

Introdução à Probabilidade e Estatística

2ª Frequência/Exame de época normal — 15 de junho de 2015

2º Semestre 2014/15 – Eng. Civil, Eng. das Energias Renováveis, Eng. Geológica, Eng. Informática e Eng. Mecatrónica

Nome: _____ Número: _____

Todas as folhas de teste entregues devem estar devidamente identificadas e numeradas (se entregar 3 folhas, devem estar numeradas 1/3, 2/3 e 3/3).

A resolução do teste deve ser entregue com Ex. 1 a 3 separados de Ex. 4 a 7.

Os alunos que realizam a 2ª FREQUÊNCIA devem resolver os exercícios 4 a 7. Os alunos que realizam o EXAME devem resolver os exercícios 1 a 7.

1. No site da Ciência Viva encontramos informação relativa ao programa Ciência Viva no Laboratório - Ocupação Científica de Jovens nas Férias (OCJF). Trata-se de um programa destinado a alunos do ensino secundário, que proporciona aos participantes a realização estágios em laboratórios e centros de investigação de todo o país. Os dados relativos ao número de instituições científicas (X) e ao número de alunos (Y) que têm participado anualmente, desde 1997, foram analisados com o auxílio do *software* SPSS, tendo-se obtido os seguintes *outputs*:

| Statistics | | | |
|------------------------|----|------------------------------|------------------------|
| | | Número anual de Instituições | Número anual de alunos |
| N | | 18 | 18 |
| Kurtosis | | -0,454 | -0,846 |
| Std. Error of Kurtosis | | 1,038 | 1,038 |
| Range | | 83 | 1034 |
| Minimum | | 11 | 130 |
| Sum | | 1091 | 13214 |
| Percentiles | 10 | 23,60 | 243,40 |
| | 15 | 37,75 | 318,90 |
| | 25 | 44,00 | 472,50 |
| | 50 | 62,00 | 747,50 |
| | 75 | 83,25 | 1023,75 |
| | 85 | 88,30 | 1068,95 |

Considere ainda que $\sum_{i=1}^{18} x_i^2 = 75409$ e $CV_Y = 42,68\%$. Com base na informação disponibilizada, responda às seguintes questões:

- (a) Determine e interprete:
- A média e a mediana do número anual de alunos que participaram no OCJF.
 - O desvio-padrão e o máximo do número anual de instituições científicas que participaram no OCJF.
- (b) Complete as seguintes afirmações:
- Em 85% dos anos, o número anual de alunos que participaram no OCJF foi superior ou igual a alunos;
 - Em cerca de% dos anos participaram anualmente no OCJF entre 243 e 748 alunos.
- (c) Como classifica a forma distribuição dos dados referentes ao número anual de instituições científicas no OCJF, quanto à assimetria e achatamento? Justifique a sua resposta.
- (d) Diga, justificando, se a seguinte afirmação é verdadeira ou falsa: “O conjunto de dados relativo ao número de instituições que participaram no OCJF é mais homogêneo que o conjunto de dados relativo ao número de alunos que participaram no OCJF.”

2. Seja X uma variável aleatória com a seguinte função de probabilidade,

| | | | | |
|--------|-----|-----|---|----|
| x | 1-k | k-1 | k | 2k |
| $f(x)$ | p | 3p | p | p |

- Sabendo que $E[X] = 8$, determine o valor de p e k .
(Caso não tenha conseguido resolver a alínea a) considere $k = 10$ e $p = 1/6$)
 - Calcule $\text{Var}(X)$.
 - Determine o menor valor x tal que $P[X \leq x] \geq 0.75$.
 - Considere a variável aleatória $Y = X^3$ e construa a função de probabilidade conjunta de (X, Y) .
3. Um fabricante de computadores garante a substituição, por novos, de todos os computadores que avariarem durante o primeiro ano após a data da compra. Admita que o tempo de vida destes computadores é uma variável aleatória que segue uma distribuição Normal com variância 0.7225 anos².
- Admitindo que em 98,3% dos casos o tempo de vida dos computadores excede metade do tempo médio de vida, determine qual o tempo médio de vida destes computadores.
 - Escolhido um computador, ao acaso, qual é a probabilidade de:
 - ter um tempo de vida superior a 1.5 anos e inferior a 4?
 - o fabricante poder ter que o substituir?

