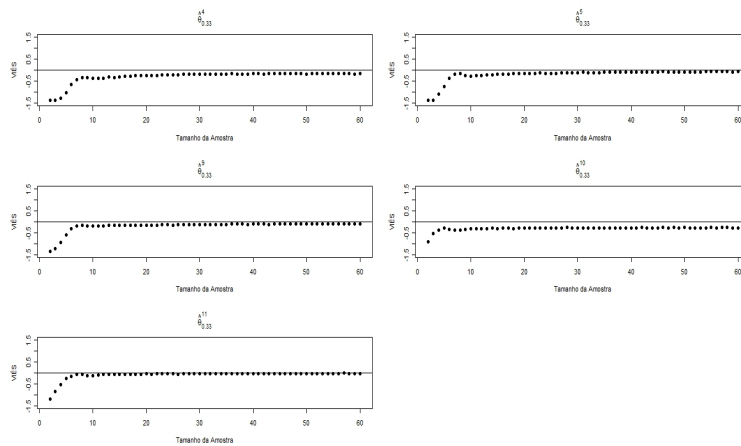


Figura 206 Viés para cada tamanho amostral e os estimadores $\hat{\theta}_{\Gamma}^4$, $\hat{\theta}_{\Gamma}^5$, $\hat{\theta}_{\Gamma}^9$, $\hat{\theta}_{\Gamma}^{10}$ e $\hat{\theta}_{\Gamma}^{11}$, amostras geradas considerando uma **Distribuição Logística** com $\mu = 2,25$ e $\sigma^2 = 1$ para $\Gamma = 0,33$ e o **DNM**

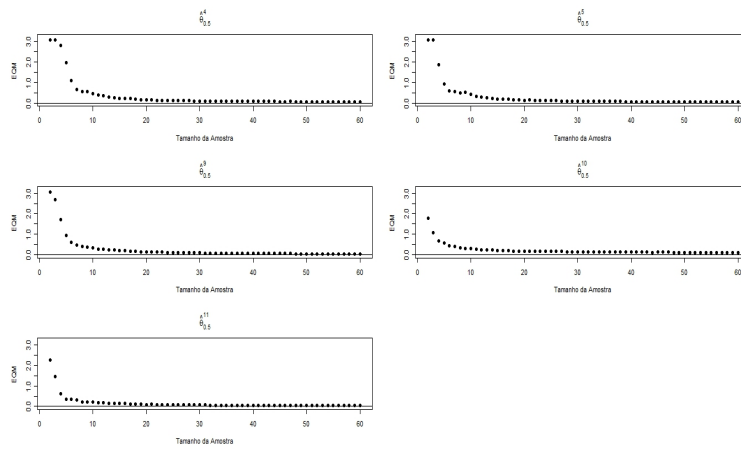


Figura 207 EQM para cada tamanho amostral e os estimadores $\hat{\theta}_{\Gamma}^4$, $\hat{\theta}_{\Gamma}^5$, $\hat{\theta}_{\Gamma}^9$, $\hat{\theta}_{\Gamma}^{10}$ e $\hat{\theta}_{\Gamma}^{11}$, amostras geradas considerando uma **Distribuição Logística** com $\mu = 2,25$ e $\sigma^2 = 1$ para $\Gamma = 0,50$ e o **DNM**

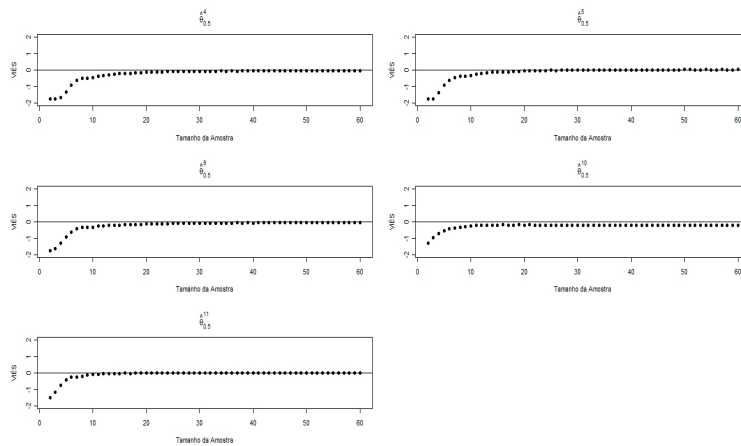


Figura 208 Viés para cada tamanho amostral e os estimadores $\hat{\theta}_{\Gamma}^4$, $\hat{\theta}_{\Gamma}^5$, $\hat{\theta}_{\Gamma}^9$, $\hat{\theta}_{\Gamma}^{10}$ e $\hat{\theta}_{\Gamma}^{11}$, amostras geradas considerando uma **Distribuição Logística** com $\mu = 2,25$ e $\sigma^2 = 1$ para $\Gamma = 0,50$ e o **DNM**

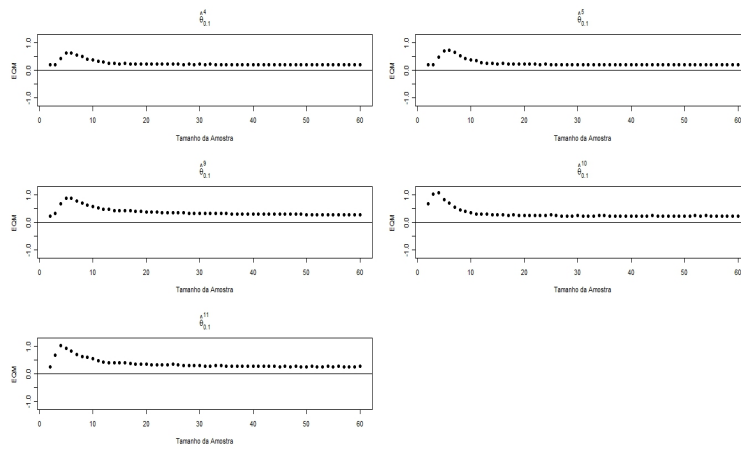


Figura 209 EQM para cada tamanho amostral e os estimadores $\hat{\theta}_{\Gamma}^4$, $\hat{\theta}_{\Gamma}^5$, $\hat{\theta}_{\Gamma}^9$, $\hat{\theta}_{\Gamma}^{10}$ e $\hat{\theta}_{\Gamma}^{11}$, amostras geradas considerando uma **Distribuição Logística** com $\mu = 1,25$ e $\sigma^2 = 1$ para $\Gamma = 0,10$ e o **DNM**

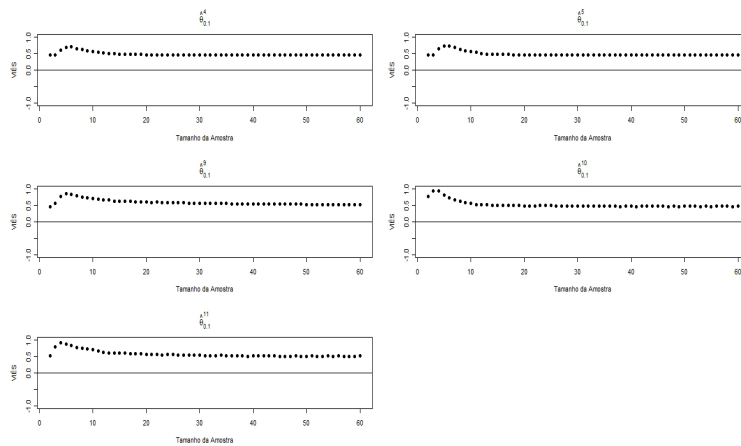


Figura 210 Viés para cada tamanho amostral e os estimadores $\hat{\theta}_{\Gamma}^4$, $\hat{\theta}_{\Gamma}^5$, $\hat{\theta}_{\Gamma}^9$, $\hat{\theta}_{\Gamma}^{10}$ e $\hat{\theta}_{\Gamma}^{11}$, amostras geradas considerando uma **Distribuição Logística** com $\mu = 1,25$ e $\sigma^2 = 1$ para $\Gamma = 0,10$ e o **DNM**

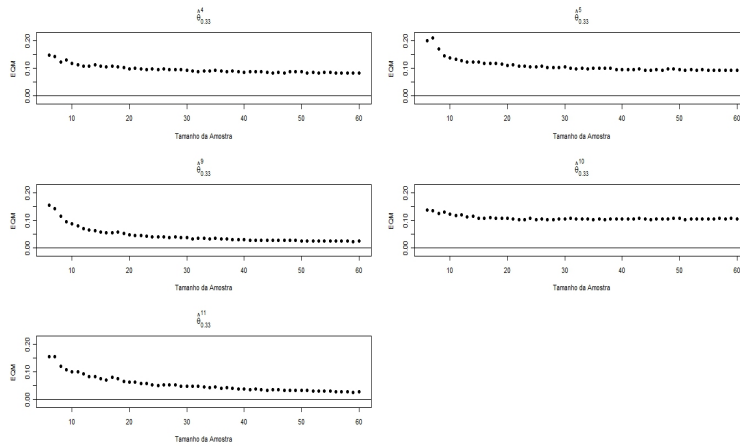


Figura 211 EQM para cada tamanho amostral e os estimadores $\hat{\theta}_{\Gamma}^4$, $\hat{\theta}_{\Gamma}^5$, $\hat{\theta}_{\Gamma}^9$, $\hat{\theta}_{\Gamma}^{10}$ e $\hat{\theta}_{\Gamma}^{11}$, amostras geradas considerando uma **Distribuição Logística** com $\mu = 1,25$ e $\sigma^2 = 1$ para $\Gamma = 0,33$ e o **DNM**

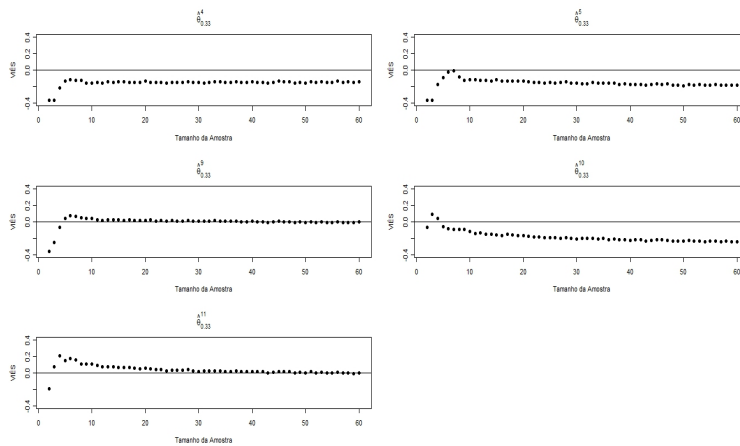


Figura 212 Viés para cada tamanho amostral e os estimadores $\hat{\theta}_{\Gamma}^4$, $\hat{\theta}_{\Gamma}^5$, $\hat{\theta}_{\Gamma}^9$, $\hat{\theta}_{\Gamma}^{10}$ e $\hat{\theta}_{\Gamma}^{11}$, amostras geradas considerando uma **Distribuição Logística** com $\mu = 1,25$ e $\sigma^2 = 1$ para $\Gamma = 0,33$ e o **DNM**

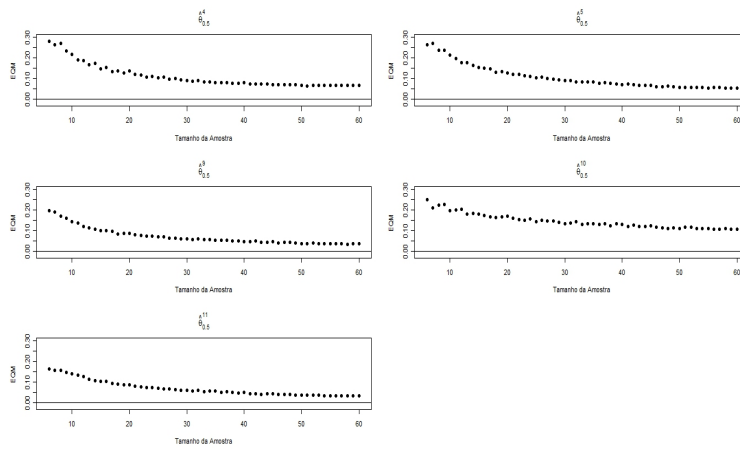


Figura 213 EQM para cada tamanho amostral e os estimadores $\hat{\theta}_{\Gamma}^4$, $\hat{\theta}_{\Gamma}^5$, $\hat{\theta}_{\Gamma}^9$, $\hat{\theta}_{\Gamma}^{10}$ e $\hat{\theta}_{\Gamma}^{11}$, amostras geradas considerando uma **Distribuição Logística** com $\mu = 1,25$ e $\sigma^2 = 1$ para $\Gamma = 0,50$ e o **DNM**

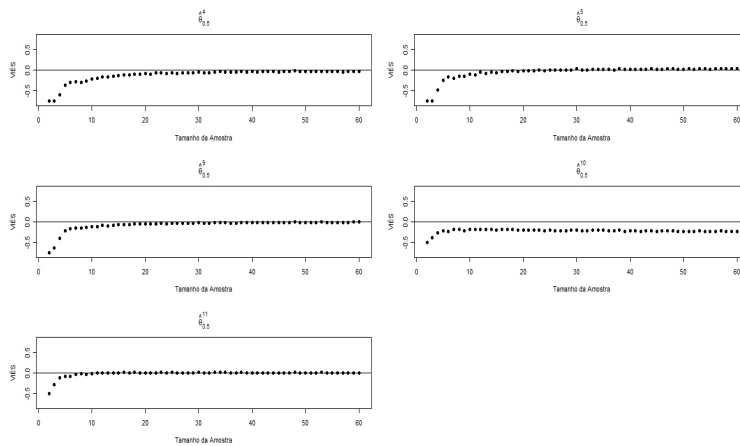


Figura 214 Viés para cada tamanho amostral e os estimadores $\hat{\theta}_{\Gamma}^4$, $\hat{\theta}_{\Gamma}^5$, $\hat{\theta}_{\Gamma}^9$, $\hat{\theta}_{\Gamma}^{10}$ e $\hat{\theta}_{\Gamma}^{11}$, amostras geradas considerando uma **Distribuição Logística** com $\mu = 1,25$ e $\sigma^2 = 1$ para $\Gamma = 0,50$ e o **DNM**

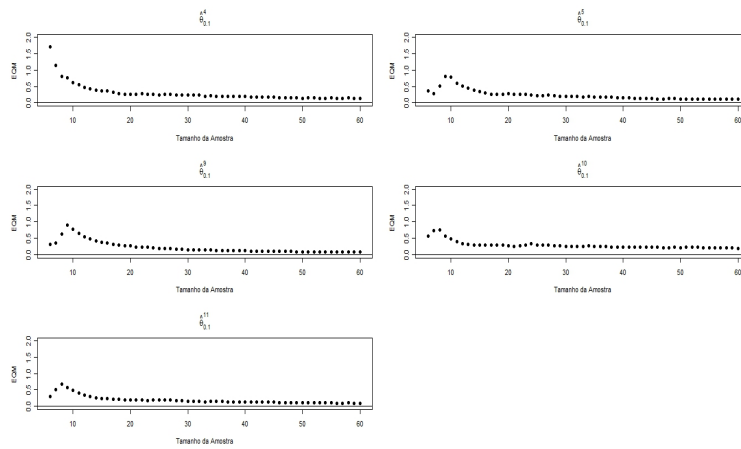


Figura 215 EQM para cada tamanho amostral e os estimadores $\hat{\theta}_{\Gamma}^4$, $\hat{\theta}_{\Gamma}^5$, $\hat{\theta}_{\Gamma}^9$, $\hat{\theta}_{\Gamma}^{10}$ e $\hat{\theta}_{\Gamma}^{11}$, amostras geradas considerando uma **Distribuição Logística** com $\mu = 3,25$ e $\sigma^2 = 1$ para $\Gamma = 0,10$ e o **DNM**

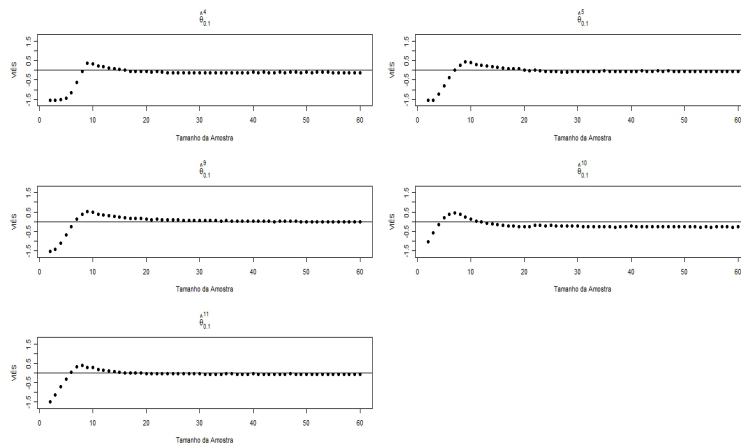


Figura 216 Viés para cada tamanho amostral e os estimadores $\hat{\theta}_{\Gamma}^4$, $\hat{\theta}_{\Gamma}^5$, $\hat{\theta}_{\Gamma}^9$, $\hat{\theta}_{\Gamma}^{10}$ e $\hat{\theta}_{\Gamma}^{11}$, amostras geradas considerando uma **Distribuição Logística** com $\mu = 3,25$ e $\sigma^2 = 1$ para $\Gamma = 0,10$ e o **DNM**

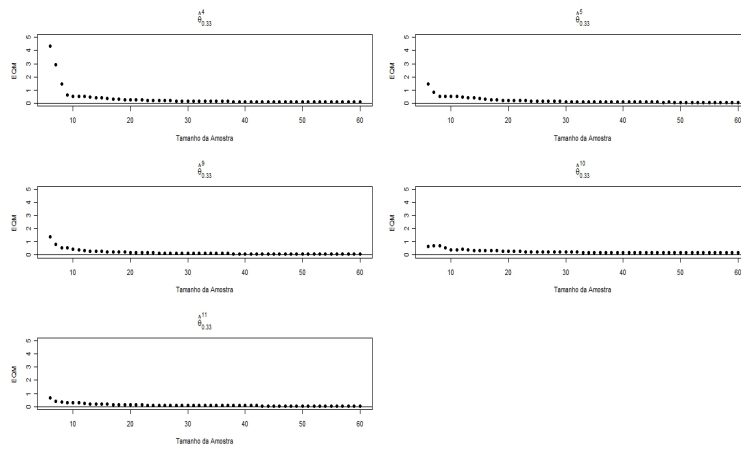


Figura 217 EQM para cada tamanho amostral e os estimadores $\hat{\theta}_{\Gamma}^4$, $\hat{\theta}_{\Gamma}^5$, $\hat{\theta}_{\Gamma}^9$, $\hat{\theta}_{\Gamma}^{10}$ e $\hat{\theta}_{\Gamma}^{11}$, amostras geradas considerando uma **Distribuição Logística** com $\mu = 3,25$ e $\sigma^2 = 1$ para $\Gamma = 0,33$ e o DNM

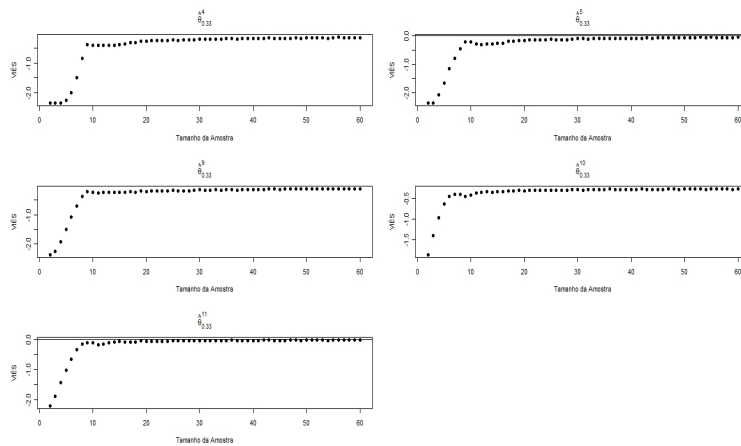


Figura 218 Viés para cada tamanho amostral e os estimadores $\hat{\theta}_{\Gamma}^4$, $\hat{\theta}_{\Gamma}^5$, $\hat{\theta}_{\Gamma}^9$, $\hat{\theta}_{\Gamma}^{10}$ e $\hat{\theta}_{\Gamma}^{11}$, amostras geradas considerando uma **Distribuição Logística** com $\mu = 3,25$ e $\sigma^2 = 1$ para $\Gamma = 0,33$ e o DNM

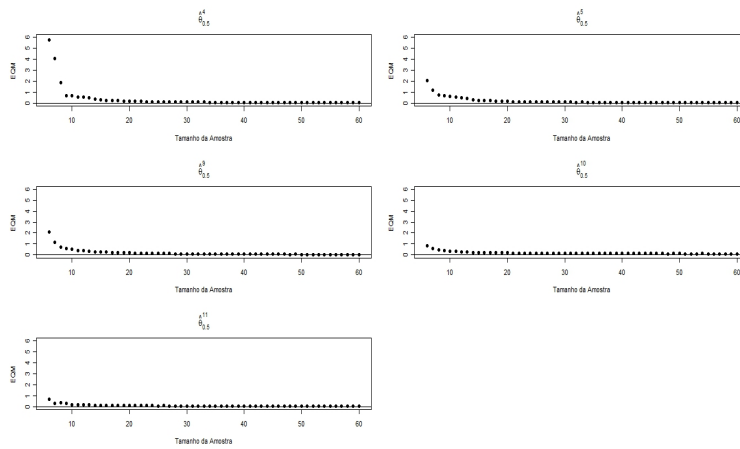


Figura 219 EQM para cada tamanho amostral e os estimadores $\hat{\theta}_{\Gamma}^4$, $\hat{\theta}_{\Gamma}^5$, $\hat{\theta}_{\Gamma}^9$, $\hat{\theta}_{\Gamma}^{10}$ e $\hat{\theta}_{\Gamma}^{11}$, amostras geradas considerando uma **Distribuição Logística** com $\mu = 3,25$ e $\sigma^2 = 1$ para $\Gamma = 0,50$ e o **DNM**

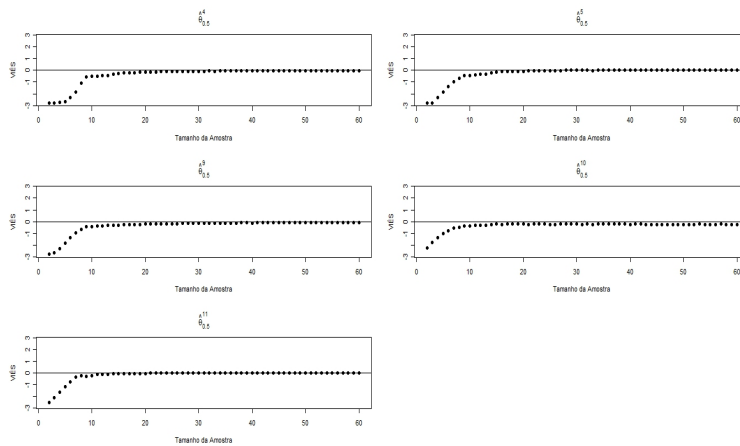


Figura 220 Viés para cada tamanho amostral e os estimadores $\hat{\theta}_{\Gamma}^4$, $\hat{\theta}_{\Gamma}^5$, $\hat{\theta}_{\Gamma}^9$, $\hat{\theta}_{\Gamma}^{10}$ e $\hat{\theta}_{\Gamma}^{11}$, amostras geradas considerando uma **Distribuição Logística** com $\mu = 3,25$ e $\sigma^2 = 1$ para $\Gamma = 0,50$ e o **DNM**

APÊNDICE J

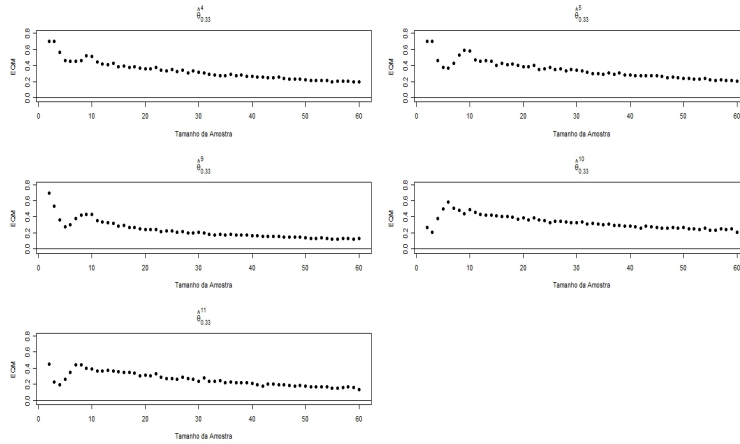


Figura 221 EQM para cada tamanho amostral e os estimadores $\hat{\theta}_{\Gamma}^4$, $\hat{\theta}_{\Gamma}^5$, $\hat{\theta}_{\Gamma}^9$, $\hat{\theta}_{\Gamma}^{10}$ e $\hat{\theta}_{\Gamma}^{11}$, amostras geradas considerando uma **Distribuição Exponencial** com parâmetro $\lambda = 0,30$, para $\Gamma = 0,33$ e o **DNM**

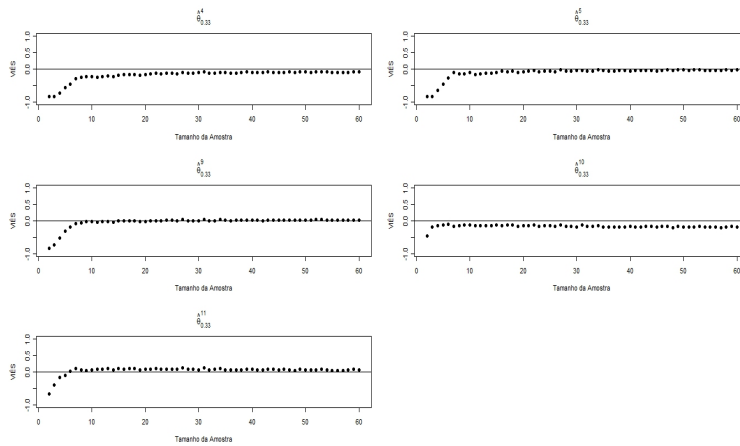


Figura 222 Viés para cada tamanho amostral e os estimadores $\hat{\theta}_{\Gamma}^4$, $\hat{\theta}_{\Gamma}^5$, $\hat{\theta}_{\Gamma}^9$, $\hat{\theta}_{\Gamma}^{10}$ e $\hat{\theta}_{\Gamma}^{11}$, amostras geradas considerando uma **Distribuição Exponencial** com parâmetro $\lambda = 0,30$, para $\Gamma = 0,33$ e o **DNM**

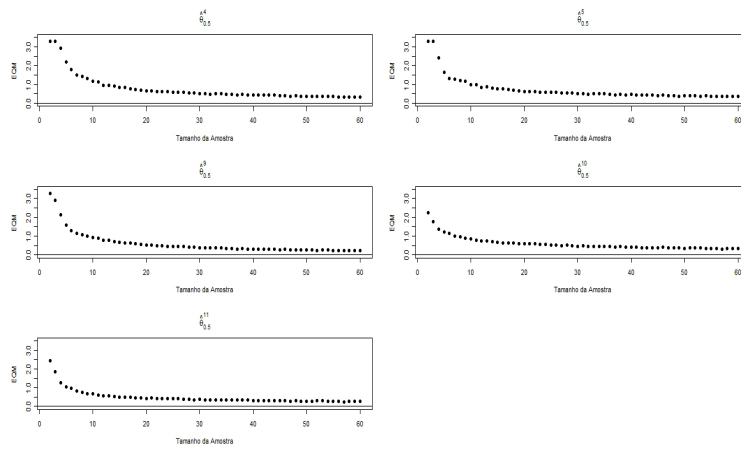


Figura 223 EQM para cada tamanho amostral e os estimadores $\hat{\theta}_{\Gamma}^4$, $\hat{\theta}_{\Gamma}^5$, $\hat{\theta}_{\Gamma}^9$, $\hat{\theta}_{\Gamma}^{10}$ e $\hat{\theta}_{\Gamma}^{11}$, amostras geradas considerando uma **Distribuição Exponencial** com parâmetro $\lambda = 0,30$, para $\Gamma = 0,50$ e o **DNM**

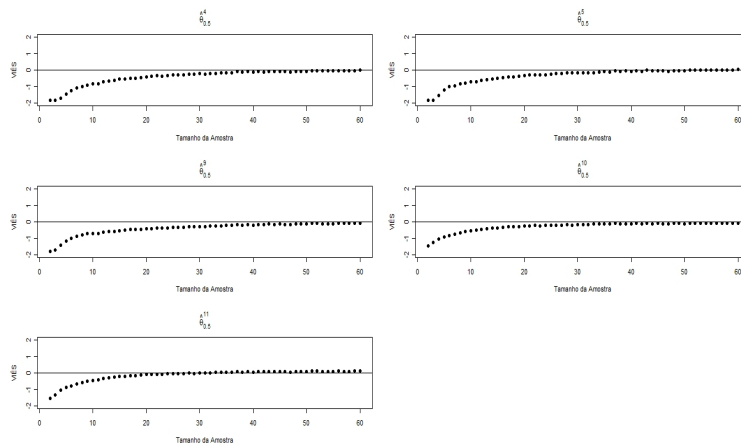


Figura 224 Viés para cada tamanho amostral e os estimadores $\hat{\theta}_{\Gamma}^4$, $\hat{\theta}_{\Gamma}^5$, $\hat{\theta}_{\Gamma}^9$, $\hat{\theta}_{\Gamma}^{10}$ e $\hat{\theta}_{\Gamma}^{11}$, amostras geradas considerando uma **Distribuição Exponencial** com parâmetro $\lambda = 0,30$, para $\Gamma = 0,50$ e o **DNM**

