ESTATÍSTICA

Michelle Hanne Soares de Andrade
michellehanne.andrade@gmail.com

Teorema do Limite Central

Distribuição Amostral

Duas amostragens oriundas da mesma população quase sempre terão estatísticas diferentes.

Diferentes amostragens produzirão amostras com estatísticas distintas.

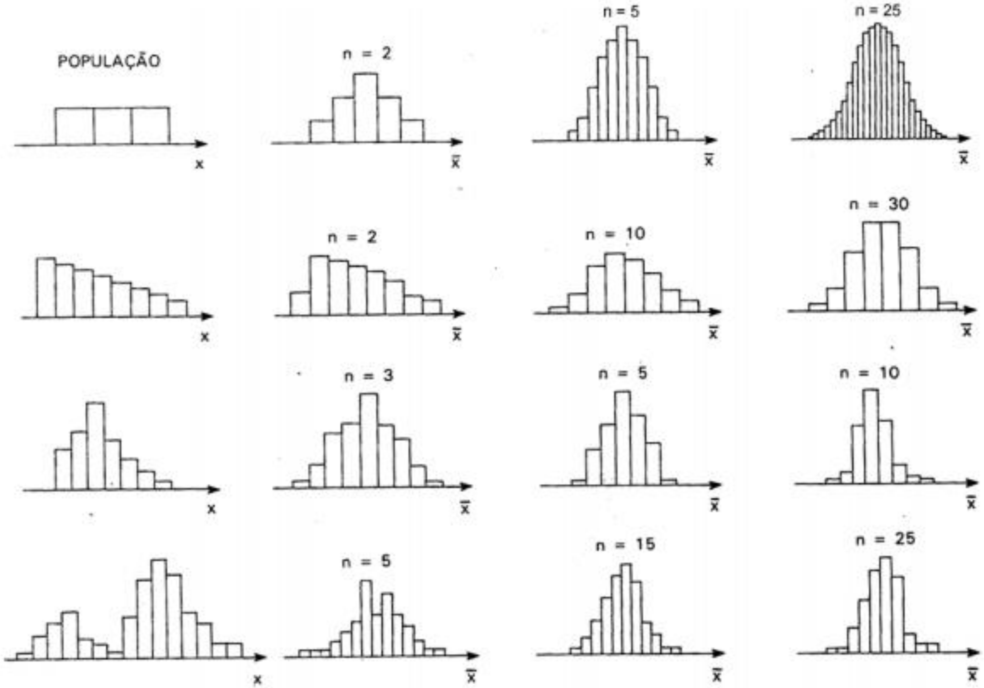
Amostragens são probabilísticas, portanto, estatísticas baseadas nas amostragens também o são.

Se as características da amostragem e a composição da população são conhecidas, a probabilidade de cada resultado pode ser determinada.

Teorema do Limite Central

- Quando o tamanho da amostra (***n***) aumenta, independente da forma de distribuição da população, a distribuição amostral da média da amostra ($\overline{x}$) converge para uma distribuição normal.

Histogramas de distribuição da média para amostras de algumas populações



Fonte: BUSSAB & MORETTIN. ***Estatística Básica***. São Paulo, Atual, 3ª edição, 1986, pp. 197.

Teorema do Limite Central

- **Se a média de uma amostra for um estimador razoável não será necessário conhecer a f.d.p. da população, pois a distribuição de probabilidades da média das amostras será aproximadamente uma normal.**
- A média das distribuições amostrais será igual à da população (μ) e a sua **variância será dada por $\boldsymbol{\sigma^2 / n}$.** Matematicamente:

$$E(\bar{x}) = \mu \qquad Var(\bar{x}) = \frac{\sigma^2}{n} \qquad \sigma(\bar{x}) = \frac{\sigma}{\sqrt{n}}$$

Distribuição Amostral da Média

- Seja X uma variável aleatória (v.a.) com média μ e variância σ^2 , e seja (X1 , X2 ,..., Xn) uma Amostra Aleatória Simples (AAS) de X, de tamanho n, então:

$$\mathrm{E}(\overline{x}) = \mu$$

$$\sigma_{\overline{X}}^2 = \frac{\sigma^2}{n} \Rightarrow \sigma_{\overline{X}} = \frac{\sigma}{\sqrt{n}}$$

- Ou seja, quanto maior a amostra, menor o desvio padrão da distribuição amostral.

