University of California, Santa Barbara

PSTAT 122

Final Exam

Author:
Emily Lu

Supervisor:
Professor John Hsu

June 11th, 2020

Abstract

This report is split into two components. The first component is based on an experiment that a farmer conducted to determine which of the six new brands of fertilizer produces the most wheat (in bushels) in one season. The second component is based on an experiment that an engineer conducted to determine which plate materials and temperature levels would result in the best observed battery life (in hours). For both components, we will be interested in analyzing the resulting data and achieving its respective objective.

Contents

1 Part I: Fertilizer and Wheat Production

1.1 Introduction

Six new brands (A, B, C, D, E, and F) of fertilizer that a farmer can use to grow crops just came on the market. Before deciding which brand he should use permanently for all crops, a farmer decided to experiment for one season. To do so, he randomly assign each fertilizer to five one-acre tracts of land that he used to grow wheat. After one season, he then recorded the production of wheat (in bushels) for each acre for six brands of fertilizer. Given this data, we are now interested in analyzing which fertilizer on average produced the most wheats. Our objectives will be to test hypotheses about the treatment means, i.e. the average produced wheat from the ith brand of fertilizer, and to estimate them.

1.2 Data and Model Description

The following table gives the production of wheat (in bushels) for each acre for six brands of fertilizer.

Fertilizer A	Fertilizer B	Fertilizer C	Fertilizer D	Fertilizer E	Fertilizer F
73	75	69	68	69	80
70	82	67	77	63	87
71	82	76	76	65	86
65	80	74	73	65	84
76	86	74	71	58	78

This is an example of a single-factor experiment with $a = 6$ treatments (brands) of the factor (fertilizer) and $n = 5$ replicates (observations under each brand of fertilizer). We could describe the observations from the experiment with the following model:

$$y_{ij} = \mu_i + \epsilon_{ij} \begin{cases} i = 1, 2, ..., 6 \\ j = 1, 2, ..., 5 \end{cases} \tag{1}$$

where y_{ij} is the ijth observation, μ_i is the mean of the ith treatment or brand of fertilizer, and ϵ_{ij} is a random error component that accounts for all other sources of variability in the experiment. For convenience, we will think of errors as having mean zero so that $E(y_{ij}) = \mu_i$.

Equation (1) is known as the means model. An alternative way to describe our model for the data is to define

$$\mu_i = \mu + \tau_i, \ i = 1, 2, ..., 6$$

so that equation (1) becomes

$$y_{ij} = \mu + \epsilon_{ij} \begin{cases} i = 1, 2, ..., 6 \\ j = 1, 2, ..., 5 \end{cases} . \tag{2}$$

This model is known as the effects (or more specifically, single-factor analysis of variance) model where μ is the overall mean and τ_i is the ith treatment effect. For this experiment, we assume this model to be a fixed model effects model.

1.3 Analysis of Fixed Effects Model

We are interested in testing the equality of the 6 treatment means; that is

$$E(y_{ij}) = \mu + \tau_i = \mu_i, \; i = 1, 2, ...6.$$

In other words, we want to see if the 6 fertilizers produced the same number of wheat bushels on average. Our hypotheses are

$$H_0 \;:\; \mu_1 = \mu_2 = ... = \mu_6 \quad (3)$$
$$H_1 \;:\; \mu_i \neq \mu_j \text{ for at least one pair } (i, j) \quad (4)$$

Alternatively and similarly, we could test if the treatment effects are zero with the following hypotheses:

$$H_0 \;:\; \tau_1 = \tau_2 = ... = \tau_6 = 0 \quad (5)$$
$$H_1 \;:\; \tau_i \neq 0 \text{ for at least one } i \quad (6)$$

The procedure for testing the equality of $a = 6$ treatment means is the analysis of variance. Using the R code 1.1 in the Appendix, we get the following analysis of variance (ANOVA) table below:

Source of Variation	Sum of Squares	Degrees of Freedom	Mean Squares	F-statistic	Pr(>F)
between treatments	1220	5	244.00	16	5.78e-07
residual/error	366	24	15.25		

Based on our ANOVA table, we could gather the following:

- total corrected sum of squares: $\text{SS}_\text{T} = 1220 + 366 = 1586$ which is a measure of overall variability in the data,
- sum of squares due to treatments: $\text{SS}_\text{Treatments} = 1220$
- sum of squares due to error: $\text{SS}_\text{E} = 366$.

