Design and analysis of experiments
Lecture 4

Jakob G. Rasmussen
Department of Mathematics
Aalborg University
Denmark

The first handin

- Hand-in exercise 1 is available at the homepage now.
- Suggested hand-in date: Monday, 7th of October.
- Solo or in groups of 2-3 (I prefer groups!).
- E-mail me a pdf-file with results and figures - remember to make 1 or 2 lines of text to conclude in each part of the exercise.
- Don't send me a text-file with R-commands (you can include the commands used in the answer if you like, but don't expect me to run your code).
- Figures can be exported from R:
  - `pdf("path/filename.pdf")` ,
  - (plotting command(s) here...)
  - `dev.off()`
- Other commands than `pdf()`, such as `postscript()`, for other file formats are also available

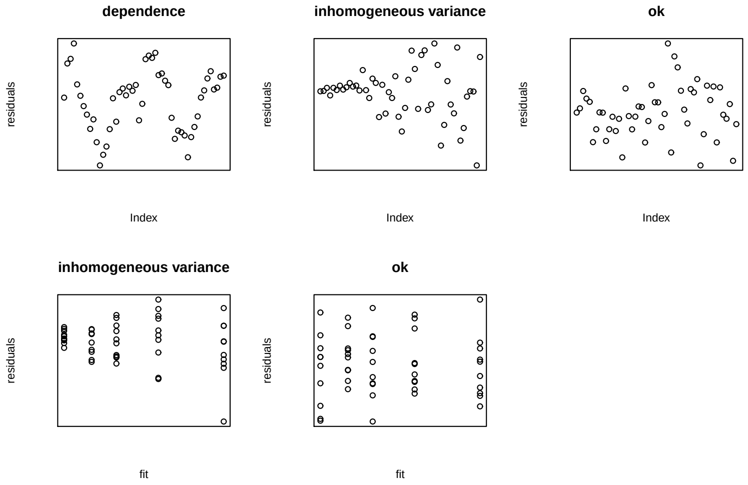
Model checking, residuals

- Most models you see in this course (fx the model in ANOVA) contain some sort of normally distributed theoretical error term ϵ_i describing the deviations from the perfect model.
- For a dataset, we can calculate the practical counterpart $r_i = y_i - \hat{y}_i$ - this is called a residual.
- In ANOVA, we have theoretical errors ϵ_{ij} (stochastic variables), and residuals $r_{ij} = y_{ij} - \bar{y}_{i\bullet}$ (realisations of ϵ_{ij} if the model was perfect).
- If the model fits the data well, the residuals should behave as realisations of the theoretical error terms, i.e. they should fulfill the assumptions of the error terms, typically
  - Normal distribution
  - Homogeneous variance
  - Independence
- Any statistical analysis should contain some sort of model checking to validate the fit of the models used!

Model checking

- Normality:
  - Histogram
  - QQ-plot
  - Kolmogorov-Smirnov test
- Variance homogeneity:
  - Scatter plot, residuals vs time (or index number) - no tendencies of different variances should be visible.
  - Scatter plot, residuals vs fitted values (or group number) - same as above.
  - Bartlett's test or Levene's test
- Independence:
  - The scatter plots are useful here too

Residual plots - examples



Coping with variance inhomogeneity

- A common way of reducing variance inhomogeneity is to transform the data.
- For example, we can work with $\sqrt{y_i}$ or $\log(y_i)$ instead of y_i.
- Page 87 in the book shows a method of finding a suitable transformation.
- Beware:
  - When we are working with tranformed data, the conclusions also apply to the transformed data.
  - The transformation may destroy other properties while fixing the variance inhomogeneity - fx if y_i is approximately normally distributed, then $\log(y_i)$ may well not be.

