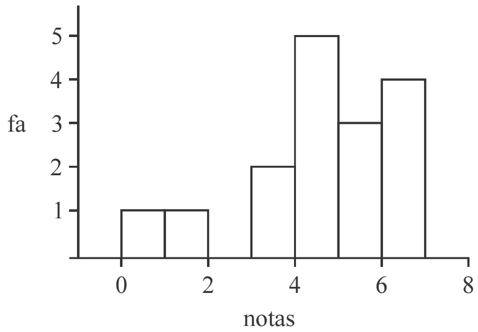
CONHECIMENTOS ESPECÍFICOS



Considerando que o histograma acima apresenta a distribuição das notas finais dos estudantes matriculados em determinada disciplina, e que fa representa a frequência absoluta, julgue os seguintes itens acerca de estatística descritiva.

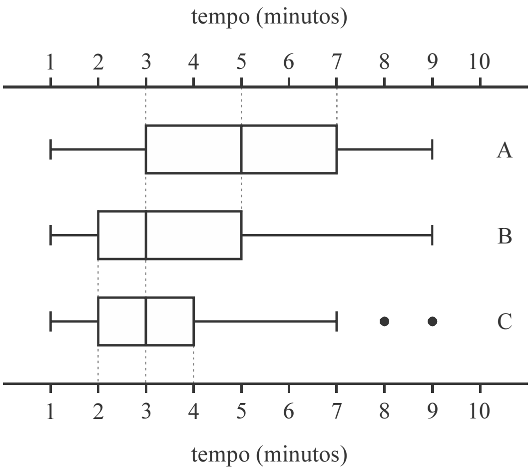
51 A moda da distribuição das notas é superior a 5,5.

52 A média das notas dos estudantes é superior a 4 e inferior a 5,5.

53 A mediana da distribuição das notas é igual ou superior a 4.

54 A amplitude total da distribuição das notas é igual a 6.

55 Nessa disciplina, o percentual de estudantes com notas inferiores a 2,5 é igual a 12,5%.



Os diagramas acima apresentam, esquematicamente, as distribuições dos tempos de execução, em minutos, de determinada tarefa administrativa sob três condições distintas de trabalho *A*, *B* e *C*. O coeficiente de variação quartil, expresso por CVQ = $\frac{IQ}{Q_3 + Q_1} \times 100\%$, em que Q_1 e Q_3 são respectivamente os quartis inferior e superior e *IQ* representa o intervalo interquartílico, é uma medida descritiva útil para a comparação dessas distribuições. Com base nessas informações, julgue os itens de **56** a **61**.

56 A média dos tempos observados sob a condição *A* é igual a 5 minutos.

57 De acordo com esses diagramas esquemáticos, o maior coeficiente de variação quartil encontra-se na distribuição *A*.

58 Entre as três distribuições, a distribuição *C* é a que exibe o maior intervalo interquartílico.

59 As formas dos diagramas de *box-plot* sugerem que as distribuições *B* e *C* possuem assimetria positiva, enquanto a distribuição *A* possui assimetria nula.

60 As distribuições *B* e *C* possuem os mesmos valores para os quartis Q_1 e Q_2, e o quartil superior em *B* corresponde ao quartil central (Q_2) da distribuição *A*.

61 O diagrama de *box-plot* correspondente à distribuição C indica a presença de dois ou mais valores atípicos (*outliers*).

RASCUNHO

Seja (Ω, A, P) um espaço de probabilidade, em que Ω é um conjunto (não vazio) que denota o espaço amostral, A é uma classe de subconjuntos de Ω e P representa uma medida de probabilidade. Considerando que os eventos aleatórios $B \in A$ e $C \in A$ sejam independentes e que $P(B) = 0{,}4$ e $P(C) = 0{,}6$, julgue os itens subsequentes.

62 $P(B|C) < P(C|B)$.

63 Como $P(B) < P(C)$, é correto concluir que $B \subset C$, ou seja, o evento B implica C.

64 Os eventos B e C são uma partição do espaço amostral Ω, pois $P(B \cup C) = 1$.

65 Sabendo-se que $P(B) = 1 - P(C)$, é correto concluir que C representa o evento "não B", de modo que C é o complementar do evento B relativamente ao espaço amostral Ω.

Tendo em vista que o número diário X de recursos administrativos protocolados em certa repartição pública segue uma distribuição de Poisson com taxa igual a $\ell n10$ processos por dia, julgue os itens que se seguem.

