

Análise do Poder do teste e definição de tamanho de amostra

BIE 5793

Princípios de Planejamento e Análise de Dados em Ecologia

Qual o tamanho ideal de uma amostra?

- Muitos palpites baseados em experiência

Gotelli & Ellison = Regra dos 10

Manly (2008)/ Quinn & Keough (2002) = Atenção para n abaixo de 20

- Técnica analítica que permite obter uma estimativa para alguns tipos de testes estatísticos -> Baseada na **Análise de Poder do Teste**

Análise de Poder do Teste (Power Analysis ou Power Test)

Poder estatístico:

Medida da confiança em detectar um efeito, se ele existe.

- Essa análise permite estimar o tamanho de amostra ideal para detectar com confiança um dado efeito que é considerado relevante
- O que é um **efeito relevante**?
- Populações reais sempre vão diferir em alguma quantidade, mas precisamos definir se essa quantidade importa ou não.

Como definir?

- Experiência do grupo de pesquisa
- Estudos anteriores

No nosso exemplo das sementes, vamos considerar que **55** é o número médio de sementes por planta em toda a região.

-> Se as populações da face sul tiverem em média **1** semente a mais que as populações da face norte, isso é importante para nós?

-> E se a diferença for **2**?

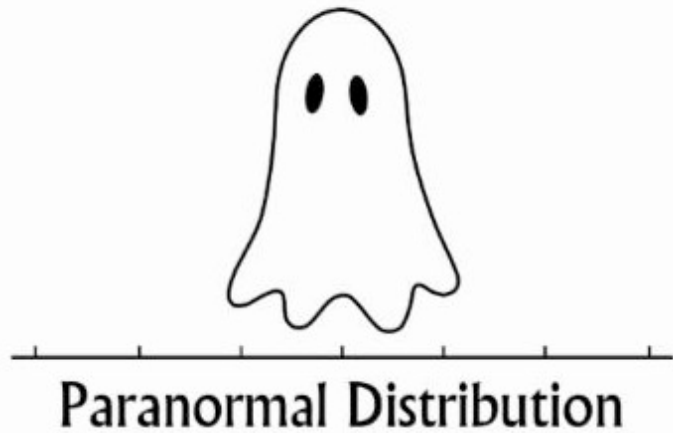
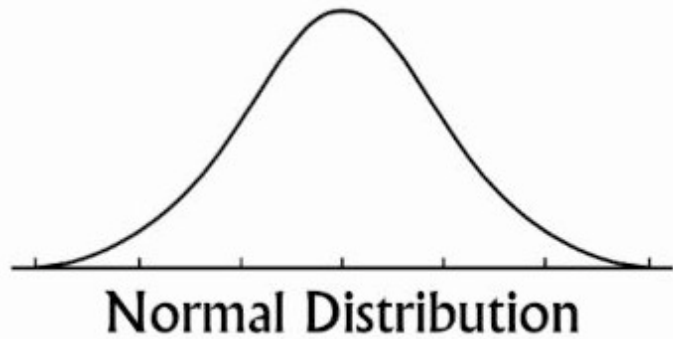
-> E se a diferença for **10**?

Por que essa pergunta? Esse valor ajuda a definir o tamanho amostral adequado

Quanto maior a precisão desejada, maior o número de amostras necessário para detectar essa diferença.

De onde vem essa estimativa?

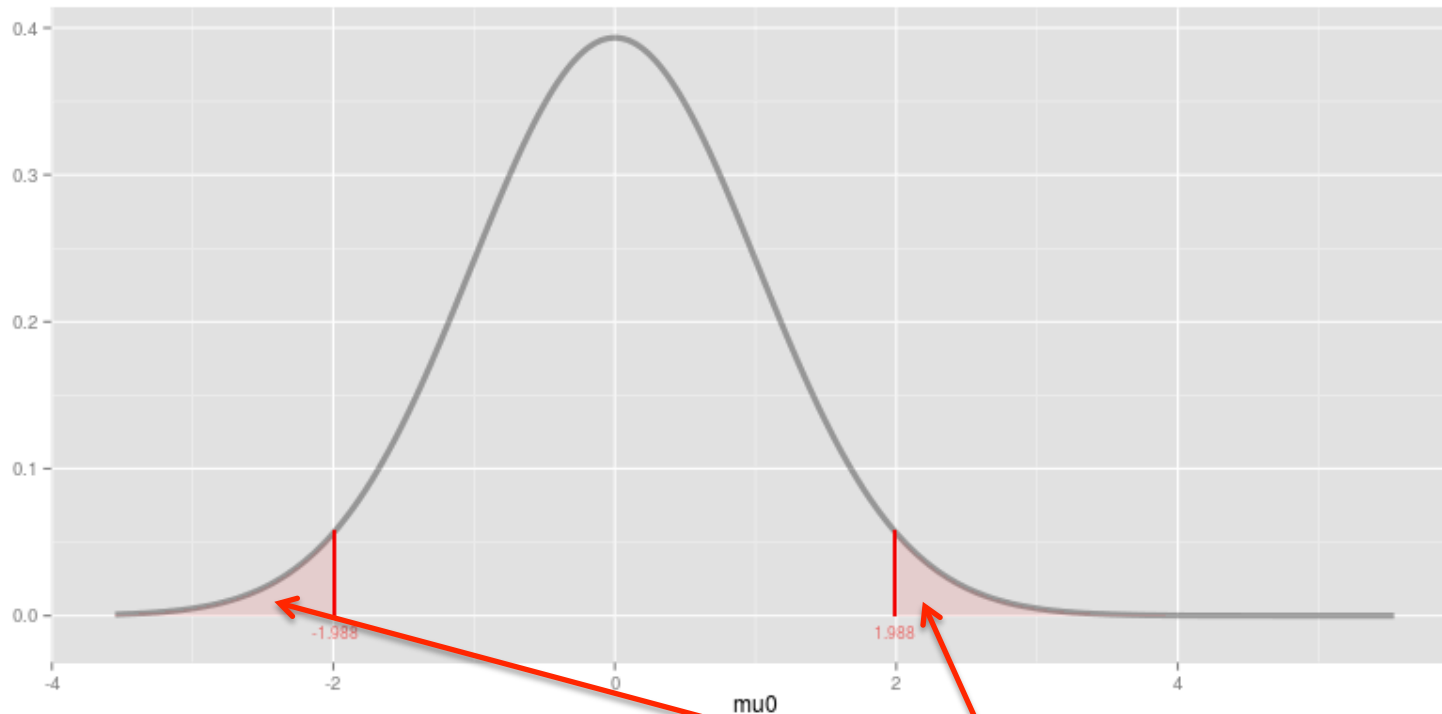
Voltando lá nas Distribuições de Probabilidades...



e Tipos de Erros...

