STAT 4100 HW 6

Cody Frisby

February 26, 2016

1) Chemical Process, Temperature and Pressure

A) Model The model for this experiment is a two factor, three level factorial experiment where we have 2 replicates. The statistical model is represented by

$$y_{ijk} = \mu + \tau_i + \beta_j + (\tau\beta)_{ij} + \varepsilon_{ijk} \begin{cases} i = 1, 2, 3 \\ j = 1, 2, 3 \\ k = 1, 2 \end{cases}$$

where μ is the overall mean of the response: **yield**, τ_i is the ith effect from **temperature**, β_j is the jth effect from **pressure**, $(\tau\beta)_{ij}$ is the effect from the interaction of **pressure** and **temperature**, and ε_{ijk} are the errors. Here, like before, we assume the errors are independent, identically distributed $\sim N(0, \sigma^2)$.

B) ANOVA

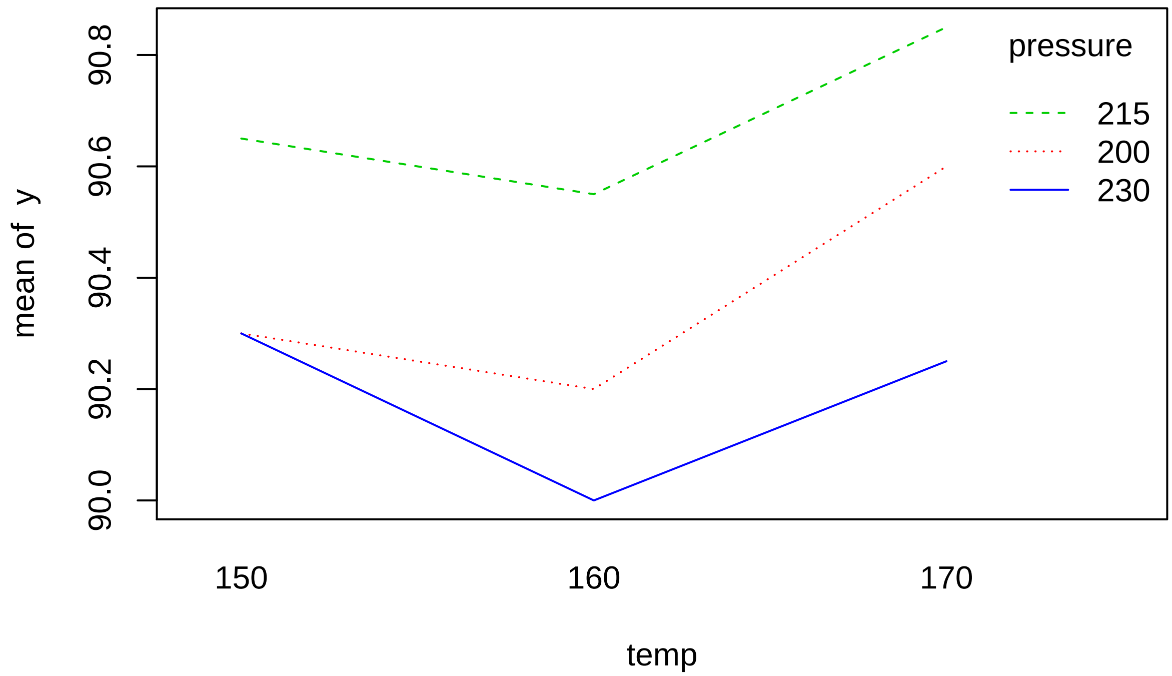
Here I fill the ANOVA table. The R code where I compute it by hand is at the end of this document.

