

## Course 02441: Applied Statistics and Statistical Software

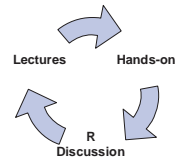
Lasse Engbo Christiansen, IMM

Course 02441 - Applied statistics and statistical software

1

## Today's program: outline

- Analysis of Variance (ANOVA)
- Hands-on exercises



Course 02441 - Applied statistics and statistical software

2

## Methods covered in the course

- Descriptive statistics Day I
- Comparing treatment means (t-test and non-parametric tests) Day I
- Multiple regression analysis Day II
- Analysis of variance Day III
- Analysis of proportions and counts Day IV
- The general linear model Day V

Course 02441 - Applied statistics and statistical software

3

## One-way ANOVA

- We consider a model on the form

$$Y_{ij} = \alpha_i + \varepsilon_{ij}$$

Course 02441 - Applied statistics and statistical software

4

## One-way ANOVA

We consider the model:

$$Y_{ij} = \alpha_i + \varepsilon_{ij}, \quad i = 1..k$$

The total variation, SST, may be partitioned into:

$$SST = SS(Tr) + SS(error)$$

Course 02441 - Applied statistics and statistical software

5

## One-way ANOVA

$$SST = SS(Tr) + SS(error)$$

The test statistics,  $F$ , is:

$$F = \frac{SS(Tr)/(k-1)}{SS(error)/(N-k)} \sim F(k-1, N-k)$$

Course 02441 - Applied statistics and statistical software

6

## Assumptions for ANOVA

- The observations within each group are normally distributed
- The variances in each group are equal
- Non-parametric methods exist when departure from normality
- Transformation of data may be useful when variances are not equal

Course 02441 - Applied statistics and statistical software

7

## Example

- The 'folate' data frame has 22 rows and 2 columns. It contains data on red cell folate levels in patients receiving three different methods of ventilation during anesthesia.
- ventilation a factor with levels:
  - 'N2O+O2,24h': 50% nitrous oxide and 50% oxygen, continuously for 24~hours;
  - 'N2O+O2,op': 50% nitrous oxide and 50% oxygen, only during operation;
  - 'O2,24h': no nitrous oxide, but 35-50% oxygen for 24~hours.

Course 02441 - Applied statistics and statistical software

8

## Two-way ANOVA

- We consider a model on the form

$$Y_{ijk} = \alpha_i + \beta_j + \alpha\beta_{ij} + \varepsilon_{ijk}$$

Course 02441 - Applied statistics and statistical software

9

## Two-way ANOVA

Corresponding to the model

$$Y_{ijk} = \alpha_i + \beta_j + \alpha\beta_{ij} + \varepsilon_{ijk}$$

The total variation, SST, may be partitioned into:

$$SST = SS(Tr) + SS(BI) + SS(ab) + SS(error)$$

Course 02441 - Applied statistics and statistical software

10

## Example

- The 'heart.rate' data frame has 36 rows and 3 columns. It contains data for nine patients with congestive heart failure before and shortly after administration of enalaprilat, in a balanced two-way layout.
- Variables:
  - hr a numeric vector. Heart rate in beats per minute.
  - subj a factor with levels '1' to '9'.
  - time a factor with levels '0' (before), '30', '60', and '120' (minutes after administration).

Course 02441 - Applied statistics and statistical software

11