Fiona Nyamari Kwamboka

BSC. Biostatistics 3.2

SCM224-1096/2022

STA 2404 Non-Parametric Methods

user 2025-03-05

*# ASSIGNMENT II*

*#Question 1*

b11 <- 10

b12 <- 20

b21 <- 15

b22 <- 25

observed\_table <- **matrix**(**c**(b11, b12, b21, b22), nrow = 2, byrow = TRUE)

**print**(observed\_table)

chi\_test <-

**chisq.test**(observed\_table, correct = FALSE)

**print**(chi\_test)

##

## Pearson's Chi-squared test ##

## data: observed\_table

## X-squared = 0.12963, df = 1, p-value = 0.7188

*# Question 2*

**set.seed**(123)

height <- **round**(**runif**(40,150,200),1) age <- **sample** (18**:**60,40,replace= TRUE) MASS <- **round** (**runif**(40,50,100),1)

data <- **data.frame**(Age = age, Height\_cm = height, Mass\_kg = MASS) data

|  |  |  |  |
| --- | --- | --- | --- |
| ## |  | [,1] | [,2] |
| ## | [1,] | 10 | 20 |
| ## | [2,] | 15 | 25 |

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| ## |  | Age | Height\_cm | Mass\_kg |
| ## | 1 | 32 | 164.4 | 56.5 |
| ## | 2 | 49 | 189.4 | 82.7 |
| ## | 3 | 59 | 170.4 | 67.2 |
| ## | 4 | 24 | 194.2 | 82.8 |

|  |  |  |  |  |
| --- | --- | --- | --- | --- |
| ## | 5 | 26 | 197.0 | 66.0 |
| ## | 6 | 58 | 152.3 | 59.4 |
| ## | 7 | 27 | 176.4 | 89.1 |
| ## | 8 | 40 | 194.6 | 54.7 |
| ## | 9 | 44 | 177.6 | 73.3 |
| ## | 10 | 24 | 172.8 | 75.6 |
| ## | 11 | 44 | 197.8 | 80.0 |
| ## | 12 | 49 | 172.7 | 66.6 |
| ## | 13 | 55 | 183.9 | 74.4 |
| ## | 14 | 42 | 178.6 | 97.7 |
| ## | 15 | 51 | 155.1 | 74.1 |
| ## | 16 | 46 | 195.0 | 94.5 |
| ## | 17 | 22 | 162.3 | 95.7 |
| ## | 18 | 25 | 152.1 | 80.4 |
| ## | 19 | 29 | 166.4 | 70.5 |
| ## | 20 | 30 | 197.7 | 57.4 |
| ## | 21 | 35 | 194.5 | 96.8 |
| ## | 22 | 50 | 184.6 | 65.1 |
| ## | 23 | 44 | 182.0 | 53.0 |
| ## | 24 | 42 | 199.7 | 97.4 |
| ## | 25 | 55 | 182.8 | 86.0 |
| ## | 26 | 38 | 185.4 | 57.1 |
| ## | 27 | 32 | 177.2 | 77.5 |
| ## | 28 | 58 | 179.7 | 97.7 |
| ## | 29 | 43 | 164.5 | 79.3 |
| ## | 30 | 48 | 157.4 | 70.2 |
| ## | 31 | 33 | 198.2 | 82.4 |
| ## | 32 | 47 | 195.1 | 66.0 |
| ## | 33 | 23 | 184.5 | 65.4 |
| ## | 34 | 60 | 189.8 | 61.0 |
| ## | 35 | 25 | 151.2 | 68.5 |
| ## | 36 | 39 | 173.9 | 99.2 |
| ## | 37 | 39 | 187.9 | 57.7 |
| ## | 38 | 56 | 160.8 | 54.6 |
| ## | 39 | 48 | 165.9 | 57.1 |
| ## | 40 | 34 | 161.6 | 84.5 |
| *# wilcoxon signed rank test # i)*  test\_result <- **wilcox.test** (age, mu = 24, alternative = "two.sided",exact=FALSE)  test\_result  ##  ## Wilcoxon signed rank test with continuity correction ##  ## data: age  ## V = 734.5, p-value = 1.344e-07  ## alternative hypothesis: true location is not equal to 24 | | | | |

**if**(test\_result**$**p.value**<**0.05){

**print**("reject the null hypothesis: the median age is significantly different from 24")

} **else** {

**print** ("fail to reject the null hypothesis: no significant difference from 24")

}

## [1] "reject the null hypothesis: the median age is significantly different from 24"

*# ii)*

test\_result2 <- **wilcox.test** (height, mu = 140, alternative = "two.sided",exact=FALSE)

test\_result2

##

## Wilcoxon signed rank test with continuity correction ##

## data: height

## V = 820, p-value = 3.708e-08

## alternative hypothesis: true location is not equal to 140

**if**(test\_result2**$**p.value**<**0.05){

**print**("reject the null hypothesis: the median height is significantly different from 140")

} **else** {

**print** ("fail to reject the null hypothesis: no significant difference from 140")

}

## [1] "reject the null hypothesis: the median height is significantly different from 140"

*# iii)*

test\_result3 <- **wilcox.test** (MASS, mu = 58, alternative = "two.sided",exact=FALSE)

test\_result3

##

## Wilcoxon signed rank test with continuity correction ##

## data: MASS

## V = 777, p-value = 8.376e-07

## alternative hypothesis: true location is not equal to 58

**if**(test\_result3**$**p.value**<**0.05){

**print**("reject the null hypothesis: the median mass is significantly different from 58")

} **else** {

**print** ("fail to reject the null hypothesis: no significant difference from

58")

