

Análise do poder estatístico 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 do Poder ou Teste do Poder**

Análise do poder ou Teste do poder (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
- Populações reais sempre vão diferir em alguma quantidade, mas precisamos definir se essa quantidade importa ou não.

Como definir o que é um efeito relevante?

- Experiência do grupo de pesquisa
- Estudos anteriores

Vamos considerar que no nosso exemplo das sementes o número médio de sementes por planta, considerando toda a região é **55**.

-> 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**?

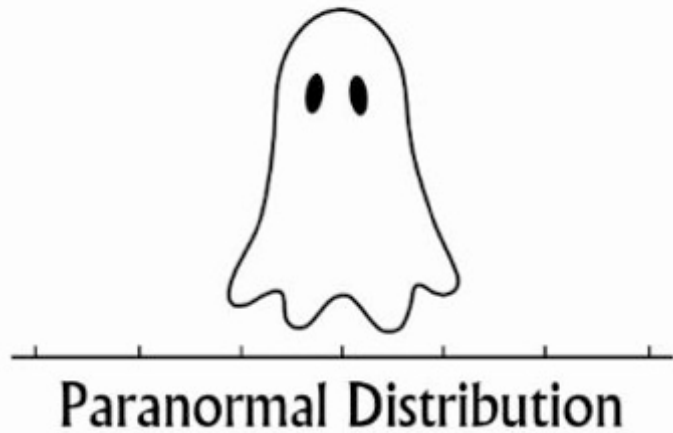
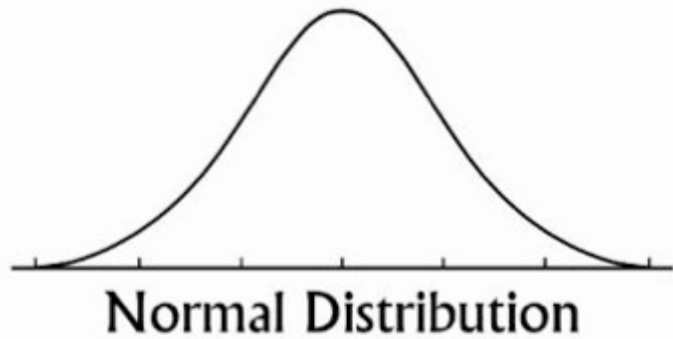
-> E se a diferença for **10**?

Por que essa pergunta?

Quanto maior a precisão desejada (ou seja, quanto menor a diferença), maior o número de amostras necessário para detectar essa diferença.

De onde vem essa estimativa?

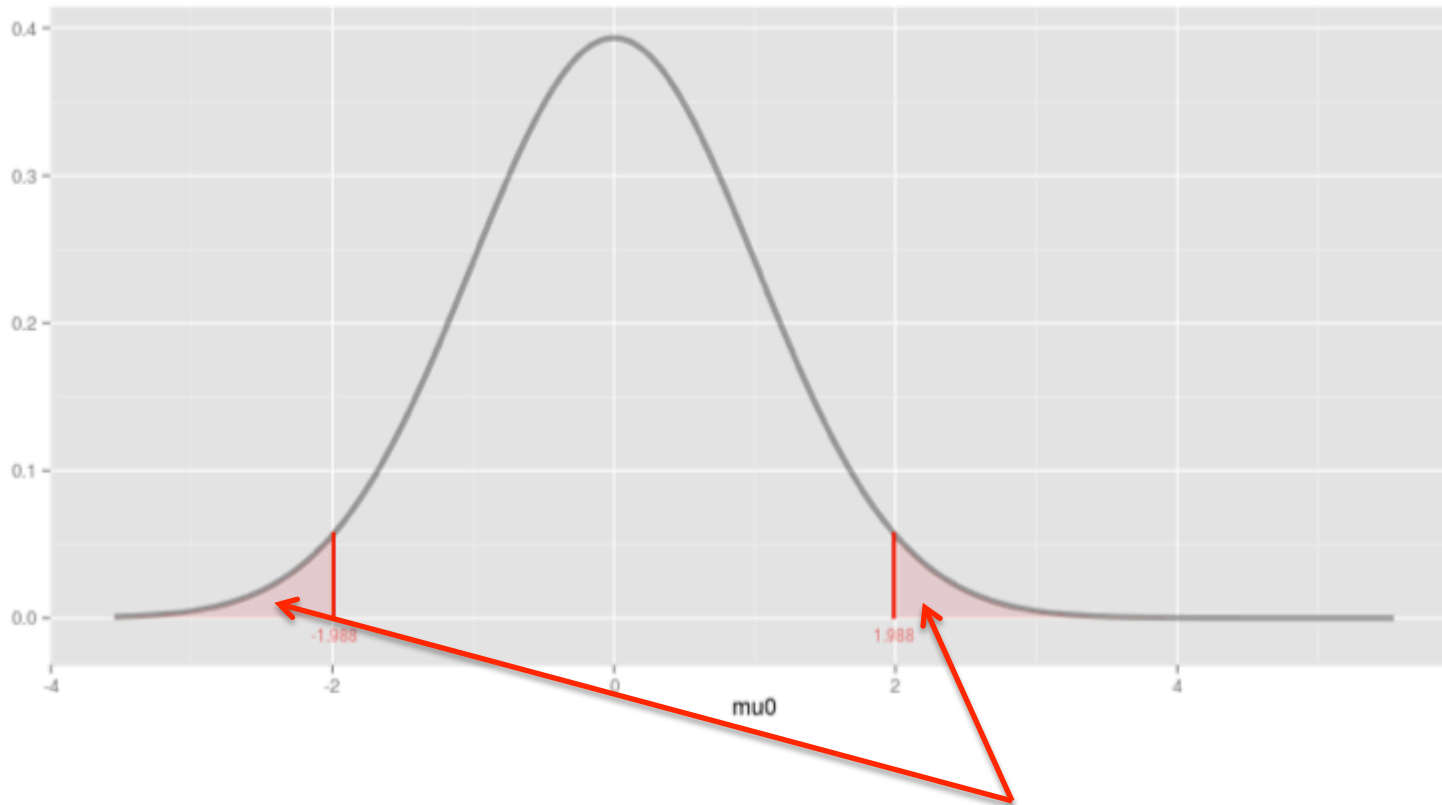
Voltando lá nas Distribuições de Probabilidades...



e Tipos de Erros...



