UFRJ - CCMN - IM - Departamento de Métodos Estatísticos
Probabilidade e Estatística - Estatística

Prova Final - Gabarito 04-12-2014

Atenção: Não serão aceitas respostas sem justificativa: as expressões que levaram a alguma resposta numérica devem ser indicadas nos espaços apropriados.

---

1. A proporção de álcool em certo produto pode ser considerada uma variável aleatória X com a seguinte função densidade de probabilidade:

$$f(x) = \begin{cases} 20x^3(1-x) & \text{se } 0 < x < 1 \\ 0 & \text{caso contrário} \end{cases}$$

(a) Calcule o valor esperado e a variância de X.

(b) Suponha que o preço de venda do produto dependa da proporção de álcool. Se $1/3 < X < 2/3$ o produto é vendido por A_1 reais/l, caso contrário por A_2 reais/l. Expresse em função de A_1, A_2, eB a distribuição de probabilidades do lucro líquido por litro de produto, supondo que o custo por litro é de B reais.

2. Suponha que o peso X de uma pessoa selecionada ao acaso de uma certa população distribui-se normalmente com média μ e variância σ^2. Suponha também que $P(X < 80) = 1/2$ e $P(X < 60) = 1/4$.

(a) Obtenha μ e σ^2 e calcule $P(X \geq 100)$.

(b) Dentre todas as pessoas pesando no mínimo 100 quilos, qual percentual pesará mais de 120 quilos?

(c) Tomamos uma amostra de 5 pessoas dessa população. Qual a probabilidade de que exatamente 3 delas pesem mais do que 60 quilos?

3. Seja X uma v.a. com média μ e variância σ^2. Seja $X_1, \ldots, X_n$ uma amostra aleatória de X. A partir desta amostra aleatória, podemos construir o estimador $\widetilde{X} := \frac{1}{n-1} \sum_{i=1}^{n} X_i$ para a média.

(a) Calcule o viés deste estimador.

(b) Calcule o erro quadrático médio deste estimador.

(c) Ao comparar $\widetilde{X}$ e $\bar{X}$, qual deles é mais indicado para estimar μ? Justifique.

4. Suponha que um professor do curso de engenharia, com objetivo de estudar o tempo, em anos, que os alunos levam para concluírem o curso, selecionou aleatoriamente 144 ex-alunos. Nessa amostra observou-se que a soma dos tempo gastos até a conclusão do curso foi de 750 anos e que a variância amostral teve valor 4 anos^2. Resolva os itens abaixo:

(a) Determine o intervalo de 96% de confiança para a média do tempo que os alunos do curso de engenharia levam até concluírem o curso.

(b) Suponha que o professor resolvesse construir um intervalo de 99% de confiança para essa média, mas com a mesma amplitude do intervalo obtido no item (a). Obtenha um valor aproximado para o novo tamanho amostral que seria necessário para construir tal intervalo.

(c) Sabendo que desses 144 ex-alunos, 36 conseguiram se formar no prazo ideal, determine o intervalo de 93% de confiança, utilizando a abordagem conservadora, para a proporção de alunos do curso de engenharia que conseguem se formar no prazo considerado ideal.

Gabarito:

Questão 1:

(a)

$$\begin{aligned}
E(X) &= \int_0^1 x[20x^3(1-x)]dx = \int_0^1 20x^4(1-x)dx = \frac{2}{3} \\
V(X) &= E(X^2) - E^2(X) \\
E(X^2) &= \int_0^1 x^2[20x^3(1-x)]dx = \int_0^1 20x^5(1-x)dx = \frac{10}{21} \\
&\downarrow \\
V(X) &= \frac{10}{21} - \left(\frac{2}{3}\right)^2 = \frac{2}{63} = 0,0317
\end{aligned}$$

(b) Calculando a probabilidade:

$$\begin{aligned}
P\left(\frac{1}{3} < X < \frac{2}{3}\right) = P\left(X \in (\frac{1}{3}, \frac{2}{3})\right) &= \int_{\frac{1}{3}}^{\frac{2}{3}} [20x^3(1-x)]dx = 0.4156 \\
P\left(X \notin (\frac{1}{3}, \frac{2}{3})\right) &= 1 - P\left(X \in (\frac{1}{3}, \frac{2}{3})\right) = 1 - 0.4156 = 0,5844
\end{aligned}$$

Assim o Lucro líquido por litro de produto será uma variável aleatória de modo que

$$L = \begin{cases} A_1 - B & \text{com probabilidde 0,4156} \\ A2 - B & \text{com probabilidde 0,5844} \end{cases}$$

Questão 2:

(a) $X \sim N(\mu, \sigma^2)$. $\quad P(X < 80) = \frac{1}{2} \Rightarrow \mu = 80$

$P(X < 60) = 0.25 \Rightarrow P\left(Z < \frac{60-80}{\sigma}\right) = 0.25 \Rightarrow \frac{-20}{\sigma} = -0.68 \Rightarrow \sigma = 29.65$

