Última Lista de Exercícios

1. Para comparar o peso de um componente fornecido por 4 diferentes fornecedores, uma amostra de cada fornecedor foi colhida e analisada. Os dados colhidos estão ao lado. Ao nível de 5% de significância, existe evidência estatística de que o peso médio varia entre os fornecedores?

Amostras			
A	B	C	D
2,4	3,1	3,0	1,9
1,0	3,3	2,9	2,7
2,4	3,5	3,7	1,7
2,5	3,7	2,6	1,8
3,0	2,2	3,4	3,6
3,1	4,9	4,6	3,0

2. Deseja-se comparar a produtividade de 4 equipamentos distintos. A produção de cada equipamento nos últimos dias (turnos de trabalho) foi anotada e os dados colhidos estão ao lado. Nem todos os equipamentos operaram nos últimos turnos, por isto as amostras têm tamanho diferentes. A produtividade dos equipamentos é a mesma ou ela varia?

Amostras		
A	B	C
3,3	2,6	3,1
3,3	3,3	2,6
3,3	4	3,2
1,9	4,7	
2,9	4	
	3,4	
	2,9	

3. Foi feita uma análise de variância para comparar as médias de 5 grupos. Para cada um dos grupos foi colhida uma amostra de 20 elementos e os dados foram analisados através de análise variância. O resultado da ANOVA está na tabela abaixo. Contudo, maior parte da informação, que foi transcrita manualmente, estava ilegível. Com os dados legíveis, é possível rejeitar a hipótese nula de igualdade das médias?

<i>Fonte da variação</i>	<i>SQ</i>	<i>gl</i>	<i>MQ</i>	<i>F</i>	<i>F crítico</i>
Entre grupos	2,68				
Dentro dos grupos					
Total	63,27				

4. Um recrutador quer identificar o melhor aluno para um estágio e decidiu que deveria fazê-lo através das notas obtidas nas disciplinas de graduação do ano anterior. Contudo, como são alunos do terceiro ano (e que não deveriam estar fazendo estágio) e de modalidades distintas de engenharia, eles têm apenas duas disciplinas comuns cursadas no segundo ano – Cálculo e Estatística. As

notas de cada um dos alunos estão abaixo. Existe diferença de nota entre alunos ou disciplinas?

Aluno	Disciplinas	
	Cálculo III	Estatística
Nanico	6,6	6,1
Pescoçudo	7,8	5
Pindoca	8,3	7,1

5. Quatro funcionários foram escolhidos para avaliar quatro métodos de trabalho. Cada funcionário executou uma determinada tarefa uma única vez para cada método (executou a mesma tarefa quatro vezes, portanto). O tempo de execução da tarefa para cada funcionário utilizando cada um dos quatro métodos foi anotado e está na tabela abaixo (16 observações = 4×4). Ao nível de 5% de significância, é possível afirmar que:
- (a) Os métodos não produzem o mesmo resultado?
 (b) Os funcionários não possuem o mesmo desempenho?

Funcionário	Método de Trabalho			
	A	B	C	D
1	23,3	27	30	33
2	28,9	16,4	22,2	33
3	17	29,3	20,7	24,8
4	24,5	27,5	20	39,2

6. O estudo anterior foi repetido em outra unidade da empresa. Mas desta vez comparou-se 5 métodos de trabalho e utilizou-se 10 funcionários. Os dados colhidos foram anotados em uma tabela onde as linhas continham os valores de cada funcionário em cada um dos 5 métodos, e as colunas continha os valores de um método para cada um dos 10 funcionários. O resultado da análise de variância foi parcialmente perdido. A tabela abaixo mostra os dados disponíveis. Ao nível de 5% de significância, é possível afirmar que:
- (a) Os métodos não produzem o mesmo resultado?
 (b) Os funcionários não possuem o mesmo desempenho?

<i>Fonte da variação</i>	<i>SQ</i>	<i>gl</i>	<i>MQ</i>	<i>F</i>	<i>F crítico</i>
Linhas	94,24				
Colunas	729,63				
Erro					
Total	1794,62				

7. Para avaliar o desempenho de diferentes materiais na produção de diferentes peças plásticas, um engenheiro produziu em laboratório quatro peças de cada combinação peça-material, e depois de uma avaliação com seus colegas de equipe, deu uma nota para qualidade geral da peça. Estas notas estão na tabela abaixo. Ao nível de 5% de significância, é possível afirmar que:
- (a) As peças apresentam diferentes graus de dificuldade de produção?
 (b) Os materiais têm diferentes impactos na qualidade de peça produzida?
 (c) Existe interação entre material e tipo de peça produzida?

