Giovana Fumer *Resolvido no caderno. Os exercícios selecionados a canto foram entregues em 27/09/2010 (2a feira)

MAE 5834 - Estatística Avançada I

Lista de Exercícios 2 – 2º semestre de 2010 – Prof. Silvia L.P. Ferrari

1. Seja $(X_1, \ldots, X_n)$ uma amostra aleatória de $X \sim \mathrm{N}(\theta, \sigma^2)$, onde θ e σ^2 são desconhecidos. Seja $\widehat{\sigma}^2$ um estimador não viesado de σ^2. Falso ou verdadeiro: $\mathrm{var}(\widehat{\sigma}^2) \geq 2\sigma^4/(n-1)$? Justifique.

2. **Proposição:** Sejam $\mathcal{P} = \{P_\theta,\ \theta \in \Omega\}$ uma família de distribuições e $\gamma = g(\theta) \in \mathcal{R}$ o estimando. Sempre uma das seguintes afirmações é verdadeira.

   i. Não existe nenhum estimador não viesado de γ.

   ii. Existe apenas um estimador não viesado de γ.

   iii. Existem infinitos estimadores não viesados de γ.

   Demos exemplos em sala de aula dos casos (i.) e (ii.) acima. Complete a demonstração.

3. Seja $X = (X_1, \ldots, X_n)$ uma amostra aleatória de uma distribuição $\mathrm{N}(\theta, 1)$, onde $\theta = 0$ ou 1. Considere estimadores de θ que assumam apenas os valores 0 e 1. Mostre que não existe nenhum estimador não viciado de θ.

4. Sejam $X_1, \ldots, X_n$ uma amostra aleatória de uma distribuição $\mathrm{U}(-\theta, \theta)$, $\theta > 0$.

   (a) Mostre que a família de distribuições $\mathcal{P} = \{\mathrm{U}(-\theta, \theta),\ \theta > 0\}$ é uma família de escala.

   (b) Mostre que $T = \max(|X_1|, \ldots, |X_n|)$ é uma estatística suficiente minimal para θ.

   (c) Encontre o estimador não viesado de variância uniformemente mínima de θ.

5. Prove ou dê contra-exemplo. Seja X variável aleatória com distribuição P_θ pertencente a uma família de distribuições $\mathcal{P} = \{P_\theta,\ \theta \in \Omega\}$ e sejam $g_1(\theta)$ e $g_2(\theta)$ funções de θ a valores reais. Sejam $\delta_1(X)$ e $\delta_2(X)$ estimadores de $g_1(\theta)$ e $g_2(\theta)$ respectivamente, com $\mathrm{E}(\delta_1(X)^2) < \infty$ e $\mathrm{E}(\delta_2(X)^2) < \infty$. Se estes estimadores são não viciados de variância uniformemente mínima (ENVVUM) então $a\delta_1(X) + b\delta_2(X)$ tem variância finita e é ENVVUM de $ag_1(\theta) + bg_2(\theta)$. Aqui a e b são números reais fixados.

6. Seja X uma variável aleatória que toma os valores $-1, 0, 1, 2, 3$ com probabilidades $P(X = -1) = 2pq$ e $P(X = k) = p^k q^{3-k}$, para $k = 0, 1, 2, 3$.

   (a) Verifique que esta é uma distribuição de probabilidade.

   (b) Obtenha o estimador não viciado de variância localmente mínima (ENVVLM) em p_0 de

   i. p

   ii. pq

   e verifique, para cada caso, se o estimador é ENVVUM.

7. Se T tem distribuição binomial $b(p, n)$ com $n > 3$, use o Método 1 para encontrar o ENVVUM de p^3.

8. Suponha que $X_1, \ldots, X_n$ sejam variáveis aleatórias independentes com distribuição $\mathrm{N}(\theta_i, 1)$, para $i = 1, \ldots, n$. Seja $\omega = n^{-1} \sum_{i=1}^n \theta_i^2$. Obtenha o ENVVUM e o EMV de ω e compare seus erros quadráticos médios.

9. Prove o Teorema 1.11, p. 88, TPE.

10. Seja $(X_1, \ldots, X_n)$ uma amostra aleatória de uma distribuição discreta com função de probabilidade $f_{\theta,j}$, em que $\theta > 0$ e $j = 1, 2$; $f_{\theta,1}$ é a distribuição de Poisson de média θ e $f_{\theta,2}$ é a distribuição geométrica de parâmetro $\theta/(1+\theta)$, isto é,

$$f_{\theta,2}(x) = \left[1 - \frac{\theta}{1+\theta}\right]\left[\frac{\theta}{1+\theta}\right]^x, \quad x = 0, 1, 2, \ldots.$$

Existe ENVVUM de θ? Justifique.

11. (a) Seja δ um estimador não viesado de $g(\theta)$ e considere perda quadrática. Seja $\delta^* = \delta + B$, em que B é uma variável aleatória independente de δ e que não é nula com probabilidade 1. Mostre que δ^* tem risco uniformemente maior que o de δ.

(b) **Definição:** Um estimador δ de $g(\theta)$ é dito ser *não viesado por risco* (*risk-unbiased*) se satisfaz

$$\mathrm{E}_\theta(L(\theta, \delta)) \leq \mathrm{E}_\theta(L(\theta', \delta)), \text{ para todo } \theta' \neq \theta.$$

Considere função de perda da forma $L(\theta, d) = \rho(d - \theta)$ e seja δ um estimador de $g(\theta)$ *não viesado por risco*. Mostre que δ^* definido acima tem risco uniformemente maior ou igual ao de δ.

12. Suponha que $X_1, \ldots, X_n$ tenham distribuição de Poisson $P(\lambda)$. Use o Método 1 para encontrar o ENVVUM de

(a) λ^k para qualquer inteiro k positivo;

(b) $e^{-\lambda}$.

13. Resolva o problema anterior (parte b) pelo Método 2 usando o fato de que um estimador não viciado de $e^{-\lambda}$ é: $\delta = 1$, se $X_1 = 0$, e $\delta = 0$, caso contrário.

14. Se $X_1, \ldots, X_n$ são iid com distribuição $\mathrm{N}(\zeta, \sigma^2)$, com σ conhecido, encontre o ENVVUM de ζ^2, ζ^3 e de ζ^4. Sugestão: Para calcular $\mathrm{E}(\overline{X}^k)$, escreva $\overline{X} = Y + \zeta$, onde $Y \sim \mathrm{N}(0, \sigma^2/n)$ e expanda $\mathrm{E}[(Y + \zeta)^k]$.

15. Resolva o problema anterior admitindo que σ é desconhecido.

16. Seja $(X_1, \ldots, X_n)$ uma amostra casual simples de $X \sim \mathrm{E}(a, b)$ com função densidade de probabilidade dada por

$$f(x; a, b) = b^{-1} \exp\{-(x - a)/b\} I_{(a,\infty)}(x)$$

para $a \in (-\infty, +\infty)$, $b > 0$. Assuma que $b = 1$.

(a) Mostre que $X_{(1)} = \min_{i \in \{1,\ldots,n\}} X_i$ é uma estatística suficiente e completa.

(b) Mostre que $X_{(1)}$ e S^2, a variância amostral, são variáveis aleatórias independentes.

17. Seja $(X_1, \ldots, X_n)$ uma amostra casual simples de $X \sim \mathrm{E}(a, b)$. Mostre que $X_{(1)} = \min_{i \in \{1,\ldots,n\}} X_i$ e $\sum [X_i - X_{(1)}]$ são independentes e têm distribuições $\mathrm{E}(a, b/n)$ e $b\mathrm{Gama}(n-1, 1)$. Sugestão: Se $a = 0$ e $b = 1$, as variáveis $Y_i = (n - i + 1)[X_{(i)} - X_{(i-1)}]$, para $i = 2, \ldots, n$, são iid e têm distribuição $E(0, 1)$.

18. Suponha que X_i, para $i = 1, \ldots, n$, sejam v.a's com distribuição $\mathrm{N}(\alpha + \beta t_i, \sigma^2)$, onde α, β e σ^2 são desconhecidos, e os t's são constantes conhecidas não todas nulas. Encontre o ENVVUM de α e β.

19. Sejam $X_1, \ldots, X_m$ e $Y_1, \ldots, Y_n$ variáveis aleatórias independentes com distribuições $U(0, \theta)$ e $U(0, \theta')$, respectivamente. Se $n > 1$, determine o ENVVUM de θ/θ'. ✓

20. Seja $(X_1, \ldots, X_n)$, $n \geq 2$, uma amostra aleatória de tamanho n de uma distribuição discreta $P_{\theta,j}$, em que $\theta \in (0, 1)$, $j = 1, 2$, sendo $P_{\theta,1}$ a distribuição de Poisson com média θ e $P_{\theta,2}$ a distribuição de Bernoulli de parâmetro θ.

   (a) Mostre que $T = \sum_{i=1}^{n} X_i$ não é uma estatística suficiente.

   (b) Encontre um ENVVUM de θ.

21. Seja $X \sim N(\mu, \sigma^2)$. Mostre que não existe estimador não viciado de σ^2 quando μ é desconhecido. OBS: $n=1$
(Sugestão: para $\sigma^2 = a$ fixado, mostre que X é uma estatística suficiente completa para μ).

22. Sejam $Z_1, \ldots, Z_n$ variáveis aleatórias i.i.d. com distribuição exponencial de média θ, i.e. com densidade $\theta^{-1} \exp(-z\theta^{-1})$, $z \geq 0$; $\theta > 0$. Sejam $x_1, \ldots, x_n$ constantes positivas. Suponha que $Y_1, \ldots, Y_n$ sejam uma amostra observável que segue o modelo

$$Y_i = \beta x_i + Z_i, \quad i = 1, \ldots, n,$$

em que $\beta > 0$, sendo β e θ desconhecidos. ✓

   (a) Encontre o estimador de máxima verossimilhança $(\widehat{\beta}, \widehat{\theta})$ de (β, θ).

   (b) $\widehat{\beta}$ e $\widehat{\theta}$ são estimadores não viciados de β e de θ respectivamente?

   (c) Obtenha a distribuição, a média e a variância do estimador de máxima verossimilhança de β.

23. Sejam $X_1, \ldots, X_n$ n observações independentes de uma distribuição exponencial $E(a, b)$, $a \in (-\infty, 0]$ e $b > 0$ é conhecido.

   (a) Mostre que $X_{(1)} = \min_{i \in \{1,\ldots,n\}} X_i$ é uma estatística suficiente, mas não é completa. ✓

   (b) Encontre o ENVVUM de a.
   Sugestão: Considere $g(X_{(1)}) = (cX_{(1)} + d) I_{(-\infty,0]}(X_{(1)})$, com c e d constantes. Mostre que $g(X_{(1)})$ é não correlacionado com os estimadores não viciados de zero.

24. Seja X uma variável aleatória com função de probabilidade

$$p_\theta(x) = P_\theta(X = x) = \frac{a(x)\theta^x}{C(\theta)}, \quad x = 0, 1, \ldots;$$

→ potência para dar 1

$\theta > 0$, $a(x) \geq 0$, $\sum_{x=0}^{\infty} a(x)\theta^x = C(\theta) < \infty$ (distribuição de série de potências).

   (a) Mostre que esta distribuição faz parte da família exponencial unidimensional e mostre que a função geradora de momentos de X é $M_X(u) = C(\theta e^u)/C(\theta)$.

   (b) Mostre que se $a(x) > 0$ para todo $x = 0, 1, \ldots$, então θ^r é estimável para qualquer r inteiro positivo e seu estimador não viciado de variância uniformemente mínima baseado em uma amostra $X_1, \ldots, X_n$ de observações independentes da distribuição de X é $\delta(T)$, sendo $T = \sum_{i=1}^{n} X_i$ e

$$\delta(t) = \begin{cases} 0, & t = 0, \ldots, r-1, \\ \frac{A(t-r,n)}{A(t,n)}, & t = r, r+1, \ldots, \end{cases}$$

   onde $A(t, n)$ é o coeficiente de θ^t na expansão em série de $[C(\theta)]^n$.

   Casella 3.6 p.105 TPE

   (c) Se $X_1, \ldots, X_n$ são observações independentes da distribuição de Poisson de média θ, $\theta > 0$, utilize o resultado em (b) para encontrar o ENVVUM de θ^r, r inteiro positivo.

(d) Mostre que a distribuição de Poisson truncada à direita em b ($b \geq 1$, inteiro) e parâmetro θ é um caso especial da distribuição de série de potências. (Se X tem distribuição de Poisson truncada à direita em b e parâmetro θ, então $\mathrm{P}_\theta(X = x) = \mathrm{P}_\theta(Y = x | Y \leq b)$, $x = 0, 1, \ldots, b$, onde Y tem distribuição de Poisson de parâmetro θ.

(e) Mostre que se X é uma única observação de uma distribuição de Poisson truncada à direita em b ($b \geq 1$, inteiro), então θ não é estimável. Esta afirmação contradiz à do item (b)? Por que?

* Ver Casella — parte de suficiência e completitude.

Lista 2 — Avançado I (1)

* Passar a limpo

Redução por suficiência

4ª Questão

$|X| \sim \frac{1}{\theta} I_{(0,\theta)}(x)$

$|X|_{(n)} \sim \frac{n!}{(n-1)!}\left(\frac{x}{\theta}\right)^{n-1} \cdot \frac{1}{\theta} \cdot I_{(0,\theta)}(x) = \frac{n}{\theta}\left(\frac{x}{\theta}\right)^{n-1} I_{(0,\theta)}(x)$

$E(g(|X|)) = 0 \Rightarrow \int_0^\theta g(|X|) \cdot \frac{n}{\theta}\left(\frac{x}{\theta}\right)^{n-1} dx = 0 \quad \forall \theta > 0$

$g(\theta) \cdot \frac{n}{\theta}\left(\frac{\theta}{\theta}\right)^{n-1} = 0$

$g(\theta) = 0 \ \forall \theta > 0$

$\boxed{g(T) \equiv 0}$

$|X|_{(n)}$ é completa

$X_1 \dots X_n$ a.a. de $U(-\theta,\theta)$, $\theta > 0$

(A) $\mathcal{P} = \{U(-\theta,\theta),\ \theta > 0\}$ é família de escala, pois considerando $Z \sim U(-1,1)$, qualquer distribuição em $\mathcal{P}$ é a distribuição de algum $X = bZ$, $b > 0$

$T = \left(\frac{n+1}{n}\right) \cdot |Y|_{(n)} = \left(1+\frac{1}{n}\right) \cdot |Y|_{(n)}$ ENVVUM de θ

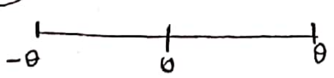
(B) $f(\underset{\sim}{x}|\theta) \overset{iid}{=} \prod_{i=1}^{n} I_{(-\theta,\theta)}(x_i) = \prod_{i=1}^{n} I_{[0,\theta)}(\underbrace{|x_i|}_{y_i}) = I_{[0,\theta)}(y_{(n)})$,

onde $y_{(n)} = \max_i |x_i|$. $E(|y|_{(n)}) = \int_0^\theta x \frac{n}{\theta} \cdot \left(\frac{x}{\theta}\right)^{n-1} dx = \frac{1}{\theta^n} \int_0^\theta x\, n (x)^{n-1} dx = \frac{n}{\theta^n} \int_0^\theta x^n = \frac{n}{\theta^n} \frac{x^{n+1}}{n+1}\Big|_0^\theta = \frac{n}{n+1}\theta$

Pelo Critério da Fatoração, $\max\{|x_i|, i=1 \dots n\}$ é suficiente para θ

$\max_i |x_i|$ também é estatística minimal, pois dadas duas amostras $\underset{\sim}{z}$ e $\underset{\sim}{w}$, temos que $\frac{f(\underset{\sim}{z}|\theta)}{f(\underset{\sim}{w}|\theta)}$ independe de θ se e somente se $\max_i |z_i| = \max_i |w_i|$. Consideramos aqui que a amostra $\underset{\sim}{w}$ não é composta somente por zeros para evitarmos indeterminação em $\frac{f(\underset{\sim}{z}|\theta)}{f(\underset{\sim}{w}|\theta)}$

(C) $\max\{|X_i|\} = |X|_{(n)}$



$y = |X| \sim U(0,\theta)$

$Y_{(n)}$ é suficiente e completa

$E(|y|_{(n)}) = \frac{(n+1)}{n+1}\theta$

$T = \frac{n+1}{n} \cdot |Y|_{(n)}$ é ENVVUM para θ.

5ª Questão

$\mathcal{P} = \{ P_\theta, \theta \in \Omega \}$

$E(\delta_1(x)) = g_1(\theta)$ e $E(\delta_1(x))^2 < \infty$, $\delta_1(x)$ ENVVUM de $g_1(\theta)$

$E(\delta_2(x)) = g_2(\theta)$ e $E(\delta_2(x))^2 < \infty$, $\delta_2(x)$ ENVVUM de $g_2(\theta)$

Portanto:

$E(a\delta_1(x) + b\delta_2(x)) = a g_1(\theta) + b g_2(\theta)$ (não viciado)

Para provar que é ENVVUM, basta agora provarmos que ele é não correlacionado com qualquer estimador $U(x)$ tal que $E_\theta(U(x)) = 0$ $\forall \theta \in \Omega$. Entretanto:

$$E\left[(a\delta_1(x) + b\delta_2(x)) U(x) \right] = a \underbrace{E\left[\delta_1(x) . U(x) \right]}_{\text{Não correlac. com } U(x)} + b \underbrace{E\left[\delta_2(x) U(x) \right]}_{\text{Não correlac. com } U(x)} \quad (1)$$

Como $\delta_1(x)$ e $\delta_2(x)$ são ENVVUM, o lado direito de (1) é igual a 0

Finalmente, a $Var(a\delta_1(x) + b\delta_2(x)) < \infty$, pois:

$$Var(a\delta_1(x) + b\delta_2(x)) = \underbrace{a^2 Var(\delta_1(x))}_{\text{finito, por hipótese}} + 2ab\, Cov(\delta_1(x), \delta_2(x)) +$$

$$+ \underbrace{b^2 Var(\delta_2(x))}_{\text{finito por hipótese}}.$$

Pela desigualdade da covariância: $0 \leq Cov^2(\delta_1(x), \delta_2(x)) \leq \underbrace{Var(\delta_1(x)) Var(\delta_2(x))}_{\text{finito}}$

Portanto: $Var(a\delta_1(x) + b\delta_2(x))$ é finita

Conclusão: $a\delta_1(x) + b\delta_2(x)$ é ENVVUM de $a g_1(\theta) + b g_2(\theta)$ e tem variância finita para $\underline{a}$ e $\underline{b}$ números reais fixados

8ª Questão

$X_i \sim N(\theta_i, 1)$, $i = 1 \dots n$, $\theta_i \in \mathbb{R}$

$W = \frac{1}{n} \sum_{i=1}^{n} \theta_i^2$

(A) A densidade conjunta é dada por:

$$f(\underset{\sim}{x}, \underset{\sim}{\theta}) \overset{indep}{=} \prod_{i=1}^{n} \left(\frac{1}{\sqrt{2\pi}}\right)^n \exp\left\{-\frac{(x_i-\theta_i)^2}{2}\right\} =$$

$$= \underbrace{\left(\frac{1}{\sqrt{2\pi}}\right)^n e^{-\frac{\sum_{i=1}^{n} x_i^2}{2}}}_{h(x)} e^{-\frac{-2\sum_{i=1}^{n} x_i\theta_i + \sum_{i=1}^{n} \theta_i^2}{2}} =$$

$$= h(x) \exp\left\{\sum_{i=1}^{n} x_i\theta_i - \frac{\sum_{i=1}^{n}\theta_i^2}{2}\right\} \Rightarrow$$ Família exponencial de posto completo sendo a amostra uma estatística suficiente e completa.

Como $1 = E_\theta X_i^2 - \theta_i^2 \Rightarrow \theta_i^2 = E_\theta X_i^2 - 1$, segue-se que:

$$\boxed{\delta_1(\underset{\sim}{X}) = \frac{1}{n}\sum_{i=1}^{n}(X_i^2 - 1) = \frac{\sum_{i=1}^{n} X_i^2}{n} - 1}$$ é não viciado para W.

Como também é função da estatística completa e suficiente é ENVVUM para W.

Quanto ao EMV de W, temos que o EMV de θ_i é X_i. Por invariância (ou por definição), o EMV de θ_i^2 é X_i^2 e o EMV de W é

$$\boxed{\delta_2(X) = \frac{1}{n}\sum_{i=1}^{n} X_i^2}$$

Questão 20

$(X_1 \dots X_n)$, $n \geq 2$, a.a. de $P_{\theta,j}$, $\theta \in (0,1)$ e $j = 1,2$, sendo

$$P_{\theta,1}(x|\theta) = \frac{\theta^x e^{-\theta}}{x!}, \ \underline{x = 0,1,2\dots} \quad e \quad P_{\theta,2}(x|\theta) = \theta^x (1-\theta)^{1-x}, \ \underline{x = 0,1}$$

Portanto: $\mathcal{P} = \{ P_\phi, \ \phi = (\theta, j), \ \theta \in (0,1), \ j \in \{1,2\} \}$

④ $$P(\underset{\sim}{x}|\theta,j) = \left\{ \prod_{i=1}^{n} \frac{\theta^{x_i} e^{-\theta}}{x_i!} I_{\{0,1,2\dots\}}(x_i) \right\}^{2-j} \times \left\{ \prod_{i=1}^{n} \theta^{x_i}(1-\theta)^{1-x_i} I_{\{0,1\}}(x_i) \right\}^{j-1} =$$

$$= \left\{ \frac{1}{\prod_{i=1}^{n} x_i!} \right\}^{2-j} \left\{ e^{-n\theta} \theta^{\sum_{i=1}^n x_i} \prod_{i=1}^{n} I_{\{0,1,2\dots\}}(x_i) \right\}^{2-j} \times \left\{ \theta^{\sum_{i=1}^n x_i} (1-\theta)^{n-\sum_{i=1}^n x_i} \prod_{i=1}^{n} I_{\{0,1\}}(x_i) \right\}^{j-1}$$

Portanto, as estatísticas conjuntamente suficientes são:

$$\left(\sum_{i=1}^n x_i, \ \prod_{i=1}^n I_{\{0,1,2\dots\}}(x_i), \ \prod_{i=1}^n I_{\{0,1\}}(x_i), \ \frac{1}{\prod_{i=1}^n x_i!} \right) = (T_1(x), T_2(x), T_3(x), T_4(x))$$

As duas distribuições (Poisson e Binomial), apesar de compartilharem a estatística $T_1(\underset{\sim}{x})$, possuem suportes diferentes e essa informação não está presente em $T_1(\underset{\sim}{x})$, mas em $T_2(\underset{\sim}{x})$ e $T_3(\underset{\sim}{x})$. Portanto, $\sum_{i=1}^n x_i$ não é suficiente para $\phi = (\theta, j)$, por não conter informação a respeito de j.

De fato, calculando a densidade condicional em $T(\underset{\sim}{x}) = \sum_{i=1}^n x_i = t$ e considerando que $T \sim$ Poisson de média $n\theta$ para $j=1$ e $T \sim$ Binomial (θ, n) para $j=2$, temos:

$$P_\phi(\underset{\sim}{x}|T=t) = \left(\frac{\dfrac{\cancel{e^{-n\theta}}}{\prod_{i=1}^n x_i!} \cancel{\theta^{\sum_{i=1}^n x_i}} \prod_{i=1}^n I_{\{0,1,2\dots\}}(x_i)}{\dfrac{\cancel{e^{-n\theta}}}{(\sum_{i=1}^n x_i)!} (n\cancel{\theta})^{\sum_{i=1}^n x_i} I_{\{0,1,2\dots\}}(\sum_{i=1}^n x_i)} \right)^{2-j} \times \left(\frac{\cancel{\theta^{\sum_{i=1}^n x_i}} \cancel{(1-\theta)}^{n-\sum x_i} \prod_{i=1}^n I_{\{0,1\}}(x_i)}{\binom{n}{\sum x_i} \cancel{(\theta)}^{\sum_{i=1}^n x_i} (1-\cancel{n\theta})^{n-\sum_{i=1}^n x_i} I_{\{0,1\dots n\}}(\sum_{i=1}^n x_i)} \right)^{j}$$

que depende de j, que é um dos parâmetros que, juntamente com θ, identifica os elementos de $\mathcal{P}$. Portanto, $\sum_{i=1}^{n} x_i$ não é suficiente para $\phi = (\theta, j)$

Na expressão (i) da folha anterior, supusemos $I_{\{0,1,2,\dots\}}\left(\sum_{i=1}^{n} x_i\right)$ e $I_{\{0,1,\dots,n\}}\left(\sum_{i=1}^{n} x_i\right)$ diferentes de 0.

(B) Tanto na distribuição binomial$^{(j=2)}$ quanto na Poisson$^{(j=1)}$, consideradas isoladamente, $T(x) = \sum_{i=1}^{n} x_i$ é estatística suficiente e completa(*) e $\frac{T(x)}{n} = \delta(x)$ é estimador não viciado de θ. Portanto, para $j=1$ e $j=2$, $\delta(x) = \frac{\sum_{i=1}^{n} x_i}{n}$ é ENVVUM para $\theta \Rightarrow \delta(x)$ é ENVVUM de θ em $\mathcal{P}$.

(*) Tal fato ocorre porque as duas distribuições, isoladamente, pertencem à uma família exponencial de posto completo.

Assim, $h(X_{(1)})$ não é identicamente nula na família das $E[a,b]$, $a \in (-\infty, 0]$ e $b > 0$ conhecido, mas possui esperança igual a 0 $\Rightarrow$ $X_{(1)}$ não é completa.

(B) Procuraremos por um estimador não viciado de a, função da estatística suficiente $X_{(1)}$ e que seja não correlacionado com estimadores de 0 na família $E[a,b]$ acima.

Seguindo a sugestão, seja $\delta(X_{(1)}) = (c X_{(1)} + d)\, I_{(-\infty,0]}(X_{(1)})$.

$$E[\delta(X_{(1)})] = \int_a^0 (cx + d) \underbrace{\frac{n}{b} e^{-\frac{n}{b}(x-a)}}_{\text{dens. de } X_{(1)}} dx =$$

$$= e^{\frac{na}{b}} \left\{ c \int_a^0 x \frac{n}{b} e^{-\frac{n}{b}x} dx + d \int_a^0 \frac{n}{b} e^{-\frac{n}{b}x} dx \right\} =$$

$$= e^{na/b} \left\{ c \left(a e^{-na/b} + \frac{b e^{-na/b}}{n} - \frac{b}{n} \right) + d \left(e^{-na/b} - 1 \right) \right\}$$

$$E[\delta(X_{(1)})] = a \iff c = 1 \text{ e } d = -\frac{b}{n}$$

$$\cancel{e^{\frac{na}{b}} \left\{ a e^{-\frac{na}{b}} + \frac{b e^{-\frac{na}{b}}}{n} - \frac{b}{n} - \frac{b}{n} e^{-\frac{na}{b}} + \right.}$$

Portanto: $\boxed{\delta(X_{(1)}) = \left(X_{(1)} - \frac{b}{n}\right) I_{(-\infty,0]}(X_{(1)})}$

$$E(X_{(1)}) = \int_a^{+\infty} x \cdot \frac{n}{b} e^{-\frac{n}{b}(x-a)} dx = e^{\frac{na}{b}} \int_a^{+\infty} x \frac{n}{b} e^{-\frac{nx}{b}} dx = e^{\frac{na}{b}} \left[x \cdot -e^{-\frac{nx}{b}} \Big|_a^{\infty} - \int_a^{\infty} -e^{-\frac{nx}{b}} dx \right] =$$

$u = x$, $du = dx$, $v = \int \frac{n}{b} e^{-\frac{nx}{b}} dx = -e^{-\frac{nx}{b}}$

$$= \underbrace{e^{\frac{na}{b}} \cdot a \cdot e^{-\frac{na}{b}} + \frac{b}{n} \cdot -e^{-\frac{nx}{b}} \Big|_a^{+\infty}}_{=(a + \frac{b}{n})} =$$

(10)

Para provarmos que $\delta(X_{(1)}) = \left(X_{(1)} - \frac{b\cancel{n}}{n\cancel{b}}\right) \times I_{(-\infty,0)}(X_{(1)})$ é ENVVUM de $\underline{a}$ basta provarmos que ele é não correlacionado com qualquer estimador não viciado de 0. Pela redução pela suficiência, basta procurarmos por estimadores que sejam função de $\delta(X_{(1)})$.

Assim, fazendo $X_{(1)} = y$:

$$0 = E_a[h(y)] = \int_a^{\infty} h(y)\, f(y|a)\, dy = \int_a^{0} h(y)\, f(y|a)\, dy + \int_0^{\infty} h(y)\, f(y|a)\, dy$$

$$\underbrace{\cancel{e^{\frac{na}{b}}} \int_0^{a} h(y)\, \frac{n}{b}\, e^{-\frac{n}{b}y}\, dy}_{\text{função de } a} = \underbrace{\cancel{e^{\frac{na}{b}}} \int_0^{\infty} h(y)\, \frac{n}{b}\, e^{-\frac{n}{b}y}\, dy}_{\text{constante em relação a } a}$$

Derivando em relação a "a":

$$h(a)\, \frac{n}{b}\, e^{-\frac{n}{b}a} = 0 \iff \boxed{h(a) = 0 \quad \forall\, a \in (-\infty, 0] \text{ e } \int_0^{\infty} h(y) f_a(y|a)\, dy = 0}$$

Portanto, os estimadores não viciados de 0 são funções de $X_{(1)}$ tais que são identicamente nulas $\forall X_{(1)} < 0$. (*) Entretanto, δ é tal que $\delta(X_{(1)}) = 0 \ \forall X_{(1)} > 0$. Portanto, $h(X_{(1)}) \times \delta(X_{(1)}) = 0$.

Assim: $E_a[h(X_{(1)}) \times \delta(X_{(1)})] = E_a[0] = 0 = E_a[h(X_{(1)})] \times E_a[\delta(X_{(1)})] = 0 \times a = 0$

Portanto, $\delta(X_{(1)})$ é ENVVUM de a.

(*) e ainda tais que $\int_0^{\infty} h(x)\, \frac{n}{b}\, e^{-\frac{n}{b}(x-a)}\, dx = 0$

C) Se $X \sim U(-\theta, \theta)$ então $Y = |X| \sim U(0, \theta)$ pois:

$F_Y(y) = P(Y \leq y) = P(|X| \leq y) = P(-y \leq X \leq y) = P(X \leq y) - P(X \leq -y) =$

$= \int_{-\theta}^{y} \frac{1}{2\theta} \, da - \int_{-\theta}^{-y} \frac{1}{2\theta} \, da = \frac{1}{2\theta} a \Big|_{-\theta}^{y} - \frac{1}{2\theta} a \Big|_{-\theta}^{-y} = \frac{1}{2\theta}(y+\theta) - \frac{1}{2\theta}(-y+\theta) = \frac{1}{2\theta}[y + \cancel{\theta} + y - \cancel{\theta}] = \frac{\cancel{2}y}{\cancel{2}\theta}$

$= \frac{y}{\theta}$

$$f_Y(y) = \frac{dF_Y(y)}{dy} = \frac{1}{\theta} I_{[0,\theta]}(y)$$

Se $Y = |X| \sim U(0, \theta)$; a distribuição do máximo dos módulos será dada por:

$Z = |X|_{(n)}$

$$f_Z(z) = \frac{n!}{(n-1)!} \left(\frac{z}{\theta}\right)^{n-1} \cdot \frac{1}{\theta} I_{[0,\theta]}(z) = \frac{n}{\theta} \left(\frac{z}{\theta}\right)^{n-1} I_{[0,\theta]}(z)$$

Além disso,

$$E(Z) = \int_0^\theta z \cdot \frac{n}{\theta} \left(\frac{z}{\theta}\right)^{n-1} dz = \int_0^\theta z \cdot \frac{n}{\theta} \cdot \frac{z^{n-1}}{\theta^{n-1}} dz = \frac{n}{\theta^n} \int_0^\theta z^n dz =$$

$$= \frac{n}{\theta^n} \cdot \frac{z^{n+1}}{n+1} \Big|_0^\theta = \frac{n}{\theta^n} \cdot \frac{\theta^{n+1}}{n+1} = \underbrace{\frac{n}{n+1}}_{\text{viciado}} \cdot \theta. \quad (*)$$

Pelo item (B) tem-se que a estatística $T = |X|_{(n)}$ é suficiente para θ.
Vamos verificar se essa estatística é completa. Dessa forma, tem-se que:

$$E_\theta(g(t)) = 0 \iff \int_0^\theta g(t) \cdot \frac{n}{\theta} \left(\frac{t}{\theta}\right)^{n-1} dt = 0 \quad \forall \theta > 0$$

$$\frac{d}{d\theta} \int_0^\theta g(t) \cdot \frac{n}{\theta} \left(\frac{t}{\theta}\right)^{n-1} dt = 0$$

$$g(\theta) \cdot \frac{n}{\theta} \cdot \left(\frac{\theta}{\theta}\right)^{n-1} = 0 \Rightarrow g(\theta) \cdot n \cdot \theta^{-1} = 0 \xRightarrow{n \cdot \theta^{-1} \neq 0} g(\theta) = 0, \ \forall \theta > 0 \Rightarrow \boxed{g(t) = 0} \quad (**)$$

$\therefore T = |X|_{(n)}$ é uma estatística completa para θ.

Portanto, um ENVVUM para θ será dado por $T^* = \frac{n+1}{n} \cdot |X|_{(n)}$ ou $T^* = \left(1 + \frac{1}{n}\right) |X|_{(n)}$; pois será um estimador não viesado para θ (ver (*)) e função de uma estatística suficiente e completa $T = |X|_{(n)}$ para θ (fato verificado pelo item (B) e por (**))

(14) a)

$\Omega_0 = [1, \infty)$

$P_\theta(\underline{x}) = \theta^{-n} \prod I_{(0,\theta)}(x_i) \qquad \theta \geq 1$

Supondo $x_1 \dots x_n$ t.q. $X_{(n)} < 1$ então

$L(\theta) = \theta^{-n} \, I_{[1,\infty)}(\theta)$ é max

qdo $\theta = 1$. Assim o EMV de θ é $\hat{\theta} = 1$

Suponha agora q' $x_1 \dots X_n$ a.a. t.q. $X_{(n)} \geq 1$.

então

$L(\theta) = \theta^{-n} \, I_{[X_{(n)},\infty)}(\theta)$ é máx

qdo $\theta = X_{(n)}$. Assim o EMV de θ é $\hat{\theta} = X_{(n)}$

EMV de θ é

$$\hat{\theta} = \begin{cases} 1 & X_{(n)} < 1 \\ X_{(n)} & X_{(n)} \geq 1 \end{cases}$$

15/09/2020

Lista de Exercícios 2

1. Seja $(X_1, \ldots, X_n)$ uma amostra aleatória de $X \sim N(\theta, \sigma^2)$, onde θ e σ^2 são desconhecidos. Seja $\hat{\sigma}^2$ um estimador não viesado de σ^2. Falso ou verdadeiro: $var(\hat{\sigma}^2) \geq \frac{2\sigma^4}{n-1}$? Justifique

$X \sim N(\theta, \sigma^2)$

$$f_x(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{-\frac{1}{2\sigma^2}(x-\theta)^2\right\}$$

$$p_x(x) = \prod_{i=1}^{n} f_{x_i}(x_i) = \prod_{i=1}^{n} \frac{1}{2\pi\sigma^2} \exp\left\{-\frac{1}{2\sigma^2}(x_i-\theta)^2\right\} =$$

$$= \exp\left\{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\theta)^2 - \frac{n}{2}\log 2\pi\sigma^2\right\} =$$

$$= \exp\left\{-\frac{1}{2\sigma^2}\Sigma(x_i^2 - 2x_i\theta + \theta^2) - \frac{n}{2}\log 2\pi\sigma^2\right\} =$$

$$= \exp\left\{-\frac{1}{2\sigma^2}\Sigma x_i^2 + \frac{\Sigma x_i\theta}{\sigma^2} - \frac{n\theta^2}{2\sigma^2} - \frac{n}{2}\log 2\pi\sigma^2\right\}$$

$$l(\underset{\sim}{u}) = -\frac{1}{2\sigma^2}\Sigma x_i^2 + \frac{\Sigma x_i\theta}{\sigma^2} - \frac{n\theta^2}{2\sigma^2} - \frac{n}{2}\log 2\pi\sigma^2$$

$\underset{\sim}{u} = (\theta, \sigma^2)$

$$\frac{\partial l(u)}{\partial\theta} = \frac{\Sigma x_i}{\sigma^2} - \frac{\cancel{2}n\theta}{\cancel{2}\sigma^2} = 0 \qquad \frac{\partial^2 l(u)}{\partial\theta^2} = -\frac{2n}{2\sigma^2} = -\frac{n}{\sigma^2} < 0$$

$$\Sigma x_i = n\theta$$

$$\theta = \Sigma x_i / n \implies \hat{\theta} = \bar{X}$$

$$\frac{\partial l}{\partial\sigma^2} = +\frac{1}{2\sigma^4}\Sigma x_i^2 - \frac{\Sigma x_i\theta}{\sigma^4} + \frac{n\theta^2}{2\sigma^4} - \frac{n}{2} \cdot \frac{1}{\cancel{2\pi}\sigma^2} \cdot \cancel{2\pi} = 0$$

$$\Sigma x_i^2 - 2\Sigma x_i\hat{\theta} + n\hat{\theta}^2 - n\hat{\sigma}^2 = 0$$

$$\Sigma x_i^2 - n2\frac{\Sigma x_i}{n}\frac{\Sigma x_i}{n} + n\bar{x}^2 = n\hat{\sigma}^2$$

$$n\sigma^2 = \Sigma x_i^2 - 2n\bar{x}^2 + n\bar{x}^2$$

$$n\sigma^2 = \Sigma x_i^2 - n\bar{x}^2 \implies \hat{\sigma}^2 = \frac{\Sigma X_i^2 - n\bar{X}^2}{n}$$

$$\sigma^2 = \frac{\Sigma x_i^2 - n\bar{x}^2}{n}$$

15/09/2010

$$E(\hat{\sigma}^2) = E\left(\frac{\sum X_i^2 - n\bar{X}^2}{n}\right) =$$

$$= \frac{1}{n}\left[E(\sum X_i^2) - n\,E(\bar{X}^2)\right] = \frac{1}{n}E(\sum X_i^2) - E(\bar{X}^2) =$$

(*) Cálculos auxiliares:

$X \sim N(\theta, \sigma^2)$

$Var(X) = E(X^2) - [E(X)]^2$

$\sigma^2 + \theta^2 = E X^2$

$\bar{X} \sim N(\theta, \frac{\sigma^2}{n})$

$Var\,\bar{X} = E(\bar{X}^2) - (E\bar{X})^2$

$\frac{\sigma^2}{n} = E(\bar{X}^2) - \theta^2$

$\Rightarrow E(\bar{X}^2) = \frac{\sigma^2}{n} + \theta^2$

$$= EX^2 - E(\bar{X}^2) = \sigma^2 + \cancel{\theta^2} - \frac{\sigma^2}{n} - \cancel{\theta^2} = \frac{n\sigma^2 - \sigma^2}{n} = \frac{(n-1)}{n}\sigma^2$$

$=$

$$E(\hat{\sigma}^2) = \frac{n-1}{n}\sigma^2$$

$$E\left(\frac{\hat{\sigma}^2 \cdot n}{(n-1)}\right) = \sigma^2$$

$$E\left(\frac{\frac{\sum X_i^2 - n\bar{X}^2}{\cancel{n}} \cdot \cancel{n}}{n-1}\right) = \sigma^2$$

$$E\left(\frac{\sum X_i^2 - n\bar{X}^2}{n-1}\right) = \sigma^2$$

$E(S^2) = \sigma^2$, em que $S^2 = \frac{\sum X_i^2 - n\bar{X}^2}{n-1}$ é um ENV para σ^2

$$E\left(\frac{\sum X_i^2 - n\bar{X}^2}{\sigma^2}\right) = n-1$$

$$E\left(\frac{\sum(X_i - \bar{X})^2}{\sigma^2}\right) = n-1 \quad e \quad \frac{\sum(X_i - \bar{X})^2}{\sigma^2} \sim \chi^2_{n-1}$$

$$Var\left(\frac{\sum(X_i - \bar{X})^2}{\sigma^2}\right) = 2(n-1)$$

15/09/2010

$$Var\left(\frac{\frac{\sum (X_i-\bar{X})^2}{n-1}}{\frac{\sigma^2}{n-1}}\right) = 2(n-1)$$

$$Var\left(\frac{\frac{S^2}{1}}{\frac{\sigma^2}{n-1}}\right) = 2(n-1)$$

$$Var\left(\frac{(n-1)\,S^2}{\sigma^2}\right) = 2(n-1)$$

$$\frac{(n-1)^{\cancel{2}}}{\sigma^4}\, Var\, S^2 = 2\cancel{(n-1)}$$

$$\boxed{Var\, S^2 = \frac{2\sigma^4}{n-1}}$$

S^2 é um estimador não viesado para $\hat{\sigma}^2$ e sua variância é exatamente igual a $\frac{2\sigma^4}{n-1}$.

