# Chapter 13

# Design of Experiments—Factorial Designs

Designs are often described by the number of factors. Chapter 6, “One-Way Analysis of Variance”, discusses designs with one factor. Chapter 12, “Two-Way Analysis of Variance”, discusses designs with two factors. More generally, we speak of “three-way” or “higher-way” designs and talk about main effects (one factor), two-way interactions (two factors), three-way interactions, four-way interactions, and so forth. Factors can have crossed or nested relationships. A factor can be fixed or random. When interaction is significant, its nature must be carefully investigated. If higher-order interactions, meaning those involving more than two factors, can be assumed to be negligible, it is often possible to design experiments that require observations on only a fraction of all possible treatment combinations.

Section 13.1 discusses a three-way ANOVA design with a covariate and polynomial contrasts. Section 13.2 introduces Latin squares. Section 13.3 introduces simple effects for interaction analyses. Section 13.4 discusses a nested factorial experiment with both crossed and nesting relationships among the factors. Section 13.6.1 discusses the SAS terminology for types of sums of squares used in sequential and conditional ANOVA tables. Related topics are discussed in Chapter 14.

## 13.1 A Three-Way ANOVA—Muscle Data

Cochran and Cox (1957) report on an experiment to assess the effect of electrical stimulation to prevent the atrophy of muscle tissue in rats. The dataset is available as `data(cc176)`. The response `wt.d` is the weight of the treated muscle. There were three fixed factors: the number of treatments daily, `n.treat`, 1, 3, or 6; the duration of treatments in `minutes`, 1, 2, 3, or 5; and the four types of `current` used. A concomitant variable, the weight of corresponding muscle on the opposite untreated side of the rat, `wt.n`, was also made available. There were two replications of the entire experiment.



R.M. Heiberger, B. Holland, *Statistical Analysis and Data Display*,
Springer Texts in Statistics, DOI 10.1007/978-1-4939-2122-5_13

The analysis is constructed with code in `HHscriptnames(13)`. The data are plotted in Figure 13.1. The ANCOVA and adjusted means are in Table 13.1. Also included in Table 13.1 is a partitioning of the 2 degrees of freedom for `n.treats` into linear and quadratic components, taking account of the unequal spacing of the quantitative levels of `n.treats`.

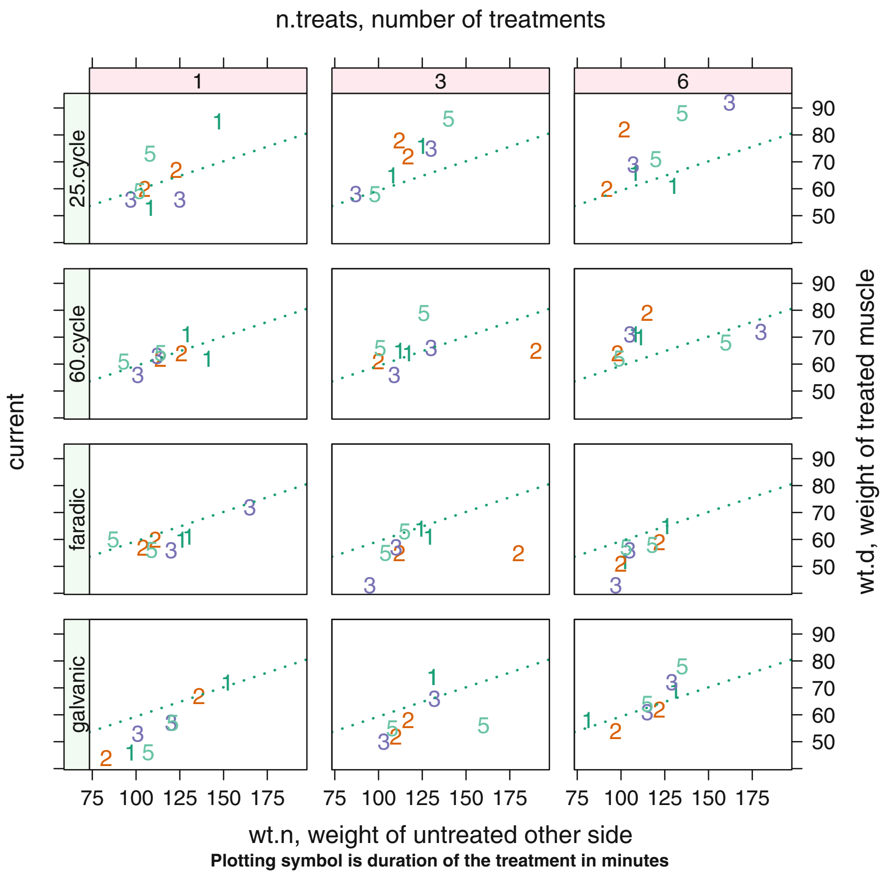



**Fig. 13.1** Muscle data. The response variable `wt.d` is plotted against the covariate `wt.n` within each `current`×`n.treats` experimental condition. The plotting symbol is the duration of the treatment in minutes. The ANCOVA and adjusted means are in Table 13.1. We also plotted the common regression line (ignoring experimental conditions) of the response against the covariate. The presence of a covariate `wt.n` effect is evident from the graph by noting that the points in all panels approximate the uphill slope of the regression slope. The absence of a `minutes` effect is evident since there is no systematic pattern among the plotting symbols.

Table 13.1 suggests that after adjusting for the concomitant variable `wt.n`, there are no significant interactions and the effect of `minutes` is not significant. This table shows that `n.treat` contributes significantly to explaining `wt.d`,

**Table 13.1** Muscle data. ANCOVA and adjusted means. The covariate `wt.n`, the linear effect of `n.treats`, and the `current` are the significant treatment effects. We show the calculation of `y.adj`, the response variable adjusted for the covariate, and the adjusted means.

```
> ## y=wt.d with x=wt.n as covariate
> ## (get essentially the same ANOVA as the approximate (y-bx)^2
> ## ANOVA table in Cochran and Cox)
> cc176.aov <- aov(wt.d ~ rep + wt.n + n.treats*minutes*current,
+                  data=cc176)

> ## summary(cc176.aov)
> summary(cc176.aov,
+         split=list(n.treats=list(n.treats.lin=1,
+                                  n.treats.quad=2)),
+         expand.split=FALSE)
                          Df Sum Sq Mean Sq F value  Pr(>F)
rep                        1    605     605   12.58 0.00091 ***
wt.n                       1   1334    1334   27.74 3.6e-06 ***
n.treats                   2    439     219    4.56 0.01557 *
  n.treats: n.treats.lin   1    438     438    9.11 0.00413 **
  n.treats: n.treats.quad  1      1       1    0.01 0.91048
minutes                    3    184      61    1.28 0.29409
current                    3   2114     705   14.66 7.8e-07 ***
n.treats:minutes           6    198      33    0.69 0.66051
n.treats:current           6    492      82    1.70 0.14163
minutes:current            9    383      43    0.88 0.54627
n.treats:minutes:current  18   1022      57    1.18 0.31542
Residuals                 46   2212      48
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

> ##
> ## adjust y for x
> cc176$y.adj <- cc176$wt.d  -
+   (cc176$wt.n - mean(cc176$wt.n))*coef(cc176.aov)["wt.n"]

> ## duplicate CC Table 5.17
> cc176.means <- tapply(cc176$y.adj,
+                       cc176[,c("current","n.treats")], mean)

> cc176.means
          n.treats
current        1     3     6
  galvanic 56.03 59.08 65.29
  faradic  59.95 55.79 57.27
  60.cycle 63.26 63.92 68.58
  25.cycle 64.47 71.78 73.20

> apply(cc176.means, 1, mean)
galvanic  faradic 60.cycle 25.cycle
   60.13    57.67    65.25    69.82
```

$p$–value = .0000036. The visible upward trend in all panels of Figure 13.1 suggests that response `wt.d` increases linearly with `n.treats` and differs according to the type of `current` used. The response variable `y.adj` in Figure 13.2 is constructed by adjusting the response `wt.d` for the covariable `wt.n`. We see in both Figures 13.1 and 13.2 a larger response when `current` is 25 `cycle` than when `current` is at one of its other three levels. Inclusion of `wt.n` reinforces these conclusions. The parallel traces in Figure 13.2 correspond to the absence of interaction between `n.treat` and `current`, a finding also suggested by the large $p$-value for this interaction in Table 13.1.

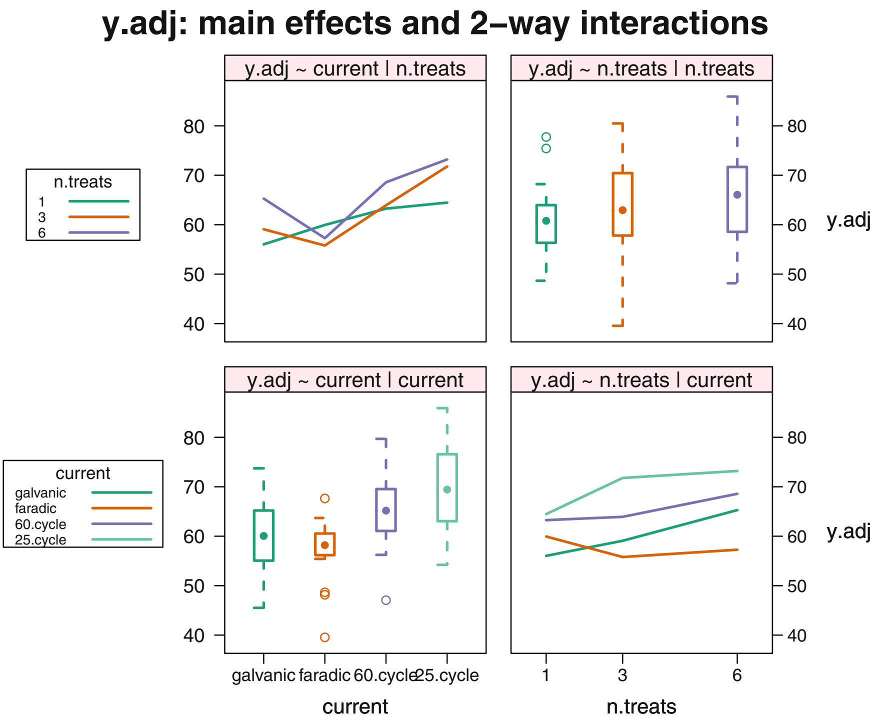


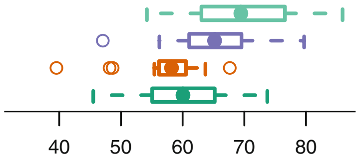

**Fig. 13.2** Muscle data. Two-way interactions of significant main effects from the ANCOVA in Table 13.1. The adjusted response `y.adj` increases linearly with `n.treats` and differs according to the type of current used.

In Table 13.2 we display a microplot of horizontal boxplots that compares the distributions of responses for each level of the factor `current`. This is the same set of boxplots that appears in the y.adj ~ current | current panel of Figure 13.2.

Boxplots capture comparative information better than numbers. They don't have to take much space, therefore they can fit into tables of numbers and satisfy both the convention (perhaps mandated) of displaying numbers and the legibility of displaying graphs. We call the plots *microplots* when they are deliberately sized to fit into a table of numbers without interfering with the overall layout. When small plots are placed into a table of numbers, they can carry the same or more information per $cm^2$ as the numbers themselves.

