

UM MODELO LINEAR NÃO-PARAMÉTRICO COM (POSSÍVEL) ERRO ASSIMÉTRICO

Este exemplar corresponde a redação final da tese devidamente corrigida e defendida pelo Sr. Vicente Garibay Cancho e aprovada pela comissão julgadora.

Campinas, 23 de Novembro de 1994.



Prof. Dr. Belmer Garcia Negrillo.
Orientador

Dissertação apresentada ao Instituto de Matemática, Estatística e Ciência da computação, UNICAMP, como requisito parcial para obtenção do Título de Mestre em Estatística.

UNIVERSIDADE ESTADUAL DE CAMPINAS
DEPARTAMENTO DE ESTATÍSTICA - IMECC

***UM MODELO LINEAR NÃO-PARAMÉTRICO
COM (POSSÍVEL) ERRO ASSIMÉTRICO***

Vicente Garibay Cancho 

Orientação

Prof. Dr. Belmer Garcia Negrillo 

Dissertação apresentada ao instituto de Matemática, Estatística e Ciência da Computação da
Universidade Estadual de Campinas para obtenção de título de Mestre em Estatística

Campinas - S. P.
1994

*À memória de meus pais
Rosendo e Eulogia.*

AGRADECIMENTOS

Gostaria de expressar meus agradecimentos

- *Aos professores do Departamento de Estatística da UNICAMP pela contribuição à minha formação.*
- *Ao professor Belmer pela amizade e orientação.*
- *Ao Prof. Dr. Irwen Guadalupe pela amizade, e importantes sugestões.*
- *À minha família pelo apoio nos momentos difíceis.*
- *À Rissa pela correção do português e amizade.*
- *Aos amigos que conquistei durante o mestrado, pelo companheirismo e apoio.*
- *A CAPES pelo apoio financeiro.*
- *À Gladys por sua compreensão e apoio constante.*

Conteúdo

1	INTRODUÇÃO	3
1.1	Visão Geral	3
1.2	Suposições	10
1.3	Estratégias Básicas e Lemas	14
1.4	Notações e Lemas Adicionais	20
2	LINEARIDADE DO GRADIENTE DE DISPERSÃO	28
2.1	Propriedade	28
2.2	Discussão	39
3	ESTIMAÇÃO DE DENSIDADES	41
3.1	Caso De Uma Amostra	41
3.2	Resíduos e Largura da Janela Aleatória	43
4	ESTIMAÇÃO DE UMA FUNCIONAL DE ESCALA	54
4.1	Introdução	54
4.2	Normalidade Assintótica no Caso de uma Amostra	56
4.3	Resíduos e Largura da Janela Aleatória	65
4.4	Cálculos	75
4.4.1	Escolha da Largura da Janela	75
4.4.2	Estimativa de θ Baseado no Comprimento do Intervalo de Con- fiança Aproximado numa Amostra.	78
4.4.3	Estimativa de θ Baseada no Comprimento do Intervalo de Con- fiança Aproximado para duas Amostras	80

5	INTERCEPTO	84
5.1	Normalidade Assintótica de $\begin{pmatrix} \hat{\alpha} \\ \hat{\beta} \end{pmatrix}$	84
5.2	Estimação de $f(0)$	93
6	Considerações finais	96
	REFERÊNCIAS BIBLIOGRÁFICAS	98

Capítulo 1

INTRODUÇÃO

1.1 Visão Geral

Os modelos lineares são ferramentas estatísticas extensivamente utilizadas na pesquisa empírica. Aplicações de análise de regressão, análise de variância e técnicas da análise de covariância etc., abundam em ciências biológicas, físicas e sociais, assim como também na pesquisa agrícola e industrial. A análise clássica dos modelos lineares se baseia principalmente em dois métodos: de mínimos quadrados, e de máxima verossimilhança. O primeiro método usualmente produz estimativas computacionalmente simples, embora não seja muito robusto aos *outliers*, erros grosseiros, e às distribuições com caudas pesadas. A análise pelo segundo método pressupõe o conhecimento da distribuição subjacente, e as estimativas determinadas por este método conservam suas propriedades (assintóticas) sempre que a distribuição subjacente for verdadeira. Em casos extremos este segundo procedimento pode produzir estimativas totalmente inconsistentes, por exemplo quando assume-se normalidade da distribuição subjacente sendo Cauchy a verdadeira distribuição.

Na análise clássica, o procedimento de inferência pressupõe o conhecimento da distribuição subjacente, geralmente supõe-se normalidade, mas muitos experimentos não satisfazem esta condição. Como foi mencionado anteriormente, uma suposição errada da distribuição subjacente, pode conduzir conclusões totalmente erradas.

Mckean e Hettmansperger (1976) propuseram uma análise não-paramétrica baseada nos postos dos resíduos supondo a simetria da distribuição dos erros.

Considera-se o modelo linear geral não-paramétrico

$$Y = 1_n \alpha_0 + X \beta_0 + e, \quad (1.1)$$

onde $Y = (Y_1, \dots, Y_n)'$ é o vetor de respostas $n \times 1$, 1_n é o vetor $n \times 1$ de 1's, X é a matriz de desenho $n \times p$ de posto completo, α_0 é o intercepto desconhecido (escalar), β_0 é o vetor $p \times 1$ de parâmetros de regressão desconhecidos, e $e = (e_1, e_2, \dots, e_n)'$ é o vetor aleatório $n \times 1$, com componentes independentes e identicamente distribuídos, com função de distribuição acumulada $F(\cdot)$ e função de densidade $f(\cdot)$.

O procedimento de inferência para β_0 que consideraremos é baseado na medida de dispersão proposta por Jaeckel (1972), a qual define-se para o vetor $v = (v_1, v_2, \dots, v_n)'$ como :

$$D(v) = \sum_{i=1}^n a(i) v_{(i)}, \quad (1.2)$$

onde $v_{(1)} \leq v_{(2)} \leq \dots \leq v_{(n)}$ são os elementos ordenados de v e $a(i)$ é definida em termos de uma função geradora de escores $\phi(\cdot)$, $a(i) = \phi(i/(n+1))$. Além disso, se $R_1, \dots, R_n$ são os postos de $v_1, v_2, \dots, v_n$, então podemos reescrever (1.2) como:

$$D(v) = \sum_{i=1}^n a(R_i) v_i. \quad (1.3)$$

A função $\phi(\cdot)$ deve ser não decrescente no intervalo $(0, 1)$ e deve satisfazer

$$\phi(u) = -\phi(1-u) \quad e \quad (1.4)$$

$$\int_0^1 \phi(u)^2 du = 1, \quad (1.5)$$

esta última condição é por conveniência para formulação de vários resultados. Qualquer função não constante, quadrado integrável em $(0, 1)$ pode ser padronizada de maneira que (1.5) seja válido sem afetar as propriedades dos procedimentos resultantes.

Uma importante função geradora de escores é

$$\phi(u) = \sqrt{12}(u - 1/2). \quad (1.6)$$

Os escores gerados por (1.6) são conhecidos como escores de Wilcoxon, eles atribuem uma ponderação a cada elemento de u , proporcional à diferença entre seu posto, e o posto médio. Procedimentos baseados nestes escores, em essência generalizam o procedimento para duas amostras de Mann-Whitney-Wilcoxon e herdam suas propriedades de eficiência assintótica.

Mckean e Scharader (1980) adaptaram os resultados de Jaeckel para mostrar que a função de dispersão é uma pseudo-norma, isto é;

$$D(kv) = |k| D(v), \quad (1.7)$$

$$D(w + v) \leq D(v) + D(w), \quad e \quad (1.8)$$

$$D(v) = 0 \quad \text{se e só se} \quad v_1 = v_2 = \dots = v_n = 0. \quad (1.9)$$

Assim, $D(Y_i - \alpha - x'_i\beta)$ é uma medida razoável de dispersão dos resíduos, $(Y_i - \alpha - x'_i\beta)$ para quaisquer α e β dados. Aqui x'_i é a i -ésima linha da matriz de desenho X . Como $D(Y - \alpha 1_n - X\beta) = D(Y - X\beta)$, uma estimativa de α_0 não pode depender desta medida de dispersão, já que esta medida não dá informação do intercepto. E dado que estamos interessados primeiramente na estimação de β_0 , por conveniência escrevemos $D(\beta)$ ao invés de $D(Y - X\beta)$.

Jaeckel (1972) propõe uma estimativa de β_0 pela minimização de $D(\beta)$. Ele mostrou que $D(\beta)$ é uma função não negativa, convexa e contínua em β , de modo que sempre existe pelo menos uma estimativa. Não obstante, a estimativa não precisa ser definida única, Jaeckel mostrou que $n^{1/2}$ vezes o diâmetro do conjunto de pontos que minimiza $D(\beta)$ converge para zero em probabilidade. Previamente Jureckova (1971) tinha proposto uma estimativa de β_0 , fazendo com que a gradiente de $D(\beta)$ seja aproximadamente igual

a zero. Além disso, ela tinha mostrado que:

$$\sqrt{n}(\hat{\beta} - \beta_0) \xrightarrow{D} Z \sim NMV(0, \tau^2 \Sigma^{-1}), \quad (1.10)$$

onde NMV denota a distribuição normal multivariada,

$$\tau^{-1} = - \int_0^1 \phi(u) \left[\frac{f'(F^{-1}(u))}{f(F^{-1}(u))} \right] du \quad (1.11)$$

e Σ é positiva definida e é o limite de $n^{-1}X'_C X_C$ quando $n \rightarrow \infty$.

A matriz de desenho centralizada é dada por

$$X_C = X - 1_n \bar{x}', \quad (1.12)$$

onde $\bar{x}'$ é o vetor linha $1 \times p$ de médias de X .

Jaekel (1972) mostrou que $n^{1/2}$ vezes a diferença entre sua estimativa e a de Jurckova converge para zero em probabilidade. Assim sendo, estas duas estimativas são assintoticamente equivalentes e têm a mesma distribuição assintótica. Neste trabalho não fazemos distinção entre estas duas.

Suponhamos que desejamos testar as seguintes hipóteses

$$H_0 : H\beta_0 = \gamma \text{ contra } H_A : H\beta_0 \neq \gamma, \quad (1.13)$$

onde H é uma matriz $q \times p$ de posto completo $q \leq p$, e γ é um vetor $q \times 1$ completamente especificado; frequentemente $\gamma = 0$. Então, sob a hipótese nula, a variável aleatória

$$Q = \tau^{-2} (H\hat{\beta} - \gamma)' (H (X'_C X_C)^{-1} H')^{-1} (H\hat{\beta} - \gamma) \quad (1.14)$$

converge em distribuição para uma variável aleatória qui-quadrado com q graus de liberdade. Se τ é conhecido, o teste que rejeita H_0 quando $Q > \chi^2_\alpha(q)$ é consistente para alternativas fixas, onde $\chi^2_\alpha(q)$ é o valor critico determinado da distribuição qui-quadrado

com q graus de liberdade com α fixado.

Outra abordagem para testar as hipóteses em (1.13) depende do ajustamento do modelo completo ($\hat{\beta}_{com}$) e do modelo reduzido ($\hat{\beta}_{red}$) pela hipótese nula. Aqui requer-se que $q < p$, McKean e Hettmansperger (1976) mostraram que sob a hipótese nula,

$$D^* = \frac{2 [D(\hat{\beta}_{red}) - D(\hat{\beta}_{com})]}{\tau}, \quad (1.15)$$

tem também uma distribuição assintótica qui-quadrado com q graus de liberdade. Além disso, eles mostraram que se τ é conhecido o teste que rejeita H_0 quando $D^* > \chi^2_\alpha(q)$ é consistente para alternativas fixas.

Na prática, ainda não conhecemos τ . Portanto, se desejamos utilizar Q ou D^* para testar hipóteses enfrentamos o problema de estimação de τ . Adicionalmente, muitas aplicações de inferência relacionadas com o intercepto α_0 são desejadas. Como mencionamos anteriormente, a função de dispersão de Jaekel não pode ser utilizada para este propósito.

McKean e Hettmansperger (1976, 1978) propuseram superar essas duas dificuldades pela aplicação do procedimento posto-sinal (*signed-rank*) numa amostra dos resíduos $(Y_i - x'_i \hat{\beta})$. O preço para obter esta solução foi a suposição de simetria da distribuição dos erros.

Define-se a função geradora de escores para uma amostra, $\phi^+(\cdot)$ correspondente a $\phi(\cdot)$ por,

$$\phi^+(u) = \phi([u + 1] / 2). \quad (1.16)$$

A restrição $\phi(u) = -\phi(1 - u)$ dada em (1.4) e a de simetria de $f(\cdot)$ em torno de zero significa que

$$(\tau^+)^{-1} = - \int_0^1 \phi^+(u) \left[\frac{f'(F^{-1}([u + 1]/2))}{f(F^{-1}([u + 1]/2))} \right] du$$

é igual a τ^{-1} dada em (1.11). Isto foi estudado por Lehmann (1963) e Sen (1966). Para uma amostra aleatória com distribuição simétrica, eles propuseram uma estimativa consistente de $\tau^+ = \tau$ baseada no comprimento do intervalo de $100(1 - \alpha)\%$ de confiança

para o centro de simetria, proveniente da estatística do teste posto-sinal com a função geradora de escores $\phi^+(\cdot)$ (1.16). Particularmente

$$\tau^* = \frac{\sqrt{n}(\hat{\alpha}_U - \hat{\alpha}_L)}{2Z_{\alpha/2}} \quad (1.17)$$

onde $(\hat{\alpha}_L, \hat{\alpha}_U)$ é o intervalo de $100(1-\alpha)\%$ de confiança para o centro de simetria e $Z_{\alpha/2}$ é o valor crítico superior da distribuição normal padrão, onde esta quantidade converge em probabilidade para τ . McKean e Hettmansperger (1976) mostraram que τ^* calculada nos resíduos, $(Y_i - x_i'\hat{\beta})$, é também consistente para τ . Na Seção 4.4 fazemos alguns cálculos que fornecem alguns critérios sobre a necessidade da suposição de simetria quando a estimativa (1.17) é usada.

Finalmente, McKean e Hettmansperger (1978) propuseram α^* como a estimativa do intercepto obtida pela aplicação do procedimento posto-sinal numa amostra com função geradora de escores $\phi^+(\cdot)$ (1.16) para os resíduos, $(Y_i - x_i'\hat{\beta})$. Novamente, assumindo simetria na distribuição dos erros eles mostraram

$$\sqrt{n} \begin{bmatrix} \alpha^* - \alpha_0 \\ \hat{\beta} - \beta_0 \end{bmatrix} \xrightarrow{D} Z \sim NMV(0, \tau^2 \Lambda^{-1}), \quad (1.18)$$

onde Λ é positiva definida e é o limite de $\sqrt{n}(1_n, X)'(1_n, X)$ quando $n \rightarrow \infty$. Portanto um teste consistente das hipóteses

$$H_0: H \begin{pmatrix} \alpha_0 \\ \beta_0 \end{pmatrix} = \gamma \quad \text{contra} \quad H_A: H \begin{pmatrix} \alpha_0 \\ \beta_0 \end{pmatrix} \neq \gamma, \quad (1.19)$$

onde H é uma matriz $q \times (p+1)$ de posto completo com $q \leq p+1$ e γ um vetor $q \times 1$ completamente especificado, pode ser baseado em

$$Q^* = \tau^{-2} \left(H \begin{pmatrix} \alpha^* \\ \hat{\beta} \end{pmatrix} - \gamma \right)' \left(H ((1_n, X)'(1_n, X))^{-1} H' \right)^{-1} \left(H \begin{pmatrix} \alpha^* \\ \hat{\beta} \end{pmatrix} - \gamma \right). \quad (1.20)$$

Sob a hipótese nula, esta variável aleatória tem uma distribuição assintótica qui-quadrado com q graus de liberdade. Assim, para uma amostra grande, o teste aproximado de tamanho α rejeita H_0 quando $Q^* > \chi^2_\alpha(q)$. Mais uma vez precisamos de uma estimativa consistente de τ para aplicar este teste.

Este trabalho trata do problema relacionado à inferência baseado em postos para modelos lineares gerais com (possíveis) erros assimétricos. Especificamente uma aproximação para a estimação da funcional de escala τ (1.11) é desenvolvida, a qual difere da aproximação dada por McKean e Hettmansperger (1976), que requer simetria da distribuição subjacente dos erros. Essa funcional de escala τ é de grande interesse se é aplicada aos procedimentos baseados em postos, já que ela aparece na estrutura da matriz de variância-covariância assintótica das estimativas do vetor de parâmetros de regressão, e no denominador da estatística do teste para hipótese sobre esse vetor. Desta forma, ela desempenha o mesmo papel que σ^2 no procedimento usual de mínimos quadrados. Uma estimativa do intercepto é desenvolvida sem a suposição de simetria da distribuição dos erros. Adicionalmente, é mostrado que as propriedades assintóticas do procedimento de inferência baseado em postos são válidas com menos condições na matriz de desenho do que previamente exigiam.

O desenvolvimento das propriedades dos procedimentos de inferência discutido anteriormente, tem como fundamentação o importante teorema de Jureckova (1971). Este teorema estabelece a propriedade de linearidade assintótica do gradiente de dispersão num modelo linear geral. No Capítulo 2, esta propriedade de linearidade é obtida usando-se condições menos complicadas na matriz de desenho, mas somente para escores de Wilcoxon (1.6). Assim, os usuários, preocupados na verificação das suposições, estarão um pouco menos indecisos ao uso dos procedimentos baseados em postos.

Em algumas áreas de aplicação, por exemplo experimentos biomédicos, são comuns as distribuições assimétricas e frequentemente os usuários realizam alguma transformação nos dados com a finalidade de obter distribuição simétrica. Enquanto as propriedades assintóticas de $\hat{\beta}$, Q (1.14), e D (1.15) não dependem da simetria da distribuição subja-

cente dos erros, a estimativa τ^* (1.17) da funcional de escala τ precisa da suposição de simetria portanto o uso de Q ou D^* em teste de hipótese sobre β_0 , exige uma estimativa consistente de τ e tal uso até agora exige a suposição de simetria. A funcional de escala τ é importante porque aparece na matriz de variância-covariância assintótica de $\hat{\beta}$. Já que frequentemente estamos interessados no desvio padrão das estimativas pontuais de β . Além disso, uma estimativa prévia do intercepto α_0 (McKean e Hettmansperger, 1978), requer a suposição de simetria .

Schuster (1974), Schweder (1975), e outros (Seção 4.1) estudaram estimativas de τ^{-1} , no caso de uma amostra baseada na estimação da função de densidade do tipo Rosemblatt (1956). Estas estimativas de τ^{-1} não exigem a suposição de simetria da distribuição. No Capítulo 4 discutimos tais estimativas para o caso de escores de Wilcoxon (1.6) e obtemos alguns resultados adicionais para o caso de uma amostra. Além disso, mostraremos que tal estimativa é sempre consistente quando calculada sobre os resíduos após o ajuste do modelo linear. Alguns cálculos são realizados para sugerir possíveis ajuste em pequenas amostras da estimativa proposta e obter alguma relação com as outras estimativas.

A estimação de densidades mencionada anteriormente é discutida no Capítulo 3. Mostramos que as estimativas obtidas são consistentes quando calculamos sobre os resíduos (não necessariamente resultante de uma estimativa baseada em postos).

No Capítulo 5 propomos uma simple estimativa $\hat{\alpha}$ para o intercepto α_0 como a mediana dos resíduos, $(Y_i - x_i'\hat{\beta})$, e derivamos a distribuição multivariada assintótica de $(\hat{\alpha}, \hat{\beta})'$ sem a suposição de simetria. Adicionalmente é discutida uma estimativa de $f(0)$, a qual aparece na matriz de variância-covariância assintótica de $(\hat{\alpha}, \hat{\beta})'$.

1.2 Suposições

Como foi tratado na Seção 1.1, o desenvolvimento das propriedades assintóticas da estimativa $\hat{\beta}$ e das variáveis aleatórias Q (1.14) e D^* (1.13) dependem fortemente da linearidade assintótica do gradiente da dispersão (Jureckova, 1971). Assim, os autores

Jaeckel (1972) e McKean e Hettmansperger (1976) que estabeleceram estas propriedades adotaram as suposições técnicas de Jureckova. A mais forte dessas suposições é sobre a matriz de desenho (ver suposição X3). A razão desta suposição é discutida na Seção 1.3. Ampliando o resultado de Jureckova (1969), para o modelo linear multiparamétrico, Kraff e Van Eeden (1972) fizeram uma suposição diferente sobre a matriz de desenho; (ver suposição X4). Eles conseguiram um resultado comparável à própria extensão de Jureckova (1971). Portanto, o resultado de linearidade assintótica do gradiente de dispersão de Jureckova deve permanecer válido sob as condições de Kraft e Van Eeden.

Neste trabalho, adotamos a maior parte das suposições de Jureckova, mas no Capítulo 2 mostraremos que para escores de Wilcoxon (1.6) podemos desistir de X3 e ainda obter a propriedade de linearidade assintótica do gradiente de dispersão, inicialmente provada por Jureckova (1971). Alguns autores fizeram desenvolvimentos teóricos posteriores baseados em X3 somente através desta propriedade de linearidade assintótica do gradiente, portanto, os resultados teóricos continuam válidos sem a suposição X3.

Finalmente, listamos as suposições de Jureckova assim como também as suposições feitas pelos outros autores e resumimos os requerimentos de vários procedimentos que foram discutidos na Seção 1.1. Denotaremos as suposições sobre a função de distribuição por F , aquelas sobre a matriz de desenho por X e as suposições sobre a função geradora de escores por S .

Suposição F1 : $F(\cdot)$ tem densidade $f(\cdot)$ com informação de Fisher finita, isto é, $f(\cdot)$ é absolutamente contínua e $\int [f'(y) / f(y)]^2 f(y) dy < \infty$. Sem perda de generalidade $F(0) = 1/2$.

Suposição F2: $f(\cdot)$ é uma função par, isto é, $f(y) = f(-y)$.

Suposição X1: A matriz de desenho $[1, X]$ tem posto completo.

Suposição X2: Σ é positiva definida e é o limite de $n^{-1} X'_C X_C$ quando $n \rightarrow \infty$.

Suposição X3 (Jureckova, 1971): $X = X^{(1)} + X^{(2)}$, $X^{(1)}$ tem colunas não decrescentes, $X^{(2)}$ tem colunas não crescentes. Esta decomposição é um pouco diferente

daquela que aparece em Jureckova, mas as duas versões são equivalentes.

Para $C^{(m)} = X^{(m)} - 1\bar{X}^{(m)}$ ($m = 1, 2$) e para cada $j = 1, \dots, p$ uma das duas condições seguintes devem ser satisfeitas:

(a) a j -ésima coluna de $C^{(m)}$ é zero para todas as linhas exceto em um número n finito; ou

(b) a j -ésima coluna de $C^{(m)}$ é diferente de zero para todas as linhas exceto em um número n finito, $n^{-1} \sum_{i=1}^n (c_{ij}^{(m)})^2$ é limitado e as colunas satisfazem a condição de Noether

$$\lim_{n \rightarrow \infty} \frac{\max (c_{ij}^{(m)})^2}{\sum_{i=1}^n (c_{ij}^{(m)})^2} = 0.$$

Suposição X4 (Kraft e Van Eeden, 1972): Cada coluna de X_C (1.12) satisfaz a condição de Noether

$$\lim_{n \rightarrow \infty} \frac{\max x_{ijc}^2}{\sum_{i=1}^n x_{ijc}^2} = 0$$

para algum par de colunas $x_c^{(j)}$ e $x_c^{(k)}$ ($j \neq k$) de X_C . Existe $\gamma_{jk} \neq 0$ e N tal que $n > N$ implica que $x_c^{(j)}$ e $x_c^{(j)} + \gamma_{jk} x_c^{(k)}$ são ordenados similarmente.

Dois vetores são ordenados similarmente se $(z_i - z_j)(w_i - w_j) \geq 0$ para cada par (i, j) .

Suposição S1: A função geradora de escores $\phi(\cdot)$ é uma função não constante quadrado integrável em $(0, 1)$.

Os escores podem ser obtidos como

$$a(i) = \phi(i / (N + 1)) \quad \text{ou} \quad a(i) = E(\phi(U^{(i)}))$$

onde $U^{(i)}$, é o i -ésima estatística de ordem de uma amostra aleatória de uma distribuição uniforme definida no intervalo $(0, 1)$ de tamanho n .

Suposição S2: $\phi(\cdot)$ é a diferença de duas funções geradoras de escores que satisfazem os requerimentos de S1 e os escores são gerados como em S1.

Suposição S3: $\int_0^1 \phi(u)du = 0$ e $\int_0^1 \phi^2(u)du = 1$.

A função geradora de escores que satisfaz S1 ou S2 pode ser sempre padronizada sem afetar as propriedades da estimativa resultante e a estatística de teste.

Suposição S4: $\phi(\cdot)$ é tal que $\sum_{i=1}^n a(i) = 0$ para todo n .

Suposição S5: $\phi(u) = -\phi(1-u)$.

Note que a suposição S5 é uma condição suficiente para S4.

Agora resumiremos as suposições requeridas de vários procedimentos: a normalidade assintótica da estimativa proposta por Jureckova (1971) requer F1, X1-X3, S2 e S3. Jaeckel(1972) requer de F1, X1-X3, S1, S3 e S4 para estabelecer a equivalência assintótica de sua estimativa com a estimativa de Jureckova. Como a propriedade assintótica da variável Q (1.14) depende unicamente da distribuição de $\hat{\beta}$, qualquer um dos dois conjuntos de suposições será suficiente, dependendo de qual destas estimativas esteja sendo usada. As propriedades assintóticas da variável D^* (1.15) são derivadas diretamente do resultado de Jaeckel (1972); portanto, F1 X1-X3, S1, S3, e S4 são necessárias. No Capítulo 2 mostramos que para escores de Wilcoxon, X3 pode ser eliminada de cada uma das listas anteriores. No caso de Q ou D^* , na prática, é exigida uma estimativa de τ . A prova da consistência de τ^* (1.17) requer das suposições F1 e S5. No Capítulo 4 desenvolve-se uma estimativa alternativa para o caso de escores de Wilcoxon sem esta condição. Finalmente a normalidade assintótica de $(\alpha^*, \hat{\beta})'$ onde α^* é a estimativa de α_0 proposta por McKean e Hettmansperger (1978) também depende de F2 e S5. No Capítulo 5 propomos uma estimativa de α_0 que não exige estas suposições.

1.3 Estratégias Básicas e Lemas

Nesta seção descrevemos uma estratégia para resolver um tipo de problema que encontraremos repetidamente nos capítulos seguintes.