Se S^2 é um estimador não-viciado de σ^2, a totalidade de estimadores não viciados é dada por:

$$\hat{\sigma}^2 = S^2 - U$$

onde U representa qualquer estimador não viciado do "zero", ou seja, U satisfaz $E_\theta(U) = 0$, para todo $\theta \in \Omega$.

$$R(\theta, \hat{\sigma}^2) = Var(\hat{\sigma}^2) = Var(S^2 - U) = E[(S^2-U)^2] - [E(S^2-U)]^2 =$$

$$= E[(S^2-U)^2] - [E(S^2)]^2 = E[(S^2-U)^2] - [\sigma^2]^2 = E[(S^2-U)^2] - \sigma^4 \quad (*)$$

A distribuição normal pertence a família exponencial de posto completo, e portanto, $T = (\bar{X}, S^2)$ são estatísticas suficientes e completas (e minimais) e então $\bar{X}$ é ENVVUM para θ e S^2 é ENVVUM para σ^2.

Deste modo, tem-se que:

$$Var(S^2 - U) = Var\, S^2 + Var(U) - 2\, Cov(S^2, U)$$

$$Cov(S^2, U) = E(S^2 U) - E(S^2).E(U) = 0 - 0 = 0$$

S^2 é ENVVUM

15/09/2010

E então:

$$V(S^2-U) = Var\, S^2 + Var\, U = \frac{2\sigma^4}{n-1} + Var(U)$$

$$[E(S^2-U)]^2 = [ES^2 - EU]^2 = [ES^2 - 0]^2 = [ES^2]^2 = (\sigma^2)^2 = \sigma^4$$

Voltando a (*) tem-se que:

$$Var(\hat{\sigma}^2) = E[(S^2-U)^2] - \sigma^4 = Var(S^2-U) + [E(S^2-U)]^2 - \sigma^4 =$$

$$= \frac{2\sigma^4}{n-1} + Var(U) \cancel{+\sigma^4 - \sigma^4} = \frac{2\sigma^4}{n-1} + Var(U)$$

Como $Var(U) \geq 0$, então

$$\boxed{Var(\hat{\sigma}^2) \geq \frac{2\sigma^4}{n-1}}$$

∴ a afirmação é verdadeira.

Outra forma: $Var(S^2) = \frac{2\sigma^4}{n-1}$

Se S^2 é o ENVVUM, então ele é único e qualquer outro estimador não viesado de $\hat{\sigma}^2$ terá uma variância maior do que a dele.

Portanto,

$$\boxed{Var(\hat{\sigma}^2) \geq \frac{2\sigma^4}{n-1}}$$

→ unbiased

U - estimável: encontra-se um estimador não viciado para o que você quer estimar

17/09/2010

→ ver outra resolução adiante (na próxima folha)

2) **Proposição:** Sejam $P = \{P_\theta,\ \theta \in \Omega\}$ uma família de distribuições e $\gamma = g(\theta) \in \mathbb{R}$ o estimando. Sempre uma das seguintes afirmações é verdadeira.

i) Não existe nenhum estimador não viesado de γ. Exemplo do estimador $\frac{1}{p}$ na bin.

ii) Existe apenas um estimador não viesado de γ. Caso dos estimadores não viciados de zero. → tem-se estatística suficiente e completa

iii) Existem infinitos estimadores não viesados de γ. Sempre tem-se uma estatística não completa

Demos exemplos em sala de aula dos casos (i) e (ii) acima.

Complete a demonstração

Em sala foram comentados os casos i) e ii), então para completar a demonstração, tem-se o iii), o qual retrata:

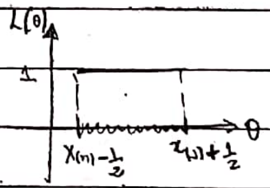
iii) Existem infinitos estimadores não viesados de γ.

Consideremos o caso no qual $X_1, \dots, X_n$ iid; $X_i \sim U(\theta - \frac{1}{2}, \theta + \frac{1}{2})$, $\theta \in \mathbb{R}$.

$$L(\theta, x) = \prod_{i=1}^{n} I_{[\theta-\frac{1}{2}, \theta+\frac{1}{2}]}(x_i) = \prod_{i=1}^{n} I_{[\theta-\frac{1}{2}, +\infty)}(x_i)\, I_{(-\infty, \theta+\frac{1}{2}]}(x_i) = I_{[\theta-\frac{1}{2}, +\infty)}(x_{(1)}) \cdot I_{(-\infty, \theta+\frac{1}{2}]}(x_{(n)}) =$$

$$\theta - \frac{1}{2} < x_{(1)} < +\infty \quad -\infty < \theta < x_{(1)} + \frac{1}{2}$$

$$-\infty < x_{(n)} < \theta + \frac{1}{2} \quad x_{(n)} - \frac{1}{2} < \theta < +\infty$$



$$= I_{(-\infty, x_{(1)}+\frac{1}{2})}(\theta) \cdot I_{(x_{(n)}-\frac{1}{2}, +\infty)}(\theta) \qquad \boxed{X_{(n)} - \frac{1}{2} \leq T(X) \leq X_{(1)} + \frac{1}{2}} \quad \text{EMV para } \theta.$$

Pelo exercício 15 da lista 1 tem-se que:

$$E(X_{(1)}) = +\theta - \frac{1}{2} + \frac{1}{n+1} \quad \text{e} \quad E(X_{(n)}) = \theta + \frac{1}{2} - \frac{1}{n+1}$$

Consideremos $T_1 = \dfrac{X_{(1)} + X_{(n)}}{2}$

$$E(T_1) = \frac{1}{2} E(X_{(1)} + X_{(n)}) = \frac{1}{2}\left[\theta - \cancel{\frac{1}{2}} + \cancel{\frac{1}{n+1}} + \theta + \cancel{\frac{1}{2}} - \cancel{\frac{1}{n+1}}\right] = \frac{2\theta}{2} = \theta$$

$\therefore T$ é não viesado para θ.

$$T_2 = X_{(1)} + \frac{1}{2} - \frac{1}{n+1}$$

$$E(T_2) = E(X_{(1)}) + \frac{1}{2} - \frac{1}{n+1} = \theta - \cancel{\frac{1}{2}} + \cancel{\frac{1}{n+1}} + \cancel{\frac{1}{2}} - \cancel{\frac{1}{n+1}} = \theta.$$

17/09/2010

Desse modo, podemos gerar infinitos estimadores não viesados de γ, $\gamma = g(\theta)$, $\theta \in \mathbb{R}$.

i) O caso de se estimar $\frac{1}{p}$ na $b(n,p)$:

ii) $E(\delta(X)) = g(\theta)$

$\delta'(X) = \delta(X) + U(X)$ (não viciado)

$E(U(X)) = 0$

Se $U(X)$ único $\Rightarrow$ $\delta(X)$ é único

X for estatística suficiente e completa

iii) Basta a função não ser completa $\Rightarrow$ $\delta(X)$ não será único.

01/10/2010

3. Seja $X = (X_1, \ldots, X_n)$ uma amostra aleatória de uma distribuição $N(\theta, 1)$, onde $\theta = 0$ ou 1. Considere estimadores de θ que assumam apenas os valores 0 e 1. Mostre que não existe nenhum estimador não viciado de θ.

Vamos supor, por absurdo, que exista um estimador não viciado de θ, então:

$$E_\theta(\delta(X)) = \theta \text{, em que}$$

$$\delta(X) = \begin{cases} 0 & \text{se } x \in B^c \\ 1 & \text{se } x \in B \end{cases}$$

Então

$$E(\delta(X)) = 0 \cdot P(X \in B^c) + 1 \cdot P(X \in B) = P(\delta(X) = 1) = \theta$$

$$\boxed{P(\delta(X) = 1) = \theta}$$

Como $\theta = 0$ ou $\theta = 1$, tem-se que:

$$\underbrace{P(\delta(X) = 1) = 0}\quad \text{ou} \quad \underbrace{P(\delta(X) = 1) = 1}_{\text{ABSURDO!}}$$

Se acerta o zero ele erra o um, se ele acerta o um, ele erra o zero.

Outro modo:

$$\delta(X) = \begin{cases} 0 & \text{se } x \in B^c \\ 1 & \text{se } x \in B \end{cases} ; \quad B: \{ \mathcal{J}_\theta;\ \theta = 1 \text{ ou } \theta = 0 \}$$

$$E(\delta(X)) = 0 \cdot P(X \in B^c) + 1 \cdot P(X \in B) = P(X \in B) = \begin{cases} P(\delta(X) = 1) \\ P(\delta(X) = 0) \end{cases}$$

$0 < P(\delta(X) = 1) < 1$ é viesado para θ pois $\theta = 1$ ou $\theta = 0$

$P(\delta(X) = 1) = 0 \Rightarrow$ é viesado para 1

$P(\delta(X) = 1) = 1 \Rightarrow$ é viesado para 0

$\therefore$ sempre é viesado

17/09/2010

4. Sejam $X_1, \dots, X_n$ uma amostra aleatória de uma distribuição $U(-\theta, \theta)$, $\theta > 0$.

(a) Mostre que a família de distribuições $P = \{U(-\theta, \theta), \theta > 0\}$ é uma família de escala.

$X \sim U(-\theta, \theta)$

$$f_X(x) = \frac{1}{\theta - (-\theta)} \cdot I_{(-\theta,\theta)}(x) = \frac{1}{2\theta} \cdot I_{(-\theta,\theta)}(x)$$

Seja $X = \theta U$, $U \sim U[-1; 1]$, $\theta > 0$

$u = \frac{x}{\theta}$

$$f_X(x) = \frac{1}{\theta} \cdot \frac{1}{2} I_{[-\theta,\theta]}(x) = \frac{1}{2\theta} I_{[-\theta,\theta]}(x)$$

$du = \frac{1}{\theta} dx$

$\frac{du}{dx} = \frac{1}{\theta}$

$$F_X(x) = P(X \leq x) = P(\theta U \leq x) = P\left(U \leq \frac{x}{\theta}\right) = \int_{-1}^{\frac{x}{\theta}} \frac{1}{2} \, da =$$

$$= \frac{1}{2} a \Big|_{-1}^{\frac{x}{\theta}} = \frac{1}{2}\left(\frac{x}{\theta} + 1\right)$$

$$f_X(x) = \frac{1}{2\theta} I_{(-\theta,0)}(x)$$

$$P(X \leq x) = P(\theta U \leq x) = P\left(U \leq \frac{x}{\theta}\right) = F_U\left(\frac{x}{\theta}\right)$$

$\therefore X$ pertence à uma família de escala.

17/09/2010

(b) Mostre que $T = \max(|X_1|, \dots, |X_n|)$ é uma estatística suficiente minimal para θ.

$-\theta \quad 0 \quad +\theta \qquad |x| < \theta$

$$f_X(x) = \frac{1}{2\theta} I_{(-\theta,\theta)}(x)$$

$$p_\theta(x) = \prod_{i=1}^{n} f_{X_i}(x_i) = \prod_{i=1}^{n} \frac{1}{2\theta} I_{(-\theta,\theta)}(x_i) = \boxed{\begin{array}{c} -\theta < x_i < \theta \\ -|x_i| < \theta \end{array}}$$

$$= \frac{1}{(2\theta)^n} \prod_{i=1}^{n} I_{[0,\theta)}(|x_i|) = \underbrace{\frac{1}{\theta^n} I_{[0,\theta)}(|X|_{(n)})}_{g_\theta(T(X))} \cdot \underbrace{\frac{1}{2^n}}_{h(x)}$$

Pelo Critério de Fatoração tem-se que $g_\theta(T(X)) = \frac{1}{\theta^n} I_{[0,\theta)}(|X|_{(n)})$ e $h(x) = \frac{1}{2^n}$; assim $T(X) = |X|_{(n)} = \max(|X_1|, |X_2|, \dots, |X_n|)$

é uma estatística suficiente para θ.

Tem-se ainda que para duas amostras distintas $X_1, \dots, X_n$ e $Y_1, \dots, Y_n$

$$\frac{p_\theta(x)}{p_\theta(y)} = \frac{\frac{1}{\theta^n} \cdot I_{[0,\theta)}(|x|_{(n)}) \cdot \frac{1}{2^n}}{\frac{1}{\theta^n} I_{[0,\theta)}(|y|_{(n)}) \frac{1}{2^n}}$$ seja independente de θ, se e somente se, (sendo que as amostras y não possui valores zero.)

$$|x|_{(n)} = |y|_{(n)} \iff T(X) = T(y)$$

$\therefore$ $T(X) = \max(|X_1|, |X_2|, \dots, |X_n|)$ é uma estatística suficiente e minimal para θ.

$\max|X_i| = |X|_{(n)}$ $\quad |X|_{(n)}$

(A) $-10 \quad -8 \quad 0 \quad 9 \quad 10$ $\quad X_{(n)} = 9 \quad |X|_{(n)} = 9$

(B) $-9 \quad 8$ $\quad X_{(n)} = 8 \quad |X|_{(n)} = 9$

25/09/2010

$\max_{i=1,\dots,n}\{|X_i|\} = |X|_{(n)}$

(c) Encontre o estimador não viesado de variância uniformemente mínima para θ.

Se $X_i \sim U(-\theta, \theta)$, X simétrica $\Rightarrow |X| \sim U(0,\theta)$, pois

$Y = |X|$

$F_Y(y) = P(Y \leq y) = P(|X| \leq y) = P(-y < X < y) =$

$= P(X<y) - P(X<-y) = \int_{-\theta}^{y} \frac{1}{2\theta}\,dx - \int_{-\theta}^{-y} \frac{1}{2\theta}\,dx =$

$= \frac{1}{2\theta}(y+\theta) - \left[\frac{1}{2\theta}(-y+\theta)\right] = \frac{1}{2\theta}\left[y + \cancel{\theta} + y - \cancel{\theta}\right] = \frac{2y}{2\theta} = \frac{1}{\theta}y$

$$f_Y(y) = \frac{1}{\theta} I_{(0,\theta)}(y)$$

E a distribuição do máximo dos módulos será dada por:

$$f_{|X|_n}(|x|_n) = \frac{n!}{(n-1)!}\left(\frac{x}{\theta}\right)^{n-1} \frac{1}{\theta} I_{(0,\theta)}(x) =$$

$$= \frac{n}{\theta}\left(\frac{x}{\theta}\right)^{n-1} I_{(0,\theta)}(x)$$

Vamos verificar se a estatística $T = \max(|X_1|, |X_2| \dots |X_n|)$ é completa.

$E_\theta(g(T)) = 0 \implies \int_0^\theta g(t)\,\frac{n}{\theta}\left(\frac{t}{\theta}\right)^{n-1} dt = 0 \quad \forall \theta > 0$

$$\frac{d}{d\theta}\int_0^\theta g(t)\,\frac{n}{\theta}\left(\frac{t}{\theta}\right)^{n-1} dt = 0$$

$$g(\theta)\,\frac{n}{\theta}\left(\frac{\theta}{\theta}\right)^{n-1} = 0$$

$\theta^{-1} n \neq 0$

$g(\theta)\, n\, \theta^{-1} = 0 \iff g(\theta) = 0, \quad \forall \theta > 0 \implies$

$\therefore T = |X|_{(n)}$ é $\implies \boxed{g(t) \equiv 0}$

$Z = |X|_n$ uma estatística suficiente e completa.

$E(Z_n) = \int_0^\theta z\,\frac{n}{\theta}\left(\frac{z}{\theta}\right)^{n-1} dz = \int_0^\theta z\,\frac{n}{\theta}\cdot\frac{z^{n-1}}{\theta^{n-1}}\,dz = \frac{n}{\theta^n}\int_0^\theta z^n\,dz =$

$= \frac{n}{\theta^n}\,\frac{z^{n+1}}{n+1}\Big|_0^\theta = \frac{n}{\theta^n}\,\frac{\theta^{n+1}}{n+1} = \frac{n}{n+1}\theta.$ Um ENVVUM para θ será dado por $T^* = \frac{n+1}{n}|X|_n = \left(1+\frac{1}{n}\right)|X|_{(n)}$, pois é ENV e função de uma estatística suficiente e completa $T = |X|_{(n)}$

17/09/2010

5. Prove ou dê contra-exemplo. Seja X variável aleatória com distribuição P_θ pertencente a uma família de distribuições $P = \{P_\theta, \theta \in \Omega\}$ e sejam $g_1(\theta)$ e $g_2(\theta)$ funções de θ a valores reais. Sejam $\delta_1(X)$ e $\delta_2(X)$ estimadores de $g_1(\theta)$ e $g_2(\theta)$ respectivamente, com $E(\delta_1(X)^2) < \infty$ e $E(\delta_2(X)^2) < \infty$. Se estes estimadores são não viciados de variância uniformemente mínima (ENVVUM) então $a\delta_1(X) + b\delta_2(X)$ tem variância finita e é ENVVUM de $ag_1(\theta) + bg_2(\theta)$. Aqui a e b são números reais fixados.

Hipóteses: $\delta_1(X)$ e $\delta_2(X)$ ENVVUM de $g_1(\theta)$ e $g_2(\theta)$, ou seja,

$E_\theta(\delta_1(X)) = g_1(\theta)$ $\quad$ $var_\theta(\delta_1(X)) < \infty$ $\quad$ $cov(\delta_1(X), U) = 0$.

$E_\theta(\delta_2(X)) = g_2(\theta)$ $\quad$ $var_\theta(\delta_2(X)) < \infty$ $\quad$ $cov(\delta_2(X), U) = 0$,

sendo U um estimador não viesado do zero.

Assim,

$E(a\delta_1(X) + b\delta_2(X)) = aE(\delta_1(X)) + bE(\delta_2(X)) = a\,g_1(\theta) + bg_2(\theta)$,

$\forall a, b \in \mathbb{R}$ fixados.

$\therefore$ $a\delta_1(X)$ e $b\delta_2(X)$ é ENV para $ag_1(\theta) + bg_2(\theta)$

$var(a\delta_1(X) + b\delta_2(X)) = var(a\delta_1(X)) + var(b\delta_2(X)) + 2\,cov(a\delta_1(X), b\delta_2(X)) =$

$= a^2\,var(\delta_1(X)) + b^2\,var(\delta_2(X)) + 2ab\,cov(\delta_1(X), \delta_2(X))$

Por hipótese, tem-se que:

$var(\delta_1(X)) < \infty$ e $var(\delta_2(X)) < \infty$

$\boxed{\rho^2 \leq 1 \quad \frac{cov^2(x,y)}{var\,x\,var\,y} \leq 1}$

Sabe-se ainda que:

$$cov(\delta_1(X), \delta_2(X)) \leq [\underbrace{var(\delta_1(X))}_{<\infty} . \underbrace{var(\delta_2(X))}_{<\infty}]^{\frac{1}{2}}$$

$cov(\delta_1(X), \delta_2(X)) < \infty$

$\therefore$ $var(a\delta_1(X) + b\delta_2(X)) < \infty$.

(b) Obtenha um estimador não viciado de variância localmente mínima (ENVVLM) em p_0 de i) p ii) pq e verifique, para cada caso, se o estimador é ENVVUM.

U é estimador não viciado de zero $\Leftrightarrow$

$$\sum_{k=-1}^{3} U(k)\; P(X=k)=0$$

$$U(-1)\,2pq+U(0)\,q^3+U(1)pq^2+U(2)\,p^2q+U(3)\,p^3=0$$

$$U(-1)\,2p(1-p)+U(0)(1-p)^3+U(1)\,p(1-p)^2+U(2)\,p^2(1-p)+U(3)\,p^3=0$$

$$U(-1)(2p-2p^2)+U(0)(1-3p+3p^2-p^3)+U(1)(p-2p^2+p^3)+U(2)(p^2-p^3)+U(3)p^3=0$$

$$\boxed{2pU(-1)}\boxed{-2p^2U(-1)}+U(0)-3pU(0)\boxed{+3p^2U(0)}\boxed{-p^3U(0)}+pU(1)-2p^2U(1)\boxed{+p^3U(1)}+$$
$$+p^2U(2)-p^3U(2)+p^3U(3)=0$$

$$\boxed{U(0)=0}$$

$$2U(-1)-3U(0)+U(1)=0 \qquad 2U(-1)+U(1)=0$$

$$-2U(-1)+3U(0)-2U(1)+U(2)=0 \qquad -2U(-1)-2U(1)+U(2)=0$$

$$-U(0)+U(1)-U(2)+U(3)=0 \qquad +U(1)-U(2)+U(3)=0$$

$$\boxed{U(1)=-2U(-1)}$$

$$2U(-1)-2[-2U(-1)]+U(2)=0$$

$$-2U(-1)+4U(-1)+U(2)=0 \Rightarrow \boxed{U(2)=-2U(-1)}$$

$$U(3)=-U(1)+U(2)$$

$$U(3)=+2U(-1)-2U(-1)$$

$$\boxed{U(3)=0}$$

$$U(0)=U(3)=0$$

$$U(1)=U(2)=-2U(-1)$$

Se tomarmos $U(-1)=a$, teremos que $U(K)$ tal que $E(U(K))=0$ tem a forma de:

$$U=U(k)=\begin{cases} a & \text{se } k=-1\\ 0 & \text{se } k=0 \text{ e } 3\\ -2a & \text{se } k=1 \text{ e } 2\end{cases}$$

ii) Um estimador não viesado para pq será dado por:

$$\delta_0=\begin{cases}\frac{1}{2} & \text{se } X=-1\\ 0 & \text{c.c.}\end{cases}$$

18/09/2010

$E_p(\delta_0) = \frac{1}{2} \cdot P(X=-1) = \frac{1}{\cancel{2}} \cdot \cancel{2}pq = pq$.

$\therefore \delta_0(X) = \frac{1}{2}$ é ENV para pq.

Vamos considerar perda quadrática: o risco (variância, pois é ENV) é minimizado em um valor fixo p_0, minimizando:

$$f(a) = \sum_{k=-1}^{3} P_{p_0}(X=k)\left[\delta_0(k) - u(k)\right]^2$$

$$f(a) = P(X=-1) \cdot \left[\tfrac{1}{2} - a\right]^2 + \cancel{P(X=0)\left[0-0\right]^2} + P(X=1)(0+2a)^2 + P(X=2) \cdot (0+2a)^2 + \cancel{P(X=3)(0-0)^2} =$$

$$f(a) = 2p_0q_0\left(\frac{1}{2} - a\right)^2 + p_0q_0^2\,4a^2 + p_0^2q_0\,4a^2$$

$$f(a) = 2p_0q_0\left(\frac{1}{4} - a + a^2\right) + p_0q_0^2\,4a^2 + p_0^2q_0\,4a^2$$

$$f(a) = \frac{p_0q_0}{2} - 2ap_0q_0 + 2a^2p_0q_0 + 4a^2p_0q_0^2 + 4a^2p_0^2q_0$$

$$f'(a) = -2p_0q_0 + 4ap_0q_0 + 8ap_0q_0^2 + 8ap_0^2q_0 \quad (*)$$

$$f''(a) = 4p_0q_0 + 8p_0q_0^2 + 8p_0^2q_0 > 0 \quad \text{ponto mínimo.}$$

Então $(*)$ $f'(a) = 0 \Leftrightarrow -2p_0q_0 + 4ap_0q_0 + 8ap_0q_0^2 + 8ap_0^2q_0 = 0$ $(\div p_0q_0)$

$$-2 + 4a + 8q_0a + 8ap_0 = 0$$

$$-2 + 4a + 8a(1-p_0) + 8ap_0 = 0$$

$$-2 + 4a + 8a - \cancel{8ap_0} + \cancel{8ap_0} = 0$$

$$-2 = -12a \Rightarrow \boxed{a = +\frac{1}{6}}$$

Note que $a = \frac{1}{6}$ independe de p_0 e assim o estimador

$\delta = \frac{1}{2}$ ... ENV ... para pq.

18/09/2010

23/09/2010

$$\delta(k) = \frac{1}{2} - U(k), \text{ onde } U(k) = \begin{cases} \frac{1}{6} & \text{se } k=-1 \\ 0 & \text{se } k = 0 \text{ e } 3 \\ -\frac{1}{3} & \text{se } k = 1 \text{ e } 2 \end{cases}$$ é o

estimador não viciado de variância localmente mínima para $p = p_0$. Tem-se ainda que $a = \frac{1}{6}$ minimiza $f(a)$ $\forall\ 0<p<1$ logo $\delta(k)$ é estimador não viciado de variância uniformemente mínima.

i) Para que se tenha um estimador não viciado para p, deve-se ter:

$$E_p(\delta(X)) = p$$

$$\delta(-1).P(X=-1) + \delta(0).P(X=0) + \delta(1).P(X=1) + \delta(2).P(X=2) + \delta(3).P(X=3) = p$$

$$\delta(-1).2pq + \delta(0).q^3 + \delta(1).pq^2 + \delta(2)p^2q + \delta(3)p^3 = p$$

$$\delta(-1).2p(1-p) + \delta(0).(1-p)^3 + \delta(1)p.(1-p)^2 + \delta(2)p^2(1-p) + \delta(3)p^3 = p$$

$$2p\,\delta(-1) - 2p^2\delta(-1) + \delta(0) - 3p\,\delta(0) + 3p^2\delta(0) - \delta(0)p^3 + p\,\delta(1) - 2p^2\delta(1) + p^3\delta(1) +$$
$$+ \delta(2)p^2 - \delta(2)p^3 + \delta(3)p^3 = p$$

$$\boxed{\delta(0) = 0.} \qquad 2\delta(-1) - 3\delta(0) + \delta(1) = 1$$

$$\boxed{\delta(1) = 1 - 2\delta(-1)}$$

$$-2\delta(-1) + \cancel{3\delta(0)} - 2\delta(1) + \delta(2) = 0$$

$$-2\delta(-1) - 2[1 - 2\delta(-1)] + \delta(2) = 0 \qquad -2(\delta(-1)) - 2 + 4\delta(-1) = -\delta(2)$$

$$-\cancel{\delta(0)} + \delta(1) - \delta(2) + \delta(3) = 0 \qquad \boxed{\delta(2) = +2 - 2\delta(-1)}$$

$$1 - 2\delta(-1) - (2 - 2\delta(-1)) + \delta(3) = 0$$

$$1 - \cancel{2\delta(-1)} - 2 + \cancel{2\delta(-1)} + \delta(3) = 0 \Rightarrow \boxed{\delta(3) = 1}$$

Se tomarmos $\delta(-1) = b$, teremos que:
$$\delta(k) = \begin{cases} b & \text{se } k=-1 \\ 0 & \text{se } k=0 \\ 1-2b & \text{se } k=1 \\ 2-2b & \text{se } k=2 \\ 1 & \text{se } k=3 \end{cases}$$
e este será um estimador não viesado para p.

Minimizando o risco considerando perda quadrática para p_0 fixo:

$$f(a) = \sum_{k=-1}^{3} P(X=k).[\delta(k) - U(k)]^2 = P(X=-1)(b-a)^2 + \cancel{P(X=0).(0-0)^2} + P(X=1).(1-2b+2a)^2 +$$
$$+ P(X=2).(2-2b+2a)^2 + P(X=3).(1-0)^2 = 2p_0q_0(b^2 - 2ab + a^2) + p_0q_0^2((1-2b)^2 + 2(1-2b)2a + 4a^2) +$$
$$+ p_0^2q_0((2-2b)^2 + 2(2-2b).2a + 4a^2) + p_0^3$$

$$f'(a) = -4p_0q_0b + 4p_0q_0a + 4(1-2b)p_0q_0^2 + 8a\,p_0q_0^2 + 4(2-2b)p_0^2q_0 + 8a\,p_0^2q_0$$

$$f''(a) = 4p_0q_0 + 8p_0q_0^2 + 8p_0^2q_0 > 0 \text{ ponto mínimo}$$

$$f'(a) = 0 \Rightarrow -4p_0q_0b + 4p_0q_0a + 4p_0q_0^2 - 8b\,p_0q_0^2 + 8a\,p_0q_0^2 + 8p_0^2q_0 - 8b\,p_0^2q_0 + 8a\,p_0^2q_0 = 0$$

$$-b + a + q_0 - 2bq_0 + 2aq_0 + 2p_0 - 2bp_0 + 2p_0 = 0$$

$$-b + a + 1 - p_0 - 2b(1-p_0) + 2a(1-p_0) + 2p_0 - 2bp_0 + 2p_0 = 0 \qquad (3 - 2p_0)a = -3p_0 + 3b - 1$$

$$-b + a + 1 - p_0 - 2b + \cancel{2bp_0} + 2a - 2ap_0 + 2p_0 - \cancel{2bp_0} + 2p_0 = 0 \qquad \boxed{a = \frac{-3p_0 + 3b - 1}{3 - 2p_0}}$$

Conclusão na próx. página.

Conclusão:

(*) $a = a_0 = -\dfrac{3p_0 + 3b - 1}{3 - 2p_0}$, a_0 depende de p_0 e o estimador $\delta = \delta_0 + a_0 X$ é não viciado e tem variância mínima em $p = p_0$.

Então este estimador será ENVVLM e não uniformemente / localmente. / /

7. Se T tem distribuição binomial $b(p,n)$ com $n > 3$, use o Método 1 para encontrar o ENVVUM de p^3.

Método I $\quad E_\theta(g(T)) = g(\theta)$

$g(p) = p^3$: $\quad \dfrac{T}{n}$ é ENV para p.

Então, considere o estimador T^3

$E_p(T^3) = E_p(T^3)$ (*)

$T \sim b(n,p)$

$M_T(t) = (1-p+pe^t)^n$

$M'_T(t) = n(1-p+pe^t)^{n-1} \cdot pe^t$

$M'_T(0) = ET = n(\cancel{1-p+p.e^0})^{n-1} \cdot p.e^0 = np$

$M''_T(t) = n(n-1)(1-p+pe^t)^{n-2} \cdot (pe^t)^2 + n(1-p+pe^t)^{n-1} \cdot pe^t$

$M''_T(0) = ET^2 = n(n-1)(1-\cancel{p}+\cancel{p}e^0)^{n-2} \cdot (p.e^0)^2 + n(1-\cancel{p}+\cancel{p.e^0})^{n-1} \cdot p.e^0 =$

$= n.(n-1)p^2 + np$

$var(T) = ET^2 - [ET]^2 = n(n-1)p^2 + np - n^2p^2 = (n^2-n)p^2 + np - n^2p^2 = \cancel{n^2p^2} - np^2 + np - \cancel{n^2p^2} =$

$= np(1-p)$.

$M''_T(t) = n(n-1)(1-p+pe^t)^{n-2} \cdot (pe^t)^2 + n(1-p+pe^t)^{n-1} \cdot pe^t$

$M'''_T(t) = n(n-1).(n-2).(1-p+pe^t)^{n-3} \cdot (pe^t)^3 + n(n-1)(1-p+pe^t)^{n-2} \cdot 2(p.e^t).(p.e^t) +$
$+ n(n-1)(1-p+pe^t)^{n-2} \cdot (pe^t)^2 + n(1-p+pe^t)^{n-1} \cdot pe^t$

$M'''_T(0) = ET^3 = n(n-1).(n-2).(1-p+p.e^0)^{n-3} \cdot (p.e^0)^3 + n(n-1)(1-p+pe^0)^{n-2} \cdot 2(p.e^0)^2 +$
$+ n(n-1)(1-p+p.e^0)^{n-2} \cdot (p.e^0)^2 + n(1-p+p.e^0)^{n-1} \cdot p.e^0 =$
$= n(n-1)(n-2)p^3 + n(n-1).2p^2 + n(n-1).p^2 + np =$
$= n.(n-1).(n-2)p^3 + 3n(n-1)p^2 + np$

$$\text{Var}\, T = E(T^2) - [ET]^2$$
$$np(1-p) = E(T^2) - (np)^2$$
$$E(T^2) = np - np^2 + n^2p^2$$

$$E_p(T^3) = n(n-1)(n-2)p^3 + 3n(n-1)p^2 + np$$

Seja $Q = \frac{1}{n(n-1)(n-2)}\left[T^3 - 3T^2 + 2T\right]$ um estimador para p^3.

$$E_p(Q) = E\left(\frac{1}{n(n-1)(n-2)}\left[T^3 - 3T^2 + 2T\right]\right) =$$

$$= \frac{1}{n(n-1)(n-2)}\left[E(T^3) - 3E(T^2) + 2ET\right] =$$

$$= \frac{1}{n(n-1)(n-2)}\left[n(n-1)(n-2)p^3 + 3n(n-1)p^2 + np - 3\left[np - np^2 + n^2p^2\right] + 2np\right] =$$

$$= \frac{1}{\cancel{n(n-1)(n-2)}}\left[\cancel{n(n-1)(n-2)}p^3 \cancel{+3n^2p^2} \cancel{-3np^2} \cancel{+np} \cancel{-3np} \cancel{+3np^2} \cancel{-3n^2p^2} \cancel{+2np}\right] = p^3$$

$$E(Q) = p^3$$

$\therefore Q = \frac{1}{n(n-1)(n-2)}\left[T^3 - 3T^2 + 2T\right]$ é um estimador não viciado de p^3

Como é uma função da estatística suficiente completa $\frac{T}{n}$, então Q é também um ENVVUM para p^3.

↳ esse fato se dá pela distribuição pertencer à família exponencial de posto completo.

17/09/2020

Outra forma:

A densidade é dada por:

$$P(T=t) = \binom{n}{t} p^t q^{n-t} = \qquad t=0,1,\dots,n \quad 0\leq p\leq 1,\ n\geq 3$$

$$= \exp\{\log p^t \cdot q^{n-t}\}\binom{n}{t} = \exp\{\log p^t + \log q^{n-t}\}\binom{n}{t} =$$

$$= \exp\{t\log p + (n-t)\log q\}\binom{n}{t} = \exp\{t\log p + n\log q - t\log q\}\binom{n}{t} =$$

$$= \exp\{t\log \frac{p}{q} + n\log q\}\binom{n}{t} \qquad \eta = \log\frac{p}{q} \qquad b(\eta) = -n\log q \qquad c(t) = \binom{n}{t}$$

Como visto, a distribuição pertence a família exponencial e o domínio da função de η contém retângulos $K=1$ dimensionais, desse modo, a família é de posto completo e a estatística $T(t)$, é uma estatística suficiente e completa (e minimal)

Método 1: Se T é uma estatística suficiente e completa, o ENVVUM de ∀ função estimável $g(p)$ é unicamente determinado por $E_p[\delta(T)] = g(p)$

Como p^3 é função de p, seja $g(p)=p^3$, então tem-se

$$E_p(\delta(T)) = g(p) = p^3 \qquad \forall \quad 0\leq p\leq 1$$

$$\sum_{t=0}^{n} \delta(t)\cdot\binom{n}{t} p^t q^{n-t} = p^3 \qquad (*)$$

Se $\rho = \frac{p}{q}$, então $\rho = \frac{p}{1-p} \Rightarrow \rho - \rho p = p$

$$\rho = p + \rho p \Rightarrow \boxed{p = \frac{\rho}{\rho+1}} \qquad q = 1-p = 1-\frac{\rho}{\rho+1} \quad \boxed{q = \frac{1}{\rho+1}}$$

Substituindo em $(*)$ tem-se que:

$$\sum_{t=0}^{n} \delta(t)\binom{n}{t}\left(\frac{\rho}{\rho+1}\right)^t\left(\frac{1}{\rho+1}\right)^{n-t} = \left(\frac{\rho}{\rho+1}\right)^3$$

17/09/2010

$$\sum_{t=0}^{n} \delta(t)\binom{n}{t} p^t \left(\frac{1}{1+p}\right)^n = \left(\frac{p}{1+p}\right)^3$$

$$\sum_{t=0}^{n} \delta(t)\binom{n}{t} p^t = \left(\frac{p}{1+p}\right)^3 \cdot (1+p)^n$$

$$\sum_{t=0}^{n} \delta(t)\binom{n}{t} p^t = p^3 (1+p)^{n-3} \qquad (**)$$

<u>Binômio de Newton</u>: $(a+b)^n = \sum_{j=0}^{n} \binom{n}{j} a^j b^{n-j}$

Então:

$$(1+p)^{n-3} = \sum_{j=0}^{n-3} \binom{n-3}{j} 1^j p^{n-3-j}$$

$$p^3 (1-p)^{n-3} = \sum_{j=0}^{n-3} \binom{n-3}{j} p^{n-j} \qquad (\cdot)$$

Voltando (**) e substituindo $(\cdot)$, tem-se:

$$\sum_{t=0}^{n} \delta(t)\binom{n}{t} p^t = \sum_{j=0}^{n-3} \binom{n-3}{j} p^{n-j}$$

ou seja,

$$\delta(0)\binom{n}{0}p^0 + \delta(1)\binom{n}{1}p^1 + \dots + \delta(n)\binom{n}{n}p^n = \binom{n-3}{0}p^{n-0} + \binom{n-3}{1}p^{n-1} + \dots + \binom{n-3}{n-3}p^{\cancel{n}-\cancel{n}+3}$$

$$\delta(0)\binom{n}{0}p^0 + \delta(1)\binom{n}{1}p^1 + \delta(2)\binom{n}{2}p^2 = 0$$

$$\delta(3)\binom{n}{3} = \binom{n-3}{n-3} \qquad \delta(4)\binom{n}{4} = \binom{n-3}{n-4}, \dots \quad \delta(n)\binom{n}{n} = \binom{n-3}{0}$$

$$\delta(t) = \begin{cases} 0 & \text{se } t = 0,1,2 \\ \dfrac{\binom{n-3}{n-t}}{\binom{n}{t}} & \text{se } t = 3,4,\dots,n \end{cases}$$

17/09/2010

Portanto, se $T \sim bin(p, n)$, $n > 3$, o ENVVUM para p^3 é:

$$\delta(T) = \frac{\binom{n-3}{n-T}}{\binom{n}{T}} \quad \text{se } T = 3, 4, \ldots n$$

Note que: $$\delta(T) = \frac{\binom{n-3}{n-T}}{\binom{n}{T}} = \frac{\frac{(n-3)!}{(n-T)!(\cancel{n}-3-\cancel{n}+T)!}}{\frac{n!}{\cancel{(n-T)!}\,T!}} = \frac{(n-3)!\ T!}{n!\ (T-3)!} =$$

$$= \frac{\cancel{(n-3)!}\ T.(T-1).(T-2).\cancel{(T-3)!}}{n.(n-1).(n-2).\cancel{(n-3)!}\ \cancel{(T-3)!}} = \frac{(T^2-T)\ (T-2)}{n(n-1).(n-2)} = \frac{T^3 - 2T^2 - T^2 + 2T}{n(n-1)(n-2)}$$

$$\delta(T) = \frac{1}{n(n-1)(n-2)} \left[T^3 - 3T^2 + 2T\right]$$

que coincide com o valor encontrado pelo exercício feito de outro modo.

