STAT 481- Project 2

Jakub Dlugosz

1 Introduction

We want to establish a model where Oil Pattern(Factor A) is fixed and Bowler(Factor B) is random.

2 Model

$$Y_{ijk} = \mu_i + \alpha_i + \beta_j + (\alpha\beta)_{ij} + \epsilon_{ijk} \tag{1}$$

$$Score = \mu_i + OilPattern + Bowler + (OilPattern * Bowler) + \epsilon_{ijk} \tag{2}$$

3 Constraints

$$\sum_{n=1}^{\alpha} \alpha i = 0 \tag{3}$$

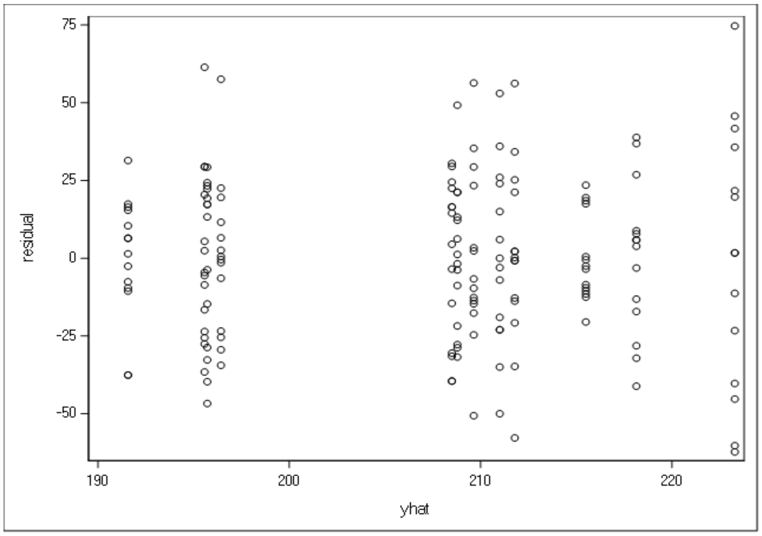
$$\beta_j \overset{iid}{\sim} N(0, \frac{a-1}{a}\sigma^2 \alpha_\beta) \tag{4}$$

$$\epsilon \overset{iid}{\sim} N(0, \sigma^2) \tag{5}$$

$$(\alpha\beta)_{ij} \overset{iid}{\sim} N(0, \sigma^2) \tag{6}$$

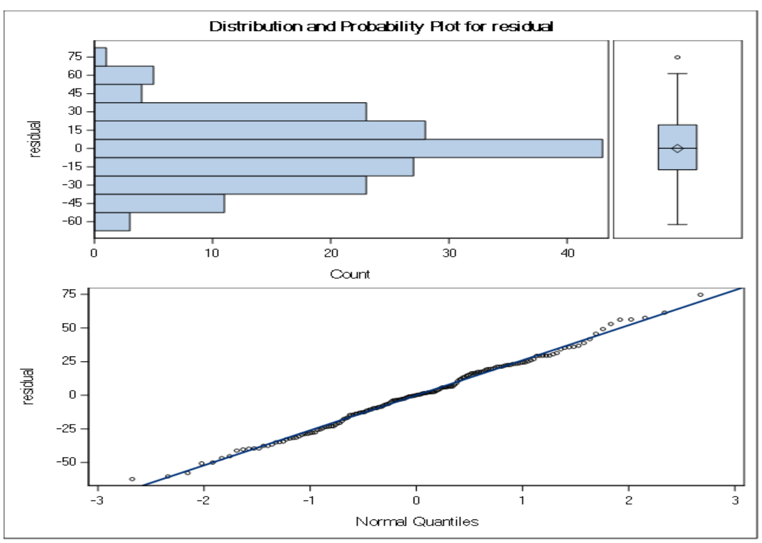
$$\beta_j, (\alpha\beta)_{ij}, \epsilon_{ijk}, \text{are pairwise independent} \tag{7}$$

4 Equal Variance Check



Based on the scatter plot above, we see that the residuals are plotted against fitted values. We can see that the points on the scatter plot are random with no pattern. Because of this randomness of points, we can conclude that this model satisfies the equal variances assumption.

