# STAT 6338: ADVANCED STATISTICAL METHODS II

Homework 4

## 1 Problem 20.2

a) Form Figure 1, it can be observed that there are main effect for *week* (Factor A) and *location* (Factor B). This is because location 1 & 4 are above location 2 & 3, and final week values are higher than that of midterm week. It also appears that there could be a slight interaction between week and *location*. However, since there is only one observation per treatment, the moderate lack of parallelism in the response lines could simply be the result of random effects within each treatment cell.

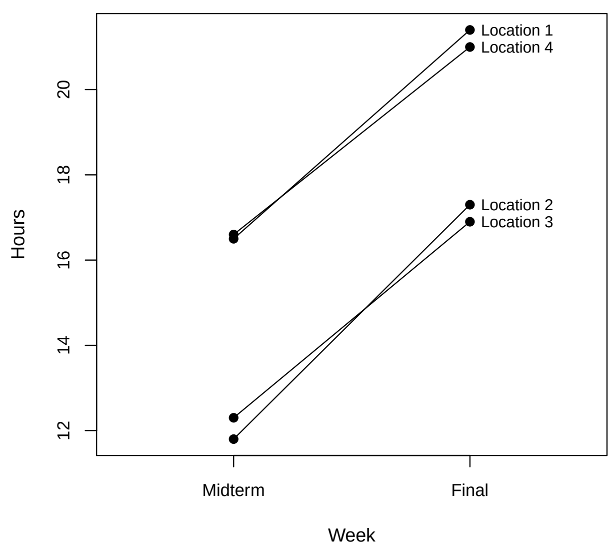


Figure 1: *Plot of number of hours each terminal was not in use during the week at the four locations and for the two different weeks*

b) ANOVA Table:

| **Source** | **df** | **SS** | **MS** | ***F* value** | ***p*-value** |
|---|---|---|---|---|---|
| Location (A) | 3 | 37.005 | 12.335 | 107.26 | $1.503 \times 10^{-3}$ |
| Week (B) | 1 | 47.045 | 47.045 | 409.09 | $2.642 \times 10^{-4}$ |
| Residuals | 3 | 0.345 | 0.115 | | |
| Total | 7 | 84.395 | | | |

**Test for Location Effect**

Hypotheses:

$H_0 : \alpha_1 = \alpha_2 = \alpha_3\alpha_4 = 0$ (Location has no effect on number of hours)

$H_a :$ at least one $\alpha_i \neq 0$ for $i = 1, 2, 3, 4$ (Location has effect)

at $\alpha = 5\%$ significance level.

Test Statistics:

$$F^* = \frac{MSA}{MSAB} = \frac{12.335}{0.115} = 107.26$$

Decision: If $F^* > F_{0.95,3,3} = 9.2766$ reject $H_0$, otherwise fail to reject $H_0$.

Conclusion: Since $F^* > 9.2766$, we reject $H_0$ and conclude that location has effect on number of hours each terminal was not in use during the week, at 5% significance level. $p - value = 1.503 \times 10^{-3}$.

**Test for Week Effect**

Hypotheses:

$H_0 : \beta_1 = \beta_2 = 0$ (Week has no effect on number of hours)

$H_a :$ at least one $\beta_i \neq 0$ for $i = 1, 2$ (Week has effect)

at $\alpha = 5\%$ significance level.

Test Statistics:

$$F^* = \frac{MSB}{MSAB} = \frac{47.045}{0.115} = 409.09$$

Decision: If $F^* > F_{0.95,1,3} = 10.1280$ reject $H_0$, otherwise fail to reject $H_0$.

Conclusion: Since $F^* > 10.1280$, we reject $H_0$ and conclude that week has effect on number of hours each terminal was not in use, at 5% significance level. $p - value = 2.642 \times 10^{-4}$.

Upper bound for the family level of significance, using Kimball inequality

$$\alpha \leq 1 - (1 - \alpha)^2 = 1 - (0.95)^2 = 0.0975$$

## 2 Problem 20.4

$$Y_{ij} = \mu_{..} + \alpha_i + \beta_j + D\alpha_i\beta_j + \varepsilon_{ij}$$

$$\hat{D} = \frac{\sum_i \sum_j (\bar{Y}_{i\cdot} - \bar{Y}_{\cdot\cdot})(\bar{Y}_{\cdot j} - \bar{Y}_{\cdot\cdot})Y_{ij}}{\sum_i (\bar{Y}_{i\cdot} - \bar{Y}_{\cdot\cdot})^2 \sum_j (\bar{Y}_{\cdot j} - \bar{Y}_{\cdot\cdot})^2} = \frac{-4.1346}{217.6125} = -0.019$$

$$SSAB^* = \sum_i \sum_j \hat{D}^2 (\bar{Y}_{i\cdot} - \bar{Y}_{\cdot\cdot})^2 \sum_j (\bar{Y}_{\cdot j} - \bar{Y}_{\cdot\cdot})^2 = \frac{(-4.1346)^2}{217.6125} = 0.0786$$

$$SSRem^* = SSTO - SSA - SSB - SSAB^* = 84.395 - 37.005 - 47.045 - 0.0786 = 0.2664$$

Hypotheses:
$H_0 : D = 0$ (No interaction present)
$H_a : D \neq 0$ (Interactions $D\alpha_i\beta_j$ present)
at 2.5% significance level.
Test Statistic:

$$F^* = \frac{SSAB^*/1}{SSRem^*/(ab - a - b)} = \frac{0.0786}{0.2664/2} = 0.59$$

Decision Rule: If $F^* > F_{(0.975,1,2)} = 38.5063$, then reject $H_0$, otherwise fail to reject $H_0$.
Conclusion: Since $F^* < 38.5063$, we fail to reject the null hypothesis and conclude that there is no interaction between the location of the terminal and the week of the semester.

