Lista 2

Exercício 1. *Suponha que* $X \sim N(7, 25)$*. Em uma amostra de tamanho* $N = 100$*, calcule* $Pr\,(6 < X < 8)$ *e* $Pr\left(6 < \bar{X} < 8\right)$*. Explique por que as probabilidades diferem.*

Resposta: Para uma amostra de tamanho $N = 100$, temos que

$$\bar{X} \sim N\left(7, \frac{25}{100}\right) \iff \bar{X} \sim N\left(7, \frac{1}{4}\right)$$

Logo, as probabilidades são:

$$Pr[6 < X < 8] = Pr\left[-\frac{1}{5} < \frac{x-7}{5} < \frac{1}{5}\right] = 0,159$$
$$Pr[6 < \bar{X} < 8] = Pr\left[-2 < \frac{\bar{x}-7}{5} < 2\right] = 0,954$$

Diferem por conta da menor variância da média. Conforme N cresce, valores da média próximos à esperança ficam cada vez mais prováveis (princípio LGN). □

Exercício 2. *Suponha que dois estatísticos queiram estimar um parâmetro* $\theta > 0$ *desconhecido. O estatístico A observa apenas uma realização* $x = 2$ *de uma variável aleatória* $X \sim Gamma(3, \theta)$*, ao passo que o estatístico B observa uma realização* $y = 3$ *de* $Y \sim Poisson(2\theta)$*. Mostre que as funções de verossimilhança determinadas pelas duas situações são proporcionais e, portanto, terão mesmo valor máximo. Qual o valor de* $\hat{\theta}$ *obtido por ambos nesse caso?*

Resposta: A densidade de $X \sim \Gamma(3, \theta)$ é dada por

$$f(x) = \frac{\theta^3}{\Gamma(3)} x^2 e^{-\theta x},$$

então, para $x = 2$, segue que

$$f(2) = \frac{\theta^3}{\Gamma(3)} 2^2 e^{-2\theta} = \frac{4}{\Gamma(3)} \theta^3 e^{-2\theta}.$$

Já a densidade de $Y \sim Poisson(2\theta)$ é

$$g(y) = \frac{e^{-2\theta}(2\theta)^y}{y},$$

que, para $y = 3$, temos

$$g(3) = \frac{4}{3}\theta^3 e^{-2\theta} \quad \propto f(2)$$

Precisamos mostrar que as funções de verossimilhança proporcionais, tais que

$$\arg\max_{\theta>0} k\theta^3 e^{-2\theta} \quad = \quad \hat{\theta} \quad = \quad \arg\max_{\theta>0} 2\ln\theta - 2\theta$$

Resolvendo por CPO, temos:

$$\begin{aligned} \frac{\partial k\theta^3 e^{-2\theta}}{\partial\theta} &= ke^{-2\theta}\theta^2(2\theta-3) = 0 \\ \hat{\theta} &= \frac{3}{2} \quad \text{(pois } \theta > 0) \end{aligned}$$

$$\begin{aligned} \frac{\partial 3\ln\theta - 2\theta}{\partial\theta} &= \frac{3}{\theta} - 2 = 0 \\ \hat{\theta} &= \frac{3}{2} \end{aligned}$$

□

Exercício 3. *Suponha que Y_1 e Y_2 sejam variáveis aleatórias independentes, com esperança comum, μ, mas variâncias distintas (e conhecidas), σ_1^2 e σ_2^2. Você deseja estimar μ através de uma média ponderada de Y_1 e Y_2, isto é, $\hat{\mu} = \alpha Y_1 + (1-\alpha)Y_2$. Mostre que $\hat{\mu}$ é não viesado e obtenha o valor de α que minimiza a variância de $\hat{\mu}$. Elabore uma intuição para seu resultado.*

Resposta:

$$\begin{aligned} E(\hat{\mu}) &= \alpha E(Y_1) + (1-\alpha)E(Y_2) = \mu(\alpha + 1 - \alpha) = \mu \\ Var(\hat{\mu}) &= \alpha^2\sigma_1^2 + (1-\alpha)^2\sigma_2^2 \end{aligned}$$

Para encontrar o valor de α que minimiza a variância de $\hat{\mu}$, iremos solucionar

$$\hat{\hat{\alpha}} = \arg\max_{\alpha\in[0,1]} Var(\hat{\mu})$$

por CPO:

$$\begin{aligned} 2\hat{\hat{\alpha}}\sigma_1^2 - 2(1-\hat{\hat{\alpha}})\sigma_2^2 = 0 \\ \hat{\hat{\alpha}} = \frac{\sigma_2^2}{\sigma_1^2 + \sigma_2^2} \end{aligned}$$

Intuição:

- V.A. com menor variância tem maior peso
- ainda assim, ambas têm peso positivo, pois conseguem conteúdo novo (independente) de informação sobre μ.

□

Exercício 4. *Suponha que* $X \sim \exp(\theta)$, *isto é, que* $f(x) = \theta e^{-\theta x}$.

a) Qual a esperança de X*? Como você construiria um estimador de método dos momentos de* θ *usando essa informação?*

b) Qual a variância de X*? Como você construiria um estimador de método dos momentos de* θ *usando essa informação?*

c) Você seria capaz de construir um estimador de θ *que utilizasse ao mesmo tempo as informações de (a) e (b)?*

d) Qual o estimador de Máxima Verossimilhança para θ*? Ele é melhor ou pior do que os propostos nos itens (a)-(c)? Por quê?*

Resposta:

a) $E(X) = \theta^{-1} \implies$ análogo amostral: $\hat{\theta} = \frac{1}{\bar{x}}$

b) $V(X) = \theta^{-1} \implies$ análogo amostral: $\hat{\theta} = \frac{1}{\frac{1}{N}\sum(x_i - \bar{x})^2}$

c) De (a) e (b), temos na população:

$$\theta - E^{-1}(X) = 0 \qquad \text{e} \qquad \theta - V^{-1}(X) = 0$$

Seja

$$M = \begin{bmatrix} \theta - {}^{1}\!/_{\bar{x}} \\ \theta - {}^{1}\!/_{\hat{V}(X)} \end{bmatrix},$$

então quero encontrar $\hat{\theta}$ que maximize a chance de $M = [0 \quad 0]'$.

