ST553 HW5

Nick Sun

May 6, 2019

Problem 8.4

We are conducting an experiment on circuit board delamination rates found in pacemakers. We have three factors with three levels each:

1. Firing Profile Time (8 vs 13 hours)
2. Furnace airflow (low vs high)
3. Laser (old vs new)

Each treatment combination has two replicates for a total of 16 experimental units.

(a) Give factor effects parameterization

Our parameterization will be:

$$y_{ijkl} = \mu + \alpha_i + \beta_j + \gamma_k + (\alpha\beta)_{ij} + (\alpha\gamma)_{ik} + (\beta\gamma)_{jk} + (\alpha\beta\gamma)_{ijk} + \epsilon_{ijkl}$$

where:

- μ is the overall mean
- α_i is the effect of the *ith* firing profile time
- β_j is the effect of the *jth* airflow
- γ_k is the effect of the *kth* laser
- $(\alpha\beta)_{ij}$ is the interaction effect of the *ith* firing profile time and *jth* airflow
- $(\alpha\gamma)_{ik}$ is the interaction effect of the *ith* firing profile time and *kth* laser
- $(\beta\gamma)_{jk}$ is the interaction effect of the *jth* airflow and *kth* laser
- $(\alpha\beta\gamma)_{ijk}$ is the three way interaction effect of the *ith* firing profile time, *jth* airflow, and *kth* laser
- y_{ijkl} is the delamination rate of the lth replicate of the treatment combination with the ith firing profile, jth furnace airflow, and kth laser
- ϵ_{ijkl} is the random error associated with the *ijkl*th replicate

(b) List all assumptions of the model including the sum to zero constraints

$$\epsilon_{ijkl} \sim N(0, \sigma^2)$$

with the following coefficient constraints

$$0 = \sum_i \alpha_i = \sum_j \beta_j = \sum_k \gamma_k =$$
$$\sum_i (\alpha\beta)_{ij} = \sum_j (\alpha\beta)_{ij} = \sum_i (\alpha\gamma)_{ik} = \sum_k (\alpha\gamma)_{ik} = \sum_j (\beta\gamma)_{kj} = \sum_k (\beta\gamma)_{kj} =$$
$$\sum_i (\alpha\beta\gamma)_{ijk} = \sum_j (\alpha\beta\gamma)_{ijk} = \sum_k (\alpha\beta\gamma)_{ijk}$$

(c) Write down the parameter vector β and the unique rows of the design matrix $\mathbf{X}$

Our parameter vector is

$$\begin{pmatrix} \mu \\ \alpha_1 \\ \beta_1 \\ \gamma_1 \\ (\alpha\beta)_{11} \\ (\alpha\gamma)_{11} \\ (\beta\gamma)_{11} \\ (\alpha\beta\gamma)_{111} \end{pmatrix}$$

and our entire model equation (only unique rows displayed) with design matrix $\mathbf{X}$ is

$$\begin{pmatrix} y_{111} \\ y_{211} \\ y_{121} \\ y_{112} \\ y_{221} \\ y_{212} \\ y_{122} \\ y_{222} \end{pmatrix} = \begin{pmatrix} 1 & 1 & 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & -1 & 1 & 1 & -1 & -1 & 1 & -1 \\ 1 & 1 & -1 & 1 & -1 & 1 & -1 & -1 \\ 1 & 1 & 1 & -1 & 1 & -1 & -1 & -1 \\ 1 & -1 & -1 & 1 & 1 & -1 & -1 & 1 \\ 1 & -1 & 1 & -1 & -1 & 1 & -1 & 1 \\ 1 & 1 & -1 & -1 & -1 & -1 & 1 & 1 \\ 1 & -1 & -1 & -1 & 1 & 1 & 1 & -1 \end{pmatrix} \begin{pmatrix} \mu \\ \alpha_1 \\ \beta_1 \\ \gamma_1 \\ (\alpha\beta)_{11} \\ (\alpha\gamma)_{11} \\ (\beta\gamma)_{11} \\ (\alpha\beta\gamma)_{111} \end{pmatrix} + \text{vector of random errors} \epsilon_{ijkl}$$

and since there are two replicates, each unique row appears twice.

