UFRJ - CCMN - IM - Departamento de Métodos Estatísticos

Avaliação Final de Probabilidade e Estatística **29-11-2019**
Atenção: Não serão aceitas respostas sem justificativa.
Resolver as questões nos espaços apropriados.

Q1) O custo de fábrica de um determinado produto é R$500,00 e ele é vendido por R$1.000,00 em uma certa loja. Quando um cliente compra esse produto, ele tem direito a 6 meses de garantia e caso o produto apresente defeito nesse prazo, ele trocará por um novo (gerando um prejuízo de R$500,00 para a loja). Assuma que essa troca só pode ser feita uma vez por cliente e considere que o tempo até esse produto apresentar defeito (em anos) é uma variável aleatória X, cuja função densidade de probabilidade é apresentada abaixo:

$$f(x) = \begin{cases} \frac{x^2}{2}, & \text{para } 0 \leq x \leq 1 \\ \frac{13-3x}{20}, & \text{para } 1 < x \leq \frac{13}{3} \\ 0, & \text{caso contrário} \end{cases}$$

(a) Qual o valor esperado do lucro para a loja com a venda de uma unidade desse produto?

(b) Considere que a loja oferecerá a opção do cliente pagar um adicional de R$200,00 por uma garantia extendida, que aumenta em 1 ano o prazo para troca em caso de defeito. O valor esperado do lucro para a loja com a venda de uma unidade com garantia extendida será maior ou menor do que com a garantia padrão? Justifique.

(c) Calcule qual tempo adicional máximo que a loja pode oferecer na garantia extendida de modo que essa opção não leve a um lucro esperado inferior ao de uma venda sem garantia extendida.

Q2) Admita que o atendimento, em minutos, num caixa de um banco seja uma v.a. T, com densidade

$$f(t) = 0,25te^{-0,5t},\ t \geq 0 \ \text{ e } \ f(t) = 0, \quad t < 0.$$

(a) Mostre que, de fato, $f(t)$ é uma função de densidade.

(b) Calcule a função de distribuição acumulada do tempo de atendimento, $F(t)$. Calcule também $P(T > 2)$.

(c) Suponha agora que o tempo de atendimento do outro caixa, X, seja uma v.a. com densidade

$$f(x) = 2e^{-2x},\ x \geq 0 \ \text{ e } \ f(x) = 0, \quad x < 0.$$

Calcule $P(X > 2)$ e $P(2 < X \leq 3)$.

Q3) Para estimar a média μ de uma população, foram propostos três estimadores não viesados, $\hat{\mu}_1$, $\hat{\mu}_2$ e $\hat{\mu}_3$, obtidos a partir de três amostras independentes, tal que $Var(\hat{\mu}_1) = 2Var(\hat{\mu}_2) = 3Var(\hat{\mu}_3)$. Considere os seguintes estimadores ponderados de μ:

$$T_1 = \frac{\hat{\mu}_1 + 2\hat{\mu}_2 + 3\hat{\mu}_3}{6},\ T_2 = \frac{3\hat{\mu}_1 + 2\hat{\mu}_2 - \hat{\mu}_3}{4} \text{ e } T_3 = \frac{\hat{\mu}_1 + \hat{\mu}_2}{2}$$

(a) Determine se os estimadores T_1, T_2 e T_3 são ou não viesados.

(b) Qual estimador possui o menor erro quadrático médio (EQM)?

Q4) Uma empresa fabricante de luvas, antes de lançar um novo modelo de luvas no mercado, deseja realizar um estudo sobre o tamanho das mãos dos indivíduos da população alvo (distância do dedo mindinho ao polegar da mão aberta) a partir de uma amostra aleatória de tamanho n. Estudos anteriores mostram que é razoável assumir que essa medida se comporta segundo uma distribuição normal.

(a) Qual deve ser o valor de n de forma que o intervalo de confiança de 95% para o tamanho médio das mãos dos indivíduos da população, μ, tenha amplitude menor ou igual a $0,2\sigma$? (σ corresponde ao desvio padrão populacional.)

