**CONTRIBUIÇÕES À TEORIA DA
PREVISÃO EM POPULAÇÕES FINITAS**

Mônica Carneiro Sandoval

**TESE DEFENDIDA
NO
INSTITUTO DE MATEMÁTICA E ESTATÍSTICA
DA
UNIVERSIDADE DE SÃO PAULO
PARA OBTENÇÃO DO GRAU DE DOUTOR
EM
ESTATÍSTICA**

Área de concentração: Estatística

Orientador: Prof. Dr. Heleno Bolfarine

– São Paulo, setembro de 1993 –

*À minha filha
Mariana*

AGRADECIMENTOS

Esta tese é a realização de um ideal. Para alcançá-lo, foi preciso muito esforço, muita dedicação e muita perseverança. Mas todo meu trabalho não teria sido possível sem ajuda, sem incentivo, sem companheirismo.

Por isso, agradeço,

ao Professor Heleno Bolfarine, que, mais uma vez, me orientou com dedicação e empenho, transmitindo-me seus conhecimentos e respondendo com presteza às minhas solicitações,

ao Professor Victor Hugo Salinas Torres, por suas valiosas sugestões na elaboração do Capítulo 2,

ao Professor José Galvão Leite, por sua contribuição na fase inicial do trabalho,

ao Walter V. Fernandes, responsável pela digitação.

Estendo meus agradecimentos a todos aqueles que colaboraram de uma maneira ou de outra para a realização deste trabalho.

RESUMO

Dedicamos a primeira parte desta tese à previsão ótima da função distribuição sob modelos de superpopulação Gaussianos. Derivamos o previsor ótimo e sua distribuição assintótica sob o modelo de locação, e para alguns modelos particulares é obtida uma aproximação assintótica para a variância preditiva dos previsores ótimos. Estudos de Monte Carlo ilustram comparações entre os previsores ótimos e alguns outros propostos na literatura.

Na segunda parte desta tese, consideramos o problema de previsão em populações finitas sob modelos de superpopulação com erros nas variáveis. Inicialmente, assumindo o modelo de locação com erro de mensuração, são obtidos os previsores Bayesianos e Bayesianos empíricos do total e da variância populacionais. Extensões são feitas para o modelo de superpopulação “não homogêneo” considerado por Tiwari e Lahiri (1989). Alguns estudos de simulação que ilustram o efeito do erro de medida nos previsores Bayesianos empíricos propostos, também são apresentados. Posteriormente, assumindo que o modelo com erro de mensuração é um modelo de regressão linear estrutural, consideramos previsores do tipo regressão para o total populacional e encontramos suas distribuições assintóticas. Ainda sob esse modelo, mas adotando o enfoque condicional com parametrização ortogonal, derivamos o previsor de mínimos quadrados condicional e o previsor Bayesiano sob uma priori não informativa para os parâmetros ortogonais. Para finalizar, estendemos os modelos considerados até então, considerando modelos de superpopulação com erros nas variáveis contendo várias variáveis regressoras. Previsores do total populacional são propostos e seus comportamentos assintóticos estudados.

ABSTRACT

We dedicate the first part of this thesis to optimal prediction of the finite population distribution function under Gaussian superpopulation models. We derive the optimal predictor and its asymptotic distribution under a Gaussian location model. For some particular models, large sample approximations for the predictive variances of the optimal predictors are derived. Monte Carlo studies illustrate comparisons between the optimal predictors and some other predictors proposed in the literature.

In the second part of this thesis, we consider the prediction problem in finite populations under error in variables superpopulation models. First, assuming the location measurement error model, the Bayes predictors and the empirical Bayes predictors of the population total and population variance are obtained. Extensions are made to the nonhomogeneous superpopulation model considered by Tiwari and Lahiri (1989). Some simulation studies that illustrate the effect of the measurement errors in the proposed empirical Bayes predictors are also reported. Later, assuming that the measurement error model is a structural linear regression model, we consider regression predictors for the population total and find their asymptotic distributions. Yet, under this model, but adopting the conditional approach with orthogonal parametrization we derive the conditional minimum square error predictor and the Bayesian predictor under a noninformative prior to the orthogonal parameters. To conclude, we extend the models considered up to now, considering error in variables superpopulation models with various regression variables. Predictors of the population total are proposed and their asymptotic behaviour studied.

ÍNDICE

CAPÍTULO 1	INTRODUÇÃO	1
CAPÍTULO 2	PREVISÃO DA FUNÇÃO DISTRIBUIÇÃO DE UMA POPULAÇÃO FINITA	8
2.1	Introdução	8
2.2	Previsores ótimos de $F_N(t)$	11
2.3	σ^2 conhecido	13
2.3.1	O modelo de locação simples	14
2.3.2	O modelo de regressão linear	17
2.4	σ^2 desconhecido	21
2.4.1	O modelo de locação simples	22
2.4.2	O modelo de regressão linear	26
2.5	Estudos de simulação	30
CAPÍTULO 3	PREVISÃO BAYESIANA EMPÍRICA SOB UM MODELO DE LOCAÇÃO COM ERRO NAS VARIÁVEIS	40
3.1	Introdução	40
3.2	Previsão Bayesiana de T e S^2	43
3.3	Previsão Bayesiana empírica	45
3.4	Risco de Bayes de $\hat{T}_{EB}$ e $\hat{S}_{EB}^2$	49
3.5	O modelo de locação “não homogêneo” com erro nas variáveis	57
3.5.1	Previsão Bayesiana de T e S^2	58
3.5.2	Previsão Bayesiana empírica	61
3.5.3	Risco de Bayes de $\hat{T}_{EB}$ e $\hat{S}_{EB}^2$	66
3.6	Ilustração numérica	77

CAPÍTULO 4	PREVISÃO DO TOTAL POPULACIONAL SOB UM	
	MODELO DE REGRESSÃO COM ERRO NAS VARIÁVEIS	.. 81
4.1	Introdução	81
4.2	Previsores do tipo regressão	83
4.3	O enfoque condicional	91
4.3.1	O previsor de mínimos quadrados condicional	92
4.3.2	O previsor Bayesiano	94
4.4	O modelo com várias variáveis auxiliares	97
CONCLUSÕES		103
APÊNDICE 1		105
APÊNDICE 2		110
APÊNDICE 3		116
BIBLIOGRAFIA		118

CAPÍTULO 1

INTRODUÇÃO

O problema de previsão em populações finitas tem sido estudado até agora fundamentalmente através de dois enfoques: amostragem clássica e modelos de superpopulação.

Na amostragem clássica os métodos de estimação não são construídos sobre a base da inferência estatística; toda inferência está baseada exclusivamente na distribuição de probabilidades criada pelo planejamento amostral. Sob esta abordagem, a única relação existente entre as unidades na amostra e fora dela é que há uma probabilidade positiva de que todas pertençam à amostra, o que não parece ser suficiente para que inferências sobre a parte desconhecida da população sejam baseadas na parte observada. Segundo alguns autores, a freqüente utilização desse princípio deve-se ao fato de que a qualidade da inferência depende apenas do processo de seleção que é criado e controlado pelo pesquisador. Nas últimas décadas diversos pesquisadores (por exemplo, Godambe (1955, 1966)

e Basu (1969, 1971)) vêm procurando estudar a amostragem clássica do ponto de vista da inferência estatística, mostrando as limitações dos modelos probabilísticos sob os quais essa teoria se apoia. Estabeleceu-se que na amostragem clássica não existe estimador não viciado e de variância mínima da média populacional na classe dos estimadores lineares. Derivou-se a função de verossimilhança e mostrou-se que esta é não informativa.

A abordagem de modelos de superpopulação assume que os valores desconhecidos dos elementos da população são realizações de variáveis aleatórias e admite a existência de um modelo relacionando as variáveis de interesse. O modelo adotado é a ligação natural entre as unidades fora da amostra e as unidades amostradas. Sob essa abordagem, o mecanismo de seleção da amostra não tem nenhuma participação na inferência, ao contrário da abordagem clássica, onde a inferência é impossível sem a aleatorização. O uso dessa abordagem permite, então, uma análise estatística para amostras não aleatórias e possibilita tomarmos intencionalmente uma amostra satisfazendo certas condições desejáveis. Na abordagem de superpopulação os resultados obtidos dependem exclusivamente da amostra observada e da validade do modelo adotado, de onde surgem algumas críticas, já que a validade do modelo nunca pode ser perfeitamente comprovada na prática.

Um número grande de artigos sobre previsão em populações finitas apareceram na literatura nas últimas décadas. Podemos citar alguns artigos anteriores a 1970, mas os trabalhos que praticamente introduziram o enfoque de superpopulação foram Royall (1970) e Royall e Herson (1973a, b).

Smith (1976) faz uma revisão de todo o desenvolvimento histórico do problema de inferência em populações finitas, tratando de ambas as abordagens, a clássica e a de modelos de superpopulação, concluindo que a inferência em populações finitas deve dar mais atenção à estrutura da população e, menos ao planejamento amostral.

Um trabalho de grande importância é Cassel, Sarndall e Wretman (1977) por ter sido o primeiro texto de inferência em populações finitas que desenvolve a abordagem de

superpopulação.

Sob essa abordagem, toda inferência é feita com base no modelo adotado, de modo que a má especificação do modelo pode acarretar sérios erros. Assim, muitos autores se preocupam com o problema de robustez para garantirem a qualidade da inferência, quando o modelo adotado apresentar falhas em alguma de suas suposições.

Royall e Herson (1973a) derivam o previsor linear ótimo para o total populacional sob o modelo de regressão polinomial e introduzem condições de robustez para que o previsor mantenha sua otimalidade sob um modelo polinomial de grau superior. Royall e Herson (1973b) consideram a estratificação como uma outra técnica para proteger contra falhas no modelo. Scott, Brewer e Ho (1978) estendem os resultados de Royall e Herson (1973a), usando uma função de variância geral. Tallis (1978) prova uma condição necessária e suficiente para a robustez do previsor expansão. Sob o modelo linear geral, Pereira e Rodrigues (1983) encontram o previsor linear ótimo de uma função linear qualquer das unidades populacionais e estudam sua otimalidade sob modelos contendo outras variáveis auxiliares, além das presentes no modelo inicial. Rodrigues, Bolfarine e Rogatko (1985) consideram a suposição adicional de normalidade e obtêm previsores ótimos para funções lineares e quadráticas das unidades populacionais, em particular para o total e para a variância populacionais. Considerando o modelo linear geral com matriz de covariâncias não diagonal, Royall (1976a) obtém o previsor linear ótimo do total populacional e, supondo um modelo de superpopulação Gaussiano, Tam (1987) mostra que esse previsor é não viciado e de variância mínima. O problema de estimação da variância preditiva dos previsores do total populacional é discutido por Royall e Eberhart (1975), Royall e Cumberland (1978) e (1981a, b).

Um dos poucos artigos sobre função de verossimilhança em populações finitas é Royall (1976b) que trata da previsão de funções lineares das unidades populacionais, assumindo um modelo de superpopulação normal. Zacks e Bolfarine (1990b) derivam previsores de

máxima verossimilhança em populações finitas, mas adotam uma definição diferente de Royall (1976b).

De acordo com Sarndal e Wright (1984) é importante que os previsores tenham formas simples e intuitivas. Um predictor simples e intuitivo para o total populacional é o predictor projeção simples, e Bolfarine e Rodrigues (1988) apresentam condições sob as quais este predictor coincide com o predictor linear ótimo do total populacional.

Previsores equivariantes do total e da variância populacionais são obtidos por Bolfarine (1989) e Zacks e Bolfarine (1990a).

Um dos primeiros trabalhos que introduziram a perspectiva Bayesiana dentro do enfoque de superpopulação, foi Ericson (1969). O problema de previsão do total populacional, inclusive considerando que o modelo assumido possa estar com erro, é abordado em Royall e Pfefferman (1982), Smouse (1984), Bolfarine, Pereira e Rodrigues (1987), Rodrigues (1988), Ericson (1988). Para a variância populacional, os trabalhos mais recentes do ponto de vista Bayesiano são Zacks (1981), Zacks e Solomon (1981) e Bolfarine, Pereira e Rodrigues (1987), entre outros.

A previsão Bayesiana empírica do total e da variância populacionais é considerada por Ghosh e Meeden (1986), Ghosh e Lahiri (1987a e b), Tiwari e Lahiri (1989) e Lahiri e Peddada (1992).

A previsão do coeficiente de regressão populacional também tem sido estudada na literatura por vários autores, dentre os quais destacamos, Fuller (1975), Hartley e Sielken (1975) e Sarndal (1982), Elian (1991).

Uma referência completa que trata do problema de previsão em populações finitas sob o enfoque de superpopulação, é Bolfarine e Zacks (1992).

Na teoria de populações finitas em geral, é dada ênfase à estimação (ou previsão) da média, do total e da variância. Livros básicos de amostragem clássica (Cochran (1977), Kish (1965)) quase não abordam o problema de estimação da função distribuição popu-

lacional.

Como vemos, a previsão do total populacional e da variância populacional através da abordagem de modelos de superpopulação, também tem sido muito estudada. No entanto, na prática, a estimação da função distribuição é um objetivo importante, particularmente quando o principal interesse da pesquisa é identificar subgrupos na população cujos valores para variáveis particulares caem bem abaixo ou bem acima da média populacional. Então, recentemente, diversos autores têm-se dedicado à previsão da função distribuição de uma população finita sob diferentes enfoques. Entre outros, podemos citar, Sedransk e Sedransk (1979), Cohen e Kuo (1985a e b), Francisco e Fuller (1986), Chambers e Dunstan (1986), Kuk (1988), Kuo (1988), Dunstan e Chambers (1989), Rao et al. (1990), Chambers, Dorfman e Hall (1992), Dorfman (1993), Chambers, Dorfman e Wehrly (1993). Portanto, uma das nossas contribuições é a obtenção de previsores ótimos da função distribuição populacional sob a abordagem de modelos de superpopulação.

Dedicamos, então, o Capítulo 2 à previsão ótima da função distribuição populacional sob um modelo de superpopulação normal. Inicialmente, consideramos conhecida a matriz de covariâncias e, adotando o modelo de locação, encontramos o previsor ótimo, sua distribuição assintótica e uma expressão para a eficiência relativa assintótica da função distribuição empírica em relação ao previsor ótimo. O previsor ótimo também é obtido sob o modelo de regressão linear, e para alguns modelos particulares é obtida uma aproximação assintótica para sua variância preditiva. Posteriormente, realizamos um estudo semelhante a esse considerando desconhecida a matriz de covariâncias. Finalizando o capítulo, apresentamos alguns resultados de estudos de simulação, comparando os previsores ótimos obtidos com alguns previsores propostos na literatura.

Os modelos de superpopulação, considerados nos artigos citados, assumem que as variáveis são medidas sem erro. Na prática, não é realista supor que as variáveis não apresentam erros de mensuração, e o mais comum é não termos acesso aos seus verdadeiros

valores. Logo, essa suposição é muitas vezes violada, particularmente em ciências sociais e biológicas.

Existem muitos trabalhos sobre o problema de estimação de parâmetros, quando algumas das variáveis envolvidas contêm erros de medida. Entre outros, podemos citar Sprent (1966) que propõe um método baseado no método de mínimos quadrados generalizados, para estimação dos coeficientes de regressão linear. Lindley (1966) e Lindley e El-Sayad (1968) foram os primeiros a apresentar uma análise Bayesiana para esse problema. Outros trabalhos Bayesianos podem ser encontrados em Zellner (1971) e Reilly e Patino-Leal (1981). Kendall e Stuart (1961), Moran (1971) e Fuller (1987) apresentam bibliografias extensas sob os modelos de regressão linear funcional e estrutural. Fuller (1987) escreveu o livro com o objetivo de aumentar o uso de técnicas estatísticas sob a presença de erros de mensuração, incluindo um grande número de exemplos reais. Wong (1989) discute e compara métodos baseados na verossimilhança, a fim de obter intervalos de confiança para o coeficiente angular no modelo de regressão linear, quando ambas as variáveis são medidas com erro. Rodrigues e Cordani (1990) e Bolfarine e Cordani (1993), utilizando parametrização ortogonal, obtêm estimadores de máxima verossimilhança dos parâmetros do modelo, quando o coeficiente de confiabilidade (Fuller (1987)) é suposto conhecido.

De acordo com Fuller (1975), existem poucos trabalhos de previsão em populações finitas sob modelos com erros nas variáveis. Recentemente, Bolfarine e Rodrigues (1990) consideraram o problema de previsão em populações finitas sob o modelo de superpopulação funcional, adotando uma perspectiva Bayesiana. Bolfarine (1990) considera como modelo de superpopulação o modelo de regressão onde ambas as variáveis são medidas com erro, estuda as propriedades de alguns previsores clássicos e propõe previsores Bayesianos para o total e para variância populacionais.

Nos Capítulos 3 e 4 pretendemos contribuir com o problema de previsão em populações finitas, considerando modelos com erros de mensuração. No Capítulo 3, assumimos o mo-

delo de locação com erro nas variáveis, encontramos os previsores Bayesianos para o total e a variância populacionais, adotando uma priori informativa e, a partir daí, consideramos a previsão Bayesiana empírica, utilizando o enfoque proposto por Ghosh e Meeden (1986). Ainda nesse capítulo, os previsores Bayesianos e Bayesianos empíricos são obtidos para o caso do modelo de locação “não homogêneo” com erro nas variáveis. Para finalizar, alguns estudos de simulação são feitos. No Capítulo 4, o interesse é a previsão do total populacional sob um modelo de regressão linear estrutural (Fuller (1987)). Consideramos previsores do tipo regressão, encontramos suas distribuições assintóticas e a seguir apresentamos algumas comparações entre os previsores considerados. Posteriormente, adotamos o enfoque condicional (Bolfarine e Rodrigues (1990)) com parametrização ortogonal para obter o preditor de mínimos quadrados condicional e o preditor Bayesiano sob uma priori não informativa para os parâmetros ortogonais. Consideramos modelos de superpopulação com erros nas variáveis, contendo várias variáveis regressoras, estendendo assim os modelos considerados até então. Previsores do total populacional são propostos para esses casos, e o comportamento assintótico de tais previsores é estudado.

Finalmente, apresentamos nossas conclusões e sugerimos diversas pesquisas que ainda poderão ser desenvolvidas.

CAPÍTULO 2

PREVISÃO DA FUNÇÃO DISTRIBUIÇÃO DE UMA POPULAÇÃO FINITA

2.1 INTRODUÇÃO

Neste capítulo consideramos a previsão ótima da função distribuição de uma população finita sob modelos de regressão linear com erros distribuídos normalmente.

Consideremos uma população finita $\mathcal{P} = \{1, \dots, N\}$, onde N é conhecido. Associadas com a k -ésima unidade de $\mathcal{P}$ existem $p + 1$ quantidades $y_k, x_{k1}, \dots, x_{kp}$ que são todas conhecidas, exceto y_k , $k = 1, \dots, N$. A quantidade y_k é considerada como sendo uma realização de uma variável aleatória Y_k , $k = 1, \dots, N$. Como y_k e Y_k são desconhecidas, não faremos distinção entre elas e usaremos sempre y_k para representar uma ou outra. Seja $\mathbf{y} = (y_1, \dots, y_N)'$ e $\mathbf{X} = (\mathbf{X}_1, \dots, \mathbf{X}_N)'$ uma matriz conhecida $N \times p$, onde $\mathbf{X}_j = (x_{j1}, \dots, x_{jp})'$, $j = 1, \dots, N$. Relacionando os dois conjuntos de variáveis $\mathbf{y}$ e $\mathbf{X}$,

consideramos o modelo linear

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \mathbf{e} , \quad (2.1)$$

onde $E[\mathbf{e}] = \mathbf{0}$, $\text{Var}[\mathbf{e}] = \sigma^2 \mathbf{W}$, sendo $\mathbf{W}$ uma matriz diagonal conhecida. Seja $\psi = (\boldsymbol{\beta}, \sigma^2)$. Depois de selecionada a amostra s de tamanho n , podemos reescrever $\mathbf{y}$, $\mathbf{X}$ e $\mathbf{W}$ de forma que

$$\mathbf{y} = \begin{pmatrix} \mathbf{y}_s \\ \mathbf{y}_r \end{pmatrix} , \quad \mathbf{X} = \begin{pmatrix} \mathbf{X}_s \\ \mathbf{X}_r \end{pmatrix} , \quad \mathbf{W} = \begin{pmatrix} \mathbf{W}_s & \mathbf{0} \\ \mathbf{0} & \mathbf{W}_r \end{pmatrix} ,$$

onde $\mathbf{y}_s$ corresponde aos elementos de $\mathbf{y}$ que foram observados na amostra e $\mathbf{y}_r$ aos que não pertencem à amostra, $\mathbf{X}_s$ é a matriz $n \times p$ de variáveis auxiliares e $\sigma^2 \mathbf{W}_s$ é a matriz $n \times n$ de covariâncias, ambas associadas aos elementos da amostra, $\mathbf{X}_r$ e $\sigma^2 \mathbf{W}_r$ são as correspondentes matrizes de variáveis auxiliares e de covariâncias relacionadas aos elementos que estão fora da amostra.

A função distribuição de uma população finita é definida por

$$F_N(t) = \frac{1}{N} \sum_{j=1}^N \Delta(t - y_j) ,$$

onde

$$\Delta(x) = \begin{cases} 1 & , x \geq 0 \\ 0 & , \text{caso contrário} . \end{cases}$$

A previsão de $F_N(t)$ tem sido objeto de estudo de muitos artigos na literatura estatística recente. Sedransk e Sedransk (1979) mostram o uso da função distribuição empírica para planejamentos estratificados. Cohen e Kuo (1985a e b) estudam propriedades da função distribuição empírica do ponto de vista da teoria da decisão. Mostram que a função distribuição empírica é admissível e obtêm estratégias minimax para estimar a função de distribuição populacional. Algumas propriedades assintóticas são estudadas por Francisco e Fuller (1986). A função distribuição empírica, estimador usual de $F_N(t)$ sob amostragem clássica, não utiliza informação auxiliar no estágio de estimação, como os estimadores razão e regressão, na estimação do total populacional. Sob o enfoque de superpopulação, Chambers e Dunstan (1986) propõem um previsor para $F_N(t)$

que incorpora diretamente no processo de estimação informação auxiliar, suposta conhecida para todos os elementos da população. Posteriormente, Dunstan e Chambers (1989) propõem extensões para o caso em que a informação auxiliar é limitada. Sob a abordagem clássica, Rao et al. (1990) obtêm estimadores razão e diferença para $F_N(t)$ para incorporar informação fornecida por variáveis auxiliares. Recentemente, Chambers et al. (1992), Dorfman (1993), fazem comparações dos estimadores propostos por Chambers e Dunstan (1986) e Rao et al. (1990). Chambers, Dorfman e Wehrly (1993) definem um previsor robusto para $F_N(t)$.

Derivamos, neste capítulo, previsores ótimos para $F_N(t)$ sob o modelo (2.1), com a suposição adicional de que o vetor e é normalmente distribuído e estudamos alguns resultados assintóticos desses previsores. Mesmo em populações finitas existem muitos trabalhos sobre teoria assintótica. Sob a abordagem clássica, Cochran (1977) utiliza, na construção de intervalos de confiança, a aproximação normal para a média amostral, enquanto que Scott e Wu (1981) apresentam resultados assintóticos importantes para os estimadores razão e regressão. Sob a abordagem de modelos de superpopulação, podemos citar Fuller (1975, 1987), Bolfarine e Zacks (1992) e Elian (1991).

Em populações finitas temos que considerar de que forma o tamanho da amostra, n , e o tamanho da população, N , tendem ao infinito. Fuller (1987) e Bolfarine e Zacks (1992) trabalham com uma seqüência de populações

$$\{\mathcal{P}_\nu, \nu = 1, 2, \dots\},$$

onde $\mathcal{P}_\nu = \{1, 2, \dots, N_\nu\}$ corresponde à ν -ésima população satisfazendo $N_\nu \leq N_{\nu+1}$. Dessa ν -ésima população, é selecionada uma amostra de tamanho n_ν ($n_\nu < n_{\nu+1}$) e quando $\nu \rightarrow \infty$ (ν tende a infinito), $N_\nu \rightarrow \infty$, $n_\nu \rightarrow \infty$, de modo que $N_\nu - n_\nu \rightarrow \infty$ e $f_\nu = \frac{n_\nu}{N_\nu} \rightarrow f$, com $0 < f < 1$. Esse é o conceito de convergência em populações finitas que adotamos neste capítulo.

Na seção 2.2 apresentamos alguns resultados gerais sobre a previsão de $F_N(t)$. Na

seção 2.3 consideramos a previsão de $F_N(t)$ sob o modelo (2.1) com σ^2 conhecido. Inicialmente, derivamos o melhor predictor não viciado e sua distribuição assintótica sob o modelo de locação simples. Derivamos também uma expressão para a eficiência relativa assintótica da função distribuição empírica em relação ao predictor ótimo. Posteriormente, consideramos o modelo de regressão linear e encontramos o predictor ótimo para $F_N(t)$. Uma aproximação assintótica para a sua variância preditiva foi obtida para alguns modelos particulares. Previsão ótima de $F_N(t)$ sob o modelo (2.1) com σ^2 desconhecido foi considerada na seção 2.4. Na seção 2.5 apresentamos os resultados de estudos de Monte Carlo baseados em populações simuladas, para ilustrar o comportamento dos predictors ótimos em relação a alguns predictors propostos na literatura.

2.2 PREVISORES ÓTIMOS DE $F_N(t)$

Nesta seção apresentamos resultados gerais sobre a previsão de $F_N(t)$ sob o modelo (2.1).

Como foi considerado em Rodrigues et al. (1985), depois que a amostra $s \subset \mathcal{P}$ foi selecionada, geralmente é possível expressar a quantidade de interesse $\theta(\mathbf{y})$ como

$$\theta(\mathbf{y}) = \theta_s + \theta_{sr} ,$$

onde θ_s é uma função de $\mathbf{y}_s = (y_i, i \in s)'$ e θ_{sr} é uma função de $\mathbf{y}_s$ e $\mathbf{y}_r = (y_i, i \notin s)'$.

Neste capítulo, a quantidade de interesse é

$$\theta(\mathbf{y}) = F_N(t) ,$$

sendo

$$\theta_s = \frac{n}{N} \sum_{i \in s} \frac{\Delta(t - y_i)}{n}$$

e

$$\theta_{sr} = \left(1 - \frac{n}{N}\right) \sum_{i \notin s} \frac{\Delta(t - y_i)}{N - n} .$$

Logo, sendo os valores y_i observados para a amostra s , θ_s é conhecido, e o problema de previsão de $F_N(t)$ se resume em prever θ_{sr} , ou seja, a parcela correspondente às unidades fora da amostra.

Os previsores de $F_N(t)$ que consideramos, são da forma

$$\hat{\theta} = \hat{F}_N(t) = \frac{n}{N} F_{N_s}(t) + \hat{\theta}_{sr} ,$$

onde

$$F_{N_s}(t) = \frac{1}{n} \sum_{i \in s} \Delta(t - y_i) \quad (2.2)$$

é a função distribuição empírica e $\hat{\theta}_{sr}$ é um predictor de θ_{sr} baseado na amostra y_s .

De acordo com Rodrigues et al. (1985), o predictor $\hat{F}_N(t)$ de $F_N(t)$ é ψ -não viciado sob o modelo (2.1) se

$$E_\psi[\hat{F}_N(t) - F_N(t)] = 0 \quad (2.3)$$

para todo ψ . Além disso, o predictor $\hat{F}_N(t)$ é ψ -ótimo se tiver o menor erro quadrático médio, ou seja,

$$E_\psi[\hat{F}_N(t) - F_N(t)]^2 \leq E_\psi[\hat{F}_N^*(t) - F_N(t)]^2$$

para qualquer predictor ψ -não viciado $\hat{F}_N^*(t)$.

Considerando o modelo de superpopulação (2.1) com a suposição adicional de que os erros se distribuem normalmente, temos que a distribuição conjunta de $\mathbf{y}$ pertence a uma família paramétrica que admite estatísticas preditivas suficientes e completas (como será definida abaixo), e então podemos obter previsores ótimos, isto é, o melhor predictor não viciado de $F_N(t)$ pode ser derivado. O método utilizado é uma extensão dos Teoremas Rao-Blackwell e Lehmann-Scheffé (Bolfarine e Zacks, 1992).

Definição 2.1. Seja $\mathcal{F} = \{F_\psi; \psi \in \Psi\}$ uma família de distribuições conjuntas de $\mathbf{y}' = (y'_s, y'_r)$ indexada pelo parâmetro ψ no espaço Ψ . A estatística $S = S(\mathbf{y}_s)$ é uma estatística totalmente suficiente em relação a $\mathcal{F}$ se:

(i) a distribuição condicional de $\mathbf{y}_s$ dado S é independente de ψ ;

(ii) $\mathbf{y}_s$ e $\mathbf{y}_r$ são condicionalmente independentes, dado S .

A estatística totalmente suficiente é completa se a família induzida $\mathcal{F}^S$ das distribuições amostrais de S é completa (Lehmann, 1983, pg. 46).

O teorema a seguir é uma consequência de um resultado mais geral que aparece em Rodrigues et al. (1985).

Teorema 2.1. *Seja S uma estatística totalmente suficiente e completa. Seja $\hat{F}_N(t) = \frac{n}{N}F_{N_s}(t) + \hat{\theta}_{sr}$ um predictor não viciado de $F_N(t)$. Então, sob o modelo (2.1), o melhor predictor não viciado (BUP) de $F_N(t)$ é*

$$\hat{F}_{BU}(t) = \frac{n}{N}F_{N_s}(t) + \hat{\theta}_{sr}(S) ,$$

onde $F_{N_s}(t)$ é dado em (2.2) e

$$\hat{\theta}_{sr}(S) = E_{\psi}[\hat{\theta}_{sr}|S] .$$

Um predictor não viciado de $F_N(t)$ tal que $\hat{\theta}_{sr}$ depende da amostra $\mathbf{y}_s$ somente através de S , é, de acordo com o Teorema 2.1, o BUP. Como no modelo (2.1) a matriz W é diagonal, se S é suficiente é também totalmente suficiente, pois a condição (ii) da Definição 2.1 é satisfeita.

2.3 σ^2 CONHECIDO

Nesta seção consideramos o modelo (2.1) com a suposição adicional de que os erros se distribuem normalmente e vamos supor que σ^2 é conhecido. Nesse caso, a estatística

totalmente suficiente e completa é

$$\hat{\beta}_s = (\mathbf{X}'_s \mathbf{W}_s^{-1} \mathbf{X}_s)^{-1} \mathbf{X}'_s \mathbf{W}_s^{-1} \mathbf{y}_s \quad (2.4)$$

e os previsores derivados nesta seção são baseados em $\hat{\beta}_s$. Inicialmente, trabalhamos sob o modelo de locação simples e depois, então, sob o modelo de regressão linear.

2.3.1 O MODELO DE LOCAÇÃO SIMPLES

O modelo de locação simples segue de (2.1), sendo

$$\mathbf{X} = \mathbf{1}_N \quad \text{e} \quad \mathbf{W} = \mathbf{I}_N, \quad (2.5)$$

onde $\mathbf{1}_N$ é um vetor de uns de dimensão N e $\mathbf{I}_N$ é a matriz identidade N -dimensional. Nesse caso, β é um parâmetro de locação desconhecido. De acordo com (2.4), a estatística totalmente suficiente e completa é $\hat{\beta}_s = \bar{y}_s$ (média da amostra) e o preditor $\hat{F}_{BU}(t)$ será baseado nela.

Teorema 2.2. *Sob o modelo especificado por (2.5) com erros distribuídos normalmente,*

$$\hat{F}_{BU}(t) = \frac{n}{N} F_{N_s}(t) + \left(1 - \frac{n}{N}\right) \Phi \left(\sqrt{\frac{n}{n-1}} \left(\frac{t - \bar{y}_s}{\sigma} \right) \right), \quad (2.6)$$

onde $\Phi(\cdot)$ é a função distribuição da $N(0,1)$.

Prova. Pelo Teorema 2.1 é suficiente mostrarmos que $\hat{F}_{BU}(t)$ é um preditor não viciado de $F_N(t)$, já que $\hat{\theta}_{sr}$ é função da estatística totalmente suficiente e completa $\bar{y}_s$. Mas, de acordo com Lehmann (1983) ou Olkin e Ghurye (1969), temos

$$E_\psi \left[\Phi \left(\sqrt{\frac{n}{n-1}} \left(\frac{t - \bar{y}_s}{\sigma} \right) \right) \right] = E_\psi [\Delta(t - y_i)] = \Phi \left(\frac{t - \beta}{\sigma} \right),$$

o que prova que $\hat{F}_{BU}(t)$ é não viciado, isto é, satisfaz (2.3).

No próximo teorema encontramos a distribuição assintótica do preditor ótimo, $\hat{F}_{BU}(t)$, considerando o conceito de convergência apresentado na seção 2.1.

Teorema 2.3. Considere uma seqüência de populações finitas $\{\mathcal{P}_\nu, \nu \geq 1\}$, satisfazendo $N_\nu \leq N_{\nu+1}$. Uma amostra de tamanho n_ν ($n_\nu < n_{\nu+1}$) é selecionada da população $\mathcal{P}_\nu$, de modo que, quando $\nu \rightarrow \infty$, $N_\nu - n_\nu \rightarrow \infty$ e $f_\nu = \frac{n_\nu}{N_\nu} \rightarrow f$, onde $0 < f < 1$. Então, quando $\nu \rightarrow \infty$,

$$\begin{aligned} & \sqrt{n_\nu}(\hat{F}_{BU_\nu}(t) - F_{N_\nu}(t)) \\ & \xrightarrow{\mathcal{D}} N\left(0, (1-f)^2 \left[\phi^2\left(\frac{t-\beta}{\sigma}\right) + \frac{f}{1-f} \Phi\left(\frac{t-\beta}{\sigma}\right) \left(1 - \Phi\left(\frac{t-\beta}{\sigma}\right)\right) \right] \right), \end{aligned} \quad (2.7)$$

onde “ $\xrightarrow{\mathcal{D}}$ ” significa convergência em distribuição e $\phi(\cdot)$ é a função densidade da distribuição normal padrão.

Prova. Para simplificar a notação, eliminamos o índice ν . Podemos escrever

$$\begin{aligned} \sqrt{n}(\hat{F}_{BU}(t) - F_N(t)) &= \left(1 - \frac{n}{N}\right) \left\{ \sqrt{n} \left[\Phi\left(\sqrt{\frac{n}{n-1}} \left(\frac{t - \bar{y}_s}{\sigma}\right)\right) - \Phi\left(\frac{t - \beta}{\sigma}\right) \right] \right. \\ &\quad \left. - \sqrt{\frac{n}{N-n}} \sqrt{N-n} \left[\sum_{i \notin s} \frac{\Delta(t - y_i)}{N-n} - \Phi\left(\frac{t - \beta}{\sigma}\right) \right] \right\}. \end{aligned}$$

Utilizando-se o teorema do limite central para variáveis aleatórias independentes e identicamente distribuídas (Serfling, 1980), segue que

$$\sqrt{N-n} \left[\sum_{i \notin s} \frac{\Delta(t - y_i)}{N-n} - \Phi\left(\frac{t - \beta}{\sigma}\right) \right] \xrightarrow{\mathcal{D}} N\left(0, \Phi\left(\frac{t - \beta}{\sigma}\right) \left(1 - \Phi\left(\frac{t - \beta}{\sigma}\right)\right)\right). \quad (2.8)$$

Além disso, temos que

$$\sqrt{\frac{n}{n-1}} = \left(1 - \frac{1}{n}\right)^{-1/2} = 1 + \frac{1}{2n} + O(n^{-2}).$$

Logo, utilizando-se o Teorema 5.2 (a) do Capítulo 2, Lehmann (1983), segue

$$\sqrt{n} \left[\Phi\left(\sqrt{\frac{n}{n-1}} \left(\frac{t - \bar{y}_s}{\sigma}\right)\right) - \Phi\left(\frac{t - \beta}{\sigma}\right) \right] \xrightarrow{\mathcal{D}} N\left(0, \phi^2\left(\frac{t - \beta}{\sigma}\right)\right). \quad (2.9)$$

Como os termos da esquerda de (2.8) e (2.9) são independentes, o resultado (2.7) segue, e concluímos a prova.

O estimador usual para $F_N(t)$ sob amostragem aleatória simples sem reposição é a função distribuição empírica, $F_{N_s}(t)$, dada em (2.2). De acordo com (2.3), temos que $F_{N_s}(t)$ é um previsor não viciado de $F_N(t)$ e pode ser considerado como um previsor não paramétrico. Isto é, pode ser utilizado em situações onde há pouco ou nenhum conhecimento sobre a forma da distribuição do vetor de erros $\mathbf{e}$. Além disso, esse é um previsor ótimo, quando o modelo considerado está na classe de todas as distribuições contínuas. Esse resultado se verifica, porque, nesse caso, a estatística totalmente suficiente e completa é o vetor das estatísticas de ordem e $F_{N_s}(t)$ sendo simétrico em $\mathbf{y}_s$, certamente, é uma função dessa estatística. Uma aplicação direta do teorema do limite central para variáveis aleatórias independentes e identicamente distribuídas fornece

$$\sqrt{n_\nu}(F_{N_{s,\nu}}(t) - F_{N_\nu}(t)) \xrightarrow{\mathcal{D}} N\left(0, (1-f)\Phi\left(\frac{t-\beta}{\sigma}\right)\left(1 - \Phi\left(\frac{t-\beta}{\sigma}\right)\right)\right). \quad (2.10)$$

De (2.7) e (2.10) segue que a eficiência relativa assintótica (ARE - razão entre as variâncias preditivas assintóticas) de $F_{N_s}(t)$ em relação a $\hat{F}_{BU}(t)$ é

$$\text{ARE}(F_{N_s}(t); \hat{F}_{BU}(t)) = f + \frac{(1-f)\phi^2\left(\frac{t-\beta}{\sigma}\right)}{\Phi\left(\frac{t-\beta}{\sigma}\right)\left(1 - \Phi\left(\frac{t-\beta}{\sigma}\right)\right)}. \quad (2.11)$$

Para $t = \beta$, temos

$$\text{ARE}(F_{N_s}(t); \hat{F}_{BU}(t)) = f + (1-f)0.637.$$

Além disso, podemos mostrar que essa eficiência assintótica é uma função decrescente de $|t - \beta|$ e quando $|t - \beta| \rightarrow \infty$,

$$\text{ARE}(F_{N_s}(t); \hat{F}_{BU}(t)) \rightarrow f.$$

Logo, a eficiência acima pode ser tão pequena quanto f .

2.3.2 O MODELO DE REGRESSÃO LINEAR

Nesta seção encontramos o melhor previsor não viciado para alguns modelos de regressão. Inicialmente, consideramos o caso onde $\mathbf{W}$ é a matriz identidade. Depois, consideramos uma matriz de covariâncias mais geral. Então, assumimos inicialmente o modelo (2.1) sendo

$$\mathbf{X} = \begin{pmatrix} x_{11} & \cdots & x_{1p} \\ \vdots & & \\ x_{N1} & \cdots & x_{Np} \end{pmatrix} \quad \text{e} \quad \mathbf{W} = \mathbf{I}_N \quad (2.12)$$

e erros distribuídos normalmente.

Sob o modelo (2.1) com a especificação (2.12) segue de (2.4) que a estatística totalmente suficiente e completa é

$$\hat{\boldsymbol{\beta}}_s = (\mathbf{X}'_s \mathbf{X}_s)^{-1} \mathbf{X}'_s \mathbf{y}_s. \quad (2.13)$$

No próximo teorema encontramos o BUP de $F_N(t)$.

Teorema 2.4. *Sob o modelo especificado em (2.12) com erros normalmente distribuídos,*

$$\hat{F}_{BU}(t) = \frac{n}{N} F_{N_s}(t) + \left(1 - \frac{n}{N}\right) \sum_{i \notin s} \frac{1}{N - n} \Phi \left(\frac{t - \mathbf{x}'_i \hat{\boldsymbol{\beta}}_s}{\sigma \sqrt{1 - \mathbf{x}'_i (\mathbf{X}'_s \mathbf{X}_s)^{-1} \mathbf{x}_i}} \right)$$

se $\mathbf{x}'_i (\mathbf{X}'_s \mathbf{X}_s)^{-1} \mathbf{x}_i < 1$ para todo $i \notin s$.

Prova. De acordo com Olkin e Ghurye (1969),

$$E \left[\Phi \left(\frac{t - \mathbf{x}'_i \hat{\boldsymbol{\beta}}_s}{\sigma \sqrt{1 - \mathbf{x}'_i (\mathbf{X}'_s \mathbf{X}_s)^{-1} \mathbf{x}_i}} \right) \right] = E[\Delta(t - y_i)]$$

se $\mathbf{x}'_i (\mathbf{X}'_s \mathbf{X}_s)^{-1} \mathbf{x}_i < 1$, o que mostra que $\hat{F}_{BU}(t)$ é um previsor não viciado de $F_N(t)$, segundo a Definição 2.4. Como $\hat{\boldsymbol{\theta}}_{sr}$ é uma função da estatística totalmente suficiente e completa, o resultado segue a partir do Teorema 2.1.

Exemplo 2.1. Consideremos o modelo de regressão simples onde $\mathbf{X}' = (x_1, \dots, x_N)$ e $\mathbf{W} = \mathbf{I}_N$, ou seja,

$$y_i = x_i\beta + e_i ,$$

onde $e_i \sim N(0, \sigma^2)$, $i = 1, \dots, N$. Nesse caso, temos de (2.13) que

$$\hat{\beta}_s = \frac{\sum_{i \in s} x_i y_i}{\sum_{i \in s} x_i^2} .$$

Segue do Teorema 2.4 que o BUP de $F_N(t)$ é dado por

$$\hat{F}_{BU}(t) = \frac{n}{N} F_{N_s}(t) + \left(1 - \frac{n}{N}\right) \sum_{i \notin s} \frac{1}{N - n} \Phi \left(\frac{t - \hat{\beta}_s x_i}{\sigma \sqrt{1 - x_i^2 / \sum_{j \in s} x_j^2}} \right)$$

se $\sum_{j \in s} x_j^2 > x_i^2$ para todo $i \notin s$.

