

Universidade Federal de Viçosa
Departamento de Estatística

Prova Seletiva para o Programa de Pós-Graduação em Estatística Aplicada e Biometria.

Nível Mestrado - dezembro/2012

Nome: _____

Assinatura: _____.

Número do CPF ou RG: _____.

- Interpretar corretamente as questões é parte da avaliação.
- Indique todos os cálculos organizadamente.
- São 6 questões em páginas numeradas de 1 a 7, total de 100 pontos.
- O tempo máximo para realização desta prova é de 4 horas.
- **BOA SORTE !!**

1 (20%). Em um experimento com salsinha foram testados 2 tratamentos (T_1 , o controle sem pré-resfriamento e T_2 , o uso de pré-resfriamento com água gelada) e 7 épocas ou tempos de avaliação ($TE = 0, 1, 2, 3, 4, 5$ e 6 dias após o tratamento). O delineamento experimental foi o inteiramente casualizado no esquema de parcelas subdivididas, com 4 repetições. Foi analisada a variável dependente "teor relativo de água", submetido inicialmente à análise de variâncias (ANOVA) conforme o quadro a seguir.

a.(15%) Complete o quadro da ANOVA.

b.(5%) Indique os testes F de interesse com as respectivas hipóteses de nulidade.

FV	GL	SQ	QM	F_c
Tratamentos (T)		176,8612		
Resíduo (a)		30,8443		
TE		4954,7409		
Interação T \times TE		111,3411		
Resíduo (b)				
Total		5654,1774		

2 (10%). Um experimento foi conduzido com o intuito de comparar a produção (t/ha) de oito famílias de milho pipoca. Devido à heterogeneidade da área experimental, foram considerados quatro blocos. Cada bloco recebeu todas as famílias, cuja ocupação dentro de cada bloco foi dada por meio de sorteio. Quais foram **os princípios básicos da experimentação empregados neste experimento?** Justifique a utilização de cada um deles.

3 (30%). Considere as 5 doses de um adubo (X , em centenas de kg por hectare) e os totais de produção (Y , em toneladas por hectare) obtidos em um estudo objetivando-se ajustar a equação de regressão linear simples $Y_i = \beta_0 + \beta_1 X_i + \varepsilon_i$.

X	0	1	2	3	4
Y	2,0	2,2	3,0	2,8	4,0

- (8%) Calcule e interprete o coeficiente de regressão.
- (8%) Calcule e interprete a constante da regressão.
- (8%) Calcule e interprete o coeficiente de determinação.
- (6%) Estime a produção média, em toneladas por hectare, quando se deseja aplicar 500 kg do adubo. Comente o resultado.

Algumas dicas:

$$\hat{Y}_i = \hat{\beta}_0 + \hat{\beta}_1 X_i \quad \hat{\beta}_1 = \frac{SPD_{XY}}{SQD_X} \quad \hat{\beta}_0 = \bar{Y} - \hat{\beta}_1 \bar{X} \quad r^2(\%) = \frac{SQ_{\text{regressão}}}{SQ_{\text{total}}} 100\%$$

$$\hat{\varepsilon}_i = Y_i - \hat{Y}_i \quad SQ_{\text{total}} = SQD_Y \quad SQ_{\text{regressão}} = \frac{(SPD_{XY})^2}{SQD_X} = \hat{\beta}_1^2 SQD_X = \hat{\beta}_1 SPD_{XY}$$

4 (10%). Amostras da profundidade efetiva do sistema radicular (metros) de duas variedades de soja (S1 e S2) são apresentadas na tabela abaixo.

Variedade	Amostras					
S1	0,50	0,52	0,53	0,50	0,52	0,50
S2	0,45	0,49	0,46	0,52	0,38	0,52

- a.(5%) Para qual variedade as amostras são mais homogêneas? Justifique.
- b.(5%) Sabendo-se que houve um erro de mensuração nas avaliações da variedade S2, deve-se efetuar uma correção dada pela adição da constante 0,06 a cada observação. Considerando-se esta correção, o resultado do item **a.** seria alterado? Justifique utilizando-se as **propriedades de média e variância.**

5 (15%). Um avião desapareceu e acredita-se que ele possa ter caído em uma dentre 3 possíveis regiões, com igual probabilidade. Considere que a probabilidade condicional de uma expedição de busca na i -ésima região não encontrar o avião, dado que ele está na i -ésima região, seja igual a P_i , $i = 1, 2, 3$. Pede-se: dado que uma expedição de busca na região 1 não encontrou o avião, calcule a probabilidade condicional,

- a. de que o avião esteja na região 1.
- b. de que o avião esteja na região 2.

6 (15%). Suponha que 40% dos funcionários de uma empresa sejam favoráveis a uma greve geral. Se uma amostra aleatória formada por 8 funcionários for investigada (entrevistados um a um, separadamente e sigilosamente), qual será a probabilidade de obtermos no máximo 3 respostas favoráveis à greve?

a.(8%) Indique os cálculos pelo modelo binomial.

b.(7%) Calcule com a aproximação pelo modelo Poisson.

Universidade Federal de Viçosa
Departamento de Estatística

Prova Seletiva para o Programa de Pós-Graduação em Estatística Aplicada e Biometria.

Nível Mestrado - Junho/2013

Nome: _____

Assinatura: _____

Número do CPF ou RG: _____

- Interpretar corretamente as questões é parte da avaliação.
- Indique todos os cálculos organizadamente.
- São 5 questões em páginas numeradas de 1 a 6, total de 100 pontos.
- O tempo máximo para realização desta prova é de 4 horas.
- **BOA SORTE !!**

1 (20%). Em um centro de pesquisa, foi instalado um experimento no delineamento em blocos casualizados completos, com quatro repetições. O objetivo do experimento era avaliar a influência de cinco tipos de gramíneas sobre o consumo diário de sal de bovinos (g/dia). Os blocos foram formados em função da classe de peso dos animais.

Classes de Peso	1	2	3	4
Total	276	278	287	339

Sabendo-se que,

$$SQ_{\text{Total}} = 652 \quad \text{e que} \quad SQ_{\text{Gramínea}} = 80,$$

e que,

$$F_{0,01}(5; 11) = 5,32; \quad F_{0,01}(4; 12) = 5,41; \quad F_{0,01}(4; 11) = 5,67; \quad \text{e} \quad F_{0,01}(3; 12) = 5,95,$$

faça a ANOVA e conclua se há efeito significativo do tipo de gramínea sobre o consumo diário de sal, em nível de 1% de significância.

2 (20%). Para avaliar hábitos e atitudes de estudo existe um índice denominado SSHA (Survey of Study Habits and Attitudes Brown, W. F. e Holtzman, W. H. (1967)) cujos valores oscilam entre 0 e 200. Estudantes de ambos os sexos foram pesquisados e os respectivos valores de SSHA são apresentados abaixo. Acredita-se que a 5% o SSHA médio do sexo Feminino (F) seja superior ao do sexo Masculino (M) .

F: 154; 109; 137; 115; 152; 140; 154; 178; 101 e 148.

M: 108; 140; 114; 91; 115; 126; 92; 169; 146 e 104.

Para F: $\sum_i X_i = 1388$ e $\sum_i X_i^2 = 197800$

Para M: $\sum_i Y_i = 1205$ e $\sum_i Y_i^2 = 150799$

- (a) Sugira uma sequência de testes de hipóteses, “coerentes segundo a teoria”, para verificar essa suposição do problema.
- (b) Os testes de normalidade para “F” e “M” foram executados no programa R, conforme resultados abaixo. A hipótese testada é: H_0 : a amostra é de uma população normal. Interprete-os CONFORME o conhecimento do significado do “p-value” que em português é traduzido como valor-p ou nível crítico ou probabilidade de significância.
Shapiro-Wilk normality test statistics = W
data F: W = 0.9388, p-value = 0.5394
data M: W = 0.9387, p-value = 0.5388
- (c) Conclua sobre a suposição do problema. Apresente os cálculos e a decisão caso estivessem disponíveis os valores tabelados necessários.
- (d) Que tipo de erro pode estar sendo cometido e com que probabilidade? Detalhe o máximo possível.
- (e) Como fica a questão do Poder do Teste nesse caso? Detalhe o máximo possível.

3 (20%). Considere uma amostra de n variáveis aleatórias independentes e identicamente distribuídas,

$$X_1, X_2, \dots, X_n \stackrel{i.i.d.}{\sim} E(X_i) = \mu, V(X_i) = \sigma^2.$$

Considere também as seguintes variáveis aleatórias, em que l_i e m_i são constantes,

$$Y = \sum_{i=1}^n l_i X_i \quad Z = \sum_{i=1}^n m_i X_i \quad \bar{X} = \frac{\sum_{i=1}^n X_i}{n}$$

(a) Calcule $\text{COV}(X_i - \bar{X}, X_j - \bar{X})$, para $i \neq j$ quaisquer números inteiros do conjunto $1, 2, \dots, n$.

(b) Mostre que se $\sum_{i=1}^n l_i m_i = 0$ então Y e Z são variáveis aleatórias não correlacionadas ($\rho_{Y,Z} = 0$).

4 (20%). (exame ANPEC 2002 - prova de estatística) Uma companhia de seguros tem 400 segurados de certo tipo. O prêmio do seguro é R\$ 1000 por ano. A seguradora indeniza R\$ 8000 a cada segurado por sinistro. A probabilidade de ocorrência de sinistro por segurado é igual a 0,1 por ano e os custos fixos da seguradora são de R\$ 8000 por ano. Calcule a probabilidade da seguradora ter prejuízo em um dado ano. Atenção: apenas indique todo o procedimento do cálculo e quais modelos de probabilidade seriam empregados, isto é, como aproximar o modelo binomial pelo normal padrão.

5 (20%). O processo de Controle da Qualidade de uma linha de produção examina, uma após a outra, as peças fabricadas. Seja 0,05 a probabilidade de se achar uma peça defeituosa em cada exame. Se duas peças defeituosas forem achadas até a quinta peça ser examinada, a linha é parada para se detectar e corrigir as causas do defeito. Se isto não ocorrer até a quinta peça ser examinada a produção não é paralisada. Pede-se: calcule a probabilidade de a paralisação ocorrer antes que a quinta peça seja examinada.

Universidade Federal de Viçosa
Departamento de Estatística

Prova Seletiva para o Programa de Pós-Graduação em Estatística Aplicada e Biometria.

Nível Mestrado - 22/novembro/2013

Nome: _____

Assinatura: _____.

Número do CPF ou RG: _____.

- Interpretar corretamente as questões é parte da avaliação.
- Indique todos os cálculos organizadamente.
- São 5 questões em páginas numeradas de 1 a 8, total de 100 pontos.
- O tempo máximo para realização desta prova é de 4 horas.
- **BOA SORTE !!**

1 (20%). Sabe-se que os pesos de coelhos (Y) de uma determinada espécie é uma variável aleatória normalmente distribuída com média λ quilogramas (quilos ou kg) e variância igual a seis kg^2 . Admita que λ é a média de uma distribuição Poisson em que $P(X = 1) = P(X = 3)$. Pedese: calcule a probabilidade de que o peso de um coelho selecionado aleatoriamente esteja entre três e cinco quilos.

Tabela 1: Áreas de uma distribuição normal padrão, $P(0 \leq Z \leq z)$. Tabela parcialmente reproduzida, adaptada de Costa Neto, P. L. O. Estatística, Editora Edgard Blucher.

z	0,00	0,01	0,02	0,03	0,04	0,05	0,06	0,07	0,08	0,09
0,0	0,0000	0,0040	0,0080	0,0120	0,0160	0,0199	0,0239	0,0279	0,0319	0,0359
0,1	0,0398	0,0438	0,0478	0,0517	0,0557	0,0596	0,0636	0,0675	0,0714	0,0753
0,2	0,0793	0,0832	0,0871	0,0910	0,0948	0,0987	0,1026	0,1064	0,1103	0,1141
0,3	0,1179	0,1217	0,1255	0,1293	0,1331	0,1368	0,1406	0,1443	0,1480	0,1517
1,0	0,3413	0,3438	0,3461	0,3485	0,3508	0,3531	0,3554	0,3577	0,3599	0,3621
1,1	0,3643	0,3665	0,3686	0,3708	0,3729	0,3749	0,3770	0,3790	0,3810	0,3830
1,2	0,3849	0,3869	0,3888	0,3907	0,3925	0,3944	0,3962	0,3980	0,3997	0,4015
1,3	0,4032	0,4049	0,4066	0,4082	0,4099	0,4115	0,4131	0,4147	0,4162	0,4177
1,4	0,4192	0,4207	0,4222	0,4236	0,4251	0,4265	0,4279	0,4292	0,4006	0,4319
1,5	0,4332	0,4345	0,4357	0,4370	0,4382	0,4394	0,4406	0,4418	0,4429	0,4441

2 (20%). Considerando a realização de um teste de hipóteses, pede-se:

a. (7%) Defina erro tipo I e erro tipo II.

b. (4%) Defina valor-p, também denominado de nível crítico ou probabilidade de significância.

c. (3%) Se for obtido para um teste de hipóteses um valor-p igual a 0,75, qual deve ser a decisão do teste considerando-se um nível de significância igual a 0,05?

d. (6%) É possível obter um valor-p igual ao nível de significância? Explique.

3 (20%). Abaixo estão listados as alturas (em polegadas) e o faturamento anual (em milhões de U\$), respectivamente da esquerda para a direita, das modelos Michelle Alves, Paris Hilton, Naomi Campbell, Gisele Bundchen, Heidi Klum, Adriana Lima, Kate Moss, Alessandra Ambrósio e Gretchen.

Altura	70	70,5	68	65	70	70	70	70	71
Faturamento	117	119	105	115	119	127	113	123	115

Considere que o faturamento anual das modelos se relaciona com a altura segundo o modelo de regressão linear simples. Pede-se:

a. (5%) Determine a equação de regressão ajustada.

b. (5%) Interprete a estimativa do coeficiente de regressão.

c. (5%) Encontre o valor estimado para o faturamento anual médio de uma modelo cuja altura é de 72 polegadas. Comente sobre esta estimativa.

d. (5%) Calcule o coeficiente de determinação e interprete-o.

4 (20%). Em teoria de probabilidade são apresentados três conceitos, ou maneiras de se atribuir probabilidades a eventos: conceito clássico (ou probabilidade a priori), conceito empírico (ou probabilidade a posteriori) e conceito moderno (ou probabilidade axiomática). Discuta como se atribui um valor de probabilidade a certo evento usando cada um desses conceitos e apresente pontos importantes para que tais formas de cálculo sejam válidas.

5 (20%). Em um experimento instalado no delineamento inteiramente ao acaso para avaliar 3 variedades de soja, com 3 repetições, foram obtidos os seguintes dados de produtividade (t/ha).

