MIDTERN

TzuhuiLi

2024-11-08

Table of contents

2.1 單樣本檢定

介紹

檢驗單個樣本的中位數或分佈是否等於某個指定值。

適用情境：1. 資料不符合常態分布。2. 樣本數量有限。

符號檢定 (Sign Test)

```r
# 定義學生成績
Scores <- c(78, 82, 74, 69, 88, 92, 81, 76, 84, 73, 77, 85, 80,
            79, 71, 90, 83, 75, 70, 86)

# 創建標記向量
marks <- ifelse(Scores > 80, "-", ifelse(Scores <80,"+","tie"))
# 成績大於 80 的標記為"-"
# 小於等於 80 的標記為"+"

# 將成績和標記合併為資料框
result <- data.frame(Scores, marks)
# 將資料框轉置
transposed_result <- as.data.frame(t(result))
```

```r
# 添加行名
rownames(transposed_result) <- c("Scores", "Marks")
colnames(transposed_result) <- c(1:20)

library(knitr)
```

```
Warning: package 'knitr' was built under R version 4.4.1
```

```r
kable(transposed_result)
```

	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
Scores	78	82	74	69	88	92	81	76	84	73	77	85	80	79	71	90	83	75	70	86
Marks	+	-	+	+	-	-	-	+	-	+	+	-	tie	+	+	-	-	+	+	-

```r
# 進行符號檢定
library(DescTools)
```

```
Warning: package 'DescTools' was built under R version 4.4.1
```

```r
sign_test <- SignTest(Scores , mu = 80)
print(sign_test)
```

```
    One-sample Sign-Test

data:  Scores
S = 9, number of differences = 19, p-value = 1
alternative hypothesis: true median is not equal to 80
95.9 percent confidence interval:
 75 84
sample estimates:
median of the differences
                    79.5
```

Wilcoxon 符號秩檢定

```r
# 計算與中位數（80）的絕對差
Scores_abs <- abs(Scores - 80)

# 將結果存入資料框
results <- data.frame(Score = Scores, Difference = Scores_abs)

# 去掉絕對差為 0 的資料
results <- results[results$Difference != 0, ]
```

```r
# 排名,並為小於中位數的加負號
results$Rank <- rank(results$Difference)
results$SignedRank <- ifelse(results$Score < 80, -results$Rank, results$Rank)

# 計算加總正符號
sum_positive_ranks <- sum(results$SignedRank[results$SignedRank > 0])
transposed_results <- as.data.frame(t(results))
rownames(transposed_results) <- c("Scores", "Zi","Rank","Ri")



# 顯示結果表格
library(knitr)
library(kableExtra)
```

Warning: package 'kableExtra' was built under R version 4.4.1

```r
kable(transposed_results, caption = "Wilcoxon Signed Rank Test Results"
      , full_width = TRUE)
```

Table 2: Wilcoxon Signed Rank Test Results

	1	2	3	4	5	6	7	8	9	10	11	12	14	15	16	17	18	19	20
Scores	78.0	82.0	74.0	69	88	92	81.0	76.0	84.0	73	77.0	85.0	79.0	71	90.0	83.0	75.0	70.0	86.0
Zi	2.0	2.0	6.0	11	8	12	1.0	4.0	4.0	7	3.0	5.0	1.0	9	10.0	3.0	5.0	10.0	6.0
Rank	3.5	3.5	11.5	18	14	19	1.5	7.5	7.5	13	5.5	9.5	1.5	15	16.5	5.5	9.5	16.5	11.5
Ri	-	3.5	-	-	14	19	1.5	-	7.5	-	-	9.5	-	-	16.5	5.5	-	-	11.5

```r
# 顯示加總正符號
cat(" 加總正符號的值為:", sum_positive_ranks)
```

加總正符號的值為: 88.5

```r
wilcox_test<- wilcox.test(Scores,mu = 80)
```

Warning in wilcox.test.default(Scores, mu = 80): cannot compute exact p-value
with ties

Warning in wilcox.test.default(Scores, mu = 80): cannot compute exact p-value
with zeroes

```r
print(wilcox_test)
```

```
	Wilcoxon signed rank test with continuity correction

data:  Scores
V = 88.5, p-value = 0.8091
alternative hypothesis: true location is not equal to 80
```

Kolmogorov-Smirnov 檢定

```r
ks_test <- ks.test(Scores , "pnorm", mean = mean(Scores), sd = sd(Scores))
print(ks_test)
```

```
    Exact one-sample Kolmogorov-Smirnov test

data:  Scores
D = 0.058642, p-value = 1
alternative hypothesis: two-sided
```

Run test

```r
Temperatures = c(15.2, 15.5, 15.3, 15.6, 15.4, 15.7, 15.6, 15.8,
                 15.9, 16.0, 15.8, 15.7, 15.5, 15.3, 15.6, 15.8,
                 16.1, 16.2, 16.3, 16.4)

library(tseries)
```

```
Warning: package 'tseries' was built under R version 4.4.1

Registered S3 method overwritten by 'quantmod':
  method            from
  as.zoo.data.frame zoo
```

```r
run_test <- runs.test(factor(Temperatures > median(Temperatures)))
print(run_test)
```

```
    Runs Test

data:  factor(Temperatures > median(Temperatures))
Standard Normal = -3.2042, p-value = 0.001355
alternative hypothesis: two.sided
```

Trend test

```r
library(randtests)
```

```
Attaching package: 'randtests'
```

The following object is masked from 'package:tseries':

    runs.test

```r
# 對 Wind 進行趨勢檢定
trend_test <- cox.stuart.test(Temperatures)
print(trend_test)
```

    Cox Stuart test

data: Temperatures
statistic = 9, n = 10, p-value = 0.02148
alternative hypothesis: non randomness