

Comparação de amostras

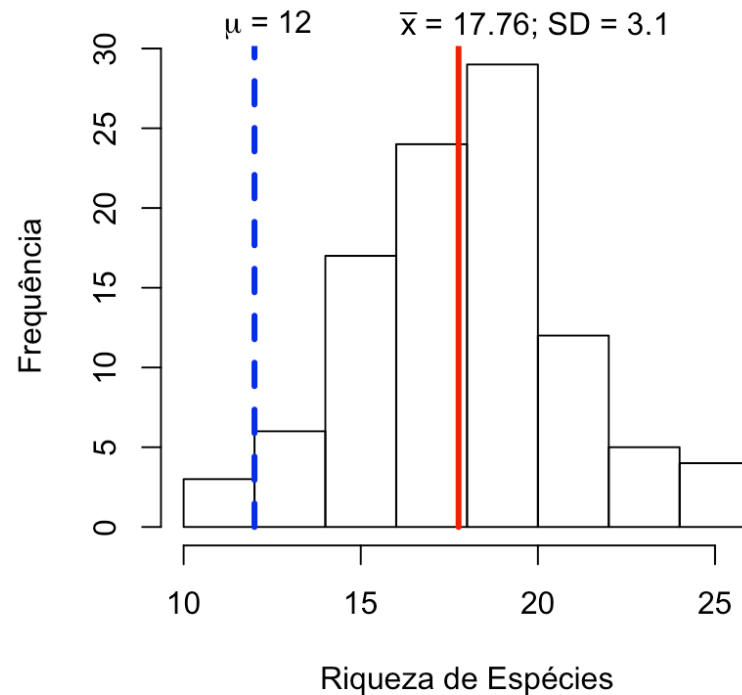
Testes de hipóteses baseados em uma ou mais amostras

Nicholas dos A. C. Marino

github.com/nacmarino

Imagine...

Em uma dada localidade, existem, em média, 17.76 espécies de larvas de insetos aquáticos por poça, com um erro padrão de 3.1. Entretanto, ao amostrar uma poça nesta mesma região, você encontrou apenas 12 espécies destes organismos. Isto fornece evidência para afirmar de que o valor que você encontrou é menor do que aquele esperado?



Conteúdo da Aula

1. Testes de hipóteses com uma amostra
 - Distribuição normal padrão
 - Distribuição t de Student
2. Testes de hipóteses com duas amostras
 - Teste-t de Student
3. Testes de hipóteses com mais de duas amostras
 - ANOVA
 - Pós-testes

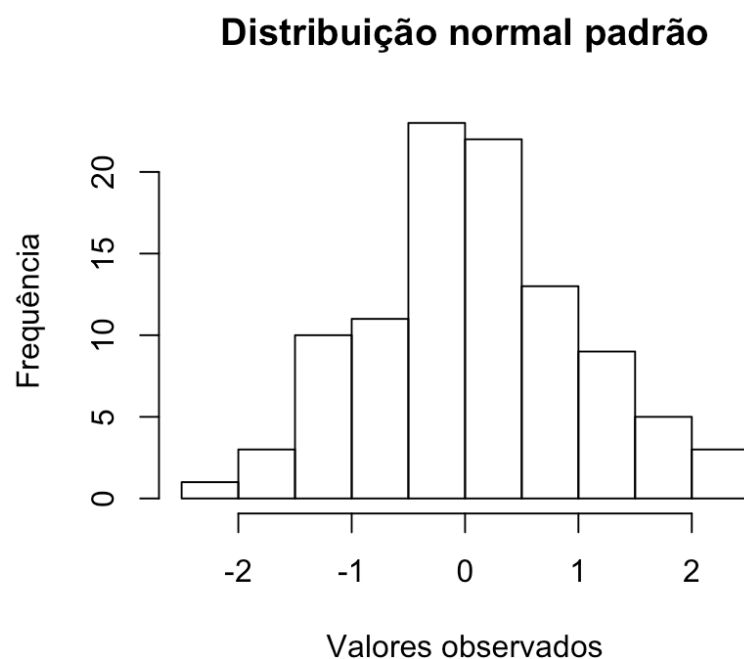
Testes de hipóteses com uma amostra

- Um parâmetro pode ser estimado a partir de uma amostra e sumariado na forma de um valor ou um intervalo de valores possíveis. Entretanto, às vezes, o objetivo de um estudo pode não ser em estimar tal parâmetro, mas decidir qual afirmação sobre um dado parâmetro é o correto.
 - H_a : A riqueza de espécies nas poças no seu estudo é **menor** do que o valor esperado; ou,
 - H_0 : A riqueza de espécies nas poças no seu estudo é **maior ou igual** ao valor esperado
- Podemos decidir qual das duas afirmações é mais sustentada por um conjunto de dados através de um valor de **teste estatístico** (*test statistic*).
 - Valor baseado em uma distribuição de probabilidade com propriedades conhecidas, e que será usado para discriminar entre as duas *hipóteses*, nos ajudando a decidir se a hipótese nula (H_0) deve ser rejeitada ou não.

A distribuição normal padrão

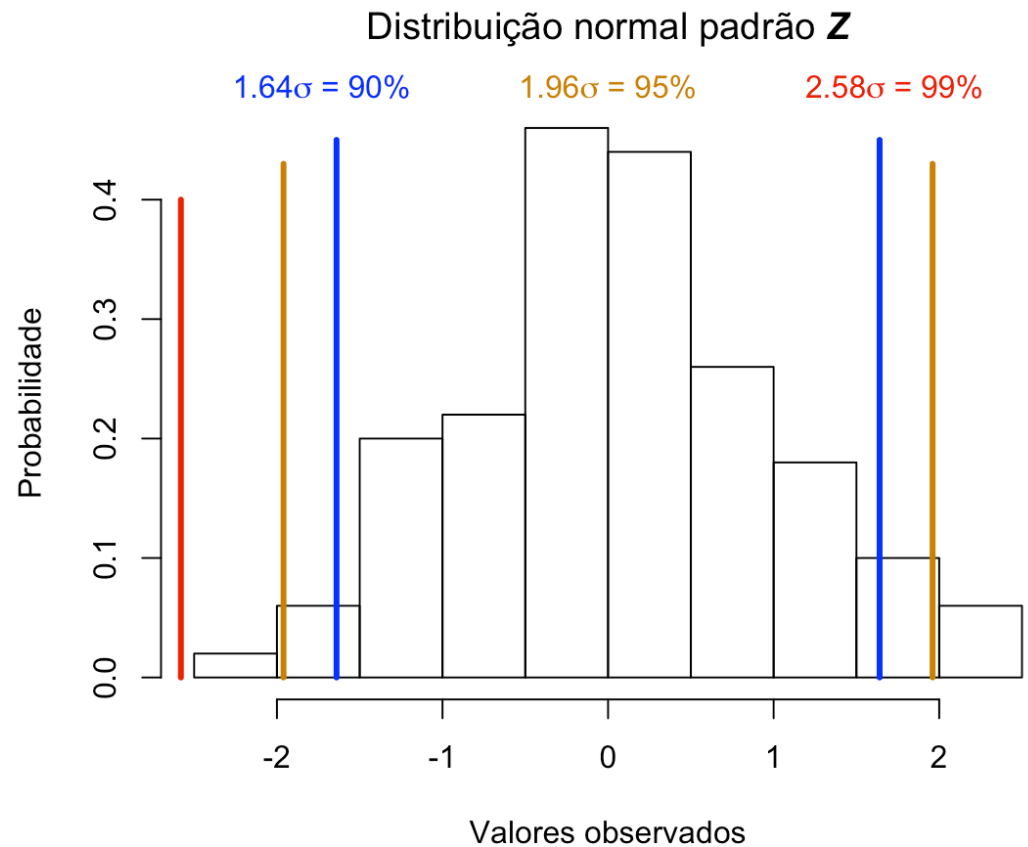
- É uma caso especial da distribuição normal, onde o valor da média (\bar{x}) de uma amostra é igual a 0 e o seu desvio padrão (σ) é igual a 1.

```
set.seed(123)  
hist(x = rnorm(n = 100, mean = 0, sd = 1))
```



A distribuição normal padrão

- De acordo com a distribuição normal padrão Z , quanto maior a diferença padronizada entre um dado valor e a média, menor a probabilidade daquele valor ser observado naquela amostra.



