

## CONCURSO PÚBLICO EDITAL Nº 4 / 2021

### CARGO

## ESTATÍSTICO

### INSTRUÇÕES AOS CANDIDATOS

- \* Verifique se este caderno contém 25 questões. Caso não contenha, solicite imediatamente ao fiscal de sala outro caderno.
- \* Você dispõe de 2 horas para responder a todas as questões.
- \* Para cada questão existe apenas uma resposta certa.
- \* Poderá utilizar a grade ao final do caderno para marcar previamente as respostas.
- \* Transcreva as respostas para o cartão resposta, preenchendo totalmente o círculo com caneta esferográfica com tinta preta ou azul escuro, não sendo permitido o uso de caneta porosa ou corretivo líquido.
- \* O telefone celular desligado, controle remoto e relógio devem estar dentro do envelope devidamente identificado e lacrado.
- \* Para se dirigir ao fiscal, erga o braço e aguarde o atendimento.
- \* Não é permitido o uso de qualquer tipo de aparelho eletrônico dentro do prédio de provas, mesmo após a entrega da prova.
- \* Durante a realização da prova não será permitido o uso de livros, manuais, impressos, anotações, máquinas calculadoras, agendas eletrônicas ou similares, telefone celular de qualquer tipo, BIP, MP3 *player* ou similar, gravador ou qualquer outro receptor de dados ou mensagens, qualquer tipo de controle remoto, protetor auricular, fones de ouvido, prótese auditiva, óculos com lentes escuras, relógio ou qualquer acessório na cabeça.
- \* É proibido fumar no interior do prédio de provas.
- \* O cartão resposta, se danificado pelo candidato não será substituído.
- \* A entrega da prova só poderá ocorrer depois de transcorrida uma hora do horário de início.
- \* Ao terminar a prova, deverá ser entregue, obrigatoriamente, ao fiscal de sala, seu cartão resposta devidamente assinado, podendo levar consigo o caderno de questões.
- \* Após a entrega da prova, o candidato deverá retirar-se imediatamente do prédio de aplicação da prova, não sendo permitido, nesse local, o uso dos sanitários.
- \* Será excluído do concurso o candidato que agir com incorreção ou descortesia com qualquer pessoa da equipe encarregada da aplicação das provas ou candidato participante do processo.
- \* Os dois últimos candidatos que permanecerem em sala de prova, só poderão retirar-se conjuntamente e após sua assinatura na ata de presença.



1. A FURG deseja selecionar uma amostra de estudantes da graduação do campus Carreiros para avaliar os espaços de lazer e a convivência na Instituição, com o intuito de planejar melhorias na infraestrutura. Para isso, fez-se um sorteio dos cursos que participarão da pesquisa, e todos os estudantes dos cursos selecionados devem participar da pesquisa. Diante desse contexto, qual técnica de amostragem deverá ser utilizada?

- a) Amostragem Aleatória Simples.
  - b) Amostragem Aleatória Estratificada.
  - c) Amostragem por Conglomerados.
  - d) Amostragem Sistemática.
  - e) Amostragem por Conveniência.
- 

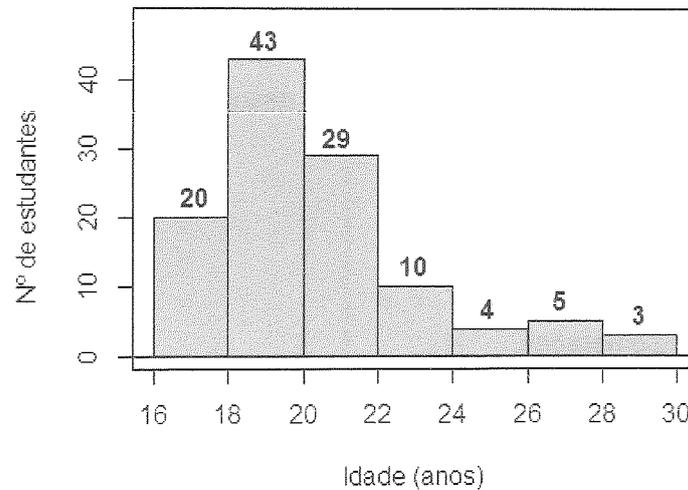
2. Com relação às técnicas de amostragem, analise as seguintes afirmativas.

