**2028: Basic Statistical Methods**
**Solutions - Homework 5**

1. **Hypothesis Testing for the Population Mean**

(a) a The parameter of interest is the true mean coefficient of restitution. The hypothesis test is: $H_0 : \mu = 0.635 \leftrightarrow H_1 : \mu > 0.635$. Let $\mu_0 = 0.635$. The test statistic is

$$t = \frac{\bar{x} - \mu_0}{s/\sqrt{n}} = -5.16$$

and compare it to the t-quantile $t_{0.05,39} = 1.685$. We therefore fail to reject the null hypothesis since $t = -5.16 < t_{0.05,39} = 1.685$. There is not sufficient evidence to conclude that the true mean coefficient of restitution is greater than 0.635 at $\alpha = 0.05$.
When $H_0$ is true, $T = \frac{\bar{x}-0.635}{s/\sqrt{n}} \sim t_{39}$
p-value=$P(T > t) = 0.9999963 > \alpha = 0.05 \Rightarrow$ We cannot reject the null hypothesis.

c The hypothesis test is: $H_0 : \mu = 0.635 \leftrightarrow H_1 : \mu > 0.635$.
If $H_1$ is true, that is $\mu = 0.64$, and then $\frac{\bar{x}-0.635}{s/\sqrt{n}} = \frac{\bar{x}-0.64+0.005}{s/\sqrt{n}} = T + \frac{0.005\sqrt{n}}{s}$, where $T \sim t_{39}$

$$\begin{aligned}
\text{power of test} &= P(\text{reject } H_0|H_1 \text{ is true}) = 1 - P(\text{accept } H_0|H_1 \text{ is true}) \\
&= 1 - P(\frac{\bar{x} - 0.635}{s/\sqrt{n}} \leq t_{0.005,39}|\mu = 0.64) = 1 - P(T + \frac{0.005\sqrt{n}}{s} \leq t_{0.005,39}) \\
&= 1 - P(T \leq t_{0.005,39} - \frac{0.005\sqrt{n}}{s}) \\
&\approx 1 - 0.23 = 0.77
\end{aligned}$$

d We've known from c that power of test $= 1 - P(T \leq t_{0.005,39} - \frac{0.005\sqrt{n}}{s})$, so

$$1 - P(T \leq t_{0.005,39} - \frac{0.005\sqrt{n}}{s}) \geq 0.75 \Leftrightarrow P(T \leq t_{0.005,39} - \frac{0.005\sqrt{n}}{s}) \leq 0.25$$

$$\Rightarrow t_{0.005,39} - \frac{0.005\sqrt{n}}{s} \leq -0.681 \Rightarrow n \geq (\frac{s(t_{0.005,39} + 0.681)}{0.005})^2 \approx 38$$

e The lower confidence bound is

$$\bar{x} - t_{0.05,n-1}\frac{s}{\sqrt{n}} = 0.6209.$$

Because $0.635 > 0.6209$ we fail to reject the null hypothesis.

(b) i. The appropriate hypotheses are $H_0 : \mu = 75$ vs $H_A : \mu < 75$, since we want to strongly support an improvement in average drying time. Only when $H_0$ is rejected, the additive is declared successful and used.

ii. The observed data come from a random sample $X_1, \dots, X_{25}$ normally distributed $N(\mu, 9^2)$, where $\mu$ is the average drying time with the additive. Therefore, assuming the null hypothesis, the sampling distribution of $\bar{X}$ is

$$\bar{X} \sim N(75, \frac{9^2}{25}).$$

We compute the type I error as:

$$\begin{aligned}\alpha =& P(\bar{X} \le 70.8 \text{ when } \bar{X} \sim N(75, (1.8)^2)) = P\left(\frac{\bar{X}-75}{1.8} \le \frac{70.8-75}{1.8}\right)\\ =& P(Z \le -2.33) = .01.\end{aligned}$$

Therefore, the type I error is .01.

iii. The type II error is computed as follows:

$$\begin{aligned}\beta(72) =& P(\text{type II error when } \mu = 72) = P(H_0 \text{ is not rejected when it is false because } \mu = 72)\\ =& P(\bar{X} > 70.8 \text{ when } \bar{X} \sim N(72, (1.8)^2)) = P\left(\frac{\bar{X}-72}{1.8} > \frac{70.8-72}{1.8}\right)\\ =& 1 - P(Z \le -0.67) = 1 - .2514 = .7486.\end{aligned}$$

2. **Hypothesis Testing for the Proportion Parameter**

(a) a The hypothesis test is: $H_0 : p = 0.78 \leftrightarrow H_1 : p > 0.78$.
p-value= $P(Z > \frac{289-0.78n}{\sqrt{0.78(1-0.78)n}}) = 1 - \Phi(\frac{289-0.78\times350}{\sqrt{0.78(1-0.78)\times350}}) \approx 0.02 < 0.05$
When $\alpha = 0.05$, we can reject the null hypothesis. So the success rate for PN is greater than the historical success rate. The p-value is 0.02.

b The confidence interval of $p$ is

$$\frac{\widehat{p}-p}{\sqrt{\widehat{p}(1-\widehat{p})/n}} \le z_\alpha \Leftrightarrow p \ge \widehat{p} - z_\alpha\sqrt{\widehat{p}(1-\widehat{p})/n} \Leftrightarrow p > 0.792$$

Since 0.78 is not in the confidence interval, we can reject the null hypothesis. The conclusion is the same as a.

