

CONHECIMENTOS ESPECÍFICOS

Supondo que o custo unitário X de um processo de execução fiscal na justiça federal seja descrito por uma distribuição exponencial com média igual a R\$ 5.000, julgue os próximos itens.

- 51 Para cada $r = 1, 2, 3, \dots$, o momento de ordem r da variável X , $E[X^r]$, é tal que $E[X^r] = (\text{R\$ } 5.000)^r$.
- 52 O coeficiente de variação de X é igual a 1.
- 53 A variável aleatória $Y = e^{-X}$ segue a distribuição Beta.
- 54 $P(X > 5.000 | X > 1.000) < P(X > 5.000)$.
- 55 A mediana da distribuição do custo unitário X é inferior a R\$ 5.000.

A quantidade de clientes atendidos em cada minuto pelos empregados 1 e 2 em um balcão de atendimentos é expressa por $T = Y_1 + Y_2$, em que Y_1 = quantidade de clientes atendidos (por minuto) pelo empregado 1, e Y_2 = quantidade de clientes atendidos (por minuto) pelo empregado 2.

Considerando que, nessa situação hipotética, Y_1 e Y_2 sejam variáveis aleatórias independentes, seguindo uma mesma distribuição Y , cuja função de probabilidade é $P(Y = y) = 0,1 \times 0,9^y$, para $y = 0, 1, 2, \dots$, julgue os seguintes itens.

- 56 A variância de Y é inferior a 87.
- 57 As variáveis aleatórias Y_1 e Y_2 possuem assimetrias negativas.
- 58 Se $H = \min(Y_1, Y_2)$ é o menor entre Y_1 e Y_2 , então, para $h = 0, 1, 2, \dots$, $P(H = h) = 0,19 \times 0,81^h$.
- 59 Se $E[T]$ = quantidade média de clientes atendidos em cada minuto por esses dois empregados, então $E[T] < 17$.
- 60 A soma T segue uma distribuição binomial negativa.

Supondo que Z seja uma distribuição normal padrão, considere as seguintes transformações de variáveis aleatórias: $W = 1 - Z$ e $V = Z^2 - W^2 + 1$. A respeito dessas variáveis aleatórias, julgue os itens a seguir.

- 61 A variável aleatória V segue distribuição normal.
- 62 A covariância entre W e Z é igual a -1.
- 63 A variância da variável aleatória V é igual a 2.
- 64 A variável aleatória W segue distribuição normal com variância unitária.

Em determinado tribunal, a probabilidade de extinção de um processo judicial com julgamento de mérito é $P(A \cap B) = 0,05$, e a probabilidade de extinção de um processo judicial sem julgamento de mérito é $P(A \cap \bar{B}) = 0,15$, em que os eventos \bar{A} e \bar{B} são eventos mutuamente excludentes e denotam, respectivamente, os eventos complementares dos eventos A e B .

Com referência a essa situação hipotética, julgue os itens que se seguem.

- 65 $P(B) = 0,25$.
- 66 $P(A|\bar{B}) > 0,80$.
- 67 $P(\bar{A} \cup B) > P(\bar{A})$.

A respeito do total amostral $T_n = X_1 + X_2 + \dots + X_n$, em que X_1, X_2, \dots, X_n é uma amostra aleatória simples retirada de uma distribuição gama com média μ e desvio padrão σ , julgue os próximos itens.

- 68 O valor esperado do total amostral T_n é igual a μ .
- 69 O total amostral T_n segue distribuição gama com desvio padrão $n \times \sigma$.

70 Se n , o tamanho da amostra, aumenta, então a razão

$$\left[\frac{T_n - n \times \mu}{\sigma} \right]^2 \text{ converge quase certamente a uma distribuição}$$

qui-quadrado com 1 grau de liberdade.

