MIDTERN

TzuhuiLi

2024-11-12

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）

```
# 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))

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

library(knitr)

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	+	+	-	-	+	+	-

```
# 進行符號檢定
library(DescTools)
```

```
Warning: package 'DescTools' was built under R version 4.3.3
```

```
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

```
# 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, ]

# 排名，並為小於中位數的加負號
```

```r
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 4.3.3
```

```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	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

```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

```
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
```

Shapiro-Wilk Test

```
# 執行 Shapiro-Wilk 檢定
shapiro_test_result <- shapiro.test(Scores)
# 顯示檢定結果
print(shapiro_test_result)
```

```

    Shapiro-Wilk normality test

data:  Scores
W = 0.97828, p-value = 0.91
```

Kolmogorov-Smirnov 和 Shapiro-Wilk 檢定的檢定力比較

```
# 載入必要的套件
library(knitr)

# 定義檢定力比較函數
compare_power <- function(n, num_simulations = 1000, alpha = 0.05) {
  # 初始化結果儲存變量
  ks_power <- 0
  shapiro_power <- 0

  # 模擬過程
  for (i in 1:num_simulations) {
    # 生成符合非正態分布的資料（例如，指數分布）
    non_normal_data <- rexp(n, rate = 1)

    # Kolmogorov-Smirnov 檢定
    ks_test_non_normal <- ks.test(non_normal_data, "pnorm"
                              , mean(non_normal_data), sd(non_normal_data))

    # Shapiro-Wilk 檢定
    shapiro_test_non_normal <- shapiro.test(non_normal_data)

    # 判斷是否成功拒絕原假設
    if (ks_test_non_normal$p.value < alpha) ks_power <- ks_power + 1
    if (shapiro_test_non_normal$p.value < alpha) shapiro_power <- shapiro_power + 1
  }

  # 計算檢定力
  ks_power <- ks_power / num_simulations
  shapiro_power <- shapiro_power / num_simulations

  return(c(KS_Power = ks_power, Shapiro_Power = shapiro_power))
}

# 設定不同的樣本大小
sample_sizes <- c(20, 50, 100, 200,300)
results <- data.frame(Sample_Size = sample_sizes)

# 計算不同樣本大小的檢定力
for (n in sample_sizes) {
  power_result <- compare_power(n)
  results[results$Sample_Size == n, "KS_Power"] <- power_result["KS_Power"]
  results[results$Sample_Size == n, "Shapiro_Power"] <- power_result["Shapiro_Power"]
}

# 顯示結果表格
kable(results, caption = " 不同樣本大小下 Kolmogorov-Smirnov 和 Shapiro-Wilk 檢定的檢定力比較
      , digits = 3)
```

Table 3: 不同樣本大小下 Kolmogorov-Smirnov 和 Shapiro-Wilk 檢定的檢定力比較

Sample_Size	KS_Power	Shapiro_Power
20	0.044	0.826
50	0.350	1.000
100	0.907	1.000
200	1.000	1.000
300	1.000	1.000

Run test

```
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)
```

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

```
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

```
library(randtests)
```

```
Warning: package 'randtests' was built under R version 4.3.3


Attaching package: 'randtests'


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

    runs.test
```

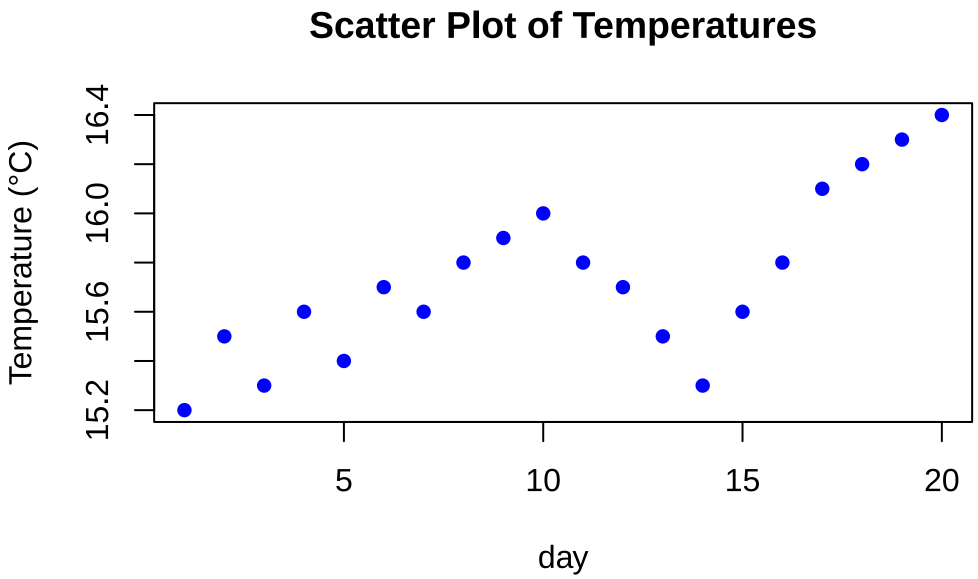
```
trend_test <- cox.stuart.test(Temperatures)
print(trend_test)
```

```

    Cox Stuart test

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

```
# 繪製散佈圖
plot(Temperatures,
     main = "Scatter Plot of Temperatures",
     xlab = "day",
     ylab = "Temperature (°C)",
     pch = 16,      # 使用實心點標記
     col = "blue") # 設置點的顏色
```


Scatter Plot of Temperatures
Temperature (°C)
15.2
15.6
16.0
16.4
5
10
15
20
day