

Estimação por mínimos quadrados ponderados e generalizados

Prof. Caio Azevedo

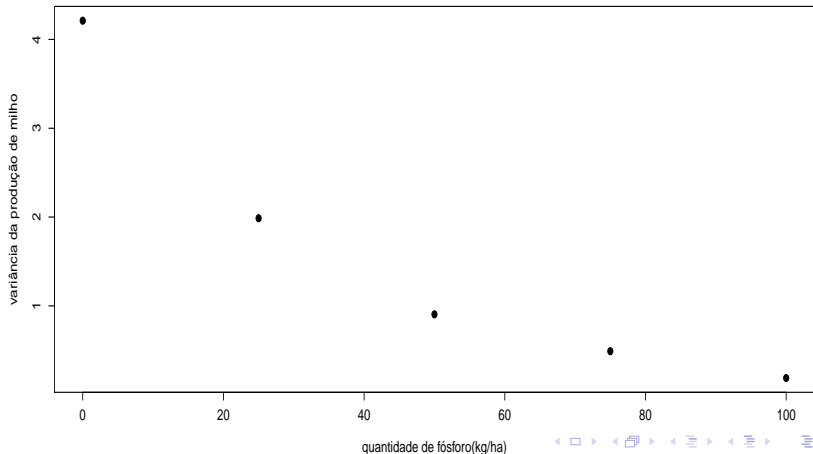
Exemplo 3: efeito do fósforo na produção de milho

Nível de fósforo	Média	DP	CV(%)	Var.	Mínimo	Máximo
0 kg/ha	4,647	2,052	44,156	4,211	2,380	6,770
25 kg/ha	8,255	1,410	17,076	1,987	6,150	9,100
50 kg/ha	8,300	0,951	11,461	0,905	6,920	9,070
75 kg/ha	9,350	0,700	7,482	0,489	8,660	10,240
100 kg/ha	9,640	0,433	4,492	0,188	9,140	10,170

Ao que parece, pela análise descritiva, as variâncias entre os grupos (níveis de fósforo) não são as mesmas. Ou seja, a suposição de homocedasticidade parece estar sendo violada.



Variância da produção de milho em função da quantidade de fósforo



Forma matricial do MNL

$$\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\xi}$$

$$\mathbf{Y} = \begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_n \end{bmatrix}, \mathbf{X} = \begin{bmatrix} X_{11} & \dots & X_{1p} \\ X_{21} & \dots & X_{2p} \\ \vdots & \ddots & \vdots \\ X_{n1} & \dots & X_{np} \end{bmatrix}, \boldsymbol{\beta} = \begin{bmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_p \end{bmatrix}, \boldsymbol{\xi} = \begin{bmatrix} \xi_1 \\ \xi_2 \\ \vdots \\ \xi_n \end{bmatrix}$$

- Suponha agora que: $\boldsymbol{\xi} \sim N_n(\mathbf{0}, \boldsymbol{\Sigma})$, em que:

$$\boldsymbol{\Sigma} = \begin{bmatrix} \sigma_1^2 & 0 & \dots & 0 \\ 0 & \sigma_2^2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \sigma_n^2 \end{bmatrix} \quad (1)$$

Ausência de homocedasticidade e presença de correlação

- Ou seja, embora continuemos assumindo independência entre as observações, a variância já não mais é considerada constante ao longo das observações (heterocedasticidade).
- Suponha, por enquanto, que Σ seja conhecida.
- Defina

em que
$$\mathbf{Z} = \mathbf{W}\boldsymbol{\beta} + \boldsymbol{\delta} \quad (2)$$

$$\boldsymbol{\Sigma} = \boldsymbol{\Sigma}^{1/2}(\boldsymbol{\Sigma}')^{1/2} \text{(decomposição de Cholesky)}$$

$$\mathbf{Z} = \boldsymbol{\Sigma}^{-1/2} \mathbf{Y}$$

$$\mathbf{W} = \boldsymbol{\Sigma}^{-1/2} \mathbf{X}$$

$$\boldsymbol{\delta} = \boldsymbol{\Sigma}^{-1/2} \boldsymbol{\xi}$$

Cont.

- Assim, $\mathcal{E}(\delta) = \Sigma^{-1/2}\mathcal{E}(\xi) = \mathbf{0}$ e $\text{Cov}(\delta) = \Sigma^{-1/2}\Sigma(\Sigma')^{-1/2} = \mathbf{I}$.
- Portanto, a estrutura apresentada em (2) corresponde à um modelo de regressão normal linear homocedástico.
- Logo, o estimador de MQO de β , sob o modelo (2), é dado por:

$$\hat{\beta}_W = (\mathbf{W}'\mathbf{W})^{-1}\mathbf{W}'\mathbf{Z} = (\mathbf{X}'\Sigma^{-1}\mathbf{X})^{-1}\mathbf{X}'\Sigma^{-1}\mathbf{Y}$$

assim

$$\hat{\beta}_W \sim N_p(\beta, (\mathbf{X}'\Sigma^{-1}\mathbf{X})^{-1})$$

Cont.

