CONHECIMENTOS ESPECÍFICOS

	2003	2004	2005	2006	2007	2008	2009	2010
X	39,0	39,5	39,5	39,0	39,5	41,5	42,0	42,0
Y	46,5	65,5	86,0	100,0	121,0	150,5	174,0	203,0

A tabela acima mostra as quantidades, em milhões de unidades, de linhas de telefones fixos (X) e de celulares (Y), em determinada região do país, de 2003 a 2010. Tendo como referência os dados dessa tabela, julgue os itens que se seguem.

51 O intervalo entre quartis (ou intervalo interquartílico) da distribuição Y é superior a $108{,}5 \times 10^6$.

52 Em relação às estatísticas de posição da variável X, observa-se que $Mo = Q_2 < Me$, em que Mo é a moda, Q_2 corresponde ao segundo quartil e Me é a média amostral. Essa relação sugere que a distribuição de X possui assimetria positiva (ou à direita).

53 O diagrama de dispersão de X *versus* Y é a representação gráfica da distribuição de probabilidade conjunta entre essas variáveis.

Em um espaço de probabilidade (Ω, Ψ, P), Ω representa o espaço amostral, Ψ é a álgebra de eventos e P é a medida de probabilidade. A respeito dos eventos não vazios A e B em $(\Omega, \Psi, \mathrm{P})$, julgue os itens seguintes.

54 Se A e B forem eventos independentes e equiprováveis com $P(A) = P(B) = 0{,}1$, então $P(A \cup B) < 0{,}20$.

55 Se $P(A) \leq P(B)$, então $\mathrm{A} \subset \mathrm{B}$.

A distribuição do número de erros (Y) registrados em um sistema computacional, do instante $T = 0$ até o instante $T = t$, é descrita pela distribuição de probabilidade condicional na forma $P(Y = k \mid X = t) = (1 - e^{-t})e^{-kt}$, em que $k = 0, 1, 2, ...$ representa uma possível realização da variável aleatória Y e X representa uma distribuição exponencial com média unitária.

Com base nessas informações, julgue os itens subsecutivos.

56 Se μ é a média da variável aleatória Y, então $0 < \mu < \infty$.

57 Para alguma constante positiva γ e para alguma medida de posição μ, a variável transformada $Z = \gamma \times (Y - \mu)$ terá média nula e variância unitária.

58 A distribuição de probabilidades da variável aleatória Y é dada por $P(Y = k) = \dfrac{1}{(k+2)(k+1)}$, em que $k = 0, 1, 2, ...$

59 A média condicional $\mathrm{E}(Y \mid X = t)$ é igual a $(e^t - 1)^{-1}$, em que $t > 0$.

RASCUNHO

Em métodos estatísticos e estudos estatísticos por simulações computacionais, a transformação de variável é um recurso que permite resolver problemas de não normalidade e de heterocedasticidade. Acerca de transformação de variáveis, julgue os itens seguintes.

60 A transformação de Box-Müller permite gerar duas distribuições normais independentes, com base em duas distribuições uniformes independentes.

61 Considere que X siga a distribuição contínua assimétrica, em torno da média, e possua mediana nula. Nessa situação, a transformação de Box-Cox $Y = 2\sqrt{X} - 2$ produzirá variável transformada, que seguirá a distribuição normal univariada.

62 Suponha que, no intervalo [0, 1], U seja uniforme e contínua e $Y = -\ell n U$. Nessa situação, a variância da variável transformada Y será inferior à da variável U.

63 Considere a transformação $Y = \sqrt{X}$, em que a variável aleatória X segue a distribuição qui-quadrado com 1 grau de liberdade. Nesse caso, é correto afirmar que Y segue a distribuição normal padrão.

64 Suponha que, no intervalo [0, 1], a variável aleatória U seja uniforme e contínua. Nesse caso, se $Y = -\ell n\left(\dfrac{U}{1-U}\right)$, então Y seguirá a distribuição logística.

De uma grande população X, será retirada, aleatoriamente, uma amostra simples de tamanho n para que seja estimada a taxa média m de satisfação do cliente. Considerando que a variância dessa população seja igual a 5 e que a média amostral ($\bar{X}$) seja o estimador não tendencioso da taxa m, julgue os itens a seguir.

65 De acordo com o teorema limite central, o erro de estimação $\varepsilon = \bar{X} - m$ converge em distribuição para a normal, com média zero e variância 5.

66 Segundo a lei fraca dos grandes números, para qualquer amostra de tamanho superior a 100, tem-se que $P(|\bar{X} - m| < 1) \geq 0{,}95$.