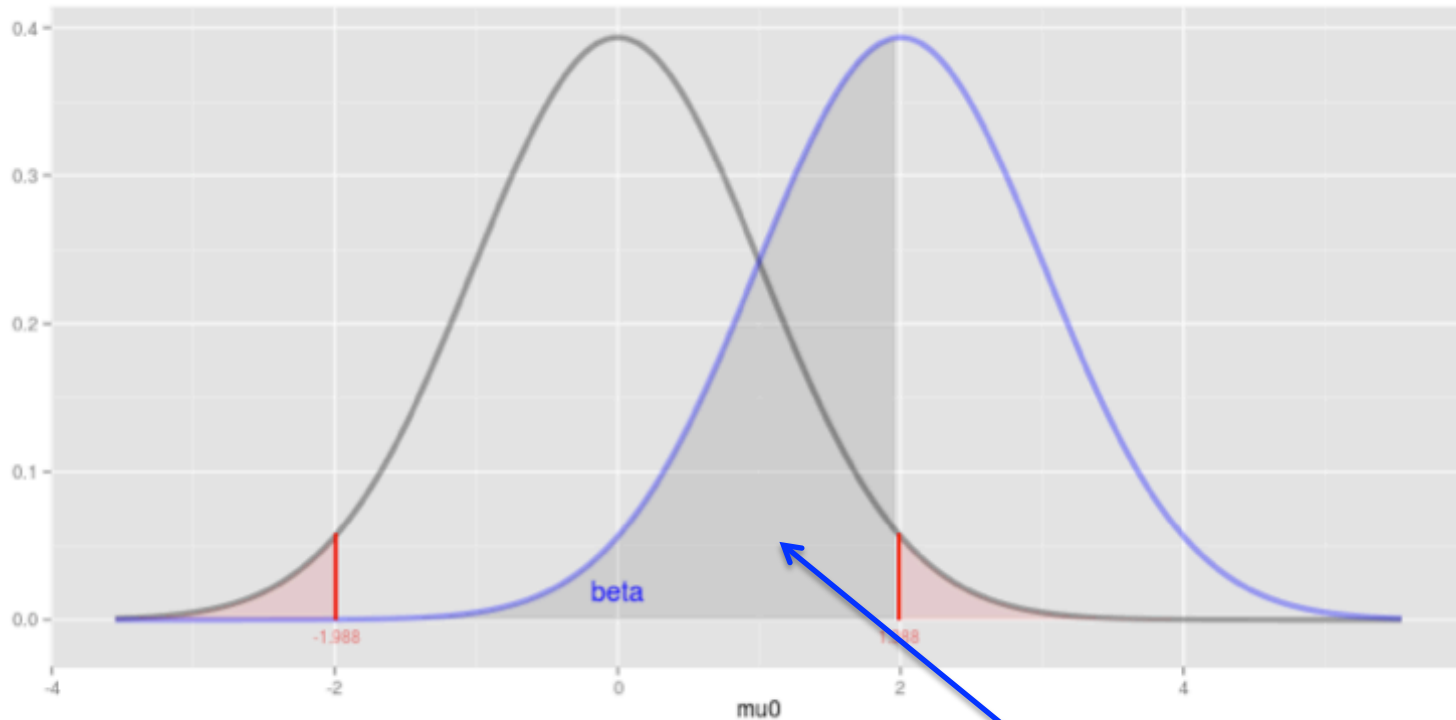
Considere uma distribuição de probabilidades para uma **Hipótese Nula** (centrada em zero*), com desvio padrão = 6



Erro Tipo I = alfa
(Rejeitar H_0 quando ela é verdadeira)

* por exemplo, em relação à diferença entre duas médias

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

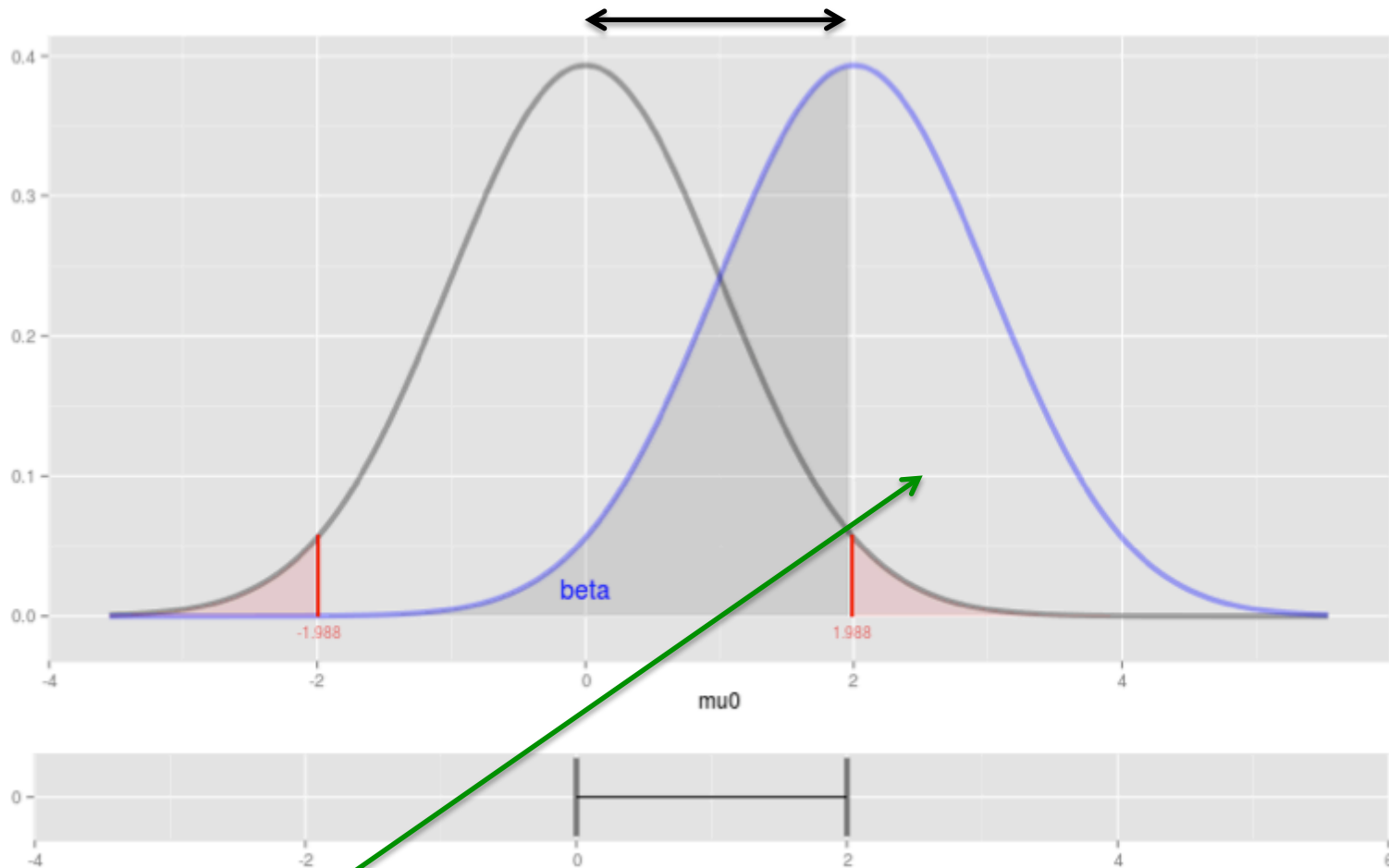


Em cinza escuro, a área ocupada por beta define o **Erro Tipo II (beta)**

Beta é a probabilidade de cometer o **Erro Tipo II - Aceitar H_0 quando ela não é verdadeira.**