A distribuição normal padrão

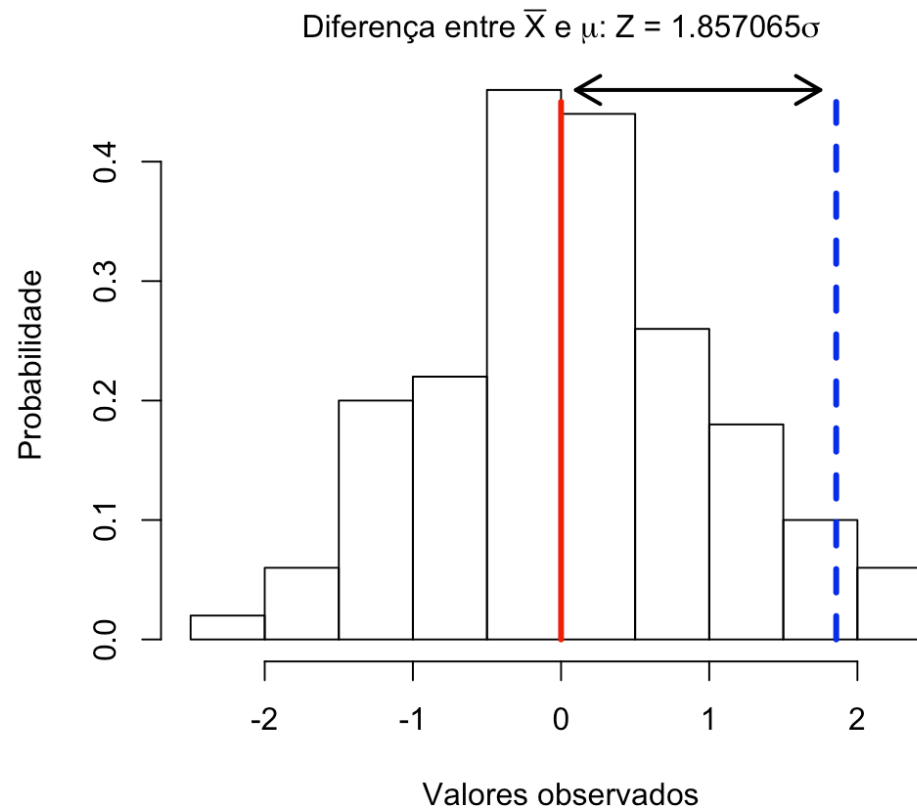
- Podemos usar o valor de Z como um teste estatístico para testar a H_0 de que a riqueza de espécies observada é maior ou igual àquela esperada.
- O valor do teste estatístico, Z , é:

$$Z = \frac{\bar{X} - \mu}{\sigma/\sqrt{n}}$$

- De acordo com o exemplo do início da aula:
 - Número de espécies médio nas poças: $\bar{X} = 17.76$;
 - Erro padrão da amostra: $\sigma/\sqrt{n} = 3.1$
 - Valor observado: $\mu = 12$
 - $Z = (17.76 - 12)/(3.1) = 1.858065$

No exemplo do início da aula

- A nossa estimativa está distante cerca de 1.85 unidades de erro padrão do valor reportado anteriormente.

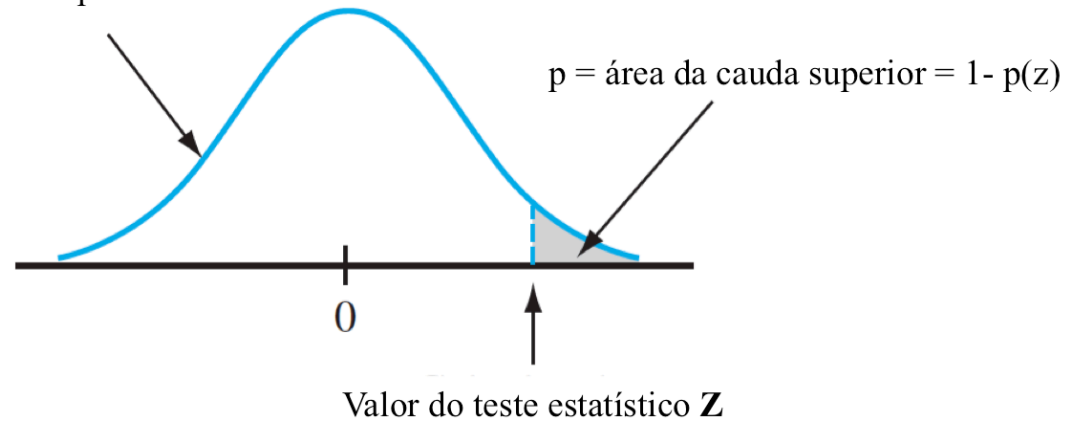


- Com o valor calculado do teste estatístico, Z , podemos testar a hipótese nula H_0 de que a riqueza de espécies observada é maior ou igual ao valor esperado.
- Testamos esta hipótese calculando a **probabilidade p de obtermos um valor de Z tão extremo quanto aquele observado assumindo que a hipótese nula é verdadeira.**
- Para tanto, devemos determinar qual o tipo de teste de hipótese devemos testar
 - Unicaudal: $H_a > H_0$ **OU** $H_a < H_0$
 - Bicaudal: $H_a \neq H_0$

Testes unicaudais

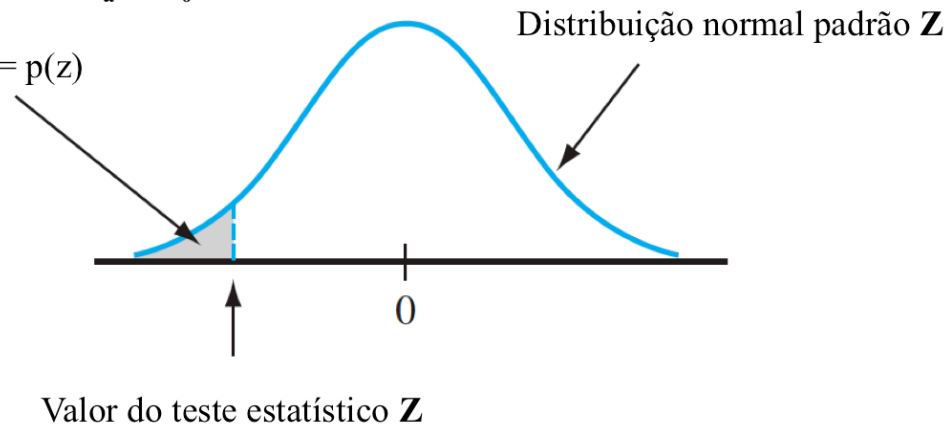
TESTE DA CAUDA SUPERIOR: $H_a > H_0$

Distribuição normal padrão Z



TESTE DA CAUDA INFERIOR: $H_a < H_0$

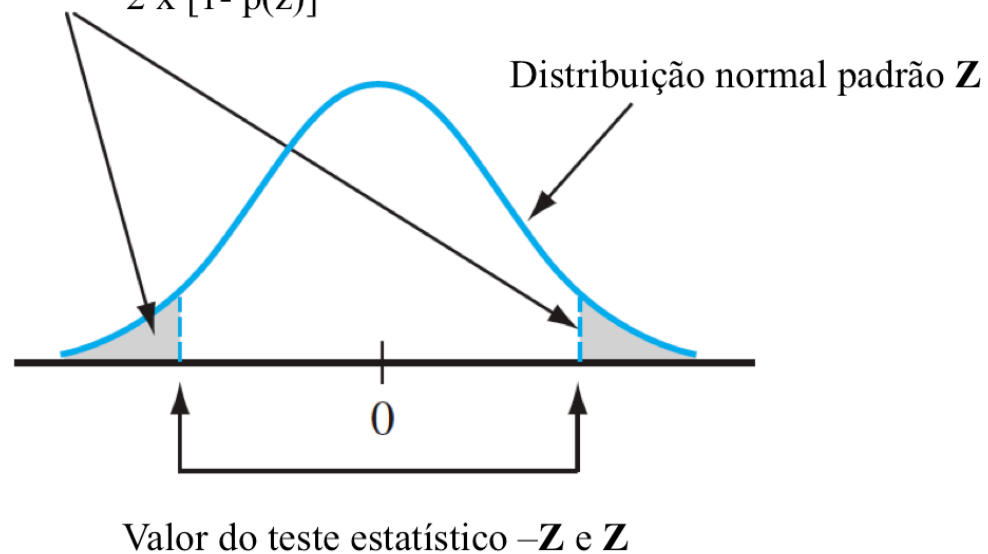
p = área da cauda superior = p(z)