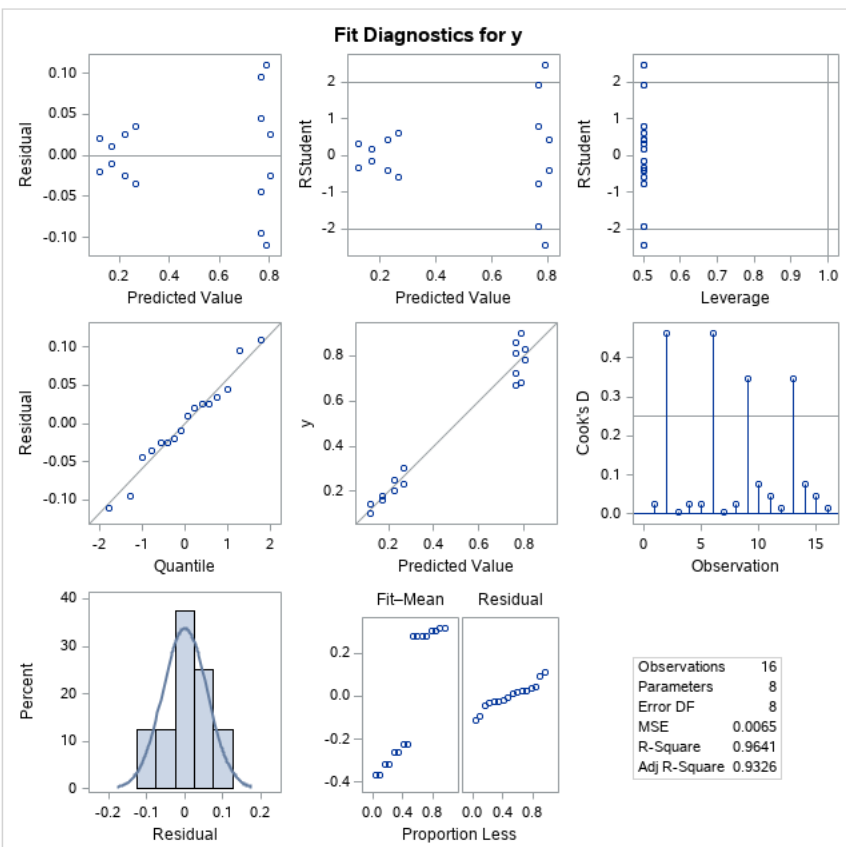
(d) Fit the model in SAS

Source	DF	Sum of Squares	Mean Square	F Value	Pr > F
Model	7	1.40119375	0.20017054	30.65	<.0001
Error	8	0.05225000	0.00653125		
Corrected Total	15	1.45344375			

Source	DF	Type I SS	Mean Square	F Value	Pr > F
flow	1	0.00275625	0.00275625	0.42	0.5341
temp	1	1.37475625	1.37475625	210.49	<.0001
flow*temp	1	0.00140625	0.00140625	0.22	0.6550
laser	1	0.00140625	0.00140625	0.22	0.6550
flow*laser	1	0.00855625	0.00855625	1.31	0.2855
temp*laser	1	0.00075625	0.00075625	0.12	0.7424
flow*temp*laser	1	0.01155625	0.01155625	1.77	0.2201

Source	DF	Type III SS	Mean Square	F Value	Pr > F
flow	1	0.00275625	0.00275625	0.42	0.5341
temp	1	1.37475625	1.37475625	210.49	<.0001
flow*temp	1	0.00140625	0.00140625	0.22	0.6550
laser	1	0.00140625	0.00140625	0.22	0.6550
flow*laser	1	0.00855625	0.00855625	1.31	0.2855
temp*laser	1	0.00075625	0.00075625	0.12	0.7424
flow*temp*laser	1	0.01155625	0.01155625	1.77	0.2201

(e) Produce a panel of diagnostic plots



The one glaring issue is heteroskedasticity. Different treatment groups appear to have different variances which could potentially be problematic when we are trying to estimate σ^2. The QQ, histogram, and R-F plots looks great though - pretty normal looking and a relatively low spread of residuals relative to the fits. In all, I would say that the assumptions for this model are mostly met - in practice, it's probably fine.

