無母數統計

廖芷萱、李敏榕、詹雅鈞、謝沛恩、李姿慧

Group 5

內容

Weekly planner
Monday
Tuesday
Sunday
To do
CLAIM THE life—
that YOU DESIRE

1.無母數統計的基本概念

李敏榕

有母數 vs 無母數

	有母數	無母數
假設	須對母體分配做假設	不須對母體的分配做假設
數據類型	適用於連續型數據(量的資料)	適用於類別型數據（質的資料）
適用情境	當樣本量大且數據近似常態分佈	當樣本量小、數據分佈不明或有極端值
計算方法	需要計算平均值和標準差等參數	依賴於排序和排名的計算(rank)
解釋	統計結果通常可以解釋為樣本估計的母體參數	統計結果通常是基於排名或序列解釋較為間接

優點

1. 沒有分配的假設
2. 小樣本仍然適用
3. 對測量誤差較不敏感（只考慮rank，對outlier較不敏感）

缺點

1. 母體分布已知時，與有母數統計相比，其統計分析效益較差
2. 較難計算信賴區間
3. 若許多測量值相等(tie)，容易造成variance被高估

無母數統計

有母數統計	無母數統計
Independent Sample t-test	Mann-Whitney U test
Paired Sample t-test	Wilcoxon signed rank test
One-way ANOVA	Kruskal-Wallis test
Pearson Correlation	Spearman's ρ, Kendall's τ

Independent Sample t-test vs Mann-Whitney U test

	Independent Sample t-test	Mann-Whitney U test
purpose	whether the two groups come from the same population	
null hypothesis	$\mu_1 = \mu_2$	$M_1 = M_2$
data	follows normal distribution	does not follow normal distribution

How to test normal distribution?

1. Kolmogorov-Smirnov Test
2. Shapiro-Wilks Test

One-way ANOVA vs Kruskal-Wallis test

	One-way ANOVA	Kruskal-Wallis test
purpose	whether the groups come from the same population	
null hypothesis	$\mu_1 = \mu_2 = ... = \mu_n$	$M_1 = M_2 = ... = M_n$
data	follows normal distribution	does not follow normal distribution

單組 -> Mann-Whitney U test

多組 -> Kruskal-Wallis test

Paired Sample t-test vs Wilcoxon signed rank test

	Paired Sample t-test	Wilcoxon signed rank test
purpose	whether there is a significant difference between two related datasets	
null hypothesis	$\mu_d = 0$	M = specific value
data	follows normal distribution	does not follow normal distribution
caculate	one set of data subtracted from the other	ranks assigned to the data

Pearson Correlation vs Spearman's ϱ, Kendall's τ

	Pearson Correlation	Spearman's ρ	Kendall's τ
purpose	whether there is a significant difference between two related datasets		
null hypothesis	$r = 0$	$\rho=0$	$\tau = 0$
data	interval scale or ratio scale	ordinal scale	

Which should we use?

1. There is more than one method of statistical analysis

2. Parametric methods may lead to significant results in some cases, while nonparametric methods may result in more significant results

—> Nonparametric methods is always valid, but not always efficient
Parametric methods is always efficient, but not always valid

1 Reference

1. https://www.researchgate.net/publication/228457648
2. https://www.researchgate.net/publication/289442433
3. Banda Gerald, Tailoka Frank Patson. Parametric and Nonparametric Tests: A Brief Review. International Journal of Statistical Distributions and Applications. Vol. 7, No. 3, 2021, pp. 78-82. doi: 10.11648/j.ijsd.20210703.12
4. https://ekja.org/journal/view.php?doi=10.4097/kjae.2016.69.1.8
5. Tutorials in Quantitative Methods for Psychology 2008, vol. 4(1), p. 13-20.
6. https://www.researchgate.net/publication/274635625
7. https://www.researchgate.net/publication/329698415
8. https://pmc.ncbi.nlm.nih.gov/articles/PMC4754273/

2. 無母數檢定方法

返回目錄頁

2. 無母數檢定方法的分類

主要以單樣本、雙樣本及多樣本檢定分類做介紹

2.1 單樣本檢定

2.2 雙樣本檢定

2.3 多樣本檢定

2.4 類別資料檢定

2.1 單樣本檢定

李姿慧

返回目錄頁

2.1 單樣本檢定

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：

- 資料不符合常態分布。
- 樣本數量有限。

1. One Sample Sign Test
2. One Sample Wilcoxon Signed-Rank Test
3. Kolmogorov-Smirnov Test
4. Shapiro-Wilk Test
5. Randomness Test / Run Test
6. Trend Test / Cox and Stuart Test

2.1 單樣本檢定

One Sample Sign Test

定義：檢驗樣本中位數是否等於某個特定值。

原理：根據樣本中每個觀測值與中位數的差異符號進行檢定。

優點：對資料的數值大小不敏感，只需關注正負號，簡單易用。

缺點：樣本量小時，檢定力相對較低。

$$H_0 : M_0 = m_0$$

$$H_1 : M_0 \neq m_0$$

2.1 單樣本檢定

One Sample Sign Test

H0: M0=80

```
# 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)
```

library(DescTools)

```
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
```

程式碼是將Xi >M0 設為"+"

p-value>0.05,不拒絕H0
沒有顯著證據證明
學生成績中位數不是80分

One Sample Sign Test

例子:檢驗20位學生成績中位數是否是80分

H0: M0=80

	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	+	+	-	-	+	+	-

T = number of "+" =10

n = 20 - 1 =19

通過查表 $\alpha = 0.05$

$$P(Y \leq 5) = 0.032 \approx 0.025$$

因為 $T > 5$ ，不拒絕H0。

20位學生成績中位數可能是80分。

One Sample Sign Test

Test Statistic:
Let T be the number of positive signs in the sample.

$$T = \sum_{i=1}^{n} I(X_i < m_0)$$

where I is the indicator function. The null distribution of T is Bin (n, $p = \frac{1}{2}$).

$$n = n' - \text{numbers of ''tie''}$$

where n' is the sample size.
Let

$$Y \sim \text{Bin}(n, p = \frac{1}{2})$$

To find r

$$P(Y \leq r) = \alpha_1 \approx \frac{\alpha}{2}$$

when n >25

$$r = \frac{1}{2}\left(n + z_{\alpha/2} \cdot \sqrt{n}\right)$$

Decision Rule:
If $\boldsymbol{T}$ is less than or equal to $\boldsymbol{r}$ or is greater than or equal to $\boldsymbol{n-r}$, we will reject H0.

2.1 單樣本檢定

$$H_0 : M_0 = m_0$$

$$H_1 : M_0 \neq m_0$$

One Sample Wilcoxon Signed-Rank Test

目的：檢驗樣本的中位數是否等於某個特定值。

原理：根據樣本中每個觀測值的秩進行檢定。不但考慮到差異值的正負號，也同時考慮到差異值的大小。

假設數據分布必須連續且對稱

優點：檢定力較Sign Test更強，適合數量較少的樣本。

缺點：若有過多相同值（ties）會影響檢定結果。

2.1 單樣本檢定

One Sample Wilcoxon Signed-Rank Test

例子:檢驗20位學生成績中位數是否是80分

H0: M0=80

	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

W = 88.5

通過查表臨界值是39，W>39 ，不拒絕H0。

20位學生成績中位數可能是80分。

	alpha values						
n	0.001	0.005	0.01	0.025	0.05	0.10	0.20
5	--	--	--	--	--	0	2
6	--	--	--	--	0	2	3
18	14	23	27	34	40	47	55
19	18	27	32	39	46	53	62
20	21	32	37	45	52	60	69

One Sample Wilcoxon Signed-Rank Test

library(DescTools)

例子:檢驗20位學生成績中位數是否是80分

H0: M0=80

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

```
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
```

p-value>0.05
沒有顯著證據證明
學生成績中位數不是80分

One Sample Wilcoxon Signed-Rank Test

Procedure:

1. Calculate the Differences

For each observation X_i, calculate the absolute difference from m_0:

$$Z_i = |X_i - m_0|$$

2. Exclude Zero Differences

Remove any pairs where $Z_i = 0$ to avoid ties.

3. Rank the Absolute Differences

Rank the non-zero values Z_i from smallest to largest, assigning each Z_i a rank R_i. For tied ranks, assign the average rank to each tied value.

4. Assign Signs to Ranks

Assign a positive sign to the ranks of positive differences $X_i > m_0$ and a negative sign to the ranks of negative differences $X_i < m_0$.

One Sample Wilcoxon Signed-Rank Test

5. Compute the Test Statistic W

Let W be the sum of the ranks with positive signs:

$$W = \sum_{X_i > m_0} R_i$$

6. Determine the Critical Value

Use the Wilcoxon signed-rank test critical value table or approximate the critical value for large n with the normal distribution.

Decision Rule:

Reject H_0 if W is less than or equal to the critical value (or, for a two-tailed test, if W is outside the acceptance region).

Kolmogorov-Smirnov Test

目的：檢驗樣本分布是否與特定分布一致。

原理：比較樣本的經驗累積分布函數（ECDF）與理論分布函數的差異。

優點：適合檢測樣本分布是否符合特定理論分布，能夠處理連續分布且無需假設正態分布。

缺點：樣本量小時檢定力較低，而樣本量過大時對小差異過於敏感，導致Type Ⅰ Error的增加。

Kolmogorov-Smirnov Test

H0:檢驗20位學生成績分布是否是常態分配

```
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
```

p-value=1>0.05
沒有顯著證據學生成
績不是常態分佈

Kolmogorov-Smirnov Test

Null Hypothesis:

The sample distribution does not significantly differ from a specified theoretical distribution:

$$H_0 : F(x) = F_0(x)$$

$$H_1 : F(x) \neq F_0(x)$$

where $F(x)$ is the cumulative distribution function (CDF) of the sample, and $F_0(x)$ is the CDF of the theoretical distribution.

Kolmogorov-Smirnov Test

Test Statistic:

Calculate the empirical cumulative distribution function $S(x)$:

$$S(x) = \frac{\text{number of } X_i\text{'s that are } \leq x}{n}, \quad x \in R$$

The test statistic D is then defined as:

$$D = \max|F_0(x) - S(x)|$$

Decision Rule:

Use the observed D value and compare it to the critical value obtained from the Kolmogorov-Smirnov table. Reject H_0 if D_{obs} exceeds the $(1-\alpha)$th quantile of the distribution of D at the desired significance level α.

Shapiro-Wilk Test

目的：檢驗樣本是否來自常態分佈。

原理：計算樣本數據的排列次序值和期望值的線性組合之間的關係。檢定統計量 W 越接近 1，樣本越可能是正態分布。

優點：對小樣本數據具有良好的檢定力，適合樣本量較小的資料（通常在 50 以下）。

缺點：僅適用於檢定常態分布。當樣本量過大（如超過 2000）時，也會對非常小的偏差產生顯著結果，增加Type 1 error。

Shapiro-Wilk Test

H0:檢驗20位學生成績分布是否是常態分配

```
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
```

p-value=0.91>0.05
沒有顯著證據學生成績不是常態分佈

Shapiro-Wilk Test

Test Statistic:

The Shapiro-Wilk test statistic W is calculated as:

$$W = \frac{\left(\sum_{i=1}^{k} a_i (X_{(n-i+1)} - X_{(i)})\right)^2}{\sum_{i=1}^{n} \left(X_i - \bar{X}\right)^2} \qquad \bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$$

where:

- (i) $k = n/2$ if n is even
- (ii) $k = (n\text{-}1)/2$ if n is odd, and $X_{(k+1)}$ is not involved in the calculation of W
- $X_{(i)}$ represents the i-th ordered sample value.
- a_i are the weights based on the sample size n and the assumption of normality. The values of a_i are tabulated in Shapiro-Wilk Table1 for n = 2,3,...,50.