Peça	Material		
	A	B	C
1	5,3	5,7	6,4
	6,3	5,4	7,9
	6,4	6	7,1
	8,3	5,9	8,5
2	5,6	6,2	8,6
	7,8	7,2	9,5
	9,5	6,1	4,8
	6,2	6,2	6
3	6,5	5,7	5,2
	7,3	7,3	8
	7,8	5,5	7,3
	7,1	4,8	7,4
4	7,5	8,1	6,5
	7,8	6,6	6,3
	7,8	5,8	7,6
	8,1	6,3	6,2

8. Para comparar o efeito da combinação adubo e tipo de semente na produção de certo tipo de planta, um agrônomo criou vários canteiros onde variava a esta combinação. Foram testados 3 adubos e 4 tipos de sementes. Para cada combinação, foram feitas 5 repetições. Com os dados da amostra é abaixo, é possível afirmar ao nível de 5% de significância que:
- Os adubos não geram resultados iguais?
  - Os diferentes tipos de semente não geram resultados iguais?
  - Há interação entre adido e tipo de semente?

Fonte da variação	SQ	gl	MQ	F	F crítico
Amostra					
Colunas	3036,593				
Interações	6078,262				
Dentro	2457,124				
Total	15210,23				

9. Péter trabalha com *advertisement*. Na última semana, ele veiculou o anúncio de 4 clientes em 5 canais da internet, e mediu o número de clicks (e o custo por click) obtido em cada um dos anúncios. Os dados estão abaixo. Ao nível de 5% de significância, é possível afirmar que:

- (a) O desempenho dos clientes não foi o mesmo na última semana?
- (b) O desempenho dos canais não foi o mesmo na última semana?
- (c) Existe iteração entre cliente e canal?

Canal	Cliente			
	A	B	C	D
Mail MKT	0,79	0,87	0,70	0,90
	0,65	0,76	0,80	0,64
	0,68	0,77	0,47	0,71
	0,86	0,98	0,84	0,60
	0,63	0,81	0,65	0,65
	0,66	0,98	0,83	0,55
	0,69	0,92	0,91	0,85
Portal Rugby	0,60	0,75	0,92	0,52
	0,69	0,77	0,79	0,63
	0,25	1,02	0,59	0,66
	0,53	0,82	0,87	0,44
	0,58	0,71	0,72	0,64
	0,52	0,99	0,58	0,48
	0,51	0,72	0,48	0,71
Portal de Economia	0,41	0,77	0,92	0,62
	0,77	0,71	0,79	0,75
	0,54	0,81	0,81	0,50
	0,61	0,71	0,70	0,84
	0,75	0,61	0,77	0,64
	0,48	0,96	0,69	0,88
	0,29	0,71	0,71	0,58
Portal de Esportes	0,89	0,89	0,88	0,85
	1,02	0,81	0,94	1,03
	0,75	1,09	0,82	0,97
	0,90	0,94	0,87	0,92
	0,68	0,84	0,63	0,76
	0,54	0,99	0,87	0,95
	0,88	1,03	1,02	0,78
Site de Busca	0,67	1,08	0,77	0,62
	0,90	0,85	0,98	1,07
	0,75	1,19	0,92	0,54
	0,82	0,80	0,61	0,58
	0,56	1,04	0,81	0,58
	0,62	1,13	0,69	0,77
	0,68	0,86	0,91	0,72

10. A tabela ao lado mostra o resultado de uma pesquisa sobre religião realizada em 3 bairros da cidade. Existe alguma evidência de que a religião predominante muda em cada bairro?

Bairro	Religião		
	Católicos	Evangélicos	Ateus
A	30	12	40
B	12	7	16
C	8	5	12

11. No banco de dados customer_suvery_200 que utilizamos no laboratório, os clientes dos restaurantes eram classificados, entre outras variáveis, por faixa etária e renda. A tabela abaixo mostra a classificação cruzada entre estas duas variáveis. Elas são independentes? Ao afirmar que elas não independentes, qual a probabilidade de estar cometendo um erro?

