

# EMPRESA BRASILEIRA DE PESQUISA AGROPECUÁRIA (EMBRAPA)

Processo Seletivo Público

Nome do candidato:

Número do documento de identidade:

Número de inscrição:

Sala:

Seqüencial:

## PESQUISADOR III

## Área

## Subárea



**Embrapa**

## TARDE

## PROVAS OBJETIVAS

Aplicação: 2/4/2006

**LEIA COM ATENÇÃO AS INSTRUÇÕES ABAIXO.**

- 1 Ao receber este caderno, confira inicialmente os seus dados pessoais transcritos acima. Em seguida, verifique se ele contém cem itens, correspondentes às provas objetivas, corretamente ordenados de 1 a 100.
- 2 Caso os dados pessoais constantes neste caderno não correspondam aos seus, ou, ainda, caso o caderno esteja incompleto ou tenha qualquer defeito, solicite ao fiscal de sala mais próximo que tome as providências cabíveis.
- 3 Nos itens das provas objetivas, recomenda-se não marcar ao acaso: em cada item, se a resposta divergir do gabarito oficial definitivo, o candidato receberá pontuação negativa, conforme consta em edital.
- 4 Não utilize lápis, lapiseira, borracha e(ou) qualquer material de consulta que não seja fornecido pelo CESPE.
- 5 Não se comunique com outros candidatos nem se levante sem autorização do chefe de sala.
- 6 A duração das provas — objetivas e discursiva — é de **quatro horas e trinta minutos**, já incluído o tempo destinado à identificação — que será feita no decorrer das provas —, ao preenchimento da folha de respostas e à transcrição dos textos definitivos para o caderno de textos definitivos de prova discursiva.
- 7 Ao terminar as provas, chame o fiscal de sala mais próximo, devolva-lhe a sua folha de respostas e o caderno de textos definitivos da prova discursiva e deixe o local de provas.
- 8 A desobediência a qualquer uma das determinações constantes no presente caderno, na folha de rascunho, na folha de respostas ou no caderno de textos definitivos da prova discursiva poderá implicar a anulação das suas provas.

#### AGENDA

- I **4/4/2006**, após as 19 h (horário de Brasília) – Gabaritos oficiais preliminares das provas objetivas: Internet — [www.cespe.unb.br/concursos/embrapa2006](http://www.cespe.unb.br/concursos/embrapa2006).
- II **5 e 6/4/2006** – Recursos (provas objetivas): exclusivamente no Sistema Eletrônico de Interposição de Recurso, Internet — [www.cespe.unb.br/concursos/embrapa2006](http://www.cespe.unb.br/concursos/embrapa2006) —, mediante instruções e formulários que estarão disponíveis nesse endereço.
- III **3/5/2006** – Resultado final das provas objetivas e resultado provisório da prova discursiva: Diário Oficial da União e Internet — [www.cespe.unb.br/concursos/embrapa2006](http://www.cespe.unb.br/concursos/embrapa2006).
- IV **4 e 5/5/2006** – Recursos (prova discursiva): exclusivamente no Sistema Eletrônico de Interposição de Recurso, Internet — [www.cespe.unb.br/concursos/embrapa2006](http://www.cespe.unb.br/concursos/embrapa2006) —, mediante instruções e formulários que estarão disponíveis nesse endereço.
- V **17/5/2006** – Resultado final da prova discursiva e convocação para a entrega de documentos para a avaliação de títulos: locais mencionados no item III.

#### OBSERVAÇÕES

- Não serão objeto de conhecimento recursos em desacordo com o item 13 do Edital n.º 1/2006 – EMBRAPA, de 31/1/2006.
- Informações adicionais: telefone 0(XX) 61 3448-0100; Internet — [www.cespe.unb.br/concursos/embrapa2006](http://www.cespe.unb.br/concursos/embrapa2006).
- É permitida a reprodução deste material apenas para fins didáticos, desde que citada a fonte.

**CESPEUnB**  
UNIVERSIDADE DE BRASÍLIA

- De acordo com o comando a que cada um dos itens de **1 a 100** se refira, marque, na **folha de respostas**, para cada item: o campo designado com o código **C**, caso julgue o item **CERTO**; ou o campo designado com o código **E**, caso julgue o item **ERRADO**. A ausência de marcação ou a marcação de ambos os campos não serão apenadas, ou seja, não receberão pontuação negativa. Para as devidas marcações, use a folha de rascunho e, posteriormente, a **folha de respostas**, único documento válido para a correção das suas provas.

## CONHECIMENTOS BÁSICOS

1 A cada dia que passa, mais me convenço de que  
vivemos num mundo inventado, e não apenas no que se  
refere às tecnologias e valores que sustentam a vida, mas  
4 no que diz respeito à história ou histórias que constituem  
nosso passado.

Ferreira Gullar. *Folha de S. Paulo*. 29/1/2006 (com adaptações).

