Exam 2

Cody Frisby

March 16, 2016

1) Depth of Cut and Feed Rate

A) Model The model for this experiment is a two factor, three & four level factorial experiment where we have 3 replicates. The statistical model is represented by

$$y_{ijk} = \mu + \tau_i + \beta_j + (\tau\beta)_{ij} + \varepsilon_{ijk} \begin{cases} i = 1, 2, 3 \\ j = 1, 2, 3, 4 \\ k = 1, 2, 3 \end{cases}$$

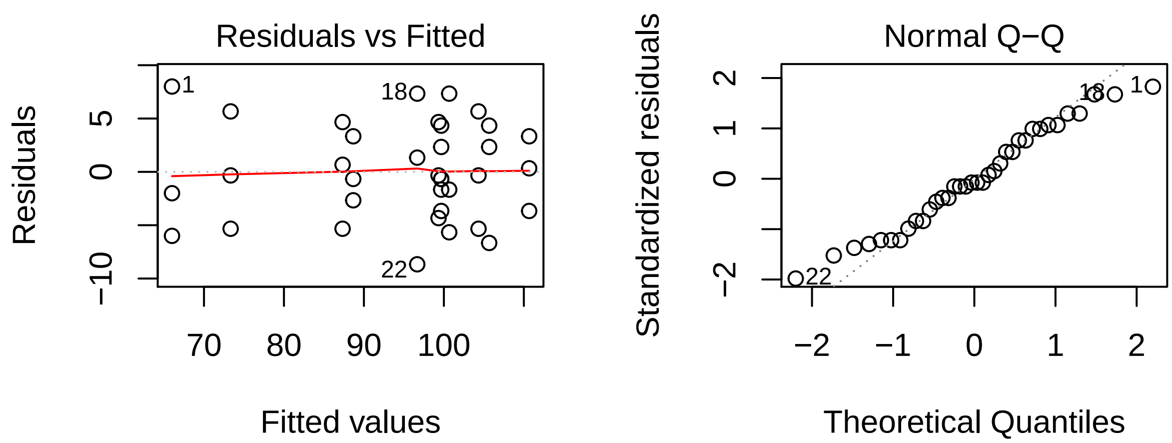
where μ is the overall mean of the response: **finish**, τ_i is the ith effect from **feed rate**, β_j is the jth effect from **depth of cut**, $(\tau\beta)_{ij}$ is the effect from the interaction of **feed rate** and **depth of cut**, and ε_{ijk} are the errors. Here we assume the errors are independent,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

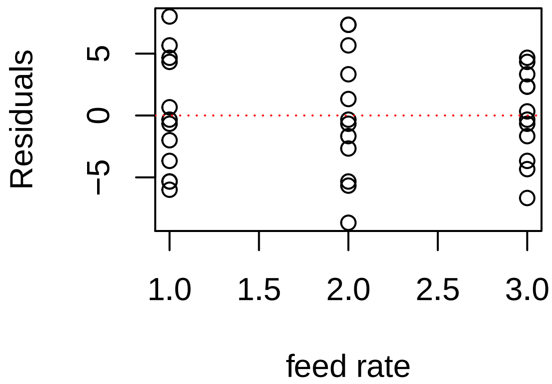
B) ANOVA We construct an ANOVA table based on the above model and model assumptions.

