UNIVERSIDADE FEDERAL DE LAVRAS
DEPARTAMENTO DE ESTATÍSTICA
Programa de Pós-Graduação em Estatística e Experimentação Agropecuária

Prova do Processo Seletivo para Doutorado 2017-1

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) Na teoria de regressão linear simples, utiliza-se o método de quadrados mínimos, que consiste em: dados os pontos (x_i, y_i), com $i = 1, 2, \ldots, n$, uma reta $y = a\,x + b$ é obtida pela minimização da função de duas variáveis $S(a,b) = \sum_{i=1}^{n} \left[y_i - (a\,x_i + b) \right]^2$, isto é, resolvendo o sistema

$$\begin{cases} \dfrac{\partial S}{\partial a} = 0 \\ \dfrac{\partial S}{\partial b} = 0 \end{cases}$$

que resulta em $a = \dfrac{n \sum_{i=1}^{n} x_i y_i - \sum_{i=1}^{n} x_i \sum_{i=1}^{n} y_i}{n \sum_{i=1}^{n} x_i^2 - \left(\sum_{i=1}^{n} x_i \right)^2}$ e $b = \dfrac{1}{n} \left(\sum_{i=1}^{n} y_i - a \sum_{i=1}^{n} x_i \right)$.

Utilizando método semelhante, obter a reta da forma $y = a\,x$ que minimiza $S(a) = \sum_{i=1}^{n} (y_i - a\,x_i)^2$.

$$\begin{aligned} S(a) &= \sum_{i=1}^{n} \left[y_i - a\,x_i \right]^2 \\ \frac{d}{da} S(a) &= 2 \sum_{i=1}^{n} (y_i - a\,x_i)(-x_i) = 2 \left[-\sum_{i=1}^{n} x_i y_i + a \sum_{i=1}^{n} x_i^2 \right] \\ \frac{d}{da} S(a) &= 0 \quad \Rightarrow \quad -\sum_{i=1}^{n} x_i y_i + a \sum_{i=1}^{n} x_i^2 = 0 \\ &\Rightarrow \quad a = \frac{\sum_{i=1}^{n} x_i y_i}{\sum_{i=1}^{n} x_i^2} \qquad \textbf{(1 ponto)} \end{aligned}$$

Pontuação parcial: caso apenas explicite a função $S(a)$ **(0,3 pontos)**

2) A esperança de uma variável aleatória contínua X é dada por uma integral do tipo $E[X]=\int_{-\infty}^{+\infty} x\, f(x)\,dx$, em que $f(x)$ é a densidade da variável aleatória X, isto é, $\int_{-\infty}^{+\infty} f(x)\,dx=1$. Sabendo que $\int_{-\infty}^{+\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(x-3)^2}\,dx=1$, determine $E[X]$ para o caso em que a densidade de X é $f(x)=\frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}(x-3)^2}$.

**** QUESTÃO ANULADA ****

Todos os candidatos receberam 1,0 ponto pela questão

3) Uma cópula é uma função $C:[0,1]^2 \to [0,1]$ com as seguintes propriedades:

$$P1: \quad C(1,v)=v \qquad C(u,1)=u \qquad C(0,v)=C(u,0)=0$$

$$P2: \quad C(b,d)-C(a,d)-C(b,c)+C(a,c)\ge 0\text{, para } 0\le a<b\le 1 \text{ e } 0\le c<d\le 1.$$

A família, a um parâmetro θ, de cópulas de Gumbel é definida como $C(u,v)=\exp\left\{-\left[(-\ln u)^{\theta}+(-\ln v)^{\theta}\right]^{1/\theta}\right\}$, com $\theta\ge 1$.

Para $\theta=1$, calcule $\int_0^1 C(u,v)\,dv$

$$\theta=1 \quad \Rightarrow \quad C(u,v)=\exp\left\{-\left[(-\ln u)^{1}+(-\ln v)^{1}\right]^{1/1}\right\}$$

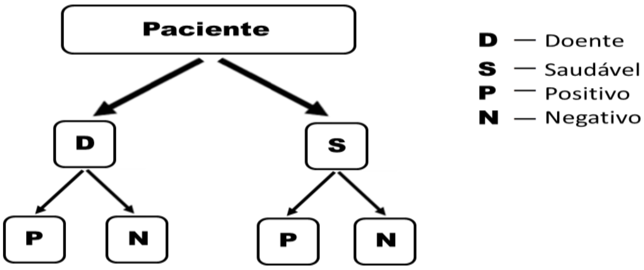
$$=\exp\{\ln u+\ln v\}=\exp\{\ln uv\}=uv$$

$$\int_0^1 C(u,v)\,dv \qquad =\int_0^1 uv\,dv=u\left[\frac{v^2}{2}\right]_0^1=\frac{u}{2} \qquad \textbf{(1 ponto)}$$

Pontuação parcial: caso apenas substitua corretamente o valor de $\theta=1$ (0,3 pts)

4) Em um exame clínico para detectar uma determinada doença, podem ocorrer os seguintes erros: o resultado do exame é negativo dado que a pessoa está doente (tal fato é denominado como falso negativo) ou o resultado do exame é positivo dado que a pessoa não está doente (falso positivo). Suponha que a probabilidade de um falso negativo seja *q* e a probabilidade de um falso positivo seja *p*. Um médico recebe o resultado do exame de um paciente como positivo. Sabendo que 1% da população apresenta a doença, qual é a probabilidade do paciente não estar doente. Considere $p = 3\%$ e $q = 2\%$.

