

**INFERÊNCIA BASEADA EM INTERVALOS  
ESTATÍSTICOS APLICADA À SEGURANÇA  
PÚBLICA**

**VERÔNICA YUMI KATAOKA**

**2005**

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

Kataoka, Verônica Yumi

Inferência baseada em intervalos estatísticos aplicada à Segurança  
Pública / Verônica Yumi Kataoka. -- Lavras : UFLA, 2005.

141 p. : il.

Orientador: Marcelo Silva de Oliveira

Dissertação (Mestrado) – UFLA.

Bibliografia.

1. Intervalo estatístico. 2. Segurança pública. 3. Gestão da qualidade. I.  
Universidade Federal de Lavras. II. Título.

CDD-363.1  
-519.54

**VERÔNICA YUMI KATAOKA**

**INFERÊNCIA BASEADA EM INTERVALOS ESTATÍSTICOS  
APLICADA À SEGURANÇA PÚBLICA**

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. Dr. Marcelo Silva de Oliveira

LAVRAS  
MINAS GERAIS - BRASIL  
2005

**VERÔNICA YUMI KATAOKA**

**INFERÊNCIA BASEADA EM INTERVALOS ESTATÍSTICOS  
APLICADA À SEGURANÇA PÚBLICA**

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".

APROVADA em 22 de dezembro de 2005

Prof. Dr. Joel Augusto Muniz UFLA

Prof. Dr. João Domingos Scalon UFLA

Prof. Dr<sup>a</sup> Thelma Sáfadi UFLA

Prof. Dr. Marcelo Silva de Oliveira

UFLA

(Orientador)

LAVRAS

MINAS GERAIS - BRASIL

Ao meu esposo, Rubens, e aos meus filhos, Alexander e Felipe,  
pelo amor e compreensão,

**OFEREÇO**

Aos meus pais e irmãos que, mesmo à distância, sempre deram forças para que eu  
continuasse caminhando e superando todas as dificuldades,

**DEDICO.**

## AGRADECIMENTOS

Ao meu esposo, Rubens, e aos meus filhos, Alexander e Felipe, que souberam ter paciência para compreender as minhas freqüentes ausências e por terem sonhado comigo a conquista deste título.

Aos meus pais, Zilda e Osamu, que souberam conduzir com muita sabedoria a minha formação como cidadã e sempre deram prioridade à educação.

Aos meus irmãos Patrícia, Flávio, Elisabete, Sheila e Frederico, fiéis companheiros em todos os momentos da minha vida.

Ao Prof. Marcelo, pela paciência, amizade e atenção durante esses dois anos, tendo um papel decisivo na condução dos trabalhos.

Aos meus grandes amigos do Mestrado, Charles, Claudiney, Elias, Eustáquio, Josiane, Lívia, Nádia, Paulinho, Rejane e Vanessa, que foram decisivos para que eu tivesse forças para concluir esta etapa. Agradeço sempre a Deus por ter colocado no meu caminho essas pessoas, pois acabamos formando uma grande família. Espero que a nossa amizade, mesmo que alguns estejam longe, seja eterna.

À Graziela (também da minha turma), tinha que dedicar um parágrafo especial, porque é assim que ela entrou na minha vida. Nosso convívio não se restringiu somente a “infinitas” horas de estudo, foi muito além. Portanto, com certeza, ela foi uma grande amiga e incentivadora durante esse processo, e espero que a nossa amizade perdure por muitos e muitos anos.

A todos os meus amigos de Petrolina (PE) e Juazeiro (BA), Marco e Ana, Rose e Betânia, que, mesmo de longe, com certeza torceram pelo meu sucesso os quais sempre trarei perto no meu coração.

Aos amigos que fiz nesta cidade, Francisca, Dalva, Dona Itinha, Adilson e Luciane, que me deram apoio em todos os momentos e já são importantes na minha vida.

Aos demais colegas de Mestrado e Doutorado, pelas palavras de incentivo e apoio nas dificuldades.

Aos professores e técnico-administrativos do DEX, pela competência, amizade e companheirismo.

À 6ª Região do Comando da Polícia Militar e ao 8º Batalhão, por disponibilizarem os dados e pelo apoio, dando oportunidade de desenvolver um trabalho com um assunto de interesse da sociedade.

À CAPES, pela concessão da bolsa de estudos.

Por fim, a Deus, que na verdade é o princípio, pois sem ele nada do que foi dito anteriormente teria sido possível.

## SUMÁRIO

RESUMO.....	i
ABSTRACT.....	ii
1 INTRODUÇÃO.....	1
2 REFERENCIAL TEÓRICO.....	3
2.1 Violência e criminalidade.....	3
2.1.1 Panorama geral.....	3
2.1.2 Tipos de crimes violentos.....	6
2.1.3 Comensurabilidade dos dados da Polícia Militar de Minas Gerais (PMMG) no período de 1991 a 2004.....	7
2.1.3.1 Mudanças em termos doutrinários e operacionais.....	7
2.1.3.2 Mudanças em termos jurídicos.....	9
2.1.4 Distribuição de Poisson: número de crimes violentos versus índice de criminalidade violenta.....	10
2.2 Gestão da Qualidade na Segurança Pública .....	16
2.2.1 Panorama nacional .....	16
2.2.2 Panorama mineiro.....	19
2.3 Intervalos estatísticos.....	24
2.3.1 Generalidades.....	24
2.3.2 Intervalos de confiança.....	27
2.3.2.1 Definição .....	27
2.3.2.2 Interpretação.....	27
2.3.2.3 Exemplo.....	28
2.3.3 Intervalos de predição .....	30
2.3.3.1. Definição.....	30
2.3.3.2. Interpretação.....	31

2.3.4 Intervalos de tolerância.....	31
2.3.4.1 Definição.....	31
2.3.4.2 Interpretação.....	33
2.4 Intervalos estatísticos para uma variável aleatória que obedece ao modelo $Y=\mu+\varepsilon$ com distribuição Normal (Modelo Simples Normal).....	34
2.4.1 Distribuição Normal.....	34
2.4.2 Modelo $Y=\mu+\varepsilon$ com distribuição Normal.....	35
2.4.3 Intervalo de confiança para a média populacional ( $\mu$ ) com variância $\sigma^2$ desconhecida.....	36
2.4.4 Intervalo de predição para a média de $m$ futuras observações ( $\bar{X}_m$ ).....	37
2.4.5 Intervalo de predição para uma observação futura.....	37
2.4.6 Intervalo de tolerância para conter pelo menos $\gamma\%$ da população com um nível de confiança de $100(1-\alpha)\%$ .....	38
2.4.7 Comparação ente os intervalos no modelo simples Normal.....	40
2.5 Regressão linear simples.....	42
2.5.1 Modelo de regressão linear simples (MRLS).....	42
2.5.2 Estimação de $\beta_0$ e $\beta_1$ empregando o método dos quadrados mínimos.....	43
2.5.3 Diagnóstico das pressuposições básicas.....	45
2.5.4 Intervalos estatísticos no contexto da regressão.....	49
2.5.4.1 Estimação da variância de $\hat{Y}_i$ .....	49
2.5.4.2 Intervalo de confiança bilateral para a média ( $E(Y_n)$ ).....	51
2.5.4.3 Intervalo de predição bilateral para uma observação futura.....	52

2.5.4.4	Intervalo de tolerância unilateral para conter uma proporção $\gamma$ da população.....	53
2.5.4.5	Intervalo de tolerância bilateral para conter pelo menos uma proporção $\gamma$ da população.....	54
3	MATERIAL E MÉTODOS.....	56
3.1	Dados.....	56
3.2	Estudo preliminar.....	57
3.3	Comportamento do número de crimes violentos contra a pessoa e contra o patrimônio no período de 1991 a 2004 em Lavras – MG.....	58
3.3.1	Testes de Aderência $\chi^2$ para distribuição uniforme discreta.....	58
3.3.2	Regressão linear simples.....	59
3.3.2.1	Modelo estatístico.....	59
3.3.2.2	Diagnóstico das pressuposições básicas.....	60
3.3.2.3	Intervalos estatísticos no contexto da regressão.....	60
3.3.3	Distribuição de Poisson.....	62
3.4	Inferência sobre a amostra de uma população Normal.....	63
3.5	Software estatístico utilizado.....	64
4	RESULTADOS E DISCUSSÃO.....	65
4.1	Estudo preliminar: número de habitantes versus número de crimes violentos .....	65
4.2	Comportamento do número de crimes violentos contra a pessoa e contra o patrimônio no período de 1991 a 2004 em Lavras – MG.....	67
4.2.1	Testes de aderência $\chi^2$ .....	67
4.2.2	Regressão linear simples.....	71
4.2.2.1	Modelo estatístico e diagnóstico das pressuposições básicas.....	71
4.2.2.2	Intervalos estatísticos no contexto da regressão.....	80
4.2.2.2.1	Construção.....	80

4.2.2.2.2 Interpretação.....	83
4.2.2.2.3 Resultados parciais do ano de 2005.....	86
4.2.3 Distribuição de Poisson: número de crimes violentos.....	87
4.3 Comportamento do número de crimes violentos contra a pessoa e contra o patrimônio no período de 2000 a 2004.....	89
4.3.1 Regressão linear simples.....	89
4.3.2 Intervalos estatísticos no contexto da regressão.....	93
4.3.2.1 Construção .....	93
4.3.2.2 Comparação dos intervalos projetados para o ano de 2005 (h=6) com os resultados parciais.....	94
4.4 Inferência sobre a amostra de uma população Normal.....	96
5 CONCLUSÕES.....	100
5.1 Conclusões sobre a análise dos dados.....	100
5.2 Conclusões sobre metodologias estatísticas.....	101
6 REFERÊNCIAS BIBLIOGRÁFICAS.....	103
ANEXOS.....	112

## RESUMO

KATAOKA, Verônica Yumi. **Inferência baseada em intervalos estatísticos aplicada à Segurança Pública**. LAVRAS: UFLA, 2005. 141p. (Dissertação - Mestrado em Agronomia/Estatística e Experimentação Agropecuária)\*

Um dos maiores problemas que afligem os cidadãos e preocupam a opinião pública em todo o mundo é a criminalidade, principalmente porque vem quase sempre acompanhada de muita violência, tendo efeitos econômicos, sociais e políticos. Em contrapartida, os órgãos de Segurança Pública vêm investindo em ações preventivas, através de planejamentos estratégicos mais eficientes, baseados em transformação de dados brutos em informações significativas. Sendo assim, o presente estudo tem por finalidade propiciar uma construção metodológica e uma análise científica no banco de dados da Polícia Militar de Minas Gerais, com ênfase no estudo da criminalidade violenta do município de Lavras – MG (período de 1991 a 2004), por meio da utilização de inferência baseada em intervalos estatísticos. Os crimes violentos foram divididos em: contra a pessoa (CVPES) e contra o patrimônio (CVPT), sendo transformados também em índice de criminalidade violenta, ICPVES e ICPVT, respectivamente. As análises foram feitas considerando esses quatro tipos de dados no contexto da regressão linear simples, sendo construídos intervalos de Confiança, Predição, Tolerância Unilateral e Bilateral, em dois períodos: de 1991 a 2004 (I) e de 2000 a 2004 (II). Os limites dos intervalos estatísticos para o ano de 2005 foram comparados com os dados reais (até outubro de 2005). Além desse procedimento, foi retirada uma amostra aleatória de 30 municípios mineiros, no ano de 2003, e com faixa populacional de 14000 a 17500 habitantes, sendo avaliado o ICV considerando um modelo simples Normal, com a construção também de alguns intervalos estatísticos. Para o período I, por exemplo, verificou-se que o ICPVES aumenta na ordem de 0,021 ao ano, a partir de 1991 ( $k=1$ ), já para o período II, o aumento é da ordem de 0,0118 ao ano. Verificou-se, também, com 95% de confiança, que para o período I o intervalo de tolerância para conter o ICPVT de pelo menos 95% dos municípios com mesmo porte populacional de Lavras no 1º semestre do ano de 2005 estará entre 0,27 e 0,72. Já para o período II, estará entre 0,068 e 0,969. Em relação à amostra dos 30 municípios mineiros, pode-se observar que, com 95% de confiança, o intervalo de tolerância para conter o ICV de pelo menos 95% dos municípios com mesmo porte populacional no ano de 2003 era de 0 a 2,251. Uma das conclusões a que se chegou é que os resultados dos intervalos estatísticos construídos no contexto do modelo simples normal e da regressão são ferramentas úteis para auxiliar a Gestão da Qualidade na Segurança Pública, no sentido de que possibilitam a projeção de valores futuros, que podem auxiliar um planejamento estratégico mais eficiente, a semelhança das cartas - controle no Controle Estatístico de Processo (CEP). Além disso, a utilidade dessa metodologia verifica-se até mesmo quando “falha” no sentido de que uma observação fora dos limites de confiança, predição e tolerância ainda servem para os propósitos da gestão.

---

\* Orientador: Prof Dr. Marcelo Silva de Oliveira - UFLA.

## ABSTRACT

KATAOKA, Verônica Yumi. Inference based on statistical intervals applied to Public Security. Lavras: UFLA, 2005. 141p. (Dissertation - Master in Agronomy/Statistics and Agricultural Experimentation) \*

One of the greatest problems that afflict citizens and worry public opinion all over the world is the criminality, mostly because it usually comes with violence, having economic, social and political effects. As opposed, the Public Security organs invest in preventive actions and more efficient strategic planning, based on transformation of brute raw in significant information. This study purposes propitiate a methodological construction and a scientific analysis on the data bench of the Military Police of Minas Gerais, with emphasis in the violent criminality in Lavras – MG (from 1991 to 2004), by inference based on statistical intervals. The violent crimes were divided into: against person (CVPES) and against patrimony (CVPT), also being transformed in index of violent criminality, ICVPES and ICVPT, respectively. The analyses were done considering these four kinds of data in the context of the simple linear regression, being built Confidence intervals, Prediction intervals, Unilateral and Bilateral Tolerance intervals, for two periods: from 1991 to 2004 (I) and from 2000 to 2004 (II). The limits of the statistical intervals for year of 2005 were compared with the real data (by October 2005). Beyond of this procedure, it was drawn a random sample of 30 cities of Minas Gerais, from 2003, with population varying from 14000 to 17500 inhabitants, being evaluated the ICV considering a Normal simple model and being constructed statistical intervals. For the period I, for example, it verifies that ICVPES increases in the order of 0.021 a year, starting from 1991 ( $k=1$ ), while for the period II, the increase belongs to the order of 0.0118 a year. It verifies, as well, with 95% of confidence, that period I tolerance interval to contain ICVPT with, at least, 95% of the municipals with same population of Lavras in the 1st year's semester of 2005 will be between 0.27 and 0.72. For the period II, it will be between 0.068 and 0.969. Regarding the sample of the 30 cities of Minas Gerais, it can note that, with 95% of confidence, the tolerance interval contains ICV of at least 95% of the municipals with same population in year of 2003 was from 0 to 2.251. One of the conclusions to what arrived is that the results of the statistical intervals built in the context of the normal simple model and of the regression are useful tools to assist the Quality Management in Public Security, in the sense that they enable the projection of future values, that they can assist a more efficient strategic planning, the similarity letters control in the Statistical Control of Process (CEP). Moreover, the utility of this methodology even verifies when “failure” in the sense that an observation outside the limits confidence, prediction and tolerance still serve for the management needs.

---

\* Adviser: Prof Dr. Marcelo Silva de Oliveira - UFLA.

# 1 INTRODUÇÃO

Segundo a Organização Mundial da Saúde, anualmente, o custo da violência no Brasil corresponde a 10,5 por cento do PIB -Produto Interno Bruto- e os gastos com problemas de saúde relacionados à violência chegam a quase 2 por cento do PIB. Esses números estão no relatório divulgado em 2003 pela entidade em Viena, na abertura da Sétima Conferência Mundial para a Promoção de Segurança e Prevenção de Violência (SSPSE, 2003). Esses dados refletem a pertinência da preocupação da sociedade brasileira com os problemas relacionados à criminalidade e violência.

Diante dessa situação, poder-se-ia esperar que os órgãos de Segurança Pública estivessem preocupados mais em tomar medidas repressoras no combate à criminalidade do que em entendê-la, visto a urgência que as estatísticas demandam. Contudo, num panorama nacional, percebe-se uma mudança na concepção de gestão, traduzida por ações baseadas em planejamentos estratégicos mais eficientes, a fim de gerenciar o modo como se intervém no fenômeno do crime.

Por sua vez, os planejamentos desenvolvidos têm encontrado suporte na atuação mais intensa da Secretaria Nacional de Segurança Pública, através de vários projetos como, por exemplo, a criação do Sistema Nacional de Estatísticas e Justiça Criminal; pela participação da sociedade por meio da criação do Conselho Comunitário e da adoção da estratégia de Polícia Comunitária; bem como pela realização de diversos trabalhos científicos, que estão propiciando a transformação de dados brutos de criminalidade em informações significativas.

Devido à complexidade do comportamento da criminalidade, é interessante o desenvolvimento de metodologias científicas que possam contribuir, mesmo que paulatinamente, em planejamentos estratégicos futuros.

Portanto, o objetivo deste trabalho é propiciar uma construção metodológica e uma análise científica no banco de dados da Polícia Militar de Minas Gerais, com ênfase no estudo da criminalidade violenta do município de Lavras – MG (período de 1991 a 2004), através da utilização de inferência baseada em intervalos estatísticos.

Para cumprir essa meta, o trabalho foi subdividido em cinco seções. Na primeira, apresenta-se uma breve introdução e o objetivo do estudo. Na segunda, o referencial teórico, que trata de tópicos como Violência e Criminalidade, Gestão da Qualidade na Segurança Pública, Intervalos Estatísticos, Modelo Simples Normal e Regressão Linear Simples. Na terceira são descritos o Material e Métodos. Na quarta, os Resultados e Discussão e, na quinta, as conclusões sobre os dados e sobre as metodologias estatísticas. Além disso, tem-se o Anexo A, com figuras complementares, o Anexo B, composto de tabelas auxiliares, e o Anexo C, com todas as rotinas desenvolvidas no software estatístico R<sup>®</sup> v.2.1.1 (R, 2005).

## **2 REFERENCIAL TEÓRICO**

### **2.1 Violência e criminalidade**

#### **2.1.1 Panorama geral**

Segundo Beato (1998a), sociólogos são bastante familiarizados com duas teorias contrastantes na definição das causas da criminalidade: (i) uma delas nos diz que criminalidade e violência são fenômenos cuja origem se deve essencialmente a fatores de natureza econômica, tais como privação de oportunidades, desigualdade social e marginalização, os quais seriam estímulos decisivos para o comportamento criminoso; (ii) a outra teoria credita ao delinqüente e aos atos criminosos uma agressão ao consenso moral e normativo da sociedade; um baixo grau de integração moral produziria o fenômeno do crime. Conseqüentemente, a punição do crime é uma necessidade imperiosa para o restabelecimento dos valores centrais do núcleo normativo.

Ao contrário do que é proposta em pilhas de produção intelectual e pesquisa sistemática, a correlação a ser estabelecida para a explicação do crime não é com a pobreza, mas com a riqueza. Isso porque a prosperidade termina por ensejar um incremento nas oportunidades para a ação criminosa, na medida em que fornece alvos viáveis e compensadores, bem como dificulta os mecanismos tradicionais de controle social e vigilância (Beato, 1998a).

De acordo com Cohen & Felson, citados por Beato (1998a): “É irônico que os mesmos fatores que incrementam as oportunidades para desfrutarmos dos benefícios da vida podem igualmente incrementar a oportunidade para as violações predatórias [...] Ao invés de tomarmos os crimes predatórios

simplesmente como indicadores de colapso social, podemos concebê-los como subprodutos da liberdade e da prosperidade tal como se manifestam nas atividades rotineiras de nosso cotidiano”.

Segundo Assis, citado por Ribas & Costa (2004), a violência está muito mais associada à desigualdade social do que à miséria em si. Há países, como a Índia ou alguns da África, que são extremamente pobres, mas que são muito menos violentos. A disparidade social brasileira é fator determinante da sociedade violenta que temos.

Rocha, citado por Lôbo (2004), afirma: "Como a violência afeta mais os pobres, é usual fazer um raciocínio simplista de que a pobreza é a principal causadora da violência entre os jovens, mas isso não é verdade". "O fato de ser pobre não significa que a pessoa será violenta. Temos inúmeros exemplos de atos violentos praticados por jovens de classe média".

De um lado, jovens brancos, bem vestidos, com um bom nível de escolaridade e trabalhando com carteira assinada. De outro, jovens negros, maltrapilhos, analfabetos e trabalhando na informalidade para comprar comida. O quadro de extrema desigualdade citado no exemplo acima, tão comum no Brasil, está entre as principais causas da violência entre jovens, segundo um estudo divulgado pelo Instituto de Pesquisas Econômicas Aplicadas - IPEA (Lôbo, 2004).

De acordo com Lôbo (2004), apesar de ser um agravante das situações de violência, os números divulgados pela pesquisa do IPEA mostram que a pobreza não é preponderante para o comportamento violento, mas sim a desigualdade social.

Para Araújo & Faljnzyber (2001), a criminalidade é um problema social, econômico e político importante que deve ser enfrentado. É um problema social, pois afeta diretamente a qualidade e expectativa de vida das populações. É um problema econômico porque, de um lado, a sua intensidade está associada

às condições econômicas e, de outro, limita o potencial de desenvolvimento das nações. Finalmente, a criminalidade é um problema político, já que as ações necessárias para combater o crime envolvem a participação ativa dos governos e a alocação de recursos públicos escassos em detrimento de outros objetivos de políticas públicas.

Beato (1998b) ressalta que poucos problemas sociais mobilizam tanto a opinião pública como a criminalidade e a violência. Não é para menos. Este é um daqueles problemas que afeta toda a população, independentemente de classe, raça, credo religioso, sexo ou estado civil. São conseqüências que se refletem tanto no imaginário cotidiano das pessoas, como nas cifras extraordinárias a respeito dos custos diretos da criminalidade violenta. Receosas de serem vítimas de violência, as pessoas adotam precauções e comportamentos defensivos na forma de seguros, sistemas de segurança eletrônicos, cães de guarda, segurança privada, grades e muros altos, alarmes, etc (já se disse que o presídio tornou-se modelo de qualidade residencial no Brasil).

Outro fator importante na discussão sobre criminalidade é o fato de que o Direito Penal brasileiro não tem atuado na origem do delito e o seu poder intimidatório está muito aquém do que se alardeia, mesmo que os órgãos repressores tenham estrutura para bem desempenhar suas funções (Barandier, 2002).

Segundo Barandier (2002), num país como o Brasil, em que a distribuição de renda está entre as piores do mundo, a única via para a redução da criminalidade violenta é buscar o equilíbrio social, garantir uma vida digna às crianças e aos jovens, dar-lhes um horizonte.

### 2.1.2 Tipos de crimes violentos

Os crimes violentos são divididos em três categorias:

- Contra a pessoa: homicídio tentado e consumado (B04); seqüestro e cárcere privado (B05).
- Contra o patrimônio: roubo consumado (C05), roubo à mão armada (C09), latrocínio<sup>1</sup> (C13) e extorsão mediante seqüestro (C17).
- Contra os Costumes: estupro tentado e consumado (D04), que não foi analisado neste trabalho.

Cada um dos tipos de crimes, bem como as suas subdivisões encontra-se na Tabela 1A (AnexoA).

Segundo Beato (1998a), é importante destacar a diferença entre as taxas de homicídio e de delitos contra a propriedade: “As taxas de homicídio distribuem-se igualmente entre as cidades de todos os tamanhos, o que indica que o tamanho da população não exerce influência sobre as taxas deste tipo de crime violento. A explicação para isto é que homicídios não dependem da seleção de alvos viáveis, já que envolvem, em boa parte dos casos, situações entre pessoas conhecidas, ou discussões casuais e freqüentemente regadas a altas doses de álcool; apenas uma pequena parcela dos homicídios envolve motivações instrumentais. Os crimes de roubo e roubo à mão armada, ao contrário, exigem o anonimato e a disponibilidade de alvos que somente cidades com certo porte podem garantir”.

Se os homicídios constituem a face mais dramática da violência urbana, os delitos contra o patrimônio são os fenômenos que mais se apresentam no cotidiano de grande parte da população nesses centros. Pode-se dizer que a percepção de grande parte das populações nos grandes centros urbanos é moldada por esse tipo de incidente. Indivíduos que andam pelas ruas dos

---

<sup>1</sup> Homicídio com o objetivo de roubo, ou roubo à mão armada.

grandes centros urbanos e são assaltados talvez sejam a expressão mais comum do fenômeno da violência urbana (Beato, 2001a).

Naturalmente, a forma de atuação e de prevenção da Polícia Militar para esses dois tipos de crimes violentos é também diferente. Observando – se os dados do município de Lavras – MG, no período de 1991 a 2004, julgou-se conveniente avaliar, respectivamente, o número de crimes violentos contra a pessoa (CVPES) e o número de crimes violentos contra o patrimônio (CVPT).

### **2.1.3 Comensurabilidade dos dados da Polícia Militar de Minas Gerais (PMMG) no período de 1991 a 2004**

Neste trabalho, identificaram-se algumas mudanças em termos jurídicos, doutrinários e operacionais, que podem ter alterado a comensurabilidade dos dados da Polícia Militar no período de 1991 a 2004.

Um aumento “imprevisível” no número de crimes violentos, principalmente contra o Patrimônio, nesse período, pode ter tido como causa alguns fatos que serão relatados a seguir.

#### **2.1.3.1 Mudanças em termos doutrinários e operacionais**

(i) Criação dos Conselhos Comunitários de Segurança Pública (CONSEP)

Segundo Putnam, citado por Souza & Gouvêa (1995?), a capacidade de organização de uma comunidade é fator determinante para seu progresso. Assim, quando as pessoas passam a se relacionar com outros cidadãos, seus problemas comuns tendem a ser encarados e compreendidos de maneira mais racional. O grupo passa a acreditar em sua própria capacidade de ação e medidas concretas substituem o antigo conformismo, a apatia e o medo.

Com a criação desses conselhos, percebe-se a situação em que o “cidadão” conversa com o “cidadão” e, através dessa atuação “informal”, tem-se

como resultado um maior registro das denúncias, que, por conseguinte, aumentam os índices de criminalidade.

(ii) Adoção da estratégia de Polícia Comunitária – por um lado, já foi detectado, empiricamente, que a atuação desse tipo de polícia faz aumentar os índices de criminalidade, por permitir uma maior interação da polícia com a comunidade. Por outro lado, percebe-se que esse tipo de atuação mais intensa (no sentido de proximidade) pode provocar uma redução nos crimes ao longo do tempo, justamente porque ocorre a atuação sobre as causas. É uma mudança na filosofia do emprego operacional da Polícia.

Vale salientar, que, no caso específico de Lavras, os itens (i) e (ii) foram implementados a partir de 2002. Na seção 2.2.1 será retomada a discussão da importância dessa reforma para a Gestão da Qualidade na Segurança Pública.

(iii) A Polícia Militar vem percebendo um aumento no número dos chamados “falsos crimes”, em que o cidadão simula uma infração inexistente. Por exemplo, no caso dos roubos, os objetivos podem ser acobertar perdas ou desvios de dinheiro. A Polícia, em muitos casos, não tem controle sob essa situação, é a palavra do “ofendido”, e, no processo de investigação, nem sempre se consegue detectar esse tipo de fraude.

Marques (2002) relata que dos 30 casos de seqüestro-relâmpago denunciados no ano de 2002 em Brasília, a Polícia Civil descobriu quatro falsas comunicações, três delas com o objetivo de esconder traição amorosa.

(iv) No período de 1991 a 2004 aconteceram três mudanças na codificação das ocorrências dos crimes violentos, conforme apresentado no Quadro 1

QUADRO 1 Mudanças no código de ocorrência dos tipos de crimes violentos em três períodos, de 1991 a 1995, 1996 a 2001 e a partir de 2002

1991 a 1995	1996 a 2001	A partir de 2002
B04 – homicídio consumado	B04.001 – homicídio consumado	B04.001 – homicídio consumado
B05 – homicídio tentado	B04.002 - homicídio tentado	B04.002 – homicídio tentado
C05 – roubo	C05.000 – roubo consumado	B05.000 – seqüestro e cárcere privado
C09 – roubo a mão armada	C09.000 – roubo a mão armada consumado	C05.000 – roubo consumado
D04 – estupro	D04.001 – estupro consumado	C09.000 – roubo a mão armada consumado
	D04.002 – estupro tentado	C13.000 – latrocínio
		C17.000 – extorsão mediante seqüestro
		D04.001 – estupro consumado
		D04.002 – estupro tentado

Os dados nos períodos de 1991 -1995 e de 1996 - 2001 foram obtidos, respectivamente, no Livro de Código CISS 0044\_1 (Batitucci, 2005a) e CIS 0044\_2 (Batitucci, 2005b) e a partir de 2002 no Memorando Nr 34.598.4 /EMPM (PMMG, 2001).

Observando o Quadro 1, verifica-se que alguns crimes que eram considerados violentos passaram a não ser mais (ex.: o roubo tentado até 1995), e vice-versa (ex.: seqüestro e cárcere privado a partir de 2002), além de mudanças de código (ex.: homicídio tentado que passou de B05 para B04.001).

### 2.1.3.2 Mudanças em termos jurídicos

A lei 9.099 foi criada em 26/09/1995 e dispõe sobre os Juizados Especiais Cíveis e Criminais (Brasil, 1995).

Os Juizados Especiais Criminais<sup>2</sup>, em linhas gerais, têm competência para conciliação, processo e julgamento das causas criminais menos complexas, como está descrito no artigo 61 (Cap. III).

Art. 61. Consideram-se infrações penais de menor potencial ofensivo, para os efeitos desta Lei, as contravenções penais e os crimes a que a lei comine pena máxima não superior a um ano, excetuado os casos em que a lei preveja procedimento especial.

Segundo a percepção do comando do 8º Batalhão da Polícia Militar de Minas Gerais, essa Lei pode ter afetado indiretamente o aumento dos crimes violentos, isto é, nas categorias desse tipo de crime as penas são maiores do que 4 anos, mas o que vem ocorrendo é que, como os pequenos delitos (ex.: furtos) são “assistidos” por essa Lei, a sensação de impunidade do infrator vai aumentado, e isso pode levá-lo, no futuro, a praticar delitos mais graves.

Portanto, pelo que foi exposto nesta seção, parece que o mais adequado seria analisar o comportamento dos crimes violentos no município de Lavras – MG somente a partir de 2000.

#### **2.1.4 Distribuição de Poisson: número de crimes violentos versus índice de criminalidade violenta**

Em 1830, Siméon Denis Poisson desenvolveu uma distribuição, que recebeu o seu nome, baseada na distribuição binomial. Ela a usou para mostrar como a variação inerente em decisões do júri afetou as inferências que poderiam ser feitas sobre a “probabilidade da convicção” nas cortes francesas (Maltz, 1996).

---

<sup>2</sup> Vale salientar que esse tipo de Juizado foi instalado no município de Lavras – MG a partir de 2000.

Matematicamente falando, a distribuição de Poisson trata do caso limite da Binomial quando o número de provas  $n$  ( $n > 0$ ) tende para o infinito, a probabilidade do evento  $p$  ( $0 < p < 1$ ) em cada prova tende para zero e o produto  $np$  tende para uma constante  $\lambda$ , isto é, trata-se de uma distribuição Binomial em que:

- a)  $n \rightarrow \infty, p \rightarrow 0$
- b)  $np \rightarrow \lambda > 0$

Portanto, seja uma variável aleatória  $Y$  com valores não negativos, partindo da expressão do modelo binomial, tem-se:

$$P[Y = y] = \frac{n!}{y!(n-y)!} p^y (1-p)^{n-y} = \frac{n(n-1)\dots(n-y+1)}{y!} p^y (1-p)^{n-y} \quad (1)$$

Definindo  $\lambda = np$ , tem-se que  $p = \lambda/n$  e  $(1-p) = (n-\lambda)/n$ . Substituindo os valores em 1:

$$P[Y = y] = \frac{n(n-1)\dots(n-y+1)}{y!} \left[\frac{\lambda}{n}\right]^y \left[\frac{n-\lambda}{n}\right]^{n-y}$$

$$P[Y = y] = \frac{\lambda^y n(n-1)\dots(n-y+1)}{y! n^y} \left[1 - \frac{\lambda}{n}\right]^n \left[1 - \frac{\lambda}{n}\right]^{-y} \quad (2)$$

Observe-se que:

$$\frac{n(n-1)\dots(n-y+1)}{n^y} = \frac{(n-1)\dots[n-(y-1)]}{n^{y-1}} = \frac{(n-1)}{n} \frac{(n-2)}{n} \dots$$

$$\dots \frac{[n-(y-1)]}{n} = \left[1 - \frac{1}{n}\right] \left[1 - \frac{2}{n}\right] \dots \left[1 - \frac{y-1}{n}\right] \quad (3)$$

E, assim, substituindo (3) em (2):

$$P[Y = y] = \frac{\lambda^y}{y!} \left[1 - \frac{1}{n}\right] \left[1 - \frac{2}{n}\right] \dots \left[1 - \frac{y-1}{n}\right] \left[1 - \frac{\lambda}{n}\right]^n \left[1 - \frac{\lambda}{n}\right]^{-y} \quad (4)$$

Fazendo  $n \rightarrow \infty$ , tem-se que os termos do tipo  $\left[1 - \frac{a}{n}\right]$  tendem a 1, inclusive

$\left[1 - \frac{\lambda}{n}\right]^{-y}$ . O termo  $\left[1 - \frac{\lambda}{n}\right]^n$ , pela própria definição do número neperiano, tende a  $e^{-\lambda}$  quando  $n \rightarrow \infty$ . E, assim, a distribuição de Poisson pode ser apresentada como:

$$P[Y = y] = \frac{\lambda^y e^{-\lambda}}{y!} \quad (5)$$

Esse modelo corresponde realmente a uma distribuição de probabilidade, observando-se que:

$$\sum_{y=0}^{\infty} P[Y = y] = \sum_{y=0}^{\infty} \frac{\lambda^y e^{-\lambda}}{y!} = e^{-\lambda} e^{\lambda} = 1 \quad (6)$$

A esperança de Y é dada por:

$$E[Y] = \sum_{y=0}^{\infty} y \frac{\lambda^y e^{-\lambda}}{y!} = \sum_{y=1}^{\infty} \frac{\lambda^y e^{-\lambda}}{(y-1)!} \quad (7)$$

Definindo  $t=y-1$  e substituindo em 7, tem-se:

$$E[Y] = \sum_{t=0}^{\infty} \frac{\lambda^{t+1} e^{-\lambda}}{t!} = \lambda \sum_{t=0}^{\infty} \frac{\lambda^t e^{-\lambda}}{t!} = \lambda \cdot 1 = \lambda \quad (8)$$

observando-se que  $t$  tem distribuição de Poisson. Portanto, o parâmetro  $\lambda$  corresponde ao valor médio da variável aleatória. Quanto à variância de  $Y$ , tem-se que:

$$E[Y^2] = \sum_{y=0}^{\infty} y^2 \frac{\lambda^y e^{-\lambda}}{y!} = \sum_{y=0}^{\infty} y \frac{\lambda^y e^{-\lambda}}{(y-1)!} \quad (9)$$

Novamente definindo  $t=y-1$  e substituindo em 9, tem-se:

$$E[Y^2] = \sum_{y=0}^{\infty} (t+1) \frac{\lambda^{t+1} e^{-\lambda}}{t!} = \lambda \sum_{y=0}^{\infty} t \frac{\lambda^t e^{-\lambda}}{t!} + \lambda \sum_{y=0}^{\infty} \frac{\lambda^t e^{-\lambda}}{t!} = \lambda^2 + \lambda \quad (10)$$

E, assim:

$$V(Y) = E(Y)^2 - [E(Y)]^2 = \lambda^2 + \lambda - \lambda^2 = \lambda \quad (11)$$

A distribuição de Poisson fornece um modelo realístico para muitos fenômenos aleatórios. Desde que os valores da variável aleatória de Poisson sejam inteiros não negativos, então qualquer fenômeno aleatório que envolva algum tipo de contagem terá a distribuição de Poisson como um modelo possível, pelo menos em princípio. (Mood et al., 1974).