Seja W uma matriz positiva-definida de pesos, então o estimador é dado por

$$\hat{\theta} = \arg\min_{\theta > 0} M'(\theta) W M(\theta)$$

d) A máxima verossimilhança é dada por

$$\mathscr{L}(\theta) = \prod_{i=1}^{N} \theta e^{-\theta x_i}$$

Logo, o estimador é

$$\hat{\theta} = \arg\max \left\{ N \ln \theta - \theta \sum x_i \right\}$$

Por CPO:

$$\hat{\theta} = \frac{1}{\bar{x}}$$

- Atinge o limite inferior de Cramer-Rao

- A fórmula é igual ao do item (a), mas as informações à priori de $f(x)$ é que garante que $\bar{x}^{-1}$ atinge o limite inferior de Cramer-Rao.

$\square$

Exercício 5. *Suponha que* $f(x) = (\alpha + 1)\, x^{\alpha}$ *para* $x \in [0, 1]$.

a) Proponha um estimador de método dos momentos para α *que use* $E(X)$ *como informação obtida à priori.*

b) Se sua amostra for $X = [0,1 \quad 0,4 \quad 0,2 \quad 0,7 \quad 0,3]'$, *qual seria seu estimador de* α *pela estratégia proposta em (a)?*

Resposta:

a)

$$E(X) = \int_0^1 x f(x)\, dx = (a+1) \int_0^1 x^{\alpha+1}\, dx = \left[\frac{\alpha+1}{\alpha+2}.x^{\alpha+2}\right]_0^1 = \frac{\alpha+1}{\alpha+2}$$

Logo,

$$\begin{aligned} \frac{\hat{\alpha}+1}{\alpha \hat{+} 2} = \bar{x} \qquad \Longleftrightarrow \qquad \hat{\alpha}+1 &= \bar{x}(\hat{\alpha}+2) \\ (1-\bar{x})\hat{\alpha} &= 2\bar{x}-1 \\ \hat{\alpha} &= \frac{2\bar{x}-1}{1-\bar{x}} \end{aligned}$$

b)

$$\begin{aligned} X &= [0,1 \quad 0,4 \quad 0,2 \quad 0,7 \quad 0,3]' \Longrightarrow \bar{x} = \frac{17}{50} \\ \hat{\alpha} &= \frac{2\bar{x}-1}{1-\bar{x}} = \frac{\frac{34}{50}-1}{1-\frac{17}{50}} = -\frac{16}{33} \end{aligned}$$

$\square$

Exercício 6. *Suponha que* $X \sim N(\mu, \sigma^2)$, *e que você tenha acesso a uma amostra aleatória* $X_1, ..., X_N$ *de realizações de* X. *Mostre como sua estimativa de máxima verossimilhança de* μ *depende de* σ.

Resposta: Não depende. $\square$

Exercício 7. *Suponha que* $f(x) = \frac{1}{b}; \quad 0 < X < b$. *Você tem acesso a uma amostra aleatória de tamanho* N *de* X.

a) Obtenha um estimador de máxima verossimilhança para b.

b) Obtenha um estimador de método dos momentos para b.

Resposta: $X \sim U[0, b]$

a)

$$f(x) = \frac{1}{b}\mathbb{1}(0, b),$$

em que $\mathbb{1}$ é uma função indicadora (função lógica que assume valor 1 se argumento é verdadeiro e 0 em caso contrário)

Logo,

$$\mathscr{L}(b) = \left(\frac{1}{b}\right)^N \prod_{i=1}^{N} \mathbb{1}_{0 \leq X_i \leq b} = \begin{cases} \left(\frac{1}{b}\right)^N, & \text{se } \min\{X_i\} \geq 0 \text{ e } \max\{X_i\} \leq b \\ 0, & \text{caso contrário.} \end{cases}$$

Resposta: $\hat{b}^{MV} = \max\{X_i\}$.

Para ver isto, considere inicialmente $\tilde{b} < \max\{X_i\}$. Então, para $x^* \equiv \max\{X_i\}$, teremos

$$Pr(X = x^* | \tilde{b}) = 0$$

Então,

$$\mathscr{L}(\tilde{b}) = 0,$$

pois basta um elemento do produtório ser nulo para que ele todo seja nulo. Agora, considere $\tilde{b} > \max\{X_i\}$. Nesse caso:

$$\mathscr{L}(\tilde{b}) = \frac{1}{\tilde{b}^N} < \frac{1}{\max\{X_i\}^N},$$

pois $\max\{X_i\}^N < b^N$.

b)

$$f(x) = \frac{1}{b} : \quad 0 < x < b \quad \Longrightarrow \quad E(X) = \frac{b}{2} \text{ ou } b = 2E(X)$$
$$\text{MM} : \quad \hat{b} = 2\bar{X}$$

□

Exercício 8. *(Revisão da aula) Mostre que...*

a) O erro quadrático médio, $E\left[\hat{\theta} - \theta_0\right]^2$*, equivale à soma da variância e do quadrado do viés, isto é,* $V\left(\hat{\theta}\right) + \left[E\left(\hat{\theta}\right) - \theta_0\right]^2$.