Sugestão: caso queira, utilize uma árvore de probabilidades



Seja *D* a variável aleatória "presença ou ausência da doença no paciente", com

$$D = \begin{cases} 0 & \text{ausência da doença} \\ 1 & \text{presença da doença} \end{cases}$$

Seja *R* a variável aleatória "resultado do exame", com

$$R = \begin{cases} 0 & \text{resultado negativo} \\ 1 & \text{resultado positivo} \end{cases}$$

Para um paciente ao acaso as seguintes probabilidades são dadas:

$$\begin{cases} P\left[D=0\right]=0,99 \\ P\left[D=1\right]=0,01 \end{cases} \qquad \begin{cases} P\left[R=0 \middle| D=1\right]=0,02 \\ P\left[R=1 \middle| D=0\right]=0,03 \end{cases}$$

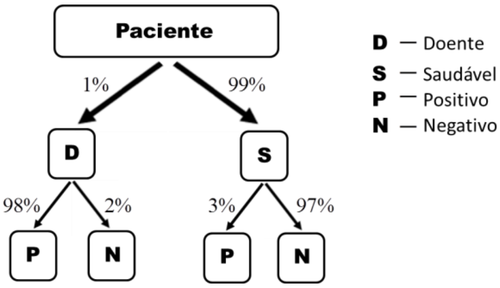
$$\Rightarrow \qquad \begin{cases} P\left[R=1 \middle| D=1\right]=0,98 \\ P\left[R=0 \middle| D=0\right]=0,97 \end{cases}$$

Queremos $\quad P\left[D=0 \middle| R=1\right]=?$

Usando a fórmula de Bayes:

$$
\begin{aligned}
P\left[D=0 \middle| R=1\right] &= \frac{P\left[R=1 \middle| D=0\right]P\left[D=0\right]}{P\left[R=1\right]} \\
&= \frac{P\left[R=1 \middle| D=0\right]P\left[D=0\right]}{P\left[R=1 \middle| D=0\right]P\left[D=0\right]+P\left[R=1 \middle| D=1\right]P\left[D=1\right]} \\
&= \frac{(0,03)(0,99)}{(0,03)(0,99)+(0,98)(0,01)} = \frac{0,0297}{0,0297+0,0098} \\
&= \frac{297}{395} = 0,7519 = 75,19\% \qquad \textbf{(1 ponto)}
\end{aligned}
$$

A resolução utilizando a árvore de probabilidades seria:



Observe que existem 4 caminhos entre o paciente e o resultado do exame:

		ESTADO DE SAÚDE		RESULTADO DO EXAME
Caminho 1: Paciente	$\xrightarrow{1\%}$	DOENTE	$\xrightarrow{98\%}$	POSITIVO
Caminho 2: Paciente	$\xrightarrow{1\%}$	DOENTE	$\xrightarrow{2\%}$	NEGATIVO
Caminho 3: Paciente	$\xrightarrow{99\%}$	SAUDÁVEL	$\xrightarrow{3\%}$	POSITIVO
Caminho 4: Paciente	$\xrightarrow{99\%}$	SAUDÁVEL	$\xrightarrow{97\%}$	NEGATIVO

Um exame positivo pode acontecer ou pelo caminho 1 ou pelo caminho 3. Portanto:

$$P\left[R=1\right]=(0,01)(0,98)+(0,99)(0,03)=0,0098+0,0297=0,0395$$

O evento "paciente saudável e exame positivo" é o caminho 3 :

$$P\left[D=0,R=1\right]=(0,99)(0,03)=0,0297$$

Portanto a probabilidade do evento condicional "paciente saudável dado que o resultado é positivo" é dada por:

$$P\left[R=1 \middle| D=0\right]=\frac{P\left[D=0,R=1\right]}{P\left[R=1\right]}=\frac{0,0297}{0,0395}=0,7519=75,19\%$$

Pontuação parcial: calcule corretamente apenas as probabilidades conjuntas: **(0,4 pontos)**

5) Se a variável aleatória X tem densidade exponencial $f_X(x) = \frac{1}{3}e^{-\frac{1}{3}x} I_{(0,+\infty)}(x)$,

determine a probabilidade de $X < 1$ ou $X > 2$, ou seja, $P\left[(X<1)\cup(X>2)\right]$.