Shapiro-Wilk Test

Decision Rule:

Given a significance level α, the α-th quantile W_α are tabulated in Shapiro-Wilk Table2. If the sample is drawn from a normal distribution, W should be close to 1. If W is smaller than the critical value W_α, reject the null hypothesis H_0 and conclude that the data does not follow a normal distribution.

The coefficients and critical values can be found in the Shapiro-Wilk table available at:

https://real-statistics.com/statistics-tables/shapiro-wilk-table/

For a data set with $n > 50$, we may use statistical software package.

檢定	作用	適用的樣本大小	選擇建議
Kolmogorov-Smirnov Test	檢定分布	50<n<1000 通常適合較大的樣本，在小樣本情況下，檢定力較弱且不太穩定。	適合用於需要檢測其他特定分布（如均勻分布、指數分布等）的情況，且常用於較大樣本的情境，但在常態性檢測的精確性上較不如Shapiro-Wilk Test。
Shapiro-Wilk Test	檢定是否是常態分佈	n<50 但在2000以內仍能發揮良好作用。 對於大樣本，檢定的靈敏度過高。	在小樣本檢測正態性時更具檢定力，因此更適合在樣本量不大或僅需檢測常態分布的情況。

Kolmogorov-Smirnov Test vs Shapiro-Wilk Test 檢定力比較

H0:資料是常態分佈
H1:資料不是常態分布

只要是檢定常態分佈，Shapiro-Wilk的檢定力都比Kolmogorov-Smirnov較高

Sample_Size	KS_Power	Shapiro_Power
20	0.052	0.826
50	0.350	0.999
100	0.920	1.000
200	1.000	1.000
300	1.000	1.000

Randomness Test / Run Test

目的：檢查數據序列是否具有隨機性，即是否是隨機分布的。

應用場景：適用於時間序列數據或至少是Ordinal資料，以判定資料序列中是否有顯著的模式或趨勢。

缺點：無法檢測數據的具體分布或模式，只能確認是否為隨機性。

H0:The sequence is random

Randomness Test / Run Test

library(tseries)

例子:某地區連續 20 天的日均氣溫(攝氏度)
H0: 日均氣溫數據具有隨機性,無明顯的系統性或規律性。

```
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)
```

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

如果資料是連續型資料,常用的方法是**使用資料的中位數**作為切分點。

```
	Runs Test

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

p-value<0.05
有顯著證據氣溫不是隨機序列

Randomness Test / Run Test

H0:The sequence is random

Procedure

1. Set a Threshold: Choose a threshold value (e.g., mean or median) to categorize the continuous variable into two groups.
2. Binarize the Data: Label each observation as ”success” (above threshold) or ”failure” (below threshold).
3. Count the Runs: Calculate the number of runs (U) in the binary sequence.
U :序列中成功和失敗的變化次數
$n1$:number of succes
$n2$:number of failure

Randomness Test / Run Test

H0:The sequence is random

Decision Rule:

Find the lower and upper critical values r_1 and r_2 from the table for a significance level $\alpha = 0.05$. Reject H_0 if $U \leq r_1$ or $U \geq r_2$.

When either n_1 or n_2 is greater than 20, use the test statistic:

$$Z = \frac{U - \frac{2n_1n_2}{n_1+n_2+1}}{\sqrt{\frac{2n_1n_2(2n_1n_2-n_1-n_2)}{(n_1+n_2)^2(n_1+n_2-1)}}} \sim N(0,1).$$

2.1 單樣本檢定

Trend Test / Cox and Stuart Test

目的：檢測時間序列資料中是否存在線性趨勢。

應用場景：時間序列資料的趨勢檢定，例如觀察隨時間變化的空氣品質、溫度或銷售數據等。

優點：專注於檢測數據中的線性趨勢，特別是環境變遷、經濟指標等資料。

缺點：僅能檢測線性趨勢，對於非線性變化的數據無法有效檢測。

H0: There is no trend in the data.

H1: There is either an upward or a downward trend.

Trend Test / Cox and Stuart Test

library(randtests)

H0: 日均氣溫沒有顯著趨勢，數據隨時間的變化是隨機的，
沒有明顯的上升或下降趨勢。

cox.stuart.test(x, alternative="two.sided")

```
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
```

p-value <0.05 拒絕H0
結論: 有明顯的上升或下降趨勢

Trend Test / Cox and Stuart Test

H0: There is no trend in the data.

Procedure

Divide the data into two halves. For each pair of observations (X_i, X_{i+c}), where $c = \frac{n'}{2}$ if n' is even, and $c = \frac{n'+1}{2}$ if n' is odd. Assign to each (X_i, X_{i+c}) a "+" symbol if $X_i < X_{i+c}$, a "-" symbol if $X_i > X_{i+c}$, and a "tie" symbol otherwise. Let n be the number of non-tied pairs of data.

Test Statistic:

$$T = \text{total number of } +\text{'s}$$

and the null distribution of T is Bin (n, $p = \frac{1}{2}$). The test hypothesis and the rest of the procedure is identical to that of the sign test.

檢定	作用	假設與限制
One Sample Sign Test	檢定中位數	只考慮了數據的符號（正或負）而忽略了數據的大小，因此檢定力較低。
One Sample Wilcoxon Signed-Rank Test	檢定中位數或平均數	假設數據分布必須連續且對稱，且樣本中的相同數值（ties）不可超過10%。
Run Test	檢定隨機性	適用於時間序列數據或至少是Ordinal資料。只適用於檢測隨機性，無法得知數據是否符合某種具體分布。
Trend Test	檢定趨勢	資料必須為時間序列資料，且僅能檢測線性趨勢，無法適用於非線性趨勢的檢測。

2 Reference

[1]ScienceDirect. (n.d.). Wilcoxon signed-rank test. Retrieved November 5, 2024, from https://www.sciencedirect.com/topics/mathematics/wilcoxon-signed-rank-test

[2]Wang, C. (2018, July 7). *醫學統計學*. https://bookdown.org/ccwang/medical_statistics6/

[3]OpenAI. (2024). *ChatGPT* [大型語言模型]. https://chat.openai.com

[4]Hsing-Ming Chang. (n.d.). *Chapter 2: Inference on p of BIN(n,p), Test and Confidence Interval for Median, Test for Randomness and Test for Trend* [未出版的講義]. NCKU.

[5]Hsing-Ming Chang. (n.d.). *Chapter 3: Test of Goodness of Fit (One Sample)* [未出版的講義].NCKU.

程式碼和檢定方法整理：

https://github.com/h24119055/Statistical-Consulting/blob/main/MIDTERN_coding.pdf

2.2 配對樣本與兩獨立樣本檢定

詹雅鈞

返回目錄頁

配對樣本或兩相關樣本

當兩樣本不是獨立，而是有成對兩變數(Xi, Yi) :

檢定成對兩變數的中位數是否相同:

- Wilcoxon signed ranks test
- Fisher's randomization analysis for matched pairs

配對樣本或兩相關樣本

Wilcoxon signed ranks test

- Wilcoxon signed ranks test用於比較**成對樣本**的**中位數**，適合樣本不符合常態分布或樣本數量較少時，通常用於測試相同樣本在不同條件下的差異

- 方法：計算每對觀測值的差異 *Di* = *Yi* - *Xi* 並排序 |*Di*| 後，對 *Di* 的正、負差異分別求和得到 *T+* 和 *T-*：

$$T^{+} = \sum_{\{i:D_i>0\}} \text{rank}\,(|\,D_i\,|) \quad ; \quad T^{-} = \sum_{\{i:D_i<0\}} \text{rank}\,(|\,D_i\,|)$$

再根據假設決定統計量 *T* 為 ***T+***（*H0: E[Di] ≥ 0*）、***T-***（*H0: E[Di] ≤ 0*）或 ***min(T+, T-)***（H0: E[Di] = 0）

配對樣本或兩相關樣本

Wilcoxon signed ranks test

- 範例資料集:包含10隻小鼠治療前後的體重，如右表
- 程式碼 wilcox_test()

```
#轉換成long data
mice.long <- mice2 %>%
  gather(key = "group", value = "weight", before, after)
#Wilcoxon signed ranks test
mice.long  %>% wilcox_test(weight ~ group, paired = TRUE )
```

➢ **comparisons**：一個包含成對比較組合的列表，用於指定想要比較的組，例如 comparisons = list(c("A", "B"), c("B", "C")) 用來比較 A、B 和 B、C

➢ **ref.group**：設定為某一組別，則該組別會作為參考組，其餘組比較，若設定 ref.group = "all" 則進行多組比較

➢ **paired**：設定paired =TRUE 進行配對檢定

output:

```
# A tibble: 1 × 7
  .y.    group1 group2    n1    n2 statistic       p
* <chr>  <chr>  <chr>  <int> <int>     <dbl>   <dbl>
1 weight after  before    10    10        55 0.00195
```

id	before	after
1	187.2	429.5
2	194.2	404.4
3	231.7	405.6
4	200.5	397.2
5	201.7	377.9
6	235	445.8
7	208.7	408.4
8	172.4	337
9	184.6	414.3
10	189.6	380.3

配對樣本或兩相關樣本

Wilcoxon signed ranks test

● p -value 僅表示結果是否具有統計顯著性，而不是效應的直接指標，所以可以用效果量（Effect Size）做為評估效果大小的指標

● wilcox_effsize()用Z值計算效應量 r，公式為： $r = \frac{z}{\sqrt{n}}$ n:樣本總數

通常對r的解釋：0.10 ~ 0.3 表示影響小，0.30 ~ 0.5 表示影響中等，r ≥ 0.5 表示影響大

```
mice2.long  %>% wilcox_effsize(weight ~ group, paired = TRUE)
```

output:

```
  .y.    group1 group2 effsize    n1    n2 magnitude
* <chr>  <chr>  <chr>    <dbl> <int> <int> <ord>
1 weight after  before   0.886    10    10 large
```

配對樣本或兩相關樣本

Fisher's randomization analysis for matched pairs

- 定義Di 為每對觀測值的差異，檢定中位數是否相同等同於檢定 $H_0:E[D] = 0$，在虛無假設$H_0: E[D_i] = 0$ ($M_1=M_2$) 成立的情況下，每對觀測值的差異 $D_i = Y_i - X_i$ 的正負號是隨機分配的

- 檢定步驟:

1. 計算每對觀測值的差異 $D_i = Y_i - X_i$
2. 檢定統計量 T 是所有正差異的總和： $T = \Sigma\, \delta_i * |D_i|$

   若 $D_i > 0$，$\delta_i = 1$，否則 $\delta_i = 0$
3. 忽略 Di 的原始正負號，考慮所有 |Di| 的 2^n 種可能的正負號分配方式 ，建立虛無假設為真的情況下統計量 T 的分佈

兩獨立樣本檢定

檢定中位數是否相同

- Mood's median test
- Mann Whitney U Test

檢定兩個獨立樣本之母體變異數是否相同

- Siegel-Tukey test
- Ansari-Bradley test

檢定兩個獨立樣本之分佈是否相同

- 兩樣本Kolmogorov-Smirnov Test
- wald-wolfowitz runs test

兩獨立樣本檢定

Mood's median test

- 用於檢定兩組數據的中位數是否相同，適合小樣本和具有許多的數據，非對稱分布、有極端值、許多ties值或等級尺度的資料

- 檢定步驟:
  1. 將兩個樣本的所有觀測值合併成一個樣本
  2. 計算合併樣本的中位數 M
  3. 根據每個觀測值是屬於樣本 1 還是樣本 2，以及其值是高於還是低於合併樣本的中位數， 建立 2x2 列聯表(如右)並利用此表進行卡方檢定

	母體 1	母體 2	列總和
大於M	$O11$	$O12$	$C1$
小於M	$O21$	$O22$	$C2$
總和	$n1$	$n2$	n

兩獨立樣本檢定

Mood's median test

- 範例：資料集有兩位講者與演講的得分，檢定兩位演講者的得分是否有差異

程式碼:

```
median_test(Likert ~ Speaker,
            data = Data,
            distribution="exact")
```

distribution="exact"：使用卡方精確分布

output:

```
      Exact Two-Sample Brown-Mood Median Test

data:  Likert by Speaker (Piglet, Pooh)
Z = -3.4871, p-value = 0.001093
alternative hypothesis: true mu is not equal to 0
```

Speaker	Likert
Pooh	3
Pooh	5
Pooh	4
Pooh	4
Pooh	4

⋮

Speaker	Likert
Piglet	2
Piglet	2
Piglet	3
Piglet	2

兩獨立樣本檢定

檢定兩獨立樣本之母體中位數

Mann Whitney U Test (Wilcoxon Rank Sum Test)

- 適合未知分布且小樣本，是t檢定的替代方法

檢定步驟:

1. 將兩個樣本合併並由數值到大小排序
2. 為合併樣本的每個觀測值標記排序，從1至 $n1+n2$ ，若出現相同值(ties)，則給予排序值平均
3. 計算其中一個樣本的排序和(Rank Sum)，並計算檢定統計量： $T=\sum_{i=1}^{n_1} R(X_i)-\frac{n_1(n_1+1)}{2}$

決策規則: 依據選定的顯著水準 α 查表得臨界值 w_α

檢定類型	H_0	H_a	決策
左尾檢定	$E[X]<E[Y]$	$F_X(x)>F_Y(x)$	若 $T\le w_\alpha$，拒絕 H_0
右尾檢定	$E[X]>E[Y]$	$F_X(x)<F_Y(x)$	若 $T\ge w_{1-\alpha}$，拒絕 H_0
雙尾檢定	$E[X]=E[Y]$	$F_X(x)\neq F_Y(x)$	若 $T\le w_{\alpha/2}$ 或 $T\ge w_{1-\alpha/2}$，拒絕 H_0

兩獨立樣本檢定

Mann Whitney U Test (Wilcoxon Rank Sum Test)

- 若 n_1 或 n_2 大於20，可以用近似服從標準常態分佈 ：

$$Z=\frac{T-\frac{n_1n_2}{2}}{\sqrt{\frac{n_1n_2(n_1+n_2+1)}{12}}}$$

數據中有相同值時，用修正因子調整檢定統計量 Z ：

$$Z=\frac{T-\frac{n_1n_2}{2}}{\sqrt{\frac{n_1n_2(n_1+n_2+1)}{12}-\frac{n_1n_2\sum t^3-\sum t}{12(n_1+n_2)(n_1+n_2-1)}}}$$

t :有相同排序值的數量

- 範例:同上個資料集，檢定兩位演講者的得分是否有差異

```
wilcox.test(Likert ~ Speaker,
            data=Data, alternative = "two.sided")
```

output:

```
        Wilcoxon rank sum test with continuity correction

data:  Likert by Speaker
W = 5, p-value = 0.0004713
alternative hypothesis: true location shift is not equal to 0

警告訊息：
於 wilcox.test.default(x = DATA[[1L]], y = DATA[[2L]], ...) :
  無法精確計算帶連結的 p 值
```

此警告訊息出現通常是因為資料中存在ties值，影響了精確 p-value的計算

兩獨立樣本檢定

Siegel-Tukey test

- Siegel-Tukey test 假設兩組的中位數相等，虛無假設下檢定統計量的分佈與Wilcoxon rank-sum test相同，不同之處是給予排序的方式:

  Wilcoxon rank-sum排序: {1,2,3,....N}

  Siegel-Tukey排序:{1, 4, 5,..., 6, 3, 2}

  所以如果另一組有較小的變異，那該組Rank大多高於另一組，可應用於研究不同實驗組結果的穩定性或不同治療方法的波動情形

- 範例為PlantGrowth資料集中對照組與處理組1的植物重量

  檢定H0：兩組樣本具有相同的變異數

ctrl	**4.17**	**5.58**	**5.18**	**6.11**	**4.50**	**4.61**	**5.17**	**4.53**	**5.33**	**5.14**
trt1	**4.81**	**4.17**	**4.41**	**3.59**	**5.87**	**3.83**	**6.03**	**4.89**	**4.32**	**4.69**

兩獨立樣本檢定

Siegel-Tukey test

- 程式碼:
  使用siegel.tukey()

```
library(ANSM5)
# Siegel-Tukey
siegel.tukey(ctrl, trt1)
```

參數設定:

mean.shift：是否使用均值偏移(預設FALSE)

cont.corr：是否使用連續性校正(預設TRUE)，

output:

```
Siegel-Tukey test using median shift for group1 and group2

H0: samples have the same variance
H1: samples have different variances

Statistic for exact test:
114 (rank sum from group1), 96 (rank sum from group2)
59 (Mann-Whitney U from group1), 41 (Mann-Whitney U from group2)
Exact p-value: 0.52885
```

兩組的rank sum

根據 Siegel-Tukey排序計算的Mann-Whitney U 統計量

兩獨立樣本檢定

Ansari-Bradley test

- Ansari-Bradley 檢定的排序方式避免Siegel-Tukey排序方向造成的偏差，使結果更具穩健性

- 方法:對合併後的觀測值按大小排列，並將最小值和最大值分別給Rank 1，第二小和第二大值給Rank 2，以此類推，並計算檢定統計量T：

$$T=\sum_{i=1}^{n_x} R_i \qquad \text{（第一個樣本的等級總和）}$$

兩獨立樣本檢定

Ansari-Bradley test

- 程式碼:使用ansari.test()

```
ansari.test(ctrl, trt1)
```

參數說明:

exact：是否計算精確 p-value，小樣本建議設為 TRUE，設定FALSE則使用近似常態
未指定時根據樣本情況自動選擇

output:

檢定統計量

```
        Ansari-Bradley test

data:  ctrl and trt1
AB = 58.5, p-value = 0.5948
alternative hypothesis: true ratio of scales is not equal to 1

警告訊息：
於 ansari.test.default(ctrl, trt1) :  無法精確計算帶連結的 p 值
```

兩獨立樣本檢定

Two sample Kolmogorov-Smirnov Test

- 和單樣本 Kolmogorov-Smirnov 檢定的原理相似，主要區別在於比較的對象
- 合併樣本並排序後，為每個樣本分別計算經驗分布函數(EDF)後，找出兩個EDF間的**最大垂直距離**作為檢定統計量 D
- 範例：從iris資料集選取Setosa、Versicolor 兩個品種的資料，並比較它們Sepal.Length的分佈是否相同：

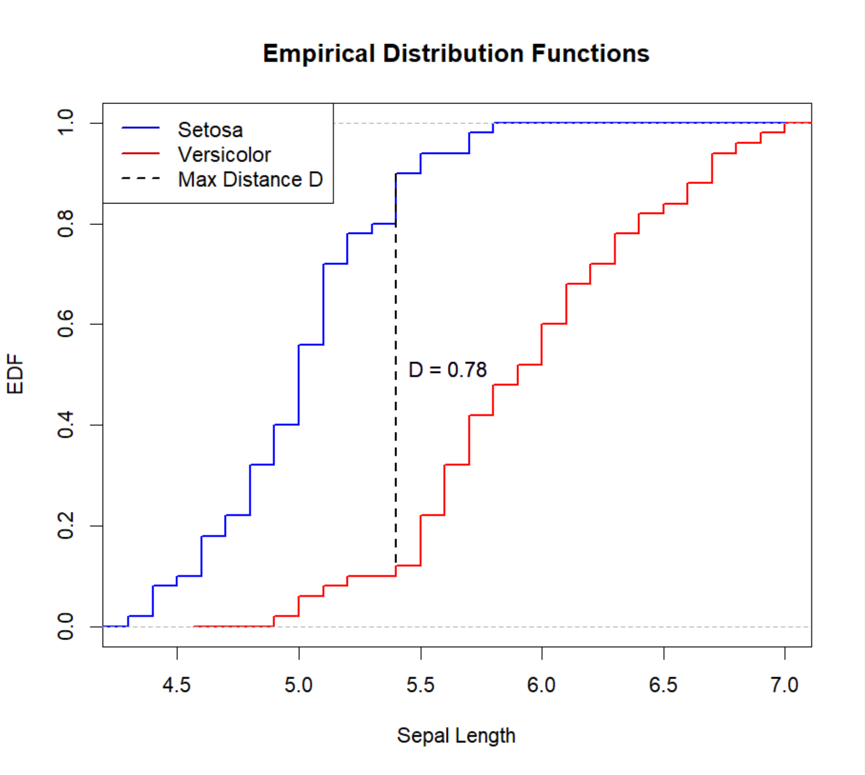
```
setosa <- iris$Sepal.Length[iris$Species == "setosa"]
versicolor <- iris$Sepal.Length[iris$Species == "versicolor"]

ks.test(setosa, versicolor)
```

兩獨立樣本檢定

Two sample Kolmogorov-Smirnov Test

- output:

```
        Exact two-sample Kolmogorov-Smirnov test

data:  setosa and versicolor
D = 0.78, p-value = 4.219e-15
alternative hypothesis: two-sided
```



兩獨立樣本檢定

Wald-wolfowitz runs test

- Wald-wolfowitz runs test透過計算合併樣本**標記變化的次數**，檢測兩組樣本在隨機性、位置和離散程度上的差異，以此檢定是否來自相同的分布

- 方法:
  將兩組樣本合併，按數值大小排序並為每個觀測值標記來源，並計算排序序列中的number of runs(標記變化的次數)
  例如：AABBABBA，number of run = 5 （AA, BB, A, BB, A）

  檢定統計量 U = number of run + 1

  當樣本數較小($n1 + n2 < 50$)時，依顯著水準 α 查表決定臨界值；樣本數較大時，使用近似標準常態計算 p-value

兩獨立樣本檢定

Wald-wolfowitz runs test

- 範例：從iris資料集選取Setosa、Versicolor 兩個品種的資料，並比較它們Sepal.Length的分布是否相同：

```
data_combined <- data.frame(
  value = c(setosa, versicolor),
  group = c(rep(1, length(setosa)), rep(0, length(versicolor)))
)

#依數值進行排序
data_sorted <- data_combined[order(data_combined$value), ]

#執行檢定
runs.test(data_sorted$group)
```

- output:

```
data:  data_sorted$group
statistic = -7.0356, runs = 16, n1 = 50, n2 = 50, n = 100, p-value = 1.984e-12
alternative hypothesis: nonrandomness
```

總整理

樣本類型	目的	檢定方法	條件與限制
配對樣本	檢定平均數或中位數是否有差異	Wilcoxon signed ranks test	假設差異值Di的分佈對稱
		Fisher's randomization analysis for matched pairs	
兩獨立樣本	檢定兩樣本是否來自相同的分佈	Two sample Kolmogorov-Smirnov Test	
		Wald-wolfowitz runs test	

總整理

樣本類型	目的	檢定方法	條件與限制
兩獨立樣本	檢定兩樣本是否有相同中位數	Mood's median test	
		Mann Whitney U Test	假設分佈形狀相同
	檢定兩樣本的離散程度是否有差異	Siegel-Tukey Test	假設兩組的中位數相等
		Ansari-Bradley Test	假設兩組的中位數相等

2 Reference

1. https://bookdown.org/xiangyun/data-analysis-in-action/common-statistical-tests.html#sec-two-samples
2. https://www.ibm.com/docs/zh-tw/spss-statistics/saas?topic=tests-two-independent-samples-test-types
3. https://rcompanion.org/handbook/F_05.html
4. https://www.kaggle.com/discussions/general/419726
5. https://rcompanion.org/handbook/F_04.html
6. https://www.datanovia.com/en/lessons/wilcoxon-test-in-r/
7. https://nelsonchiou.blogspot.com/2018/07/nonparametric-statistics-step-by-step_35.html
8. https://openpress.usask.ca/introtoappliedstatsforpsych/chapter/16-3-paired-sample-sign-test/
9. Hsing-Ming Chang. (n.d.). Chapter 4: Tests for Two Independent Samples
10. Hsing-Ming Chang. (n.d.). Chapter 5: Tests for Matched-Pairs or Two Related Samples
11. https://support.minitab.com/zh-cn/minitab/help-and-how-to/statistics/nonparametrics/how-to/mood-s-median-test/before-you-start/overview/

程式碼:

https://github.com/ChanYaJun/nonparametric-tests-for-paired-samples-and-two-independent-samples

2.3 多樣本檢定

謝沛恩

Kruskal-Wallis Test

目的：檢定多組獨立樣本的中位數或分佈是否完全相同。適用於非常態分配、小樣本、有順序的資料

假設檢定：H0:各組中位數皆相等

H1:至少有一個中位數不相等

2.3 多樣本檢定

Kruskal-Wallis Test

檢定步驟：

1. 排名：將所有樣本中的資料從小到大排序，給出排名順序。

   若重複的值，則排名取平均。

2. 計算排名總和：計算各組的排名值總和(Ri)

3. 檢定統計量：

$$H=\frac{12}{N(N+1)}\sum_{i=1}^{k}\frac{R_i^2}{n_i}-3(N+1)$$

H：服從卡方分配
N：總個數
k：組數
ni：樣本個數

4. 結論：當H值大於臨界值（或P值小於顯著水準）時，表示

   各組中位數不完全相同。

Kruskal-Wallis Test

（例子）：比較三個城市的工作滿意度是否相同

City A	7	8	5	5	7
City B	6	5	4	8	9
City C	6	5	8	9	8

假設檢定：

H_0：三個城市的工作滿意度相同$(M_{city\,A} = M_{city\,C} = M_{city\,C})$

H_1：三個城市的工作滿意度不完全相同(至少有一個中位數不相同)

2.3 多樣本檢定

Kruskal-Wallis Test

City A	7	8	5	5	7
City B	6	5	4	8	9
City C	6	5	8	9	8

（例子）：比較三個城市的工作滿意度

1. 排名：

Score	4	5	5	5	5	6	6	7	7	8	8	8	8	9	9
排名	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15
調整後	1	3.5	3.5	3.5	3.5	6.5	6.5	8.5	8.5	11.5	11.5	11.5	11.5	14.5	14.5

2. 計算排名總和：

$$R_1 = 8.5 + 11.5 + 3.5 + 3.5 + 8.5 = 35.5$$

$$R_2 = 6.5 + 3.5 + 1 + 11.5 + 14.5 = 37$$

$$R_3 = 6.5 + 3.5 + 11.5 + 14.5 + 11.5 = 47.5$$

2.3 多樣本檢定

City A	7	8	5	5	7
City B	6	5	4	8	9
City C	6	5	8	9	8

Kruskal-Wallis Test

(例子)：比較三個城市的工作滿意度

3. 檢定統計量：

$$H=\frac{12}{15(15+1)}\left(\frac{35.5^2}{5}+\frac{37^2}{5}+\frac{47.5^2}{5}\right)-3(15+1)=0.855$$

4. 結論：H<5.991，不拒絕虛無假設，表示三個城市的工作滿意度相同。

Kruskal-Wallis Test

（例子）：比較三個城市的工作滿意度

```
> # 建立數據集
> satisfaction <- data.frame(
+   score = c(7, 8, 5, 5, 7, 6, 5, 4, 8, 9, 6, 5, 8, 9, 8),
+   city = factor(c(rep("CityA", 5), rep("CityB", 5), rep("CityC", 5)))
+ )
> kruskal.test(score ~ city, data = satisfaction)

	Kruskal-Wallis rank sum test

data:  score by city
Kruskal-Wallis chi-squared = 0.89162, df = 2, p-value = 0.6403
```

結論：P-value>0.05，不拒絕虛無假設，表示三個城市的工作滿意度相同。

Levene's Test

目的：檢定多組獨立樣本的變異數是否完全相等。
適用於非常態分配、有異常值的資料

假設檢定：H0：每個變異數相等
H1：只少有一個變異數不相同

Levene's Test

檢定步驟：

1. 計算中心值：計算每個樣本的中位數。

2. 計算每個值與組內中心值的相差絕對值：

$$D_{ij} = \left|Y_{ij} - median_i\right|, Y_{ij}\text{為第 i 組中第 j 個觀察值}$$

3. 將相差絕對值進行單因子ANOVA檢定，以檢測不同組別之間的偏差值是否顯著不同。

4. 結論：當檢定統計量大於臨界值（或P值小於顯著水準）時，表示各組變異數有顯著差異。

Levene's Test

(例子)：有個資料有三家醫院的患者住院天數，來比較這些醫院之間患者住院天數的變異數是否相同

```
> # 不同醫院的住院天數
> hospital_A <- c(5, 7, 8, 6, 7)
> hospital_B <- c(6, 7, 8, 9, 7)
> hospital_C <- c(4, 6, 8, 5, 6)
```

假設檢定：

H_0：三間醫院的住院天數變異數相同。

H_1：三間醫院的住院天數變異數不完全相同。

2.3 多樣本檢定

Levene's Test

(例子)：有個資料有三家醫院的患者住院天數，來比較這些醫院之間患者住院天數的變異數是否相同

```
> # 將數據整理成 dataframe
> data <- data.frame(
+   days = c(hospital_A, hospital_B, hospital_C),
+   hospital = factor(rep(c("A", "B", "C"), each = 5))
+ )
> # 基於中位數的 Levene's Test
> library(car)
> leveneTest(days ~ hospital, data = data, center = median)
Levene's Test for Homogeneity of Variance (center = median)
      Df F value Pr(>F)
group  2  0.0833 0.9206
      12
```

結論：P-value>0.05，不拒絕虛無假設，表示三家醫院的患者住院天數變異數至少有一家不同。

Jonckheere-Terpstra Test

目的： 檢定多組獨立樣本的趨勢變化。
適用於未知分配(非常態分配)、有順序的資料

假設檢定： H0：各組的分布隨組別的變化無趨勢
H1：各組的分布隨組別的變化有增加或下降的趨勢(雙尾)

Jonckheere-Terpstra Test

檢定步驟：

1. 排序組別：按照變數的分組進行排序。

2. 比較每組的每個觀測值：兩兩作比對，比較每組的每個觀測值。若數值在預期方向上符合，則增加一個正的計數，否則為負。

3. 檢定統計量J值：計算所有符合趨勢的數量，將其作為檢定統計量，J值越高，越能支持存在趨勢變化的假設。

4. 結論：根據檢定統計量J和樣本數，計算檢定的P-value。如果P-value小於顯著性水準，拒絕虛無假設，表示資料有顯著增加或下降的趨勢。

Jonckheere-Terpstra Test

(例子)：檢定不同治療強度（輕度、中度、重度）之間的住院天數是否有變化的趨勢(增加或減少)。

```
> # 不同治療強度的住院天數
> mild <- c(5, 4, 6, 5, 6)
> moderate <- c(6, 7, 6, 7, 8)
> severe <- c(8, 9, 8, 10, 11)
```

假設檢定：

H_0：不同治療強度之間的住院天數無增加或減少的趨勢。

H_1：不同治療強度之間的住院天數有增加或減少的趨勢。

Jonckheere-Terpstra Test

（例子）：檢定不同治療強度（輕度、中度、重度）之間的住院天數是否有變化的趨勢(增加或減少)。

```
> # 整理數據
> Data <- data.frame(
+   days = c(mild, moderate, severe),
+   treatment = factor(rep(c("Mild", "Moderate", "Severe"), each = 5), or
dered = TRUE)
+ )
> # 進行 Jonckheere-Terpstra 檢定
> library(DescTools)
> JonckheereTerpstraTest(days ~ treatment, data = Data)

        Jonckheere-Terpstra test

data:  days by treatment
JT = 72, p-value = 0.0002673
alternative hypothesis: two.sided
```

單尾：alternative = "increasing"
alternative = "decreasing"

結論：P-value<0.05，拒絕虛無假設，表示不同治療強度的住院天數有顯著增加或減少的趨勢。

Friedman Test

目的：檢定多組配對樣本的中位數是否完全相等。

適用於非常態分配、小樣本、有順序的資料

假設檢定：H0:各組中位數皆相等

H1:至少有一個中位數不相等

2.3 多樣本檢定

Friedman Test

檢定步驟：

1. 排名：對每個個體的觀察值在每個條件內進行排序，並給出排名值。

2. 計算排名總和：將每個條件的排名值加總(Ri)。

3. 檢定統計量：

$$F_r = \frac{12}{nk(k+1)}\left(\sum_{j=1}^{k} R_j^2\right) - 3n\,(k+1)$$

Fr：服從卡方分配
n：樣本數
k：組數

4. 結論：當H值大於臨界值（或P值小於顯著水準）時，表示各組中位數有顯著差異。

2.3 多樣本檢定

Friedman Test

（例子）：比較學生們的三次考試成績之間是否有差異

	Student 1	Student 2	Student 3	Student 4	Student 5
Exam1	76	82	90	85	78
Exam2	80	84	89	87	82
Exam3	78	88	92	84	79

假設檢定：

H_0：學生的三次考試成績之間無差異

H_1：學生的三次考試成績之間有差異

2.3 多樣本檢定

	Student 1	Student 2	Student 3	Student 4	Student 5
Exam1	76	82	90	85	78
Exam2	80	84	89	87	82
Exam3	78	88	92	84	79

Friedman Test

(例子)：比較學生們的三次考試成績之間是否有差異

1. 排名：

	Student 1	Student 2	Student 3	Student 4	Student 5
Exam1	1	1	2	2	1
Exam2	3	2	1	3	3
Exam3	2	3	3	1	2

2. 計算排名總和：

$$R_1 = 1 + 1 + 2 + 2 + 1 = 7$$
$$R_2 = 3 + 2 + 1 + 3 + 3 = 12$$
$$R_3 = 2 + 3 + 3 + 1 + 2 = 11$$

2.3 多樣本檢定

Friedman Test

	Student 1	Student 2	Student 3	Student 4	Student 5
Exam1	76	82	90	85	78
Exam2	80	84	89	87	82
Exam3	78	88	92	84	79

(例子)：比較學生們的三次考試成績之間是否有差異

3. 檢定統計量：

$$F = \frac{12}{(5)(3)(3+1)}(7^2 + 12^2 + 11^2) - (3)(5)(3+1) = 2.8$$

4. 結論：F<5.991，不拒絕虛無假設，表示學生的三次考試成績之間無顯著差異。

Friedman Test

（例子）：比較學生們的三次考試成績之間是否有差異

```
> # Friedman 檢定
> friedman.test(score ~ exam | student, data = scores_long)