Variedades	Repetições		
	1	2	3
UFV-1	2,0	2,3	2,4
UFV-2	3,0	3,1	3,8
UFV-3	5,0	5,4	5,2

Pede-se:

a. (15%) Apresente o quadro da ANOVA e conclua para $\alpha = 1\%$. Dado $F(1\%) = 10,92$.

b. (5%) O teste F é conclusivo? Justifique sua resposta.

Universidade Federal de Viçosa
Departamento de Estatística

Prova Seletiva para o Programa de Pós-Graduação em Estatística Aplicada e Biometria.

Nível Mestrado - Novembro/2014

Nome: _____

Assinatura: _____

Número do CPF ou RG: _____

- Interpretar corretamente as questões é parte da avaliação.
- Indique todos os cálculos organizadamente. UTILIZE DUAS CASAS DECIMAIS nos cálculos.
- São 5 questões em páginas numeradas de 1 a 9, total de 100 pontos.
- O tempo máximo para realização desta prova é de 4 horas.
- **BOA SORTE !!**

1 (20%). Considere os dados apresentados no Quadro 1 a seguir, relativos a um experimento instalado no delineamento em blocos ao acaso, com 4 repetições, em esquema fatorial para avaliar o efeito de 3 níveis do fator A e 2 níveis do fator B.

Quadro 1 - Totais para tratamentos.

Fator A (4)	Fator B		Totais (8)
	B1	B2	
A1	60	80	140
A2	85	100	185
A3	100	110	210
Totais (12)	245	290	535 (24)

a.(8%) Calcule a soma de quadrados para a interação entre os fatores A e B ($SQ_{A \times B}$).

b.(6%) Apresente o quadro parcial da Anova com somente as Fontes de Variação e respectivos Graus de Liberdade.

c.(6%) Sabendo-se que a interação entre os fatores foi não significativa, aplique o teste de Scheffé a 1% de significância para o contraste:

$$C = m_{A_1} + m_{A_2} - 2m_{A_3}$$

Dados:

$$\text{QMResíduo} = 6,24; F_{1\%} = 6,36; S = \sqrt{(I - 1) \times F_{\alpha} \times \widehat{V}(\widehat{C})};$$

$$\widehat{V}(\widehat{C}) = \text{QMResíduo} \times \sum_i \frac{a_i^2}{r_i}$$

2 (20%). Considere a variável aleatória discreta X com a seguinte distribuição de probabilidade apresentada na sua forma tabular,

x	1	2	3
$P(x)$	0,50	0,25	$P(3)$

a.(6%) Calcule o valor para $P(3)$.

b.(6%) Faça a representação gráfica da distribuição de probabilidades de X .

c.(8%) Obtenha a Função de distribuição acumulada de X , geralmente representada por $F(x)$ e faça seu gráfico.

3 (20%). Para esta questão utilize a Tabela 1 apresentada na página a seguir. Sabendo-se que o tempo necessário para o atendimento de uma pessoa em um guichê de um banco tem distribuição Normal com média $\mu = 130$ segundos e desvio padrão $\sigma = 45$ segundos, pede-se:

a.(6%) Qual a probabilidade de que um indivíduo selecionado aleatoriamente requeira entre 2 e 3 minutos no guichê?

b.(8%) Dentro de que período de tempo os 20% de pessoas que realizam as transações mais rapidamente terminam seus negócios?

c.(6%) Qual é o tempo mínimo necessário para o atendimento dos 5% de indivíduos com as transações mais demoradas?

Tabela 1: Áreas de uma distribuição normal padrão entre $z=0$ e um valor positivo de z . As áreas para os valores de z negativos são obtidas por simetria.

z	0,00	0,01	0,02	0,03	0,04	0,05	0,06	0,07	0,08	0,09
0,0	0,0000	0,0040	0,0080	0,0120	0,0160	0,0199	0,0239	0,0279	0,0319	0,0359
0,1	0,0398	0,0438	0,0478	0,0517	0,0557	0,0596	0,0636	0,0675	0,0714	0,0753
0,2	0,0793	0,0832	0,0871	0,0910	0,0948	0,0987	0,1026	0,1064	0,1103	0,1141
0,3	0,1179	0,1217	0,1255	0,1293	0,1331	0,1368	0,1406	0,1443	0,1480	0,1517
0,4	0,1554	0,1591	0,1628	0,1664	0,1700	0,1736	0,1772	0,1808	0,1844	0,1879
0,5	0,1915	0,1950	0,1985	0,2019	0,2054	0,2088	0,2123	0,2157	0,2190	0,2224
0,6	0,2257	0,2291	0,2324	0,2357	0,2389	0,2422	0,2454	0,2486	0,2517	0,2549
0,7	0,2580	0,2611	0,2642	0,2673	0,2703	0,2734	0,2764	0,2794	0,2823	0,2852
0,8	0,2881	0,2910	0,2939	0,2967	0,2995	0,3023	0,3051	0,3078	0,3106	0,3133
0,9	0,3159	0,3186	0,3212	0,3238	0,3264	0,3289	0,3315	0,3340	0,3365	0,3389
1,0	0,3413	0,3438	0,3461	0,3485	0,3508	0,3531	0,3554	0,3577	0,3599	0,3621
1,1	0,3643	0,3665	0,3686	0,3708	0,3729	0,3749	0,3770	0,3790	0,3810	0,3830
1,2	0,3849	0,3869	0,3888	0,3907	0,3925	0,3944	0,3962	0,3980	0,3997	0,4015
1,3	0,4032	0,4049	0,4066	0,4082	0,4099	0,4115	0,4131	0,4147	0,4162	0,4177
1,4	0,4192	0,4207	0,4222	0,4236	0,4251	0,4265	0,4279	0,4292	0,4006	0,4319
1,5	0,4332	0,4345	0,4357	0,4370	0,4382	0,4394	0,4406	0,4418	0,4429	0,4441
1,6	0,4452	0,4463	0,4474	0,4484	0,4495	0,4505	0,4515	0,4525	0,4535	0,4545
1,7	0,4554	0,4564	0,4573	0,4582	0,4591	0,4599	0,4608	0,4616	0,4625	0,4633
1,8	0,4641	0,4649	0,4656	0,4664	0,4671	0,4678	0,4686	0,4693	0,4699	0,4706
1,9	0,4713	0,4719	0,4726	0,4732	0,4738	0,4744	0,4750	0,4756	0,4761	0,4767
2,0	0,4772	0,4778	0,4783	0,4788	0,4793	0,4798	0,4803	0,4808	0,4812	0,4817
2,1	0,4821	0,4826	0,4830	0,4834	0,4838	0,4842	0,4846	0,4850	0,4854	0,4857
2,2	0,4861	0,4864	0,4868	0,4871	0,4875	0,4878	0,4881	0,4884	0,4887	0,4890
2,3	0,4893	0,4896	0,4898	0,4901	0,4904	0,4906	0,4909	0,4911	0,4913	0,4916
2,4	0,4918	0,4920	0,4922	0,4925	0,4927	0,4929	0,4931	0,4932	0,4934	0,4936
2,5	0,4938	0,4940	0,4941	0,4943	0,4945	0,4946	0,4948	0,4949	0,4951	0,4952
2,6	0,4953	0,4955	0,4956	0,4957	0,4959	0,4960	0,4961	0,4962	0,4963	0,4964
2,7	0,4965	0,4966	0,4967	0,4968	0,4969	0,4970	0,4971	0,4972	0,4973	0,4974
2,8	0,4974	0,4975	0,4976	0,4977	0,4977	0,4978	0,4979	0,4979	0,4980	0,4981
2,9	0,4981	0,4982	0,4982	0,4983	0,4984	0,4984	0,4985	0,4985	0,4986	0,4986
3,0	0,4987	0,4987	0,4987	0,4988	0,4988	0,4989	0,4989	0,4989	0,4990	0,4990

Adaptada de Costa Neto, P. L. O. Estatística, Editora Edgard Blucher.

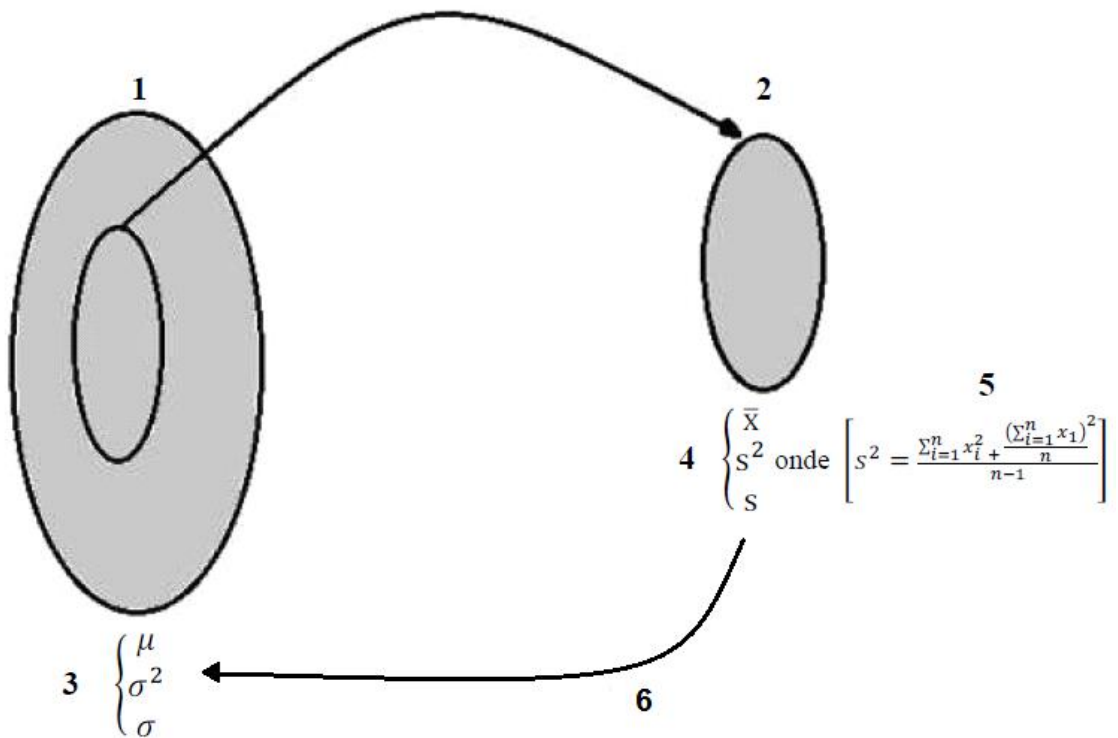
4 (20%). Dada $f(x) = x e^{-x^2}$, para $x \geq 0$ e, $f(x) = 0$ para outros valores de x , responda:

a.(12%) Verifique se $f(x)$ é uma função densidade de probabilidade (f.d.p.).

b.(8%) Caso $f(x)$ não seja uma f.d.p., verifique se é possível obter uma f.d.p. a partir de uma transformação em $f(x)$.

5 (20%). O esquema abaixo está relacionado a um teste de hipóteses, em que,

$$S^2 = \frac{SQD_X}{n-1}, \quad SQD_X = \sum_{i=1}^n x_i^2 - \frac{(\sum_{i=1}^n x_i)^2}{n}.$$



a.(8%) Denomine, **no esquema acima**, o que representa cada um dos pontos de 1 a 6.

b.(6%) Descreva com detalhes porque se aplica testes de hipóteses em estatística.

c.(6%) Explique como se toma uma decisão baseada em um teste de hipóteses e o que pode ocorrer com esta decisão tomada.

Universidade Federal de Viçosa
Departamento de Estatística

Prova seletiva para o Programa de Pós-Graduação em Estatística Aplicada e Biometria.

Nível Mestrado - 12/Junho/2015

Nome: _____

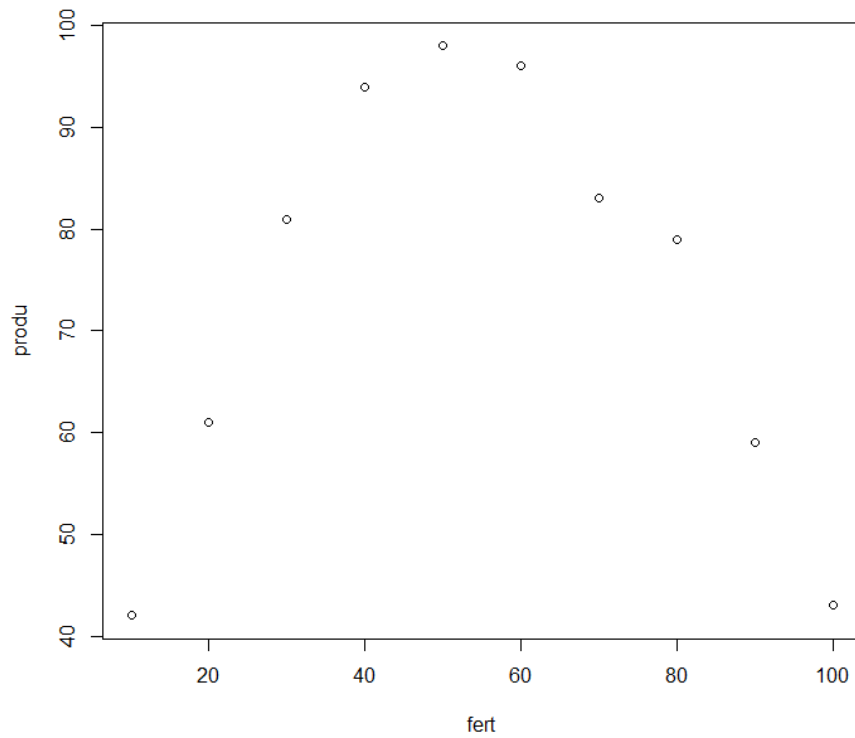
Assinatura: _____

Número do CPF ou RG: _____

- Interpretar corretamente as questões é parte da avaliação.
- Indique todos os cálculos organizadamente.
- São 5 (CINCO) questões em páginas numeradas de 1 a 9, total de 100 pontos.
- O tempo máximo para realização desta prova é de 4 horas.
- **BOA SORTE !!**

1 (20%). Considere que um médico está diante do seguinte dilema: “se ele estiver 80% confiante de que seu paciente esteja com a doença (evento D), então ele recomenda cirurgia; por outro lado, se ele não tiver esta confiança de que o paciente esteja doente, então ele recomenda exames adicionais que são caros e às vezes dolorosos”. Para o paciente *Fulano* o médico estava apenas 60% confiante de que ele estivesse doente e portanto pediu os exames adicionais. Estes exames adicionais sempre resultam em positivo (evento $+$) quando o paciente tem a doença, e, nunca resultam em positivo quando o paciente não tem a doença. Entretanto, se o paciente for diabético, então os exames resultam em 30% de resultados positivos para a doença, quando o paciente não está doente, ou seja, 30% de resultados falso positivo. No caso do paciente *Fulano* o resultado do teste foi $+$, portanto o médico iria recomendar cirurgia, quando pela primeira vez, *Fulano* informou ao médico que ele é diabético. Pede-se: Diante dessas informações, qual deverá ser a decisão do médico? Recomendar que *Fulano* seja submetido à cirurgia ou não? Responda com base no cálculo da probabilidade condicional apropriada.

