UFV- CCE - DET

EST 105 - 1^a avaliação - 1^0 semestre de 2016 - 9/abr/16

Nome: ______________________________ **Matrícula:** ____________________

Assinatura: ______________________________. Favor apresentar documento com foto.

- São 5 questões e formulário em páginas numeradas de 1 a 8, total de 30 pontos, FAVOR CONFERIR ANTES DE INICIAR.
- ATENÇÃO: Assinale (X) a seguir em qual turma está matriculado (sua nota será divulgada no sistema SAPIENS).

```
-------------------------------------------------------------------------------
      TURMA    HORÁRIO     SALA    PROFESSOR
-------------------------------------------------------------------------------
(    ) T1 3ª 08-10 5ª 10-12 PVB300 Sebastião
(    ) T2 3ª 10-12 6ª 08-10 PVB109 Sebastião
(    ) T3 3ª 14-16 5ª 16-18 PVB109 CHOS
(    ) T4 2ª 14-16 4ª 16-18 PVB107 Policarpo
(    ) T5 2ª 18:30-20:10 4ª 20:30-22:10 PVB208 Camila
(    ) T6 4ª 14-16 6ª 16-18 PVA361 CHOS
(    ) T7 2ª 16-18 5ª 14-16 PVB307 Sebast/Policarpo
(    ) T10 3ª 18:30-20:10 5ª 20:30-22:10 PVA361 Camila
(    ) T20 = EST085 T1 2ª16 PVA134 e T2 5ª 18:30 PVA126 Leísa(monitor II)
--------------------------------------------------------------------------------
```

- Interpretar corretamente as questões é parte da avaliação, portanto não é permitido questionamentos durante a prova !
- É OBRIGATÓRIO APRESENTAR OS CÁLCULOS organizadamente, para ter direito à revisão.
- PODE UTILIZAR A CALCULADORA, porém mostre os valores utilizados na fórmula.
- BOA SORTE e BOA PROVA !!!.

FORMULÁRIO

$$\overline{X} = \frac{\sum_{i=1}^{n} X_i}{n} \quad \text{ou} \quad \overline{X} = \frac{\sum_{i=1}^{k} f_i X_i}{\sum_{i=1}^{k} f_i} \qquad Md_X = \frac{X_{\left(\frac{n}{2}\right)} + X_{\left(\frac{n}{2}+1\right)}}{2} \quad \text{ou} \quad Md_X = X_{\left(\frac{n+1}{2}\right)}$$

$$\overline{X}_H = \frac{n}{\sum_{i=1}^{n} \frac{1}{X_i}} \quad \text{ou} \quad \overline{X}_H = \frac{\sum_{i=1}^{k} f_i}{\sum_{i=1}^{k} \frac{f_i}{X_i}} \qquad \overline{X}_G = \sqrt[n]{\prod_{i=1}^{n} X_i} \quad \text{ou} \quad \overline{X}_G = \sqrt[\sum f_i]{\prod_{i=1}^{k} X_i^{f_i}}$$

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

$$S_X^2 = \frac{SQD_X}{n-1} \quad \text{ou} \quad S_X^2 = \frac{SQD_X}{\sum_{i=1}^{k} f_i - 1}$$

$$S_X = \sqrt{S_X^2} \quad S(\overline{X}) = \frac{S_X}{\sqrt{n}} \quad CV_X(\%) = \frac{S_X}{\overline{X}} 100\%$$

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

$$Y_i = \beta_0 + \beta_1 X_i + \varepsilon_i \qquad \widehat{\varepsilon}_i = Y_i - \widehat{Y}_i$$

$$\widehat{Y}_i = \widehat{\beta}_0 + \widehat{\beta}_1 X_i = b_0 + b_1 X_i \qquad b_1 = \widehat{\beta}_1 = \frac{SPD_{XY}}{SQD_X} = r_{XY} \frac{S_Y}{S_X} \qquad b_0 = \widehat{\beta}_0 = \overline{Y} - \widehat{\beta}_1 \overline{X}$$

$$r^2(\%) = \frac{SQ\text{regressão}}{SQ\text{total}} 100\%$$

$$SQ\text{regressão} = \widehat{\beta}_1^2 SQD_X = \widehat{\beta}_1 SPD_{XY} = (SPD_{XY})^2 / SQD_X \qquad SQ\text{total} = SQD_Y$$