66 A distribuição de Poisson não possui memória, pois $P(X = k|X \geq 1) = P(X = k - 1)$, em que $k \geq 1$.

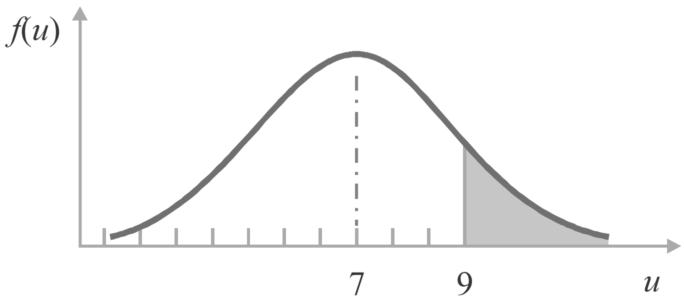
67 A moda da distribuição X é igual a zero.

68 Considerando o coeficiente de assimetria que se define pelo terceiro momento central, é correto afirmar que a distribuição de Poisson exibe assimetria positiva.

69 Em determinado dia, a probabilidade de não haver recurso protocolado é igual ou inferior a 0,1.

70 O desvio padrão da distribuição de X é inferior a $\ell n10$ processos por dia.

71 A distribuição do número diário de recursos administrativos apresenta coeficiente de variação igual a 1.



O consumo mensal de arroz (U, em toneladas) em determinado restaurante universitário segue uma distribuição normal, conforme a figura acima. Com base nessas informações, e considerando que $P(U > 9 \text{ ton}) = 0{,}16$ e $P(Z < 1) = 0{,}84$, em que Z representa a distribuição normal padrão, julgue os itens subsecutivos.

72 A média e o desvio padrão da distribuição U são, respectivamente, iguais a 7 ton e 4 ton.

73 Em determinado mês, a probabilidade de ocorrer o evento [5 ton < U < 9 ton] é inferior a 0,70.

74 $P(U > 9 \text{ ton} \mid U > 7 \text{ ton}) = 0{,}32$.

75 A probabilidade de ocorrer o evento [U < 4 ton] é inferior a 0,16.

76 Com base na figura, é correto afirmar que

$$P(X = 5) = P(X = 9) < P(X = 7).$$

Considerando que, em um circuito elétrico, a corrente I siga uma distribuição uniforme no intervalo (0,1) e que a potência W desse circuito seja expressa por $W = I^2$, julgue os itens a seguir relativos às transformações de variáveis.

77 A distribuição da variável aleatória é simétrica em torno da média.

78 A função de distribuição acumulada de W é $F(w) = P(W \leq w) = w^2$, em que $0 \leq w \leq 1$.

79 A mediana de W é igual a 0,25.

80 A variância de W é maior que 0,10.

81 A distribuição da potência possui função de densidade na forma $f(w) = 3w^2$, em que $0 \leq w \leq 1$.

82 Nesse circuito, a potência esperada é igual a $\frac{1}{3}$.

A lei dos grandes números é um importante resultado teórico que permite o estudo das propriedades de estimadores estatísticos, como, por exemplo, a média amostral . Considerando que μ representa a média populacional e que o desvio padrão populacional σ^2 seja finito, julgue os itens subsequentes a respeito desse assunto.

83 Se n é o tamanho da amostra, então na versão fraca da lei dos grandes números, $P(|\overline{X} - \mu| > \varepsilon) < \frac{1}{n}\left(\frac{\sigma}{\varepsilon}\right)^2$ para todo $\varepsilon > 0$.

84 Considere que um estimador T converge em média quadrática para um parâmetro τ à medida que o tamanho da amostra aumenta. Nessas condições, é correto afirmar que a lei fraca dos grandes números se aplica para esse estimador.

85 Segundo a lei forte dos grandes números, à medida que o tamanho da amostra aumenta, a estatística $\overline{X}$ converge em probabilidade para a média μ.

RASCUNHO

Considere que uma a1mostra aleatória simples com reposição, representada por $X_1, X_2, \cdots, X_n$, tenha sido retirada de uma grande população de estudantes para a avaliação da opinião sobre a qualidade dos serviços de transporte coletivo, em que

$$X_k = \begin{cases} 1, \text{se o estudante } k \text{ se mostrou satisfeito com os serviços;} \\ 0, \text{se o estudante } k \text{ se mostrou insatisfeito com os serviços.} \end{cases}$$

Com respeito ao total de satisfeitos na amostra, $Y_n = X_1 + X_2 + \cdots + X_n$, julgue os próximos itens.

86 À medida que o tamanho da amostra aumenta, a distribuição amostral da contagem Y_n se aproxima de uma distribuição normal padrão.

87 Segundo o teorema limite central, $\lim_{n\to\infty} \frac{Y_n}{n} = p$.

88 A estatística Y_n segue uma distribuição binomial com parâmetros n e p, em que p representa a fração populacional de estudantes satisfeitos com os serviços de transporte.

Julgue os itens que se seguem a respeito de propriedades de estimadores.