5 Normality Check



Based on the Q-Q plot for the residuals, we see that most of the points fall on the line for this plot.
Shapiro-Wilk Test:

$$H_0 : \epsilon_{ijk} = 0 \text{ versus } H_1 : \epsilon_{ijk} = 0 \tag{8}$$

Tests for Normality

Test		Statistic		p-Values
Shapiro-Wilk	W	0.994644	Pr <W	0.8041
Kolmogorov-Smirnov	D	0.039577	Pr >D	>0.1500
Cramer-von Mises	W-Sq	0.039696	Pr >W-Sq	>0.2500
Anderson-Darling	A-Sq	0.259462	Pr >A-Sq	>0.2500

Based on the p-value of the Shapiro-Wilk Test, which in this case it is 0.8041, since the p-value > alpha= 0.05, we fail to reject H_0. In other words, we can conclude that the residuals for this model are normally distributed.

6 Hypothesis Testing for Main Effects and Interactions

Source	DF	Type I SS	Mean Square	F Value	Pr >F
Oil_Pattern	3	6925.827381	2308.609127	3.18	0.0256
Bowler	2	6069.142857	3034.571429	4.18	0.0170
Oil_Pattern*Bowler	6	2691.047619	448.507937	0.62	0.7155

-Hypothesis Testing for Interactions:

$$H_0 : \sigma^2_{\alpha\beta} = 0 \text{ versus } H_1 : \sigma^2_{\alpha\beta} > 0 \tag{9}$$

Test Statistic: F= 0.062
P-Value=0.7155
Since our p-value is greater than alpha=0.05, 0.07155>0.05, we fail to reject H_0. We can conclude that there is not enough evidence to suggest that there are significant interactions effects between Oil Patterns and Bowlers.

-Hypothesis Testing for Factor B: Bowler

$$H_0 : \sigma^2_{\beta} = 0 \text{ versus } H_1 : \sigma^2_{\beta} > 0 \tag{10}$$

Test Statistic: F=4.18
P-Value= 0.0170
Since our p-value<alpha, 0.0170 < 0.05, we can reject H_0 in favor of H_1. We can conclude that there are factor B main effects present in this model.

Tests of Hypotheses for Oil_Pattern*Bowler					
Source	DF	Type III SS	Mean Square	F Value	Pr >F
Oil_Pattern	3	6925.827381	2308.609127	5.15	0.0426

-Hypothesis Testing for Main Factor A :Oil Pattern

$$H_0: \text{ all } \alpha_i = 0 \text{ versus } H_1: \text{ at least one } \alpha_i \neq 0 \tag{11}$$

Test Statistic: F=5.15
P-value= 0.0426
Since out p-value < alpha, 0.0426 < 0.05, we reject H_0 in favor of H_1. We can conclude that there are factor A main effects present for this model.

7 Point Estimates for Variances

$$\hat{\sigma_\beta}^2 = s_\beta^2 = \frac{MSB - MSE}{na} = \frac{3034.571429 - 725.4785}{(14)(3)} = 54.97 \tag{12}$$

$$\hat{\sigma}_{\alpha\beta}^2 = \frac{MSAB - MSE}{n} = \frac{448.507937 - 725.4785}{14} = -19.78 \approx 0 \tag{13}$$

As we can see $\hat{\sigma}_\beta^2 > \hat{\sigma}_{\alpha\beta}^2$, we can conclude that factor B has a greater effect than the interaction term.

8 Conclusions

-Based on the hypothesis tests we performed, we can conclude that there are appears to be no significant effects with the interaction terms. However, we do see that our fixed effect, Oil Pattern, has significant effect on the model along with our random variable, Bowler, has a significant effect on this mixed model. In other words, the oil patterns have a great significant impact on the score for each game. While the bowlers have an impact on score in terms of greater variance on the scores per game since it is a random effect.

9 Code

```
PROC Import  Datafile="/folders/myshortcuts/myfolder/project2.xlsx"
                                  Out=WORK.Game
                                  DBMS=XLSX
                                  Replace;
Run;
```

```
Proc Print Data=WORK.Game; Run;
PROC GLM DATA = Work.Game;
CLASS Oil_Pattern Bowler;
MODEL Score= Oil_Pattern Bowler Oil_Pattern*Bowler;
RANDOM Bowler Oil_Pattern*Bowler;
TEST H = Oil_Pattern E = Oil_Pattern*Bowler;

RUN;

Proc GLM Data=WORK.Game;
        Class Oil_Pattern Bowler;
        MODEL Score= Oil_Pattern Bowler Oil_Pattern*Bowler;
        Random Bowler Oil_Pattern*Bowler;
        Output Out=results R= residual P=yhat;
Run;

Proc SGPLOT Data = results;
        Scatter X= yhat Y= residual;
Run;

PROC UNIVARIATE DATA = results NORMAL PLOT;
        VAR residual;
RUN;
```