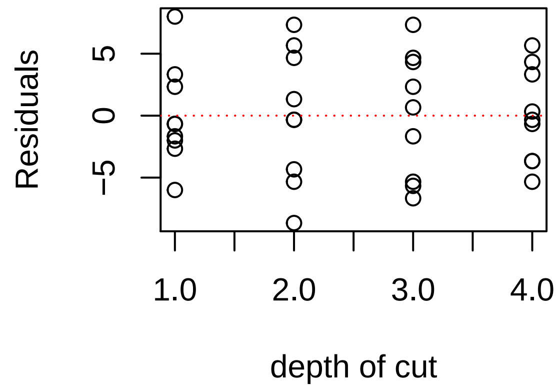
Source	Sum Sq	Df	Mean Sq	F value	Pr(>F)
feed	3160.5	2	1580.25	55.0183752	1.0860459×10^{-9}
depth	2125.1111111	3	708.3703704	24.6627982	1.6520005×10^{-7}
feed:depth	557.0555556	6	92.8425926	3.2324307	0.017973
Error	689.3333333	24	28.7222222		
Total	6532	35			

We conclude that both factors have a significant effect on the response as does their interaction.

C) Model Adequacy







There doesn't appear to be any violations of our models assumptions. Normality of the errors looks OK as well as equality of variances.

D) Point Estimates and Confidence Intervals - Feed Rate The point estimate for when feed is equal to **0.25** is 97.5833333. We calculate a confidence interval by

$$\hat{y} \pm t_{\alpha/2:df} \sqrt{\frac{MS_E}{n}}$$

$$97.5833333 \pm 2.0638986 \sqrt{\frac{28.7222222}{3}}$$

which results in a 95% confidence interval of $[91.1972183, 103.9694484]$ for $feed = 0.25$. We would do the same for when feed rate is equal to 0.3, 103.8333333, resulting in a 95% confidence interval of $[97.4472183, 110.2194484]$.

E) Point Estimates and Confidence Intervals - Depth of Cut Differences Similarly to part D, we first estimate the points when **depth of cut** is equal to 0.2 and 0.25, 97.8888889 and 104.8888889 respectively. We then calculate a confidence interval using the same formula as above but taking the difference between the two point estimates and multiplying MS_E by 2.

$$97.8888889 - 104.8888889 \pm (2.0638986) \sqrt{\frac{2 * 28.7222222}{3}}$$

$$[-16.0313305, 2.0313305]$$

2) Density of Particle Board

A) There are no replicates, *ie* $n = 1$, so the error degrees of freedom would result in zero since $abc(n-1) = 0$ for any abc.

B) Main Effect Target Estimate The main effect estimate for target is computed by taking the mean of y when target is equal to 48 minus the mean of y when target is equal to 42.

$$\frac{437.5}{9} - \frac{393.9}{9} = 4.8444444$$

C) ANOVA *Here, I display the ANOVA table for the model from this problem. The calculations were done in R by "hand". I will display the code at the end of this document.*

Source	Sum Sq	Df	Mean Sq	F value	Pr(>F)
target	105.6088889	1	105.6088889	254.3941117	2.3920374×10^{-7}
resin	52.1244444	2	26.0622222	62.7795249	1.2872612×10^{-5}
slash	9.7377778	2	4.8688889	11.7283372	0.0041832
resin:slash	2.6655556	4	0.6663889	1.6052191	0.2633865
Error	3.3211111	9	0.4151389		
Total	173.4577778	17			

Testing our model, at $\alpha = 0.01$, we conclude that all the effects are significant *except* the interaction of resin & slash.

D) "Sweet Spot" The best treatment combinations for minimizing the response would be resin = 6, slash = 0, target = 42.

3) Experiment with three factors with two levels and 4 replicates.

A table listing the main effects of a, b, and c respectively.

main.a	main.b	main.c
21	141.75	186.5

B) AB interaction effect The estimate for the AB interaction is 6.75.

C) ANOVA Table Here, using a computer, we fill in the ANOVA table.

Source	Sum Sq	Df	Mean Sq	F value
A	3528	1	3528	10.3384615
B	1.607445×10^5	1	1.607445×10^5	471.0461538
C	2.78258×10^5	1	2.78258×10^5	815.4080586
AB	364.5	1	364.5	1.0681319
AC	840.5	1	840.5	2.4630037
BC	364.5	1	364.5	1.0681319
ABC	112.5	1	112.5	0.3296703
Error	8190	24	341.25	
Total	466150	31		

Critical F value for $\alpha = 0.01$ is 7.8228706, so for any F value greater than this we would conclude that effect is significant. From the above ANOVA table it appears we have three, they are the main effects A, B, and C.

