UAb

ESTATÍSTICA APLICADA I | 21041

ATIVIDADE FORMATIVA 1

Período de realização

Decorre de 3 a 9 de Novembro de 2020

Temas:

- Introdução à Inferência Estatística
- Estimação Pontual
- Estimação por Intervalos de Confiança

Objetivos

Os objetivos da atividade formativa insirem sobre os objetivos dos Temas 1, 2 e 3:

- Conceitos das principais distribuições discretas e contínuas;
- Objetivos da Inferência Estatística
- Estimadores pontuais e suas propriedades
- Método da Máxima Verosimilhança
- Principais distribuições amostrais
- Intervalos de confiança para: propoções; diferença de proporções; médias; diferença de médias; variância e desvio padrão; razão de variâncias
- Implementação em **R** dos conceitos anteriores

Descrição da Atividade (Exercícios a Resolver)

1. Considere dois universos normais:

$$X_1 \sim N(\mu_1; \sigma_1)$$

$$X_2 \sim N(\mu_2; \sigma_2)$$

e amostras aleatórias independentes de cada um dos universos (de tamanhos n_1 e n_2 respetivamente).
Deduza a distribuição amostral da estatística $T = \bar{X}_1 + \bar{X}_2$.
Que poderia concluir se fosse desconhecida a distribuição dos universos?

2. Para o parâmetro θ de certa população, foram indicados dois estimadores: $\hat{\theta}_1$ e $\hat{\theta}_2$. Diga qual preferiria, sabendo que:

$$E\left[\hat{\theta}_1\right] = \frac{n+1}{n}\theta \quad VAR\left[\hat{\theta}_1\right] = \frac{k}{n}$$

$$E\left[\hat{\theta}_2\right] = \frac{n+1}{n}\theta \quad VAR\left[\hat{\theta}_2\right] = \frac{k}{n+3}$$

onde k é uma constante e n o tamanho da amostra.

3. Seja a estatística:

$$T = \frac{(n-1)X_1 + X_n}{n}$$

definida com base numa amostra aleatória de dimensão n, recolhida de uma população normal.

3.1 Verifique se T constitui um estimador não enviesado ou centrado para a média da população.

3.2 Será T um estimador consistente para aquele parâmetro da população?

4. Derive o estimador de máxima verosimilhança para o parâmetro p de uma população com distribuição de Bernoulli, calculados a partir de uma amostra de dimensão N.

5. Encontre os estimadores de máxima verosimilhança para os parâmetros μ e σ^2 de uma população normal, calculados a partir de uma amostra de dimensão N.

6. A variável aleatória X segue uma distribuição com função densidade de probabilidade dada por

$$f(x) = \begin{cases} \frac{(x-\alpha)}{2}, & se \;\; \alpha < x < \alpha + 2 \\ 0, & no \;\; caso \;\; contrário \end{cases}$$

onde α é um parâmetro desconhecido.
Recorrendo ao método de máxima verosimilhança, estime o parâmetro α a partir da seguinte amostra aleatória, constituída por 10 observações

3.5 ; 4.3 ; 2.8 ; 4.5 ; 2.9 ; 3.3 ; 3.8 ; 2.9 ; 4.0 ; 3.9

7. O tempo que uma máquina leva a executar certa operação em cada peça produzida é sujeita a variações. Para verificar se as condições de funcionamento estão dentro das normas, registou-se 12 vezes o referido tempo.
Os resultados (em segundos) foram os seguintes:

29	33	36	35
36	40	32	37
31	35	30	36

Construa um intervalo de confiança a 95% para o tempo médio de execução da tarefa pela máquina em análise.
NOTA: suponha que se pode aceitar que o tempo de execução da tarefa segue uma distribuição aproximadamente normal.

8. Um fabricante produz peças de diâmetro especificado em 100 mm. Querendo estimar o verdadeiro diâmetro num grande lote, selecionou 25 peças ao acaso, que depois de medidas forneceram os seguintes valores:

$$\sum x_i = 2530\,mm \qquad \sum (x_i - \bar{x})^2 = 384\,mm^2$$

8.1 Identifique o universo a ser objeto de estudo.

8.2 Apresente um estimativa para o diâmetro médio do lote.

8.3 Faça o mesmo através de um intervalo com um grau de confiança de 99%. Que vantagens resultam da identificação de um intervalo de confiança, em relação à estimativa pontual na alínea 6.2?

8.4 Quantas peças deveriam ser incluídas na amostra, se se pretendesse aumentar a precisão do intervalo - reduzir a sua amplitude para 3 mm?

9. Pretende-se avaliar a proporção p de indivíduos com uma certa doença numa população com um erro que não exceda $\pm 2\%$ para um coeficiente de confiança de 95%.
Qual a dimensão da amostra a recolher, se:

9.1 nada se sabe sobre p;

9.2 se admitir que p é da ordem de grandeza 10%.

Exercícios Computacionais em R

10. Elabore uma rotina e implemente-a no R, para gerar 100 números pseudo-aleatórios com distribuição:

10.1 $X \sim Poisson(\lambda = 2)$

10.2 $X \sim U\,[0, 1]$ (Distribuição Uniforme)

10.3 $X \sim N(5, 5)$ (Distribuição Normal)

11. Considere a seguinte amostra:

109 118 119 121 121 134 111 125 137 121 133 111 118 102 112 121 109 117 114
105 132 122 132 134 109 112 115 121 122 120 128 119 116 114 111 102 120 122

Determine um intervalo de confiança para a média da população de onde esta amostra foi retirada. Considere $\alpha = 1\%$.

12. Uma empresa de transportes consome em média, 1000 litros de biodiesel, por dia. O consumo diário durante a última semana foi de 995, 1015 , 985 , 1004 , 907, 1002, 976 litros.
Determine um I.C. a 99% para variância do consumo diário.

Recursos
Pode utilizar todos os recursos disponibilizados na página da unidade curricular e da bibliografia recomendada.

FIM