**Faculdade de Filosofia, Letras e Ciências Humanas**
Pós-Graduação em Ciência Política
Graduação em Ciências Sociais

# FLS 5028 - Métodos Quantitativos e Técnicas de Pesquisa em Ciência Política
# FLP 0406 -Métodos e Técnicas de Pesquisa em Ciência Política

## Aula 9 – Testes de Hipóteses I

Glauco Peres da Silva

1º sem / 2021

# Inferência Estatística

Estimação de Ponto e de Intervalo – Como utilizamos dados amostrais para estimar valores dos parâmetros populacionais?

Teste de Hipótese -- Como usamos métodos estatísticos para tester hipóteses?

# Estimação

*Objetivo*: Como utilizamos dados amostrais para estimar valores dos parâmetros populacionais?

**Estimativa de Ponto**: Um único valor estatístico que é o “melhor palpite” para o valor do parâmetro.

**Estimativa de Intervalo**: Um intervalo de números em torno do ponto estimado, que apresenta um “nível de confiança” de conter o verdadeiro parâmetro populacional. É chamado de ***Intervalo de confiança***.

# Erro amostral

Definição: O erro de uma estatística que ocorre quando nos baseamos em uma amostra para prever o parâmetro populacional.

# Estimação

Problema: O Erro padrão é desconhecido ($\sigma$ é também um parâmetro). É estimado pela substituição de $\sigma$ por sua estimativa a partir dos dados amostrais.

| Proporção | Estimativa de Ponto |
|---|---|
| $\pi = proporção$ | $\hat{\pi} = proporção$ |
| $\sigma = \sqrt{\pi(1-\pi)}$ | $s = \sqrt{\hat{\pi}(1-\hat{\pi})}$ |
| $\sigma_{\hat{\pi}} = \frac{\sqrt{\pi(1-\pi)}}{\sqrt{n}}$ | $s_{\hat{\pi}} = \frac{\sqrt{\hat{\pi}(1-\hat{\pi})}}{\sqrt{n}}$ |

## Intervalo de Confiança para uma Proporção

A distribuição amostral de uma proporção amostral de um grande número de amostras aleatórias é aproximadamente normal (Teorema do Limite Central)

Assim, com probabilidade de 0,95, a proporção amostral $\hat{\pi}$ estará a 1,96 erros padrões da proporção populacional π:

$$\hat{\pi}\ estar\acute{a}\ entre\ \ \pi\ - 1{,}96\sigma_{\hat{\pi}}\ e\ \ \pi + 1{,}96\sigma_{\hat{\pi}}$$

Teremos 0,95 de probabilidade que :

$$\hat{\pi}\ - 1{,}96\sigma_{\hat{\pi}}\ at\acute{e}\,\hat{\pi} + 1{,}96\sigma_{\hat{\pi}}\ contenha\ \pi$$

Uma vez que uma amostra é selecionada, temos 95% de confiança.

Este é um IC para uma população de proporção π (quase)

## Intervalo de Confiança para uma Proporção

Complicação: O verdadeiro erro padrão depende de um parâmetro desconhecido!

Efetivamente, estimamos

$$\sigma_{\hat{\pi}} = \sqrt{\frac{\pi(1-\pi)}{n}} \text{ por } s_{\hat{\pi}} = \sqrt{\frac{\hat{\pi}(1-\hat{\pi})}{n}}$$

E então encontramos o IC com 95% usando a fórmula

$$\hat{\pi} - 1{,}96 s_{\hat{\pi}} \ at\acute{e}\, \hat{\pi} + 1{,}96 s_{\hat{\pi}}$$

## Intervalo de Confiança

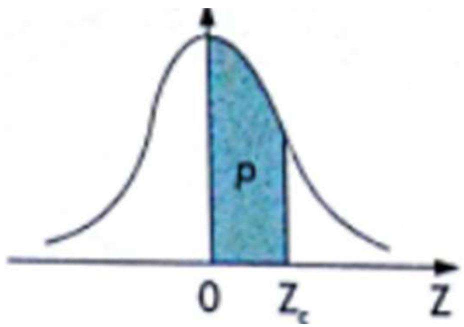
Se repetidamente tirarmos amostras aleatórias de algum tamanho fixo *n* e calcularmos para cada uma um IC com 95%, no longo prazo, em 95% dos IC's encontraríamos a proporção populacional π.