2 (20%). Considere o diagrama de dispersão para produção de plantas (prod) e fertilidade do solo (fert). O diagrama de dispersão desse conjunto de dados sugere um modelo de ajuste do tipo $Y_i = \beta_0 + \beta_1 X_i + \beta_2 X_i^2 + \epsilon_i$, em que Y_i : “prod” e X_i : “fert”.



a) Este modelo de regressão é do tipo Regressão Linear Simples, Regressão Linear Múltipla, Regressão Não Linear? Justifique sua resposta.

b) A estimativa para o coeficiente de correlação linear de Pearson (r) entre as variáveis foi $-0,0436$. Interprete esse valor, contextualizando a resposta.

c) Geralmente, ouve-se falar que o coeficiente de determinação (R^2) é estimado utilizando-se o quadrado do coeficiente de correlação linear de Pearson (r^2). Para o modelo sugerido pelo diagrama de dispersão, o coeficiente de determinação pode ser estimado pelo quadrado do coeficiente de correlação linear de Pearson? Justifique sua resposta.

3 (20%). Defina:

a) Variável aleatória discreta.

b) Variável aleatória contínua.

c) Função densidade de probabilidade.

d) Função de probabilidade.

e) Considere a seguinte variável aleatória bidimensional discreta com a seguinte distribuição de probabilidade conjunta.

X	Y			
	1	2	4	5
2	0,20	0,10	0,10	0,20
3	0,10	0,10	0,10	0,10

Dado que $V(X) = E(X^2) - [E(X)]^2$ e que $Cov(X, Y) = E(XY) - E(X)E(Y)$, determine:

i) $E(X + Y)$.

ii) $V(2X - 3Y - 8)$.

4 (20%). Considere os dados abaixo, relativos a duas amostras, cada uma com 51 elementos, sendo:

X : peso de grãos de milho por espiga (g);

Y : peso de grãos de feijão por planta (g).

Dados:

$$\bar{X} = 51,45 \text{ g} \quad \bar{Y} = 42,30 \text{ g} \quad SQD_X = 3.200 \text{ g}^2 \quad SQD_Y = 2.450 \text{ g}^2$$

a) Qual das duas amostras é mais uniforme ou homogênea? Justifique sua resposta.

b) Qual das duas amostras apresentou estimativa de peso médio mais precisa? Justifique sua resposta.

5 (20%). Um professor de uma disciplina de Pós-Graduação com 90 estudantes matriculados, deseja avaliar os efeitos de seis metodologias de ensino em relação ao aprendizado dos conceitos de testes de hipóteses. Contudo, esses 90 estudantes são oriundos de seis diferentes cursos de Graduação (Agronomia, Matemática, Biologia, Engenharia Civil, Engenharia Florestal e Engenharia Agrícola).

Além disso, esses mesmos 90 estudantes pertencem a seis diferentes programas de Pós-Graduação (Estatística, Entomologia, Fitotecnia, Genética e Melhoramento, Zootecnia e Veterinária).

Baseado nessas informações o professor da disciplina resolveu instalar um experimento da seguinte forma:

- i) - Classificou cada estudante de acordo com o curso de Graduação e o programa de Pós-Graduação.
- ii) - Separou um conjunto de 36 estudantes de tal forma que havia um único estudante representante de cada uma das combinações possíveis entre curso de graduação e programa de pós-graduação
- iii)- Distribuiu ao acaso as seis metodologias de ensino aos 36 estudantes com as restrições de que cada metodologia fosse avaliada em cada grupo de curso de graduação e em cada grupo de programa de pós-graduação uma única vez.
- iv)- Após a aplicação das metodologias de ensino, os 36 estudantes foram então avaliados em uma mesma prova, recebendo uma nota na escala de 0 a 100.

Com base nestas informações responda:

- a) Qual(is) foi(ram) o(s) fator(es) em estudo nesse experimento? Justifique a sua resposta.

b) Quais foram os princípios básicos da experimentação utilizados? Justifique a sua resposta.

c) Qual foi o delineamento experimental utilizado? Justifique a sua resposta.

Universidade Federal de Viçosa
Departamento de Estatística

Prova seletiva para o Programa de Pós-Graduação em Estatística Aplicada e Biometria.

Nível Mestrado - 20/Novembro/2015

Nome: _____

Assinatura: _____

Número do CPF ou RG: _____

- Interpretar corretamente as questões é parte da avaliação.
- Indique todos os cálculos organizadamente.
- São SEIS (6) questões em páginas numeradas de 1 a 8, total de 100 pontos.
- O tempo máximo para realização desta prova é de 4 horas.
- **BOA SORTE !!**

Formulário

$$SQD_X = \sum_{i=1}^n X_i^2 - \frac{\left(\sum_{i=1}^n X_i\right)^2}{n} \quad \text{ou} \quad SQD_X = \sum_{i=1}^k f_i X_i^2 - \frac{\left(\sum_{i=1}^k f_i X_i\right)^2}{\sum_{i=1}^k f_i}$$

$$\hat{\rho}_{XY} = r_{XY} = \frac{SPD_{XY}}{\sqrt{SQD_X SQD_Y}} \quad SPD_{XY} = \sum_{i=1}^n X_i Y_i - \frac{\left(\sum_{i=1}^n X_i\right) \left(\sum_{i=1}^n Y_i\right)}{n}$$

$$P(A_i|B) = \frac{P(A_i \cap B)}{P(B)} = \frac{P(A_i)P(B|A_i)}{\sum_j P(A_j)P(B|A_j)}, \quad P(B) > 0$$

1) (20%). Suponha que um pesquisador deseje estudar os efeitos que a temperatura do ambiente de criação e a ração de engorda têm no peso final dos animais na época do abate. Para isto, o pesquisador, planeja realizar um experimento com três temperaturas e quatro rações. Em sua instituição de pesquisa, o pesquisador possui a sua disposição réplicas em tamanho reduzido de ambientes de criação utilizados comercialmente. As 15 réplicas foram construídas da mesma forma e, portanto, homogêneas entre si. Cada réplica possui quatro partes, tal que cada parte permite a criação de apenas um único animal. Portanto, na instalação do experimento, o pesquisador pretende distribuir inteiramente ao acaso as temperaturas aos ambientes de criação, de tal forma que cada temperatura seja designada a cinco réplicas. Posteriormente, as quatro rações são designadas ao acaso às partes que compõem cada réplica. Até o momento da mensuração do peso dos animais, cada réplica do ambiente será mantida na temperatura que foi destinada a receber e o animal acondicionado em determinada parte do ambiente receberá a ração designada àquela parte que se encontra. Com base nessas informações, responda as seguintes perguntas apresentando a respectiva justificativa.

a) Quais são os fatores em estudo neste experimento?

b) Existe fator perturbador neste experimento?

c) Qual é o delineamento experimental utilizado neste planejamento?

d) Qual foi o esquema experimental utilizado?

2) (10%). O quadro abaixo foi preparado para o estudo da relação linear entre as variáveis, o peso (Y), o número de cáries (X) até a idade atual e a altura (Z) de 8 meninos, considerando (Y) como variável resposta.

	Meninos							
	1	2	3	4	5	6	7	8
Peso (Y)	29	32	24	33	27	26	34	26
Cáries (X)	8	10	6	11	8	7	10	9
Altura (Z)	1,44	1,49	1,24	1,57	1,29	1,27	1,39	1,22

a) Calcule o coeficiente de correlação e determinação entre Y e X ;

b) Calcule o coeficiente de correlação e determinação entre Y e Z ;

c) Com base nos conhecimentos de regressão linear simples, qual das variáveis explicativas você escolheria para o ajuste do modelo? Justifique.

3) (10%). Considere a realização dos testes de comparação múltipla, (TUKEY, DUNCAN e DUNNET), e responda:

a) Qual a hipótese que está sendo testada em cada um deles?

b) Qual a diferença entre os testes?

4) (20%). Considere a seguinte função densidade de probabilidade conjunta

$$f(x) = \begin{cases} \frac{1}{8}x(x-y), & 0 < x < 2, \quad -x < y < x \\ 0, & \text{caso contrário.} \end{cases}$$

Determine $P(0 \leq X \leq 1)$.

5) (20%) Sabe-se que 1% das mulheres com 40 anos de idade têm câncer de mama. Por outro lado, a mamografia indica câncer de mama em 9,6% das mulheres com 40 anos de idade e que não têm câncer de mama (resultado falso +); enquanto que 20% das mulheres com 40 anos de idade e que têm câncer de mama, não são diagnosticadas na mamografia (resultado falso -). Pede-se:

a.(10%)Qual é a probabilidade de uma mulher com 40 anos de idade receber um resultado negativo (-) de mamografia?

b.(10%)Se uma mulher com 40 anos de idade receber um resultado positivo (+) de mamografia, indicando que ela tem câncer de mama, qual é a probabilidade dela realmente estar com câncer de mama?

6) (20%). Considere os dados apresentados a seguir, relativos a um experimento instalado no delineamento em blocos ao acaso, com 4 repetições, em esquema fatorial para avaliar o efeito de 3 níveis do fator A e 2 níveis do fator B .

Tabela 1: Totais para tratamentos

(4)	B_1	B_2	Totais (8)
A_1	60	80	140
A_2	85	100	185
A_3	100	110	210
Totais (12)	245	290	535 (24)

$$SQ_{\text{Total}} = 688,17 \text{ e } SQ_{\text{Blocos}} = 243,12.$$

a) (10%). Execute a ANOVA para testar a significância do efeito da interação entre os fatores. Considere o $F_{\text{Tabelado}} = 3,68$ e utilize duas casas decimais nos cálculos.

b) (5%). Em um esquema fatorial, explique como deve ser feito o estudo dos fatores A e B nas duas situações abaixo:

Situação 1) Interação significativa

Situação 2) Interação não significativa

c) (5%) Considerando a avaliação de dois fatores, A_i para $i = 1, 2, \dots, I$ e B_j para $j = 1, 2, \dots, J$, apresente as diferenças entre os esquemas fatorial e parcelas subdivididas, quanto à casualização dos fatores nas unidades experimentais.

Universidade Federal de Viçosa
Departamento de Estatística

Prova seletiva para o Programa de Pós-Graduação em Estatística Aplicada e Biometria.

Nível Mestrado - 7/Novembro/2016

Nome: _____

Assinatura: _____

Número do CPF ou RG: _____

- Interpretar corretamente as questões é parte da avaliação.
- Indique todos os cálculos organizadamente.
- São SEIS (6) questões em páginas numeradas de 1 a 10, total de 100 pontos.
- Nas páginas 9 e 10 temos formulário e tabela.
- O tempo máximo para realização desta prova é de 4 horas.
- **BOA SORTE !!**



1) (20%) Um dado não viciado é lançado duas vezes de forma independente. Sejam X e Y variáveis aleatórias que representem, respectivamente, os números obtidos no 1º e 2º lançamento. Pede-se:

a) (4%) $P(X = Y)$;

b) (7%) Considerando $W = |X - Y|$, apresente a distribuição de probabilidade para W e informe qual a probabilidade de W ser igual a 3.



c) (9%) Considerando as informações da letra **b)**, pede-se:

i) (3%) O que é variável aleatória? Qual é a variável aleatória do item **b)**?

ii) (3%) O que é espaço amostral e qual é o espaço amostral relacionado ao problema?

iii) (3%) O que é evento? Qual é o evento de interesse no item **b)**? Qual definição de probabilidade pode ser utilizada para a obtenção da probabilidade de interesse?

2) (20%) Um pesquisador deseja verificar os efeitos de dois tipos de comunicação persuasiva (a_1 = unilateral com apenas argumentos favoráveis ao objetivo da comunicação e a_2 = bilateral com argumentos prós e contras ao objetivo da comunicação). Tendo desconfiado que o nível intelectual da audiência ao qual a comunicação persuasiva se destina influi nos resultados, ele decidiu introduzir esse fator no estudo, composto de pessoas de baixo nível (b_1) e de alto nível intelectual (b_2). Nesse caso, tem-se um experimento fatorial 2×2 , em que a variável-resposta Y é o grau de mudança da atitude verificada no sentido da comunicação persuasiva. Cada um dos quatro tratamentos foram avaliados por cinco indivíduos homogêneos, cujo experimento foi instalado sob o delineamento inteiramente casualizado (DIC). De acordo com os resultados e considerando-se satisfeitas as pressuposições da análise de variância, qual(is) é(são) o(s) melhor(es) tipo(s) de comunicação(ões) persuasiva(s) (a_1 e, ou, a_2), ou seja, aquele(s) que proporciona(m) maior(es) média(s) paramétrica(s) de Y , para pessoas de baixo (b_1) e alto (b_2) níveis intelectuais, separadamente ($\alpha = 0,05$)? **Por que?**

	b_1					b_2				
a_1	9	9	8	8	9	6	5	5	6	5
a_2	6	4	5	6	4	3	3	2	2	3

	Totais			Médias			
	b_1	b_2		b_1	b_2		
a_1	43	27	70	a_1	8,6	5,4	7,0
a_2	25	13	38	a_2	5,0	2,6	3,8
	68	40	108		6,8	4,0	5,4



3) (15%) Um engenheiro precisa planejar uma viga de concreto armado de um edifício. O concreto armado é obtido pela adição de um tipo de concreto a uma estrutura armada de aço. Existem três tipos de concreto e quatro tipos de estrutura armada que podem, a princípio, serem utilizadas para a construção da viga. Contudo, ele deseja, selecionar o tipo de concreto e a estrutura armada que proporcionem maior resistência à tração. Com essa finalidade, ele planeja construir 2 corpos de prova para cada uma das doze combinações possíveis entre tipos de concreto e de estrutura armada. Esses corpos de prova serão homogêneos, pois possuirão composições idênticas, exceto pela combinação, tipo de concreto e de estrutura armada, que serão inteiramente ao acaso designados a receber. Após um período de maturação do concreto armado, cada corpo de prova será submetido a um teste para mensurar a sua resistência à tração. Responda e justifique cada uma das perguntas abaixo

a) (2%) Qual(is) é(são) o(s) fator(es) em estudo neste planejamento?

b) (2%) Quais são os tratamentos nesse planejamento?

c) (2%) Qual é a unidade experimental nesse planejamento?

d) (2%) Existe fator perturbador neste planejamento?

e) (3%) Quais princípios básicos da experimentação constam no planejamento experimental?

f) (2%) Qual é o delineamento experimental utilizado neste planejamento?

g) (2%) Qual é o esquema experimental utilizado neste planejamento?

--

Nota

4) (10%) Para cada uma das estatísticas descritivas abaixo, apresente o intervalo de valores que cada uma pode assumir em \mathbb{R} .

a) (1%) Média aritmética;

b) (1%) Variância;

c) (2%) Coeficiente de variação;

d) (2%) Coeficiente de correlação.

