

NRES 798 — Statistical Methods for Ecologists  
Chapter 10: Categorical predictors, ANOVA

Oscar García

March 27, 2013

**Contents**

<b>1</b>	<b>Linear models and ANOVA</b>	<b>2</b>
1.1	Dummy variables . . . . .	2
1.2	Contrasts . . . . .	4
1.3	Notation . . . . .	4
<b>2</b>	<b>Assumptions</b>	<b>6</b>
<b>3</b>	<b>Hypothesis tests and ANOVA</b>	<b>7</b>
3.1	Historical . . . . .	7
3.2	Hypotheses and $F$ tests . . . . .	7
<b>4</b>	<b>ANOVA models</b>	<b>9</b>
4.1	One-way layout . . . . .	10
4.2	Two-way layout, no interaction . . . . .	11
4.3	Two-way with interaction . . . . .	12
4.4	Factorial designs . . . . .	12
4.5	Nested designs . . . . .	13
4.6	ANCOVA . . . . .	14
<b>5</b>	<b>Multiple comparisons</b>	<b>14</b>

# 1 Linear models and ANOVA

The *analysis of variance* (ANOVA) is a conventional way of arranging computations for testing hypotheses about linear models with categorical predictors. Estimation is secondary, through calculating several so-called “means”. ANOVA is connected with experimental design, and works well mainly for *balanced* data, where the numbers of observations for each combination of factors are the same. It does simplify hand calculation in those cases.

The modern trend, implemented in statistical packages like *R* and *SAS*, is to treat these problems in a unified way within a general linear models framework<sup>1</sup>. Traditional ANOVA tables can still be produced if required. The approach applies to balanced or unbalanced data. There is a straightforward generalization to mixtures of categorical and quantitative variables, the area known as *analysis of covariance* (ANCOVA).

The material in this section is important for understanding the output from statistical packages, and in particular from *R*.

## 1.1 Dummy variables

Let us use as an example a small data set similar to that of the ant nest counts used before:

```
> ants
  nests habitat
1     9  field
2    12  field
3     9 forest
4     6 forest
5     4 forest
6    10  scrub
```

There is a numerical response variable, **nests**, and a categorical predictor (factor) **habitat** with 3 possible values (levels).

Remember that a linear regression model is something like

$$Y_j = \beta_0 + \beta_1 x_{1j} + \beta_2 x_{2j} + \beta_3 x_{3j} + \varepsilon_j . \quad (1)$$

---

<sup>1</sup> Not a *generalized linear model* (GLM). GLMs are a special type of nonlinear model, typically used when the response is not normal as an alternative to using transformations. In *R* they are implemented by the function `glm`.

For the observations  $j = 1, 2, \dots, n$ , it relates the response  $Y_j$  to the quantitative predictors  $x_{1j}, x_{2j}$ , and  $x_{3j}$ . More generally, the model may or may not include the intercept  $\beta_0$ . Or one can write

$$Y_j = \beta_0 x_{0j} + \beta_1 x_{1j} + \beta_2 x_{2j} + \beta_3 x_{3j} + \varepsilon_j ,$$

where  $x_{0j}$  is an implicit column of 1's if  $\beta_0$  is an intercept.

We could write a similar model to relate the number of nests to habitat in our example. Let  $i$  be the habitat type, with  $i = 1$  for *field*,  $i = 2$  for *forest*, and  $i = 3$  for *scrub*. Define  $x_{ij}$  as 1 if the observation  $j$  is in habitat  $i$ , and 0 otherwise. The regression data would look like

	nests	field	forest	scrub
1	9	1	0	0
2	12	1	0	0
3	9	0	1	0
4	6	0	1	0
5	4	0	1	0
6	10	0	0	1