Testes bicaudais

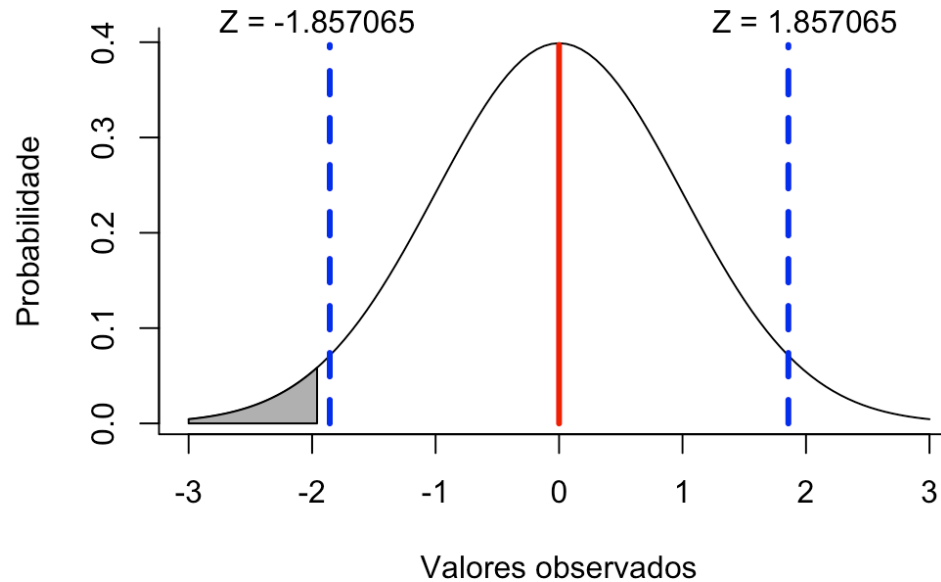
TESTE BICAUDAL: $H_a \neq H_0$

p = soma da área das duas caudas =
 $2 \times [1 - p(z)]$



No nosso exemplo...

- Nossa hipótese alternativa H_a é que a riqueza de espécies observada ($\mu = 12$) é menor do que a riqueza de espécies esperado ($\bar{x} = 17.76$): $H_a < H_0$.
- Apesar do nosso valor calculado de Z ser positivo ($Z = 1.857065$), nosso teste de hipótese 'pede' que ele seja negativo. Como as caudas da distribuição normal padrão são simétricas, podemos multiplicar o valor de Z por -1 para testar a hipótese nula e buscar o valor de p do teste estatístico.



- Podemos procurar o valor de p na tabela de distribuição de tais valores para a distribuição normal padrão.

- Ou podemos usar alguma outra ferramenta que calcule este valor para nós.
- No R, a função `pnorm` pode ser usada para isso.

```
pnorm(q = -1.857065)
```

```
## [1] 0.03165095
```

- Ou, de forma direta...

```
pnorm(q = 12, mean = 17.76, sd = 3.1)
```

```
## [1] 0.03157993
```

Exercício 1

1. Compare o resultado das duas linhas de comando abaixo. Qual a diferença entre elas e por que isso ocorre?

```
pnorm(q = 1.857065)
```

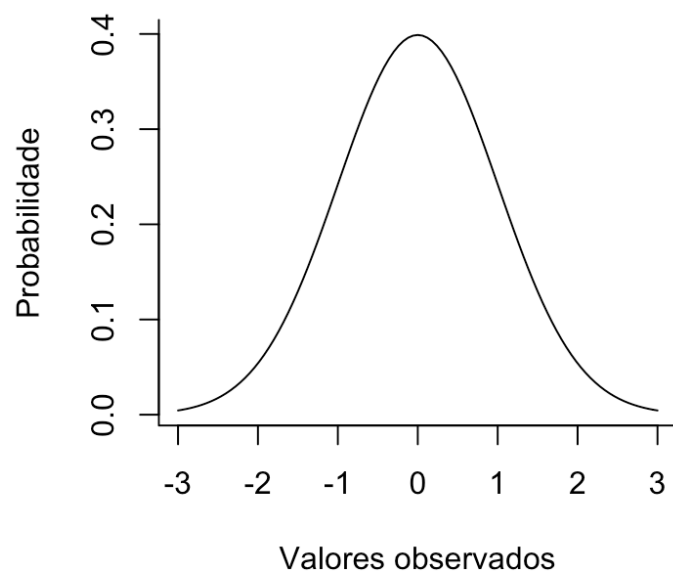
```
pnorm(q = 1.857065, lower.tail = FALSE)
```

A distribuição t de Student

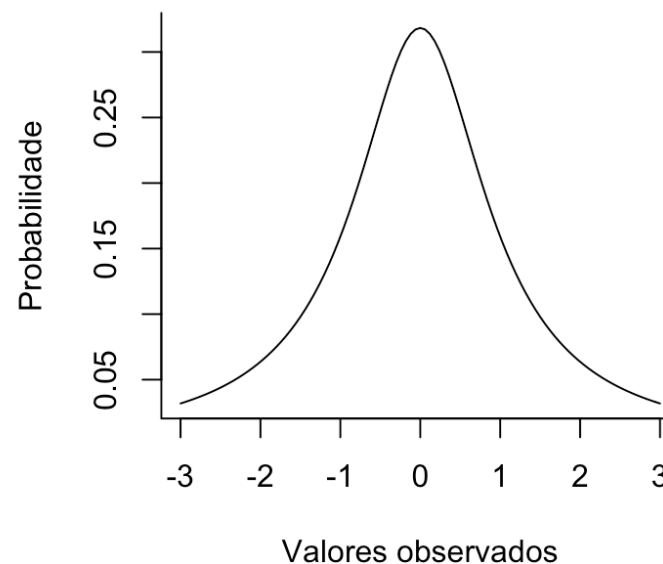
- Muito usada para testar *diferenças*, especialmente quando:
 - As observações de uma amostra seguem uma distribuição normal;
 - O tamanho da amostra é pequeno (normalmente < 40 observações); e,
 - O desvio padrão da população é desconhecido (não confundir com o desvio padrão da amostra!).
- Ao contrário da distribuição (padrão) normal, é descrita por apenas um parâmetro: os **graus de liberdade**.
- No R, você pode criar um conjunto de observações que sigam esta distribuição usando: `rt(n, df)`.
- Distribuição de densidade é muito similar àquela da distribuição normal, mas curva é mais 'magra'.

