MATH 3080 Lab Project 6

Aini Liang

February 15, 2018

- Problem 1

*Remember: I expect to see commentary either in the text, in the code with comments created using `#`, or (preferably) both! **Failing to do so may result in lost points!***

Problem 1

As part of a study of its sheet metal assembly process, a major automobile manufacturer uses sensors that record the deviation from the nominal thickness (millimeters) at 4 locations on a car. Data on 30 cars are given in the file `car_body.csv` (Source: Data Courtesy of Darek Ceglarek).

STEP 1 (Data Preparation) Download the data file `car_body.csv` from our course webpage. Read the file into R, and reshape the data by using `stack()` function. You may use the variable names 'V1, V2, V3, V4' as treatment names. (Hint: see codes of previous lab.)

STEP 2 (ANOVA) Perform an ANOVA to test for the null hypothesis \(H_0\): \(\mu_1=\mu_2=\mu_3=\mu_4\) vs. \(H_A\): \(H_0\) is not true at \(\alpha=0.05\). Check all the appropriate assumptions.

STEP 3 (Tukey's method) Use Tukey's method to find which individual locations seem to show a cause for concern.

```
# Your code here

mydata = read.csv("car_body.csv", header = TRUE)
mydata
```

```
##       V1    V2    V3    V4
## 1  -0.12  0.36  0.40  0.25
## 2  -0.60 -0.35  0.04 -0.28
## 3  -0.13  0.05  0.84  0.61
## 4  -0.46 -0.37  0.30  0.00
## 5  -0.46 -0.24  0.37  0.13
## 6  -0.46 -0.16  0.07  0.10
## 7  -0.46 -0.24  0.13  0.02
## 8  -0.13  0.05 -0.01  0.09
## 9  -0.31 -0.16 -0.20  0.23
## 10 -0.37 -0.24  0.37  0.21
## 11 -1.08 -0.83 -0.81  0.05
## 12 -0.42 -0.30  0.37 -0.58
## 13 -0.31  0.10 -0.24  0.24
## 14 -0.14  0.06  0.18 -0.50
## 15 -0.61 -0.35 -0.24  0.75
## 16 -0.61 -0.30 -0.20 -0.21
## 17 -0.84 -0.35 -0.14 -0.22
## 18 -0.96 -0.85  0.19 -0.18
## 19 -0.90 -0.34 -0.78 -0.15
## 20 -0.46  0.36  0.24 -0.58
## 21 -0.90 -0.59  0.13  0.13
## 22 -0.61 -0.50 -0.34 -0.58
## 23 -0.61 -0.20 -0.58 -0.20
## 24 -0.46 -0.30 -0.10 -0.10
## 25 -0.60 -0.35 -0.45  0.37
## 26 -0.60 -0.36 -0.34 -0.11
## 27 -0.31  0.35 -0.45 -0.10
## 28 -0.60 -0.25 -0.42  0.28
## 29 -0.31  0.25 -0.34 -0.24
## 30 -0.36 -0.16  0.15 -0.38
```

```
stack_all <- stack(mydata)
stack_all
```

```
##     values ind
## 1    -0.12  V1
## 2    -0.60  V1
## 3    -0.13  V1
## 4    -0.46  V1
```

```
## 5     -0.46  V1
## 6     -0.46  V1
## 7     -0.46  V1
## 8     -0.13  V1
## 9     -0.31  V1
## 10    -0.37  V1
## 11    -1.08  V1
## 12    -0.42  V1
## 13    -0.31  V1
## 14    -0.14  V1
## 15    -0.61  V1
## 16    -0.61  V1
## 17    -0.84  V1
## 18    -0.96  V1
## 19    -0.90  V1
## 20    -0.46  V1
## 21    -0.90  V1
## 22    -0.61  V1
## 23    -0.61  V1
## 24    -0.46  V1
## 25    -0.60  V1
## 26    -0.60  V1
## 27    -0.31  V1
## 28    -0.60  V1
## 29    -0.31  V1
## 30    -0.36  V1
## 31     0.36  V2
## 32    -0.35  V2
## 33     0.05  V2
## 34    -0.37  V2
## 35    -0.24  V2
## 36    -0.16  V2
## 37    -0.24  V2
## 38     0.05  V2
## 39    -0.16  V2
## 40    -0.24  V2
## 41    -0.83  V2
## 42    -0.30  V2
## 43     0.10  V2
## 44     0.06  V2
## 45    -0.35  V2
## 46    -0.30  V2
## 47    -0.35  V2
## 48    -0.85  V2
## 49    -0.34  V2
## 50     0.36  V2
## 51    -0.59  V2
## 52    -0.50  V2
## 53    -0.20  V2
## 54    -0.30  V2
## 55    -0.35  V2
## 56    -0.36  V2
## 57     0.35  V2
## 58    -0.25  V2
## 59     0.25  V2
## 60    -0.16  V2
## 61     0.40  V3
## 62     0.04  V3
## 63     0.84  V3
## 64     0.30  V3
## 65     0.37  V3
## 66     0.07  V3
## 67     0.13  V3
## 68    -0.01  V3
## 69    -0.20  V3
## 70     0.37  V3
## 71    -0.81  V3
## 72     0.37  V3
## 73    -0.24  V3
## 74     0.18  V3
## 75    -0.24  V3
## 76    -0.20  V3
## 77    -0.14  V3
```

```
## 78    0.19  V3
## 79   -0.78  V3
## 80    0.24  V3
## 81    0.13  V3
## 82   -0.34  V3
## 83   -0.58  V3
## 84   -0.10  V3
## 85   -0.45  V3
## 86   -0.34  V3
## 87   -0.45  V3
## 88   -0.42  V3
## 89   -0.34  V3
## 90    0.15  V3
## 91    0.25  V4
## 92   -0.28  V4
## 93    0.61  V4
## 94    0.00  V4
## 95    0.13  V4
## 96    0.10  V4
## 97    0.02  V4
## 98    0.09  V4
## 99    0.23  V4
## 100   0.21  V4
## 101   0.05  V4
## 102  -0.58  V4
## 103   0.24  V4
## 104  -0.50  V4
## 105   0.75  V4
## 106  -0.21  V4
## 107  -0.22  V4
## 108  -0.18  V4
## 109  -0.15  V4
## 110  -0.58  V4
## 111   0.13  V4
## 112  -0.58  V4
## 113  -0.20  V4
## 114  -0.10  V4
## 115   0.37  V4
## 116  -0.11  V4
## 117  -0.10  V4
## 118   0.28  V4
## 119  -0.24  V4
## 120  -0.38  V4
```