Julgue os seguintes itens a respeito do fragmento de texto acima.

- 1 Dado que o termo “A cada dia” (l.1) mantém correspondência semântica com a expressão **Todos os dias**, esta pode substituí-la, sem prejudicar a coerência e a correção gramatical do texto.
- 2 A presença da preposição “de” (l.1) antes da conjunção “que” indica que esta introduz uma oração, de função substantiva, que complementa o verbo **convencer**.
- 3 Na linha 2, a repetição do termo “inventado” depois de “e” produziria redundância, que funcionaria como ênfase para a idéia expressa, sem prejudicar a coerência argumentativa ou a correção gramatical do texto.

1 Nenhum autor sintetiza melhor do que Pico della  
Mirandola a imagem que os pensadores italianos tinham  
do homem. Ele nos fala de maneira eloqüente desse ser  
4 cheio de potencialidades que, ao contrário dos outros  
animais, recebeu de Deus o direito de forjar o próprio  
rosto.

7 Um homem assim constituído certamente não  
pode compreender sua passagem pelo mundo da mesma  
forma que os homens a compreendiam em plena Idade  
10 Média. Pico sugere não só que não somos prisioneiros de  
nenhuma lei externa, mas que podemos escolher nossa  
própria natureza, moldando-a segundo nossos desejos e  
13 criando nossas próprias leis. O que essa imagem de um  
homem totalmente livre tem de encantadora não é a  
descoberta do caráter não necessário das escolhas que  
16 operamos no mundo, os medievais já sabiam disso, e sim  
o fato de que o pensador italiano sugere que podemos  
inventar uma natureza e também sua forma de estar no  
19 mundo. Em outras palavras, somos os inventores de nossa  
natureza e também de nossa história.

Newton Bignotto, *O círculo e a linha*. In: Adauto  
Novaes (org.), *Tempo e história* (com adaptações).

Com base no texto acima, julgue os seguintes itens.

- 4 Para que esse texto faça parte de um documento oficial, respeitando-se as normas do padrão culto da língua portuguesa, a contração da preposição em “do” (l.1) deveria ser omitida ou, ao menos, o artigo nela presente deveria ser retirado.
- 5 A expressão “ser cheio de potencialidades” (l.3-4), que qualifica “Pico della Mirandola”, evidencia a admiração do autor do texto por esse “pensador italiano” (l.17).
- 6 Preservam-se a coerência e a correção gramatical do texto ao se utilizar, alternativamente, a preposição **a** no lugar da preposição “de” em “o direito de forjar” (l.5).

7 Apesar de, no texto, serem focalizadas reflexões sobre situações e eventos passados e empregados vários verbos no pretérito, o que se diz a respeito de “Um homem assim constituído” (l.7) está no presente, pois, se a oração fosse considerada de pretérito, o verbo seria grafado **podia** ou **pôde**, em vez de “pode” (l.8).

8 O pronome “a” (l.9) está empregado no feminino singular porque se refere a “sua passagem pelo mundo” (l.8).

9 Nas linhas 10 e 11, a repetição da conjunção “que” indica que a forma verbal “sugere” é complementada pelas duas orações iniciadas por essa conjunção, sendo, nesse caso, facultativa a segunda ocorrência da conjunção.

10 Pela natureza da informação, a oração “os medievais já sabiam disso” (l.16) poderia ser iniciada pela conjunção **pois** e demarcada pelo sinal de parênteses ou de duplo travessão no lugar de vírgulas.

11 Para se evitar o uso excessivo da palavra **que**, poderia ser feita a substituição do trecho “o fato de que o pensador” (l.17) por **o fato do pensador**, sem prejuízo da argumentação e correção gramatical do texto.

1 O futuro da humanidade será cada vez mais movido pelo uso intensivo de conhecimento, quer na atividade produtiva quer em atividades cotidianas, como entretenimento e convívio.

4 Criar mecanismos capazes de não apenas estimular a produção como também sua ampla difusão é tarefa para todos, devendo ocupar opção preferencial na agenda das empresas  
7 contemporâneas.

Compartilhar conhecimentos é um fato social tão relevante quanto o próprio ato de criação de novos  
10 conhecimentos, principalmente se levarmos em conta que boa parte da produção destes resulta da combinação de saberes já existentes e utilizados parcialmente ou com outros fins.

A gestão do conhecimento na Eletro nuclear, **UnB**  
*Revista*, maio/jun./jul./2005, p. 58 (com adaptações).

Julgue os seguintes itens a respeito do emprego das estruturas linguísticas do texto acima.

12 Para enfatizar a idéia de “uso intensivo de conhecimento” (l.2), uma redação para o primeiro período do texto, preservando-se a coerência e a correção, seria: O uso intensivo do conhecimento moverá cada vez mais o futuro da humanidade.

