Components of variance

- Example: 1-way layout/k-group comparison/ 1-way ANOVA
- $y_{tr} = \mu + \alpha_t + \epsilon_{tr}, \quad r = 1, \dots, R; t = 1, \dots, T, \quad \epsilon_{tr} \sim (0, \sigma^2)$
- $\hat{\alpha}_t = \bar{y}_{t.} - \bar{y}_{..}, \quad$ under constraint $\sum \alpha_t = 0$
- $\text{var}(\bar{y}_{t.} - \bar{y}_{s.}) = \frac{2\sigma^2}{R}$
- $\hat{\sigma}^2 = s^2 = \sum_{r,t}(y_{tr} - \bar{y}_{t.})^2 / T(R-1)$
- ANOVA:

Components of variance

- Example: 1-way layout/k-group comparison/ 1-way ANOVA
- $y_{tr} = \mu + \alpha_t + \epsilon_{tr}, \quad r = 1, \ldots, R; t = 1, \ldots, T, \quad \epsilon_{tr} \sim (0, \sigma^2)$
- $\hat{\alpha}_t = \bar{y}_{t.} - \bar{y}_{..}$, under constraint $\sum \alpha_t = 0$
- $\text{var}(\bar{y}_{t.} - \bar{y}_{s.}) = \frac{2\sigma^2}{R}$
- $\hat{\sigma}^2 = s^2 = \sum_{r,t}(y_{tr} - \bar{y}_{t.})^2 / T(R-1)$
- ANOVA:

Source	df	SS	E(MS)
betw	$T-1$	$\sum_{t,r}(\bar{y}_{t.} - \bar{y}_{..})^2$	$\sigma^2 + *$
with	$T(R-1)$	$\sum_{t,r}(y_{tr} - \bar{y}_{t.})^2$	σ^2

$R\sum\alpha_t^2/(T-1)$ $*$ (ntbc)

Change model parameterization

- $y_{tr} = \mu + b_t + \epsilon_{tr}, \quad \epsilon_{tr} \sim (0, \sigma^2), \quad b_t \sim (0, \sigma_b^2)$ (ind.)
- $SS_b =$
- ANOVA unchanged
- $E(MS_b) =$

$\mathrm{cor}(y_{tr}, y_{ts})$

Change model parameterization

$\sum(\bar{x}_i - \bar{\bar{x}})^2 \sim \chi^2_{n-1}$

- $y_{tr} = \mu + b_t + \epsilon_{tr}, \quad \epsilon_{tr} \sim (0, \sigma^2), \quad b_t \sim (0, \sigma_b^2)$
- $SS_b = \sum_t \sum_r (\bar{y}_{t.} - \bar{y}_{..})^2 = \sum_t R(\bar{y}_{t.} - \bar{y}_{..})^2$

$\sim (\sigma^2 + R\sigma_b^2)\chi^2_{T-1}$

- ANOVA unchanged
- $E(MS_b) =$

$\text{var}\, \bar{y}_{t.} = \text{var} \frac{\sum y_{tr}}{R^2}$

$= \text{var} \frac{\sum(b_t + \epsilon_{tr})}{R^2}$

$= \ldots =$

$\text{cor}(y_{tr}, y_{ts})$

Change model parameterization

$-\frac{1}{T-1} \le \rho \le 1$

- $y_{tr} = \mu + b_t + \epsilon_{tr}, \quad \epsilon_{tr} \sim (0, \sigma^2), \quad b_t \sim (0, \sigma_b^2)$

Intraclass correlation

$R \frac{\sum \alpha_t^2}{T-1}$

- $SS_b =$
- ANOVA unchanged
- $E(MS_b) =$

		E(MS)
betw	$T-1$	$\sigma^2 + R\sigma_b^2$
with	$T(R-1)$	σ^2

$\frac{\sigma_b^2}{\sigma_b^2 + \sigma^2} = \rho$

$\text{cor}(y_{tr}, y_{ts}) = \frac{\text{cov}(y_{tr}, y_{ts})}{\sqrt{v_1}\sqrt{v_2}} = \frac{\text{cov}(b_t + \epsilon_{tr}, b_t + \epsilon_{ts})}{}$