$$\begin{aligned} P\left[(X<1)\cup(X>2)\right] &= 1 - P[1 \le X \le 2] \\ &= 1 - \int_1^2 \frac{1}{3}e^{-\frac{1}{3}x}dx = 1 + \left[e^{-\frac{1}{3}x}\right]_1^2 \\ &= 1 + \left[e^{-\frac{2}{3}} - e^{-\frac{1}{3}}\right] = 0,7969 = 79,69\% \quad \textbf{(1 ponto)} \end{aligned}$$

Pontuação parcial: caso apenas transforme probabilidades laterais em intervalar (0,3 pontos)

6) Seja $X_1,\ldots,X_{10}$ uma amostra aleatória de uma população $normal(\mu,\sigma^2)$. Sabe-se que a quantidade aleatória $Q = \dfrac{\sum_{j=1}^{10}\left(X_j - \bar{X}\right)^2}{\sigma^2}$ tem distribuição qui-quadrado com 9 graus de liberdade. Observada uma amostra, $\sum_{j=1}^{10}\left(x_j - \bar{x}\right)^2 = 81$. Sabendo que $P[Q \le 2,09] = 0,01$ e que $P[Q \le 16,9] = 0,95$, determine um intervalo de confiança a 94% para a variância σ^2.

$$\begin{cases} P[Q \le 2,09] = 0,01 \\ P[Q \le 16,9] = 0,95 \end{cases} \Rightarrow \quad P[2,09 < Q \le 16,9] = P[Q \le 16,9] - P[Q \le 2,09]$$

$$= 0,95 - 0,01 = 0,94$$

$$\Rightarrow \quad P\left[2,09 < \frac{\sum_{j=1}^{10}\left(X_j - \bar{X}\right)^2}{\sigma^2} \le 16,9\right] = 0,94$$

$$\Rightarrow \quad P\left[\frac{1}{2,09} \ge \frac{\sigma^2}{81} > \frac{1}{16,9}\right] = P\left[\frac{81}{2,09} \ge \sigma^2 > \frac{81}{16,9}\right] = 0,94$$

$$P\left[38,76 \ge \sigma^2 > 4,79\right] = P\left[4,79 < \sigma^2 \le 38,76\right] = 0,94$$

Assim, o intervalo de confiança é dado por: $(4,79;38,76]$. **(1 ponto)**

<u>Obs.:</u> Como a distribuição é contínua e, portanto, a probabilidade de um ponto é nula, o intervalo de confiança pode ser expresso também por:

$$(4,79\ ;\ 38,76) \text{ ou } [4,79\ ;\ 38,76] \text{ ou } [4,79\ ;\ 38,76).$$

Pontuação parcial:

- **Caso obtenha apenas o intervalo para a quantidade pivotal: (0,2 pts)**
- **Caso faça apenas o pivotamento: (0,4 pts)**

7) Realizada uma amostra aleatória $x_1,\ldots,x_n$ de uma população exponencial $f_X(x)=\lambda e^{-\lambda x} I_{(0,+\infty)}(x)$, obter o estimador de máxima verossimilhança para o parâmetro λ.

Função de verossimilhança: $L(\lambda|x_1,\ldots,x_n)=\prod_{j=1}^{n}\lambda e^{-\lambda x_j}=\lambda^n e^{-\lambda\sum_{j=1}^{n}x_j}$

Assim, a derivada é dada por:

$$\frac{d}{d\lambda}L(\lambda|x_1,\ldots,x_n)=n\lambda^{n-1}e^{-\lambda\sum_{j=1}^{n}x_j}-\lambda^n\sum_{j=1}^{n}x_j\,e^{-\lambda\sum_{j=1}^{n}x_j}$$
$$=\lambda^{n-1}e^{-\lambda\sum_{j=1}^{n}x_j}\left(n-\lambda\sum_{j=1}^{n}x_j\right)$$

Igualando a zero, tem-se:

$$\frac{d}{d\lambda}L(\lambda|x_1,\ldots,x_n)=0 \qquad \Rightarrow \qquad n-\lambda\sum_{j=1}^{n}x_j=0$$
$$\Rightarrow \qquad \lambda=\frac{n}{\sum_{j=1}^{n}x_j}$$

O estimador de máxima verossimilhança: $\hat{\lambda}=\frac{n}{\sum_{j=1}^{n}x_j}$ **(1 ponto)**

Pontuação parcial: caso defina apenas a função de verossimilhança (0,4 pts)

8) Um pesquisador executou um experimento, utilizando o delineamento inteiramente cassualizado com 6 repetições, onde foram comparadas 3 cultivares de feijão. Os resultados foram:

Cultivares	Produção média ($\hat{m}_i$)
1. ESAL-2	2,80
2. IAC-18	1,44
3. IAC-38	2,25

a) Suponha que o pesquisador esteja interessado em responder às seguintes perguntas:

i. Será que a cultivar ESAL-2 apresenta, em média, produção maior do que as cultivares IAC?

ii. Existe diferença significativa entre as produções das duas cultivares IAC?