APÊNDICE L

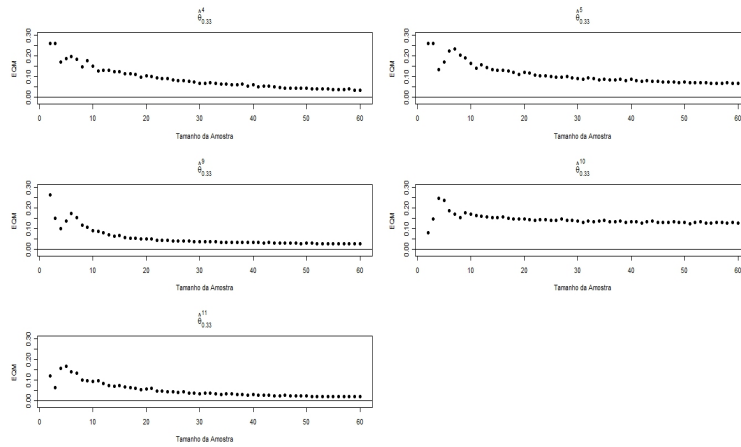


Figura 225 EQM para cada tamanho amostral e os estimadores $\hat{\theta}_{\Gamma}^4$, $\hat{\theta}_{\Gamma}^5$, $\hat{\theta}_{\Gamma}^9$, $\hat{\theta}_{\Gamma}^{10}$ e $\hat{\theta}_{\Gamma}^{11}$, amostras geradas considerando uma **Distribuição Gama** com parâmetros $\alpha = 3$ e $\lambda = 0,50$, para $\Gamma = 0,33$ e o **DNM**

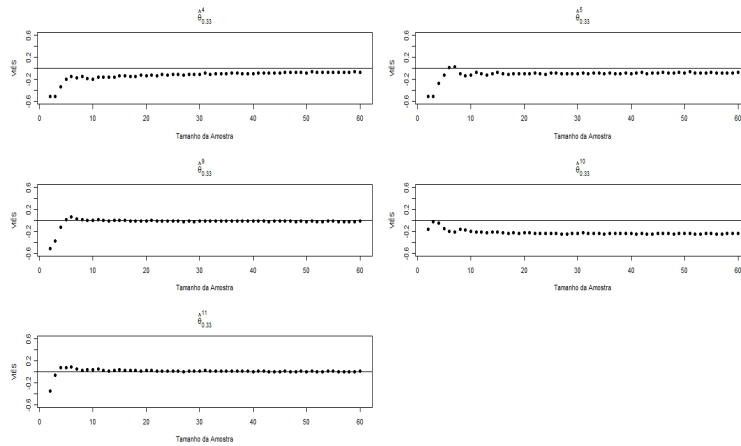


Figura 226 Viés para cada tamanho amostral e os estimadores $\hat{\theta}_{\Gamma}^4$, $\hat{\theta}_{\Gamma}^5$, $\hat{\theta}_{\Gamma}^9$, $\hat{\theta}_{\Gamma}^{10}$ e $\hat{\theta}_{\Gamma}^{11}$, amostras geradas considerando uma **Distribuição Gama** com parâmetros $\alpha = 3$ e $\lambda = 0,50$, para $\Gamma = 0,33$ e o **DNM**

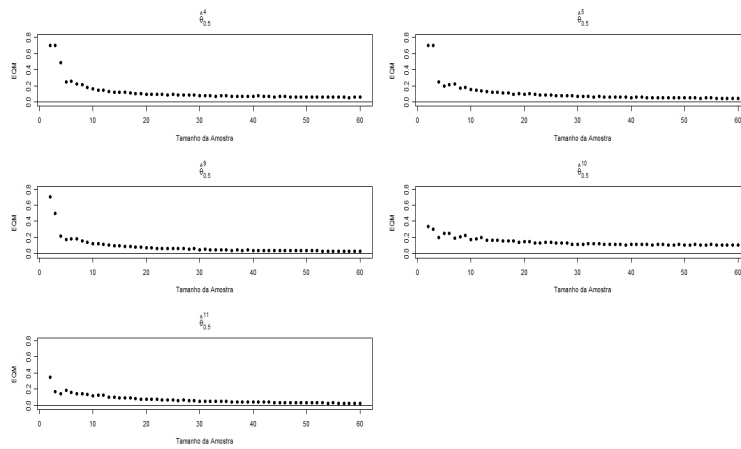


Figura 227 EQM para cada tamanho amostral e os estimadores $\hat{\theta}_{\Gamma}^4$, $\hat{\theta}_{\Gamma}^5$, $\hat{\theta}_{\Gamma}^9$, $\hat{\theta}_{\Gamma}^{10}$ e $\hat{\theta}_{\Gamma}^{11}$, amostras geradas considerando uma **Distribuição Gama** com parâmetros $\alpha = 3$ e $\lambda = 0,50$, para $\Gamma = 0,50$ e o **DNM**

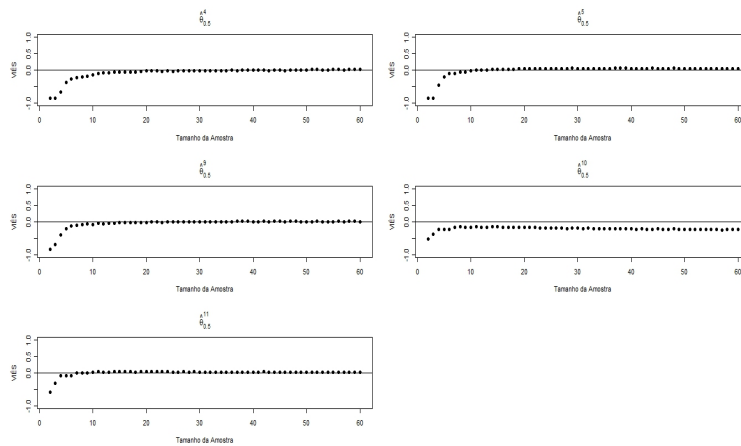


Figura 228 Viés para cada tamanho amostral e os estimadores $\hat{\theta}_{\Gamma}^4$, $\hat{\theta}_{\Gamma}^5$, $\hat{\theta}_{\Gamma}^9$, $\hat{\theta}_{\Gamma}^{10}$ e $\hat{\theta}_{\Gamma}^{11}$, amostras geradas considerando uma **Distribuição Gama** com parâmetros $\alpha = 3$ e $\lambda = 0,50$, para $\Gamma = 0,50$ e o **DNM**

APÊNDICE M

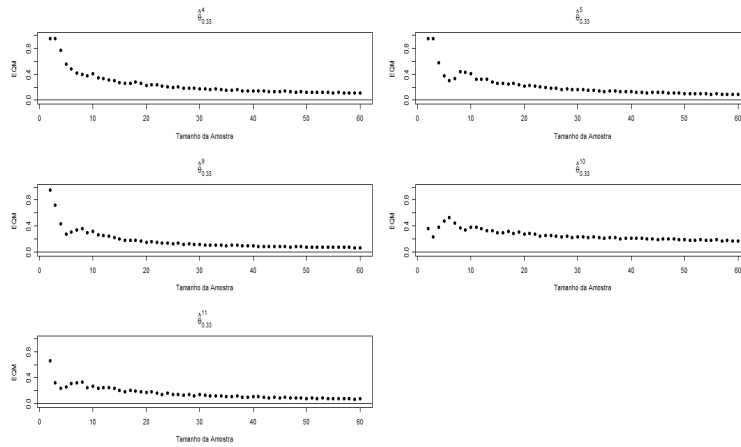


Figura 229 EQM para cada tamanho amostral e os estimadores $\hat{\theta}_{\Gamma}^4$, $\hat{\theta}_{\Gamma}^5$, $\hat{\theta}_{\Gamma}^9$, $\hat{\theta}_{\Gamma}^{10}$ e $\hat{\theta}_{\Gamma}^{11}$, amostras geradas considerando uma **Distribuição Gama** com parâmetros $\alpha = 2$ e $\lambda = 1,25$, para $\Gamma = 0,33$ e o **DNM**

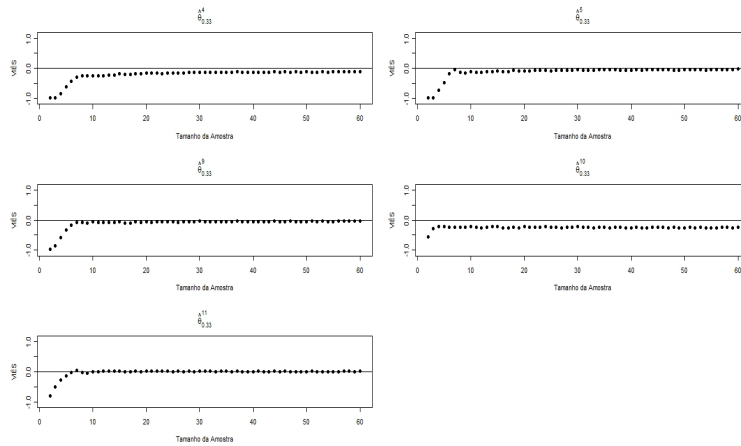


Figura 230 Viés para cada tamanho amostral e os estimadores $\hat{\theta}_{\Gamma}^4$, $\hat{\theta}_{\Gamma}^5$, $\hat{\theta}_{\Gamma}^9$, $\hat{\theta}_{\Gamma}^{10}$ e $\hat{\theta}_{\Gamma}^{11}$, amostras geradas considerando uma **Distribuição Gama** com parâmetros $\alpha = 2$ e $\lambda = 1,25$, para $\Gamma = 0,33$ e o **DNM**

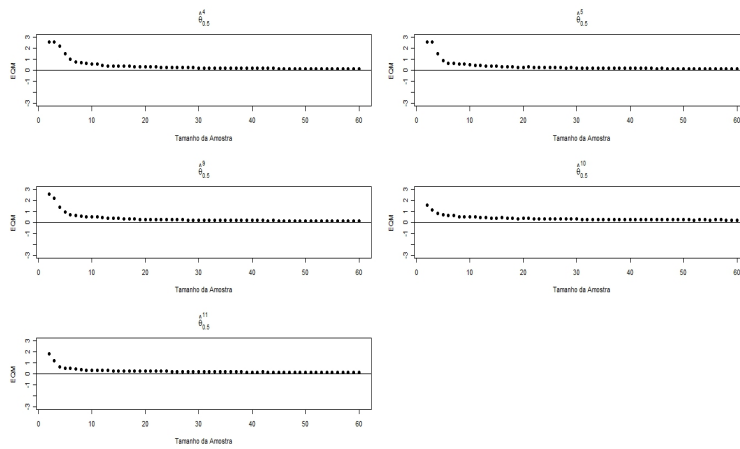


Figura 231 EQM para cada tamanho amostral e os estimadores $\hat{\theta}_{\Gamma}^4$, $\hat{\theta}_{\Gamma}^5$, $\hat{\theta}_{\Gamma}^9$, $\hat{\theta}_{\Gamma}^{10}$ e $\hat{\theta}_{\Gamma}^{11}$, amostras geradas considerando uma **Distribuição Gama** com parâmetros $\alpha = 2$ e $\lambda = 1,25$, para $\Gamma = 0,50$ e o **DNM**

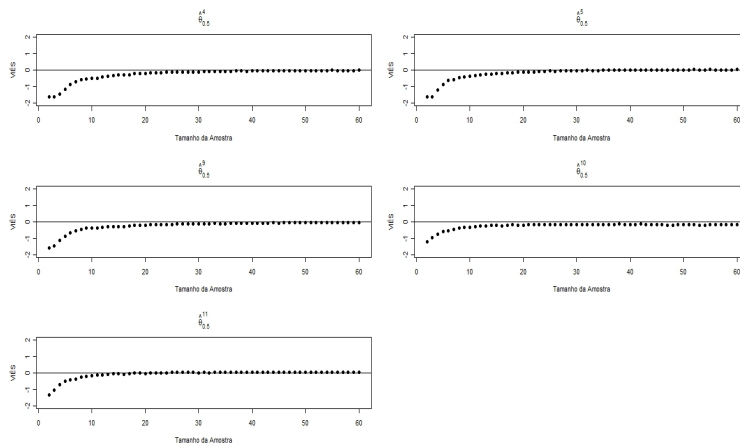


Figura 232 Viés para cada tamanho amostral e os estimadores $\hat{\theta}_{\Gamma}^4$, $\hat{\theta}_{\Gamma}^5$, $\hat{\theta}_{\Gamma}^9$, $\hat{\theta}_{\Gamma}^{10}$ e $\hat{\theta}_{\Gamma}^{11}$, amostras geradas considerando uma **Distribuição Gama** com parâmetros $\alpha = 2$ e $\lambda = 1,25$, para $\Gamma = 0,50$ e o **DNM**

APÊNDICE N

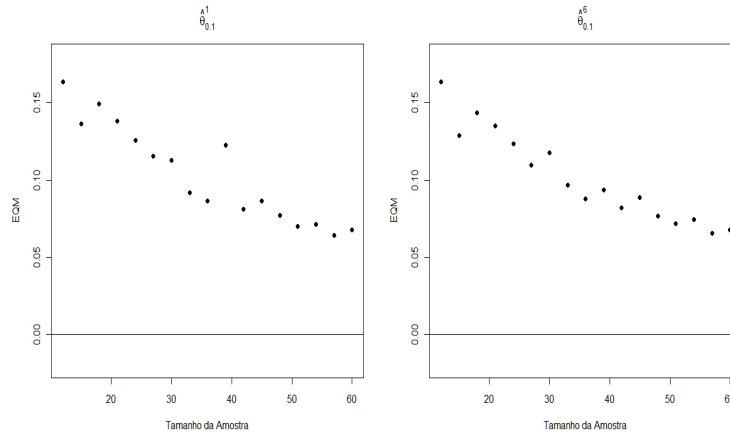


Figura 233 EQM para cada tamanho amostral e os estimadores $\hat{\theta}_{\Gamma}^1$ e $\hat{\theta}_{\Gamma}^6$, amostras geradas considerando uma **Distribuição Normal** com $\mu = 2,25$ e $\sigma^2 = 1$ para $\Gamma = 0,10$ e o **DILW**

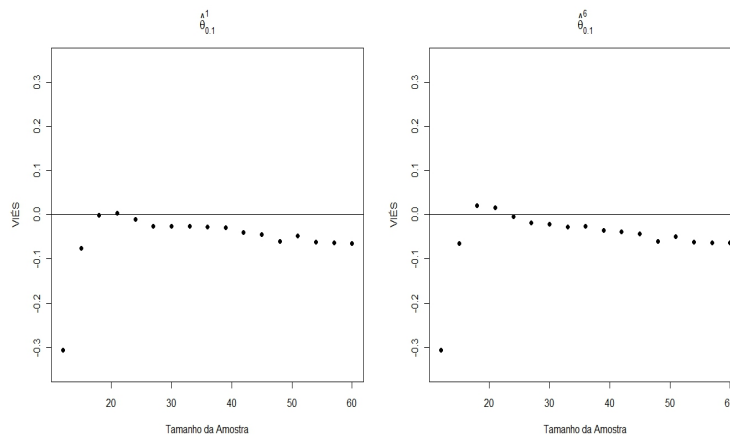


Figura 234 Viés para cada tamanho amostral e os estimadores $\hat{\theta}_{\Gamma}^1$ e $\hat{\theta}_{\Gamma}^6$, amostras geradas considerando uma **Distribuição Normal** com $\mu = 2,25$ e $\sigma^2 = 1$ para $\Gamma = 0,10$ e o **DILW**

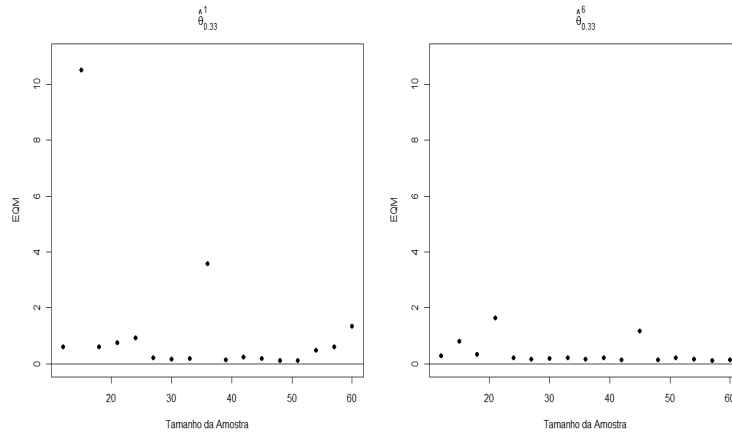


Figura 235 EQM para cada tamanho amostral e os estimadores $\hat{\theta}_{\Gamma}^1$ e $\hat{\theta}_{\Gamma}^6$, amostras geradas considerando uma **Distribuição Normal** com $\mu = 2,25$ e $\sigma^2 = 1$ para $\Gamma = 0,33$ e o **DILW**

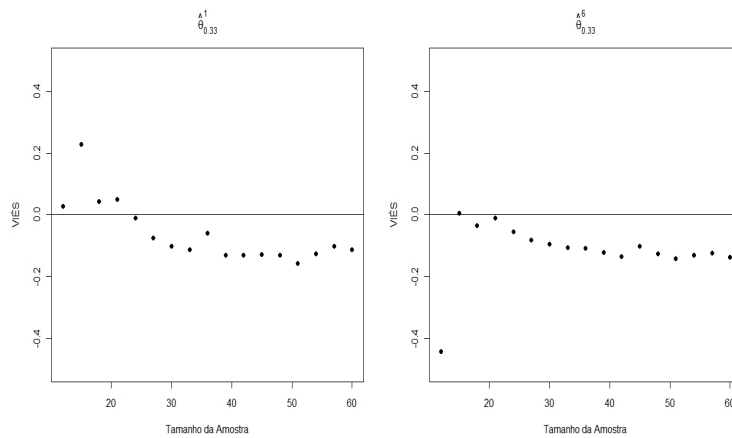


Figura 236 Viés para cada tamanho amostral e os estimadores $\hat{\theta}_{\Gamma}^1$ e $\hat{\theta}_{\Gamma}^6$, amostras geradas considerando uma **Distribuição Normal** com $\mu = 2,25$ e $\sigma^2 = 1$ para $\Gamma = 0,33$ e o **DILW**

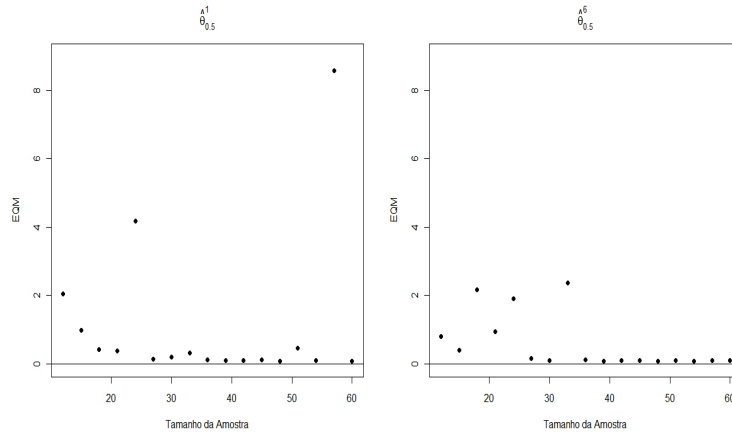


Figura 237 EQM para cada tamanho amostral e os estimadores $\hat{\theta}_\Gamma^1$ e $\hat{\theta}_\Gamma^6$, amostras geradas considerando uma **Distribuição Normal** com $\mu = 2,25$ e $\sigma^2 = 1$ para $\Gamma = 0,50$ e o **DILW**

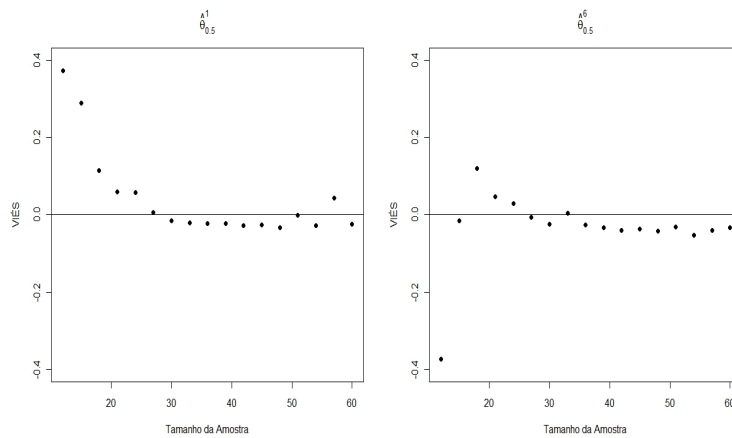


Figura 238 Viés para cada tamanho amostral e os estimadores $\hat{\theta}_\Gamma^1$ e $\hat{\theta}_\Gamma^6$, amostras geradas considerando uma **Distribuição Normal** com $\mu = 2,25$ e $\sigma^2 = 1$ para $\Gamma = 0,50$ e o **DILW**

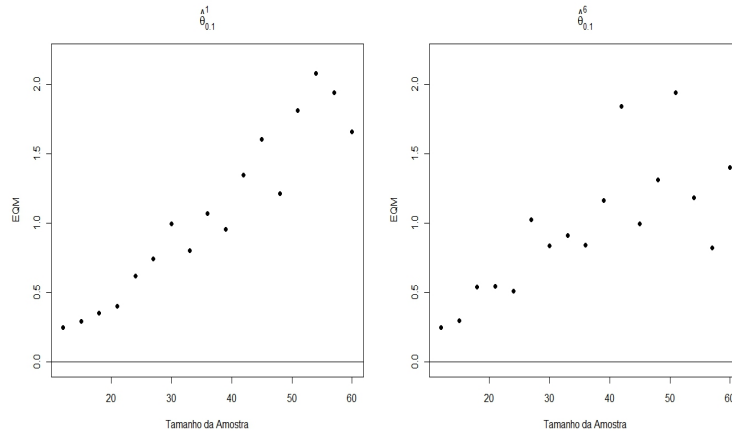


Figura 239 EQM para cada tamanho amostral e os estimadores $\hat{\theta}_{\Gamma}^1$ e $\hat{\theta}_{\Gamma}^6$, amostras geradas considerando uma **Distribuição Normal** com $\mu = 1,25$ e $\sigma^2 = 1$ para $\Gamma = 0,10$ e o **DILW**

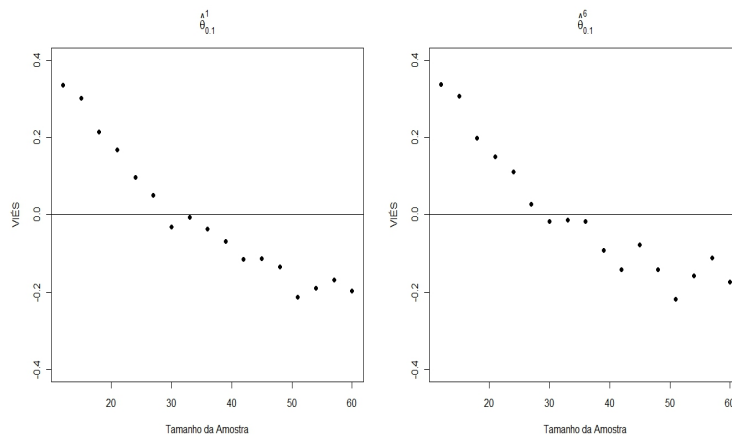


Figura 240 Viés para cada tamanho amostral e os estimadores $\hat{\theta}_{\Gamma}^1$ e $\hat{\theta}_{\Gamma}^6$, amostras geradas considerando uma **Distribuição Normal** com $\mu = 1,25$ e $\sigma^2 = 1$ para $\Gamma = 0,10$ e o **DILW**

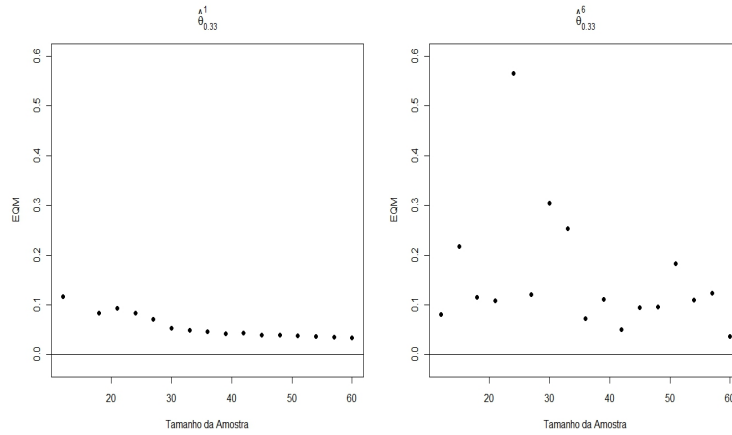


Figura 241 EQM para cada tamanho amostral e os estimadores $\hat{\theta}_{\Gamma}^1$ e $\hat{\theta}_{\Gamma}^6$, amostras geradas considerando uma **Distribuição Normal** com $\mu = 1,25$ e $\sigma^2 = 1$ para $\Gamma = 0,33$ e o **DILW**

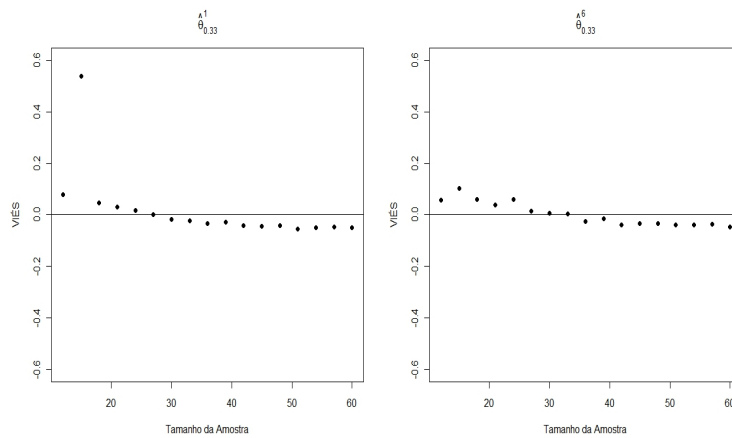


Figura 242 Viés para cada tamanho amostral e os estimadores $\hat{\theta}_{\Gamma}^1$ e $\hat{\theta}_{\Gamma}^6$, amostras geradas considerando uma **Distribuição Normal** com $\mu = 1,25$ e $\sigma^2 = 1$ para $\Gamma = 0,33$ e o **DILW**

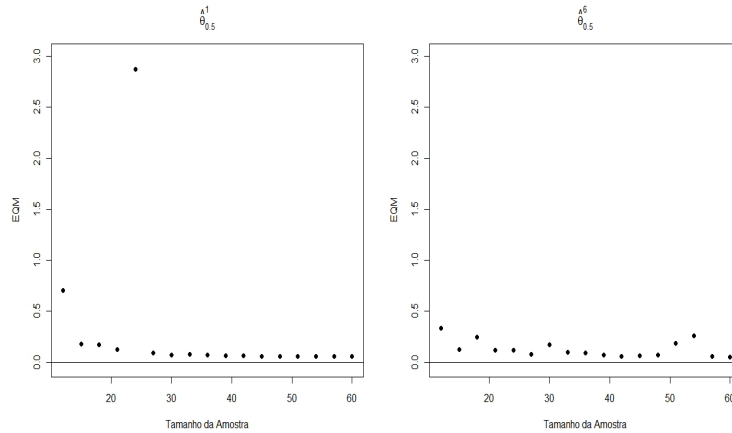


Figura 243 EQM para cada tamanho amostral e os estimadores $\hat{\theta}_{\Gamma}^1$ e $\hat{\theta}_{\Gamma}^6$, amostras geradas considerando uma **Distribuição Normal** com $\mu = 1,25$ e $\sigma^2 = 1$ para $\Gamma = 0,50$ e o **DILW**

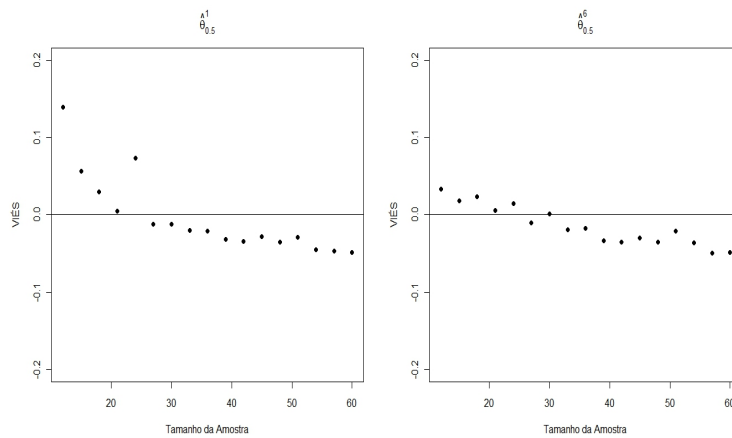


Figura 244 Viés para cada tamanho amostral e os estimadores $\hat{\theta}_{\Gamma}^1$ e $\hat{\theta}_{\Gamma}^6$, amostras geradas considerando uma **Distribuição Normal** com $\mu = 1,25$ e $\sigma^2 = 1$ para $\Gamma = 0,50$ e o **DILW**