If the additive model is not appropriate (Tueky test indicates the presence of interaction effects), interaction effects should be removed or made unimportant by transforming the response variable. If no transformation can be found to make the interactions unimportant, an approximate method of analysis can be employed

## 3 Problem 23.6

a) ANOVA model:

$$Y_{ijk} = \mu_{\cdot\cdot} + \alpha_i + \beta_j + (\alpha\beta)_{ij} + \varepsilon_{ijk}$$

Equivalent regression model:

$$Y_{ijk} = \mu_{\cdot\cdot} + \alpha_1 X_{ijk1} + \alpha_2 X_{ijk2} + \beta_1 X_{ijk3} + (\alpha\beta)_{11} X_{ijk1} X_{ijk3} + (\alpha\beta)_{21} X_{ijk2} X_{ijk3} + \varepsilon_{ijk}$$

$$X_{ijk1} = \begin{cases} 1, & \text{if case from level 1 for factor A (Young)} \\ -1, & \text{if case from level 3 for factor A (Elderly)} \\ 0, & \text{otherwise} \end{cases}$$

$$X_{ijk2} = \begin{cases} 1, & \text{if case from level 2 for factor A (Middle)} \\ -1, & \text{if case from level 3 for factor A (Elderly)} \\ 0, & \text{otherwise} \end{cases}$$

$$X_{ijk3} = \begin{cases} 1, & \text{if case from level 1 for factor B (Male)} \\ -1, & \text{if case from level 2 for factor B (Female)} \end{cases}$$

b)

$$\boldsymbol{\beta} = \begin{bmatrix} \mu_{..} & \alpha_1 & \alpha_2 & \beta_1 & (\alpha\beta)_{11} & (\alpha\beta)_{21} \end{bmatrix}'$$

| $A$ | $B$ | Freq | | $X_{ijk1}$ | $X_{ijk2}$ | $X_{ijk3}$ | $X_{ijk1}X_{ijk3}$ | $X_{ijk2}X_{ijk3}$ |
|---|---|---|---|---|---|---|---|---|
| 1 | 1 | 6 | 1 | 1 | 0 | 1 | 1 | 0 |
| 1 | 2 | 6 | 1 | 1 | 0 | -1 | -1 | 0 |
| 2 | 1 | 5 | 1 | 0 | 1 | 1 | 0 | 1 |
| 2 | 2 | 6 | 1 | 0 | 1 | -1 | 0 | -1 |
| 3 | 1 | 6 | 1 | -1 | -1 | 1 | -1 | -1 |
| 3 | 2 | 5 | 1 | -1 | -1 | -1 | 1 | 1 |

c) $X\beta$:

| $A$ | $B$ | $X_{ij}\boldsymbol{\beta}$ | | |
|---|---|---|---|---|
| 1 | 1 | $\mu_{..} + \alpha_1 + \beta_1 + (\alpha\beta)_{11}$ | | |
| 1 | 2 | $\mu_{..} + \alpha_1 - \beta_1 - (\alpha\beta)_{11}$ | $=$ | $\mu_{..} + \alpha_1 + \beta_2 + (\alpha\beta)_{12}$ |
| 2 | 1 | $\mu_{..} + \alpha_2 + \beta_1 + (\alpha\beta)_{12}$ | | |
| 2 | 2 | $\mu_{..} + \alpha_2 - \beta_1 - (\alpha\beta)_{12}$ | $=$ | $\mu_{..} + \alpha_2 + \beta_2 + (\alpha\beta)_{22}$ |
| 3 | 1 | $\mu_{..} - \alpha_1 - \alpha_2 + \beta_1 - (\alpha\beta)_{11} - (\alpha\beta)_{12}$ | $=$ | $\mu_{..} + \alpha_3 + \beta_1 + (\alpha\beta)_{31}$ |
| 3 | 2 | $\mu_{..} - \alpha_1 - \alpha_2 - \beta_1 + (\alpha\beta)_{11} + (\alpha\beta)_{12}$ | $=$ | $\mu_{..} + \alpha_3 + \beta_2 + (\alpha\beta)_{32}$ |

d) The reduced regression model for testing for interaction effects:

$$Y_{ijk} = \mu_{..} + \alpha_1 X_{ijk1} + \alpha_2 X_{ijk2} + \beta_1 X_{ijk3} + \varepsilon_{ijk}$$

e) The fitted **full model** is

$$\hat{Y} = 23.5667 - 2.0667X_1 + 4.1667X_2 + 0.3667X_3 - 0.2X_1X_3 - 0.3X_2X_3$$

and $SSE(F) = 71.333$ with $df_F = 28$.
The fitted **reduced model** is

$$\hat{Y} = 23.5909 - 2.0909X_1 + 4.1691X_2 + 0.3602X_3$$

and $SSE(R) = 75.521$ with $df_R = 30$.

Hypothesis:
$H_0 : (\alpha\beta)_{11} = (\alpha\beta)_{21} = 0$ (There is no interaction effect)

$H_a : (\alpha\beta)_{11} \neq 0$ or $(\alpha\beta)_{21} \neq 0$ (There is interaction effect)
at $\alpha = 5\%$ significance level.
Test Statistic:

$$F^* = \frac{(SSE(R) - SSE(F))/(df_R - df_F)}{SSE(F)/df_F} = \frac{(75.521 - 71.333)/(30 - 28)}{71.333/28} = 0.8219$$

Decision Rule: If $F^* > F_{(0.95,2,28)} = 3.3404$, we reject $H_0$, otherwise we fail to reject $H_0$.
Conclusion: Since $F^* < 3.3404$, we fail to reject $H_0$ and conclude that, at 5% significance level, there is no interaction effect between age and gender. The $p - value = 0.4499$.

f) Fitted reduced model to test **age effect**:

$$\hat{Y} = 23.5 + 0.1768X_3 - 0.0101X_1X_3 - 0.4949X_2X_3$$

and $SSE(R) = 359.94$ with $df_R = 30$.