(b) a The hypothesis test is: $H_0 : p = 0.1 \leftrightarrow H_1 : p > 0.1$.
p-value= $P(Z > \frac{16-0.1n}{\sqrt{0.1(1-0.1)n}}) = 1 - \Phi(\frac{16-0.1\times200}{\sqrt{0.1(1-0.1)\times200}}) \approx 0.827 > 0.01$
When $\alpha = 0.01$, we fail to reject the null hypothesis. So this finding doesn't support the researcher's claim. The p-value is 0.827.

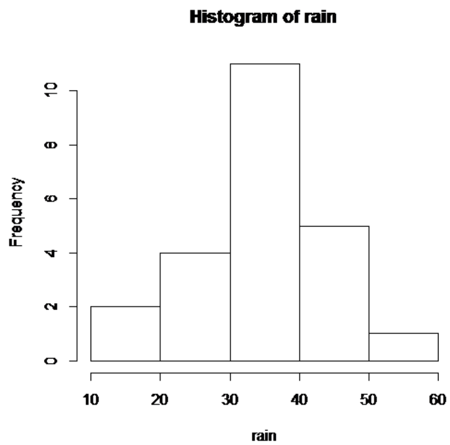
b The confidence interval is

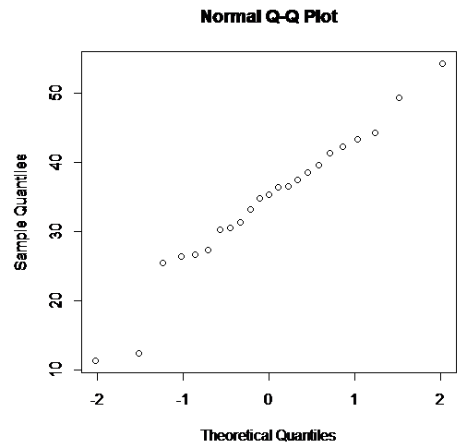
$$\frac{\widehat{p}-p}{\sqrt{\widehat{p}(1-\widehat{p})/n}} \le z_\alpha \Leftrightarrow p \ge \widehat{p} - z_\alpha\sqrt{\widehat{p}(1-\widehat{p})/n} \Leftrightarrow p > 0.0354$$

Since 0.1 is in the confidence interval, we cannot reject the null hypothesis. The conclusion is the same as a.

3. **Computer Problem**

(a) The data appears to be roughly normal.




(b) The two-sided 95% confidence interval for the mean rainfall is (29.88, 38.71).
Since 35 is included, the answer is no.

(c)

$$H_0 : \mu = 30 \leftrightarrow H_A : \mu > 30$$

(d) The one-sided confidence interval is (30.63744, $\infty$).
We conclude that the mean rainfall is above 30 because the lower bound for the confidence interval is above 30.

(e) The p-value is 0.02805.
We would reject $H_0$ for a 90% confidence interval but not for a 99% confidence interval.

```
R Code:
#read in the data
data = read.table("rain.txt")
rain = as.numeric(data[,1])
#Question 1
hist(rain)
qqnorm(rain)
#Question 2
#do the computations yourself
mu = mean(rain)
sd = sd(rain)
n = 23
a = .05
l = mu - qt(1-a/2,n-2)*sd/sqrt(n)
u = mu + qt(1-a/2,n-2)*sd/sqrt(n)
#or use t.test, like in question 4
t = t.test(rain, conf.level = .95, mu = 35)
> t
```

```
One Sample t-test
data:  rain
t = -0.3323, df = 22, p-value = 0.7428
alternative hypothesis: true mean is not equal to 35
95 percent confidence interval:
29.87810 38.70712
sample estimates:
mean of x
34.29261
#Question 3, 4 and 5
t2 = t.test(rain, conf.level=.95,mu=30,alternative=c("greater"))
>t2
One Sample t-test
data:  rain
t = 2.0166, df = 22, p-value = 0.02805
alternative hypothesis: true mean is greater than 30
95 percent confidence interval:
30.63744      Inf
sample estimates:
mean of x
 34.29261
```