Em um tribunal, entre os processos que aguardam julgamento, foi selecionada aleatoriamente uma amostra contendo 30 processos. Para cada processo da amostra que estivesse há mais de 5 anos aguardando julgamento, foi atribuído o valor 1; para cada um dos outros, foi atribuído o valor 0. Os dados da amostra são os seguintes:

1 1 0 0 1 0 1 1 0 1 1 1 1 0 1 1 1 0 1 0 1 0 1 0 1 1 1 0 1 1

A proporção populacional de processos que aguardam julgamento há mais de 5 anos foi denotada por p ; a proporção amostral de processos que aguardam julgamento há mais de 5 anos foi representada por \hat{p} .

Com referência a essas informações, julgue os itens a seguir, considerando que, para a distribuição normal padrão Z , $P(Z > 1,28) = 0,10$; $P(Z > 1,645) = 0,05$; e $P(Z > 1,96) = 0,025$.

- 71 O teste t de Student seria apropriado para testar se, nesse tribunal, p é maior que 50%, com 29 graus de liberdade.
- 72 Em um teste unilateral à direita, cujo objetivo seja testar se metade dos processos levam, em média, mais de 5 anos para serem julgados, o valor crítico de processos aguardando julgamento por mais de 5 anos, na amostra de 30 processos, seria superior a 20 processos, considerando 10% de significância.
- 73 O teste não paramétrico de Wilcoxon seria uma alternativa para testar se p é maior que 50%.
- 74 A variância da proporção amostral \hat{p} sob a hipótese nula $H_0: p = 0,5$ é menor que 0,1.
- 75 Estima-se que, nesse tribunal, $p > 60\%$.

A tabela a seguir foi usada para verificar se existe alguma relação entre a variável p = quantidade de páginas de um processo e a variável t = tempo necessário para a conclusão desse processo.

p	t	
	$t \leq 1$ ano	$t > 1$ ano
$p < 50$	50 processos	30 processos
$50 \leq p \leq 100$	40 processos	70 processos
$p > 100$	10 processos	100 processos

Considerando que X_j^2 denota a distribuição qui-quadrado com j graus de liberdade e que $P(X_1^2 > 3,84) = 0,05$, $P(X_2^2 > 5,99) = 0,05$, $P(X_3^2 > 7,81) = 0,05$, $P(X_1^2 > 2,71) = 0,10$, $P(X_2^2 > 4,60) = 0,10$, $P(X_3^2 > 6,25) = 0,10$, julgue os itens que se seguem, tendo como referência as informações na tabela.

- 76 O teste não paramétrico de Friedman pode ser aplicado nos dados com 298 graus de liberdade.
- 77 Caso fossem aplicados o teste exato de Fisher e o teste qui-quadrado convencional, as conclusões seriam diferentes devido ao fato de o teste exato de Fisher ser exato e o teste qui-quadrado ser aproximado.
- 78 A amostra utilizada para o estudo contém mais de 290 processos.
- 79 Se for utilizado o teste qui-quadrado para verificar se existe associação entre as variáveis referidas, então o grau de liberdade do referido teste será igual a 2.
- 80 Se, a partir do teste de independência entre as referidas variáveis, o valor calculado da estatística qui-quadrado for superior a 60, então será correto concluir com 95% de confiança que existe associação entre essas variáveis.

Diversos processos buscam reparação financeira por danos morais. A tabela seguinte mostra os valores, em reais, buscados em 10 processos — numerados de 1 a 10 — de reparação por danos morais, selecionados aleatoriamente em um tribunal.

processo	valor
1	3.700
2	3.200
3	2.500
4	2.100
5	3.000
6	5.200
7	5.000
8	4.000
9	3.200
10	3.100

A partir dessas informações e sabendo que os dados seguem uma distribuição normal, julgue os itens subsequentes.

- 81 Se μ = estimativa pontual para a média dos valores buscados como reparação por danos morais no referido tribunal, então $3.000 < \mu < 3.300$.