Suponhamos que estamos interessados no comportamento assintótico de $T(U, V)$, onde $U \xrightarrow{P} 0$. Além disso, suponhamos conhecer o comportamento assintótico de $T(0, V)$. Por exemplo, podemos estar interessados em alguma função $W(\cdot)$ dos resíduos, $\hat{e}_i = Y_i - \hat{\alpha} - x'_i \hat{\beta}$, depois que ajustamos um modelo linear. Como

$$\begin{aligned}\hat{e}_i &= (Y_i - \alpha_0 - x'_i \beta_0) - (\hat{\alpha} - \alpha_0 - x'_i (\hat{\beta} - \beta_0)) \\ &= e_i - (\hat{\alpha} - \alpha_0 - x'_i (\hat{\beta} - \beta_0)),\end{aligned}$$

podemos considerar $W(\cdot)$ como $W \left[\begin{pmatrix} \hat{\alpha} - \alpha_0 \\ \hat{\beta} - \beta_0 \end{pmatrix}; e \right]$, com $\begin{pmatrix} \hat{\alpha} - \alpha_0 \\ \hat{\beta} - \beta_0 \end{pmatrix}$, que converge para zero em probabilidade, e $e = (e_1, \dots, e_n)'$ que fazem os papéis de U e V respectivamente. Então estamos aptos para determinar seu comportamento assintótico dado que $W(0, e)$ é uma função dos erros e_i 's, que são independentes e identicamente distribuídos. Em geral, se mostramos que

$$T(U, V) - T(0, V) \xrightarrow{P} 0, \quad (1.21)$$

então podemos transferir o comportamento de $T(0, V)$ para $T(U, V)$. Talvez a forma mais direta de estabelecer (1.21) é mostrando que $E[T(U, V) - T(0, V)]^2$ converge para zero quando $n \rightarrow \infty$. Não obstante, em muitos casos de interesse, o problema de cálculo desse momento é impossível; assim seria necessária alguma aproximação menos direta.

Para o propósito desta seção usamos a norma $\|\Delta\| = \max_i |\Delta_i|$ e para simplificar a notação escrevemos frequentemente como $T(U)$ ao invés de $T(U, V)$.

Se $T(b)$ é particularmente "bem comportado", o seguinte lema resolve o problema.

Lema 1.3.a: Suponha-se $U \xrightarrow{P} 0$ e $T(b)$ equicontínuo em $b = 0$. Então

$$T(U, V) - T(0, V) \xrightarrow{P} 0$$

Prova: Sejam $\epsilon > 0$ e $\gamma > 0$ dados. Como $T(b)$ é equicontínuo em 0, existe um $\delta > 0$ tal que, para todo n $\|b\| \leq \delta$, implica $|T(b) - T(0)| < \epsilon$. Como $U \xrightarrow{P} 0$, existe um N tal que $n \geq N$ a qual implica $P\{\|U\| \leq \delta\} \geq 1 - \gamma$. Mas $\|U\| \leq \delta$ implica que $|T(U) - T(0)| < \epsilon$. Portanto para $n \geq N$ implica

$$P\{|T(U) - T(0)| < \epsilon\} \geq 1 - \gamma.$$

Desta forma o lema está demonstrado.

Infelizmente, muitas estatísticas de interesse não são uniformemente contínuas, sem falar de equicontinuidade. Outra maneira natural de abordar o problema é provando que

$$\sup_{\|b\| \leq B} |T(b/n^{1/2}) - T(0)| \xrightarrow{P} 0. \quad (1.22)$$

Então, se $n^{1/2}(U)$ fosse limitada em probabilidade, poderíamos ter

$$T(U) - T(0) \xrightarrow{P} 0.$$

Uma primeira tentativa para estabelecer (1.22) é obter

$$E [T(b/n^{1/2}) - T(0)]^2 \rightarrow 0$$

uniformemente para $\|b\| \leq B < \infty$. Isto serve para mostrar que, para $\forall \epsilon > 0$,

$$\sup_{\|b\| \leq B} P\{|T(b/n^{1/2}) - T(0)| \geq \epsilon\} \rightarrow 0, \quad (1.23)$$

através da desigualdade de Chebychev. Mas, geralmente, (1.23) não implica (1.22) e

assim não responde a nossa necessidade.

Alguns artifícios úteis, embora indiretos, para estabelecer (1.22) são encontrados nos trabalhos de Jureckova (1969, 1971). O seguinte lema, para o caso em que b é unidimensional, é uma ajuda importante.

Lema 1.3.b: Suponha que podemos escrever

$$T(b, V) = S(b, V) + c_n b \quad (1.24)$$

(para simplificar a notação, seja $S(b) = S(b, V)$),

onde $S(b)$ é monótona em b e c_n é uma constante positiva tal que $|c_n/n^{1/2}| \leq c < \infty$. Então,

$$T(b/n^{1/2}, V) - T(0, V) \xrightarrow{P} 0 \quad (1.25)$$

para todo b fixado implica que cada $B < \infty$ fixo

$$\sup_{|b| \leq B} |T(b/n^{1/2}) - T(0)| \xrightarrow{P} 0, \text{ quando } n \rightarrow \infty. \quad (1.26)$$

Prova: Sejam $\epsilon > 0$, $\gamma > 0$ dados. Particionamos $[-B, B]$ em d intervalos: $-B \leq b_0 < b_1 < \dots < b_d = B$ tal que $c|b_i - b_{i-1}| < \epsilon/2$ para $i = 1, \dots, d$. Como existem um número grande b_i podemos achar um N tal que

$$P \left\{ \max_i |T(b_i/n^{1/2}) - T(0)| < \epsilon/2 \right\} \geq 1 - \gamma,$$

para todo $n \geq N$.

Suponha que $S(b)$ é não decrescente se $|b| \leq B$ então $b_{i-1} \leq b \leq b_i$ para algum i .

Se $T(b/n^{1/2}) - T(0) \geq 0$, então

$$|T(b/n^{1/2}) - T(0)| = T(b/n^{1/2}) - T(0) = S(b/n^{1/2}) + c_n b/n^{1/2} - T(0) \quad (1.27)$$

$$\begin{aligned}
&\leq S(b_i/n^{1/2}) + c_n b_i/n^{1/2} + c_n b/n^{1/2} - c_n b_i/n^{1/2} - T(0) \\
&\leq \left| S(b_i/n^{1/2}) + c_n b_i/n^{1/2} - T(0) \right| + |b - b_i| \times |c_n|/n^{1/2} \\
&\leq \max_i \left| T(b_i/n^{1/2}) - T(0) \right| + \epsilon/2.
\end{aligned}$$

Similarmente, se $T(b/n^{1/2}) - T(0) < 0$ então

$$\begin{aligned}
\left| T(b/n^{1/2}) - T(0) \right| &= -T(b/n^{1/2}) + T(0) = -S(b/n^{1/2}) - c_n b/n^{1/2} + T(0) \quad (1.28) \\
&\leq -S(b_{i-1}/n^{1/2}) - c_n b_{i-1}/n^{1/2} + T(0) + c_n b_{i-1}/n^{1/2} - c_n b_i/n^{1/2} \\
&\leq \left| S(b_{i-1}/n^{1/2}) + c_n b_{i-1}/n^{1/2} - T(0) \right| + |b_{i-1} - b| \times |c_n|/n^{1/2} \\
&\leq \max_i \left| T(b_i/n^{1/2}) - T(0) \right| + \epsilon/2
\end{aligned}$$

Argumento similar é valido quando $S(b)$ é não crescente. Assim

$$\left| T(b/n^{1/2}) - T(0) \right| \leq \max_i \left| T(b_i/n^{1/2}) - T(0) \right| + \epsilon/2 \quad (1.29)$$

para todo b , tal que $|b| \leq B$. Portanto

$$\sup_{|b| \leq B} \left| T(b/n^{1/2}) - T(0) \right| \leq \max_i \left| T(b_i/n^{1/2}) - T(0) \right| + \epsilon/2, \quad (1.30)$$

de maneira que

$$P \left\{ \sup_{|b| \leq B} \left| T(b/n^{1/2}) - T(0) \right| < \epsilon \right\} \geq 1 - \gamma,$$

com isto o lema está provado.

Observe que esta prova permite a possibilidade de que $S(b)$ seja não decrescente para algum b , e não crescente para outros.

Este lema é facilmente extendido para o caso de um vetor $b_{p \times 1}$.

Lema 1.3.c: Suponha que podemos escrever

$$T(b, V) = S(b, V) + c'_n b \quad (1.31)$$

onde $S(b)$ é monótona em cada um dos componentes de b e c_n é um vetor $p \times 1$ de constante tal que $\|c_n/n^{1/2}\| \leq c < \infty$, então

$$T(b/n^{1/2}) - T(0) \xrightarrow{p} 0 \quad (1.32)$$

para todo b fixado, implica que cada $B < \infty$

$$\sup_{\|b\| \leq B} |T(b/n^{1/2}) - T(0)| \xrightarrow{p} 0. \quad (1.33)$$

Como em muitas situações de interesse conhecemos que $n^{1/2}(U)$ converge em distribuição para alguma variável, $n^{1/2}(U)$ é tipicamente limitado em probabilidade. Então se podemos escrever $T(b)$ na forma necessária (1.31), podemos mostrar que (1.32) é satisfeito provando que

$$E [T(b/n^{1/2}) - T(0)]^2 \longrightarrow 0$$

para todo b fixado. Então uma aplicação do Lema 1.3.c com os lemas seguintes dá o resultado desejado (1.21).

Lema 1.3.d: Se $n^{1/2}(U)$ é limitado em probabilidade. Então sob as condições do lema 1.3.c

$$T(U, V) - T(0, V) \xrightarrow{p} 0$$

Prova: Como $n^{1/2}(U)$ é limitado em probabilidade, para cada $\gamma > 0$,

existe um $B < \infty$ e N_1 , tal que para $n \geq N_1$, implica que

$$P \left\{ \|n^{1/2}(U)\| \leq B \right\} \geq 1 - \gamma/2.$$

Do Lema 1.3.c, para cada $\epsilon > 0$ existe um N_2 tal que para $n \geq N_2$, então

$$P \left\{ \sup_{\|b\| \leq B} |T(b/n^{1/2}) - T(0)| \leq \epsilon \right\} \geq 1 - \gamma/2.$$

Seja $N = \max\{N_1, N_2\}$; então para $n \geq N$

$$P \left\{ \|n^{1/2}(U)\| \leq B, \sup_{\|b\| \leq B} |T(b/n^{1/2}) - T(0)| < \epsilon \right\} \geq 1 - \gamma.$$

Assim

$$P \{ |T(U) - T(0)| \leq \epsilon \} \geq 1 - \gamma,$$

e o qual lema esta provado.

Se $T(b)$ não pode ser escrito diretamente na forma (1.31) requerida para aplicar o Lema 1.3.c, mas é possível escrever $T(b) = T_1(b) + T_2(b) + \dots + T_k(b)$, onde cada $T_m(b)$, $m = 1, 2, \dots, k$ é da forma como é exigida em (1.31). Então se

$$T_m(b/n^{1/2}) - T_m(0) \xrightarrow{P} 0$$

para todo b fixado e $m = 1, 2, \dots, k$, o Lema 1.3.d produz $T_m(U) - T_m(0) \xrightarrow{P} 0$ para cada $m = 1, 2, \dots, k$. Mas

$$T(U) - T(0) = \sum_{m=1}^k (T_m(U) - T_m(0)).$$

Outro artifício empregado por Jureckova (1971) envolve a obtenção de $T^*(b'_1 b'_2)'$ de modo que $T^*(b' b')' = T(b)$ e tal que $T^*(b'_1 b'_2)'$ é uma função que toma valores de um

vetor $2p \times 1$, e satisfaz as condições do Lema 1.3.c. Neste caso o lema 1.3.d , produz

$$T^*(U'U')' - T^*(0'0')' \xrightarrow{p} 0,$$

já que $n^{1/2}(U'U')'$ é limitado em probabilidade. Mas

$$T^*(U'U')' - T^*(0'0')' = T(U) - T(0),$$

o qual converge para zero em probabilidade quando $n \rightarrow \infty$.

Em alguns casos é necessário aplicar ambas técnicas. A prova do resultado de linearidade assintótica em Jureckova (1971) é um bom exemplo com relação a essa prova poderíamos mencionar a necessidade de aplicar estes artifícios, enquanto conservamos outras propriedades, motivantes da adoção da suposições mais forte X3 (Seção 1.2) sobre a matriz de desenho. Especificamente, a estatística de interesse nessa prova pode ser particionada em partes, que satisfazem (1.31) sem X3, mas somente com a propriedade de continuidade (Hajek e Sidak, 1967, Capítulo VI). Que parece essencial para provar (1.32). No Capítulo 2 mostraremos que, para escores de Wilcoxon podemos provar (1.32) sem recorrer à continuidade.

1.4 Notações e Lemas Adicionais

Nesta seção, provaremos alguns lemas aos quais freqüentemente nos referiremos. Adicionalmente definiremos algumas notações a serem usadas nos capítulos seguintes.

Començaremos pela consideração da distribuição de $e_i - e_j$, $i \neq j$. Dado que os e_i 's são i.i.d., esta distribuição é independente de i e j , e portanto a distribuição de $e_1 - e_2$ é a mesma que $e_2 - e_1$, logo a distribuição é simétrica em torno de zero. A função de distribuição acumulada $G(\cdot)$ de $z = e_i - e_j$ é definida por

$$\begin{aligned} G(z) &= P(e_1 - e_2 \leq z) = \int_{-\infty}^{\infty} \int_{-\infty}^{e_2+z} f(e_1) f(e_2) de_1 de_2 \\ &= \int_{-\infty}^{\infty} F(e+z) f(e) de. \end{aligned} \quad (1.34)$$

Os seguintes lemas estabelecem a existência das primeiras três derivadas e os limites de $G(z)$.

Lema 1.4.a

Sob a suposição F1 (Seção 1.2)

(i) A densidade de $G(z)$ é

$$g(z) = \int_{-\infty}^{\infty} f(e+z) f(e) de; \quad (1.35)$$

(ii) $g(z)$ é limitada e absolutamente contínua com derivada

$$g'(z) = \int_{-\infty}^{\infty} f'(e+z) f(e) de; \quad (1.36)$$

(iii) $g'(z)$ é limitada e absolutamente continua com derivada

$$g''(z) = - \int_{-\infty}^{\infty} f''(e+z) f'(e) de \quad (1.37)$$

e

(iv) $g''(z)$ é limitada.

Prova:

Usamos as idéias básicas da prova do Teorema 1.2.3 de Hajek e Sidak (1967).

(i) Usando a continuidade absoluta de $F(e)$, reescrevemos $G(z)$ (1.34) como

$$G(z) = \int_{-\infty}^{\infty} \int_{-\infty}^z f(e+t) f(e) dt de.$$

Como o integrando é não negativo, podemos trocar a ordem da integração para obter

$$G(z) = \int_{-\infty}^z \int_{-\infty}^{\infty} f(e+t)f(e)dtde.$$

Portanto $G(z)$ é absolutamente contínua com derivada (1.35).

(ii) Dado que $f(e)$ é absolutamente contínua, podemos reescrever $g(z)$ (1.35) como

$$g(z) = \int_{-\infty}^{\infty} \int_{-\infty}^z f'(e+t)f(e)dtde \quad (1.38)$$

então tem-se

$$\begin{aligned} g(z) &\leq \int_{-\infty}^{\infty} \int_{-\infty}^z |f'(e+t)| f(e)dtde \\ &\leq \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} |f'(e+t)| f(e)dtde \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} |f'(t)| dt f(e)de; \end{aligned} \quad (1.39)$$

assim pela desigualdade de Jensen e a suposição F1.

$$\begin{aligned} \left[\int_{-\infty}^{\infty} |f'(t)| dt \right]^2 &= \left[\int_{-\infty}^{\infty} |f'(t)/f(t)| f(t) dt \right]^2 \\ &\leq \left[\int_{-\infty}^{\infty} |f'(t)/f(t)|^2 f(t) dt \right] < \infty. \end{aligned} \quad (1.40)$$

Usando isto em (1.39), temos

$$g(z) \leq \left[\int_{-\infty}^{\infty} |f'(t)/f(t)|^2 f(t) dt \right]^{1/2} < \infty, \quad (1.41)$$

de maneira que $g(z)$ é limitada. Além disso de (1.39) e (1.40) podemos concluir que o integrando (1.38) é integrável no espaço produto de modo que, pelo Teorema de Fubini (Haaser & Sullivan, 1972, p.134) podemos trocar a ordem da integração e obter

$$g(z) = \int_{-\infty}^z \int_{-\infty}^{\infty} f'(e+t)f(e)dtde. \quad (1.42)$$

Portanto $g(z)$ é absolutamente continua com derivada (1.36). De (1.40) temos.

$$\begin{aligned} f(e) &= \int_{-\infty}^e f'(t) dt \leq \int_{-\infty}^e |f'(t)| dt \\ &\leq \int_{-\infty}^{\infty} |f'(t)| dt < \infty, \end{aligned} \quad (1.43)$$

portanto $f(e)$ é limitado.

(iii) É fácil ver que $g'(z)$ (1.36) é limitada, e tendo em consideração (1.40) e (1.43), temos

$$\begin{aligned} |g'(z)| &= \left| \int_{-\infty}^{\infty} f'(e+z) f(z) de \right| \\ &\leq \int_{-\infty}^{\infty} |f'(e+z)| f(e) de \\ &\leq \left[\sup_e f(e) \right] \int_{-\infty}^{\infty} |f'(e)| de < \infty. \end{aligned} \quad (1.44)$$

A seguir, consideramos uma mudança de variável em (1.36), $y = e + z$ e usando a continuidade absoluta de $f(e)$;

$$g'(z) = \int_{-\infty}^{\infty} f'(y) f(y-z) dy = - \int_{-\infty}^{\infty} \int_{-\infty}^z f'(y) f'(y-t) dt dy. \quad (1.45)$$

Novamente usando (1.40)

$$\begin{aligned} \int_{-\infty}^{\infty} \int_{-\infty}^z f'(y) f'(y-t) dt dy &\leq \int_{-\infty}^{\infty} \int_{-\infty}^z |f'(y) f'(y-t)| dt dy \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} |f'(-t)| dt |f'(y)| dy < \infty, \end{aligned} \quad (1.46)$$

de modo que podemos trocar a ordem da integração no lado direito de (1.45) para obter

$$g'(z) = - \int_{-\infty}^z \int_{-\infty}^{\infty} f'(y) f'(y-t) dt dy \quad (1.47)$$

fazendo outra mudança de variáveis, $e = y - t$, produz

$$g'(z) = - \int_{-\infty}^z \int_{-\infty}^{\infty} f'(e+t) f'(e) dt dt, \quad (1.48)$$

portanto $g'(z)$ é absolutamente contínua com derivada dada por (1.37).

(iv) Primeiramente, mostraremos que $f'(e)$ é quadrado integrável. Multiplicando e dividindo o integrando $[f'(e)]^2$ por $f(e)$, obtemos

$$\int_{-\infty}^{\infty} [f'(e)]^2 de = \int_{-\infty}^{\infty} \left\{ \frac{f'(e)^2}{f(e)} \right\} f(e) de \quad (1.49)$$

$$\begin{aligned} &\leq [\sup_e f(e)] \int_{-\infty}^{\infty} [f'(e)]^2 / f(e) de \\ &= [\sup_e f(e)] \int_{-\infty}^{\infty} [f'(e)/f(e)]^2 f(e) de < \infty, \end{aligned}$$

por (1.43) e a suposição F1. Agora considere-se

$$|g''(z)| = \left| - \int_{-\infty}^{\infty} f'(e+z) f'(e) de \right| \quad (1.50)$$

$$\begin{aligned} &\leq \int_{-\infty}^{\infty} |f'(e+z) f'(e)| de \\ &\leq \left[\int_{-\infty}^{\infty} [f'(e+z)]^2 de \int_{-\infty}^{\infty} [f'(e)]^2 de \right]^{1/2} \\ &= \int_{-\infty}^{\infty} [f'(e)]^2 de < \infty \end{aligned}$$

pelo uso da desigualdade Cauchy-Schwarz e (1.49).

Assim o lema está provado. Observe-se que obtivemos algumas propriedades de $f(\cdot)$ e $f'(\cdot)$ que são resumidas no seguinte corolário.

Corolário 1.4.1

Sob a suposição F1 (Seção 1.2)

- (i) $f(\cdot)$ é limitada
- (ii) $\int_{-\infty}^{\infty} f^r(y) dy < \infty$, para $r \geq 1$;
- (iii) $\int_{-\infty}^{\infty} |f'(y)| dy < \infty$, e
- (iv) $\int_{-\infty}^{\infty} |f'(y)|^2 dy < \infty$

Prova

(i) Ver (1.43)

(ii) $\int_{-\infty}^{\infty} f^r(y) dy \leq [\sup_y f^{r-1}(y)] \int_{-\infty}^{\infty} f(y) dy < \infty$

(iii) Ver (1.40)

(iv) Ver (1.49).

A seguir definimos algumas notações. Em todo este trabalho usaremos a norma de máximo dada por:

$$\|\delta\| = \max_q |\delta_q|; \quad q = 1, 2, \dots, p, \quad (1.51)$$

onde δ é um vetor $p \times 1$. Denotamos o espaço euclidiano p -dimensional por R^p . Para um escalar Z definimos a parte positiva como Z^+ e a parte negativa como Z^- , isto é,

$$\begin{aligned} Z^+ &= Z, & \text{se } Z &\geq 0, \\ &= 0, & \text{se } Z < 0; \end{aligned} \quad (1.52)$$

$$\begin{aligned} Z^- &= Z, & \text{se } Z < 0, \\ &= 0, & \text{se } Z \geq 0. \end{aligned}$$

Estenderemos esta notação para vetores e matrizes, sobre cada um de seus componentes básicos. E freqüentemente usamos a função indicadora $I\{\text{expressão}\}$, a qual é igual a um se a expressão é verdadeira, zero se for falsa.

Agora voltamos nossa atenção para a matriz de desenho. Modificamos a suposição X2 como segue.

Suposição X5: $n^{-1}[1_n, X]'[1_n, X] \rightarrow \Lambda$, quando $n \rightarrow \infty$, é uma matriz positiva definida.

O Lema seguinte mostra que esta suposição é uma extensão de X2.

Lema 1.4.b

$$X5 \rightarrow X2.$$

Prova: $n^{-1}[1_n, X]'[1_n, X] = \begin{bmatrix} 1 & \bar{X}' \\ \bar{X} & n^{-1}X'X \end{bmatrix}$, onde $\bar{X}'$ é o vetor linha $1 \times p$ das médias

de X . Particionando Λ convenientemente:

$$\Lambda = \begin{bmatrix} 1 & \mu'_x \\ \mu_x & \psi \end{bmatrix}.$$

Seja $A_n = \begin{bmatrix} 1 & \bar{X}' \\ 0 & I \end{bmatrix}$, onde I é a matriz identidade $p \times p$. Note que A_n é invertível e

$$A = \lim_{n \rightarrow \infty} A_n = \begin{bmatrix} 1 & \mu'_x \\ 0 & I \end{bmatrix},$$

que também é invertível. Então temos

$$A'_n \begin{bmatrix} 1 & \bar{X}' \\ \bar{X} & n^{-1}X'X \end{bmatrix} A_n = \begin{bmatrix} 1 & 0 \\ 0 & n^{-1}X'_cX_c \end{bmatrix} \rightarrow A'\Lambda A \quad (1.53)$$

quando $n \rightarrow \infty$. X_c é a matriz de desenho centralizada (1.12). Como

$$A'\Lambda A = \begin{bmatrix} 1 & 0 \\ 0 & \psi - \mu_x \mu'_x \end{bmatrix},$$

como A é invertível $A'\Lambda A$ é positiva definida; assim $\Sigma = \psi - \mu_x \mu'_x$ é também positiva definida. Portanto $n^{-1}X'_cX_c \rightarrow \Sigma$ quando $n \rightarrow \infty$.

A seguir mostraremos que certa soma é limitada.

Lema 1.4.c: Sob a suposição X2, $N^{-1} \sum_{i=1}^n |c_{ij}|$ é limitada onde $c_{ij} = x_{ij} - \bar{x}_j$ é o (i,j) -ésimo elemento de X_c .

Prova: Usando a desigualdade de Jensen, temos

$$\left[n^{-1} \sum_{i=1}^n |c_{ij}| \right]^2 \leq n^{-1} \sum_{i=1}^n c_{ij}^2$$

A quantidade no lado direito é o (i,j) -ésimo elemento de $n^{-1}X_c'X_c$, o qual converge para Σ quando $n \rightarrow \infty$.

Finalmente, a seguinte generalização do Lema de Arnold (1980) nos permite derivar alguns resultados importantes.

Lema 1.4.d: Seja a_i uma seqüência de números tal que $n^{-1} \sum_{i=1}^n a_i$ converge quando $n \rightarrow \infty$. Então, $n^{-1} \max_i |a_i| \rightarrow 0$ quando $n \rightarrow \infty$, $i = 1, \dots, n$.

Lema 1.4.e: Sob suposição X5, $n^{-1} \max_i x_{ij}^2 \rightarrow 0$ então $n^{-1} \max_i c_{ij}^2 \rightarrow 0$ quando $n \rightarrow \infty$.

Prova: Pela suposição, $n^{-1} \sum_{i=1}^n x_{ij}^2$ converge quando $n \rightarrow \infty$, então pelo Lema 1.4.d $n^{-1} \max_i x_{ij}^2 \rightarrow 0$. Além disso $c_{ij}^2 = (x_{ij} - \bar{x}_j)^2 = (x_{ij}^2 - 2x_{ij}\bar{x}_j + \bar{x}_j^2)$ e como $\bar{x}_j$ converge quando $n \rightarrow \infty$ (ver a prova do Lema 1.4.b), existe um $B < \infty$ tal que $|\bar{x}_j| \leq B$ para todo n . Portanto,

$$n^{-1} \max_i c_{ij}^2 \leq n^{-1} \max_i x_{ij}^2 + 2n^{-1} \max_i |x_{ij}| B + n^{-1} B^2 \quad (1.54)$$

O lado direito de (1.54) claramente converge para zero quando $n \rightarrow \infty$.

Em resumo, estes lemas adicionais são importantes, pois eles dão as condições necessárias para a prova dos teoremas enunciados nos seguintes capítulos.

Capítulo 2

LINEARIDADE DO GRADIENTE DE DISPERSÃO

2.1 Propriedade

Nesta seção estabelecemos a propriedade de linearidade assintótica do gradiente de dispersão, primeiramente provada por Jureckova (1971). Apesar de considerar só o caso de escores de Wilcoxon (1.6), a condição sobre a matriz de desenho é bastante simplificada. Na Seção 2.2 discutiremos brevemente as implicações da presente seção.