Given that the between treatments mean square (244) is many times larger than the error mean square (15.25), this means that it is unlikely for the treatment means to be equal.

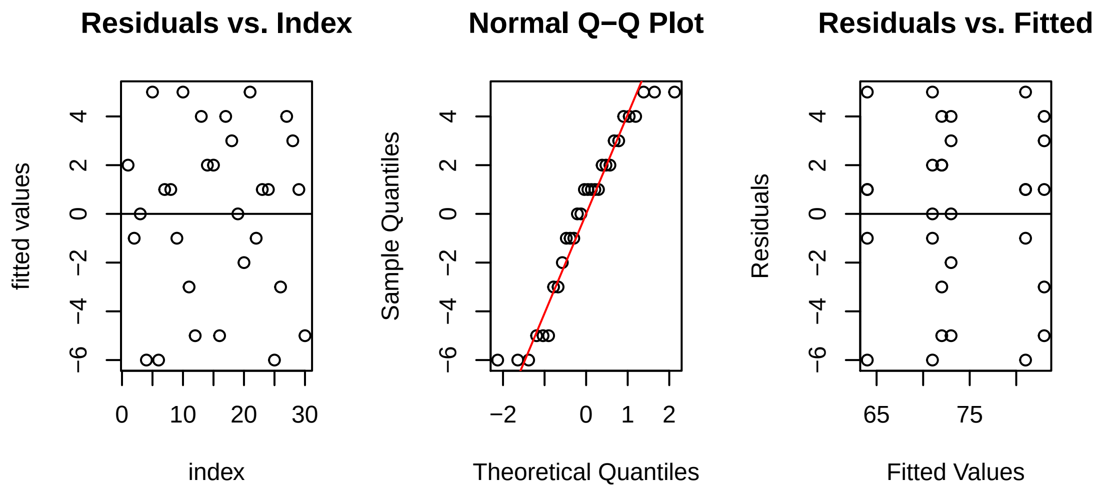
More formally, we could compute the F-statistic ($\text{F} = 244/15.25 = 16$) and compare it to the upper-tail percentage point of the $\text{F}_{0.05,5,24}$ distribution. Since $\text{F} > \text{F}_{0.05,5,24} = 2.620654$, we would reject H_0 and conclude that the treatment means differ. Additionally, since the p-value (5.78e-07) is less than $\alpha = 0.05$, this also tells us that at a significance level of $\alpha = 0.05$, there is sufficient evidence to suggest that treatment means differ. In other words, the brand of the fertilizer significantly affects the average production of wheat.

1.4 Model Adequacy Checking

Before we could use partitioning to test formally for no differences in treatment means, we need to check to check if certain assumptions are satisfied. More specifically, these assumptions are that the observations are adequately described by the model,

$$y_{ij} = \mu + \tau_i + \epsilon_{ij}$$

and that the errors are normally and independently distributed with mean zero and constant but unknown variance σ^2. If the model satisfies the normality, equal variances, and independence assumptions, then the model is adequate/stable. To do so, we used the R code 1.2 from the appendix to produce the following plots:



Since the points fall approximately along the diagonal line on the Q-Q plot above, the normality assumption is satisfied. As for the Residuals vs. Fitted plot and the Residual vs. Index plot, the residuals appear to be structureless without any obvious pattern for both which means that the equal variances and independence assumptions are satisfied.

1.5 Estimation of Model Parameters

We are now interested in the estimating the parameters of our single-factor model,

$$y_{ij} = \mu + \tau_i + \epsilon_{ij}$$

and the confidence intervals on the treatment means. In general, the reasonable estimates of the overall mean and treatment effects are given by

$$\hat{\mu} = \bar{y}_{..} \tag{7}$$
$$\hat{\tau}_i = \bar{y}_{i.} - \bar{y}_{..}, \; i = 1, 2, ..., 6 \tag{8}$$

For our particular dataset and using the R code 1.3 in the Appendix, we estimated our overall mean and treatment effects to be:

$$\hat{\mu} = 74 \tag{9}$$
$$\hat{\tau}_1 = 71 - 74 = -3 \tag{10}$$
$$\hat{\tau}_2 = 81 - 74 = 7 \tag{11}$$
$$\hat{\tau}_3 = 72 - 74 = -2 \tag{12}$$
$$\hat{\tau}_4 = 73 - 74 = -1 \tag{13}$$
$$\hat{\tau}_5 = 64 - 74 = -10 \tag{14}$$
$$\hat{\tau}_6 = 83 - 74 = 9 \tag{15}$$