- 82 Uma forma de verificar a normalidade dos dados seria pelo teste de Kolmogorov-Smirnov, calculando-se o valor crítico, para n pequeno ($n < 10$), pela aproximação

$$D_{\alpha(2),n} = \sqrt{\frac{-\ln(\frac{\alpha}{2})}{2n}}$$

em que α é o nível de significância do teste.

- 83 Caso seja de interesse testar, por exemplo, se a média dos valores é diferente de 3.500, para calcular o p -valor do teste no referido estudo é suficiente multiplicar a $P(\bar{x} > 3.500)$ por 2, em que \bar{x} é a média amostral.
- 84 Nessa situação, se for possível usar o teste t de Student, então esse teste teria 9 graus de liberdade.
- 85 Na situação em questão, em que os dados seguem uma distribuição normal, o teste não paramétrico de Wilcoxon é menos poderoso que o teste t de Student.

Considerando que \hat{Y} seja uma variável resposta ajustada por um modelo de regressão em função de uma variável explicativa X , que x_1, \dots, x_n representem as réplicas de X e que $\hat{\alpha}$ e $\hat{\beta}$ sejam as estimativas dos parâmetros do modelo, julgue os itens a seguir.

- 86 No método de mínimos quadrados, a condição de estimativas não viesadas significa que os erros terão variância positiva.
- 87 Um modelo na forma $\hat{Y} = \hat{\alpha} \cdot \hat{\beta}^X$, em que $\hat{\alpha} > 0$ e $\hat{\beta} > 0$, apesar de ser exponencial em sua estrutura original, é linearizável, podendo ser tratado pelos métodos de regressão linear.
- 88 Em um modelo linear $\hat{Y} = \hat{\alpha} + \hat{\beta}X$, com coeficientes obtidos pelo método dos mínimos quadrados ordinários, sendo $\hat{\alpha} > 0$, a média dos valores estimados de Y é igual à média dos valores de X multiplicados por $\hat{\beta}$.
- 89 Em um modelo linear $\hat{Y} = \hat{\alpha} + \hat{\beta}X$, a hipótese de homoscedasticidade significa que a variância dos erros deve ser constante, e o valor esperado dos erros deve ser zero.
- 90 No modelo linear $Y = \alpha + \beta X + e$, considere que para cada valor x_i de X corresponda um erro e_i , que é uma variável aleatória. Nessa situação, a hipótese de erros não autocorrelacionados implica que $cov(e_i, e_j) = 0$ para $i \neq j$.

Em um modelo de regressão linear simples na forma $y_i = a + bx_i + \varepsilon_i$, em que a e b são constantes reais não nulas, y_i representa a resposta da i -ésima observação a um estímulo x_i e ε_i é o erro aleatório correspondente, para $i = 1, \dots, n$, considere que $\sum_i (x_i - \bar{x})^2 = 10$, em que $\bar{x} = (x_1 + \dots + x_n)/n$, e que o desvio padrão de cada ε_i seja igual a 10, para $i = 1, \dots, n$.

A respeito dessa situação hipotética, julgue os itens que se seguem.

- 91 A heteroscedasticidade é um problema que surge quando o valor esperado dos erros não é zero.
- 92 Se \hat{b} representar o estimador de mínimos quadrados ordinários do coeficiente b , então $var[\hat{b}] = 10$.

A equação seguinte foi obtida de um modelo de regressão linear múltipla ajustado sobre 12 amostras, em que cada valor entre parênteses abaixo do coeficiente representa o erro-padrão desse coeficiente, e representa o erro, D é uma variável *dummy* que assume o valor 0 caso não ocorra determinado evento e 1 caso ocorra, e X_1 e X_2 são duas variáveis regressoras.

$$Y = 340 + 150 X_1 + 540 X_2 + 89 D + e$$

(40) (5) (45) (20)

A tabela de análise de variância (ANOVA) proporcionada pelo referido modelo é apresentada a seguir.

fonte de variação	soma de quadrados	graus de liberdade	quadrados médios	teste F
modelo	2.000.000	a	c	F_C
erro	16.000	8	2.000	
total		11	d	

Com base nas informações e na tabela apresentadas, julgue os próximos itens.