**Table 13.2** Muscle data: Distribution statistics and boxplots for adjusted weights. The statistics show only a few values. The boxplot shows the entire distribution.

| Treatment | Min | m−sd | Mean | m+sd | Max | boxplot |
|---|---|---|---|---|---|---|
| 25.cycle | 54.21 | 61.04 | 69.82 | 78.59 | 85.92 | |
| 60.cycle | 47.07 | 57.93 | 65.25 | 72.58 | 79.69 | |
| faradic | 39.55 | 51.98 | 57.67 | 63.37 | 67.64 | |
| galvanic | 45.52 | 52.82 | 60.13 | 67.44 | 73.72 | |




A display comparable to Figure 13.3 could be used to determine the nature of a 3-way interaction. Such an interaction does not exist in this example.

Figure 13.4 shows four different models for the relationship between the response `wt.d` and the covariate `wt.n`. The figure is similar to Figure 10.12 which showed four sets of panels for the simpler dataset with only one factor. The overall conclusion is that the relation between `wt.d` and `wt.n` differs according to the levels of `current` and `n.treat`. More detail appears in the caption of this figure.

Figure 13.5 is a Tukey procedure MMC plot examining the six pairwise differences among the four levels of `current`. As summarized in the caption of this figure, four of these six differences are declared statistically significant. Inspection of Figure 13.5 and the means of the levels of `current` in Table 13.1 reveals that `25.cycle` and `60.cycle` current, indistinguishable from each other, correspond to significantly greater treated muscle weight `wt.d` than either `galvanic` or `faradic` current.

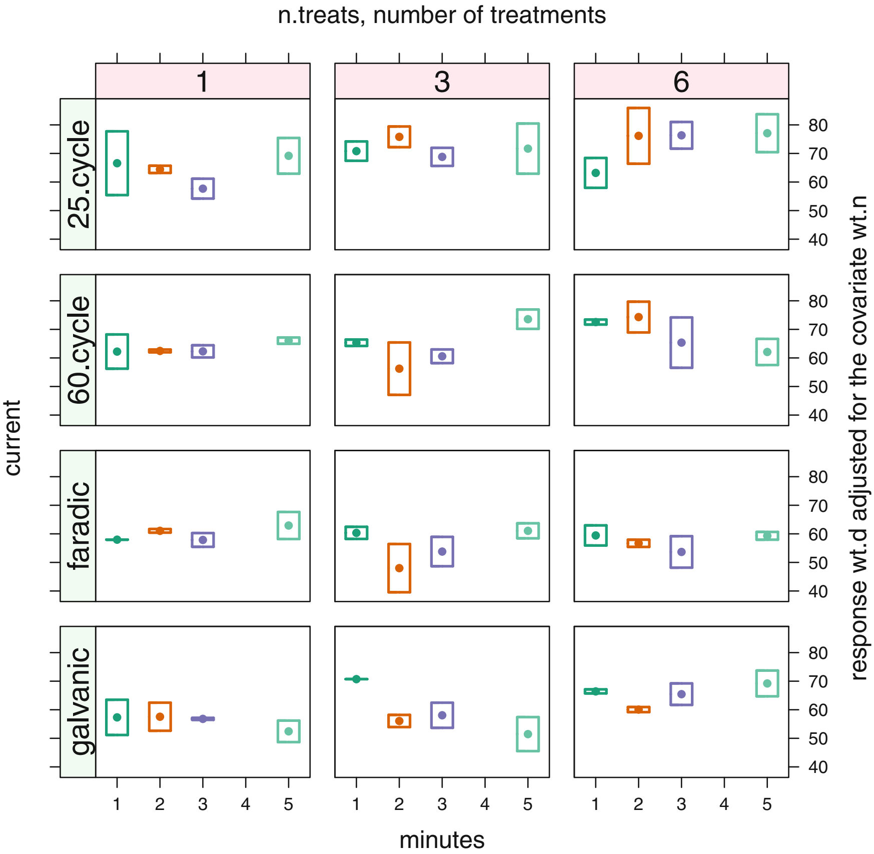



**Fig. 13.3** Muscle data. Three-way interactions of all effects. One of the (3! = 6) possible orderings. The three-way interaction is not significant in this example. If there were a significant three-way interaction, the patterns in boxplots in adjacent rows and columns would not be the same. For example, we note a hint of a difference in the `y.adj ~ minutes` behavior across panels. It has a negative slope in the `galvanic ~ 3` panel and a positive slope in the `faradic ~ 3` panel, but a positive slope in the `galvanic ~ 6` panel and a negative slope in the `faradic ~ 6` panel. The ANOVA table tells us these differences in slope are not significant. These boxplots are all based on samples of size 2. Such boxplots are a well-defined but uncustomary way to display such a sample.

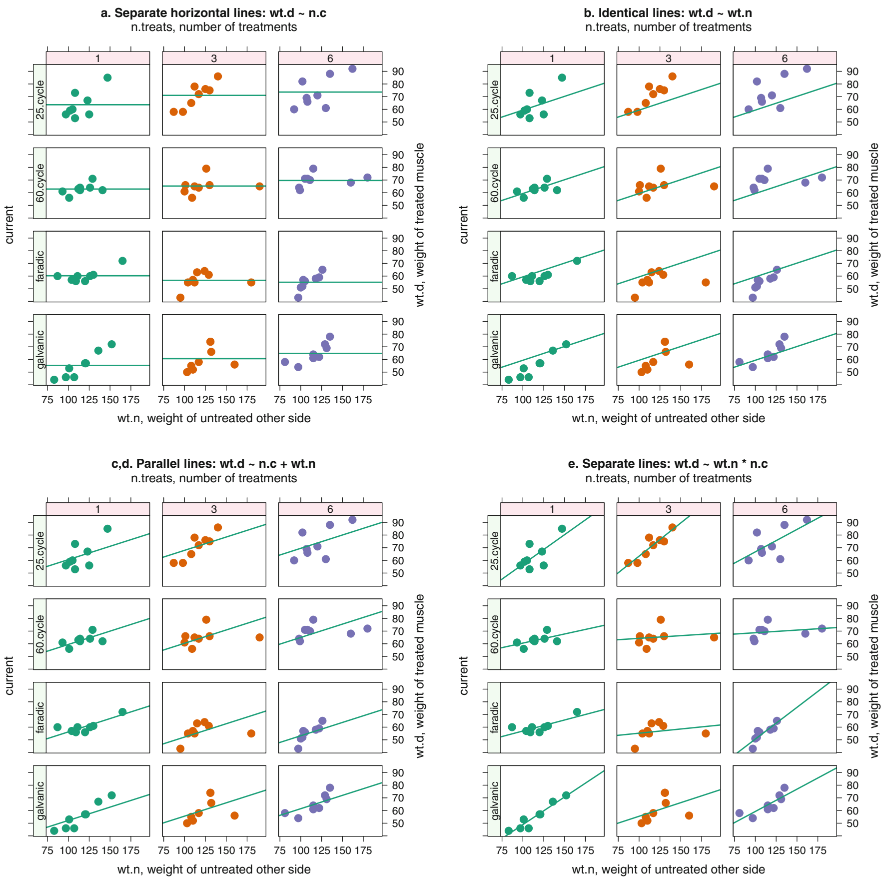



**Fig. 13.4** Muscle data. ANCOVA plots with four different models. Panel a ignores `wt.n` and shows the average value of `wt.d`. Panel b fits a common regression of `wt.d` on `wt.n` on all combinations of `n.treat` and `current` and differs from Figure 13.1 only in its choice of plot symbol. Panel c,d allows for different intercepts but forces common slopes. The difference in intercepts corresponds to the small *p*-value for `wt.n` in Table 13.1. Panel e shows distinct regressions of `wt.d` on `wt.n` for each combination of `current` and `n.treat`. It suggests that the relationship between `wt.d` and `wt.n` differs according to the levels of `n.treat` and `current`. The term `n.c` in the title for the graphs is an abbreviation for the interaction `(n.treats*current)`.