A probabilidade de que o IC *não* contenha π é chamado de **nível de significância**, e é chamado de α.

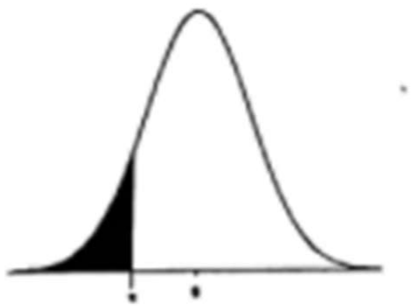
| $(1-\alpha)100\%$ | $\alpha$ | $\alpha/2$ | $z_{\alpha/2}$ |
|---|---|---|---|
| 90% | .10 | .050 | 1.645 |
| 95% | .05 | .025 | 1.96 |
| 99% | .01 | .005 | 2.58 |

# Intervalo de Confiança

O valor da tabela do Z-score varia de acordo com a padronização de cada tabela. Exemplos:







Os softwares atualmente apresentam o resultado final corretamente, mas o uso da tabela pode ser dificultado pela diferença em sua organização. Porém, o resultado é sempre o mesmo

## Estimação

Problema: O Erro padrão é desconhecido ($\sigma$ é também um parâmetro). É estimado pela substituição de $\sigma$ por sua estimativa a partir dos dados amostrais.

| **População** | **Estimativas de ponto da Distribuição Amostral** |
|---|---|
| $\mu$ | $\hat{\mu} = \bar{y} = \frac{1}{n}\sum y_i$ |
| $\sigma_y = \sqrt{(y-\mu)^2}$ | $\hat{\sigma}_y = s_y = \sqrt{\frac{\sum_{i=1}^{n}(y_i-\bar{y})^2}{n-1}}$ |
| $\sigma_{\bar{y}} = \frac{\sigma_y}{\sqrt{n}}$ | $\hat{\sigma}_{\bar{y}} = \mathrm{s}_{\bar{y}} = \frac{\mathrm{s_y}}{\sqrt{\mathrm{n}}}$ |

## Intervalo de Confiança para a média populacional

Em amostras aleatórias grandes,

A média amostral tem aproximadamente uma distribuição normal em torno da média com média $\bar{y}$ e erro padrão :

$$s_{\bar{y}} = \frac{s_y}{\sqrt{n}}$$

Um intervalo de confiança de 95% para $\mu$:

$$\bar{y} \pm 1{,}96 S_{\bar{y}}, \text{ onde } \bar{y} \pm 1{,}96 \frac{S_y}{\sqrt{n}}$$

# Distribuições normais

Normalização:

