PSTAT 122 - HW1

Emily Lu (Perm ID: 4707147)

April 19, 2020

2.25 The following are the burning times (in minutes) of chemical flares of two different formulations. The design engineers are interested in both the mean and variance of the burning times.

Type 1	Type 2
65	64
81	71
57	83
66	59
82	65
82	56
67	69
59	74
75	82
70	79

(a) Test the hypothesis that the 2 variances are equal. Use $\alpha = 0.05$.

Sol.: $H_0 : \sigma_1^2 = \sigma_2^2$ vs. $H_1 : \sigma_2^2 \neq \sigma_2^2$.

```
n1 <- length(type1)
n2 <- length(type2)

# sample variances
s12 <- sum((type1 - mean(type1))^2)/(n1 - 1)
s22 <- sum((type2 - mean(type2))^2)/(n2 - 1)

# critical F
F_stat <- s12/s22

sprintf('Critical F is %s.', F_stat)
```

```
## [1] "Critical F is 0.978216818642351."
```

```
sprintf('Observed F is %s.', qf(0.025, 9, 9))
```

```
## [1] "Observed F is 0.248385854694455."
```

```
# checking results with off-the-shelf R function
var.test(type1, type2)
```

```
##
##  F test to compare two variances
##
## data:  type1 and type2
## F = 0.97822, num df = 9, denom df = 9, p-value = 0.9744
## alternative hypothesis: true ratio of variances is not equal to 1
## 95 percent confidence interval:
##  0.2429752 3.9382952
## sample estimates:
```

```
## ratio of variances
##          0.9782168
```

Using the F-test, we find that F=0.97822 and the observed F statistic is 0.24839. Since our observed F statistic is less than our critical F, we fail to reject H_0 and could conclude that the variances are equal.

(b) Using the results of (a), test the hypothesis that the mean burning times are equal. Use $\alpha = 0.05$. What is the p-value for this test?

Sol. $H_0 : \mu_1 = \mu_2$ vs. $H_1 : \mu_1 \neq \mu_2$

```r
ybar1 <- mean(type1)
ybar2 <- mean(type2)

sp <- sqrt(((n1 - 1)*s12 + (n2 - 1)*s22)/(n1 + n2 - 2))
t_stat <- (ybar1 - ybar2)/(sp*(sqrt(1/n1 + 1/n2)))

pval <- 2*pt(t_stat, n1 + n2 - 2, lower.tail = F)
sprintf('p-value is %s.', pval)
```

```
## [1] "p-value is 0.962238784477904."
```

```r
# checking results with off-the-shelf R function
t.test(type1, type2, var.equal = T)
```

```
##
##  Two Sample t-test
##
## data:  type1 and type2
## t = 0.048008, df = 18, p-value = 0.9622
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -8.552441  8.952441
## sample estimates:
## mean of x mean of y
##      70.4      70.2
```

Using the two-sided t-test, we find our p-value to be 0.9622 which is greater than $\alpha = 0.05$. Therefore, we fail to reject H_0 and could conclude that the meaning burning times are equal.

(c) Discuss the role of the normality assumption in this problem. Check the assumption of normality for both types of flares.