Antes de formular o teorema precisamos estabelecer algumas notações. Seja $C = X_C$, a matriz de desenho centralizada (1.12), denotaremos a i -ésima linha desta matriz por c_i . Então para escores de Wilcoxon a medida de dispersão dos resíduos pode ser reescrita como:

$$\begin{aligned} D(\beta) &= \sum_{i=1}^n a[R(Y_i - x'_i\beta)](Y_i - x'_i\beta) \\ &= \sum_{i=1}^n a[R(Y_i - c'_i\beta)](Y_i - c'_i\beta) \\ &= \sqrt{12} \sum_{i=1}^n \left[\frac{R(Y_i - c'_i\beta)}{n+1} - \frac{1}{2} \right] (Y_i - c'_i\beta), \end{aligned} \tag{2.1}$$

onde $R(Y_i - c'_i\beta)$ é o posto de $Y_i - c'_i\beta$ entre $Y_1 - c'_1\beta, \dots, Y_n - c'_n\beta$. Então o gradiente de $D(\beta)$ (derivadas parciais) existe, exceto sobre um conjunto de pontos de medida zero em

R^p . Quando o gradiente existe, o j -ésimo componente é dado por:

$$\begin{aligned}\frac{\partial D(\beta)}{\partial \beta_j} &= \sqrt{12} \sum_{i=1}^n \left[\frac{R(Y_i - c'_i \beta)}{n+1} - \frac{1}{2} \right] (-c_{ij}) \\ &= -\sqrt{12}(n+1)^{-1} \sum_{i=1}^n c_{ij} [R(Y_i - c'_i \beta)],\end{aligned}\quad (2.2)$$

para $j = 1, \dots, p$, a segunda igualdade resulta de $\sum_{i=1}^n c_{ij} = 0$. Denotaremos o negativo do j -ésimo elemento deste gradiente por

$$S_j(\beta) = \sqrt{12}(n+1)^{-1} \sum_{i=1}^n c_{ij} [R(Y_i - c'_i \beta)], \quad (2.3)$$

Dado que os postos em (2.3) são invariantes à traslação, podemos reescrever $S_j(\beta)$ convenientemente:

$$S_j(\beta) = \sqrt{12}[n^{1/2}(n+1)]^{-1} \sum_{i=1}^n c_{ij} [R(e_i - c'_i(\beta - \beta_0))], \quad (2.4)$$

onde $e_i = Y_i - c'_i \beta_0$. Observe que os e_i 's são i.i.d. com função de distribuição acumulada $F(\cdot)$.

Finalmente, define-se uma função $S_j^*(b)$ para $b \in R^p$ por:

$$S_j^*(b) = \sqrt{12}[n^{1/2}(n+1)]^{-1} \sum_{i=1}^n c_{ij} [R(e_i - c'_i b)]. \quad (2.5)$$

Observe que

$$S_j^*(0) = S_j(\beta_0) \quad e \quad S_j^*(\beta - \beta_0) = S_j(\beta). \quad (2.6)$$

O objetivo é estabelecer o seguinte teorema, que prova a linearidade assintótica do gradiente da dispersão para os escores de Wilcoxon sem a suposição $X3$. Jureckova (1971) provou esta linearidade considerando $X3$, além de outras suposições para escores gerais.

Teorema 2.1.1 : Sob as suposições F1, X1, e X5 (seções 1.2 e 1.4) para cada $B < \infty$, fixo

$$\sup_{\|b\| \leq B} \left| S_j^*(b/n^{1/2}) - S_j^*(0) + \sqrt{12}\theta(C'c^{(j)})'b/(n+1) \right| \xrightarrow{P} 0 \quad (2.7)$$

onde $c^{(j)}$ é a j -ésima coluna de C , e $j = 1, 2, \dots, p$, $\theta = \int f^2(x)dx$, e $Cc^{(j)}$ é simplesmente a j -ésima coluna de $C'C$.

Prova: Seguindo as ideias básicas de Aubuchon, J. A. e Hettmansperger, T. P. (1989),

utilizaremos o artifício da Seção 1.3 com o objetivo de aplicar o Lema 1.3.c. O argumento requer uma notação adicional, que definimos agora. Seja D_{ki} um vetor em R^{2p} dado por,

$$D_{ki} = \begin{bmatrix} (c_k - c_i)^+ \\ (c_k - c_i)^- \end{bmatrix} = \begin{bmatrix} (x_k - x_i)^+ \\ (x_k - x_i)^- \end{bmatrix}, \quad (2.8)$$

onde $(c_k - c_i)^+$ é a parte positiva do vetor $(c_k - c_i)$ e $(c_k - c_i)^-$ sua parte negativa, de modo que $(c_k - c_i) = (c_k - c_i)^+ + (c_k - c_i)^-$. Além disso, sejam $M_{j1}(\delta)$ e $M_{j2}(\delta)$ funções de R^{2p} em R definidas como segue:

$$M_{j1}(\delta) = \sqrt{12}[n^{1/2}(n+1)]^{-1} \sum_{i=1}^n c_{ij}^+ \sum_{k=1}^n I\{e_k - e_i \leq D'_{ki}\delta\}, \quad (2.9)$$

$$e \quad M_{j2}(\delta) = \sqrt{12}[n^{1/2}(n+1)]^{-1} \sum_{i=1}^n c_{ij}^- \sum_{k=1}^n I\{e_k - e_i \leq D'_{ki}\delta\}.$$

Por exemplo, se $c_{ij}^+ \geq 0$ e o sinal do coeficiente de δ_q em $D'_{ki}\delta$ é o mesmo para todo k e i , então $M_{j1}(\delta)$ e $M_{j2}(\delta)$ são monótonas nas componentes de δ , não decrescentes em algumas componentes, não crescentes nas outras. Também, sejam T_{j1} e T_{j2} vetores em R^{2p} dados por:

$$T_{j1} = -[n(n+1)]^{-1} \sum_{i=1}^n c_{ij}^+ \sum_{k=1}^n D_{ki}, \quad e \quad (2.10)$$

$$T_{j2} = -[n(n+1)]^{-1} \sum_{i=1}^n c_{ij} \sum_{k=1}^n D_{ki}$$

observamos que, embora M_{j1} , M_{j2} , T_{j1} e T_{j2} dependam de n , é suprimido da notação.

Se $b \in R^p$, encontramos

$$\begin{aligned} (T_{j1} + T_{j2}) \begin{pmatrix} b \\ b \end{pmatrix} &= -[n(n+1)]^{-1} \sum_{i=1}^n c_{ij} \sum_{k=1}^n D'_{ki} \begin{pmatrix} b \\ b \end{pmatrix} \\ &= -[n(n+1)]^{-1} \sum_{i=1}^n c_{ij} \sum_{k=1}^n [(c_k - c_i)^+' b + (c_k - c_i)^-' b] \\ &= -[n(n+1)]^{-1} \sum_{i=1}^n c_{ij} \sum_{k=1}^n [(c_k - c_i)^+ + (c_k - c_i)^-]' b \\ &= -[n(n+1)]^{-1} \sum_{i=1}^n c_{ij} \sum_{k=1}^n (c_k - c_i)' b \\ &= -[n(n+1)]^{-1} \sum_{i=1}^n c_{ij} \left[\sum_{k=1}^n c'_k b - n c'_i b \right] \\ &= [(n+1)]^{-1} \sum_{i=1}^n c_{ij} c'_i b - [n(n+1)]^{-1} \sum_{i=1}^n c_{ij} \left[\sum_{k=1}^n c'_k b \right] \\ &= \sum_{i=1}^n c_{ij} c'_i b / (n+1) = C' c^{(j)} b / (n+1). \end{aligned} \tag{2.11}$$

Retornando ao problema, temos a seguinte decomposição para a quantidade de interesse em (2.7):

$$S_j^*(b/n^{1/2}) - S_j^*(0) + \sqrt{12}\theta(C' c^{(j)})' b / (n+1) \tag{2.12}$$

$$= [M_{j1}(\delta/n^{1/2}) - M_{j1}(0) + \sqrt{12}\theta T'_{j1} \delta] + [M_{j2}(\delta/n^{1/2}) - M_{j2}(0) + \sqrt{12}\theta T'_{j2} \delta],$$

para $\delta = \begin{pmatrix} b \\ b \end{pmatrix}$. A prova do teorema segue a aplicação do Lema 1.3.c para as duas quantidades em (2.12). O seguinte lema estabelece uma das condições necessárias para a aplicação de esse lema.

Lema 2.1.a

$E[M_{jm}(\delta/n^{1/2}) - M_{jm}(0) + \sqrt{12\theta}T'_{jm}\delta]^2 \rightarrow 0$, quando $n \rightarrow \infty$, para qualquer δ fixo em R^{2p} e para $m = 1, 2$.

Prova:

Procederemos pelo cálculo da esperança para $m = 1$.

$$\begin{aligned}
 & E[M_{j1}(\delta/n^{1/2}) - M_{j1}(0) + \sqrt{12\theta}T'_{j1}\delta]^2 \quad (2.13) \\
 &= 12[n(n+1)^2]^{-1} E \left[\sum_{i=1}^n c_{ij}^+ \sum_{k=1}^n \left[I\{e_k - e_i \leq D'_{ki}\delta/n^{1/2}\} - I\{e_k - e_i \leq 0\} - \theta D'_{ki}\delta/n^{1/2} \right] \right. \\
 & \quad \times \left. \sum_{t=1}^n c_{ij}^+ \sum_{w=1}^n \left[I\{e_w - e_t \leq D'_{wt}\delta/n^{1/2}\} - I\{e_w - e_t \leq 0\} - \theta D'_{wt}\delta/n^{1/2} \right] \right] \\
 &= 12[n(n+1)^2]^{-1} \sum_{i \neq k} \sum_{t \neq w} c_{ij}^+ c_{ij}^+ E[(\cdot)(\cdot)],
 \end{aligned}$$

já que $D_{ii} = 0$.

Os termos nesta quádrupla somatória podem ser divididos em três grupos onde cada um será tratado separadamente. O primeiro tem termos com dois pares de sub-índice iguais, isto é, $i = t$ e $k = w$ ou $i = w$ e $k = t$; como i e k podem ser escolhidos de $n(n-1)$ maneiras, então neste grupo temos $2n(n-1)$ termos. O segundo grupo em um par de sub-índices iguais, isto é, $i = t$ ou $i = w$ ou $k = t$ ou $k = w$; portanto três distintos sub-índices podem ser escolhidos de $n(n-1)(n-2)$ formas, desta forma neste grupo existem $4n(n-1)(n-2)$ termos. Finalmente, o terceiro não tem sub-índices iguais, e temos $n(n-1)(n-2)(n-3)$ termos neste grupo. Estas contagens estão resumidas da seguinte forma

<u>Grupo</u>	<u>Descrição</u>	<u>Nº de Termos</u>
1	Dois pares de sub-índices iguais	$2n(n-1)$
2	Um par de sub-índices iguais	$4n(n-1)(n-2)$
3	Todos os sub-índices diferentes	$n(n-1)(n-2)(n-3)$
Total		$n^2(n-1)^2$

Denotaremos a soma sobre Grupo i , como $\sum_{G_i}$, $i = 1, 2$ e 3 .

Agora consideramos o Grupo 1 em (2.13) e mostraremos

$$12[n(n+1)^2]^{-1} \sum_{G_1} c_{ij}^+ c_{ij}^+ E[(\cdot)(\cdot)] \leq n^{-1} \max_i (c_{ij}^2) O(1) \rightarrow 0, \quad (2.14)$$

quando $n \rightarrow \infty$, onde $O(1)$ é limitado. É fácil de verificar que os termos dentro da esperança são limitados uniformemente com probabilidade um. Pois a função indicadora é limitada, e

$$|\theta D'_{ki} \delta / n^{1/2}| \leq \theta 2p \|D_{ki}\| \|\delta\| / n^{1/2}.$$

Mas

$$\|D_{ki}\| = \max_q |x_{kq} - x_{iq}| \leq 2 \max_{k,q} |x_{kq}|,$$

onde $q = 1, \dots, p$. Então temos que

$$|\theta D'_{ki} \delta / n^{1/2}| \leq 2\theta p \|\delta\| \max_q |x_{kq} - x_{iq}| / n^{1/2} \quad (2.15)$$

$$\leq 4\theta p \|\delta\| \max_q \max_k |x_{kq}| / n^{1/2}$$

onde $k = 1, \dots, n$. Mas $\max_k |x_{kq}| / n^{1/2} \rightarrow 0$ (Lema 1.4.e). Portanto, este termo em qualquer dos multiplicandos em (2.13) converge para zero uniformemente. Por outro lado existem termos de ordem $O(n^2)$ em $\sum_{G_1}$, e $c_{ij}^+ c_{ij}^+ \leq \max_i (c_{ij}^2)$, portanto temos (2.14) como desejamos.

Considerando agora o Grupo 2, o fazendo referência a (2.13) podemos escrever,

$$12[n(n+1)^2]^{-1} \sum_{G_2} c_{ij}^+ c_{ij}^+ E[(\cdot)(\cdot)] \quad (2.16)$$

$$\begin{aligned} &= 12[n(n+1)^2]^{-1} \sum_{G_2} c_{ij}^+ c_{ij}^+ E \left[\left(I\{e_k - e_i \leq D'_{ki} \delta / n^{1/2}\} - I\{e_k - e_i \leq 0\} \right) \right. \\ &\quad \times \left. \left(I\{e_w - e_t \leq D'_{wt} \delta / n^{1/2}\} - I\{e_w - e_t \leq 0\} \right) \right] \end{aligned}$$

$$\begin{aligned}
& +12[n(n+1)^2]^{-1} \sum_{G2} c_{ij}^+ c_{ij}^+ E \left[-\theta D'_{ki} \delta / n^{1/2} \left(I\{e_w - e_t \leq D'_{wt} \delta / n^{1/2}\} \right. \right. \\
& \left. \left. - I\{e_w - e_t \leq 0\} - \theta D'_{wt} \delta / n^{1/2} \right) - \theta D'_{wt} \delta / n^{1/2} \left(I\{e_k - e_i \leq D'_{ki} \delta / n^{1/2}\} \right. \right. \\
& \left. \left. - I\{e_k - e_i \leq 0\} - \theta D'_{ki} \delta / n^{1/2} \right) - \theta^2 (D'_{ki} \delta) (D'_{wt} \delta) / n \right] \\
& = Q_1 + Q_2.
\end{aligned}$$

Como

$$|Q_1| \leq 12[n(n+1)^2]^{-1} \sum_{G2} c_{ij}^+ c_{ij}^+ E[|.| \times |.|] \quad (2.17)$$

$$|Q_1| \leq 12[n(n+1)^2]^{-1} \sum_{G2} c_{ij}^+ c_{ij}^+ [E|.| \times E|.]^{1/2},$$

pela desigualdade de Cauchy-Schwarz, note que em (2.17) $|I\{.\} - I\{.\}|$ é igual ao seu quadrado. Além disso

$$\begin{aligned}
& E \left| I\{e_k - e_i \leq D'_{ki} \delta / n^{1/2}\} - I\{e_k - e_i \leq 0\} \right| \quad (2.18) \\
& = \left| G(D'_{ki} \delta / n^{1/2}) - G(0) \right| \\
& = \left| D'_{ki} \delta / n^{1/2} \right| G'(\xi_{ki}),
\end{aligned}$$

onde $|\xi_{ki}| \leq |D'_{ki} \delta / n^{1/2}|$. Como $G'(\cdot)$ é limitada (Lema 1.4.a) e $|D'_{ki} \delta / n^{1/2}|$ converge para zero uniformemente para todo (k, i) , quando $n \rightarrow \infty$, pela análise do Grupo 1. Além disso,

$$n^{-3} \sum_{G2} c_{ij}^+ c_{ij}^+ \leq n^{-2} 4 \sum_{i=1}^n \sum_{j=1}^n c_{ij}^+ c_{ij}^+ \leq \left[n^{-1} 2 \sum_{i=1}^n |c_{ij}| \right]^2, \quad (2.19)$$

a qual é limitada (Lema 1.4.c). Assim o lado direito de (2.17) converge para zero quando $n \rightarrow \infty$. Portanto Q_1 converge para zero.

Agora consideremos Q_2 ,

$$Q_2 = 12[n(n+1)^2]^{-1} \sum_{G2} c_{ij}^+ c_{ij}^+ \left[-\theta D'_{ki} \delta / n^{1/2} \left(G(D'_{wt} \delta / n^{1/2}) - G(0) - \theta D'_{wt} \delta / n^{1/2} \right) \right] \quad (2.20)$$

$$\begin{aligned}
& -\theta D'_{wt}\delta/n^{1/2} \left(G(D'_{ki}\delta/n^{1/2}) - G(0) - \theta D'_{ki}\delta/n^{1/2} \right) - \theta^2 (D'_{ki}\delta)(D'_{wt}\delta)/n \Big] \\
& = 12[n(n+1)^2]^{-1} \sum_{G2} c_{ij}^+ c_{ij}^+ \left[-\theta D'_{ki}\delta/n^{1/2} (D'_{wt}\delta/n^{1/2})^2 G''(\xi_{wt})/2 \right. \\
& \quad \left. - \theta D'_{wt}\delta/n^{1/2} \left((D'_{ki}\delta/n^{1/2})^2 G''(\xi_{ki})/2 \right) - \theta^2 (D'_{ki}\delta)(D'_{wt}\delta)/n \right]
\end{aligned}$$

onde $|\xi_{wt}| \leq |D'_{wt}\delta/n^{1/2}|$, lembre que $G'(0) = \int_{-\infty}^{\infty} f^2(y)dy = \theta$. Como $G''(\cdot)$ é limitada (Lema 1.4.a), $n^{-3} \sum_{G2} c_{ij}^+ c_{ij}^+$ é limitada (2.19), e $|D'_{ki}\delta/n^{1/2}| \rightarrow 0$ uniformemente em todo (k, i) , portanto temos que $Q_2 \rightarrow 0$. Voltando para (2.16) vemos que a contribuição dos termos no Grupo 2 converge para zero quando $n \rightarrow \infty$.

Finalmente considerando os termos do Grupo 3, os multiplicandos são independentes, portanto, a esperança do produto é a esperança dos fatores:

$$\begin{aligned}
& 12[n(n+1)^2]^{-1} \sum_{G3} c_{ij}^+ c_{ij}^+ E[(\cdot)(\cdot)] \tag{2.21} \\
& = 12[n(n+1)^2]^{-1} \sum_{G3} c_{ij}^+ c_{ij}^+ E(\cdot)E(\cdot).
\end{aligned}$$

Considerando uma das esperanças separadamente:

$$\begin{aligned}
& E \left[I\{e_k - e_i \leq D'_{ki}\delta/n^{1/2}\} - I\{e_k - e_i \leq 0\} - \theta D'_{ki}\delta/n^{1/2} \right] \tag{2.22} \\
& = G(D'_{ki}\delta/n^{1/2}) - G(0) - G'(0)D'_{ki}\delta/n^{1/2} \\
& = [D'_{ki}\delta/n^{1/2}]^2 G''(\xi_{ki})/2,
\end{aligned}$$

onde $|\xi_{ki}| \leq |D'_{ki}\delta/n^{1/2}|$. Assim o valor absoluto da somatória em (2.21) é:

$$\begin{aligned}
& \left| 12[n(n+1)^2]^{-1} \sum_{G3} c_{ij}^+ c_{ij}^+ \left[[D'_{ki}\delta/n^{1/2}]^2 G''(\xi_{ki})/2 \right] \times \left[[D'_{wt}\delta/n^{1/2}]^2 G''(\xi_{wt})/2 \right] \right| \tag{2.23} \\
& \leq 3[\sup_z G''(z)]^2 \left[n^{-1/2} \max_i |c_{ij}| \right]^2 [n(n+1)]^{-2} \sum_{G3} (D'_{ki}\delta)^2 (D'_{wt}\delta)^2,
\end{aligned}$$

como $G'''(\cdot)$ é limitada (Lema 1.4.a), e $n^{-1/2} \max_i |c_{ij}|$ converge para zero quando $n \rightarrow \infty$ (Lema 1.4.e), e

$$\begin{aligned}
n^{-4} \sum_{G3} [D'_{ki} \delta]^2 [D'_{wt} \delta]^2 &\leq \left[n^{-2} \sum_{k=1}^n \sum_{i=1}^n [D'_{ki} \delta]^2 \right]^2 \\
&= \left[n^{-2} \sum_{k=1}^n \sum_{i=1}^n \left[\sum_{q=1}^{2p} D_{kiq} \delta q \right]^2 \right]^2 \\
&\leq \left[n^{-2} \sum_{k=1}^n \sum_{i=1}^n \left[\sum_{q=1}^{2p} D_{kiq}^2 \right] \left[\sum_{q=1}^{2p} \delta_q^2 \right] \right]^2 \\
&\leq \left[2p \|\delta\|^2 n^{-2} \sum_{k=1}^n \sum_{i=1}^n \sum_{q=1}^p [x_{kq} - x_{iq}]^2 \right]^2 \\
&= \left[2p \|\delta\|^2 \sum_{q=1}^p n^{-2} \sum_{k=1}^n \sum_{i=1}^n [x_{kq} - x_{iq}]^2 \right]^2,
\end{aligned} \tag{2.24}$$

a qual é claramente limitada, já que ambos termos, $n^{-1} \sum_{i=1}^n x_{iq}^2$ e $n^{-2} \sum_{i=1}^n \sum_{k=1}^n x_{kq} x_{iq}$, convergem. Tomando isto em conta, temos que o lado direito de (2.23) converge para zero quando $n \rightarrow \infty$.

Assim mostramos que a soma sobre os termos em cada um dos três grupos converge para zero. Com isto demonstra-se que a esperança no lado direito de (2.13) converge para zero. Portanto estabelecemos o Lema 2.1.a para $m = 1$ e a prova para $m = 2$ é completamente análoga.

Mais um passo preliminar é necessário antes de completar a prova do Teorema 2.1.1.

Lema 2.1.b

$\|T_{jm}\|$ é limitado para $m = 1, 2$.

Prova:

$$\left(\frac{n+1}{n}\right) \|T_{j1}\| = n^{-2} \left\| \sum_{k=1}^n \sum_{i=1}^n c_{ij}^+ \begin{bmatrix} (c_k - c_i)^+ \\ (c_k - c_i)^- \end{bmatrix} \right\| \quad (2.25)$$

$$\leq n^{-2} \max_q \sum_{k=1}^n \sum_{i=1}^n |c_{ij}| \times |c_{kq} - c_{iq}|;$$

para $q = 1, \dots, p$. Como existem no máximo p elementos diferentes de zero em T_{j1} , ambas, $|(c_{kq} - c_{iq})^+|$ e $|(c_{kq} - c_{iq})^-|$ são menores ou iguais que $|(c_{kq} - c_{iq})|$.

Portanto

$$\begin{aligned} \left(\frac{n+1}{n}\right) \|T_{j1}\| &\leq \max_q \left[n^{-2} \sum_{k=1}^n \sum_{i=1}^n |c_{ij}| \times |c_{kq}| + n^{-2} \sum_{k=1}^n \sum_{i=1}^n |c_{ij}| \times |c_{iq}| \right] \\ &= \max_q \left[\left(n^{-1} \sum_{i=1}^n |c_{ij}| \right) \left(n^{-1} \sum_{k=1}^n |c_{kq}| \right) + n^{-1} \sum_{i=1}^n |c_{ij}| \times |c_{iq}| \right] \\ &\leq \max_q \left[\left(n^{-1} \sum_{i=1}^n |c_{ij}| \right) \left(n^{-1} \sum_{k=1}^n |c_{kq}| \right) + \left(n^{-1} \sum_{i=1}^n c_{ij}^2 \right)^{1/2} \left(n^{-1} \sum_{i=2}^n c_{iq}^2 \right)^{1/2} \right], \end{aligned} \quad (2.26)$$

isto pela desigualdade de Cauchy - Schwarz pois as somatórias são todos limitadas (Lema 1.4.b e 1.4.c). Portanto, está provado o lema para $m = 1$. E a prova para T_{j2} é muito similar.

Agora estamos em condições de aplicar o Lema 1.3.c para cada uma das duas quantidades no lado direito de (2.12). O Lema 2.1.a implica que

$$M_{jm}(\delta/n^{1/2}) - M_{jm}(0) + \sqrt{12}\theta T'_{jm} \delta \xrightarrow{p} 0,$$

para qualquer $\delta \in R^{2p}$, $m = 1, 2$. Já que $M_{jm}(b/n^{1/2}) - M_{jm}(0)$ é monótona em qualquer componentes de δ , e por Lema 2.1.b temos que $\|T_{jm}\|$ é limitada, portanto o Lema 1.3.c se aplica para qualquer $B < \infty$ e produz

$$\sup_{\|\delta\| \leq B} \left| M_{jm}(\delta/n^{1/2}) - M_{jm}(0) + \sqrt{12}\theta T'_{jm} \delta \right| \xrightarrow{p} 0, \quad (2.27)$$

para $m = 1, 2$. Mas para $A = \{\delta \in R^{2p}, \delta = (b' \ b')', b \in R^p \text{ e } \|\delta\| \leq B\}$,

$$\sup_{\|\delta\| \leq B} \left| S_j^*(b/n^{1/2}) - S_j^*(0) + \sqrt{12}\theta(C'c^{(j)})'b/(n+1) \right| \quad (2.28)$$

$$\begin{aligned} &= \sup_A \sum_{m=1}^2 \left| M_{jm}(\delta/n^{1/2}) - M_{jm}(0) + \sqrt{12}\theta T'_{jm} \delta \right| \\ &\leq \sum_{m=1}^2 \sup_{\|\delta\| \leq B} \left| M_{jm}(\delta/n^{1/2}) - M_{jm}(0) + \sqrt{12}\theta T'_{jm} \delta \right| \xrightarrow{P} 0; \end{aligned}$$

assim o teorema está provado.

Podemos apresentar este resultado na forma que corresponde mais diretamente ao Teorema 3.1 de Jureckova (1971).

Corolário 2.1.1. Se as condições do Teorema 2.1.1 são satisfeitas, então

$$\sup_{n^{1/2}\|(\beta - \beta_0)\| \leq B} \left| S_j(\beta) - S(\beta_0) + \sqrt{12}\theta(\sigma^{(j)})'(\beta - \beta_0) \right| \xrightarrow{P} 0, \quad (2.29)$$

onde $\sigma^{(j)}$ é a j -ésima coluna de Σ .

Prova:

Note que

$$\begin{aligned} &\sup_{\|b\| \leq B} \left| (C'c^{(j)})b/(n+1) - \sigma^{(j)}b \right| \quad (2.30) \\ &\leq p \left\| (C'c^{(j)})/(n+1) - \sigma^{(j)} \right\| \times B \rightarrow 0 \end{aligned}$$

quando $n \rightarrow \infty$, além disso seja $b = n^{1/2}(\beta - \beta_0)$, e lembre de (2.6), $S_j^*(0) = S_j(\beta_0)$ e $S_j^*(\beta - \beta_0) = S_j(\beta)$. Portanto o corolário está provado.

