Section 7.2 Two-way ANOVA with random effect(s)

1. Model with Two Random Effects

- The factors in higher-way ANOVAs can again be considered fixed or random, depending on the context of the study. For each factor:

- Are the levels of that factor of direct interest? Or do they just represent some larger “population” of levels that could have been included?

- If the study were to be conducted again, would the exact same levels of that factor be used again? Or would other levels be used?

Examples

Example 1.

Potassium measurements across labs Ten commercial laboratories across the UK were chosen by administrators of the National Quality Control Scheme.

- They are interested in differences in potassium measurements from serum samples across labs. Ten serum samples are created, each of which contains a pre-determined quantity of potassium. Each specimen is divided into ten equal portions.
- One portion from each specimen is sent to each lab in a completely randomized design.

Example 2.

Research nursing staff need to be trained to measure sub-scapular skin fold with calipers for an ongoing single-center clinical trial.

- Within the center, b patients are randomly chosen; a random sample of a nurses at the center measure each patient in a randomized order.
- Multiple measurements may be taken r times.

ANOVA Model with two random effects (balanced design)

$$Y_{ijk} = \mu + \alpha_i + \beta_j + (\alpha\beta)_{ij} + \varepsilon_{ijk},$$

where $\varepsilon_{ijk} \sim^{iid} N\left(0, \sigma^2\right), i = 1,\ldots,a,\ j = 1,\ldots,b, k = 1,..,n,$ and

$$\alpha_i \sim^{iid} N\left(0, \sigma_\alpha^2\right),\ \beta_j \sim^{iid} N\left(0, \sigma_\beta^2\right),\ (\alpha\beta)_{ij} \sim^{iid} N\left(0, \sigma_{\alpha\beta}^2\right), \forall i, j,$$

and $\alpha_i, \beta_j, (\alpha\beta)_{ij}, \varepsilon_{ijk}$ are all independent.

$$EY_{ijk} = \mu, \forall i, j, k$$

The covariance structure of the response:

$$Cov\left(Y_{ijk}, Y_{i'jk'}\right) = \sigma_\beta^2, \forall i \neq i',$$

$$Cov\left(Y_{ijk}, Y_{ij'k'}\right) = \sigma_\alpha^2, \forall j \neq j',$$

$$Cov\left(Y_{ijk}, Y_{ijk'}\right) = \sigma_\alpha^2 + \sigma_\beta^2 + \sigma_{\alpha\beta}^2, \forall k \neq k',$$

$$\mathrm{Var}\left(Y_{ijk}\right) = \sigma^2 + \sigma_\alpha^2 + \sigma_\beta^2 + \sigma_{\alpha\beta}^2$$

$$Cov\left(Y_{ijk}, Y_{i'j'k'}\right) = 0, i \neq i', j \neq j', k \neq k'.$$

1. Deviation of response from the grand mean in (i, j) - th cell:

$$\left(Y_{ijk} - \overline{Y}_{...}\right) = \left(\overline{Y}_{ij.} - \overline{Y}_{...}\right) + \left(Y_{ijk} - \overline{Y}_{ij.}\right),$$

$$\text{SSTO} = \sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{n}\left(Y_{ijk} - \overline{Y}_{...}\right)^2$$

$$\text{SSTR} = \sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{n}\left(\overline{Y}_{ij.} - \overline{Y}_{...}\right)^2 = n\sum_{i=1}^{a}\sum_{j=1}^{b}\left(\overline{Y}_{ij.} - \overline{Y}_{...}\right)^2$$

$$\text{SSE} = \sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{n}\left(Y_{ijk} - \overline{Y}_{ij.}\right)^2$$

2. Deviation Decomposition of the (i, j) - th cell mean:

$$\left(\overline{Y}_{ij.} - \overline{Y}_{...}\right) = \left(\overline{Y}_{i..} - \overline{Y}_{...}\right) + \left(\overline{Y}_{.j.} - \overline{Y}_{...}\right) + \left(\overline{Y}_{ij.} - \overline{Y}_{i..} - \overline{Y}_{.j.} + \overline{Y}_{...}\right)$$

$$\text{SSA} = \sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{n}\left(\overline{Y}_{i..} - \overline{Y}_{...}\right)^2 = bn\sum_{i=1}^{a}\left(\overline{Y}_{i..} - \overline{Y}_{...}\right)^2$$

$$\text{SSB} = \sum_{i=1}^{a}\sum_{j=1}^{b}\sum_{k=1}^{n}\left(\overline{Y}_{.j.} - \overline{Y}_{...}\right)^2 = an\sum_{j=1}^{b}\left(\overline{Y}_{.j.} - \overline{Y}_{...}\right)^2$$

$$\text{SSAB} = \text{SSTR} - \text{SSA} - \text{SSB} = \left(\text{SSTO} - \text{SSE}\right) - \text{SSA} - \text{SSB}$$

ANOVA with two random effects

Source	SS	DF	MS	F
A	SSA	a-1	MSA = SSA / (a-1)	F_A = MSA / MSAB
B	SSB	b-1	MSB = SSB / (b-1)	F_B = MSB / MSAB
AB	SSAB	(a-1)(b-1)	MSAB = SSAB / ((a-1)(b-1))	F_{AB} = MSAB / MSE
Error	SSE	ab(n-1)	MSE = SSE/(ab(n-1))	
Total	SSTO	abn-1		