67 Para garantir a convergência em probabilidade da média amostral para a taxa média m, a população pesquisada deverá ser, necessariamente, gaussiana ou normal.

68 De acordo com a lei forte dos grandes números, quase certamente, a média amostral converge para o valor m, desde que n seja finito e suficientemente grande.

A amostra aleatória simples X_1, X_2, ..., X_{100} foi retirada, sem reposição, de população de tamanho 10 mil, em que X_1 representa a primeira observação, X_2, a segunda, e assim sucessivamente. O valor esperado e o desvio padrão de X_1 são, respectivamente, iguais a 10 e 5. Considerando essas informações e que $S = X_1 + X_2 + ... + X_{100}$, julgue os itens subsequentes.

69 O valor esperado de S é igual a 1.000.

70 A variância de S é inferior a 2.500.

71 A distribuição amostral exata da soma S é hipergeométrica.

72 As variáveis aleatórias X_1, X_2, ..., X_{100} são independentes e identicamente distribuídas.

RASCUNHO

Determinado estudo sobre a distribuição X das flutuações das intensidades de sinais registrados em receptor móvel considerou a função de probabilidade acumulada na forma $F(x) = P(X \leq x) = \frac{x^3}{V^3}$, em que V representa um parâmetro desconhecido e $0 < x \leq V$. Com base em amostra aleatória simples, de tamanho 20, foram obtidas as seguintes estatísticas descritivas acerca da distribuição X: média amostral = 5; mediana amostral = 6; máximo = 7.

Considerando essas informações, julgue os próximos itens.

73 Com base na média amostral, é correto afirmar que a estimativa de momentos do parâmetro V está entre 6,5 e 6,8.

74 A estimativa de máxima verossimilhança da variância da distribuição X é igual a $\frac{147}{80}$.

75 A estimativa de máxima verossimilhança do parâmetro V é superior a 6,8.

A respeito de inferência estatística, julgue os itens que se seguem.

76 Considere que T_1 e T_2 sejam estimadores não viciados de um mesmo parâmetro e que as variâncias var(T_1) e var(T_2) sejam tais que var(T_1) < var(T_2). Nesse caso, o estimador T_1 é mais eficiente que T_2.

77 Considerando a amostra aleatória simples X_1, X_2, X_3, retirada de determinada distribuição de Bernoulli, com parâmetro p desconhecido, é correto afirmar que $X_1 + X_2X_3$ é estatística suficiente.

78 Se T for estimador cujo erro quadrático médio (*mean-squared error*) é igual a sua variância, então, nesse caso, T é estimador não viciado.

79 Considere que determinado estimador E seja não viciado e que sua variância seja var(E) = $k\ n$, em que k é uma constante positiva e n, o tamanho da amostra. Nesse caso, E é um estimador consistente.

Acerca da estatística $Q^2 = \sum_{j=n}^{n} (X_j - \bar{X})^2$, em que $\bar{X} = \frac{X_1 + \cdots + X_n}{n}$ e $X_1, X_2, ..., X_n$ representa uma amostra aleatória simples retirada de uma distribuição normal, com média 3 e variância igual a 4, julgue os itens a seguir.

80 Q^2 e $\bar{X}$ são independentes.

81 A razão $\frac{(\bar{X} - 3)\sqrt{n(n-1)}}{\sqrt{Q^2}}$ define a distribuição t de Student, com $n - 1$ graus de liberdade.

82 A distribuição amostral da estatística $\frac{Q^2}{n-1}$ é qui-quadrado, com $n - 1$ graus de liberdade.

RASCUNHO

Um estudo foi realizado para identificar a percepção dos analistas de mercado a respeito do clima organizacional de determinada empresa de telecomunicações. Com base nos resultados dessa pesquisa, deseja-se testar a hipótese nula H_0: $\theta = 0,5$, contra a hipótese alternativa H_1: $\theta \neq 0,5$, em que θ é o parâmetro de interesse. Considerando essas informações, julgue itens consecutivos.

83 A região de rejeição do teste corresponde a um intervalo de confiança para θ.

84 O nível de significância do teste é a probabilidade de que seja rejeitada a hipótese nula quando, seguramente, ela é verdadeira.

85 O teste em questão é bilateral e contempla duas hipóteses simples: θ é igual ou diferente de 0,5.

Uma variável aleatória contínua X tem a função densidade de probabilidade $f(x) = (r + 1)x^r$ no intervalo [0, 1], sendo $(x_1, ..., x_n)$ uma amostra de X.

A respeito de estimadores de máxima verossimilhança (MV), julgue os itens seguintes.