(c) Se uma empresa adquirir 25 computadores àquele fabricante, qual a probabilidade de menos de 3 desses computadores serem substituídos?

4. O site EUROSTAT apresenta informação anual acerca da percentagem de indivíduos, com idades compreendidas entre os 16 e os 74 anos, com acesso à internet em casa. Foram considerados 12 anos de observações referentes a Itália e a Portugal e realizada uma análise estatística com recurso ao *software* SPSS. Alguns dos resultados obtidos são apresentados abaixo.

| Descriptives | | | | | |
|---|----------|----------------------------------|-------------|--------|------------|
| | País | Statistic | | | Std. Error |
| % de indivíduos, com idades entre os 16 e os 74 anos, com acesso à internet em casa | ITÁLIA | Mean | | 34,50 | 3,545 |
| | | 95% Confidence Interval for Mean | Lower Bound | 26,70 | |
| | | | Upper Bound | 42,30 | |
| | | Std. Deviation | | 12,281 | |
| | PORTUGAL | Mean | | 32,83 | 4,651 |
| | | 95% Confidence Interval for Mean | Lower Bound | 22,60 | |
| | | | Upper Bound | 43,07 | |
| | | Std. Deviation | | 16,112 | |

| Tests of Normality | | | | | | | |
|---|----------|--------------------|----|------|--------------|----|------|
| | País | Kolmogorov-Smirnov | | | Shapiro-Wilk | | |
| | | Statistic | df | Sig. | Statistic | df | Sig. |
| % de indivíduos, com idades entre os 16 e os 74 anos, com acesso à internet em casa | ITÁLIA | ,146 | 12 | ,200 | ,909 | 12 | ,206 |
| | PORTUGAL | ,146 | 12 | ,200 | ,923 | 12 | ,312 |

| Independent Samples Test | | | | | | |
|---|-----------------------------|---|------|------------------------------|-----------------------|--|
| | | Levene's Test for Equality of Variances | | t-test for Equality of Means | | |
| | | F | Sig. | Sig. (2-tailed) | Std. Error Difference | 95% Confidence Interval of the Difference Upper |
| % de indivíduos, com idades entre os 16 e os 74 anos, com acesso à internet em casa | Equal variances assumed | 1,595 | ,220 | ,778 | 5,848 | 13,795 |
| | Equal variances not assumed | | | ,779 | 5,848 | 13,845 |

- (a) Ao nível de significância de 10%, podemos considerar que, no nosso País, a percentagem média de indivíduos com acesso à internet em casa é de 50%?
- (b) Teste, ao nível de significância de 5%, se existe evidência de que a variância da percentagem de indivíduos de Itália com acesso à internet excede 144.
- (c) Calcule o *p-value* correspondente ao teste da alínea anterior.
- (d) Diga, ao nível de significância de 10%, se existem diferenças significativas entre Itália e Portugal no que refere à percentagem média de indivíduos com acesso à internet.

5. O programa Ciência Viva no Laboratório - Ocupação Científica de Jovens nas Férias proporciona aos participantes a oportunidade de aproximação à realidade da investigação científica e tecnológica em diversas áreas científicas. Alguns dos estágios disponibilizados são nas áreas científicas de Biologia, Informática, Física, Química e Robótica.

