

CONHECIMENTOS ESPECÍFICOS

RASCUNHO

Uma cadeia de Markov é denominada irredutível (ou ergódica) caso qualquer estado possa ser transformado em qualquer outro estado, não necessariamente em um único passo. Uma cadeia de Markov com matriz de transição P é regular caso exista um número inteiro positivo n tal que todos os elementos da matriz potência P^n sejam estritamente positivos.

Julgue os seguintes itens a respeito desses conceitos.

- 51 Algum elemento da matriz de transição P de uma cadeia de Markov regular pode ser zero.
- 52 Considere que, na matriz P mostrada abaixo, cada elemento $p_{i,j}$ represente a probabilidade de transição do estado i para o estado j .

$$P = \begin{pmatrix} 0,1 & 0,2 & 0,2 & 0,3 & 0,2 \\ 0,5 & 0 & 0,2 & 0 & 0,3 \\ 0 & 0 & 0,4 & 0 & 0,6 \\ 0,1 & 0,1 & 0,3 & 0,2 & 0,3 \\ 0 & 0 & 0,3 & 0 & 0,7 \end{pmatrix}$$

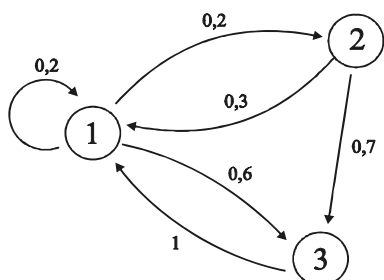
Em face dessas informações, é correto afirmar que a matriz P é a matriz de transição de uma cadeia de Markov irredutível.

- 53 Considere que, na matriz P mostrada a seguir, cada elemento $p_{i,j}$ represente a probabilidade de transição do estado i para o estado j .

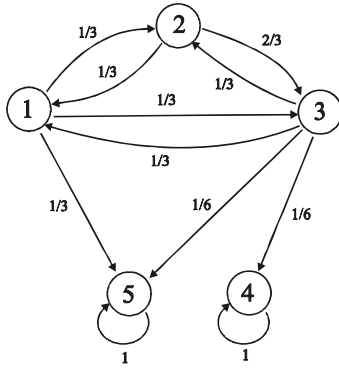
$$P = \begin{pmatrix} 0 & 0,3 & 0 & 0,3 & 0,4 \\ 0,3 & 0 & 0,7 & 0 & 0 \\ 0 & 0,4 & 0 & 0,3 & 0,3 \\ 0,2 & 0 & 0,8 & 0 & 0 \\ 0,5 & 0 & 0,5 & 0 & 0 \end{pmatrix}$$

A partir dessas informações, é correto afirmar que a matriz P é a matriz de transição de uma cadeia de Markov regular.

- 54 O dígrafo abaixo representa uma cadeia de Markov regular.



RASCUNHO



Considerando que uma cadeia de Markov seja representada pelo dígrafo ilustrado acima, julgue os itens a seguir.

- 55 Se o processo inicia-se no estado 1, 2 ou 3, então, a probabilidade de ser absorvido no estado 4 é $5/24$, $7/24$ ou $1/2$, respectivamente.
- 56 A cadeia de Markov representada pelo dígrafo acima é absorvente e a matriz de transição P , na forma canônica, tem a representação mostrada a seguir, em que cada elemento $p_{i,j}$ representa a probabilidade de transição do estado i para o estado j .

$$P = \begin{pmatrix} Q & R \\ 0 & I \end{pmatrix} = \begin{pmatrix} 0 & 1/3 & 1/3 & 0 & 1/3 \\ 1/3 & 0 & 2/3 & 0 & 0 \\ 1/3 & 1/3 & 0 & 1/6 & 1/6 \\ 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

- 57 Para a cadeia de Markov representada pelo dígrafo mostrado acima, a matriz fundamental é expressa por

$$N = \begin{pmatrix} 7/4 & 1 & 5/4 \\ 5/4 & 2 & 7/4 \\ 1 & 1 & 2 \end{pmatrix}.$$

- 58 Se o processo inicia-se no estado 1, 2 ou 3, então, o número esperado de passos até a absorção é 4, 5 ou 5, respectivamente.

Considere que uma variável aleatória bidimensional (X, Y) possua função de densidade conjunta $f(x, y) = 2$, se $0 \leq x \leq y \leq 1$; e $f(x, y) = 0$ para outros valores de x e y . Com base nessas informações, julgue os próximos itens.