Source	Sum Sq	Df	Mean Sq	F value	Pr(>F)
temp	0.3011111	2	0.1505556	8.46875	0.0085392
pressure	0.7677778	2	0.3838889	21.59375	3.6731788×10^{-4}
temp:pressure	0.0688889	4	0.0172222	0.96875	0.4700058
Error	0.16	9	0.0177778		
Total	1.2977778	17			

We would conclude that the main effects **temperature** and **pressure** have a significant effect on yield where the interaction of the two does not. Here I display the data used to calculate $SS_{subgroups}$ as a step for calculating SS_{AB}.

temp	pressure	sum(y)
150	200	180.6
150	215	181.3
150	230	180.6
160	200	180.4
160	215	181.1
160	230	180.0
170	200	181.2
170	215	181.7
170	230	180.5

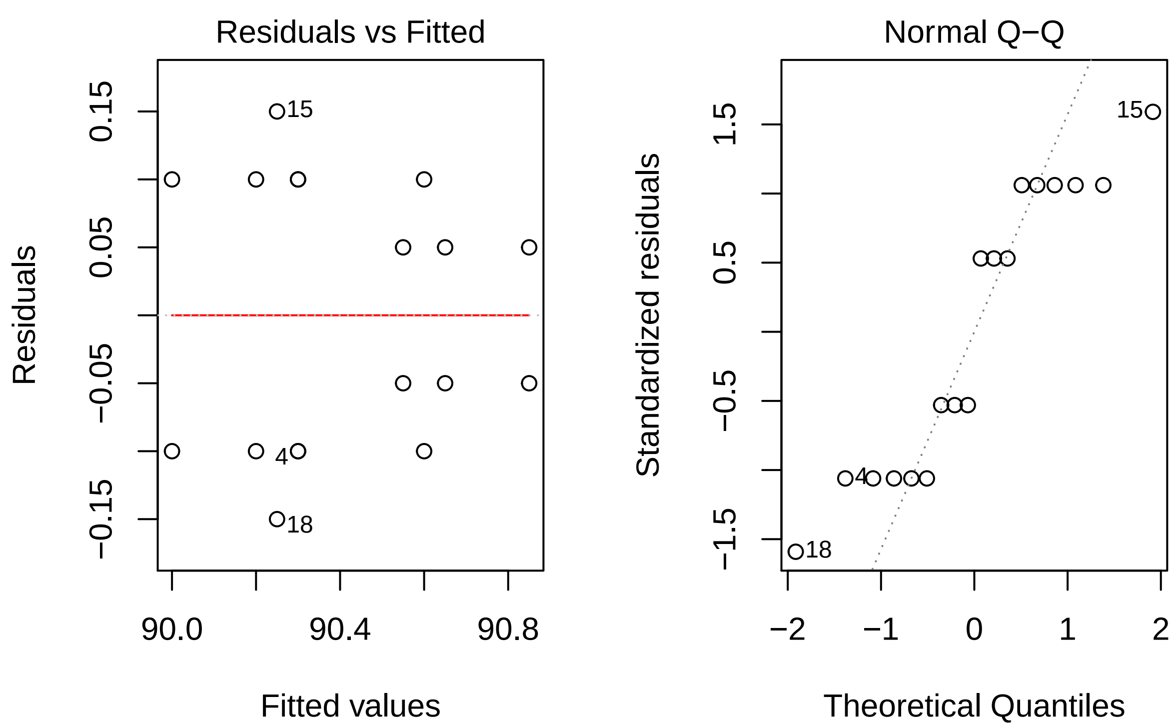
C)Interaction plot



Yes, this appears to be consistent with our F test from the ANOVA table above. There does not appear to be any interactions between **pressure** and **temperatrue**.

D) Model Adequacy

Here we plot the residuals vs fitted and qqplot of the residuals.



The model assumptions are suspect. The QQplot is questionable and looking at the residuals vs fitted plot there may not be equality of variances. Our conclusions and model may need reexamination.

2) Glass Experiment

Here we have 3^2 experiment, 2 factors with 3 levels. The number of replicates, n, is 3.

A) Model

The statistical model is represented by

$$y_{ijk} = \mu + \tau_i + \beta_j + (\tau\beta)_{ij} + \varepsilon_{ijk} \begin{cases} i = 1,2,3 \\ j = 1,2,3 \\ k = 1,2,3 \end{cases}$$

where μ is the overall mean of the response: **light output**, τ_i is the ith effect from **glass type**, β_j is the jth effect from **temperature**, $(\tau\beta)_{ij}$ is the effect from the interaction of **glass type** and **temperature**, and ε_{ijk} are the errors. Here, like before, we assume the errors are independent, identically distributed $\sim N(0, \sigma^2)$.

