

Universidade Federal de Goiás

Bioestatística

Prof. Adriano Sanches Melo - Dep. Ecologia – ICB
asm.adrimelo@gmail.com

Página do curso: <http://www.ecologia.ufrgs.br/~adrimelo/bioestat>

Mais um probleminha...

1. Você está num quarto escuro e quer pegar um par de meias;
2. Na gaveta existem 3 meias idênticas exceto pela cor. Duas azuis e uma vermelha;
3. Ao pegar duas meias, a probabilidade de pegar 2 meias da mesma cor (azul) é maior, semelhante ou menor do que pegar cores diferentes?

RESPOSTA →

Teste de hipóteses

Um pouco sobre o acaso:

Provérbios 16,33: “A sorte é lançada ao colo; mas toda a decisão a partir daí cabe ao Senhor”

→ não é bem o que acreditamos em Ciência...

Provérbios 18,18: “O sorteio provoca o cessar da disputa e decide entre os poderosos”

→ ainda fazemos muito uso disto!

O que é um resultado raro?

Podemos avaliar uma determinada probabilidade
com apenas 1 evento?

Quando sei que um dado resultado é raro?

Podemos calcular a probabilidade daquele resultado acontecer

Exemplo:

Qual probabilidade de sortear 5 caras em sequência?

$$\frac{1}{2} * \frac{1}{2} * \frac{1}{2} * \frac{1}{2} * \frac{1}{2} = \frac{1}{32} = 0,0312 = 3,12\%$$

Podemos pensar em alguma outra forma de determinar a probabilidade, mesmo que tenhamos apenas uma estimativa?

1. Jogue uma moeda 5 vezes e anote quantas caras saíram;
2. Repita o passo acima 1000 vezes e em cada um anote o n° de caras
3. O que você espera como resultado?

[1] 3 0 2 3 2 2 3 2 2 2 4 1 1 3 1 2 5 3 5 3 2 2 2 2 1 1 2 2 1 2 5 3 4 1 3 4 2
[38] 1 0 1 2 1 3 1 1 2 2 3 3 2 3 2 3 3 2 1 3 1 3 2 4 2 3 3 3 3 5 3 3 2 2 0
[75] 4 3 2 2 2 1 3 1 1 3 3 3 2 2 3 4 4 2 3 3 3 2 4 4 2 2 5 2 2 3 3 2 3 2 2 0 4
[112] 2 3 1 2 2 1 4 4 2 3 5 1 5 2 4 3 2 2 3 3 2 2 1 2 5 4 3 2 3 3 4 3 2 3 1 4 2
[149] 2 2 2 5 1 5 1 1 1 3 1 3 3 3 3 3 4 2 4 2 3 2 4 5 3 2 2 2 1 4 1 2 1 4 2 4 3
[186] 3 1 0 4 3 4 3 4 5 2 3 3 3 5 2 2 2 2 2 3 3 0 4 3 2 2 4 2 3 3 2 5 0 1 2 4 2
[223] 4 2 3 4 2 1 1 3 0 1 2 4 3 3 3 1 2 2 3 4 5 2 2 4 2 3 2 3 3 3 1 2 0 3 2 2 3
[260] 2 3 4 2 1 3 4 3 2 2 3 3 1 2 1 3 4 3 0 3 3 3 3 1 2 2 2 4 1 1 3 1 4 1 3 2 2
[297] 3 3 2 4 4 2 2 3 3 3 2 2 2 3 5 4 2 3 3 1 3 4 4 3 4 3 4 3 3 2 3 2 2 3 4 3 1
[334] 4 3 3 2 2 3 3 3 3 3 3 3 1 4 1 2 4 3 3 4 2 2 3 2 3 3 1 4 3 1 1 3 3 3 2 3 2
[371] 2 4 3 0 0 2 4 2 2 4 2 3 2 1 1 1 2 3 2 3 2 3 3 3 2 2 2 4 2 4 0 2 2 1 4 2 2
[408] 1 2 2 3 3 2 2 3 2 2 2 3 3 4 3 2 3 2 3 4 3 0 4 2 2 1 3 1 3 4 4 1 2 2 3 3 2
[445] 1 2 3 3 3 2 2 3 2 5 1 3 5 1 2 4 2 3 1 4 2 3 0 1 0 0 2 3 4 3 2 2 2 2 2 2 3
[482] 2 4 2 2 4 2 1 2 2 2 2 1 4 3 2 3 3 3 4 3 3 3 2 2 2 1 3 5 2 3 2 2 1 1 2 3 3
[519] 3 2 2 2 3 2 4 2 3 4 4 1 5 3 2 1 3 2 3 4 1 1 2 2 3 0 2 2 2 3 2 2 2 1 4 3 1
[556] 1 2 2 3 1 2 5 2 2 3 2 1 2 3 2 2 3 2 2 2 2 3 3 1 1 3 0 3 1 3 2 1 3 2 2 2 2
[593] 3 3 2 4 4 4 3 3 2 4 3 2 1 3 2 0 4 2 4 3 3 2 2 3 2 1 1 2 1 3 4 1 2 2 2 2 5
[630] 1 2 3 3 1 3 1 3 4 2 3 3 1 1 5 1 5 4 3 3 4 2 1 2 2 2 4 3 2 1 4 1 3 2 2 1 3
[667] 2 2 1 2 1 1 2 2 3 3 1 4 4 3 3 1 1 4 2 2 5 3 2 2 1 3 2 4 3 2 0 3 0 2 2 4 2
[704] 1 3 1 2 2 3 4 3 0 2 4 4 2 2 3 3 1 3 4 4 2 2 5 4 2 4 3 3 3 2 3 2 4 2 3 2 4
[741] 1 2 2 3 1 2 2 5 4 4 3 2 1 0 2 4 2 1 4 4 1 4 2 3 4 4 2 2 1 3 3 1 3 3 1 3 4
[778] 3 1 4 1 0 3 5 4 3 4 1 1 3 1 5 0 2 1 3 3 4 2 2 1 2 1 2 2 3 4 2 3 3 1 2 3 2
[815] 2 4 2 3 2 2 3 1 1 3 3 3 3 4 4 4 3 3 3 2 2 2 3 3 3 1 0 1 4 4 3 2 1 2 3 3 4
[852] 1 3 2 1 2 2 4 3 4 2 3 3 2 5 4 4 2 2 1 1 2 3 3 1 4 3 3 4 4 2 2 3 2 2 3 5 2
[889] 3 1 2 3 3 3 5 1 2 3 2 3 3 4 2 3 3 2 3 2 1 2 1 4 4 2 2 2 3 3 3 2 1 2 3 4 1
[926] 3 2 2 0 3 2 1 4 2 1 4 1 4 0 2 1 2 3 2 2 3 3 2 3 4 4 3 1 5 2 1 3 2 3 3 4 1
[963] 4 4 5 1 3 3 4 2 4 4 2 3 2 3 2 2 5 1 3 2 1 2 2 2 4 2 2 3 3 2 2 3 1 0 3 4 3
[1000] 3