**Table 13.3** Muscle data. ANCOVA with the simpler model using only the significant terms from Table 13.1 plus an additional interaction. We show the MMC plot for the `current` effect in Figure 13.5.

```
>   cc176t <- cc176

>   for (i in names(cc176t))
+     if (is.factor(cc176t[[i]]))
+       contrasts(cc176t[[i]]) <-
+         contr.treatment(length(levels(cc176t[[i]])))

>   sapply(cc176t, class)
$wt.d
[1] "numeric"

$wt.n
[1] "numeric"

$n.treats
[1] "positioned" "ordered"    "factor"

$current
[1] "positioned" "ordered"    "factor"

$minutes
[1] "positioned" "ordered"    "factor"

$rep
[1] "factor"


>   cc176t.aov <- aov(wt.d ~ rep + wt.n + n.treats + wt.n*current,
+                      data=cc176t)

>   summary(cc176t.aov)
             Df Sum Sq Mean Sq F value  Pr(>F)
rep           1    605     605   14.19 0.00030 ***
wt.n          1   1334    1334   31.29 2.6e-07 ***
n.treats      2    439     219    5.15 0.00776 **
current       3   2114     705   16.53 1.5e-08 ***
wt.n:current  3    867     289    6.78 0.00038 ***
Residuals    85   3624      43
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

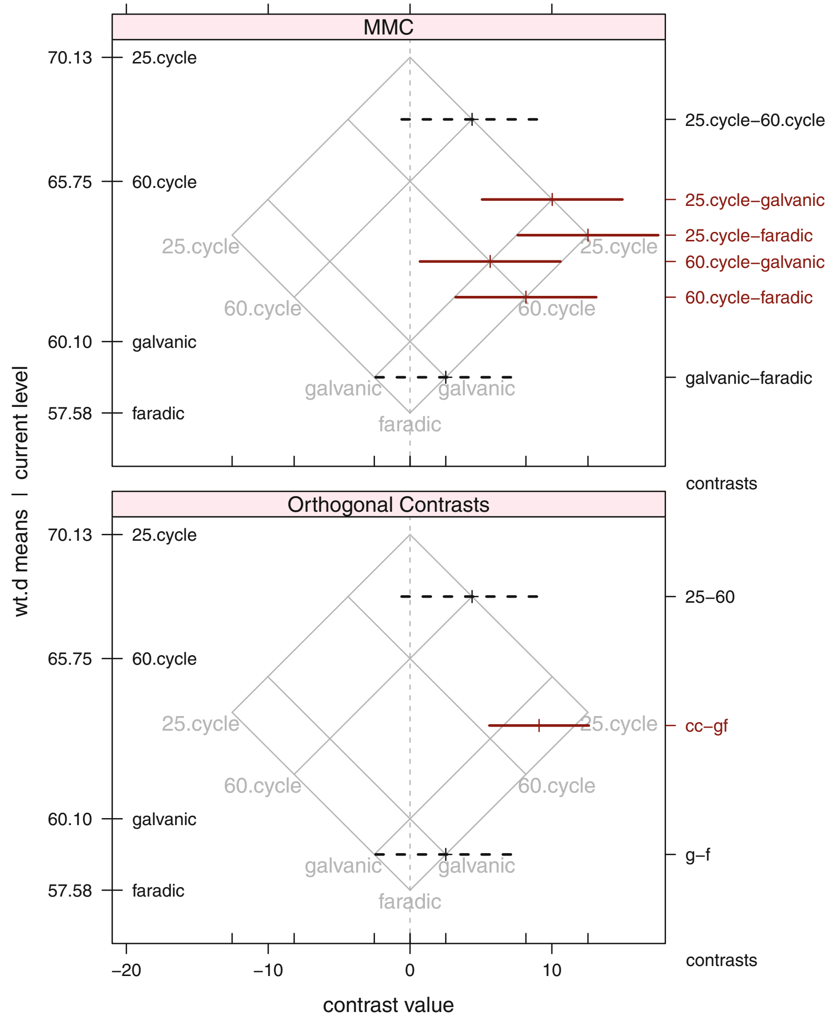



**Fig. 13.5** Muscle data. MMC plot for the analysis in Table 13.3. The top panel shows the Tukey 95% intervals for all pairwise contrasts of treatment means `wt.d` adjusted for `wt.n` at the four levels of `current`. The Tukey procedure leads to the conclusions illustrated by the orthogonal contrasts in the bottom panel. The adjusted treatment means of both of `25.cycle` and `60.cycle` exceed both of `galvanic` and `faradic`; `25.cycle` is indistinguishable from `60.cycle`; and `galvanic` is indistinguishable from `faradic`. The Tiebreaker plot is not needed in this example.

## 13.2 Latin Square Designs

This design is useful when we have three factors having the same number, say $r$, of levels and the factors do not interact. Although there are $r^3$ treatment combinations, the Latin Square design permits us to run the experiment with a carefully chosen

**Table 13.4** Sample 4 × 4 Latin square design. The rows represent tire positions: LF is Left-Front, RF is Right-Front, LR is Left-Rear, and RR is Right-Rear.

| | Car | | | |
|---|---|---|---|---|
| Position | 1 | 2 | 3 | 4 |
| LF | C | D | A | B |
| RF | B | C | D | A |
| LR | A | B | C | D |
| RR | D | A | B | C |

**Table 13.5** Sample ANOVA for 4 × 4 Latin square design.

| Source | df | Sum of Sq | Mean Sq | $F$ Value | Pr(> $F$) |
|---|---|---|---|---|---|
| Row | $r-1$ | $\text{SS}_{\text{Row}}$ | $\text{MS}_{\text{Row}}$ | | |
| Column | $r-1$ | $\text{SS}_{\text{Col}}$ | $\text{MS}_{\text{Col}}$ | | |
| Treatment | $r-1$ | $\text{SS}_{\text{Trt}}$ | $\text{MS}_{\text{Trt}}$ | $\text{MS}_{\text{Trt}}/\text{MS}_{\text{Res}}$ | $1-\mathcal{F}_{\text{df}_{\text{Trt}},\text{df}_{\text{Res}}}(F)$ |
| Residual | $(r-1)(r-2)$ | $\text{SS}_{\text{Res}}$ | $\text{MS}_{\text{Res}}$ | | |
| Total | $r^2-1$ | $\text{SS}_{\text{Total}}$ | | | |

subset of $r^2$ of them while retaining the ability to conduct tests on all main effects. A Latin square is a square array of $r$ Latin letters $A, B, C, \ldots$ such that each letter appears exactly once in each row and once in each column. Typically, the treatment factor is associated with these letters, and both *row* and *column* are blocking type factors. For example, if an experiment is run at $r$ selected times of day on each of $r$ days, then each row could represent one of the days, and each column one of the selected times. As another example, displayed in Table 13.4, if we have four cars available to compare the wear of four brands of tire, the rows of the square could represent the wheel position on the car, the columns represent the selected car, and the letters the tire brands.

The basic structure of the ANOVA table is in Table 13.5. Since we are using the Row and Columns factors as blocks, there is no test for those terms in the table. The only test we are justified in making is the test of the Treatment. The purpose of including the Row and Column factors is to pick up some of the Total Sum of Squares and thereby reduce the size of the residual mean square.

The arithmetic of the Latin square design depends on the assumption of no interaction between Row, Column, and Treatment. The arithmetic of the interaction of Row and Column gives $(r-1)^2$ df to the interaction and 0 df for an error term. By assuming no interaction, we gain the ability to split the $(r-1)^2$ df into two components: Treatment with $r-1$ df and Error with $(r-1)^2-(r-1)=(r-1)(r-2)$ df. If the no-interaction assumptions hold, this is a very efficient design.

Almost always, $5 \le r \le 8$, for if $r < 5$ there are too few df for error, and one is unlikely to encounter situations where one has three factors each having $r > 8$ levels, two of which are blocking factors. However, it is possible to run an experiment containing several 3×3 squares or several 4×4 squares, each of which is considered a block, in order to achieve sufficient error df.

Catalogs of Latin squares appear in Cochran and Cox (1957) and elsewhere. In practice, one selects a square from a catalog and randomizes it by randomly assigning levels of one of the blocking factors to the rows of the square, randomly assigning levels of the other blocking factor to the columns of the square, and then randomly assigning treatment levels to the letters.

### *13.2.1 Example—Latin Square*

The dataset `data(tires)`, from Hicks and Turner (1999, page 115), is displayed in Table 13.6 alongside the original Latin square. A boxplot of the data is in Figure 13.6.

An initial ANOVA run in Table 13.7 revealed significant differences among cars and brands, but not among positions. Here $r = 4$, allowing just 6 df for estimating error. Hence the denominator df of the $F$-tests is also 6, which as discussed in Section 5.4.4 implies that these tests have little power. Nevertheless, the differences in this example among `cars` and `brands` are large enough for the $F$-tests to detect them.

**Table 13.6** Latin square of tire wear experiment. The Latin square from Table 13.4 is repeated here. On the left with letters and on the right with the observed response values.

| | Car | | | | | Car | | | |
|---|---|---|---|---|---|---|---|---|---|
| Position | 1 | 2 | 3 | 4 | Position | 1 | 2 | 3 | 4 |
| LF | C | D | A | B | LF | 12 | 11 | 13 | 8 |
| RF | B | C | D | A | RF | 14 | 12 | 11 | 13 |
| LR | A | B | C | D | LR | 17 | 14 | 10 | 9 |
| RR | D | A | B | C | RR | 13 | 14 | 13 | 9 |

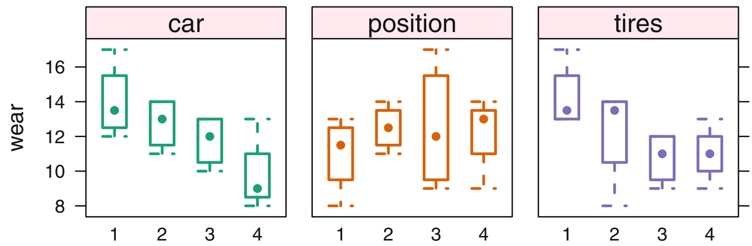



**Fig. 13.6** Tires data. Boxplots of the response variables `wear` against the three factors.

To learn about the nature of the brand differences, we reran with a request for Tukey multiple comparisons tests on the brand means (Tables 13.7 and 13.8 and Figure 13.7). We find that brand 1 had significantly greater wear than brands 3 and 4, but the improvement in wear of brand 1 over brand 2 was not significant. We also see that cars 1, 2, and 3 all had significantly greater wear than car 4; no significant difference in tire wear was detected among cars 1, 2, and 3.

In this example the primary interest is studying the differences between brands of tires. Both car and position are blocking factors. We assume different cars will have different effects on tires because each person who owns a car drives different routes and puts the car through different wear patterns. We know there are differences in position on the car. Front tires are used for steering, rear tires just follow. In some cars only the front tires get power directly from the engine. In other cars only the rear tires, and in 4-wheel drive vehicles both front and rear tires get power. The goal of the Latin square experiment is to reduce the residual sum of squares by absorbing some of the variation into known blocking factors. This makes the comparisons of interest, those on brand, more precise because they can be made with a smaller standard deviation (based on the residual mean square).

**Table 13.7** Latin square design. `tires` data. Differences in the blocks `car` are large, justifying blocking. Differences in `brand` are significant. We investigate further in Table 13.8.

```
>   data(tires)

>   tires.aov <- aov(wear ~ car + position + brand, data=tires)

>   summary(tires.aov)
            Df Sum Sq Mean Sq F value  Pr(>F)
car          3  38.69  12.896  14.395 0.00378 **
position     3   6.19   2.062   2.302 0.17695
brand        3  30.69  10.229  11.419 0.00683 **
Residuals    6   5.38   0.896
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

>   tapply(tires$wear, tires$car, "mean")
    1     2     3     4
14.00 12.75 11.75  9.75

>   tapply(tires$wear, tires$position, "mean")
    1     2     3     4
11.00 12.50 12.50 12.25

>   tapply(tires$wear, tires$brand, "mean")
    1     2     3     4
14.25 12.25 10.75 11.00
```

**Table 13.8** Continuation of analysis in Table 13.7. Latin square design. `tires` data. Brand 1 shows significantly greater mean `wear` than brand 3 or brand 4.

```
>   tires.mmc.brand <- mmc(tires.aov, linfct=mcp(brand="Tukey"))

>   ## print(tires.mmc.brand)
>   brand.lmat <- cbind("1-43" =c( 2, 0,-1,-1),
+                       "4-3"  =c( 0, 0,-1, 1),
+                       "143-2"=c( 1,-3, 1, 1))

>   dimnames(brand.lmat)[[1]] <- levels(tires$brand)

>   tires.mmc.brand <- mmc(tires.aov, linfct=mcp(brand="Tukey"),
+                          focus.lmat=brand.lmat)

>   print(tires.mmc.brand)
Tukey contrasts
Fit: aov(formula = wear ~ car + position + brand, data = tires)
Estimated Quantile = 3.462
95% family-wise confidence level
$mca
    estimate stderr   lower upper height
1-2     2.00 0.6693 -0.3173 4.317  13.25
1-4     3.25 0.6693  0.9327 5.567  12.62
1-3     3.50 0.6693  1.1827 5.817  12.50
2-4     1.25 0.6693 -1.0673 3.567  11.63
2-3     1.50 0.6693 -0.8173 3.817  11.50
4-3     0.25 0.6693 -2.0673 2.567  10.88
$none
  estimate stderr  lower upper height
1    14.25 0.4732 12.611 15.89  14.25
2    12.25 0.4732 10.611 13.89  12.25
4    11.00 0.4732  9.361 12.64  11.00
3    10.75 0.4732  9.111 12.39  10.75
$lmat
      estimate stderr  lower upper height
1-43     3.375 0.5796  1.368 5.382  12.56
2-143    0.250 0.5465 -1.642 2.142  12.13
4-3      0.250 0.6693 -2.067 2.567  10.88

>   contrasts(tires$brand) <- brand.lmat

>   tires.aov <- aov(wear ~ car + position + brand, data=tires)

