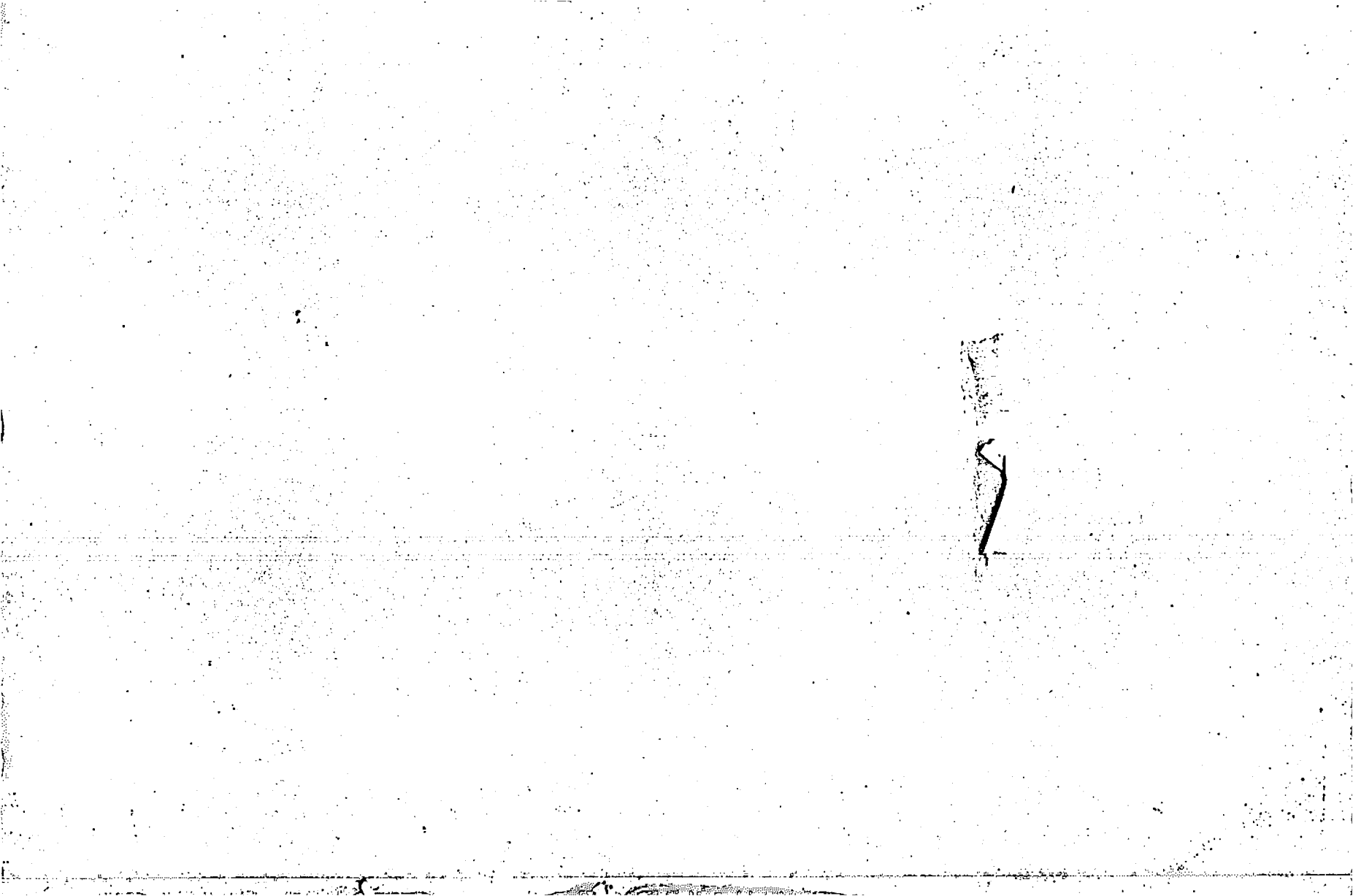




**CONTROLE ESTATÍSTICO DA QUALIDADE  
DE SEMENTES DE SOJA ARMAZENADAS.  
UM ESTUDO VIA PREDIÇÃO DE  
QUANTIDADES ALEÁTORIAS**

**RAFAELA CARDOSO ALVES**

**2004**



**RAFAELA CARDOSO ALVES**

**CONTROLE ESTATÍSTICO DA QUALIDADE DE SEMENTES  
DE SOJA ARMAZENADAS. UM ESTUDO VIA PREDIÇÃO DE  
QUANTIDADES ALEATÓRIAS**

Dissertação apresentada à Universidade Federal de Lavras como parte das exigências do Curso de Mestrado em Agronomia, área de concentração em Estatística e Experimentação Agropecuária, para a obtenção do título de "Mestre".

Orientador

Prof. Mário Javier Ferrua Vivanco

Co-orientador

Maria Laene Moreira de Carvalho

MINAS GERAIS - BRASIL

2004

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

**Alves, Rafaela Cardoso**

**Controle estatístico da qualidade de sementes de soja armazenadas: um estudo via predição de quantidades aleatórias / Rafaela Cardoso Alves. -- Lavras : UFLA, 2004.**

**60 p. : il.**

**Orientador: Mario Javier Ferrua Vivanco.**

**Dissertação (Mestrado) – UFLA.**

**Bibliografia.**

**1. Predição de qualidade aleatória. 2. Distribuição Weibull. 3. Viabilidade de semente de soja. I. Universidade Federal de Lavras. II. Título.**

**CDD-631.521**

**-633.3421**

**RAFAELA CARDOSO ALVES**

**CONTROLE ESTATÍSTICO DA QUALIDADE DE SEMENTES DE  
SOJA ARMAZENADAS. UM ESTUDO VIA PREDIÇÃO DE  
QUANTIDADES ALEATÓRIAS**

Dissertação apresentada à Universidade Federal de Lavras como parte das exigências do Curso de Pós-graduação em Estatística e Experimentação Agropecuária, para obtenção do título de "Mestre".

APROVADA em 20 de fevereiro de 2004

Marcelo Silva de Oliveira

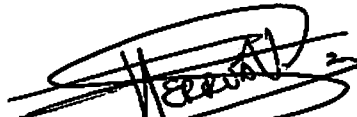
UFLA

Maria Laene Moreira Carvalho

UFLA

Rogério de M. Costa Pinto

UFU



Prof. Mário Javier Ferrua Vivanco

DEX/UFLA

(Orientador)

LAVRAS  
MINAS GERAIS - BRASIL

## **Dedico**

**Ao meu pai, Edson Queiroz Alves e minha mãe, Rosângela Cardoso Alves pelo grande amor, apoio e incentivo. Me educando de forma que, eu pudesse alcançar meus objetivos. E por todos os sacrifícios feitos para que eu pudesse chegar até aqui.**

**As minhas irmãs, Fabíola, Roberta e Eduarda, que sempre me ajudaram e são acima de tudo minhas melhores amigas.**

## **AGRADECIMENTOS**

A Deus, por realizar este sonho e pela força dada neste período tão importante de minha vida.

Ao Prof. Doutor Mário Javier Ferrua Vivanco, pela orientação, compreensão e paciência, importantes na realização deste trabalho.

Aos meus companheiros de república, Vânia, José Waldemar, Elisângela, Janser e Gisele pelo companheirismo e amizade, fundamentais durante o curso.

Aos meus colegas de curso, Carla, Denismar, Eliana, Fabiano, Francisco, Janser, José Waldemar, Leite, Neia, Renata e Vânia pelos bons momentos que passamos juntos.

Ao Prof. Fortunato Silva de Menezes pela colaboração no desenvolvimento computacional deste trabalho.

A Prof. Maria Laene Moreira de Carvalho pela colaboração na realização do experimento com sementes de soja.

Ao Prof. Daniel Furtado Ferreira pela ajuda e boa vontade demonstrada.

Aos professores do DEX, Augusto, Daniel, Delly, Eduardo, Joel, Júlio, Lucas, Luiz Henrique, Marcelo, Paulo e Thelma pelos ensinamentos.

Às funcionárias do DEX, pela disposição em nos atender.

À Universidade Federal de Lavras e ao Programa de Pós-Graduação em Agronomia / Estatística e Experimentação agropecuária., pela oportunidade de realizar este curso.

A CAPES, pela concessão da bolsa de estudos.

Enfim, a todos que contribuíram para a realização deste trabalho.

# SUMÁRIO

|  | Página |
|--|--------|
| GLOSSÁRIO.....   | i      |
| RESUMO.....  | ii     |
| ABSTRACT.....  | iii    |
| 1 INTRODUÇÃO.....  | 1      |
| 2 REFERÊNCIAL TEÓRICO.....   | 4      |
| 3 MATERIAL E MÉTODOS.....  | 16     |
| 3.1 Obtenção do tempo até a perda da viabilidade da semente.....   | 18     |
| 3.2 Determinação da distribuição do tempo até a perda da viabilidade da semente.....                           | 21     |
| 3.3 Predição do número futuro de sub-amostras inviáveis.....   | 21     |
| 3.4 Determinação da probabilidade de cobertura para intervalos de predição bilateral.....                      | 24     |
| 4 RESULTADOS e DISCUSSÃO.....  | 26     |
| 4.1 Adequação do modelo (dados obtidos Via Simulação).....   | 26     |
| 4.2 Predição Pontual e Intervalar para o número futuro de sementes não viáveis utilizando dados simulados..... | 42     |
| 4.3 Adequação do modelo (dados obtidos Via Experimento).....   | 49     |
| 4.4 Predição Pontual e Intervalar para o número futuro de sementes não viáveis utilizando dados reais.....     | 51     |
| 4.5 Determinação da probabilidade de cobertura para intervalos de predição bilateral.....                      | 52     |
| 5.CONCLUSÕES.....  | 54     |
| REFERÊNCIAS BIBLIOGRÁFICAS.....  | 55     |
| ANEXO.....   | 57     |



## Glossário

|   |                |
|---|----------------|
| Estimativa de máxima verossimilhança de $q$ .....             | $\hat{q}$      |
| Número de amostras.....                                       | $n$            |
| Número de sub-amostras falhas no período $[0,4]$ semanas..... | $X$            |
| Parâmetro escala da distribuição Weibull.....                 | $\hat{\eta}$   |
| Parâmetro forma da distribuição Weibull.....                  | $\hat{\beta}$  |
| Preditor pontual.....   | $\hat{Y}$      |
| Preditor intervalar.....                                      | $[Y, \bar{Y}]$ |
| Probabilidade de haver falha no período $[0, t_c]$ .....      | $p$            |
| Probabilidade de haver falha no período $[t_c, t_w]$ .....    | $q$            |
| Probabilidade de haver falha no período $[t_w, \infty]$ ..... | $r$            |
| Probabilidade de inviabilidade inicial.....                   | $p_{di}$       |
| Probabilidade de inviabilidade máxima permitida.....          | $P$            |
| Tamanho da amostra.....                                       | $n$            |
| Tamanho da sub-amostra.....                                   | $n_1$          |

## RESUMO

Alves, Rafaela Cardoso. **Controle Estatístico da Qualidade de sementes de soja armazenadas. Um Estudo Via Predição de Quantidades Aleatórias.** 2004 60p. Dissertação (Mestrado em Estatística e Experimentação Agropecuária) – Universidade Federal de Lavras, MG.<sup>1</sup>

No período de armazenamento das sementes, geralmente ocorre uma deterioração destas sementes, o que pode acarretar prejuízos na comercialização. Assim, torna-se necessário a utilização de uma metodologia adequada para prever o número de sementes inviáveis em um determinado tempo.

O objetivo deste trabalho foi estabelecer uma metodologia para avaliar a viabilidade (potencial de germinação) das sementes armazenadas, a partir da predição pontual e intervalar do número de sementes inviáveis num intervalo de tempo futuro. Para realizar este estudo, foram utilizados dois tipos de dados: reais e simulados. Os dados reais foram obtidos em experimento no Laboratório de Análise de Sementes da UFLA. O teste de tetrazólio (teste destrutivo) foi utilizado para avaliar a viabilidade das sementes de soja em dois tipos de cultivares diferentes, M.Soy e Serra Negra. Pela metodologia usada, constatou-se que os dados reais e os dados obtidos via simulação seguem uma distribuição Weibull e que é possível conhecer a distribuição dos tempos de vida, quando se trabalha com ensaios destrutivos. O cálculo da probabilidade de cobertura comprovou que os intervalos de predição, para o número de sub-amostras falhas em um intervalo de tempo futuro contidos nos resultados, são confiáveis.

---

<sup>1</sup> Orientador: Prof. Dr. Mário Javier Ferrua Vivanco – UFLA.  
Co-orientadora: Prof. Dra Maria Laene Moreira Carvalho - UFLA.

## **Abstract**

**Alves, Rafaela Cardoso. Statistical Quality Control in Stowage Soy bean Seed: A Random Quantities Prediction Study. Lavras: UFLA, 2004. 60p. (Dissertation – Master Program in Agronomy / Major in Statistics and Agricultural Experimentation)<sup>1</sup> – Federal University of Lavras**

An appropriate methodology is needed in the study the number of failure prediction in the time. This study can be applied to evaluate the seed lots unviability in stowage period. The aim of this work is to present an adequate in methodology to estimate the range of unviable seed in the time future interval. The study was conducted using real data as well as simulated ones. The real data were count data of tetrazólio test results obtained at UFLA'S evaluation seed laboratory for two soy bean varieties( M.Soy and Serra Negra). The life time data were modelated by a Weibull distribution. The coverage probability evaluated censured the efficiency of methodology in the prediction of future failure interval.

---

<sup>1</sup> Guidance Committee: Mário Javier Ferrua Vivanco (Orientador)  
Maria Laene Moreira de Carvalho (Co-Orientadora)

# 1 INTRODUÇÃO

Estudos de Predição de Quantidades Aleatórias em que se deseje prever o número futuro de falhas em um determinado tempo são de grande importância e ajudam em várias áreas da ciência. Uma aplicação destes estudos pode ser feita na avaliação da deterioração (perda da qualidade inicial da semente) de lotes de sementes em um determinado período de armazenamento, que são importantes na tomada de decisões quanto à compra ou venda destes lotes. Esta deterioração ocorre por vários fatores que afetam a qualidade da semente. Os principais fatores são: danos mecânicos, resultantes de impactos físicos durante a colheita, trilha, secagem, beneficiamento, transporte e semeadura.; deterioração por umidade, que resulta da exposição das sementes a ciclos alternados de condições ambientais úmidas e secas na fase final de maturação, antes da colheita; danos causados por percevejos. Danos por seca e altas temperaturas; danos de secagem; danos de geada.

Na avaliação da qualidade dos lotes foi utilizado o método do teste de tetrazólio que, além de avaliar a viabilidade, fornece o diagnóstico das possíveis causas responsáveis pela redução da qualidade.

A análise da viabilidade das sementes, isto é, se existe potencial de germinação, pode ser feita a partir do conhecimento do número de sementes inviáveis, no intervalo de tempo compreendido entre o início do armazenamento até um tempo censurado, sendo assim possível prever o número de sementes não viáveis, num intervalo de tempo futuro preestabelecido. Para tal utilizam-se Predições Pontual e Intervalar em dados obtidos via experimento e via simulação.

Vários trabalhos relacionados com a predição de quantidades aleatórias são encontrados na literatura: intervalos de confiança e teste de hipóteses para a

razão de duas proporções de uma população multinomial realizado por Nelson (1972), o mesmo em 1973 apresenta também um trabalho sobre intervalos de predição. Uma boa discussão sobre a distribuição Weibull pode ser encontrada em Nelson (1985). Escobar e Meeker (1999), descrevem métodos para a construção de limites e intervalos de predição, utilizando dados de novas amostras e dados dentro de amostras. Nelson (2000), citado por Nordman e Meeker (2002), propôs um método para determinar o preditor pontual e três para determinar intervalos de predição para o número futuro de tubos falhados em um trocador de calor. Para cada procedimento de predição intervalar proposto por Nelson (2000), Nordman e Meeker (2002), avaliaram a probabilidade de cobertura. Entretanto, estes trabalhos estão relacionados com dados de engenharia, não sendo encontrada literatura relacionada a estudos de viabilidade de sementes.

Portanto, o objetivo desta pesquisa é propor um método estatístico para prever o potencial de armazenamento de lotes de sementes, a partir da predição do número de sementes viáveis num determinado intervalo de tempo futuro.

O estudo envolve a determinação da distribuição probabilística do tempo até a deterioração das sementes (via simulação), a avaliação da probabilidade de cobertura do preditor intervalar (via simulação) e a comparação de resultados: simulação versus dados reais (obtidos experimentalmente).

No capítulo 2 apresenta-se o “estado da arte”, relacionado com a pesquisa.

O capítulo 3 propõe uma metodologia para a determinação do número de sementes inviáveis em um intervalo de tempo futuro. Esta metodologia também poderá ser aplicada em outras áreas da ciência, em que seja necessário predizer valores de quantidades aleatórias.

Os resultados obtidos no capítulo 4 são de aplicação direta na comercialização de sementes, sendo apresentados em tabelas e gráficos, fornecendo condições para se chegar a conclusões claras e objetivas.

## 2 REFERENCIAL TEÓRICO

Quando o estudo de interesse trata da predição de variáveis aleatórias, pode-se falar em estudo via predição de quantidades aleatórias.

Um exemplo da necessidade de se predizer quantidades aleatórias é o caso das sementes em que se deseja avaliar a viabilidade das sementes armazenadas, isto é, número de sementes deterioradas num determinado intervalo de tempo futuro. Tal predição é de interesse para empresas que comercializam sementes, agricultores e pecuaristas, os quais freqüentemente se defrontam com problemas na escolha dos lotes que pretendem adquirir.

Esta predição de quantidades aleatórias pode ser feita através de Preditores Pontuais e Intervalares.

Preditores Pontuais são usados quando se é necessário encontrar um valor numérico único que esteja satisfatoriamente próximo do verdadeiro valor da variável aleatória, ou seja, obtemos um único número que representa o valor mais plausível da variável aleatória. Sabemos que a predição pontual quase nunca é exatamente igual ao verdadeiro valor. Então, para melhorar a estimação também é desejável encontrar um Preditor Intervalar, o qual é um intervalo de valores satisfatórios para a variável aleatória. Baseando-se em dados amostrais, é possível construir um intervalo de valores para predizer a variável aleatória.

A predição de quantidades aleatórias descreve informações futuras baseadas em observações ou dados presentes ou passados. Tal predição baseia-se em modelos probabilísticos que fornecem intervalos ou regiões de interesse para estes dados futuros, sendo usados na obtenção de preditores pontuais e intervalares.

Neste trabalho será estudada a predição futura de quantidades aleatórias. Estudos deste tipo são realizados atualmente na engenharia, geralmente quando

se quer prever o número de componentes de um determinado equipamento, que falharão num intervalo de tempo futuro.