Inference

- Under $H_0 : \sigma_b^2 = 0$: $E(MS_w) = \sigma^2$ $\tilde{\sigma}^2 = MS_w$

- Estimation of σ^2 and σ_b^2 $E(MS_b) = \sigma^2 + R\sigma_b^2$

m.o.m. $\frac{1}{R}(MS_b - MS_w) = \tilde{\sigma}_b^2$

- Estimation of σ_b^2/σ^2 using F distribution

- Estimation of μ:

- See Example 9.14: Exercise: verify CI for ratio

Inference

- Under $H_0 : \sigma_b^2 = 0$:

$$\frac{MS_b / (\sigma^2 + R\sigma_b^2)}{MS_w / \sigma^2} \sim F_{T-1,\, T(R-1)}$$

- Estimation of σ^2 and σ_b^2

- Estimation of σ_b^2/σ^2 using F distribution

p. 450

90% CI

$(-0.01, 1.34)$

- Estimation of μ:

- See Example 9.14: Exercise: verify CI for ratio

Inference

- Under $H_0 : \sigma_b^2 = 0$:

- Estimation of σ^2 and σ_b^2

$n+bc$

- Estimation of σ_b^2/σ^2 using F distribution

- Estimation of μ: $E(\bar{y}_{..}) = \mu \quad \text{var}\, \bar{y}_{..} = \frac{\sigma^2 + R\sigma_b^2}{RT}$

- See Example 9.14: Exercise: verify CI for ratio

Inference

- Under $H_0 : \sigma_b^2 = 0$:

- Estimation of σ^2 and σ_b^2

- Estimation of σ_b^2/σ^2 using F distribution

- Estimation of μ:

- See Example 9.14: Exercise: verify CI for ratio

Fixed or random effects?

- depends on context
- one rule of thumb: if factor levels are a random sample of all possible levels
- which is of interest: mean or variance?
- nested factors are often modelled with random effects
- are levels of factor in one group same as levels of factors in another group?

Fixed or random effects?

sch   cc 1   2   3   4

1   0

2

3

- depends on context
- one rule of thumb:
- which is of interest: mean or variance?

4

- nested factors are often modelled with random effects

5

- are levels of factor in one group same as levels of factors in another group?

Example: several nested levels of variation p.450

- response: success of a surgical procedure ("measured on ...")
- patients surgeons hospitals
- $y_{hsp} =$ $\mu + b_h + e_{hs} + \varepsilon_{hsp}$

$E(y_{hsp}) = \mu$ all random

$\mu + b_h$ if $e, \varepsilon \sim (\quad)$, $\sim$

$\mu + b_h + e_{hs}$ $\varepsilon \sim$

... fixed or random?

- $y_{hsp} = \mu + b_h + e_{hs} + \epsilon_{hsp}$

$p = 1, \dots, P$

$s = \dots$

$h =$

- $E(y_{hsp}) =$

est σ^2, σ_e^2, σ_b^2 under suitable ass$^{=}$

$$\tilde{\sigma}^2 = \frac{\sum_{i=1}^{n} (x_i - \bar{x})^2}{n-1} \qquad var(\tilde{\sigma}^2) = \frac{2\sigma^4}{n-1}$$

- See Table 9.23 and columns of expected mean squares

Example: randomized blocks with replications

- $y_{tbr} = \mu + \alpha_t + \beta_b + (\alpha\beta)tb + \epsilon_{tbr}$

$\bar{r}$ blocks

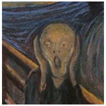
E(MS)

- ANOVA:

		fixed	rand.
trt	$T-1$	$\sigma^2 + \sum \alpha_t^2 \frac{BR}{T-1}$	$\sigma^2 + R\sigma^2_{AB} + (1)$
blk	$B-1$	$\sigma^2 + \sum \beta_b^2 \frac{TR}{B-1}$	$\sigma^2 + RT\sigma^2_B$
t x b	$(T-1)(B-1)$	$\sigma^2 + \sum_{t.b} (\alpha\beta)^2_{tb} \cdot \frac{R}{(T-1)(B-1)}$	$\sigma^2 + R\sigma^2_{AB}$
within cell	$TB(R-1)$	σ^2	σ^2