- 93 Dado o valor crítico da estatística t de Student para 8 graus de liberdade a 5% de significância, $t_{8,5\%} = 2,3$, rejeita-se a hipótese de que cada um dos coeficientes da regressão seja nulo.
- 94 O valor de a reflete a quantidade de variáveis explicativas, e deve ser igual a 3.
- 95 A soma dos quadrados totais é igual a 2.016.000.
- 96 Fixando-se determinado ponto (X_1, X_2) , a ocorrência do evento representado por D faz que a estimativa de Y diminua em mais de 80 unidades.
- 97 O coeficiente de determinação ajustado dessa regressão, \bar{R}^2 , é maior que o coeficiente de determinação R^2 .

Considerando um modelo de regressão linear com erros heteroscedásticos, julgue os itens seguintes.

- 98 Para corrigir a heteroscedasticidade, como regra geral, é suficiente fazer a regressão da variável dependente em função das raízes quadradas das variáveis independentes.
- 99 Para um modelo de regressão linear na forma $Y = \alpha + \beta X + e$, em que Y representa a variável resposta, X é a variável regressora, e e denota o erro aleatório, o teste de Goldfeld–Quandt consiste em fazer duas regressões: uma com os maiores valores de X e outra com os menores valores de X , e verificar se as variâncias são distintas.
- 100 Para um modelo de regressão linear múltiplo, o teste de White permite detectar a heteroscedasticidade a partir da regressão de cada erro estimado da regressão original com as variáveis explicativas e seus inversos.

A respeito da autocorrelação dos erros de um modelo de regressão linear, julgue os itens subsequentes.

- 101 A autocorrelação dos erros, desde que não seja unitária em termos absolutos, insere um viés nas estimativas da variável dependente.
- 102 O teste de Durbin–Watson é um teste que permite identificar a autocorrelação serial de primeira ordem.
- 103 Como regra geral, a presença de autocorrelação dos erros é um problema que não pode ser corrigido, de modo que a modelagem por regressão deve ser abandonada quando detectado esse problema.
- 104 Ocorre autocorrelação dos erros caso os erros da regressão sigam um processo autorregressivo de ordem 1, ou seja, um AR(1).
- 105 Na presença de autocorrelação de erros, o estimador mais eficiente da regressão por mínimos quadrados ordinários continua sendo BLUE (*best linear unbiased estimator*), ou seja, melhor estimador linear não viesado.

Para um estudo sobre a gestão de riscos jurídicos em determinado tribunal, um analista efetuará simulações de Monte Carlo com base em realizações de variáveis aleatórias contínuas Y (exponencial, com média m), U (uniforme no intervalo $[0,1]$) e Q (qui-quadrado, com k graus de liberdade).

Considerando que Y , U e Q sejam mutuamente independentes, julgue os próximos itens.

- 106 Suponha que Y_1, Y_2, \dots, Y_n sejam n realizações independentes retiradas de uma distribuição exponencial com média m . Nessa situação, a média $\frac{1}{n} \sum_{i=1}^n \cos(Y_i)$ representa uma estimativa da integral $\frac{1}{m} \int_0^{+\infty} \cos(t) \exp(-\frac{t}{m}) dt$.
- 107 O produto $\text{sen}(2\pi U) \times \sqrt{Y}$ segue uma distribuição normal com média nula.
- 108 Realizações G de uma distribuição gama com média $2m$ podem ser obtidas com base na transformação $G = Y - m \times \ln(U)$.
- 109 A transformação \sqrt{Q} proporciona uma realização da distribuição normal padrão.
- 110 Caso W seja uma realização retirada de uma distribuição normal com média nula e variância k , será correto afirmar que o produto $W \times Q^{-1/2}$ é realização de uma distribuição t de Student com k graus de liberdade.