(f) Find point estimates for all of your parameters, including σ^2

We first need to compute the sample means for all of the possible cell combinations.

We can calculate this in SAS using `lsmeans` option in `PROC GLM`. This outputs a bunch of tables containing the sample means of different treatment combinations.

flow	y LSMEAN
1	0.50125000
2	0.47500000

temp	y LSMEAN
1	0.78125000
2	0.19500000

flow	temp	y LSMEAN
1	1	0.78500000
1	2	0.21750000
2	1	0.77750000
2	2	0.17250000

laser	y LSMEAN
1	0.49750000
2	0.47875000

flow	laser	y LSMEAN
1	1	0.48750000
1	2	0.51500000
2	1	0.50750000
2	2	0.44250000

flow	temp	laser	y LSMEAN
1	1	1	0.80500000
1	1	2	0.76500000
1	2	1	0.17000000
1	2	2	0.26500000
2	1	1	0.79000000
2	1	2	0.76500000
2	2	1	0.22500000
2	2	2	0.12000000

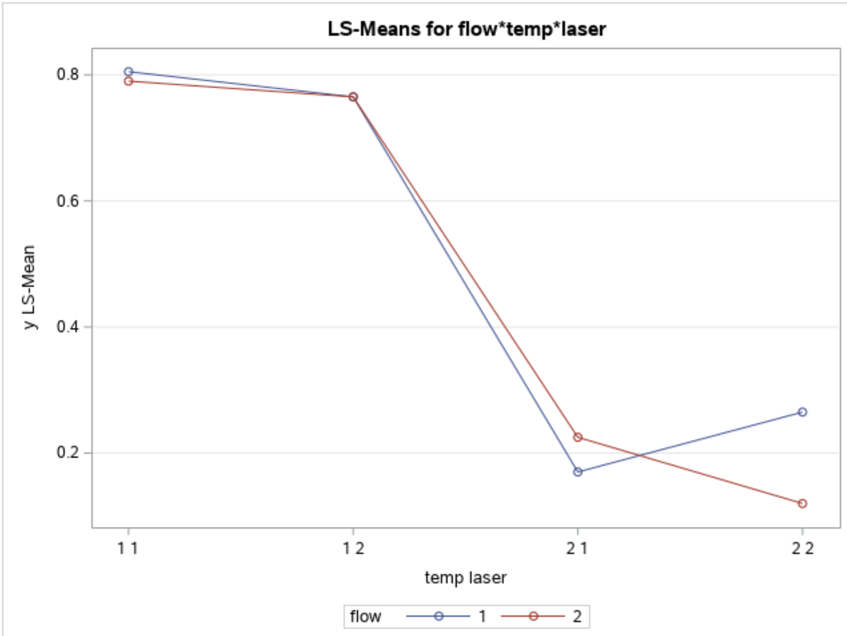
With all the estimated $\bar{y}$s, we can estimate the model parameters:

$$\hat{\mu} = .488$$
$$\hat{\alpha}_1 = .78 - .488 = .293$$
$$\hat{\beta}_1 = .5 - .488 = .013$$
$$\hat{\gamma}_1 = .497 - .488 = .009$$
$$(\hat{\alpha\beta})_{11} = .785 - .781 - .5 + .488 = -.009$$
$$(\hat{\alpha\gamma})_{11} = .797 - .781 - .497 + .488 = .006$$
$$(\hat{\beta\gamma})_{11} = .487 - .5 - .497 + .488 = -.023$$
$$(\alpha\hat{\beta}\gamma)_{11} = .805 - .785 - .797 - .487 + .781 + .5 + .497 - .488 = .0268$$
$$\hat{\sigma}^2 = .006$$