Um trabalho que guarda certa relação com a predição de quantidades aleatórias é o realizado por Nelson (1972), que apresenta intervalos de confiança e testes de hipóteses para a razão de proporções de uma população multinomial, usando amostra aleatória simples de uma população infinita. O autor apresenta três exemplos (estudo da preferência do consumidor, escolha de votos, tempo de vida de uma moeda em circulação), ilustrando o uso de intervalos de confiança para a razão de proporções.

Nelson (1973), apresenta intervalos de predição de amostras futuras contendo resultados de predição simples (de uma população com  $N$  unidades, todas elas com igual probabilidade de seleção) e intervalos de predição simultâneos e suas definições. Estes intervalos de predição para uma amostra futura são apresentados para populações: normal, binomial, poisson e exponencial. Também são discutidos intervalos de predição não paramétricos, usando diferentes exemplos para cada população apresentada. O mesmo autor sugere que usando outras distribuições como Weibull, Valor Extremo, Lognormal e Gama também é possível o cálculo do intervalo de predição.

Em particular, uma interessante discussão sobre a distribuição Weibull pode-se encontrar em Nelson (1985), que apresenta estimadores pontuais e limites de confiança para dados censurados múltiplos e desenvolve a teoria base para metodologias particularmente usadas, quando se tem pouca ou nenhuma falha.

Escobar e Meeker (1999), descrevem métodos para a construção de limites e intervalos de predição, assumindo predição de novas amostras e predição dentro de amostras, para resultados futuros, usando dados de vida censurados.



Conceitualmente, faz-se necessária à distinção entre predição de novas amostras e predição dentro de amostras.

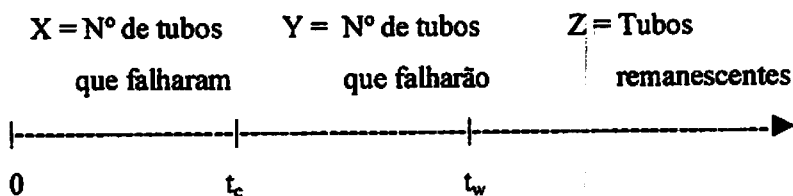
Para a predição de novas amostras, dados de uma amostra passada são usadas para fazer a predição de uma unidade desejável. Usando dados prévios é possível prever: tempo para falha de um novo item, tempo até K falhas em amostra futura com M unidades, número de falhas no tempo  $t_w$  (tempo futuro), em amostra futura com M unidades.

Para a predição dentro de amostras, o evento baseia-se em dados do começo do processo e, a partir destes dados, faz-se a predição: tempo da próxima falha, tempo até K falhas adicionais, número de falhas adicionais em um intervalo de tempo futuro desejado [ $t_w$ ,  $t_w$ ].

Contudo, o interesse é focalizar, ora nos limites de predição superior, ora nos limites de predição inferior, isto é, por exemplo: o custo de garantia máximo é mais importante que o mínimo e o tempo de falha, no começo de uma produção, é mais importante que o tempo de falha final.

Alguns estudos foram realizados, para desenvolver modelos estatísticos para predição do estado das tubulações em redes de distribuição de água. A partir do histórico de falhas foi possível prever o número futuro de falhas em cada rede. Essas predições foram usadas para tomar decisões sobre manutenção. As decisões a serem tomadas tinham como objetivo responder à seguinte pergunta: deve-se apenas consertar os pontos de falha (trincas) ou deve-se trocar a tubulação?

Nelson (2000), citado por Nordman e MeeKer (2002), desenvolveu uma metodologia para prever, mediante um preditor pontual, o número futuro de tubos a falhar em componentes de Trocadores de Calor. Neste estudo, assumiu-se que a variável aleatória T: tempo até um tubo falhar, segue uma Distribuição Weibull.



**Figura 1:** Esquematização do número de tubos falhados de trocadores de calor, em cada intervalo de tempo.

Segundo o esquema apresentado na figura 1, a predição do número futuro de falhas no intervalo de tempo  $[t_c, t_w]$  foi feita considerando a informação sobre o número de tubos falhados no intervalo  $[0, t_c]$ . É possível demonstrar que a distribuição das variáveis: número de tubos falhados ( $X$ ), número futuro de falhas ( $Y$ ) e número de tubos restantes ( $Z$ ) é trinomial, isto é,

$$P(x, y, z; p, q, r) = \frac{N!}{x!y!z!} p^x q^y r^z \quad (1.1)$$

Sendo,

$p = P(\text{haver falha no período } [0, t_c])$

$q = P(\text{haver falha no período } [t_c, t_w])$

$r = P(\text{haver falha no período } [t_w, \infty])$

A variável aleatória  $T$  (tempo até um tubo falhar) segue uma distribuição Weibull cuja função de densidade de probabilidade e função de distribuição acumulada são dadas, respectivamente, por:

$$f_t(T) = \frac{\beta}{\eta} \left(\frac{x}{\eta}\right)^{\beta-1} \exp\left[-\left(\frac{x}{\eta}\right)^\beta\right]; \quad x \geq 0$$

$$F(T) = 1 - \exp\left[-\left(\frac{x}{\eta}\right)^\beta\right]; \quad x > 0$$

em que,

$\beta > 0$  parâmetro forma.

$\eta > 0$  parâmetro escala.

Pode-se expressar p, q e r da seguinte maneira:

$$p = P(T \leq t_c) = 1 - \exp\left[-\left(\frac{t_c}{\eta}\right)^\beta\right] \quad (1.2)$$

$$\begin{aligned} q &= P(t_c \leq T \leq t_w) = P(T \leq t_w) - P(T \leq t_c) \\ &= \exp\left[-\left(\frac{t_c}{\eta}\right)^\beta\right] - \exp\left[-\left(\frac{t_w}{\eta}\right)^\beta\right] \end{aligned} \quad (1.3)$$

$$r = 1 - p - q = \exp\left[-\left(\frac{t_w}{\eta}\right)^\beta\right] \quad (1.4)$$

A partir das relações (1.2), (1.3), (1.4), Nelson (2000) propôs um método para determinar o preditor pontual e três métodos para determinar intervalos de predição para o número futuro de falhas em Trocadores de Calor.

No método do preditor pontual primeiramente estima-se o parâmetro escala desconhecido ( $\eta$ ) e, em seguida, é possível predizer Y. Esta predição é

lógica, pois,  $q$  é a probabilidade de haver falha no intervalo  $[t_c, t_w]$ . Como  $\hat{\eta}$  já foi estimado e  $\beta$  é conhecido (por estudos realizados anteriormente com o mesmo propósito, é possível conhecer o valor de  $\beta$ ), basta calcular  $\hat{q}$  e, então, obtém-se o preditor pontual  $\hat{Y} = N * \hat{q}$ .

Os três métodos de predição intervalar são: o Método da Razão de Probabilidade, o Método da Razão de Probabilidade Simplificada e o Método da Razão de Verossimilhança.

No primeiro método, o autor baseou-se na razão  $p/q$  (probabilidade de um tubo falhar em um certo intervalo de tempo, pela probabilidade de um tubo falhar em um intervalo de tempo futuro desejado).

Como,

$$\frac{q}{p} = \frac{(1 - \exp[-(t_w / \eta)^\beta]) - (1 - \exp[-(t_c / \eta)^\beta])}{1 - \exp[-(t_w / \eta)^\beta]}$$

Chamando  $A = (t_w / \eta)^\beta$  e  $B = (t_c / \eta)^\beta$

$$\frac{q}{p} = \frac{(1 - \exp(-A)) - (1 - \exp(-B))}{1 - \exp(-B)}$$

Para  $p$  pequeno tem-se  $\exp(-B) = 1$  para isso  $B$  tem que assumir um valor pequeno. Do mesmo modo, para  $q$  pequeno  $1 - \exp(-B) = 1 - \exp(-A)$ . Como  $B$  assume um valor pequeno necessariamente  $A$  também assumirá.

Portanto:

$$P(gl(Y, X, \alpha_1) \leq p/q \leq gu(Y, X, \alpha_2)) \geq 1 - \alpha$$

Simplificando os  $\eta$ , tem-se:

$$\frac{p}{q} = \frac{1}{\left(\frac{t_w}{t_c}\right)^\beta - 1}$$

O intervalo de confiança para a razão de probabilidade  $p/q$  foi dado por Nelson(1972), em que:

$$gl(Y, X, \alpha_1) = \frac{X}{[(Y+1)F(1-\alpha_1, 2Y+2, 2X)]}, \quad X \neq 0$$

$$= 0, \quad X = 0$$

$$gu(Y, X, \alpha_2) = \frac{[(X+10)F(1-\alpha_2, 2X+2, 2Y)]}{Y}$$

Sendo  $F(\gamma; m, n)$ ,  $\gamma$  é o quantil da distribuição F com graus de liberdade  $m$ (numerador) e  $n$ (denominador).

$gl(Y, X, \alpha_1)$  e  $gu(Y, X, \alpha_2)$  são respectivamente o limite inferior e o limite superior da razão  $p/q$  e  $\alpha_1 + \alpha_2 = \alpha$ . Sendo assim, podemos escrever:

$$P(gl(Y, X, \alpha_1) \leq p/q \leq gu(Y, X, \alpha_2)) \geq 1 - \alpha \quad (1.2)$$

Denotado por:

$Y_{pr}$  : o menor valor que  $Y$  assume, satisfazendo a desigualdade esquerda de (1.5).

$\tilde{Y}_{pr}$  : o maior valor que  $Y$  assume, satisfazendo a desigualdade direita de (1.5).

Assim,  $Y_{pr}$  e  $\tilde{Y}_{pr}$  formam o intervalo de predição para  $Y$ .

No segundo método, Nelson considera o número futuro de falhas ( $Y$ ) como sendo um valor bastante grande ( $Y \rightarrow \infty$ ), para  $X$  fixo, de modo que foi



possível calcular os limites de predição unilateral inferior e superior  $(\underline{Y}_{spr}, \tilde{Y}_{spr})$

no qual :

$$\underline{Y}_{spr} = [0, 5((t_w / t_c)^\beta - 1) \chi_{2x}^2 - 1].$$

$$\tilde{Y}_{spr} = [0, 5((t_w / t_c)^\beta - 1) \chi_{2x+2}^2].$$

Assim  $(\underline{Y}_{spr}, \tilde{Y}_{spr})$  formam o intervalo de predição para Y.

O terceiro método trata da razão de verossimilhança, sendo uma aproximação mais geral na qual, a partir da verossimilhança multinomial, obteve-se o intervalo de predição da razão de verossimilhança para Y,  $(\underline{Y}_{LR}, \tilde{Y}_{LR})$ .

Este método será detalhado no capítulo 3, pois é parte da metodologia que será usada neste trabalho. Porém, devido a sua grande utilização, faz-se necessário explicar os conceitos teóricos que fundamentam o teste da razão de verossimilhança.

O método da máxima verossimilhança (ML) é um método que apresenta certas vantagens, especialmente quando os parâmetros a serem determinados originam-se de uma população com dados censurados.

Seja  $x_1, x_2, \dots, x_n$  uma amostra aleatória de tamanho n, com função densidade probabilidade,  $f_x(x; \theta)$ , em que  $\theta$  é parâmetro desconhecido. A função de verossimilhança (L) desta amostra aleatória é a densidade conjunta das n variáveis aleatórias.

Então:

$$L = \prod_{i=1}^n f_{x_i}(x_i, \theta)$$

O estimador de máxima verossimilhança (MLE) de  $\theta$ , é  $\hat{\theta}$ : valor de  $\theta$  que maximiza  $L$  ou, equivalentemente, o logaritmo de  $L$ . O MLE de  $\theta$  é a solução de:

$$\frac{\partial \log L}{\partial \theta} = 0$$

A razão de verossimilhança será representada por:

$$\lambda_{H_0} = \frac{f(x; \theta_0)}{f(x; \hat{\theta})}$$

Em que

$$H_0 : \theta = \theta_0$$

**Teorema:** Se  $\hat{\theta}$  tem distribuição normal assintótica então,

$$\lim_{n \rightarrow \infty} P(-2 \log \lambda_{H_0} < \chi^2 | \theta_0) = \frac{1}{\sqrt{2\pi}} \int_0^{\chi^2} u^{-\frac{1}{2}} e^{-\frac{1}{2}u} du$$

Se  $H_0$  é verdadeiro,  $-2 \log \lambda_{H_0}$ ,  $n = 1, 2, \dots$  converge para uma distribuição  $\chi_1^2$ .

**Prova:** Como  $\hat{\theta}$  converge para  $\theta_0$  quando  $n \rightarrow \infty$ , para  $\varepsilon > 0$  arbitrário existe

$n_\varepsilon$  tal que a probabilidade excede  $1 - \varepsilon$  na seguinte equação:

$$\log f(x; \theta_0) = \log f(x; \hat{\theta}) + \frac{1}{2} \left[ \frac{\partial^2}{\partial \theta^2} (\log f(x; \theta)) \right] * (\theta_0 - \hat{\theta})^2 \quad (1.3)$$

Então a equação (1.6) pode ser reescrito como:

$$-2 \log \left( \frac{f(x; \theta_0)}{f(x; \hat{\theta})} \right) = -1/n \left[ \frac{\partial^2}{\partial \theta^2} (\log f(x; \theta)) \right] * [\sqrt{n} (\theta_0 - \hat{\theta})]^2 \quad (1.4)$$

O lado esquerdo de (1.7) é  $-2 \log \lambda_{H_0}$ . O fato de a probabilidade (1.7) exceder  $1 - \epsilon$  para todo  $n > n_\epsilon$  implica que o lado esquerdo e direito de (1.7) são seqüências de variáveis aleatórias para  $n = 1, 2, \dots$  os quais convergem em distribuição(  $x_n$  converge em distribuição para  $x$ , quando  $n \rightarrow \infty$ , se  $F_n(x) \rightarrow F(x)$  para todo  $x$  ponto de continuidade de  $F$ ) para  $\chi_1^2$ . A expressão

$$-1/n \left[ \partial^2 / \partial \theta^2 (\log f(x; \theta)) \right] \quad (1.5)$$

converge em probabilidade( Para valores grandes de  $n$  ( $n \rightarrow \infty$ ) as variáveis  $x_n$  e  $x$  são aproximadamente iguais com probabilidade bem alta); para  $B^2(\theta_0; \theta_0)$  em

$$\text{que: } B^2(\theta_0; \theta_0) = \int_{-\infty}^{\infty} (S(x, \theta_0))^2 \partial F(x; \theta_0) \text{ e } S(x; \theta_0) = \frac{\partial}{\partial \theta_0} \log \partial F(x; \theta_0)$$

Além disso,  $\sqrt{n}B(\theta_0, \theta_0)(\theta_0 - \hat{\theta})$  converge para  $N(0,1)$ . Por isso, a seqüência de variáveis aleatórias do lado direito de (1.7) converge para uma distribuição  $\chi_1^2$ . Então, concluímos que a expressão do lado esquerdo de (1.7), isto é,  $-2 \log \lambda_{H_0}$  converge para uma distribuição  $\chi_1^2$ .

Quando  $\theta$  é  $r$ -dimensional  $-2 \log \lambda_{H_0}$  converge para  $\chi_r^2$  e a prova é similar ( wilks.S.S 1962).

Nordman e Meeker (2002), avaliaram a probabilidade de cobertura para cada procedimento de predição intervalar, proposto por Nelson (2000), concluindo que o procedimento mais adequado é o da razão de verossimilhança.

Dois aspectos devem ser ressaltados em relação aos estudos apresentados acima por Nelson (1972,2000) e Nordman e Meeker (2002):

1.A distribuição assumida para o tempo até a falha é a Distribuição Weibull.



## **2-Os Teste que avaliam se um tubo está deteriorado ou não são NÃO-DESTRUTIVOS.**

Segundo a Associação Brasileira de Ensaio Não Destrutivos (ABENDE), os Ensaio Não Destrutivos (END) são realizados em materiais, acabados ou semi acabados, para verificar a existência ou não de descontinuidades ou defeitos, através de princípios físicos definidos, sem alterar suas características físicas, químicas, mecânicas ou dimensionais e sem interferir em seu uso posterior.

Constituem uma das principais ferramentas do controle de qualidade de materiais e produtos, contribuindo para a qualidade, a redução dos custos e aumentar a confiabilidade da inspeção.

São utilizados na fabricação, montagem, inspeção em serviço e manutenção, sendo largamente aplicados em soldas, fundidos, forjados, laminados, plásticos, concreto, entre outros, nos setores petróleo/petroquímico, nuclear, aeroespacial, siderúrgico, ferroviário, naval, eletromecânico, e automotivo.

Os END incluem métodos capazes de proporcionar informações a respeito do teor de defeitos de um determinado produto, das características tecnológicas de um material, ou ainda, da monitoração da degradação em serviços de componentes, equipamentos e estruturas.

Os métodos mais usuais de END são: ensaio visual, líquido penetrante, partículas magnéticas, ultra-som, radiografia (Raios-X e Gama), correntes parasitas, análise de vibrações, termografia, emissão acústica, estanqueidade e análise de deformações.

Até aqui foram apresentados estudos em que a coleta da informação é proveniente de experimentos não destrutivos, porém, o que fazer quando a coleta vem de experimentos destrutivos?

Na literatura das ciências agrárias, particularmente na área de SEMENTES, segundo Ellis (1984), o tempo até uma semente deteriorar segue uma Distribuição normal. Além disso, os testes mais comuns que avaliam se uma semente está deteriorada ou não são DESTRUTIVOS (Marcos Filho, J. et al, 1987).

Ensaio destrutivo, como o próprio nome diz, destroem o elemento avaliado e, portanto, não mantêm a estrutura em condições idênticas às existentes antes do ensaio, isto é, geralmente inutiliza-se a estrutura ou modelo ao qual está sendo aplicado, ficando esses métodos restritos à pesquisa ou à utilização de amostras representativas.

No capítulo 3 será estabelecida uma metodologia para prever quantidades aleatórias quando os ensaios são destrutivos, como ocorre no caso da avaliação da viabilidade de sementes.

### 3 MATERIAL E MÉTODOS

Este trabalho foi realizado por meio de simulações e experimento com dados reais para determinação do número de sementes viáveis em um intervalo de tempo futuro.

O programa utilizado para a realização das simulações, foi desenvolvido em linguagem C(Kernighan and Richte. 1988). Essa escolha ocorreu devido à fácil adaptação à metodologia na obtenção dos tempos de vida desejados. O SAS<sup>®</sup>(SAS INSTITUTE,1990), de uso restrito a assinantes, é um pacote estatístico que possui várias rotinas (procedimentos). Neste trabalho, utilizaram-se os procedimentos Lifetest e Lifereg. O primeiro usado para encontrar a distribuição probabilística do tempo até a deterioração da semente e o segundo para a estimação de parâmetros. Neste procedimento os parâmetros são estimados por máxima verossimilhança, usando o algoritmo de Newton-Raphson.