Uma amostra aleatória simples Y_1, Y_2, \dots, Y_n , retirada de uma população Bernoulli, é tal que

$$P(Y_k = y) = \theta^y (1 - \theta)^{1-y},$$

para $y = 0$ ou 1 , $0 < \theta < 1$ e $k = 1, 2, \dots, n$. O objetivo é efetuar inferências acerca do parâmetro θ mediante aplicação de métodos computacionais.

Considerando que para $r \geq 0$, $\hat{\theta}^{(r)}$ represente a estimativa de θ obtida na r -ésima iteração de um algoritmo de estimação, julgue os seguintes itens.

- 111 O método de Monte Carlo via cadeia de Markov (MCMC) pertence à classe de algoritmos de estimação não sequencial, em que $\{\hat{\theta}^{(r)}; r \geq 0\}$ forma um conjunto de valores mutuamente independentes. Excluindo-se o valor inicial $\hat{\theta}^{(0)}$, uma estimativa do parâmetro θ é dada por $[\hat{\theta}^{(1)} + \dots + \hat{\theta}^{(q)}]/q$, na qual q representa um valor suficientemente grande.

- 112 No algoritmo de Metropolis-Hastings tem-se a forma iterativa $f(\hat{\theta}^{(r+1)}) = f(\hat{\theta}^{(r)}) + \epsilon_r$, na qual f representa a função de densidade *a priori* de θ , e $\epsilon_r > 0$ representa um incremento aleatório. Nesse algoritmo, a probabilidade de aceitação do valor proposto $\hat{\theta}^{(r+1)}$ como uma estimativa viável para o parâmetro de interesse é constante.
- 113 O amostrador de Gibbs, um algoritmo sequencial de Monte Carlo, permite simular a distribuição *a priori* do parâmetro θ , desde que a forma funcional da sua função de densidade, $f(\theta)$, seja conhecida.
- 114 O método HPD (*high probability density*) é um algoritmo que proporciona um intervalo de confiança clássico (frequentista) para o parâmetro θ .

Espaço livre

Um estudo acerca do tempo (x , em anos) de guarda de autos findos em determinada seção judiciária considerou uma amostragem aleatória estratificada. A população consiste de uma listagem de autos findos, que foi segmentada em quatro estratos, segundo a classe de cada processo (as classes foram estabelecidas por resolução de autoridade judiciária). A tabela a seguir mostra os tamanhos populacionais (N) e amostrais (n), a média amostral (\bar{x}) e a variância amostral dos tempos (s^2) correspondentes a cada estrato.

estratos	tamanhos populacionais (N)	tamanhos amostrais (n)	\bar{x}	s^2
A	30.000	300	20	3
B	40.000	400	15	16
C	50.000	500	10	5
D	80.000	800	5	8
total	200.000	2.000	–	–

Considerando que o objetivo do estudo seja estimar o tempo médio populacional (em anos) de guarda dos autos findos, julgue os itens a seguir.

- 115 No estudo em questão foi aplicada uma amostragem aleatória estratificada com alocação proporcional ao tamanho dos estratos.
- 116 A fração amostral utilizada no estudo em tela foi igual ou superior a 10%.
- 117 A estimativa do tempo médio populacional da guarda dos autos findos é maior ou igual a 12 anos.
- 118 O erro padrão referente à estimação do tempo médio x no estrato D foi menor ou igual a 0,1.
- 119 No desenho amostral em tela há duas unidades amostrais: a primeira (unidade primária) corresponde à classe de cada processo, e a segunda (unidade secundária) refere-se a auto findo presente na listagem.
- 120 Combinando-se todos os estratos envolvidos, a estimativa da variância do tempo médio amostral da guarda dos autos findos é inferior a 0,005 ano².