(g) State the full and reduced models for your seven type III F-tests

	Full	Reduced
flow	$y_{ijkl} = \mu + \alpha_i + \beta_j + \gamma_k + (\alpha\beta)_{ij} + (\alpha\gamma)_{ik} + (\beta\gamma)_{jk} + (\alpha\beta\gamma)_{ijk} + \epsilon_{ijkl}$	$y_{ijkl} = \mu + \alpha_i + \gamma_k + (\alpha\beta)_{ij} + (\alpha\gamma)_{ik} + (\beta\gamma)_{jk} + (\alpha\beta\gamma)_{ijk} + \epsilon_{ijkl}$
temp	$y_{ijkl} = \mu + \alpha_i + \beta_j + \gamma_k + (\alpha\beta)_{ij} + (\alpha\gamma)_{ik} + (\beta\gamma)_{jk} + (\alpha\beta\gamma)_{ijk} + \epsilon_{ijkl}$	$y_{ijkl} = \mu + \beta_j + \gamma_k + (\alpha\beta)_{ij} + (\alpha\gamma)_{ik} + (\beta\gamma)_{jk} + (\alpha\beta\gamma)_{ijk} + \epsilon_{ijkl}$
flow* temp	$y_{ijkl} = \mu + \alpha_i + \beta_j + \gamma_k + (\alpha\beta)_{ij} + (\alpha\gamma)_{ik} + (\beta\gamma)_{jk} + (\alpha\beta\gamma)_{ijk} + \epsilon_{ijkl}$	$y_{ijkl} = \mu + \alpha_i + \beta_j + \gamma_k + (\alpha\gamma)_{ik} + (\beta\gamma)_{jk} + (\alpha\beta\gamma)_{ijk} + \epsilon_{ijkl}$
laser	$y_{ijkl} = \mu + \alpha_i + \beta_j + \gamma_k + (\alpha\beta)_{ij} + (\alpha\gamma)_{ik} + (\beta\gamma)_{jk} + (\alpha\beta\gamma)_{ijk} + \epsilon_{ijkl}$	$y_{ijkl} = \mu + \alpha_i + \beta_j + (\alpha\beta)_{ij} + (\alpha\gamma)_{ik} + (\beta\gamma)_{jk} + (\alpha\beta\gamma)_{ijk} + \epsilon_{ijkl}$
flow* laser	$y_{ijkl} = \mu + \alpha_i + \beta_j + \gamma_k + (\alpha\beta)_{ij} + (\alpha\gamma)_{ik} + (\beta\gamma)_{jk} + (\alpha\beta\gamma)_{ijk} + \epsilon_{ijkl}$	$y_{ijkl} = \mu + \alpha_i + \beta_j + \gamma_k + (\alpha\beta)_{ij} + (\alpha\gamma)_{ik} + (\alpha\beta\gamma)_{ijk} + \epsilon_{ijkl}$
temp* laser	$y_{ijkl} = \mu + \alpha_i + \beta_j + \gamma_k + (\alpha\beta)_{ij} + (\alpha\gamma)_{ik} + (\beta\gamma)_{jk} + (\alpha\beta\gamma)_{ijk} + \epsilon_{ijkl}$	$y_{ijkl} = \mu + \alpha_i + \beta_j + \gamma_k + (\alpha\beta)_{ij} + (\beta\gamma)_{jk} + (\alpha\beta\gamma)_{ijk} + \epsilon_{ijkl}$
flow* temp* laser	$y_{ijkl} = \mu + \alpha_i + \beta_j + \gamma_k + (\alpha\beta)_{ij} + (\alpha\gamma)_{ik} + (\beta\gamma)_{jk} + (\alpha\beta\gamma)_{ijk} + \epsilon_{ijkl}$	$y_{ijkl} = \mu+\alpha_i+\beta_j+\gamma_k+(\alpha\beta)_{ij}+(\alpha\gamma)_{ik}+(\beta\gamma)_{jk}+\epsilon_{ijkl}$