>   summary(tires.aov, split=list(brand=list("1-43"=1, rest=2:3)))
              Df Sum Sq Mean Sq F value Pr(>F)
car            3   38.7   12.90   14.40 0.0038
position       3    6.2    2.06    2.30 0.1769
brand          3   30.7   10.23   11.42 0.0068
  brand: 1-43  1   30.4   30.37   33.91 0.0011
  brand: rest  2    0.3    0.16    0.17 0.8441
Residuals      6    5.4    0.90
```

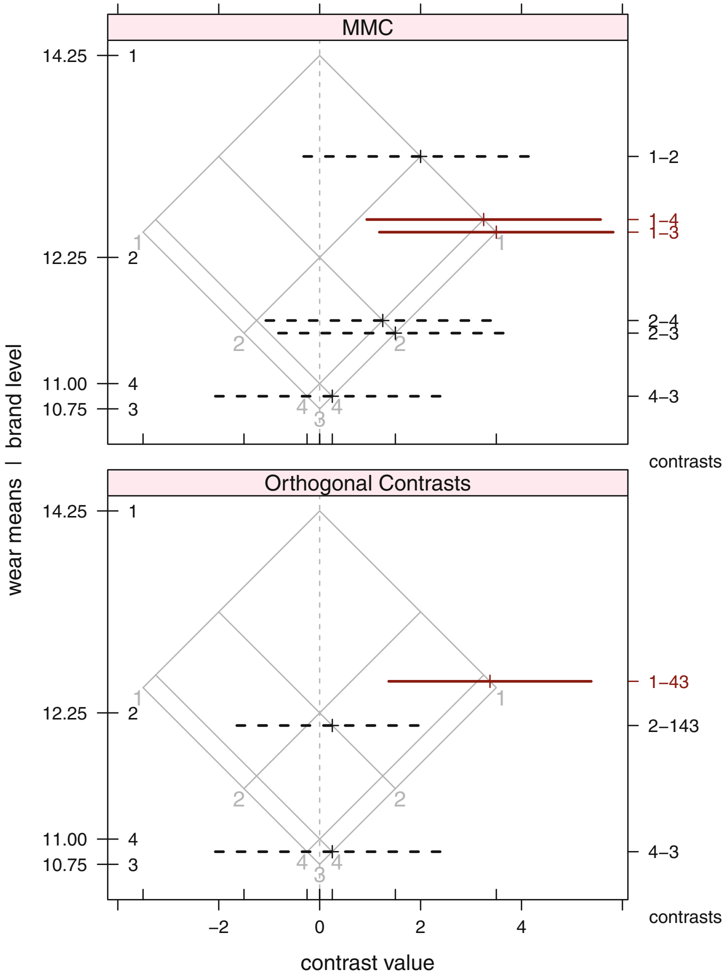



**Fig. 13.7** Tires data. The top panel shows the Tukey 95% intervals for all pairwise contrasts of wear means at the four levels of `brand`. The Tukey procedure leads to the conclusions illustrated by the orthogonal contrasts in the bottom panel. Brand 1 shows significantly greater mean `wear` than brand 3 or brand 4. The Tiebreaker plot is not needed in this example.

The results of the $F$-test on a blocking factor are not ordinarily presented in the discussion because block differences are expected, and multiple comparisons on block means are not usually performed. Nevertheless, when blocks are significant, it is an indication that the blocking was worthwhile. Most experimental design texts contain formulas for the efficiency attributable to blocking in Latin square, randomized complete block, and other experimental designs; see, for example, Cochran and Cox (1957) (Section 4.37).

We continue with this example in Exercise 13.6 where we use dummy variables to illustrate the linear dependence of the treatment (`brand`) sum of squares on the interaction of the two blocking factors.

## 13.3 Simple Effects for Interaction Analyses

When a low $p$-value for an interaction term in an ANOVA table leads the analyst to believes that interaction exists, it is necessary to study the nature of the interaction.

We re-emphasize that in this situation, tests of the main effects of the factors comprising the interaction are inappropriate. Instead we seek to analyze the *simple effects*, which we now define. An analysis of simple effects, along with interaction plots of cell means which we've previously discussed, are the correct tools for investigating interaction. We note which simple effects are appreciable, either by examining confidence intervals on the simple effects or by testing whether the simple effects are zero. Since such activities involve simultaneous inferences, it is desirable to give attention to use either simultaneous confidence levels or a familywise error rate for simultaneous tests. The importance of studying individual simple effects rather than the overall interaction effect is comparable to the ecological fallacy introduced in Section 4.2 and the cautions resulting from Simpson's paradox, discussed in Section 15.3.

We confine attention here to the case of two factors, say $A$ and $B$, at $a$ and $b$ levels, respectively. Continuing with the notation of Equation (12.1), let $\mu_{ij}$ denote the mean response of the treatment combination where $A$ is at its $i^{\text{th}}$ level and $B$ is at its $j^{\text{th}}$ level. Then a simple effect for $B$ is a pairwise comparison of levels of $B$ at a particular level of $A$, for example, $\mu_{12} - \mu_{13}$. Similarly, a simple effect for $A$ is a pairwise comparison of levels of $A$ at a particular level of $B$, such as $\mu_{32} - \mu_{12}$.

This assumes that the levels of a factor for which we are calculating simple effects are qualitative in nature. If instead the levels of factor $B$ are quantitative, then a different analysis is called for, namely, a comparison of comparable polynomial contrasts of the cell means at each level of factor $A$. This analysis is superior to performing $a$ separate one-way analyses comparing the levels of $B$ because we are pooling the information from all levels of $A$ to estimate the common error variance. This enables us to compare the levels of $B$ with maximum available power.

In experiments with three or more factors, there is a potential for 3-factor interaction. If there are three factors $(A, B, C)$ that interact, this may be interpreted as saying that the nature of the $AB$ interaction varies according to the particular level of factor $C$. An analysis of such an interaction is more complicated than is the case for two interacting factors.

### 13.3.1 Example—The `filmcoat` Data

We illustrate the use of simple effects in analyzing interaction with the dataset `data(filmcoat)` from Iman (1994, pp. 768–778).

### 13.3.2 Study Objectives

> Chemical vapor deposition is a process used in the semiconductor industry to deposit thin films of silicon dioxide and photoresist on substrates of wafers as they are manufactured. The films must be as thin as possible and have a uniform thickness, which is measured by a process called infrared interference.
>
> A process engineer evaluated a low-pressure chemical vapor deposition (LPCVD) process that reduces costs and increases productivity. The engineer set up an experiment to study the effect of chamber temperature and pressure on film thickness. Three temperatures and three pressures were selected to represent the low, medium, and high levels of operating conditions for both factors. The experiment was conducted by randomly selecting one of the temperature–pressure combinations and determining the thickness of the film coating after processing is completed. This experiment was repeated three times with each temperature–pressure combination. The engineer wanted to determine the joint effect of temperature and pressure on the mean film thickness. The response was thickness (in Ångström units) of film coatings applied to wafers.

### 13.3.3 Data Description

`temprt`: temperature: low, medium, and high levels

`pressure`: pressure: low, medium, and high levels

`coat`: thickness of film coat

The data are displayed in Table 13.9 and Figure 13.8. The table of means and ANOVA table are in Table 13.10. The plots of the means and interactions are in Figure 13.9.

We observe that the `temprt` × `pressure` interaction is moderately significant. Therefore, conclusions about which level of `temprt` tends to minimize `coat` depend on the level of `pressure`. This statement is supported by Figure 13.9, which suggests that for low and high `pressure`, the response `coat` is minimized at medium `temprt` while for medium `pressure`, `coat` is minimized at high `temprt`. In addition, it is suggested that for low and medium `temprt`, `coat` is minimized at low `pressure` while for high `temprt`, `coat` is minimized at medium `pressure`.

**Table 13.9** `filmcoat` data. Thickness of film coat at various settings of temperature and pressure.

| | Pressure | | |
|---|---|---|---|
| Temperature | Low | Medium | High |
| Low | 42, 43, 39 | 45, 43, 45 | 45, 44, 47 |
| Medium | 36, 34, 37 | 39, 39, 37 | 40, 42, 38 |
| High | 38, 37, 37 | 35, 36, 33 | 40, 41, 42 |

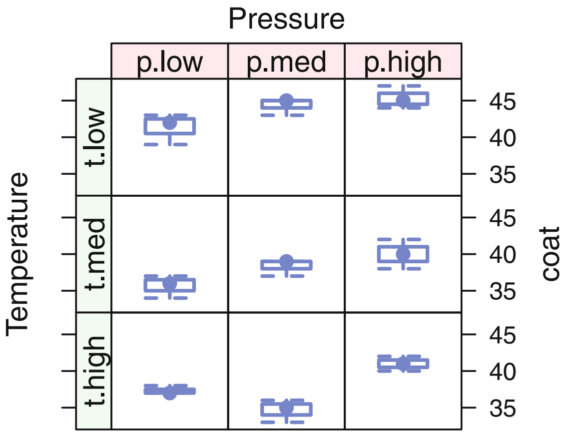



**Fig. 13.8** a Filmcoat data. Each box is based on three data points. The interaction between temperature and pressure is very visible in the `t.high ~ p.med` panel. Along the top two rows of the plot, the boxes move up from left to right. In the bottom row, the second box is below the other two.

### *13.3.4 Data Analysis*

These informal visual impressions are formally investigated by examining the simultaneous confidence intervals on the simple effects displayed in Figures 13.10 and 13.11. Control of the simultaneous confidence level at 95% within each of the two sets of nine intervals is maintained by using simulation-generated critical points for this procedure as recommended by Edwards and Berry (1987). These simultaneous confidence intervals are produced in R with the `glht` function in the **multcomp** package. The simultaneous confidence of the collection of all 18 confidence intervals is closer to 90%.

**Table 13.10** Means and ANOVA table for `filmcoat` data. The moderately significant interaction between pressure and temperature requires that we examine simple effects rather than main effects.

```
>   reshape2::acast(filmcoat, temprt ~ pressure, mean,
+                   value.var="coat", margins=TRUE)
           p.low    p.med   p.high    (all)
t.low  41.33333 44.33333 45.33333 43.66667
t.med  35.66667 38.33333 40.00000 38.00000
t.high 37.33333 34.66667 41.00000 37.66667
(all)  38.11111 39.11111 42.11111 39.77778

>   film.aov1 <- aov(coat ~ temprt*pressure, data=filmcoat)

>   summary(film.aov1)
                Df Sum Sq Mean Sq F value   Pr(>F)
temprt           2 204.67  102.33  47.638 6.46e-08 ***
pressure         2  78.00   39.00  18.155 4.83e-05 ***
temprt:pressure  4  37.33    9.33   4.345   0.0124 *
Residuals       18  38.67    2.15
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

We examine which of these intervals excludes zero and for those that do, whether the interval lies above or below zero. In Figure 13.10, we see that at medium `pressure`,

- the confidence interval on mean `coat` at low `temprt` minus mean `coat` at high `temprt` lies entirely above zero,
- the confidence interval on mean `coat` at medium `temprt` minus mean `coat` at high `temprt` is closer to zero, but is also entirely above zero.

From these two statements we can conclude that at medium `pressure`, mean `coat` is minimized at high `temprt`. This is the only firm conclusion we can draw about `coat` minimization because many of the other intervals in these two figures overlap zero, indicating nonsignificant differences of means. We are unable to formally confirm our other graphical impressions that for low and high `pressure`, the response `coat` is minimized at medium `temprt`. Nor are we able to confirm our initial graphical impressions about levels of `pressure` that minimize `coat` at each level of `temprt`.

In summary, while we can make some confident assertions about differences in coating between some of the combinations of temperature and pressure, it is not possible to infer from these data an overall recommendation of the optimal combination of temperature and pressure. It is possible that a larger experiment would have led to such a conclusion.