Para a aplicação da metodologia proposta neste trabalho utilizaram-se dois diferentes sacos de sementes de soja para que se pudesse saber o comportamento de ambas e assim compará-las: a cultivar M-Soy com porcentagem inicial de viabilidade 15% e a cultivar Serra Negra com porcentagem inicial de viabilidade 10%. Ambas mantidas em condições normais de armazenamento.

O experimento com dados reais foi desenvolvido no Laboratório de Análises de Sementes da UFLA, utilizando-se o teste de tetrazólio, conforme metodologia descrita por França Neto et al 1998, para avaliar a viabilidade das sementes de soja ao longo do armazenamento. O teste de tetrazólio baseia-se na atividade das enzimas desidrogenases, as quais catalizam as reações respiratórias nas mitocôndrias. Quando a semente é imersa na solução de tetrazólio (sal de

tetrazólio), esta é difundida através dos tecidos, ocorrendo nas células vivas uma reação que resulta na formação de um composto vermelho, isto é, a semente ficará vermelha, caso haja atividade respiratória nas mitocôndrias. Portanto, a coloração resultante da reação é uma indicação positiva de viabilidade. Tecidos não viáveis não reagem e conseqüentemente não são coloridos. Esta análise da condição da semente é feita individualmente. Cada semente analisada individualmente é classificada como viável ou não viável.

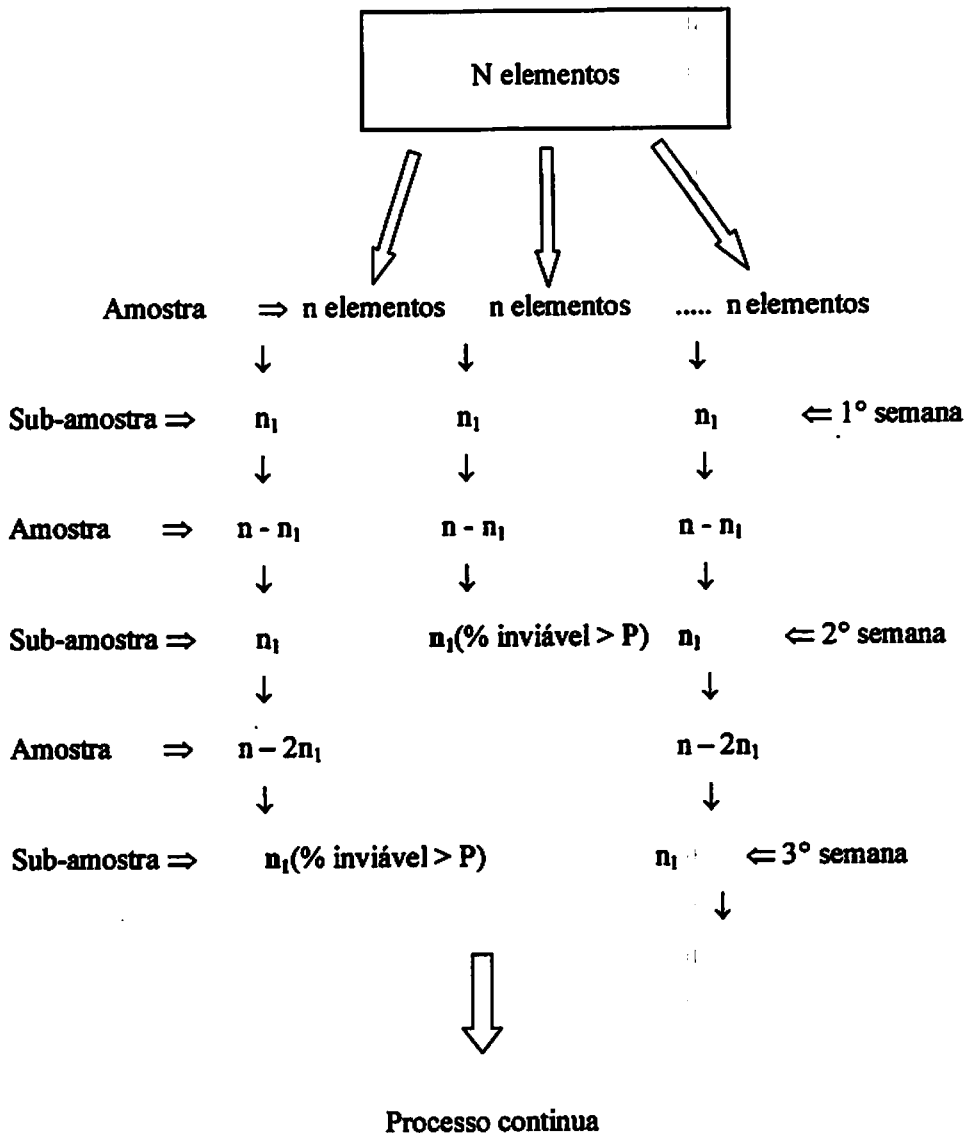
Dentro dos diversos métodos de avaliação da qualidade, adotados pela indústria de sementes no Brasil, o teste de tetrazólio tem se destacado principalmente para a soja, devido à sua rapidez e precisão. Para a realização do teste de tetrazólios algumas etapas devem ser seguidas conforme sugerido por França Neto et al. (1998): pré-condicionamento (as sementes devem ser embaladas em papel de germinação umedecido e mantidas nestas condições por um período de 16 horas, na temperatura de 25°C.), coloração (após o pré-condicionamento, as sementes são colocadas em frascos, sendo totalmente submersas na solução de tetrazólio, devem permanecer assim a uma temperatura de 30°C por aproximadamente 150 a 180 minutos.) e lavagem (alcançada a coloração, as sementes são retiradas do ambiente e são, em seguida, lavadas com água comum e devem ser mantidas submersas em água até o momento da avaliação. Caso as amostras não sejam avaliadas de imediato, devem ser mantidas em refrigerador, por até 12 horas.).

Como se desejava conhecer a distribuição probabilística dos dados correspondentes ao tempo até a deterioração da semente (T), a seguinte metodologia foi proposta:

### 3.1 Obtenção do Tempo até a Deterioração da Semente (T)

A metodologia seguida para obtenção de T está esquematizada na Figura (3.1) e é descrita a seguir.

Supondo sementes ensacadas para a comercialização, divide-se o conteúdo de um saco de sementes em “M” amostras com “n” sementes cada uma. De cada amostra retira-se, aleatória e sistematicamente, sub-amostras de tamanho “ $n_1$ ”, isto é, ter-se-ia agora “M” sub-amostras. Em cada sub-amostra aplica-se o Teste de Tetrazólio para determinar a porcentagem de sementes inviáveis. Se a porcentagem ultrapassar um determinado valor “P”, preestabelecido, então se considera a sub-amostra como inviável e registra-se o tempo em que foi realizado o teste (que em geral é dado em semanas). Uma sub-amostra considerada inviável, implica parar de retirar mais sub-amostras da amostra respectiva. Uma sub-amostra considerada viável implica na seleção de uma nova sub-amostra de tamanho “ $n_1$ ”, da amostra respectiva. Assim, aplica-se o Teste de Tetrazólio nestas novas sub-amostras e determina-se a porcentagem de sementes deterioradas que ultrapassam “P”, continua-se sub-amostrando e aplicando o Teste de Tetrazólio. O procedimento continua até que a porcentagem de sementes deterioradas da sub-amostra ultrapasse o valor “P”(no qual o dado registrado para a amostra correspondente será o tempo em que o teste foi aplicado) ou até terminar com as sementes (sendo que, neste caso, o tempo a ser registrado será um dado censurado). Dessa forma, é possível obter “M” dados, correspondente ao tempo até a perda da viabilidade de uma sub-amostra, com os quais foi possível determinar a respectiva distribuição de probabilidade.



**FIGURA 3.1 – Procedimento para obtenção do tempo até a perda da viabilidade do lote semente**

Para a realização do experimento foram utilizadas sementes de soja em que, avaliou-se um saco da cultivar M.Soy e um saco da cultivar Serra Negra, ambos com aproximadamente 120000 sementes cada. Utilizou-se: número de amostras ( $n_a$ ) = 60, tamanho da amostra ( $n$ ) = 2000, tamanho da sub-amostra ( $n_1$ ) = 10. Através do teste de tetrazólio, determinou-se a viabilidade inicial do lote que para a cultivar M.Soy foi de 90% ou  $p_{di} = 0,10$  e para cultivar Serra Negra foi de 85% ou  $p_{di} = 0,15$ . Esta viabilidade inicial do lote foi calculada usando o teste de tetrazólio em amostra representativa das duas cultivares. A probabilidade de perda da viabilidade máxima permitida (P) para as sub-amostras foi 0,15, pois, a germinação do lote na comercialização da semente de soja fiscalizada tem que ser maior ou igual a 85%.

A avaliação das sementes foi feita via sub-amostras, por que o teste de tetrazólio é destrutivo, isto é, perde-se o elemento analisado e, portanto, não é possível registrar os tempos até a inviabilidade das sementes, a menos que não fosse adotado o procedimento alternativo, proposto neste trabalho.

Sendo assim, a variável tempo até uma sub-amostra falhar foi considerada equivalente à variável tempo até a não viabilidade da semente.

Após a aplicação da metodologia foi possível obter o tempo até a perda da viabilidade da sub-amostra (T), isto é, em qual semana cada sub-amostra falhou. Caso uma amostra não perca a viabilidade até a última semana, T será considerado censurado.

### 3.2 Determinação da Distribuição do Tempo até a Perda da Viabilidade da Semente

Depois de se obter T, o próximo passo será a determinação da distribuição de T. Utilizando-se os dados (obtidos via simulação), foi possível determinar a distribuição de probabilidades de T. O procedimento lifetest do SAS® (SAS INSTITUTE,1990) foi utilizado para determinar a distribuição dos dados da variável tempo até a perda da viabilidade das sementes.

### 3.3 Predição do Número Futuro de Sub-amostras inviáveis

Uma vez determinada a distribuição de probabilidades, o seguinte passo foi prever o número de sub-amostras deterioradas, por saco, em um intervalo de tempo. O esquema apresentado na figura (3.2) ilustra o procedimento:

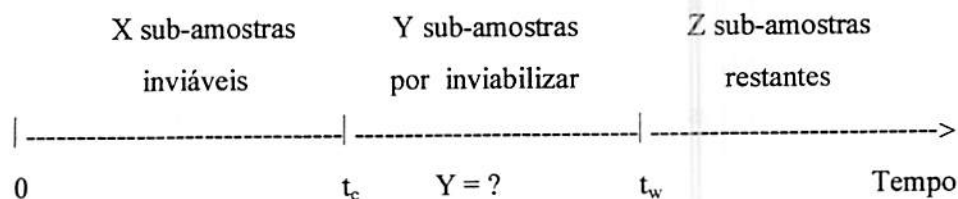


FIGURA 3.2 – Predição do número futuro de sub-amostras inviáveis

Supondo que 1 saco de sementes contenha no total “N” sub-amostras, temos que  $X + Y + Z = N$  em que,  $(X,Y,Z)$  segue uma distribuição Trinomial com probabilidades:

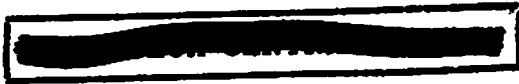
$p$  = probabilidade de uma sub-amostra falhar no intervalo de tempo  $[0,t_c)$ .

$q$  = probabilidade de uma sub-amostra falhar no intervalo de tempo  $[t_c, t_w)$ .

$r$  = probabilidade de uma sub-amostra falhar no intervalo de tempo  $[t_w, \infty)$ .

Os valores de  $p$ ,  $q$  e  $r$  dependerão da distribuição dos dados.





Obtidos T, distribuição de T e valores de p, q e r as predições pontual e intervalar do número de sementes inviáveis em um intervalo de tempo futuro foram calculadas, segundo Nelson(2000), citado por Nordman e Meeker (2002), da seguinte forma:

**Predição Pontual:** O número de sub-amostras fálhas em um intervalo do tempo futuro  $[t_0, t_w]$  será dado por:

$$\hat{Y} = N * \hat{q} \tag{3.3.1}$$

em que N é o numero total de sub-amostras e  $\hat{q}$ , o estimador de máxima verossimilhança de q.

**Predição Intervalar:** A predição intervalar é feita usando-se o procedimento da razão de verossimilhança. Sendo que (x,y,z) tem distribuição Trinomial com parâmetros p,q e r.

As hipóteses a serem testadas serão:

$$H_0: \theta = \begin{pmatrix} p \\ q \\ r \end{pmatrix} \text{ assume valores relacionados com a distribuição dos dados.}$$

$$H_1: \theta = \begin{pmatrix} p \\ q \\ r \end{pmatrix} \text{ não assume valores relacionados com a distribuição dos}$$

dados.

Sendo que  $(X,Y,Z) \sim \text{Trinomial}(p, q, r)$ , então a função densidade é dada por:

$$f_{X,Y,Z}(x,y,z) = \frac{N!}{x!y!z!} p^x q^y r^z \tag{3.3.2}$$

A verossimilhança multinomial irrestrita dos parâmetros  $p$  e  $q$  é dada por:

$$L(p, q, x, y) = \frac{N!}{x!y!(N-x-y)!} p^x q^y (1-p-q)^{N-x-y} \quad (3.3.3)$$

Maximizando  $L$ , ou seja, substituindo  $p$  e  $q$  por seus respectivos estimadores, tem-se:

$$L^*(x, y) \quad (3.3.4)$$

A função de verossimilhança restrita, em que  $p, q$  e  $r$  dependem da distribuição dos dados, é representada por:

$$K(\theta; x, y) \quad (3.3.5)$$

Maximizando  $\theta$  e substituindo em  $K(\theta; x, y)$ , tem-se:

$$K^*(x, y) \quad (3.3.6)$$

Assim a razão de verossimilhança será dada por:

$$\lambda = \frac{K^*(x, y)}{L^*(x, y)} \quad (3.3.7)$$

Aplicando-se  $-2 \log$  na razão de verossimilhança  $\lambda$ , tem-se:

$$Q(x, y) = -2 \log (K^*(x, y) - L^*(x, y)) \quad (3.3.8)$$

em que  $Q(x,y)$  segue aproximadamente uma distribuição Qui-quadrado com 1 grau de liberdade, como foi demonstrado na seção 2.

Logo,

$$\Pr ( Q(x, y) \leq \chi^2_{(1-\alpha; 1)} ) \cong 1 - \alpha \quad (3.3.9)$$

Em seguida, é possível prever  $Y$ , dado o número de falhas  $X$  no intervalo de tempo  $[0, t_c]$ , ao nível  $100(1 - \alpha)\%$ . O menor valor de  $y$  ( $\underline{Y}$ ) que satisfaz a equação (3.3.9) será considerado o limite inferior para o intervalo de predição para  $Y$ . O maior valor de  $y$  ( $\tilde{Y}$ ) que satisfaz a equação (3.3.9) será considerado o limite superior para o intervalo de predição para  $Y$ .

### **3.4 Cálculo da Probabilidade de Cobertura para Intervalos de Predição Bilateral.**

Para a obtenção da probabilidade de cobertura do preditor intervalar, usou-se o procedimento descrito a seguir:

Realizaram-se várias simulações considerando os seguintes parâmetros: número de amostras = 120, tamanho da amostra = 1000, tamanho da sub-amostra = 10,  $pdi = 0,10$  e  $P = 0,15$ . Com esses parâmetros, determinaram-se os intervalos de predição  $[\underline{Y}; \tilde{Y}]$ , com nível de confiança igual a 95%, para cada simulação realizada.

Sendo  $x$  como número de sub-amostras falhas no intervalo  $[0, t_c]$ , então,  $Y$  tem Distribuição Binomial com parâmetros  $N-x$  e  $\pi$ , em que  $\pi$  é dado pela seguinte expressão:

$$\pi = \frac{q}{q-r}$$

Logo,  $\pi$  é a probabilidade condicional em que a sub-amostra seja considerada inviável no intervalo de tempo  $[t_s, t_w]$ , dado que não perdeu a viabilidade até  $t_c$ .

Conhecendo  $x$  e obtendo  $[Y; \tilde{Y}]$ , a probabilidade de cobertura condicional (ao valor de  $x$ ) para o intervalo de predição  $[Y; \tilde{Y}]$ , com nível de confiança nominal de 95%, será dado por:

$$PC[IP(1-\alpha) | x] = \Pr[\underline{Y} < Y < \tilde{Y} | x]$$

Sendo  $(1-\alpha)$ : igual a 0.95

A probabilidade de cobertura não condicional (ao valor de  $x$ ) será dada pela seguinte expressão:

$$PC[IP(1-\alpha)] = E_x\{PC[IP(1-\alpha) | x]\}$$

Esta expressão representa, a média das probabilidades de cobertura obtidas com os dados de cada simulação.

No próximo capítulo, serão apresentados os resultados obtidos, procedendo conforme a metodologia proposta.

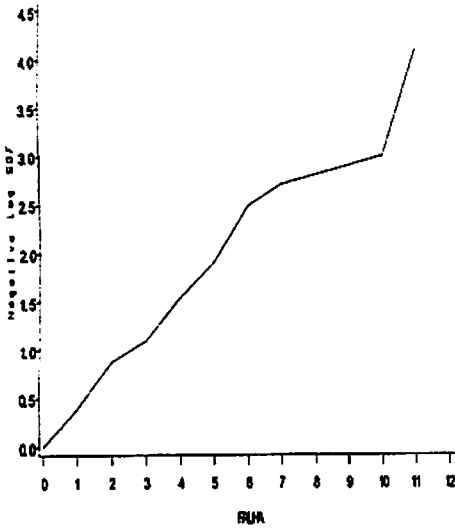
## 4 RESULTADOS E DISCUSSÃO

Neste capítulo serão apresentadas tabelas e gráficos dos resultados do número de sub-amostras futuras que seriam consideradas inviáveis no intervalo de tempo [4,16] semanas. Em virtude de serem inúmeros os resultados provenientes de dados obtidos via simulação e experimento real, foram selecionados alguns para serem apresentados. Os gráficos mostram qual distribuição probabilística melhor se adequa aos dados e as tabelas contem os valores de  $\hat{\beta}$  (estimativa do parâmetro forma da distribuição weibull),  $\hat{\eta}$  (estimativa do parâmetro escala da distribuição weibull),  $\hat{q}$  (estimativa de máxima verossimilhança de  $q$ ),  $x$  (número de sub-amostras falhas em [0,4] semanas, preditor pontual ( $\hat{Y}$ ) e preditor intervalar  $[Y; \bar{Y}]$ .