Interprete em e) e f) os seguintes resultados:

e) (2%) “O coeficiente de correlação entre peso e produção é igual a zero”;

f) (2%) “A amplitude total e a variância de um conjunto de valores são iguais a zero”.

5) (20%) Digamos que o preço de um modelo de motocicleta usada está linearmente relacionado ao último dígito do ano de fabricação (para 93 utilizou-se 3, para 94 utilizou-se 4, etc.). A tabela a seguir apresenta os valores do preço (em milhares de reais, $R\$ \times 1000$) e o respectivo ano de fabricação.

Moto (i)	1	2	3	4	5	6
Ano	3	4	5	7	8	9
Preço	6,3	7,0	8,2	9,0	10,5	12,0

Através do método dos mínimos quadrados para a estimação dos parâmetros da Regressão Linear Simples entre as variáveis acima, responda:

a) (7%) estime e **interprete** β_1 ;

b) (7%) estime e **interprete** β_0 ;

c) (6%) estime β_0 e β_1 utilizando a notação matricial, sabendo-se que $\hat{\beta} = (X^t X)^{-1} X^t Y$,

$$(X^t X)^{-1} = \begin{bmatrix} 1,5 & -0,2 \\ -0,2 & 0,04 \end{bmatrix} \text{ e } X = \begin{bmatrix} 1 & 3 \\ 1 & 4 \\ \vdots & \vdots \\ 1 & 9 \end{bmatrix}.$$



Nota

6) (15%) Sabe-se que existem quatro tipos sanguíneos, a saber: **A**, **B**, **AB** e **O**, além disso existem dois fatores Rh que cada um dos tipos sanguíneos pode ter: o positivo (**Rh**⁺) e o negativo (**Rh**⁻). Uma pessoa tem um único tipo sanguíneo e um único fator Rh.

Um grupo de 200 pessoas possui a seguinte composição quanto ao grupo sanguíneo: 55 são do tipo **A**, 45 são do tipo **B**, 70 são do tipo **AB** e 30 do tipo **O**. Adicionalmente, sabe-se que 15 destas pessoas do grupo sanguíneo **A**, 10 do grupo sanguíneo **B**, 20 do grupo sanguíneo **AB** e 5 do grupo sanguíneo **O**, são pessoas com fator **Rh**⁻.

Pede-se: utilize a regra ou teorema de Bayes para calcular a probabilidade condicional de se selecionar aleatoriamente do grupo de 200 pessoas, uma pessoa com tipo sanguíneo **A** ou **B**, dado que a pessoa selecionada é fator **Rh**⁻.



Formulário

$SPD_{XY} = \sum_{i=1}^n X_i Y_i - \frac{\left(\sum_{i=1}^n X_i\right) \left(\sum_{i=1}^n Y_i\right)}{n}$	$\hat{Y}_i = \hat{\beta}_0 + \hat{\beta}_1 X_i$
$SQD_X = \sum_{i=1}^n X_i^2 - \frac{\left(\sum_{i=1}^n X_i\right)^2}{n}$	$\hat{\beta}_1 = \frac{SPD_{XY}}{SQD_X}$
$\hat{\beta}_0 = \bar{Y} - \hat{\beta}_1 \bar{X}$	$P[B_j A] = \frac{P[A B_j]P[B_j]}{\sum_{i=1}^n P[A B_i]P[B_i]}$
$P[A B] = \frac{P[A \cap B]}{P[B]}$	$P[A \cup B] = P[A] + P[B] - P[A \cap B]$

Tabela 1: Valores de F ao nível de 5% de probabilidade. $P(F_{n_1, n_2} \geq F) = 0,05$. Número de graus de liberdade: n_1 = numerador e n_2 = denominador

n_2	n_1																						
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	20	24	30	40	60	120	∞
1	161,4	199,5	215,7	224,6	230,2	234,0	236,8	238,9	240,5	241,9	243,0	243,9	244,4	245,0	245,9	246,0	248,0	249,1	250,1	251,1	252,2	253,3	254,3
2	18,51	19,00	19,16	19,25	19,30	19,33	19,35	19,37	19,38	19,40	19,40	19,41	19,42	19,42	19,43	19,43	19,45	19,45	19,46	19,47	19,48	19,49	19,50
3	10,13	9,55	9,28	9,12	9,01	8,94	8,89	8,85	8,81	8,79	8,76	8,74	8,72	8,71	8,70	8,69	8,66	8,64	8,62	9,59	8,57	8,55	8,53
4	7,71	6,94	6,59	6,39	6,26	6,16	6,09	6,04	6,00	5,96	5,93	5,91	5,89	5,87	5,86	5,84	5,80	5,77	5,75	5,72	5,69	5,66	5,63
5	6,61	5,79	5,41	5,19	5,05	4,95	4,88	4,82	4,77	4,74	4,70	4,68	4,66	4,64	4,62	4,60	4,56	4,53	4,50	4,46	4,43	4,40	4,36
6	5,99	5,14	4,76	4,53	4,39	4,28	4,21	4,15	4,10	4,06	4,03	4,00	3,98	3,96	3,94	3,92	3,87	3,84	3,81	3,77	3,74	3,70	3,67
7	5,59	4,74	4,35	4,12	3,97	3,87	3,79	3,73	3,68	3,64	3,60	3,57	3,55	3,52	3,51	3,49	3,44	3,41	3,38	3,34	3,30	3,27	3,23
8	5,32	4,46	4,07	3,84	3,69	3,58	3,50	3,44	3,39	3,35	3,31	3,28	3,25	3,23	3,22	3,20	3,15	3,12	3,08	3,04	3,01	2,97	2,93
9	5,12	4,26	3,86	3,63	3,48	3,37	3,29	3,23	3,18	3,14	3,10	3,07	3,04	3,02	3,01	2,98	2,94	2,90	2,86	2,83	2,79	2,75	2,71
10	4,96	4,10	3,71	3,48	3,33	3,22	3,14	3,07	3,02	2,98	2,94	2,91	2,88	2,86	2,85	2,82	2,77	2,74	2,70	2,66	2,62	2,58	2,54
11	4,84	3,98	3,59	3,36	3,20	3,09	3,01	2,95	2,90	2,85	2,82	2,79	2,76	2,74	2,72	2,70	2,65	2,61	2,57	2,53	2,49	2,45	2,40
12	4,75	3,89	3,49	3,26	3,11	3,00	2,91	2,85	2,80	2,75	2,72	2,69	2,66	2,64	2,62	2,60	2,54	2,51	2,47	2,43	2,38	2,34	2,30
13	4,67	3,81	3,41	3,18	3,03	2,92	2,83	2,77	2,71	2,67	2,63	2,60	2,57	2,55	2,53	2,51	2,46	2,42	2,38	2,34	2,30	2,25	2,21
14	4,60	3,74	3,34	3,11	2,96	2,85	2,76	2,70	2,65	2,60	2,56	2,53	2,50	2,48	2,46	2,44	2,39	2,35	2,31	2,27	2,22	2,18	2,13
15	4,54	3,68	3,29	3,06	2,90	2,79	2,71	2,64	2,59	2,54	2,51	2,48	2,45	2,43	2,40	2,39	2,33	2,29	2,25	2,20	2,16	2,11	2,07
16	4,49	3,63	3,24	3,01	2,85	2,74	2,66	2,59	2,54	2,49	2,45	2,42	2,39	2,37	2,35	2,33	2,28	2,24	2,19	2,15	2,11	2,06	2,01
17	4,45	3,59	3,20	2,96	2,81	2,70	2,61	2,55	2,49	2,45	2,41	2,38	2,35	2,33	2,31	2,29	2,23	2,19	2,15	2,10	2,06	2,01	1,96
18	4,41	3,55	3,16	2,93	2,77	2,66	2,58	2,51	2,46	2,41	2,37	2,34	2,31	2,29	2,27	2,25	2,19	2,15	2,11	2,06	2,02	1,97	1,92
19	4,38	3,52	3,13	2,90	2,74	2,63	2,54	2,48	2,42	2,38	2,34	2,31	2,28	2,26	2,23	2,21	2,16	2,11	2,07	2,03	1,98	1,93	1,88
20	4,35	3,49	3,10	2,87	2,71	2,60	2,51	2,45	2,39	2,35	2,31	2,28	2,25	2,23	2,20	2,18	2,12	2,08	2,04	1,99	1,95	1,90	1,84
21	4,32	3,47	3,07	2,84	2,68	2,57	2,49	2,42	2,37	2,32	2,28	2,25	2,22	2,20	2,18	2,15	2,10	2,05	2,01	1,96	1,92	1,87	1,81
22	4,30	3,44	3,05	2,82	2,66	2,55	2,46	2,40	2,34	2,30	2,26	2,23	2,20	2,18	2,15	2,13	2,07	2,03	1,98	1,94	1,89	1,84	1,78
23	4,28	3,42	3,03	2,80	2,64	2,53	2,44	2,37	2,32	2,27	2,24	2,20	2,17	2,14	2,13	2,10	2,05	2,01	1,96	1,91	1,86	1,81	1,76
24	4,26	3,40	3,01	2,78	2,62	2,51	2,42	2,36	2,30	2,25	2,22	2,18	2,15	2,13	2,11	2,09	2,03	1,98	1,94	1,89	1,84	1,79	1,73
25	4,24	3,39	2,99	2,76	2,60	2,49	2,40	2,34	2,28	2,24	2,20	2,16	2,13	2,11	2,09	2,06	2,01	1,96	1,92	1,87	1,82	1,77	1,71
26	4,23	3,37	2,98	2,74	2,59	2,47	2,39	2,32	2,27	2,22	2,18	2,15	2,12	2,10	2,07	2,05	1,99	1,95	1,90	1,85	1,80	1,75	1,69
27	4,21	3,35	2,96	2,73	2,57	2,46	2,37	2,31	2,25	2,20	2,16	2,13	2,10	2,08	2,06	2,03	1,97	1,93	1,88	1,84	1,79	1,73	1,67
28	4,20	3,34	2,95	2,71	2,56	2,45	2,36	2,29	2,24	2,19	2,15	2,12	2,09	2,06	2,04	2,02	1,96	1,91	1,87	1,82	1,77	1,71	1,65
29	4,18	3,33	2,93	2,70	2,55	2,43	2,35	2,28	2,22	2,18	2,14	2,10	2,07	2,05	2,03	2,00	1,94	1,90	1,85	1,81	1,75	1,70	1,64
30	4,17	3,32	2,92	2,69	2,53	2,42	2,33	2,27	2,21	2,16	2,12	2,09	2,06	2,04	2,01	1,99	1,93	1,89	1,84	1,79	1,74	1,68	1,62
40	4,08	3,23	2,84	2,61	2,45	2,34	2,25	2,18	2,12	2,08	2,04	2,00	1,97	1,95	1,92	1,90	1,84	1,79	1,74	1,69	1,64	1,58	1,51
60	4,00	3,15	2,76	2,53	2,37	2,25	2,17	2,10	2,04	1,99	1,95	1,92	1,89	1,86	1,84	1,81	1,75	1,70	1,65	1,59	1,53	1,47	1,39
120	3,92	3,07	2,68	2,45	2,29	2,17	2,09	2,02	1,96	1,91	1,86	1,83	1,80	1,77	1,75	1,73	1,66	1,61	1,55	1,50	1,43	1,35	1,25
∞	3,84	3,00	2,60	2,37	2,21	2,10	2,01	1,94	1,88	1,83	1,79	1,75	1,72	1,69	1,67	1,64	1,57	1,52	1,46	1,39	1,32	1,22	1,00

Adaptada de Bussab, W. O. e Morettin, P. A. Estatística Básica - Métodos Quantitativos, Editora Atual.

Universidade Federal de Viçosa
Departamento de Estatística

Prova seletiva para o Programa de Pós-Graduação em Estatística Aplicada e Biometria.

Nível Mestrado - 31/Maio/2017

Nome: _____

Assinatura: _____

Número do CPF ou RG: _____

- Interpretar corretamente as questões é parte da avaliação.
- Indique todos os cálculos organizadamente.
- São SEIS (6) questões em páginas numeradas de 1 a 9, total de 100 pontos.
- Na página 9 temos formulário.
- O tempo máximo para realização desta prova é de 4 horas.
- **BOA SORTE !!**



1) (20%) Dados X_1, X_2, \dots, X_n os n valores amostrais de uma variável X . Ao multiplicar cada elemento por uma constante k , sendo $k \in \mathbb{R}$, o que acontece com cada uma das medidas amostrais abaixo: **(Justifique sua resposta)**

a) Média

b) Variância

c) Desvio padrão

d) Coeficiente de variação

2) (20%) A tabela abaixo apresenta dados relativos ao número de manutenções realizadas, em um determinado mês, em 6 máquinas de uma indústria e o número de peças produzidas com defeito por estas máquinas. Suponha que o número de peças produzidas com defeito por uma máquina esteja linearmente relacionado ao número de manutenções realizadas.

Máquina (i)	1	2	3	4	5	6
Número de manutenções	0	3	6	7	9	10
Nº de peças com defeito	20	15	13	11	10	8

a) Ajuste e apresente uma equação de regressão linear simples para explicar o número de peças produzidas com defeito em função do número de manutenções realizadas nas máquinas. Faça as devidas interpretações.

b) Calcule e interprete o coeficiente de determinação.



Nota

c) Qual é o número esperado de peças produzidas com defeito por máquinas nas quais foram realizadas 7 e 11 manutenções, respectivamente? Comente sobre estas estimativas.

d) Forneça o desvio da regressão para 7 manutenções da máquina.



3) (15%) Um sistema eletrônico possui quatro componentes designados 1, 2, 3 e 4. Quando este sistema é utilizado por um período de tempo, poderá ocorrer uma falha deste sistema **se o componente 1 falhar, ou, se pelo menos dois dos componentes 2, 3 e 4 falharem**. Seja A_i o seguinte evento,

$A_i = \{\text{o } i\text{-ésimo componente falha durante o período de utilização do sistema eletrônico}\}.$

Admita que os eventos A_1, A_2, A_3 e A_4 sejam mutuamente independentes com $P(A_i) = 0,05$ para $i = 1, 2, 3, 4$.

Pede-se: calcule a probabilidade de ocorrer uma falha neste sistema eletrônico durante o período de utilização.