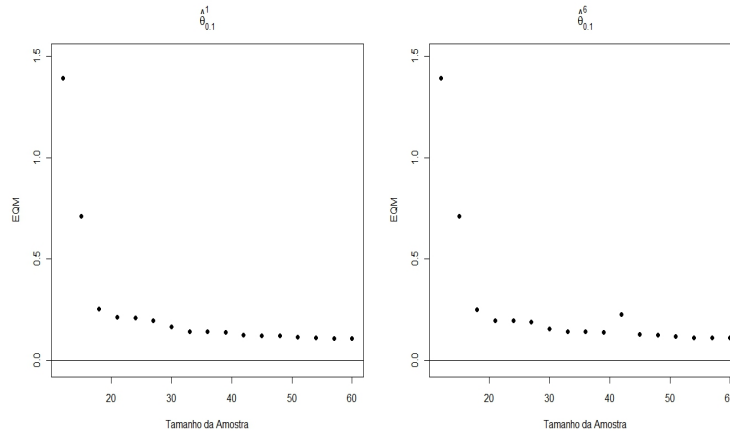


Figura 245 EQM para cada tamanho amostral e os estimadores $\hat{\theta}_{\Gamma}^1$ e $\hat{\theta}_{\Gamma}^6$, amostras geradas considerando uma **Distribuição Normal** com $\mu = 3,25$ e $\sigma^2 = 1$ para $\Gamma = 0,10$ e o **DILW**

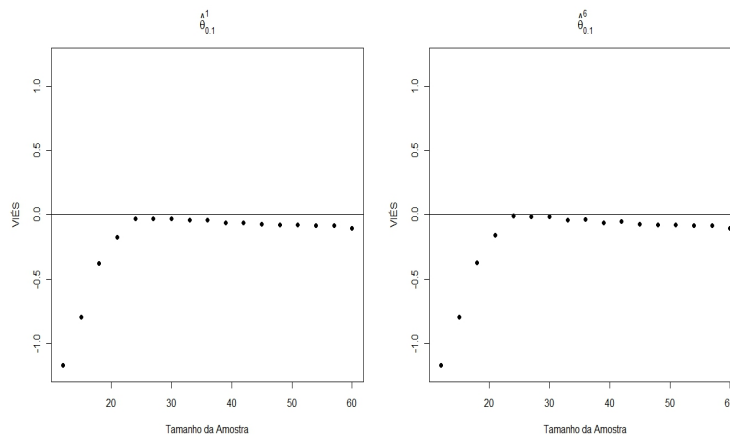


Figura 246 Viés para cada tamanho amostral e os estimadores $\hat{\theta}_{\Gamma}^1$ e $\hat{\theta}_{\Gamma}^6$, amostras geradas considerando uma **Distribuição Normal** com $\mu = 3,25$ e $\sigma^2 = 1$ para $\Gamma = 0,10$ e o **DILW**

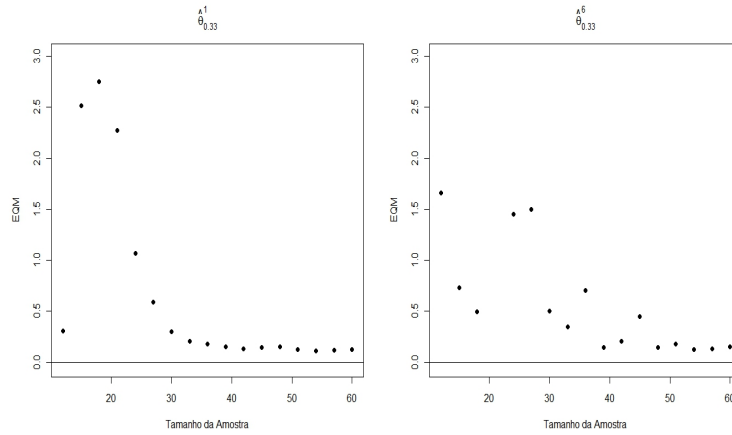


Figura 247 EQM para cada tamanho amostral e os estimadores $\hat{\theta}_{\Gamma}^1$ e $\hat{\theta}_{\Gamma}^6$, amostras geradas considerando uma **Distribuição Normal** com $\mu = 3,25$ e $\sigma^2 = 1$ para $\Gamma = 0,33$ e o **DILW**

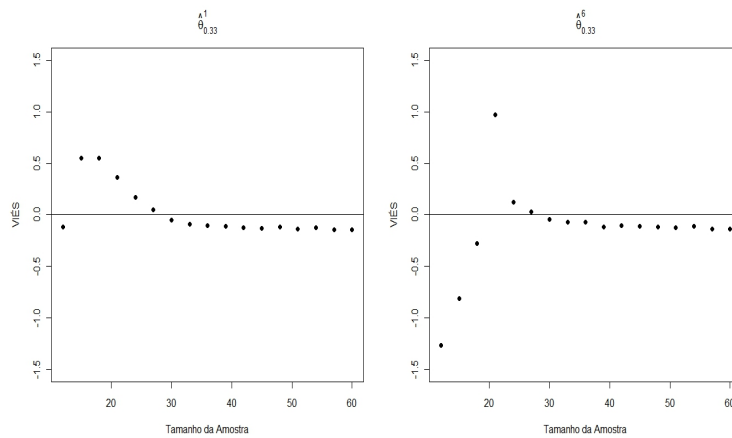


Figura 248 Viés para cada tamanho amostral e os estimadores $\hat{\theta}_{\Gamma}^1$ e $\hat{\theta}_{\Gamma}^6$, amostras geradas considerando uma **Distribuição Normal** com $\mu = 3,25$ e $\sigma^2 = 1$ para $\Gamma = 0,33$ e o **DILW**

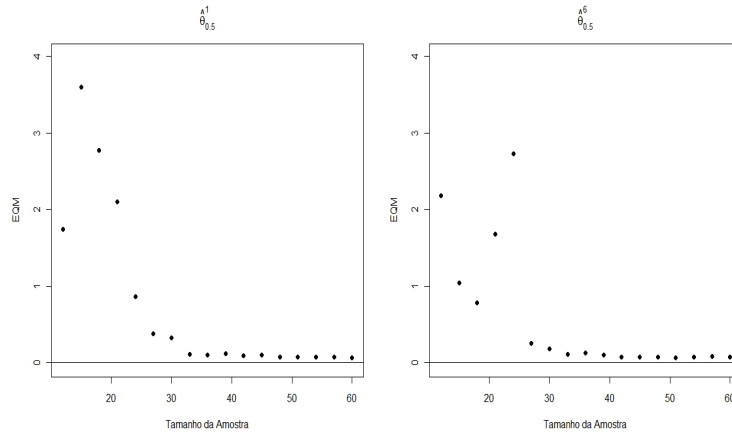


Figura 249 EQM para cada tamanho amostral e os estimadores $\hat{\theta}_{\Gamma}^1$ e $\hat{\theta}_{\Gamma}^6$, amostras geradas considerando uma **Distribuição Normal** com $\mu = 3,25$ e $\sigma^2 = 1$ para $\Gamma = 0,50$ e o **DILW**

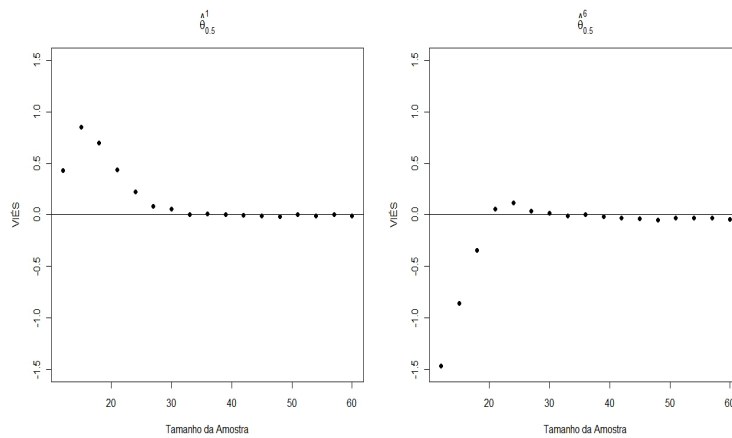


Figura 250 Viés para cada tamanho amostral e os estimadores $\hat{\theta}_{\Gamma}^1$ e $\hat{\theta}_{\Gamma}^6$, amostras geradas considerando uma **Distribuição Normal** com $\mu = 3,25$ e $\sigma^2 = 1$ para $\Gamma = 0,50$ e o **DILW**

APÊNDICE O

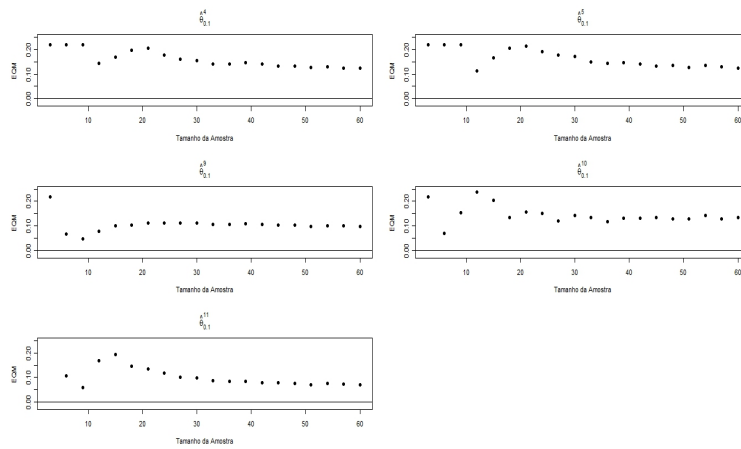


Figura 251 EQM para cada tamanho amostral e os estimadores $\hat{\theta}_{\Gamma}^4$, $\hat{\theta}_{\Gamma}^5$, $\hat{\theta}_{\Gamma}^9$, $\hat{\theta}_{\Gamma}^{10}$ e $\hat{\theta}_{\Gamma}^{11}$, amostras geradas considerando uma **Distribuição Normal** com $\mu = 2,25$ e $\sigma^2 = 1$ para $\Gamma = 0,10$ e o **DILW**

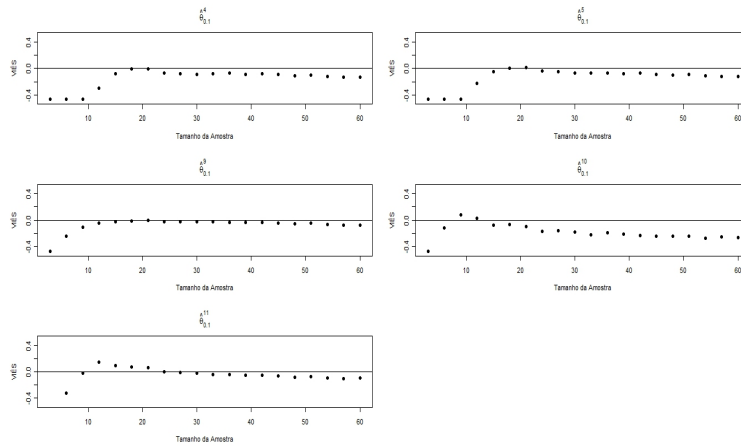


Figura 252 Viés para cada tamanho amostral e os estimadores $\hat{\theta}_{\Gamma}^4$, $\hat{\theta}_{\Gamma}^5$, $\hat{\theta}_{\Gamma}^9$, $\hat{\theta}_{\Gamma}^{10}$ e $\hat{\theta}_{\Gamma}^{11}$, amostras geradas considerando uma **Distribuição Normal** com $\mu = 2,25$ e $\sigma^2 = 1$ para $\Gamma = 0,10$ e o **DILW**

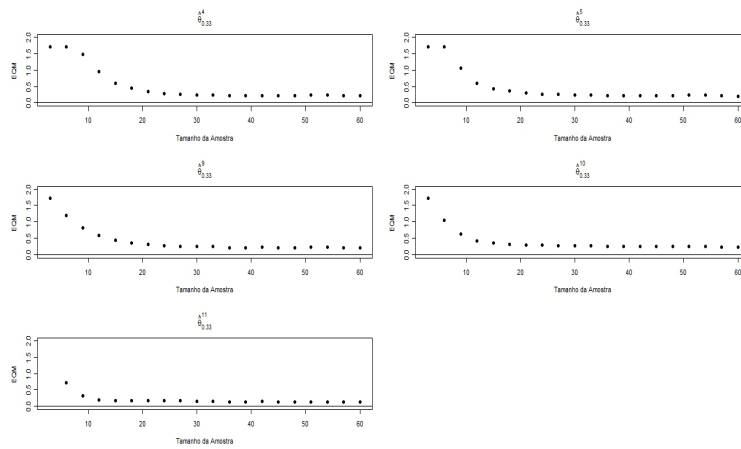


Figura 253 EQM para cada tamanho amostral e os estimadores $\hat{\theta}_{\Gamma}^4$, $\hat{\theta}_{\Gamma}^5$, $\hat{\theta}_{\Gamma}^9$, $\hat{\theta}_{\Gamma}^{10}$ e $\hat{\theta}_{\Gamma}^{11}$, amostras geradas considerando uma **Distribuição Normal** com $\mu = 2,25$ e $\sigma^2 = 1$ para $\Gamma = 0,33$ e o **DILW**

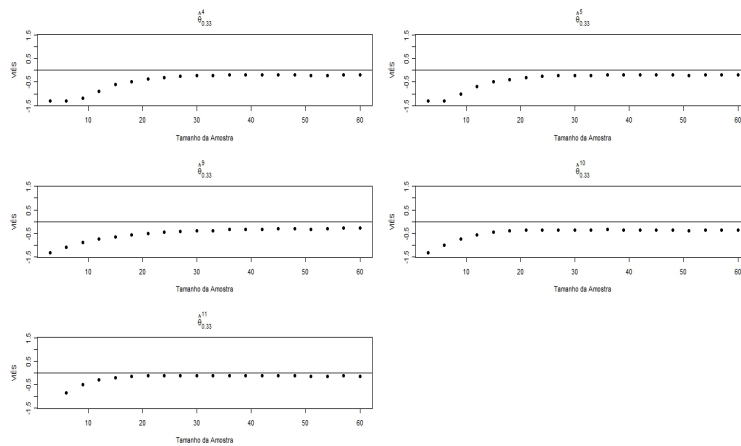


Figura 254 Viés para cada tamanho amostral e os estimadores $\hat{\theta}_{\Gamma}^4$, $\hat{\theta}_{\Gamma}^5$, $\hat{\theta}_{\Gamma}^9$, $\hat{\theta}_{\Gamma}^{10}$ e $\hat{\theta}_{\Gamma}^{11}$, amostras geradas considerando uma **Distribuição Normal** com $\mu = 2,25$ e $\sigma^2 = 1$ para $\Gamma = 0,33$ e o **DILW**

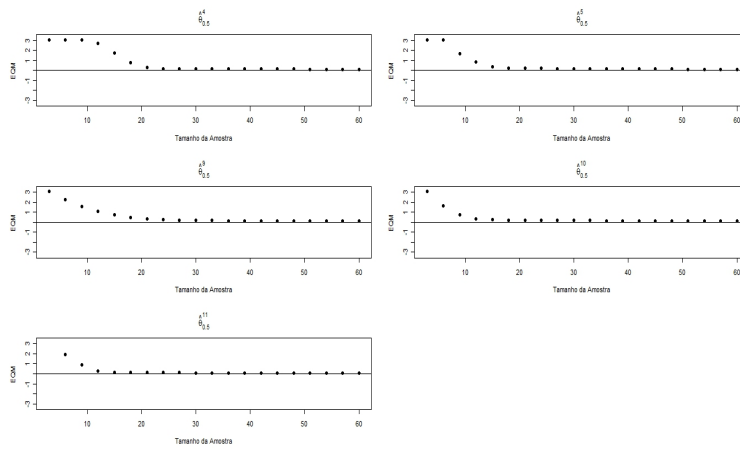


Figura 255 EQM para cada tamanho amostral e os estimadores $\hat{\theta}_{\Gamma}^4$, $\hat{\theta}_{\Gamma}^5$, $\hat{\theta}_{\Gamma}^9$, $\hat{\theta}_{\Gamma}^{10}$ e $\hat{\theta}_{\Gamma}^{11}$, amostras geradas considerando uma **Distribuição Normal** com $\mu = 2,25$ e $\sigma^2 = 1$ para $\Gamma = 0,50$ e o **DILW**

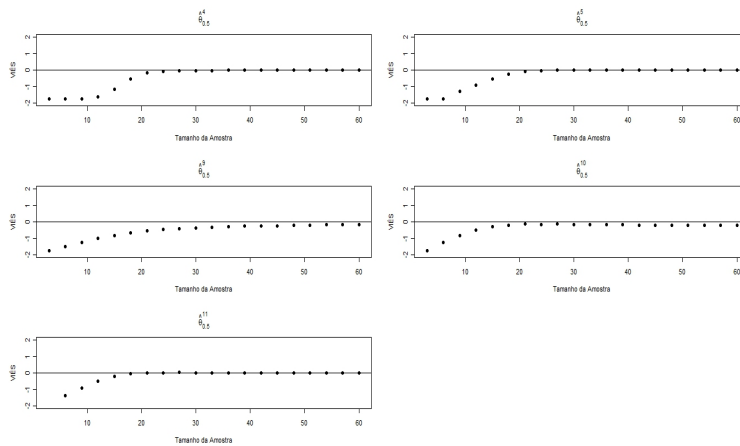


Figura 256 Viés para cada tamanho amostral e os estimadores $\hat{\theta}_{\Gamma}^4$, $\hat{\theta}_{\Gamma}^5$, $\hat{\theta}_{\Gamma}^9$, $\hat{\theta}_{\Gamma}^{10}$ e $\hat{\theta}_{\Gamma}^{11}$, amostras geradas considerando uma **Distribuição Normal** com $\mu = 2,25$ e $\sigma^2 = 1$ para $\Gamma = 0,50$ e o **DILW**

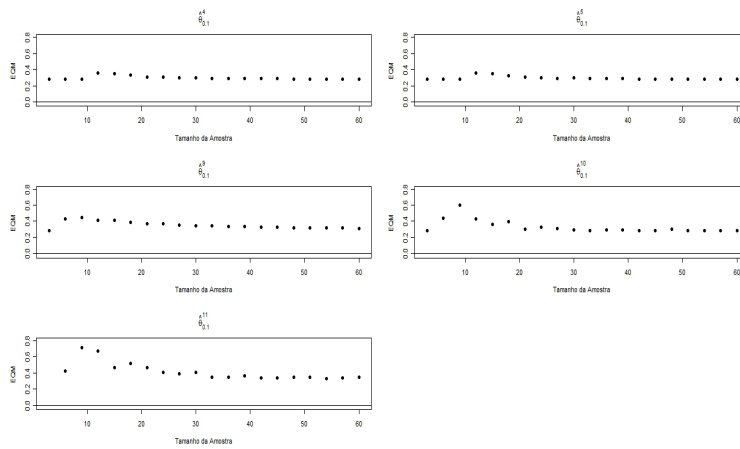


Figura 257 EQM para cada tamanho amostral e os estimadores $\hat{\theta}_{\Gamma}^4$, $\hat{\theta}_{\Gamma}^5$, $\hat{\theta}_{\Gamma}^9$, $\hat{\theta}_{\Gamma}^{10}$ e $\hat{\theta}_{\Gamma}^{11}$, amostras geradas considerando uma **Distribuição Normal** com $\mu = 1,25$ e $\sigma^2 = 1$ para $\Gamma = 0,10$ e o **DILW**

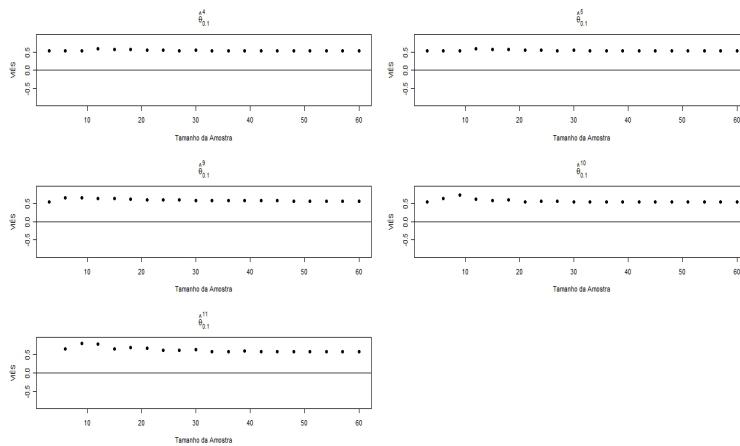


Figura 258 Viés para cada tamanho amostral e os estimadores $\hat{\theta}_{\Gamma}^4$, $\hat{\theta}_{\Gamma}^5$, $\hat{\theta}_{\Gamma}^9$, $\hat{\theta}_{\Gamma}^{10}$ e $\hat{\theta}_{\Gamma}^{11}$, amostras geradas considerando uma **Distribuição Normal** com $\mu = 1,25$ e $\sigma^2 = 1$ para $\Gamma = 0,10$ e o **DILW**

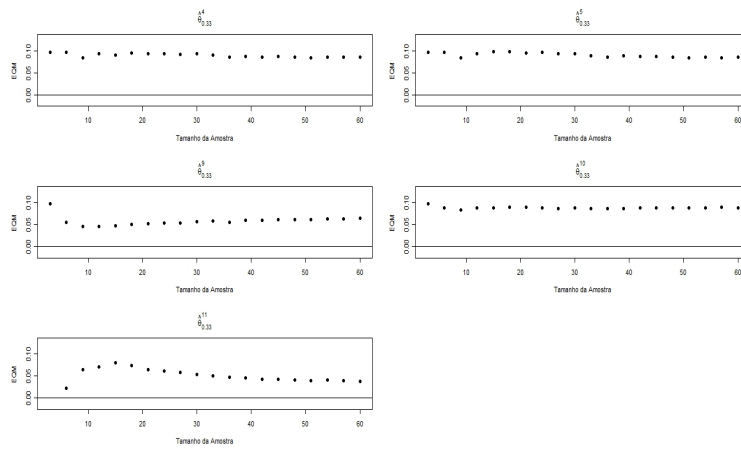


Figura 259 EQM para cada tamanho amostral e os estimadores $\hat{\theta}_{\Gamma}^4$, $\hat{\theta}_{\Gamma}^5$, $\hat{\theta}_{\Gamma}^9$, $\hat{\theta}_{\Gamma}^{10}$ e $\hat{\theta}_{\Gamma}^{11}$, amostras geradas considerando uma **Distribuição Normal** com $\mu = 1,25$ e $\sigma^2 = 1$ para $\Gamma = 0,33$ e o **DILW**

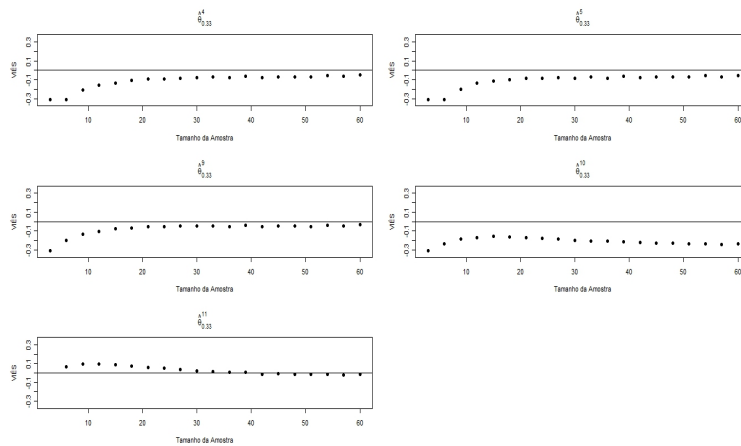


Figura 260 Viés para cada tamanho amostral e os estimadores $\hat{\theta}_{\Gamma}^4$, $\hat{\theta}_{\Gamma}^5$, $\hat{\theta}_{\Gamma}^9$, $\hat{\theta}_{\Gamma}^{10}$ e $\hat{\theta}_{\Gamma}^{11}$, amostras geradas considerando uma **Distribuição Normal** com $\mu = 1,25$ e $\sigma^2 = 1$ para $\Gamma = 0,33$ e o **DILW**

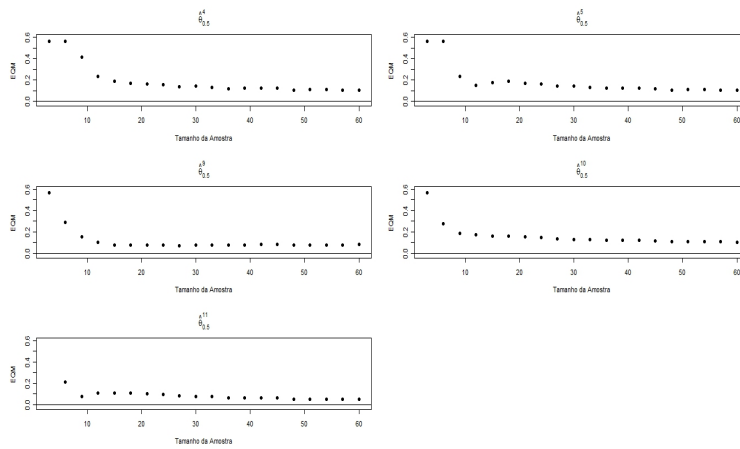


Figura 261 EQM para cada tamanho amostral e os estimadores $\hat{\theta}_{\Gamma}^4$, $\hat{\theta}_{\Gamma}^5$, $\hat{\theta}_{\Gamma}^9$, $\hat{\theta}_{\Gamma}^{10}$ e $\hat{\theta}_{\Gamma}^{11}$, amostras geradas considerando uma **Distribuição Normal** com $\mu = 1,25$ e $\sigma^2 = 1$ para $\Gamma = 0,50$ e o **DILW**

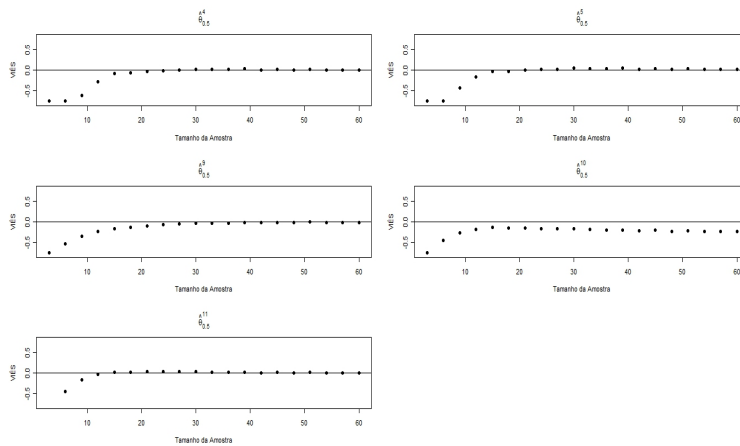


Figura 262 Viés para cada tamanho amostral e os estimadores $\hat{\theta}_{\Gamma}^4$, $\hat{\theta}_{\Gamma}^5$, $\hat{\theta}_{\Gamma}^9$, $\hat{\theta}_{\Gamma}^{10}$ e $\hat{\theta}_{\Gamma}^{11}$, amostras geradas considerando uma **Distribuição Normal** com $\mu = 1,25$ e $\sigma^2 = 1$ para $\Gamma = 0,50$ e o **DILW**

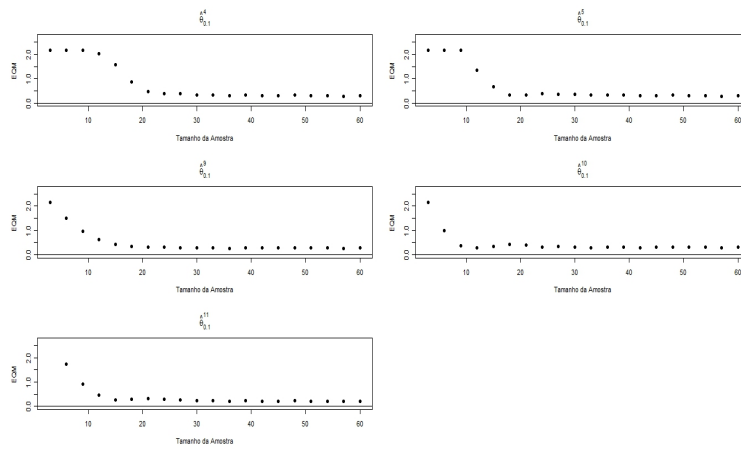


Figura 263 EQM para cada tamanho amostral e os estimadores $\hat{\theta}_{\Gamma}^4$, $\hat{\theta}_{\Gamma}^5$, $\hat{\theta}_{\Gamma}^9$, $\hat{\theta}_{\Gamma}^{10}$ e $\hat{\theta}_{\Gamma}^{11}$, amostras geradas considerando uma **Distribuição Normal** com $\mu = 3,25$ e $\sigma^2 = 1$ para $\Gamma = 0,10$ e o **DILW**

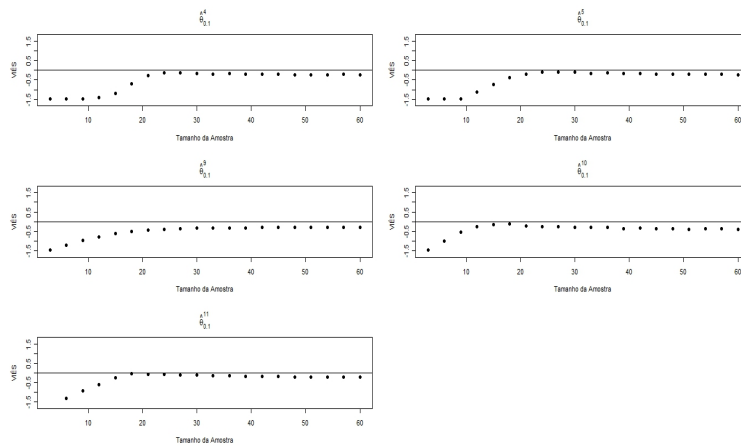


Figura 264 Viés para cada tamanho amostral e os estimadores $\hat{\theta}_{\Gamma}^4$, $\hat{\theta}_{\Gamma}^5$, $\hat{\theta}_{\Gamma}^9$, $\hat{\theta}_{\Gamma}^{10}$ e $\hat{\theta}_{\Gamma}^{11}$, amostras geradas considerando uma **Distribuição Normal** com $\mu = 3,25$ e $\sigma^2 = 1$ para $\Gamma = 0,10$ e o **DILW**