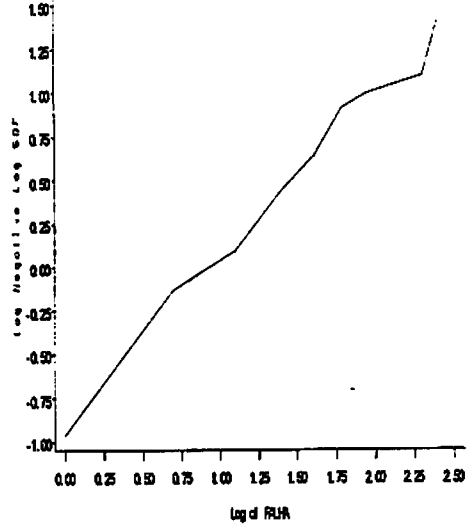
### 4.1 Adequação do modelo (dados obtidos via simulação)

Nas figuras de 1 a 14 foi possível perceber que o tempo, até a perda de viabilidade de uma sub-amostra nos dados simulados, segue o modelo probabilístico Weibull, pois, quanto mais o gráfico se aproximar de uma linha reta, melhor ajustados estarão os dados, este é, método dos gráficos de linearização. Nos gráficos denominados por (a) e (c) tem-se na abcissa número de falhas em cada semana e na ordenada os dados ajustados à distribuição normal e nos gráficos denominados por (b) e (d) tem-se na abcissa o logaritmo do número de falhas em cada semana e na ordenada os dados ajustados à distribuição weibull. Os dados utilizados para o ajuste foram: Número de amostras ( $n_a$ ) = 60 e 120, tamanho da amostra ( $n$ ) = 2000 e 1000, respectivamente. Utilizando para cada caso: probabilidade da inviabilidade

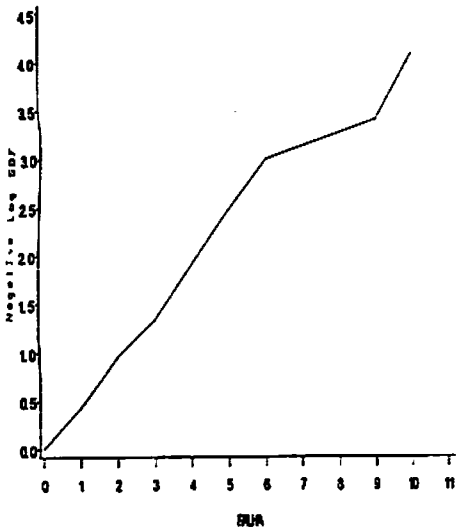
inicial(pdi) = 0,10 e 0,15; probabilidade da inviabilidade(p) = 0,15 e 0,15 respectivamente; tamanho da sub-amostra( $n_1$ ) = 3, 5, 7, 10, 15, 17, 20.



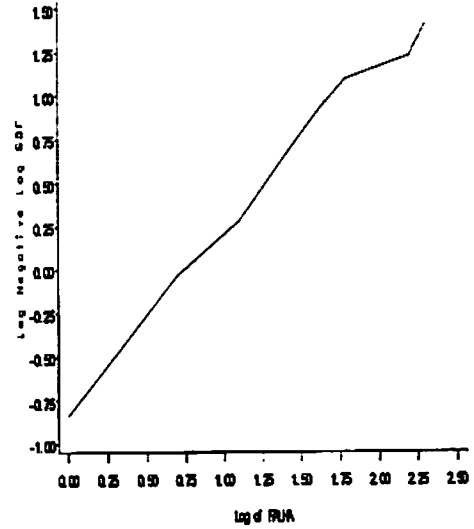
(a)



(b)

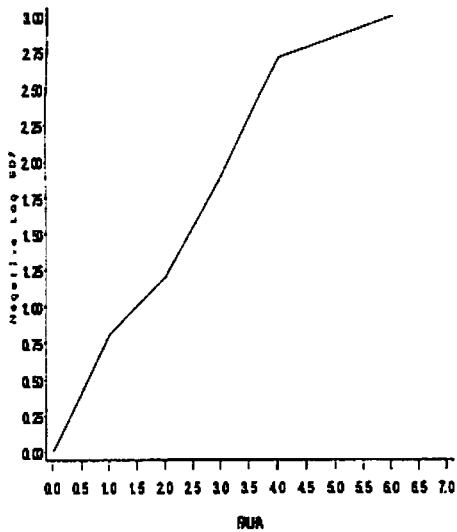


(c)

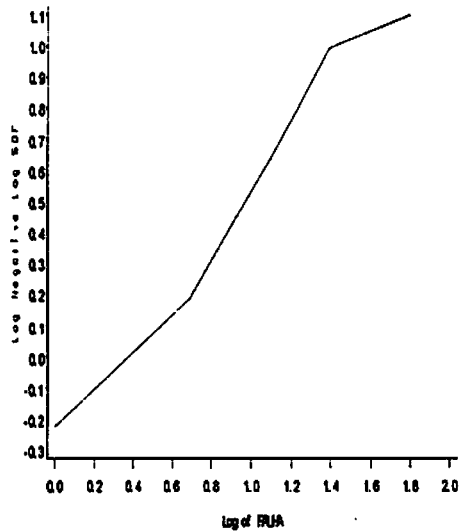


(d)

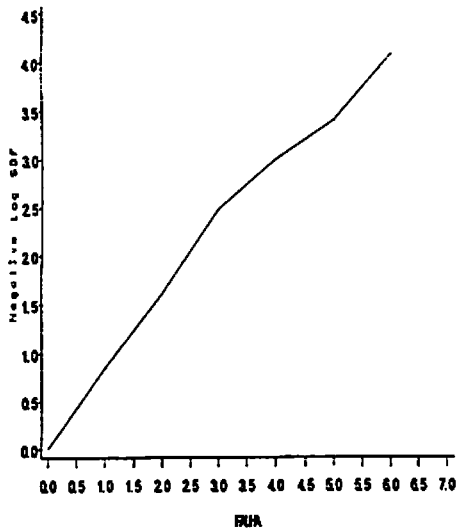
Figura 1 Gráficos de probabilidade normal(a e c) e weibull(b e d) para  $n=60$ ,  $n_1=3$ ,  $pdi=0.10$ (a e b),  $pdi=0.15$ (c e d) e  $p=0.15$



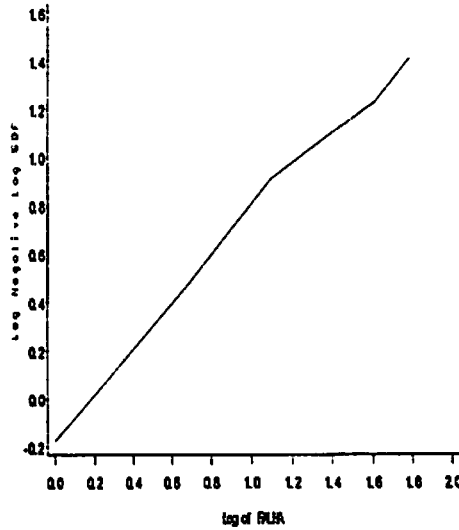
(a)



(b)

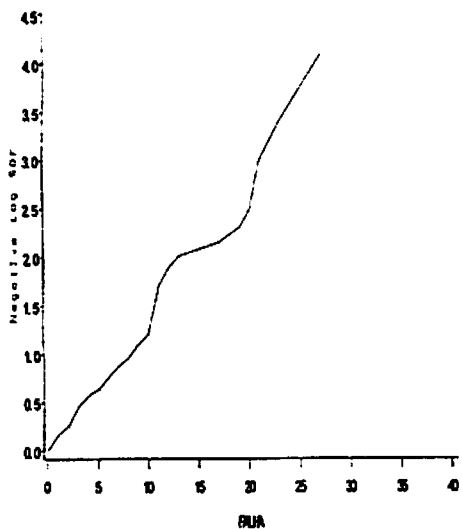


(c)

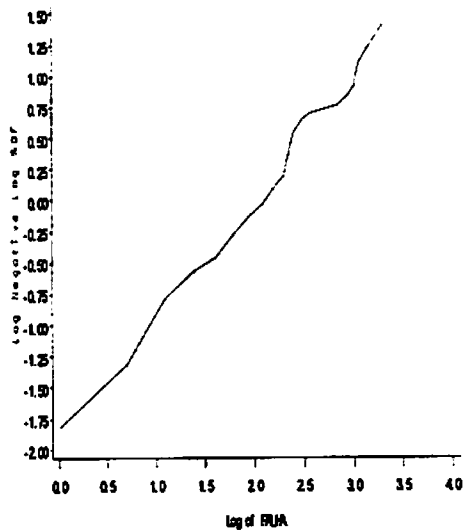


(d)

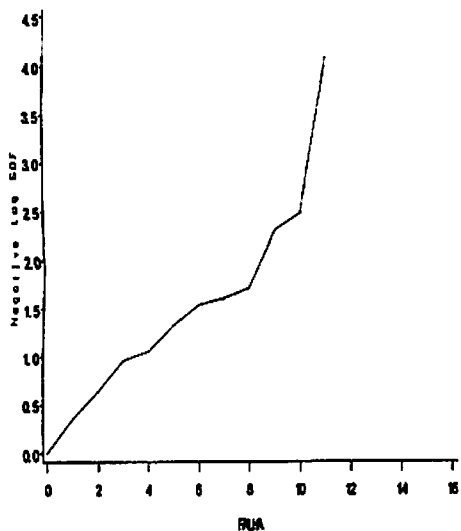
Figura 2 Gráficos de probabilidade normal(a e c) e weibull(b e d) para  $n_a=60$ ,  $n_1=5$ ,  $pdi=0.10$ (a e b),  $pdi=0.15$ (c e d) e  $p=0.15$



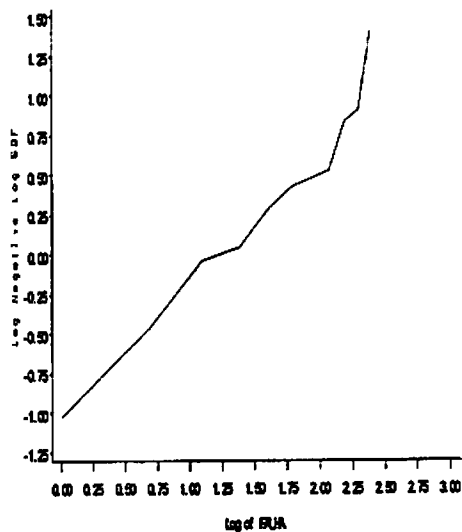
(a)



(b)



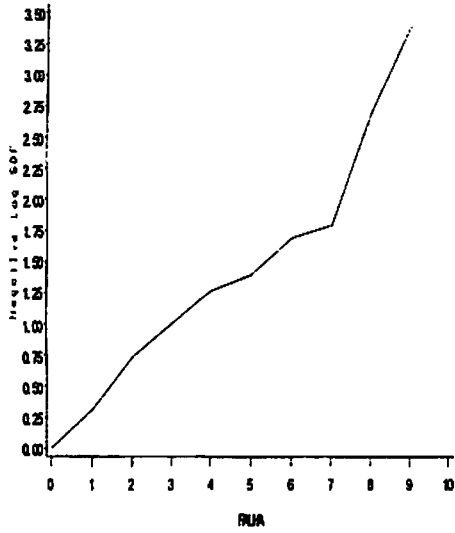
(c)



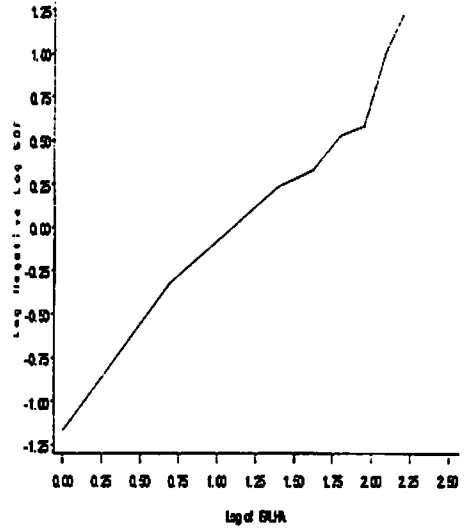
(d)

Figura 3 Gráficos de probabilidade normal(a e c) e weibull(b e d) para  $n=60$ ,  $n_1=7$ ,  $pdi=0.10$ (a e b),  $pdi=0.15$ (c e d) e  $p=0.15$

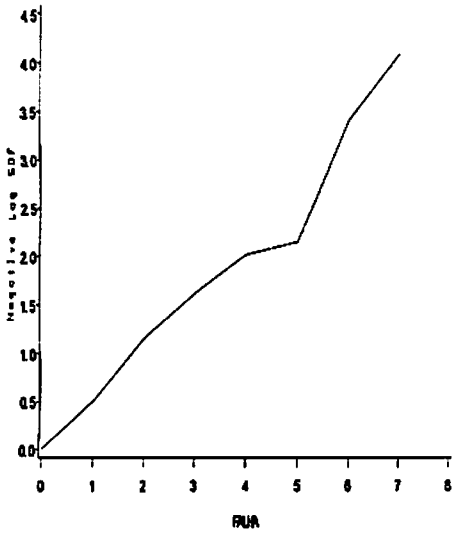




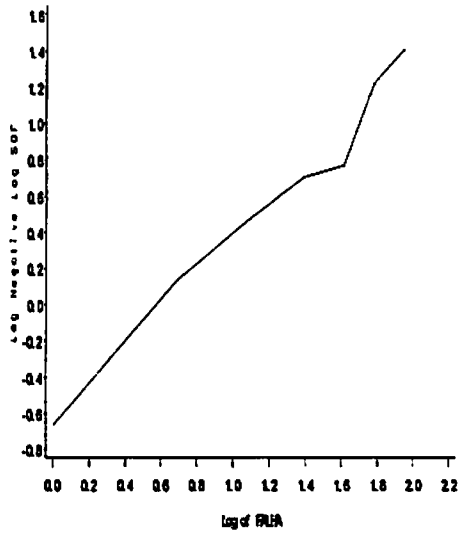
(a)



(b)

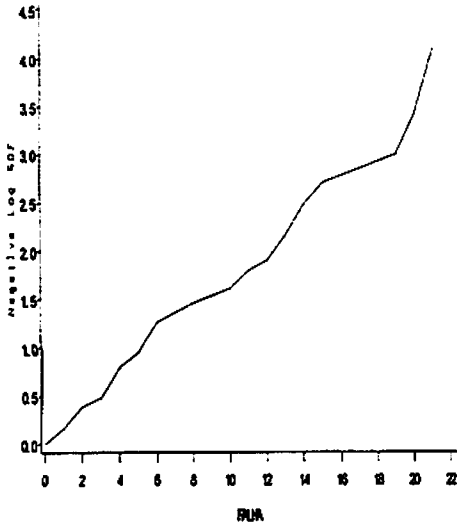


(c)

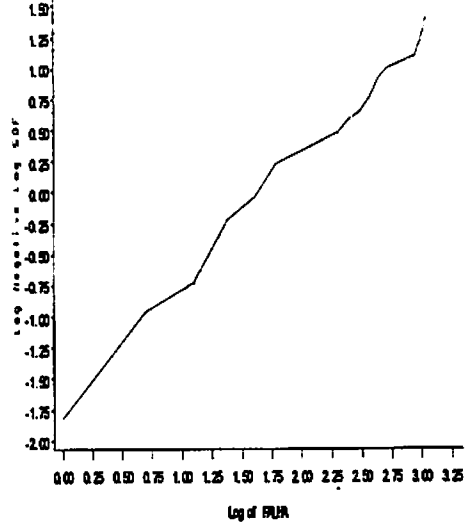


(d)

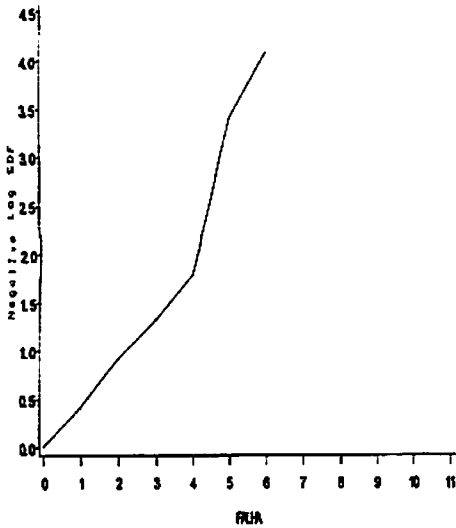
Figura 4 Gráficos de probabilidade normal(a e c) e weibull(b e d) para  $n_a=60$ ,  $n_1=10$ ,  $p_{di}=0.10$ (a e b),  $p_{di}=0.15$ (c e d) e  $p=0.15$



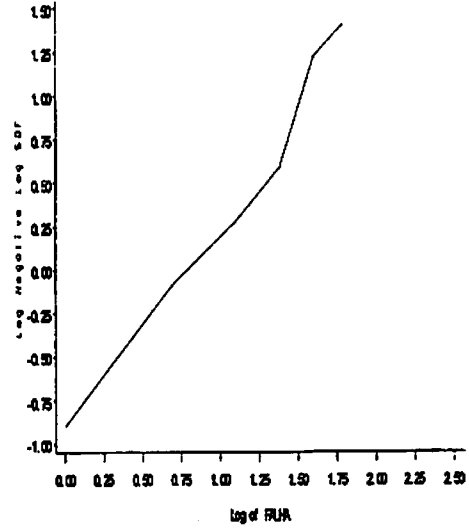
(a)



(b)

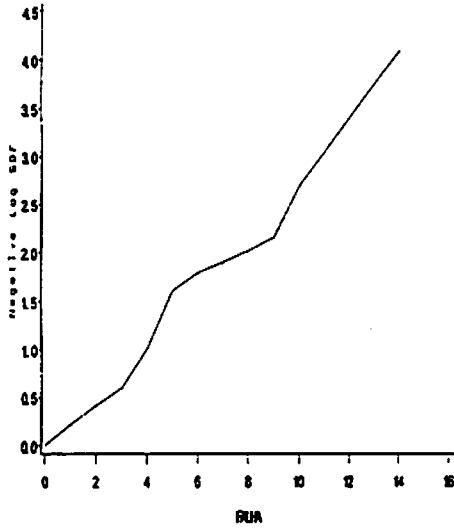


(c)