If  $Y_j$  is the number of nests, the model equations are

	nests	field	forest	scrub	
$Y_1 =$	$\beta_0$	$+\beta_1$			$+\varepsilon_1$
$Y_2 =$	$\beta_0$	$+\beta_1$			$+\varepsilon_2$
$Y_3 =$	$\beta_0$		$+\beta_2$		$+\varepsilon_3$
$Y_4 =$	$\beta_0$		$+\beta_2$		$+\varepsilon_4$
$Y_5 =$	$\beta_0$		$+\beta_2$		$+\varepsilon_5$
$Y_6 =$	$\beta_0$			$+\beta_3$	$+\varepsilon_6$

or

$$Y_j = \beta_0 + \beta_i + \varepsilon_j . \tag{2}$$

These 0/1 indicators are called *dummy variables*. Calculations for unbalanced designs, which are not easily handled in ANOVA tables, used to be done in this way as a multiple regression (1) with the  $x$ 's being dummy variables (after dealing with a slight complication to be discussed in a moment). This was also a good method for ANCOVA, including both ordinary numerical variables and dummy variables. Nowadays, statistical packages use this method internally for both balanced and unbalanced data. They make it a little easier for the user, however, by taking care of the dummy variable generation behind the scenes.

## 1.2 Contrasts

The slight complication with the dummy variable regression above is that there is redundancy in the parameters. Different combinations of parameter values give the same prediction. We can see that the columns for `field`, `forest` and `scrub` add up to a column of 1's, the same as the implicit column corresponding to  $\beta_0$ . In linear algebra it is said that the four column vectors are not linearly independent, and the system does not have a unique solution (not to be confused with statistical independence). Another way of seeing this is to notice that in eq. (2) one could add any constant to  $\beta_0$  and subtract it from all the other  $\beta_i$  without changing the  $Y_j$ . Or, we see that the mean for level  $i$  is  $\beta_0 + \beta_i$ , and choosing any arbitrary value for  $\beta_0$  would not make any essential difference in the model; what matters are the sums  $\beta_0 + \beta_i$ .

Perhaps the simplest way of solving the problem in this case would be to suppress the intercept. A regression without intercept gives coefficients  $\beta'_i$  representing the mean within level  $i$  that are equal to  $\beta_0 + \beta_i$  in the original model. But this would not work if there were another factor, e.g., `site`. The dummy columns for the levels of `site` would add up to 1, equal to the sum of the columns of `habitat`.

As an alternative, one could leave the intercept in and drop any one of the columns from each factor. By default, *R* omits the first level from each factor (levels go in alphabetical order unless told otherwise). When interpreting the output, note that in this instance  $\beta_0$  in eq. (2) is the expected  $Y_j$  for the first level, and the other  $\beta_i$  are differences between level  $i$  and level 1.

In general, it is necessary to impose some linear constraints on the model parameters, called *contrasts*. We have seen the simple constraints  $\beta_0 = 0$ , and  $\beta_1 = 0$ . Another contrast commonly used is shown below.

The *R* behaviour just described is for unordered factors. With factors that have been specified as *ordered*, the default contrasts represent polynomial terms: constant, linear, quadratic, etc. We ignore these for now.

## 1.3 Notation

A slight variation on eq. (2) is used to describe models with categorical predictors (“ANOVA models”). Instead of  $Y_j$ , one uses  $Y_{ij}$  to indicate replicate  $j$  within factor level  $i$ , reflecting the alternative data tabulation

Habitat	Replicate		
	1	2	3
field	9	12	
forest	9	6	4
scrub	10		

Then the model is

$$Y_{ij} = \mu + \alpha_i + \varepsilon_{ij} . \quad (3)$$

The parameter  $\mu$  is an over-all mean, and  $\alpha_i$  is the *effect* of the level  $i$  of the factor. It is assumed that the  $\alpha_i$  add up to 0, that is, the effects are above and below the over-all mean  $\mu$ . This provides a contrast that makes the parameters well-defined. For the observed values one can write

$$y_{ij} = \hat{\mu} + \hat{\alpha}_i + e_{ij} .$$

With two factors one might have

nests				Site		
habitat	site			Habitat	east	west
1	9	field	east	field	9	12
2	12	field	west	forest	9, 4	6
3	9	forest	east	scrub		10
4	6	forest	west			
5	4	forest	east			
6	10	scrub	west			

and

$$Y_{ijk} = \mu + \alpha_i + \beta_j + \varepsilon_{ijk} . \quad (4)$$

Additional terms may include *interactions* and special components for nested designs. An interaction  $(\alpha\beta)_{ij}$  corresponds to a product of variables in regression, more on that later.