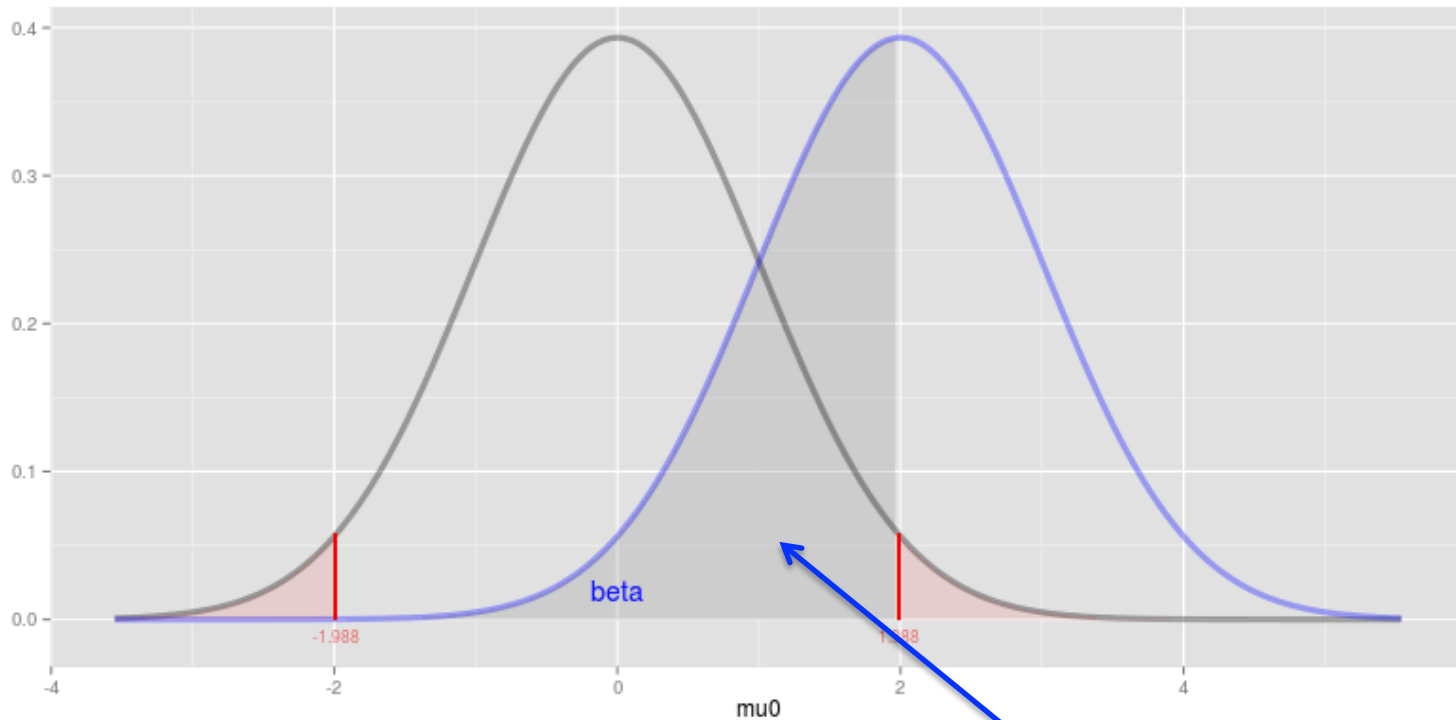


Considere uma distribuição de probabilidades para uma **Hipótese Nula** (centrada em zero), com desvio padrão = 6