        Friedman rank sum test

data:  score and exam and student
Friedman chi-squared = 2.8, df = 2, p-value = 0.2466
```

結論：P-value>0.05，不拒絕虛無假設，表示學生的三次考試成績之間無顯著差異。

2.3 多樣本檢定

檢定	作用	條件與限制
Kruskal-Wallis Test	檢定中位數或分布	適用連續或有順序資料。 獨立樣本 分布形狀應大致相同
Levene's Test	檢定變異數	適用連續資料。 獨立樣本 假設分布形狀相同
Jonckheere-Terpstra Test	檢定趨勢	適用連續或有順序的資料。 獨立樣本
Friedman Test	檢定中位數	適用連續或有順序的資料。 配對樣本

2.4 類別資料的檢定

謝沛恩

2.4 類別資料的檢定

卡方檢定

目的：1. 檢定兩類別變數是否獨立。(關聯性檢定)

2. 檢定兩個或兩個以上的母體某一特性的分配(比例)是否相同。(適合度檢定)

條件要求：1. 一般要求每個期望頻數至少為5（特別是2x2的表格），否則檢定結果可能不準確。

2. 適用於樣本量較大、資料分布未知或偏離常態分布的情況。

卡方檢定

檢定步驟：

1. 建立列聯表：將變數的各類別個數建立成列聯表。

2. 計算期望頻率數：根據行和列的總數計算期望頻率數。 $E_{ij} = \frac{R_{ij} \times C_{ij}}{N}$

3. 檢定統計量： $\chi^2 = \sum \frac{(O_{ij} - E_{ij})^2}{E_{ij}} \sim {\chi^2}_{\alpha,((r-1)(c-1))}$

r：列數
c：行數

4. 結論：當檢定統計量大於臨界值（或p值小於顯著水準）時，表示各組間有顯著相關。

卡方檢定

（例子）：檢定性別和抽菸是否有相關

建立列聯表：

性別/抽菸	有	無	
男	**30**	**10**	40
女	**20**	**40**	60
	50	50	100

假設檢定：

H_0：性別和吸菸無關

H_1：性別和吸菸有關

2.4 類別資料的檢定

卡方檢定

（例子）：檢定性別和抽菸是否有相關

計算期望值：

性別/抽菸	有	無	
男	**30** (20)	**10** (20)	40
女	**20** (30)	**40** (30)	60
	50	50	100

檢定統計量：

$$\chi^2 = \sum \frac{(O_{ij} - E_{ij})^2}{E_{ij}} = \frac{(30-20)^2}{20} + \frac{(10-20)^2}{20} + \frac{(20-30)^2}{30} + \frac{(40-30)^2}{30}$$

$$= 16.667$$

卡方檢定

（例子）：檢定性別和抽菸是否有相關

結論：檢定統計量>3.841，拒絕虛無假設，表示性別和抽菸有顯著相關（不獨立）。

卡方檢定

(例子)：檢定性別和抽菸是否有相關

```
> # 卡方檢定
> # 數據
> gender_smoke <- matrix(c(30, 10, 20, 40), nrow = 2, byrow = TRUE)
> dimnames(gender_smoke) <- list(Gender = c("Male", "Female"), Smoke = c
("Yes", "No"))
> chisq.test(gender_smoke)

        Pearson's Chi-squared test with Yates' continuity correction

data:  gender_smoke
X-squared = 15.042, df = 1, p-value = 0.0001052
```

結論：檢定統計量>3.841，P-value<0.05，拒絕虛無假設，表示性別和抽菸有顯著相關(不獨立)。

Fisher's exact Test

目的：檢定小樣本兩類別變數是否獨立。

條件要求：1. 適合樣本數較少，列聯表中期望值小於5的情況。

2. 沒有像卡方檢定那樣需要期望值限制，因此更適合小樣本的準確檢定。

Fisher's exact Test

檢定步驟：

1. 建立列聯表

2x2 列聯表	A	B	
變數 1	a	b	a+b
變數 2	c	d	c+d
	a+c	b+d	a+b+c+d

2. Fisher's exact機率公式：$P = \dfrac{\binom{a+b}{a} \cdot \binom{c+d}{c}}{\binom{N}{a+c}}$

3.結論：當P值小於顯著水準時，表示各組間有顯著相關。

Fisher's exact Test

（例子）：檢定性別和抽菸是否有相關(觀察個數較少)

建立列聯表：

性別/抽菸	有	無	
男	**2**	**7**	9
女	**8**	**5**	13
	10	12	22

假設檢定：

H_0：性別和吸菸無關

H_1：性別和吸菸有關

2.4 類別資料的檢定

性別/抽菸	有	無	
男	**2**	**7**	9
女	**8**	**5**	13
	10	12	22

Fisher’s exact Test

（例子）：檢定性別和抽菸是否有相關(觀察個數較少）

Fisher’s exact機率公式：

$$P=\frac{\binom{9}{2}\binom{13}{8}}{\binom{22}{10}}=\frac{36(1287)}{646646}=0.072$$

結論：P-value>0.05，不拒絕虛無假設，表示性別和抽菸無顯著相關(獨立）。

Fisher's exact Test

(例子)：檢定性別和抽菸是否有相關(觀察次數變少)

```
> # 小樣本數據
> small_sample <- matrix(c(2, 7, 8, 5), nrow = 2, byrow = TRUE)
> dimnames(small_sample) <- list(Gender = c("Male", "Female"), Smoke = c
("Yes", "No"))
> fisher.test(small_sample)

        Fisher's Exact Test for Count Data

data:  small_sample
p-value = 0.09907
alternative hypothesis: true odds ratio is not equal to 1
95 percent confidence interval:
 0.01410547 1.60171447
sample estimates:
odds ratio
  0.194317
```

結論：P-value>0.05，不拒絕虛無假設，表示性別和抽菸無顯著相關(獨立)。

2.4 類別資料的檢定

	卡方檢定	Fisher's exact 檢定
適用樣本量	較大樣本，期望值至少大於5	小樣本，特別適用於2x2或3x3的列聯表
期望值要求	期望值大於5	無限制
計算方式	根據觀察值與期望值之間的差異計算卡方統計量	根據精確概率計算P值

McNemar's Test

目的：適用於2x2配對二元資料，檢定配對樣本間比例是否顯著不同。

應用：比較前後變化是否有差異

假設檢定：

H_0：配對樣本在處理前後沒有變化(b = c)

H_1：配對樣本在處理前後有變化(b ≠ c)

McNemar’s Test

檢定步驟：

1. 建立列聯表

2x2 列聯表	類別 Y 是 (After Yes)	類別 Y 否 (After No)
類別 X 是 (Before Yes)	a	b
類別 X 否 (Before No)	c	d

2. 檢定統計量：

$$\chi^2 = \frac{(b-c)^2}{b+c}$$

3.結論：當檢定統計量大於臨界值時，表示配對樣本在處理前後有顯著改變。

McNemar’s Test

(例子)：檢定在新政策實施前後，對政策的支持情況是否有改變

建立列聯表：

2x2 列聯表	政策實施後支持	政策實施後不支持
政策實施前支持	20	10
政策實施前不支持	5	15

假設檢定：

H_0：在新政策實施前後，對政策的支持情況無改變

H_1：在新政策實施前後，對政策的支持情況有改變

McNemar's Test

（例子）：檢定在新政策實施前後，對政策的支持情況是否有改變

檢定統計量：

$$\chi^2 = \frac{(b-c)^2}{b+c} = \frac{(10-5)^2}{10+5} = 1.667$$

結論：檢定統計量<3.841，不拒絕虛無假設，表示在新政策實施前後，對政策的支持情況無顯著改變。

McNemar’s Test

（例子）：檢定在新政策實施前後，對政策的支持情況是否有改變

```
> # 建立 2x2 列聯表
> data <- matrix(c(20, 10, 5, 15), nrow = 2, byrow = TRUE)
> rownames(data) <- c("Before", "After")
> colnames(data) <- c("Support", "Not Support")
> # 進行 McNemar's Test
> mcnemar.test(data)

        McNemar's Chi-squared test with continuity correction

data:  data
McNemar's chi-squared = 1.0667, df = 1, p-value = 0.3017
```

結論：檢定統計量<3.841，P-value>0.05，不拒絕虛無假設，表示在新政策實施前後，對政策的支持情況無顯著改變。

2.4 類別資料的檢定

檢定	作用	條件與限制
卡方檢定	檢定變數獨立性或分布一致	適用類別資料 獨立樣本 期望值大於5
Fisher’s exact Test	檢定變數獨立性	適用類別資料、小樣本 獨立樣本
McNemar’s Test	檢定前後變化是否有差異	適用類別資料 配對樣本

2 Reference

https://dasanlin888.pixnet.net/blog/post/34467986

https://www.yongxi-stat.com/kruskal-wallis-h-test/

https://www.technologynetworks.com/informatics/articles/the-kruskal-wallis-test-370025

https://www.yongxi-stat.com/friedman-test/

https://www.ibm.com/docs/zh-tw/spss-statistics/beta?topic=tests-friedman-test

https://www.yongxi-stat.com/fishers-exact-test-r/

https://en.wikipedia.org/wiki/Levene%27s_test

https://statistics.laerd.com/spss-tutorials/jonckheere-terpstra-test-using-spss-statistics.php

程式碼
https://github.com/H24119021/STATISTICAL-CONSULTING/tree/Midterm

3 無母數估計方法

廖芷萱

3

3 無母數估計方法介紹

3.1 Correlation：

3.1.1 Spearman 等級相關係數

3.1.2 Kendall's Tau 相關係數

3.1.3 Theil-Sen 估計量

3.1.4 無母數迴歸

3.2 核密度估計（Kernel Density Estimation, KDE）

3

3.1 Correlation

廖芷萱

3.1.1

Spearman
等級相關係數

廖芷萱

3

Spearman 等級相關係數

目的：衡量兩個變數之間的單調關係的非參數統計方法。

*它不要求數據符合正態分布，適合於非線性關係或等級數據。

假設檢定：

H_0: X_i and Y_i are mutually independent

H_a: X_i and Y_i are either positively or negatively correlated

3

步驟：（假設資料為二元隨機變數 X 和 Y）

1. **將X和Y根據n對觀察值分配**
2. **將數據中的每個變量進行排名，若遇到相同值分配相同的平均排名**
3. **計算Spearman correlation (ρ)**

$$\rho = \frac{\sum_{i=1}^{n} R(X_i)R(Y_i) - n\left(\frac{n+1}{2}\right)^2}{\sqrt{\left[\sum_{i=1}^{n}[R(X_i)]^2 - n\left(\frac{n+1}{2}\right)^2\right]\left[\sum_{i=1}^{n}[R(Y_i)]^2 - n\left(\frac{n+1}{2}\right)^2\right]}}.$$

1. **根據計算結果/做鑒定，看兩個變數之間的單調關係**

~ 如果ρ接近 1 表示強正相關，接近 -1 表示強負相關，0 則表示無相關（但不意味著沒有關聯，可能存在複雜的非線性關係）。

~ 若 p 值小於 α，則說明相關性顯著，可以推斷出變數之間確實存在單調關係。

3

（實例） 1973 年 5 月到 9 月在美國紐約市（華盛頓廣場）所收集的空氣質量數據