13 Preservam-se a coerência da argumentação e a correção gramatical do texto ao enumerar as atividades movidas pelo uso do conhecimento, empregando-se as conjunções **seja ... seja** ou **tanto ... quanto**, em lugar de “quer ... quer” (l.2-3).

14 Na linha 4, o deslocamento do advérbio “apenas” para depois de “capazes” preservaria a coerência e a correção gramatical do texto, com a vantagem de enfatizar os dois objetivos dos mecanismos propostos: estimular e difundir conhecimento.

15 Na linha 12, subentende-se do desenvolvimento das idéias no texto que a expressão “com outros fins” estabelece um paralelo com “parcialmente” porque as duas expressões modificam a ação de **utilizar**.

**Text related to items from 16 to 25.**

**1 Globalization:** The globalization of all aspects of the food and fiber system is having a sudden and surprising impact on American agriculture. Profound changes are seen worldwide from competitive markets around the world, from diseases not limited to national boundaries, to population growth and evolving diets. These changes have led to a dramatically new trade environment, threats of exotic diseases and pests to domestic production, and international controversies over the use of biotechnology. To remain competitive, the food and agriculture sector needs to take these developments into consideration.

**Information Access and Communication:** The explosion of information technology, the worldwide use of the Internet, and the major advancements of cyberspace communications are changing the way private industry, government, and individuals conduct daily business. Vast amounts of information soon will be available in “real time,” more people from around the world will be able to retrieve the information, and advanced computer software will make the information more useful and meaningful. Advancements in communication technology offer benefits and opportunities for everyone involved in the American food and agriculture sector.

**Workforce:** A very important employment issue is the need to recruit and retain a highly skilled and technically well trained Federal workforce. The relatively low U.S. unemployment rate makes recruitment highly competitive. This competitive environment is expected to require more employer emphasis on recruitment, retention, student employment, upward mobility, and training/retraining programs. The public sector will need to recruit a diversity of people and to maintain a highly qualified and technically competent workforce. Expanding job opportunities for women and minorities in science and engineering will help to tap the Nation’s human potential.

**Technology:** Advances in technology — such as bioengineering, precision agriculture, remote sensing, and decision modeling — can enable agricultural production to enhance nutrition, protect the environment, and continue to make the food supply safe. Biotechnology offers great promise for increasing production efficiency, improving food quality, and enhancing nutritional value. However, concerns about genetically modified organisms (GMO) have had a marked impact on international exports of affected commodities, and prompted questions about the potential benefits and risks. Precision agriculture, remote sensing, and decision modeling will increase production efficiency and mitigate adverse environmental impacts on agriculture. Public concern about food safety has led to new rapid detection technologies that, when fully implemented, will make the food supply increasingly safer.

Internet: < <http://www.ars.usda.gov/aboutus/docs.htm?docid=1800>> (with adaptations).

Based on the previous text, judge the following items.

- 16** The globalization of food production is bringing about dramatic changes worldwide.
- 17** The use of biotechnology is thought to be an undeniable advancement as far as global changes are concerned.
- 18** Never before has information technology been undergoing such deep changes all over the world.
- 19** Due to the shortage of US workforce, it is not all too easy to hire highly qualified personnel to confront present-day agricultural challenges.
- 20** Bioengineering, precision agriculture, remote sensing, and decision modeling are some of the factors required to make safe the food supply.
- 21** GMO caused an uncontroversial impact on international agricultural transactions.
- 22** Three factors — precision agriculture, remote sensing, and decision modeling — can actually prevent environmental disasters.

In the text,

- 23** “take these developments into consideration” (ℓ.10-11) is the same as **take these developments for granted**.
- 24** “retain” (ℓ.25) is synonymous with **detaim**.
- 25** “However” (ℓ.42) means **Nevertheless**.

Julgue os itens a seguir, acerca de noções e conceitos de estatística e de tratamento de dados estatísticos.

- 26** Considere-se que, em um experimento, um pesquisador calculou a altura média dos animais que sofriam de uma doença específica. O valor obtido foi igual a 156 cm e o desvio-padrão foi de 5 cm. Investigações posteriores, entretanto, revelaram que todas as medidas feitas estavam 2 cm maiores que os valores verdadeiros. Nessa situação, os valores corretos para a média e o desvio-padrão são 1,54 m e 0,05 m, respectivamente.
- 27** Considere-se que, em experimentos com ratos de laboratório, são registrados o sexo e o peso desses animais. O peso é freqüentemente expresso em gramas e o valor registrado é arredondado para o número inteiro mais próximo. Nesse contexto, quanto à classificação dessas variáveis, é correto afirmar que sexo é uma variável discreta e peso é uma variável contínua.
- 28** Considere a seguinte situação hipotética. Em um experimento, os números de carrapatos observados em cada um dos animais de um grupo, dispostos em ordem crescente, foram iguais a: 4, 7, 7, 9, 10, 11, 13, 15, 15, 17, 17, 19, 19, 20. Após terem sido calculadas a média, a mediana e a moda, um erro foi descoberto: um dos animais com 15 carrapatos tinha, na realidade, 17. Nessa situação, é correto afirmar que, entre as medidas de tendência citadas, apenas a média deve-se alterar após a correção do dado.
- 29** Considere que, na pesagem de 30 peixes, obteve-se a média de 30 g e desvio-padrão igual a 2 g. Depois de completar a pesagem, verificou-se que a balança estava descalibrada e que os pesos obtidos estavam 2 g abaixo do peso real. Nessa situação, é correto afirmar que o erro relativo dos dados deve diminuir após a correção dos dados.

