RANDOM-EFFECTS MODEL

Example

A manufacturer was developing a new spectrophotometer for use in medical clinical laboratories. A critical component of instrument performance is the consistency of measurements from day to day and among machines.

The scientist who developed the instrument wanted to know if the variability of measurements among machines operated over several days was within acceptable standards for clinical applications.

The scientist set up a factorial experiment with "machines" and "days" as factors. Four machines were to be tested on four separate days in a 4×4 arrangement.

Two serum samples were randomly assigned to each of the four machines on each of the four days in a completely randomized design.

Our goal will be to assess the amount of variability in lipid measurements that can be attributed to machine and to day.

A manufacturer was developing a new spectrophotometer for use in medical clinical laboratories. A critical component of instrument performance is the consistency of measurements from day to day and among machines.

The scientist who developed the instrument wanted to know if the variability of measurements among machines operated over several days was within acceptable standards for clinical applications.

The scientist set up a factorial experiment with "machines" and "days" as factors. Four machines were to be tested on four separate days in a 4 by 4 arrangement.

Two serum samples were randomly assigned to each of the four machines on each of the four days in a completely randomized design.

Our goal will be to assess the amount of variability in lipid measurements that can be attributed to machine and to day.

Treatment layout

		Machine 1	2	3	4	
Day	1					y(ijk) is the kth measurement on the jth machine and ith day
	2	(2 msmts per day/machine machine combination)				
	3					
	4					

This table shows the treatment layout in this experiment.

Random effects

Because the four machines can be thought of as **a random sample from a larger population** of machines, and because we are **not interested in drawing conclusions about the specific machines chosen**, machine can be thought of as a ***random effect***.

Similarly, it makes sense to model day as a random effect: the days can be thought of as **a random sample from a larger population** of days, and we are not interested in drawing conclusions about the specific days chosen.

Our goal will to be to assess the amount of variability in lipid measurements that can be attributed to machine and to day.

Because the four machines can be thought of as a random sample from a larger population of machines, and because we are not interested in drawing conclusions about the specific machines chosen, machine can be thought of as a random effect.

Similarly, it makes sense to model day as a random effect: the days can be thought of as a random sample from a larger population of days, and we are not interested in drawing conclusions about the specific days chosen.

Our goal will to be to assess the amount of variability in lipid measurements that can be attributed to machine and to day.

Random-effect factors versus fixed-effect factors

Random effects (or random-effect factors) can be a difficult concept to comphrehend initially. To determine whether a factor should be treated as a random-effect factor or a fixed-effect factor, we ask these two questions:

1. Are the factor levels in the experiment a sample from a larger population? Or are they already representing all possible levels?
2. Do we want to generalize the experimental conclusions to levels not seen in the experiment?

If the answer is "yes" to both questions, the factor should be treated as an random-effect factor. For example

- In clinical studies, sex and race are usually considered fixed-effect factors.
- In the lipid measurements example, the machines used in the experiment will not the be machines used in labs. The days when the experiment was performed will not be the same days that the machines will be used in real applications. So both "machine" and "day" were treated as random-effect factors.

Random effects (or random-effect factors) can be a difficult concept to comphrehend initially. To determine whether a factor should be treated as a random-effect factor or a fixed-effect factor, we ask these two questions:

1.Are the factor levels in the experiment a sample from a larger population? Or are they already representing all possible levels?

1.Do we want to generalize the experimental conclusions to levels not seen in the experiment?

If the answer is "yes" to both questions, the factor should be treated as an random-effect factor. For example

In clinical studies, sex and race are usually considered fixed-effect factors.

In the lipid measurements example, the machines used in the experiment will not the be machines used in labs. The days when the experiment was performed will not be the same days that the machines will be used in real applications. So both "machine" and "day" were treated as random-effect factors.

Random effects or fixed effects: recap

In the lipids measurements example, why treating the machines and days as random effects?

We ask a simple question: when we make measurements in practice, will we be using the same machines and on the same days as in the experiment?

No. The days and machines in the experiment were selected as **a representation of a larger collections of days and machines**.

Just to recap: in the lipids measurements example, why treating the machines and days as random effects?