- 59 O valor do coeficiente de correlação entre X e Y está entre 0 e $1/3$.
- 60 As variáveis X e Y são independentes.
- 61 As densidades marginais de X e Y são $g(x) = 2(1 - x)$, para $x \in [0, 1]$, e $h(y) = 2y$, para $y \in [0, 1]$, respectivamente.
- 62 A variável aleatória X tem valor esperado igual a $1/3$ e variância igual a $1/18$.
- 63 A média e a variância da variável aleatória Y são, respectivamente, iguais a $2/3$ e $1/16$.
- 64 A covariância entre X e Y é expressa por $\frac{1 - 4E(X)E(Y)}{4}$.

amostra <i>i</i>	1	2	3	4	5	6
X_i (diâmetro em 0,1 mm)	5	5	6	8	8	10
Y_i (resistência, em kg)	15	10	15	25	25	30

$\bar{x} = 7; \bar{y} = 20; \sum x_i y_i = 915; \sum x_i^2 = 314; \sum y_i^2 = 2.700$

A tabela e as estatísticas mostradas acima correspondem ao estudo realizado por um engenheiro acerca da resistência Y (em kg) à tração de 6 fios de determinado material, considerando-se os respectivos diâmetros X (em 0,1 mm).

Considerando essas informações e um modelo de regressão linear simples na forma $Y = a + bX + \epsilon$, em que ϵ representa o erro aleatório com média 0 e desvio padrão σ , julgue os itens que se seguem a respeito de regressão e correlação.

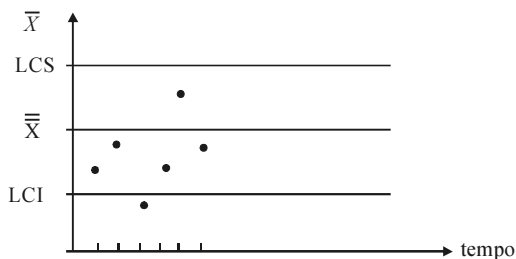
- 65 A soma total corrigida dos quadrados da variável Y é igual a 200.
- 66 A soma dos quadrados dos resíduos (variações não explicadas) é inferior a 20.
- 67 O coeficiente de determinação é superior a 90%.
- 68 O módulo do coeficiente de correlação entre X e Y é a raiz quadrada do coeficiente de determinação.
- 69 As estimativas de mínimos quadrados ordinários para os coeficientes do modelo de regressão linear simples são

$\hat{b} = 15/4$ e $\hat{a} = -25/4$.

Uma empresa produz leite em pó e o comercializa em latas. Se a máquina que enche as latas estiver em boas condições de funcionamento, a quantidade X de leite em pó em uma lata tem valor esperado de $\mu = 500$ g e desvio padrão de $\sigma = 10$ g. Para controlar o funcionamento dessa máquina, toma-se uma amostra de 9 latas a cada duas horas de funcionamento e seja calculada a quantidade líquida média \bar{X} da amostra.

Com base nessa situação e sabendo-se que $\Phi(3) = 0,9987$, em que Φ representa a função de distribuição acumulada da distribuição normal padronizada, julgue os itens de 70 a 73 a respeito do controle de qualidade 6-sigma.

- 70 Um gráfico de controle para a média do processo consiste em uma representação da evolução temporal da quantidade média amostral \bar{X} , conforme o exemplo mostrado na figura abaixo. Nessa figura, as três linhas horizontais representam, respectivamente, o limite superior de controle (LCS), o limite inferior (LCI) e a média dos valores \bar{X} . Tradicionalmente, a distância escolhida entre cada uma das linhas LCS e LCI e a linha do meio é três vezes o desvio padrão de \bar{X} . Quando é observado um valor \bar{X} fora dos limites, o processo é considerado fora do controle.



- 71 Na situação apresentada, os limites de controle LCS e LCI são 490 g e 510 g, respectivamente.
- 72 Considere que a máquina esteja em boas condições de funcionamento e que sejam observadas 10 médias amostrais \bar{X} . Nessa situação, a probabilidade de serem encontradas pelo menos 8 médias amostrais \bar{X} abaixo da linha central \bar{X} é superior a 0,06.
- 73 Caso a máquina esteja em boas condições de funcionamento, a probabilidade de que uma média amostral encontre-se dentro dos limites de controle será de 95%.

