

-- CONHECIMENTOS ESPECÍFICOS --

Supondo que N seja uma variável aleatória discreta com função de probabilidade $P(N = n) = 0,5^{n+1}$, em que $n \in \{0,1,2, \dots\}$, e considerando outra variável aleatória discreta X que segue a distribuição condicional na forma $P(X = x|N = n) = \binom{n}{x} 0,5^n$, na qual $0 \leq x \leq n$, julgue os seguintes itens.

- 51 A variável aleatória N possui variância igual a 2.
 52 A curva de regressão da variável aleatória X sobre o ponto $N = n$ é dada pela média condicional $E(X|N = n) = 0,25n$.
 53 O valor esperado de X é igual a 1.
 54 $P(X = 0) = \frac{2}{3}$.
 55 A variância de X é igual a 0,75.

Considerando que W represente uma variável aleatória absolutamente contínua tal que $P(W \geq w) = e^{-2w}$, para $w \geq 0$, e $P(W \geq w) = 1$, para $w < 0$, e definindo a variável aleatória discreta T tal que $P(T = t) = P(t \leq W \leq t + 1)$, para $t \in \{0,1,2,3, \dots\}$, julgue os seguintes itens.

- 56 $P(T > 0) = P(W \geq 1)$.
 57 Se U for uma variável aleatória uniforme contínua no intervalo $[0,1]$ e se $f_U(u)$ denota sua função de densidade de probabilidade, então $P(W \geq w|T = t) = f_U(u)$, em que $u = w - t$, para quaisquer $w \geq 0$ e $t \in \{0,1,2,3, \dots\}$.
 58 É correto afirmar que $P(T = 1) = P(T = 0) \times P(W \leq 1)$.
 59 O valor esperado de T é igual a 0,5.
 60 $P(T = 3|W \geq 3) = P(W \leq 1)$.

Sabendo que Z_1 e Z_2 são cópias independentes de uma distribuição normal padrão e considerando a razão $R = \frac{Z_1}{Z_2}$, julgue os próximos itens.

- 61 A diferença $D = Z_1 - Z_2$ e a soma $S = Z_1 + Z_2$ são variáveis aleatórias normais independentes.
 62 $P(R^2 > 1) = 0,5$.
 63 A soma dos quadrados $Z_1^2 + Z_2^2$ se distribui conforme uma distribuição χ^2 com dois graus de liberdade.
 64 A média da distribuição da razão R é igual a 1.
 65 A variância do quadrado da razão, R^2 , é igual a 1.

Considerando que X seja uma variável aleatória absolutamente contínua cuja função de distribuição acumulada seja representada por $F(x)$, na qual $x \in \mathbf{R}$, julgue os próximos itens, referentes à transformação $Y = F(X)$.

- 66 $P(Y < 0,95) = 0,90$.
 67 $E[\ln(Y)] \geq -0,5$.
 68 A variância de Y é igual a 0,25.
 69 A mediana de Y é igual a 0,5.
 70 Y^4 segue uma distribuição beta.

Considerando uma amostra aleatória simples X_1, \dots, X_n retirada de uma população normal com média 1 e variância 2, julgue os seguintes itens, acerca da soma ponderada $S_n = \sum_{k=1}^n 0,5^k X_k$.

- 71 A razão $\frac{S_n - 1}{\sqrt{2/n}}$ segue uma distribuição t de Student com $n - 1$ graus de liberdade.
 72 $\lim_{n \rightarrow \infty} E[S_n] = 1$.
 73 A variância de S_n tende para zero à medida que n aumenta.
 74 S_n converge em probabilidade para o valor da média populacional.
 75 S_n segue uma distribuição normal para qualquer tamanho de amostra $n \geq 1$.