Hypothesis:
$H_0 : \alpha_1 = \alpha_2 = 0$ (There is no age main effect)
$H_a : \alpha_1 \neq 0$ or $\alpha_2 \neq 0$ (There is age main effect)
at $\alpha = 5\%$ significance level.
Test Statistic:

$$F^* = \frac{(SSE(R) - SSE(F))/(df_R - df_F)}{SSE(F)/df_F} = \frac{(359.94 - 71.333)/(30 - 28)}{71.333/28} = 56.6428$$

Decision Rule: If $F^* > F_{(0.95,2,28)} = 3.3404$, we reject $H_0$, otherwise we fail to reject $H_0$.
Conclusion: Since $F^* > 3.3404$, we reject $H_0$ and conclude that, at 5% significance level, there is age main effect on the cash offer. The $p - value = 1.4416 \times 10^{-10}$.

Fitted reduced model to test **gender effect**:

$$\hat{Y} = 23.5667 - 2.0667X_1 + 4.1323X_2 - 0.1771X_1X_3 - 0.3115X_2X_3$$

and $SSE(R) = 75.871$ with $df_R = 29$.

Hypothesis:
$H_0 : \beta_1 = 0$ (There is no gender main effect)
$H_a : \beta_1 \neq 0$ (There is gender main effect)
at $\alpha = 5\%$ significance level.

Test Statistic:

$$F^* = \frac{(SSE(R) - SSE(F))/(df_R - df_F)}{SSE(F)/df_F} = \frac{(75.871 - 71.333)/(29 - 28)}{71.333/28} = 1.7813$$

Decision Rule: If $F^* > F_{(0.95,1,28)} = 4.196$, we reject $H_0$, otherwise we fail to reject $H_0$.
Conclusion: Since $F^* < 4.196$, we fail to reject $H_0$ and conclude that, at 5% significance level, there is no gender main effect on the cash offer. The $p - value = 0.187$.

g) $\hat{\alpha}_1 = -2.0667$, $\hat{\alpha}_2 = 4.1667$, $s^2\{\hat{\alpha}_1\} = 0.1462$, $s^2\{\hat{\alpha}_2\} = 0.1534$, $s\{\hat{\alpha}_1, \hat{\alpha}_2\} = -0.0734$

$$\begin{aligned}
\hat{D}_1 &= \hat{\alpha}_1 - \hat{\alpha}_2 = -6.2334 \\
\hat{D}_2 &= \hat{\alpha}_1 - \hat{\alpha}_3 = \hat{\alpha}_1 - (-\hat{\alpha}_1 - \hat{\alpha}_2) = 2\hat{\alpha}_1 + \hat{\alpha}_2 = 0.0333 \\
\hat{D}_3 &= \hat{\alpha}_2 - \hat{\alpha}_3 = \hat{\alpha}_2 - (-\hat{\alpha}_1 - \hat{\alpha}_2) = \hat{\alpha}_1 + 2\hat{\alpha}_2 = 6.2667 \\
s^2\{\hat{D}_1\} &= s^2\{\hat{\alpha}_1\} + s^2\{\hat{\alpha}_2\} - 2s\{\hat{\alpha}_1, \hat{\alpha}_2\} = 0.1462 + 0.1534 - 2(-0.0734) = 0.4463 \\
s^2\{\hat{D}_2\} &= 4s^2\{\hat{\alpha}_1\} + s^2\{\hat{\alpha}_2\} + 4s\{\hat{\alpha}_1, \hat{\alpha}_2\} = 4(0.1462) + 0.1534 + 4(-0.0734) = 0.4448 \\
s^2\{\hat{D}_3\} &= s^2\{\hat{\alpha}_1\} + 4s^2\{\hat{\alpha}_2\} + 4s\{\hat{\alpha}_1, \hat{\alpha}_2\} = 0.1462 + 4(0.1534) + 4(-0.0734) = 0.4461
\end{aligned}$$

$$\begin{aligned}
\text{Bonferroni: } B &= t_{(1-\alpha/2g; n_T - ab)} = t_{(0.9833;28)} = 2.2383 \\
\text{Scheffé: } S &= \sqrt{(a-1)F_{(1-\alpha; a-1, n_T - ab)}} = \sqrt{2F_{(0.9;2,28)}} = \sqrt{2(2.5028)} = 2.2373 \\
\text{Tukey } T &= \frac{1}{\sqrt{2}} q_{(1-\alpha; a, n_T - ab)} = \frac{1}{\sqrt{2}} q_{(0.9;3,28)} = \frac{1}{\sqrt{2}}(3.0257) = 2.1395
\end{aligned}$$

