Lecture notes on Experiment Design & Data Analysis

Design of Engineering Experiments – Experiments with Random Factors

- Text reference, Chapter 13, Pg. 484
- Previous chapters have considered **fixed** factors
  - A specific set of factor levels is chosen for the experiment
  - Inference confined to those levels
  - Often **quantitative** factors are fixed (why?)
- When factor levels are chosen at random from a larger population of potential levels, the factor is **random**
  - Inference is about the entire population of levels
  - Industrial applications include measurement system studies

Random Effects Models

- Example 13-1 (pg. 487) – weaving fabric on looms
- Response variable is strength
- Interest focuses on determining if there is difference in strength due to the different looms
- However, the weave room contains many (100s) looms
- Solution – select a (random) **sample** of the looms, obtain fabric from each
- Consequently, “looms” is a **random factor**
- See data, Table 13-1; looks like standard single-factor experiment with $a = 4$ & $n = 4$

Table 13-1 Strength Data for Example 13-1

Looms	Observations 1	2	3	4	$y_{i.}$
1	98	97	99	96	390
2	91	90	93	92	366
3	96	95	97	95	383
4	95	96	99	98	388
					$1527 = y_{..}$

Table 13.1 (p. 487)

Strength Data for Example 13.1

Random Effects Models

- The usual single factor ANOVA model is

$$y_{ij} = \mu + \tau_i + \varepsilon_{ij} \begin{cases} i = 1, 2, \ldots, a \\ j = 1, 2, \ldots, n \end{cases}$$

- Now both the error term and the treatment effects are random variables:

$$\varepsilon_{ij} \text{ is } NID(0, \sigma^2) \text{ and } \tau_i \text{ is } NID(0, \sigma_\tau^2)$$

- **Variance components**: $V(y_{ij}) = \sigma^2 + \sigma_\tau^2$

Relevant Hypotheses in the Random Effects (or Components of Variance) Model

- In the fixed effects model we test equality of treatment means
- This is no longer appropriate because the treatments are randomly selected
  - the individual ones we happen to have are not of specific interest
  - we are interested in the **population** of treatments
- The appropriate hypotheses are

$$H_0 : \sigma_\tau^2 = 0$$

$$H_1 : \sigma_\tau^2 > 0$$

Testing Hypotheses - Random Effects Model

- The standard ANOVA partition of the total sum of squares still works; leads to usual ANOVA display
- Form of the hypothesis test depends on the **expected mean squares**

$$E(MS_E) = \sigma^2 \text{ and } E(MS_{Treatments}) = \sigma^2 + n\sigma_\tau^2$$

- Therefore, the appropriate test statistic is

$$F_0 = MS_{Treatments} / MS_E$$

Estimating the Variance Components

- Use the **ANOVA** method; equate expected mean squares to their observed values:

$$\sigma^2 = MS_E \text{ and } \sigma^2 + n\sigma_\tau^2 = MS_{Treatment}$$

$$\hat{\sigma}_\tau^2 = \frac{MS_{Treatment} - MS_E}{n}$$

$$\hat{\sigma}^2 = MS_E$$

- Potential problems with these estimators
  - Negative estimates (woops!)
  - They are moment estimators & don't have best statistical properties

Confidence Intervals on the Variance Components

- Easy to find a 100(1-α)% CI on σ^2:

$$\frac{(N-a)MS_E}{\chi^2_{\alpha/2,N-a}} \le \sigma^2 \le \frac{(N-a)MS_E}{\chi^2_{1-(\alpha/2),N-a}}$$

- a 100(1-α)% CI on intraclass correlation coefficient $\frac{\sigma_\tau^2}{\sigma_\tau^2+\sigma^2}$

$$\frac{L}{1+L} \le \frac{\sigma_\tau^2}{\sigma_\tau^2+\sigma^2} \le \frac{U}{U+1}$$

$$L=\frac{1}{n}\left(\frac{MS_{treatment}}{MS_E}\frac{1}{F_{\alpha/2,a-1,N-a}}-1\right), U=\frac{1}{n}\left(\frac{MS_{treatment}}{MS_E}\frac{1}{F_{1-\alpha/2,a-1,N-a}}-1\right)$$

R-code and Output

```
###one-way random effects model####
fiber=read.table("textile.txt",header=T)
attach(fiber)
Loom=factor(Loom)
aov1=aov(Strength~Loom)
(sum1=summary(aov1)[[1]])
```

	Df	Sum Sq	Mean Sq	F value	Pr(>F)	
Loom	3	89.188	29.729	15.681	0.0001878	***
Residuals	12	22.750	1.896			

 # The F test of no treatment effects is valid.

```
# The variance estimate for $\sigma^2$ is valid
sigma2.E=sum1[2,3]
> sigma2.E
[1] 1.895833

# Estimate the variance of the $\tau_i$, i.e. $sigma_tau^2:
n=sum1[2,1]/(sum1[1,1]+1)+1
sigma2.A=(sum1[1,3]-sigma2.E)/n
> sigma2.A
[1] 6.958333

# # Calculate the intra-class correlation:
# $\rho=\sigma^2_{\tau}/(\sigma^2_{\tau}+\sigma^2)$
 rho=sigma2.A/(sigma2.A+sigma2.E)
> rho
[1] 0.7858824
```

```
# An 95% confidence interval estimation for $\rho$
alpha=0.05
ci1=sum1[1,4]/qf(c(1-alpha/2,alpha/2),n-1,sum1[2,1])-1
ci1=ci1/(ci1+n)
>ci1
[1] 0.3850736 0.9824420
```

```
# An 95% interval estimation for $\sigma^2$
Alpha=0.05
ci2=sum1[2,2]/qchisq(c(1-alpha/2,alpha/2),sum1[2,1])

> ci2
[1] 0.9748608 5.1660065
```

Lecture notes on Experiment Design & Data Analysis

Model Assumption Checking