8. Suponha que $X_1, \dots, X_n$ sejam variáveis aleatórias independentes com distribuição $N(\theta_i, 1)$, para $i = 1, \dots, n$. Seja $w = n^{-1} \sum_{i=1}^{n} \theta_i^2$. Obtenha o ENVVUM e o EMV de w e compare seus erros quadráticos médios

$$f_{X_i}(x_i) = \frac{1}{\sqrt{2\pi}} \exp\left\{ -\frac{1}{2}(x_i - \theta_i)^2 \right\} I_{\mathbb{R}}(x_i)$$

$$p_X(x) = \prod_{i=1}^{n} f_{X_i}(x_i) = \frac{1}{(2\pi)^{\frac{n}{2}}} \exp\left\{ -\frac{1}{2} \sum_{i=1}^{n} (x_i - \theta_i)^2 \right\} =$$

$$p_X(x) = \exp\left\{ -\frac{1}{2} \sum_{i=1}^{n} (x_i^2 - 2x_i\theta_i + \theta_i^2) - \frac{n}{2} \log 2\pi \right\}$$

A estatística suficiente neste caso é a amostra $\sum X_i \theta_i$

$$p_X(x) = \exp\left\{ -\frac{\sum_{i=1}^{n} x_i^2}{2} + \sum_{i=1}^{n} x_i\theta_i - \frac{\sum_{i=1}^{n} \theta_i^2}{2} - \frac{n}{2} \log 2\pi \right\}$$

A cada X_i você tem uma amostra completa.

$$l(\theta_i) = -\frac{\sum x_i^2}{2} + \sum x_i \theta_i - \frac{\sum \theta_i^2}{2} - \frac{n}{2} \log 2\pi$$

$$\frac{dl(\theta)}{d\theta_i} = x_i - \theta_i \implies \hat{\theta}_i = X_i$$

$$\frac{d^2 l_i(\theta)}{d\theta_i^2} = -1 < 0 \text{ ponto máximo}$$

Pelo Princípio da invariância dos EMV, tem-se que:

$$\hat{\theta}_i = X_i$$

$$\hat{w} = \frac{1}{n} \sum_{i=1}^{n} \hat{\theta}_i^2 = \frac{1}{n} \sum_{i=1}^{n} X_i^2$$

$$E(\hat{w}) = E\left(\frac{1}{n} \sum_{i=1}^{n} X_i^2 \right) = \frac{1}{n} E\left(\sum X_i^2 \right) = \frac{\sum \theta_i^2 + n}{n} = \frac{\sum \theta_i^2}{n} + 1 = w + 1$$

$\hat{w}_1 = \hat{w} - 1$ é um ENV de $g(\theta) = \frac{1}{n} \sum_{i=1}^{n} \theta_i^2$. A distribuição normal pertence à família exponencial de posto completo, então a estatística $T = X_i$ é suficiente completa e minimal para θ_i. Como $\hat{w}_1$ é uma função da estatística suficiente e completa de uma $g(\theta)$,

$\hat{W}_1$ é um ENVVUM de $g(\theta)$

Teorema:

Sejam $X_i \sim N(\mu_i, \sigma^2)$ $i=1,2,\dots,n$, v.a. independentes. Nessas condições:

$$U = \sum_{i=1}^{n} \frac{X_i^2}{\sigma^2} \sim \chi^2_{n,\lambda}$$

onde $\lambda = \frac{\sum_{i=1}^{n} \mu_i^2}{2\sigma^2}$

$E(U) = n + 2\lambda$

$Var(U) = 2(n+4\lambda)$

$X_i \sim N(\mu_i, 1)$

$$Y = \sum X_i^2 \sim \chi^2_{n,\lambda}, \quad \lambda = \frac{1}{2}\sum_{i=1}^{2} \mu_i^2$$

$E(Y) = n + 2\lambda$

$Var(Y) = 2(n+4\lambda)$

Então:

$$E(\hat{W}) = E\left(\frac{\sum X_i^2}{n}\right) = \frac{1}{n} E(\sum X_i^2) = \frac{1}{n}\cdot\left(n + 2\cdot\frac{1}{2}\sum_{i=1}^{n}\theta_i^2\right) = 1 + \frac{\sum_{i=1}^{n}\theta_i^2}{n} = 1 + W$$

$$Var(\hat{W}) = Var\left(\frac{\sum X_i^2}{n}\right) = \frac{1}{n^2} Var(\sum X_i^2) = \frac{1}{n^2} 2\cdot\left(n + 4\cdot\frac{1}{2}\sum_{i=1}^{n}\theta_i^2\right) =$$

$$= \frac{1}{n^2} 2\left(n + 2\sum_{i=1}^{n}\theta_i^2\right) = \frac{2\cancel{n}}{n^{\cancel{2}}} + \frac{4}{n^2}\sum_{i=1}^{n}\theta_i^2 = \frac{2}{n} + \frac{4}{n^2}\sum_{i=1}^{n}\theta_i^2$$

$$Viés(\hat{W}) = E(\hat{W}) - g(\theta) = \frac{\sum\theta_i^2}{n} + 1 - \frac{\sum\theta_i^2}{n} = 1$$

$$Var(\hat{W}_1) = Var(\hat{W} - 1) = Var(\hat{W}) = \frac{2}{n} + \frac{4}{n^2}\sum_{i=1}^{n}\theta_i^2$$

Comparando os erros quadráticos médios, tem-se que:

$$\hat{\Psi} = \frac{\sum X_i^2}{n} \quad e \quad \hat{\Psi}_1 = \frac{\sum X_i^2 - 1}{n}$$

$$EQM(\hat{\Psi}) = Var(\hat{\Psi}) + Viés^2(\hat{\Psi}) = \frac{2}{n} + \frac{4}{n^2}\sum_{i=1}^{n}\theta_i^2 + 1$$

$$EQM(\hat{\Psi}_1) = Var(\hat{\Psi}_1) + \underbrace{\cancel{Viés^2(\Psi_1)}}_{ENV} \xrightarrow{0} = Var(\Psi_1) = \frac{2}{n} + \frac{4}{n^2}\sum_{i=1}^{n}\theta_i^2$$

$$\therefore EQM(\hat{\Psi}_1) < EQM(\hat{\Psi})$$

Pela lógica, poderíamos considerar que Ψ_1 seria o melhor.

apenas cuidado com conclusões gerais!

18/09/2010

estimador da função $g(\theta) = \frac{\sum \theta_i^2}{n}$ porque ele é não viesado e apresenta um erro quadrático médio menor quando comparado a outro estimador, porém, ele pode tomar valores negativos sempre que $\frac{\sum X_i^2}{n} < 1$, o que não é interessante para se estimar $g(\theta)$. Então, o estimador viesado alternativo $\hat{\psi} = \frac{\sum X_i^2}{n}$, que será sempre positivo, seria mais adequado para se estimar $g(\theta)$.

* Discussão sobre erro: depende do objetivo)

9. Prove o Teorema 1.11, p. 88, TPE.

Seja X distribuído de acordo com a família de distribuições em $\mathcal{P} = \{P_\theta, \theta \in \Omega\}$, e suponha que T é uma estatística suficiente e completa para $\mathcal{P}$.

a) Para toda função estimável (não viesado) $g(\theta)$, existe um estimador não viesado que minimiza uniformemente o risco para alguma função de perda $L(\theta, d)$ a qual é convexa no seu segundo argumento, portanto, este estimador é ENVVUM.

b) O ENVVUM de (a) é o único estimador não viesado o qual é uma função de T, que tem o único estimador não viesado com risco mínimo, e seu risco é finito e L é estritamente convexa em d.

* A prova a seguir é feita considerando $L(\theta, d)$ como a função de perda quadrática.

Seja $\delta(T)$ uma função de uma estatística suficiente e completa T, tal que $E(\delta(T)) = g(\theta)$, $\forall \theta \in \Omega$.

Seja $\phi(T)$ uma outra função de T com $E(\phi(T)) = g(\theta)$

Então,

$$E[\delta(T) - \phi(T)] = E(\delta(T)) - E(\phi(T)) = g(\theta) - g(\theta) = 0$$

Então, se $\delta(T) - \phi(T) = g(T)$, tem-se que:

$$E(\delta(T) - \phi(T)) = E(g(T)) = 0 \implies$$

se a estatística é completa, tem-se que:

$$g(T) = 0 \text{ q.c} \Rightarrow \delta(T) = \phi(T) \text{ q.c}$$ e então tem-se que a função estimável que minima o estimando $g(\theta)$ é única. (*)

Além disso, vamos considerar W outro estimador não viciado para $g(\theta)$. $[E_\theta(W) = g(\theta)]$

Se W não é função apenas de T:

$$\phi(T) = E(W|T)$$

$$E\phi(T) = E(E(W|T)) = EW = g(\theta)$$

$\therefore \phi(T)$ é não viciado para $g(\theta)$ e $Var(\phi(T)) \leq Var(W)$ (por Rao-Blackwell.)

Mas por (*), $\phi(T) = \delta(T)$ q.c $\Rightarrow Var(\delta(T)) \leq Var(W)$ p/ $\forall \theta \in \Omega$.

18/09/2010

Refazendo a prova, utilizando uma função de perda convexa qualquer:

Seja T uma estatística suficiente e completa para θ. Seja $E(\gamma(T)) = g(\theta)$, $\forall \theta \in \Omega$, uma função $\gamma(T)$ estimável do estimando $g(\theta)$.

Tomemos $\phi(T)$ uma outra função de T com

$E(\phi(T)) = g(\theta)$

Então

$$E[\gamma(T) - \phi(T)] = E(\gamma(T)) - E(\phi(T)) = g(\theta) - g(\theta) = 0$$

Se fizermos $g(T) = \gamma(T) - \phi(T)$

Tem-se que:

$$E(g(T)) = 0 \xrightarrow{T \text{ é completa}} g(T) = 0 \text{ q.c. em } P$$

$$\Updownarrow$$

$$\gamma(T) - \phi(T) = 0 \quad \text{q.c. em } P$$

$$\Updownarrow$$

$$\gamma(T) = \phi(T) \quad \text{q.c. em } P \quad (*)$$

$\therefore$ Se existe uma função estimável $\gamma(T)$ do estimando $g(\theta)$, a qual é função de uma estatística T suficiente e completa, então essa função deve ser única.

Suponha W um outro estimador não viciado de $g(\theta)$, isto é: $E_\theta(W) = g(\theta)$, $\forall \theta \in \Omega$.

Suponha que W não é função apenas de T (estatística suficiente em P)

Se W é um estimador de um estimando $g(\theta)$ e se $L(\theta, d)$ é uma função de perda estritamente convexa de d. Então, se W tem média finita e risco finito, então seja:

$\phi(T) = E(W|T)$

$\therefore E(\phi(T)) = E(E(W|T)) = E(W) = g(\theta)$, então por Rao ...

/ /

Blackwell, o risco do estimador $\phi(T)$ satisfaz

$$R(\theta, \phi) < R(\theta, W)$$

a menos que $W(X) = \phi(T)$ com probabilidade 1.

Mas por (*) tem-se que $\gamma(T) = \phi(T)$; então:

$$R(\theta, \gamma) < R(\theta, W) \quad \text{para } \forall \theta \in \Omega.$$

L convexa (1)

estritamente convexa (2)

T suficiente — completa — unicidade

$E_\theta(\delta(X)) = g(\theta)$

$\delta_2 = \delta' + U, \quad E_\theta(U) = 0$

$\delta_2(T) = \delta(T) + U(T) \to 0$

redução por suficiência

$E_\theta(\delta'(X) \mid T) = \delta'(T)$ não viciado

$R(\delta') \leq R(\delta)$ (1)

$<$ (2)

$$E(X) = \frac{q}{p} = \frac{\frac{\theta}{1+\theta}}{1-\frac{\theta}{1+\theta}} = \frac{\frac{\theta}{\cancel{1+\theta}}}{\frac{1+\cancel{\theta}-\cancel{\theta}}{\cancel{1+\theta}}} = \theta$$

/ /

10. Seja $(X_1, \ldots, X_n)$ uma amostra aleatória de uma distribuição discreta com função de probabilidade $f_{\theta,j}$, em que $\theta > 0$ e $j = 1, 2$; $f_{\theta,1}$ é a distribuição de Poisson de média θ e $f_{\theta,2}$ é a distribuição geométrica de parâmetro $\frac{\theta}{1+\theta}$, isto é,

$$f_{\theta,2}(x) = \left[1 - \frac{\theta}{1+\theta}\right]\left[\frac{\theta}{1+\theta}\right]^x, \; x = 0, 1, 2, \ldots$$

Existe ENVVUM de θ? Justifique.

$X_1, \ldots, X_n$ a.a. com função de probabilidade $f_{\theta,j}$, $\theta > 0$ e $j = 1, 2$

i) $f_{\theta,1}(x) = \frac{e^{-\theta} \cdot \theta^x}{x!}$ $\quad x = 0, 1, 2, \ldots$

ii) $f_{\theta,2}(x) = \left(1 - \frac{\theta}{1+\theta}\right)\left(\frac{\theta}{1+\theta}\right)^x$ $\quad x = 0, 1, 2, \ldots$

i) Se $X \sim P(\theta)$, tem-se que sua distribuição de probabilidade será dada por:

$f_{\theta,1}(x) = \frac{e^{-\theta} \cdot \theta^x}{x!}$. Para uma a.a. tem-se que:

$$p_\theta(x) = \prod_{i=1}^n f_{\theta,1}(x_i) = \prod_{i=1}^n \frac{e^{-\theta} \theta^{x_i}}{x_i!} = \frac{e^{-n\theta} \cdot \theta^{\sum x_i}}{\prod x_i!} = \exp\left\{-n\theta + \log \theta^{\sum x_i}\right\} \frac{1}{\prod_{i=1}^n x_i} =$$

$$= \exp\left\{\sum x_i \log \theta - n\theta\right\} \frac{1}{\prod_{i=1}^n x_i!}$$

$\therefore$ A distribuição pertence à família exponencial, na qual $\eta = \log \theta$, $b(\eta) = n e^\eta$ e $c(x_i) = \left(\prod_{i=1}^n x_i!\right)^{-1}$. O domínio da função $\eta = \log \theta$ ($\theta = e^\eta$) contém retângulos $K=1$ dimensionais, e portanto, a família exponencial é de posto completo, e além disso, a estatística $T = \sum_{i=1}^n X_i$ será uma estatística suficiente completa (e minimal) para θ.

Então, se tomarmos $T^* = \bar{X}$, temos que:

$$E(T^*) = E(\bar{X}) = E\left(\frac{\sum X_i}{n}\right) = \frac{1}{n} E\left(\sum X_i\right) \overset{\text{iid}}{=} \frac{\cancel{n}}{\cancel{n}} E(X_1) = \theta.$$

Como $T^* = \bar{X}$ é não viciado para θ e função de uma estatística suficiente e completa para θ, então $T^* = \bar{X}$ é ENVVUM para θ.

$$E(X) = \frac{q}{p} = \frac{\frac{\theta}{1+\theta}}{1-\frac{\theta}{1+\theta}} = \frac{\frac{\theta}{\cancel{1+\theta}}}{\frac{1+\cancel{\theta}-\cancel{\theta}}{\cancel{1+\theta}}} = \theta$$

18/09/2020

ii) Se $X \sim G(\theta)$ sua função de densidade de probabilidade será dada por

$$f_\theta(x) = \left[1-\frac{\theta}{1-\theta}\right]\left[\frac{\theta}{1+\theta}\right]^x, \quad x=0,1,2...$$

Para uma a.a. tem-se que:

$$p_X(x) = \prod_{i=1}^{n} f_\theta(x_i) = \prod_{i=1}^{n}\left[1-\frac{\theta}{1-\theta}\right]\left[\frac{\theta}{1+\theta}\right]^{x_i} =$$

$$= \left[1-\frac{\theta}{1-\theta}\right]^n\left[\frac{\theta}{1+\theta}\right]^{\sum x_i} = \exp\left\{\sum x_i \log\left(\frac{\theta}{1+\theta}\right) + n\log\left(1-\frac{\theta}{1+\theta}\right)\right\}$$

∴ A distribuição pertence à família exponencial, na qual tem-se que $\eta = \log\frac{\theta}{1+\theta}$ e $b(\eta) = -n\log(1-e^\eta)$; observa-se que o domínio da função $\eta = \log\frac{\theta}{1+\theta}$ contém retângulos $k=1$ dimensionais e, portanto a família é de posto completo. Desse modo, a estatística $T=\sum_{i=1}^{n} X_i$ será uma estatística suficiente completa (e minimal) para θ.

Tomando $T^* = \bar{X}$, tem-se que:

$$E(T^*) = E(\bar{X}) = E\left(\frac{\sum X_i}{n}\right) = \frac{1}{n}E\left(\sum X_i\right) \overset{X_i\ iid}{=} \frac{\cancel{n}}{\cancel{n}}E(X) = \theta \quad *$$

Como T^* é não viciado para θ e é função de uma estatística suficiente e completa para θ, então T^* é ENVVUM para θ.

Conclusão: se $X \sim$ Poisson ou se $X \sim$ Geométrica, em ambos, o ENVVUM existe e é único, e será dado por

$$T = \frac{\sum X_i}{n} \Rightarrow \boxed{T = \bar{X}}$$

∴ O ENVVUM existe e é dado pela média amostral.

$p = \frac{\theta}{1+\theta}$ | $\delta = \frac{\sum X_i}{n}$ $E(\delta) = \frac{1}{n}EX$ $E(X) = \sum_{x=0}^{\infty} x(1-p)p^x = \sum_{x=1}^{\infty} x(1-p)p^x = (1-p)\sum_{x=1}^{\infty} x p^x \Rightarrow \frac{EX}{p} = (1-p)\sum_{x=1}^{n} x p^{x-1} \Rightarrow \frac{EX}{p} = (1-p)\sum_{x=1}^{\infty}\frac{dp^x}{dp} \Rightarrow$

$1-p = 1-\frac{\theta}{1+\theta} = \frac{1+\cancel{\theta}-\cancel{\theta}}{1+\theta} = \frac{1}{1+\theta}$

$$\Rightarrow \frac{E(X)}{p} = (1-p)\cdot\frac{d}{dp}\sum_{x=1}^{\infty} p^x \Rightarrow \frac{E(X)}{p} = (1-p)\cdot\frac{d}{dp}\left(\frac{p}{1-p}\right) = \cancel{(1-p)}\frac{(1-p+p)}{(1-p)^{\cancel{2}}} = \frac{1}{1-p}$$

Teorema do convergência dominada

$$EX = \frac{p}{1-p} = \theta$$

18 / 09/2010

11 (a) Seja δ um estimador não viesado de $g(\theta)$ e considere perda quadrática. Seja $\delta^* = \delta + B$, em que B é uma variável aleatória independente de δ e que não é nula com probabilidade 1. Mostre que δ^* tem risco uniformemente maior que o de δ.

δ é estimador não viesado de $g(\theta) \Rightarrow E_\theta(\delta(X)) = g(\theta)$ — não viesado

* Perda quadrática: $R(\theta, \delta(X)) = E\{(\delta(X) - g(\theta))^2\} = Var_\theta(\delta(X)) + \underbrace{Víes^2(\delta(X))}_{\geq 0} =$

$= Var(\delta(X))$

Seja,

$\delta^* = \delta + B$; B ind. δ e $P(B \neq 0) = 1$

Então

$$E(\delta^*) = E(\delta + B) = E(\delta) + E(B) = g(\theta) + E(B)$$

$\therefore$ o estimador δ^* é viesado para $g(\theta)$

Cálculo do viés de δ^*:

$$E(\delta^*) - g(\theta) = g(\theta) + E(B) - g(\theta) = E(B)$$

$$Var(\delta^*) = Var(\delta + B) = Var(\delta) + Var(B) + 2\,Cov(\delta, B) \underset{\text{por hipótese}}{\overset{\delta \text{ e } B \text{ independentes}}{=}} Var(\delta) + Var(B)$$

$$R(\theta, \delta^*(X)) = E\{(\delta^*(X) - g(\theta))^2\} = Var(\delta^*) + Víes^2(\delta^*) =$$

$$= Var(\delta) + Var(B) + E^2(B)$$

Como $P(B \neq 0) = 1$, então $Var(B) > 0$ e $E^2(B) > 0$, então

$$R(\theta, \delta) = Var(\delta)$$

$$R(\theta, \delta^*) = Var(\delta) + \underbrace{Var(B)}_{>0} + \underbrace{E^2(B)}_{>0} = R(\theta, \delta) + S, \text{ em que}$$

$$S = Var(B) + E^2(B) > 0$$

$$\therefore R(\theta, \delta^*) > R(\theta, \delta)$$

$\therefore$ δ^* tem risco uniformemente maior do que δ.

Outro modo:

$$R_\theta(\delta^*) = E(\delta + B - g)^2 = E((\delta - g) + B)^2 = E\Big(\underbrace{(\delta - g)^2}_{R(\delta)} + 2(\delta - g)B + \underbrace{B^2}_{\geq 0}\Big) =$$

$$= R_\theta(\delta) + \underbrace{2E_\theta\{(\delta - g(\theta))B\}}_{\text{independentes}} + E(B^2) = R_\theta(\delta) + K$$

$=$ $K \geq 0$

$$\underbrace{E(\delta - g(\theta))}_{=0 \text{ não viciado}} \cdot E(B) \qquad R_\theta(\delta^*) \geq R_\theta(\delta).$$

18/09/2020

(b) Definição: Um estimador δ de $g(\theta)$ é dito ser não viesado por risco (risk-unbiased) se satisfaz:

$$\underbrace{E_\theta(L(\theta,\delta))}_{R_\theta(\delta)} \leq \underbrace{E_\theta(L(\theta',\delta))}_{R_\theta(\delta^*)}, \text{ para todo } \theta' \neq \theta$$

Considere a função de perda da forma $L(\theta,d)=\rho(d-\theta)$ (função diferença) e seja δ um estimador de $g(\theta)$ não viesado por risco. Mostre que δ^* definido acima tem risco uniformemente maior ou igual ao de δ.

$$R_\theta(\delta) = \int_{-\infty}^{\infty} \underbrace{L(\theta,\delta(X))}_{\rho(\delta(X)-\theta)} f(x|\theta)\,dx \leq \int_{-\infty}^{+\infty} \underbrace{L(\theta',\delta(X))}_{\rho(\delta(X)-\theta')} f(x,\theta)\,dx \quad \theta'\neq\theta$$

$\boxed{\theta'=\theta+\mathcal{E}}$, $\mathcal{E}\neq 0$ $\quad \boxed{P(\mathcal{E}\neq 0)=1}$

$$\rho(\delta(X)-\theta-\mathcal{E}) = \rho((\underbrace{\delta(X)-\mathcal{E}}_{\delta^*})-\theta)$$

$$R_\theta(\delta^*) = \int_{-\infty}^{+\infty} L(\theta,\delta^*(x)) f(x|\theta)\,dx = \int_{-\infty}^{+\infty} \rho(\underbrace{\delta+B}_{\delta^*}-\theta)\,dx \underset{(**)}{\geq} R_\theta(\delta) \quad \forall\theta$$

B é uma v.a. independente de δ tal que $P(B\neq 0)=1$

O B faz o papel do $\mathcal{E}$, então vale a desigualdade $(**)$

18/09/2010

12. Suponha que $X_1, \ldots, X_n$ tenham distribuição de Poisson $P(\lambda)$. Use o método 1 para encontrar o ENVVUM de:

(a) λ^K para qualquer inteiro K positivo.

$$f_{X_i}(x_i) = \frac{e^{-\lambda} \cdot \lambda^{x_i}}{x_i!}$$

$$p_X(x) = \prod_{i=1}^{n} f_{X_i}(x_i) = \prod_{i=1}^{n} \frac{e^{-\lambda} \cdot \lambda^{x_i}}{x_i!} = \prod_{i=1}^{n} \exp\{-\lambda + \log \lambda^{x_i}\} \frac{1}{x_i!} =$$

$$= \exp\{\Sigma x_i \log \lambda - n\lambda\} \cdot \frac{1}{\prod_{i=1}^{n} x_i!}$$

$\eta = \log \lambda$ e $T = \Sigma X_i$ é uma estatística suficiente e completa para λ, pois esta estatística está associada à uma distribuição que pertence a família exponencial de posto completo (o domínio de $\eta = \log \lambda$ contém retângulos $K=1$ dimensional).

$\therefore T = \Sigma X_i$ é uma estatística suficiente e completa.

Se T é estatística suficiente e completa, o ENVVUM de $\forall$ função estimável $g(\lambda)$ é unicamente determinado pela equação: $T = \Sigma X_i \sim P(n\lambda)$

$$E_\lambda[\delta(T)] = g(\lambda)$$

Então.

$$E_\lambda[\delta(T)] = \lambda^K, \; K \text{ positivo}$$

$$\sum_{t=0}^{+\infty} \delta(t) \cdot \frac{e^{-n\lambda} \cdot (n\lambda)^t}{t!} = \lambda^K$$

$$e^{-n\lambda} \cdot \sum_{t=0}^{+\infty} \delta(t) \cdot \frac{(n\lambda)^t}{t!} = \lambda^K \qquad e^{n\lambda} = 1 + \frac{n\lambda}{1} + \frac{(n\lambda)^2}{2!} + \frac{(n\lambda)^3}{3!} + \ldots$$

$$\sum_{t=0}^{+\infty} \delta(t) \cdot \frac{(n\lambda)^t}{t!} = \underbrace{e^{n\lambda} \cdot \lambda^K}_{(*)} \qquad (*)\; e^{n\lambda} = \sum_{i=0}^{+\infty} \frac{(n\lambda)^i}{i!}$$

$$\sum_{t=0}^{+\infty} \delta(t) \cdot \frac{(n\lambda)^t}{t!} = \lambda^K \cdot \sum_{i=0}^{+\infty} \frac{(n\lambda)^i}{i!}$$

18/09/2010

$$\sum_{t=0}^{+\infty} \delta(t) \cdot \frac{(n\lambda)^t}{t!} = \lambda^k \cdot \sum_{i=0}^{+\infty} \frac{(n\lambda)^i}{i!}$$

$$\sum_{t=0}^{+\infty} \delta(t) \cdot \frac{(n\lambda)^t}{t!} = \sum_{i=0}^{+\infty} \lambda^k \cdot \frac{(n\lambda)^i}{i!}$$

$$\sum_{t=0}^{+\infty} \delta(t) \cdot \frac{(n\lambda)^t}{t!} = \sum_{i=0}^{+\infty} \frac{n^i \cdot \lambda^{k+i}}{i!}$$

Fazendo $t = i + k \Rightarrow i = t - k$, então $i = 0 \Rightarrow t = k$

$$\sum_{t=0}^{+\infty} \delta(t) \cdot \frac{(n\lambda)^t}{t!} = \sum_{t=k}^{+\infty} \frac{n^{t-k} \lambda^t}{(t-k)!}$$

$$\sum_{t=0}^{+\infty} \delta(t) \cdot \frac{(n\lambda)^t}{t!} = \sum_{t=k}^{+\infty} \frac{(n\lambda)^t}{n^k (t-k)!}$$

$$\delta(0) \cdot \frac{(n\lambda)^0}{0!} + \delta(1) \frac{(n\lambda)^0}{0!} + \dots + \delta(k) \frac{(n\lambda)^k}{k!} + \delta(k+1) \cdot \frac{(n\lambda)^{k+1}}{(k+1)!} + \dots = \frac{(n\lambda)^k}{n^k (k-k)!} + \frac{(n\lambda)^{k+1}}{n^k (k+1-k)!} + \dots$$

Comparando os termos das séries tem-se que:

$$\delta(0) = \delta(1) = \dots = \delta(k-1) = 0$$

$$\frac{\delta(k)}{\underbrace{k!}_{t}} = \frac{1}{n^k (\underbrace{k-k}_{t})!} \qquad \frac{\delta(k+1)}{(k+1)!} = \frac{1}{n^k (\underbrace{k+1-k}_{t})!}, \dots$$

Desse modo;

$$\delta(t) = \begin{cases} 0 & \text{se } t = 0, 1, \dots, k-1 \\ \dfrac{t!}{n^k (t-k)!} & \text{se } t = k, k+1, \dots \end{cases}$$

Portanto, o ENVVUM de λ^k é $\delta(T) = \dfrac{T!}{n^k (T-k)!}$, $T = k, k+1, \dots$

18/09/2010

b) $e^{-\lambda}$

Se T é estatística suficiente e completa, o ENVVUM de ∀ função estimável $g(\lambda)$ é unicamente determinado pela equação:

$$E_\lambda(\delta(T)) = g(\lambda) \quad \text{para todo } \lambda \in \Omega.$$

Então:

$$E_\lambda(\delta(T)) = e^{-\lambda}$$

$$\sum_{t=0}^{+\infty} \delta(t) \cdot \frac{e^{-n\lambda}(n\lambda)^t}{t!} = e^{-\lambda}$$

$$\sum_{t=0}^{+\infty} \delta(t) \cdot \frac{(n\lambda)^t}{t!} = e^{\lambda(n-1)}$$

$$\sum_{t=0}^{+\infty} \delta(t) \cdot \frac{(n\lambda)^t}{t!} = \sum_{t=0}^{+\infty} \frac{[\lambda(n-1)]^t}{t!}$$

$$\delta(0)\frac{(n\lambda)^0}{0!} + \delta(1)\frac{(n\lambda)^1}{1!} + \delta(2)\frac{(n\lambda)^2}{2!} + \delta(3)\frac{(n\lambda)^3}{3!} + \delta(4)\frac{(n\lambda)^4}{4!} + \dots =$$

$$= \frac{(\lambda(n-1))^0}{0!} + \frac{(\lambda(n-1))^1}{1!} + \frac{[\lambda(n-1)]^2}{2!} + \frac{[\lambda(n-1)]^3}{3!} + \dots$$

$$\delta(0) = 1 = \left(\frac{n-1}{n}\right)^0$$

$$\delta(1)\, n\lambda = \lambda n - \lambda$$

$$\delta(1) = \frac{\cancel{\lambda} n - \cancel{\lambda}}{n\cancel{\lambda}} = \left(\frac{n-1}{n}\right)$$

$$\frac{\delta(2)(n\lambda)^2}{\cancel{2!}} = \frac{\cancel{\lambda^2}(n^2 - 2n - 1)}{\cancel{2!}}$$

$$\delta(2) = \frac{(n-1)^2}{n^2} = \left(\frac{n}{n} - \frac{1}{n}\right)^2 = \left(1 - \frac{1}{n}\right)^2$$

$$\frac{\delta(3)(n^3\cancel{\lambda^3})}{\cancel{3!}} = \frac{\cancel{\lambda^3}(n-1)^3}{\cancel{3!}}$$

$$\delta(3) = \left(1 - \frac{1}{n}\right)^3$$

$$\boxed{\delta(T) = \left(1 - \frac{1}{n}\right)^{\sum_{i=1} X_i}}$$

$$\delta(t) = \left(1 - \frac{1}{n}\right)^t, \quad t = 0, 1, 2, \dots$$

$$\left(1 - \frac{1}{n}\right)^t = \left(1 - \frac{1}{n}\right)^{\sum X_i} = \Bigg[\underbrace{\left(1 - \frac{1}{n}\right)^n}_{e^{-1}}\Bigg]^{\bar{X}_n} \approx e^{-\bar{X}_n}$$

$n \to +\infty$

EMV de λ

13. Resolva o problema anterior (parte b) pelo método 2 usando o fato de que um estimador não viciado de $e^{-\lambda}$ é: $\delta = 1$, se $X_1 = 0$, e $\delta = 0$, caso contrário.

$$\delta(X) = \begin{cases} 1 & \text{se } X_1 = 0 \\ 0 & \text{c.c.} \end{cases}$$

$$E(\delta(X)) = 1 \cdot P(X_1 = 0) + 0 \cdot P(X_1 \neq 0) = 1 \cdot P(X_1 = 0) = \frac{\cancel{\lambda^0} e^{-\lambda}}{\cancel{0!}} = e^{-\lambda} \quad \text{ENV}$$

para $g(\lambda) = e^{-\lambda}$

Além disso, tem-se que $T = \sum_{i=1}^{n} X_i$ é estatística suficiente e completa para λ.

$T \sim P(n\lambda)$

$$E\left(\delta \,\Big|\, \sum_{i=1}^{n} X_i = t\right) = 1 \cdot P\left(X_1 = 0 \,\Big|\, \sum_{i=1}^{n} X_i = t\right) = \frac{P\left(X_1 = 0, \sum_{i=1}^{n} X_i = t\right)}{P\left(\sum_{i=1}^{n} X_i = t\right)} = \frac{P\left(\sum X_i = t \,/\, X_1 = 0\right) \cdot P(X_1 = 0)}{P\left(\sum_{i=1}^{n} X_i = t\right)}$$

$$= \frac{P\left(\sum_{i=2}^{n} X_i = t\right) \cdot P(X_1 = 0)}{P\left(\sum_{i=1}^{n} X_i = t\right)} =$$

$$= \frac{\dfrac{e^{-[(n-1)\lambda]} \cdot [(n-1)\lambda]^t}{t!} \, e^{-\lambda}}{\dfrac{e^{-n\lambda} \cdot (n\lambda)^t}{\cancel{t!}}} = \frac{\cancel{e^{-n\lambda}} \cdot \cancel{e^{\lambda}} \cdot (n-1)^t \cdot \cancel{\lambda^t} \cdot \cancel{e^{-\lambda}}}{\cancel{e^{-n\lambda}} \cdot n^t \cdot \cancel{\lambda^t}} = \left(\frac{n-1}{n}\right)^t = \underbrace{\left(1 - \frac{1}{n}\right)^t}_{\text{mesmo estimador dado no item anterior}}$$

$X_i \sim P(\lambda)$

$\sum_{i=1}^{n} X_i \sim P(n\lambda)$

$\sum_{i=2}^{n} X_i \sim P((n-1)\lambda)$

$$\boxed{T = \left(1 - \frac{1}{n}\right)^{\sum_{i=1}^{n} X_i}}$$

$$E\left(\delta(X) \,\Big|\, T = \sum_{i=1}^{n} X_i = t\right) = E\left(I(X_1 = 0) \,\Big|\, \sum_{i=1}^{n} X_i = t\right) = P\left(X_1 = 0 \,\Big|\, \sum X_i = t\right)$$

$$= \frac{P\left([X_1 = 0] \cap \left[\sum_{i=1}^{n} X_i = t\right]\right)}{P\left(\sum_{i=1}^{n} X_i = t\right)} = \frac{P\left(\sum_{i=1}^{n} X_i = t \,/\, P(X_1 = 0)\right) \cdot P(X_1 = 0)}{P\left(\sum_{i=1}^{n} X_i = t\right)} =$$

$$= \frac{P\left(\sum_{i=2}^{n} X_i = t\right) \cdot P(X_1 = 0)}{P\left(\sum_{i=1}^{n} X_i = t\right)}$$

14. Se $X_1,\dots,X_n$ são iid com distribuição $N(\theta,\sigma^2)$, com σ^2 conhecido, encontre o ENVVUM de θ^2, θ^3 e de θ^4. Sugestão: Para calcular $E(\bar{X}^k)$, escreva $\bar{X} = Y+\theta$, onde $Y\sim N(0,\frac{\sigma^2}{n})$ e expanda $E[(Y+\theta)^k]$

$$X \sim N(\theta,\sigma^2)$$

$$f_X(x) = \prod_{i=1}^{n} f_{X_i}(x_i) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{-\frac{1}{2\sigma^2}(x_i-\theta)^2\right\} =$$

$$= \frac{1}{(2\pi\sigma^2)^{\frac{n}{2}}} \exp\left\{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i-\theta)^2\right\} = \exp\left\{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i^2-2x_i\theta+\theta^2) - \frac{n}{2}\log(2\pi\sigma^2)\right\} = \exp\left\{-\frac{\sum_{i=1}^{n}x_i^2}{2\sigma^2} + \frac{\sum x_i\theta}{\sigma^2} - \frac{n\theta^2}{2\sigma^2} - \frac{n}{2}\log(2\pi\sigma^2)\right\} =$$

$$= \exp\left\{+\frac{\sum x_i\theta}{\sigma^2} - \frac{n\theta^2}{2\sigma^2} + \frac{\sum x_i^2}{2\sigma^2} - \frac{n}{2}\log(2\pi\sigma^2)\right\}$$

A distribuição pertence à família exponencial de posto completo pois $\eta = \frac{\theta}{\sigma^2}$ contém no seu domínio retângulos $k=1$ dimensionais e a estatística associada é $T=\sum_{i=1}^{n}X_i$, que é uma estatística suficiente e completa para θ. Como $T^*=\bar{X}$ é função 1 a 1 de T, então $\hat{T}=\bar{X}$ também é estatística suficiente e completa para θ.

Deseja-se estimar $g_1(\theta)=\theta^2$; $g_2(\theta)=\theta^3$ e $g_3(\theta)=\theta^4$.

Um candidato a ser estimador de g_1, seria $T_1 = \bar{X}^2$.

Escrevendo $\bar{X} = Y+\theta$, onde $Y\sim N(0,\frac{\sigma^2}{n})$, tem-se que:

$$E(\bar{X}^2) = E((Y+\theta)^2) = E(Y^2+2Y\theta+\theta^2) = E(Y^2)+2\theta E(Y)+\theta^2 =$$

$$= \frac{\sigma^2}{n}+2\theta 0+\theta^2 = \frac{\sigma^2}{n}+\theta^2 \Rightarrow E\left(\bar{X}^2-\frac{\sigma^2}{n}\right)=\theta^2$$

Então, tem-se que $\boxed{T_2 = \bar{X}^2-\frac{\sigma^2}{n}}$ é ENV de θ^2. Como T_2 é função de uma estatística suficiente e completa de θ, tem-se que T_2 é ENVVUM para $g_1(\theta)=\theta^2$.

23/09/2010

Para estimar $g_2(\theta)=\theta^3$, vamos verificar o estimador $T_3=\bar{X}^3$

$$E(\bar{X}^3) = E[(y+\theta)^3] = E[y^3+3y^2\theta+3y\theta^2+\theta^3] =$$

$$= E(y^3)+3\theta E(y^2)+3\theta E(y)+\theta^3 = \cancel{0+}3\theta\frac{\sigma^2}{n}\cancel{+3\theta\cdot 0}+\theta^3 =$$

$$y \sim N\left(0,\frac{\sigma^2}{n}\right) \qquad = +3\theta\frac{\sigma^2}{n}+\theta^3$$

$$M_y(t)=e^{\frac{\sigma^2t^2}{2n}}$$

$$M_y'(t) = e^{\frac{\sigma^2t^2}{2n}}\cdot\frac{\cancel{2}t\sigma^2}{\cancel{2}n} = e^{\frac{\sigma^2t^2}{2n}}\cdot\frac{t\sigma^2}{n}$$

$$E(y)=M_y'(t)\Big|_{t=0} = e^{\frac{0\cdot\sigma^2}{2}}\cdot\frac{0\cdot\sigma^2}{2n}=0$$

$$M_y''(t)= e^{\frac{\sigma^2t^2}{2n}}\left(\frac{t\sigma^2}{n}\right)^2 + e^{\frac{\sigma^2t^2}{2n}}\frac{\sigma^2}{n}$$

$$E(y^2)=M_y''(t)\Big|_{t=0} = e^{\frac{0\sigma^2}{2n}}\left(\frac{0\sigma^2}{n}\right)^2+e^{\frac{0\sigma^2}{2n}}\frac{\sigma^2}{n}=\frac{\sigma^2}{n}$$

$$M_y'''(t)= e^{\frac{\sigma^2t^2}{2n}}\left(\frac{t\sigma^2}{n}\right)^3 + e^{\frac{\sigma^2t^2}{2n}}\frac{2t\sigma^4}{n^2}+e^{\frac{\sigma^2t^2}{2n}}\left(\frac{\sigma^2 t}{n}\right)\frac{\sigma^2}{n}$$

$$E(y^3)=M_y'''(t)\Big|_{t=0} = e^{\frac{0\sigma^2}{2n}}\left(\frac{0\cdot\sigma^2}{n}\right)^3+e^{\frac{0\sigma^2}{2n}}\frac{2\cdot 0\cdot\sigma^2}{n}+e^{\frac{0\sigma^2}{2n}}\left(\frac{\sigma^2 0}{n}\right)\frac{\sigma^2}{n}=0$$

$$E\left(\bar{X}^3-3\bar{X}\frac{\sigma^2}{n}\right) = E(\bar{X}^3)-3\frac{\sigma^2}{n}E(\bar{X}) = \left(+3\theta\frac{\sigma^2}{n}+\theta^3\right)-\frac{3\sigma^2\theta}{n}=$$

$$=\cancel{\frac{3\theta\sigma^2}{n}}+\theta^3-\cancel{\frac{3\sigma^2\theta}{n}}=\theta^3$$

$$\boxed{T_4=\bar{X}^3-3\frac{\sigma^2}{n}\bar{X}} \text{ é um ENV de } g_2(\theta)=\theta^3.$$

Além disso, é uma função de uma estatística suficiente e completa dada por $T=\bar{X}$, logo T_4 é ENVVUM para $g_2(\theta)=\theta^3$.