Distribuição Amostral da Média -Exemplo

n	Amplitude das amostras				Das médias	
	médias		desvios-padrão			
	mín	máx	mín	máx	média	dpad
15	1,550	1,777	0,085	0,239	1,672	0,039
30	1,610	1,738	0,101	0,213	1,672	0,027
60	1,624	1,720	0,114	0,182	1,671	0,019
150	1,643	1,703	0,132	0,168	1,669	0,012

População = $X \sim N(1{,}67;\ 0{,}15^2)$

Parece haver alguma relação entre o desvio-padrão das médias e o tamanho da amostra (n)?

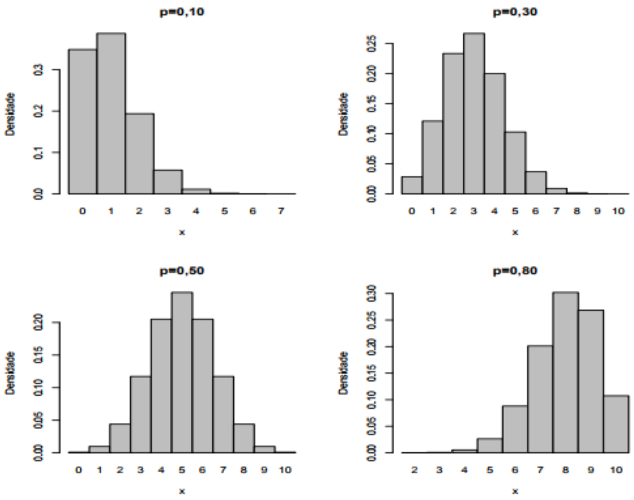
Distribuição Binomial

- **Serão exibidos construídos histogramas para a distribuição de X ∼ B(n, p) variando-se o número de ensaios n e também a probabilidade de sucesso p.**

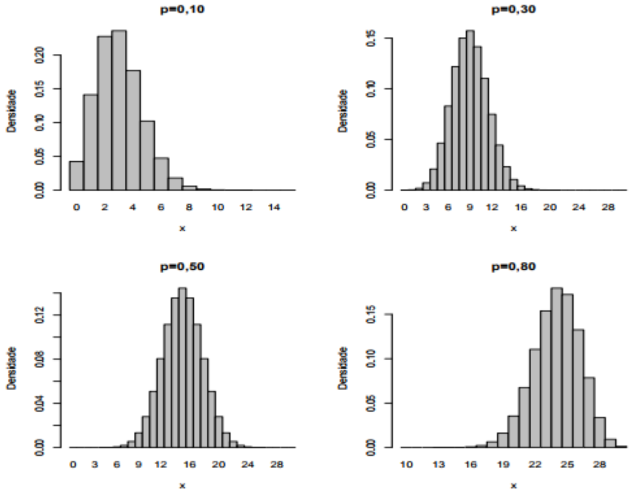
- Nota-se pelos gráficos que à medida que n cresce a distribuição de X ∼ B(n, p) se aproxima da distribuição de Y $\sim N(\mu_X, {\sigma^2}_x)$ em que $\mu x = np$ e ${\sigma^2}_x = np(1 - p)$.