- Dizemos que $\hat{\beta}_W$ é o estimador de mínimos quadrados ponderados de β , pois ele é o vetor que minimiza a soma de quadrados ponderada (exercício):

$$Q_W = (\mathbf{Y} - \mathbf{X}\beta)' \mathbf{\Sigma}^{-1} (\mathbf{Y} - \mathbf{X}\beta)$$

- Em geral, considera-se que:

$$\mathbf{\Sigma} = \sigma^2 \begin{bmatrix} w_1^{-1} & 0 & \dots & 0 \\ 0 & w_2^{-1} & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & w_n^{-1} \end{bmatrix} = \sigma^2 \mathbf{W}^{-1} \quad (3)$$

em que \mathbf{W} é a matriz de pesos (conhecida) e σ^2 é desconhecido.

Cont.

- Sob heterocedasticidade temos que:
 - A variância de cada componente de $\hat{\beta}_W$ tende a ser menor do que a variância de cada componente de $\hat{\beta}$.
 - A metodologia ANOVA (da maneira como ela foi apresentada) não pode ser utilizada.
 - A metodologia baseada na estatística de Wald para testar ($H_0 : \mathbf{C}\beta = \mathbf{M}$ vs $H_1 : \mathbf{C}\beta \neq \mathbf{M}$) pode ser ainda utilizada, de modo semelhante ao caso anterior, com a seguinte modificação

$$Q_W = \left(\mathbf{C}\hat{\beta}_W - \mathbf{M} \right)' \left(\mathbf{C}(\mathbf{X}'\boldsymbol{\Sigma}^{-1}\mathbf{X})^{-1}\mathbf{C}' \right)^{-1} \left(\mathbf{C}\hat{\beta}_W - \mathbf{M} \right);$$

de modo que se $\boldsymbol{\Sigma}$ for conhecida, então $Q_W \sim \chi^2_{(c)}$ (exercício), sob H_0 .

Cont.

e $c = \text{rank}(\mathbf{C})$.

- Ainda para testar $H_0 : \mathbf{C}\boldsymbol{\beta} = \mathbf{M}$ vs $H_1 : \mathbf{C}\boldsymbol{\beta} \neq \mathbf{M}$, (se $\boldsymbol{\Sigma} = \sigma^2 \mathbf{W}^{-1}$), com σ^2 desconhecido e \mathbf{W} conhecido,

$$Q_W^* = \frac{1}{\hat{\sigma}_W^2 c} \left(\mathbf{C}\hat{\boldsymbol{\beta}}_W - \mathbf{M} \right)' \left(\mathbf{C}(\mathbf{X}'\mathbf{W}\mathbf{X})^{-1}\mathbf{C}' \right)^{-1} \left(\mathbf{C}\hat{\boldsymbol{\beta}}_W - \mathbf{M} \right)$$

em que, sob H_0 , $Q_W^* \sim F_{(c, n-p)}$

- O desenvolvimento é análogo ao caso homocedástico, utilizando-se (exercício)

$$\hat{\sigma}_W^2 = \frac{1}{n-p} (\mathbf{Y} - \mathbf{X}\hat{\boldsymbol{\beta}}_W)' (\mathbf{Y} - \mathbf{X}\hat{\boldsymbol{\beta}}_W). \quad (4)$$

ao invés de $\hat{\sigma}^2 = \frac{1}{n-p} (\mathbf{Y} - \mathbf{X}\hat{\boldsymbol{\beta}})' (\mathbf{Y} - \mathbf{X}\hat{\boldsymbol{\beta}})$

Comparação entre os estimadores MQO e MQP sob heterocedasticidade (exercício)

- MQO:

$$\hat{\beta} \sim N_p(\beta, (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\Sigma\mathbf{X}(\mathbf{X}'\mathbf{X})^{-1}). \quad (5)$$

Contudo, a obtenção das estimativas de MQO através de qualquer pacote computacional retornará

$$\text{Cov}(\hat{\beta}) = \sigma^2(\mathbf{X}'\mathbf{X})^{-1}(\text{estimada}) \neq (\mathbf{X}'\mathbf{X})^{-1}\mathbf{X}'\Sigma\mathbf{X}'(\mathbf{X}'\mathbf{X})^{-1}(\text{verdadeira}).$$

- MQP: $\hat{\beta}_W \sim N_p(\beta, (\mathbf{X}'\Sigma^{-1}\mathbf{X})^{-1})$. Obtendo, corretamente, as estimativas de MQO através de qualquer pacote computacional, obteremos uma estimativa apropriada de $\text{Cov}(\hat{\beta})$.

Alguns resultados relativos aos estimadores de MQP

- Se Σ for geral, equação (1), mas conhecida, então:
 - $IC(\beta_j, \gamma) = \left[\hat{\beta}_j - z_{\frac{1+\gamma}{2}} \sqrt{\psi_j}; \hat{\beta}_j + z_{\frac{1+\gamma}{2}} \sqrt{\psi_j} \right]$, em que $P(X \leq z_{\frac{1+\gamma}{2}}) = \frac{1+\gamma}{2}$, $X \sim N(0, 1)$ e ψ_j é o j-ésimo elemento da diagonal principal da matriz $(\mathbf{X}'\Sigma^{-1}\mathbf{X})^{-1}$.
 - Para testar $H_0 : \beta_j = \beta_{j0}$ vs $H_1 : \beta_j \neq \beta_{j0}$, para algum j, em que β_{j0} é um valor fixado. Estatística do teste: $Z_t = \frac{\hat{\beta}_j - \beta_{j0}}{\sqrt{\psi_j}}$, em que $\hat{\beta}_j$ é o estimador de MQP de β_j . Assim, rejeita-se H_0 se $|z_t| \geq z_c$, em que z_t é o valor calculado da estatística Z_t e $P(X \geq z_c | H_0) = \alpha/2$. De modo equivalente, rejeita-se H_0 se p-valor $\leq \alpha$, em que p-valor = $2P(X \geq |z_t| | H_0)$, $X \sim N(0, 1)$.