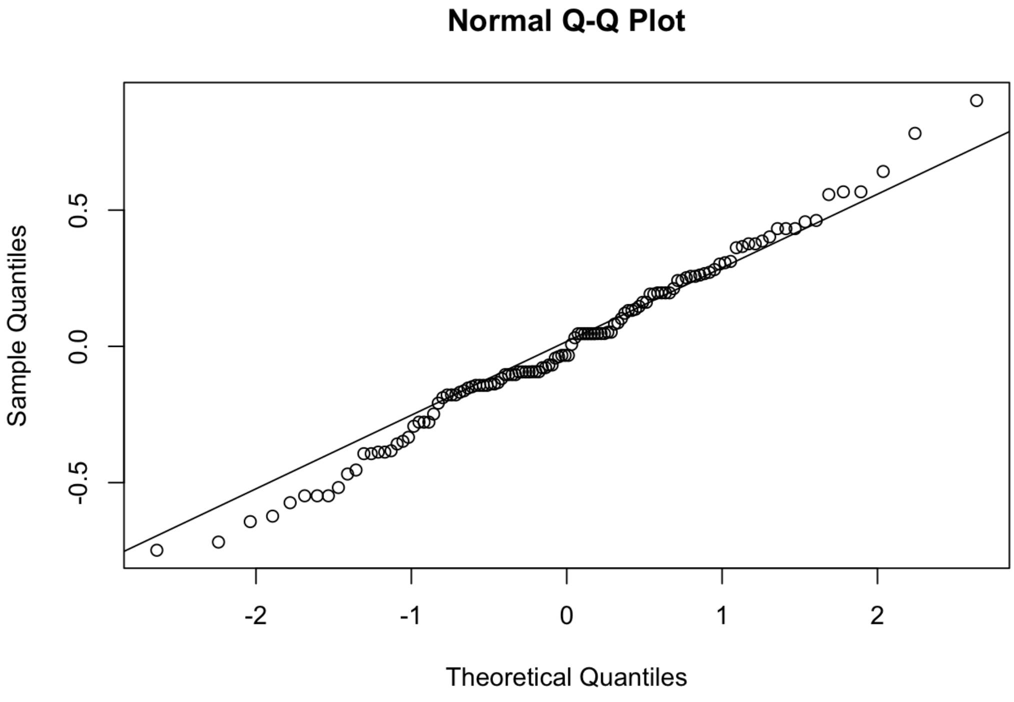
```
names(stack_all) <- c("body", "car")
stack_all
```

```
##       body car
## 1   -0.12  V1
## 2   -0.60  V1
## 3   -0.13  V1
## 4   -0.46  V1
## 5   -0.46  V1
## 6   -0.46  V1
## 7   -0.46  V1
## 8   -0.13  V1
## 9   -0.31  V1
## 10  -0.37  V1
## 11  -1.08  V1
## 12  -0.42  V1
## 13  -0.31  V1
## 14  -0.14  V1
## 15  -0.61  V1
## 16  -0.61  V1
## 17  -0.84  V1
## 18  -0.96  V1
## 19  -0.90  V1
## 20  -0.46  V1
## 21  -0.90  V1
## 22  -0.61  V1
## 23  -0.61  V1
```

```
## 24  -0.46  V1
## 25  -0.60  V1
## 26  -0.60  V1
## 27  -0.31  V1
## 28  -0.60  V1
## 29  -0.31  V1
## 30  -0.36  V1
## 31   0.36  V2
## 32  -0.35  V2
## 33   0.05  V2
## 34  -0.37  V2
## 35  -0.24  V2
## 36  -0.16  V2
## 37  -0.24  V2
## 38   0.05  V2
## 39  -0.16  V2
## 40  -0.24  V2
## 41  -0.83  V2
## 42  -0.30  V2
## 43   0.10  V2
## 44   0.06  V2
## 45  -0.35  V2
## 46  -0.30  V2
## 47  -0.35  V2
## 48  -0.85  V2
## 49  -0.34  V2
## 50   0.36  V2
## 51  -0.59  V2
## 52  -0.50  V2
## 53  -0.20  V2
## 54  -0.30  V2
## 55  -0.35  V2
## 56  -0.36  V2
## 57   0.35  V2
## 58  -0.25  V2
## 59   0.25  V2
## 60  -0.16  V2
## 61   0.40  V3
## 62   0.04  V3
## 63   0.84  V3
## 64   0.30  V3
## 65   0.37  V3
## 66   0.07  V3
## 67   0.13  V3
## 68  -0.01  V3
## 69  -0.20  V3
## 70   0.37  V3
## 71  -0.81  V3
## 72   0.37  V3
## 73  -0.24  V3
## 74   0.18  V3
## 75  -0.24  V3
## 76  -0.20  V3
## 77  -0.14  V3
## 78   0.19  V3
## 79  -0.78  V3
## 80   0.24  V3
## 81   0.13  V3
## 82  -0.34  V3
## 83  -0.58  V3
## 84  -0.10  V3
## 85  -0.45  V3
## 86  -0.34  V3
## 87  -0.45  V3
## 88  -0.42  V3
## 89  -0.34  V3
## 90   0.15  V3
## 91   0.25  V4
## 92  -0.28  V4
## 93   0.61  V4
## 94   0.00  V4
## 95   0.13  V4
## 96   0.10  V4
```

```