Tabulação

0	1	2	3	4	5
30	152	335	303	145	35

Probabilidade $\sim 35/1000 = 0,035 = 3,5\%$

SE eu repetir o sorteio mais 1000 vezes, o resultado é o mesmo?

[1] 3 3 5 1 4 1 3 5 4 4 1 4 5 3 1 5 1 3 2 3 4 5 1 4 2 4 2 1 2 4 4 3 1 3 1 3 1
[38] 2 1 1 4 1 4 0 1 4 3 2 3 3 2 3 1 2 1 5 3 2 4 1 2 1 3 4 3 4 1 3 3 2 3 2 3 2
[75] 2 2 4 3 1 4 2 5 2 1 4 1 2 4 2 2 3 3 4 0 4 2 2 4 3 3 5 2 3 3 1 1 1 4 0 3 3
[112] 4 4 1 3 1 5 4 4 3 3 5 4 3 3 3 2 3 4 1 3 1 1 4 2 3 2 3 2 2 2 2 1 4 3 4 2 2
[149] 3 3 5 1 4 2 2 3 1 3 4 2 1 2 4 2 1 4 4 3 1 1 3 3 1 1 2 1 5 3 2 2 2 3 1 3 4
[186] 2 0 3 3 3 3 3 2 1 3 1 3 2 4 2 0 3 3 1 1 2 1 0 1 2 4 2 3 2 5 5 1 2 4 2 3 2
[223] 4 2 3 2 3 1 3 2 2 4 3 1 2 4 2 0 3 3 4 1 4 3 2 2 3 3 2 2 3 1 4 3 2 2 4 4 2
[260] 2 0 2 3 2 4 1 1 3 3 1 2 2 3 3 2 0 3 2 1 3 4 3 4 2 2 2 3 5 1 2 3 3 3 2 1 2
[297] 2 3 3 2 1 1 3 3 1 1 3 1 4 3 4 2 4 3 3 2 1 2 1 2 3 3 3 4 5 2 2 4 3 3 4 2 1
[334] 1 2 0 3 3 3 2 4 2 4 3 4 2 3 4 4 1 3 4 4 4 0 4 3 5 1 1 2 3 3 3 4 1 4 5 2 3
[371] 4 2 3 2 2 2 2 1 3 4 1 3 3 1 3 5 2 5 1 3 3 2 3 3 3 3 3 2 1 5 2 4 2 4 2 3 2
[408] 1 4 2 4 4 3 1 4 2 2 4 2 2 2 2 3 3 3 3 2 3 3 4 4 0 2 3 3 3 0 1 2 2 3 3 2 2
[445] 2 1 2 1 0 3 5 3 1 2 4 3 3 2 1 1 0 2 1 2 3 2 4 0 3 4 4 2 4 4 3 2 2 2 3 1 2
[482] 2 2 3 3 4 2 2 3 4 2 1 2 3 0 3 4 2 2 3 3 1 2 3 4 1 2 4 2 3 1 1 2 3 2 3 3 3
[519] 2 2 4 2 1 3 0 2 2 2 2 4 1 4 2 4 3 0 3 4 2 2 1 4 2 1 4 3 3 0 3 3 1 1 3 3 2
[556] 2 3 3 3 2 3 3 0 1 4 3 3 0 3 4 5 4 0 3 3 4 2 1 1 3 2 2 2 4 0 2 2 2 4 4 2 3
[593] 3 1 2 3 3 3 4 3 5 3 2 3 1 2 2 2 2 3 4 0 3 2 5 3 2 5 3 1 3 2 3 4 4 3 4 2 2
[630] 3 4 4 3 1 3 2 2 4 3 0 3 1 1 5 3 2 1 2 3 1 1 1 3 4 3 3 0 4 2 3 4 2 3 5 2 3
[667] 0 3 1 2 1 1 0 3 4 2 2 4 3 4 4 4 4 2 3 4 1 4 2 1 3 3 3 2 4 2 2 2 4 2 2 3 3
[704] 1 3 2 3 2 2 2 2 3 3 5 1 3 1 0 3 3 3 4 3 4 3 2 2 1 2 1 3 2 2 3 1 2 3 4 4 2
[741] 3 1 3 1 2 4 3 3 3 3 2 3 2 2 5 2 1 2 2 3 1 4 3 3 2 4 3 2 3 2 0 2 3 0 3 3 4
[778] 2 2 3 2 2 3 5 3 3 1 3 2 1 1 3 3 3 1 1 3 3 1 3 2 4 2 2 2 3 5 3 2 0 4 2 1 1
[815] 1 2 3 2 0 3 3 0 3 2 3 3 2 2 2 1 2 2 3 4 2 4 1 3 2 3 3 3 2 3 3 3 3 3 3 4 2
[852] 2 3 1 1 3 5 1 1 3 3 4 2 3 2 1 4 3 2 1 4 3 1 3 2 1 3 2 4 2 3 3 3 4 3 3 5 3
[889] 3 2 2 2 3 1 3 5 2 3 3 3 2 2 2 4 0 3 1 2 2 3 3 0 3 3 2 3 1 2 0 2 2 2 2 2 2
[926] 0 2 2 2 2 1 2 0 3 2 2 3 1 3 3 3 1 1 1 1 2 3 3 1 4 3 2 3 3 4 1 2 2 3 2 2 3
[963] 4 3 3 2 2 3 3 2 1 2 1 3 2 3 2 2 1 2 1 4 2 3 3 2 1 2 2 4 3 2 4 2 3 3 1 4 2
[1000] 5