90% family confidence intervals:

$$\begin{array}{lclcl}
\hat{D}_1 \pm Ts\{\hat{D}_1\} & \Longrightarrow & -6.2334 \pm 2.1395\sqrt{0.4463} & \Longrightarrow & -7.6628 \leq D_1 \leq -4.8040 \\
\hat{D}_2 \pm Ts\{\hat{D}_2\} & \Longrightarrow & 0.0333 \pm 2.1395\sqrt{0.4448} & \Longrightarrow & -1.3936 \leq D_2 \leq 1.4602 \\
\hat{D}_3 \pm Ts\{\hat{D}_3\} & \Longrightarrow & 6.2667 \pm 2.1395\sqrt{0.4461} & \Longrightarrow & 4.8060 \leq D_3 \leq 7.7274
\end{array}$$

h)

$$\hat{L} = 0.3\bar{Y}_{12\cdot} + 0.6\bar{Y}_{22\cdot} + 0.1\bar{Y}_{32\cdot} = 0.3(21.3333) + 0.6(27.6667) + 0.1(20.6) = 25.06$$

$$s^2\{\hat{L}\} = MSE \sum_{i=1}^{3} \frac{c_i^2}{n_{i2}} = \frac{71.333}{28}\left(\frac{0.3^2}{6} + \frac{0.6^2}{6} + \frac{0.1^2}{5}\right) = 0.1962; \quad s\{\hat{L}\} = 0.4429$$

$$t_{(0.975,28)} = 2.0484$$

95% confidence interval for mean cash offer for this population is

$$\hat{L} \pm t_{(0.975,28)} s\{\hat{L}\} \quad \Longrightarrow \quad 25.06 \pm 2.0484(0.4429) \quad \Longrightarrow \quad 24.1528 \leq L \leq 25.9673$$

# 4 Problem 24.4

a)

$$
\begin{aligned}
\alpha_1 &= \mu_{1\cdot\cdot} - \mu_{\cdots} = 138.0 - 131.5 = 6.5 \\
\alpha_2 &= \mu_{2\cdot\cdot} - \mu_{\cdots} = 131.5 - 131.5 = 0 \\
\alpha_3 &= \mu_{3\cdot\cdot} - \mu_{\cdots} = 125.0 - 131.5 = -6.5
\end{aligned}
$$

b)

$$
\begin{aligned}
\beta_2 &= \mu_{\cdot 2\cdot} - \mu_{\cdots} = 134.0 - 131.5 = 2.5 \\
\gamma_1 &= \mu_{\cdot\cdot 1} - \mu_{\cdots} = 128.5 - 131.5 = -3.0
\end{aligned}
$$

c)

$$
\begin{aligned}
(\alpha\beta)_{12} &= \mu_{12\cdot} - \mu_{1\cdot\cdot} - \mu_{\cdot 2\cdot} + \mu_{\cdots} = 141.0 - 138.0 - 134.0 + 131.5 = 0.5 \\
(\alpha\gamma)_{21} &= \mu_{2\cdot 1} - \mu_{2\cdot\cdot} - \mu_{\cdot\cdot 1} + \mu_{\cdots} = 128.0 - 131.5 - 134.5 + 131.5 = -0.5 \\
(\beta\gamma)_{12} &= \mu_{\cdot 12} - \mu_{\cdot 1\cdot} - \mu_{\cdot\cdot 2} + \mu_{\cdots} = 132.0 + 2.5 - 3 + 131.5 = 0
\end{aligned}
$$

d)

$$
\begin{aligned}
(\alpha\beta\gamma)_{ijk} &= \mu_{ijk} - \mu_{ij\cdot} - \mu_{i\cdot k} - \mu_{\cdot jk} + \mu_{i\cdot\cdot} + \mu_{\cdot j\cdot} + \mu_{\cdot\cdot k} - \mu_{\cdots} \\
(\alpha\beta\gamma)_{111} &= 130 - 135 - 134 - 126 + 138 + 129 + 128.5 - 131.5 = -1 \\
(\alpha\beta\gamma)_{322} &= 131 - 128 - 126.5 - 137 + 125 + 134 + 134.5 - 131.5 = 1.5
\end{aligned}
$$

# 5 Problem 24.6

a) Residuals, $e_{ijk}$:

| | $k=1$ | | $k=2$ | |
|---|---|---|---|---|
| | $j=1$ | $j=2$ | $j=1$ | $j=2$ |
| $i=1$ | 3.7667 | 1.1667 | −0.5000 | −1.0333 |
| | −3.9333 | −1.6333 | 0.4000 | 1.3667 |
| | 0.1667 | 0.4667 | 0.1000 | −0.3333 |
| $i=2$ | −1.7000 | −0.8333 | 1.1333 | −0.5667 |
| | 1.1000 | 1.3667 | −1.6667 | 1.7333 |
| | 0.6000 | −0.5333 | 0.5333 | −1.1667 |