1.(4 pontos) Dado que,

$$\sum_{j=1}^{10} Y_j = 18 \quad \text{e} \quad \sum_{\substack{i=5 \\ i\neq 10,20}}^{30} X_i = 20$$

utilize as propriedades de somatório e calcule: $\displaystyle\sum_{j=1}^{10} \sum_{\substack{i=5 \\ i\neq 10,20}}^{30} (X_i - 5 + 3Y_j)$.

$NT_j = 10$ e $NT_i = (30-5)+1-2 = 24$

$$\begin{aligned}
\sum_{j=1}^{10} \sum_{\substack{i=5 \\ i\neq 10,20}}^{30} (X_i - 5 + 3Y_j) &= \sum_j \sum_i X_i - \sum_j \sum_i 5 + \sum_j \sum_i 3Y_j \\
&= 10 \sum_i X_i - 10 \times 24 \times 5 + 3 \times 24 \sum_j Y_j \\
&= 10 \times 20 - 1200 + 72 \times 18 \\
&= 200 - 1200 + 1296 \\
&= 296
\end{aligned}$$

2.(4 pontos) Dado que,

$$\sum_{s=1}^{n} s = \frac{n(n+1)}{2}, \qquad \sum_{s=1}^{n} s^2 = \frac{n(n+1)(2n+1)}{6} \qquad \text{e} \qquad \sum_{s=1}^{n} s^3 = \left[\frac{n(n+1)}{2}\right]^2,$$

utilize as propriedades de somatório e calcule: $\displaystyle\sum_{k=1}^{5}\sum_{i=1}^{20}\sum_{j=1}^{10}\left[(i-1)^2(j+k)\right]$.

Solução 1:

$$\begin{aligned}
\sum_{k=1}^{5}\sum_{i=1}^{20}\sum_{j=1}^{10}\left[(i-1)^2(j+k)\right] &= \sum_{i}(i-1)^2\sum_{j}\sum_{k}(j+k) = \sum_{i}\left(i^2-2i+1\right)\left(5\sum_{j} j + 10\sum_{k} k\right) \\
&= \left(\sum_{i} i^2 - 2\sum_{i} i + 20\right)\left(5\sum_{j} j + 10\sum_{k} k\right) \\
&= \left(\frac{20\times21\times41}{6} - \cancel{2}\left(\frac{20\times21}{\cancel{2}}\right) + 20\right)\left(5\times\frac{10\times11}{2} + \cancel{10}^{5}\times\frac{5\times6}{\cancel{2}}\right) \\
&= (2870-420+20)(275+150) = 2470\times425 \\
&= 1049750
\end{aligned}$$

Solução 2:

$$\begin{aligned}
\sum_{k=1}^{5}\sum_{i=1}^{20}\sum_{j=1}^{10}\left[(i-1)^2(j+k)\right] &= \sum_{k}\sum_{i}\sum_{j}\left[\left(i^2-2i+1\right)(j+k)\right] \\
&= \sum_{k}\sum_{i}\sum_{j}\left[\left(i^2j+i^2k-2ij-2ik+j+k\right)\right] \\
&= 5\sum_{i} i^2\sum_{j} j + 10\sum_{i} i^2\sum_{k} k - 2\times5\sum_{i} i\sum_{j} j \\
&- 2\times10\sum_{i} i\sum_{k} k + 5\times20\sum_{j} j + 20\times10\sum_{k} k \\
&= 5\times2870\times55 + 10\times2870\times15 - 10\times210\times55 \\
&- 20\times210\times15 + 100\times55 + 200\times15 \\
&= 789250 + 430500 - 115500 - 63000 + 5500 + 3000 \\
&= 1049750
\end{aligned}$$

