Outline

1. One-Way ANOVA--- Fixed Effects
   ① One-Way ANOVA
   ② Post-hoc Test After One-Way ANOVA
2. Two-Way ANOVA
3. The Kruskal-Wallis Test
4. One-Way ANOVA--- Random Eeffects
5. Meta Analysis
6. Mixed Model

Random-Effects One-way ANOVA

Example: Endocrinology

- Source: blood samples were obtained from a subset of nurses and serum levels of various hormones were recorded.

- Each blood sample was split into two equal aliquots, which were sent in a blinded fashion to one laboratory for analysis.

- The goal of the study was to assess how much variation in the analyses was attributable to between-person vs. within-person variation (or estimate the overall mean of the population).

Reproducibility data for plasma estradiol (pg/mL), Nurses' Health Study

Subject	Replicate 1	Replicate 2	Absolute value of difference between replicates	Mean value
1	25.5	30.4	4.9	27.95
2	11.1	15.0	3.9	13.05
3	8.0	8.1	0.1	8.05
4	20.7	16.9	3.8	18.80
5	5.8	8.4	2.6	7.10

ANOVA model

- One-way ANOVA

$$y_{ij} = \mu + \alpha_i + \varepsilon_{ij}$$

- y_{ij}: jth replicate for the ith subject
- α_i: between-subject variability;
  - $\alpha_i \sim N(0, {\sigma_A}^2)$, σ_A^2 is the population variance effect
- e_{ij}: within-subject variability;
  - $e_{ij} \sim N(0, \sigma^2)$, σ^2 is the measurement variance
  - independent of α_i and any of the other e_{ij}
- Hierarchical structure

Testing in One-way ANOVA

- An important goal in the random-effects ANOVA is to test

$$H_0: \sigma_A^2 = 0 \quad vs \quad H_1: \sigma_A^2 > 0$$

- Under H_0, there is no underlying between-subject variation; all variation seen between individual subjects is attributable to within-person variation(or “noise”).

- Under H_1 there is a true underlying difference among means for individual subjects.

Contd.

- Can we use the same F statistics ($\mathrm{MSB}/\mathrm{MSW}$)for fixed effect models?
- Easy to check

$$E(\mathrm{MSW}) = \sigma^2$$
$$E(\mathrm{MSB}) = \sigma^2 + n_0\sigma_A^2$$

where

$$n_0 = \begin{cases} n_1 = \cdots = n_k, & \text{balanced case} \\ \left(\sum n_l - \sum n_l^2 / \sum n_l\right)/(k-1), & \text{unbalanced case} \end{cases}$$

- If H_0 is true($\sigma_A^2 = 0$), then F will be small;
- If H_1 is true($\sigma_A^2 > 0$), then F will be large; ==> F test will work

Inference for variance components

1. the unbiased estimate of σ^2 is given by MSW

2. And since

$$E\left(\frac{\text{MSB}-\text{MSW}}{n_0}\right) = \frac{\sigma^2 + n_0\sigma_A^2 - \sigma^2}{n_0} = \sigma_A^2$$

$\Rightarrow$an unbiased estimate of σ_A^2 is given by $\hat{\sigma}_A^2 = (\text{MSB} - \text{MSW})/n_0$.

$\Rightarrow$Problem: this estimate can be negative. If it is, set to 0.

Coefficient of variation

- Another parameter that is often of interest in reproducibility studies, such as the Endocrinology example, is the Coefficient of variation(CV).

- Generally speaking, *CV*s of <20% are desirable, whereas *CV*s of >30% are undesirable.

$$CV = 100\% \times \frac{\text{within-person standard deviation}}{\text{within-person mean}}$$

- based on the assumption that the standard deviation is independent of the mean value. If not, see next slide…

Estimation of the CV in Reproducibility Studies

- Suppose we have *k* subjects enrolled in a reproducibility study where there are n_i replicates for the *i*th subject, $i = 1, \ldots, k$. To estimate the *CV*,

① Apply the ln transformation to each of the values.
② Estimate the between- and within-subject variance components using a one-way random-effects model ANOVA.
③ The *CV* in the *original scale* is estimated by

$$100\% \times \sqrt{\mathrm{MSW}}$$

The Intraclass Correlation

- In some instances, correlation between variables that are not readily distinguishable from each other.

- In Endocrinology example, the replicate-sample determinations are indistinguishable from each other, since each plasma sample was split into two halves at random.