- 30 Considere que se deseja fazer uma amostragem de 5 animais, sem reposição, de uma população de 50 e que cada animal tenha recebido uma numeração seqüencial de dois dígitos: 01, 02, (...), 50. Nessa situação hipotética, considerando a linha de uma tabela de números aleatórios mostrada abaixo, a escolha dos animais 11, 36, 23, 23 e 08 é apropriada para a amostragem desejada.

11362	35692	96237	90842	46843	62719	64049	17823
-------	-------	-------	-------	-------	-------	-------	-------

- 31 Verificou-se que, depois de tabular os dados de um experimento, houve necessidade de estabilizar a variância, porque a variância do tratamento na escala original variou diretamente com a média. Nesse caso, deve-se considerar uma transformação logarítmica dos dados.
- 32 Considere que, para determinar o grau de confiança na técnica de ultra-som para verificação de prenhez, tenham sido feitos 280 exames cujos resultados encontram-se resumidos na tabela abaixo. Nessa situação, caso um novo exame seja realizado e acuse que o animal não está prenhe, então, a probabilidade de erro tipo II é de  $\frac{15}{280}$ .

		situação real	
		não prenhez	prenhez
resultado do exame	não prenhez	131	15
	prenhez	9	125

- 33 É possível calcular a variância de um conjunto de dados sem determinar previamente o valor médio desse conjunto.

Classe (g)	Frequência
0 - 5	8
5 -10	2
10-15	6
15-20	8
20-25	5
25-30	5
30-35	0
35-40	1

A tabela acima ilustra a classificação por peso, em gramas, de uma amostra com 35 peixes. Considerando essas informações, julgue os itens seguintes.

- 34 O valor que se deve associar à terceira classe da tabela é de 15 g.
- 35 A média dos valores agrupados é igual a 15,35 g.

Visando investigar o efeito da salinidade do solo no crescimento de determinada vegetação, um pesquisador selecionou tipos diferentes de solo para constituírem os blocos e aplicou seis níveis diferentes de tratamento com sal (15, 20, 30, 35, 45 e 50 ppm). O resultado da análise de variância é mostrado abaixo.

fonte	gl	SQ	QM	F
modelo			111,39	13,73
erro		121,7		
total corrigido	23	1012,8		
trat		664,4		16,38
bloco	3			9,31

T test (DMS) para variável: BIOMASSA

Alpha= 0,05; gl= 15; EMS= 8,1118;  
valor crítico para T= 2,13;  
Diferença Mínima Significativa= 4,2926

	Média	T-Grupo	N
Trat.			
20	18,100	A	4
15	14,150	A	4
30	7,475	B	4
35	6,000	BC	4
45	5,775	BC	4
50	3,075	C	4

Em relação a essa situação hipotética e aos dados da análise de variância apresentados acima, julgue os itens a seguir.

- 36 O experimento apresenta 8 blocos.
- 37 A diferença no resultado de dois tratamentos distintos só é significativa se as suas respectivas médias diferirem em mais de 8,118.
- 38 O experimento apresenta quatro repetições por tratamento, uma por bloco.
- 39 A produção média de biomassa para o tratamento com 30 ppm não foi significativamente diferente daquela com 35 ppm e esta, por sua vez, não foi significativamente diferente daquela com 50 ppm. Assim, não há evidência de que existe diferença significativa entre os tratamentos com 30 ppm e com 50 ppm de sal.
- 40 Nessa análise, o erro possui 15 graus de liberdade.
- 41 O nível de significância para determinar o DMS foi de 1%.
- 42 O quadrado médio de blocos é igual a 226,7.
- 43 Se o F da Tabela para 8 e 15 graus de liberdade no numerador e no denominador, respectivamente, com  $\alpha = 0,01$ , foi igual a 4,00, é correto concluir que houve diferença significativa entre os tratamentos ao nível de 1%.
- 44 A maior produção de biomassa ocorreu no nível de salinidade de 20 ppm.
- 45 O desdobramento dos graus de liberdade de tratamento pode fornecer informações sobre regressões de até o 5.º grau.

A metodologia científica é um conjunto de abordagens, técnicas e processos utilizados pela ciência para formular e resolver problemas de aquisição objetiva do conhecimento, de maneira sistemática. Com referência a esse assunto, julgue os itens subsequentes.

