

# MAD-CB



## Erros Tipo I e II e Poder Estatístico

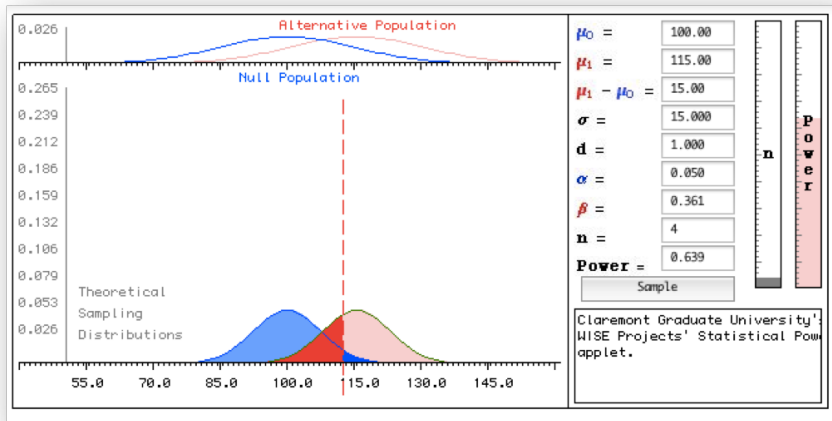
- Hipótese Nula ( $H_0$ ): parâmetro = valor teórico
- Hipótese Alternativa ( $H_1$ ):  $\geq$ ,  $\neq$ ,  $\leq$  valor teórico
- Estatística de teste calculada dos dados assumindo  $H_0$  verdade
- Valor  $p$  – probabilidade de observar valor teórico testado ou um valor mais extremo
- Valores  $p$  pequenos mostram evidência contra  $H_0$
- Nível de significância define limite para “quão pequeno é pequeno” de um valor  $p$
- Anotado como  $\alpha$  (letra grega “alfa”)

# Poder de Um Teste Estatístico

- $\alpha$  fornece uma ideia do desempenho do teste com amostras múltiplas
- Se  $H_0$  é verdade e  $\alpha = 0.01$ 
  - ▶ Em 1% das amostras, rejeitaremos  $H_0$  erroneamente
  - ▶ Em 99% das amostras, não rejeitaremos  $H_0$  (decisão correta)
- Queremos rejeitar  $H_0$  quando é realmente falsa
- Se  $\alpha$  é pequeno demais, podemos nunca rejeitar  $H_0$  mesmo se o valor correto do parâmetro seja muito diferente da  $H_0$
- **Poder** de um teste é a probabilidade de tomar a decisão correta (i.e., rejeitar  $H_0$  quando é realmente falsa)
- Um grau de poder mais alto de um teste significa que o teste fica mais sensível

# Poder – Applet

- Wise Power Applet – Claremont Graduate University, Claremont, CA
- [http://http://wise1.cgu.edu/power\\_applet/power.asp](http://http://wise1.cgu.edu/power_applet/power.asp)



*Poder é a probabilidade de rejeitar corretamente uma hipótese nula falsa quando uma hipótese alternativa é verdade*

- Permite determinar a probabilidade de um teste de significância estatística rejeitar a hipótese nula
- Permite calculo de número de casos ( $n$ ) que seria necessária na amostra para conseguir um certo patamar de poder.
- Mais simples: permite que determinamos *a priori* a probabilidade que tomaremos a decisão certa

## 2 Tipos de Erros Estatísticos

- **Tipo I** – Rejeitar  $H_0$  quando **verdade**

- ▶ Ocorre com probabilidade  $\alpha$
- ▶ Uma pessoa inocente condenada falsamente de um crime

- **Tipo II** – Não rejeitar  $H_0$  quando **falso**

- ▶ Ocorre com probabilidade  $\beta$
- ▶  $\beta = 1 - \text{poder}$
- ▶ Um criminal liberado em erro



# Resultados Possíveis de Um Teste de Hipótese

Conclusão do Teste	Estado de Natureza		
		Nula Verdadeira	Nula Falsa
	Não rejeitar $H_0$	Correta $p = 1 - \alpha$	Erro Tipo II $p = \beta$
	Rejeitar $H_0$	Erro Tipo I $p = \alpha$	Correta $p = 1 - \beta$

# Poder e o Tamanho de Amostra

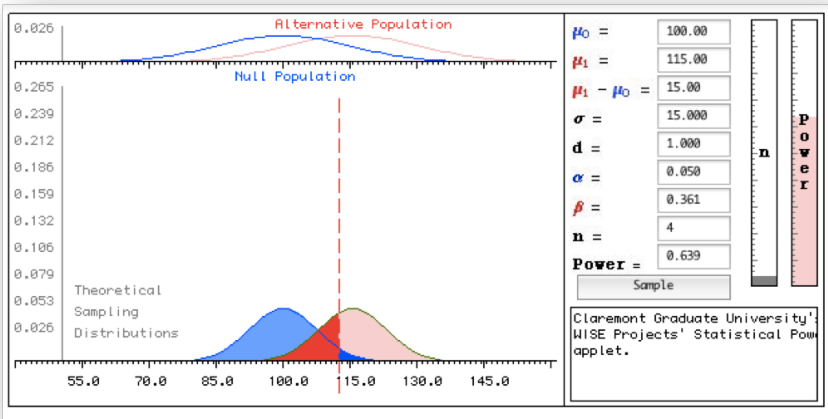
- Pesquisador presta atenção a 4 valores importantes:
  - ▶ Tamanho da amostra
  - ▶ Nível de significância ( $\alpha$ )
  - ▶ Poder
  - ▶ Tamanho de efeito
- Se você conhece três, pode calcular o quarto
- $n$  e  $\alpha$  são sob seu controle direto

