Handout 7

Random effects model

We will consider a one-factor model where the factor effects are random. The following example is useful.

Coil winding machines: A plant contains a large number of coil winding machines. A production analyst studied a certain characteristics of the wound coils produced by these machines by selecting four machines at random and then choosing 10 coils at random from the day's output of each selected machine. The results follow

	j									
i	1	2	3	4	5	6	7	8	9	10
1	205	204	207	203	108	206	109	205	207	206
2	201	204	198	202	209	207	199	206	205	204
3	198	204	196	201	199	203	202	198	202	197
4	210	209	214	215	211	208	210	209	211	210

The model is

$$Y_{ij} = \mu_i + \varepsilon_{ij}, j = 1, ..., n, i = 1, ..., r,$$

where ε_{ij}'s are independent $N(0, \sigma^2)$. Here μ_i's are independent $N(\mu., \sigma_\mu^2)$, and $\{\varepsilon_{ij}\}$ and $\{\mu_i\}$ are independent. The total number of observations is $n_T = nr$. Note that

$$E(Y_{ij}) = \mu., \; Var(Y_{ij}) = Var(\mu_i) + Var(\varepsilon_{ij}) = \sigma_\mu^2 + \sigma^2,$$
$$Cov(Y_{ij}, Y_{ij'}) = \sigma_\mu^2, j \neq j', \quad Cov(Y_{ij}, Y_{i'j'}) = 0, \, i \neq i'.$$

Note that the variability of Y's has two components: σ_μ^2, variability due to machines, and σ^2, variability of ε_{ij}. It is customary to call σ_μ^2 and σ^2 them variance components (of Y_{ij}). Also note that for any i, Y_{ij}'s are not independent. This is clearly different from the one factor model where the factor is not random (chapter 16). The quantity $\sigma_\mu^2/(\sigma_\mu^2 + \sigma^2)$ is the proportion of variability in the coil characteristics that is due to the variability in machines. This ratio is called the intraclass correlation since $Corr(Y_{i,j}, Y_{ij'}) = \sigma_\mu^2/(\sigma_\mu^2 + \sigma^2)$, $j \neq j'$.

We can rewrite the random effects model as a random factor effects model as given below

$$Y_{ij} = \mu. + \tau_i + \varepsilon_{ij},$$

where $\mu.$ is the overall mean, ε_{ij}'s are independent $N(0, \sigma^2)$, τ_i's are independent $N(0, \sigma_\mu^2)$, and $\{\varepsilon_{ij}\}$ and $\{\tau_i\}$ are independent.

Notations and sums of squares etc. are the same as in chapter 16. However some of the results are different. So we have

$$\begin{aligned} SSTO &= \sum\sum(Y_{ij} - \bar{Y}_{..})^2,\ df = nr-1, \\ SSTR &= n\sum(\bar{Y}_{i.} - \bar{Y}_{..})^2,\ df = r-1, MSTR = \frac{SSTR}{r-1}, \\ SSE &= \sum\sum(Y_{ij} - \bar{Y}_{i.})^2, df = n_T - r = (n-1)r, MSE = \frac{SSE}{(n-1)r}. \end{aligned}$$

Here is an important fact

Fact 1 (i) $E(\bar{Y}_{..}) = \mu_{.}$, (ii) $Var(\bar{Y}_{..}) = \frac{n\sigma_\mu^2+\sigma^2}{nr}$, (iii) $E(MSE) = \sigma^2$, (iv) $E(MSTR) = n\sigma_\mu^2 + \sigma^2$.

Remark: This above fact tells us a few things.

(a) $\bar{Y}_{..}$ estimates $\mu_{.}$, (b) MSE estimates σ^2, (c) $Var(\bar{Y}_{..})$ is estimated by $s^2(\bar{Y}_{..}) = \frac{MSTR}{nr}$,

(d) Estimate of σ_μ^2 is given by $s_\mu^2 = \frac{MSTR-MSE}{n}$. Note that if $MSTR$ is smaller than MSE, then s_μ^2 is taken to be equal to zero.

ANOVA Table for "Coil winding machines" data

Source	df	SS	MS	F	p-value
Machine	$r-1=3$	602.50	200.83	28.09	0.0000
Error	$(n-1)r=36$	257.40	7.15		
Total	$nr-1=39$	859.90			

Hypothesis testing.