Distribuição normal padrão vs t de Student

Distribuição Normal Padrão Z

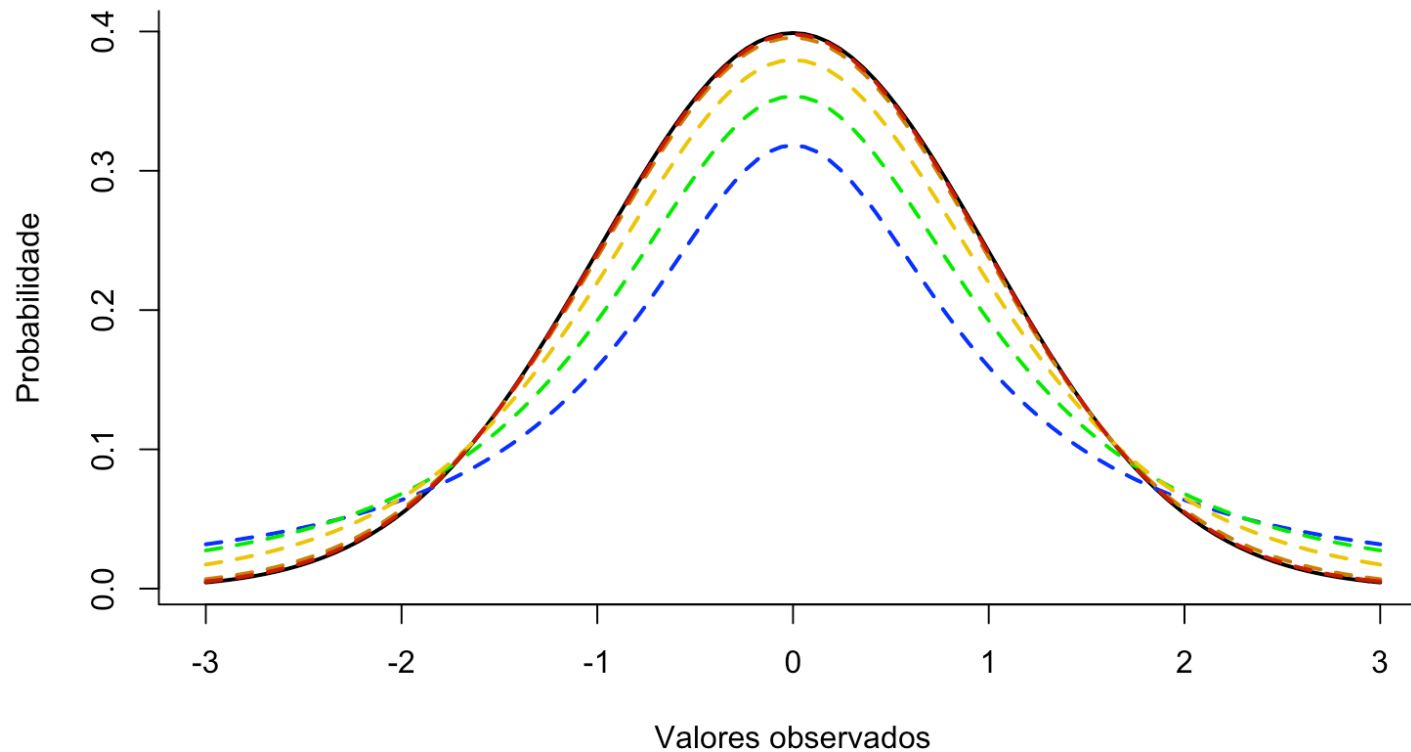


Distribuição t de Student



A distribuição t de Student

- Quanto maior o número de graus de liberdade, mais próximo a distribuição t de Student se aproxima da distribuição (normal) padrão.



Teste de hipóteses com uma amostra usando a distribuição t de Student

- O teste estatístico que segue a distribuição t de Student, T é calculado de forma muito similar ao valor de Z :

$$T = \frac{\bar{X} - \mu}{S/\sqrt{n}}$$

- Aqui, o desvio padrão da população, σ , é substituído pelo desvio padrão da amostral, S .
- Além disso, a obtenção da probabilidade de observarmos um valor do teste estatístico T mais extremo do que aquele esperado caso a hipótese nula seja verdadeira é dependente dos graus de liberdade, calculados com $n - 1$.

Teste de hipóteses com uma amostra usando a distribuição t de Student

Exercício 2

Ao amostrar 10 fragmentos florestais de tamanho similar, você obtém as seguintes estimativas de densidade para uma espécie de árvore:

```
abundancia <- c(11, 16, 43, 21, 21, 45, 26, 1, 9, 13)
```

Em uma paisagem não-fragmentada, esta mesma espécie apresenta uma densidade aproximada de 10 indivíduos para uma área similar àquela dos fragmentos. Neste sentido, poderíamos afirmar que a densidade registrada nos fragmentados é maior do que nesta área contínua?

- H_a : densidade nos fragmentos é > 10 ;
- H_0 : densidade nos fragmentos é ≤ 10 ;
- Calcule o valor de t e se a função `pt` para calcular o valor de p .

Exercício 2

A solução do exercício é:

```
## calculando o valor de t
valor_t <- (mean(abundancia) - 10)/(sd(abundancia)/sqrt(10))
valor_t

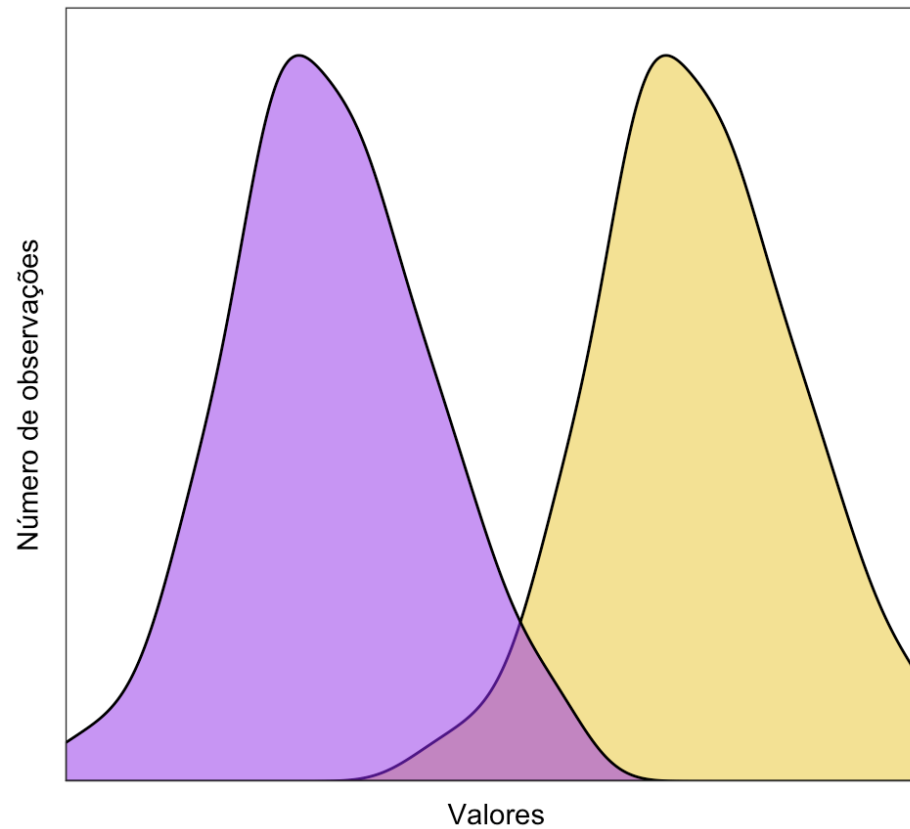
## [1] 2.359508

## calculando o valor de p
### como queremos determinar se o valor de acordo com Ha é > do que de
### acordo com H0, lower.tail = FALSE
pt(q = valor_t, df = 9, lower.tail = FALSE)

## [1] 0.02131788
```

Testes de hipóteses com duas amostras

Como determinar se dois conjuntos de amostras tomadas de uma população diferem?



O teste-t de Student

- É uma análise estatística paramétrica muito utilizada para determinar se duas amostras são diferentes;
 - Análise paramétrica: parâmetros estimados a partir dos dados pertencem à uma família de distribuição de probabilidade.
- Testa a hipótese nula de que os valores médios das observações vindas de duas amostras não diferem entre si; ou seja:
 - $H_0: \bar{x}_1 - \bar{x}_2 = 0$ **OU** $\bar{x}_1 = \bar{x}_2$
- Pressupostos:
 - Amostras seguem uma distribuição normal;
 - Amostras balanceadas;
 - Variâncias não diferem entre as amostras;
 - Observações em cada amostra e entre amostras são independentes entre si.

O teste-t de Student

- Quando todos os pressupostos do teste-t de Student são atendidos, podemos calcular o valor do teste estatístico, t , como a diferença entre as médias das duas amostras, padronizada pela medida agrupada do erro na estimativa das duas médias.

$$t = \frac{\overline{X}_1 - \overline{X}_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}}$$

- O numerador representa a magnitude da diferença entre as duas médias: quanto maior esta diferença, maior o valor de t ;
- O denominador representa a magnitude da variabilidade existente nas duas amostras: quanto maior a variabilidade, menor o valor de t .

O teste-t de Student

- Uma vez que tenhamos calculado o valor do teste estatístico t , podemos determinar a probabilidade p de obter um valor tão extremo quanto aquele observado para o teste estatístico caso a hipótese nula seja verdadeira - de maneira similar ao que fizemos para o teste t para uma amostra.
- Aqui, no entanto, o número de graus de liberdade utilizados para o teste de hipótese é $n_1 - 1 + n_2 - 1$.