- 46 A hipótese é o conjunto de teses que explicam um fenômeno, enquanto tese é uma afirmação comprovada sobre algum fenômeno e teoria corresponde a uma afirmação ainda não comprovada sobre algum fenômeno.
- 47 Na comparação entre estudos quantitativos e qualitativos, o quadro teórico e as hipóteses devem ser mais rigorosamente definidos nos qualitativos.
- 48 Na pesquisa *ex-post-facto*, há uma investigação sistemática e empírica em que o pesquisador não tem controle direto sobre as variáveis independentes, porque já ocorreram suas manifestações ou porque estas são intrinsecamente não manipuláveis. Nesse caso, são feitas inferências sobre as relações entre variáveis em observação direta, a partir da variação concomitante entre as variáveis independentes e dependentes.
- 49 Em estudos observacionais, a casualização é um fator imprescindível.
- 50 Considere que, na concepção de um experimento para determinar a eficácia de uma vacina contra clostridiose, selecionou-se um lote uniforme de ratos, com mesma idade, que foram divididos aleatoriamente em dois grupos (A e B). Nos animais do grupo A foi injetada vacina, enquanto nos animais do grupo B nada foi injetado. Os ratos dos dois grupos foram alojados em gaiolas individuais e mantidos sob as mesmas condições de alimentação, temperatura e fotoperíodo. Nessa situação, foi cometida pelo menos uma falha grave na concepção do experimento.

RASCUNHO

## CONHECIMENTOS ESPECÍFICOS

Um pesquisador, para avaliar o desenvolvimento de mudas de uma certa planta frutífera ao longo do tempo em função da adubação do substrato, estabeleceu quatro níveis de concentração ( $a_1, a_2, a_3, a_4$ ) de um fertilizante X, e quatro níveis de concentração ( $b_1, b_2, b_3, b_4$ ) de um fertilizante Y. Ele combinou essas quatro possibilidades para cada fertilizante, o que resultou em 16 combinações diferentes; em seguida, plantou  $n$  mudas para cada combinação de fertilizantes X e Y, todas com características idênticas e selecionadas ao acaso. Cada muda foi monitorada ao longo do tempo e suas alturas foram registradas 15, 30, 60 e 90 dias após o plantio. A estrutura de correlação das alturas das mudas ao longo do tempo é estimada pela seguinte matriz de covariância.

$$\begin{bmatrix} 0,8 & 0,7 & 0,4 & 0,2 \\ 0,7 & 1,5 & 0,8 & 0,6 \\ 0,4 & 0,8 & 2,5 & 3,0 \\ 0,2 & 0,6 & 3,0 & 5,0 \end{bmatrix}$$

Considerando a situação hipotética descrita acima e que não houve perdas de dados, julgue os itens a seguir.

- 51 O experimento pode ser denominado como fatorial  $4 \times 4 \times 4$ , em que os dois primeiros fatores são os níveis de concentração dos fertilizantes X e Y e o terceiro fator é o tempo, cujos níveis são os dias de registro das alturas das mudas após o plantio.
- 52 A situação descrita não permite utilizar um modelo de análise de variância multivariada.
- 53 Para avaliar os perfis das alturas médias por meio de curvas de crescimento em função do tempo, pode-se usar os modelos polinomiais de Potthoff e Roy.
- 54 Os dados mostrados na matriz de covariância permitem concluir que o pesquisador observou uma crescente variabilidade das alturas das plantas ao longo do tempo.
- 55 A situação descrita permite utilizar um modelo de efeitos fixos e aleatórios. Essa classe de modelo também é conhecida como modelo hierárquico ou em dois estágios.
- 56 A coleta de dados descrita pode ser corretamente classificada como coleta de dados longitudinais.
- 57 Considere que  $n = 1$  e que o pesquisador esteja interessado apenas nos resultados observados 90 dias após o plantio. Nessa situação, o delineamento é do tipo *split-plot*.
- 58 Considere que  $n = 1$  e que o pesquisador esteja interessado em avaliar os resultados observados 30 e 90 dias após o plantio. Nessa situação, o delineamento é do tipo *quadrados latinos*.
- 59 A matriz de covariância apresenta uma estrutura de correlação do tipo simetria composta (*compound symmetry*).

