**13.10 R 语言应用**

**# 陈文贤 着 《大话统计学》 清华大学出版社 2022年**

if(!require(BSDA)){install.packages("BSDA")}; library(BSDA)

if(!require(snpar)){install.packages("snpar")}; library(snpar)

if(!require(randtests)){install.packages("randtests")}; library(randtests)

if(!require(PairedData)){install.packages("PairedData")}; library(PairedData)

if(!require(DescTools)){install.packages("DescTools")}; library(DescTools)

if(!require(devtools)){install.packages("devtools")}; library(devtools) ; par(mfrow=c(1,1))

if(!require(tidyverse)){install.packages("tidyverse")}; library(tidyverse)

if(!require(rstatix)){install.packages("rstatix")}; library(rstatix)

if(!require(ggpubr)){install.packages("ggpubr")}; library(ggpubr)

if(!require(dplyr)){install.packages("dplyr")}; library(dplyr)

# 符号检验 Sign test例题13.1

x <- c(22, 30, 31, 40, 37, 25, 29, 14, 30, 17, 23, 32, 20, 40, 28, 26, 33, 25, 34, 21)

options(digits = 4) ; median(x)

SIGN.test(x, md = 25, alternative = "greater")

# 符号检验 Sign test例题13.2

x <- c(750, 860, 950, 830, 750, 680, 720, 810)

y <- c(850, 880, 930, 860, 800, 740, 760, 800)

z <- x - y ; SIGN.test(z, md = 0, alternative = "two.sided")

SIGN.test(x, y, alternative = "two.sided") # 符号检验

binom.test(2, 8) # 二项分布检验 binom.test(2, 8, p=0.05, alternative="less")

# 秩和检验 Wilcoxon rank sum test例题13.3

x<- c(100, 94, 95, 81, 103, 97, 90, 102, 91, 98, 88, 87)

wilcox.test(x, mu=90, alternative = "two.sided") # 秩和检验

# 游程检验 Run test例题13.4

x <- c(1, 12, 8, 7, 3, 10, 5, 7, 15, 12, 9, 18, 12, 17, 1, 7, 18, 6, 14, 11, 5, 2)

runs.test(x) ; RunsTest(x) ; RunsTest(x > median(x))

plot( (x < median(x)) - 0.5, type="s", ylim=c(-1,1) ) + abline(h=0)

RunsTest(x, exact=TRUE) # 游程检验 R分布

RunsTest(x, exact=FALSE) # 游程检验 正态分布

# 游程检验 Run test例题13.5

A <- c(73,67,72,46,83,75,62,90,95) ; B <- c(71,47,68,87,77,92,65,86,79,57)

RunsTest(A, B, exact=TRUE) # 游程检验 R分布

RunsTest(A, B, exact=FALSE) # 游程检验 正态分布

# 秩和检验 Wilcoxon rank sum test例题13.5

wilcox.test(A, B, alternative = "two.sided") # 秩和检验

# KW 检验 Kruskal-Wallis test例题13.6

data <- read.csv("C:/大话统计学 网络资源/StatData/Chap13\_6.csv", head=TRUE) ; data

options(digits = 4) ; as.factor(data$Control) ; summary(data$Y)

group\_by(data, Control) %>% summarise( count = n(), mean = mean(Y), sd = sd(Y),

median = median(Y), IQR = IQR(Y) )

ggboxplot(data, x = "Control", y = "Y", color = "Control", palette = c("#00AFBB",

"#E7B800", "#FC4E07"), order = c("A", "B", "C"), main = "例题13.6", ylab = "Y",

xlab = "Control")

ggline(data, x = "Control", y = "Y", add = c("mean\_se", "jitter"), order = c("A", "B", "C"),

ylab = "Y", xlab = "Control", main = "例题13.6")

kruskal.test(Y ~ Control, data = data) # KW 检验kruskal test

# cor(X, Y, method = "pearson") ; cor(X, Y, method = "kendall")

# cor(X, Y, method = "spearman") ; cor.test(X, Y, method="spearman")

# shapiro.test(X) ; shapiro.test(Y) # 检验 X, Y 是否正态分布

# DescTools::LillieTest(X) ; nortest::lillie.test(X) # Lilliefors检验 X 是否正态分布

# Friedman检验 Friedman test例题13.7

Times <- matrix(c(4.7, 9.4, 6.3, 3.5, 7.6, 5.1, 0.1, 5.3, 1.8, 1.6, 6.2, 3.6),

nrow = 4, byrow = TRUE, dimnames = list(1 :4, c("A", "B", "C")))

friedman.test(Times)

# Spearman 相关系数 例题13.9

X <- c(2.7,4.1,3.5,2.7,2.2,3.9,2.2,2.2,3.5,2.2)

Y <- c( 28, 19, 22 ,30, 26, 19, 38 ,40, 26, 45)

cor(X, Y, method = "spearman")

cor(X, Y, method = "spearman")

cor(X, Y, method = "pearson")

cor(X, Y, method = "kendall")

cor(X, Y, method = "spearman")

cor.test(X, Y, method="spearman")

shapiro.test(X) ; shapiro.test(Y) # 检验 X, Y 是否正态分布

DescTools::LillieTest(X)

nortest::lillie.test(X) # Lilliefors检验 X 是否正态分布

# 综合检验 例题13.10

data <- read.csv("C:/大话统计学 网络资源/StatData/Chap13\_10.csv", header=TRUE)

data <- data %>% gather(key = "time", value = "score", t1, t2, t3) %>%

convert\_as\_factor(id, time)

data %>% group\_by(time) %>% get\_summary\_stats(score, type = "common")

ggboxplot(data, x = "time", y = "score", add = "jitter", color = "time",

palette = c("blue", "red", "purple") , main = "例题13.10" )

# 以下三个检验调用library(stats)

kruskal.test(score ~ time, data = data) # KW 检验 Kruskal-Wallis test

pairwise.wilcox.test(x=data$score, g=data$time, p.adjust.method = "BH")

friedman.test(score ~ time |id, data = data) # Friedman 检验

# 以下三个检验调用library(rstatix)

res.fried <- data %>% friedman\_test(score ~ time |id) ; res.fried # Friedman 检验

pwc <- data %>% wilcox\_test(score ~ time, paired = TRUE, p.adjust.method =

"bonferroni") ; pwc # Wilcoxon test 符号秩检验

data %>% sign\_test(score ~ time, p.adjust.method = "bonferroni") # Sign test 符号检验

pwc <- pwc %>% add\_xy\_position(x = "time")

ggboxplot(data, x = "time", y = "score", add = "point", color = "time", palette =

c("blue", "red", "purple") , main = "例题13.10") + stat\_pvalue\_manual(pwc, hide.ns =

TRUE) + labs( subtitle = get\_test\_label(res.fried, detailed = TRUE), caption =

get\_pwc\_label(pwc) )