R

- R-demo, part 1
- Exercise 1

Contrasts

- ANOVA tells us if one or more treatments differ from the rest, but sometimes we need to know more in more details which treatments differs from which other treatments
- For example, in the etch data (p.87) we may test whether whether the two high levels are different or the low power levels are different from the high power levels:

$$H_0 : \mu_3 = \mu_4$$
$$H_1 : \mu_3 \neq \mu_4$$

$$H_0 : \mu_1 + \mu_2 = \mu_3 + \mu_4$$
$$H_1 : \mu_1 + \mu_2 \neq \mu_3 + \mu_4$$

Definition of contrasts

- Model:

$$y_{ij} = \mu_i + \epsilon_{ij} = \mu + \alpha_i + \epsilon_{ij}, \qquad \epsilon_{ij} \sim N(0, \sigma^2) \quad \text{(independent)}$$

- Population contrast (a parameter):

$$\Gamma = \sum_{i=1}^{a} c_i \mu_i = \sum_{i=1}^{a} c_i \alpha_i, \qquad \sum_{i=1}^{a} c_i = 0$$

- Hypothesis:

$$H_0 : \Gamma = 0$$
$$H_1 : \Gamma \neq 0$$

- Example:

$$H_0 : \mu_3 = \mu_4, \qquad (c_1, c_2, c_3, c_4) = (0, 0, 1, -1)$$
$$H_0 : \mu_1 + \mu_2 = \mu_3 + \mu_4, \qquad (c_1, c_2, c_3, c_4) = (1, 1, -1, -1)$$

Sample contrasts

- Sample contrast (an estimate of the population contrast):

$$C = \sum_{i=1}^{a} c_i \bar{y}_{i\bullet} = \sum_{i=1}^{a} c_i (\bar{y}_{i\bullet} - \bar{y}_{\bullet\bullet})$$

- Mean:
  $\mathbb{E}[C] = \mathbb{E}[\sum_{i=1}^{a} c_i \bar{y}_{i\bullet}] = \sum_{i=1}^{a} c_i \mathbb{E}[\bar{y}_{i\bullet}] = \sum_{i=1}^{a} c_i \mu_i = \Gamma$
- Variance:
  $\text{Var}(C) = \text{Var}(\sum_{i=1}^{a} c_i \bar{y}_{i\bullet}) = \sum_{i=1}^{a} c_i^2 \text{Var}(\bar{y}_{i\bullet}) = \sigma^2 \sum_{i=1}^{a} \frac{c_i^2}{n_i}$
- Distribution:

$$C \sim N\left(\Gamma, \sigma^2 \sum_{i=1}^{a} \frac{c_i^2}{n_i}\right) \Rightarrow \frac{C - \Gamma}{\sigma \sqrt{\sum_{i=1}^{a} \frac{c_i^2}{n_i}}} \sim N(0, 1)$$

- Since σ is unknown, we use $s = \sqrt{MS_E}$ instead:

$$t_0 = \frac{\sum c_i \bar{y}_{i\bullet} - \sum c_i \mu_i}{s \sqrt{\sum \frac{c_i^2}{n_i}}} \sim t_{\nu_E}$$

Contrasts, t-test & confidence interval

- Test statistic:
$$t_0 = \frac{\sum c_i \bar{y}_{i\bullet}}{s\sqrt{\sum \frac{c_i^2}{n_i}}} \sim t_{\nu_E}$$
We accept H_0 if $t_0 \in [-t_{\nu_E;\alpha/2}, t_{\nu_E;\alpha/2}]$
- Confidence interval:
$$\Gamma = \left[C \pm t_{\nu_E;\alpha/2}\sqrt{MS_E \sum \frac{c_i^2}{n_i}}\right]$$

Contrast, F-test

- There is also an F-test as an alternative (the t-test and the F-test are equivalent)
- Note that

$$\frac{(C-\Gamma)^2}{\sum \frac{c_i^2}{n_i}} \sim \sigma^2\chi_1^2$$

- Under H_0:

$$SS_C = \frac{C^2}{\sum \frac{c_i^2}{n_i}} = \frac{(\sum c_i\bar{y}_{i\bullet})^2}{\sum \frac{c_i^2}{n_i}} \sim \sigma^2\chi_1^2$$

- Test statistic:

$$F_0 = t_0^2 = \frac{(\sum c_i\bar{y}_{i\bullet})^2}{s^2\sum \frac{c_i^2}{n_i}} = \frac{MS_C}{MS_E} \sim F_{1,\nu_e} \qquad MS_C = \frac{SS_C}{1}$$