Alguns resultados relativos aos estimadores de MQP

- Se $\Sigma = \sigma^2 \mathbf{W}^{-1}$, com \mathbf{W} conhecido e σ^2 desconhecido, equação (3), então:
 - $IC(\beta_j, \gamma) = \left[\hat{\beta}_j - t_{\frac{1+\gamma}{2}} \sqrt{\hat{\sigma}_W^2 \psi_j}; \hat{\beta}_j + t_{\frac{1+\gamma}{2}} \sqrt{\hat{\sigma}_W^2 \psi_j} \right]$, em que $P(X \leq t_{\frac{1+\gamma}{2}}) = \frac{1+\gamma}{2}$, $X \sim t_{(n-p)}$ e ψ_j é o j -ésimo elemento da diagonal principal da matriz $(\mathbf{X}'\mathbf{W}\mathbf{X})^{-1}$ e $\hat{\sigma}_W^2$ é dado em (4).
 - Para testar $H_0 : \beta_j = \beta_{j0}$ vs $H_1 : \beta_j \neq \beta_{j0}$, para algum j , em que β_{j0} é um valor fixado. Estatística do teste: $T_t = \frac{\hat{\beta}_j - \beta_{j0}}{\sqrt{\hat{\sigma}_W^2 \psi_j}}$, em que $\hat{\beta}_j$ é o estimador de MQP de β_j . Assim, rejeita-se H_0 se $|t_t| \geq t_c$, em que t_t é o valor calculado da estatística T_t e $P(X \geq t_c | H_0) = \alpha/2$. De modo equivalente, rejeita-se H_0 se p-valor $\leq \alpha$, em que p-valor = $2P(X \geq |t_t| | H_0)$, $X \sim t_{(n-p)}$.

Comentários

- Na prática, utilizamos um estimador consistente de \mathbf{W} , ou seja :

$$\widehat{\mathbf{W}} \xrightarrow[n \rightarrow \infty]{P} \mathbf{W}.$$

- A estimação dos pesos \mathbf{W} depende da situação. Eles são inversamente proporcionais as variâncias de cada observação.
- Assim, os resultados dos slides 9 e 12, passam a ser assintóticas, substituindo-se \mathbf{W} por $\widehat{\mathbf{W}}$.
- Na função “lm” (do programa R), as estimativas de \mathbf{W} devem ser inseridas de modo a representarem o inverso das estimativas das variâncias.

Comparação entre os estimadores MQO e MQP sob heterocedasticidade

- Em princípio, a comparação entre as variâncias dos estimadores, para cada parâmetro, entre os dois métodos não é simples.
- Contudo, o mais importante é: na presença do heterocedasticidade, a utilização dos estimadores de MQO produzirá (um estimador) para a matriz de covariâncias (e conseqüentemente para os erros-padrão), de seus estimadores, diferente das respectivas verdadeiras fórmulas.
- MQO corrigido: utilizar a fórmula (5) para calcular os erros-padrão associados as estimativas de MQO (estimativas pontuais não mudam).

Voltando ao Exemplo 3

- Vamos considerar que $\Sigma = \sigma^2 \mathbf{W}^{-1}$ em que σ^2 é estimado usando (4). Para obter os pesos w_1, \dots, w_n :
 - Calculamos a variância amostral $s_j^2, j = 1, 2, 3, 4, 5$, da produção de milho para cada grupo (nível de fósforo).
 - Ajustamos o modelo
$$s_j^2 = \alpha_0 + \alpha_1 x_j + \alpha_2 x_j^2 + \xi_j, \xi_j \stackrel{i.i.d.}{\sim} N(0, \psi), j = 1, \dots, 5$$
 em que x_j é o j -ésimo nível de fósforo.
 - $\tilde{w}_i = (\tilde{\alpha}_0 + \tilde{\alpha}_1 x_i + \tilde{\alpha}_2 x_i^2)^{-1}, i = 1, \dots, 20$, em que x_i é o nível de fósforo associado à i -ésima observação e $\tilde{\alpha}_k, k = 0, 1, 2$ é a respectiva estimativa de MQO.

MQO

Parâmetro	Est.	EP	IC(95%)	Estat. t	p-valor
β_0	5,8225	0,5615	[4,6429 ; 7,0021]	10,3701	< 0,0001
β_1	0,0443	0,0092	[0,0251 ; 0,0636]	4,8338	0,0001

MQO (corrigido)

Parâmetro	Est.	EP	IC(95%)	Estat. t	p-valor
β_0	5,8225	0,9933	[3,7358 ; 7,9092]	5,8621	<0,0001
β_1	0,0443	0,0130	[0,0170 ; 0,0716]	3,4119	0,0031

MQP

Parâmetro	Est.	EP	IC(95%)	Estat. t	p-valor
β_0	6,7089	0,6255	[5,3947 ; 8,0231]	10,7249	<0,0001
β_1	0,0317	0,0077	[0,0154 ; 0,0479]	4,0885	0,0007

MQO

Parâmetro	Est.	EP	IC(95%)	Estat. t	p-valor
β_0	5,0182	0,6122	[3,7266 ; 6,3098]	8,1971	< 0,0001
β_1	0,1087	0,0290	[0,0475 ; 0,1699]	3,7460	0,0016
β_2	-0,0006	0,0003	[-0,0012 ; -0,0001]	-2,3132	0,0335