- I. Na amostragem por conglomerado, temos a alocação proporcional e a alocação uniforme das observações.
- II. A Amostragem sistemática é um tipo de amostragem não probabilística e indicada quando a população está ordenada de tal forma, que cada elemento seja identificado unicamente pela sua posição.
- III. Indicamos a amostragem estratificada quando a população se divide em grupos, sendo que a variável de interesse seja homogênea dentro desses grupos e heterogênea entre eles.
- IV. A amostragem aleatória simples é indicada quando a população é heterogênea em relação à variável de interesse.

Está correto o que se afirma em:

- a) Apenas IV
- b) Apenas I e III
- c) Apenas II e III
- d) Apenas III
- e) Apenas I, III e IV

3. O histograma apresenta a distribuição de frequência da idade de estudantes que ingressaram no curso de Engenharia em 2021. Para sua construção, foi utilizado o intervalo (a, b].



Avalie as afirmações com base no histograma.

- I. O valor da mediana para a variável idade é 21 anos.
- II. A probabilidade de um estudante selecionado ao acaso ter mais de 24 anos é  $2/19$ .
- III. O coeficiente de assimetria de Pearson será positivo.
- IV. O valor modal está na 2ª classe.

Está correto o que se afirma em:

- a) Apenas II, III e IV
- b) Apenas I e III
- c) Apenas II e IV
- d) Apenas I, II e III
- e) Todas as afirmações

4. O Exame Nacional de Desempenho do Estudante (ENADE) é composto por questões de conhecimentos gerais (CG) e de conhecimentos específicos (CE) do curso de graduação no qual o estudante está matriculado. Considere o desempenho de um estudante do curso de Matemática Aplicada que participou do último ENADE e obteve 70 pontos na prova de CG e 70 pontos na prova de CE. Sabe-se que a média nacional para o curso na prova CG é igual a 54, com desvio padrão igual a 21, e para a prova CE obteve-se média nacional igual a 47, com desvio padrão igual a 22. O que é possível concluir sobre o desempenho desse estudante da Matemática Aplicada?

- a) O desempenho relativo do estudante na prova CG foi inferior ao da prova CE.
- b) O desempenho do estudante na prova CG apresentou maior variabilidade que na prova CE.
- c) O coeficiente de variação das notas da prova de CG foi maior do que na prova CE.
- d) O desempenho relativo do estudante foi o mesmo nas duas provas.
- e) O desempenho do estudante na prova CG está a mais de um desvio padrão da média nacional.

5. Sendo  $P(A)=x$ ,  $P(B)=y$  e  $P(A \cap B)=z$ , então:

I)  $P(\bar{A} \cup \bar{B}) = 1 - z$

II)  $P(\bar{A} \cup B) = 1 - x + z$

III)  $P(\bar{A} \cap \bar{B}) = x + y - z$

As afirmações corretas são:

- a) Apenas I
- b) Apenas III
- c) Todas as afirmações
- d) Apenas I e II
- e) Nenhuma das afirmações

6. Dados do último anuário da FURG mostram que a instituição conta, atualmente, com um total de 09 Casas do Estudante Universitário – CEUs em funcionamento, para atendimento aos estudantes de graduação e de pós-graduação em situação de vulnerabilidade socioeconômica. Ao todo, são oferecidas 457 vagas de moradia estudantil, sendo que 420 vagas são da FURG – Rio Grande e 37 são nos demais campi da FURG. Sabe-se que 320 estudantes das CEUs Rio Grande recebem algum tipo de bolsa acadêmica, já dos outros campi, 15 recebem bolsa acadêmica. Ao selecionar um estudante aleatoriamente que vive numa casa de estudante da FURG, qual é a probabilidade de este ser de outro campus e não receber bolsa acadêmica?

- a)  $22/37$
- b)  $37/457$
- c)  $22/457$
- d)  $15/457$
- e)  $15/37$

7. Analise as afirmativas abaixo:

- I. A rejeição da hipótese nula em um teste indica que o p-valor do teste é idêntico ao poder do teste.
- II. O erro do tipo II consiste em não rejeitar a hipótese nula quando ela é falsa.
- III. Em testes de hipóteses estatísticos, o resultado é estatisticamente significativo quando o p-valor observado é menor que o nível de significância definido para o estudo.
- IV. O p-valor representa a probabilidade de rejeitar a hipótese nula, quando ela é verdadeira.