Obtenha um grupo de contrastes mutuamente ortogonais que sejam adequados para proceder as comparações de interesse do pesquisador e obtenha a estimativa de cada um dos contrastes.

(i) $Y_1 = 2m_1 - m_2 - m_3 \quad \Rightarrow \quad \hat{Y}_1 = 2 \cdot 2{,}80 - 1{,}44 - 2{,}25 = 1{,}91$ **(0,1 pts)**

(ii) $Y_2 = m_2 - m_3 \quad \Rightarrow \quad \hat{Y}_2 = 1{,}44 - 2{,}25 = -0{,}81$ **(0,1 pts)**

b) Desse mesmo experimento, é fornecido a você as seguintes informações:

- O teste F da análise de variância foi significativo;
- *QMErro* = 0,61.

Aplique o teste F aos contrastes formulados por você. Formule as hipóteses e conclua corretamente.

Tabela da Análise de Variância dos contrastes

FV	GL	SQ	QM	F	
ESAL vs IAC	1	3,6481	3,6481	$5{,}98^{*}$	**(0,05 pts)**
IAC-18 vs IAC-38	1	1,9683	1,9683	$3{,}23^{ns}$	**(0,05 pts)**
Erro	15	-	0,61		

*Significativo a 5%; ^{ns} Não significativo.

$SQY_1 = \dfrac{\hat{Y}_1^2}{\sum_{i=1}^{3} c_i^2} \cdot r = \dfrac{1{,}91^2}{6} \cdot 6 = 3{,}6481$ **(0,1 pts)**

$SQY_2 = \dfrac{\hat{Y}_2^2}{\sum_{i=1}^{3} c_i^2} \cdot r = \dfrac{(-0{,}81)^2}{2} \cdot 6 = 1{,}9683$ **(0,1 pts)**

$F_{1\%}(1{,}15) = 8{,}68 \; ; \; F_{5\%}(1{,}15) = 4{,}54$ **(0,1 pts)**

- <u>Teste F</u>:

(i) Contraste 1: **(0,2 pts)**

$\begin{cases} H_0: Y_1 = 0 \\ H_a: Y_1 \neq 0 \end{cases}$

$F_c > F_{5\%} \Rightarrow$ Rejeita-se H_0 ao nível de 5% de probabilidade pelo teste F. Portanto, o contraste Y_1 é estatisticamente diferente de zero, ou seja, a produção da cultivar ESAL-2 é, em média, superior à produção das cultivares IAC.

(ii) Contraste 2: **(0,2 pts)**

$\begin{cases} H_0: Y_2 = 0 \\ H_a: Y_2 \neq 0 \end{cases}$

$F_c < F_{5\%} \Rightarrow$ Não se rejeita H_0 ao nível de 5% de probabilidade pelo teste F. Portanto, o contraste Y_1 é estatisticamente nulo, ou seja, não existe diferença significativa entre as produções das duas cultivares IAC.

<u>DADOS:</u> $Y = c_1\mu_1 + c_2\mu_2 + \cdots + c_I\mu_I$; $SQY = \frac{\hat{Y}^2}{\sum_i c_i^2} \cdot r$, em que r é o número de repetições.

9) Para se verificar a aceitação de duas bebidas derivadas do leite (iogurte e bebida láctea) com sabores de morango, pêssego, abacaxi e côco, foi instalado um experimento fatorial 2x4, utilizando um delineamento inteiramente casualizado com 4 repetições. A pesquisa foi conduzida em um colégio e a aceitação foi avaliada por meio da escala hedônica facial, classificando-se o alimento em terrível, não muito bom, indiferente, bom e excelente, numa escala de 1 a 5. Pede-se:

a) Apresente o modelo estatístico referente a esse experimento, descrevendo adequadamente cada termo do modelo. **(0,2 pts)**

$$y_{ijk} = \mu + \alpha_i + \beta_j + \alpha\beta_{ij} + e_{ijk} \ ,$$

em que

y_{ijk} é a nota obtida na repetição k da bebida derivada do leite i e do sabor j;
μ é uma constante associada a cada observação;
α_i é o efeito da *i-ésima* bebida derivada do leite, com $i = 1, 2$;
β_j é o efeito do *j-ésimo* sabor, com $j = 1, 2, 3, 4$;
$\alpha\beta_{ij}$ é o efeito da interação entre os fatores tipo de bebida e sabor;
e_{ijk} é o erro associado a cada observação, sendo independentes com $e_{ijk} \sim N(0, \sigma^2)$.

b) Desse experimento é apresentado a você as seguintes informações:

i. De acordo com o teste F da Análise de Variância, verificou-se que os fatores em estudo não atuam independentemente;

ii. Na Análise de Variância obteve-se $QMErro = 0{,}21$;

iii. O quadro auxiliar com os totais das notas de cada tratamento é:

