UNIVERSIDADE FEDERAL DE LAVRAS
DEPARTAMENTO DE ESTATÍSTICA

Programa de Pós-Graduação em Estatística e Experimentação Agropecuária

Prova do Processo Seletivo para o Doutorado 2018-02

Nº de inscrição do candidato: ______________________

- Utilizar APENAS o número de inscrição para identificar a sua prova;
- A interpretação das questões é parte da avaliação;
- Indique todos os cálculos organizadamente;
- São DEZ (10) questões, valendo UM (1) ponto cada, totalizando 10 pontos;
- O tempo máximo para a realização desta prova é de 4 horas;
- A prova pode ser feita a lápis ou caneta (azul ou preta) e é permitido utilizar somente a calculadora.
- BOA SORTE !!

1) Considere uma variável aleatória X com densidade de probabilidade $f(x,\alpha)$, em que α é um parâmetro desconhecido. Existem vários métodos para se obter, a partir de $f(x,\alpha)$, outras distribuições de probabilidade. Um desses métodos consiste em supor uma densidade de probabilidade $g(\alpha)$ para o parâmetro α e definir uma nova densidade por

$$h(x)=\int_D f(x,\alpha)g(\alpha)d\alpha$$

em que a integral é sobre o domínio D de definição do parâmetro α.

Mostre que $h(x)$ é realmente uma densidade de probabilidade.

2) Um outro método de se construir distribuições, ainda na mesma linha da questão anterior, consiste em: se $F(x)$ é uma função de distribuição de probabilidade, sob certas condições, $G(x)$ definida como $G(x)=\left(F(x)\right)^n$ é uma nova distribuição de probabilidade.

Se X tem distribuição exponencial de parâmetro $\alpha>0$, determine a expressão para a densidade de probabilidade correspondente à nova distribuição definida por $G(x)=\left(F(x)\right)^2$

3) Se $\mathbf{x}\in\mathbb{R}^n$ e $\mathbf{y}\in\mathbb{R}^n$, define-se o produto interno desses vetores como $\langle\mathbf{x},\mathbf{y}\rangle=\sum_{j=1}^{n}x_j y_j$.

Observe que $\langle\mathbf{x},\mathbf{x}\rangle=\sum_{j=1}^{n}x_j^2=\|\mathbf{x}\|^2$. A projeção ortogonal do vetor $\mathbf{x}$ sobre o vetor de constantes $(1,1,...,1)=\mathbf{1}_n$ é dada por

$$\frac{\langle\mathbf{x},\mathbf{1}_n\rangle}{n}\mathbf{1}_n=\left(\frac{\sum_{j=1}^{n}x_j}{n}\right)\mathbf{1}_n=\bar{x}\,\mathbf{1}_n=(\bar{x},\bar{x},...,\bar{x})=\bar{\mathbf{x}}.$$

A diferença $\mathbf{x}_c=\mathbf{x}-\bar{\mathbf{x}}$ é chamada de resíduo e $\|\mathbf{x}\|^2=\|\bar{\mathbf{x}}\|^2+\|\mathbf{x}_c\|^2$.

Mostre que, para um real qualquer $\alpha\neq 0$, o produto interno $\langle\mathbf{x}-\alpha\mathbf{1}_n,\mathbf{x}-\alpha\mathbf{1}_n\rangle=\|\mathbf{x}-\alpha\mathbf{1}_n\|^2$ é mínimo se $\alpha=\bar{x}$.

4) Considerando que X e Y são variáveis aleatórias absolutamente contínuas conjuntamente com função de distribuição $F_{X,Y}(x,y)=\frac{1}{2}x^2y+\frac{1}{2}xy^3$, para $0\leq x\leq 1$ e $0\leq y\leq 1$, encontrar as funções densidades de probabilidade marginais de X e Y.

5) Para uma função ser função de probabilidade legítima ela tem de ser não negativa e somar 1 em relação ao conjunto suporte da variável aleatória X subjacente. Se X é uma binomial, provar que a função

$$p_X(x)=P(X=x)=\binom{n}{x}p^x(1-p)^{n-x}$$

é uma função de probabilidade legítima, para $0<p<1$ e $x=0, 1, ..., n$.

6) Um dos métodos de inferência mais utilizados é o método dos momentos. Se $X_1,...,X_n$ é uma amostra de uma população, então o método consiste em igualar os momentos populacionais com os momentos amostrais.

$$E\left[X^k\right]=\frac{1}{n}\sum_{i=1}^{n}X_i^k$$

Se um parâmetro é função de momentos populacionais, seu estimador é dado pela mesma função aplicada aos momentos amostrais correspondentes.