Estudos realizados na área de criminalidade já vêm relacionando alguns eventos a uma distribuição de Poisson:

- Greene & Stollmack (1981), citado por Maltz (1996), e Cohen et al. (1998) supuseram, em seus estudos, que cada criminoso experimenta  $k$  prisões em um ano de acordo com uma mesma distribuição de Poisson.
- Mukherjee & Carcach (1998), baseados nas suposições de que a probabilidade da vitimização é a mesma para todas as pessoas (casas), e que a probabilidade de ser vitimizado não depende do número de vitimizações precedentes, é possível ajustar um modelo simples de Poisson para estes casos. Sob este modelo, as pessoas (casas) que experimentam um número elevado de vitimizações são meramente “azarados”.
- Assunção et al. (2001) supôs que o número de crimes em áreas do centro comercial de Belo Horizonte seguia uma distribuição de Poisson. Nesse estudo, a contagem do número de crimes foi feita um ano antes e um ano depois das instalações de Postos de Observação e Vigilância.

Dependendo do cálculo a ser realizado, pode ser necessário aproximar a distribuição de Poisson pela distribuição Normal ( $N(\lambda, \lambda)$ ), com função de densidade a ser definida na seção 2.4.1, desde que seja respeitada a condição de que a média ( $\lambda$ ) seja maior ou igual a 15 e utilizada a correção de continuidade (Guerra & Donaire, 1982).

Ainda segundo esses autores, tal correção consiste em considerar qual a probabilidade de ocorrer um valor genérico  $y_i$  na distribuição de Poisson (lembrando que a variável aleatória  $Y$  aqui é sempre um número inteiro) e equivale a calcular a probabilidade de ocorrer um valor entre  $(y_i - 0,5)$  e  $(y_i + 0,5)$  na distribuição Normal.

Neste trabalho, foi admitido que o número de crimes violentos (CV) anual (ou semestral) seguia uma distribuição de Poisson, uma vez que a ocorrência do evento era sucesso ou fracasso e o  $n$  pôde ser considerado infinito, pois num intervalo de um ano (ou semestre) existiam muitíssimos tempos de

oportunidades de acontecer o crime. As outras duas condições, que seriam a média se manter constante por intervalo de tempo e a independência dos ensaios, foram supostas como sendo verdadeiras, mas com a ressalva de que estudos futuros deveriam ser conduzidos para a confirmação dessas afirmações.

Ainda em relação a este trabalho, o CV foi transformado em índice de criminalidade violenta (ICV), de acordo com a expressão 12 (Diretriz nº 01 da PMMG, 2002):

$$ICV = \frac{CV \times 1000}{n^\circ \text{ de habitantes}} \quad (12)$$

Nota-se que, pela expressão 12, esse índice representa o quociente entre duas variáveis aleatórias, o numerador igual ao CV, a princípio seguindo uma distribuição de Poisson, e o denominador igual ao nº de habitantes. Mood et al. (1974) avaliaram a distribuição do quociente e apresentaram o seguinte resultado relevante:

- Sejam X e Y variáveis aleatórias contínuas distribuídas conjuntamente com densidade  $f_{X,Y}(x,y)$ . Considere uma variável aleatória definida a partir de X e Y como  $U=X/Y$ , então  $f_U(u) = \int_{-\infty}^{\infty} |y| f_{X,Y}(uy, y) dy, \quad -\infty < u < \infty$ .

Neste estudo não foi avaliada a distribuição do quociente, já que não era objetivo identificar qual distribuição poderia ser adequada para explicar o índice de criminalidade violenta.

Apesar desse complicador, o índice de criminalidade violenta foi estudado por ser considerado um padrão da Polícia Militar, que possibilita a comparação, nesse caso, entre municípios (além de facilitar o estudo da variação desse índice num período - mês, semestre, ou ano - em uma mesma cidade), e

pela possibilidade de se beneficiar do Teorema Central do Limite, para contornar o problema distribucional.

Para ambos os casos, CV (mesmo para situações em que  $\lambda \leq 15$ ) ou ICV, as análises foram feitas considerando uma aproximação pela distribuição Normal.

## **2.2 Gestão da Qualidade na Segurança Pública**

### **2.2.1 Panorama nacional**

Como relata Oliveira (2005?), o governo atual redefiniu as atribuições da Secretaria Nacional de Segurança Pública (Senasp), sendo que a mesma passou a ser responsável pela criação e a implantação da Política Nacional de Segurança Pública, cuja principal diretriz é o redesenho do aparelho policial brasileiro. Com isso, o Fundo Nacional de Segurança Pública tornou-se instrumento indutor da Política Nacional de Segurança e passou a privilegiar planos que contenham planejamento, metas, avaliação e monitoramento. Neste cenário de mudanças, o governo federal assumiu, de forma explícita, parte da responsabilidade na gestão do Sistema Brasileiro de Segurança Pública.

A Senasp, em 2003, lançou o Sistema Nacional de Estatísticas de Segurança Pública e Justiça Criminal, com o objetivo de mudar a forma de coleta de dados sobre criminalidade no país e criar um banco de dados estatísticos com informações de segurança pública e justiça de todos os Estados brasileiros, o que permitiu (ou permitirá) o aperfeiçoamento do planejamento de políticas nacionais de segurança. Assim, foram criadas condições para implantar a gestão como princípio de administração das organizações de segurança pública no Brasil.

Outra iniciativa da Senasp foi criar, em 2004, o curso de Gestão em Segurança Pública, que foi planejado para possuir quatro módulos diferentes:

Gestão de Processos em Segurança Pública, Gestão da Informação em Segurança Pública, Gestão em Segurança Pública Utilizando Geoprocessamento e Experiências Exitosas de Gestão em Segurança Pública. A criação desse curso decorreu da necessidade de capacitar os profissionais de segurança pública em relação aos princípios de gestão e também sensibilizar esses profissionais em relação à importância da informação e de todos os processos envolvendo a coleta, registro, sistematização e disseminação de informação. A sensibilização desses profissionais seria, portanto, fundamental para garantir o sucesso do Sistema Nacional de Estatísticas de Segurança Pública e Justiça Criminal, pois traria a legitimidade necessária ao sistema e garantiria que as informações fossem produzidas e utilizadas em toda a sua plenitude com o rigor metodológico necessário.

De acordo com Silva (2002), um dos mais graves problemas, hoje, no Brasil, refere-se ao aumento significativo e generalizado da violência. Estudiosos e especialistas no assunto reconhecem que as ações nessa área devem partir da construção de uma nova concepção de segurança pública, que tradicionalmente tem sido tratada somente como uma questão de polícia. Entretanto, as experiências evidenciam que nem todas as questões de segurança pública são problemas exclusivamente policiais. A promoção de uma segurança pública democrática requer a ação de vários atores sociais e depende da articulação de políticas sociais implementadas em parceria entre governos estaduais e municipais e que envolvam a participação dos cidadãos.

A participação comunitária, como estratégia mestra, constitui uma verdadeira medida para o alcance em profundidade dos graves problemas relacionados à criminalidade e violência. A comunidade tem uma função útil a desempenhar, na medida em que desenvolva o aspecto associativo e mobilize os cidadãos para ombrearem os órgãos e entidades governamentais, que possuem um excedente de tarefas e um déficit de recursos para cumprir de forma isolada

suas obrigações constitucionais, no esforço em prol da tranquilidade pública (Diretriz nº 01 da PMMG, 2002).

Como exemplo da participação comunitária, encontra-se a adoção em vários Estados brasileiros da estratégia de Polícia Comunitária e da criação dos Conselhos Comunitários de Segurança.

A idéia de uma polícia orientada para a solução de problemas e melhoria da qualidade de vida das comunidades, nos moldes da filosofia de polícia comunitária, foi introduzida no Brasil nos anos 80, tendo como principal precursor e defensor o coronel da Polícia Militar do Estado do Rio de Janeiro Carlos Nazareth Cerqueira (Beato, 2001b).

Estudiosos, pesquisadores e ativistas identificam no policiamento comunitário a solução mais imediata e coerente para os males que afligem os cidadãos. Está demonstrado, que não apenas no Brasil mas em todo o mundo, a decisão da comunidade em participar do processo da segurança pública é condição necessária para se atingir resultados satisfatórios em termos de contenção da criminalidade violenta. A meta do policiamento comunitário é a redução do crime, da desordem e do medo, por meio do exame cuidadoso das características dos problemas locais (da vizinhança), aplicando, então, as soluções apropriadas (Diretriz nº 01 da PMMG, 2002).

Na visão de Souza & Gouvêa (1995?), os Conselhos Comunitários de Segurança são grupos de pessoas do mesmo bairro ou município que se reúnem para discutir, analisar, planejar e acompanhar a solução de seus problemas comunitários de segurança e desenvolver campanhas educativas, estreitando os laços de entendimento e cooperação entre as várias lideranças locais.

Em Minas Gerais, esses Conselhos foram designados como Conselhos Comunitários de Segurança Pública (CONSEP), que, segundo a Diretriz nº 01 da PMMG (2002), tem por objetivo desenvolver programas de prevenção da criminalidade com a participação da comunidade, cumprindo a função de

planejar, junto com a polícia, as estratégias de policiamento, enfatizando-se a prevenção.

O geoprocessamento vem sendo, também, uma ferramenta útil na Gestão da Qualidade na Segurança Pública. Segundo Alves (2004), esse recurso, que há dez anos era utilizado com maior frequência para promover análises ambientais, ganhou outras aplicações. Atualmente, ele também serve ao estudo de variados fenômenos relacionados ao espaço urbano. Um deles, a violência, talvez seja o que mais preocupe os moradores das grandes e médias cidades brasileiras.

Como resultado do trabalho conjunto do Instituto Nacional de Pesquisas Espaciais (Inpe) e da Universidade Federal de Minas Gerais (UFMG), em parceria com a Senasp e a Presidência da República, foi lançado pelo Ministério da Justiça, em 2004, o Terra Crime, que é um software livre de georeferenciamento, o qual mapeia áreas de incidência criminal. A partir da análise do mapa e dos boletins, o Terra Crime permite o monitoramento e facilita o controle da criminalidade em ambiente urbano, pois possibilita a visualização ágil de estatísticas sobre crimes (Ministério da Justiça, 2003).

### **2.2.2 Panorama mineiro**

De acordo com Viégas (2003), na mesma velocidade em que a sociedade evolui, a criminalidade avança sobre todos os segmentos e camadas da população. Em contrapartida, as ações de prevenção e combate não seguem o mesmo ritmo, sobretudo porque, atualmente, esse tipo de trabalho demanda “mais cérebro e menos músculos”. Além dos pesquisadores, o comando da Polícia Militar de Minas Gerais também reconhece a importância e a necessidade da pesquisa e do estudo científico como ferramentas essenciais para a elaboração de políticas, estratégias e ações de prevenção e combate à

criminalidade. Essa mudança na elaboração de políticas na área de segurança pública possibilitou uma aproximação entre a Polícia Militar e órgãos de pesquisa e proporcionou uma análise preditiva da criminalidade e da violência. Até então, predominava a apresentação apenas de informações sobre o que já havia ocorrido (análise descritiva).

“Antes, a segurança pública era à base da força. Hoje, porém, ela é realizada com base no conhecimento científico e tecnológico”, salienta Beato, citado por Viégas (2003).

Segundo o tenente-coronel Renato Vieira de Souza, citado por Viégas (2003), a pesquisa científica é fundamental na elaboração do planejamento estratégico da Polícia Militar de Minas Gerais. A partir de 1990, a Polícia Militar intensificou a aplicação dos estudos científicos e a tecnologia como ferramenta de prevenção e combate à criminalidade e à violência em Minas Gerais.

O processo de mudança da Polícia Militar, que antes se pautava mais por ações de repressão, também é constatado pelos cursos de formação de praças (soldados, cabos e sargentos) e oficiais. O currículo inclui matérias como Geoprocessamento, Estatística, Sociologia, Filosofia e Direitos Humanos, segundo Viégas (2003), o qual ainda afirma que a Polícia Militar substituiu o empirismo pela pesquisa científica.

De acordo com Beato (2001b), em 1999, o Comando de Policiamento da Capital da Polícia de Minas Gerais desenhou um programa mais amplo, denominado "Polícia de Resultados", que passou a orientar o planejamento das ações de polícia na cidade, tendo por base o mapeamento da criminalidade e o fortalecimento do atendimento descentralizado às demandas da comunidade. A utilização do geoprocessamento e a criação dos Conselhos Comunitários de Segurança - CONSEP's, tornaram-se os pilares para a viabilização dessa nova

política de polícia. Seu principal objetivo é desenvolver programas de prevenção da criminalidade com a participação da comunidade.

Ainda segundo esse autor, para cada uma das 25 companhias de polícia há um Conselho Comunitário de Segurança que cumpre a função de planejar, junto com a polícia, estratégias de policiamento e programas adequados às necessidades da população da sua área. Os membros básicos de cada conselho são os comandantes das companhias de Polícia Militar, representantes da prefeitura de Belo Horizonte, das associações e de outras entidades (Beato, 2001b).

Outro fato importante foi o lançamento, pela Secretaria de Estado de Defesa Social de Minas Gerais, em 2004, do Boletim de Informações Criminais, com periodicidade trimestral, compondo, assim, o fornecimento de uma estatística criminal confiável e de qualidade, configurando-se como uma ferramenta gerencial importante e auxiliando no planejamento estratégico, tático e inclusive operacional, das ações dos órgãos que compõem o Sistema de Defesa Social em Minas Gerais.

É meta do governo atual de Minas Gerais a sistematização e a compatibilização das bases de dados de segurança pública do Estado, que serão plenamente alcançadas com a implantação do Sistema Integrado de Defesa Social - SIDS. Trata-se de um sistema modular e integrado, que permitirá a gestão das informações de Defesa Social relacionadas às ocorrências policiais e de bombeiros, à investigação policial, ao processo judicial e à execução penal, respeitadas as atribuições legais dos órgãos que o compõem. Além disso, o SIDS tem o objetivo de articular a integração das informações criminais ao planejamento operacional das organizações envolvidas.

O Núcleo de Estudos em Segurança Pública (NESP), da Fundação João Pinheiro, tem desempenhando um papel importante nesse processo de

reestruturação da Segurança pública, por meio de uma série de atividades que vem desenvolvendo, dentre elas se destacando as do período de 2001-2002 como seguem (Fundação João Pinheiro, 2005):

- Organização Policial e Combate à Criminalidade Violenta (1997-2001) – trata-se de uma iniciativa para aplicar as informações da pesquisa científica no geoprocessamento pela Polícia Militar. Dentro desse projeto também foi realizada uma pesquisa entre os policiais da PMMG, com o objetivo de desenvolver um estudo da organização e de seus componentes, seus valores, seus problemas, etc.
- Plano Mineiro de Direitos Humanos (2001).
- Projeto de Combate à Violência no Hipercentro de Belo Horizonte (2001-).
- Projeto de Desenvolvimento do Plano de Segurança Pública do Município de Belo Horizonte (2001-)

Esse mesmo Núcleo vem realizando, também, nos últimos anos, os seguintes cursos de Formação de Recursos Humanos:

- CEGESP - Curso de Especialização em Gestão Estratégica de Segurança Pública (PMMG) - tem como objetivo desenvolver conhecimentos, habilidades e atitudes dos oficiais maiores e tenentes-coronéis da Polícia Militar, como gerentes dos níveis estratégicos da organização, habilitando-os a cargos e funções superiores dos níveis de comando, assessoria e ao cargo de coronel.
- CESP – Curso de Especialização em Segurança Pública (PMMG).
- Curso de Estatística e Análise Criminal (Ministério da Justiça) – tem como objetivo desenvolver conhecimentos estatísticos e habilidades analíticas em profissionais da segurança pública que atuem nas áreas de produção estatística, análise criminal e planejamento operacional.

Nesses últimos anos, o Centro de Estudos de Criminalidade e Segurança Pública (Crisp) e o Laboratório de Estatística Espacial (LESTE) da Universidade Federal de Minas Gerais (UFMG) apresentaram os seguintes trabalhos ligados à área de Segurança Pública (UFMG, 2005):

- Regionalização do Estado de Minas Gerais Segundo as Taxas de Criminalidade Violenta e de Mortes por Homicídios no Período de 1996 a 2000 (Castro et al., 2004).
- Comparação de Dados de Homicídios entre Dois Sistemas de Informação em Minas Gerais (Castro et al., 2003).
- Atlas da Criminalidade de Belo Horizonte (Crisp, 2000-2002).
- Criação de uma Biblioteca de Sub-rotinas de Análise Estatísticas Integrada ao Sistema de Informação da Polícia Militar de Minas Gerais (Crisp, 2001-). De acordo com o resumo, este trabalho consiste em consultoria e assessoria para criação de uma biblioteca de sub-rotinas de análise estatística integrada ao sistema de informações a ser adquirido pela PMMG. Essa biblioteca será capaz de gerar análises estatísticas de dados georeferenciados das ocorrências de crime e atuação policial para auxiliar a tomada de decisões no combate e controle da criminalidade do Estado de Minas Gerais. A principal característica do sistema em desenvolvimento é reunir, num só elemento, dados provenientes de diversas fontes, tornar fácil sua manipulação e atualização, e apresentá-los já no formato de relatórios de análise estatística, permitindo comparações e análises.

A Universidade Federal de Lavras (UFLA), em convênio com a 6ª Região da Polícia Militar de Minas Gerais (6ª RPMMG) e com o 8º Batalhão da Polícia Militar de Minas Gerais (8º BPMMG), nos dois últimos anos, apresentou, também, alguns trabalhos ligados à área de Segurança Pública:

- Comparação dos índices de criminalidade entre municípios de mesmo porte populacional do Estado de Minas Gerais (Kataoka et al., 2004a);
- Análise estatística da influência do índice de desenvolvimento humano municipal no índice de criminalidade do Estado de Minas Gerais no ano de 2000 (Kataoka et al., 2004b);
- Uso de Séries Temporais na análise de dados de Segurança Pública no município de Ijaci (MG) (Teixeira & Sáfyadi, 2004);
- Uma proposta para o auxílio na educação continuada da Polícia Militar (Abílio et al., 2005);
- Aplicação da Estatística Básica na Gestão da Segurança Pública (Araújo et al., 2005);
- Um estudo preliminar da Segurança Pública na Região Sul de Minas Gerais (Teixeira et. al., 2005).

Diante do exposto, percebe-se que, através de metodologias estatísticas, a Polícia Militar, especialmente a de Minas Gerais, vem tentando obter respostas aos seus questionamentos, promovendo, assim, uma mudança na concepção da Gestão da Qualidade na Segurança Pública, com tomadas de decisões baseadas em análises científicas.

Neste trabalho, para avaliar os dados de criminalidade violenta em Minas Gerais, foi utilizada inferência baseada em intervalos estatísticos.

## **2.3 Intervalos estatísticos**

### **2.3.1 Generalidades**

Chama-se amostra aleatória de dimensão  $n$  a um conjunto de  $n$  variáveis aleatórias  $Y_1, Y_2, \dots, Y_n$  independentes e identicamente distribuídas (i.i.d). A fim de seguir a ortodoxia da simbologia estatística, deve-se dizer que antes da amostragem ser realizada, tem-se  $n$  variáveis aleatórias  $Y_1, Y_2, \dots, Y_n$ ; depois de

efetuada a amostragem, tem-se um conjunto de dados que constituem a amostra observada  $y_1, y_2, \dots, y_n$ .

Como citado em Guerra & Donaire (1982), o processo de indução que se pretende realizar sobre uma população pode ser feito, a partir de uma amostra, de duas maneiras:

1<sup>a</sup>) Estimação: o processo que usa os resultados extraídos de uma amostra para produzir inferências sobre a população.

Existem dois tipos de estimação:

- I) Por ponto: quando, a partir da amostra, procura-se obter um único valor de certo parâmetro  $\theta$  (função dos valores populacionais). Exemplo:  $\bar{Y}$  (média amostral) é uma estatística (função de valores amostrais) usada para fazer uma estimação por ponto do parâmetro  $\mu$  (média populacional).
- II) Por intervalo: quando, a partir da amostra, procura-se construir um intervalo de variação  $[T_1, T_2]$  (em que  $T_1$  e  $T_2$  são estatísticas) com certa probabilidade de conter o verdadeiro valor do parâmetro populacional  $\theta$ . Exemplo: obter um intervalo que contenha  $\sigma^2$  (variância populacional) com 95% de probabilidade, a partir de  $S^2$  (variância amostral).

2<sup>a</sup>) Testes de Hipótese: é o processo que usa os resultados extraídos da amostra para testar valores de certos parâmetros da população (testes paramétricos) ou para testar a natureza da distribuição da população (testes não-paramétricos ou de aderência).

Outros dois conceitos importantes são o de estimador e de estimativa. Segundo Mood et al. (1974), o estimador ( $\hat{\theta}$ ) é qualquer estatística (função conhecida de valores aleatórios observados que é por si só uma variável

aleatória) cujos valores são usados para estimar  $\tau(\theta)$ , em que  $\tau(\cdot)$  é uma função do parâmetro  $\theta$ . Já a estimativa é o valor numérico obtido pelo estimador numa certa amostra. Então, retomando o exemplo da estimação por ponto da média populacional ( $\mu$ ), o estimador pelo método dos quadrados mínimos seria dado por  $\hat{\mu} = \bar{Y} = \sum_{i=1}^n Y_i / n$  (representado por letras maiúsculas) e a estimativa por  $\bar{y} = \sum_{i=1}^n y_i / n$  (letras minúsculas).

Neste trabalho, foi dada uma atenção especial para a estimação por intervalo.

Hahn & Meeker (1991) afirmam que, quando se tenta retirar conclusões sobre a população de uma amostra aleatória selecionada apropriadamente, intervalos estatísticos desempenham um papel central na quantificação da incerteza, e fornecem um importante suplemento na estimação pontual.

Guerra & Donaire (1982) relatam que os estimadores são variáveis aleatórias, muitas vezes contínuas e as estimativas obtidas quase certamente serão distintas do valor do parâmetro (a probabilidade é, em geral, praticamente nula), ou seja, quase certamente estar-se-á cometendo um erro de estimação. Sendo assim, surge a idéia de se construir um intervalo em torno da estimativa por ponto, de modo que esse intervalo tenha uma probabilidade conhecida de conter o verdadeiro valor do parâmetro.

Vários intervalos estatísticos podem ser calculados de uma amostra de dados. O intervalo apropriado depende da aplicação específica e pode-se fazer uso de intervalos estatísticos unilaterais ou bilaterais.

## 2.3.2 Intervalos de confiança

### 2.3.2.1 Definição

Os intervalos de confiança quantificam o conhecimento, ou a falta dele, sobre um parâmetro ou alguma característica de uma população baseada em uma amostra aleatória (Hahn e Meeker, 1991).

Portanto, um intervalo de confiança pode ser definido como:

Seja  $(Y_1, Y_2, \dots, Y_n)$  uma amostra aleatória de tamanho  $n$ , de densidade  $f(y; \theta)$ , em que o parâmetro  $\theta$  pode ser um escalar ou um vetor (dois ou mais parâmetros) e seja  $\tau(\theta)$  uma função do parâmetro  $\theta$  a cuja estimação interessa proceder. Em particular, se  $\tau(\theta) = \theta$ , cai-se na estimação do próprio parâmetro, que foi o caso utilizado neste trabalho (a simbologia  $\tau(\theta)$  será “abandonada” em detrimento de utilizar apenas  $\theta$ ).

Sejam  $T_1 = T_1(Y_1, Y_2, \dots, Y_n)$ ,  $T_2 = T_2(Y_1, Y_2, \dots, Y_n)$ ,  $T_1 < T_2$ , duas estatísticas tais que:

$$P(T_1 < \theta < T_2) = 1 - \alpha \quad 0 < \alpha < 1 \quad (13)$$

Assim,  $1 - \alpha$  traduz o coeficiente de confiança que se tem em que uma particular amostra dá origem a um intervalo de confiança bilateral  $(t_1, t_2)$ , que compreenda o verdadeiro valor do parâmetro  $\theta$  (para maiores detalhes ver Murteira, 1990).

### 2.3.2.2 Interpretação

Um intervalo de confiança a  $100(1 - \alpha)\%$  para uma quantidade desconhecida  $\theta$  pode ser interpretado da seguinte forma: se repetidamente forem calculados tais intervalos  $(t_1, t_2)$  de muitas amostras aleatórias independentes,

mas sob as mesmas condições, em  $100(1-\alpha)\%$  dos intervalos devem, ao longo do tempo, conter corretamente o verdadeiro valor de  $\theta$ .

Na prática, comumente, mas menos precisamente, um intervalo de confiança bilateral é descrito por uma declaração tal como: com  $100(1-\alpha)\%$  de confiança pode-se afirmar que o verdadeiro valor de  $\theta$  está contido no intervalo entre  $T_1$  e  $T_2$ . De fato, o intervalo observado contém ou não  $\theta$ . Então  $100(1-\alpha)\%$  refere-se ao procedimento para construção do intervalo estatístico, e não ao intervalo observado por si só (Hahn & Meeker, 1991).

### **2.3.2.3 Exemplo: Intervalo de confiança (IC) bilateral $100(1-\alpha)\%$ para a média $\mu$ de uma população Normal com variância $\sigma^2$ conhecida**

Para esse tipo de intervalo, as referidas estatísticas  $T_1$  e  $T_2$  seriam respectivamente  $\bar{Y} - z_{\alpha/2}\sigma_{\bar{Y}}$  e  $\bar{Y} + z_{\alpha/2}\sigma_{\bar{Y}}$ , portanto o IC é construído de acordo com a seguinte expressão:

$$P(T_1 < \theta < T_2) = P(\bar{Y} - z_{\alpha/2}\sigma_{\bar{Y}} < \mu < \bar{Y} + z_{\alpha/2}\sigma_{\bar{Y}}) = 1 - \alpha \quad (14)$$

em que:

$\bar{Y}$  = média amostral;

$z_{\alpha/2}$  = quantil ( $\alpha/2$ ) de uma distribuição Normal Padrão;

$\sigma_{\bar{Y}}$  = desvio padrão da média amostral:  $\sigma_{\bar{Y}} = \frac{\sigma}{\sqrt{n}}$ .

A interpretação para esse tipo de intervalo pode ser feita da seguinte forma: se forem selecionadas 100 amostras da mesma população e tamanho  $n$ , obter-se-ão 100 estimativas de  $\bar{Y}$ , com as quais serão construídos 100 IC para  $\mu$ , sendo que os valores de  $z_{\alpha/2}$  e  $\sigma_{\bar{Y}}$  serão os mesmos para todas as amostras (fixos).

Se  $\alpha=5\%$ , pode-se esperar que 95 dos IC contenham o verdadeiro valor de  $\mu$  e 5 não contenham o valor de  $\mu$ , como ilustra a Figura 1 (Morettin, 2000).

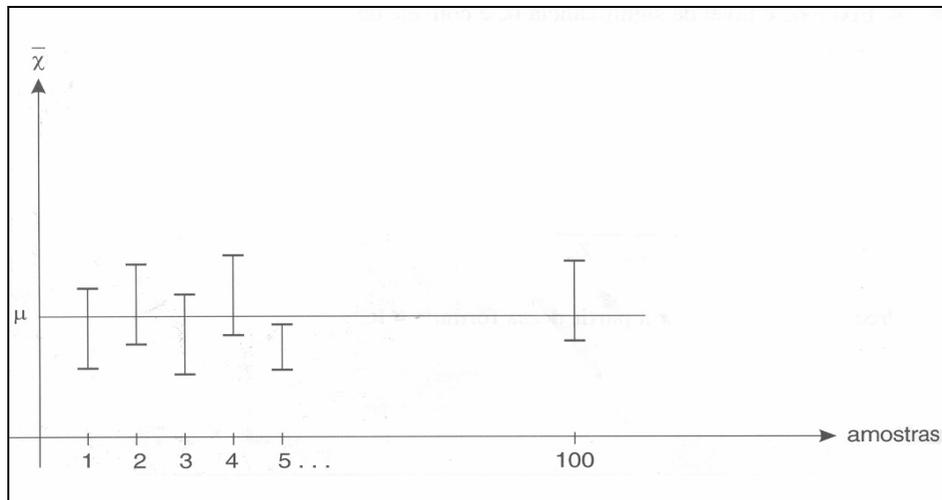


FIGURA 1 Representação genérica dos IC para  $\mu$  de 100 amostras da mesma população e tamanho  $n$ .

Observando a Figura 1, verifica-se que o IC da 5ª amostra não conteria o verdadeiro valor de  $\mu$ , sendo que ainda haveria mais 4 amostras nessa mesma situação até completar o total de 100 amostras.

Logo, em uma amostra qualquer, a probabilidade de que o IC contenha o verdadeiro valor da média ( $\mu$ ) é de 95%. Portanto, corre-se o risco de 5% de que ele não contenha esse valor.

Vale salientar que essa forma de interpretar um intervalo (sendo que a Figura 1 é específica para esse exemplo) pode ser utilizada para os outros intervalos que serão discutidos a seguir, resguardando as devidas particularidades de cada um.

### 2.3.3 Intervalos de predição

#### 2.3.3.1. Definição

Um intervalo de predição contém uma ou mais futuras observações, ou alguma função de tais futuras observações, de uma população amostrada previamente (Hahn & Meeker, 1991).

Então, um intervalo de predição pode ser definido da seguinte forma:

Seja  $(Y_1, Y_2, \dots, Y_n)$  uma amostra aleatória (passada) de tamanho  $n$ , de densidade  $f(\cdot; \theta)$  em que o parâmetro  $\theta$  pode ser um escalar ou um vetor.

Sejam  $T_1 = T_1(Y_1, Y_2, \dots, Y_n)$  e  $T_2 = T_2(Y_1, Y_2, \dots, Y_n)$ ,  $T_1 < T_2$  duas estatísticas baseadas na amostra observada. Seja  $(X_1, X_2, \dots, X_m)$  uma amostra aleatória futura de tamanho  $m$  da mesma distribuição  $f(\cdot; \theta)$ . Suponha-se que a amostra futura seja retirada independentemente da amostra passada. Seja  $T = T(X_1, X_2, \dots, X_m)$  uma estatística baseada na amostra futura. Agora, se  $T_1$  e  $T_2$  são determinadas, então:

$$P(T_1 < T < T_2) = 1 - \alpha \quad 0 < \alpha < 1 \quad (15)$$

Assim,  $1 - \alpha$  traduz o coeficiente de confiança que se tem, em que uma particular amostra passada dá origem a um intervalo de predição bilateral  $(t_1, t_2)$ , que compreenda o verdadeiro valor de  $T$  de uma amostra futura (para maiores detalhes ver Patel, 1989).

De acordo com Fertig et al. (1980), tais intervalos têm importância prática em projetos de confiabilidade. Patel (1989) comenta que um intervalo de predição é um intervalo o qual usa resultados de uma amostra passada para conter resultados de uma amostra futura de uma mesma população com uma probabilidade especificada.

Segundo Whitmore (1986), livros de Estatística Básica invariavelmente introduzem o tópico de predição no contexto de regressão simples para um modelo bivariado normal. Ainda segundo esse autor, deveria ser mais apropriado, contudo, introduzir o tópico em conjunção com a inferência univariada normal.

### **2.3.3.2. Interpretação**

A interpretação de um intervalo de predição com grau de confiança de  $100(1-\alpha)\%$  para conter, por exemplo, uma observação futura é dada a seguir: se repetidamente forem calculados tais intervalos  $(t_1, t_2)$  de muitos pares de amostras aleatórias independentes (computados dos dados da primeira amostra para conter um valor de uma segunda amostra futura) , em  $100(1-\alpha)\%$  dos intervalos deve, ao longo do tempo, conter corretamente o valor da observação futura. Como no caso do intervalo de confiança, o  $100(1-\alpha)\%$  refere-se ao procedimento usado para construir o intervalo de predição e não para qualquer intervalo que é computado (Hahn & Meeker, 1991).

Os intervalos de predição têm geralmente comprimentos mais largos quando comparados com os intervalos de confiança porque existem duas fontes de variações envolvidas na construção deles. A primeira delas é que parâmetros desconhecidos precisam ser estimados usando dados passados e a segunda, que uma amostra futura traz uma variação aleatória adicional (Patel, 1989).

### **2.3.4 Intervalos de tolerância**

#### **2.3.4.1 Definição**

Como citado na seção 2.3.2, os intervalos de predição são necessários para prever uma unidade futura ou um pequeno número delas. No caso em que se deseja tirar conclusões de um número relativamente grande de observações

futuras, baseadas em dados de uma amostra aleatória da população de interesse, conceitualmente, pode-se também construir intervalos de predição para tais situações como, por exemplo, intervalo de predição para conter todos os 100, 1000 ou qualquer  $m$ . Contudo, tais intervalos freqüentemente são muito longos, além disso, pode ser que o número de unidades futuras de interesse seja desconhecido ou conceitualmente infinito. Então, nestes casos, podem ser calculados intervalos de tolerância. Tais intervalos contêm pelo menos uma proporção especificada de unidades de uma população amostrada, com um grau de confiança pré-estabelecido de  $100(1-\alpha)\%$  (Hahn & Meeker, 1991).

Portanto, um intervalo de tolerância para conter pelo menos uma proporção  $\gamma$  da população pode ser definido da seguinte forma:

Seja  $(Y_1, Y_2, \dots, Y_n)$  uma amostra aleatória de tamanho  $n$ , de densidade  $f(y; \theta)$ , em que o parâmetro  $\theta$  pode ser um escalar ou um vetor.