Temos que

$$\begin{aligned} & \text{Var} [\hat{F}_{BU}(t) - F_N(t)] \\ &= \left(1 - \frac{n}{N}\right)^2 \left\{ \frac{1}{(N - n)^2} \sum_{i \notin s} \sum_{k \notin s} \text{Cov} \left[\Phi \left(\frac{t - \hat{\beta}_s x_i}{\sigma \sqrt{1 - x_i^2 / \sum_{j \in s} x_j^2}} \right); \Phi \left(\frac{t - \hat{\beta}_s x_k}{\sigma \sqrt{1 - x_k^2 / \sum_{j \in s} x_j^2}} \right) \right] \right. \\ & \quad \left. + \frac{1}{(N - n)^2} \sum_{i \notin s} \Phi \left(\frac{t - \beta x_i}{\sigma} \right) \left[1 - \Phi \left(\frac{t - \beta x_i}{\sigma} \right) \right] \right\} . \end{aligned}$$

Seja

$$\Phi \left(\frac{t - \hat{\beta}_s x_i}{\sigma \sqrt{1 - x_i^2 / \sum_{j \in s} x_j^2}} \right) = g(T_n^i, n) \quad \text{e} \quad \theta^i = \frac{t - \beta x_i}{\sigma} ,$$

onde

$$T_n^i = \frac{t - \hat{\beta}_s x_i}{\sigma \sqrt{1 - x_i^2 / \sum_{j \in s} x_j^2}} = c_n^i \left(\frac{t - \hat{\beta}_s x_i}{\sigma} \right) .$$

Podemos utilizar os Teoremas A.2.1 e A.2.3 do Apêndice 2 para obter uma aproximação para a variância preditiva de $\hat{F}_{BU}(t)$. Para isso, basta verificarmos as condições exigidas pelos teoremas sobre a função g e a estatística T_n^i definidas acima. Fixando $q = 2$ nos Teoremas A.2.1 e A.2.3 e da forma como definimos a função g , temos que para todo

n, g é uma função limitada com as três primeiras derivadas contínuas e limitadas sobre $[\theta^i - \delta, \theta^i + \delta] \times N$. Como $\hat{\beta}_s \sim N(\beta; \sigma^2 / \sum_{j \in s} x_j^2)$, então a sequência $\{T_n^i\}$ de estatísticas tem até o 6º momento finito. Além disso,

$$\begin{aligned} E(\hat{\beta}_s - \beta)^j &= 0 \quad \text{se } j = 1, 3, 5, \\ E(\hat{\beta}_s - \beta)^2 &= \frac{\sigma^2}{\sum_{j \in s} x_j^2}, \quad E(\hat{\beta}_s - \beta)^j = O(n^{-j/2}), \quad j = 4, 6 \quad \text{e} \\ c_n^i &= \frac{1}{\sqrt{1 - x_i^2 / \sum_{j \in s} x_j^2}} = 1 + \frac{x_i^2}{2 \sum_{j \in s} x_j^2} + O(n^{-2}). \end{aligned}$$

Logo, podemos verificar que

$$E|T_n^i - \theta^i|^6 = E \left[(c_n^i - 1) \frac{t}{\sigma} - \frac{x_i}{\sigma} (c_n^i (\hat{\beta}_s - \beta) + (c_n^i - 1) \beta) \right]^6 = O(n^{-3}). \quad (2.14)$$

Dessa forma, estão satisfeitas as condições e da mesma maneira que em (2.14), temos

$$\begin{aligned} E[T_n^i - \theta^i] &= \frac{x_i^2}{\sum_{j \in s} x_j^2} \left(\frac{t - \beta x_i}{\sigma} \right) + O(n^{-2}), \\ E[T_n^i - \theta^i]^2 &= \frac{x_i^2}{\sum_{j \in s} x_j^2} + O(n^{-2}), \\ E[T_n^i - \theta^i]^3 &= O(n^{-2}), \\ E[(T_n^i - \theta^i)(T_n^l - \theta^l)] &= \frac{x_i x_l}{\sum_{j \in s} x_j^2} + O(n^{-2}), \\ E[(T_n^i - \theta^i)^2 (T_n^l - \theta^l)] &= O(n^{-2}) \quad \text{e} \\ E[(T_n^i - \theta^i)^2 (T_n^l - \theta^l)^2] &= O(n^{-2}). \end{aligned}$$

Portanto, pelos Teoremas A.2.1 e A.2.3, uma aproximação para a variância preditiva de $\hat{F}_{BU}(t)$ é dada por

$$\begin{aligned} &\text{Var} [\hat{F}_{BU}(t) - F_N(t)] \\ &= \left(1 - \frac{n}{N}\right)^2 \left\{ \frac{1}{(N-n)^2} \sum_{i \notin s} \sum_{k \notin s} \phi \left(\frac{t - \beta x_i}{\sigma} \right) \phi \left(\frac{t - \beta x_k}{\sigma} \right) \frac{x_i x_k}{\sum_{j \in s} x_j^2} \right. \\ &\quad \left. + \frac{1}{(N-n)^2} \sum_{i \notin s} \Phi \left(\frac{t - \beta x_i}{\sigma} \right) \left(1 - \Phi \left(\frac{t - \beta x_i}{\sigma} \right) \right) \right\} + O(n^{-2}) \end{aligned} \quad (2.15)$$

Notamos que se $x_i = 1$, $i = 1, \dots, N$, ou seja, se o modelo é de locação e ν é grande, a variância preditiva (2.15) coincide com a variância assintótica (2.7) obtida no Teorema 2.3.

Suponhamos agora que o modelo (2.1) é tal que

$$\mathbf{X} = \begin{pmatrix} x_{11} & \cdots & x_{1p} \\ \vdots & & \vdots \\ x_{N1} & \cdots & x_{Np} \end{pmatrix} \quad \text{e} \quad \mathbf{W} = \text{diag}(v_1, \dots, v_N) \quad (2.16)$$

com e distribuído normalmente.

Sob o modelo (2.1) com matriz de covariâncias especificada em (2.16), segue de (2.4) que a estatística totalmente suficiente e completa é

$$\hat{\beta}_s = (\mathbf{X}'_s \mathbf{W}_s^{-1} \mathbf{X}_s)^{-1} \mathbf{X}'_s \mathbf{W}_s^{-1} \mathbf{y}_s. \quad (2.17)$$

Apesar dos resultados de Olkin e Ghurye (1969) serem para vetores de variáveis aleatórias independentes e com variância constante, podemos utilizar o previsor derivado no Teorema 2.4 para obtermos o previsor de $F_N(t)$ sob a situação mais geral especificada em (2.16).

Teorema 2.5. *Sob o modelo (2.1) com $\mathbf{X}$ e $\mathbf{W}$ especificados em (2.16) e com erros distribuídos normalmente*

$$\hat{F}_{BU}(t) = \frac{n}{N} F_{N_s}(t) + \left(1 - \frac{n}{N}\right) \sum_{i \notin s} \frac{1}{N - n} \Phi \left(\frac{t - \mathbf{x}'_i \hat{\beta}_s}{\sigma \sqrt{v_i} \sqrt{1 - \frac{1}{v_i} \mathbf{x}'_i (\mathbf{X}'_s \mathbf{W}_s^{-1} \mathbf{X}_s)^{-1} \mathbf{x}_i}} \right)$$

se $\frac{1}{v_i} \mathbf{x}'_i (\mathbf{X}'_s \mathbf{W}_s^{-1} \mathbf{X}_s)^{-1} \mathbf{x}_i < 1$ para todo $i \notin s$ e onde $\hat{\beta}_s$ é dado em (2.17).

Prova. A prova segue diretamente do Teorema 2.4 fazendo a transformação $y_i^* = y_i / v_i^{1/2}$, $x_i^* = x_i / v_i^{1/2}$ e $e_i^* = e_i / v_i^{1/2}$.

No exemplo a seguir consideramos um importante caso particular.

Exemplo 2.2. Para o caso particular do modelo especificado em (2.16) onde $\mathbf{X} = (x_1, \dots, x_N)$ e $v_i = x_i$, $i = 1, \dots, N$, segue de (2.17) que

$$\hat{\beta}_s = \bar{y}_s / \bar{x}_s ,$$

onde $\bar{y}_s = \frac{1}{n} \sum_{i \in s} y_i$ e $\bar{x}_s = \frac{1}{n} \sum_{i \in s} x_i$, logo $\hat{\beta}_s$ é o estimador razão usual. Segue do Teorema 2.5 que o BUP de $F_N(t)$ é dado por

$$\hat{F}_{BU}(t) = \frac{n}{N} F_{N_s}(t) + \left(1 - \frac{n}{N}\right) \sum_{i \notin s} \frac{1}{N - n} \Phi \left(\frac{t - \hat{\beta}_s x_i}{\sigma \sqrt{x_i(1 - x_i / \sum_{j \in s} x_j)}} \right)$$

se $\sum_{j \in s} x_j > x_i$ para todo $i \notin s$.

Como no Exemplo 2.1, considerando

$$g(T_n^i, n) = \Phi \left(\frac{t - \hat{\beta}_s x_i}{\sigma \sqrt{x_i(1 - x_i / \sum_{j \in s} x_j)}} \right) \quad \text{e} \quad \theta^i = \frac{t - \beta x_i}{\sigma \sqrt{x_i}}$$

onde $T_n^i = \frac{t - \hat{\beta}_s x_i}{\sigma \sqrt{x_i(1 - x_i / \sum_{j \in s} x_j)}} = c_n^i \left(\frac{t - \hat{\beta}_s x_i}{\sigma \sqrt{x_i}} \right)$, derivamos, a partir dos Teoremas A.2.1 e A.2.3, uma aproximação para a variância preditiva de $\hat{F}_{BU}(t)$ que é dada por

$$\begin{aligned} & \text{Var} [\hat{F}_{BU}(t) - F_N(t)] \\ &= \left(1 - \frac{n}{N}\right)^2 \left\{ \frac{1}{(N - n)^2} \sum_{i \notin s} \sum_{k \notin s} \phi \left(\frac{t - \beta x_i}{\sigma \sqrt{x_i}} \right) \phi \left(\frac{t - \beta x_k}{\sigma \sqrt{x_k}} \right) \frac{\sqrt{x_i x_k}}{\sum_{j \in s} x_j} \right. \\ & \quad \left. + \frac{1}{(N - n)^2} \sum_{i \notin s} \Phi \left(\frac{t - \beta x_i}{\sigma \sqrt{x_i}} \right) \left(1 - \Phi \left(\frac{t - \beta x_i}{\sigma \sqrt{x_i}} \right) \right) \right\} + O(n^{-2}) \end{aligned} \quad (2.18)$$

2.4 σ^2 DESCONHECIDO

Nesta seção consideramos a previsão ótima da função distribuição da população finita sob o modelo (2.1), mas vamos supor que σ^2 é desconhecido. Nesse caso, a estatística totalmente suficiente e completa é

$$(\hat{\beta}_s, S_y^2) \quad (2.19)$$

onde

$$\hat{\beta}_s = (\mathbf{X}_s' \mathbf{W}_s^{-1} \mathbf{X}_s)^{-1} \mathbf{X}_s' \mathbf{W}_s^{-1} \mathbf{y}_s$$

e

$$S_y^2 = (\mathbf{y}_s - \mathbf{X}_s \hat{\beta}_s)' \mathbf{W}_s^{-1} (\mathbf{y}_s - \mathbf{X}_s \hat{\beta}_s) / (n - p) .$$

Os previsores derivados a seguir serão baseados na estatística $S = (\hat{\beta}_s, S_y^2)$. Como na seção 2.3, inicialmente trabalhamos sob o modelo de locação simples e depois então, sob o modelo de regressão linear.

2.4.1 O MODELO DE LOCAÇÃO SIMPLES

Como já vimos na seção 2.3.1, no modelo de locação simples temos

$$\mathbf{X} = \mathbf{1}_N \quad \text{e} \quad \mathbf{W} = \mathbf{I}_N . \quad (2.20)$$

De acordo com (2.19), a estatística totalmente suficiente e completa é $(\hat{\beta}_s, S_y^2)$, onde

$$\hat{\beta}_s = \bar{y}_s$$

e

$$S_y^2 = \frac{1}{n-1} \sum_{i \in s} (y_i - \bar{y}_s)^2$$

e o preditor $\hat{F}_{BU}(t)$ será baseado nela.

Teorema 2.6. *Sob o modelo especificado por (2.1) e (2.20) com erros normalmente distribuídos,*

$$\hat{F}_{BU}(t) = \frac{n}{N} F_{N_s}(t) + \left(1 - \frac{n}{N}\right) T_{n-2}(U_n(t))$$

onde

$$U_n(t) = \frac{\sqrt{n-2}(t - \bar{y}_s)}{\sqrt{\frac{(n-1)^2}{n} S_y^2 - (t - \bar{y}_s)^2}}$$

e $T_{n-2}(\cdot)$ é a função distribuição da t -Student com $n-2$ graus de liberdade.

Prova. De acordo com Olkin e Ghurye (1969), O'Reilly (1974), segue que

$$E_\psi[T_{n-2}(U_n(t))] = E_\psi[\Delta(t - y_i)] = \Phi\left(\frac{t - \beta}{\sigma}\right)$$

o que prova que $\hat{F}_{BU}(t)$ é não viciado. Como $\hat{\theta}_{sr}$ é função da estatística totalmente suficiente e completa $(\hat{\beta}_s, S_y^2)$, pelo Teorema 2.1, a prova está completa.

Assumindo o conceito de convergência apresentado na seção 2.1, derivamos a distribuição assintótica do previsor ótimo no teorema a seguir.

Teorema 2.7. *Consideremos uma seqüência de populações finitas $\{\mathcal{P}_\nu, \nu \geq 1\}$, satisfazendo $N_\nu \leq N_{\nu+1}$. Da população $\mathcal{P}_\nu$, uma amostra de tamanho n_ν ($n_\nu < n_{\nu+1}$) é selecionada de modo que, quando $\nu \rightarrow \infty$*

$$N_\nu - n_\nu \rightarrow \infty \quad e \quad f_\nu = \frac{n_\nu}{N_\nu} \rightarrow f, \quad \text{onde } 0 < f < 1.$$

Então, quando $\nu \rightarrow \infty$,

$$\begin{aligned} \sqrt{n_\nu}(\hat{F}_{BU_\nu}(t) - F_{N_\nu}(t)) &\xrightarrow{\mathcal{D}} N\left(0, (1-f)^2 \left[\left(1 + \frac{(t-\beta)^2}{2\sigma^2}\right) \phi^2\left(\frac{t-\beta}{\sigma}\right) \right. \right. \\ &\quad \left. \left. + \frac{f}{1-f} \Phi\left(\frac{t-\beta}{\sigma}\right) \left(1 - \Phi\left(\frac{t-\beta}{\sigma}\right)\right) \right] \right) \end{aligned} \quad (2.21)$$

Prova. Para simplificar a notação, eliminamos o índice ν . Podemos escrever

$$\begin{aligned} \sqrt{n}(\hat{F}_{BU}(t) - F_N(t)) &= \left(1 - \frac{n}{N}\right) \left\{ \sqrt{n} \left[T_{n-2}(U_n(t)) - \Phi\left(\frac{t-\beta}{\sigma}\right) \right] \right. \\ &\quad \left. - \sqrt{\frac{n}{N-n}} \sqrt{N-n} \left[\sum_{j \notin s} \frac{\Delta(t - y_j)}{N-n} - \Phi\left(\frac{t-\beta}{\sigma}\right) \right] \right\}. \end{aligned} \quad (2.22)$$

Utilizando-se o Teorema do Limite Central para variáveis aleatórias independentes e identicamente distribuídas (Serfling, 1980), segue que

$$\sqrt{N-n} \left[\sum_{j \notin s} \frac{\Delta(t - y_j)}{N-n} - \Phi\left(\frac{t-\beta}{\sigma}\right) \right] \xrightarrow{\mathcal{D}} N\left(0, \Phi\left(\frac{t-\beta}{\sigma}\right) \left(1 - \Phi\left(\frac{t-\beta}{\sigma}\right)\right)\right). \quad (2.23)$$

Além disso, podemos mostrar, utilizando o Teorema A.1.1, que

$$\sqrt{n} \left[T_{n-2}(U_n(t)) - \Phi \left(\frac{t - \beta}{\sigma} \right) \right] = \sqrt{n} \left[\Phi(U_n(t)) - \Phi \left(\frac{t - \beta}{\sigma} \right) \right] + O_p(n^{-1/2}) \quad (2.24)$$

e pelo Teorema A.1.2

$$\sqrt{n} \left(U_n(t) - \frac{t - \beta}{\sigma} \right) \xrightarrow{d} N \left(0, 1 + \frac{(t - \beta)^2}{2\sigma^2} \right). \quad (2.25)$$

Portanto, aplicando-se o método delta (Serfling, 1980), segue de (2.24) e (2.25) que

$$\sqrt{n} \left[T_{n-2}(U_n(t)) - \Phi \left(\frac{t - \beta}{\sigma} \right) \right] \xrightarrow{d} N \left(0, \left(1 + \frac{(t - \beta)^2}{2\sigma^2} \right) \phi^2 \left(\frac{t - \beta}{\sigma} \right) \right). \quad (2.26)$$

Utilizando-se (2.23) e (2.26) em (2.22), concluímos a prova.

Observando o melhor previsor não viciado de $F_N(t)$ dado em (2.6) para o caso em que σ^2 é conhecido, podemos definir, no caso de σ^2 desconhecido, o seguinte previsor alternativo

$$\hat{F}_N^*(t) = \frac{n}{N} F_{N_s} + \left(1 - \frac{n}{N} \right) \Phi \left(\sqrt{\frac{n}{n-1}} \left(\frac{t - \bar{y}_s}{S_y} \right) \right).$$

O resultado a seguir estabelece uma importante relação entre $\hat{F}_{BU}(t)$ e $\hat{F}_N^*(t)$.

Teorema 2.8. *Sob as suposições discutidas no Teorema 2.7*

$$\hat{F}_{BU_\nu}(t) - \hat{F}_{N_\nu}^*(t) \xrightarrow{\text{q.c.}} 0$$

onde “ $\xrightarrow{\text{q.c.}}$ ” significa convergência quase certa.

Prova. Notemos que

$$\begin{aligned} \hat{F}_{BU}(t) - \hat{F}_N^*(t) &= \left(1 - \frac{n}{N} \right) \left[T_{n-2}(U_n(t)) - \Phi \left(\frac{t - \beta}{\sigma} \right) \right] \\ &\quad - \left(1 - \frac{n}{N} \right) \left[\Phi \left(\sqrt{\frac{n}{n-1}} \left(\frac{t - \bar{y}_s}{S_y} \right) \right) - \Phi \left(\frac{t - \beta}{\sigma} \right) \right]. \end{aligned} \quad (2.27)$$

Temos que

$$\Phi \left(\sqrt{\frac{n}{n-1}} \left(\frac{t - \bar{y}_s}{S_y} \right) \right) \xrightarrow{\text{q.c.}} \Phi \left(\frac{t - \beta}{\sigma} \right) \quad (2.28)$$

e como $U_n(t) \xrightarrow{\text{q.c.}} \frac{t - \beta}{\sigma}$ logo

$$\Phi(U_n(t)) \xrightarrow{\text{q.c.}} \Phi\left(\frac{t - \beta}{\sigma}\right) . \quad (2.29)$$

Além disso,

$$T_{n-2}(x) \xrightarrow{\text{q.c.}} \Phi(x) \quad (\text{uniformemente})$$

e, portanto,

$$T_{n-2}(U_n(t)) - \Phi(U_n(t)) \xrightarrow{\text{q.c.}} 0 . \quad (2.30)$$

Podemos escrever

$$|T_{n-2}(U_n(t)) - \Phi\left(\frac{t - \beta}{\sigma}\right)| \leq |T_{n-2}(U_n(t)) - \Phi(U_n(t))| + |\Phi(U_n(t)) - \Phi\left(\frac{t - \beta}{\sigma}\right)| ,$$

logo, de (2.29) e (2.30) segue que

$$T_{n-2}(U_n(t)) - \Phi\left(\frac{t - \beta}{\sigma}\right) \xrightarrow{\text{q.c.}} 0 . \quad (2.31)$$

Utilizando-se (2.28) e (2.31) em (2.27) concluimos a prova.

De forma similar à do Teorema 2.7, podemos mostrar que a distribuição assintótica de $\sqrt{n_\nu}(\hat{F}_{N_\nu}^*(t) - F_{N_\nu}(t))$ é dada por (2.21), isto é, $\hat{F}_N^*(t)$ e $\hat{F}_{BU}(t)$ têm a mesma distribuição preditiva assintótica. Já vimos na seção 2.3.1 que a função distribuição empírica (2.2) é o estimador usual para $F_N(t)$ sob amostragem aleatória simples. Temos que a eficiência relativa assintótica de $F_{N_s}(t)$ com relação a $\hat{F}_{BU}(t)$ é, a partir de (2.10) e (2.21), dada por

$$\text{ARE}(F_{N_s}(t); \hat{F}_{BU}(t)) = f + \frac{(1 - f) \left(1 + \frac{(t - \beta)^2}{2\sigma^2}\right) \phi^2\left(\frac{t - \beta}{\sigma}\right)}{\Phi\left(\frac{t - \beta}{\sigma}\right) \left(1 - \Phi\left(\frac{t - \beta}{\sigma}\right)\right)} . \quad (2.32)$$

Para $t = \beta$,

$$\text{ARE}(F_{N_s}(t), \hat{F}_{BU}(t)) = f + (1 - f)0.637 .$$

Através da comparação de (2.11), (2.32) e da Figura 2.1, temos que a eficiência de $F_{N_s}(t)$ em relação ao previsor ótimo é maior para σ^2 desconhecido.

Além disso, podemos verificar que, quando $|t - \beta| \rightarrow \infty$, $\text{ARE}(F_{N_s}(t), \hat{F}_{BU}(t)) \rightarrow f$, o mesmo que ocorre para σ^2 conhecido.

Figura 2.1. Eficiência relativa assintótica de $F_{N_s}(t)$ com relação a $\hat{F}_{BU}(t)$ para $f = 0,10$

2.4.2 O MODELO DE REGRESSÃO LINEAR

Suponhamos agora o modelo 2.1, sendo

$$\mathbf{X} = \begin{pmatrix} x_{11} & \cdots & x_{1p} \\ \vdots & & \vdots \\ x_{N1} & \cdots & x_{Np} \end{pmatrix} \quad \text{e} \quad \mathbf{W} = \text{diag}(v_1, \dots, v_N) \quad (2.33)$$

onde $v_i > 0$, $i = 1, \dots, N$. Nesse caso, a estatística totalmente suficiente e completa é dada por (2.19), e o BUP de $F_N(t)$, apresentado no teorema a seguir, será baseado nela.

Teorema 2.9. *Sob o modelo (2.1) com $\mathbf{X}$ e $\mathbf{W}$ dados em (2.33) e σ^2 desconhecido, segue que o BUP de $F_N(t)$ é dado por*

$$\hat{F}_{BU}(t) = \frac{n}{N} F_{N_s}(t) + \left(1 - \frac{n}{N}\right) \sum_{i \notin s} \frac{1}{N - n} T_{n-p-1}(U_n^i(t))$$

onde

$$U_n^i(t) = \frac{\sqrt{n-p-1}(t - \mathbf{x}_i' \hat{\boldsymbol{\beta}}_s)}{\sqrt{v_i} \sqrt{\left(1 - \frac{1}{v_i} \mathbf{x}_i' (\mathbf{X}_s' \mathbf{W}_s^{-1} \mathbf{X}_s)^{-1} \mathbf{x}_i\right) (n-p) S_y^2 - \frac{1}{v_i} (t - \mathbf{x}_i' \hat{\boldsymbol{\beta}}_s)^2}}$$

e T_ν é a função distribuição da t -Student com ν graus de liberdade.

Prova. Inicialmente consideramos o caso de variância constante, isto é, $v_i = 1$, $i = 1, \dots, N$. De acordo com Olkin e Ghurye (1969),

$$E_\psi[T_{n-p-1}(U_n^i(t))] = E_\psi[\Delta(t - y_i)] = \Phi\left(\frac{t - \mathbf{x}_i' \boldsymbol{\beta}}{\sigma}\right),$$

o que mostra que $\hat{F}_{BU}(t)$ é um previsor não viciado de $F_N(t)$. Como $\hat{\theta}_{sr}$ é uma função da estatística totalmente suficiente e completa, o resultado segue do Teorema 2.1. Apesar dos resultados de Olkin e Ghurye (1969) serem para vetores aleatórios com variância constante, a prova para o caso $\mathbf{W} = \text{diag}(v_1, \dots, v_N)$ segue diretamente do caso $\mathbf{W} = \mathbf{I}_N$ fazendo-se a transformação $y_i^* = y_i/v_i^{1/2}$, $x_i^* = x_i/v_i^{1/2}$ e $e_i^* = e_i/v_i^{1/2}$, $i = 1, \dots, N$.

A seguir apresentamos dois exemplos onde consideramos casos particulares do modelo especificado por (2.33).

Exemplo 2.3. Consideremos o modelo de regressão simples onde $\mathbf{X}' = (x_1, \dots, x_N)$ e $\mathbf{W} = \mathbf{I}_N$, isto é,

$$y_i = x_i \beta + e_i$$

e $e_i \sim N(0, \sigma^2)$, com σ^2 desconhecido, $i = 1, \dots, N$. Nesse caso, segue de (2.19) que

$$\hat{\beta}_s = \frac{\sum_{i \in s} x_i y_i}{\sum_{i \in s} x_i^2} \quad \text{e} \quad S_y^2 = \frac{1}{n-1} \left[\sum_{i \in s} y_i^2 - \frac{(\sum_{i \in s} x_i y_i)^2}{\sum_{i \in s} x_i^2} \right].$$

A partir do Teorema 2.9 temos que

$$\hat{F}_{BU}(t) = \frac{n}{N} F_{N_s}(t) + \left(1 - \frac{n}{N}\right) \sum_{i \notin s} \frac{1}{N-n} T_{n-2}(U_n^i(t))$$

onde

$$U_n^i(t) = \frac{\sqrt{n-2}(t - x_i \hat{\beta}_s)}{\sqrt{\left(1 - \frac{x_i^2}{\sum_{j \in s} x_j^2}\right) (n-1) S_y^2 - (t - x_i \hat{\beta}_s)^2}} .$$

Utilizando-se o Teorema A.2.2 e o Teorema A.2.4, podemos obter uma aproximação para a variância preditiva de $\hat{F}_{BU}(t)$. Seja

$$\begin{aligned} T_{n-2}(U_n^i(t)) &= g(T_{1n}^i, T_{2n}^i, n) = T_{n-2} \left(\frac{t - x_i T_{1n}^i}{(T_{2n}^i)^{1/2}} \right) , \\ g(\beta, \sigma^2, n) &= T_{n-2} \left(\frac{t - x_i \beta}{\sigma} \right) \end{aligned}$$

onde

$$\begin{aligned} T_{1n}^i &= \hat{\beta}_s , \\ T_{2n}^i &= \left(1 - \frac{x_i^2}{\sum_{j \in s} x_j^2} \right) \left(\frac{n-1}{n-2} \right) S_y^2 - \frac{(t - x_i \hat{\beta}_s)^2}{n-2} \quad e \\ \boldsymbol{\theta} &= (\beta, \sigma^2) . \end{aligned}$$

Dessa forma, a função g é limitada sobre $\mathbb{R}^2 \times N$ e, para todo n , g é uma função diferenciável até a ordem 3 com relação a t_1 e t_2 no intervalo $K = [\beta - \delta, \beta + \delta] \times [\sigma^2 - \eta, \sigma^2 + \eta]$, δ e $\eta > 0$. Além disso, todas essas derivadas são limitadas sobre $K \times N$.

Já vimos no Exemplo 2.1 que $T_{1n}^i = \hat{\beta}_s \sim N(\beta, \sigma^2 / \sum_{j \in s} x_j^2)$, logo T_{1n}^i tem momentos finitos e $E|T_{1n}^i - \beta|^6 = O(n^{-3})$. Como $(n-1) \frac{S_y^2}{\sigma^2} \sim \chi_{n-1}^2$, temos que

$$\begin{aligned} E(T_{2n}^i - \sigma^2)^6 &= E \left\{ \left(1 - \frac{x_i^2}{\sum_{j \in s} x_j^2} \right) \left(\frac{n-1}{n-2} \right) (S_y^2 - \sigma^2) \right. \\ &\quad \left. + \left[\left(1 - \frac{x_i^2}{\sum_{j \in s} x_j^2} \right) \left(\frac{n-1}{n-2} \right) - 1 \right] \sigma^2 - \left(\frac{t - x_i \hat{\beta}_s}{n-2} \right) \right\}^6 = O(n^{-3}) . \end{aligned}$$

Portanto, estão satisfeitas as condições para utilizarmos o Teorema A.2.2 e o Teorema A.2.4 e como

$$\begin{aligned} \text{Cov} [(T_{2n}^i - \sigma^2), (T_{2n}^i - \sigma^2)] &= \frac{2\sigma^4}{n-1} + O(n^{-2}) , \\ \text{Cov} [(T_{1n}^i - \beta)^j, (T_{2n}^i - \sigma^2)^k] &= O(n^{-2}) , \quad j = 1, 2 , \quad k = 1, 2 , \quad j + k \neq 4 , \end{aligned}$$

$$\text{Cov}[(T_{1n}^i - \beta), (T_{1n}^i - \beta)(T_{2n}^i - \sigma^2)] = O(n^{-2}) ,$$

$$\text{Cov}[(T_{2n}^i - \sigma^2), (T_{1n}^i - \beta)(T_{2n}^i - \sigma^2)] = O(n^{-2}) ,$$

$$\text{Cov}[(T_{1n}^i - \beta), (T_{1n}^i - \beta^2)] = 0 \quad e$$

$$\text{Cov}[(T_{2n}^i - \sigma^n), (T_{2n}^i - \sigma^2)^2] = O(n^{-2}) ,$$

segue que

$$\begin{aligned} & \text{Var}[\hat{F}_{BU}(t) - F_N(t)] \\ = & \left(1 - \frac{n}{N}\right)^2 \left\{ \frac{1}{(N-n)^2} \sum_{i \notin s} \sum_{k \notin s} t_{n-2} \left(\frac{t - x_i \beta}{\sigma} \right) t_{n-2} \left(\frac{t - x_k \beta}{\sigma} \right) \left[\frac{x_i x_k}{\sum_{j \in s} x_j^2} \right. \right. \\ & \left. \left. + \frac{(t - x_i \beta)(t - x_k \beta)}{2\sigma^2(n-1)} \right] + \frac{1}{(N-n)^2} \sum_{i \notin s} \Phi \left(\frac{t - x_i \beta}{\sigma} \right) \left[1 - \Phi \left(\frac{t - x_i \beta}{\sigma} \right) \right] \right\} + O(n^{-2}) \end{aligned} \quad (2.34)$$

onde t_{n-2} denota a função densidade da distribuição t -Student com $n - 2$ graus de liberdade.

Exemplo 2.4. No caso especial onde $\mathbf{X}' = (x_1, \dots, x_N)$ e $v_i = x_i$, $i = 1, \dots, N$, segue de (2.19) que a estatística totalmente suficiente e completa $(\hat{\beta}_s, S_y^2)$ é tal que

$$\hat{\beta}_s = \frac{\bar{y}_s}{\bar{x}_s} \quad e \quad S_y^2 = \frac{1}{n-1} \left[\sum_{i \in s} \frac{y_i^2}{x_i} - \frac{(\sum_{i \in s} y_i)^2}{\sum_{i \in s} x_i} \right] .$$

Nesse caso, $\hat{\beta}_s$ é o estimador razão usual. Pelo Teorema 2.9 temos que o BUP de $F_N(t)$ é dado por

$$\hat{F}_{BU}(t) = \frac{n}{N} F_{N_s}(t) + \left(1 - \frac{n}{N}\right) \sum_{i \notin s} \frac{1}{N-n} T_{n-2}(U_n^i(t))$$

onde

$$U_n^i(t) = \frac{\sqrt{n-2}(t - x_i \hat{\beta}_s)}{\sqrt{x_i} \sqrt{\left(1 - \frac{x_i}{\sum_{j \in s} x_j}\right) (n-1) S_y^2 - \frac{(t - x_i \hat{\beta}_s)^2}{x_i}}} .$$

Como no exemplo anterior, considerando

$$g(T_{1n}^i, T_{2n}^i, n) = T_{n-2} \left(\frac{t - x_i T_{1n}^i}{\sqrt{x_i} \sqrt{T_{2n}^i}} \right) ,$$

$$g(\beta, \sigma^2, n) = T_{n-2} \left(\frac{t - x_i \beta}{\sigma \sqrt{x_i}} \right)$$

onde

$$\begin{aligned} T_{1n}^i &= \hat{\beta}_s, \\ T_{2n}^i &= \left(1 - \frac{x_i}{\sum_{j \in s} x_j}\right) \left(\frac{n-1}{n-2}\right) S_y^2 - \frac{(t - x_i \hat{\beta}_s)}{(n-2)x_i} \quad e \\ \theta &= (\beta, \sigma^2), \end{aligned}$$

derivamos, a partir do Teorema A.2.2 e do Teorema A.2.4, uma aproximação para a variância preditiva que é dada por

$$\begin{aligned} &\text{Var}_\psi [\hat{F}_{BU}(t) - F_N(t)] \\ &= \left(1 - \frac{n}{N}\right)^2 \left\{ \frac{1}{(N-n)^2} \sum_{i \notin s} \sum_{k \notin s} t_{n-2} \left(\frac{t - x_i \beta}{\sigma \sqrt{x_i}} \right) t_{n-2} \left(\frac{t - x_k \beta}{\sigma \sqrt{x_k}} \right) \left[\frac{\sqrt{x_i x_k}}{\sum_{j \in s} x_j} \right. \right. \\ &\quad \left. \left. + \frac{(t - x_i \beta)(t - x_k \beta)}{2\sqrt{x_i x_k} \sigma^2 (n-1)} \right] + \frac{1}{(N-n)^2} \sum_{i \notin s} \Phi \left(\frac{t - x_i \beta}{\sigma \sqrt{x_i}} \right) \left[1 - \Phi \left(\frac{t - x_i \beta}{\sigma \sqrt{x_i}} \right) \right] \right\} + O(n^{-2}) \end{aligned} \quad (2.35)$$

Notemos que se $x_i = 1$, $i = 1, \dots, N$, as variâncias preditivas (2.34) e (2.35) coincidem.

2.5 ESTUDOS DE SIMULAÇÃO

Nesta seção apresentamos alguns estudos de simulação onde comparamos os previsores propostos com outros previsores que aparecem na literatura. O computador utilizado foi o Unisys A15 do Centro de Computação Eletrônica da USP.

Chambers e Dunstan (1986) consideram o modelo (2.1) com $\mathbf{X} = (x_1, \dots, x_N)$, $\mathbf{W} = \text{diag}(v_1, \dots, v_N)$, mas sem fazer nenhuma suposição sobre a forma da distribuição do vetor de erros $\mathbf{e}$. Partindo de métodos intuitivos, introduzem o seguinte predictor para $F_N(t)$

$$\hat{F}_{CD}(t) = \frac{n}{N} F_{N_s}(t) + \frac{1}{Nn} \sum_{i \notin s} \sum_{j \in s} \Delta \left[\left(\frac{t - \hat{\beta}_s x_i}{\sqrt{v_i}} \right) - \left(\frac{y_j - \hat{\beta}_s x_j}{\sqrt{v_j}} \right) \right] \quad (2.36)$$

onde

$$\hat{\beta}_s = \frac{\sum_{i \in s} x_i y_i / v_i}{\sum_{i \in s} x_i^2 / v_i}$$

é o estimador usual de mínimos quadrados de β . Para um modelo de superpopulação Gaussiano, o previsor de $F_N(t)$ obtido utilizando-se o enfoque de Chambers e Dunstan, é dado por

$$\hat{F}_{CD}(t) = \frac{n}{N} F_{N_s}(t) + \left(1 - \frac{n}{N}\right) \sum_{i \notin s} \frac{1}{N - n} \Phi \left(\frac{t - x_i \hat{\beta}_s}{\sigma \sqrt{v_i}} \right)$$

que é semelhante ao BUP dado no Teorema 2.5, especialmente para n grande. Apesar de $\hat{F}_{CD}$ dado em (2.36) ser um previsor simples e intuitivo, sua variância preditiva assintótica é uma expressão complicada e difícil de ser tratada. Notemos também que $\hat{F}_{BU}$ e $\hat{F}_{CD}$ não podem ser comparados diretamente, pois $\hat{F}_{CD}$ não é um previsor não viciado. Os estudos de simulação feitos por Chambers e Dunstan (1986) e os estudos apresentados a seguir, indicam que o vício do previsor $\hat{F}_{CD}$ não afeta seriamente o erro quadrático médio, já que em termos de erro quadrático médio, seu comportamento é bem melhor que o da função distribuição empírica F_{N_s} dada em (2.2). Além disso, temos que variações de $\bar{x}_s$, média amostral, não afetam o vício do previsor $\hat{F}_{CD}$. Outra característica de $\hat{F}_{CD}$ dada em (2.36) é que, se os valores de x_i forem próximos e $v_i = 1$, então $\hat{F}_{CD} \equiv F_{N_s}$. Se esse for o caso, a eficiência relativa assintótica de $\hat{F}_{CD}$ em relação a $\hat{F}_{BU}$ será próxima da eficiência relativa dada pela expressão (2.11). Mas existem situações, como os estudos de simulação apresentados a seguir mostram, onde essa eficiência relativa é próxima de 1. Os estudos de simulação mostram também que $\hat{F}_{BU}$ e $\hat{F}_{CD}$ são bem melhores, em termos de erro quadrático médio, do que F_{N_s} .

Sob a abordagem clássica onde a inferência está baseada exclusivamente na distribuição de probabilidades criada pelo planejamento amostral, o estimador usual de $F_N(t)$ é

$$\hat{F}_{N_s}(t) = \frac{\sum_{i \in s} \pi_i^{-1} \Delta(t - y_i)}{\sum_{i \in s} \pi_i^{-1}} \quad (2.37)$$

onde π_i , $i = 1, \dots, n$ é a probabilidade de inclusão do elemento i na amostra s . Nos estudos de simulação abaixo, $\pi_i = n/N$, $i = 1, \dots, n$, o que corresponde à amostragem casual simples e, portanto, o estimador (2.37) se reduz à função distribuição empírica dada

em (2.2). Um problema com o estimador (2.37) é a incorporação de variáveis auxiliares. Na abordagem clássica isso pode ser feito no estágio do planejamento, através de uma definição apropriada das probabilidades de inclusão π_i . No entanto, muitas vezes isso não é possível, e gostaríamos de introduzir a informação auxiliar no estágio de estimação. Um exemplo disso é o uso dos estimadores razão e regressão para estimar o total populacional sob amostragem casual simples.

A partir daí, Rao et al. (1990) propuseram estimadores razão e diferença para a função distribuição populacional, utilizando-se informação auxiliar x_i conhecida para todos os elementos da população. O estimador razão proposto é dado por

$$\hat{F}_r(t) = \frac{1}{N} \left\{ \sum_{i=1}^N \Delta(t - \hat{R}x_i) \right\} \left\{ \frac{\sum_{i \in s} \pi_i^{-1} \Delta(t - y_i)}{\sum_{i \in s} \pi_i^{-1} \Delta(t - \hat{R}x_i)} \right\} \quad (2.38)$$

onde $\hat{R} = \frac{\sum_{i \in s} y_i / \pi_i}{\sum_{i \in s} x_i / \pi_i}$.

Um outro estimador proposto é o estimador diferença que é dado por

$$\hat{F}_d(t) = \frac{1}{N} \left\{ \sum_{i \in s} \pi_i^{-1} \Delta(t - y_i) + \left[\sum_{i=1}^N \Delta(t - \hat{R}x_i) - \sum_{i \in s} \pi_i^{-1} \Delta(t - \hat{R}x_i) \right] \right\} . \quad (2.39)$$

Os estimadores (2.38) e (2.39) são assintoticamente não viciados para o planejamento, mas são viciados para o modelo (2.1). Logo, Rao et al. (1990) propuseram um estimador alternativo que é assintoticamente não viciado, tanto para o planejamento como para o modelo. Esse estimador é o estimador diferença modificado, dado por

$$\hat{F}_{dm}(t) = \frac{1}{N} \left\{ \sum_{i \in s} \pi_i^{-1} \Delta(t - y_i) + \left(\sum_{i=1}^N \hat{G}_i - \sum_{i \in s} \pi_i^{-1} \hat{G}_{ic} \right) \right\} \quad (2.40)$$

onde

$$\begin{aligned} \hat{G}_i &= \left(\sum_{j \in s} \pi_j^{-1} \right)^{-1} \cdot \left[\sum_{j \in s} \pi_j^{-1} \Delta(x_i^{-1/2}(t - \hat{R}x_i) - \hat{u}_j) \right] , \\ \hat{G}_{ic} &= \left(\sum_{j \in s} \pi_i / \pi_{ij} \right)^{-1} \cdot \left[\sum_{j \in s} (\pi_i / \pi_{ij}) \Delta(x_i^{-1/2}(t - \hat{R}x_i) - \hat{u}_j) \right] , \\ \hat{u}_j &= x_j^{-1/2}(y_j - \hat{R}x_j) \quad e \end{aligned}$$

π_{ij}/π_i é a probabilidade condicional de seleção dos elementos i e j dado que $i \in s$.

Para as populações normais geradas, o comportamento dos previsores (2.38), (2.39) e (2.40) em termos do erro quadrático médio é sempre pior do que os previsores baseados em modelo, como mostram as Tabelas 2.1, 2.2 e 2.3. Além disso, o predictor razão $\hat{F}_r$ se comporta bem pior do que F_{N_s} . As Tabelas 2.1, 2.2 e 2.3 mostram também que $\hat{F}_d$ e $\hat{F}_{dm}$ são melhores que $\hat{F}_r$ e F_{N_s} , mas, conforme observamos acima, suas eficiências são sempre inferiores às dos previsores baseados em modelos, $\hat{F}_{BU}$ e $\hat{F}_{CD}$.

Simulação A. Consideremos uma população finita de tamanho $N = 100$, gerada de acordo com o modelo

$$y_i = 3x_i + e_i ,$$

$i = 1, \dots, 100$, onde os e_i são independentes e seguem uma distribuição $N(0, \sigma^2)$, com $\sigma = 8,0$. Os x_i foram gerados de acordo com uma distribuição $N(50, 25)$. Dessa população foram selecionadas 1000 amostras casuais simples de tamanho $n = 10$. Para essa população, $F_N(t_1) = 0,25$, $F_N(t_2) = 0,5$ e $F_N(t_3) = 0,75$ implica que $t_1 = 141,64$, $t_2 = 149,41$ e $t_3 = 162,04$. Com base nas amostras selecionadas, o erro quadrático médio de qualquer predictor $\hat{F}_N$ é estimado por

$$EQM(\hat{F}_N(t)) = \frac{1}{1000} \sum_{i=1}^{1000} (\hat{F}_{iN}(t) - F_N(t))^2 .$$

A Tabela 2.1 apresenta os resultados da simulação em termos das razões entre os erros quadráticos médios dos previsores considerados acima. Estamos supondo σ^2 conhecido, logo consideramos o predictor ótimo dado no Exemplo 2.1.