23/09/2010

Para estimar $g_3(\mathcal{G}) = \mathcal{G}^4$, vamos verificar o estimador $T_4 = \bar{X}^4$

$E(\bar{X}^4) = E((y+\mathcal{G})^4) = E[(y+\mathcal{G}).(y^3+3y^2\mathcal{G}+3y\mathcal{G}^2+\mathcal{G}^3)] =$

$= E(y^4+3y^3\mathcal{G}+3y^2\mathcal{G}^2+y\mathcal{G}^3+\mathcal{G}y^3+3y^2\mathcal{G}^2+3y\mathcal{G}^3+\mathcal{G}^4) =$

$= E(y^4)+3\mathcal{G}E(y^3)+3\mathcal{G}^2E(y^2)+\mathcal{G}^3E(y)+\mathcal{G}E(y^3)+3\mathcal{G}^2E(y^2)+3\mathcal{G}^3E(y)+\mathcal{G}^4 \overset{*}{=}$

$M_y^{IV}(t) = e^{\frac{\sigma^2t^2}{2n}}\cancel{\left(\frac{t\sigma^2}{n}\right)^4} + e^{\frac{\sigma^2t^2}{2n}}.3\cancel{\left(\frac{t\sigma^2}{n}\right)^2} + e^{\frac{\sigma^2t^2}{2n}}.\cancel{\frac{\sigma^2 2t}{2n}}.\frac{2t\sigma^4}{n^2} + e^{\frac{\sigma^2t^2}{n}}.\frac{2\sigma^4}{n^2} +$

$+ e^{\frac{\sigma^2t^2}{2n}}.\frac{\sigma^2 2t}{2n}.\frac{\sigma^2 t}{n}.\frac{\sigma^2}{n} + e^{\frac{\sigma^2t^2}{2n}}.\frac{\sigma^2}{n}.\frac{\sigma^2}{n}$

$E(y^4) = M_y^{IV}(0) = \frac{2\sigma^4}{n^2}+\frac{\sigma^4}{n^2} = \frac{3\sigma^4}{n^2}$

$\overset{*}{=} \frac{2\sigma^4}{n^2}+\frac{\sigma^4}{n^2}+\cancel{3\mathcal{G}.0}+3.\mathcal{G}^2.\frac{\sigma^2}{n}+\cancel{\mathcal{G}^3.0}+\cancel{\mathcal{G}.0}+3\mathcal{G}^2.\frac{\sigma^2}{n}+\cancel{3.\mathcal{G}^3.0}+\mathcal{G}^4 =$

$= \frac{3\sigma^4}{n^2}+6\mathcal{G}^2\frac{\sigma^2}{n}+\mathcal{G}^4$

$E\left(\bar{X}^4+\frac{3\sigma^4}{n^2}-\frac{6\sigma^2\bar{X}^2}{n}\right) = E(\bar{X}^4)+\frac{3\sigma^4}{n^2}-\frac{6\sigma^2}{n}E(\bar{X}^2) =$

$= \frac{3\sigma^4}{n^2}+6\mathcal{G}^2\frac{\sigma^2}{n}+\mathcal{G}^4+\frac{3\sigma^4}{n^2}-\frac{6\sigma^2}{n}\left[+\mathcal{G}^2+\frac{\sigma^2}{n}\right] =$

$= \cancel{\frac{3\sigma^4}{n^2}}+\cancel{\frac{6\mathcal{G}^2\sigma^2}{n}}+\mathcal{G}^4+\cancel{\frac{3\sigma^4}{n^2}}-\cancel{\frac{6\sigma^2\mathcal{G}^2}{n}}-\cancel{\frac{6\sigma^4}{n^2}} = \mathcal{G}^4$

Então $T_5 = \bar{X}^4+\frac{3\sigma^4}{n^2}-\frac{6\sigma^2\bar{X}^2}{n}$ é um ENV de $g_3(\mathcal{G}) = \mathcal{G}^4$.

Como T_5 é função de uma estatística suficiente e completa ($T^* = \bar{X}$) para $\mathcal{G}$, então T_5 é ENVVUM para $g(\mathcal{G}) = \mathcal{G}^4$.

Outra forma:

$$\bar{X} \sim N\left(\mu, \frac{\sigma^2}{n}\right)$$

σ^2 conhecido.

$$\bar{X} = y + \mu$$

$$y \sim N\left(0, \frac{\sigma^2}{n}\right)$$

$$E(\bar{X}^2) = E(y+\mu)^2 = E(y^2 + 2\mu y + \mu^2) = \frac{\sigma^2}{n} + \mu^2$$

$\delta = \bar{X}^2 - \frac{\sigma^2}{n}$ é estimador ENVVUM para μ^2

$Z \sim N(0,1)$

$N(0,1) \to EX^3 = 0$

$N(0,1) \to EX^4 = 3$

$$E(\bar{X})^3 = \cancel{Ey^3}^{0} + 3E(y^2\mu) + 3(\cancel{y\mu^2})^{0} + E(\mu^3)$$

$$Y = \frac{\sigma}{\sqrt{n}} Z$$

$$E(\bar{X})^3 = 3\mu \frac{\sigma^2}{n} + \mu^3$$

$$\delta = \bar{X}^3 - 3\bar{X}\frac{\sigma^2}{n}$$

$$E(y^4) = E\left(\frac{\sigma}{\sqrt{n}} Z\right)^4 = \frac{\sigma^4}{n^2} EZ^4 = \frac{3\sigma^4}{n^2}$$

23/09/2019

15. Resolva o problema anterior admitindo que σ é desconhecido.

Como $X \sim N(\xi, \sigma^2)$, pelo exercício 1, sabe-se que $S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2$ é um ENV de σ^2, ou seja, $E_{\sigma^2}(S^2) = \sigma^2$, então, para estimarmos as funções dadas, tem-se que:

i) $g_1(\xi) = \xi^2$

$$T_2 = \bar{X}^2 - \frac{S^2}{n}$$

$T = (\bar{X}, S^2)$ é uma estatística suficiente e completa para estimar $T = (\xi, \sigma^2)$, então segue analogamente ao encontrado no exercício 14.

ii) $g_2(\xi) = \xi^3$

$$T_3 = \bar{X}^3 - \frac{3S^2 \cdot \bar{X}}{n}$$

(aqui continua sendo válido, pois por Basu, S^2 é independente de $\bar{X}$.)
($\bar{X}$ - estatística suficiente e completa e S^2 - estatística ancilar)

iii) $g_3(\xi) = \xi^4$

$$Var(S^2) = E(S^4) - [E(S^2)]^2$$

Pelo exercício 1;

$$\frac{2\sigma^4}{n-1} = E(S^4) - [\sigma^2]^2$$

$$E(S^4) = \frac{2\sigma^4}{n-1} + \sigma^4 = \frac{2\sigma^4 + n\sigma^4 - \sigma^4}{n-1} = \left(\frac{n+1}{n-1}\right)\sigma^4$$

Pelo exercício 3:

$$T_4 = \frac{3\sigma^4}{n^2} + \frac{6\sigma^2\bar{X}^2}{n} + \bar{X}^4$$

Novamente
* Basu

Então

$$T_5 = \frac{3}{n^2}\left(\frac{n+1}{n-1}\right)S^4\left(\frac{n-1}{n+1}\right) + \frac{6S^2\bar{X}^2}{n} + \bar{X}^4$$

SÃO DOMINGOS

$E(S^2)^2 \overset{?}{=} \sigma^4$

01/10/2010

$\frac{(n-1)}{\sigma^2} S^2 \sim \chi^2_{n-1}$

$S^2 = \frac{\sigma^2}{n-1} \chi^2_{n-1}$

$\frac{\sigma^4}{(n-1)^2} 2(n-1) = Var(S^2) = ES^4 - (ES^2)^2$

$\left(\frac{\sigma^2}{\cancel{n-1}} \cancel{n-1}\right)$

$\frac{2\sigma^4}{n-1} = ES^4 - \sigma^4$

$\frac{2\sigma^4}{n-1} + \sigma^4 = ES^4$

$ES^4 = \frac{(n+1)}{(n-1)} \sigma^4$

$E\left(\frac{n-1}{n+1} S^4\right) = \sigma^4$

16 a) Completitude

$* \ \forall a \in \mathbb{R}, \ \int_a^{+\infty} \underbrace{f(x) e^{-nx}}_{h(x)} dx = 0$

$\int_{a_1}^{a_2} h(x)\, dx = 0$

$a_1 \quad a_2$

$(\ A \ )$

$\int_A h(x)\, dx = 0$

$\Downarrow$

$h(x) = 0$ q.c

$\underbrace{\int_{a_1}^{\infty} h(x)\, dx}_{=0} = \int_{a_1}^{a_2} h(x)\, dx + \underbrace{\int_{a_2}^{\infty} h(x)\, dx}_{=0}$

$\Downarrow$

$= 0$

16. Seja $(X_1, \dots, X_n)$ uma amostra casual simples de $X \sim E(a,b)^2$ com função densidade de probabilidade dada por:

$$f(x; a, b) = b^{-1} \exp\left\{ -\frac{(x-a)}{b} \right\} I_{(a,+\infty)}(x)$$

para $a \in (-\infty, +\infty)$, $b>0$. Assuma que $b=1$.

(a) Mostre que $X_{(1)} = \min_{i \in \{1,\dots,n\}} X_i$ é uma estatística suficiente e completa.

Assumindo $b=1$, tem-se que a função densidade de probabilidade será dada por:

$$f(x; a) = \exp\{-(x-a)\} I_{(a,+\infty)}(x).$$

i) Suficiência:

Considerando uma amostra aleatória simples $(X_1, \dots, X_n)$ na qual $X_i \sim E(a,1)$ tem-se que a densidade conjunta será dada por:

$$p_X(x) = \prod_{i=1}^{n} f_{X_i}(x_i) = \prod_{i=1}^{n} \exp\{-(x_i - a)\} I_{(a,+\infty)}(x_i) =$$

$$= \exp\left\{ -\sum_{i=1}^{n} x_i + na \right\} \prod_{i=1}^{n} I_{(a,+\infty)}(x_i) \qquad a < x_i < +\infty$$

$$= \exp\left\{ -\sum_{i=1}^{n} x_i + na \right\} I_{(a,+\infty)}(x_{(1)}) = \underbrace{\exp\{na\} I_{(a,+\infty)}(x_{(1)})}_{g_a(T(x))} \cdot \underbrace{\exp\left\{ -\sum_{i=1}^{n} x_i \right\}}_{h(x)}$$

Pelo Critério da Fatoração, tem-se que $g_a(T(X)) = \exp\{na\} I_{(a,+\infty)}(x_{(1)})$ e $h(x) = \exp\left\{ -\sum_{i=1}^{n} x_i \right\}$. Desse modo, a estatística suficiente será dada por $T = X_{(1)} = \min_{i \in \{1,\dots,n\}} X_i$.

ii) Completa

* Distribuição do Mínimo

$f(X_{(1)},\dots,X_{(n)}) = n!\, f(X_1)\dots f(X_n)$

$X_{(1)}$ $X_{(n)}$

20/09/2010

$f(X_{(n)}) = \frac{n!}{(n-1)!} [P(X_i < X)]^{n-1} f(X) = n\,[F(X)]^{n-1} f(X)$

combinações anteriores

seja $y = X_{(1)}$

$F_{X_{(n)}}(X) = [F(X)]^n$

$n\,[F(X)]^{n-1} f(x)$

$F_Y(y) = P(Y \le y) = P(X_{(1)} \le y) = 1 - P(X_{(1)} > y) =$

$= 1 - [P(X_1 > y, X_2 > y, \dots, X_n > y] \overset{X_i\text{'s ind}}{=} 1 - [P(X_1 > y).P(X_2 > y)\dots P(X_n > y)] \overset{X_i\text{'s ident.}}{=}$

$= 1 - (P(X_1 > y))^n = 1 - \left[\int_y^{+\infty} e^{-(z-a)} dz\right]^n =$

$= 1 - \left[\int_y^{+\infty} e^{-z} e^{a} dz\right]^n = 1 - \left[e^a \int_y^{+\infty} e^{-z} dz\right]^n =$

$= 1 - \left[e^a \left. -e^{-z}\right|_y^{+\infty}\right]^n = 1 - [e^a e^{-y}]^n = 1 - e^{n(a-y)}$

$f_Y(y) = n\, e^{-n(y-a)}\, I_{(a,+\infty)}(y)$ $\quad X_{(1)} \sim E(a, \frac{1}{n})$

Para que uma estatística seja completa devemos ter:

$E(g(t)) = 0 \Rightarrow g(t) \equiv 0$

$\int_A g(X) f_{X_{(1)}}(x|\theta)\,dx = 0 \quad \forall A \in$ Borelianos $\Leftrightarrow$

$\Leftrightarrow g(X_{(1)}) = 0$ q.c. em P onde $P = \{E(\theta, 1),\ \theta \in \Theta\} \Leftrightarrow$

$\Leftrightarrow X_{(1)}$ é completa para a $\in \Theta$

Então

$\int_a^{+\infty} g(y).n.e^{n(a-y)} dy = 0$

$\int_a^{+\infty} g(y)\, e^{-ny} dy = 0, \ \forall a \in \mathbb{R} \Leftrightarrow \int_A g(y)\, e^{-ny} dy = 0$, A pertence aos Borelianos $\Leftrightarrow$

$g(y).e^{-ny} \equiv 0$ q.c. em $P \Leftrightarrow g(y) \equiv 0$ q.c. em P

$\therefore X_{(1)} = \min_{i \in \{1,\dots,n\}} X_i$ é uma estatística suficiente e completa em $P = \{E(a,1), a \in \mathbb{R}\}$

*Ver na folha anterior

A

a_1 a_2 $+\infty$

logo essa tem que ser zero

$\int_a^{+\infty} g(y)\, e^{-ny} dy = 0 \Rightarrow \int_{a_1}^{a_2} g(y).e^{-ny} dy + \int_{a_2}^{+\infty} g(y)\, e^{-ny} dy = \int_a^{+\infty} g(y) e^{-ny} dy = 0$

0 0

Mood graybill estatísticas de ordem e distribuições completo

$$X_{(1)} \quad (4) \quad X_{(5)}\ X \quad (1) \quad X_{(7)}\ Y \quad (n-7) \quad X_{(n)}$$

$$f_{X_{(5)},X_{(7)}}(X,Y) = \frac{n!}{4!\,1!\,(n-7)!}[F(X)]^4 f(X)\,[F(Y)-F(X)] \cdot [1-F(Y)]^{n-7}$$

20/09/2010

(b) Mostre que $X_{(1)}$ e S^2, a variância amostral, são variáveis aleatórias independentes

Pelo item a) tem-se que $X_{(1)}$ é uma estatística suficiente e completa.

Seja S^2 a variância amostral, dado por:

$$S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i-\bar{X})^2$$

$$\boxed{f_{X_{(1)}}(X) = \frac{n!}{(n-1)!}f(x)[1-F(X)]^{n-1}}$$

Densidade do Mínimo

Mas $X_i \sim E(a,1)$; que pertence a uma família de locação. Desse modo, X_i pode ser reescrito como:

$X_i = U_i + a, \quad U \sim E(1)$

Além disso,

$$\bar{X} = \frac{\Sigma X_i}{n} = \frac{X_1+X_2+\dots+X_n}{n}$$

Cada X_i acima pertence a uma família de locação; então a soma acima poderia ser reescrito como:

$$\bar{X} = \frac{\Sigma X_i}{n} = \frac{X_1+X_2+\dots+X_n}{n} = \underbrace{\frac{U_1+a+U_2+a+U_3+a+\dots+U_n+a}{n}}_{X_i=U_i+a,\ U_i\sim Exp(1)} = \frac{\sum_{i=1}^{n} U_i}{n} + \frac{na}{n}$$

$$\bar{X} = \frac{\Sigma U_i}{n} + a$$

Desse modo, a diferença $X_i - \bar{X}$ pode ser escrito da forma:

$$X_i - \bar{X} = U_i + a - \left(\frac{\sum_{i=1}^{n} U_i}{n} + a\right) = U_i + \cancel{a} - \bar{U} - \cancel{a} = U_i - \bar{U}$$

A diferença $U_i - \bar{U}$ independe de a, ou seja, $X_i - \bar{X}$ independe de a, e assim, $S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i-\bar{X})^2$ independe de a, e portanto, S^2 é uma estatística ancilar para a.

Desse modo, pelo Teorema de Basu, como $X_{(1)}$ é uma estatística suficiente e completa e S^2 é uma estatística ancilar, $X_{(1)}$ e S^2 são v.a. independentes.

21/09/2010

17. Seja $(X_1, \dots, X_n)$ uma amostra casual simples de $X \sim E(a,b)$. Mostre que $X_{(1)} = \min_{i \in \{1,\dots,n\}} X_i$ e $\sum (X_i - X_{(1)})$ são independentes e têm distribuições $E(a, \frac{b}{n})$ e b Gama $(n-1, 1)$. Sugestão: se $a=0$ e $b=1$, as variáveis $Y_i = (n-i+1)[X_{(i)} - X_{(i-1)}]$ para $i=2,\dots,n$ são iid e têm distribuição $E(0,1)$.

Se $X \sim E(a,b)$ sua densidade é da forma:

$$f_X(x) = \frac{1}{b} \exp\left\{-\frac{(x-a)}{b}\right\} I_{(a,+\infty)}(x) \qquad b>0, \text{ conhecido.}$$

Para uma amostra aleatória $i=1,\dots,n$, tem-se que:

$$p_X(x) = \prod_{i=1}^{n} f_{X_i}(x_i) = \prod_{i=1}^{n} \frac{1}{b} \exp\left\{-\frac{(x_i-a)}{b}\right\} I_{(a,+\infty)}(x_i) =$$

$$= \frac{1}{b^n} \exp\left\{-\sum_{i=1}^{n} \frac{(x_i-a)}{b}\right\} \prod_{i=1}^{n} I_{(a,+\infty)}(x_i) = \quad a < x_i < +\infty$$

$$= \frac{1}{b^n} \exp\left\{-\sum_{i=1}^{n} \frac{x_i}{b} + \frac{na}{b}\right\} I_{(a,+\infty)}(x_{(1)}) =$$

$$= \underbrace{\exp\left\{\frac{na}{b}\right\} I_{(a,+\infty)}(x_{(1)})}_{g_a(T(x))} \underbrace{\frac{1}{b} \exp\left\{-\frac{\sum x_i}{b}\right\}}_{h(x)}$$

Pelo Critério da fatoração, tem-se que: $g_a(T(X)) = \exp\left\{\frac{na}{b}\right\} I_{(a,+\infty)}(x_{(1)})$ e $h(x) = \frac{1}{b} \exp\left\{-\frac{\sum x_i}{b}\right\}$. Desse modo, $T = X_{(1)} = \min\{X_1, \dots, X_n\}$ é uma estatística suficiente para a, supondo $b>0$, conhecido.

Cálculo da distribuição do mínimo. $Y = X_{(1)}$

$$F_Y(y) = P(Y \le y) = P(X_{(1)} \le y) = 1 - P(X_{(1)} > y) = 1 - [P(X_1 > y, X_2 > y, \dots, X_n > y)] =$$

$$= 1 - [P(X_1 > y) \cdot P(X_2 > y) \dots P(X_n > y)] = 1 - [P(X_1 > y)]^n = 1 - \left[\int_y^{+\infty} \frac{1}{b} e^{-\frac{(z-a)}{b}} dz\right]^n =$$

$$= 1 - \left[\frac{e^{\frac{a}{b}}}{b} \int_y^{+\infty} e^{-\frac{z}{b}} dz\right]^n = 1 - \frac{e^{\frac{a}{b}}}{\cancel{b}} \cancel{b} \cdot -e^{-\frac{z}{b}}\Big|_y^{+\infty} = 1 - e^{\frac{na}{b} - \frac{ny}{b}}$$

$$f_Y(y) = \frac{d F_Y(y)}{dy} = \frac{n}{b} e^{-\frac{n}{b}(y-a)}$$

/ /

$$f_Y(y) = \frac{n}{b} e^{-\frac{n}{b}(y-a)} \, I_{(a,\infty)}(y)$$

$$X_{(1)} \sim E\left(a, \frac{b}{n}\right)$$

$$E(g(t)) = 0 \Rightarrow g(t) \equiv 0$$

$$\Downarrow$$

$$\int_a^{+\infty} g(t) \cdot \frac{n}{b} e^{-\frac{n}{b}(t-a)} \, dt = 0$$

Como $e^{-\frac{n}{b}(t-a)} > 0$, então $g(t) = 0$

$\therefore$ $X_{(1)}$ é estatística suficiente e completa para a.
(mais detalhes 16(a))

Se $a=0$ e $b=1$

$$f_X(x) = e^{-x} \, I_{(0,+\infty)}(x) \qquad X_3 - X_{(1)} = X_{(3)} - X_{(2)} + X_{(2)} - X_{(1)}$$

$$\sum_{i=2}^{n} (X_i - X_{(1)}) = (X_2 - X_{(1)}) + (X_3 - X_{(1)}) + \dots + (X_n - X_{(1)})$$

$Y_i = (n-i+1)\,[X_{(i)} - X_{(i-1)}]$ para $i = 2, \dots n$

$Y_2 = (n-2+1)\,[X_{(2)} - X_{(2-1)}] = (n-1) \cdot [X_{(2)} - X_{(1)}]$

$Y_3 = (n-3+1)\,[X_{(3)} - X_{(3-1)}] = (n-2) \cdot [X_{(3)} - X_{(2)}]$

$Y_4 = (n-4+1)\,[X_{(4)} - X_{(4-1)}] = (n-3)\,[X_{(4)} - X_{(3)}]$

$\vdots$

$Y_n = (n-n+1)\,[X_{(n)} - X_{(n-1)}] = 1 \cdot [X_{(n)} - X_{(n-1)}]$

$$Y_i \sim E(0,1)$$

$$\sum_{i=1}^{n} Y_i \sim \text{Gama}(n-1; 1)$$

$(X_i - X_{(1)})$ é ancilar pois X_i pode ser escrito da forma

$X_i = b \cdot U_i + a$, $U_i \sim E(0,1)$

e $X_{(1)}$ pode ser escrito da forma

$X_{(1)} = \frac{b}{n} \cdot U_1 + a$, $U_i \sim E(0,1)$

18. Suponha que X_i, para $i = 1, \dots, n$, sejam v.a's com distribuição $N(\alpha + \beta t_i, \sigma^2)$, onde α, β e σ^2 são desconhecidos, e os t's são constantes conhecidos não todos nulos. Encontre o ENVVUM de α e β.

$X_i \sim N(\alpha + \beta t_i, \sigma^2)$ $\quad$ α, β e σ^2 desconhecidos

t's constantes

$$f_{X_i}(x_i) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{-\frac{1}{2\sigma^2}(x_i - \alpha - \beta t_i)^2\right\}$$

$$p_X(x) = \prod_{i=1}^{n} f_{X_i}(x_i) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{-\frac{1}{2\sigma^2}(x_i - \alpha - \beta t_i)^2\right\} =$$

$$= \frac{1}{(2\pi\sigma^2)^{\frac{n}{2}}} \exp\left\{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i - \alpha - \beta t_i)^2\right\} =$$

$$= \exp\left\{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i - (\alpha + \beta t_i))^2 - \frac{n}{2}\log 2\pi\sigma^2\right\} =$$

$$= \exp\left\{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}\left(x_i^2 - 2x_i(\alpha + \beta t_i) + (\alpha + \beta t_i)^2\right) - \frac{n}{2}\log 2\pi\sigma^2\right\} =$$

$$= \exp\left\{-\frac{1}{2\sigma^2}\sum_{i=1}^{n}\left(x_i^2 - 2x_i\alpha + 2\beta x_i t_i + \alpha^2 + 2\alpha\beta t_i + \beta^2 t_i^2\right) - \frac{n}{2}\log 2\pi\sigma^2\right\} =$$

$$= \exp\left\{-\frac{1}{2\sigma^2}\left(\sum_{i=1}^{n} x_i^2 - 2\alpha\sum x_i - 2\beta\sum x_i t_i + n\alpha^2 + 2\alpha\beta\sum t_i + \beta^2\sum t_i^2\right) - \frac{n}{2}\log 2\pi\sigma^2\right\}$$

$$= \exp\left\{-\frac{1}{2\sigma^2}\sum x_i^2 + \frac{\alpha\sum x_i}{\sigma^2} + \frac{\beta}{\sigma^2}\sum x_i t_i - \frac{n\alpha^2}{2\sigma^2} - \frac{\alpha\beta\sum t_i}{\sigma^2} - \frac{\beta^2\sum t_i^2}{2\sigma^2} - \frac{n}{2}\log 2\pi\sigma^2\right\}$$

$\eta_1 = \frac{-1}{2\sigma^2}$ $\quad$ $T_1 = \sum X_i^2$

$\eta_2 = \frac{\alpha}{\sigma^2}$ $\quad$ $T_2 = \sum X_i$

$\eta_3 = \frac{\beta}{\sigma^2}$ $\quad$ $T_3 = \sum X_i t_i$

A distribuição pertence a família exponencial tridimensional, no qual o domínio das funções η's $\in \overline{\mathbb{R}}\times\mathbb{R}\times\mathbb{R}$. e assim, a família exponencial é de posto completo e a estatística $T = \left(\sum_{i=1}^{n} X_i^2, \sum_{i=1}^{n} X_i, \sum_{i=1}^{n} X_i t_i\right)$ é suficiente completa e minimal.

22/09/2010 Encontrando estimadores pelo método da máxima verossimilhança

$$f_X(x) = \exp\left\{ -\frac{1}{2\sigma^2}\sum x_i^2 + \frac{\alpha\sum x_i}{\sigma^2} + \frac{\beta}{\sigma^2}\sum x_i t_i - \frac{n\alpha^2}{2\sigma^2} - \frac{\alpha\beta\sum t_i}{\sigma^2} - \frac{\beta^2\sum t_i^2}{2\sigma^2} - \frac{n}{2}\log 2\pi\sigma^2 \right\}$$

$\theta = (\alpha, \beta, \sigma^2)$

$$l(\theta) = -\frac{1}{2\sigma^2}\sum x_i^2 + \frac{\alpha\sum x_i}{\sigma^2} + \frac{\beta}{\sigma^2}\sum x_i t_i - \frac{n\alpha^2}{2\sigma^2} - \frac{\alpha\beta\sum t_i}{\sigma^2} - \frac{\beta^2\sum t_i^2}{2\sigma^2} - \frac{n}{2}\log 2\pi\sigma^2$$

$$\frac{\partial l(\theta)}{\partial\beta} = \frac{\sum x_i t_i}{\sigma^2} - \frac{\alpha\sum t_i}{\sigma^2} - \frac{\cancel{2}\beta\sum t_i^2}{\cancel{2}\sigma^2} = 0$$

$$\frac{\partial l(\theta)}{\partial\alpha} = \frac{\sum x_i}{\sigma^2} - \frac{\cancel{2}n\alpha}{\cancel{2}\sigma^2} - \frac{\beta\sum t_i}{\sigma^2} = 0$$

$$\sum x_i - \beta\sum t_i = n\alpha$$

$$\alpha = \frac{\sum x_i}{n} - \hat{\beta}\frac{\sum t_i}{n}$$

$$\boxed{\hat{\alpha} = \bar{X} - \hat{\beta}\bar{t}}$$

$$\frac{\sum x_i t_i}{\sigma^2} - \frac{\alpha\sum t_i}{\sigma^2} - \frac{\beta\sum t_i^2}{\sigma^2} = 0 \Rightarrow \sum x_i t_i - \sum t_i(\bar{X} - \beta\bar{t}) - \beta\sum t_i^2 = 0$$

$$\sum x_i t_i - \frac{\sum t_i \sum x_i}{n} + \frac{\beta\sum t\sum t}{n} - \beta\sum t_i^2 = 0$$

$$\boxed{\hat{\beta} = \frac{\sum X_i t_i - n\bar{t}\bar{X}}{\sum t_i^2 - n\bar{t}^2}}$$

(Supondo hessiana negativa definida)

$$E(\hat{\beta}) = E\left(\frac{\sum X_i t_i - \bar{t}\bar{X}}{\sum t_i^2 - n\bar{t}^2}\right) = \frac{1}{\sum t_i^2 - n\bar{t}^2}\left[E\left(\sum X_i t_i\right) - E(n\bar{t}\bar{X})\right] =$$

$$= \frac{1}{\sum t_i^2 - n\bar{t}^2}\left(\cancel{\alpha\sum t_i} + \beta\sum t_i^2 - \cancel{\alpha\sum t_i} - \beta n\bar{t}^2\right) = \beta \cdot \frac{\cancel{(\sum t_i^2 - n\bar{t}^2)}}{\cancel{\sum t_i^2 - n\bar{t}^2}} = \beta$$

$$E(\hat{\alpha}) = E(\bar{X} - \hat{\beta}\bar{t}) = E(\bar{X}) - E(\hat{\beta}\bar{t}) = \alpha + \beta\bar{t} - \beta\bar{t} = \alpha$$

Os EMV são dados por $\hat{\alpha} = \bar{X} - \hat{\beta}\bar{t}$ e $\hat{\beta} = \frac{\sum X_i t_i - n\bar{t}\bar{X}}{\sum t_i^2 - n\bar{t}^2}$ são não viesados e são funções da estatística suficiente e completa $T = (\sum X_i^2, \sum X_i, \sum X_i t_i)$ então $\hat{\alpha}$ e $\hat{\beta}$ são ENVVUM de α e β, respectivamente.

21/09/2010

19. Sejam $X_1, \dots, X_m$ e $Y_1, \dots, Y_n$ variáveis aleatórias independentes com distribuições $U(0,\theta)$ e $U(0,\theta')$, respectivamente se $n>1$, determine o ENVVUM de $\frac{\theta}{\theta'}$

Se $X_i \sim U(0,\theta)$.

$$f_{\underline{X}}(\underline{x}) = \prod_{i=1}^{n} f_{X_i}(x_i) = \prod_{i=1}^{n} \frac{1}{\theta} I_{(0,\theta)}(x_i) = \frac{1}{\theta^n} \prod_{i=1}^{n} I_{(0,\theta)}(x_i) \; \underset{0<x_i<\theta}{=} \; \frac{1}{\theta^n} I_{(0,\theta)}(x_{(n)}) \quad ①$$

Pelo Critério da fatoração, tem-se que $X_{(n)} = \max\{X_1, \dots, X_n\}$ é uma estatística suficiente para θ.

$$P(X_{(n)} < x) = P(X_1 < x, X_2 < x, \dots, X_n < n) = [P(X_1 < x)]^n = \left[\int_0^x \frac{1}{\theta}\, da\right]^n = \left[\frac{x}{\theta}\right]^n$$

$$f_{X_{(n)}}(x_{(n)}) = \frac{1}{\theta^n} \cdot n\, x_{(n)}^{n-1} I_{(0,\theta)}(x_{(n)})$$

$$E(X_{(n)}) = \int_0^\theta x \cdot \frac{1}{\theta^n} \cdot n\, x^{n-1}\, dx = \frac{n}{\theta^n}\int_0^\theta x^n\, dx = \frac{n}{\theta^n} \cdot \frac{x^{n+1}}{n+1}\Big|_0^\theta =$$

$$= \frac{n}{\theta^n} \cdot \frac{\theta^{n+1}}{n+1} = \frac{n}{n+1} \cdot \theta \quad (*) \qquad T = \frac{n+1}{n} \cdot X_{(n)} \text{ é um ENV de } \theta.$$

* Verificar se $X_{(n)}$ é uma estatística completa para θ.

$$E_\theta(g(t)) = 0 \Rightarrow g(t) \equiv 0$$

$$\int_0^\theta g(t) \cdot \frac{1}{\theta^n} n\, t^{n-1}\, dt = 0 \underset{\forall \theta>0}{\Rightarrow} \frac{d}{d\theta}\int_0^\theta g(t) \frac{1}{\theta^n} n \cdot t^{n-1}\, dt = 0 \Leftrightarrow g(\theta)\frac{1}{\theta^n} n\, \theta^{n-1} = 0 \Leftarrow$$

$$\Leftrightarrow g(\theta) \cdot n \cdot \theta^{-1} = 0 \underset{n \cdot \theta^{-1} \neq 0}{\Longrightarrow} g(\theta) \equiv 0 \; \forall \theta > 0 \Rightarrow g(t) = 0 \qquad \therefore T = X_{(n)} \text{ é uma estatística completa} \quad ②$$

Portanto, $T^* = \frac{n+1}{n} X_{(n)}$ é um ENVVUM para θ pois é não viciado (ver (*)) e função de uma estatística suficiente e completa $T = X_{(n)}$ para θ. (ver ① e ②).

Se $Y \sim U(0,\theta')$, então $Z = \frac{1}{Y}$ terá uma densidade da forma: $\frac{1}{\theta'} < z < +\infty$

$$F_Z(z) = P(Z \leq z) = P\left(\frac{1}{Y} \leq z\right) = P\left(Y \geq \frac{1}{z}\right) = \int_{\frac{1}{z}}^{\theta'} \frac{1}{\theta'}\, da =$$

$$= \frac{1}{\theta'} a \Big|_{\frac{1}{z}}^{\theta'} = \frac{1}{\theta'}\left(\theta' - \frac{1}{z}\right) \qquad f_Z(z) = \frac{1}{\theta'} \cdot \frac{1}{z^2} I_{(\frac{1}{\theta'}, +\infty)}(z)$$

$P(Z < z) =$
$(z - \frac{1}{z} < \dots)$

/ /

$$f_Z(z) = \frac{1}{\theta'} \frac{1}{z^2} I_{(\frac{1}{\theta'}, +\infty)}(z)$$

$$f_{\underline{Z}}(\underline{z}) = \prod_{i=1}^{n} f_{Z_i}(z_i) = \frac{1}{\theta'^n} \cdot \frac{1}{\prod_{i=1}^{n} z^2} I_{(\frac{1}{\theta'}, +\infty)}(z_{(1)})$$

Pelo Critério da fatoração tem-se que o mínimo é estatística suficiente para θ'

$$F_{Z_{(1)}}(z_{(1)}) = P(Z_{(1)} < z) = 1 - P(Z_{(1)} > z) = 1 - [P(Z_i > z)]^n = 1 - [1 - P(Z \leq z)]^n$$

$$= 1 - \left[1 - \int_{\frac{1}{\theta'}}^{z} \frac{1}{\theta'} \frac{1}{a^2} da\right]^n = 1 - \left[1 - \frac{1}{\theta'} \int_{\frac{1}{\theta'}}^{z} a^{-2} da\right]^n =$$

$$= 1 - \left[1 - \frac{1}{\theta'} \left.\frac{a^{-1}}{-1}\right|_{\frac{1}{\theta'}}^{z}\right]^n =$$

$$= 1 - \left[1 - \frac{1}{\theta'} \left[-\frac{1}{z} + \frac{1}{\frac{1}{\theta'}}\right]\right]^n = 1 - \left[\cancel{1} + \frac{1}{\theta' z} - \cancel{1}\right]^n = 1 - \left[\frac{1}{\theta' z}\right]^n$$

$$f_{Z_{(1)}}(z) = \frac{-1}{\theta'^n} \cdot -n \, (z)^{-n-1} = \frac{n}{\theta'^n} \, z^{-n-1} I_{(\frac{1}{\theta'}, \infty)}(z)$$

$$E(Z_{(1)}) = \int_{\frac{1}{\theta'}}^{+\infty} z \frac{n}{\theta'^n} z^{-n-1} dz = \frac{n}{\theta'^n} \left.\frac{z^{-n+1}}{-n+1}\right|_{\frac{1}{\theta'}}^{+\infty} = -\frac{n}{-n+1} \frac{((\theta')^{-1})^{-n+1}}{\theta'^n} = \frac{-n}{-n+1} \frac{1}{\theta'}$$

$$T_2^* = \frac{n-1}{n} Z_{(1)} \text{ é ENV de } \frac{1}{\theta'}$$

Tem-se ainda que $(Z_{(1)})$ é completa pois:

$$E_{\theta'}(g(t)) = 0 \Rightarrow g(t) \equiv 0 \Rightarrow \int_{\frac{1}{\theta'}}^{+\infty} g(t) \cdot \frac{n}{\theta'^n} t^{-n-1} dt = 0 \Longleftrightarrow g\left(\frac{1}{\theta'}\right) \cdot \frac{n}{\theta'^n} \left(\frac{1}{\theta'}\right)^{-n-1} = 0$$

$$g\left(\tfrac{1}{\theta'}\right) = 0 \;\; \forall \tfrac{1}{\theta'} > 0$$

Como $T_2^* = \frac{n-1}{n} \frac{1}{Z_{(1)}'}$ é ENV de $\frac{1}{\theta'}$ e é uma função da estatística suficiente e completa para θ', então ele será ENVVUM para θ'.

Tem-se ainda que: $E\left(\frac{T_1^*}{T_2^*}\right) \underset{X \perp Y}{=} E(T_1^*) \cdot E\left(\frac{1}{T_2^*}\right) = \theta \cdot \frac{1}{\theta'} = \frac{\theta}{\theta'}$

Então: $T^* = \frac{T_1^*}{T_2^*} = \frac{n+1}{n} X_{(n)} \cdot \frac{n-1}{n} Z_{(1)}$ será o ENVVUM para $\frac{\theta}{\theta'}$.

23/09/2010

20. Seja $(X_1, \dots, X_n)$, $n \geq 2$, uma amostra aleatória de tamanho n de uma distribuição discreta $P_{\theta,j}$, em que $\boxed{\theta \in (0,1)}$, $j = 1, 2$, sendo $P_{\theta,1}$ a distribuição de Poisson com média θ e $P_{\theta,2}$ a distribuição de Bernoulli de parâmetro θ.

(a) Mostre que $T = \sum_{i=1}^{n} X_i$ não é uma estatística suficiente.

$$P_{\theta,j}(x) = \begin{cases} P_{\theta,1}(x) = \dfrac{e^{-\theta}\,\theta^{x}}{x!} & x = 0,1,2,\dots \\ P_{\theta,2}(x) = \theta^{x}(1-\theta)^{1-x} & x = 0,1 \end{cases}$$

$$P(\underline{X} \mid \theta_j) = \left\{\prod_{i=1}^{n} \frac{\theta^{x_i}\, e^{-\theta}}{x_i!} I_{(0,1,2,\dots)}(x_i)\right\}^{2-j} \left\{\prod_{i=1}^{n} \theta^{x_i}(1-\theta)^{1-x_i} I_{(0,1)}(x_i)\right\}^{j-1} =$$

$$\left\{\prod_{i=1}^{n} \frac{1}{x_i!}\, \theta^{\sum x_i}\, e^{-n\theta} \prod_{i=1}^{n} I_{(0,1,2,\dots)}(x_i)\right\}^{2-j} \left\{\theta^{\sum x_i}(1-\theta)^{n-\sum x_i} \prod_{i=1}^{n} I_{(0,1)}(x_i)\right\}^{j-1}$$

Portanto, as estatísticas conjuntamente suficientes são:

$$\left(\sum_{i=1}^{n} X_i,\ \prod_{i=1}^{n} I_{(0,1,2,\dots)}(x_i),\ \prod_{i=1}^{n} I_{(0,1)}(x_i),\ \prod_{i=1}^{n} \frac{1}{x_i!}\right) = (T_1(X), T_2(X), T_3(X), T_4(X))$$

Note que a estatística $T = \sum X_i$ não é suficiente para definir em qual família de distribuição θ está sendo definido. Então, para que fosse suficiente, precisaríamos do par de estatísticas $J=1$ ou $J=2$ (para definir em qual distribuição θ está inserido), e $T = \sum_{i=1}^{n} X_i$, somente $T = \sum_{i=1}^{n} X_i$ não é suficiente para definir θ, ou seja, ele não é uma estatística suficiente para θ. (não reune todas as informações suficientes para substituir todas as informações sobre θ por ele.)

/ /

Fazendo pela definição de suficiência,

$T \sim Poisson(n\theta)$ e $T \sim Binomial(\theta, n)$

$$P(X \mid T=t) = \left\{ \frac{\frac{e^{-n\theta}}{\prod x_i!} \theta^{\sum x_i} \prod_{i=1}^{n} I_{\{0,1,2,\dots\}}(x_i)}{\frac{e^{-n\theta}}{\sum x_i!} (n\theta)^{\sum x_i} \prod_{i=1}^{n} I_{\{0,1,2,\dots\}}(x_i)} \right\}^{2-j} \left\{ \frac{\theta^{\sum x_i}(1-\theta)^{n-\sum x_i} \prod I_{\{0,1\}}(x_i)}{\binom{n}{\sum x_i} \theta^{\sum x_i}(1-\theta)^{1-\sum x_i} I_{\{0,1,\dots,n\}}(\sum x_i)} \right\}^{j-1}$$

que depende de j, que é um dos parâmetros conjuntamente para θ, identifica os elementos de P. Portanto $\sum X$ não é suficiente para θ.

