QSCI 482 Story 7

Simon Hans Edasi

2023-11-28

1a. Read the data into an R object and use the table() function to check that you have a balanced two-factor ANOVA design.

Balanced design will have same n for each factor.

```
df <- read.csv('7 Guineapigdata.csv')
df$supp <- as.factor(df$supp)
df$dose <- as.factor(df$dose)
table(df$supp)
```

```
##
## OrangeJuice        Pill
##          30          30
```

```
table(df$dose)
```

```
##
## Dose0.5   Dose1   Dose2
##      20      20      20
```

1b. For the rest of the assignment (for grading simplicity), choose the dose to be Factor A and the supplement type to be Factor B. Test the underlying assumptions of the two-factor ANOVA, using Levene's Test (short way around) for homogeneity of variances. Write down your conclusions.

Underlying assumptions:

1. Variances are the same

H_0 : Variances across factors are equal

H_A : Variances across factors are not equal

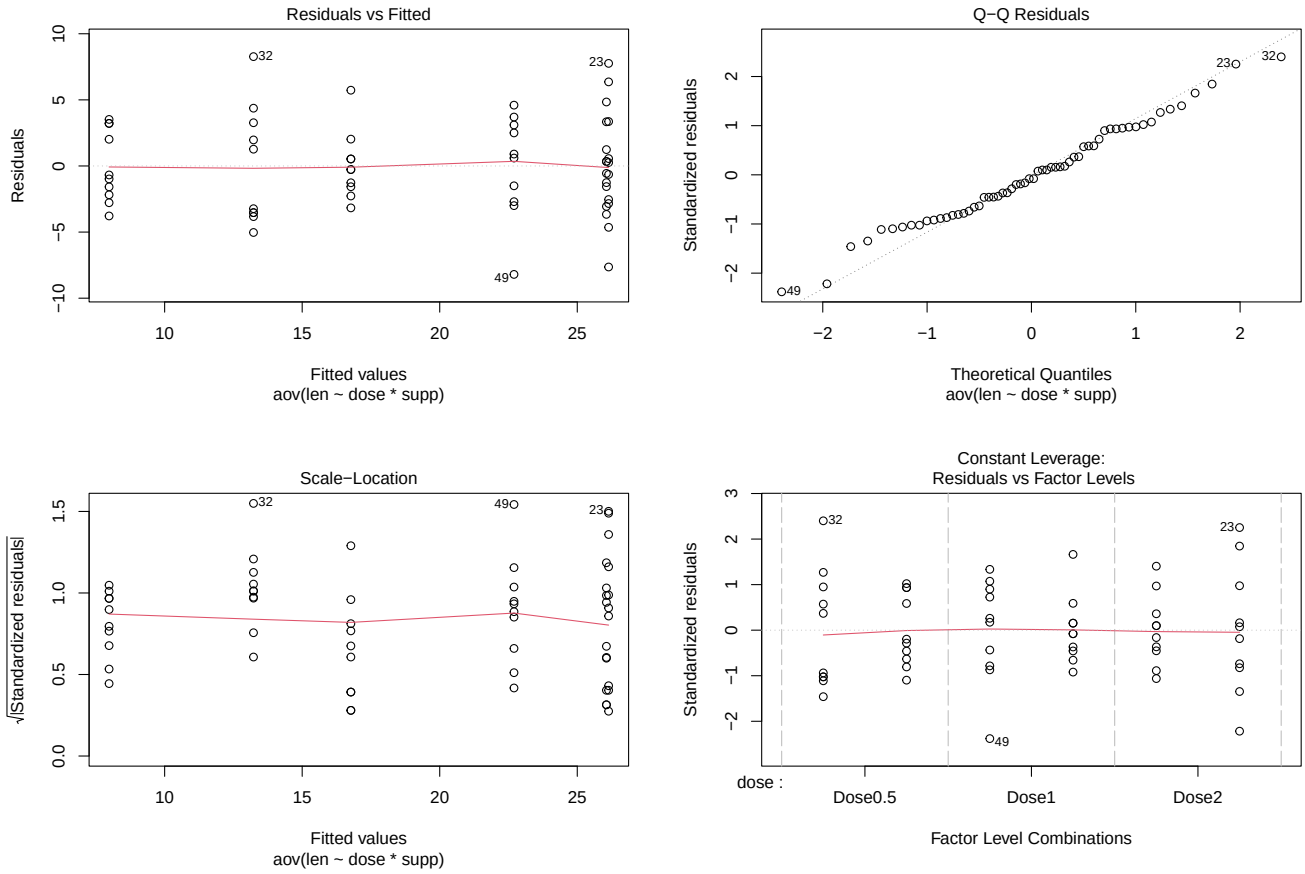
```
leveneTest(len ~ dose*supp,data = df)
```

```
## Levene's Test for Homogeneity of Variance (center = median)
##       Df F value Pr(>F)
## group  5  1.7086 0.1484
##       54
```

We fail to reject the null hypothesis that variances are equal.

2. Data in each cell comes from a normally distributed population

```
aov.res <- aov(len ~ dose*supp, data = df)
plot(aov.res)
```



Residuals appear normally distributed, we appear to be good here.

1c. Run a two-factor ANOVA on the data the short way around. Copy and paste the output table in R that includes the degrees of freedom, mean square, p-values etc. Interpret the resulting p-values in your own words.

```
summary(aov.res)
```

```
##             Df Sum Sq Mean Sq F value   Pr(>F)
## dose         2 2426.4  1213.2  92.000  < 2e-16 ***
## supp         1  205.3   205.3  15.572 0.000231 ***
## dose:supp    2  108.3    54.2   4.107 0.021860 *
## Residuals   54  712.1    13.2
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

1d. Provide the plot for standardized residuals, the Q-Q plot, and run a Shapiro-Wilks test (short-way around), to check whether the data are really normally distributed. Paste the plots into this document, and explain your conclusions about whether the data are normally distributed.