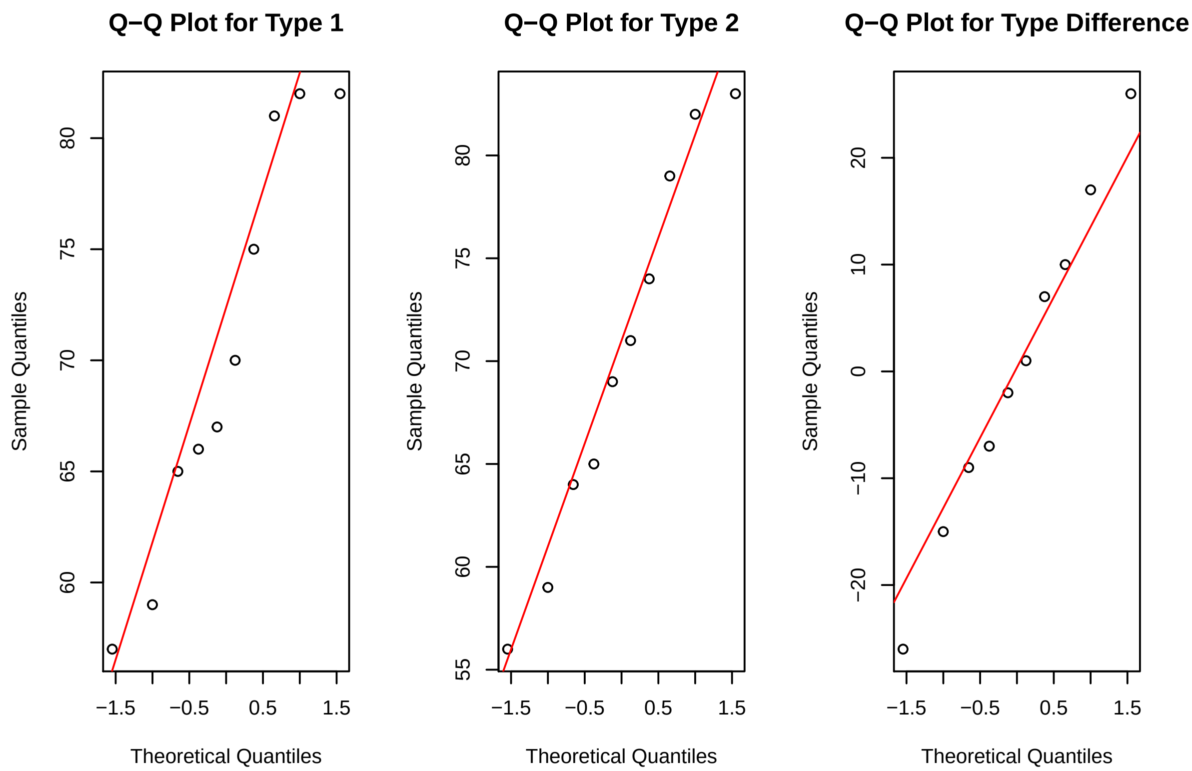
Sol.: The role of the normality assumption in this problem is the equality of variances for the two group. Based on our part a results, the variances of the two groups are equal and thus, the normality assumption would be satisfied. However, to further examine the normality assumption, we plotted a Q-Q plot below.

```r
par(mfrow = c(1,3))

qqnorm(type1, main = 'Q-Q Plot for Type 1')
qqline(type1, col = 'red')

qqnorm(type2, main = 'Q-Q Plot for Type 2')
qqline(type2, col = 'red')

type_diff <- type1 - type2
qqnorm(type_diff, main = 'Q-Q Plot for Type Difference')
qqline(type_diff, col = 'red')
```



Although there are some minor divergences from the normal lines, the overall trend for both type of flares seems to be normally distributed. Additionally, since a paired t-test was performed, a Q-Q plot of the type differences was also plotted to determine the assumption of normality. The Q-Q plot of the differences looks better than the Q-Q plots of individual ones and follows the diagonal line more closely, thus the normality assumption is reasonable for paired t-test.

2.29 The diameter of a ball bearing was measured by 12 inspectors, each using two different kinds of calipers. The results are as follows:

Caliper 1	Caliper 2
0.265	0.264
0.265	0.265
0.266	0.264
0.267	0.266
0.267	0.267
0.265	0.268
0.267	0.264
0.267	0.265
0.265	0.265
0.268	0.267
0.268	0.268
0.265	0.269

(a) Is there a significance difference between the means of the population of measurements from which the two samples were selected? Use $\alpha = 0.05$.

