Experimental Design and Analysis HW03 Solution

Problem 1

The one-way fixed effect model for the pulp experiment is $\Big(i \in \{A,B,C,D\},\ j \in \{1,\dots,5\}\Big)$:

$$y_{ij} = \eta + \tau_i + \varepsilon_{ij}$$

where

$$\begin{cases} \eta & \text{is the } \textbf{grand mean} \\ \tau_i & \text{is the } \textbf{treatment (operator) effect} \\ \varepsilon_{ij} \overset{i.i.d.}{\sim} \mathcal{N}(0,\sigma^2) & \text{is the } \textbf{random error} \end{cases} .$$

The summary of the fitted model is:

```
##             Df Sum Sq Mean Sq F value Pr(>F)
## operator     3   1.34  0.4467   4.204 0.0226 *
## Residuals   16   1.70  0.1062
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

From the summary, we see that $N = 20,\ k = 4,\ n_1 = n_2 = n_3 = n_4 = 5,\ \hat{\sigma}^2 = 0.10625.$

We may use ***bonfCI*** *function and* ***tukeyCI*** *function in the* ***asbio*** *package to obtain the following results.*

The 95% ***Bonferroni*** *simultaneous confidence interval for* $\tau_i - \tau_j$ *is*

$$\overline{y}_{i\cdot} - \overline{y}_{j\cdot} \pm t_{\left(N-k,\ 0.05\big/2\binom{4}{2}\right)} \hat{\sigma} \sqrt{\frac{1}{n_i} + \frac{1}{n_j}}$$

```
##
## 95% Bonferroni confidence intervals
##
##           Diff    Lower   Upper Decision Adj. p-value
## muA-muB  0.18 -0.44018 0.80018   FTR H0            1
## muA-muC -0.38 -1.00018 0.24018   FTR H0     0.503359
## muB-muC -0.56 -1.18018 0.06018   FTR H0     0.091504
```

```
## muA-muD -0.44 -1.06018 0.18018   FTR H0      0.291823
## muB-muD -0.62 -1.24018 0.00018   FTR H0      0.050093
## muC-muD -0.06 -0.68018 0.56018   FTR H0             1
```

*The length of 95% **Bonferroni** simultaneous confidence interval is*

$$2 \times t_{\left(N-k,\ 0.05/2\binom{4}{2}\right)} \hat{\sigma} \sqrt{\frac{1}{n_i} + \frac{1}{n_j}} = 1.24036782211.$$

---

*The 95% **Tukey** simultaneous confidence interval for $\tau_i - \tau_j$ is*

$$\overline{y}_{i\cdot} - \overline{y}_{j\cdot} \pm \frac{1}{\sqrt{2}} q_{(k,\ N-k,\ \alpha)} \hat{\sigma} \sqrt{\frac{1}{n_i} + \frac{1}{n_j}}$$

```
##
## 95% Tukey-Kramer confidence intervals
##
##           Diff    Lower    Upper  Decision Adj. p-value
## muA-muB  0.18 -0.40981  0.76981    FTR H0      0.818543
## muA-muC -0.38 -0.96981  0.20981    FTR H0      0.290304
## muB-muC -0.56 -1.14981  0.02981    FTR H0      0.065794
## muA-muD -0.44 -1.02981  0.14981    FTR H0      0.184479
## muB-muD -0.62 -1.20981 -0.03019 Reject H0      0.037669
## muC-muD -0.06 -0.64981  0.52981    FTR H0      0.991078
```

*The length of 95% **Tukey** simultaneous confidence interval is*

$$2 \times \frac{1}{\sqrt{2}} q_{(k,\ N-k,\ \alpha)} \hat{\sigma} \sqrt{\frac{1}{n_i} + \frac{1}{n_j}} = 1.17962869298.$$

*Hence, the **Tukey** method gives the shorter intervals.*

Problem 2

(a)

```
##
## 99% Bonferroni confidence intervals
##
##           Diff    Lower   Upper Decision Adj. p-value
## muA-muB  0.18 -0.59772 0.95772   FTR H0             1
## muA-muC -0.38 -1.15772 0.39772   FTR H0      0.503359
## muB-muC -0.56 -1.33772 0.21772   FTR H0      0.091504
## muA-muD -0.44 -1.21772 0.33772   FTR H0      0.291823
## muB-muD -0.62 -1.39772 0.15772   FTR H0      0.050093
## muC-muD -0.06 -0.83772 0.71772   FTR H0             1
```