$P(X \geq 100) = P\left(Z \geq \frac{100-80}{29.65}\right) = P(Z \geq 0.675) \approx P(Z \geq 0.68) = 0.25$

(b) $P(X > 120 \mid X > 100) = \frac{P(X>120,\ X>100)}{P(X>100)} = \frac{P(X>120)}{0.25} = \frac{P(Z>1.35)}{0.25} = \frac{0.0885}{0.25} = 0.354$

(c) $Y \sim \text{Binomial}(5, 0.75)$; $\quad P(Y = 3) = \binom{5}{3} 0.75^3 0.25^2 = 0.264$

Questão 3:

(a) Calculando a esperança,

$$E(\widetilde{X}) = E\left(\frac{1}{n-1}\sum_{i=1}^{n} E(X_i)\right) = \frac{1}{n-1}\sum_{i=1}^{n} E(X_i) = \frac{1}{n-1}\sum_{i=1}^{n} \mu = \frac{n\mu}{n-1}$$

Como o viés $B(\widetilde{X}) = E(\widetilde{X} - \mu) = E(\widetilde{X}) - \mu$, obtemos que

$$B(\widetilde{X}) = \frac{n\mu}{n-1} - \mu = \frac{\mu}{n-1}$$

(b) Sabemos que $EQM(\widetilde{X}) = E[(\widetilde{X} - \mu)^2] = Vaar(\widetilde{X}) + [B(\widetilde{X})]^2$

$$Var(\widetilde{X}) = Var\left(\frac{1}{n-1}\sum_{i=1}^{n} X_i\right) = \frac{1}{(n-1)^2} Var\left(\sum_{i-1}^{n} X_i\right) \overset{independência}{=} \frac{1}{(n-1)^2}\sum_{i-1}^{n} Var\left(X_i\right) = \frac{n\sigma^2}{(n-1)^2}$$

Por outro lado, do item anterior,

$$B(\widetilde{X}) = \frac{\mu}{n-1}.$$

Segue que

$$EQM(\widetilde{X}) = \frac{n\sigma^2 + \mu^2}{(n-1)^2}.$$

(c) Lembrando que $B(\bar{X}) = 0$ e $EQM(\bar{X}) = \frac{\sigma^2}{n}$, concluímos que $\bar{X}$ é um estimador melhor que $\widetilde{X}$ para a média populacional μ (segundo o critério do erro quadrático médio) pois para todo $n \geq 2$,

$$EQM(\widetilde{X}) > EQM(\bar{X}).$$

Questão 4:

(a) Seja $(X_1, ..., X_{144})$ a amostra coletada. Temos $\bar{X} = \sum_{i=1}^{144} \frac{X_i}{144} \cong 5,208$ anos. Como o tamanho amostral é grande (maior que 30), será construido o intervalo utilizando aproximação para normal.

$$IC(\mu, 1-\alpha) = \bar{X} \pm Z_{\left(1-\frac{\alpha}{2}\right)} \frac{S}{\sqrt{n}} \quad \Rightarrow \quad IC(\mu, 96\%) = 5,208 \pm Z_{\left(1-\frac{0,04}{2}\right)} \frac{\sqrt{4}}{\sqrt{144}} = 5,208 \pm Z_{(0,98)} \frac{2}{12} =$$

$$= 5,208 \pm 2,05 \times 0,1667 = 5,208 \pm 0,3417 = [4,8663 \; ; \; 5,5497]$$

(b) Para ter mesma amplitude é preciso que o termo após o sinal de $\pm$ tenha o mesmo valor de antes, então:

$$Z_{\left(1-\frac{0,01}{2}\right)} \frac{\sqrt{4}}{\sqrt{n}} = 0,3417 \quad \Leftrightarrow \quad Z_{(0,995)} \frac{2}{\sqrt{n}} = 0,3417 \quad \Leftrightarrow \quad 2,575 \times \frac{2}{\sqrt{n}} = 0,3417 \quad \Leftrightarrow \quad 5,15 = 0,3417\sqrt{n}$$

$$\sqrt{n} = \frac{5,15}{0,3417} \quad \Rightarrow \quad n = 15,0717^2 = 227,1561$$

Como n precisa ser inteiro, segue que o tamanho amostral mínimo será n = 228.

(c) Como o intervalo será feito pela abordagem conservadora, vamos substituir $p(1-p)$ na expressão do intervalo por $\frac{1}{4}$. Com isso temos:

$$IC(p, 1-\alpha) = \hat{p} \pm Z_{\left(1-\frac{\alpha}{2}\right)} \sqrt{\frac{p(1-p)}{n}} \quad \Rightarrow \quad IC(p, 93\%) = \frac{36}{144} \pm Z_{\left(1-\frac{0,07}{2}\right)} \sqrt{\frac{1}{4 \times 144}} = 0,25 \pm Z_{(0,965)} \sqrt{\frac{1}{576}} =$$

$$= 0,25 \pm 1,81 \times \frac{1}{24} = 0,25 \pm 0,0754 = [0,1746 \; ; \; 0,3254]$$