Comparison of groups
One-way ANOVA

Ib Skovgaard & Claus Ekstrøm
E-mail: ims@life.ku.dk

Program

- Comparison of two groups: overview
- Comparison of more than two groups: one-way ANOVA
  - Data: antibiotics and decomposition of organic material
  - Statistical model
  - Estimation and confidence intervals
  - Comparison of the groups (test)
  - Pairwise comparisons

Antibiotics and the decomposition of organic material

Data
- Five types antibiotics and a control
- 36 heifers allocated to 6 treatment groups. Feed added antibiotics
- Dung suspended in bags in the ground
- Amount of organic material measured after 8 weeks.
- For spiramycin: only four measurements available

Problem:
- Does the antibiotics affect the decomposition of organic material?

Comparison of two samples: overview

<table>
<thead>
<tr>
<th>x, y indep.?</th>
<th>Same sd.?</th>
<th>R</th>
</tr>
</thead>
<tbody>
<tr>
<td>Example?</td>
<td>Yes</td>
<td>t.test(x,y, var.equal=T)</td>
</tr>
<tr>
<td>Example?</td>
<td>Yes</td>
<td>t.test(x,y)</td>
</tr>
<tr>
<td>Example?</td>
<td>No</td>
<td>t.test(x,y, paired=T)</td>
</tr>
</tbody>
</table>

For comparison of two groups some form of t-test may be used.

What about comparison of three or more groups?
One-way ANOVA!
Group means and standard deviations

<table>
<thead>
<tr>
<th>Type</th>
<th>n</th>
<th>(\bar{y}_j)</th>
<th>s_j</th>
</tr>
</thead>
<tbody>
<tr>
<td>Control</td>
<td>6</td>
<td>2.603</td>
<td>0.119</td>
</tr>
<tr>
<td>(\alpha)-cyperm.</td>
<td>6</td>
<td>2.895</td>
<td>0.117</td>
</tr>
<tr>
<td>Enrofloxacin</td>
<td>6</td>
<td>2.710</td>
<td>0.162</td>
</tr>
<tr>
<td>Fenbendaz.</td>
<td>6</td>
<td>2.833</td>
<td>0.124</td>
</tr>
<tr>
<td>Ivermectin</td>
<td>6</td>
<td>3.002</td>
<td>0.109</td>
</tr>
<tr>
<td>Spiramycin</td>
<td>4</td>
<td>2.855</td>
<td>0.054</td>
</tr>
</tbody>
</table>

Pooled estimate of the standard deviation:

\[ s = \sqrt{\frac{1}{28} (5 \cdot s_1^2 + \cdots + 3 \cdot s_6^2)} = \sqrt{\frac{1}{34} \cdot 6 \sum_{i=1}^{n} (y_i - \bar{y}_{g(i)})^2} = 0.1217 \]

Statistical model

Recall that \(g(i)\) denotes the group for observation \(i\). For example

\[ g(1) = \cdots = g(6) = \text{control}, \quad g(31) = \cdots = g(34) = \text{Spiramycin} \]

\[ g(1) = \cdots = g(6) = 1, \quad g(31) = \cdots = g(34) = 6. \]

Statistical model: \(y_1, \ldots, y_{34}\) are independent and

\[ y_i \sim N(\alpha_{g(i)}, \sigma^2) \]

Parameters: \(\alpha_1, \ldots, \alpha_6\) and \(\sigma\).

Equivalently:

\[ y_i = \alpha_{g(i)} + e_i, \quad e_1, \ldots, e_{34} \sim N(0, \sigma^2) \text{ independent} \]

Estimation and confidence intervals

Statistical model:

\[ y_i = \alpha_{g(i)} + e_i, \quad e_1, \ldots, e_{34} \sim N(0, \sigma^2) \text{ independent} \]

Parameters: \(\alpha_1, \ldots, \alpha_k\) and \(\sigma\). In particular, we are interested in differences, \(\alpha_j - \alpha_l\)!