- Thus a special type of correlation is needed between repeated measures on the same subject, called an intraclass correlation coefficient (the magnitude of ${\sigma_A}^2$ comparing to σ^2).

The Intraclass Correlation Coefficient

Definition: the correlation between two replicates from the same subject—i.e., between y_{ij} and y_{il} where $j \neq l$ and $1 \leq j \leq n_i, 1 \leq l \leq n_i$

denoted by ρ_I

Interpretation of Intraclass Correlation

- $\rho_I < 0.4$ indicates poor reproducibility
- $0.4 \leq \rho_I < 0.75$ indicates fair to good reproducibility
- $\rho_I \geq 0.75$ indicates excellent reproducibility

Contd.

- If y_{ij} follows a one-way random-effects ANOVA model,

$$\rho_I = \sigma_A^2/(\sigma_A^2 + \sigma^2)$$

  - 1 indicating perfect reproducibility
  - 0 indicating no reproducibility at all

- A point estimate of ρ_I is given by

$$\hat{\rho}_I = \max[\hat{\sigma}_A^2/(\hat{\sigma}_A^2 + \hat{\sigma}^2), 0]$$

Back to Endocrinology

```
> rep <- c(25.5, 30.4, 11.1, 15.0, 8.0, 8.1,
+ 20.7, 16.9, 5.8, 8.4)
> rep <- log(rep)
> sub <- factor(rep(1:5,c(2,2,2,2,2),
+ labels = c("Sub1","Sub2","Sub3","Sub4","Sub5")))
> plasma.lm <- lm(rep ~ sub)
> print(anova(plasma.lm))
Analysis of Variance Table

Response: rep
          Df  Sum Sq Mean Sq F value  Pr(>F)
Sub        4 2.65775 0.66444  22.146 0.002221 **
Residuals  5 0.15001 0.03000
---
```

significant differences among the underlying mean ln(plasma estradiol) values for different subjects.

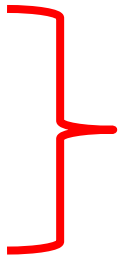
Contd.

```
> # Estimate sigma^2_A
> MSB <- anova(plasma.lm)$Mean[1]
> MSW <- anova(plasma.lm)$Mean[2]

> n <- 2
> sigma.A.2 <- (MSB - MSW)/n

> print(sigma.A.2)
[1] 0.3172171

> print(MSW)
[1] 0.03000244
```



Thus the between-person variance is about 10 times as large as the within-person variance, which indicates good reproducibility.

Contd.

```
> ## estimation of the coefficent of variation
> # the correct way
> sqrt(MSW) # in the original scale
[1] 0.1732121

> # incorrect way
> sqrt(MSW)/mean(rep) # CV for ln scale
[1] 0.06744243

> aov.raw <- anova(lm(repl ~ sub))
> sqrt(aov.raw$Mean[2])/mean(repl) # for raw scale
 # not appropriate because the standard deviation
      #appears to increase as the mean increases
[1] 0.1639928
```

Contd.

```
> ## the Intraclass Correlation Coefficient
> sigma.A.2/(sigma.A.2+MSW)
# p_I
[1] 0.9135923
```

Thus there is excellent reproducibility for ln(plasma estradiol).

Meta-Analysis

REF: Prognostic immune-related gene models for breast cancer: a pooled analysis, Oncotargets and Therapy, 2017

Motivation

- We have examined methods of analysis for a single study.
- However, often more than one investigation is performed to study a particular research question, often by different research groups.
- Many of the studies are small and individually are likely to yield nonsignificant results.
- The question is, What is the appropriate way to combine evidence across all the studies so as to reduce sampling error and increase the power of the investigation and, in some instances, to resolve the inconsistencies among the study results?

Meta-analysis

	True value of targeted parameter(unknwon)	Results from the i-th study	
		Estimation	variance
Study 1	θ_1	y_1	${\sigma_1}^2$
Study 2	θ_2	y_2	${\sigma_2}^2$
…	…	…	…
Study K	θ_K	y_K	${\sigma_K}^2$

Known from the literature

$$y_i = \theta_i + \epsilon_i,$$

- $\epsilon_i \sim N(0, {\sigma_i}^2)$, ϵ_i: within-study variation
- Fixed-effect model, $\theta_1 = \theta_2 = \ldots = \theta_K = \theta$,

$$y_i = \theta + \epsilon_i \sim N(\theta, {\sigma_i}^2)$$

- Random-effect model, θ_i i.i.d.$\sim N(\theta, {\sigma_A}^2)$,

$$y_i = \theta_i + \epsilon_i \sim N\left(\theta, {\sigma_A}^2 + {\sigma_i}^2\right),$$

where ${\sigma_A}^2$ is unknown and needs to be estimated.