Para essa H1, beta é aproximadamente 0,48
Isso representa uma chance de 48% de aceitarmos H_0 erroneamente

O **PODER ESTATÍSTICO** é dado por $[1 - \text{Beta}]$



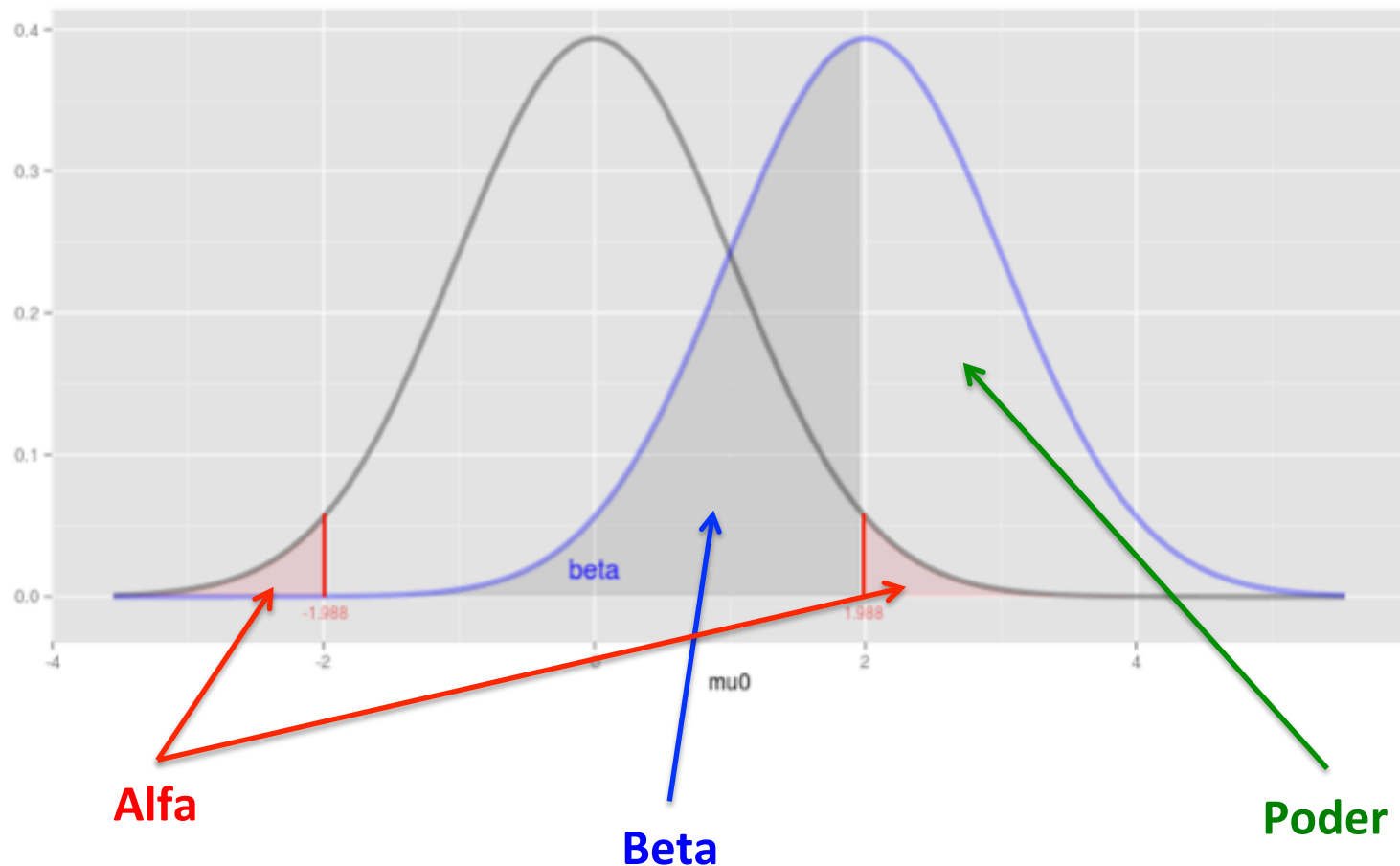
O **poder estatístico** é a probabilidade representada pelo restante da área abaixo da curva azul, **excluindo-se a área do Erro Tipo II (ou beta)**

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



Poder estatístico

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

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

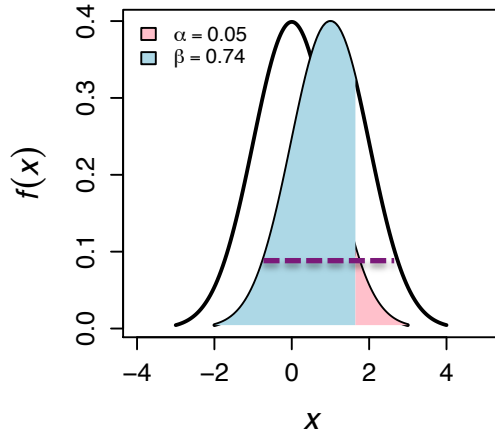
Correlação

Etc...

