# Plano Aula 15 e 16

## Inferência Estatística

Essa semana veremos resultados e extensões de Probabilidade que terão aplicação nas próximas semanas.

- Estatística descritica $\times$ inferência estatística;
  - população e amostra: parâmetros ($\mu$, $\sigma^2$, $\pi$, ...) $\times$ estatísticas ($\overline{x}$, $s^2$, $p$, ...).

Definição **Estatística**: é qualquer valor obtido em função da amostra. Exemplo, $\overline{x}$, $s^2$, $p$, ...

### Distribuição amostral (Bussab e Morettin - Seção 10.7)

*"Toda função de variáveis aleatórias (v.a.s) é uma v.a."*

Definição **Distribuição Amostral**: é a distribuição de probabilidade de uma estatística.

- **Exemplo 1**: Seja $X$ a v.a. que denota o número de livros que a população de monitores do curso 'Probabilidade e Estatística' lêem por semestre. Suponha que no último semestre foram lidos 5, 7, 4. Se não soubéssemos essa informação e decidíssemos observar uma amostra de tamanho $n = 2$ para saber a média de livros lidos $\overline{X}$.
  - Quais as possíveis amostras? Cada amostra pode gerar um $\overline{x}$ diferente;
  - Os valores de média calculados com cada amostra formam a distribuição amostral de $\overline{X}$.

**Lembrando: Amostra aleatória simples (a.a.s.) = v.a. idependentes e identicamente distribuídas (i.i.d.)**

Definição **A.A.S**: Seja $X_1, X_2, \ldots, X_n$ uma a.a.s. de tamanho $n$ de $X \sim f(x;\theta)$, então $X_1 \sim f(x;\theta)$, ..., $X_n \sim f(x;\theta)$ e $X_i$ e $X_j$ são independentes para todo $i \neq j$.

- **Exemplo 2**: Seja $X$ a duração de vida de um tipo de lâmpada, tal que $X \sim Normal(\mu, \sigma^2)$. Também assuma que $X_1, X_2, \ldots, X_n$ são uma a.a. de tamanho $n$ de $X$ e $E(X) = \mu$.
  - Média amostral $\overline{X} = \frac{X_1+X_2+\ldots+X_n}{n} = \frac{1}{n}\sum_{i=1}^{n} X_i$ é aleatória.
  - Depois de observada a amostra $(X_1, \ldots, X_n) = (x_1, \ldots, x_n)$ escrevemos a estatística $\overline{x} = \frac{1}{n}\sum_{i=1}^{n} x_i$.

Definição **Erro padrão**: é o desvio padrão de uma estatística. Exemplo, erro padrão da *média amostral* é $\sigma_{\overline{X}} = \sqrt{Var(\overline{X})}$.

- **Exemplo 3**: ... Proporção amostral $\hat{p} = \frac{1}{n}\sum_{i=1}^{n} X_i$, para $X_i \in \{0, 1\}$. ...

**Teorema central do limite (Bussab e Morettin - Seção 10.8)**

*"Garante que uma média amostral se aproxima do seu valor esperado à medida que o tamanho da amostra aumenta (dadas algumas condições...)"*

- **Teorema 10.2** e **Corolário 10.1**
- Aplicativo que ilustra o TCL - https://brunamdalmoro.shinyapps.io/TCL_medias/

---

## Ler slides e ver vídeos da semana 8.

## Fazer lista de exercícios 2-3.

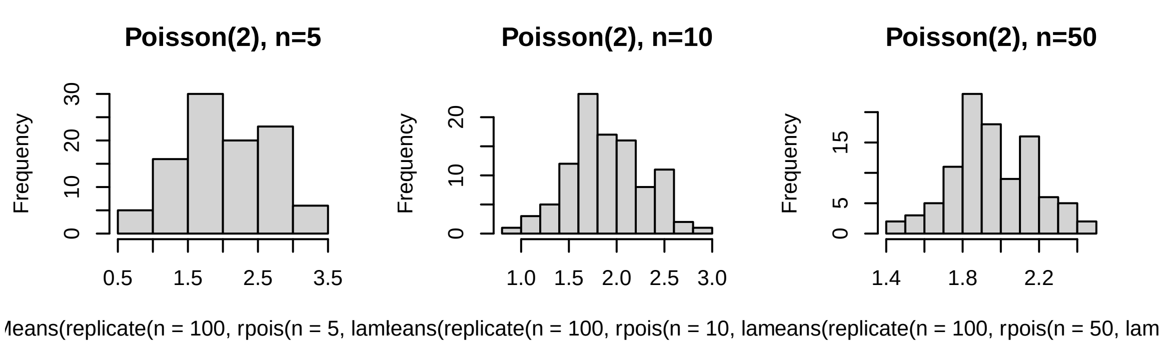
---

### Ilustração do TCL no `R`

No `R` é possível gerar amostras, calcular a mádia de cada a mostra e plotar o histograma:
(*usamos* `replicate` *para gerar 100 amostras de tamanho n= 25, 50 e 100*)

- a.a. de $X \sim Poisson(2)$

```
par(mfrow=c(1,3))
hist( colMeans( replicate( n = 100, rpois( n = 5, lambda = 2))), main="Poisson(2), n=5")
hist( colMeans( replicate( n = 100, rpois( n = 10, lambda = 2))), main="Poisson(2), n=10")
hist( colMeans( replicate( n = 100, rpois( n = 50, lambda = 2))), main="Poisson(2), n=50")
```



- $X \sim Uniforme(0, 1)$

```
par(mfrow=c(1,3))
hist( colMeans( replicate( n = 100, runif( n = 5, min = 0, max = 1))), main="Uniforme(0,1), n=5")
hist( colMeans( replicate( n = 100, runif( n = 10, min = 0, max = 1))), main="Uniforme(0,1), n=10")
hist( colMeans( replicate( n = 100, runif( n = 50, min = 0, max = 1))), main="Uniforme(0,1), n=50")
```