2.2 Discussão

Talvez a prova da Seção 2.1 seja menos elegante do que a original dada por Jureckova (1969, 1971) e certamente o resultado é menos geral já que tratamos somente dos escores de Wilcoxon; mas eliminamos a suposição $X3$ que é a mais complicada sobre a matriz de desenho. Os métodos da prova de Jureckova utilizam esta suposição de uma maneira básica. Embora McKean e Hettmansperger (1976) considerem como satisfeita esta suposição em quase qualquer aplicação, eles não apresentam nenhuma justificativa para esta afirmação. Assim na prática é melhor ficar sem esta suposição.

Notamos também que a prova do Teorema 4.1 de Jureckova (1971), que estabelece a normalidade assintótica da estimativa proposta para β_0 , conta com a suposição $X3$ somente através do Teorema 3.1 (linearidade assintótica do gradiente de dispersão). Assim se estimamos β_0 pela raiz aproximada do gradiente de $D(\beta)$, conservamos a normalidade assintótica sem a suposição de $X3$. A que formulada no seguinte teorema.

Teorema 2.2.1. Sob as suposições $F1, X1, X2$, e $X5$ (Seção 1.2 e 1.4),

$$n^{1/2}(\hat{\beta} - \beta_0) \xrightarrow{D} Z \sim NMV \left(0, \tau^2 \Sigma^{-1} \right)$$

onde $\hat{\beta}$ é obtida pela solução (aproximada) de

$$S_j(\beta) \doteq 0, \quad j = 1, 2, \dots, p$$

e $S_j(\beta)$ é definido por (2.3).

Prova:

Veja Jureckova (1971).

Adicionalmente, o trabalho de Jaeckel (1972) relaciona a estimativa de β_0 , baseada na minimização da dispersão dos resíduos, com a estimativa de Jureckova (1971) sem usar explicitamente a suposição $X3$, a qual é mostrada usando somente a linearidade assintótica do gradiente de dispersão e a normalidade assintótica de $\hat{\beta}$. E o mesmo é usado

para o teste de hipótese aproximado de β_0 desenvolvido por McKean e Hettmansperger (1976). A prova da consistência de τ^* (1.17) para o parâmetro de escala τ não depende diretamente desta linearidade.

Van Eeden (1972) e Kraft e Van Eeden (1972) utilizaram um resultado análogo para escores posto-sinal. Estes autores fizeram suposições sobre a matriz de desenho de complexidade similar a $X3$, embora os detalhes não apareçam na literatura. Assim os resultados obtidos por estes autores devem permanecer válidos sob $X3$. Além disso supomos que para escores de Wilcoxon, τ^* (1.17) é consistente para τ sem $X3$. O mesmo pode ser dito para a estimativa de α_0 estudada por McKean e Hettmansperger (1978). Finalmente, o desenvolvimento das propriedades da estimativa da janela de τ (Capítulo 4) e da mediana dos resíduos não centralizados $(Y_i - x'_i \hat{\beta})$, como uma estimativa de α_0 (Capítulo 5) não usam esta suposição.

Resumindo, se usamos os escores de Wilcoxon, podemos empregar os procedimentos baseados em postos, estudados por Jureckova (1971), Jaeckel (1972), e McKean e Hettmansperger (1976, 1978) sem a suposição de $X3$ sobre a matriz de desenho.

Capítulo 3

ESTIMAÇÃO DE DENSIDADES

3.1 Caso De Uma Amostra

No presente capítulo apresentaremos uma breve revisão da estimação de densidade. Pois no próximo capítulo, fazemos uso da idéia de estimação de densidade com a finalidade de estimar a densidade envolvida numa integral τ (1.11). Particularmente, estamos interessados na estimativa proposta por Rosemblatt (1956) e estudadas ainda por Parzen (1962) e conhecida como núcleo-estimadores ou estimativa de janela. Amplos estudos sobre esta abordagem de estimação de densidade além de outros podem ser encontrados em Wegman (1972) e Bean e Tsokos (1980). Aqui definimos uma estimativa de janela e mencionamos algumas de suas propriedades.

Suponha que $Y_1, Y_2, \dots, Y_n$ é uma amostra aleatória de uma função de distribuição acumulada $F(\cdot)$ com densidade $f(\cdot)$. Seja $W(\cdot)$ uma densidade quadrado integrável, e h_n uma seqüência de constantes positivas tal que $h_n \rightarrow 0$, e $nh_n \rightarrow \infty$. Então se $f(\cdot)$ é contínua em y ,

$$f_n(y) = (nh_n)^{-1} \sum_{i=1}^n W\left(\frac{y - Y_i}{h_n}\right) \quad (3.1)$$

é uma estimativa consistente de $f(y)$ (Rosemblatt, 1956). A densidade $W(\cdot)$ é chamada janela ou núcleo; h_n é conhecido como o parâmetro de suavização ou largura da janela (bandwidth).

Os núcleos mais utilizados são:

$$W(t) = \frac{15}{16}(1 - t^2)I\{|t| < 1\} \quad (\text{Biquadrado}),$$

$$W(t) = (1 - |t|)I\{|t| < 1\} \quad (\text{Triangular}),$$

$$W(t) = \frac{1}{\sqrt{2\pi}} \exp\{-t^2/2\} \quad (\text{Gaussiano}), \text{ e}$$

$$W(t) = I\{|t| < 1/2\} \quad (\text{Retangular}).$$

Vários pesquisadores estudaram as propriedades assintóticas de $f_n(y)$, incluindo consistência forte e normalidade assintótica. Uma ótima escolha de h_n poderia ser considerando que erro quadrático integrando médio (EQIM)

$$EQIM = E \left[\int_{-\infty}^{\infty} (f_n(y) - f(y))^2 dy \right]. \quad (3.2)$$

seja mínimo. Outras possibilidades foram estudadas por Bean e Tsokos (1982).

Se $W(\cdot)$ é uma densidade uniforme no intervalo $(-1/2, 1/2)$ o núcleo-estimador (3.1) apresenta a forma intuitiva

$$\begin{aligned} f_n(y) &= (nh_n)^{-1} \sum_{i=1}^n I\{-h_n/2 < y - Y_i < h_n/2\} \\ &= (nh_n)^{-1} \sum_{i=1}^n [I\{Y_i < y + h_n/2\} - I\{Y_i < y - h_n/2\}] \\ &= h_n^{-1} [F_n(y + h_n/2) - F_n(y - h_n/2)] \end{aligned} \quad (3.3)$$

onde $F_n(\cdot)$ é a função de distribuição empírica. Assim, $f_n(y)$ é uma derivada empírica da função de distribuição.

3.2 Resíduos e Largura da Janela Aleatória

Nesta seção mostraremos que a estimativa de densidade da forma (3.1) mesmo calculada sobre os resíduos após o ajuste do modelo linear ainda é consistente. Também mostraremos que a largura da janela pode ser uma variável aleatória. Para o objetivo desta seção não diferenciaremos o intercepto de qualquer outro coeficiente de regressão, conseqüentemente a notação aqui difere dos outros capítulos.

Suponha que observamos o vetor $Y = (Y_1, Y_2, \dots, Y_n)'$ de ordem $n \times 1$, com

$$Y = X\beta_0 + e, \quad (3.4)$$

onde X é uma matriz de desenho $n \times p$ de posto completo, β_0 é o vetor $p \times 1$ de coeficientes de regressão, e $e = (e_1, e_2, \dots, e_n)'$ é o vetor de erros aleatórios $n \times 1$, com componentes i.i.d., com função de distribuição absolutamente contínua e tendo $F(0) = 1/2$.

Substituímos as suposições $X1$ e $X5$ por suposições equivalentes para considerar o modelo linear (3.4).

Suposição X6. X tem posto completo.

Suposição X7. $n^{-1}X'X \rightarrow \Lambda$, quando $n \rightarrow \infty$, é uma matriz positiva definida.

Devemos também colocar algumas suposições referentes à forma da janela, largura da janela e assim como à estimativa de β_0 .

Suposição W1. $W(z)$ é uma densidade unimodal, simétrica em torno de zero e satisfaz

$$\int_{-\infty}^{\infty} W^2(u) du < \infty.$$

Suposição H1. h_n é uma seqüência de constantes positivas da forma $h_n = n^{-r}h$, onde $r > 0$ e h é uma constante positiva e fixa. Além disso $0 < r < 1/2$, $(nh_n^2)^{-1}$ é limitada e $nh_n \rightarrow \infty$, como mencionamos na Seção 3.1.

Suposição H2. $\hat{h}_n = n^{-r} \hat{h}$, onde $n^q(\hat{h} - h)$ é limitado em probabilidade para algum $q > 0$.

Suposição B1. $n^{1/2}(\hat{\beta} - \beta_0)$ é limitada em probabilidade (isto é satisfeito desde que $n^{1/2}(\hat{\beta} - \beta_0)$ tenha uma distribuição limite.).

Agora estabelecemos algumas notações.

Definimos

$$\hat{f}(z, \Delta, v) = (nh_n)^{-1} \sum_{i=1}^n W([e_i - z - x'_i \Delta]/k_n) \quad (3.5)$$

onde Δ é um vetor $p \times 1$, $e_i = Y_i - x'_i \beta_0$, x'_i é a i -ésima linha de X e

$$k_n = n^{-q-r} v + h_n. \quad (3.6)$$

A sequência k_n faz o papel de $\hat{h}_n$ (largura da janela aleatória) na prova do seguinte teorema. Note que $n^{q+r}(k_n - h_n) = v$ faz o papel de $n^q(\hat{h} - h)$, e pela suposição H2 é limitada em probabilidade. Além disso

$$\hat{f}(z, 0, 0) = (nh_n)^{-1} \sum_{i=1}^n W([e_i - z]/h_n)$$

é o núcleo-estimador de $f(z)$ baseado em e_i e na largura da janela não aleatória h_n , enquanto

$$\hat{f}(z, (\hat{\beta} - \beta_0), n^q(\hat{h} - h)) = (nh_n)^{-1} \sum_{i=1}^n W([e_i - z - x'_i(\hat{\beta} - \beta_0)]/\hat{h}_n)$$

é um núcleo-estimador de $f(z)$ baseado nos $\hat{e}_i = Y_i - x'_i \hat{\beta}$, e largura da janela aleatória, $\hat{h}_n$. Isto segue facilmente, já que

$$\begin{aligned} e_i - z - x'_i(\hat{\beta} - \beta_0) &= Y_i - x'_i \beta_0 - z - x'_i \hat{\beta} + x'_i \beta_0 \\ &= Y_i - x'_i \hat{\beta} - z = \hat{e}_i - z \end{aligned} \quad (3.7)$$

e

$$k_n = n^{-q-\tau} n^{+q}(\hat{h} - h) + h_n = n^{-\tau} \hat{h} - h_n + h_n = \hat{h}_n. \quad (3.8)$$

Usando a estratégia da Seção 1.3 provaremos que $f_n(z)$ é ainda consistente quando calculada sobre os resíduos após o ajustamento de um modelo linear geral.

Teorema 3.2.1. Se as suposições F1, X6, X7, W1, H1, H2 e B1 são satisfeitas então

$$\hat{f}(z, \hat{\beta} - \beta_0, n^q(\hat{h} - h)) \xrightarrow{p} f(z) \quad (3.9)$$

para qualquer $z \in R$

Prova: Usaremos as sugestões dadas em Aubuchon e Hettmansperger (1989)

Particionamos X em suas partes positiva e negativa: $X = X^+ + X^-$. Seja $A = [X^+, X^-]$, e a'_i a i -ésima linha de A . Então $a_i \in R^{2p}$.

Particionamos também a densidade unimodal $W(\cdot)$ em duas partes:

$$\begin{aligned} W_1(z) &= W(z), & \text{se} & & z \leq 0, \\ &= W(0), & \text{se} & & z > 0; \end{aligned} \quad (3.10)$$

e

$$\begin{aligned} W_2(z) &= 0, & \text{se} & & z \leq 0, \\ &= W(z) - W(0), & \text{se} & & z > 0. \end{aligned}$$

Como $W(z) = W_1(z) + W_2(z)$; $W_1(z)$ e $W_2(z)$ são monótonas em z . Define-se

$$f^*(z, \delta, v) = (nh_n)^{-1} \sum_{i=1}^n W([e_i - z - a'_i \delta]/k_n) \quad (3.11)$$

onde k_n está relacionada com v por (3.6). E se $\delta' = (\Delta' \Delta')$, então

$$f^*(z, \delta, v) = \hat{f}(z, \Delta, v).$$

Além disso, particionamos $f^*(z, \delta, v)$ em duas partes .

$$\begin{aligned} f^*(z, \delta, v) &= (nh_n)^{-1} \sum_{i=1}^n [W_1([e_i - z - a'_i \delta]/k_n) + W(0)f(z)a'_i \delta] \\ &\quad + (nh_n)^{-1} \sum_{i=1}^n [W_2([e_i - z - a'_i \delta]/k_n) - W(0)f(z)a'_i \delta] \\ &= f_1^*(z, \delta, v) + f_2^*(z, \delta, v). \end{aligned} \quad (3.12)$$

Agora o principal problema é mostrar que

$$f_m^*(z, \delta/n^{1/2}, v) - f_m^*(z, 0, 0) \xrightarrow{p} 0, \quad (3.13)$$

para cada δ, v fixos, e para $m = 1, 2$. Observa-se que

$$f_m^*(z, \delta/n^{1/2}, v) - f_m^*(z, 0, 0)$$

satisfaz as condições do Lema 1.3.c, que essencialmente completara a prova.

Estabeleceremos (3.13) por meio do seguinte lema.

Lema 3.2.1. Sob as condições do Teorema 3.1.1, para $m = 1, 2$,

$$E \left[f_m^*(z, \delta/n^{1/2}, v) - f_m^*(z, 0, 0) \right]^2 \longrightarrow 0,$$

para qualquer $\delta \in R^{2p}$ e v .

Prova:

Para $m = 1$, temos

$$E \left[f_1^*(z, \delta/n^{1/2}, v) - f_1^*(z, 0, 0) \right]^2 \quad (3.14)$$

$$= (nh_n)^{-2} \sum_{i=1}^n E \left[W_1([e_i - z - a'_i \delta/n^{1/2}]/k_n) + W(0)f(z)a'_i \delta/n^{1/2} - W_1([e_i - z]/h_n) \right]^2$$

$$\begin{aligned}
& + (nh_n)^{-2} \sum_{i \neq j} E \left[\left(W_1([e_i - z - a'_i \delta / n^{1/2}] / k_n) + W(0)f(z)a'_i \delta / n^{1/2} - W_1([e_i - z] / h_n) \right) \right. \\
& \times \left. \left(W_1([e_j - z - a'_j \delta / n^{1/2}] / k_n) + W(0)f(z)a'_j \delta / n^{1/2} - W_1([e_j - z] / h_n) \right) \right] \\
& = Q_1 + Q_2.
\end{aligned}$$

Usando a definição (3.10) de $W_1(\cdot)$

$$Q_1 = (nh_n)^{-2} \sum_{i=1}^n E \left[W([e_i - z - a'_i \delta / n^{1/2}] / k_n) I\{e_i \leq z + a'_i \delta / n^{1/2}\} - W([e_i - z] / h_n) I\{e_i \leq z\} \right] \quad (3.15)$$

$$\begin{aligned}
& + W(0) \left(I\{e_i > z + a'_i \delta / n^{1/2}\} - I\{e_i > z\} + f(z)a'_i \delta / n^{1/2} \right)^2 \\
& = (nh_n)^{-2} \sum_{i=1}^n \left[\int_{-\infty}^{z+a'_i \delta / n^{1/2}} W^2([y - z - a'_i \delta / n^{1/2}] / k_n) f(y) dy + \int_{-\infty}^z W^2([y - z] / h_n) f(y) dy \right. \\
& - 2 \int_{-\infty}^{\min\{z, z+a'_i \delta / n^{1/2}\}} W([y - z - a'_i \delta / n^{1/2}] / k_n) W([y - z] / h_n) f(y) dy \\
& + 2f(z)W(0)a'_i \delta / n^{1/2} \left(\int_{-\infty}^{z+a'_i \delta / n^{1/2}} W([y - z - a'_i \delta / n^{1/2}] / k_n) f(y) dy \right. \\
& - \left. \int_{-\infty}^z W([y - z] / h_n) f(y) dy \right) + 2f(z)W^2(0)a'_i \delta / n^{1/2} (1 - F(z + a'_i \delta / n^{1/2}) - 1 + F(z)) \\
& + f^2(z)W^2(0)(a'_i \delta)^2 / n + W^2(0) |F(z + a'_i \delta / n^{1/2}) - F(z)| \\
& - 2W(0) \left| \int_{-\infty}^{z+a'_i \delta / n^{1/2}} W([y - z] / h_n) f(y) dy \right| \\
& \left. - 2W(0) \left| \int_{-\infty}^{z+a'_i \delta / n^{1/2}} W([y - z + a'_i \delta / n^{1/2}] / k_n) f(y) dy \right| \right].
\end{aligned}$$

Fazendo a mudança de variável $u = [y - z - a'_i \delta / n^{1/2}] / k_n$ na primeira integral, e mudanças similares nas outras, temos que

$$\begin{aligned}
Q_1 & = (nh_n)^{-2} \sum_{i=1}^n \left[k_n \int_{-\infty}^0 W^2(u) f(k_n u + z + a'_i \delta / n^{1/2}) du + h_n \int_{-\infty}^z W^2(u) f(h_n u + z) du \right. \\
& - 2 \min\{h_n \int_{-\infty}^0 W(u h_n / k_n - a'_i \delta / [n^{1/2} k_n]) W(u) f(h_n u + z) du,
\end{aligned}$$

$$\begin{aligned}
& k_n \int_{-\infty}^0 W(u) W(uk_n/h_n + a'_i \delta / [h_n n^{1/2}]) f(k_n u + z + a'_i \delta / n^{1/2}) du \} \\
& + 2f(z) W(0) a'_i \delta / n^{1/2} \left(k_n \int_{-\infty}^0 W(u) f(k_n u + z + a'_i \delta / n^{1/2}) du - h_n \int_{-\infty}^0 W(u) f(h_n u + z) du \right) \\
& - 2f(z) W^2(0) a'_i \delta / n^{1/2} (a'_i \delta f(\xi_i) / n^{1/2}) + f^2(z) W^2(0) (a'_i \delta)^2 / n \\
& + W^2(0) |a'_i \delta f(\xi_i) / n^{1/2}| - 2W(0) \left| h_n \int_0^{a_i \delta / [h_n n^{1/2}]} W(u) f(h_n u + z) du \right| \\
& - 2W(0) \left| k_n \int_{-a'_i \delta / [k_n / n^{1/2}]}^0 W(u) f(k_n u + z + a'_i \delta / n^{1/2}) du \right| ,
\end{aligned}$$

onde $|\xi_i - z| \leq |a'_i \delta / n^{1/2}|$. Agora usando o fato que $W(\cdot)$ é quadrado integrável (W1) e $f(\cdot)$ é limitada (Corolário 1.4.1), temos para a primeira integral:

$$k_n \int_{-\infty}^0 W^2(u) f(k_n u + z + a'_i \delta / n^{1/2}) du \leq k_n [\sup_y f(y)] \int_{-\infty}^{\infty} W^2(u) du = k_n O(1), \quad (3.16)$$

onde $O(1)$ é limitada. Similarmente, a segunda, sexta e sétima integral são todas $h_n O(1)$; a quinta e oitava são $k_n O(1)$. Para a terceira integral temos

$$\begin{aligned}
& h_n \int_{-\infty}^0 W(u h_n / k_n - a'_i \delta / [n^{1/2} k_n]) W(u) f(h_n u + z) du \\
& \leq h_n [\sup_y f(y)] \left[\int_{-\infty}^{\infty} W^2(h_n u / k_n) du \int_{-\infty}^{\infty} W^2(u) du \right]^{1/2} \\
& = h_n [k_n / h_n]^{1/2} [\sup_y f(y)] \int_{-\infty}^{\infty} W^2(u) du \\
& = (h_n k_n)^{1/2} O(1)
\end{aligned} \tag{3.17}$$

Através de uma análise semelhante, podemos ver que a quarta integral é também $(h_n k_n)^{1/2} O(1)$. Assim todos os oito termos envolvidos na integral em (3.15) convergem para zero uniformemente em todo i . Além disso, temos:

$$\begin{aligned}
|a'_i \delta / n^{1/2}| & \leq p \|a_i\| \times \|\delta\| / n^{1/2} \\
& \leq p \|\delta\| \max_{ij} \|x_{ij}\| / n^{1/2} \rightarrow 0
\end{aligned} \tag{3.18}$$

quando $n \rightarrow \infty$, onde o *máximo* é sobre $i = 1, 2, \dots, n$, e $j = 1, 2, \dots, p$. A convergência para zero segue-se da suposição X7 e o Lema 1.4.d. Assim, lembrando que $f(\cdot)$ é limitada, vemos que os três termos não envolvidos nas integrais em (3.15) todos convergem para zero uniformemente para todo i . Então n^{-1} vezes a somatória em (3.15) converge para zero. Já que pela suposição H1, $(nh_n^2)^{-1}$ é limitada, portanto Q_1 converge para zero.

Agora voltemos nossa atenção para Q_2 (3.14). Como e_i e e_j , $i \neq j$, são independentes temos que a esperança do produto é o produto das esperanças de cada fator.

$$Q_2 = (nh_n)^{-2} \sum_{i \neq j} E \left[W_1([e_i - z - a'_i \delta / n^{1/2}] / k_n) + W(0)f(z)a'_i \delta / n^{1/2} - W_1([e_i - z] / h_n) \right] \quad (3.19)$$

$$\times E \left[W_1([e_j - z - a'_j \delta / n^{1/2}] / k_n) + W(0)f(z)a'_j \delta / n^{1/2} - W_1([e_j - z] / h_n) \right].$$

Considerando uma das esperanças separadamente, e usando a definição (3.10) de $W_1(\cdot)$, temos :

$$\begin{aligned} & E \left(W_1([e_i - z - a'_i \delta / n^{1/2}] / k_n) + W(0)f(z)a'_i \delta / n^{1/2} - W_1([e_i - z] / h_n) \right) \quad (3.20) \\ &= \int_{-\infty}^{z + a'_i \delta / n^{1/2}} W([y - z - a'_i \delta / n^{1/2}] / k_n) f(y) dy + W(0)[1 - F(z + a'_i \delta / n^{1/2})] \\ &+ W(0)f(z)a'_i \delta / n^{1/2} - \int_{-\infty}^z W([y - z] / h_n) f(y) dy - W(0)[1 - F(z)] \\ &= k_n \int_{-\infty}^0 W(u) f(k_n u + z + a'_i \delta / n^{1/2}) du - h_n \int_{-\infty}^0 W(u) f(h_n u + z) du \\ &- W(0)[F(z + a'_i \delta / n^{1/2}) - F(z) - f(z)a'_i \delta / n^{1/2}]. \end{aligned}$$

Lembrando que $f(\cdot)$ é limitada (Corolário 1.4.1), o Teorema de Convergência Dominada fornece

$$\begin{aligned} & \lim_{n \rightarrow \infty} \int_{-\infty}^0 W(u) f(k_n u + z + a'_i \delta / n^{1/2}) du \quad (3.21) \\ &= \int_{-\infty}^0 W(u) \lim_{n \rightarrow \infty} f(k_n u + z + a'_i \delta / n^{1/2}) du \end{aligned}$$

$$= f(z) \int_{-\infty}^0 W(u) du = \frac{f(z)}{2}.$$

Similarmente,

$$\lim_{n \rightarrow \infty} \int_{-\infty}^0 W(u) f(h_n u + z) du = f(z) \int_{-\infty}^0 W(u) du = \frac{f(z)}{2}.$$

Aplicando a expansão em séries de Taylor temos

$$F(z + a'_i \delta / n^{1/2}) - F(z) - f(z) a'_i \delta / n^{1/2} = a'_i \delta (f(\xi_i) - f(z)) / n^{1/2} \quad (3.22)$$

onde $|\xi_i - z| \leq |a'_i \delta / n^{1/2}|$, por (3.18) converge para zero para todo i . Assim a esperança em (3.20) pode ser escrita da seguinte forma

$$(k_n - h_n) \frac{f(z)}{2} + a'_i \delta o(1) / n^{1/2}, \quad (3.23)$$

onde $o(1)$ converge para zero quando $n \rightarrow \infty$.

Como Q_2 (3.19) pode ser escrita como

$$Q_2 = \left[(nh_n)^{-1} \sum_{i=1}^n E[.] \right]^2 - (nh_n)^{-2} \sum_{i=1}^n E^2[.]. \quad (3.24)$$

Como $E^2[.] \leq E[.]^2$, e o fato de que Q_1 (3.15) converge para zero implica que $(nh_n)^{-2} \sum_{i=1}^n E^2[.]$ também converge para zero. Mostramos justamente que

$$(nh_n)^{-1} \sum_{i=1}^n E[.] \quad (3.25)$$

$$= (nh_n)^{-1} \sum_{i=1}^n \left[(k_n - h_n) \frac{f(z)}{2} + a'_i \delta o(1) / n^{1/2} \right].$$

Usando (3.6) e H1,

$$(k_n - h_n) / h_n = n^{-q-r} v / [n^{-r} h] = n^{-q} v / h,$$

que converge para zero. Também temos que $(h_n n^{1/2})^{-1}$ é limitada (H1). E,

$$\begin{aligned} n^{-1} \sum_{i=1}^n |a_i \delta| &\leq \|\delta\| \sum_{i=1}^p n^{-1} \sum_{i=1}^n |x_{ij}| \\ &\leq \|\delta\| \sum_{j=1}^p \left[n^{-1} \sum_{i=1}^n x_{ij}^2 \right]^{1/2} < \infty. \end{aligned} \quad (3.26)$$

Portanto, podemos notar que a expressão em (3.25) converge para zero quando $n \rightarrow \infty$.

Retornando a (3.24) podemos dizer que Q_2 também converge para zero.

Dado que os termos Q_1 e Q_2 em (3.14), convergem para zero, está mostrado a prova do Lema 3.2.1 para $m = 1$; e a prova para $m = 2$ é análoga.