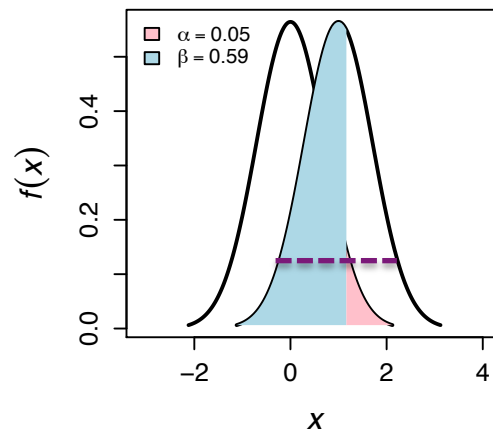
Mas a direção das relações entre essas medidas são as mesmas para qualquer teste estatístico

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

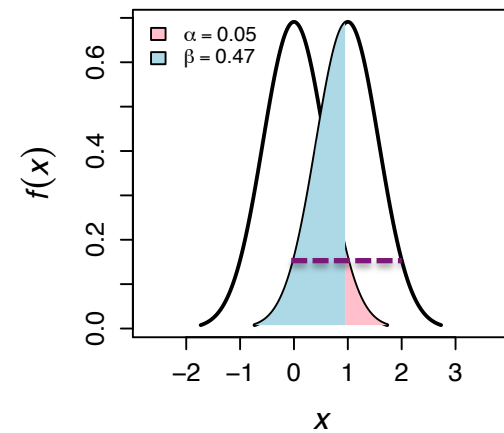
$1 - \beta = 0.26$



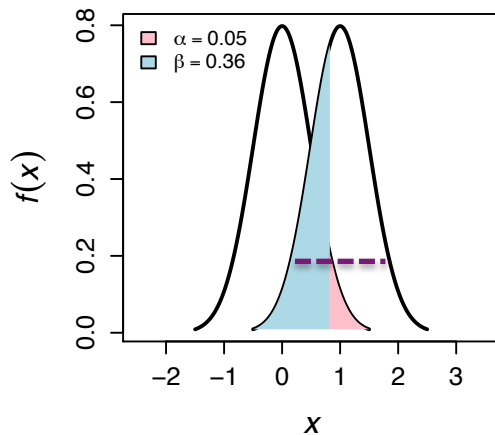
$1 - \beta = 0.41$



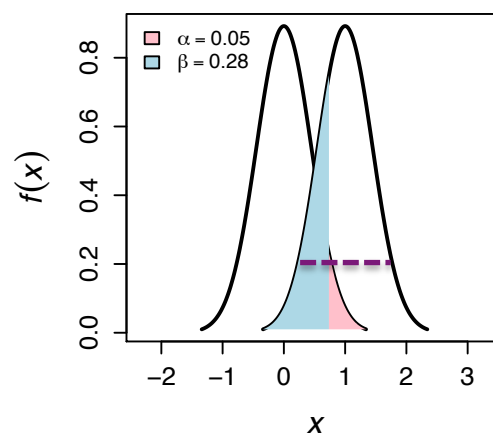
$1 - \beta = 0.53$



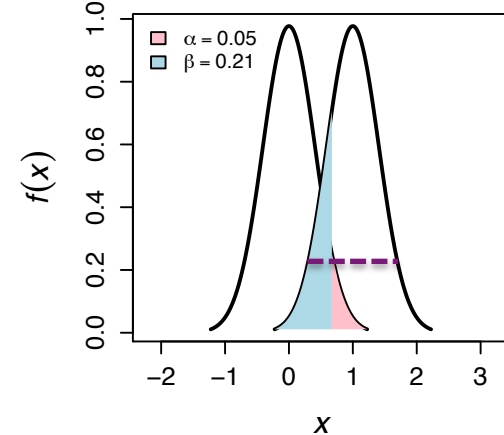
$1 - \beta = 0.64$



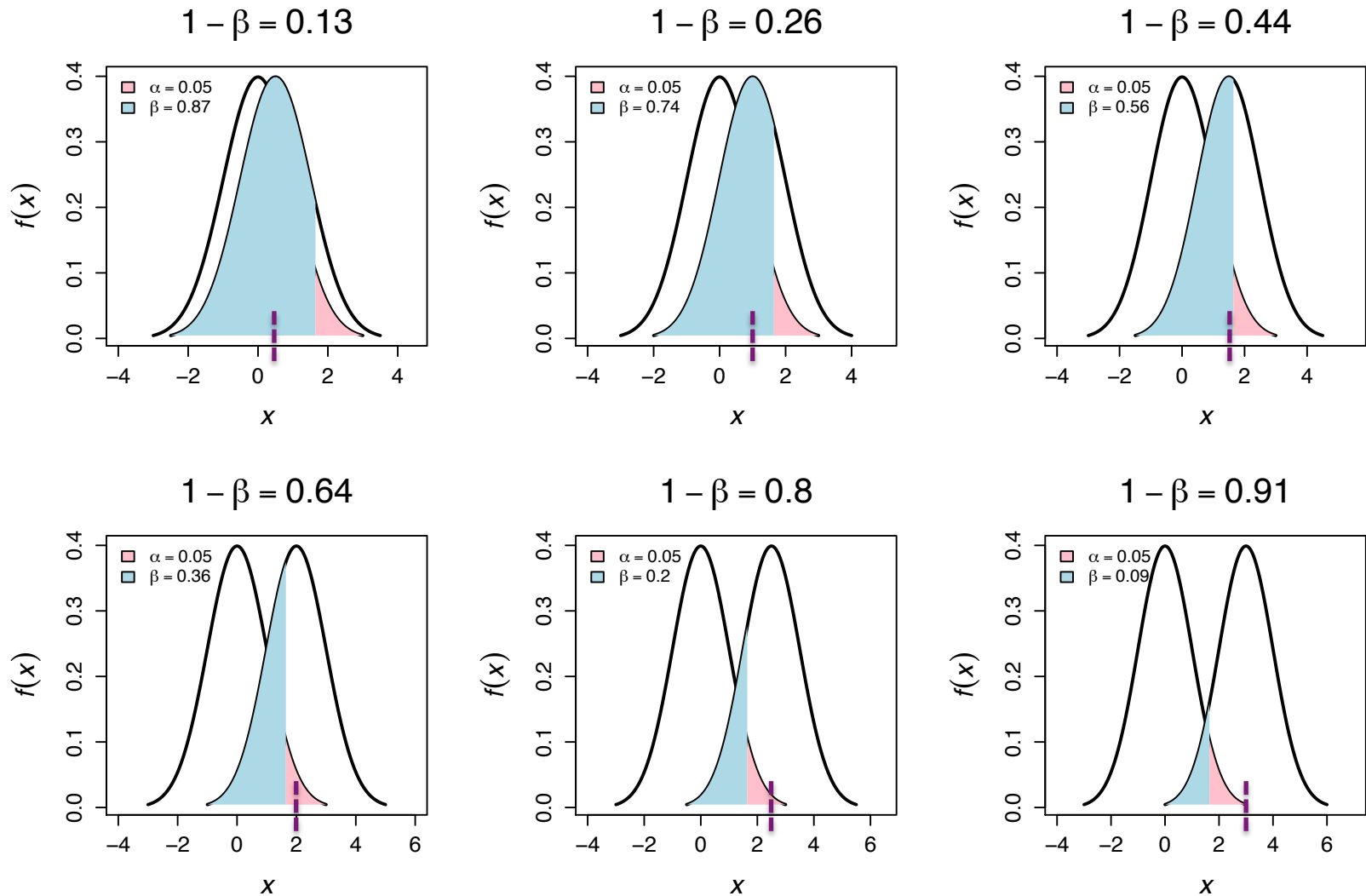
$1 - \beta = 0.72$



$1 - \beta = 0.79$



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