Tipo de bebida	Sabor				Totais
	Morango	Pêssego	Abacaxi	Côco	
Iogurte	17	15	14	17	63
Bebida láctea	15	11	12	9	47
Totais	32	26	26	26	110

Suponhamos que o pesquisador esteja interessado em saber se houve diferença significativa na aceitação dos diferentes sabores de bebida láctea e qual(is) desses sabor(es) teve(tiveram) melhor aceitação. É possível responder a essas perguntas do pesquisador com as informações apresentadas? Se não for possível, justifique. Se for possível, responda ao pesquisador, utilizando os testes F e o teste de Student-Newman-Keuls (SNK), caso sejam necessários, concluindo corretamente.

Sim, basta fazer o estudo do desdobramento de sabores dentro de bebida láctea, como se segue.

Tabela da Análise de Variância **(0,3 pts)**

FV	GL	SQ	QM	F
Sabor (bebida láctea)	3	4,6875	1,5625	7,44**
Erro	24	-	0,21	

** Significatico a 1%

$$SQSabor(beb.láctea) = \frac{1}{4}(15^2 + 11^2 + 12^2 + 9^2) - \frac{47^2}{16} = 4{,}6875$$

- <u>Teste F</u>: **(0,2 pts)**

$$\begin{cases} H_0: m_{\text{Morango}} = m_{\text{Pêssego}} = m_{\text{Abacaxi}} = m_{\text{Côco}} \\ H_a: \text{Existe pelo menos uma diferença entre as médias dos sabores de bebida láctea} \end{cases}$$

$F_{1\%}(3, 24) = 4{,}72$; $F_{5\%}(3, 24) = 3{,}01$

$F_c < F_{1\%} \Rightarrow$ Rejeita-se H_0 ao nível de 1% de probabilidade pelo teste F. Portanto, existe diferença significativa na aceitação dos diferentes sabores de bebida láctea.

- <u>Teste de SNK (1%)</u>: **(0,3 pts)**

$\hat{m}_{\text{Morango}} = 3{,}75$ a
$\hat{m}_{\text{Abacaxi}} = 3{,}00$ ab
$\hat{m}_{Pêssego} = 2{,}75$ ab
$\hat{m}_{\text{Côco}} = 2{,}25$ b

$$DMS_4 = 4{,}91\sqrt{\frac{0{,}21}{4}} = 1{,}13;\ \ DMS_3 = 4{,}55\sqrt{\frac{0{,}21}{4}} = 1{,}04;\ \ DMS_2 = 3{,}96\sqrt{\frac{0{,}21}{4}} = 0{,}91$$

Médias seguidas de pelo menos uma mesma letra não diferem entre si ao nível de 1% de probabilidade pelo teste de SNK. Portanto, os sabores de bebida láctea que apresentaram melhor aceitação foram Morango, Abacaxi e Pêssego.

ou

- <u>Teste de SNK (5%)</u>:

$\hat{m}_{\text{Morango}} = 3{,}75$ a
$\hat{m}_{\text{Abacaxi}} = 3{,}00$ b
$\hat{m}_{Pêssego} = 2{,}75$ b
$\hat{m}_{\text{Côco}} = 2{,}25$ b

$$DMS_4 = 3{,}90\sqrt{\frac{0{,}21}{4}} = 0{,}89;\ \ DMS_3 = 3{,}53\sqrt{\frac{0{,}21}{4}} = 0{,}81;\ \ DMS_2 = 2{,}92\sqrt{\frac{0{,}21}{4}} = 0{,}67$$

Médias seguidas de pelo menos uma mesma letra não diferem entre si ao nível de 5% de probabilidade pelo teste de SNK. Portanto, o sabor de bebida láctea que apresentou melhor aceitação foi Morango.

<u>DADOS:</u> $DMS_k = K_k = q_{(\alpha;\,k;\,\text{GLErro})}\sqrt{\frac{QMErro}{r}}$

10) Um experimento foi executado para verificar o desgaste de pneus de diferentes modelos e marcas. O experimento foi conduzido utilizando-se o delineamento em blocos casualizados, medindo-se o desgaste de quatro tipos de pneus (1. Perilla 1; 2. Perilla 2; 3. Feristane 1; 4. Goodday 1), após um certo tempo de uso e quilometragem rodada. Cada parcela consistiu de um conjunto de quatro pneus e os desgastes médios desses pneus, dado em *mm*, foram:

	Blocos					
Pneu	1	2	3	4	5	Totais
1. Perilla 1	2,5	2,5	1,8	1,9	2,2	10,9
2. Perilla 2	1,5	3,1	1,2	3,4	2,3	11,5
3. Feristane 1	2,0	1,9	3,1	3,0	2,5	12,5
4. Goodday 1	1,9	2,9	2,0	2,8	2,4	12,0
Totais	7,9	10,4	8,1	11,1	9,4	46,9