Agora estamos prontos para completar a prova do Teorema 3.2.1. Para $f_1^*(.)$ definida em (3.12),

$$f_1^*(z, \delta/n^{1/2}, v) - f_1^*(z, 0, 0) \quad (3.27)$$

$$\begin{aligned} &= (nh_n)^{-1} \sum_{i=1}^n \left[W_1([e_i - z - a'_i \delta/n^{1/2}]/k_n) - W_1([e_i - z]/h_n) \right] \\ &\quad + W(0)f(z)(nh_n)^{-1} \sum_{i=1}^n a'_i \delta/n^{1/2} \end{aligned}$$

satisfaz as condições do Lema 1.3.c. Isto é,

$$(nh_n)^{-1} \sum_{i=1}^n \left[W_1([e_i - z - a'_i \delta/n^{1/2}]/k_n) - W_1([e_i - z]/h_n) \right]$$

é monótona nos componentes de δ e em v (relembre que $k_n = n^{-q-r}v + h_n$). Pela suposição H1, $(n^{1/2}h_n)^{-1}$ é limitada e

$$\begin{aligned} \left\| n^{-1} \sum_{i=1}^n a_i \right\| &\leq \max_j n^{-1} \sum_{i=1}^n |x_{ij}| \\ &\leq \max_j \left[n^{-1} \sum_{i=1}^n x_{ij}^2 \right]^{1/2} < \infty, \end{aligned} \quad (3.28)$$

onde $j = 1, \dots, p$. Ainda, conhecemos, que $\|W(0)f(z)(nh_n)^{-1} \sum_{i=1}^n a_i/n^{1/2}\|$ é limitada. Portanto, o Lema 3.2.1, através do Lema 1.3.c, implica que

$$\sup |f_1^*(z, \delta/n^{1/2}, v) - f_1^*(z, 0, 0)| \xrightarrow{p} 0,$$

onde *supremo* é sobre $\|\delta\| \leq B$ e $|v| \leq B$. O mesmo é válido para $f_2^*(.)$.

Como $n^{1/2}(\hat{\beta} - \beta_0)$ e $n^q(\hat{h} - h)$ são limitados em probabilidade. Portanto, uma aplicação do lema 1.3.d para $m = 1, 2$ produz

$$f_m^* \left(z, \begin{bmatrix} \hat{\beta} - \beta_0 \\ \hat{\beta} - \beta_0 \end{bmatrix}, n^q(\hat{h} - h) \right) - f_m^*(z, 0, 0) \xrightarrow{p} 0$$

quando $\rightarrow \infty$.

Mas

$$\hat{f}(z, \hat{\beta} - \beta_0, n^q(\hat{h} - h)) - \hat{f}(z, 0, 0) = \sum_{m=1}^2 \left[f_m^* \left(z, \begin{bmatrix} \hat{\beta} - \beta_0 \\ \hat{\beta} - \beta_0 \end{bmatrix}, n^q(\hat{h} - h) \right) - f_m^*(z, 0, 0) \right] \xrightarrow{p} 0.$$

E como

$$\hat{f}(z, 0, 0) \xrightarrow{p} f(z),$$

então $\hat{f}(z, (\hat{\beta} - \beta_0), n^q(\hat{h} - h))$ converge em probabilidade para $f(z)$. Com isto completamos a prova do Teorema 3.2.1.

Agora pelo teorema de Slutsky (Randles e Wolfe 1979, p.424) tem-se o seguinte corolário.

Corolário 3.2.1. Sob as suposições do Teorema 3.2.1.

$$(n\hat{h}_n)^{-1} \sum_{i=1}^n W([Y_i - x_i'\hat{\beta} - z]/\hat{h}_n) \xrightarrow{p} f(z)$$

Prova: Sob a suposição H2, $h_n/\hat{h}_n \xrightarrow{p} 1$. E o teorema estabelece que

$$(nh_n)^{-1} \sum_{i=1}^n W([Y_i - x_i' \hat{\beta} - z]/\hat{h}_n) \xrightarrow{p} f(z).$$

E pelo o Teorema de Slutsky temos que o produto também converge em probabilidade para $f(z)$.

Capítulo 4

ESTIMAÇÃO DE UMA FUNCIONAL DE ESCALA

4.1 Introdução

Vários autores consideraram a possibilidade de estimação de $\tau^{-1} = - \int_0^1 \phi(u) \left[\frac{f'(F^{-1}(u))}{f(F^{-1}(u))} \right] du$ (1.11) no caso de uma amostra, usando estimação de densidades. Se $\phi(u)$ é contínua em $(0, 1)$, existe a 1ª derivada em todo u , e $\lim_{y \rightarrow \pm\infty} f(y)\phi(F(y)) = 0$, então fazendo uma mudança de variável e integrando por partes podemos escrever como $\tau^{-1} = \int_{-\infty}^{\infty} \phi'(F(y))f^2(y)dy$. Para escores de Wilcoxon, na qual limitamos nossa atenção temos que $\phi'(u) = \sqrt{12}$, portanto o problema se reduz à estimação de $\theta = \int_{-\infty}^{\infty} f^2(y)dy$. Bhattacharyya e Roussas (1969), mostraram que uma estimativa consistente é dada por $\int_{-\infty}^{\infty} f_n^2(y)dy$, onde $f_n(y)$ é um núcleo-estimador de $f(y)$. Schuster (1974) estudou convergência forte para a estimativa $\int_{-\infty}^{\infty} f_n(y)dF_n(y)$, onde $F_n(\cdot)$ é a função de distribuição empírica usual. Ahmad (1976) obteve outros resultados de convergência forte para esta segunda estimativa e estabeleceu sua normalidade assintótica. Schweder (1975) derivou a normalidade assintótica de $\int_{-\infty}^{\infty} \phi'(F_n(y))f_n(y)dF_n(y)$; que pode ser relevante para funções escores gerais. Cheng e Serfling (1981) estenderam a abordagem de Bhattacharyya e Roussas para a estimação de $\int_{-\infty}^{\infty} \Psi(y)\phi[F(y)]f^2(y)dy$ por $\int_{-\infty}^{\infty} \Psi(y)\phi[\hat{F}(y)]f_n^2(y)dy$, onde $\hat{F}(y) = \int_{-\infty}^y f_n(z)dz$. Eles obtiveram vários resultados de convergência forte.

Na Seção 4.2, estabelecemos a normalidade assintótica da estimativa $\hat{\theta} = \int_{-\infty}^{\infty} f_n(y)dF_n(y)$

proposta por Schuster (1974) sob algumas condições diferentes das requeridas por Schweder (1975) e Ahmad (1976). Na Seção 4.3 mostraremos que $\hat{\theta}$ é consistente para θ quando é calculada sobre os resíduos após ajuste de um modelo linear geral. Finalmente, na Seção 4.4 alguns cálculos são feitos para comparar $\hat{\theta}$ com outras estimativas de θ , e sugerimos como pode ser escolhida a largura da janela na prática.

Antes de prosseguir, mostraremos que a estimativa estudada por Battacharyya e Roussas (1969) é um caso especial da estimativa considerada por Schuster (1974), Schweder (1975) e Ahmad (1976).

Teorema 4.1.1. Qualquer estimativa do tipo Bhattcharyya e Roussas pode ser expressa como uma estimativa do Schuster.

Prova:

Seja

$$\hat{\theta}_{BR} = \int_{-\infty}^{\infty} f_n^2(y) dy; \quad \text{onde} \quad f_n(y) = (nh_n)^{-1} \sum_{i=1}^n W([Y_i - y]/h_n)$$

para uma densidade $W(\cdot)$, simétrica em torno de zero. Então temos :

$$\begin{aligned} \hat{\theta}_{BR} &= \int_{-\infty}^{\infty} (nh_n)^{-2} \sum_{i=1}^n \sum_{j=1}^n W([Y_i - y]/h_n) W([Y_j - y]/h_n) dy \\ &= (nh_n)^{-2} \sum_{i=1}^n \sum_{j=1}^n \int_{-\infty}^{\infty} W([Y_i - y]/h_n) W([Y_j - y]/h_n) dy. \end{aligned} \quad (4.1)$$

Seja $z = (Y_j - y)/h_n$; então $dz = -dy/h_n$ e $-y = h_n z - Y_j$.

Assim, temos

$$\begin{aligned} \hat{\theta}_{BR} &= n^{-2} h_n^{-1} \sum_{i=1}^n \sum_{j=1}^n \int_{-\infty}^{\infty} W(z + [Y_i - Y_j]/h_n) W(z) dz \\ &= n^{-2} h_n^{-1} \sum_{i=1}^n \sum_{j=1}^n W^*([Y_i - Y_j]/h_n) \end{aligned} \quad (4.2)$$

onde $W^*(z)$ é a convolução de $W(z)$ com $W(-z)$. Como, $W(\cdot)$ é simétrica em torno de zero, esta é a mesma convolução de $W(z)$ consigo mesma. Finalmente, podemos escrever que

$$\hat{\theta}_{BR} = \int_{-\infty}^{\infty} f_n^*(y) dF_n(y) \quad (4.3)$$

onde $f_n^*(y) = (nh_n)^{-1} \sum_{i=1}^n W^*([Y_i - y]/h_n)$ e $F_n(y)$ é a função de distribuição empírica usual. Esta última expressão (4.3) é identificada como a estimativa do tipo Schuster.

Como nem toda densidade $W(\cdot)$ pode ser escrita como a convolução com ela mesma, não podemos reverter o argumento anterior. Além disso a relação parece ser válida só para o caso de escores de Wilcoxon.

4.2 Normalidade Assintótica no Caso de uma Amostra

As provas da normalidade assintótica do núcleo-estimador de θ requerem que $h_n^2 n \rightarrow \infty$ (Schweder, 1975, Ahmad, 1976). Apresentaremos uma prova usando o argumento de projeção, que permite relaxar esta exigência para $h_n n \rightarrow \infty$. Como descobriremos através dos cálculos na Seção 4.4 isto pode ser conveniente para estimativas em amostras pequenas.

Como descrevemos na Seção 4.1, a estimativa de θ é a forma

$$\hat{\theta} = \int_{-\infty}^{\infty} f_n(y) dF_n(y)$$

a qual pode ser expressada como

$$\begin{aligned} \hat{\theta} &= (n^2 h_n)^{-1} \sum_{i=1}^n \sum_{j=1}^n W([Y_i - Y_j]/h_n) \\ &= W(0)(nh_n)^{-1} + (n^2 h_n)^{-1} \sum_{i \neq j} W([Y_i - Y_j]/h_n). \end{aligned}$$

É conveniente considerar uma pequena modificação na estimativa, irrelevante para seu comportamento assintótico. Os termos $i = j$ não são aleatórios; portanto os eliminamos

enquanto ajustamos o divisor convenientemente.

Definimos

$$\hat{\theta} = (n(n-1)h_n)^{-1} \sum_{i \neq j} W([Y_i - Y_j]/h_n) \quad (4.4)$$

Suponha que $Y_1, Y_2, \dots, Y_n$, é uma amostra aleatória de uma função de distribuição contínua $F(\cdot)$. Como $\hat{\theta}$ é claramente invariante quanto a uma variação de locação, podemos tomar $F(0) = 1/2$ sem perda de generalidade.

O seguinte teorema prova a normalidade assintótica de $\hat{\theta}$.

Teorema 4.2.1. Se a suposição F1 é válida, $W(\cdot)$ uma densidade quadrado integrável, simétrica em torno de zero com segundo momento finito, $h_n \rightarrow 0$ e $nh_n \rightarrow \infty$.

Então

$$V = n^{1/2}(\hat{\theta} - \theta_n) \xrightarrow{D} Z \sim N(0, \sigma^2), \quad (4.5)$$

onde

$$\theta_n = E(\hat{\theta}) \quad (4.6)$$

e

$$\sigma^2 = 4 \left[\int_{-\infty}^{\infty} f^3(y) dy - \left(\int_{-\infty}^{\infty} f^2(y) dy \right)^2 \right]. \quad (4.7)$$

Prova:

Usaremos um argumento de projeção (Hajek, 1968) para estabelecer este resultado. Assim consideramos $V_p = \sum_{k=1}^n E[V | Y_k]$, como a projeção de V sobre o espaço da soma das variáveis aleatórias i.i.d.

Primeiramente mostraremos que

$$V_p \xrightarrow{D} Z \sim N(0, \sigma^2); \quad (4.8)$$

logo mostraremos que

$$E[V - V_p]^2 = \text{Var}(V) - \text{Var}(V_p) \longrightarrow 0 \quad (4.9)$$

quando $n \rightarrow \infty$, que implica $V - V_p \xrightarrow{p} 0$ e estabelece (4.5). A igualdade em (4.9) segue da definição de V_p (Hajek, 1968).

No argumento fazemos uso da função de distribuição $G(\cdot)$ de $Y_i - Y_j$, $i \neq j$. Esta distribuição é discutida em detalhe no Lema 1.4.a que estabelece propriedades de $G(\cdot)$, assim como as propriedades de suas três primeiras derivadas, $g(\cdot)$, $g'(\cdot)$, e $g''(\cdot)$ sob a suposição $F1$ em $F(\cdot)$.

Començaremos a prova pelo cálculo de θ_n (4.6). Referindo-nos à definição (4.4) de $\hat{\theta}$, temos :

$$\begin{aligned} \theta_n &= E \left[(n(n-1)h_n)^{-1} \sum_{i \neq j} W([Y_i - Y_j]/h_n) \right] \\ &= h_n^{-1} \int_{-\infty}^{\infty} W(z/h_n) g(z) dz \\ &= \int_{-\infty}^{\infty} W(u) g(h_n u) du. \end{aligned} \quad (4.10)$$

A seguir calculamos os termos que contribuem em V_p :

$$E[V \mid Y_k = y_k] = n^{1/2} E \left[(n(n-1)h_n)^{-1} \sum_{i \neq j} W([Y_i - Y_j]/h_n) - \theta_n \mid Y_k = y_k \right].$$

Dos $n(n-1)$ termos da soma, $2(n-1)$ envolvem Y_k e como $W(\cdot)$ é simétrica, a esperança de tais termos é a mesma se $i = k$ ou $j = k$. O resto dos termos que não envolvem Y_k são $n(n-1) - 2(n-1) = (n-1)(n-2)$. As esperanças destes termos são as mesmas, pois a distribuição de $Y_i - Y_j$, $i \neq j$, é independente de i e j . Assim temos

$$E[V \mid Y_k = y_k] = n^{1/2} (n(n-1)h_n)^{-1} 2(n-1) E[W([Y - y_k]/h_n)] \quad (4.11)$$

$$+ n^{1/2} (n(n-1)h_n)^{-1} (n-2)(n-1) E[W([Y_i - Y_j]/h_n)] - n^{1/2} \theta_n$$

$$\begin{aligned}
&= 2(n^{1/2}h_n)^{-1} \int_{-\infty}^{\infty} W([y - y_k]/h_n) f(y) dy - 2\theta_n/n^{1/2} \\
&= 2 \left[\int_{-\infty}^{\infty} W(u) f(y_k + h_n u) du - \theta_n \right] / n^{1/2}.
\end{aligned}$$

Portanto a projeção de V é

$$V_p = \sum_{k=1}^n 2 \left[\int_{-\infty}^{\infty} W(u) f(Y_k + h_n u) du - \theta_n \right] / n^{1/2}. \quad (4.12)$$

Agora determinaremos a distribuição limite de V_p , usando o Teorema Central do Limite de Lindeberg (Randles e Wolfe 1979 p. 422).

Seja T_{kn} o k -ésimo termo de V_p (4.12) de maneira que $V_p = \sum_{i=1}^n T_{kn}$. Podemos provar que $E[T_{kn}] = 0$, e

$$Var(T_{kn}) = 4n^{-1} \int_{-\infty}^{\infty} \left[\int_{-\infty}^{\infty} W(u) f(y + h_n u) du \right]^2 f(y) dy - 4n^{-1} \theta_n^2. \quad (4.13)$$

Seja

$$\sigma_n^2 = \sum_{k=1}^n Var(T_{kn}) = n Var(T_{kn}), \quad (4.14)$$

para aplicar o Teorema Central do Limite de Lindenberg, devemos mostrar para todo $\epsilon > 0$;

$$Q = \sigma_n^{-2} \sum_{k=1}^n \int_{-\infty}^{\infty} I\{|t| > \epsilon \sigma_n\} t^2 dF_{kn}(t) \longrightarrow 0 \quad (4.15)$$

quando $n \rightarrow \infty$, onde $F_{kn}(\cdot)$ é a função de distribuição acumulada de T_{kn} . Com esta finalidade mostraremos que existe um $B < \infty$, tal que com probabilidade um $|T_{kn}| \leq n^{1/2}B$. Agora,

$$\begin{aligned}
|T_{kn}| &= 2n^{-1/2} \left| \int_{-\infty}^{\infty} W(u) f(Y_k + h_n u) du - \theta_n \right| \\
&\leq 2n^{-1/2} \int_{-\infty}^{\infty} W(u) f(Y_k + h_n u) du + 2n^{1/2} \theta_n \\
&= M_1 + M_2
\end{aligned} \quad (4.16)$$

usando o fato de que $f(\cdot)$ é limitada (Corolário 1.4.1), temos

$$\begin{aligned} M_1 &\leq 2n^{-1/2} \left[\sup_y f(y) \right] \int_{-\infty}^{\infty} W(u) du \\ &= 2[\sup_y f(y)]/n^{1/2} < \infty. \end{aligned}$$

Pelo Lema 1.4.a $g(\cdot)$ é limitada, assim

$$\begin{aligned} M_2 &= 2n^{-1/2} \int_{-\infty}^{\infty} W(u)g(h_n u) du \\ &\leq 2[\sup_u g(u)]/n^{1/2} < \infty \end{aligned} \tag{4.17}$$

Portanto $T_{kn} \leq n^{-1/2}B$, onde B é dado por

$$B = 2[\sup_y f(y)] + 2[\sup_u g(u)] \tag{4.18}$$

usando este fato em (4.15), temos

$$Q \leq \sigma_n^{-2} \sum_{k=1}^n B^2 n^{-1} \int_{-\infty}^{\infty} I\{n^{1/2}B > \epsilon \sigma_n\} dF_{kn}(t). \tag{4.19}$$

Agora, se σ_n^2 (4.14) é limitada longe de zero, o lado direito de (4.19) converge para zero quando $n \rightarrow \infty$. De fato, mostraremos que $\sigma_n^2 \rightarrow \sigma^2$, onde σ^2 é dado por (4.7).

Primeiramente consideramos

$$\theta_n = \int_{-\infty}^{\infty} W(u)g(h_n u) du.$$

O integrando é limitado por $[\sup_u g(u)]W(u)$ que é integrável. Além disso,

$$W(u)g(h_n u) \rightarrow g(0)W(u)$$

quando $n \rightarrow \infty$. Pelo Teorema de Convergência Dominada temos

$$\theta_n \rightarrow g(0) \int_{-\infty}^{\infty} W(u) du = g(0).$$

Mas,

$$g(0) = \int_{-\infty}^{\infty} f^2(y) dy.$$

A seguir consideramos que

$$\begin{aligned} M_n &= \int_{-\infty}^{\infty} \left[\int_{-\infty}^{\infty} W(u) f(y + h_n u) du \right]^2 f(y) dy \\ &= \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} W(u) W(v) f(y + h_n u) f(y + h_n v) f(y) du dv dy. \end{aligned} \quad (4.20)$$

O integrando é limitado por $[\sup_z f(z)]^2 W(u) W(v) f(y)$, que é integrável. Além disso, o integrando converge para $W(u) W(v) f^3(y)$ quando $n \rightarrow \infty$. Assim, novamente, podemos aplicar o Teorema de Convergência Dominada para obter

$$M_n \rightarrow \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} W(u) W(v) f^3(y) du dv dy = \int_{-\infty}^{\infty} f^3(y) dy,$$

quando $n \rightarrow \infty$. Assim temos,

$$\begin{aligned} \sigma_n^2 &= 4 \int_{-\infty}^{\infty} \left[\int_{-\infty}^{\infty} W(u) f(y + h_n u) du \right]^2 f(y) dy - 4\theta_n^2 \\ &\rightarrow 4 \int_{-\infty}^{\infty} f^3(y) dy - 4 \left[\int_{-\infty}^{\infty} f^2(y) dy \right]^2 = \sigma^2. \end{aligned} \quad (4.21)$$

Agora voltemos para (4.19):

$$Q \leq \sigma_n^{-2} B^2 I \left\{ n^{-1/2} B > \epsilon \sigma_n \right\} \rightarrow 0 \quad (4.22)$$

quando $n \rightarrow \infty$, dado que σ_n^2 é limitada longe de zero. Portanto, podemos aplicar o Teorema Central do Limite de Lindeberg

$$\frac{\sum_{k=1}^n T_{kn}}{\sigma_n} \xrightarrow{D} Z \sim N(0,1). \quad (4.23)$$

De (4.21), $\sigma_n/\sigma \rightarrow 1$, quando $n \rightarrow \infty$. Portanto, podemos aplicar o Teorema de Slutsky (Randles e Wolfe 1979 p.429) em (4.23) e obter:

$$\left(\frac{\sigma_n}{\sigma}\right) \frac{\sum_{k=1}^n T_{kn}}{\sigma_n} = \frac{\sum_{k=1}^n T_{kn}}{\sigma} \xrightarrow{D} Z \sim N(0,1), \quad (4.24)$$

ou

$$V_p = \sum_{k=1}^n T_{kn} \xrightarrow{D} Z \sim N(0, \sigma^2)$$

Assim temos (4.8) como era desejado. E agora falta mostrar que $Var(V) - Var(V_p) \rightarrow 0$ (4.9). Como já foi mostrado que $Var(V_p) \rightarrow \sigma^2$, é suficiente mostrar que $Var(V) \rightarrow \sigma^2$.

Tendo que

$$Var(V) = nE(\hat{\theta}^2) - n\theta_n^2 \quad (4.25)$$

$$= n[n(n-1)h_n]^{-2} \sum_{i \neq j} \sum_{k \neq t} E[W([Y_i - Y_j]/h_n) \times W([Y_k - Y_t]/h_n)] - n\theta_n^2$$

Como no Lema 2.1.a, os termos desta somatória quádrupla pode ser particionada em três grupos.

<u>Grupo</u>	<u>Descrição</u>	<u>Nº de termos</u>
1	Dois pares iguais de sub-índices	$2n(n-1)$
2	Um par de sub-índices iguais	$4n(n-1)(n-2)$
3	Todos os sub-índices diferentes	$\frac{n(n-1)(n-2)(n-3)}{6}$
Total		$n^2(n-1)^2$

Cada grupo será considerando separadamente.

A soma sobre os termos no Grupo 1 é dado por,

$$2n(n-1)n[n(n-1)h_n]^{-2}E[W^2(z/h_n)] \quad (4.26)$$

$$\begin{aligned}
&= 2[(n-1)h_n^2]^{-1} \int_{-\infty}^{\infty} W^2([z/h_n])g(z)dz \\
&= 2[(n-1)h_n]^{-1} \int_{-\infty}^{\infty} W^2(u)g(h_n u)du = [nh_n]^{-1}O(1) \rightarrow 0
\end{aligned}$$

quando $n \rightarrow \infty$, onde $O(1)$ é limitada, pois $g(\cdot)$ é limitada e $W(\cdot)$ é assumida quadrado integrável. Assumimos também $nh_n \rightarrow \infty$, com a finalidade de que $(nh_n)^{-1} \rightarrow 0$.

Para a soma sobre os termos do Grupo 2 temos:

$$\begin{aligned}
&4n(n-1)(n-2)n[n(n-1)h_n]^{-2}E[W([Y_2 - Y_1]/h_n)W([Y_3 - Y_1]/h_n)] \quad (4.27) \\
&= 4(n-2)[(n-1)h_n^2]^{-1} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} W([Y_2 - Y_1]/h_n)W([Y_3 - Y_1]/h_n) \\
&\quad f(y_2)f(y_3)f(y_1)dy_3dy_2dy_1 \\
&= 4(n-2)[(n-1)h_n^2]^{-1} \int_{-\infty}^{\infty} \left[\int_{-\infty}^{\infty} W([Y_2 - Y_1]/h_n)f(y_2)dy_2 \right]^2 f(y_1)dy_1 \\
&= 4(n-2)(n-1)^{-1} \int_{-\infty}^{\infty} \left[\int_{-\infty}^{\infty} W(u)f(y_1 + h_n u)du \right]^2 f(y_1)dy_1 \rightarrow 4 \int_{-\infty}^{\infty} f^3(y)dy,
\end{aligned}$$

quando $n \rightarrow \infty$, isto foi mostrada como parte da convergência, no argumento $\sigma_n^2 \rightarrow \sigma^2$ (4.21).

Finalmente, a soma dos termos no Grupo 3 é:

$$\begin{aligned}
&n(n-1)(n-2)(n-3)n[n(n-1)h_n]^{-2}E^2[W(z/h_n)] \quad (4.28) \\
&= (n-2)(n-3)(n-1)^{-1}\theta_n^2
\end{aligned}$$

por (4.10).

Combinando (4.26), (4.27) e (4.28) com (4.25) podemos escrever

$$\begin{aligned}
Var(V) &= o(1) + 4 \int_{-\infty}^{\infty} f^3(y)dy + (n-2)(n-3)(n-1)^{-1}\theta_n^2 - n\theta_n^2 \\
&= 4 \int_{-\infty}^{\infty} f^3(y)dy + (4n-6)(n-1)^{-1}\theta_n^2 + o(1) \rightarrow \sigma^2
\end{aligned}$$

quando $n \rightarrow \infty$, onde $o(1) = [nh_n]^{-1}O(1)$ converge para zero quando $n \rightarrow \infty$. Assim

$$\theta_n \longrightarrow \theta = \int_{-\infty}^{\infty} f^2(y)dy \quad \text{quando } n \rightarrow \infty.$$

Desta maneira estabelecemos (4.9), com isso a prova do Teorema 4.2.1 está completada. Este resultado é compatível ao Teorema 2.2 de Ahmad (1976). Ao invés da suposição F1, ele assumiu que $\int_{-\infty}^{\infty} f^4(y)dy < \infty$. Em seu artigo ele faz outras suposições que permitem substituir θ_n com θ .

Teorema 4.2.2. Adicionalmente às suposições do Teorema 4.2.1, suponhamos que $W(\cdot)$ tem segundo momento finito e que $n^{1/2}h_n^2 \rightarrow 0$. Então

$$\sqrt{n}(\hat{\theta} - \theta) \xrightarrow{D} Z \sim N(0, \sigma^2)$$

Prova: Fazemos uma expansão de Taylor dentro da integral definida em θ_n (4.6) e (4.10):

$$\begin{aligned} \theta_n &= \int_{-\infty}^{\infty} W(u)g(h_n u)du = \int_{-\infty}^{\infty} W(u) \left[g(0) + h_n u g'(0) + h_n^2 u^2 g''(\epsilon)/2 \right] du \\ &= g(0) \int_{-\infty}^{\infty} W(u)du + h_n g'(0) \int_{-\infty}^{\infty} u W(u)du + h_n^2 g''(\epsilon)/2 \int_{-\infty}^{\infty} u^2 W(u)du \end{aligned} \quad (4.29)$$

onde $|\epsilon| \leq |h_n u|$, pelo Lema 1.4.a $g''(\cdot)$ é limitada; como $W(\cdot)$ é simétrica e seu primeiro momento é zero e por suposição $\int_{-\infty}^{\infty} u^2 W(u)du$ é finita, temos

$$\theta_n = g(0) + h_n^2 O(1) = \theta + h_n^2 O(1) \quad (4.30)$$

onde $O(1)$ é limitada. Portanto,

$$\begin{aligned} \sqrt{n}(\hat{\theta} - \theta) &= \sqrt{n}(\hat{\theta} - \theta_n) + \sqrt{n}(\theta_n - \theta) \\ &= \sqrt{n}(\hat{\theta} - \theta_n) + o(1), \end{aligned} \quad (4.31)$$

como assumimos $n^{1/2}h_n^2 \rightarrow 0$. Uma aplicação do Teorema de Slutsky (Randles e Wolfe 1979, p.424) completa a prova.