De acordo com um estudo realizado em anos anteriores, sabe-se que 36% dos estágios oferecidos eram na área de Biologia, 5% na área de Informática, 21% na área de Física, 22% na área de Química e 16% na área de Robótica. Com base na análise de dados mais recentes, pretende-se averiguar se a distribuição do número de estágios nas referidas áreas se manteve inalterado, para tal realizou-se um teste de hipóteses cujos resultados foram os seguintes: 3

| Estágios por Área Científica | | | | Test Statistics | |
|------------------------------|------------|------------|----------|---|------------------------------|
| | Observed N | Expected N | Residual | | Estágios por Área Científica |
| Biologia | A | 37,4 | -4,4 | Chi-Square | 5,347 ^a |
| Física | 17 | B | -4,8 | df | G |
| Informática | 8 | C | D | Asymp. Sig. | 0,253 |
| Química | E | 22,9 | F | a. 0 cells (0,0%) have expected frequencies less than 5. The minimum expected cell frequency is 5,2. | |
| Robótica | 16 | 16,6 | -0,6 | | |
| Total | 104 | | | | |

- (a) Determine os valores de A a G.
- (b) Indique, justificando convenientemente, se pode concluir que a distribuição do número de estágios nas referidas áreas sofreu alterações significativas.
6. A evolução de 278 pacientes com nomofobia (forte angústia causada pela incapacidade de comunicação por meio de telemóveis ou computadores) foi acompanhada por uma equipa médica. Os pacientes foram agrupados em duas faixas etárias: “até aos 25 anos” e “com mais de 25 anos”. Após 3 anos de um tratamento de recuperação, os pacientes foram classificados como: “Recuperado”, “Não recuperado e ainda em tratamento” ou “Não recuperado e desistiu do tratamento”. Os dados resultantes foram analisados com o auxílio do *software* SPSS, tendo-se obtido os seguintes resultados:

| Faixa etária * Classificação dos pacientes Crosstabulation | | | | | |
|--|-----------------|----------------|-----------------------------|--------------------------------------|---|
| | | | Classificação dos pacientes | | |
| | | | Recuperado | Não recuperado e ainda em tratamento | Não recuperado e desistiu do tratamento |
| Faixa etária | Até aos 25 anos | Count | 56 | 72 | 52 |
| | | Expected Count | 58,3 | ? | ? |
| | Mais de 25 anos | Count | 34 | 21 | 43 |
| | | Expected Count | 31,7 | ? | ? |
| Total | | Count | 90 | 93 | 95 |

- (a) Diga qual o teste de hipóteses que permite concluir se o estado dos pacientes ao fim de 3 anos de tratamento é independente da idade dos mesmos e indique as hipóteses nula e alternativa.
- (b) Efetue o teste de hipóteses que indicou e interprete o resultado.

7. Pretende-se modelar o resultado de um determinado teste em função do valor do Quociente de Inteligência (QI). Foram disponibilizados os valores dos QI e dos respectivos resultados obtidos no teste por 19 indivíduos. Para tal, realizou-se uma breve análise de regressão linear simples com recurso ao *software* SPSS, apresentando-se alguns dos resultados obtidos nas tabelas que se seguem:

Descriptive Statistics

| | Mean | Std. Deviation |
|---------------------------|--------|----------------|
| Resultado do teste | 70,79 | 17,031 |
| Quociente de Inteligência | 108,68 | 15,603 |

Coefficients^a

| Model | | Unstandardized Coefficients | | t | Sig. | 99,0% Confidence Interval for B | |
|-------|---------------------------|-----------------------------|------------|--------|------|---------------------------------|-------|
| | | B | Std. Error | | | | |
| 1 | (Constant) | -43,193 | 8,055 | | ,000 | | |
| | Quociente de Inteligência | | ,073 | 14,288 | | 0,836 | 1,261 |

a. Dependent Variable: Resultado do teste

- Apresente a equação da reta de regressão estimada e interprete os seus coeficientes.
- Determine e interprete os valores dos coeficientes de correlação e de determinação.
- Ao nível de significância de 1% pode concluir que a reta de regressão passa pela origem?
- Para um nível de significância de 5%, verifique se o QI de um indivíduo influencia linearmente de forma significativa o resultado do teste.
- Sabendo que o resultado do teste de um indivíduo com um QI de 115 foi 85, calcule o resíduo de estimação.

Bom Trabalho!