We ask a simple question: when we make measurements in practice, will we be using the same machines and on the same days as in the experiment? The answer is No.

The days and machines in the experiment were selected as **a representation of a larger collections (populations) of days and machines**.

The model

The two-factor random-effects model for a completely randomized design: $y_{ijk} = \mu + a_i + b_j + (ab)_{ij} + e_{ijk}$, where

- $\mu =$ over all mean
- $a_i \sim N(0, \sigma_a^2)$, effect of ith level of row factor
- $b_j \sim N(0, \sigma_b^2)$, effect of jth level of column factor
- $(ab)_{ij} \sim N(0, \sigma_{ab}^2)$, ijth interaction effect
- $e_{ijk} \sim N(0, \sigma^2)$, random error (within day/machine)

$i = 1, \ldots, a$, $j = 1, \ldots, b$, and $k = 1, \ldots, r$.

Note that all terms in the model except μ are random variables! σ_a^2, σ_b^2, σ_{ab}^2 and σ^2 are called ***variance components***.

All of the effects are assumed to be **independent**. Therefore,

$$\text{Var } y_{ijk} = \sigma_a^2 + \sigma_b^2 + \sigma_{ab}^2 + \sigma^2.$$

The model equation for the two-factor random-effects model look similar to the two-factor fixed-effects model.

But note that (a i), (b j) and (a b i j) are all random! They are assumed to be normally distributed with mean 0 and variances (sigma a squared), (sigma b squared) and (sigma ab squared) respectively.

(sigma a squared), (sigma b squared) and (sigma ab squared) are called random components.

All of the random effects are assumed to be independent.

These random components are usually the main interest of inference in random-effects models: we want to attribute the total variance to different sources. Unlike in fixed-effect models, in random-effects models, we are often not interested in the specific mean levels in each treatment groups.

Covention

When writing a model equation,

- we use roman letters, a_i, b_j, for random-effects terms.
- we use greek letters, α_i, β_j, for fixed-effects terms.

When writing a model equation, we follow the convention of using roman letters, (a i), (b j) for random-effects terms and using greek letters, (alpha i), (beta j), for fixed-effects terms.

Sum of squares and the expected values of mean squares

SSA , SSB , SSAB , and SSE are calculated as in the fixed-effects case, but test statistics are different in some cases.

Treating the factors as random will affect **the expected values of the sum of squares** (since we have more random terms in the model), which in turn will affect the denominators to use in different F-tests.

It can be shown (Kuehl, p. 234) that

$$E(\text{MSA}) = \sigma^2 + r\sigma^2_{ab} + br\sigma^2_a$$
$$E(\text{MSB}) = \sigma^2 + r\sigma^2_{ab} + ar\sigma^2_b$$
$$E(\text{MSAB}) = \sigma^2 + r\sigma^2_{ab}$$
$$E(\text{MSE}) = \sigma^2$$

We can summarize information regarding the partitioning of the sum of squares mean squares, degrees of freedom, and F-tests compactly in an ANOVA table.

SSA , SSB , SSAB , and SSE are calculated as in the fixed-effects case, but test statistics are different in some cases.

Treating the factors as random will affect the expected values of the sum of squares since we have more random terms in the model, which in turn will affect the denominators to use in different F-tests.

Here we list the expected values of MSA, MSB, MSAB and MSE.

We can summarize information regarding the partitioning of the sum of squares mean squares, degrees of freedom, and F-tests compactly in an ANOVA table.

The ANOVA table for a two-factor random-effects model

Source of Variation	Degrees of Freedom	Sum of Squares	Mean Square	Expected Mean Square	F-test
A	$a-1$	SSA	MSA	$\sigma^2 + r\sigma^2_{ab} + br\sigma^2_a$	MSA/MSAB
B	$b-1$	SSB	MSB	$\sigma^2 + r\sigma^2_{ab} + ar\sigma^2_b$	MSB/MSAB
$A \times B$	$(a-1)(b-1)$	SSAB	MSAB	$\sigma^2 + r\sigma^2_{ab}$	MSAB/MSE
Error	$ab(r-1)$	SSE	MSE	σ^2	
Total	$abr-1$	SS Total			

Pay particular attention to the "Expected Mean Square" column: it suggests the correct dominator to use in a F-test.