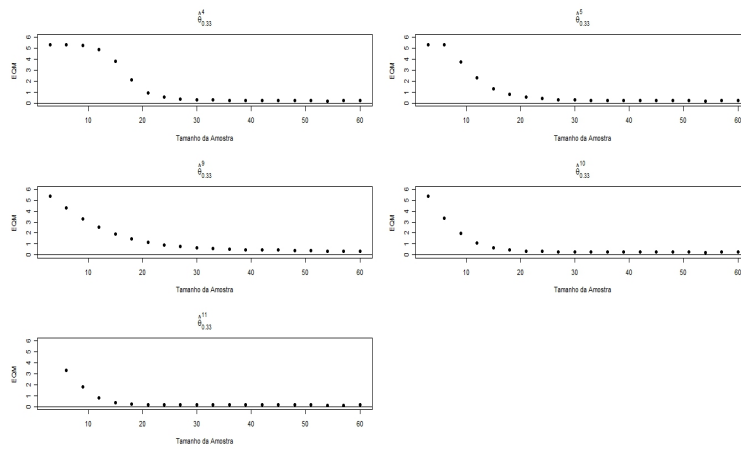


Figura 265 EQM para cada tamanho amostral e os estimadores $\hat{\theta}_{\Gamma}^4$, $\hat{\theta}_{\Gamma}^5$, $\hat{\theta}_{\Gamma}^9$, $\hat{\theta}_{\Gamma}^{10}$ e $\hat{\theta}_{\Gamma}^{11}$, amostras geradas considerando uma **Distribuição Normal** com $\mu = 3,25$ e $\sigma^2 = 1$ para $\Gamma = 0,33$ e o **DILW**

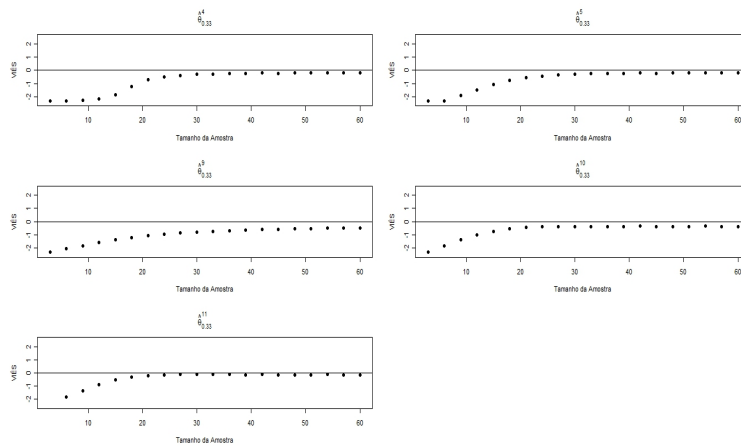


Figura 266 Viés para cada tamanho amostral e os estimadores $\hat{\theta}_{\Gamma}^4$, $\hat{\theta}_{\Gamma}^5$, $\hat{\theta}_{\Gamma}^9$, $\hat{\theta}_{\Gamma}^{10}$ e $\hat{\theta}_{\Gamma}^{11}$, amostras geradas considerando uma **Distribuição Normal** com $\mu = 3,25$ e $\sigma^2 = 1$ para $\Gamma = 0,33$ e o **DILW**

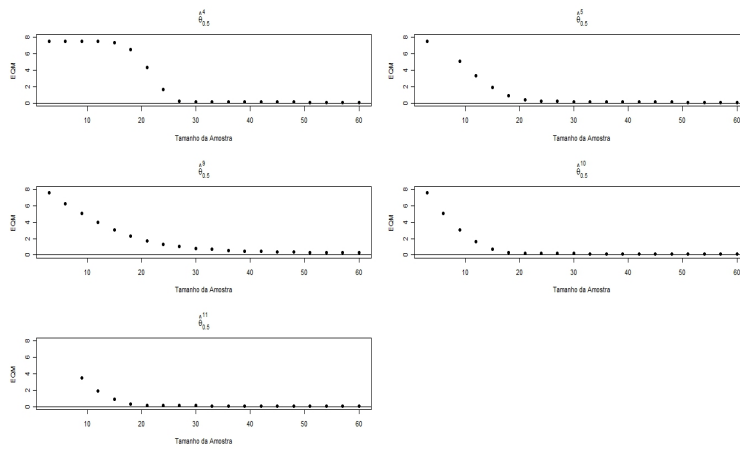


Figura 267 EQM para cada tamanho amostral e os estimadores $\hat{\theta}_{\Gamma}^4$, $\hat{\theta}_{\Gamma}^5$, $\hat{\theta}_{\Gamma}^9$, $\hat{\theta}_{\Gamma}^{10}$ e $\hat{\theta}_{\Gamma}^{11}$, amostras geradas considerando uma **Distribuição Normal** com $\mu = 3,25$ e $\sigma^2 = 1$ para $\Gamma = 0,50$ e o **DILW**

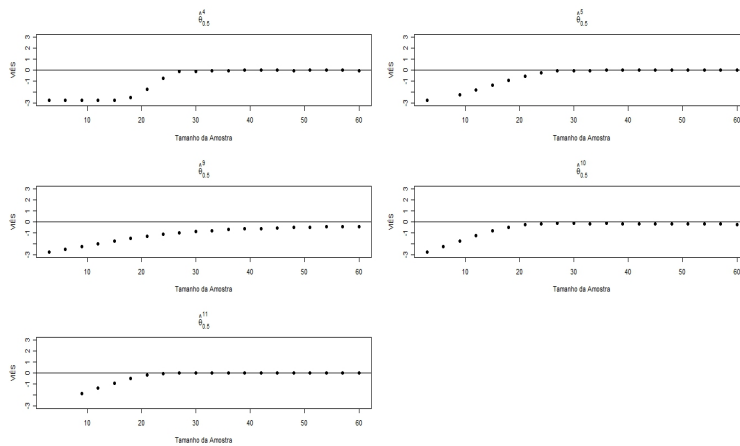


Figura 268 Viés para cada tamanho amostral e os estimadores $\hat{\theta}_{\Gamma}^4$, $\hat{\theta}_{\Gamma}^5$, $\hat{\theta}_{\Gamma}^9$, $\hat{\theta}_{\Gamma}^{10}$ e $\hat{\theta}_{\Gamma}^{11}$, amostras geradas considerando uma **Distribuição Normal** com $\mu = 3,25$ e $\sigma^2 = 1$ para $\Gamma = 0,50$ e o **DILW**

APÊNDICE P

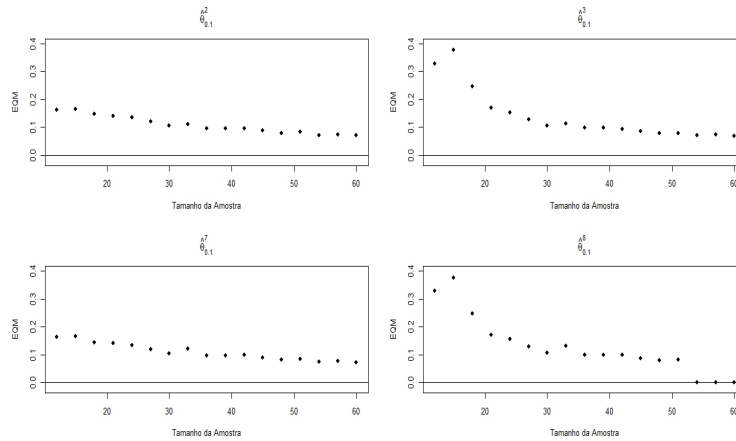


Figura 269 EQM para cada tamanho amostral e os estimadores $\hat{\theta}_{\Gamma}^2$, $\hat{\theta}_{\Gamma}^3$, $\hat{\theta}_{\Gamma}^7$ e $\hat{\theta}_{\Gamma}^8$, amostras geradas considerando uma **Distribuição Logística** com $\mu = 2,25$ e $\sigma^2 = 1$ para $\Gamma = 0,10$ e o **DILW**

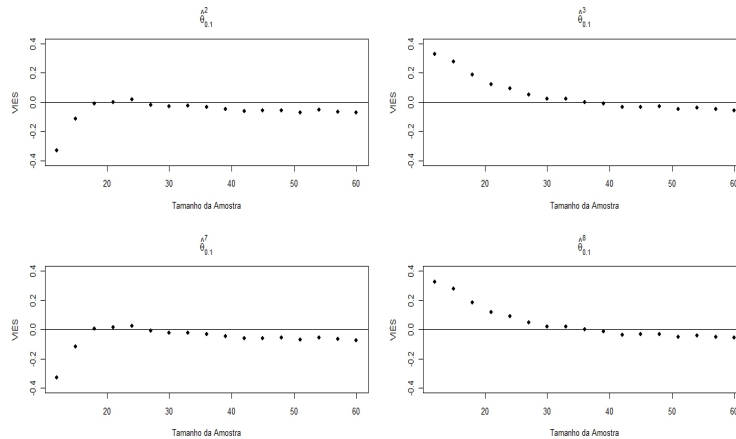


Figura 270 Viés para cada tamanho amostral e os estimadores $\hat{\theta}_{\Gamma}^2$, $\hat{\theta}_{\Gamma}^3$, $\hat{\theta}_{\Gamma}^7$ e $\hat{\theta}_{\Gamma}^8$, amostras geradas considerando uma **Distribuição Logística** com $\mu = 2,25$ e $\sigma^2 = 1$ para $\Gamma = 0,10$ e o **DILW**

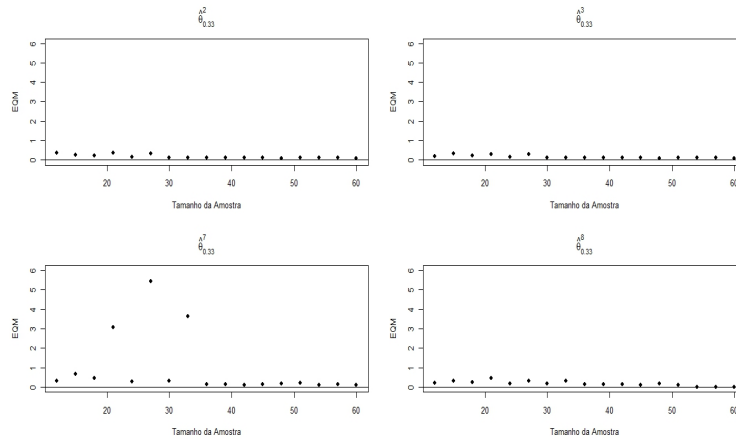


Figura 271 EQM para cada tamanho amostral e os estimadores $\hat{\theta}_{\Gamma}^2$, $\hat{\theta}_{\Gamma}^3$, $\hat{\theta}_{\Gamma}^7$ e $\hat{\theta}_{\Gamma}^8$, amostras geradas considerando uma **Distribuição Logística** com $\mu = 2,25$ e $\sigma^2 = 1$ para $\Gamma = 0,33$ e o **DILW**

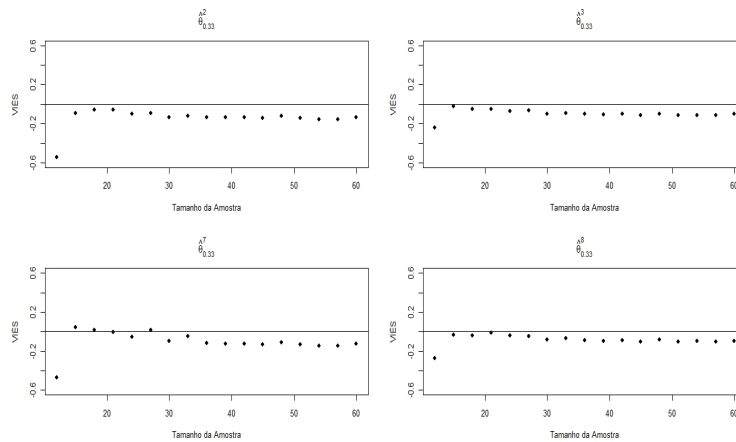


Figura 272 Viés para cada tamanho amostral e os estimadores $\hat{\theta}_{\Gamma}^2$, $\hat{\theta}_{\Gamma}^3$, $\hat{\theta}_{\Gamma}^7$ e $\hat{\theta}_{\Gamma}^8$, amostras geradas considerando uma **Distribuição Logística** com $\mu = 2,25$ e $\sigma^2 = 1$ para $\Gamma = 0,33$ e o **DILW**

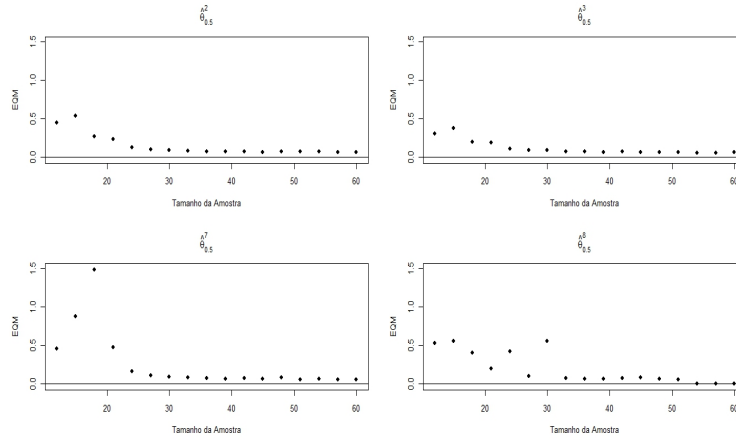


Figura 273 EQM para cada tamanho amostral e os estimadores $\hat{\theta}_{\Gamma}^2$, $\hat{\theta}_{\Gamma}^3$, $\hat{\theta}_{\Gamma}^7$ e $\hat{\theta}_{\Gamma}^8$, amostras geradas considerando uma **Distribuição Logística** com $\mu = 2,25$ e $\sigma^2 = 1$ para $\Gamma = 0,50$ e o **DILW**

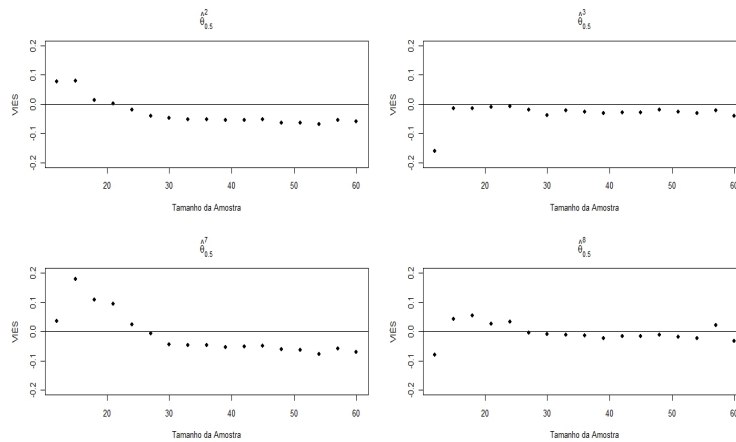


Figura 274 Viés para cada tamanho amostral e os estimadores $\hat{\theta}_{\Gamma}^2$, $\hat{\theta}_{\Gamma}^3$, $\hat{\theta}_{\Gamma}^7$ e $\hat{\theta}_{\Gamma}^8$, amostras geradas considerando uma **Distribuição Logística** com $\mu = 2,25$ e $\sigma^2 = 1$ para $\Gamma = 0,50$ e o **DILW**

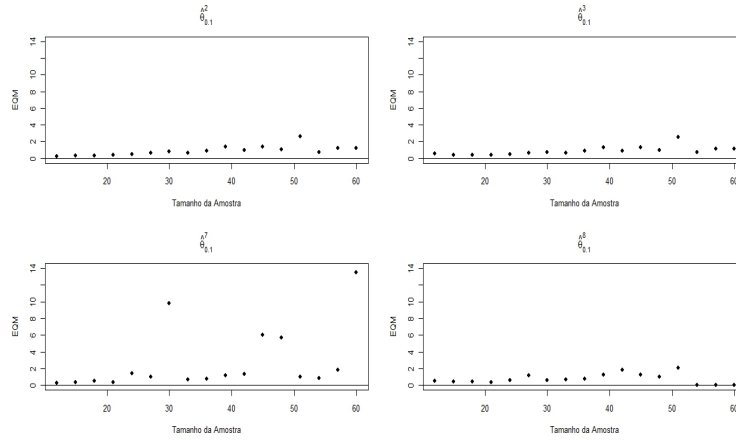


Figura 275 EQM para cada tamanho amostral e os estimadores $\hat{\theta}_{\Gamma}^2$, $\hat{\theta}_{\Gamma}^3$, $\hat{\theta}_{\Gamma}^7$ e $\hat{\theta}_{\Gamma}^8$, amostras geradas considerando uma **Distribuição Logística** com $\mu = 1,25$ e $\sigma^2 = 1$ para $\Gamma = 0,10$ e o **DILW**

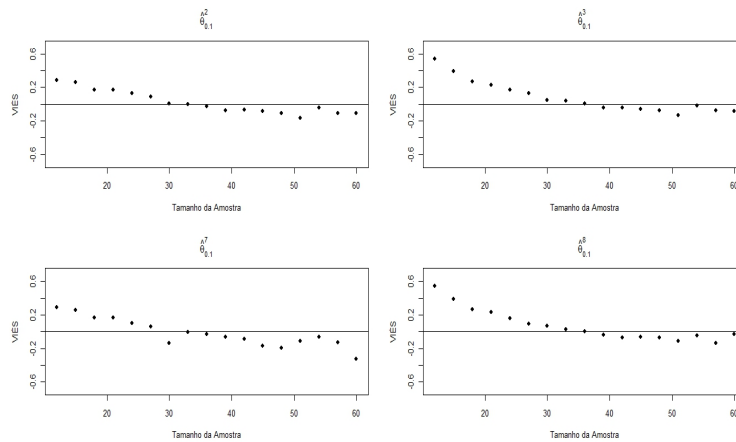


Figura 276 Viés para cada tamanho amostral e os estimadores $\hat{\theta}_{\Gamma}^2$, $\hat{\theta}_{\Gamma}^3$, $\hat{\theta}_{\Gamma}^7$ e $\hat{\theta}_{\Gamma}^8$, amostras geradas considerando uma **Distribuição Logística** com $\mu = 1,25$ e $\sigma^2 = 1$ para $\Gamma = 0,10$ e o **DILW**

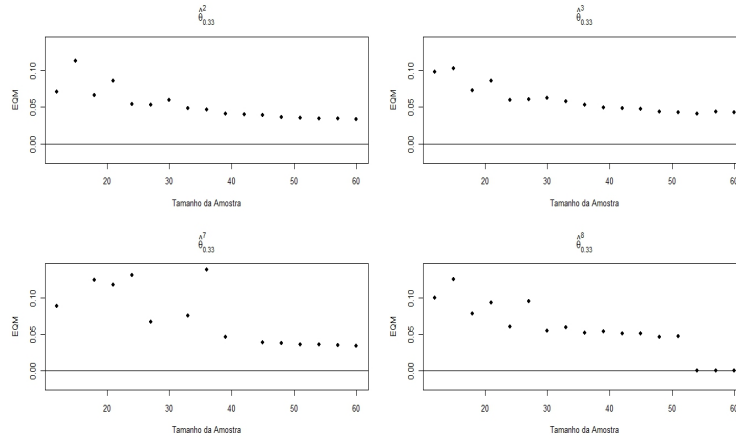


Figura 277 EQM para cada tamanho amostral e os estimadores $\hat{\theta}_{\Gamma}^2$, $\hat{\theta}_{\Gamma}^3$, $\hat{\theta}_{\Gamma}^7$ e $\hat{\theta}_{\Gamma}^8$, amostras geradas considerando uma **Distribuição Logística** com $\mu = 1,25$ e $\sigma^2 = 1$ para $\Gamma = 0,33$ e o **DILW**

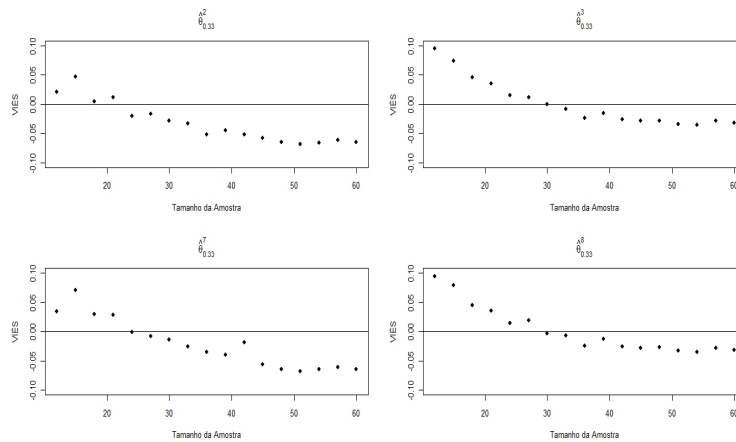


Figura 278 Viés para cada tamanho amostral e os estimadores $\hat{\theta}_{\Gamma}^2$, $\hat{\theta}_{\Gamma}^3$, $\hat{\theta}_{\Gamma}^7$ e $\hat{\theta}_{\Gamma}^8$, amostras geradas considerando uma **Distribuição Logística** com $\mu = 1,25$ e $\sigma^2 = 1$ para $\Gamma = 0,33$ e o **DILW**

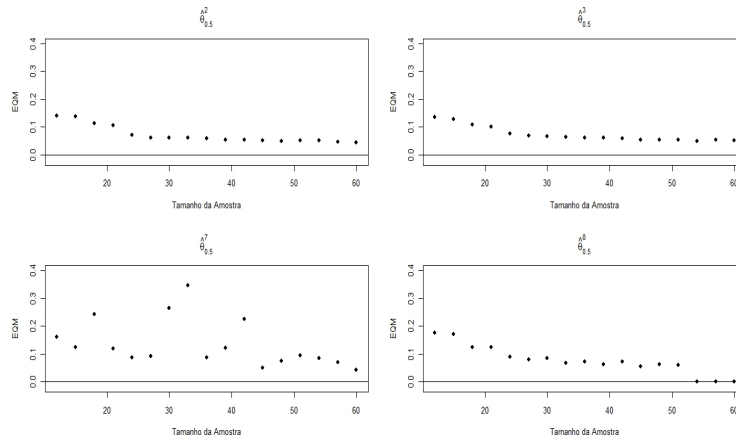


Figura 279 EQM para cada tamanho amostral e os estimadores $\hat{\theta}_{\Gamma}^2$, $\hat{\theta}_{\Gamma}^3$, $\hat{\theta}_{\Gamma}^7$ e $\hat{\theta}_{\Gamma}^8$, amostras geradas considerando uma **Distribuição Logística** com $\mu = 1,25$ e $\sigma^2 = 1$ para $\Gamma = 0,50$ e o **DILW**

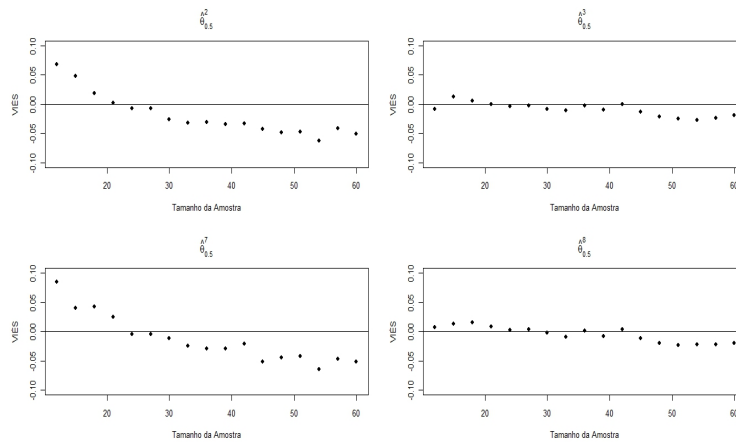


Figura 280 Viés para cada tamanho amostral e os estimadores $\hat{\theta}_{\Gamma}^2$, $\hat{\theta}_{\Gamma}^3$, $\hat{\theta}_{\Gamma}^7$ e $\hat{\theta}_{\Gamma}^8$, amostras geradas considerando uma **Distribuição Logística** com $\mu = 1,25$ e $\sigma^2 = 1$ para $\Gamma = 0,50$ e o **DILW**

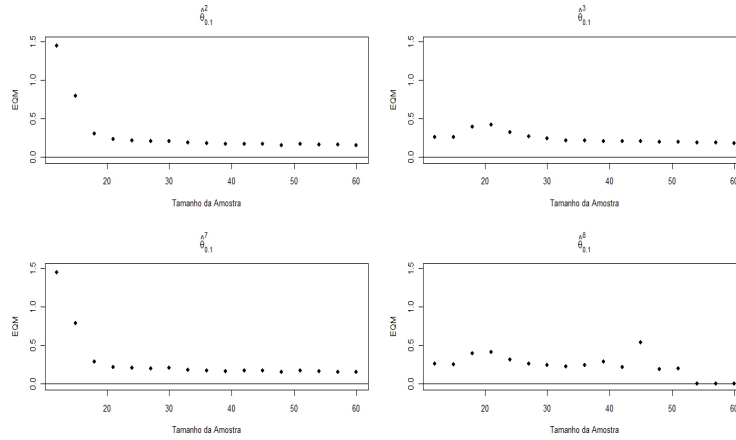


Figura 281 EQM para cada tamanho amostral e os estimadores $\hat{\theta}_{\Gamma}^2$, $\hat{\theta}_{\Gamma}^3$, $\hat{\theta}_{\Gamma}^7$ e $\hat{\theta}_{\Gamma}^8$, amostras geradas considerando uma **Distribuição Logística** com $\mu = 3,25$ e $\sigma^2 = 1$ para $\Gamma = 0,10$ e o **DILW**

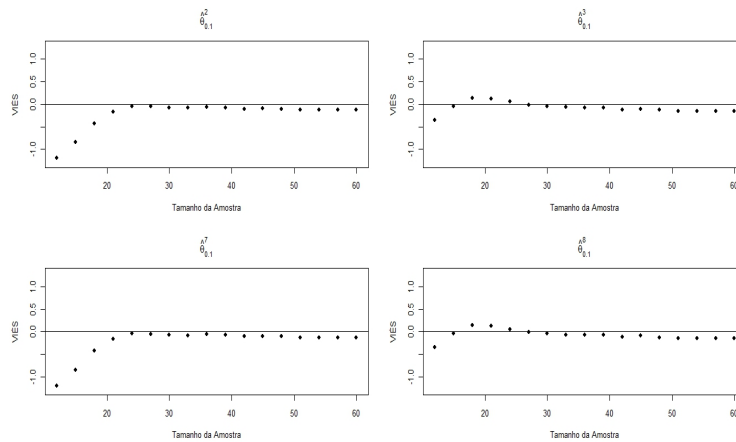


Figura 282 Viés para cada tamanho amostral e os estimadores $\hat{\theta}_{\Gamma}^2$, $\hat{\theta}_{\Gamma}^3$, $\hat{\theta}_{\Gamma}^7$ e $\hat{\theta}_{\Gamma}^8$, amostras geradas considerando uma **Distribuição Logística** com $\mu = 3,25$ e $\sigma^2 = 1$ para $\Gamma = 0,10$ e o **DILW**

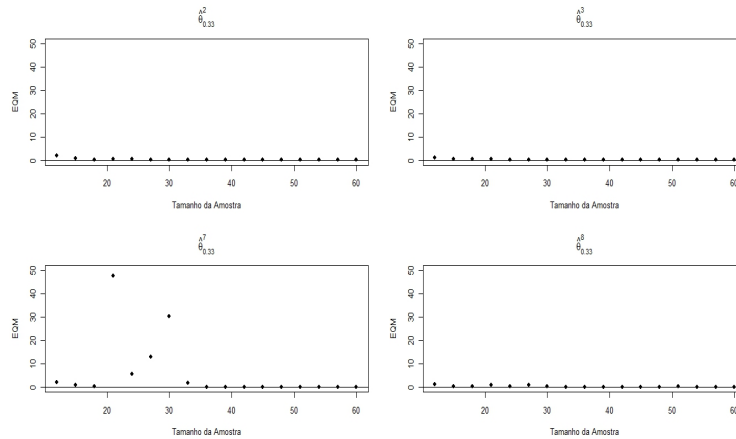


Figura 283 EQM para cada tamanho amostral e os estimadores $\hat{\theta}_{\Gamma}^2$, $\hat{\theta}_{\Gamma}^3$, $\hat{\theta}_{\Gamma}^7$ e $\hat{\theta}_{\Gamma}^8$, amostras geradas considerando uma **Distribuição Logística** com $\mu = 3,25$ e $\sigma^2 = 1$ para $\Gamma = 0,33$ e o **DILW**

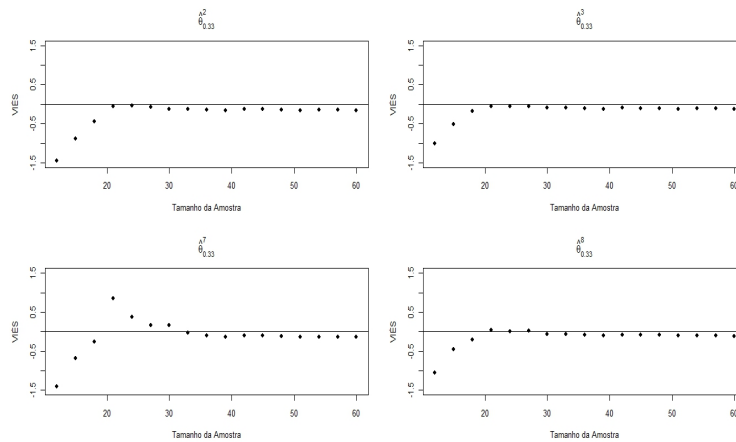


Figura 284 Viés para cada tamanho amostral e os estimadores $\hat{\theta}_{\Gamma}^2$, $\hat{\theta}_{\Gamma}^3$, $\hat{\theta}_{\Gamma}^7$ e $\hat{\theta}_{\Gamma}^8$, amostras geradas considerando uma **Distribuição Logística** com $\mu = 3,25$ e $\sigma^2 = 1$ para $\Gamma = 0,33$ e o **DILW**