Statistical packages like *R* specify the model with formulae based on a notation proposed by Wilkinson and Rogers in 1973. Model (3) becomes

$$\text{nests} \sim \text{habitat}$$

and (4) is

$$\text{nests} \sim \text{habitat} + \text{site}$$

As with the intercept in regression, the mean  $\mu$  is implicit. Computations in *R* are done with the general linear model function `lm` or with the analysis of

variance function `aov`, e.g., `lm(nests ~ habitat, ants)`. These functions are essentially equivalent, differing in how the outputs are displayed by default. Function `aov` has some additional argument options for nested design models.

The type of contrasts to be used can be set as a global option, or in the optional argument `contrasts` of `lm` or `aov`. The default `"contr.treatment"` works as described in Section 1.2. For our one-factor example, in the `lm` output the intercept is  $\hat{\mu} + \hat{\alpha}_1$ , `habitatforest` is  $\hat{\alpha}_2 - \hat{\alpha}_1$ , and `habitatscrub` is  $\hat{\alpha}_3 - \hat{\alpha}_1$ .

Sometimes it may be convenient to produce output in a format closer to eq. (3) or (4). With contrasts set to `"contr.sum"`, `lm` gives  $\hat{\mu}$  as the intercept,  $\hat{\alpha}_1$  as `habitat1`, and  $\hat{\alpha}_2$  as `habitat2`. Because the  $\alpha$ 's add up to 0,  $\hat{\alpha}_3 = -(\hat{\alpha}_1 + \hat{\alpha}_2)$ .

## 2 Assumptions

Like in any analysis, or perhaps in any scientific investigation, conclusions are only valid if the assumed model is true. Or more precisely, they would be valid if the model were true, since true models do not exist. Pragmatically, one hopes that if the model is “good enough”, then any conclusions derived from it might be good enough.

Models like those in equations (3) and (4) include a number of assumptions:

1. The model is linear. More specifically, the effects  $\alpha$ ,  $\beta$ , etc., possibly including additive interaction or other complex terms, and the residual error  $\varepsilon$ , they all contribute to the response as a sum. Like in regression, sometimes a transformation of  $Y$  can make the assumption more plausible. For instance, multiplicative effects become additive for  $\log Y$ . The linearity assumption is somewhat less restrictive than in regression, because nonlinear transformations of zeroes or ones do not make any essential difference.
2. The random variable  $\varepsilon$  is normal, with mean 0 and a variance that does not depend on the factors. Again, transformations can help, although achieving at the same time additivity, normality and homoscedasticity with the one transformation might be too much to ask.

3. Simple random sampling. That is, the observations are statistically independent, and have the same distribution as the population. It is the role of careful sampling or experimental design and randomization to strive for this.

## 3 Hypothesis tests and ANOVA

### 3.1 Historical

ANOVA computes  $F$ -ratios for testing hypotheses about factors (predictors). The null hypothesis is that certain factor in the model has no effect. That is, that all the coefficients for the factor (one for each level) are zero.

Calculations are traditionally arranged in tables, where the sums of squares for the numerator and denominator of the  $F$ -ratio arise from combining sums of squares for various groups of factor levels. Some sums of squares are obtained from differences, facilitating hand computations. We saw an example in the ANOVA for linear regression.