To find the estimates of each treatment mean, we could compute the 95 percent confidence interval on the mean of each treatment. In general, the 100(1-α) percent confidence interval on the ith treatment mean μ_i is

$$\bar{y}_{i.} - t_{\alpha/2,N-a}\sqrt{\frac{MS_E}{n}} \leq \mu_i \leq \bar{y}_{i.} + t_{\alpha/2,N-a}\sqrt{\frac{MS_E}{n}}.$$

Using the formula above, we computed the following 95 percent confidence intervals:

$$67.39556 \leq \mu_1 \leq 74.60444 \tag{16}$$
$$77.39556 \leq \mu_2 \leq 84.60444 \tag{17}$$
$$68.39556 \leq \mu_3 \leq 75.60444 \tag{18}$$
$$69.39556 \leq \mu_4 \leq 76.60444 \tag{19}$$
$$60.39556 \leq \mu_5 \leq 67.60444 \tag{20}$$
$$79.39556 \leq \mu_5 \leq 86.60444 \tag{21}$$

1.6 Comparing Pairs of Treatment Means

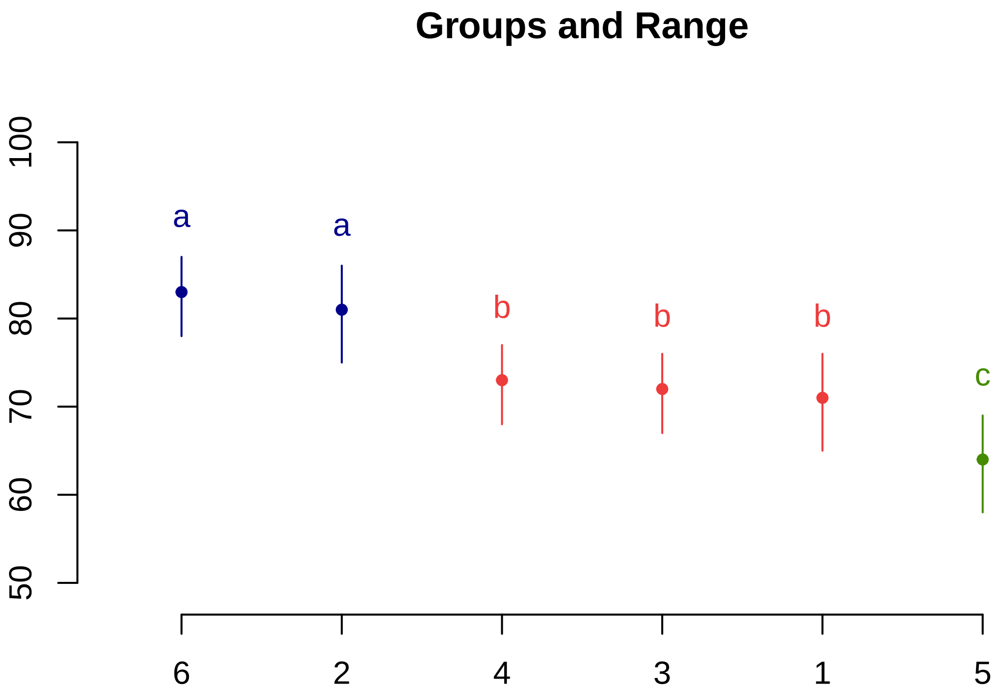
From our analysis of the fixed model, we could conclude that the brand of the fertilizer significantly affects the average production of wheat. In other words, we know that that some fertilizer brands produce different production rates than others, but we don't know which ones actually cause this difference. Thus, to find out, we could test the differences between all pairs of treatment means using the Fisher Least Significant Difference (LSD) Method. This procedure uses the t-statistic for testing the null hypotheses,

$$H_0 : \mu_i = \mu_j,\ i \neq j.$$

For the Fisher LSD method, we would simply compare the observed differences between each pair of averages to the corresponding least significant difference

$$\text{LSD} = t_{\alpha/2,N-a}\sqrt{\frac{2MS_E}{n}}.$$

If $|\bar{y}_{i.} - \bar{y}_{j.}| > \text{LSD}$, then we could conclude that the population means μ_i and μ_j differ.