MQO (corrigido)

Parâmetro	Est.	EP	IC(95%)	Estat. t	p-valor
β_0	5,0182	1,2088	[2,4679 ; 7,5685]	4,1515	0,0007
β_1	0,1087	0,0448	[0,0141 ; 0,2033]	2,4238	0,0268
β_2	-0,0006	0,0004	[-0,0014 ; 0,0001]	-1,7782	0,0933

MQP

Parâmetro	Est.	EP	IC(95%)	Estat. t	p-valor
β_0	5,3844	0,8910	[3,5047 ; 7,2642	6,0434	0,0000
β_1	0,0853	0,0283	[0,0256 ; 0,1450]	3,0158	0,0078
β_2	-0,0004	0,0002	[-0,0009 ; 0,0000]	-1,9615	0,0664

Conclusões

- Nesse caso, pelo que vimos até agora no curso, devemos estimar os parâmetros por MQP.
- Passos seguintes: além de verificar a qualidade do ajuste de cada modelo, linear e quadrático, e compará-los, devemos apresentar as estimativas (pontuais e intervalares) e as previsões individuais (pontuais e intervalares) da produção de milho para cada nível de fósforo (exercício).
- Verificação da qualidade do ajuste. Deve-se adaptar os resíduos vistos para o caso anterior, bem como os respectivos programas (diag_norm; diag2_norm; envel_norm). Veja referências constantes no programa

Exemplo 11: Estudo (longitudinal) da eficácia de escovas de dentes

- Considere o estudo na área de Odontopediatria realizado na Faculdade de Odontologia da Universidade de São Paulo por Célia Regina Rodrigues e Symonne Parizotto.
- O objetivo é comparar duas escovas de dente (convencional e monobloco) com respeito à redução de um índice de placa bacteriana (IPB).
- Os valores obtidos correspondem a índices de placa bacteriana medidos nos dentes posteriores (pré-molares e molares) antes e depois da escovação dental de 32 crianças entre 4 e 6 anos de idade.
- O tipo de escova tende a ser melhor quanto maior for sua “capacidade de remoção” da placa bacteriana.

Dados

Tipo de escova	Sessão	Antes	Depois	Indivíduo
CT	1	1,05	1,00	1
CT	2	1,13	0,84	1
CT	3	1,15	0,86	1
CT	4	1,13	0,94	1
⋮	⋮	⋮	⋮	⋮
CT	1	1,4	1,12	16
CT	2	1,25	0,67	16
CT	3	1,5	1,10	16
CT	4	1,5	1,22	16
MT	1	1,66	1,63	17
MT	2	1,36	1,16	17
MT	3	1,52	0,88	17
MT	4	1,41	1,20	17
⋮	⋮	⋮	⋮	⋮
⋮	⋮	⋮	⋮	⋮
MT	1	1,15	1,00	32
MT	2	1,23	1,11	32
MT	3	1,15	1,07	32
MT	4	1,26	1,00	32



Medidas repetidas

- Este conjunto de dados se enquadra no que chamamos de **Medidas repetidas**.
- De fato, ele corresponde à uma estrutura chamada de **Dados longitudinais**.
- **Medidas repetidas**: quando medimos o(s) mesmo(s) indivíduo(s) em mais de uma condição de avaliação (tempo, distância, peso etc).
- **Dados longitudinais**: são medidas repetidas em que a condição de avaliação não pode ser aleatorizada (tempo, por exemplo).

Medidas repetidas & séries temporais

- Mesmo quando a condição de avaliação é o tempo, os dados do tipo **medidas repetidas** diferem-se daqueles que chamamos de **séries temporais**.
- **Medidas repetidas:** em geral temos muitos indivíduos e poucos instantes de avaliação.
- **Séries temporais:** poucos indivíduos e muitos instantes de avaliação.
- No caso em questão espera-se observar correlações entre as medidas (IPB) feitas nas mesmas crianças.

Dados

Por simplicidade, vamos considerar apenas duas sessões (instantes de avaliação)

Tipo de escova	Sessão	Antes	Depois	Indivíduo
CT	1	1,05	1,00	1
CT	2	1,13	0,84	1
⋮	⋮	⋮	⋮	⋮
CT	1	1,4	1,12	16
CT	2	1,25	0,67	16
MT	1	1,66	1,63	17
MT	2	1,36	1,16	17
⋮	⋮	⋮	⋮	⋮
MT	1	1,15	1,00	32
MT	2	1,23	1,11	32

Questões de interesse

- Comparar os desempenhos do tipo de escova em cada tempo e ao longo do tempo.
- O quanto cada escova é eficaz em reduzir o IPB.
- Problema: as observações (IPB pós escovação) não podem mais ser consideradas não-correlacionadas (sob normalidade seriam independentes).
- Com efeito, a correlação entre os IPB's pós escovação entre as duas sessões são: 0,612 (global), 0,661 (escova convencional) e 0,541 (escova monobloco).