86 O estimador de MV de r será negativo se e somente se $x_1 \cdot ... \cdot x_n < e^{-n}$.

87 A função de verossimilhança é $L(r) = (r + 1)^n (x_1 \dots x_n)^r$.

88 Para determinar o estimador de MV, é suficiente maximizar a função de verossimilhança ou minimizar o logaritmo dessa função.

89 O estimador de MV do parâmetro r é igual a $\dfrac{n}{\ln(x_1 \cdot ... \cdot x_n)} - 1$.

O número X de realizações de determinado experimento necessárias para obter o primeiro sucesso segue a distribuição geométrica $P(X = k) = p(1 - p)^{k-1}$. Considerando $(x_1, ..., x_n)$ uma amostra de X, julgue os itens subsequentes.

90 Se, após realizadas cinco séries do experimento, cada série tiver terminado com o primeiro sucesso e os números de experimentos, em cada série, tiverem sido 4, 7, 6, 5 e 3, então o estimador de máxima verossimilhança para p é igual a 0,2.

91 O estimador de máxima verossimilhança para o parâmetro p é dado por $\hat{p} = \dfrac{1}{\bar{x}}$, em que $\bar{x}$ é a média amostral.

RASCUNHO

estação	X	Y
1	200	20
2	600	15
3	800	10
4	1.400	5
5	2.000	0

A tabela acima mostra os resultados da temperatura, Y, em graus Celsius, obtidos a partir de um estudo realizado por um meteorologista em cinco diferentes estações, situadas em altitudes diferentes, X, em metros. As medições foram feitas no mesmo horário e no mesmo dia, e os dados da tabela satisfazem as relações a seguir.

$$\sum_{i=1}^{5} x_i = 5.000$$

$$\sum_{i=1}^{5} y_i = 50$$

$$\sum_{i=1}^{5} x_i^2 = 7.000.000$$

$$\sum_{i=1}^{5} y_i^2 = 750$$

$$\sum_{i=1}^{5} x_i y_i = 28.000$$

Considerando que o modelo de regressão linear simples $Y = a + bX + \varepsilon$, em que ε represente o erro aleatório, tenha sido ajustado aos dados, julgue os itens que se seguem.

92 Os estimadores de mínimos quadrados para os parâmetros a e b são não viciados.

93 Para permitir um teste de hipóteses ou a construção de um intervalo de confiança para os parâmetros a e b, é necessário supor que as temperaturas observadas sejam distribuídas normalmente. Além disso, para a construção do intervalo de confiança, utilizam-se estatísticas com distribuição t de Student, com $n - 2$ e $n - 1$ graus de liberdade para os parâmetros a e b, respectivamente.

94 Nesse modelo, se ε for uma variável aleatória com valor esperado zero e variância constante, os parâmetros a e b serão denominados, respectivamente, inclinação e intercepto.

95 As equações a seguir, com $n = 5$, são as equações normais do modelo.

$$an + b\sum_{i=1}^{n} x_i = \sum_{i=1}^{n} y_i$$

$$a\sum_{i=1}^{n} x_i + b\sum_{i=1}^{n} x_i^2 = \sum_{i=1}^{n} y_i^2$$

96 Os estimadores de mínimos quadrados para a e b são $\hat{a} = 21$ e $\hat{b} = -0{,}011$.

97 O valor absoluto do terceiro resíduo é 2,2.

RASCUNHO

X	Y
0	$y_1 = 80$
1	$y_2 = 70$
2	$y_3 = 50$
3	$y_4 = 40$
4	$y_5 = 30$

A tabela acima mostra o resultado do estudo efetuado por certa empresa automobilística a respeito do preço de determinado modelo de veículo, Y, em R$ mil, em função da idade, X, em anos. O correspondente modelo de regressão linear simples foi determinado na forma $Y = 80.000 - 13.000\,X + \varepsilon$, em que o erro aleatório ε tem desvio padrão de R$ 5.000,00. O preço médio dos veículos é $\bar{y} = 54.000$ e a soma dos quadrados total é $\text{SQT} = \sum_{i=1}^{5}(y_i - \bar{y})^2 = 1.720 \times 10^6$.

Com base na tabela e nas informações apresentadas, julgue os itens seguintes a respeito desse modelo.

98 O coeficiente de determinação é superior a 90%.

99 A soma dos quadrados de regressão é inferior a 1.420×10^6.

100 O preço Y é uma variável aleatória com valor esperado igual a $80.000 - 13.000\,X$ e variância de 25 milhões.

101 Para um veículo com 2,5 anos de idade, o preço estimado pelo modelo é igual a R$ 45.000,00.

102 Considere que os parâmetros determinados sejam os verdadeiros parâmetros populacionais. Nessa situação, o preço de um veículo com 3 anos de idade está entre R$ 41.000,00 e R$ 43.500,00, com probabilidade $\Phi(0{,}5) - 0{,}5$, em que $\Phi(x)$ é a função de distribuição acumulada da distribuição normal padronizada.