Exemplo

A riqueza de espécies difere entre ilhas costeiras e ilhas oceânicas?

```
## carregando o conjunto de dados que vamos trabalhar
```

```
ilhas <- read_tsv(file = "data/ilhas.txt")
```

```
# qual o numero de observacoes em cada uma das duas categorias de tipos de ilhas?
```

```
table(ilhas$ilha)
```

```
##
```

```
## costeira oceanica
```

```
##          40          40
```

Exemplo

- Calcular a riqueza de espécies média entre os dois tipos de ilha;

```
media <- as.vector(by(data = ilhas$riqueza, INDICES = ilhas$ilha, mean))  
by(data = ilhas$riqueza, INDICES = ilhas$ilha, mean)
```

```
## ilhas$ilha: costeira  
## [1] 29.525  
## -----  
## ilhas$ilha: oceanica  
## [1] 15.425
```

- Calcular o desvio padrão da riqueza de espécies entre os dois tipos de ilha.

```
desvio <- as.vector(by(data = ilhas$riqueza, INDICES = ilhas$ilha, sd))  
by(data = ilhas$riqueza, INDICES = ilhas$ilha, sd)
```

```
## ilhas$ilha: costeira  
## [1] 10.79646  
## -----  
## ilhas$ilha: oceanica  
## [1] 16.51866
```

Exemplo

- Calculando o valor de t

```
## diferenca entre medias
diff_medias <- (media[1] - media[2])

## variabilidade total
var_total <- sqrt(((desvio[1]^2)/40) + ((desvio[2]^2)/40))

## valor de t
valor_t <- diff_medias/var_total
valor_t

## [1] 4.51892
```

- Testando a nossa hipótese: $\bar{x}_{\text{costeira}} \neq \bar{x}_{\text{oceanica}}$

```
## como não determinamos se a diferença é positiva ou negativa, usaremos um teste bicaudal
2 * (1 - pt(q = valor_t, df = ((40 - 1) + (40 - 1))))

## [1] 2.180675e-05
```

Simplificando o teste-t

- No R, basta usar:

```
t.test(riqueza ~ ilha, data = ilhas, var.equal = TRUE, alternative = "two.sided")
```

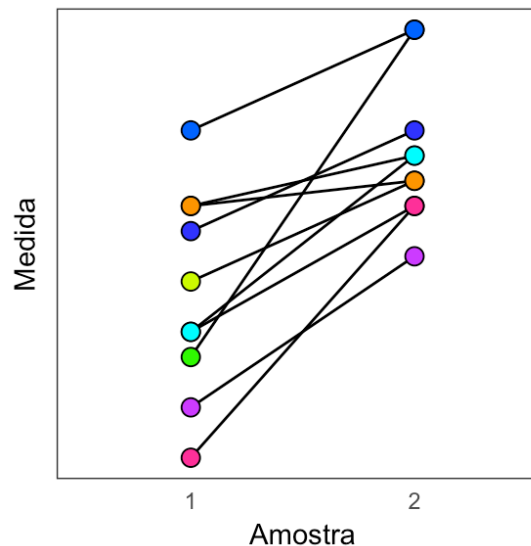
```
##  
## Two Sample t-test  
##  
## data:  riqueza by ilha  
## t = 4.5189, df = 78, p-value = 2.181e-05  
## alternative hypothesis: true difference in means is not equal to 0  
## 95 percent confidence interval:  
##    7.888131 20.311869  
## sample estimates:  
## mean in group costeira mean in group oceanica  
##                29.525                15.425
```

Exercício 3

E se quisséssemos testar se as ilhas costeiras têm uma riqueza de espécies maior do que as ilhas oceânicas, como deveríamos fazer a análise?

Teste-t pareado

- Às vezes, nossa pergunta principal pode envolver comparações entre observações que **não** são independentes - nestes casos, podemos usar um **teste-t pareado**:
 - Design Before-After-Control-Impact (BACI);
 - Medidas nos mesmos indivíduos/populações em dois momentos diferentes;



Teste-t pareado

- Testa a hipótese nula de que não há diferença entre cada par de observações é zero: $\bar{D} = 0$.

```
riachos <- read_tsv(file = "data/riachos.txt")
riachos$diferenca <- riachos$montante - riachos$jusante
riachos
```

```
## # A tibble: 10 x 4
##   riacho   montante jusante diferenca
##   <chr>     <int>   <int>     <int>
## 1 riacho1      13      8         5
## 2 riacho2      14     10         4
## 3 riacho3      20      7        13
## 4 riacho4      15     13         2
## 5 riacho5      15      8         7
## 6 riacho6      20     16         4
## 7 riacho7      16     12         4
## 8 riacho8      11      5         6
## 9 riacho9      13      3        10
## 10 riacho10     14     13         1
```

Teste-t pareado

- De forma análoga ao teste-t padrão, o valor do teste estatístico t é obtido através da razão entre a magnitude média da diferença entre os pares de observações e a sua variabilidade.

$$t = \frac{\overline{D}}{\sqrt{\frac{s_D^2}{n}}}$$

- O número de graus de liberdade utilizados para o teste de hipóteses é $n - 1$, onde n é o número de pares independentes.
- No R, podemos fazer um teste-t pareado utilizando o argumento `paired = TRUE` dentro da função `t.test`.

Exemplo de teste-t pareado

- Podemos fazer um teste-t pareado no R utilizando:

```
t.test(x = riachos$montante, y = riachos$jusante, paired = TRUE, var.equal = TRUE)
```

```
##  
## Paired t-test  
##  
## data: riachos$montante and riachos$jusante  
## t = 4.8824, df = 9, p-value = 0.0008685  
## alternative hypothesis: true difference in means is not equal to 0  
## 95 percent confidence interval:  
## 3.005358 8.194642  
## sample estimates:  
## mean of the differences  
## 5.6
```

- Note a diferença para: `t.test(x = riachos$montante, y = riachos$jusante, paired = FALSE, var.equal = TRUE)`.

Teste-t quando as variâncias não são homogêneas

- Também é possível utilizar um teste-t quando as variâncias entre as duas amostras não são homogêneas;
- O teste-t com variâncias não homogêneas também é conhecido como teste de Welch;
- Neste caso, todos os pressupostos do teste-t padrão continuam sendo válidas, mas os graus de liberdade utilizados para o teste de hipóteses são ajustados:

$$\text{d.f.} = \frac{\left(\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2} \right)^2}{\frac{(s_1^2/n_1)^2}{n_1-1} + \frac{(s_2^2/n_2)^2}{n_2-1}}.$$

Teste de Welch no R

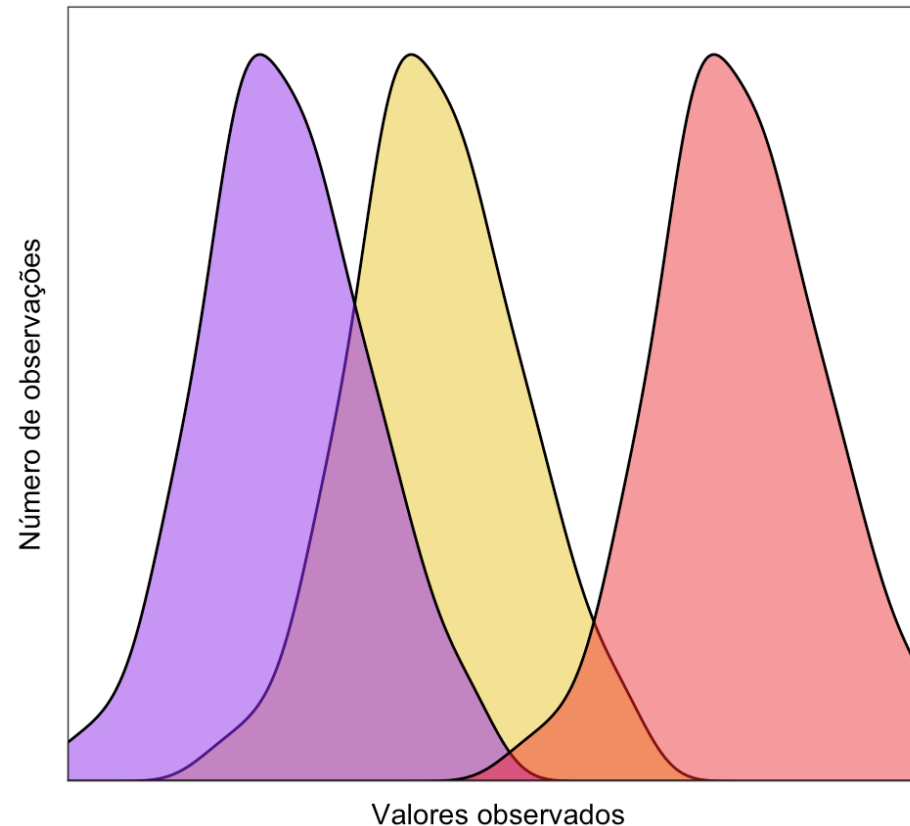
- O teste de Welch bicaudal é o padrão assumido pelo R quando você não especifica nenhum outro argumento na função `t.test`.

```
t.test(riqueza ~ ilha, data = ilhas)
```

```
##  
## Welch Two Sample t-test  
##  
## data:  riqueza by ilha  
## t = 4.5189, df = 67.178, p-value = 2.585e-05  
## alternative hypothesis: true difference in means is not equal to 0  
## 95 percent confidence interval:  
##    7.87233 20.32767  
## sample estimates:  
## mean in group costeira mean in group oceanica  
##                29.525                15.425
```

Testes de hipóteses com mais de duas amostras

E quando temos mais de duas amostras? Podemos fazer múltiplos testes-t?