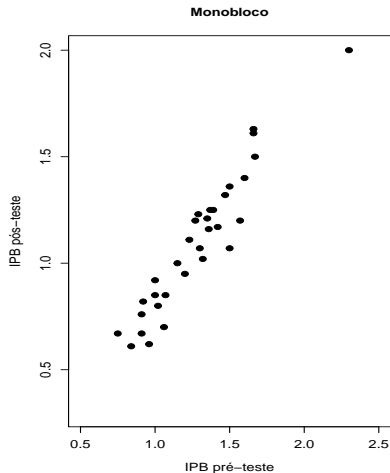
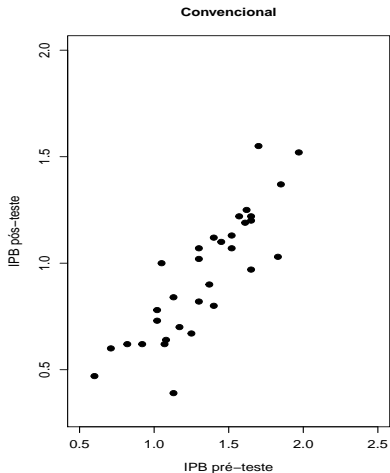
Medidas descritivas (pré-teste)

Tipo de escova	Sessão	Média	DP	Var.	CV(%)	Mínimo	Máximo	n
CT	1	1,31	0,35	0,12	26,40	0,71	1,97	16
CT	2	1,35	0,34	0,12	25,20	0,60	1,85	16
MT	1	1,33	0,38	0,14	28,17	0,75	2,30	16
MT	2	1,23	0,25	0,06	20,67	0,84	1,60	16

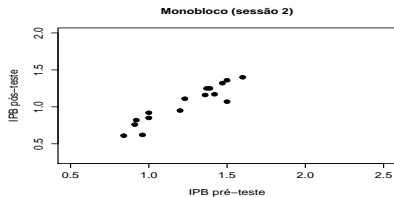
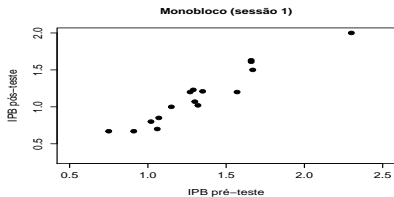
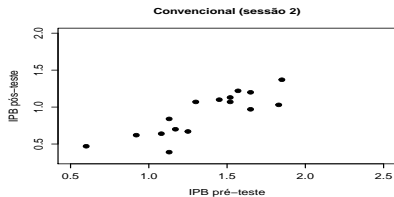
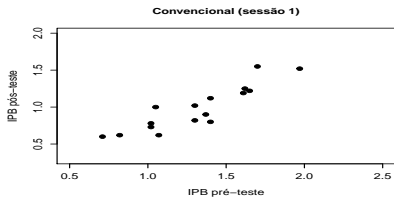
Medidas descritivas (pós-teste)

Tipo de escova	Sessão	Média	DP	Var.	CV(%)	Mínimo	Máximo	n
CT	1	0,98	0,30	0,09	30,87	0,60	1,55	16
CT	2	0,91	0,29	0,08	32,17	0,39	1,37	16
MT	1	1,15	0,39	0,15	33,55	0,67	2,00	16
MT	2	1,04	0,26	0,06	24,64	0,61	1,40	16

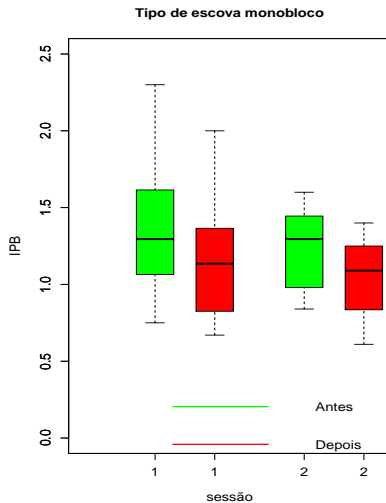
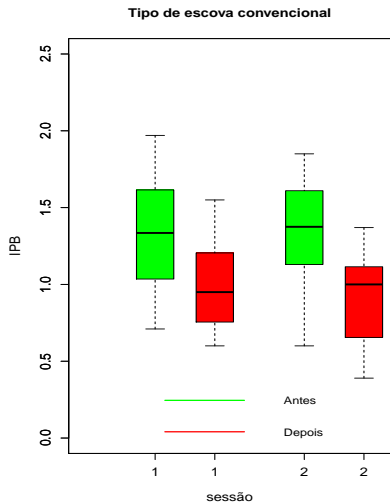
Gráf. de dispersão IPB (pré x pós) por tipo de escova



Gráf. de dispersão IPB (pré x pós) por tipo de escova e sessão



Box-plots do IPB



Modelo 1

$$Y_{ijk} = \alpha_{ij} + \beta_{ij}x_{ijk} + \xi_{ijk},$$

$i = 1$ (convencional (CT)), 2 (monobloco (MT)) (tipo de escova), $j = 1, 2$ (sessão), $k = 1, 2, \dots, 16$ (indivíduo)

- x_{ijk} : é o IPB da criança k , submetida ao tipo de escova i , na sessão j , antes da escovação.
- Y_{ijk} : é o IPB da criança k , submetida ao tipo de escova i , na sessão j , depois da escovação.
- α_{ij} : é o IPB pós escovação esperado quando se utiliza a escova do tipo i na j -ésima sessão.
- β_{ij} : é o incremento (positivo ou negativo) no IPB esperado pós escovação para crianças submetidas ao tipo de escova i na sessão j , para o aumento em uma unidade no IPB pré escovação.

Modelo 1 (cont.)

