STP 531

Applied Analysis of Variance

Homework 7

Nathan A. Nguyen

26 April 2021

Question 25.5

Part A

Suppose now that the six machine were randomly selected from a population of machines eligible for the study and consider model (25.1).

$\mu.$ denotes the mean carton weight for all 20 cartons for all machines in the study.

σ_μ^2 denotes the variability between of the mean carton fill between the 6 machines.

σ^2 denotes the overall mean carton fill variance on the whole

$\sigma^2\{Y_{ij}\}$ denotes the variability between all observations among the 6 machines, regardless of which machine was used.

Question 25.7

Part A

Table 1: ANOVA Table

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
Brand	5	854.5292	170.9058333	238.7112	0
Residuals	42	30.0700	0.7159524	NA	NA

$$\begin{aligned}
H_0 &: \sigma_\mu^2 = 0 \\
H_a &: \sigma_\mu^2 > 0 \\
F^* &= \frac{MSTR}{MSE} = \frac{170.906}{0.716} \approx 283.71 \\
F_{critical} &= F(0.99; 5, 42) \approx 3.49 \\
p &\approx 0^+ \\
&\text{since } F^* > F_{critical}, \text{ reject } H_0
\end{aligned}$$

Since $F^* > F_{critical}$ and $p = 1.085 * 10^{-29} << 0.01 = \alpha$, there is sufficient evidence to reject the null hypothesis. The mean sodium content is not the same across all brands sold.

Part B

$$\begin{aligned}
&\bar{Y}_{..} \pm t(0.995, 5) s\{\bar{Y}_{..}\} \\
&\hat{\mu}_{.} = \bar{Y}_{..} \approx 17.63 \\
s^2\{\bar{Y}..\} &= \frac{MSTR}{rn} = \frac{170.906}{6(8)} = 3.56 \\
s\{\bar{Y}..\} &= 1.89
\end{aligned}$$

Table 2: 99% CI for mean sodium content for all brands

	value
Lower Limit	10.02076
Upper Limit	25.23758

Question 25.8

$$\sigma^2 \approx E\{MSE\} \to 0.716$$
$$s_\mu^2 = \frac{MSTR - MSE}{n} = \frac{170.906 - 0.716}{8} \approx 21.27$$

Part B

Part C

$$\frac{r(n-1)MSE}{\chi^2[1-\alpha/2; r(n-1)]} \leq \sigma^2 \leq \frac{r(n-1)MSE}{\chi^2[\alpha/2; r(n-1)]}$$
$$\frac{42(0.716)}{69.336} \leq \sigma^2 \leq \frac{42(0.716)}{22.138}$$

Table 3: 99% CI

	value
Lower Limit	0.4337141
Upper Limit	1.3583599

Part D

$$H_0 : \sigma_\mu^2 \leq 2\sigma^2$$
$$H_a : \sigma_\mu^2 > 2\sigma^2$$
$$F^* = \frac{MSTR}{2n+1} \div MSE$$
$$F^* = \frac{170.906}{2(8)+1} \div 0.716 \approx 14.04$$
$$F_{critical} = F(0.99; 5; 42) \approx 3.48$$
$$\text{Since } F^* > F_{critical}, \text{ reject } H_0$$

Since $F^* \approx 14 > 3.48 \approx F_{critical}$ there is sufficient evidence to reject the null hypothesis in favour of the alternative. There is reason to suspect that the mean sodium content between brands is greater than twice as within brands.

Question 25.11

For (25.42) Factor A is fixed and Factor B is random $\sum_i (\alpha\beta)_{ij} = 0 \ \forall j$ and $\sum_j (\alpha\beta)_{ij} \neq 0 \ \forall j$ because when summing over i, we are summing over the fixed factor levels for factor A. $\sum_i (\alpha)_i = 0$ it self, so summing over interactions with i index will result in the sum-zero restriction. The interactions when summed over i will be independent.

On the other hand, when we sum the interactions over j, we are not restricted to sum-zero because of the correlated nature of the observations within the same factor j-th factor level of B. The sum of the interactions of j will not be independent when they come from the same level of random factor B.

Question 25.17

Part A

$$
\begin{aligned}
H_0 &: \sigma^2_{\alpha\beta} = 0 \\
H_a &: \sigma^2_{\alpha\beta} > 0 \\
F^* &= \frac{MSAB}{MSE} = \frac{0.309}{4.823} \approx 0.06 \\
F_{critical} &= F(0.95; 6, 36) \approx 2.36 \\
\text{Since } F^* &< F_{critical} \text{ fail to reject } H_0 \\
p &\approx 0.999
\end{aligned}
$$

Table 4: ANOVA Table

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
FactorA	2	150.387917	75.1939583	15.5909719	0.0000133
FactorB	3	152.851667	50.9505556	10.5642621	0.0000398
FactorA:FactorB	6	1.852083	0.3086806	0.0640029	0.9988283
Residuals	36	173.625000	4.8229167	NA	NA

Since $F^* < F_{critical}$ and $p > 0.05$ we fail the reject the null hypothesis. There is sufficient evidence to suggest that there are no meaningful interactions present.