```
> library(Hmisc)
> data(airquality)
> str(airquality)
'data.frame':   153 obs. of  6 variables:
 $ Ozone  : int  41 36 12 18 NA 28 23 19 8 NA ...
 $ Solar.R: int  190 118 149 313 NA NA 299 99 19 194 ...
 $ Wind   : num  7.4 8 12.6 11.5 14.3 14.9 8.6 13.8 20.1 8.6 ...
 $ Temp   : int  67 72 74 62 56 66 65 59 61 69 ...
 $ Month  : int  5 5 5 5 5 5 5 5 5 5 ...
 $ Day    : int  1 2 3 4 5 6 7 8 9 10 ...
```

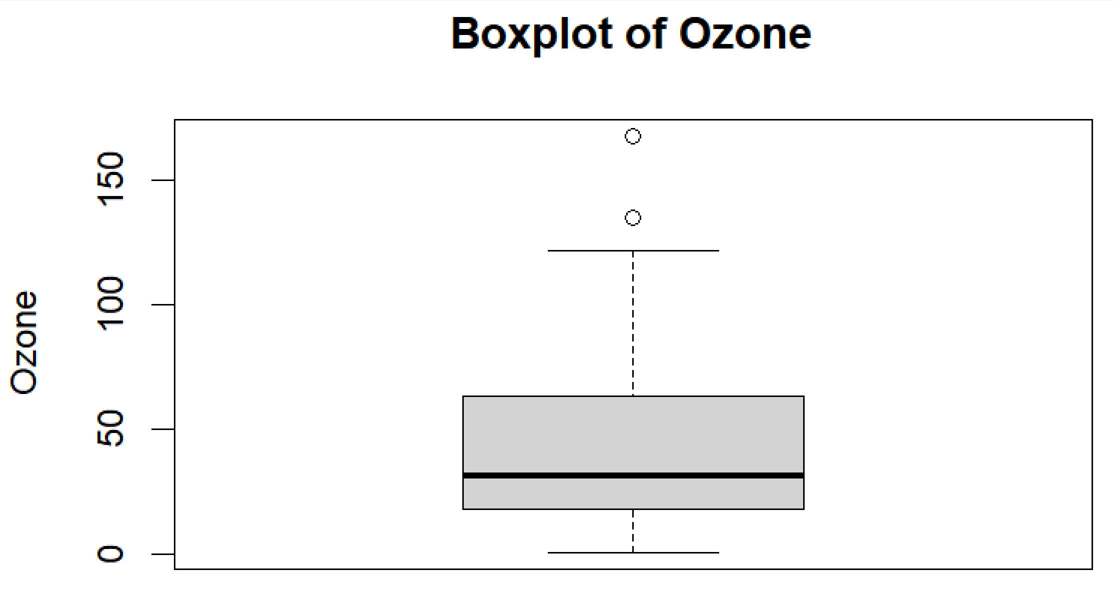
Ozone : 每十億分濃度的臭氧濃度（parts per billion, ppb）；
Solar.R : 每日的太陽能量值（每平方公分測得的能量）；
Wind : 每小時風速（mph）； Temp : 每日最高氣溫（華氏度）；
Month : 5月到9月； Day : 每月日期

3

```
> describe(airquality)
airquality

 6  Variables      153  Observations
---------------------------------------------------------------------------
Ozone
       n  missing distinct     Info     Mean      Gmd      .05      .10      .25
     116       37       67    0.999    42.13    35.28     7.75    11.00    18.00
     .50      .75      .90      .95
   31.50    63.25    87.00   108.50

lowest :   1   4   6   7   8, highest: 115 118 122 135 168
---------------------------------------------------------------------------
Temp
       n  missing distinct     Info     Mean      Gmd      .05      .10      .25
     153        0       40    0.999    77.88    10.74     60.2     64.2     72.0
     .50      .75      .90      .95
    79.0     85.0     90.0     92.0

lowest : 56 57 58 59 61, highest: 92 93 94 96 97
```

3

```
> # 使用箱型圖檢測離群值
> boxplot(airquality$Ozone, main = "Boxplot of Ozone", ylab = "Ozone")
```



3

H_0: 每十億分濃度的臭氧濃度和每日最高氣溫之間有單調相關性($\rho=0$)

H_a: 每十億分濃度的臭氧濃度和每日最高氣溫之間有單調相關性($\rho\neq0$)

```
> spearman_test <- cor.test(airquality$Ozone, airquality$Temp, method = "spearman")
Warning message:
In cor.test.default(airquality$Ozone, airquality$Temp, method = "spearman") :
  Cannot compute exact p-value with ties
> spearman_test

	Spearman's rank correlation rho

data:  airquality$Ozone and airquality$Temp
S = 58778, p-value < 2.2e-16
alternative hypothesis: true rho is not equal to 0
sample estimates:
     rho
0.774043
```

若設α=0.05, p-value < α => reject Ho

每十億分濃度的臭氧濃度和每日最高氣溫之間存在單調相關性

ρ 接近 1 表示每十億分濃度的臭氧濃度和每日最高氣溫之間存在強正相關

3.2.2

Kendall's Tau

相關係數

廖芷萱

3

Kendall's Tau 相關係數

目的：衡量兩個變數之間的單調關係的非參數統計方法。

*它適用於Ordinal scale的數據

假設檢定:

H_0: X and Y are independent

H_a: pairs of observations tend to be discordant, or tend to be concordant

3

步驟： （假設資料為二元隨機變數 X 和 Y）

1. **將 X 和 Y 根據n對觀察值分配**
2. **對比每對觀測值:** (Xi,Yi) v.s. (Xj,Yj)

   若Xi<Xj同時Yi<Yj 或Xi>Xj同時Yi>Yj ： Nc +1

   若Xi<Xj同時Yi>Yj 或Xi>Xj同時Yi<Yj ： Nd +1

   若Xi=Xj 或 Yi=Yj： Nc+0.5 和 Nd+0.5

1. **計算 Kendall's Tau (τ)**

$$\tau = \frac{N_c - N_d}{N_c + N_d},$$

1. **根據計算結果/做鑒定，看兩個變數之間的單調關係**

~ 如果τ接近 1 表示強正相關，接近 -1 表示強負相關，0 則表示無相關。

~ 若 p 值小於 α，則說明相關性顯著，可以推斷出變數之間確實存在單調關係。

3

H_0: 每十億分濃度的臭氧濃度和每日最高氣溫之間有單調相關性($\rho=0$)

H_a: 每十億分濃度的臭氧濃度和每日最高氣溫之間有單調相關性($\rho\neq0$)

```
> # Kendall等級相關檢定
> kendall_test <- cor.test(airquality$Ozone, airquality$Temp, method = "kendall")
> kendall_test

        Kendall's rank correlation tau

data:  airquality$Ozone and airquality$Temp
z = 9.1599, p-value < 2.2e-16
alternative hypothesis: true tau is not equal to 0
sample estimates:
      tau
0.5862988
```

若設$\alpha=0.05$, p-value $<\alpha$ => reject Ho

每十億分濃度的臭氧濃度和每日最高氣溫之間存在單調相關性

$\tau>0$ 表示每十億分濃度的臭氧濃度和每日最高氣溫之間存在正相關

3

對比	Spearman 等級相關係數	Kendall's Tau 相關係數
計算方式	基於變數的排名計算	計算排名之間的一致性程度
處理平局的方式	將平局的排名分配給相同的排名值，可能**影響精確性**。	**更直接考慮平局的影響**，提供穩健的估計。
適用情況	適合分析單調關係，特別是**存在極端值或異常值**的情況。	更適合**小樣本或有較多平局**的情況。
結果解釋	結果範圍一樣在-1和1之間，但Kendall’s Tau 相關係數的計算，通常 Kendall 的值會比 Spearman 的值小，尤其是在樣本較大時。	

3.1.3

Theil-Sen 估計量

廖芷萱

3

Theil-Sen 估計量

目的： 用於線性回歸的非參數估計方法。

* 特別適合在數據包含異常值（outliers）時使用。

* 使用中位數而非平均值來計算斜率

適用的數據：

→ 數據中可能包含異常值。

→ 誤差分佈未知或非常態。

→ 需要在穩健性和效率之間取得平衡。

3

步驟：（假設資料為二元隨機變數 X 和 Y）

1. 將 X 由小到大排序，並與 Y 根據1-n搭配
2. 計算每一對

$$S_{ij} = \frac{Y_j - Y_i}{X_{[j]} - X_{[i]}}$$

, $i<j$

1. β估計值：

$$\hat{\beta} = \mathrm{median}\{S_{ij},\ i,j = 1,\ldots,n,\ i < j\}$$

1. α估計值：

$$\hat{\alpha} = \bar{Y} - \hat{\beta}\bar{X}$$

1. 根據α和β估計

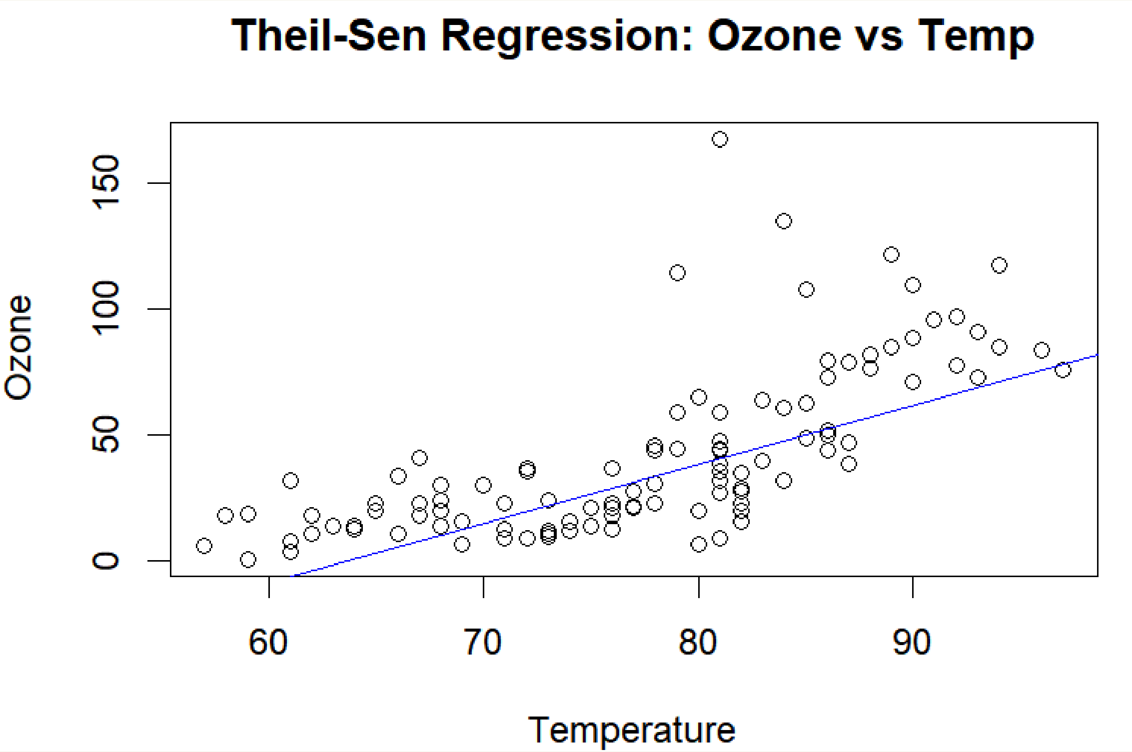
$$Y_i = \alpha + \beta X_i + \epsilon_i,$$

3

```
> # 去除NA值 (Theil-Sen不能處理缺失值)
> airquality_clean <- na.omit(airquality)
> theilsen_model <- mblm(Ozone ~ Temp, data = airquality_clean)
> summary(theilsen_model)

Call:
mblm(formula = Ozone ~ Temp, dataframe = airquality_clean)

Residuals:
    Min      1Q  Median      3Q     Max 
-31.846  -8.513   8.154  19.487 127.154 