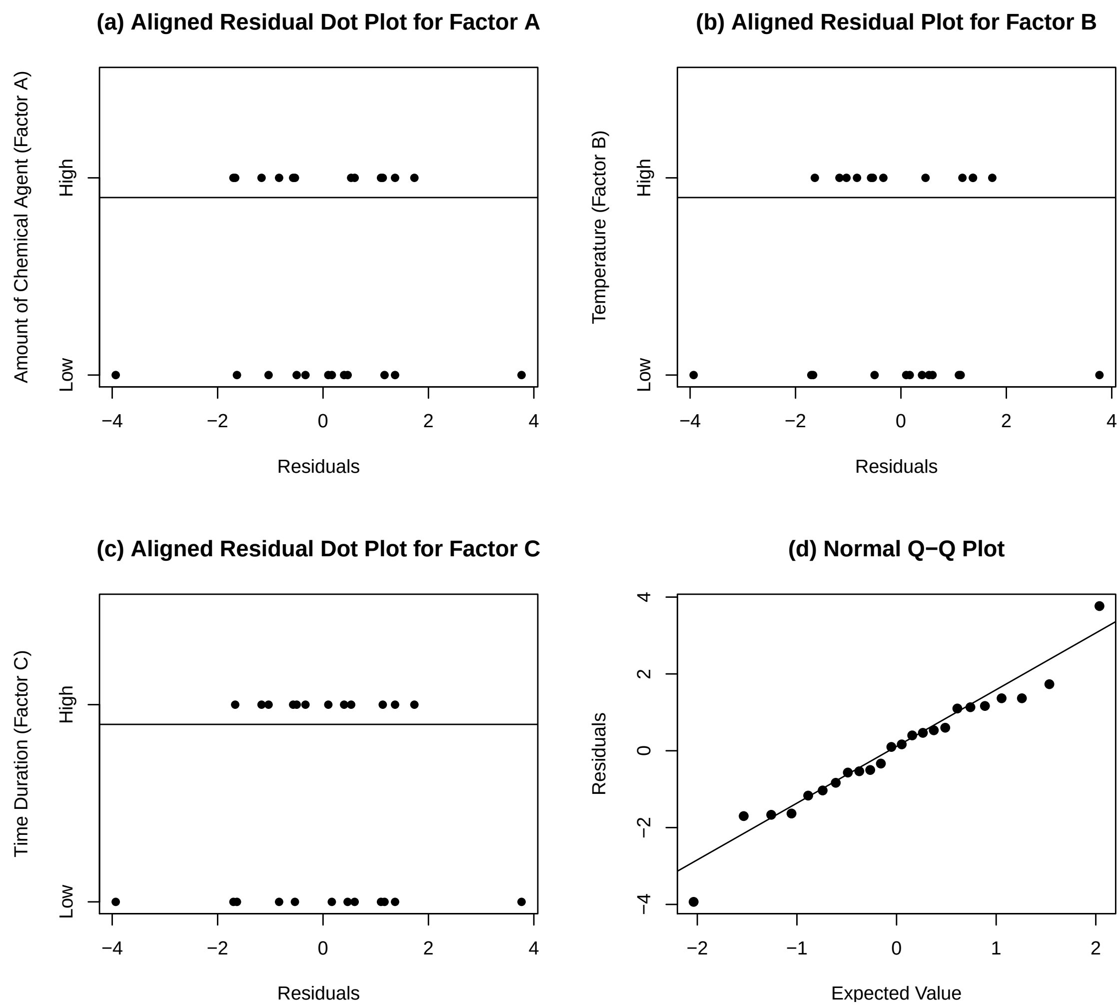


Figure 2: *Aligned residual dot plots for each factor, and normal probability plot of the residuals.*

From Figure 2, the residuals do not have constant variance across all factors. High level of each factor has less variation in residuals, while low level of each factor has more variation in their residuals. Since the constancy of variance assumption is violated, an ANOVA model is not appropriate.

b) Coefficient of correlation between ordered residuals and their expected values under normality is 0.9734. The normality assumption appears to be reasonable.

# 6 Problem 24.7

a) From Figure 3, there seems to be no three-factor interactions and no two-factor interactions. However, it can be seen that there are main effects for all three factors.

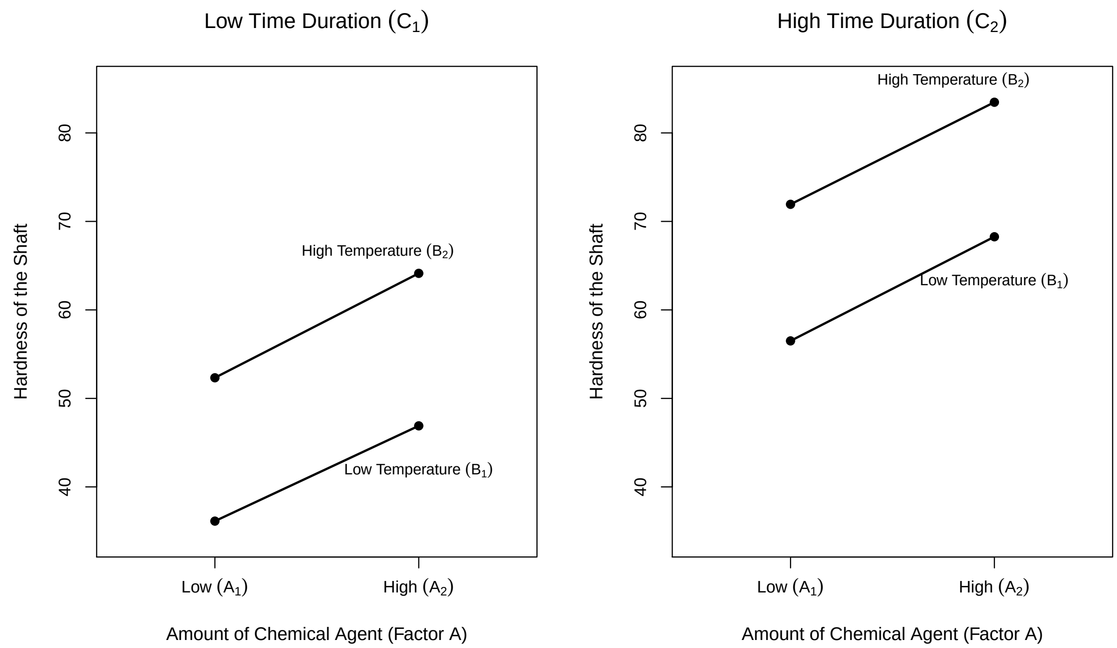


Figure 3: *Amount of chemical agent and temperature plots of the estimated treatment means*

b) ANOVA Table:

| **Source** | **df** | **SS** | **MS** | ***F*-value** | ***p*-value** |
|---|---|---|---|---|---|
| Treatment | 7 | 4772.26 | 681.75 | 202.9021 | $2.054\times10^{-14}$ |
| $A$ (Chemical) | 1 | 788.91 | 788.91 | 234.8810 | $5.533\times10^{-11}$ |
| $B$ (Temperature) | 1 | 1539.20 | 1539.20 | 458.2662 | $3.346\times10^{-13}$ |
| $C$ (Time Duration) | 1 | 2440.17 | 2440.17 | 726.5104 | $9.218\times10^{-15}$ |
| $AB$ interactions | 1 | 0.24 | 0.24 | 0.0715 | 0.7926 |
| $AC$ interactions | 1 | 0.20 | 0.20 | 0.0600 | 0.8095 |
| $BC$ interactions | 1 | 2.94 | 2.94 | 0.8753 | 0.3634 |
| $ABC$ interactions | 1 | 0.60 | 0.60 | 0.1791 | 0.6778 |
| Residuals | 16 | 53.74 | 3.36 | | |
| Total | 23 | 4826.00 | | | |

c) **Test for Three-Factor Interactions**

Hypotheses:

$H_0$ : all $(\alpha\beta\gamma)_{ijk} = 0$ (There is no three-factor interaction)

$H_a$ : at least one $(\alpha\beta\gamma)_{ijk} \neq 0$ (There are three-factor interactions)

at $\alpha = 2.5\%$ significance level.

Test Statistics:

$$F^* = \frac{MSABC}{MSE} = \frac{0.6}{3.36} = 0.1791$$

Decision: If $F^* > F_{0.975,1,16} = 6.1151$ reject $H_0$, otherwise fail to reject $H_0$.

Conclusion: Since $F^* < 6.1151$, we fail to reject $H_0$ and conclude that there is no three-factor interaction, at 2.5% significance level. $p - value = 0.6778$.

d) **Test for Chemical & Temperature ($AB$) Interactions**

Hypotheses:

$H_0$ : all $(\alpha\beta)_{ij} = 0$ (There is no $AB$ interaction)

$H_a$ : at least one $(\alpha\beta)_{ij} \neq 0$ (There are $AB$ interactions)

at $\alpha = 2.5\%$ significance level.

Test Statistics:

$$F^* = \frac{MSAB}{MSE} = \frac{0.24}{3.36} = 0.0715$$

Decision: If $F^* > F_{0.975,1,16} = 6.1151$ reject $H_0$, otherwise fail to reject $H_0$.

Conclusion: Since $F^* < 6.1151$, we fail to reject $H_0$ and conclude that there is no interaction between amount of chemical agent and temperature of the hardening process on the outside hardness of the shaft, at 2.5% significance level. $p - value = 0.7926$.

**Test for Chemical & Time Duration ($AC$) Interactions**

Hypotheses:

$H_0$ : all $(\alpha\gamma)_{ik} = 0$ (There is no $AC$ interaction)

$H_a$ : at least one $(\alpha\gamma)_{ik} \neq 0$ (There are $AC$ interactions)

at $\alpha = 2.5\%$ significance level.

Test Statistics:

$$F^* = \frac{MSAC}{MSE} = \frac{0.2}{3.36} = 0.06$$

Decision: If $F^* > F_{0.975,1,16} = 6.1151$ reject $H_0$, otherwise fail to reject $H_0$.

Conclusion: Since $F^* < 6.1151$, we fail to reject $H_0$ and conclude that there is no interaction between amount of chemical agent and time duration of the hardening process on the outside hardness of the shaft, at 2.5% significance level. $p - value = 0.8095$.

**Test for Temperature & Time Duration ($BC$) Interactions**

Hypotheses:

$H_0:$ all $(\beta\gamma)_{jk} = 0$ (There is no $BC$ interaction)
$H_a:$ at least one $(\beta\gamma)_{jk} \neq 0$ (There are $BC$ interactions)
at $\alpha = 2.5\%$ significance level.
Test Statistics:

$$F^* = \frac{MSBC}{MSE} = \frac{2.94}{3.36} = 0.8753$$

Decision: If $F^* > F_{0.975,1,16} = 6.1151$ reject $H_0$, otherwise fail to reject $H_0$.
Conclusion: Since $F^* < 6.1151$, we fail to reject $H_0$ and conclude that there is no interaction between the temperature and time duration of the hardening process on the outside hardness of the shaft, at 2.5% significance level. $p - value = 0.3634$.

e) **Test for Chemical ($A$) Main Effect**
Hypotheses:
$H_0:$ $\alpha_1 = \alpha_2 = 0$ (Chemical agent has no effect)
$H_a:$ $\alpha_1 \neq 0$ or $\alpha_2 \neq 0$ (Chemical agent has effect)
at $\alpha = 2.5\%$ significance level.
Test Statistics:

$$F^* = \frac{MSA}{MSE} = \frac{788.91}{3.36} = 234.881$$

Decision: If $F^* > F_{0.975,1,16} = 6.1151$ reject $H_0$, otherwise fail to reject $H_0$.
Conclusion: Since $F^* > 6.1151$, we reject $H_0$ and conclude that the amount of chemical agent added to alloy in a molten state has an effect on the outside hardness of the shaft, at 2.5% significance level. $p - value = 5.533 \times 10^{-11}$.

**Test for Temperature ($B$) Main Effect**
Hypotheses:
$H_0:$ $\beta_1 = \beta_2 = 0$ (Temperature has no effect)
$H_a:$ $\beta_1 \neq 0$ or $\beta_2 \neq 0$ (Temperature has effect)
at $\alpha = 2.5\%$ significance level.
Test Statistics:

$$F^* = \frac{MSB}{MSE} = \frac{1539.2}{3.36} = 458.2662$$

Decision: If $F^* > F_{0.975,1,16} = 6.1151$ reject $H_0$, otherwise fail to reject $H_0$.
Conclusion: Since $F^* > 6.1151$, we reject $H_0$ and conclude that the temperature of the hardening process has an effect on the outside hardness of the shaft, at 2.5% significance level. $p - value = 3.346 \times 10^{-13}$.