B) ANOVA and hypothesis test

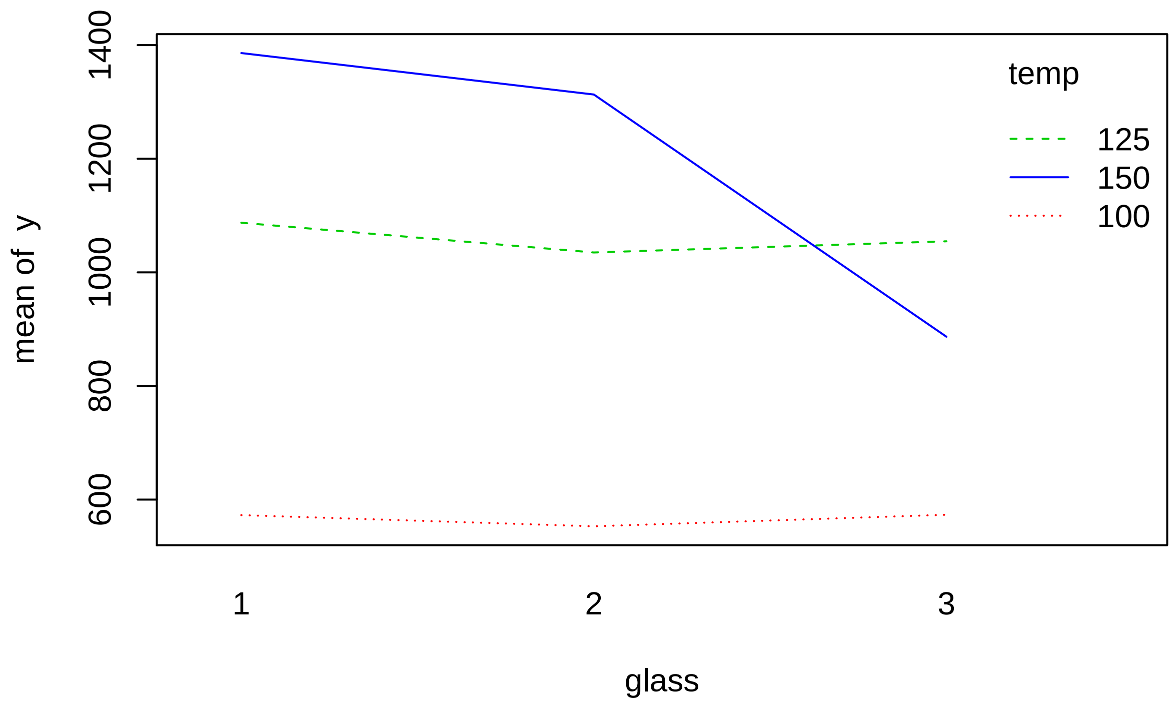
Here, we use the data to fill an ANOVA table and run hypothesis tests on the main effects and interaction.

Source	Sum Sq	Df	Mean Sq	F value	Pr(>F)
glass type	1.5086452×10^5	2	7.5432259×10^4	206.3705543	$3.8857806 \times 10^{-13}$
temperature	1.9703345×10^6	2	9.8516726×10^5	2695.2594994	0
glass:temp	2.905517×10^5	4	7.2637926×10^4	198.7257068	1.254552×10^{-14}
Error	6579.3333333	18	365.5185185		
Total	2.4183301×10^6	26			

We can conclude there is a significant effect from the main effects and the interaction of the **glass type** and **temperature**.