Estão corretas:

- a) Apenas I, II e III
  - b) Apenas II e III
  - c) Apenas I e IV
  - d) Apenas II e III
  - e) Todas as alternativas
- 

8. Analise as afirmativas sobre inferência estatística.

- I. É característica de um teste paramétrico só poder ser aplicado se a variável for medida em escala de intervalos ou em escala razão.
- II. A estatística de Mann-Whitney para teste de hipótese entre duas amostras independentes é aplicável somente para análise de variáveis quantitativas.
- III. O teste não paramétrico alternativo para o teste ANOVA two-way é o Teste Friedman.
- IV. Um teste que compara a distribuição real dos dados (amostra) com uma distribuição normal gerada por uma média e um desvio padrão supostamente conhecidos (populacionais) é o Teste Kolmogorov-Smirnov.

Estão corretas as alternativas:

- a) Apenas I e IV
  - b) Apenas III e IV
  - c) Apenas II e III
  - d) Apenas I e III
  - e) Apenas I, III e IV
-

9. Problemas relacionados à saúde mental de estudantes gera preocupação e têm se agravado por conta da pandemia de COVID-19. Pensando nisso, um pesquisador deseja avaliar os níveis de ansiedade de estudantes de um determinado curso por meio de um teste validado e desenvolvido para esse fim. Os resultados indicam que, quanto maior é o escore no teste, maior é o nível de ansiedade atual do(a) estudante, que pode variar de 0 a 63. O pesquisador deseja verificar se, em média, os escores de ansiedade das mulheres difere dos escores dos homens. Assumindo a normalidade dos dados, veja, a seguir, os resultados da análise aplicada.

F test to compare two variances

data: escore by Sexo

F = 1.9795, num df = 52, denom df = 66, p-value = 0.008963

alternative hypothesis: true ratio of variances is not equal to 1

95 percent confidence interval:

1.187507 3.353569

sample estimates:

ratio of variances

1.979538

Welch Two Sample t-test

data: escore by Sexo

t = 4.6015, df = 90.48, p-value = 0.00001361

alternative hypothesis: true difference in means is not equal to 0

95 percent confidence interval:

4.036285 10.168727

sample estimates:

mean in group F mean in group M

17.49057 10.38806

Que procedimento foi adotado para saber se há diferença entre o sexo dos estudantes, quanto ao escore de ansiedade?

- a) Igualdade das médias para duas amostras independentes, considerando as variâncias iguais e usando a distribuição t-Student para a estatística do teste.
- b) Igualdade das médias para duas amostras independentes, considerando as variâncias diferentes e usando a distribuição Z para a estatística do teste.
- c) Igualdade das médias para duas amostras dependentes e usando a distribuição t-Student para a estatística do teste.
- d) Os dados são insuficientes para conclusão do teste de hipóteses.
- e) Igualdade das médias para duas amostras independentes, considerando as variâncias diferentes e usando a distribuição t-Student para a estatística do teste.

10. Uma pesquisa realizada com 70 discentes de um curso de especialização EAD da FURG teve por objetivo analisar o perfil dos que buscam essa modalidade de ensino. Dentre as perguntas de interesse, os pesquisadores avaliaram a relação entre o tempo disponível para estudo e com quem os discentes preferem tirar suas dúvidas ao longo do curso. Os resultados foram:

Tempo para estudo	Com quem tira dúvidas			Total
	Colegas	Professores	Tutores	
Apenas nos finais de semana	19	7	13	39
Vários dias da semana	7	14	10	31
Total	26	21	23	70

Considerando que o valor calculado de 7,44 e escolhendo entre os seguintes valores críticos da distribuição qui-quadrado, ao nível de 5% de significância,  $X^2_{(2g.l.)} = 5,99$  ou  $X^2_{(3g.l.)} = 7,81$  ou  $X^2_{(6g.l.)} = 12,59$ . Conclua sobre a relação entre o tempo de estudo e com quem prefere tirar suas dúvidas.

- A proporção de discentes que preferem tirar dúvidas com o professor é significativamente maior independente do tempo de estudo.
- Não há evidências amostrais para concluir quanto à existência de relação significativa entre tempo de estudo e com quem os discentes preferem tirar suas dúvidas.
- Existe relação significativa entre tempo de estudo e com quem os discentes preferem tirar suas dúvidas.
- Não existe relação significativa entre tempo de estudo e com quem os discentes preferem tirar suas dúvidas.
- Não é possível concluir algo com os dados apresentados.