Pede-se:

a) Quais as hipóteses serão avaliadas com o teste F da Análise de Variância?

$$\begin{cases} H_0: m_{\text{Perilla 1}} = m_{\text{Perilla 2}} = m_{\text{Feristane 1}} = m_{\text{Goodday 1}} \\ H_a: \text{Existe pelo menos uma diferença entre as médias dos pneus} \end{cases}$$ **(0,15 pts)**

b) Efetuar a Análise de Variância e concluir corretamente, discutindo sobre a precisão do experimento. (**DADO:** $\boldsymbol{SQTotal}$ **=6,45**)

Tabela da Análise de Variância **(0,35 pts)**

FV	GL	SQ	QM	F
Blocos	4	1,9570	0,4892	1,39 ^{ns}
Tratamentos	3	0,2815	0,0938	0,27 ^{ns}
Erro	12	4,2110	0,3509	
Total	19	6,4495		

^{ns} Não significativo

$F_{5\%}(3, 12) = 3{,}49$

$F_c < F_{5\%} \Rightarrow$ Não se rejeita H_0 ao nível de 5% de probabilidade pelo teste F. Portanto, não existe diferença significativa no desgaste das diferentes marcas e modelos de pneus.

- Precisão do experimento: **(0,10 pts)**

$$CV = \frac{\sqrt{QMErro}}{\hat{m}} \cdot 100 = \frac{\sqrt{0{,}3509}}{2{,}345} \cdot 100 = 26{,}63\%$$

O experimento é de baixa precisão, uma vez que o CV está entre 20 e 30%.

c) Apresente o modelo linear de Gauss-Markov, na forma matricial, desse experimento e apresente os vetores e a matriz desse modelo. **(0,20 pts)**

$$\boldsymbol{Y} = \boldsymbol{X\beta} + \boldsymbol{\varepsilon}\,,$$

em que

$\boldsymbol{Y}$ é o vetor de observações, de ordem 20x1;
$\boldsymbol{X}$ é a matriz de incidência dos efeitos do modelo, de ordem 20x10;
$\boldsymbol{\beta}$ é o vetor de efeitos do modelo, de ordem 10x1;
$\boldsymbol{\varepsilon}$ é o vetor de erros, de ordem 20x1, de $\boldsymbol{\varepsilon} \sim N(\boldsymbol{\phi}, \boldsymbol{I}\sigma^2)$.

As matrizes e vetores são dados por:

$$\boldsymbol{Y} = \boldsymbol{X\beta} + \boldsymbol{\varepsilon}$$

$$\begin{bmatrix} 2{,}5 \\ 2{,}5 \\ 1{,}8 \\ 1{,}9 \\ 2{,}2 \\ 1{,}5 \\ 3{,}1 \\ 1{,}2 \\ 3{,}4 \\ 2{,}3 \\ 2{,}0 \\ 1{,}9 \\ 3{,}1 \\ 3{,}0 \\ 2{,}5 \\ 1{,}9 \\ 2{,}9 \\ 2{,}0 \\ 2{,}8 \\ 2{,}4 \end{bmatrix} = \begin{bmatrix} 1 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 1 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 1 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 1 & 0 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 1 \\ 1 & 0 & 0 & 0 & 1 & 1 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 1 & 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 1 & 0 & 0 & 1 & 0 & 0 \\ 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 1 & 0 \\ 1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} \mu \\ \alpha_1 \\ \alpha_2 \\ \alpha_3 \\ \alpha_4 \\ \beta_1 \\ \beta_2 \\ \beta_3 \\ \beta_4 \\ \beta_5 \end{bmatrix} + \begin{bmatrix} e_{11} \\ e_{12} \\ e_{13} \\ e_{14} \\ e_{15} \\ e_{21} \\ e_{22} \\ e_{23} \\ e_{24} \\ e_{25} \\ e_{31} \\ e_{32} \\ e_{33} \\ e_{34} \\ e_{35} \\ e_{41} \\ e_{42} \\ e_{43} \\ e_{44} \\ e_{45} \end{bmatrix}$$

d) Defina as matrizes e os vetores do sistema de equações normais (*SEN*) $\boldsymbol{X}'\boldsymbol{X}\boldsymbol{\beta} = \boldsymbol{X}'\boldsymbol{y}$ referente a esse experimento. **(0,20 pts)**

$$\boldsymbol{X}'\boldsymbol{X}\boldsymbol{\beta} = \boldsymbol{X}'\boldsymbol{Y},$$