Sejam  $T_1 = T_1(Y_1, Y_2, \dots, Y_n)$ ,  $T_2 = T_2(Y_1, Y_2, \dots, Y_n)$ ,  $T_1 < T_2$  duas estatísticas. Então,  $P_Y(T_1 < Y < T_2)$  é conhecida como a cobertura do intervalo  $[T_1, T_2]$ .

Sejam  $\gamma$  e  $1-\alpha$  duas constantes tais que  $0 < \gamma < 1$  e  $0 < 1-\alpha < 1$ . Assim se  $T_1$  e  $T_2$  são determinados tais que:

$$P\left[P_Y(T_1 < Y < T_2) \geq \gamma\right] = 1 - \alpha \quad (16)$$

Dessa forma,  $1-\alpha$  traduz o coeficiente de confiança que se tem em que uma particular amostra dá origem a um intervalo de tolerância bilateral  $(t_1, t_2)$  que contenha pelo menos uma proporção  $\gamma$  da população (para maiores detalhes ver Patel, 1986).

Tietjen & Johnson (1979) ressaltam que, no controle estatístico da qualidade, a intenção é expor um intervalo onde se espera que uma fração

especificada de uma amostra estatística esteja contida quando o processo está sendo controlado. O intervalo de tolerância ajusta essa descrição satisfatoriamente. Essa observação é particularmente lúcida para mostrar a predominância dos intervalos de confiança (IC) sobre os intervalos de predição e de tolerância, mesmo em situação em que o IC não é o mais indicado, como neste caso.

Hahn & Chandra (1981) comentam que, algumas vezes, em aplicação de controle de qualidade requer-se a obtenção de intervalos de tolerância para as distribuições Binomial e Poisson. Segundo Patel (1986), entre os vários problemas estatísticos que aparecem no processo de controle de qualidade de produção em massa, é importante determinar os limites de tolerância em que a variabilidade do produto é conhecida por conter fatores aleatórios.

O problema de determinar limites de tolerância foi tratado inicialmente por Wilks (1941) dentro de uma abordagem não paramétrica. Já em 1942, esse mesmo autor trabalhou o caso paramétrico, sendo que, em 1943, Wald estendeu os resultados de Wilks para distribuições multivariadas. Mas somente em 1946, com Wald & Wolfowitz, é que foram realizados os primeiros estudos de limites de tolerância para uma distribuição normal, sendo até hoje o artigo mais citado em qualquer referência sobre o assunto, o que motivou ter sido utilizado como procedimento neste trabalho.

#### **2.3.4.2 Interpretação**

A interpretação de um intervalo de tolerância com grau de confiança de  $100(1-\alpha)\%$  para conter pelo menos uma proporção  $\gamma$  da população é dada a seguir: se repetidamente forem calculados tais intervalos  $(t_1, t_2)$  de muitos grupos independentes de amostras aleatórias, em  $100(1-\alpha)\%$  dos intervalos

devem, ao longo do tempo, corretamente ser incluídos pelo menos  $100\% \gamma$  dos valores da população (Hahn & Meeker, 1991)

Como os intervalos de confiança e predição, o  $100(1-\alpha)\%$  refere-se ao procedimento de construção do intervalo de tolerância e não a qualquer intervalo particular que é computado. A proporção atual da população contida dentro do intervalo de tolerância é desconhecida porque essa proporção depende de parâmetros desconhecidos.

## **2.4 Intervalos estatísticos para uma variável aleatória que obedece ao modelo $Y=\mu+\epsilon$ com distribuição Normal (Modelo simples Normal)**

### **2.4.1 Distribuição Normal**

A distribuição Normal é a mais importante distribuição de probabilidade em estatística. As razões para isso são muitas, mas algumas delas podem ser enumeradas: a) a maioria das técnicas empregadas na estatística é baseada na distribuição normal; b) inúmeros são os fenômenos aleatórios cujos comportamentos podem ser descritos precisamente ou de forma aproximada pelo modelo probabilístico normal; c) a distribuição normal é a forma limitante de muitas outras distribuições de probabilidades, como consequência do teorema central do limite; d) muitas estatísticas apresentam distribuições assintóticas normais, ou seja, a distribuição da estatística se aproxima da normal à medida que o tamanho da amostra cresce (Ferreira, 2005).

Uma variável aleatória  $Y$  é definida por ser distribuída normalmente se a sua densidade é dada por:

$$f(y, \mu, \sigma) = \frac{1}{\sqrt{2\pi}\sigma} \exp \left[ -\frac{1}{2} \left( \frac{y - \mu_y}{\sigma_y} \right)^2 \right] \quad (17)$$
$$-\infty < y < \infty, \quad -\infty < \mu_y < \infty, \quad \sigma_y > 0$$

em que  $\mu$  é a média e  $\sigma_y$  é o desvio padrão.

Como ressaltam Hahn & Meeker (1991), a sensibilidade de um intervalo para a suposição de normalidade depende do tipo específico de intervalo, da população ou processo que está sendo estudado, e do tamanho da amostra, pelo menos de 20 a 30 observações.

#### 2.4.2 Modelo $Y=\mu+\epsilon$ com distribuição Normal

Seja  $Y$  uma variável aleatória de interesse. Pode-se usar a esperança de  $Y$  igual a  $\mu$  ( $E[Y]=\mu_y$ ) e a variância de  $Y$  igual a  $\sigma_y^2$  ( $\text{Var}[Y]=\sigma_y^2$ ) para descrever de maneira resumida a sua natureza, expressando, então, a variável  $Y$  como uma soma  $\mu+\epsilon$ , em que  $\epsilon$  é uma variável aleatória com esperança igual a zero e variância igual à variância de  $Y$ ,  $\sigma_y^2$ . Assim,  $\epsilon$  representa os inúmeros fatores que, conjuntamente, fazem as observações de  $Y$  oscilarem em torno de  $\mu_y$ . Esse modo de expressar  $Y$  pode ser chamado de modelo simples (Charnet et al., 1999).

No caso particular de  $Y$  ter distribuição normal ( $N(\mu_y, \sigma_y^2)$ ) o modelo simples é caracterizado por:

$$Y = \mu + \epsilon$$

$$\mu_y : \text{constante}$$

$$\epsilon \sim N(0, \sigma_y^2)$$

Vários tipos de intervalos estatísticos podem ser construídos considerando-se o modelo simples normal, contudo, neste trabalho, foram citados apenas alguns deles.

### 2.4.3 Intervalo de Confiança para a média populacional ( $\mu$ ) com variância $\sigma^2$ desconhecida

Seja uma amostra aleatória  $(Y_1, Y_2, \dots, Y_n)$  de tamanho  $n$  e as estatísticas

$$T_1 = \bar{Y}_n - t_{(\alpha/2; n-1)} \frac{S_n}{\sqrt{n}} \text{ e } T_2 = \bar{Y}_n + t_{(\alpha/2; n-1)} \frac{S_n}{\sqrt{n}}, \text{ um intervalo de confiança}$$

bilateral  $100(1-\alpha)\%$  para o parâmetro  $\mu$  pode ser descrito pela expressão 18:

$$P\left[T_1 \leq \theta \leq T_2\right] = P\left[\bar{Y}_n - t_{(\alpha/2; n-1)} \frac{S_n}{\sqrt{n}} \leq \mu \leq \bar{Y}_n + t_{(\alpha/2; n-1)} \frac{S_n}{\sqrt{n}}\right] = 1-\alpha \quad (18)$$

em que:

$$\bar{Y}_n = \frac{\sum_{i=1}^n Y_i}{n} \text{ (média amostral);}$$

$$S_n = \sqrt{\frac{\sum_{i=1}^n (Y_i - \bar{Y}_n)^2}{n-1}} \text{ (desvio padrão amostral);}$$

$t_{(\alpha/2; n-1)}$  =  $(\alpha/2)$  quantil da distribuição  $t$  com  $n-1$  graus de liberdade.

Sendo  $\bar{Y}_n$  e  $S_n$  estimadores, respectivamente, da média e do desvio padrão populacionais pelo método dos quadrados mínimos.

Um intervalo de confiança para a média populacional não é sensível para desvios de normalidade na população amostrada, a menos que a amostra seja bem pequena e esses desvios pronunciados, sendo então considerado um intervalo robusto, justificado pelo Teorema Central do Limite (Hahn & Meeker, 1991).

#### 2.4.4 Intervalo de predição para a média de m futuras observações ( $\bar{X}_m$ )

Seja  $(Y_1, Y_2, \dots, Y_n)$  uma amostra aleatória de tamanho n,  $(X_1, X_2, \dots, X_m)$  uma amostra aleatória futura de tamanho m de uma mesma distribuição e as estatísticas  $T_1 = \bar{Y}_n - t_{(\alpha/2; n-1)} S_n \sqrt{\frac{1}{m} + \frac{1}{n}}$  e  $T_2 = \bar{Y}_n + t_{(\alpha/2; n-1)} S_n \sqrt{\frac{1}{m} + \frac{1}{n}}$ ; um intervalo de predição bilateral com grau de confiança  $100(1-\alpha)\%$  para a média de m futuras observações pode ser descrito pela expressão 19:

$$P[T_1 \leq T \leq T_2] = P\left[ \bar{Y}_n - t_{(\alpha/2; n-1)} S_n \sqrt{\frac{1}{m} + \frac{1}{n}} \leq \bar{X}_m \leq \bar{Y}_n + t_{(\alpha/2; n-1)} S_n \sqrt{\frac{1}{m} + \frac{1}{n}} \right] = 1 - \alpha \quad (19)$$

Da mesma forma que o intervalo de confiança para a média, esses intervalos são bastante insensíveis para desvios de normalidade, a menos que a amostra dada ou a amostra futura sejam muito pequenas, ou com desvios acentuados.

#### 2.4.5 Intervalo de predição para uma observação futura

Um intervalo de predição para uma observação futura simples é um intervalo que, com um grau de confiança especificado, conterá a próxima (ou alguma outra especificada) observação selecionada aleatoriamente da população, sob as mesmas condições da amostra inicial.

Esse intervalo é um caso particular de (19) para  $m=1$  e pode ser descrito pela expressão 20:

$$P[T_1 \leq T \leq T_2] = P\left[ \bar{Y}_n - t_{(\alpha/2; n-1)} S_n \sqrt{1 + \frac{1}{n}} \leq X_1 \leq \bar{Y}_n + t_{(\alpha/2; n-1)} S_n \sqrt{1 + \frac{1}{n}} \right] = 1 - \alpha \quad (20)$$

#### 2.4.6 Intervalo de tolerância para conter pelo menos $\gamma\%$ da população com um grau de confiança de $100(1-\alpha)\%$

Para construção dos intervalos de tolerância, considerando uma distribuição normal, é necessário calcular primeiro o fator de tolerância ( $k$ ).

Odeh (1978) determinou valores exatos para  $k$ -fatores de tolerância usando integração numérica. Contudo, vários procedimentos foram desenvolvidos para o cálculo aproximado de  $k$ , como, por exemplo:

- I) Aproximação de Wald & Wolfowitz (1946);
- II) Aproximação de Weissberg & Beaty (1960);
- III) Aproximação de Howe (1969).

Para uma amostra aleatória de tamanho  $n$ , o procedimento I só pode ser utilizado nos casos em que  $v$  (graus de liberdade) é igual a  $n-1$ . Já os procedimentos II e III podem ser utilizados quando  $v \neq n-1$ , que ocorrem nos cálculos dos limites de tolerância no contexto da regressão (em que  $v$  está em função do número de parâmetros envolvidos -  $v=n-p$ ), em modelos de efeitos aleatórios, em dados com erros de mensuração, entre outros.

Contudo, quando  $v$  for muito maior do que  $n$ , os procedimentos II e III são bem menos precisos do que o método exato (para maiores detalhes veja Eberhardt et al., 1989).

Neste trabalho, foi utilizada a aproximação de Wald & Wolfowitz (1946), como comentado na seção 2.3.3.1. Portanto, seja, uma amostra aleatória  $(Y_1, Y_2, \dots, Y_n)$  de tamanho  $n$  e as estatísticas

$$T_1 = \bar{Y}_n - kS_n \sqrt{\frac{n-1}{\chi^2_{(1-\alpha; n-1)}}} \text{ e } T_2 = \bar{Y}_n + kS_n \sqrt{\frac{n-1}{\chi^2_{(1-\alpha; n-1)}}};$$

um intervalo de tolerância bilateral para conter pelo menos  $\gamma\%$  da população com um grau de confiança de  $100(1-\alpha)\%$  pode ser descrito pela expressão 21:

$$P \left[ P_Y \left( \bar{Y}_n - kS_n \sqrt{\frac{n-1}{\chi_{(1-\alpha; n-1)}^2}} < Y < \bar{Y}_n + kS_n \sqrt{\frac{n-1}{\chi_{(1-\alpha; n-1)}^2}} \right) \geq \gamma \right] = 1 - \alpha \quad (21)$$

em que  $k$  é o fator de tolerância, sendo a raiz da equação 22.

$$\frac{1}{\sqrt{2\pi}} \int_{1/\sqrt{n-k}}^{1/\sqrt{n+k}} e^{-t^2/2} dt = \gamma \quad (22)$$

Segundo Wald & Wolfowitz (1946), os limites de tolerância calculados pela expressão 21 diferem dos valores exatos pela magnitude na ordem de  $O\left(\frac{1}{n^2}\right)$ , isto é, essa diferença aproxima-se de zero rapidamente (na velocidade de  $n^2$ ) à medida que  $n \rightarrow \infty$ . Como um exemplo ilustrativo, se  $n$ , o tamanho da amostra, dobra, a diferença fica quatro vezes menor; se triplica, nove vezes menor, e assim por diante.

Os intervalos de tolerância tendem a envolver inferências sobre as caudas das distribuições, justamente onde os desvios de normalidade são mais pronunciados. Por essa razão, tais intervalos devem estar seriamente comprometidos quando a distribuição básica não é normal, especialmente para altos níveis de confiança e quando a proporção da população a ser contida dentro do intervalo está próxima de 1 (Hahn & Meeker, 1991).

Em 1984, Canavos & Koutrouvelis examinaram a robustez dos limites de tolerância da expressão 21 para distribuição normal baseados em simulação computacional, em que as distribuições gama e  $t$  foram usadas como modelos geradores. Os resultados indicaram que os limites de tolerância normal são sensíveis a desvios de normalidade quando  $k$ -fatores de tolerância são selecionados para cobertura excedendo a 0,9. Razoável robustez é achado para cobertura de 0,9, bem como para distribuições básicas que não tem caudas extremamente pesadas e nem altamente enviesadas.

### 2.4.7 Comparação entre os intervalos no modelo simples normal

Para o modelo simples normal ( $Y = \mu + \varepsilon$ ) pode-se verificar que o intervalo de tolerância (IT) é uma conjugação do intervalo de predição (IP) com o intervalo de confiança (IC), isto é, o IT é uma “correção” que se faz no IP por causa da incerteza da média dada pelo IC. Então, para verificar tal afirmação, pode-se aumentar o tamanho da amostra, fazendo com que tenda ao infinito ( $n \rightarrow \infty$ ) e, dessa forma, o intervalo de confiança para a média populacional ( $\mu_Y$ ), utilizando a expressão 18, vai tender para o próprio valor de  $\mu_Y$ , uma vez que:

$$\frac{S_n}{\sqrt{n}} \rightarrow 0$$

$$\bar{Y} \rightarrow \mu_Y$$

Logo,

$$\bar{Y}_n \pm t_{(\alpha/2; n-1)} \frac{S_n}{\sqrt{n}} \rightarrow \mu_Y \quad (23)$$

Já no caso do intervalo de predição para uma observação futura, considerando a expressão 20, tem-se que:

$$\frac{1}{\sqrt{n}} \rightarrow 0$$

$$t_{(\alpha/2; n-1)} \rightarrow Z_{(\alpha/2)}$$

$$S_n \rightarrow \sigma_Y$$

Portanto,

$$\bar{Y}_n \pm t_{(\alpha/2; n-1)} S_n \sqrt{1 + \frac{1}{n}} \rightarrow \mu_Y \pm Z_{(\alpha/2)} \sigma_Y \quad (24)$$

Para o intervalo de tolerância para conter pelo menos uma proporção  $\gamma$  da população, primeiramente deve-se determinar o fator de tolerância ( $k$ ). Então, através da expressão 22, tem-se que:

$$\frac{1}{\sqrt{n}} \rightarrow 0$$

$k \rightarrow z_{\gamma/2}$  (quantil  $\gamma/2$  de uma distribuição  $N\sim(0,1)$ )

Logo,

$$\frac{1}{\sqrt{2\pi}} \int_{\frac{1}{\sqrt{n} \rightarrow \infty}}^{\frac{1}{\sqrt{n}} + k} e^{-t^2/2} dt \rightarrow \frac{1}{\sqrt{2\pi}} \int_{-k}^k e^{-t^2/2} dt = \gamma \quad (25)$$

E, utilizando a expressão 21:

$$\sqrt{\frac{n-1}{\chi^2_{(1-\alpha; n-1)}}} \rightarrow 1$$

Então,

$$\bar{Y}_n \pm k S_n \sqrt{\frac{n-1}{\chi^2_{(1-\alpha; n-1)}}} \rightarrow \mu_Y \pm z_{(\gamma/2)} \sigma_Y \quad (26)$$

Verifica-se, portanto, que, se o valor da proporção  $\gamma$  do intervalo de tolerância for muito próximo ao grau de confiança  $(1-\alpha)$  do intervalo de predição, as expressões 24 e 26 são praticamente idênticas, uma vez que  $z_{(\alpha/2)} \cong z_{(\gamma/2)}$ .

Estes argumentos se prestam, pois, para justificar afirmação de que o intervalo de tolerância é uma “correção” do intervalo de predição pela incorporação da incerteza da média expressada pelo intervalo de confiança.

## **2.5 Regressão linear simples**

Retomando o modelo simples  $Y = \mu + \varepsilon$  (definido na seção 2.4.2), considere-se agora, a existência de uma outra variável,  $X$ , com alguma relação com a variável  $Y$ , o que sugere uma maneira alternativa de estudar  $Y$ , tendo como base informações sobre  $X$  (Charnet et al., 1999).

A regressão é exatamente a busca de um modelo que estabelece uma relação funcional entre duas variáveis. Especificamente, os métodos que estabelecem relações lineares para os parâmetros de um modelo entre duas ou mais variáveis são denominados métodos de regressão linear (Ferreira, 2005).

Os primeiros trabalhos com regressão começaram a ser desenvolvidos a partir de 1877, com Francis Galton estudando o coeficiente de reversão (atualmente regressão) e Adcock com a avaliação da regressão simples com erro na variável independente, sendo que em 1887, Francis Galton lançou o livro intitulado Teoria da Regressão (Cordeiro, 2002).

Uma regressão linear é considerada múltipla quando existem “ $p$ ” parâmetros envolvidos no modelo. Então, quando “ $p$ ” for igual a 2, será considerado um caso particular, denominado regressão linear simples.

Neste trabalho, foi analisada a relação funcional linear entre o número de crimes violentos (CV) e o ano, bem como o índice de criminalidade violenta (ICV) e o ano, no período de 1991 a 2004, utilizando, para tal fim, o modelo de regressão linear simples, que será definido a seguir (seção 2.5.1).

### **2.5.1 Modelo de regressão linear simples (MRLS)**

De acordo com Charnet et al. (1999), podem-se considerar duas maneiras para obtenção de uma amostra: valores de  $X$  prefixados e, para estes valores, obtenção de observações independentes de  $Y$ , ou obtenção de uma amostra aleatória de  $(X, Y)$  bivariada. Em ambos os casos, tem-se uma amostra de tamanho  $n$ , sendo  $x_1, x_2, \dots, x_n$  os valores prefixados de  $X$ , ou os valores

observados de  $X$ , e  $y_1, y_2, \dots, y_n$  os correspondentes valores observados de  $Y$  (variáveis aleatórias).

Neste trabalho, o valor da variável “ano” foi prefixado, ou seja, não foi aleatória. Então, para todas as expressões a serem apresentadas, será considerado esse caso.

Um modelo de regressão linear simples pode ser apresentado na forma geral pela equação 27 (Ferreira, 2005):

$$Y_i = \beta_0 + \beta_1 x_i + \varepsilon_i \quad (27)$$

em que:

$1 \leq i \leq n$  (tamanho da amostra);

$Y_i$ : variável dependente (ou variável resposta) associada à  $i$ -ésima observação;

$x_i$ : variável independente (ou variável auxiliar) associada à  $i$ -ésima observação;

$\beta_0$ : coeficiente linear ou intercepto;

$\beta_1$ : coeficiente de regressão;

$\varepsilon_i$ : erro aleatório associado à  $i$ -ésima observação.

### **2.5.2 Estimação de $\beta_0$ e $\beta_1$ empregando o método dos quadrados mínimos ordinários**

Para a estimação dos parâmetros, comumente, são empregados o método dos quadrados mínimos e o método da máxima verossimilhança, que conduzem aos mesmos estimadores. A ressalva que deve ser feita é que, no caso do método da máxima verossimilhança, deve-se assumir a normalidade para a distribuição dos erros, sendo, então, mais restritivo do que o método dos quadrados mínimos, que não necessita de nenhuma pressuposição (Ferreira, 2005)

O método dos quadrados mínimos determina os estimadores de  $\beta_0$  e  $\beta_1$  que minimizam a soma de quadrados dos erros. Então, utilizando (27), tem-se:

$$S = \sum_{i=1}^n \epsilon_i^2 = \sum_{i=1}^n [Y_i - (\beta_0 + \beta_1 x_i)]^2 \quad (28)$$

em que S terá o seu mínimo quando as derivadas parciais em relação à  $\beta_0$  e  $\beta_1$  forem nulas.

$$\frac{\partial S}{\partial \beta_0} = -2 \sum_{i=1}^n [Y_i - (\beta_0 + \beta_1 x_i)] \quad (29)$$

$$\frac{\partial S}{\partial \beta_1} = -2 \sum_{i=1}^n [Y_i - (\beta_0 + \beta_1 x_i)] (x_i) \quad (30)$$

igualando 29 e 30 a zero, obtêm-se os estimadores 31 e 32:

$$\hat{\beta}_0 = \bar{Y} - \hat{\beta}_1 \bar{x} \quad (31)$$

$$\hat{\beta}_1 = \frac{\sum_{i=1}^n Y_i (x_i - \bar{x})}{\sum_{i=1}^n (x_i - \bar{x})^2} \quad (32)$$

em que  $\bar{x}$  e  $\bar{Y}$  representam, respectivamente, as médias da variável independente e dependente.

Se o teste da segunda derivada for feito, serão obtidos como resultados para  $\frac{\partial^2 S}{\partial \beta_0^2}$ ,  $\frac{\partial^2 S}{\partial \beta_1^2}$  e  $\frac{\partial^2 S}{\partial \beta_0 \partial \beta_1}$ , respectivamente,  $2n$ ,  $2 \sum_{i=1}^n x_i^2$  e  $2 \sum_{i=1}^n x_i$ . Logo, as expressões 33 e 34 serão sempre positivas, indicando que realmente S tem mínimo em  $(\hat{\beta}_0, \hat{\beta}_1)$ :

$$\frac{\partial^2 S}{\partial \beta_0^2} \frac{\partial^2 S}{\partial \beta_1^2} - \left[ \frac{\partial^2 S}{\partial \beta_0 \partial \beta_1} \right]^2 = 4n \left[ \sum_{i=1}^n (x_i - \bar{x})^2 \right] \quad (33)$$

$$\frac{\partial^2 S}{\partial \beta_0^2} + \frac{\partial^2 S}{\partial \beta_1^2} = 2n + 2 \sum_{i=1}^n x_i^2 \quad (34)$$

Pode ser demonstrado que os estimadores de quadrados mínimos são lineares não tendenciosos de variância mínima (Hoffmann e Vieira, 1998).

### 2.5.3 Diagnóstico das pressuposições básicas

Segundo Ferreira (2005), se inferências forem realizadas no MRLS, os erros (variável aleatória  $\varepsilon_i$ ) deverão assumir um modelo específico de probabilidade e devem ser homocedásticos (ter variância constante).

Neste trabalho, foi pressuposto que o modelo de probabilidade do erro era o modelo normal, com  $E(\varepsilon_i)=0$  e  $\text{Var}(\varepsilon_i)=\sigma^2$ , isto é,  $\varepsilon_i \sim N(0, \sigma^2)$ ; por conseguinte, supõe-se, também, que os erros não eram correlacionados ( $E[\varepsilon_i \varepsilon_j]=0$  para  $i \neq j$ ).

O diagnóstico das pressuposições sobre os erros pode ser feito por meio dos seus estimadores, denominados resíduos ( $\hat{\varepsilon}_i = \hat{e}_i = Y_i - \hat{Y}_i$ ), utilizando a análise visual do esboço de alguns gráficos e/ou testes estatísticos descritos a seguir.

(i) Pressuposição I: os erros têm variância constante

Para um dado valor de  $x$ , a variância do erro  $\varepsilon$  é sempre  $\sigma^2$ , denominado variância residual, isto é,  $E[\varepsilon^2]=\sigma^2$ . Nesse caso, o erro é dito homocedástico.

Para a verificação dessa pressuposição pode-se fazer análise gráfica, esboçando os valores de  $y$  ajustados versus os resíduos, pois, se a variância for constante, os resíduos se distribuem aleatoriamente em torno da média zero, como pode ser observado na Figura 2:

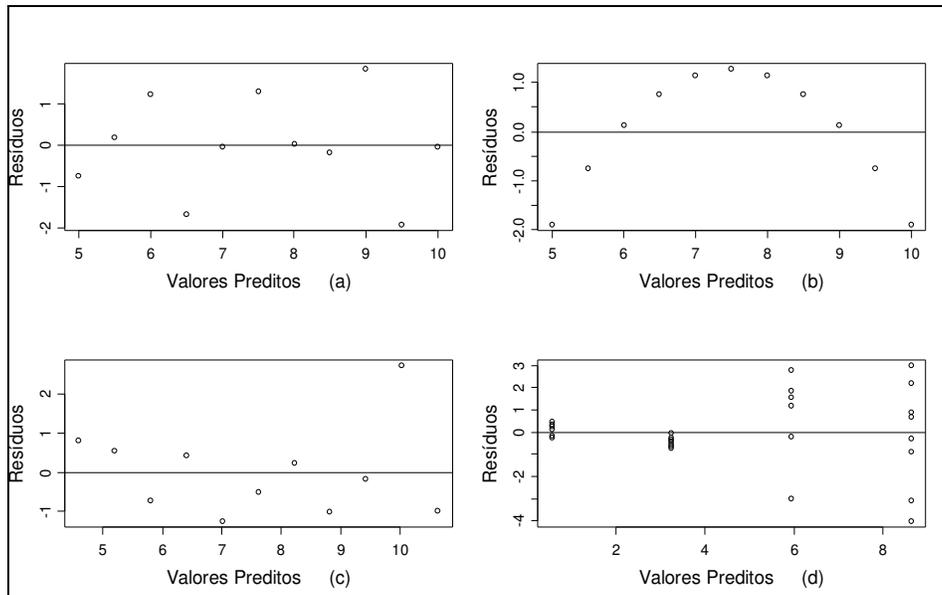


FIGURA 2 Gráficos dos resíduos versus os valores preditos para quatro exemplos: (a) situação ideal; (b) modelo não linear; (c) elemento atípico; (d) heterocedasticidade.

Na Figura 2, o gráfico (a) representa justamente a situação ideal, o que sugere que a variância é constante; (b) sugere uma mudança na forma estrutural do modelo ajustado por não apresentar um comportamento linear; (c) apresenta um elemento atípico que pode ser considerado um outlier (valor discrepante), e que talvez deva ser retirado para uma melhor adequação do modelo; (d) indica que os resíduos não devem ter variância constante, ou seja, são heterocedásticos; nesse caso, pode ser necessária uma transformação na variável resposta, ou a utilização do método dos mínimos quadrados ponderados (Hahn & Meeker, 1991).

(ii) Pressuposição II: normalidade dos erros

Os erros devem ter distribuição Normal ( $\varepsilon \sim N(0, \sigma^2)$ ), sendo que essa pressuposição é necessária para que possam ser utilizadas as distribuições de t e F para testar a hipótese a respeito dos valores dos parâmetros  $\beta_0$  e  $\beta_1$  ou construir intervalos de confiança, por exemplo, para  $E[Y_h]$ .

Para testar a suposição de normalidade de uma variável aleatória por meio de uma amostra aleatória  $X_1, X_2, \dots, X_n$  pode ser utilizado o teste de Shapiro-Wilk proposto em 1965, com a estatística W (Ferreira, 2005) apresentada em 35:

$$W = \frac{\left( \sum_{i=1}^n a_i x_{(i)} \right)^2}{\sum_{i=1}^n (x_i - \bar{x})^2} \quad (35)$$

A expressão 35 é dependente do coeficiente  $a_i$ , que é denominado por melhor estimador linear não-viesado normalizado do valor esperado das estatísticas de ordem da distribuição normal padrão. A determinação exata desses coeficientes é dependente da determinação exata dos valores esperados e da matriz de variâncias e covariâncias das estatísticas de ordem da distribuição normal padrão (Ferreira, 2005).

Para avaliar ainda essa pressuposição, pode ser esboçado o gráfico dos quantis teóricos da distribuição Normal Padrão versus os resíduos padronizados. Então, se os pontos estiverem próximos da reta da 1ª bissetriz (sendo as escalas dos 2 eixos cartesianos as mesmas), este fato é um indicativo para a normalidade dos resíduos.

(iii) Pressuposição III: os erros são independentes

O erro de uma observação não é correlacionado com o erro em outra observação, isto é,  $E[\varepsilon_i \varepsilon_j] = 0$  para  $i \neq j$ .

Em 1950, foi desenvolvido o teste de Durbin e Watson, que avalia a presença ou não de correlação serial ou autocorrelação, utilizando a estatística descrita em 36 (Draper & Smith, 1998). A hipótese de nulidade testada é que a autocorrelação dos resíduos é igual a zero:

$$d = \frac{\sum_{i=2}^n (e_i - e_{i-1})^2}{\sum_{i=1}^n e_i^2} \quad (36)$$

em que  $e_i = y_i - \hat{y}_i$  é a estimativa dos resíduos (já definido na seção 2.5.3).

Esse teste é conduzido da seguinte maneira: compara-se  $d$  ou  $4-d$  (o que estiver mais próximo de zero) com os valores de  $d_L$  e  $d_U$  presentes em tabela específica, podendo ocorrer as situações apresentadas no Quadro 2, conforme indica Draper & Smith (1998).

QUADRO 2 Interpretação da comparação dos valores de  $d$  ou  $4-d$  com  $d_L$  e  $d_U$  para o teste de Durbin-Watson.

Situações	Interpretação
$d < d_L$	correlação em série positiva
$d > d_U$ ou $4-d > d_U$	correlação em série não indicada
$4-d < d_L$	correlação em série negativa
$d_L < d < d_u$ ou $d_L < 4-d < d_U$	teste inconclusivo

Quando esse teste é realizado automaticamente no software estatístico R<sup>®</sup> v2.1.1 (R, 2005), com o comando “dwtest” (ver apêndice), são apresentados como resultados o valor da estatística e o p-valor, não havendo, portanto, necessidade do uso de tabelas.

Se os erros forem correlacionados isso indica que os valores observados não são estatisticamente independentes. Então, pode ser apropriada a utilização dos métodos dos quadrados mínimos generalizados ou a análise de séries temporais (Hahn & Meeker, 1991).

Assim, após a verificação das pressuposições básicas, é possível construir intervalos estatísticos para uma observação futura, para  $E(Y_i)$ , ou uma determinada proporção da população, além de outros.

#### **2.5.4 Intervalos estatísticos no contexto da regressão**

Como comentado na seção anterior, a validade dos intervalos estatísticos depende da adequação do modelo assumido. No modelo de regressão linear simples geralmente (é o caso neste trabalho) assume-se que a variável resposta observada (Y) segue uma distribuição normal.

Porém, antes de um estudo mais detalhado sobre os intervalos, é necessário determinar o estimador da variância de  $\hat{Y}_i$ , bem como redefinir o modelo estatístico. Vale lembrar que, neste trabalho, o valor de x (variável independente) foi prefixado.

##### **2.5.4.1 Estimação da variância de $\hat{Y}_i$**

Para estimar a variância de  $\hat{Y}_i$  utilizando (27) é necessário saber que (Hoffmann & Vieira, 1998):

$$V(\hat{\beta}_0) = \sigma^2 \left[ \frac{1}{n} + \frac{\bar{x}^2}{\sum_{i=1}^n (\bar{x}_i - \bar{x})^2} \right], \quad V(\hat{\beta}_1) = \frac{\sigma^2}{\sum_{i=1}^n (x_i - \bar{x})^2} \quad \text{e} \quad \text{cov}(\hat{\beta}_0, \hat{\beta}_1) = -\frac{\sigma^2 \bar{x}}{\sum_{i=1}^n (x_i - \bar{x})^2}$$

Com os resultados anteriores, pode-se verificar que o estimador da variância de  $\hat{Y}_i$  é dado pela expressão 37:

$$\begin{aligned} V(\hat{Y}_i) &= V(\hat{\beta}_0 + \hat{\beta}_1 x_i) = V(\hat{\beta}_0) + x_i^2 V(\hat{\beta}_1) + 2x_i \text{cov}(\hat{\beta}_0, \hat{\beta}_1) \\ V(\hat{Y}_i) &= \sigma^2 \left[ \frac{1}{n} + \frac{(x_i - \bar{x})^2}{\sum_{i=1}^n (x_i - \bar{x})^2} \right] \\ \hat{V}(\hat{Y}_i) &= S^2(\hat{Y}_i) = S^2 \left[ \frac{1}{n} + \frac{(x_i - \bar{x})^2}{\sum_{i=1}^n (x_i - \bar{x})^2} \right] \end{aligned} \quad (37)$$

em que  $S^2$  é o estimador de quadrados mínimos de  $\sigma^2$  (variância do erro) dado pela expressão 38:

$$S^2 = \frac{\sum_{i=1}^n (y_i - \hat{y}_i)^2}{(n-2)} \quad (38)$$

Muitas vezes, para a construção dos intervalos estatísticos, pretende-se estimar um valor de  $Y$  correspondente a um valor de  $x$  que não existe na amostra. Se for reservado o índice  $i$  para indicar os elementos pertencentes à amostra, deve-se introduzir um outro índice ( $h$ ) para indicar outros valores de  $X$ . O novo valor  $x_h$ , pode coincidir ou não com um dos valores ( $x_i$ ) da amostra.

Portanto, a partir de (27) e (37), o modelo estatístico e o estimador da variância de  $\hat{Y}_h$  passam a ser apresentados pelas expressões 39 e 40, respectivamente:

$$\hat{Y}_h = \hat{\beta}_0 + \hat{\beta}_1 x_h \quad (39)$$

$$S^2(\hat{Y}_h) = S^2 \left[ \frac{1}{n} + \frac{(x_h - \bar{x})^2}{\sum_{i=1}^n (x_i - \bar{x})^2} \right] = S^2 A \quad (40)$$

em que  $A = \frac{1}{n} + \frac{(x_h - \bar{x})^2}{\sum_{i=1}^n (x_i - \bar{x})^2}$ .