C) Interaction Plots

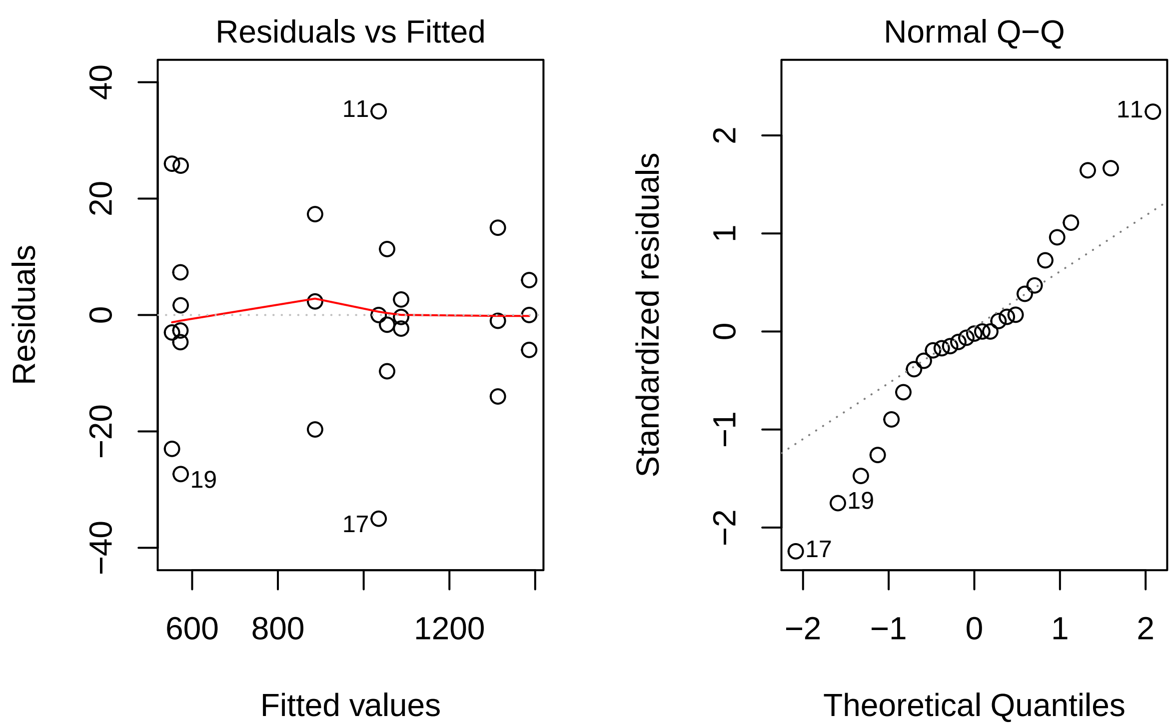
Here, we use R to create an interaction plot



There is some interaction here between factors and large differences on the response from the main effects. This plot, visually, agrees with our F test from the ANOVA table above.

D) Model Adequacy

Here we analyze the residuals from the above model and comment on the the adequacy of the assumptions.