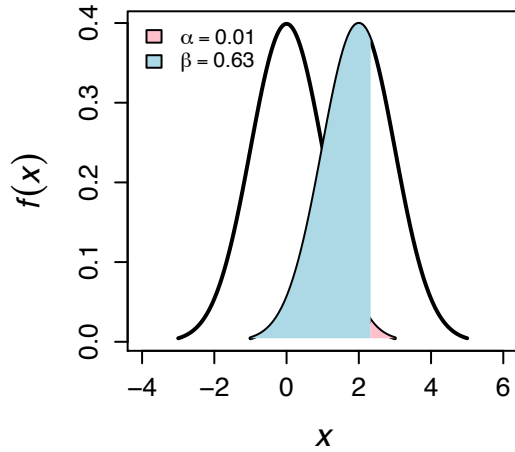


Nathaniel Helwig (2017)

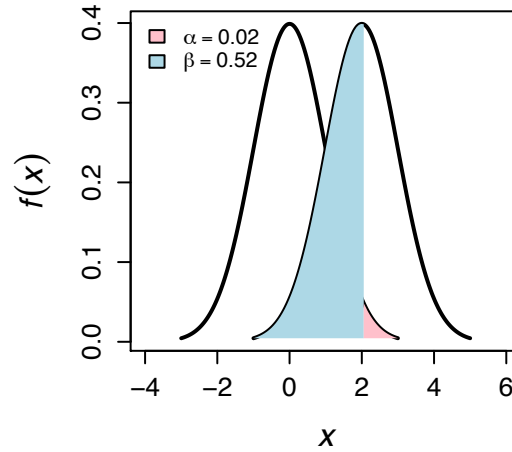
No nosso exemplo, o efeito seria a diferença de médias (em uma distribuição t)

Quanto maior o valor de alfa, maior o Poder (1 - Beta)

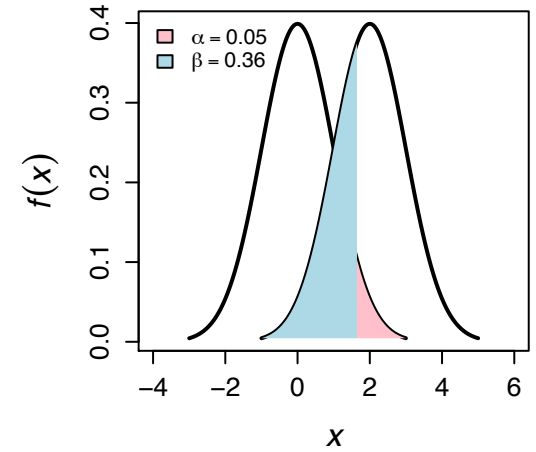
$1 - \beta = 0.37$



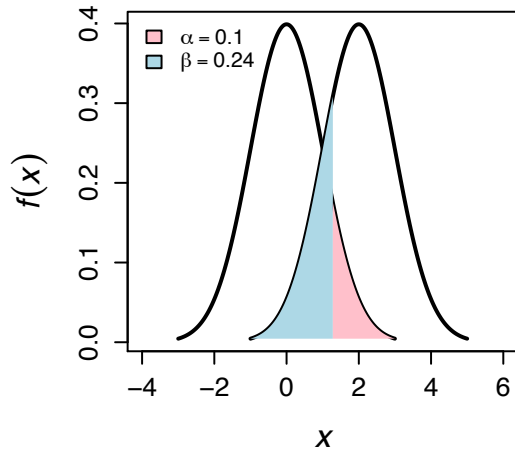
$1 - \beta = 0.48$



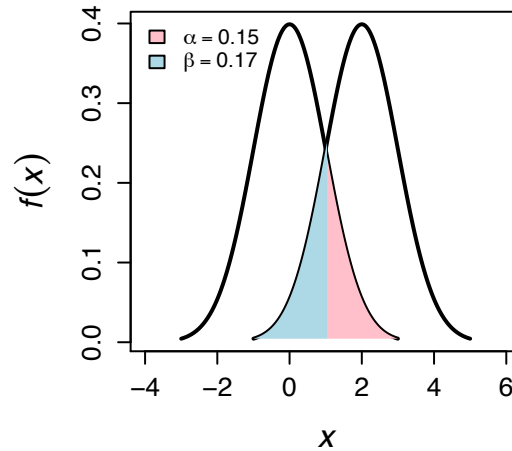
$1 - \beta = 0.64$



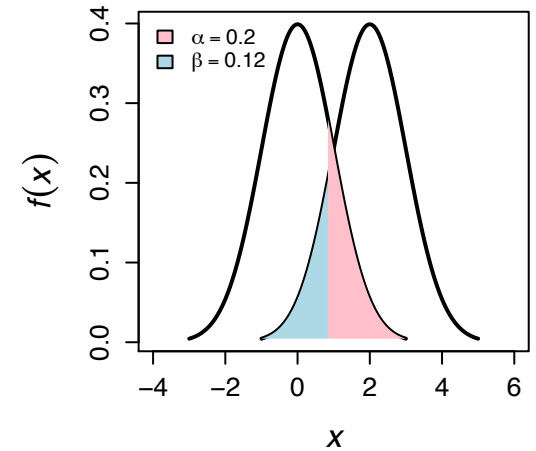
$1 - \beta = 0.76$



$1 - \beta = 0.83$



$1 - \beta = 0.88$

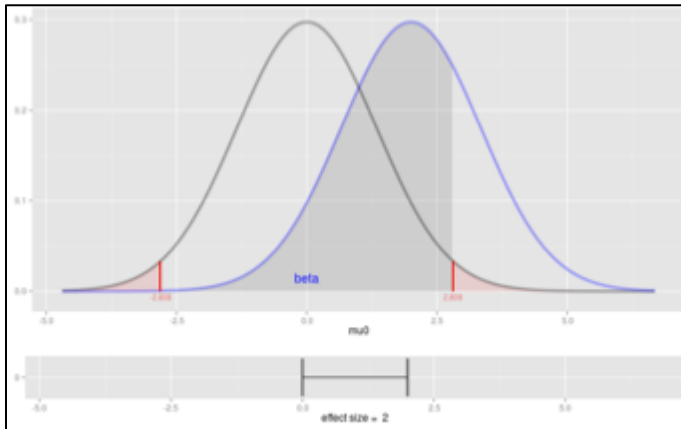


Nathaniel Helwig (2017)