Tabela 2.1. Razões entre os EQMs estimados

$F_N(t)$	0,25	0,50	0,75
$\frac{\widehat{EQM}(\widehat{F}_{BU}(t))}{\widehat{EQM}(\widehat{F}_{CD}(t))}$	0,9597	0,9768	1,0235
$\frac{\widehat{EQM}(\widehat{F}_{BU}(t))}{\widehat{EQM}(\widehat{F}_{N_s}(t))}$	0,2342	0,2058	0,1273
$\frac{\widehat{EQM}(\widehat{F}_{BU}(t))}{\widehat{EQM}(\widehat{F}_r(t))}$	0,2936	0,1405	0,4013
$\frac{\widehat{EQM}(\widehat{F}_{BU}(t))}{\widehat{EQM}(\widehat{F}_d(t))}$	0,5090	0,2639	0,4235
$\frac{\widehat{EQM}(\widehat{F}_{BU}(t))}{\widehat{EQM}(\widehat{F}_{dm}(t))}$	0,6267	0,4020	0,5141

Notamos que $\widehat{F}_{BU}$ e $\widehat{F}_{CD}$ são muito mais eficientes que F_{N_s} , $\widehat{F}_r$, $\widehat{F}_d$ e $\widehat{F}_{dm}$. Observamos também que o predictor $\widehat{F}_{CD}$ se comporta bem com relação ao predictor ótimo $\widehat{F}_{BU}$. Chambers e Dunstan (1986) notaram um aumento pequeno no vício de $\widehat{F}_{CD}$ comparado com o de F_{N_s} . A conclusão geral é que, sob modelos de superpopulação normais, os predictors baseados no modelo apresentam uma melhora significativa sobre os predictors baseados no planejamento. A Figura 2.2 ilustra a relação existente entre a média $\bar{x}_s$ e o erro médio dos predictors $\widehat{F}_{BU}$ e F_{N_s} . Utilizamos a mesma técnica que foi considerada por Chambers e Dunstan (1986) com relação ao predictor $\widehat{F}_{CD}$. As 1000 amostras selecionadas foram divididas em 20 grupos de 50 amostras cada, ordenadas de acordo com $\bar{x}_s$. Ou seja, no primeiro grupo temos as 50 amostras com os menores valores de $\bar{x}_s$, no segundo grupo temos amostras com os próximos 50 valores de $\bar{x}_s$ e, finalmente, no último grupo temos as amostras com os 50 maiores valores de $\bar{x}_s$. O vício médio no primeiro grupo (por exemplo) de um predictor $\widehat{F}_N$ é dado por

$$\sum \frac{\widehat{F}_N(t) - F_N(t)}{50}$$

onde essa $\sum$ é feita sobre as 50 amostras do primeiro grupo.

Figura 2.2. Variação do erro médio de $\hat{F}_{BU}(t_2)$ e $F_{N_s}(t_2)$

Notamos pela Figura 2.2 que o vício médio de $\hat{F}_{BU}(t)$ não é afetado por mudanças em $\bar{x}_s$. Chambers e Dunstan (1986) observaram o mesmo comportamento com relação ao previsor $\hat{F}_{CD}$.

Nesta simulação foram utilizadas as subrotinas *GGNML* e *MDNOR* do IMSL.

Simulação B. Consideremos agora uma população finita de tamanho $N = 100$ gerada de acordo com o modelo

$$y_i = 3x_i + e_i$$

onde os e_i são independentes com distribuição $N(0, \sigma^2 x_i)$, $\sigma = 8,0$, $i = 1, \dots, N$. Os x_i foram gerados de acordo com uma distribuição uniforme no intervalo $(10, 200)$. Dessa população foram selecionadas 1000 amostras casuais simples de tamanho $n = 10$. Para essa população, $F_N(t_1) = 0,25$, $F_N(t_2) = 0,5$ e $F_N(t_3) = 0,75$ implica $t_1 = 201,25$, $t_2 = 317,88$ e $t_3 = 432,33$. A Tabela 2.2 apresenta as razões entre os erros quadráticos médios estimados dos previsores $\hat{F}_{BU}$, $\hat{F}_{CD}$, F_{N_s} , $\hat{F}_r$, $\hat{F}_d$ e $\hat{F}_{dm}$. Aqui também estamos supondo σ^2 conhecido, logo, consideramos $\hat{F}_{BU}$ dado no Exemplo 2.2.

Tabela 2.2. Razões entre os EQMs estimados

$F_N(t)$	0,25	0,50	0,75
$\frac{\widehat{EQM}(\hat{F}_{BU}(t))}{\widehat{EQM}(\hat{F}_{CD}(t))}$	0,8415	0,9019	1,0240
$\frac{\widehat{EQM}(\hat{F}_{BU}(t))}{\widehat{EQM}(\hat{F}_{N_s}(t))}$	0,0912	0,1263	0,2512
$\frac{\widehat{EQM}(\hat{F}_{BU}(t))}{\widehat{EQM}(\hat{F}_r(t))}$	0,0807	0,1145	0,2249
$\frac{\widehat{EQM}(\hat{F}_{BU}(t))}{\widehat{EQM}(\hat{F}_d(t))}$	0,1598	0,2159	0,3514
$\frac{\widehat{EQM}(\hat{F}_{BU}(t))}{\widehat{EQM}(\hat{F}_{dm}(t))}$	0,2333	0,3209	0,5589

Como na Simulação A, notamos que $\hat{F}_{BU}$ e $\hat{F}_{CD}$ são muito mais eficientes que os demais previsores. Além disso, apesar do predictor $\hat{F}_r$ incorporar informação da variável auxiliar, ele é menos eficiente que o predictor F_{N_s} .

A Figura 2.3 mostra o vício médio dos previsores $F_{N_s}(t_3)$ e $\hat{F}_{BU}(t_3)$ em função de $\bar{x}_s$, e podemos observar o mesmo comportamento da Figura 2.2.

A Figura 2.4 mostra o comportamento da variância preditiva dada em (2.18) com relação a variações em $\bar{x}_s$. Ela foi baseada em 1000 amostras casuais simples de tamanho $n = 10$, divididas em 20 grupos de 50 amostras cada, conforme já explicado na Simulação A.

Figura 2.3. Variação do erro médio de $\hat{F}_{BU}(t_3)$ e $F_{N_s}(t_3)$

Figura 2.4. Variância preditiva de $\hat{F}_{BU}(t_3)$

Notamos claramente que, à medida que a média amostral $\bar{x}_s$ cresce, a variância preditiva decresce. Logo, parece que a amostra ótima para ser usada com o previsor ótimo sob o modelo de superpopulação considerado no Exemplo 2.2, é aquela que maximiza $\sum_{i \in s} x_i$. É bem conhecido que essa amostra é também ótima para prever o total populacional sob o modelo em questão.

Nesta simulação foram utilizadas as subrotinas *GGNML* e *MDNOR* do IMSL.

Simulação C. Consideremos uma população finita de tamanho $N = 100$ gerada de acordo com o modelo

$$y_i = 0,5x_i + e_i ,$$

$i = 1, \dots, 100$, onde os e_i são independentes e seguem uma distribuição $N(0, \sigma^2)$ com $\sigma = 6,0$. Nessa simulação os x_i foram gerados de acordo com uma distribuição $N(80, 64)$. Dessa população foram selecionadas 1.000 amostras casuais simples de tamanho $n = 10$. Para essa população, $F_N(t_1) = 0,25$, $F_N(t_2) = 0,50$ e $F_N(t_3) = 0,75$ implica que $t_1 = 35,59$, $t_2 = 39,54$ e $t_3 = 45,21$. Agora estamos supondo σ^2 desconhecido, logo, consideramos o previsor ótimo dado no Exemplo 2.3. A Tabela 2.3 apresenta os resultados da simulação em termos das razões entre os erros quadráticos médios dos previsores propostos anteriormente.

Observamos novamente que, sob modelos de superpopulação normais, os previsores baseados no modelo, $\hat{F}_{BU}$ e $\hat{F}_{CD}$, são muito mais eficientes que os previsores baseados no planejamento, F_{N_s} , $\hat{F}_r$, $\hat{F}_d$ e $\hat{F}_{dm}$.

Nesta simulação foram utilizadas as subrotinas *GGNML* e *MDTD* do IMSL.

Tabela 2.3. Razões entre os EQMs estimados

$F_N(t)$	0,25	0,50	0,75
$\frac{\widehat{EQM}(\widehat{F}_{BU}(t))}{\widehat{EQM}(\widehat{F}_{CD}(t))}$	1,0048	0,9547	0,9752
$\frac{\widehat{EQM}(\widehat{F}_{BU}(t))}{\widehat{EQM}(\widehat{F}_{N_s}(t))}$	0,4751	0,5028	0,4132
$\frac{\widehat{EQM}(\widehat{F}_{BU}(t))}{\widehat{EQM}(\widehat{F}_r(t))}$	0,2082	0,2523	0,5034
$\frac{\widehat{EQM}(\widehat{F}_{BU}(t))}{\widehat{EQM}(\widehat{F}_d(t))}$	0,3865	0,4893	0,5030
$\frac{\widehat{EQM}(\widehat{F}_{BU}(t))}{\widehat{EQM}(\widehat{F}_{dm}(t))}$	0,5408	0,7307	0,6440

CAPÍTULO 3

PREVISÃO BAYESIANA EMPÍRICA SOB UM MODELO DE LOCAÇÃO COM ERRO NAS VARIÁVEIS

3.1 INTRODUÇÃO

Neste capítulo consideramos o problema da previsão Bayesiana empírica do total populacional e da variância populacional sob um modelo de locação com erro nas variáveis, utilizando-se o enfoque proposto por Ghosh e Meeden (1986).

O método Bayesiano empírico foi introduzido por Robbins (1955) no contexto de estimação não paramétrica. A partir daí, o método foi explorado com sucesso numa série de artigos por Efron e Morris (1973, 1975, 1977) num contexto paramétrico. Morris (1983) listou várias áreas onde os métodos Bayesianos empíricos são usados. A idéia básica da análise Bayesiana empírica é a de retirar o máximo da amostra, ou seja, utilizar informação

de fontes similares juntamente com a fonte direta de informação. Em populações finitas existem várias situações onde esses métodos podem ser razoáveis. Por exemplo, em pesquisas que são repetidas, como as domiciliares, onde temos à disposição não somente os dados correntes, mas também dados dos experimentos passados. Outro exemplo são os problemas de estimação em pequenas áreas, cujas estimativas de uma certa área podem ser melhoradas incorporando informações de áreas similares. Ghosh e Meeden (1986) consideram a estimação Bayesiana empírica da média de uma população finita assumindo um modelo de superpopulação normal. A suposição de normalidade é relaxada em Ghosh e Lahiri (1987a e 1987b) que assumem a suposição de “linearidade posterior”, e, além disso, consideram a estimação simultânea das médias e variâncias de vários estratos. Tiwari e Lahiri (1989) consideram a previsão Bayesiana empírica da média assumindo um modelo “não homogêneo”. E mais recentemente, Lahiri e Peddada (1992) estudam o problema da estimação Bayesiana e Bayesiana empírica da média populacional usando informação de variáveis auxiliares.

Bolfarine (1991) considera o modelo de locação com erro nas variáveis e obtém previsores Bayesianos para o total e variância populacionais sob priores não informativas. Na seção 3.2 esses resultados são estendidos para o caso de priores informativas.

Consideremos uma população finita $\mathcal{P} = \{1, 2, \dots, N\}$, onde N é conhecido. Associado com a k -ésima unidade de $\mathcal{P}$, temos uma quantidade desconhecida y_k , $k = 1, \dots, N$. Neste caso, o modelo de locação com erros nas variáveis é definido por

$$\begin{aligned} Y_k &= y_k + v_k \\ y_k &= \theta + e_k \end{aligned} \tag{3.1}$$

onde e_k, v_k são independentes, $e_k \sim N(0, \sigma_e^2)$, $v_k \sim N(0, \sigma_v^2)$, $k = 1, \dots, N$, e consideremos que a distribuição a priori de θ é normal com média μ e variância σ^2 . A principal idéia por trás desse modelo é que as quantidades y_k não são observadas diretamente, e todas as inferências sobre as quantidades de interesse são baseadas em Y_k que são observáveis e

não em y_k que não o são. Um exemplo de tal situação é o caso onde y_k é a quantidade de nitrogênio no solo num local k . Valores de Y_k , neste caso, podem ser obtidos por análise de laboratório e são somente estimativas dos verdadeiros valores de y_k . Fuller (1990) considera uma versão multivariada do modelo com erro descrito acima, sendo dada ênfase para a previsão da quantidade não observada y_k dentro do contexto de população infinita, com σ_v^2 considerado conhecido.

Neste capítulo, nosso interesse é obter informação sobre o total populacional, $T = \sum_{k=1}^N y_k$ e a variância populacional $S^2 = \sum_{k=1}^N (y_k - \bar{y})^2 / N$, onde $\bar{y} = T/N$. Para isso, uma amostra s de tamanho n é selecionada de $\mathcal{P}$ de acordo com algum plano amostral. Notemos que as duas quantidades populacionais acima são definidas com relação a y_k , $k = 1, \dots, N$, que não são observados, mesmo depois de selecionada a amostra s . Podemos escrever $T = \sum_{k \in s} y_k + \sum_{k \in r} y_k = T_s + T_r$, onde $r = \mathcal{P} - s$ corresponde às unidades que estão fora da amostra. Sob o modelo de locação simples (Rodrigues et al. (1985)), temos que T_s é conhecido depois de selecionada a amostra, mas sob o modelo (3.1) tanto T_s como T_r permanecem desconhecidos mesmo depois de selecionada a amostra, então ambos precisam ser previstos. O mesmo acontece com a variância populacional S^2 .

Na seção 3.2 apresentamos a previsão Bayesiana de T e S^2 sob perda quadrática, considerando uma distribuição a priori informativa para θ . Na seção 3.3 nos dedicamos à previsão Bayesiana empírica de T e S^2 , com relação à mesma função de perda. O comportamento assintótico do risco dos previsores Bayesianos empíricos obtidos na seção 3.3 é considerado na seção 3.4. Na seção 3.5 os resultados das seções anteriores são estendidos para o caso do modelo de locação “não homogêneo” com erro nas variáveis. E, finalmente, a seção 3.6 apresenta estudos de Monte Carlo, onde o efeito da razão $\lambda = \sigma_e^2 / \sigma_v^2$ no comportamento dos previsores Bayesianos são investigados.

3.2 PREVISÃO BAYESIANA DE T E S^2

Nesta seção encontramos os previsores Bayesianos de T e S^2 sob função de perda quadrática. Considerando o modelo definido na seção 3.1, podemos verificar que a distribuição a posteriori de θ dada a amostra observada $\mathbf{Y}_s$ é

$$\theta/\mathbf{Y}_s \sim N\left(\frac{n\sigma^2}{n\sigma^2 + \sigma_e^2 + \sigma_v^2}\bar{Y}_s + \frac{\sigma_e^2 + \sigma_v^2}{n\sigma^2 + \sigma_e^2 + \sigma_v^2}\mu; \frac{\sigma^2(\sigma_e^2 + \sigma_v^2)}{n\sigma^2 + \sigma_e^2 + \sigma_v^2}\right)$$

onde $\bar{Y}_s = \sum_{i \in s} Y_i/n$ é a média amostral. Como

$$E[y_i|\theta, \mathbf{Y}_s] = \begin{cases} \theta & , i \notin s \\ \theta + \frac{\sigma_e^2}{\sigma_e^2 + \sigma_v^2}(Y_i - \theta) & , i \in s \end{cases} \quad (3.2)$$

segue que o preditor Bayesiano de T é dado por

$$\begin{aligned} \hat{T}_B &= E[T|\mathbf{Y}_s] = E\left[\sum_{i \in s} y_i|\mathbf{Y}_s\right] + E\left[\sum_{i \notin s} y_i|\mathbf{Y}_s\right] \\ &= E\left[E\left(\sum_{i \in s} y_i|\theta, \mathbf{Y}_s\right)|\mathbf{Y}_s\right] + E\left[E\left(\sum_{i \notin s} y_i|\theta, \mathbf{Y}_s\right)|\mathbf{Y}_s\right] \\ &= NE[\theta|\mathbf{Y}_s] + \frac{n\sigma_e^2}{\sigma_e^2 + \sigma_v^2}[\bar{Y}_s - E(\theta|\mathbf{Y}_s)] \\ &= N[(1 - B)\bar{Y}_s + B\mu] + \frac{n\sigma_e^2}{\sigma_e^2 + \sigma_v^2}B(\bar{Y}_s - \mu), \end{aligned} \quad (3.3)$$

onde $B = \frac{M}{M + n}$ e $M = \frac{\sigma_e^2 + \sigma_v^2}{\sigma^2}$.

Notamos que esse preditor se reduz ao preditor expansão quando $\sigma^2 \rightarrow \infty$.

Em relação à variância populacional, de acordo com Bolfarine e Zacks (1991), podemos escrever

$$S^2 = \frac{1}{N}\mathbf{y}'(\mathbf{I}_N - \frac{1}{N}\mathbf{J}_N)\mathbf{y}$$

onde $\mathbf{y} = (y_1, \dots, y_N)'$, $\mathbf{I}_N$ é a matriz identidade $N \times N$, $\mathbf{J}_N = \mathbf{1}_N\mathbf{1}_N'$ e $\mathbf{1}_N$ é o vetor de uns de dimensão N . Utilizando-se (3.2) e como

$$\text{Var}[y_i|\theta, \mathbf{Y}_s] = \begin{cases} \sigma_e^2 & , i \notin s \\ \frac{\sigma_e^2\sigma_v^2}{\sigma_e^2 + \sigma_v^2} & , i \in s \end{cases}$$

temos que

$$\begin{aligned} \Sigma &= \text{Var}(\mathbf{y}|\mathbf{Y}_s) = E\{\text{Var}[\mathbf{y}|\theta, \mathbf{Y}_s]|\mathbf{Y}_s\} + \text{Var}\{E[\mathbf{y}|\theta, \mathbf{Y}_s]|\mathbf{Y}_s\} \\ &= \begin{pmatrix} \frac{\sigma_e^2 \sigma_v^2}{\sigma_e^2 + \sigma_v^2} \mathbf{I}_n & \mathbf{0} \\ \mathbf{0} & \sigma_e^2 \mathbf{I}_{N-n} \end{pmatrix} + \begin{pmatrix} \left(\frac{\sigma_v^2}{\sigma_e^2 + \sigma_v^2}\right)^2 \sigma^2 B \mathbf{J}_n & \frac{\sigma_v^2}{\sigma_e^2 + \sigma_v^2} \sigma^2 B \mathbf{1}_n \mathbf{1}_{N-n}' \\ \frac{\sigma_v^2}{\sigma_e^2 + \sigma_v^2} \sigma^2 B \mathbf{1}_{N-n} \mathbf{1}_n' & \sigma^2 B \mathbf{J}_{N-n} \end{pmatrix}. \end{aligned}$$

Logo,

$$\text{tr}[(\mathbf{I}_N - \frac{1}{N} \mathbf{J}_N) \Sigma] = \frac{\sigma_e^2 \sigma_v^2}{\sigma_e^2 + \sigma_v^2} \left[(n-1) + \frac{N-n}{N} B \right] + \frac{N-n}{N} (N-B) \sigma_e^2. \quad (3.4)$$

Utilizando-se o fato de $E[\mathbf{y}|\mathbf{Y}_s] = E\{E[\mathbf{y}|\theta, \mathbf{Y}_s]|\mathbf{Y}_s\}$ e por (3.2) verificamos que

$$E[y_i|\mathbf{Y}_s] = \begin{cases} (1-B)\bar{Y}_s + B\mu & , i \notin s \\ \frac{\sigma_v^2}{\sigma_e^2 + \sigma_v^2} [(1-B)\bar{Y}_s + B\mu] + \frac{\sigma_e^2}{\sigma_e^2 + \sigma_v^2} Y_i & , i \in s \end{cases}$$

Mostramos, depois de algumas manipulações algébricas, que

$$E[\mathbf{y}'|\mathbf{Y}_s](\mathbf{I}_N - \frac{1}{N} \mathbf{J}_N)E[\mathbf{y}|\mathbf{Y}_s] = \left(\frac{\sigma_e^2}{\sigma_e^2 + \sigma_v^2}\right)^2 \left[n s_Y^2 + \frac{n(N-n)}{N} B^2 (\bar{Y}_s - \mu)^2 \right] \quad (3.5)$$

onde $s_Y^2 = \sum_{i \in s} (Y_i - \bar{Y}_s)^2 / n$. Bolfarine e Zacks (1991) mostram que o previsor Bayesiano de S^2 sob perda quadrática é

$$\hat{S}_B^2 = E[S^2|\mathbf{Y}_s] = \frac{1}{N} \text{tr}[(\mathbf{I}_N - \frac{1}{N} \mathbf{J}_N) \Sigma] + \frac{1}{N} E[\mathbf{y}'|\mathbf{Y}_s](\mathbf{I}_N - \frac{1}{N} \mathbf{J}_N)E[\mathbf{y}|\mathbf{Y}_s].$$

Segue de (3.4) e (3.5) que

$$\begin{aligned} \hat{S}_B^2 &= \frac{1}{N} \left\{ \frac{\sigma_e^2 \sigma_v^2}{\sigma_e^2 + \sigma_v^2} \left[(n-1) + \frac{(N-n)}{N} B \right] + \frac{(N-n)}{N} (N-B) \sigma_e^2 \right. \\ &\quad \left. + \left(\frac{\sigma_e^2}{\sigma_e^2 + \sigma_v^2}\right)^2 \left[n s_Y^2 + \frac{n(N-n)}{N} B^2 (\bar{Y}_s - \mu)^2 \right] \right\}. \end{aligned} \quad (3.6)$$

Notamos que o previsor (3.6) se reduz ao previsor (5) dado em Bolfarine (1991) quando $\sigma^2 \rightarrow \infty$.

3.3 PREVISÃO BAYESIANA EMPÍRICA

Na análise Bayesiana empírica os parâmetros μ , σ^2 , σ_e^2 e σ_v^2 são desconhecidos e precisam ser estimados. O enfoque proposto por Ghosh e Meeden (1986) assume que estamos no m -ésimo estágio de um processo de amostragem, e que a amostra foi previamente repetida $m - 1$ vezes. O tamanho da população no j -ésimo estágio do experimento é denotado por N_j e nesse estágio, associamos à l -ésima unidade uma quantidade desconhecida $y_l^{(j)}$ gerada pelo modelo

$$y_l^{(j)} = \theta^{(j)} + e_l^{(j)}$$

$$Y_l^{(j)} = y_l^{(j)} + v_l^{(j)}, \quad (3.7)$$

$l = 1, \dots, N_j$, $j = 1, \dots, m$, onde $\theta^{(j)}$, $e_l^{(j)}$ e $v_l^{(j)}$ são todos independentes e $\theta^{(j)} \sim N(\mu, \sigma^2)$, $e_l^{(j)} \sim N(0, \sigma_e^2)$ e $v_l^{(j)} \sim N(0, \sigma_v^2)$. A principal idéia por trás do modelo (3.7) é que a quantidade $y_l^{(j)}$ não é observada diretamente (por exemplo, é medida com erro) e a quantidade $Y_l^{(j)}$ é um estimador de $y_l^{(j)}$. Uma amostra s_j de tamanho n_j é selecionada no j -ésimo estágio. Definindo $B_j = M/(M + n_j)$, $j = 1, \dots, m$, segue de (3.3) e (3.6) que no estágio m , os previsores Bayesianos de $T^{(m)} = \sum_{l=1}^{N_m} y_l^{(m)}$ e $S^{2(m)} = \sum_{l=1}^{N_m} (y_l^{(m)} - \bar{y}^{(m)})^2 / N_m$ onde $\bar{y}^{(m)} = T^{(m)} / N_m$ são, respectivamente,

$$\hat{T}_B = N_m[(1 - B_m)\bar{Y}_{sm} + B_m\mu] + \frac{n_m\sigma_e^2}{\sigma_e^2 + \sigma_v^2}B_m(\bar{Y}_{sm} - \mu)$$

e

$$\begin{aligned} \hat{S}_B^2 = & \frac{1}{N_m} \left\{ \frac{\sigma_e^2\sigma_v^2}{\sigma_e^2 + \sigma_v^2} \left[(n_m - 1) + \frac{N_m - n_m}{N_m}B_m \right] + \frac{N_m - n_m}{N_m}\sigma_e^2(N_m - B_m) \right. \\ & \left. + \left(\frac{\sigma_e^2}{\sigma_e^2 + \sigma_v^2} \right)^2 \left[n_m s_Y^{2(m)} + \frac{n_m(N_m - n_m)}{N_m}B_m^2(\bar{Y}_{sm} - \mu)^2 \right] \right\} \end{aligned}$$

onde $\bar{Y}_{sm} = \sum_{l \in s_m} Y_l^{(m)} / n_m$ e $s_Y^{2(m)} = \sum_{l \in s_m} (Y_l^{(m)} - \bar{Y}_{sm})^2 / n_m$.

Consideremos que M , μ , σ_e^2 , σ_v^2 e σ^2 são desconhecidos, mas a razão $\lambda = \sigma_e^2 / \sigma_v^2$ é conhecida. Logo, a razão $\sigma_e^2 / (\sigma_e^2 + \sigma_v^2) = \lambda / (\lambda + 1)$ é também conhecida. Fuller (1990) considerou σ_v^2 como conhecida.

Seja $\mathbf{Y}_{s_j} = (Y_1^{(j)}, \dots, Y_{n_j}^{(j)})$, então $\mathbf{Y}_{s_1}, \dots, \mathbf{Y}_{s_m}$ são mutuamente independentes com

$$\mathbf{Y}_{s_j} \sim N_{n_j}(\mu \mathbf{1}_{n_j}; (\sigma_e^2 + \sigma_v^2) \mathbf{I}_{n_j} + \sigma^2 \mathbf{1}_{n_j} \mathbf{1}_{n_j}') ,$$

$j = 1, \dots, m$. Seja $n_T = \sum_{j=1}^m n_j$ e $\bar{Y}_{s.} = \sum_{j=1}^m n_j \bar{Y}_{s_j} / n_T$. Vamos supor que

a) $\inf_j n_j \geq 2$,

b) $\sup_j n_j = k < \infty$ e

c) $\sup_j N_j < N_0 < \infty$,

$j = 1, \dots, m$. Vamos definir

$$\text{BMS} = \sum_{j=1}^m n_j (\bar{Y}_{s_j} - \bar{Y}_{s.})^2 / (m-1) \quad \text{e}$$

$$\text{WMS} = \sum_{j=1}^m \sum_{l=1}^{n_j} (Y_l^{(j)} - \bar{Y}_{s_j})^2 / (n_T - m) .$$

Notemos que BMS e WMS são independentes. O resultado que apresentamos a seguir está relacionado com a estimação de M . Sua prova foi omitida, e uma prova similar pode ser vista em Ghosh e Meeden (1986).

Lema 3.1. Sob as suposições acima,

$$E(\text{WMS}) = \sigma_e^2 + \sigma_v^2$$

e

$$E(\text{BMS}) = \sigma_e^2 + \sigma_v^2 + g(m-1)^{-1} \sigma^2 .$$

Além disso,

$$\text{Var}(\text{WMS}) = 2(\sigma_e^2 + \sigma_v^2)^2 / (n_T - m)$$

e

$$\text{Var}(\text{BMS}) = 2(m-1)^{-2} \left[\sum_{j=1}^m \sigma_j^4 + \frac{\left(\sum_{j=1}^m n_j \sigma_j^2 \right)^2}{n_T^2} - 2 \frac{\sum_{j=1}^m n_j \sigma_j^4}{n_T} \right]$$

onde

$$\sigma_j^2 = \sigma_e^2 + \sigma_v^2 + n_j \sigma^2 \quad \text{e} \quad g = n_T - \sum_{j=1}^m n_j^2 / n_T .$$

Segue do Lema 3.1 que WMS é um estimador consistente de $\sigma_e^2 + \sigma_v^2$, ou seja, $\text{WMS} \xrightarrow{P} \sigma_e^2 + \sigma_v^2$, quando $m \rightarrow \infty$, onde “ $\xrightarrow{P}$ ” denota convergência em probabilidade. Ainda pelo Lema 3.1 e sob as suposições acima, temos

$$\begin{aligned} \text{(i)} \quad & \sum_{j=1}^m \sigma_j^4 = \sum_{j=1}^m (\sigma_e^2 + \sigma_v^2 + n_j \sigma^2)^2 \leq m(\sigma_e^2 + \sigma_v^2 + k\sigma^2)^2, \\ \text{(ii)} \quad & \frac{\left(\sum_{j=1}^m n_j \sigma_j^2\right)^2}{n_T^2} \leq (\max_j \sigma_j^2)^2 = \max_j \sigma_j^4, \\ \text{(iii)} \quad & g = n_T - \frac{\sum_{j=1}^m n_j^2}{n_T} = \frac{2 \sum_{1 \leq j \leq j' \leq m} n_j n_{j'}}{n_T} \geq \frac{4(m-1)}{k}. \end{aligned}$$

Então, por (i), (ii) e (iii), temos que $(\text{BMS} - \text{WMS})g^{-1}(m-1)$ é um estimador consistente de σ^2 . Portanto, a partir do Lema 3.1, segue que um estimador consistente para $M^{-1} = \sigma^2/(\sigma_e^2 + \sigma_v^2)$ é dado por $\left(\frac{\text{BMS}}{\text{WMS}} - 1\right)(m-1)g^{-1}$. No entanto, esse estimador pode ser negativo, então o estimador proposto é

$$\widehat{M}^{-1} = \max \left\{ 0, \left(\frac{\text{BMS}}{\text{WMS}} - 1 \right) (m-1)g^{-1} \right\},$$

que continua sendo consistente.

Como estimador de $B_j = M/(M + n_j) = (1 + n_j M^{-1})^{-1}$ consideramos

$$\widehat{B}_j = (1 + n_j \widehat{M}^{-1})^{-1},$$

$j = 1, \dots, m$. O próximo lema será utilizado na prova dos Teoremas 3.2 e 3.3 que são os principais resultados deste capítulo.

Lema 3.2. Sob as suposições a) e b) apresentadas anteriormente,

$$\max_j |\widehat{B}_j - B_j| \xrightarrow{P} 0$$

quando $m \rightarrow \infty$.

Prova. $|\widehat{B}_j - B_j| = \frac{n_j |\widehat{M}^{-1} - M^{-1}|}{(1 + n_j \widehat{M}^{-1})(1 + n_j M^{-1})} \leq k |\widehat{M}^{-1} - M^{-1}|.$

Logo, $\max_j |\hat{B}_j - B_j| \leq k|\hat{M}^{-1} - M^{-1}| \xrightarrow{P} 0$, pois $\hat{M}^{-1} \xrightarrow{P} M^{-1}$.

A seguir consideramos a estimação de μ . Temos que a densidade conjunta de $(Y_{s_1}, \dots, Y_{s_m})$ é

$$\begin{aligned} h(Y_{s_1}, \dots, Y_{s_m}) &\propto \exp \left[-\frac{1}{2} \sum_{j=1}^m (Y_{s_j} - \mu \mathbf{1}_{n_j})' (\sigma_e^2 + \sigma_v^2)^{-1} \left(\mathbf{I}_{n_j} - \frac{\sigma^2 \mathbf{1}_{n_j} \mathbf{1}_{n_j}'}{\sigma_e^2 + \sigma_v^2 + n_j \sigma^2} \right) (Y_{s_j} - \mu \mathbf{1}_{n_j}) \right] \\ &= \exp \left\{ -\frac{1}{2(\sigma_e^2 + \sigma_v^2)} \sum_{j=1}^m \left[\sum_{l=1}^{n_j} (Y_l^{(j)} - \mu)^2 - \frac{\sigma^2}{\sigma_e^2 + \sigma_v^2 + n_j \sigma^2} \left(\sum_{l=1}^{n_j} (Y_l^{(j)} - \mu) \right)^2 \right] \right\} \\ &= \exp \left\{ -\frac{1}{2(\sigma_e^2 + \sigma_v^2)} \sum_{j=1}^m \left[\sum_{l=1}^{n_j} (Y_l^{(j)} - \bar{Y}_{s_j})^2 + n_j (\mu - \bar{Y}_{s_j})^2 - \frac{\sigma^2}{\sigma_e^2 + \sigma_v^2 + n_j \sigma^2} n_j^2 (\bar{Y}_{s_j} - \mu)^2 \right] \right\} \\ &= \exp \left\{ -\frac{1}{2(\sigma_e^2 + \sigma_v^2)} \sum_{j=1}^m \sum_{l=1}^{n_j} (Y_l^{(j)} - \bar{Y}_{s_j})^2 - \frac{1}{2} \sum_{j=1}^m \frac{n_j}{\sigma_e^2 + \sigma_v^2 + n_j \sigma^2} (\mu - \bar{Y}_{s_j})^2 \right\}. \end{aligned}$$

Segue que o estimador de máxima verossimilhança de μ com B_j , $j = 1, \dots, m$ conhecido é dado por

$$\mu^* = \frac{\sum_{j=1}^m (1 - B_j) \bar{Y}_{s_j}}{\sum_{j=1}^m (1 - B_j)}.$$

Notemos que a condição $\hat{M}^{-1} = 0$ é equivalente a $\hat{B}_1 = \dots = \hat{B}_m = 1$. Então, como estimadores de μ quando B_j é desconhecido, propomos

$$\hat{\mu} = \begin{cases} \frac{\sum_{j=1}^m (1 - \hat{B}_j) \bar{Y}_{s_j}}{\sum_{j=1}^m (1 - \hat{B}_j)}, & \hat{M}^{-1} \neq 0, \\ m^{-1} \sum_{j=1}^m \bar{Y}_{s_j}, & \hat{M}^{-1} = 0. \end{cases}$$

Concluindo, temos que os previsores Bayesianos empíricos de $T^{(m)}$ e $S^{2(m)}$, total e variância populacionais do estágio m , são respectivamente,

$$\hat{T}_{EB} = N_m [(1 - \hat{B}_m) \bar{Y}_{s_m} + \hat{B}_m \hat{\mu}] + n_m \frac{\lambda}{\lambda + 1} \hat{B}_m (\bar{Y}_{s_m} - \hat{\mu})$$

e

$$\begin{aligned} \hat{S}_{EB}^2 &= \frac{1}{N_m} \left\{ \frac{\lambda \text{WMS}}{(\lambda + 1)^2} \left[(n_m - 1) + \frac{(N_m - n_m)}{N_m} \hat{B}_m \right] + \left(\frac{N_m - n_m}{N_m} \right) \frac{\lambda \text{WMS}}{(\lambda + 1)} (N_m - \hat{B}_m) \right. \\ &\quad \left. + \left(\frac{\lambda}{\lambda + 1} \right)^2 \left[n_m s_Y^{2(m)} + \frac{n_m (N_m - n_m)}{N_m} \hat{B}_m^2 (\bar{Y}_{s_m} - \hat{\mu})^2 \right] \right\}. \end{aligned}$$

3.4 RISCO DE BAYES DE $\hat{T}_{EB}$ E $\hat{S}_{EB}^2$

Denotaremos por $r(\xi, e)$ o risco de Bayes (sob perda quadrática) de um previsor e com respeito ao modelo ξ dado em (3.7). A seguir comparamos o previsor $\hat{T}_{EB}$ com os previsores clássicos $\hat{T}_0 = N_m \bar{Y}_{sm}$ e $\hat{T}_1 = N_m \frac{\sum_{j=1}^m n_j \bar{Y}_{sj}}{n_T}$.

Ao invés de compararmos diretamente o risco de Bayes de $\hat{T}_{EB}$ e $\hat{T}$ (onde $\hat{T} = \hat{T}_0$ e $\hat{T}_1$), faremos a comparação em termos de

$$\text{RSL}(\xi; \hat{T}_{EB}, \hat{T}) = \frac{r(\xi, \hat{T}_{EB}) - r(\xi, \hat{T}_B)}{r(\xi, \hat{T}) - r(\xi, \hat{T}_B)}$$

e a importância desse risco relativo (RSL) como uma maneira de comparar previsões é discutida em Efron e Morris (1973) e Ghosh e Meeden (1986). O RSL pode ser visto como uma possível melhora sobre $\hat{T}$ que é perdida ao utilizarmos $\hat{T}_{EB}$ ao invés do previsor Bayesiano $\hat{T}_B$. Estudaremos o comportamento assintótico do RSL quando $m \rightarrow \infty$. O próximo teorema será importante na prova do principal resultado desta seção.

Teorema 3.1.

$$\text{RSL}(\xi; \hat{T}_{EB}, \hat{T}_0) = \frac{E\{[(B_m - \hat{B}_m)(\bar{Y}_{sm} - \mu) - \hat{B}_m(\mu - \hat{\mu})][N_m - n_m\lambda/(\lambda + 1)]\}^2}{B_m^2[N_m - n_m\lambda/(\lambda + 1)]^2 E(\bar{Y}_{sm} - \mu)^2} \quad (3.8)$$

e

$$\text{RSL}(\xi; \hat{T}_{EB}, \hat{T}_1) = \frac{E\{[(B_m - \hat{B}_m)(\bar{Y}_{sm} - \mu) - \hat{B}_m(\mu - \hat{\mu})][N_m - n_m\lambda/(\lambda + 1)]\}^2}{E(N_m \bar{Y}_s - \hat{T}_B)^2} \quad (3.9)$$

Prova. De acordo com o Lema 3 de Ghosh e Meeden (1986) podemos escrever

$$\text{RSL}(\xi; \hat{T}_{EB}, \hat{T}_0) = \frac{E[\hat{T}_{EB} - \hat{T}_B]^2}{E[\hat{T}_0 - \hat{T}_B]^2} \quad \text{e} \quad \text{RSL}(\xi; \hat{T}_{EB}, \hat{T}_1) = \frac{E[\hat{T}_{EB} - \hat{T}_B]^2}{E[\hat{T}_1 - \hat{T}_B]^2}. \quad (3.10)$$

Verificamos que

$$\begin{aligned}
E[\hat{T}_{EB} - \hat{T}_B]^2 &= E\left\{N_m[(1 - \hat{B}_m)\bar{Y}_{s_m} + \hat{B}_m\hat{\mu}] + n_m \frac{\lambda}{\lambda + 1} \hat{B}_m(\bar{Y}_{s_m} - \hat{\mu}) \right. \\
&\quad \left. - N_m[(1 - B_m)\bar{Y}_{s_m} + B_m\mu] - n_m \frac{\lambda}{\lambda + 1} B_m(\bar{Y}_{s_m} - \mu)\right\}^2 \\
&= E\{N_m(\hat{B}_m\hat{\mu} - \hat{B}_m\bar{Y}_{s_m} + B_m\bar{Y}_{s_m} - B_m\mu + \hat{B}_m\mu - \hat{B}_m\mu) \\
&\quad + n_m \frac{\lambda}{\lambda + 1} [\hat{B}_m(\bar{Y}_{s_m} - \hat{\mu}) - B_m(\bar{Y}_{s_m} - \mu) + \hat{B}_m(\bar{Y}_{s_m} - \mu) - \hat{B}_m(\bar{Y}_{s_m} - \mu)]\}^2 \\
&= E\{[(B_m - \hat{B}_m)(\bar{Y}_{s_m} - \mu) - \hat{B}_m(\mu - \hat{\mu})][(N_m - n_m\lambda/(\lambda + 1))]\}^2 \quad (3.11)
\end{aligned}$$

Temos também,

$$\begin{aligned}
E[\hat{T}_0 - \hat{T}_B]^2 &= E\left\{N_m\bar{Y}_{s_m} - N_m[(1 - B_m)\bar{Y}_{s_m} + B_m\mu] - n_m \frac{\lambda}{\lambda + 1} B_m(\bar{Y}_{s_m} - \mu)\right\}^2 \\
&= B_m^2 [N_m - n_m\lambda/(\lambda + 1)]^2 E(\bar{Y}_{s_m} - \mu)^2 \quad (3.12)
\end{aligned}$$

e

$$E[\hat{T}_1 - \hat{T}_B]^2 = E[N_m\bar{Y}_s - \hat{T}_B]^2. \quad (3.13)$$

Portanto, (3.8) e (3.9) segue substituindo (3.11), (3.12) e (3.13) em (3.10).

Teorema 3.2. *Sob as suposições a), b) e c) temos*

$$\text{RSL}(\xi; \hat{T}_{EB}, \hat{T}) \rightarrow 0$$

quando $m \rightarrow \infty$, sendo $\hat{T} = \hat{T}_0$ e $\hat{T}_1$.

Prova. Notemos que

$$E[\bar{Y}_{s_m} - \mu]^2 = \text{Var}[\bar{Y}_{s_m}] = \sigma^2 + n_m^{-1}(\sigma_e^2 + \sigma_v^2) > \sigma^2 > 0. \quad (3.14)$$

Além disso, podemos escrever

$$\begin{aligned}
E[N_m\bar{Y}_s - \hat{T}_B]^2 &= E\left\{N_m\bar{Y}_s - N_m[(1 - B_m)\bar{Y}_{s_m} + B_m\mu] - n_m \frac{\sigma_e^2}{\sigma_e^2 + \sigma_v^2} B_m(\bar{Y}_{s_m} - \mu)\right\}^2 \\
&= E\left\{N_m(\bar{Y}_s - \mu) - (\bar{Y}_{s_m} - \mu) \left[N_m(1 - B_m) + n_m \frac{\sigma_e^2}{\sigma_e^2 + \sigma_v^2} B_m\right]\right\}^2
\end{aligned}$$

$$\begin{aligned}
&= N_m^2 E[\bar{Y}_{s.} - \mu]^2 + \left[N_m(1 - B_m) + n_m \frac{\sigma_e^2}{\sigma_e^2 + \sigma_v^2} B_m \right]^2 E[\bar{Y}_{s_m} - \mu]^2 \\
&\quad - 2N_m \left[N_m(1 - B_m) + n_m \frac{\sigma_e^2}{\sigma_e^2 + \sigma_v^2} B_m \right] E[\bar{Y}_{s.} - \mu][\bar{Y}_{s_m} - \mu] \quad (3.15)
\end{aligned}$$

onde

$$\begin{aligned}
E[\bar{Y}_{s.} - \mu]^2 &= \text{Var}[\bar{Y}_{s.}] = \frac{1}{n_T^2} \sum_{j=1}^m n_j^2 [\sigma^2 + (\sigma_e^2 + \sigma_v^2)/n_j] \\
&\leq \frac{1}{(2m)^2} \sum_{j=1}^m k^2 (\sigma^2 + \sigma_e^2 + \sigma_v^2) = O(m^{-1}) .
\end{aligned}$$

Então, como $N_m < N_0 < \infty$ para todo m ,

$$N_m^2 E[\bar{Y}_{s.} - \mu]^2 = O(m^{-1}) .$$

Além disso,

$$\begin{aligned}
E[(\bar{Y}_{s.} - \mu)(\bar{Y}_{s_m} - \mu)] &= E \left[\frac{1}{n_T} \bar{Y}_{s_m} \sum_{j=1}^m n_j \bar{Y}_{s_j} \right] - \mu^2 \\
&= \frac{1}{n_T} \left[\sum_{j=1}^{m-1} n_j \mu^2 + n_m E(\bar{Y}_{s_m}^2) \right] - \mu^2 = \frac{1}{n_T} \{ \mu^2 n_T + n_m [\sigma^2 + (\sigma_e^2 + \sigma_v^2)/n_m] \} - \mu^2 \\
&\leq \frac{k(\sigma^2 + \sigma_e^2 + \sigma_v^2)}{2m} = O(m^{-1})
\end{aligned}$$

o que implica

$$2N_m \left[N_m(1 - B_m) + n_m \frac{\sigma_e^2}{\sigma_e^2 + \sigma_v^2} B_m \right] E[\bar{Y}_{s.} - \mu][\bar{Y}_{s_m} - \mu] = O(m^{-1}) .$$

Logo, existe m_0 tal que

$$E[N_m \bar{Y}_{s.} - \hat{T}_B]^2 > \left(\frac{\sigma_e^2}{\sigma_e^2 + \sigma_v^2 + k\sigma^2} \right)^2 \sigma^2 \quad (3.16)$$

para todo $m > m_0$.