em que

$$\boldsymbol{X}'\boldsymbol{X} = \begin{bmatrix} 20 & 5 & 5 & 5 & 5 & 4 & 4 & 4 & 4 & 4 \\ 5 & 5 & 0 & 0 & 0 & 1 & 1 & 1 & 1 & 1 \\ 5 & 0 & 5 & 0 & 0 & 1 & 1 & 1 & 1 & 1 \\ 5 & 0 & 0 & 5 & 0 & 1 & 1 & 1 & 1 & 1 \\ 5 & 0 & 0 & 0 & 5 & 1 & 1 & 1 & 1 & 1 \\ 4 & 1 & 1 & 1 & 1 & 4 & 0 & 0 & 0 & 0 \\ 4 & 1 & 1 & 1 & 1 & 0 & 4 & 0 & 0 & 0 \\ 4 & 1 & 1 & 1 & 1 & 0 & 0 & 4 & 0 & 0 \\ 4 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 4 & 0 \\ 4 & 1 & 1 & 1 & 1 & 0 & 0 & 0 & 0 & 4 \end{bmatrix}; \ \boldsymbol{\beta} = \begin{bmatrix} \mu \\ \alpha_1 \\ \alpha_2 \\ \alpha_3 \\ \alpha_4 \\ \beta_1 \\ \beta_2 \\ \beta_3 \\ \beta_4 \\ \beta_5 \end{bmatrix}; \ \boldsymbol{X}'\boldsymbol{Y} = \begin{bmatrix} 46,9 \\ 10,9 \\ 11,6 \\ 12,5 \\ 12,0 \\ 7,9 \\ 10,4 \\ 8,1 \\ 11,1 \\ 9,4 \end{bmatrix}$$

TABELAS

Tabela 1 – Limites unilaterais de F ao nível de 5% de probabilidade

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

n_1 = número de graus de liberdade do numerador e n_2 = número de graus de liberdade do denominador

Tabela 2 – Limites unilaterais de F ao nível de 1% de probabilidade

$n_2 \backslash n_1$	1	2	3	4	5	6	7	8	9	10
1	4052	5000	5403	5625	5764	5859	5928	5982	6022	6056
2	98,50	99,00	99,17	99,25	99,30	99,33	99,36	99,37	99,39	99,40
3	34,12	30,82	29,46	28,71	28,24	27,91	27,67	27,49	27,35	27,23
4	21,20	18,00	16,69	15,98	15,52	15,21	14,98	14,80	14,66	14,55
5	16,26	13,27	12,06	11,39	10,97	10,67	10,46	10,29	10,16	10,05
6	13,75	10,92	9,78	9,15	8,75	8,47	8,26	8,10	7,98	7,87
7	12,25	9,55	8,45	7,85	8,46	7,19	6,99	6,84	6,72	6,62
8	11,26	8,65	7,59	7,01	6,63	6,37	6,18	6,03	5,91	5,81
9	10,56	8,02	6,99	6,42	6,06	5,80	5,61	5,47	5,35	5,26
10	10,04	7,56	6,55	5,99	5,64	5,39	5,20	5,06	4,94	4,85
11	9,65	7,21	6,22	5,67	5,32	5,07	4,89	4,74	4,63	4,54
12	9,33	6,93	5,95	5,41	5,06	4,82	4,64	4,50	4,39	4,30
13	9,07	6,70	5,74	5,21	4,86	4,62	4,44	4,30	4,19	4,10
14	8,86	6,51	5,56	5,04	4,69	4,46	4,28	4,14	4,03	3,94
15	8,68	6,36	5,42	4,89	4,56	4,32	4,14	4,00	3,89	3,80
16	8,53	6,23	5,29	4,77	4,44	4,20	4,03	3,89	3,78	3,69
17	8,40	6,11	5,18	4,67	4,34	4,10	3,93	3,79	3,68	3,59
18	8,29	6,01	5,09	4,58	4,25	4,01	3,84	3,71	3,60	3,51
19	8,18	5,93	5,01	4,50	4,17	3,94	3,77	3,63	3,52	3,43
20	8,10	5,85	4,94	4,43	4,10	3,87	3,70	3,56	3,46	3,37
21	8,02	5,78	4,87	4,37	4,04	3,81	3,64	3,51	3,40	3,31
22	7,95	5,72	4,82	4,31	3,99	3,76	3,59	3,45	3,35	3,26
23	7,88	5,66	4,76	4,26	3,94	3,71	3,54	3,41	3,30	3,21
24	7,82	5,61	4,72	4,22	3,90	3,67	3,50	3,36	3,26	3,17
25	7,77	5,57	4,68	4,18	3,85	3,63	3,46	3,32	3,22	3,13

n_1 = número de graus de liberdade do numerador e n_2 = número de graus de liberdade do denominador

Tabela 3 - Valores da amplitude total estudentizada (q), para uso no teste de Tukey, ao nível de 5% de probabilidade.