Espaço livre

Para verificar o número x de famílias que seriam atendidas com a instalação de cisternas, foi feito um levantamento por amostragem aleatória simples em 20 municípios, tendo sido obtidos os resultados a seguir.

$$\sum_{i=1}^{20} x_i = 90$$

$$\sum_{i=1}^{20} x_i^2 = 460$$

Com base nessas informações e sabendo que os dados apresentam distribuição normal, julgue os itens que se seguem, considerando que:

$$P(Z > 1,96) = 0,025,$$

$$P(Z > 1,645) = 0,05,$$

$$P(T_{19} > 2,093) = 0,025,$$

$$P(T_{19} > 1,729) = 0,05,$$

$$P(T_{20} > 2,086) = 0,025,$$

$$P(T_{20} > 1,725) = 0,05,$$

em que Z denota uma variável aleatória normal padrão e T_g representa uma distribuição t de *Student* com g graus de liberdade.

- 76** Caso se pretenda testar que a média populacional é igual a 5, então a estatística do teste teria 20 graus de liberdade.
- 77** Em média, a instalação das cisternas beneficiaria mais de 4 famílias por município.
- 78** O limite inferior do intervalo de 95% de confiança para a média populacional é menor que 4 e o limite superior é maior que 5.
- 79** A estimativa da variância da média estimada é maior que 0,2.
- 80** Caso fosse necessário escolher entre a média amostral e a mediana amostral para a estimação da média populacional, a mediana seria a melhor escolha visto que ela melhor representa a tendência central dos dados, além de possuir as quatro propriedades desejáveis de um estimador: não-viesado, suficiente, consistente e eficiente.

Espaço livre

Na tabela a seguir, são apresentados os dados coletados em um estudo realizado com o objetivo de verificar a associação entre a escolaridade dos produtores rurais e a participação destes em programas de auxílio financeiro do governo federal.

participa de programas	escolaridade		
	até nível fundamental	nível médio	nível superior
sim	20	10	5
não	12	20	30

Com base nessas informações e nos conceitos de testes de hipóteses, julgue os itens subsequentes.

- 81** No total foram avaliadas mais de 95 pessoas.
- 82** A estatística qui-quadrado do teste de independência entre as variáveis “participa de programas” e “escolaridade” possui 2 graus de liberdade.
- 83** No teste qui-quadrado para a avaliação da independência entre as variáveis “participa de programas” e “escolaridade”, a hipótese nula desse teste é que tais variáveis sejam mutuamente dependentes.
- 84** Se o p -valor do teste qui-quadrado de independência for inferior a 1%, então, para um nível de significância de 5%, a hipótese nula deverá ser rejeitada.
- 85** Caso o critério escolaridade fosse dividido em apenas duas categorias — “nível médio” e “nível superior” —, então a estatística Z (normal padrão) que compara a proporção de pessoas que participam de algum programa de auxílio financeiro do governo federal com a das pessoas que possuem nível superior é igual à estatística qui-quadrado, uma vez que esta é uma extensão do teste de comparação de duas proporções.

Com base em estudos de alguns pesquisadores, acredita-se que os açudes de certa região brasileira possuem profundidade média igual ou inferior a 1,5 m. A fim de verificar essa informação, um instituto independente, que acredita que a profundidade média é maior que o relatado, mediu as profundidades de uma amostra de 10 açudes, escolhidos aleatoriamente, obtendo os dados a seguir, em metros.

2,0	1,0	1,5	1,0	1,5	2,0	1,0	1,5	1,0	0,5
-----	-----	-----	-----	-----	-----	-----	-----	-----	-----

Com base nessas informações e nos conceitos de testes de hipóteses, julgue os itens seguintes.

- 86** No caso de a média populacional ser $\mu = 2$, basta calcular poder = $1 - P(\bar{x} \geq 1,5 | \mu = 2)$ para se determinar o poder do teste, independentemente do nível de significância do teste.
- 87** Quanto mais distante a média hipotética estiver da média populacional verdadeira, maior será o poder do teste.
- 88** A profundidade média dos açudes, conforme a amostra, é superior a 1,5 m.
- 89** Uma hipótese alternativa de um teste para a média populacional poderia ser escrita como $H_1: \bar{x} \leq 1,5$.
- 90** Considerando-se que os dados seguem uma distribuição normal e sabendo-se que $\sum_{i=1}^{10} (x_i - \bar{x})^2 = 2,1$, é correto afirmar que a estatística do teste t de *Student* para a média populacional é inferior a 1, em valor absoluto.