Distribuição Binomial

Histogramas $B(n, p)$ para $n = 10$

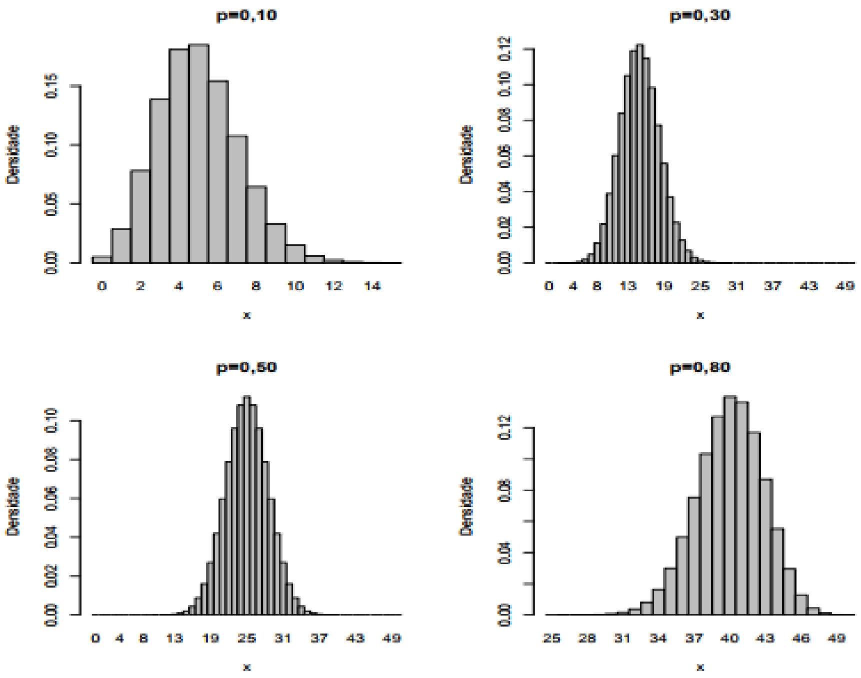


Histogramas $B(n, p)$ para $n = 30$



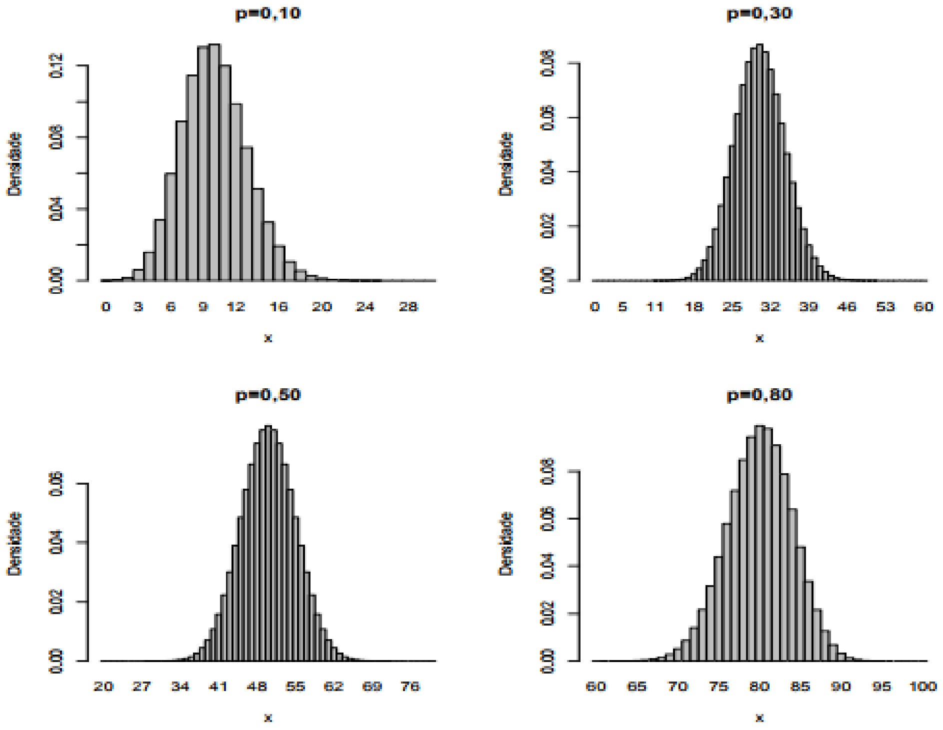
Distribuição Binomial

Histogramas B(n, p) para $n = 50$



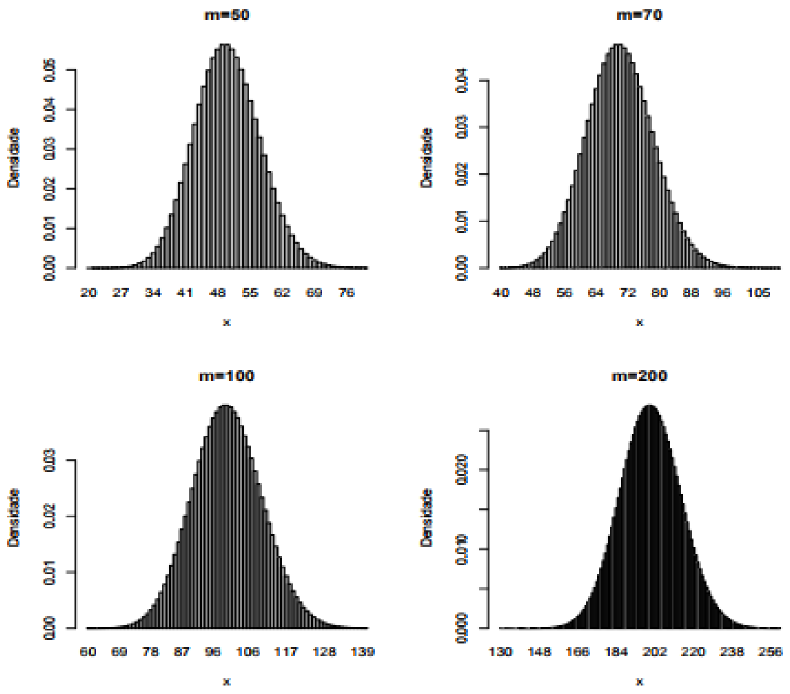
Distribuição Binomial

Histogramas B(n, p) para $n = 100$



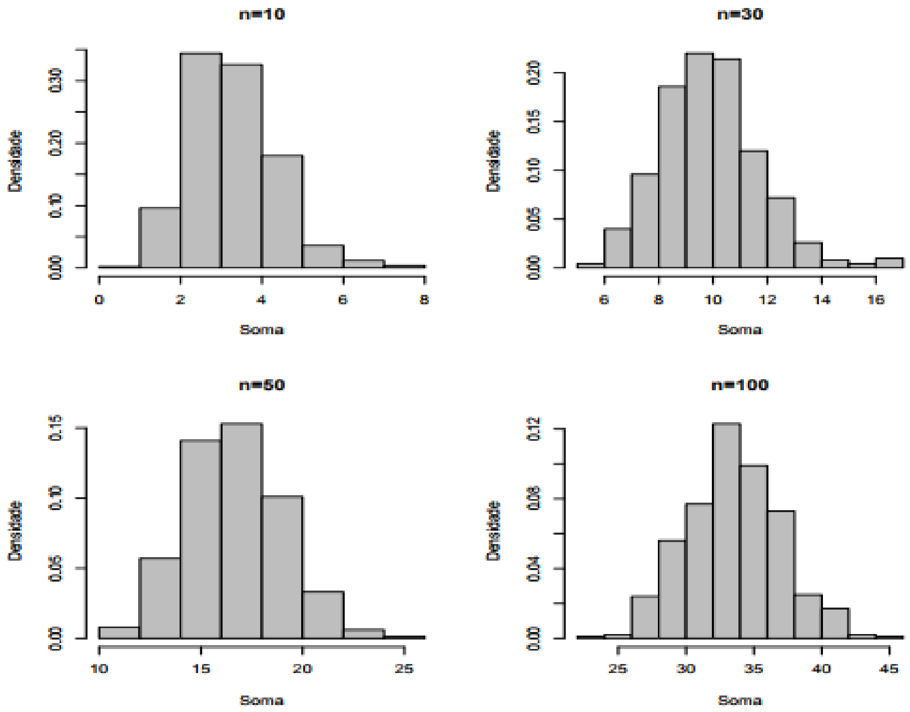
Distribuição Poisson

Histogramas P(m)



Distribuição Exponencial

Histogramas Soma de Exponenciais com $\lambda = 3$



Teorema do Limite Central

Enunciado para a Soma Amostral

Para variáveis aleatórias $X_1, \dots, X_n$ independentes e com mesma distribuição de média μ e variância σ^2 finitas, a distribuição da soma

$$X = X_1 + \cdots + X_n$$

se aproxima à medida que n cresce da distribuição de $Y \sim \mathrm{N}(\mu_X, \sigma_X^2)$, em que $\mu_X = n\mu$ e $\sigma_X^2 = n\sigma^2$.