#### 2.5.4.2 Intervalo de confiança bilateral para a média ( $E(Y_h)$ )

Podem ser construídos intervalos de confiança para os parâmetros, para o valor esperado ou a média da variável resposta e ainda para a variância  $\sigma^2$  das observações para um dado conjunto de  $x_i$ 's.

Porém, neste trabalho, foi abordado apenas o caso do intervalo de confiança para a média,  $E(Y_h) = \beta_0 + \beta_1(x_h)$ , sendo que o mesmo pode ser calculado por meio da expressão 41 (Hoffmann & Vieira, 1998):

$$P \left[ \hat{Y}_h - t_{(\alpha/2; n-2)} S(\hat{Y}_h) < E[Y_h] < \hat{Y}_h + t_{(\alpha/2; n-2)} S(\hat{Y}_h) \right] = 1 - \alpha \quad (41)$$

em que:

$t_{(\alpha/2; n-2)}$  = quantil ( $\alpha/2$ ) da distribuição t com n-2 graus de liberdade.

Deve-se observar o comportamento do intervalo de confiança no caso da extrapolação, sendo que essa ocorre quando se tenta prever a resposta para

valores do preditor que estão fora dos limites da amostra. Segundo Faraway (2005), existem dois tipos de extrapolação:

(i) Extrapolação quantitativa: deve ser checado se o novo  $x_h$  está dentro da extensão dos dados originais. Se não, a predição pode não ser realista. Intervalos de confiança para predições são mais largos à medida que se move para fora da amostra. Contudo, esse alargamento pode não refletir a possibilidade que a estrutura do modelo pode mudar. A relação pode tornar-se não linear para fora da extensão da amostra.

(ii) Extrapolação qualitativa: o novo  $x_h$  foi retirado da mesma população que a amostra original? Se o modelo foi construído no passado e este foi usado para predições futuras, muitas vezes é difícil julgar se as condições permanecem constantes.

De acordo com Hoffmann & Vieira (1998), ao se fazer uma extrapolação, há, ainda, um outro problema mais sério que o crescimento da amplitude do intervalo de confiança (e também do intervalo de predição). É o fato de que freqüentemente o modelo (linear) ajustado é razoável para o intervalo coberto pela amostra, mas é absolutamente inapropriado para extrapolação.

#### 2.5.4.3 Intervalo de predição bilateral para uma observação futura

Para a construção do Intervalo de Predição (IP) para uma observação futura pode ser utilizada a expressão 42 (Hoffmann & Vieira, 1998):

$$P\left[\hat{Y}_h - t_{(1-\alpha/2;n-2)}\left[(1+A)S^2\right]^{1/2} < Y_h < \hat{Y}_h + t_{(1-\alpha/2;n-2)}\left[(1+A)S^2\right]^{1/2}\right] = 1 - \alpha \quad (42)$$

De acordo com Hoffmann & Vieira (1998), o conceito de intervalo de previsão é análogo ao de intervalo de confiança, com a diferença de que,

enquanto o intervalo de confiança se refere a uma constante (por exemplo,  $E(Y_h)$ ), o intervalo de predição se refere a uma variável aleatória ( $Y_h$ , no caso).

Hahn & Meeker (1991) fazem referências a literaturas para o estudo do intervalo de predição no contexto da regressão. Por exemplo:

- No capítulo 3 de Draper & Smith (1998) e nos capítulos 3 e 7 de Neter et al. (1990) são apresentados métodos para o cálculo do intervalo de predição para a média de  $m$  futuras observações;
- Lieberman (1961) mostra uma aproximação dos intervalos de predição simultâneos para conter todas as  $m$  futuras observações para um modelo de regressão.

#### **2.5.4.4 Intervalo de tolerância unilateral para conter uma proporção $\gamma$ da população.**

Para a construção dos intervalos de tolerância unilateral tanto inferior como superior podem ser utilizadas, respectivamente, as expressões 43 e 44, descritas por Kendal & Stuart (1979).

$$P\left[P_{Y_h}\left(Y_h > \hat{Y}_h - g_\gamma S\right) = \gamma\right] = 1 - \alpha \quad (43)$$

$$P\left[P_{Y_h}\left(Y_h < \hat{Y}_h + g_\gamma S\right) = \gamma\right] = 1 - \alpha \quad (44)$$

em que :

$$g_\gamma = -(A)^{1/2} t_{(1-\alpha, n-2, \delta)} ;$$

$t_{(1-\alpha, n-2, \delta)}$  = quantil  $(1 - \alpha)$  de uma distribuição t com n-2 graus de liberdade e parâmetro  $\delta$  de não centralidade;

$$\text{delta} = \delta = -\frac{N_\gamma}{A};$$

$N_\gamma$  = quantil  $\gamma$  de uma distribuição  $N \sim (0,1)$  ;

$$A = \left[ \frac{1}{n} + \frac{(X_h - \bar{X})^2}{\sum_{i=1}^n (X_i - \bar{X})^2} \right], \text{ de acordo com (40).}$$

O limite de tolerância inferior determina o valor no qual acima dele está contida uma proporção  $\gamma$  da população; já para o caso do limite superior fica determinado o valor abaixo.

#### 2.5.4.5 Intervalo de tolerância bilateral para conter pelo menos uma proporção $\gamma$ da população

Para a construção do intervalo de tolerância bilateral de acordo com o procedimento não simultâneo de Wallis (1951) pode ser utilizada a expressão 45:

$$P \left[ P_{Y_h} \left( \hat{Y}_h - kS < Y_h < \hat{Y}_h + kS \right) \geq \gamma \right] = 1 - \alpha \quad (45)$$

em que:

$$k = r \sqrt{\frac{n-2}{\chi_{(1-\alpha; n-2)}^2}} ;$$

$$r = N_\gamma \left( 1 + \frac{1}{2N} - \frac{2N_\gamma^2 - 3}{24N^2} \right) ;$$

$$N' = \frac{1}{A^2}.$$

Hahn & Meeker (1991) fazem referências a literaturas para o estudo do intervalo de tolerância no contexto da regressão. Por exemplo:

- Lieberman & Miller (1963) fornecem tabelas dos fatores de tolerância e exemplo de intervalos de tolerância simultâneo para a distribuição da variável resposta a partir de uma série de condições da variável independente;
- No capítulo 3 de Miller (1981), em Bowden (1968), em Turner & Bowden (1977, 1979) e em Limam & Thomas (1988), são descritos outros métodos para construir limites de tolerância.

## 3 MATERIAL E MÉTODOS

### 3.1 Dados

Foram utilizadas três bases de dados para a realização deste trabalho.

A primeira foi cedida pelo 8º batalhão da Polícia Militar de Minas Gerais (8º BPMMG). Corresponde aos valores mensais por tipo de crimes violentos ocorridos na cidade de Lavras – MG, no período de 1991 até outubro de 2005, bem como o número de crimes violentos (CV) anuais, no período de 1997 a 2003, nos demais municípios mineiros.

A segunda foi obtida no *site* do Instituto Brasileiro de Geografia e Estatística (IBGE). Refere-se ao número de habitantes por município no Estado de Minas Gerais, no período de 1991 a 2004. Vale salientar que os dados dos períodos de 1992 a 1995, 1997 a 1999 e 2001 a 2005 são estimativas calculadas por meio de fórmulas específicas pelo IBGE, sempre divulgadas no mês de julho de cada ano.

A terceira base de dados foi obtida no *site* do Centro de Informações Sociais – CIS (Batitucci, 2005c), estando disponíveis os valores mensais por tipo de crimes violentos nos municípios mineiros, no período de 1991 a 2001. Neste trabalho, foram utilizadas apenas as informações até o ano de 1996.

É bom lembrar que, como já citado na seção 2.1.3, durante o período de 1991 a 1999, aconteceram algumas mudanças em termos jurídicos, doutrinários e operacionais que podem ter alterado a comensurabilidade da base de dados da Polícia Militar e, conseqüentemente, das análises estatísticas.

Portanto, um modelo estatístico ajustado, incluindo esse período, pode propiciar discrepâncias nos intervalos estatísticos construídos e valores preditos em relação aos dados reais.

Seria, então, recomendável trabalhar apenas com os dados de 2000 a 2004. Porém, para algumas análises estatísticas no contexto da regressão, uma amostra de tamanho 5 pode gerar informações também não confiáveis.

Em função de todas essas restrições, neste trabalho, foi analisado o número de crimes violentos para os dois períodos, de 1991 a 2004 e de 2000 a 2004, para o município de Lavras – MG.

### **3.2 Estudo preliminar**

Para verificar a necessidade de ajustar um modelo estatístico que explicasse o número de crimes violentos (CV) no município de Lavras em função da população, foi realizado um estudo envolvendo o número de habitantes e o CV de alguns municípios mineiros. Foram separados, a cada ano, no período de 1991 a 2003, os municípios mineiros que tinham número de habitantes distribuídos em três faixas: 62.000 a 68.000, 72.000 a 78.000 e 82.000 a 88.000.

A escolha dessas três faixas teve como objetivo selecionar municípios com grandeza populacional similar à cidade de Lavras, uma vez que, durante esse período, a população dessa cidade variou entre 65.000 a 83.000. Além disso, a estratificação possibilitou a criação de classes homogêneas. A amplitude de cada faixa foi determinada à partir de um equilíbrio entre o tamanho da amostra e a homogeneidade do estrato.

Diante dessas informações, foram calculadas as médias do número de crimes violentos anuais e do número de habitantes para cada faixa populacional.

O próximo passo foi avaliar o “comportamento” do número de crimes violentos com a manutenção de uma população relativamente constante ao longo dos anos, ou seja, estudaram-se as seguintes questões: (i) o CV se manteria também constante? (ii) Ou aumentaria? Para tal fim, foi analisado, em conjunto, um gráfico representando as médias de CV anuais para cada faixa populacional.

Após esse estudo preliminar, foram feitas as análises, considerando-se somente o número de crimes violentos no município de Lavras – MG, de 1991 a 2004, e, posteriormente, para o período de 2000 a 2004.

### **3.3 Comportamento do número de crimes violentos contra a pessoa e contra o patrimônio no período de 1991 a 2004 em Lavras – MG**

Os crimes violentos foram subdivididos em duas categorias: contra a pessoa e contra o patrimônio (como foi informado na seção 2.1.2). E as análises passaram a ser feitas considerando-se esses dois tipos de crimes.

#### **3.3.1 Testes de Aderência $\chi^2$ para distribuição uniforme discreta**

Nesta etapa, foi avaliado o comportamento mensal dos crimes violentos contra a pessoa (CVPES) e contra o patrimônio (CVPT) ao longo dos meses. Para tal fim, foram aplicados testes de aderência  $\chi^2$ . A hipótese testada era de que o CVPES ou o CVPT de cada mês, ao longo do ano, ocorriam com igual frequência. A estatística do teste de aderência está descrita pela expressão 46 (Bussab & Morettin, 2003):

$$\chi_c^2 = \sum_{i=1}^n \frac{(O_i - E_i)^2}{E_i} \quad (46)$$

Em (25)  $O_i$  é o  $i$ -ésimo dado observado,  $E_i$  é o  $i$ -ésimo dado esperado numa amostra de tamanho  $n$  e  $\chi_c^2$  com graus de liberdade igual a  $n-1$ . Pelos resultados obtidos, foi necessário separar por semestre os dados de CVPT e proceder a novos testes de aderência.

Como apresentado na discussão procedida na seção 2.1.4, o CVPES e o CVPT foram transformados em índice de criminalidade violenta contra a pessoa (ICVPES) e contra o patrimônio (ICVPT), respectivamente, de acordo com a expressão 12.

A partir da sistematização dos dados procurou-se ajustar modelos de regressão linear simples para explicar o CV ou o ICV em função do ano.

### 3.3.2 Regressão linear simples

#### 3.3.2.1 Modelo estatístico

Para os quatro tipos de dados (CVPES, CVPT, ICPES e ICPT), os modelos estatísticos utilizados para o ajuste, no período de 1991 a 2004, foram adaptados da expressão 11, ou seja,  $Y_k = \beta_0 + \beta_1 k + \epsilon_k$ , em que:

$k = \{1, 2, \dots, 14\}$  - correspondente, respectivamente, aos anos de  $\{1991, 1992, \dots, 2004\}$ ;

$Y_k$  = variável resposta (CV ou ICV) para cada um dos anos  $k$ ;

$\beta_0$  = parâmetro, denominado de coeficiente linear ou intercepto;

$\beta_1$  = parâmetro, denominado de coeficiente de regressão;

$\epsilon_k$  = erros aleatórios para cada um dos anos  $k$ .

Em seguida, foram estimados os parâmetros  $\beta_0$  e  $\beta_1$  por meio das expressões 31 e 32 (seção 2.5.2) e a variância do erro pela expressão 38 (seção 2.5.4.1), utilizando-se o método dos quadrados mínimos e o coeficiente de determinação ajustado foi calculado pela expressão 47:

$$R^2_{\text{ajust.}} = 1 - \left[ \frac{n-1}{n-2} \right] \left[ 1 - \frac{\sum_{i=1}^n (\hat{y}_i - \bar{y})^2}{\sum_{i=1}^n (y_i - \bar{y})^2} \right] \quad (47)$$

### 3.3.2.2 Diagnóstico das pressuposições básicas

Como comentado na seção 2.5.3, neste trabalho, presumiu-se que  $\text{Var}(\varepsilon_k) = \sigma^2$ ,  $\varepsilon_k \sim N(0, \sigma^2)$  e que os erros não eram correlacionados, sendo feito o diagnóstico dessas pressuposições por meio dos resíduos, utilizando-se a análise visual do esboço de alguns gráficos e/ou testes estatísticos citados a seguir:

(i) Pressuposição I: os erros têm variância constante

Essa pressuposição tem por princípio a verificação da homocedasticidade dos resíduos (isto é,  $E[\varepsilon_k^2] = \sigma^2$ ). Nesse caso, optou-se por esboçar o gráfico dos valores de  $y$  ajustados versus os resíduos.

(ii) Pressuposição II: normalidade dos erros

Para verificar se os erros tinham distribuição Normal ( $\varepsilon \sim N(0, \sigma^2)$ ), utilizaram-se dois procedimentos: o esboço do gráfico dos quantis teóricos da distribuição Normal Padrão versus os resíduos padronizados e o teste de Shapiro-Wilk (estatística descrita na seção 2.5.3).

(iii) Pressuposição III: os erros são independentes

Para detectar se o erro de uma observação era não correlacionado com o erro de uma outra observação ( $E[\varepsilon_i \varepsilon_j] = 0$  para  $i \neq j$ ) foi utilizado o teste de Durbin-Watson (estatística descrita na seção 2.5.3).

### 3.3.2.3 Intervalos estatísticos no contexto da regressão

Para o cálculo dos limites inferiores e superiores foi necessário estimar valores de CVPES, CVPT, ICPES e ICVPT correspondentes a um ano ( $k$ ) que não existia na amostra. Por esse motivo, foi feita uma alteração na simbologia e,

portanto, a variável ano (vale reforçar que os valores foram prefixados, não sendo considerada uma variável aleatória) passou a ser representada pela letra h. E o novo valor h coincidia ou não com um dos valores k da amostra. Portanto, o modelo estatístico, lembrando que Y pode ser o CV ou o ICV, passou a ser representado pela expressão 49:

$$Y_h = \beta_0 + \beta_1 h + \varepsilon_h \quad (49)$$

Para este estudo, o grau de confiança  $(1-\alpha)$  foi considerado igual a 0,95, sendo utilizadas as expressões descritas na seção 2.5.4 para a construção dos seguintes intervalos estatísticos:

- Intervalo de confiança bilateral para a  $E(Y_h) = \beta_0 + \beta_1 h$ ;
- Intervalo de predição bilateral para uma observação futura ( $Y_h$ );
- Intervalo de tolerância unilateral (inferior e superior) para conter o CV ou o ICV de 95% ( $\gamma\%$ ) da população;
- Intervalo de tolerância bilateral para conter o CV ou o ICV de pelo menos 95% ( $\gamma\%$ ) da população.

Para facilitar uma análise conjunta, os quatro intervalos foram esboçados em um único gráfico com projeção para 20, anos contados a partir de 1991, para os quatro tipos de dados, tendo sido feito um estudo em particular para o ano de 2005, comparando-se os intervalos estatísticos construídos e o valor predito com os dados reais disponíveis (até outubro de 2005).

As mesmas análises, com exceção dos testes de aderência, foram repetidas, considerando-se, agora, o período de 2000 a 2004 e, no caso dos intervalos estatísticos, foram feitas projeções para 7 anos, contados a partir de 2000.

Neste trabalho, para a construção dos limites do intervalo de tolerância unilateral foi necessário determinar o quantil da distribuição t não central de

acordo com ano (h) requerido. Para tanto utilizou-se um processo iterativo de tentativa e erro, apoiado computacionalmente, uma vez que o comando “qt” do software estatístico R® (R, 2005) não permite a inclusão do parâmetro de não centralidade. Os resultados desse processo estão apresentados nas Tabelas 2A e 3A (Anexo A), respectivamente, para os períodos de 1991 a 2004 e de 2000 a 2004.

Importa ressaltar que o software estatístico SISVAR® v4.0 (Ferreira, 2000) calcula esse quantil de forma direta, rápida e com alto grau de precisão.

Após a análise dos dados de Lavras – MG, no contexto da regressão, avaliou-se a construção e a interpretação de intervalos estatísticos para o caso de uma amostra, considerando-se a população com distribuição normal.

### 3.3.3 Distribuição de Poisson

Como analisado na seção 2.1.4, é razoável supor que, pelo menos à princípio, o número de crimes violentos anuais segue uma distribuição de Poisson. Então, baseando-se nos resultados da regressão (valores preditos) e tomando-se como exemplo o número de crimes violentos contra a pessoa (CVPES), foi possível comparar a probabilidade, por um lado, considerando-se a distribuição de Poisson (expressão 5) desse tipo de crime ser maior do que o seu valor observado e, por outro, considerando-se a aproximação pela Normal ( $N(\lambda, \lambda)$ ), para cada ano (k), no período de 1991 a 2004, de acordo com a expressão 48:

$$P(\text{CVPES} \geq \text{CVPES}_{\text{obs}_k}) = \sum_{\text{CVPES}_{\text{obs}_k}}^{\infty} \frac{e^{-\lambda} \lambda^{\text{CVPES}}}{(\text{CVPES})!} \quad (48)$$

em que:

$\text{CVPES}_{\text{obs}_k}$  = número de crimes violentos observados no ano k;

$\lambda = \text{CVPES}_k$  = valor predito para o CVPES no ano k.

È bom salientar que a correção de continuidade, a que se refere na seção 2.1.4, utilizada para essa situação foi a de  $\text{CVPESobs}_k - 0,5$ .

### **3.4 Inferência sobre a amostra de uma população Normal**

Com a finalidade de se apresentar um exemplo ilustrativo para a aplicação do modelo simples Normal, buscou-se, na lista de municípios brasileiros no ano de 2003, um estrato de pequena amplitude que contivesse pelo menos 50 municípios mineiros, a fim de não prejudicar a exemplificação em decorrência de uma amostra pequena. Após um exame da lista, observou-se que o estrato de 14.000 a 17.500 habitantes, representado por 413 municípios brasileiros, atendia a esse critério, e o estrato foi, então, tomado como exemplo.

Deve-se observar que esse estrato não foi utilizado para comparação com os estratos de 62.000 a 68.000, 72.000 a 78.000 e 82.000 a 88.000 habitantes, anteriormente definidos na seção 3.1, sendo apenas, como já mencionado, uma aplicação do modelo simples Normal.

Então, tomando como referência os municípios de Minas Gerais, foi testada a suposição de normalidade do ICV total através do teste de Shapiro Wilk. Com nível de significância de 5%, a hipótese de normalidade não foi rejeitada. (ver os resultados na seção 4.4).

Após esse processo inicial, determinou-se a retirada de uma amostra aleatória de tamanho 30, dos 50 municípios mineiros, para se construir alguns intervalos estatísticos. Obviamente, na prática, não haveria sentido selecionar 30 de 50 em detrimento de usar todos os 50. Apesar disso, nesta dissertação, esse expediente foi usado para fins de ilustração de uma metodologia amostral.

Foram construídos, também, quatro tipos de Intervalos estatísticos bilaterais considerando-se o modelo  $Y = \mu + \varepsilon$  com distribuição Normal e as fórmulas descritas na seção 2.4:

- Intervalo de confiança para a média populacional com variância  $\sigma^2$  desconhecida;
- Intervalo de predição para uma observação futura;
- Intervalo de predição para a média de m futuras observações;
- Intervalo de tolerância para conter uma determinada proporção  $\gamma$  da população – aproximação de Wald & Wolfowitz (1946).

Neste trabalho, foi utilizado um processo iterativo de tentativa e erro, apoiado computacionalmente, para determinação do fator “k” da expressão 22, para o intervalo de tolerância pela aproximação de Wald & Wolfowitz, isto é, o valor que mais aproximava o resultado da integração ao valor de  $\gamma$  (proporção desejada da população) de uma distribuição normal padrão.

### **3.5 Software estatístico utilizado**

Para todas as análises realizadas foi utilizado o software estatístico R<sup>®</sup> v2.1.1 (R, 2005), cujas as rotinas estão disponíveis no Anexo C.

## 4 RESULTADOS E DISCUSSÃO

### 4.1 Estudo preliminar: número de habitantes versus número de crimes violentos

Os resultados correspondentes às médias do número de crimes violentos e do número de habitantes para cada faixa populacional nos anos de 1991 a 2003 dos municípios mineiros estão apresentados na Tabela 1, em que:

Ano = 1, 2,...,13 correspondem ,respectivamente, a 1991, 1992,...,2003;

Faixa = 1, 2 e 3 representam o número de habitantes entre 62000 a 68000, 72000 a 78000 e 82000 a 88000, respectivamente.

n = número de municípios envolvidos<sup>3</sup> no estudo.

TABELA 1 – Média dos números de crimes violentos e de habitantes de cada faixa populacional dos municípios mineiros nos anos de 1991 a 2003.

Ano	Faixa 1	CV	n	Faixa 2	CV	n	Faixa 3	CV	n
1	65119	43	4	73496	62	3	85311	80	5
2	66517	45	7	74820	48	2	85939	78	6
3	64259	44	5	73408	79	2	86245	80	6
4	65607	43	6	73675	60	3	85746	105	3
5	66542	43	6	74691	60	3	86384	87	2
6	64473	48	3	76299	72	6	87621	83	2
7	65400	73	5	75134	69	4	85896	191	1
8	64850	76	5	75568	87	6	86735	183	1
9	63410	101	5	75781	94	7	84326	125	3
10	63438	118	8	75527	143	4	83995	81	2
11	66276	126	8	75968	150	6	86235	144	2
12	66289	115	5	76353	149	3	86036	137	1
13	66185	119	4	76376	167	3	84829	148	2

<sup>3</sup> Alguns municípios não foram incluídos no estudo por terem um CV considerado muito elevado (“prováveis outliers”), sendo, ao todo: 15, na faixa 1; 2, na faixa 2 e 3, na faixa 3.

Com os dados da Tabela 1, foi avaliado o comportamento do número de crimes violentos ao longo dos anos para cada faixa populacional, como pode ser observado na Figura 3.

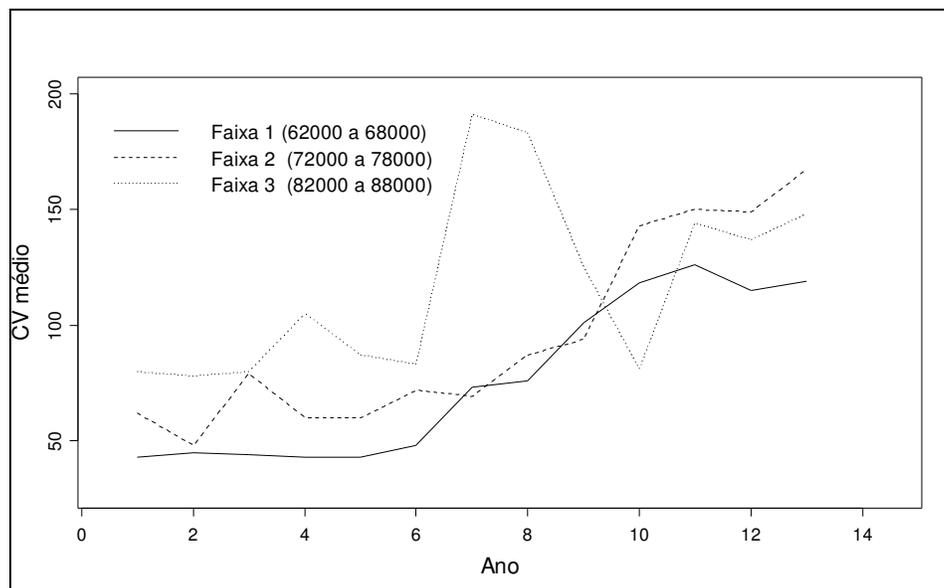


FIGURA 3 Representação gráfica do ano versus o CV médio dos municípios mineiros para cada faixa populacional.

De acordo com a Figura 3, verifica-se que, mesmo mantendo-se em cada faixa o número de habitantes aproximadamente constante, o número de crimes violentos continuou ao longo dos anos, apresentando uma tendência de crescimento, com exceção de alguns pontos como, por exemplo, os anos de 1992 e de 1994 da faixa 2.

Diante desse fato, levantou-se a suposição de que o crescimento populacional não era o fator único para explicar o aumento do número de crimes

violentos. Dessa forma, não há justificativa para promover o ajuste de uma equação em função apenas do número de habitantes. Portanto, será apresentado, nas próximas seções, como foi incorporado esse fator aos modelos estatísticos com o cálculo do ICV.

Para um aprofundamento dos estudos, foram utilizados apenas os dados de crimes violentos ocorridos no município de Lavras – MG, no período de 1991 a 2004.

#### **4.2 Comportamento do número de crimes violentos contra a pessoa e contra o patrimônio no período de 1991 a 2004 em Lavras - MG**

Os crimes violentos foram subdivididos em duas categorias: contra a pessoa e contra o patrimônio (como foi informado, na seção 2.1.2, o comportamento e o tipo de ação envolvida no combate desses crimes é diferente). O primeiro estudo realizado avaliou se esses tipos de crimes ocorriam com igual frequência nos meses do ano, por meio de testes de aderência  $\chi^2$ .

##### **4.2.1 Testes de aderência $\chi^2$**

O número de crimes violentos contra a pessoa (CVPES) e contra o patrimônio (CVPT) mensais ocorridos no município de Lavras – MG, no período de 1991 a 2004, estão apresentados nas Tabelas 2 e 3, respectivamente, e foram obtidos por meio do somatório  $\sum_m x_{mjk}$ , em que:

\*  $1 \leq m \leq 3$  = tipos de crimes violentos contra a pessoa (Tabela 1A – Anexo A) e

$\sum_m x_{mjk} = xps_{jk}$ , no caso da Tabela 2;

\*  $4 \leq m \leq 42$  = tipos de crimes violentos contra o patrimônio (Tabela 1A – Anexo A) e  $\sum_m x_{mjk} = xpt_{jk}$ , no caso da Tabela 3;

\*  $1 \leq j \leq 12$  = meses de janeiro, fevereiro, ..., dezembro;

\*  $1 \leq k \leq 13$  = anos de 1991, 1992, ..., 2004;

\*  $w_k$  = número de habitantes no ano k.

Para ilustrar melhor a simbologia adotada anteriormente, pode-se tomar, por exemplo,  $x_{ps_{11}} = 1$  (primeiro valor da Tabela 2) como sendo o resultado do somatório dos três tipos de crimes violentos contra a pessoa no mês de janeiro do ano de 1991 e  $w_1 = 65893$  (primeira linha e última coluna da Tabela 2) como sendo o número de habitantes no ano de 1991, no município de Lavras – MG.

TABELA 2 – Número de habitantes e crimes violentos contra a pessoa no município de Lavras - MG, no período de 1991 a 2004

Ano(k)	Mês(j)												Total	$w_k$
	1	2	3	4	5	6	7	8	9	10	11	12		
1	1	0	1	0	1	2	0	2	1	1	0	1	10	65893
2	0	0	0	1	2	1	1	0	0	0	0	0	5	67233
3	0	0	0	0	2	0	0	1	2	1	3	2	11	68517
4	0	1	0	1	0	0	0	1	0	0	0	0	3	69704
5	1	0	1	1	2	1	1	0	0	0	2	2	11	70871
6	0	0	1	0	0	1	1	2	0	0	0	0	5	72947
7	2	1	3	2	0	1	0	0	1	2	1	3	16	74945
8	0	2	2	2	1	0	4	0	1	1	2	0	15	76633
9	0	2	1	1	0	0	0	2	1	0	3	2	12	78320
10	3	0	0	0	0	0	3	0	1	1	0	0	8	78772
11	0	0	1	1	2	2	3	0	1	2	2	3	17	80338
12	3	3	2	2	0	1	1	2	2	1	1	4	22	81482
13	3	1	2	1	1	2	0	2	2	2	1	0	17	82740
14	3	2	2	1	1	1	1	3	0	4	1	2	21	85380
Total	16	12	16	13	12	12	15	15	12	15	16	19	173	

TABELA 3 – Número de habitantes e crimes violentos contra o patrimônio no município de Lavras - MG, no período de 1991 a 2004.

Ano(k)	Mês(j)												Total	w <sub>k</sub>
	1	2	3	4	5	6	7	8	9	10	11	12		
1	0	1	2	2	0	0	1	0	1	1	5	0	13	65893
2	1	2	3	1	0	2	2	1	3	1	0	6	22	67233
3	3	3	2	2	1	0	1	6	3	2	0	1	24	68517
4	2	0	0	1	0	1	3	2	5	3	0	4	21	69704
5	1	4	2	2	2	4	3	1	3	1	1	1	25	70871
6	0	1	1	3	1	0	1	1	0	3	1	1	13	72947
7	2	2	6	2	2	0	3	4	8	6	1	2	38	74945
8	1	3	3	6	8	4	3	3	1	1	2	2	37	76633
9	8	1	4	0	4	1	4	1	1	9	5	7	45	78320
10	5	4	6	4	4	4	6	5	7	8	7	5	65	78772
11	7	6	1	6	7	5	9	3	5	7	7	5	68	80338
12	1	6	4	7	7	5	9	11	10	13	12	9	94	81482
13	10	4	11	6	4	9	11	9	7	4	11	16	102	82740
14	6	7	3	7	5	10	7	17	7	8	10	12	99	85380
Total	47	44	48	49	45	45	63	64	61	67	62	71	666	

Com os dados das Tabelas 2 e 3 foram testadas as hipóteses de que CVPES e CVPT ocorriam com igual frequência em todos os meses do ano. Para tal, utilizou-se o teste de aderência  $\chi^2$  (Bussab & Morettin, 2003) com nível de significância ( $\alpha$ ) de 5% descrita pela expressão 28 (adaptação da fórmula 27):

$$\chi_c^2 = \sum_{j=1}^{12} \frac{\left( \sum_{k=1}^{12} x_{jk} - \left( \sum_{j=1}^{12} \sum_{k=1}^{12} x_{jk} / j \right) \right)^2}{\sum_{j=1}^{12} \sum_{k=1}^{12} x_{jk} / j} \quad (28)$$

A hipótese nula ( $H_0$ ) foi testada contra a hipótese alternativa ( $H_1$ ), enunciadas da seguinte forma:

$$H_0 : p_1 = p_2 = \dots = p_{12} = \frac{1}{12},$$

$$H_1 : p_j \neq \frac{1}{12}, \text{ para pelo menos um } j.$$

TABELA 4 Resultados dos testes de aderência dos totais de CVPES e CVPT entre os meses de janeiro e dezembro, no período de 1991 a 2004, em Lavras - MG

Teste de aderência	Estatística $\chi^2$	gl*	p-valor
CVPES	3,8092	11	0,9752
CVPT	20,3114	11	0,0413

\*gl = graus de liberdade

A partir dos resultados apresentados na Tabela 4, observar-se que, no caso do CVPES, a hipótese de nulidade não foi rejeitada, uma vez que o p-valor do teste foi maior que  $\alpha$ . Ou seja, com nível de significância de 5%, pode-se afirmar que o número de crimes violentos contra a pessoa ocorre com igual frequência em todos os meses do ano.

Já no caso do CVPT, a hipótese de nulidade foi rejeitada. Então, foram testadas, ainda, mais duas hipóteses: que o CVPT ocorria com igual frequência nos meses de janeiro a junho do ano (1º semestre) e de julho a dezembro (2º semestre), ou seja, no primeiro caso utilizando a expressão 27 com  $1 \leq j \leq 6$  e hipótese  $H_0 : p_1 = p_2 = \dots = p_6 = \frac{1}{6}$ , e, no segundo, com  $7 \leq j \leq 12$ .

Os resultados desses testes estão apresentados na Tabela 5.

TABELA 5 - Resultados dos testes de aderência dos totais de crimes violentos contra o patrimônio a cada semestre, no período de 1991 a 2004, em Lavras - MG

Teste	Estatística $\chi^2$	gl*	p-valor
CVPT de janeiro a junho (1º semestre)	0,4173	5	0,9948
CVPT de julho a dezembro (2º semestre)	1,0154	5	0,9613

\*gl = graus de liberdade

De acordo com as Tabelas 3, 4 e 5, pode-se afirmar, com um nível de significância de 5%, que o número de crimes violentos contra o patrimônio tem comportamento diferenciado de acordo com o semestre do ano.

#### 4.2.2 Regressão linear simples

##### 4.2.2.1 Modelo estatístico e diagnóstico das pressuposições básicas

Diante dos resultados das seções anteriores, foram ajustados modelos de regressão linear simples do tipo  $Y_k = \beta_0 + \beta_1 k + \varepsilon_k$  ( $1 \leq k \leq 14$ ), descrito na seção 3.3.1. A variável resposta ( $Y_k$ ) assumiu uma simbologia diferente para cada um dos quatro tipos de dados:

- I) Número de crimes violentos anual contra a pessoa (CVPES<sub>k</sub>);
- II) Número de crimes violentos contra o patrimônio a cada semestre (CVPT<sub>1k</sub> e CVPT<sub>2k</sub>);
- III) Índice de criminalidade violenta anual contra a pessoa (ICVPES<sub>k</sub>), calculado de acordo com a expressão 5, adaptada da seguinte forma:

$$ICVPES_k = \frac{\sum_{j=1}^{12} xps_{jk} \cdot 1000}{w_k}$$

IV) Índice de criminalidade violenta contra o patrimônio a cada semestre (ICVPT<sub>1k</sub> e ICVPT<sub>2k</sub>), calculados também de acordo com a expressão 5 e adaptados a seguir:

$$\text{ICVPT}_{1k} = \frac{\sum_{j=1}^6 \text{xpt}_{jk} \cdot 1000}{w_k} \quad \text{e} \quad \text{ICVPT}_{2k} = \frac{\sum_{j=7}^{12} \text{xpt}_{jk} \cdot 1000}{w_k}$$

Convém lembrar que as justificativas para terem sido feitas as análises tanto com o número de crimes violentos quanto com os seus índices já foram explicitadas na seção 2.1.4.