```
##
## 99% Tukey-Kramer confidence intervals
##
##          Diff    Lower   Upper Decision Adj. p-value
## muA-muB  0.18 -0.57684 0.93684   FTR H0     0.818543
## muA-muC -0.38 -1.13684 0.37684   FTR H0     0.290304
## muB-muC -0.56 -1.31684 0.19684   FTR H0     0.065794
## muA-muD -0.44 -1.19684 0.31684   FTR H0     0.184479
## muB-muD -0.62 -1.37684 0.13684   FTR H0     0.037669
## muC-muD -0.06 -0.81684 0.69684   FTR H0     0.991078
```

Since all the confidence intervals contain 0, neither the Bonferroni method nor the Tukey method declares any pair of treatments as different at the 0.01 level.

(b)

The results are not contradictory because the one-way ANOVA is performed at the 0.05 level, whereas the Bonferroni and Tukey multiple comparisons are conducted at the 0.01 level, which is more stringent.

In Section 2.1, the one-way ANOVA rejects H_0 *at the 0.05 level, which implies that at least one pair of treatments is significantly different. Hence, we do multiple comparisons to identify which pairs differ.*

However, we do multiple comparisons at the 0.01 level now, which is more stringent to reject H_0 *than at the 0.05 level.*

Because the differences are not large enough to be identified at the more stringent manner, the tests may fail to identify any significant pairwise differences even when some true differences do exist.

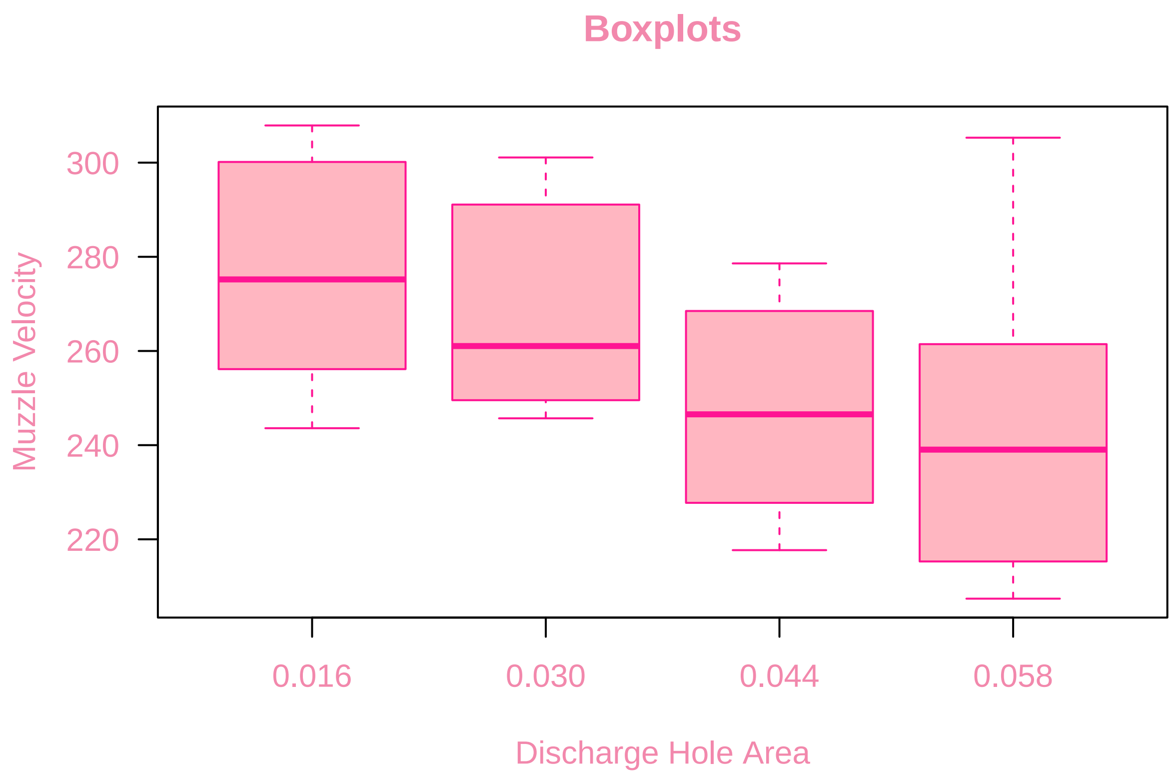
If we do the multiple comparisons at the 0.05 level, then some pairs of oprators are declared to be significantly different under both Bonferroni and Tukey methods. (*It still makes sense if none of them are declared to be significantly different.*)

(c)

The p-value for the one-way ANOVA F-test is 0.02.

Therefore, at the 0.01 significance level we fail to reject H_0(*since* $0.02 > 0.01$). *Because the overall test is not significant at 0.01, there is no evidence of differences among treatment means at this level; consequently, we would not expect any pairwise difference to be declared significant by the Bonferroni or Tukey procedure at the 0.01 level. Hence, we can reach the same conclusion as in (a) without performing the multiple comparisons.*

Problem 3



From the figure, we can indeed observe that as the discharge hole area increases, the muzzle velocity decreases.

(a)

The one-way fixed effect model for the muzzle velocity experiment is $\Big(i \in \{A, B, C, D\},\ j \in \{1, \dots, 16\}\Big)$:

$$y_{ij} = \eta + \tau_i + \varepsilon_{ij}$$

where

$$\begin{cases} \eta & \text{is the } \textbf{grand mean} \\ \tau_i & \text{is the } \textbf{treatment (discharge hole area) effect} \\ \varepsilon_{ij} \overset{i.i.d.}{\sim} \mathcal{N}(0, \sigma^2) & \text{is the } \textbf{random error} \end{cases}.$$

Using the F-test to test

$$\begin{cases} H_0 : \tau_A = \tau_B = \tau_C = \tau_D \\ H_1 : \text{at least one pair of the treatments is different} \end{cases}$$