b) Os seguintes testes são assintoticamente equivalentes para $H_0 : \beta = b$ *no contexto de uma estimação por máxima verossimilhança:*

- $w_{LR}(\hat{\theta}) = -\frac{2}{N}\ln\left[\frac{\mathscr{L}(b)}{\mathscr{L}(\hat{\theta})}\right]$
- $w_{\lambda}(\widetilde{\theta}) = \left[\sqrt{N}\frac{1}{N}S_N\left(b\right)\right]'\left[\mathscr{I}\left(b\right)\right]^{-1}\left[\sqrt{N}\frac{1}{N}S_N\left(b\right)\right] = \lambda'\left[\mathscr{I}\left(b\right)\right]^{-1}\lambda$
- $w(\hat{\theta}) = \left[\hat{\theta} - b\right]'\left[N^{-1}R\mathscr{I}^{-1}(\hat{\theta})R'\right]^{-1}\left[\hat{\theta} - b\right]$

<u>*Resposta*</u>:

a)

$$
\begin{aligned}
E\left[(\hat{\theta}-\theta_0)^2\right] &= E\left[\left(\left(\hat{\theta}-E(\hat{\theta})\right)+\left(E(\hat{\theta})-\theta_0\right)\right)^2\right] \\
&= E\left[\left(\hat{\theta}-E(\hat{\theta})\right)^2\right]+\left(E(\hat{\theta})-\theta_0\right)^2+2E\left(\hat{\theta}-E(\hat{\theta}\right)\left(E(\hat{\theta})-\theta_0\right)
\end{aligned}
$$

Note que $E(\hat{\theta})$ é constante, pois $E(a) = a$. Como $E\left(\hat{\theta} - E(\hat{\theta})\right) = E(\hat{\theta}) - E(\hat{\theta}) = 0$, então o último termo é igual a zero. Logo,

$$
E\left[(\hat{\theta}-\theta_0)^2\right] = E\left[\left(\hat{\theta}-E(\hat{\theta})\right)^2\right]+\left(E(\hat{\theta})-\theta_0\right)^2
$$

b) Seja $\hat{\theta}$ o estimador de MV e $H_0 : \theta_0 = \tilde{\theta}$

i) <u>Wald</u>: $w(\hat{\theta}) = (\hat{\theta}-\theta_0)'\mathscr{I}(\theta_0)(\hat{\theta}-\theta_0) \overset{A}{\sim} \chi^2_{(G)}$

ii) <u>LR</u>:

$$
w_{LR}(\hat{\theta}) = -2\ln\left[\frac{\mathscr{L}(\theta_0)}{\mathscr{L}(\hat{\theta})}\right] = -2\left[\ell(\theta_0)-\ell(\hat{\theta})\right]
$$

Note que, por expansão de Taylor, temos:

$$
\begin{aligned}
\frac{1}{N}\ell_N(\hat{\theta}) &\approx \frac{1}{N}\ell_N(\theta_0)+\frac{1}{N}\ell'_N(\theta_0)(\hat{\theta}-\theta_0)+\frac{1}{2}(\hat{\theta}-\theta_0)^{\intercal}\frac{1}{N}\ell''_N(\theta_0)(\hat{\theta}-\theta_0) \\
\frac{1}{N}\left[\ell_N(\hat{\theta})-\ell_N(\theta_0)\right] &\approx \frac{1}{N}\ell'_N(\theta_0)(\hat{\theta}-\theta_0)+\frac{1}{2}(\hat{\theta}-\theta_0)^{\intercal}\frac{1}{N}\ell''_N(\theta_0)(\hat{\theta}-\theta_0)
\end{aligned}
$$

Como

$$
\frac{1}{N}\ell'_N(\theta_0)(\hat{\theta}-\theta_0) \overset{P}{\to} \cancelto{0}{E[S(\theta_0)]}.(\hat{\theta}-\theta_0)
$$

e

$$
\frac{1}{N}\ell''_N(\theta_0)(\hat{\theta}-\theta_0) \overset{P}{\to} \mathscr{I}(\theta_0),
$$

então

$$
\begin{aligned}
\frac{1}{N}\left[\ell_N(\hat{\theta})-\ell_N(\theta_0)\right] &\approx 0+\frac{1}{2}(\hat{\theta}-\theta_0)^{\intercal}\mathscr{I}(\hat{\theta})(\hat{\theta}-\theta_0) \\
\frac{2}{N}\left[\ell_N(\hat{\theta})-\ell_N(\theta_0)\right] &\approx (\hat{\theta}-\theta_0)^{\intercal}\mathscr{I}(\hat{\theta})(\hat{\theta}-\theta_0)
\end{aligned}
$$

Como $-\frac{2}{N}\left[\ell_N(\hat{\theta}) - \ell_N(\theta_0)\right] = \frac{2}{N}\left[\ell_N(\hat{\theta}) - \ell_N(\theta_0)\right]$, segue que

$$-\frac{2}{N}\left[\ell_N(\hat{\theta}) - \ell_N(\theta_0)\right] \approx (\hat{\theta} - \theta_0)^\intercal \mathscr{I}(\hat{\theta})(\hat{\theta} - \theta_0) = w(\hat{\theta})$$

iii) LM:

$$w_\lambda(\hat{\theta}) = \left[\sqrt{N}\frac{1}{N}S_N(\hat{\theta})\right]' \left[\mathscr{I}(\hat{\theta})\right]^{-1} \left[\sqrt{N}\frac{1}{N}S_N(\hat{\theta})\right]$$