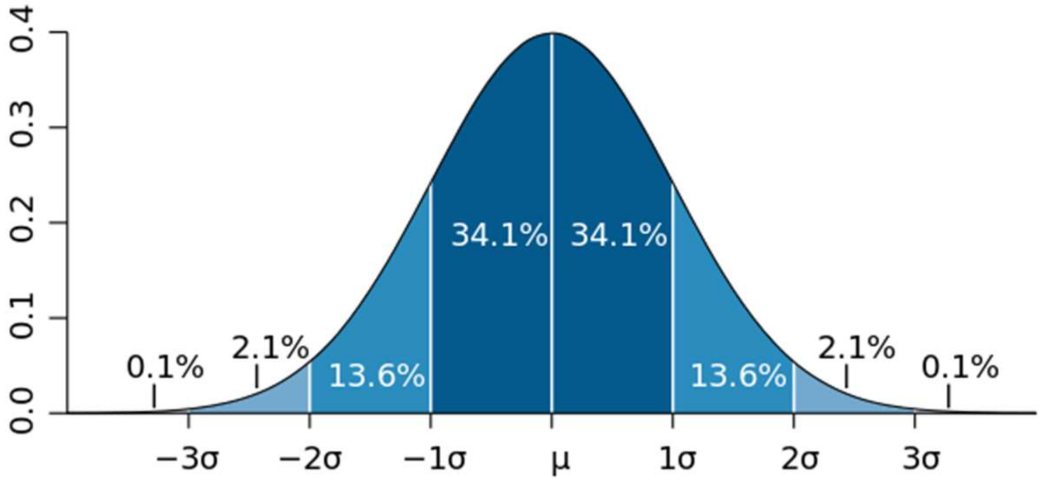
População

$$Z = \frac{Y - \mu}{\sigma} \sim N(0,1)$$

Amostra

$$Z = \frac{y - \bar{y}}{S_{\bar{y}}} \sim \mathrm{N}(0, 1)$$

# Distribuição Normal




$$P(\mu - \sigma \leq y \leq \mu + \sigma) \approx 0.68$$
$$P(\mu - 2\sigma \leq y \leq \mu + 2\sigma) \approx 0.95$$
$$P(\mu - 3\sigma \leq y \leq \mu + 3\sigma) \approx 0.997$$

# Intervalo de Confiança para a Média da População

Para "*n* grande," $S_y$ é uma boa estimativa de σ (e TLC vale).

Para *n* pequeno, *substituir* σ por sua estimativa $S_y$ introduz um erro extra, e o IC não é largo o suficiente, a não ser que substituamos *z*-score por um "*t*-score" maior.

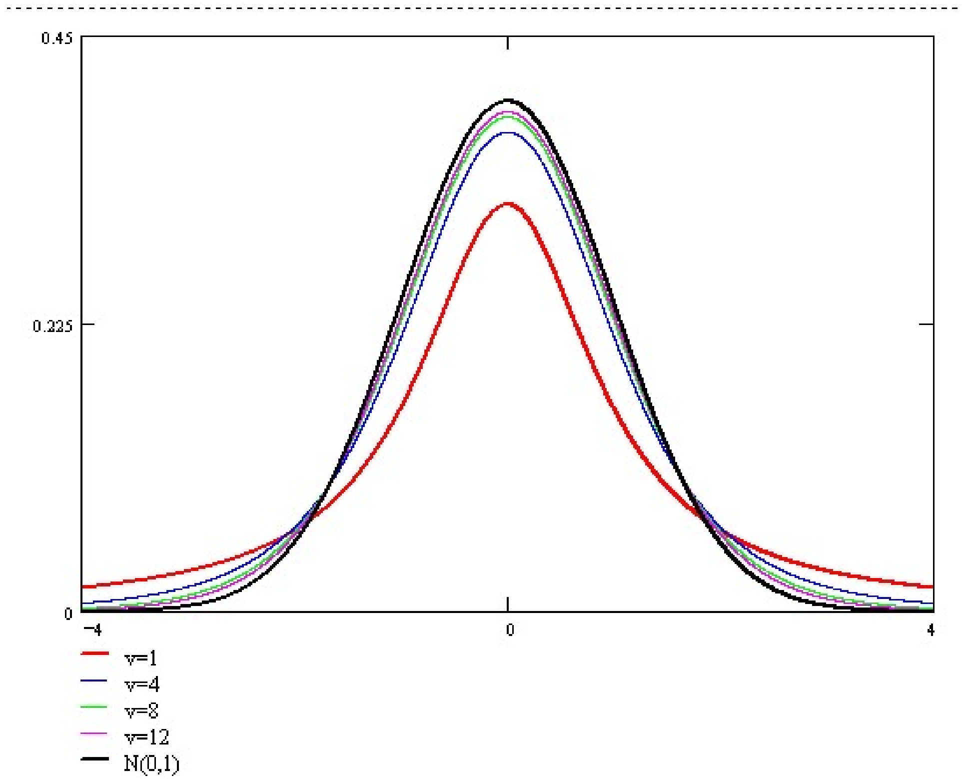
O formato preciso depende dos **graus de liberdade** (*gl*). Para inferências sobre a média, *gl = n – 1*