No que diz respeito à estimação pontual e intervalar, julgue os seguintes itens.

- 91** Os intervalos de confiança e de credibilidade propiciam a mesma interpretação do ponto de vista estatístico.
- 92** Se $Var(\bar{x}) = \frac{\sigma^2}{n}$ e $Var(M_d) = \frac{\sigma^2 \pi}{2n}$, então a média (\bar{x}) e a mediana (M_d) são igualmente eficientes.
- 93** Caso um intervalo de 95% de confiança para determinada característica média seja dado por [22,5; 27,5], então a estimativa pontual é igual a 25.
- 94** Se um intervalo de 95% de confiança para uma dada proporção for dado por [0,25; 0,35], então o erro amostral será de 5%.
- 95** Se um intervalo de 95% de confiança para determinada característica média for dado por [30; 50], então, do ponto de vista frequentista, a probabilidade de a média populacional estar nesse intervalo é de 90%.

Considerando um modelo de regressão linear simples na forma $\hat{M} = -10 + 3 \times N$, em que N denote o número de mãos por cacho produzido por um tipo de bananeira e \hat{M} represente a massa do cacho ajustada pelo método de mínimos quadrados ordinários, e sabendo que seu coeficiente de explicação é igual a 0,90 e que o tamanho da amostra é de 102 cachos, julgue os próximos itens.

- 96** 90% da variação do número de mãos por cacho pode ser explicada pela massa do cacho.
- 97** A estimativa da variância do coeficiente angular do modelo ajustado é igual a 0,01.
- 98** O desvio padrão das massas dos cachos é inferior ao desvio padrão dos números de mãos por cacho.
- 99** A correlação entre \hat{M} e N é igual a 1.
- 100** As estimativas dos coeficientes do modelo invertido $\hat{N} = \frac{10}{3} + \frac{M}{3}$ resultam da aplicação do critério de mínimos quadrados ordinários.

fonte de variação	soma dos quadrados
regressão	281
erro	39
total	320

A tabela de análise de variância precedente diz respeito ao ajuste de um modelo de regressão linear na forma $y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \varepsilon$, em que β_0 , β_1 e β_2 denotam os coeficientes do modelo e ε se refere a um erro aleatório normal com média zero e desvio padrão σ . O ajuste é feito pelo critério de mínimos quadrados ordinários e o tamanho da amostra é igual a 81.

Com base na situação hipotética apresentada, julgue os seguintes itens.

- 101** O valor da razão F do teste linear geral, cuja hipótese nula é $H_0: \beta_0 = \beta_1 = \beta_2 = 0$, é igual ou inferior a 10.
- 102** A correlação linear entre y e x_1 é superior a 0,95.
- 103** O coeficiente de explicação ajustado é igual a 0,875.
- 104** A estimativa de σ é igual ou superior a 1.
- 105** O desvio padrão amostral da variável resposta é igual a 2.

Um órgão governamental está desenvolvendo um estudo para avaliar a opinião dos moradores da região Nordeste do Brasil sobre a mudança de sua qualidade de vida após a transposição do Rio São Francisco. Na coleta de dados da pesquisa, feita por meio de questionário, os moradores participantes devem responder apenas se estão satisfeitos ou insatisfeitos com a transposição.

Com base nessa situação hipotética e nos conceitos de amostragem e considerando que $P(Z > 1,96) = 0,025$, $P(Z > 1,645) = 0,05$ e $P(Z > 1,28) = 0,10$, em que Z representa a distribuição normal padrão, julgue os seguintes itens.