## 97   0.02  V4
## 98   0.09  V4
## 99   0.23  V4
## 100  0.21  V4
## 101  0.05  V4
## 102 -0.58  V4
## 103  0.24  V4
## 104 -0.50  V4
## 105  0.75  V4
## 106 -0.21  V4
## 107 -0.22  V4
## 108 -0.18  V4
## 109 -0.15  V4
## 110 -0.58  V4
## 111  0.13  V4
## 112 -0.58  V4
## 113 -0.20  V4
## 114 -0.10  V4
## 115  0.37  V4
## 116 -0.11  V4
## 117 -0.10  V4
## 118  0.28  V4
## 119 -0.24  V4
## 120 -0.38  V4
```

```
# STEP 2 (ANOVA) Perform an ANOVA to test for the null hypothesis Ho:
# mu1=mu2=mu3=mu4 vs. Ha: Ho is not true at alpha=0.05. Check all the
# appropriate assumptions.
attach(mydata)
dev = c(V1 - mean(V1), V2 - mean(V2), V3 - mean(V3), V4 - mean(V4))
dev = sort(dev)
qqnorm(dev)  # a built-in function to generate qq-plots
qqline(dev)  # add a straight line to the qq-plots
```




```
shapiro.test(V1)
```

```
##
##  Shapiro-Wilk normality test
##
## data:  V1
## W = 0.94164, p-value = 0.1007
```

```
shapiro.test(V2)  ## we donot sure how to determine this group if the Normality by step by step test.
```

```
##
##  Shapiro-Wilk normality test
##
## data:  V2
## W = 0.93309, p-value = 0.05934
```

```
shapiro.test(V3)
```

```
##
##  Shapiro-Wilk normality test
##
## data:  V3
## W = 0.97856, p-value = 0.7862
```

```
shapiro.test(V4)
```

```
##
##  Shapiro-Wilk normality test
##
## data:  V4
## W = 0.97124, p-value = 0.5737
```

```
# it is Normality. If it is not normality,we can not to do anova

detach(mydata)
bartlett.test(body ~ car, data = stack_all)
```

```
##
##  Bartlett test of homogeneity of variances
##
## data:  body by car
## Bartlett's K-squared = 5.0695, df = 3, p-value = 0.1668
```

```
# P-value is larger than a=0.05, so the null that the variances in each of
# the groups (samples) are the same

# if the varuances are the same, we can use anova to analysis the data.
# Then, we can use Tukey for more detail.

anova <- aov(body ~ car, data = stack_all)
summary(anova)
```