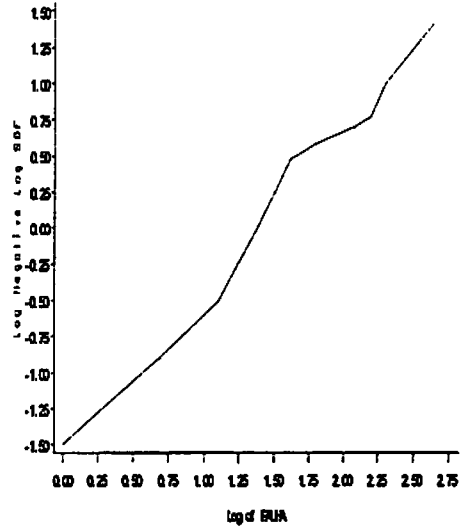


(d)

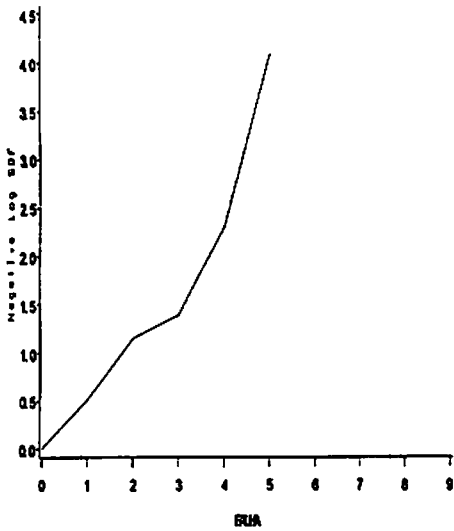
Figura 5 Gráficos de probabilidade normal(a e c) e weibull(b e d) para  $n_a=60$ ,  $n_1=15$ ,  $pdi=0.10$ (a e b),  $pdi=0.15$ (c e d) e  $p=0.15$



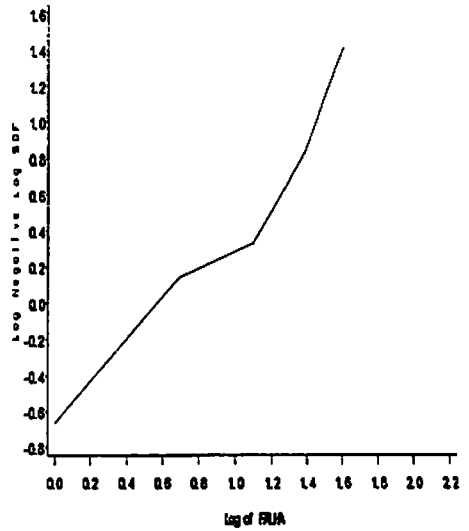
(a)



(b)

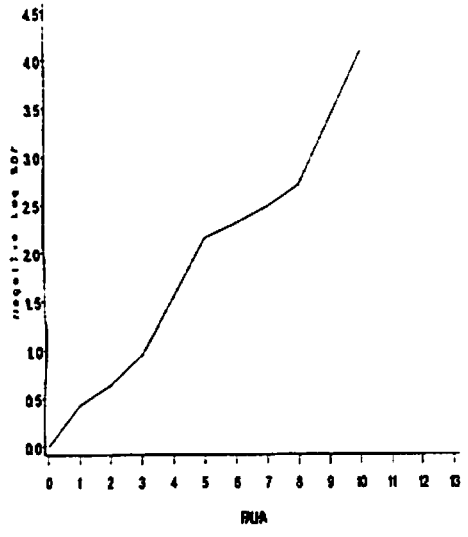
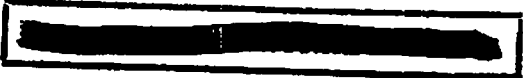


(c)

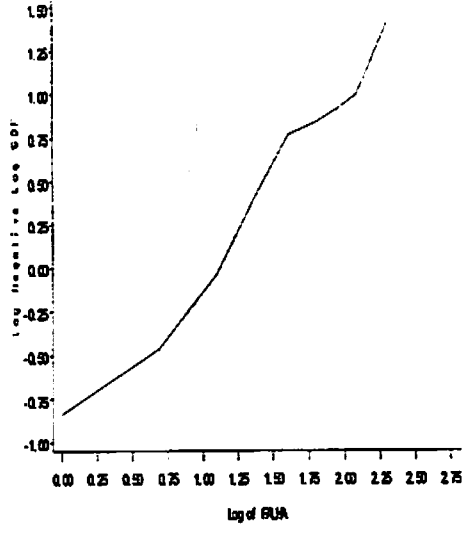


(d)

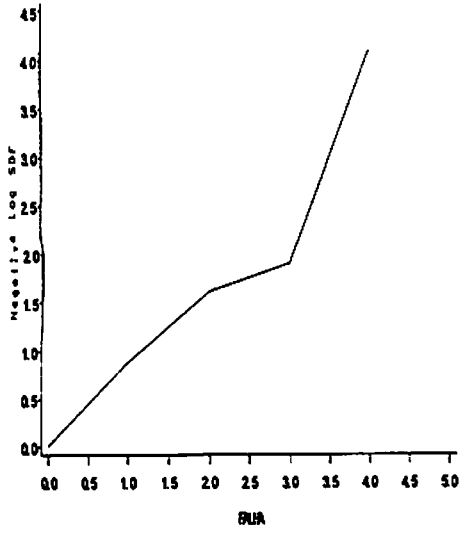
Figura 6 Gráficos de probabilidade normal(a e c) e weibull(b e d) para  $n=60$ ,  $n_1=17$ ,  $pdi=0.10$ (a e b),  $pdi=0.15$ (c e d) e  $p=0.15$



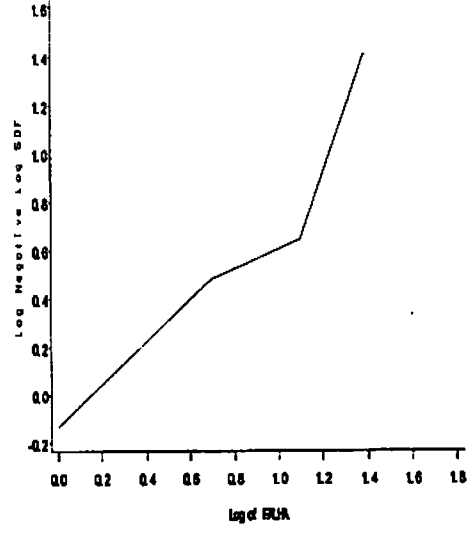
(a)



(b)



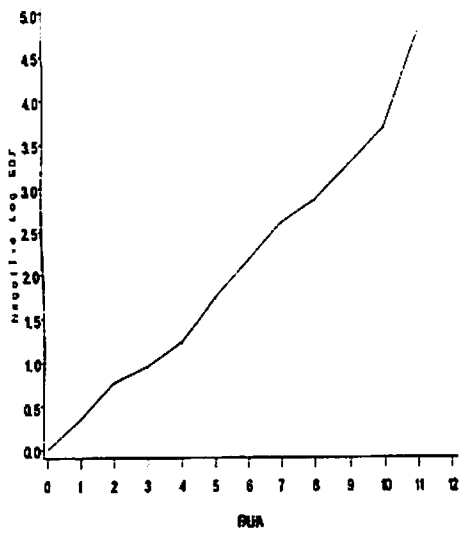
(c)



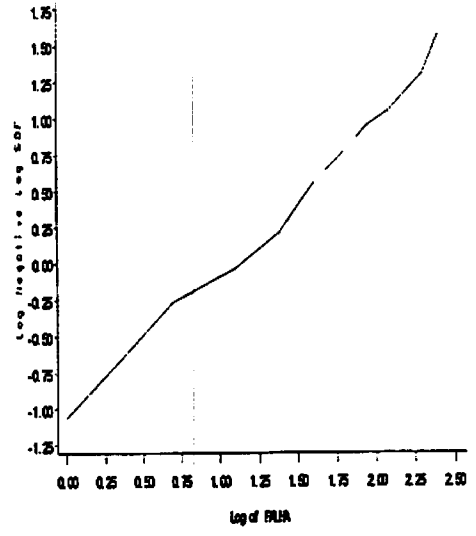
(d)

Figura 7 Gráficos de probabilidade normal(a e c) e weibull(b e d) para  $n_a=60$ ,  $n_1=20$ ,  $pdi=0.10$ (a e b),  $pdi=0.15$ (c e d) e  $p=0.15$

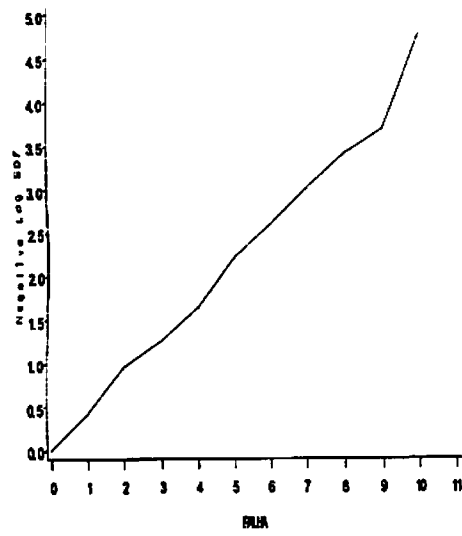
Nas figuras 1 a 7 pode-se observar que, ao avaliar 60 amostras, não foi possível afirmar que a distribuição normal ajusta-se bem ao conjunto de dados de tempo de vida de sub-amostras de sementes, já que Ellis (1984) afirma em seu trabalho que a deterioração das sementes segue uma distribuição normal. Não há uma diferença evidente entre as distribuições normal e Weibull, embora nas figuras 1, 4 e 5 exista uma tendência dos dados à distribuição Weibull.



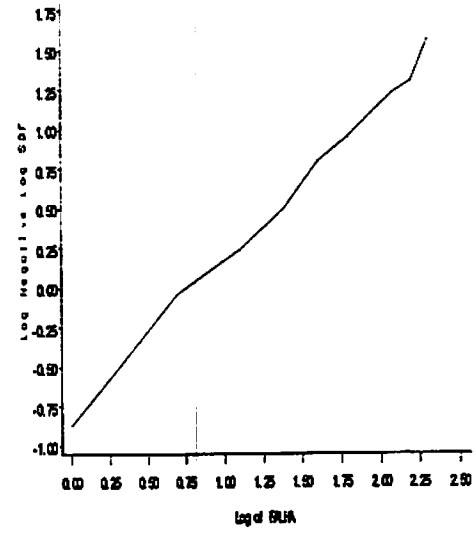
(a)



(b)

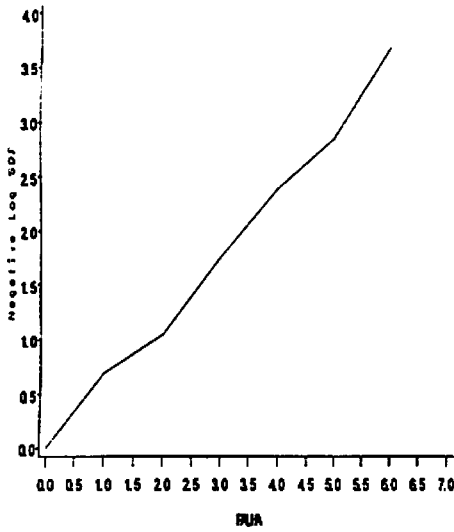


(c)

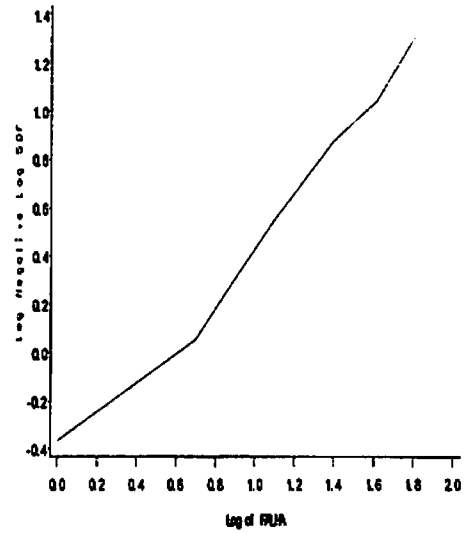


(d)

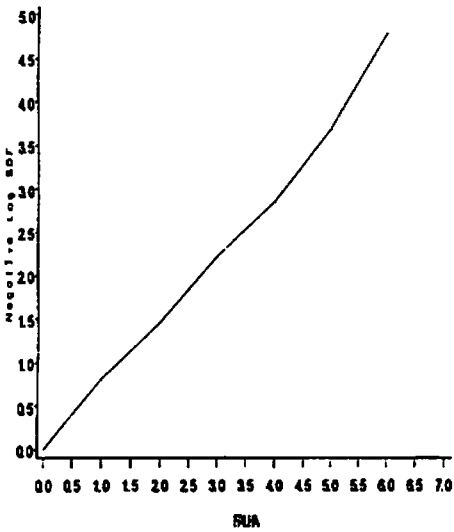
Figura 8 Gráficos de probabilidade normal(a e c) e weibull(b e d) para  $n_a=120$ ,  $n_1=3$ ,  $pdi=0.10$ (a e b),  $pdi=0.15$ (c e d) e  $p=0.15$



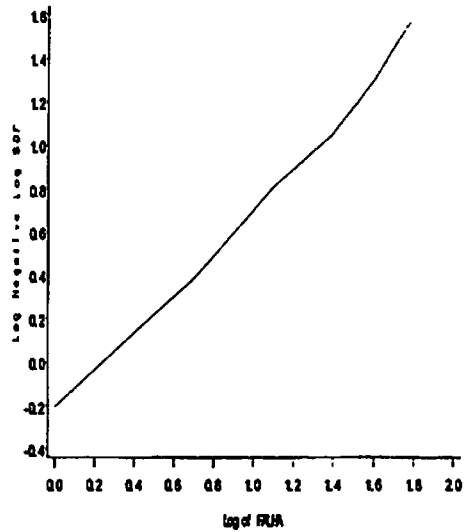
(a)



(b)

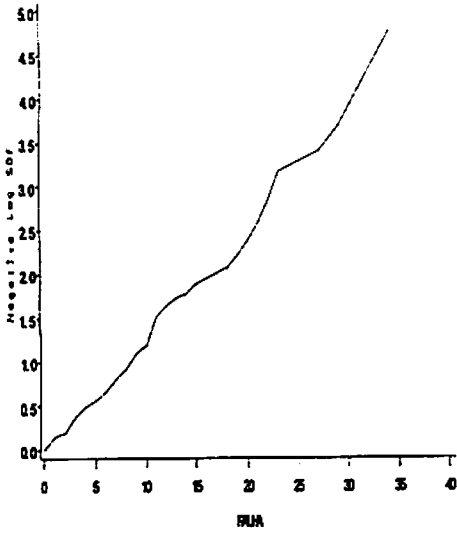


(c)

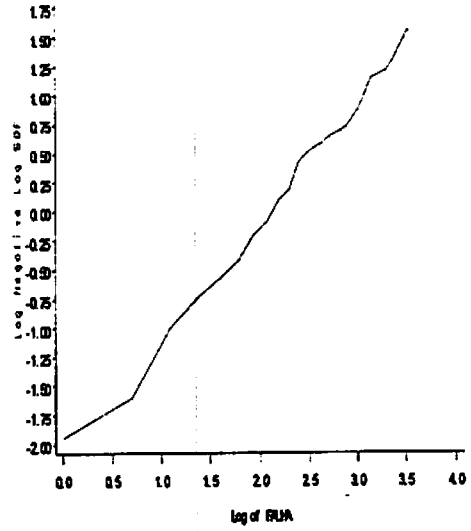


(d)

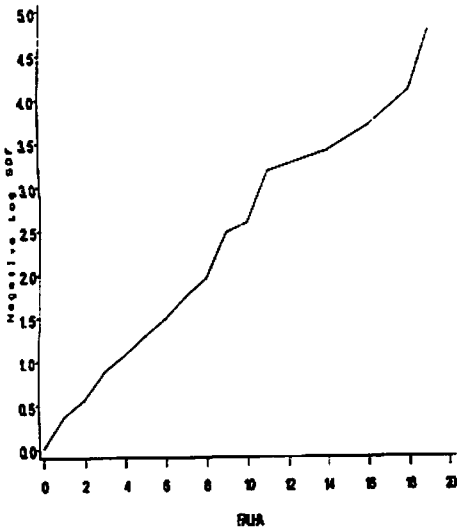
**Figura 9 Gráficos de probabilidade normal(a e c) e weibull(b e d) para  $n_a=120$ ,  $n_1=5$ ,  $p_{di}=0.10$ (a e b),  $p_{di}=0.15$ (c e d) e  $p=0.15$**



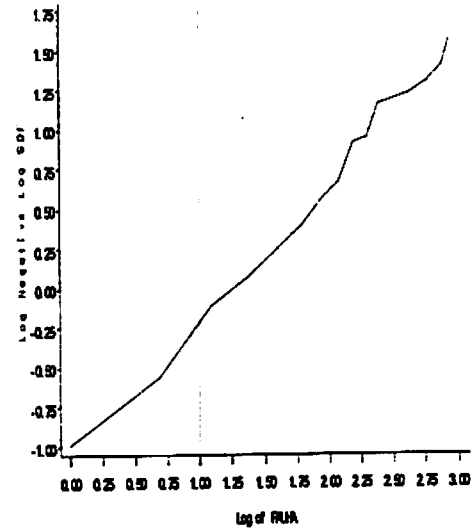
(a)



(b)



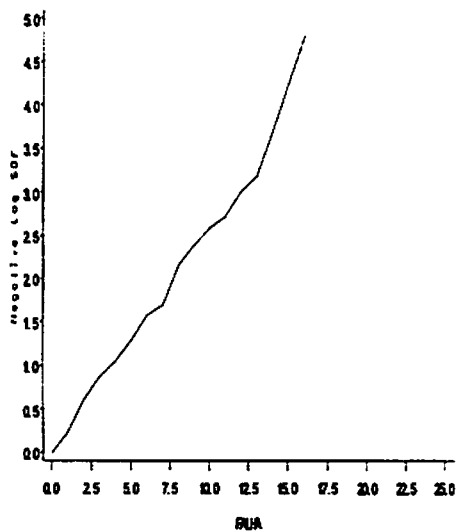
(c)



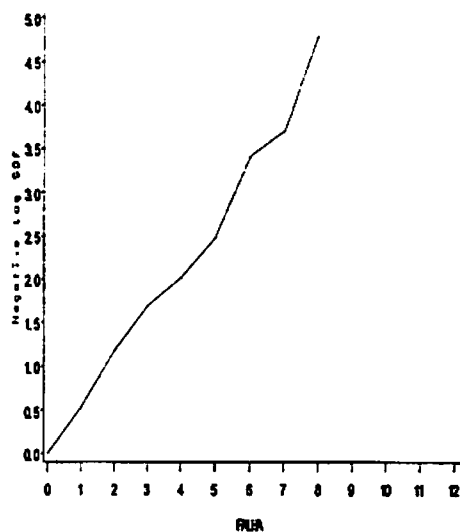
(d)