Estimates and standard errors:

\[ \hat{\alpha}_j = \bar{y}_j; \quad \text{SE}(\hat{\alpha}_j) = s \sqrt{1/n_j} = s/\sqrt{n_j} \]

\[ \hat{\alpha}_j - \hat{\alpha}_l = \bar{y}_j - \bar{y}_l; \quad \text{SE}(\hat{\alpha}_j - \hat{\alpha}_l) = s \sqrt{1/n_j + 1/n_l} \]

\[ \hat{\sigma} = s \]

Confidence intervals from the usual recipe:

\[ \text{estimate} \pm t_{0.975, n-k} \cdot \text{SE(estimate)} \]

NB. The pooled \(s\) is used, also when comparing two groups!
One-way ANOVA in R

Fitting the model:

```r
> model1 <- lm(org~factor(type))
> summary(model1)
```

R chooses a reference group — the first in alphabetic order — and estimates changes compared to that group.

We prefer the control group as the reference group:

```r
> type <- relevel(type, ref="Control")
> model1 <- lm(org~factor(type))
> summary(model1)
```

Output from `summary(model1)`:

| Estimate | Std. Error | t value | Pr(>|t|) |
|----------|------------|---------|----------|
| (Intercept) | 2.60333 | 0.04970 | 52.379 | < 2e-16 *** |
| factor(type)Alfacyp | 0.29167 | 0.07029 | 4.150 | 0.000281 *** |
| factor(type)Enroflox | 0.10667 | 0.07029 | 1.518 | 0.140338 |
| factor(type)Fenbenda | 0.23000 | 0.07029 | 3.272 | 0.002834 ** |
| factor(type)Ivermect | 0.39833 | 0.07029 | 5.667 | 4.5e-06 *** |
| factor(type)Spiramyc | 0.25167 | 0.07858 | 3.202 | 0.003384 ** |

Residual standard error: 0.1217 on 28 degrees of freedom

Interpretations:

- Estimate and CI for $\alpha_{\text{cont}}, \alpha_{\text{Fenb}} - \alpha_{\text{cont}}$ and $\alpha_{\text{Fenb}}$? Estimat for $\sigma$?
- Why are the SE's not the same?

If we prefer to see the group means:

```r
> model2 <- lm(org~factor(type)-1)
> summary(model2)
```

| Estimate | Std. Error | t value | Pr(>|t|) |
|----------|------------|---------|----------|
| factor(type)Control | 2.60333 | 0.04970 | 52.38 | <2e-16 *** |
| factor(type)Alfacyp | 2.89500 | 0.04970 | 58.25 | <2e-16 *** |
| factor(type)Enroflox | 2.71000 | 0.04970 | 54.53 | <2e-16 *** |
| factor(type)Fenbenda | 2.83333 | 0.04970 | 57.01 | <2e-16 *** |
| factor(type)Ivermect | 2.70167 | 0.04970 | 58.25 | <2e-16 *** |
| factor(type)Spiramyc | 2.85500 | 0.06087 | 46.90 | <2e-16 *** |

Residual standard error: 0.1217 on 28 degrees of freedom

Hypothesis. Variation within and between groups

Hypothesis, $H_0: \alpha_1 = \cdots = \alpha_k$.

Alternative, $H_A$: at least two $\alpha$'s are different.

- Variation within groups — points around the lines
  $SS_e = \sum_{i=1}^n (y_i - \bar{y}(g(i)))^2$
- Variation between groups — Lines around the dashed line
  $SS_{\text{grp}} = \sum_{j=1}^k n_j (\bar{y}_j - \bar{y})^2$
- Test statistic
  $F = \frac{MS_{\text{grp}}}{MS_e} = \frac{SS_{\text{grp}}/(k-1)}{SS_e/(n-k)}$
Comparison af alle groupsne

Do not use model2 for this — only model1

> anova(model1)

<table>
<thead>
<tr>
<th>Df</th>
<th>Sum Sq</th>
<th>Mean Sq</th>
<th>F value</th>
<th>Pr(&gt;F)</th>
</tr>
</thead>
<tbody>
<tr>
<td>factor(type)</td>
<td>5</td>
<td>0.59082</td>
<td>0.11816</td>
<td>7.9726</td>
</tr>
<tr>
<td>Residuals</td>
<td>28</td>
<td>0.41500</td>
<td>0.01482</td>
<td></td>
</tr>
</tbody>
</table>