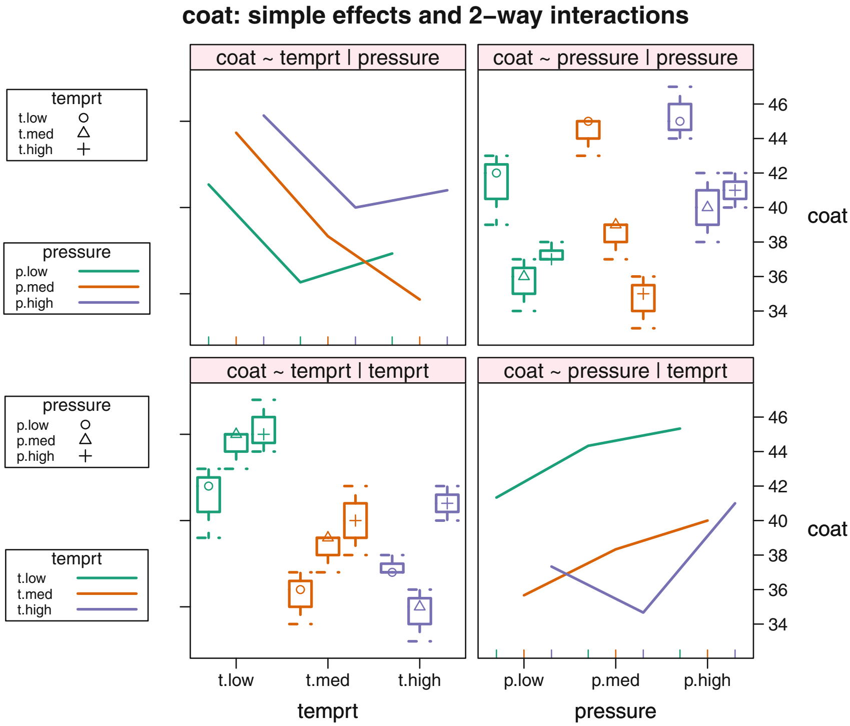



**Fig. 13.9** Simple effects plot for `filmcoat` data. At medium `pressure`, mean `coat` is minimized at high `temprt`. No other firm conclusions can be drawn because many of the simultaneous confidence intervals in Figures 13.10 and 13.11 overlap zero. We saw in Section 13.3 that main effects are not defined in the presence of interaction. The simple effects plot here shows individual boxplots for each level of pressure conditioned on the level of temperature in the lower left panel. The plot shows individual boxplots for each level of temperature conditioned on the level of pressure in the upper right panel.

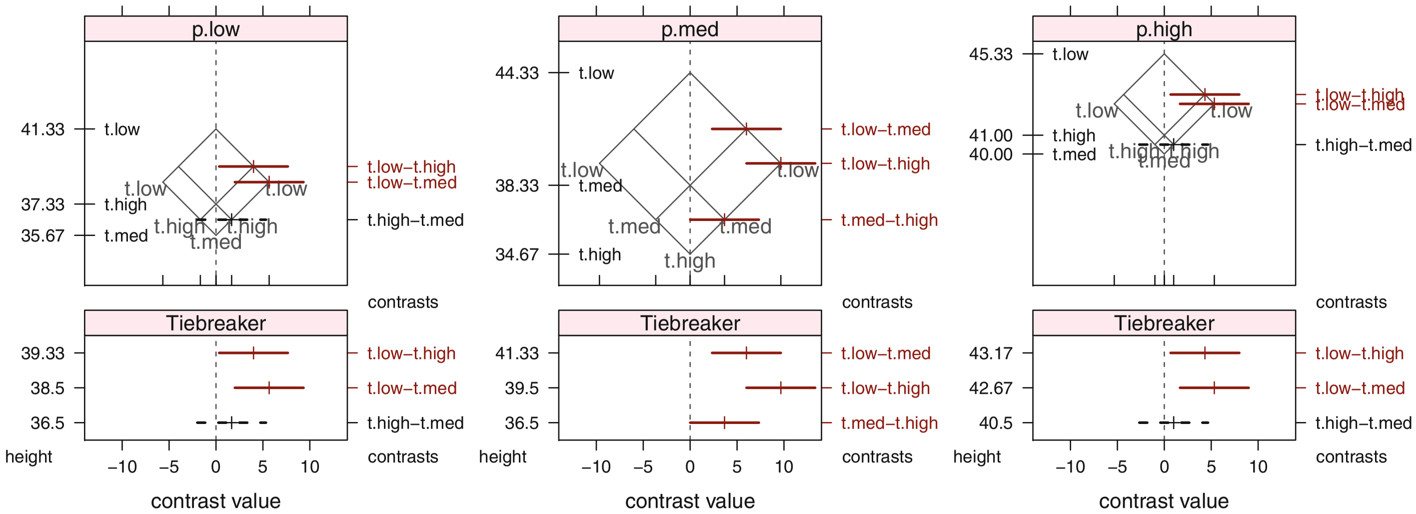



**Fig. 13.10** Simultaneous 95% simulation-based confidence intervals for the `filmcoat` data: Simple effects of temperature at each level of pressure. The Tiebreaker plots in the bottom panels are imperative for this example.

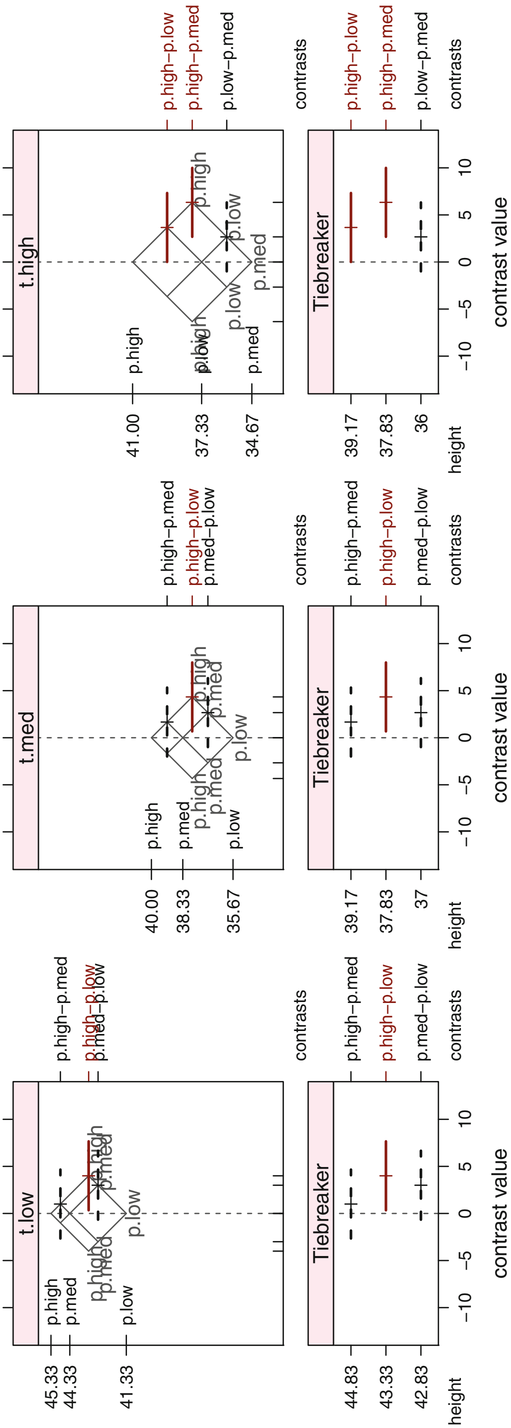



**Fig. 13.11** Simultaneous 95% simulation-based confidence intervals for the `filmcoat` data: Simple effects of pressure at each level of temperature. The Tiebreaker plots in the bottom panels are imperative for this example.

## 13.4 Nested Factorial Experiment

Thus far we have considered situations where the relationships among the factors are either completely crossed or completely nested. It is also possible to have an experiment with three or more factors having both crossed and nesting relationships. Such an arrangement is called a nested factorial experiment.

### *13.4.1 Example—Gunload Data*

We illustrate one possible arrangement with an example taken from Hicks and Turner (1999). It was desired to improve the number of rounds per minute that could be fired from a large naval gun. There are two levels of loading `method`, 1=new and 2=old, and three `groups` defining the physiques of the loaders, 1=slight, 2=average, 3=heavy. From each of these groups the experimenter selected three equal sized `teams` of men. Thus there are three teams of men having slight build, three teams of men having average build, and three teams of men having heavy build. Using both of the two methods, each of the nine teams fired the gun on two separate occasions. It is seen that `team` is nested within `group`, and that `method` is crossed with both `group` and `team` within `group`. The factors `method` and `group` are fixed, while `team` is a random factor. The data are contained in the dataset `data(gunload)`. We display the data in Figure 13.12.

If all three factors were fixed factors, then the residual mean square would serve as the denominator for all analysis of variance table $F$-tests on main effects and interactions. When at least one factor is random, the $F$-test denominators are sometimes another mean square in the ANOVA table. As with our use of Table 12.8 to determine the correct denominator for an analysis with two crossed factors where one or both could be random factors, we construct Table 13.11 to aid in our analysis of the `gunload` data. The table is constructed by writing the sums of squares as quadratic forms in the $Y_{ijkl}$ defined in Equation (13.1), and using Equation (I.6) in Appendix I for finding the expected values of these quadratic forms. Then as with Table 12.8, the ANOVA $F$-test of any effect uses as the denominator the mean square having expectation identical to the expected mean square of the effect, apart from the term for the effect itself.

In Table 13.11, we have three factors, which we call M, G, and T for easy association with `method`, `group`, and `team` in the `gunload` example. We use the corresponding Greek letters $\theta$ (for *meTHod*), $\gamma$, and $\tau$ for the population effects. Here T is nested within G, and M is crossed with both G and T. In this table, the factors M, G, and T have $m, g$, and $t$ levels, respectively, and the common number of replications of each treatment combination is $n$. In the `gunload` example, there are $n = 2$ occasions.

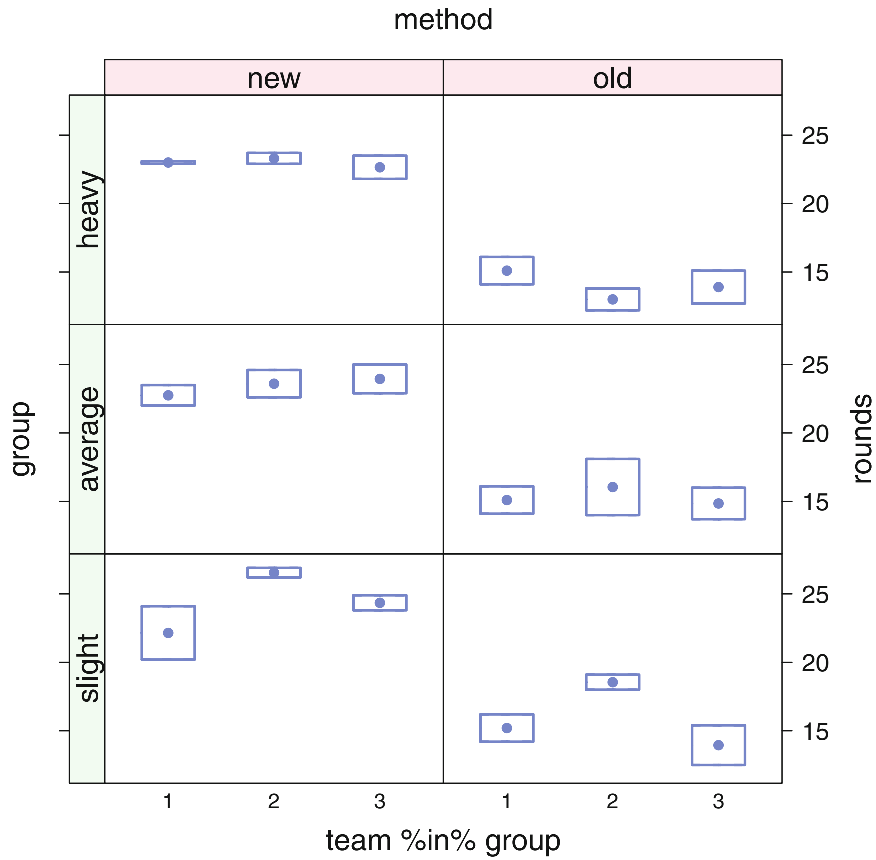



**Fig. 13.12** Boxplot of gunload data. The response `rounds` is higher for the new `method` than the old `method` and does not appear to differ across the three physique `groups`. These findings are consistent with the small $p$-value for `method` and the large $p$-value for `group` in Table 13.13.