- 106** Caso o órgão governamental tenha o cadastro das pessoas participantes do estudo, então o plano de amostragem aleatória simples será o mais indicado, pois a existência desse cadastro por si só permite excluir outras possibilidades de planos amostrais, como a amostragem estratificada e sistemática.
- 107** Caso o órgão governamental tenha o cadastro dos participantes do estudo e esse cadastro esteja ordenado pelo CPF ou por ordem alfabética, então os resultados (estimativa pontual e variância) das técnicas amostragem aleatória simples e sistemática podem ser equivalentes.
- 108** Como a opinião dos moradores é uma variável qualitativa, apenas variando nas categorias de “satisfeito” ou “insatisfeito”, então, para efeitos de cálculo do tamanho amostral de uma amostragem aleatória simples, a variância máxima seria igual a 0,25.
- 109** Dada a possibilidade de os dados das pessoas estarem alinhados à unidade da Federação (UF) e dado o fato de não se ter ideia da proporção de pessoas “satisfeitas” ou “insatisfeitas” com a transposição do Rio São Francisco, então o tamanho da amostra de uma amostragem estratificada por UF, com alocação proporcional, seria menor que o tamanho da amostra de uma amostragem aleatória simples.
- 110** Considerando-se que a pesquisa tenha um custo elevado e que não se tem ideia da proporção de pessoas “satisfeitas” ou “insatisfeitas” com a transposição do Rio São Francisco, o tamanho da amostra mínima de uma amostragem aleatória simples, com 95% de confiança e 5% de margem de erro, seria superior a 400 pessoas.
- 111** Caso o questionário aplicado na pesquisa também contemple perguntas de variável quantitativa, como renda e número de dependentes, então o tamanho amostral de uma amostragem aleatória simples, se for utilizada uma proporção de 0,5 para as variáveis dicotômicas, será maior que o tamanho amostral de uma amostragem aleatória simples caso o coeficiente de variação das variáveis quantitativas seja inferior a 50%.
- 112** Considere que o banco de dados disponível para o sorteio da amostra esteja ordenado por alguma variável altamente correlacionada com a opinião dos moradores sobre a mudança em sua qualidade de vida após a transposição do Rio São Francisco. Nesse caso, a variância da amostragem sistemática seria tão ou mais eficiente que uma amostragem aleatória simples.

Em um estudo, foram investigadas diversas características de famílias em uma amostra de 400 domicílios, a fim de o governo melhor direcionar as políticas sociais desenvolvidas em prol da população, que é de 2 milhões de pessoas.

Com base nessa situação hipotética, julgue os itens subsequentes.

- 113** O fato de serem selecionados domicílios e o de as características a serem investigadas pertencerem a pessoas fazem que o plano conglomerado seja o mais indicado no estudo, uma vez que cada domicílio é um conglomerado de pessoas.
- 114** Suponha que uma das variáveis investigadas no estudo seja a renda e que o erro padrão para essa variável seja igual a 10, calculado como se fosse uma amostragem aleatória simples nessa mesma pesquisa; suponha, ainda, que o efeito do planejamento seja igual a 2. Nesse caso, a variância da amostragem por conglomerados para essa mesma variável seria inferior a 100.
- 115** Se todas as pessoas dos domicílios da amostra forem entrevistadas, ou apenas uma amostra dessas pessoas em cada domicílio for entrevistada, então haverá uma amostragem conglomerada em 1 estágio.
- 116** O peso amostral do i -ésimo domicílio amostrado, sabendo-se que nele moram 5 pessoas e fazendo-se uma seleção com probabilidade proporcional ao tamanho, é inferior a 100.
- 117** Em uma amostragem por conglomerados em dois estágios, em que as unidades primárias de seleção são selecionadas com probabilidade proporcional ao tamanho, uma forma de fazer o plano amostral autoponderado, ou seja, com o mesmo peso para todas as unidades, é selecionar uma quantidade fixa de unidades secundárias de seleção.
- 118** Se os domicílios amostrados tivessem 2,5 moradores, em média, então o peso amostral do plano de amostragem aleatória simples seria superior a 1.000.
- 119** Se o efeito do planejamento for igual a 2 e se os domicílios amostrados tiverem 2,5 moradores, em média, então o tamanho de uma amostra aleatória simples com os mesmos parâmetros do estudo domiciliar será superior a 1.000 pessoas.
- 120** Ao se calcular o tamanho da amostra em uma amostragem por conglomerados em dois estágios, não existe diferença em se calcular, em primeiro lugar, ou a quantidade de unidades primárias de seleção ou a quantidade de unidades secundárias de seleção.

Espaço livre