4) (10%) Uma pesquisa na área de tecnologia de alimentos tem o objetivo de verificar os efeitos de dois fatores (A = método de fabricação e B = tipo de sangue do suíno) sobre a cor da mortadela (Y), medida numa escala de 1 a 110 pontos. Os quatro níveis de A foram casualizados às parcelas sob o delineamento em blocos casualizados com duas repetições e, os três níveis de B , foram distribuídos inteiramente ao acaso às subparcelas. Sabe-se que quanto maiores forem a coloração e o custo, menos comercial será a mortadela. De acordo com os resultados dos testes F da ANOVA e do Tukey, se necessário, e considerando-se satisfeitas as pressuposições da ANOVA, qual é o melhor tratamento ($\alpha = 0,05$)? Por quê?

	a_1	a_2	a_3	a_4	b_1	b_2	b_3
Custo	5,23	5,06	4,85	4,37	1,23	1,06	0,85

	b_1	b_2	b_3	
a_1	61,5	83,5	92,0	79,00
a_2	83,5	88,5	95,0	89,00
a_3	84,0	86,0	94,0	88,00
a_4	64,0	66,0	71,0	67,00
	73,25	81,00	88,00	80,75

FV	GL	SQ	QM	f_{cal}	f_{tab}
Bloco		1.176,0			10,13
A					9,28
Res(a)		1.053,3			
B					4,46
AxB					3,58
Res(b)		572,7			
Total		5966,5			
		$q_A = 6,83$			$q_B = 4,04$



Nota

5) (20%) A ANOVA parcial oriunda dos dados de um experimento instalado segundo o delineamento em blocos casualizados é apresentada na Tabela 1.

Tabela 1: Anova parcial

FV	GL
Blocos	
Tratamentos	3
Resíduo	12
Total	

Suponha que ao instalar esse experimento o pesquisador tenha utilizado corretamente os princípios básicos da experimentação necessários para a instalação do mesmo.

Com base nessas informações, responda e JUSTIFIQUE cada uma das perguntas a seguir.

a) Quantas repetições foram utilizadas nesse experimento?

b) Quais princípios básicos da experimentação foram utilizados nesse experimento?

c) Qual o modelo estatístico para realizar a ANOVA para esse experimento?



Nota

6) (15%) Seja X_1, X_2, \dots, X_n , uma amostra aleatória de uma variável aleatória com distribuição de probabilidade $N(\mu, \sigma^2)$. Pede-se:

a) Qual a distribuição de probabilidade da média aritmética? Por que?

b) Qual a distribuição de probabilidade da variável aleatória $Z = \frac{\bar{X} - \mu}{\sqrt{\sigma^2/n}}$? Por que?



Formulário

$\bar{X} = \frac{\sum_{i=1}^n X_i}{n}$	$S_X^2 = \frac{\sum_{i=1}^n (X_i - \bar{X})^2}{n-1} = \frac{\sum_{i=1}^n X_i^2 - \frac{\left(\sum_{i=1}^n X_i\right)^2}{n}}{n-1}$	
$S(X) = \sqrt{\widehat{V}(X)}$	$CV(X)\% = \frac{S(X)}{\bar{X}} \times 100\%$	$SQD_X = \sum_{i=1}^n X_i^2 - \frac{\left(\sum_{i=1}^n X_i\right)^2}{n}$
$SPD_{XY} = \sum_{i=1}^n X_i Y_i - \frac{\left(\sum_{i=1}^n X_i\right) \left(\sum_{i=1}^n Y_i\right)}{n}$		$r_{XY} = \frac{SPD_{XY}}{\sqrt{SQD_X \times SQD_Y}}$
$\widehat{Y}_i = \widehat{\beta}_0 + \widehat{\beta}_1 X_i$	$\widehat{\beta}_1 = \frac{SPD_{XY}}{SQD_X}$	$\widehat{\beta}_0 = \bar{Y} - \widehat{\beta}_1 \bar{X}$
$SQ_{\text{Regressão}} = \frac{(SPD_{XY})^2}{SQD_X} = \widehat{\beta}_1^2 SQD_X = \widehat{\beta}_1 SPD_{XY}$		$SQ_{\text{Total}} = SQD_Y$
$\widehat{\epsilon}_i = Y_i - \widehat{Y}_i$	$r^2(\%) = \frac{SQ_{\text{Regressão}}}{SQ_{\text{Total}}} \times 100\%$	

Universidade Federal de Viçosa
Departamento de Estatística

Prova seletiva para o Programa de Pós-Graduação em Estatística Aplicada e Biometria.

Nível Mestrado - 8/Novembro/2017

Número de inscrição: _____

- Interpretar corretamente as questões é parte da avaliação.
- Indique todos os cálculos organizadamente.
- É permitida a utilização de calculadora científica;
- São SEIS (6) questões em páginas numeradas de 1 a 10, total de 100 pontos.
- O tempo máximo para realização desta prova é de 4 horas.
- **BOA SORTE !!**

1) (20%) Um horticultor realizou um experimento para avaliar o efeito do tipo de material (madeira ou plástico) e volume (30, 40, 50, 60 e 70 litros) da embalagem em relação a danificação dos frutos de tomates colhidos em sua propriedade e transportado para uma feira livre. Uma amostra de frutos de tomates homogêneos e sem danos foi separada para realizar o experimento. Ele então encheu duas caixas de cada uma das combinações possíveis entre material e volume da embalagem com frutos de tomate. Como nenhum dos seus dois caminhões cabia as vinte embalagens e também não estavam com suspensões de amortecimento em iguais condições, ele decidiu que cada um dos caminhões transportaria uma embalagem de cada uma das combinações possíveis entre o material e volume da embalagem. Vale lembrar que condições insatisfatórias da suspensão de amortecimento podem danificar o fruto de tomate. Após o percurso, o número de frutos de tomates danificados foi obtido para cada embalagem.

Com base nessas informações, responda e **JUSTIFIQUE** cada uma das perguntas a seguir.

a) (04%) Qual(is) foi(ram) o(s) fator(es) em estudo?

b) (04%) Qual foi o número de repetições?

c) (04%) Qual foi o delineamento experimental utilizado?

d) (04%) Qual foi o esquema experimental utilizado?

e) (04%) Qual foi a variável resposta?



Nota

2) (20%) Dois jogadores A e B disputam um jogo de azar. O jogo consiste no lançamento de um dado não-simétrico de seis lados pelo Juiz, sendo que o jogador A vence a rodada se sair um número menor ou igual a 3 e o jogador B vence a rodada se sair um número maior que 3 na face superior do dado. O vencedor do jogo é aquele que primeiro vencer seis rodadas. Considere que a probabilidade de sair qualquer número par $\{2, 4, 6\}$ é $2p$ e de sair qualquer número ímpar $\{1, 3, 5\}$ é p . Pede-se:

a) (05%) Encontre o valor de p .

b) (05%) Calcule a probabilidade de cada jogador vencer uma única rodada.



Nota

c) (10%) Se o jogo fosse interrompido quando o jogador A havia vencido 5 rodadas e o jogador B 3 rodadas, calcule a probabilidade de cada jogador vencer o jogo.

3) (15%) Responda as seguintes questões justificando sua resposta.

a) (03%) A média aritmética pode ser zero? Pode ser negativa? Justifique sua resposta.

b) (03%) A mediana pode ser zero? Pode ser negativa? Justifique sua resposta.

c) (03%) A amplitude total pode ser zero? Pode ser negativa? Justifique sua resposta.

d) (03%) O desvio padrão pode ser zero? Pode ser negativo? Justifique sua resposta.

e) (03%) Suponha que calculando-se o coeficiente de correlação amostral obteve-se $r_{XY} = 0$. Dê a interpretação deste resultado.

4) (10%) A função densidade de probabilidade da variável aleatória X é

$$f(x) = \begin{cases} h + \frac{h}{k}(x - a), & \text{se } a - k < x < a \\ h - \frac{h}{k}(x - a), & \text{se } a \leq x < a + k. \\ 0, & \text{caso contrário} \end{cases}$$

a) (05%) Expresse h em função de a e k para que f seja realmente uma função densidade de probabilidade.

b) (05%) Determine o primeiro quartil (q_1) em função de a e k .

Dica: $F(q_1) = 1/4$, sendo $F(\cdot)$ a função de distribuição acumulada.

5) (20%) Uma empresa forneceu um curso para treinamento de 6 novos funcionários e três meses depois mediu a produtividade dos mesmos. Seja X o número de horas de curso (conforme tabela abaixo) e Y a produtividade em unidades produzidas por dia.

i	1	2	3	4	5	6
X_i	22	25	15	19	22	18

Neste estudo, um dos objetivos era avaliar o efeito do treinamento sobre a produtividade dos novos funcionários, a partir do ajuste da equação de regressão linear simples abaixo:

$$Y_i = \beta_0 + \beta_1 X_i + \varepsilon_i.$$

Pede-se:

a) (10%) Obtenha os estimadores de Mínimos Quadrados Ordinários dos parâmetros do modelo (β_0 e β_1).

b) (05%) Após o ajuste pelo método dos mínimos quadrados, as produtividades estimadas para os funcionários 1 e 6 foram, $\hat{Y}_1 = 37,60$ e $\hat{Y}_6 = 32,40$, respectivamente. Com base nessas informações, quais foram as estimativas dos parâmetros do modelo? Interprete-as.

c) (05%) Estime a produtividade média de funcionários que participaram de 20 e 25 horas de curso. Comente sobre as suas estimativas.

6) (15%) Um engenheiro agrônomo quer testar a seguinte hipótese: plantas de milho irrigadas com menor quantidade de água e adubadas com maior quantidade de fertilizante produzem mais. Para isso, ele realizou um experimento em parcelas subdivididas, em que as quantidades de água (A) ($a_1 =$ baixa e $a_2 =$ alta) foram casualizadas às parcelas sob o DBC com quatro repetições (c_1, c_2, c_3 e c_4), e as quantidades de fertilizante (B) ($b_1 =$ baixa e $b_2 =$ alta) foram distribuídas inteiramente ao acaso às subparcelas dentro de cada parcela. No total, foram obtidas 16 unidades experimentais. De acordo com os resultados e considerando-se satisfeitas as pressuposições da análise de variância, o agrônomo conseguiu provar a sua hipótese? Se necessário, aplique o teste de Tukey ($\alpha = 0,05$). Se sim, por quê? Se não, qual deve ser a recomendação para aumentar a produção (Y) e por quê?

A	B	Bloco	Y
1	1	1	25,4
1	1	2	32,9
1	1	3	26,1
1	1	4	35,2
1	2	1	25,6
1	2	2	34,7
1	2	3	25,9
1	2	4	35,8
2	1	1	14,6
2	1	2	25,2
2	1	3	14,3
2	1	4	25,5
2	2	1	15,5
2	2	2	24,7
2	2	3	15,6
2	2	4	24,8

Quadros de Totais

	b_1	b_2	Bloco	
a_1	119,6	122,0	c_1	81,1
a_2	79,6	80,6	c_2	117,5
	199,2	202,6	c_3	81,9
			c_4	121,3
			401,8	



FV	GL	SQ	QM	f_{cal}	f_{tab}
Bloco					9,28
A					10,13
Res(a)					
Parcela		777,7375			
B					5,99
A×B					5,99
Res(b)					
Total		781,1975			
		$q_{tab}(2; 3) = 4,50$			$q_{tab}(2; 6) = 3,46$

Universidade Federal de Viçosa
Departamento de Estatística

Prova seletiva para o Programa de Pós-Graduação em Estatística Aplicada e Biometria.

Nível Mestrado - 11/junho/2018

Nome: _____

Assinatura: _____

Número do CPF ou RG: _____

- Interpretar corretamente as questões é parte da avaliação.
- Indique todos os cálculos organizadamente.
- É permitida a utilização de calculadora;
- São SEIS (6) questões em páginas numeradas de 1 a 13, total de 100 pontos.
- Nas páginas 11, 12 e 13 temos formulário e tabelas.
- O tempo máximo para realização desta prova é de 4 horas.
- **BOA SORTE !!**



1) (20%) Faça o planejamento de um experimento considerando o que se descreve abaixo.

O problema: Uma empresa que produz alimentos para vacas leiteiras desenvolveu 5 novos tipos de rações (A, B, C, D, E). Porém, para que essas novas rações possam ser colocadas no mercado é preciso que elas sejam testadas nos animais. A empresa possui recursos para comercializar apenas três rações, preferencialmente aquelas que proporcionarem maior produção de leite.

As condições: O técnico responsável observou que para a realização da pesquisa, existiam na empresa poucos animais de mesma raça e além disso, eles não tinham idades semelhantes dentro da mesma raça. Entretanto, foi verificado que poder-se-ia agrupar animais de diferentes raças com a mesma idade.

Seu trabalho: Planeje o experimento, de forma que o técnico possa avaliar a produção de leite, sem que o efeito de raça e idade dos animais influencie de forma significativa nos resultados. Detalhe o experimento descrevendo:

a) (5%) Qual o delineamento experimental deve ser utilizado e qual o seu modelo estatístico.

b) (5%) Qual deve ser o número de repetições e qual é a unidade experimental a ser utilizada no experimento.

c) (5%) Mostre como deve ser feita a análise estatística do delineamento utilizado e a indicação das rações a serem comercializadas. Utilize tabela, gráficos, etc.

d) (5%) Forneça um croqui do experimento.

2) (20%) A umidade relativa do ar de um determinado local interfere diretamente na qualidade de vida dos habitantes. Baixas umidades podem gerar problemas respiratórios, sangramentos nasais e desidratação. Altas umidades podem provocar tonturas e proliferação de fungos. Um dos principais fatores que afetam a umidade é a temperatura. Estudos meteorológicos foram realizados nas cidades de Viçosa-MG e Belo Horizonte-MG durante 15 e 20 dias, respectivamente, obtendo os seguintes resultados para a temperatura ($^{\circ}\text{C}$) e umidade relativa do ar (%):

Cidades	Variáveis climatológicas	Dias	Média	Desvio-padrão
Viçosa	Temperatura	15	19,10	5,00
	Umidade relativa do ar		62,04	10,00
Belo Horizonte	Temperatura	20	20,50	5,50
	Umidade relativa do ar		67,4	8,00

Pede-se:

a) (5%) Em qual das cidades a média da temperatura foi estimada com maior precisão? Justifique.

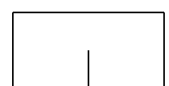
b) (5%) Baseando-se na cidade de Viçosa, qual das duas amostras (temperatura e umidade relativa do ar) é mais homogênea?



Nota

c) (5%) Baseando-se na cidade de Belo Horizonte, considere que a covariância entre a temperatura e a umidade relativa do ar é igual a 26,50. Calcule o coeficiente de correlação entre a temperatura e a umidade relativa do ar.

d) (5%) Neste estudo a temperatura foi medida em graus Celsius, caso se deseje representar a temperatura em graus Fahrenheit, como ficaria a resposta do item **c)**? (Para converter graus Celsius em graus Fahrenheit, multiplique por 1,80 e adicione 32, ou seja, $^{\circ}F = 1,80 ^{\circ}C + 32$).



3) (10%) Em relação aos princípios básicos que devem ser considerados no planejamento de experimentos, avalie as afirmações a seguir.

- I. A repetição refere-se ao número de unidades experimentais que receberão os tratamentos diferentes.
- II. A repetição possibilita que os erros experimentais sejam estimados.
- III. A casualização garante que os erros experimentais sejam independentes.
- IV. A ideia básica do controle local é a partição do conjunto total de unidades experimentais em subconjuntos (blocos) com a maior homogeneidade possível.
- V. Para a instalação de um experimento, é necessário, obrigatoriamente, estabelecer os três princípios básicos.