Part B

Factor A

$$
\begin{aligned}
H_0 &: \sigma^2_{\alpha} = 0 \\
H_a &: \sigma^2_{\alpha} \neq 0 \\
F^* &= \frac{MSA}{MSAB} = \frac{75.194}{0.309} \approx 243.35 \\
F_{critical} &= F(0.95; 2, 6) \approx 5.14 \\
p &= 0^+ \\
\text{Since } F^* &> F_{critical}, \text{ reject } H_0
\end{aligned}
$$

Factor B

$$
\begin{aligned}
H_0 &: \sigma^2_{\beta} = 0 \\
H_a &: \sigma^2_{\beta} \neq 0 \\
F^* &= \frac{MSB}{MSE} = \frac{50.951}{4.823} \approx 10.56 \\
F_{critical} &= F(0.95; 3, 36) \approx 2.87 \\
p &= 0^+ \\
\text{Since } F^* &> F_{critical}, \text{ reject } H_0
\end{aligned}
$$

When testing for main effects for both Factor A and Factor B, we rejected the null hypothesis for both cases. In both cases $F^* > F_{critical}$ and $p < 0.05 = \alpha$ respectively. There is sufficient evidence to conclude that there a main effects present. Operator 5 has an interesting looking curve.

Part D

$$
\begin{aligned}
\hat{\mu_{2.}} &= \bar{Y}_{2..} \approx 76.8 \\
s^2\{\hat{\mu_{2.}}\} &= c1MSAB + c2MSB \approx 1.704 \\
s\{\hat{\mu_{2.}}\} &\approx 1.037 \\
df_{sat} &= \frac{\left(\frac{a-1}{nab}MSAB + \frac{1}{nab}MSB\right)^2}{\frac{\left(\frac{a-1}{nab}MSAB\right)^2}{(a-1)(b-1)} + \frac{\left(\frac{1}{nab}MSB\right)^2}{b-1}} \quad (25.65) \\
df_{sat} &\approx 3.702 \\
t &= t(0.975; 3) \approx 3.182 \\
\hat{\mu_{2.}} &\pm t(0.975; 3)s\{\hat{\mu_{2.}}\}
\end{aligned}
$$

Table 5: 95% CI Satterthwaite Approx.

	value
Lower Limit	73.50136
Upper Limit	80.09864

With 95% confidence, we conclude that the mean market value for 8 coats on the pearls fall somewhere between 73.50 and 80.10 units (73.50136 $\leq \mu_2 \leq$ 80.09864.

Observations	48
Dependent variable	Pearls
Type	Mixed effects linear regression

AIC	223.66
BIC	234.88
Pseudo-R^2 (fixed effects)	0.32
Pseudo-R^2 (total)	0.60

Fixed Effects

	Est.	S.E.	t val.	d.f.	p
(Intercept)	73.11	0.98	74.53	5.82	0.00
FactorA2	3.69	0.71	5.22	44.00	0.00
FactorA3	3.82	0.71	5.41	44.00	0.00

p values calculated using Satterthwaite d.f.

Random Effects

Group	Parameter	Std. Dev.
FactorA:FactorB	(Intercept)	0.00
FactorB	(Intercept)	1.69
Residual		2.00

Grouping Variables		
Group	# groups	ICC
FactorA:FactorB	12	0.00
FactorB	4	0.42

Question 25.17 Matrix

$$Y_{ijk} = \mu_{..} + \alpha_i + \beta_j + (\alpha\beta)_{ij} + \varepsilon_{ijk} \quad (25.42)$$

$a = 3,\ b = 4,\ n = 4 \Rightarrow (3)(4)(4) = 48 = n_T$

$$\underset{(N\times 1)}{\underset{\sim}{Y}} = \underset{(N\times p)}{X}\ \underset{(p\times 1)}{\underset{\sim}{\beta}} + \underset{(N\times q)}{Z}\ \underset{(q\times 1)}{\underset{\sim}{\gamma}} + \underset{(N\times 1)}{\underset{\sim}{\varepsilon}}$$

$\underset{(48\times 1)}{\underset{\sim}{Y}} \equiv$ Vector of ith subject.

