Data simulation seminar Ian Hussey

Andreas Szukics

2024-02-21

What is the distribution of the p value under the null hypothesis? Normal distribution? -> answer in a simulation

```r
replicate(# 3. repeat 1 & 2 many times ('iterations')
  n= 10000,
  expr = t.test( # 2. analyse data with a statistical method
    x = rnorm(n=50, mean = 0, sd = 1), # 1. generate pseudo-random data set with known properties
    y = rnorm(n = 50, mean = 0, sd = 1)
    )$p.value
) |>
  hist(main = "Distribution of p values under the null hypothesis",
       xlab = "p") # 4. collect and aggregate results across iterations
```


Distribution of p values under the null hypothesis