Note que, por expansão de Taylor, temos que

$$\frac{1}{N}S_N(\hat{\theta}) \approx \frac{1}{N}S_N(\theta_0) + \frac{1}{N}S'_N(\theta_0)(\hat{\theta} - \theta_0)$$

Como

$$\frac{1}{N}S_N(\theta_0) \xrightarrow{P} E[S(\theta_0)] = 0$$

e

$$\frac{1}{N}S'_N(\theta_0) \xrightarrow{P} E[S'(\theta_0)] = E[\mathscr{H}_\ell(\theta_0)],$$

então

$$\frac{1}{N}S_N(\hat{\theta}) \approx 0 + E[\mathscr{H}_\ell(\theta_0)](\hat{\theta} - \theta_0)$$

Substituindo em $w_\lambda(\hat{\theta})$, segue que

$$\begin{aligned} w_\lambda(\hat{\theta}) &\approx \left[\sqrt{N}E[\mathscr{H}_\ell(\theta_0)](\hat{\theta} - \theta_0)\right]' \left[\mathscr{I}(\hat{\theta})\right]^{-1} \left[\sqrt{N}E[\mathscr{H}_\ell(\theta_0)](\hat{\theta} - \theta_0)\right] \\ &\approx N\left[E[\mathscr{H}_\ell(\theta_0)](\hat{\theta} - \theta_0)\right]' \left[\mathscr{I}(\hat{\theta})\right]^{-1} \left[E[\mathscr{H}_\ell(\theta_0)](\hat{\theta} - \theta_0)\right] \end{aligned}$$

Dado que $E[\mathscr{H}_\ell(\theta_0)] = -\mathscr{I}(\theta_0)$, temos

$$\begin{aligned} w_\lambda(\hat{\theta}) &\approx N\left[-\mathscr{I}(\theta_0)(\hat{\theta} - \theta_0)\right]' \left[\mathscr{I}(\hat{\theta})\right]^{-1} \left[-\mathscr{I}(\theta_0)(\hat{\theta} - \theta_0)\right] \\ &\approx N(\hat{\theta} - \theta_0)' \left[\mathscr{I}(\theta_0)\right]' \underbrace{\left[\mathscr{I}(\hat{\theta})\right]^{-1} \mathscr{I}(\theta_0)}_{I}(\hat{\theta} - \theta_0) \\ &\approx N(\hat{\theta} - \theta_0)'\mathscr{I}(\theta_0)(\hat{\theta} - \theta_0) = w(\hat{\theta}) \end{aligned}$$

□

Exercício 9. *Suponha que você tenha acesso a uma amostra de tamanho N de uma V.A. Y tal que $Y_i = \beta + \varepsilon_i$, onde $Pr(\varepsilon = 1) = Pr(\varepsilon = -1) = \frac{1}{2}$.*

a) Mostre que $\bar{Y}$ é um estimador não-viesado para β.

b) *Suponha que você saiba a priori que $\beta = 2$ ou $\beta = 0$, e queira apenas estimar qual destes dois valores é o verdadeiro. Considere o estimador alternativo:*

- $\hat{\beta} = 0$ *se ao menos um* $Y_i = -1$
- $\hat{\beta} = 2$ *se ao menos um* $Y_i = 3$
- $\hat{\beta} = \bar{Y}$ *em caso contrário*

Qual estimador é melhor, $\bar{Y}$ *ou* $\hat{\beta}$*? Por quê?*

Resposta:

$$E[\bar{Y}] = E\left[\frac{1}{N}\sum E(Y_i)\right] = \frac{1}{N}\sum E(Y_i)\frac{1}{N}\sum \beta + E(\varepsilon_i) = \beta + \frac{1}{N}\sum E(\varepsilon_i)$$

a) $E(\varepsilon_i) = \frac{1}{2}(-1) + \frac{1}{2}(1) = 0$

b) Y_i apenas possui 3 valores: $\{-1, 1, 3\}$:

- se $Y_i = -1 \implies$ para algum i, $\hat{\beta}$ precisa ser 0
- se $Y_i = 3 \implies$ para algum i, $\hat{\beta}$ precisa ser 2
- se $Y_i = 1 \implies$ para todo i, $\hat{\beta}$ será $\bar{Y}$

Portanto, $\hat{\beta}$ é tão bom quanto $\bar{Y}$ se $\bar{Y} = 1$ e é melhor nos demais casos.

□

Exercício 10. *Suponha que em uma amostra de 21 observações de $X \sim \exp(\theta)$, você saiba que a média das 20 primeiras é 6 e que a última é maior do que 15. Como você construiria um estimador de máxima verossimilhança para θ? Qual a resposta para $\hat{\theta}$?*

Resposta:

$$\mathscr{L}(\theta) = \prod_{i=1}^{20} \theta e^{-\theta X_i} . \underbrace{\left(e^{-15\theta}\right)}_{Pr(X>15)}$$

$$\begin{aligned}
\ell(\theta) &= \sum_{i=1}^{20} [\ln(\theta) - \theta X_i] - 15\theta \\
&= 20\ln(\theta) - 20.\theta.\frac{1}{20}\sum_{i=1}^{20} X_i - 15\theta \\
&= 20\ln\theta - 20.\theta.6 - 15\theta = 20\ln\theta - 135\theta
\end{aligned}$$

Maximizando por CPO:

$$\frac{20}{\hat{\theta}} - 135 = 0 \iff \hat{\theta} = \frac{20}{135} = \frac{4}{27}$$

□

Exercício 11. *Suponha que* $y = a+bx+\varepsilon$*, e que* $E[\varepsilon|X] = 0$*. Mostre que* $\hat{a}^{MQO} = \bar{y} - \frac{\widehat{cov}(y,x)}{\widehat{var}(x)}\bar{x}$ *é uma estimativa consistente de* a*.*

Resposta: Note que $\hat{a} = \frac{1}{N}\sum_{i=1}^{N} Y_i - \hat{b}\frac{1}{N}\sum X_i$. Vimos em sala de aula que $\hat{b} = \frac{\widehat{cov}(y,x)}{\widehat{var}(x)} \xrightarrow{P} b$ e sabemos pela Lei dos Grandes Números que $\frac{1}{N}\sum_{i=1}^{N} Y_i \xrightarrow{P} E(Y)$ e que $\frac{1}{N}\sum_{i=1}^{N} X_i \xrightarrow{P} E(X)$. Por Slutsky, $\hat{a} \xrightarrow{P} E(Y) - bE(X)$.