Análise de Variância (ANOVA)

- É uma análise estatística utilizada para testar diferenças entre os valores das médias de duas ou mais amostras/tratamentos de uma única vez;
- Testa a hipótese nula de que não existem diferenças nos valores das médias entre as amostras/tratamentos:

$$H_0: \mu_1 - \mu_2 - \mu_3 - \dots - \mu_n = 0$$

OU

$$H_0: \mu_1 = \mu_2 = \mu_3 = \dots = \mu_n$$

- Não te diz qual média difere de qual, apenas que pelo menos uma das médias é diferente das demais - descobrir qual delas é uma tarefa para os pós-testes.

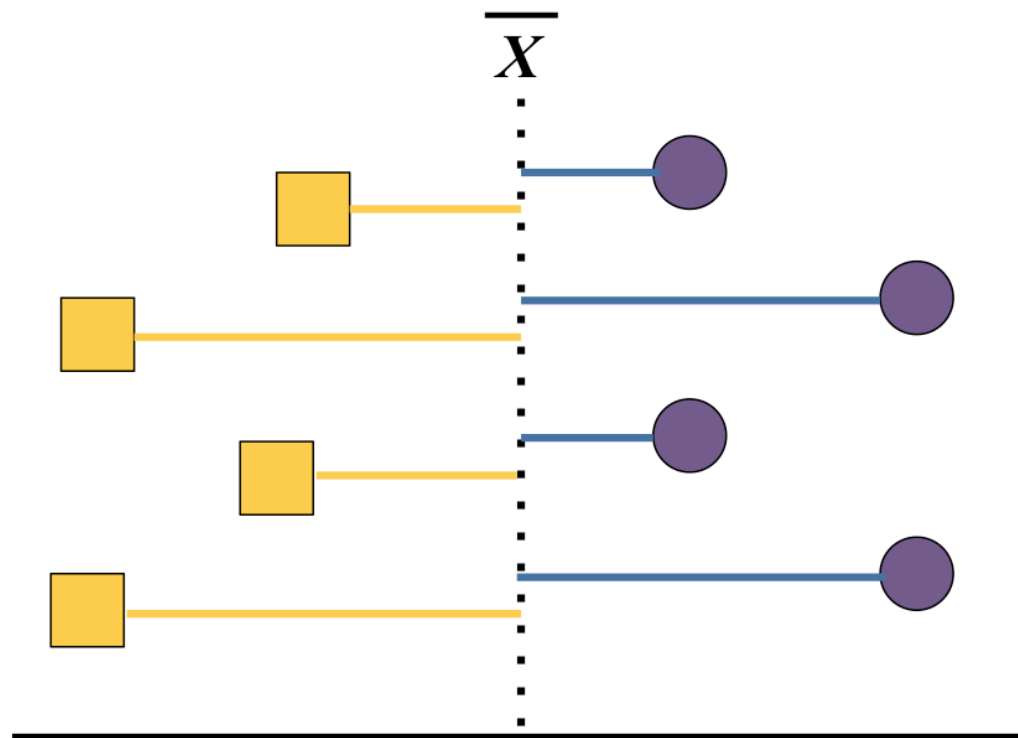
Principais pressupostos da ANOVA

1. **Resíduos** seguem uma distribuição normal;
 - Transformação da variável resposta: log, sqrt, asin(sqrt), raiz cúbica, raiz quarta, inverso, logit...
 - Algumas pessoas suportam que pode ser feita;
 - Outras pessoas suportam que não deve ser feita;
 - Para mim, depende da sua pergunta e objetivo.
2. Variâncias entre amostras/tratamentos são homogêneas (**homocedasticidade das variâncias**);
3. Valores das amostras/tratamentos são fixos, normalmente categorias;
4. Observações são independentes, espacialmente e temporalmente.

A mecânica da ANOVA

- Mas se é uma análise de **variância**, porque testa uma hipótese nula sobre as **médias**?
- A ANOVA calcula:
 - A diferença total entre cada observação e a média total do conjunto de dados composto por várias amostras/tratamentos (**total**): $x_i - \bar{x}$
 - A diferença total entre cada observação pertencente à uma amostra/tratamento e o valor da média daquele tratamento (**within-subject**): $x_i - \bar{x}_j$; e,
 - A diferença total entre o valor da média de cada tratamento e o valor da média de todo o conjunto de dados (**between-subject**): $\bar{x}_j - \bar{x}$.
- O problema das diferenças: seu somatório tende a ser zero ou muito próximo dele, pois os desvios positivos 'anulam' os desvios negativos;
- A solução é elevar cada uma destas diferenças ao quadrado - ou seja, utilizar as variâncias como medida de diferenças entre médias:
 - Total: $\sum (x_i - \bar{x})^2$
 - Within-subject: $\sum (x_i - \bar{x}_j)^2$
 - Between-subject: $\sum (\bar{x}_j - \bar{x})^2$

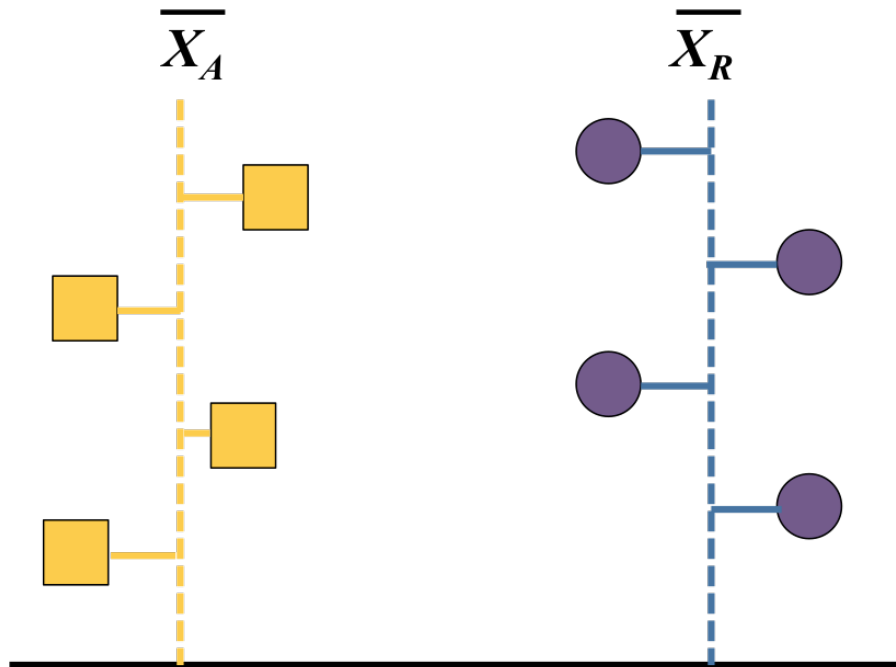
A mecânica da ANOVA



Calcular a média entre todo o conjunto de dados \bar{X} e, então, a diferença entre cada observação x_i e a média geral para depois realizar a soma. Este valor será a soma dos quadrados total, SS_T .