De (3.14), (3.16) e Teorema 3.1 para mostrarmos que $\text{RSL}(\xi; \hat{T}_{EB}, \hat{T}) \rightarrow 0$, basta verificarmos que

$$E\{[(B_m - \hat{B}_m)(\bar{Y}_{s_m} - \mu) - \hat{B}_m(\mu - \hat{\mu})][N_m - n_m \lambda/(\lambda + 1)]\}^2 \rightarrow 0 \quad (3.17)$$

quando $m \rightarrow \infty$.

Para provarmos (3.17) é suficiente mostrarmos que

$$A_m = [(B_m - \hat{B}_m)(\bar{Y}_{s_m} - \mu) - \hat{B}_m(\mu - \hat{\mu})][N_m - n_m\lambda/(\lambda + 1)] \xrightarrow{P} 0$$

quando $m \rightarrow \infty$ e que A_m^2 é uniformemente integrável em m (ver Apêndice 3).

Como $\bar{Y}_{s_m} - \mu \sim N(0; \sigma^2 + (\sigma_e^2 + \sigma_v^2)/n_m)$, então

$$|\bar{Y}_{s_m} - \mu| \stackrel{d}{=} [\sigma^2 + (\sigma_e^2 + \sigma_v^2)/n_m]^{1/2} \chi_1 \stackrel{\text{ST}}{\leq} (\sigma^2 + \sigma_e^2 + \sigma_v^2)^{1/2} \chi_1 = O_p(1)$$

onde a notação $\stackrel{\text{ST}}{\leq}$ é usada para denotar estocasticamente menor e “ $\stackrel{d}{=}$ ” para denotarmos igualdade em distribuição.

Pelo Lema 3.2 temos

$$|B_m - \hat{B}_m| |\bar{Y}_{s_m} - \mu| \leq \max_j |B_j - \hat{B}_j| |\bar{Y}_{s_m} - \mu| = o_p(1)$$

o que implica que

$$(B_m - \hat{B}_m)(\bar{Y}_{s_m} - \mu) \xrightarrow{P} 0 \quad (3.18)$$

quando $m \rightarrow \infty$. Além disso, como $0 \leq \hat{B}_m \leq 1$ e $(\hat{\mu} - \mu) \xrightarrow{P} 0$ (Ghosh e Meeden (1986)) temos

$$\hat{B}_m(\hat{\mu} - \mu) \xrightarrow{P} 0. \quad (3.19)$$

Como $N_m - n_m\lambda/(\lambda + 1) = O(1)$, segue de (3.18) e (3.19) que $A_m \xrightarrow{P} 0$ quando $m \rightarrow \infty$.

Agora provamos a integrabilidade uniforme de A_m^2 . Inicialmente, notemos que

$$\begin{aligned} & \{[(B_m - \hat{B}_m)(\bar{Y}_{s_m} - \mu) - \hat{B}_m(\mu - \hat{\mu})](N_m - n_m\lambda/(\lambda + 1))\}^2 \\ & \leq 2[N_m - n_m\lambda/(\lambda + 1)]^2 [(B_m - \hat{B}_m)^2 (\bar{Y}_{s_m} - \mu)^2 + \hat{B}_m^2 (\mu - \hat{\mu})^2] \\ & \leq 2N_m^2 [(\bar{Y}_{s_m} - \mu)^2 + (\mu - \hat{\mu})^2], \end{aligned}$$

onde

$$(\bar{Y}_{s_m} - \mu)^2 \stackrel{d}{=} [\sigma^2 + (\sigma_e^2 + \sigma_v^2)n_m^{-1}] \chi_1^2 \stackrel{\text{ST}}{\leq} (\sigma^2 + \sigma_e^2 + \sigma_v^2) \chi_1^2$$

então $(\bar{Y}_{s_m} - \mu)^2$ é uniformemente integrável em m (Apêndice 3).

Além disso,

$$(\hat{\mu} - \mu)^2 = \left[m^{-1} \sum_{j=1}^m (\bar{Y}_{s_j} - \mu) \right]^2 I_{\{\hat{M}^{-1}=0\}} + \frac{\left[\sum_{j=1}^m (1 - \hat{B}_j)(\bar{Y}_{s_j} - \mu) \right]^2}{\left[\sum_{j=1}^m (1 - \hat{B}_j) \right]^2} I_{\{\hat{M}^{-1}>0\}} .$$

Para $\hat{M}^{-1} > 0$ temos

$$2(\hat{M} + 2)^{-1} \leq 1 - \hat{B}_j = n_j(\hat{M} + n_j)^{-1} \leq k(\hat{M} + 2)^{-1} .$$

Então

$$\begin{aligned} (\hat{\mu} - \mu)^2 &\leq m^{-1} \sum_{j=1}^m (\bar{Y}_{s_j} - \mu)^2 + \frac{1}{4} k^2 m^{-1} \sum_{j=1}^m (\bar{Y}_{s_j} - \mu)^2 \\ &\stackrel{\text{ST}}{\leq} (\sigma^2 + \sigma_e^2 + \sigma_v^2) m^{-1} \sum_{j=1}^m U_j + \frac{1}{4} k^2 (\sigma^2 + \sigma_e^2 + \sigma_v^2) m^{-1} \sum_{j=1}^m U_j \end{aligned}$$

onde $U_j \sim \chi_1^2$, para todo j . Segue, portanto, a integrabilidade uniforme de A_m^2 em m .

Observando a definição do RSL e pelos resultados dos Teoremas 3.1 e 3.2 temos que $r(\xi; \hat{T}_{EB}) - r(\xi; \hat{T}_B) \rightarrow 0$ quando $m \rightarrow \infty$, então podemos dizer que o previsor Bayesiano empírico para o total populacional é assintoticamente ótimo no sentido de Robbins (1955). No próximo teorema mostramos que o previsor Bayesiano empírico para a variância populacional também é assintoticamente ótimo no sentido de Robbins (1955).

Teorema 3.3. *Sob as suposições a), b) e c) segue que*

$$r(\xi; \hat{S}_{EB}^2) - r(\xi; \hat{S}_B^2) \rightarrow 0$$

quando $m \rightarrow \infty$.

Prova. Utilizando-se o Lema 3 de Ghosh e Meeden (1986) temos

$$r(\xi, \hat{S}_{EB}^2) - r(\xi, \hat{S}_B^2) = E[\hat{S}_{EB}^2 - \hat{S}_B^2]^2$$

$$\begin{aligned}
= & E \left[\frac{\lambda WMS (n_m - 1)}{(\lambda + 1)^2 N_m} + \frac{\lambda WMS (N_m - n_m)}{(\lambda + 1)^2 N_m^2} \hat{B}_m + \frac{(N_m - n_m)}{N_m^2} \frac{\lambda WMS}{\lambda + 1} (N_m - \hat{B}_m) \right. \\
& + \left(\frac{\lambda}{\lambda + 1} \right)^2 \frac{(N_m - n_m)}{N_m} n_m \hat{B}_m^2 (\bar{Y}_{sm} - \hat{\mu})^2 - \frac{\sigma_e^2}{\lambda + 1} \frac{(n_m - 1)}{N_m} - \frac{\sigma_e^2}{\lambda + 1} \frac{(N_m - n_m)}{N_m^2} B_m \\
& \left. - \frac{(N_m - n_m)}{N_m^2} \sigma_e^2 (N_m - B_m) - \left(\frac{\lambda}{\lambda + 1} \right)^2 \frac{(N_m - n_m)}{N_m^2} n_m B_m^2 (\bar{Y}_{sm} - \mu)^2 \right]^2
\end{aligned}$$

Agora, somando e subtraindo, $\frac{\sigma_e^2}{\lambda+1} \frac{(N_m-n_m)}{N_m^2} \hat{B}_m$ e $\frac{N_m-n_m}{N_m^2} \sigma_e^2 (N_m - \hat{B}_m)$, temos

$$\begin{aligned}
& r(\xi, \hat{S}_{EB}^2) - r(\xi, \hat{S}_B^2) \\
= & E \left\{ \frac{(n_m - 1)}{(\lambda + 1) N_m} \left(\frac{\lambda WMS}{\lambda + 1} - \sigma_e^2 \right) + \frac{(N_m - n_m)}{(\lambda + 1) N_m^2} \hat{B}_m \left(\frac{\lambda WMS}{\lambda + 1} - \sigma_e^2 \right) \right. \\
& + \frac{\sigma_e^2}{\lambda + 1} \frac{(N_m - n_m)}{N_m^2} (\hat{B}_m - B_m) + \frac{(N_m - n_m)}{N_m^2} (N_m - \hat{B}_m) \left(\frac{\lambda WMS}{\lambda + 1} - \sigma_e^2 \right) \\
& - \frac{(N_m - n_m)}{N_m^2} \sigma_e^2 (\hat{B}_m - B_m) + \left(\frac{\lambda}{\lambda + 1} \right)^2 \frac{(N_m - n_m)}{N_m} n_m [(\bar{Y}_{sm} - \mu)^2 (\hat{B}_m^2 - B_m^2) \\
& + \hat{B}_m^2 (\hat{\mu} - \mu)^2 - 2 \hat{B}_m (\bar{Y}_{sm} - \mu) (\hat{\mu} - \mu)] \left. \right\}^2 \\
= & E \left\{ \left(\frac{\lambda WMS}{\lambda + 1} - \sigma_e^2 \right) \left[\frac{n_m - 1}{N_m (\lambda + 1)} + \frac{N_m - n_m}{N_m} \right] - \sigma_e^2 \frac{(N_m - n_m)}{N_m^2} \frac{\lambda}{\lambda + 1} (\hat{B}_m - B_m) \right. \\
& - \frac{(N_m - n_m)}{N_m^2} \hat{B}_m \left(\frac{\lambda WMS}{\lambda + 1} - \sigma_e^2 \right) \frac{\lambda}{\lambda + 1} + \left(\frac{\lambda}{\lambda + 1} \right)^2 \frac{(N_m - n_m)}{N_m} n_m [\hat{B}_m^2 (\hat{\mu} - \mu)^2 \\
& + (\bar{Y}_{sm} - \mu)^2 (\hat{B}_m^2 - B_m^2) - 2 \hat{B}_m^2 (\bar{Y}_{sm} - \mu) (\hat{\mu} - \mu)] \left. \right\}^2
\end{aligned}$$

$$\begin{aligned}
&\leq 4E \left\{ \left[\frac{n_m-1}{N_m(\lambda+1)} + \frac{N_m-n_m}{N_m} \right]^2 \left(\frac{\lambda WMS}{\lambda+1} - \sigma_e^2 \right)^2 + \left[\sigma_e^2 \frac{(N_m-n_m)}{N_m^2} \frac{\lambda}{\lambda+1} \right]^2 (\hat{B}_m - B_m)^2 \right. \\
&\quad + \left[\frac{(N_m-n_m)}{N_m^2} \frac{\lambda}{\lambda+1} \right]^2 \hat{B}_m^2 \left(\frac{\lambda WMS}{\lambda+1} - \sigma_e^2 \right)^2 + \left[\left(\frac{\lambda}{\lambda+1} \right)^2 \frac{(N_m-n_m)}{N_m^2} n_m \right]^2 [\hat{B}_m^2 (\hat{\mu} - \mu)^2 \\
&\quad \left. + (\bar{Y}_{sm} - \mu)^2 (\hat{B}_m^2 - B_m^2) - 2\hat{B}_m^2 (\bar{Y}_{sm} - \mu)(\hat{\mu} - \mu) \right]^2 \Big\} \\
&\leq 4 \left\{ 2E \left(\frac{\lambda WMS}{\lambda+1} - \sigma_e^2 \right)^2 + (\sigma_e^2)^2 E(\hat{B}_m - B_m)^2 + E[(\bar{Y}_{sm} - \mu)^2 (\hat{B}_m^2 - B_m^2) \right. \\
&\quad \left. + \hat{B}_m^2 (\hat{\mu} - \mu)^2 - 2\hat{B}_m^2 (\bar{Y}_{sm} - \mu)(\hat{\mu} - \mu)]^2 \right\}. \tag{3.20}
\end{aligned}$$

Mas

$$E(\hat{B}_m - B_m)^2 \leq E[k^2(\hat{M}^{-1} - M^{-1})^2] = k^2 \{ \text{Var}(\hat{M}^{-1}) + [E(\hat{M}^{-1}) - M^{-1}]^2 \} \rightarrow 0 \tag{3.21}$$

quando $m \rightarrow \infty$ pois $\hat{M}^{-1}$ é um estimador consistente de M^{-1} . Como $\text{Var} \left[\frac{\lambda WMS}{\lambda+1} \right] = \left[\frac{\lambda}{\lambda+1} \right]^2 \text{Var}(WMS) \rightarrow 0$ segue que

$$E \left[\frac{\lambda WMS}{\lambda+1} - \sigma_e^2 \right]^2 \rightarrow 0 \quad \text{quando } m \rightarrow \infty. \tag{3.22}$$

Além disso,

$$\begin{aligned}
&(\bar{Y}_{sm} - \mu)^2 |\hat{B}_m^2 - B_m^2| \leq (\bar{Y}_{sm} - \mu)^2 [(\hat{B}_m - B_m)^2 + 2B_m |\hat{B}_m - B_m|] \\
&\leq (\bar{Y}_{sm} - \mu)^2 [\max_j (\hat{B}_j - B_j)^2 + 2 \max_j |\hat{B}_j - B_j|] = o_p(1) \tag{3.23}
\end{aligned}$$

Temos também que

$$\hat{B}_m^2 (\hat{\mu} - \mu)^2 \leq (\hat{\mu} - \mu)^2 = o_p(1) \tag{3.24}$$

e

$$2\hat{B}_m^2 (\bar{Y}_{sm} - \mu)(\hat{\mu} - \mu) \leq 2(\bar{Y}_{sm} - \mu)(\hat{\mu} - \mu) = o_p(1). \tag{3.25}$$

Então, de (3.23), (3.24) e (3.25) segue

$$(\bar{Y}_{s_m} - \mu)^2(\hat{B}_m^2 - B_m^2) + \hat{B}_m^2(\hat{\mu} - \mu)^2 - 2\hat{B}_m^2(\bar{Y}_{s_m} - \mu)(\hat{\mu} - \mu) \xrightarrow{P} 0 \quad (3.26)$$

quando $m \rightarrow \infty$. Para provar que a esperança em (3.20) converge para zero, falta provarmos a integrabilidade uniforme de

$$[(\bar{Y}_{s_m} - \mu)^2(\hat{B}_m^2 - B_m^2) + \hat{B}_m^2(\hat{\mu} - \mu)^2 - 2\hat{B}_m^2(\bar{Y}_{s_m} - \mu)(\hat{\mu} - \mu)]^2. \quad (3.27)$$

Mas

$$\begin{aligned} & [(\bar{Y}_{s_m} - \mu)^2(\hat{B}_m^2 - B_m^2) + \hat{B}_m^2(\hat{\mu} - \mu)^2 - 2\hat{B}_m^2(\bar{Y}_{s_m} - \mu)(\hat{\mu} - \mu)]^2 \\ & \leq 3[(\bar{Y}_{s_m} - \mu)^4(\hat{B}_m^2 - B_m^2)^2 + \hat{B}_m^4(\hat{\mu} - \mu)^4 + 4\hat{B}_m^4(\bar{Y}_{s_m} - \mu)^2(\hat{\mu} - \mu)^2] \\ & \leq 3[(\bar{Y}_{s_m} - \mu)^4 + (\hat{\mu} - \mu)^4 + 4(\bar{Y}_{s_m} - \mu)^2(\hat{\mu} - \mu)^2] \end{aligned}$$

onde

$$(\bar{Y}_{s_m} - \mu)^4 = [(\bar{Y}_{s_m} - \mu)^2]^2 \stackrel{d}{=} [\sigma^2 + (\sigma_e^2 + \sigma_v^2)n_m^{-1}]^2(\chi_1^2)^2 \stackrel{\text{ST}}{\leq} (\sigma^2 + \sigma_e^2 + \sigma_v^2)^2(\chi_1^2)^2$$

que é uniformemente integrável em m . Além disso,

$$\begin{aligned} & (\hat{\mu} - \mu)^4 = [(\hat{\mu} - \mu)^2]^2 \\ & = \left\{ \left[m^{-1} \sum_{j=1}^m (\bar{Y}_{s_j} - \mu) \right]^2 I_{\{\hat{M}^{-1}=0\}} + \frac{[\sum_{j=1}^m (1 - \hat{B}_j)(\bar{Y}_{s_j} - \mu)]^2}{[\sum_{j=1}^m (1 - \hat{B}_j)]^2} I_{\{\hat{M}^{-1}>0\}} \right\}^2 \\ & \leq 2 \left\{ \left\{ \left[m^{-1} \sum_{j=1}^m (\bar{Y}_{s_j} - \mu) \right]^2 \right\}^2 I_{\{\hat{M}^{-1}=0\}} + \left\{ \frac{[\sum_{j=1}^m (1 - \hat{B}_j)(\bar{Y}_{s_j} - \mu)]^2}{[\sum_{j=1}^m (1 - \hat{B}_j)]^2} \right\}^2 I_{\{\hat{M}^{-1}>0\}} \right\} \\ & \leq 2 \left\{ \left[m^{-1} \sum_{j=1}^m (\bar{Y}_{s_j} - \mu)^2 \right]^2 I_{\{\hat{M}^{-1}=0\}} + \left[\frac{k^2}{4} m^{-1} \sum_{j=1}^m (\bar{Y}_{s_j} - \mu)^2 \right]^2 I_{\{\hat{M}^{-1}>0\}} \right\} \\ & \leq 2 \left[\frac{k^2}{4} m^{-1} \sum_{j=1}^m (\bar{Y}_{s_j} - \mu)^2 \right]^2 \leq \frac{k^4}{8} m^{-1} \sum_{j=1}^m [(\bar{Y}_{s_j} - \mu)^2]^2 \\ & \stackrel{\text{ST}}{\leq} \frac{k^4}{8} (\sigma^2 + \sigma_e^2 + \sigma_v^2)^2 \left(m^{-1} \sum_{j=1}^m U_j^2 \right) \end{aligned}$$

onde $U_j \sim \chi_1^2$, portanto $(\hat{\mu} - \mu)^4$ é uniformemente integrável em m .

Finalmente,

$$\begin{aligned} (\bar{Y}_{s_m} - \mu)^2 (\hat{\mu} - \mu)^2 &\leq (\bar{Y}_{s_m} - \mu)^2 \frac{k^2}{4} m^{-1} \sum_{j=1}^m (\bar{Y}_{s_j} - \mu)^2 \\ &= \frac{k^2}{4} m^{-1} \sum_{j=1}^m (\bar{Y}_{s_j} - \mu)^2 (\bar{Y}_{s_m} - \mu)^2 \stackrel{\text{ST}}{\leq} \frac{k^2}{4} (\sigma^2 + \sigma_e^2 + \sigma_v^2)^2 m^{-1} \sum_{j=1}^m U_j U_m \end{aligned}$$

onde $U_j \sim \chi_1^2$, $U_m \sim \chi_1^2$ e U_j e U_m são independentes para $j \neq m$, o que prova a integrabilidade de (3.27). Juntamente com (3.21), (3.22) e (3.26) concluímos a prova.

Consideramos a previsão do total e da variância populacionais do estágio m , mas se o interesse for prever simultaneamente totais e variâncias de vários estágios, os previsores obtidos podem ser generalizados, e resultados similares aos Teoremas 3.2 e 3.3 podem ser obtidos (Ghosh e Lahiri (1987a) e (1987b)).

3.5 O MODELO DE LOCAÇÃO “NÃO HOMOGÊNEO” COM ERRO NAS VARIÁVEIS

Nesta seção consideramos uma versão com erro nas variáveis do modelo de locação “não homogêneo” considerado por Tiwari e Lahiri (1989). Suponhamos que instrumentos de dois laboratórios sejam utilizados para medir uma determinada característica e que suspeitamos que os dois conjuntos de observações tenham a mesma média, mas variâncias diferentes. Nesse caso, o modelo considerado na seção 3.1 não se aplica. Um modelo apropriado pode ser

$$y_i = \theta + e_i$$

$$Y_i = y_i + v_i$$

$j = 1, \dots, N$, onde e_i , v_i e θ são independentes sendo $\theta \sim N(\mu, \sigma^2)$, $e_1, \dots, e_{N_1} \sim N(0, \sigma_e^2)$, $e_{N_1+1}, \dots, e_N \sim N(0, c\sigma_e^2)$ e $v_i \sim N(0, \sigma_v^2)$, $1 \leq N_1 < N$ e sem perda de generalidade assumimos $0 < c < 1$. A amostra observada é denotada por $\mathbf{Y}_s = (\mathbf{Y}_{s_1}, \mathbf{Y}_{s_2})'$,

onde $\mathbf{Y}_{s_1} = (Y_1, \dots, Y_{n_1})'$ são observados entre os primeiros N_1 valores de $\mathbf{Y}$ e $\mathbf{Y}_{s_2} = (Y_{N_1+1}, \dots, Y_{N_1+n_2})'$ são observados entre os últimos N_2 valores de $\mathbf{Y}$.

3.5.1 PREVISÃO BAYESIANA DE T E S^2

Dada a amostra observada $\mathbf{Y}_s$, e considerando a priori $\theta \sim N(\mu, \sigma^2)$, segue que a distribuição a posteriori de θ é

$$\theta | \mathbf{Y}_s \sim N(\mu_s; \sigma_s^2) \quad (3.28)$$

onde

$$\mu_s = \frac{n_1 \sigma^2 (c\sigma_e^2 + \sigma_v^2) \bar{Y}_{s_1} + n_2 \sigma^2 (\sigma_e^2 + \sigma_v^2) \bar{Y}_{s_2} + (\sigma_e^2 + \sigma_v^2)(c\sigma_e^2 + \sigma_v^2)\mu}{n_1 \sigma^2 (c\sigma_e^2 + \sigma_v^2) + n_2 \sigma^2 (\sigma_e^2 + \sigma_v^2) + (\sigma_e^2 + \sigma_v^2)(c\sigma_e^2 + \sigma_v^2)}$$

e

$$\sigma_s^2 = \frac{\sigma^2 (\sigma_e^2 + \sigma_v^2)(c\sigma_e^2 + \sigma_v^2)}{n_1 \sigma^2 (c\sigma_e^2 + \sigma_v^2) + n_2 \sigma^2 (\sigma_e^2 + \sigma_v^2) + (\sigma_e^2 + \sigma_v^2)(c\sigma_e^2 + \sigma_v^2)}$$

com $\bar{Y}_{s_1} = \sum_{i=1}^{n_1} Y_i / n_1$ e $\bar{Y}_{s_2} = \sum_{i=N_1+1}^{N_1+n_2} Y_i / n_2$. Como

$$E[y_i | \theta, \mathbf{Y}_s] = \begin{cases} \theta & , \quad i \notin s \\ \theta + \frac{\sigma_e^2}{\sigma_e^2 + \sigma_v^2} (Y_i - \theta) & , \quad i = 1, \dots, n_1 \\ \theta + \frac{c\sigma_e^2}{c\sigma_e^2 + \sigma_v^2} (Y_i - \theta) & , \quad i = N_1 + 1, \dots, N_1 + n_2 \end{cases} \quad (3.29)$$

segue que o previsor Bayesiano de T é dado por

$$\begin{aligned} \hat{T}_B &= E[T | \mathbf{Y}_s] = E \left[\sum_{i=1}^{n_1} y_i / \mathbf{Y}_s \right] + E \left[\sum_{i=N_1+1}^{N_1+n_2} y_i / \mathbf{Y}_s \right] + E \left[\sum_{i \notin s} y_i / \mathbf{Y}_s \right] \\ &= E \left[E \left(\sum_{i=1}^{n_1} y_i | \theta, \mathbf{Y}_s \right) / \mathbf{Y}_s \right] + E \left[E \left(\sum_{i=N_1+1}^{N_1+n_2} y_i | \theta, \mathbf{Y}_s \right) / \mathbf{Y}_s \right] \\ &\quad + E \left[E \left(\sum_{i \notin s} y_i | \theta, \mathbf{Y}_s \right) / \mathbf{Y}_s \right] \end{aligned} \quad (3.30)$$

$$= N[B_1 \bar{Y}_{s_1} + B_2 \bar{Y}_{s_2} + (1 - B_1 - B_2)\mu] + \frac{n_1 \sigma_e^2}{\sigma_e^2 + \sigma_v^2} [(1 - B_1) \bar{Y}_{s_1} - B_2 \bar{Y}_{s_2}]$$

$$-(1 - B_1 - B_2)\mu] + \frac{n_2 c \sigma_e^2}{c \sigma_e^2 + \sigma_v^2} [(1 - B_2) \bar{Y}_{s_2} - B_1 \bar{Y}_{s_1} - (1 - B_1 - B_2)\mu]$$

onde

$$B_1 = \frac{n_1 M_2}{n_1 M_2 + n_2 M_1 + M_1 M_2}, \quad B_2 = \frac{n_2 M_1}{n_1 M_2 + n_2 M_1 + M_1 M_2}$$

com

$$M_1 = \frac{\sigma_e^2 + \sigma_v^2}{\sigma^2} \quad \text{e} \quad M_2 = \frac{c \sigma_e^2 + \sigma_v^2}{\sigma^2}.$$

De acordo com a seção 3.2 o previsor Bayesiano de S^2 sob perda quadrática é

$$\hat{S}_B^2 = \frac{1}{N} \text{tr}[(\mathbf{I}_N - \frac{1}{N} \mathbf{J}_N) \boldsymbol{\Sigma}] + \frac{1}{N} E[\mathbf{y}' | \mathbf{Y}_s] (\mathbf{I}_N - \frac{1}{N} \mathbf{J}_N) E[\mathbf{y} | \mathbf{Y}_s]$$

onde $\boldsymbol{\Sigma} = \text{Var}(\mathbf{y} | \mathbf{Y}_s)$.

De (3.28), (3.29) e como

$$\text{Var}(y_i | \theta, \mathbf{Y}_s) = \begin{cases} \frac{\sigma_e^2 \sigma_v^2}{\sigma_e^2 + \sigma_v^2} & , \quad i = 1, \dots, n_1 \\ \sigma_e^2 & , \quad i = n_1 + 1, \dots, N_1 \\ \frac{c \sigma_e^2 \sigma_v^2}{c \sigma_e^2 + \sigma_v^2} & , \quad i = N_1 + 1, \dots, N_1 + n_2 \\ c \sigma_e^2 & , \quad i = N_1 + n_2 + 1, \dots, N \end{cases}$$

temos

$$\boldsymbol{\Sigma} = \begin{pmatrix} \frac{\sigma_e^2 \sigma_v^2}{\sigma_e^2 + \sigma_v^2} \mathbf{I}_{n_1} & \mathbf{0} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \sigma_e^2 \mathbf{I}_{N_1 - n_1} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \frac{c \sigma_e^2 \sigma_v^2}{c \sigma_e^2 + \sigma_v^2} \mathbf{I}_{n_2} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{0} & c \sigma_e^2 \mathbf{I}_{N - N_1 - n_2} \end{pmatrix} +$$

$$\begin{pmatrix} \left(\frac{\sigma_v^2}{\sigma_e^2 + \sigma_v^2}\right)^2 \sigma^2 B_{12} \mathbf{J}_{n_1} & \left(\frac{\sigma_v^2}{\sigma_e^2 + \sigma_v^2}\right)^2 \sigma^2 B_{12} \mathbf{1}_{n_1} \mathbf{1}_{N_1 - n_1}' & \frac{(\sigma_v^2) \sigma^2 B_{12}}{(\sigma_e^2 + \sigma_v^2)(c \sigma_e^2 + \sigma_v^2)} \mathbf{1}_{n_1} \mathbf{1}_{n_2}' & \frac{\sigma_v^2}{\sigma_e^2 + \sigma_v^2} \sigma^2 B_{12} \mathbf{1}_{n_1} \mathbf{1}_{N - N_1 - n_2}' \\ \left(\frac{\sigma_v^2}{\sigma_e^2 + \sigma_v^2}\right)^2 \sigma^2 B_{12} \mathbf{1}_{N_1 - n_1} \mathbf{1}_{n_1}' & \sigma^2 B_{12} \mathbf{J}_{N_1 - n_1} & \left(\frac{\sigma_v^2}{c \sigma_e^2 + \sigma_v^2}\right)^2 \sigma^2 B_{12} \mathbf{1}_{N_1 - n_1} \mathbf{1}_{n_2}' & \sigma^2 B_{12} \mathbf{1}_{N_1 - n_1} \mathbf{1}_{N - N_1 - n_2}' \\ \frac{(\sigma_v^2) \sigma^2 B_{12}}{(\sigma_e^2 + \sigma_v^2)(c \sigma_e^2 + \sigma_v^2)} \mathbf{1}_{n_2} \mathbf{1}_{n_1}' & \left(\frac{\sigma_v^2}{c \sigma_e^2 + \sigma_v^2}\right)^2 \sigma^2 B_{12} \mathbf{1}_{n_2} \mathbf{1}_{N_1 - n_1}' & \left(\frac{\sigma_v^2}{c \sigma_e^2 + \sigma_v^2}\right)^2 \sigma^2 B_{12} \mathbf{J}_{n_2} & \left(\frac{\sigma_v^2}{c \sigma_e^2 + \sigma_v^2}\right)^2 \sigma^2 B_{12} \mathbf{1}_{n_2} \mathbf{1}_{N - N_1 - n_2}' \\ \left(\frac{\sigma_v^2}{\sigma_e^2 + \sigma_v^2}\right)^2 \sigma^2 B_{12} \mathbf{1}_{N - N_1 - n_2} \mathbf{1}_{n_1}' & \sigma^2 B_{12} \mathbf{1}_{N - N_1 - n_2} \mathbf{1}_{N_1 - n_1}' & \left(\frac{\sigma_v^2}{c \sigma_e^2 + \sigma_v^2}\right)^2 \sigma^2 B_{12} \mathbf{1}_{N - N_1 - n_2} \mathbf{1}_{n_2}' & \sigma^2 B_{12} \mathbf{J}_{N - N_1 - n_1} \end{pmatrix}$$

onde $B_{12} = 1 - B_1 - B_2$. Logo

$$\begin{aligned}
\text{tr} \left[\left(\mathbf{I}_N - \frac{1}{N} \mathbf{J}_N \right) \Sigma \right] &= \frac{\sigma_e^2 \sigma_v^2}{\sigma_e^2 + \sigma_v^2} \left[(n_1 - B_1) - \frac{n_1}{N} (1 - B_1 - B_2) \right] \\
&+ \frac{c \sigma_e^2 \sigma_v^2}{c \sigma_e^2 + \sigma_v^2} \left[(n_2 - B_2) - \frac{n_2}{N} (1 - B_1 - B_2) \right] \\
&+ \frac{N_1 - n_1}{N} \sigma_e^2 [N - (1 - B_1 - c B_2)] \\
&+ \frac{N - N_1 - n_2}{N} c \sigma_e^2 [N - (1 - c^{-1} B_1 - B_2)]
\end{aligned} \tag{3.31}$$

Utilizando-se (3.28) temos que

$$E[y_i | \mathbf{Y}_s] = \begin{cases} B_1 \bar{Y}_{s_1} + B_2 \bar{Y}_{s_2} + B_{12} \mu, & i \notin S \\ \frac{\sigma_v^2}{\sigma_e^2 + \sigma_v^2} [B_1 \bar{Y}_{s_1} + B_2 \bar{Y}_{s_2} + B_{12} \mu] + \frac{\sigma_e^2}{\sigma_e^2 + \sigma_v^2} Y_i, & i = 1, \dots, n_1 \\ \frac{\sigma_v^2}{c \sigma_e^2 + \sigma_v^2} [B_1 \bar{Y}_{s_1} + B_2 \bar{Y}_{s_2} + B_{12} \mu] + \frac{c \sigma_e^2}{c \sigma_e^2 + \sigma_v^2} Y_i, & i = N_1 + 1, \dots, N_1 + n_2 \end{cases}$$

e verificamos que

$$\begin{aligned}
&E[\mathbf{y}' | \mathbf{Y}_s] \left(\mathbf{I}_N - \frac{1}{N} \mathbf{J}_N \right) E[\mathbf{y} | \mathbf{Y}_s] \\
&= \left(\frac{\sigma_e^2}{\sigma_e^2 + \sigma_v^2} \right)^2 \left\{ n_1 s_{Y_1}^2 + n_1 \left(1 - \frac{n_1}{N} \right) [(1 - B_1) \bar{Y}_{s_1} - B_2 \bar{Y}_{s_2} - (1 - B_1 - B_2) \mu]^2 \right\} \\
&+ \left(\frac{c \sigma_e^2}{c \sigma_e^2 + \sigma_v^2} \right)^2 \left\{ n_2 s_{Y_2}^2 + n_2 \left(1 - \frac{n_2}{N} \right) [(1 - B_2) \bar{Y}_{s_2} - B_1 \bar{Y}_{s_1} - (1 - B_1 - B_2) \mu]^2 \right\} \\
&- \frac{2n_1 n_2}{N} \left(\frac{c \sigma_e^2}{c \sigma_e^2 + \sigma_v^2} \right) \left(\frac{\sigma_e^2}{\sigma_e^2 + \sigma_v^2} \right) [(1 - B_1) \bar{Y}_{s_1} - B_2 \bar{Y}_{s_2} - (1 - B_1 - B_2) \mu] [-B_1 \bar{Y}_{s_1} \\
&+ (1 - B_2) \bar{Y}_{s_2} - (1 - B_1 - B_2) \mu]
\end{aligned} \tag{3.32}$$

onde

$$s_{Y_1}^2 = \frac{1}{n_1} \sum_{i=1}^{n_1} (Y_i - \bar{Y}_{s_1})^2 \quad \text{e} \quad s_{Y_2}^2 = \frac{1}{n_2} \sum_{i=N_1+1}^{N_1+n_2} (Y_i - \bar{Y}_{s_2})^2.$$

Portanto, de (3.31) e (3.32) segue que o previsor Bayesiano de S^2 é

$$\begin{aligned}
\widehat{S}_B^2 &= \frac{1}{N} \left\{ \frac{\sigma_e^2 \sigma_v^2}{\sigma_e^2 + \sigma_v^2} \left[(n_1 - B_1) - \frac{n_1}{N} (1 - B_1 - B_2) \right] + \frac{c \sigma_e^2 \sigma_v^2}{c \sigma_e^2 + \sigma_v^2} \left[(n_2 - B_2) \right. \right. \\
&\quad \left. \left. - \frac{n_2}{N} (1 - B_1 - B_2) \right] + \left(\frac{N_1 - n_1}{N} \right) \sigma_e^2 [N - (1 - B_1 - c B_2)] \right. \\
&\quad \left. + \left(\frac{N - N_1 - n_2}{N} \right) c \sigma_e^2 [N - (1 - c^{-1} B_1 - B_2)] \right. \\
&\quad \left. + \left(\frac{\sigma_e^2}{\sigma_e^2 + \sigma_v^2} \right)^2 \left\{ n_1 s_{Y_1}^2 + n_1 \left(1 - \frac{n_1}{N} \right) [(1 - B_1) \bar{Y}_{s_1} - B_2 \bar{Y}_{s_2} - (1 - B_1 - B_2) \mu]^2 \right\} \right. \\
&\quad \left. + \left(\frac{c \sigma_e^2}{c \sigma_e^2 + \sigma_v^2} \right)^2 \left\{ n_2 s_{Y_2}^2 + n_2 \left(1 - \frac{n_2}{N} \right) [(1 - B_2) \bar{Y}_{s_2} - B_1 \bar{Y}_{s_1} - (1 - B_1 - B_2) \mu]^2 \right\} \right. \\
&\quad \left. - 2 \frac{n_1 n_2}{N} \left(\frac{c \sigma_e^2}{c \sigma_e^2 + \sigma_v^2} \right) \left(\frac{\sigma_e^2}{\sigma_e^2 + \sigma_v^2} \right) [(1 - B_1) \bar{Y}_{s_1} - B_2 \bar{Y}_{s_2} - (1 - B_1 - B_2) \mu] \right. \\
&\quad \left. [-B_1 \bar{Y}_{s_1} + (1 - B_2) \bar{Y}_{s_2} - (1 - B_1 - B_2) \mu] \right\}
\end{aligned} \tag{3.33}$$

3.5.2 PREVISÃO BAYESIANA EMPÍRICA

Como na seção 3.3, consideramos agora uma análise Bayesiana empírica para o modelo não homogêneo. O tamanho da população no estágio j é denotado por N_j e nesse estágio consideramos o modelo

$$\begin{aligned}
y_l^{(j)} &= \theta^{(j)} + e_l^{(j)} \\
Y_l^{(j)} &= y_l^{(j)} + v_l^{(j)}
\end{aligned} \tag{3.34}$$

onde $\theta^{(j)}$, $e_l^{(j)}$ e $v_l^{(j)}$ são todos independentes sendo $\theta^{(j)} \sim N(\mu, \sigma^2)$, $v_l^{(j)} \sim N(0, \sigma_v^2)$, $e_1^{(j)}, \dots, e_{N_{1j}}^{(j)} \sim N(0, \sigma_e^2)$ e $e_{N_{1j}+1}^{(j)}, \dots, e_{N_j}^{(j)} \sim N(0, c \sigma_e^2)$, $1 \leq N_{1j} < N_j$ e $0 < c < 1$.

Os previsores Bayesianos de $T^{(m)}$ e $S^{2(m)}$, total e variância populacionais no estágio m , seguem de (3.30) e (3.33), respectivamente, com B_1 , B_2 , $\bar{Y}_{s_1}$, $\bar{Y}_{s_2}$, $s_{Y_1}^2$ e $s_{Y_2}^2$ substituídos por

$$\begin{aligned} B_{1m} &= \frac{n_{1m}M_2}{n_{1m}M_2 + n_{2m}M_1 + M_1M_2}, \quad B_{2m} = \frac{n_{2m}M_1}{n_{1m}M_2 + n_{2m}M_1 + M_1M_2}, \\ \bar{Y}_{s_1}^{(m)} &= \frac{1}{n_{1m}} \sum_{l=1}^{n_{1m}} Y_l^{(m)}, \quad \bar{Y}_{s_2}^{(m)} = \frac{1}{n_{2m}} \sum_{l=N_{1m}+1}^{N_{1m}+n_{2m}} Y_l^{(m)}, \\ s_{Y_1}^{2(m)} &= \frac{1}{n_{1m}} \sum_{l=1}^{n_{1m}} (Y_l^{(m)} - \bar{Y}_{s_1}^{(m)})^2, \quad s_{Y_2}^{2(m)} = \frac{1}{n_{2m}} \sum_{l=N_{1m}+1}^{N_{1m}+n_{2m}} (Y_l^{(m)} - \bar{Y}_{s_2}^{(m)})^2. \end{aligned}$$

A amostra observada no estágio j é $\mathbf{Y}_s^{(j)} = (\mathbf{Y}_{s_1}^{(j)}, \mathbf{Y}_{s_2}^{(j)})'$ onde $\mathbf{Y}_{s_1}^{(j)} = (Y_1^{(j)}, \dots, Y_{n_{1j}}^{(j)})'$ e $\mathbf{Y}_{s_2}^{(j)} = (Y_{N_{1j}+1}^{(j)}, \dots, Y_{N_{1j}+n_{2j}}^{(j)})'$. Consideremos que M_1 , M_2 e μ são desconhecidos, mas a razão $\lambda = \sigma_e^2/\sigma_v^2$ e c são conhecidos. Temos que, $\mathbf{Y}_s^{(1)}, \dots, \mathbf{Y}_s^{(m)}$ são independentes e

$$\mathbf{Y}_s^{(j)} \sim N \left(\mu \mathbf{1}_{n_j}; \begin{pmatrix} (\sigma_e^2 + \sigma_v^2) \mathbf{I}_{n_{1j}} & 0 \\ 0 & (c\sigma_e^2 + \sigma_v^2) \mathbf{I}_{n_{2j}} \end{pmatrix} + \sigma^2 \mathbf{1}_{n_j} \mathbf{1}_{n_j}' \right),$$

onde $\mathbf{1}_{n_j}$ é um vetor de uns de dimensão n_j e $\mathbf{I}_{n_{ij}}$ é a matriz identidade de dimensão n_{ij} . Notemos que $n_j = n_{1j} + n_{2j}$, é o tamanho total da amostra no estágio j . Seja $n_{1T} = \sum_{j=1}^m n_{1j}$, $n_{2T} = \sum_{j=1}^m n_{2j}$, $n_T = n_{1T} + n_{2T}$ e $\bar{Y}_s = \sum_{j=1}^m n_j \bar{Y}_s^{(j)} / n_T$, onde $\bar{Y}_s^{(j)} = \sum_{l \in s_j} Y_l^{(j)} / n_j$ e s_j é a amostra completa selecionada no estágio j . Vamos supor que

- a) $\inf_{j \geq 1} n_{ij} \geq 2$,
- b) $\sup_{j \geq 1} n_{ij} = k < \infty$,
- c) $\sup_{j \geq 1} N_j < N_0 < \infty$,

$i = 1, 2$. Seja

$$\begin{aligned} \text{BMS} &= (m-1)^{-1} \sum_{j=1}^m n_j (\bar{Y}_s^{(j)} - \bar{Y}_s)^2, \\ \text{WMS}_1 &= (n_{1T} - m)^{-1} \sum_{j=1}^m \sum_{l=1}^{n_{1j}} (Y_l^{(j)} - \bar{Y}_{s_1}^{(j)})^2 \quad \text{e} \\ \text{WMS}_2 &= (n_{2T} - m)^{-1} \sum_{j=1}^m \sum_{l=N_{1j}+1}^{N_{1j}+n_{2j}} (Y_l^{(j)} - \bar{Y}_{s_2}^{(j)})^2. \end{aligned}$$

Notemos que BMS, WMS₁ e WMS₂ são independentes. O resultado a seguir apresenta a média e a variância das quantidades acima. A prova é similar à do Lema 3.1 e será omitida.