Portanto, foram calculadas, para cada um dos quatro tipos de dados, as estimativas dos parâmetros  $\beta_0$  e  $\beta_1$ , com os seus respectivos p-valores, ou seja, simultaneamente, com a estimação, com nível de significância de 5%, a hipótese nula ( $H_0$ ) será testada contra a hipótese alternativa ( $H_1$ ) enunciadas da seguinte forma:

$$\begin{array}{ll} H_0 : \beta_0 = 0 & \text{e} \quad H_0 : \beta_1 = 0 \\ H_1 : \beta_0 \neq 0 & \quad H_1 : \beta_1 \neq 0 \end{array}$$

Foram calculados, também, as estimativas da variância residual ( $\hat{\sigma}^2 = s^2$ ) e os coeficientes de determinação ajustado ( $R^2$  ajust.). Além disso foram realizados, com nível de significância de 5%, os testes de Shapiro–Wilk (para verificar a pressuposição de normalidade dos resíduos) e de Durbin-Watson (com a hipótese de nulidade se que os erros não são correlacionados), cujos resultados estão apresentados na Tabela 6.

TABELA 6 Resumo das análises no contexto da regressão para os dados de CVPES (1), CVPT (2), ICPES (3) e ICPPT (4), no período de 1991 a 2004, no município de Lavras – MG.

D A D O S	Tipo de Análise	Modelo Estatístico	Estimativas					R <sup>2</sup> Ajust.	Testes			
			$\beta_0$	p- valor	$\beta_1$	p-valor	$\sigma^2$		Shapiro-Wilk		Durbin- Watson	
									W	p- valor	DW	p- valor
1	Anual	$CVPEs_k = \beta_0 + \beta_1 k + \epsilon_k$	4,494	0,078	1,048	0,002	17,10	0,512	0,92	0,19	2,49	0,74
2	1º sem	$CVPT_{1k} = \beta_0 + \beta_1 k + \epsilon_{1k}$	-1,374	0,655	2,831	<0,001	28,22	0,830	0,96	0,70	2,26	0,57
	2º sem	$CVPT_{2k} = \beta_0 + \beta_1 k + \epsilon_{2k}$	-4,659	0,429	4,316	<0,001	101,5	0,758	0,96	0,79	1,23	0,03
3	Anual	$ICVPES_k = \beta_0 + \beta_1 k + \epsilon_k$	0,081	0,021	0,010	0,013	0,003	0,364	0,89	0,11	2,61	0,81
4	1º sem	$ICVPT_{1k} = \beta_0 + \beta_1 k + \epsilon_{1k}$	0,012	0,276	0,032	<0,001	0,004	0,798	0,95	0,63	2,51	0,75
	2º sem	$ICVPT_{2k} = \beta_0 + \beta_1 k + \epsilon_{2k}$	-0,016	0,822	0,049	<0,001	0,015	0,729	0,96	0,79	1,35	0,05

\* Testes realizados considerando-se um nível de significância de 5%.

A interpretação do valor de  $\beta_0$  por si só, para o caso deste trabalho, não tem efeito prático, uma vez que seria o valor de CV ou ICV para o ano 0 (1990), que está fora do período estudado. Mas, obviamente, ele tem importância para o modelo, por isso, apesar dos p-valores altos, optou-se por correr o risco de se cometer o erro tipo I (é aquele que se comete ao rejeitar  $H_0$ , quando essa hipótese é correta) em detrimento de ajustar uma equação “forçosamente” pela da origem. Isso porque, entre outras, uma das consequências relevantes neste tipo de modelo é o fato de que a soma dos resíduos seria diferente de zero.

Para o valor estimado de  $\beta_1$ , a sua interpretação pode ser feita tomando-se como exemplo o modelo ajustado  $CVPEs_k = 4,494 + 1,048(k)$ , o que significa dizer que o número de crimes violentos contra a pessoa aumenta na ordem de 1,048 ao ano, a partir de 1991.

Pelos valores dos coeficientes de determinação ajustados, principalmente para os dados do tipo CVPES e ICVPES, o modelo linear parece não ser o ideal para explicar esse tipo de crime, sendo que essa discussão será retomada a seguir. Contudo, para efeito prático, tomando-se como exemplo o valor de 0,830 para o  $CVPT_1$ , pode-se dizer que 83% da variação observada, nos valores de CVPT do 1º semestre, é explicada pela variação de k (ano).

Ainda com os resultados da Tabela 6, e para todas as quatro situações, foi possível concluir, também, que não foi rejeitada a hipótese de nulidade para o teste de Shapiro-Wilk, ou seja, com nível de significância de 5%, pode-se afirmar que os resíduos seguem uma distribuição Normal.

No caso do teste de Durbin-Watson, de uma forma geral, a hipótese de nulidade não foi rejeitada, podendo-se afirmar, com nível de significância de 5%, que os resíduos não são correlacionados, ou seja, os erros são independentes. As exceções foram para o CVPT e o ICVPT do 2º semestre, em que os p-valores ficaram próximos do nível de significância de 5%, dando margem a uma possível violação dessa pressuposição. Com esses resultados,

poder-se-ia pensar em ajustar um modelo com autocorrelação nos resíduos, porém não era objetivo deste estudo, podendo ser um indicativo para trabalhos futuros, já que, em muitos casos, variáveis que são analisados ao longo do tempo, apresentam uma dependência das observações de um ano para o outro.

Com as equações ajustadas foi possível esboçar os gráficos, no mesmo plano, do ano versus os valores observados e os valores preditos para cada tipo de dado, como mostram as Figuras 4 a 7.

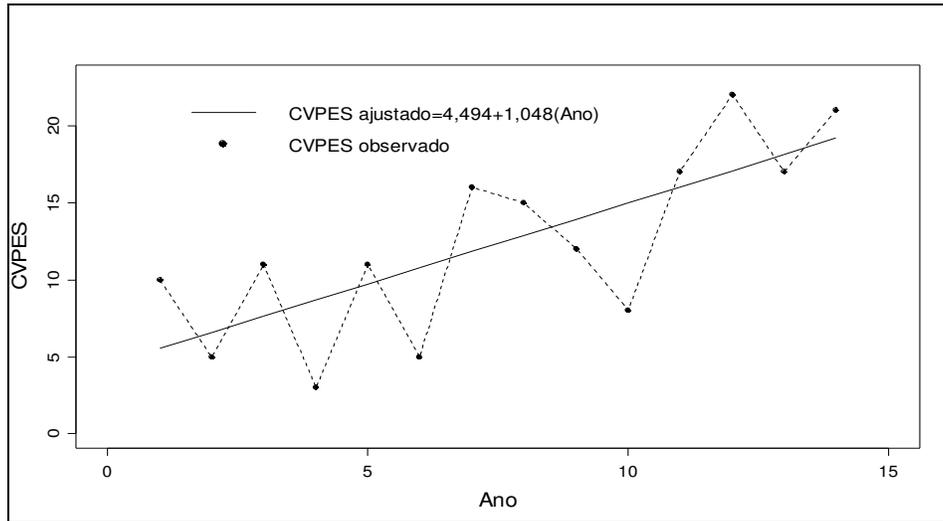


FIGURA 4 Gráfico do ajuste da equação de regressão para o número de crimes violentos contra a pessoa (CVPES), no período de 1991 a 2004, em Lavras - MG

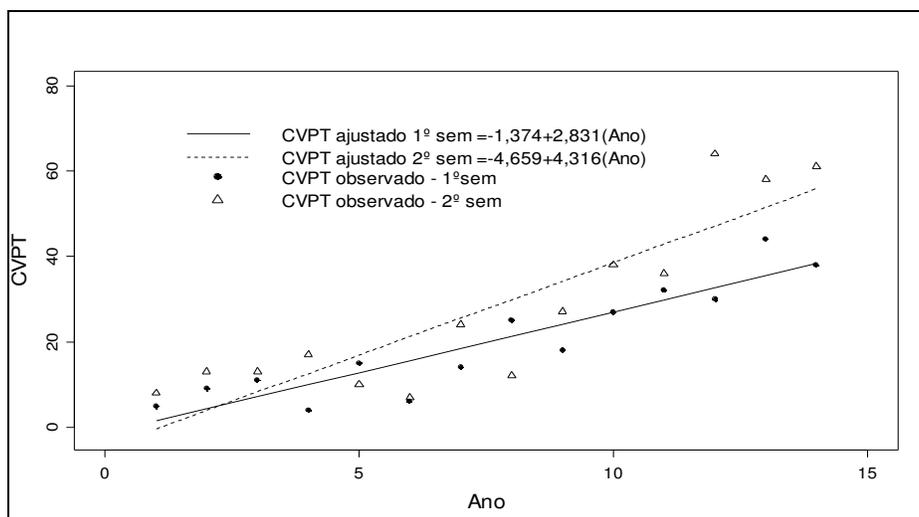


FIGURA 5 Gráfico para comparação do ajuste da equação de regressão, em cada semestre, para o número de crimes violentos contra o patrimônio (CVPT), no período de 1991 a 2004, em Lavras – MG

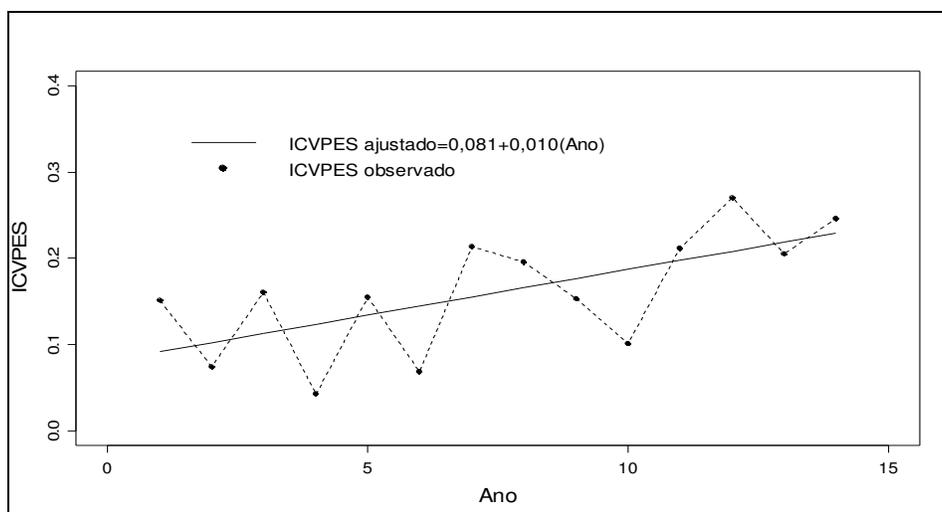


FIGURA 6 Gráfico do ajuste da equação de regressão para o índice de criminalidade violenta contra a pessoa (ICVPES), no período de 1991 a 2004, em Lavras - MG

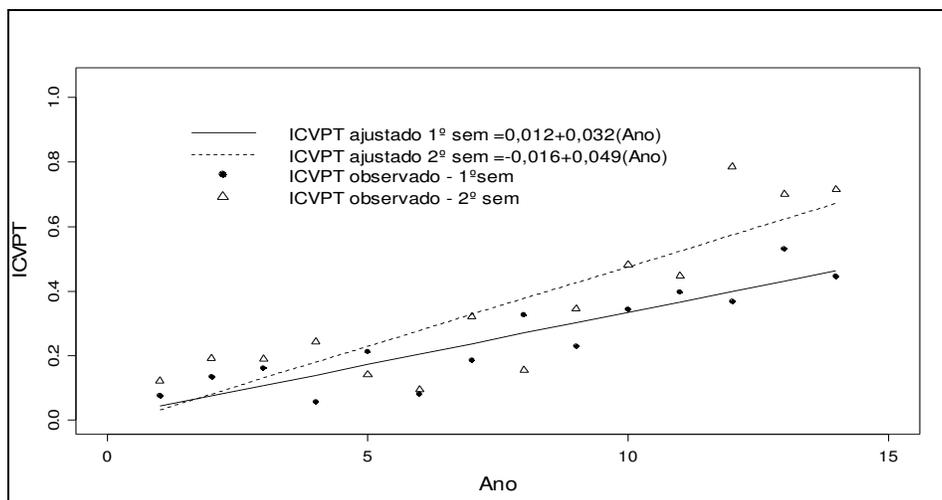
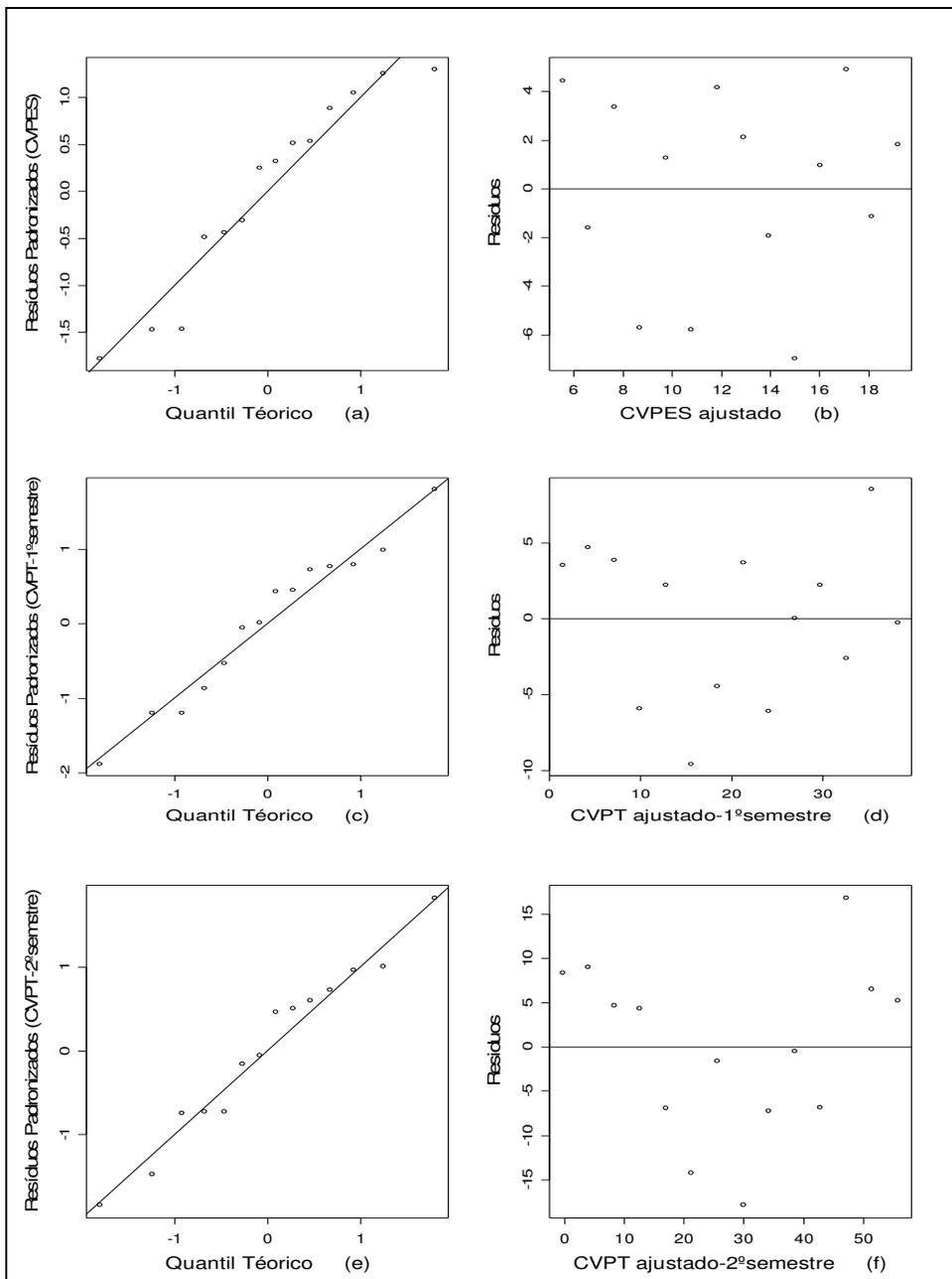


FIGURA 7 Gráfico para comparação do ajuste da equação de regressão, em cada semestre, para o índice de criminalidade violenta contra o patrimônio (ICVPT), no período de 1991 a 2004, em Lavras - MG

A partir das Figuras 4 e 6, pode-se retomar a discussão acerca da não adequação de um ajuste linear para o número de crimes violentos contra a pessoa ou o seu índice, uma vez que a linha pontilhada, que une os valores observados, indica uma possível sazonalidade para esse tipo de crime. Observa-se, também, que, na maioria das vezes, há uma alternância entre aumento e diminuição, ou seja, parece que a ação policial é mais ostensiva em um ano posterior a um aumento, o que causaria uma diminuição. Portanto, em um estudo futuro, poderia ser recomendável um ajuste dos modelos via Séries Temporais.

Já no caso das Figuras 5 e 7, observa-se uma tendência mais linear para os dados do 1º semestre que para os do 2º semestre.

Como foi descrito na seção 3.2.2.2, para o diagnóstico das pressuposições de homocedasticidade e normalidade dos erros (além do teste de Shapiro-Wilk) foram utilizados procedimentos gráficos para cada um dos tipos de dados, como mostra a Figura 8.



Continuação

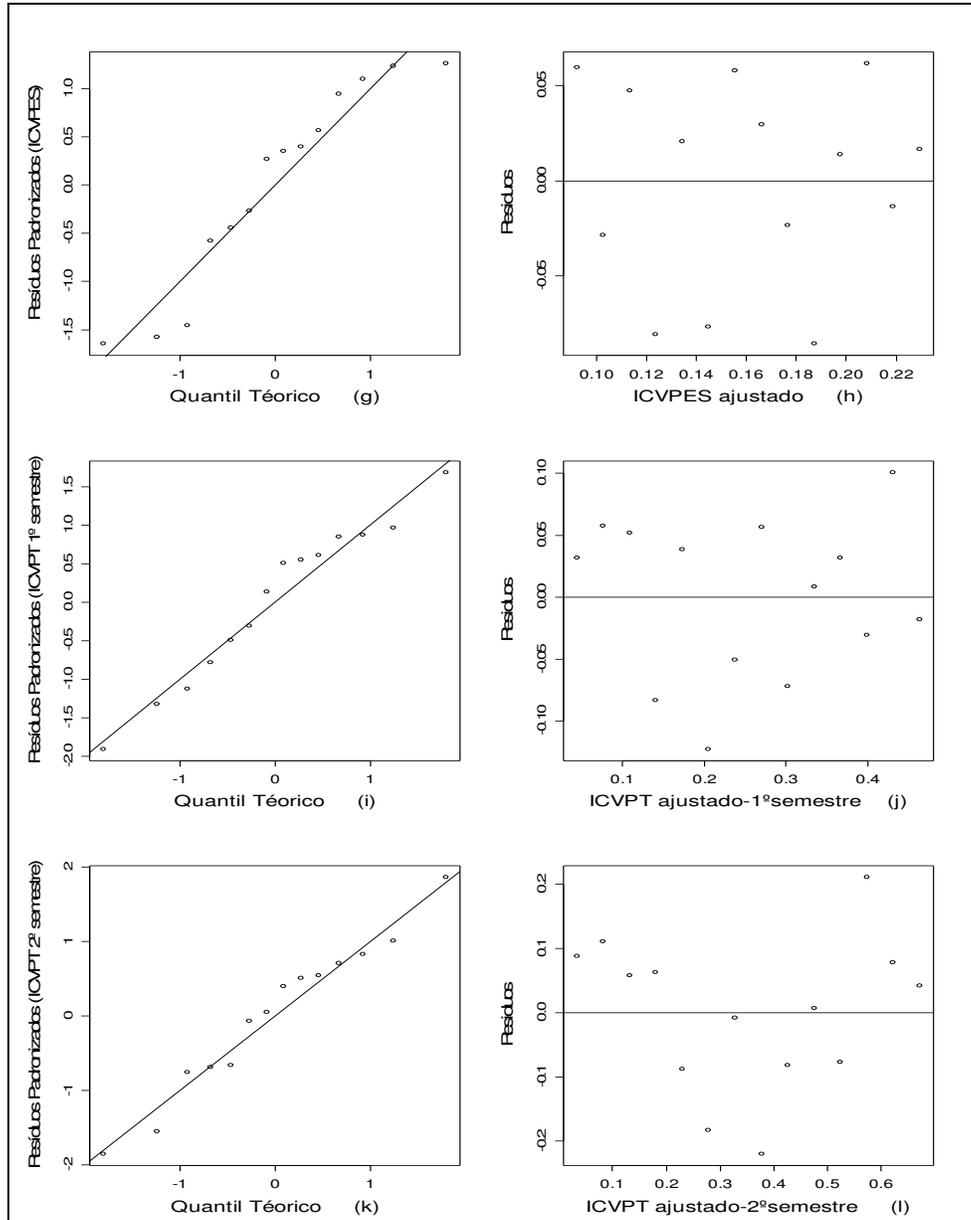


FIGURA 8 Gráficos para verificação das suposições de normalidade (a, c, e, g, i, k) e homoscedasticidade dos resíduos (b, d, f, h, j, l) dos CVPES, CVPT, ICVPES e ICVPT, no período de 1991 a 2004

Analisando-se a Figura 8, verifica-se que, no geral, as pressuposições de normalidade e homocedasticidade dos resíduos não são violadas, com exceção dos gráficos (f) e (l) que, visualmente, apresentam uma leve tendência a heterocedasticidade dos erros para, respectivamente, o CVPT e o ICVPT do 2º semestre.

#### **4.2.2.2 Intervalos estatísticos no contexto da regressão**

##### **4.2.2.2.1 Construção**

Após as análises apresentadas nas seções anteriores, foram construídos quatro tipos de intervalos estatísticos, com as expressões descritas na seção 2.5.4:

- Intervalo de confiança bilateral para  $E(Y_h)$  ;
- Intervalo de predição bilateral para uma observação futura;
- Limites de tolerância unilateral para conter uma proporção de 95% da população;
- Intervalo de tolerância bilateral para conter pelo menos uma proporção de 95% da população (procedimento não simultâneo de Wallis, 1951).

Foram calculados os limites inferiores e superiores de cada um dos intervalos estatísticos, para cada uma dos tipos de dados, com uma projeção de 20 anos, contados a partir de 1991.

Os valores aproximados desses limites estão apresentados nas Tabelas 4B a 9B (Anexo B), sendo que os mesmos foram esboçados em um mesmo plano, para cada um dos tipos de dados, conforme representado nas Figuras 9 a 12.

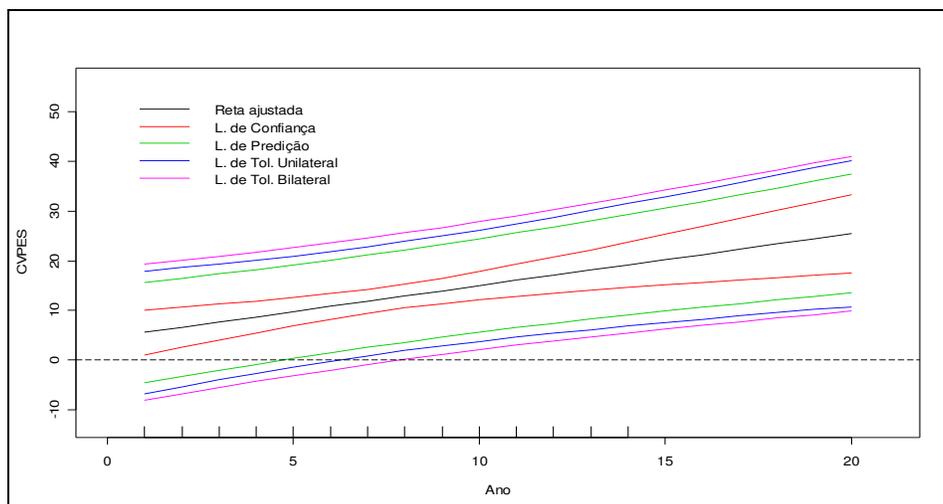


FIGURA 9 Comparação entre os quatro tipos de intervalos estatísticos para número de crimes violentos contra a pessoa (CVPES) anual, com uma projeção de 20 anos contados a partir de 1991

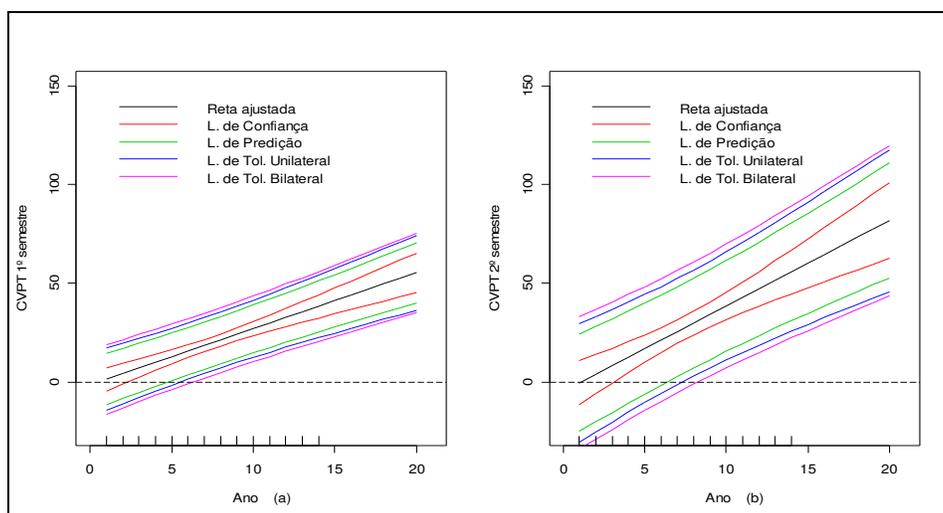


FIGURA 10 Comparação entre os quatro tipos de intervalos estatísticos para o número de crimes violentos contra o patrimônio (CVPT) do 1º semestre (a) e do 2º semestre (b), com uma projeção de 20 anos contados a partir de 1991

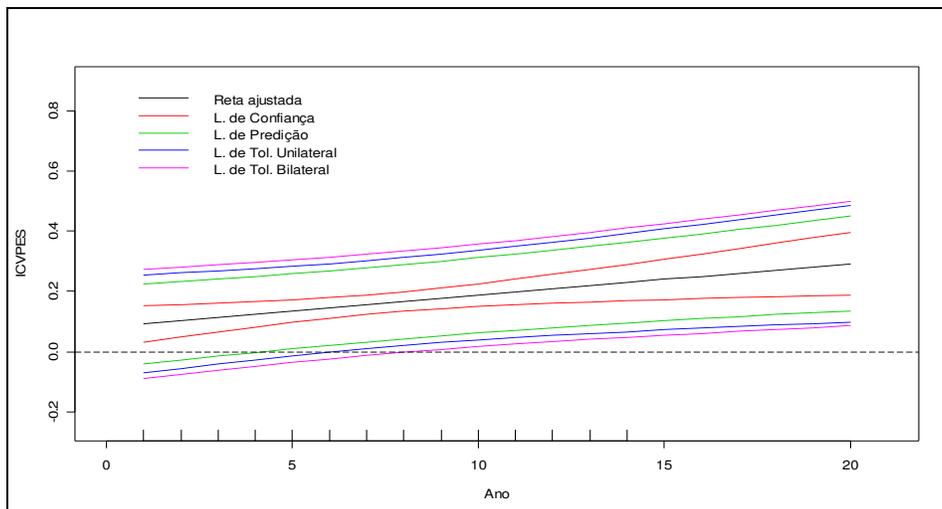


FIGURA 11 Comparação entre os quatro tipos de intervalos estatísticos para o índice de criminalidade violenta contra a pessoa (ICVPES) anual, com uma projeção de 20 anos contados a partir de 1991

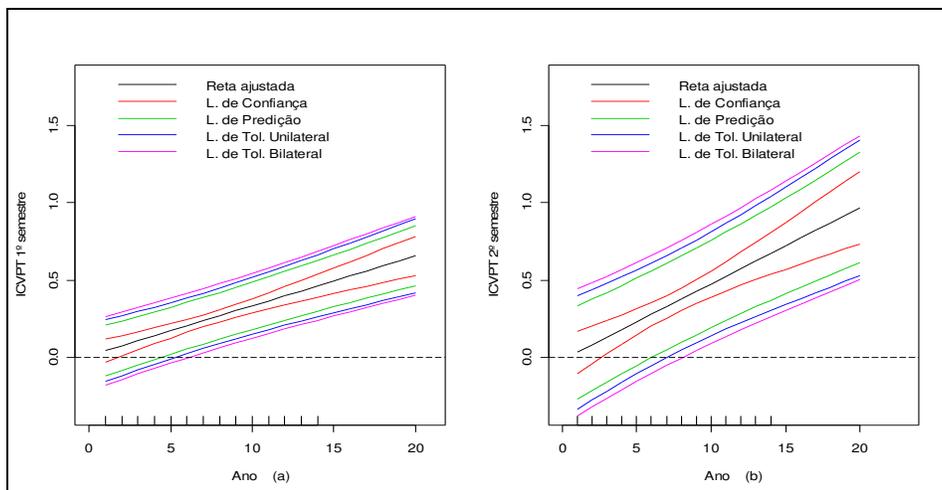


FIGURA 12 Comparação entre os quatro tipos de intervalos estatísticos para o índice de criminalidade violenta contra o patrimônio (ICVPT) do 1º semestre (a) e do 2º semestre (b), com uma projeção de 20 anos contados a partir de 1991

De acordo com as Figuras de 9 a 12 pode - se observar que o intervalo com maior comprimento é sempre o de Tolerância bilateral e com o menor comprimento de Confiança. Tais resultados estão de acordo com a literatura, como já citados nas seções 2.3.2.2 e 2.3.3.2. Isso implica dizer que os intervalos de predição têm comprimentos mais longos quando comparados aos de confiança porque existem duas fontes de variações envolvidas na construção deles: parâmetros desconhecidos precisam ser estimados usando dados passados; uma amostra futura traz uma variação aleatória adicional.

Já os intervalos de tolerância são ainda mais longos porque a proporção atual da população contida dentro destes é desconhecida, porque depende de parâmetros também desconhecidos.

#### **4.2.2.2.2 Interpretação**

Tomando-se como exemplo o ano 15, os valores dos limites superiores e inferiores dos intervalos estatísticos, para cada uma dos tipos de dados, estão apresentados na Tabela 7.

Considerando - se os resultados da Tabela 7 e tomando - se como exemplo o ICVPT, pode-se interpretar cada intervalo de acordo com o Quadro 3.

TABELA 7 Limites inferiores e superiores dos intervalos de Confiança (LC), Predição (LP), Tolerância Unilateral (LTU) e Bilateral (LTB) para o ano de 2005 (h = 15) do CVPES, CVPT, ICPES e ICVPT (1991-2004)

Dados	Tipo de Análise	Valor predito Ano h=15	Intervalos Estatísticos*							
			L.C.		L.P.		L.T.U.		L.T.B.	
			Inf.	Sup.	Inf.	Sup.	Inf.	Sup.	Inf.	Sup.
CVPES	Anual	20,22	15,13	25,31	9,87	30,57	7,55	32,89	6,23	34,21
CVPT	1º sem	41,09	34,55	47,62	27,80	54,38	24,81	57,37	23,11	59,07
	2º sem	60,09	47,70	72,48	34,88	85,30	29,21	90,96	25,99	94,19
ICVPES	Anual	0,24	0,17	0,31	0,11	0,38	0,05	0,43	0,07	0,50
ICVPT	1º sem	0,50	0,41	0,58	0,33	0,66	0,29	0,70	0,27	0,72
	2º sem	0,72	0,57	0,87	0,41	1,03	0,34	1,10	0,30	1,14

\* Intervalos estatísticos com grau de confiança de 95%.

QUADRO 3 Interpretação dos intervalos de Confiança, Predição, Tolerância Unilateral e Bilateral para o ano de 2005 (h = 15) do ICVPT (1991-2004)

Tipo de Intervalo		Interpretação
Intervalo de Confiança Bilateral para a média		Pode-se afirmar, com 95% de confiança, que o ICVPT médio no 1º semestre do ano de 2005 para cidades com o mesmo porte populacional de Lavras - MG está entre 0,41 e 0,58; no 2º semestre, entre 0,57 e 0,87.
Intervalo de Predição Bilateral		Pode-se afirmar, com 95% de confiança, que o valor do ICVPT predito para Lavras no 1º semestre do ano de 2005 está entre 0,33 e 0,66; no 2º semestre, entre 0,41 e 1,03.
Limites de Tolerância Unilateral	Inferior	Pode-se afirmar, com 95% de confiança, que 95% dos municípios com o mesmo porte populacional de Lavras no 1º semestre do ano de 2005 têm ICVPT acima de 0,29 e no 2º semestre, acima de 0,34.
	Superior	A diferença para o caso inferior é que, em vez de ser ICVPT acima, é o ICVPT abaixo. Então, no 1º semestre, é abaixo de 0,70 e no 2º, abaixo de 1,10.
Intervalo de Tolerância Bilateral		Pode-se afirmar, com 95% de confiança, que o intervalo para conter o ICV de pelo menos 95% dos municípios com o mesmo porte populacional de Lavras no 1º semestre do ano de 2005 está entre 0,27 e 0,72; já no 2º semestre, entre 0,30 e 1,14.

Os resultados dos intervalos estatísticos, no contexto da regressão, podem auxiliar no planejamento estratégico da Polícia Militar, por exemplo, se os valores observados para qualquer um dos tipos de crimes violentos ultrapassarem os limites de tolerância bilateral (que são aqueles de maior comprimento). Essa informação estatística pode servir como uma indicação de que alguma etapa do processo não está funcionando adequadamente.

Tal utilização é semelhante ao uso de cartas-controle no Controle Estatístico de Processo (CEP), tão disseminadas no uso industrial, aproximando, então, a Gestão da Segurança Pública da Gestão da Qualidade.

No entanto, deve-se ressaltar que os resultados dos intervalos deste trabalho devem se utilizados com cautela, em função da fragilidade (comensurabilidade) do banco de dados anterior a 2000, o que já foi comentado, e considerando-se que será discutido na próxima seção.

#### 4.2.2.2.3 Resultados parciais do ano de 2005

Foram fornecidos, pelo 8º batalhão da Polícia Militar – MG, os dados de CVPES e CVPT até outubro do corrente ano em Lavras - MG, sendo que estes foram transformados também em ICPES e ICVPT, respectivamente, e apresentados na Tabela 8. É bom ressaltar que a estimativa para o número de habitantes deste município, no ano de 2005, é de 86.871.

TABELA 8 Quadro resumo dos CVPES, CVPT, ICPES e ICVPT até outubro de 2005 no município de Lavras - MG.