If we want to test that there is no machine effect, i.e., test $H_0 : \sigma_\mu^2 = 0$ against $H_1 : \sigma_\mu^2 \neq 0$, then the F-statistic is $F^* = \frac{MSTR}{MSE}$. The degrees of freedom associated with this test are $(r-1,(n-1)r)$. So we reject H_0 if $F^* > F(1\text{-}\alpha; r-1,(n-1)r)$, where α is the given level of significance.

For the coil winding data: we have:

$$\begin{aligned} r = 4, n = 10, n_T = nr = 40, \\ \bar{Y}_{1.} = 205.9, \bar{Y}_{2.} = 203.6, \bar{Y}_{3.} = 200.0, \bar{Y}_{4.} = 210.7, \bar{Y}_{..} = 205.05, \\ SSTO = 859.90, SSTR = 602.50, SSE = 257.40, \\ MSTR = 200.83, MSE = 7.15 \end{aligned}$$

So if we want to test $H_0 : \sigma_\mu^2 = 0$ against $H_1 : \sigma_\mu^2 \neq 0$, then the F-statistic is $F^* = \frac{MSTR}{MSE} = 28.09$. The degrees of freedom for the F-test are $(3,36)$. The p-value is close to zero.

Clearly we cannot reject H_0. So we conclude that the machine effect exist.

Estimation of $\mu_{.}$

Estimate of $\mu.$ is $\bar{Y}..$. Estimate of $Var(\bar{Y}..)$ is given by $s^2(\bar{Y}..) = \frac{MSTR}{nr}$. So a $(1-\alpha)100\%$ confidence interval for $\mu.$ is $\bar{Y}.. \pm t(1-\alpha/2; r-1)s(\bar{Y}..)$.

For the coil winding data,

$$\bar{Y}.. = 205.05, s(\bar{Y}..) = \sqrt{\frac{MSTR}{nr}} = \sqrt{\frac{200.83}{40}} = 2.2407.$$

So a 95% confidence interval for $\mu.$ is given by

$$\bar{Y}.. \pm t(.975; 3)s(\bar{Y}..), \text{ i.e., } 205.05 \pm (3.1824)(2.2407), \text{ i.e., } 205.05 \pm 7.13, \text{ i.e., } (192.82, 207.18).$$

Estimation of $\sigma_\mu^2/(\sigma_\mu^2+\sigma^2)$.

Estimates of σ_μ^2 is $s_\mu^2 = \frac{MSTR-MSE}{n} = \frac{200.83-7.15}{10} = 19.368$ and an estimate of σ^2 is $MSE = 7.15$. So an estimate of $\sigma_\mu^2/(\sigma_\mu^2+\sigma^2)$ is given by

$$\frac{s_\mu^2}{s_\mu^2 + MSE} = \frac{19.368}{19.368 + 7.15} = .7304.$$

This tells us that about 73% of the variability in coil characteristics is due to variability in machines.

We can construct a $(1-\alpha)100\%$ confidence interval for $\sigma_\mu^2/(\sigma_\mu^2+\sigma^2)$ of the form (L^*, U^*), where L^* and U^* are given below. Let

$$L = \frac{1}{n}\left[\frac{F^*}{F(1-\alpha/2; r-1, (n-1)r)} - 1\right], \quad U = \frac{1}{n}\left[\frac{F^*}{F(\alpha/2; r-1, (n-1)r)} - 1\right],$$
$$L^* = \frac{L}{L+1}, U^* = \frac{U}{U+1},$$

where $F^* = \frac{MSTR}{MSE}$. For our data, $F^* = 28.0881$. We want to construct a 95% confidence interval for $\sigma_\mu^2/(\sigma_\mu^2+\sigma^2)$. Note that

$$F(1-\alpha/2; r-1, (n-1)r) = F(.975; 3, 16) = 3.5047,$$
$$F(\alpha/2; r-1, (n-1)r) = F(.05; 3, 36) = \frac{1}{F(.975; 36, 3)} = \frac{1}{14.2508} = .0712.$$

So we have

$$L = \frac{1}{n}\left[\frac{F^*}{F(1-\alpha/2; r-1, (n-1)r)} - 1\right] = \frac{1}{10}\left[\frac{28.0881}{3.5047} - 1\right] = .70144,$$
$$U = \frac{1}{n}\left[\frac{F^*}{F(\alpha/2; r-1, (n-1)r)} - 1\right] = \frac{1}{10}\left[\frac{28.0881}{.0712} - 1\right] = 39.3496,$$
$$L^* = \frac{L}{L+1} = \frac{.70144}{.70144+1} = .4123, U^* = \frac{U}{U+1} = \frac{39.3496}{39.3496+1} = .9752.$$

So a 95% confidence interval for $\sigma_\mu^2/(\sigma_\mu^2+\sigma^2)$ is given by $(.412, .975)$.