Um pesquisador desejava avaliar a distribuição do tempo de funcionamento até a ocorrência da primeira falha quando um equipamento agrícola operava continuamente. Como o equipamento poderia funcionar por um longo período de tempo sem falhar, o pesquisador fixou o tempo máximo de observação em 30 dias. Para esse estudo, foram escolhidos, ao acaso, 100 desses equipamentos e, ao final, produziram-se as tabelas a seguir. Os problemas inesperados, mencionados na tabela, foram causados por motivos diversos que forçaram a interrupção do experimento em 10 dos 100 equipamentos.

evento	número de equipamentos
foi observada a primeira falha antes de 30 dias	60
não foi observada a primeira falha antes de 30 dias	30
ocorrência de problemas inesperados	10

intervalo de tempo (dias)	n.º sob risco (a)	n.º de falhas (b)	n.º de censuras (c)	estimativa da probabilidade de falha = $b/(a-c/2)$
1 a 5	100	5	5	0,05
6 a 10	90	10	0	0,11
11 a 15	80	20	5	0,26
16 a 20	55	15	0	0,27
21 a 25	40	5	0	0,12
26 a 30	35	5	30	0,25

No que se refere à situação hipotética acima, julgue os itens subsequentes.

- 60 A situação descrita apresenta um estudo de confiabilidade com mecanismos de censura aleatória do tipo II.
- 61 Como o pesquisador deseja avaliar a distribuição do tempo de funcionamento até a ocorrência da primeira falha, apenas as 60 observações de falha antes dos 30 dias devem ser utilizadas no estudo. As observações restantes devem ser descartadas, pois não representam o tempo de funcionamento até a primeira falha.
- 62 O texto apresenta uma estimativa de Kaplan-Meier da confiabilidade do equipamento.
- 63 A estimativa da probabilidade de um equipamento falhar entre 11 e 15 dias de funcionamento é igual a 0,26.
- 64 Considerando as estimativas de probabilidade de falha apresentadas, a estimativa da função de confiabilidade no tempo de 12 dias é menor que 0,65.

Um pesquisador está desenvolvendo um estudo sobre um sistema hidrológico-meteorológico. Parte desse estudo consiste na modelagem da vazão diária ( $x_t$ ) de um rio, em  $\text{m}^3/\text{s}$ , em função de vazões passadas e de variáveis meteorológicas. Nesse estudo, o pesquisador coletou dados sobre a precipitação diária ( $y_t$ ), em mm, e a temperatura média do dia ( $z_t$ ), em  $^\circ\text{C}$ . Os dados meteorológicos foram coletados em uma estação nas proximidades do rio e a vazão foi observada em uma das represas existentes ao longo do rio. Como o estudo está em fase inicial, o pesquisador dispõe apenas de dados referentes ao ano de 2005 e uma parte de 2006 ( $n = 400$  observações). Um relatório parcial emitido pelo pesquisador apresenta três modelos preliminares, mostrados abaixo.

- I  $x_t = 9 - 0,1y_t + 0,2z_t + 0,1z_{t-1} + 15\varepsilon_t$ ,  
em que  $\varepsilon_t$  é o erro aleatório no instante  $t$  com distribuição normal padrão; o valor do critério de informação de Akaike (AIC) nesse modelo foi igual a 3.
- II  $x_t = 1,7 + 0,3x_{t-1} + 0,1y_{t-1} + 0,1z_t - 0,05z_{t-1} + 1,5\varepsilon_t$ ,  
em que  $\varepsilon_t$  é o erro aleatório no instante  $t$  com distribuição normal padrão; o valor do critério de informação de Akaike (AIC) nesse modelo foi igual a 1.
- III Se  $x_{t-1} \leq 15$ , então  $x_t = 0,5 + 1,2x_{t-1} + 0,08y_{t-1} + 0,3z_t + \varepsilon_t$ ,  
se  $x_{t-1} > 15$ , então  $x_t = 7 + 1,1x_{t-1} + 0,3y_{t-1} + 0,5z_t + 4\varepsilon_t$ ,  
em que  $\varepsilon_t$  é o erro aleatório no instante  $t$  com distribuição normal padrão; o valor do critério de informação de Akaike (AIC) nesse modelo foi igual a 0,5.

De acordo com o relatório, os métodos usuais de diagnóstico dos resíduos não rejeitam a hipótese de aleatoriedade residual. Além disso, tanto a função de autocorrelação amostral como a função de autocorrelação parcial amostral dos resíduos não apresentaram valores diferentes de zero ao nível de significância igual a 5%. Nos modelos apresentados, a estatística de Ljung-ox e de McLeod-Li foram iguais a 25 e 15, respectivamente.

Acerca dessa situação hipotética, julgue os próximos itens.