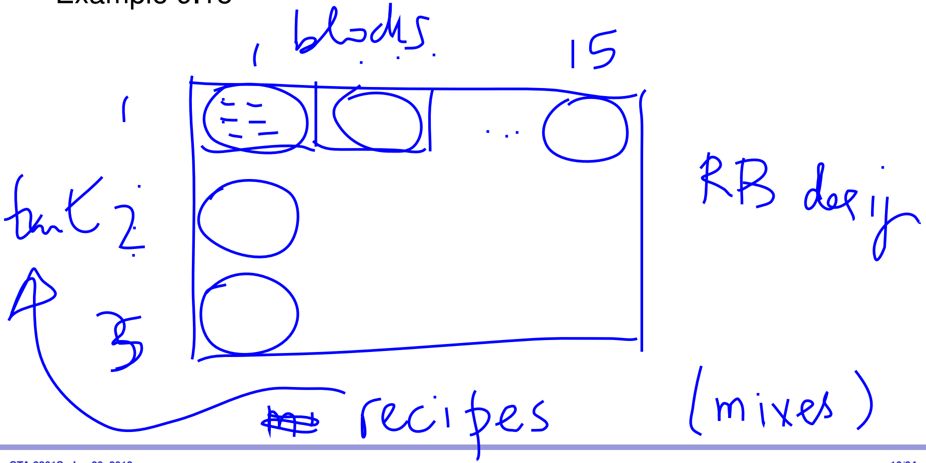
$\beta \sim (0, \sigma^2_B)$ $(\alpha\beta) \sim (0, \sigma^2_{AB})$

MS_t / MS_{AB}

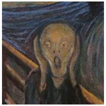


Split plot experiments

One design, often RB, at 'whole plot' level
Second design, often with random effects, at subplot level
Example 9.15



Split plot experiments



p 455

One design, often RB, at 'whole plot' level
Second design, often with random effects, at subplot level
Example 9.15

whole plots is 1 expt

sub plots is another

> 2 levels of var =

Anova	1	F	2
rec	MS_r	MS_r / MS_{mr}	temp
mix	MS_m		temp×rec
r × mix	MS_{mr}		residual

Linear mixed effects models

v. general

- $y = X\beta + Zb + \epsilon$

  y: $n \times 1$ X: $n \times p$ β: $p \times 1$ fixed eff.

  $Z_{n \times q}$ ← design

  $b_{q \times 1}$ ← random

- Assumptions:

- $y \mid b \sim$

- $y \sim$

- See Example 9.16 – note imbalance

Linear mixed effects models

- $y = X\beta + Zb + \epsilon$

  $n \times 1$

  $\underset{\sim}{\epsilon} \perp \underline{b}$

- Assumptions: $\underset{\sim}{\epsilon} \sim N(0, \Omega)$; $\underset{\sim}{b} \sim N_q(0, \Omega_b)$

  $(\sim N(0, \sigma^2 I))$

- $y \mid b \sim$

- $y \sim$

- See Example 9.16 – note imbalance

Linear mixed effects models

- $y = X\beta + Zb + \epsilon$ app$^{\text{n}}$ - multi level model
  - repeated measures
  - longitudinal data
  - components of var-
- Assumptions:
- $y \mid b \sim N(X\beta + Zb, \Omega)$ (ntbc)
- $y \sim N(X\beta, Z\Omega_b Z^T + \Omega)$
- See Example 9.16 – note imbalance

... Example 9.16

Example 9.16 (Longitudinal data) A short longitudinal study has one individual allocated to the treatment and two to the control, with observations

$$y_{1j} = \beta_0 + b_1 + \varepsilon_{1j}, \quad y_{21} = \beta_0 + b_2 + \varepsilon_{21}, \quad y_{3j} = \beta_0 + \beta_1 + b_3 + \varepsilon_{3j}, \quad j = 1, 2.$$

Thus there are two measurements on the first and third individuals, and just one on the second. The b_j represent variation among individuals and the ε_{ij} variation between measures on the same individuals. If the b's and ε's are all mutually independent with variances σ_b^2 and σ^2, then

$$\begin{pmatrix} y_{11} \\ y_{12} \\ y_{21} \\ y_{31} \\ y_{32} \end{pmatrix} = \begin{pmatrix} 1 & 0 \\ 1 & 0 \\ 1 & 0 \\ 1 & 1 \\ 1 & 1 \end{pmatrix} \begin{pmatrix} \beta_0 \\ \beta_1 \end{pmatrix} + \begin{pmatrix} 1 & 0 & 0 \\ 1 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} b_1 \\ b_2 \\ b_3 \end{pmatrix} + \begin{pmatrix} \varepsilon_{11} \\ \varepsilon_{12} \\ \varepsilon_{21} \\ \varepsilon_{31} \\ \varepsilon_{32} \end{pmatrix},$$