In general, with more than one factor ANOVA really works well only for balanced data from designed experiments, where the numbers of replicates in all groups are the same. With unbalanced data things do not add up, literally. Some of the product terms that appear when squaring partitioned differences do not vanish, so that component sums of squares do not add to the total. With unbalanced data, the interpretation of  $F$ -ratio tests also becomes more difficult.

It can be argued that the motivation behind the ANOVA table is largely obsolete, only the  $F$ -ratios with their degrees of freedom and/or the  $p$ -values are useful. Some reviewers and journals, however, still insist on the conventional tables, and statistical software can print them out. One possible justification for the ANOVA table is that the *Sources* column describes the model that was assumed in the analysis. However, more explicit model descriptions in the form of equations (3) and (4) are gaining ground.

### 3.2 Hypotheses and $F$ tests

The usual ANOVA null hypothesis is that some factor has no effect, e.g., that *all* the  $\alpha_i$  in (3) or (4) are 0. The alternative is a model containing

those effects:

$$H_0 : Y_{ijk} = \mu + \beta_j + \varepsilon_{ijk} , \quad H_a : Y_{ijk} = \mu + \alpha_i + \beta_j + \varepsilon_{ijk} .$$

An appropriate  $F$ -ratio statistic can be built with the RSS obtained from fitting the  $H_0$  and  $H_a$  models, say  $\text{RSS}_0$  and  $\text{RSS}_a$ , and the respective degrees of freedom:

$$F = \frac{(\text{RSS}_0 - \text{RSS}_a)/(\text{df}_0 - \text{df}_a)}{\text{RSS}_a/\text{df}_a} . \quad (5)$$

It measures how much worse the fit gets when dropping  $\alpha$ . If  $H_0$  is true, then this  $F$ -ratio has an  $F$  distribution with  $\text{df}_0 - \text{df}_a$  and  $\text{df}_a$  degrees of freedom. The more important the factor  $\alpha$  is, the worse the fit of  $H_0$  becomes, and the larger the value of  $F$ . The null hypothesis is therefore rejected if the observed  $F$  is larger than the quantile corresponding to the chosen significance level. Or if the  $p$ -value for  $F$  is smaller than the significance level.

In ANOVA the denominator of the  $F$ -ratio above is called the residual mean square, or mean square within groups. It is the variance of the residuals. The numerator is called the mean square due to the factor  $\alpha$ .

Yes, very often that is the wrong question. A real problem with many published statistical analyses. One may know that there must be *some* effect, and the relevant question is how large the effect may be. An estimation problem, answered by the parameter estimates in the linear model. Or the fact that a factor affects the response may be uninteresting, but we might like to know about differences among levels; a problem of multiple comparisons, to be discussed later.

In  $R$ ,  $F$  could be obtained as shown here, from the results of `lm` applied to the  $H_0$  and  $H_a$  models. Alternatively, given the two fitted models as arguments, the `anova` function produces a small model comparison ANOVA table that contains the numbers required:

```
> H0 <- lm(nests ~ site, ants2)
> Ha <- lm(nests ~ habitat + site, ants2)
> anova(H0, Ha)
Analysis of Variance Table

Model 1: nests ~ site
Model 2: nests ~ habitat + site
  Res.Df    RSS Df Sum of Sq    F Pr(>F)
1       4 35.333
```

```
2      2 16.000  2      19.333 1.2083 0.4528
```

The hypothesis test is only valid if one model is a special case of (it is contained in) the other one (they are *nested*), and if both were fitted to the same data. Apart from that, this is very general, and applies also to linear and nonlinear regression models, and to analysis of covariance. Although for testing significance of individual variables in regression it might be easier to use the  $t$  values from the output; the tests are equivalent, the  $F$ -ratio is the square of the  $t$  statistic. Granted, it is not the most efficient approach computationally, it wastes many milliseconds of computer time!

