Statistics 512: Solution to Homework#8

Use the dataset from Problem 16.11 described on page 725 of KNNL (`ch16pr11a.dat`). This analysis will be continued on the next homework set, so be sure to save your SAS code.

1. Give a table of sample sizes, means, and standard deviations for the six different filling machines.

    Solution: The SAS output is as follows:

    ```
                            The MEANS Procedure
                         Analysis Variable : wtdev
                 N
    machine     Obs     N        Mean         Std Dev        Minimum        Maximum
    ----------------------------------------------------------------------------------
       1         20     20     0.0735000      0.1925255     -0.2700000      0.3900000
       2         20     20     0.1905000      0.1854007     -0.1200000      0.5300000
       3         20     20     0.4600000      0.1646368      0.2000000      0.7800000
       4         20     20     0.3655000      0.1653537      0.0100000      0.5800000
       5         20     20     0.1250000      0.1727334     -0.2000000      0.3500000
       6         20     20     0.1515000      0.1734867     -0.0900000      0.4700000
    ----------------------------------------------------------------------------------
    ```

    (a) Based on this table, does the constant variance assumption appear to hold?

    Solution: The standard deviations range from 0.165 to 0.192 which is not a very big range; thus the constant variance assumption appears to hold.

    (b) Based on this table, do the six machines appear to give the same average amount of detergent?

    Solution: The means range from 0.07 to 0.46 which is quite a big range, so it appears there are some differences among the machines.

2. Make a plot of the means (with `i=join`) overlaid on a plot of the individual observations (with `i=none`) versus filling machine number.

    Solution: See Figure 1.

    (a) Based on this plot, does the constant variance assumption appear to hold?

    Solution: Yes, the variance appears approximately constant.

    (b) Based on this plot, do the six machines appear to give the same average amount of detergent?

    Solution: No, the means appear to vary.

3. Examine the question of whether or not the six machines give the same average amount of detergent.

    (a) Write the cell means model for this analysis. State the null and alternative hypotheses in terms of the cell means model parameters, give the test statistic with degrees of freedom, the p-value and your conclusion.
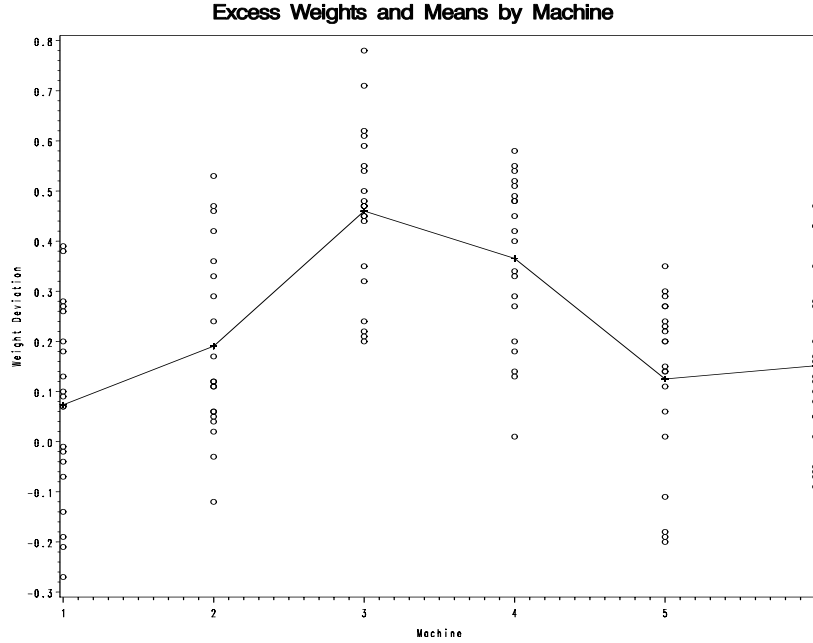
Figure 1: Scatterplot for Problem 2

Solution: The model is

$$Y_{i,j} = \mu_i + \epsilon_{i,j}.$$

The null hypothesis is

$$H_0 : \mu_1 = \mu_2 = \mu_3 = \mu_4 = \mu_5 = \mu_6,$$

while the alternative hypothesis is

$$H_A : \text{ at least one of the } \mu_i \text{ is not the same.}$$

The F-ratio is 14.78 which, when compared to an F-distribution with 5 and 114 degrees of freedom, gives a p-value of 3.7×10^{-11}. As a result, we reject H_0 and conclude that at least one of the machines has a different mean from the others.

(b) Write the factor effects model for this analysis. State the null and alternative hypotheses in terms of the factor effects model parameters, give the test statistic with degrees of freedom, the p-value and your conclusion.

Solution: The model is

$$Y_{i,j} = \mu + \tau_i + \epsilon_{i,j}.$$

The null hypothesis is

$$H_0 : \tau_1 = \tau_2 = \tau_3 = \tau_4 = \tau_5 = \tau_6 = 0,$$

while the alternative hypothesis is

$$H_A : \text{ at least one of the } \tau_i \text{ is not zero.}$$

2

The F-ratio is 14.78 which, when compared to an F-distribution with 5 and 114 degrees of freedom, gives a p-value of 3.7×10^{-11}. As a result, we reject H_0 and conclude that at least one of the machines has a different mean from the others.

```
                              The GLM Procedure
Dependent Variable: wtdev
                                 Sum of
Source                    DF     Squares      Mean Square    F Value    Pr > F
Model                      5    2.28934667     0.45786933     14.78     <.0001
Error                    114    3.53060000     0.03097018
Corrected Total          119    5.81994667

             R-Square     Coeff Var      Root MSE     wtdev Mean
             0.393362     77.29873       0.175983      0.227667
```