11. Considerando a estimação de um intervalo de confiança para a média de uma população com variância desconhecida. Ao aumentar 4 vezes o tamanho da amostra coletada, assumindo os demais parâmetros inalterados, a amplitude do intervalo, em relação ao tamanho de amostra original, é:

- Multiplicada por 2
- Dividida por 2
- Dividida por 4
- Multiplicada por 16
- Dividida por 16

12. O parâmetro populacional  $\theta$  tem distribuição normal com média e variância igual a  $\theta$  e 9, respectivamente. Uma amostra de tamanho  $n = 16$  é extraída dessa população, obtendo-se  $\hat{\theta} = 10$ , onde  $\hat{\theta}$  é o estimador de  $\theta$ .

Sendo  $\phi(z)$  a função de distribuição acumulada da normal-padrão, considere:

$$\phi(1,5) \cong 0,95 \text{ e } \phi(2) \cong 0,975$$

Sob essas condições, o intervalo de confiança para  $\theta$  a partir da amostra coletada, com 95% de confiança, será:

- a) [8,5; 11,5]
- b) [9; 11]
- c) [9,5; 10,5]
- d) [10; 11]
- e) [8; 12]

13. Realizando uma regressão linear múltipla no software R, obtivemos o seguinte retorno:

Call: lm(formula = Y ~ + X1+ X2 + X3 + X4 + X5)				
Residuals:				
Min	1Q	Median	3Q	Max
-0.046524	-0.013137	-0.006239	0.013234	0.058643
Coefficients				
	Estimate	Std. Error	t value	Pr(> t )
(Intercept)	0.0318259	0.0330181	0.964	0.34471
X1	0.1396939	0.0548606	2.546	0.01772 *
X2	-0.0008814	0.0003267	-2.698	0.01258 *
X3	0.0001583	0.0002211	0.716	0.48086
X4	-0.0660166	0.0241977	-2.728	0.01172 *
X5	-0.0811859	0.0263812	-3.077	0.00516 **
Signif. codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1				
Residual standard error: 0.02513 on 24 degrees of freedom				
Multiple R-squared: 0.5526, Adjusted R-squared: 0.4594				
F-statistic: 5.93 on 5 and 24 DF, p-value: 0.001049				

Sob essas condições, e considerando nível de significância de 5%, quais variáveis da regressão possuem coeficientes estatisticamente diferentes de zero?

- a) X1, X2 e X4
- b) X5
- c) Intercepto e X3
- d) X1, X2, X4 e X5
- e) Todas as variáveis

14. Se realizada uma Análise de Componentes Principais calculada por intermédio da decomposição espectral de uma matriz de covariância, é correto afirmar que:

- a) Os autovalores são os componentes principais, enquanto seus autovetores correspondentes são uma indicação direta da explicação da variância do respectivo autovalor.
- b) Caso a Análise de Componentes principais fosse calculada utilizando a matriz de correlação os resultados seriam iguais.
- c) Não é possível calcular os autovetores e autovalores, uma vez que a Análise de Componentes Principais pode ser realizada apenas por meio de uma matriz de correlação.
- d) Os valores da matriz de autovalores resultante sempre estão em ordem crescente.
- e) Os autovetores são os componentes principais, enquanto seus autovalores correspondentes são uma indicação direta da explicação da variância do respectivo autovetor.

15. Considere a Variável Aleatória Contínua  $X$ , com função densidade de probabilidade dada por  $f(X) = x + 2$ , e definida no intervalo  $0 < X < 3$ . Qual a esperança dessa variável aleatória contínua?

- a) 27
- b) 18
- c) 9
- d) 81
- e) 6

16. Disciplinas com maior taxa de reprovação despendem uma atenção especial das Universidades Federais. Para mensurar a quantidade de turmas ofertadas para essas disciplinas, é de interesse das Universidades modelar a quantidade de vezes que cada aluno precisa se matricular na disciplina até sua aprovação. Considerando que as chances de aprovação são iguais para todos os alunos, qual é o modelo probabilístico adequado a essa situação?