- 65 O modelo I tem a forma de regressão linear.
- 66 O modelo II tem a forma de função de transferência.
- 67 O modelo III é conhecido como modelo auto-regressivo bilinear de ordem 1 ou BL-AR(1).
- 68 O erro quadrático médio (*mean squared error*) do modelo I é maior ou igual a 100.
- 69 Pelo critério AIC, o melhor modelo entre os três que foram apresentados é o I, por ser o mais parcimonioso (tem o menor número de parâmetros).
- 70 No modelo III, o ciclo produzido pelas estações do ano (componente sazonal) é composto por dois regimes: um para vazões maiores que  $15 \text{ m}^3/\text{s}$  e outro para vazões menores que  $15 \text{ m}^3/\text{s}$ .
- 71 A função de autocorrelação amostral dos resíduos apresenta valores entre  $-0,1$  e  $+0,1$ , incluindo estes.
- 72 Para o diagnóstico de resíduos, o teste de Durbin-Watson, em geral, não é equivalente ao teste de Ljung-Box.
- 73 O teste de Ljung-Box testa se os resíduos são autocorrelacionados. A estatística desse teste consiste, basicamente, de uma soma dos quadrados das autocorrelações amostrais entre os resíduos.
- 74 O teste de McLeod-Li aplica-se para resíduos independentes, identicamente distribuídos e não necessariamente normais. Para a realização desse teste, os resíduos são elevados ao quadrado e, sobre a série de resíduos ao quadrado, aplica-se o mesmo procedimento aplicado no teste de Ljung-Box.

Um pesquisador precisa avaliar os efeitos de uma nova substância contra uma doença que ataca um tipo de animal. Para isso separou os animais em dois grupos, A e B. Cada grupo era formado por  $n$  animais selecionados ao acaso, todos portadores da doença. Os animais do grupo A foram tratados com um medicamento usual, enquanto os do grupo B foram tratados com a nova substância. Porém, como o pesquisador não sabia qual era a dose adequada dessa nova substância, ele aplicou diferentes doses para cada animal do grupo B. Entre outros resultados, o pesquisador desejava saber se a nova substância era mais eficaz que o medicamento usual e, em caso afirmativo, quantas vezes a nova substância era mais eficaz que este e em que dose essa nova substância deveria ser aplicada de modo a produzir efeitos positivos.

Com relação à situação hipotética relatada acima, julgue os itens a seguir.

- 75 A situação apresentada descreve um estudo de coorte prospectivo.
- 76 Considere que o pesquisador esteja interessado apenas nos resultados observados nos animais do grupo B. Nesse caso, o experimento é denominado *cross-over*.
- 77 Considere que o pesquisador forme um novo grupo C, formado por animais do grupo B que foram submetidos a uma mesma dose da nova substância. Considere também que ele deseje comparar os resultados observados para o grupo A com os do grupo C. Essa situação caracteriza um estudo de caso-controle.

RASCUNHO

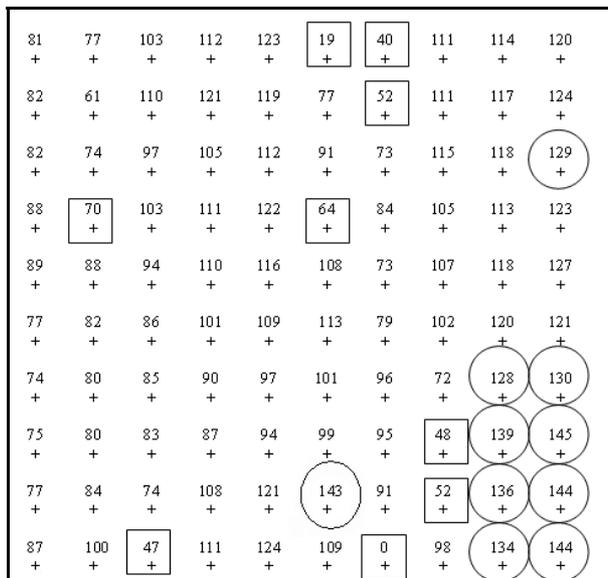


Figura I

92,3	99,3	88,6	103,1
+	+	+	+
17,1	26,0	30,9	25,7
91,1	102,6	98,3	106,7
+	+	+	+
12,2	13,6	17,7	18,5
86,3	98,3	94,3	106,3
+	+	+	+
9,1	10,3	17,4	26,5
83,9	98,3	90,0	103,3
+	+	+	+
14,5	21,5	32,9	41,3

Figura II

A figura I, acima, mostra um conjunto de valores, os mais baixos indicados por quadrados e os mais altos indicados por círculos. A figura II mostra os valores médios — valores acima do sinal “+” — e os seus respectivos desvios-padrão — valores abaixo do sinal “+” —, obtidos após a execução de uma máscara móvel com dimensão 4x4. Com base nesses dados, julgue os seguintes itens.

- 78** Na área de amostragem, existe uma tendência de os valores mais baixos ficarem distribuídos na direção norte-sul.
- 79** O uso da máscaras móveis retangulares usualmente facilita os cálculos computacionais.
- 80** A faixa de variação dos valores médios na área de amostragem é maior que a faixa de variação dos desvios-padrão.
- 81** Se um programa computacional fosse utilizado para gerar a curva de isovalores com intervalo de 10 unidades, o resultado final apresentaria isolinhas variando de 10 a 150 unidades.