$$SS_T = \sum (X_i - \bar{X})^2$$

A mecânica da ANOVA

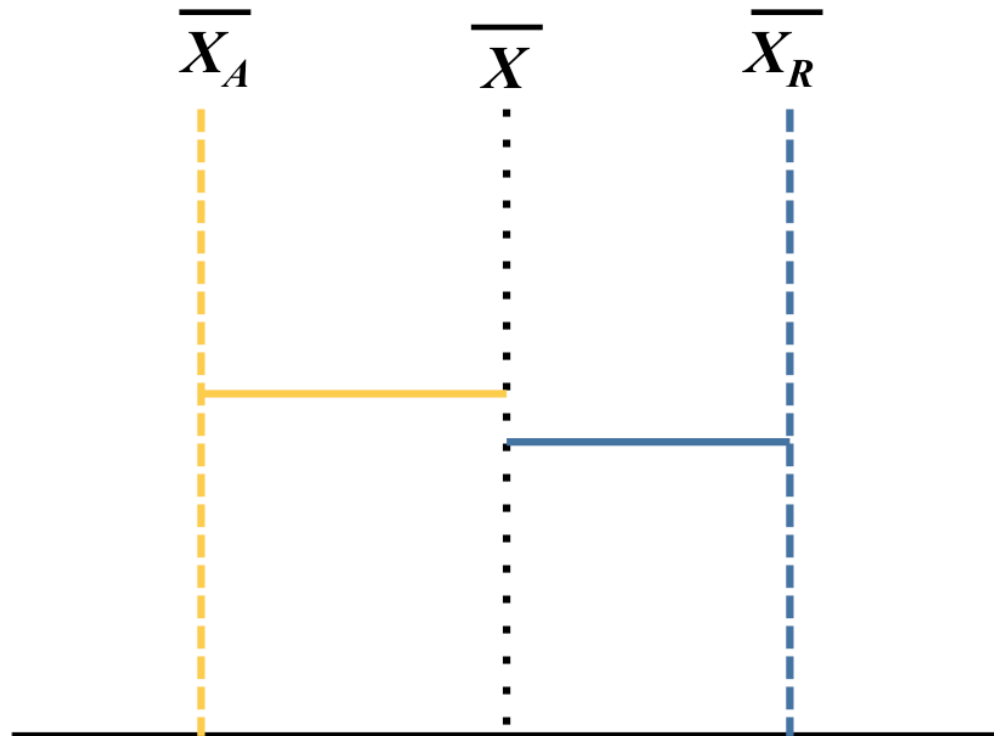


Calcular a média de cada tratamento e, então a diferença entre cada observação em um tratamento e sua média e somar:

- Tratamento A: $\sum (X_{Ai} - \overline{X}_A)^2$
- Tratamento R: $\sum (X_{Ri} - \overline{X}_R)^2$

Estes valores serão, respectivamente, SS_A e SS_R . Somaremos estes dois valores e teremos o SS_W – a Soma dos Quadrados dentro dos Tratamentos.

A mecânica da ANOVA



Calcular a diferença entre o valor médio de cada tratamento e o valor da média de todo o conjunto de dados, e depois somar os dois valores:

- Tratamento A: $(\bar{X}_A - \bar{X})^2$
- Tratamento B: $(\bar{X}_B - \bar{X})^2$

A soma destes dois valores representa a soma dos quadrados entre os tratamentos, SS_B .

A mecânica da ANOVA

- Ao final do processo, acabamos com três tipos de soma dos quadrados, que se relacionam na forma:

$$SS_T = SS_B + SS_W$$

- Em outras palavras, a variabilidade total em um conjunto de dados (SS_T) é particionada entre o que pode ser explicado pela diferença existente entre amostras/tratamentos (SS_B) e pela variação natural existente dentro de cada tratamento ou a variação *residual* (SS_W).
- Uma vez que se consiga quantificar duas destas quantidades, a terceira pode ser determinada por subtração ou soma.

A mecânica da ANOVA

- Precisamos também considerar o efeito no número de amostras/tratamentos e observações no cálculo da Soma dos Quadrados:
 - Quanto maior o número de amostras/tratamentos, maior tende a ser o SS_B ;
 - Quanto maior o número de observações, maior tende a ser o SS_W .
- Por conta disso, cada um destes termos deve ser ponderado pelo seu número de graus de liberdade:
 - SS_B : $n - 1$, onde n aqui é o número de amostras/tratamentos;
 - SS_W : $k - n$, onde k aqui é o número total de observações;
 - SS_T : $k - 1$
- A divisão de cada uma destas somas dos quadrados pelos seus respectivos graus de liberdade dá origem à uma medida conhecida com os Quadrados Médios (MSE_B , MSE_W e MSE_T , respectivamente).
- Mas qual o tipo de valor de teste estatístico utilizado pela ANOVA?

A distribuição de probabilidade F

- É uma família de distribuição de probabilidade utilizada para descrever razões entre medidas de variância ou fontes de variabilidade em um conjunto de dados.
- O valor do teste estatístico F é obtido como: $F = \frac{\text{num } df}{\text{den } df}$
- Os dois parâmetros utilizados para calcular este teste estatístico são:
 - **num df**: número de graus de liberdade no numerador
 - **den df**: número de graus de liberdade no denominador
- Uma variável que assuma uma distribuição F não pode apresentar valores negativos; portanto, valores que seguem a distribuição F são sem $0 > x > +\infty$.

Exercício 4

- Utilize o código abaixo para entender de que forma funciona distribuição de probabilidade de F.

```
## criando uma função para plotar diferentes valores de num df e den df
```

```
distF <- function(numdf, dendf) {  
  curve(df(x, df1 = numdf, df2 = dendf), xlim = c(0, 10),  
        xlab = "Valor de F", ylab = "Probabilidade")  
}
```

```
## criando um modo de realizarmos manipulações automatizadas em cima do gráfico
```

```
## gerado pela função distF
```

```
library(manipulate)  
manipulate(distF(numdf, dendf),  
           numdf = slider(2, 100, step = 1),  
           dendf = slider(2, 100, step = 1))
```


A distribuição F como teste estatístico para a ANOVA

- Em analogia à distribuição, calculamos um valor de teste estatístico F para cada termo em uma ANOVA que represente a variabilidade em função a diferenças entre amostras/tratamentos;
- Este valor do teste estatístico é calculado como: $F = \frac{MSE_B}{MSE_W}$;
- Quanto maior a diferença média entre amostras/tratamentos (MSE_B) em relação àquela diferença média que não pode ser explicada por ela (MSE_W), maior será o valor de F;
- Como vimos no exercício, quanto maior o valor de F, maior a probabilidade de rejeitarmos a hipótese nula de que não existem diferenças nos valores das médias entre as amostras/tratamentos.

Exemplo

- No R, podemos fazer uma ANOVA através da função `aov` ou, de preferência, associando a função `lm` à função `anova`.

```
## criando um modelo
```

```
modelo <- lm(riqueza ~ arquipelago, data = ilhas)
```

```
## rodando uma ANOVA no modelo
```

```
anova(modelo)
```

```
## Analysis of Variance Table
```

```
##
```

```
## Response: riqueza
```

```
##           Df  Sum Sq Mean Sq F value    Pr(>F)
```

```
## arquipelago  2   3548.1   1774.0    8.7475 0.0003773 ***
```

```
## Residuals   77  15615.9    202.8
```

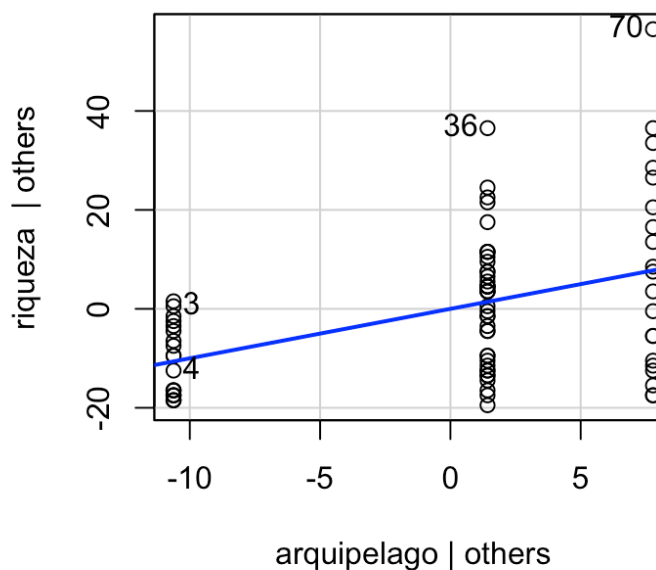
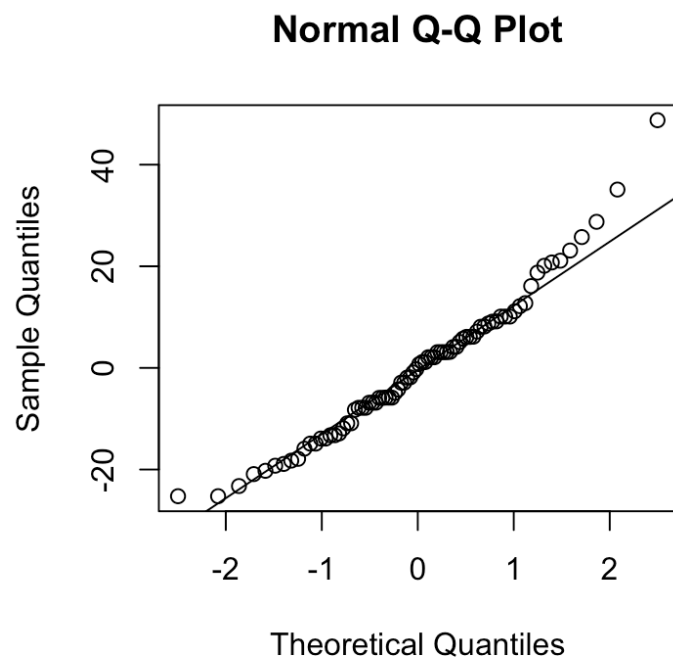
```
## ---
```

```
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Um modelo criado é um modelo final?

- A resposta é **não**: para que tenhamos um modelo final, ele precisa ser validado de acordo com os pressupostos da análise em questão e, também, pelo bom senso.
- A função `plot`, pode nos ajudar a avaliar a adequabilidade de um modelo aos dados.