Enquanto à distribuição, F1 é suficiente para provar este resultado. Ahmad (1976), para provar este mesmo resultado, supôs que $f''(\cdot)$ existe e é limitada.

4.3 Resíduos e Largura da Janela Aleatória

Nesta seção mostraremos que a estimativa da janela (4.4) de $\theta = \int_{-\infty}^{\infty} f^2(y)dy$ é ainda consistente quando é calculada sobre os resíduos após ajuste de um modelo linear. Também mostraremos que, sob condições apropriadas, a largura da janela h_n pode ser uma variável aleatória. Do ponto de vista prático, este último fato é de grande importância. Por exemplo, podemos incorporar alguma estimativa de escala para largura da janela, de modo que, a estimativa de θ seja equivariante em escala. A prova, mais uma vez, se baseia na estratégia da Seção 1.3. Um artifício adicional é usado como na Seção 3.2. Primeiramente provaremos um teorema que permita o uso da largura da janela aleatória em $W(\cdot)$. Então uma aplicação do Teorema de Slutsky, como no Capítulo 3, nós permite usar uma largura da janela aleatória no divisor e na frente da somatória de (4.4) satisfatoriamente.

Define-se

$$\hat{\theta}(\Delta; v) = [n(n-1)h_n]^{-1} \sum_{i \neq j} W([e_i - e_j - (x_i - x_j)' \Delta] / k_n)$$

onde Δ é um vetor $p \times 1$, x_i é a i -ésima da matriz de desenho X , e $e_i = Y_i - x_i' \beta$, e k_n está relacionado com v por:

$$k_n = h^{-r-q} v + h_n. \quad (4.32)$$

Nota-se que

$$\hat{\theta}(0; 0) = [n(n-1)h_n]^{-1} \sum_{i \neq j} W([e_i - e_j] / h_n)$$

é a estimativa baseada em e_i 's variáveis aleatórias i.i.d., com largura da janela h_n não aleatória.

Mas,

$$\hat{\theta}(\hat{\beta} - \beta_0; n^{q+r}(\hat{h}_n - h_n)) = [n(n-1)h_n]^{-1} \sum_{i \neq j} W([e_i - e_j - (x_i - x_j)'(\hat{\beta} - \beta_0)]/\hat{h}_n)$$

é baseado sobre os resíduos, $\hat{e}_i = Y_i - x_i'\hat{\beta}$, e largura da janela aleatória, $\hat{h}_n$, dentro de $W(\cdot)$. Na Seção 4.4 determinamos algum $h_n = n^{-r}h$, que depende de um parâmetro de escala e a forma da distribuição do erro, que tenha um bom comportamento em pequenas amostras. Então na prática, deveríamos estimar h_n usando algum $\hat{h}_n = n^{-r}\hat{h}$ que dependa das observações.

Teorema 4.3.1. Se F1, X1 e X5 são válidas, $n^{1/2}(\hat{\beta} - \beta_0)$ é limitado em probabilidade, e $W(\cdot)$ é uma densidade quadrado integrável, simétrica, com segundo momento finito. Seja $h_n = n^{-r}h$, onde $0 < r \leq 1/2$ e h é fixo. Suponha $\hat{h}_n = n^{-r}\hat{h}$, e $n^q(\hat{h} - h)$ é limitada em probabilidade para algum $q > 0$. Então,

$$\hat{\theta}(\hat{\beta} - \beta_0; n^q(\hat{h} - h)) \xrightarrow{p} \theta \quad (4.33)$$

Prova: Necessitamos considerar a notação dada no Capítulo 3, seja

$$\begin{aligned} W_1(z) &= W(z), & \text{se } z \leq 0, \\ &= W(0), & \text{se } z > 0; \end{aligned} \quad (4.34)$$

e

$$\begin{aligned} W_2(z) &= 0, & \text{se } z \leq 0, \\ &= W(z) - W(0), & \text{se } z > 0. \end{aligned}$$

Então $W(z) = W_1(z) + W_2(z)$; $W_1(z)$ e $W_2(z)$ são monótonas em z . Também,

como no Capítulo 2, seja

$$D_{ij} = \begin{bmatrix} (x_i - x_j)^+ \\ (x_i - x_j)^- \end{bmatrix} \quad (4.35)$$

onde D_{ij} é um vetor $2p \times 1$; e $(x_i - x_j)^+$ é a parte positiva de $(x_i - x_j)$, $(x_i - x_j)^-$ a parte negativa da mesma, além disso $(x_i - x_j) = (x_i - x_j)^+ + (x_i - x_j)^-$.

Então definimos

$$\theta^*(\delta; v) = [n(n-1)h_n]^{-1} \sum_{i \neq j} \sum W([e_i - e_j - D'_{ij}\delta]/k_n),$$

onde δ é um vetor $2p \times 1$, e k_n é definido por (4.33). Observe que se $\delta' = (\Delta'\Delta')$ temos:

$$\theta^*(\delta; v) = \hat{\theta}(\Delta; v). \quad (4.36)$$

Finalmente particionamos $\theta^*(\delta; v)$ usando (4.34):

$$\begin{aligned} \theta^*(\delta; v) &= [n(n-1)h_n]^{-1} \sum_{i \neq j} \sum \left[W_1 \left([e_i - e_j - D'_{ij}\delta] / k_n \right) + W(0)D'_{ij}\delta \theta \right] \\ &\quad + [n(n-1)h_n]^{-1} \sum_{i \neq j} \sum \left[W_2 \left([e_i - e_j - D'_{ij}\delta] / k_n \right) - W(0)D'_{ij}\delta \theta \right] \\ &= \theta_1^*(\delta; v) + \theta_2^*(\delta; v), \end{aligned} \quad (4.37)$$

A quantidade $W(0)D'_{ij}\delta \theta$ é adicionada e subtraída na expressão acima, e tem uma função importante na prova de

$$\theta_m(\delta/n^{1/2}; v) - \theta_m(0; 0) \xrightarrow{p} 0, \quad (4.38)$$

para $m = 1, 2$. Essa prova é feito usando o lema seguinte.

Lema 4.3.2. Para $m = 1, 2$

$$E \left[\theta_m(\delta/n^{1/2}; v) - \theta_m(0; 0) \right]^2 \rightarrow 0$$

Prova: Consideremos o caso de $m = 1$

$$E \left[\theta_1(\delta/n^{1/2}; v) - \theta_1(0; 0) \right]^2 \quad (4.39)$$

$$= [n(n-1)h_n]^{-2} \sum_{i \neq j} \sum_{s \neq t} E \left[\left(W(z_{ij}/k_n) I\{z_{ij} \leq 0\} - W(z_{ij}^*/h_n) I\{z_{ij}^* \leq 0\} \right. \right. \\ \left. \left. + W(0) \left(I\{z_{ij} > 0\} - I\{z_{ij}^* > 0\} + \theta D'_{ij} \delta/n^{1/2} \right) \right) \left(W(z_{st}/k_n) I\{z_{st} \leq 0\} \right. \right. \\ \left. \left. - W(z_{st}^*/h_n) I\{z_{st}^* \leq 0\} + W(0) \left(I\{z_{st} > 0\} - I\{z_{st}^* > 0\} + \theta D'_{st} \delta/n^{1/2} \right) \right) \right],$$

onde $z_{ij} = e_i - e_j - D'_{ij} \delta/n^{1/2}$ e $z_{ij}^* = e_i - e_j$.

Como na prova do Lema 2.1.a; particionamos os termos desta somatória quádrupla em três grupos

<u>Grupo</u>	<u>Descrição</u>	<u>Nº de termos</u>
1	Dois pares iguais de sub-índices	$2n(n-1)$
2	Um par de sub-índices iguais	$4n(n-1)(n-2)$
3	Todos os sub-índices diferentes	$\frac{n(n-1)(n-2)(n-3)}{2}$
Total		$n^2(n-1)^2$

Trataremos cada grupo separadamente.

Considerando os termos no Grupo 1. Os termos neste grupo são uniformemente limitados, portanto a soma é

$$2n(n-1)[n(n-1)h_n]^{-2} O(1) = (nh_n)^{-2} O(1) \rightarrow 0$$

quando $n \rightarrow \infty$, onde $O(1)$ é limitado.

Particionamos a soma de termos no Grupo 2 em três partes

$$[n(n-1)h_n]^{-2} \sum_{G_2} E \left[W(0) \theta D'_{ij} \delta/n^{1/2} \left(W(z_{st}/k_n) I\{z_{st} \leq 0\} \right. \right. \\ \left. \left. - W(z_{st}^*/h_n) I\{z_{st}^* \leq 0\} + W(0) \left(I\{z_{st} > 0\} - I\{z_{st}^* > 0\} + \theta D'_{ij} \delta/n^{1/2} \right) \right) \right] \quad (4.40)$$

$$\begin{aligned}
& + W(0)\theta D'_{st}\delta/n^{1/2} \left(W(z_{ij}/k_n)I\{z_{ij} \leq 0\} - W(z_{ij}^*/h_n)I\{z_{ij}^* \leq 0\} \right. \\
& + W(0) \left(I\{z_{ij} > 0\} - I\{z_{ij}^* > 0\} + \theta D'_{ij}\delta/n^{1/2} \right) - W^2(0)\theta^2(D'_{ij}\delta)(D'_{st}\delta)/n \Big] \\
& + [n(n-1)h_n]^{-2} \sum_{G_2} E \left[\left(W(z_{ij}/k_n)I\{z_{ij} \leq 0\} - W(z_{ij}^*/k_n)I\{z_{ij}^* \leq 0\} \right) \right. \\
& \times \left. \left(W(z_{st}/k_n)I\{z_{st} \leq 0\} - W(z_{st}^*/k_n)I\{z_{st}^* \leq 0\} \right) \right] \\
& + W^2(0)[n(n-1)h_n]^{-2} \sum_{G_2} E \left[\left(I\{z_{ij} > 0\} - I\{z_{ij}^* > 0\} \right) \left(I\{z_{st} > 0\} - I\{z_{st}^* > 0\} \right) \right] \\
& = Q_1 + Q_2 + Q_3.
\end{aligned}$$

Consideremos Q_1 :

$$\begin{aligned}
Q_1 &= W(0)\theta[n(n-1)h_n]^{-1}n^{-1/2} \sum_{G_2} \left[D'_{ij}\delta \left(\int_{-\infty}^{D'_{st}\delta/n^{1/2}} W([r - D'_{st}\delta/n^{1/2}]/k_n)g(r)dr \right. \right. \\
& - \int_{-\infty}^0 W(r/h_n)g(r)dr + W(0)(1 - G(D'_{st}\delta/n^{1/2}) - 1 + G(0) + \theta D'_{st}\delta/n^{1/2}) \Big) \\
& + D'_{st}\delta \left(\int_{-\infty}^{D'_{ij}\delta/n^{1/2}} W([r - D'_{ij}\delta/n^{1/2}]/k_n)g(r)dr - \int_{-\infty}^0 W(r/h_n)g(r)dr \right. \\
& \left. + W(0)(1 - G(D'_{ij}\delta/n^{1/2}) - 1 + G(0) + \theta D'_{ij}\delta/n^{1/2}) - W(0)\theta(D'_{ij}\delta)(D'_{st}\delta)/n \right]^{1/2} \\
& = Q_{11} + Q_{12} + Q_{13}.
\end{aligned}$$

Somente trabalharemos com a primeira e terceira porções desta soma, pois a segunda soma é semelhante a primeira. Em Q_{11} fazemos uma mudança de variável, $u = [r - D'_{st}\delta/n^{1/2}]/k_n$; na primeira integral e na segunda $u = r/h_n$, e relembremos que $G'(0) = \theta$. Assim temos

$$\begin{aligned}
Q_{11} &= W(0)\theta[n(n-1)h_n]^{-2}n^{-1/2} \sum_{G_2} \left[D'_{ij}\delta \left(k_n \int_{-\infty}^0 W(u)g([k_n u + D'_{st}\delta/n^{1/2}])du \right. \right. \\
& \left. - h_n \int_{-\infty}^0 W(u)g(h_n u)du - W(0)(D'_{st}\delta/n^{1/2})^2 G''(\xi_{st})/2 \right], \quad (4.41)
\end{aligned}$$

onde $|\xi_{st}| \leq |D'_{st}\delta/n^{1/2}|$. Como $g(\cdot)$ e $G''(\cdot)$ são limitadas (Lema 1.4.a) e de (2.15) $D'_{st}\delta/n^{1/2} \rightarrow 0$, uniformemente em todo s, t , quando $n \rightarrow \infty$. Então,

$$Q_{11} = W(0)\theta[n(n-1)h_n]^{-2} \sum_{G_2} (D'_{ij}\delta/n)(k_n O(1) + h_n O(1) + o(1)) \rightarrow 0 \quad (4.42)$$

quando $n \rightarrow \infty$, pois $(nh_n^2)^{-1}$ é limitada, e existem termos da ordem $O(n^3)$ na somatória. Da mesma maneira podemos obter $Q_{12} \rightarrow 0$, quando $n \rightarrow \infty$. Agora consideremos Q_{13} :

$$Q_{13} = W^2(0)\theta^2[n(n-1)h_n^2]^{-1} \sum_{G_2} (D'_{ij}\delta)(D'_{st}\delta)/n \rightarrow 0,$$

quando $n \rightarrow \infty$, pois $(nh_n^2)^{-1}$ é limitado e $D'_{ij}\delta/n^{1/2}$ converge para zero uniformemente para todo i, j . Assim temos que $Q_1 \rightarrow 0$.

A seguir trabalhamos com Q_2 , fazendo $j = t$, por exemplo,

$$Q_2(j=t) = [n(n-1)h_n]^{-2} \sum_{\substack{G_2 \\ j=t}} \int_{-\infty}^{\infty} \int_{-\infty}^{e_j - D'_{ij}\delta/n^{1/2}} \int_{-\infty}^{e_j - D'_{sj}\delta/n^{1/2}} W([e_i - e_j - D'_{ij}\delta/n^{1/2}]/k_n) \quad (4.43)$$

$$\times W([e_s - e_j - D'_{sj}\delta/n^{1/2}]/k_n) f(e_s) f(e_i) f(e_j) de_s de_i de_j$$

$$- [n(n-1)h_n]^{-2} \sum_{\substack{G_2 \\ j=t}} \int_{-\infty}^{\infty} \int_{-\infty}^{e_j + D'_{ij}\delta/n^{1/2}} \int_{-\infty}^{e_j} W([e_i - e_j - D'_{ij}\delta/n^{1/2}]/k_n)$$

$$\times W([e_s - e_j]/h_n) f(e_s) f(e_i) f(e_j) de_s de_i de_j$$

$$- [n(n-1)h_n]^{-2} \sum_{\substack{G_2 \\ j=t}} \int_{-\infty}^{\infty} \int_{-\infty}^{e_j} \int_{-\infty}^{e_j + D'_{sj}\delta/n^{1/2}} W([e_i - e_j]/h_n) W([e_s - e_j - D'_{sj}\delta/n^{1/2}]/k_n)$$

$$\times f(e_i) f(e_s) f(e_j) de_i de_s de_j$$

$$+ [n(n-1)h_n]^{-2} \sum_{\substack{G_2 \\ j=t}} \int_{-\infty}^{\infty} \int_{-\infty}^{e_j} \int_{-\infty}^{e_j} W([e_i - e_j]/h_n) W([e_s - e_j]/h_n) f(e_s) f(e_i) f(e_j) de_s de_i de_j$$

$$= Q_{21} + Q_{22} + Q_{23} + Q_{24}(j=t).$$

Em $Q_{21}(j=t)$, sejam $u_1 = [e_i - e_j - D'_{ij}\delta/n^{1/2}]/k_n$ e $u_2 = [e_s - e_j - D'_{sj}\delta/n^{1/2}]/k_n$.

Então temos

$$Q_{21}(j = t) = [n(n-1)h_n]^{-2} \sum_{\substack{G_2 \\ j=t}} k_n^2 \int_{-\infty}^{\infty} \int_{-\infty}^0 \int_{-\infty}^0 W(u_1)W(u_2) \quad (4.44)$$

$$\times f(e_j + k_n u_1 + D'_{ij} \delta / n^{1/2}) f(e_j + k_n u_2 + D'_{sj} \delta / n^{1/2}) f(e_j) f(u_1) f(u_2) du_1 du_2 de_j \longrightarrow 0$$

quando $n \rightarrow \infty$, já que a integral é limitada por

$$[\sup_y f(y)]^2 \int_{-\infty}^{\infty} \int_{-\infty}^0 \int_{-\infty}^0 W(u_1)W(u_2) f(e_j) du_1 du_2 de_j = [\sup_y f(y)]^3 / 4,$$

$(nh_n^2)^{-1}$ é limitado, e $k_n^2 \rightarrow 0$; substituições similares podem ser feitas em Q_{22} , Q_{23} , e Q_{24} , para $j = t$ e para cada uma das outras possibilidades, $i = s$, $i = t$, e $j = s$. Assim $Q_2 \rightarrow 0$, quando $n \rightarrow \infty$.

Finalmente, consideremos Q_3 , $j = t$, por exemplo:

$$Q_3(j = t) = [n(n-1)h_n]^{-2} \sum_{\substack{G_2 \\ j=t}} \left[\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} I\{e_s > e_j + D'_{ij} \delta / n^{1/2}\} \quad (4.45)$$

$$\begin{aligned} & \times I\{e_i > e_j + D'_{ij} \delta / n^{1/2}\} f(e_s) f(e_i) f(e_j) de_s de_i de_j \\ & - \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} I\{e_s > e_j\} I\{e_i > e_j + D'_{ij} \delta / n^{1/2}\} f(e_s) f(e_i) f(e_j) de_s de_i de_j \\ & - \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} I\{e_s > e_j + D'_{ij} \delta / n^{1/2}\} I\{e_i > e_j\} f(e_s) f(e_i) f(e_j) de_s de_i de_j \\ & + \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} I\{e_s > e_j\} I\{e_i > e_j\} f(e_s) f(e_i) f(e_j) de_s de_i de_j \end{aligned}$$

Em cada uma das três primeiras integrais, o integrando é limitado por $f(e_s) f(e_i) f(e_j)$, que é integrável. Portanto, podemos aplicar Teorema de Convergência Dominada em cada uma das integrais, as quais convergem para

$$\int_{-\infty}^{\infty} \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} I\{e_s > e_j\} I\{e_i > e_j\} f(e_s) f(e_i) f(e_j) de_s de_i de_j,$$

e como $D'_{ij}\delta/n^{1/2} \rightarrow 0$, uniformemente em todo i, j . Portanto

$$Q_3(j=t) = [n(n-1)h_n]^{-2} \sum_{\substack{G_2 \\ j=t}} o(1) \rightarrow 0, \text{ quando } n \rightarrow \infty,$$

como $(nh_n^2)^{-1}$ é limitada, e existem termos da ordem $O(n^3)$ na somatória. Argumentos similares são válidos para os outros casos: $i=s$, $i=t$, e $j=t$. Assim $Q_3 \rightarrow 0$, desta forma os termos do grupo 2 convergem para zero.

Agora consideremos a soma sobre o Grupo 3:

$$[n(n-1)h_n]^{-2} \sum_{G_3} E[(\cdot)(\cdot)] = [n(n-1)h_n]^{-2} \sum_{G_3} E[\cdot]E[\cdot], \quad (4.46)$$

pois os multiplicandos são independentes porque os sub-índices são todos distintos. Considerando uma das esperanças, mencionada em (4.39)

$$\begin{aligned} & E \left[W(z_{ij}/k_n) I\{z_{ij} \leq 0\} - W(z_{ij}^*/h_n) I\{z_{ij}^* \leq 0\} \right. \\ & \left. + W(0)(I\{z_{ij} > 0\} - I\{z_{ij}^* > 0\} + \theta D'_{ij}\delta/n^{1/2}) \right] \\ &= \int_{-\infty}^{D'_{ij}\delta/n^{1/2}} W([r - D'_{ij}\delta/n^{1/2}]/k_n) g(r) dr - \int_{-\infty}^0 W(r/h_n) g(r) dr \\ & \quad + W(0)[1 - G(D'_{ij}\delta/n^{1/2}) - 1 + G(0) + \theta D'_{ij}\delta/n^{1/2}] \\ &= k_n \int_{-\infty}^0 W(u) g(k_n u + D'_{ij}\delta/n^{1/2}) du - h_n \int_{-\infty}^0 W(u) g(h_n u) du \\ & \quad - W(0)(D'_{ij}\delta/n^{1/2})^2 G''(\xi_{ij})/2, \end{aligned} \quad (4.47)$$

onde $|\xi_{ij}| \leq |D'_{ij}\delta/n^{1/2}|$, lembrando que $G'(0) = \theta$. Fazemos uma expansão de Taylor de $g(\cdot)$ em torno de zero, em cada integral, reexpressamos a esperança como

$$k_n \int_{-\infty}^0 W(u) \left((g(0) + (k_n u + D'_{ij}\delta/n^{1/2})g'(0) + (k_n u + D'_{ij}\delta/n^{1/2})^2 g''(\zeta_{ij})/2) \right) du \quad (4.48)$$

$$-h_n \int_{-\infty}^0 W(u)(g(0) + h_n u g'(0) + (h_n u)^2 g''(\varsigma)/2) du - W(0)(D'_{ij}\delta/n^{1/2})^2 G''(\xi_{ij})/2,$$

onde $|\varsigma_{ij}| \leq |k_n u + D'_{ij}\delta/n^{1/2}|$ e $|\varsigma| \leq |k_n u|$. Como $g'(0) = 0$ e $g''(\cdot)$ é limitada, além disso $\int_{-\infty}^0 W(u) du = 1/2$, $\int_{-\infty}^0 |u| W(u) du < \infty$ e $\int_{-\infty}^0 u^2 W(u) du < \infty$.

Portanto podemos escrever

$$[n(n-1)h_n]^{-1} \sum_{i \neq j} \sum |E[.]| \quad (4.49)$$

$$\leq [n(n-1)h_n]^{-1} \sum_{i \neq j} \sum \left[|k_n - h_n| g(0)/2 + |k_n|^3 O(1) + k_n^2 |D'_{ij}\delta/n^{1/2}| O(1) \right.$$

$$\left. + |k_n(D'_{ij}\delta/n^{1/2})| O(1) + h_n^3 O(1) + (D'_{ij}\delta/n^{1/2})^2 O(1) \right] \rightarrow 0$$

quando $n \rightarrow \infty$. Pois $|k_n - h_n|/h_n \rightarrow 0$, $|k_n|^3/h_n \rightarrow 0$, $k_n^2/h_n \rightarrow 0$, $|D'_{ij}\delta/n^{1/2}| \rightarrow 0$ uniformemente para todo i, j , $n^{-2} \sum_{i \neq j} \sum (D'_{ij}\delta/n^{1/2})$ é limitada (ver (2.23)) e $(nh_n)^{-1} \rightarrow 0$. Mas

$$\left| [n(n-1)h_n]^{-2} \sum_{G3} E[.]E[.] \right| \leq \left[[n(n-1)h_n]^{-1} \sum_{i \neq j} \sum |E(.)| \right]^2 \rightarrow 0, \text{ quando } n \rightarrow \infty.$$

Portanto a soma dos termos no Grupo 3 converge para zero.

Um argumento similar é válido para $\theta_2^*(\delta/n^{1/2}; v) - \theta_2^*(0; 0)$ e com isto completamos a prova do Lema 4.3.a.

Agora relembando que (4.37)

$$\begin{aligned} \theta_1^*(\delta/n^{1/2}; v) - \theta_1^*(0; 0) &= [n(n-1)h_n]^{-1} \sum_{i \neq j} \sum \left[W_1([e_i - e_j - D'_{ij}\delta/n^{1/2}]/k_n) - W_1([e_i - e_j]/h_n) \right] \\ &\quad + [n(n-1)h_n]^{-1} \sum_{i \neq j} \sum W(0) \theta D'_{ij}\delta/n^{1/2} \end{aligned}$$

A primeira somatória é monótona nas componentes de δ e em v (ver (4.32), para a relação de v com k_n). Por (2.23) $\left\| [n(n-1)]^{-1} \sum_{i \neq j} \sum D'_{ij}\delta \right\|$ é limitada e pela suposição temos que $[n^{1/2}h_n]^{-1}$ é limitado.

Portanto, podemos aplicar Lema 1.3.c e

$$\sup \left| \theta_m^*(\delta/n^{1/2}; v) - \theta_m^*(0; 0) \right| \xrightarrow{P} 0 \quad (4.50)$$

onde *supremo* é sobre $\|\delta\| \leq B$ e $|v| \leq B$. Pela suposição $n^{1/2}(\hat{\beta} - \beta_0) \text{ e } n^q(\hat{h} - h)$ são limitados em probabilidade. Portanto, uma aplicação do Lema 1.3.d para $m = 1, 2$ produz

$$\theta_m^* \left(\begin{pmatrix} \hat{\beta} - \beta_0 \\ \hat{\beta} - \beta_0 \end{pmatrix}; n^q(\hat{h} - h) \right) - \theta_m^*(0; 0) \xrightarrow{P} 0.$$

Mas

$$\hat{\theta} \left[(\hat{\beta} - \beta_0); n^q(\hat{h} - h) \right] - \hat{\theta}(0; 0) = \sum_{m=1}^2 \left[\theta_m^* \left(\begin{pmatrix} \hat{\beta} - \beta_0 \\ \hat{\beta} - \beta_0 \end{pmatrix}; n^q(\hat{h} - h) \right) - \theta_m^*(0; 0) \right] \xrightarrow{P} 0$$

com o qual o teorema está provado.

Uma aplicação do Teorema de Slutsky produz o seguinte corolário

Corolário 4.3.1 . Sobre as condições do Teorema 4.3.1

$$[n(n-1)\hat{h}_n]^{-1} \sum_{i \neq j} W([\hat{e}_i - \hat{e}_j]/\hat{h}_n) \xrightarrow{P} \theta, \text{ quando } n \rightarrow \infty.$$

Prova: Sabemos que $h_n/\hat{h}_n \xrightarrow{P} 1$ e do Teorema 4.3.1 temos que

$$[n(n-1)h_n]^{-1} \sum_{i \neq j} W([\hat{e}_i - \hat{e}_j]/\hat{h}_n) \xrightarrow{P} \theta.$$

Portanto o produto converge em probabilidade para θ .

4.4 Cálculos

4.4.1 Escolha da Largura da Janela

Nesta subseção, desenvolvemos algumas escolhas da largura da janela h_n e discutimos a possibilidade de correção do vício.