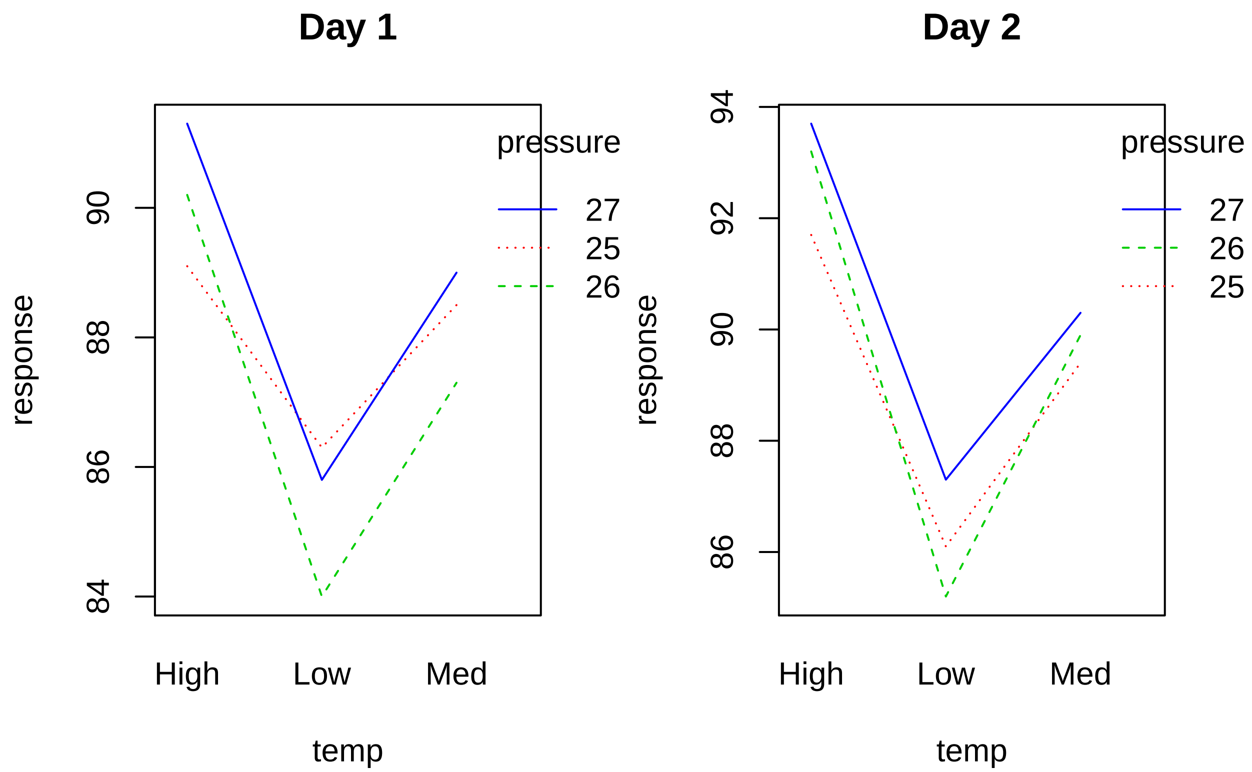


There appears to be a serious violation of the normality of the model errors assumption looking at the qqplot of the residuals. There may also be some non equality of variances from the residuals vs fitted plot but nothing to be too concerned about. The normality assumption appears to be violated and will need reexamination. Our conclusions above may not be valid considering this violation.

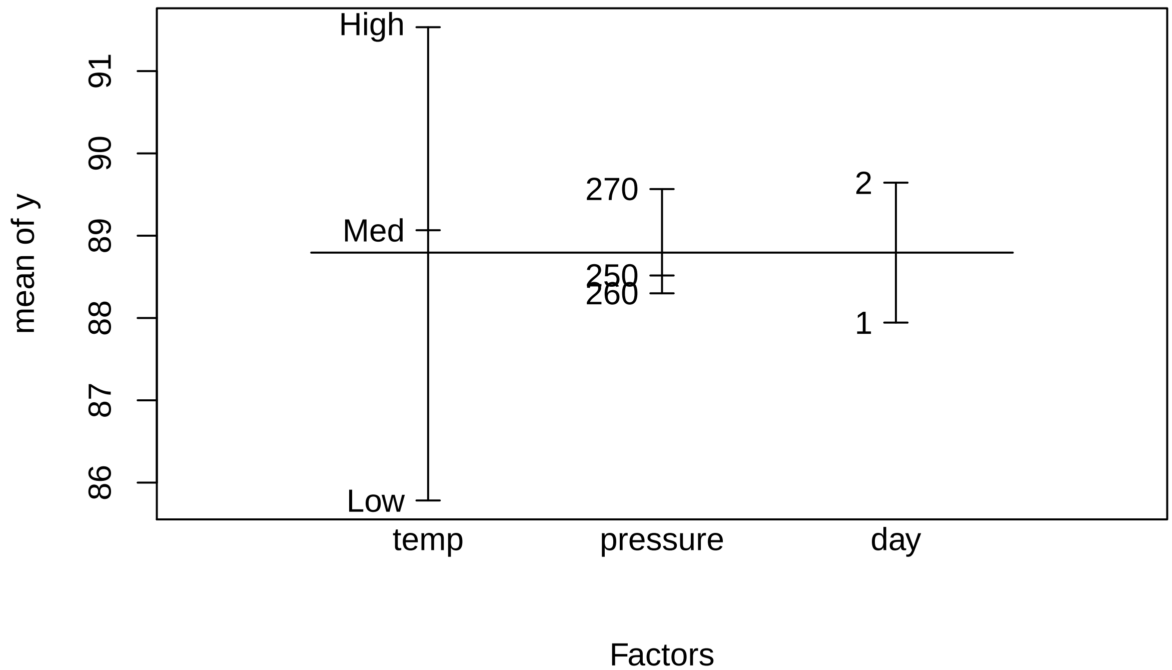
3) Factorial with Blocking

A) Interaction Plots

Here, I plot two interaction plots, seperating by block (day).