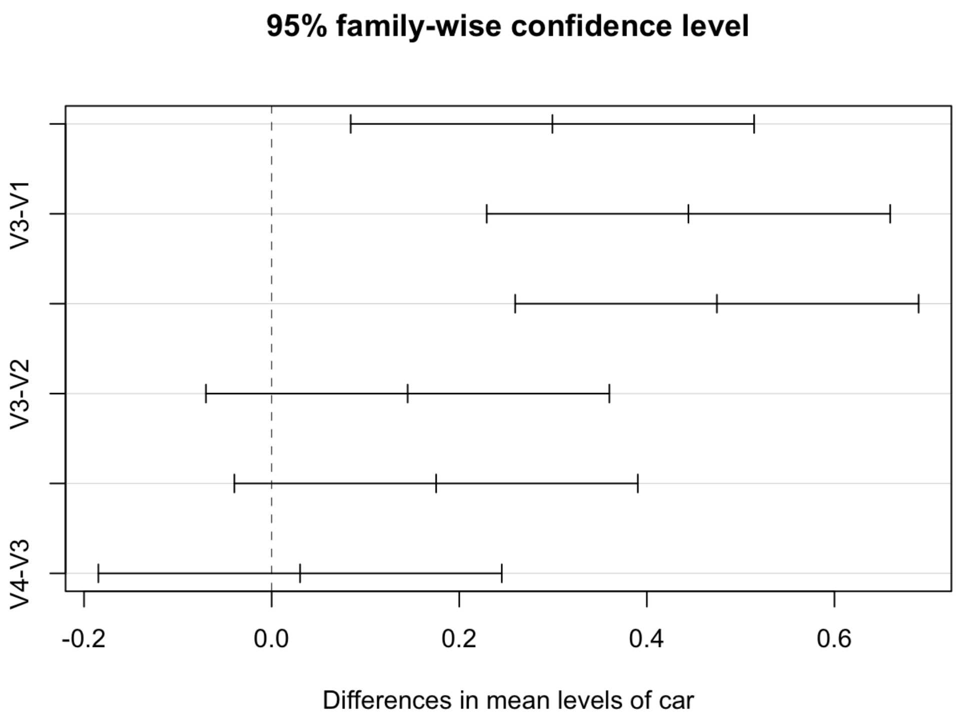
```
##              Df Sum Sq Mean Sq F value   Pr(>F)
## car           3  4.238  1.4126   13.84 8.94e-08 ***
## Residuals   116 11.841  0.1021
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

```
# The p-value is less than a=0.05, so we reject H0.

# STEP 3 (Tukey's method) Use Tukey's method to find which individual
# locations seem to show a cause for concern
tk <- TukeyHSD(anova, ordered = T)
tk
```

```
##   Tukey multiple comparisons of means
##     95% family-wise confidence level
##     factor levels have been ordered
##
## Fit: aov(formula = body ~ car, data = stack_all)
##
## $car
##              diff         lwr       upr     p adj
## V2-V1 0.29933333  0.08429847 0.5143682 0.0023727
## V3-V1 0.44433333  0.22929847 0.6593682 0.0000023
## V4-V1 0.47466667  0.25963180 0.6897015 0.0000004
## V3-V2 0.14500000 -0.07003486 0.3600349 0.2990726
## V4-V2 0.17533333 -0.03970153 0.3903682 0.1511715
## V4-V3 0.03033333 -0.18470153 0.2453682 0.9829339
```

```
plot(tk)
```




```
# V1 ~ v2, v3 and v4 are significant difference.  V2 ~ v3 and v4 are
# siginificant.  V3 ~ v4 is siginificant.
```