103 A soma dos quadrados dos erros entre preços teóricos e observados é corretamente obtida por $\sum_{i=1}^{5}(y_i - \hat{y}_1)^2 = 30 \times 10^6$.

No universo $U = \{1, 2, ..., N\}$, por meio de amostragem aleatória simples com reposição, seleciona-se uma amostra de tamanho n. Considerando que f_i seja o número de vezes em que a unidade i do universo aparece na amostra, julgue os itens a seguir.

104 Para $N = 100$ e $n = 10$, a variância da variável aleatória f_i é igual a 0,099.

105 Se $N = 100$ e $n = 10$, então a probabilidade de que a amostra contenha o elemento 1 é exatamente igual a 10%.

106 Se $N = 100$ e $n = 10$, então a probabilidade de que a amostra contenha os elementos 1 e 2 pode ser corretamente expressa por $1 - 0{,}99^{10} + 0{,}98^{10}$.

107 Se $N = 10$ e $n = 5$, então a probabilidade de que a amostra contenha uma vez o elemento 2, duas vezes o elemento 5 e duas vezes o elemento 8 é igual a 0,0003.

108 Toda variável aleatória f_i é binomialmente distribuída com valor esperado $\frac{n}{N}$.

RASCUNHO

No universo $U = \{1, 2, ..., N\}$, por meio de amostragem aleatória simples sem reposição, seleciona-se uma amostra de tamanho n. Considerando que f_i seja o número de vezes em que a unidade i do universo aparece na amostra, julgue os itens subsequentes.

109 O valor esperado da variável aleatória f_i é igual a $\frac{n}{N}$.

110 A variância da variável aleatória f_i é igual a $\frac{1}{N} \times (1 - \frac{n}{N})$.

111 A probabilidade de que a amostra contenha os elementos 1, 2 e 3 é igual a $\frac{n \cdot (n-1) \cdot (n-2)}{N \cdot (N-1) \cdot (N-2)}$.

112 A variável f_i segue uma distribuição de Bernoulli, isto é, $P(f_i = 1) = \frac{n}{N}$ e $P(f_i = 0) = 1 - \frac{n}{N}$.

Em uma população de 8 pessoas — $U = \{1, ..., 8\}$ —, com pesos — $D = \{40, 50, 60, 80, 80, 90, 100, 100\}$ — medidos em quilogramas, o peso médio da amostra é igual a 75 kg. A população foi dividida nos estratos $U_A = \{1, 2, 3\}$ e $U_B = \{4, 5, 6, 7, 8\}$, com pesos $D_A = \{40, 50, 60\}$ e $D_B = \{80, 80, 90, 100, 100\}$, em que as variâncias desses estratos são $\sigma_A^2 = \frac{200}{3}$ e $\sigma_B^2 = 80$, respectivamente.

Julgue os itens a seguir no que se refere a essa população.

113 A variância das médias dos estratos (variância entre estratos) é igual a 375.

114 Se, em cada estrato, for escolhida, aleatoriamente, uma amostra de tamanho 2, então as variâncias das médias amostrais serão $V(\bar{y}_A) = \frac{100}{3}$ e $V(\bar{y}_B) = 40$.

115 Uma estimativa não viciada para o peso médio populacional é $\frac{3}{8}\bar{y}_A + \frac{5}{8}\bar{y}_B$ e sua variância é igual a $\frac{1.300}{64}$.

116 A média das variâncias dos estratos (variância dentro) é igual a $\frac{75}{2}$.

Julgue os itens seguintes a respeito do tamanho de uma amostra.

117 A amostra de 400 indivíduos em populações de qualquer tamanho é suficiente se o erro amostral for de 3%.

118 Se o erro amostral tolerável for de 4%, então uma amostra aleatória simples da população de determinada cidade, com 400.000 habitantes, precisa conter mais de 700 pessoas.

Determinada população consiste de apenas 12 pessoas, com as seguintes alturas, em centímetros: 150, 170, 165, 175, 160, 180, 140, 175, 170, 160, 145, 190. Tendo como referência essa população e as alturas das pessoas que fazem parte dela, julgue os itens a seguir em relação a amostragem.

119 As amostras sistemáticas de tamanho 3 têm as alturas médias 160 cm, 170 cm, 150 cm e 180 cm.

120 A amostragem sistemática pode ser considerada um caso específico de amostragem por conglomerados, em que os conglomerados têm tamanhos iguais e duas amostras são selecionadas para observação.

RASCUNHO