Test statistic

\[
F = \frac{MS_{grp}}{MS_e} = \frac{SS_{grp}/(k-1)}{SS_e/(n-k)}
\]

Large values of \( F \) are in disagreement with the hypothesis. Hence, the \( p \)-value is

\[ p = P(F \geq F_{obs}) = P(F \geq 7.97) = 0.00009 \]

There is overwhelming evidence that the hypothesis is not true.

How did we get the \( p \)-value?

The \( F \)-distribution

If the hypothesis is true, then the \( F \)-test statistic is \( F \)-distributed with \((k-1,n-k)\) degrees of freedom.

\[
p = P(F \geq 7.97) = 0.00009
\]

F-probabilities and quantiles in R:

\[
> pf(7.97, df1=5, df2=28)
\]

\[
[1] 0.9999102
\]

\[
> qf(0.95, df1=5, df2=28)
\]

\[
[1] 2.558128
\]

Pairwise comparisons

Suppose we want to compare the control group (1) with the Fenbendazole group (4): \( \alpha_4 - \alpha_1 \).

Estimate and its standard error:

\[
\hat{\alpha}_4 - \hat{\alpha}_1 = 2.833; \quad SE(\hat{\alpha}_4 - \hat{\alpha}_1) = 0.07029
\]

- Confidence interval for \( \alpha_4 - \alpha_1 \)?
- Test for the hypothesis \( H_0: \alpha_5 = \alpha_4 \)?
- Do all the groups differ significantly from the control group?

LSD-value: least significant difference

A confidence interval for the difference \( \alpha_j - \alpha_l \) is

\[
\hat{\alpha}_j - \hat{\alpha}_l \pm \text{LSD}
\]

where

\[
\text{LSD}_{j,l} = t_{0.975,n-k} \cdot SE(\hat{\alpha}_j - \hat{\alpha}_l) = t_{0.975,n-k} \cdot s \cdot \sqrt{1/n_j + 1/n_l}.
\]

A \( t \)-test for the hypothesis that the difference is zero uses the test statistic

\[
T = \frac{|\alpha_j - \alpha_l|}{\text{SE}(\hat{\alpha}_j - \hat{\alpha}_l)}
\]

which is \( t \)-distributed with \( n-k \) degrees of freedom.

LSD for control and fenbend.: \( 2.048 \cdot 0.1217 \cdot \sqrt{1/6 + 1/6} = 0.144 \)

If all group sizes are the same, then so are the LSD-values:

\[
\text{LSD} = t_{0.975,n-k} \cdot s \cdot \sqrt{2/n'}
\]
Conclusion

Different effects of the different types has been shown with high degree of certainty ($p < 0.0001$)

For all types except Enrofloxacin the amount of organic material is significantly higher than for the control group.

These statements should be supplemented by estimates and confidence intervals for $\alpha$'s and/or for differences to the control group.

Multiple comparisons

Any time we make a test a type I error may occur. The risk depends on the level of significance — often 5%.

One test: risk of type I error: 5%
By $m$ independent tests:

$$1 - 0.95^m$$

Summary: one-way ANOVA

- Statistical model: normal distribution with same SD in the groups; independence
- Estimation: group means and pooled SD
- Confidence interval: $\text{estimate} \pm t_{0.975,n-k} \cdot \text{SE(estimate)}$
- Hypothesis of equal group means tested by $F = \text{MS}_{\text{grp}}/\text{MS}_e$
- Pairwise comparisons conducted “within” the model, using all the observations to estimate the SD.

With only two groups, $t$-tests suffice. Different versions:
- Paired or unpaired?
- If unpaired: same SD or not?

Lecture summary: main points

- One-way ANOVA
- Assumptions for one-way ANOVA
- Hypotheses for one-way ANOVA
- Test statistic and the $F$-fordelingen