Figura 10 Gráficos de probabilidade normal(a e c) e weibull(b e d) para  $n_a=120$ ,  $n_1=7$ ,  $p_{di}=0.10$ (a e b),  $p_{di}=0.15$ (c e d) e  $p=0.15$

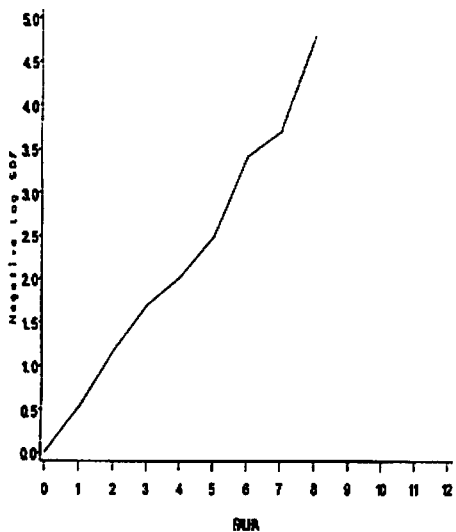




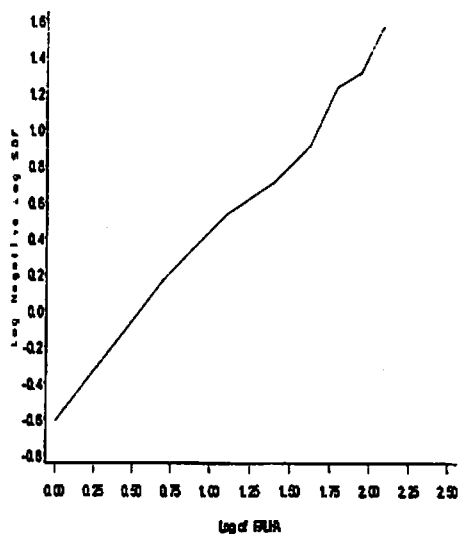
(a)



(b)

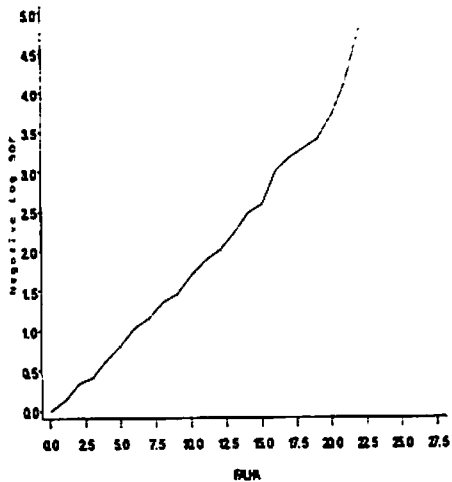


(c)

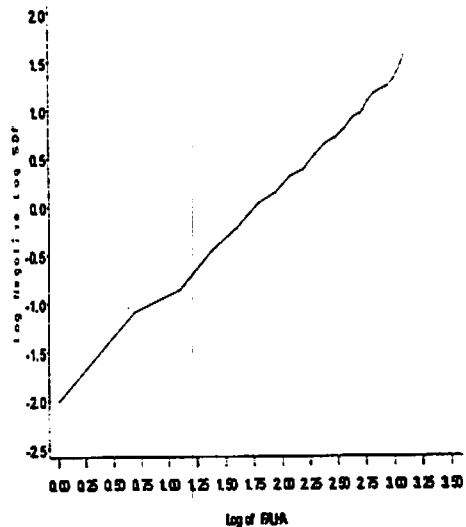


(d)

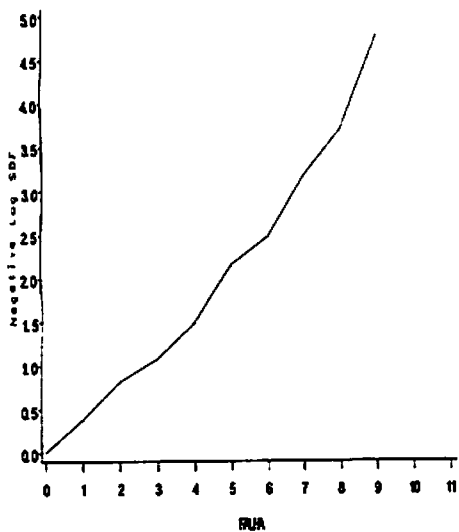
Figura 11 Gráficos de probabilidade normal(a e c) e weibull(b e d) para  $n_a=120$ ,  $n_1=10$ ,  $p_{di}=0.10$ (a e b),  $p_{di}=0.15$ (c e d) e  $p=0.15$



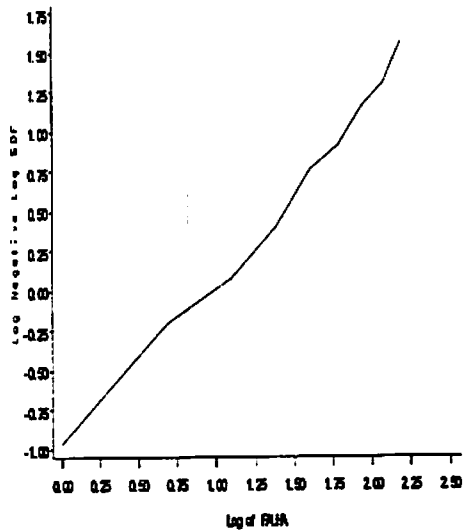
(a)



(b)

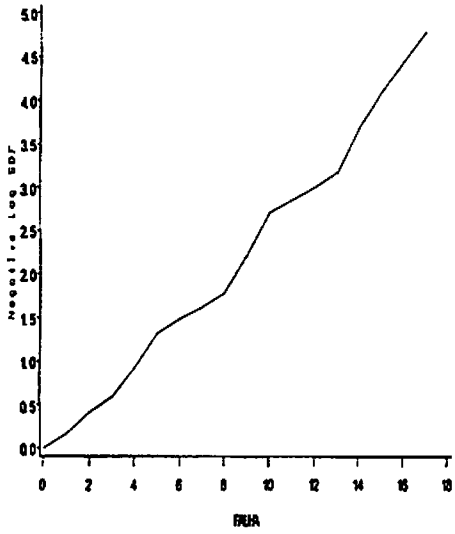


(c)

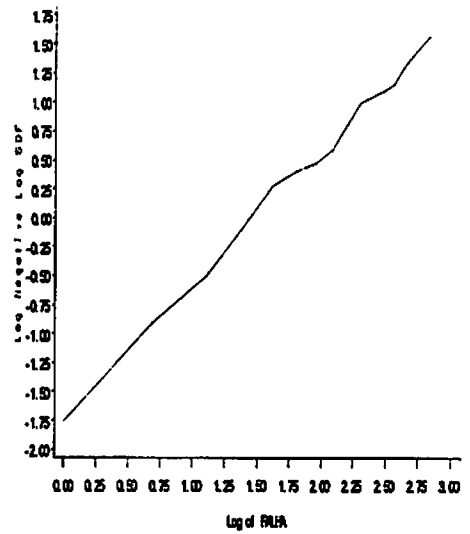


(d)

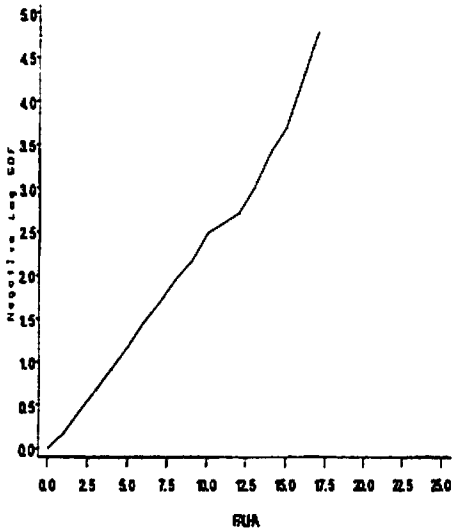
Figura 12 Gráficos de probabilidade normal(a e c) e weibull(b e d) para  $n_a=120$ ,  $n_1=15$ ,  $p_{di}=0.10$ (a e b),  $p_{di}=0.15$ (c e d) e  $p=0.15$



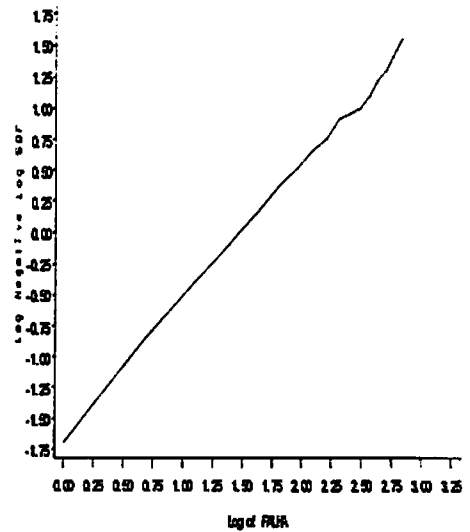
(a)



(b)

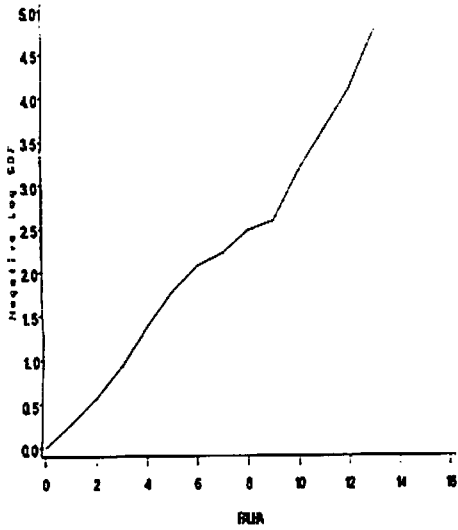


(c)

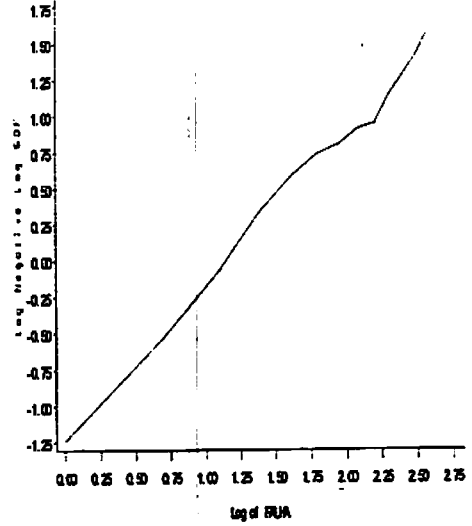


(d)

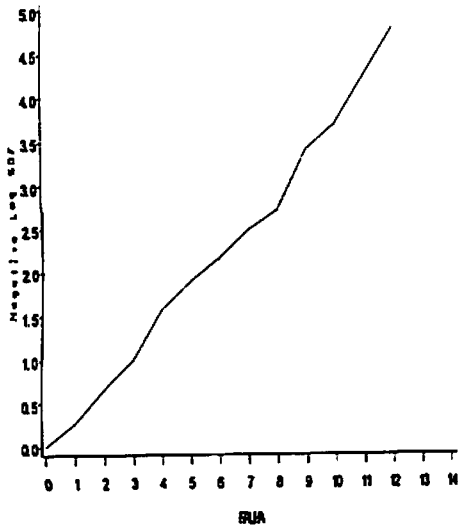
Figura 13 Gráficos de probabilidade normal(a e c) e weibull(b e d) para  $n_a=120$ ,  $n_1=17$ ,  $p_{di}=0.10$ (a e b),  $p_{di}=0.15$ (c e d) e  $p=0.15$



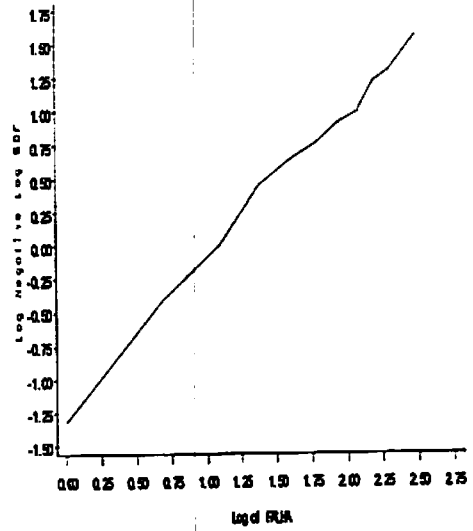
(a)



(b)



(c)



(d)

Figura 14 Gráficos de probabilidade normal(a e c) e weibull(b e d) para  $n_a=120$ ,  $n_1=20$ ,  $p_{di}=0.10$ (a e b),  $p_{di}=0.15$ (c e d) e  $p=0.15$

Pode-se observar que para número de amostra 120, a tendência à distribuição Weibull é muito melhor que a distribuição normal, sendo mais evidente nas figuras 12,13 e 14, isto é, quando o tamanho da sub-amostra aumentou.

#### **4.2 Predição Pontual e Intervalar para número futuro de sementes inviáveis utilizando dados simulados**

Sendo a distribuição Weibull a que melhor se ajusta aos dados é possível obter os valores de p,q e r que serão dados pelas seguintes expressões:

$$p = 1 - \exp\left[-\left(\frac{t_c}{\eta}\right)^\beta\right] \quad (4.1.1)$$

$$q = \exp\left[-\left(\frac{t_c}{\eta}\right)^\beta\right] - \exp\left[-\left(\frac{t_w}{\eta}\right)^\beta\right] \quad (4.1.2)$$

$$r = \exp\left[-\left(\frac{t_w}{\eta}\right)^\beta\right] \quad (4.1.3)$$

Como  $\eta$  e  $\beta$  são os parâmetros de escala e de forma, respectivamente, da distribuição Weibull, é preciso estimá-los para se obter os preditores pontual e intervalar, assim:

$\beta$  foi estimado usando o proc lifereg do SAS® (SAS INSTITUTE, 1990). Após estimar  $\beta$ , o valor de  $\eta$  foi estimado a partir da seguinte relação:

$$p = 1 - \exp\left[-\left(\frac{t_c}{\eta}\right)^\beta\right] \rightarrow \eta = \frac{t_c}{[-\log(1-p)]^{1/\beta}}$$

### Estimador de Máxima Verossimilhança de p

Sabendo que x é número de sub-amostras falhas em  $[0, t_c]$ .

$$\frac{\partial \log L}{\partial \eta} = x \left(\frac{1}{p}\right) \frac{\partial p}{\partial \eta} + (N-x) \left(\frac{1}{1-p}\right) (-1) \frac{\partial p}{\partial \eta} P(X=x) = C_{N-x} p^x (1-p)^{N-x}$$

$$L = C_{N-x} p^x (1-p)^{N-x}$$

$$\log L = \log C_{N-x} + x \log p + (N-x) \log(1-p)$$

Derivando ambos lados em relação a  $\eta$ .

$$\frac{\partial \log L}{\partial \eta} = 0 + x \frac{\partial \log p}{\partial p} \frac{\partial p}{\partial \eta} + (N-x) \frac{\partial \log(1-p)}{\partial p} \frac{\partial p}{\partial \eta}$$

$$\frac{\partial \log L}{\partial \eta} = x \left(\frac{1}{p}\right) \frac{\partial p}{\partial \eta} + (N-x) \left(\frac{1}{1-p}\right) (-1) \frac{\partial p}{\partial \eta}$$

Igualando a zero tem-se:

$$\frac{1-p}{p} = \frac{N-x}{x} \rightarrow \hat{p} = \frac{x}{N}$$

Então,

$$\hat{\eta} = \frac{t_c}{[-\log\left(1-\frac{x}{N}\right)]^{1/\beta}} \quad (4.1.4)$$

Estimando-se também q ,por máxima verossimilhança, obtém-se:

$$\hat{q} = \exp\left[-\left(\frac{t_c}{\hat{\eta}}\right)^\beta\right] - \exp\left[-\left(\frac{t_w}{\hat{\eta}}\right)^\beta\right]$$

Substituindo o valor de  $\hat{\eta}$  tem-se:

$$\hat{q} = \left[1 - \frac{x}{N}\right] - \left[1 - \frac{x}{N}\right] \left(\frac{t_w}{t_c}\right)^\beta \quad (4.1.5)$$

Como foi mencionado na seção 3.1, segundo Nelson (2000), o preditor pontual é dado pela seguinte equação:

$$\hat{Y} = N * \hat{q}$$

Como na predição intervalar utilizou-se o procedimento da Razão de Verossimilhança temos:

A função de verossimilhança multinomial irrestrita será dada por:

$$L^*(x, y) = \frac{N!}{x!y!(N-x-y)!} \left(\frac{x}{N}\right)^x \left(\frac{y}{N}\right)^y \left(1 - \frac{x}{N} - \frac{y}{N}\right)^{N-x-y} \quad (4.1.6)$$

A função de verossimilhança restrita, sob o modelo Weibull, será dada por:

$$K^*(x,y) = \frac{N!}{x!y!(N-x-y)!} \left[ 1 - e^{-\left(\frac{t_c}{\hat{\eta}}\right)^\beta} \right]^x * \left[ e^{-\left(\frac{t_c}{\hat{\eta}}\right)^\beta} - e^{-\left(\frac{t_w}{\hat{\eta}}\right)^\beta} \right]^y * \left[ e^{-\left(\frac{t_w}{\hat{\eta}}\right)^\beta} \right]^{N-x-y} \quad (4.1.7)$$

Sendo x um valor conhecido, denominou-se a Razão de verossimilhança por  $Q(x,y)$ , então:

$$Q(x,y) = -2 \left[ \log(K^*(y)) - \log(L^*(y)) \right] \leq \chi^2_{(0.95;1)} (= 3,841) \quad (4.1.8)$$

Os valores de y, que satisfizerem a inequação acima, formarão o intervalo de predição.

As tabelas 1,2,3 e 4 apresentam os valores de  $\hat{\beta}$ ,  $\hat{\eta}$ ,  $\hat{q}$ , x, predição pontual ( $\hat{Y}$ ) e predição intervalar  $[Y; \tilde{Y}]$ , variando: número de amostras(na), tamanho da amostra(n), probabilidade de deterioração inicial(pdi), probabilidade de deterioração(p) e tamanho da sub-amostra( $n_1$ ).

Em que x é o número de sub-amostra falhas em  $[0, t_c]$  semanas e y é o número de sub-amostras que falharão em  $[t_c, t_w]$  semanas. Considerando  $[0,t_c]=[0,4]$ , isto é, as 4 primeiras semanas de armazenamento da semente(1 mês) e  $[t_c,t_w]=[4,16]$ , isto é, as 12 semanas seguintes(3 meses).