3.(5 pontos) Nos itens a seguir assinale (V) se estiver inteiramente correto, ou, assinale (F) caso contrário e **indique e corrija onde estiver errado** (1 ponto cada item).

a.(**F**) A média geométrica dos valores $\{3,\ 7,\ 12,\ 25\ \}$ é igual a $10,91$.

$$\bar{X}_G = \sqrt[4]{6300} \cong 8,91$$

A média geométrica dos valores $\{3,\ 7,\ 12,\ 25\ \}$ é igual a $8,91$.

b.(**F**) Para a amostra de valores $\{10, 10, 10, 10, 15, 15, 2, 2, 2, 2, 0, 30, 30\}$, o valor mediano é igual a 15.

Solução:

$$n = 13 \Rightarrow Md_X = X_{\left(\frac{13+1}{2}\right)} = X_{(7)} = 10$$

X_i	f_i	$\sum f_i$	
0	1	1	
2	4	5	
10	4	9	$\Rightarrow X_{(7)} = 10$
15	2	11	
30	2	13	

Observação: ocorreu um erro de impressão na questão, de modo que a vírgula separando o 15 do 2, ficou como um 15.2. Portanto, pode haver resposta com $n = 12$ com o valor 15.2 na amostra. Neste caso,

$$Md_X = \frac{X_{(6)} + X_{(7)}}{2} = \frac{10 + 10}{2} = 10.$$

c.(**F**) A média harmônica dos valores $\{60,\ 75,\ 90\ \}$ é igual a 75.

$$\bar{X}_G = \frac{3}{\frac{1}{60} + \frac{1}{75} + \frac{1}{90}} \cong 72,97.$$

A média harmônica dos valores $\{60,\ 75,\ 90\ \}$ é igual a $72,97$.

d.(**F**) O coeficiente de determinação simples, r_{XY}, estima o grau de relacionamento linear entre as variáveis X e Y. O coeficiente de correlação (ou correlação linear ou de Pearson), r_{XY}, estima o grau de relacionamento linear entre as variáveis X e Y.

e.(**F**) A Estatística Descritiva inclui a elaboração de tabelas e gráficos e também um resumo dos dados por meio de medidas descritivas de posição e de dispersão. Trata também das condições sob as quais as inferências sao válidas.

A Estatística Descritiva inclui a elaboração de tabelas e gráficos e também um resumo dos dados por meio de medidas descritivas de posição e de dispersão. Não trata das condições sob as quais as inferências sao válidas.

ou

A Estatística Descritiva inclui a elaboração de tabelas e gráficos e também um resumo dos dados por meio de medidas descritivas de posição e de dispersão.

4.(10 pontos) A tabela seguinte mostra valores para o resíduo de cloro (Y em partes por milhão) em uma piscina em vários momentos (número de horas, X), após ter sido tratada com produtos químicos (obtida de Simon & Freund, Estatística Aplicada, 9^a ed.)

i	1	2	3	4	5	6	7
Número de horas (X_i)	0	2	4	6	8	10	12
Resíduo de Cloro (Y_i)	2,2	1,8	1,5	1,4	1,1	1,1	0,90

Pede-se:

a.(1 pt) Utilize sua calculadora para informar as seguintes somas:

$$\sum X = 42 \qquad \sum X^2 = 364 \qquad \sum Y = 10$$

$$\sum Y^2 = 15,52 \qquad \sum XY = 48,6$$

b.(2 pts) Apresente o modelo de regressão linear simples (RLS) ajustado que permita estimar o valor médio do resíduo de cloro como uma função do número de horas após o tratamento (aproxime a segunda casa após a vírgula, por ex: 0,925=0,93 ou 4,106=4,11).

