CONHECIMENTOS ESPECÍFICOS

Em um pequeno aquário natural, há 15 peixes: quatro da espécie A, cinco da espécie B e seis da espécie C. Para avaliação de uma possível contaminação ambiental, cinco peixes desse aquário serão selecionados aleatoriamente e sacrificados para estudos em laboratório. Com base nessas informações, julgue os itens a seguir, que versam sobre cálculos de probabilidades.

51 A probabilidade de que na amostra haja, no máximo, um peixe da espécie C é superior a 0,25.

52 Se no vetor aleatório $\mathbf{X} = (X_1, X_2, X_3)$, as variáveis aleatórias X_1, X_2 e X_3 representam, respectivamente, as quantidades de peixes das espécies A, B e C na amostra, então o vetor **X** segue uma distribuição multinomial.

53 A probabilidade de serem encontrados, na amostra, 3 peixes da espécie A e 2 peixes da espécie C é superior a 0,03.

54 A probabilidade de que todos os peixes selecionados na amostra sejam da mesma espécie é inferior a 0,003.

Considerando, no espaço tridimensional, o plano Π que contém os pontos $A = (1, 2, 0)$, $B = (0, 1, 1)$ e $C = (-2, 2, 2)$ e a reta R que passa pelos pontos B e C, julgue os itens seguintes.

55 Considere a reta S que passa pela origem O = (0, 0, 0) e é perpendicular ao plano Π. Nesse caso, o ponto de intersecção de S com Π é o ponto de coordenadas $\left(\frac{4}{7}, \frac{2}{7}, \frac{6}{7}\right)$.

56 A distância do ponto A à reta R é igual a $\frac{\sqrt{23}}{3}$ unidades de comprimento.

Um indicador de qualidade ambiental X é uma variável aleatória contínua e um pesquisador deseja testar se a distribuição desse indicador tem função de densidade de probabilidade expressa por $f(x) = 0{,}25 + 1{,}5x$, se $0 \le x \le 1$; e $f(x) = 0$ para os demais valores de x. Para usar o teste qui-quadrado de aderência de Pearson, o pesquisador definiu os eventos A_1, ..., A_5, em que $A_i : \frac{i-1}{5} \le X < \frac{i}{5}$, para $i = 1, 2, 3$ e 4, e $A_5 : 0{,}8 \le X \le 1$. A partir de uma amostra aleatória simples de $n = 100$ realizações de X, os respectivos números das ocorrências n_1, ..., n_5 desses eventos A_1, ..., A_5 foram iguais a 12, 15, 18, 25, 30.

Com base nessas informações, julgue os próximos itens, considerando que a distribuição qui-quadrado com 4 graus de liberdade satisfaça $P(D^2 \le 9{,}488) = 0{,}95$ e que a estatística desse teste seja dada por $D^2 = \sum_{i=1}^{5} \frac{(n_i - np_{i0})^2}{np_{i0}}$, em que $p_{i0} = P(A_i)$ representa a probabilidade de ocorrência do evento A_i sob a hipótese nula.

57 Com base na amostra observada, adotando-se um nível de significância de 5%, não há evidências estatísticas que permitem rejeitar a hipótese nula.

58 Sob a hipótese nula, tem-se que $p_{i0} = \frac{1+2i}{35}$.

59 O valor da estatística do teste foi superior a 3 e inferior a 4.

RASCUNHO

Considere que X seja uma variável aleatória discreta tal que $p_i = P(X = x_i) > 0$, em que x_i representa um possível valor de X, $i = 1, 2, ...,$ e $p_1 + p_2 + ... = 1$. Para simular realizações de X, considere o seguinte algoritmo:

- Gere um valor aleatório u de uma distribuição uniforme no intervalo [0, 1];
- Efetue as comparações:
  se $u < p_1$, faça $x = x_1$ e pare; se $u > p_1$, continue;
  se $p_1 \leq u < p_1 + p_2$, faça $x = x_2$ e pare; caso contrário, continue;
  ....
  se $p_1 + ... + p_{k-1} \leq u < p_1 + ... + p_k$, faça $x = x_k$ e pare; caso contrário continue;
  ....