It appears there is some interaction between the factors and an effect from **pressure** and **temperature**. How much, or how significant, we would need to fit a model and run statistical tests on the main effects and interactions.
I like to look at a plot of the model. I use R to draw the plot here:



One thing we should keep in mind is that we assume no interactions between block and all other terms. Here, blocks, is represented as δ_k. Symbolically we are assuming there is no $(\tau\delta)_{ik}$, $(\beta\delta)_{jk}$, $(\tau\beta\delta)_{ijk}$. These interactions are contained in the error term, SS_{errors}. Additionaly, we consider

$$\sigma_\delta^2 = \frac{MS_{blocks} - MS_E}{ab}$$

B) Analysis of the data

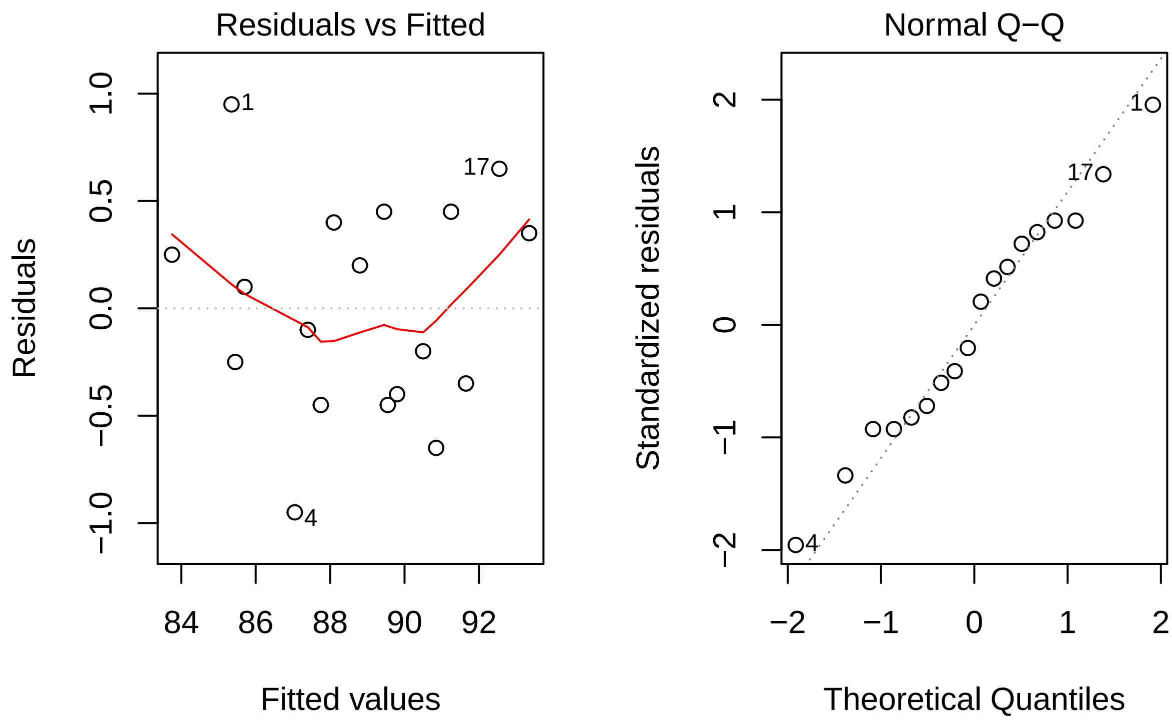
The calculations by hand are shown in the R code at the end of this document.