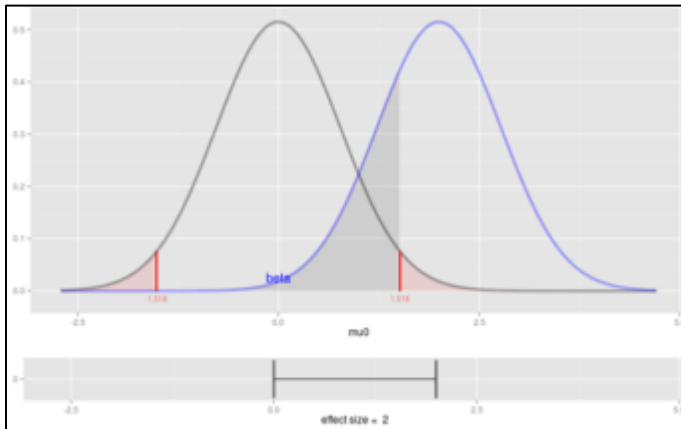
alfa e Beta são negativamente relacionados

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

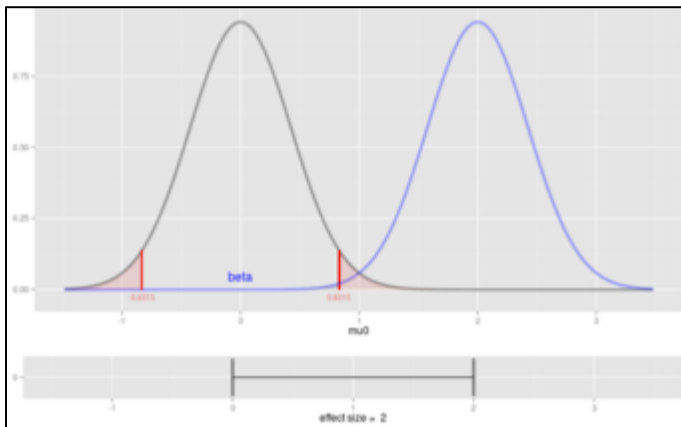
Quanto maior o tamanho da amostra (**n**), maior o Poder (1-Beta)



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)

Outros usos, além de calcular tamanho de amostra e poder estatístico:

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.

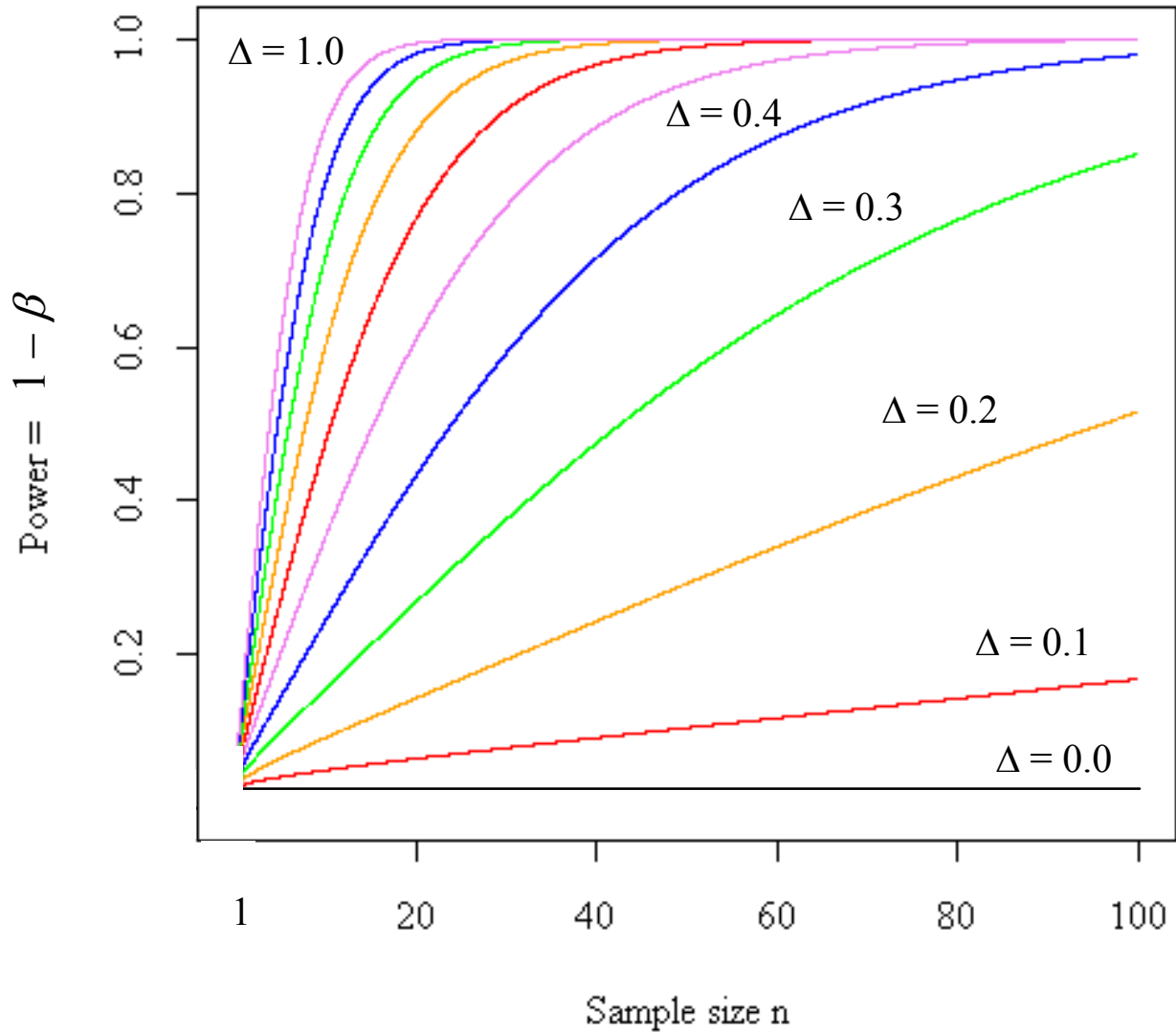
Melhora a discussão dos resultados, levando em consideração a quantidade e a qualidade dos dados e também as propriedades essenciais do sistema de estudo (BB)

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**

Uma abordagem interessante é plotar uma **curva de efeitos mínimos** em função do tamanho da amostra



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

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...