(h) Produce an interaction plot for the three way interaction



(i) State the null and alternative hypotheses used to test for the three way interaction

$H_0 : (\alpha\beta\gamma)_{ijk} = 0$ for $i, j, k \in \{1, 2\}$ $H_A :$ at least one $(\alpha\beta\gamma)_{ijk} \neq 0$ for $i, j, k \in \{1, 2\}$

The p-value from this hypothesis test is .22 so we fail to reject our null hypothesis at the $\alpha = .05$ level. This tells us that the delamination rate does not depend on the three way interaction between flow, temperature, and laser type.

(j) Estimate the difference in mean delamination fraction for the two firing profile times

Here we want to estimate $\mu_{1..} - \mu_{2..}$ which we can do using the sample means $\bar{y}_{1...} - \bar{2}... = .781 - .195 = .586$

The standard error for this difference in estimates is found using the formula:

$$\sqrt{\hat{\sigma}^2\left(\frac{1}{n} + \frac{1}{n}\right)} = \sqrt{.0065 * 2\left(\frac{1}{8}\right)} = .0404$$

so our resulting confidence interval is (.493, .679).

We are 95% confident that the true difference between the mean delamination proportion in the two firing profiles is between .493 and .679

The appropriate SAS output found using the `estimate` and `clparm` in `PROC GLM` is provided below:

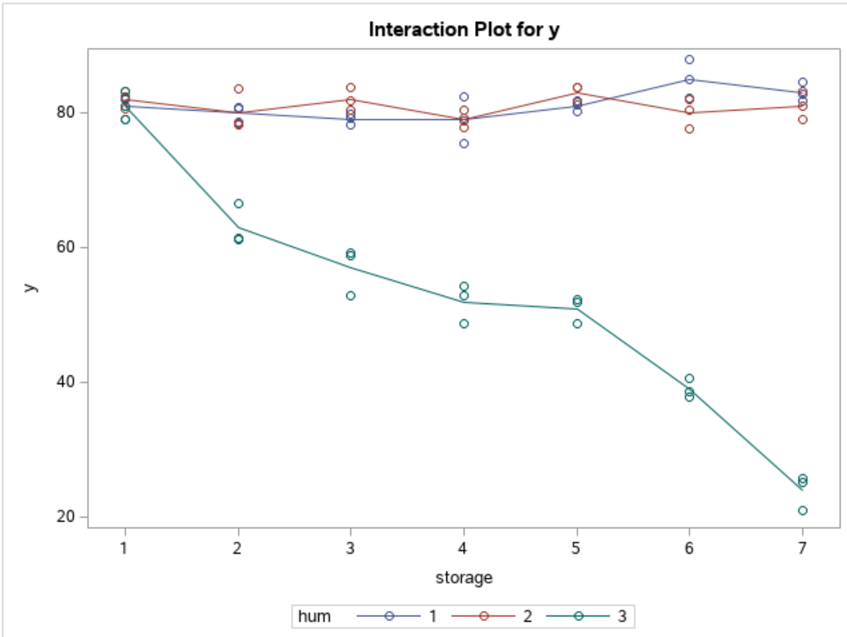
Parameter	Estimate	Standard Error	t Value	Pr > |t|	95% Confidence Limits	
8 vs 13 hr	0.58625000	0.04040807	14.51	<.0001	0.49306881	0.67943119

(k) Does furnace airflow affect mean delamination fraction?

$H_0 : \beta_j = 0$ for j = 1, 2 $H_A :$ at least one of $\beta_j \neq 0$

We get an F-statistic here of .42 and a resulting large p-value of .534 so we fail to reject the null hypothesis. The airflow of the furnace does not affect mean delamination rate.

Problem 8.7



For 0% and 32% humidity, the effects of remain relatively constant regardless of the length of storage. For 45% humidity, seed viability decreases as the storage length increases suggesting heavily that there is an interaction between this level humidity and storage which is further corroborated by the low p-value.