Source	Sum Sq	Df	Mean Sq	F value	Pr(>F)
blocks	13.005	1	13.005	24.48	0.0011241
glass type	99.8544444	2	49.9272222	93.9806536	2.7776571×10^{-6}
temperature	5.5077778	2	2.7538889	5.1837908	0.0359875
glass:temp	4.4522222	4	1.1130556	2.0951634	0.1733137
Error	4.25	9	0.53125		
Total	127.0694444	17			

We would conclude that **temp** and **pressure** have a significant effect on the response but not the interaction of the two. Also, it appears that the days did have an effect, it was good to block by day. This reduced our error term that would have otherwise been much larger had we not blocked by day.

C) Model Adequacy

We need to examine the residual plots from the above model.



We are safe in concluding that the model assumptions are valid. There is nothing we need be concerned about regaring normailty and equality of variance.

R Code:

```
# 1
temp <- c(rep(150,6), rep(160,6), rep(170,6))
pressure <- rep(c(200,215,230), 6)
y <- c(90.4, 90.7, 90.2, 90.2, 90.6, 90.4, 90.1,
       90.5, 89.9, 90.3, 90.6, 90.1, 90.5, 90.8, 90.4, 90.7, 90.9, 90.1)
D <- data.frame(y, temp, pressure)
D2 <- D
library(dplyr)
# aov function in R needs the predictors to be factors
D2$temp <- as.factor(D2$temp); D2$pressure <- as.factor(D2$pressure)
fit <- aov(y ~ .+temp*pressure, D2)
# build the ANOVA table by "hand"
y.i <- D %>%
  group_by(temp) %>%
  summarise_each(funs(sum), y)
y.j <- D %>%
  group_by(pressure) %>%
  summarise_each(funs(sum), y)
yij. <- D %>%
  group_by(temp, pressure) %>%
  summarise_each(funs(sum), y)
y.. <- sum(y)
a <- 3; b <- 3; n <- 2; N <- a*b*n
```

```r
ssa <- (1/(b*n))*sum(y.i$y^2) - y..^2/N
ssb <- (1/(a*n))*sum(y.j$y^2) - y..^2/N
sssub <- (1/n)*sum(yij.$y^2) - y..^2/N
ssab <- sssub - ssa - ssb
sst <- sum(y^2) - y..^2/N
sse <- sst - ssab - ssa - ssb
msa <- ssa/(a-1); msb <- ssb/(b-1); msab <- ssab/((a-1)*(b-1))
mse <- sse/(a*b*(n-1))
Fa <- msa/mse; Fb <- msb/mse; Fab <- msab/mse
df <- a*b*(n-1)
p.Fa <- 1 - pf(Fa, a-1, df); p.Fb <- 1 - pf(Fb, b-1, df)
p.Fab <- 1 - pf(Fab, (a-1)*(b-1), df)
A <- yij.; names(A) <- c("temp", "pressure", "sum(y)")
knitr::kable(A)
interaction.plot(temp, pressure, y, col = 2:4)
par(mfrow=c(1,2))
plot(fit, which = c(1,2))
# 2
# get the data into R.  There isn't a file so we will type it in
y <- c(580, 1090, 1392, 568, 1087, 1380, 570, 1085, 1386, 550, 1070,
       1328, 530, 1035, 1312, 579, 1000, 1299, 546, 1045, 867, 575,
       1053, 904, 599, 1066, 889)
glass <- as.factor(c(rep(1,9), rep(2,9), rep(3,9)))
temp <- as.factor(rep(c(100,125,150), 9))
E <- data.frame(y, glass, temp)
library(dplyr)
# build the ANOVA table by "hand"
y.i <- E %>%
  group_by(glass) %>%
  summarise_each(funs(sum), y)
y.j <- E %>%
  group_by(temp) %>%
  summarise_each(funs(sum), y)
yij. <- E %>%