Meta-analysis: Estimation

- Fixed-effect model, $y_i \sim N(\theta, {\sigma_i}^2)$

MLE of θ: $\hat{\theta} = \frac{\sum_{i=1}^{K} w_i y_i}{\sum_{i=1}^{K} w_i}$,

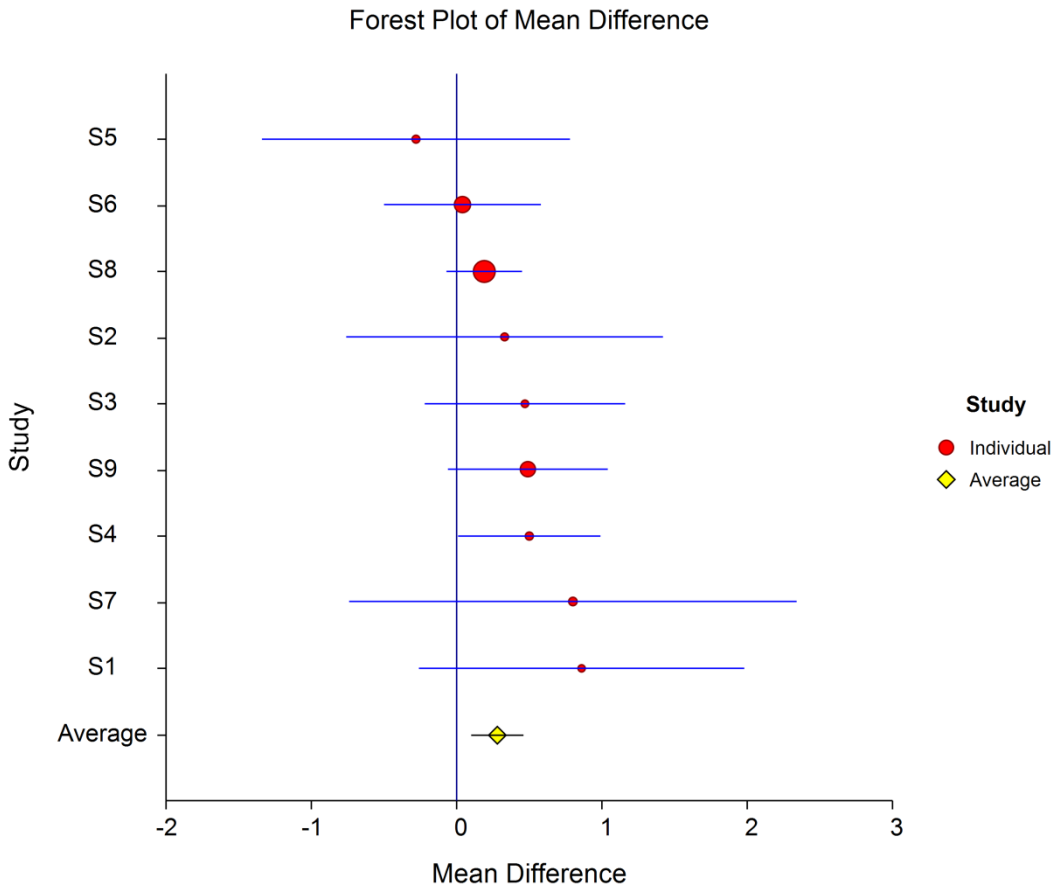
where $w_i = \frac{1}{{\sigma_i}^2}$

- Random-effect model

MLE of θ: $\hat{\theta} = \frac{\sum_{i=1}^{K} {w_i}^* y_i}{\sum_{i=1}^{K} {w_i}^*}$,

where ${w_i}^* = \frac{1}{\widehat{\sigma_A}^2 + {\sigma_i}^2}$ and

$\widehat{\sigma_A}^2$ is the MLE of ${\sigma_A}^2$.