In this ANOVA table for a two-factor random-effects model, pay particular attention to the last two columns: the expected mean squares, and the F-tests.

We see that for testing A effect or B effect, we use MSAB as the denominator in the F-test, not the MSE.

The null hypotheses being tested by the F-test are also different from the what tested in fixed-effects models.

Hypothesis tests in random-effects models

Note that in random-effects models, we focus more on the random components σ_a^2, σ_b^2, σ_{ab}^2, rather than of the mean parameters.

The null hypotheses to test are often

- $H_0: \sigma_{ab}^2 = 0$: no interaction effect, or
- $H_0: \sigma_a^2 = 0$ or $H_0: \sigma_b^2 = 0$, no main effects.

Note that in random-effects models, we focus more on the random components (sigma a squared), (sigma b squared), (sigma a b squared), rather than of the mean parameters.

The null hypotheses to test are often

(sigma a b squared equals 0), which means no interaction effect, or

(sigma a squared equals 0), (sigma b squared equals 0) which correspond to no main effects.

Test no interaction effect

To test $H_0: \sigma_{ab}^2 = 0$ (no interaction), compare $F_0 = \text{MSAB}/\text{MSE}$ to an $F_{(a-1)(b-1),ab(r-1)}$ distribution (one can find the numbers of d.f. from the ANOVA table).

From the ANOVA table,

$$E(\text{MSAB}) = \sigma^2 + r\sigma_{ab}^2$$

- If H_0 in true, $E(\text{MSAB}) = E(\text{MSE}) = \sigma^2$.
- If H_0 is false, $E(\text{MSAB}) > E(\text{MSE})$, and F_0 tends to be large.

To test the null hypothesis of no interaction or (sigma a b squared equals 0), we compare MSAB to MSE using a F-test with numbers of d.f. (a - 1, b - 1).

Note that from the ANOVA table, the expected value of MSAB is (sigma squared) plus (sigma ab squared).

If (H naught) is true, (sigma ab squared) equals 0, then the expected value of MSAB equals the expected value of MSE.

If the null hypothesis is false, the expected value of MSAB is greater than the expected value of MSE.

Test main effects

To test $H_0: \sigma_a^2 = 0$, compare $F_0 = MSA/MSAB$ to $F_{a-1,(a-1)(b-1)}$.

Note that **we use MSAB as the denominator here**, not MSE as in the fixed-effects case.

Why? From the ANOVA table,

$$E(\text{MSA}) = \sigma^2 + r\sigma_{ab}^2 + br\sigma_a^2$$

If $H_0: \sigma_a^2 = 0$ is true,

$$E(MSA) = E(MSAB) = \sigma^2 + r\sigma_{ab}^2 \neq E(MSE).$$

Similarly, to test $H_0: \sigma_b^2 = 0$, compare $F_0 = MSB/MSAB$ to $F_{b-1,(a-1)(b-1)}$.

To analyze random effects models in R function using `lm`, one may need to compute the F-tests manually: the lm function will not be able to tell whether a factor is random-effect factor or a fixed-effect factor, and the ANOVA output assumes a fixed-effects model.

To test the main effect of factor A, we use the ratio of MSA to MSAB as the F-test statistic.

Note that we use MSAB as the denominator here, not MSE as in the fixed-effects case.

The reason to use MSAB as the denominator is that when (H naught) is true, the expected of MSA is the same as the expected value of MSAB, not that of MSE.

Similarly, to test the main effect of factor B, we use the ratio of MSB to MSAB as the F-test statistic.

To analyze random effects models in R function using `lm`, one may need to compute the F-tests manually: the lm function will not be able to tell whether a factor is random-effect factor or a fixed-effect factor, and the ANOVA output assumes a fixed-effects model.