Lema 3.3. *Sob as suposições feitas acima*

$$\begin{aligned} E[\text{WMS}_1] &= \sigma_e^2 + \sigma_v^2, \quad \text{Var}[\text{WMS}_1] = \frac{2(\sigma_e^2 + \sigma_v^2)^2}{n_{1T} - m} \\ E[\text{WMS}_2] &= c\sigma_e^2 + \sigma_v^2, \quad \text{Var}[\text{WMS}_2] = \frac{2(c\sigma_e^2 + \sigma_v^2)^2}{n_{2T} - m} \\ E[\text{BMS}] &= (m-1)^{-1} \left[\sigma^2 g + (\sigma_e^2 + \sigma_v^2) \sum_{j=1}^m \left(1 - \frac{n_j}{n_T}\right) \frac{n_{1j}}{n_j} + (c\sigma_e^2 + \sigma_v^2) \sum_{j=1}^m \left(1 - \frac{n_j}{n_T}\right) \frac{n_{2j}}{n_j} \right] \\ \text{Var}[\text{BMS}] &= 2(m-1)^{-2} \left[\sum_{j=1}^m \sigma_j^4 + \frac{\left(\sum_{j=1}^m n_j \sigma_j^2\right)^2}{n_T^2} - 2 \frac{\sum_{j=1}^m n_j \sigma_j^4}{n_T} \right] \end{aligned}$$

$$\text{onde } \sigma_j^2 = \frac{n_{1j}(\sigma_e^2 + \sigma_v^2) + n_{2j}(c\sigma_e^2 + \sigma_v^2)}{n_j} + n_j \sigma^2 \text{ e } g = n_T - \sum_{j=1}^m n_j^2 / n_T.$$

Do Lema 3.3 e usando as suposições a), b) e c) temos que WMS₁ e WMS₂ são, respectivamente, estimadores consistentes de $\sigma_e^2 + \sigma_v^2$ e $c\sigma_e^2 + \sigma_v^2$. Além disso, temos que

$$g^{-1} \left[(m-1) \text{BMS} - \text{WMS}_1 \sum_{j=1}^m \left(1 - \frac{n_j}{n_T}\right) \frac{n_{1j}}{n_j} - \text{WMS}_2 \sum_{j=1}^m \left(1 - \frac{n_j}{n_T}\right) \frac{n_{2j}}{n_j} \right]$$

é um estimador não viciado e consistente de σ^2 . Portanto, propomos os seguintes estimadores consistentes para M_1^{-1} e M_2^{-1} .

$$\widehat{M}_1^{-1} = \max \left\{ 0; g^{-1} \left[(m-1) \frac{\text{BMS}}{\text{WMS}_1} - \frac{\text{WMS}_2}{\text{WMS}_1} \sum_{j=1}^m \left(1 - \frac{n_j}{n_T}\right) \frac{n_{2j}}{n_T} - \sum_{j=1}^m \left(1 - \frac{n_j}{n_T}\right) \frac{n_{1j}}{n_T} \right] \right\}$$

e

$$\widehat{M}_2^{-1} = \max \left\{ 0; g^{-1} \left[(m-1) \frac{\text{BMS}}{\text{WMS}_2} - \frac{\text{WMS}_1}{\text{WMS}_2} \sum_{j=1}^m \left(1 - \frac{n_j}{n_T}\right) \frac{n_{1j}}{n_T} - \sum_{j=1}^m \left(1 - \frac{n_j}{n_T}\right) \frac{n_{2j}}{n_T} \right] \right\}.$$

Então, temos que B_{1j} e B_{2j} são estimados por

$$\widehat{B}_{1j} = \frac{n_{1j} \widehat{M}_1^{-1}}{n_{1j} \widehat{M}_1^{-1} + n_{2j} \widehat{M}_2^{-1} + 1} \quad \text{e} \quad \widehat{B}_{2j} = \frac{n_{2j} \widehat{M}_2^{-1}}{n_{2j} \widehat{M}_2^{-1} + n_{1j} \widehat{M}_1^{-1} + 1}$$

respectivamente. Provamos o Lema 3.4 que será útil mais tarde.

Lema 3.4. *Sob as suposições a), b) e c) apresentadas acima*

$$\max_j |\hat{B}_{1j} - B_{1j}| \xrightarrow{P} 0 \quad (3.35)$$

$$\max_j |\hat{B}_{2j} - B_{2j}| \xrightarrow{P} 0 \quad (3.36)$$

quando $m \rightarrow \infty$.

Prova. Podemos escrever

$$\begin{aligned} |\hat{B}_{1j} - B_{1j}| &= \left| \frac{n_{1j}\hat{M}_1^{-1}(1 + n_{1j}M_1^{-1} + n_{2j}M_2^{-1}) - n_{1j}M_1^{-1}(1 + n_{1j}\hat{M}_1^{-1} + n_{2j}\hat{M}_2^{-1})}{(n_{1j}\hat{M}_1^{-1} + n_{2j}\hat{M}_2^{-1} + 1)(n_{1j}M_1^{-1} + n_{2j}M_2^{-1} + 1)} \right| \\ &\leq |n_{1j}(\hat{M}_1^{-1} - M_1^{-1}) + n_{1j}n_{2j}(\hat{M}_1^{-1}M_2^{-1} - M_1^{-1}\hat{M}_2^{-1})| \\ &\leq k|\hat{M}_1^{-1} - M_1^{-1}| + k^2|\hat{M}_1^{-1}M_2^{-1} - M_1^{-1}\hat{M}_2^{-1}|. \end{aligned}$$

Então

$$\max_j |\hat{B}_{1j} - B_{1j}| \leq k|\hat{M}_1^{-1} - M_1^{-1}| + k^2|\hat{M}_1^{-1}M_2^{-1} - M_1^{-1}\hat{M}_2^{-1}| \xrightarrow{P} 0$$

pois $\hat{M}_1^{-1} \xrightarrow{P} M_1^{-1}$ e $\hat{M}_2^{-1} \xrightarrow{P} M_2^{-1}$, quando $m \rightarrow \infty$.

A prova de (3.36) é análoga a de (3.35).

Como μ é desconhecido, precisamos estimá-lo a partir dos dados. Temos que a densidade conjunta de $(\mathbf{Y}_s^{(1)}, \dots, \mathbf{Y}_s^{(m)})$ é proporcional a

$$\begin{aligned} h(\mathbf{Y}_s^{(1)}, \dots, \mathbf{Y}_s^{(m)}) &\propto \exp \left[-\frac{1}{2} \sum_{j=1}^m (\mathbf{Y}_s^{(j)} - \mu \mathbf{1}_{n_j})' \left(\begin{pmatrix} (\sigma_e^2 + \sigma_v^2)\mathbf{I}_{n_{1j}} & 0 \\ 0 & (c\sigma_e^2 + \sigma_v^2)\mathbf{I}_{n_{2j}} \end{pmatrix} \right. \right. \\ &\quad \left. \left. + \sigma^2 \mathbf{1}_{n_j} \mathbf{1}_{n_j}' \right)^{-1} (\mathbf{Y}_s^{(j)} - \mu \mathbf{1}_{n_j}) \right] \end{aligned}$$

e depois de algumas manipulações algébricas segue que

$$\begin{aligned}
h(\mathbf{Y}_s^{(1)}, \dots, \mathbf{Y}_s^{(m)}) &\propto \exp \left[-\frac{1}{2} \sum_{j=1}^m \sum_{l=1}^{n_{1j}} \frac{(Y_l^{(j)} - \bar{Y}_{s_1}^{(j)})^2}{\sigma_e^2 + \sigma_v^2} - \frac{1}{2} \sum_{j=1}^m \frac{n_{1j}(\bar{Y}_{s_1}^{(j)} - \mu)^2}{\sigma_e^2 + \sigma_v^2} \right. \\
&\quad - \frac{1}{2} \sum_{j=1}^m \sum_{l=N_{1j}+1}^{N_{1j}+n_{2j}} \frac{(Y_l^{(j)} - \bar{Y}_{s_2}^{(j)})^2}{c\sigma_e^2 + \sigma_v^2} - \frac{1}{2} \sum_{j=1}^m \frac{n_{2j}(\bar{Y}_{s_2}^{(j)} - \mu)^2}{c\sigma_e^2 + \sigma_v^2} \\
&\quad + \frac{\sigma^2(c\sigma_e^2 + \sigma_v^2)}{2(\sigma_e^2 + \sigma_v^2)} \sum_{j=1}^m \frac{n_{1j}^2(\bar{Y}_{s_1}^{(j)} - \mu)^2}{\bar{\sigma}_j^2} + \frac{\sigma^2(\sigma_e^2 + \sigma_v^2)}{2(c\sigma_e^2 + \sigma_v^2)} \sum_{j=1}^m \frac{n_{2j}^2(\bar{Y}_{s_2}^{(j)} - \mu)^2}{\bar{\sigma}_j^2} \\
&\quad \left. + \sigma^2 \sum_{j=1}^m \frac{n_{1j}n_{2j}(\bar{Y}_{s_1}^{(j)} - \mu)(\bar{Y}_{s_2}^{(j)} - \mu)}{\bar{\sigma}_j^2} \right]
\end{aligned}$$

onde $\bar{\sigma}_j^2 = (c\sigma_e^2 + \sigma_v^2)(\sigma_e^2 + \sigma_v^2 + n_{1j}\sigma^2) + n_{2j}\sigma^2(\sigma_e^2 + \sigma_v^2)$.

Então, sendo M_1 e M_2 conhecidos, segue que o estimador de máxima verossimilhança de μ é dado por

$$\hat{\mu}^* = \frac{\sum_{j=1}^m (B_{1j} \bar{Y}_{s_1}^{(j)} + B_{2j} \bar{Y}_{s_2}^{(j)})}{\sum_{j=1}^m (B_{1j} + B_{2j})}.$$

Notemos que se $\widehat{M}_1^{-1} = 0$ então $\widehat{B}_{1j} = 0$ e se $\widehat{M}_2^{-1} = 0$ então $\widehat{B}_{2j} = 0$, $j = 1, \dots, m$. Portanto, quando M_1 e M_2 são desconhecidos, motivados por $\hat{\mu}^*$, propomos o seguinte estimador para μ .

$$\hat{\mu} = \begin{cases} \frac{\sum_{j=1}^m (\widehat{B}_{1j} \bar{Y}_{s_1}^{(j)} + \widehat{B}_{2j} \bar{Y}_{s_2}^{(j)})}{\sum_{j=1}^m (\widehat{B}_{1j} + \widehat{B}_{2j})} & , \text{ se } \widehat{M}_1^{-1} \neq 0 \text{ e } \widehat{M}_2^{-1} \neq 0 \\ m^{-1} \sum_{j=1}^m \left(\frac{n_{1j}}{n_{1j} + c^{-1}n_{2j}} \bar{Y}_{s_1}^{(j)} + \frac{c^{-1}n_{2j}}{n_{1j} + c^{-1}n_{2j}} \bar{Y}_{s_2}^{(j)} \right) & , \text{ se } \widehat{M}_1^{-1} = 0 \text{ ou } \widehat{M}_2^{-1} = 0 \end{cases}.$$

Concluindo, temos que o previsor Bayesiano empírico de $T^{(m)}$ é dado por

$$\begin{aligned}
\hat{T}_{EB} &= N_m [\widehat{B}_{1m} \bar{Y}_{s_1}^{(m)} + \widehat{B}_{2m} \bar{Y}_{s_2}^{(m)} + (1 - \widehat{B}_{1m} - \widehat{B}_{2m}) \hat{\mu}] \\
&\quad + \frac{n_{1m}\lambda}{\lambda + 1} [(1 - \widehat{B}_{1m}) \bar{Y}_{s_1}^{(m)} - \widehat{B}_{2m} \bar{Y}_{s_2}^{(m)} - (1 - \widehat{B}_{1m} - \widehat{B}_{2m}) \hat{\mu}] \\
&\quad + \frac{n_{2m}c\lambda}{c\lambda + 1} [(1 - \widehat{B}_{2m}) \bar{Y}_{s_2}^{(m)} - \widehat{B}_{1m} \bar{Y}_{s_1}^{(m)} - (1 - \widehat{B}_{1m} - \widehat{B}_{2m}) \hat{\mu}].
\end{aligned}$$

Da mesma forma, com c conhecido, temos que o previsor Bayesiano empírico de $S^{(m)}$

é dado por

$$\begin{aligned}
\hat{S}_{EB}^2 &= \frac{1}{N_m} \left\{ \frac{\lambda \text{WMS}_1}{(\lambda + 1)^2} \left[(n_{1m} - \hat{B}_{1m}) - \frac{n_{1m}}{N_m} (1 - \hat{B}_{1m} - \hat{B}_{2m}) \right] \right. \\
&\quad + \frac{c\lambda \text{WMS}_2}{(c\lambda + 1)^2} \left[(n_{2m} - \hat{B}_{2m}) - \frac{n_{2m}}{N_m} (1 - \hat{B}_{1m} - \hat{B}_{2m}) \right] \\
&\quad + \left(\frac{N_{1m} - n_{1m}}{N_m} \right) \frac{\lambda \text{WMS}_1}{\lambda + 1} \left[N_m - (1 - \hat{B}_{1m} - c\hat{B}_{2m}) \right] \\
&\quad + \left(\frac{N_m - N_{1m} - n_{2m}}{N_m} \right) \frac{c\lambda \text{WMS}_2}{c\lambda + 1} \left[N_m - (1 - c^{-1}\hat{B}_{1m} - \hat{B}_{2m}) \right] \\
&\quad + \left(\frac{\lambda}{\lambda + 1} \right)^2 \left\{ n_{1m} s_{Y_1}^{2(m)} + n_{1m} \left(1 - \frac{n_{1m}}{N_m} \right) \left[(1 - \hat{B}_{1m}) \bar{Y}_{s_1}^{(m)} - \hat{B}_{2m} \bar{Y}_{s_2}^{(m)} \right. \right. \\
&\quad \left. \left. - (1 - \hat{B}_{1m} - \hat{B}_{2m}) \hat{\mu} \right]^2 \right\} + \left(\frac{c\lambda}{c\lambda + 1} \right)^2 \left\{ n_{2m} s_{Y_2}^{2(m)} + n_{2m} \left(1 - \frac{n_{2m}}{N_m} \right) \left[-\hat{B}_{1m} \bar{Y}_{s_1}^{(m)} \right. \right. \\
&\quad \left. \left. + (1 - \hat{B}_{2m}) \bar{Y}_{s_2}^{(m)} - (1 - \hat{B}_{1m} - \hat{B}_{2m}) \hat{\mu} \right]^2 \right\} \\
&\quad - \frac{2n_{1m}n_{2m}}{N_m} \left(\frac{\lambda}{\lambda + 1} \right) \left(\frac{c\lambda}{c\lambda + 1} \right) \left[(1 - \hat{B}_{1m}) \bar{Y}_{s_1}^{(m)} - \hat{B}_{2m} \bar{Y}_{s_2}^{(m)} \right. \\
&\quad \left. - (1 - \hat{B}_{1m} - \hat{B}_{2m}) \hat{\mu} \right] \left[(1 - \hat{B}_{2m}) \bar{Y}_{s_2}^{(m)} - \hat{B}_{1m} \bar{Y}_{s_1}^{(m)} - (1 - \hat{B}_{1m} - \hat{B}_{2m}) \hat{\mu} \right] \Big\} .
\end{aligned}$$

3.5.3 RISCO DE BAYES DE $\hat{T}_{EB}$ E $\hat{S}_{EB}^2$

Como na seção 3.4, denotamos por $r(\xi, e)$ o risco de Bayes, sob perda quadrática, do previsor e em relação ao modelo ξ , dado em 3.5.2. No Teorema 3.4 mostramos que o previsor Bayesiano empírico $\hat{T}_{EB}$ é assintoticamente ótimo no sentido de Robbins (1955).

Teorema 3.4. *Sob as suposições a), b) e c)*

$$r(\xi, \hat{T}_{EB}) - r(\xi, \hat{T}_B) \rightarrow 0$$

quando $m \rightarrow \infty$.

Prova. Utilizando-se o Lema 3 de Ghosh e Meeden (1986), temos que

$$\begin{aligned} r(\xi, \hat{T}_{EB}) - r(\xi, \hat{T}_B) &= E[\hat{T}_{EB} - \hat{T}_B]^2 \\ &= E\left\{ \left(N_m - \frac{n_{1m}\lambda}{\lambda+1} - \frac{n_{2m}c\lambda}{c\lambda+1} \right) [(\hat{B}_{1m} - B_{1m})(\bar{Y}_{s_1}^{(m)} - \mu) \right. \\ &\quad \left. + (\hat{B}_{2m} - B_{2m})(\bar{Y}_{s_2}^{(m)} - \mu) + (1 - \hat{B}_{1m} - \hat{B}_{2m})(\hat{\mu} - \mu)] \right\}^2. \end{aligned}$$

Para provar o teorema, temos que mostrar que essa última esperança tende a zero quando $m \rightarrow \infty$. Da mesma forma que no Teorema 3.2 mostramos que

$$|\hat{B}_{1m} - B_{1m}| |\bar{Y}_{s_1}^{(m)} - \mu| = o_p(1)$$

e

$$|\hat{B}_{2m} - B_{2m}| |\bar{Y}_{s_2}^{(m)} - \mu| = o_p(1)$$

então $(\hat{B}_{1m} - B_{1m})(\bar{Y}_{s_1}^{(m)} - \mu) \xrightarrow{P} 0$ e $(\hat{B}_{2m} - B_{2m})(\bar{Y}_{s_2}^{(m)} - \mu) \xrightarrow{P} 0$ quando $m \rightarrow \infty$.

Além disso, para todo $\varepsilon > 0$,

$$\begin{aligned} P(|\hat{\mu} - \mu| > \varepsilon) &\leq P(\hat{M}_1^{-1} = 0 \text{ ou } \hat{M}_2^{-1} = 0) + P(|\hat{\mu} - \mu| > \varepsilon, \hat{M}_1^{-1} \neq 0; \hat{M}_2^{-1} \neq 0) \\ &\leq P(\hat{M}_1^{-1} = 0) + P(\hat{M}_2^{-1} = 0) + P(|\hat{\mu} - \mu| > \varepsilon, \hat{M}_1^{-1} \neq 0; \hat{M}_2^{-1} \neq 0) \end{aligned}$$

mas $P(\hat{M}_1^{-1} = 0) \rightarrow 0$ e $P(\hat{M}_2^{-1} = 0) \rightarrow 0$ pois $\hat{M}_1^{-1} \xrightarrow{P} M_1^{-1}$ e $\hat{M}_2^{-1} \xrightarrow{P} M_2^{-1}$ sendo $M_1^{-1} > 0$ e $M_2^{-1} > 0$.

Para $\hat{M}_1^{-1} \neq 0$ e $\hat{M}_2^{-1} \neq 0$ temos que

$$\hat{\mu} - \mu = \frac{\sum_{j=1}^m [\hat{B}_{1j}(\bar{Y}_{s_1}^{(j)} - \mu) + \hat{B}_{2j}(\bar{Y}_{s_2}^{(j)} - \mu)]}{\sum_{j=1}^m (\hat{B}_{1j} + \hat{B}_{2j})}$$

e

$$\sum_{j=1}^m (\hat{B}_{1j} + \hat{B}_{2j}) \geq \frac{2m(\widehat{M}_1 + \widehat{M}_2)}{2(\widehat{M}_1 + \widehat{M}_2) + \widehat{M}_1 \widehat{M}_2}$$

Temos então

$$\begin{aligned} |\hat{\mu} - \mu| &\leq \frac{2(\widehat{M}_1 + \widehat{M}_2) + \widehat{M}_1 \widehat{M}_2}{2m(\widehat{M}_1 + \widehat{M}_2)} \left| \sum_{j=1}^m [\hat{B}_{1j}(\bar{Y}_{s_1}^{(j)} - \mu) + \hat{B}_{2j}(\bar{Y}_{s_2}^{(j)} - \mu)] \right| \\ &\leq \frac{2(\widehat{M}_1 + \widehat{M}_2) + \widehat{M}_1 \widehat{M}_2}{2m(\widehat{M}_1 + \widehat{M}_2)} \left[\sum_{j=1}^m |\hat{B}_{1j} - B_{1j}| |\bar{Y}_{s_1}^{(j)} - \mu| + \left| \sum_{j=1}^m B_{1j}(\bar{Y}_{s_1}^{(j)} - \mu) \right| \right. \\ &\quad \left. + \sum_{j=1}^m |\hat{B}_{2j} - B_{2j}| |\bar{Y}_{s_2}^{(j)} - \mu| + \left| \sum_{j=1}^m B_{2j}(\bar{Y}_{s_2}^{(j)} - \mu) \right| \right] \\ &\leq \frac{2(\widehat{M}_1 + \widehat{M}_2) + \widehat{M}_1 \widehat{M}_2}{2(\widehat{M}_1 + \widehat{M}_2)} \left[\max_j |\hat{B}_{1j} - B_{1j}| \frac{1}{m} \sum_{j=1}^m |\bar{Y}_{s_1}^{(j)} - \mu| + \frac{1}{m} \left| \sum_{j=1}^m B_{1j}(\bar{Y}_{s_1}^{(j)} - \mu) \right| \right. \\ &\quad \left. + \max_j |\hat{B}_{2j} - B_{2j}| \frac{1}{m} \sum_{j=1}^m |\bar{Y}_{s_2}^{(j)} - \mu| + \frac{1}{m} \left| \sum_{j=1}^m B_{2j}(\bar{Y}_{s_2}^{(j)} - \mu) \right| \right]. \end{aligned} \quad (3.37)$$

Mas

$$\frac{1}{m} \sum_{j=1}^m |\bar{Y}_{s_1}^{(j)} - \mu| \stackrel{d}{=} \frac{1}{m} \sum_{j=1}^m [\sigma^2 + (\sigma_e^2 + \sigma_v^2) n_{1j}^{-1}] U_j \stackrel{\text{ST}}{\leq} (\sigma^2 + \sigma_e^2 + \sigma_v^2) \frac{1}{m} \sum_{j=1}^m U_j = O_p(1) \quad (3.38)$$

onde U_j iid χ_1 .

Analogamente,

$$\frac{1}{m} \sum_{j=1}^m |\bar{Y}_{s_2}^{(j)} - \mu| = O_p(1). \quad (3.39)$$

Além disso,

$$\begin{aligned} E \left[m^{-1} \left| \sum_{j=1}^m B_{1j}(\bar{Y}_{s_1}^{(j)} - \mu) \right|^2 \right] &= m^{-2} \sum_{j=1}^m B_{1j}^2 E[\bar{Y}_{s_1}^{(j)} - \mu]^2 \\ &= m^{-2} \sum_{j=1}^m B_{1j}^2 [\sigma^2 + (\sigma_e^2 + \sigma_v^2) n_{1j}^{-1}] \leq m^{-2} (\sigma^2 + \sigma_e^2 + \sigma_v^2) \sum_{j=1}^m B_{1j}^2 \\ &\leq m^{-1} (\sigma^2 + \sigma_e^2 + \sigma_v^2) = O(m^{-1}), \end{aligned}$$

portanto

$$m^{-1} \left| \sum_{j=1}^m B_{1j}(\bar{Y}_{s_1}^{(j)} - \mu) \right| = o_p(1) \quad (3.40)$$

e de maneira análoga,

$$m^{-1} \left| \sum_{j=1}^m B_{1j}(\bar{Y}_{s_1}^{(j)} - \mu) \right| = o_p(1) . \quad (3.41)$$

Utilizando-se (3.38) – (3.41) e o Lema 3.4 na expressão (3.37) segue que

$$P(|\hat{\mu} - \mu| > \varepsilon) \rightarrow 0$$

quando $m \rightarrow \infty$. Isso conclui a primeira parte da prova, isto é,

$$\begin{aligned} & \left(N_m - \frac{n_{1m}\lambda}{\lambda+1} - \frac{n_{2m}c\lambda}{c\lambda+1} \right) [(\hat{B}_{1m} - B_{1m})(\bar{Y}_{s_1}^{(m)} - \mu) + (\hat{B}_{2m} - B_{2m})(\bar{Y}_{s_2}^{(m)} - \mu) \\ & + (1 - \hat{B}_{1m} - \hat{B}_{2m})(\hat{\mu} - \mu)] \xrightarrow{P} 0 \end{aligned}$$

quando $m \rightarrow \infty$. Falta mostrarmos que

$$[(\hat{B}_{1m} - B_{1m})(\bar{Y}_{s_1}^{(m)} - \mu) + (\hat{B}_{2m} - B_{2m})(\bar{Y}_{s_2}^{(m)} - \mu) + (1 - \hat{B}_{1m} - \hat{B}_{2m})(\hat{\mu} - \mu)]^2$$

é uniformemente integrável em m . Mas

$$\begin{aligned} & [(\hat{B}_{1m} - B_{1m})(\bar{Y}_{s_1}^{(m)} - \mu) + (\hat{B}_{2m} - B_{2m})(\bar{Y}_{s_2}^{(m)} - \mu) + (1 - \hat{B}_{1m} - \hat{B}_{2m})(\hat{\mu} - \mu)]^2 \\ & \leq 3[(\hat{B}_{1m} - B_{1m})^2(\bar{Y}_{s_1}^{(m)} - \mu)^2 + (\hat{B}_{2m} - B_{2m})^2(\bar{Y}_{s_2}^{(m)} - \mu)^2 + (1 - \hat{B}_{1m} - \hat{B}_{2m})^2(\hat{\mu} - \mu)^2] \\ & \leq 3[(\bar{Y}_{s_1}^{(m)} - \mu)^2 + (\bar{Y}_{s_2}^{(m)} - \mu)^2 + (\hat{\mu} - \mu)^2] . \end{aligned}$$

Como

$$(\bar{Y}_{s_1}^{(m)} - \mu)^2 \stackrel{d}{=} [\sigma^2 + (\sigma_e^2 + \sigma_v^2)n_{1m}^{-1}]\chi_1^2 \stackrel{\text{ST}}{\leq} (\sigma^2 + \sigma_e^2 + \sigma_v^2)\chi_1^2$$

segue que $(\bar{Y}_{s_1}^{(m)} - \mu)^2$ é uniformemente integrável em m . De maneira análoga, mostramos que $(\bar{Y}_{s_2}^{(m)} - \mu)^2$ também é uniformemente integrável em m .

Para concluirmos a prova falta mostrarmos que $(\hat{\mu} - \mu)^2$ é uniformemente integrável em m . Temos que

$$(\hat{\mu} - \mu)^2 = \left\{ m^{-1} \sum_{j=1}^m \left[\frac{n_{1j}}{n_{1j} + n_{2j}} (\bar{Y}_{s_1}^{(j)} - \mu) + \frac{c^{-1}n_{2j}}{n_{1j} + c^{-1}n_{2j}} (\bar{Y}_{s_2}^{(j)} - \mu) \right] \right\}^2 I_{\{\hat{M}_1^{-1}=0 \text{ ou } \hat{M}_2^{-1}=0\}}$$

$$+ \left\{ \frac{\left[\sum_{j=1}^m [\hat{B}_{1j}(\bar{Y}_{s_1}^{(j)} - \mu) + \hat{B}_{2j}(\bar{Y}_{s_2}^{(j)} - \mu)] \right]^2}{\left[\sum_{j=1}^m (\hat{B}_{1j} + \hat{B}_{2j}) \right]^2} \right\} I_{\{\hat{M}_1^{-1} > 0 \text{ e } \hat{M}_2^{-1} > 0\}}. \quad (3.42)$$

Denotamos por A o segundo termo do lado direito de (3.42) e B o primeiro. Então

$$\begin{aligned} A &\leq \frac{\left[\sum_{j=1}^m \frac{k\hat{M}_2(\bar{Y}_{s_1}^{(j)} - \mu) + k\hat{M}_1(\bar{Y}_{s_2}^{(j)} - \mu)}{2(\hat{M}_1 + \hat{M}_2) + \hat{M}_1\hat{M}_2} \right]^2}{\left[\frac{2(\hat{M}_1 + \hat{M}_2)m}{2(\hat{M}_1 + \hat{M}_2) + \hat{M}_1\hat{M}_2} \right]^2} \\ &= \frac{m^{-2}k^2}{4(\hat{M}_1 + \hat{M}_2)^2} \left\{ \sum_{j=1}^m [\hat{M}_2(\bar{Y}_{s_1}^{(j)} - \mu)] + \sum_{j=1}^m \hat{M}_1(\bar{Y}_{s_2}^{(j)} - \mu) \right\}^2 \\ &= \frac{k^2}{4} \left\{ \frac{\hat{M}_2}{\hat{M}_1 + \hat{M}_2} \sum_{j=1}^m \frac{(\bar{Y}_{s_1}^{(j)} - \mu)}{m} + \frac{\hat{M}_1}{\hat{M}_1 + \hat{M}_2} \sum_{j=1}^m \frac{(\bar{Y}_{s_2}^{(j)} - \mu)}{m} \right\}^2 \\ &\leq \frac{k^2}{4} \left[\frac{1}{m} \sum_{j=1}^m (\bar{Y}_{s_1}^{(j)} - \mu + \bar{Y}_{s_2}^{(j)} - \mu) \right]^2 \\ &\leq \frac{k^2}{2} \left[m^{-1} \sum_{j=1}^m (\bar{Y}_{s_1}^{(j)} - \mu)^2 + m^{-1} \sum_{j=1}^m (\bar{Y}_{s_2}^{(j)} - \mu)^2 \right] \\ &\stackrel{\text{ST}}{\leq} \frac{k^2}{2} \left[(\sigma^2 + \sigma_e^2 + \sigma_v^2)m^{-1} \sum_{j=1}^m U_j + (\sigma^2 + c\sigma_e^2 + \sigma_v^2)m^{-1} \sum_{j=1}^m U_j \right] \end{aligned}$$

onde $U_j \sim \chi_1^2$, portanto A é uniformemente integrável em m . Além disso, se $\hat{M}_1^{-1} = 0$ ou $\hat{M}_2^{-1} = 0$, então

$$\begin{aligned} B &\leq m^{-1} \sum_{j=1}^m \left[\frac{n_{1j}}{n_{1j} + n_{2j}} (\bar{Y}_{s_1}^{(j)} - \mu) + \frac{c^{-1}n_{2j}}{n_{1j} + c^{-1}n_{2j}} (\bar{Y}_{s_2}^{(j)} - \mu) \right]^2 \\ &\leq 2m^{-1} \sum_{j=1}^m [(\bar{Y}_{s_1}^{(j)} - \mu)^2 + (\bar{Y}_{s_2}^{(j)} - \mu)^2] \\ &\stackrel{\text{ST}}{\leq} 2 \left[(\sigma^2 + \sigma_e^2 + \sigma_v^2)m^{-1} \sum_{j=1}^m U_j + (\sigma^2 + c\sigma_e^2 + \sigma_v^2)m^{-1} \sum_{j=1}^m U_j \right] \end{aligned}$$

que também é uniformemente integrável em m . E isso conclui a prova.

No próximo teorema mostraremos a otimalidade assintótica do previsor $\hat{S}_{EB}^2$.

Teorema 3.5. *Sob as suposições a), b) e c)*

$$r(\xi, \hat{S}_{EB}^2) - r(\xi, \hat{S}_B^2) \rightarrow 0$$

quando $m \rightarrow \infty$.

Prova. Utilizando o Lema 3 de Ghosh e Meeden (1986) e depois de algumas manipulações algébricas segue que

$$\begin{aligned} r(\xi, \hat{S}_{EB}^2) - r(\xi, \hat{S}_B^2) &= E[\hat{S}_{EB}^2 - \hat{S}_B^2]^2 \\ &= \frac{1}{N_m^2} E \left\{ \frac{1}{\lambda + 1} \left[(n_{1m} - \hat{B}_{1m}) - \frac{n_{1m}}{N_m} (1 - \hat{B}_{1m} - \hat{B}_{2m}) \right] \left[\frac{\lambda}{\lambda + 1} \text{WMS}_1 - \sigma_e^2 \right] \right. \\ &\quad + \frac{1}{c\lambda + 1} \left[(n_{2m} - \hat{B}_{2m}) - \frac{n_{2m}}{N_m} (1 - \hat{B}_{1m} - \hat{B}_{2m}) \right] \left[\frac{c\lambda}{c\lambda + 1} \text{WMS}_2 - c\sigma_e^2 \right] \\ &\quad + \frac{\sigma_e^2}{\lambda + 1} \left[\left(1 - \frac{n_{1m}}{N_m} \right) (B_{1m} - \hat{B}_{1m}) + \frac{n_{1m}}{N_m} (\hat{B}_{2m} - B_{2m}) \right] \\ &\quad + \frac{c\sigma_e^2}{c\lambda + 1} \left[\left(1 - \frac{n_{2m}}{N_m} \right) (B_{2m} - \hat{B}_{2m}) + \frac{n_{2m}}{N_m} (\hat{B}_{1m} - B_{1m}) \right] \\ &\quad + \left(\frac{N_{1m} - n_{1m}}{N_m} \right) \left\{ [N_m - (1 - \hat{B}_{1m} - c\hat{B}_{2m})] \left[\frac{\lambda}{\lambda + 1} \text{WMS}_1 - \sigma_e^2 \right] \right. \\ &\quad \left. + \sigma_e^2 (\hat{B}_{1m} - B_{1m}) + c\sigma_e^2 (\hat{B}_{2m} - B_{2m}) \right\} \\ &\quad + \left(\frac{N_m - N_{1m} - n_{2m}}{N_m} \right) \left\{ \left[N_m - \left(1 - \frac{\hat{B}_{1m}}{c} - \hat{B}_{2m} \right) \right] \left[\frac{c\lambda}{c\lambda + 1} \text{WMS}_2 - c\sigma_e^2 \right] \right. \\ &\quad \left. + \sigma_e^2 (\hat{B}_{1m} - B_{1m}) + c\sigma_e^2 (\hat{B}_{2m} - B_{2m}) \right\} \\ &\quad + \left[\left(\frac{\lambda}{\lambda + 1} \right)^2 n_{1m} \left(1 - \frac{n_{1m}}{N_m} \right) \right] \left\{ [(\hat{B}_{1m}^2 - B_{1m}^2) - 2(\hat{B}_{1m} - B_{1m})] (\bar{Y}_{s_1}^{(m)} - \mu)^2 \right. \\ &\quad + (\hat{B}_{2m}^2 - B_{2m}^2) (\bar{Y}_{s_2}^{(m)} - \mu)^2 + 2[\hat{B}_{2m}(\hat{B}_{1m} - B_{1m}) - (1 - B_{1m})(\hat{B}_{2m} - B_{2m})] \\ &\quad + (\bar{Y}_{s_1}^{(m)} - \mu)(\bar{Y}_{s_2}^{(m)} - \mu) - 2(1 - \hat{B}_{1m})(1 - \hat{B}_{1m} - \hat{B}_{2m})(\bar{Y}_{s_1}^{(m)} - \mu)(\hat{\mu} - \mu) \\ &\quad \left. + 2\hat{B}_{2m}(1 - \hat{B}_{1m} - \hat{B}_{2m})(\bar{Y}_{s_2}^{(m)} - \mu)(\hat{\mu} - \mu) + (1 - \hat{B}_{1m} - \hat{B}_{2m})^2 (\hat{\mu} - \mu)^2 \right\} \end{aligned}$$

$$\begin{aligned}
& + \left[\left(\frac{c\lambda}{c\lambda + 1} \right)^2 n_{2m} \left(1 - \frac{n_{2m}}{N_m} \right) \right] \left\{ (\hat{B}_{1m}^2 - B_{1m}^2)(\bar{Y}_{s_1}^{(m)} - \mu)^2 + [(\hat{B}_{2m}^2 - B_{2m}^2) \right. \\
& - 2(\hat{B}_{2m} - B_{2m})](\bar{Y}_{s_2}^{(m)} - \mu)^2 + 2[\hat{B}_{1m}(\hat{B}_{2m} - B_{2m}) - (1 - B_{2m})(\hat{B}_{1m} - B_{1m})] \\
& (\bar{Y}_{s_1}^{(m)} - \mu)(\bar{Y}_{s_2}^{(m)} - \mu) + 2\hat{B}_{1m}(1 - \hat{B}_{1m} - \hat{B}_{2m})(\bar{Y}_{s_1}^{(m)} - \mu)(\hat{\mu} - \mu) \\
& - 2(1 - \hat{B}_{2m})(1 - \hat{B}_{1m} - \hat{B}_{2m})(\bar{Y}_{s_2}^{(m)} - \mu)(\hat{\mu} - \mu) + (1 - \hat{B}_{1m} - \hat{B}_{2m})^2(\hat{\mu} - \mu)^2 \} \\
& + \frac{2n_{1m}n_{2m}}{N_m} \left(\frac{\lambda}{\lambda + 1} \right) \left(\frac{c\lambda}{c\lambda + 1} \right) \{ [(\hat{B}_{1m}^2 - B_{1m}^2) + (\hat{B}_{1m} - B_{1m})](\bar{Y}_{s_1}^{(m)} - \mu)^2 \\
& + [(\hat{B}_{2m}^2 - B_{2m}^2) + (\hat{B}_{2m} - B_{2m})](\bar{Y}_{s_2}^{(m)} - \mu)^2 + [(\hat{B}_{1m} - B_{1m}) + (\hat{B}_{2m} - B_{2m}) \\
& - 2B_{2m}(\hat{B}_{1m} - B_{1m}) - 2\hat{B}_{1m}(\hat{B}_{2m} - B_{2m})](\bar{Y}_{s_1}^{(m)} - \mu)(\bar{Y}_{s_2}^{(m)} - \mu) \\
& + (1 - \hat{B}_{1m} - \hat{B}_{2m})(1 - 2\hat{B}_{1m})(\bar{Y}_{s_1}^{(m)} - \mu)(\hat{\mu} - \mu) \\
& + (1 - \hat{B}_{1m} - \hat{B}_{2m})(1 - 2\hat{B}_{2m})(\bar{Y}_{s_2}^{(m)} - \mu)(\hat{\mu} - \mu) - (1 - \hat{B}_{1m} - \hat{B}_{2m})^2(\hat{\mu} - \mu)^2 \} \Big\}^2 \\
\leq & \frac{9}{N_m^2} \left\{ E \left\{ \frac{1}{(\lambda + 1)^2} \left[(n_{1m} - \hat{B}_{1m}) - \frac{n_{1m}}{N_m} (1 - \hat{B}_{1m} - \hat{B}_{2m}) \right]^2 \left[\frac{\lambda}{\lambda + 1} \text{WMS}_1 - \sigma_e^2 \right]^2 \right\} \right. \\
& + E \left\{ \frac{1}{(c\lambda + 1)^2} \left[(n_{2m} - \hat{B}_{2m}) - \frac{n_{2m}}{N_m} (1 - \hat{B}_{1m} - \hat{B}_{2m}) \right]^2 \left[\frac{c\lambda}{c\lambda + 1} \text{WMS}_2 - c\sigma_e^2 \right]^2 \right\} \\
& + 2 \left(\frac{\sigma_e^2}{\lambda + 1} \right)^2 \left[\left(1 - \frac{n_{1m}}{N_m} \right)^2 E(B_{1m} - \hat{B}_{1m})^2 + \left(\frac{n_{1m}}{N_m} \right)^2 E(\hat{B}_{2m} - B_{2m})^2 \right] \\
& + 2 \left(\frac{c\sigma_e^2}{c\lambda + 1} \right)^2 \left[\left(1 - \frac{n_{2m}}{N_m} \right)^2 E(B_{2m} - \hat{B}_{2m})^2 + \left(\frac{n_{2m}}{N_m} \right)^2 E(\hat{B}_{1m} - B_{1m})^2 \right] \\
& + 3 \left(\frac{N_{1m} - n_{1m}}{N_m} \right)^2 \left\{ E \left\{ [N_m - (1 - \hat{B}_{1m} - c\hat{B}_{2m})]^2 \left[\frac{\lambda}{\lambda + 1} \text{WMS}_1 - \sigma_e^2 \right]^2 \right\} \right. \\
& + \sigma_e^4 E(\hat{B}_{1m} - B_{1m})^2 + c^2 \sigma_e^4 E(\hat{B}_{2m} - B_{2m})^2 \} \\
& + 3 \left(\frac{N_m - N_{1m} - n_{2m}}{N_m} \right)^2 \left\{ E \left\{ [N_m - (1 - \frac{\hat{B}_{1m}}{c} - \hat{B}_{2m})]^2 \left[\frac{c\lambda}{c\lambda + 1} \text{WMS}_2 - c\sigma_e^2 \right]^2 \right\} \right. \\
& + \sigma_e^4 E(\hat{B}_{1m} - B_{1m})^2 + c^2 \sigma_e^4 E(\hat{B}_{2m} - B_{2m})^2 \} \\
& + \left[\left(\frac{\lambda}{\lambda + 1} \right)^2 n_{1m} \left(1 - \frac{n_{1m}}{N_m} \right) \right]^2 E \{ [(\hat{B}_{1m}^2 - B_{1m}^2) - 2(\hat{B}_{1m} - B_{1m})](\bar{Y}_{s_1}^{(m)} - \mu)^2 \\
& + (\hat{B}_{2m}^2 - B_{2m}^2)(\bar{Y}_{s_2}^{(m)} - \mu)^2 + 2[\hat{B}_{2m}(\hat{B}_{1m} - B_{1m}) - (1 - B_{1m}) \\
& (\hat{B}_{2m} - B_{2m})](\bar{Y}_{s_1}^{(m)} - \mu)(\bar{Y}_{s_2}^{(m)} - \mu) - 2(1 - \hat{B}_{1m})(1 - \hat{B}_{1m} - \hat{B}_{2m})(\bar{Y}_{s_1}^{(m)} - \mu)(\hat{\mu} - \mu) \\
& + 2\hat{B}_{2m}(1 - \hat{B}_{1m} - \hat{B}_{2m})(\bar{Y}_{s_2}^{(m)} - \mu)(\hat{\mu} - \mu) + (1 - \hat{B}_{1m} - \hat{B}_{2m})^2(\hat{\mu} - \mu)^2 \}^2 \\
& + \left[\left(\frac{c\lambda}{c\lambda + 1} \right)^2 n_{2m} \left(1 - \frac{n_{2m}}{N_m} \right) \right]^2 E \{ (\hat{B}_{1m}^2 - B_{1m}^2)(\bar{Y}_{s_1}^{(m)} - \mu)^2 + [(\hat{B}_{2m}^2 - B_{2m}^2)
\end{aligned}$$

$$\begin{aligned}
& -2(\hat{B}_{2m} - B_{2m})(\bar{Y}_{s_2}^{(m)} - \mu)^2 + 2(\hat{B}_{1m}(\hat{B}_{2m} - B_{2m}) - (1 - B_{2m})(\hat{B}_{1m} - B_{1m})) \\
& (\bar{Y}_{s_1}^{(m)} - \mu)(\bar{Y}_{s_2}^{(m)} - \mu) + 2\hat{B}_{1m}(1 - \hat{B}_{1m} - \hat{B}_{2m})(\bar{Y}_{s_1}^{(m)} - \mu)(\hat{\mu} - \mu) \\
& -2(1 - \hat{B}_{2m})(1 - \hat{B}_{1m} - \hat{B}_{2m})(\bar{Y}_{s_2}^{(m)} - \mu)(\hat{\mu} - \mu) + (1 - \hat{B}_{1m} - \hat{B}_{2m})^2(\hat{\mu} - \mu)^2\} \\
& + \left[\frac{2n_{1m}n_{2m}}{N_m} \left(\frac{\lambda}{\lambda + 1} \right) \left(\frac{c\lambda}{c\lambda + 1} \right) \right]^2 E\{[(\hat{B}_{1m}^2 - B_{1m}^2) + (\hat{B}_{1m} - B_{1m})](\bar{Y}_{s_1}^{(m)} - \mu)^2 \\
& + [(\hat{B}_{2m}^2 - B_{2m}^2) + (\hat{B}_{2m} - B_{2m})](\bar{Y}_{s_2}^{(m)} - \mu)^2 + [(\hat{B}_{1m} - B_{1m}) + (\hat{B}_{2m} - B_{2m}) \\
& - 2B_{2m}(\hat{B}_{1m} - B_{1m}) - 2\hat{B}_{1m}(\hat{B}_{2m} - B_{2m})](\bar{Y}_{s_1}^{(m)} - \mu)(\bar{Y}_{s_2}^{(m)} - \mu) \\
& + (1 - \hat{B}_{1m} - \hat{B}_{2m})(1 - 2\hat{B}_{1m})(\bar{Y}_{s_1}^{(m)} - \mu)(\hat{\mu} - \mu) + (1 - \hat{B}_{1m} - \hat{B}_{2m}) \\
& (1 - 2\hat{B}_{2m})(\bar{Y}_{s_2}^{(m)} - \mu)(\hat{\mu} - \mu) - (1 - \hat{B}_{1m} - \hat{B}_{2m})^2(\hat{\mu} - \mu)^2\}^2 \tag{3.43}
\end{aligned}$$

Mas

$$\begin{aligned}
E\left(\frac{\lambda}{\lambda + 1} \text{WMS}_1 - \sigma_e^2\right)^2 &= \text{Var}\left(\frac{\lambda}{\lambda + 1} \text{WMS}_1\right) \rightarrow 0 \quad \text{quando } m \rightarrow \infty, \\
E\left(\frac{c\lambda}{c\lambda + 1} \text{WMS}_2 - c\sigma_e^2\right)^2 &= \text{Var}\left(\frac{c\lambda}{c\lambda + 1} \text{WMS}_2\right) \rightarrow 0 \quad \text{quando } m \rightarrow \infty.
\end{aligned}$$

Utilizando-se o desenvolvimento da prova do Lema 3.4 temos que

$$\begin{aligned}
E(\hat{B}_{1m} - B_{1m})^2 &\leq E[k|\hat{M}_1^{-1} - M_1^{-1}| + k^2|\hat{M}_1^{-1}M_2^{-1} - \hat{M}_1^{-1}\hat{M}_2^{-1}|]^2 \\
&\leq E[k|\hat{M}_1^{-1} - M_1^{-1}| + k^2|M_2^{-1}(\hat{M}_1^{-1} - M_1^{-1}) - M_1^{-1}(\hat{M}_2^{-1} - M_2^{-1})|^2] \\
&\leq E[(k + k^2M_2^{-1})|\hat{M}_1^{-1} - M_1^{-1}| + k^2M_1^{-1}|\hat{M}_2^{-1} - M_2^{-1}|]^2 \\
&\leq 2(k + k^2M_2^{-1})^2 E(\hat{M}_1^{-1} - M_1^{-1})^2 + 2k^4M_1^{-2} E(\hat{M}_2^{-1} - M_2^{-1})^2 \rightarrow 0 \quad \text{quando } m \rightarrow \infty
\end{aligned}$$

e de maneira análoga

$$E(\hat{B}_{2m} - B_{2m})^2 \rightarrow 0 \quad \text{quando } m \rightarrow \infty.$$

Com relação aos três últimos termos de (3.43), para mostrarmos que as esperanças vão para zero, temos que mostrar que seus argumentos convergem para zero em probabilidade e também que seus quadrados são uniformemente integráveis em m . Isto pode ser verificado

utilizando-se argumentos similares aos considerados nas provas dos Teoremas 3.2, 3.3 e 3.4.