- $\beta = (\alpha_{11}, \alpha_{12}, \alpha_{21}, \alpha_{22}, \beta_{11}, \beta_{12}, \beta_{21}, \beta_{22})'$
- $\xi \sim N_{64}(\mathbf{0}, \Sigma)$.

$$\Sigma = \sigma^2 \begin{bmatrix} 1 & \rho & 0 & 0 & \dots & 0 \\ \rho & 1 & 0 & 0 & \dots & 0 \\ 0 & 0 & 1 & \rho & \dots & 0 \\ 0 & 0 & \rho & 1 & \dots & 0 \\ \vdots & \vdots & \vdots & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & 1 & \rho \\ 0 & 0 & 0 & 0 & \rho & 1 \end{bmatrix} = \mathbf{I}_{32} \otimes \sigma^2 \begin{bmatrix} 1 & \rho \\ \rho & 1 \end{bmatrix}$$

Forma matricial do MNL

$$\mathbf{Y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\xi}$$

$$\mathbf{Y} = \begin{bmatrix} Y_1 \\ Y_2 \\ \vdots \\ Y_n \end{bmatrix}, \mathbf{X} = \begin{bmatrix} X_{11} & \dots & X_{1p} \\ X_{21} & \dots & X_{2p} \\ \vdots & \ddots & \vdots \\ X_{n1} & \dots & X_{np} \end{bmatrix}, \boldsymbol{\beta} = \begin{bmatrix} \beta_1 \\ \beta_2 \\ \vdots \\ \beta_p \end{bmatrix}, \boldsymbol{\xi} = \begin{bmatrix} \xi_1 \\ \xi_2 \\ \vdots \\ \xi_n \end{bmatrix}$$

- Suponha agora que: $\boldsymbol{\xi} \sim N_n(\mathbf{0}, \boldsymbol{\Sigma})$, em que:

$$\boldsymbol{\Sigma} = \begin{bmatrix} \sigma_1^2 & \sigma_{12} & \dots & \sigma_{1n} \\ \sigma_{12} & \sigma_2^2 & \dots & \sigma_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ \sigma_{1n} & \sigma_{2n} & \dots & \sigma_n^2 \end{bmatrix} \quad (6)$$

Estimador de mínimos quadrados generalizados (MQG)

- Os desenvolvimentos relativos aos estimadores de mínimos quadrados generalizados (MQG) são semelhantes aqueles apresentados para a obtenção dos estimadores de MQP.
- Devemos utilizar a matriz (6) ao invés da matriz (1).
- No entanto, sempre consideraremos Σ desconhecido.
- Podemos provar que $\hat{\beta}_G = (\mathbf{X}'\Sigma^{-1}\mathbf{X})^{-1}\mathbf{X}'\Sigma^{-1}\mathbf{Y}$ e $\hat{\beta}_G \sim N_p(\beta_G, (\mathbf{X}'\Sigma^{-1}\mathbf{X})^{-1})$ (exercício).

Estimador de mínimos quadrados generalizados (MQG)

- Sob a estrutura (6) não é possível utilizar a metodologia ANOVA (como foi apresentada anteriormente).
- Na presença de correlações não nulas, a utilização dos estimadores de MQO/MQP produzirá (um estimador) para a matriz de covariâncias (e conseqüentemente para os erros-padrão), de seus estimadores, diferente das verdadeiras fórmulas.
- Além disso, mesmo corrigindo as estimativas das matrizes de covariâncias dos estimadores de MQO e MQP elas, em geral, fornecerão erros-padrão de seus estimadores maiores do que aqueles associados aos estimadores de MQG.

Estimador de mínimos quadrados generalizados (MQG)

- A metodologia baseada na estatística de Wald para testar $(H_0 : \mathbf{C}\boldsymbol{\beta} = \mathbf{M} \text{ vs } H_1 : \mathbf{C}\boldsymbol{\beta} \neq \mathbf{M})$ pode ser ainda utilizada, de modo semelhante ao caso anterior, com a seguinte modificação

$$Q_W = \left(\mathbf{C}\hat{\boldsymbol{\beta}}_W - \mathbf{M} \right)' \left(\mathbf{C}(\mathbf{X}'\hat{\boldsymbol{\Sigma}}^{-1}\mathbf{X})^{-1}\mathbf{C}' \right)^{-1} \left(\mathbf{C}\hat{\boldsymbol{\beta}}_W - \mathbf{M} \right)$$

em que $Q_W \approx \chi_{(c)}^2$, para n suficientemente grande, sob H_0 em que $\hat{\boldsymbol{\Sigma}}$ é um estimador consistente de $\boldsymbol{\Sigma}$ ($\hat{\boldsymbol{\Sigma}} \xrightarrow[n \rightarrow \infty]{P} \boldsymbol{\Sigma}$).

Alguns resultados relativos aos estimadores de MQP

- Além disso, temos que (resultados assintóticos):

- $IC(\beta_j, \gamma) = \left[\hat{\beta}_j - z_{\frac{1+\gamma}{2}} \sqrt{\hat{\psi}_j}; \hat{\beta}_j + z_{\frac{1+\gamma}{2}} \sqrt{\hat{\psi}_j} \right]$, em que

$P(X \leq z_{\frac{1+\gamma}{2}}) = \frac{1+\gamma}{2}$, $X \approx N(0, 1)$ e $\hat{\psi}_j$ é o j -ésimo elemento da diagonal principal da matriz $(\mathbf{X}'\hat{\Sigma}^{-1}\mathbf{X})^{-1}$.