CVPES	ICVPES	CVPT		ICVPT	
		1º sem	2º sem	1º sem	2º sem
9	0,1036	98	50	1,128	0,576

Confrontando os resultados das Tabelas 7 e 8 para o CVPT e o ICVPT do 1º semestre, verifica-se que nenhum dos intervalos estatísticos construídos inclui os valores observados. Já no 2º semestre, pelo comportamento parcial, é possível que esse panorama se reverta.

Nos casos de CVPES e ICVPEs, os valores observados parciais estão contidos, pelo menos por enquanto, nos intervalos de tolerância unilateral e bilateral.

Contudo, esses resultados um pouco discrepantes já eram quase que esperados, pois, como já foi avaliado nas seções 2.1.3 e 3.1.1, foram detectados alguns possíveis problemas de comensurabilidade no banco de dados de 1991 a 1999.

Foi proposta, então, a análise dos dados para o período de 2000 a 2004 com o intuito apenas de dar início a um estudo prospectivo, já que uma amostra de tamanho 5, sob o contexto da regressão, pode apresentar resultados não muito confiáveis.

#### **4.2.3 Distribuição de Poisson**

Com os resultados da regressão, e tomando como exemplo o número de crimes violentos contra a pessoa (CVPES), foi possível comparar a probabilidade desse tipo de crime ser maior do que o seu valor observado, considerando - se a distribuição de Poisson (calculada de acordo com a expressão 26) com a probabilidade considerando a aproximação pela Normal ( $N(\lambda, \lambda)$ ) para cada ano ( $k$ ), no período de 1991 a 2004, cujos os resultados estão apresentados na Tabela 9.

TABELA 9 Cálculo da probabilidade do  $CVPEs \geq CVPEs$  observado utilizando a distribuição de Poisson e a aproximação pela Normal, para os dados no período de 1991 a 2004 no município de Lavras - MG

Ano (k)	CVPEs Obser	$\lambda = \hat{Y}_k = CVPEs_k$	Probabilidade ( $CVPEs \geq CVPEs_{obs}$ )	
			Poisson	Normal
1	10	5,54	0,056	0,046
2	5	6,59	0,786	0,792
3	11	7,64	0,150	0,150
4	3	8,69	0,992	0,982
5	11	9,74	0,384	0,404
6	5	10,78	0,982	0,972
7	16	11,83	0,143	0,143
8	15	12,88	0,313	0,326
9	12	13,93	0,734	0,742
10	8	14,98	0,982	0,973
11	17	16,03	0,437	0,453
12	22	17,07	0,142	0,142
13	17	18,12	0,635	0,648
14	21	19,17	0,368	0,381

Com os resultados da Tabela 9, pode - se verificar que os valores da probabilidade, considerando - se a distribuição de Poisson e a aproximação pela Normal, são muito próximos, mesmo para os casos em que  $\lambda < 15$ .

Tomando - se como exemplo o ano 10 (2000) do CVPEs, é possível fazer a seguinte interpretação: existia a probabilidade de 98,2% e 97,3% considerando - se as distribuições de Poisson e Normal, respectivamente, de que o número de crimes violentos contra a pessoa na cidade de Lavras no ano de 2000 fosse maior ou igual a 8, ou seja, analisando sob esse ponto de vista, o CVPEs estaria num patamar abaixo do esperado com alta probabilidade (rever a Figura 4 para o ano de 2000).

Portanto, a avaliação pela distribuição de Poisson, ou até mesmo pela aproximação da Normal, associada com a equação de regressão pode dar um panorama geral do comportamento do número de crimes violentos.

### **4.3 Comportamento do número de crimes violentos contra a pessoa e contra o patrimônio no período de 2000 a 2004**

#### **4.3.1 Regressão linear simples**

Da mesma forma que na seção 4.2.3.1, foram ajustados modelos de regressão linear simples do tipo  $Y_k = \beta_0 + \beta_1 k + \varepsilon_k$  ( $1 \leq k \leq 5$ ), descritos na seção 3.3.1, para os mesmos tipos de dados:  $CVPEs_k$ ,  $CVPT_{1k}$ ,  $CVPT_{2k}$ ;  $ICVPES_k$ ,  $ICVPT_{1k}$  e  $ICVPT_{2k}$ .

Foram calculadas, então, para cada um desses tipos de dados, as estimativas dos parâmetros  $\beta_0$  e  $\beta_1$  (testadas sob as mesmas hipóteses) e da variância residual ( $\hat{\sigma}^2 = s^2$ ), bem como os coeficientes de determinação ajustados ( $R^2$  ajust.). Foram realizados, também, com nível de significância de 5%, os testes de Shapiro–Wilk e de Durbin-Watson; cujos resultados estão apresentados na Tabela 10.

TABELA 10 Resumo das análises no contexto da regressão para os dados de CVPES (1), CVPT (2), ICVPEs (3) e ICVPT (4), no período de 2000 a 2004, no município de Lavras – MG

D A D O S	Tipo de Análise	Modelo Estatístico	Estimativas					R <sup>2</sup> Ajust.	Testes			
			$\beta_0$	P- valor	$\beta_1$	P- valor	$\sigma^2$		Shapiro-Wilk		Durbin- Watson	
									W	P- valor	DW	P- valor
			1	Anual	$CVPEs_k = \beta_0 + \beta_1 k + \epsilon_k$	9,200	0,132		2,600	0,149	18,13	0,405
2	1º sem	$CVPT_{1k} = \beta_0 + \beta_1 k + \epsilon_{1k}$	24,000	0,017	3,400	0,111	23,07	0,501	0,94	0,66	3,36	0,91
	2º sem	$CVPT_{2k} = \beta_0 + \beta_1 k + \epsilon_{2k}$	31,000	0,048	6,800	0,099	82,93	0,534	0,92	0,51	2,83	0,68
3	Anual	$ICVPEs_k = \beta_0 + \beta_1 k + \epsilon_k$	0,122	0,118	0,028	0,194	0,002	0,307	0,93	0,58	1,89	0,17
4	1º sem	$ICVPT_{1k} = \beta_0 + \beta_1 k + \epsilon_{1k}$	0,316	0,015	0,034	0,172	0,003	0,353	0,92	0,52	0,70	0,04
	2º sem	$ICVPT_{2k} = \beta_0 + \beta_1 k + \epsilon_{2k}$	0,411	0,042	0,072	0,143	0,013	0,419	0,91	0,49	0,89	0,07

\* Testes realizados considerando – se um nível de significância de 5%.

As interpretações para os resultados das estimativas de  $\beta_0$  e  $\beta_1$  e do coeficiente de determinação ajustado da Tabela 10 são as mesmas feitas para a Tabela 6. Porém, os resultados dos testes de Shapiro-Wilk e de Durbin-Watson foram apresentados mais para efeito ilustrativo, uma vez que, como já foi dito, com uma amostra de tamanho 5, esses testes podem estar comprometidos.

Com as equações ajustadas foi possível esboçar o gráfico, no mesmo plano, do ano versus os valores observados e os valores preditos para cada situação, como mostram as Figuras 13 a 16.

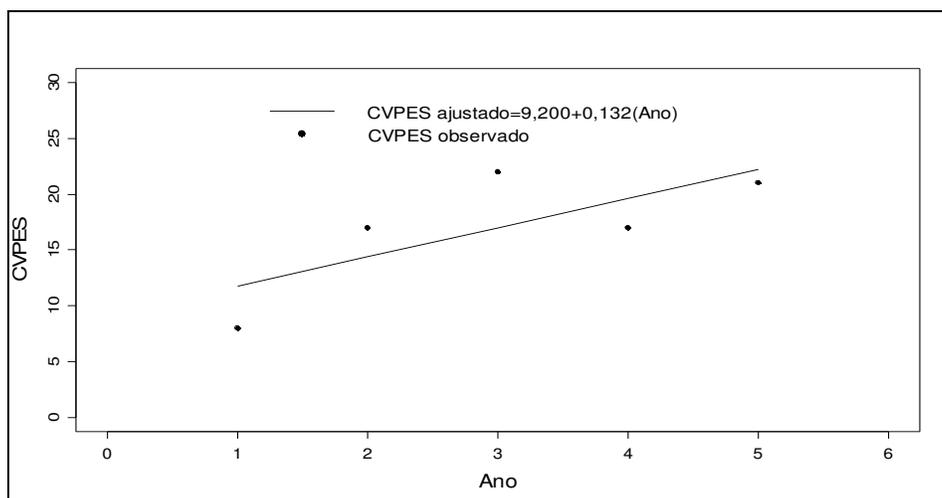


FIGURA 13 Gráfico do ajuste da equação de regressão para o número de crimes violentos contra a pessoa (CVPES), no período de 2000 a 2004, em Lavras - MG

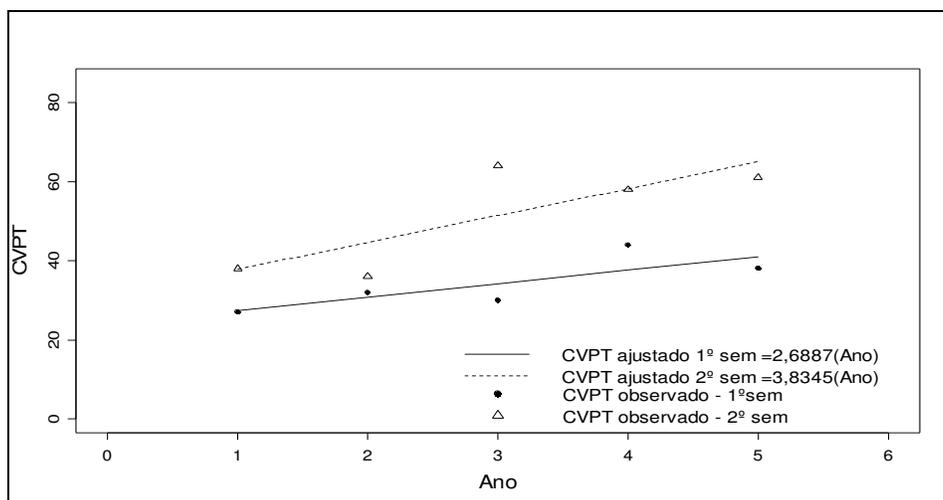


FIGURA 14 Gráfico para comparação do ajuste da equação de regressão em cada semestre para o número de crimes violentos contra o patrimônio (CVPT), no período de 2000 a 2004, em Lavras - MG

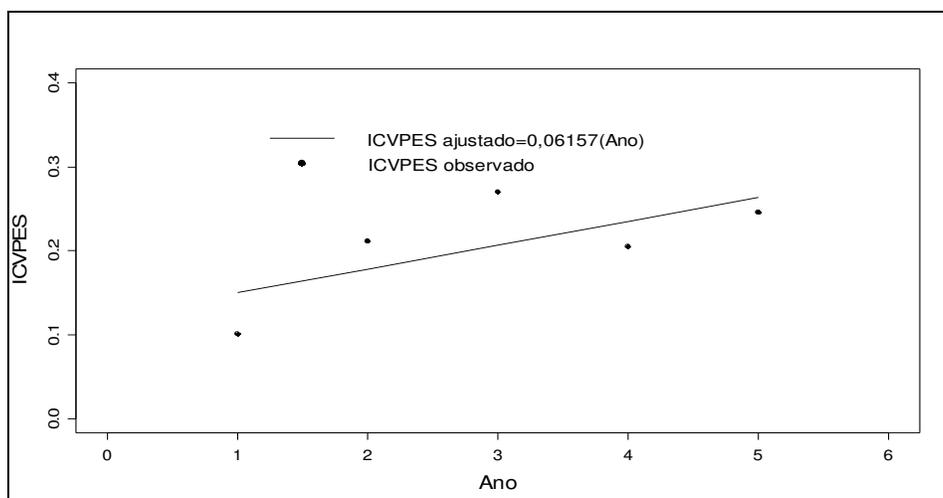


FIGURA 15 Gráfico do ajuste da equação de regressão para o índice de criminalidade violenta contra a pessoa (ICVPES), no período de 2000 a 2004, em Lavras - MG

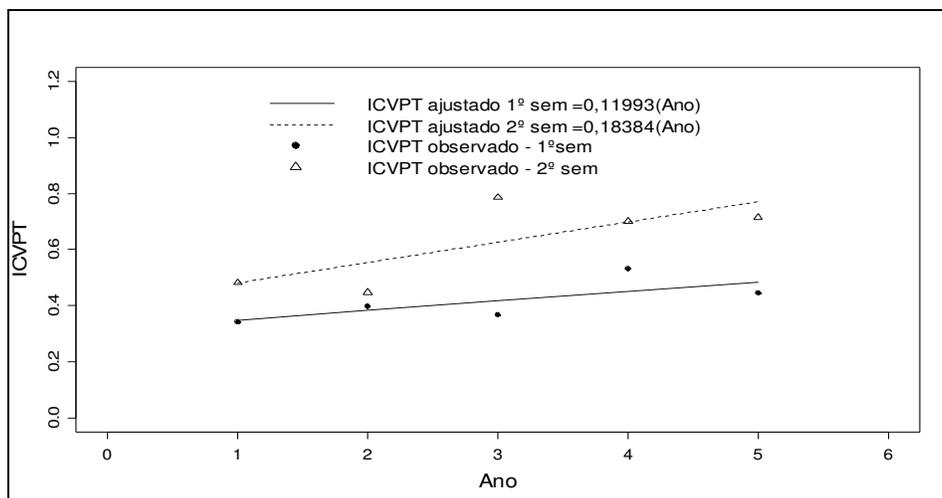


FIGURA 16 Gráfico para comparação do ajuste da equação de regressão em cada semestre para o índice de criminalidade violenta contra o patrimônio (ICVPT), no período de 2000 a 2004, em Lavras - MG

### 4.3.2 Intervalos estatísticos no contexto da regressão

#### 4.3.2.1 Construção

Da mesma forma que na seção 4.2.3.2.1 foram construídos quatro tipos de intervalos estatísticos com as fórmulas descritas na seção 2.5.4:

- Intervalo de confiança bilateral para  $E(Y_h)$  ;
- Intervalo de predição bilateral para uma observação futura;
- Limites de tolerância unilateral para conter uma proporção de 95% da população;
- Intervalo de tolerância bilateral para conter pelo menos uma proporção de 95% da população (procedimento não simultâneo de Wallis, 1951).

Foram, também, calculados os limites inferiores e superiores de cada um dos intervalos estatísticos para cada uma das situações, com uma projeção de 7 anos contados a partir de 2000. Os valores aproximados desses limites estão apresentados nas Tabelas 10B a 15B (Anexo B), e foram esboçados num mesmo plano, para cada um dos tipos de dados, conforme representado nas Figuras 1A a 4A (Anexo A).

#### **4.3.2.2 Comparação dos intervalos projetados para o ano de 2005 (h=6) com os resultados parciais.**

Os valores dos limites superiores e inferiores dos intervalos estatísticos para o ano de 2005, em cada uma das situações, estão apresentados na Tabela 11 a seguir.

TABELA 11 Limites inferiores e superiores dos intervalos de Confiança (LC), Predição (LP), Tolerância Unilateral (LTU) e Bilateral (LTB) para o ano de 2005 (h = 6) de CVPES, CVPT, ICPVES e ICVPT (2000-2004)

Dados	Tipo de Análise	Valor predito Ano h=6	Intervalos Estatísticos*							
			L.C.	L.C	L.P.	L.P.	L.T.U.	L.T.U.	L.T.B.	L.T.B.
			Inf.	Sup.	Inf.	Sup.	Inf.	Sup.	Inf.	Sup.
CVPES	Anual	24,80	10,59	39,01	5,16	44,44	0	50,56	0	56,82
CVPT	1º sem	44,40	28,37	60,43	22,25	66,55	15,35	73,45	8,28	80,52
	2º sem	71,80	41,40	102,20	29,80	113,80	16,71	126,89	3,32	140,28
ICVPES	Anual	0,292	0,113	0,471	0,044	0,539	0	0,616	0	0,695
ICVPT	1º sem	0,519	0,319	0,719	0,242	0,795	0,156	0,881	0,068	0,969
	2º sem	0,841	0,457	1,225	0,311	1,372	0,145	1,537	0	1,707

\* Intervalos estatísticos com grau de confiança de 95%.

Comparando os resultados da Tabela 11 com os dados da Tabela 8 verifica-se uma adequação melhor do que na avaliação com a Tabela 7. No caso do CVPES e dos ICVPEs, os valores observados parciais estão contidos nos intervalos, com exceção do intervalo de confiança. Para o CVPT e o ICVPT no 1º semestre, os valores observados continuam fora dos intervalos, mas para o 2º semestre é possível que a abrangência seja total.

Contudo, convém salientar, que os estudos para esse período (2000 a 2004) são preliminares e ilustrativos, e que uma série maior de tempo deverá ser analisada para se observar o comportamento dos crimes.

Após os estudos no contexto da regressão, partiu-se para uma avaliação do comportamento dos intervalos estatísticos para um estrato de uma população com distribuição Normal.

#### **4.4 Inferência sobre a amostra de uma população Normal**

Após o levantamento do número de municípios brasileiros (413) que tinham população entre 14000 a 17500 habitantes no ano de 2003, foi testada a suposição de normalidade do ICV dos 50 municípios mineiros nesta mesma faixa populacional, por meio do teste de Shapiro - Wilk. Com nível de significância de 5%, a hipótese de nulidade não foi rejeitada, uma vez que a estatística W foi igual a 0,9714 e o p-valor igual a 0,2631.

Apesar de ter sido aceita a hipótese de normalidade dos dados, pode-se verificar, pela Figura 17, que, provavelmente, existem populações diferentes reunidas por grupos. Portanto, para um estudo futuro mais aprofundado, seria interessante estratificar esses municípios, possivelmente por regiões, de onde seriam retiradas sub-amostras de cada estrato.

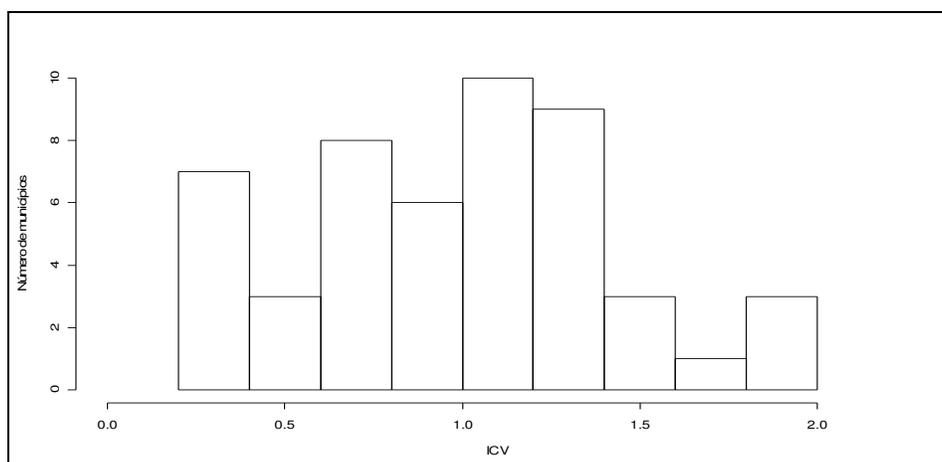


FIGURA 17 Histograma do ICV versus o número de municípios com população entre 14000 a 17500, no ano de 2003, em Minas Gerais

Mas, para efeito didático, foi admitida a suposição de normalidade, inclusive com extrapolação dos resultados para a população dos 413 municípios brasileiros. Isso porque não se pode afirmar, com esses estudos preliminares, se Minas Gerais é ou não uma amostra representativa do Brasil no tangente à questão dos crimes violentos. Então, dos 50 municípios mineiros foi retirada uma amostra aleatória de tamanho  $n = 30$  (Tabela 16B – Anexo B) com  $\bar{x} = 1,053$  e  $s = 0,397$ .

A partir desses dados foram construídos quatro tipos de intervalos estatísticos bilaterais, considerando – se a distribuição populacional Normal, com as fórmulas descritas na seção 2.4 e grau de confiança  $(1-\alpha)$  pré-estabelecido de 0,95. Os resultados, bem como as interpretações, estão apresentados, no Quadro4.

QUADRO 4 Intervalos estatísticos bilaterais para amostra de tamanho 30 dos municípios mineiros no ano de 2003 com número de habitantes entre 14000 a 17500

Tipos de Intervalos	Resultado e Interpretação
Confiança para média populacional	Com 95% de confiança, o intervalo para conter o ICV médio dos municípios com mesmo porte populacional, no ano de 2003, é de 0,810 a 1,123.
Predição para a média de 30 futuras observações	Com 95% de confiança, o intervalo de predição para conter o ICV médio de 30 observações futuras de municípios com mesmo porte populacional está entre 0,749 e 1,289.
Predição para uma observação futura	Com 95% de confiança, o intervalo para conter o ICV de uma observação futura, ou seja, de um município com mesmo porte populacional no ano de 2003 é de 0,059 a 1,979.
Tolerância para conter pelo menos 95% da população – Aproximação de Wald e Wolfowitz (1946)	Com 95% de confiança, o intervalo para conter o ICV de pelo menos 95% dos municípios com mesmo porte populacional, no ano de 2003, é de 0 a 2,251.

No caso do Intervalo de Tolerância para determinar os valores de  $k$  e  $\gamma$  da expressão 22 pela aproximação de Wald & Wolfowitz (1946) foi utilizado o processo iterativo de tentativa e erro, apoiado computacionalmente, e os valores encontrados foram, respectivamente, iguais a 2,008027 e 0,9500027.

Observa-se que o intervalo de tolerância (0 a 2,251) é mais amplo e pode ser entendido como o intervalo de predição para uma observação futura (0,059 a 1,979) corrigido pela incerteza da média dada pelo intervalo de confiança (0,810 a 1,123), conforme apresentado na seção 2.4.7.

Os resultados dos Intervalos Estatísticos construídos no contexto de um estrato têm uma importância prática menor do que no contexto da regressão, pois as interpretações são limitadas para o ano de 2003 sob as mesmas condições, com exceção dos de Predição, que, mesmo assim, não levam em consideração possíveis tendências temporais. Contudo, vale ressaltar que as ferramentas estatísticas a serem utilizadas dependem muito da base de dados existente.

## 5 CONCLUSÕES

### 5.1 Conclusões sobre a análise dos dados

1. O crescimento populacional não é o fator único para explicar o aumento no número de crimes violentos. O passar dos anos mostra um aumento dos crimes, mesmo para cidades numa mesma faixa populacional.
2. O número de crimes violentos contra o Patrimônio no município de Lavras – MG apresenta comportamento diferenciado para o 1º e para o 2º semestres.
3. Para o período de 1991 a 2004 verifica-se que o índice de criminalidade violenta contra a Pessoa aumenta na ordem de 0,021 ao ano, a partir de 1991 ( $k=1$ ). Já para o período de 2000 a 2004, o aumento é da ordem de 0,0118 ao ano.
4. Para o período de 1991 a 2004 pode-se afirmar, com 95% de confiança, que o intervalo de tolerância para conter o ICVPT de pelo menos 95% dos municípios com mesmo porte populacional de Lavras, no 1º semestre do ano de 2005 está entre 0,27 e 0,72. Já para o período de 2000 a 2004 está entre 0,068 e 0,969.
5. O número de crimes violentos contra o patrimônio, no 1º semestre de 2005, no município de Lavras – MG foi igual a 98, não estando incluído em nenhum dos intervalos estatísticos construídos.

## 5.2 Conclusões sobre metodologias estatísticas

1 As aproximações da distribuição de Poisson pela Normal para o cálculo de probabilidade, mesmo para as situações em que o  $\lambda < 15$ , foram satisfatórias, sendo que a maior diferença foi da ordem de 2%.

2 As análises feitas no contexto de um estrato (modelo simples  $Y = \mu + \varepsilon$ ) são menos informativas do que as feitas no contexto da regressão.

3 Na comparação desenvolvida entre os intervalos estatísticos, em ordem crescente do comprimento do intervalo, têm-se: o de confiança, o de predição, o de tolerância unilateral, o de tolerância bilateral; sendo este último o mais largo, porque a proporção atual da população contida dentro destes é desconhecida, uma vez que depende de parâmetros também desconhecidos. Essa seqüência é explicada pelo entendimento de que o intervalo de tolerância é a conjugação do intervalo de predição com o intervalo de confiança.

4 Os resultados dos intervalos estatísticos, construídos no contexto do modelo simples normal e da regressão, são ferramentas úteis para auxiliar a Gestão da Qualidade na Segurança Pública, no sentido de que possibilitam a projeção de valores futuros, que podem auxiliar num planejamento estratégico mais eficiente, à semelhança das cartas controle no Controle Estatístico de Processo (CEP).

5 A utilidade desta metodologia verifica-se, portanto, até mesmo quando “falha” no sentido de que uma observação mesmo fora dos limites de confiança, predição e tolerância, ainda servem para os propósitos da gestão.

6 Sugere-se o estabelecimento de uma linha de pesquisa para a PMMG, uma vez que os ajustes feitos com os dados disponíveis acabaram tendo uma interpretação de caráter mais teórico, por problemas de comensurabilidade na base de dados.

## 6 REFERÊNCIAS BIBLIOGRÁFICAS

ABÍLIO, R. S.; ARAÚJO, F. P. A. de; SANTOS, A. C. dos; NICOLAU, L. A. **Universidade Cooperativa de Segurança Pública Módulo on line: uma proposta para o auxílio na educação continuada da Polícia Militar.** Sbsi, Natal, v. 1, n. 1, 2005.

ALVES FILHO, M. Geoprocessamento é usado no estudo da violência. **Jornal da Unicamp**, Campinas, 26 abr. 2004. Disponível em: <[http://www.unicamp.br/unicamp/unicamp\\_hoje/ju/abril2004/ju249pag12.html](http://www.unicamp.br/unicamp/unicamp_hoje/ju/abril2004/ju249pag12.html)> Acesso em 09 dez. 2005.

ARAÚJO JÚNIOR, A. F. de; FAJNZYLBBER, P. **O que causa a criminalidade violenta no brasil?** uma análise a partir do modelo econômico do crime: 1981 a 1996. Universidade Federal de Minas Gerais: Faculdade de Ciências Econômicas. Texto para discussão nº 62, Belo Horizonte, 2001. Disponível em: <<http://www.cedeplar.ufmg.br/pesquisas/td/TD%20162.pdf>>. Acesso em: 06 dez. 2005.

ARAÚJO, F. P. de; NICOLAU, L. A.; OLIVEIRA, M. S. de; SANTOS, A. C. dos. Aplicação da Estatística Básica na Gestão da Segurança Pública. In: REUNIÃO DA REGIÃO BRASILEIRA DA SOCIEDADE INTERNACIONAL DE BIOMETRIA, 50., 2005, Londrina. **Programas e resumos...** Londrina, PR: UEL, 2005. 1CD-ROM.

ASSUNÇÃO, R. M.; FAGUNDES F. F.; BEATO FILHO, C. A Spatial Model to Evaluate Impact Effects of Observation and Surveillance Police Stations in a Large Brazilian City. In: ANNUAL INTERNATIONAL CRIME MAPPING RESEARCH CONFERENCE, 1., 2001, Dallas Texas. Disponível em: <<http://www.ojp.usdoj.gov/nij/maps/dallas2001/Assuncao.doc>>. Acesso em: 4 dez. 2005.

BARANDIER, M. **A solução imediata.** JB Online, 16 abr. 2002. Disponível em: <<http://jbonline.terra.com.br/papel/opiniaio/2002/04/15/joropi20020415002.html>>. Acesso em: 10 dez. 2005.

BATITUCCI, E. C. **Ocorrências registradas pela Polícia Militar de Minas Gerais, 1991-2001**: Livro de Código CIS 0044\_1 (1991 -1 955). Belo Horizonte: Fundação João Pinheiro; Núcleo de Estudos de Segurança Pública, 2001. In: Consórcio de Informações Sociais, 2005a. Disponível em: <<http://www.nadd.prp.usp.br/cis/index.aspx>>. Acesso em: 10 out. 2005.

BATITUCCI, E. C. **Ocorrências registradas pela Polícia Militar de Minas Gerais, 1991-2001**: Livro de Código CIS 0044\_2 (1996 – 2001). Belo Horizonte: Fundação João Pinheiro; Núcleo de Estudos de Segurança Pública, 2001. In: Consórcio de Informações Sociais, 2005b. Disponível em: <<http://www.nadd.prp.usp.br/cis/index.aspx>>. Acesso em: 10 out. 2005.

BATITUCCI, E. C. **Ocorrências registradas pela Polícia Militar de Minas Gerais, 1991-2001 (Banco de dados)**. Belo Horizonte: Fundação João Pinheiro; Núcleo de Estudos de Segurança Pública, 2001. In: Consórcio de Informações Sociais, 2005c. Disponível em: <<http://www.nadd.prp.usp.br/cis/index.aspx>>. Acesso em: 10 out. 2005.

BEATO FILHO, C. C. Determinantes da criminalidade em Minas Gerais. **Revista Brasileira de Ciências Sociais**, São Paulo, v. 13, n. 37, p. 74-87, 1998a. Disponível em: [www.scielo.br](http://www.scielo.br). Acesso em: 11 nov. 2005.

BEATO FILHO, C. C. Políticas públicas de segurança: equidade, eficiência e accountability. In: SEMINÁRIO A REFORMA DA ADMINISTRAÇÃO PÚBLICA NO BRASIL: possibilidades e obstáculos, 1998b, Belo Horizonte. Fundação Joaquim Nabuco. Disponível em: <<http://www.crisp.ufmg.br/polpub.pdf>>. Acesso em: 07 dez. 2005.

BEATO FILHO, C. C. **Crime e políticas sociais na América Latina**. CRISP, Belo Horizonte, Informativo ano 0, n.1, 2001a. Disponível em: <<http://www.crisp.ufmg.br/INFO1.pdf>>. Acesso em: 07 dez. 2005.

BEATO FILHO, C. C. **Reinventar a polícia**: a implementação de um Programa de Policiamento Comunitário. CRISP, Belo Horizonte, Informativo ano 0, n.2, 2001b. Disponível em: <<http://www.crisp.ufmg.br/INFO2.pdf>>. Acesso em: 08 dez. 2005.

BOLETIM DE INFORMAÇÕES CRIMINAIS. **Editorial**. Governo do Estado de Minas Gerais: Secretária de Estado de Defesa Social, Belo Horizonte, n. 01, 2004.

BOWDEN, D. C. Tolerance Intervals in regression (query). **Technometrics**, Alexandria, v.10, n.1, p. 207-209, 1968.

BRASIL. Lei 9.099, 26 /09/1995. Dispõe sobre os Juizados Especiais Cíveis e Criminais. Brasília: Presidência República, 1995. Disponível em: <[https://www.planalto.gov.br/ccivil\\_03/Leis/L9099.htm](https://www.planalto.gov.br/ccivil_03/Leis/L9099.htm)>. Acesso em: 09 dez. 2005.

BUSSAB, W. O.; MORETTIN, P. A. **Estatística básica**. 5. ed. São Paulo: Saraiva, 2003. 526 p.

CANAVOS, G. C.; KOUTROUVELIS, I. A. The robustness of two-sided tolerance limits for a Normal Distribution. **Journal of Quality Technology**, Wisconsin Avenue, v. 16, n. 3, p. 144-149, 1984.

CASTRO, M. S.; ASSUNÇÃO, R. M.; DURANTE, M. O. Comparação de Dados de Homicídios entre Dois Sistemas de Informação, Minas Gerais. **Revista de Saúde Pública**, São Paulo, v. 37, n. 2, p. 168-176, 2003.

CASTRO, M. S. M.; SILVA, B. F. A.; ASSUNÇÃO, R. M. Regionalização do Estado de Minas Gerais segundo Taxas de Criminalidade Violenta e de Mortes por Homicídios no Período de 1996 a 2000. **Cadernos de Saúde Pública**, Rio de Janeiro, v. 20, p. 109-118, 2004.

CHARNET, R.; FREIRE, C. A. de L.; CHARNET, E. M. R.; BONVINO, H. **Análise de modelos de regressão linear com aplicações**. Campinas: Editora da Unicamp, 1999. 353 p.

COHEN, J.; NAGIN, D.; WALLSTROM, G.; WASSERMAN, L. Hierarchical Bayesian Analysis of Arrest Rates. **Journal of the American Statistical Association**, Alexandria, v. 93, n. 444, p. 260-1270, 1998. Disponível em: <<http://www.amstat.org/publications/jasa/cohen.pdf>> Acesso em: 04 dez. 2005.

CORDEIRO, G. M. Cronologia de alguns conceitos e fatos importantes da Estatística. **Boletim da Associação Brasileira de Estatística**, Rio de Janeiro, n. 50, 2002. Disponível em: <[www.ime.usp.br/~abe/bolet50.pdf](http://www.ime.usp.br/~abe/bolet50.pdf)>. Acesso em: 10 nov. 2005.

DRAPPER, N. R.; SMITH, H. **Applied regression analysis**. 3. ed. New York: J. Wiley, 1998. 380 p. 706 p.

EBERHARDT, K. R.; MEE, R. W.; REEVE, C. P. Computing factors for exact two-sided tolerance limits for a Normal Distribution. **Communication in Statistics – Simulation and Computation**, New York, v.18, n.1, p. 397-413, 1989.

FARAWAY, J. J. **Linear models with R**. Flórida: Chapman & Hall, 2005. v. 6, p. 229.

FERREIRA, D. F. Análises estatísticas por meio do Sisvar para Windows versão 4.0. In: REUNIÃO DA REGIÃO BRASILEIRA DA SOCIEDADE INTERNACIONAL DE BIOMETRIA, 45., 2000, São Carlos. **Programas e resumos...** São Carlos, SP: UFSCAR, 2000. p.255-258.

FERREIRA, D. F. **Estatística básica**. Lavras: Editora UFLA, 2005. 664 p.

FERTIG, K. W.; MANN, N. R. On Constructing Prediction Intervals for Samples from a Weibull or Extreme Value Distribution. **Technometrics**, Alexandria, v. 22, n. 4, p. 567-573, 1980.

FUNDAÇÃO JOÃO PINHEIRO. **Núcleo de Estudos em Segurança Pública: Produtos**. Minas Gerais. Disponível em: [http://www.fjp.gov.br/produtos/cees/nesp/Download/ra\\_01\\_02.pdf](http://www.fjp.gov.br/produtos/cees/nesp/Download/ra_01_02.pdf). Acesso em: 20 jan. 2005.

GUERRA, M. J.; DONAIRE, D. **Estatística indutiva**. 2. ed. São Paulo: Livraria Ciência e Tecnologia Editora, 1982. 311 p.

HAHN, G. J.; CHANDRA, R. Tolerance Intervals for Poisson and Binomial Variables. **Journal of Quality Technology**, Wisconsin Avenue, v. 13, n. 2, p. 100-110, 1981.

HAHN, G. J.; MEEKER W. Q. **Statistical Intervals: A Guide for Practitioners**. New York: J. Wiley, 1991. 392 p.

