**Exemplo de teste de hipóteses com a distribuição de Student (t)**

**Uso:** Esta distribuição permite o teste de hipóteses entre médias de dois grupos amostrais, tanto para n < 30 quanto para n>= 30, em que se aproxima da distribuição normal.

Para ilustrar, suponha o seguinte exercício. Um pesquisador mediu **a duração da sílaba tônica [do]** de oito repetições dos pares de frases “O bor**dô** chinês derramou-se pela mesa.” e “O bor**dô** xucro derramou-se pela mesa.”, para as quais se troca a oxítona “chinês” (frase de controle) pela paroxítona “xucro” (frase de teste). Deseja-se avaliar se há uma diferença significativa na duração da sílaba tônica [do] entre as frases de controle e de teste devido à proximidade das sílabas tônicas na frase de teste. Usando as durações das linhas abaixo, o sistema de teste de hipóteses H0: μ1 = μ2  contra Ha: μ1 ≠ μ2, decida se há diferença (para α = 2 %) entre as médias das duas linhas de duração da tabela abaixo (Não há diferença estatística entre as variâncias dos grupos de amostrais pelo teste F).

**Tabela:** *Duração em milissegundos da sílaba [do] diante de palavra*

*dissilábica oxítona ou paroxítona, como especificado acima.*

|  |  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- | --- |
| **antes de paroxítona** | 161 | 174 | 166 | 170 | 147 | 175 | 177 | 185 |
| **antes de oxítona** | 166 | 143 | 152 | 123 | 145 | 154 | 176 | 149 |

O conjunto de hipóteses é este:

H0: μ1 = μ2

Ha: μ1 ≠ μ2

α = 0,02

É um teste t de variáveis independentes, para o qual texperimento = (x̄1 - x̄2)/√[(s21 + s21)/n]. Primeiramente vamos criar um vetor para cada um dos grupos de valores de duração, um diante de oxítona (durox), outro diante de paroxítona (durparox):

durparox <- c(161, 174, 166, 170, 147, 175, 177,185)

durox <- c(166, 143, 152, 123, 145, 154, 176, 149)

Calculemos então o valor de *texp* no R, fazendo:

texp <- (mean(durox)-mean(durparox))/sqrt((var(durox)+var(durparox))/8)

O valor é -2.653252. Como o teste é bicaudal (Ha: μ1 ≠ μ2), o valor da probabilidade abaixo deve ser duplicado. O número de graus de liberdade é (8-1)+(8-1) = 14.

p <- 2\*pt (texp, 14)

O resultado, 0.01890868, é a probabilidade, de admitindo a hipótese nula, de os dois grupos terem a mesma média. Como é menor que o nível de significância, rejeita-se H0: os dois grupos são vinculados a populações distintas. A probabilidade de erro é o valor de p.

O mesmo se obtém usando a função qnorm, que informa o intervalo de confiança dos valores individuais, mas para tanto, como é bicaudal, deve-se dividir

limiar <- qt(0.01, 14)

O resultado é **-2.624494**. Como a distribuição é simétrica, os valores de t a 98% de confiança estão entre esse valor e seu valor positivo: **2.624494**. Como *texp* está fora desse intervalo, rejeita-se a hipótese nula: a duração da sílaba [do] é maior diante da palavra “xucro”, por proximidade do acento lexical da mesma.

IMPORTANTE: O mesmo resultado pode ainda ser obtido com o teste de hipóteses t, no R:

t.test(durox,durparox,var.equal=T)

A expressão *var.equal=T* informa que as variâncias são estatisticamente iguais (isto é, aceita-se a hipótese nula num teste F). O teste bicaudal é o default. Para unicaudais escreva: alternative = "less" ou alternative ="greater", respecitvamente para Ha: μ1 < μ2 e Ha: μ1 > μ2). Se as variâncias não forem iguais, omita a expressão acima.

Observe abaixo o resultado da execução do teste acima em termos de informação: as variáveis, o valor de t, o número de graus de liberdade, o valor de p. Qual a hipótese alternativa, o intervalor de confiança das diferenças de média, e as próprias médias.

data: durox and durparox

t = -2.6533, df = 14, p-value = 0.01891

alternative hypothesis: true difference in means is not equal to 0

95 percent confidence interval:

-33.228643 -3.521357

sample estimates:

mean of x mean of y

151.000 169.375