With balanced data, the simplest and most direct way is to obtain an ANOVA table for  $H_a$  with `anova(Ha)`, or as

```
> summary(aov(nests ~ habitat + site, ants2))
              Df Sum Sq Mean Sq F value Pr(>F)
habitat       2  24.167  12.083   1.510  0.398
site          1   1.167   1.167   0.146  0.739
Residuals    2  16.000   8.000
```

It should give the  $F$ -ratios and  $p$ -values for testing both factors. However, in this example the data is unbalanced and the numbers are wrong. Or at least are not what one would expect, and change if one changes the order of `habitat` and `site`.

## 4 ANOVA models

The “bestiary” of ANOVA models can be understood in terms of the general model just discussed. The purpose of the ANOVA tables is to facilitate the manual computation of the  $F$ -ratio in equation (5). Examples of many of them are found in the textbook.

It is a good idea, and it is becoming standard, to show explicitly the structure of the model with the notation of equations (3) and (4). In  $R$  the  $F$ -ratio and  $p$ -value can always be obtained from `anova` applied to the null and alternative models fitted with `lm` or with `aov`. If required, for balanced data one can obtain the traditional table with `anova(lm(.))` or `summary(aov(.))` for the alternative model, and pretend that it was done the old way.

Very frequently the hypothesis test is uninteresting. The fact that certain treatments must have *some* effect may be obvious, the real question is how

large the effects might be. This is an estimation problem, that ANOVA handles largely as an afterthought. The expected responses to combinations of factor levels are estimated by so-called *means*, the averages of the observations within each group.

An alternative to calculating group averages, that produces the same result, is to substitute the parameter estimates from `lm` into the model equation. output, by summing the intercept and the appropriate parameter estimates. For instance, for a model of the form  $Y_{ijk} = \alpha_i + \beta_j + \varepsilon_{ijk}$  such as `nests ~ habitat + site`, using the `sum` contrast option:

Habitat	Site	
	East	West
Field	(Intercept) + habitat1 + site1	(Intercept) + habitat1 - site1
Forest	(Intercept) + habitat2 + site1	(Intercept) + habitat2 - site1
Scrub	(Intercept) - habitat1 - habitat2 + site1	(Intercept) - habitat1 - habitat2 - site1

The last level that is not shown in the `lm` output have been substituted in terms of the other levels, using  $\sum \alpha_i = \sum \beta_j = 0$ .

With the default `treatment` contrast, which omits the first level from each factor, it is a little simpler:

Habitat	Site	
	East	West
Field	(Intercept)	(Intercept) + sitewest
Forest	(Intercept) + habitatforest	(Intercept) + habitatforest + sitewest
Scrub	(Intercept) + habitatscrub	(Intercept) + habitatscrub + sitewest

In addition, one may be interested in testing differences between means for a number of group pairs. These *multiple comparisons* problems are discussed later.

## 4.1 One-way layout

Or *one-factor* ANOVA. This is the model

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

that we have already seen.

## 4.2 Two-way layout, no interaction

Or *two-factor*. The *no replicates* variant has one observation for each combination of levels from factors  $\alpha$  and  $\beta$ :

$$Y_{ij} = \mu + \alpha_i + \beta_j + \varepsilon_{ij} .$$

The *replicated* variant has the same number of observations (replicates) for each combination of levels:

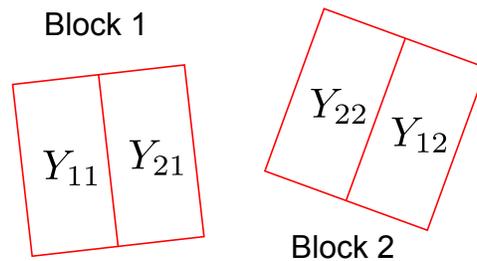
$$Y_{ijk} = \mu + \alpha_i + \beta_j + \varepsilon_{ijk}$$

(eg. (4), balanced).

Depending on the objectives and design, factors or factor levels may be called *treatments* if one is interested in their effects (e.g., habitat), or *blocks*, if not (e.g., site).