Confidence interval for σ_μ^2/σ^2

A $(1-\alpha)100\%$ confidence interval for σ_μ^2/σ^2 is given by (L, U) where L and U are as given above. Note that this confidence interval can be used to carry out some tests. If we want to test $H_0 : \sigma_\mu^2 = 4\sigma^2$ against $H_1 : \sigma_\mu^2 \neq 4\sigma^2$, at a level α, then it is equivalent to testing $H_0 : \sigma_\mu^2/\sigma^2 = 4$ against $H_1 \sigma_\mu^2/\sigma^2 \neq 4$. In order to carry out this test we may construct a $(1-\alpha)100\%$ confidence interval for σ_μ^2/σ^2 and then check if 4 is inside this interval. If it is not, we may reject H_0.

Similarly, if we want to test $H_0 : \sigma_\mu^2/\sigma^2 = 4$ against $H_1 :\ \sigma_\mu^2/\sigma^2 > 4$, we need to construct a $(1-2\alpha)100\%$ confidence interval for σ_μ^2/σ^2. If 4 is smaller than the lower bound of this interval, then we will reject H_0.

Best linear unbiased predictor (BLUP) of τ_i.

Suppose that we wish to predict τ_i. We call this prediction since τ_i is supposed to be random. Consider a predictor of the form $H = d_1\bar{Y}_1. + \cdots + d_r\bar{Y}_r.$, $d_1, ..., d_r$ are constant to be chosen. Ideally we should consider a predictor of the form $\sum\sum c_{ij}Y_{ij}$, but heuristic arguments suggest that we may restrict ourselves to linear combinations of $\bar{Y}_1., ..., \bar{Y}_r.$ since in the fixed effects case $(\bar{Y}_1., ..., \bar{Y}_r., SSE)$ is sufficient for $(\mu.., \tau_1, .., \tau_r, \sigma^2)$ and SSE contains no information on $(\mu.., \tau_1, .., \tau_r)$ (see Remark below).If we restrict ourselves to unbiased estimators and find that estimate which has the smallest mean square error $E(H-\tau_i)^2 = [E(H-\tau_i)]^2 + Var(H-\tau_i)$. If we restrict ourselves unbiased predictors, i.e., $E(H-\tau_i) = 0$, then we need to minimize $Var(H-\tau_i)$ subject to the condition that $E(H-\tau_i) = 0$. Solution to this restricted minimization is unique and it is called the best linear unbiased predictor (BLUP), and it given by $H^* = w(\bar{Y}_{i\cdot} - \bar{Y}_{\cdot\cdot})$, where $w = n/(n + \sigma^2/\sigma_\mu^2)$, assuming σ_μ^2 and σ_μ^2 are known. Here is a result.

Fact 2. Consider a linear predictor of τ_i of the form $H = d_1\bar{Y}_1. + \cdots + d_r\bar{Y}_r..$ satisfying the constraint $E(H-\tau_i) = 0$. Then $E(H-\tau_i)^2$ is minimized when $d_i = w(1-1/r)$ and $d_{i'} = -w/r$, $i' \neq i$. The best predictor τ_i is of the form $w(\bar{Y}_{i\cdot} - \bar{Y}_{\cdot\cdot})$.

Note that the constant $0 \leq w \leq 1$. We can easily obtain an estimate of this by substituting estimated of σ_μ^2 and σ^2 in w. Note that $w = n\sigma_\mu^2/(n\sigma_\mu^2 + \sigma^2)$Thus estimated BLUP of τ_i is $\hat{\tau}_i = \hat{w}(\bar{Y}_{i\cdot} - \bar{Y}_{\cdot\cdot})$, where $\hat{w} = ns_\mu^2/MSTR$.

For our data, suppose we wish to obtain the BLUP of τ_4. Now

$$\hat{w} = ns_\mu^2/MSTR = (10)(19.368)/200.83 = 0.9644.$$

Estimated BLUP of τ_4 is

$$\hat{\tau}_4 = \hat{w}(\bar{Y}_4. - \bar{Y}_{\cdot\cdot}) = (0.9644)(210.7 - 205.05) = (0.9644)(5.65) = 5.449.$$

If it were a fixed effects case, then estimate of τ_2 be $\bar{Y}_4. - \bar{Y}_{\cdot\cdot} = 5.65$, but in the random case the BLUP of τ_4 is estimated to be equal to 5.45.