Teorema do Limite Central

Aproximação para *n* Grande

$$\begin{aligned} P(a \leq X \leq b) &\cong P(a \leq Y \leq b) \\ &= P\left(\frac{a - n\mu}{\sigma\sqrt{n}} \leq Z \leq \frac{b - n\mu}{\sigma\sqrt{n}}\right), \end{aligned}$$

em que $Z \sim N(0, 1)$.

Observação: correção de continuidade pode ser aplicada apenas para variáveis aleatórias discretas, tais como binomial e Poisson.

Teorema do Limite Central

Enunciado para a Média Amostral

Para variáveis aleatórias $X_1, \ldots, X_n$ independentes e com mesma distribuição de média μ e variância σ^2 finitas, a distribuição da média amostral

$$\bar{X} = \frac{X_1 + \cdots + X_n}{n}$$

se aproxima à medida que n cresce da distribuição de $Y \sim N(\mu_{\bar{X}}, \sigma^2_{\bar{X}})$, em que $\mu_{\bar{X}} = \mu$ e $\sigma^2_{\bar{X}} = \frac{\sigma^2}{n}$.

Teorema do Limite Central

Aproximação para *n* Grande

$$
\begin{aligned}
P(a \leq \bar{X} \leq b) &\cong P(a \leq Y \leq b) \\
&= P\left(\frac{a-\mu}{\sigma/\sqrt{n}} \leq Z \leq \frac{b-\mu}{\sigma/\sqrt{n}}\right),
\end{aligned}
$$

em que $Z \sim N(0, 1)$.

Teorema do Limite Central

Exemplo 1

Uma loja recebe em média 16 clientes por dia com desvio padrão de 4 clientes. Calcule aproximadamente a probabilidade de num período de 30 dias a loja receber mais do que 500 clientes. Calcule também a probabilidade aproximada de nesse mesmo período a média de clientes ultrapassar a 18 clientes.

Teorema do Limite Central

Dados do Problema

Seja U:número de clientes que a loja recebe num dia. Temos que

- $\mathrm{E}(U) = \mu = 16$
- $\mathrm{Var}(U) = \sigma^2 = 4^2 = 16$

Soma Amostral

Seja X:número de clientes que a loja recebe em 30 dias. Temos que

- $\mu_X = n \times \mu = 30 \times 16 = 480$
- $\sigma_X^2 = n \times \sigma^2 = 30 \times 16 = 480$
- $\sigma_X = \sqrt{480} \cong 21,91$

Teorema do Limite Central

Média Amostral

Seja $\bar{X}$:número médio de clientes que a loja recebe em 30 dias.
Temos que

- $\mu_{\bar{X}} = \mu = 16$
- $\sigma^2_{\bar{X}} = \frac{\sigma^2}{n} = \frac{16}{30} \cong 0,533$
- $\sigma_{\bar{X}} = \sqrt{0,533} \cong 0,73$

Teorema do Limite Central

Cálculo da Probabilidade

A probabilidade da loja receber mais do que 500 clientes em 30 dias fica dada por

$$
\begin{aligned}
P(X \geq 501) &\cong P\left(Z \geq \frac{501 - \mu_X}{\sigma_X}\right) \\
&= P\left(Z \geq \frac{501 - 480}{21,91}\right) \\
&= P(Z \geq 0,96) \\
&= 1 - P(Z \leq 0,96) \\
&= 1 - A(0,96) \\
&= 1 - 0,8315 \\
&= 0,1685(16,85\%).
\end{aligned}
$$

Teorema do Limite Central

Cálculo da Probabilidade

A probabilidade da média de clientes ultrapassar 18 clientes em 30 dias fica dada por

$$
\begin{aligned}
P(\bar{X} > 18) &\cong P\left(Z > \frac{18 - \mu_{\bar{X}}}{\sigma_{\bar{X}}}\right) \\
&= P\left(Z > \frac{18 - 16}{0,73}\right) \\
&= P(Z > 2,74) \\
&= 1 - P(Z \leq 2,74) \\
&= 1 - A(2,74) \\
&= 1 - 0.9969 \\
&= 0,0031(0,31\%).
\end{aligned}
$$