Agora, tomando a esperança em $y = a + bx + \varepsilon$, segue que:

$$\begin{aligned}
E(Y) &= a + bE(X) + E(\varepsilon) \\
E(Y) &= a + bE(X) + 0 && (E(\varepsilon|X) = 0 \implies E(\varepsilon) = 0) \\
a &= E(Y) - bE(X)
\end{aligned}$$

Logo, $\hat{a} \xrightarrow{P} a$.
$\square$

Exercício 12. *Uma variável* X *tem distribuição normal, mas você não sabe se* $N(1,3)$ *ou se* $N(2,3)$*. A partir de uma amostra aleatória com* $N = 10$*, você formula um teste onde* $H_0 : \mu = 1$ *e* $H_1 : \mu = 2$*. Pensando em um teste de* H_0*:*

a) Qual o ponto de corte para H_0 *coerente com um erro do tipo I de no máximo 10%?*

b) No caso acima, qual a probabilidade do erro do tipo II?

c) Como suas respostas mudam se $N = 100$*? Mantendo a probabilidade de erro do tipo I fixa, o que você aprende sobre o comportamento do ponto de corte quando* N *aumenta?*

Resposta:

a) Erro do Tipo I = Pr(Rejeitar $H_0|H_0$ verdadeira)
Sob H_0: $\bar{X} \sim N\left(1, \frac{3}{10}\right)$ e c^* resolve

$$Pr[\bar{X} \geq c^*] \leq 0,1$$
$$Pr\left[\frac{\bar{X}-1}{\sqrt{3/10}} \geq \frac{c^*-1}{\sqrt{3/10}}\right] \leq 0,1$$

Sejam $z \equiv \frac{\bar{X}-1}{\sqrt{3/10}}$ e $z^* \equiv \frac{c^*-1}{\sqrt{3/10}}$, logo

$$Pr[z \geq z^*] \leq 0,1 \implies z^* = 1,28$$

Portanto,

$$z^* = 1,28 \iff \frac{c^*-1}{\sqrt{3/10}} = 1,28 \iff c^* = 1,70$$

b) Erro do Tipo II = Pr(Aceitar $H_0|H_1$ verdadeira)

$$Pr\left[\frac{\bar{x}-2}{\sqrt{3/10}} \leq \frac{c^*-2}{\sqrt{3/10}}\right] = Pr[z \leq -0,55] = 0,71$$

c)

$$Pr\left[z \geq \frac{c^*-1}{\sqrt{3/100}}\right]$$

Seja $z^* \equiv \frac{c^*-1}{\sqrt{3/100}}$, logo

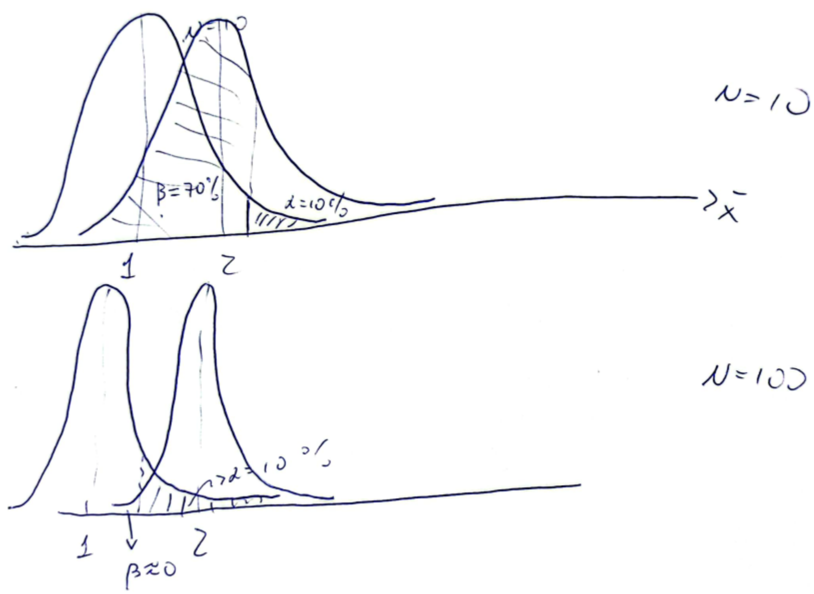
$$Pr[z \geq z^*] \leq 0,1 \implies z^* = 1,28$$

Portanto,

$$z^* = \frac{c^*-1}{\sqrt{3/100}} = \frac{10(c^*-1)}{\sqrt{3}} \implies c^* = 0,22$$

Assim,

$$Pr\left[z \leq \frac{c^*-2}{\sqrt{3/100}}\right] = Pr\left[z \leq \frac{-1,78}{\sqrt{3/100}}\right] = Pr\left[z \leq -10,28\right] \approx 0$$