$n_2 \backslash n_1$	2	3	4	5	6	7	8	9	10
1	17,97	26,98	32,82	37,08	40,41	43,12	45,40	47,36	49,07
2	6,09	8,33	9,80	10,88	11,74	12,44	13,03	13,54	13,99
3	4,50	5,91	6,83	7,50	8,04	8,48	8,85	9,18	9,46
4	3,93	5,04	5,76	6,29	6,71	7,05	7,35	7,60	7,83
5	3,64	4,60	5,22	5,67	6,03	6,33	6,58	6,80	7,00
6	3,46	4,34	4,90	5,31	5,63	5,90	6,12	6,32	6,49
7	3,34	4,17	4,68	5,06	5,36	5,61	5,82	6,00	6,16
8	3,26	4,04	4,53	4,89	5,17	5,40	5,60	5,77	5,92
9	3,20	3,95	4,42	4,76	5,02	5,24	5,43	5,60	5,74
10	3,15	3,88	4,33	4,65	4,91	5,12	5,31	5,46	5,60
11	3,11	3,82	4,26	4,57	4,82	5,03	5,20	5,35	5,49
12	3,08	3,77	4,20	4,51	4,75	4,95	5,12	5,27	5,40
13	3,06	3,74	4,15	4,45	4,69	4,89	5,05	5,19	5,32
14	3,03	3,70	4,11	4,41	4,64	4,83	4,99	5,13	5,25
15	3,01	3,67	4,08	4,37	4,60	4,78	4,94	5,08	5,20
16	3,00	3,65	4,05	4,33	4,56	4,74	4,90	5,03	5,15
17	2,98	3,63	4,02	4,30	4,52	4,71	4,86	4,99	5,11
18	2,97	3,61	4,00	4,28	4,50	4,67	4,82	4,96	5,07
19	2,96	3,59	3,98	4,25	4,47	4,65	4,79	4,92	5,04
20	2,95	3,58	3,96	4,23	4,45	4,62	4,77	4,90	5,01
24	2,92	3,53	3,90	4,17	4,37	4,54	4,68	4,81	4,92
30	2,89	3,49	3,85	4,10	4,30	4,46	4,60	4,72	4,82
40	2,86	3,44	3,79	4,04	4,23	4,39	4,52	4,64	4,74
60	2,83	3,40	3,74	3,98	4,16	4,31	4,44	4,55	4,65
120	2,80	3,36	3,69	3,92	4,10	4,24	4,36	4,47	4,56

n_1 = número de níveis do fator em estudo e n_2 = número de graus de liberdade do resíduo

Tabela 4 - Valores da amplitude total estudentizada (q), para uso no teste de Tukey, ao nível de 1% de probabilidade.

$n_2 \backslash n_1$	2	3	4	5	6	7	8	9	10
1	90,03	135,0	164,3	185,6	202,2	215,8	227,2	237,0	245,6
2	14,04	19,02	22,29	24,72	26,63	28,20	29,53	30,68	31,69
3	8,26	10,62	12,17	13,33	14,24	15,00	15,64	16,20	16,69
4	6,51	8,12	9,17	9,96	10,58	11,10	11,55	11,93	12,27
5	5,70	6,98	7,80	8,42	8,91	9,32	9,70	9,97	10,24
6	5,24	6,33	7,03	7,56	7,97	8,32	8,61	8,87	9,10
7	4,95	5,92	6,54	7,00	7,37	7,68	7,94	8,17	8,37
8	4,75	5,64	6,20	6,62	6,96	7,24	7,47	7,68	7,86
9	4,60	5,43	5,96	6,35	6,66	6,92	7,13	7,32	7,50
10	4,48	5,27	5,77	6,14	6,43	6,67	6,88	7,06	7,21
11	4,39	5,15	5,62	5,97	6,25	6,48	6,67	6,84	6,99
12	4,32	5,05	5,50	5,84	6,10	6,32	6,51	6,67	6,81
13	4,26	4,96	5,40	5,73	5,98	6,19	6,37	6,53	6,67
14	4,21	4,90	5,32	5,63	5,88	6,08	6,26	6,41	6,54
15	4,17	4,84	5,25	5,56	5,80	5,99	6,16	6,31	6,44
16	4,13	4,79	5,19	5,49	5,72	5,92	6,08	6,22	6,35
17	4,10	4,74	5,14	5,43	5,66	5,85	6,01	6,15	6,27
18	4,07	4,70	5,09	5,38	5,60	5,79	5,94	6,08	6,20
19	4,05	4,67	5,05	5,33	5,55	5,74	5,89	6,02	6,14
20	4,02	4,64	5,02	5,29	5,51	5,69	5,84	5,97	6,09
24	3,96	4,55	4,91	5,17	5,37	5,54	5,68	5,81	5,92
30	3,89	4,46	4,80	5,05	5,24	5,40	5,54	5,65	5,76
40	3,82	4,37	4,70	4,93	5,11	5,26	5,39	5,50	5,60
60	3,76	4,28	4,60	4,82	4,99	5,13	5,25	5,36	5,45
120	3,70	4,20	4,50	4,71	4,87	5,00	5,12	5,21	5,30

n_1 = número de níveis do fator em estudo e n_2 = número de graus de liberdade do resíduo