89 Um estimador somente será consistente se também for não viciado.

90 Os estimadores para a média e a variância de uma distribuição Normal obtidos pelo método dos momentos são, respectivamente, $\hat{\mu} = \overline{X}$ e $\hat{\sigma}^2 = \frac{1}{n-1}\left[\sum_{i=1}^{n} X_i^2 - n\cdot\overline{X}^2\right]$.

Acerca de intervalos de confiança e de credibilidade, julgue os itens subsequentes.

91 Considere duas amostras provenientes da mesma população, para as quais os intervalos de confiança para um parâmetro θ sejam, respectivamente, $\mathcal{J}_1 = [a, b]$ e $\mathcal{J}_2 = [c, d]$. No teste de hipóteses H_0: θ = 1 *versus* H_1: θ ≠ 1, caso a hipótese nula seja rejeitada na primeira amostra, mas não na segunda, é correto afirmar que $a \le \theta \le c$.

92 Não se pode definir um intervalo $\mathcal{J} = [a, b]$ de credibilidade HPD (*highest probability density*) para o parâmetro aleatório θ, tal que $P(\theta \le a) \ne P(\theta \le b)$.

93 Se $\mathcal{J}_1$ for o intervalo de confiança de tamanho 1 – α para o parâmetro θ e, se $\mathcal{J}_2$ for o intervalo de credibilidade 1 – α para o mesmo parâmetro, então, após selecionar a amostra, $P(\theta \in \mathcal{J}_1) = P(\theta \in \mathcal{J}_2)$.

94 Em geral, os intervalos de confiança são obtidos com base em uma quantidade pivotal apropriada que segue uma distribuição normal padrão.

No que se refere a testes de hipóteses, julgue os itens subsecutivos.

95 O tamanho amostral influencia o poder do teste e o nível de significância.

96 O teste de razão de verossimilhanças generalizadas (TRVG) é uma alternativa ao teste qui-quadrado de Pearson para a avaliação da independência em tabelas de contingência. Sabendo-se que o TRVG considera uma distribuição multinomial, é correto afirmar que a distribuição assintótica da sua estatística do teste possui número de graus de liberdade diferente do número de graus de liberdade da distribuição do teste de Pearson.

97 O poder de um teste de hipóteses tende a diminuir à medida que o nível de significância decresce.

Julgue o item abaixo, sobre a relação entre intervalo de confiança e teste de hipóteses.

98 Considere que o intervalo de confiança [3, 8] seja usado para testar as hipóteses H_0: $\mu = \mu_0$ *versus* H_1: $\mu > \mu_0$. Nesse cenário, a hipótese nula será rejeitada somente se $\mu_0 > 8$.

Com relação aos estimadores de mínimos quadrados e de máxima verossimilhança, julgue os itens seguintes.

99 Se o estimador de mínimos quadrados para os coeficientes de um modelo linear coincidir com o respectivo estimador de máxima verossimilhança, então a distribuição da variável resposta será Normal.

100 Se a amostra $X_1, X_2, \cdots, X_n$ for formada por observações dependentes, então a função de verossimilhança será igual a $L(\theta; X_1, \ldots, X_n) = \prod_{i=1}^{n} f(X_i|\theta)$.

101 Não há garantias de que o estimador de máxima verossimilhança seja não viesado.

Acerca dos modelos de regressão linear, julgue os itens a seguir.

102 O intercepto do modelo de regressão linear simples $y_i = \alpha + \beta x_i + \epsilon_i$, $\epsilon_i \sim N(0;\sigma^2)$ depende apenas da média de x e y para ser calculado.

103 Considere que um modelo linear múltiplo com interação seja dado por $Y_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i} - \beta_{12} X_{1i} X_{2i} + \epsilon_i$, $\epsilon_i \sim N(0;\sigma^2)$, em que $E(Y_i \mid X_{1i} = 0, X_{2i} = 0) = E(Y_i \mid X_{1i} = 1, X_{2i} = \kappa)$, $\kappa < \infty$, $\beta_1 > 0$, $\beta_2 > 0$, $\beta_{12} > 0$. Nessa situação, $\beta_2 \ne \beta_{12}$.

104 Considere que Y seja uma variável binária e Z seja definida por $Z = \ln\left(\frac{p}{1-p}\right) = \beta_0 + \beta_1 X$ em que $p = P(Y = 1)$ e X é uma covariável. Considere ainda que X assuma valores inteiros positivos, que $\beta_0 = 1$ e $\beta_1 = 0{,}2$ e que 2,72 e 7,39 sejam os valores aproximados, respectivamente, de e e e^2. Nessa situação, é correto afirmar que a chance de $Y = 1$ quando $X = 10$ é superior a 5 vezes a chance correspondente quando $X = 0$.