(b) Em uma amostra de 400 indivíduos, observou-se média de 17,5 cm e variância amostral de 25 cm^2. Obtenha um intervalo de 95% de confiança para o tamanho médio das mãos dos indivíduos da população. Interprete-o.

(c) Um intervalo com uma confiança maior teria amplitude maior ou menor? Justifique.

Boa Prova!

Solução

Q1) (a) Vamos denotar por L o lucro de uma venda. Note que L é uma variável aleatória discreta que só pode assumir os valores -500 (caso o produto apresente defeito em menos de $\frac{1}{2}$ anos) e 500 (caso o produto não apresente defeito antes de $\frac{1}{2}$ anos). Vamos então calcular primeiro $P(L=-500)$ e $P(L=500)$:

$$P(L=-500) = P\left(X < \frac{1}{2}\right) = \int_0^{\frac{1}{2}} f(x)dx = \int_0^{\frac{1}{2}} \frac{x^2}{2}dx = \left[\frac{x^3}{6}\right]_0^{\frac{1}{2}} = \frac{1}{48} \approx 0,021.$$

$$P(L=500) = 1 - P(L=-500) = \frac{47}{48} \approx 0,979.$$

Logo, temos que:

$$E[L] = -500 \times \frac{1}{48} + 500 \times \frac{47}{48} \approx 479.$$

Então, o lucro médio de uma venda será de $R\$479,00$.

(b) Vamos denotar por L_G o lucro de uma venda com garantia extendida. Note que L_G é uma variável aleatória discreta que só pode assumir os valores -500 (caso o produto apresente defeito em menos de $\frac{3}{2}$ anos) e 700 (caso o produto não apresente defeito antes de $\frac{3}{2}$ anos). Vamos então calcular $P(L_G=-500)$ e $P(L_G=700)$:

$$P(L_G=-500) = P\left(X < \frac{3}{2}\right) = \int_0^{\frac{3}{2}} f(x)dx = \int_0^1 \frac{x^2}{2}dx + \int_1^{\frac{3}{2}} \frac{1}{20}(13-3x)dx = \left[\frac{x^3}{6}\right]_0^1 + \frac{1}{20}\left[13x - \frac{3x^2}{2}\right]_1^{\frac{3}{2}} =$$

$$= \frac{1}{6} + \frac{1}{20}\left(\frac{39}{2} - \frac{27}{8} - 13 + \frac{3}{2}\right) = \frac{1}{6} + 0,23125 \approx 0,398.$$

$$P(L_G=700) = 1 - P(L_G=-500) \approx 1 - 0,398 = 0,602.$$

Logo, temos que:

$$E[L_G] = -500 \times 0,398 + 700 \times 0,602 \approx 222,4.$$

Então, o lucro médio de uma venda com a garantia extendida será de $R\$222,40$, que é menor do que o lucro esperado sem a garantia extendida (calculado no item (a)).

(c) Vamos denotar por y o tempo máximo que a garantia extendida pode oferecer para não ser desvantajosa. Pela resposta do item (a), sabemos que para encontrar y devemos resolver a seguinte equação:

$$-500P(X<y) + 700(1-P(X<y)) = 479 \quad \Rightarrow \quad -500P(X<y) - 700P(X<y) = -221$$

$$P(X<y) = \frac{221}{1200} \approx 0,1842.$$

Então, resta apenas obter o quantil 0,1842 da distribuição de X. Como

$$P(X < 1) = \int_0^1 \frac{x^2}{2}dx = \left[\frac{x^3}{6}\right]_0^1 = \frac{1}{6} \approx 0,1667,$$

temos que $P(1 < X < y) = 0,1842 - 0,1667 = 0,0175$. Logo:

$$\int_1^y \frac{1}{20}(13-3x)dx = 0,0175 \quad \Rightarrow \quad \left[13x - \frac{3x^2}{2}\right]_1^y = 0,35 \quad \Rightarrow \quad -1,5y^2 + 13y - 11,5 = 0,35$$

$$-1,5y^2 + 13y - 11,85 = 0.$$

Resolvendo essa equação de segundo grau, obtemos $y \approx 1,035$ ou $y \approx 7,631$. Como 7,631 anos não faz sentido, segue que o tempo máximo que pode ser oferecido na garantia extendida será de $1,035$ anos, que equivale a aproximadamente 378 dias. Por fim, o tempo adicional máximo que a loja pode oferecer é $0,535$ anos, que equivale a aproximadamente 195 dias.