Multiple contrasts - orthogonality

- Consider two contrasts:

$$\Gamma = \sum_{i=1}^{a} c_i \mu_i, \qquad \Delta = \sum_{i=1}^{a} d_i \mu_i$$

- These are called orthogonal if $\sum c_i d_i n_i = 0$
- Example of set of orthogonal contrasts (assuming all $n_i = n$):

$$C_1 = \begin{pmatrix} 1 \\ -1 \\ 0 \\ 0 \end{pmatrix}, \qquad C_2 = \begin{pmatrix} 1 \\ 1 \\ -1 \\ -1 \end{pmatrix}, \qquad C_3 = \begin{pmatrix} 0 \\ 0 \\ 1 \\ -1 \end{pmatrix}$$

- Orthogonal constrasts lead to independent tests.
- Note that you can make a set of $a - 1$ orthogonal contrasts.

Using (orthogonal) contrasts

- Preplanned comparisons: the contrasts should be chosen before the experiment.
- It is tempting to look at the data first, and then test particularly large differences, but this is incorrect!
- If we do not know which contrasts we want to test beforehand, we can test them all - Scheffé's Method.

Scheffé's Method - testing all contrasts

- Hypothesis:

$$H_0 : \text{All contrast are } 0$$
$$H_1 : \text{Not all contrast are } 0$$

- There are infinitely many contrasts, so we look at the worst contrast. Under H_0:

$$\max_C \frac{(\sum c_i \bar{y}_{i\bullet})^2}{\sum \frac{c_i^2}{n_i}} \sim \sigma^2 \chi^2_{\nu_A}$$

- Thus

$$\max_C \frac{(\sum c_i \bar{y}_{i\bullet})^2 / \nu_A}{MS_E \sum \frac{c_i^2}{n_i}} \sim F_{\nu_A, \nu_E}$$

- This suggests that a contrast is 0 if

$$\frac{(\sum c_i \bar{y}_{i\bullet})^2 / \nu_A}{MS_E \sum \frac{c_i^2}{n_i}} < F_{\nu_A, \nu_E; \alpha}$$

Pairwise comparisons

- If we are only interested in pairwise comparisons, we have simpler methods than contrasts.
- Three tests:
  - Tukey's test
  - Fisher's least significant difference method
  - Dunnett's test

Tukey's test

- In Tukey's test we compare each pair (i, j) of treatments to see which ones differ significantly.
- Hypothesis:

$$H_0^{ij} : \mu_i = \mu_j$$
$$H_1^{ij} : \mu_i \neq \mu_j$$

- Studentized range statistic:

$$q = \frac{\bar{y}_{\text{max}} - \bar{y}_{\text{min}}}{s/\sqrt{n}}, \qquad s^2 = MS_E$$

Tukey's test

- First calculate the Tukeyspan

$$T_\alpha = \frac{q_\alpha(a, \nu_E)}{\sqrt{2}} \sqrt{MS_E(1/n_i + 1/n_j)}$$

- Then we calculate each difference $\bar{y}_i - \bar{y}_j$
- If $|\bar{y}_i - \bar{y}_j| > T_\alpha$, we reject H_0^{ij}
- Note that the error rate α is experimentwise (in the balanced case, otherwise it less than α), so ALL pairwise comparisons should be made.

Fisher's LSD test

- Fisher's LSD method has the same hypothesis and procedure as Tukeys test.
- Fisherspan:

$$LSD = t_{\nu_E;\alpha/2}\sqrt{MS_E(1/n_i + 1/n_j)}$$

- If $|\bar{y}_i - \bar{y}_j| > LSD$, we reject H_0^{ij}
- In Fisher's LSD method the error rate α is for each individual test of pairs - not the whole experiment.

Dunnett's test - case vs. control

- In Dunnett's test all treatments are compared to a control group
- $i = 1, \ldots, a-1$ are treatments, $i = a$ is control group - only hypotheses $\mu_i - \mu_a = 0$ are considered.
- Dunnettspan:

$$D_\alpha = d_\alpha(a-1, \nu_E)\sqrt{MS_E(1/n_i + 1/n_j)}$$

- The error rate α is for the joined test.

R

- R-demo, part 2
- Exercise 2