Coefficients:
            Estimate     MAD V value Pr(>|V|)    
(Intercept) -148.154  86.590      19   <2e-16 ***
Temp           2.333   1.050    6214   <2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 25.51 on 109 degrees of freedom
```

Ozone = - 148.154 + 2.333 Temp

3



從這個圖中，我們可以觀察到每十億分濃度的臭氧濃度和每日最高氣溫之間是**有一種正相關的趨勢**。

隨著溫度的上升，臭氧濃度有增高的趨勢。

3.1.4

局部加權回歸平滑

Loess

廖芷萱

3

局部加權回歸平滑 Loess

目的：让数据自身揭示其潜在的模式的非參數迴歸方法。

* 透過在資料的局部區域內擬合簡單模型來平滑資料並估計趨勢。

適用的數據：

➔非線性關係的數據
➔具有局部變化的數據
➔包含異常值的數據

3

```
> loess_model <- loess(Ozone ~ Temp, data = airquality_clean)
> plot(airquality_clean$Temp, airquality_clean$Ozone, main = "Loess Fit: Ozone vs Tem
p",
+      xlab = "Temperature", ylab = "Ozone")
> lines(airquality_clean$Temp, predict(loess_model, newdata = airquality_clean), col =
"red")
```

可調整參數：span & degree

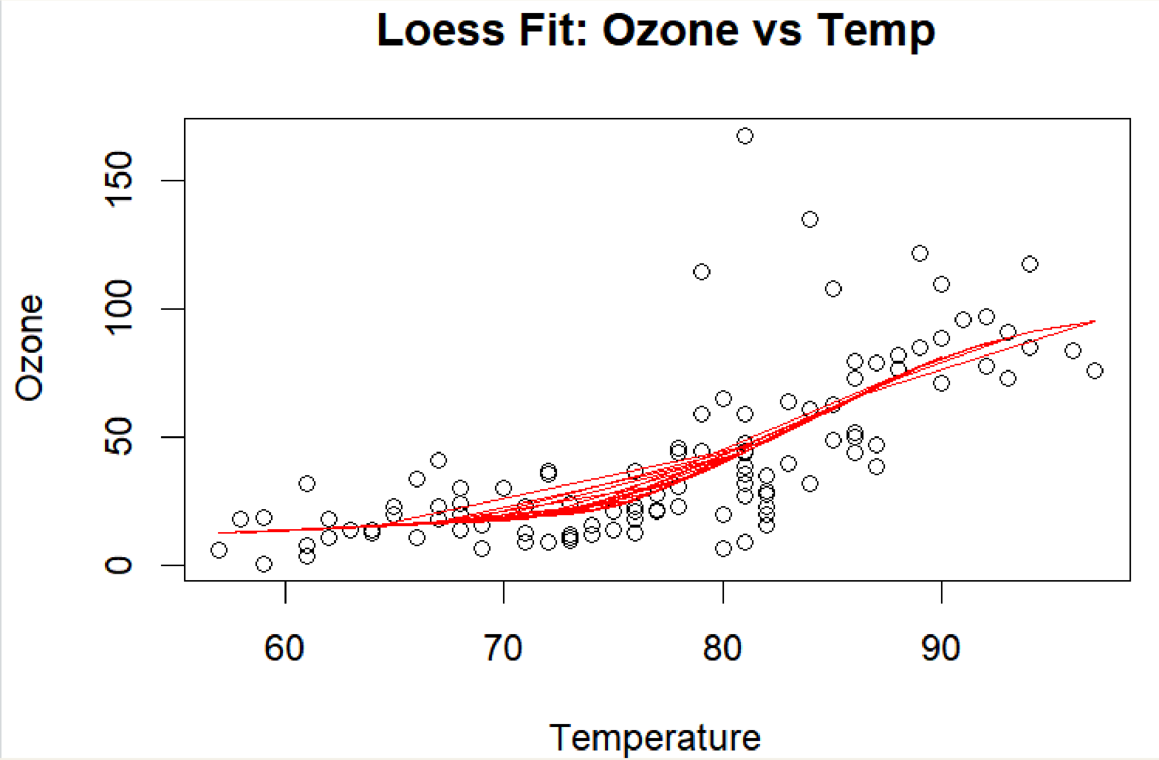
平滑跨度 (span):

- 決定在局部加權回歸中使用多少數據點

多項式階數 (degree):

- 決定了**擬合曲線的複雜度**。
- 建議使用一階多項式 (p = 1)，因為它在計算效率和靈活性之間取得了良好的平衡。

3



從這個圖中，我們可以觀察到隨著溫度的上升，臭氧濃度**整體上先緩慢上升**，直到溫度超過約80華氏度後，**臭氧濃度開始迅速上升**。

3

對比	Theil-Sen 估計量	局部加權回歸平滑 Loess
適用場景	當數據中**含有異常值或極端值**時，Theil-Sen回歸通常表現更好。	適用於**複雜的非線性數據**，能夠進行平滑並擬合局部模式。
計算性能	計算上較為簡單，特別是在需要對抗異常值時。	計算複雜，對**大數據集而言可能不夠高效**。
優勢	非參數方法，**不需要對數據進行假設**	對局部數據變化更為敏感，能夠**捕捉到數據的局部模式**，让数据自身揭示其潛在的模式。
缺點	- 計算上可能**不如最小二乘回歸效率高**，尤其是在處理大量數據時。 - **無法提供估計的置信區間**。	- 比Theil-Sen回歸**對異常值更加敏感** - **不適合於高維數據**，容易過擬合。

3 Reference

Wayne W. Daniel, Applied Nonparametric Statistics, Second Edition.

Sept. 2023 – Jan., 2024, Hsing-Ming Chang, STAT3003 Nonparametric Statistics Lecture Notes

Theil-Sen Estimators in a Multiple Linear Regression Model Xin Dang , Hanxiang Peng , Xueqin Wang and Heping Zhang University of Mississippi and Yale University https://citeseerx.ist.psu.edu/document?repid=rep1&type=pdf&doi=63167c5dbb9bae6f0a269237a9b6a28fa7e1ac20

Nonparametric Regression: Lowess/Loess，GEOG 414/514:Advanced Geographic Data Analysis Scatter-diagram smoothing，Adapted by Ronaldo Dias https://www.ime.unicamp.br/~dias/loess.pdf

Wilcox, R. R. (2012). "Introduction to Robust Estimation and Hypothesis Testing." Academic Press.

3.2 核密度估計

KDE

廖芷萱

3

核密度估計（Kernel Density Estimation, KDE）

目的：估計一個連續變數的真實 pdf。

（實例）美國黃石國家公園的 Old Faithful 間歇泉的噴發數據

```
> library(Hmisc)
> data("faithful")
> str(faithful)
'data.frame':   272 obs. of  2 variables:
 $ eruptions: num  3.6 1.8 3.33 2.28 4.53 ...
 $ waiting  : num  79 54 74 62 85 55 88 85 51 85 ...
```

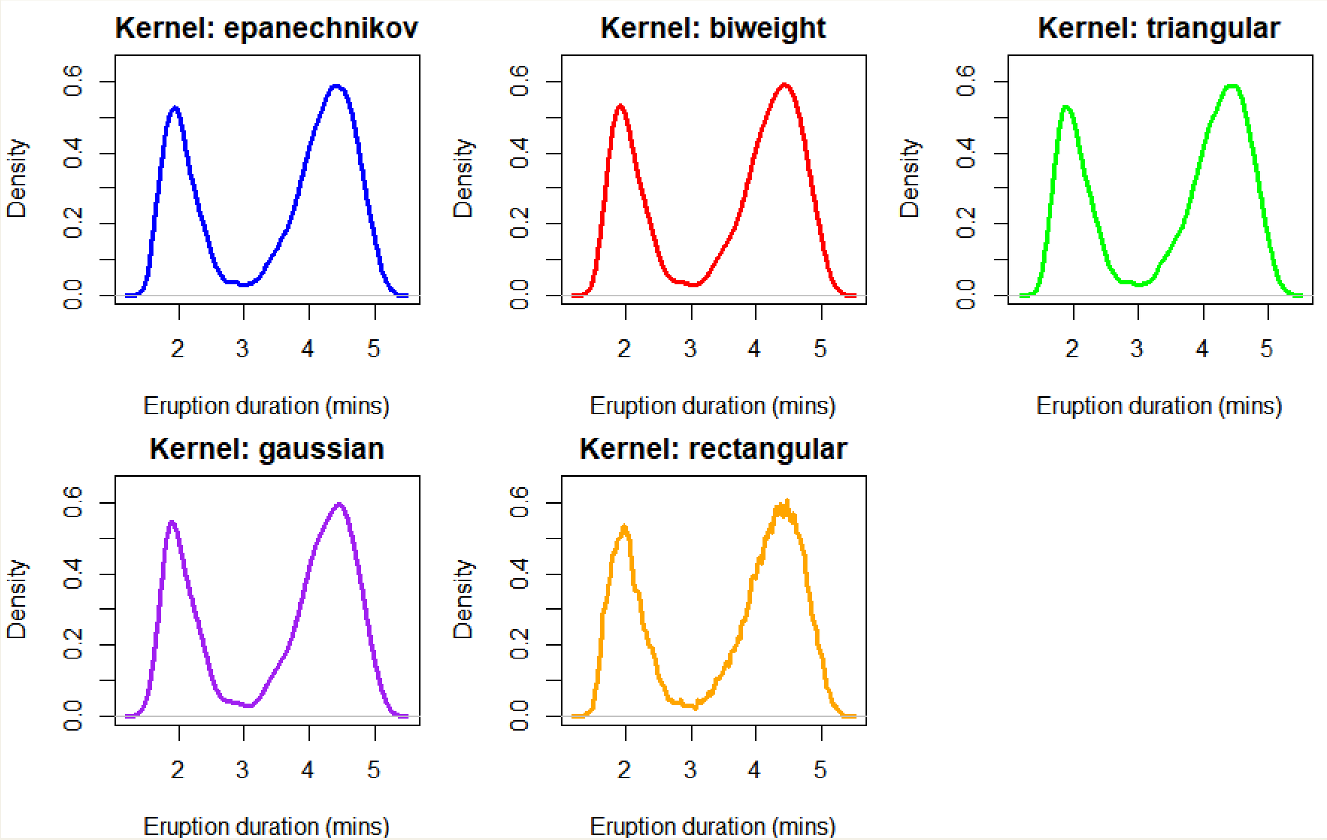
eruptions：每次噴發的持續時間（分鐘）; waiting：噴發之間的等待時間（分鐘）

3