Outra resolução: $(X_1, \dots X_n)$, $n \geq 2$ a.a. de $P_{\theta,j}$, $\theta \in (0,1)$, $j = 1, 2$, sendo
$P_{\theta_1}(x|\theta) = \frac{\theta^x e^{-\theta}}{x!}$, $x = 0,1,2,\dots$ e $P_{\theta,2}(X|\theta) = \theta^x (1-\theta)^{1-x}$, $x = 0,1,\dots$

Portanto: $\mathcal{P} = \{P_\phi, \ \phi = (\theta, j) : \theta \in (0,1), \ j \in \{1,2\}\}$

$$P(\underline{X} \mid \theta, j) = \left\{ \prod_{i=1}^{n} \frac{\theta^{x_i} e^{-\theta}}{x_i!} I_{\{0,1,2,\dots\}}(x_i) \right\}^{2-j} \left\{ \prod_{i=1}^{n} \theta^{x_i}(1-\theta)^{1-x_i} I_{\{0,1\}}(x_i) \right\}^{j-1} =$$

$$= \left\{ \frac{1}{\prod_{i=1}^{n} x_i!} \right\}^{2-j} \left\{ e^{-n\theta} \theta^{\sum_{i=1}^{n} x_i} \prod_{i=1}^{n} I_{\{0,1,2,\dots\}}(x_i) \right\}^{2-j} \left\{ \theta^{\sum x_i} (1-\theta)^{n-\sum x_i} \prod_{i=1}^{n} I_{\{0,1\}}(x_i) \right\}^{j-1}$$

Portanto, as estatísticas conjuntamente suficientes são: $\left(\sum_{i=1}^{n} X_i, \prod_{i=1}^{n} I_{\{0,1,2,\dots\}}(X_i), \prod_{i=1}^{n} I_{\{0,1\}}(X_i), \frac{1}{\prod_{i=1}^{n} X_i!} \right) = (T_1(X), T_2(X), T_3(X), T_4(X))$. As distribuições (Poisson e Binomial), apesar de compartilharem a estatística $T_1(X)$, possuem suportes diferentes e essa informação não está presente em $T_1(X)$ mas em $T_2(X)$ e $T_3(X)$. Portanto $\sum_{i=1}^{n} X_i$ não é suficiente para $\phi = (\theta, j)$, por não conter informação a respeito de j. De fato, calculando a densidade condicional em $T(X) = \sum_{i=1}^{n} X_i = t$ e considerando que $T \sim Poisson$, de média $n\theta$ para $j=1$ e $T \sim Binomial(\theta, n)$ para $j=2$, temos:

$$P_\phi(X \mid T=t) = \left\{ \frac{\frac{\cancel{e^{-n\theta}}}{\prod_{i=1}^{n} x_i!} \cancel{\theta^{\sum_{i=1}^{n} x_i}} \prod_{i=1}^{n} I_{\{0,1,2,\dots\}}(x_i)}{\frac{\cancel{e^{-n\theta}}}{(\sum_{i=1}^{n} X_i)!} (n\theta)^{\sum_{i=1}^{n} X_i} I_{\{0,1,2,\dots\}}(\sum X_i)} \right\}^{2-j} \left\{ \frac{\cancel{\theta^{\sum_{i=1}^{n} X_i}} (1-\theta)^{n-\sum x_i} \prod_{i=1}^{n} I_{\{0,1\}}(x_i)}{\binom{n}{\sum x_i} \cancel{\theta^{\sum_{i=1}^{n} X_i}} (1-\theta)^{n-\sum X_i} I_{\{0,1,\dots,n\}}(\sum X_i)} \right\}^{j-1} \quad (1)$$

que depende de j, que é um dos parâmetros que juntamente com θ identifica os elementos de $\mathcal{P}$. Portanto, $\sum X_i$ não é suficiente para $\phi = (\theta, j)$. Na expressão (1) supomos $I_{\{0,1,2,\dots\}}(\sum X_i)$ e $I_{\{0,1,\dots,n\}}(\sum X_i) \neq 0$.

(b) Encontre um ENVVUM de θ.

Consideremos o caso no qual $j=1$. Neste caso, a amostra segue uma distribuição de Poisson com média θ, ou seja:

$$p_{\theta,1}(x) = \prod_{i=1}^{n} f_{X_i}(x_i) = \prod_{i=1}^{n} \frac{e^{-\theta} \cdot \theta^{x_i}}{x_i!} = \exp\left\{-n\theta + \sum x_i \ln\theta - \ln \prod x_i!\right\} \quad x_i = 0,1,2,\ldots$$

Note-se que a distribuição pertence a família exponencial, com $\eta = \ln\theta$. O domínio da função η contém retângulos $k=1$ dimensionais, então a família é de posto completo e consequentemente, a estatística $T_1 = \sum_{i=1}^{n} X_i$ é estatística suficiente completa (e minimal). Como $\bar{X}$ é função 1 a 1 de T_1, então $T=\bar{X}$ também é estatística suficiente completa (e minimal) para θ.

Analogamente, para o caso onde $j=2$, a amostra segue uma distribuição de Bernoulli com média θ, ou seja,

$$p_{\theta,2}(x) = \prod_{i=1}^{n} f_{X_i}(x_i) = \prod_{i=1}^{n} \theta^{x_i}(1-\theta)^{1-x_i}; \quad x_i = 0,1.$$

$$= \exp\left\{\sum x_i \ln\theta + (n-\sum x_i)\ln(1-\theta)\right\} =$$

$$= \exp\left\{+\sum x_i \ln\left(\frac{\theta}{1-\theta}\right) + n\ln(1-\theta)\right\}, \quad x_i = 0,1$$

Novamente, a distribuição pertence a família exponencial, com $\eta = \ln\frac{\theta}{1-\theta}$, uma função que contém no seu domínio retângulos $k=1$ dimensionais (segmento de reta) e portanto, a família é de posto completo, e a estatística $T_2 = \sum_{i=1}^{n} X_i$ é estatística suficiente completa (e minimal) para θ. Como $\bar{X}$ é função 1 a 1 de T_2, então $T=\bar{X}$ é estatística suficiente completa (e minimal) para θ.

Como para $\forall j$, $j=1,2$, a estatística $T=\bar{X}$ é uma função de uma estatística suficiente e completa para θ, $T=\bar{X}$ é o ENVVUM para θ.

/ /

Outra forma:

Tanto na distribuição binomial $(j=2)$ quanto na Poisson $(j=1)$, considerados isoladamente, $T(X) = \sum X_i$ é estatística suficiente e completa (pelo fato das duas distribuições, isoladamente, pertencerem à uma família exponencial de posto completo) e $T(X) = \delta(X)$ é estimador não viciado de θ. Portanto, para $j=1$ e $j=2$, $\delta(X) = \frac{\sum_{i=1}^{n} X_i}{n}$ é ENNVUM para $\theta \Rightarrow \delta(X)$ é ENNVUM de θ em $\mathcal{P}$.

23/09/2010

21. Seja $X \sim N(\mu, \sigma^2)$. Mostre que não existe estimador não viciado de σ^2 quando μ é desconhecido. (Sugestão: para $\sigma^2 = a$ fixado, mostre que X é uma estatística suficiente completa para μ).

Para $\sigma^2 = a$ fixado, tem-se que:

$X \sim N(\mu, a)$ e sua fdp é da forma:

$$f_X(x) = \frac{1}{\sqrt{2\pi a}} \exp\left\{ -\frac{1}{2a}(x-\mu)^2 \right\} I_{\mathbb{R}}(x) =$$

$$= \exp\left\{ -\frac{1}{2a}(x^2 - 2x\mu + \mu^2) \right\} \frac{1}{\sqrt{2\pi a}} =$$

$$= \exp\left\{ +\frac{\mu}{a} x - \frac{\mu^2}{2a} - \frac{x^2}{2a} \right\} \frac{1}{\sqrt{2\pi a}}$$

$$\eta = \frac{\mu}{a} \qquad b(\eta) = +\frac{\eta^2 a^2}{2a} = \frac{\eta^2 a}{2}$$

$\mu = \eta a$

Observa-se que se tem uma única variável aleatória X, então faz sentido não se conseguir estimar a variância, uma vez que nem se pode falar de variabilidade

A distribuição pertence à família exponencial com $k=1$ parâmetros, na qual $\eta = \frac{\mu}{a}$ e seu domínio contém retângulos $k=1$ dimensionais, fazendo da família uma família de posto completo. Desse modo, a estatística associada para estimar μ; $T = X$ será uma estatística suficiente completa (e minimal).

Para mostrarmos que não existe um estimador não viciado de σ^2 quando μ é desconhecido, como tem-se só uma observação, então qualquer estimador de σ^2 é função de $T=X$. Por absurdo, vamos supor que existe um ENV de σ^2, então: $E_{\sigma^2}(T) = \sigma^2 \Rightarrow$

$$\Rightarrow E(T) - \sigma^2 = 0 \Rightarrow$$

$$\Rightarrow E(T - \sigma^2) = 0 \Rightarrow T \text{ é suficiente e completa } \xRightarrow{\forall \mu}$$

$$\Rightarrow T - \sigma^2 \equiv 0 \Rightarrow T = \sigma^2 \quad \text{ABSURDO!}$$

A estimativa não pode assumir o valor do estimando

$\therefore \nexists$ T. não viesado para σ^2 quando μ é desconhecido.

22/09/2010

22. Sejam $Z_1, \ldots, Z_n$ variáveis aleatórias iid com distribuição exponencial de média θ, isto é, com densidade $\theta^{-1} \exp(-z\theta^{-1})$, $z \geq 0$; $\theta > 0$. Sejam $x_1, \ldots, x_n$ constantes positivas. Suponha que $Y_1, \ldots, Y_n$ sejam uma amostra observável que segue o modelo:

$$Y_i = \beta x_i + Z_i, \quad i = 1, \ldots, n, \text{ em que } \beta > 0, \text{ sendo } \beta \text{ e } \theta \text{ desconhecidos.}$$

(a) Encontre o estimador de máxima verossimilhança $(\hat{\beta}, \hat{\theta})$ de (β, θ)

$$Y_i = \beta x_i + Z_i \qquad Z_i \sim E(\theta)$$

$$f_{Z_i}(z_i) = \frac{1}{\theta} \exp\left\{-\frac{z}{\theta}\right\} I_{(0,+\infty)}(z)$$

$$y_i - \beta x_i = z_i$$
$$dy_i = dz_i$$

$$f_{Y_i}(y_i) = \frac{1}{\theta} \exp\left\{-\frac{(y_i - \beta x_i)}{\theta}\right\} I_{(\beta x_i, +\infty)}(y_i)$$

$$f_{\underset{\sim}{Y}}(\underset{\sim}{y}) = \prod_{i=1}^{n} f_{Y_i}(y_i) = \prod_{i=1}^{n} \frac{1}{\theta} \exp\left\{-\frac{(y_i - \beta x_i)}{\theta}\right\} I_{(\beta x_i, +\infty)}(y_i) =$$

$$f_{\underset{\sim}{Y}}(\underset{\sim}{y}) = \frac{1}{\theta^n} \exp\left\{-\frac{\sum y_i}{\theta}\right\} \exp\left\{+\frac{\sum x_i \beta}{\theta}\right\} \prod_{i=1}^{n} I_{(\beta x_i, +\infty)}(y_i) \qquad \beta x_i < y_i < +\infty$$

$$< \beta < \frac{y_i}{x_i} < \infty$$

constantes positivas

Considerando β fixado, tem-se que:

$$f_{\underset{\sim}{Y}}(\underset{\sim}{y}) = \exp\left\{-\frac{\sum y_i}{\theta} + \frac{\sum x_i \beta}{\theta} - n \log \theta\right\}$$

$$l(\theta) = -\frac{\sum y_i}{\theta} + \frac{\sum x_i \beta}{\theta} - n \log \theta$$

(Supondo hessiana negativa definida)

$$\frac{dl(\theta)}{d\theta} = +\frac{\sum y_i}{\theta^2} - \frac{\sum x_i \beta}{\theta^2} - \frac{n}{\theta} = 0$$

$$\sum y_i - \sum x_i \beta - n\theta = 0 \Rightarrow \theta = \frac{\sum y_i}{n} - \frac{\sum x_i \beta}{n}$$

$$\boxed{\hat{\theta} = \bar{y}_i - \bar{x}\hat{\beta}}$$

Para θ fixado, tem-se que:

$$f_{\underset{\sim}{Y}}(\underset{\sim}{y}) = \frac{1}{\theta^n} \exp\left\{-\frac{\sum y_i}{\theta}\right\} \exp\left\{\frac{\sum x_i \beta}{\theta}\right\} I_{\left(\frac{y_{(1)}}{x_i}, +\infty\right)}(\beta)$$

$$\boxed{\hat{\beta} = \frac{y_{(1)}}{x_i} = \frac{1}{x_i} \min\{y_1, \ldots, y_n\}}$$

$\int u\,dv = uv - \int v\,du$

24/09/2010

(b) $\hat{\beta}$ e $\hat{\theta}$ são estimadores não viciados de β e de θ, respectivamente?

$\hat{\beta} = Y_{(1)} = z$

$F_z(z) = P(Z \leq z) = P(Y_{(1)} < z) = 1 - P(Y_{(1)} > z) = 1 - [P(Y_1 > z)]^n =$

$= 1 - [1 - P(Y_1 \leq z)]^n = 1 - \left[1 - \int_{\beta x_i}^{z} \frac{1}{\theta} e^{-\frac{(a_i - \beta x_i)}{\theta}} da\right]^n =$

$= 1 - \left[1 - \frac{e^{\frac{\beta x_i}{\theta}}}{\theta} \int_{\beta x_i}^{z} e^{-\frac{a_i}{\theta}} da\right]^n = 1 - \left[1 - \frac{e^{\frac{\beta x_i}{\theta}}}{\cancel{\theta}} \cdot \cancel{\theta} \cdot -e^{\frac{-a_i}{\theta}}\Big|_{\beta x_i}^{z}\right]^n =$

$= 1 - \left[1 - e^{\frac{\beta x_i}{\theta}} \cdot \left(-e^{-\frac{z_i}{\theta}} + e^{-\frac{\beta x_i}{\theta}}\right)\right]^n = 1 - \left[\cancel{1} + e^{\frac{\beta x_i - z_i}{\theta}} - \cancel{1}\right]^n \quad 1 - e^{\frac{n}{\theta}[\beta x_i - z_i]}$

$f_z(z) = -e^{\frac{n}{\theta}[\beta x_i - z_i]} \cdot \frac{-n}{\theta} = \frac{n}{\theta} e^{-\frac{n}{\theta}(z_i - \beta x_i)} I_{(\beta x_i, +\infty)}(z_i)$

$z = Y_{(1)} \sim E\left(\beta x_i, \frac{\theta}{n}\right)$

$E(\hat{\beta}) = E\left(\frac{y_1}{x_i}\right) = \frac{1}{x_i} E(y_1) = \frac{1}{x_i} \int_{\beta x_i}^{+\infty} y_i \frac{n}{\theta} e^{-\frac{n}{\theta}(y_i - \beta x_i)} dy_i =$

$y = u \qquad dv = \frac{n}{\theta} e^{-\frac{n}{\theta}(y_i - \beta x_i)} dy$

$dy = du \qquad v = \int \frac{n}{\theta} e^{-\frac{n}{\theta}y_i + \frac{n}{\theta}\beta x_i} dy = \frac{n}{\theta} e^{\frac{n}{\theta}\beta x_i} \int e^{-\frac{n}{\theta}y_i} dy = \frac{\cancel{n}}{\cancel{\theta}} e^{+\frac{n}{\theta}\beta x_i} \frac{\cancel{\theta}}{\cancel{n}} - e^{-\frac{n}{\theta}y_i} = -e^{\frac{n}{\theta}\beta x_i - \frac{n}{\theta}y_i}$

$= \frac{1}{x_i}\left[y \cdot -e^{\frac{n}{\theta}\beta x_i - \frac{n}{\theta}y_i}\Big|_{\beta x_i}^{+\infty} - \int -e^{\frac{n}{\theta}\beta x_i - \frac{n}{\theta}y_i} dy_i\right] =$

$= \frac{\beta x_i}{x_i} + \frac{1}{x_i} e^{\frac{n}{\theta}\beta x_i} \int_{\beta x_i}^{+\infty} e^{-\frac{n}{\theta}y_i} dy_i = \beta + \frac{1}{x_i} e^{\frac{n}{\theta}\beta x_i} \frac{\theta}{n} - e^{-\frac{n}{\theta}y_i}\Big|_{\beta x_i}^{+\infty} = \beta + \frac{\theta}{n x_i}$

$\hat{\beta}$ é viciado para β, então $T = Y_{(1)} - \frac{\hat{\theta}}{n x_i}$ é não viciado para β.

$E(\hat{\theta}) = E(\bar{y} - \bar{x}\hat{\beta}) = E(\bar{y}) - \bar{x}\beta + \frac{\bar{x}\theta}{n x_i} = E\left(\frac{\sum y_i}{n}\right) - \bar{x}\beta + \frac{\bar{x}\theta}{n x_i} =$

$= \frac{E(\sum y_i)}{n} - \bar{x}\hat{\beta} + \frac{\bar{x}\theta}{n x_i} = \cancel{\bar{x}\beta} - \cancel{\bar{x}\beta} + \frac{\bar{x}\theta}{n x_i} \quad \frac{\bar{x}\theta}{n x_i}$

$\hat{\theta}$ é viciado para θ, então $T^* = \frac{n x_i}{\bar{x}}(\bar{y} - \bar{X}\hat{\beta})$ é não viciado para θ.

24/09/2010

c) Obtenha a distribuição, a média e a variância do estimador de máxima verossimilhança de β.

$w = \frac{z_i}{x_i} \Rightarrow w = \frac{\beta x_i}{x_i}$

$Z_i = Y_{(1)} \sim E\left(\beta x_i, \frac{\theta}{n}\right) \qquad \hat{\beta} = \frac{Y_{(1)}}{x_i}$

$W = \frac{Y_{(1)}}{x_i} = \frac{Z_i}{x_i} \qquad z_i = w \cdot x_i \qquad \frac{dz_i}{dw} = x_i$

$dz_i = dw \cdot x_i$

$f_{W_i}(w) = \frac{n}{\theta} e^{-\frac{n}{\theta}(x_i w_i - \beta x_i)} \cdot x_i = \frac{n x_i}{\theta} e^{-\frac{n x_i}{\theta}(w_i - \beta)} I_{(\beta, \infty)}(w_i)$

$W = \frac{Y_{(1)}}{x_i} \sim E\left(\beta, \frac{\theta}{n x_i}\right)$

$E(W) = \int_{\beta}^{+\infty} w \cdot \frac{n x_i}{\theta} e^{-\frac{n x_i}{\theta}(w-\beta)} dw = w \cdot -e^{-\frac{n x_i}{\theta}(w-\beta)} \Big|_{\beta}^{\infty} - \int_{\beta}^{+\infty} -e^{-\frac{n x_i}{\theta}(w-\beta)} dw =$

$u = w \qquad dv = \frac{n x_i}{\theta} e^{-\frac{n x_i}{\theta}(w-\beta)} dw$

$du = dw$

$v = \int \frac{n x_i}{\theta} e^{-\frac{n x_i}{\theta}(w-\beta)} dw = \cancel{\frac{n x_i}{\theta}} \cdot \cancel{\frac{\theta}{n x_i}} - e^{-\frac{n x_i}{\theta}(w-\beta)}$

$= \beta + \frac{\theta}{n x_i} - e^{\frac{n x_i}{\theta}(w-\beta)} \Big|_{\beta}^{-\infty} = \beta + \frac{\theta}{n x_i}$

$E(W^2) = \int_{\beta}^{+\infty} w^2 \cdot \frac{n x_i}{\theta} e^{-\frac{n x_i}{\theta}(w-\beta)} dw = w^2 \cdot -e^{-\frac{n x_i}{\theta}(w-\beta)} \Big|_{\beta} - \int_{\beta}^{+\infty} 2w \; -e^{-\frac{n x_i}{\theta}(w-\beta)} dw =$

$u = w^2 \qquad dv = n x_i \, e^{-\frac{n x_i}{\theta}(w-\beta)} dw$

$du = 2w\,dw \qquad v = -e^{\theta - \frac{n x_i}{\theta}(w-\beta)}$

$= \beta^2 + 2 \cdot \frac{\theta}{n x_i} \int_{\beta}^{+\infty} w \cdot \frac{n x_i}{\theta} e^{-\frac{n x_i}{\theta}(w-\beta)} dw = \beta^2 + 2 \frac{\theta}{n x_i}\left(\beta + \frac{\theta}{n x_i}\right) = \beta^2 + \frac{2\beta\theta}{n x_i} + \frac{2\theta^2}{n^2 x_i^2}$

$Var(W) = EW^2 - [EW]^2 =$

$= \beta^2 + \frac{2\beta\theta}{n x_i} + \frac{2\theta^2}{n^2 x_i^2} - \left(\beta + \frac{\theta}{n x_i}\right)^2 = \cancel{\beta^2} + \cancel{\frac{2\beta\theta}{n x_i}} + \frac{2\theta^2}{n^2 x_i^2} - \cancel{\beta^2} - \cancel{\frac{2\beta\theta}{n x_i}} - \frac{\theta^2}{n^2 x_i^2} = \frac{\theta^2}{n^2 x_i^2}$

$E(W) = \beta + \frac{\theta}{n x_i} \qquad Var(W) = \frac{\theta^2}{n^2 x_i^2}$

não consigo usar os teoremas feito no exercício (16)

$A \ne$ $a_1=+1$ $a_2=2$ não consigo mostrar para este teorema

23. Sejam $X_1, \dots, X_n$, n observações independentes de uma distribuição exponencial $E(a,b)$, $a \in (-\infty, 0]$ e $b>0$ é conhecido.

(a) Mostre que $X_{(1)} = \min\limits_{i \in \{1,\dots,n\}} X_i$ é uma estatística suficiente, mas não é completa.

i) Suficiência

$X_i \sim E(a,b)$, $a \in (-\infty, 0]$, $b>0$ conhecido, então

$$f_{X_i}(x_i) = \frac{1}{b} \exp\left\{-\frac{(x_i-a)}{b}\right\} I_{[a,+\infty)}(x_i)$$

Então:

$$p_X(x) = \prod_{i=1}^{n} f_{X_i}(x_i) = \prod_{i=1}^{n} \frac{1}{b} \exp\left\{-\frac{(x_i-a)}{b}\right\} I_{[a,+\infty)}(x_i) =$$

$$= \frac{1}{b^n} \exp\left\{-\sum_{i=1}^{n} \frac{(x_i-a)}{b}\right\} \prod_{i=1}^{n} I_{(a,+\infty)}(x_i) =$$

$a \le x_i < +\infty$

$$= \frac{1}{b^n} \exp\left\{-\frac{\sum x_i}{b} + \frac{na}{b}\right\} I_{[a,+\infty)}(x_{(1)}) = \underbrace{\exp\left\{\frac{na}{b}\right\} I_{[a,+\infty)}(x_{(1)})}_{g_a(T(x))} \underbrace{\frac{1}{b^n} \exp\left\{-\frac{\sum x_i}{b}\right\}}_{h(x)}$$

Pelo Critério da fatoração, tem-se que

$$g_a(T(X)) = \exp\left\{\frac{na}{b}\right\} \cdot I_{[a,+\infty)}(x_{(1)}) \quad \text{e} \quad h(x) = \frac{1}{b^n} \exp\left\{-\frac{\sum x_i}{b}\right\}$$

Desse modo, $T = X_{(1)} = \min\limits_{i \in \{1,\dots,n\}} X_i$ é estatística suficiente para a.

b) Completitividade. Seja $Y = X_{(1)}$, então:

$$F_Y(y) = P(Y \le y) = P(X_{(1)} \le y) = 1 - P(X_{(1)} > y) =$$

$$= 1 - [P(X_1 > y, X_2 > y, \dots, X_n > y)] = 1 - [P(X_1 > y)]^n =$$

$$= 1 - \left[\int_y^{+\infty} \frac{1}{b} \exp\left\{-\frac{(z-a)}{b}\right\} dz\right]^n = 1 - \left[\frac{e^{\frac{a}{b}}}{b} \int_y^{+\infty} e^{-\frac{z}{b}} dz\right]^n = 1 - \left[\frac{e^{\frac{a}{b}}}{\cancel{b}} \cdot \cancel{b} \cdot e^{-\frac{z}{b}}\right]_y$$

$$= 1 - e^{n\left[\frac{a}{b} - \frac{y}{b}\right]} = 1 - e^{-\frac{n}{b}(y-a)}$$

$\int u\,dv = uv - \int v\,du$

25/09/2010

$$f_Y(y) = \frac{dF_Y(y)}{dy} = \frac{n}{b} e^{-\frac{n}{b}(y-a)} I_{[a,+\infty)}(y)$$

$$X_{(1)} = y \sim E\left(a, \frac{b}{n}\right)$$

Seja $g(T) = \begin{cases} 0 & \text{se } T \leq 0 \\ \frac{n}{b} - e^{-t\left(1-\frac{n}{b}\right)} & T > 0 \end{cases}$

$$E(g(t)) = \int_0^{+\infty} \left(\frac{n}{b} - e^{-t\left(1-\frac{n}{b}\right)}\right) \cdot \frac{n}{b} e^{-\frac{n}{b}(t-a)}\, dt =$$

$$= \frac{n}{b}\int_0^{+\infty} \frac{n}{b} e^{-\frac{n}{b}(t-a)}\, dt - \int_0^{+\infty} \frac{n}{b} e^{-t+t\cancel{\frac{n}{b}}-\cancel{\frac{n}{b}t}+\frac{na}{b}}\, dt =$$

$$= \frac{n}{b} e^{\frac{na}{b}} \int_0^{+\infty} \frac{n}{b} e^{-\frac{nt}{b}}\, dt - \frac{n}{b} e^{\frac{na}{b}} \int_0^{+\infty} e^{-t}\, dt =$$

$$= \frac{n}{b} e^{\frac{na}{b}} \Bigg[\underbrace{\int_0^{+\infty} \frac{n}{b} e^{-\frac{nt}{b}}\, dt}_{\substack{T \sim E\left(\frac{b}{n}\right) \\ \text{fdp}}} - \underbrace{\int_0^{+\infty} e^{-t}\, dt}_{\substack{T \sim E(1) \\ \text{fdp}}} \Bigg] = \frac{n}{b} e^{\frac{na}{b}} [1-1] =$$

$$= \frac{n}{b} e^{\frac{na}{b}}\, 0 = 0 \qquad g(T) = \begin{cases} 0 & \text{se } T \leq 0 \\ \frac{n}{b} - e^{-T\left(1-\frac{n}{b}\right)} & T > 0 \end{cases} \qquad E(g(T)) = 0 \not\Rightarrow g(T) \equiv 0$$

$\therefore T = X_{(1)}$ não é completa

Outro modo: Seja $h(X_{(1)})$ tal que $E_a[h(X_{(1)})] = 0$

$$E_a[h(X_{(1)})] = 0 \Leftrightarrow \int_a^0 h(x)\, f_{X_{(1)}}(x|a)\, dx + \int_0^{+\infty} h(x)\, f_{X_{(1)}}(x|a)\, dx = 0 \Leftrightarrow$$

$$\cancel{\frac{n}{b}} e^{\frac{na}{b}} \int_a^0 h(x)\, e^{-\frac{nx}{b}}\, dx + \cancel{\frac{n}{b}} e^{\frac{na}{b}} \int_0^{+\infty} h(x)\, e^{-\frac{nx}{b}}\, dx = 0$$

Se tomamos $\begin{cases} 0 & \text{se } X < 0 \\ X - \frac{b}{n} & \text{se } X \geq 0 \end{cases}$ então

$$E_a[h(X_{(1)})] = \int_a^0 0 \cdot f_a(x|a)\, dx + \underbrace{\int_0^{\infty} x\, f_a(x|a)\, dx - \int_0^{\infty} \frac{b}{n} f_a(x|a)\, dx}_{EX \quad X \sim E\left(a, \frac{b}{n}\right)} = 0$$

$E_{(a)}[h(X_{(1)})] = 0$ e $g(X_{(1)}) \neq 0$.

$\frac{b}{n} \qquad - \frac{b}{n} \times 1$ derivada $\quad 0$

25/09/2010

(b) Encontre o ENVVUM de a (sugestão: Considere $g(X_{(1)}) = (c X_{(1)} + d) I_{(-\infty,0]}(X_{(1)})$, com c e d constantes. Mostre que $g(X_{(1)})$ é não correlacionado com os estimadores não viciados de zero.

[Procuramos um estimador não viciado de a, função da estatística suficiente $X_{(1)}$ e que seja não correlacionado com estimadores de zero na família $E[a,b]$.

Seja $g(X_{(1)}) = (c X_{(1)} + d) \cdot I_{(-\infty,0]}(X_{(1)})$

$$E(g(X_{(1)})) = E\left[c X_{(1)} + d\right] = \int_a^0 (cx+d) \cdot \underbrace{\frac{n}{b} e^{-\frac{n}{b}(x-a)}}_{\text{densidade de } X_{(1)}} dx =$$

$$= e^{\frac{na}{b}} \left[c \int_a^0 x \frac{n}{b} e^{-\frac{n}{b}x} dx + d \int_a^0 \frac{n}{b} e^{-\frac{nx}{b}} dx \right] = (*)$$

$$\int_a^0 x \frac{n}{b} e^{-\frac{n}{b}x} dx = x \cdot -e^{-\frac{nx}{b}} \Big|_a^0 - \int_a^0 -e^{-\frac{n}{b}x} dx = a e^{-\frac{na}{b}} + \int_a^0 e^{-\frac{nx}{b}} dx = a e^{-\frac{na}{b}} + \frac{b}{n} - e^{-\frac{n}{b}}$$

$u = x \qquad dv = \frac{n}{b} e^{-\frac{n}{b}x} dx$

$du = dx \qquad v = \int \frac{n}{b} e^{-\frac{nx}{b}} dx = \frac{n}{b} \int e^{-\frac{n}{b}x} dx = \frac{\cancel{n}}{\cancel{b}} \frac{\cancel{b}}{\cancel{n}} \cdot -e^{-\frac{n x}{b}} = -e^{-\frac{n}{b}x}$

$$= a e^{-\frac{na}{b}} - \frac{b}{n} + \frac{b}{n} e^{-\frac{na}{b}}$$

$$\int_a^0 \frac{n}{b} e^{-\frac{nx}{b}} dx = \frac{\cancel{n}}{\cancel{b}} \cdot \frac{\cancel{b}}{\cancel{n}} \cdot -e^{-\frac{nx}{b}} \Big|_a^0 = -1 + e^{-\frac{na}{b}}$$

$$= (*) = e^{\frac{na}{b}} \left[c \left(a e^{-\frac{na}{b}} - \frac{b}{n} + \frac{b}{n} e^{-\frac{na}{b}} \right) + d \left(-1 + e^{-\frac{na}{b}} \right) \right] =$$

$$= E(g(X_{(1)})) = a \iff c = 1 \quad d = -\frac{b}{n}$$

$$= g((X_{(1)})) = \left(X_{(1)} - \frac{b}{n} \right) I_{(-\infty,0]}(X_{(1)})$$

Para provarmos que $g(X_{(1)}) = \left(X_{(1)} - \frac{b}{n} \right) I_{(-\infty,0]}(X_{(1)})$ é ENVVUM de a, basta provarmos que ele é não correlacionado com qualquer estimador não viciado de zero. Pela redução pela suficiência, basta procurarmos por estimadores que sejam

25/09/2010

função de $g(X_{(1)})$.

Assim, tomando $y = X_{(1)}$

$E_a[h(y)] = 0$

$$0 = \int_a^{+\infty} h(y) \cdot f(y|a)\, dy = \int_a^0 h(y)\, f(y|a)\, dy + \int_0^{\infty} h(y)\, f(y|a)\, dy \Rightarrow$$

$$-\int_a^0 h(y)\, f(y|a)\, dy = \int_0^{\infty} h(y)\, f(y|a)\, dy$$

$$-\int_0^a h(y) \cdot \frac{n}{b}\, e^{-\frac{n}{b}(y-a)}\, dy = \int_0^{\infty} h(y)\, \frac{n}{b}\, e^{-\frac{n}{b}(y-a)}\, dy$$

$$\cancel{\frac{n}{b}\, e^{\frac{na}{b}}} \int_0^a h(y)\, e^{-\frac{ny}{b}}\, dy = \cancel{\frac{n}{b}\, e^{\frac{na}{b}}} \int_0^{+\infty} h(y)\, e^{-\frac{ny}{b}}\, dy$$

$$\underbrace{\int_0^a h(y)\, e^{-\frac{ny}{b}}\, dy}_{\text{função de } a} = \underbrace{\int_0^{+\infty} h(y) \cdot e^{-\frac{ny}{b}}\, dy}_{\text{constante em relação a } a}$$

$$\frac{d}{da} \int_0^a h(y)\, e^{-\frac{ny}{b}}\, dy = \frac{d}{da} \int_0^{+\infty} h(y)\, e^{-\frac{ny}{b}}\, dy$$

$$h(a)\, e^{-\frac{na}{b}} = 0 \iff h(a) = 0, \ \forall a \in (-\infty, 0] \text{ e } \int_0^{+\infty} h(x)\, f_a(y|a)\, dy = 0$$

Assim, os Estimadores não viciados de 0 são funções de $X_{(1)}$ tais que são identicamente nulos $\forall X_{(1)} \leq 0$ e $\int_0^{+\infty} h(x)\, e^{-\frac{nx}{b}}\, dx = 0$.
Entretanto $g(X_{(1)}) = 0$ $\forall (X_{(1)}) > 0$. Portanto $h(X_{(1)}) \times g(X_{(1)}) = 0$.

Desse modo,

$$E_a[h(X_{(1)}) \cdot g(X_{(1)})] = E_a[0] = 0 = E_a[h(X_{(1)})] \times E[g(X_{(1)})] = 0 \quad a = a$$

Portanto, $g(X_{(1)})$ é o ENVVUM de a

24. Seja X uma variável aleatória com função de probabilidade:

$$p_\theta(x) = P_\theta(X=x) = \frac{a(x)\,\theta^x}{C(\theta)}, \quad x = 0, 1, \dots$$

$\theta > 0$, $a(x) \geq 0$, $\sum_{x=0}^{+\infty} a(x)\,\theta^x = C(\theta) < \infty$ (distribuição de série de potências

(a) Mostre que esta distribuição faz parte da família exponencial unidimensional e mostre que a função geradora de momentos de X é $M_X(u) = \frac{C(\theta e^u)}{C(\theta)}$.

$$p_\theta(x) = P_\theta(X=x) = \frac{a(x)\,\theta^x}{C(\theta)} = \exp\left\{\log \frac{\theta^x}{C(\theta)}\right\} a(x) =$$

$$= \exp\{x \cdot \log\theta - \log C(\theta)\}\, a(x).$$

$\eta = \log\theta$ $\quad T = X \quad B(\eta) = \log C(e^\eta) \quad c(x) = a(x)$

$\theta = e^\eta$

∴ a distribuição pertence a família exponencial a um parâmetro.

$$M_X(u) = \frac{\exp(A(\eta+u))}{\exp(A(\eta))}$$

$$M_X(u) = \frac{\exp(B(\eta+u))}{\exp(B(\eta))} = \frac{e^{\log C(e^{\eta+u})}}{e^{\log C(e^\eta)}} = \frac{C(e^{\eta+u})}{C(e^\eta)} = \frac{C(e^\eta \cdot e^u)}{C(e^\eta)} \overset{e^\eta=\theta}{=}$$

$$= \frac{C(\theta e^u)}{C(\theta)}$$

$$\boxed{M_X(u) = \frac{C(\theta e^u)}{C(\theta)}}$$

(b) Mostre que se $a(x) > 0$ para todo $x = 0, 1, \ldots$, então θ^r é estimável para qualquer r inteiro positivo e seu estimador não viciado de variância uniformemente mínima baseado em uma amostra $X_1, \ldots, X_n$ de observações independentes da distribuição de X é $\delta(T)$ sendo $T = \sum_{i=1}^{n} X_i$ e

$$\delta(t) = \begin{cases} 0 & t = 0, \ldots, r-1, \\ \dfrac{A(t-r, n)}{A(t, n)} & t = r, r+1, \ldots \end{cases}$$

onde $A(t,n)$ é o coeficiente de θ^t na expansão em série de $[C(\theta)]^n$.

Se T é estatística suficiente completa, o ENVVUM de $\forall$ função estimável $g(\theta)$ é unicamente determinado por: $E_\theta[\delta(T)] = g(\theta)$ para todo $\theta \in \Omega$.

Então

$$E_\theta[\delta(T)] = \theta^r$$

$$\sum_{x=0}^{+\infty} \frac{\delta(x) . a(x) . \theta^x}{C(\theta)} = \theta^r$$

$$\sum_{x=0}^{+\infty} \delta(x)\, a(x)\, \theta^x = C(\theta) . \theta^r$$

Pelo enunciado temos que $C(\theta) = \sum_{x=0}^{+\infty} a(x)\, \theta^x$, então

$$\sum_{x=0}^{+\infty} \delta(x) . a(x)\, \theta^x = \sum_{x=0}^{+\infty} a(x) . \theta^x . \theta^r$$

$$\sum_{x=0}^{+\infty} \delta(x)\, a(x) . \theta^x = \sum_{x=0}^{+\infty} a(x) . \theta^{x+r}$$

$k = x + r \Rightarrow x = k - r$ e assim:

$$\sum_{x=0}^{+\infty} \delta(x)\, a(x)\, \theta^x = \sum_{k=r}^{+\infty} a(k-r) . \theta^k$$

$$\sum_{x=0}^{r-1} \delta(x) . a(x) . \theta^x + \sum_{x=r}^{+\infty} \delta(x)\, a(x) . \theta^x = \sum_{k=r}^{+\infty} a(k-r) . \theta^k$$

25/09/2010

Daí:

$\delta(x)\,.\underbrace{a(x)}_{>0 \text{ hipótese}} = 0 \quad , \quad x = 0, 1, \ldots, r-1$

$\delta(x) = 0 \quad$ quando $\quad x = 0, 1, \ldots, r-1$

e

$\delta(x)\,.\,a(x) = a(x-r) \quad$ quando $\quad x = r, r+1, \ldots$

$\delta(x) = \dfrac{a(x-r)}{a(x)}$

ou seja, $\delta(x) = \begin{cases} \dfrac{a(x-r)}{a(x)} & \text{se } x = r, r+1, \ldots \\ 0 & \text{se } x = 0, \ldots, r-1 \end{cases}$

$\therefore \theta^r$ é estimável, $\forall r \in \mathbb{Z}^+$

Como a distribuição de X pertence a uma família exponencial, então $T(X) = \sum_{i=1}^{n} X_i$ é uma estatística suficiente completa pois $\Omega = \{\theta \in \mathbb{R} : \theta > 0\}$ contém retângulos unidimensionais e com isso a família é de posto completo.

Tem-se ainda que:

$$P(T = t) = \frac{A(t, n)\,\theta^t}{[C(\theta)]^n}$$

onde $A(t, n)$ é o coeficiente de θ^t na expansão de série de potências de $[C(\theta)]^n$

Assim, análogo ao caso de uma única observação X temos que para $X_1, \ldots, X_n$

$$\delta(T) = \begin{cases} 0 & \text{se } T = 0, 1, \ldots, r-1 \\ \dfrac{A(T-r, n)}{A(T)} & \text{se } T = r, r+1, \ldots \end{cases}$$ é ENV de θ^r

$\delta(T)$ é ENVVUM de θ^r.

25/09/2010

(c) Se $X_1, \ldots, X_n$ são observações importantes da distribuição de Poisson de média θ, $\theta > 0$, utilize o resultado em (b) para encontrar o ENVVUM de θ^r, r inteiro positivo.

Temos que

$$p_\theta(x) = \frac{\theta^x e^{-\theta}}{x!} = \frac{(x!)^{-1}\,\theta^x}{e^\theta}, \quad x = 0, 1, \ldots$$

Fazendo $a(x) = (x!)^{-1}$ e $c(\theta) = e^\theta = \sum_{x=0}^{+\infty} \frac{\theta^x}{x!}$ $\quad x = 0, 1, \ldots$

então

$$p_\theta(x) = \frac{a(x) \cdot \theta^x}{C(\theta)}$$

A distribuição de X pertence a família exponencial. Então $T(X) = \sum X_i$ é estatística suficiente. Além disso, $T(X) = \sum X_i$ é completa pois $\Omega = \{\theta \in \mathbb{R} : \theta > 0\}$ contém retângulos unidimensionais e com isso a família é de posto completo.

Pelo item (b), temos que:

$$\delta(T) = \begin{cases} 0 & \text{se } T = 0, \ldots, r-1, \\ \dfrac{A(T-r, n)}{A(T, n)} & \text{se } T = r, r+1, \ldots \end{cases}$$

onde $A(T,n)$ é o coeficiente de θ^t na expansão em série de $[C(\theta)]^n$, é o ENVVUM de θ^r

Como $T = \sum X_i \sim \text{Poisson}(n\theta)$ então

$$p_\theta(t) = \frac{(n\theta)^t e^{-n\theta}}{t!} = \frac{(t!)^{-1} n^t \theta^t}{e^{n\theta}} =$$

Então $A(t,n) = (t!)^{-1} n^t$ e $[C(\theta)]^n = [e^\theta]^n$

$$p_\theta(t) = \frac{A(t,n)\,\theta^t}{[C(\theta)]^n}$$

25/09/2010

Daí:

$A(t-r, n) = [(t-r)!]^{-1} \, n^{t-r}$ e então

$$S(t) = \begin{cases} 0 & t = 0, \ldots, r-1 \\ \dfrac{t!}{(t-r)!} \dfrac{1}{n^r} & t = r, r+1, \ldots \end{cases}$$

Logo, $S(T)$ é ENVVUM para θ^r, $r \in \mathbb{Z}^+$

d) Mostre que a distribuição de Poisson truncada à direita em b ($b \geq 1$, inteiro) e parâmetro θ é um caso especial da distribuição de séries de potências. (se X tem distribuição de Poisson truncada à direita em b e parâmetro θ, então $P_\theta(X = x) = P_\theta(Y = x \mid Y \leq b)$, $x = 0, 1, \ldots b$, onde Y tem distribuição de Poisson de parâmetro θ.