RASCUNHO

Considerando os métodos de inferência para os parâmetros do modelo de regressão, julgue os próximos itens.

105 Considerando um gráfico da distância de Cook para cada observação amostral que resultou de um ajuste por regressão linear, as observações influentes são aquelas que apresentam pequenas distâncias de Cook.

106 Uma medida de alavanca de um modelo de regressão é tal que $\phi_{ij} = \dfrac{\tau_{ij}^2}{\sum \tau_{ij}^2}$ em que τ_{ij} é o resíduo do modelo de regressão da variável X_i explicada pelas demais variáveis independentes do modelo para a observação j. Supondo um modelo de regressão com 2 variáveis dependentes, no qual apenas a k-ésima observação amostral seja influente, se $\phi_{1k} > \phi_{2k}$ então o valor X_{2k} tem um impacto maior que o valor X_{1k} na influência da k-ésima observação.

107 Um critério utilizado para se verificar a qualidade de ajuste de um modelo de regressão é o AIC (critério de informação de Akaike), que é dado por $AIC = 2(k - \ell(b; X))$, em que k é o número de parâmetros do modelo e $\ell(b; X)$ é a log-verossimilhança $\ell(\beta; X)$ calculada em $\beta = b$. Considerando a classe dos modelos com $k = \kappa$ parâmetros, então o AIC será mínimo se b for o estimador de máxima verossimilhança para β.

Com relação à análise de variância para verificação da qualidade de ajuste de um modelo de regressão, julgue os itens seguintes.

108 Em uma tabela de análise de variância para a qualidade de ajuste do seguinte modelo de regressão $Y_i = \beta_0 + \beta_1 X_1 + \beta_2 X_2 + \epsilon_i$, $\epsilon_i \sim N(0; \sigma^2)$ se a hipótese nula for rejeitada, então $\beta_0 = 0$ mas $\beta_1 \neq 0$, $\beta_2 \neq 0$.

109 Em um modelo de regressão linear simples, o quadrado médio associado ao modelo é menor que a respectiva soma de quadrados. O mesmo ocorre com o quadrado médio dos resíduos em comparação com a soma de quadrado dos resíduos.

110 O coeficiente de correlação múltipla R^2 pode ser calculado dividindo a soma de quadrados do resíduo pela soma de quadrado total.

111 Considerando um modelo de regressão no qual a média da variável resposta é aproximadamente zero, se o coeficiente de correlação múltipla (R^2) tende a 1, então $\sum \hat{y}_i^2 \rightarrow \sum y_i^2$.

A respeito dos métodos de análise de resíduos do modelo de regressão, julgue os itens subsequentes.

112 A suposição de homocedasticidade pode ser verificada através de um gráfico de resíduos.

113 Na análise de resíduos de um modelo de regressão, o diagrama de dispersão entre os resíduos do modelo ajustado e os valores preditos para a variável resposta permitem avaliar a ocorrência de heterocedasticidade.

Acerca do modelo $Y_i = \beta_0 + \beta_1 X_{1i} + \beta_2 X_{2i}^2 + \epsilon_i$, $\epsilon_i \sim N(0; \sigma^2)$ julgue os itens subsecutivos.

114 Se as variáveis X_1 e X_2 possuírem correlação próxima a 1, então os parâmetros β_1 e β_2 serão linearmente independentes.

115 A variável $Z_i = \exp(Y_i)$ tem uma distribuição log-Normal.

116 Esse modelo é linear.

Julgue os itens a seguir, relativos às técnicas de amostragem.

117 No cálculo do tamanho amostral para a comparação de proporções através da expressão $n \geq \frac{1}{4}\left(\frac{z}{m}\right)^2$ em que m é a margem máxima de erro pretendida e z é o quantil da normal padrão que define a significância do teste, a tendência é que as amostras sejam maiores que aquelas calculadas considerando uma abordagem não conservativa.

118 Na amostragem aleatória simples sem reposição (AASc), a probabilidade de seleção de elementos é praticamente igual à probabilidade de seleção caso a amostragem seja com reposição.

119 Na amostragem estratificada, a variância dentro dos estratos deve ser pequena, enquanto a variância entre os estratos deve ser grande. Na amostragem por conglomerados, por outro lado, é regra geral que a variância dentro dos conglomerados seja maior que a variância entre os conglomerados.

120 Uma dúvida comum entre as pessoas ao observarem os resultados de uma pesquisa eleitoral é acerca da validade dos resultados obtidos com base em uma amostra muito menor frente ao tamanho da população. De fato, essa dúvida procede, pois o tamanho populacional é um dado relevante no cálculo do tamanho mínimo de uma amostra.

RASCUNHO

RASCUNHO