D) Factor Conclusions It appears the main effects have a significant effect on the response while their interaction with each other does not. B and C appear to have a stronger affect than A.

E) Best for Maximizing yield I would say a+ b+ c+ since the main effects of the three treatments are all positive.

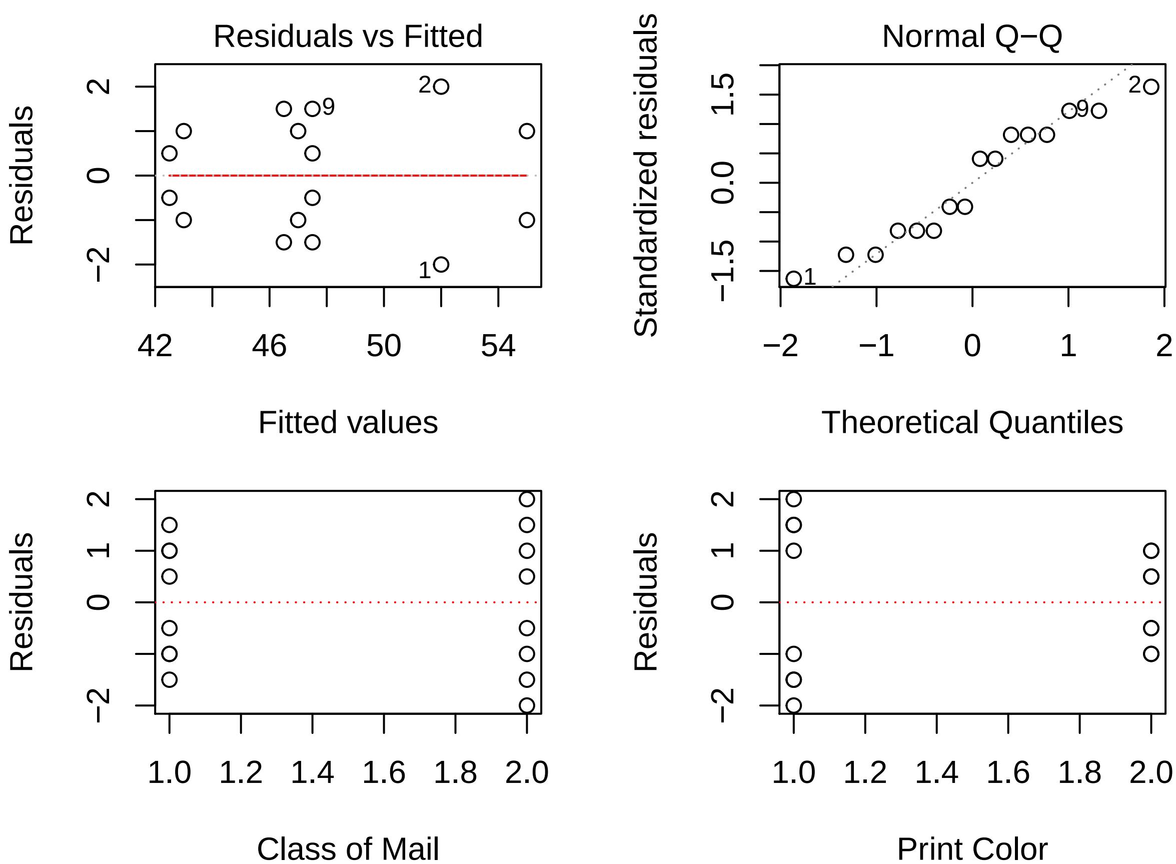
4) Marketing print experiment

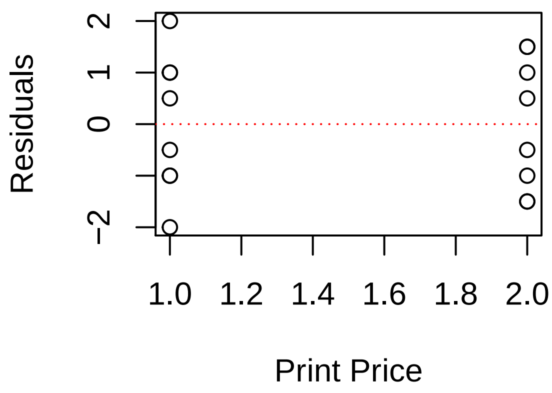
A) ANOVA Here, we look at an ANOVA table based on the model of main effects and interaction effects between all three variables. Here, $n = 2$, we have 2 replicates and 2^k, where $k = 3$, factor/level combinations.

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
A	1	12.25	12.25	4.083333	0.0779708
B	1	2.25	2.25	0.750000	0.4116944
C	1	36.00	36.00	12.000000	0.0085163
A:B	1	42.25	42.25	14.083333	0.0056019
A:C	1	100.00	100.00	33.333333	0.0004176
B:C	1	49.00	49.00	16.333333	0.0037282
A:B:C	1	4.00	4.00	1.333333	0.2815369
Residuals	8	24.00	3.00	NA	NA

It appears that the most significant factor is the *interaction* between **price** and **class** of the direct mail. Also significant are main effect C (price), AB interaction, and BC interaction.

B) Model Adequacy Lets take a look at the model residuals.





Model assumptions are OK. There's nothing that need concern us from the residual plots.

C) Recomendations If we wish to maximize orders then the reccomendation would be to print in color and send first class. If there is a desire to have a lower price ($19.95) while still maximizing orders, then the reccomendation would be to print in black and white and send 3rd class.

