STAT537: Statistics for Research I: HW#4

Due on Sep. 15, 2016

Dr. Schmidhammer TR 11:10am – 12:25pm

Wenqiang Feng

Contents

Problem 1

Ott 5.66 abc

Solution. (a) **Using a graphical display, determine whether the data appear to be a random sample from a normal distribution.**

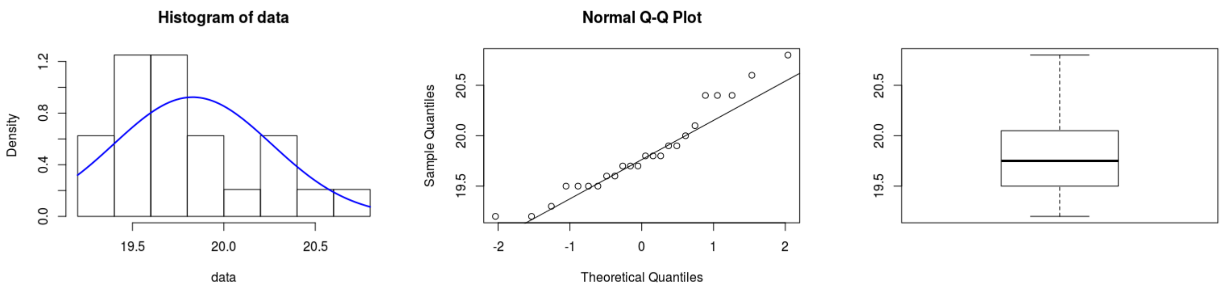

Figure 1: Histogram, Normal Q-Q and Boxplot of the data.

```
Shapiro-Wilk normality test

data:  data
W = 0.93774, p-value = 0.1452
> skewness(data)
     Rate
0.6270016
> kurtosis(data)
    Rate
2.623399
```

The plots in Figure.1 indicate that the data is a right-skewed but only slightly, since there are no outliers indicated on the box plot. And skewness here is 0.62 which confirms that the data is slightly skewed to the right. For the kurtosis, we have 2.623399 implying that the distribution of the data is platykurtic, since the computed value is less than 3.

(b) **Estimate the mean dissolution rate for the batch of tablets, for both a point estimate and a 99% confidence interval.**

Since $n = 24$, $\bar{X} = 19.82917$ and $\sigma = 0.4318607$, the formula for 99% confidence interval is given by

$$\bar{X} \pm t_{\alpha/2}\left(\frac{\sigma}{\sqrt{n}}\right)$$

Therefore 99% confidence interval is $[19.58169, 20.07664]$.

(c) **Is there significant evidence that the batch of pills has a mean dissolution rate less than 20 mg (80% of the labeled amount in the tablets)? Use a $\alpha = .01$.**

Let $H_0 : \mu \geq 20$, $H_\alpha : \mu < 20$, $n = 24$, $\bar{X} = 19.82917$, $\sigma = 0.4318607$, $\alpha = 0.01$, so the p-value is

$$p - value = P(t \leq \frac{\bar{X} - 20}{\sigma/\sqrt{n}}) = P(t \leq -1.937914) = 0.03250285 > 0.01$$

Hence, there is no sufficient information to reject the Null Hypothesis. Therefore, there is no sufficient information to conclude that the average dissolution rate is less than 20.

□

Problem 2

Ott 5.68

Solution. Let $H_0 : \mu \leq 25$, $H_\alpha : \mu > 25$, $n = 15, \bar{X} = 28.2$, $\sigma = 11.44053$, $\alpha = 0.05$, so the p-value is

$$p - value = P(t \geq \frac{\bar{X} - 25}{\sigma/\sqrt{n}}) = P(t \leq 1.083302) = 0.1484898 > 0.05$$

Hence, there is no sufficient information to reject the Null Hypothesis. Therefore, there is no sufficient information to conclude that the average time to fill an order is greater than 25 minutes. □

Problem 3

Ott 5.72

Solution. (a) **95% confidence interval:** $n = 35, \bar{X} = 30.51429$, $\sigma = 12.35831$, $\alpha = 0.05$, the formula for 99% confidence interval is given by

$$\bar{X} \pm t_{\alpha/2}\left(\frac{\sigma}{\sqrt{n}}\right)$$

Therefore 95% confidence interval is $[26.26906, 34.75951]$.

(b) **99% confidence interval:** $n = 35, \bar{X} = 30.51429$, $\sigma = 12.35831$, $\alpha = 0.01$, the formula for 99% confidence interval is given by

$$\bar{X} \pm t_{\alpha/2}\left(\frac{\sigma}{\sqrt{n}}\right)$$

Therefore 99% confidence interval is $[24.81485, 36.21372]$. And the average exercise capacity for healthy male inductees in located in the 99% confidence interval.

(c) **How would your interval change using a 99% confidence interval?** It seems that 99% confidence interval end points are is pushed outward compared with the 95% confidence interval. Hence 99% confidence interval contains 95% confidence interval.

□

Appendix

R code for HW#4

Listing 1: R Source code for HW#4

```r
rm(list = ls())
# set the path or enverionment
setwd("/home/feng/Dropbox/UTK_Course/Stat537/Excel/CH05")


#install.packages("readxl") # CRAN version
library(readxl)
#install.packages("moments")
library(moments)
rawdata = read_excel("ex5-66.xls",sheet = 1)
#attach(rawdata)

data =as.matrix(rawdata)
```

```
#hist(data)
hist(data, prob=TRUE)            # prob=TRUE for probabilities not counts
#lines(density(data))             # add a density estimate with defaults
#lines(density(data, adjust=2), lty="dotted")   # add another "smoother" density
curve(dnorm(x, mean=mean(data), sd=sd(data)), add=TRUE, col="blue", lwd=2)

boxplot(data)
shapiro.test(data)
skewness(data)
kurtosis(data)
qqnorm(data);qqline(data)

#
n =nrow(data)
n
mu=mean(data)
mu
sigma = sd(data)
sigma

#error <- qnorm(0.995)*sigma/sqrt(n)
error = qt(.995, df=n-1)*sigma/sqrt(n)
ci = c(mu-error,  mu+error)
ci

#
t=(mu-20)/(sigma/sqrt(n))
t
p= pt(t,df=n-1)
p

##########################################################
# Ott 5.68
rawdata = read_excel("ex5-68.xls",sheet = 1)
data = as.matrix(rawdata)
n =nrow(data)
n
mu=mean(data)
mu
sigma = sd(data)
sigma

t=(mu-25)/(sigma/sqrt(n))
t
p = 1-pt(t,df=n-1)
p

##########################################################
# Ott 5.68
rawdata = read_excel("ex5-72.xls",sheet = 1)
data = as.matrix(rawdata)
n =nrow(data)
```

```
n
mu=mean(data)
mu
sigma = sd(data)
sigma

# 95% C.I.
error = qt(.975, df=n-1)*sigma/sqrt(n)
ci = c(mu-error,  mu+error)
ci

# 99% C.I.
error = qt(.995, df=n-1)*sigma/sqrt(n)
ci = c(mu-error,  mu+error)
ci
```