UFRJ - CCMN - IM - Departamento de Métodos Estatísticos

Probabilidade e Estatística

Prova #2 **08-12-2016**

Atenção: Não serão aceitas respostas sem justificativa.

1. Num avião com 100 lugares embarcam passageiros, de forma aleatória, com peso médio de 75Kg e desvio padrão de 12Kg. A bagagem de cada passageiro tem média 25Kg e desvio padrão de 6Kg . Suponha que os pesos e os pertences das pessoas são v.a.'s independentes.

   (a) Sabendo que o limite de peso recomendado para a aeronave é de 10,2 Toneladas, use a distribuição da soma dos pesos das cem pessoas (com suas bagagens) para calcular a probabilidade de que o peso no avião não exceda esse limite.

   (b) Na construção de um novo modelo de aeronave (também com 100 lugares), qual deve ser o limite de carga para que se garanta, com 98% de confiança, que a soma dos pesos das pessoas com seus pertences, dessa população de usuários, não ultrapasse este limite?

   (c) Foi observado que a probabilidade de uma pessoa que faz uma reserva comparecer para viajar é de 0,85. Se forem feitas 110 reservas, qual a probabilidade de que todos os passageiros que comparecem consigam, de fato, viajar? (Explique todos os passos de sua solução!)

2. Considere um conjunto formado por 10 casais. A tabela abaixo contem as idades dessas 20 pessoas.

i = Número da observação	1	2	3	4	5	6	7	8	9	10
x_i = idade da mulher	30	31	32	33	34	35	36	37	38	39
y_i = idade do homem	32	35	35	y_4	42	42	45	y_8	49	52

   Note que não são fornecidas as idades y_4 e y_8 dos homens, para os casais 4 e 8.

   (a) Calcule o desvio padrão amostral da idade das mulheres.

   (b) Sabendo que $35 \leq y_4 \leq 42$ e $45 \leq y_8 \leq 49$, forneça a expressão matemática para a distância interquartil da idade dos homens, em função de y_4 e y_8.

   (c) Determine os valores das idades y_4 e y_8, sabendo que a distância interquartil da idade dos homens é igual a 12 e a covariância entre x e y é igual a $\frac{61}{3}$ anos^2.

   OBS: Para facilitar os cálculos são fornecidos os somatórios abaixo, todos eles referentes apenas às observações em que $i \neq 4$ e $i \neq 8$: $\sum x_i = 275$; $\sum y_i = 332$; $\sum x_i^2 = 9527$; $\sum y_i^2 = 14132$ e $\sum x_i y_i = 11573$.

3. Um pesquisador deseja estimar a proporção de ratos nos quais se desenvolve um certo tipo de tumor quando submetidos a radiação. Ele deseja que sua estimativa não se desvie da proporção verdadeira por mais de 0,04 com uma probabilidade de pelo menos 95%.

   (a) Quantos animais ele precisa examinar para satisfazer essa exigência?

   (b) Para quanto seria possível diminuir o tamanho da amostra utilizando a informação adicional de que em geral esse tipo de radiação não afeta mais que 22% dos ratos?

4. Uma editora anuncia que os livros editados por ela não contém mais que 4% de páginas com pelo menos um erro tipográfico. Seja p a proporção de páginas com pelo menos um erro tipográfico de um certo livro a ser editado e suponha que nesse livro 100 páginas sejam sorteadas para serem inspecionadas e para serem usadas em um teste de hipóteses, cuja conclusão determina se o livro editado será ou não lançado no mercado.

   (a) Formule um procedimento de teste de hipótese para o problema indicando: as hipóteses a serem testadas, a estatística de teste e sua distribuição sob a hipótese nula e a regra de decisão com nível de significância de 3%.

   (b) Com base na regra de decisão obtida no item (a), calcule a probabilidade de ser lançado no mercado um livro em que 5% de páginas contém erros tipográficos.

   (c) Se forem observadas 9% das páginas com pelo menos um erro, qual o p-valor? Interprete este resultado.

Boa prova!

SOLUÇÕES

1. Sejam X$_i$=peso do i-ésimo passageiro e Y$_i$=peso da bagagem do i-ésimo passageiro.
T= peso total no avião (passageiros mais bagagens)

(a) T=$\sum_{i=1}^{100}(X_i+Y_i)$. Como $n=100$ (grande), pelo TCL e devido ao resultado "soma de Normais é Normal", $T \sim N(\mu_T;\sigma_T^2)$, onde: $\mu_T = 100\times 75 + 100 \times 25 = 10000$ e $\sigma_T^2 = 100 \times 12^2 + 100 \times 6^2 = 134,16^2$
Assim, $P(T < 10200) = \Phi(\frac{10200-10000}{134,16}) = \Phi(1,49) = 0,9319$

(b) $P(T < lim) = 0,98 \Rightarrow \frac{lim-10000}{134,16} = 2,05 \Rightarrow lim = 134,16 * 2.05 + 10000 = 10275Kg$

(c) Seja N= Número de pessoas com reserva que comparecem para viajar.
N $\sim$ Bin(n=110; p=0,85). Como np(1-p)=14.05 $\geq$ 3 poderemos usar a aproximação da Binomial pela Normal.
Seja N$_c$ a v.a. continua correspondente a N.
N$_c$ $\sim N(np = 93,5;\ np(1-p) = 14,05 = 3,74^2)$. Assim,
$P(N \leq 100) = P(N_c < 100,5) = \Phi(\frac{100,5-93,5}{3,74}) = \Phi(1,87) = 0,97$

2. **(a)** $s_x = \sqrt{\frac{(9527+33^2+37^2)-\frac{(275+33+37)^2}{10}}{9}} = \sqrt{\frac{82,5}{9}} = 3,03$ anos.

