CONCURSO PÚBLICO
MINISTÉRIO DA EDUCAÇÃO
SECRETARIA DE EDUCAÇÃO PROFISSIONAL E TECNOLÓGICA
INSTITUTO FEDERAL DE EDUCAÇÃO, CIÊNCIA E TECNOLOGIA DO RIO DE JANEIRO

EDITAL Nº 006/2022

PADRÃO DE RESPOSTAS DA PROVA DISCURSIVA REALIZADA DOMINGO, 15 DE MAIO DE 2022.
PRAZO PARA RECURSO CONTRA O PADRÃO DE RESPOSTAS: 16 E 17 DE MAIO DE 2022, NO ENDEREÇO ELETRÔNICO:
http://www.selecon.org.br

PADRÃO DE RESPOSTAS PRELIMINAR

RIO – 03

ESTATÍSTICA
Estatística; Probabilidade e Estatística

Nº DA QUESTÃO	Espera-se que o candidato(a) desenvolva os aspectos/conteúdos propostos a seguir.
1	O candidato deverá desenvolver o(s) conteúdo(s) com base nos seguintes aspectos: Apresentar probabilisticamente o desenvolvimento da questão. Mostrar conhecimento de probabilidade condicional. Mostrar conhecimento de Teorema de Bayes. (Cada item da questão 1 **vale 2,5 pontos)**

Solução da Questão 1:

Primeiramente vamos definir os eventos, portanto, temos:

Cumprir as metas – CM
Não cumprir as metas – NCM
Participar do curso de treinamento – CT
Não participar do curso de treinamento – NCT

a) P(NCM) = P(NCM|CT)P(CT) + P(NCM|NCT)P(NCT)
= 0,18*0,70 + 0,45*0,30
= 0,126 +0,135
= 0,261 ou 26,1%

b) P(CM) = P(CM|CT)P(CT) + P(CM|NCT)P(NCT)
= 0,82*0,70 + 0,55*0,30
= 0,574 +0,165
= 0,739 ou 73,9%

Ou

P(CM)= 1- P(CM')
= 1- P(NCM)
= 1- 0,261 = 0,739 ou 73,9%

c) P(CT| CM) = $\frac{P(CM|CT)P(CT)}{\mathrm{P(CM)}} = \frac{0,82*0,70}{0,739} = \frac{0,574}{0,739}$ =0,7767 OU 77,67%

P(CM) = 1 - P(NCM) = 1- 0,261 =0,739

d) P(NCT| NCM) = $\frac{P(NCM|NCT)P(NCT)}{\mathrm{P(NCM)}} = \frac{0,135}{0,261} = \frac{0,574}{0,739}$ =0,5172 OU 51,72%

Total previsto de linhas para a resposta final do(a) candidato(a): **1 lauda**

2

O candidato deverá desenvolver o(s) conteúdo(s) com base nos seguintes aspectos:

a) Demonstrar que há correlação entre as duas variáveis através do cálculo do coeficiente de correlação **(5 pontos)**

b) Desenvolver a equação da regressão linear **(3 pontos)**

c) Calcular a proporção da variação total da relação entre as variáveis **(2 pontos)**

Solução da Questão 2:

a) Usando a Fórmula do Coeficiente de Correlação Linear

$$r = \frac{n\sum xy - (\sum x)(\sum y)}{\sqrt{n(\sum x^2) - (\sum x)^2}\sqrt{n(\sum y^2) - (\sum y)^2}}$$

Note que a variável *x* é usada para o INPC e a variável *y* é usada para a tarifa do metrô.

x (INPC)	*y* (Metrô)	x^2	y^2	*xy*
1,9	3,70	3,61	13,69	7,03
2,6	4,10	6,76	16,81	10,66
2,1	4,30	4,41	18,49	9,03
2,5	4,30	6,25	18,49	10,75
2,8	4,60	7,84	21,16	12,88
2,6	5,00	6,76	25,00	13,00
3,1	5,80	9,61	33,64	17,98
3,5	6,50	12,25	42,25	22,75
$\sum x = 21{,}1$	$\sum y = 38{,}3$	$\sum x^2 = 57{,}49$	$\sum y^2 = 189{,}53$	$\sum xy = 104{,}08$

Usando os valores da fórmula e da Tabela acima, r é calculado como segue:

$$r = \frac{n\sum xy - (\sum x)(\sum y)}{\sqrt{n(\sum x^2) - (\sum x)^2}\sqrt{n(\sum y^2) - (\sum y)^2}}$$

$$= \frac{8(104{,}08) - (21{,}1)(38{,}3)}{\sqrt{8(57{,}49) - (21{,}1)^2}\sqrt{8(189{,}53) - (38{,}3)^2}}$$

$$= \frac{832{,}64 - 808{,}13}{\sqrt{459{,}92 - 445{,}21}\sqrt{1516{,}24 - 1466{,}89}}$$

$$= \frac{24{,}51}{\sqrt{14{,}71}\sqrt{49{,}35}} = 0{,}9097$$

Sim, há uma correlação linear entre as duas variáveis.