# Distribuição t-Student

| | *Nível de Confiança* | | | |
|---|---|---|---|---|
| | 90% | 95% | 98% | 99% |
| gl | t.050 | t.025 | t.010 | t.005 |
| 1 | 6.314 | 12.706 | 31.821 | 63.657 |
| 10 | 1.812 | 2.228 | 2.764 | 3.169 |
| 16 | 1.746 | 2.120 | 2.583 | 2.921 |
| 30 | 1.697 | 2.042 | 2.457 | 2.750 |
| 100 | 1.660 | 1.984 | 2.364 | 2.626 |
| infinito | 1.645 | 1.960 | 2.326 | 2.576 |

gl = ∞ corresponde a distribuição normal padrão

# Distribuição t-Student



# Intervalos de Confiança

O intervalo de confidança é ele próprio uma quantidade aleatória, sujeita a variabilidade amostral.

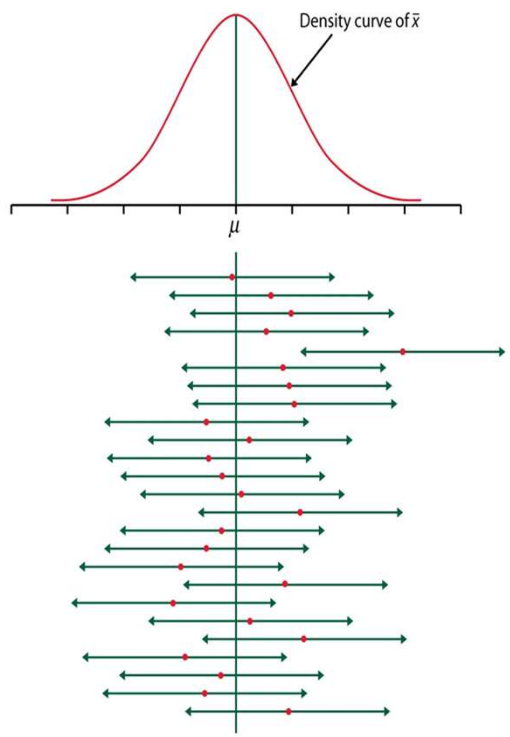
Maior grau de confiança exige IC mais largos.

Maiores *n* produzem ICs mais estreitos.

# Intervalos de Confiança

O intervalo de confidança é ele próprio uma quantidade aleatória, sujeita a variabilidade amostral.

No "longo prazo," 95% de ICs de 95% para uma média populaciona μ conterá o real valor de μ (No gráfico, cada linha mostra um IC para uma amostra particular com sua própria média amostral, extraída de uma curva de *distribuição amostral* de médias amostrais possíveis).



# Testes de Significância

Um teste de significância usa dados para avaliar uma hipótese ao comparar uma estimativa de um ponto de um parâmetro com valores preditos pela hipótese.

1. Pressupostos (tipo de distribuição de probabilidade, amostra aleatória, $\alpha = 1\%, 5\%, 10\%$)
2. Hipóteses (*hipótese nula* e *hipótese alternativa*. São denotadas $H_0$ e $H_1$, respectivamente)
3. Estatística de teste (o número de erros padrão entre a estimativa amostral do ponto e a $H_0$ do valor do parâmetro)
4. Uma probabilidade mensura a evidência sobre $H_0$. (Se o P-valor não é suficientemente pequeno, nós não rejeitamos $H_0$ (então, $H_0$ não é necessariamente verdadeiro, mas é plausível)
5. Conclusão