R Code:

```
feed <- c(rep(0.2,12), rep(0.25,12), rep(0.3,12))
depth <- rep(c(0.15,0.18,0.2, 0.25), 3)
y <- c(74,79,82,99,64,68,88,104,60,73,92,96,92,98,99,104,86,104,108,110,88,88,95,99,99,104,108,114,98,9
D <- data.frame(y, feed, depth)
D$feed <- as.factor(D$feed)
D$depth <- as.factor(D$depth)
library(dplyr)
y.i <- D %>%
  group_by(feed) %>%
  summarise_each(funs(sum), y)
y.j <- D %>%
  group_by(depth) %>%
  summarise_each(funs(sum), y)
yij. <- D %>%
  group_by(feed,depth) %>%
  summarise_each(funs(sum), y)
y.. <- sum(y)
a <- 3; b <- 4; 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)
```

```r
fit.1 <- aov(y ~ feed*depth, D)
#knitr::kable(anova(fit))
par(mfrow=c(1,2))
plot(fit.1, which = c(1,2))
par(mfrow=c(1,2))
plot(fit.1$residuals~as.numeric(D$feed), xlab = "feed rate",
     ylab = "Residuals"); abline(h=0, lty = 3, col = "red")
plot(fit.1$residuals~as.numeric(D$depth), xlab = "depth of cut",
     ylab = "Residuals"); abline(h=0, lty = 3, col = "red")
est.25 <- mean(D[D$feed == 0.25, 1])
est.3 <- mean(D[D$feed == 0.3, 1])
t0 <- qt(0.975, a*b*(n-1))
cv <- sqrt(mse/n)
cv2 <- sqrt(2*mse/n)
estd.2 <- mean(D[D$depth == 0.2, 1])
estd.25 <- mean(D[D$depth == 0.25, 1])
est <- estd.2 - estd.25
# get the data into R.  There isn't a file so we will type it in
y <- c(40.9,41.9,42,44.4,46.2,48.4,42.8,43.9,44.8,48.2,48.6,50.7,
       45.4,46,46.2,49.9,50.8,50.3)
resin <- as.factor(c(rep(6,6), rep(9,6), rep(12,6)))
slash <- as.factor(rep(c(0,25,50), 6))
target <- as.factor(rep(c(rep(42,3),rep(48,3)), 3))
G <- data.frame(y, resin, slash, target)
est.t <- mean(G[G$target == 48,1]) - mean(G[G$target == 42,1])
tar.48 <- G[G$target == 48,1]
tar.42 <- G[G$target == 42,1]
library(dplyr)
# using R to calculate the model
fit.G <- aov(y ~ target+resin*slash, data = G)
# build the ANOVA table by "hand"
y..k <- G %>% # blocking
  group_by(target) %>%
  summarise_each(funs(sum), y)
y.i <- G %>% # rows