□

Exercício 13. *A decisão a respeito de publicar um artigo é tipicamente baseada em pareceres de especialistas solicitados pelo editor de uma revista. Em estudo encomendado pela American Economic Association, referente a submissões à American Economic Review, suspeitou-se que as taxas de aceitação de artigos eram maiores quando a submissão não era cega, ou seja, quando os pareceristas sabiam quem estava submetendo o artigo. Neste estudo, um número substancialmente grande de artigos cegos e não-cegos foram aleatoriamente submetidos a diferentes pareceristas para o propósito do estudo. Os dados foram:*

Taxas de aceitação
(erro padrão estimado entre parênteses)

Amostra cega	Amostra não-cega	
10,6%	14,1%	Total
(1,1)	(1,4)	
10,0%	11,2%	Autoras mulheres
(3,0)	(3,5)	
11,0%	15,0%	Autores homens
(1,2)	(1,5)	

a) A aleatorização foi bem sucedida? Como você testaria o sucesso do procedimento de aleatorização?

b) Teste a hipótese nula de que a submissão ser cega não influencia o resultado de aceitação do artigo.

c) Calcule o p-valor do teste na parte (b)

d) Repita o teste em (b) para as amostras de homens e mulheres separadamente.

e) Existe discriminação contra as mulheres? Construa seu teste de hipótese e apresente suas conclusões.

Resposta:

a) Aqui há poucas variáveis para um exercício caprichado, mas um diagnóstico comum é saber se a distribuição de covariadas, X, é a mesma entre os grupos de tratamento e controle (teste de balanceamento). No caso, a única informação disponível é o gênero do autor. Um possível teste de balanceamento seria calcular a razão homens/mulheres nas duas amostras e testar se são iguais.

b) Aqui, a nossa V.A. X para cada submissão é Bernoulli(p), de modo que

$$Y = \sum_{i=1}^{n} X_i \quad \text{é Binomial}(n, p)$$

Podemos testar na Binomial, mas como o n é grande, podemos aproximar em uma Normal:

$$\sum X_i \sim N(np, np(1-p)) \implies \bar{X} \sim N\left(p; \frac{p(1-p)}{n}\right)$$

Logo, para a diferença entre as médias temos

$$\bar{X}_A - \bar{X}_B \sim N(p_A - p_B; \sigma_A^2 + \sigma_B^2)$$

Vamos testar na amostra total a hipótese nula de que a média de aceitação entre as amostras cega e não-cega são iguais

$$H_0 : p_A = p_B \sim N(0; 1,1^2 + 1,4^2) = N(0; 3,17)$$

Logo,

$$z = \frac{p_A - p_B}{\sqrt{3,17}} = \frac{10,6 - 14,1}{\sqrt{3,17}} = \frac{3,5}{\sqrt{3,17}} = -1,966$$

Como $|z| = 1,966 > 1,96$, então rejeita-se H_0 a 5%.

c) p-valor $= 2.\Phi(-1,966) = 2.[1 - \Phi(1,966)] = 0,0493$

d) Mulheres:

$$z = \frac{10 - 11,2}{\sqrt{3^2 + 3,5^2}} = \frac{-1,2}{\sqrt{9 + 12,25}} = -0,26 \implies \text{não rejeita } H_0 \text{ a } 5\%$$

Homens:

$$z = \frac{11 - 15}{\sqrt{1,2^2 + 1,5^2}} = \frac{-4}{\sqrt{3,69}} = 2,08 \implies \text{rejeita } H_0 \text{ a } 5\%$$

e) Os resultados sugerem que apenas os homens melhoram sua taxa de aceitação ao ter seus nomes revelados. Um teste mais caprichado seria comparar os ganhos de ambos:

$$\Delta_h = +4\% \qquad \Delta_m = +1,2\% \qquad H_0 : \Delta_h = \Delta_m$$

Logo,

$$\Delta_h - \Delta_m \sim N(0; 3^2 + 3,5\% + 1,2^2 + 1,5^2) \approx N(0; 25)$$

Portanto,

$$z = \frac{4 - 1,2}{\sqrt{25}} = \frac{2,8}{5} = 0,56 \implies \text{não rejeita } H_0 \text{ a } 5\%$$

□

Exercício 14. *Suponha que $X \sim N(\mu, \sigma^2)$, e que você tenha acesso a uma amostra aleatória com N realizações de X. Considere as seguintes estatísticas:*

$$\bar{X} = \frac{1}{N}\sum_{i=1}^{N} X_i, \qquad e \qquad X^* = \frac{1}{N}\sum_{i=1}^{N}\left(X_i - \bar{X}\right)^2$$

a) Mostre que $\bar{X} \sim N(\mu, \sigma^2/N)$ e $\frac{NX^*}{\sigma^2} \sim \chi^2_{N-1}$

b) Mostre que $\bar{X}$ e X^* são independentes.

Resposta:

Resultados anteriores:

- Seja A matriz simétrica com N autovetores linearmente indep.:

- equação canônica:

$$Av = \lambda v : v \neq \vec{0} \implies AQ = Q\Lambda,$$

em que

$$\Lambda = \begin{bmatrix} \lambda_1 & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & \lambda_N \end{bmatrix} \quad \text{e} \quad Q = \begin{bmatrix} | & & | \\ v_1 & \dots & v_N \\ | & & | \end{bmatrix},$$

sendo Q uma matriz ortogonal tal que $v_i \perp\!\!\!\perp v_j, \forall i \neq j$. Logo,

$$A = Q\Lambda Q^{-1} \quad \text{e} \quad \Lambda = Q^{-1}AQ$$

- Se A é simétrica, então $Q^{-1} = Q^\intercal \implies M = Q\Lambda Q^\intercal$

- Para M idempotente e simétrica (ou seja, $M^\intercal = M$ e $MM = M$):
  - autovalores de M são sempre 0 e 1. Faça

  $$AAv = A\lambda v \implies Av\lambda Av = \lambda^2 v \implies \lambda v = \lambda^2 v.$$

  Como $v \neq 0$, então $\lambda = \lambda^2 \implies \lambda \in \{0, 1\}$
  - $rank(M) = m = N$, então

  $$\Lambda_{k \times k} = \begin{bmatrix} I_m & 0 \\ 0 & 0 \end{bmatrix}$$

  e, sem perda de generalidade, ordene os autovalores por magnitude.