Tabulação

0	1	2	3	4	5
33	153	307	293	169	45

Probabilidade $\sim 45/1000 = 0,045 = 4,5\%$

Podemos aumentar o número de sorteios!
Com 10.000

Tabulação

0	1	2	3	4	5
320	1585	3109	3137	1538	311

Probabilidade $\sim 311/10000 = 0,031 = 3,11\%$

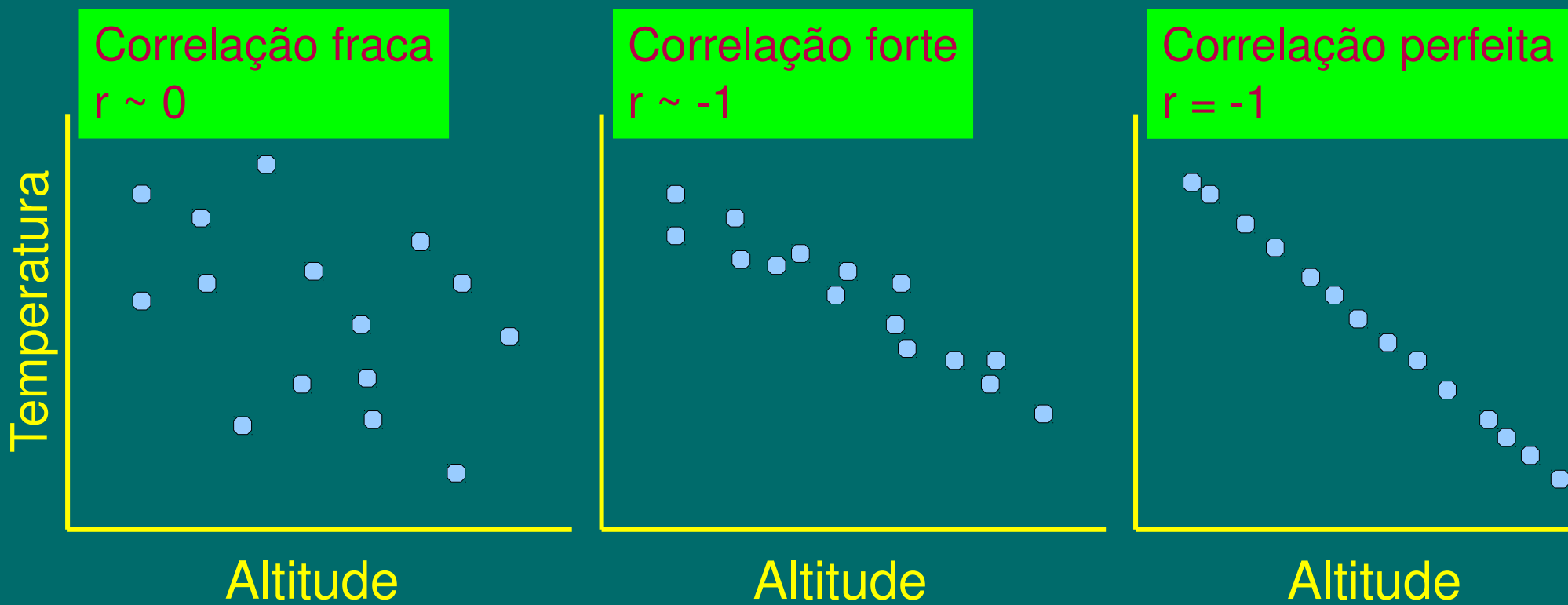
Exercício com Planilha Eletrônica

Exercício com R

Correlação

$$r = \frac{\sum (XY) - \frac{(\sum X)(\sum Y)}{n}}{\sqrt{[\sum X^2 - \frac{(\sum X)^2}{n}][\sum Y^2 - \frac{(\sum Y)^2}{n}]}}$$

Coeficiente de correlação de Pearson
Varia de -1 a 1



O valor de r que obtive com meus dados é alto ou baixo?

Em outras palavras, qual a probabilidade de obter um valor de r tão alto quanto o meu simplesmente devido ao acaso?