RASCUNHO

amostra	1	2	3	4	5	6	7	8	9	10
peças defeituosas (em %)	9	11	8	6	7	8	10	18	13	10

RASCUNHO

A tabela acima mostra o percentual de peças defeituosas encontradas em 10 amostras aleatórias simples independentes, cada uma de tamanho igual a 100. Com base nessas informações, julgue os itens a seguir, relativos ao controle de qualidade para atributos.

- 74 Durante o monitoramento das 10 amostras, o processo esteve sempre sob controle.
- 75 Os limites de controle do gráfico, referente a tabela acima, são iguais a 4% e 16%, respectivamente.

Julgue os seguintes itens a respeito de controle estatístico de qualidade.

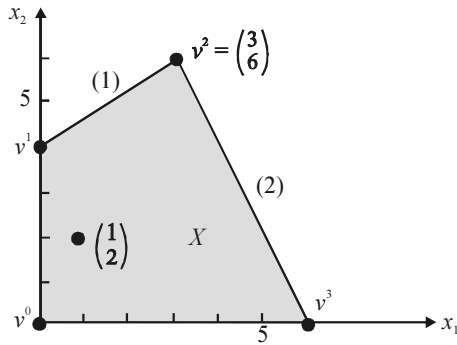
- 76 O gráfico de controle MMEP considera uma média ponderada de observações amostrais passadas e presentes. Os pesos crescem com a idade da observação, atribuindo-se maior peso às observações mais antigas que ocorreram no início do processo.
- 77 A principal desvantagem dos gráficos de controle básicos de Shewhart é a não sensibilidade a pequenas mudanças no processo. O gráfico de controle de somas cumulativas (CUSUM) e o gráfico de controle da média móvel exponencialmente ponderada (MMEP) são mais sensíveis a pequenas mudanças no processo.

Considere o seguinte problema de programação linear:

$$\begin{aligned} \max z &= x_1 + x_2 \\ \text{sujeito a} \\ x_1 &\leq 4; \quad (1) \\ x_2 &\leq 3; \quad (2) \\ x_1 + x_2 &\leq 5; \quad (3) \\ -x_1 + x_2 &\leq 3; \quad (4) \\ x_1, x_2 &\geq 0. \quad (5) \end{aligned}$$

Julgue os itens a seguir, a respeito desse problema.

- 78 O método simplex original de Dantzig resolve o problema em duas iterações, terminando necessariamente com o vértice ótimo $\begin{pmatrix} 4 \\ 1 \end{pmatrix}$.
- 79 Os vértices $\begin{pmatrix} 4 \\ 1 \end{pmatrix}$ e $\begin{pmatrix} 2 \\ 3 \end{pmatrix}$ do conjunto das soluções viáveis são soluções dualmente degeneradas, isto é, o custo reduzido de pelo menos uma variável não básica é zero.
- 80 A restrição (4) é fortemente redundante.
- 81 Existem infinitas soluções ótimas, isto é, o conjunto de soluções ótimas é o segmento de reta juntando os vértices $\begin{pmatrix} 2 \\ 3 \end{pmatrix}$ e $\begin{pmatrix} 4 \\ 1 \end{pmatrix}$.



A figura acima ilustra o conjunto X de soluções viáveis do seguinte problema de programação não-linear:

$$\begin{aligned} \max z &= (x_1 - 1)^2 + (x_2 - 2)^2 \\ \text{sujeito a} \\ -2x_1 + 3x_2 &\leq 12; & (1) \\ 2x_1 + x_2 &\leq 12; & (2) \\ x_1, x_2 &\geq 0. & (3) \end{aligned}$$

Com base nessas informações, julgue os próximos itens.

- 82 O valor máximo de z no problema apresentado é inferior a 25.
- 83 Os vértices v^0 e v^2 do conjunto X representam pontos de máximos locais (não globais) do problema.
- 84 O vértice $v^3 = \begin{pmatrix} 6 \\ 0 \end{pmatrix}$ é o único ponto de máximo global do problema.
- 85 Para todo ponto viável $x \neq \begin{pmatrix} 1 \\ 2 \end{pmatrix}$, o vetor $d = x - \begin{pmatrix} 1 \\ 2 \end{pmatrix}$ é uma direção de descida máxima.