- Se $Y_p \sim N(0, \sigma^2 I)$, então $\frac{Y'Y}{\sigma^2} \sim \chi^2_{(p)}$
  - considere $Q^\intercal y = v$, transformação linear de y tal que
    * $E(v) = 0$
    * $var(v) = Q\sigma^2 I Q^\intercal = \sigma^2 QQ^\intercal = \sigma^2 I$ (Q ortogonal $\implies QQ^\intercal = I$)
  - agora, tome $\frac{y'My}{\sigma^2} = \frac{v'QMQ'v}{\sigma^2} = \frac{1}{\sigma^2}v' \begin{bmatrix} I_m & 0 \\ 0 & 0 \end{bmatrix} = \frac{1}{\sigma^2}\sum_{i=1}^m v_i^2 \sim \chi^2_{(m)}$

a)
- $\bar{x}\frac{1}{N}\sum x_i \to$ Transformação linear de $\bar{x}$:

$$\left.\begin{array}{ll} E(\bar{x}) & = \frac{1}{N}\sum E(x_i) = \frac{N\mu}{N} = \mu \\ V(\bar{x}) & = \frac{1}{N^2}\sum V(x_i) = \frac{N\sigma^2}{N^2} = \frac{\sigma^2}{N} \end{array}\right\} \implies \bar{x} \sim N\left(\mu, \frac{\sigma^2}{N}\right)$$

- Defina $\varepsilon \equiv x - \mu \implies \varepsilon \sim N(0, 1)$ e $\hat{\varepsilon} \equiv x - \bar{x} \implies M\vec{x} = \hat{\varepsilon}$
- $x^* = \frac{1}{N}\hat{\varepsilon}'\hat{\varepsilon}$, como $\vec{\varepsilon} = \vec{x} - \mu$, então $M\vec{\varepsilon} = M\vec{x} - M\mu$. Note que $M\mu = 0$, logo

$$M\vec{\varepsilon} = M\vec{x}$$

- $x^* = \frac{1}{N}(M\varepsilon)'(M\varepsilon) = \frac{1}{N}\varepsilon'\mu'M\varepsilon = \frac{1}{N}\varepsilon'MM\varepsilon = \frac{1}{N}\varepsilon'M\varepsilon$

- Usando o resultado anterior:

$$N\frac{x^*}{\sigma^2} = \frac{\varepsilon'}{\sigma}M\frac{\varepsilon}{\sigma} \sim \chi^2_{(rankM)}$$

$$\begin{aligned} rank(M) &= Tr(M) = Tr(I - \tilde{M}) = Tr(I) - Tr(\tilde{M}) \\ &= N - Tr[\iota(\iota'\iota)^{-1}\iota'] = N - Tr[(\iota'\iota)^{-1}\iota'\iota] \\ &= N - Tr(1) = N - 1 \end{aligned}$$

b) Primeiro, lembre que, para $\vec{x}$ normal, temos que $cor(x_i, x_j) = 0$ implica independência. Sejam $x_1 = \bar{x} = \frac{1}{N}\sum x$ e $x_2 = \hat{\varepsilon}_i = x_i - \bar{x}$ para qualquer i. Segue que

$$\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} \sim N\left(\begin{bmatrix} \mu \\ 0 \end{bmatrix}; \begin{bmatrix} \frac{\sigma^2}{N} & 0 \\ 0 & \frac{N-1}{N}\sigma^2 \end{bmatrix}\right)$$

Logo,

$$x^* = \frac{\sum \hat{\varepsilon}^2}{N} = g(\hat{\varepsilon}_1, ..., \hat{\varepsilon}_N).$$

Se todo $\hat{\varepsilon} \perp\!\!\!\perp \bar{x}$, então

$$g(\hat{\varepsilon}_1, ..., \hat{\varepsilon}_N) \perp\!\!\!\perp \bar{x}.$$

□

Exercício 15. *Considere o modelo $y = bx + \varepsilon$. Suponha que $E(\varepsilon) = 0$ e $E(\varepsilon x) = 0$.*

a) Construa um estimador de método dos momentos para b que utilize $E(\varepsilon) = 0$.

b) Construa um estimador de método dos momentos para b que utilize $E(\varepsilon x) = 0$.

c) Se $\varepsilon|X \sim N(0, \sigma^2)$, qual dos dois estimadores é melhor? Justifique.