É correto apenas o que se afirma em:

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

4) (15%) A probabilidade de que um componente eletrônico de um computador falhe antes de mil horas de funcionamento é: 0,05, se o componente for da marca A ; 0,10 se o componente for da marca B ; e 0,15 se o componente for da marca C . Numa loja de manutenção, 50% dos componentes do estoque são da marca A , 20% da marca B e 30% da marca C . Um componente é escolhido ao acaso para o conserto de um computador. Determine a probabilidade de que ele funcione por mais de mil horas.

5) (15%) Seja X uma variável aleatória contínua (v.a.c) com a seguinte função densidade de probabilidade,

$$f(x) = \begin{cases} k, & -2 \leq x < 0 \\ k + \frac{3x}{125}, & 0 \leq x \leq 5 \\ 0, & \text{para outros valores de } x \end{cases}$$

a) (5%) Determine o valor de k ;

b) (5%) Encontre o valor de c , tal que $P(X \geq c) = 0,85$;

c) (5%) Se Y é uma variável aleatória definida como $Y = 8 + \frac{1}{2}X$, qual é a $V(Y)$?

6) (20%) São dados os seguintes valores obtidos de uma amostra aleatória com 5 observações.

X	1	2	4	6	12
Y	5	8	10	12	25

Admite-se que as variáveis X e Y estão relacionadas de acordo com o modelo $Y_i = \beta_0 + \beta_1 X_i + e_i$, em que, os e_i são erros independentes com distribuição normal de média zero e variância σ^2 .

a) (5%) Determine a reta da regressão de Y em relação a X e dê a interpretação do coeficiente de regressão.

b) (5%) Calcule o coeficiente de determinação e dê sua interpretação.

c) (5%) Construir um intervalo de confiança para o coeficiente de regressão ao nível de confiança de 90%.

d) (5%) Teste as hipóteses $H_0 : \beta_0 - 2\beta_1 = 0$ versus $H_a : \beta_0 - 2\beta_1 \neq 0$, ao nível de 5% de significância.

Formulário

$SQD_X = \sum_{i=1}^n X_i^2 - \frac{\left(\sum_{i=1}^n X_i\right)^2}{n}$	$SPD_{XY} = \sum_{i=1}^n X_i Y_i - \frac{\left(\sum_{i=1}^n X_i\right)\left(\sum_{i=1}^n Y_i\right)}{n}$	
$E[X] = \int x f(x) dx$	$V[X] = E[X^2] - (E[X])^2$	$CV(\%) = \frac{s}{\bar{X}} 100\%$
$S^2 = \frac{SQD_X}{n-1}$	$\widehat{V}(\widehat{\beta}_0) = \left(\frac{1}{n} + \frac{\bar{X}^2}{SQD_X}\right) QMR$	$\widehat{V}(\widehat{\beta}_1) = \frac{QMR}{SQD_X}$
$\widehat{Y}_i = \widehat{\beta}_0 + \widehat{\beta}_1 X_i$	$\widehat{\beta}_1 = \frac{SPD_{XY}}{SQD_X}$	$\widehat{\beta}_0 = \bar{Y} - \widehat{\beta}_1 \bar{X}$
$SQ_{\text{Regressão}} = \frac{(SPD_{XY})^2}{SQD_X} = \widehat{\beta}_1^2 SQD_X = \widehat{\beta}_1 SPD_{XY}$	$SQ_{\text{Total}} = SQD_Y$	
$\widehat{\epsilon}_i = Y_i - \widehat{Y}_i$	$r^2(\%) = \frac{SQ_{\text{Regressão}}}{SQ_{\text{Total}}} \times 100\%$	$r_{XY} = \frac{SPD_{XY}}{\sqrt{SQD_X \times SQD_Y}}$
$\widehat{V}(\widehat{Y}_n) = \left(\frac{1}{n} + \frac{(X_n - \bar{X})^2}{SQD_X}\right) QMR$	$F = \frac{r^2(n-2)}{1-r^2}$	$IC_{1-\alpha}(\mu) : \bar{X} \pm t_{\frac{\alpha}{2}} \frac{s}{\sqrt{n}}$
$t^2 = F(1, \text{gl do resíduo})$	$IC_{1-\alpha}(\beta_i) : \widehat{\beta}_i \pm t_{\frac{\alpha}{2}} \sqrt{\widehat{V}(\widehat{\beta}_i)}$	$t_{cal} = \frac{\widehat{\beta}_i - \beta_i}{\sqrt{\widehat{V}(\widehat{\beta}_i)}}$
$F_{cal} = \frac{\geq s^2}{< s^2}$		

Tabela 1: Valores de F ao nível de 5% de probabilidade. $P(F_{n_1, n_2} \geq F) = 0,05$. Número de graus de liberdade: $n_1 =$ numerador e $n_2 =$ denominador

n_2	n_1																						
	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	20	24	30	40	60	120	∞
1	161,4	199,5	215,7	224,6	230,2	234,0	236,8	238,9	240,5	241,9	243,0	243,9	244,4	245,0	245,9	246,0	248,0	249,1	250,1	251,1	252,2	253,3	254,3
2	18,51	19,00	19,16	19,25	19,30	19,33	19,35	19,37	19,38	19,40	19,40	19,41	19,42	19,42	19,43	19,43	19,45	19,45	19,46	19,47	19,48	19,49	19,50
3	10,13	9,55	9,28	9,12	9,01	8,94	8,89	8,85	8,81	8,79	8,76	8,74	8,72	8,71	8,70	8,69	8,66	8,64	8,62	9,59	8,57	8,55	8,53
4	7,71	6,94	6,59	6,39	6,26	6,16	6,09	6,04	6,00	5,96	5,93	5,91	5,89	5,87	5,86	5,84	5,80	5,77	5,75	5,72	5,69	5,66	5,63
5	6,61	5,79	5,41	5,19	5,05	4,95	4,88	4,82	4,77	4,74	4,70	4,68	4,66	4,64	4,62	4,60	4,56	4,53	4,50	4,46	4,43	4,40	4,36
6	5,99	5,14	4,76	4,53	4,39	4,28	4,21	4,15	4,10	4,06	4,03	4,00	3,98	3,96	3,94	3,92	3,87	3,84	3,81	3,77	3,74	3,70	3,67
7	5,59	4,74	4,35	4,12	3,97	3,87	3,79	3,73	3,68	3,64	3,60	3,57	3,55	3,52	3,51	3,49	3,44	3,41	3,38	3,34	3,30	3,27	3,23
8	5,32	4,46	4,07	3,84	3,69	3,58	3,50	3,44	3,39	3,35	3,31	3,28	3,25	3,23	3,22	3,20	3,15	3,12	3,08	3,04	3,01	2,97	2,93
9	5,12	4,26	3,86	3,63	3,48	3,37	3,29	3,23	3,18	3,14	3,10	3,07	3,04	3,02	3,01	2,98	2,94	2,90	2,86	2,83	2,79	2,75	2,71
10	4,96	4,10	3,71	3,48	3,33	3,22	3,14	3,07	3,02	2,98	2,94	2,91	2,88	2,86	2,85	2,82	2,77	2,74	2,70	2,66	2,62	2,58	2,54
11	4,84	3,98	3,59	3,36	3,20	3,09	3,01	2,95	2,90	2,85	2,82	2,79	2,76	2,74	2,72	2,70	2,65	2,61	2,57	2,53	2,49	2,45	2,40
12	4,75	3,89	3,49	3,26	3,11	3,00	2,91	2,85	2,80	2,75	2,72	2,69	2,66	2,64	2,62	2,60	2,54	2,51	2,47	2,43	2,38	2,34	2,30
13	4,67	3,81	3,41	3,18	3,03	2,92	2,83	2,77	2,71	2,67	2,63	2,60	2,57	2,55	2,53	2,51	2,46	2,42	2,38	2,34	2,30	2,25	2,21
14	4,60	3,74	3,34	3,11	2,96	2,85	2,76	2,70	2,65	2,60	2,56	2,53	2,50	2,48	2,46	2,44	2,39	2,35	2,31	2,27	2,22	2,18	2,13
15	4,54	3,68	3,29	3,06	2,90	2,79	2,71	2,64	2,59	2,54	2,51	2,48	2,45	2,43	2,40	2,39	2,33	2,29	2,25	2,20	2,16	2,11	2,07
16	4,49	3,63	3,24	3,01	2,85	2,74	2,66	2,59	2,54	2,49	2,45	2,42	2,39	2,37	2,35	2,33	2,28	2,24	2,19	2,15	2,11	2,06	2,01
17	4,45	3,59	3,20	2,96	2,81	2,70	2,61	2,55	2,49	2,45	2,41	2,38	2,35	2,33	2,31	2,29	2,23	2,19	2,15	2,10	2,06	2,01	1,96
18	4,41	3,55	3,16	2,93	2,77	2,66	2,58	2,51	2,46	2,41	2,37	2,34	2,31	2,29	2,27	2,25	2,19	2,15	2,11	2,06	2,02	1,97	1,92
19	4,38	3,52	3,13	2,90	2,74	2,63	2,54	2,48	2,42	2,38	2,34	2,31	2,28	2,26	2,23	2,21	2,16	2,11	2,07	2,03	1,98	1,93	1,88
20	4,35	3,49	3,10	2,87	2,71	2,60	2,51	2,45	2,39	2,35	2,31	2,28	2,25	2,23	2,20	2,18	2,12	2,08	2,04	1,99	1,95	1,90	1,84
21	4,32	3,47	3,07	2,84	2,68	2,57	2,49	2,42	2,37	2,32	2,28	2,25	2,22	2,20	2,18	2,15	2,10	2,05	2,01	1,96	1,92	1,87	1,81
22	4,30	3,44	3,05	2,82	2,66	2,55	2,46	2,40	2,34	2,30	2,26	2,23	2,20	2,18	2,15	2,13	2,07	2,03	1,98	1,94	1,89	1,84	1,78
23	4,28	3,42	3,03	2,80	2,64	2,53	2,44	2,37	2,32	2,27	2,24	2,20	2,17	2,14	2,13	2,10	2,05	2,01	1,96	1,91	1,86	1,81	1,76
24	4,26	3,40	3,01	2,78	2,62	2,51	2,42	2,36	2,30	2,25	2,22	2,18	2,15	2,13	2,11	2,09	2,03	1,98	1,94	1,89	1,84	1,79	1,73
25	4,24	3,39	2,99	2,76	2,60	2,49	2,40	2,34	2,28	2,24	2,20	2,16	2,13	2,11	2,09	2,06	2,01	1,96	1,92	1,87	1,82	1,77	1,71
26	4,23	3,37	2,98	2,74	2,59	2,47	2,39	2,32	2,27	2,22	2,18	2,15	2,12	2,10	2,07	2,05	1,99	1,95	1,90	1,85	1,80	1,75	1,69
27	4,21	3,35	2,96	2,73	2,57	2,46	2,37	2,31	2,25	2,20	2,16	2,13	2,10	2,08	2,06	2,03	1,97	1,93	1,88	1,84	1,79	1,73	1,67
28	4,20	3,34	2,95	2,71	2,56	2,45	2,36	2,29	2,24	2,19	2,15	2,12	2,09	2,06	2,04	2,02	1,96	1,91	1,87	1,82	1,77	1,71	1,65
29	4,18	3,33	2,93	2,70	2,55	2,43	2,35	2,28	2,22	2,18	2,14	2,10	2,07	2,05	2,03	2,00	1,94	1,90	1,85	1,81	1,75	1,70	1,64
30	4,17	3,32	2,92	2,69	2,53	2,42	2,33	2,27	2,21	2,16	2,12	2,09	2,06	2,04	2,01	1,99	1,93	1,89	1,84	1,79	1,74	1,68	1,62
40	4,08	3,23	2,84	2,61	2,45	2,34	2,25	2,18	2,12	2,08	2,04	2,00	1,97	1,95	1,92	1,90	1,84	1,79	1,74	1,69	1,64	1,58	1,51
60	4,00	3,15	2,76	2,53	2,37	2,25	2,17	2,10	2,04	1,99	1,95	1,92	1,89	1,86	1,84	1,81	1,75	1,70	1,65	1,59	1,53	1,47	1,39
120	3,92	3,07	2,68	2,45	2,29	2,17	2,09	2,02	1,96	1,91	1,86	1,83	1,80	1,77	1,75	1,73	1,66	1,61	1,55	1,50	1,43	1,35	1,25
∞	3,84	3,00	2,60	2,37	2,21	2,10	2,01	1,94	1,88	1,83	1,79	1,75	1,72	1,69	1,67	1,64	1,57	1,52	1,46	1,39	1,32	1,22	1,00

Adaptada de Bussab, W. O. e Morettin, P. A. Estatística Básica - Métodos Quantitativos, Editora Atual.

Tabela 2: Tabela da distribuição t de Student para $P[-t_c \leq T \leq t_c] = 1 - \alpha$, na primeira coluna temos os graus de liberdade (ν) e na primeira linha o nível de significância.