The statistical model associated with this analysis may be written as

$$Y_{ijkl} = \mu + \gamma_i + \tau_{j(i)} + \theta_k + (\gamma\theta)_{ik} + (\tau\theta)_{jk(i)} + \epsilon_{ijkl} \tag{13.1}$$

Here $\gamma_i$ is the effect of `group` level $i$, $\tau_{j(i)}$ is the effect of level $j$ of `team` nested within level $i$ of `group`, $\theta_k$ represents level $k$ of `method`, $(\gamma\theta)_{ik}$ represents the interaction of `group` and `method`, $(\tau\theta)_{jk(i)}$ represents the interaction of `team` and `method` within `group` level $i$, and $\epsilon_{ijkl}$ is the residual error.

For ease of presentation, we use the convention that

$$\sigma_A^2 = \frac{\sum_i \alpha_i^2}{a-1} \quad \text{if A is a fixed factor}$$

and

$$\sigma^2_{AB} = \frac{\sum_{ij}(\alpha\beta)^2_{ij}}{(a-1)(b-1)} \quad \text{if A and B are both fixed factors.}$$

We use the indicator function $I_A$ defined as 1 if A is a random factor or 0 if A is a fixed factor. (This is not the same convention we used in Table 6.4. There we used the notation $\sigma^2_A$ for random effects and $\kappa^2_A$ for fixed effects.)

We illustrate the use of Table 13.11 by considering the test of the main effect for `group`, factor G. Since in this example `method` is fixed and `team` is random, we have $I_M = 0$ and $I_T = 1$. Therefore, the expected value of the mean square for G is

$$\sigma^2 + nm\sigma^2_T + nmt\sigma^2_G \tag{13.2}$$

and the expected value of the mean square for T nested in G is

$$\sigma^2 + nm\sigma^2_T \tag{13.3}$$

The ratio of these mean squares is appropriate for testing equality of the levels of `group`, factor G, because the corresponding ratio of these expected mean squares exceeds one if and only if $\sigma^2_G > 0$. Use of the residual mean square as the denominator of the $F$-test would be inappropriate because such a ratio would exceed one if there is a G effect, a T (`team`) effect, or both effects.

Note that if instead `method` were a random factor and `team` were a fixed factor, the pattern of expected mean squares would be quite different from those in Table 13.11, with different denominators appropriate for some of the ANOVA $F$-tests.

The model specifications for the sums of squares in the gunload example are shown for both R and SAS in Table 13.12. Discussion of the operators in this table appears in Section 13.5 and Table 13.18.

We overrode the default choices of the denominator mean squares for the $F$-tests for `method`, `group`, and `method*group`. These new choices are necessitated by the facts that one of the factors is random and there are both mixing and crossing of factors. Our conclusions here are that after correcting for both loaders' physiques and other person-to-person differences, the two methods have significantly different loading speeds. The new method averaged 23.59 rounds per minute compared with 15.08 rounds per minute for the old method. The analysis also shows a secondary finding that loading times do not differ significantly across physique `groups`.

**Table 13.11** Expected mean squares for a three-factor nested factorial ANOVA. See also Tables 6.4 and 12.8.

| Source | df | Expected Mean Square | |
|---|---|---|---|
| G | $g-1$ | $\sigma^2 + nI_M\sigma^2_{TM} + ntI_M\sigma^2_{GM}$ | $+ nmI_T\sigma^2_T + nmt\sigma^2_G$ |
| T within G | $g(t-1)$ | $\sigma^2 + nI_M\sigma^2_{TM}$ | $+ nmI_T\sigma^2_T$ |
| M | $m-1$ | $\sigma^2 + nI_T\sigma^2_{TM} + ntI_G\sigma^2_{GM} + ngt\sigma^2_M$ | |
| GM | $(m-1)(g-1)$ | $\sigma^2 + nI_T\sigma^2_{TM} + nt\quad\sigma^2_{GM}$ | |
| TM within G | $g(m-1)(t-1)$ | $\sigma^2 + n\quad\sigma^2_{TM}$ | |
| Residual | $mgt(n-1)$ | $\sigma^2$ | |
| Total | $mgtn-1$ | | |

**Table 13.12** Nested factorial model specifications in R and SAS. Specifications for the sums of squares for the gunload example are shown here. The tests are specified separately. In Table 13.13 we show use of the `Error` function in R.

| | | | | | | | |
|---|---|---|---|---|---|---|---|
| Algebra | $Y_{ijkl} = \mu +$ | $\gamma_i +$ | $\tau_{j(i)}$ | $+ \theta_k$ | $+ (\gamma\theta)_{ik} +$ | $(\tau\theta)_{jk(i)}$ | $+ \epsilon_{ijkl}$ |
| R | `Y ~` | `G` | `+ T%in%G` | `+ M` | `+ G:M` | `+ T:M%in%G` | |
| SAS | `Y =` | `G` | `T(G)` | `M` | `G*M` | `T*M(G)` | |

## 13.4.2 Example—Turkey Data (Continued)

We continue the discussion of the turkey data `data(turkey)` from Section 6.8.

Contrasts of the form used in Table 6.8 are so important in the design of experiments and in their analysis that we have a simple terminology and notation to describe them. In this experiment there are three distinct factors, one with two levels and two with three levels:

`trt.vs.control:` with levels `control` and `treatment`

`additive:` with levels `control`, `A`, and `B`

`amount:` with levels `0`, `1`, and `2`

occurring in the pattern shown in Table 13.14. The algebraic formula describing the model is

$$Y_{mijk} = \mu + \tau_m + \alpha_i + \beta_j + (\alpha\beta)_{ij} + \epsilon_{ijk} = \mu_{ij} + \epsilon_{mijk} \tag{13.4}$$

**Table 13.13** Gunload data. The $F$-tests that appear without the `Error` function are incorrect. Here we produce correct tests by using the `Error` function to override the default choice of denominators of $F$-tests.

```
> gunload.aov <-
+     aov(rounds ~ method*group + Error((team %in% group)/method),
+         data=gunload)

> summary(gunload.aov)

Error: team:group
          Df Sum Sq Mean Sq F value Pr(>F)
group      2  16.05   8.026   1.227  0.358
Residuals  6  39.26   6.543

Error: team:group:method
             Df Sum Sq Mean Sq F value   Pr(>F)
method        1  652.0   652.0 364.841 1.33e-06 ***
method:group  2    1.2     0.6   0.332     0.73
Residuals     6   10.7     1.8
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Error: Within
          Df Sum Sq Mean Sq F value Pr(>F)
Residuals 18  41.59   2.311

> model.tables(gunload.aov, type="means")
Tables of means
Grand mean

19.33333

 method
method
   new    old
23.589 15.078

 group
group
 slight average   heavy
 20.125  19.383  18.492

 method:group
      group
method slight average heavy
   new 24.350 23.433  22.983
   old 15.900 15.333  14.000
```

Several issues are raised here to be discussed. How do these factors relate to each other? How does describing a design in terms of the factors specify the analysis?

Factors can be related in several ways (see Table 13.18). In the turkey example, we illustrate two relations: crossing and nesting.

crossing: Every level of `additive` appears at every level of `amount`. In this example, Additive A appears at Amounts 1 and 2, as does Additive B.

nesting: Some Additive–Amount combinations (A1, A2, B1, B2) appear in only the treatment level of `trt.vs.control`. Other Additive–Amount combinations (control-0) appear in only the control level of `trt.vs.control`. The factors `additive` and `amount` are then said to be *nested* within the factor `trt.vs.control`.

When we add these factors to the dataset, for example with commands in Table 13.15, we can write a much simpler model formula that automatically produces the easily readable ANOVA table in Table 13.16. Notice that the four 1-degree-of-freedom sums of squares in Table 13.16 are a decomposition of the 4-degree-of-freedom sum of squares in Table 6.7. The significance of the corresponding $F$-test in Table 6.7 is a rationale for producing and interpreting Table 13.16. We illustrate the structure with the table of means in Table 13.17 and the boxplots in Figure 13.13. See the discussion on orientation of boxplots in Section 13.A.

**Table 13.14** Factor structure for turkey data.

| Treatment | Level 0 | Level 1 | Level 2 | Trt.vs.Cont | Treatment | Level 0 | Level 1 | Level 2 |
|---|---|---|---|---|---|---|---|---|
| control | × | | | C | control | × | | |
| A | | × | × | T | A | | × | × |
| B | | × | × | T | B | | × | × |

In the turkey example there does not seem to be serious interaction ($p \approx .06$). In other situations the interaction dominates the analysis. An example with prominent interaction is the analysis of the *Rhizobium* clover data in Section 12.14.7.

**Table 13.15** R commands to create factors for the turkey data introduced in Section 6.8.

```
> data(turkey)

> turkey[c(1,7,13,19,25),]
      diet wt.gain
1  control     4.1
7       A1     5.2
13      A2     6.3
19      B1     6.5
25      B2     9.5

> turkey$trt.vs.control <-
+    factor(rep(c("control","treatment"), c(6,24)))

> contrasts(turkey$trt.vs.control) <- c(4,-1)

> turkey$additive <- factor(rep(c("control","A","B"), c(6,12,12)),
+                            levels=c("control","A","B"))

> contrasts(turkey$additive) <- c(0,1,-1)

> turkey$amount <- factor(rep(c(0,1,2,1,2), c(6,6,6,6,6)))

> contrasts(turkey$amount) <- c(0,1,-1)

> turkey[c(1,7,13,19,25),]
      diet wt.gain trt.vs.control additive amount
1  control     4.1        control  control      0
7       A1     5.2      treatment        A      1
13      A2     6.3      treatment        A      2
19      B1     6.5      treatment        B      1
25      B2     9.5      treatment        B      2
```

**Table 13.16** ANOVA for turkey data with the crossing of the `additive` and `amount` factors nested within the `trt.vs.control` factor. Interaction is borderline nonsignificant. The main effects of `additive` and `amount` nested within `trt.vs.control` are significant.

```
> turkey3.aov <- aov(wt.gain ~ trt.vs.control / (additive*amount),
+                     data=turkey, x=TRUE)

> summary(turkey3.aov)
                               Df Sum Sq Mean Sq F value   Pr(>F)
trt.vs.control                  1  56.58   56.58 179.395 6.58e-13 ***
trt.vs.control:additive         1  22.81   22.81  72.337 7.56e-09 ***
trt.vs.control:amount           1  22.43   22.43  71.105 8.88e-09 ***
trt.vs.control:additive:amount  1   1.21    1.21   3.852   0.0609 .
Residuals                      25   7.88    0.32
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

**Table 13.17** Means for turkey data.