Resposta:

$$y = bx + \varepsilon \implies \varepsilon = y - bx$$

a)

$$E(\varepsilon) = 0 \implies \frac{1}{N}\sum y_i - \hat{b}x_i = 0 \implies \hat{b}_1 = \frac{\bar{y}}{\bar{x}}$$

b)

$$E(\varepsilon x) = 0 \implies \frac{1}{N}\sum (y_i - \hat{b}x_i)x_i = 0 \implies \hat{b}_2 = \frac{\sum y_i x_i}{\sum x_i^2}$$

c) Para $\hat{b}_1$:

$$\hat{b}_1 = \frac{1}{N}\sum\left(\frac{y_i}{\bar{x}}\right) = \frac{1}{N}\sum\frac{bx_i + \varepsilon_i}{\bar{x}} = b + \frac{\bar{\varepsilon}}{\bar{x}}$$

Se $\varepsilon|X \sim N(0,\sigma^2)$, então $E(\hat{b}_1) = b$ e

$$Var(\hat{b}_1) = Var\left(\frac{\bar{\varepsilon}}{\bar{x}}\right) = E\left[Var(\hat{b}_1|\bar{x})\right] = E\left[\frac{\sigma^2}{N\bar{x}^2}\right]$$

Para $\hat{b}_2$:

$$\hat{b}_2 = \frac{\sum y_i x_i}{\sum x_i^2} = \frac{\sum (bx_i + \varepsilon)x_i}{\sum x_i^2} = \frac{\sum bx_i^2}{\sum x_i^2} + \frac{\sum \varepsilon x_i}{\sum x_i^2} = b + \frac{\sum \varepsilon x_i}{\sum x_i^2}$$

Logo,

$$\begin{aligned}
E(\hat{b}_2) &= E\left(b + \frac{\sum \varepsilon_i x_i}{\sum x_i^2}\right) = b \\
Var(\hat{b}_2) &= Var\left(\frac{\sum \varepsilon_i x_i}{\sum x_i^2}\right) = E\left[Var\left(\frac{\sum \varepsilon_i x_i}{\sum x_i^2}\right)\bigg| x\right] \\
&= E\left[\frac{1}{\left(\sum x_i^2\right)^2}\sigma^2 \sum x_i^2\right] = E\left[\frac{\sigma^2}{\sum x_i^2}\right] \\
&= E\left[\frac{\sigma^2}{N\left(\widehat{var}(x) + \bar{x}^2\right)}\right] < E\left[\frac{\sigma^2}{N\bar{x}^2}\right] = Var(\hat{b}_1)
\end{aligned}$$

Portanto, ambos estimadores são não-viesados ($\hat{b}_1 = \hat{b}_2 = b$), mas $Var(\hat{b}_2) < Var(\hat{b}_1)$.

□

Exercício 16. *Sejam β_0^{MQO}, β_1^{MQO} o intercepto e inclinação de uma estimação por MQO do modelo $y = \beta_0 + \beta_0 x + \varepsilon$, em amostra de tamanho N. Para constantes não-nulas c_1, c_2, construa novas variáveis $\check{y} = c_1 y$ e $\check{x} = c_2 x$.*

a) Sejam $\widetilde{\beta}_0, \widetilde{\beta}_1$ o intercepto e inclinação de uma regressão por MQO de $\check{y}$ em $\check{x}$. Mostre que $\widetilde{\beta}_1 = \frac{c_1}{c_2}\ \beta_1^{MQO}$ e $\widetilde{\beta}_0 = c_1\ \beta_0^{MQO}$

b) Agora, sejam $\widetilde{\beta}_0, \widetilde{\beta}_1$ o intercepto e inclinação de uma regressão por MQO de $c_1 + y$ em $c_2 + x$. Mostre que $\widetilde{\beta}_1 = \beta_1^{MQO}$ e $\widetilde{\beta}_0 = \beta_0^{MQO} + c_1 - c_2\beta_1^{MQO}$

Resposta:

a)

$$\begin{aligned}
\tilde{\beta}_0 &= \bar{\check{y}} - \tilde{\beta}_1 \bar{\check{x}} \\
\tilde{\beta}_1 &= \frac{\widehat{cov}(\check{x}, \check{y})}{\widehat{var}(\check{x})} = \frac{\widehat{cov}(c_2 x, c_1 y)}{\widehat{var}(c_2 x)} = \frac{c_2 c_1}{c_2^2}\frac{\widehat{cov}(x,y)}{\widehat{var}(x)} = \frac{c_1}{c_2}\hat{\beta}_1^{MQO}
\end{aligned}$$

Substituindo $\tilde{\beta}_1$ em $\tilde{\beta}_0$, segue que

$$\tilde{\beta}_0 = \bar{\check{y}} - \tilde{\beta}_1 \bar{\check{x}} = c_1 \bar{y} - \frac{c_1}{c_2}\hat{\beta}_1^{MQO} c_2 \bar{x} = c_1\left[\bar{y} - \hat{\beta}_1^{MQO}\bar{x}\right] = c_1 \hat{\beta}_0^{MQO}$$

b)

$$\tilde{\beta}_1 = \frac{\widehat{cov}(\check{x}, \check{y})}{\widehat{var}(\check{x})} = \frac{\widehat{cov}(c_2 + x, c_1 + y)}{\widehat{var}(c_2 + x)} \overset{(a)}{=} \frac{\widehat{cov}(x, y)}{\widehat{var}(x)} = \hat{\beta}_1^{MQO}$$

(a): somar constantes não altera os segundos momentos

Substituindo $\tilde{\beta}_1$ em $\tilde{\beta}_0$, segue que

$$\begin{aligned}\tilde{\beta}_0 &= \bar{\check{y}} - \tilde{\beta}_1 \bar{\check{x}} = (\overline{c_1 + y}) - \tilde{\beta}_1(\overline{c_2 + x}) = (\overline{c_1 + y}) - \tilde{\beta}_1^{MQO}(\overline{c_2 + x}) \\ &= \left(\bar{y} - \hat{\beta}_1^{MQO}\bar{x}\right) + c_1 - c_2\hat{\beta}_1^{MQO} = \hat{\beta}_0^{MQO} + c_1 - c_2\hat{\beta}_1^{MQO}\end{aligned}$$

□