HOFFMANN, V.; VIEIRA, S. **Análise de regressão: uma introdução à Econometria**. 3. ed. São Paulo: Editora Hucitec, 1998. 379 p.

HOWE, W. G. Two-sided Tolerance Limits for Normal Distribution – some improvements. **Journal of the American Statistical Association**, Alexandria, v. 64, n. 326, p. 610-620, 1969.

KATAOKA, V. Y.; SILVA, C. S. F. da; OLIVEIRA, M. S. de. Comparação do Índice de Criminalidade entre Municípios de mesmo porte populacional do Estado de Minas Gerais 2004. In: ENCONTRO MINEIRO DE ESTATÍSTICA, 6., 2004, Belo Horizonte. **Anais....** Belo Horizonte: UFMG, 2004a.

KATAOKA, V. Y.; SILVA, C. S. F. da; OLIVEIRA, M. S. de; PEDROZO, D. O. Análise estatística da influência do índice de desenvolvimento humano municipal no índice de criminalidade do estado de Minas Gerais no ano de 2000. In: CONGRESSO DA APG, 13., 2004, Lavras. **Resumos....** Lavras: UFLA, 2004b. 1CD-ROM.

KENDALL, M. G.; STUART, A. **The advanced theory of statistics**. 4. ed. New York: Macmillan Publishing Company, 1979. v. 2, 748 p.

LIEBERMAN, G. J. Prediction regions for several predictions from a single regression line. **Technometrics**, Alexandria, v. 3, n. 1, p. 21-27, 1961.

LIEBERMAN, G. J.; MILLER JÚNIOR, R. G. Simultaneous tolerance intervals in regression. **Biometrika**, London, v. 50, n. 1/2, p. 155-168, 1963.

LIMAM, M. M. T.; THOMAS, D. R. Simultaneous tolerance intervals for the linear regression model. **Journal of the American Statistical Association**, Alexandria, v. 83, n. 403, p. 885-888, 1988.

LÔBO, I. **Desigualdade social é a grande causa da violência entre jovens, afirma pesquisa do Ipea**. Agência Brasil, 2004. Disponível em: <[http://www.radiobras.gov.br/materia\\_i\\_2004.php?materia=193123&q=1](http://www.radiobras.gov.br/materia_i_2004.php?materia=193123&q=1)>. Acesso em: 10 dez. 2005.

MALTZ, M. From Poisson to the Present: Applying Operations Research to Problems of Crime and Justice. **Journal of Quantitative Criminology**, New York, v. 12, n.1, p 1-61, 1996. Disponível em: <<http://www.tigger.uic.edu/~mikem/Poisson.PDF>>. Acesso em: 04 dez. 2005.

MARQUES, H. **Seqüestro vira álibi de marido fujão**. JB Online, 28 maio 2002. Disponível em: <<http://jbonline.terra.com.br/jb/papel/brasil/2002/05/27/jorbra20020527001.html>>. Acesso em: 10 dez. 2005.

MILLER JÚNIOR, R. G. **Simultaneous statistical inference**. 2. ed. New York: Springer, 1981. 299 p.

MINISTÉRIO DA JUSTIÇA. **Ministério da Justiça lança software livre Terra Crime para mapear áreas de incidência criminal.** Brasília, 2003. Disponível em: [http://www.mj.gov.br/noticias/2003/dezembro/RLS151203-terra\\_crime.htm](http://www.mj.gov.br/noticias/2003/dezembro/RLS151203-terra_crime.htm). Acesso em: 17 fev. 2005.

MOOD, A. M.; GRAYBILL, F. A.; BOES, D. C. **Introduction to the theory of statistics.** 3. ed. Singapore: McGraw-Hill International, 1974. 480 p.

MORETIN, L. G. **Estatística básica: inferência.** São Paulo: Pearson Education do Brasil, 2000. v. 2, p. 182.

MUKHERJEE, S.; CARCACH, C. Repeat Victimisation in Australia. Research and Public Policy Series. **Australian Institute of Criminology**, Canberra, n. 15, p. 1-40, 1998. Disponível em: <http://www.aic.gov.au/publications/rpp/15/RPP15.pdf>. Acesso em: 04 dez. 2005.

MURTEIRA, B. J. F. de. **Probabilidade e estatística.** 2. ed. Torres Vedras: McGraw-Hill, 1990. v. 2, 480 p.

NETER, J.; WASSERMAN, W.; KUTNER, M. H. **Applied linear statistical models.** 3. ed. Homewood: R. D. Irwin, 1990. 1184 p.

ODEH, R. E. Tables of two-sided tolerance factors for a Normal Distribution. **Communications in Statistics B**, New York, v. 7, n. 2, p. 183-201, 1978.

OLIVEIRA, A. Os desafios impostos pela violência e criminalidade são inumeráveis. Governo, Estados e empresas devem se unir em prol da segurança. **Revista Security**, São Paulo, n. 39, p. 16- , 2005(?). Disponível em: [http://www.cipanet.com.br/materia\\_capa.asp?id=5&n=39](http://www.cipanet.com.br/materia_capa.asp?id=5&n=39). Acesso em: 02 nov. 2005

PATEL, J. K. Tolerance intervals - a review. **Communications in Statistics – Theory and Methods**, New York, v. 15, n. 9, p. 2719-2762, 1986.

PATEL, J. K. Prediction intervals - a review. **Communications in Statistics – Theory and Methods**, New York, v. 18, n. 7, p. 2393-2465, 1989.

POLÍCIA MILITAR DE MINAS GERAIS. **Memorando Nr 34.598.4 /2001/EMPM.** Belo Horizonte, 27 de dez. 2001.

POLÍCIA MILITAR DE MINAS GERAIS. **Diretriz para produção de serviços de segurança pública nº 01/ 2.002**. Belo Horizonte, p. 205, 2002.

R Development Core Team (2005). **R: A language and environment for statistical computing**. R Foundation for Statistical Computing, Vienna, Áustria. ISBN 3-900051-07-0. Disponível em: <<http://www.R-project.org>>. Acesso em: 2005.

RIBAS, F.; COSTA, A. C. **Boletim**: Rádio pela Infância. UNICEF, n. 68, 2004. Disponível em: <[http://www.andi.org.br/noticias/templates/boletins/template\\_radiopi.asp?articleid=4650&zoneid=25](http://www.andi.org.br/noticias/templates/boletins/template_radiopi.asp?articleid=4650&zoneid=25)>. Acesso em: 10 dez. 2005.

SECRETARIA DE ESTADO DA SEGURANÇA PÚBLICA DO GOVERNO DE SERGIPE (SSPSE). **Violência no Brasil custa 10 por cento do Produto Interno Bruto**. Sergipe, 2003. Disponível em: <<http://www.ssp.se.gov.br/Homepages/hppadraosp.nsf/0/b24919aeb9644b3c03256eb50057fe55?OpenDocument>>. Acesso em: 18 ago. 2004.

SILVA, L. da. **Conselhos Comunitários de Defesa Social**. Originalmente publicado em: 20 Experiências de Gestão Pública e Cidadania, São Paulo, 2002. Disponível em: <<http://inovando.fgvsp.br/conteudo/documentos/20experiencias2002/2%20-%20conselhos%20comunitarios%20defesa%20social.pdf>>. Acesso em: 20 nov. 2005.

SOUZA, R. R.; GOUVÊA, R. N. **A Importância dos Conselhos Comunitários de segurança - CONSEGs - no contextos da segurança pública e no fortalecimento dos Direitos Humanos e da Cidadania**. Enciclopédia Digital Direitos Humanos II :DHnet - Rede de Direitos Humanos & Cultura, 1995(?). Disponível em: <[http://www.dhnet.org.br/oficinas/goias/importancia\\_conselho\\_comu.html](http://www.dhnet.org.br/oficinas/goias/importancia_conselho_comu.html)>. Acesso em: 22 nov. 2005.

TEXEIRA, J. M.; SAFÁDI, T. Uso de Séries Temporais na análise de dados de Segurança Pública no município de Ijací. In: CONGRESSO DA APG, 13., 2004, Lavras. **Resumos...** Lavras: UFLA, 2004. 1CD-ROM.

TEXEIRA, J. M.; SAFÁDI, T.; FERREIRA, D. F. Um estudo preliminar da Segurança Pública na região Sul de Minas Gerais. In: REUNIÃO DA REGIÃO BRASILEIRA DA SOCIEDADE INTERNACIONAL DE BIOMETRIA, 50.,

2005, Londrina. **Programas e resumos...** Londrina, PR: UEL, 2005. 1CD-ROM.

TIETJEN, G. L.; JOHNSON, M. E. Exact Statistical Tolerance Limits for Sample Variances. **Technometrics**, Alexandria, v. 21, n. 1, p. 107-110, 1979.

TURNER, D. L.; BOWDEN, D. C. Simultaneous confidence bands for percentil lines in the general linear model. **Journal of the American Statistical Association**, Alexandria, v. 72, n. 360, p. 886-889, 1977.

TURNER, D. L.; BOWDEN, D. C. Sharp confidence bands for percentil lines and tolerance bands for simple linear models. **Journal of the American Statistical Association**, Alexandria, v. 74, n. 368, p. 885-888, 1979.

UNIVERSIDADE FEDERAL DE MINAS GERAIS. **Laboratório de Estatística Espacial**: Projetos Aplicados. Disponível em: <http://www.est.ufmg.br/leste/projetos.html#Projetos%20Aplicados>. Acesso em: 28 jan. 2005.

VIÉGAS, F. **Violência e Crime**. Disponível em: <http://revista.fapemig.br/15/sociologia.html>. Acesso em: 28 jan. 2005.

WALD, A. An extension of Wilk's method for setting tolerance limits. **Annals of Mathematical Statistics**, Hayward, v. 14, p. 45-55, 1943.

WALD, A.; WOLFOWITZ, J. Tolerance limits for a normal distribution. **Annals of Mathematical Statistics**, Hayward, v.17, p.208-215, 1946.

WALLIS, W.A. Tolerance Intervals for linear regression. Proc. **Second Berkeley Symposium**. Berkeley: University of Califórnia Press, 1951. p. 43-51.

WEISSEBERG, A.; BEATTY, G. H. Tables of Tolerance – Limit factors for Normal Distribution. **Technometrics**, Alexandria, v. 2, n. 4, p. 483-500, 1960.

WHITMORE, G. A. Prediction Limits for a Univariate Normal Observation. **The American Statistician**, Washington, v. 40, n. 2, p. 141-143, 1986.

WILKS, S. S. Determination of sample sizes for setting tolerance. **Annals of Mathematical Statistics**, Hayward, v. 12, p. 91-96, 1941.

WILKS, S. S. Statistical prediction with special reference to the problem of tolerance limits. **Annals of Mathematical Statistics**, Hayward, v. 13, p. 400-409, 1942.

## ANEXOS

<b>ANEXO A</b>	<b>Página</b>
FIGURA 1A Comparação entre os quatro tipos de intervalos estatísticos para CVPES anual com uma projeção de 7 anos contados a partir de 2000.....	115
FIGURA 2A Comparação entre os quatro tipos de intervalos estatísticos para CVPT de cada semestre com uma projeção de 7 anos contados a partir de 2000.....	115
FIGURA 3A Comparação entre os quatro tipos de intervalos estatísticos para ICVPES anual com uma projeção de 7 anos contados a partir de 2000.....	116
FIGURA 4A Comparação entre os quatro tipos de intervalos estatísticos para ICVPT de cada semestre com uma projeção de 7 anos contados a partir de 2000.....	116
<b>ANEXO B</b>	<b>Página</b>
TABELA 1B Tipos de Crimes Violentos.....	117
TABELA 2B Valores aproximados do quantil (t) da distribuição t não central para cada ano (h), utilizados na construção dos intervalos de tolerância unilateral considerando o período de 1991 a 2010.....	119
TABELA 3B Valores aproximados do quantil (t) da distribuição t não central para cada ano (h), utilizados na construção dos intervalos de tolerância unilateral considerando o período de 2000 a 2007.....	119
TABELA 4B Limites inferiores e superiores dos quatros intervalos estatísticos para o CVPES com projeção de 20 anos contados a partir de 1991.....	120

TABELA 5B Limites inferiores e superiores dos quatro intervalos estatísticos para o CVPT 1º semestre com projeção de 20 anos contados a partir de 1991.....121

TABELA 6B Limites inferiores e superiores dos quatro intervalos estatísticos para o CVPT 2º semestre com projeção de 20 anos contados a partir de 1991.....122

TABELA 7B Limites inferiores e superiores dos quatro intervalos estatísticos para o ICPES com projeção de 20 anos contados a partir de 1991.....123

TABELA 8B Limites inferiores e superiores dos quatro intervalos estatísticos para o ICVPT 1º semestre com projeção de 20 anos contados a partir de 1991.....124

TABELA 9B Limites inferiores e superiores dos quatro intervalos estatísticos para o ICVPT 2º semestre com projeção de 20 anos contados a partir de 1991.....125

TABELA 10B Limites inferiores e superiores dos quatro intervalos estatísticos para o CVPS com projeção de 7 anos contados a partir de 2000.....126

TABELA 11B Limites inferiores e superiores dos quatro intervalos estatísticos para o CVPT 1º semestre com projeção de 7 anos contados a partir de 2000...126

TABELA 12B Limites inferiores e superiores dos quatro intervalos estatísticos para o CVPT 2º semestre com projeção de 7 anos contados a partir de 2000...127

TABELA 13B Limites inferiores e superiores dos quatros intervalos estatísticos para o ICVPES com projeção de 7 anos contados a partir de 2000.....127

TABELA 14B Limites inferiores e superiores dos quatros intervalos estatísticos para o ICVPT 1º semestre com projeção de 7 anos contados a partir de 2000.....128

TABELA 15B Limites inferiores e superiores dos quatros intervalos estatísticos para o ICVPT 2º semestre com projeção de 7 anos contados a partir de 2000.....128

TABELA 16B ICV da amostra de tamanho 30 dos municípios mineiros no ano de 2003 com número de habitantes entre 14000 a 17500.....129

<b>ANEXO C</b>	<b>Página</b>
Rotinas no R (v 2.1.1).....	130

## ANEXO A

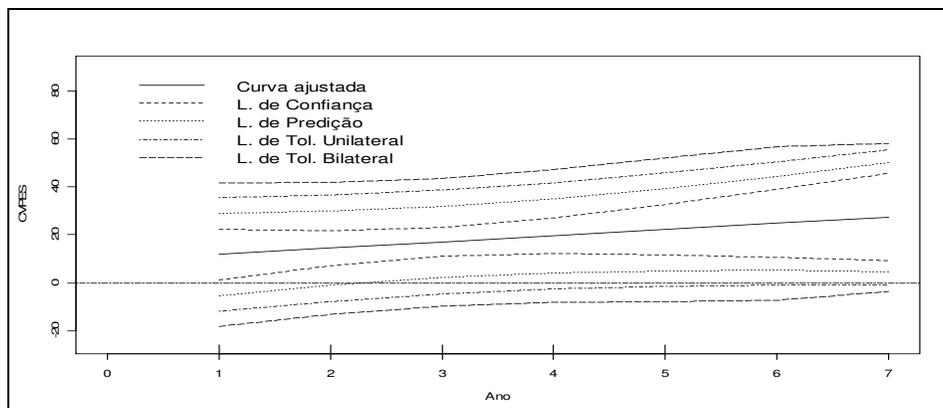


FIGURA 1A Comparação entre os quatro tipos de intervalos estatísticos para CVPEs anual com uma projeção de 7 anos contados a partir de 2000.

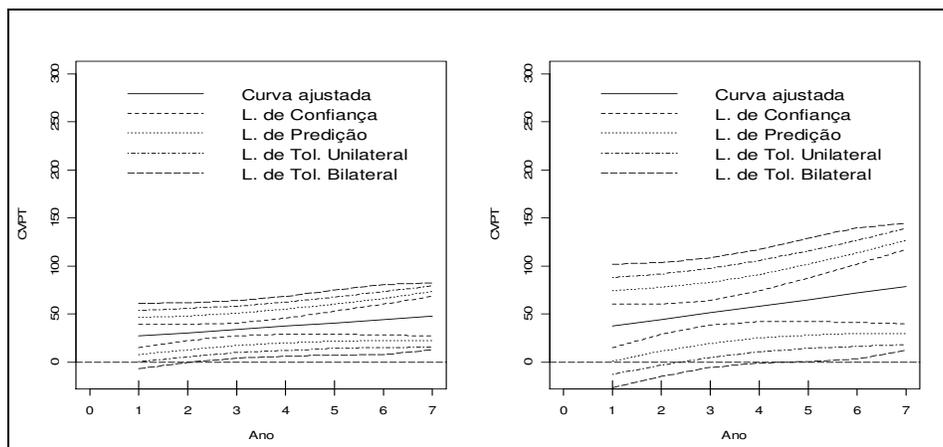


FIGURA 2A Comparação entre os quatro tipos de intervalos estatísticos para CVPT de cada semestre com uma projeção de 7 anos contados a partir de 2000

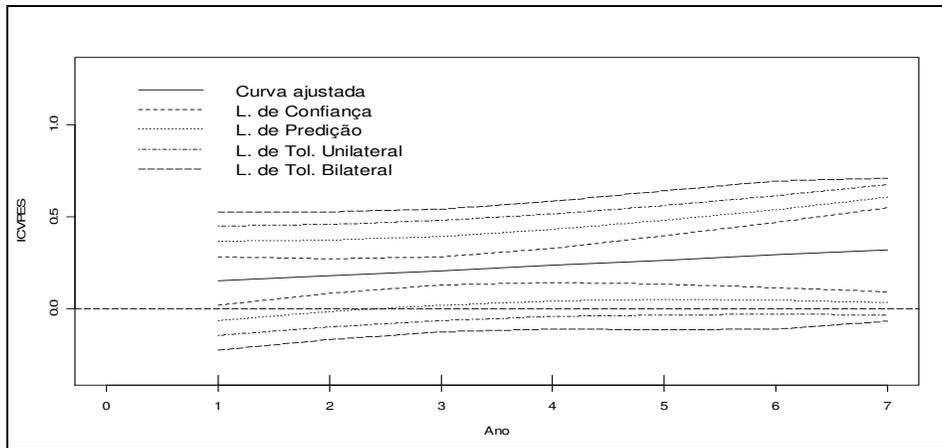


FIGURA 3A Comparação entre os quatro tipos de intervalos estatísticos para ICVPES anual com uma projeção de 7 anos contados a partir de 2000

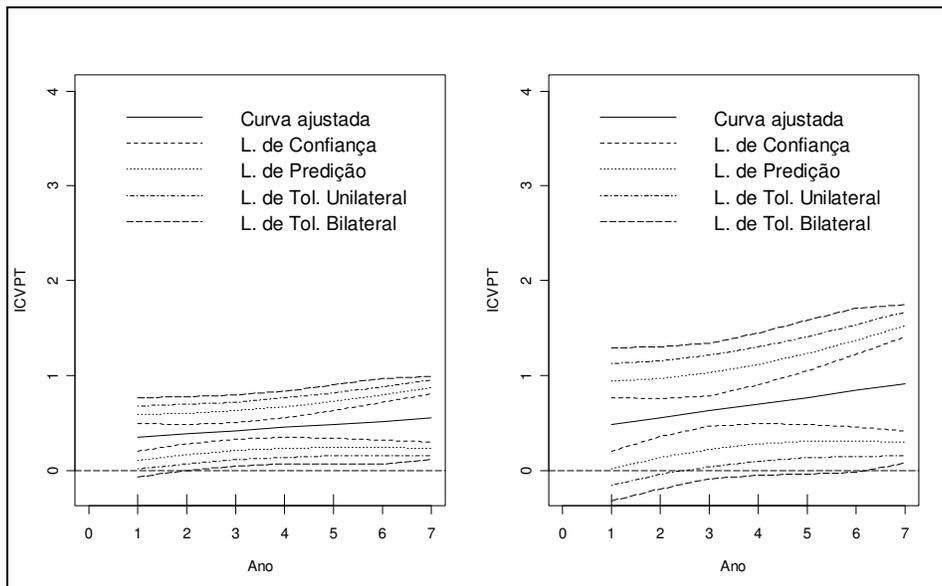


FIGURA 4A Comparação entre os quatro tipos de intervalos estatísticos para ICVPT de cada semestre com uma projeção de 7 anos contados a partir de 2000

## ANEXO B

TABELA 1B Tipos de Crimes Violentos

Cód natureza	m	Descrição da natureza
B04001	1	Homicídio tentado
B04002	2	Homicídio consumado
B05000	3	Seqüestro e cárcere privado
C05001	4	Roubo consumado a residência urbana
C05002	5	Roubo consumado a estabelecimento bancário/valores
C05004	6	Roubo consumado a ônibus/coletivo
C05007	7	Roubo consumado a drogaria/farmácia
C05009	8	Roubo consumado a supermercado/mercearia
C05010	9	Roubo consumado a prédio comercial
C05014	10	Roubo consumado a residência rural/sítio/chácara
C05015	11	Roubo consumado a estabelecimento público
C05016	12	Roubo consumado a depósitos em geral
C05017	13	Roubo consumado a garagem de ônibus/coletivo
C05018	14	Roubo consumado de veículo automotor
C05020	15	Roubo consumado a táxi
C05023	16	Roubo consumado a caminhão de cigarros
C05025	17	Roubo consumado de carga transportada
C05027	18	Roubo consumado a transeunte
C05028	19	Roubo consumado de bicicleta
C05031	20	Roubo consumado a posto de abastecimento de combustíveis
C05099	21	Outros roubos consumados
C09001	22	Roubo à mão armada consumado a residência urbana
C09004	23	Roubo à mão armada consumado a ônibus/coletivo
C09005	24	Roubo à mão armada consumado a casa lotérica
C09007	25	Roubo à mão armada consumado a drogaria/farmácia
C09008	26	Roubo à mão armada consumado a padaria
C09009	27	Roubo à mão armada consumado a supermercado/mercearia
C09010	28	Roubo à mão armada consumado a prédio comercial
C09013	29	Roubo à mão armada consumado a galeria de arte
C09014	30	Roubo à mão armada consumado a residência rural/sítio/chácara
C09016	31	Roubo à mão armada consumado a depósitos em geral
C09018	32	Roubo à mão armada consumado de veículo automotor
C09019	33	Roubo à mão armada consumado a passageiro de ônibus/coletivo

(continuação)

Cód natureza	m	Descrição da natureza
C09020	34	Roubo à mão armada consumado a táxi
C09024	35	Roubo à mão armada consumado a caminhão de cargas diversas
C09026	36	Roubo à mão armada consumado de veículo automotor e sua carga transportada
C09027	37	Roubo à mão armada consumado a transeunte
C09028	38	Roubo à mão armada consumado a bicicleta
C09031	39	Roubo à mão armada consumado a posto de abastecimento de Combustíveis
C09099	40	Outros roubos à mão armada consumados
C10	41	Roubo de veículo
C11	42	Roubo de veículo a mão armada
C13*	43	Latrocínio
C17*	44	Extorsão Mediante Seqüestro
D04001**	45	Estupro tentado
D04002**	46	Estupro consumado

\* Durante o período de 1991 a 2004, não houve nenhuma ocorrência dos tipos de crimes violentos C13 e C17 no município de Lavras – MG.

\*\* Estão classificados na categoria de crimes violentos contra o costume, portanto não foram considerados neste trabalho.

TABELA 2B Valores aproximados do quantil (t) da distribuição t não central para cada ano (h), utilizados na construção dos intervalos de tolerância unilateral considerando o período de 1991 a 2010

Ano (h)	Quantil (t)						
1	-5,87	6	-9,39	11	-7,81	16	-5,06
2	-6,41	7	-9,88	12	-7,05	17	-4,75
3	-7,05	8	-9,88	13	-6,41	18	-4,49
4	-7,81	9	-9,39	14	-5,87	19	-4,27
5	-8,63	10	-8,63	15	-5,43	20	-4,08

TABELA 3B Valores aproximados do quantil (t) da distribuição t não central para cada ano (h), utilizados na construção dos intervalos de tolerância unilateral considerando o período de 2000 a 2007

Ano (h)	Quantil (t)						
1	-7,17	3	-11,36	5	-7,17	7	-4,94
2	-9,50	4	-9,50	6	-5,77		

TABELA 4B Limites inferiores e superiores dos quatros intervalos estatísticos para o CVPES com projeção de 20 anos contados a partir de 1991

Ano (h)	CVPES obs.	CVPES ajust.	L.C. Inf.	L.C. Sup.	L.P. Inf.	L.P. Sup.	L.T.U. Inf.	L.T.U. Sup.	L.T.B. Inf.	L.T.B. Sup.
1	10	5,54	0,97	10,11	-4,56	15,64	-6,77	17,86	-8,16	19,24
2	5	6,59	2,52	10,66	-3,30	16,48	-5,39	18,57	-6,84	20,03
3	11	7,64	4,03	11,25	-2,07	17,34	-4,05	19,32	-5,56	20,84
4	3	8,69	5,50	11,88	-0,87	18,25	-2,74	20,12	-4,32	21,70
5	11	9,74	6,90	12,57	0,29	19,18	-1,49	20,96	-3,13	22,60
6	5	10,78	8,22	13,35	1,42	20,15	-0,29	21,86	-1,98	23,55
7	16	11,83	9,41	14,26	2,50	21,16	0,83	22,84	-0,88	24,55
8	15	12,88	10,45	15,31	3,55	22,21	1,88	23,89	0,17	25,60
9	12	13,93	11,36	16,50	4,56	23,30	2,85	25,01	1,17	26,69
10	8	14,98	12,14	17,81	5,53	24,42	3,75	26,20	2,11	27,84
11	17	16,03	12,84	19,22	6,47	25,58	4,60	27,45	3,01	29,04
12	22	17,07	13,47	20,68	7,37	26,78	5,39	28,76	3,87	30,28
13	17	18,12	14,05	22,20	8,24	28,01	6,14	30,11	4,69	31,56
14	21	19,17	14,60	23,74	9,07	29,27	6,86	31,49	5,47	32,87
15		20,22	15,13	25,31	9,87	30,57	7,55	32,89	6,23	34,21
16		21,27	15,65	26,89	10,65	31,89	8,22	34,32	6,96	35,58
17		22,32	16,15	28,48	11,40	33,23	8,87	35,76	7,69	36,95
18		23,36	16,65	30,08	12,13	34,60	9,51	37,22	8,41	38,32
19		24,41	17,13	31,69	12,83	36,00	10,14	38,69	9,15	39,68
20		25,46	17,62	33,31	13,52	37,41	10,76	40,17	9,90	41,02

TABELA 5B Limites inferiores e superiores dos quatros intervalos estatísticos para o CVPT 1º semestre com projeção de 20 anos contados a partir de 1991

Ano (h)	CVPT obs.	CVPT ajust.	L.C. Inf.	L.C. Sup.	L.P. Inf.	L.P. Sup.	L.T.U. Inf.	L.T.U. Sup.	L.T.B. Inf.	L.T.B. Sup.
1	5	1,46	-4,41	7,33	-11,52	14,44	-14,36	17,28	-16,15	19,06
2	9	4,29	-0,95	9,52	-8,42	16,99	-11,11	19,68	-12,97	21,55
3	11	7,12	2,48	11,76	-5,35	19,59	-7,89	22,13	-9,85	24,08
4	4	9,95	5,85	14,05	-2,33	22,23	-4,73	24,63	-6,77	26,67
5	15	12,78	9,14	16,42	0,65	24,91	-1,64	27,20	-3,75	29,31
6	6	15,61	12,31	18,91	3,57	27,65	1,38	29,85	-0,79	32,01
7	14	18,44	15,32	21,56	6,45	30,43	4,30	32,58	2,11	34,78
8	25	21,27	18,16	24,39	9,28	33,26	7,13	35,41	4,94	37,61
9	18	24,10	20,80	27,40	12,07	36,14	9,87	38,34	7,70	40,50
10	27	26,93	23,29	30,57	14,80	39,07	12,51	41,35	10,41	43,46
11	32	29,76	25,67	33,86	17,49	42,04	15,08	44,45	13,05	46,48
12	30	32,60	27,96	37,23	20,13	45,07	17,58	47,61	15,63	49,56
13	44	35,43	30,19	40,66	22,72	48,13	20,03	50,82	18,16	52,69
14	38	38,26	32,39	44,13	25,28	51,24	22,44	54,08	20,65	55,86
15		41,09	34,55	47,62	27,80	54,38	24,81	57,37	23,11	59,07
16		43,92	36,70	51,14	30,28	57,56	27,15	60,69	25,54	62,30
17		46,75	38,83	54,67	32,72	60,77	29,47	64,03	27,95	65,55
18		49,58	40,95	58,21	35,14	64,02	31,78	67,38	30,37	68,79
19		52,41	43,06	61,76	37,53	67,29	34,07	70,75	32,79	72,03
20		55,24	45,16	65,32	39,89	70,59	36,35	74,14	35,25	75,23

TABELA 6B Limites inferiores e superiores dos quatros intervalos estatísticos para o CVPT 2º semestre com projeção de 20 anos contados a partir de 1991

Ano (h)	CVPT obs.	CVPT ajust.	L.C. Inf.	L.C. Sup.	L.P. Inf.	L.P. Sup.	L.T.U. Inf.	L.T.U. Sup.	L.T.B. Inf.	L.T.B. Sup.
1	8	-0,34	-11,47	10,79	-24,96	24,27	-30,35	29,66	-33,73	33,04
2	13	3,97	-5,95	13,90	-20,12	28,06	-25,22	33,17	-28,76	36,71
3	13	8,29	-0,50	17,08	-15,36	31,94	-20,18	36,76	-23,88	40,46
4	17	12,61	4,84	20,38	-10,68	35,89	-15,24	40,45	-19,10	44,31
5	10	16,92	10,02	23,83	-6,09	39,93	-10,42	44,27	-14,42	48,27
6	7	21,24	14,98	27,50	-1,59	44,07	-5,76	48,23	-9,86	52,34
7	24	25,56	19,64	31,47	2,82	48,29	-1,26	52,37	-5,42	56,53
8	12	29,87	23,96	35,78	7,14	52,61	3,06	56,69	-1,10	60,85
9	27	34,19	27,93	40,45	11,36	57,02	7,19	61,18	3,09	65,29
10	38	38,51	31,60	45,41	15,49	61,52	11,16	65,85	7,16	69,85
11	36	42,82	35,05	50,59	19,54	66,11	14,98	70,67	11,12	74,53
12	64	47,14	38,35	55,93	23,49	70,79	18,67	75,61	14,97	79,31
13	58	51,45	41,53	61,38	27,36	75,55	22,26	80,65	18,72	84,19
14	61	55,77	44,64	66,90	31,16	80,38	25,77	85,77	22,39	89,16
15		60,09	47,70	72,48	34,88	85,30	29,21	90,96	25,99	94,19
16		64,40	50,71	78,10	38,53	90,28	32,60	96,20	29,55	99,26
17		68,72	53,70	83,74	42,12	95,32	35,96	101,48	33,07	104,37
18		73,04	56,67	89,41	45,65	100,42	39,28	106,80	36,60	109,47
19		77,35	59,62	95,09	49,13	105,57	42,57	112,14	40,15	114,55
20		81,67	62,56	100,79	52,56	110,78	45,84	117,50	43,76	119,58

TABELA 7B Limites inferiores e superiores dos quatros intervalos estatísticos para o ICVPES com projeção de 20 anos contados a partir de 1991

Ano (h)	ICV obs.	ICV ajust.	L.C. Inf.	L.C. Sup.	L.P. Inf.	L.P. Sup.	L.T.U. Inf.	L.T.U. Sup.	L.T.B. Inf.	L.T.B. Sup.
1	0,152	0,092	0,032	0,153	-0,042	0,226	-0,089	0,274	-0,174	0,212
2	0,074	0,103	0,049	0,157	-0,028	0,234	-0,075	0,281	-0,155	0,231
3	0,161	0,113	0,066	0,161	-0,015	0,242	-0,062	0,288	-0,136	0,250
4	0,043	0,124	0,082	0,166	-0,003	0,250	-0,048	0,296	-0,118	0,270
5	0,155	0,134	0,097	0,172	0,009	0,259	-0,036	0,305	-0,100	0,290
6	0,069	0,145	0,111	0,179	0,021	0,269	-0,024	0,314	-0,082	0,310
7	0,213	0,155	0,123	0,188	0,032	0,279	-0,013	0,324	-0,064	0,330
8	0,196	0,166	0,134	0,198	0,042	0,290	-0,002	0,334	-0,047	0,350
9	0,153	0,177	0,143	0,211	0,053	0,301	0,008	0,346	-0,029	0,371
10	0,102	0,187	0,150	0,225	0,062	0,312	0,017	0,357	-0,012	0,392
11	0,212	0,198	0,155	0,240	0,071	0,324	0,025	0,370	0,005	0,412
12	0,270	0,208	0,160	0,256	0,080	0,337	0,033	0,383	0,022	0,433
13	0,205	0,219	0,165	0,273	0,088	0,350	0,041	0,397	0,039	0,454
14	0,246	0,229	0,169	0,290	0,096	0,363	0,048	0,411	0,056	0,476
15		0,240	0,173	0,307	0,103	0,377	0,055	0,425	0,073	0,497
16		0,250	0,176	0,325	0,110	0,391	0,061	0,440	0,089	0,518
17		0,261	0,179	0,343	0,116	0,405	0,067	0,455	0,106	0,540
18		0,272	0,183	0,360	0,123	0,420	0,074	0,469	0,122	0,561
19		0,282	0,186	0,378	0,129	0,435	0,080	0,484	0,139	0,583
20		0,293	0,189	0,396	0,134	0,451	0,087	0,499	0,155	0,605

TABELA 8B Limites inferiores e superiores dos quatros intervalos estatísticos para o ICVPT 1º semestre com projeção de 20 anos contados a partir de 1991

Ano (h)	ICVPT obs.	ICVPT ajust.	L.C. Inf.	L.C. Sup.	L.P. Inf.	L.P. Sup.	L.T.U. Inf.	L.T.U. Sup.	L.T.B. Inf.	L.T.B. Sup.
1	0,076	0,044	-0,030	0,118	-0,120	0,208	-0,156	0,244	-0,178	0,267
2	0,134	0,076	0,010	0,142	-0,084	0,237	-0,118	0,271	-0,142	0,294
3	0,161	0,109	0,050	0,167	-0,049	0,266	-0,081	0,298	-0,106	0,323
4	0,057	0,141	0,089	0,193	-0,014	0,296	-0,045	0,326	-0,071	0,352
5	0,212	0,173	0,127	0,219	0,020	0,326	-0,009	0,355	-0,036	0,382
6	0,082	0,205	0,164	0,247	0,053	0,357	0,025	0,385	-0,002	0,413
7	0,187	0,237	0,198	0,277	0,086	0,389	0,059	0,416	0,031	0,444
8	0,326	0,270	0,230	0,309	0,118	0,421	0,091	0,448	0,063	0,476
9	0,230	0,302	0,260	0,344	0,150	0,454	0,122	0,482	0,095	0,509
10	0,343	0,334	0,288	0,380	0,181	0,488	0,152	0,516	0,125	0,543
11	0,398	0,366	0,315	0,418	0,211	0,522	0,181	0,552	0,155	0,578
12	0,368	0,399	0,340	0,457	0,241	0,556	0,209	0,588	0,184	0,613
13	0,532	0,431	0,365	0,497	0,270	0,591	0,236	0,625	0,213	0,649
14	0,445	0,463	0,389	0,537	0,299	0,627	0,263	0,663	0,241	0,686
15		0,495	0,413	0,578	0,327	0,663	0,290	0,701	0,268	0,723
16		0,528	0,436	0,619	0,355	0,700	0,316	0,740	0,295	0,760
17		0,560	0,460	0,660	0,383	0,737	0,342	0,778	0,322	0,797
18		0,592	0,483	0,701	0,410	0,775	0,367	0,817	0,349	0,835
19		0,624	0,506	0,743	0,436	0,812	0,393	0,856	0,376	0,872
20		0,657	0,529	0,784	0,463	0,851	0,418	0,895	0,404	0,909