Sol.: Let $\mu_d = \mu_1 - \mu_2$. $H_0 : \mu_d = 0$ vs. $H_1 : \mu_d \neq 0$

```r
n1 <- length(caliper1)
n2 <- length(caliper2)

S1 <- sum((caliper1 - mean(caliper1))^2)/(n1 - 1)
S2 <- sum((caliper2 - mean(caliper2))^2)/(n2 - 1)

sp <- sqrt(((n1-1)*S1 + (n2-1)*S2)/(n1+n2-2))
t_stat <- (mean(caliper1) - mean(caliper2))/(sp*sqrt(1/n1+1/n2))

sprintf('Critical t is %s.', t_stat)
```

```
## [1] "Critical t is 0.405190207776022."
```

```r
sprintf('Observed t is %s.', qt(.025, n1 + n2-2, lower.tail = F))
```

```
## [1] "Observed t is 2.07387306790403."
```

```r
var.test(caliper1, caliper2)
```

```
##
##  F test to compare two variances
##
## data:  caliper1 and caliper2
## F = 0.47794, num df = 11, denom df = 11, p-value = 0.2364
## alternative hypothesis: true ratio of variances is not equal to 1
## 95 percent confidence interval:
##  0.1375885 1.6602238
## sample estimates:
## ratio of variances
##          0.4779412
```

```r
sprintf('Observed F is %s.', qf(0.025, n1-1, n2-1))
```

```
## [1] "Observed F is 0.287877557984599."
```

```r
# checking results with off-the-shelf R function
# From our F-test above, we could assume equal variance since
# critical F > observed F
t.test(caliper1, caliper2, var.equal = T)
```

```
##
##  Two Sample t-test
##
## data:  caliper1 and caliper2
## t = 0.40519, df = 22, p-value = 0.6893
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -0.001029568  0.001529568
## sample estimates:
## mean of x mean of y
##   0.26625   0.26600
```

Using the two-sided t-test, we found our critical t-statistic of 0.40519 to be less than our observed t-statistic of 2.073873. Therefore, we would fail to reject the null hypothesis and could conclude that there is not a significant difference between the means of the population of measurements from which the two samples were selected.

(b) Find the p-value for the test in part(a).

```r
pval <- 2*pt(t_stat, n1 + n2 - 2, lower.tail = F)
sprintf('p-value is %s.', pval)
```

```
## [1] "p-value is 0.689250891741353."
```

```r
# checking results with off-the-shelf R function
# see explanation in (a) on why equal variance is assumed
t.test(caliper1, caliper2, var.equal = T)
```

```
##
##  Two Sample t-test
##
## data:  caliper1 and caliper2
## t = 0.40519, df = 22, p-value = 0.6893
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -0.001029568  0.001529568
## sample estimates:
## mean of x mean of y
##   0.26625   0.26600
```

According to our t-test above, the p-value for the test in part(a) is 0.6893 which is greater than $\alpha = 0.05$. Therefore, we would fail to reject H_0.

(c) Construct a 95% CI on the difference in mean diameter measurements for the two types of calipers.

Sol.:

```r
diff <- mean(caliper1) - mean(caliper2)
ts <- qt(.025, n1 + n2-2, lower.tail = F)

x <- diff - ts*sp*sqrt(1/n1 + 1/n2)
y <- diff + ts*sp*sqrt(1/n1 + 1/n2)

sprintf('The 95-percent CI is (%.5f, %.5f).', x, y)
```

```
## [1] "The 95-percent CI is (-0.00103, 0.00153)."
```

```r
# checking results with off-the-shelf R function
# see explanation in (a) on why equal variance is assumed
t.test(caliper1, caliper2, var.equal = T)
```

```
##
##  Two Sample t-test
##
## data:  caliper1 and caliper2
## t = 0.40519, df = 22, p-value = 0.6893
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -0.001029568  0.001529568
## sample estimates:
## mean of x mean of y
##   0.26625   0.26600
```

As shown above, the 95% confidence interval (CI) on the difference in mean diameter measurements for the two types of calipers is (-0.00103, 0.00153).