		X24 -- Income					Total
		\$20 - \$35,000 annually	\$35 - \$50,000 annually	\$50 - \$75,000 annually	\$75 - \$100,000 annually	> \$100,000 annually	
X23 -- Age	18 - 25	9	7	5	3	5	29
	26 - 34	2	12	20	21	3	58
	35 - 49	2	8	18	13	1	42
	50 - 59	1	6	4	15	34	60
	60 and Over	0	0	0	3	8	11
Total		14	33	47	55	51	200

12. Perguntou-se para 100 pessoas qual era seu time preferido e sua cor preferida. Estas duas variáveis são independentes? Use um nível de significância de 5%.

Obs	Time	Cor
1	Palmeiras	Azul
2	São Paulo	Cinza
3	Palmeiras	Verde
4	Corinthians	Azul
5	Santos	Azul
6	São Paulo	Verde
7	Santos	Azul
8	São Paulo	Azul
9	Corinthians	Verde
10	Corinthians	Verde
11	Santos	Azul
12	Palmeiras	Verde
13	Corinthians	Verde
14	Palmeiras	Verde
15	Palmeiras	Azul
16	Corinthians	Verde
17	Santos	Cinza
18	São Paulo	Cinza
19	São Paulo	Cinza
20	Corinthians	Verde
21	Corinthians	Cinza
22	Palmeiras	Verde
23	São Paulo	Verde
24	Palmeiras	Azul
25	Corinthians	Verde
26	Santos	Verde
27	Santos	Azul
28	Palmeiras	Verde
29	Palmeiras	Azul
30	Corinthians	Cinza
31	Corinthians	Verde
32	Santos	Azul
33	Palmeiras	Verde

34	Palmeiras	Azul
35	Santos	Verde
36	Corinthians	Cinza
37	Corinthians	Verde
38	Corinthians	Cinza
39	Palmeiras	Verde
40	Palmeiras	Verde
41	Corinthians	Verde
42	Corinthians	Azul
43	São Paulo	Azul
44	São Paulo	Cinza
45	Santos	Azul
46	Corinthians	Verde
47	Santos	Azul
48	Santos	Verde
49	Corinthians	Cinza
50	Palmeiras	Azul
51	Palmeiras	Cinza
52	Corinthians	Cinza
53	São Paulo	Cinza
54	Santos	Verde
55	Palmeiras	Verde
56	São Paulo	Verde
57	Palmeiras	Verde
58	Corinthians	Verde
59	São Paulo	Azul
60	Palmeiras	Cinza
61	São Paulo	Cinza
62	Palmeiras	Azul
63	São Paulo	Cinza
64	São Paulo	Cinza
65	Palmeiras	Verde
66	Corinthians	Cinza
67	São Paulo	Cinza

68	São Paulo	Cinza
69	Palmeiras	Verde
70	Palmeiras	Azul
71	Palmeiras	Verde
72	Palmeiras	Verde
73	Palmeiras	Azul
74	Santos	Cinza
75	São Paulo	Cinza
76	São Paulo	Cinza
77	Corinthians	Azul
78	Palmeiras	Verde
79	Santos	Verde
80	São Paulo	Cinza
81	Palmeiras	Cinza
82	Palmeiras	Verde
83	Palmeiras	Azul
84	Santos	Azul
85	São Paulo	Cinza
86	São Paulo	Verde
87	Santos	Azul
88	Corinthians	Azul
89	Santos	Azul
90	Palmeiras	Azul
91	Palmeiras	Cinza
92	Santos	Azul
93	São Paulo	Verde
94	Palmeiras	Azul
95	Corinthians	Verde
96	Corinthians	Cinza
97	São Paulo	Verde
98	São Paulo	Verde
99	Corinthians	Verde
100	Palmeiras	Azul

13. Considere a amostra abaixo. É possível afirmar que os dados vêm de uma população com distribuição Normal? Utilize nível de significância de 1%.

25,7	26,3	25,5	23,9	20,7
25,3	24,7	20,1	22,7	28,7
18,6	24,9	28,4	28,8	22,4
26,1	16,6	20,2	22	21,5
25,1	21,9	23,4	23,1	22,6

14. Considere a amostra abaixo. É possível afirmar que os dados vêm de uma população com distribuição Exponencial? Utilize nível de significância de 5%.

38,7	17,9	9,7	4,8	11,7	22,6
86,4	12,8	37,8	22,3	100,5	27,6
18,8	16,2	53,3	116,7	12,1	15,6
7,1	5,9	24,5	8,1	2,2	6,6
6,1	42,7	15,9	27,7	21,6	24,6

15. Considere a amostra abaixo. É possível afirmar que os dados vêm de uma população com distribuição de Poisson? Utilize nível de significância de 5%.