ν	Nível de probabilidade bilateral								
	0, 20	0, 10	0, 05	0, 04	0, 02	0, 01	0, 005	0, 002	0, 001
1	3,0777	6,3138	12,7062	15,8945	31,8205	63,6567	127,3213	318,3088	636,6192
2	1,8856	2,9200	4,3027	4,8487	6,9646	9,9248	14,0890	22,3271	31,5991
3	1,6377	2,3534	3,1824	3,4819	4,5407	5,8409	7,4533	10,2145	12,9240
4	1,5332	2,1318	2,7764	2,9985	3,7469	4,6041	5,5976	7,1732	8,6103
5	1,4759	2,0150	2,5706	2,7565	3,3649	4,0321	4,7733	5,8934	6,8688
6	1,4398	1,9432	2,4469	2,6122	3,1427	3,7074	4,3168	5,2076	5,9588
7	1,4149	1,8946	2,3646	2,5168	2,9980	3,4995	4,0293	4,7853	5,4079
8	1,3968	1,8595	2,3060	2,4490	2,8965	3,3554	3,8325	4,5008	5,0413
9	1,3830	1,8331	2,2622	2,3984	2,8214	3,2498	3,6897	4,2968	4,7809
10	1,3722	1,8125	2,2281	2,3593	2,7638	3,1693	3,5814	4,1437	4,5869
11	1,3634	1,7959	2,2010	2,3281	2,7181	3,1058	3,4966	4,0247	4,4370
12	1,3562	1,7823	2,1788	2,3027	2,6810	3,0545	3,4284	3,9296	4,3178
13	1,3502	1,7709	2,1604	2,2816	2,6503	3,0123	3,3725	3,8520	4,2208
14	1,3450	1,7613	2,1448	2,2638	2,6245	2,9768	3,3257	3,7874	4,1405
15	1,3406	1,7531	2,1314	2,2485	2,6025	2,9467	3,2860	3,7328	4,0728
16	1,3368	1,7459	2,1199	2,2354	2,5835	2,9208	3,2520	3,6862	4,0150
17	1,3334	1,7396	2,1098	2,2238	2,5669	2,8982	3,2224	3,6458	3,9651
18	1,3304	1,7341	2,1009	2,2137	2,5524	2,8784	3,1966	3,6105	3,9216
19	1,3277	1,7291	2,0930	2,2047	2,5395	2,8609	3,1737	3,5794	3,8834
20	1,3253	1,7247	2,0860	2,1967	2,5280	2,8453	3,1534	3,5518	3,8495
21	1,3232	1,7207	2,0796	2,1894	2,5176	2,8314	3,1352	3,5272	3,8193
22	1,3212	1,7171	2,0739	2,1829	2,5083	2,8188	3,1188	3,5050	3,7921
23	1,3195	1,7139	2,0687	2,1770	2,4999	2,8073	3,1040	3,4850	3,7676
24	1,3178	1,7109	2,0639	2,1715	2,4922	2,7969	3,0905	3,4668	3,7454
25	1,3163	1,7081	2,0595	2,1666	2,4851	2,7874	3,0782	3,4502	3,7251
26	1,3150	1,7056	2,0555	2,1620	2,4786	2,7787	3,0669	3,4350	3,7066
27	1,3137	1,7033	2,0518	2,1578	2,4727	2,7707	3,0565	3,4210	3,6896
28	1,3125	1,7011	2,0484	2,1539	2,4671	2,7633	3,0469	3,4082	3,6739
29	1,3114	1,6991	2,0452	2,1503	2,4620	2,7564	3,0380	3,3962	3,6594
30	1,3104	1,6973	2,0423	2,1470	2,4573	2,7500	3,0298	3,3852	3,6460
31	1,3095	1,6955	2,0395	2,1438	2,4528	2,7440	3,0221	3,3749	3,6335
32	1,3086	1,6939	2,0369	2,1409	2,4487	2,7385	3,0149	3,3653	3,6218
33	1,3077	1,6924	2,0345	2,1382	2,4448	2,7333	3,0082	3,3563	3,6109
34	1,3070	1,6909	2,0322	2,1356	2,4411	2,7284	3,0020	3,3479	3,6007
35	1,3062	1,6896	2,0301	2,1332	2,4377	2,7238	2,9960	3,3400	3,5911
36	1,3055	1,6883	2,0281	2,1309	2,4345	2,7195	2,9905	3,3326	3,5821
37	1,3049	1,6871	2,0262	2,1287	2,4314	2,7154	2,9852	3,3256	3,5737
38	1,3042	1,6860	2,0244	2,1267	2,4286	2,7116	2,9803	3,3190	3,5657
39	1,3036	1,6849	2,0227	2,1247	2,4258	2,7079	2,9756	3,3128	3,5581
40	1,3031	1,6839	2,0211	2,1229	2,4233	2,7045	2,9712	3,3069	3,5510
50	1,2987	1,6759	2,0086	2,1087	2,4033	2,6778	2,9370	3,2614	3,4960
60	1,2958	1,6706	2,0003	2,0994	2,3901	2,6603	2,9146	3,2317	3,4602
120	1,2886	1,6577	1,9799	2,0763	2,3578	2,6174	2,8599	3,1595	3,3735
$+\infty$	1,2816	1,6449	1,9600	2,0537	2,3263	2,5758	2,8070	3,0902	3,2905



Universidade Federal de Viçosa

Departamento de Estatística

Prova seletiva para o Programa de Pós-Graduação em Estatística Aplicada e Biometria.

Nível Mestrado - 19/Novembro/2018

Número de inscrição:

--	--	--	--	--	--	--	--	--	--

- Interpretar corretamente as questões é parte da avaliação.
- Indique todos os cálculos organizadamente.
- É permitida a utilização de calculadora científica;
- São SEIS (6) questões em páginas numeradas de 1 a 11, total de 100 pontos.
- O tempo máximo para realização desta prova é de 4 horas.
- **BOA SORTE !!**

--

1) (20%) Considerando as informações apresentadas nas Tabelas 1 e 2, responda as questões apresentadas a seguir.

Tabela 1: Resultados da análise de variância

FV	GL	SQ	QM	F	$F_{0,05}$
Blocos	4	660,80			
Tratamentos		381,40			3,49
Resíduo					
Total	19	1155,80			

Tabela 2: Médias ($\bar{y}_{i.}$) de cada um dos i tratamentos

Médias
$\bar{y}_{4.} = 86,0$
$\bar{y}_{3.} = 84,8$
$\bar{y}_{2.} = 82,8$
$\bar{y}_{1.} = 74,8$

a) (05%) Qual é o modelo estatístico para essa ANOVA? Descreva cada um dos termos presentes no modelo.

b) (05%) Quais são as hipóteses estatísticas associadas a fonte de variação tratamentos?



Nota

c) (05%) Complete a ANOVA e conclua a respeito do teste F para a fonte de variação tratamentos.

d) (05%) Supondo que, para $\alpha = 5\%$, a diferença mínima significativa pelo teste de Tukey seja $\Delta = 5,77$. Qual tratamento proporciona menor média?

2) (20%) Abaixo estão listadas algumas informações sobre a variável aleatória X :

i) A função densidade de probabilidade de X é dada por:

$$f(x) = \begin{cases} ax + b, & 0 \leq x \leq c \\ 0, & \text{para outros valores de } x \end{cases}$$

em que $c \geq 1$;

ii) A função de distribuição acumulada de X é denotada por $F(x)$, sendo que, o valor da função nos pontos $\frac{1}{2}$ e 1 são, respectivamente, iguais a:

$$F\left(\frac{1}{2}\right) = \frac{7}{16} \text{ e } F(1) = \frac{3}{4}.$$

Utilize os itens i) e ii) e responda:

a) (10%) Encontre os valores das constantes a e b .



Nota

b) (10%) Encontre o limite superior (c) do intervalo $0 \leq x \leq c$.

3) (15%) Considere o modelo $Y_i = \beta_0 + \frac{\beta_1}{X_i} + e_i$, para $i = 1, \dots, n$; $E[e_i] = 0$; $E[e_i e_j] = 0$, para $i \neq j$ e $E[e_i^2] = \sigma^2$.

a) (**10%**) Obtenha, de acordo com o método dos mínimos quadrados, a fórmula para estimar os parâmetros do modelo.

b) (05%) Determine as estimativas dos parâmetros β_0 e β_1 , com base nos seguintes dados:

Y	9	8	8	6
X	12	16	18	24

4) (10%) A probabilidade de que um homem esteja vivo daqui a 30 anos é $\frac{2}{5}$; a de sua mulher é $\frac{2}{3}$. Determinar a probabilidade de que daqui a 30 anos:

a) (2,5%) Ambos estejam vivos;

b) (2,5%) Somente o homem esteja vivo;

c) (2,5%) Nenhum esteja vivo;

d) (2,5%) Pelos menos um esteja vivo.



Nota

5) (20%) Na estatística aplicada quando se tem mais de dois tratamentos qualitativos, é comum a realização de uma análise de variância para verificar se ocorre ou não diferenças entre estes tratamentos. Entretanto, para realizar esta análise deve-se verificar um conjunto de pressuposições, entre as quais temos a distribuição normal dos erros, a homogeneidade de variâncias dos erros, etc., e também, realizar uma análise descritiva das observações dos tratamentos. Em um experimento ao usar um delineamento inteiramente casualizado (DIC), a estimativa destes erros em cada tratamento, é o mesmo que calcular um desvio em relação à média do tratamento, isto é, $\hat{e}_{ij} = y_{ij} - \bar{y}_i$, em que $i = 1, 2, \dots, I$ tratamentos e $j = 1, 2, \dots, k$ repetições. Considerando esta premissa e a Figura 1 abaixo, responda:

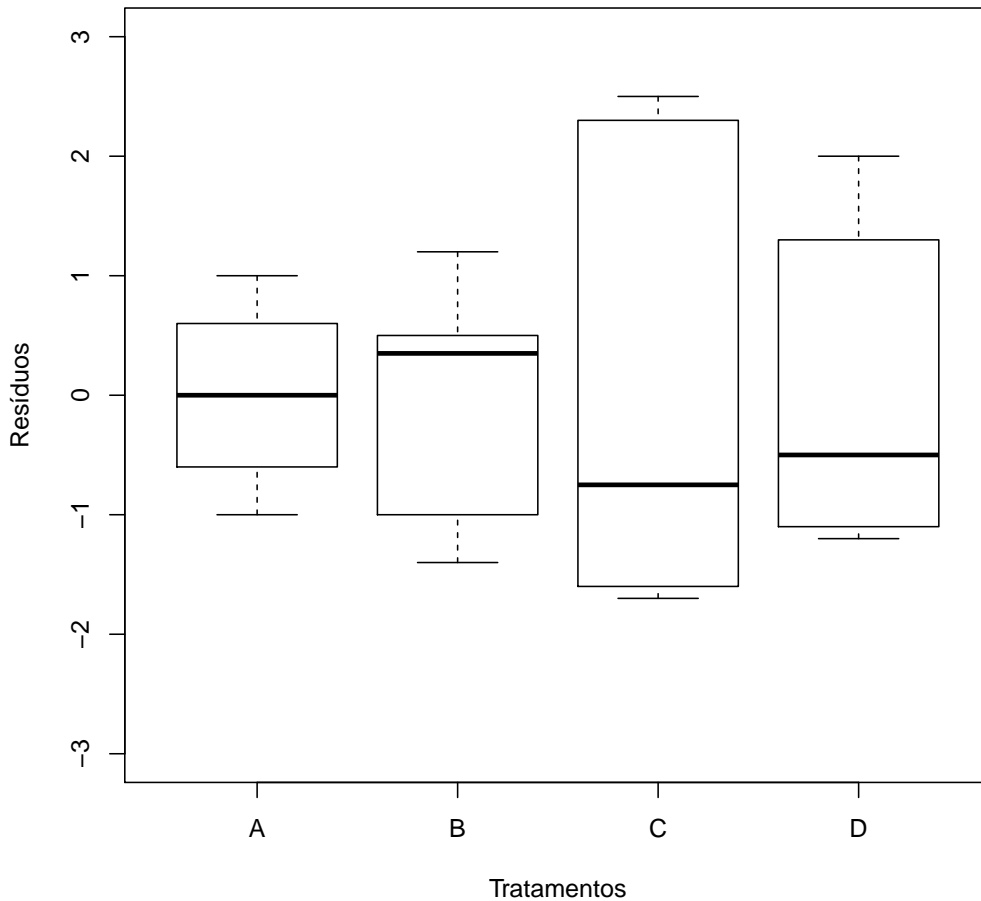


Figura 1: Box-plot para as estimativas dos erros das observações de cada tratamentos de um experimento realizado no delineamento inteiramente casualizado (DIC).

a) (5%) Existe homogeneidade de variâncias nas estimativas dos erros entre os tratamentos? Explique.



Nota

b) (5%) Qual é a média das estimativas dos erros em cada tratamento? Explique.

c) (5%) Em qual tratamento pode-se inferir que os erros se aproximam de uma distribuição normal? Por que?

d) (5%) Classifique o tipo de assimetria que está ocorrendo no tratamento *C*.

6) (15%) Para um experimento balanceado, o modelo estatístico sob o delineamento em blocos casualizados (DBC) é definido por $y_{ij} = \mu + \tau_i + \omega_j + \varepsilon_{ij}$ em que: y_{ij} é o valor observado de Y no tratamento x_i , para $i = 1, 2, \dots, t$ e no bloco b_j , para $j = 1, 2, \dots, r$; $\mu = \mu_Y$ é a média geral populacional de Y ; $\tau_i = \mu_{Y_i} - \mu_Y$ é o efeito do tratamento x_i ; $\omega_j = \mu_{Y_j} - \mu_Y$ é o efeito do bloco b_j ; $\varepsilon_{ij} = y_{ij} - \mu_{Y_i} - \mu_{Y_j} + \mu_Y$ é o erro experimental associado ao valor observado y_{ij} ; μ_{Y_i} é a média populacional de Y no tratamento x_i e; μ_{Y_j} é a média populacional de Y no bloco b_j . Desse modo e de acordo com os dados coletados em dois experimentos instalados sob o DBC, com três tratamentos (x_1, x_2, x_3) e três repetições (b_1, b_2, b_3) por tratamento, qual deles proporcionou o maior quadrado médio do resíduo (QMRes)? Por quê?

Experimento 1				
	x_1	x_2	x_3	$\bar{y}_{.j}$
b_1	10,3	20,4	30,8	20,5
b_2	20,9	30,5	40,7	30,7
b_3	30,6	40,2	50,1	40,3
\bar{y}_i	20,6	30,4	40,5	30,5

Experimento 2				
	x_1	x_2	x_3	$\bar{y}_{.j}$
b_1	10,3	20,4	30,8	20,5
b_2	40,6	40,2	40,7	40,5
b_3	50,1	40,2	30,6	40,3
\bar{y}_i	33,7	33,6	34	33,8



Universidade Federal de Viçosa
Departamento de Estatística

Prova seletiva para o Programa de Pós-Graduação em Estatística Aplicada e Biometria.

Nível Mestrado - 11/Junho/2019

Número de inscrição:

--	--	--	--	--	--	--	--	--	--

- Interpretar corretamente as questões é parte da avaliação.
- Indique todos os cálculos organizadamente.
- É permitida a utilização de calculadora científica;
- São SEIS (6) questões em páginas numeradas de 1 a 11, total de 100 pontos.
- O tempo máximo para realização desta prova é de 4 horas.
- **BOA SORTE !!**

--

1) (15%)

a) (07%) Cite e defina os três princípios básicos da experimentação;

b) (08%) Descreva um experimento, que não seja um dos apresentados em outras questões da prova, onde foram utilizados os três princípios básicos da experimentação, informando os tratamentos, a variável resposta, o delineamento experimental e as unidades experimentais. Na descrição do experimento, resalte como foi feita a casualização dos tratamentos.



2) (15%) Os estudantes em uma pré-seleção para empresa júnior são classificados em $2x\%$ ótimos, $12x\%$ bons e $6x\%$ regulares. Um segundo teste é proposto para classificar os estudantes em aprovado ou reprovado. Com base na classificação anterior, foram obtidas as seguintes probabilidades condicionais no teste:

- i) daqueles considerados ótimos, $19x\%$ foram aprovados;
- ii) daqueles considerados bons, $16x\%$ foram aprovados;
- iii) daqueles considerados regulares, $2x\%$ foram aprovados;

Pede-se:

- a) (05%) Calcule o valor de x ;



Nota

b) (05%) Calcule a probabilidade de um estudante aprovado no teste pertencer à classe regular.

c) (05%) Calcule a probabilidade de um estudante não aprovado no teste pertencer à classe ótimo ou à classe bom.