Test of Homogeneity

- Some investigators feel the (random effect) procedure should only be used if there is no significant heterogeneity among the k study-specific θ_is

- So to test

  $H_0: \theta_1 = \cdots = \theta_k = \theta$ vs H_1: at least two θ_is are different
- The test statistic

$$Q_w = \sum_{i=1}^{k} w_i (y_i - \hat{\theta})^2$$

- $Q_w \sim \chi^2_{k-1}$ under H_0

Mixed model

Mixed model

- Two-way ANOVA model-balanced design with one fixed effect and one random effect

$$y_{ijk} = \mu + \alpha_i + \beta_j + (\alpha\beta)_{ij} + e_{ijk}$$

- y_{ijk}: kth replicate for the ith subject at the time point j
- α_i: a random effect; $\alpha_i \sim N\left(0, {\sigma_A}^2\right)$
- β_j: a fixed effect
- $(\alpha\beta)_{ij}$: interaction effect
- e_{ij}: error term; $e_{ij} \sim N(0, \sigma^2)$

Example: Ophthalmology

- A study was performed among dry eye patients comparing an active drug with placebo.

- Measurement: tear break-up time (TBUT, the time it takes for a tear to dissolve); a smaller number indicates worse symptoms (more tearing).

- The TBUT was measured at baseline (bas), immediately after drop instillation (im), and at 5 minutes (pst5), 10 minutes (pst10), and 15 minutes (pst15) after instillation.

Contd.

Endocrinology

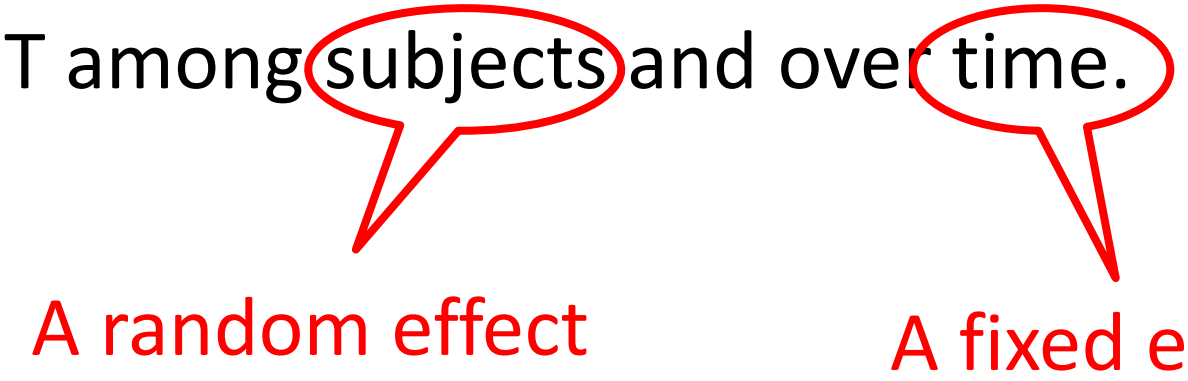
Table 12.24 Tear break-up time over 15 minutes among 14 dry eye patients in the placebo group of a clinical trial after a 3 second nonblink period

	od bas 1	od bas 2	od im 1	od im 2	od pst5 1	od pst5 2	od pst10 1	od pst10 2	od pst15 1	od pst15 2
1	5.44	4.59	4.34	5.31	4.81	6.53	6.00	4.63	6.47	7.03
2	3.28	3.00	10.87	19.06	13.34	12.31	10.34	9.71	5.81	7.25
3	3.18	2.43	14.78	16.28	12.53	16.84	5.53	6.68	6.78	6.43
4	2.47	1.40	11.12	6.44	8.46	3.84	1.93	2.46	2.62	3.21
5	4.40	4.90	12.93	14.84	7.43	9.78	5.93	6.81	6.28	7.65
6	4.93	4.87	10.56	11.71	4.53	5.50	5.37	3.75	4.56	3.31
7	7.21	6.15	14.34	15.50	10.56	10.87	10.43	8.57	3.15	2.78
8	4.93	4.15	15.31	14.00	7.51	6.59	4.01	3.59	3.90	3.62
9	3.18	2.84	6.90	5.75	8.09	7.03	5.31	6.81	4.09	4.06
10	3.47	1.37	6.91	3.57	7.06	5.09	1.53	1.06	2.96	1.34
11	9.46	8.50	12.03	10.75	17.03	14.93	12.31	14.62	13.06	15.09
12	3.03	3.12	7.12	7.19	5.68	4.32	2.41	3.10	4.47	4.25
13	2.47	2.62	18.97	10.60	2.06	2.66	1.87	2.91	2.22	2.40
14	2.66	2.32	12.97	14.81	5.03	3.03	2.35	1.31	1.32	1.19

Contd.