A variância populacional é

$$\sigma^2=E\left[\left(X-E\left[X\right]\right)^2\right]=E\left[X^2-2X\,E\left[X\right]+\left(E\left[X\right]\right)^2\right]$$

A partir dessa fórmula obtenha o estimador de σ^2 pelo método dos momentos

7) Considere a distribuição de probabilidade dada por

$$f\left(x;\alpha\right)=\alpha\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}\left(x+1\right)^2}+\left(1-\alpha\right)\frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}\left(x-1\right)^2} \qquad \left(0<\alpha<1\right)$$

Calcule o estimador de máxima verossimilhança para α, considerando a amostra $x_1=-1$ e $x_2=1$.

8) Um pesquisador pretende instalar um experimento para comparar dois sistemas de plantio usando máquinas (1 – plantio convencional, 2 – plantio direto) e seis variedades de milho (A,B,C,D, E, F). O experimento será instalado em blocos ao acaso, com três repetições.
   i) Faça o esquema de campo ilustrando a casualização dos tratamentos em um bloco considerando que o experimento será instalado em esquema fatorial;

ii) Faça o esquema de campo ilustrando a casualização dos tratamentos em um bloco, considerando que o experimento será instalado em esquema de parcela subdividida, com o fator "**Sistema de plantio**" na parcela e o fator "**Variedade**" na subparcela;

iii) Faça o esquema de campo ilustrando a casualização dos tratamentos em um bloco, considerando que o experimento será instalado em esquema de parcela subdividida, com o fator "**Variedade**" na parcela e o fator "**Sistema de plantio**" na subparcela;

iv) Apresente a estrutura da análise de variância (Fontes de Variação e Número de Graus de Liberdade) para os três planejamentos.

9) Um pesquisador da área de forragicultura avaliou o efeito de doses de fósforo no crescimento e desenvolvimento do capim "Coastcross". O estudo envolveu a aplicação de fósforo por hectare nos seguintes níveis: 0, 60, 120 e 240 kg de P_2O_5/ha.

A produção de matéria seca, em kg/ha, respondeu aos níveis de fósforo segundo a equação:

$$\hat{y}_i = 3786{,}9100 + 8{,}0540X_i - 0{,}0229X_i^2$$

$$R^2 = 92{,}83\%$$

sendo: $\mathbf{X_i}$ o nível de fósforo, em kg/ha de P_2O_5 e $\mathbf{y_i}$ a produção de matéria seca, em kg/ha.

i) Faça o estudo matemático da equação, informando a adubação que promoveu a produção máxima de matéria seca. Estime a produção máxima;

ii) Interprete o coeficiente de determinação (R^2)

10) Um experimento teve por objetivo comparar a produtividade de cinco cultivares de alho (A, B, C, D, E). Adotou-se o delineamento em quadrado latino para controlar variações na fertilidade do solo (Linhas) e a heterogeneidade nos tamanhos dos bulbos de alho no plantio - Colunas (1-florão, 2-graúdo, 3-médio, 4-pequeno, 5-miúdo). A parcela experimental teve área de $2m^2$ (1mx2m). Os resultados obtidos, em t/ha, foram os seguintes:

Linhas	Colunas/Tamanhos dos Bulbos					Totais Linhas
	1	**2**	**3**	**4**	**5**	
1	11 **B**	4 **D**	7 **C**	9 **E**	9 **A**	**40**
2	12 **A**	9 **B**	7 **D**	9 **C**	15 **E**	**52**
3	8 **D**	7 **C**	12 **E**	5 **A**	8 **B**	**40**
4	17 **C**	12 **E**	12 **A**	10 **B**	7 **D**	**58**
5	12 **E**	7 **A**	9 **B**	8 **D**	10 **C**	**46**
Totais Colunas	**60**	**39**	**47**	**41**	**49**	**G = 236**

Tratamentos	Totais (T_i)	Médias ($\bar{y}_i$)
A	45	9,0
B	47	9,4
C	50	10,0
D	34	6,8
E	60	12,0

Os resultados forneceram SQTotal = 210,2.

i) Faça a análise de variância, aplique o teste F e discuta os resultados. Utilize o valor 4,0 para F tabelado com $\alpha = 5\%$.

ii) Avalie a precisão do experimento.