(b) Posição(Q1) $= \frac{10+3}{4} = 3,25$ e Posição(Q3) $= \frac{3\times 10+1}{4} = 7{,}75$. Portanto:
$DIQ(y) = Q3(y) - Q1(y) = \left[\frac{y_{(7)}+3y_{(8)}}{4}\right] - \left[\frac{3y_{(3)}+y_{(4)}}{4}\right] = \left[\frac{45+3y_{(8)}}{4}\right] - \left[\frac{3\times 35+y_{(4)}}{4}\right] = \frac{3y_8-y_4-60}{4}$.
OBS: Decorre das desigualdades fornecidas: $35 \leq y_4 \leq 42$ e $45 \leq y_8 \leq 49$, que os dados relativos à idade dos homens já estão dispostos em ordem crescente. Por isso é que foi possível substituir $y_{(4)}$ por y_4 e $y_{(8)}$ por y_8.

(c) Como DIQ(y) = 12, decorre que $3y_8 - y_4 = 108$ (I). Por outro lado,
$\frac{61}{3} = s_{xy} = \frac{(11573+33y_4+37y_8)-\frac{(275+33+37)\times(332+y_4+y_8)}{10}}{9}$.
Logo, $\frac{61\times 9}{3} = 183 = 11573 + 33y_4 + 37y_8 - \frac{345}{10}(332 + y_4 + y_8)$. Daí decorre que $5y_8 - 3y_4 = 128$ (II).
Resolvendo o sistema formado pelas equações (I) e (II), chegamos a $y_4 = 39$ anos e $y_8 = 49$ anos.

3. **(a)** Pelo enunciado acima temos:

- Tolerância: $\epsilon = 0,04$.
- Coeficiente de confiança: $P(|Erro| < \epsilon) = 0,95$.
- Logo, pela tabela da distribuição Normal Padrão, temos que z é tal que $\Phi(z) = 0,975$, portanto, $z = 1,96$. Como não temos uma informação preliminar sobre p, devemos utilizar p=0,5, que maximiza p(1-p). Assim, podemos calcular o tamanho da amostra da seguinte forma:

$$n = \left(\frac{z}{\epsilon}\right)^2 p(1-p) = \left(\frac{1,96}{0,04}\right)^2 0,25 = 600,25 \approx 601$$

Logo, para que o erro cometido na estimação da proporção de ratos nos quais se desenvolve certo tipo de tumor quando submetidos a radiação seja no máximo 0,04 com probabilidade igual a 0,95, o pesquisador precisa examinar 601 animais.

(b) Se p for no máximo 22%, o tamanho da amostra será:

$$n = \left(\frac{z}{\epsilon}\right)^2 p(1-p) = \left(\frac{1,96}{0,04}\right)^2 0,22 \times 0,78 = 412,02 \approx 413.$$

Logo, se p for no máximo 22%, para que o erro cometido na estimação da proporção de ratos nos quais se desenvolve certo tipo de tumor quando submetidos a radiação seja no máximo 0,04 com probabilidade igual a 0,95, o pesquisador precisa examinar apenas 413 animais.

4. **(a)** Seja p, a verdadeira proporção de páginas com pelo menos um erro tipográfico nos livros editados por esta editora.

   (i) $H_0 : p \leq 0,04$ vs. $H_1 : p > 0,04$.

   (ii) A estatística teste é $Z = \frac{\hat{p}-0,04}{\sqrt{\frac{0,04\times 0,96}{100}}}$ e sua distribuição assintótica sob a hipótese nula é Normal(0,1).

   (iii) Como $\alpha = 0,03$, rejeita-se H_0 se $z > 1,8808$. Ou ainda se $\frac{\hat{p}-0,04}{\sqrt{\frac{0,04\times 0,96}{100}}} > 1,8808$, logo, se $\hat{p} > 0,0768$. Caso contrário, aceita-se H_0.

   (b) $P(\widehat{p} < 0.0768 \mid p = 0,05) = P(Z < \frac{0,0768-0,05}{\frac{\sqrt{(0,05*0,95)}}{10}}) = P(Z < 1,23) = 0,89 = 89\%$.

   (c) p-valor $= P(\widehat{p} > 0,09 \mid H_0) = P(Z > \frac{0,09-0,04}{\sqrt{\frac{0,04\times 0,96}{100}}}) = P(Z > 2,551) = 0,0054 = 0,54\%$. Os dados fortemente contradizem a hipótese nula, ou indicam a rejeição.