and this fits into formulation (9.12) with $\Omega_b = \sigma_b^2 I_3$ and $\Omega = \sigma^2 I_5$. Here ψ comprises the scalar σ_b^2/σ^2, and hence the variance matrix

$$\Omega + Z\Omega_b Z^{\mathrm{T}} = \begin{pmatrix} \sigma_b^2 + \sigma^2 & \sigma_b^2 & 0 & 0 & 0 \\ \sigma_b^2 & \sigma_b^2 + \sigma^2 & 0 & 0 & 0 \\ 0 & 0 & \sigma_b^2 + \sigma^2 & 0 & 0 \\ 0 & 0 & 0 & \sigma_b^2 + \sigma^2 & \sigma_b^2 \\ 0 & 0 & 0 & \sigma_b^2 & \sigma_b^2 + \sigma^2 \end{pmatrix}$$

may be written as

Inference

- $y \sim N(X\beta, \underline{Z\Omega_b Z^T + \Omega}) = N(X\beta, \sigma^2\Upsilon^{-1})$ ψ par. in Υ

- log-likelihood function

$$\ell(\beta, \sigma^2, \psi) = -\frac{n}{2}\log\sigma^2 + \frac{1}{2}\log|\Upsilon| - \frac{1}{2\sigma^2}(y - X\beta)^T\Upsilon(y - X\beta)$$

- constrained m.l.e.'s

$$\hat\beta_\psi, \hat\sigma^2_\psi : \hat\beta_\psi = (X^T\Upsilon X)^{-1}X^T\Upsilon y \quad \hat\sigma^2_\psi = \frac{1}{n}(y - X\hat\beta_\psi)^T\Upsilon(y - X\hat\beta)$$

- REML

$$\ell(\hat\beta_\psi, \hat\sigma^2_\psi, \psi) =: \ell_p(\psi)$$

Inference

- $y \sim N(X\beta, Z\Omega_b Z^T + \Omega) = N(X\beta, \sigma^2\Upsilon^{-1})$
- log-likelihood function $-\frac{1}{2\sigma^2}(y - X\beta)^T \Upsilon (y - X\beta)$
$-\frac{n}{2}\log\sigma^2 + \frac{1}{2}\log|\Upsilon|$
- constrained m.l.e.'s
- REML

$$\ell_{REML}(\beta, \sigma^2, \psi) = \ell(\beta, \sigma^2, \psi) + \frac{p}{2}\log\sigma^2$$
$$-\frac{1}{2}\log|X^T \Upsilon X|$$

Inference

- $y \sim N(X\beta, Z\Omega_b Z^T + \Omega) = N(X\beta, \sigma^2\Upsilon^{-1})$
- log-likelihood function

$$\tilde{\sigma}^2_{REML} = \frac{(y - X\hat{\beta})\hat{\Upsilon}(y - X\hat{\beta})}{n-p}$$

- constrained m.l.e.'s
- REML

restr'd max. lik.

fixes the ~~problems~~ divisor in max lik

est. of σ^2, & ψ

Inference

- $y \sim N(X\beta, Z\Omega_b Z^T + \Omega) = N(X\beta, \sigma^2\Upsilon^{-1})$ $\hat{\beta}$ $\hat{\sigma}^2$ $\hat{\Omega}$ $\hat{\Omega}_b$ ...
- prediction:

$$E(b \mid y) = (Z^T\Omega^{-1}Z + \Omega_b^{-1})^{-1}Z^T\Omega^{-1}(y - X\beta)$$

$$\text{var}(b \mid y) = (Z^T\Omega^{-1}Z + \Omega_b^{-1})^{-1}$$

- $\tilde{b} = (Z^T\hat{\Omega}^{-1}Z + \hat{\Omega}_b \ldots)^{-1}Z^T\hat{\Omega}^{-1}(y - X\hat{\beta})$

Example 9.17

$\sim (0, \sigma_b^2)$

- $y_{ij} = \mu + b_i + \epsilon_{ij}, \quad j = 1, \dots n_i \; ; i = 1, \dots, q$ 1-way