Assumiremos agora que c é desconhecido. Consideramos como um estimador de c ,

$$\hat{c}^* = \left[(n_{1T} - m)^{-1} (n_{1T} - m - 2) \frac{\text{WMS}_2}{\text{WMS}_1} \right] (1 + \lambda^{-1}) - \lambda^{-1}.$$

Mas, como $\frac{\text{WMS}_2/c\sigma_e^2 + \sigma_v^2}{\text{WMS}_1/\sigma_e^2 + \sigma_v^2} \sim F_{(n_{2T}-m, n_{1T}-m)}$ temos que $E[\hat{c}^*] = c$ e

$$\text{Var}[\hat{c}^*] = \frac{2(c + \lambda^{-1})^2(n_{2T} - 2m - 2)}{(n_{2T} - m)(n_{1T} - m - 4)} \leq \frac{2(c + \lambda^{-1})^2 k}{m - 4} \rightarrow 0 \quad \text{quando } m \rightarrow \infty,$$

segue então que $\hat{c}^*$ é um estimador consistente de c . Notemos que $\hat{c}^*$ pode não estar entre 0 e 1, então propomos

$$\hat{c} = \begin{cases} 0 & , \hat{c}^* < 0 \\ \hat{c}^* & , 0 \leq \hat{c}^* \leq 1 \\ 1 & , \text{se } \hat{c}^* > 1 \end{cases} \quad (3.44)$$

que também é consistente. Neste caso, o previsor Bayesiano empírico de $T^{(m)}$, que denotaremos por $\hat{\hat{T}}_{EB}$, e de $S^{2(m)}$, que denotaremos por $\hat{\hat{S}}_{EB}^2$, podem ser obtidos do previsor $\hat{T}_{EB}$ e $\hat{S}_{EB}^2$ substituindo-se c por $\hat{c}$. A seguir estabelecemos a otimalidade assintótica de $\hat{\hat{T}}_{EB}$ e $\hat{\hat{S}}_{EB}^2$.

Teorema 3.6. *Sob as suposições a), b) e c)*

$$(i) \quad r(\xi, \hat{\hat{T}}_{EB}) - r(\xi, \hat{T}_B) \rightarrow 0$$

e

$$(ii) \quad r(\xi, \hat{\hat{S}}_{EB}^2) - r(\xi, \hat{S}_B^2) \rightarrow 0$$

quando $m \rightarrow \infty$.

Prova. (i) Notemos que

$$r(\xi, \hat{\hat{T}}_{EB}) - r(\xi, \hat{T}_B) = E[\hat{\hat{T}}_{EB} - \hat{T}_B]^2 = E[A_m^2]$$

onde

$$A_m^2 = \left\{ \left(N_m - n_{1m} \frac{\lambda}{\lambda + 1} - \frac{n_{2m} \hat{c} \lambda}{\hat{c} \lambda + 1} \right) [(\hat{B}_{1m} - B_{1m})(\bar{Y}_{s_1}^{(m)} - \mu) \right.$$

$$\begin{aligned}
& +(\hat{B}_{2m} - B_{2m})(\bar{Y}_{s_2}^{(m)} - \mu) + (1 - \hat{B}_{1m} - \hat{B}_{2m})(\hat{\mu} - \mu)] \\
& - n_{2m} \left(\frac{\hat{c}\lambda}{\hat{c}\lambda + 1} - \frac{c\lambda}{c\lambda + 1} \right) [B_{1m}(\bar{Y}_{s_1}^{(m)} - \mu) - (1 - B_{2m})(\bar{Y}_{s_2}^{(m)} - \mu)] \Big\}^2
\end{aligned}$$

onde $\hat{\mu}$ é obtido substituindo-se c por $\hat{c}$ em $\hat{\mu}$.

Como na prova do Teorema 3.4,

$$[(\hat{B}_{1m} - B_{1m})(\bar{Y}_{s_1}^{(m)} - \mu) + (\hat{B}_{2m} - B_{2m})(\bar{Y}_{s_2}^{(m)} - \mu) + (1 - \hat{B}_{1m} - \hat{B}_{2m})(\hat{\mu} - \mu)] \xrightarrow{P} 0 .$$

Além disso,

$$\frac{\hat{c}\lambda}{\hat{c}\lambda + 1} \xrightarrow{P} \frac{c\lambda}{c\lambda + 1} .$$

Logo, $A_m \xrightarrow{P} 0$ quando $m \rightarrow \infty$. Para concluirmos a prova, falta mostrarmos que A_m^2 é uniformemente integrável em m . Mas, podemos mostrar que

$$\begin{aligned}
A_m^2 \leq & 2 \left\{ 3 \left(N_m - \frac{n_{1m}\lambda}{\lambda + 1} - \frac{n_{2m}\hat{c}\lambda}{\hat{c}\lambda + 1} \right)^2 [(\bar{Y}_{s_1}^{(m)} - \mu)^2 + (\bar{Y}_{s_2}^{(m)} - \mu)^2 + (\hat{\mu} - \mu)^2] \right. \\
& \left. + 2k^2 [(\bar{Y}_{s_1}^{(m)} - \mu)^2 + (\bar{Y}_{s_2}^{(m)} - \mu)^2] \right\}
\end{aligned}$$

de onde concluímos (ver prova do Teorema 3.4) que A_m^2 é uniformemente integrável em m , e então a prova de (i) está completa.

(ii) Temos que

$$r(\xi, \hat{\hat{S}}_{EB}) - r(\xi, \hat{S}_B) = E[\hat{\hat{S}}_{EB} - \hat{S}_B]^2 = E[B_m^2]$$

onde

$$\begin{aligned}
B_m^2 = & \frac{1}{N_m} \left\{ \frac{1}{\lambda + 1} \left[(n_{1m} - \hat{B}_{1m}) - \frac{n_{1m}}{N_m} (1 - \hat{B}_{1m} - \hat{B}_{2m}) \right] \left(\frac{\lambda \text{WMS}_1}{\lambda + 1} - \sigma_e^2 \right) \right. \\
& + \frac{1}{c\lambda + 1} \left[(n_{2m} - \hat{B}_{2m}) - \frac{n_{2m}}{N_m} (1 - \hat{B}_{1m} - \hat{B}_{2m}) \right] \left(\frac{c\lambda \text{WMS}_2}{c\lambda + 1} - c\sigma_e^2 \right) \\
& + \frac{\sigma_e^2}{\lambda + 1} \left[\left(1 - \frac{n_{1m}}{N_m} \right) (B_{1m} - \hat{B}_{1m}) + \frac{n_{1m}}{N_m} (\hat{B}_{2m} - B_{2m}) \right] \\
& \left. + \frac{c\sigma_e^2}{c\lambda + 1} \left[\left(1 - \frac{n_{2m}}{N_m} \right) (B_{2m} - \hat{B}_{2m}) + \frac{n_{2m}}{N_m} (\hat{B}_{1m} - B_{1m}) \right] \right\}
\end{aligned}$$

$$\begin{aligned}
& + \left(\frac{N_{1m} - n_{1m}}{N_m} \right) \left\{ [N_m - (1 - \hat{B}_{1m} - \hat{c}\hat{B}_{2m})] \left[\frac{\lambda}{\lambda + 1} \text{WMS}_1 - \sigma_e^2 \right] \right. \\
& + \hat{\sigma}_e^2 (\hat{B}_{1m} - B_{1m}) + \hat{c}\sigma_e^2 (\hat{B}_{2m} - B_{2m}) \left. \right\} \\
& + \left(\frac{N_m - N_{1m} - n_{2m}}{N_m} \right) \left\{ \left[N_m - \left(1 - \frac{\hat{B}_{1m}}{\hat{c}} - \hat{B}_{2m} \right) \right] \left(\frac{c\lambda}{c\lambda + 1} \text{WMS}_2 - c\sigma_e^2 \right) \right. \\
& + \frac{c\sigma_e^2}{\hat{c}} (\hat{B}_{1m} - B_{1m}) + c\sigma_e^2 (\hat{B}_{2m} - B_{2m}) \left. \right\} \\
& + \left[\left(\frac{\lambda}{\lambda + 1} \right)^2 n_{1m} \left(1 - \frac{n_{1m}}{N_m} \right) \right] \left\{ [(\hat{B}_{1m}^2 - B_{1m}^2) - 2(\hat{B}_{1m} - B_{1m})](\bar{Y}_{s_1}^{(m)} - \mu)^2 \right. \\
& + (\hat{B}_{2m}^2 - B_{2m}^2)(\bar{Y}_{s_2}^{(m)} - \mu)^2 + 2[\hat{B}_{2m}(\hat{B}_{1m} - B_{1m}) - (1 - B_{1m})(\hat{B}_{2m} - B_{2m})] \\
& (\bar{Y}_{s_1}^{(m)} - \mu)(\bar{Y}_{s_2}^{(m)} - \mu) - 2(1 - \hat{B}_{2m})(1 - \hat{B}_{1m} - \hat{B}_{2m})(\bar{Y}_{s_1}^{(m)} - \mu)(\hat{\mu} - \mu) \\
& + 2\hat{B}_{2m}(1 - \hat{B}_{1m} - \hat{B}_{2m})(\bar{Y}_{s_2}^{(m)} - \mu)(\hat{\mu} - \mu) + (1 - \hat{B}_{1m} - \hat{B}_{2m})^2(\hat{\mu} - \mu)^2 \left. \right\} \\
& + \left[\left(\frac{\hat{c}\lambda}{\hat{c}\lambda + 1} \right)^2 n_{2m} \left(1 - \frac{n_{2m}}{N_m} \right) \right] \left\{ (\hat{B}_{1m}^2 - B_{1m}^2)(\bar{Y}_{s_1}^{(m)} - \mu)^2 \right. \\
& + [(\hat{B}_{2m}^2 - B_{2m}^2) - 2(\hat{B}_{2m} - B_{2m})](\bar{Y}_{s_2}^{(m)} - \mu)^2 + 2[\hat{B}_{1m}(\hat{B}_{2m} - B_{2m}) \\
& - (1 - B_{2m})(\hat{B}_{1m} - B_{1m})](\bar{Y}_{s_1}^{(m)} - \mu)(\bar{Y}_{s_2}^{(m)} - \mu) \\
& + 2\hat{B}_{1m}(1 - \hat{B}_{1m} - \hat{B}_{2m})(\bar{Y}_{s_1}^{(m)} - \mu)(\hat{\mu} - \mu) \\
& - 2(1 - \hat{B}_{2m})(1 - \hat{B}_{1m} - \hat{B}_{2m})(\bar{Y}_{s_2}^{(m)} - \mu)(\hat{\mu} - \mu) + (1 - \hat{B}_{1m} - \hat{B}_{2m})^2(\hat{\mu} - \mu)^2 \left. \right\} \\
& + 2\frac{n_{1m}n_{2m}}{N_m} \left(\frac{\lambda}{\lambda + 1} \right) \left(\frac{\hat{c}\lambda}{\hat{c}\lambda + 1} \right) \left\{ [(\hat{B}_{1m}^2 - B_{1m}^2) + (\hat{B}_{1m} - B_{1m})](\bar{Y}_{s_1}^{(m)} - \mu)^2 \right. \\
& + [(\hat{B}_{2m}^2 - B_{2m}^2) + (\hat{B}_{2m} - B_{2m})](\bar{Y}_{s_2}^{(m)} - \mu)^2 \\
& + [(\hat{B}_{1m} - B_{1m}) + (\hat{B}_{2m} - B_{2m}) - 2B_{2m}(\hat{B}_{1m} - B_{1m}) - 2\hat{B}_{1m}(\hat{B}_{2m} - B_{2m})] \\
& (\bar{Y}_{s_1}^{(m)} - \mu)(\bar{Y}_{s_2}^{(m)} - \mu) + (1 - \hat{B}_{1m} - \hat{B}_{2m})(1 - 2\hat{B}_{1m})(\bar{Y}_{s_1}^{(m)} - \mu)(\hat{\mu} - \mu) \\
& + (1 - \hat{B}_{1m} - \hat{B}_{2m})(1 - 2\hat{B}_{2m})(\bar{Y}_{s_2}^{(m)} - \mu)(\hat{\mu} - \mu) - (1 - \hat{B}_{1m} - \hat{B}_{2m})^2(\hat{\mu} - \mu)^2 \left. \right\} \\
& + \left[(n_{2m} - \hat{B}_{2m}) - \frac{n_{2m}}{N_m} (1 - \hat{B}_{1m} - \hat{B}_{2m}) \right] \left(\frac{\hat{c}\lambda}{(\hat{c}\lambda + 1)^2} - \frac{c\lambda}{(c\lambda + 1)^2} \right) \text{WMS}_2 \\
& + \left(\frac{N_{1m} - n_{1m}}{N_m} \right) \sigma_e^2 B_2 (\hat{c} - c) + \left(\frac{N_m - N_{1m} - n_{2m}}{N_m} \right) c\sigma_e^2 B_1 (\hat{c}^{-1} - c^{-1}) \\
& + \left(\frac{N_m - N_{1m} - n_{2m}}{N_m} \right) \left[N_m - \left(1 - \frac{\hat{B}_{1m}}{\hat{c}} - \hat{B}_{2m} \right) \right] \left(\frac{\hat{c}\lambda}{\hat{c}\lambda + 1} - \frac{c\lambda}{c\lambda + 1} \right) \text{WMS}_2
\end{aligned}$$

$$\begin{aligned}
& + \left[\left(\frac{\hat{c}\lambda}{c\lambda + 1} \right)^2 - \left(\frac{c\lambda}{c\lambda + 1} \right)^2 \right] \left\{ n_{2m} s_{y_2}^2 + n_{2m} \left(1 - \frac{n_{2m}}{N_m} \right) [(1 - B_{2m})(\bar{Y}_{s_2}^{(m)} - \mu) \right. \\
& - B_{1m}(\bar{Y}_{s_1}^{(m)} - \mu)]^2 \left. \right\} + \frac{2n_{1m}n_{2m}}{N_m} \frac{\lambda}{\lambda + 1} \left(\frac{c\lambda}{c\lambda + 1} - \frac{\hat{c}\lambda}{\hat{c}\lambda + 1} \right) [(1 - B_{1m})(\bar{Y}_{s_1}^{(m)} - \mu) \\
& - B_{2m}(\bar{Y}_{s_2}^{(m)} - \mu)][(1 - B_{2m})(\bar{Y}_{s_2}^{(m)} - \mu) - B_{1m}(\bar{Y}_{s_1}^{(m)} - \mu)] \left. \right\}^2 .
\end{aligned}$$

A prova de (ii) segue da prova do Teorema 3.5 e também do fato de $\hat{c} \xrightarrow{P} c$.

3.6 ILUSTRAÇÃO NUMÉRICA

Para ilustrar o comportamento dos previsores Bayesianos empíricos $\hat{T}_{EB}$ e $\hat{S}_{EB}$ derivados nas seções 3.3 e 3.5, apresentamos agora estudos de Monte Carlo. O computador utilizado foi o Unisys A15 do Centro de Computação Eletrônica da USP. Consideramos inicialmente o modelo especificado em (3.7) com $\theta^{(j)} \sim N(20, 16)$, $e_i^{(j)} \sim N(0, 4)$ e $v_i^{(j)} \sim N(0, \sigma_v^2)$ onde $\sigma_v^2 = 1, 0; 0, 25; 0, 04$ então $\lambda = 9, 36$ e 224 , $i = 1, \dots, N_j$, $j = 1, \dots, m$. Foi utilizada a subrotina *GGNML* do IMSL. Tomamos $m = 5, 10, 15, 20$ e 50 e N_j distribuído uniformemente no intervalo $(100, 200)$. O tamanho amostral n_j no estágio j é aproximadamente 10% do valor de N_j . A Tabela 3.1 apresenta os valores de (N_j, n_j) usados no estudo gerado para $m = 10$.

Tabela 3.1. Valores de N_j e n_j para $m = 10$

N_j	143	154	141	146	143	155	148	124	129	175
n_j	14	15	14	14	14	15	14	12	12	17

O total e a variância populacionais no estágio m estão apresentados na Tabela 3.2.

Tabela 3.2. Valores de $T^{(m)}$ e $S_y^{2(m)}$

m	$T^{(m)}$	$S_y^{2(m)}$
5	2406.17	9.50
10	2881.89	7.97
15	2244.46	9.28
20	3831.31	9.81
50	2065.71	7.33

Para obtermos cada entrada da Tabela 3.3, geramos 1000 amostras casuais simples sem reposição para cada valor de λ e m e calculamos

$$\text{ERD}(\hat{T}_{EB}, \hat{T}_B) = \frac{1}{1000} \sum_{i=1}^{1000} (\hat{T}_{EB}^{(i)} - \hat{T}_B^{(i)})^2,$$

isto é, uma estimativa da diferença dos riscos dos previsores $\hat{T}_{EB}$ e $\hat{T}_B$ (dados na seção 3.3), onde $\hat{T}_{EB}^{(i)}$ e $\hat{T}_B^{(i)}$ correspondem aos valores dos previsores de $\hat{T}_{EB}$ e $\hat{T}_B$, calculados a partir da i -ésima amostra gerada.

A Tabela 3.3 apresenta entre parêntesis os valores correspondentes ao $\text{ERD}(\hat{S}_{EB}^2, \hat{S}_B^2)$ calculados a partir das amostras geradas.

Tabela 3.3. ERD de $\hat{T}_{EB}$ e $\hat{S}_{EB}^2$ (entre parêntesis) para diferentes valores de m e λ .

m	$\lambda = 9$	$\lambda = 36$	$\lambda = 224$
5	1860,12 (3,73)	1377,51 (2,82)	1187,27 (2,32)
10	1476,79 (1,55)	1143,89 (1,22)	1028,09 (1,09)
15	631,81 (0,56)	482,03 (0,52)	429,87 (0,53)
20	45,64 (0,96)	40,21 (0,85)	39,86 (0,78)
50	1,56 (0,18)	1,34 (0,18)	1,32 (0,19)

Consideramos agora o modelo não homogêneo (3.34) descrito na seção 3.5.2, com $\theta^{(j)} \sim N(20, 16)$, $e_i^{(j)} \sim N(0, 9)$, $i = 1, \dots, N_{1j}$, $v_i^{(j)} \sim N(0, \sigma_v^2)$, $i = 1, \dots, N_j$, com σ_v^2 escolhido de forma que $\lambda = 9, 36$ e 224 . Os valores de m são os mesmos que anteriormente e consideramos cinco valores de c : 0,5 para todos os valores de λ e 0, 1; 0,3; 0,7 e 0,9 para $\lambda = 36$. Os valores N_{1j} e N_{2j} foram gerados de distribuições uniformes no

intervalo (50, 100) e (60, 120), respectivamente. Os tamanhos amostrais considerados são aproximadamente 10% dos valores gerados de N_{1j} e N_{2j} . Como ilustração, a Tabela 3.4 apresenta os valores de $T^{(m)}$ e $S_y^{2(m)}$ para $c = 0, 5$.

Tabela 3.4. Valores de $T^{(m)}$ e $S_y^{2(m)}$ para $c = 0, 5$

m	5	10	15	20	50
$T^{(m)}$	2982,0	2844,5	2345,8	4403,1	2509,8
$S_y^{2(m)}$	5,84	6,20	7,37	7,07	6,70

Foram geradas 1000 amostras e foram calculados $\text{ERD}(\hat{T}_{EB}, \hat{T}_B)$ e $\text{ERD}(\hat{S}_{EB}^2, \hat{S}_B^2)$ para cada uma das combinações de m, λ e c descritas acima, onde $\hat{T}_{EB}, \hat{T}_B, \hat{S}_{EB}^2$ e $\hat{S}_B^2$ são dados na seção 3.5. A Tabela 3.5 apresenta os valores de ERD para $c = 0, 5$ e todos os valores de λ . A Tabela 3.6 apresenta os valores de ERD para $c = 0, 1; 0, 3; 0, 7$ e $0, 9$ com $\lambda = 36$.

Tabela 3.5. ERD de $\hat{T}_{EB}$ e $\hat{S}_{EB}^2$ para $c = 0, 5$

m	$\lambda = 9$	$\lambda = 36$	$\lambda = 224$
5	940,83 (1,98)	641,63 (1,57)	539,22 (1,35)
10	829,26 (1,21)	591,54 (1,07)	510,44 (0,96)
15	390,63 (0,50)	293,09 (0,41)	260,69 (0,37)
20	97,48 (0,36)	75,83 (0,33)	68,87 (0,31)
50	12,65 (0,08)	10,40 (0,08)	9,83 (0,07)

Tabela 3.6. ERD de $\hat{T}_{EB}$ e $\hat{S}_{EB}^2$ para $c = 0,1$ e $0,9$ e $\lambda = 36$

m	$c = 0,1$	$c = 0,3$	$c = 0,7$	$c = 0,9$
5	71,79 (1,24)	325,72 (1,37)	969,68 (1,86)	1291,35 (2,22)
10	94,85 (0,96)	345,82 (1,01)	803,32 (1,15)	981,57 (1,25)
15	39,07 (0,32)	161,16 (0,35)	416,35 (0,50)	527,53 (0,61)
20	12,47 (0,21)	45,37 (0,25)	101,37 (0,43)	122,82 (0,56)
50	1,35 (0,05)	5,68 (0,06)	14,72 (0,10)	18,47 (0,13)

A partir dos estudos de simulação feitos acima, notamos que os ERDs convergem para zero razoavelmente rápido com m . Além disso, valores grandes de λ parecem afetar o risco dos previsores $\hat{T}_{EB}$ e $\hat{S}_{EB}^2$, principalmente para valores pequenos de m e isso é menos notado à medida que m cresce. As Tabelas 3.5 e 3.6 mostram também que os ERDs crescem com c , especialmente para pequenos valores de m .

Ainda sob o modelo não homogêneo com $c = 0,5$ e $\lambda = 36$ elaboramos a Tabela 3.7 com os valores de ERD considerando c desconhecido, e portanto utilizando os previsores $\hat{\hat{T}}_{EB}$ e $\hat{\hat{S}}_{EB}$.

Tabela 3.7. ERD de $\hat{\hat{T}}_{EB}$ e $\hat{\hat{S}}_{EB}^2$ para $c = 0,5$ e $\lambda = 36$

m	ERD
5	640,64 (1,49)
10	591,36 (1,02)
15	292,55 (0,40)
20	75,36 (0,32)
50	10,31 (0,07)

Comparando essa tabela com a Tabela 3.5, notamos que o comportamento do ERD não se alterou ao supormos c desconhecido e ao utilizarmos o previsor dado em (3.44).

CAPÍTULO 4

PREVISÃO DO TOTAL POPULACIONAL SOB UM MODELO DE REGRESSÃO COM ERRO NAS VARIÁVEIS

4.1 INTRODUÇÃO

Neste capítulo nosso interesse é a previsão do total populacional, considerando um modelo de regressão linear estrutural (Fuller (1987)).

A população finita é denotada por $\mathcal{P} = \{1, \dots, N\}$, onde N é conhecido. Associado com a k -ésima unidade existe um vetor desconhecido (y_k, x_k) , $k = 1, \dots, N$. Para obtermos informação sobre o total populacional $T = \sum_{i=1}^N y_i$, uma amostra s de tamanho n é selecionada de $\mathcal{P}$ de acordo com algum plano amostral. Consideremos que a população $\mathcal{P}$

foi gerada de acordo com o modelo definido pelas equações

$$\begin{aligned} y_k &= \alpha + \beta x_k + e_k \\ X_k &= x_k + u_k \end{aligned} \quad (4.1)$$

para $k = 1, \dots, N$, com as suposições de que e_k e u_k são variáveis aleatórias normais e independentes, com média zero e variâncias σ_e^2 e σ_u^2 , respectivamente. Supondo x_k ($k = 1, \dots, N$) como sendo variável aleatória com distribuição normal, com média μ_x e variância σ_x^2 , temos caracterizado o chamado modelo estrutural. Então, temos que

$$\begin{pmatrix} e_i \\ u_i \\ x_i \end{pmatrix} \sim N \left(\begin{pmatrix} 0 \\ 0 \\ \mu_x \end{pmatrix} ; \begin{pmatrix} \sigma_e^2 & 0 & 0 \\ 0 & \sigma_u^2 & 0 \\ 0 & 0 & \sigma_x^2 \end{pmatrix} \right)$$

onde faremos as suposições adicionais de que x_k são independentes entre si e também independentes de e_k e u_k , para todo $k = 1, \dots, N$. Para este modelo, temos que $x_1, x_2, \dots, x_N$ não são observados diretamente. As quantidades observadas são y_k para $k \in s$ e X_k para $k = 1, \dots, N$. A partir do modelo (4.1) e com as suposições feitas, temos que (y_k, X_k) são independentes e tem distribuição normal bivariada, isto é,

$$\begin{pmatrix} y_k \\ X_k \end{pmatrix} \sim N \left(\begin{pmatrix} \alpha + \beta \mu_x \\ \mu_x \end{pmatrix} ; \begin{pmatrix} \beta^2 \sigma_x^2 + \sigma_e^2 & \beta \sigma_x^2 \\ \beta \sigma_x^2 & \sigma_x^2 + \sigma_u^2 \end{pmatrix} \right). \quad (4.2)$$

Na seção 4.2 consideramos alguns previsores do tipo regressão e encontramos suas distribuições assintóticas. Através da eficiência relativa assintótica são feitas algumas comparações entre os previsores considerados. Também são obtidos alguns resultados assintóticos para o modelo sem intercepto. Na seção 4.3 adotamos o enfoque condicional (Bolfarine e Rodrigues (1990)) e utilizamos uma parametrização ortogonal para facilitar as inferências. A partir daí, obtemos o preditor de mínimos quadrados condicionais e, adotando uma priori não informativa para os parâmetros ortogonais, obtemos o preditor Bayesiano. Generalizamos o modelo (4.1), isto é, consideramos um modelo com várias variáveis auxiliares na seção 4.4. Utilizando resultados de Fuller (1987), definimos um preditor do tipo regressão para o total e estudamos seu comportamento assintótico. Re-

sultados assintóticos em modelos de regressão para populações finitas, onde o interesse é o coeficiente de regressão β , são considerados em Fuller (1975).

4.2 PREVISORES DO TIPO REGRESSÃO

Nesta seção consideramos uma classe geral de previsores para o total populacional que, sob o modelo de superpopulação (4.1), pode ser definida como

$$\hat{T}_R = N\bar{y}_s + (N - n)(\bar{X}_r - \bar{X}_s)\hat{\beta} \quad (4.3)$$

onde $\bar{X}_s = \frac{1}{n} \sum_{i \in s} X_i$, $\bar{X}_r = \frac{1}{N-n} \sum_{i \notin s} X_i$, $\bar{y}_s = \frac{1}{n} \sum_{i \in s} y_i$ e $\hat{\beta}$ é um estimador de β .

Se o estimador $\hat{\beta}$ for independente de $\bar{X}_s$, então o preditor regressão (4.3) é não viciado, pois

$$E[\hat{T}_R - T] = (N - n)\{E(\bar{X}_s - \bar{X}_r)E(\hat{\beta}) - E(\bar{y}_s - \bar{y}_r)\} = 0.$$

Notamos de (4.3) que diferentes previsores do tipo regressão podem ser definidos considerando-se diferentes estimadores de β .

Se $\sigma_u^2 = 0$ no modelo (4.1), ou seja, se x_i foi medido sem erro, é bem conhecido que o estimador de mínimos quadrados ordinários $[\sum_{i \in s} (x_i - \bar{x})^2]^{-1} \sum_{i \in s} (x_i - \bar{x})(y_i - \bar{y})$ é não viciado para β e de variância mínima. Então, o primeiro estimador considerado será

$$\hat{\beta}_1 = \frac{S_{XY}}{S_X^2}$$

onde

$$S_{XY} = \frac{1}{n-1} \sum_{i \in s} (X_i - \bar{X}_s)(y_i - \bar{y}_s)$$

e

$$S_X^2 = \frac{1}{n-1} \sum_{i \in s} (X_i - \bar{X}_s)^2$$

que é o estimador usual do coeficiente de regressão computado a partir dos dados observados.

Mas, sob o modelo (4.1), $E(\hat{\beta}_1) = k_x \beta$ onde $k_x = \frac{\sigma_x^2}{\sigma_x^2 + \sigma_u^2}$, sendo $\sigma_x^2 = \sigma_x^2 + \sigma_u^2$, então se conhecermos a razão k_x , usualmente chamada de fator de atenuação (Fuller (1987)), é possível construirmos um estimador não viciado para β , isto é,

$$\hat{\beta}_2 = k_x^{-1} \frac{S_{XY}}{S_X^2}.$$

A denominação fator de atenuação segue do fato de $k_x \in [0, 1]$. Dependendo do contexto podemos encontrar outras denominações para k_x : coeficiente de atenuação, razão de confiabilidade, herdabilidade, comunalidade, etc.. O fator de atenuação representa o quanto da variabilidade de X_i é explicada pela variabilidade de x_i . Se $\sigma_u^2 = 0$ então $k_x = 1$ e $E(\hat{\beta}_1) = \beta$, que é o modelo sem erro na variável independente. Se k_x for pequeno, por exemplo, 0,10, então $E(\hat{\beta}_1) = 0,10\beta$ e se for grande, por exemplo, 0,90, então $E(\hat{\beta}_1) = 0,90\beta$. Segundo Fuller (1987), existe um número grande de situações, particularmente em psicologia, biologia e sociologia, onde k_x é tão bem estimado que pode ser tratado como sendo conhecido. Fuller (1987) apresenta uma tabela com valores de k_x para diversas variáveis sendo o maior valor apresentado 0,98 para a variável sexo, e o menor 0,58 para índice de pobreza. Isto é bastante razoável, uma vez que respostas sobre sexo não são passíveis de erros tão grandes como podem ser aquelas relativas ao índice de pobreza.

O modelo estrutural dado em (4.1) não é identificável, no sentido de que dois diferentes vetores de parâmetros podem dar origem à mesma distribuição amostral. Temos um modelo com seis parâmetros $(\mu_x, \alpha, \beta, \sigma_u^2, \sigma_e^2, \sigma_x^2)$ e o conjunto de estatísticas suficientes associado à distribuição normal bivariada (4.2) estima os parâmetros $(E(X), E(Y), \text{Var}(X), \text{Var}(Y), \text{Cov}(X, Y))$. Desta forma, não é possível encontrar solução de máxima verossimilhança para β , daí a necessidade de reduzirmos o número de parâmetros desconhecidos, colocando alguma suposição adicional.

A seguir analisamos dois tipos de informação adicional. Supondo que a variância do

erro de medida, σ_u^2 , é conhecida, Fuller (1987) mostrou que o estimador obtido através do método dos momentos (método equivalente ao método de máxima verossimilhança sob normalidade) é

$$\hat{\beta}_3 = (S_X^2 - \sigma_u^2)^{-1} S_{XY} .$$

Supondo a razão $\delta = \sigma_e^2/\sigma_u^2$ conhecida, o modelo (4.1) é chamado, na literatura, de modelo clássico com erro nas variáveis, e o estimador obtido através do método dos momentos por Fuller (1987) é

$$\hat{\beta}_4 = \frac{S_Y^2 - \delta S_X^2 + \sqrt{(S_Y^2 - \delta S_X^2)^2 + 4\delta S_{XY}^2}}{2S_{XY}} .$$

onde $S_Y^2 = \frac{1}{n-1} \sum_{i \in S} (y_i - \bar{y}_s)^2$.

Em situações de laboratório, por exemplo, parece realista o uso dessa suposição, considerando que os pesquisadores conhecem a relação entre as variabilidades associadas aos erros de medida produzidos pelos seus aparelhos.

No Teorema 4.1 apresentamos a distribuição assintótica dos estimadores propostos acima. Para obtenção desse resultado utiliza-se o fato dos momentos amostrais S_X^2 , S_{XY} e S_Y^2 serem não viciados para $\sigma_X^2 = \sigma_x^2 + \sigma_u^2$, $\beta\sigma_x^2$ e $\beta^2\sigma_x^2 + \sigma_e^2$, respectivamente, e terem variâncias de ordem n^{-1} .

Teorema 4.1. *Sob o modelo (4.1) temos que quando $n \rightarrow \infty$*

- (i) $\sqrt{n}(\hat{\beta}_1 - k_x\beta) \xrightarrow{D} N(0; \sigma_X^{-4}[\sigma_X^2(\sigma_e^2 + \beta^2\sigma_u^2) - \beta^2\sigma_u^4])$,
- (ii) $\sqrt{n}(\hat{\beta}_2 - \beta) \xrightarrow{D} N(0; \sigma_x^{-4}[\sigma_X^2(\sigma_e^2 + \beta^2\sigma_u^2) - \beta^2\sigma_u^4])$,
- (iii) $\sqrt{n}(\hat{\beta}_3 - \beta) \xrightarrow{D} N(0; \sigma_x^{-4}[\sigma_X^2(\sigma_e^2 + \beta^2\sigma_u^2) + \beta^2\sigma_u^4])$,
- (iv) $\sqrt{n}(\hat{\beta}_4 - \beta) \xrightarrow{D} N(0; \sigma_x^{-4}[\sigma_X^2(\sigma_e^2 + \beta^2\sigma_u^2) - \beta^2\sigma_u^2])$.

Prova. Fuller (1987) apresenta nos Teoremas 1.2.1 e 1.3.1 a prova de (iii) e (iv) e de maneira similar provamos a seguir (ii).

Expandindo $\hat{\beta}_2$ em série de Taylor em torno dos valores populacionais, $\sigma_{XY}^2 = \beta\sigma_x^2$ e $\sigma_X^2 = \sigma_x^2 + \sigma_u^2$ temos

$$\hat{\beta}_2 = \frac{\sigma_{XY}^2}{k_x \sigma_X^2} + \frac{1}{k_x \sigma_X^2} (S_{XY} - \sigma_{XY}^2) - \frac{\sigma_{XY}^2}{k_x (\sigma_X^2)^2} (S_X^2 - \sigma_X^2) + O_p(n^{-1}) .$$

Logo,

$$\sqrt{n}(\hat{\beta}_2 - \beta) = (\sigma_x^2)^{-1} \sqrt{n} [S_{XY} - \beta k_x S_X^2 - \sigma_{XY}^2 + \beta k_x \sigma_X^2] + O_p(n^{-1/2}) .$$

Sendo $r_i = y_i - \gamma_0 - \gamma_1 X_i$, onde γ_0 é uma constante e $\gamma_1 = \beta \frac{\sigma_x^2}{\sigma_X^2}$, podemos escrever a expressão acima como

$$\sqrt{n}(\hat{\beta}_2 - \beta) = (\sigma_x^2)^{-1} \sqrt{n} \left[\frac{1}{n-1} \sum_{i \in s} (X_i - \bar{X})(r_i - \bar{r}) \right] + O_p(n^{-1/2}) .$$

Depois de algumas manipulações algébricas temos

$$\begin{aligned} & \sqrt{n} \left[\frac{1}{n-1} \sum_{i \in s} (X_i - \bar{X})(r_i - \bar{r}) \right] \\ &= \frac{n}{n-1} \frac{1}{\sqrt{n}} \sum_{i \in s} (x_i - \mu_x + u_i) [e_i - \beta k_x u_i - \beta(k_x - 1)(x_i - \mu_x)] \\ & \quad - \frac{\sqrt{n}}{n-1} \{ n[\bar{e}_s - \beta k_x \bar{u}_s - \beta(k_x - 1)(\bar{x}_s - \mu_x)](\bar{x}_s + \bar{u}_s - \mu_x) \} . \end{aligned}$$

Mas, $(x_i - \mu_x + u_i)[e_i - \beta k_x u_i - \beta(k_x - 1)(x_i - \mu_x)]$, $i \in s$, são variáveis aleatórias independentes com

$$E\{(x_i - \mu_x + u_i)[e_i - \beta k_x u_i - \beta(k_x - 1)(x_i - \mu_x)]\} = 0 \quad \text{e}$$

$$\text{Var} \{(x_i - \mu_x + u_i)[e_i - \beta k_x u_i - \beta(k_x - 1)(x_i - \mu_x)]\} = \sigma_X^2 (\sigma_e^2 + \beta^2 \sigma_u^2) - \beta^2 \sigma_u^4 .$$

Então, pelo Teorema do Limite Central,

$$\frac{n}{n-1} \frac{1}{\sqrt{n}} \sum_{i \in s} (x_i - \mu_x + u_i) [e_i - \beta k_x u_i - \beta(k_x - 1)(x_i - \mu_x)] \xrightarrow{\mathcal{D}} N(0; \sigma_X^2 (\sigma_e^2 + \beta^2 \sigma_u^2) - \beta^2 \sigma_u^4) . \quad (4.4)$$

Além disso, para $\forall \varepsilon > 0$

$$\begin{aligned}
& P \left(\left| \frac{\sqrt{n}}{n-1} \{n[\bar{e}_s - \beta k_x \bar{u}_s - \beta(k_x - 1)(\bar{x}_s - \mu_x)](\bar{x}_s + \bar{u}_s - \mu_x)\} \right| \geq \varepsilon \right) \\
& \leq \frac{E \left\{ \frac{\sqrt{n}}{n-1} n[\bar{e}_s - \beta k_x \bar{u}_s - \beta(k_x - 1)(\bar{x}_s - \mu_x)](\bar{x}_s + \bar{u}_s - \mu_x) \right\}^2}{\varepsilon^2} \\
& = \frac{n^3}{(n-1)^2 \varepsilon^2} \text{Var} \{ [\bar{e}_s - \beta k_x \bar{u}_s - \beta(k_x - 1)(\bar{x}_s - \mu_x)](\bar{x}_s + \bar{u}_s - \mu_x) \} \\
& = \frac{n^3}{(n-1)^2} \frac{[\sigma_X^2(\sigma_e^2 + \beta^2 \sigma_u^2) - \beta^2 \sigma_u^4]}{n^2} \rightarrow 0,
\end{aligned}$$

de modo que

$$\frac{\sqrt{n}}{n-1} \{n[\bar{e}_s - \beta k_x \bar{u}_s - \beta(k_x - 1)(\bar{x}_s - \mu_x)](\bar{x}_s + \bar{u}_s - \mu_x)\} \xrightarrow{P} 0. \quad (4.5)$$

De (4.4) e (4.5) fica provado (ii).

Para provarmos (i), basta verificar que

$$\sqrt{n}(\hat{\beta}_1 - k_x \beta) = k_x \sqrt{n}(\hat{\beta}_2 - \beta).$$

Comparando as distribuições no Teorema 4.1 vemos que a variância da distribuição assintótica de $\hat{\beta}_2$ obtido com k_x conhecido é a mesma de $\hat{\beta}_4$ obtido com $\delta = \sigma_e^2/\sigma_u^2$ conhecida e são menores que a variância da distribuição assintótica de $\hat{\beta}_3$ obtido com σ_u^2 conhecido.

A seguir vamos obter a distribuição assintótica do previsor $\hat{T}_R$ dado em (4.3). Adotaremos o conceito de convergência em populações finitas apresentado no Capítulo 2.

