

Homework 2: Hypothesis Test

2151294 马威

目录

1	step 0: 环境准备	1
2	step 1: 穿孔数目的地区差异	2
3	step 2: 穿孔数目差异检验	2
4	step 3: 纹身差异检验	3
5	step 4: 海盗上的大学与眼罩佩戴情况关系检验	3

海盗一直是一个神秘的身份，人们对于海盗的印象大多停留在影视作品上。这种神秘也引发了一些刻板印象，例如海盗们非常喜欢在身上穿洞洞，这对于海盗是一种时尚宣言。可有人认为欧洲海盗比美国海盗会有更多的身体穿孔，诸如此类的刻板印象还有很多。但幸好有研究者收集到了关于海盗的一些数据，我们可以通过分析来回答关于海盗的种种迷思。

1 step 0: 环境准备

```
# 装载必要的包
library(ggplot2)

# 清除当前镜像中的数据
rm(list = ls())

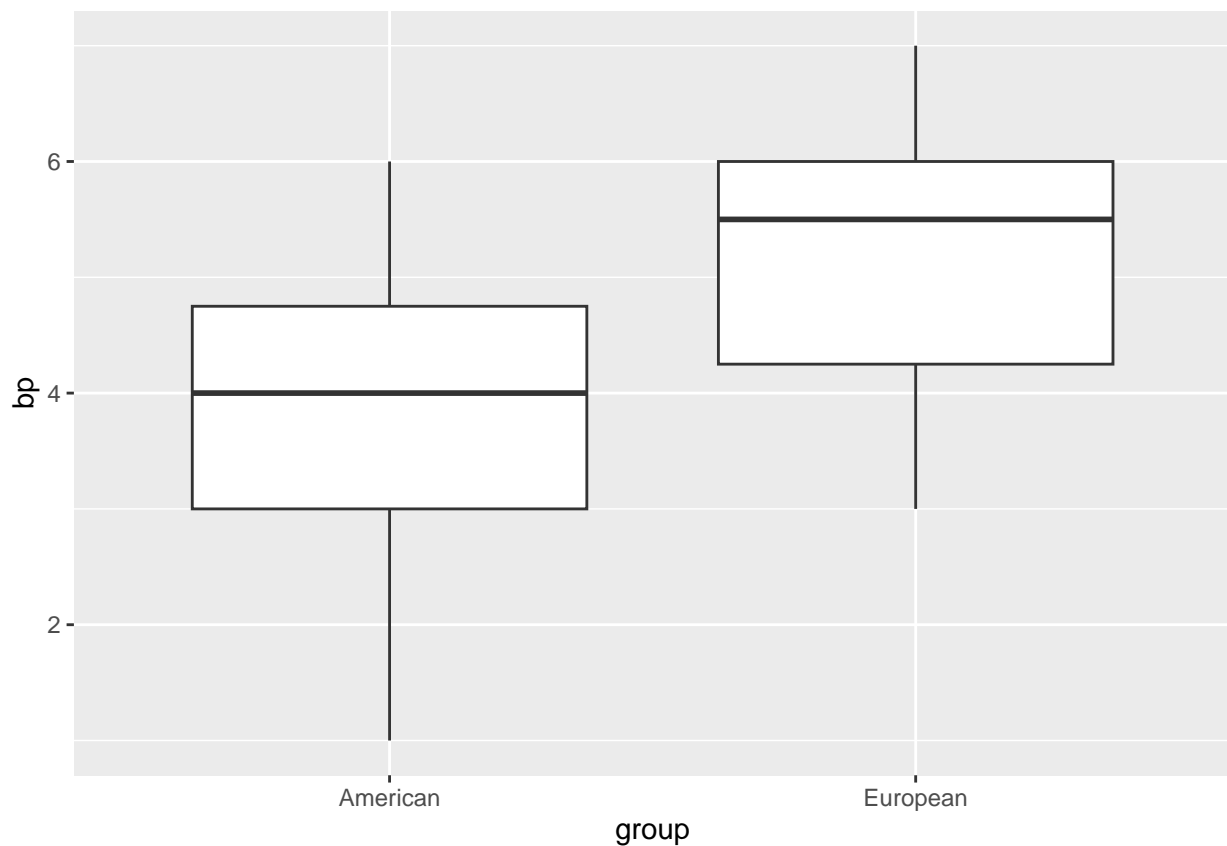
# 从镜像中装载所需数据
load("1 Hypothesis_ pirates.RData")

# Body piercing data
american.bp <- c(3, 5, 2, 1, 4, 4, 6, 3, 5, 4)
european.bp <- c(6, 5, 7, 7, 6, 3, 4, 6, 5, 4)
```

```
# Store data in a dataframe
bp.survey <- data.frame("bp" = c(american.bp, european.bp),
                        "group" = rep(c("American", "European"), each = 10),
                        stringsAsFactors = FALSE)
```

2 step 1: 穿孔数目的地区差异

```
# 根据数据框画出两个地区海盗穿孔数的箱线图
ggplot(bp.survey, aes(x = group, y = bp)) + geom_boxplot()
```



可以看出，欧洲海盗穿孔数总体上明显高于美国海盗

3 step 2: 穿孔数目差异检验

```
# 使用 t 检验判断美国和欧洲海盗穿孔数目是否有差异
# H0: 无差异
# H1: 有差异
```

```
p.test <- t.test(american.bp, european.bp, alternative = "two.sided")
cat(p.test$p.value)
```

```
## 0.02140251
```

p 值小于 0.05, 因此在 0.95 的置信水平下, 拒绝 H_0 , 即认为美国和欧洲海盗穿孔数目有差异

4 step 3: 纹身差异检验

取 29 和 30 岁海盗的纹身情况

```
age29.tattoos <- pirates[pirates$age == 29,][["tattoos"]]
age30.tattoos <- pirates[pirates$age == 30,][["tattoos"]]
```

使用 t 检验判断 29 岁和 30 岁海盗的纹身情况是否有差异

H_0 : 无差异

H_1 : 有差异

```
t.test <- t.test(age29.tattoos, age30.tattoos, alternative = "two.sided")
cat(t.test$p.value)
```

```
## 0.7910686
```

p 值大于 0.05, 因此在 0.95 的置信水平下, 无法拒绝 H_0 , 即认为 29 岁和 30 岁海盗的纹身情况无差异

5 step 4: 海盗上的大学与眼罩佩戴情况关系检验

使用卡方检验判断海盗上的大学与眼罩佩戴情况是否有关联

H_0 : 无关联

H_1 : 有关联

```
c.test <- chisq.test(pirates$college, pirates$eyepatch)
cat(c.test$p.value)
```

```
## 1
```

p 值明显大于 0.05, 因此在 0.95 的置信水平下, 无法拒绝 H_0 , 即认为海盗上的大学与眼罩佩戴情况无关联