Figura 285 EQM para cada tamanho amostral e os estimadores $\hat{\theta}_{\Gamma}^2$, $\hat{\theta}_{\Gamma}^3$, $\hat{\theta}_{\Gamma}^7$ e $\hat{\theta}_{\Gamma}^8$, amostras geradas considerando uma **Distribuição Logística** com $\mu = 3,25$ e $\sigma^2 = 1$ para $\Gamma = 0,50$ e o **DILW**

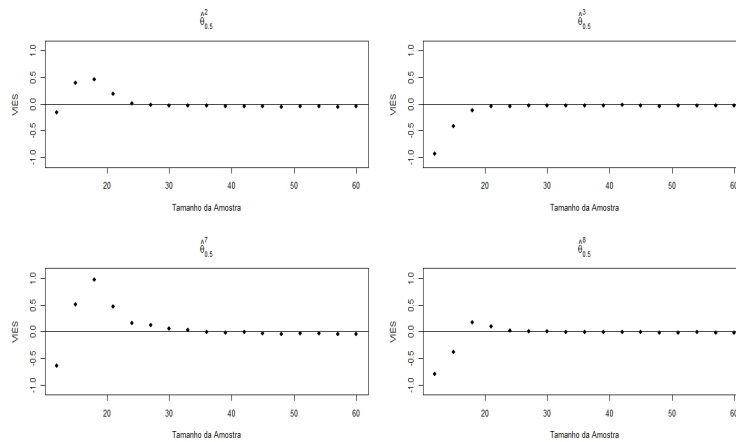


Figura 286 Viés para cada tamanho amostral e os estimadores $\hat{\theta}_{\Gamma}^2$, $\hat{\theta}_{\Gamma}^3$, $\hat{\theta}_{\Gamma}^7$ e $\hat{\theta}_{\Gamma}^8$, amostras geradas considerando uma **Distribuição Logística** com $\mu = 3,25$ e $\sigma^2 = 1$ para $\Gamma = 0,50$ e o **DILW**

APÊNDICE Q

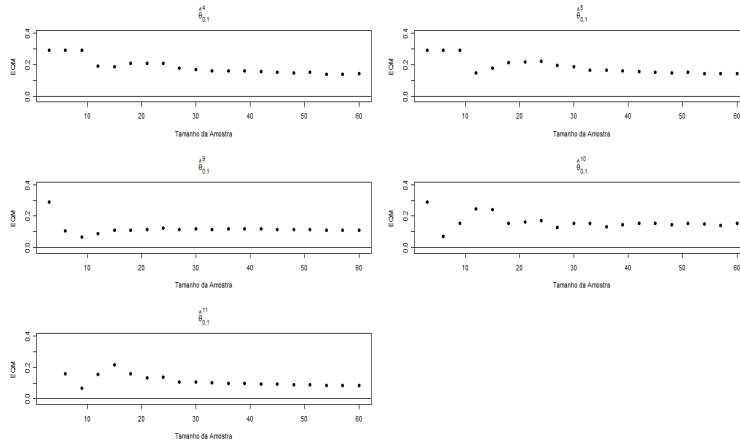


Figura 287 EQM para cada tamanho amostral e os estimadores $\hat{\theta}_{\Gamma}^4$, $\hat{\theta}_{\Gamma}^5$, $\hat{\theta}_{\Gamma}^9$, $\hat{\theta}_{\Gamma}^{10}$ e $\hat{\theta}_{\Gamma}^{11}$, amostras geradas considerando uma **Distribuição Logística** com $\mu = 2,25$ e $\sigma^2 = 1$ para $\Gamma = 0,10$ e o **DILW**

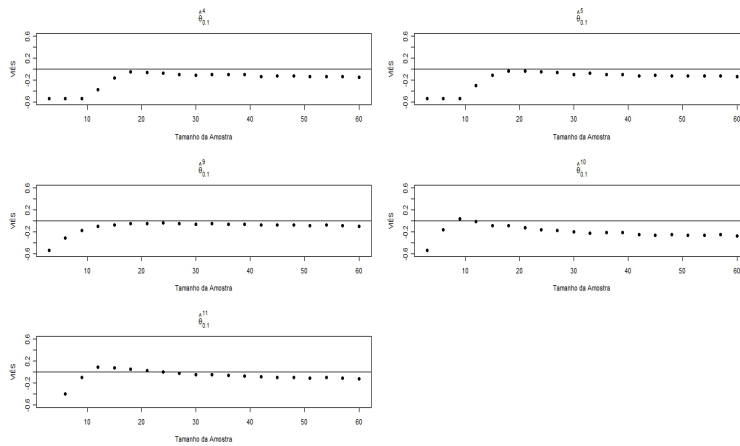


Figura 288 Viés para cada tamanho amostral e os estimadores $\hat{\theta}_{\Gamma}^4$, $\hat{\theta}_{\Gamma}^5$, $\hat{\theta}_{\Gamma}^9$, $\hat{\theta}_{\Gamma}^{10}$ e $\hat{\theta}_{\Gamma}^{11}$, amostras geradas considerando uma **Distribuição Logística** com $\mu = 2,25$ e $\sigma^2 = 1$ para $\Gamma = 0,10$ e o **DILW**

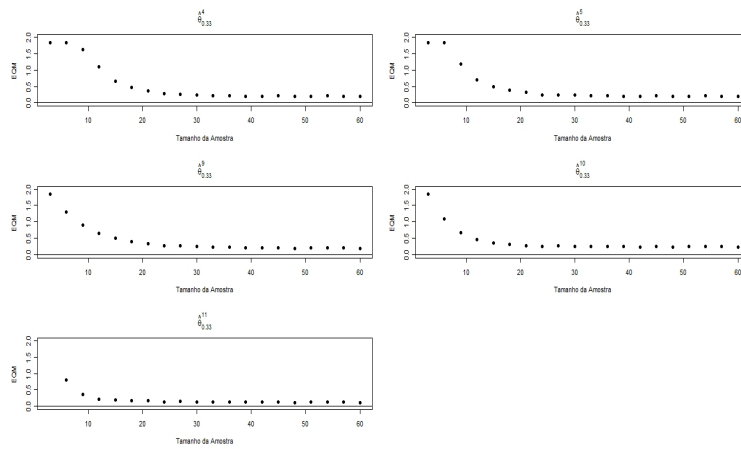


Figura 289 EQM para cada tamanho amostral e os estimadores $\hat{\theta}_{\Gamma}^4$, $\hat{\theta}_{\Gamma}^5$, $\hat{\theta}_{\Gamma}^9$, $\hat{\theta}_{\Gamma}^{10}$ e $\hat{\theta}_{\Gamma}^{11}$, amostras geradas considerando uma **Distribuição Logística** com $\mu = 2,25$ e $\sigma^2 = 1$ para $\Gamma = 0,33$ e o **DILW**

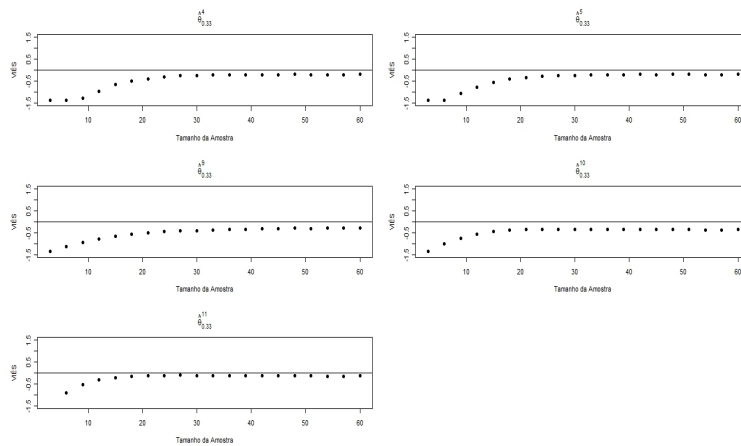


Figura 290 Viés para cada tamanho amostral e os estimadores $\hat{\theta}_{\Gamma}^4$, $\hat{\theta}_{\Gamma}^5$, $\hat{\theta}_{\Gamma}^9$, $\hat{\theta}_{\Gamma}^{10}$ e $\hat{\theta}_{\Gamma}^{11}$, amostras geradas considerando uma **Distribuição Logística** com $\mu = 2,25$ e $\sigma^2 = 1$ para $\Gamma = 0,33$ e o **DILW**

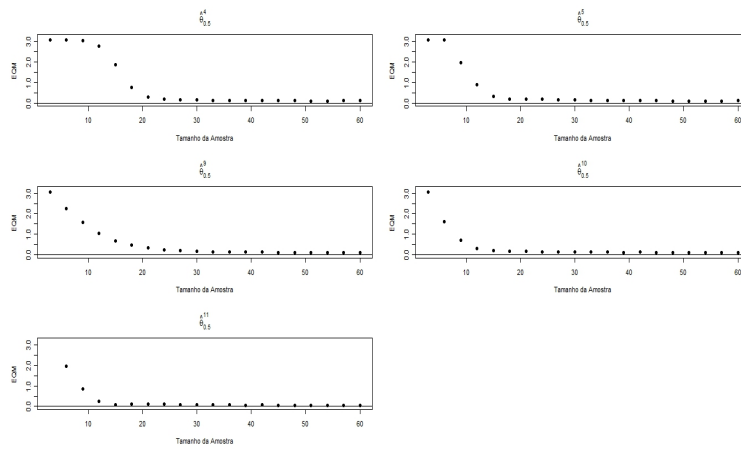


Figura 291 EQM para cada tamanho amostral e os estimadores $\hat{\theta}_{\Gamma}^4$, $\hat{\theta}_{\Gamma}^5$, $\hat{\theta}_{\Gamma}^9$, $\hat{\theta}_{\Gamma}^{10}$ e $\hat{\theta}_{\Gamma}^{11}$, amostras geradas considerando uma **Distribuição Logística** com $\mu = 2,25$ e $\sigma^2 = 1$ para $\Gamma = 0,50$ e o **DILW**

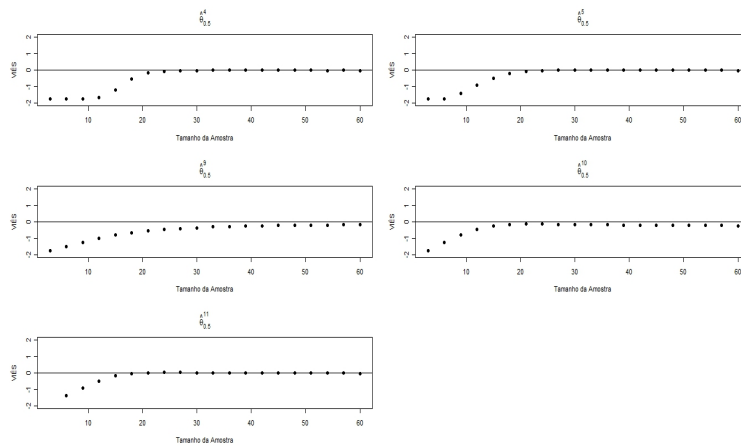


Figura 292 Viés para cada tamanho amostral e os estimadores $\hat{\theta}_{\Gamma}^4$, $\hat{\theta}_{\Gamma}^5$, $\hat{\theta}_{\Gamma}^9$, $\hat{\theta}_{\Gamma}^{10}$ e $\hat{\theta}_{\Gamma}^{11}$, amostras geradas considerando uma **Distribuição Logística** com $\mu = 2,25$ e $\sigma^2 = 1$ para $\Gamma = 0,50$ e o **DILW**

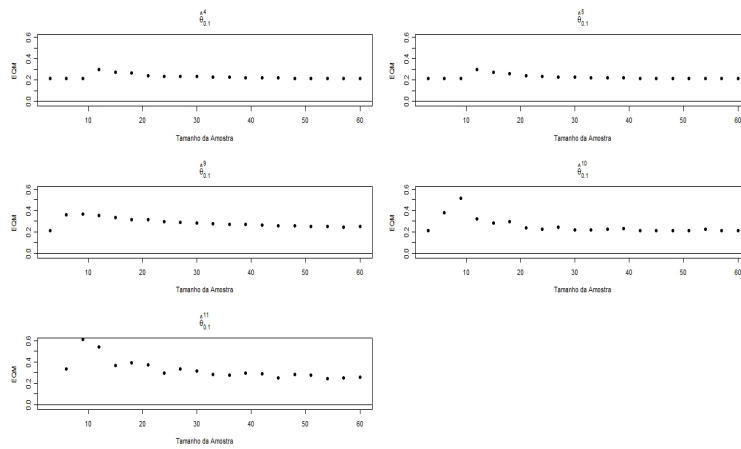


Figura 293 EQM para cada tamanho amostral e os estimadores $\hat{\theta}_{\Gamma}^4$, $\hat{\theta}_{\Gamma}^5$, $\hat{\theta}_{\Gamma}^9$, $\hat{\theta}_{\Gamma}^{10}$ e $\hat{\theta}_{\Gamma}^{11}$, amostras geradas considerando uma **Distribuição Logística** com $\mu = 1,25$ e $\sigma^2 = 1$ para $\Gamma = 0,10$ e o **DILW**

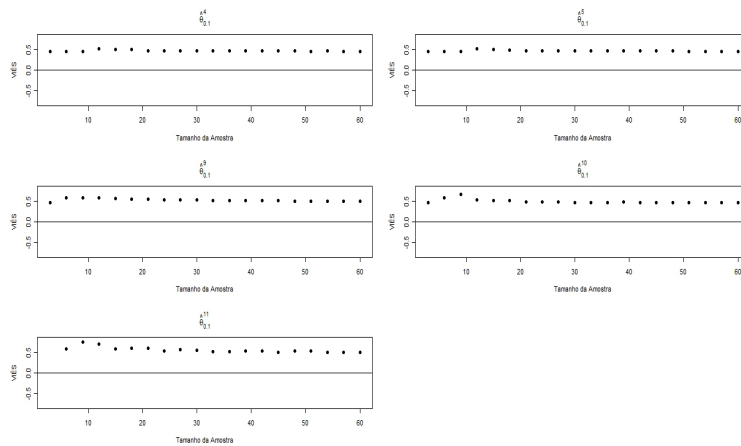


Figura 294 Viés para cada tamanho amostral e os estimadores $\hat{\theta}_{\Gamma}^4$, $\hat{\theta}_{\Gamma}^5$, $\hat{\theta}_{\Gamma}^9$, $\hat{\theta}_{\Gamma}^{10}$ e $\hat{\theta}_{\Gamma}^{11}$, amostras geradas considerando uma **Distribuição Logística** com $\mu = 1,25$ e $\sigma^2 = 1$ para $\Gamma = 0,10$ e o **DILW**

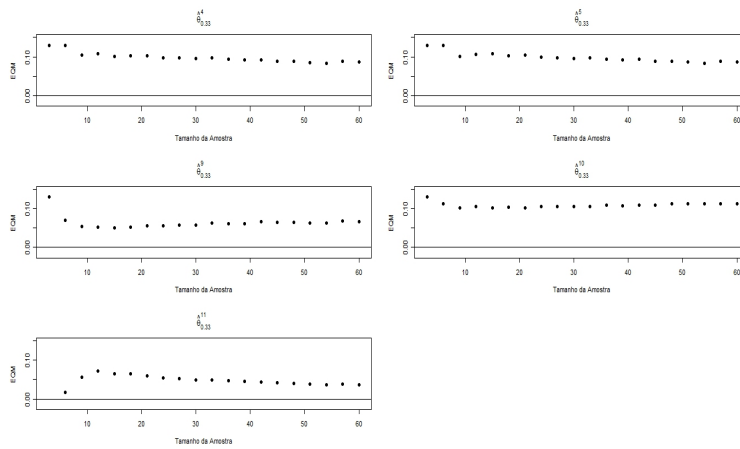


Figura 295 EQM para cada tamanho amostral e os estimadores $\hat{\theta}_{\Gamma}^4$, $\hat{\theta}_{\Gamma}^5$, $\hat{\theta}_{\Gamma}^9$, $\hat{\theta}_{\Gamma}^{10}$ e $\hat{\theta}_{\Gamma}^{11}$, amostras geradas considerando uma **Distribuição Logística** com $\mu = 1,25$ e $\sigma^2 = 1$ para $\Gamma = 0,33$ e o **DILW**

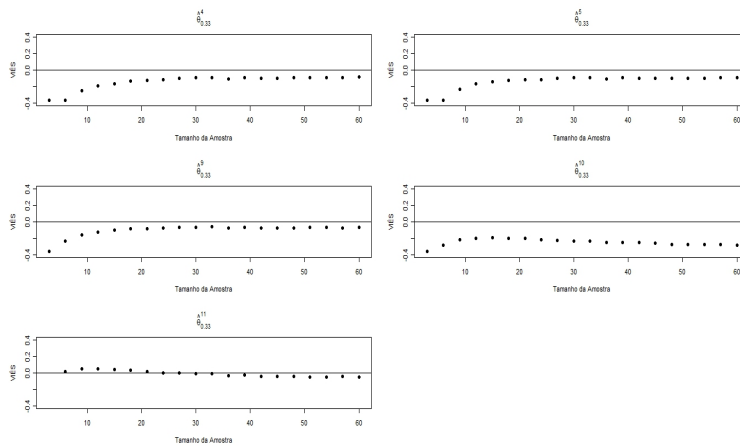


Figura 296 Viés para cada tamanho amostral e os estimadores $\hat{\theta}_{\Gamma}^4$, $\hat{\theta}_{\Gamma}^5$, $\hat{\theta}_{\Gamma}^9$, $\hat{\theta}_{\Gamma}^{10}$ e $\hat{\theta}_{\Gamma}^{11}$, amostras geradas considerando uma **Distribuição Logística** com $\mu = 1,25$ e $\sigma^2 = 1$ para $\Gamma = 0,33$ e o **DILW**

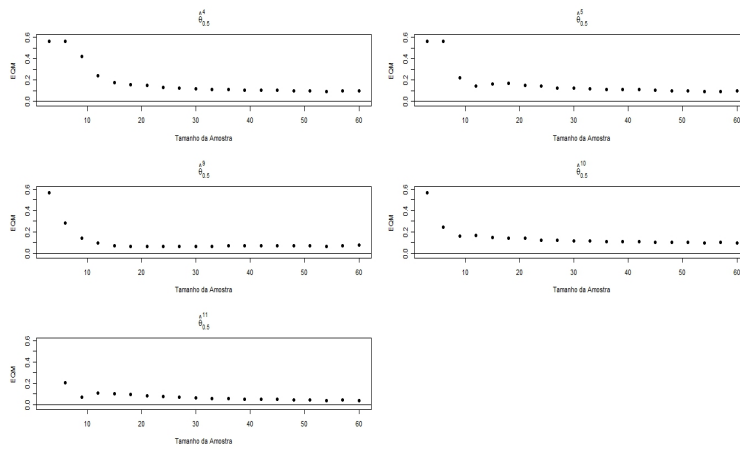


Figura 297 EQM para cada tamanho amostral e os estimadores $\hat{\theta}_{\Gamma}^4$, $\hat{\theta}_{\Gamma}^5$, $\hat{\theta}_{\Gamma}^9$, $\hat{\theta}_{\Gamma}^{10}$ e $\hat{\theta}_{\Gamma}^{11}$, amostras geradas considerando uma **Distribuição Logística** com $\mu = 1,25$ e $\sigma^2 = 1$ para $\Gamma = 0,50$ e o **DILW**

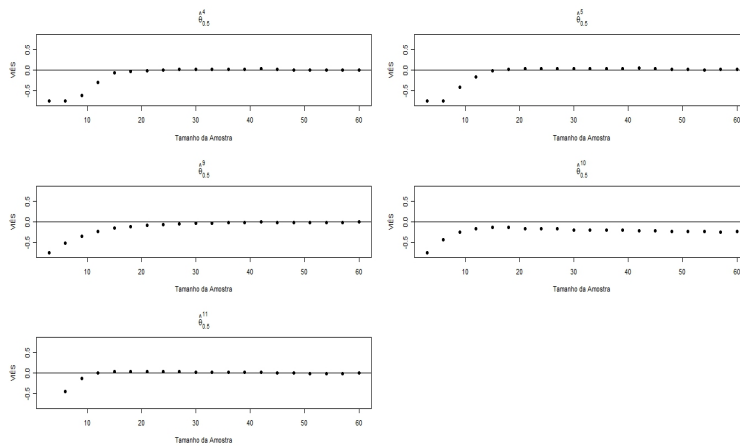


Figura 298 Viés para cada tamanho amostral e os estimadores $\hat{\theta}_{\Gamma}^4$, $\hat{\theta}_{\Gamma}^5$, $\hat{\theta}_{\Gamma}^9$, $\hat{\theta}_{\Gamma}^{10}$ e $\hat{\theta}_{\Gamma}^{11}$, amostras geradas considerando uma **Distribuição Logística** com $\mu = 1,25$ e $\sigma^2 = 1$ para $\Gamma = 0,50$ e o **DILW**

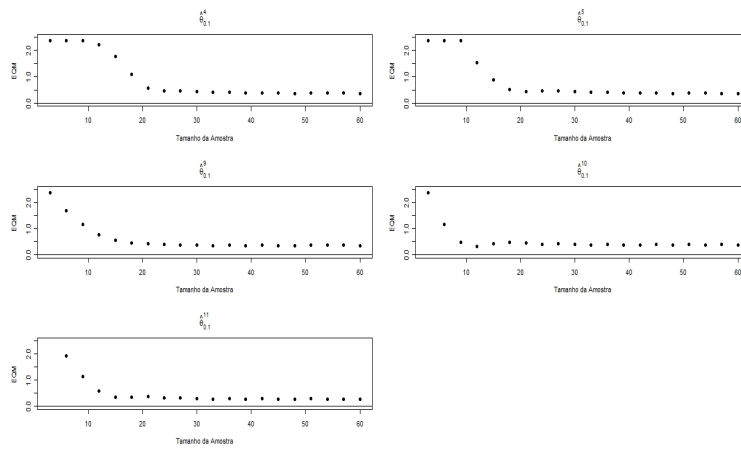


Figura 299 EQM para cada tamanho amostral e os estimadores $\hat{\theta}_{\Gamma}^4$, $\hat{\theta}_{\Gamma}^5$, $\hat{\theta}_{\Gamma}^9$, $\hat{\theta}_{\Gamma}^{10}$ e $\hat{\theta}_{\Gamma}^{11}$, amostras geradas considerando uma **Distribuição Logística** com $\mu = 3,25$ e $\sigma^2 = 1$ para $\Gamma = 0,10$ e o **DILW**

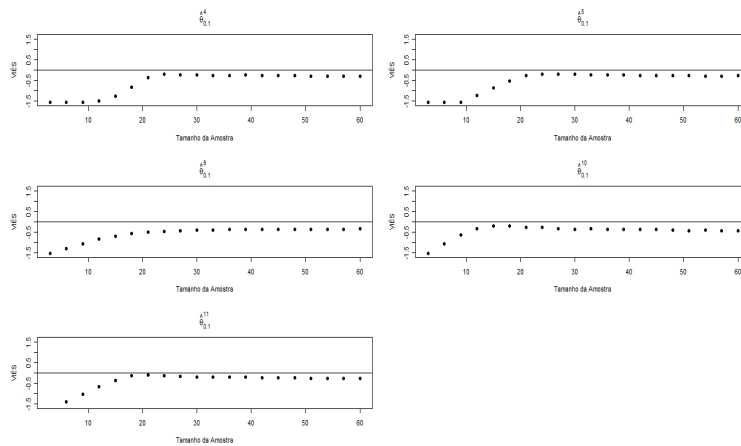


Figura 300 Viés para cada tamanho amostral e os estimadores $\hat{\theta}_{\Gamma}^4$, $\hat{\theta}_{\Gamma}^5$, $\hat{\theta}_{\Gamma}^9$, $\hat{\theta}_{\Gamma}^{10}$ e $\hat{\theta}_{\Gamma}^{11}$, amostras geradas considerando uma **Distribuição Logística** com $\mu = 3,25$ e $\sigma^2 = 1$ para $\Gamma = 0,10$ e o **DILW**

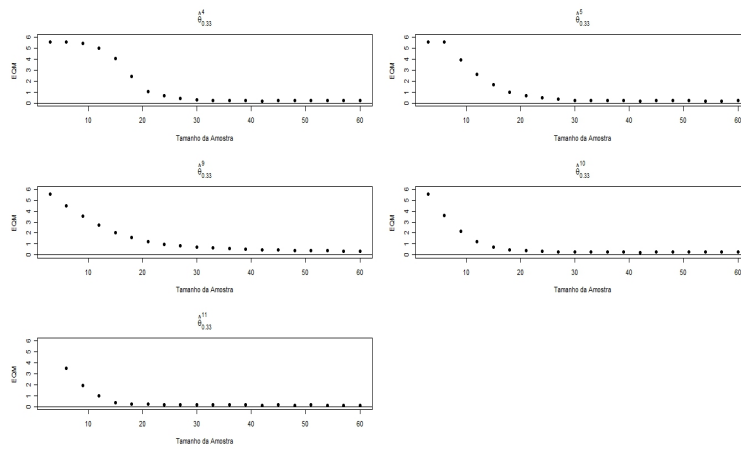


Figura 301 EQM para cada tamanho amostral e os estimadores $\hat{\theta}_{\Gamma}^4$, $\hat{\theta}_{\Gamma}^5$, $\hat{\theta}_{\Gamma}^9$, $\hat{\theta}_{\Gamma}^{10}$ e $\hat{\theta}_{\Gamma}^{11}$, amostras geradas considerando uma **Distribuição Logística** com $\mu = 3,25$ e $\sigma^2 = 1$ para $\Gamma = 0,33$ e o **DILW**

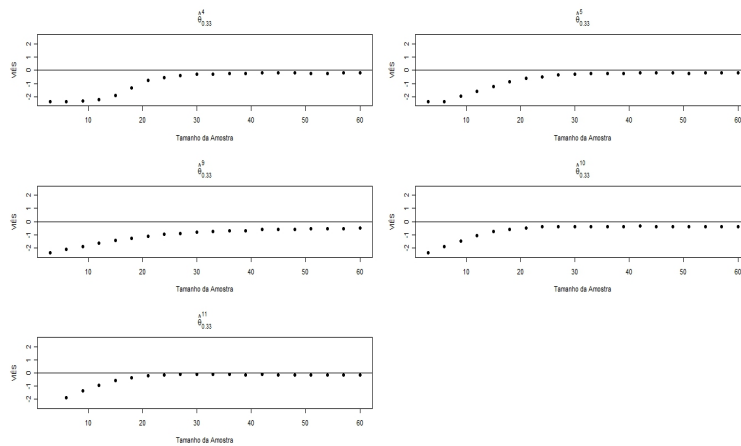


Figura 302 Viés para cada tamanho amostral e os estimadores $\hat{\theta}_{\Gamma}^4$, $\hat{\theta}_{\Gamma}^5$, $\hat{\theta}_{\Gamma}^9$, $\hat{\theta}_{\Gamma}^{10}$ e $\hat{\theta}_{\Gamma}^{11}$, amostras geradas considerando uma **Distribuição Logística** com $\mu = 3,25$ e $\sigma^2 = 1$ para $\Gamma = 0,33$ e o **DILW**

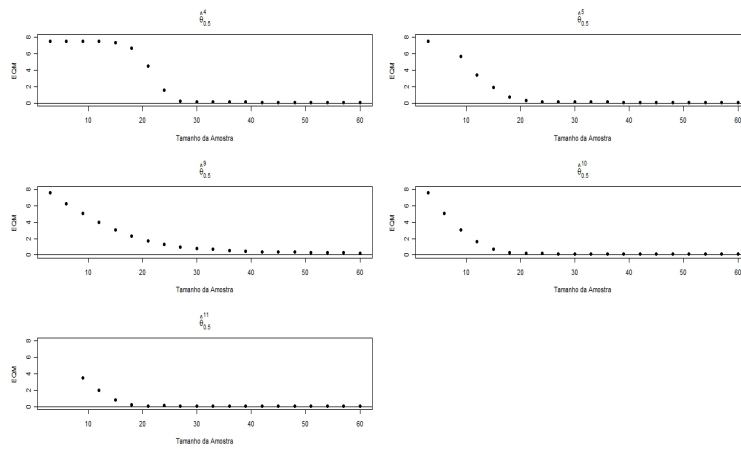


Figura 303 EQM para cada tamanho amostral e os estimadores $\hat{\theta}_{\Gamma}^4$, $\hat{\theta}_{\Gamma}^5$, $\hat{\theta}_{\Gamma}^9$, $\hat{\theta}_{\Gamma}^{10}$ e $\hat{\theta}_{\Gamma}^{11}$, amostras geradas considerando uma **Distribuição Logística** com $\mu = 3,25$ e $\sigma^2 = 1$ para $\Gamma = 0,50$ e o **DILW**