The LSD plot above shows that the fertilizer brand #6 results in the most production. Additionally, the Fisher LSD procedure, from the R code 1.4 in the Appendix, also indicates that brand #6 is significantly different from other brands and its average production exceeds that of the other brands.

1.7 Conclusion

In conclusion, based on the residual plots, I found that the experiment results does represent an adequate model so my analysis of variance could be adopted. Thus, based on my analysis of variance tests, I found that the brand of fertilizer does affect the average production of wheat bushels. In fact, based on the Fisher Least Significant Difference test, we could see that the fertilizer brand #6 is the recommended fertilizer since it is significantly different from other brands and it produces the highest average production among the average production of the other brands.

2 Part II: Battery Life

2.1 Introduction

An engineer is designing a battery for use. He has three possible plate materials to use. He is aware that temperature may affect the effective battery life and decides to test all three plate materials at three temperature levels (15, 70 and 125°F). After testing all three plate, he then recorded the resulting observed battery life (in hours). Given this data, we are now interested in finding out what effects do the material types and temperatures have on the battery life and if possible, which choice of material would give a uniformly long battery life regardless of the temperature. Finding a material alternative that is not greatly affected by temperature is important for engineering because doing so would allow the engineer to produce a battery that is robust to temperature variation in the field. To find achieve our objectives, we will formulate and test hypotheses about the treatment effects such as the significant effects of each material type or temperature or the interaction between the material type and temperature.

2.2 Data and Model Description

The experiment and the resulting observed battery life (in hours) are given below.

	Temperature (°F)					
Material Type	15		70		125	
I	147	153	182	178	128	124
II	157	161	144	150	121	119
III	132	126	156	162	109	107

This experiment is a 3^2 factorial design since there are 3 levels (of temperature) and 2 factors (plate material and temperature). We could model the obserations by the following effects model:

$$y_{ijk} = \mu_i + \epsilon_{ijk} \begin{cases} i = 1, 2, 3 \\ j = 1, 2, 3 \\ k = 1, 2 \end{cases} \tag{22}$$

where the mean of the ijth cell is

$$\mu_{ij} = \mu + \tau_i + \beta_j + (\tau\beta)_{ij}.$$

2.3 Two-Factor Analysis of Variance

Since both row and column factors, material type (factor A) and temperature (factor B), are of equal interest, we are interested in testing hypotheses about the equality of material type effects, say

$$H_0 \ : \ \tau_1 = \tau_2 = \tau_3 = 0 \tag{23}$$
$$H_1 \ : \ \text{at least one } \tau_i \neq 0 \tag{24}$$

and the equality of temperature effects, say

$$H_0 \ : \ \beta_1 = \beta_2 = \beta_3 = 0 \tag{25}$$
$$H_1 \ : \ \text{at least one } \beta_i \neq 0. \tag{26}$$

We are also interested in finding out whether the material type and temperature effects interact so we wish to test

$$H_0 : (\tau\beta)_{ij} = 0 \text{ for all } i, j \tag{27}$$
$$H_1 : \text{at least one } (\tau\beta)_{ij} \neq 0 \tag{28}$$

Using the R code 2.1 in the Appendix, we get the following analysis of variance (ANOVA) table below:

Source	Sum of Squares	degrees of freedom	Mean Squares	F-statistic	Pr(>F)
Material type	1200	2	600.0	54.0	9.71e-06
Temperature	5952	2	2976.0	267.8	9.58e-09
Interaction	1200	4	300.0	27.0	5.00e-05
Error	100	9	11.1		
Total	8452	17			

Based on the ANOVA table above, we could gather the following three conclusions:

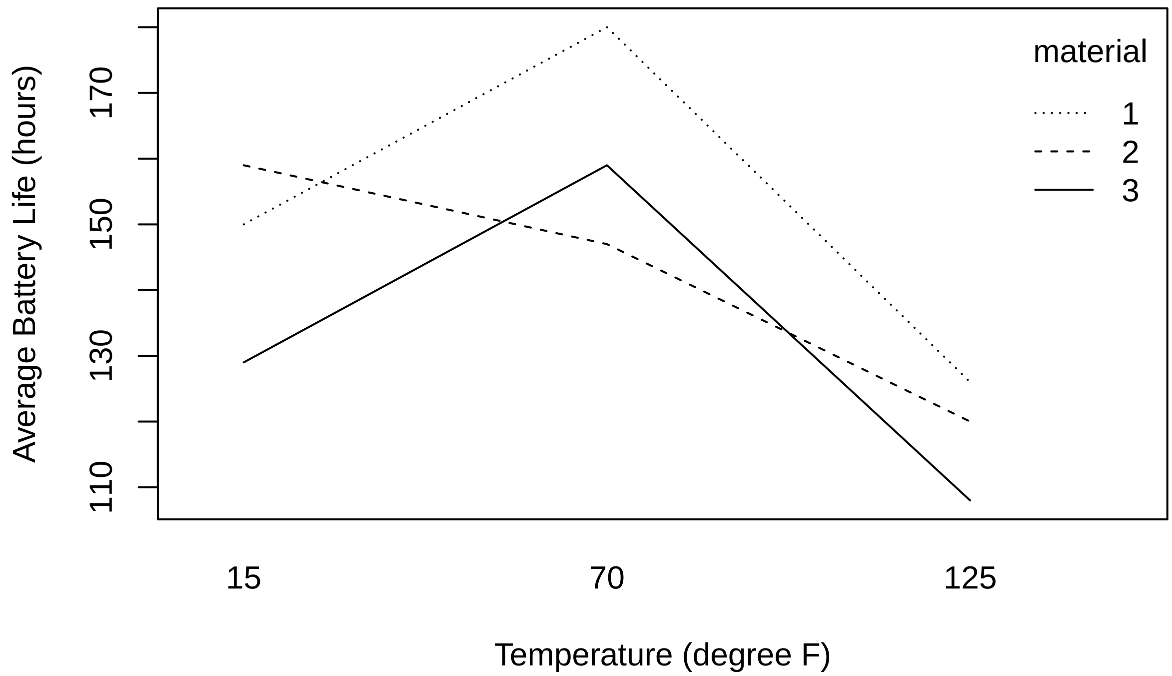
- the main effects of material type are significant since the p-value (9.71e-06) for material type is less than $\alpha = 0.05$,
- the main effects of temperature are significant since the p-value (9.58e-09) for temperature is less than $\alpha = 0.05$, and
- there is a significant interaction between the material type and temperature since the p-value (5.00e-05) for the interaction is less than $\alpha = 0.05$.

Additionally, given

$$R^2 = \frac{SS_{Model}}{SS_{Total}} \tag{29}$$
$$= \frac{SS_{Material} + SS_{Temperature} + SS_{Interaction}}{SS_{Total}} \tag{30}$$
$$= \frac{1200 + 5952 + 1200}{8452} \tag{31}$$
$$= 0.9881685, \tag{32}$$

this means that about 98.82 percent the variability in the battery life is explained by the plate material in the battery, the temperature, and the material type-temperature interaction.

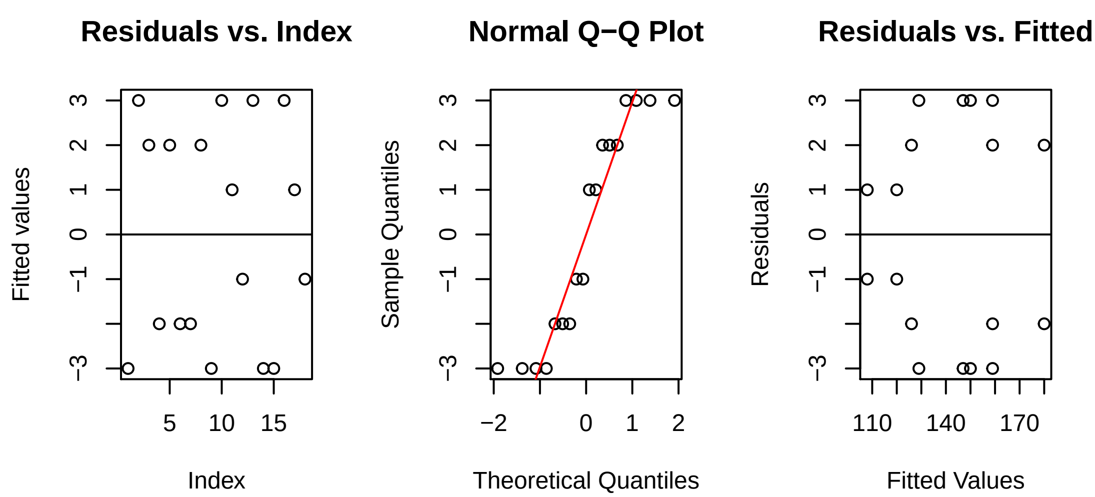
To get a better interpretation of the results of this experiment, we could graph the average responses at each treatment combination using the R code 2.2 in the Appendix.