Endocrinology

- We wish to test whether there are differences in TBUT among subjects and over time.



- This type of design is called a mixed model

Analysis of Mixed model

Table 12.25 Analysis of two-way ANOVA table with one fixed effect and one random effect

Effect	SS	df	MS	*F* Stat	df
Row (random)	$RSS = \sum_{i=1}^{r} y_{i.}^2 / (nc) - y_{...}^2 / (nrc)$	$r-1$	$RMS = RSS/(r-1)$	RMS / EMS	$r-1, rc(n-1)$
Column (fixed)	$CSS = \sum_{j=1}^{c} y_{.j}^2 / (nr) - y_{...}^2 / (nrc)$	$c-1$	$CMS = CSS/(c-1)$	CMS / IMS	$c-1, (r-1)(c-1)$
Interaction	$ISS = \sum_{i=1}^{r}\sum_{j=1}^{c} y_{ij.}^2 / n - RSS - CSS - y_{...}^2 / (nrc)$	$(r-1)(c-1)$	$IMS = ISS/[(r-1)(c-1)]$	IMS / EMS	$(r-1)(c-1), rc(n-1)$
Error	$ESS = \sum_{i=1}^{r}\sum_{j=1}^{c}(y_{ijk} - \bar{y}_{ij})^2$	$rc(n-1)$	$EMS = ESS/[rc(n-1)]$		

General two way ANOVA

Table 12.26 Computation of the F statistics for tests of significance in a two-factor ANOVA with replication

Hypothesized effect	Model I (factors A and B both fixed)	Model II (factors A and B both random)	Model III (factor A random; factor B fixed)
Factor A	$\frac{\text{factor } A \text{ MS}}{\text{error } MS}$	$\frac{\text{factor } A \text{ MS}}{A \times B \text{ MS}}$	$\frac{\text{factor } A \text{ MS}}{\text{error MS}}$
Factor B	$\frac{\text{factor } B \text{ MS}}{\text{error MS}}$	$\frac{\text{factor } B \text{ MS}}{A \times B \text{ MS}}$	$\frac{\text{factor } B \text{ MS}}{A \times B \text{ MS}}$
$A \times B$ interaction	$\frac{A \times B \text{ MS}}{\text{error MS}}$	$\frac{A \times B \text{ MS}}{\text{error MS}}$	$\frac{A \times B \text{ MS}}{\text{error MS}}$

Back to:
Ophthalmology

```
> # read data into R

> tear <- read.table("tear.txt",header= F)
> tea <- as.matrix(tear[,2:11])
> tear <- read.table("tear.txt",header= F)
> tea <- as.matrix(tear[,2:11])
> tear.m <- as.vector(tea)
> time <- factor(rep(1:5,rep(14*2,5),
+ labels = paste("time.",1:5,sep="")))
> sub <- factor(rep(1:14,10,label =
+         paste("sub.",1:14,sep="")))
> tear.lm <- lm(tear.m ~ sub+time+sub*time)
```

R codes

```
> summary(tear.aov <- aov(tear.lm))
            Df Sum Sq Mean Sq F value    Pr(>F)
sub         13 897.58  69.045  32.271 < 2.2e-16 ***
time         4 964.64 241.159 112.715 < 2.2e-16 ***
sub:time    52 655.52  12.606   5.892 9.696e-12 ***
Residuals   70 149.77   2.140
--- Signif. codes: 0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05
‘.’ 0.1 ‘ ’ 1
```

Another method for multiple comparison , like FDR and Bonferroni

```
> TukeyHSD(tear.aov, "time")
  Tukey multiple comparisons of means
    95% family-wise confidence level

Fit: aov(formula = tear.lm)
```

R codes

```
$time
           diff        lwr        upr     p adj
2-1   7.2353571  6.1406971  8.3300171 0.0000000
3-1   3.9667857  2.8721257  5.0614457 0.0000000
4-1   1.3914286  0.2967686  2.4860886 0.0058827
5-1   0.8903571 -0.2043029  1.9850171 0.1645008
3-2  -3.2685714 -4.3632314 -2.1739114 0.0000000
4-2  -5.8439286 -6.9385886 -4.7492686 0.0000000
5-2  -6.3450000 -7.4396600 -5.2503400 0.0000000
4-3  -2.5753571 -3.6700171 -1.4806971 0.0000001
5-3  -3.0764286 -4.1710886 -1.9817686 0.0000000
5-4  -0.5010714 -1.5957314  0.5935886 0.7031749
```