Erro Tipo I = alfa

Agora, em azul, uma **distribuição para H1**, com média 2 e desvio padrão = 6

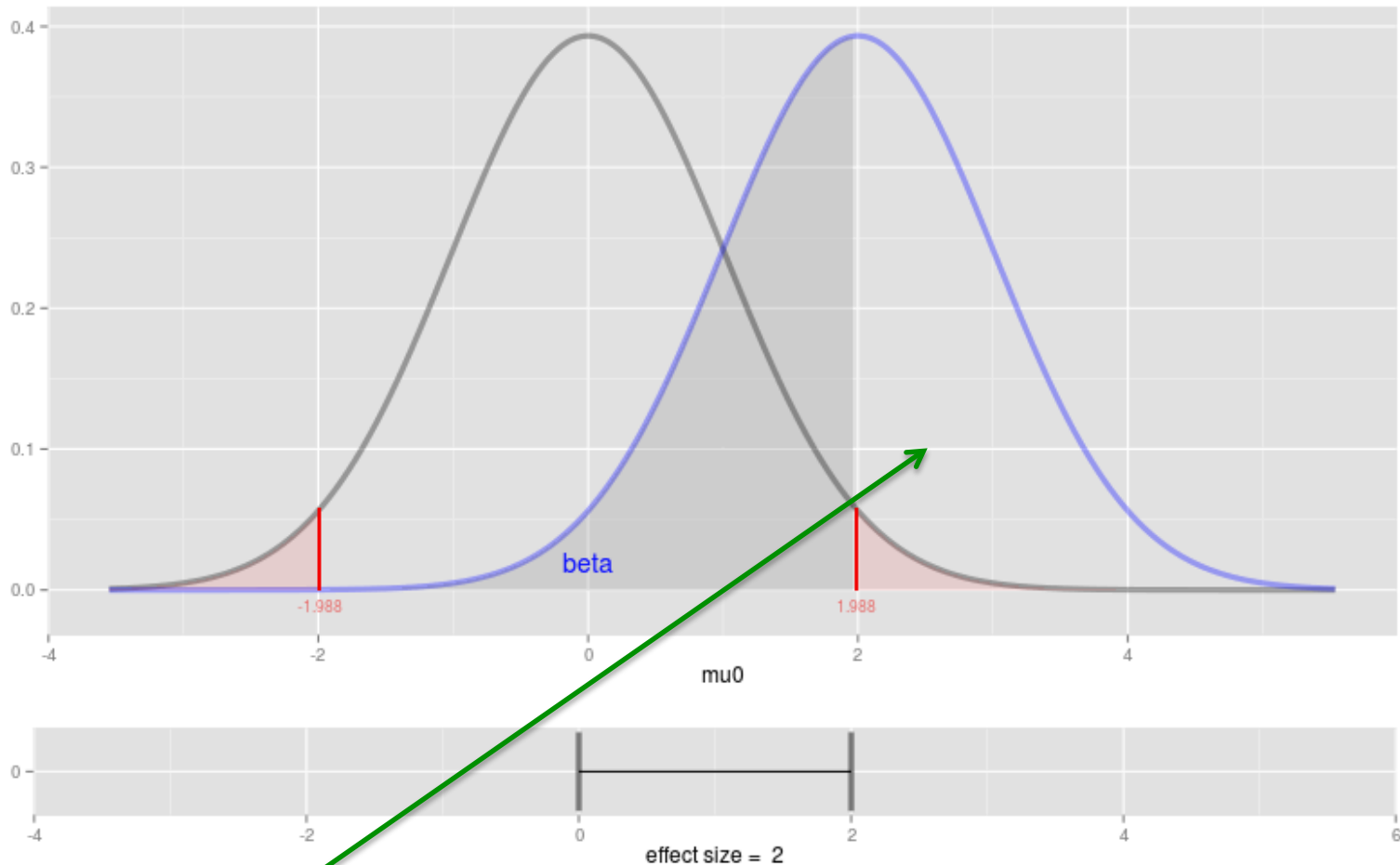


Em cinza escuro, a área ocupada por beta define o **Erro Tipo II (beta)**

Beta é a probabilidade de cometermos o erro Tipo II - Aceitar H_0 quando H_1 é verdadeira.

Para essa H_1 , beta é aproximadamente 0,48
Isso representa uma chance de 48% de aceitarmos H_0 erroneamente

O **PODER** do teste é dado por **1 - Beta**



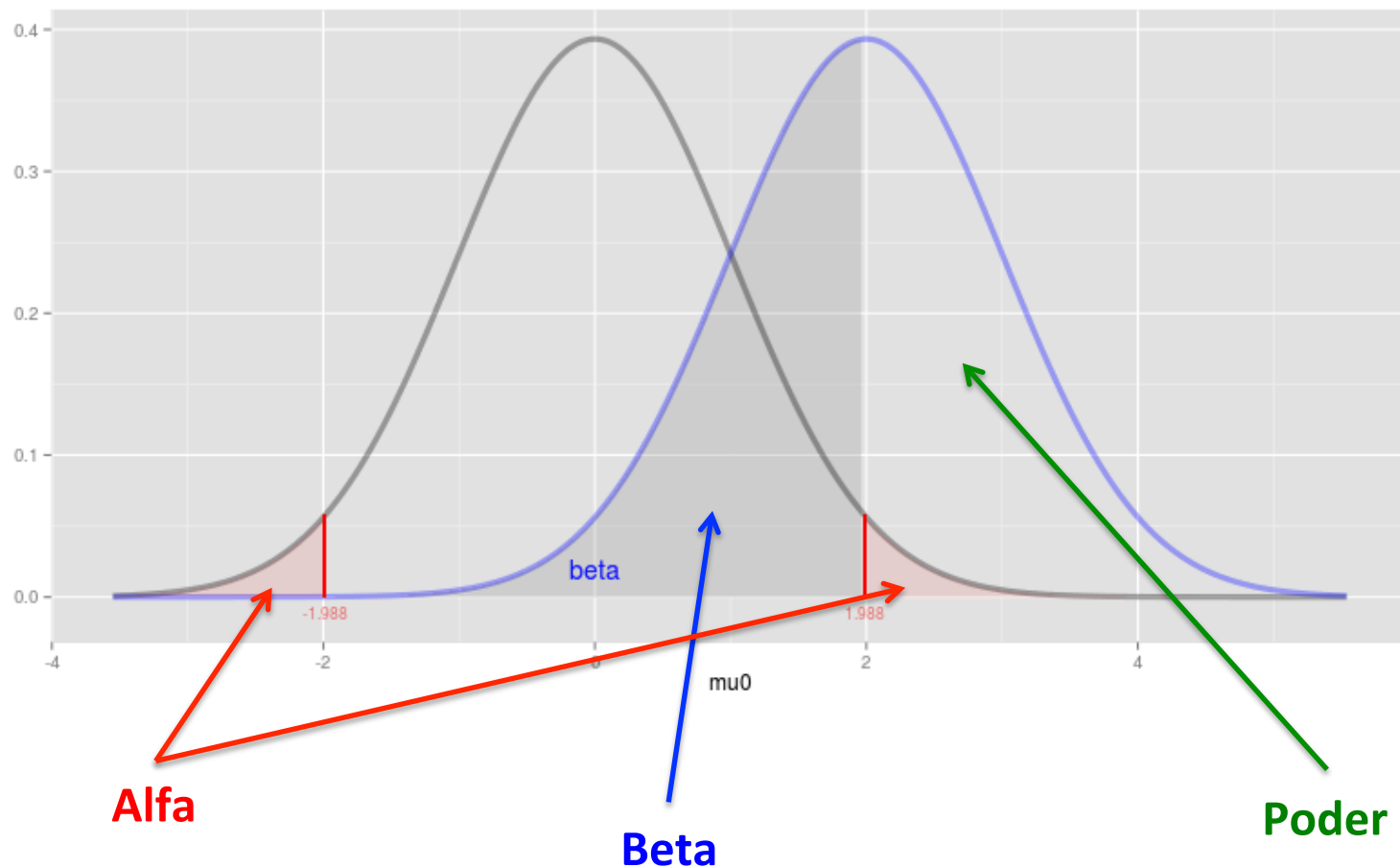
O **poder do teste** é a probabilidade representada pelo restante da área abaixo da curva azul, **excluindo-se a área do Erro Tipo II (ou beta)**

Resumindo

Alfa = $P(\text{Rejeitar } H_0 \mid H_0 \text{ seja verdadeira})$ ou **P(Erro Tipo I)**

Beta = $P(\text{Aceitar } H_0 \mid H_1 \text{ seja verdadeira})$ ou **P(Erro Tipo II)**

Poder do teste = $P(\text{Rejeitar } H_0 \mid H_1 \text{ seja verdadeira})$ ou **1 - Beta**



Poder do teste

$$\text{Power} \propto \frac{ES \alpha \sqrt{n}}{\sigma}$$

Onde:

ES (Effect Size) = Tamanho do efeito que queremos detectar

n = tamanho da amostra

σ (sigma) = desvio padrão

α (alfa) = nível de significância a ser considerado

Note o símbolo de **proporcional** -> depende do teste estatístico a ser analisado

O cálculo exato da relação entre essas medidas **depende do tipo de teste estatístico:**