- a) Binomial
- b) Binomial Hipergeométrica
- c) Geométrica
- d) Poisson
- e) Hipergeométrica

17. Em um hipotético estudo, deseja-se associar determinada característica com uma certa doença. Por intermédio de um estudo caso-controle, obteve-se os seguintes resultados:

	Com Doença	Sem Doença
Com Característica	48	36
Sem Característica	24	90

Sob essas condições, podemos afirmar que:

- a) Possuir a característica diminui a chance de ter a doença em 1,25 vezes, quando comparado a não possuir a característica.
- b) Possuir a característica aumenta a chance de ter a doença em 1,05 vezes, quando comparado a não possuir a característica.
- c) Possuir a característica aumenta a chance de ter a doença em 5 vezes, quando comparado a não possuir a característica.
- d) Possuir a característica aumenta a chance de ter a doença em 2 vezes, quando comparado a não possuir a característica.
- e) Possuir a característica diminui a chance de ter a doença em 2 vezes, quando comparado a não possuir a característica.

18. Uma nova pesquisa deseja avaliar a possível associação entre fumar e desenvolver câncer de pulmão. Para tanto, os pesquisadores selecionaram 1000 casos aleatórios de câncer pulmonar registrados em hospitais da região e determinado se tais pacientes fumavam ou não. Foram também selecionados outros 1000 pacientes desses mesmos hospitais, porém que estavam internados por outras doenças que não câncer. O desenho desse estudo foi:

- a) Caso-Controle
- b) Coorte
- c) Transversal
- d) Experimental
- e) Nenhuma das anteriores

19. Seja o modelo autorregressivo AR(p), com ordem  $p=1$  dado por  $\tilde{Z}_t = \phi\tilde{Z}_{t-1} + a_t$  de maneira que  $\tilde{Z}_t$  depende apenas de  $\tilde{Z}_{t-1}$  e do ruído no instante t, é correto afirmar que

- a) O processo é estacionário se  $\psi(B)$  convergir para  $|B| \leq 1$  dado  $\psi(B) = \sum_{j=0}^{\infty} \phi^j B^j = [\phi(B)]^{-1} = (1 - \phi B)^{-1}$
- b) processo é estacionário se  $\psi(B)$  convergir para  $|B| > 1$  dado  $\psi(B) = \sum_{j=0}^{\infty} \phi^j B^j = [\phi(B)]^{-1} = (1 - \phi B)^{-1}$
- c) O processo é sempre estacionário, pois  $\sum_{j=0}^{\infty} \phi^j B^j = [\phi(B)]^{-1} = (1 - \phi B)^{-1}$
- d) O processo é nunca invertível, pois  $\pi(B) = \phi(B) = 1 - \phi B$
- e) O processo é estacionário se a raiz  $B = \phi^{(-1)}$  da equação  $\pi(B) = \phi(B) = 1 - \phi B$ , cair dentro do círculo unitário.

20. Seja um processo ARMA(p,q) com função de autocorrelação definida por  $\rho_k \neq 0$  e função de autocorrelação parcial definida por  $\phi_{kk} \neq 0$  com  $k = 1, 2, 3, \dots$

- I.  $\rho_k \neq 0$  para  $k=1, 2$  e  $3$  e  $\rho_k = 0$  para todos os outros valores de  $k$ .
- II.  $\phi_{kk}$  é dado por exponenciais e/ou senoides amortecidas

Com base nas afirmativas (I) e (II), determine a ordem p e q desse processo

- a) ARMA(3,0)
- b) ARMA(2,1)
- c) ARMA(1,2)
- d) ARMA(0,3)
- e) ARMA(2,0)

21. Seja  $\pi(\theta) \sim \text{Beta}(\alpha, \beta)$   $\alpha, \beta > 0$  uma distribuição à priori e  $f(x|\theta) = \text{Bin}(n, \theta)$  uma função de verossimilhança,

Onde

$$\text{Beta}(\alpha, \beta) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} \theta^{\alpha-1} (1-\theta)^{\beta-1} \text{ e } \text{Bin}\left(\begin{matrix} n \\ x \end{matrix}\right) \theta^x (1-\theta)^{n-x}$$

portanto, visto que as distribuições Beta e Binomial pertencem à mesma família, ou seja, são conjugadas, a distribuição posterior que se procura é dada por

- a)  $\pi(x|\theta) \sim \text{Beta}(x + \alpha, n - x + \beta)$
- b)  $\pi(x|\theta) \sim \text{Beta}(x + \alpha - 1, n - x + \beta)$
- c)  $\pi(x|\theta) \sim \text{Bin}(x + \alpha, n - x + \beta)$
- d)  $\pi(\theta|x) \sim \text{Bin}(x, n - x)$
- e)  $\pi(\theta|x) \sim \text{Beta}(x + \alpha, n - x + \beta)$