Remark: (i) Note that $\hat{w}$ should be in between 0 and 1. In case, the method given above it is negative, then the estimate of w should be set to 0.

(ii) It is easy to see that $SSE = \sum\sum(\varepsilon_{ij} - \bar{\varepsilon}_{i\cdot})^2$. Clearly this is free of $\mu_{\cdot\cdot}$ and τ's. Thus we may guess that SSE may not aid in getting a predictor for BLUP of τ's.

(iii) Ideally we should a general linear combinations such as $H = \sum\sum c_{ij}Y_{ij}$ in order to predict τ_i. It turns out that the best linear unbiased predictor of this general form is the same as the one given in Fact 2 above.

Some important features of the BLUP:

(i) if $\sigma_\mu^2 \to \infty$, $w \to 1$ and in that case $H^* \approx \bar{Y}_{i\cdot} - \bar{Y}_{\cdot\cdot}$, the estimate in the fixed effects case,

(ii) if $\sigma_\mu^2 \to 0$, $w \to 0$, and in that case $H^* \approx 0$ and $\tau_i \approx 0$ for all i, the model is $Y_{ij} \approx \mu_{\cdot\cdot} + \varepsilon_{ij}$.

This indicates that the random effects model is, in some sense, is between the following two models: $Y_{ij} = \mu_{\cdot\cdot} + \varepsilon_{ij}$ and, the fixed effects model, $Y_{ij} = \mu_{\cdot\cdot} + \tau_i + \varepsilon_{ij}$, α_i's fixed.

Appendix: Some proofs.

Proof of Fact 1: We will always use the identity: for any sequence of number $\{z_1, ..., z_m\}$,

$$\sum(z_i - \bar{z})^2 = \sum z_i^2 - m\bar{z}^2.$$

(i) This proof is trivial.

(ii) Denoting $\bar{\tau} = \sum\tau_i/r$ and $\bar{\varepsilon} = \sum\sum\varepsilon_{ij}/(nr)$, we can write

$$\bar{Y}_{\cdot\cdot} = \mu_{\cdot\cdot} + \bar{\tau} + \bar{\varepsilon}.$$

Since $\{\tau_i\}$ and $\{\varepsilon_{ij}\}$ are independent, $\bar{\tau}$ and $\bar{\varepsilon}$ are also independent. Hence

$$Var(\bar{Y}_{\cdot\cdot}) = Var(\bar{\tau}) + Var(\bar{\varepsilon}) = \sigma_\mu^2/r + \sigma^2/(nr) = \frac{n\sigma_\mu^2 + \sigma^2}{nr}.$$

(iii) Note that we can write

$$\begin{aligned} Y_{ij} - \bar{Y}_{i\cdot} &= \varepsilon_{ij} - \bar{\varepsilon}_{i\cdot}, \text{ and check that} \\ SSE &= \sum\sum(\varepsilon_{ij} - \bar{\varepsilon}_{i\cdot})^2 \\ &= \sum_i\left\{\sum_j(\varepsilon_{ij} - \bar{\varepsilon}_{i\cdot})^2\right\} = \sum_i\left\{\sum_j\varepsilon_{ij}^2 - n\bar{\varepsilon}_{i\cdot}^2\right\} \\ &= \sum\sum\varepsilon_{ij}^2 - n\sum_i\bar{\varepsilon}_{i\cdot}^2. \end{aligned}$$

Now $E(\varepsilon_{ij}^2) = \sigma^2$, $E(\bar{\varepsilon}_{i\cdot}^2) = \sigma^2/n$ and hence we have

$$\begin{aligned} E(SSE) &= nr\sigma^2 - n(r)(\sigma^2/n) = (n-1)r\sigma^2, \text{ and} \\ E(MSE) &= E(SSE/((n-1)r)) = \sigma^2. \end{aligned}$$

(iv) Note that

$$\begin{aligned}\bar{Y}_{i\cdot} - \bar{Y}_{\cdot\cdot} &= \tau_i - \bar{\tau} + \bar{\varepsilon}_{i\cdot} - \bar{\varepsilon}_{\cdot\cdot}, \quad \text{and} \\ SSTR &= n\sum(\bar{Y}_{i\cdot} - \bar{Y}_{\cdot\cdot})^2 = n\sum(\tau_i - \bar{\tau} + \bar{\varepsilon}_{i\cdot} - \bar{\varepsilon}_{\cdot\cdot})^2 \\ &= n\sum(\tau_i - \bar{\tau})^2 + n\sum(\bar{\varepsilon}_{i\cdot} - \bar{\varepsilon}_{\cdot\cdot})^2 + 2n\sum(\tau_i - \bar{\tau})(\bar{\varepsilon}_{i\cdot} - \bar{\varepsilon}_{\cdot\cdot}).\end{aligned}$$