**Test for Time Duration ($C$) Main Effect**
Hypotheses:
$H_0:$ $\gamma_1 = \gamma_2 = 0$ (Time duration has no effect)

$H_a: \gamma_1 \neq 0$ or $\gamma_2 \neq 0$ (Time duration has effect)
at $\alpha = 2.5\%$ significance level.
Test Statistics:

$$F^* = \frac{MSC}{MSE} = \frac{2440.17}{3.36} = 726.5104$$

Decision: If $F^* > F_{0.975,1,16} = 6.1151$ reject $H_0$, otherwise fail to reject $H_0$.
Conclusion: Since $F^* > 6.1151$, we reject $H_0$ and conclude that the time duration of the hardening process has an effect on the outside hardness of the shaft, at 2.5% significance level. $p - value = 9.218 \times 10^{-15}$.

f) The seven separate $F$ tests for factor effects led to the following conclusions:

   i) There are no three-factor interactions.

   ii) There are no two-factor interactions.

   iii) Main effects for chemical agent, time duration, and temperature are present. Higher levels for each factor results in higher values for outside hardness of the shaft.

   Upper bound for the family level of significance (using Kimball inequality):

$$\alpha \leq 1 - (0.975)^7 = 0.1624$$

g) The results in part (f) confirms my graphic analysis in part (a).

# 7 Problem 24.19

$$\begin{aligned}
\sum_{i=1}^{a} (\alpha\beta\gamma)_{ijk} &= \sum_{i=1}^{a} (\mu_{ijk} - \mu_{ij\cdot} - \mu_{i\cdot k} - \mu_{\cdot jk} + \mu_{i\cdot\cdot} + \mu_{\cdot j\cdot} + \mu_{\cdot\cdot k} - \mu_{\cdot\cdot\cdot}) \\
&= a\mu_{\cdot jk} - a\mu_{\cdot j\cdot} - a\mu_{\cdot\cdot k} - a\mu_{\cdot jk} + a\mu_{\cdot\cdot\cdot} + a\mu_{\cdot j\cdot} + a\mu_{\cdot\cdot k} - a\mu_{\cdot\cdot\cdot} \\
&= 0
\end{aligned}$$

# 8 Problem 24.20

| **Source** | **SS** | **df** | **MS** | **E{MS}** |
|---|---|---|---|---|
| Factor $A$ | $SSA$ | $a-1$ | $MSA$ | $\sigma^2 + bc\frac{\sum \alpha_i^2}{a-1}$ |
| Factor $B$ | $SSB$ | $b-1$ | $MSB$ | $\sigma^2 + ac\frac{\sum \beta_j^2}{b-1}$ |
| Factor $C$ | $SSC$ | $c-1$ | $MSC$ | $\sigma^2 + ab\frac{\sum \gamma_k^2}{c-1}$ |
| $AB$ interactions | $SSAB$ | $(a-1)(b-1)$ | $MSAB$ | $\sigma^2 + c\frac{\sum\sum(\alpha\beta)_{ij}^2}{(a-1)(b-1)}$ |
| $AC$ interactions | $SSAC$ | $(a-1)(c-1)$ | $MSAC$ | $\sigma^2 + b\frac{\sum\sum(\alpha\gamma)_{ik}^2}{(a-1)(c-1)}$ |
| $BC$ interactions | $SSBC$ | $(b-1)(c-1)$ | $MSBC$ | $\sigma^2 + a\frac{\sum\sum(\beta\gamma)_{jk}^2}{(b-1)(c-1)}$ |
| Error | $SSABC$ | $(a-1)(b-1)(c-1)$ | $MSABC$ | $\sigma^2$ |
| Total | $SSTO$ | $abc-1$ | | |

$$
\begin{aligned}
SSA &= bc\sum_i \left(\bar{Y}_{i\cdot\cdot} - \bar{Y}_{\cdot\cdot\cdot}\right)^2 \\
SSB &= ac\sum_j \left(\bar{Y}_{\cdot j\cdot} - \bar{Y}_{\cdot\cdot\cdot}\right)^2 \\
SSC &= ab\sum_k \left(\bar{Y}_{\cdot\cdot k} - \bar{Y}_{\cdot\cdot\cdot}\right)^2 \\
SSAB &= c\sum_i\sum_j \left(\bar{Y}_{ij\cdot} - \bar{Y}_{i\cdot\cdot} - \bar{Y}_{\cdot j\cdot} + \bar{Y}_{\cdot\cdot\cdot}\right)^2 \\
SSAC &= b\sum_i\sum_k \left(\bar{Y}_{i\cdot k} - \bar{Y}_{i\cdot\cdot} - \bar{Y}_{\cdot\cdot k} + \bar{Y}_{\cdot\cdot\cdot}\right)^2 \\
SSBC &= a\sum_j\sum_k \left(\bar{Y}_{\cdot jk} - \bar{Y}_{\cdot j\cdot} - \bar{Y}_{\cdot\cdot k} + \bar{Y}_{\cdot\cdot\cdot}\right)^2 \\
SSABC &= \sum_i\sum_j\sum_k \left(\bar{Y}_{ijk} - \bar{Y}_{ij\cdot} - \bar{Y}_{i\cdot k} - \bar{Y}_{\cdot jk} + \bar{Y}_{i\cdot\cdot} + \bar{Y}_{\cdot j\cdot} + \bar{Y}_{\cdot\cdot k} - \bar{Y}_{\cdot\cdot\cdot}\right)^2 \\
SSTO &= \sum_i\sum_j\sum_k \left(Y_{ijk} - \bar{Y}_{\cdot\cdot\cdot}\right)^2
\end{aligned}
$$

