# a3q1

*Mushi Wang*

*11/06/2020*

1.

```r
options(contrasts = c('contr.sum', 'contr.copy'))
dia1 = c(80, 83, 83, 85)
dia2 = c(75, 75, 79, 79)
dia3 = c(74, 73, 76, 77)
dia4 = c(67, 72, 74, 74)
dia5 = c(62, 62, 67, 69)
dia6 = c(60, 61, 64, 66)
y = c(dia1, dia2, dia3, dia4, dia5, dia6)
x = as.factor(c(rep(1, 4), rep(2, 4), rep(3, 4), rep(4, 4), rep(5, 4), rep(6, 4)))
```

(a) We want to use F-test.

$H_0 : \tau_1 = \tau_2 = \tau_3 = \tau_4 = \tau_5 = \tau_6 = 0$, $H_a$ : at least one is not 0

```r
model = lm(y ~ x)
summary(model)
```

```
##
## Call:
## lm(formula = y ~ x)
##
## Residuals:
##    Min     1Q Median     3Q    Max
##  -4.75  -2.00   0.25   2.00   4.00
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)  72.3750     0.5533 130.807  < 2e-16 ***
## x1           10.3750     1.2372   8.386 1.24e-07 ***
## x2            4.6250     1.2372   3.738   0.0015 **
## x3            2.6250     1.2372   2.122   0.0480 *
## x4           -0.6250     1.2372  -0.505   0.6196
## x5           -7.3750     1.2372  -5.961 1.22e-05 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 2.711 on 18 degrees of freedom
## Multiple R-squared:  0.8955, Adjusted R-squared:  0.8665
## F-statistic: 30.85 on 5 and 18 DF,  p-value: 3.16e-08
```

From summary, F-statistic is 30.85 with degree of freedoms of 5 and 18. The p-value is $3.16 \times 10^{-8}$. So we have tons of evidence to reject $H_0$. Hence the size of the orifice affects the mean percentage of radon released.

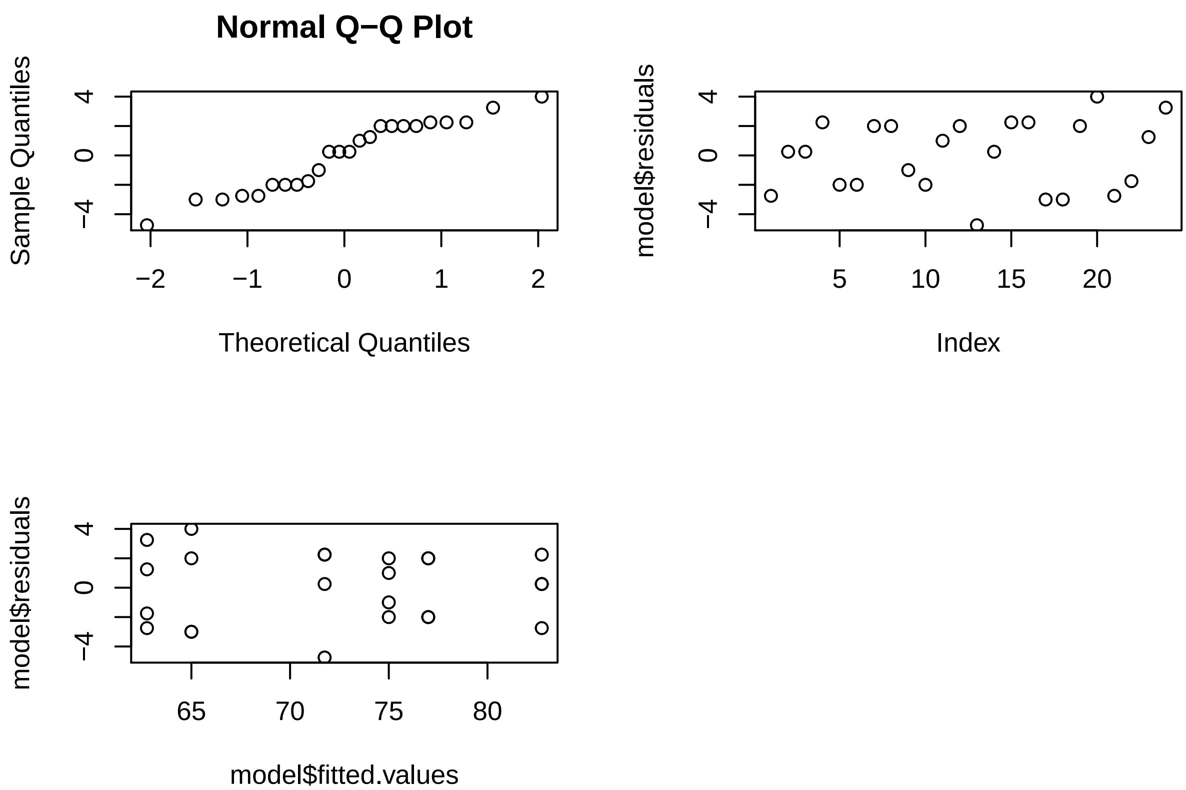
(b)

```r
model$residuals
```

```
##     1     2     3     4     5     6     7     8     9    10    11    12
## -2.75  0.25  0.25  2.25 -2.00 -2.00  2.00  2.00 -1.00 -2.00  1.00  2.00
```

```
##    13    14    15    16    17    18    19    20    21    22    23    24
## -4.75  0.25  2.25  2.25 -3.00 -3.00  2.00  4.00 -2.75 -1.75  1.25  3.25
```

```r
par(mfrow=c(2,2))
qqnorm(model$residuals)
plot(model$residuals)
plot(model$fitted.values, model$residuals)
```




In Q-Q plot, the residuals lie on a straight line reasonably well. And in the other plots, the residuals lies in a band between 0 and relatively random and there is no obvious patterns.

(c) The 95% confidence interval for $\hat{\mu} + \tau_5$ is $\hat{\mu} + \hat{\tau_5} \pm c\sqrt{\frac{\hat{\sigma}^2}{2r}}$ where $c \sim t_{18}$.

```r
qt(0.975, 18)
```

```
## [1] 2.100922
```

The CI is, $72.3750 - 7.3750 \pm 2.1\frac{2.711}{\sqrt{2 \times 4}} = (62.98719, 67.01281)$