}

## [1] "reject the null hypothesis: the median mass is significantly different from 58"

*#sign test # i)*

n\_positive <- **sum**(age **>**24) n\_negative <- **sum**(age **<**24) n <- n\_positive **+** n\_negative

test\_result4 <- **binom.test**(n\_positive,n,p = 0.5, alternative = "two.sided") test\_result4

##

## Exact binomial test ##

## data: n\_positive and n

## number of successes = 36, number of trials = 38, p-value = 5.399e-09 ## alternative hypothesis: true probability of success is not equal to 0.5 ## 95 percent confidence interval:

## 0.8225094 0.9935613

## sample estimates:

## probability of success ## 0.9473684

*# ii)*

n\_positive <- **sum**(height **>**140) n\_negative <- **sum**(height **<**140) n <- n\_positive **+** n\_negative

test\_result5 <- **binom.test**(n\_positive,n,p = 0.5, alternative = "two.sided") test\_result5

##

## Exact binomial test ##

## data: n\_positive and n

## number of successes = 40, number of trials = 40, p-value = 1.819e-12 ## alternative hypothesis: true probability of success is not equal to 0.5 ## 95 percent confidence interval:

## 0.9119027 1.0000000

## sample estimates:

## probability of success ## 1

*# iii)*

n\_positive <- **sum**(MASS **>**58) n\_negative <- **sum**(MASS **<**58) n <- n\_positive **+** n\_negative

test\_result6 <- **binom.test**(n\_positive,n,p = 0.5, alternative = "two.sided") test\_result6

##

## Exact binomial test ##

## data: n\_positive and n

## number of successes = 32, number of trials = 40, p-value = 0.0001822 ## alternative hypothesis: true probability of success is not equal to 0.5 ## 95 percent confidence interval:

## 0.6435220 0.9094776

## sample estimates:

## probability of success ## 0.8

*#Question 3*

**set.seed**(42)

systolic\_bp <- **rnorm**(40, mean = 130, sd =10) systolic\_bp

## [1] 143.7096 124.3530 133.6313 136.3286 134.0427 128.9388 145.1152

129.0534

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| ## | [9] 150.1842 | 129.3729 | 143.0487 | 152.8665 | 116.1114 | 127.2121 | 128.6668 |
| 136.3595 | |  |  |  |  |  |  |
| ## [17] 127.1575 | | 103.4354 | 105.5953 | 143.2011 | 126.9336 | 112.1869 | 128.2808 |
| 142.1467 | |  |  |  |  |  |  |
| ## [25] 148.9519 | | 125.6953 | 127.4273 | 112.3684 | 134.6010 | 123.6001 | 134.5545 |
| 137.0484 | |  |  |  |  |  |  |
| ## [33] 140.3510 | | 123.9107 | 135.0496 | 112.8299 | 122.1554 | 121.4909 | 105.8579 |
| 130.3612 | |  |  |  |  |  |  |

median\_hypothesis <- 125 median\_hypothesis

## [1] 125

positive\_signs <- **sum**(systolic\_bp **>** median\_hypothesis) negative\_signs <- **sum**(systolic\_bp **<** median\_hypothesis)

sign\_test\_p <- **binom.test**(**min**(positive\_signs, negative\_signs), n = positive\_signs **+** negative\_signs,

p = 0.5, alternative = "two.sided")**$**p.value

**print**(**paste**("Sign Test p-value:", sign\_test\_p))

## [1] "Sign Test p-value: 0.0165890033749747"

*# Interpretation*

"If p-value <0.05, reject H0 and Median systolic BP is significantly different from 125mmHg

If p-value >0.05, fail to reject H0 and No significant difference"

## [1] "If p-value <0.05, reject H0 and Median systolic BP is significantly different from 125mmHg\n If p-value >0.05, fail to reject H0 and No significant difference"

*# Question 4*

**set.seed**(42)

systolic\_bp <- **rnorm**(40, mean = 130, sd =10) systolic\_bp

## [1] 143.7096 124.3530 133.6313 136.3286 134.0427 128.9388 145.1152

129.0534

|  |  |  |  |  |  |  |  |
| --- | --- | --- | --- | --- | --- | --- | --- |
| ## | [9] 150.1842 | 129.3729 | 143.0487 | 152.8665 | 116.1114 | 127.2121 | 128.6668 |
| 136.3595 | |  |  |  |  |  |  |
| ## [17] 127.1575 | | 103.4354 | 105.5953 | 143.2011 | 126.9336 | 112.1869 | 128.2808 |
| 142.1467 | |  |  |  |  |  |  |
| ## [25] 148.9519 | | 125.6953 | 127.4273 | 112.3684 | 134.6010 | 123.6001 | 134.5545 |
| 137.0484 | |  |  |  |  |  |  |
| ## [33] 140.3510 | | 123.9107 | 135.0496 | 112.8299 | 122.1554 | 121.4909 | 105.8579 |
| 130.3612 | |  |  |  |  |  |  |

median\_hypothesis <- 80 median\_hypothesis

## [1] 80

positive\_signs <- **sum**(systolic\_bp **>** median\_hypothesis) negative\_signs <- **sum**(systolic\_bp **<** median\_hypothesis)

sign\_test\_p <- **binom.test**(**min**(positive\_signs, negative\_signs), n = positive\_signs **+** negative\_signs,

p = 0.5, alternative = "two.sided")**$**p.value

**print**(**paste**("Sign Test p-value:", sign\_test\_p))

## [1] "Sign Test p-value: 1.81898940354586e-12"

*# Interpretation*

"If p-value <0.05, reject H0 and Median systolic BP is significantly different from 80mmHg

If p-value >0.05, fail to reject H0 and No significant difference"

## [1] "If p-value <0.05, reject H0 and Median systolic BP is significantly different from 80mmHg\n If p-value >0.05, fail to reject H0 and No significant difference"