Expectations:

$$E(MSA) = \sigma^2 + n\sigma^2_{\alpha\beta} + bn\sigma^2_{\alpha},$$

$$E(MSB) = \sigma^2 + n\sigma^2_{\alpha\beta} + an\sigma^2_{\beta},$$

$$E(MSAB) = \sigma^2 + n\sigma^2_{\alpha\beta}$$

$$E(MSE) = \sigma^2,$$

Under $H_0 : \sigma^2_{\alpha} = 0 \quad \text{vs.} \quad H_1 : \sigma^2_{\alpha} > 0$

$$F_A \sim F((a-1),(a-1)(b-1));$$

under $H_0 : \sigma^2_{\beta} = 0 \quad \text{vs.} \quad H_1 : \sigma^2_{\beta} > 0$

$$F_B \sim F((b-1),(a-1)(b-1));$$

under $H_0 : \sigma^2_{\alpha\beta} = 0 \quad \text{vs.} \quad H_1 : \sigma^2_{\alpha\beta} > 0$

$$F_{AB} \sim F((a-1)(b-1), ab(n-1)).$$

Variance Components Estimation

Estimators of the variance components:

$$\hat{\sigma}^2 = MSE,\ \hat{\sigma}^2_{\alpha\beta} = \frac{MSAB - MSE}{n},\ \hat{\sigma}^2_{\beta} = \frac{MSB - MSAB}{an},\quad \hat{\sigma}^2_{\alpha} = \frac{MSA - MSAB}{bn}$$

EXAMPLE 1: A Measurement Systems Capability Study

Statistically designed experiments are frequently used to investigate the sources of ability that affect a system. A common industrial application is to use a designed experiment to study the components of variability in a measurement system. These studies are often called gauge capability studies or gauge repeatability and reproducibility (R&R) studies because these are the components of variability that are of interest (for more discussion of gauge R&R studies.

An instrument or gauge is used to measure a critical dimension on a part. Twenty parts have been selected from the production process, and three randomly selected operators measure each part twice with this gauge. The order in which the measurements are made is completely randomized, so this is a two-factor factorial experiment with design factor parts and operators, with two replications. Both parts and operators are random factors.

Source	DF	Type I SS	Mean Square	F Value	Pr > F
part	19	1185.425000	62.390789	62.92	<.0001
operator	2	2.616667	1.308333	1.32	0.2750
part*operator	38	27.050000	0.711842	0.72	0.8614

2. Model with fixed and random effects

Mixed - effects ANOVA Model

$$Y_{ijk} = \mu + \alpha_i + \beta_j + (\alpha\beta)_{ij} + \varepsilon_{ijk},$$

where $\varepsilon_{ijk} \sim^{iid} N\left(0, \sigma^2\right), i = 1,\ldots,a,\ j = 1,\ldots,b, k = 1,..,n,$ and

fixed factor A : $\sum_{i=1}^{a} \alpha_i = 0,$ random factor B : $\beta_j \sim^{iid} N\left(0, \sigma_\beta^2\right),$

interaction effect : $(\alpha\beta)_{ij} \sim N\left(0, \frac{a-1}{a}\sigma_{\alpha\beta}^2\right)$ subject to restrictions

$$\begin{cases} \sum_i (\alpha\beta)_{ij} = 0 \\ Cov\left((\alpha\beta)_{ij}, (\alpha\beta)_{i'j}\right) = \frac{-1}{a}\sigma_{\alpha\beta}^2, \forall i \neq i' \end{cases}$$

and $\beta_j, (\alpha\beta)_{ij}, \varepsilon_{ijk}$ are pairwise independent.

(1). $E\left(Y_{ijk}\right) = \mu + \alpha_i, \mathrm{Var}\left(Y_{ijk}\right) = \sigma^2 + \sigma_\beta^2 + \frac{a-1}{a}\sigma_{\alpha\beta}^2$

(2). $Cov\left(Y_{ijk}, Y_{i'jk'}\right) = \sigma_\beta^2 - \frac{1}{a}\sigma_{\alpha\beta}^2, \forall i \neq i',$

$$Cov\left(Y_{ijk}, Y_{ijk'}\right) = \sigma_\beta^2 + \frac{a-1}{a}\sigma_{\alpha\beta}^2, \forall k \neq k',$$

$$Cov\left(Y_{ijk}, Y_{i'j'k'}\right) = 0, i \neq i', j \neq j', k \neq k'.$$

ANOVA with mixed effects

Source	SS	DF	MS	F
A	SSA	a-1	MSA = SSA / (a-1)	F_A = MSA / MSAB
B	SSB	b-1	MSB = SSB / (b-1)	F_B = MSB / MSE
AB	SSAB	(a-1)(b-1)	MSAB = SSAB / ((a-1)(b-1))	F_{AB} = MSAB / MSE
Error	SSE	ab(n-1)	MSE = SSE/(ab(n-1))	
Total	SSTO	abn-1		