Problema:

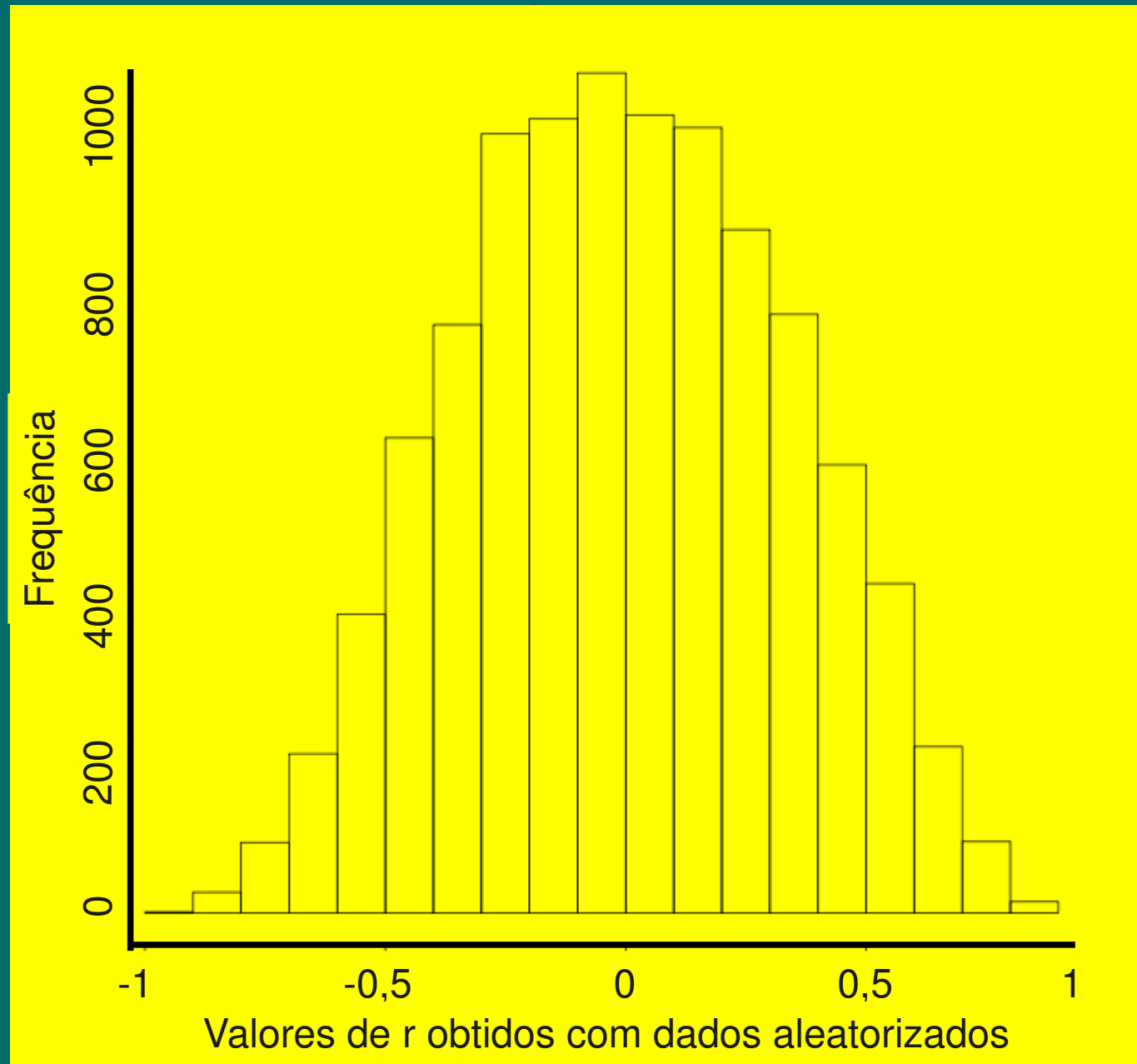
Diferente do sorteio da moeda, aqui existe outro fator que afeta o resultado: tamanho da amostra

Solução: Aleatorizar os dados e calcular o r muitas vezes!

Os valores de r são contínuos e portanto nossa tabulação não é tão eficaz. Podemos fazer um histograma:

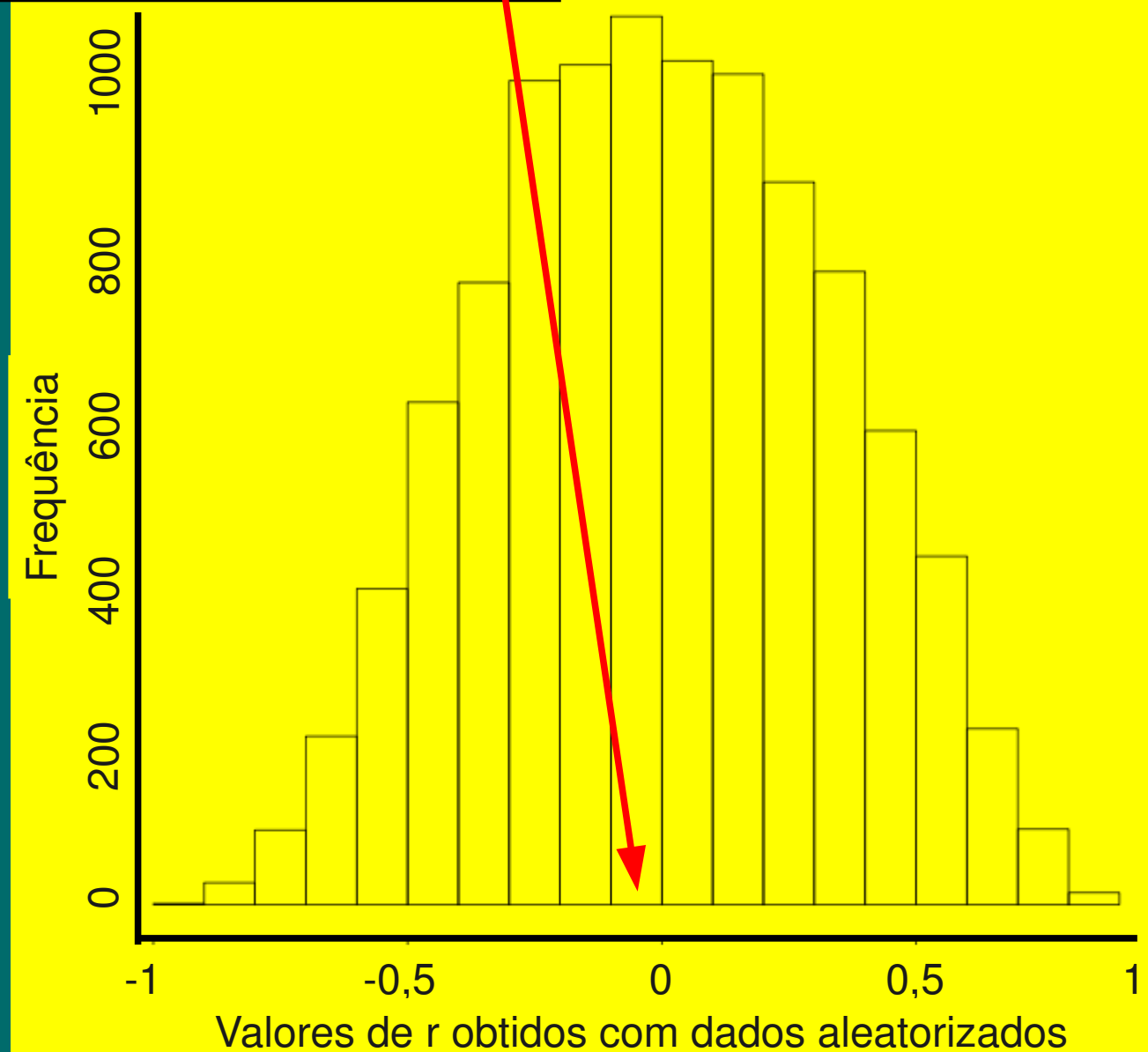
Dados:

X	Y
1	114
2	-2
3	-150
4	463
5	-1124
6	136
7	547
8	-543
9	-156
10	-351



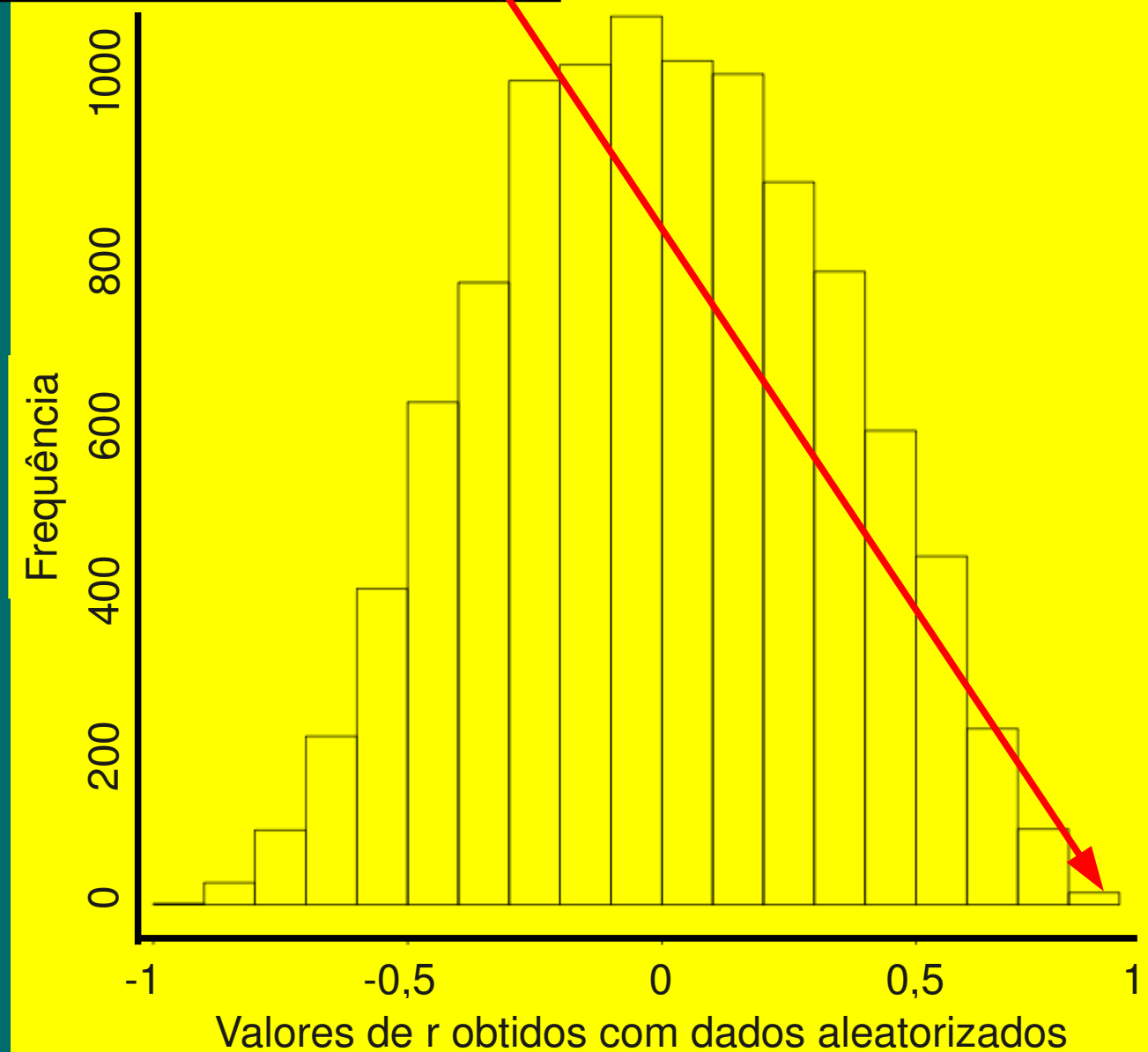
Os valores de r são contínuos e portanto nossa tabulação não é tão eficaz. Podemos fazer um histograma:

O que você conclui se o seu r estiver aqui?



Os valores de r são contínuos e portanto nossa tabulação não é tão eficaz. Podemos fazer um histograma:

O que você conclui se o seu r estiver aqui?



Podemos ser mais diretos e expressar o resultado com uma probabilidade:

$$p = \frac{(\text{frequência com que } |esp| > |obs|) + 1}{\text{número de aleatorizações} + 1}$$

Teste bicaudais e ...

A fórmula anterior é válida para testes bicaudais.

Neste caso, concluimos que deve existir mais do que o acaso quando nosso valor de estatística observada é maior do que o esperado, independente do sinal (módulo).

A priori, não esperamos que a correlação seja positiva ou negativa. Se for extrema, independente se negativa ou positiva, aceitamos o resultado.

Mas nem sempre é assim...

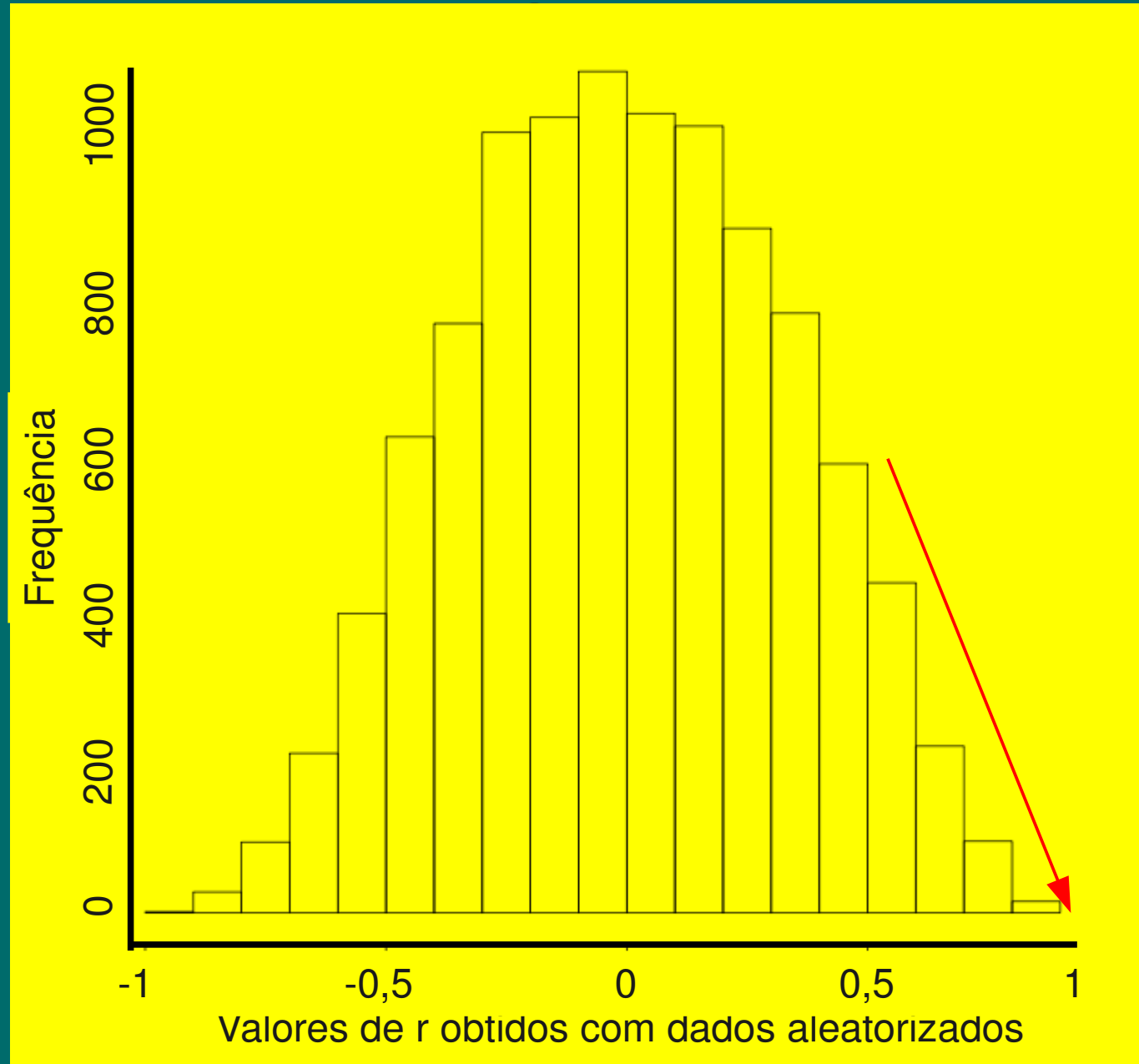
Testes monocaudais

Em algumas situações temos hipóteses mais específicas.