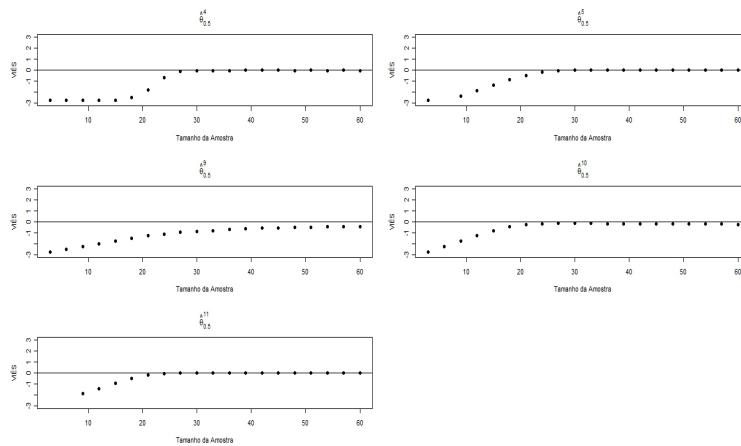


Figura 304 Viés para cada tamanho amostral e os estimadores $\hat{\theta}_{\Gamma}^4$, $\hat{\theta}_{\Gamma}^5$, $\hat{\theta}_{\Gamma}^9$, $\hat{\theta}_{\Gamma}^{10}$ e $\hat{\theta}_{\Gamma}^{11}$, amostras geradas considerando uma **Distribuição Logística** com $\mu = 3,25$ e $\sigma^2 = 1$ para $\Gamma = 0,50$ e o **DILW**

APÊNDICE R

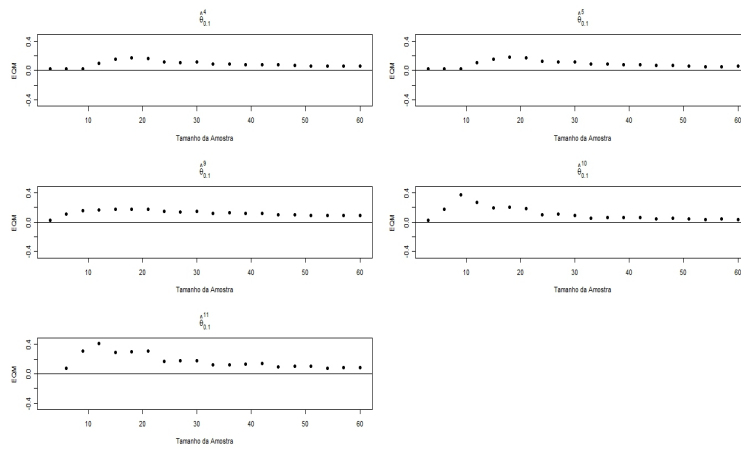


Figura 305 EQM para cada tamanho amostral e os estimadores $\hat{\theta}_{\Gamma}^4$, $\hat{\theta}_{\Gamma}^5$, $\hat{\theta}_{\Gamma}^9$, $\hat{\theta}_{\Gamma}^{10}$ e $\hat{\theta}_{\Gamma}^{11}$, amostras geradas considerando uma **Distribuição Exponencial** com parâmetro $\lambda = 0,30$, para $\Gamma = 0,10$ e o **DILW**

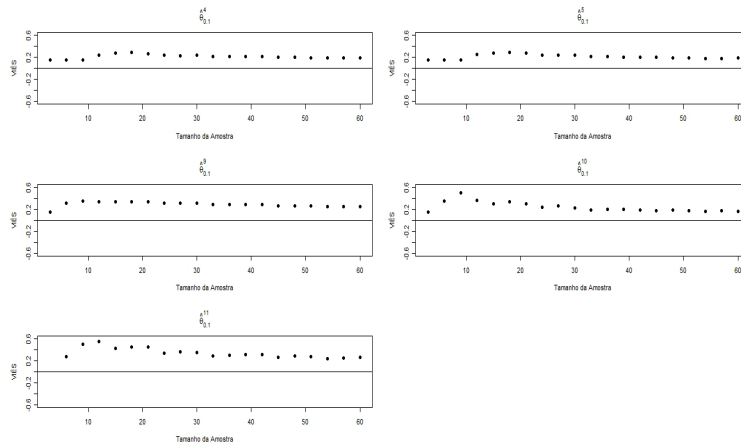


Figura 306 Viés para cada tamanho amostral e os estimadores $\hat{\theta}_{\Gamma}^4$, $\hat{\theta}_{\Gamma}^5$, $\hat{\theta}_{\Gamma}^9$, $\hat{\theta}_{\Gamma}^{10}$ e $\hat{\theta}_{\Gamma}^{11}$, amostras geradas considerando uma **Distribuição Exponencial** com parâmetro $\lambda = 0,30$, para $\Gamma = 0,10$ e o **DILW**

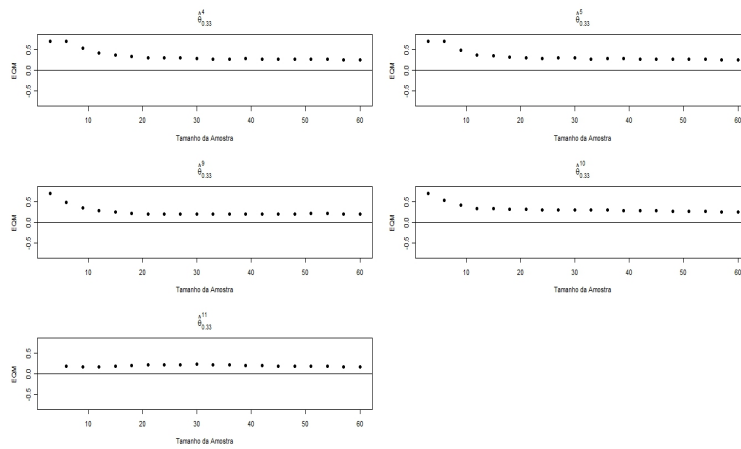


Figura 307 EQM para cada tamanho amostral e os estimadores $\hat{\theta}_{\Gamma}^4$, $\hat{\theta}_{\Gamma}^5$, $\hat{\theta}_{\Gamma}^9$, $\hat{\theta}_{\Gamma}^{10}$ e $\hat{\theta}_{\Gamma}^{11}$, amostras geradas considerando uma **Distribuição Exponencial** com parâmetro $\lambda = 0,30$, para $\Gamma = 0,33$ e o **DILW**

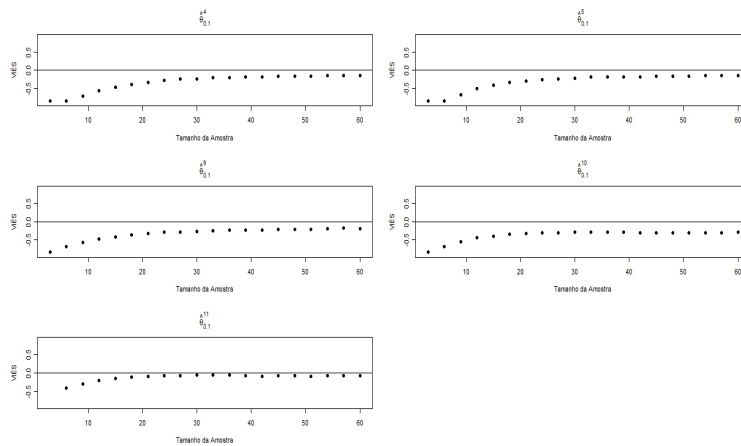


Figura 308 Viés para cada tamanho amostral e os estimadores $\hat{\theta}_{\Gamma}^4$, $\hat{\theta}_{\Gamma}^5$, $\hat{\theta}_{\Gamma}^9$, $\hat{\theta}_{\Gamma}^{10}$ e $\hat{\theta}_{\Gamma}^{11}$, amostras geradas considerando uma **Distribuição Exponencial** com parâmetro $\lambda = 0,30$, para $\Gamma = 0,33$ e o **DILW**

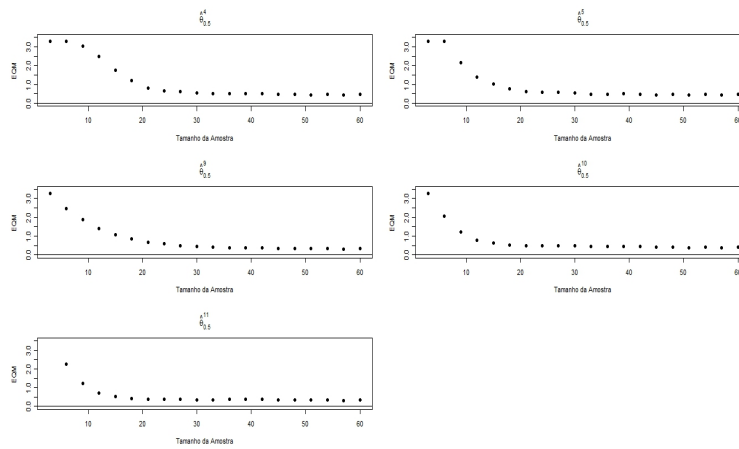


Figura 309 EQM para cada tamanho amostral e os estimadores $\hat{\theta}_{\Gamma}^4$, $\hat{\theta}_{\Gamma}^5$, $\hat{\theta}_{\Gamma}^9$, $\hat{\theta}_{\Gamma}^{10}$ e $\hat{\theta}_{\Gamma}^{11}$, amostras geradas considerando uma **Distribuição Exponencial** com parâmetro $\lambda = 0,30$, para $\Gamma = 0,50$ e o **DILW**

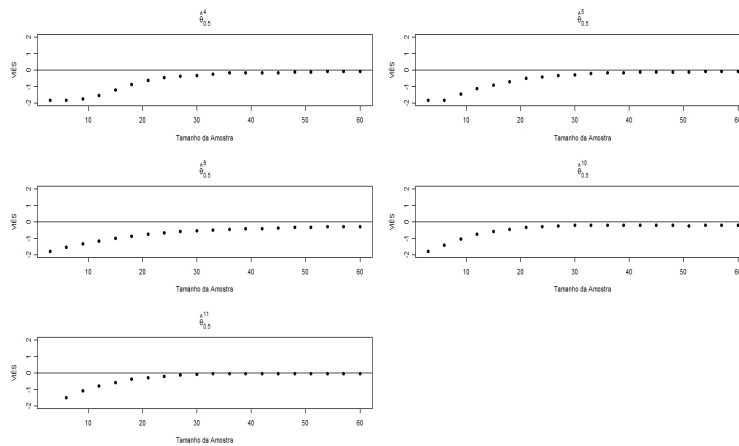


Figura 310 Viés para cada tamanho amostral e os estimadores $\hat{\theta}_{\Gamma}^4$, $\hat{\theta}_{\Gamma}^5$, $\hat{\theta}_{\Gamma}^9$, $\hat{\theta}_{\Gamma}^{10}$ e $\hat{\theta}_{\Gamma}^{11}$, amostras geradas considerando uma **Distribuição Exponencial** com parâmetro $\lambda = 0,30$, para $\Gamma = 0,50$ e o **DILW**

APÊNDICE S

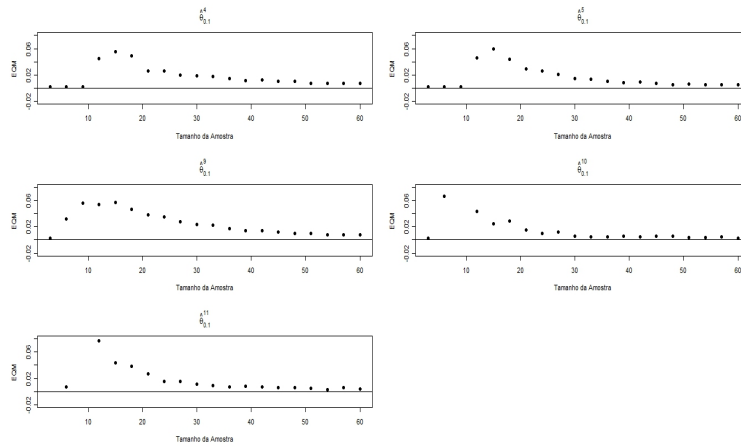


Figura 311 EQM para cada tamanho amostral e os estimadores $\hat{\theta}_{\Gamma}^4$, $\hat{\theta}_{\Gamma}^5$, $\hat{\theta}_{\Gamma}^9$, $\hat{\theta}_{\Gamma}^{10}$ e $\hat{\theta}_{\Gamma}^{11}$, amostras geradas considerando uma **Distribuição Gama** com parâmetros $\alpha = 3$ e $\lambda = 0,50$, para $\Gamma = 0,10$ e o **DILW**

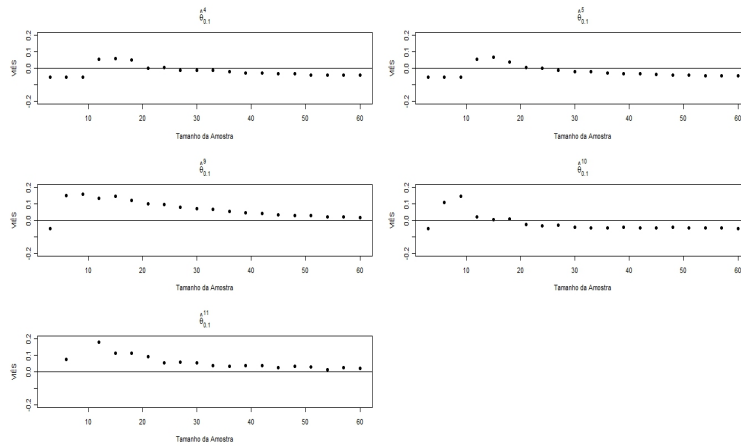


Figura 312 Viés para cada tamanho amostral e os estimadores $\hat{\theta}_{\Gamma}^4$, $\hat{\theta}_{\Gamma}^5$, $\hat{\theta}_{\Gamma}^9$, $\hat{\theta}_{\Gamma}^{10}$ e $\hat{\theta}_{\Gamma}^{11}$, amostras geradas considerando uma **Distribuição Gama** com parâmetros $\alpha = 3$ e $\lambda = 0,50$, para $\Gamma = 0,10$ e o **DILW**

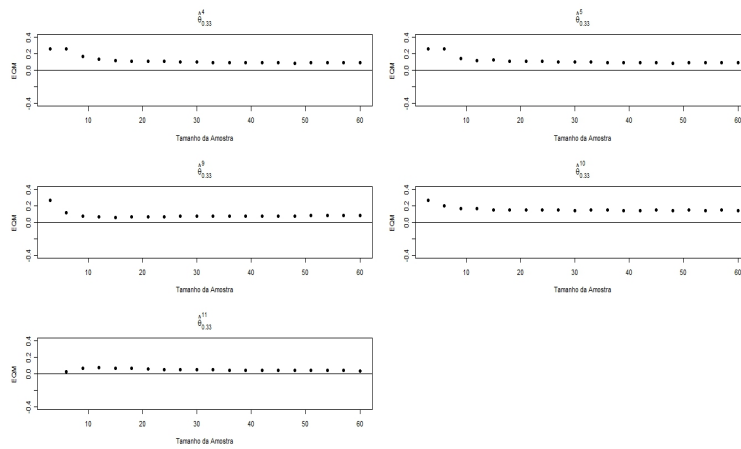


Figura 313 EQM para cada tamanho amostral e os estimadores $\hat{\theta}_{\Gamma}^4$, $\hat{\theta}_{\Gamma}^5$, $\hat{\theta}_{\Gamma}^9$, $\hat{\theta}_{\Gamma}^{10}$ e $\hat{\theta}_{\Gamma}^{11}$, amostras geradas considerando uma **Distribuição Gama** com parâmetros $\alpha = 3$ e $\lambda = 0,50$, para $\Gamma = 0,33$ e o **DILW**

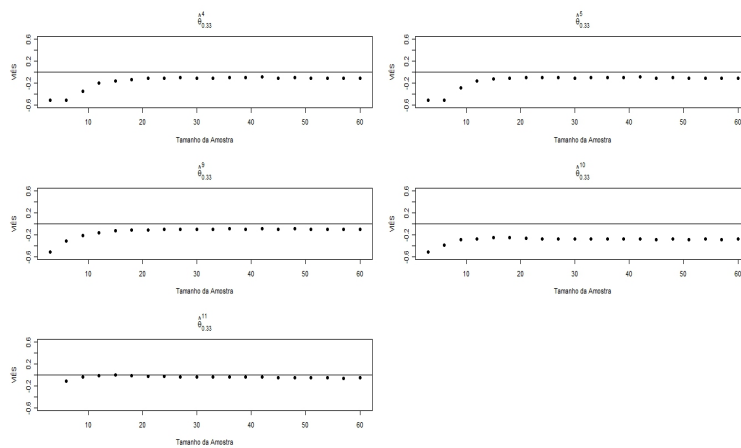


Figura 314 Viés para cada tamanho amostral e os estimadores $\hat{\theta}_{\Gamma}^4$, $\hat{\theta}_{\Gamma}^5$, $\hat{\theta}_{\Gamma}^9$, $\hat{\theta}_{\Gamma}^{10}$ e $\hat{\theta}_{\Gamma}^{11}$, amostras geradas considerando uma **Distribuição Gama** com parâmetros $\alpha = 3$ e $\lambda = 0,50$, para $\Gamma = 0,33$ e o **DILW**

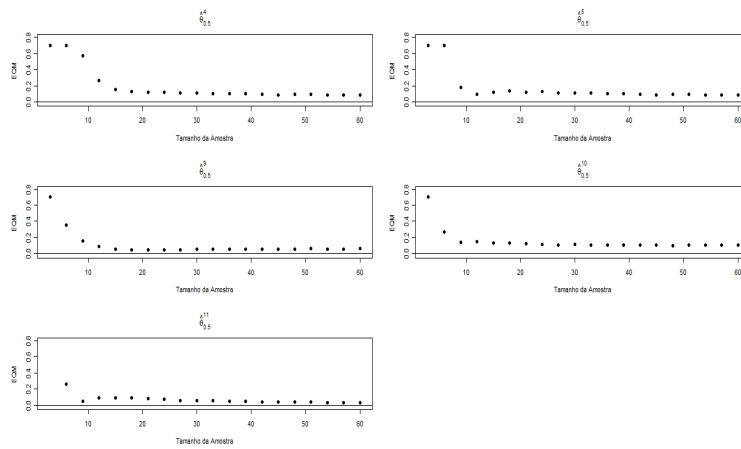


Figura 315 EQM para cada tamanho amostral e os estimadores $\hat{\theta}_{\Gamma}^4$, $\hat{\theta}_{\Gamma}^5$, $\hat{\theta}_{\Gamma}^9$, $\hat{\theta}_{\Gamma}^{10}$ e $\hat{\theta}_{\Gamma}^{11}$, amostras geradas considerando uma **Distribuição Gama** com parâmetros $\alpha = 3$ e $\lambda = 0,50$, para $\Gamma = 0,50$ e o **DILW**

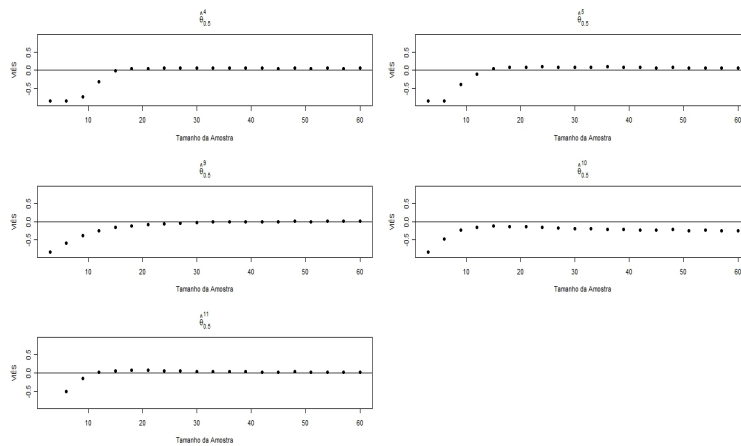


Figura 316 Viés para cada tamanho amostral e os estimadores $\hat{\theta}_{\Gamma}^4$, $\hat{\theta}_{\Gamma}^5$, $\hat{\theta}_{\Gamma}^9$, $\hat{\theta}_{\Gamma}^{10}$ e $\hat{\theta}_{\Gamma}^{11}$, amostras geradas considerando uma **Distribuição Gama** com parâmetros $\alpha = 3$ e $\lambda = 0,50$, para $\Gamma = 0,50$ e o **DILW**

APÊNDICE T

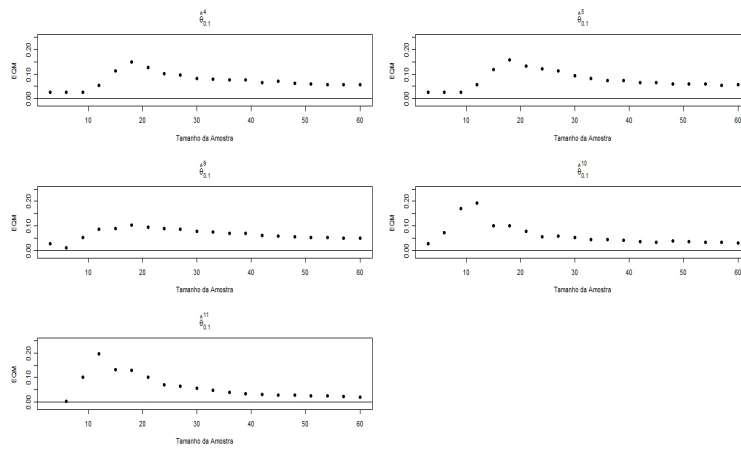


Figura 317 EQM para cada tamanho amostral e os estimadores $\hat{\theta}_{\Gamma}^4$, $\hat{\theta}_{\Gamma}^5$, $\hat{\theta}_{\Gamma}^9$, $\hat{\theta}_{\Gamma}^{10}$ e $\hat{\theta}_{\Gamma}^{11}$, amostras geradas considerando uma **Distribuição Gama** com parâmetros $\alpha = 2$ e $\lambda = 1,25$, para $\Gamma = 0,10$ e o **DILW**

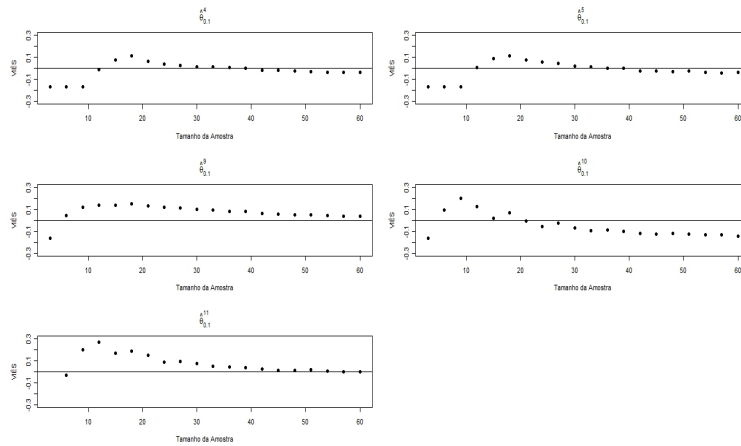


Figura 318 Viés para cada tamanho amostral e os estimadores $\hat{\theta}_{\Gamma}^4$, $\hat{\theta}_{\Gamma}^5$, $\hat{\theta}_{\Gamma}^9$, $\hat{\theta}_{\Gamma}^{10}$ e $\hat{\theta}_{\Gamma}^{11}$, amostras geradas considerando uma **Distribuição Gama** com parâmetros $\alpha = 2$ e $\lambda = 1,25$, para $\Gamma = 0,10$ e o **DILW**

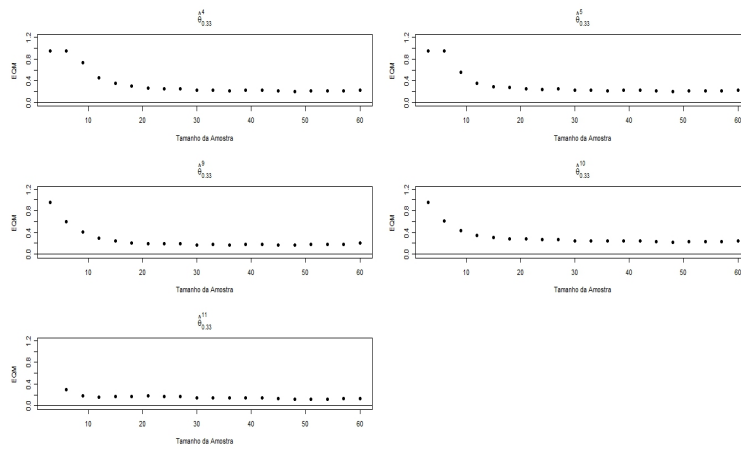


Figura 319 EQM para cada tamanho amostral e os estimadores $\hat{\theta}_{\Gamma}^4$, $\hat{\theta}_{\Gamma}^5$, $\hat{\theta}_{\Gamma}^9$, $\hat{\theta}_{\Gamma}^{10}$ e $\hat{\theta}_{\Gamma}^{11}$, amostras geradas considerando uma **Distribuição Gama** com parâmetros $\alpha = 2$ e $\lambda = 1,25$, para $\Gamma = 0,33$ e o **DILW**

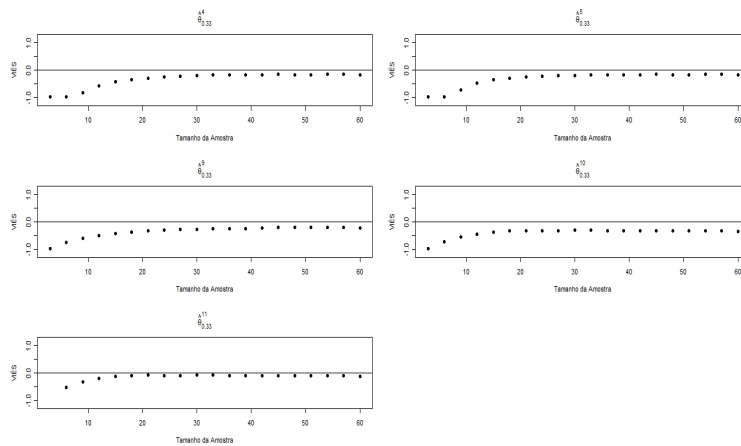


Figura 320 Viés para cada tamanho amostral e os estimadores $\hat{\theta}_{\Gamma}^4$, $\hat{\theta}_{\Gamma}^5$, $\hat{\theta}_{\Gamma}^9$, $\hat{\theta}_{\Gamma}^{10}$ e $\hat{\theta}_{\Gamma}^{11}$, amostras geradas considerando uma **Distribuição Gama** com parâmetros $\alpha = 2$ e $\lambda = 1,25$, para $\Gamma = 0,33$ e o **DILW**

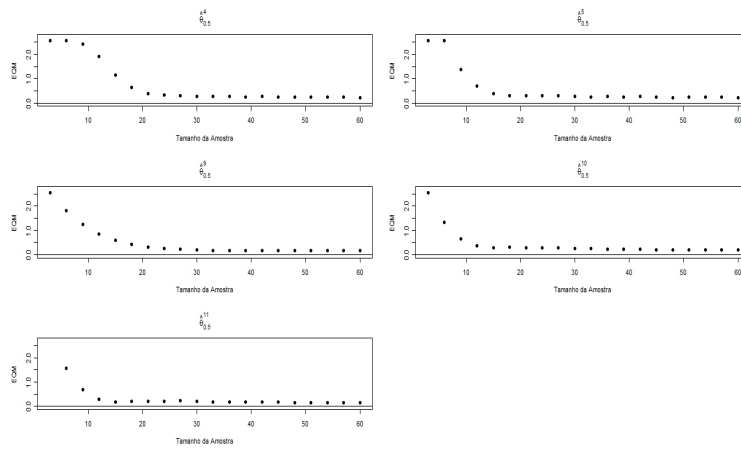


Figura 321 EQM para cada tamanho amostral e os estimadores $\hat{\theta}_{\Gamma}^4$, $\hat{\theta}_{\Gamma}^5$, $\hat{\theta}_{\Gamma}^9$, $\hat{\theta}_{\Gamma}^{10}$ e $\hat{\theta}_{\Gamma}^{11}$, amostras geradas considerando uma **Distribuição Gama** com parâmetros $\alpha = 2$ e $\lambda = 1,25$, para $\Gamma = 0,50$ e o **DILW**

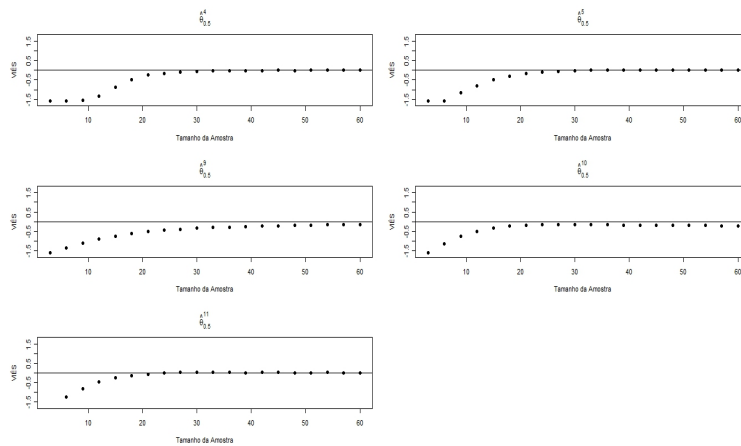


Figura 322 Viés para cada tamanho amostral e os estimadores $\hat{\theta}_{\Gamma}^4$, $\hat{\theta}_{\Gamma}^5$, $\hat{\theta}_{\Gamma}^9$, $\hat{\theta}_{\Gamma}^{10}$ e $\hat{\theta}_{\Gamma}^{11}$, amostras geradas considerando uma **Distribuição Gama** com parâmetros $\alpha = 2$ e $\lambda = 1,25$, para $\Gamma = 0,50$ e o **DILW**