Expectations of MSs:

$$E(MSA) = \sigma^2 + nb\frac{\sum \alpha_i^2}{(a-1)} + n\sigma_{\alpha\beta}^2,$$

$$E(MSB) = \sigma^2 + na\sigma_\beta^2,$$

$$E(MSAB) = \sigma^2 + n\sigma_{\alpha\beta}^2$$

$$E(MSE) = \sigma^2,$$

Under $H_0: \alpha_i = 0, \forall i$ vs. H_1: at least one $\alpha_i \neq 0$.

$$F_A \sim F((a-1),(a-1)(b-1));$$

under $H_0: \sigma_\beta^2 = 0$ vs. $H_1: \sigma_\beta^2 > 0$

$$F_B \sim F((b-1), ab(n-1));$$

under $H_0: \sigma_{\alpha\beta}^2 = 0$ vs. $H_1: \sigma_{\alpha\beta}^2 > 0$

$$F_{AB} \sim F((a-1)(b-1), ab(n-1)).$$

Estimation of fixed Effects in Mixed Model

1. Least Square Estimator of i-th mean and effect of fixed factor A :

$$\hat{\mu}_i = \overline{Y}_{i\cdot\cdot}, \quad \hat{\alpha}_i = \overline{Y}_{i\cdot\cdot} - \overline{Y}_{\cdot\cdot\cdot}, \quad i = 1, \ldots, a$$

2. Variance of fixed effects :

$$\operatorname{Var}(\hat{\alpha}_i) = \frac{\sigma^2 + n\sigma^2_{\alpha\beta}}{bn} = \frac{E(MSAB)}{bn}.$$

Squared of standard error : $s^2(\hat{\alpha}_i) = MSAB/(bn)$

3. Confidence Interval for α_i : $\left(\overline{Y}_{i\cdot\cdot} - \overline{Y}_{\cdot\cdot\cdot}\right) \pm t_{\frac{\alpha}{2}}((a-1)(b-1)) \cdot \sqrt{\dfrac{MSAB}{bn}}$

Follow-up Test :

(1). Pariwise comparison – C.I. for α_i and α_j :

$$\left(\overline{Y}_{i\cdot\cdot} - \overline{Y}_{j\cdot\cdot}\right) \pm t_{\frac{\alpha}{2}}((a-1)(b-1)) \cdot \sqrt{2\frac{MSAB}{bn}}$$

(2). Multiple Comparison – Tukey's simultaneous C.I. for $(\alpha_i - \alpha_j)$, $\forall$i, j = 1,..., a,

$$\left(\overline{Y}_{i\cdot\cdot} - \overline{Y}_{j\cdot\cdot}\right) \pm q_{\alpha}(a, (a-1)(b-1)) \cdot \sqrt{\frac{MSAB}{bn}}$$

(3). Confidence Interval for a contrast $L = \sum c_i \alpha_i$:

$$\sum c_i \overline{Y}_{i\cdot\cdot} \pm t_{\frac{\alpha}{2}}((a-1)(b-1)) \cdot \sqrt{\frac{MSAB}{bn} \sum c_i^2}$$

Variance Component Estimation in Mixed Model

Variance estimator for random effect and interaction effect :

$$\hat{\sigma}_{\beta}^{2} = \frac{MSB - MSE}{an}, \hat{\sigma}_{\alpha\beta}^{2} = \frac{MSAB - MSE}{n}$$

Example 2 : A company training program tries to investigate the effets of four different training methods (factor A). Five instructors (factor B) are selected at random to use all training methods in different classes. Four classes are assigned to each training - instructor combination. Response variable is the mean improvement per student in the class at the end of the training program.

Source	SS	DF	MS	F
A (training)	42.1			
B (instructor)	53.9			
AB interaction	46.7			
Error	126.4			
Total	269.1			

SAS code: Example 1 (cont). Capability Study

TABLE 25.5 Expected Mean Squares for Balanced Two-Factor ANOVA Models.

Mean Square	df	Fixed ANOVA Model (A and B fixed)	Random ANOVA Model (A and B random)	Mixed ANOVA Model (A fixed, B random)
MSA	$a-1$	$\sigma^2 + nb\frac{\sum \alpha_i^2}{a-1}$	$\sigma^2 + nb\sigma_\alpha^2 + n\sigma_{\alpha\beta}^2$	$\sigma^2 + nb\frac{\sum \alpha_i^2}{a-1} + n\sigma_{\alpha\beta}^2$
MSB	$b-1$	$\sigma^2 + na\frac{\sum \beta_j^2}{b-1}$	$\sigma^2 + na\sigma_\beta^2 + n\sigma_{\alpha\beta}^2$	$\sigma^2 + na\sigma_\beta^2$
$MSAB$	$(a-1)(b-1)$	$\sigma^2 + n\frac{\sum\sum(\alpha\beta)_{ij}^2}{(a-1)(b-1)}$	$\sigma^2 + n\sigma_{\alpha\beta}^2$	$\sigma^2 + n\sigma_{\alpha\beta}^2$
MSE	$(n-1)ab$	σ^2	σ^2	σ^2