Teste t com uma amostra

Teste t com duas amostras *

ANOVA

ANOVA com dois fatores

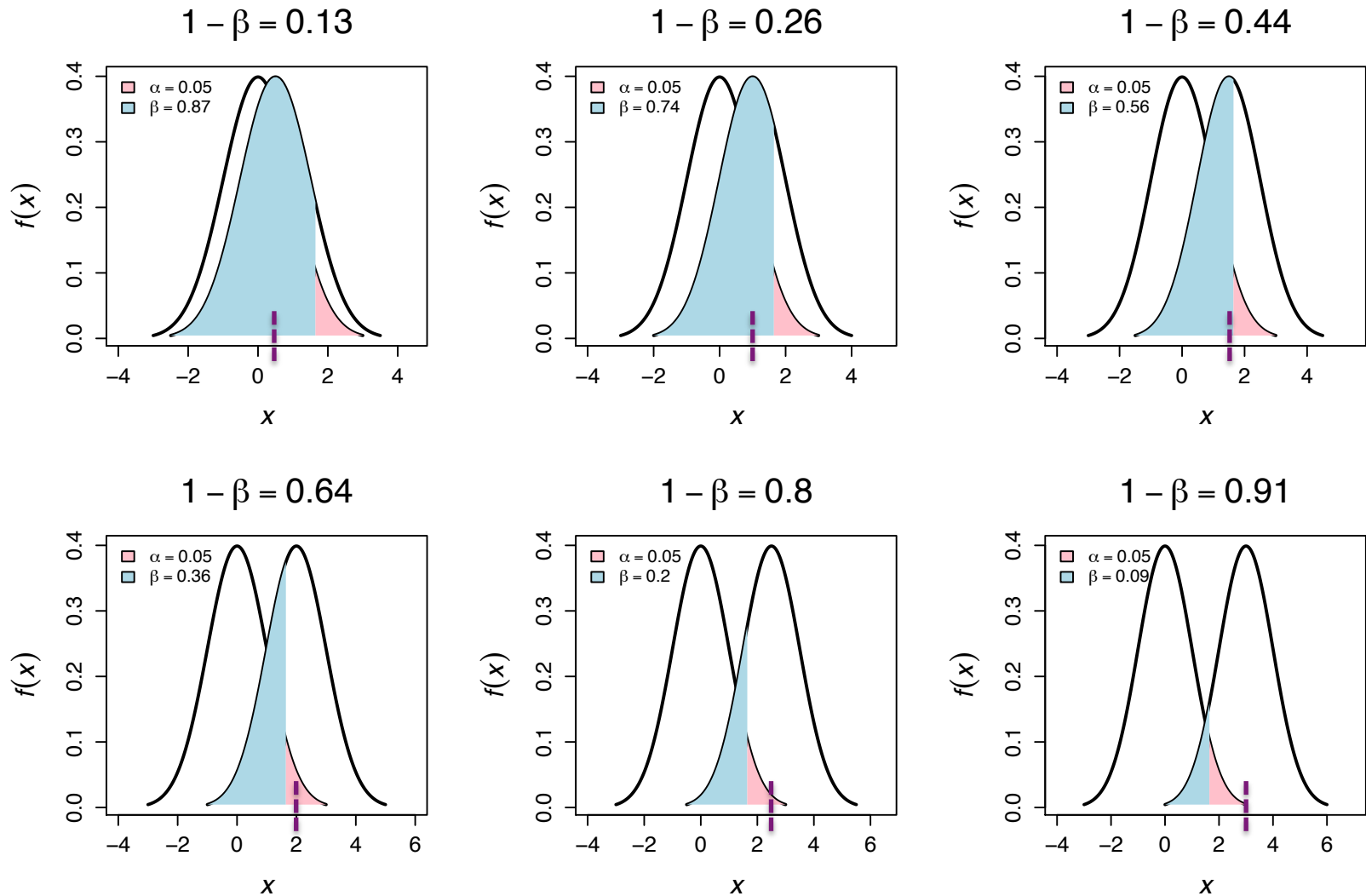
Regressão linear

Correlação

Etc...

Mas a direção das relações entre essas medidas são as mesmas para qualquer teste estatístico