  group_by(glass, temp) %>%
  summarise_each(funs(sum), y)
y.. <- sum(y)
a <- 3; b <- 3; n <- 3; N <- a*b*n
ssa <- (1/(b*n))*sum(y.i$y^2) - y..^2/N
ssb <- (1/(a*n))*sum(y.j$y^2) - y..^2/N
sssub <- (1/n)*sum(yij.$y^2) - y..^2/N
ssab <- sssub - ssa - ssb
sst <- sum(y^2) - y..^2/N
sse <- sst - ssab - ssa - ssb
msa <- ssa/(a-1); msb <- ssb/(b-1); msab <- ssab/((a-1)*(b-1))
mse <- sse/(a*b*(n-1))
Fa <- msa/mse; Fb <- msb/mse; Fab <- msab/mse
df <- a*b*(n-1)
p.Fa <- 1 - pf(Fa, a-1, df); p.Fb <- 1 - pf(Fb, b-1, df)
p.Fab <- 1 - pf(Fab, (a-1)*(b-1), df)
fit2 <- aov(y ~ glass*temp, E)
interaction.plot(glass, temp, y, col = 2:4)
par(mfrow=c(1,2))
```

```
plot(fit2, which = c(1,2))
# 3
# get the data into R.  There isn't a file so we will type it in
y <- c(86.3, 84, 85.8, 86.1, 85.2, 87.3, 88.5, 87.3, 89, 89.4, 89.9,
       90.3, 89.1, 90.2, 91.3, 91.7, 93.2, 93.7)
temp <- as.factor(c(rep("Low",6), rep("Med",6), rep("High",6)))
pressure <- as.factor(rep(c(250,260,270), 6))
day <- as.factor(rep(c(1,1,1,2,2,2), 3))
G <- data.frame(y, temp, pressure, day)
day1 <- subset(G, G$day == "1")
day2 <- subset(G, G$day == "2")
par(mfrow=c(1,2))
interaction.plot(day1$temp, day1$pressure, day1$y,
                 col = 2:4, xlab = "temp", ylab="response",
                 trace.label = "pressure", main = "Day 1")
interaction.plot(day2$temp, day2$pressure, day2$y, col = 2:4,
                 xlab = "temp", ylab="response",
                 trace.label = "pressure", main = "Day 2")
plot.design(G)
# using R to calculate the model
fit.G <- aov(y ~ day+temp*pressure, G)
# build the ANOVA table by "hand"
y..k <- G %>% # blocking
  group_by(day) %>%
  summarise_each(funs(sum), y)
y.i <- G %>% # rows
  group_by(temp) %>%
  summarise_each(funs(sum), y)
y.j <- G %>% # columns
  group_by(pressure) %>%
  summarise_each(funs(sum), y)
yij. <- G %>% # subgroups
  group_by(temp, pressure) %>%
  summarise_each(funs(sum), y)
y.. <- sum(y) # sum of all obs
a <- 3; b <- 3; n <- 2; N <- a*b*n
blocks <- (1/(a*b))*sum(y..k$y^2) - y..^2/N #ss due to blocks
ssa <- (1/(b*n))*sum(y.i$y^2) - y..^2/N #ss due to row
ssb <- (1/(a*n))*sum(y.j$y^2) - y..^2/N # ss due to col
sssub <- (1/n)*sum(yij.$y^2) - y..^2/N # ss sub for interaction
ssab <- sssub - ssa - ssb # ss due to interaction
sst <- sum(y^2) - y..^2/N # total sum of squares
sse <- sst - ssab - ssa - ssb - blocks # ss errors
msa <- ssa/(a-1); msb <- ssb/(b-1); msab <- ssab/((a-1)*(b-1))
msblocks <- blocks/(n-1)
mse <- sse/((a*b-1)*(n-1))
sigdelta <- (msblocks - mse)/(a*b)
Fa <- msa/mse; Fb <- msb/mse; Fab <- msab/mse; Fbl <- msblocks/mse
df <- (a*b-1)*(n-1)
p.Fa <- 1 - pf(Fa, a-1, df); p.Fb <- 1 - pf(Fb, b-1, df)
p.Fab <- 1 - pf(Fab, (a-1)*(b-1), df)
p.Fbl <- 1 - pf(Fbl, n-1, df)
```