```
##              Df Sum Sq Mean Sq F value   Pr(>F)
## factor        3  13515    4505   7.067 0.000381 ***
```

```
## Residuals   60  38250     638
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Since the p-value is much smaller than 0.05, we reject H_0 *and do the multiple comparisons by Tukey method*

```
##
## 95% Tukey-Kramer confidence intervals
##
##              Diff    Lower    Upper  Decision Adj. p-value
## muA-muB      7.95 -15.6393  31.5393    FTR H0     0.809747
## muA-muC 30.56875  6.97945 54.15805 Reject H0     0.005991
## muB-muC 22.61875 -0.97055 46.20805    FTR H0     0.064808
## muA-muD 34.18125 10.59195 57.77055 Reject H0      0.00172
## muB-muD 26.23125  2.64195 49.82055 Reject H0     0.023551
## muC-muD   3.6125 -19.9768  27.2018    FTR H0     0.977394
```

The results indicate there are significant differences between (A&C), (A&D), (B&D).

(b)

Note that the four quantitative levels $(0.016, 0.030, 0.044, 0.058)$ *are equally spaced.*
Following the same notations as in page 60 of the textbook, set $m = 0.037,\ \Delta = 0.014$.

The orthogonal polynomials of degree 1 and degree 2 are: (See Appendix G of textbook for the coefficients)

$$\begin{cases} P_1(x) = 2\left(\dfrac{x - 0.037}{0.014}\right) \\ P_2(x) = \left(\dfrac{x - 0.037}{0.014}\right)^2 - \left(\dfrac{4^2 - 1}{12}\right) \end{cases}$$

Then we can use the orthogonal polynomials to model the linear and quadratic effects:

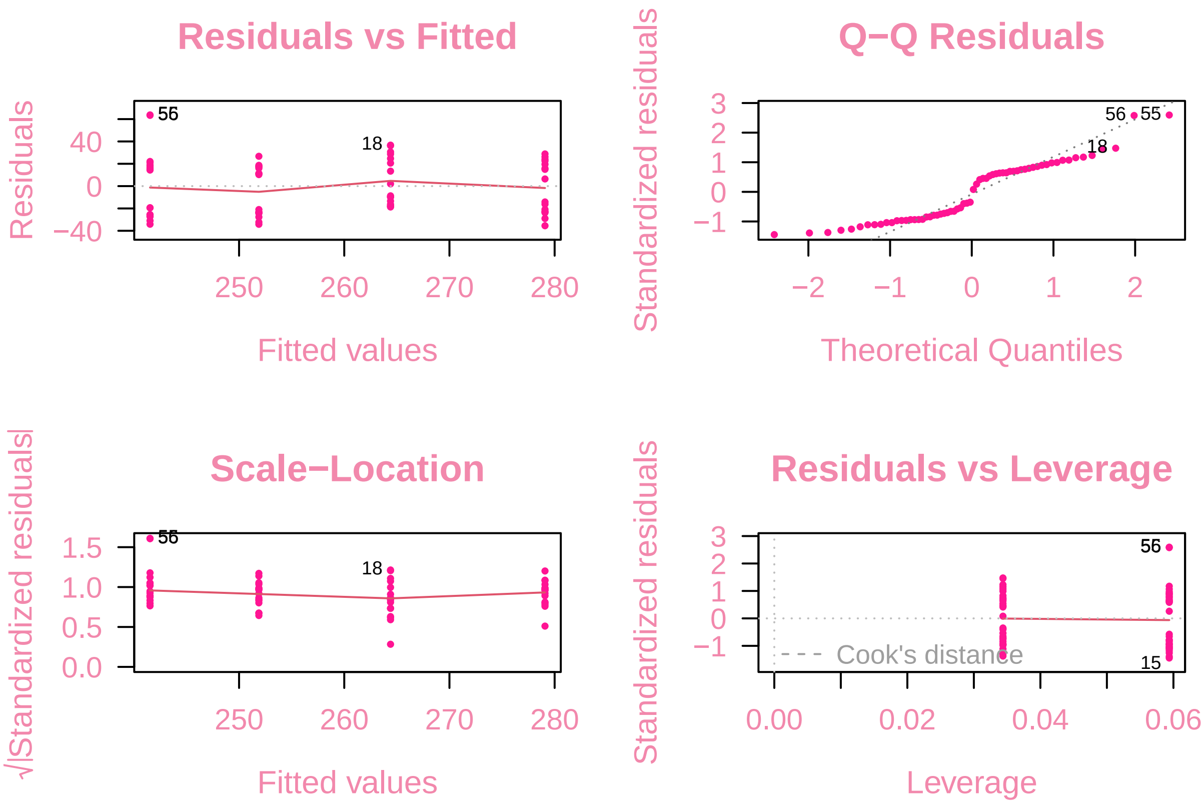
$$y = \beta_0 + \beta_1 \frac{P_1(x)}{\sqrt{20}} + \beta_2 \frac{P_2(x)}{\sqrt{4}} + \varepsilon,\ \text{where } \varepsilon \sim \mathcal{N}(0, \sigma^2)$$

```
##
## Call:
## lm(formula = velocity ~ I(P1(area)/sqrt(20)) + I(P2(area)/sqrt(4)),
##     data = mv)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -35.484 -23.133  -3.349  18.953  63.765
##
## Coefficients:
##                      Estimate Std. Error t value Pr(>|t|)
## (Intercept)           259.225      3.167  81.851  < 2e-16 ***
## I(P1(area)/sqrt(20))  -27.987      6.334  -4.419 4.15e-05 ***
## I(P2(area)/sqrt(4))     2.169      6.334   0.342    0.733
## ---
```

```
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
## Residual standard error: 25.34 on 61 degrees of freedom
## Multiple R-squared:  0.2436, Adjusted R-squared:  0.2188
## F-statistic:  9.82 on 2 and 61 DF,  p-value: 0.0002008
```

The results indicate that the discharge hole area has a strong negative (as expected) linear effect but no quadratic effect on the muzzle velocity.

Do some diagnostics:



Based on the residual plots, there is no evidence of non-constant variance or mean-curvature. However, the residuals are divided into two groups and the QQ plot shows departure from normality, which means there is at least an important factor that has not been added to the model. (As expected, since these data are obtained by collapsing and adapting three-way layout data.)

Problem 4

(a)

The one-way fixed effect model for the packing machine experiment is $\Big(i \in \{1,2,3\},\ j \in \{1,\dots,20\}\Big)$:

$$y_{ij} = \eta + \tau_i + \varepsilon_{ij}$$

where

$$\begin{cases} \eta & \textit{is the } \textbf{\textit{grand mean}} \\ \tau_i & \textit{is the } \textbf{\textit{treatment}} \ (\textbf{\textit{machine}}) \ \textbf{\textit{effect}} \\ \varepsilon_{ij} \overset{i.i.d.}{\sim} \mathcal{N}(0,\sigma^2) & \textit{is the } \textbf{\textit{random error}} \end{cases}.$$