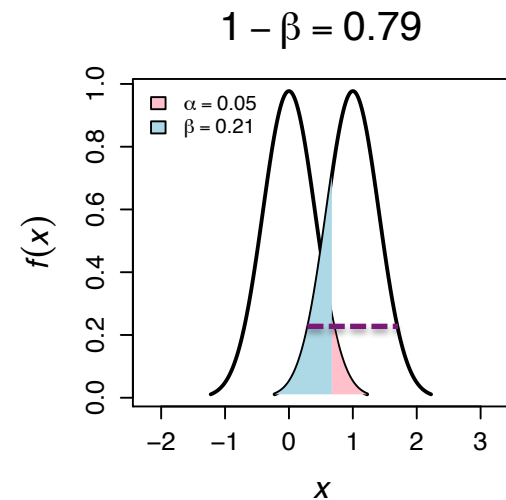
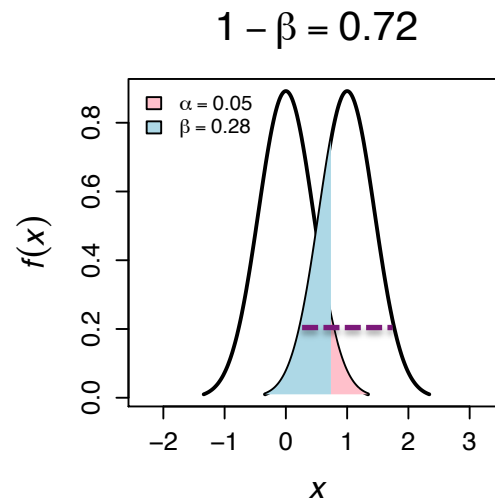
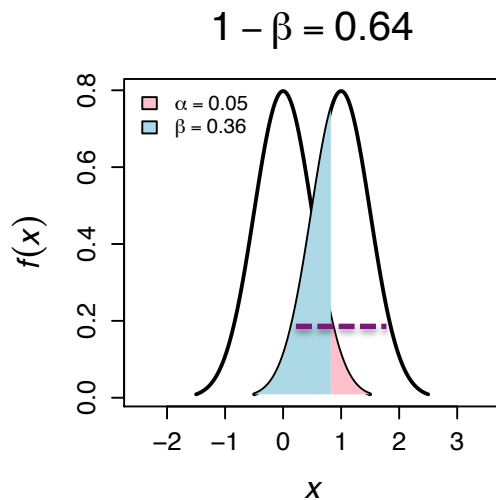
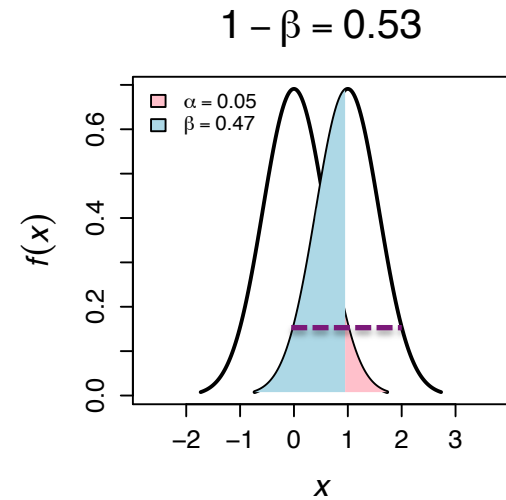
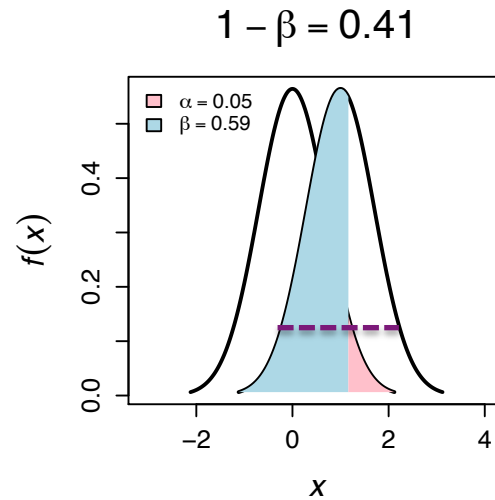
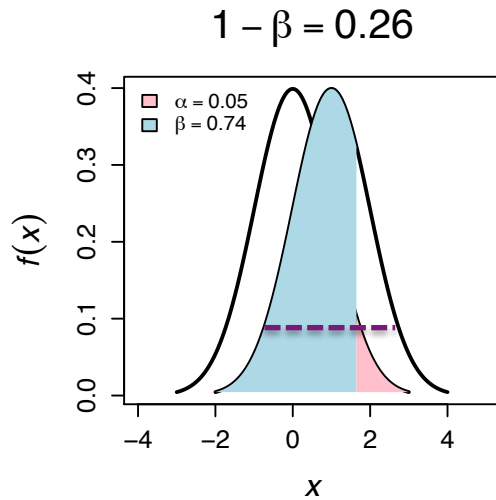
Quanto maior o Tamanho do Efeito (**ES**), maior o Poder ($1 - \beta$)



Nathaniel Helwig (2017)

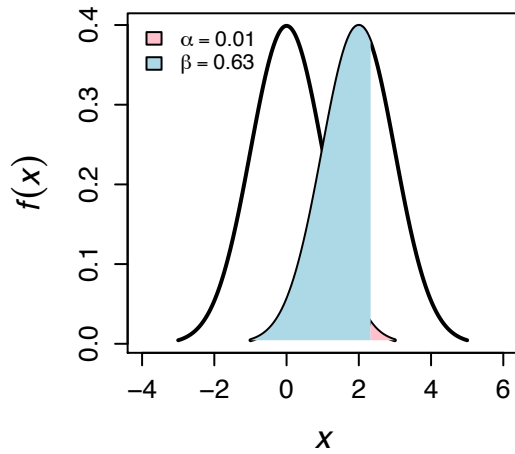
No nosso exemplo, o efeito seria a diferença de médias (em uma distribuição t)

Quanto menor o Desvio Padrão (sigma), maior o Poder (1- Beta)

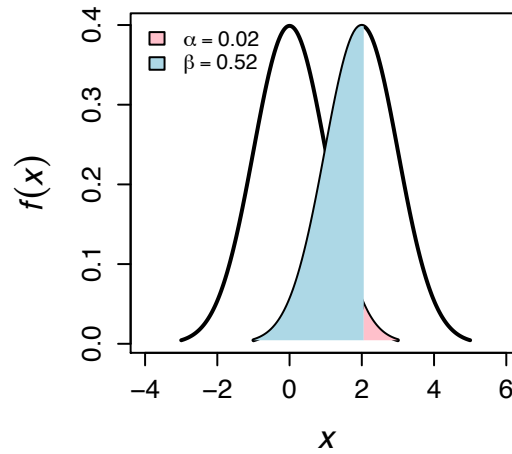


Quanto maior o valor de alfa, maior o Poder ($1 - \beta$)