Teorema 4.2. *Sob o modelo (4.1) e considerando as convergências em distribuição dos estimadores $\hat{\beta}_1$, $\hat{\beta}_2$, $\hat{\beta}_3$ e $\hat{\beta}_4$ dados no Teorema 4.1, temos que quando $\nu \rightarrow \infty$*

(i) para $\hat{\beta} = \hat{\beta}_2, \hat{\beta}_3$ e $\hat{\beta}_4$,

$$\sqrt{n_\nu} \frac{(\hat{T}_{R_\nu} - T_\nu)}{N_\nu} \xrightarrow{D} N(0; (1-f)(\beta^2 \sigma_u^2 + \sigma_e^2)); \quad (4.6)$$

(ii) para $\hat{\beta} = \hat{\beta}_1$,

$$\sqrt{n_\nu} \frac{(\hat{T}_{R_\nu} - T)}{N_\nu} \xrightarrow{\mathcal{D}} N(0; (1-f)(k_x \beta^2 \sigma_u^2 + \sigma_e^2)) . \quad (4.7)$$

Prova. (i) Como

$$\begin{aligned} \hat{T}_R - T &= (N - n)[(\bar{X}_r - \bar{X}_s)\hat{\beta} + (\bar{y}_s - \bar{y}_r)] \\ &= (N - n)(\bar{X}_r - \bar{X}_s)(\hat{\beta} - \beta) + (N - n)[\bar{y}_s - \alpha - \beta\bar{X}_s - (\bar{y}_r - \alpha - \beta\bar{X}_r)] \end{aligned}$$

podemos escrever

$$\begin{aligned} \sqrt{n} \frac{(\hat{T}_R - T)}{N} &= \left(1 - \frac{n}{N}\right) (\bar{X}_r - \bar{X}_s) \sqrt{n} (\hat{\beta} - \beta) + \left(1 - \frac{n}{N}\right) \frac{1}{\sqrt{n}} \sum_{i \in s} (y_i - \alpha - \beta X_i) \\ &\quad - \sqrt{\frac{n}{N}} \sqrt{1 - \frac{n}{N}} \frac{1}{\sqrt{N - n}} \sum_{i \notin s} (y_i - \alpha - \beta X_i) . \end{aligned}$$

Mas

- $\bar{X}_r - \bar{X}_s = o_p(1)$, pois $\bar{X}_r \xrightarrow{P} \mu_x$ e $\bar{X}_s \xrightarrow{P} \mu_x$;
- $\frac{1}{\sqrt{n}} \sum_{i \in s} (y_i - \alpha - \beta X_i) = \frac{1}{\sqrt{n}} \sum_{i \in s} (e_i - \beta u_i) \xrightarrow{\mathcal{D}} N(0; \beta^2 \sigma_u^2 + \sigma_e^2)$;
- $\frac{1}{\sqrt{N - n}} \sum_{i \notin s} (y_i - \alpha - \beta X_i) = \frac{1}{\sqrt{N - n}} \sum_{i \notin s} (e_i - \beta u_i) \xrightarrow{\mathcal{D}} N(0; \beta^2 \sigma_u^2 + \sigma_e^2)$

e como por suposição $\sqrt{n}(\hat{\beta} - \beta) = O_p(1)$ concluímos que

$$\sqrt{n} \frac{(\hat{T}_R - T)}{N} \xrightarrow{\mathcal{D}} N(0; (1-f)(\beta^2 \sigma_u^2 + \sigma_e^2)) .$$

(ii) Como

$$\begin{aligned} \hat{T}_R - T &= (N - n)[(\bar{X}_r - \bar{X}_s)\hat{\beta}_1 + (\bar{y}_s - \bar{y}_r)] \\ &= (N - n)(\bar{X}_r - \bar{X}_s)(\hat{\beta}_1 - k_x \beta) + (N - n)[\bar{y}_s - \alpha - k_x \bar{X}_s \beta \\ &\quad - (\bar{y}_r - \alpha - k_x \bar{X}_r \beta)] \\ &= (N - n)(\bar{X}_r - \bar{X}_s)(\hat{\beta}_1 - k_x \beta) + (N - n)[\bar{e}_s - k_x \bar{u}_s \beta + (1 - k_x)(\bar{x}_s - \mu_x) \beta \\ &\quad - (\bar{e}_r - k_x \bar{u}_r \beta + (1 - k_x)(\bar{x}_r - \mu_x) \beta)] . \end{aligned}$$

Mas, $\bar{e}_s, \bar{\mu}_s, \bar{x}_s, \bar{u}_r, \bar{e}_r, \bar{x}_r$ são independentes, logo

$$\begin{aligned}\sqrt{n}[\bar{e}_s - k_x \bar{u}_s \beta + (1 - k_x)(\bar{x}_s - \mu_x)\beta] &\xrightarrow{D} N(0; \sigma_e^2 + k_x \beta^2 \sigma_u^2) \text{ e} \\ \sqrt{N-n}[\bar{e}_r - k_x \bar{u}_r \beta + (1 - k_x)(\bar{x}_r - \mu_x)\beta] &\xrightarrow{D} N(0; \sigma_e^2 + k_x \beta^2 \sigma_u^2)\end{aligned}$$

e como em (i) $\frac{(N-n)}{N}(\bar{X}_r - \bar{X}_s)\sqrt{n}(\hat{\beta}_1 - k_x \beta) \xrightarrow{P} 0$, fica provado (ii).

De (4.6) e (4.7) segue que a eficiência relativa assintótica de $\hat{T}_{R1}$ (isto é, o previsor regressão com $\hat{\beta} = \hat{\beta}_1$) em relação a $\hat{T}_{R2}$ (isto é, previsor regressão com $\hat{\beta} = \hat{\beta}_2, \hat{\beta}_3, \hat{\beta}_4$ é

$$e_{12} = \text{ARE}(\hat{T}_{R1}, \hat{T}_{R2}) = \frac{\sigma_e^2 + \beta^2 \sigma_u^2}{\sigma_e^2 + k_x \beta^2 \sigma_u^2} = \frac{\delta + \beta^2}{\delta + k_x \beta^2}$$

Podemos concluir que independentemente do estimador $\hat{\beta}_1$ ser viciado para β , o previsor $\hat{T}_{R1}$ é não viciado para T e é assintoticamente mais eficiente que os previsores do tipo regressão obtidos com $\hat{\beta} = \hat{\beta}_2, \hat{\beta}_3$ e $\hat{\beta}_4$.

Na Tabela 4.1 apresentamos alguns valores da eficiência assintótica relativa e_{12} para o caso em que $\beta = \delta = 1,0$ e vários valores de k_x .

Tabela 4.1. Valores de e_{12} com $\beta = \delta = 1,0$

k_x	0,1	0,3	0,5	0,7	0,9
e_{12}	1,82	1,54	1,33	1,18	1,05

Temos portanto que $\hat{T}_R$ com $\hat{\beta} = \hat{\beta}_1$ (estimador regressão) pode ser bem mais eficiente que os outros previsores.

Para o modelo (4.1) com $\alpha = 0$ temos que a classe de previsores considerada é da forma

$$\hat{T}_R = n\bar{y}_s + (N-n)\hat{\beta}\bar{X}_r \quad (4.8)$$

onde $\hat{\beta}$ é um estimador de β .

Consideremos inicialmente $\mu_x = 0$ e como estimadores de β os estimadores $\hat{\beta}_i$, considerados acima, mas com $S_X^2 = \sum_{i \in s} X_i^2/n$, $S_Y^2 = \sum_{i \in s} y_i^2/n$ e $S_{XY} = \sum_{i \in s} X_i y_i/n$. Como

no Teorema 4.1, o Teorema 4.3 apresenta as distribuições assintóticas dos estimadores de β .

Teorema 4.3. *Sob o modelo 4.1 com $\alpha = 0$ e $\mu_x = 0$ temos para $n \rightarrow \infty$,*

$$\begin{aligned}\sqrt{n}(\hat{\beta}_1 - k_x \beta) &\xrightarrow{\mathcal{D}} N(0, \sigma_X^{-4}[(\beta^2 \sigma_u^2 + \sigma_e^2) \sigma_X^2 - \beta^2 \sigma_u^4]) , \\ \sqrt{n}(\hat{\beta}_2 - \beta) &\xrightarrow{\mathcal{D}} N(0, \sigma_x^{-4}[(\beta^2 \sigma_u^2 + \sigma_e^2) \sigma_X^2 - \beta^2 \sigma_u^4]) , \\ \sqrt{n}(\hat{\beta}_3 - \beta) &\xrightarrow{\mathcal{D}} N(0, \sigma_x^{-4}[(\beta^2 \sigma_u^2 + \sigma_e^2) \sigma_X^2 + \beta^2 \sigma_u^4]) , \\ \sqrt{n}(\hat{\beta}_4 - \beta) &\xrightarrow{\mathcal{D}} N(0, \sigma_x^{-4}[(\beta^2 \sigma_u^2 + \sigma_e^2) \sigma_X^2 - \beta^2 \sigma_u^4]) .\end{aligned}$$

No Teorema 4.4 encontramos a distribuição assintótica do previsor $\hat{T}_R$ dado em (4.8).

Teorema 4.4. *Sob o modelo (4.1) com $\alpha = 0$, $\mu_x = 0$ e considerando as convergências em distribuição dos estimadores $\hat{\beta}_1$, $\hat{\beta}_2$, $\hat{\beta}_3$ e $\hat{\beta}_4$ dadas no Teorema 4.3 temos que quando $\nu \rightarrow \infty$*

(i) *para $\hat{\beta} = \hat{\beta}_2$, $\hat{\beta}_3$ e $\hat{\beta}_4$,*

$$\sqrt{n_\nu} \frac{(\hat{T}_{R_\nu} - T_\nu)}{N_\nu} \xrightarrow{\mathcal{D}} N(0, f(1-f)(\sigma_e^2 + \beta^2 \sigma_u^2)) ;$$

(ii) *para $\hat{\beta} = \hat{\beta}_1$,*

$$\sqrt{n_\nu} \frac{(\hat{T}_{R_\nu} - T_\nu)}{N_\nu} \xrightarrow{\mathcal{D}} N(0, f(1-f)(\sigma_e^2 + \beta^2 k_x \sigma_u^2)) .$$

A prova é similar à do Teorema 4.2. Portanto, a eficiência assintótica de $\hat{T}_{R1}$ com relação aos outros previsores é ainda como dada por e_{12} acima, ou seja $\hat{T}_{R1}$ é o previsor assintoticamente mais eficiente.

Por outro lado, no caso em que $\alpha = 0$ e μ_x é conhecido e não nulo, temos novamente o modelo (4.1) com equações

$$y_i = \alpha' + \beta x_i' + e_i \quad \text{e} \quad X_i' = x_i' + u_i ,$$

onde $\alpha' = \mu_x \beta$, $x'_i = x_i - \mu_x$ e $X'_i = X_i - \mu_x$, de modo que voltamos a um modelo com intercepto não nulo. Devemos portanto considerar o predictor $\hat{T}_R$ como definido em (4.3) com β_i , $i = 1, 2, 3, 4$ calculados com $S_X^2 = \sum_{i \in s} (X_i - \bar{X}_s)^2 / (n - 1)$, $S_{XY} = \sum_{i \in s} (X_i - \bar{X}_s)(y_i - \bar{y}_s) / (n - 1)$ e $S_Y^2 = \sum_{i \in s} (y_i - \bar{y}_s)^2 / (n - 1)$ e novamente as distribuições assintóticas de $\hat{T}_R$ são como dadas no Teorema 4.2, com $\hat{T}_{R1}$ sendo ainda o predictor assintoticamente mais eficiente. No caso em que μ_x é desconhecido e $\alpha = 0$, podemos considerar novamente o modelo (4.1) com um intercepto desconhecido, sendo neste caso, $\alpha' = \bar{X} \beta$, $x'_i = x_i - \bar{X}$ e $X'_i = X_i - \bar{X}$, onde $\bar{X} = \frac{1}{N} \sum_{i=1}^N X_i$, de modo que o resultado do Teorema 4.2 se aplica e novamente o predictor $\hat{T}_{R1}$ é assintoticamente mais eficiente que os outros previsores considerados acima.

4.3 O ENFOQUE CONDICIONAL

Nesta seção, ao invés de trabalharmos com a distribuição conjunta de (y_i, X_i) (descrita em (4.2)), adotaremos um outro enfoque onde trabalharemos com a distribuição condicional de y_i dado X_i . Como mencionado na seção 4.2, existem muitas situações onde a razão $k_x = \sigma_x^2 / \sigma_X^2$ é conhecida. Desta forma, podemos escrever $\sigma_u^2 = k \sigma_x^2$ onde $k = \frac{(1-k_x)}{k_x}$. Bolfarine e Cordani (1993) transformaram o vetor de parâmetros original $(\mu_x, \alpha, \beta, \sigma_x^2, \sigma_e^2)$ num novo vetor de parâmetros $(\beta, \lambda_0, \lambda_1, \lambda_2, \lambda_3)$ onde

$$\begin{aligned}\lambda_0 &= \mu_x, \\ \lambda_1 &= \alpha + \beta \mu_x, \\ \lambda_2 &= (k + 1) \sigma_e^2 + k \lambda_3 \beta^2, \\ \lambda_3 &= \sigma_x^2.\end{aligned}$$

Segundo Cox e Reid (1987) as inferências sobre β são simplificadas utilizando-se essa parametrização onde β é ortogonal aos parâmetros λ_0 , λ_1 , λ_2 e λ_3 .

Nessa nova parametrização temos que a distribuição conjunta de (y_i, X_i) é normal

bivariada com vetor de médias e matriz de covariâncias dados, respectivamente, por

$$\begin{pmatrix} \lambda_1 \\ \lambda_0 \end{pmatrix} ; \quad \begin{pmatrix} \frac{\lambda_2 + \lambda_3 \beta^2}{k+1} & \beta \lambda_3 \\ \beta \lambda_3 & \lambda_3(k+1) \end{pmatrix} .$$

Então, a distribuição condicional de y_i dado X_i é tal que

$$y_i | X_i \sim N \left(\alpha' + \beta Z_i; \frac{\lambda_2}{k+1} \right) \quad (4.9)$$

onde

$$\alpha' = \lambda_1 - \frac{\lambda_0 \beta}{k+1}$$

e

$$Z_i = \frac{X_i}{k+1} ,$$

$i = 1, \dots, n$. A partir dessa distribuição condicional obteremos a seguir o previsor de mínimos quadrados e o previsor Bayesiano para o total populacional.

4.3.1 O PREVISOR DE MÍNIMOS QUADRADOS CONDICIONAL

Considerando o modelo condicional (4.9) obtido depois da parametrização ortogonal, temos que o previsor de mínimos quadrados de T , supondo k_x conhecido, é

$$\hat{T}_c = N\bar{y}_s + (N - n)(\bar{Z}_r - \bar{Z}_s)\hat{\beta} \quad (4.10)$$

$$\text{onde } \hat{\beta} = \frac{\sum_{i \in s} (y_i - \bar{y}_s)(Z_i - \bar{Z}_s)}{\sum_{i \in s} (Z_i - \bar{Z}_s)^2} .$$

De acordo com Bolfarine (1990) temos que

$$E[\hat{T}_c - T/\mathbf{X}] = 0 ,$$

ou seja, $\hat{T}_c$ é não viciado e

$$\begin{aligned} E[(\hat{T}_c - T)^2/\mathbf{X}] &= \text{Var}[\hat{T}_c - T/\mathbf{X}] \\ &= N \frac{(1-f)}{f} \{ \sigma_e^2 + \beta^2 k_x \sigma_u^2 \} \left\{ 1 + n(1-f) \frac{(\bar{Z}_r - \bar{Z}_s)^2}{\sum_{i \in s} (Z_i - \bar{Z}_s)^2} \right\} . \end{aligned} \quad (4.11)$$

Notemos que a variância preditiva condicional (4.11) de $\hat{T}_c$ é mínima se tomarmos uma amostra satisfazendo

$$\overline{Z}_r = \overline{Z}_s . \quad (4.12)$$

Uma amostra satisfazendo essa condição é conhecida como amostra balanceada. Veja por exemplo Royall e Herson (1973) e Bolfarine et al. (1987). Se a condição (4.12) é satisfeita, ou seja, se a amostra for balanceada, $\hat{T}_c = N\overline{y}_s$ (independente de k_x). Notamos que a condição (4.12) é satisfeita se, e somente se, $\overline{X}_r = \overline{X}_s$, ou seja, se a amostra for balanceada com respeito à variável X , eliminando a necessidade do conhecimento de k_x .

No caso do modelo (4.1) com $\alpha = 0$ e μ_x conhecido e considerando a razão k_x conhecida, o novo vetor de parâmetros obtido por Bolfarine e Cordani (1993), após a transformação ortogonal, é $(\beta, \lambda_0, \lambda_1)$, onde

$$\lambda_0 = (k + 1)\sigma_e^2 + k\lambda_1\beta^2$$

e

$$\lambda_1 = \sigma_x^2 .$$

A distribuição condicional de y_i dado X_i nessa nova parametrização é tal que

$$y_i \mid X_i \sim N(\beta Z_i, k_x \lambda_0) , \quad (4.13)$$

onde $Z_i = \mu_x + k_x(X_i - \mu_x)$.

Logo, sendo μ_x e k_x conhecidos, o previsor de mínimos quadrados condicional é dado por

$$\hat{T}_c = n\overline{y}_s + (N - n)\hat{\beta}\overline{Z}_r \quad (4.14)$$

onde $\hat{\beta} = \frac{\sum_{i \in s} Z_i y_i}{\sum_{i \in s} Z_i^2}$.

Podemos verificar que esse previsor é não viciado sob o modelo condicional (4.13) e

$$\text{Var} [\hat{T}_c - T/\mathbf{X}] = k_x \lambda_0 (N - n) \left[1 + (N - n) \frac{\overline{Z}_r^2}{\sum_{i \in s} Z_i^2} \right] . \quad (4.15)$$

Portanto, uma amostra maximizando $\sum_{i \in s} Z_i^2$ é ótima, pois, nesse caso, a variância preditiva condicional (4.15) é mínima.

4.3.2 O PREVISOR BAYESIANO

Nesta seção vamos obter o previsor Bayesiano de T considerando uma priori não informativa. De acordo com Bolfarine et al. (1987), o previsor de Bayes do total populacional T sob o modelo condicional (4.9) usando perda quadrática é dado por

$$\hat{T}_B = E[T/\mathbf{y}_s, \mathbf{X}_s] = n\bar{y}_s + (N - n)\hat{\alpha}' + \hat{\beta} \sum_{i \notin s} Z_i \quad (4.16)$$

onde $\hat{\alpha}'$ e $\hat{\beta}$ são médias das distribuições marginais a posteriori de α' e β , respectivamente.

E a variância a posteriori de T é

$$\begin{aligned} \text{Var}[T/\mathbf{y}_s, \mathbf{X}] &= E\{\text{Var}[T/\mathbf{y}_s, \alpha', \beta, \lambda_2]/\mathbf{y}_s, \mathbf{X}\} + \text{Var}\{E[T/\mathbf{y}_s, \alpha', \beta, \lambda_2]/\mathbf{y}_s, \mathbf{X}\} \\ &= (N - n)k_x E[\lambda_2/\mathbf{y}_s, \mathbf{X}] + (N - n)^2 \text{Var}[\alpha'/\mathbf{y}_s, \mathbf{X}] \\ &\quad + \text{Var}[\beta/\mathbf{y}_s, \mathbf{X}] \left(\sum_{i \notin s} Z_i \right)^2 + 2(N - n) \sum_{i \notin s} Z_i \text{Cov}[\alpha', \beta/\mathbf{y}_s, \mathbf{X}]. \end{aligned} \quad (4.17)$$

Considerando a priori não informativa

$$p(\alpha', \beta, \lambda_2) \propto \frac{1}{\lambda_2}$$

temos que a densidade marginal a posteriori de β é

$$\begin{aligned} p(\beta/\mathbf{y}_s, \mathbf{X}_s) &\propto \int_0^\infty \int_{-\infty}^\infty \left(\frac{k+1}{\lambda_2} \right)^{n/2} \exp \left\{ -\frac{k+1}{2\lambda_2} \sum_{i \in s} (y_i - \alpha' - \beta Z_i)^2 \right\} \frac{1}{\lambda_2} d\alpha' d\lambda_2 \\ &\propto \int_0^\infty \left[\frac{1}{\lambda_2^{\frac{n+1}{2}}} \exp \left\{ -\frac{k+1}{2\lambda_2} \left[\sum_{i \in s} y_i^2 + \beta^2 \sum_{i \in s} Z_i^2 - 2\beta \sum_{i \in s} y_i Z_i \right] \right\} \right. \\ &\quad \left. \int_{-\infty}^\infty \exp \left\{ -\frac{n(k+1)}{2\lambda_2} [\lambda_1^2 + 2\beta\alpha'\bar{Z} - 2\alpha'\bar{y}] \right\} d\alpha' \right] d\lambda_2 \\ &\propto \int_0^\infty \frac{1}{\lambda_2^{\frac{n+1}{2}}} \exp \left\{ -\frac{k+1}{2\lambda_2} \left[\sum_{i \in s} y_i^2 + \beta^2 \sum_{i \in s} Z_i^2 - 2\beta \sum_{i \in s} y_i Z_i \right] \right\} \end{aligned}$$

$$\begin{aligned}
& \exp \left\{ \frac{n(k+1)}{2\lambda_2} (\bar{y} - \beta \bar{Z})^2 \right\} d\lambda_2 \\
& \propto \int_0^\infty \frac{1}{\lambda_2^{\frac{n+1}{2}}} \exp \left\{ -\frac{k+1}{2\lambda_2} \sum_{i \in s} [(y_i - \bar{y}) - \beta(Z_i - \bar{Z})]^2 \right\} d\lambda_2 \\
& \propto \left\{ \sum_{i \in s} [(y_i - \bar{y}) - \beta(Z_i - \bar{Z})]^2 \right\}^{-\left(\frac{n-1}{2}\right)} \quad (4.18)
\end{aligned}$$

Depois de algumas manipulações algébricas, podemos escrever a densidade marginal a posteriori (4.18) como

$$p(\beta | \mathbf{y}_s, \mathbf{X}) \propto \left\{ 1 + \frac{(\beta - \hat{\beta})^2 \sum_{i \in s} (Z_i - \bar{Z})^2}{(n-2)k_x \hat{\lambda}_2} \right\}^{-\left(\frac{n-1}{2}\right)} \quad (4.19)$$

onde

$$\hat{\beta} = \frac{\sum_{i \in s} (y_i - \bar{y})(Z_i - \bar{Z})}{\sum_{i \in s} (Z_i - \bar{Z})^2}$$

e

$$\hat{\lambda}_2 = \frac{\sum_{i \in s} [(y_i - \bar{y}) - \hat{\beta}(Z_i - \bar{Z})]^2}{(n-2)k_x}.$$

De acordo com (4.19) vemos que

$$\frac{(\beta - \hat{\beta}) \left[\sum_{i \in s} (Z_i - \bar{Z})^2 \right]^{1/2}}{(k_x \hat{\lambda}_2)^{1/2}}$$

segue uma distribuição t com $n-2$ graus de liberdade e portanto

$$E[\beta | \mathbf{y}_s, \mathbf{X}] = \hat{\beta}$$

e

$$\text{Var}[\beta | \mathbf{y}_s, \mathbf{X}] = \frac{(n-2)}{(n-4)} \frac{k_x \hat{\lambda}_2}{\sum_{i \in s} (Z_i - \bar{Z})^2}.$$

Além disso, verificamos que

$$E[\alpha' | \mathbf{y}_s, \mathbf{X}] = \bar{y}_s - \hat{\beta} \bar{Z}_s, \quad (4.20)$$

$$\text{Var}[\alpha' | \mathbf{y}_s, \mathbf{X}] = \left(\frac{n-2}{n-4} \right) k_x \hat{\lambda}_2 \left(\frac{\bar{Z}_s^2}{\sum_{i \in s} (Z_i - \bar{Z}_s)^2} + \frac{1}{n} \right), \quad (4.21)$$

$$E[\lambda_2 \mid \mathbf{y}_s, \mathbf{X}] = \frac{n-2}{n-4} \hat{\lambda}_2 \quad \text{e} \quad (4.22)$$

$$\text{Cov}[\alpha', \beta \mid \mathbf{y}_s, \mathbf{X}] = - \left(\frac{n-2}{n-4} \right) \frac{k_x \hat{\lambda}_2 \bar{Z}_s}{\sum_{i \in s} (Z_i - \bar{Z}_s)^2} . \quad (4.23)$$

De (4.16), (4.17), (4.20) - (4.23) concluímos que o previsor Bayesiano de T é

$$\hat{T}_B = N \bar{y}_s + (N - n) \hat{\beta} [\bar{Z}_r - \bar{Z}_s] \quad (4.24)$$

e a variância a posteriori de T é

$$\text{Var}[T \mid \mathbf{y}_s, \mathbf{X}] = \left(\frac{n-2}{n-4} \right) k_x \hat{\lambda}_2 \left[\frac{N(N-n)}{n} + (N-n)^2 \frac{(\bar{Z}_r - \bar{Z}_s)^2}{\sum_{i \in s} (Z_i - \bar{Z}_s)^2} \right] . \quad (4.25)$$

Podemos observar que o previsor Bayesiano $\hat{T}_B$ dado em (4.24) coincide com o previsor $\hat{T}_c$ dado em (4.10). Além disso, através da variância a posteriori de T (4.25) vemos que a amostra balanceada é ótima.

Em (4.13) temos a distribuição condicional de y_i dado X_i após a transformação ortogonal sob o modelo (4.1) com $\alpha = 0$ e supondo μ_x e k_x conhecidos. Considerando a priori não informativa

$$p(\beta, \lambda_0) \propto \frac{1}{\lambda_0} ,$$

Bolfarine e Cordani (1993) mostraram que a densidade a posteriori de β é

$$p(\beta \mid \mathbf{y}_s, \mathbf{Z}_s) \propto \left(1 + \frac{(\beta - \hat{\beta})^2 \sum_{i \in s} Z_i^2}{(n-1) k_x \hat{\lambda}_0} \right)^{-n/2}$$

onde

$$\hat{\beta} = \frac{\sum_{i \in s} y_i Z_i}{\sum_{i \in s} Z_i^2} \quad \text{e} \quad \hat{\lambda}_0 = \frac{\sum_{i=1}^n (y_i - \hat{\beta} Z_i)^2}{(n-1) k_x}$$

logo o previsor de Bayes de T é dado por

$$\hat{T}_B = n \bar{y}_s + \hat{\beta} \sum_{i \in s} Z_i ,$$

sendo $\hat{\beta}$ a média a posteriori. Nesse caso também, o previsor coincide com o previsor de mínimos quadrados dado em (4.14). Além disso, a variância a posteriori de T é

$$\text{Var}[T \mid \mathbf{y}_s, \mathbf{Z}_s] = \frac{n-1}{n-3} k_x \hat{\lambda}_0 \left\{ (N-n) + \frac{(\sum_{i \in s} Z_i)^2}{\sum_{i \in s} Z_i^2} \right\}$$

e, portanto, é minimizada, se $\sum_{i \in s} Z_i^2$ for máxima.

4.4 O MODELO COM VÁRIAS VARIÁVEIS AUXILIARES

Nesta seção vamos generalizar o modelo (4.1). Consideraremos um modelo contendo um vetor de variáveis auxiliares.

Seja $\mathcal{P} = \{1, \dots, N\}$ uma população finita de tamanho N . Associadas com a k -ésima unidade de $\mathcal{P}$ existem $p+1$ quantidades desconhecidas $y_k, x_{k_1}, \dots, x_{k_p}$. Relacionando essas variáveis consideramos uma generalização do modelo de regressão com erro nas variáveis dado em (4.1), isto é,

$$y_k = \alpha + \mathbf{x}'_k \boldsymbol{\beta} + e_k \quad (4.26)$$

mas $\mathbf{x}'_k$ não é observado diretamente de modo que

$$X_{kj} = x_{kj} + u_{kj} ,$$

para $k = 1, \dots, N$ e $j = 1, \dots, p$, com a suposição de que

$$\begin{pmatrix} e_k \\ \mathbf{x}_k \\ \mathbf{u}_k \end{pmatrix} \sim N \left(\begin{pmatrix} 0 \\ \boldsymbol{\mu}_x \\ \mathbf{0} \end{pmatrix} ; \begin{pmatrix} \sigma_e^2 & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \boldsymbol{\Sigma}_{xx} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \boldsymbol{\Sigma}_{uu} \end{pmatrix} \right) ,$$

onde $\boldsymbol{\Sigma}_{xx}$ e $\boldsymbol{\Sigma}_{uu}$ são matrizes de dimensão $p \times p$.

Considerando a matriz $\boldsymbol{\Sigma}_{uu}$ conhecida, Fuller (1987) obteve estimadores de máxima verossimilhança de α e do vetor $\boldsymbol{\beta}$ dados por

$$\hat{\alpha} = \bar{y}_s - \bar{\mathbf{X}}'_s \hat{\boldsymbol{\beta}} \quad \text{e} \quad \hat{\boldsymbol{\beta}} = (\mathbf{M}_{XX} - \boldsymbol{\Sigma}_{uu})^{-1} \mathbf{M}_{XY}$$

onde

$$\mathbf{M}_{XX} = \frac{1}{n-1} \sum_{k \in s} (\mathbf{X}_k - \bar{\mathbf{X}}_s)(\mathbf{X}_k - \bar{\mathbf{X}}_s)'$$

e

$$\mathbf{M}_{XY} = \frac{1}{n-1} \sum_{k \in s} (\mathbf{X}_k - \bar{\mathbf{X}}_s)(y_k - \bar{y}_s) ,$$

com $\mathbf{X}'_k = (X_{k_1}, \dots, X_{k_p})$, $k \in s$, $\bar{\mathbf{X}}_s = \sum_{k \in s} \mathbf{X}_k / n$.

Observando o estimador de β , vemos que é uma extensão do estimador $\hat{\beta}_2$ dado na seção 4.2. No teorema a seguir temos a distribuição assintótica do estimador $\hat{\beta}$.

Teorema 4.5. *Sob o modelo (4.26) temos que*

$$\sqrt{n}(\hat{\beta} - \beta) \xrightarrow{D} N(0, \mathbf{G})$$

quando $n \rightarrow \infty$, onde

$$\mathbf{G} = \Sigma_{xx}^{-1} \left[(\beta' \Sigma_{uu} \beta + \sigma_e^2) (\Sigma_{xx} + \Sigma_{uu} + \Sigma_{uu} \beta \beta' \Sigma_{uu}) \right] \Sigma_{xx}^{-1}.$$

Prova. Utilizando o fato que $\mathbf{X}_k = \mathbf{x}_k + \mathbf{u}_k$, podemos escrever $y_k = \alpha + \mathbf{X}'_k \beta + v_k$ onde $v_k = e_k - \mathbf{u}'_k \beta$. Logo,

$$\begin{aligned} \hat{\beta} &= (\mathbf{M}_{XX} - \Sigma_{uu})^{-1} (\mathbf{M}_{XX} \beta + \mathbf{M}_{Xv} - \Sigma_{uu} \beta + \Sigma_{uu} \beta) \\ &= \beta + (\mathbf{M}_{XX} - \Sigma_{uu})^{-1} (\mathbf{M}_{Xv} + \Sigma_{uu} \beta) \end{aligned}$$

onde

$$\mathbf{M}_{Xv} = \frac{1}{n-1} \sum_{k \in s} (\mathbf{X}_k - \bar{\mathbf{X}}_s)(v_k - \bar{v}_s).$$

Temos também que

- $\mathbf{M}_{XX} = \frac{1}{n-1} \sum_{k \in s} [(\mathbf{x}_k - \bar{\mathbf{x}}_s)(\mathbf{x}_k - \bar{\mathbf{x}}_s)' + (\mathbf{x}_k - \bar{\mathbf{x}}_s)(\mathbf{u}_k - \bar{\mathbf{u}}_s)' + (\mathbf{u}_k - \bar{\mathbf{u}}_s)(\mathbf{x}_k - \bar{\mathbf{x}}_s)' + (\mathbf{u}_k - \bar{\mathbf{u}}_s)(\mathbf{u}_k - \bar{\mathbf{u}}_s)']$,
- $\frac{1}{n-1} \sum_{k \in s} (\mathbf{x}_k - \bar{\mathbf{x}}_s)(\mathbf{x}_k - \bar{\mathbf{x}}_s)' = \Sigma_{xx} + O_p(n^{-1/2})$,
- $\frac{1}{n-1} \sum_{k \in s} (\mathbf{x}_k - \bar{\mathbf{x}}_s)(\mathbf{u}_k - \bar{\mathbf{u}}_s)' = O_p(n^{-1/2})$ e
- $\frac{1}{n-1} \sum_{k \in s} (\mathbf{u}_k - \bar{\mathbf{u}}_s)(\mathbf{u}_k - \bar{\mathbf{u}}_s)' = \Sigma_{uu} + O_p(n^{-1/2})$.

Logo

$$\mathbf{M}_{XX} - \Sigma_{uu} = \Sigma_{xx} + O_p(n^{-1/2})$$

e como $\mathbf{M}_{Xv} + \Sigma_{uu}\boldsymbol{\beta} = O_p(n^{-1/2})$, temos que

$$\hat{\boldsymbol{\beta}} - \boldsymbol{\beta} = (\Sigma_{xx})^{-1}(\mathbf{M}_{Xv} + \Sigma_{uu}\boldsymbol{\beta}) + O_p(n^{-1}) . \quad (4.27)$$

Consideremos a combinação linear

$$\begin{aligned} \sqrt{n}\boldsymbol{\delta}'(\mathbf{M}_{Xv} + \Sigma_{uu}\boldsymbol{\beta}) &= n^{-1/2} \sum_{k=1}^n \sum_{j=1}^p \delta_j [(x_{kj} - \mu_{xj})v_k + u_{kj}v_k + \beta_j\sigma_{u_j}^2] + O_p(n^{-1}) \\ &= n^{-1/2} \sum_{k=1}^n g_k + O_p(n^{-1}) \end{aligned}$$

onde $\sigma_{u_j}^2$ é o elemento de índice $j \times j$ da matriz Σ_{uu} , $\boldsymbol{\delta}' = (\delta_1, \dots, \delta_p)$ é um vetor arbitrário tal que $\boldsymbol{\delta}'\boldsymbol{\delta} \neq 0$.

As variáveis aleatórias,

$$g_k = \sum_{j=1}^p \delta_j [(x_{kj} - \mu_{xj})v_k + u_{kj}v_k + \beta_j\sigma_{u_j}^2]$$

são independentes com média zero e variância

$$E[g_k^2] = \boldsymbol{\delta}'[(\Sigma_{xx} + \Sigma_{uu})\sigma_v^2 + \Sigma_{uu}\boldsymbol{\beta}\boldsymbol{\beta}'\Sigma_{uu}]\boldsymbol{\delta} ,$$

onde $\sigma_v^2 = \sigma_e^2 + \boldsymbol{\beta}'\Sigma_{uu}\boldsymbol{\beta}$.

Logo

$$n^{-1/2} \sum_{k=1}^n g_k \xrightarrow{\mathcal{D}} N(0; \boldsymbol{\delta}'[(\Sigma_{xx} + \Sigma_{uu})(\sigma_e^2 + \boldsymbol{\beta}'\Sigma_{uu}\boldsymbol{\beta}) + \Sigma_{uu}\boldsymbol{\beta}\boldsymbol{\beta}'\Sigma_{uu}]\boldsymbol{\delta}) .$$

Como $\boldsymbol{\delta}$ é arbitrário e $\boldsymbol{\delta} \neq \mathbf{0}$,

$$\sqrt{n}(\mathbf{M}_{Xv} + \Sigma_{uu}\boldsymbol{\beta}) \xrightarrow{\mathcal{D}} N(0; (\Sigma_{xx} + \Sigma_{uu})(\sigma_e^2 + \boldsymbol{\beta}'\Sigma_{uu}\boldsymbol{\beta}) + \Sigma_{uu}\boldsymbol{\beta}\boldsymbol{\beta}'\Sigma_{uu}) .$$

Utilizando esse resultado em (4.27), concluimos a prova.

Sob o modelo (4.26) o previsor do tipo regressão para T é dado por

$$\hat{T}_R = N\bar{y}_s + (N - n) \sum_{j=1}^p (\bar{X}_{r_j} - \bar{X}_{s_j})\hat{\beta}_j , \quad (4.28)$$

com $\bar{X}_{r_j} = \sum_{k \in s} X_{k_j} / N - n$ e $\bar{X}_{s_j} = \sum_{k \in s} X_{k_j} / n$.

No teorema a seguir temos a distribuição assintótica do previsor $\hat{T}_R$.

Teorema 4.6. *Sob o modelo (4.26) e considerando o previsor $\hat{T}_R$ dado em (4.28) temos*

$$\sqrt{n_\nu} \frac{(\hat{T}_{R_\nu} - T)}{N_\nu} \xrightarrow{\mathcal{D}} N(0; (1-f)(\sigma_e^2 + \boldsymbol{\beta}' \boldsymbol{\Sigma}_{uu} \boldsymbol{\beta}))$$

quando $\nu \rightarrow \infty$.

Prova. Podemos escrever

$$\begin{aligned} \sqrt{n} \frac{(\hat{T}_R - T)}{N} &= \left(1 - \frac{n}{N}\right) (\bar{\mathbf{X}}_r - \bar{\mathbf{X}}_s)' \sqrt{n} (\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}) \\ &\quad + \left(1 - \frac{n}{N}\right) \sqrt{n} (\bar{y}_s - \alpha - \bar{\mathbf{X}}_s' \boldsymbol{\beta}) - \sqrt{\frac{n}{N}} \sqrt{1 - \frac{n}{N}} \sqrt{N - n} (\bar{y}_r - \alpha - \bar{\mathbf{X}}_r' \boldsymbol{\beta}) . \end{aligned}$$

Mas

- $(\bar{\mathbf{X}}_r - \bar{\mathbf{X}}_s) \xrightarrow{P} 0$ pois $\bar{\mathbf{X}}_r \xrightarrow{P} \boldsymbol{\mu}_x$ e $\bar{\mathbf{X}}_s \xrightarrow{P} \boldsymbol{\mu}_x$
- $\sqrt{n}(\hat{\boldsymbol{\beta}} - \boldsymbol{\beta}) = O_p(1)$ pelo Teorema 4
- $\sqrt{n}(\bar{y}_s - \alpha - \bar{\mathbf{X}}_s' \boldsymbol{\beta}) = \sqrt{n}(\bar{e}_s - \bar{\mathbf{u}}_s' \boldsymbol{\beta}) \xrightarrow{\mathcal{D}} N(0, \sigma_e^2 + \boldsymbol{\beta}' \boldsymbol{\Sigma}_{uu} \boldsymbol{\beta})$
- $\sqrt{N - n}(\bar{y}_r - \alpha - \bar{\mathbf{X}}_r' \boldsymbol{\beta}) = \sqrt{N - n}(\bar{e}_r - \bar{\mathbf{u}}_r' \boldsymbol{\beta}) \xrightarrow{\mathcal{D}} N(0, \sigma_e^2 + \boldsymbol{\beta}' \boldsymbol{\Sigma}_{uu} \boldsymbol{\beta})$.

Concluimos, então, que

$$\sqrt{n} \frac{(\hat{T}_R - T)}{N} \xrightarrow{\mathcal{D}} N(0; (1-f)(\sigma_e^2 + \boldsymbol{\beta}' \boldsymbol{\Sigma}_{uu} \boldsymbol{\beta})) .$$

Considerando o modelo sem intercepto, $\alpha = 0$, temos que o previsor para T é dado por

$$\hat{T}_R = n \bar{y}_s + \mathbf{1}_r' X_r \hat{\boldsymbol{\beta}}$$

onde

$$\hat{\boldsymbol{\beta}} = (\mathbf{M}_{XX} - \boldsymbol{\Sigma}_{uu})^{-1} \mathbf{M}_{XY}$$

com $\mathbf{M}_{XX} = \frac{1}{n} \sum_{k \in s} \mathbf{X}_k \mathbf{X}_k'$ e $\mathbf{M}_{XY} = \frac{1}{n} \sum_{k \in s} \mathbf{X}_k y_k$.

Nesse caso, a distribuição assintótica de $\hat{\beta}$ é tal que

$$\sqrt{n}(\hat{\beta} - \beta) \xrightarrow{\mathcal{D}} N(\mathbf{0}, \mathbf{G})$$

onde

$$\mathbf{G} = (\Sigma_{xx} + \mu_x \mu_x')^{-1} [(\beta' \Sigma_{uu} \beta + \sigma_e^2)(\Sigma_{xx} + \Sigma_{uu} + \mu_x \mu_x') + \Sigma_{uu} \beta \beta' \Sigma_{uu}] (\Sigma_{xx} + \mu_x \mu_x')^{-1}$$

e no teorema a seguir temos a distribuição assintótica do previsor $\hat{T}_R$.

Teorema 4.7. *Considerando o modelo (4.26) com $\alpha = 0$ temos*

$$\sqrt{n_\nu} \frac{(\hat{T}_{R_\nu} - T)}{N_\nu} \xrightarrow{\mathcal{D}} N(0; (1-f)^2 \mu_x' \mathbf{G} \mu_x + f(1-f)(\sigma_e^2 + \beta' \Sigma_{uu} \beta)) ,$$

quando $\nu \rightarrow \infty$.

Prova. Podemos escrever

$$\begin{aligned} \sqrt{n} \frac{(\hat{T}_R - T)}{N} &= \left(1 - \frac{n}{N}\right) \frac{\mathbf{1}_r' \mathbf{X}_r}{N-n} \sqrt{n}(\hat{\beta} - \beta) - \sqrt{\frac{n}{N}} \sqrt{1 - \frac{n}{N}} \mathbf{1}_r' \frac{(\mathbf{y}_r - \mathbf{X}_r \beta)}{\sqrt{N-n}} \\ &= \left(1 - \frac{n}{N}\right) \mu_x \sqrt{n}(\hat{\beta} - \beta) - \sqrt{\frac{n}{N}} \sqrt{1 - \frac{n}{N}} \mathbf{1}_r' \frac{(\mathbf{y}_r - \mathbf{X}_r \beta)}{\sqrt{N-n}} + O_p(n^{-1/2}) . \end{aligned}$$

Mas

$$\mathbf{1}_r' \frac{(\mathbf{y}_r - \mathbf{X}_r \beta)}{\sqrt{N-n}} = \frac{1}{\sqrt{N-n}} (\mathbf{1}_r' \mathbf{e}_r - \mathbf{1}_r' \mathbf{u}_r \beta) ,$$

e como

$$\begin{aligned} \frac{\mathbf{1}_r' \mathbf{e}_r}{\sqrt{N-n}} &\xrightarrow{\mathcal{D}} N(0, \sigma_e^2) , \\ \frac{\mathbf{1}_r' \mathbf{u}_r \beta}{\sqrt{N-n}} &\xrightarrow{\mathcal{D}} N(0, \beta' \Sigma_{uu} \beta) \quad \text{e} \\ \mu_x \sqrt{n}(\hat{\beta} - \beta) &\xrightarrow{\mathcal{D}} N(0, \mu_x' \mathbf{G} \mu_x) \end{aligned}$$

segue o resultado.

Como na seção 2, podemos observar que a variância da distribuição assintótica do previsor do tipo regressão para o total independe da variância assintótica do estimador $\hat{\beta}$ para o modelo com intercepto, o que não ocorre para o modelo sem intercepto.

CONCLUSÕES

Essa tese pode ser dividida em duas partes: a primeira consiste na previsão da função distribuição populacional, e a segunda, na previsão do total e da variância populacionais sob modelos com erros nas variáveis.

Com relação à primeira parte, a literatura recente tem dado atenção a esse problema tanto do ponto de vista da amostragem clássica, como também sob o enfoque de superpopulação. Considerando, então, modelos de superpopulação Gaussianos, encontramos o previsor ótimo, e para alguns casos importantes derivamos uma aproximação assintótica para a variância destes previsores. Os resultados indicam que o previsor ótimo e também o previsor proposto por Chambers e Dunstan (1986), ambos baseados em modelos, podem ser consideravelmente mais eficientes que os estimadores baseados no planejamento, se o modelo normal for adequado. Como todos os resultados obtidos dependem da validade dos modelos adotados, seria de interesse em trabalhos futuros estudar o comportamento dos previsores ótimos, quando o modelo apresentar falhas em alguma de suas suposições.

Na segunda parte, consideramos modelos com erros nas variáveis, pois em muitas

situações não é realista supor que as variáveis são medidas sem erro. Inicialmente, encontramos os previsores Bayesianos e Bayesianos empíricos para o total e a variância populacionais sob o modelo de locação com erro de mensuração. Mostramos que os previsores Bayesianos empíricos obtidos são assintoticamente ótimos. Como sugestão para um trabalho futuro propomos um estudo semelhante ao que foi feito no Capítulo 3, adotando um modelo de regressão incluindo informação de variáveis auxiliares. Para finalizar, considerando o modelo de regressão linear estrutural, estudamos o comportamento dos previsores do tipo regressão para o total populacional e, sob o enfoque condicional, encontramos o previsor de mínimos quadrados e o previsor Bayesiano. Considerando uma única variável auxiliar, concluímos que, apesar do estimador de mínimos quadrados ordinários ser viciado para β , o respectivo previsor regressão é assintoticamente não viciado para o total populacional e pode ser assintoticamente bem mais eficiente que os outros previsores propostos. Sob o enfoque condicional concluímos que o previsor de mínimos quadrados e o Bayesiano coincidem, e que para o modelo com intercepto a amostra balanceada é ótima. Dando continuidade ao que foi feito no Capítulo 4, seria de interesse para os modelos de superpopulação com erros nas variáveis contendo várias variáveis regressoras, considerar outros previsores para o total populacional e estudar seus comportamentos assintóticos. Dentre estes pretendemos considerar previsores baseados em estimadores de mínimos quadrados, quando várias variáveis regressoras estão presentes no modelo.