Q2) $T =$ atendimento (em minutos) do caixa do banco, $\quad f(t) = 0,25te^{-0,5t},\ t \geq 0$.

(a) $f(t)$ é uma função de densidade,
i) $f(t) \geq 0$, pois, $t \geq 0$ e $e^{-0,5t} \geq 0$
ii) $\int_0^\infty 0,25te^{-0,5t}dt$=1. Usando integração por partes,
$u = 0,25t \quad du = 0,25dt$ e $dv = e^{-0,5t}dt \quad v = -2e^{-0,5t}$. Assim,

$$\int_0^\infty 0,25te^{-0,5t}dt = \left[-0,5te^{-0,5t}\right]_0^\infty + \int_0^\infty 0,5e^{-0,5t}dt = 0 - \left[e^{-0,5t}\right]_0^\infty = 1.$$

(b) A função de distribuição acumulada do tempo de atendimento,
$F(t) = P(T \leq t) = \int_0^t 0,25xe^{-0,5x}dx$. Usando a mesma integração por partes,
$u = 0,25t \quad du = 0,25dt$ e $dv = e^{-0,5t}dt \quad v = -2e^{-0,5t}$. Assim,
$\int_0^t 0,25xe^{-0,5x}dx = \left[-0,5xe^{-0,5x}\right]_0^t + \int_0^t 0,5e^{-0,5x}dx = -0,5te^{-0,5t} - \left[e^{-0,5x}\right]_0^t$
Assim, $F(t) = 1 - 0,5te^{-0,5t} - e^{-0,5t}$
$P(T > 2) = 1 - F(2) = 0,5 \times 2 \times e^{-1} + e^{-1} = 2e^{-1} = 0,7357$.

(c) Suponha agora que o tempo de atendimento, X, seja uma v.a. $f(x) = 2e^{-2x},\ x \geq 0$. Então, X$\sim Exp(\lambda = 2)$.
$P(X > 2) = 1 - F(2) = e^{-4} = 0,0183$
$P(2 < X \leq 3) = F(3) - F(2) = e^{-4} - e^{-6} = 0,0183 - 0,0025 = 0,0158$.

Q3) (a)

$$E[T_1] = E\left[\frac{\hat{\mu}_1 + 2\hat{\mu}_2 + 3\hat{\mu}_3}{6}\right] = \frac{1}{6}\left(E[\hat{\mu}_1] + 2E[\hat{\mu}_2] + 3E[\hat{\mu}_3]\right) = \frac{1}{6}\left(\mu + 2\mu + 3\mu\right) = \mu, \text{ assim } T_1 \text{ é não viesado.}$$

$$E[T_2] = E\left[\frac{3\hat{\mu}_1 + 2\hat{\mu}_2 - \hat{\mu}_3}{4}\right] = \frac{1}{4}\left(3E[\hat{\mu}_1] + 2E[\hat{\mu}_2] - E[\hat{\mu}_3]\right) = \frac{1}{4}\left(3\mu + 2\mu - \mu\right) = \mu, \text{ assim } T_2 \text{ é não viesado.}$$

$$E[T_3] = E\left[\frac{\hat{\mu}_1 + \hat{\mu}_2}{2}\right] = \frac{1}{2}\left(E[\hat{\mu}_1] + E[\hat{\mu}_1]\right) = \frac{1}{2}(\mu + \mu) = \mu, \text{ assim } T_3 \text{ é não viesado.}$$