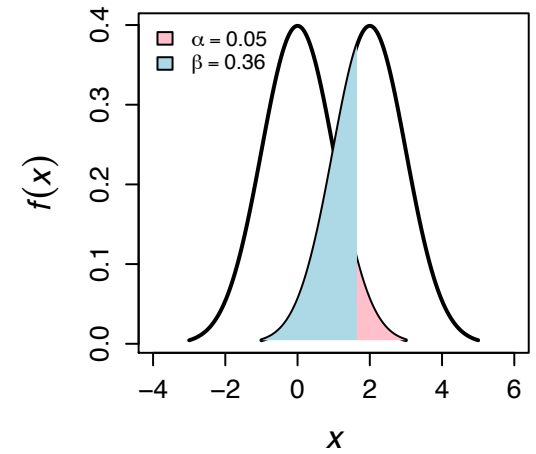
$1 - \beta = 0.37$



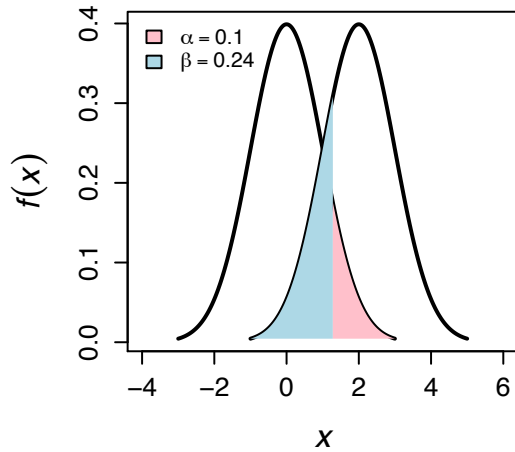
$1 - \beta = 0.48$



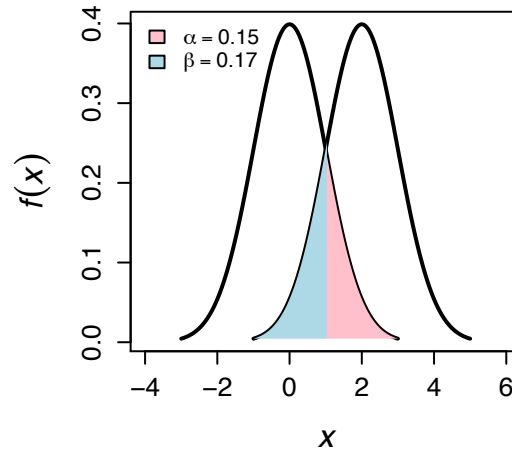
$1 - \beta = 0.64$



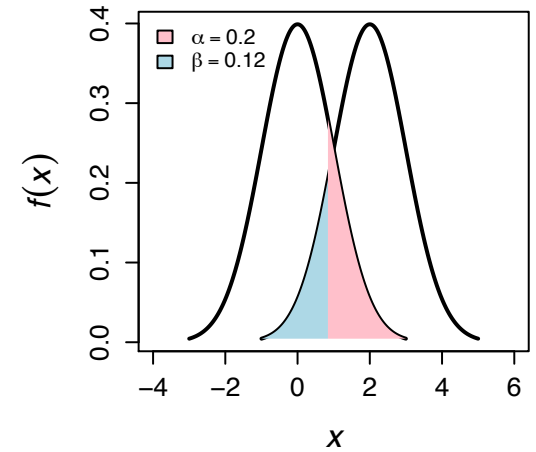
$1 - \beta = 0.76$



$1 - \beta = 0.83$



$1 - \beta = 0.88$

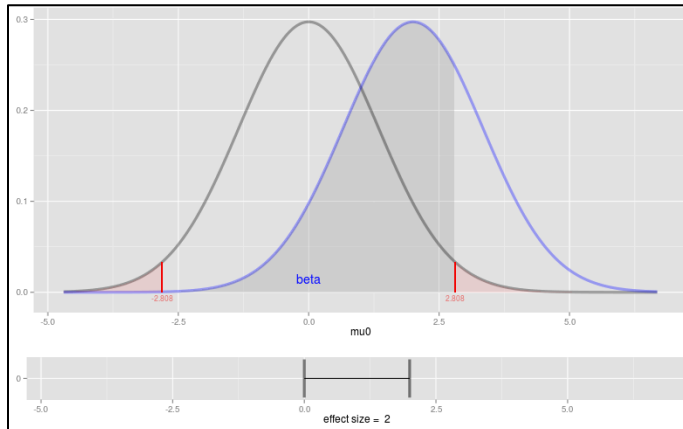


Nathaniel Helwig (2017)

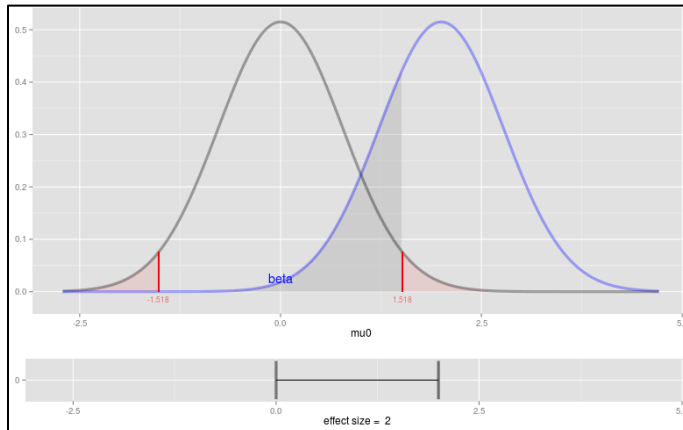
alfa e Beta são negativamente relacionados

Altos valores de alfa (menos restritivos) -> menor Beta -> maior Poder

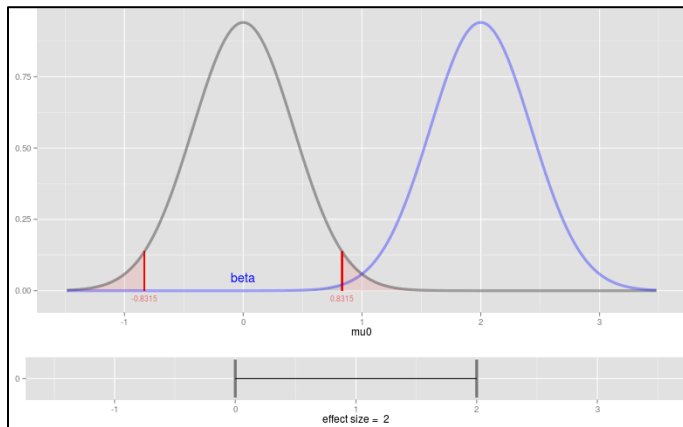
Quanto maior o tamanho da amostra (n), maior o Poder (1-Beta)



n=20 Poder = 0,28



n=60 Poder = 0,74



n=200 Poder = 0,99

Para cada tipo de teste estatístico é feito um cálculo diferente, mas envolvendo basicamente essas medidas ou suas variantes:

Poder (1 - Beta) = Confiança para estimar um dado efeito

ES (Effect Size) = Tamanho do efeito que queremos detectar

n = Tamanho da amostra

σ (sigma) = Desvio padrão

α (alfa) = Nível de significância a ser considerado

Como todas essas medidas estão relacionadas, se você quiser saber o valor de uma delas, basta estimar ou saber os valores das outras!!

Algumas convenções:

alfa = 0.05

Beta = 0.20

Poder = 1 - Beta = 0.8

Pausa para a Ecologia...

As estimativas do **tamanho do efeito (ES)** e do **desvio padrão (sigma)** devem ser baseadas em:

- Experiências prévias
- Literatura existente

O **desvio padrão (sigma)** pode ser estimado a partir de um projeto piloto



No nosso exemplo das sementes das plantas, o que seriam **tamanhos de efeito e desvios padrão** aceitáveis?

Lembrando que nosso **efeito** é a diferença de médias (é sobre ela que calcularíamos a estatística t)

Os cálculos envolvidos nesse método podem ser bem complexos, mas felizmente temos programas e páginas que calculam tudo!

Bom para visualização, mas não estima o número de amostras diretamente:

<https://casertamarco.shinyapps.io/power/>

Bem simples e rápido:

<https://www.stat.ubc.ca/~rollin/stats/ssize/n2.html>

Para vários tipos de testes:

<http://powerandsamplesize.com/Calculators/>

Função no R para testes baseados na distribuição t:

`power.t.test`

Pacote ***pwr*** no R tem várias outras funções para vários outros testes

Outros usos, além de calcular tamanho de amostra e poder do teste:

Análise do poder do teste *a posteriori*

Resultados não-significativos (H_0 não foi rejeitada) -> avaliação se o estudo tinha um **poder** suficiente para detectar o **efeito** estimado.