Os argumentos na Seção 4.3 mostraram que o núcleo-estimador calculado sobre os resíduos usando largura da janela aleatória é consistente para θ . Assim, qualquer escolha de $W(\cdot)$ e h_n , que satisfaçam as condições dessa seção, resultam uma estimativa aceitável, dada uma amostra suficientemente grande. Para uma aplicação, não poderíamos decidir quão grande poderia ser uma amostra para que fosse suficientemente grande. Em qualquer caso, raramente temos um tamanho de amostra ilimitada. Com a finalidade de suavizar este problema, escolheremos h_n de maneira que as estimativas calculadas em pequenas amostras tenham um "bom comportamento", quando estas são medidas pelo vício ou pela variância.

Embora estejamos interessados na estimativa calculada a partir dos resíduos, os cálculos são feitos numa amostra aleatória. Isto infelizmente se torna necessário, já que os cálculos de momentos no primeiro caso é muito complicado.

Observe que escolhemos tratar os termos da diagonal ($i = j$) em

$$\hat{\theta} = n^{-2} h_n^{-1} \sum_{i=1}^n \sum_{j=1}^n W([Y_i - Y_j]/h_n) \quad (4.51)$$

separadamente, e talvez possamos ignorar totalmente, já que eles contribuiriam meramente como uma parte não aleatória na estimativa e é irrelevante para seu comportamento assintótico. Estes termos podem ser melhorados através de uma correção do vício, pois eles afetam a variância da estimativa. Além disso dividimos a soma restante por $n(n-1)h_n$ ao invés de $n^2 h_n$, com a finalidade de evitar uma confusão nos cálculos. Assim consideremos,

$$\hat{\theta} = [n(n-1)h_n]^{-1} \sum_{i \neq j} W([Y_i - Y_j]/h_n), \quad (4.52)$$

deixando aberta a possibilidade de adicionar correção ao vício. Nos cálculos que seguem, assumimos que $Y_1, Y_2, \dots, Y_n$ são variáveis aleatórias i.i.d., com função de distribuição acumulada $F(\cdot)$ satisfazendo F1. Faremos uso das propriedades de $G(\cdot)$ função de distribuição acumulada de $Y_i - Y_j$ ($i \neq j$), discutida na Seção 1.4. Considerando a esperança de $\hat{\theta}$ (4.52):

$$\begin{aligned} E[\hat{\theta}] &= h_n^{-1} \int_{-\infty}^{\infty} W(z/h_n) g(z) dz \\ &= \int_{-\infty}^{\infty} W(u) g(h_n u) du \\ &= \int_{-\infty}^{\infty} W(u) (g(0) + h_n u g'(0) + (h_n u)^2 g''(\epsilon)/2) du \end{aligned} \quad (4.53)$$

onde $|\epsilon| \leq |h_n u|$. Pelo lema do final desta subseção podemos escrever que

$$\begin{aligned} E[\hat{\theta}] &= \int_{-\infty}^{\infty} W(u) (g(0) + h_n u g'(0) + (h_n u)^2 g''(0)/2) du + h_n^2 o(1) \\ &= g(0) \int_{-\infty}^{\infty} W(u) du + h_n g'(0) \int_{-\infty}^{\infty} u W(u) du + (h_n)^2 g''(0)/2 \int_{-\infty}^{\infty} u^2 W(u) du + h_n^2 o(1) \end{aligned} \quad (4.54)$$

como $W(\cdot)$ é simétrica em torno de zero e $g(0) = \theta$, então o vício de $\hat{\theta}$ é:

$$\text{vício}(\hat{\theta}) = 2^{-1} h_n^2 g''(0) \int_{-\infty}^{\infty} u^2 W(u) du + o(h_n^2). \quad (4.55)$$

Como

$$g''(0) = - \int_{-\infty}^{\infty} [f'(y)]^2 dy, \quad (4.56)$$

$\hat{\theta}$ converge sobestimadamente a θ . E, se estamos interessados no vício devemos fazer que h_n convirja para zero para zero tão rápido quanto possível. Relembrando que o argumento da Seção 4.3 requer que $(h_n^2 n)^{-1}$ seja limitado não podemos deixar que h_n convirja para zero tão rápido do que $n^{-1/2}$.

Suponha que $h_n = kn^{-1/2}$, onde k ainda é desconhecido. Uma forma de obter k

poderia ser modificando a estimativa $\hat{\theta}$ pela adição de uma correção ao vício; por exemplo,

$$\hat{\theta} = (nk)^{-1} + [n(n-1)h_n]^{-1} \sum_{i \neq j} W([Y_i - Y_j]/h_n). \quad (4.57)$$

Ignorando os termos de ordem $o(h_n^2)$ no vício temos

$$vicio(\hat{\theta}) = (nk)^{-1} + k^2(2n)^{-1}g''(0) \int_{-\infty}^{\infty} u^2 W(u) du, \quad (4.58)$$

fazendo esta expressão igual a zero e resolvendo para k , obtemos

$$k = \left[-1/2g''(0) \int_{-\infty}^{\infty} u^2 W(u) du \right]^{-1/3}. \quad (4.59)$$

Assim, se tomamos $h_n = kn^{-1/2}$ o vício é de ordem $o(n^{-1})$, a qual parece razoavelmente apropriado.

Schweder(1975) sugere a largura da janela para o caso em que θ é estimado por

$$\hat{\theta} = (nh_n)^{-1} + [n(n-1)h_n]^{-1} \sum_{i \neq j} W([Y_i - Y_j]/h_n). \quad (4.60)$$

Então o vício ignorando os termos de ordem $o(h_n^2)$, é

$$vicio(\hat{\theta}) = (nh_n)^{-1} + h_n^2 g''(0) \int_{-\infty}^{\infty} \frac{u^2}{2} W(u) du. \quad (4.61)$$

Então fazendo esta expressão igual zero, resulta uma raiz cúbica real

$$h_n = \left[-1/2g''(0) \int_{-\infty}^{\infty} u^2 W(u) du \right]^{-1/3} n. \quad (4.62)$$

Observe que $h_n \rightarrow 0$ à razão de $n^{-1/3}$; de maneira que o resto do vício é de ordem $o(n^{-2/3})$.

Em (4.59) não conhecemos $g''(0)$ (4.56) pois $\int_{-\infty}^{\infty} [f'(y)]^2 dy$ é desconhecido. Portanto, suponhamos que

$$f(y) = \sigma^{-1} f_1\left(\frac{y}{\sigma}\right),$$

para algum parâmetro de escala σ . Então em (4.56) temos que

$$g''(0) = \sigma^{-3} \int_{-\infty}^{\infty} [f'_1(y)]^2 dy.$$

Assim,

$$k = \sigma \left[\int_{-\infty}^{\infty} u^2 W(u) du \int_{-\infty}^{\infty} [f'_1(y)]^2 dy / 2 \right]^{-1/3}. \quad (4.63)$$

Temos separadamente duas componentes em (4.63): um parâmetro de escala σ e a forma padronizada de $f(\cdot)$ denotada por $f_1(\cdot)$. Então podemos estimar estas duas componentes das observações; uma simples aproximação de σ pode ser um parâmetro de escala robusto tal como o desvio absoluto da média ou da amplitude interquartilica, enquanto que assumimos uma distribuição para o cálculo de $\int_{-\infty}^{\infty} [f'_1(y)]^2 dy$.

Agora provaremos um lema que dá uma medida para o cálculo do vício.

Lema 4.4.a. Se $|\xi| \leq |h_n u|$ e $g''(\cdot)$ contínua em zero, então

$$\int_{-\infty}^{\infty} u^2 W(u) g''(\xi) du \rightarrow g''(0) \int_{-\infty}^{\infty} u^2 W(u) du$$

Prova: Pelo Lema 1.4.a $g''(\cdot)$ é limitado. Assim $|u^2 W(u) g''(\xi)| \leq \sup_z |g''(z)| u^2 W(u)$, é integrável. Portanto o Teorema de Convergência Dominada pode-se aplicar para produzir o resultado desejado.

Note que se $f'(\cdot)$ é limitada e contínua, ou se $f''(\cdot)$ existe e é absolutamente integrável, então $g''(\cdot)$ é contínua em zero.

4.4.2 Estimativa de θ Baseado no Comprimento do Intervalo de Confiança Aproximado numa Amostra.

Nesta subseção seguimos o ponto de vista de Schweder (1975) na estimativa de θ baseado no comprimento do intervalo de confiança de Hodges-Lehmann para o centro de simetria, numa amostra, que pode ser visto como um estimativa da janela (núcleo-

estimador). Esta investigação esclarece a necessidade da simetria de $f(\cdot)$ em torno de zero, quando esta aproximação é usada. Suponha que temos uma amostra aleatória $Y_1 + \alpha_0, Y_2 + \alpha_0, \dots, Y_n + \alpha_0$ de tamanho n , com função de distribuição acumulada $F(\cdot)$ que satisfaça F1.

Seja

$$\begin{aligned} T(\alpha) &= [n(n+1)]^{-1} \sum_{i \leq j} I\{Y_i + Y_j > 2\alpha\} \\ &= [2n(n+1)]^{-1} \sum_{i=1}^n \sum_{j=1}^n I\{Y_i + Y_j > 2\alpha\} + [2n(n+1)]^{-1} \sum_{i=1}^n I\{Y_i > \alpha\} \\ &= [2n(n+1)]^{-1} \sum_{i=1}^n \sum_{j=1}^n I\{Y_i + Y_j > 2\alpha\} + O(n^{-1}), \end{aligned} \quad (4.64)$$

onde $O(n^{-1})$ é limitada. Escolhemos $\hat{\alpha}_L$ e $\hat{\alpha}_U$ tal que

$$(12n)^{1/2} (T(\hat{\alpha}_L) - 1/4) = Z_{\alpha/2} \quad (4.65)$$

e

$$(12n)^{1/2} (T(\hat{\alpha}_U) - 1/4) = -Z_{\alpha/2},$$

onde $Z_{\alpha/2}$ é o valor crítico superior da distribuição normal padronizada. Então,

$$\hat{\theta}_1 = 2Z_{\alpha/2} / [(12n)^{1/2} (\hat{\alpha}_U - \hat{\alpha}_L)] \quad (4.66)$$

é a estimativa usual de θ . Sejam $h_n = 2(\hat{\alpha}_U - \hat{\alpha}_L)$ e $\hat{\alpha} = (\hat{\alpha}_L + \hat{\alpha}_U)/2$, e assim $h_n/2 = 2(\hat{\alpha}_U - \hat{\alpha}) = 2(\hat{\alpha} - \hat{\alpha}_L)$, onde $\hat{\alpha}$ é estimativa de α_0 . Agora subtraindo a segunda linha da primeira em (4.65) obtemos

$$2Z_{\alpha/2} = (12n)^{1/2} (T(\hat{\alpha}_L) - T(\hat{\alpha}_U)). \quad (4.67)$$

Substituindo isto em (4.66), temos

$$\begin{aligned}
\hat{\theta}_1 &\doteq (T(\hat{\alpha}_L) - T(\hat{\alpha}_U))/(\hat{\alpha}_U - \hat{\alpha}_L) \\
&= 2h_n^{-1}[2n(n+1)]^{-1} \sum_{i=1}^n \sum_{j=1}^n [I\{Y_i + Y_j > 2\hat{\alpha}_L\} - I\{Y_i + Y_j > 2\hat{\alpha}_U\}] \\
&= [n(n+1)h_n]^{-1} \sum_{i=1}^n \sum_{j=1}^n I\{-h_n/2 < Y_i + Y_j - 2\hat{\alpha} \leq h_n/2\} \\
&\doteq [n(n+1)h_n]^{-1} \sum_{i=1}^n \sum_{j=1}^n I\{|Y_i + Y_j - 2\hat{\alpha}| > h_n/2\}
\end{aligned} \tag{4.68}$$

onde a aproximação aparece por não considerar a porção $O(n^{-1})$ de $T(\hat{\alpha}_L)$ e $T(\hat{\alpha}_U)$. Agora observamos que $\hat{\theta}_1$ é essencialmente

$$\theta_1^* = \int f_n(-y) dF_n(y) \tag{4.69}$$

onde $f_n(\cdot)$ é um núcleo-estimador da densidade $f(\cdot)$ baseado sobre os resíduos $Y_i - \hat{\alpha}$, e $F_n(\cdot)$ é uma função de distribuição empírica destes mesmos resíduos. Mas observe que $\hat{\theta}_1$ será consistente para $\int f(-y) dF(y)$, que é igual a $\theta = \int f(y) dF(y)$ somente se $f(\cdot)$ é simétrica em torno de zero. Portanto, a menos que a distribuição do erro subjacente seja simétrico, a estimativa de θ baseada sobre a amplitude do intervalo de confiança em uma amostra, será calculada erradamente.

4.4.3 Estimativa de θ Baseada no Comprimento do Intervalo de Confiança Aproximado para duas Amostras

Draper (1988) estudou a estimativa de θ baseado no comprimento do intervalo de confiança de Hodges-Lehmann para o parâmetro de locação baseado em duas amostras. Esta aproximação requer replicações na matriz de desenho, mas evita o requerimento da simetria da distribuição dos erros. Essencialmente, para cada par de grupo de réplicas, uma estimativa de θ é calculada baseada no comprimento do intervalo de confiança de

Hodges-Lehmann para a mudança entre os dois grupos. Então estas estimativas de θ são combinadas com a finalidade de obter uma estimativa total. Nesta subseção consideramos o modelo linear mais simples para que este modelo seja aplicado, num problema de duas amostras. Fazemos alguns cálculos que ajudam a ver a relação entre esta aproximação e o núcleo-estimador.

Suponha que temos uma amostra aleatória $Y_1, Y_2, \dots, Y_n$ com $Y_1, Y_2, \dots, Y_m$ e $Y_{m+1} - \Delta, Y_{m+2} - \Delta, \dots, Y_n - \Delta$, todos i.i.d. com função de distribuição $F(\cdot)$ e sem perda de generalidade consideramos $F(0) = 1/2$.

Seja $V(\Delta)$ a estatística do teste Mann-Whitey-Wilcoxon para Δ :

$$V(\Delta) = [m(n-1)]^{-1} \sum_{i=1}^m \sum_{j=m+1}^n I\{Y_j - Y_i > \Delta\}, \quad (4.70)$$

e escolhemos $\hat{\Delta}_L$ e $\hat{\Delta}_U$ tais que:

$$V(\hat{\Delta}_L) - 1/2 = Z_{\alpha/2} [(n+1)/\{12m(n-m)\}]$$

e

$$V(\hat{\Delta}_U) - 1/2 = -Z_{\alpha/2} [(n+1)/\{12m(n-m)\}].$$

Então uma estimativa consistente de θ é dada por

$$\hat{\theta}_2 = Z_{\alpha/2} / [3m(n-m)/\{12m(n-m)\}^{1/2}(\hat{\Delta}_U - \hat{\Delta}_L)], \quad (4.71)$$

(Lehmann, 1963).

Sejam $\hat{h}_n = \hat{\Delta}_U - \hat{\Delta}_L$ e $\hat{\Delta} = (\hat{\Delta}_L + \hat{\Delta}_U) / 2$. Observamos que $\hat{h}_n \rightarrow 0$; e $\hat{\Delta}$ é uma estimativa de Δ , tal que $n^{1/2}(\hat{\Delta} - \Delta)$ é limitada em probabilidade. Então podemos reescrever $\hat{\theta}_2$ como segue:

$$\hat{\theta}_2 = [V(\hat{\Delta}_L) - V(\hat{\Delta}_U)] / [2\{(n+1)/[12m(n-m)]\}^{1/2}\{3m(n-m)/(n+1)\}^{1/2}\hat{h}_n] \quad (4.72)$$

$$\begin{aligned}
&= [V(\hat{\Delta}_L) - V(\hat{\Delta}_U)] / \hat{h}_n \\
&= [m(n-m)\hat{h}_n]^{-1} \sum_{i=1}^m \sum_{j=m+1}^n [I\{Y_j - Y_i > \hat{\Delta}_L\} - I\{Y_j - Y_i > \hat{\Delta}_U\}] \\
&= [m(n-m)\hat{h}_n]^{-1} \sum_{i=1}^m \sum_{j=m+1}^n [I\{\hat{\Delta}_L < Y_j - Y_i < \hat{\Delta}_U\}] \\
&\doteq [m(n-m)\hat{h}_n]^{-1} \sum_{i=1}^m \sum_{j=m+1}^n [I\{|Y_j - Y_i - \hat{\Delta}| < \hat{h}_n/2\}].
\end{aligned}$$

Notamos que $\hat{\theta}_2$ é bastante similar ao núcleo-estimador calculado nos resíduos após o ajustamento do modelo linear.

De fato, podemos escrever

$$\hat{\theta}_2 \doteq \int_{-\infty}^{\infty} f_m(y) dF_{n-m}(y), \quad (4.73)$$

onde $f_m(y)$ é uma estimativa com janela retangular na primeira amostra, usando a largura da janela aleatória $\hat{h}_n$, e $F_{n-m}(y)$ é uma função de distribuição baseada sobre resíduos da segunda amostra. Naturalmente, podemos também escrever

$$\hat{\theta}_2 \doteq \int_{-\infty}^{\infty} f_{n-m}(y) dF_m(y), \quad (4.74)$$

onde a estimação de densidades é calculada sobre os resíduos da segunda amostra e $F_m(y)$ é a função de distribuição da primeira amostra.

Além disso, a estimativa de janela retangular no intervalo $(-1/2, 1/2)$ de θ , baseada nos resíduos, pode ser particionada como segue

$$\begin{aligned}
\hat{\theta}_w &= (n^2 h_n)^{-1} \sum_{i=1}^n \sum_{j=1}^n I\{|r_j - r_i| < h_n/2\} \\
&= (n^2 h_n)^{-1} \left[2 \sum_{i=1}^m \sum_{j=m+1}^n I\{|Y_j - \hat{\Delta} - Y_i| < h_n/2\} \right.
\end{aligned} \quad (4.75)$$

$$+ \sum_{i=1}^m \sum_{j=1}^m I\{|Y_j - Y_i| < h_n/2\} + \sum_{i=m+1}^n \sum_{j=m+1}^n I\{|Y_j - Y_i| < h_n/2\} \Bigg] ,$$

onde $r_i = Y_i$ se $1 \leq i \leq m$ e $r_i = Y_i - \hat{\Delta}$, para $m+1 \leq i \leq n$. Então temos

$$\hat{\theta}_w = [2m(n-m)n^{-2}][m(n-m)]^{-1} \sum_{i=1}^m \sum_{j=m+1}^n I\{|Y_j - \hat{\Delta} - Y_i| < h_n/2\} \quad (4.76)$$

$$+ [m^2 n^{-2}][m^2 h_n]^{-1} \sum_{i=1}^m \sum_{j=1}^m I\{|Y_j - Y_i| < h_n/2\}$$

$$+ [(n-m)^2 n^{-2}][(n-m)^2 h_n]^{-1} \sum_{i=m+1}^n \sum_{j=m+1}^n I\{|Y_j - Y_i| < h_n/2\}.$$

Assim $\hat{\theta}_w$ é uma soma poderada das três estimativas: a primeira, é essencialmente $\hat{\theta}_2$; a segunda e a terceira são estimativas de janela em amostras separadas (na primeira e na segunda amostra respectivamente). Podemos afirmar que $\hat{\theta}_w$ está usando mais informação dos dados do que $\hat{\theta}_2$.

Capítulo 5

INTERCEPTO

5.1 Normalidade Assintótica de $\begin{pmatrix} \hat{\alpha} \\ \hat{\beta} \end{pmatrix}$

Nesta seção definimos uma estimativa do intercepto, sem a suposição de simetria da distribuição dos erros e mostraremos que a distribuição de $\begin{pmatrix} \hat{\alpha} \\ \hat{\beta} \end{pmatrix}$ é assintoticamente normal. Sejam α_0 e β_0 os valores verdadeiros de α e β respectivamente. Define-se $\alpha_0^* = \alpha_0 + \bar{x}'\beta_0$, onde $\bar{x}'$ é o vetor linha $1 \times p$ de médias de X . Particionando a matriz de desenho centralizada X_c (1.12) em suas partes, positivas e negativas, $X_c = X_c^+ + X_c^-$, e seja $D = [X_c^+, X_c^-]$ com a i -ésima linha d_i' .

Agora definimos

$$F_n(a; \Delta) = n^{-1} \sum_{i=1}^n I\{e_i - (x_i - \bar{x})'\Delta - a \leq 0\}. \quad (5.1)$$

Observe que esta é uma função não decrescente em a . Portanto definimos a estimativa $\hat{\alpha}^*$ ou $\hat{\alpha}$ por:

$$F_n((\hat{\alpha}^* - \alpha_0^*); (\hat{\beta} - \beta_0)) \doteq 1/2 \quad (5.2)$$

e

$$\hat{\alpha}_0 = \hat{\alpha}_0^* - \bar{x}'\hat{\beta}.$$

Então $\hat{\alpha}_0^*$ é a mediana dos $Y_i - (x_i - \bar{x})'\hat{\beta}$, e $\hat{\alpha}$ é a mediana dos $Y_i - x_i'\hat{\beta}$. Suponha-se que com probabilidade um não há empates nos resíduos; assim para valores pares de n , a igualdade em (5.2) é exata, enquanto para valores ímpares de n pode ser igual a $\frac{(n+1)}{2n}$.

O seguinte teorema mostra a normalidade assintótica de $\begin{pmatrix} \hat{\alpha} \\ \hat{\beta} \end{pmatrix}$ sem a suposição da simetria da distribuição dos erros.

Teorema 5.1.1. Se F1, X1, e X5 são válidas, e $\hat{\beta}$ é a estimativa baseada em postos definida no Capítulo 1. Então para escores de Wilcoxon,

$$n^{1/2} \begin{pmatrix} \hat{\alpha} - \alpha_0 \\ \hat{\beta} - \beta_0 \end{pmatrix} \xrightarrow{D} Z \sim NMV(0, V) \quad (5.3)$$

com matriz de variância-covariância

$$V = \begin{bmatrix} (4f^2(0))^{-1} + \tau^2 \mu'_x \Sigma^{-1} \mu_x & -\tau^2 \mu'_x \Sigma^{-1} \\ -\tau^2 \mu'_x \Sigma^{-1} & \tau^2 \Sigma^{-1} \end{bmatrix} \quad (5.4)$$

onde $\Sigma = \lim_{n \rightarrow \infty} n^{-1} X'_c X_c$ e $\mu_x = \lim_{n \rightarrow \infty} \bar{x}$.

Prova:

A prova do teorema será feita através de uma serie de lemas.

Lema 5.1.a. Seja $\{a_n\}$ uma seqüência limitada . Então:

$$n^{1/2} [F_n(a_n; (\hat{\beta} - \beta_0)) - F_n(a_n; 0)] \xrightarrow{p} 0 \quad (5.5)$$

(note que a_n pode ou não depender de n).

Prova: Primeiro definimos

$$F_n^*(a; \delta) = n^{-1} \sum_{i=1}^n I\{e_i - d'_i \delta - a \leq 0\}; \quad (5.6)$$

notamos que esta é monótona nas componentes de δ . O seguinte argumento mostra que

$$n^{1/2} \left[F_n^*(a_n; \delta/n^{1/2}) - F_n^*(a_n; 0) - n^{-1} \sum_{i=1}^n f(a_n) d'_i \delta / n^{1/2} \right] \xrightarrow{p} 0, \quad (5.7)$$

com $|a_n| \leq A$. Calculamos o segundo momento desta quantidade:

$$\begin{aligned}
& nE \left[F_n^*(a_n; \delta/n^{1/2}) - F_n^*(a_n; 0) - n^{-1} \sum_{i=1}^n f(a_n) d'_i \delta/n^{1/2} \right]^2 \quad (5.8) \\
& n^{-1} E \left[\sum_{i=1}^n \left[I\{e_i - d'_i \delta/n^{1/2} - a_n \leq 0\} - I\{e_i - a_n \leq 0\} - f(a_n) d'_i \delta/n^{1/2} \right] \right]^2 \\
& = n^{-1} \sum_{i=1}^n E \left[I\{e_i - d'_i \delta/n^{1/2} - a_n \leq 0\} - I\{e_i - a_n \leq 0\} - f(a_n) d'_i \delta/n^{1/2} \right]^2 \\
& + n^{-1} \sum_{i \neq j} \sum E \left[I\{e_i - d'_i \delta/n^{1/2} - a_n \leq 0\} - I\{e_i - a_n \leq 0\} - f(a_n) d'_i \delta/n^{1/2} \right] \\
& \times E \left[I\{e_j - d'_j \delta/n^{1/2} - a_n \leq 0\} - I\{e_j - a_n \leq 0\} - f(a_n) d'_j \delta/n^{1/2} \right] \\
& = Q_1 + Q_2.
\end{aligned}$$

Primeiro, consideremos Q_1 :

$$\begin{aligned}
Q_1 &= n^{-1} \sum_{i=1}^n E \left[I\{e_i - d'_i \delta/n^{1/2} - a_n \leq 0\} - I\{e_i - a_n \leq 0\} - f(a_n) d'_i \delta/n^{1/2} \right]^2 \quad (5.9) \\
&= n^{-1} \sum_{i=1}^n E \left[\left(I\{e_i - d'_i \delta/n^{1/2} - a_n \leq 0\} - I\{e_i - a_n \leq 0\} \right)^2 \right. \\
&\quad \left. - 2f(a_n) \left(d'_i \delta/n^{1/2} \right) \left(I\{e_i - d'_i \delta/n^{1/2} - a_n \leq 0\} - I\{e_i - a_n \leq 0\} \right) + n^{-1} f^2(a_n) \left(d'_i \delta/n^{1/2} \right)^2 \right] \\
&= n^{-1} \sum_{i=1}^n \left[\left| F(a_n + d'_i \delta/n^{1/2}) - F(a_n) \right| - 2f(a_n) \left(d'_i \delta/n^{1/2} \right) \left(F(d'_i \delta/n^{1/2} - a_n) - F(a_n) \right) \right. \\
&\quad \left. + n^{-1} f^2(a_n) \left(d'_i \delta/n^{1/2} \right)^2 \right].
\end{aligned}$$

Como

$$|d'_i \delta/n^{1/2}| \leq p \|\delta\| \max_{i,j} |x_{ij} - \bar{x}_{.j}| / n^{1/2} \longrightarrow 0 \quad (5.10)$$

quando $n \rightarrow \infty$, uniformemente em todo i, j (Lema 1.4.a). Além disso, $F(\cdot)$ é uniformemente contínua em $[-2A, 2A]$, e pelo Corolário 1.4.1 $f(\cdot)$ é limitada. Assim as somatórias

em (5.9) convergem para zero uniformemente para todo i , em consequencia $Q_1 \rightarrow 0$. A seguir consideramos Q_2 .

$$Q_2 = n^{-1} \sum_{i \neq j} \sum \left[F(a_n + d'_i \delta / n^{1/2}) - F(a_n) - d'_i \delta / n^{1/2} f(a_n) \right] \quad (5.11)$$

$$\times \left[F(d'_j \delta / n^{1/2} + a_n) - F(a_n) - d'_j \delta / n^{1/2} f(a_n) \right].$$

Como

$$n^{-1/2} \sum_{i=1}^n \left| F(d'_i \delta / n^{1/2} + a_n) - F(a_n) - f(a_n) d'_i \delta / n^{1/2} \right| \quad (5.12)$$

$$= n^{-1/2} \sum_{i=1}^n \left| d'_i \delta / n^{1/2} F(\epsilon_i) - d'_i \delta / n^{1/2} n^{-1/2} f(a_n) \right|$$

$$= n^{-1} \sum_{i=1}^n \left| d'_i \delta / n^{1/2} \right| \times |F(\epsilon_i) - f(a_n)|$$

onde $|\epsilon_i - a_n| \leq |d'_i \delta / n^{1/2}| \rightarrow 0$, uniformemente em todo i . Como $f(\cdot)$ é uniformemente contínua em $[-2A, 2A]$, isto implica que $|F(\epsilon_i) - f(a_n)| \rightarrow 0$, quando $n \rightarrow \infty$, uniformemente em todo i .