- Para testar $H_0 : \beta_j = \beta_{j0}$ vs $H_1 : \beta_j \neq \beta_{j0}$, para algum j , em que β_{j0} é um valor fixado. Estatística do teste $Z_t = \frac{\hat{\beta}_j - \beta_{j0}}{\sqrt{\hat{\psi}_j}}$, em que $\hat{\beta}_j$ é o estimador de MQP de β_j . Assim, rejeita-se H_0 se $|z_t| \geq z_c$, em que z_t é o valor calculado da estatística Z_t e $P(X \geq z_c | H_0) = \alpha/2$. De modo equivalente, rejeita-se H_0 se p-valor $\leq \alpha$, em que p-valor = $2P(X \geq |z_t| | H_0)$, $X \approx N(0, 1)$

Voltando ao Exemplo 11

- Para cada modelo: estimou-se σ^2 através de

$$\hat{\sigma}^2 = \frac{1}{n-p} (\mathbf{Y} - \mathbf{X}\hat{\beta})' (\mathbf{Y} - \mathbf{X}\hat{\beta}) \quad (\hat{\beta} \text{ é o estimador de MQO}).$$

- $\tilde{\rho} = 0,612$ (correlação amostral).
- Ambos os estimadores são consistentes.

Resultados (MQO sem correção)

Param.	Estim.	EP	IC 95%	Estat. t	p-valor
α_{11}	-0,048	0,131	[-0,310 ; 0,215]	-0,362	0,7189
α_{12}	-0,161	0,124	[-0,409 ; 0,086]	-1,308	0,1964
α_{21}	-0,070	0,137	[-0,345 ; 0,204]	-0,5128	0,6101
α_{22}	-0,118	0,166	[-0,449 ; 0,214]	-0,7097	0,4808
β_{11}	0,785	0,097	[0,591 ; 0,979]	8,112	< 0,0001
β_{12}	0,981	0,090	[0,802 ; 1,160]	10,988	< 0,0001
β_{21}	0,722	0,099	[0,525 ; 0,920]	7,327	< 0,0001
β_{22}	0,941	0,132	[0,676 ; 1,205]	7,122	< 0,0001

Resultados (MQG)

Param.	Estim.	EP	IC 95%	Estat. Z	p-valor
$\alpha_{11}(S1 - CT)$	0,157	0,110	[-0,059 ; 0,373]	1,422	0,1550
$\alpha_{12}(S1 - MT)$	-0,155	0,106	[-0,362 ; 0,052]	-1,470	0,1417
$\alpha_{21}(S2 - CT)$	-0,009	0,115	[-0,235 ; 0,217]	-0,076	0,9392
$\alpha_{22}(S2 - MT)$	-0,126	0,140	[-0,401 ; 0,150]	-0,893	0,3717
$\beta_{11}(S1 - CT)$	0,630	0,080	[0,472 ; 0,787]	7,836	< 0,0001
$\beta_{12}(S1 - MT)$	0,976	0,075	[0,829 ; 1,124]	12,981	< 0,0001
$\beta_{21}(S2 - CT)$	0,677	0,082	[0,516 ; 0,837]	8,271	< 0,0001
$\beta_{22}(S2 - MT)$	0,947	0,111	[0,729 ; 1,165]	8,513	< 0,0001

Os resultados indicam que nenhum dos interceptos é diferente de zero e que alguns coeficiente angulares parecem ser iguais entre si.

Modelo 2

$$Y_{ijk} = \beta_{ij}x_{ijk} + \xi_{ijk},$$

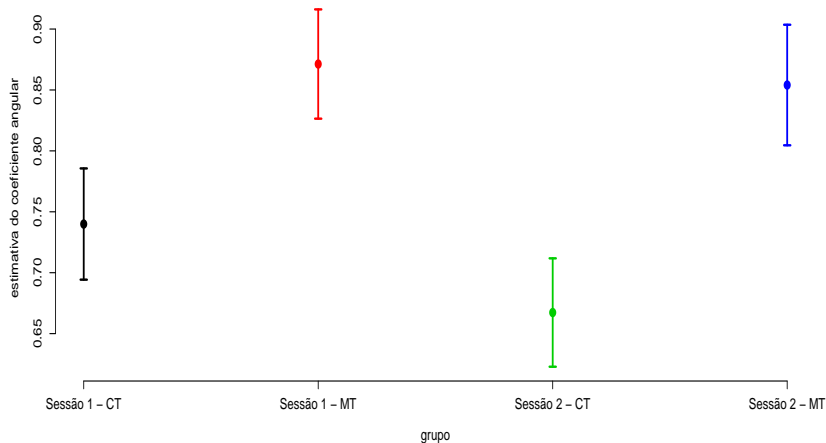
$i = 1$ (convencional (CT)), 2 (monobloco (MT)) (tipo de escova), $j = 1, 2$ (sessão), $k = 1, 2, \dots, 16$ (indivíduo)

- x_{ijk} : é o IPB da criança k , submetida ao tipo de escova i , na sessão j , antes da escovação.
- Y_{ijk} : é o IPB da criança k , submetida ao tipo de escova i , na sessão j , depois da escovação.
- β_{ij} : é a diminuição (se $\beta_{ij} \in (0, 1)$) ou o aumento (se $\beta_{ij} > 1$) no IPB esperado para crianças submetidas ao tipo de escova i na sessão j .
- $\xi \sim N_{64}(\mathbf{0}, \Sigma)$ com Σ igual ao anterior.

