A two-way factorial design

Example: How do certain toxic agents affect survival time?

Response variable: survival time

Set up: 48 animals were randomly allocated to the 12 combinations of 3 poisons and 4 treatments, 4 animals to each cell.
$\rightarrow 3 \times 4$ factorial design.

The effects of both poison and treatment have to be considered also keeping in mind that they may interact.
$\rightarrow$ two-way ANOVA

ANOVA model including interaction between treatment and poison

$$\eta_{ti} = \eta + \tau_t + \pi_i + \omega_{ti},$$

where

- η is the overall mean
- τ_t is the treatment effect (mean increment in survival time associated with the treatment t, $t = A, B, C, D$)
- π_i is the poison effect (mean increment in survival time associated with the poison i, $i = I, II, III$)
- ω_{ti} is the interaction effect.

Conclusions based on the two-way ANOVA (Table 8.2, p. 319)

- Treatment and poison have significant effect (at 0.1% level) on survival time.
- No significant interaction effect (p-value 0.11), i.e. difference in survival times between treatments does not depend on the poison.

After some further investigation (Table 8.3a), it seems that

- Poison III results always in the shortest survival times.
- Poison II together with treatment A or C results in smaller values than poison I but not together with treatment B or D

Remark: Model for the observations is

$$y_{tij} = \bar{y} + \tau_t + \pi_i + \omega_{ti} + \epsilon_{tij},$$

where ϵ_{tij} is the random error term.

F-test is based on the assumption that the errors ϵ_{tij} are independent and $N(0, \sigma)$-distributed. Note especially that the variance is assumed to be constant.

In the example, variances are not equal (Table 8.3b).

Two kinds of variance inhomogeneity

- inherent inhomogeneity: for example, a smaller variance achieved by an experienced person than by an inexperienced person
- transformable inhomogeneity: untransformed observations give rise to an unnecessarily complicated model with non-constant variance and (possibly) unnecessary interaction

Variance stabilizing transformations

When standard deviation σ is a function of the mean η, often one can find a data transformation that has more constant variance than the original data. For example,

1. σ is proportional to η
   $\rightarrow$ $Y = \log y$ would stabilize the variance
2. σ is proportional to η^{α}
   $\rightarrow$ $Y = y^{\lambda}$, where $\lambda = 1 - \alpha$, would stabilize the variance.

   For example, for Poisson distributed data, where $\eta = \sigma^2$, i.e.

   $$\sigma = \sqrt{\eta} = \eta^{0.5}$$

   ($\alpha = \lambda = 0.5$), the transformation $Y = \sqrt{y}$ stabilizes the variance.

How to find α?

1. Take one set of experimental conditions, say treatment A and poison I (this cell is denoted by j below), and assume that under these conditions

$$\sigma_j \propto \eta_j^{\alpha}.$$

2. Then, $\log \sigma_j = \text{constant} + \alpha \log \eta_j$, and $\log \sigma_j$ plotted against $\log \eta_j$ would give a straight line with slope α.
3. In practise, σ_j and η_j would be replaced by their estimates s_j and $\bar{y}_j$, respectively.

Interpretation of the results after the transformation (example continues)

- Effect of poison and treatment on survival time even more significant than before the transformation of the data.
- No significant interaction between poison and treatment: Effects of poison and treatment are approximately additive when measured as rates of failure ($Y = \frac{1}{y}$)
- Variances more equal after the transformation than before.

Some remarks:

- Transformed data and results based on them may be hard to interpret (e.g. log(kg) instead of kg)
- Other reasons to transform the data (other than obtaining constant variance)
  - to make data approximately normally distributed (reduce skewness)
  - to obtain a linear relationship
  - to obtain an additive relationship ($y = a + bx$ instead of $y = ax^b$, needed in ANOVA)