Considerando essas informações, julgue os itens que se seguem, acerca de estatística computacional.

60 Considere que os valores possíveis $x_1, x_2,.., x_n$ sejam ordenados segundo suas probabilidades de ocorrência, de modo que se tenha $P(X = y_1) > P(X = y_2) > ... > P(X = y_n)$ na nova sequência de valores possíveis $y_1, y_2, ..., y_n$. Nessa situação, o número esperado de comparações será mínimo.

61 Suponha que a variável aleatória X possua uma quantidade infinita de valores possíveis. Nessa situação, é possível que a quantidade de comparações seja infinita, isto é, não haja parada.

62 Suponha que a distribuição de X seja dada conforme a tabela abaixo. Nesse caso, o número médio de comparações até a parada é igual a 2,5.

x_i	1	2	3	4
$p_i = P(X = x_i)$	0,2	0,15	0,25	0,4

Julgue os próximos itens, a respeito de análise matemática.

63 $\frac{d}{dx}\exp(2\,\mathrm{sen}\,x^n) = 2nx^{n-1}\cos x^n \exp(2\,\mathrm{sen}\,x^n)$.

64 $\int x^2 e^x dx = e^x(x^2 - 2x + 2) + \mathrm{c}$, em que c é uma constante real.

65 $\frac{d}{dx}\frac{\mathrm{sen}x^2}{\cos x} = x\cos x + \frac{\mathrm{sen}x\ \mathrm{sen}x^2}{\cos x^2}$.

Determinada população, sem movimentos migratórios, é descrita pela relação $p_t = p_0(1 + r)^t$, em que t, p_t, p_0 e r representam, respectivamente, o tempo, a população no instante t, a população no instante $t = 0$, e a taxa de crescimento. Considerando esse modelo populacional, julgue os itens a seguir, que versam sobre demografia.

66 Se o tempo t for expresso em anos, se, em 2000, a população era formada por 100 milhões de indivíduos e por 121 milhões em 2010, então, em 2005, a população era formada por menos de 109 milhões de indivíduos.

67 A partir dos valores t, p_0 e p_t, a taxa de crescimento r é estimada pela fórmula $r = \left(\frac{p_t}{p_0}\right)^t - 1$.

68 Se a taxa de crescimento for constante, então o instante t em que a população p_t será o dobro da população inicial p_0 é expresso por $t = \frac{\ln 2}{\ln(1+r)}$.

RASCUNHO

O método da bissecção, ou da dicotomia, permite determinar a raiz aproximada x^* de uma função contínua f: $[a, b] \rightarrow \mathbf{R}$, em que $f(a) < 0 < f(b)$. Acerca desse método numérico, julgue os itens subsecutivos.

69 O método da bissecção pode também ser utilizado para minimizar uma função continuamente diferenciável f: $[a, b] \rightarrow$ R, tal que $f'(a) < 0 < f'(b)$. Nessa situação, o método converge sempre para um ponto de mínimo ou para um ponto de máximo de f.

70 Esse método consiste na construção de uma sequência de intervalos $I_k = [a_k, b_k]$ contendo x^*. Por indução, constrói-se o intervalo I_{k+1} a partir do intervalo I_k da seguinte forma: o primeiro intervalo é definido por $I_0 = [a_0, b_0] = [a, b]$. Encontrado o intervalo I_k, determina-se o valor $f(m_k)$, em que $m_k = \dfrac{a_k + b_k}{2}$. Se $f(m_k) = 0$, $I_{k+1} = I_k$ e $x^* = m_k$. Se $f(m_k) < 0$, define-se $I_{k+1} = [m_k, b_k]$; se $f(m_k) > 0$, define-se $I_{k+1} = [a_k, m_k]$.

