



Laboratório
Nacional de
Computação
Científica

Programa de Pós-Graduação em
Modelagem Computacional

Trabalho

GA030 - Estatística

Prof.: Marcio Borges
email: *mrborges@lncc.br*
Sala: 2A40

Petrópolis-RJ
28 de outubro de 2024

-
1. Após abordarmos a *Lei dos Grandes Números* e o *Teorema do Limite Central*, chegamos a um ponto crucial do curso: a estimação de parâmetros (desconhecidos) associados à distribuição de probabilidade de uma variável aleatória.

O presente trabalho tem como objetivo a fixação das ideias introduzidas até aqui. Para isso, utilizaremos dados armazenados em quatro arquivos, que contêm amostras de diferentes variáveis aleatórias, conforme a Tabela 1.

Tabela 1: Descrição das variáveis aleatórias.

| Variável | Arquivo | Distribuição |
|-------------------------------------|-------------------|--------------|
| $Q \sim \mathbb{N}(0, 2)$ | <i>data1q.dat</i> | Normal |
| $X \sim \mathbb{U}[-1, 1]$ | <i>data1x.dat</i> | Uniforme |
| $Y \sim \mathbb{E}(\lambda = 0,05)$ | <i>data1y.dat</i> | Exponencial |
| $T \sim \mathbb{B}(15, 0,40)$ | <i>data1t.dat</i> | Binomial |

- (a) Dado que conhecemos a distribuição de probabilidades de cada variável aleatória e os parâmetros que as caracterizam (Tabela 1), calcule a expectância e a variância (teóricas) de cada uma delas, usando as definições que vimos em aula.
- (b) Utilize o R (ou outro programa) para ler cada arquivo e calcule estimativas para a média e a variância do conjunto de dados (usando todos os dados disponíveis nos arquivos). Em seguida, compare com os resultados obtidos no exercício anterior. Faça comentários.
- (c) Construa os histogramas com as frequências relativas de cada uma das variáveis, verificando se estes são condizentes com os modelos teóricos (Tabela 1).
- (d) Considere cada uma das amostras das variáveis aleatórias, contidas nos arquivos, e suas diferentes distribuições de probabilidades. Tome **amostras aleatórias** de tamanho n ($n = 5, 10$ e 50) de cada uma das variáveis aleatórias e construa as variáveis aleatórias (**estatísticas**):

- *média amostral*: $\bar{W}^{(n)} = \frac{1}{n} \sum_{i=1}^n W_i$
- *variância amostral*: $S_W^{2(n)} = \frac{1}{n-1} \sum_{i=1}^n (W_i - \bar{W}_n)^2$

onde $W = Q, X, Y$ ou T . Use 10000 amostras simples (pontos amostrais) para gerar as variáveis aleatórias **média amostral** e **variância amostral**. Obs.: Lembre-se das características que as **amostras aleatórias** devem ter. Apresente o código.

- (e) Usando o código da questão anterior, construa os histogramas de frequências das variáveis aleatórias **média amostral** e **variância amostral**, para os diferentes valores de n e compare com as distribuições teóricas esperadas para estas variáveis. Faça isso para as variáveis (Q, X, Y e T).
- (f) Compare os histogramas, para os diferentes valores de n , e discuta os resultados.
-