A geoestatística é uma ferramenta básica na pesquisa agropecuária. Ela é utilizada com frequência, por exemplo, na geração de mapas de produtividade, de declividade — a partir de dados pontuais de elevação — e de curvas de nível, bem como na espacialização de dados de precipitação, provenientes de estações pluviométricas. Julgue os itens seguintes, relativos a conceitos e aplicações da geoestatística na pesquisa agropecuária.

- 82** Uma das diferenças básicas entre a *krigagem* e outros métodos de interpolação é a forma com que os pesos são atribuídos às diferentes amostras.
- 83** Nas interpolações que envolvem o método de Thiessen, o tamanho das células depende somente da densidade dos dados amostrados.
- 84** As versões mais recentes — por exemplo, versão 4.2 — do aplicativo brasileiro de domínio público denominado SPRING, que se destina ao processamento de informações georreferenciadas, trazem opção para análise geoestatística, o que não acontece com o aplicativo comercial internacional denominado ArcView GIS 3.2.

Dados topográficos são largamente utilizados em gestão territorial, manejo de recursos naturais e zoneamentos ecológico-econômicos. Com relação a esse tema, julgue os itens subsequentes.

- 85** Não é possível espacializar, por meio de técnicas de geoestatísticas, os dados interferométricos conhecidos como SRTM (*Shuttle Radar Topography Mission*), obtidos pela NASA, devido à incompatibilidade entre o formato usado na codificação desses dados e o formato comumente usado em aplicativos comerciais.
- 86** Em um sistema de informações geográficas, os dados de elevação costumam ser convertidos para declividade e aspecto. A declividade pode ser dada em graus ou em porcentagem. O aspecto está relacionado com direção ou orientação da declividade.
- 87** Atualmente, dados de altimetria são normalmente obtidos no campo por meio de bússolas.
- 88** As grades triangulares permitem uma boa representação de relevos acidentados, mas não são as mais adequadas para visualização em 3D.
- 89** Na triangulação de Delaunay, utiliza-se o critério de maximização dos ângulos mínimos de cada triângulo. Isso significa que a rede triangular final deve conter triângulos que sejam o mais próximo possível de triângulos equiláteros.

Acerca das distribuições de probabilidade e do teorema do limite Central, julgue os itens seguintes.

- 90** Se uma variável aleatória  $X$  tem distribuição binomial com parâmetros  $p = 0,1$  e  $n = 100$ , então  $P\{X \leq a\} \approx \Phi\left(\frac{a - 10}{\sqrt{9}}\right)$  em que  $\Phi(\bullet)$  é a função de distribuição normal padrão.
- 91** Pode-se utilizar a distribuição binomial para, por exemplo, calcular a probabilidade de se encontrar  $k$  sementes que sofreram mutações genéticas, em um lote de  $n$  sementes selecionadas ao acaso sem reposição.
- 92** Se a variável aleatória  $X$  tem distribuição normal  $(\mu, \sigma^2)$ , então  $P(x \leq \mu + 2\sigma) = P(|x - \mu| \leq 2\sigma)$ .

93 Se a variável aleatória  $W$  tem distribuição exponencial com parâmetro  $\alpha$ :  $f_{\text{exp}}(w|\alpha) = \alpha^{-1}e^{(-\alpha^{-1}w)}$  se  $0 < w < \infty$ , então  $P(w > \alpha) = e^{-1}$ .

Julgue os seguintes itens, com relação ao modelo de regressão linear para dados seccionais:

$$y_i = \beta_0 + \beta_1 x_{1i} + \beta_2 x_{2i} + \dots + \beta_k x_{ki} + u_i, \text{ com } i = 1, \dots, n.$$

- 94 Os coeficientes de inclinação não se alteram quando as unidades de medida de  $y$  e dos  $x$  são multiplicadas por uma constante, como ocorre, por exemplo, quando os valores dessas variáveis são convertidos de uma unidade para outra.
- 95 Sempre que o modelo tiver pelo menos uma variável explicativa além do intercepto, o coeficiente de determinação  $R^2$  será maior que o  $R^2$  ajustado.
- 96 A inclusão de regressores irrelevantes no modelo resulta em coeficientes estimados não-viciados, mas ineficientes.
- 97 Na presença de heterocedasticidade, o estimador de mínimos quadrados (MQO) é viesado e não se pode confiar nos procedimentos de testes usuais (testes  $F$  e  $t$ ), já que os estimadores são também ineficientes.
- 98 O estimador de mínimos quadrados (MQO), assim como a média aritmética, é muito sensível à presença de observações discrepantes (*outliers*).

Com relação a *data mining* (mineração de dados), julgue os próximos itens.

- 99 Em modelos de classificação, ocorre *overfitting* quando o número de erros cometidos no grupo de dados usado para treinar (ajustar) o modelo é muito pequeno e o número de erros de generalização é grande.
- 100 Na análise de *cluster*,  $K$ -médias é uma técnica de clusterização hierárquica (aninhada), que visa encontrar o número de clusters ( $K$ ) especificados pelo usuário, representados por seus centróides.