(b) De forma geral $EQM(T) = Var(T) + B(T)^2$, como os estimadores são não viesados, temos $B(T) = 0$, assim:

$$\begin{aligned}
EQM(T_1) &= Var\left(\frac{\hat{\mu}_1 + 2\hat{\mu}_2 + 3\hat{\mu}_3}{6}\right) \\
&= \frac{1}{36}\left(Var(\hat{\mu}_1) + 4Var(\hat{\mu}_2) + 9Var(\hat{\mu}_3)\right) \\
&= \frac{1}{36}\left(Var(\hat{\mu}_1) + 4\frac{Var(\hat{\mu}_1)}{2} + 9\frac{Var(\hat{\mu}_1)}{3}\right) \\
&= \frac{Var(\hat{\mu}_1)}{6}
\end{aligned}$$

$$\begin{aligned}
EQM(T_2) &= Var\left(\frac{3\hat{\mu}_1 + 2\hat{\mu}_2 - \hat{\mu}_3}{4}\right) \\
&= \frac{1}{16}\left(9Var(\hat{\mu}_1) + 4Var(\hat{\mu}_2) + Var(\hat{\mu}_3)\right) \\
&= \frac{1}{16}\left(9Var(\hat{\mu}_1) + 4\frac{Var(\hat{\mu}_1)}{2} + \frac{Var(\hat{\mu}_1)}{3}\right) \\
&= \frac{17}{24}Var(\hat{\mu}_1)
\end{aligned}$$

$$\begin{aligned}
EQM(T_3) &= Var\left(\frac{\hat{\mu}_1 + \hat{\mu}_2}{2}\right) \\
&= \frac{1}{4}\left(Var(\hat{\mu}_1) + Var(\hat{\mu}_2)\right) \\
&= \frac{1}{4}\left(Var(\hat{\mu}_1) + \frac{Var(\hat{\mu}_1)}{2}\right) \\
&= \frac{3}{8}Var(\hat{\mu}_1)
\end{aligned}$$

O estimador como o menor EQM é T_1.

Q4) (a) O intervalo de confiança de 95% para a média populacional μ é dado por

$$IC(\mu; 0,95) = \left(\bar{x} - z_{0,975}\frac{\sigma}{\sqrt{n}}, \bar{x} + z_{0,975}\frac{\sigma}{\sqrt{n}}\right).$$

Portanto, sua amplitude é $2z_{0,975}\frac{\sigma}{\sqrt{n}}$, que deve ser menor ou igual a 2σ. Assim,

$$\begin{aligned}
2z_{0,975}\frac{\sigma}{\sqrt{n}} &\leq 0,2\sigma \\
\Rightarrow \sqrt{n} &\geq \frac{2}{0,2}z_{0,975} \\
\Rightarrow n &\geq \left(\frac{2}{0,2}z_{0,975}\right)^2 = \left(\frac{2}{0,2}1,96\right)^2 = 384,16 \approx 385.
\end{aligned}$$

(b) Observou-se $\bar{x} = 17,5$ e $s^2 = 25 \Rightarrow s = 5$. Considerando $n = 400$ grande o suficiente, pelo TCL a distribuição t de Student se aproxima da distribuição normal padrão e, assim, o intervalo de confiança pode ser calculado como

$$\begin{aligned} IC(\mu; 0,95) &= \left(\bar{x} - z_{0,975}\frac{s}{\sqrt{n}}, \bar{x} + z_{0,975}\frac{s}{\sqrt{n}}\right) \\ &= \left(17,5 - 1,96 \times \frac{5}{\sqrt{400}}; 17,5 + 1,96 \times \frac{5}{\sqrt{400}}\right) \\ &= (17,01; 17,99). \end{aligned}$$

Com 95% de confiança, o intervalo acima compreende μ. Em outras palavras, espera-se que em 95% das amostras, o intervalos construídos a partir delas compreendam o valor da média populacional.

(c) Um intervalo com maior nível de confiança teria amplitude maior. Uma vez que mais confiança é exigida, há mais incerteza e o quantil $z_{1-\frac{\alpha}{2}}$ é maior, produzindo um intervalo com maior amplitude.