O método *simplex* é um dos principais métodos de resolução do problema de programação linear:

$$\text{Maximizar } z = c^T\mathbf{x}, \text{ sujeito a}$$
$$\mathbf{Ax} \leq \mathbf{b},$$
$$\mathbf{x} \geq 0,$$

em que c^T é um vetor transposto, **x** e **b** são vetores e **A** representa uma matriz. Os cálculos podem ser convenientemente realizados em um *tableau simplex*, que será transformado por pivoteamentos até que a solução ótima seja encontrada. Suponha que o *tableau* inicial seja escrito na forma

VB	$x_1, ..., x_n$	$s_1, ..., s_m$	b
s_1 $\vdots$ s_m	A	I	b_1 $\vdots$ b_m
Z	−c	0	0

em que $s_1, ..., s_m$ denotam as variáveis de folga, a primeira coluna contém as variáveis básicas (VB), Z representa o valor objetivo. Suponha, ainda, que, no *tableau* inicial, os componentes b_i sejam todos não negativos. A partir das informações acima, julgue os itens subsequentes.

71 O algoritmo termina quando todos os elementos da última linha do *tableau* forem não negativos. A existência de zeros nessa linha indica que há varias soluções ótimas.

72 Se pelo menos um dos elementos da última coluna de um *tableau* for nulo, então a solução atual será degenerada, isto é, várias bases são associadas ao vértice atual, e o próximo pivoteamento deve produzir uma solução básica com valor objetivo menor.

73 Se o algoritmo *simplex* (na versão original de Dantzig) for aplicado ao problema

$$\text{Maximizar } z = x_1 + 2x_2, \text{ sujeito a}$$
$$-x_1 + x_2 \leq 2,$$
$$x_1 + x_2 \leq 4,$$
$$x_1, x_2 \geq 0,$$

a solução ótima $(x_1, x_2) = (1, 3)$ é determinada depois de duas iterações e o valor ótimo de z é $z = 7$.

RASCUNHO

As lâmpadas fabricadas por A têm durabilidade média de 1.400 horas com desvio padrão de 200 horas; as do fabricante B têm durabilidade média de 1.200 horas com desvio padrão de 100 horas. Considerando a retirada de uma amostra aleatória simples de 125 lâmpadas de cada um dos fabricantes A e B, julgue os itens a seguir.

74 É praticamente impossível que a durabilidade média amostral das lâmpadas do fabricante B seja superior à durabilidade média amostral das lâmpadas produzidas pelo fabricante A.

75 A diferença $\overline{X}_A - \overline{X}_B$ entre os tempos médios amostrais de durabilidade das lâmpadas dos fabricantes A e B segue uma distribuição aproximadamente normal com média igual a 200 horas e desvio padrão inferior a 15 horas.

76 A probabilidade de a durabilidade média amostral das lâmpadas do fabricante A ser superior à durabilidade média do fabricante B em pelo menos 220 horas é igual a $\phi(1)$, em que ϕ é a função de distribuição acumulada da distribuição normal padrão.

X	Y
−2	5
−1	1
0	0
2	4
3	10

Considerando o ajuste do modelo de regressão linear na forma $Y = aX + bX^2 + \varepsilon$ pelo método de mínimos quadrados ordinários, com base nos dados apresentados na tabela acima, em que ε representa o termo aleatório, julgue o item abaixo.

77 Para os valores da tabela, o sistema de equações normais é

$$18a + 26b = 27$$
$$26a + 114b = 127.$$

Em determinado jogo, ganha-se R$ 1,00 com probabilidade p e perde-se R$ 1,00 com probabilidade $1 - p$. Um jogador inicia o jogo com o capital de i reais, continuando no jogo até acumular a quantia de T reais ou até perder todo o seu capital. Se w_i é a probabilidade de o jogo terminar com sucesso do jogador que dispõe do capital de i reais no início de uma rodada, julgue os itens seguintes.

78 Se $p \neq 0{,}5$, então as probabilidades w_i são dadas por $w_i = \dfrac{q^i - 1}{q^T - 1}$, em que $q = \dfrac{1-p}{p}$.

79 Se $p = 0{,}5$, então $w_i = \dfrac{i}{T}$.

80 Se T = 5 reais, $p = 0{,}6$ e $i = 3$ reais, então a probabilidade de o jogador finalizar o jogo com sucesso é inferior a 60%.

81 O desenvolvimento do jogo representa uma cadeia de Markov com estados absorventes.

82 As probabilidades w_i satisfazem as relações $w_0 = 0$, $w_T = 1$ e $w_i = (1-p)w_{i+1} + pw_{i-1}$, para $i = 1, ..., T - 1$.