Observou-se nos resultados apresentados nas tabelas 1 e 2 que, ao variar o valor da probabilidade de inabilidade inicial, existe uma influência

| $n_1$      | x  | $\hat{\beta}$ | $\hat{\eta}$ | $\hat{q}$ | $\hat{Y}$ | $[\bar{X}; Y]$ |
|------------|----|---------------|--------------|-----------|-----------|----------------|
| $n_1 = 3$  | 51 | 0.71          | 1.62         | 0.14      | 8         | [7;9]          |
| $n_1 = 5$  | 57 | 0.62          | 0.68         | 0.05      | 2         | 2              |
| $n_1 = 7$  | 39 | 0.82          | 3.8          | 0.31      | 18        | [16;20]        |
| $n_1 = 10$ | 52 | 0.66          | 1.38         | 0.12      | 7         | [6;8]          |
| $n_1 = 15$ | 50 | 0.63          | 1.59         | 0.15      | 9         | [7;10]         |
| $n_1 = 17$ | 54 | 0.62          | 1.04         | 0.09      | 5         | [4;6]          |
| $n_1 = 20$ | 59 | 0.68          | 0.5          | 0.02      | 1         | 1              |

$n = 2000, pdi = 0.15$  e  $p = 0.15$

TABELA 2. Valores das predições pontual ( $\hat{Y}$ ) e intervalar  $[\bar{X}; Y]$  para  $n_a = 60$ ,

| $n_1$      | x  | $\hat{\beta}$ | $\hat{\eta}$ | $\hat{q}$ | $\hat{Y}$ | $[\bar{X}; Y]$ |
|------------|----|---------------|--------------|-----------|-----------|----------------|
| $n_1 = 3$  | 47 | 0.75          | 2.27         | 0.2       | 12        | [10;12]        |
| $n_1 = 5$  | 56 | 0.6           | 0.76         | 0.07      | 4         | 4              |
| $n_1 = 7$  | 26 | 0.87          | 7.7          | 0.41      | 25        | [21;30]        |
| $n_1 = 10$ | 43 | 0.72          | 2.87         | 0.24      | 14        | [12;16]        |
| $n_1 = 15$ | 32 | 0.8           | 5.6          | 0.37      | 21        | [18;25]        |
| $n_1 = 17$ | 38 | 0.7           | 3.98         | 0.29      | 17        | [14;20]        |
| $n_1 = 20$ | 47 | 0.74          | 2.2          | 0.2       | 12        | [10;13]        |

$n = 2000, pdi = 0.10$  e  $p = 0.15$

TABELA 1. Valores das predições pontual ( $\hat{Y}$ ) e intervalar  $[\bar{X}; Y]$  para  $n_a = 60$ ,

significativa no número de sub-amostras falhas em [0,4] semanas (x) que, conseqüentemente afetará  $\hat{Y}$  e  $[Y; \bar{Y}]$ . A partir da sub-amostra 7, esta diferença se mostrou mais acentuada; com isso, na tabela 2 poder-se-ia afirmar que todas as 60 amostras falharão no decorrer de 4 meses, o mesmo não acontecendo na tabela 1.

Pode-se concluir então que os 5% de diferença na probabilidade de inviabilidade inicial influenciam significativamente no tempo de viabilidade da semente.

Em relação aos intervalos de predição que apresentaram apenas um valor e não limites inferior e superior como o esperado, justifica-se, pois Y é uma variável aleatória quantitativa contínua, isto é, apresenta valores reais. Porém, como Y é o número de sub-amostras que falharão em um intervalo de tempo futuro, este valor deverá ser um número natural. Por exemplo, na tabela 1,  $n_1 = 5$ , o intervalo encontrado foi [3,6;4,78], portanto considerou-se que 4 sub-amostras falharão no intervalo de tempo [4,16] semanas.

TABELA 3. Valores das predições pontual ( $\hat{Y}$ ) e intervalar  $[Y; \bar{Y}]$  para  $n_a = 120$ ,  $n = 1000$ ,  $p_{di} = 0.10$  e  $p = 0.15$

|            | x   | $\hat{\beta}$ | $\hat{\eta}$ | $\hat{q}$ | $\hat{Y}$ | $[Y; \bar{Y}]$ |
|------------|-----|---------------|--------------|-----------|-----------|----------------|
| $n_1 = 3$  | 84  | 0.72          | 3.1          | 0.26      | 31        | [28;34]        |
| $n_1 = 5$  | 109 | 0.65          | 1            | 0.1       | 11        | [11;13]        |
| $n_1 = 7$  | 47  | 0.84          | 9.2          | 0.4       | 49        | [42;56]        |
| $n_1 = 10$ | 78  | 0.81          | 3.8          | 0.31      | 37        | [34;40]        |
| $n_1 = 15$ | 56  | 0.77          | 7.3          | 0.37      | 45        | [38;51]        |
| $n_1 = 17$ | 72  | 0.72          | 4.5          | 0.31      | 38        | [33;43]        |
| $n_1 = 20$ | 90  | 0.73          | 2.6          | 0.23      | 27        | [24;29]        |

TABELA 4. Valores das predições pontual ( $\hat{Y}$ ) e intervalar [ $\underline{Y}; \tilde{Y}$ ] para  $n_a = 120$ ,  
 $n = 1000$ ,  $pdi = 0.15$  e  $p = 0.15$

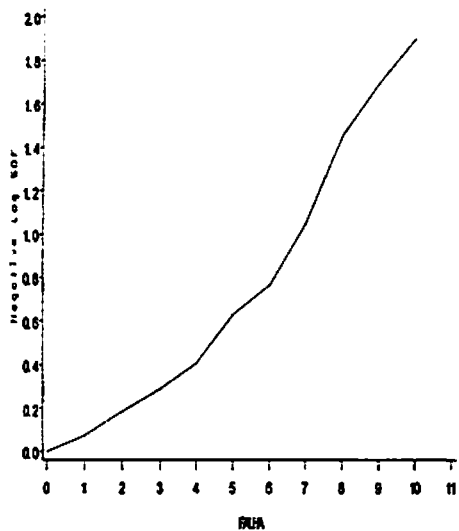
|            | x   | $\hat{\beta}$ | $\hat{\eta}$ | $\hat{q}$ | $\hat{Y}$ | $[\underline{Y}; \tilde{Y}]$ |
|------------|-----|---------------|--------------|-----------|-----------|------------------------------|
| $n_1 = 3$  | 113 | 0.61          | 0.72         | 0.06      | 7         | 7                            |
| $n_1 = 5$  | 97  | 0.70          | 1.95         | 0.18      | 21        | [19;23]                      |
| $n_1 = 7$  | 79  | 0.85          | 3.7          | 0.31      | 36        | [34;40]                      |
| $n_1 = 10$ | 104 | 0.69          | 1.45         | 0.12      | 15        | [13;16]                      |
| $n_1 = 15$ | 94  | 0.67          | 2.12         | 0.2       | 23        | [21;25]                      |
| $n_1 = 17$ | 72  | 0.77          | 4.5          | 0.33      | 39        | [35;44]                      |
| $n_1 = 20$ | 95  | 0.69          | 2.1          | 0.19      | 23        | [20;25]                      |

Comparando as tabelas 3 e 4 observa-se o mesmo comportamento em relação às tabelas 1 e 2. entretanto constatou-se que a porcentagem de sub-amostras falhas(X) no intervalo [0,4] é menor, quando o número da amostra é igual a 120, isto é, quando  $n_1 = 15$ . Obteve-se nas tabelas 1 e 2,  $X = 32$  e  $X = 50$ , que correspondem a 53,2% e 83,3% de amostras inviáveis respectivamente no total de 60 amostras. Nas tabelas 3 e 4 para  $n_1 = 15$ ,  $X = 56$  e  $X = 94$  que respectivamente são 46,6% e 78,3% de amostras inviáveis. Portanto, os dados obtidos nos levam a acreditar que  $n_a = 120$  seja mais apropriada para a realização de estudos deste tipo. Utilizando a probabilidade de cobertura proposta na metodologia, é possível decidir com maior precisão ou obter a confirmação sobre qual o número de amostras mais apropriado.

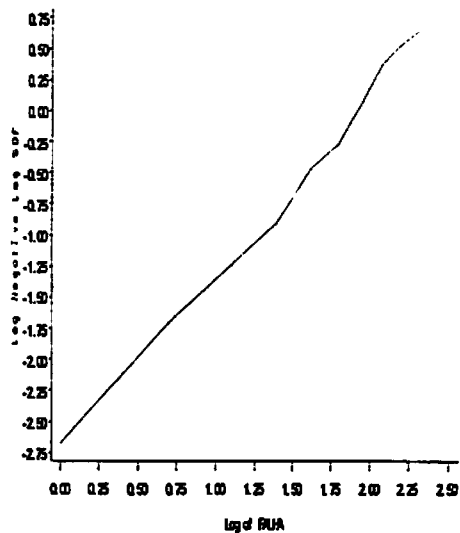
### 4.3 Adequação do modelo (dados obtidos via experimento)

Para a obtenção dos dados reais foi realizado um experimento no Laboratório de Análise de Sementes da UFLA. Neste experimento, foi avaliado, a viabilidade de semente de soja, aplicando a metodologia proposta neste trabalho. Foram avaliadas duas cultivares diferentes M.Soy e Serra Negra. O método utilizado na análise das sementes foi o teste de tetrazólio, que se baseia na análise da condição de cada semente individualmente. Cada semente foi classificada como viável ou não viável.

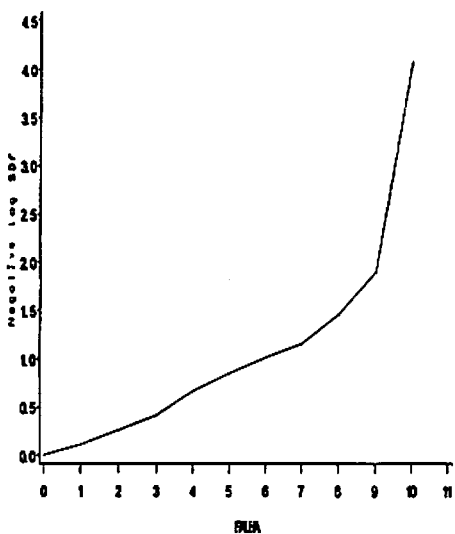
A figura 15 mostra que, pelo método dos gráficos de linearização, foi possível perceber que o tempo, até a perda da viabilidade de uma sub-amostra dos dados experimentais, segue o modelo probabilístico Weibull. Os dados utilizados para o ajuste foram: número de amostras ( $n_a$ )=60, tamanho da amostra( $n$ )=2000, probabilidade de inviabilidade( $p$ )=0,20, tamanho da sub-amostra ( $n_1$ )=10 e probabilidade de inviabilidade inicial( $p_{di}$ )=0,10 para a cultivar M.Soy e probabilidade de inviabilidade inicial( $p_{di}$ )=0,15 para a cultivar Serra Negra.



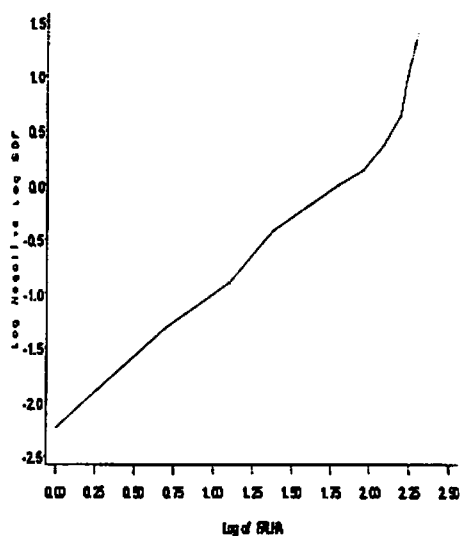
(a)



(b)



(c)



(d)

**Figura 15. Gráficos de probabilidade normal(a) e weibull(b) para Cultivar M.Soy , Gráficos de probabilidade normal(c) e weibull(d) para Cultivar Serra Negra**

Na figura 15 pode-se observar que os dados reais também se ajustam melhor a uma distribuição Weibull, ficando evidente que a cultivar M.Soy, com probabilidade de inviabilidade inicial 0,10, possui melhor ajuste que a cultivar Serra Negra com  $pdi = 0,15$ .

#### 4.4 Predição Pontual e Intervalar para número futuro de sementes não viáveis utilizando dados reais

Abaixo, as tabelas contendo valores de  $\hat{\beta}$ ,  $\hat{\eta}$ ,  $\hat{q}$ ,  $x$ , predição pontual ( $\hat{Y}$ ) e predição intervalar  $[Y; \bar{Y}]$ , variando: número de amostras( $n_a$ ), tamanho da amostra( $n$ ), probabilidade de inviabilidade inicial( $pdi$ ), probabilidade de inviabilidade( $p$ ) e tamanho da sub-amostra( $n_1$ ).

Em que  $x$  é o número de sub-amostra falhas em  $[0, t_c]$  semanas e  $Y$  é o número de sub-amostras que falharão em  $[t_w, t_w]$  semanas. Considerando neste caso do experimento real,  $[0, t_c]=[0,4]$  e  $[t_w, t_w]=[4,10]$ , isto é, as 6 semanas seguintes(1 mês e meio).

TABELA 5. Valores das predições pontual ( $\hat{Y}$ ) e intervalar  $[Y; \bar{Y}]$  das cultivares M.Soy e Serra Negra para  $n_a = 60$ ,  $n = 2000$ ,  $pdi = 0.10$  e  $p = 0.20$  e  $n_a = 60$ ,  $n = 2000$ ,  $pdi = 0.15$  e  $p = 0.20$  respectivamente.

|                      | $x$ | $\hat{\beta}$ | $\hat{\eta}$ | $\hat{q}$ | $\hat{Y}$ | $[Y; \bar{Y}]$ |
|----------------------|-----|---------------|--------------|-----------|-----------|----------------|
| $n_1=10$ M.Soy       | 16  | 0.55          | 33.61        | 0.13      | 8         | [4;13]         |
| $n_1=10$ Serra Negra | 21  | 0.56          | 18           | 0.16      | 10        | [5;15]         |

#### 4.5 Determinação da Probabilidade de Cobertura do Intervalo de Predição Bilateral.

Na Tabela 6, são apresentados os valores das probabilidades de cobertura condicionais do intervalo de predição,  $[\underline{Y}; \bar{Y}]$ , para o número de sub-amostras falhas no intervalo de tempo futuro [4,16] semanas, para um nível de confiança nominal de 95%.

TABELA 6. Valores das probabilidades de cobertura dos dados simulados para  $n_a = 120$ ,  $n = 1000$ ,  $p_{di} = 0,10$  e  $p = 0,15$

| Simulação | Número de sub-amostras Falhas se [0,4] semanas | $[\underline{Y}; \bar{Y}]$ | PC[IP(95%)   x] |
|-----------|--|----------------------------|-----------------|
| 1         | 101  | [16,19]                    | 0.9888          |
| 2         | 79   | [33,40]                    | 0.9646          |
| 3         | 75   | [32,41]                    | 0.9237          |
| 4         | 80   | [31,38]                    | 0.9435          |
| 5         | 83   | [28,36]                    | 0.9803          |
| 6         | 76   | [30,40]                    | 0.9653          |
| 7         | 93   | [24,27]                    | 0.9692          |
| 8         | 83   | [31,36]                    | 0.9277          |
| 9         | 81   | [30,37]                    | 0.9456          |
| 10        | 75   | [20,32]                    | 0.9510          |
| 11        | 82   | [32,37]                    | 0.9285          |
| 12        | 84   | [30,35]                    | 0.9254          |
| 13        | 80   | [31,38]                    | 0.9446          |
| 14        | 82   | [31,37]                    | 0.9539          |
| 15        | 75   | [32,41]                    | 0.9483          |
| 16        | 78   | [34,40]                    | 0.9163          |
| 17        | 93   | [23,27]                    | 0.9895          |
| 18        | 84   | [27,34]                    | 0.9467          |
| 19        | 98   | [18,22]                    | 0.9805          |

| Simulação | Número de sub-amostras falhas se [0,4] semanas | $[Y; \tilde{Y}]$ | PC[IP(95%)   x] |
|-----------|--|------------------|-----------------|
| 20        | 80   | [29,35]          | 0.9554          |
| 21        | 75   | [32,41]          | 0.9237          |
| 22        | 74   | [29,38]          | 0.9622          |
| 23        | 75   | [20,32]          | 0.9510          |
| 24        | 84   | [27,34]          | 0.9467          |
| 25        | 77   | [33,39]          | 0.9263          |
| 26        | 81   | [30,37]          | 0.9456          |
| 27        | 75   | [22,32]          | 0.9610          |
| 28        | 76   | [30,40]          | 0.9653          |
| 29        | 81   | [28,36]          | 0.9823          |
| 30        | 78   | [33,40]          | 0.9636          |
| 31        | 84   | [27,34]          | 0.9467          |
| 32        | 100  | [15,19]          | 0.9828          |
| 33        | 76   | [32,39]          | 0.9193          |
| 34        | 93   | [23,27]          | 0.9895          |
| 35        | 80   | [29,36]          | 0.9456          |
| 36        | 101  | [16,19]          | 0.9888          |
| 37        | 80   | [29,35]          | 0.9554          |
| 38        | 83   | [33,38]          | 0.9266          |
| 39        | 90   | [21,25]          | 0.9792          |
| 40        | 81   | [26,34]          | 0.9823          |

Com os resultados das probabilidades de cobertura condicionais, apresentados na Tabela 6, a probabilidade de cobertura não condicional será:

$$E_x\{PC[IP(95\%)] | x\} = 0.9548$$

Observou-se que a probabilidade de cobertura não condicional é muito próxima do valor nominal, isto é, o intervalo de predição  $[Y; \tilde{Y}]$  para a quantidade aleatória  $Y =$  número de sub-amostras falhas no intervalo de tempo  $[t_c, t_w]$ , demonstra ser confiável.



## 5 CONCLUSÕES

- É possível conhecer a distribuição dos tempos de vida quando se trabalha com ensaios destrutivos (caso das sementes)
- O delineamento amostral, envolvendo sub-amostras falhas apresentou-se como uma alternativa viável quando não é possível realizar a medição da variável em estudo, nas próprias unidades amostrais.
- A predição intervalar do número de sub-amostras falhas no intervalo de tempo  $[t_0, t_w]$  mostrou-se confiável, conforme demonstra a análise da probabilidade de cobertura.
- No estudo via simulação, encontrou-se que a Distribuição Weibull pode ser uma boa alternativa para ajustá-la à variável tempo até a perda da viabilidade das sementes do lote.
- No estudo com dados reais, ao comparar os valores das estimativas pontual e intervalar do número de sub-amostras de sementes falhas, estas se mostraram aceitáveis.
- A predição do número futuro de sub-amostras de sementes falhas permitirá a redução dos riscos de utilização de lotes com baixa qualidade fisiológica.