Reforço (para editores) de que a ausência de diferenças é um resultado importante.

Uso reverso do poder do teste.

Identificação do "tamanho mínimo de efeito detectável" (MDES - Minimum detectable effect size)

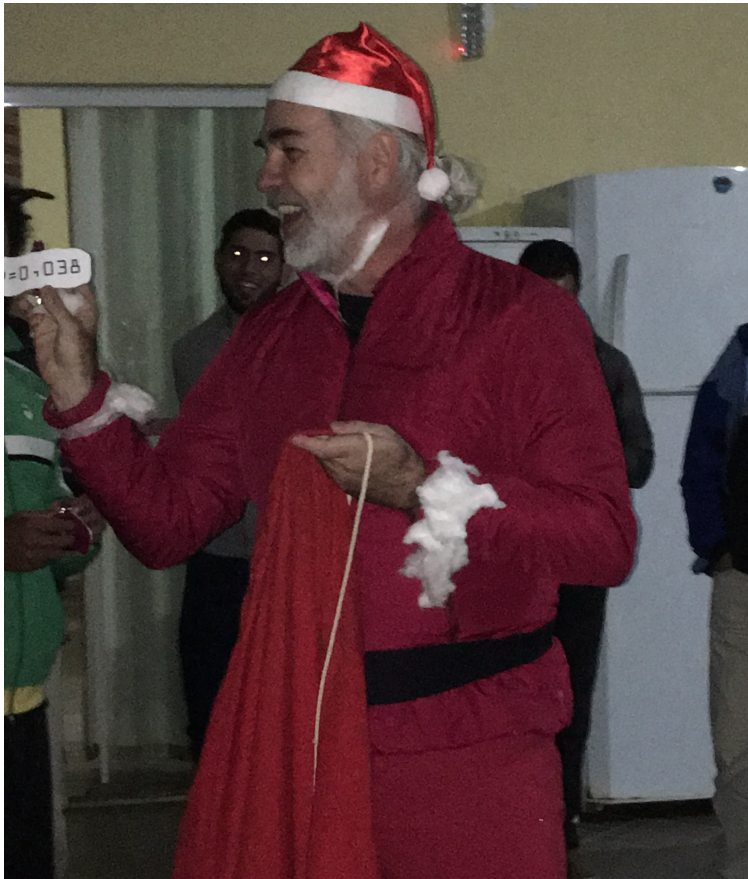
Se existe restrição ao tamanho de amostras -> usa os valores conhecidos para estimar esse **tamanho mínimo de efeito**

O ideal é plotar uma **curva de efeitos mínimos** em função do tamanho de amostra

Em geral, a recomendação principal é : **Cuidado com amostras pequenas!**

Porém, **amostras grandes demais** também podem ser problemáticas!

Se o seu sonho é obter um $p < 0.05$, seus problemas acabaram...

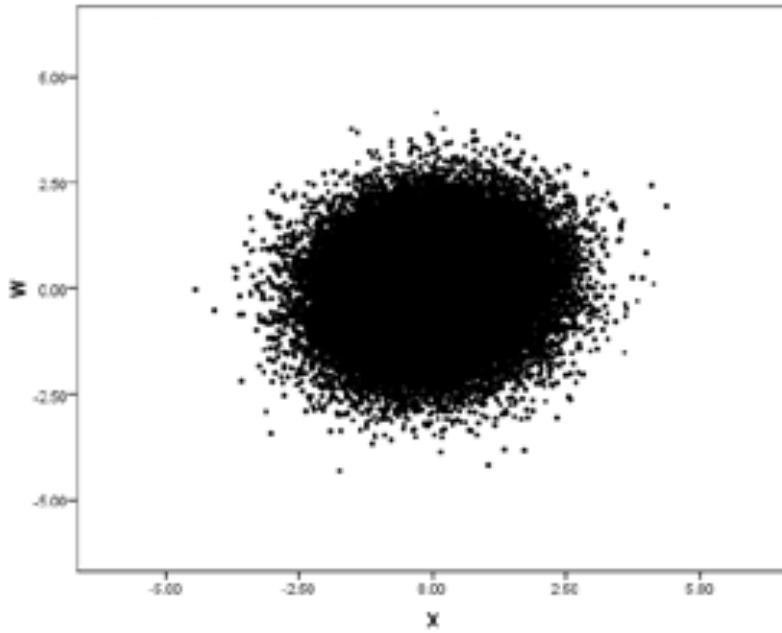


Trabalhe com um $n \gg$ que 300

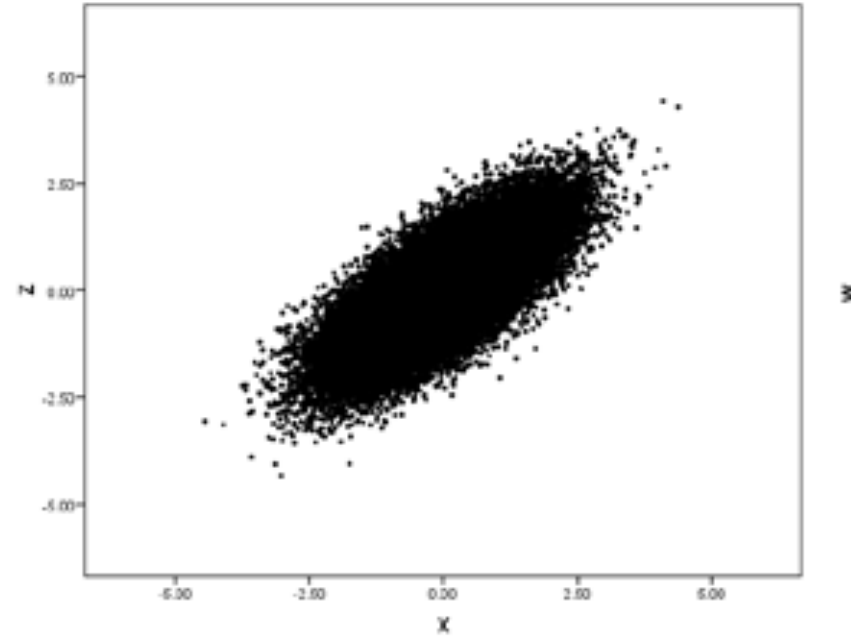
Suas chances aumentam muito!!!