Além disso,

$$n^{-1} \sum_{i=1}^n |d'_i \delta| \leq \sum_{j=1}^p \|\delta\| \quad n^{-1} \sum_{i=1}^n |x_{ij} - \bar{x}_{.j}|, \quad (5.13)$$

que é limitada (Lema 1.4.c). Portanto a quantidade em (5.12) converge para zero quando $n \rightarrow \infty$. Mas de (5.11);

$$|Q_2| \leq n^{-1} \sum_{i \neq j} \sum \left| F(d'_i \delta / n^{1/2} + a_n) - F(a_n) - f(a_n) d'_i \delta / n^{1/2} \right| \quad (5.14)$$

$$\times \left| F(d'_j \delta / n^{1/2} + a_n) - F(a_n) - f(a_n) d'_j \delta / n^{1/2} \right|$$

$$\leq \left[n^{1/2} \sum_{i=1}^n \left| F(d'_i \delta / n^{1/2} + a_n) - F(a_n) - f(a_n) d'_i \delta / n^{1/2} \right| \right]^2 \rightarrow 0$$

quando $n \rightarrow \infty$. Então $Q_2 \rightarrow 0$. E com isto temos estabelecemos (5.7). Já que

$$\left\| n^{-1} \sum_{i=1}^n f(a_n) d_i \right\| \leq f(a_n) \max_j n^{-1} \sum_{i=1}^n |x_{ij} - \bar{x}_{.j}| \quad (5.15)$$

é limitado para $j = 1, \dots, p$, portanto o Lema 1.3.c pode ser aplicado para provar que:

$$\sup_{\|\delta\| \leq B} n^{1/2} \left| F_n^*(a_n; \delta/n^{1/2}) - F_n^*(a_n; 0) - f(a_n) n^{-1} \sum_{i=1}^n d_i' \delta/n^{1/2} \right| \xrightarrow{P} 0. \quad (5.16)$$

Como $n^{1/2} \begin{pmatrix} \hat{\beta} - \beta_0 \\ \hat{\beta} - \beta_0 \end{pmatrix}$ é limitada em probabilidade, temos

$$n^{1/2} \left| F_n(a_n; (\hat{\beta} - \beta_0)) - F_n(a_n; 0) - f(a_n) n^{-1} \sum_{i=1}^n (x_i - \bar{x})' (\hat{\beta} - \beta_0) \right| \xrightarrow{P} 0. \quad (5.17)$$

Mas $\sum_{i=1}^n (x_i - \bar{x}) = 0$, com o qual o Lema 5.1.a, esta provado.

Lema 5.1.b

$$n^{1/2} [F_n(a/n^{1/2}; 0) - F_n(0; 0) - f(0)a/n^{1/2}] \xrightarrow{P} 0$$

Prova:

$$\begin{aligned} & nE [F_n(a/n^{1/2}; 0) - F_n(0; 0) - f(0)a/n^{1/2}]^2 \quad (5.18) \\ &= n^{-1} E \left[\sum_{i=1}^n (I\{e_i - a/n^{1/2} \leq 0\} - I\{e_i \leq 0\} - f(0)a/n^{1/2}) \right]^2 \\ &= n^{-1} \sum_{i=1}^n E [I\{e_i - a/n^{1/2} \leq 0\} - I\{e_i \leq 0\} - f(0)a/n^{1/2}]^2 \\ &+ n^{-1} \sum_{i \neq j} E [I\{e_i - a/n^{1/2} \leq 0\} - I\{e_i \leq 0\} - f(0)a/n^{1/2}] \\ &\quad \times E [I\{e_j - a/n^{1/2} \leq 0\} - I\{e_j \leq 0\} - f(0)a/n^{1/2}] \\ &= Q_1 + Q_2. \end{aligned}$$

Considerando Q_1 :

$$\begin{aligned}
 Q_1 &= n^{-1} \sum_{i=1}^n E \left[\left(I\{e_i - a/n^{1/2} \leq 0\} - I\{e_i \leq 0\} \right)^2 \right. \\
 &\quad \left. - 2(I\{e_i - a/n^{1/2} \leq 0\} - I\{e_i \leq 0\})f(0)a/n^{1/2} + f^2(0)a^2/n \right] \\
 &= n^{-1} \sum_{i=1}^n \left[|F(a/n^{1/2}) - F(0)| - 2(F(a/n^{1/2}) - F(0))f(0)a/n^{1/2} + f^2(0)a^2/n \right]
 \end{aligned} \tag{5.19}$$

Como $F(\cdot)$ é contínua, todos os somandos convergem para zero uniformemente, portanto $Q_1 \rightarrow 0$. A seguir consideramos Q_2 :

$$\begin{aligned}
 Q_2 &= n^{-1} \sum_{i \neq j} \sum \left[F(a/n^{1/2}) - F(0) - f(0)a/n^{1/2} \right]^2 \\
 &\leq \left[n^{1/2} \sum_{i=1}^n |F(a/n^{1/2}) - F(0) - f(0)a/n^{1/2}| \right]^2 \\
 &= \left[n^{-1} \sum_{i=1}^n |a_i| \times |f(\epsilon) - f(0)| \right]^2,
 \end{aligned} \tag{5.20}$$

onde $|\epsilon| \leq |a/n^{1/2}| \rightarrow 0$, como $f(\cdot)$ é contínua em zero, os somandos, no lado direito convergem para zero uniformemente; e $Q_2 \rightarrow 0$ também. Assim o Lema 5.1.b está provado.

Lema 5.1.c

$$\sup_{|a| \leq A} n^{1/2} \left| F_n(a/n^{1/2}; (\hat{\beta} - \beta_0)) - F_n(0; 0) - f(0)a/n^{1/2} \right| \xrightarrow{p} 0$$

Prova:

$$\begin{aligned}
 &n^{1/2} \left| F_n(a/n^{1/2}; (\hat{\beta} - \beta_0)) - F_n(0; 0) - f(0)a/n^{1/2} \right| \\
 &\leq n^{1/2} \left| F_n(a/n^{1/2}; (\hat{\beta} - \beta_0)) - F(a/n^{1/2}; 0) \right| \\
 &\quad + n^{1/2} \left| F_n(a/n^{1/2}; 0) - F(0; 0) - f(0)a/n^{1/2} \right| \xrightarrow{p} 0,
 \end{aligned} \tag{5.21}$$

pelos Lemas 5.1.a , Lema 5.1.b; pois $|a_n| = |a/n^{1/2}| \leq A$. Como

$$n^{1/2} [F_n(a/n^{1/2}; 0) - F(0; 0)]$$

é monótona em a , e $f(0)$ é limitada . Então podemos aplicar o Lema 1.3.b para obter o resultado desejado.

Lema 5.1.d. $n^{1/2}(\hat{\alpha}^* - \alpha_0^*)$ é limitada em probabilidade.

Prova: Mostraremos que $n^{1/2}(\hat{\alpha}^* - \alpha_0^*)$ é limitada superiormente em probabilidade.

Para o caso que seja limitada inferiormente a prova é similar.

Seja $\epsilon > 0$ dado, sabemos que

$$n^{1/2}(F_n(0; 0) - 1/2) \xrightarrow{D} N(0, 1/4).$$

Portanto existe um $M > 0$ tal que para um n suficientemente grande

$$P \{n^{1/2}(F_n(0; 0) - 1/2) \leq M\} < \epsilon/2.$$

Definimos um $M^* = (-M + \epsilon)/f(0) > 0$. Se $n^{1/2}(\hat{\alpha}^* - \alpha_0^*) > M^*$ então

$$F_n(M^*/n^{1/2}; (\hat{\beta} - \beta_0)) \leq 1/2,$$

como $F_n(a; \Delta)$ é não decrescente em a . Portanto.

$$P \{n^{1/2}(\hat{\alpha}^* - \alpha_0^*) > M^*\} \tag{5.22}$$

$$\begin{aligned} &\leq P \{n^{1/2} (F_n(M^*/n^{1/2}; (\hat{\beta} - \beta_0)) - 1/2) \leq 0\} \\ &= P \{n^{1/2} (F_n(M^*/n^{1/2}; (\hat{\beta} - \beta_0)) - F_n(0; 0) - M^*f(0)/n^{1/2}) \\ &\quad + n^{1/2} (F_n(0; 0) - 1/2 + M^*f(0)/n^{1/2}) \leq 0\} \end{aligned}$$

$$\leq P \left\{ n^{1/2} \left| F_n(M^*/n^{1/2}; (\hat{\beta} - \beta_0)) - F_n(0; 0) - M^* f(0)/n^{1/2} \right| \geq \epsilon \right\} \\ + P \left\{ n^{1/2} \left(F_n(0; 0) - 1/2 + M^* f(0)/n^{1/2} \right) \leq \epsilon \right\}$$

Pelo Lema 5.1.c, a primeira probabilidade é menor o igual a $\epsilon/2$ para um n suficientemente grande. A segunda é igual a

$$P \left\{ n^{1/2} (F_n(0; 0) - 1/2) \leq M \right\} \leq \epsilon/2$$

para uma escolha adequada de M .

Desta forma o Lema é estabelecido.

Os Lema 5.1.c e o Lema 5.1.d implicam que

$$n^{1/2} \left[F_n((\hat{\alpha}^* - \alpha_0^*); (\hat{\beta} - \beta_0)) - F_n(0; 0) - (\hat{\alpha}^* - \alpha_0^*) f(0) \right] \xrightarrow{P} 0.$$

Pela definição, $F_n((\hat{\alpha}^* - \alpha_0^*); (\hat{\beta} - \beta_0)) = 1/2$, então

$$n^{1/2} [1/2 - F_n(0; 0) - (\hat{\alpha}^* - \alpha_0^*) f(0)] \xrightarrow{P} 0.$$

Lema 5.1.e

$$S = n^{1/2} \begin{bmatrix} (\hat{\alpha}^* - \alpha_0^*) \\ \tau^{-1} n^{-1} X_c' X_c (\hat{\beta} - \beta_0) \end{bmatrix} \xrightarrow{D} Z \sim NMV(0, C), \quad (5.23)$$

onde a matriz de variância - covariância é dada por

$$C = \begin{bmatrix} (2f(0))^{-2} & 0 \\ 0 & \Sigma \end{bmatrix}, \quad (5.24)$$

Prova: É conhecido que

$$(\tau n^{1/2})^{-1} X_c' X_c (\hat{\beta} - \beta_0) - n^{-1/2} X_c' \phi(F(e)) \xrightarrow{p} 0 \quad (5.25)$$

onde $\phi(F(e))$ é o vetor com a i -ésima componente $\phi(F(e_i))$. Assim podemos escrever

$$n^{1/2} \begin{bmatrix} (F_n(0;0) - 1/2) / f(0) \\ n^{-1} X_c' \phi(F(e)) \end{bmatrix} - S \xrightarrow{p} 0 \quad (5.26)$$

O j -ésimo elemento do vetor $p \times 1$

$$E[(F_n(0;0) - 1/2) X_c' \phi(F(e))]$$

é

$$\begin{aligned} & E \left[\sum_{i=1}^n (I\{e_i \leq 0\} - 1/2) \sum_{k=1}^n (x_{kj} - \bar{x}_{.j}) \phi(F(e_k)) \right] \\ &= \sum_{k=1}^n (x_{kj} - \bar{x}_{.j}) \left[\sum_{i=1}^n E(I\{e_i \leq 0\} - 1/2) \phi(F(e_k)) \right] = 0, \end{aligned} \quad (5.27)$$

como

$$\sum_{i=1}^n E(I\{e_i \leq 0\} - 1/2) \phi(F(e_k)) = E(I\{e_k \leq 0\} - 1/2) \phi(F(e_k))$$

não depende de k e $\sum_{k=1}^n (x_{kj} - \bar{x}_{.j}) = 0$. Portanto

$$n^{1/2} (F_n(0;0) - 1/2) / f(0)$$

e os elementos de $n^{-1/2} X_c' \phi(F(e))$ são não correlacionados. Isto, combinado com o fato que

$$n^{-1/2} X_c' \phi(F(e)) \xrightarrow{D} Z \sim NMV(0, \Sigma)$$

produz o resultado. Ver Lehman (1975), Teorema 18, p.390.

Finalmente, estamos prontos para provar o teorema. Definimos

$$T = \begin{bmatrix} 1 & -\tau \bar{x}'(n^{-1} X_c' X_c)^{-1} \\ 0 & \tau(n^{-1} X_c' X_c)^{-1} \end{bmatrix}, \quad (5.28)$$

Então

$$TS = n^{1/2} \begin{bmatrix} \hat{\alpha} - \alpha_0 \\ \hat{\beta} - \beta_0 \end{bmatrix},$$

e TS é assintoticamente normal multivariada com matriz de variância - covariância dado por:

$$\lim_{n \rightarrow \infty} T \begin{bmatrix} (2f(0))^{-2} & 0 \\ 0 & \Sigma \end{bmatrix} T' = \begin{bmatrix} (4f^2(0))^{-1} + \tau^2 \mu_x' \Sigma^{-1} \mu_x & -\tau^2 \mu_x' \Sigma^{-1} \\ -\tau^2 \mu_x' \Sigma^{-1} & \tau^2 \Sigma^{-1} \end{bmatrix} = V$$

com em (5.4).

Note-se que se desejamos uma estimativa dos erros padronizados associados a $\hat{\alpha}$ e $\hat{\beta}$, necessitamos uma estimativa consistente de $f(0)$ e τ . Brevemente discutiremos a estimação de $f(0)$ na Seção 5.2. No Capítulo 4 estudamos uma estimativa de τ a qual não requer simetria de $F(\cdot)$, mas somente para o caso de Escores de Wilcoxon. Se outros escores são usados, podemos utilizar τ^* (1.17), que exige de simetria de $F(\cdot)$ e adicionalmente necessitamos impor a suposição de X3 na matriz de desenho.

5.2 Estimação de $f(0)$

O uso da estimativa de $\hat{\alpha}$ definida na Seção 5.1 requer que estimemos $f(0)$, já que esta quantidade aparece na matriz de variância-covariância de $\begin{pmatrix} \hat{\alpha} \\ \hat{\beta} \end{pmatrix}$. Nesta seção mostraremos simplifcadamente duas possibilidades para a estimação de $f(0)$.

Pode-se usar a estimação de densidades baseada nos resíduos, $\hat{e}_i = Y_i - \hat{\alpha} - x_i' \hat{\beta}$:

$$\hat{f}(0) = (n \hat{h}_n)^{-1} \sum_{i=1}^n W(\hat{e}_i / \hat{h}_n), \quad (5.29)$$

onde $\hat{h}_n$ é a largura da janela baseada nas observações. O resultado do Capítulo 3 mostra que estimativa (5.29) é consistente para $f(0)$. O problema de escolha de h_n , para obter uma estimativa que tenha um bom comportamento em pequenas amostras deve ser tratado. Para este propósito, Rosemblat (1956) fez com que os termos principais da expansão de Taylor do erro quadrático médio fossem iguais a zero e determinou

$$h_n = \left[(9f(0))/(2(f''(0))^2) \right]^{1/5} n^{-1/5}. \quad (5.30)$$

Naturalmente não conhecemos $f(0)$ e $f''(0)$ na prática. Se padronizamos $f(0)$ pela remoção de um parâmetro de escala: $f(y) = \sigma f_1(\sigma^{-1}y)$, então podemos escrever

$$h_n = \sigma \left[(9f_1(0))/(2(f_1''(0))^2) \right]^{1/5} n^{-1/5}. \quad (5.31)$$

Então, podemos supor uma distribuição subjacente e calcular $f_1(0)$, $f_1''(0)$ e também estimar σ das observações usando uma padronização apropriada de alguma estimativa de escala resistente. A estimação de σ será feita considerando que estimativa de $f(0)$ (5.30) seja equivariante.

Uma outra possibilidade de estimar $f(0)$ é aplicando a estimativa estudada por Block e Gastwirth (1968), para os resíduos não centralizados, $Y_i - x_i' \hat{\beta}$. Suponha $T_1 < T_2 < \dots < T_n$ uma amostra aleatória ordenada com função de distribuição acumulada $F(t - \alpha)$; $F(0) = 1/2$, com densidade positiva $f(t - \alpha)$, tendo primeira derivada contínua numa vizinhança de α . Então

$$S_{mn} = n(2m)^{-1} (T_{[n/2]+m} - T_{[n/2]-m+1}), \quad (5.32)$$

onde $[.]$ é a função maior inteiro, é consistente para $1/f(0)$ sob a condição, de $m = o(n)$ e $m \rightarrow \infty$, quando $n \rightarrow \infty$. Pela minimização do erro quadrático médio assintótico, Block

e Gastwirth (1968) encontraram que uma escolha ótima de m é

$$m = cn^{4/5}, \quad (5.33)$$

onde c é uma constante que depende da distribuição subjacente. Estimação de c e a consistência de S_{mn} quando é calculada sobre os resíduos requer outra investigação.

Um caso especial de S_{mn} é a estimativa de $f(0)$ baseada no comprimento do intervalo de confiança da mediana, obtida pela inversa do teste sinal (*sign-test*). Suponha que (T_L, T_U) é o intervalo de $100(1 - \alpha)\%$ de confiança, onde L e U podem ser obtidos das tabelas exatas da distribuição normal ou da aproximação:

$$L = \left[\frac{n}{2} - \frac{n^{1/2} Z_{\alpha/2}}{2} \right] \quad (5.34)$$

e

$$U = \left[\frac{n}{2} + \frac{n^{1/2} Z_{\alpha/2}}{2} \right].$$

Então

$$\hat{f}(0) = \frac{n^{1/2} (T_U - T_L)}{2Z_{\alpha/2}}.$$

A escolha ótima do coeficiente de confiança e seu comportamento quando é calculada sobre os resíduos ainda devem ser pesquisados.

Capítulo 6

Considerações finais

As principais vantagens do uso de testes não-paramétricos devem-se ao fato de serem menos exigentes e mais eficientes que os métodos paramétricos, quando os dados da população não têm distribuição normal. São úteis no caso em que é difícil de estabelecer uma escala de valores quantitativos para os dados e geralmente os testes não-paramétricos, baseados em postos, são robustos aos valores discrepantes, por isso podemos afirmar que os métodos não-paramétricos podem ser uma complementação aos métodos paramétricos.

O procedimento de inferência baseado em postos para β_0 no modelo linear é válido sem a suposição $X3$ sobre a matriz de desenho, se os escores de Wilcoxon são usados. Talvez este resultado pode ser obtido para outras funções geradoras de escores.

A densidade estimada a partir dos resíduos após o ajustamento de um modelo linear é ainda consistente.

A estimativa de janela (ou núcleo-estimador) de $\theta = \int f^2(y)dy$, baseada nos resíduos após o ajuste do modelo linear não exige a suposição de simetria de $F(\cdot)$. A normalidade assintótica de $\hat{\theta}$ no caso de uma amostra aleatória foi estabelecida sob condições diferentes

que Schweder (1975) e Ahmad (1976). A estimativa de θ é consistente quando é calculada sobre os resíduos usando largura da janela aleatória. Alguns cálculos são feitos com a finalidade de obter uma escolha ótima de h_n .

Uma estimativa de α_0 é definida, a qual não requer simetria de $F(\cdot)$. Além disso mostramos a normalidade assintótica de $\left(\hat{\alpha}\right)'$ e mencionamos duas possibilidades de estimação de $f(0)$, a qual aparece na matriz de variância-covariância de $\left(\hat{\alpha}\right)'$; essas estimativas requerem outros estudos.

Referências Bibliográficas

- [1] Adiche, J. N. (1967). Estimates of regression parameters based on rank tests. *Ann. Math. Stat.* 38: 884-893.
- [2] Adiche, J. N. (1978). Rank tests of sub-hypotheses in the general linear regression. *Ann. Stat. Stat.* 6: 1012-1026.
- [3] Ahmad, I. A. (1976). On asymptotic properties an estimates of functional of a probability density. *Scand. Actuarial. J.* 4: 176-181.
- [4] Aubuchon, J.C. and Hettmansperger, T. P. (1989). Rank-based inference for linear models: assymetrics erros. *Statistics & Probability Letters.* 8: 97-107.
- [5] Arnold, S. F (1980). Asymptotic validity of of teste for the ordinary model and the multiple correlation model. *J. Am. Stat.Assoc.* 75: 890-894.
- [6] Bean, S. J. and Tsokos, C. P. (1980). Developoments in nonparametrics density estimation. *International Statiscal Review.* 48: 267-287.
- [7] Bean, S. J. and Tsokos, C. P. (1982). Bandwidth selection procedures for kernel density estimates. *Comm. Statist.- Theory Metth.* 11: 1045-1069.
- [8] Bhattacharyya, P. K. and Roussas, G. G. (1969). Estimation of certain functional of a probability function density. *Skand. Aktuar. Tidskr.* 52: 431-444.
- [9] Block, D. A. and Gastwirth, J. L. (1968). On a simple estimate of the reciprocal of the density function. *Ann. Math. Statist.* 39: 1083-1085.
- [10] Ckeng, K. F. and Serfling, R. J. (1981). On estimation of class of efficacy-related parameters. *Scand. Actuarial J.* 5: 83-92

- [11] Draper, D. (1988). Rank- based robust analysis of linear models. I. Exposition and review. *Statistical Science* 3: 239-271.
- [12] Hájek, J. (1968). Asymptotic normality of simple linear rank statistics under alternatives. *Ann. Math. Statist.* 39: 325-346.
- [13] Hájek, J. (1969). *Nonparametric Statistics*. Holden-Day, San Francisco.
- [14] Hájek, J. and Sidak, Z. (1967) *Theory of rank tests*. Academic Press, New York.
- [15] Hettmansperger, T. P. and McKean, J. W. (1977). A robust alternative based on ranks to least squares in analyzing linear models. *Technometrics*. 19: 275-284.
- [16] Hettmansperger, T. P. and McKean, J. W. (1983). A geometric interpretation of inference based on ranks in the linear model. *J. Am. Statist. Assoc.* 78: 885-886.
- [17] Jaeckel, L. A. (1972). Estimating regression coefficients by minimizing the dispersion of the residuals. *Ann. Math. Statist.* 43: 1449-1458.
- [18] Jureckova, J. (1969). Asymptotic linearity of rank statistic in regression parameter. *Ann. Math. Statist.* 40: 1889-1900.
- [19] Jureckova, J. (1971). Nonparametric estimate of regression coefficients. *Ann. Math. Statist.* 40: 1328-1338.
- [20] Kraft, C. H. and Van Eeden, C. (1972). Linearized rank estimate and signed-rank estimate for the general linear model. *Ann. Math. Statist.* 43: 42-57.
- [21] Lehmann, E. L. (1963). Nonparametric confidence intervals for shift parameter. *Ann. Math. Statist.* 34: 1507-1512.
- [22] Lehmann, E. L. (1975). *Nonparametrics: Statistical methods based on rank*. Holden-Day, San Francisco.
- [23] McKean, J. W. and Hettmansperger, T. P. (1976). Tests of hypotheses based on ranks in the models linear general. *Comm. Statist.-Theor. Meth.* 5: 652-709.

- [24] McKean, J. W. and Hettmansperger, T. P. (1978). A robust analysis of the general linear model on the step R-estimates. *Biometrics*. 65: 571-579.
- [25] McKean, J. W. and Scharader, R. M. (1980). The geometry of robust procedures in linear models. *J. R. Statist. Soc. (B)* 42: 366-371.
- [26] McKean, J. W. and Scharader, R. M. (1984). A comparison of methods for estimating the sample median. *Comm. Statist. -Simul. Comput.* 6: 751-773.
- [27] Parzen, E. (1962). On estimation of a probability density function and mode. *Ann. Math. Statist.* 33: 1065-1076.
- [28] Puri, M. L. Sen, P.K. (1985). *Nonparametric methods in general linear models*. John Wiley & Sons.
- [29] Randles, R. H. and Wolfe, D. A. (1979). *Introduction to the theory of nonparametric statistics*. John Wiley & Son, New York.
- [30] Rosenblatt, M. (1956). Remarks on some nonparametric estimates of density function. *Ann. Math. Statist.* 27: 832-837.
- [31] Schrader, R. M. and McKean, J. W. (1977). Robust analysis of variance. *Comm. Statist.- Theor. Meth.* 9: 879-894.
- [32] Schrader, R. M. and Hettmansperger, T. P. (1980). Robust analysis of variance based upon a likelihood ratio criterion. *Biometrika* 67: 93-101.
- [33] Schuster, E. (1974). On the rate of convergence of an estimate of functional of a probability density. *Scand. Actuarial J.* 1: 103-107.
- [34] Schweder, T. (1975). Window estimation of the asymptotic variance of rank estimators of location. *Scand. J. Statist.* 2: 113-126.
- [35] Sen, P. K. (1966). On a distribution-free method of estimating asymptotic efficiency of the a class of nonparametric tests. *Ann. Math. Statist.* 43: 791-802.

- [36] Sen, P. K. (1968). Estimates of the regression coefficient based on Kendall's tau. *J. Am. Stat. Assoc.* 63: 1379-1389.
- [37] Serfling, R. J. (1980). *Approximation theorems of mathematical statistics*. Wiley, New York.
- [38] Silverman, B. W. (1986). *Density estimation for statistics and data analysis*: Chapman and Hall.
- [39] Shirley, E. A. (1981). A distribution-free method for analysis of covariance based on ranked data. *Applied Statistics* 30: 158- 162.
- [40] Van Eeden, C. (1972). An analogue, for signed rank statistics, of Jureckova's asymptotic linearity theorem for rank statistics. *Ann. Math. Statist.* 43: 791-802.
- [41] Wegman, E. J. (1972). Nonparametric probability density estimation: I. A summary of available method. *Technometrics* 14: 533-546.