APÊNDICE U

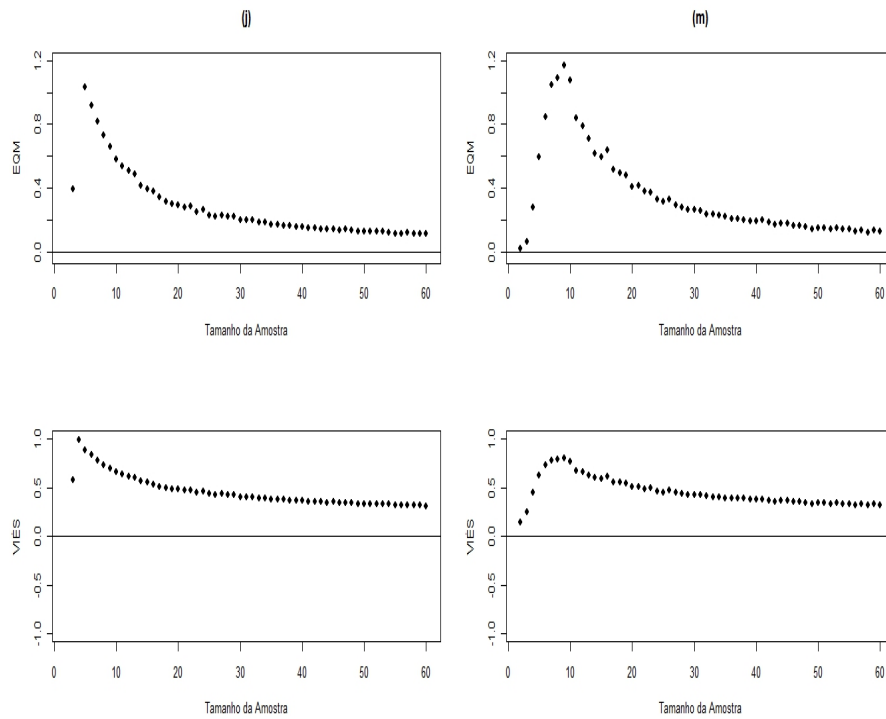


Figura 323 EQM e Viés para cada tamanho amostral e o estimador $\hat{\theta}_{\Gamma}^g$, amostras geradas considerando uma **Distribuição Exponencial** com parâmetro $\lambda = 0,30$, para $\Gamma = 0,10$. Cenários (j) e (m) apresentados na Tabela 10, referentes ao **DBIS** e **DNM** respectivamente

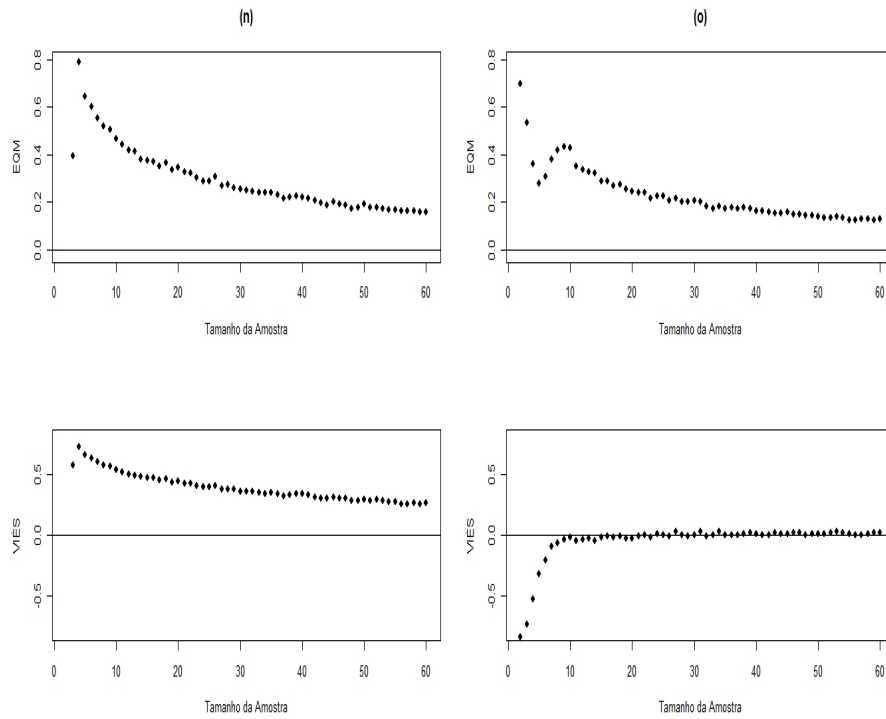


Figura 324 EQM e Viés para cada tamanho amostral e o estimador $\hat{\theta}_{\Gamma}^g$, amostras geradas considerando uma **Distribuição Exponencial** com parâmetro $\lambda = 0,30$, para $\Gamma = 0,33$. Cenários (n) e (o) apresentados na Tabela 10, referentes ao **DBIS** e **DNM** respectivamente

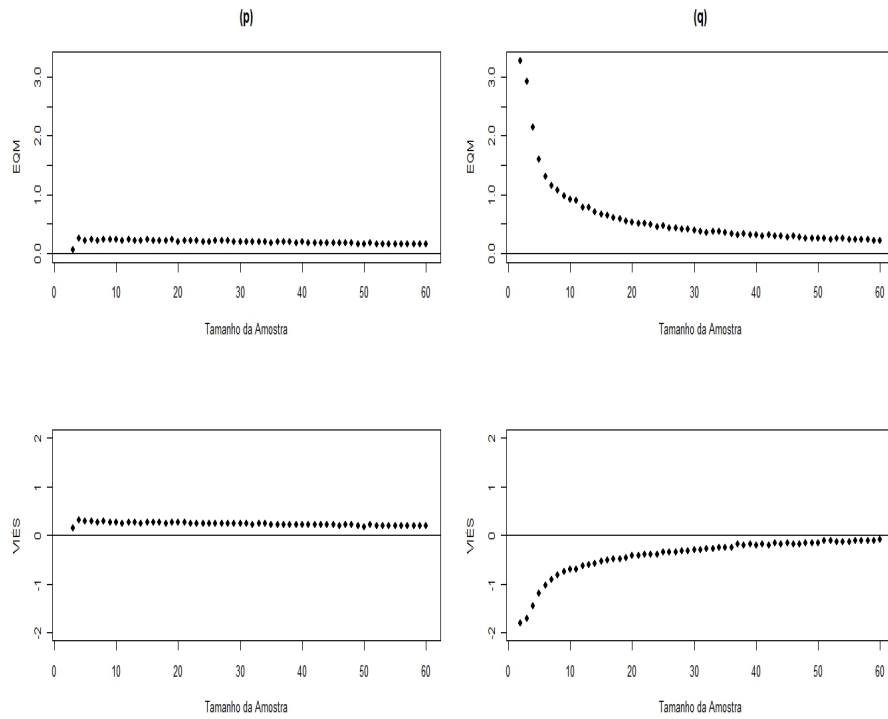


Figura 325 EQM e Viés para cada tamanho amostral e o estimador $\hat{\theta}_{\Gamma}^g$, amostras geradas considerando uma **Distribuição Exponencial** com parâmetro $\lambda = 0,30$, para $\Gamma = 0,50$. Cenários (p) e (q) apresentados na Tabela 10, referentes ao **DBIS** e **DNM** respectivamente

APÊNDICE V

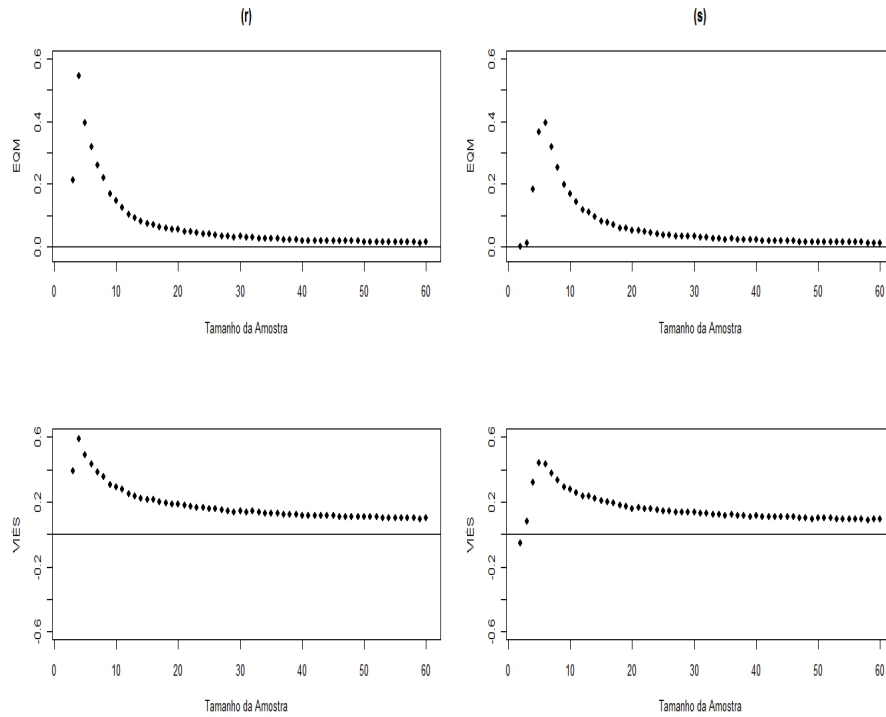


Figura 326 EQM e Viés para cada tamanho amostral e o estimador $\hat{\theta}_{\Gamma}^g$, amostras geradas considerando uma **Distribuição Gama** com parâmetros $\alpha = 3$ e $\lambda = 0,50$, para $\Gamma = 0,10$. Cenários (r) e (s) apresentados na Tabela 10, referentes ao **DBIS** e **DNM** respectivamente

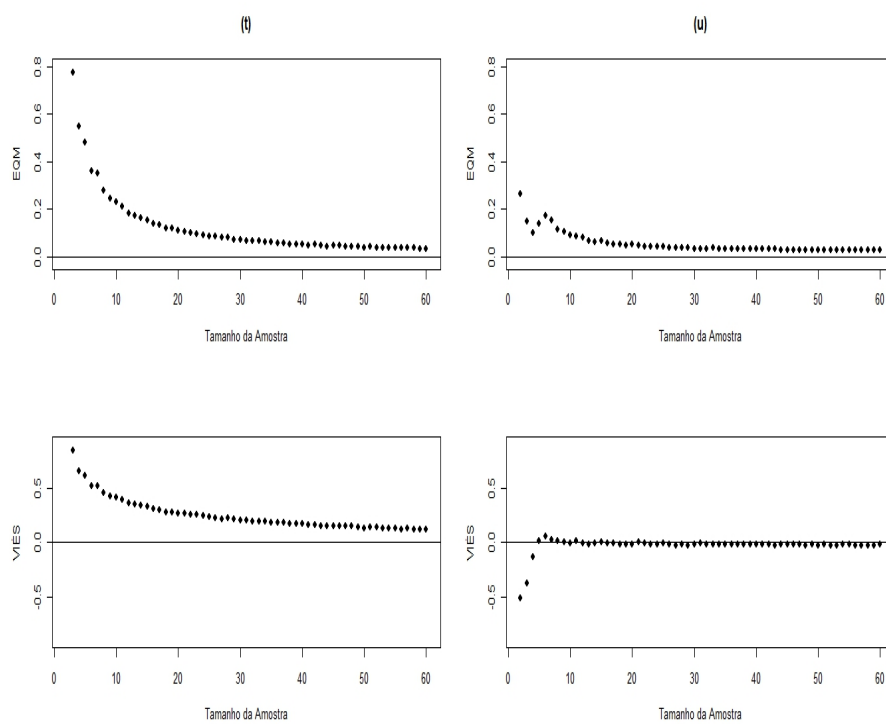


Figura 327 EQM e Viés para cada tamanho amostral e o estimador $\hat{\theta}_{\Gamma}^g$, amostras geradas considerando uma **Distribuição Gama** com parâmetros $\alpha = 3$ e $\lambda = 0,50$, para $\Gamma = 0,33$. Cenários (t) e (u) apresentados na Tabela 10, referentes ao **DBIS** e **DNM** respectivamente

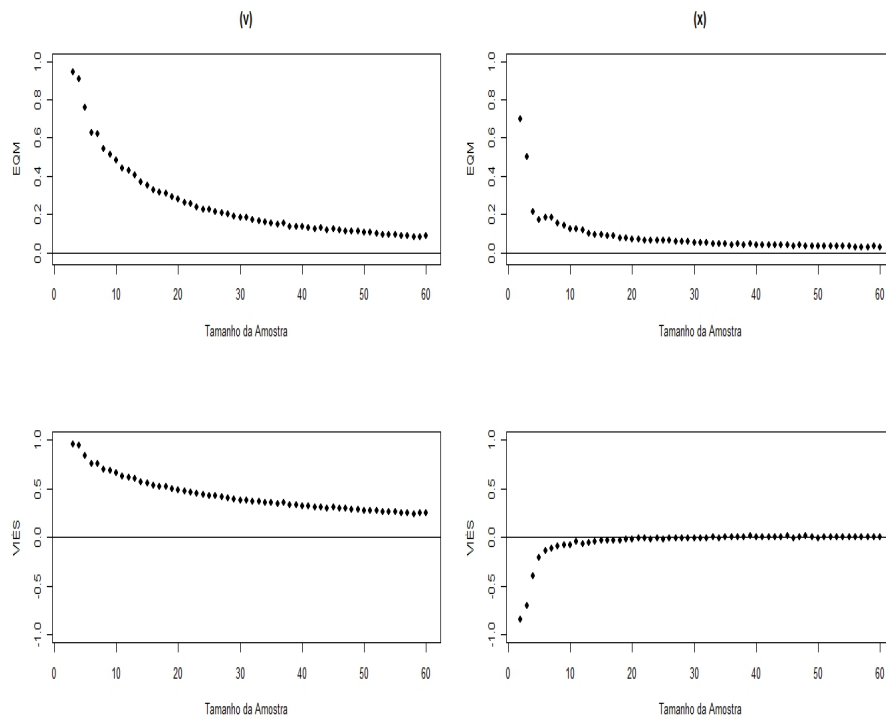


Figura 328 EQM e Viés para cada tamanho amostral e o estimador $\hat{\theta}_{\Gamma}^g$, amostras geradas considerando uma **Distribuição Gama** com parâmetros $\alpha = 3$ e $\lambda = 0,50$, para $\Gamma = 0,50$. Cenários (v) e (x) apresentados na Tabela 10, referentes ao **DBIS** e **DNM** respectivamente

APÊNDICE X

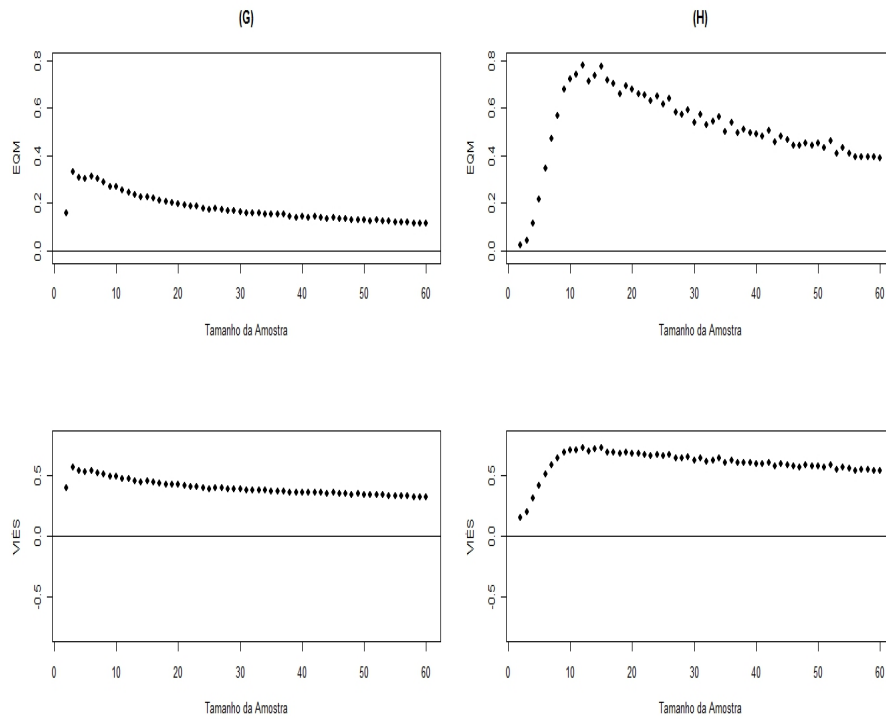


Figura 329 EQM e Viés para cada tamanho amostral e o estimador $\hat{\theta}_{\Gamma}^9$, amostras geradas considerando uma **Distribuição Exponencial** com parâmetro $\lambda = 0,30$, para $\Gamma = 0,10$. Cenários (G) e (H) apresentados na Tabela 11, referentes ao **DBIS** e **DNM** respectivamente

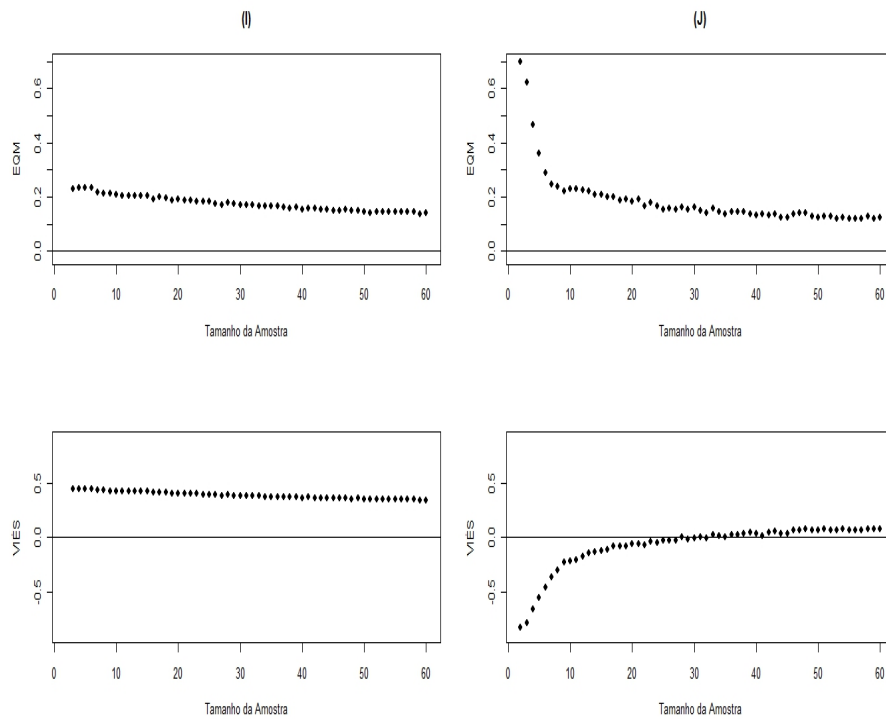


Figura 330 EQM e Viés para cada tamanho amostral e o estimador $\hat{\theta}_{\Gamma}^g$, amostras geradas considerando uma **Distribuição Exponencial** com parâmetro $\lambda = 0,30$, para $\Gamma = 0,33$. Cenários **(I)** e **(J)** apresentados na Tabela 11, referentes ao **DBIS** e **DNM** respectivamente

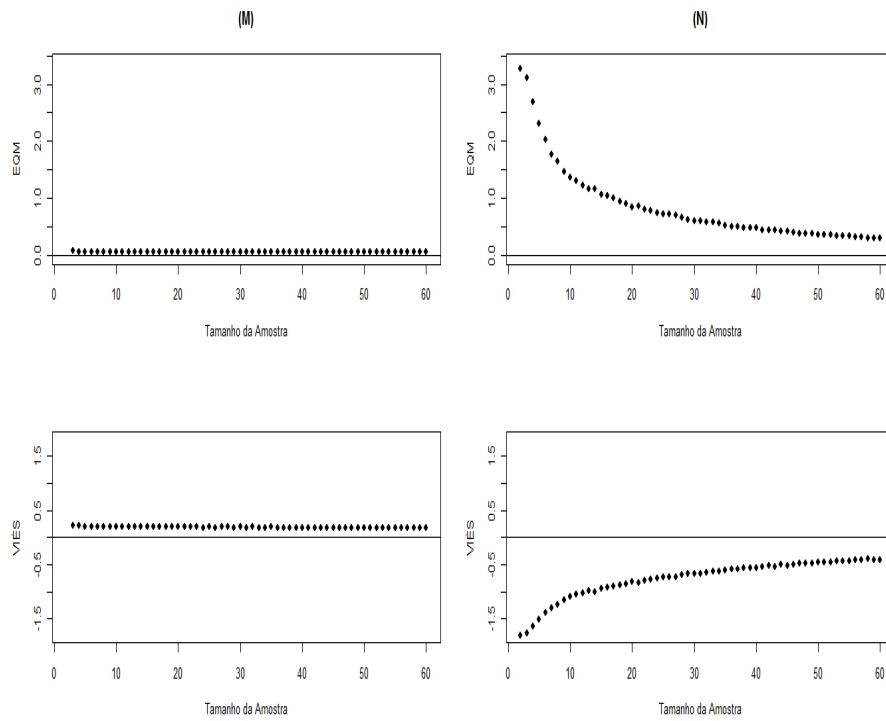


Figura 331 EQM e Viés para cada tamanho amostral e o estimador $\hat{\theta}_{\Gamma}^g$, amostras geradas considerando uma **Distribuição Exponencial** com parâmetro $\lambda = 0,30$, para $\Gamma = 0,50$. Cenários (M) e (N) apresentados na Tabela 11, referentes ao **DBIS** e **DNM** respectivamente

APÊNDICE Z

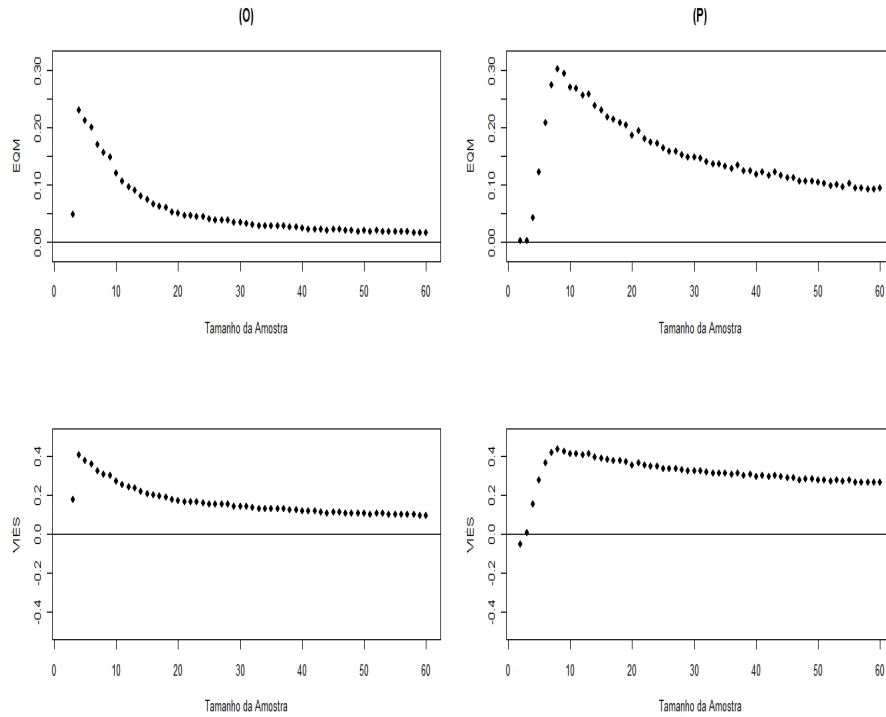


Figura 332 EQM e Viés para cada tamanho amostral e o estimador $\hat{\theta}_{\Gamma}^9$, amostras geradas considerando uma **Distribuição Gama** com parâmetros $\alpha = 3$ e $\lambda = 0,50$, para $\Gamma = 0,10$. Cenários (O) e (P) apresentados na Tabela 11, referentes ao **DBIS** e **DNM** respectivamente

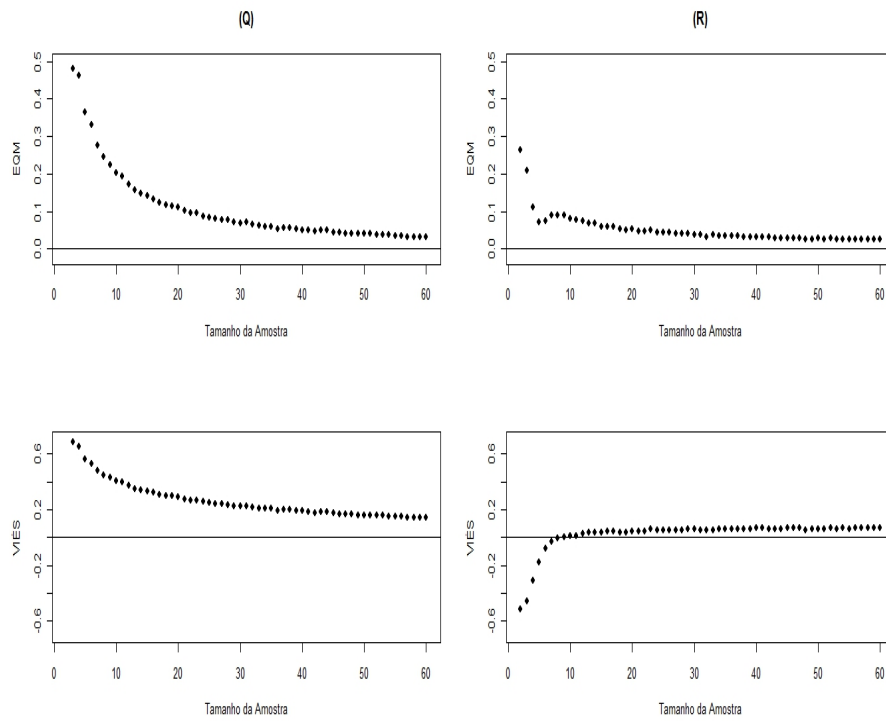


Figura 333 EQM e Viés para cada tamanho amostral e o estimador $\hat{\theta}_{\Gamma}^g$, amostras geradas considerando uma **Distribuição Gama** com parâmetros $\alpha = 3$ e $\lambda = 0,50$, para $\Gamma = 0,33$. Cenários (Q) e (R) apresentados na Tabela 11, referentes ao **DBIS** e **DNM** respectivamente

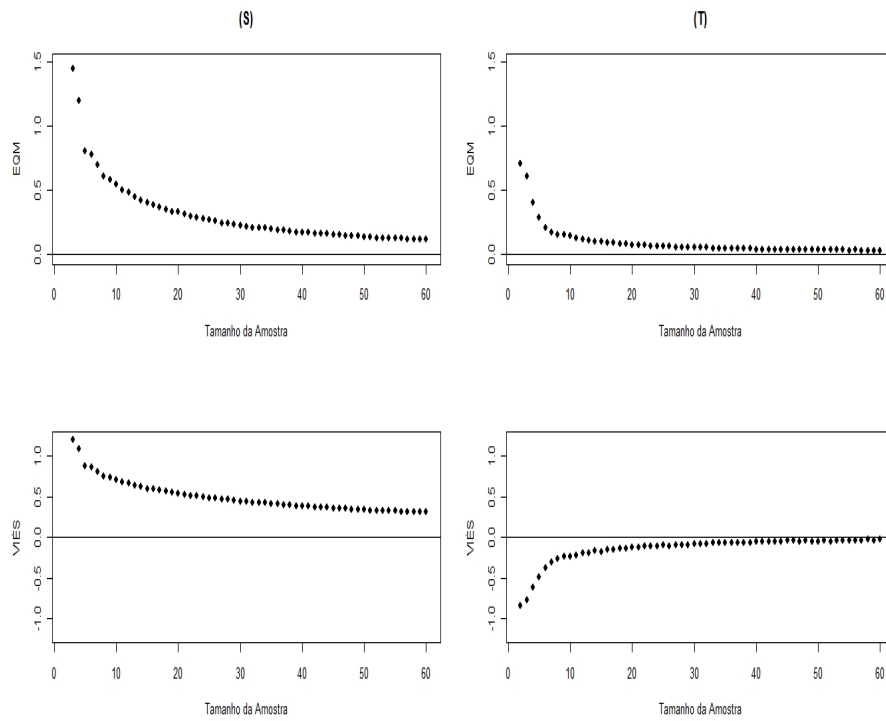


Figura 334 EQM e Viés para cada tamanho amostral e o estimador $\hat{\theta}_{\Gamma}^g$, amostras geradas considerando uma **Distribuição Gama** com parâmetros $\alpha = 3$ e $\lambda = 0,50$, para $\Gamma = 0,50$. Cenários (S) e (T) apresentados na Tabela 11, referentes ao **DBIS** e **DNM** respectivamente

APÊNDICE AA

```
# Simulando o Delineamento da Moeda Viciada (BCD);
# K: tamanho do passeio;
# p: vetor com as probabilidades;
# j: início do passeio;
# y: vetor de doses;
# Gama: Toxicidade alvo, determinada pelo pesquisador.
```

```
BCD=function(y,p,K,j,Gama){
  t=length(y)
  sucesso=c(rep(0,t))
  fracasso=c(rep(0,t))

  if(Gama<=0.5){
    b=Gama/(1-Gama)
    for(i in 1:K){
      w=p[j]
      u=runif(1)
      if(u<=w){
        sucesso[j]=sucesso[j]+1
        j=j-1
        if(j<1){
          j=1
        }
      }
    }
  } else{
    fracasso[j]=fracasso[j]+1
    m=runif(1)
    if(m<=b){
      j=j+1
      if(j>t){
        j=t
      }
    }
  } else{
    j=j
  }
}
```

```

        }
    }
}
else{
    b=(1-Gama)/Gama
    for(i in 1:K){
        w=p[j]
        u=runif(1)
        if(u>w){
            # aux=c(aux,y[j])
            fracasso[j]=fracasso[j]+1
            j=j+1
            if(j>t){
                j=t
            }
        }
    }
    else{
        # aux=c(aux,y[j])
        sucesso[j]=sucesso[j]+1
        m=runif(1)
        if(m<=b){
            j=j-1
            if(j<1){
                j=1
            }
        }
        else{
            j=j
        }
    }
}
}
tabela=matrix(c(y,sucesso,fracasso),length(y),3)
write(t(tabela),"cenarioBCD_27.txt",3,append=TRUE)
}