Como $Y \sim P(\theta)$, então

$$P(Y \leq b) = \sum_{y=0}^{b} \frac{\theta^y e^{-\theta}}{y!}$$

Tem-se que

$$P_\theta(X = x) = P_\theta(Y = x \mid Y \leq b) \quad x = 0, 1, \ldots, b$$

$$= \frac{\theta^x \cdot e^{-\theta}}{x!} \frac{1}{P(Y \leq b)} \quad x = 0, 1, \ldots b$$

$$= \frac{\theta^x \cancel{e^{-\theta}}}{x!} \frac{1}{\sum_{y=0}^{b} \frac{\theta^y \cancel{e^{-\theta}}}{y!}}$$

$$= \frac{\theta^x}{x!} \left[\sum_{y=0}^{b} \frac{\theta^y}{y!} \right]^{-1} \quad x = 0, 1, \ldots, b$$

Fazendo $a(x) = (x!)^{-1}$ e $C(\theta) = \sum_{y=0}^{b} \frac{\theta^y}{y!}$ onde $x = 0, 1, \ldots, b$, tem-se que a distribuição de Poisson truncada à direita e parâmetro θ é um caso especial de série de potências.

25/09/2020

(e) Mostre que se X é uma única observação de uma distribuição de Poisson truncada à direita em b ($b \geq 1$, inteiro), então θ não é estimável. Esta afirmação contradiz à do item (b)? Por que?

Sabe-se que g é estimável se existe um estimador não viciado de g. Assim, para mostrarmos que θ não é estimável, basta mostrarmos que não existe um estimador não viciado de θ.

Pelo item (d) tem-se que

$$P_\theta(X=x) = \frac{\theta^x}{x!}\left(\sum_{y=0}^{b}\frac{\theta^y}{y!}\right)^{-1} \quad x=0,1,\dots,b$$

então $E_\theta[\delta(X)] = \theta \iff$

$$\sum_{x=0}^{b}\delta(x)\frac{\theta^x}{x!}\left(\sum_{y=0}^{b}\frac{\theta^y}{y!}\right)^{-1} = \theta$$

$$\sum_{x=0}^{b}\delta(x)\frac{\theta^x}{x!} = \theta.\sum_{y=0}^{b}\frac{\theta^y}{y!}$$

$$\sum_{x=0}^{b}\delta(X)\frac{\theta^x}{x!} = \sum_{y=0}^{b}\frac{\theta^{y+1}}{y!}$$

$$\sum_{x=0}^{b}\delta(X)\frac{\theta^x}{x!} = \sum_{y=0}^{b}\frac{(y+1)\,\theta^{y+1}}{(y+1)!}$$

Fazendo $u=y+1$, se $y=0 \Rightarrow u=1$, se $y=b \Rightarrow u=b+1$ e então

$$\sum_{x=0}^{b}\delta(X).\frac{\theta^x}{x!} = \sum_{u=1}^{b+1}\frac{u.\theta^u}{u!} = \sum_{u=1}^{b+1}\frac{u\,\theta^u}{u!}+0$$

$$\sum_{x=0}^{b}\delta(X).\frac{\theta^x}{x!} = \sum_{u=0}^{b+1}\frac{u\,\theta^u}{u!} \quad (*)$$

Observamos que do lado esquerdo de $(*)$ há um polinômio

25/09/2010

de grau $b+1$ em θ enquanto que do lado direito há um polinômio de grau $b+2$ em θ.

$$\therefore \nexists \; \delta(X) \text{ que satisfaça } (*)$$

$\therefore \nexists$ ENV de θ, então θ não é estimável

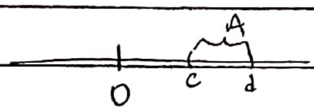
No item (d) vimos que a distribuição de Poisson truncada à direita de b e parâmetro θ é um caso especial da distribuição de série de potências onde $a(x) = (x!)^{-1} > 0$. Logo, isto contradiz a afirmação do item (b):

"Se $a(X) > 0 \;\forall\; x = 0, 1, \ldots$, então θ^r é estimável para $\forall r \in \mathbb{Z}^+$."

Borelianos

$a \in \mathbb{R}$

$$0 = \int_a^{\infty} h(x)\, f(x)\, dx = 0$$



$$0 = \int_a^c + \int_c^d + \int_d^{\infty}$$

$\forall a \in \mathbb{R}$

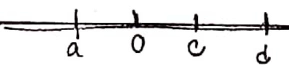
$$\int_c^{\infty} = \int_c^d + \int_d^{\infty}$$

$$0 \quad\quad 0 \quad\quad 0$$

$$\int_a^{\infty} = \int_a^c + \int_a^{+\infty}$$

$$0 \quad\quad 0$$

$$\int_a^{+\infty} h(x)\, f(x)\, dx = 0$$



$-\infty < a < 0$

$$0 = \int_a^{\infty} = \int_a^0 + \int_0^{\infty}$$

$$0$$

$$\int_0^{\infty} = \int_0^c + \int_c^d + \int_d^{\infty} = 0$$

não consigo usar o argumento para $a > 0$, não posso argumentar desse modo, só posso fazer isso até o zero.

/ /.

funciona

$$0 = \int_a^{\infty} h(x)\, f(x)\, dx$$

$h(x) \geq 0$ q.c.en P

$a \geq 0$ $\qquad P(h(x)=0) = 1$

0 c d

$$0 = \int_a^{\infty} = \int_a^{c} + \int_c^{d} + \underbrace{\int_d^{\infty}}_{0}$$

$$\int_c^{\infty} = \int_c^{d} + \int_d^{\infty}$$

$\int_c^{\infty} \to 0$, $\int_c^{d} \to 0$, $\int_d^{\infty} \to 0$

Lista 2 (24 b)

$(n-1) \supset$ θ^2

$(n-2) \supset$ $\theta, \theta, \ldots /$

b). $\left(\sum_{x=0}^{n} a(x)\,\theta^x\right)^n = \underbrace{[a(0)]^n \cdot \theta^0}_{P(\sum X_i = 0)} + \underbrace{C_n^1 [a(0)]^{n-1} \cdot a(1)\,\theta^1}_{P(\sum X = 1)} + \underbrace{\ldots}_{P(\sum X = 2)} \quad \theta^2$

$[C(\theta)]^n$

$(a(0)\,\theta^0 + a(1)\,\theta^1 + \ldots)\,(a(0)\,\theta^0 + a(1)\cdot\theta^1 + \ldots)\,(a(0)\cdot\theta^0 + a(1)\,\theta^1 + \ldots)$

$\theta^r = \sum_{t=0}^{+\infty} \delta(t)\; P(T=t) \quad ; \quad T = \sum_{i=1}^{n} X_i \quad$ (método 1)

$= 0 + \frac{P(0)}{P(r)} \cdot P(r) \cdot \theta^r + \frac{P(1)}{P(r+1)} \; P(r+1) \cdot \theta^{r+1} + \ldots =$

$= \theta^r \left(\underbrace{P(0) + P(1) \cdot \theta + P(2)\,\theta^2 + \ldots}_{=1} \right) = \theta^r$

$r = 5$

$\delta(0) = \delta(1) = \delta(2) = \delta(3) = \delta(4) = 0$

$\delta(5) = \frac{A(5-5)}{A(5)} = \frac{A(0)}{A(5)}\,\theta^5$

$\delta(6) = \frac{A(6-5)}{A(6)} = \frac{A(1)}{A(6)}\,\theta^6$

$\delta(7) = \frac{A(7-5)}{A(7)} = \frac{A(2)}{A(7)}\,\theta^7$

Tatiane FNMS.

MAE 5834 - Estatística Avançada I

Lista de Exercícios 2 – 2º semestre de 2005 – Prof. Silvia L.P. Ferrari

ok 1. Seja $X = (X_1, \ldots, X_n)$ um vetor de variáveis aleatórias independentes tais que $X_i \sim \mathrm{N}(d_i\theta, \theta^2)$, $d_i \neq 0$, para $i = 1, \ldots, n$, $\theta > 0$. Sejam

$$\tilde{X} = \frac{\sum_{i=1}^n d_i X_i}{\sum_{i=1}^n d_i^2} \quad \text{e} \quad S^2 = \sum_{i=1}^n \frac{(X_i - d_i\tilde{X})^2}{n-1}.$$

Mostre que entre todos os estimadores não viciados de θ da forma $\sum_{i=1}^n v_i X_i$, com $v_i \in (-\infty, +\infty)$ para $i = 1, \ldots, n$, $\tilde{X}$ é o que tem variância uniformemente mínima.

ok 2. Prove ou dê contra-exemplo. Seja X variável aleatória com distribuição P_θ pertencente a uma família de distribuições $\mathcal{P} = \{P_\theta,\ \theta \in \Omega\}$ e sejam $g_1(\theta)$ e $g_2(\theta)$ funções de θ a valores reais. Sejam $\delta_1(X)$ e $\delta_2(X)$ estimadores de $g_1(\theta)$ e $g_2(\theta)$ respectivamente, com $\mathrm{E}(\delta_1(X)^2) < \infty$ e $\mathrm{E}(\delta_2(X)^2) < \infty$. Se estes estimadores são não viciados de variância uniformemente mínima (ENVVUM) então $a\delta_1(X) + b\delta_2(X)$ tem variância finita e é ENVVUM de $ag_1(\theta) + bg_2(\theta)$. Aqui a e b são números reais fixados.

ok 3. Seja X uma variável aleatória que toma os valores $-1, 0, 1, 2, 3$ com probabilidades $P(X = -1) = 2pq$ e $P(X = k) = p^k q^{3-k}$, para $k = 0, 1, 2, 3$.

ok (a) Verifique que esta é uma distribuição de probabilidade.

ok (b) Obtenha o estimador não viciado de variância localmente mínima (ENVVLM) em p_0 de

i. p

ii. pq

e verifique, para cada caso, se o estimador é ENVVUM.

ok 4. Prove o Teorema 1.11, p. 88, TPE.

ok 5. Se T tem distribuição binomial $b(p, n)$ com $n > 3$, use o Método 1 para encontrar o ENVVUM de p^3.

6. Suponha que $X_1, \ldots, X_n$ independentes tenham distribuição de Poisson $P(\lambda)$. Use o Método 1 para encontrar o ENVVUM de

ok (a) λ^k para qualquer inteiro k; $\quad X \sim P(\lambda) \Rightarrow p_\lambda(x) = \frac{e^{-\lambda}\lambda^x}{x!}$

(b) $e^{-\lambda}$.

± 7. Resolva o problema 6(b) pelo Método 2 usando independentes o fato de que um estimador não viciado de $e^{-\lambda}$ é: $\delta = 1$, se $X_1 = 0$, e $\delta = 0$, caso contrário.

ok 8. Se $X_1, \ldots, X_n$ são iid com distribuição $\mathrm{N}(\zeta, \sigma^2)$, com σ conhecido, encontre o ENVVUM de ζ^2, ζ^3 e de ζ^4. Sugestão: Para calcular $\mathrm{E}(\overline{X}^k)$, escreva $\overline{X} = Y + \zeta$, onde $Y \sim \mathrm{N}(0, \sigma^2/n)$ e expanda $\mathrm{E}[(Y+\zeta)^k]$. [Ver Problema 2.4, p. 132, TPE].

9. Resolva o problema (8) admitindo que σ é desconhecido.

ok 10. Suponha que $X_1, \ldots, X_n$ sejam variáveis aleatórias independentes com distribuição $\mathrm{N}(\theta_i, 1)$, para $i = 1, \ldots, n$. Seja $\omega = n^{-1}\sum_{i=1}^n \theta_i^2$.

ok (a) Obtenha o ENVVUM de ω e sua variância.

ok (b) Mostre que se $a(x) > 0$ para todo $x = 0, 1, \ldots$, então θ^r é estimável para qualquer r inteiro positivo e seu estimador não viciado de variância uniformemente mínima baseado em uma amostra $X_1, \ldots, X_n$ de observações independentes da distribuição de X é $\delta(T)$, sendo $T = \sum_{i=1}^{n} X_i$ e

$$\delta(t) = \begin{cases} 0, & t = 0, \ldots, r-1, \\ \frac{A(t-r,n)}{A(t,n)}, & t = r, r+1, \ldots, \end{cases}$$

onde $A(t, n)$ é o coeficiente de θ^t na expansão em série de $[C(\theta)]^n$.

ok (c) Se $X_1, \ldots, X_n$ são observações independentes da distribuição de Poisson de média θ, $\theta > 0$, utilize o resultado em (b) para encontrar o ENVVUM de θ^r, r inteiro positivo.

ok (d) Mostre que a distribuição de Poisson truncada à direita em b ($b \geq 1$, inteiro) e parâmetro θ é um caso especial da distribuição de série de potências. (Se X tem distribuição de Poisson truncada à direita em b e parâmetro θ, então $\mathrm{P}_\theta(X = x) = \mathrm{P}_\theta(Y = x | Y \leq b)$, $x = 0, 1, \ldots, b$, onde Y tem distribuição de Poisson de parâmetro θ.

ok (e) Mostre que se X é uma única observação de uma distribuição de Poisson truncada à direita em b ($b \geq 1$, inteiro), então θ não é estimável. Esta afirmação contradiz à do item (b)? Por que?

17. Seja $(X_1, \ldots, X_n)$ uma amostra aleatória de tamanho n de uma distribuição de Bernoulli de parâmetro θ ($0 \leq \theta \leq 1$). Seja $H = (\sum_{i=1}^{n} X_i / n)^2$ o estimador de máxima verossimilhança de θ^2. Seja

$$H^* = nH - \frac{n-1}{n} \sum_{i=1}^{n} H^{(i)}$$

o estimador "jackknife" de θ^2, onde $H^{(i)}$ é calculado como H exceto que X_i é excluída da amostra.

ok (a) Mostre que o estimador de máxima verossimilhança de θ^2 é um estimador viciado mas H^* é não viciado.

(b) H^* é o ENVVUM de θ^2? Se for, prove. Se não for, encontre o ENVVUM.

ok 18. Seja $(X_1, \ldots, X_n)$ uma amostra casual simples de $X \sim$ Exponencial deslocada, ie., com função densidade de probabilidade dada por

$$f(x; a) = \exp\{-(x-a)\} I_{(a,\infty)}(x)$$

para $a \in (-\infty, +\infty)$.

ok (a) Mostre que $X_{(1)} = \min_{i \in \{1,\ldots,n\}} X_i$ é uma estatística suficiente e completa.

ok (b) Mostre que $X_{(1)}$ e S^2, a variância amostral, são variáveis aleatórias independentes.

19. Seja $(X_1, \ldots, X_n)$ uma amostra casual simples de $X \sim \mathrm{E}(a, b)$ com função densidade de probabilidade dada por

$Y_1 \sim E(0,1)$

$$f(x; a, b) = \exp\{-(x-a)/b\} I_{(a,\infty)}(x)$$

para $a \in (-\infty, +\infty)$, $b > 0$. Mostre que $X_{(1)} = \min_{i \in \{1,\ldots,n\}} X_i$ e $\sum [X_i - X_{(1)}]$ são independentes e têm distribuições $\mathrm{E}(a, b/n)$ e $b\mathrm{Gamma}(n-1, 1)$. Sugestão: Se $a = 0$ e $b = 1$, as variáveis $Y_i = (n-i+1)[X_{(i)} - X_{(i-1)}]$, para $i = 2, \ldots, n$, são iid e têm distribuição $E(0, 1)$.

$\Rightarrow \sum Y_i = \sum(X_i - X_{(1)})$, mas $Y_i \sim E(0,1) \Rightarrow \sum Y_i \sim$ Gama $(n-1,1)$

ok 20. Seja $X \sim \mathrm{N}(\mu, \sigma^2)$. Mostre que não existe estimador não viciado de σ^2 quando μ é desconhecido. (Sugestão: para $\sigma^2 = a$ fixado, mostre que X é uma estatística suficiente completa para μ).

1) Seja $X = (X_1, \dots, X_n)$ um vetor de variáveis aleatórias independentes tais que $X_i \sim N(d_i\theta, \theta^2)$, $d_i \neq 0$, para $i = 1, \dots, n$, $\theta > 0$. Sejam

$$\tilde{X} = \frac{\sum_{i=1}^{n} d_i X_i}{\sum_{i=1}^{n} d_i^2} \quad e \quad S^2 = \sum_{i=1}^{n} \frac{(X_i - d_i \tilde{X})^2}{n-1}$$

mostre que entre todos os estimadores não viciados de θ da forma $\sum_{i=1}^{n} v_i X_i$, com $v_i \in (-\infty, \infty)$ para $i = 1, \dots, n$, $\tilde{X}$ é o que tem variância uniformemente mínima.

1º) Calcular o valor esperado de $\tilde{X}$ e sua variância

$$E(\tilde{X}) = \frac{1}{\sum_{i=1}^{n} d_i^2} \sum^{n} d_i E(X_i) = \frac{1}{\cancel{\sum_{i=1}^{n} d_i^2}} \cancel{\sum_{i=1}^{n} d_i^2}\,\theta = \theta$$

$$Var(\tilde{X}) = \frac{1}{(\sum_{i=1}^{n} d_i^2)^2} \sum_{i=1}^{n} d_i^2 \, Var(X_i) = \frac{\theta^2}{\sum_{i=1}^{n} d_i^2}$$

2º) Calcular o valor esperado de $\sum^{n} v_i X_i$ e sua variância

$$E\left[\sum_{i=1}^{n} v_i X_i\right] = \sum_{i=1}^{n} v_i E(X_i) = \theta \sum_{i=1}^{n} v_i d_i$$

para que $\sum_{i=1}^{n} v_i X_i$ seja um estimador não viciado para θ devemos ter $\sum_{i=1}^{n} v_i d_i = 1$.

$$Var\left[\sum_{i=1}^{n} v_i X_i\right] = \sum_{i=1}^{n} v_i^2 \, Var(X_i) = \theta^2 \sum_{i=1}^{n} v_i^2$$

3º) Mostrar que entre todos os estimadores n. viciados de θ da forma $\sum_{i=1}^{n} v_i X_i$, $\tilde{X}$ é o que tem variância uniformemente mínima

Ou seja, devemos mostrar que o v_i é igual a $\frac{d_i}{\sum d_i^2}$

Seja X^* um estimador qualquer de θ da forma $\sum_{i=1}^{n} v_i d_i$, devemos encontrar v_i tal que

$$Var(\tilde{X}) \leq Var(X^*)$$

Para encontrar v_i que minimiza $Var(X^*)$, utilizaremos o multiplicador de Lagrange, considerando a restrição $\sum_{i=1}^{n} v_i d_i = 1$

Considere $f(\underset{\sim}{v}) = \theta^2 \sum_{i=1}^{n} v_i^2 - \lambda\left(\sum_{i=1}^{n} v_i d_i - 1\right)$, $\underset{\sim}{v} = (v_1, \dots, v_n)$

temos que

$$\frac{\partial f(\underset{\sim}{v})}{\partial v_i} = 2\theta^2 v_i - \lambda d_i$$

Quando $\dfrac{\partial f(\underset{\sim}{v})}{\partial v_i} = 0$ temos

$$2\theta^2 v_i - \lambda d_i = 0$$ multiplica por d_i dos 2 lados

$$2\theta^2 v_i d_i - \lambda d_i^2 = 0$$

Daí

$$2\theta^2 \sum_{i=1}^{n} v_i d_i - \sum_{i=1}^{n} \lambda d_i^2 = 0$$

$\sum_{i=1}^{n} v_i d_i = 1 \iff \dfrac{\partial f(\underset{\sim}{v})}{\partial \lambda} = 0$

$$2\theta^2 - \lambda \sum_{i=1}^{n} d_i^2 = 0$$

$$\lambda = \frac{2\theta^2}{\sum_{i=1}^{n} d_i^2}$$

Logo

$$2\theta^2 v_i - \left(\frac{2\theta^2}{\sum d_i^2}\right) d_i = 0$$

$$\cancel{2\theta^2} v_i = \frac{\cancel{2\theta^2}\, d_i}{\sum d_i^2}$$

$$\boxed{v_i = \frac{d_i}{\sum d_i^2}}$$

Logo $v_i = \dfrac{d_i}{\sum_{i=1}^{n} d_i^2}$ minimiza $Var(X^*)$ $\forall\, \theta$

Daí,

$$Var\left[\sum v_i X_i\right] = Var\left\{\sum_{i=1}^{n} \frac{d_i}{\sum_{i=1}^{n} d_i^2} X_i\right\} = Var\left\{\frac{\sum_{i=1}^{n} d_i X_i}{\sum_{i=1}^{n} d_i^2}\right\}$$

ou seja, quando $v_i = \dfrac{d_i}{\sum_{i=1}^{n} d_i^2}$ temos

$$\text{Var}\left\{ \sum_{i=1}^{n} v_i X_i \right\} = \text{Var}(\tilde{X})$$

Mas mostramos anteriormente que v_i da forma acima minimiza a variância. Logo $\tilde{X}$ é o estimador que tem variância uniformemente ($\forall \theta$) mínima.

C2) Prove ou dê contra-exemplo. Seja X variável aleatória com distribuição P_θ pertencente a uma família de distribuições $\mathcal{P} = \{P_\theta, \theta \in \Omega\}$ e sejam $g_1(\theta)$ e $g_2(\theta)$ funções de θ a valores reais. Sejam $\delta_1(X)$ e $\delta_2(X)$ estimadores de $g_1(\theta)$ e $g_2(\theta)$ respectivamente, com $E[\delta_1(X)^2] < \infty$ e $E(\delta_2(X)^2) < \infty$. Se estes estimadores são não viciados de variância uniformemente mínima (ENVVUM) então $a\delta_1(X) + b\delta_2(X)$ tem variância finita e é ENVVUM de $ag_1(\theta) + bg_2(\theta)$. Aqui a e b são nºs reais fixados.

Seja $\mathcal{E}$ a classe dos estimadores com variância finita e

$$\mathcal{U} = \{U \in \mathcal{E} : E_\theta[U] = 0 \; ; \; \forall \theta \in \Omega\} \subset \mathcal{E}$$

a classe dos estimadores não viciados do zero.

Por um lema sabemos que: Uma condição necessária e suficiente para que um estimador $\delta(X)$ não viciado de $g(\theta)$ seja ENVVUM é que $\delta(X) \in \mathcal{E}$ e $Cov(\delta(X), U) = 0 \quad \forall U \in \mathcal{U}$, onde $E_\theta[\delta^2(X)] < \infty$

Por hipótese temos que $\delta_1(X)$ e $\delta_2(X)$ são ENVVUM de $g_1(\theta)$ e $g_2(\theta)$ respectivamente, então, pelo lema acima temos que:

$E_\theta[\delta_1(X)] = g_1(\theta)$, $Var_\theta[\delta_1(X)] < \infty$, $Cov_\theta[\delta_1(X), U] = 0$ e

$E_\theta[\delta_1(X)] = g_2(\theta)$, $Var_\theta[\delta_2(X)] < \infty$, $Cov_\theta[\delta_2(X), U] = 0 \quad \forall U \in \mathcal{U}$,

$\forall \theta \in \Omega$.

Quero mostrar que $a\delta_1(X) + b\delta_2(X)$ é ENVVUM de $ag_1(\theta) + bg_2(\theta)$ ou seja, devo mostrar as propriedades seguintes e aplicar o lema acima:

a) $E[a\delta_1(X) + b\delta_2(X)] = ag_1(\theta) + bg_2(\theta)$

b) $Var_\theta[a\delta_1(X) + b\delta_2(X)] < \infty$

c) $Cov_\theta[a\delta_1(X) + b\delta_2(X), U] = 0 \quad \forall U \in \mathcal{U}$.

a) $E[a\delta_1(X)+b\delta_2(X)] = E[a\delta_1(X)] + E[b\delta_2(X)]$

$= a\,E[\delta_1(X)] + b\,E(\delta_2(X))$

$= a\,g_1(\theta) + b\,g_2(\theta)$

$\forall\, a,b \in \mathbb{R}$ fixados

b) $Var[a\delta_1(X)+b\delta_2(X)] = E[(a\delta_1(X)+b\delta_2(X))^2] - \{E[a\delta_1(X)+b\delta_2(X)]\}^2$

$= E[a^2\delta_1(X)^2 + b^2\delta_2(X) + 2ab\,\delta_1(X)\delta_2(X)] - \{aE[\delta_1(X)] + bE[\delta_2(X)]\}^2$

$= a^2E[\delta_1(x)^2] + b^2E[\delta_2(X)] + 2ab\,E[\delta_1(X)\delta_2(X)] - a^2\{E(\delta_1(X))\}^2 -$
$b^2\{E[\delta_2(X)]\}^2 - 2ab\,E[\delta_1(X)]E[\delta_2(X)]$

$= a^2\{E[\delta_1(X)^2] - E^2[\delta_1(X)]\} + b^2\{E[\delta_2(X)^2] - E^2[\delta_2(X)]\}$
$+\, 2ab\{E[\delta_1(X)\delta_2(X)] - E[\delta_1(X)]E[\delta_2(X)]\}$

$= a^2\,Var[\delta_1(X)] + b^2\,Var[\delta_2(X)] + 2\,Cov\{\delta_1(X),\delta_2(X)\}$

Sabemos que $Var[\delta_1(X)] < \infty$ e $Var[\delta_2(X)] < \infty$. Logo para mostrar que

$$Var[a\delta_1(X)+b\delta_2(X)] < \infty$$

resta mostrar que $Cov\{\delta_1(X),\delta_2(X)\} < \infty$.

Mas pela desigualdade de Hölder $(p=q=2)$ temos

$$E_\theta[\delta_1(X)\delta_2(X)] \leq \{\underbrace{E(\delta_1(X)^2)}_{<\infty}\underbrace{E(\delta_2(X)^2)}_{<\infty}\}^{1/2} < \infty \quad \forall\, \theta \in \Omega.$$

como $E(S_1(x))E(S_1(x))$ tambem é finito (por hipótese) logo

$$Cov[S_1(X), S_2(X)] < \infty$$

$$\therefore \; Var[aS_1(X) + bS_2(X)] < \infty$$

c) $Cov[aS_1(X) + bS_2(X), U] = a\,Cov[S_1(X), U] + b\,Cov[S_2(X), U]$

$$= 0$$

pois por hipótese $Cov[S_1(X), U] = Cov[S_2(X), U] = 0 \quad \forall\, \theta \in \Omega$.

$\therefore\; aS_1(X) + bS_2(X)$ é ENVVUM de $ag_1(\theta) + bg_2(\theta)$

$\forall\, a, b \in \mathbb{R}$ fixados.

5

03) Seja X uma variável aleatória que toma os valores $-1, 0, 1, 2,$ 3 com probabilidades $P(X=-1) = 2pq$ e $P(X=x) = p^k q^{3-k}$ para $k=0,1,2,$ 3.

(a) Verifique que esta é uma distribuição de probabilidade.

(i) Como $0 < p, q < 1$ logo $P(X=k) \geq 0 \quad \forall \, k = -1, 0, 1, 2, 3$.

(ii) mostrar que $\sum_{k=-1}^{3} P(X=k) = 1$

Temos que

$$\sum_{k=-1}^{3} P(X=k) = P(X=-1) + P(X=0) + P(X=1) + P(X=2) + P(X=3)$$

$$= 2pq + p^0q^3 + p^1q^2 + p^2q^1 + p^3q^0$$

$$= 2pq + q^3 + pq^2 + p^2q + p^3$$

$$= 2p(1-p) + (1-p)^2(1-p) + p(1-p)^2 + p^2(1-p) + p^3$$

$$= 2p - 2p^2 + (1-2p+p^2)(1-p) + p(1-2p+p^2) + p^2 - \cancel{p^3} + \cancel{p^3}$$

$$= 2p - \cancel{2p^2} + 1 - p - 2p + \cancel{2p^2} + p^2 - p^3 + p - 2p^2 + p^3 + p^2$$

$$= 1$$

$\therefore$ $P(X=k)$ é uma distribuição de probabilidade

(b) Obtenha o estimador não viciado de variância localmente mínima (ENVVLM) em p_0 de

(i) p

(ii) pq

e verifique, para cada caso, se o estimador é ENVVUM

Temos um lema que diz: Se δ_0 é qualquer estimador não viciado de $g(\theta)$, a totalidade de estimadores não viciados é dado por

$$\delta = \delta_0 - U,$$

onde U representa qualquer estimador não viciado do "zero", ou seja, U satisfaz

$$E_\theta(U) = 0 \quad \forall\, \theta \in \Omega$$

Solução:

(i) a idéia aqui é procurar um estimador não viesado δ_0 e encontrar um estimador δ da forma do lema acima que minimize o vesei uniformemente em p.

Considere um estimador de p da forma abaixo

$$\delta_0(x) = \delta_0 = \begin{cases} 0 & \text{se } x = -1 \text{ e } 0 \\ 1 & \text{se } x = 1, 3 \\ 2 & \text{se } x = 2 \end{cases}$$

Daí,

$$E(\delta_0) = 0 \cdot P(X=-1) + 0\,P(X=0) + 1\,P(X=1) + 1\,P(X=3) + 2P(X=2)$$

$$= p(1-p)^2 + p^3(1-p)^0 + 2p^2(1-p)$$

$$= p(1-2p+p^2) + p^3 + 2p^2 - 2p^3$$

$$= p - 2p^2 + p^3 + p^3 + 2p^2 - 2p^3$$

$\Rightarrow E(\delta_0) = p$. Logo δ_0 é um estimador não viesado de p.

Agora vamos encontrar U tal que $E(U) = 0$ para que possamos posteriormente encontrar δ no lema acima. Temos que

$$E[U(k)] = \sum_{k=-1}^{3} u(x)\, P(X=k) = u(-1)P(X=-1) + u(0)P(X=0) + u(1)P(X=1) + u(2)P(X=2) + u(3)P(X=3)$$

$$= u(-1)\, 2pq + u(0)\, q^3 + u(1)\, p(1-p)^2 + u(2)\, p^2(1-p) + u(3)\, p^3$$

$$= u(-1)\, 2pq + u(0)\, q^3 + u(1)\, p(1-p)^2 + u(2)\, p^2(1-p) + u(3)\, p^3$$

$$= u(-1)[2p - 2p^2] + u(0)(1-p)^2(1-p) + u(1)(p(1-2p+p^2)) + u(2)(p^2 - p^3) + u(3)p^3$$

$$E[U(k)] = 2p\,u(-1) - 2p^2u(-1) + u(0)(1-2p+p^2)(1-p) + u(1)(p-2p^2+p^3) + p^2u(2) - p^3u(2) + p^3u(3)$$

$$= 2p\,u(-1) - 2p^2u(-1) + u(0)(1-2p+p^2-p+2p^2-p^3) + pu(1) - 2p^2u(1) + p^3u(1) + p^2u(2) - p^3u(2) + p^3u(3)$$

$$= 2p\,u(-1) - 2p^2u(-1) + u(0) - 3p\,u(0) + 3p^2u(0) - p^3u(0) + pu(1) - 2p^2u(1) + p^3u(1) + p^2u(2) - p^3u(2) + p^3u(3)$$

$$= p\left[2u(-1) - 3u(0) + u(1)\right] + p^2\left[-2u(-1) + 3u(0) - 2u(1) + u(2)\right] +$$

$$+ p^3\left[-u(0) + u(1) - u(2) + u(3)\right] + u(0)$$

Supondo que $E[U(k)] = 0 \;\; \forall p$, mas isso acontece se e somente se

$$\begin{cases} u(0) = 0 \\ 2u(-1) - 3u(0) + u(1) = 0 & \Rightarrow u(1) = -2u(-1) \\ -2u(-1) + 3u(0) - 2u(1) + u(2) = 0 & \Rightarrow u(2) = 2u(1) + 2u(-1) = 2u(-1) - 4u(-1) = -2u(-1) \\ -u(0) + u(1) - u(2) + u(3) = 0 & \Rightarrow u(3) = u(2) - u(1) = -2u(-1) + 2u(-1) = 0 \end{cases}$$

$\Rightarrow$ $u(0) = u(3) = 0$ e $u(1) = u(2) = -2u(-1)$. Faça $c = u(-1)$ logo teremos que $u(k)$ tal que $E[u(k)] = 0$ tem a forma

$$U = u(k) = \begin{cases} c & \text{se } k = -1 \\ 0 & \text{se } k = 0 \text{ e } 3 \\ -2c & \text{se } k = 1 \text{ e } 2 \end{cases}$$

Digamos que a perda seja quadrática então para encontrar um estimador não viciado de variância localmente mínima em p_0 temos que minimizar

$$E_{p_0}[(\delta_0 - U)^2].$$

Temos que

$$E_{p_0}[(\delta_0 - U)^2] = \sum_{k=-1}^{3} (\delta_0(k) - u(k))^2 \, P_{p_0}(X = k)$$

$$= [\delta_0(-1) - u(-1)]^2 P_{p_0}(X=-1) + [\delta_0(0) - u(0)]^2 P_{p_0}(X=0) + [\delta_0(1) - u(1)]^2 P_{p_0}(X=1) +$$

$$[\delta(2) - u(2)]^2 P_{p_0}(X=2) + [\delta(3) - u(3)]^2 P_{p_0}(X=3)$$

$$\Rightarrow E_{p_0}\left[(\delta_0 - U)^2\right] = (-c)^2\, 2p_0(1-p_0) + [1+2c]^2\, p_0(1-p_0)^2 + [2+2c]^2\, p_0^2(1-p_0) + p_0^3$$

$$= 2p_0c^2 - 2p_0^2c^2 + (1+4c+4c^2)\,p_0(1-2p_0+p_0^2) + (4+8c+4c^2)(p_0^2-p_0^3) + p_0^3$$

$$= 2p_0c^2 - 2p_0^2c^2 + (1+4c+4c^2)(p_0-2p_0^2+p_0^3) + 4p_0^2 - 4p_0^3 + 8p_0^2c - 8p_0^3c + 4p_0^2c^2 - 4p_0^3c^2 + p_0^3$$

$$= 2p_0c^2 + 2p_0^2c^2 + p_0 - 2p_0^2 + p_0^3 + 4p_0c - 8p_0^2c + 4p_0^3c + 4p_0c^2 - 8p_0^2c^2 + 4p_0^3c^2 + 4p_0^2 - 3p_0^3 + 8p_0^2c - 8p_0^3c - 4p_0^3c^2$$

$$= 6p_0c^2 - 6p_0^2c^2 + p_0 + 2p_0^2 - 2p_0^3 + 4p_0c - 4p_0^3c$$

Faça

$$f(c) = c^2(6p_0 - 6p_0^2) + p_0 + 2p_0^2 - 2p_0^3 - c(-4p_0 + 4p_0^3)$$

Logo

$$\frac{\partial f(c)}{\partial c} = 2c(6p_0 - 6p_0^2) + 4p_0 - 4p_0^3$$

$$\frac{-3p_0-1}{3-2p_0}$$

Daí

$$\frac{\partial f(c)}{\partial c} = 0 \Leftrightarrow 2c(6p_0 - 6p_0^2) + 4p_0 - 4p_0^3 = 0$$

$$\Leftrightarrow c = \frac{-4p_0 + 4p_0^3}{12p_0 - 12p_0^2} = \frac{\cancel{4p_0}(p_0^2-1)}{\cancel{4p_0}(3-3p_0)} = -\frac{(1-p_0^2)}{3(1-p_0)} = -\frac{\cancel{(1-p_0)}(1+p_0)}{3\cancel{(1-p_0)}} = -\frac{(1+p_0)}{3}$$

Temos ainda que

$$\frac{\partial^2 f(c)}{\partial c^2} = 12p_0 - 12p_0^2 = 12p_0(1-p_0) = 12p_0q_0 > 0$$

Logo o ponto que minimiza $f(c)$ é $c = -\frac{(1+p_0)}{3}$. Daí, $\delta(K) = \delta_0(K) - U(K)$ onde

$$U(K) = \begin{cases} -\frac{(1+p_0)}{3} & \text{se } K = -1 \\ 0 & \text{se } K = 0 \text{ e } 3 \\ \frac{2}{3}(1+p_0) & \text{se } K = 1 \text{ e } 2 \end{cases}$$

é estimador não viciado de variância localmente mínima para $p = p_0$. Temos que c minimiza $f(c)$ para $p = p_0$ e não $\forall\, p$ logo $\delta(K)$ não é um estimador não viciado de variância uniformemente mínima.

(ii) Considere um estimador de pq da forma

$$\delta_0(X) = \begin{cases} \frac{1}{2} & \text{se } X=-1 \\ 0 & \text{se } X=0,1,2,3 \end{cases}$$

Daí, temos que

$$E[\delta_0(X)] = \frac{1}{2}P(X=-1) = \frac{1}{2}2pq = pq.$$

logo $\delta_0(X)$ é estimador não viciado de pq.

Sabemos, pelo item anterior, que $U(K)$ tal que $E(U(K))=0$ é dado por

$$U(K) = \begin{cases} c & \text{se } K=-1 \\ 0 & \text{se } K=0 \text{ e } 3 \\ -2c & \text{se } K=1 \text{ e } 2 \end{cases}$$

onde $c = U(-1)$.

Consideremos novamente, que a perda é quadrática, então para encontrar um ENVVLH em p_0 temos que minimizar

$$E_{p_0}[(\delta_0 - U)^2]$$

Temos que

$$E_{p_0}[(\delta_0-U)^2] = \sum_{K=-1}^{3}(\delta_0(K)-U(K))^2 P_{p_0}(X=K)$$

$$= \left[\tfrac{1}{2}-c\right]^2 2p_0q_0 + [0-0]^2 P_0(X=0) + [0+2c]^2 p_0(1-p_0)^2 + [0+2c]^2 p_0^2(1-p_0) + [0-0]p_0^3$$

$$= \left(\tfrac{1}{4}-c+c^2\right)2p_0q_0 + 4c^2\{p_0q_0^2+p_0^2q_0\}$$

$$= \frac{2p_0q_0}{4} - 2p_0q_0c + 2p_0q_0c^2 + 4c^2\{p_0q_0^2+p_0^2q_0\}$$

Faça

$$f(c) = \frac{p_0q_0}{2} - 2p_0q_0c + 4c^2\left\{p_0q_0^2+p_0^2q_0+\frac{p_0q_0}{2}\right\}$$

Daí

$$\frac{\partial f(c)}{\partial c} = -2p_0q_0 + 8c\left\{p_0q_0^2+p_0^2q_0+\frac{p_0q_0}{2}\right\}$$

logo

$$\frac{\partial f(c)}{\partial c} = 0 \iff 8c\left[p_0q_0^2 + p_0^2q_0 + \frac{p_0q_0}{2}\right] = 2p_0q_0$$

$$\iff c = \frac{\cancel{2}\,\cancel{p_0q_0}}{\cancel{8}^4\,\cancel{p_0q_0}\left[q_0 + p_0 + \frac{1}{2}\right]} \iff c = \frac{1}{4(1-\cancel{p_0}+\cancel{p_0}+1/2)} = \frac{1}{2\cancel{4}(\frac{3}{\cancel{2}})} = \frac{1}{6}$$

Temos ainda

$$\frac{\partial^2 f(c)}{\partial c} = 8\left[p_0q_0^2 + p_0^2q_0 + \frac{p_0q_0}{2}\right] = 8\,p_0q_0\left[q_0 + p_0 + 1/2\right]$$

$$= 8\,p_0q_0\,(1 + 1/2) = \cancel{8}^4 \times \frac{3}{\cancel{2}}\,p_0q_0 = 12\,p_0q_0 > 0$$

Daí, temos que o ponto onde $f(c)$ é mínima é $c = \frac{1}{6}$. Logo $S(k) = S_0(k) - U(k)$, onde $u(k) = \begin{cases} \frac{1}{6} & \text{se } k=-1 \\ 0 & \text{se } k = 0 \text{ e } 3 \\ -1/3 & \text{se } k = 1 \text{ e } 2 \end{cases}$ é o estimador não viciado de variância localmente mínima para $p = p_0$. Temos que $c = \frac{1}{c}$ minimiza $f(c)$ para $\forall\ 0 < p < 1$ logo $S(k)$ é estimador não viciado de variância uniformemente mínima.

04) Teorema: Seja X distribuída de acordo com a distribuição em $\mathcal{P} = \{ P_\theta , \theta \in \Omega \}$. Suponha que T é uma estatística suficiente e completa para $\mathcal{P}$.

(i) Para toda função estimável $g(\theta)$ existe um estimador não viciado que minimiza o risco uniformemente para qualquer função de perda convexa, este estimador é ENVVUM.

Se $g(\theta)$ é estimável então existe um estimador não viciado $\delta(X)$ de $g(\theta)$ onde $\varphi(T) = E(\delta(X)/T)$ e T é uma estatística suficiente e completa para $\mathcal{P}$. Se $L(\theta, d)$ é uma função convexa em d então pelo teorema de Rao - Blacwell

$$R(\theta, \varphi(T)) < R(\theta, \delta(x)) \qquad \forall \theta \in \Omega$$

sempre que

$$R(\theta, \delta(X)) = E[L(\theta, \delta(X))] < \infty$$

Assim $\varphi(T)$ é um estimador de risco mínimo e como $\varphi(T)$ não depende de θ mas de T então $\varphi(T)$ é ENVVUM, onde a função de perda quadrática que é estritamente convexa.