The purpose of blocks is to increase homogeneity within groups and thus reduce the RSS and increase the power of the test by making  $F$  larger (see eq. (5)). Same as in the paired  $t$ -test, where the comparisons are done between pairs of similar individuals, or on a same individual.

Blocks are implemented in *randomized block designs*, that correspond to a no-replicates model (the blocks themselves can be considered as replicates). Note the different elements: model–design–analysis; these are closely related but it is useful to keep in mind the distinctions. A typical design in a field experiment would have the treatment levels  $i$  randomized within compact blocks  $j$ :



This model applies also to some *repeated-measures* and other experimental designs.

### 4.3 Two-way with interaction

Or two-way factorial. In the models of Section 4.2 the factor effects are additive, the response differences between two levels of factor 1 are the same regardless of the level of factor 2. Same for the factor 2 effects. An *interaction* term causes the responses to be different.

The analogous in linear regression is to have  $E[Y] = \beta_0 + \beta_1x_1 + \beta_2x_2$ . The slope of  $Y$  vs.  $x_1$  is the same for any value of  $x_2$ , only the intercept changes. A product term causes the slope to vary:  $E[Y] = \beta_0 + \beta_1x_1 + \beta_2x_2 + \beta_3x_1x_2$ . With categorical variables the interaction is implemented with products of dummy variables, but as before the model can be specified with simplified formulae without worrying about the details.

The two-factor model with interaction is

$$Y_{ijk} = \mu + \alpha_i + \beta_j + (\alpha\beta)_{ij} + \varepsilon_{ijk} ,$$

where  $(\alpha\beta)_{ij}$  denotes the interaction term. Replication is needed, otherwise there is not enough information for estimating the interaction. On the computer, the Wilkinson-Rogers notation used a dot to suggest the product in the interaction, but because *dot* is a valid name character in *R* it was changed to a colon:

$$Y \sim \text{alpha} + \text{beta} + \text{alpha:beta}$$

This can also be written as

$$Y \sim \text{alpha} * \text{beta}$$

The interaction can be visualized with `interaction.plot(alpha, beta, Y)` and `interaction.plot(beta, alpha, Y)`.

Besides balance, for the ANOVA table to be correct the null hypothesis must respect the *principle of marginality*. This means that the *main effects*  $\alpha$  and  $\beta$  cannot be absent if the interaction term is included; a null hypothesis like  $Y_{ijk} = \mu + \alpha_i + (\alpha\beta)_{ij} + \varepsilon_{ijk}$  is not admissible.

### 4.4 Factorial designs

A *full factorial* is the model of Section 4.3 extended to any number of factors. E.g., for 3 factors:

$$Y_{ijkl} = \mu + \alpha_i + \beta_j + \gamma_k + (\alpha\beta)_{ij} + (\alpha\gamma)_{ik} + (\beta\gamma)_{jk} + (\alpha\beta\gamma)_{ijk} + \varepsilon_{ijkl} ,$$

and

$$Y \sim \text{alpha} + \text{beta} + \text{gamma} + \text{alpha}:\text{beta} + \text{alpha}:\text{gamma} + \text{beta}:\text{gamma} + \text{alpha}:\text{beta}:\text{gamma} .$$

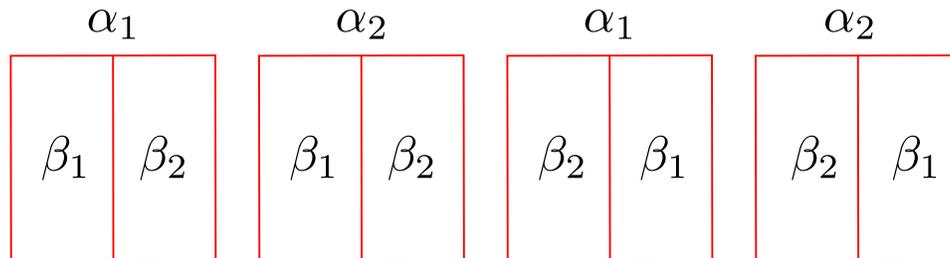
Or

$$Y \sim \text{alpha} * \text{beta} * \text{gamma} .$$

Various *incomplete-blocks* and other designs omit some of the interactions. The principle of marginality must be observed when using ANOVA tables.