```

APÊNDICE AB

```

# Simulando o Delineamento de Narayna Modificado (DNM);
# p: vetor de probabilidades para cada dose;
# j: representa o inicio do passeio;
# y: vetor de doses;
# N: tamanho do passeio;
# k: número de doses anteriores;
# Gama: Toxicidade alvo, determinada pelo pesquisador

DNM=function(y,p,N,k,j,Gama){
  t=length(y)
  aux=c(rep(0,k))
  sucesso=c(rep(0,t))
  fracasso=c(rep(0,t))
  razao=c(rep(0,t))

  for(i in 1:N){
    w=p[j]
    u=runif(1)
    {if(u<=w){
      aux=c(aux,1)
      sucesso[j]=sucesso[j]+1
    }
    else{
      aux=c(aux,0)
      fracasso[j]=fracasso[j]+1
    }}

    razao[j]=sucesso[j]/(sucesso[j]+fracasso[j])
    cts= sum(aux[(length(aux)-k+1):length(aux)])

    {if((razao[j]>Gama & cts>=1)|(razao[j]<Gama
      & cts==0)){
      {if(razao[j]>Gama & cts>=1){
        j=j-1
      }
      if(j<1){
        j=1
      }
    }
  }
}

```

```
        }}
        {if(razao[j]<Gama & cts==0){
            j=j+1
        }
        if(j>t){
            j=t
        }}
    }
    else{
        j=j
    }}
}
tabela=matrix(c(y,sucesso,fracasso),length(y),3)
write(t(tabela),"cenarioDNM_27.txt",3,append=TRUE)
}
```

APÊNDICE AC

```

# Simulando o Delineamento Isotônico
# de Leung e Wang (DILW);
# j: representa o início do passeio;
# y: vetor de doses
# N: tamanho do passeio
# j: início do passeio
# Gama: Toxicidade alvo, determinada pelo pesquisador

```

```

DILW=function(y,p,N,M,j,Gama){
  h=length(y)
  n=c(rep(0,h))
  t=c(rep(0,h))
  phat=c(rep(0,h))
  library(Iso)

  for(i in 1:N){
    w=p[j]
    aux=rbinom(1,M,w)
    t[j]=aux+t[j]
    n[j]=M+n[j]
    phat[j]=t[j]/n[j]
    pest=pava(phat,n)

    {if(pest[j]<Gama){

      {if(j==h){
        c=pest[j]
        b=pest[j]
      }
      else{
        c=pest[j+1]
        b=pest[j]
      }
    }

    {if((Gama-b)>=(c-Gama)){
      j=j+1
      if(j>h){

```

```

        j=h
    }
}

else{
    j=j
}}

}
else{
    {if(j==1){
        a=pest[1]
        b=pest[1]
    }
    else{
        a=pest[j-1]
        b=pest[j]
    }}
    {if((Gama-a)<(b-Gama)){
        j=j-1
        if(j<1){
            j=1
        }
    }
    else{
        j=j
    }}
}}

}
sucesso=t
fracasso=n-t
tabela=matrix(c(y,sucesso,fracasso),length(y),3)
write(t(tabela),"cenarioDILW_27.txt",3,append=TRUE)
}

```

APÊNDICE AD

```

# Simulando o Delineamento Bayesiano Isotônico
# Sequencial (DBIS);
# j: representa o inicio do passeio;
# y: vetor de doses;
# N: tamanho do passeio;
# k: número de doses anteriores;
# M: Número de elementos amostrais por teste;
# h: vetor de precisão populacional;
# p0: vetor com moda a priori;
# h0: vetor de precisão a priori;
# prob: vetor com as verdadeiras probabilidades;
# Gama: Toxicidade alvo, determinada pelo pesquisador;
#####
# PS.: Nos ensaios de dose-resposta todos os      #
# elementos de h são iguais a 1, a rotina aqui   #
# é mais geral.                                  #
#####

DBIS=function(y,N,M,j,Gama,h,p0,h0,k,prob){
  f=length(y)
  n=c(rep(0,f))
  t=c(rep(0,f))
  phat=p0
  aux=c(rep(0,k))
  sucesso=c(rep(0,f))
  fracasso=c(rep(0,f))

  library(Iso)

  for(i in 1:N){
    p=prob[j]

    for(i in 1:M){
      u=runif(1)
      {if(u<=p){
        aux=c(aux,1)
        sucesso[j]=sucesso[j]+1
      }
    }
  }
}

```

```

        }
        else{
            aux=c(aux,0)
            fracasso[j]=fracasso[j]+1
        }
    }
    aux2=aux[(k+1):length(aux)]
    a=sum(aux2[(length(aux2)-M+1):length(aux2)])
    t[j]=a+t[j]
    n[j]=M+n[j]
    phat[j]=t[j]/n[j]
    w=h*n+h0
    ptil=(phat*h*n+p0*h0)/(h*n+h0)
    pest=pava(ptil,w)
    cts= sum(aux[(length(aux)-k+1):length(aux)])
    {if((pest[j]>Gama & cts>=1)|(pest[j]<Gama
        & cts==0)){
        {if(pest[j]>Gama & cts>=1){
            j=j-1
        }
        if(j<1){
            j=1
        }
        {if(pest[j]<Gama & cts==0){
            j=j+1
        }
        if(j>f){
            j=f
        }
    }
    }
    else{
        j=j
    }
}
phat=pest

}
tabela=matrix(c(y,sucesso,fracasso),length(y),3)
write(t(tabela),"cenarioDBI_12.txt",3,append=TRUE)
}

```

APÊNDICE AE

```

# Criando uma função para o Estimador Theta1;
# amostra: é uma matriz com três colunas
# a primeira coluna da matriz são as doses;
# a segunda coluna os sucessos observados;
# a terceira coluna os fracassos observados;
# Gama: Toxicidade alvo, determinada pelo pesquisador;

Theta1=function(amostra,Gama){
  { b=c(0)
  for(i in 1:(length(amostra)/3)){
    if(amostra[i,2]==0 & amostra[i,3]==0){
      b=c(b,i)
    }
  }
  visita=matrix(amostra[-b[-1],1:3],,3)
  if(length(visita)==0){
    visita=matrix(amostra,,3)
  }}
library(Iso)
  doses=visita[,1]
  sucesso=visita[,2]
  ki=visita[,2]+visita[,3]
  pin=sucesso/ki
  pi=pava(pin,ki)
  {if(sum(pi)>0){
    aux=matrix(c(doses,pi),,2)
    ct = 0
    for(i in 1:length(pi)){
      if(aux[i,2]!=0 & aux[i,2]!=1)
      {
        ct = ct + 1
        if (ct == 1) {
          mor_pro = aux[i,]
        } else mor_pro=rbind(mor_pro,aux[i,])
      }
    }
    {if(ct!=0){
      mor_pro = matrix(c(mor_pro),,2)
    }
  }
}

```

```

    {if(length(mor_pro)>2){
      ct0 = 0
      for(i in 1:length(pi)){
        if(aux[i,2]==0 )
          {
            ct0 = ct0 + 1
            if (ct0 == 1) {
              pi_zero = aux[i,]
            } else pi_zero=rbind(pi_zero,aux[i,])
          }
      }
      if(ct0!=0){
        pi_zero=matrix(c(pi_zero),,2)
      }
      ct1 = 0
      for(i in 1:length(pi)){
        if(aux[i,2]==1 )
          {
            ct1 = ct1 + 1
            if (ct1 == 1) {
              pi_um = aux[i,]
            } else pi_um=rbind(pi_um,aux[i,])
          }
      }
      if(ct1!=0){
        pi_um=matrix(c(pi_um),,2)
      }
      pi2=mor_pro[,2]
      doses2=mor_pro[,1]
      if (ct0!=0){
        dose_zero=pi_zero[,1]
      }
      if (ct1!=0){
        dose_um=pi_um[,1]
      }
      probit1=qnorm(pi2)
      z=dnorm(probit1)
      p=pnorm(probit1)
      q=1-p
    }
  }

```

```

wi=c(0)
for(i in 1:(length(visita)/3)){
  if(pi[i]!=0 & pi[i]!=1 ){
    wi=c(wi,ki[i])
  }
}
peso=(wi[-1]*z^2)/(p*q)
reg1=lm(probit1~doses2,weights=peso)
beta1=as.real(reg1$coefficients[1])
beta2=as.real(reg1$coefficients[2])
if (ct0!=0){
  probit_ajut_zero=beta1+beta2*dose_zero
  pnz=probit_ajut_zero-pnorm(probit_ajut_zero)
  /dnorm(probit_ajut_zero)
}
if (ct1!=0){
  probit_ajut_um=beta1+beta2*dose_um
  pnum=probit_ajut_um+(1-pnorm(probit_ajut_um))
  /dnorm(probit_ajut_um)
}
if (ct0!=0){
  aux3=matrix(c(dose_zero,pnz),,2)
}
if (ct1!=0){
  aux4=matrix(c(dose_um,pnum),,2)
}
aux5=matrix(c(mor_pro[,1],probit1),,2)
  if (ct0!=0 & ct1!=0){
    ndado=as.matrix(rbind(aux3,aux4,aux5))
  }
  if (ct0!=0 & ct1==0){
    ndado=as.matrix(rbind(aux3,aux5))
  }
  if (ct0==0 & ct1!=0){
    ndado=as.matrix(rbind(aux4,aux5))
  }
  if (ct0==0 & ct1==0){
    ndado=aux5
  }
}

```


APÊNDICE AF

```

# Criando uma função para o Estimador Theta2;
# amostra: é uma matriz com três colunas
# a primeira coluna da matriz são as doses;
# a segunda coluna os sucessos observados;
# a terceira coluna os fracassos observados;
# Gama: Toxicidade alvo, determinada pelo pesquisador;

Theta2=function(amostra,Gama){
  { b=c(0)
  for(i in 1:(length(amostra)/3)){
    if(amostra[i,2]==0 & amostra[i,3]==0){
      b=c(b,i)
    }
  }
  visita=matrix(amostra[-b[-1],1:3],,3)
  if(length(visita)==0){
    visita=matrix(amostra,,3)
  }
library(Iso)
  doses=visita[,1]
  sucesso=visita[,2]
  ki=visita[,2]+visita[,3]
  pin=sucesso/ki
  pi=pava(pin,ki)
  {if(sum(pi)>0){
    aux=matrix(c(doses,pi,ki),,3)
    ct = 0
    for(i in 1:length(pi)){
      if(aux[i,2]!=0 & aux[i,2]!=1)
      {
        ct = ct + 1
        if (ct == 1) {
          mor_pro = aux[i,]
        } else mor_pro=rbind(mor_pro,aux[i,])
      }
    }
  }
  {if(ct!=0){
    mor_pro = matrix(c(mor_pro),,3)
  }
}

```

```

    {if(length(mor_pro)>3){
      ct0 = 0
      for(i in 1:length(pi)){
        if(aux[i,2]==0 )
          {
            ct0 = ct0 + 1
            if (ct0 == 1) {
              pi_zero = aux[i,]
            } else pi_zero=rbind(pi_zero,aux[i,])
          }
      }
      if(ct0!=0){
        pi_zero=matrix(c(pi_zero),,3)
      }
      ct1 = 0
      for(i in 1:length(pi)){
        if(aux[i,2]==1 )
          {
            ct1 = ct1 + 1
            if (ct1 == 1) {
              pi_um = aux[i,]
            } else pi_um=rbind(pi_um,aux[i,])
          }
      }
      if(ct1!=0){
        pi_um=matrix(c(pi_um),,3)
      }
      pi2=mor_pro[,2]
      doses2=mor_pro[,1]
      if (ct0!=0){
        dose_zero=pi_zero[,1]
      }
      if (ct1!=0){
        dose_um=pi_um[,1]
      }
      logit1=log(pi2/(1-pi2))
      p=mor_pro[,2]
      q=1-p
      mi=mor_pro[,3]

```

```

peso=mi*p*q
  reg1=lm(logit1~doses2,weights=peso)
  beta1=as.real(reg1$coefficients[1])
  beta2=as.real(reg1$coefficients[2])
    {if(beta2>0){
if (ct0!=0){
  logit_ajut_zero=beta1+beta2*dose_zero
  pj0=exp(logit_ajut_zero)/(1+exp(logit_ajut_zero))
  qj0=1-pj0
  pnz=logit_ajut_zero- 1/qj0
}
if (ct1!=0){
  logit_ajut_um=beta1+beta2*dose_um
  pj1=exp(logit_ajut_um)/(1+exp(logit_ajut_um))
  pnum=logit_ajut_um+1/pj1
}
if (ct0!=0){
  aux3=matrix(c(dose_zero,pnz,pj0),,3)
}
if (ct1!=0){
  aux4=matrix(c(dose_um,pnum,pj1),,3)
}
aux5=matrix(c(mor_pro[,1],logit1,mor_pro[,2]),,3)
  if (ct0!=0 & ct1!=0){
    ndado=as.matrix(rbind(aux3,aux4,aux5))
  }
  if (ct0!=0 & ct1==0){
    ndado=as.matrix(rbind(aux3,aux5))
  }
  if (ct0==0 & ct1!=0){
    ndado=as.matrix(rbind(aux4,aux5))
  }
  if (ct0==0 & ct1==0){
    ndado=aux5
  }
dosef=ndado[,1]
logitf=ndado[,2]
mif=visita[,2]+visita[,3]
pif=ndado[,3]

```



```
}}
}
else{
  theta3="SS"
}}
}
else{
  theta3="SS"
}}
list(theta3=theta3)
}
```

APÊNDICE AH

```

# Criando uma função para o Estimador Theta4;
# amostra: é uma matriz com três colunas
# a primeira coluna da matriz são as doses;
# a segunda coluna os sucessos observados;
# a terceira coluna os fracassos observados;

Theta4=function(amostra) {
  { b=c(0)
  for(i in 1:(length(amostra)/3)){
    if(amostra[i,2]==0 & amostra[i,3]==0){
      b=c(b,i)
    }
  }
  visita=matrix(amostra[-b[-1],1:3],,3)
  if(length(visita)==0){
    visita=matrix(amostra,,3)
  }
  library(Iso)
  {if(length(visita)==3){
    theta4=visita[1,1]
  }
  else{
    doses=visita[,1]
    fi=visita[,2]+visita[,3]
    res=ufit(y=fi,x=doses)
    theta4=res$mode
  }
  list(theta4=theta4)
}

```

APÊNDICE AI

```

# Criando uma função para o Estimador Theta5;
# amostra: é uma matriz com três colunas
# a primeira coluna da matriz são as doses;
# a segunda coluna os sucessos observados;
# a terceira coluna os fracassos observados;

Theta5=function(amostra){
  { b=c(0)
  for(i in 1:(length(amostra)/3)){
    if(amostra[i,2]==0 & amostra[i,3]==0){
      b=c(b,i)
    }
  }
  visita=matrix(amostra[-b[-1],1:3],,3)
  if(length(visita)==0){
    visita=matrix(amostra,,3)
  }
  {if(length(visita)==3){
    theta5=visita[1,1]
  }
  else{
  library(Iso)
  doses=visita[,1]
  fi=visita[,2]+visita[,3]
  res=ufit(y=fi,x=doses)
  fii=res$y
  mat=matrix(c(doses,fii),,2)
  dmax=res$mode
  for(i in 1:length(mat[,1])){
    if(mat[i,1]==dmax){
      j=i
    }
  }
  {if(dmax==visita[1,1]){
    dant=dmax
  }
  else{
    dant=mat[j-1,1]
  }
  }
}

```

```
    }}
    {if (dmax==visita[length(visita[,1]),1]){
        dpos=dmax
    }
    else{
        dpos=mat[j+1,1]
    }}
    for(i in 1:length(mat[,1])){
        if(mat[i,1]==dmax){
            fmax=mat[i,2]
        }
        if(mat[i,1]==dant){
            fant=mat[i,2]
        }
        if(mat[i,1]==dpos){
            fpos=mat[i,2]
        }
    }
    {if(fmax==fant & fmax==fpos & fant==fpos){
        theta5="SS"
    }
    else{
        theta5=dant+((fmax-fant)/(2*fmax-fpos-fant))
        *(dpos-dant)
    }}
    }}
    list(theta5=theta5)
}
```

APÊNDICE AJ

```

# Criando uma função para o Estimador Theta6;
# amostra: é uma matriz com três colunas
# a primeira coluna da matriz são as doses;
# a segunda coluna os sucessos observados;
# a terceira coluna os fracassos observados;
# Gama: Toxicidade alvo, determinada pelo pesquisador;

Theta6=function(amostra,Gama){
  { b=c(0)
  for(i in 1:(length(amostra)/3)){
    if(amostra[i,2]==0 & amostra[i,3]==0){
      b=c(b,i)
    }
  }
  visita=matrix(amostra[-b[-1],1:3],,3)
  if(length(visita)==0){
    visita=matrix(amostra,,3)
  }}
  doses=visita[,1]
  sucesso=visita[,2]
  ki=visita[,2]+visita[,3]
  pi=sucesso/ki
  {if(sum(pi)>0){
    aux=matrix(c(doses,pi),,2)
    ct = 0
    for(i in 1:length(pi)){
      if(aux[i,2]!=0 & aux[i,2]!=1)
      {
        ct = ct + 1
        if (ct == 1) {
          mor_pro = aux[i,]
        } else mor_pro=rbind(mor_pro,aux[i,])
      }
    }
  }
  {if(ct!=0){
    mor_pro = matrix(c(mor_pro),,2)
    {if(length(mor_pro)>2){
      ct0 = 0

```

```

for(i in 1:length(pi)){
  if(aux[i,2]==0 )
  {
    ct0 = ct0 + 1
    if (ct0 == 1) {
      pi_zero = aux[i,]
    } else pi_zero=rbind(pi_zero,aux[i,])
  }
}
if(ct0!=0){
  pi_zero=matrix(c(pi_zero),,2)
}
ct1 = 0
for(i in 1:length(pi)){
  if(aux[i,2]==1 )
  {
    ct1 = ct1 + 1
    if (ct1 == 1) {
      pi_um = aux[i,]
    } else pi_um=rbind(pi_um,aux[i,])
  }
}
if(ct1!=0){
  pi_um=matrix(c(pi_um),,2)
}
pi2=mor_pro[,2]
doses2=mor_pro[,1]
if (ct0!=0){
  dose_zero=pi_zero[,1]
}
if (ct1!=0){
  dose_um=pi_um[,1]
}
probit1=qnorm(pi2)
z=dnorm(probit1)
p=pnorm(probit1)
q=1-p
wi=c(0)
for(i in 1:(length(visita)/3)){

```

```

    if(pi[i]!=0 & pi[i]!=1 ){
      wi=c(wi,ki[i])
    }
  }
  peso=(wi[-1]*z^2)/(p*q)
  reg1=lm(probit1~doses2,weights=peso)
  beta1=as.real(reg1$coefficients[1])
  beta2=as.real(reg1$coefficients[2])
  {if(beta2>0){
if (ct0!=0){
  probit_ajut_zero=beta1+beta2*dose_zero
  pnz=probit_ajut_zero-pnorm(probit_ajut_zero)
  /dnorm(probit_ajut_zero)
}
if (ct1!=0){
  probit_ajut_um=beta1+beta2*dose_um
  pnum=probit_ajut_um+(1-pnorm(probit_ajut_um))
  /dnorm(probit_ajut_um)
}
if (ct0!=0){
  aux3=matrix(c(dose_zero,pnz),,2)
}
if (ct1!=0){
  aux4=matrix(c(dose_um,pnum),,2)
}
aux5=matrix(c(mor_pro[,1],probit1),,2)
  if (ct0!=0 & ct1!=0){
    ndado=as.matrix(rbind(aux3,aux4,aux5))
  }
  if (ct0!=0 & ct1==0){
    ndado=as.matrix(rbind(aux3,aux5))
  }
  if (ct0==0 & ct1!=0){
    ndado=as.matrix(rbind(aux4,aux5))
  }
  if (ct0==0 & ct1==0){
    ndado=aux5
  }
dosef=ndado[,1]

```


APÊNDICE AL

```

# Criando uma função para o Estimador Theta7;
# amostra: é uma matriz com três colunas
# a primeira coluna da matriz são as doses;
# a segunda coluna os sucessos observados;
# a terceira coluna os fracassos observados;
# Gama: Toxicidade alvo, determinada pelo pesquisador;

Theta7=function(amostra,Gama){
  { b=c(0)
  for(i in 1:(length(amostra)/3)){
    if(amostra[i,2]==0 & amostra[i,3]==0){
      b=c(b,i)
    }
  }
  visita=matrix(amostra[-b[-1],1:3],,3)
  if(length(visita)==0){
    visita=matrix(amostra,,3)
  }}
  doses=visita[,1]
  sucesso=visita[,2]
  ki=visita[,2]+visita[,3]
  pi=sucesso/ki
  {if(sum(pi)>0){
    aux=matrix(c(doses,pi,ki),,3)
    ct = 0
    for(i in 1:length(pi)){
      if(aux[i,2]!=0 & aux[i,2]!=1)
      {
        ct = ct + 1
        if (ct == 1) {
          mor_pro = aux[i,]
        } else mor_pro=rbind(mor_pro,aux[i,])
      }
    }
  }
  {if(ct!=0){
    mor_pro = matrix(c(mor_pro),,3)
    {if(length(mor_pro)>3){
      ct0 = 0

```

```

for(i in 1:length(pi)){
  if(aux[i,2]==0 )
  {
    ct0 = ct0 + 1
    if (ct0 == 1) {
      pi_zero = aux[i,]
    } else pi_zero=rbind(pi_zero,aux[i,])
  }
}
if(ct0!=0){
  pi_zero=matrix(c(pi_zero),,3)
}
ct1 = 0
for(i in 1:length(pi)){
  if(aux[i,2]==1 )
  {
    ct1 = ct1 + 1
    if (ct1 == 1) {
      pi_um = aux[i,]
    } else pi_um=rbind(pi_um,aux[i,])
  }
}
if(ct1!=0){
  pi_um=matrix(c(pi_um),,3)
}
pi2=mor_pro[,2]
doses2=mor_pro[,1]
if (ct0!=0){
  dose_zero=pi_zero[,1]
}
if (ct1!=0){
  dose_um=pi_um[,1]
}
logit1=log(pi2/(1-pi2))
p=mor_pro[,2]
q=1-p
mi=mor_pro[,3]
peso=mi*p*q
reg1=lm(logit1~doses2,weights=peso)

```

```

        beta1=as.real(reg1$coefficients[1])
        beta2=as.real(reg1$coefficients[2])
        {if(beta2>0){
if (ct0!=0){
    logit_ajut_zero=beta1+beta2*dose_zero
    pj0=exp(logit_ajut_zero)/(1+exp(logit_ajut_zero))
    qj0=1-pj0
    pnz=logit_ajut_zero- 1/qj0 # probit ajustado
}
if (ct1!=0){
    logit_ajut_um=beta1+beta2*dose_um
    pj1=exp(logit_ajut_um)/(1+exp(logit_ajut_um))
    pnum=logit_ajut_um+1/pj1 # logit ajustado
}
if (ct0!=0){
    aux3=matrix(c(dose_zero,pnz,pj0),,3)
}
if (ct1!=0){
    aux4=matrix(c(dose_um,pnum,pj1),,3)
}
aux5=matrix(c(mor_pro[,1],logit1,mor_pro[,2]),,3)
    if (ct0!=0 & ct1!=0){
        ndado=as.matrix(rbind(aux3,aux4,aux5))
    }
    if (ct0!=0 & ct1==0){
        ndado=as.matrix(rbind(aux3,aux5))
    }
    if (ct0==0 & ct1!=0){
        ndado=as.matrix(rbind(aux4,aux5))
    }
    if (ct0==0 & ct1==0){
        ndado=aux5
    }
dosef=ndado[,1]
logitf=ndado[,2]
mif=visita[,2]+visita[,3]
pif=ndado[,3]
qif=1-pif
pesof=mif*pif*qif

```

```
    regf=lm(logitf~dosef,weights=pesof)
    beta1f=as.real(regf$coefficients[1])
    beta2f=as.real(regf$coefficients[2])
    {if(beta2f>0.009){
      theta7=(log(Gama/(1-Gama))-beta1f)/beta2f
    }
    else{
      theta7="SS"
    }}
  }
  else{
    theta7="SS"
  }}
}
else {
  theta7="SS"
}}
}
else{
  theta7="SS"
}}
list(theta7=theta7)
}
```

APÊNDICE AM

```

# Criando uma função para o Estimador Theta8;
# amostra: é uma matriz com três colunas
# a primeira coluna da matriz são as doses;
# a segunda coluna os sucessos observados;
# a terceira coluna os fracassos observados;
# Gama: Toxicidade alvo, determinada pelo pesquisador;

Theta8=function(amostra,Gama){
  { b=c(0)
  for(i in 1:(length(amostra)/3)){
    if(amostra[i,2]==0 & amostra[i,3]==0){
      b=c(b,i)
    }
  }
  visita=matrix(amostra[-b[-1],1:3],,3)
  if(length(visita)==0){
    visita=matrix(amostra,,3)
  }
  {if(length(visita)>3){
    doses=visita[,1]
    sucesso=visita[,2]
    ki=visita[,2]+visita[,3]
    pi=visita[,2]/ki
  }
  {if(sum(pi)>0){
    resposta=cbind(sucesso,ki-sucesso)
    reg.glm=glm(resposta~doses,
    family=binomial(link="logit"))
    cof=as.vector(reg.glm$coefficients)
    beta1=cof[1]
    beta2=cof[2]
    {if(beta2>0.009){
      theta8=(log(Gama/(1-Gama))-beta1)/beta2
    }
  }
  else{
    theta8="SS"
  }
  }
  else{

```

```
theta8="SS"  
  }  
  }  
  }  
  else{  
    theta8="SS"  
  }  
  }  
list(theta8=theta8)  
}
```

APÊNDICE AN

```

# Criando uma função para o Estimador Theta9;
# amostra: é uma matriz com três colunas
# a primeira coluna da matriz são as doses;
# a segunda coluna os sucessos observados;
# a terceira coluna os fracassos observados;

Theta9=function(amostra){
  { b=c(0)
  for(i in 1:(length(amostra)/3)){
    if(amostra[i,2]==0 & amostra[i,3]==0){
      b=c(b,i)
    }
  }
  visitao=matrix(amostra[-b[-1],1:3],,3)
  if(length(visitao)==0){
    visitao=matrix(amostra,,3)
  }
  ct1=visitao[,2]+visitao[,3]
  ct2=0
  ct3=0
  for(i in 1:length(ct1)){
    ct2=ct1[i]+ct2
    if(ct2==i){
      ct3=ct3+1
    }
  }
  {if(ct3!=(length(visitao)/3)){
    visita=matrix(visitao[(ct3+1):length(ct1),1:3],,3)
    theta9=sum(visita[,1]*(visita[,2]+visita[,3]))
    /sum(visita[,2]+visita[,3])
  }
  else{
    theta9="SS"
  }
}
list(theta9=theta9)
}

```

APÊNDICE AO

```

# Criando uma função para o Estimador Theta10;
# amostra: é uma matriz com três colunas
# a primeira coluna da matriz são as doses;
# a segunda coluna os sucessos observados;
# a terceira coluna os fracassos observados;
# Gama: Toxicidade alvo, determinada pelo pesquisador;

Theta10=function(amostra,Gama){
  { b=c(0)
    for(i in 1:(length(amostra)/3)){
      if(amostra[i,2]==0 & amostra[i,3]==0){
        b=c(b,i)
      }
    }
    visita=matrix(amostra[-b[-1],1:3],,3)
    if(length(visita)==0){
      visita=matrix(amostra,,3)
    }
  }
  library(Iso)
  doses=visita[,1]
  ki=visita[,2]+visita[,3]
  pin=visita[,2]/ki
  pi=pava(y=pin,w=ki)
  mat=matrix(c(doses,pi),,2)
  ct=0
  for(i in 1:length(doses)){
    if(mat[i,2]<=Gama){
      aux=mat[i,1:2]
      ct=ct+1
    }
  }
  {if(ct!=0){
    theta10=aux[1]
  } else{ theta10="SS"
  }}
  list(theta10=theta10)
}

```

APÊNDICE AP

```

# Criando uma função para o Estimador Theta11;
# amostra: é uma matriz com três colunas
# a primeira coluna da matriz são as doses;
# a segunda coluna os sucessos observados;
# a terceira coluna os fracassos observados;
# Gama: Toxicidade alvo, determinada pelo pesquisador;

Theta11=function(amostra,Gama){
  { b=c(0)
  for(i in 1:(length(amostra)/3)){
    if(amostra[i,2]==0 & amostra[i,3]==0){
      b=c(b,i)
    }
  }
  visita=matrix(amostra[-b[-1],1:3],,3)
  if(length(visita)==0){
    visita=matrix(amostra,,3)
  }
library(Iso)
  doses=visita[,1]
  ki=visita[,2]+visita[,3]
  pin=visita[,2]/ki
  pi=pava(y=pin,w=ki)
  mat=matrix(c(doses,pi),,2)
  {if(length(mat)>2){
    ct1=0
    ct2=0
    for(i in 1:length(doses)){
      if(mat[i,2]<=Gama){
        aux1=mat[i,1:2]
        ct1=ct1+1
        {if(aux1[1]<max(visita[,1])){
          aux2=mat[i+1,1:2]
        }
        }
      }
      else{
        ct2=ct2+1
      }
    }
  }
}

```

```
    }
    {if(ct1!=0){
theta10=aux1[1]
}
else{
  theta10="SS"
}}
{if(theta10=="SS"){
  theta11="SS"
}
else{
  {if(ct2==0){
    theta11=((Gama-aux1[2])/(aux2[2]-aux1[2]))
    *(aux2[1]-theta10)+theta10
  }
  else{
    theta11="SS"
  }}
}}
}
else{
  theta11="SS"
}}
list(theta11=theta11)
}
```