(ii) O ENVVUM de (i) é o único estimador não viesado o qual é função de uma estatística suficiente, e é o único estimador não viesado com risco mínimo, desde que seu risco seja finito e L seja estritamente convexa.

Suponha que existe outro estimador não viesado de risco mínimo que denotaremos por $\varphi_1(T)$, com $L(\theta, d)$ perda convexa em d.
Então como $\varphi(T)$ e $\varphi_1(T)$ são 2 ENV de $g(\theta)$ temos

$$E[\varphi(T) - \varphi_1(T)] = 0 \qquad \forall \theta \in \Omega$$

Como T é estatística suficiente completa para a família $\mathcal{P}$. temos

$$\varphi(T) - \varphi_1(T) = 0 \qquad q.c.-\mathcal{P}.$$

Logo $\varphi(T) = \varphi_1(T)$ q.c.-P_θ $\forall \theta \in \Omega$. Assim o estimador em (6) é único, sempre que o risco seja finito e L estritamente convexa

5) Se $T \sim bin(p,n)$ com $n > 3$, use o método 1 para encontrar o ENVVUM de p^3

Temos que a densidade é dada por:

$$P[T=t] = \binom{n}{t} p^t q^{n-t} \qquad t = 0,1,\dots,n \;,\; 0 \leq p \leq 1 \;,\; n \geq 3, \dots$$

$$= \exp\left\{ \log p^t + \log q^{n-t} \right\} \binom{n}{t}$$

$$= \exp\left\{ t \log p + n \log q - t \log q \right\} \binom{n}{t}$$

$$= \exp\left\{ \log\left(\frac{p}{q}\right) t - (-n \log q) \right\} \binom{n}{t}$$

$$= \exp\left\{ \eta(p) T(t) - B(p) \right\} h(t)$$

onde $\eta(p) = \log\left(\frac{p}{1-p}\right)$, $T(t) = t$, $B(p) = -n \log(1-p)$ e $h(t) = \binom{n}{t}$

O espaço paramétrico é $\{\eta(p) : p \in [0,1]\}$. Como $\eta(p) = \log\left(\frac{p}{1-p}\right)$ logo $\{\eta(p) : p \in [0,1]\} = (0,1)$ é um aberto em $\mathbb{R}$ de interior não vazio então a estatiststística $T(t)$ é suficiente e completa para p.

Método 1: Se T é estatística suficiente e completa, o ENVVUM de $\forall$ função estimável $g(p)$ é unicamente determinado por

$$E_p[\delta(T)] = g(p) \qquad \forall \theta \in \Omega.$$

Como p^3 é função de p, Tome $g(p) = p^3$ logo usando o método 1 temos

$$E_p[\delta(T)] = p^3 \qquad \forall\; 0 \leq p \leq 1.$$

logo

$$E_p[\delta(T)] = \sum_{t=0}^{n} \delta(t) \binom{n}{t} p^t q^{n-t} = p^3 \quad (*)$$

Considere $\alpha = \frac{p}{q}$ logo $\alpha = \frac{p}{1-p}$ $\Rightarrow \alpha(1-p) = p \Rightarrow \alpha - \alpha p = p$

$\Rightarrow p + \alpha p = \alpha \Rightarrow p(1+\alpha) = \alpha \Rightarrow p = \frac{\alpha}{1+\alpha}$ e

$q = 1 - \frac{\alpha}{1+\alpha} = \frac{1+\alpha-\alpha}{1+\alpha} = \frac{1}{1+\alpha}$, daí substituindo em $(*)$

$$\sum_{t=0}^{n} \delta(t) \binom{n}{t} \alpha^t \left(\frac{1}{1+\alpha}\right)^n = \left(\frac{\alpha}{1+\alpha}\right)^3$$

$$\sum_{t=0}^{n} \delta(t) \binom{n}{t} \alpha^t = \left(\frac{\alpha}{1+\alpha}\right)^3 (1+\alpha)^n$$

ou

$$\sum_{t=0}^{n} \delta(t) \binom{n}{t} \alpha^t = \alpha^3 (1+\alpha)^{n-3} \qquad (**)$$

Pelo Binômio de Newton temos que

$$(a+b)^n = \sum_{j=0}^{n} \binom{n}{j} a^j b^{n-j}$$

Daí,

$$(1+\alpha)^{n-3} = \sum_{j=0}^{n-3} \binom{n-3}{j} 1^j \alpha^{n-3-j}$$

Logo

$$\alpha^3 (1+\alpha)^{n-3} = \sum_{j=0}^{n-3} \binom{n-3}{j} \alpha^{n-j}.$$

Voltando em $(**)$ temos

$$\sum_{t=0}^{n} \delta(t) \binom{n}{t} \alpha^t = \sum_{j=0}^{n-3} \binom{n-3}{j} \alpha^{n-j}.$$

Ou seja,

$$\delta(0)\binom{n}{0}\alpha^0+\delta(1)\binom{n}{1}\alpha^1+\dots+\delta(n)\binom{n}{n}\alpha^n=\binom{n-3}{0}\alpha^{n-0}+\binom{n-3}{1}\alpha^{n-1}+\dots+\binom{n-3}{n-3}\alpha^3$$

$$=\binom{n-3}{n-3}\alpha^3+\dots+\binom{n-3}{1}\alpha^{n-1}+\binom{n-3}{0}\alpha^n$$

Daí,

$$\delta(0)\binom{n}{0}=\delta(1)\binom{n}{1}=\delta(2)\binom{n}{2}=0$$

e

$$\delta(3)\binom{n}{3}=\binom{n-3}{n-3},\quad \delta(4)\binom{n}{4}=\binom{n-3}{n-4},\ \dots,\ \delta(n)\binom{n}{n}=\binom{n-3}{0}$$

ou ainda

$$\delta(t)=\begin{cases}0 & \text{se } t=0,1,2\\[4pt] \dfrac{\binom{n-3}{n-t}}{\binom{n}{t}} & \text{se } t=3,4,\dots,n\end{cases}$$

Portanto se $T\sim bin(p,n)$, $n>3$, o ENVVUM para p^3 é

$$\delta(T)=\frac{\binom{n-3}{n-T}}{\binom{n}{T}}\quad \text{se } T=3,4,\dots,n.$$

$$\frac{\dfrac{(n-3)!}{\cancel{(n-t)!}\,(\cancel{n}-3-\cancel{n}+t)!}}{\dfrac{n!}{t!\,\cancel{(n-t)!}}}=\frac{\dfrac{(n-3)!}{(t-3)!}}{\dfrac{n!}{t!}}$$

$$\frac{t!\,(n-3)!}{n!\,(t-3)!}=$$

$$=\frac{t\cdot(t-1)\cdot(t-2)}{n(n-1)(n-2)}=$$

$$=\frac{(t^2-t)(t-2)}{n(n-1)(n-2)}=$$

$$=\frac{t^3-2t^2-t^2+2t}{n(n-1)(n-2)}=$$

$$=\frac{t^3-3t^2+2t}{n(n-1)(n-2)}$$

6) Suponha que $X_1, \ldots, X_n$ tenham distribuição de Poisson $P(\lambda)$. Use o método 1 para encontrar o ENVVUM de

(a) λ^k para qualquer inteiro k.

Método 1: Resolvendo equações em δ. Se T é estatística suficiente completa, o ENVVUM de $\forall$ função estimável $g(\theta)$ é unicamente determinado pela equação

$$E_\theta[\delta(T)] = g(\theta) \quad \forall \theta \in \Omega.$$

1º) Mostrar que $T = \sum X_i$ é estatística suficiente e completa para λ.

A verossimilhança de $\underset{\sim}{X} = (X_1, \ldots, X_n)$ é dada por:

$$p_\lambda(\underset{\sim}{x}) = \prod_{i=1}^{n} \frac{e^{-\lambda} \lambda^{x_i}}{x_i!} = \frac{e^{-n\lambda} \lambda^{\sum_{i=1}^{n} x_i}}{\prod x_i!}$$

$$= \exp\left\{ \log\left[e^{-n\lambda} \lambda^{\sum_{i=1}^{n} x_i} \right] \right\} \left(\prod_{i=1}^{n} x_i!\right)^{-1}$$

$$= \exp\left\{ \log\lambda \sum_{i=1}^{n} x_i - n\lambda \right\} \left(\prod_{i=1}^{n} x_i!\right)^{-1}$$

$$= \exp\left\{ \eta(\lambda) T(x) - B(\lambda) \right\} h(x)$$

onde $\eta(\lambda) = \log\lambda$, $T(x) = \sum_{i=1}^{n} x_i$, $B(\lambda) = n\lambda$ e $h(x) = \left(\prod_{i=1}^{n} x_i!\right)^{-1}$

$\therefore$ Logo esta distribuição pertence a família exponencial uniparamétrica λ.

2º) Encontrar o ENVVUM para λ^k

Pelo método 1, o ENVVUM de λ^k é determinado unicamente determinado pela equação:

$$E_\lambda[\delta(T)] = \lambda^k, \quad \text{onde} \quad T = \sum_{i=1}^{n} x_i$$

Sabemos que

$$\sum_{i=1}^{n} X_i \sim \text{Poisson}(n\lambda)$$

Daí,

$$E[\delta(T)] = \sum_{t=0}^{\infty} \delta(t)\, P(X=t)$$

$$= \sum_{t=0}^{\infty} \delta(t) \frac{e^{-n\lambda}(n\lambda)^t}{t!} = \lambda^k$$

$$\Rightarrow \sum_{t=0}^{\infty} \frac{\delta(t)(n\lambda)^t}{t!} = \lambda^k e^{n\lambda}$$

mas sabemos que

$$e^{n\lambda} = \sum_{i=0}^{\infty} \frac{(n\lambda)^i}{i!}$$

Daí,

$$\sum_{t=0}^{\infty} \frac{\delta(t)(n\lambda)^t}{t!} = \lambda^k \sum_{i=0}^{\infty} \frac{(n\lambda)^i}{i!}$$

$$= \sum_{i=0}^{\infty} \frac{\lambda^k n^i \lambda^i}{i!}$$

$$= \sum_{i=0}^{\infty} \frac{n^i \lambda^{i+k}}{i!} = \sum_{i=0}^{\infty} \frac{\lambda^k (n\lambda)^i}{i!}$$

tomando $t = i + k \Rightarrow i = t - k$ logo

$$\sum_{t=0}^{\infty} \frac{\delta(t)(n\lambda)^t}{t!} = \sum_{t=k}^{\infty} \frac{n^{(t-k)}\lambda^t}{(t-k)!}$$

$$\sum_{t=0}^{\infty} \frac{\delta(t)(n\lambda)^t}{t!} = \sum_{t=k}^{\infty} \frac{(n\lambda)^t}{n^k (t-k)!}$$

$$\sum_{t=0}^{k-1} \frac{\delta(t)(n\lambda)^t}{t!} + \sum_{t=k}^{\infty} \frac{\delta(t)(n\lambda)^t}{t!} = \sum_{t=k}^{\infty} \frac{(n\lambda)^t}{n^k (t-k)!}$$

$$\frac{\delta(0)(n\lambda)^0}{0!} + \frac{\delta(1)(n\lambda)^1}{1!} + \cdots + \frac{\delta(k-1)(n\lambda)^{k-1}}{(k-1)!} + \frac{\delta(k)(n\lambda)^k}{k!} + \cdots = \frac{(n\lambda)^k}{n^k(k-k)!} + \frac{(n\lambda)^{k+1}}{n^{k+1}(k+1-k)!} + \cdots$$

Comparando os dois lados da igualdade os termos $(n\lambda)$ temos que

$\delta(0) = \delta(1) = \cdots = \delta(k-1) = 0$ e

$$\frac{\delta(k)}{k!} = \frac{1}{n^k (k-k)!} \quad , \quad \frac{\delta(k+1)}{(k+1)!} = \frac{1}{n^{k+1}(k+1-k)!} \quad , \ldots$$

ou seja,

$$\delta(t) = \begin{cases} 0 & \text{se } t = 0, \ldots, k-1 \\ \dfrac{t!}{n^k (t-k)!} & \text{se } t = k, k+1, \ldots \end{cases}$$

Portanto o ENVVUM de λ^k é

$$\delta(t) = \frac{t!}{n^k (t-k)!} \quad , \quad t = k, k+1, \ldots$$

(b) E

Como no item (a) tenho que encontrar $\delta(T)$ tal que

$$E_\lambda[\delta(T)] = e^{-\lambda} \quad , \text{ onde } \quad T = \sum_{i=1}^{n} X_i \sim \mathcal{P}(n\lambda)$$

Logo,

$$\sum_{t=0}^{\infty} \frac{\delta(t)\, e^{-n\lambda}}{t!} (n\lambda)^t = e^{-\lambda}$$

Sabemos que

$$e^{-\lambda} = \sum_{t=0}^{\infty} \frac{\lambda^t}{t!}$$

Pela parte a) $\delta_k(T) = \dfrac{T!}{n^T (T-k)!}$ é ENVVUM para λ^k

07) Resolva o problema 6(b) pelo método 2 usando o fato de que um estimador não viciado de $e^{-\lambda}$ é: $\delta = 1$ se $X_1 = 0$ e $\delta = 0$, caso contrário

Método 2: Toma-se um estimador não viciado $\delta(X)$ para $g(\theta)$ e o ENVVUM é obtido como esperança condicional de $\delta(X)$ dado T, uma estatística suficiente completa.

Solução:

Temos que

$$\delta = \begin{cases} 1 & \text{se } X_1 = 0 \\ 0 & \text{se } X_1 \neq 0 \end{cases}$$

$\Rightarrow E(\delta) = 1\, P(X_1 = 0) + 0\, P(X_1 \neq 0) = P(X_1 = 0)$

Mas

$$P(X_i = x_i) = \frac{e^{-\lambda} \lambda^{x_i}}{x_i!} \Rightarrow P(X_1 = 0) = e^{-\lambda} \quad \text{logo}$$

$$E(\delta) = e^{-\lambda}$$

$\Rightarrow \delta$ é não viciado para $e^{-\lambda}$. Temos que encontrar a esperança condicional de δ dado $T = \sum_{i=1}^{n} X_i$. a estatística suficiente completa.

$\rightarrow$ Encontrar $P(X_1 = 0 \mid \sum X_i = t)$

Mas

$$P(X_1 = 0 \mid \sum X_i = t) = \frac{P(X_1 = 0, \sum_{i=1}^{n} X_i = t)}{P(\sum_{i=1}^{n} X_i = t)} = \frac{P(\sum_{i=1}^{n} X_i = t \mid X_1 = 0) . P(X_1 = 0)}{P(\sum_{i=1}^{n} X_i = t)}$$

$$= \frac{P(\sum_{i=2}^{n} X_i = t) . P(X_1 = 0)}{P(\sum_{i=1}^{n} X_i = t)}$$

$$= \cancel{e^{\lambda}} \left(\frac{n-1}{n}\right)^t \cancel{e^{-\lambda}}$$

$$= \left(\frac{n-1}{n}\right)^t$$

$$\begin{cases} X_i \sim P(\lambda) \\ \sum_{i=1}^{n} X_i \sim P(n\lambda) \\ \sum_{i=2}^{n} X_i \sim P((n-1).\lambda) \\ P(\sum_{i=2}^{n} X_i = t) = \dfrac{e^{-(n-1)\lambda} . ((n-1)\lambda)^t}{t!} \\ P(\sum_{i=1}^{n} X_i = t) = \dfrac{e^{-n\lambda} . (n\lambda)^t}{t!} \end{cases}$$

08) Se $X_1, \ldots, X_n$ são iid com distribuição $N(\xi, \sigma^2)$, com σ conhecido encontre o ENVVUM de ξ^2, ξ^3 e de ξ^4. Sugestão: Para calcular $E(\bar{X}^k)$ escreva $\bar{X} = Y + \xi$ onde $Y \sim N(0, \sigma^2/n)$ e expanda $E[(Y+\xi)^k]$ [Ver Problema 2.4, pg 132, TPE].

1º) Mostrar que $T = \sum X_i$ é estatística suficiente e completa para ξ.

Temos que verossimilhança é dada por

$$p_\theta(\underset{\sim}{x}) = \prod_{i=1}^{n} (\sqrt{2\pi\sigma^2})^{-1} \exp\left\{-\frac{1}{2\sigma^2}(x_i - \xi)^2\right\}$$

$$= \exp\left\{-\frac{\sum x_i^2}{2\sigma^2} + \frac{\xi \sum x_i}{\sigma^2} - \frac{n\xi^2}{2\sigma^2}\right\} (2\pi\sigma^2)^{-n/2}$$

$$= \exp\left\{\frac{\xi}{\sigma^2}\sum x_i - \frac{n\xi^2}{2\sigma^2}\right\} \exp\left(-\frac{\sum x_i^2}{2\sigma^2}\right)(2\pi\sigma^2)^{-n/2}$$

$$= \exp\left\{\eta(\xi)\, T(x) - B(\xi)\right\} h(x)$$

onde $\eta(\xi) = \frac{\xi}{\sigma^2}$, $T(X) = \sum_{i=1}^{n} X_i$, $B(\xi) = \frac{\xi^2}{2\sigma^2}$ e $h(x) = \exp\left(-\frac{\sum x_i^2}{2\sigma^2}\right)(2\pi\sigma^2)^{-n/2}$

O espaço paramétrico é $\{\eta(\xi); \xi \in \mathbb{R}\}$, mas $\eta(\xi) = \frac{\xi}{\sigma^2}$, $\sigma^2 > 0$ logo $\{\eta(\xi); \xi \in \mathbb{R}\} = \mathbb{R} = (-\infty, \infty)$ um aberto de interior não vazio. Por esse fato e pelo fato de que $p_\theta(\underset{\sim}{x})$ pertence a família exponencial temos que $T(X) = \sum_i X_i$ é uma estatística suficiente e completa para ξ.

2º) Encontrar o ENVVUM para ξ^2

Se $X_i \sim N(\xi, \sigma^2) \Rightarrow \bar{X} = \frac{\sum_{i=1}^{n} X_i}{n} \sim N(\xi, \sigma^2/n)$

temos que

$$Var(\bar{X}) = E[\bar{X}^2] - E^2(\bar{X})$$

$$\frac{\sigma^2}{n} = E[\bar{X}^2] - \xi^2$$

$$\Rightarrow E(\bar{X}^2) = \frac{\sigma^2}{n} + \xi^2$$

Logo temos que $\delta_2(X) = \bar{X}^2 - \frac{\sigma^2}{n}$ é um estimador não viciado de ξ^2, ou seja,

$$E[\delta_2(X)] = E(\bar{X}^2) - \frac{\sigma^2}{n}$$

$$= \frac{\sigma^2}{n} + \xi^2 - \frac{\sigma^2}{n}$$

$$= \xi^2$$

mas temos um lema que diz que toda função estimável tem somente um estimador não viciado que é função de uma estatística suficiente e completa.

No nosso caso $\delta_2(X)$ é função da estatística suficiente e completa $T = \sum X_i$. Logo $\delta_2(X)$ é o único estimador não viciado de ξ^2 que depende de T. Por esse fato e usando outro teorema temos que $\delta_2(X)$ é o ENVVUM de ξ^2.

3º) Encontrar o ENVVUM de ξ^3.

Temos que $\bar{X} \sim N(\xi, \sigma^2/n)$ então

$$\mu_r = \begin{cases} 0 & \text{se } r \text{ ímpar} \\ \dfrac{r!\,(\sigma/\sqrt{n})^r}{(r/2)!\,2^{r/2}} & \text{se } r \text{ par} \end{cases}$$

além disso $\mu_r = E[(\bar{X}-\xi)^r]$ $\forall r$ logo

$$\mu_3 = 0 \quad \text{e} \quad \mu_3 = E[(\bar{X}-\xi)^3]$$

daí

$$E[(\bar{X}-\xi)^3] = 0$$

$$E[(\bar{X}^2 - 2\xi\bar{X} + \xi^2)(\bar{X}-\xi)] = 0$$

$$E\left[\bar{X}^3 - \xi\bar{X}^2 - 2\xi\bar{X}^2 + 2\xi^2\bar{X} + \xi^2\bar{X} - \xi^3\right] = 0$$

$$E(\bar{X}^3) - 3\xi E(\bar{X}^2) + 3\xi^2 E(\bar{X}) - \xi^3 = 0$$

$$E(\bar{X}^3) - 3\xi\left\{\frac{\sigma^2}{n} + \xi^2\right\} + 3\xi^2\xi - \xi^3 = 0$$

$$E(\bar{X}^3) - \frac{3\xi\sigma^2}{n} - \cancel{3\xi^3} + \cancel{3\xi^3} - \xi^3 = 0$$

$$E(\bar{X}^3) = +\frac{3\sigma^2}{n}\xi + \xi^3$$

Considere o estimador

$$\delta_3(X) = \bar{X}^3 - \frac{3\sigma^2}{n}\hat{\xi} = \bar{X}^3 - \frac{3\sigma^2}{n}\bar{X}$$

Temos que $\delta_3(X)$ é um estimador não viesado de ξ^3, isto é,

$$E[\delta_3(X)] = E(\bar{X}^3) - \frac{3\sigma^2}{n}E(\bar{X})$$

$$= \cancel{\frac{3\sigma^2}{n}}\xi + \xi^3 - \cancel{\frac{3\sigma^2}{n}}\xi$$

$$= \xi^3$$

Como $\delta_3(X)$ é função de $T = \sum X_i$, pela mesma explicação dada anteriormente temos que $\delta_3(X)$ é o ENVVUM de ξ^3

4º) Encontrar o ENVVUM de ξ^4

Temos que

$$E\left[(\bar{X}-\xi)^4\right] = \frac{4!\,(\sigma/\sqrt{n})^4}{2!\,2^2} = \frac{\cancel{4} \times 3 \times \cancel{2!}}{\cancel{2!}\,\cancel{4}}\,(\sigma^4/n^2) = \frac{3\sigma^4}{n^2} \quad (*)$$

Mas $(\bar{X}-\xi)^4 = (\bar{X}^2 - 2\xi\bar{X} + \xi^2)(\bar{X}^2 - 2\xi\bar{X} + \xi^2)$

$$= \bar{X}^4 - 2\xi\bar{X}^3 + \xi^2\bar{X}^2 - 2\xi\bar{X}^3 + 4\xi^2\bar{X}^2 - 2\xi^3\bar{X}$$
$$+ \xi^2\bar{X}^2 - 2\xi^3\bar{X} + \xi^4$$

Daí

$$E[(\bar{X}-\xi)^4] = E(\bar{X}^4) - 4\xi E(\bar{X}^3) + 6\xi^2 E(\bar{X}^2) - 4\xi^3 E(\bar{X}) + \xi^4$$

$$= E(\bar{X}^4) - 4\xi\left\{\frac{3\sigma^2}{n}\xi + \xi^3\right\} + 6\xi^2\left[\frac{\sigma^2}{n} + \xi^2\right] - 4\xi^4 + \xi^4$$

$$= E(\bar{X}^4) - \frac{12\sigma^2}{n}\xi^2 - 4\xi^4 + \frac{6\xi^2\sigma^2}{n} + 6\xi^4 - 3\xi^4$$

$$= E(\bar{X}^4) - \frac{6\sigma^2\xi^2}{n} - \xi^4$$

Por (*) temos

$$E(\bar{X}^4) - \frac{6\sigma^2\xi^2}{n} - \xi^4 = \frac{3\sigma^4}{n^2}$$

$$E(\bar{X}^4) = \frac{3\sigma^4}{n^2} + \frac{6\sigma^2}{n}\xi^2 + \xi^4$$

$$= \frac{\sigma^2}{n}\left[\frac{3\sigma^2}{n} + 6\xi^2\right] + \xi^4$$

$$= \frac{3\sigma^2}{n}\left[\underbrace{\frac{\sigma^2}{n} + \xi^2}_{} + \xi^2\right] + \xi^4$$

$$= \frac{3\sigma^2}{n}\left\{E(\bar{X}^2) + E(S_2(X))\right\} + \xi^4$$

$$= \frac{3\sigma^2}{n}\left\{E(\bar{X}^2) + E(\bar{X}^2) - \frac{\sigma^2}{n}\right\} + \xi^4$$

$$= \frac{6\sigma^2}{n} E(\bar{X}^2) - \frac{3\sigma^4}{n^2} + \xi^4$$

Considere o estimador

$$\delta_4(X) = \bar{X}^4 - \frac{6\sigma^2}{n}\bar{X}^2 + \frac{3\sigma^4}{n^2}$$

para ξ^4. Este estimador é não viciado para ξ^4, ou seja,

$$E[\delta_4(X)] = E[\bar{X}^4] - \frac{6\sigma^2}{n}E(\bar{X}^2) + \frac{3\sigma^4}{n^2}$$

$$= \cancel{\frac{6\sigma^2}{n}E(\bar{X}^2)} - \cancel{\frac{3\sigma^4}{n^2}} + \xi^4 + \cancel{\frac{3\sigma^4}{n^2}} - \cancel{\frac{6\sigma^2}{n}E(\bar{X}^2)}$$

$$= \xi^4$$

Como $\delta_4(X)$ é função de $T = \sum X_i$, pela mesma explicação dada anteriormente temos que

$\delta_4(X)$ é o ENVVUM de ξ^4.

(9) Resolva o problema 8 admitindo que σ é desconhecido.

Se consideramos σ desconhecido então $N(\xi, \sigma^2)$ pertence a família exponencial biparamétrica e $(\bar{X}, S^2)$ são estatísticas suficientes e completas para (ξ, σ^2) (como já foi mostrado em aula).

Temos que

$$E(\bar{X}) = \frac{1}{n}\sum_{i=1}^{n} E(X_i) = \frac{n\xi}{n} = \xi$$

e sabemos ainda que $S^2 = \frac{\sum(X_i-\bar{X})^2}{n-1}$ é não viesado para σ^2, ou seja,

$$E(S^2) = \sigma^2.$$

Como $\bar{X}$ é estatística suficiente e completa para ξ e S^2 é ancilar para ξ então usando o teorema de Basu temos que $\bar{X}$ e S^2 são independentes.

Sabemos pelo exercício 8 que

$$E(\bar{X}^2) = \frac{\sigma^2}{n} + \xi^2$$

e temos ainda que $E(S^2) = \sigma^2$ logo considere o estimador $\delta_1(X) = \bar{X}^2 - \frac{S^2}{n}$, daí

$$E[\delta_1(X)] = E(\bar{X}^2) - \frac{1}{n}E(S^2)$$

$$= \cancel{\frac{\sigma^2}{n}} + \xi^2 - \cancel{\frac{1}{n}\sigma^2}$$

$$= \xi^2$$

Logo temos que $\delta_1(X) = \bar{X}^2 - \frac{S^2}{n}$ é um estimador não viciado de ξ^2 que é função da estatística suficiente e completa $T = (\bar{X}, S^2)$. Por um lema $\delta_1(X)$ é o único estimador não viciado de ξ^2 que depende de T. Por esse fato e usando outro teorema temos que $\delta_1(X)$ é o ENVVUM de ξ^2

→ Encontrar o ENVVUM de ξ^3

Considere $\delta_2(X) = \bar{X}^3 - \frac{3\bar{X}}{n} S^2$ logo

$$E[\delta_2(X)] = E(\bar{X}^3) - \frac{3}{n} E(\bar{X} S^2)$$

$$= E(\bar{X}^3) - \frac{3}{n} E(\bar{X}) E(S^2)$$

↳ pois $\bar{X}$ e S^2 são independentes

Pelo exercício 8 temos que

$$E(\bar{X}^3) = \frac{3\sigma^2}{n} \xi + \xi^3$$

logo

$$E[\delta_2(X)] = \cancel{\frac{3}{n} \sigma^2 \xi} + \xi^3 - \cancel{\frac{3}{n} \xi \sigma^2} = \xi^3$$

Temos que $\delta_2(X)$ é um estimador não viciado de ξ^3 que depende de $T = (\bar{X}, S^2)$, pela mesma explicação dada anteriormente temos que $\delta_2(X)$ é o ENVVUM de ξ^3.

→ Encontrar o ENVVUM de ξ^4

Considere $\delta_3(X) = \left\{ \bar{X}^4 - \frac{6}{n} \bar{X}^2 S^2 + \frac{3}{n^2} \left(\frac{n-1}{n+1} \right) S^4 \right\}$, daí

$$E(\delta_3(X)) = \left\{ E(\bar{X}^4) - \frac{6}{n} E(\bar{X}^2 S^2) + \frac{3}{n^2} \left(\frac{n-1}{n+1} \right) E(S^4) \right\}$$

$$E[\delta_3(X)] = \left\{ E(\bar{X}^4) - \frac{6}{n} E(\bar{X}^2) E(S^2) + \frac{3}{n^2} \left(\frac{n-1}{n+1} \right) E(S^4) \right\}$$

sabemos que $\frac{(n-1)}{\sigma^2} S^2 \sim \chi^2_{(n-1)}$ logo

$$Var\left[\frac{(n-1)}{\sigma^2} S^2\right] = 2(n-1) \Rightarrow \frac{(n-1)^{\cancel{2}}}{\sigma^4} Var(S^2) = 2\cancel{(n-1)}$$

$$\Rightarrow Var(S^2) = \frac{2\sigma^4}{n-1}$$

Temos que

$$E(S^4) = Var(S^2) + E^2(S^2)$$

$$= \frac{2\sigma^4}{n-1} + \sigma^4$$

$$= \sigma^4\left(\frac{2}{n-1} + 1\right)$$

$$= \sigma^4\left(\frac{n+1}{n-1}\right)$$

Sabemos do exercício 8 que

$$E(\bar{X}^4) = \frac{6}{n}\sigma^2 E(\bar{X}^2) - \frac{3\sigma^4}{n^2} + \xi^4$$

$$= \frac{6}{n}\sigma^2\left(\sigma^2/n + \xi^2\right) - \frac{3\sigma^4}{n^2} + \xi^4$$

$$= \frac{6\sigma^4}{n^2} + \frac{6\sigma^2\xi^2}{n} - \frac{3\sigma^4}{n^2} + \xi^4$$

$$= \frac{3\sigma^4}{n^2} + \frac{6\sigma^2\xi^2}{n} + \xi^4$$

Logo

$$E\{S_3(X)\} = \left\{\frac{3\sigma^4}{n^2} + \frac{6\sigma^2\xi^2}{n} + \xi^4 - \frac{6}{n}\left(\frac{\sigma^2}{n} + \xi^2\right)\sigma^2 + \frac{3}{n^2}\left(\frac{n-1}{n+1}\right)\left(\frac{n+1}{n-1}\right)\sigma^4\right\}$$

$$E[\delta_3(X)] = \frac{3\sigma^4}{n^2} + \frac{6\sigma^2\xi^2}{n} + \xi^4 - \frac{6\sigma^4}{n^2} - \frac{6\sigma^2\xi^2}{n} + \frac{3\sigma^4}{n^2}$$

$$= \xi^4$$

Logo temos que $\delta_3(X)$ é um estimador não viciado de ξ^4 que depende de $T = (\bar{X}, S^2)$ pela mesma explicação dada anteriormente temos que $\delta_3(X)$ é o ENVVUM de ξ^4.

10) Suponha que $X_1, \dots, X_n$ sejam variáveis aleatórias independentes com distribuições $N(\theta_i, 1)$, para $i = 1, \dots, n$. Seja $w = n^{-1} \sum_{i=1}^{n} \theta_i^2$. 21

(a) Obtenha o ENVVUM de w e sua variância

1º) Encontrar estatística suficiente completa para a família $P = \{P_{\theta_i}(x_i), \theta_i \in \mathbb{R}\}$

Seja $\underset{\sim}{X} = (X_1, \dots, X_n)$, logo temos temos a verossimilhança:

$$p_\theta(\underset{\sim}{x}) = \prod_{i=1}^{n} (2\pi)^{-1/2} \exp\left\{-\frac{1}{2}(x_i - \theta_i)^2\right\}$$

$$= (2\pi)^{-n/2} \exp\left\{-\frac{1}{2}\sum_{i=1}^{n}(x_i - \theta_i)^2\right\}$$

$$= \exp\left\{-\frac{1}{2}\sum_{i=1}^{n}\left[x_i^2 - 2x_i\theta_i + \theta_i^2\right]\right\}(2\pi)^{-n/2}$$

$$= \exp\left\{-\frac{1}{2}\sum_{i=1}^{n} x_i^2 + \sum_{i=1}^{n} x_i\theta_i - \frac{1}{2}\sum_{i=1}^{n}\theta_i^2\right\}(2\pi)^{-n/2}$$

$$= \exp\left\{\sum_{i=1}^{n}\theta_i x_i - \frac{1}{2}\sum_{i=1}^{n}\theta_i^2\right\}\exp\left\{-\frac{1}{2}\sum_{i=1}^{n}x_i^2\right\}(2\pi)^{-n/2}$$

$$= \exp\left\{\sum_{i=1}^{n}\eta_i T(x_i) - b(\theta)\right\}h(x)$$

onde $\eta_i = \theta_i$, $T(x_i) = x_i$, $b(\theta) = \frac{1}{2}\sum_{i=1}^{n}\theta_i^2$ e $h(x) = \exp\left\{-\frac{1}{2}\sum_{i=1}^{n}x_i^2\right\}(2\pi)^{-n/2}$.

Considerando a família canônica temos que

$$A(\eta) = \frac{\sum_{i=1}^{n}\eta_i^2}{2} \Rightarrow \exp\left\{\frac{\sum \eta_i^2}{2}\right\}$$

então o espaço paramétrico canônico Δ é dado pelos valores η_i, $i = 1, \dots, n$ tal que $\exp\left\{\frac{1}{2}\sum_{i=1}^{n}\eta_i^2\right\} < \infty$ implicando que $\eta_i \in \mathbb{R}$, $i = 1, \dots, n$

Logo $\Delta = \mathbb{R}^n$, que é um aberto no $\mathbb{R}^n$. Assim contém todo

$\underset{\sim}{\eta} = (\eta_1, \dots, \eta_n)$ em algum retângulo n-dimensional. Logo $p_\theta(\underset{\sim}{x})$ pertence a família exponencial com espaço paramétrico aberto $\Theta = \{\underset{\sim}{\theta} : \underset{\sim}{\theta} \in \mathbb{R}^n\}$

daí, $T(\underset{\sim}{X}) = \{T_1(\underset{\sim}{X}), \dots, T_n(\underset{\sim}{X})\} = \{X_1, \dots, X_n\}$ é uma estatística suficiente e completa para a família $P = \{P_{\theta_i}(x_i), \theta_i \in \mathbb{R}\}$.

2º) encontrar um estimador δ não viciado para $\omega = \frac{\sum \theta_i^2}{n}$ baseado na estatística suficiente e completa $T(\underset{\sim}{X}) = (X_1, ..., X_n)$

Considere um estimador da forma

$$\delta_1(X) = \frac{\sum^n X_i^2}{n}$$

agora vamos verificar se $\delta_1(X)$ é não viciado para ω

$$E[\delta_1(X)] = \frac{1}{n} \sum_{i=1}^{n} E(X_i^2)$$

$$= \frac{1}{n} \sum_{i=1}^{n} \{Var(X_i) + E^2(X_i)\}$$

$$= \frac{1}{n} \sum_{i=1}^{n} \{1 + \theta_i^2\}$$

$$= \frac{n}{n} + \frac{\sum \theta_i^2}{n}$$

$$= 1 + \frac{\sum \theta_i^2}{n}.$$

verificamos que $\delta_1(X)$ é viciado para ω. Mas para obter um estimador δ_2 não viciado para ω basta fazer $\delta_1(X) - 1$, ou seja, considere

$$\delta_2(X) = \frac{\sum_{i=1}^{n} X_i^2}{n} - 1$$

$$E[\delta_2(X)] = \frac{1}{n} \sum E(X_i^2) - 1$$

$$= 1 + \frac{\sum \theta_i^2}{n} - 1$$

$$= \frac{\sum \theta_i^2}{n}$$

Portanto $\delta_2(X) = \frac{\sum X_i^2}{n} - 1$ é um estimador não viciado de ω; além disso é função de $T(\underset{\sim}{X})$. Temos um teorema que diz: Toda função estimável tem um e somente um estimador não viciado que é função de uma estatística suficiente e completa

Daí, temos que δ_2 é o único estimador não viciado, em função de $T(\underset{\sim}{X})$, de ω. Pela unicidade do ENVVUM temos que

$$\delta_2(X) = \frac{\sum_{i=1}^{n} X_i^2}{n} - 1 \quad \text{é o ENVVUM de } \omega = \frac{\sum_{i=1}^{n} \theta_i^2}{n}$$

3º) Encontrar a variância de $\delta_2(X)$

$$Var(\delta_2(X)) = Var\left[\frac{\sum X_i^2}{n} - 1\right]$$

$$= Var\left[\frac{\sum X_i^2}{n}\right]$$

$$= \frac{1}{n^2} Var\left(\sum X_i^2\right)$$

$$\overset{ind}{=} \frac{1}{n^2} \sum Var(X_i^2)$$

mas:

$$Var(X_i^2) = E[X_i^4] - E^2(X_i^2)$$

Sabemos que se $Y \sim N(\mu, \sigma^2) \Rightarrow \mu_r = \frac{r!\, \sigma^r}{(r/2)!\, 2^{r/2}}$ se r é par e além disso, $\mu_r = E[(Y-\mu)^r]$, logo temos que

$$E[(X_i - \theta_i)^4] = \frac{4!}{2!\, 2^2} \Leftrightarrow E[(X_i - \theta_i)^2 (X_i - \theta_i)^2] = \frac{\cancel{4} \times 3 \times \cancel{2}}{\cancel{2} \times \cancel{4}}$$

$$\Leftrightarrow E\left[(X_i^2 - 2\theta_i X_i + \theta_i^2)(X_i^2 - 2\theta_i X_i + \theta_i^2)\right] = 3$$

$$\Leftrightarrow E\left[X_i^4 - 2\theta_i X_i^3 + \theta_i^2 X_i^2 - 2\theta_i X_i^3 + 4\theta_i^2 X_i^2 - 2\theta_i^3 X_i + \theta_i^2 X_i^2 - 2\theta_i^3 X_i + \theta_i^4\right] = 3$$

$$\Rightarrow E[X_i^4] - 4\theta_i E(X_i^3) + 6\theta_i^2 E(X_i^2) - 4\theta_i^3 E(X_i) + \theta_i^4 = 3$$

$$E(X_i^4) - 4\theta_i E(X_i^3) + 6\theta_i^2 (1 + \theta_i^2) - 4\theta_i^3 \theta_i + \theta_i^4 = 3$$

$$E(X_i^4) - 4\theta_i E[X_i^3] + 6\theta_i^2 + 6\theta_i^4 - 3\theta_i^4 = 3$$

$E(X_i^4) = 3 + 4\theta_i E(X_i^3) + 6\theta_i^4 + 3\theta_i^4 \quad (*)$.

Sabemos ainda que $\mu_r = 0$ se r é ímpar logo

$\mu_3 = 0$ e $\mu_3 = E[(X_i - \theta_i)^3] = E[(X_i^2 - 2\theta_i X_i + \theta_i^2)(X_i - \theta_i)]$

$\Rightarrow E[X_i^3 - \theta_i X_i^2 - 2\theta_i X_i^2 + 2\theta_i^2 X_i + \theta_i^2 X_i - \theta_i^3] = 0$.

$\Rightarrow E(X_i^3) - 3\theta_i E(X_i^2) + 3\theta_i^2 E(X_i) + \theta_i^3 = 0$

$E(X_i^3) - 3\theta_i(1 + \theta_i^2) + 3\theta_i^3 + \theta_i^3 = 0$

$E(X_i^3) - 3\theta_i - \cancel{3\theta_i^3} + \cancel{3\theta_i^3} - \theta_i^3 = 0$

$E(X_i^3) = 3\theta_i + \theta_i^3 \quad (**)$

Substituindo $(**)$ em $(*)$

$E(X_i^4) = 3 + 4\theta_i(3\theta_i + \theta_i^3) + 6\theta_i^2 + 3\theta_i^4$

$= 3 + 12\theta_i^2 + 4\theta_i^4 - 6\theta_i^2 + 3\theta_i^4$

$= 3 + 6\theta_i^2 + \theta_i^4$

daí

$Var(X_i^2) = 3 + 6\theta_i^2 + \theta_i^4 - (1 + \theta_i^2)^2$

$= \cancel{3} + 6\cancel{\theta_i^2} + \theta_i^4 - \cancel{1} - 2\theta_i^2 - \theta_i^4$

$= 2 + 4\theta_i^2$.