Using the F-test to test

$$\begin{cases} H_0 : \tau_1 = \tau_2 = \tau_3 \\ H_1 : \textit{at least one pair of the treatments is different} \end{cases}$$

```
##             Df Sum Sq Mean Sq F value   Pr(>F)
## ind          2  31.65  15.826   59.51 1.11e-14 ***
## Residuals   57  15.16   0.266
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

Since the p-value of F-test is extremely small, we reject H_0 *with a very high confidence.*

(b)

Among the three chosen machines, at least one pair differs significantly, suggesting that some machines have systematically different mean fill weights (i.e., tend to overfill or underfill).

(c)

The one-way random effect model for the packing machine experiment is $\Big(i \in \{1,2,3\},\ j \in \{1,\dots,20\}\Big)$:

$$y_{ij} = \eta + \tau_i + \varepsilon_{ij}$$

where

$$\begin{cases} \eta & \textit{is the } \textbf{\textit{grand mean}} \\ \tau_i \overset{i.i.d.}{\sim} \mathcal{N}(0,\sigma_\tau^2) & \textit{is the } \textbf{\textit{random treatment}} \ (\textbf{\textit{machine}}) \ \textbf{\textit{effect}} \quad \textit{and } \tau_i \perp \varepsilon_{ij}. \\ \varepsilon_{ij} \overset{i.i.d.}{\sim} \mathcal{N}(0,\sigma^2) & \textit{is the } \textbf{\textit{random error}} \end{cases}$$

Using the F-test to test

$$\begin{cases} H_0 : \sigma_\tau^2 = 0 \\ H_1 : \sigma_\tau^2 > 0 \end{cases}$$

```
##              Df Sum Sq Mean Sq F value   Pr(>F)    
## ind           2  31.65  15.826   59.51 1.11e-14 ***
## Residuals    57  15.16   0.266                     
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

The ANOVA table is the same as in part (a), and the estimates of variance components are

$$\begin{cases} \hat{\sigma}^2 = MSE & = 0.26594956 \\ \hat{\sigma}_\tau^2 = \dfrac{MSTr - MSE}{n} & = 0.77801352 \end{cases}.$$

(d)

The 95% *confidence interval for* η *is*

$$\hat{\eta} \pm t_{(k-1,\ 0.05/2)} \sqrt{\frac{MSTr}{nk}}$$

Plugging $n = 20,\ k = 3,\ \hat{\eta} = \overline{y}_{..} = 50.1015,\ MSTr = 15.82622$ *into the formula above, we find that*

$$(47.8917188259,\ 52.3112811741)$$

is a 95% *confidence interval for the mean weight of the bags filled in the plant.*

Problem 5

(a)

The one-way random effect model is :

$$y_{ij} = \eta + \tau_i + \varepsilon_{ij}$$

where

$$\begin{cases} \eta & \text{is the } \textbf{grand mean} \\ \tau_i \overset{i.i.d.}{\sim} \mathcal{N}(0, \sigma_\tau^2) & \text{is the } \textbf{random treatment effect} \quad \text{and } \tau_i \perp \varepsilon_{ij}. \\ \varepsilon_{ij} \overset{i.i.d.}{\sim} \mathcal{N}(0, \sigma^2) & \text{is the } \textbf{random error} \end{cases}$$

*In **REM** τ_i is a random effect since it is a treatment selected randomly from the treatment population.*

Note that $\mathbb{E}(y_{ij}) = \mathbb{E}(\eta + \tau_i + \varepsilon_{ij}) = \eta$, so η is the sample mean under all treatments and all EU's. Hence η is called the population mean.

(b)

*Consider the pulp experiment, where the treatments are operators. If the study aims to make inference about operator performance in general—i.e., whether operators differ and what the average performance is across the whole population of operators—then a one-way **random effects model (REM)** is appropriate. In this setting, the operators in the experiment are viewed as a random sample from a larger population, and the parameter η represents the **population mean** (the average response over all operators). Therefore, η is scientifically meaningful and worth estimating with a confidence interval because it generalizes beyond the specific operators observed.*

*In contrast, if we use a one-way **fixed effects model (FEM)**, we treat the included operators as the only levels of interest (not a random sample). The grand mean parameter in the fixed effects model is then merely the average of these particular selected operators, so it does **not** represent a population average and typically does not generalize. Hence, the grand mean in the FEM is often not of primary interest, whereas η in the REM is of interest because it corresponds to the population mean.*

*The main difference between **FEM** and **REM** is the role of the treatment effects:*

$$\begin{cases} \text{In } \textbf{FEM}, & \text{the treatment effect is a parameter.} \\ \text{In } \textbf{REM}, & \text{the treatment effect is a random variable.} \end{cases}$$