  group_by(resin) %>%
  summarise_each(funs(sum), y)
y.j <- G %>% # columns
  group_by(slash) %>%
  summarise_each(funs(sum), y)
yij. <- G %>% # subgroups
  group_by(resin, slash) %>%
  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))
```

```r
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)
G$pred <- fit.G$fitted.values
one <- 425; a <- 426; b <- 1118; ab <- 1203; c <- 1283
ac <- 1396; bc <- 1670; abc <- 1807; n <- 4
# Part A, main effects
main.a <- (1/(4*n))*(a-one+ab-b+ac-c+abc-bc)
main.b <- (1/(4*n))*(b+ab+bc+abc-one-a-c-ac)
main.c <- (1/(4*n))*(c+ac+bc+abc-one-a-b-ab)
# table of the main effects
knitr::kable(cbind(main.a, main.b, main.c))
k <- 3; n <- 4
AB <- (1/(4*n))*(abc-bc+ab-b-ac+c-a+one)
ssab <- (abc-bc+ab-b-ac+c-a+one)^2 / 32
ssa <- (a-one+ab-b+ac-c+abc-bc)^2 / 32
ssb <- (b+ab+bc+abc-one-a-c-ac)^2 / 32
ssc <- (c+ac+bc+abc-one-a-b-ab)^2 / 32
ssbc <- (one+a-b-ab-c-ac+bc+abc)^2 / 32
ssac <- (one-a+b-ab-c+ac-bc+abc)^2 / 32
ssabc <- (abc-bc-ac+c-ab+b+a-one)^2 / 32
sse <- 466150-ssa-ssb-ssc-ssab-ssac-ssbc-ssabc
dfe <- 2^k*(n-1); dft <- 2^k*n - 1
mse <- sse/dfe; Fa <- ssa/mse; Fb <- ssb/mse; Fc <- ssc/mse
Fab <- ssab/mse; Fac <- ssac/mse; Fbc <- ssbc/mse; Fabc <- ssabc/mse
y <- c(50,54,44,42,46,48,42,43,49,46,48,45,47,48,56,54)
A <- as.factor(rep(c(rep(3,2), rep(1,2)), 4))
B <- as.factor(rep(c(rep("BW", 4), rep("Color",4)), 2))
C <- as.factor(rep(c(rep(19.95,8), rep(24.95,8)), 1))
H <- data.frame(y, A, B, C)
fit.H <- aov(y ~ A*B*C, data = H)
par(mfrow=c(1,2))
plot(fit.H, which = c(1,2))
par(mfrow=c(1,2))
plot(fit.H$residuals~as.numeric(H$A), xlab = "Class of Mail",
     ylab = "Residuals"); abline(h=0, lty = 3, col = "red")
plot(fit.H$residuals~as.numeric(H$B), xlab = "Print Color",
     ylab = "Residuals"); abline(h=0, lty = 3, col = "red")
plot(fit.H$residuals~as.numeric(H$C), xlab = "Print Price",
     ylab = "Residuals"); abline(h=0, lty = 3, col = "red")
```