RASCUNHO

Para o julgamento acerca da hipótese de independência entre duas variáveis X e Y em uma tabela de contingência, pode-se usar a estatística de teste $Q = \sum_{i=1}^{m} \sum_{j=1}^{n} \frac{(o_{i,j} - e_{i,j})^2}{e_{i,j}}$, em que $o_{i,j}$ e $e_{i,j}$ representam, respectivamente, as frequências observadas e esperadas sob hipótese nula de observações. Com base nessas informações, considerando que N denota o tamanho da amostra, julgue os próximos itens.

83 Assintoticamente, a estatística Q segue uma distribuição qui-quadrado com $n \times m - 1$ graus de liberdade.

84 Com relação à tabela de contingência ilustrada abaixo, correspondente ao resultado de uma pesquisa de consumo de refrigerantes, a estatística Q possui valor superior a 7.

X \ Y	1 não bebe	2 bebe pouco	3 bebe muito	total
1 (homem)	15	30	15	60
2 (mulher)	10	10	20	40
total	25	40	35	100

85 A estatística desse teste pode ser reescrita como $Q = \sum_{i=1}^{m} \sum_{j=1}^{n} \frac{o_{i,j}^2}{e_{i,j}} - N$.

86 Os valores esperados são dados por $e_{i,j} = \frac{o_{i\bullet}\, o_{\bullet j}}{N}$, em que $o_{i\bullet}$ e $o_{\bullet j}$ são, respectivamente, as somas dos elementos das linhas e das colunas de valores observados.

---

diâmetro (mm)	3	4	6	7
resistência (kg)	10	15	19	24

A tabela acima mostra os resultados acerca da resistência à tração de quatro fios de náilon com diferentes diâmetros. Com base nessas informações, julgue os seguintes itens, que versam sobre análise multivariada.

87 A matriz de correlação amostral é igual a $\begin{pmatrix} 1 & r \\ r & 1 \end{pmatrix}$, em que $r = \frac{32}{\sqrt{1.060}}$.

88 A variância amostral generalizada é inferior a 3.

89 O vetor de médias, ou centroide, é $\overline{X} = (6{,}5;\ 9{,}5;\ 12{,}5;\ 15{,}5)$.

90 A matriz de covariância amostral é $\frac{1}{3}\begin{pmatrix} 10 & 32 \\ 32 & 106 \end{pmatrix}$.

RASCUNHO

mês	1	2	3	4	5	6	7
vendas	2,5	3,1	2,8	3,7	4,2	4,0	4,4

Considerando a série temporal mostrada acima, correspondente à evolução mensal das vendas de certo produto, julgue os itens subsecutivos, acerca de séries temporais.

91 O crescimento mensal das vendas, calculado pelo método das semimédias e usando médias móveis de ordem 3, é superior a 0,3.

92 A venda prognosticada para o mês 8, com base em médias móveis de ordem 4, é igual a 3,8.

93 Para a série mostrada, as médias móveis de ordem 3 são 2,8; 3,2; 3,8; 3,966; 4,2.

---

massa corporal m (em kg)	frequências
$40 \le m < 50$	5
$50 \le m < 60$	15
$60 \le m < 70$	25
$70 \le m < 80$	30
$80 \le m < 90$	20
$90 \le m < 100$	5
total	100

A tabela acima mostra a distribuição de frequências das massas corporais (em kg) de um grupo de 100 pessoas. Considerando os coeficientes de assimetria $A_2 = \dfrac{Q_1 + Q_3 - 2Q_2}{Q_3 - Q_1}$ e de curtose $K = \dfrac{Q_3 - Q_1}{2(D_9 - D_1)}$, em que Q_i e D_i denotam os i-ésimos quartil e decil, respectivamente, julgue os itens seguintes.

