

PARTE A

GRUPO I

Ainda emocionadas com a vinda do Papa a Portugal, as irmãs do convento de Santa Catarina pretendem estimar o número médio de pessoas que chegaram ao Terreiro do Paço, por segundo, no dia da visita. As irmãs tem em sua posse uma amostra aleatória de 20 observações para o número de pessoas que chegaram por segundo (X), verificando-se $\sum x_i = 100$.

- Obtenha o estimador da máxima verosimilhança para o valor médio de X . Obtenha, ainda, a estimativa da máxima verosimilhança para a probabilidade de, em 5 segundos, terem chegado ao Terreiro do Paço exactamente 30 pessoas
- Imagine agora que lhe era proposto o seguinte estimador para o mesmo parâmetro:

$$\tilde{\lambda} = \frac{X_2 + X_3 + \dots + X_{N-2}}{N-3}$$

Compare os estimadores quanto ao seu enviesamento e eficiência relativa. Comente, ainda, as propriedades do estimador da máxima verosimilhança.

- Imagine que as irmãs acabaram de registar a entrada de 1 pessoa no Terreiro do Paço. Se elas pretenderem obter o estimador da máxima verosimilhança para o tempo médio de espera, em segundos, até voltarem a contabilizar a entrada de outra pessoa, que estimador da máxima verosimilhança sugere? Nota: Calcule-o tendo em conta a alínea a). Qual a propriedade que utilizou?

GRUPO II

O padre Bruno sugeriu às irmãs que estendessem a sua análise estatística aos restantes locais que o Papa visitou em Portugal. Considere populações normais.

- As irmãs pretendem estudar a variabilidade do número de pessoas presentes nas várias missas, através de um intervalo de confiança a 90%. Na amostra recolhida constam os seguintes dados:

$$\sum_{i=1}^{25} x_i = 2500 \qquad \sum_{i=1}^{25} x_i^2 = 251200$$

Calcule o intervalo de confiança pedido e interprete-o.

- A irmã Andreia tem em sua posse um intervalo de confiança a 90% para o número médio de pessoas que esteve presente na missa do Porto, cuja margem de erro é de 8%. Sabendo que a variância populacional é igual a 0.0645, qual foi o tamanho da amostra utilizada?
- Se a irmã Andreia desconhecesse a verdadeira variância, que prevê que fosse acontecer à margem de erro do intervalo, para a mesma amostra? Não faça cálculos para responder a esta questão.

PARTE B

GRUPO III

Numa determinada região foi feito um estudo de caracterização da população, tendo sido recolhidos, entre outros, dados relativos à altura e ao peso de 200 homens e 200 mulheres residentes nessa região. Os resultados obtidos foram (a letra A refere-se à altura, a letra P ao Peso, a letra H aos homens e a letra M às mulheres),

$$\begin{aligned}\bar{X}_{A,H} &= 180 \text{ cm}; & S'_{A,H} &= 40 \text{ cm}; & \bar{X}_{P,H} &= 70 \text{ kg}; & S'_{P,H} &= 25 \text{ kg} \\ \bar{X}_{A,M} &= 165 \text{ cm}; & S'_{A,M} &= 38 \text{ cm}; & \bar{X}_{P,M} &= 55 \text{ kg}; & S'_{P,M} &= 22 \text{ kg}\end{aligned}$$

- Será de rejeitar a hipótese nula da altura média das mulheres ser de 160 cm contra a hipótese de ser de 170 cm, a um nível de significância de 10%?
- Em relação ao teste definido na alínea anterior, considere que a regra de decisão passa a ser rejeitar a hipótese nula se a altura média das mulheres na amostra for superior a 166.5. Neste caso, calcule a probabilidade de cometer os dois tipos de erro, assim como a potência do ensaio.
- Supondo que o verdadeiro desvio padrão da variável altura das mulheres é igual a 38 cm, calcule a dimensão da amostra a recolher para que, com a regra de decisão definida em b), a probabilidade de cometer o erro tipo II seja igual a 5%.
- Será de considerar que o peso dos homens nesta região não é significativamente diferente do peso das mulheres, a um nível de significância de 1%?

GRUPO IV

Nos EUA, três programas de televisão piloto para potenciais *sitcoms* foram mostrados a audiências em quatro regiões do país – Leste, Sul, Costa Oeste e *Midwest*. Com base nas reacções das várias audiências foi obtida uma pontuação (numa escala de 0 a 100) para cada um dos programas. Alguns dos resultados encontram-se no quadro em baixo,

FONTE DE VARIAÇÃO	SOMAS DE QUADRADOS
Programa	95.2
Região	69.5
Residual	
Total	244

- Diga quais as hipóteses que se podem testar com esta análise de variância. Indique as variáveis relevantes e os pressupostos sobre os quais assentam os testes a realizar.
- Complete o quadro ANOVA e diga quais as conclusões a que chega.

PARTE C

GRUPO V

Um micro organismo foi sujeito a dois tipos de radiação, sendo o objectivo o de determinar se estas produzem ou não efeitos diferentes sobre a duração de vida deste micro organismo. Os tempos registados são os expostos na tabela seguinte, em segundos. Faça o teste que considerar apropriado para determinar se sim ou não devemos considerar que há efeitos distintos sobre a duração de vida, trabalhando com uma margem de erro de 5% (observação: não tem informação que lhe permita admitir que a população ou populações segue a distribuição normal).

Radiação alfa7		Radiação gama5	
198,1	197,0	198,5	208,5
199,0	196,0	312,5	206,5
200,9	210,8	203,8	210,5
202,0	204,0	199,5	198,2
313,0	203,3	202,6	203,5
187,0	210,0	197,5	311,5
189,0		204,5	207,5

GRUPO VI

- Imagine que está a testar que a média de uma população é igual a um determinado valor, usando a distribuição normal. Responda, justificando, às questões seguintes:
 - Pode cometer neste teste dois tipos de erros. Quais? Em que se distinguem? São independentes um do outro?
 - Porque razão neste teste a região crítica é simétrica?
 - Que é a “potência de teste”? Como se calcula?
- Comente as seguintes afirmações:
 - Os testes não paramétricos como o “sign test” ou o U de Mann-Whitney são testes sobre a igualdade de medianas;
 - Mas, por vezes, podem também ser testes sobre a média, como no caso em que as distribuições são simétricas.