TABELA 9B Limites inferiores e superiores dos quatros intervalos estatísticos para o ICVPT 2º semestre com projeção de 20 anos contados a partir de 1991

Ano (h)	ICV obs.	ICV ajust.	L.C. Inf.	L.C. Sup.	L.P Inf.	L.P. Sup.	L.T.U Inf.	L.T.U Sup.	L.T.B Inf.	L.T.B Sup.
1	0,121	0,033	-0,103	0,169	-0,268	0,335	-0,334	0,401	-0,376	0,442
2	0,193	0,082	-0,039	0,204	-0,213	0,377	-0,275	0,440	-0,319	0,483
3	0,190	0,131	0,024	0,239	-0,158	0,421	-0,217	0,480	-0,263	0,526
4	0,244	0,181	0,085	0,276	-0,105	0,466	-0,161	0,522	-0,208	0,569
5	0,141	0,230	0,145	0,314	-0,052	0,512	-0,105	0,565	-0,154	0,614
6	0,096	0,279	0,202	0,356	-0,001	0,559	-0,052	0,610	-0,102	0,660
7	0,320	0,328	0,256	0,401	0,050	0,607	0,000	0,657	-0,051	0,708
8	0,157	0,377	0,305	0,450	0,099	0,656	0,049	0,706	-0,002	0,757
9	0,345	0,427	0,350	0,503	0,147	0,706	0,096	0,757	0,045	0,808
10	0,482	0,476	0,391	0,560	0,194	0,758	0,141	0,811	0,092	0,860
11	0,448	0,525	0,430	0,620	0,240	0,810	0,184	0,866	0,136	0,913
12	0,785	0,574	0,466	0,682	0,284	0,864	0,225	0,923	0,180	0,968
13	0,701	0,623	0,502	0,745	0,328	0,918	0,266	0,981	0,222	1,024
14	0,714	0,672	0,536	0,809	0,371	0,974	0,305	1,040	0,263	1,081
15		0,722	0,570	0,873	0,413	1,030	0,343	1,100	0,304	1,139
16		0,771	0,603	0,938	0,454	1,088	0,381	1,160	0,344	1,198
17		0,820	0,636	1,004	0,494	1,146	0,418	1,221	0,383	1,257
18		0,869	0,669	1,070	0,534	1,205	0,455	1,283	0,423	1,315
19		0,918	0,701	1,136	0,572	1,264	0,492	1,344	0,462	1,374
20		0,967	0,733	1,202	0,611	1,324	0,528	1,406	0,503	1,432

TABELA 10B Limites inferiores e superiores dos quatros intervalos estatísticos para o CVPS com projeção de 7 anos contados a partir de 2000

Ano (h)	ICV obs.	ICV ajust.	L.C. Inf.	L.C. Sup.	L.P. Inf.	L.P. Sup.	L.T.U. Inf.	L.T.U. Sup.	L.T.B. Inf.	L.T.B. Sup.
1	8	11,80	1,30	22,30	-5,34	28,94	-11,85	35,45	-18,17	41,77
2	17	14,40	6,98	21,82	-1,05	29,85	-7,77	36,57	-13,20	42,00
3	22	17,00	10,94	23,06	2,15	31,85	-4,63	38,63	-9,62	43,62
4	17	19,60	12,18	27,02	4,15	35,05	-2,57	41,77	-8,00	47,20
5	21	22,20	11,70	32,70	5,06	39,34	-1,45	45,85	-7,77	52,17
6		24,80	10,59	39,01	5,16	44,44	-0,96	50,56	-7,22	56,82
7		27,40	9,22	45,58	4,72	50,08	-0,84	55,64	-3,50	58,30

TABELA 11B Limites inferiores e superiores dos quatros intervalos estatísticos para o CVPT 1º semestre com projeção de 7 anos contados a partir de 2000

Ano (h)	ICV obs.	ICV ajust.	L.C. Inf.	L.C. Sup.	L.P. Inf.	L.P. Sup.	L.T.U. Inf.	L.T.U. Sup.	L.T.B. Inf.	L.T.B. Sup.
1	27	27,40	15,56	39,24	8,07	46,73	0,73	54,07	-6,40	61,20
2	32	30,80	22,43	39,17	13,37	48,23	5,79	55,81	-0,33	61,93
3	30	34,20	27,36	41,04	17,46	50,94	9,81	58,59	4,18	64,22
4	44	37,60	29,23	45,97	20,17	55,03	12,59	62,61	6,47	68,73
5	38	41,00	29,16	52,84	21,67	60,33	14,33	67,67	7,20	74,80
6		44,40	28,37	60,43	22,25	66,55	15,35	73,45	8,28	80,52
7		47,80	27,29	68,31	22,22	73,38	15,95	79,65	12,95	82,65

TABELA 12B Limites inferiores e superiores dos quatros intervalos estatísticos para o CVPT 2º semestre com projeção de 7 anos contados a partir de 2000

Ano (h)	ICV obs.	ICV ajust.	L.C. Inf.	L.C. Sup.	L.P. Inf.	L.P. Sup.	L.T.U. Inf.	L.T.U. Sup.	L.T.B. Inf.	L.T.B. Sup.
1	38	37,80	15,35	60,25	1,14	74,46	-12,78	88,38	-26,29	101,89
2	36	44,60	28,73	60,47	11,56	77,64	-2,81	92,01	-14,42	103,62
3	64	51,40	38,44	64,36	19,65	83,15	5,15	97,65	-5,52	108,32
4	58	58,20	42,33	74,07	25,16	91,24	10,79	105,61	-0,82	117,22
5	61	65,00	42,55	87,45	28,34	101,66	14,42	115,58	0,91	129,09
6		71,80	41,40	102,20	29,80	113,80	16,71	126,89	3,32	140,28
7		78,60	39,72	117,48	30,10	127,10	18,21	138,99	12,52	144,68

TABELA 13B Limites inferiores e superiores dos quatros intervalos estatísticos para o ICV PES com projeção de 7 anos contados a partir de 2000

Ano (h)	ICV obs.	ICV ajust.	L.C. Inf.	L.C. Sup.	L.P. Inf.	L.P. Sup.	L.T.U. Inf.	L.T.U. Sup.	L.T.B. Inf.	L.T.B. Sup.
1	0,1016	0,150	0,018	0,283	-0,066	0,366	-0,148	0,448	-0,227	0,528
2	0,2116	0,179	0,085	0,272	-0,016	0,373	-0,101	0,458	-0,169	0,526
3	0,27	0,207	0,131	0,283	0,020	0,394	-0,066	0,479	-0,129	0,542
4	0,2055	0,235	0,142	0,329	0,040	0,430	-0,044	0,515	-0,113	0,583
5	0,246	0,263	0,131	0,396	0,047	0,479	-0,035	0,561	-0,114	0,641
6		0,292	0,113	0,471	0,044	0,539	-0,033	0,616	-0,112	0,695
7		0,320	0,091	0,549	0,034	0,606	-0,036	0,676	-0,069	0,709

TABELA 14B Limites inferiores e superiores dos quatros intervalos estatísticos para o ICVPT 1º semestre com projeção de 7 anos contados a partir de 2000

Ano (h)	ICV obs.	ICV ajust.	L.C. Inf.	L.C. Sup.	L.P. Inf.	L.P. Sup.	L.T.U. Inf.	L.T.U. Sup.	L.T.B. Inf.	L.T.B. Sup.
1	0,343	0,350	0,202	0,497	0,108	0,591	0,017	0,682	-0,072	0,771
2	0,398	0,383	0,279	0,488	0,166	0,601	0,071	0,695	-0,005	0,772
3	0,368	0,417	0,332	0,503	0,208	0,626	0,113	0,722	0,043	0,792
4	0,532	0,451	0,347	0,556	0,234	0,669	0,139	0,763	0,063	0,839
5	0,445	0,485	0,337	0,633	0,244	0,726	0,152	0,818	0,063	0,907
6		0,519	0,319	0,719	0,242	0,795	0,156	0,881	0,068	0,969
7		0,552	0,297	0,808	0,233	0,872	0,155	0,950	0,118	0,987

TABELA 15B Limites inferiores e superiores dos quatros intervalos estatísticos para o ICVPT 2º semestre com projeção de 7 anos contados a partir de 2000

Ano (h)	ICV obs.	ICV ajust.	L.C. Inf.	L.C. Sup.	L.P. Inf.	L.P. Sup.	L.T.U. Inf.	L.T.U. Sup.	L.T.B. Inf.	L.T.B. Sup.
1	0,482	0,483	0,199	0,767	0,020	0,946	-0,156	1,122	-0,327	1,293
2	0,448	0,555	0,354	0,755	0,137	0,972	-0,045	1,154	-0,191	1,300
3	0,785	0,626	0,462	0,790	0,225	1,027	0,042	1,211	-0,093	1,346
4	0,701	0,698	0,497	0,899	0,280	1,116	0,099	1,297	-0,048	1,444
5	0,714	0,770	0,486	1,053	0,306	1,233	0,131	1,409	-0,040	1,580
6		0,841	0,457	1,225	0,311	1,372	0,145	1,537	-0,024	1,707
7		0,913	0,422	1,404	0,300	1,526	0,150	1,676	0,078	1,748

TABELA 16B ICV da amostra de tamanho 30 dos municípios mineiros no ano de 2003 com número de habitantes entre 14000 a 17500.

Amostra	ICV								
1	0,31	7	0,64	13	1,04	19	1,16	25	1,4
2	0,34	8	0,76	14	1,05	20	1,21	26	1,44
3	0,55	9	0,83	15	1,07	21	1,25	27	1,51
4	0,58	10	0,86	16	1,1	22	1,28	28	1,59
5	0,61	11	0,89	17	1,1	23	1,28	29	1,82
6	0,62	12	0,96	18	1,14	24	1,34	30	1,85

## ANEXO C

Neste anexo estão disponíveis todas as rotinas desenvolvidas no Software estatístico R<sup>®</sup> v.2.1.1 (R, 2005), sendo que as expressões ou linhas de comando que estão em negrito são aquelas que podem ser alteradas de acordo com os dados.

### **Rotina para a figura 3 da seção 4.1**

```
PMano<-read.table("62_68C.txt",h=T)
attach(PMano)
y<-Crimes
x<-Ano
PMano1<-read.table("72_78C.txt",h=T)
attach(PMano1)
y1<-Crimes
x1<-Ano
PMano2<-read.table("82_88C.txt",h=T)
attach(PMano2)
y2<-Crimes
x2<-Ano
plot(x2,y2,xlab="Ano",ylab="CV", cex.lab=1.5, cex.axis=1.2, type="l", lty=3, ylim =
c(28,200), xlim=c(0.5,14.5))
lines(x1,y1,lty=2)
lines(x,y,lty=1)
legend (0,195,c('Faixa 1(62000 a 68000)' , 'Faixa 2(72000 a 78000)',
'Faixa 3(82000 a 88000)'), lty=c(1,2,3),cex=1.3,bty="n")

#Observação: Na utilização do comando plot deve se verificar a configuração da
resolução de tela do computador, no caso deste trabalho foi utilizado 1024x768 pixels,
caso a resolução seja outra, pode ser necessário alterar os valores dos comandos tipo
cex.lab, cex.axis,etc .#
```

### Rotinas para as tabelas 4 e 5 da seção 4.2.1

```
xtotalps<-read.table("testeaderpsCV.txt",h=F)
xtotalps<-as.matrix(xtotalps)
chisq.test(xtotalps)
xtotalpt<-read.table("testeaderptCV.txt",h=F)
xtotapt<-as.matrix(xtotalpt)
chisq.test(xtotalpt)
xpt1sem<-read.table("testeaderpt1CV.txt",h=F)
xpt1sem<-as.matrix(xpt1sem)
chisq.test(xpt1sem)
xpt2sem<-read.table("testeaderpt2CV.txt",h=F)
xpt2sem<-as.matrix(xpt2sem)
chisq.test(xpt2sem)
```

### Rotinas para a tabela 6 da seção 4.2.2.1

```
#1° exemplo) ICVPES (para o CVPES é só modificar o arquivo de dados)
#Regressão
PManops<-read.table("novops.txt",h=T)
regps<-lm(ICVPES~Ano,data=PManops)
summary(regps)
rps<-resid(regps)
shapiro.test(rps)
library(lmtest)
dwtest(ICVPES~ 0+Ano,data=PManops)

#2° Exemplo) ICVPT 1° e 2° semestres (para o CVPT é só modificar o arquivo dos
dados)
#Regressão 1° semestre
PMano1<-read.table("novopt1_6.txt",h=T)
reg1<-lm(ICV~Ano,data=PMano1)
summary(reg1)
r1<-resid(reg1)
shapiro.test(r1)
library(lmtest)
dwtest(ICV~ Ano,data=PMano1)
#Regressão 2° semestre
PMano2<-read.table("novopt7_12.txt",h=T)
reg2<-lm(ICV~Ano,data=PMano2)
summary(reg2)
r2<-resid(reg2)
shapiro.test(r2)
dwtest(ICV~Ano,data=PMano2)
```

#Observação: Para a obtenção da tabela 10 (seção 4.3.1.1) utilizaram-se os mesmos comandos, com a única diferença, que o banco de dados é referente ao período de 2000 a 2004.

#### Rotinas para as figuras 4 a 7 da seção 4.2.2.1

```
#1º exemplo) ICPVES (para o CVPES é só modificar o arquivo de dados)
#Gráfico para comparação entre o ICV contra a Pessoa ajustado e observado
PManops<-read.table("novops.txt",h=T)
regps<-lm(ICVPES~Ano,data=PManops)
attach(PManops)
yps<-ICVPES
xps<-Ano
attach(regps)
fps<-fitted(regps)
plot(xps, fps, xlab = "Ano",ylab="ICVPES", cex.lab = 1.5, cex.axis = 1.2,type=
"l",lty= 1,xlim=c(0,15), ylim=c(0,0.4))
points(xps,yps,pch=19)
legend(1.0,0.36,'ICVPES ajustado=0,0189(Ano) ',lty=1,cex=1.3,bty="n")
legend(1.6,0.33,' ICVPES observado ',cex=1.3,bty="n",pch=19,lty=0)

#2º Exemplo) ICPVT 1º e 2º semestres (para o CVPT é só modificar o arquivo dos
dados)
#Gráfico para comparação entre o 1º e 2º semestres
PMano1<-read.table("novopt1_6.txt",h=T)
reg1<-lm(ICV~Ano,data=PMano1)
attach(PMano1)
y1<-ICV
x1<-Ano
attach(reg1)
f1<-fitted(reg1)
PMano2<-read.table("novopt7_12.txt",h=T)
reg2<-lm(ICV~Ano,data=PMano2)
attach(PMano2)
y2<-ICV
x2<-Ano
attach(reg2)
f2<-fitted(reg2)
plot(x2,f2,xlab="Ano",ylab="ICVPT", cex.lab=1.5, cex.axis=1.2, type="l",lty=2,
xlim=c(0,15), ylim=c(0,1.05))
mtext("(1991-2004)",cex=1.2)
points(x2,y2,pch=24)
points(x1,y1,pch=19)
lines(x1,f1,lty=1)
legend(1,0.96,c('ICVPT ajustado 1º sem =0,0335(Ano)','ICVPT ajustado 2º sem
=0,0475(Ano)'),lty=c(1,2),cex=1.3,bty="n")
```

```
legend(1.6,0.83,c(' ICVPT observado - 1ºsem', ' ICVPT observado - 2º sem'),  
lty=c(0,0), pch=c(19,24), cex=1.3,bty="n")
```

#Observação; Para a obtenção das figuras de 13 a 16 (seção 4.3.1.1) utilizaram-se os mesmos comandos, com a única diferença, que o banco de dados é referente ao período de 2000 a 2004.

### Rotina para a figura 8 da seção 4.2.2.1

```
#Gráficos para análise dos resíduos CVPES  
PManops<-read.table("novops.txt",h=T)  
regps<-lm(CVPES~Ano,data=PManops)  
sregps<-summary(regps)  
attach(regps)  
fps<-fitted(regps)  
rps<-resid(regps)  
rinfps<-influence(regps)  
studps<-rps/(sregps$sig*sqrt(1-rinfps$hat))  
par(mfrow=c(1,2))  
qqnorm(studps,xlab="Quantil Teórico (a)", ylab="Resíduos Padronizados", main  
=" ",cex.lab=1.5,cex.main=1.8,cex.axis=1.2)  
abline(0,1)  
plot(fps,rps,xlab="CVPES ajustado (b)",ylab="Resíduos", main="", cex.lab=1.5,  
cex.main=1.8, cex.axis=1.2)  
abline(h=0)  
  
#Gráficos para análise dos resíduos CVPT- 1º semestre  
PMano1<-read.table("novopt1_6.txt",h=T)  
reg1<-lm(PT1~Ano,data=PMano1)  
sreg1<-summary(reg1)  
attach(reg1)  
f1<-fitted(reg1)  
r1<-resid(reg1)  
rinf1<-influence(reg1)  
stud1<-r1/(sreg1$sig*sqrt(1-rinf1$hat))  
par(mfrow=c(1,2))  
qqnorm(stud1,xlab="Quantil Teórico (c)",ylab="Resíduos Padronizados", main  
="", cex.lab=1.5, cex.main=1.8, cex.axis=1.2)  
abline(0,1)  
plot(f1,r1,xlab="CVPT ajustado (d)",ylab="Resíduos", main="", cex.lab=1.5,  
cex.main=1.8, cex.axis=1.2)  
abline(h=0)
```

```

#Gráficos para análise dos resíduos CVPT- 2º semestre
PMano2<-read.table("novopt7_12.txt",h=T)
reg2<-lm(PT2~Ano,data=PMano2)
sreg2<-summary(reg2)
attach(reg2)
f2<-fitted(reg2)
r2<-resid(reg2)
rinf2<-influence(reg2)
stud2<-r2/(sreg2$sig*sqrt(1-rinf2$hat))
par(mfrow=c(1,2))
qqnorm(stud2,xlab="Quantil Téorico (e)",ylab="Resíduos Padronizados",main="",
cex.lab=1.5, cex.main=1.8, cex.axis=1.2)
abline(0,1)
plot(f2,r2,xlab="CVPT ajustado (f)",ylab="Resíduos", main="", cex.lab=1.5,
cex.main=1.8, cex.axis=1.2)
abline(h=0)

#Gráficos para análise dos resíduos ICPES
PManops<-read.table("novops.txt",h=T)
regips<-lm(ICVPES~Ano,data=PManops)
sregips<-summary(regips)
attach(regips)
fips<-fitted(regips)
rips<-resid(regips)
rinfips<-influence(regips)
studips<-rips/(sregips$sig*sqrt(1-rinfips$hat))
par(mfrow=c(1,2))
qqnorm(studips,xlab="Quantil Téorico (g)",ylab="Resíduos Padronizados",main
="", cex.lab=1.5, cex.main=1.8, cex.axis=1.2)
abline(0,1)
plot(fips,rips,xlab="ICVPES ajustado (h)",ylab="Resíduos", main="", cex.lab=1.5,
cex.main=1.8, cex.axis=1.2)
abline(h=0)

#Gráficos para análise dos resíduos ICVPT - 1º semestre
PMano1i<-read.table("novopt1_6.txt",h=T)
reg1i<-lm(ICV~Ano,data=PMano1i)
sreg1i<-summary(reg1i)
attach(reg1i)
f1i<-fitted(reg1i)
r1i<-resid(reg1i)
rinf1i<-influence(reg1i)
stud1i<-r1i/(sreg1i$sig*sqrt(1-rinf1i$hat))
par(mfrow=c(1,2))
qqnorm(stud1i,xlab="Quantil Téorico (i)",ylab="Resíduos Padronizados",main=""
,cex.lab=1.5, cex.main=1.8, cex.axis=1.2)

```

```

abline(0,1)
plot(f1i,r1i,xlab="ICVPT ajustado      (j)",ylab="Residuos", main="", cex.lab=1.5,
cex.main=1.8, cex.axis=1.2)
abline(h=0)

#Gráficos para análise dos resíduos ICVPT - 2º semestre
PMano2i<-read.table("novopt7_12.txt",h=T)
reg2i<-lm(ICV~Ano,data=PMano2i)
sreg2i<-summary(reg2i)
attach(reg2i)
f2i<-fitted(reg2i)
r2i<-resid(reg2i)
rinf2i<-influence(reg2i)
stud2i<-r2i/(sreg2i$sig*sqrt(1-rinf2i$hat))
par(mfrow=c(1,2))
qqnorm(stud2i,xlab="Quantil Teórico      (k)",ylab="Resíduos Padronizados",main
="", cex.lab=1.5, cex.main=1.8, cex.axis=1.2)
abline(0,1)
plot(f2i,r2i,xlab="ICVPT ajustado      (l)",ylab="Residuos", main="", cex.lab=1.5,
cex.main=1.8, cex.axis=1.2)
abline(h=0)

```

#### **Rotinas para as figuras 9 a 12 (seção 4.2.2.2.1) e para as tabelas 7 (seção 4.2.2.2.2) e 4B a 9B (Anexo B)**

```

#1º exemplo) ICVPES (para o CVPES é só modificar o arquivo de dados e a escala
dos gráfico)
#Intervalos de Confiança e Predição - Projeção 20 anos - 1º semestre

PManops<-read.table("novops.txt",h=T)
regps<-lm(ICVPES~Ano,data=PManops)
xps<-Ano
mps<-mean(xps)
grid<-seq(1,20,1)
ps<-data.frame(Ano=grid)
psC<-predict(regps,ps,interval="confidence") #Cálculo direto do intervalo de
confiança
psP<-predict(regps,ps,interval="prediction") #Cálculo direto do intervalo de predição
# Por default o nível de confiança para Intervalos de Confiança Predição é igual a 95

#Intervalos de Tolerância Bilateral - Projeção 20 anos
compalfa<-0.95
gamma<-0.95
sqx<-sum(c((xps-mps)**2))
sqx1<-c((grid-mps)**2)
A<-c((sqx1/sqx))

```

```

n<-length(xps)
N<-c(1/((1/n)+A))
r<-c((qnorm((1+gamma)/2))*(1+1/(2*N))-((2*(qnorm((1+gamma)/2)**2))-3)/
(24*(N**2))))
k<-c(r*sqrt((n-2)/qchisq((1- compalfa),n-2)))
bps<-matrix(c(coefficients(regps)),2,1)
b0ps<-bps[1,1]
b1ps<-bps[2,1]
regps<-summary(regps)
sps<-regps$sigma
LTinfps<-matrix(c(b0ps+b1ps*grid-k*sps),length(grid),1)
LTsupps<-matrix(c(b0ps+b1ps*grid+k*sps),length(grid),1)
LTps<-cbind(LTinfps,LTsupps)

#Intervalo de Tolerância Unilateral - Projeção 20 anos
quantis<-read.table('quantispt_ps.txt',h=F) #Arquivo obtido com a rotina da tabela
2A
quantis<-as.matrix(quantis)
gp<-c(-(sqrt((1/n)+A))*quantis[,1])
LTUinfps<-matrix(c(b0ps+b1ps*grid-gp*sps),length(grid),1)
LTUsupps<-matrix(c(b0ps+b1ps*grid+gp*sps),length(grid),1)
LTUps<-cbind(LTUinfps,LTUsupps)

#Gráfico para comparação entre Intervalos - projeção 20 anos
matplot(grid,cbind(psC,psP[-1],LTUps,LTps),lty=c(1,1,1,1,1,1,1,1),
type="l",col=c(1,2,2,3,3,4,4,6,6),xlab="Ano",ylab="ICVPES",main=" ", cex.axis=1.0,
cex.main=1.5,xlim=c(0,26),ylim=c(-0.25,0.9))
rug(Ano)
abline(h=0,lty=5)
legend(0.2,0.9,c('Reta ajustada','L. de Confiança','L. de Predição', 'L. de Tol.
Unilateral','L. de Tol. Bilateral'),lty=c(1,1,1,1,1),col=c(1,2,3,4,6),cex=1,bty="n")

#2º Exemplo) ICVPT 1º e 2º semestres (para o CVPT é só modificar o arquivo dos
dados e a escala do gráfico)
#Intervalos de Confiança e Predição - Projeção 20 anos - 1º semestre
PMano1<-read.table("novopt1_6.txt",h=T)
reg1<-lm(ICV~0+Ano,data=PMano1)
x1<-Ano
m1<-mean(x1)
PMano2<-read.table("novopt7_12.txt",h=T)
reg2<-lm(ICV~0+Ano,data=PMano2)
x2<-Ano
m2<-mean(x2)
grid<-seq(1,20,1)
ps<-data.frame(Ano=grid)
psC<-predict(reg1,ps,interval="confidence")

```

```

psP<-predict(reg1,ps,interval="prediction")
#Intervalos de Tolerância Bilateral - Projeção 20 anos - 1º semestre
compalfa<-0.95
gamma<-0.95
sqx<-sum(c((x1-m1)**2))
sqx1<-c((grid-m1)**2)
A<-c(sqx1/sqx)
n<-length(x1)
N<-c(1/((1/n)+A))
r<-c((qnorm((1+gamma)/2))*(1+1/(2*N))-((2*(qnorm((1+gamma)/2)**2))-3)/
(24*(N**2)))
k<-c(r*sqrt((n-2/qchisq((1- compalfa),n-2)))
bpt<-matrix(c(coefficients(reg1)),2,1)
bpt0<-bpt[1,1]
bpt1<-bpt[2,1]
reg1<-summary(reg1)
s1<-reg1$sigma
LTinf1<-matrix(c(bpt0+bpt1*grid-k*s1),length(grid),1)
LTsup1<-matrix(c(bpt0+bpt1*grid+k*s1),length(grid),1)
LT1<-cbind(LTinf1,LTsup1)

#Intervalo de Tolerância Unilateral - Projeção 20 anos - 1º semestre
quantis<-read.table('quantispt_ps.txt',h=F) #Arquivo obtido com a rotina da tabela 2A
quantis<-as.matrix(quantis)
gp<-c(-(sqrt((1/n)+A))*quantis[,1])
LTUinf1<-matrix(c(bpt0+bpt1*grid-gp*s1),length(grid),1)
LTUsup1<-matrix(c(bpt0+bpt1*grid+gp*s1),length(grid),1)
LTU1<-cbind(LTUinf1,LTUsup1)

#Intervalos de Confiança e Predição - Projeção 20 anos - 2º semestre
ss<-data.frame(Ano=grid)
ssC<-predict(reg2,ss,interval="confidence")
ssP<-predict(reg2,ss,interval="prediction")

#Intervalos de Tolerância Bilateral - Projeção 20 anos - 2º semestre
b2pt<-matrix(c(coefficients(reg2)),2,1)
b20pt<-b2pt[1,1]
b21pt<-b2pt[2,1]
reg2<-summary(reg2)
s2<-reg2$sigma
LTinf2<-matrix(c(b20pt+b21pt*grid-k*s2),length(grid),1)
LTsup2<-matrix(c(b20pt+b21pt*grid+k*s2),length(grid),1)
LT2<-cbind(LTinf2,LTsup2)
#Intervalo de Tolerância Unilateral - Projeção 20 anos - 2º semestre
LTUinf2<-matrix(c(b20pt+b21pt*grid-gp*s2),length(grid),1)
LTUsup2<-matrix(c(b20pt+b21pt*grid+gp*s2),length(grid),1)

```

```

LTU2<-cbind(LTUinf2,LTUsup2)
#Gráfico para comparação entre Intervalos - projeção 20 anos
par(mfrow=c(1,2))
matplot(grid,cbind(psC,psP[-1],LTU1,LT1),lty=c(1,1,1,1,1,1,1,1),
type="l",col=c(1,2,2,3,3,4,4,6,6),xlab="Ano",ylab="ICVPT",main=" ", cex.axis=1.0,
cex.main=1.5, xlim=c(0,26),ylim=c(-0.35,1.8))
rug(Ano)
abline(h=0,lty=5)
legend(0,1.87,c('Reta ajustada','L. de Confiança','L. de Predição', 'L. de Tol.
Unilateral','L. de Tol. Bilateral'),lty=c(1,1,1,1,1),col=c(1,2,3,4,6),cex=1,bty="n")
matplot(grid,cbind(ssC,ssP[-1],LTU2,LT2),lty=c(1,1,1,1,1,1,1,1),
type="l",col=c(1,2,2,3,3,4,4,6,6),xlab="Ano",ylab="ICVPT",main=" ", cex.axis=1.0,
cex.main=1.5,xlim=c(0,26),ylim=c(-0.35,1.8))
rug(Ano)
abline(h=0,lty=5)
legend(0,1.87,c('Reta ajustada','L. de Confiança','L. de Predição','L. de Tol.
Unilateral','L. de Tol. Bilateral'),lty=c(1,1,1,1,1),col=c(1,2,3,4,6),cex=1,bty="n")

#Observação: Para a obtenção das tabelas 11 (Seção 4.3.2.2), 10B a 15B (Anexo B) e
as figuras 1A a 4A (Anexo A) utilizaram-se os mesmos comandos, com a única
diferença, que o banco de dados é referente ao período de 2000 a 2004, o grid passa a
ser uma seq(1,7) e a o arquivo quantispt_ps.txt é diferente e se refere aos dados da
tabela 3B.

```

#### **Rotinas para o quadro 4 (seção 4.4), para a figura 17 (seção 4.4) e para a tabela 16B (Anexo B)**

```

#Análise Exploratória do ICV de 2003 dos municípios com nº de habitantes entre
14.000 e 17.500 - figura 15
ICV2003<-read.table('ICV 2003.txt',h=T)
ICV2003<-as.matrix(ICV2003)
shapiro.test(ICV2003)
hist(ICV2003,xlab="ICV",ylab="Número de municípios",main=" ", cex.axis=1.0,
cex.main=1.5,xlim=c(0,2.2))

#Escolha de 30 dos 50 municípios classificados em ordem crescente pelo nº de
habitantes – tabela 16A
#Geração de números aleatórios
IE<-sample(ICV2003,30)
IE<-sort(IE)
write(IE,file="IE.txt") #arquivo para ser utilizado na construção dos intervalos
estatísticos

#Intervalos Estatísticos – quadro 3
# Análises preliminares - amostra de tamanho 30
IE<-read.table('IE.txt',h=T)

```

```

IE<-as.matrix(IE)
media<-mean(IE)
desvpad<-sqrt(var(IE))
n<-length(IE)
v<-n-1

# Intervalo de Confiança Bilateral para a média
compalfa<-0.95
t<-qt(((1+compalfa)/2),v)
k<-desvpad/sqrt(n)
LCmedinf<-media-(t*k)
LCmedsup<-media+(t*k)
LCmedinf
LCmedsup

#Intervalo de Predição Bilateral para conter 1 observação futura
t<-qt(((1+compalfa)/2),v)
k<-t*((1+1/n)^0.5)
LPinf<-media-k*desvpad
LPsup<-media+k*desvpad
LPinf
LPsup

#Intervalo de Predição Bilateral para conter a média de "m" futuras observações
m<-30
t<-qt(((1+compalfa)/2),v)
k<-t*((1/m+1/n)^0.5)
LPmedinf<-media-k*desvpad
LPmedsup<-media+k*desvpad
LPmedinf
LPmedsup

#Intervalo de Tolerância Bilateral - Aproximação de Wald and Wolfowitz
# Cálculo do fator de tolerância – Método iterativo
gamma<-0.95
ri<-qnorm(gamma)/2
rf<-2.6
v<-n-1
a<-c(seq(ri,rf,by=0.0001))
r=ri
cont=0
m=matrix(0,length(a),3)
while (r<=rf+0.00001)
{
x<-pnorm(1/sqrt(n)-r)
x1<-pnorm(1/sqrt(n)+r)

```

```

y<-x1-x
cont=cont+1
m[cont,1]=y
m[cont,2]=r
m[cont,3]=abs(y-gamma)
r=r+0.0001
}
for (i in 1:nrow(m))
{
if (m[i,3]<=0.00001)
{
yfinal=m[i,1]
rfinal=m[i,2]
} }

#Cálculo dos Limites de Tolerância
b<-qchisq(1-compalfa,v)
c<-sqrt(v/b)
lambada<-rfinal*c
LTinf<-media-(lambada*desvpad)
LTsup<-media+(lambada*desvpad)
lambada
LTinf
LTsup
rfinal
yfinal

```

### **Rotina para tabela 2B (AnexoB)**

```

#Cálculo dos quantis da distribuição t não central (1° e 2° semestres) – Método
iterativo
PMano1<-read.table("regano1_6.txt",h=T)
attach(PMano1)
x1<-Ano
m1<-mean(x1)
prob=0
quantil=0
grid<-seq(1,20 # Tamanho da projeção
sqx<-sum(c((x1-m12))
sqx1<-c((grid-m12)
n<-length(x1)
alfa<-0.95
gamma<-0.95
A<-sqrt((1/n)+c(sqx1/sqx))
Ngamma<-qnorm(gamma)

```

```

delta<-c(-Ngamma/A)
delta<-as.matrix(delta)
ti<--12
tf<--0.5
v<-n-2
a<-c(seq(ti,tf,by=0.0001))
t=ti
resquantis=matrix(0,nrow(delta),2)

for (u in 1:nrow(delta))
{
auxdelta=delta[u,1]
while (t<=tf+0.00001)
{
x<-pt(t,v,auxdelta)
dif=abs(x-(1-alfa))
if (dif<=0.000001)
{
prob=x
quantil=t
}
t=t+0.0001
}
resquantis[u,1]=prob
resquantis[u,2]=quantil
prob=0
quantil=0
t=ti
}
quantis<-resquantis[,2] # Resultado dos quantis para os respectivos parâmetros de não
centralidade#
write(quantis,file="quantispt_ps.txt") # Salva esses resultados num documento tipo
texto no diretório que estiver sendo trabalhado#

#Observação; Para a obtenção da tabela 3B (Anexo B) utilizaram-se os mesmos
comandos, com a única diferença de que o banco de dados é referente ao período de
2000 a 2004 e o grid passa a ser uma seq(1,7.)

```