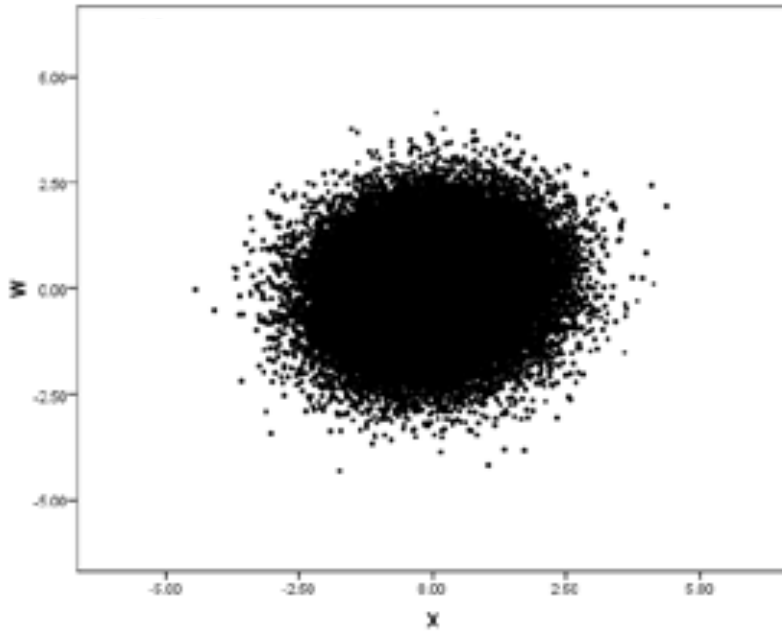


Trabalhe com um $n \gg$ que 300

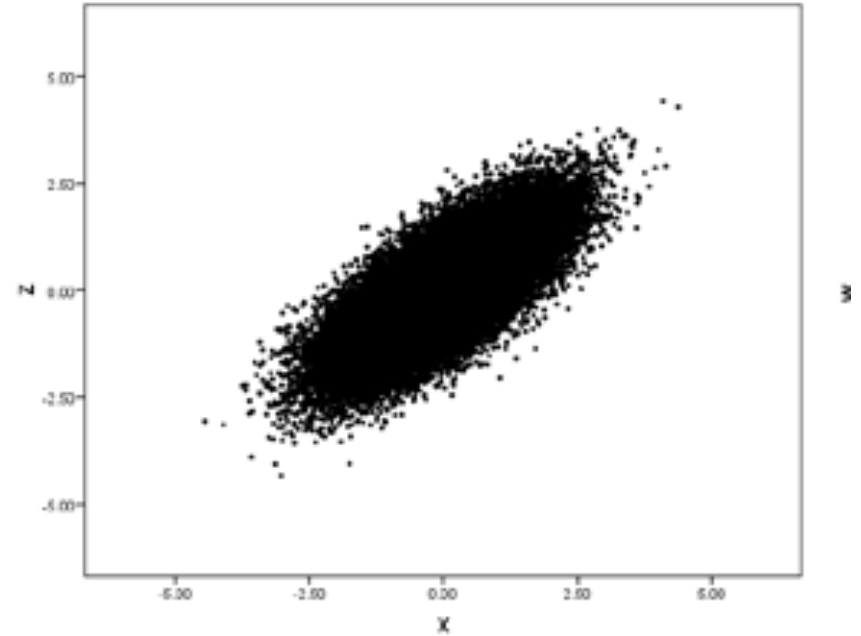
Suas chances aumentam muito!!!