```
> describe(faithful)
faithful

 2  Variables      272  Observations
------------------------------------------------------------------------------------------------------------------------
eruptions
       n  missing distinct     Info     Mean      Gmd      .05      .10      .25      .50      .75      .90      .95
     272        0      126        1    3.488    1.266    1.800    1.852    2.163    4.000    4.454    4.700    4.817

lowest : 1.6   1.667 1.7   1.733 1.75 , highest: 4.933 5     5.033 5.067 5.1
------------------------------------------------------------------------------------------------------------------------
waiting
       n  missing distinct     Info     Mean      Gmd      .05      .10      .25      .50      .75      .90      .95
     272        0       51    0.999     70.9    15.37       48       51       58       76       82       86       89

lowest : 43 45 46 47 48, highest: 91 92 93 94 96
------------------------------------------------------------------------------------------------------------------------
```

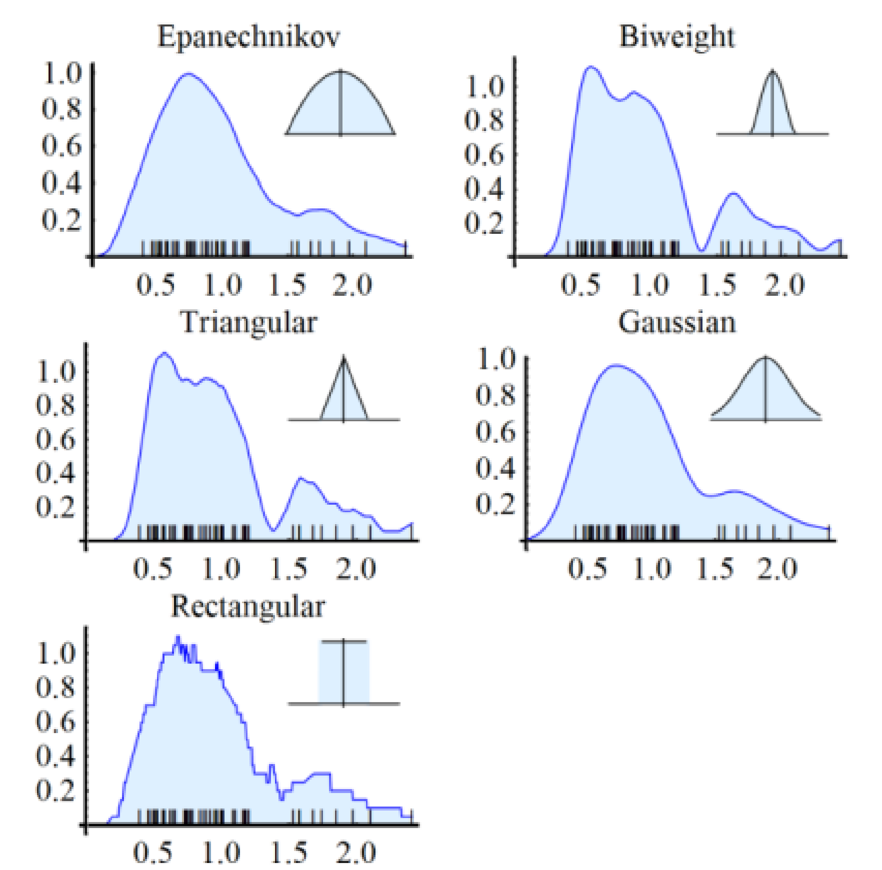
3

不同種類的核函數



Kernel: epanechnikov
Kernel: biweight
Kernel: triangular
Kernel: gaussian
Kernel: rectangular
Density
0.0
0.2
0.4
0.6
2
3
4
5
Eruption duration (mins)

3



Epanechnikov kernel:

-對有界數據表現良好

Biweight kernel:

-對有界數據處理效果較差。

Triangular and Rectangular kernels :

- 容易產生許多局部最大值
- 不建議應用

Gaussian kernel:

- 最平滑的估計
- 最常使用的kernel

3

最佳平滑參數 smoothing parameter

選擇平滑參數的方法	適用的樣本大小	優點	缺點
偏差交叉驗證 (BCV)	15-25	降低樣本變異性	可能會過度平滑密度。
Silverman **經驗法則** (SROT)	50-100	在處理偏態或雙峰態密度時表現良好。	面對具有許多特徵的密度時，可能會產生較高的偏差。
Sheather-Jones **插件** (SJDP)	大樣本和小樣本都表現不錯	在模擬研究、漸近分析和真實數據集中都表現出色。	

3

選擇平滑參數的方法	適用的樣本大小	優點	缺點
最小平方交叉驗證 (LSCV)	-	具有許多特徵的密度方面表現良好。	- 容易低估密度 - 具有較高的樣本變異性。
常態經驗法則 (NROT)	-		處理多峰態或非常態分佈的密度時，容易產生過度平滑的結果。

3

根據數據的局部密度調整平滑程度

選擇平滑參數的方法	適用的樣本大小	優點
自適應核密度估計 (AKDE)	大樣本量	- 有效地處理有界變數問題。 - 非常適合處理具有長尾的偏態變數， - 比固定帶寬方法更準確地估計密度。
自適應核密度估計+第二階段使用 Sheather-Jones 插件 (SJDP) 算法來選擇帶寬(AKSJ)	250以上	- 繼承了 AKDE 的優點。 - 在大樣本中表現出色。 -在對數常態分佈條件下表現最佳。

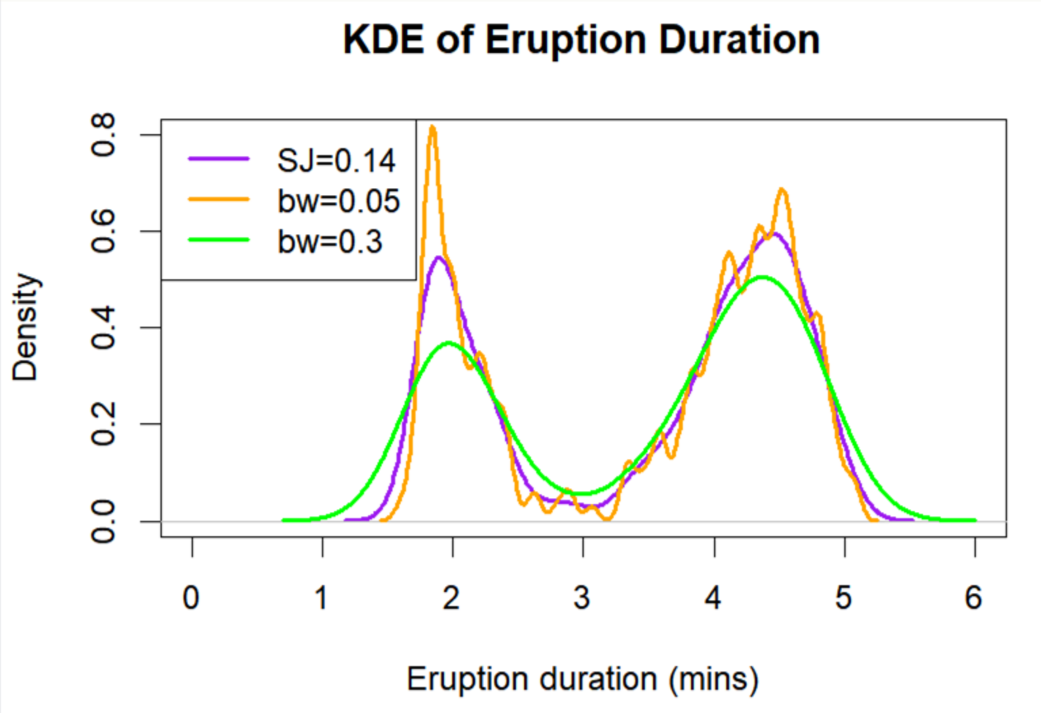
3

```
> #使用Gaussian & SJ
> kde_choosen<- density(faithful$eruptions,kernel="gaussian",bw="SJ")
> kde_choosen

Call:
	density.default(x = faithful$eruptions, bw = "SJ", kernel = "gaussian")

Data: faithful$eruptions (272 obs.);	Bandwidth 'bw' = 0.14

       x                 y            
 Min.   :1.180   Min.   :0.0001834  
 1st Qu.:2.265   1st Qu.:0.0422638  
 Median :3.350   Median :0.1709243  
 Mean   :3.350   Mean   :0.2301726  
 3rd Qu.:4.435   3rd Qu.:0.4134348  
 Max.   :5.520   Max.   :0.5945634  
```

默認平滑參數：Silverman 經驗法則(SROT) , 默認 kernel： gaussian

3

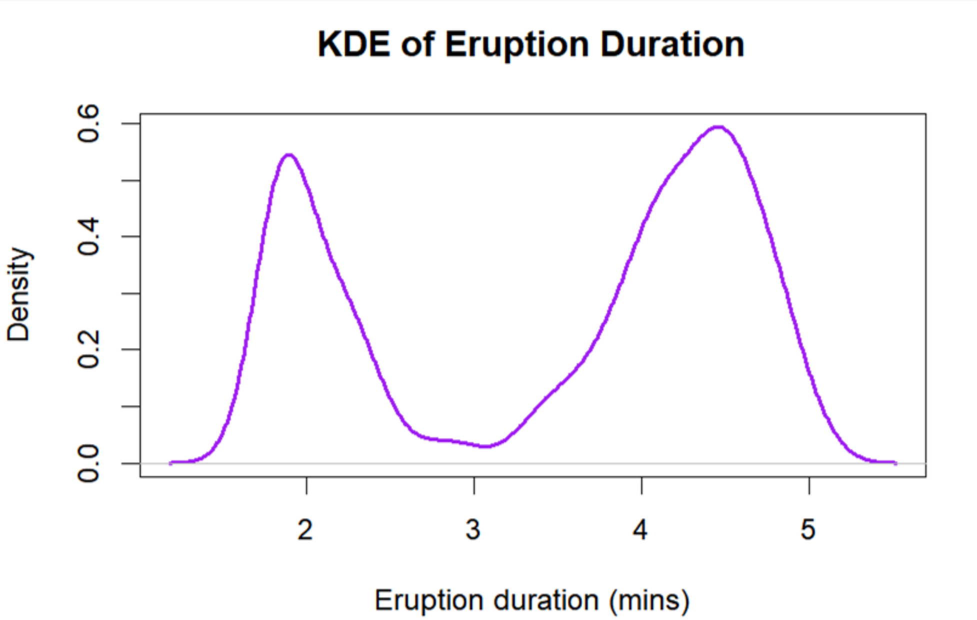
```
> plot(kde_choosen, main = "KDE of Eruption Duration",
+       xlab = "Eruption duration (mins)", ylab = "Density", col = "purple", lwd = 2,yl
im=c(0,0.8),xlim=c(0,6))
> # 添加不同帶寬的 KDE 曲線
> lines(density(faithful$eruptions, bw = 0.05), col = "orange", lwd = 2)  # 小帶寬
> lines(density(faithful$eruptions, bw = 0.3), col = "green", lwd = 2)  # 大帶寬
> legend("topleft", legend = c("SJ=0.14", "bw=0.05", "bw=0.3"),
+         col = c("purple", "orange", "green"), lwd = 2)
```



h值過小會導致估計結果過於粗糙，顯示不必要的細節；

h值過大會導致估計結果過於平滑，掩蓋數據中的重要特徵，例如多模態性

3



圖中顯示了明顯的**雙峰**，意味著數據中存在兩個主要的噴發持續時間範圍。這可能表示美國黃石國家公園的 Old Faithful 間歇泉的噴發有兩種主要模式：**一個較短的噴發時間和一個較長的噴發時間。**

3

```
> pdf_values <- data.frame(Duration = kde_choosen$x, Density = kde_choosen$y)
> head(pdf_values)
  Duration      Density
1 1.179869 0.0001833851
2 1.188363 0.0002229529
3 1.196857 0.0002696114
4 1.205350 0.0003278130
5 1.213844 0.0003968097
6 1.222338 0.0004779552
```

kde_choosen$x 是 x 軸的值（噴發持續時間）
kde_choosen$y 是估計的密度值（PDF）

3 Reference (KDE)

Kernel density estimation and its application by Stanisław Węglarczyk: https://www.itm-conferences.org/articles/itmconf/pdf/2018/08/itmconf_sam2018_00037.pdf

Harpole, J. K., Woods, C. M., Rodebaugh, T. L., Levinson, C. A., & Lenze, E. J. (2014). How bandwidth selection algorithms impact exploratory data analysis using kernel density estimation. Psychological Methods, 19(3), 428–443. https://sci-hub.st/10.1037/a0036850