Quero saber se existe correlação positiva entre biomassa e área foliar.

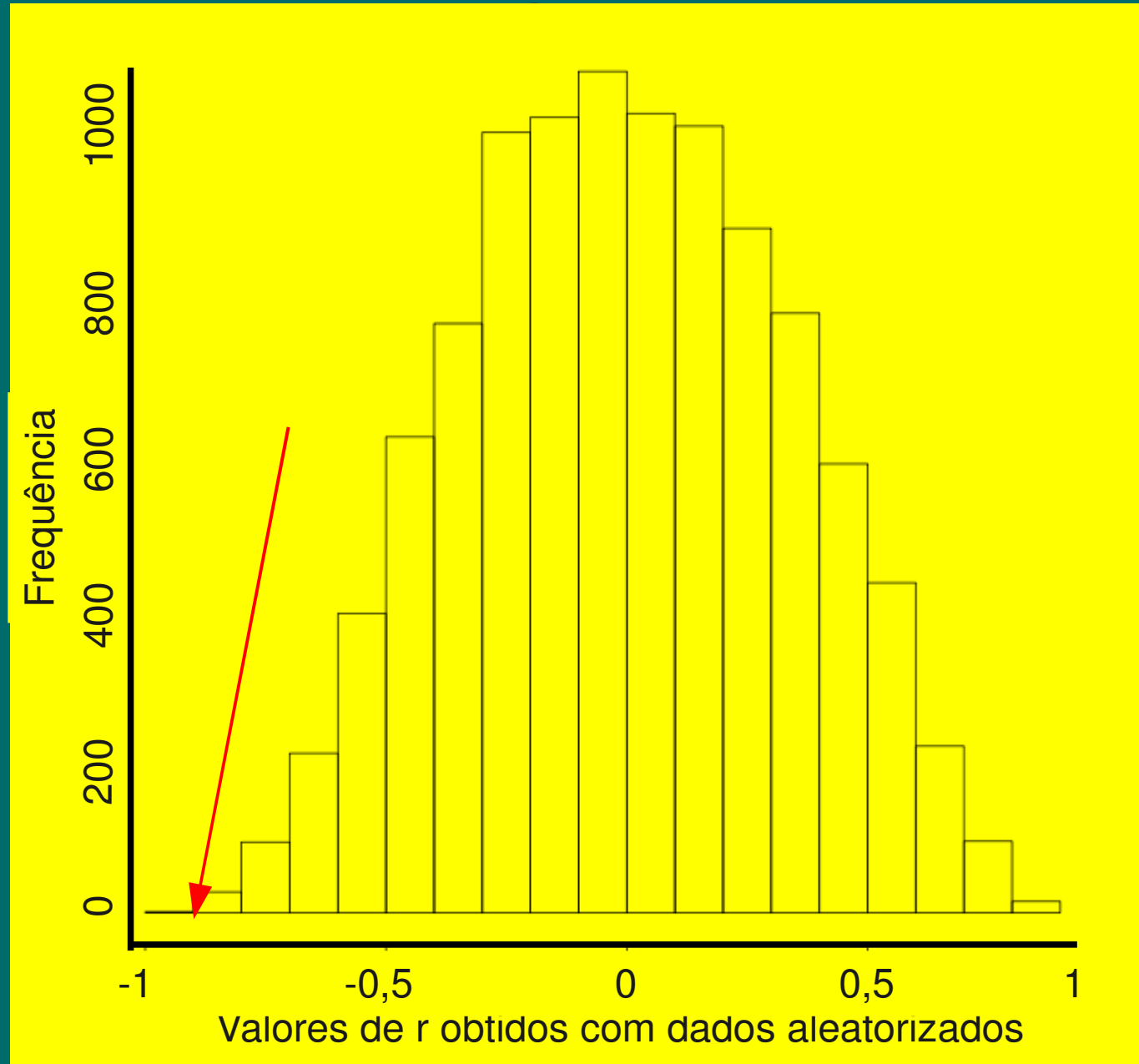
Testes monocaudais

Se o r observado foi de 0,9 e a distribuição for esta, concluo que seja significativa



Testes monocaudais

Mas se o r observado foi de $-0,9$ e a distribuição for esta, conluo que **NÃO** seja significativa



Testes monocaudais

Para um teste monocaudal, positivo, a fórmula seria:

$$p = \frac{(\text{frequência com que } esp > obs) + 1}{\text{número de aleatorizações} + 1}$$

Para um teste monocaudal, negativo, a fórmula seria:

$$p = \frac{(\text{frequência com que } esp < obs) + 1}{\text{número de aleatorizações} + 1}$$

Testes bicaudais e monocaudais

Quando uso um ou outro?

Posso escolher depois de fazer o teste?

NÃO

Posso escolher o monocaudal pois é significativo e o outro não?

NÃO

Posso escolher o monocaudal, a priori, pois acho-o mais interessante?

NÃO, você precisa justificar com argumento(s)!

Exercício com Planilha Eletrônica

(exercício aula passada)

Exercício com R

Fazer as aleatorizações no computador é fácil!