- $\Omega = \sigma^2 I_n \quad \Omega_b = \sigma_b^2 \cdot I_q$ $\underset{n \times q}{Z} \underset{q \times 1}{\tilde{b}}$

- $X = \begin{pmatrix} 1 \\ \vdots \\ \vdots \\ 1 \end{pmatrix}$ $p = 1$ $\quad Z = \begin{array}{c} n_1 \left\{ \begin{array}{c} \\ \\ \\ \end{array} \right. \\ \\ \\ \end{array} \begin{pmatrix} 1 & & \\ 1 & 0 & 0 \\ \vdots & & \\ 1 & & \\ & 1 & \\ 0 & 1 & 0 \\ & \vdots & \\ & & 1 \\ & & \vdots \\ & & 1 \end{pmatrix}$

- $\tilde{b}_i = \dfrac{\bar{y}_{i.} - \bar{y}_{..}}{\left(1 + \dfrac{\hat{\sigma}^2}{n_i \hat{\sigma}_b^2}\right)} \qquad \widehat{\mathrm{var}}(\tilde{b}_i) = \dfrac{1}{\dfrac{1}{\hat{\sigma}_b^2} + \dfrac{n_i}{\hat{\sigma}^2}}$

Example 9.18

- repeated measurements on the 30 individuals, at 5 time points
- might expect that regression relationship against time is similar for each individual, subject to random variation
- model $y_{jt} = \beta_0 + b_{j0} + (\beta_1 + b_{j1})x_{jt} + \epsilon_{jt}, \quad t = 1, \ldots, 5$
- x_{jt} takes values $0, 1, 2, 3, 4$ for $t = 1, 2, 3, 4, 5$
- same for each j
- `data(rat.growth, library="SMPracticals")`
- $(b_{j0}, b_{j1}) \dot\sim N_2(0, \Omega_b), \quad \epsilon_{jt} \dot\sim N(0, \sigma^2)$ independent
- two fixed parameters β_0, β_1
- four variance/covariance parameters: $\sigma^2_{b0}, \sigma^2_{b1}, \text{cov}(b_0, b_1), \sigma^2$

... Example 9.18

- maximum likelihood estimates of fixed effects: $\hat{\beta}_0 = 156.05(2.16), \hat{\beta}_1 = 43.27(0.73)$
- weight in week 1 is estimated to be about 156 units, and average increase per week estimated to be 43.27
- there is large variability between rats: estimated standard deviation of 10.93 for intercept, 3.53 for slope
- there is little correlation between the intercepts and slopes
- ```
   library(MASS)  # this is included the standard R distribution
  library(SMPracticals) # this has various data sets from Davison's book
  library(ellipse) # but I got an error the first time and had to download an additional p
  library(SMPracticals) # and now it works
  data(rat.growth) # for Example 9.18
  rat.growth[1:10,]  # to see what it looks like, and to see variable names
  with(rat.growth, plot( y ~ week , type="l"))
  separate.lm = lm(y ~ week + factor(rat)+ week:factor(rat), data = rat.growth) # fit sepa
  rat.mixed = lmer(y ~ week + (week|rat), data = rat.growth) # REML is the default
  summary(rat.mixed)  # compare Table 9.28
  ```
  ```

This week's study

Enlarge this image

No clear evidence Tamiflu works, study finds

CARLY WEEKS

From Thursday's Globe and Mail
Published Wednesday, Jan. 18, 2012 5:41PM EST
Last updated Wednesday, Jan. 18, 2012 5:47PM EST

Wiley Online Library Home

THE COCHRANE LIBRARY

Independent high-quality evidence for health care decision making

from The Cochrane Collaboration

LOGIN
Enter e-m
Enter pas
REME

Home > Evidence Based Medicine > Evidence-Based Medicine > Database Home > Abstract

DATABASE TOOLS
- Save to My Profile
- Recommend to Your Librarian

DATABASE MENU
- Database Home

FIND ARTICLES
- A-Z
- By Topic
- New Reviews

Intervention Review

Neuraminidase inhibitors for preventing and treating influenza in healthy adults and children

Tom Jefferson[1,*], Mark A Jones[2], Peter Doshi[3], Chris B Del Mar[4], Carl J Heneghan[5], Rokuro Hama[6], Matthew J Thompson[5]

Database Title
The Cochrane Library

Editorial Group: Cochrane Acute Respiratory Infections Group

Published Online: 18 JAN 2012