```
> print(na.print="",
+ tapply(turkey$wt.gain,
+        turkey[,c("additive","amount")],
+        mean)
+ )
         amount
additive         0   1        2
  control 3.783333
  A                5.5 6.983333
  B                7.0 9.383333
```

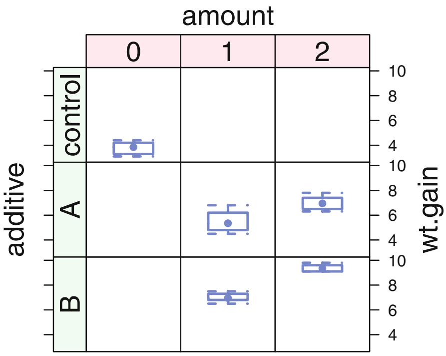



**Fig. 13.13** Turkey data with factor structure. The main effect for `amount` is visible as the consistent increase in `wt.gain` as `amount` goes from 1 to 2. The main effect for `additive` is visible as the consistent increase in `wt.gain` as `additive` goes from A to B. The hint of borderline non-significant Interaction is seen as the different slope connecting the median dots in the A row and the B row of the display. This figure shows the boxplots in vertical orientation. Figure 13.17 shows the same boxplots in horizontal orientation. We discuss orientation in Section 13.A.

## 13.5 Specification of Model Formulas

Dummy variables (discussed in Section 10.1), and the contrasts they code for, are so important that all statistical languages have constructs for describing them and the relations between them. The model specification operators in R and SAS are detailed in Table 13.18.

Let us explore the meaning of the concepts of crossed and nested factors with a set of simple examples using the data in Table 13.19. The dataset `data(abc)` has two factors, *A* with three levels and *B* with four levels. Tables 13.20 and 13.21 show visual interpretations of the structure of the dataset with several different assumptions about the relation of the factors.

We show the model formula specifications, the generated dummy variables, the ANOVA tables, and the estimated $a_i$, $b_j$, $(ab)_{ij}$, and $b_{j(i)}$ values for several differently structured models. Table 13.22 contains the complete R input for these models. We recommend that you read these examples closely and experiment with them on your computer.

The simplest set of dummy variables (that is, easiest to understand) is the set of treatment contrasts. The most frequently used is the set of sum contrasts. We show both in Table 13.23.

**Table 13.18** Model specification operators. In R, the two notations `a:b` and `b%in%a` are equivalent.

| | R | | SAS | | Algebra |
|---|---|---|---|---|---|
| | abbrev | expanded | abbrev | expanded | |
| Double index | `a:b` | | `a*b` | | $(ab)_{ij}$ |
| Sum | `a+b` | | `a b` | | $a_i + b_j$ |
| Cross | `a*b` | `a+b+a:b` | `a\|b` | `a b a*b` | $a_i + b_j + (ab)_{ij}$ |
| Nested | `b%in%a` | | `b(a)` | `a*b` | $b_{j(i)}$ |
| Nest | `a/b` | `a + b%in%a` | `a b(a)` | `a a*b` | $a_i + b_{j(i)}$ |

**Table 13.19** Sample data used to explore concepts of crossed and nested factors. Factor `A` has three levels and factor `B` has four levels. The interaction `AB` has 12 levels named by crossing the level names of `A` and `B`. The nested factor `BwA` (`B` within `A`) has 12 levels named without reference to factor `A`.

| obs | A | B | AB | BwA | y |
|---|---|---|---|---|---|
| r.w | r | w | r.w | c | 0.17 |
| r.x | r | x | r.x | d | 2.25 |
| r.y | r | y | r.y | e | −1.57 |
| r.z | r | z | r.z | f | −1.55 |
| s.w | s | w | s.w | g | −0.24 |
| s.x | s | x | s.x | h | 1.71 |
| s.y | s | y | s.y | i | 0.38 |
| s.z | s | z | s.z | j | −1.26 |
| t.w | t | w | t.w | k | 0.34 |
| t.x | t | x | t.x | l | −0.15 |
| t.y | t | y | t.y | m | −1.70 |
| t.z | t | z | t.z | n | −1.93 |

**Table 13.20** Rearrangements of abc data to show different assumptions about the relation of the factors: data, one factor, two factors crossed.

```
> data(abc)

> abc
    A B  AB BwA     y
r.w r w r.w   c -0.02
r.x r x r.x   d  1.19
r.y r y r.y   e -0.02
r.z r z r.z   f  0.23
s.w s w s.w   g  0.67
s.x s x s.x   h  1.95
s.y s y s.y   i -0.71
s.z s z s.z   j -0.40
t.w t w t.w   k -0.56
t.x t x t.x   l  0.01
t.y t y t.y   m  0.13
t.z t z t.z   n  1.19

> abc.oneway <- ## one-way
+ with(abc,
+      matrix(y, 4, 3, dimnames=list(1:4, A=unique(A)))
+      )

> abc.oneway
   A
        r     s     t
  1 -0.02  0.67 -0.56
  2  1.19  1.95  0.01
  3 -0.02 -0.71  0.13
  4  0.23 -0.40  1.19

> abc.crossed <- ## crossed
+ with(abc,
+      matrix(y, 3, 4, byrow=TRUE,
+             dimnames=list(A=unique(A), B=unique(B)))
+      )

> abc.crossed
   B
A       w    x     y     z
  r -0.02 1.19 -0.02  0.23
  s  0.67 1.95 -0.71 -0.40
  t -0.56 0.01  0.13  1.19
```

**Table 13.21** Rearrangements of abc data to show different assumptions about the relation of the factors: two factors nested, two factors doubly indexed.

```
> abc.nested <- ## nested
+ with(abc,
+      matrix(c(y[1:4],    rep(NA,8),
+               rep(NA,4), y[5:8],    rep(NA,4),
+               rep(NA,8),            y[9:12]),
+               3, 12, byrow=TRUE,
+               dimnames=list(A=unique(A), BwA=BwA))
+             )

> print(abc.nested, na.print="")
   BwA
A       c    d     e    f    g    h     i    j     k    l    m    n
  r -0.02 1.19 -0.02 0.23
  s                       0.67 1.95 -0.71 -0.4
  t                                             -0.56 0.01 0.13 1.19

> abc.double.indexed <- ## doubly-indexed
+   abc[,"y",drop=FALSE]

> abc.double.indexed
        y
r.w -0.02
r.x  1.19
r.y -0.02
r.z  0.23
s.w  0.67
s.x  1.95
s.y -0.71
s.z -0.40
t.w -0.56
t.x  0.01
t.y  0.13
t.z  1.19
```

**Table 13.22** Specification of several models using one or both variables in the abc dataset.

```
## one-way
abc.A.aov <- aov(y ~ A, data=abc)
anova(abc.A.aov)
coef(abc.A.aov)
contrasts(abc$A)
model.matrix(abc.A.aov)


## crossed: no interaction
abc.ApB.aov <- aov(y ~ A+B, data=abc)
anova(abc.ApB.aov)
coef(abc.ApB.aov)
contrasts(abc$A)
contrasts(abc$B)
model.matrix(abc.ApB.aov)


## crossed: with interaction
abc.AsB.aov <- aov(y ~ A*B, data=abc)
anova(abc.AsB.aov)
coef(abc.AsB.aov)
contrasts(abc$A)
contrasts(abc$B)
contrasts(abc$AB)
model.matrix(abc.AsB.aov)


## nested
abc.BwA.aov <- aov(y ~ A/B, data=abc)
anova(abc.BwA.aov)
coef(abc.BwA.aov)
contrasts(abc$A)
contrasts(interaction(abc$A, abc$B))
model.matrix(abc.BwA.aov)


## doubly-indexed
abc.AB.aov <- aov(y ~ AB, data=abc)
anova(abc.AB.aov)
coef(abc.AB.aov)
contrasts(abc$AB)
model.matrix(abc.AB.aov)
```

**Table 13.23** These dummy variables are constructed for a fit of the form $\hat{y}_i = m + a_i$ to a model of the form $y_{ij} = \mu + \alpha_i + \epsilon_{ij}$. With treatment contrasts, $m$ is an estimate of $\mu + \alpha_r$, $a_s$ is an estimate of $\alpha_s - \alpha_r$, and $a_t$ is an estimate of $\alpha_t - \alpha_r$. With sum contrasts, $m$ is an estimate of $\mu$, $a_1$ is an estimate of $\alpha_r$, and $a_2$ is an estimate of $\alpha_s$. There is no need for an $a_3$ because of the constraint on the parameters $\alpha_t = -(\alpha_r + \alpha_s)$.

```
> model.matrix(~A, data=abc,
+     contrasts=
+       list(A=contr.treatment))
    (Intercept) A2 A3
r.w           1  0  0
r.x           1  0  0
r.y           1  0  0
r.z           1  0  0
s.w           1  1  0
s.x           1  1  0
s.y           1  1  0
s.z           1  1  0
t.w           1  0  1
t.x           1  0  1
t.y           1  0  1
t.z           1  0  1
attr(,"assign")
[1] 0 1 1
attr(,"contrasts")
attr(,"contrasts")$A
  2 3
r 0 0
s 1 0
t 0 1
```

```
> model.matrix(~A, data=abc,
+     contrasts=
+       list(A=contr.sum))
    (Intercept) A1 A2
r.w           1  1  0
r.x           1  1  0
r.y           1  1  0
r.z           1  1  0
s.w           1  0  1
s.x           1  0  1
s.y           1  0  1
s.z           1  0  1
t.w           1 -1 -1
t.x           1 -1 -1
t.y           1 -1 -1
t.z           1 -1 -1
attr(,"assign")
[1] 0 1 1
attr(,"contrasts")
attr(,"contrasts")$A
  [,1] [,2]
r    1    0
s    0    1
t   -1   -1
```

### *13.5.1 Crossing of Two Factors*

We provide a detailed discussion here of just one model, the crossing of two factors. The dummy variables for interaction in the crossing model

| | |
|---|---|
| Algebra: | $y_i = \mu + \alpha_i + \beta_j + (\alpha\beta)_{ij} + \epsilon_{ij}$ |
| R: | `Y ~      A + B + A:B` |
| SAS: | `Y =      A   B   A*B` |

are constructed as the outer product of the rows of the dummy variables for each of the main effects. We illustrate in Table 13.24 with the sum contrasts.

**Table 13.24** Dummy variables for the interaction $(\alpha\beta)_{ij}$ constructed as the outer product of the rows of the dummy variables for the two main effects `A` and `B`. We continue with the data of Table 13.19 and the sum contrasts defined in the right side of Table 13.23. For example, in row `r.z` the value in column `A1:B1` is the product of the 1 in column `A1` and the $-1$ in column `B1`. There are two degrees of freedom, hence two dummy variables for the `A` effect. There are three dummy variables for the `B` effect. Therefore, there are $2 \times 3$ dummy variables for the `A:B` interaction.