Expectation of the cross product term $\sum(\tau_i - \bar{\tau})(\bar{\varepsilon}_{i\cdot} - \bar{\varepsilon}_{\cdot\cdot})$ is zero since τ's and ε's are independent and they have zero means. Now,

$$\begin{aligned}E\left[n\sum(\tau_i - \bar{\tau})^2\right] &= nE\left[\sum(\tau_i - \bar{\tau})^2\right] = nE\left[\sum\tau_i^2 - r\bar{\tau}^2\right] \\ &= n[r\sigma_\mu^2 - (r)(\sigma_\mu^2/r)] = n(r-1)\sigma_\mu^2.\end{aligned}$$

Note that

$$\begin{aligned}E\left[n\sum(\bar{\varepsilon}_{i\cdot} - \bar{\varepsilon}_{\cdot\cdot})^2\right] &= nE\left[\sum(\bar{\varepsilon}_{i\cdot} - \bar{\varepsilon}_{\cdot\cdot})^2\right] = nE\left[\sum\bar{\varepsilon}_{i\cdot}^2 - r\bar{\varepsilon}_{\cdot\cdot}^2\right] \\ &= n[(r)(\sigma^2/n) - (r)(\sigma^2/(nr))] = (r-1)\sigma^2.\end{aligned}$$

Hence we have,

$$\begin{aligned}E(SSTR) &= n(r-1)\sigma_\mu^2 + (r-1)\sigma^2 = [n\sigma_\mu^2 + \sigma^2](r-1) \text{ and hence} \\ E(MSTR) &= n\sigma_\mu^2 + \sigma^2.\end{aligned}$$

Proof of Fact 2.

For notational convenience we will take $i = 1$. Note that the condition $E(H - \tau_1) = 0$ means that $0 = \sum d_i E(\bar{Y}_{i\cdot}) - E(\tau_1) = \sum d_i \mu_{\cdot\cdot} - 0$. Since this equation holds for all $\mu_{\cdot\cdot}$ including $\mu_{\cdot\cdot} \neq 0$, we have $\sum d_i = 0$. Now note that

$$H = \sum d_i \bar{Y}_{i\cdot} = \sum d_i(\mu_{\cdot\cdot} + \tau_i + \bar{\varepsilon}_{i\cdot}) = \sum d_i\tau_i + \sum d_i\bar{\varepsilon}_{i\cdot}.$$

Since α's and ε's are independent of each other we have

$$\begin{aligned}Var(H - \tau_1) &= Var\left(\sum d_i\tau_i - \tau_1\right) + Var\left(\sum d_i\bar{\varepsilon}_{i\cdot}\right) \\ &= (d_1 - 1)^2\sigma_\mu^2 + \sum_{i\geq 2} d_i^2\sigma_\mu^2 + \sum d_i^2(\sigma^2/n) \\ &= \sigma_\mu^2 - 2d_1\sigma_\mu^2 + (\sigma_\mu^2 + \sigma^2/n)\sum d_i^2.\end{aligned}$$

We need to minimize $Var(H - \tau_1)$ subject to the constraint $\sum d_i = 0$. Towards this purpose, we may use the method of Lagrange multipliers to get the solution, i.e., minimize

$$Q = \sigma_\mu^2 - 2d_1\sigma_\mu^2 + (\sigma_\mu^2 + \sigma^2/n)\sum d_i^2 + 2\lambda\sum d_i$$

with respect to $d_1, ..., d_r, \lambda$ subject to the constraint $\partial Q/\partial\lambda = 0$. Note that

$$\begin{aligned}\partial Q/\partial d_1 &= -2\sigma_\mu^2 + 2(\sigma_\mu^2 + \sigma^2/n)d_1 + 2\lambda\\ \partial Q/\partial d_i &= 2(\sigma_\mu^2 + \sigma^2/n)d_i + 2\lambda,\ i \geq 2,\\ \partial Q/\partial\lambda &= \sum d_i.\end{aligned}$$

Setting these equations to zero and solving them we get (details are left as exercise)

$$d_1 = w(1 - 1/r),\ d_i = -w/r, i \geq 2.$$