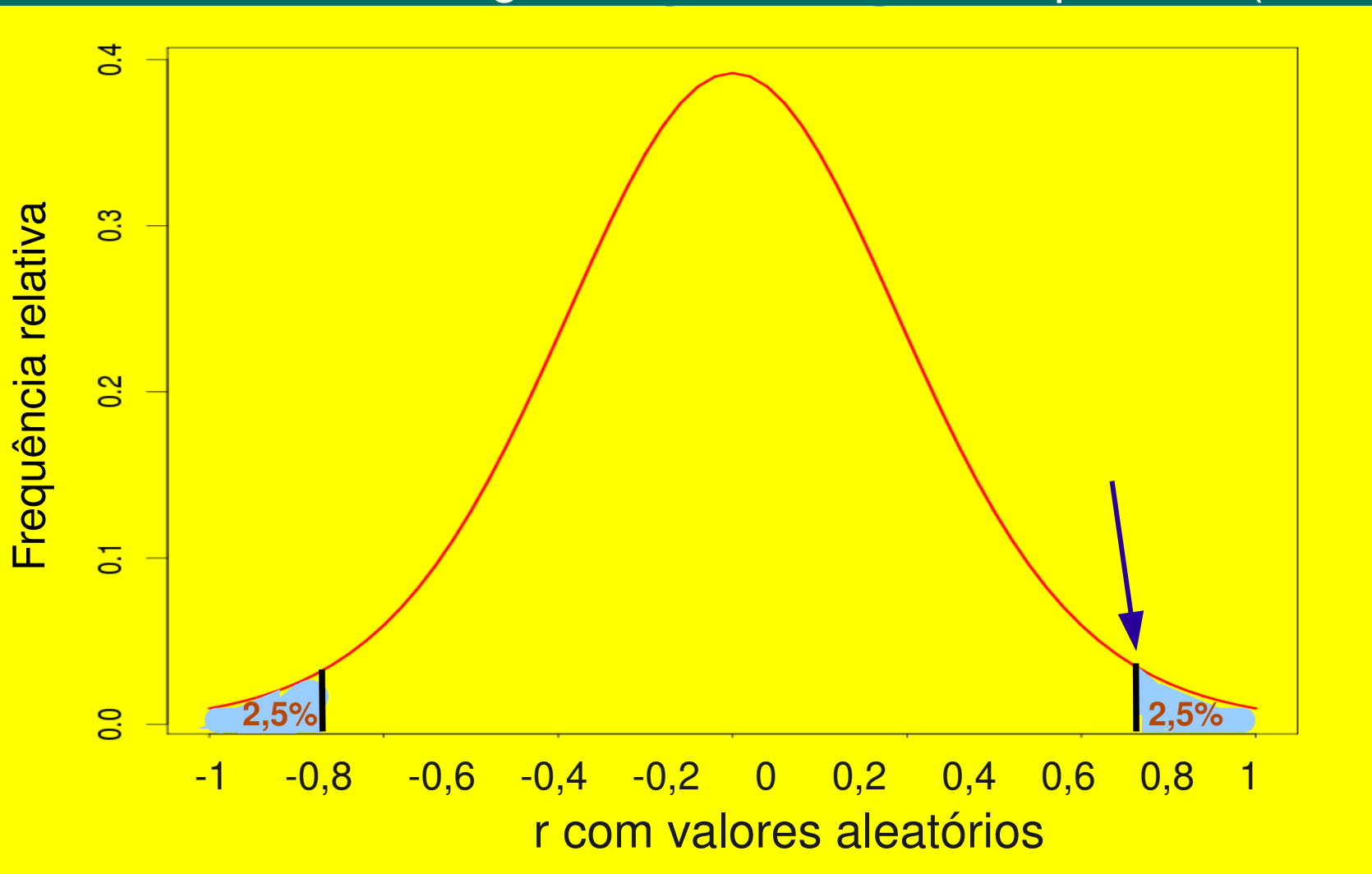
E quando não existia computador?

Fazer aleatorizações manualmente só é possível para conjuntos de dados muito pequenos.

Sugere algo?

As tabelas estatísticas

1. Pegue um conjunto de $n=5$.
2. Calcule a correlação e guarde o valor.
3. Repita o passo 2 muitas vezes e faça um histograma
4. Qual o valor 'crítico' de r de tal forma que temos 2,5% dos valores na cauda negativa e 2,5% na cauda positiva (total=5%)?



As tabelas estatísticas

O procedimento descrito anteriormente era aplicado para:

- a) várias estatísticas
- b) diferentes tamanhos amostrais
- c) diferentes níveis críticos

Veremos isto com mais detalhe no fim do curso.

Obs.: Em geral usamos uma fórmula para obter uma estatística t a partir do nosso valor de r e então obter valores críticos numa tabela t

O probleminha do início da aula...

Mais um probleminha...

1. Você está num quarto escuro e quer pegar um par de meias;
2. Na gaveta existem 3 meias idênticas exceto pela cor. Duas azuis e uma vermelha;
3. Ao pegar duas meias, a probabilidade de pegar 2 meias da mesma cor (azul) é maior, semelhante ou menor do que pegar cores diferentes?

RESPOSTA → $2 \text{ azuis} = \frac{2}{3} * \frac{1}{2} = \frac{1}{3}$

Sugestão de estudo:

Para probabilidades:

Vieira, S. 2008. Introdução à Bioestatística. 4a ed. Elsevier. (Cap. 8)

Para idéia de aleatorização:

Magnusson, W.E. & G. Mourão. 2003. Estatística *Sem* Matemática. Planta ed. (Cap. 5, pp. 39-48)