94 A distribuição desses dados possui assimetria positiva.

95 O coeficiente de curtose é inferior a 0,25.

96 Da fórmula Sturges infere-se que a moda dessa distribuição é inferior a 75 kg.

97 A massa corporal média desse grupo de pessoas é superior a 70 kg.

8	0 1 1
9	2 2 3 4
10	0 1 1 2 3
11	4 5 6 6 7
12	0 1 1 2 2 2 3 6
13	1 2 2 3 4 8
14	4 4 5 7 9
15	0 1 2 3
16	0 1

A fim de combater um possível surto de dengue em determinada cidade, os agentes de saúde registraram, por 42 dias, a quantidade X de focos de dengue na localidade. Os dados estão representados no diagrama ramo-e-folhas acima. Com base nessas informações, julgue os itens de **98** a **104**.

98 A mediana da distribuição de X é igual a 2.

99 Se o coeficiente de variação da distribuição de X foi igual a 18,5%, então o desvio padrão de X foi superior a 20.

100 Considere que a distribuição dos dados X seja platicúrtica e que m_4 e s^2 representem, respectivamente, o quarto momento central e a variância amostral de X. Nessa situação, a razão $\dfrac{m_4}{s^4}$ é superior a 3.

101 O 3.º quartil de X é igual ao 30.º percentil.

102 O 1.º decil da distribuição de X corresponde à observação que ocupa a 5.ª posição.

103 A moda e a mediana da distribuição da quantidade X são iguais.

104 A média da variável X encontra-se no mesmo ramo da mediana.

RASCUNHO

		quantidade Y de ração		
		0	1	2
quantidade X de animais vendidos	0	0,27	0,11	0,07
	1	0,17	0,10	0,04
	2	0,11	0,08	0,05

A tabela acima mostra a distribuição de probabilidade conjunta da quantidade diária X de animais domésticos vendidos e a quantidade Y de ração vendida (em kg) por dia em determinada loja. Com base nessa tabela e nos conceitos de distribuição conjunta, julgue os itens seguintes.

105 A média diária da quantidade de ração vendida é igual a 1 kg.

106 O desvio padrão de X é inferior ao valor da média de X.

107 A covariância nula entre duas variáveis quaisquer W e Z permite concluir que ambas são independentes.

108 A probabilidade de que 2 animais sejam vendidos em determinado dia é superior à probabilidade de que 2 kg de ração sejam vendidos nesse mesmo dia.

109 As variáveis X e Y são independentes.

110 A probabilidade de que nenhum animal seja vendido em determinado dia é superior a 0,4.

111 Sob a condição de que em determinado dia são vendidos 2 kg de ração, a probabilidade de 2 animais serem vendidos nesse mesmo dia é superior a 0,3.

No que se refere a bancos de dados governamentais de informações na área de saúde e os aplicativos de análise estatística e epidemiológica, julgue os próximos itens.

112 Mediante o sistema de autorizações de internação hospitalar (AIH), é possível verificar a alocação dos recursos destinados a cada hospital que integra a rede do SUS.

113 Considere que se deseje construir um mapa que mostre a incidência da dengue nos 5.564 municípios brasileiros, usando-se o *software* SAS, e que base de dados contenha os nomes dos pacientes com essa doença e os respectivos municípios registrados. Nessa situação, o cálculo da quantidade de casos observados em cada município pode ser feito mediante o uso do procedimento PROC FREQ.

114 Um mapa do estado do Rio de Janeiro pode ser elaborado no *software* EPIINFO importando-se um arquivo de coordenadas geográficas do tipo *.cgm, dispobibilizado pelo Instituto Brasileiro de Geografia e Estatística (IBGE).

115 A planilha eletrônica Excel permite o cálculo de estatísticas descritivas e a produção de mapas temáticos, sem a necessidade de se instalarem módulos adicionais.

116 Os *software* SAS e EPIINFO permitem gerar, com base nos dados do sistema de informações sobre mortalidade (SIM), mapas que descrevem a distribuição regional da taxa de mortalidade.

117 O *software* STATA é compatível com o *software* SAS no que se refere à capacidade de gerenciamento de grandes bases de dados.

118 Dados acerca da incidência de casos de AIDS são disponibilizados pelo SINASC.

119 Não se efetua análises epidemiológicas com o *software* SPSS, visto que ele foi desenvolvido apenas para análise de dados sociais.

120 É possível criar, no Excel, arquivos do tipo_.dbf, para posterior utilização no *software* STATA.

RASCUNHO