RASCUNHO

Um analista deseja inspecionar um lote de 500 pacotes com encomendas internacionais. Como essa inspeção requer a abertura de cada pacote, ele decidiu fazê-la por amostragem, selecionando n pacotes desse lote. O analista dispõe de um cadastro que permite localizar precisamente cada pacote do lote por meio de um código de identificação.

Com base nessas informações e nos conceitos de amostragem, julgue os itens a seguir.

- 86 Para se calcular o tamanho da amostra com o objetivo de se estimar a proporção de pacotes que necessitam de recolhimento de impostos, independentemente do nível de confiança desejado, o analista deverá usar a fórmula $n = 1/E^2$, em que E é o erro amostral.
- 87 Se $n = 50$ pacotes selecionados aleatoriamente, então o fator de correção para populações finitas será superior a 0,85.
- 88 Considere que o lote de pacotes seja dividido em dois estratos segundo a massa de cada pacote: o primeiro, formado por 400 pacotes que possuem massas inferiores a 1 kg, e o segundo, por 100 pacotes com massas superiores a 1 kg. Nessa situação, se o analista efetuar uma amostragem estratificada de tamanho $n = 50$ com alocação uniforme, então essa amostra deverá contemplar 40 pacotes do primeiro estrato e 10 pacotes do segundo.
- 89 Na estimação de uma proporção p por amostragem aleatória simples, a variância do estimador é máxima quando $p = 0,5$.
- 90 Se o analista desejar fazer uma amostragem intencional (não probabilística) de tamanho $n = 10$, com base em sua experiência pessoal, então, nesse caso, a variância do estimador de uma proporção p será igual a $u(1 - u)/500$, em que u é uma probabilidade *a priori* estabelecida subjetivamente pelo analista.
- 91 A disponibilidade do cadastro permite que o analista efetue uma seleção por amostragem aleatória simples ou por amostragem sistemática com base nos códigos de identificação dos pacotes.
- 92 Se o analista optar pela amostragem sistemática, a seleção de uma amostra de tamanho $n = 50$ será efetuada de 10 em 10 pacotes, e o primeiro pacote a ser inspecionado será, necessariamente, o primeiro pacote registrado no cadastro.

tipo	capacidade (c), em kg	demanda (d)
1	$c \leq 1$	14
2	$1 < c \leq 2$	20
3	$2 < c \leq 3$	16
4	$3 < c \leq 4$	10
5	$4 < c \leq 5$	9
6	$5 < c \leq 6$	5
7	$6 < c \leq 7$	5
8	$7 < c \leq 8$	4
9	$8 < c \leq 9$	2
10	$9 < c \leq 10$	1

Considere que determinado fabricante classifique suas embalagens de acordo com a capacidade de armazenamento unitário (c), em kg, de determinado produto — por exemplo, a embalagem do tipo 10 permite armazenar mais de 9 e até 10 kg do produto. Com base nessas informações e na tabela acima, que mostra a distribuição dos dez tipos de embalagens e a demanda observada em março de 2011 para cada tipo, julgue os itens que se seguem.

- 93 O valor máximo da variável capacidade foi igual a 20.
- 94 Em março de 2011, a demanda total pelas embalagens foi superior a 100.
- 95 Pelos menos metade da demanda foi por embalagens com capacidades superiores a 2 kg.
- 96 É correto afirmar que o nível de mensuração da variável c foi intervalar.
- 97 A capacidade média das embalagens demandadas foi superior a 4 kg.
- 98 A capacidade mediana das embalagens demandadas foi maior que a média da capacidade das embalagens demandadas.
- 99 Um quarto da demanda foi por embalagens com capacidades para mais de 4 kg.
- 100 O gráfico de barras verticais é adequado para representar a distribuição da variável capacidade.
- 101 É possível inferir que a distribuição da capacidade segue uma distribuição de Poisson.

i	1	2	3	4	5	6	7	8	9	10
X	10	15	8	10	9	5	6	7	11	12
Y	80	90	60	85	80	50	60	65	110	100

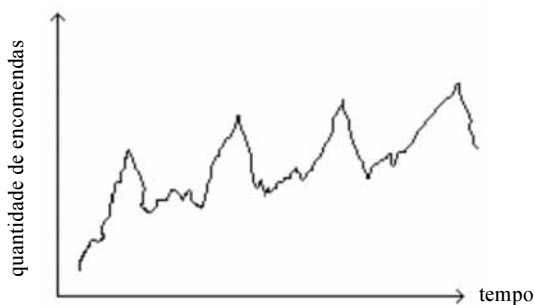
$$\sum x = 93, \sum x^2 = 945, \sum y = 780, \sum y^2 = 64.150$$

A fim de planejar o orçamento de uma grande empresa para o próximo ano, um analista selecionou uma amostra aleatória de 10 produtos (i) das empresas filiais e anotou as despesas (X) e os faturamentos (Y) totais decorrentes desses produtos (em R\$ milhões). Os resultados por ele obtidos são mostrados na tabela acima.