```
plot(modelo)
```



Equivalência entre o teste-t e a one-way ANOVA com dois níveis

- A ANOVA também pode ser usada para testar diferenças entre amostras/tratamentos quando estes possuem dois níveis - assim como você usaria um teste t.
- No geral, existe uma equivalência entre os valores de F e de t na forma: $F = t^2$.

compare

```
t.test(riqueza ~ ilha, data = ilhas)  
anova(lm(riqueza ~ ilha, data = ilhas))
```

Pós-testes relacionadas à ANOVA

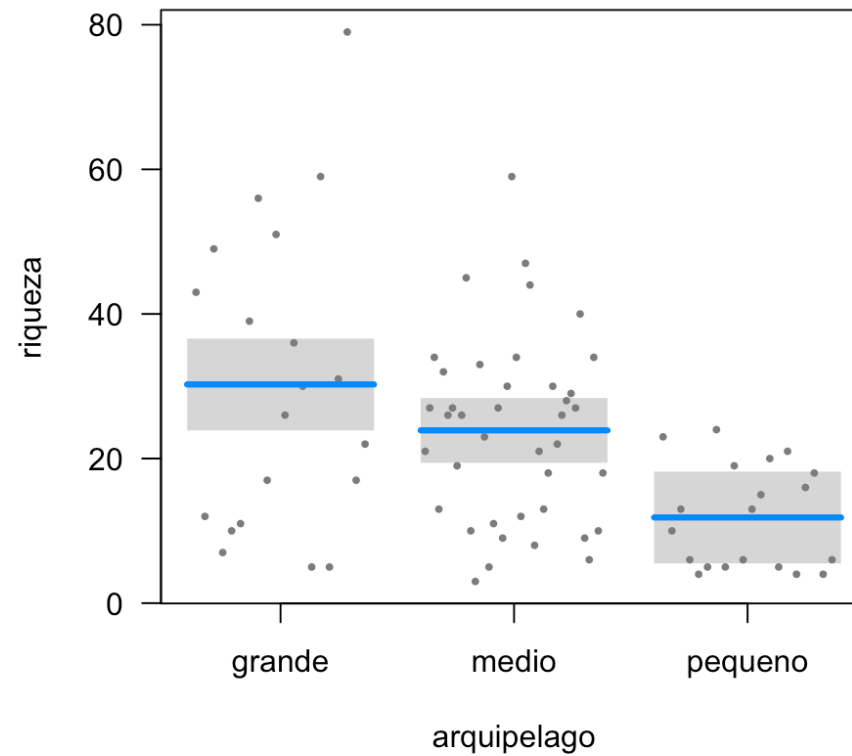
- Encontramos diferenças entre os níveis dos fator `arquipelago`, mas qual(is) destes níveis é (são) diferente(s) do que?

```
knitr::kable(anova(modelo), format = "markdown", digits = 4)
```

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
arquipelago	2	3548.05	1774.0250	8.7475	4e-04
Residuals	77	15615.90	202.8039	NA	NA

Pós-testes relacionadas à ANOVA

- A visualização dos dados já é um bom começo, mas não é o suficiente.



Uma breve visita ao **summary** do modelo

- O que cada uma destas informações está nos contando?

```
summary(modelo)
```

```
##
## Call:
## lm(formula = riqueza ~ arquipelago, data = ilhas)
##
## Residuals:
##      Min       1Q   Median       3Q      Max
## -25.250  -8.912   0.250   8.113  48.750
##
## Coefficients:
##              Estimate Std. Error t value Pr(>|t|)
## (Intercept)      30.250      3.184   9.500 1.31e-14 ***
## arquipelagomedio  -6.350      3.900  -1.628 0.107570
## arquipelagopequeno -18.400      4.503  -4.086 0.000107 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 14.24 on 77 degrees of freedom
## Multiple R-squared:  0.1851, Adjusted R-squared:  0.164
## F-statistic: 8.747 on 2 and 77 DF, p-value: 0.0003773
```

Como determinar os valores estimados para cada tratamento de um fator?

- O pacote **lsmeans** oferece diversas funcionalidades para realizar o pós-teste e outras análises mais detalhadas a partir de modelos estabelecidos.

```
lsmeans(object = modelo, specs = ~ arquipelago)
```

```
## arquipelago lsmean      SE df  lower.CL upper.CL
## grande      30.25 3.184367 77 23.909115 36.59089
## medio       23.90 2.251688 77 19.416317 28.38368
## pequeno     11.85 3.184367 77  5.509115 18.19089
##
## Confidence level used: 0.95
```


Comparações múltiplas

- Poderíamos fazer vários testes-t para fazer a comparação múltipla, mas...
 - Incorreríamos em aumentar a chance do Erro do Tipo I;
 - Teríamos um trabalho enorme em realizar cada comparação par-a-par conforme o número de tratamentos aumenta.
- Uma opção muito usada é a correção de Bonferroni, onde o valor de α é corrigido de acordo com o número de comparações sendo feitas, n :

$$\frac{\alpha}{n}$$

- Por outro lado, esta correção também pode acabar aumentando a chance do Erro Tipo II, quando o número de testes é muito grande.

Teste de Tukey

- O teste de Tukey é um dos mais conhecidos e utilizados em ecologia.
- Ele é baseado no teste t de Student, com a modificação de que a variabilidade média dentro dos tratamentos (MSE_W) é considerada no lugar da variabilidade entre as duas amostras/tratamentos sendo testadas.

$$t = \frac{\bar{X}_1 - \bar{X}_2}{\sqrt{MSE_W \left(\frac{1}{n_1} + \frac{1}{n_2} \right)}}$$

Exemplo da aplicação do teste de Tukey

- O pacote `lsmeans` oferece diversas funcionalidades para realizar o pós-teste e outras análises mais detalhadas a partir de modelos estabelecidos.

```
medias <- lsmeans(object = modelo, specs = ~ arquipelago)
contrast(medias, method = "tukey")
```

```
## contrast      estimate      SE df t.ratio p.value
## grande - medio      6.35 3.900037 77   1.628  0.2400
## grande - pequeno    18.40 4.503375 77   4.086  0.0003
## medio - pequeno     12.05 3.900037 77   3.090  0.0078
##
## P value adjustment: tukey method for comparing a family of 3 estimates
```

Contrastes

- Outra opção válida de pós-teste é a realização de **contrastes** planejados, que, normalmente, são utilizados quando temos hipóteses à priori que as justifiquem.
- O número de contrastes que normalmente podem ser feitos para os tratamentos de um mesmo fator são $n - 1$, onde n é o número de níveis dos fatores.

```
## as ilhas em arquipelagos grandes diferem das demais?; e,  
## ilhas em arquipelagos pequenos diferem das ilhas em arquipélagos médios  
contrast(medias, list("Grande vs outros" = c(1, -0.5, -0.5),  
                      "Pequeno vs Médio" = c(0, 1, -1)))
```

##	contrast	estimate	SE	df	t.ratio	p.value
##	Grande vs outros	12.375	3.734002	77	3.314	0.0014
##	Pequeno vs Médio	12.050	3.900037	77	3.090	0.0028

Pós-testes sem ajustes

- Também existe a possibilidade ignorar qualquer tipo de ajuste ao realizar os pós-testes da ANOVA.

```
contrast(medias, method = "tukey", adjust = "none")
```

##	contrast	estimate	SE	df	t.ratio	p.value
##	grande - medio	6.35	3.900037	77	1.628	0.1076
##	grande - pequeno	18.40	4.503375	77	4.086	0.0001
##	medio - pequeno	12.05	3.900037	77	3.090	0.0028

Afinal, qual pós-teste usar?

- É a pergunta de ouro, que depende da sua pergunta, da sua abordagem estatística e dos seus dados.