$X \equiv$ fixed effects design matrix

$\underset{\sim}{\beta} =$ fixed effects vector

$Z =$ Random effects design matrix

$\underset{\sim}{\gamma} =$ random effects vector

$\underset{\sim}{\varepsilon} =$ error vector.

$$\sigma^2\{\underset{\sim}{Y}\} = \begin{pmatrix} \underset{(12\times 12)}{\Sigma} & \underset{(12\times 12)}{0} \\ \underset{(12\times 12)}{0} & \underset{(12\times 12)}{\Sigma} \end{pmatrix}$$

25.1 (Extra credit)

The μ_i denote the variability for the different i-levels of the random factor variable. On the other hand the ϵ_{ij} denotes the variation with regards to the different potential values for the i-th level of the random factor any different j-th observation. There are two difference variations in the model, variability by means at the i-th level of random factor and the variability.

The distribution for each i-th level may be different, but the variance should be the same.

Appendix

Question 25.5

Part A

Question 25.7

Part A

```
data <- read.table("https://people.stat.sc.edu/hitchcock/sodiumcontent.txt", header = FALSE, col.names =

data$Brand <- as.factor(data$Brand)

model.1 <- lm(Sodium ~ Brand, data = data)
blah <- anova(model.1)
kable(blah[], caption = "ANOVA Table", format = "markdown") %>%
  kable_styling(position = "center")

F.star <- 170.906/0.716
F.crit <- qf(0.99,5,42)
p <- pf(F.star,5,42,lower.tail = FALSE)
```

Part B

```
Ybar.. <- mean(data$Sodium)
t.value <- qt(0.995,5)
Ybar..var <- 170.906/48
Ybar..STD <- sqrt(Ybar..var)
Ybar..LL <- Ybar.. - t.value*Ybar..STD
Ybar..UL <- Ybar.. + t.value*Ybar..STD

blah <- data.frame(
  value = c(Ybar..LL, Ybar..UL)
)
row.names(blah) <- c("Lower Limit","Upper Limit")
kable(blah[], caption = "99% CI for mean sodium content for all brands", format = "markdown") %>%
  kable_styling(position = "center")
```

Question 25.8

Part B

```
s2 <- (170.906-0.716)/8
```

Part C

```
MSE <- 0.716
value <- 6*(8-1)
chi.lower <- qchisq(1-(0.01/2),42)
chi.upper <- qchisq((0.01/2),42)
LL <- ((42*MSE))/chi.lower
UP <- ((42*MSE))/chi.upper
blah <- data.frame(
  value = c(LL, UP)
)
row.names(blah) <- c("Lower Limit", "Upper Limit")
kable(blah[], caption = "99% CI", format = "markdown") %>% kable_styling(position = "center")
```

Part D

```
F.star <- 170.906 / (2*8+1) * 1/MSE
F.crit <- qf(0.99,5,42)
```

Question 25.11

Question 25.17

Part A

```
data <- read.table("https://people.stat.sc.edu/hitchcock/imitationpearls.txt",
                   header = FALSE, col.names = c("Pearls", "FactorA", "FactorB", "Index"))
data$FactorA <- as.factor(data$FactorA)
data$FactorB <- as.factor(data$FactorB)

model.2 <- lm(Pearls ~ FactorA + FactorB + FactorA:FactorB, data = data)

blah <- anova(model.2)
kable(blah[], caption = "ANOVA Table", format = "markdown") %>%
  kable_styling(position = "center")

F.star <- 0.309/4.823
F.crit <- qf(0.95,6,36)
p <- pf(F.star,6,36,lower.tail=F)
```

Part B

```
# Factor A
F.star <- 75.194/0.309
F.crit <- qf(0.95,2,6)
p <- pf(F.star,2,6,lower.tail = FALSE)

# Factor B
F.star <- 50.951/4.823
F.crit <- qf(0.95,3,36)
p <- pf(F.star,3,36,lower.tail = FALSE)
```

Part D

```
means <- data %>%
  group_by(FactorA) %>%
  summarise(means = mean(Pearls))

mu2. <- 76.8
a <- 3
b <- 4
n <- 48
c1 <- (a-1)/(48)
c2 <- (1)/(48)
MSAB <- 0.309
MSB <- 50.951
temp1 <- c1*MSAB
temp2 <- c2*MSB
variance <- temp1 + temp2
std <- sqrt(variance)

nab <- 48
df <- ((temp1 + temp2)^2) / ((temp1^2)/(2*3) + (temp2^2)/(3))
t.value <- qt(0.975, df= 3)
```

```r
mu2.LL <- mu2. - t.value*std
mu2.UL <- mu2. + t.value*std

blah <- data.frame(
  value = c(mu2.LL, mu2.UL)
)

row.names(blah) <- c("Lower Limit", "Upper Limit")
kable(blah[], caption = "95% CI Satterthwaite Approx.", format = "markdown") %>%
  kable_styling(position = "center")

library(lme4); library(jtools)
data2$FactorA <- as.factor(data2$FactorA)
model.3 <- lmer(Pearls ~ FactorA + (1|FactorB) + (1|FactorA:FactorB), data = data2, REML = FALSE)
summ(model.3)
```