#### 4.5 Nested designs

In nested designs one treatment is applied “inside” another. There are two or more hierarchical levels of experimental units, sometimes called *main plots* and *subplots*. For instance, in an agricultural experiment irrigation, factor  $\alpha$ , may be difficult to control on small areas, so it is applied to large main plots. Then, each main plot is split into subplots that receive different levels on another treatment,  $\beta$ , such as plowing or fertilizing. This is called a *split-plot* design:



This is a specialized topic in experimental design that can get rather complex, with nesting at several hierarchical levels. A few more details can be found in the textbook. Within the framework that we have been discussing, a very simple model might look like

$$Y_{ijk} = \mu + \alpha_i + \beta_{j(i)} + \varepsilon_{ijk} ,$$

where  $\beta_{j(i)}$  indicates that treatment  $\beta$  is applied to subplots of the main plot receiving the treatment  $\alpha$ . The model translates into the following  $R$  formula:

$$Y \sim \text{alpha} + \text{alpha} / \text{beta} .$$

An alternative to `alpha / beta` is `beta %in% alpha`.

Another approach to nested designs represents effects of some uninteresting factors as random variables instead of parameters. *Random effects* instead of *fixed effects*. There are then other error terms in addition to the regular  $\varepsilon$ . In the *R* model formula, random expressions are specified inside a term `Error(.)`, like `Y ~ alpha + Error(beta)` as the simplest possible example. This is available in `aov`, but not in `lm`.

This second approach goes beyond the standard linear model framework, into what are called random effect methods and variance component models. More general linear mixed effect models are handled in package `lme`, and nonlinear mixed effects in package `nlme`.

## 4.6 ANCOVA

It is straightforward to combine categorical and numeric predictors in a linear model, using `lm` or `aov`. That is called *analysis of covariance* (ANCOVA), and the numeric predictors are *covariates*.

Hypotheses can always be tested by the *F*-ratio based on a reduced null model and the full alternative. In well behaved cases the ANOVA table for the full model may provide the information needed.

## 5 Multiple comparisons

OK, we rejected the null hypothesis, now what? It seems unlikely that the factor has no effect. But what about some of the factor levels, or differences between levels?

One could carry out hypothesis tests to try answering many of these questions. Specifically, we may want to test significance for the response differences among all the pairs of levels of a factor. Differences among the pairs of *means*, that is, means within groups that have the same factor levels. In Section 10 it was seen how to get these from the `lm` output. A *t*-test for the difference of means can be used.

The problem is that finding a significant difference in a battery of tests is much more likely than what the nominal significance level suggests. There is

a pretty good chance of rejecting one hypothesis at the 0.05 level in a batch of 20 tests.

The significance level  $\alpha$  is the probability of a hypothesis test rejecting the true hypothesis (Type I error). The probability of any one of  $k$  tests rejecting its hypothesis at the  $\alpha$  level, if the tests were statistically independent, would be  $k\alpha$ . This is a Bonferroni upper bound, the probability for non-independent tests is less (draw a Venn diagram). The Bonferroni method adjusts the significance level of each test to  $\alpha/k$ , so that the Type I error for the lot is at most  $\alpha$ . This method is commonly used, although obviously it can be rather conservative. The way it is usually applied is to compute an  $\alpha/k$ -level confidence interval for the mean differences, and if the interval includes 0 the means are deemed not to be significantly different.

Another popular multiple comparisons tests that works in a similar way but is less conservative, is Tukey's *honestly significant difference* (HSD). This is available in *R* as `TukeyHSD`.