20	24	17	19
19	21	24	18
24	23	18	20
23	21	23	25
21	23	32	17

16. Considere as amostras abaixo. Para cada uma delas, qual a distribuição de probabilidade que melhor representa o comportamento da variável:

- (a) Normal
- (b) Exponencial
- (c) Uniforme
- (d) Poisson

Amostra 1

34	34	28
32	36	34
25	36	31
28	30	33

Amostra 3

31	33,7	35,7
32,5	33,4	29,9
31,4	32,9	30,7
31,9	36,1	32,3

Amostra 2

32,2	37,8	30,4
32,4	36,3	35,7
33,8	34,6	32,4
33,8	37,6	30,1

Amostra 4

61,1	32,2	29,3
21,8	65,7	89,5
56,8	34,2	121,5
17,8	6,2	77,2

17. Um determinado material é entregue por dois fornecedores diferentes, e a resistência à tração muda de um fornecedor para outro. O material do primeiro fornecedor tem uma resistência média de $8,7 \text{ Kgf/mm}^2$ e desvio padrão de $2,1 \text{ Kgf/mm}^2$, e o segundo tem média de $12,5 \text{ Kgf/mm}^2$ e desvio padrão de $3,2 \text{ Kgf/mm}^2$. Houve uma falha no almoxarifado e a informação de origem de um determinado lote foi perdida. Uma amostra de 10 peças foi analisada e os resultados estão abaixo. Qual a origem mais provável do lote?

$$N = 10$$

$$\text{Média Amostral} = 10,6$$

$$\text{Desvio Padrão Amostral} = 2,2$$

18. Para realizar o teste de hipótese abaixo ao nível de significância de 2,5%, uma amostra de 20 elementos foi colhida. Qual o poder do teste para o valor da média populacional de 45?

$$H_0: \mu \geq 50$$

$$H_1: \mu < 50$$

$$\alpha = 2,5\%$$

19. Para estudar o tempo que um consumidor gasta para escolher entre vários produtos alternativos, uma amostra de 20 elementos foi colhida. Com os dados desta amostra:

Obs	x	Y
1	2	6
2	2	9
3	2	4
4	3	8
5	3	5
6	3	4
7	4	9
8	4	10
9	4	8
10	5	9
11	5	9
12	5	6
13	6	9
14	6	11
15	6	9
16	6	9
17	6	12
18	7	14
19	7	17
20	7	12

- Determine o Coeficiente e Correlação de Pearson. Ele é estatisticamente relevante?
- Construa uma reta de regressão. Esta reta possui relevância estatística?
- Construa intervalo de confiança para coeficiente angular e linear da reta de regressão.
- Quando o número de opções aumenta em 1, o tempo necessário para a escolha aumenta quanto em média?
- Construa um intervalo de confiança para o valor esperado do tempo de escolha quando o número de opções é 4
- Construa um intervalo de confiança para o tempo de escolha quando o número de opções é 4

20. Uma amostra de residências selecionadas aleatoriamente e foi anotada sua idade (X) e quanto ao preço de venda (Y). Resultou:

X	Y
1	10
2	30
3	40
4	50
5	65
6	70

- a) Estime a reta de regressão populacional;
- b) Teste, usando o coeficiente angular, se o preço de venda do imóvel diminui à medida que a idade cresce. Use 5%.

- c) Obtenha e interprete o intervalo de projeção de 95% para o preço de uma casa com 3 anos;
- d) Obtenha e interprete o intervalo de projeção de 95% para o preço médio de uma casa com 3 anos;
- e) Estime os coeficientes de correlação e determinação entre X e Y;
- f) É necessário testar a significância do coeficiente de correlação? Explique.

21. A evolução da superfície agrícola utilizada anualmente na produção de soja no Brasil, no período de 2000 a 2011

- a) Coeficiente e Correlação de Pearson – R
- b) Reta de regressão
- c) Comente a qualidade da reta e seu coeficiente de determinação
- d) Construa intervalo de confiança para coeficiente angular e linear da reta de regressão

Ano	Área Plantada (hectares)
2.000	13.693.677
2.001	13.988.351
2.002	16.376.035
2.003	18.527.544
2.004	21.601.340
2.005	23.426.756
2.006	22.082.666
2.007	20.571.393
2.008	21.252.762
2.009	21.761.782
2.010	23.339.094
2.011	24.032.410