b) A equação de regressão: $\hat{y} = b_0 + b_1 x$ é a equação típica de uma reta onde b_0 é o intercepto y e b_1 é a inclinação.

$$b_1 = \frac{n\sum xy - (\sum x)(\sum y)}{n(\sum x^2) - (\sum x)^2} \qquad b_0 = \frac{(\sum y)(\sum x^2) - (\sum x)(\sum xy)}{n(\sum x^2) - (\sum x)^2}$$

$$b_1 = \frac{n\sum xy - (\sum x)(\sum y)}{n(\sum x^2) - (\sum x)^2} = \frac{8(104{,}08) - (21{,}1)(38{,}3)}{8(57{,}49) - (21{,}1)^2} = \frac{24{,}51}{14{,}71} = 1{,}67$$

$$b_0 = \frac{(\sum y)(\sum x^2) - (\sum x)(\sum xy)}{n(\sum x^2) - (\sum x)^2} = \frac{(38{,}3)(57{,}49) - (21{,}1)(104{,}08)}{8(57{,}49) - (21{,}1)^2} = \frac{5{,}779}{14{,}71} = 0{,}39$$

$$\hat{y} = 0{,}39 + 1{,}67x$$

c) Calcular o coeficiente de determinação:

O coeficiente de determinação é $r^2 = (0{,}9097)^2 = 0{,}828$. Como r^2 é a proporção da variação total da relação, conclui-se que cerca de 82,8% da variação total em y (tarifa do metrô) pode ser explicada pela relação linear entre as duas variáveis.

Total previsto de linhas para a resposta final do(a) candidato(a): **2 laudas**

3	O candidato deverá desenvolver o(s) conteúdo(s) com base nos seguintes aspectos: Fundamentos (hipóteses) **(1,5 pontos)** Quadro da Anova **(3,5 pontos)** Conclusão sobre a ANOVA **(1,5 pontos)** Teste de Tukey **(3,5 pontos)** Nessa questão, queremos determinar se existem diferenças entre as médias dos números de caixas empacotadas por dia pelas quatro bancadas na semana analisada. Vamos ao teste da ANOVA: * As hipóteses são: $\mathrm{H_0}: \mu_{B1} = \mu_{B2} = \mu_{B3} = \mu_{B4}$ $\mathrm{H_1}$: nem todas as médias são iguais(pelo menos uma das médias difere das outras) * Dados: $n = 20\ observações\ \ e\ \ k = 4\ bancadas.$ * Procedimento da ANOVA: $C = \frac{\left(\sum\sum X_{ij}\right)^2}{n} = \frac{(4.716{,}00)^2}{20} = 1.112.032{,}80$ Soma dos Quadrados Total: $Q_t = \left(\sum\sum X_{ij}^2\right) - C = 1.115.426{,}00 - 1.112.032{,}80 = 3.393{,}2$

Soma dos Quadrados entre amostras:

$$Q_e = \sum \left[\frac{(\sum X_{ij})^2}{n_j}\right] - C = \left[\frac{5.567.560,00}{5}\right] - 1.112.032,80 = 1.479,2$$

Soma dos Quadrados do resíduo:

$$Q_r = Q_t - Q_e = 3.393,2 - 1.479,2 = 1.914,0$$

QUADRO DA ANOVA:

Fonte de Variação	Soma dos Quadrados	Graus de liberdade	Quadrados médios	Fcalc
Entre tratamentos	1.479,2	3	493,0667	4,121769
Dentro das amostras (resíduo)	1.914,0	16	119,625	
Total	3.393,2	19		

* Então para 5% de significância, $k - 1 = 3 \ e \ n - k = 16$, temos que:

$$F_{tab} = F_{(3,16)} = 3.2389 \cong 3,24$$

Como $F_{calc} > F_{tab}$, então, para o nível de significância de 5%, rejeitamos a hipótese nula que diz que as médias das quantidades de caixas empacotadas por cada bancada naquela semana não diferem entre si, logo, existe pelo menos uma média que difere das demais.

Teste de Tukey.

* Calculemos a Diferença Mínima Significativa:

$$\Delta = q_{(k,n-k)}\sqrt{\frac{S_r^2}{r}} = 4{,}05\sqrt{\frac{119{,}625}{5}} = 19{,}8098$$

Marcamos todos os valores das diferenças que são maiores que 19,8098.

	$\bar{x}_{B1} = 242$	$\bar{x}_{B2} = 246{,}4$	$\bar{x}_{B3} = 228{,}8$	$\bar{x}_{B4} = 226$
$\bar{x}_{B1} = 242$	X	4,4	13,2	16
$\bar{x}_{B2} = 246{,}4$	X	x	17,6	20,4
$\bar{x}_{B3} = 228{,}8$	X	x	X	2,8
$\bar{x}_{B4} = 226$	X	x	X	x

Conclusão: A diferença na média do número de caixas empacotadas naquela semana foi entre as bancadas 2 e 4.

$$\mu_{B2} \neq \mu_{B4}$$

Total previsto de linhas para a resposta final do(a) candidato(a): **2 laudas**