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

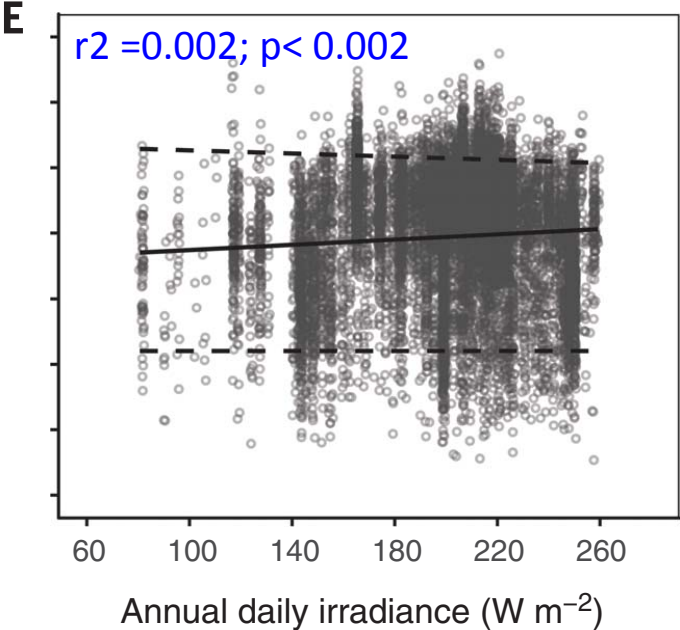
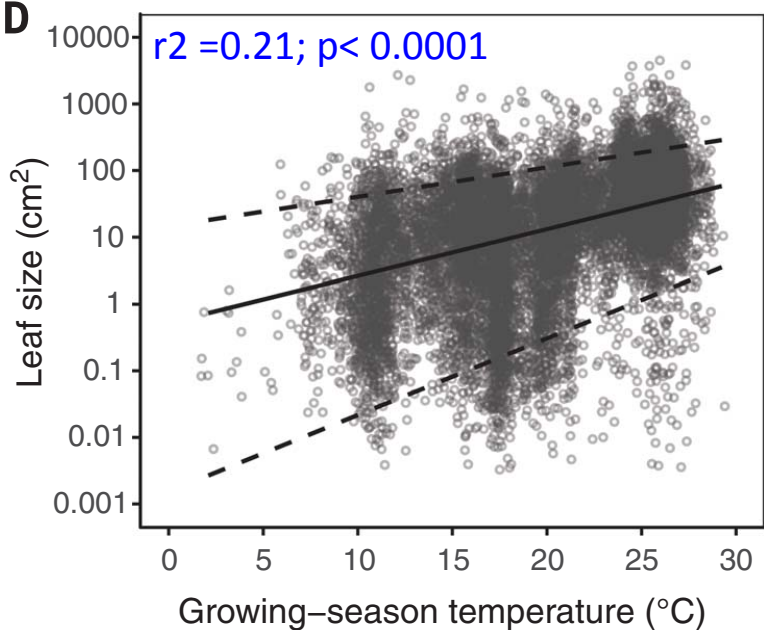
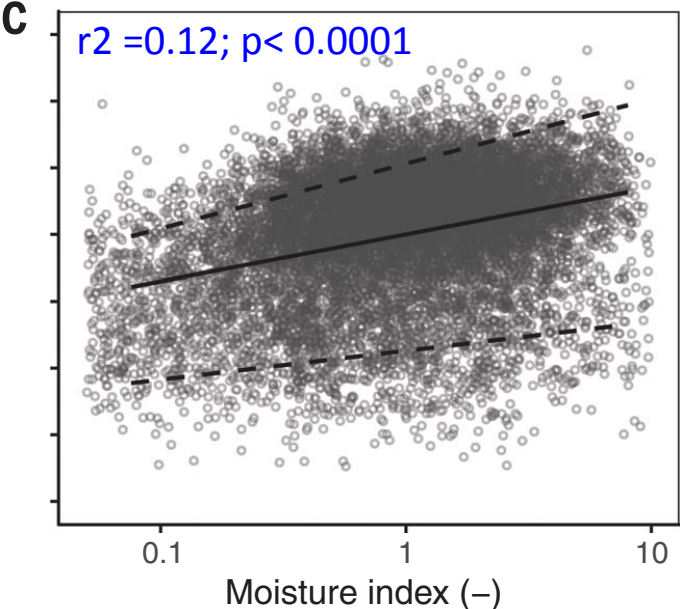
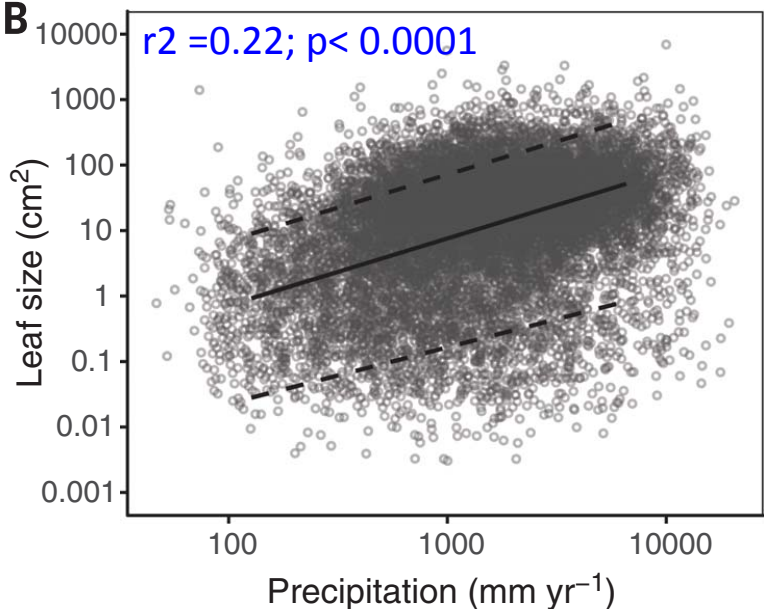


Existe uma relação
entre X e W?



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.

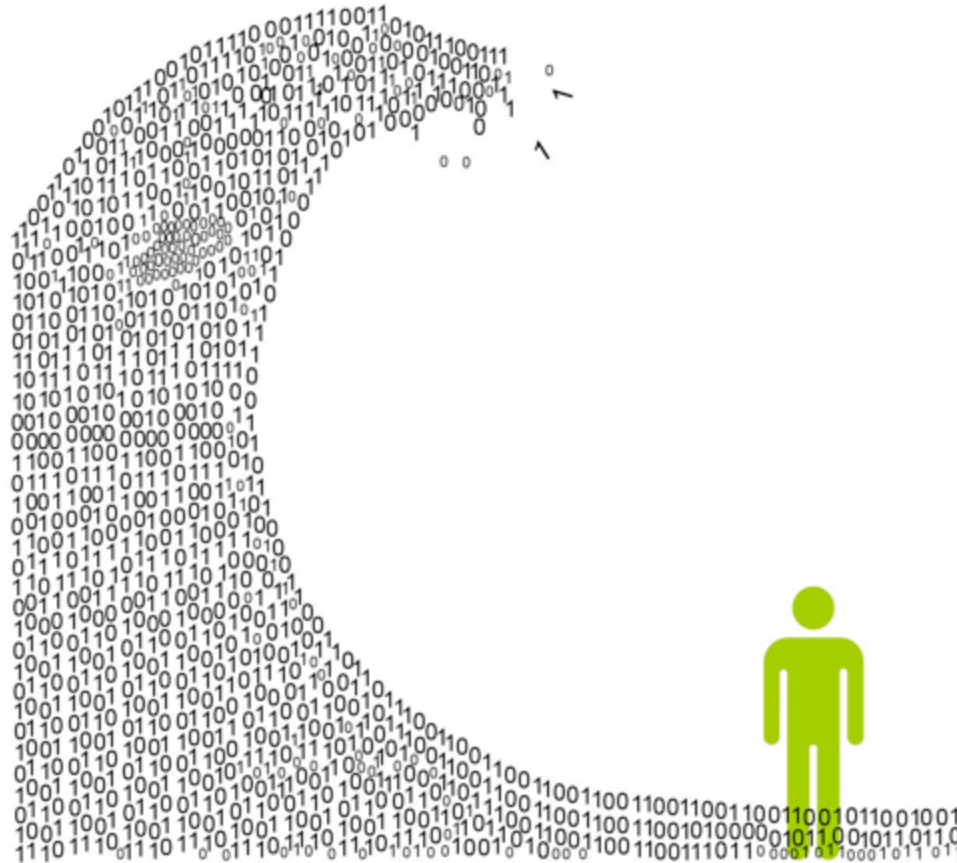


Table of Sample Sizes* for Two-Sided Tests ($\alpha = .05$)

Δ	Power				
	80%	85%	90%	95%	99%
0.1	785	898	1051	1300	1838
0.125	503	575	673	832	1176
0.15	349	400	467	578	817
0.175	257	294	344	425	600
0.2	197	225	263	325	460
0.25	126	144	169	208	294
0.3	88	100	117	145	205
0.35	65	74	86	107	150
0.4	50	57	66	82	115
0.45	39	45	52	65	91
0.5	32	36	43	52	74
0.6	24	27	30	37	52
0.7	19	21	24	29	38
0.8	15	17	19	23	31
0.9	12	14	15	19	25
1.0	10	11	13	15	21

* Shaded cells indicate that 2 was added to compensate for small n .