ANOVA table for the lipid measurements data

Table 7.2 Analysis of variance for spectrophotometric readings from four machines on each of four days

Source of Variation	*Degrees of Freedom*	*Mean Square*	*Expected Mean Square*
Day	3	$MSD = 445$	$\sigma^2 + r\sigma^2_{dm} + rb\sigma^2_d$
Machine	3	$MSM = 549$	$\sigma^2 + r\sigma^2_{dm} + ra\sigma^2_m$
Interaction	9	$MS(DM) = 87$	$\sigma^2 + r\sigma^2_{dm}$
Error	16	$MSE = 18$	σ^2

ANOVA Table

Here is the ANOVA table for the lipid data. Note that the column of expected mean squares will not be available from an R anova output.

Test the machine-day interaction in the lipid example

To test the machine-day interaction in the lipid example, we use $F_0 = MS(DM)/MSE = 87/18 = 4.83$ as the F-test statistic. The numbers of d.f. for this F-test are (9, 16) (from the ANOVA table).

The p-value is $\Pr(F_{9,16} > 4.83) = 0.003$. The machines perform inconsistently from day to day.

To test the machine-day interaction in the lipid example, we use MS(DM) over MSE as the F-test statistic. The numbers of d.f. for this F-test are (9, 16).

The p-value of the test is 0.003. The test result indicates that the machines perform inconsistently from day to day.

Estimating the variance components (random effects model)

To estimate the variance components, σ_a^2, σ_b^2 and σ_{ab}^2. We use the method of moments: we equate the mean squares to their expected values and solve for the estimates of the needed random components.

For example, since $E(\text{MSE}) = \sigma^2$, we let $\hat{\sigma}^2 = MSE$.

$E(\text{MSAB}) = \sigma^2 + r\sigma_{ab}^2$, so we let

$$MSAB = \hat{\sigma}^2 + r\hat{\sigma}_{ab}^2.$$

This equation, combined with $\hat{\sigma}^2 = MSE$, gives

$$\hat{\sigma}_{ab}^2 = \frac{\text{MSAB} - \text{MSE}}{r}.$$

Estimates of the other two variance components (Kuehl, p.235):

$$\hat{\sigma}_a^2 = \frac{\text{MSA} - \text{MSAB}}{rb}, \quad \hat{\sigma}_b^2 = \frac{\text{MSB} - \text{MSAB}}{ra}.$$

To estimate the variance components, we use the method of moments: we equate the mean squares to their expected values and solve for the estimates of the needed random components.

For example, since the expected value of MSE is (sigma squared), the error variance, we can use MSE as an estimate of (sigma squared).

Since the expected value of MSAB is (sigma squared + r times (sigma a b squared)), we let MSAB equal (sigma squared hat) + r times (sigma a b squared hat) . This equation, combined with (sigma squared hat equals MSE), gives

(sigma a b squared hat) equals (MSAB - MSE) / r.

Estimates of the other two variance components are

(sigma a squared hat) equals (MSA - MSAB) / rb, (sigma b squared hat) equals (MSB - MSAB) / ra.

The lipid-measurement example

In the lipid-measurement example (p. 235):

Source	Estimated variance	Percent of total
Days	44.8	29
Machines	57.8	37
Interaction	34.5	22
Error	18.0	12

This table summarizes the variance components estimated in the lipid-measurements example.

For each random component, it also shows the percentage of total variance that is attributable tot the corresponding random-effect factor.

In this example, the day, machine, and their interaction each contribute to a large proportion of the total variance.

Data analysis using R

See example7.1.html for more details on analyzing the lipid-measurements example using R.

See example7.1.html for more details on analyzing the lipid-measurements example using R.

Summary

- Random effects (or random-effect factors) versus fixed effects (fixed-effect factors)
- Random-effects model for a two-factor experiment with completely randomized design
- Testing the main and interaction effects in a random-effects model
  - The null hypotheses
  - The denominator to use in the F-tests
- ANOVA table for two-factor random-effects model (from a complete-randomized design)
  - Use the correct denominator in a F-test
  - Estimating the variance components

In this lecture, we discussed the differences between random-effects or random-effect factors and fixed effects or fixed-effect factors.

We studied he random-effects model for a two-factor experiment with completely randomized design.

When testing the main and interaction effects in a random-effects model, notice the differences in the null hypotheses and forms of the F-tests as compared to the fixed-effects model case

Using the ANOVA table for two-factor random-effects model, we should be able to perform the correct F-tests for the main and interaction effects, and estimate the variance components.