Assim $Var[S_2(X)] = \frac{1}{n^2}\sum_{i=1}^{n}(2 + 4\theta_i^2) = \frac{2}{n} + \frac{4}{n^2}\sum_{i=1}^{n}\theta_i^2$

(b) Obtenha o EMV de w e seu erro quadrático médio. Qual dos dois estimadores você prefere? Justifique

A log-verossimilhança é dada por:

$$l(\underset{\sim}{\theta}, \underset{\sim}{X}) = \sum_{i=1}^{n} \theta_i x_i - \frac{1}{2}\sum_{i=1}^{n}\theta_i^2 - \frac{1}{2}\sum_{i=1}^{n} x_i^2 - \frac{n}{2}\log(2\pi)$$

a função escore é dada por

$$\frac{\partial l(\theta_i, \underset{\sim}{X})}{\partial \theta_i} = 0 \qquad \frac{\partial l}{\partial \theta_i} = x_i - \theta_i.$$

Daí,

$$\frac{\partial l(\theta, \underset{\sim}{X})}{\partial \theta_i} = \sum_{i=1}^{n} x_i \cancel{- \frac{1}{2}\sum_{i=1}^{n} 2\hat{\theta}_i} = 0$$

$$\Rightarrow \sum_{i=1}^{n} \hat{\theta}_i = \sum_{i=1}^{n} x_i$$

Temos que

$$\frac{\partial^2 l(\underset{\sim}{\theta}, \underset{\sim}{X})}{\partial \theta_i^2} = -1 < 0 \quad \text{(negativo)}$$

Logo o estimador de máxima verossimilhança de $\sum \theta_i$ é $\sum x_i$
Pelo princípio da invariância do EMV segue que o EMV de $w = \frac{\sum \theta_i^2}{n}$ é

$$\hat{w} = \frac{\sum \hat{\theta}_i^2}{n} = \frac{\sum x_i^2}{n}.$$

$$\therefore \boxed{\hat{w} = \frac{\sum x_i^2}{n}}$$

→ calcular o erro quadrático médio de w

$$EQM(\hat{w}) = Var_\theta(\hat{w}) + Viés(\hat{w})$$

$$= Var_\theta(\hat{w}) + E(\hat{w}) - \hat{w}$$

$$= \frac{2}{n} + \frac{4}{n^2}\sum_{i=1}^{n}\theta_i^2 + \frac{1}{n}E\left[\sum_{i=1}^{n}\theta_i^2\right] - \frac{1}{n}\sum_{i=1}^{n}\theta_i^2$$

↳ por (a)

$$= \frac{2}{n} + \frac{4}{n^2}\sum\theta_i^2 + \frac{1}{n}\sum E(X_i^2) - \frac{1}{n}\sum\theta_i^2$$

$$= \frac{2}{n} + \left\{\frac{4}{n} - 1\right\}\frac{\sum\theta_i^2}{n} + \frac{1}{n}\sum(1+\theta_i^2)$$

$$= \frac{2}{n} + \left\{\frac{4}{n} - \cancel{1}\right\}\frac{\sum\theta_i^2}{n} + \frac{n}{n} + \cancel{\frac{\sum\theta_i^2}{n}}$$

$$= \frac{4}{n^2}\sum\theta_i^2 + \frac{2}{n} + 1 > 0$$

→ calcular o erro quadrático médio do ENVVUH $\delta_2(X)$

$$EQM(\delta_2(X)) = Var(\delta_2(X)) + \cancel{Viés(\delta_2(X))} \rightarrow 0$$

$$= \frac{4}{n^2}\sum_{i=1}^{n}\theta_i^2 + \frac{2}{n} > 0$$

Daí temos que

$$EQM[\hat{w}] = EQM[\delta_2(X)] + 1.$$

$$\Rightarrow \quad EQM[\delta_2(X)] < EQM(\hat{w})$$

Um estimador que seria preferível seria

$$\delta_2(X) = \frac{\sum X_i^2}{n} - 1$$

pois possui o menor EQM. Mas este não é um bom estimador pois pode tomar valores negativos enquanto que $w = \frac{\sum \sigma_i^2}{n} > 0$. Logo escolheríamos o EMV de w, ou seja,

$$\hat{w} = \frac{\sum X_i^2}{n}.$$

13) buição na família $\mathcal{P} = \{P_\theta, \theta \in \Omega\}$. Uma estatística $U = U(X)$ é ig-
norável se existir uma estatística suficiente $T = T(X)$ tal que
T e U são independentes. Mostre que se U é uma estatística
ignorável, então U é uma estatística ancilar.

Considere uma estatística suficiente T e U uma estatística ignorável logo pela definição de estatística ignorável temos que T e U são independentes.

Como T é estatística suficiente então $P(U|T)$ não depende de θ, mas

$$P(U|T) = \frac{P(U,T)}{P(T)}$$

$$= \frac{P(U)\,P(T)}{P(T)}$$

→ pela independência de U e T

$$= P(U)$$

Logo $P(U|T) = P(U)$, como $P(U|T)$ não depende de θ então $P(U)$ também não depende de θ.

Assim U é ancilar para θ.

14) Suponha que X_i, para $i = 1, \ldots, n$, sejam v.a's com distribuições $N(\alpha + \beta t_i, \sigma^2)$, onde α, β e σ^2 são desconhecidos, e os t's são constantes conhecidos não todos nulos. Encontre o ENVVUM de α e β

Temos que a verossimilhança é dada por

$$p_\theta(x) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{ -\frac{1}{2\sigma^2} (x_i - (\alpha + \beta t_i))^2 \right\}$$

$$= (2\pi\sigma^2)^{-n/2} \exp\left\{ -\frac{1}{2\sigma^2} \sum_{i=1}^{n} \left[x_i^2 - 2x_i(\alpha + \beta t_i) + \alpha^2 + 2\alpha\beta t_i + \beta^2 t_i^2 \right] \right\}$$

$$= (2\pi\sigma^2)^{-n/2} \exp\left\{ -\frac{1}{2\sigma^2} \sum \left[x_i^2 - 2\alpha x_i - 2\beta x_i t_i + \alpha^2 + 2\alpha\beta t_i + \beta^2 t_i^2 \right] \right\}$$

$$= (2\pi)^{-n/2} \exp\left\{ -n\log\sigma - \frac{\sum x_i^2}{2\sigma^2} + \frac{\alpha}{\sigma^2}\sum x_i + \frac{\beta}{\sigma^2}\sum x_i t_i - \frac{n\alpha^2}{2\sigma^2} - \frac{\alpha\beta}{\sigma^2}\sum t_i - \frac{\beta^2}{2\sigma^2}\sum t_i^2 \right\}$$

$$= (2\pi)^{-n/2} \exp\left\{ \frac{\alpha}{\sigma^2}\sum x_i + \frac{\beta}{\sigma^2}\sum x_i t_i - \frac{\sum x_i^2}{2\sigma^2} - \left[n\log\sigma + \frac{n\alpha^2}{2\sigma^2} + \frac{\alpha\beta}{\sigma^2}\sum t_i + \frac{\beta^2}{2\sigma^2}\sum t_i^2 \right] \right\}$$

$$= \exp\left\{ \eta_1 T_1(X) + \eta_2 T_2(X) + \eta_3 T_3(X) - B(\alpha, \beta, \sigma^2) \right\} h(x)$$

onde $\eta_1 = \frac{\alpha}{\sigma^2}$, $T_1(X) = \sum X_i = n\bar{X}$

$\eta_2 = \frac{\beta}{\sigma^2}$, $T_2(X) = \sum X_i t_i$

$\eta_3 = -\frac{1}{2\sigma^2}$, $T_3(X) = \sum X_i^2$

$B(\alpha, \beta, \sigma^2) = n\log\sigma + \frac{n\alpha^2}{2\sigma^2} + \alpha\beta\sum t_i + \frac{\beta^2}{2\sigma^2}\sum t_i^2$ e $h(x) = (2\pi)^{-n/2}$

A distribuição de X pertence a fam. exp. tridimensional c/ espaço param natural $\Lambda = \mathbb{R} \times \mathbb{R} \times \mathbb{R}_-$. Logo a fam. é de posto completo. Então $T(X) = (n\bar{X}, \sum t_i X_i, \sum X_i^2)$ é estatística suficiente completa p/ $\theta = (\alpha, \beta, \sigma^2)$.

suficientes e completa. Uma forma de encontrar o ENVVUM é calcular a esperança do EMV e corrigir o viés

$$L(\theta;x) = p_\theta(x)$$

$$l(\theta;x) = -\frac{n}{2}\log(2\pi) - \frac{\sum x_i^2}{2\sigma^2} + \frac{\alpha}{\sigma^2} n\bar{x} + \frac{\beta}{\sigma^2}\sum x_i t_i -$$

$$- \frac{\sum(\alpha+\beta t_i)^2}{2\sigma^2} - \frac{n}{2}\log\sigma^2$$

$$\frac{\partial l(\theta;x)}{\partial\alpha} = 0 \Rightarrow \frac{n\bar{x}}{\sigma^2} - \frac{\sum(\alpha+\beta t_i)}{\sigma^2} = 0$$

$$\Rightarrow \hat{\alpha} = \bar{x} - \beta\bar{t}$$

$E(\hat{\alpha}) =$

$$= E(\bar{x}-\beta\bar{t}) = E(\bar{x}) - \beta\bar{t} = \frac{1}{n}\sum(\alpha+\beta t_i) - \beta\bar{t}$$

$$= \frac{\cancel{n}\alpha+\beta\bar{t}\cancel{n}}{\cancel{n}} - \beta\bar{t} = \alpha$$

Por outro lado temos que $g = g(\beta,\sigma^2;x) = l(\hat{\alpha},\beta,\sigma^2;x)$

$$= -\frac{n}{2}\log(2\pi) - \frac{\sum x_i^2}{2\sigma^2} + \frac{\hat{\alpha}}{\sigma^2} n\bar{x} + \frac{\beta}{\sigma^2}\sum x_i t_i - \frac{\sum(\hat{\alpha}+\beta t_i)^2}{2\sigma^2} - \frac{n}{2}\log\sigma^2$$

$$= -\frac{n}{2}\log(2\pi) - \frac{\sum x_i^2}{2\sigma^2} + \frac{(\bar{x}-\beta\bar{t})n\bar{x}}{\sigma^2} + \frac{\beta}{\sigma^2}\sum x_i t_i - \frac{\sum(\bar{x}-\beta(\bar{t}-t_i))^2}{2\sigma^2}$$

$$-\frac{n}{2}\log\sigma^2 = 0$$

$$\Rightarrow \frac{\sum x_i t_i}{\sigma^2} - \frac{\bar{t}n\bar{x}}{\sigma^2} + \frac{\sum(\bar{x}-\hat{\beta}(\bar{t}-t_i))}{\sigma^2}(\bar{t}-t_i) = 0$$

$$\Rightarrow \hat{\beta} = \frac{\sum(t_i-\bar{t})X_i}{\sum(t_i-\bar{t})^2}$$

Quando fizer $E(\hat{\beta}) = \beta$. Assim como $\hat{\alpha}$ e $\hat{\beta}$ são funções de est. suf. completa e são ENV de α e β então $\hat{\alpha}$ e $\hat{\beta}$ são ENVVUM de α e β respectivamente.

16) seja X uma v.a. com fdp

$$p_\theta(x) = P_\theta(X=x) = \frac{a(x)\theta^x}{C(\theta)} \quad , x = 0,1,\dots$$

$\theta > 0$, $a(x) \geq 0$, $\sum_{x=0}^{\infty} a(x)\theta^x = C(\theta) < \infty$

(a) mostre que esta distribuição faz parte da família exponencial unidimensional e mostre que a função geradora de momentos de X é $M_X(u) = C(\theta e^u) / C(\theta)$

Temos que

$$p_\theta(x) = \exp\left\{ \log\left[\frac{a(x)\theta^x}{C(\theta)}\right]\right\}$$

$$= \exp\left\{ \log a(x) + x\log\theta - \log C(\theta)\right\}$$

$$= \exp\left\{ (\log\theta)x - \log C(\theta)\right\} \exp\{a(x)\}$$

$$= \exp\left\{ \eta(\theta)T(x) - B(\theta)\right\} h(x)$$

onde $\eta(\theta) = \log\theta$, $T(x) = x$, $B(\theta) = \log C(\theta)$ e $h(x) = e^{a(x)}$

logo temos que esta distribuição pertence a família exponencial unidimensional.

Como $\eta(\theta) = \log\theta \Rightarrow \theta = e^\eta \Rightarrow A(\eta) = B(e^\eta) = \log C(e^\eta)$

logo

$$p_\theta(x) = \exp\{\eta T(x) - A(\eta)\} h(x)$$

Sabemos que

$$M_X(u) = \frac{\exp\{A(\eta+u)\}}{\exp\{A(\eta)\}} =$$

$$= \frac{\exp\{\log C(e^{\eta+u})\}}{\exp\{\log C(e^\eta)\}}$$

$$= \frac{C(e^\eta e^u)}{C(e^\eta)} \quad \text{mas } e^\eta = \theta$$

$$= C(\theta e^u) / C(\theta)$$

3

b) Mostre que se $a(x) > 0 \; \forall \; x = 0, 1, \dots,$ então θ^r é estimável para qualquer r inteiro positivo e seu estimador não viciado de variância uniformemente mínima baseado em uma amostra $X_1, \dots, X_n$ de observações independentes da distribuição de X é $\delta(T)$, sendo $T = \sum_{i=1}^{n} X_i$ e

$$\delta(t) = \begin{cases} 0, & t = 0, \dots, r-1 \\ \dfrac{A(t-r, n)}{A(t, n)}, & t = r, r+1, \dots \end{cases} \qquad (*)$$

onde $A(t,n)$ é o coeficiente de θ^t na expansão em série de $[C(\theta)]^n$

1º) Mostrar que θ^r é estimável $\forall \; r \in \mathbb{Z}^+$

Devemos encontrar $\delta(X)$ tal que

$$E(\delta(X)) = \theta^r$$

$$\Leftrightarrow \sum_{x=0}^{\infty} \delta(x) \, p_\theta(x) = \theta^r \quad \Leftrightarrow \quad \sum_{x=0}^{\infty} \delta(x) \frac{a(x)\theta^x}{C(\theta)} = \theta^r.$$

$$\Leftrightarrow \sum_{x=0}^{\infty} \delta(x) \, a(x) \, \theta^x = C(\theta) \, \theta^r \;, \quad \text{mas } C(\theta) = \sum_{x=0}^{\infty} a(x) \theta^x < \infty$$

logo

$$\sum_{x=0}^{\infty} \delta(x) a(x) \theta^x = \sum_{x=0}^{\infty} a(x) \, \theta^x \theta^r$$

$$\sum_{x=0}^{\infty} \delta(x) a(x) \theta^x = \sum_{x=0}^{\infty} a(x) \, \theta^{x+r} \qquad , \; r \in \mathbb{Z}^+$$

Faça $k = x + r \Rightarrow x = k - r$ logo

$$\sum_{x=0}^{\infty} \delta(x) a(x) \theta^x = \sum_{k=r}^{\infty} a(k-r) \, \theta^k$$

$$\sum_{x=0}^{r-1} \delta(x) a(x) \theta^x + \sum_{x=r}^{\infty} \delta(x) \, a(x) \, \theta^x = \sum_{k=r}^{\infty} a(k-r) \theta^k$$

vai

$$\underbrace{\delta(x)a(x)}_{\geq 0} = 0 \quad \text{quando} \quad x = 0, 1, \dots, r-1$$

então $\delta(x) = 0$ quando $x = 0, 1, \dots, r-1$ e

$$\delta(x)a(x) = a(x-r) \quad \text{quando} \quad x = r, r+1, \dots$$

ou seja

$$\delta(x) = \begin{cases} \dfrac{a(x-r)}{a(x)} & \text{se } x = r, r+1, \dots \\ 0 & \text{se } x = 0, \dots, r-1 \end{cases}$$

$\therefore$ θ^r é estimável $\forall r \in \mathbb{Z}^+$.

2º) Mostrar que o ENVVUM de θ^r é dado por (*)

Como a distribuição de X pertence à família exponencial, então $T(\underset{\sim}{X}) = \sum X_i$ é estatistita suficiente completa pois $\Omega = \{\theta \in \mathbb{R} : \theta > 0\}$ contém retângulos unidimensionais e com isso a família é de posto completo.

Pelo lema 3.6, pag 105, TPE, temos que

$$P(T=t) = \frac{A(t,n)\,\theta^t}{[C(\theta)]^n}$$

onde $A(t,n)$ é o coeficiente de θ^t na expansão de série de potências de $[C(\theta)]^n$.

Assim, análogo ao caso de uma única observação X temos que para $X_1, \dots, X_n$

$$\delta(T) = \begin{cases} 0 & \text{se} \quad T = 0, 1, \dots, r-1 \\ \dfrac{A(T-r,n)}{A(T)} & \text{se} \quad T = r, r+1, \dots \end{cases} \quad \text{é ENV de } \theta^r$$

$\Rightarrow$ $\delta(T)$ é ENVVUM de θ^r

c) Se $X_1, \dots, X_n$ são observações independentes da distribuição Poisson (θ), $\theta > 0$, utilize o resultado em (b) para encontrar o ENVVUM de θ^r, r inteiro positivo.

Temos que

$$p_\theta(x) = \frac{\theta^x e^{-\theta}}{x!} = \frac{(x!)^{-1}\theta^x}{e^\theta}, \quad x = 0, 1, \dots$$

Faça $a(x) = (x!)^{-1}$ e $c(\theta) = e^\theta = \sum_{x=0}^{\infty} \frac{\theta^x}{x!}$, $x = 0, 1, \dots$

então

$$p_\theta(x) = \frac{a(x)\,\theta^x}{c(\theta)}$$

A distribuição de X pertence a família exponencial. Então $T(X) = \sum X_i$ é estatística suficiente. Além disso $T(X) = \sum X_i$ é completa pois $\Omega = \{\theta \in \mathbb{R} : \theta > 0\}$ contém retângulos unidimensionais e com isso a família é de posto completo.

Pelo item (b), temos que

$$\delta(T) = \begin{cases} 0 & \text{se } T = 0, \dots, r-1 \\ \dfrac{A(T-r, n)}{A(T, n)} & \text{se } T = r, r+1, \dots \end{cases}$$

onde $A(T, n)$ é o coeficiente de θ^t na expansão em série de $(c(\theta))^n$, é o ENVVUM de θ^r.

Como $T = \sum X_i \sim$ Poisson $(n\theta)$ então

$$p_\theta(t) = \frac{(n\theta)^t e^{-n\theta}}{t!}$$

$$= \frac{(t!)^{-1} n^t \theta^t}{e^{n\theta}}$$

Faça $A(t, n) = (t!)^{-1} n^t$ e $[c(\theta)]^n = [e^\theta]^n$

$$p_\theta(t) = \frac{A(t,n)\,\theta^t}{[c(\theta)]^n}$$

Daí,

$$A(t-r, n) = [(t-r)!]^{-1} n^{t-r}$$ e então

$$S(t) = \begin{cases} 0 & t = 0, \dots, r-1 \\ \dfrac{t!}{(t-r)!} \dfrac{1}{n^r} & t = r, r+1, \dots \end{cases}$$

Logo $S(T)$ é ENVVUM p/ θ^r, $r \in \mathbb{Z}^+$

d) Mostre que a distribuição de Poisson truncada a direita em b ($b \geq 1$, inteiro) e parâmetro θ é um caso especial da distribuição de série de potências. (se X tem distribuição de Poisson truncada à direita em b e parâmetro θ, então $P_\theta(X=x) = P_\theta(Y=x \mid Y \leq b)$ $x = 0, 1, \dots, b$, onde $Y \sim \mathcal{P}(\theta)$).

Como $Y \sim \mathcal{P}(\theta)$ então

$$P(Y \leq b) = \sum_{y=0}^{b} \frac{\theta^y e^{-\theta}}{y!}$$

Temos que

$$P(X=x) = P_\theta(Y=x \mid Y \leq b) \qquad x = 0, 1, \dots, b$$

$$= \frac{\theta^x e^{-\theta}}{x!} \frac{1}{P(Y \leq b)} \qquad x = 0, 1, \dots, b$$

$$= \frac{\theta^x \cancel{e^{-\theta}}}{x!} \left(\sum_{y=0}^{b} \frac{\theta^y \cancel{e^{-\theta}}}{y!} \right)^{-1}$$

$$= \frac{\theta^x}{x!} \left[\sum_{y=0}^{b} \frac{\theta^y}{y!} \right]^{-1} \qquad x = 0, 1, \dots, b$$

Faça $a(x) = (x!)^{-1}$ e $C(\theta) = \sum_{y=0}^{b} \frac{\theta^y}{y!}$ onde $x = 0, 1, \dots, b$

Logo a distribuição de Poisson truncada a direita e parâmetro θ é um caso especial de série de potências.

e) Mostre que se X é uma única observação de uma distribuição de Poisson truncada à direita em b (b ≥ 1 inteiro), então θ não é estimável. Esta afirmação contradiz à do item (b)? Por que?

Sabemos que se existe um estimador não viciado de g então g é estimável.

Logo para mostrarmos que θ não é estimável basta mostrarmos que não existe um estimador não viciado de θ.

Temos pelo item (d) que

$$P_\theta(X=x) = \frac{\theta^x}{x!}\left(\sum_{y=0}^{b}\frac{\theta^y}{y!}\right)^{-1} \qquad x = 0,1,\ldots,b$$

então

$$E_\theta[\delta(X)] = \theta \iff \sum_{x=0}^{b}\delta(x)\frac{\theta^x}{x!}\left(\sum_{y=0}^{b}\frac{\theta^y}{y!}\right)^{-1} = \theta$$

$$\iff \sum_{x=0}^{b}\frac{\delta(x)\,\theta^x}{x!} = \theta\sum_{y=0}^{b}\frac{\theta^y}{y!}$$

$$\sum_{x=0}^{b}\frac{\delta(x)}{x!}\theta^x = \sum_{y=0}^{b}\frac{\theta^{y+1}}{y!} = \sum_{y=0}^{b}\frac{(y+1)\,\theta^{y+1}}{(y+1)!}$$

Faça $u = y+1$, se $y=0 \Rightarrow u=1$, se $y=b \Rightarrow u=b+1$ logo

$$\sum_{x=0}^{b}\frac{\delta(x)\,\theta^x}{x!} = \sum_{u=1}^{b+1}\frac{u\,\theta^u}{u!}$$

somei zero.

$$\iff \sum_{x=0}^{b}\frac{\delta(x)}{x!}\theta^x = \sum_{u=0}^{b+1}\frac{u\,\theta^u}{u!} \qquad (*)$$

Observamos que do lado esquerdo de (*) há um polinômio de grau b+1 em θ enquanto que do lado direito há um polinômio de grau b+2 em θ.

$\therefore \nexists\ \delta(X)$ que satisfaça $(*)$

Logo $\nexists$ ENV de θ, o que implica que θ não é estimável

No item (d) vimos que a distribuição de Poisson truncada à direita em b e parâmetro θ é um caso especial da dist. de série de potências onde $a(x) = (x!)^{-1} > 0$.

Logo isto contradiz a afirmação do item (b):

"Se $a(x) > 0\ \forall\ x = 0, 1, \ldots$, então θ^r é estimável para $\forall\ r \in \mathbb{Z}^+$"

17) Seja $(X_1, \dots, X_n)$ uma amostra aleatória de tamanho n de uma distribuição de Bernoulli de parâmetro θ $(0 \leq \theta \leq 1)$. Seja $H = (\sum X_i / n)^2$ o EMV de θ^2. Seja 38

$$H^* = nH - \frac{(n-1)}{n} \sum H^{(i)}$$

o estimador "jackknife" de θ^2, onde $H^{(i)}$ é calculado como H exceto que X_i é excluída da amostra.

(a) mostre o EMV de θ^2 é um estimador viciado mas H^* é não viciado.

1º) mostrar que o EMV de θ^2 é viciado

$$E[H] = E\left\{ \frac{\sum X_i}{n} \right\}^2$$

$$= \frac{1}{n^2} E[(\sum X_i)^2]$$

mas

$$\left(\sum_{i=1}^{n} X_i\right)^2 = \sum X_i^2 + \sum_{i=1}^{n} \sum_{\substack{j=1 \\ j \neq i}}^{n} X_i X_j$$

logo

$$E[(\sum X_i)^2] = E[\sum X_i^2] + E\left[\sum_{i=1}^{n} \sum_{\substack{j=1 \\ j \neq i}}^{n} X_i X_j\right]$$

$$= \sum_{i=1}^{n} E(X_i^2) + \sum_{i=1}^{n} \sum_{\substack{j=1 \\ j \neq i}}^{n} E(X_i X_j)$$

como $X_1, \dots, X_n$ são Bernoulli temos

$$\sum_{i=1}^{n} X_i^2 = \sum_{i=1}^{n} X_i$$

temos ainda que X_i's são independentes logo

$$E(X_i X_j) = E(X_i) E(X_j)$$

daí,

$$E(H) = \frac{1}{n^2}\left\{\sum_{i=1} E(X_i) + \sum_{i=1}^{n}\sum_{\substack{j=1\\ j\neq i}}^{n} E(X_i)E(X_j)\right\}$$

$$= \frac{1}{n^2}\left\{\sum_{i=1}^{n}\theta + \sum_{i=1}^{n}\sum_{\substack{j=1\\ j\neq i}}^{n}\theta^2\right\}$$

$$= \frac{n\theta}{n^2} + \frac{n(n-1)\theta^2}{n^2} = \frac{\theta}{n} + \frac{(n-1)\theta^2}{n}$$

$$\Rightarrow \quad E(H) = \frac{n\theta^2 - \theta^2 + \theta}{n} \neq \theta^2$$

$\therefore$ H é viciado para θ^2.

2º) mostrar que H^* não é viciado para θ^2

$$E[H^*] = n\,E(H) - \frac{(n-1)}{n}\sum_{i=1}^{n} E[H^{(i)}]$$

$$= n\theta^2 - \theta^2 + \theta - \frac{(n-1)}{n}\sum_{i=1}^{n} E[H^{(i)}]$$

Como $H^{(i)}$ é calculada como H exceto que X_i é excluída da amostra

e $E(H) = \frac{\theta + (n-1)\theta^2}{n}$ então para $i = 1, \dots, n$

$$E[H^{(i)}] = \frac{\theta + (n-2)\theta^2}{n-1}$$

Daí,

$$E[H^*] = \theta + (n-1)\theta^2 - \frac{\cancel{(n-1)}}{n}\sum_{i=1}^{n}\frac{[\theta + (n-2)\theta^2]}{\cancel{n-1}}$$

$$= \cancel{\theta} + (n-1)\theta^2 - \frac{\cancel{n\theta}}{n} - \frac{\cancel{n}(n-2)\theta^2}{\cancel{n}}$$

$$= \cancel{n\theta^2} - \theta^2 - \cancel{n\theta^2} + 2\theta^2$$

$$= \theta^2$$

$\therefore H^*$ é não viciado para θ^2

b) H^* é o ENVVUM de θ^2? Se for, prove. Se não for, encontre o ENVVUM

1º) Mostrar que $T(\underset{\sim}{X}) = \sum_{i=1}^{n} X_i$ é uma estatística suficiente completa para θ.

Temos que

$$p_\theta(x) = \theta^{\sum x_i} (1-\theta)^{n-\sum x_i}$$

$$= \theta^{n\bar{x}} (1-\theta)^{n-n\bar{x}}$$

$$= \left(\frac{\theta}{1-\theta}\right)^{n\bar{x}} (1-\theta)^n$$

$$= \exp\left\{ n\bar{x} \log\left(\frac{\theta}{1-\theta}\right) + n\log(1-\theta) \right\}$$

$$= \exp\left\{ \eta(\theta) T(x) - B(\theta) \right\} h(x)$$

onde $\eta(\theta) = \log\left(\frac{\theta}{1-\theta}\right)$, $T(x) = n\bar{X}$ e $B(\theta) = -n\log(1-\theta)$ e $h(x) = 1$

$\Rightarrow e^\eta = \frac{\theta}{1-\theta} \Rightarrow \theta = \frac{e^\eta}{1+e^\eta} \Rightarrow 1-\theta = \frac{1}{1+e^\eta} \Rightarrow A(\eta) = n\log(1+e^\eta)$

$\therefore$ Esta distribuição pertence a família exponencial uniparamétrica e o espaço paramétrico $\{\eta(\theta),\ \theta \in (0,1)\} = \mathbb{R}$ um aberto de interior não vazio contendo retângulos unidimensionais logo a família é de posto completo

$\therefore T = \sum x_i = n\bar{X}$ é est. suf. completa.

Note que $H = \left(\frac{\sum X_i}{n}\right)^2 = \left(\frac{T(X)}{n}\right)^2$ e

$$H^{(i)} = \frac{(\sum X_j - X_i)^2}{(n-1)^2} = \frac{(T(X))^2 - 2X_i T(X) + X_i^2}{(n-1)^2} \Rightarrow \sum H^{(i)} = n\frac{(T(X))^2 - 2T(X)\sum X_i + \sum (X_i)^2}{(n-1)^2}$$

mas $\sum X_i^2 = \sum X_i$ logo

$$\sum_{i=1}^{n} H^{(i)} = \frac{n[T(X)]^2 - 2[T(X)]^2 + T(X)}{(n-1)^2}$$

Assim como H^* é ENV de θ^2 podemos escrevê-lo como

$$H^* = n\left(\frac{T(x)}{n}\right)^2 - \frac{\cancel{(n-1)}}{n}\,\frac{(n-2)[T(x)]^2 + T(x)}{(n-1)^{\cancel{2}}} = \frac{[T(x)]^2}{n} - \frac{(n-2)[T(x)]^2 + T(x)}{n(n-1)}$$

ou seja,

$$H^* = \frac{[T(x)]^2 - T(x)}{n(n-1)}$$

Como H^* é não viciado p/ θ^2 e é função da estatística suficiente completa $T(X)$ então H^* é ENVVUM de θ^2.

18) Seja $(X_1, \ldots, X_n)$ uma amostra casual simples de $X \sim$ Exponencial deslocada, i.e., com fdp dada por.

$$f(x;a) = \exp\{-(x-a)\} I_{(a,\infty)}(x) \qquad a \in (-\infty, \infty)$$

(a) mostre que $X_{(1)} = \min_{i \in \{1,\ldots,n\}} X_i$ é uma estatística suficiente e completa.

Temos que a verossimilhança é dada por

$$f(\underset{\sim}{x}; a) = \exp\{-\sum_{i=1}^{n}(x_i - a)\} \prod_{i=1}^{n} I_{(a,\infty)}(x) \qquad a \in (-\infty; \infty)$$

$$= \exp\{na - \sum x_i\} I_{(a,\infty)}(x_{(1)}) \qquad \text{onde } x_{(1)} = \min\{x_1, \ldots, x_n\}$$

$$= \exp\{na\} I_{(a,\infty)}(x_{(1)}) \exp\{-\sum_{i=1}^{n} x_i\}$$

$$= g_a(T(x)) h(x)$$

onde $g_a(T(x)) = e^{na} I_{(a,\infty)}(x_{(1)})$ e $h(x) = e^{-\sum_{i=1}^{n} x_i}$. Pelo critério da fatoração

$T(X) = X_{(1)}$ é estatística suficiente para $\underline{\underline{a}}$

→ mostrar que $T(X) = X_{(1)}$ é estatística suficiente completa para $\underline{\underline{a}}$.

Ou seja, se $E[f(T)] = 0 \Rightarrow f(T) \equiv 0 \quad \forall \theta$

mas precisaremos da distribuição de $X_{(1)}$ logo

$$P(X_{(1)} \leq x) = 1 - P(X_{(1)} > x)$$

$$= 1 - \prod_{i=1}^{n} P(X_i > x)$$

$$= 1 - [P(X_1 > x)]^n$$

mas

$$P(X_1 \leq x) = \int_a^x f(t;a)\,dt$$

$$= \int_a^x e^{a-t}\,dt$$

$$= \frac{e^{a-t}}{-1}\Big|_a^x = -e^{a-x} + e^{a-a} = 1 - e^{a-x} \qquad x > a.$$

Logo

$$P(X_{(1)} \leq x) = 1 - [\cancel{1} - (\cancel{1} - e^{a-x})]^n$$

$$= 1 - [e^{a-x}]^n$$

$$= 1 - e^{n(a-x)} \qquad , \; x > a$$

Daí,

$$f_{X_{(1)}}(x) = \frac{d\,P(X_{(1)} \leq x)}{dx}$$

$$= \frac{d}{dx}\left\{1 - e^{n(a-x)}\right\}$$

$$= -e^{n(a-x)}\,(-n)$$

$$= n\,e^{n(a-x)} \qquad c/\; x > a$$

ou seja

$$f_{X_{(1)}}(x) = n\,e^{-n(x-a)}\,I_{(a,\infty)}(x)$$

Logo

$X_{(1)} \sim E(a, 1/n)$ exponencial dupla.

Agora mostraremos que $X_{(1)}$ é completa

$$E[f(T)] = 0 \iff \int_a^{\infty} f(t)\, n\, e^{-n(t-a)}\, dt = 0$$

$$\iff \int_a^{\infty} f(t)\, e^{-n(t-a)}\, dt = 0 \qquad \forall \quad a \in \mathbb{R}$$

que é equivalente a

$$\int_B f(t)\, e^{-n(t-a)}\, dt = 0$$

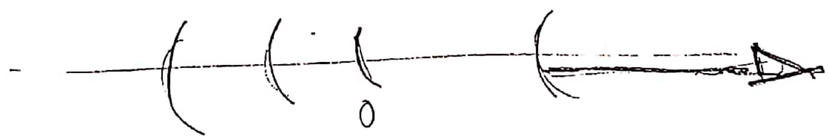
onde $B \in \mathcal{B}$ e $\mathcal{B}$ são todos os borelianos do intervalo $(a, +\infty)$.

Logo

$$f(t)\, e^{-n(t-a)} = 0$$

mas $e^{-n(t-a)} > 0$ então $f(T) = 0$ q.c.-P.

$\therefore$ $T(X) = X_{(1)}$ é estatística suficiente completa.



$E(a, b)$ $\quad f(x; a, b) = \frac{1}{b} \exp\left\{-\frac{(x-a)}{b}\right\} I_{[a,+\infty)}(x)$

b) Mostre que $X_{(1)}$ e S^2, a variância amostral, são variáveis aleatórias independentes.

Teorema de Basu: Se T é uma estatística suficiente completa para família $P = \{P_\theta, \theta \in \Omega\}$ então qualquer estatística ancilar V é independente de T.

Sabemos pela letra a) que $T(X) = X_{(1)}$ é estatística suficiente e completa para $\underline{\underline{a}}$ então se mostrarmos que S^2 é ancilar para $\underline{\underline{a}}$ teremos pelo teorema de Basu que $X_{(1)}$ e S^2 são independentes.

→ Mostrar que $S^2 = \dfrac{\sum_{i=1}^{n}(X_i-\bar{X})^2}{n-1}$ é uma estatística ancilar para $\underline{\underline{a}}$

Temos que a distribuição da exponencial deslocada pertence a família de locação, onde $U \sim Exp(1)$. Logo

$$P(X \le x) = F_U(x-a)$$

onde $U = X - a$. Então

$$S^2 = \frac{\sum^{n}(X_i-\bar{X})^2}{n-1}$$

$$= \frac{\sum_{i=1}^{n}(X_i - a - (\bar{X}-a))^2}{n-1}$$

$$= \frac{\sum_{i=1}^{n}(U_i - \bar{U})^2}{n-1}$$

onde $\bar{U} = \dfrac{\sum_{i=1}^{n} U_i}{n} = \dfrac{\sum_{i=1}^{n}(X_i - a)}{n} = \dfrac{\sum_{i=1}^{n} X_i - na}{n} = \dfrac{\sum_{i=1}^{n} X_i}{n} - a$

$$\bar{U} = \bar{X} - a.$$

Temos então:

$$S^2 = \frac{1}{n-1}\left\{\sum_{i=1}^{n}(U_i-\bar{U})^2\right\}$$ que sua dist. independe de θ.

S^2 é ancilar para $\underline{\underline{a}}$

$\therefore$ Pelo teorema de Basu temos que $X_{(1)}$ e S^2 são variáveis aleatórias independentes.

19)

Analogamente ao exercício 18 temos que

$X_{(1)}$ é uma estatística suficiente completa e $X_{(1)} \sim E(a, b/n)$

Para mostrar que $X_{(1)}$ e $\sum [X_i - X_{(1)}]$ são independentes temos que usar o teorema de Basu, para isso devemos mostrar que

$\sum [X_i - X_{(1)}]$ é ancilar

Já que $X_{(1)}$ é estatística suficiente.

Considerando a sugestão faça $a = 0$ e $b = 1$ e $Y_i \sim E(0,1)$

onde $Y_i = (n-i+1)\ [X_{(i)} - X_{(i-1)}]$

Note que

$Y_2 = (n-1)\ [X_{(2)} - X_1] = X_{(1)} - nX_{(1)}$

$Y_3 = (n-2)[X_{(3)} - X_{(2)}] = X_{(2)}$

$\vdots$

$Y_n = X_{(n)} - X_{(n-1)}$

logo $\sum_{i=1}^{n} Y_i = \sum X_{(i)} - nX_{(1)} = \sum_{i=2}^{n} X_i - nX_{(1)} = \sum_{i=1}^{n} [X_i - X_{(1)}]$

Logo $$\sum_{i=2}^{n} Y_i = \sum_{i=1}^{n} [X_i - X_{(1)}]$$

mas $Y_i \sim E(0,1) \Rightarrow \sum_{i=2}^{n} Y_i \sim \text{Gama}(n-1, 1)$.

$\therefore \sum_{i=1}^{n} [X_i - X_{(1)}] \sim \text{Gama}(n-1, 1)$ independe de a

$\Rightarrow \sum [X_i - X_{(1)}]$ é ancilar

$n X_{(1)}$

2º) Mostrar que não existe estimador não viciado de σ^2 quando μ é desconhecido.

Como temos somente uma observação, então qualquer estimador de σ^2 será função de X.

Suponha por contradição que existe $\delta(X)$ tal que

$$E[\delta(X)] = \sigma^2 \qquad \forall \mu$$

ou seja,

$$E[\delta(X) - \sigma^2] = 0$$

Mas $T = X$ é estatística suficiente e completa, então $\forall \mu$

$$E[\delta(X) - \sigma^2] = 0 \Rightarrow \delta(X) - \sigma^2 = 0 \qquad \forall \mu$$

$\Rightarrow \quad \delta(X) = \sigma^2$ que é um absurdo (o estimador ser igual ao estimando).

Portanto não existe estimador não viciado de σ^2 quando μ é desconhecido.

20) Seja $X \sim N(\mu, \sigma^2)$. Mostre que não existe estimador não viciado de σ^2 quando μ é desconhecido. (sugestão: para $\sigma^2 = a$ fixado, mostre que X é uma estatística suficiente completa para μ.

1º) Mostrar que X é uma estatística suficiente e completa para μ quando $\sigma^2 = a$ fixado.

Temos que a fção densidade de X é

$$p_\mu(x) = \frac{1}{\sqrt{2\pi\sigma^2}} \exp\left\{-\frac{1}{2\sigma^2}(x-\mu)^2\right\}$$

$$= \frac{1}{\sqrt{2\pi a}} \exp\left\{-\frac{1}{2a}(x^2 - 2\mu x + \mu^2)\right\}$$

$$= \frac{1}{\sqrt{2\pi a}} \exp\left\{-\frac{x^2}{2a} + \frac{\mu x}{a} - \frac{\mu^2}{2a}\right\}$$

$$= \exp\left\{\frac{\mu}{a}x - \frac{\mu^2}{2a}\right\} \exp\left\{-\frac{x^2}{2a}\right\} \frac{1}{\sqrt{2\pi a}}$$

$$= \exp\left\{\eta(\mu)T(x) - B(\mu)\right\} h(x)$$

onde $\eta(\mu) = \frac{\mu}{a}$, $T(x) = x$, $B(\mu) = \frac{\mu^2}{2a}$ e $h(x) = \exp\left[-\frac{x^2}{2a}\right]\frac{1}{\sqrt{2\pi a}}$

Temos que o espaço paramétrico é $\{\eta(\mu) : \mu \in \mathbb{R}\}$. Mas $\eta(\mu) = \frac{\mu}{a}$, $a > 0$ então $\{\eta(\mu) : \mu \in \mathbb{R}\} = \mathbb{R} = (-\infty, \infty)$.

Logo temos que $p_\mu(x)$ pertence a família exponencial uniparamétrica e o espaço paramétrico é os $\mathbb{R} = (-\infty, \infty)$ de interior não vazio então $T(X) = X$ é estatística suficiente e completa para μ quando $\sigma^2 = a$ fixado. Daí,

$$E_\mu[f(T)] = 0 \Rightarrow f(T) = 0, \quad q.c.-\mathcal{P}. \quad (*)$$