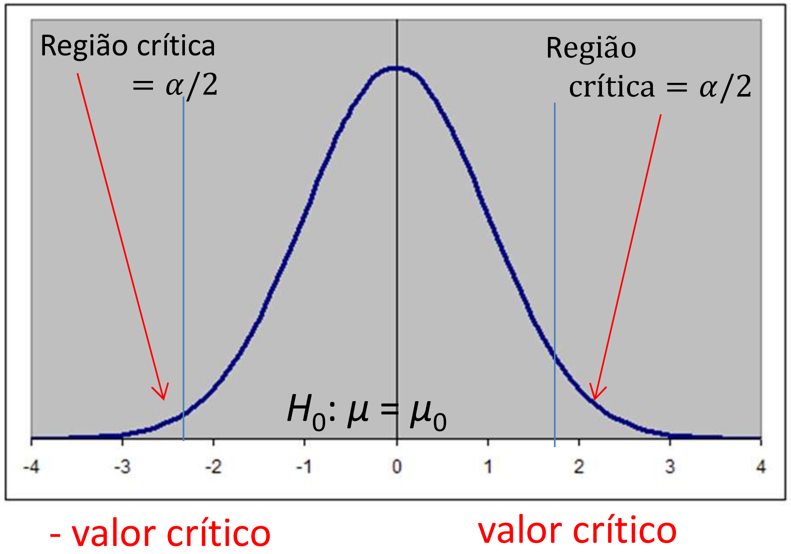
# Teste de Significância

Teste Bicaudal para a média amostral

$\alpha = 5\%$ = nível de significância do teste

tc= valor crítico da estatística de teste

$H_0$: $\mu = \mu_0$
$H_a$: $\mu \neq \mu_0$




$$t = \frac{\bar{y} - \mu_0}{s_{\bar{y}}} \quad \text{where} \quad s_{\bar{y}} = s_y/\sqrt{n}$$

p-valor = a probabilidade da estatística-t se $H_0$ é falsa.

Conclusão
se $P$-valor ≤ $\alpha$, “rejeitamos $H_0$” / se $P$-valor > $\alpha$, “não rejeitamos $H_0$”

# Teste de Significância

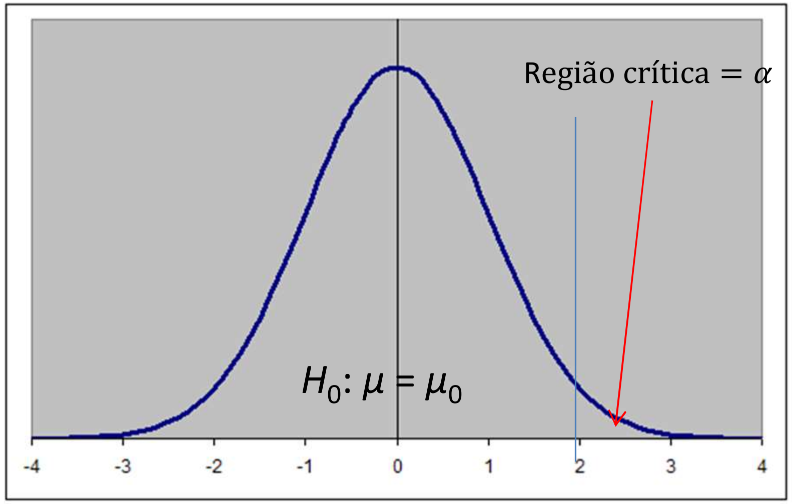
Teste monocaudal para a média amostral

$\alpha = 5\%$ = nível de significância do teste

tc= valor crítica da estatística do teste

$H_0$: $\mu = \mu_0$
$H_a$: $\mu > 0$

$$t = \frac{\bar{y} - \mu_0}{s_{\bar{y}}} \quad \text{where} \quad s_{\bar{y}} = s_y/\sqrt{n}$$




p-valor = a probabilidade da estatística-t se $H_0$ é verdadeiro.