# 9 Problem 25.13

a)

$$E\{MSA\} = \sigma^2 + bn\sigma_\alpha^2 + n\sigma_{\alpha\beta}^2 = 5 + (2)(5)(8) + (5)(6) = 115$$

$$\begin{aligned} E\{MSB\} &= \sigma^2 + an\sigma^2_\beta + n\sigma^2_{\alpha\beta} = 5 + (3)(5)(10) + (5)(6) = 185 \\ E\{MSAB\} &= \sigma^2 + n\sigma^2_{\alpha\beta} = 5 + (5)(6) = 35 \end{aligned}$$

b)

$$\begin{aligned} E\{MSA\} &= \sigma^2 + bn\sigma^2_\alpha = 5 + (2)(5)(8) = 85 \\ E\{MSB\} &= \sigma^2 + an\sigma^2_\beta = 5 + (3)(5)(10) = 155 \\ E\{MSAB\} &= \sigma^2 = 5 \end{aligned}$$

## 10 Problem 25.15

| **Source** | **df** | **SS** | **MS** |
|---|---|---|---|
| Driver (A) | 3 | 280.285 | 93.428 |
| Car (B) | 4 | 94.714 | 23.678 |
| AB Interactions | 12 | 2.447 | 0.204 |
| Residuals | 20 | 3.515 | 0.176 |
| Total | 39 | 380.9598 | |

a) Hypotheses:
$H_0 : \sigma^2_{\alpha\beta} = 0$ The two do not factors interact
$H_a : \sigma^2_{\alpha\beta} > 0$ The two factors interact
at 5% significance level.
Test Statistic:

$$F^* = \frac{MSAB}{MSE} = \frac{0.204}{0.176} = 1.16$$

Decision: If $F* > F_{(0.95,12,20)} = 2.2776$, we reject $H_0$, otherwise we fail to reject $H_0$.
Conclusion: Since $F^* = 1.16 < 2.2776$, we fail to reject the null hypothesis and conclude that there is no interaction between *driver* and *car*. The $p - value$ of the test is 0.37.

b) **Test for Factor A**
Hypotheses:
$H_0 : \sigma^2_\alpha = 0$ (There is no factor A main effect)
$H_a : \sigma^2_\alpha > 0$ (There is factor A main effect)
at 5% significance level.
Test Statistic:

$$F^* = \frac{MSA}{MSAB} = \frac{93.428}{0.204} = 457.98$$

Decision: If $F* > F_{(0.95,3,12)} = 3.4903$, we reject $H_0$, otherwise we fail to reject $H_0$.

Conclusion: Since $F^* = 457.98 > 3.4903$, we reject the null hypothesis and conclude that main effect of *driver* (factor A) is present. The $p-value$ of the test is $1.231 \times 10^{-12}$.

**Test for Factor B**

Hypotheses: $H_0 : \sigma^2_\beta = 0$ (There is no factor B main effect)
$H_a : \sigma^2_\beta > 0$ (There is factor B main effect)
at 5% significance level.
Test Statistic:

$$F^* = \frac{MSB}{MSAB} = \frac{23.678}{0.204} = 116.0686$$

Decision: If $F* > F_{(0.95,4,12)} = 3.2592$, we reject $H_0$, otherwise we fail to reject $H_0$.
Conclusion: Since $F^* = 116.0686 > 3.2592$, we reject the null hypothesis and conclude that main effect of *car* (factor B) is present. The $p-value$ of the test is $1.7521 \times 10^{-9}$.

c) Point estimate for $\sigma^2_\alpha$:

$$s^2_\alpha = \frac{MSA - MSAB}{bn} = \frac{93.42825 - 0.203875}{(5)(2)} = 9.3224$$

Point estimate for $\sigma^2_\beta$:

$$s^2_\beta = \frac{MSB - MSAB}{an} = \frac{23.6784 - 0.203875}{(4)(2)} = 2.9343$$

Factor A (*driver*) has the greater effect on gasoline consumption.

d) We haven't covered MLS.

e) Using the Scatterthwaite procedure;

$$\begin{array}{ll} c_1 = \frac{1}{na} = \frac{1}{2(4)} = 0.125 & c_2 = -\frac{1}{nb} = -\frac{1}{2(4)} = -0.125 \\ MS_1 = MSB = 23.67837 & MS_2 = MSAB = 0.203875 \\ df_1 = b - 1 = 5 - 1 = 4 & df_2 = (a-1)(b-1) = (3)(4) = 12 \end{array}$$

$$\hat{L} = c_1 MS_1 + c_2 MS_2 = 2.9343$$

$$df = \frac{(c_1 MS_1 + c_2 MS_2)^2}{\frac{(c_1 MS_1)^2}{df_1} + \frac{(c_2 MS_2)^2}{df_2}} = \frac{(2.9343)^2}{\frac{(2.9598)^2}{4} + \frac{(-0.0255)^2}{12}} = 3.9313 \approx 4$$

$$\chi^2_{0.025,4} = 0.4844 \qquad \chi^2_{0.975,4} = 11.1433$$

$$\frac{(df)\hat{L}}{\chi^2_{1-\alpha/2,df}} \leq \sigma^2_\beta \leq \frac{(df)\hat{L}}{\chi^2_{\alpha/2,df}} \implies \frac{(4)(2.9343)}{11.1433} \leq \sigma^2_\beta \leq \frac{(4)(2.9343)}{0.4844} \implies 1.0352 \leq \sigma^2_\beta \leq 23.8135$$

Wide not very accurate.