3) (15%) Um pesquisador testou seis doses de um adubo (X) em kg por hectare e registrou os respectivos totais de produção (Y) em kg por hectare. Neste estudo o pesquisador obteve uma equação de regressão linear simples ajustada dada por $\hat{Y}_i = 1850 + 11X_i$ e um coeficiente de determinação $R^2 = 0,81$. Dado que as doses de adubo foram 0; 50; 100; 120; 170 e 200 kg, pede-se:

a) (03%) Interprete a constante de regressão;

b) (03%) Interprete o coeficiente de regressão;

c) (03%) Interprete o coeficiente de determinação;

d) (03%) Calcule e interprete o coeficiente de correlação linear simples;

e) (03%) Estime a produção média, em kg por hectare, quando se aplica 180 kg de adubo.

4) (15%) Você é um candidato à seleção ao mestrado em Estatística Aplicada e Biometria e necessita **planejar um projeto de pesquisa**, para ser avaliado e mostrar o seu conhecimento em Estatística Experimental para a banca examinadora.

O problema: O trabalho consiste em avaliar 4 tipos de gasolina (A, B, C, D), que se diferenciam pelo tipo de aditivo acrescentados às mesmas, e verificar em qual delas se obtém o maior rendimento em km/L.

As condições: Para este estudo você tem a sua disposição apenas 4 carros da mesma marca, porém com quilometragens diferentes. Os testes são demorados e a pista onde estes serão realizados só poderá ser utilizada uma vez por dia por cada carro, sendo necessário, portanto, quatro dias para a realização dos testes.

Seu trabalho: Planeje o experimento de forma adequada, seguindo os princípios básicos da experimentação. Neste planejamento, responda:

a) (05%) Qual o delineamento a ser utilizado? Mostre um croqui do experimento no qual a banca examinadora poderá observar o delineamento utilizado na pesquisa.



b) (05%) Mostre como deverá ser realizada a análise de variância deste experimento mostrando o Quadro da ANOVA com as fontes de variações e graus de liberdade. Discuta também, qual o procedimento deverá ser adotado para obter a conclusão do objetivo do estudo.

c) (05%) Faça uma sugestão de como poderia aumentar o número de unidades experimentais neste estudo.



5) (20%) O t3pico “Estatística Descritiva” é básico em qualquer disciplina de estatística básica. Nesse t3pico apresentamos e discutimos sobre medidas de posição e de dispersão, além de construção e interpretação de gráficos e tabelas. Pede-se:

a) (04%) Defina uma variável e crie/invente um pequeno conjunto de dados ($n = 20$) dessa variável. Apresente os dados brutos;

b) (04%) Apresente seu conjunto de dados numa forma gráfica apropriada. Acrescente um título à sua forma gráfica apresentada.



c) (04%) Discorra sobre uma medida de posição e calcule seu valor usando seus dados;

d) (04%) Discorra sobre uma medida de dispersão e calcule seu valor usando seus dados;

e) (04%) Interprete ou discuta sobre os resultados obtidos nos itens b), c) e d).

6) (20%) Dada a variável aleatória contínua (v.a.c.) cuja função densidade de probabilidade (f.d.p.) é dada por:

$$f(x) = \begin{cases} (1 + K + x), & -1 \leq x < 0 \\ (1 + K - x), & 0 \leq x < 1 \\ 0, & \text{para outros valores de } x \end{cases},$$

responda:

a) (05%) Qual o valor de K para que $f(x)$ seja uma função densidade de probabilidade?

b) (05%) Calcule a $E[X]$ e $\text{var}[X]$;



Nota

c) (05%) Considerando-se $f(x)$, o que pode ser dito acerca da mediana de X ?

d) (05%) Reescreva a fórmula prática de $\text{var}[X]$ considerando o resultado de $E[X]$ obtido no item b).

Universidade Federal de Viçosa

Departamento de Estatística

Prova seletiva para o Programa de Pós-Graduação em Estatística Aplicada e Biometria.

Nível Mestrado - 11/Novembro/2019

Número de inscrição:

--	--	--	--	--	--	--	--	--	--

- Interpretar corretamente as questões é parte da avaliação.
- Indique todos os cálculos organizadamente.
- São SEIS (6) questões em páginas numeradas de 1 a 13, total de 100 pontos.
- Na página 13 temos tabela.
- É permitida a utilização de calculadora científica.
- O tempo máximo para realização desta prova é de 4 horas.
- **BOA SORTE!!**

Revisão	Total

1) (17%) O conhecimento básico sobre “Estatística Descritiva” é de caráter fundamental nos cursos de Estatística. Nesse tópico falamos sobre medidas de posição, também conhecidas como medidas de tendência central; medidas de dispersão; e sobre a organização dos dados brutos usando tabelas e gráficos. Pede-se:

a) (5%) Explique a vantagem de usarmos tabelas e/ou gráficos para descrever certo conjunto de dados brutos. Você deverá usar um ou mais exemplo(s) hipotético(s) para ilustrar seus argumentos.

b) (6%) Explique como deve ser montado um histograma. Sua resposta deve conter, no mínimo, comentários sobre como definir o número de classes e a amplitude da classe.

c) (6%) Comente sobre como realizar o cálculo aproximado da média e da variância dos dados organizados num histograma. Use um exemplo genérico.

2) (17%) Considere um experimento aleatório E e S um espaço amostral a ele associado. Sejam A , B e C eventos quaisquer de S com as seguintes probabilidades:

$$\begin{array}{cccc} P(A) = 0,30 & P(B) = 0,25 & P(C) = 0,25 & P(A \cup B) = 0,47 \\ P(C \cap (A \cup B)) = 0,12 & P(B \cap C) = 0,10 & P(B \cap (A \cup C)) = 0,15 & \end{array}$$

a) (10%) Encontre a probabilidade de que exatamente dois dos três eventos ocorram.

b) (7%) Encontre a probabilidade de que nenhum dos três eventos ocorra.

3) (17%) Sabidamente há diferenças entre variáveis aleatórias discretas que seguem distribuição Binomial e Poisson. Contudo em algumas situações podemos mensurar as diferenças entre as probabilidades calculadas entre estas distribuições, o que chamamos de Aproximação da Binomial pela Poisson. Assim, dado que a cada 400 lâmpadas, 2 se queimam, qual a probabilidade de que, em 600 lâmpadas, no máximo 2 se queimam?

a) (4%) Utilizando a distribuição binomial;

b) (4%) Utilizando a aproximação da distribuição binomial pela Poisson;

c) (5%) Sabe-se também que pode-se aproximar o cálculo de probabilidades da distribuição binomial pela Normal e, em geral, a correção de continuidade é efetuada para melhorar a aproximação. Calcule a probabilidade pedida utilizando a aproximação da binomial pela normal considerando a correção de continuidade.

d) (4%) Comparando-se as aproximações apresentadas em b) e em c) qual produziu menor erro? Justifique sua resposta.



4) (17%) Um pesquisador aplicou CO₂ com onze diferentes concentrações em folhas de trigo (X) à temperatura de 35 °C e observou a quantidade de CO₂ absorvida (Y) em cm³/dm²/hora, sendo que as concentrações de CO₂ aplicadas variaram de 75 a 250. Neste estudo, o pesquisador obteve a equação de regressão linear simples ajustada dada por: $\hat{Y}_i = -2,076 + 0,025X_i$. Uma tabela com o resumo da análise de variância da regressão é apresentada abaixo

FV	GL	QM	F _{cal}	Hipóteses
Regressão		22,3041		$\left\{ \begin{array}{l} H_0 : \\ H_1 : \end{array} \right.$
Resíduo		0,0631		
Total				

Pede-se:

a) (5%) Complete a tabela.

b) (4%) Calcule e interprete os coeficientes de determinação e de correlação linear simples.



Nota

c) (4%) Interprete o coeficiente de regressão.

d) (4%) Interprete a constante de regressão.

5) (16%)

a) (8%) Considere um experimento instalado para avaliar o efeito de 5 tratamentos (5 rações: A , B , C , D e E) sobre a variável-resposta GPD (ganho de peso diário, g/dia) em suínos. O pesquisador utilizou 5 animais desuniformes para o fator perturbador peso inicial. Cada ração foi designada de forma sistemática a um único animal. Os princípios básicos da experimentação foram utilizados de forma adequada? **Justifique sua resposta.**

b) (8%) Cite e explique a finalidade de cada um dos três princípios básicos da experimentação.

6) (16%) Foi realizado um experimento no esquema fatorial utilizando o delineamento inteiramente ao acaso, com 5 repetições. O objetivo foi estudar a resistência do concreto formado por uma combinação dos fatores G (agregado graúdo) e M (agregado miúdo). O fator G possuía três níveis (G_1 —pedregulho, G_2 —brita e G_3 —uma mistura de ambos) e o fator M também possuía três níveis (M_1 —areia de origem natural, M_2 —areia resultante de britagem e M_3 —uma mistura de ambas). A resistência do concreto foi avaliada submetendo um corpo de prova, feito do concreto que se quer avaliar, a um carregamento que aumenta progressivamente até a ruptura da amostra. O valor da força exercida no momento da ruptura indicou a resistência máxima que o concreto pode suportar. Pergunta-se:

a) (4%) Quanto e quais foram os tratamentos utilizados? Quantas unidades experimentais haviam no experimento?

b) (4%) Mostre o quadro de ANOVA do experimento com as fontes de variação e os graus de liberdade.

c) (4%) Qual é a estratégia para continuar a análise estatística se o teste F para a interação for não significativa?

d) (4%) Qual é a estratégia para continuar a análise estatística se o teste F para a interação for significativa?

Tabela 1: Probabilidade na distribuição normal padrão de zero a um valor positivo z_c , ou seja, $P[0 \leq Z \leq z_c] = p$.

z_c	0,00	0,01	0,02	0,03	0,04	0,05	0,06	0,07	0,08	0,09
0,0	0,0000	0,0040	0,0080	0,0120	0,0160	0,0199	0,0239	0,0279	0,0319	0,0359
0,1	0,0398	0,0438	0,0478	0,0517	0,0557	0,0596	0,0636	0,0675	0,0714	0,0753
0,2	0,0793	0,0832	0,0871	0,0910	0,0948	0,0987	0,1026	0,1064	0,1103	0,1141
0,3	0,1179	0,1217	0,1255	0,1293	0,1331	0,1368	0,1406	0,1443	0,1480	0,1517
0,4	0,1554	0,1591	0,1628	0,1664	0,1700	0,1736	0,1772	0,1808	0,1844	0,1879
0,5	0,1915	0,1950	0,1985	0,2019	0,2054	0,2088	0,2123	0,2157	0,2190	0,2224
0,6	0,2257	0,2291	0,2324	0,2357	0,2389	0,2422	0,2454	0,2486	0,2517	0,2549
0,7	0,2580	0,2611	0,2642	0,2673	0,2704	0,2734	0,2764	0,2794	0,2823	0,2852
0,8	0,2881	0,2910	0,2939	0,2967	0,2995	0,3023	0,3051	0,3078	0,3106	0,3133
0,9	0,3159	0,3186	0,3212	0,3238	0,3264	0,3289	0,3315	0,3340	0,3365	0,3389
1,0	0,3413	0,3438	0,3461	0,3485	0,3508	0,3531	0,3554	0,3577	0,3599	0,3621
1,1	0,3643	0,3665	0,3686	0,3708	0,3729	0,3749	0,3770	0,3790	0,3810	0,3830
1,2	0,3849	0,3869	0,3888	0,3907	0,3925	0,3944	0,3962	0,3980	0,3997	0,4015
1,3	0,4032	0,4049	0,4066	0,4082	0,4099	0,4115	0,4131	0,4147	0,4162	0,4177
1,4	0,4192	0,4207	0,4222	0,4236	0,4251	0,4265	0,4279	0,4292	0,4306	0,4319
1,5	0,4332	0,4345	0,4357	0,4370	0,4382	0,4394	0,4406	0,4418	0,4429	0,4441
1,6	0,4452	0,4463	0,4474	0,4484	0,4495	0,4505	0,4515	0,4525	0,4535	0,4545
1,7	0,4554	0,4564	0,4573	0,4582	0,4591	0,4599	0,4608	0,4616	0,4625	0,4633
1,8	0,4641	0,4649	0,4656	0,4664	0,4671	0,4678	0,4686	0,4693	0,4699	0,4706
1,9	0,4713	0,4719	0,4726	0,4732	0,4738	0,4744	0,4750	0,4756	0,4761	0,4767
2,0	0,4772	0,4778	0,4783	0,4788	0,4793	0,4798	0,4803	0,4808	0,4812	0,4817
2,1	0,4821	0,4826	0,4830	0,4834	0,4838	0,4842	0,4846	0,4850	0,4854	0,4857
2,2	0,4861	0,4864	0,4868	0,4871	0,4875	0,4878	0,4881	0,4884	0,4887	0,4890
2,3	0,4893	0,4896	0,4898	0,4901	0,4904	0,4906	0,4909	0,4911	0,4913	0,4916
2,4	0,4918	0,4920	0,4922	0,4925	0,4927	0,4929	0,4931	0,4932	0,4934	0,4936
2,5	0,4938	0,4940	0,4941	0,4943	0,4945	0,4946	0,4948	0,4949	0,4951	0,4952
2,6	0,4953	0,4955	0,4956	0,4957	0,4959	0,4960	0,4961	0,4962	0,4963	0,4964
2,7	0,4965	0,4966	0,4967	0,4968	0,4969	0,4970	0,4971	0,4972	0,4973	0,4974
2,8	0,4974	0,4975	0,4976	0,4977	0,4977	0,4978	0,4979	0,4979	0,4980	0,4981
2,9	0,4981	0,4982	0,4982	0,4983	0,4984	0,4984	0,4985	0,4985	0,4986	0,4986
3,0	0,4987	0,4987	0,4987	0,4988	0,4988	0,4989	0,4989	0,4989	0,4990	0,4990
3,1	0,4990	0,4991	0,4991	0,4991	0,4992	0,4992	0,4992	0,4992	0,4993	0,4993
3,2	0,4993	0,4993	0,4994	0,4994	0,4994	0,4994	0,4994	0,4995	0,4995	0,4995
3,3	0,4995	0,4995	0,4995	0,4996	0,4996	0,4996	0,4996	0,4996	0,4996	0,4997
3,4	0,4997	0,4997	0,4997	0,4997	0,4997	0,4997	0,4997	0,4997	0,4997	0,4998
3,5	0,4998	0,4998	0,4998	0,4998	0,4998	0,4998	0,4998	0,4998	0,4998	0,4998
3,6	0,4998	0,4998	0,4999	0,4999	0,4999	0,4999	0,4999	0,4999	0,4999	0,4999
3,7	0,4999	0,4999	0,4999	0,4999	0,4999	0,4999	0,4999	0,4999	0,4999	0,4999
3,8	0,4999	0,4999	0,4999	0,4999	0,4999	0,4999	0,4999	0,4999	0,4999	0,4999
3,9	0,5000	0,5000	0,5000	0,5000	0,5000	0,5000	0,5000	0,5000	0,5000	0,5000