# Decisões e Tipos de Erros

| | Decisão | |
|---|---|---|
| | Rejeitar Ho | Não rejeitar Ho |
| Condição Verdadeira | | |
| Ho verdadeiro | Erro tipo I $\alpha$ | Decisão Correta |
| Ho falso | Decisão Correta | Erro tipo II $= \beta$ |

Analogia ao Sistema Judicial
$H_0$: Acusado é *inocente*
$H_a$: Acusado é *guilty*

# Limitações dos Testes de Significância

1. Diferença entre significância estatística e significância prática
2. Testes de significância são menos úteis do que os Intervalos de Confiança
3. Interpretações errôneas dos testes de significância e dos p-*values*
    1. *Relatar resultados apenas se forem estatisticamente significativos;*
    2. *Há chance da significância ocorrer apenas por acaso;*
    3. *O p-value não significa que Ho é verdadeira;*
    4. *Efeitos verdadeiros podem ser menores do que as estatísticas relatadas.*

# Erro Tipo II

Para calcularmos o Erro Tipo II dependemos de saber qual a média real

Ex.: Experimento com os astrólogos (acertaram 40 de 116)

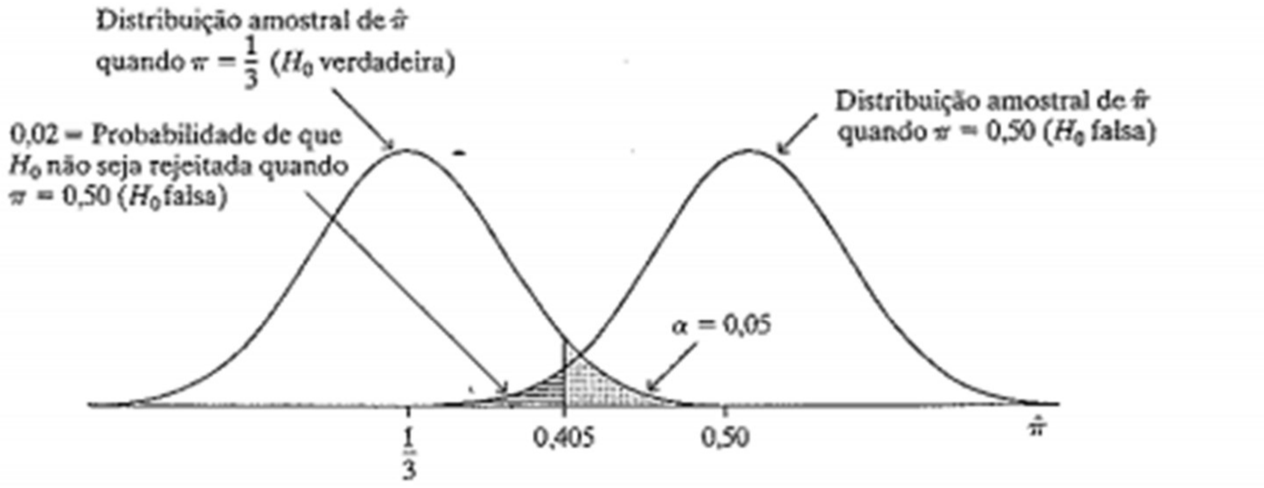



**Figura 6.11** Cálculo de P(erro do Tipo II) para testar $H_0: \pi = 1/3$ contra $H_a: \pi > 1/3$ no nível $\alpha = 0,05$, quando a proporção verdadeira é $\pi = 0,50$. Um erro do Tipo II ocorre se $\hat{\pi} < 0,405$, visto que, então, o valor-$p > 0,05$ embora $H_0$ seja falsa.

# Poder de um Teste

O poder de um teste é uma medida que nos indica a probabilidade de rejeitar Ho quando ela for falsa. Esperamos que nesta situação a probabilidade de rejeitarmos seja de fato elevada.

Poder = 1 – P(Erro tipo II)

Nem sempre conseguimos calcular o poder de um teste, pois dependemos de uma situação em que temos ciência da verdadeira média populacional.