## REFERÊNCIAS BIBLIOGRÁFICAS

ELLIS, R, H. The Meaning of Viability. In: DICRIE, J.B. LININGTON, S.H. WILLIAM, J.T. (Eds). Seed management technique for genebanks, Rome: International Boards for Plant Genetic Resources. 1984. P. 146-178.

ESCOBAR, L. A; MEEKER, W, Q. Statistical Prediction Based on Censored Life Data. *Technometrics*, v.41, p.113-124. 1999.

FRANÇA NETO, J, B.; KRZYZANOWSKI, F, C.; COSTA, N, P. da. O Teste de Tetrázólio em Sementes de Soja. Londrina: EMBRAPA-CNPSO, 1998. 72p. (EMBRAPA-CNPSO. Documentos, 116).

KERNIGHAN, B,W.; RITCHIE, D, M. *The C Programming Language*. 2. Ed. 1988.

MARCO FILHO, J; W.F. e CÍCERO, S.M. Teste se Tetrázólio. Departamento de Agricultura e Horticultura, ESALQ/USP, 40p. 1987.

NELSON, W. Statistical Methods for the Ratio of Two Multinomial Proportions. *The American Statistician*, v.26, p.22-27. 1972

NELSON, W. ;HAHN, G, J A survey of Prediction Intervals and Their Applications. *Journal of Quality Technology*, v.5, p.178-188. 1973

NELSON, W. Weibull Analysis of Reliability Data with Few or No Failures. *Journal of Quality Technology*, v.17, p.140-146. 1985.

**NELSON, W. Weibull Prediction of a Future Number of Failures. *Quality and Reliability Engineering International*, v.16, p.23-26. 2000.**

**NORDMAN, D, J; MEEKER, W, Q. Weibull Prediction Intervals for a Future Number of Failures. *Technometrics*, v.44, p.15-23. 2002.**

**SAS® INSTITUTE. SAS Procedures guide for Computer 4. Ed. Cary, NC, 1990 V.2, 997,1027 p.**

**WILKS, S, S. *Mathematical Statistics*. New York: Wiley & Sons, 1962. p 429.**

## ANEXOS

|                  | Página   |
|------------------|--|
| <b>Anexo A</b>   |  |
| <b>Tabela 1A</b> | <b>Dados do número de sementes viáveis da cultivar M.Soy, obtidos durante 10 semanas no Laboratório de Análise de Sementes da UFLA, segundo o procedimento de sub-amostragem proposto na seção(3.1).....58</b>       |
| <b>Tabela 2A</b> | <b>Dados do número de sementes viáveis da cultivar Serra Negra, obtidos durante 10 semanas no Laboratório de Análise de Sementes da UFLA, segundo o procedimento de sub-amostragem proposto na seção(3.1).....59</b> |

**Tabela 1A** Dados do número de sementes viáveis da cultivar M.Soy, obtidos durante 10 semanas no Laboratório de Análise de Sementes da UFLA, segundo o procedimento de sub-amostragem proposto na seção (3.1)

| Amostra/ Semana | 1  | 2  | 3  | 4  | 5  | 6  | 7  | 8  | 9  | 10 |
|-----------------|----|----|----|----|----|----|----|----|----|----|
| Amostra 1       | 9  | 8  | 9  | 9  | 8  | 8  | 8  | 8  | 9  | 8  |
| Amostra 2       | 9  | 9  | 8  | 9  | 9  | 9  | 8  | 7  | 8  | 10 |
| Amostra 3       | 8  | 9  | 10 | 8  | 9  | 7  | 7  | 7  | 10 | 8  |
| Amostra 4       | 10 | 8  | 7  | 8  | 10 | 9  | 10 | 9  | 8  | 9  |
| Amostra 5       | 8  | 9  | 9  | 8  | 9  | 6  | 10 | 8  | 9  | 9  |
| Amostra 6       | 9  | 10 | 9  | 9  | 9  | 10 | 7  | 8  | 10 | 10 |
| Amostra 7       | 9  | 10 | 8  | 8  | 8  | 10 | 7  | 9  | 9  | 7  |
| Amostra 8       | 9  | 9  | 8  | 9  | 8  | 10 | 9  | 6  | 9  | 6  |
| Amostra 9       | 8  | 8  | 6  | 10 | 9  | 9  | 9  | 9  | 7  | 9  |
| Amostra 10      | 10 | 9  | 10 | 10 | 7  | 7  | 9  | 6  | 9  | 9  |
| Amostra 11      | 10 | 9  | 9  | 10 | 6  | 7  | 10 | 10 | 8  | 10 |
| Amostra 12      | 9  | 8  | 9  | 8  | 7  | 10 | 10 | 9  | 9  | 10 |
| Amostra 13      | 9  | 10 | 9  | 9  | 9  | 7  | 9  | 9  | 9  | 8  |
| Amostra 14      | 10 | 10 | 10 | 9  | 9  | 9  | 6  | 10 | 9  | 8  |
| Amostra 15      | 8  | 10 | 9  | 8  | 7  | 9  | 10 | 8  | 10 | 9  |
| Amostra 16      | 8  | 9  | 8  | 8  | 10 | 10 | 6  | 7  | 10 | 8  |
| Amostra 17      | 8  | 6  | 9  | 9  | 7  | 8  | 9  | 9  | 10 | 9  |
| Amostra 18      | 10 | 9  | 9  | 9  | 9  | 6  | 8  | 9  | 10 | 9  |
| Amostra 19      | 10 | 8  | 10 | 9  | 8  | 8  | 8  | 7  | 9  | 8  |
| Amostra 20      | 9  | 9  | 9  | 10 | 7  | 8  | 7  | 7  | 9  | 8  |
| Amostra 21      | 10 | 9  | 9  | 5  | 9  | 9  | 10 | 8  | 8  | 10 |
| Amostra 22      | 9  | 8  | 10 | 10 | 9  | 9  | 7  | 9  | 8  | 9  |
| Amostra 23      | 10 | 7  | 10 | 9  | 7  | 10 | 10 | 9  | 8  | 10 |
| Amostra 24      | 8  | 9  | 7  | 9  | 9  | 10 | 8  | 10 | 10 | 10 |
| Amostra 25      | 10 | 10 | 10 | 8  | 7  | 9  | 7  | 9  | 10 | 10 |
| Amostra 26      | 7  | 9  | 10 | 8  | 10 | 10 | 8  | 10 | 10 | 9  |
| Amostra 27      | 6  | 9  | 9  | 9  | 10 | 10 | 8  | 9  | 9  | 9  |
| Amostra 28      | 9  | 9  | 8  | 7  | 9  | 9  | 10 | 9  | 8  | 8  |
| Amostra 29      | 10 | 7  | 8  | 9  | 9  | 9  | 10 | 10 | 10 | 8  |
| Amostra 30      | 9  | 9  | 6  | 9  | 8  | 8  | 9  | 10 | 10 | 10 |
| Amostra 31      | 7  | 8  | 9  | 10 | 9  | 8  | 9  | 9  | 7  | 7  |
| Amostra 32      | 10 | 9  | 9  | 10 | 9  | 6  | 10 | 9  | 7  | 9  |
| Amostra 33      | 10 | 8  | 6  | 10 | 10 | 10 | 10 | 7  | 10 | 9  |
| Amostra 34      | 9  | 9  | 10 | 7  | 9  | 10 | 10 | 9  | 10 | 8  |
| Amostra 35      | 5  | 10 | 10 | 9  | 10 | 10 | 9  | 8  | 8  | 9  |
| Amostra 36      | 9  | 9  | 8  | 8  | 10 | 9  | 9  | 7  | 8  | 10 |
| Amostra 37      | 10 | 9  | 9  | 7  | 9  | 9  | 8  | 9  | 10 | 10 |
| Amostra 38      | 9  | 9  | 9  | 10 | 9  | 8  | 8  | 10 | 9  | 10 |
| Amostra 39      | 10 | 8  | 9  | 9  | 7  | 9  | 8  | 10 | 9  | 10 |
| Amostra 40      | 9  | 8  | 8  | 9  | 6  | 8  | 9  | 9  | 10 | 7  |
| Amostra 41      | 10 | 7  | 10 | 9  | 9  | 9  | 10 | 8  | 10 | 9  |
| Amostra 42      | 9  | 8  | 10 | 6  | 9  | 10 | 10 | 8  | 10 | 8  |
| Amostra 43      | 10 | 8  | 6  | 7  | 8  | 9  | 9  | 9  | 7  | 9  |

“...continuação...”

| Amostra/ Semana | 1  | 2  | 3  | 4  | 5  | 6  | 7  | 8  | 9  | 10 |
|-----------------|----|----|----|----|----|----|----|----|----|----|
| Amostra 44      | 9  | 10 | 9  | 8  | 8  | 10 | 7  | 7  | 7  | 9  |
| Amostra 45      | 9  | 10 | 9  | 9  | 9  | 10 | 10 | 9  | 6  | 7  |
| Amostra 46      | 9  | 9  | 10 | 9  | 9  | 9  | 9  | 9  | 7  | 9  |
| Amostra 47      | 9  | 10 | 9  | 9  | 9  | 9  | 7  | 6  | 6  | 10 |
| Amostra 48      | 10 | 8  | 9  | 9  | 10 | 9  | 9  | 10 | 8  | 8  |
| Amostra 49      | 10 | 9  | 10 | 10 | 10 | 8  | 9  | 10 | 9  | 10 |
| Amostra 50      | 10 | 9  | 9  | 10 | 8  | 10 | 10 | 10 | 8  | 9  |
| Amostra 51      | 9  | 9  | 10 | 10 | 9  | 10 | 9  | 6  | 7  | 10 |
| Amostra 52      | 8  | 8  | 10 | 9  | 9  | 9  | 9  | 10 | 8  | 10 |
| Amostra 53      | 10 | 10 | 8  | 9  | 10 | 9  | 9  | 8  | 9  | 8  |
| Amostra 54      | 10 | 10 | 8  | 8  | 10 | 10 | 10 | 8  | 9  | 9  |
| Amostra 55      | 9  | 8  | 8  | 10 | 8  | 9  | 10 | 8  | 7  | 7  |
| Amostra 56      | 9  | 10 | 10 | 10 | 8  | 9  | 10 | 10 | 10 | 7  |
| Amostra 57      | 9  | 10 | 9  | 10 | 10 | 10 | 9  | 10 | 9  | 7  |
| Amostra 58      | 9  | 8  | 9  | 8  | 9  | 8  | 9  | 9  | 8  | 8  |
| Amostra 59      | 9  | 9  | 9  | 8  | 9  | 8  | 8  | 9  | 8  | 9  |
| Amostra 60      | 10 | 9  | 8  | 8  | 9  | 8  | 8  | 7  | 8  | 9  |

**Tabela 2A** Dados do número de sementes viáveis da cultivar Serra Negra, obtidos durante 10 semanas no Laboratório de Análise de Sementes da UFLA, segundo o procedimento de sub-amostragem proposto na seção(3.1)

| Amostra/Semana | 1  | 2  | 3  | 4  | 5  | 6  | 7  | 8  | 9  | 10 |
|----------------|----|----|----|----|----|----|----|----|----|----|
| Amostra 1      | 8  | 10 | 9  | 6  | 10 | 9  | 10 | 8  | 8  | 10 |
| Amostra 2      | 8  | 8  | 8  | 9  | 9  | 8  | 9  | 8  | 9  | 6  |
| Amostra 3      | 8  | 7  | 7  | 8  | 8  | 8  | 9  | 9  | 9  | 8  |
| Amostra 4      | 7  | 9  | 8  | 8  | 10 | 9  | 8  | 10 | 10 | 9  |
| Amostra 5      | 8  | 5  | 7  | 9  | 9  | 10 | 8  | 10 | 8  | 10 |
| Amostra 6      | 8  | 9  | 6  | 7  | 7  | 8  | 9  | 9  | 8  | 8  |
| Amostra 7      | 8  | 10 | 9  | 10 | 7  | 8  | 7  | 9  | 9  | 9  |
| Amostra 8      | 8  | 8  | 8  | 10 | 8  | 5  | 10 | 8  | 10 | 8  |
| Amostra 9      | 9  | 10 | 8  | 8  | 8  | 10 | 7  | 7  | 8  | 9  |
| Amostra 10     | 10 | 9  | 8  | 9  | 6  | 7  | 10 | 10 | 10 | 9  |
| Amostra 11     | 8  | 8  | 9  | 7  | 6  | 10 | 10 | 8  | 9  | 9  |
| Amostra 12     | 10 | 8  | 9  | 9  | 10 | 10 | 9  | 10 | 8  | 10 |
| Amostra 13     | 7  | 9  | 7  | 8  | 7  | 8  | 7  | 9  | 8  | 9  |
| Amostra 14     | 10 | 9  | 9  | 8  | 10 | 9  | 9  | 9  | 8  | 7  |
| Amostra 15     | 8  | 10 | 8  | 9  | 9  | 9  | 8  | 8  | 7  | 8  |
| Amostra 16     | 8  | 10 | 8  | 9  | 7  | 6  | 8  | 9  | 10 | 8  |
| Amostra 17     | 8  | 9  | 8  | 8  | 9  | 8  | 10 | 9  | 10 | 9  |
| Amostra 18     | 10 | 9  | 9  | 9  | 8  | 7  | 6  | 10 | 6  | 10 |
| Amostra 19     | 9  | 8  | 6  | 6  | 8  | 8  | 10 | 10 | 9  | 10 |
| Amostra 20     | 10 | 10 | 9  | 10 | 10 | 10 | 9  | 9  | 9  | 7  |
| Amostra 21     | 9  | 8  | 8  | 7  | 8  | 10 | 9  | 7  | 6  | 8  |
| Amostra 22     | 8  | 9  | 10 | 9  | 8  | 7  | 8  | 8  | 9  | 9  |

“...continuação...”

| Amostra/ Semana | 1  | 2  | 3  | 4  | 5  | 6  | 7  | 8  | 9  | 10 |
|-----------------|----|----|----|----|----|----|----|----|----|----|
| Amostra 23      | 8  | 10 | 9  | 9  | 9  | 10 | 8  | 8  | 7  | 10 |
| Amostra 24      | 10 | 9  | 9  | 8  | 9  | 9  | 7  | 9  | 10 | 10 |
| Amostra 25      | 8  | 9  | 8  | 9  | 8  | 9  | 9  | 7  | 7  | 9  |
| Amostra 26      | 9  | 10 | 8  | 9  | 7  | 8  | 9  | 10 | 8  | 9  |
| Amostra 27      | 6  | 8  | 7  | 7  | 10 | 10 | 10 | 6  | 8  | 8  |
| Amostra 28      | 9  | 9  | 8  | 10 | 9  | 9  | 10 | 7  | 9  | 9  |
| Amostra 29      | 8  | 6  | 7  | 6  | 8  | 9  | 9  | 8  | 9  | 9  |
| Amostra 30      | 8  | 9  | 9  | 9  | 8  | 9  | 9  | 8  | 7  | 10 |
| Amostra 31      | 10 | 8  | 9  | 8  | 10 | 8  | 8  | 9  | 9  | 10 |
| Amostra 32      | 8  | 8  | 10 | 9  | 10 | 8  | 9  | 9  | 9  | 7  |
| Amostra 33      | 9  | 10 | 9  | 9  | 10 | 8  | 9  | 9  | 8  | 6  |
| Amostra 34      | 8  | 10 | 9  | 10 | 9  | 9  | 8  | 8  | 7  | 7  |
| Amostra 35      | 8  | 9  | 7  | 8  | 9  | 10 | 8  | 10 | 8  | 8  |
| Amostra 36      | 9  | 8  | 9  | 9  | 7  | 8  | 8  | 9  | 10 | 10 |
| Amostra 37      | 8  | 8  | 8  | 9  | 8  | 7  | 10 | 8  | 9  | 10 |
| Amostra 38      | 8  | 9  | 8  | 6  | 8  | 8  | 10 | 7  | 6  | 8  |
| Amostra 39      | 10 | 10 | 8  | 10 | 9  | 9  | 9  | 10 | 10 | 7  |
| Amostra 40      | 10 | 9  | 9  | 9  | 9  | 9  | 9  | 9  | 9  | 8  |
| Amostra 41      | 9  | 10 | 9  | 9  | 10 | 9  | 8  | 6  | 8  | 8  |
| Amostra 42      | 8  | 8  | 10 | 8  | 9  | 10 | 8  | 8  | 6  | 10 |
| Amostra 43      | 10 | 8  | 8  | 9  | 8  | 10 | 8  | 9  | 9  | 10 |
| Amostra 44      | 9  | 10 | 8  | 9  | 6  | 9  | 9  | 8  | 9  | 9  |
| Amostra 45      | 8  | 7  | 8  | 10 | 10 | 9  | 8  | 8  | 8  | 9  |
| Amostra 46      | 9  | 9  | 9  | 10 | 9  | 8  | 9  | 7  | 10 | 10 |
| Amostra 47      | 9  | 8  | 9  | 7  | 9  | 8  | 9  | 10 | 9  | 8  |
| Amostra 48      | 10 | 9  | 10 | 9  | 8  | 8  | 9  | 9  | 10 | 6  |
| Amostra 49      | 6  | 8  | 8  | 8  | 9  | 9  | 10 | 8  | 10 | 9  |
| Amostra 50      | 8  | 8  | 9  | 8  | 8  | 6  | 6  | 10 | 8  | 9  |
| Amostra 51      | 8  | 8  | 9  | 7  | 10 | 10 | 10 | 8  | 9  | 9  |
| Amostra 52      | 8  | 8  | 10 | 9  | 9  | 9  | 7  | 6  | 9  | 8  |
| Amostra 53      | 10 | 10 | 9  | 9  | 9  | 8  | 8  | 8  | 9  | 10 |
| Amostra 54      | 9  | 9  | 9  | 10 | 7  | 6  | 9  | 7  | 8  | 8  |
| Amostra 55      | 8  | 9  | 10 | 10 | 10 | 9  | 10 | 7  | 10 | 9  |
| Amostra 56      | 9  | 8  | 10 | 7  | 9  | 8  | 8  | 6  | 9  | 10 |
| Amostra 57      | 7  | 7  | 6  | 8  | 8  | 7  | 9  | 10 | 8  | 7  |
| Amostra 58      | 8  | 6  | 9  | 9  | 9  | 9  | 9  | 10 | 8  | 8  |
| Amostra 59      | 8  | 8  | 9  | 9  | 9  | 9  | 10 | 9  | 9  | 7  |
| Amostra 60      | 9  | 6  | 8  | 10 | 9  | 9  | 10 | 8  | 10 | 7  |