Com base nessas informações, julgue os itens subsequentes.

- 102 A variância de X foi inferior à metade da média da variável X.
- 103 Sabendo-se que a despesa X segue uma distribuição normal com média de R\$ 10 milhões e desvio padrão de R\$ 2 milhões, espera-se que $P(\bar{X} < 10) < P(\bar{X} \leq 10)$, em que \bar{X} representa a média amostral.
- 104 Considerando-se $X \sim N(\mu, 4)$, em que μ representa a média populacional da variável X, ao se testar a hipótese nula $H_0: \mu = \text{R\$ } 10$ milhões contra a hipótese alternativa $H_1: \mu < \text{R\$ } 10$ milhões, é correto afirmar que o valor da estatística do teste z foi negativo.
- 105 Considere o teste de hipóteses $H_0: \mu = \text{R\$ } 10$ milhões *versus* $H_1: \mu < \text{R\$ } 10$ milhões, em que μ representa a média populacional da variável X, e suponha que X segue uma distribuição normal com desvio padrão igual a R\$ 2 milhões. Com base nessas informações, considerando-se o nível de significância igual a 5%, é correto afirmar que a hipótese nula não seria rejeitada.
- 106 A média das despesas X por produto foi igual a R\$ 9,3 milhões.
- 107 A estimativa do faturamento médio por produto foi superior a 8 vezes o valor da despesa média por produto.
- 108 Considere um teste de hipóteses acerca da média da variável X. Nesse caso, se todos os demais momentos da distribuição X forem desconhecidos, então a estatística apropriada para esse teste segue uma distribuição t com 9 graus de liberdade.

RASCUNHO



RASCUNHO

O gráfico acima mostra a evolução temporal da quantidade mensal de encomendas X entregues em determinada cidade. A partir dessa figura e dos conceitos de séries temporais, julgue os itens subsequentes.

- 109 Uma das suposições do modelo $AR(p)$ é que os erros aleatórios sejam ruído branco.
- 110 A série apresenta sazonalidade e tendência.
- 111 Um modelo $AR(1)$ é apropriado para representar a série temporal X.
- 112 A metodologia de Box e Jenkins se aplica somente a séries temporais estacionárias.
- 113 Se os picos da série temporal X ocorrem nos meses de dezembro, então o período sazonal a ser considerado em um modelo SARIMA é igual a 12.

Para criar um *ranking* das universidades brasileiras, um pesquisador dispõe das seguintes variáveis: X_1 = número de professores doutores; X_2 = quantidade de pesquisas publicadas em periódicos nacionais; X_3 = quantidade de pesquisas publicadas em periódicos internacionais; X_4 = área total do *campus*; X_5 = quantidade de cursos de pós-graduação.

Considerando essas informações e os conceitos de análise multivariada, julgue os itens seguintes.

- 114 Considere que, na universidade A, $X_1 = 300$ e $X_2 = 50$ e que, na universidade B, $X_1 = 400$ e $X_2 = 60$. Nesse caso, a distância euclidiana entre essas universidades é inferior a 80.
- 115 O gráfico que representa o processo de agrupamento hierárquico é denominado dendograma.
- 116 Se o pesquisador utilizar a técnica de análise de componentes principais, serão geradas 5 componentes.
- 117 A primeira componente principal representa uma média das 5 variáveis.
- 118 Cada componente principal é independente da outra, podendo a variância explicada pela segunda componente ser inferior à variância explicada pela terceira componente.
- 119 Na análise de conglomerados (*clusters*), a similaridade entre as observações pode ser medida com base na distância euclidiana ou na distância de Mahalanobis.
- 120 O objetivo principal da análise de conglomerados é maximizar a homogeneidade das observações dentro do mesmo grupo, além de maximizar a heterogeneidade entre os grupos.