- Meta normal: maximizar o poder enquanto manter o nível de significância e minimizar tamanho de amostra
- Quer maximizar probabilidade de achar um efeito verdadeiro enquanto minimizando chance de identificar um efeito que não existe
- Fazemos nossos cálculos de poder e tamanho de amostra ao início da pesquisa – no planejamento
  - ▶ Não depois – já executou experimento. Não pode tentar aumentar casos

# Poder – Diagrama



- Magnitude da diferença entre a média verdadeira na população e a média hipotetizada na nula relativa ao desvio padrão
- Quando efeito de tamanho é maior, poder maior
  - ▶ Existe menos sobreposição entre as duas curvas



- Desses valores, mais difícil para entender
- Normalmente precisa experiência com pesquisas para aplicar bem
- Tabela para guiar calouros estatísticos no uso de tamanho de efeito
- Esses são indicações gerais
  - ▶ Com experiência, vai ganhar melhor ideia

# Tabela de Tamanhos de Efeito Exemplares

	<b>Pequeno</b>	<b>Médio</b>	<b>Grande</b>
teste - t	0,2	0,5	0,8
modelo linear	0,02	0,15	0,35
proporções	0,2	0,5	0,8
Qui-quadrado ( $\chi^2$ )	0,1	0,3	0,5



## Como Calcular o Poder

# Trabalhar com os Dados de Temperatura Normal Humana

```
temps <- read_table("TempData.txt", col_names = FALSE)
colnames(temps) <- "tempC"
suppressMessages(library(psych))
summary(temps)
```

```
##      tempC
##  Min.   :35.72
##  1st Qu.:36.56
##  Median :36.83
##  Mean   :36.81
##  3rd Qu.:37.06
##  Max.   :38.22
```

```
dp <- sd(temps$tempC); paste ("Desvio Padrão =", dp) # desvio padrão
```

```
## [1] "Desvio Padrão = 0.407323976688302"
```

```
n <- length(temps$tempC); paste ("n =", n)
```

```
## [1] "n = 130"
```

# Cálculo de Poder no Pacote pwr

- Pode especificar 3 dos quatro componentes de poder; função calcula o outro

```
pwr.t.test(n = NULL, d = NULL, sig.level = 0.05, power = NULL,  
  type = c("two.sample", "one.sample", "paired"),  
  alternative = c("two.sided", "less", "greater"))
```

- $d$  = tamanho de efeito
- $\text{sig.level} = \alpha$
- Deve especificar `type` e `alternative`

# Executar Função `pwr.t.test`

```
## Se não tiver carregado o pacote pwr, precisa instalar.  
## Tira a marca de comentário na próxima linha para ativar  
# install.packages("pwr")  
## Se já tem, pode ir diretamente ao próximo comando  
library(pwr)  
pwr.t.test(n = n, d = 0.8, sig.level = 0.05,  
           type = "one.sample", alternative = "two.sided")
```

```
##  
##      One-sample t test power calculation  
##  
##              n = 130  
##              d = 0.8  
##      sig.level = 0.05  
##              power = 1  
##      alternative = two.sided
```

# Conclusão sobre Amostra de 130 Pessoas

- Poder = 1; Amostra exagerada
- Vamos testar uma amostra de 10

```
pwr.t.test(n = 10, d = 0.8, sig.level = 0.05,  
           type = "one.sample", alternative = "two.sided")
```

```
##  
##      One-sample t test power calculation  
##  
##              n = 10  
##              d = 0.8  
##      sig.level = 0.05  
##      power     = 0.6162328  
##      alternative = two.sided
```

# Um Pouco Fraco

- O que seria o tamanho certo da amostra para conseguir um poder = 0.95?

```
pwr.t.test(d = 0.80, sig.level = 0.05, power = 0.95,  
           type = "one.sample", alternative = "two.sided")
```

```
##  
##      One-sample t test power calculation  
##  
##              n = 22.32453  
##              d = 0.8  
##      sig.level = 0.05  
##              power = 0.95  
##      alternative = two.sided
```

- Uma amostra de 23 seria suficiente para ter todo o poder estatístico que você quer
- Pode contar para seu colega montando o experimento que 23 seria um número adequado.

## Considerações Finais sobre Testes de Hipóteses

- ① Não interpretar erroneamente os valores de  $p$ 
  - ▶ Valor de  $p$  não informa você se  $H_0$  é verdade
  - ▶ O que diz é: quão provável seriam os dados observados se  $H_0$  for verdade
- ② A coleta de dados é o passo chave para chegar numa conclusão correta
- ③ Sempre usar testes de 2 lados senão existe certeza que 1 dos lados não tem interesse nenhum
- ④ Significância estatística  $\neq$  significância prática



<b>Moeda – 3 Amostras</b>	<b>1</b>	<b>2</b>	<b>3</b>
Dados	6/10	58/100	576/1000
Estatística Z	0,63	1,60	4,81
Valor p	0,53	0,11	0,000002

- ⑤ Falta de significância não quer dizer que  $H_0$  verdade
  - ▶ Valores grandes de  $p$  ocorrem por causa de:
    - ▶ Sorte
    - ▶ Problemas na coleta de dados
    - ▶  $H_0$  realmente falsa
- ⑥ Especificar hipóteses antes de coletar os dados
  - ▶ Siga os passos em ordem
- ⑦ Testes de proporção e da média precisam
  - ▶ Independência das observações
  - ▶ Distribuição dos estimadores aproximadamente normal
- ⑧ Teste t muito robusto
  - ▶ Aguenta todos os outliers fora dos mais extremos