22. O intervalo de credibilidade delimita, com probabilidade especificada, um conjunto de parâmetros plausíveis para o parâmetro de acordo com sua distribuição posterior marginal. É correto afirmar que:

- a) O intervalo de credibilidade percentil 95% para  $\theta$  simbolizado por  $IC_{r_{95\%}}$  é o intervalo delimitado pelos percentis 0% e 95% da distribuição posterior marginal  $p(\theta|x)$  para  $\theta$ .
- b) O intervalo de credibilidade percentil 95% para  $\theta$  simbolizado por  $IC_{r_{95\%}}$  inclui os valores mais extremos em cada cauda da distribuição posterior marginal.
- c) O intervalo de credibilidade percentil 95% para  $\theta$  simbolizado por  $IC_{r_{95\%}}$  é o intervalo delimitado pelos percentis 2,5% e 97,5% da distribuição posterior marginal  $p(\theta|x)$  para  $\theta$ .
- d) O intervalo de maiores densidades de 95% para  $\theta$  simbolizado por  $HDI_{r_{95\%}}$  inclui os valores que representam as 5% menores densidades na distribuição posterior.
- e) O intervalo de maiores densidades de 95% para  $\theta$  simbolizado por  $HDI_{r_{95\%}}$  exclui os valores que representam as 5% menores densidades na distribuição posterior de forma igualitária, independente da simetria da distribuição posterior.

23. Seja  $T$  um estimador do parâmetro  $\theta$ , então:
- I. O estimador  $T$  é não viesado para  $\theta$  se  $E(X) = \theta, \forall \theta$
  - II. Estimativa é o valor assumido pelo estimador em uma particular amostra
  - III. Uma sequência  $\{T_n\}$  de estimadores de um parâmetro  $\theta$  é consistente se para  $\epsilon > 0$ ,  
 $P|T_n - \theta| > \epsilon \rightarrow 0, n \rightarrow \infty$
  - IV. Se  $T$  e  $T'$  são dois estimadores não viesados de um mesmo parâmetro  $\theta$  e ainda  $Var(T) < Var(T')$ , então  $T$  diz-se mais eficiente do que  $T'$

É correto afirmar que:

- a) I e III são corretas
- b) I, II e IV são corretas
- c) II e III são incorretas
- d) I, II e III são incorretas
- e) I, II, III e IV são corretas

24. Suponha  $n$  provas de Bernoulli com  $P(\text{sucesso}) = p, 0 < p < 1$  e  $X$  o número de sucessos. Sendo  $n = 6$  e obtendo quatro sucessos e 2 fracassos, a função de verossimilhança é dada por  $L(p) = p^4(1 - p)^2$ , assim o estimador de máxima verossimilhança do parâmetro  $p$  da distribuição binomial é dado por:

- a)  $\hat{p} = 2/6$
- b)  $\hat{p} = 4/6$
- c)  $\hat{p} = 2/4$
- d)  $\hat{p} = 1/3$
- e)  $\hat{p} = 3/6$

25. De acordo com a definição de Matriz de Markov, também conhecida como matriz estocástica ou de transição, observe as matrizes abaixo e assinale a alternativa correta:

I.  $\begin{bmatrix} 1/4 & 2/3 \\ 3/4 & 1/3 \end{bmatrix}$

II.  $\begin{bmatrix} 0 & 1/2 & 1/4 \\ 1/2 & 1/2 & 1/2 \\ 1/2 & 0 & 1/2 \end{bmatrix}$

III.  $\begin{bmatrix} 1/2 & 1/4 & 1 \\ 1/2 & 1/2 & 0 \\ 1/2 & 0 & 0 \end{bmatrix}$

IV.  $\begin{bmatrix} 1/4 & -2/3 \\ 3/4 & -1/3 \end{bmatrix}$

- a) As matrizes (I) e (II) são matrizes estocásticas
- b) Apenas a Matriz (I) é estocástica
- c) As matrizes (II) e (III) são estocásticas
- d) As matrizes (I) e (IV) são estocásticas
- e) Apenas a matriz (IV) é estocástica

	A	B	C	D	E
1					
2					
3					
4					
5					
6					
7					
8					
9					
10					
11					
12					
13					
14					
15					
16					
17					
18					
19					
20					
21					
22					
23					
24					
25					