From the interaction plot above, we could see that the significant interaction is indicated by the lack of parallelism among lines. In general, it appears that longer life is attained at the lower to intermediate levels of temperature regardless of the material type. For material type 1 and 3, the average battery life span is seems to have a maximum cap around the temperature of 70°F and decreases in average battery life span if the temperature is otherwise. However, for material type 2, its average battery life span is almost linear, i.e. it increases as temperatures decreases and vice versa. Overall, it seems that material type 1 has the best average battery life span for all temperatures.

2.4 Model Adequacy Checking

Before we could assume the conclusions from the ANOVA, we need to check the adequacy of the underlying model. To do so, we used the R code 2.3 to graph the Residual vs. Index, Residual vs. Fitted Values, and Q-Q plots to check for independency, equal variances, and normality.



Since the points fall approximately along the diagonal line on the Q-Q plot above, the normality assumption is satisfied. As for the Residuals vs. Fitted plot and the Residual vs. Index plot, the residuals appear to be structureless without any obvious pattern for both which means that the equal variances and independence assumptions are satisfied.

2.5 Conclusion

In conclusion, based on the residual plots, I found that the experiment results does represent an adequate model so my analysis of variance could be adopted. Thus, based on my analysis of variance tests, I found that the temperature and material type do significantly affect the battery life. In fact, based on my hypotheses testings and the interaction plots, I found that the material type 1 has the best average battery life span for all temperatures.

3 Appendix

```r
# 1.1
wheat <- data.frame(obs = c(73, 70, 71, 65, 76, 75, 82,
                82, 80, 86, 69, 67, 76, 74, 74, 68, 77, 76,
                73, 71, 69, 63, 65, 65, 58, 80, 87, 86, 84, 78),
                level = factor(rep(seq(1, 6, by = 1), each = 5)))

fit <- aov(obs ~ level, data = wheat)
#summary(fit)

# 1.2
par(mfrow = c(2, 3), cex = 0.75, oma = c(0.75, 0.75, 0.75, 0.75))

plot(fit$residuals, ylab = 'fitted values', xlab = 'index',
     main = "Residuals vs. Index")
abline(h = 0)

qqnorm(fit$residuals)
qqline(fit$residuals, col = "red")

plot(fit$fitted.values, fit$residuals,
     xlab="Fitted Values",ylab="Residuals",main="Residuals vs. Fitted")
abline(h = 0)


# 1.3 Model parameter estimations
mean(wheat$obs)
x <- aggregate(wheat$obs, list(wheat$level), mean)
t <- qt(0.05/2, 24, lower.tail = FALSE)
MSE <- 15.25
CI_lower <- x-t*sqrt(MSE/5)
CI_upper <- x+t*sqrt(MSE/5)
x; CI_lower; CI_upper


# Fisher LSD Test
library(agricolae)
plot(LSD.test(wheat$obs, wheat$level, DFerror = 24,
              MSerror = 15.25, alpha = 0.06))

# 1.4
LSD.test(wheat$obs, wheat$level, DFerror = 24, MSerror = 15.25,
         alpha = 0.06, console = T)


# 2.1
```

```r
life <- c(147, 153, 182, 178, 128, 124,
          157, 161, 144, 150, 121, 119,
          132, 126, 156, 162, 109, 107)
temp <- factor(c(rep(c(15, 70, 125), each = 2, times = 3)))
material <- factor(c(rep(1:3, each = 6)))
# cbind(temp, material, life)
battery.aov <- aov(life ~ material*temp)
#summary(battery.aov)


# 2.2
interaction.plot(temp, material, life, ylab = 'Average Battery Life (hours)',
                 xlab = 'Temperature (degree F)')


# 2.3
par(mfrow = c(2, 3), cex = 0.75, oma = c(0.75, 0.75, 0.75, 0.75))

plot(battery.aov$residuals, ylab = 'Fitted values', xlab = 'Index',
     main = "Residuals vs. Index")
abline(h = 0)

qqnorm(battery.aov$residuals)
qqline(battery.aov$residuals, col = "red")

plot(battery.aov$fitted.values, battery.aov$residuals,
     xlab="Fitted Values",ylab="Residuals",main="Residuals vs. Fitted")
abline(h = 0)
```