Estudo realizado com simulações

bpsr ■ Dalson Britto Figueiredo Filho, Ranulfo Paranhos
(2013) Enivaldo C. da Rocha, Mariana Batista
José Alexandre da Silva Jr., Manoel L. Wanderley D. Santos
and Jacira Guiro Marino

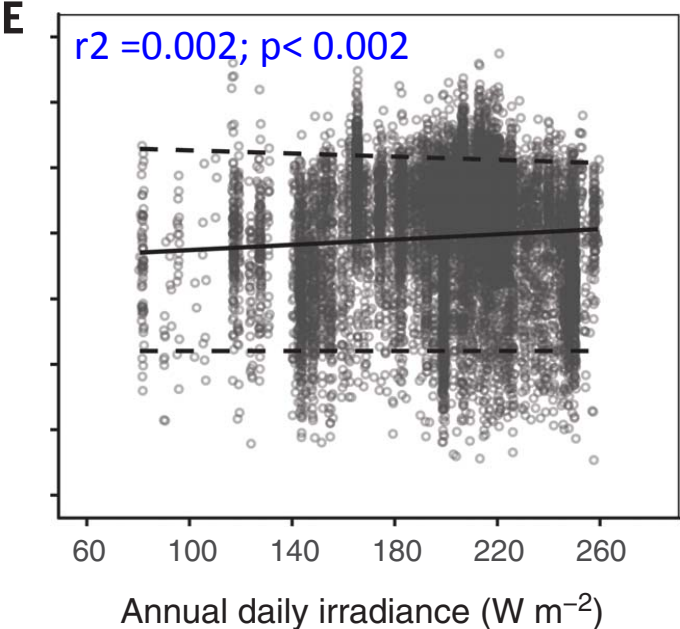
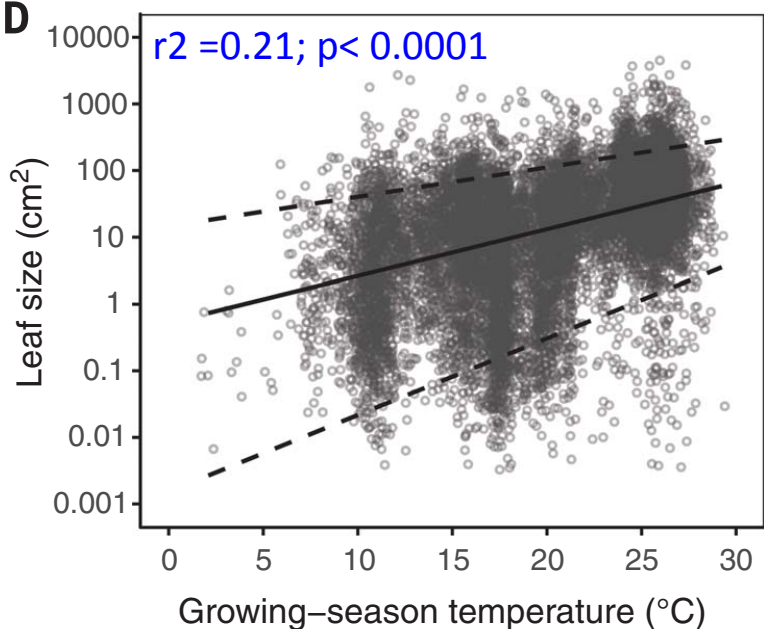
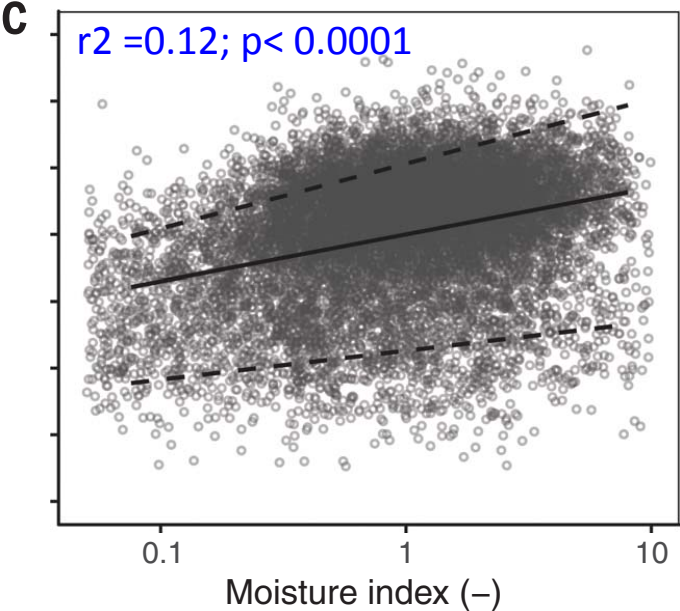
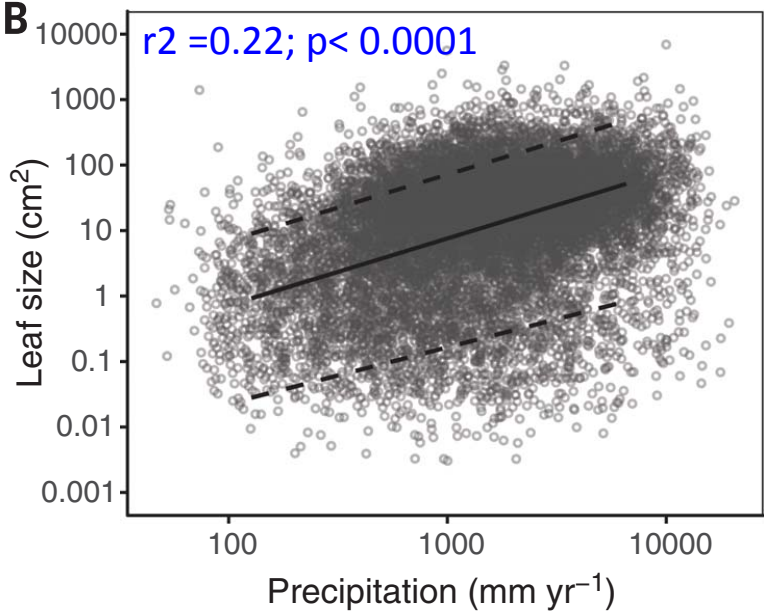


Existe uma relação
entre X e W?



Existe uma relação
entre X e Z?

Estudo com dados reais (7670 espécies)



Recomendação atual:

No caso de "bigdatas" os valores de tamanho de efeito (diferenças de médias, inclinações, ...) e os valores de variação explicada (r , r^2 , ...) devem ser estabelecidos **antes** de realizar as análises.