Resultados (MQG)

Param.	Estim.	EP	IC 95%	Estat. Z	p-valor
β_{11} (S1 - CT)	0,740	0,023	[0,694 ; 0,786]	31,770	< 0,0001
β_{12} (S1 - MT)	0,871	0,023	[0,826 ; 0,916]	38,074	< 0,0001
β_{21} (S2 - CT)	0,668	0,023	[0,623 ; 0,713]	29,402	< 0,0001
β_{22} (S2 - MT)	0,854	0,025	[0,805 ; 0,904]	33,830	< 0,0001

Estimativas pontuais e intervalares dos parâmetros



Resultados

- Igualdade simultânea entre os coeficientes angulares dos grupos Sessão 1 - CT e Sessão 2 - CT e dos grupos Sessão 1 - MT e Sessão 2 - MT ($H_0 : \beta_{11} = \beta_{21}$ e $\beta_{12} = \beta_{22}$): 12,89 (0,0016).
- Igualdade entre os coeficientes angulares dos grupos Sessão 1 - CT e Sessão 2 -CT($H_0 : \beta_{11} = \beta_{21}$): 12,26 ($< 0,0001$).
- Igualdade simultânea entre os coeficientes angulares dos grupos Sessão 1 - MT e Sessão 2 - MT ($H_0 : \beta_{12} = \beta_{22}$): 0,63 (0,4279).

Modelo 3

$$Y_{ijk} = \beta_{ij}x_{ijk} + \xi_{ijk},$$

$i = 1$ (convencional (CT)), 2 (monobloco (MT)) (tipo de escova), $j = 1, 2$ (sessão), $k = 1, 2, \dots, 16$ (indivíduo)

- x_{ijk} : é o IPB da criança k , submetida ao tipo de escova i , na sessão j , antes da escovação.
- Y_{ijk} : é o IPB da criança k , submetida ao tipo de escova i , na sessão j , depois da escovação.
- β_{ij} : é a diminuição (se $\beta_{ij} \in (0, 1)$) ou o aumento (se $\beta_{ij} > 1$) no IPB esperado para crianças submetidas ao tipo de escova i na sessão j ($\beta_{12} = \beta_{22}$).
- $\xi \sim N_{64}(\mathbf{0}, \Sigma)$ com Σ igual ao anterior.

Resultados (MQG)

Param.	Estim.	EP	IC 95%	Estat. t	p-valor
β_{11} (S1 - CT)	0,740	0,023	[0,695 ; 0,785]	31,947	< 0,0001
β_{12} (S1 & 2 - MT)	0,865	0,021	[0,823 ; 0,906]	40,738	< 0,0001
β_{21} (S2 - CT)	0,667	0,023	[0,623 ; 0,712]	29,566	< 0,0001

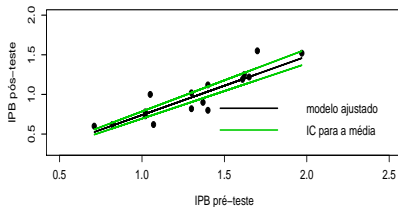
Conclusão: A escova tradicional apresenta o melhor desempenho, sendo este obtido na sessão 2. De fato, o teste para equivalência da capacidade de redução do IPB da escova CT entre as duas sessões ($H_0 : \beta_{11} = \beta_{21}$) indicou diferença: 12,4 (< 0,0001).

Resumindo

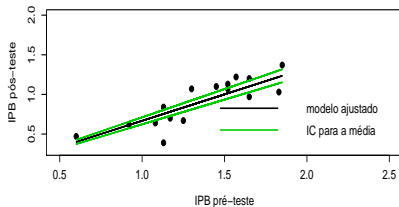
- Sob heterocedasticidade: usar um estimador consistente para Σ ou utilizar modelos de regressão heterocedásticos ($h(\sigma_i^2) = \mathbf{W}'_i\delta$).
- Sob correlações não nulas (com ou sem homocedasticidade): usar um estimador consistente para Σ , modelo mistos (hierárquicos)
http://www.ime.unicamp.br/~cnaber/Material_ADL_POS_2S_2018.htm ou de séries temporais.

Reta ajustada e IC para as médias

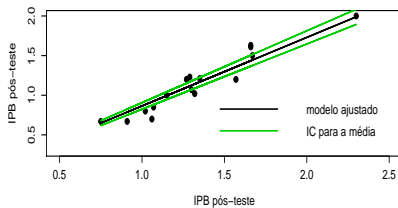
Convencional (sessão 1)



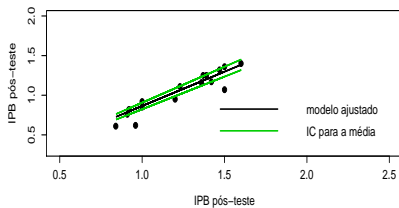
Convencional (sessão 2)



Monobloco (sessão 1)



Monobloco (sessão 2)



Comandos no R

- Mínimos quadrados ponderados: introduzir os pesos (inversamente proporcionais às variâncias das observações), através do comando “`weights`” na função “`lm`”.
- Mínimos quadrados generalizados: funções “`glm`” e “`glmer`” disponíveis no arquivo. Há também a função “`glsp`”.
- Mais sobre o assunto pode ser encontrado em http://www.ime.unicamp.br/~cnaber/Material_ADL_POS_2S_2018.htm