Pode apresentar os valores direto da calculadora.

$$b_1 = \frac{SPD_{XY}}{SQD_X} = \frac{48,6 - \frac{42\times 10}{7}}{364 - \frac{42^2}{7}} = -\frac{11,4}{112} \cong -0,102 = -0,10$$

$$b_0 = \bar{Y} - b_1\bar{X} = \frac{10}{7} - (-0,10)\,\frac{42}{7} \cong 2,028 = 2,03$$

Modelo ajustado

$$\widehat{Y}_i = 2,03 - 0,10X_i.$$

c.(2 pts) Apresente o valor estimado e o desvio da regressão quando $i = 4$.

$$i = 4 \Rightarrow X_i = X_4 = 6$$

$$\widehat{Y}_4 = 2,03 - 0,10 \times 6 = 1,43 \qquad (\widehat{Y}_{\text{calculadora}} \cong 1,428 = 1,43)$$

$$\widehat{\varepsilon}_4 = Y_4 - \widehat{Y}_4 = 1,4 - 1,43 = -0,03 \quad (\text{“negativo”})$$

d.(3 pts) Interprete as duas estimativas dos parâmetros do modelo ajustado.

$b_1 = \widehat{\beta}_1 = -0,10$. Para cada hora de acréscimo no tempo após o tratamento, estima-se decréscimo médio de $0,10$ p.p.m. de cloro.

$b_0 = \widehat{\beta}_0 = 2,03$. Para $X = 0$, no instante inicial, estima-se em média 2,03 p.p.m. de cloro.

e.(2 pts) Calcule e interprete a estimativa do coeficiente de determinação simples.

$$r^2(\%) = (r_{XY})^2 \times 100\% \cong (-0,9696)^2 \times 100\% \cong 94,01\%.$$

Pela calculadora:

$$\begin{aligned}
SQ_{\text{Reg}} &= b_1^2\text{SQD}_X \cong (-0,10)^2 \times 112 = 1,12 \qquad (1,16 \text{ pela calculadora})\\
SQ_{\text{Tot}} &= \text{SQD}_Y = 15,52 - \frac{10^2}{7} \cong 1,23 \qquad (1,23 \text{ pela calculadora})\\
r^2(\%) &= \frac{1,12}{1,23} \times 100\% = 91,01\%.
\end{aligned}$$

Porcentagem da variabilidade observada no resíduo de cloro que foi explicada pela regressão linear simples nos valores do número de horas após o tratamento.

5.(7 pontos) Na tabela abaixo são informados os valores do número de vezes que um componente falhou, para uma amostra aleatória de 50 componentes de dois fabricantes, A e B.

Número de	Fabricante											
	A						B					
Falhas	0	1	2	3	4	5	0	1	2	3	4	5
Componentes	25	12	5	4	2	2	17	15	9	6	2	1

Pede-se:

a.(3 pts) Qual fabricante apresentou uma estimativa do número médio de falhas por componente associada à uma maior precisão? Justifique sua resposta.

(Pode-se apresentar S_A e S_B da calculadora)

$S(\bar{X}) = \frac{S_X}{\sqrt{n}}$ erro padrão da média, $n_A = n_B = 50$.

$$S_X = \sqrt{S_X^2} = \text{desvio-padrão} = \sqrt{\frac{\sum X^2 - \frac{(\sum X)^2}{n}}{n-1}}$$

$$S_A = \sqrt{\frac{150 - \frac{(52)^2}{50}}{49}} \cong 1,399 \Rightarrow S(\bar{X}_A) \cong 0,198$$

$$S_B = \sqrt{\frac{162 - \frac{(64)^2}{50}}{49}} \cong 1,278 \Rightarrow S(\bar{X}_B) \cong 0,181 \Rightarrow \text{maior precisão por apresentar menor } S(\bar{X})$$

b.(4 pts) Qual fabricante apresentou uma amostra mais homogênea? Justifique sua resposta.

$$CV_X = \frac{S_X}{\bar{X}} \times 100\%, \qquad \text{coeficiente de variação}$$

$$CV_A = \frac{1,399}{52/50} \times 100\% \cong \frac{1,399}{1,04} \times 100\% \cong 134,52\%$$

$$CV_B = \frac{1,278}{64/50} \times 100\% \cong \frac{1,278}{1,28} \times 100\% \cong 99,84\% \text{ mais homogênea, pois menor CV\%}$$