APÊNDICE 1

Neste apêndice apresentamos alguns resultados assintóticos utilizados na prova do Teorema 2.7. Consideremos o modelo especificado por (2.1) e (2.20) com erros normalmente distribuídos, ou seja,

$$y_i = \beta + e_i$$

onde $e_i \sim N(0, \sigma^2)$, $i = 1, \dots, N$.

Sendo $\Phi(\cdot)$ e $T_{n-2}(\cdot)$ as funções de distribuição da distribuição normal com média zero e variância 1 e da t -Student com $(n - 2)$ graus de liberdade, respectivamente, podemos mostrar que

Teorema A.1.1. $T_{n-2}(U_n(t)) = \Phi(U_n(t)) + O_p(n^{-1})$ onde $U_n(t) = \frac{\sqrt{n-2}(t - \bar{y}_s)}{\left[\frac{(n-1)^2}{n}s_y^2 - (t - \bar{y}_s)^2\right]^{1/2}}$.

Prova. Seja $p(t)$ a função densidade da distribuição t -Student com $(n - 2)$ graus de

liberdade. Temos que

$$\log p(U_n(t)) = \log \frac{\Gamma\left(\frac{n-1}{2}\right)}{\sqrt{(n-2)\pi}\Gamma\left(\frac{n-2}{2}\right)} - \frac{n-1}{2} \log \left(1 + \frac{U_n^2(t)}{n-2}\right).$$

É fácil verificar que $\frac{nU_n^2(t)}{n-2} \xrightarrow{P} \frac{(t-\beta)^2}{\sigma^2}$, logo $\frac{nU_n^2(t)}{n-2} = O_p(1)$ e então $\frac{U_n^2(t)}{n-2} = O_p(n^{-1})$.

Pelo Corolário 5.1.5, Fuller (1976), temos que

$$\log \left(1 + \frac{U_n^2(t)}{n-2}\right) = \frac{U_n^2(t)}{n-2} + O_p(n^{-2}),$$

de modo que

$$\frac{n-1}{2} \log \left(1 + \frac{U_n^2(t)}{n-2}\right) = \frac{U_n^2(t)}{2} + O_p(n^{-1}).$$

Utilizando-se a fórmula de Stirling: $\Gamma(z) = e^{-z} z^{z-1/2} (2\pi)^{1/2} (1 + O(z^{-1}))$, temos

$$\log \frac{\Gamma\left(\frac{n-1}{2}\right)}{\sqrt{(n-2)\pi}\Gamma\left(\frac{n-2}{2}\right)} = \log \left[e^{-1/2} \left(1 + \frac{1/2}{(n-2)/2}\right)^{(n-2)/2} (2\pi)^{-1/2} \frac{(1 + O(n^{-1}))}{(1 + O(n^{-1}))} \right] \quad (A.1)$$

Mas

$$(i) \log \left(1 + \frac{x}{n}\right) = \frac{x}{n} - \frac{x^2}{2n^2} + o(n^{-2}), \text{ logo}$$

$$e^{n \log(1 + \frac{x}{n})} = \left(1 + \frac{x}{n}\right)^n = e^{x + o(n^{-1})};$$

$$(ii) e^x = 1 + x + o(|x|), x \rightarrow 0, \text{ logo}$$

$$e^{O(n^{-1})} = 1 + O(n^{-1});$$

$$(iii) \frac{1}{1+x} = 1 - x + o(|x|), x \rightarrow 0, \text{ logo}$$

$$\frac{1}{1 + O(n^{-1})} = 1 + O(n^{-1});$$

$$(iv) \log(1+x) = x + o(|x|), x \rightarrow 0, \text{ logo}$$

$$\log(1 + O(n^{-1})) = O(n^{-1}).$$

Utilizando-se (i), (ii), (iii) e (iv) em (A.1), temos

$$\log \frac{\Gamma\left(\frac{n-1}{2}\right)}{\sqrt{(n-2)\pi}\Gamma\left(\frac{n-2}{2}\right)} = -\frac{1}{2} \log 2\pi + O(n^{-1}) .$$

Portanto

$$\log p(U_n(t)) = -\frac{1}{2} \log 2\pi - \frac{U_n^2(t)}{2} + O_p(n^{-1})$$

e então

$$p(U_n(t)) = \phi(U_n(t))[1 + O_p(n^{-1})]$$

onde $\phi(\cdot)$ é a função densidade da $N(0, 1)$. Integrando a equação acima,

$$T_{n-2}(U_n(t)) = \Phi(U_n(t))[1 + O_p(n^{-1})] .$$

Teorema A.1.2. *Seja $U_n(t)$ como definido no Teorema A.1.1. Então*

$$\sqrt{n} \left[U_n(t) - \frac{t - \beta}{\sigma} \right] \xrightarrow{\mathcal{D}} N \left(0, 1 + \frac{(t - \beta)^2}{2\sigma^2} \right) .$$

Prova. Seja $\mathbf{T}_n = \left(\bar{y}_s, \frac{(n-1)^2}{n(n-2)} s_y^2 - \frac{t - \bar{y}_s}{n-2} \right)$ e $\boldsymbol{\mu} = (\beta, \sigma^2)$, temos então $\sqrt{n}(\mathbf{T}_n - \boldsymbol{\mu}) \xrightarrow{\mathcal{D}} N \left(\mathbf{0}, \begin{bmatrix} \sigma^2 & 0 \\ 0 & 2\sigma^4 \end{bmatrix} \right)$, que pode ser mostrado por Cramer-Wold (Serfling (1980)), ou seja, $\forall \mathbf{t} = (t_1, t_2)$,

$$\sqrt{n}[t_1(\bar{y}_s - \beta)] + \sqrt{n}t_2 \left[\frac{(n-1)^2}{n(n-2)} s_y^2 - \sigma^2 - \frac{t - \bar{y}_s}{n-2} \right] \xrightarrow{\mathcal{D}} N(0, \theta) ,$$

$$\text{onde } \theta = [t_1 \ t_2] \begin{bmatrix} \sigma^2 & 0 \\ 0 & 2\sigma^4 \end{bmatrix} \begin{bmatrix} t_1 \\ t_2 \end{bmatrix} = t_1^2 \sigma^2 + 2t_2^2 \sigma^4 .$$

$$\text{Seja } a_n = \frac{(n-1)^2}{n(n-2)}, \text{ logo}$$

$$a_n s_y^2 - \sigma^2 = \frac{n-1}{n} (a_n s_y^2 - \sigma^2) + \frac{1}{n} (a_n s_y^2 - \sigma^2)$$

$$\begin{aligned}
&= \frac{n-1}{n} \left\{ \frac{a_n}{n-1} \sum_{j \in s} (y_j - \beta)^2 - \frac{a_n n}{n-1} (\bar{y}_s - \beta)^2 - \sigma^2 \right\} + \frac{1}{n} (a_n s_y^2 - \sigma^2) \\
&= a_n \frac{1}{n} \sum_{j \in s} (y_j - \beta)^2 - a_n (\bar{y}_s - \beta)^2 - \frac{n-1}{n} \sigma^2 + \frac{1}{n} (a_n s_y^2 - \sigma^2) \\
&= a_n \frac{1}{n} \sum_{j \in s} X_j + \left(a_n - \frac{n-1}{n} \right) \sigma^2 - a_n (\bar{y}_s - \beta)^2 + \frac{1}{n} (a_n s_y^2 - \sigma^2)
\end{aligned}$$

onde $X_j = (y_j - \beta)^2 - \sigma^2$, $E(X_j) = 0$, $V(X_j) = 2\sigma^4$.

Desta forma,

$$\begin{aligned}
&\sqrt{n}t_1(\bar{y}_s - \beta) + \sqrt{n}t_2 \left[\frac{(n-1)^2}{n(n-2)} s_y^2 - \sigma^2 - \frac{t - \bar{y}_s}{n-2} \right] \\
&= \sqrt{n}t_1(\bar{y}_s - \beta) + \sqrt{n}t_2 \left(a_n \frac{1}{n} \sum_{j \in s} X_j + R_n \right),
\end{aligned}$$

onde $R_n = \left(a_n - \frac{n-1}{n} \right) \sigma^2 - a_n (\bar{y}_s - \beta)^2 + \frac{1}{n} (a_n s_y^2 - \sigma^2) - \frac{t - \bar{y}_s}{n-2}$ e então $\sqrt{n}R_n \xrightarrow{P} 0$.

Além disso,

$$\begin{aligned}
&\sqrt{n}t_1(\bar{y}_s - \beta) + \sqrt{n}t_2 a_n \frac{1}{n} \sum_{j \in s} X_j \\
&= \sqrt{n} \left[t_1(\bar{y}_s - \beta) + t_2 \frac{1}{n} \sum_{j \in s} X_j \right] + \sqrt{n} \left[t_2 \frac{\sum_{j \in s} X_j}{n} (a_n - 1) \right] \\
&= \sum_{j \in s} \frac{[t_1(y_j - \beta) + t_2 X_j]}{\sqrt{n}} + t_2 \sum_{j \in s} \frac{X_j}{\sqrt{n}} (a_n - 1).
\end{aligned}$$

Mas $\sum_{j \in s} \frac{[t_1(y_j - \beta) + t_2 X_j]}{\sqrt{n}} \xrightarrow{D} N(0; \theta)$, pois

$$\begin{aligned}
\text{Var} [t_1(y_j - \beta) + t_2 X_j] &= E[t_1^2(y_j - \beta)^2 + t_2^2 X_j^2 + 2t_1 t_2 (y_j - \beta)[(y_j - \beta)^2 - \sigma^2]] \\
&= t_1^2 \sigma^2 + t_2^2 2\sigma^4 = \theta.
\end{aligned}$$

Além disso,

$$\frac{1}{\sqrt{n}} \sum_{j \in s} X_j = O_p(1) \quad \text{e} \quad (a_n - 1) = o_p(1).$$

Fica assim mostrado que

$$\sqrt{n}(\mathbf{T}_n - \boldsymbol{\mu}) \xrightarrow{D} N \left(0; \begin{bmatrix} \sigma^2 & 0 \\ 0 & 2\sigma^4 \end{bmatrix} \right).$$

Seja $g : \mathbb{R} \times \mathbb{R}_+ \rightarrow \mathbb{R}$ tal que $g(x_1, x_2) = \frac{t - x_1}{x_2^{1/2}}$. Temos

$$\begin{aligned}\dot{g}_1(x_1, x_2) &= \frac{\partial g}{\partial x_1} = -\frac{1}{x_2^{1/2}} \quad \text{e} \\ \dot{g}_2(x_1, x_2) &= \frac{\partial g}{\partial x_2} = \frac{1}{2} \frac{(t - x_1)}{t_2^{3/2}}\end{aligned}$$

Pelo Teorema A, pg. 122, Serfling (1980),

$$\sqrt{n}(g(\mathbf{T}_n) - g(\boldsymbol{\mu})) \xrightarrow{\mathcal{D}} N(0; \omega^2)$$

onde $\omega^2 = \left(-\frac{1}{\sigma}, \frac{(t - \beta)}{2\sigma^3}\right) \begin{bmatrix} \sigma^2 & 0 \\ 0 & 2\sigma^4 \end{bmatrix} \begin{pmatrix} -1/\sigma \\ \frac{(t - \beta)}{2\sigma^3} \end{pmatrix} = 1 + \frac{(t - \beta)^2}{2\sigma^2}$, de onde segue o resultado.

APÊNDICE 2

Neste apêndice apresentamos alguns resultados utilizados para derivarmos as expressões das variâncias preditivas dos previsores ótimos que aparecem nos exemplos do Capítulo 2.

Hurt (1976) apresenta resultados importantes sobre expansões assintóticas da média e da variância de algumas funções de variáveis aleatórias. Enunciamos a seguir, nos Teorema A.2.1, Lema A.2.1 e Teorema A.2.2, alguns resultados de Hurt (1976).

Teorema A.2.1. *Seja $g = g(t, n)$ uma função definida em $\mathbb{R} \times N$, tal que para todo n e algum $q \geq 1$, g admite $(q + 1)$ primeiras derivadas contínuas para $t \in [\theta - \delta, \theta + \delta]$, onde $\delta > 0$ é independente de n . Suponha que g é limitada em $\mathbb{R} \times N$ e todas as derivadas $g^{(1)}, \dots, g^{(q+1)}$ são limitadas em $[\theta - \delta, \theta + \delta] \times N$. Seja $\{T_n\}$ uma sequência de estatísticas com momentos finitos até a ordem $2(q + 1)$ tal que $E|T_n - \theta|^{2(q+1)} = O(n^{-(q+1)})$. Então,*

$$\begin{aligned} & E[g(T_n, n) - g(\theta, n)] \\ &= \sum_{j=1}^q \frac{1}{j!} \left(\frac{\partial^j g}{\partial t^j} \right)_{t=\theta} E(T_n - \theta)^j + O(n^{-(q+1)/2}) \end{aligned}$$

e

$$\begin{aligned} & \text{Var} [g(T_n, n) - g(\theta, n)] \\ &= \sum_{\substack{j=1 \\ j+k \leq q+1}}^q \sum_{k=1}^q \frac{1}{j!} \frac{1}{k!} \left(\frac{\partial^j g}{\partial t^j} \right)_{t=\theta} \left(\frac{\partial^k g}{\partial t^k} \right)_{t=\theta} \text{Cov} [(T_n - \theta)^j, (T_n - \theta)^k] + O(n^{-(q+2)/2}) . \end{aligned}$$

Lema A.2.1. *Sejam $i_1, \dots, i_r$ números reais não negativos, $\sum_{k=1}^r i_k = j$, $j > 0$ e $T_1, \dots, T_r$ variáveis aleatórias. Então*

$$E[|T_1|^{i_1} \dots |T_r|^{i_r}] \leq \{[E|T_1|^j]^{i_1} \dots [E|T_r|^j]^{i_r}\}^{1/j}$$

assumindo que os momentos existem.

Teorema A.2.2. *Seja $g = g(t_1, \dots, t_r, n)$ uma função definida em $\mathbb{R}^r \times N$. Suponhamos que*

- 1) *para todo n , g é $(q+1)$ vezes totalmente diferenciável com respeito aos t_i 's no intervalo $K = \times_{i=1}^r [\theta_i - \delta_i, \theta_i + \delta_i]$, $\delta_i > 0$, δ_i independente de n ,*
- 2) *g é limitada em $\mathbb{R}^r \times N$,*
- 3) *todas as derivadas até ordem $q+1$ são limitadas em $K \times N$,*
- 4) *$\{(T_{1n}, \dots, T_{rn})\}_{n=1}^\infty$ é uma seqüência de estatísticas multidimensional tal que existem os momentos absolutos de T_{in} até a ordem $2(q+1)$,*
- 5) *para $i = 1, \dots, r$, $E|T_{in} - \theta_i|^{2(q+1)} = O(n^{-(q+1)})$.*

Então,

$$\begin{aligned} & E[g(T_{1n}, \dots, T_{rn}, n) - g(\theta_1, \dots, \theta_r, n)] \\ &= \sum_{j=1}^q \frac{1}{j!} \sum_{\substack{i_1 + \dots + i_r = j \\ i_1, \dots, i_r \geq 0}} \dots \sum \left(\frac{\partial^j g}{\partial t_1^{i_1} \dots \partial t_r^{i_r}} \right)_{t=\theta} E[(T_{1n} - \theta_1)^{i_1} \dots (T_{rn} - \theta_r)^{i_r}] + O(n^{-(q+1)/2}) , \end{aligned}$$

$$\begin{aligned}
& \text{Var} [g(T_{1n}, \dots, T_{rn}, n) - g(\theta_1, \dots, \theta_r, n)] \\
&= \sum_{\substack{j=1 \\ j+k \leq q+1}}^q \sum_{k=1}^q \frac{1}{j!} \frac{1}{k!} \sum_{\substack{i_1 + \dots + i_r = j \\ i_1, \dots, i_r \geq 0}} \dots \sum_{\substack{m_1 + \dots + m_r = k \\ m_1, \dots, m_r \geq 0}} \left(\frac{\partial^j g}{\partial t_1^{i_1} \dots \partial t_r^{i_r}} \right)_{\mathbf{t}=\boldsymbol{\theta}} \left(\frac{\partial^k g}{\partial t_1^{m_1} \dots \partial t_r^{m_r}} \right)_{\mathbf{t}=\boldsymbol{\theta}} \\
& \quad \text{Cov} [(T_{1n} - \theta_1)^{i_1} \dots (T_{rn} - \theta_r)^{i_r}, (T_{1n} - \theta_1)^{m_1} \dots (T_{rn} - \theta_r)^{m_r}] + O(n^{-(q+2)/2}) ,
\end{aligned}$$

onde $\mathbf{t} = (t_1, \dots, t_r)$, $\boldsymbol{\theta} = (\theta_1, \dots, \theta_r)$.

Baseando-se nas provas do Teorema A.2.1 e A.2.2 de Hurt (1976) enunciados acima, mostramos a seguir o

Teorema A.2.3. *Seja g uma função definida como no Teorema A.2.1. Sejam $\{T_n^i\}$ e $\{T_n^l\}$ duas seqüências de estatísticas com momentos finitos até a ordem $2(q+1)$, tais que $E|T_n^i - \theta^i|^{2(q+1)} = O(n^{-(q+1)})$ e $E|T_n^l - \theta^l|^{2(q+1)} = O(n^{-(q+1)})$. Então*

$$\begin{aligned}
& \text{Cov} [g(T_n^i, n) - g(\theta^i, n); g(T_n^l, n) - g(\theta^l, n)] \\
&= \sum_{j=1}^q \sum_{k=1}^q \frac{1}{j!} \frac{1}{k!} \left(\frac{\partial^j g}{\partial t_i^j} \right)_{t_i=\theta^i} \left(\frac{\partial^k g}{\partial t_l^k} \right)_{t_l=\theta^l} \text{Cov} [(T_n^i - \theta^i)^j, (T_n^l - \theta^l)^k] + O(n^{-(q+2)/2})
\end{aligned}$$

Prova. Denotemos

$$\begin{aligned}
J &= \text{Cov} [g(T_n^i, n) - g(\theta^i, n); g(T_n^l, n) - g(\theta^l, n)] \\
&= E\{[g(T_n^i, n) - g(\theta^i, n)][g(T_n^l, n) - g(\theta^l, n)]\} \\
&\quad - E[g(T_n^i, n) - g(\theta^i, n)]E[g(T_n^l, n) - g(\theta^l, n)] = J_1 - J_2 .
\end{aligned}$$

Usaremos a seguinte notação

$$\begin{aligned}
b_j^i &= \frac{1}{j!} \left(\frac{\partial^j g}{\partial t_i^j} \right)_{t_i=\theta^i}, \quad j = 1, \dots, q \\
b_{q+1}^i &= \frac{1}{(q+1)!} \left(\frac{\partial^{q+1} g}{\partial t_i^{q+1}} \right)_{t_i=\theta^i + \xi(t - \theta^i)}, \quad \xi \in (0, 1) .
\end{aligned}$$

Aplicando o Teorema 1, temos

$$\begin{aligned}
J_2 &= \left[\sum_{j=1}^q b_j^i E(T_n^i - \theta^i)^j + O(n^{-(q+1)/2}) \right] \left[\sum_{k=1}^q b_k^l E(T_n^l - \theta^l)^k + O(n^{-(q+1)/2}) \right] \\
&= \sum_{j=1}^q \sum_{k=1}^q b_j^i b_k^l E(T_n^i - \theta^i)^j E(T_n^l - \theta^l)^k + O(n^{-(q+1)/2}) \sum_{j=1}^q b_j^i E(T_n^i - \theta^i)^j \\
&\quad + O(n^{-(q+1)/2}) \sum_{k=1}^q b_k^l E(T_n^l - \theta^l)^k + O(n^{-(q+1)}) .
\end{aligned}$$

Então

$$J_2 = \sum_{j=1}^q \sum_{k=1}^q b_j^i b_k^l E(T_n^i - \theta^i)^j E(T_n^l - \theta^l)^k + O(n^{-(q+2)/2})$$

pois $E(T_n^i - \theta^i)^j = O(n^{-1/2})$ e $E(T_n^l - \theta^l)^j = O(n^{-1/2})$, $\forall j \geq 1$.

Seja $M = \{t : |t^i - \theta^i| < \varepsilon \text{ e } |t^l - \theta^l| < \varepsilon\}$, $M^c = \{t : |t^i - \theta^i| \geq \varepsilon \text{ ou } |t^l - \theta^l| \geq \varepsilon\}$ e F_n é a distribuição conjunta de T_n^i e T_n^l . Podemos reescrever J_1 como

$$\begin{aligned}
J_1 &= \iint_M [g(t^i, n) - g(\theta^i, n)][g(t^l, n) - g(\theta^l, n)] dF_n \\
&\quad + \iint_{M^c} [g(t^i, n) - g(\theta^i, n)][g(t^l, n) - g(\theta^l, n)] dF_n \\
&= J_{11} + J_{12}
\end{aligned}$$

Como g é limitada e, usando a desigualdade de Chebyshev, temos

$$\begin{aligned}
J_{12} &\leq c \iint_{M^c} dF_n = cP\{(T_n^i, T_n^l) \in M^c\} \leq cP(|T_n^i - \theta^i| \geq \varepsilon) \\
&\leq c \frac{E|T_n^i - \theta^i|^{q+2}}{\varepsilon^{q+2}} = O(n^{-(q+2)/2}) .
\end{aligned}$$

Expandindo g em Taylor, temos

$$\begin{aligned}
J_{11} &= \iint_M \left[\sum_{j=1}^q b_j^i (T_n^i - \theta^i)^j + b_{q+1}^i (T_n^i - \theta^i)^{q+1} \right] \left[\sum_{k=1}^q b_k^l (T_n^l - \theta^l)^k + b_{q+1}^l (T_n^l - \theta^l)^{q+1} \right] dF_n \\
&= \sum_{j=1}^q \sum_{k=1}^q b_j^i b_k^l \iint_M (T_n^i - \theta^i)^j (T_n^l - \theta^l)^k dF_n \\
&\quad + \sum_{j=1}^q b_j^i \iint_M b_{q+1}^l (T_n^i - \theta^i)^j (T_n^l - \theta^l)^{q+1} dF_n \\
&\quad + \sum_{k=1}^q b_k^l \iint_M b_{q+1}^i (T_n^i - \theta^i)^{q+1} (T_n^l - \theta^l)^k dF_n \\
&\quad + b_{q+1}^i b_{q+1}^l \iint_M (T_n^i - \theta^i)^{q+1} (T_n^l - \theta^l)^{q+1} dF_n
\end{aligned}$$

$$\begin{aligned}
& + \sum_{j=1}^q b_j^l \iint_M b_{q+1}^i (T_n^l - \theta^l)^j (T_n^i - \theta^i)^{q+1} dF_n \\
& + \iint_M b_{q+1}^i b_{q+1}^l (T_n^i - \theta^i)^{q+1} (T_n^l - \theta^l)^{q+1} dF_n \\
& = J_{111} + J_{112} + J_{113} + J_{114}
\end{aligned}$$

Como b_{q+1} é limitado, utilizando-se o Lema A.2.1 enunciado anteriormente, e pelas condições dos momentos das seqüências de estatísticas, temos

$$\begin{aligned}
& \left| \iint_M b_{q+1}^l (t^i - \theta^i)^j (t^l - \theta^l)^{q+1} dF_n \right| \leq c \iint_M |t^i - \theta^i|^j |t^l - \theta^l|^{q+1} dF_n \\
& \leq c E[|T_i - \theta^i|^j |T^l - \theta^l|^{q+1}] \leq c \{ [E|T_i - \theta^i|^{j+q+1}]^j [E|T^l - \theta^l|^{j+q+1}]^{q+1} \}^{1/j+q+1} \\
& = \{ [O(n^{-(j+q+1)/2})]^j [O(n^{-(j+q+1)/2})]^{q+1} \}^{1/j+q+1} = c O(n^{-(j+q+1)/2}) .
\end{aligned}$$

Logo $J_{112} = O(n^{-(q+2)/2})$. Da mesma forma, podemos verificar que $J_{113} = O(n^{-(q+2)/2})$ e $J_{114} = O(n^{-(q+2)/2})$. Além disso,

$$\iint_M (t^i - \theta^i)^j (t^l - \theta^l)^k dF_n = E[(T_n^i - \theta^i)^j (T_n^l - \theta^l)^k] - \iint_{M^c} (t^i - \theta^i)^j (t^l - \theta^l)^k dF_n .$$

Mas, utilizando-se as desigualdades de Cauchy e Chebyshev e o Lema 2, temos

$$\begin{aligned}
& \iint_{M^c} (t^i - \theta^i)^j (t^l - \theta^l)^k dF_n \leq E[|T_n^i - \theta^i|^{2j} |T_n^l - \theta^l|^{2k}] P\{|T_n^i, T_n^l\} \in M^c\} \\
& \leq \{ [E|T_n^i - \theta^i|^{2j+2k}]^{2j} [E|T_n^l - \theta^l|^{2j+2k}]^{2k} \}^{1/2j+2k} \frac{E|T_n^i - \theta^i|^{2q}}{\varepsilon^{2q}} = O(n^{-(q+2)})
\end{aligned}$$

logo $\iint_M (t^i - \theta^i)^j (t^l - \theta^l)^k dF_n = E[(T_n^i - \theta^i)^j (T_n^l - \theta^l)^k] + O(n^{-(q+2)/2})$. Portanto,

$$\begin{aligned}
J & = \sum_{j=1}^q \sum_{k=1}^q b_j^i b_k^l E[(T_n^i - \theta^i)^j (T_n^l - \theta^l)^k] \\
& \quad - \sum_{j=1}^q \sum_{k=1}^q b_j^i b_k^l E[(T_n^i - \theta^i)^j (T_n^l - \theta^l)^k] + O(n^{-(q+2)/2})
\end{aligned}$$

e o resultado segue substituindo-se b_j^i e b_k^l .

Teorema A.2.4. *Seja g uma função definida como no Teorema 2. Sejam $\{(T_{1n}^i, \dots, T_{rn}^i)\}$ e $\{(T_{1n}^l, \dots, T_{rn}^l)\}$ duas seqüências de estatísticas multidimensionais com momentos absolutos até a ordem $2(q+1)$, tais que $E|T_{jn}^i - \theta_j^i|^{2(q+1)} = O(n^{-(q+1)})$ e $E|T_{jn}^l - \theta_j^l|^{2(q+1)} = O(n^{-(q+1)})$ para $j = 1, \dots, r$. Então*

$$\begin{aligned} & \text{Cov} [g(T_{1n}^i, \dots, T_{rn}^i, n) - g(\theta_1^i, \dots, \theta_r^i, n); g(T_{1n}^l, \dots, T_{rn}^l, n) - g(\theta_1^l, \dots, \theta_r^l, n)] \\ &= \sum_{j=1}^q \sum_{k=1}^q \frac{1}{j!} \frac{1}{k!} \sum_{\substack{i_1 + \dots + i_r = j \\ i_1, \dots, i_r \geq 0}} \dots \sum_{\substack{m_1 + \dots + m_r = j \\ m_1, \dots, m_r \geq 0}} \left(\frac{\partial^j g}{\partial t_{1i}^{i_1} \dots \partial t_{ri}^{i_r}} \right)_{\mathbf{t}_i = \boldsymbol{\theta}^i} \left(\frac{\partial^k g}{\partial t_{1l}^{i_1} \dots \partial t_{rl}^{i_r}} \right)_{\mathbf{t}_l = \boldsymbol{\theta}^l} \\ & \text{Cov} [(T_{1n}^i - \theta_1^i)^{i_1} (T_{rn}^i - \theta_r^i)^{i_r}, (T_{1n}^l - \theta_1^l)^{i_1} \dots (T_{rn}^l - \theta_r^l)^{i_r}] + O(n^{-(q+2)/2}) . \end{aligned}$$

Prova. A prova é uma generalização do Resultado A.2.3.

APÊNDICE 3

Neste apêndice apresentamos alguns resultados da teoria assintótica (Serfling (1980)) utilizados no Capítulo 3.

A seguir temos a definição de integrabilidade uniforme.

Definição A.3.1. A seqüência de variáveis aleatórias $\{Y_n\}$ é uniformemente integrável, se

$$\lim_{c \rightarrow \infty} \sup_n E\{|Y_n|I(|Y_n| > c)\} = 0 .$$

O Lema A.3.1 enunciado a seguir fornece um par de condições de equivalência e uma condição suficiente para integrabilidade uniforme.

Lema A.3.1. (i) *Integrabilidade uniforme de $\{Y_n\}$ em $(\Omega, \mathcal{A}, P)$ é equivalente ao par de condições*

a) $\sup_n E|Y_n| < \infty;$

b) o conjunto de funções $\{Q_n\}$ definidas por $Q_n(A) = \int_A |Y_n| dP$ é uniforme e absolutamente contínua com respeito a P , ou seja, dado $\varepsilon > 0$, existe $\delta > 0$ tal que

$$P(A) < \delta \Rightarrow \sup_n |Q_n(A)| < \varepsilon .$$

(ii) É suficiente para a integrabilidade uniforme de $\{Y_n\}$ que exista uma variável aleatória Y , tal que $E|Y| < \infty$ e $P(|Y_n| \geq y) \leq P(|Y| \geq y)$ para todo $n \geq 1$ e $y > 0$.

O Teorema A.3.1 fornece condições para convergência de momentos.

Teorema A.3.1. Suponha que $X_n \xrightarrow{d} X$ e a seqüência $\{X_n^r\}$ é uniformemente integrável, onde $r > 0$. Então, $E|X|^r < \infty$, $\lim_n E\{X_n^r\} = E\{X^r\}$ e $\lim_n E|X_n|^r = E|X|^r$.

BIBLIOGRAFIA

- Basu, D. (1969). Role of sufficiency and likelihood principles in sample survey theory. *Sankhya A*, **31**: 441–454.
- Basu, D. (1971). An essay on the logical foundations of survey sampling. *Foundations of Statistical Inference*, Holt, Rinehart and Winston, Toronto: 203–242.
- Bolfarine, H. (1989). Equivariant prediction in finite populations. *Communications in Statistics, Theory and Methods*, **18**(3):927–942.
- Bolfarine, H. (1990). *Ratio and regression estimators under error in variables superpopulation*. São Paulo, IME-USP. 16p. (RT-MAE-9007).
- Bolfarine, H. (1991). Finite population prediction under error in variables superpopulation models. *Canadian Journal of Statistics*, **19**(2):191–207.
- Bolfarine, H. e Cordani, L.K. (1993). Estimation of a structural linear regression model with a known reliability ratio. *The Annals of the Institute of Statistical Mathematics*. (a ser publicado no vol. 45, 1993 ou vol. 46, 1994.)
- Bolfarine, H. & Rodrigues, J. (1990). Finite population prediction under a linear function superpopulation model: a Bayesian perspective. *Communications in Statistics, Theory and Methods*, **19**(7):2577–2594.
- Bolfarine, H. & Rodrigues, J. (1988). On the simple projection predictor in finite populations. *Australian Journal of Statistics*, **30**(3):338–341.

- Bolfarine, H. & Zacks, S. (1991). Bayes and minimax prediction in finite populations. *Journal of Statistical Planning and Inference*, **28**:139–151.
- Bolfarine, H. & Zacks, S. (1992). *Prediction theory for finite populations*. Springer-Verlag, New York.
- Bolfarine, H.; Pereira, C.A.B. & Rodrigues, J. (1987). Robust linear prediction in finite populations. A Bayesian perspective. *Sankhya B*, **49**:23–35.
- Cassel, C.M.; Sarndal, C.E. & Wretman, J.H. (1977). *Foundations of inference in survey sampling*. New York, John Wiley.
- Chambers, R.L. & Dunstan, R. (1986). Estimating distribution functions from survey data. *Biometrika*, **73**:597–604.
- Chambers, R.L.; Dorfman, A.H. & Hall, P. (1992). Properties of estimators of the finite population distribution function. *Biometrika*, **79**(3):577–582.
- Chambers, R.L.; Dorfman, A.H. & Wehrly, T.E. (1993). Bias robust estimation in finite populations using nonparametric calibration. *Journal of the American Statistical Association*, **88** (421):268–277.
- Cochran, W.G. (1977). *Sampling techniques*. New York, John Wiley.
- Cohen, M.P. & Kuo, L. (1985a). Admissibility of the empirical distribution functions. *Annals of Statistics*, **11**:262–271.
- Cohen, M.P. & Kuo, L. (1985b). Minimax sampling strategies for estimating a finite population distribution function. *Statistics and Decisions*, **3**:205–224.
- Cox, D.R. & Reid, N. (1987). Parameter orthogonality and approximate conditional inference (with discussion). *Journal of the Royal Statistical Society B*, **49**(1):1–39.
- Dorfman, A.H. (1993). A comparison of design-based and model-based estimators of the finite population distribution function. *Australian Journal of Statistics*, **35**.
- Dunstan, R. & Chambers, R.L. (1989). Estimating distribution functions from survey data with limited benchmark information. *Australian Journal of Statistics*, **31**(1):1–11.
- Efron, B. & Morris, C. (1973). Stein's estimation rule and its competitors. An empirical Bayesian approach. *Journal of the American Statistical Association*, **68**:117–130.
- Efron, B. & Morris, C. (1975). Data analysis using Stein's estimator and its generalizations. *Journal of the American Statistical Association*, **70**:311–319.
- Efron, B. & Morris, C. (1977). Stein's paradox in statistics. *Scientific American*, **236**:119–127.

- Elian, S.N. (1991). *Análise de regressão em populações finitas*. São Paulo. 229p. Tese (Doutorado). Instituto de Matemática e Estatística da USP.
- Ericson, W.A. (1969). Subjective Bayesian models in sampling finite populations. *Journal of the Royal Statistical Society B*, **31**:195–224.
- Ericson, W.A. (1988). Bayesian inference in finite populations. *Handbook of Statistics*, vol. 6, 213–246. P.R. Krishnaiah and C.R. Rao, eds. North Holland, New York.
- Francisco, C.A. & Fuller, W.A. (1986). Estimation of the distribution function with a complex survey. *Proc. Sec. Survey Res. Methods*, American Statistical Association, Washington, D.C., 37–45.
- Fuller, W.A. (1975). Regression analysis for sample surveys. *Sankhya C*, **37**:117–32.
- Fuller, W.A. (1976). *Introduction to statistical time series*. New York, John Wiley.
- Fuller, W.A. (1987). *Measurement error models*. New York, John Wiley.
- Fuller, W.A. (1990). Prediction of true values for the measurement error model. *Contemporary Mathematics*, **112**:41–57.
- Ghosh, M. & Lahiri, P. (1987a). Robust empirical Bayes estimation of means from stratified samples. *Journal of the American Statistical Association*, **82**:1153–1162.
- Ghosh, M. & Lahiri, P. (1987b). Robust empirical Bayes estimation of variances from stratified samples. *Sankhya B*, **49**:78–89.
- Ghosh, M. & Meeden, G. (1986). Empirical Bayes estimation in finite population sampling. *Journal of the American Statistical Association*, **81**:1058–1062.
- Godambe, V.P. (1955). A unified theory of sampling from finite populations. *Journal of the Royal Statistical Society B*, **17**: 269–278.
- Godambe, V.P. (1966). A new approach to sampling from finite populations, I, II. *Journal of the Royal Statistical Society B*, **28**:310–319:320–328.
- Hartley, H.O. & Sielken, R.L. Jr. (1975). A “super-population viewpoint” for finite population sampling. *Biometrics*, **31**: 411–422.
- Hurt, J. (1976). Asymptotic expansions of functions of statistics. *Aplikace Matematiky*, **21**:444–456.
- Kendall, M.G. & Stuart, A. (1961). *The advanced theory of statistics*, vol. 2. Griffin, London.
- Kish, L. (1965). *Survey sampling*. New York, John Wiley.
- Kuk, A.Y.C. (1988). Estimating of distribution functions and medians under sampling with unequal probabilities. *Biometrika*, **75**:97–104.

- Kuo, L. (1988). Classical and prediction approaches to estimating distribution functions from survey data. *Proceedings of the Section on Survey Research Methods*, American Statistical Association: 280–285.
- Lahiri, P. & Peddada, S.D. (1992). Bayes and empirical Bayes estimation of the finite population mean using auxiliary information. *Statistics & Decisions*, **10**:67–80.
- Lehmann, E.L. (1983). *Theory of point estimation*. New York, John Wiley.
- Lindley, D.V. (1966). Discussion on a generalized least squares approach to linear functional relationship. *Journal of the Royal Statistical Society B*, **28**:279–297.
- Lindley, D.V. & El Sayad, G. (1968). The Bayesian estimation of a linear functional relationship. *Journal of the Royal Statistical Society B*, **30**:190–202.
- Moran, P.A.P. (1971). Estimating structural and functional relationships. *Journal of Multivariate Analysis*, **1**:232–255.
- Morris, C. (1983). Parametric empirical Bayes inference: theory and applications (with discussion). *Journal of the American Statistical Association*, **78**:47–65.
- Olkin, I. & Ghurye, S.G. (1969). Unbiased estimators of some multivariate densities and related functions. *The Annals of Mathematical Statistics*, **40**:1261–1271.
- O'Reilly, F.J. (1974). On a criterion for a extrapolation in normal regression. *The Annals of Statistics*, **3**:219–222.
- Pereira, C.A.B. & Rodrigues, J. (1983). Robust linear prediction in finite populations. *International Statistical Review*, **51**: 293–300.
- Rao, J.N.K.; Kovar, J.G. & Mantel, H.J. (1990). On estimating distribution functions and quantiles from survey data using auxiliary information. *Biometrika*, **77**:365–375.
- Robbins, H. (1955). On empirical Bayes approach to statistics. *Proceedings of the 3rd Berkeley Symposium on Mathematical Statistics and Probability*. (vol. 6). Berkeley. University of California Press, 157–163.
- Reilly, P. & Patino-Leal, H. (1981). A Bayesian study of the error in variable models. *Technometrics*, **23**(3):221–231.
- Rodrigues, J. (1988). Some results on restricted Bayes least squares predictors for finite populations. *South African Statistics Journal*, **22**:45–53.
- Rodrigues, J.; Bolfarine, H. & Rogakto, A. (1985). A general theory of prediction in finite populations. *International Statistical Review*, **53**:239–254.

- Rodrigues, J. & Cordani, L.K. (1990). A note on likelihood estimation of a simple regression model with measurement error, via the orthogonal parametrization. *South African Statistics Journal*, **24**:177-183.
- Royall, R.M. (1970). On finite population sampling theory under certain linear regression models. *Biometrika*, **57**: 377-387.
- Royall, R.M. (1976a). The linear least squares prediction approach to two stage sampling. *Journal of the American Statistical Association*, **71**:657-664.
- Royall, R.M. (1976b). Likelihood function in finite population sampling theory. *Biometrika*, **63**:605-614.
- Royall, R.M. & Cumberland, W.G. (1978). Variance estimation in finite population sampling. *Journal of the American Statistical Association*, **73**:351-358.
- Royall, R.M. & Cumberland, W.G. (1981a). An empirical study of the ratio estimator and estimators of its variance. *Journal of the American Statistical Association*, **76**:66-77.
- Royall, R.M. & Cumberland, W.G. (1981b). The finite population linear regression estimator and estimators of its variance. An empirical study. *Journal of the American Statistical Association*, **76**:924-930.
- Royall, R.M. & Eberhardt, K.R. (1975). Variance estimates of the ratio estimator. *Sankhya C*, **37**:43-52.
- Royall, R.M. & Herson, J. (1973a). Robust estimation in finite populations I. *Journal of the American Statistical Association*, **68**:880-889.
- Royall, R.M. & Herson, J. (1973b). Robust estimation in finite populations II. Stratification on a size variable. *Journal of the American Statistical Association*, **68**:890-893.
- Royall, R.M. & Pfefferman, D. (1982). Balanced samples and robust Bayesian inference in finite population sampling. *Biometrika*, **69**:401-409.
- Sarndal, C.E. (1982). Implications of survey design for generalized regression estimation of linear functions. *Journal of Statistical Planning and Inference*, **7**:155-170.
- Sarndal, C.E. & Wright, R.L. (1984). Cosmetic forms of estimators in survey sampling. *Scandinavian Journal of Statistics*, **11**:146-156.
- Scott, A.J.; Brewer, K.R.W. & Ho, E.W.H. (1978). Finite population sampling and robust estimation. *Journal of the American Statistical Association*, **73**(362):359-361.
- Scott, A.J. & Wu, C. (1981). On the asymptotic distribution of ratio and regression estimators. *Journal of the American Statistical Association*, **76**(373):98-102.

- Sedransk, N. & Sedransk, J. (1979). Distinguishing among distributions using data from complex sample designs. *Journal of the American Statistical Association*, **74**:754–760.
- Serfling, R.J. (1980). *Approximation theorems of mathematical statistics*. New York, John Wiley.
- Smith, T.M.F. (1976). The foundations of survey sampling: a review. *Journal of the Royal Statistical Society A*, **146** (4):394–403.
- Smouse, E. (1984). A note on Bayesian least squares inference for finite population models. *Journal of the American Statistical Association*, **79**(386):390–392.
- Sprenst, P. (1966). A generalized least squares approach to linear functional relationships. *Journal of the Royal Statistical Society B*, **28**:279–297.
- Tallis, G.M. (1978). Note of robust estimation in finite populations. *Sankhya C*, **40**(2):136–138.
- Tam, S.M. (1987). Optimality of Royall's predictor under a Gaussian superpopulation model. *Biometrika*, **74**(3):659–660.
- Tiwari, R. & Lahiri, P. (1989). Empirical Bayes estimation of the mean of a nonhomogeneous finite population. *Communications in Statistics. Theory and Methods*, **18**(7):2553–2568.
- Wong, M.Y. (1989). Likelihood estimation of a simple regression model when both variables have error. *Biometrika*, **76**:141–148.
- Zacks, S. (1981). Bayes equivariant estimators of the variance of a finite population for exponential priors. *Communications in Statistics A*, **10**:427–437.
- Zacks, S. & Bolfarine, H. (1990a). *Equivariant prediction of the population variance in finite populations*. São Paulo, IME-USP. 8p. (RT-MAE-9005)
- Zacks, S. & Bolfarine, H. (1990b). *Maximum likelihood predictors in finite populations*. São Paulo, IME-USP. 15p. (RT-MAE-9004)
- Zacks, S. & Solomon, H. (1981). Bayes equivariant estimators of the variance of a finite population: part I, simple random sampling. *Communications in Statistics. Theory and Methods*, **10**: 407–426.
- Zellner, A. (1971). *An introduction to Bayesian inference in econometrics*. New York, John Wiley.