

# Estatística não paramétrica

Profa. Luciane Alcoforado

# Ementa

- <http://est.uff.br/wp-content/uploads/sites/146/2019/01/Form13-GET00155.pdf>

# Presença

- Chegar no horário
- Chamada será realizada até as 19:30, após isso será imputado falta ao aluno
- Excesso de falta reprova, observar o limite permitido.
- Reposição será para repor avaliações presenciais, trabalhos não entregues no prazo não poderão ser repostos

# Avaliação

- Avaliações periódicas (semanal)
- Listas de exercícios, trabalhos individuais e em grupo, avaliação presencial.
- VS ou Reposição ao final com todo o conteúdo
- Os exercícios e trabalhos serão disponibilizados em “tarefa no conexão uff” com prazo de entrega a ser observado pelo aluno.

# Teste de hipóteses

- Revisar conceitos.
- é o processo de inferir a partir de uma amostra se uma determinada afirmação sobre a população parece ser verdade ou não.
- Hipótese nula e alternativa
- Testes unilaterais e bilaterais
- Erros tipo I e II
- Nível de significância
- Regra de decisão

# Teste não paramétrico

- Um teste não-paramétrico é aquele cujo modelo não especifica condições sobre os parâmetros da população da qual a amostra foi obtida. Mesmo quando existem certas pressuposições, estas são mais brandas do que aquelas associadas ao testes paramétricos.

# Definição

Um método estatístico é não paramétrico se satisfizer pelo menos um dos seguintes critérios.

1. O método pode ser usado em dados com uma escala nominal de medição.
2. O método pode ser usado em dados com uma escala ordinal de medição.
3. O método pode ser usado em dados com uma escala de medida de intervalo ou razão, onde a função de distribuição da variável aleatória que produz os dados é não especificado ou especificado exceto por um número infinito de parâmetros desconhecidos.

# Escalas de Mensuração

- **Nominal:** Usa números ou símbolos para identificar elementos ou suas características em diferentes classes ou categorias. Os valores da escala não indicam ordem, hierarquia ou proporção. Desta forma a atribuição de números ou símbolos a essas categorias é arbitrária.
- **Ordinal:** Os valores na escala ordinal são números ou mesmo símbolos, como na escala nominal, mas possuem características de ordenação ou de hierarquia.
- **Intervalar:** Tem todas as características de uma escala ordinal, mas além disso envolve uma unidade de medida que contem o ponto 0 como referência. O grau de mensuração por intervalos é obtido quando podemos definir a distância (amplitude do intervalo) entre duas medidas quaisquer da escala.
- **De razão:** leva em conta a ordem, o intervalo e a razão entre duas medidas. Se uma medida  $m$  é três vezes maior que  $n$ , estas medidas estão numa escala de razão. Nesta escala o 0 (zero) é absoluto, ou seja é zero mesmo, e, por isso todas as operações aritméticas tem sentido.



# Função de distribuição empírica

- Quase sempre desconhecemos  $F(x)$ , a função de distribuição da variável aleatória  $X$

**Definition 2.2.1** Let  $X_1, X_2, \dots, X_n$  be a random sample. The **empirical distribution function**  $S(x)$  is a function of  $x$ , which equals the fraction of  $X_i$ s that are less than or equal to  $x$ ,  $-\infty < x < \infty$ .

$$S(x) = \frac{\text{number of } X_i\text{'s} \leq x}{\text{number of } X_i\text{'s}}$$

**Example 2.2.1** In a physical fitness study five boys were selected at random. They were asked to run a mile, and the time it took each of them to run the mile was recorded. The times were 6.23, 5.58, 7.06, 6.42, 5.20. The empirical distribution function  $S(x)$  is repre-

$$\begin{aligned} P(X = 5.20) &= .2, & P(X = 5.58) &= .2, & P(X = 6.23) &= .2, \\ P(X = 6.42) &= .2, & P(X = 7.06) &= .2 \end{aligned}$$

►  $E[X] = \sum_x x f(x) = 6.098$

►  $\text{Var}(X) = \sum_x (x - E[X])^2 f(x) = .424$

# Exercício

- Os dados abaixo foram gerados de uma distribuição exponencial. Ajuste a distribuição empírica e obtenha  $E(X)$ ,  $VAR(X)$  e  $DP(X)$

0.13 0.19 0.21 0.41 0.47 0.71 0.87 1.36 2.48 2.81

# Quantile amostral

**Definition 2.2.2** Let  $X_1, X_2, \dots, X_n$  be a random sample. The *p*th sample quantile is that number  $Q_p$  satisfies the two conditions:

1. The fraction of the  $X_i$ s that are less than  $Q_p$  is  $\leq p$ .
2. The fraction of the  $X_i$ s that exceed  $Q_p$  is  $\leq 1 - p$ .

$$Q_p = \begin{cases} \frac{X_{(np)} + X_{(np+1)}}{2} & \text{if } np \text{ is an integer,} \\ X_{(\lceil np \rceil)} & \text{otherwise.} \end{cases}$$

**Example 2.2.3** Six married women were selected at random and the number of children belonging to each was recorded.

0, 2, 1, 2, 3, 4

- ▶ Empirical distribution function
- ▶  $Q_{.5} = 2$ ,  $Q_{.25} = 1$ ,  $Q_{.75} = 3$
- ▶  $Q_{1/3} = \frac{x_{(2)} + x_{(3)}}{2} = 1.5$

Fim da aula 1