```
> old.width <- options(width=70)

> mm <- model.matrix(~A*B, data=abc,
+                   contrasts=list(A=contr.helmert, B=contr.helmert))

> mm[,]
    (Intercept) A1 A2 B1 B2 B3 A1:B1 A2:B1 A1:B2 A2:B2 A1:B3 A2:B3
r.w           1 -1 -1 -1 -1 -1     1     1     1     1     1     1
r.x           1 -1 -1  1 -1 -1    -1    -1     1     1     1     1
r.y           1 -1 -1  0  2 -1     0     0    -2    -2     1     1
r.z           1 -1 -1  0  0  3     0     0     0     0    -3    -3
s.w           1  1 -1 -1 -1 -1    -1     1    -1     1    -1     1
s.x           1  1 -1  1 -1 -1     1    -1    -1     1    -1     1
s.y           1  1 -1  0  2 -1     0     0     2    -2    -1     1
s.z           1  1 -1  0  0  3     0     0     0     0     3    -3
t.w           1  0  2 -1 -1 -1     0    -2     0    -2     0    -2
t.x           1  0  2  1 -1 -1     0     2     0    -2     0    -2
t.y           1  0  2  0  2 -1     0     0     0     4     0    -2
t.z           1  0  2  0  0  3     0     0     0     0     0     6

> print(AA <- mm["s.y", c("A1","A2")])
A1 A2
 1 -1

> print(BBB <- mm["s.y", c("B1","B2","B3")])
B1 B2 B3
 0  2 -1

> outer(AA, BBB)
   B1 B2 B3
A1  0  2 -1
A2  0 -2  1

> as.vector(outer(AA, BBB))
[1]  0  0  2 -2 -1  1

> mm["s.y", c("A1:B1","A2:B1","A1:B2","A2:B2","A1:B3","A2:B3")]
A1:B1 A2:B1 A1:B2 A2:B2 A1:B3 A2:B3
    0     0     2    -2    -1     1

> options(old.width)
```

### *13.5.2 Example—Dummy Variables for Crossed Factors Nested Within Another Factor—Turkey Data (Continued Again)*

A model formula specifies a set of dummy variables. Just as in one-way analysis of variance, we control the structure of the dummy variables with the contrast matrix assigned to each factor. Let us look at the dummy variables generated for us by the model formula

```
wt.gain ~ trt.vs.control / (additive*amount)
```

We do so in R by adding the argument `x=TRUE` to the `aov` statement in Table 13.16 and then displaying the `x` component of the resulting `aov` object in Table 13.25.

There are some complications in the display in Table 13.25. The generation of the $x$ matrix of dummy variables doesn't know about the actual degrees of freedom for each effect. It assumes the maximum possible if all implied cells were observed (in this example there are $2 \times 3 \times 3 = 18$ cells implied by the complete crossing of `trt.vs.control`, `additive`, and `amount`). Only five of those implied cells actually have observations. The `match` and the relabeling are used to find just the ones that matter in this example and to give them more reasonable names. See Exercise 13.10 for guidance on discovering how the `match` function is used.

The predicted value for an observation $i$ is calculated, as with any linear model, as the inner product of the regression coefficients with the dummy variables in row $i$. In this example, we predict the weight gain for observation 7 as

$$\hat{y}_7 = (\,1\ \ {-1}\ \ 1\ \ 1\ \ 1\,)\begin{pmatrix} 6.5300 \\ -0.6867 \\ -0.9750 \\ -0.9667 \\ 0.2250 \end{pmatrix} = 5.5$$

## 13.6 Sequential and Conditional Tests

When there are two or more predictors in a model, they are usually not orthogonal to each other. Therefore, the interpretation given to the relative importance of each predictor depends on the order in which they enter the model. One of the important goals of designed experiments is the choice of combinations of levels for factors that will make the dummy variables for each factor or interaction orthogonal to the others. Most of the examples in this book in Chapters 12, 13, and 14 have orthogonal effects.

When the data for an example have continuous predictors or covariates, or are classified by factors with unequal numbers of observations per cell, the effects are

**Table 13.25** Regression coefficients and dummy variables for turkey data. This table is a continuation of Tables 13.15 and 13.16.

```
> match(dimnames(coef(summary.lm(turkey3.aov)))[[1]],
+       dimnames(turkey3.aov$x)[[2]])
[1]  1  2  4  8 12

> turkey[c(1,7,13,19,25),]
      diet wt.gain trt.vs.control additive amount
1  control     4.1        control  control      0
7       A1     5.2      treatment        A      1
13      A2     6.3      treatment        A      2
19      B1     6.5      treatment        B      1
25      B2     9.5      treatment        B      2

> turkey3.coef <- summary.lm(turkey3.aov)$coef

> turkey3.x <- turkey3.aov$x

> term.names <-
+     c("(Intercept)","trt.vs.control","additive","amount",
+       "additive:amount")

> dimnames(turkey3.coef)[[1]] <- term.names

> dimnames(turkey3.x)[[2]][c(1,2,4,8,12)] <- term.names

> zapsmall(turkey3.coef)
                Estimate Std. Error   t value Pr(>|t|)
(Intercept)      6.53000    0.10253  63.68585   0.0000
trt.vs.control  -0.68667    0.05127 -13.39386   0.0000
additive        -0.97500    0.11464  -8.50510   0.0000
amount          -0.96667    0.11464  -8.43241   0.0000
additive:amount  0.22500    0.11464   1.96272   0.0609

> turkey3.x[c(1,7,13,19,25), c(1,2,4,8,12)]
   (Intercept) trt.vs.control additive amount additive:amount
1            1              4        0      0               0
7            1             -1        1      1               1
13           1             -1        1     -1              -1
19           1             -1       -1      1              -1
25           1             -1       -1     -1               1
```

usually not orthogonal. Most of the examples in Chapters 9, 10, and 11 have continuous predictor variables and therefore do not have orthogonal effects.

When effects are not orthogonal, the sequence in which they are entered into the model affects the interpretation of the effects. See Sections 9.6 (Partial $F$-Tests) and 9.13 (Residual Plots) for techniques used to investigate the relative importance of the predictors.

The sequential ANOVA table depends on the order in which the effects are entered into the model. Each row of the table is calculated under the assumption that all effects in higher rows have already been included and that all effects in lower rows have not. R normally prints the sequential ANOVA table. SAS calls the sequential ANOVA table the table of Type I sums of squares.

There are several types of conditional ANOVA tables. One of the most frequently used is Yates' weighted squares of mean, what SAS calls Type III sums of squares, in which each row of the table is calculated under the assumption that all other rows—both higher and lower in their placement in the ANOVA table—have already been included. We, and many others, have difficulty with Type III sums of squares when used with designed experiments because they violate the principle of marginality that says it is usually not meaningful to test, estimate, or interpret main effects of explanatory variables when the variables interact. The principle was stated by Nelder (1977) and strongly supported by Venables (1998).

Another method, Yates' Method of Fitting Constants, what SAS calls Type II sums of squares, makes different assumptions for each class of effect. ANOVA table rows for main effects assume all other main effects are already included in the model. ANOVA table rows for two-way interactions assume all main effects and other two-way interactions are in the model. Higher-order interactions assume all lower-order effects and interactions are already in the model, hence are consistent with the marginality principle.

## *13.6.1 SAS Terminology for Conditional Sums of Squares*

The terminologies Type I, Type II, and Type III sum of squares originated by SAS have become so widely known that they are used nowadays even outside the context of interpreting SAS listing files. In order that readers be able to request and interpret SAS analysis of variance presentations, we provide here more details on these types in two contexts, the context of designed experiments having two factors and the context of regression analysis with continuous predictor variables.

Suppose the response is `Y`, the two factors are `A` and `B`, and the SAS model statement reads

```
Model Y = A B A*B;
```

Since no particular types of sums of squares were requested, SAS provides by default Types I and III. If the user wishes to override the default, particular types can be requested as illustrated here:

```
Model Y = A B A*B /ss1 ss2;
```

The sequential (Type I) sum of squares for each effect is the portion of model sum of squares attributable to that effect above and beyond what is attributable to all effects listed prior to it in the expanded model statement. It is conditional on all the previous terms already being in the model. Thus in the illustration, the Type I sum of squares for `B` is the marginal contribution of factor `B` conditional on factor `A` already being in the model. Use of this sum of squares is appropriate if a model containing factor `A` without factor `B` makes sense, but a model containing factor `B` makes no sense unless the model already includes factor `A`.

For each main effect in the model statement, the Type II sum of squares is the marginal contribution of that effect beyond the sum of squares attributable to all other main effects in the model statement. The Type II sum of squares for `A*B` is the portion of model sum of squares attributable to this interaction after the main effects `A` and `B` are already in the model. Yates (1934) gave the name *method of fitting constants* to what is now called Type II sums of squares. In the absence of interaction, this method produces the maximum power tests for the main effects.

Note that while the Type I sums of squares for `A`, `B` and `A*B` add up to the model sum of squares, the Type II sums of squares for these three effects are not in general an orthogonal partitioning of the model sum of squares and hence do not in general sum to the model sum of squares. An exception occurs when the data are balanced (for example, each of the *ab* cells contain the same number of observations); in this case the Type I and Type II sums of squares coincide.

Type III sums of squares can be used with the above model statement provided that each of the *ab* cells contains at least one observation. The Type III sum of squares for any effect, including A*B, is adjusted for *all* other effects in the model. If the sampling is balanced, the Type III sums of squares for main effects coincide with the sums of squares for Type I and Type II. The Type III partitioning provides what is known as the Yates' weighted squares of means analysis of unbalanced data; see Searle (1971).

Type IV sums of squares coincide with Type III sums of squares when all cells contain observations. This partitioning is used when some cells are empty, a situation we do not pursue in this text. The Type IV sum of squares partitioning is not unique, a feature that makes many analysts uncomfortable with their use.

The nomenclature Type I and Type III (as well as Type II and Type IV) was originated by SAS in Goodnight (1978) and summarized in SAS Institute, Inc. (1999).

### 13.6.2 *Example—Application to Clover Data*

In Table 12.12 we showed the two-way ANOVA table for the clover data. The data in `data(rhiz.clover)` is balanced. As a consequence all three forms of conditional sums of squares give exactly the same results (as long as proper contrasts (with columns orthogonal to the constant column) are used for all factors. The treatment contrasts (the default) are not proper contrasts and may show strange results).

In Table 13.26 we illustrate the three conditional sets of sums of squares using an unbalanced dataset, constructed by deleting two observations from the balanced clover data. The data are displayed in Figure 13.14.

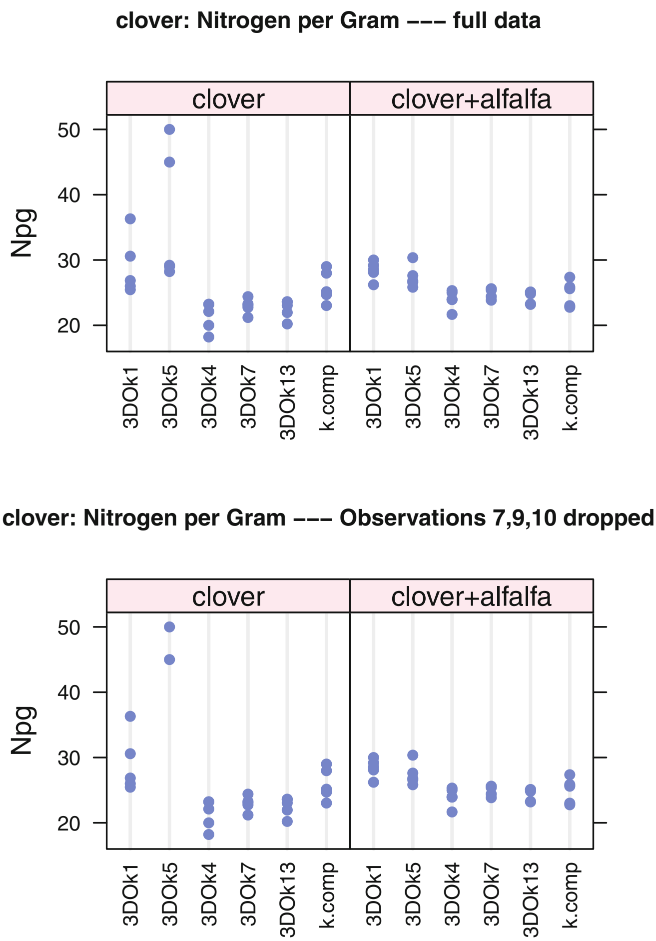



**Fig. 13.14** Full clover dataset on top. Two observations removed on the bottom.

**Table 13.26** Display of conditional sums of squares (Sequential/Type I, Method of Fitting Constants/Type II, and Weighted Squares of Means/Type III) for an unbalanced dataset. The `car::Anova` function requires contrasts orthogonal to the constant column. Compare the `Sum Sq` values for `strain` and `comb`. The left column shows the model formula `Npg ~ strain * comb` and the right column shows `Npg ~ comb * strain`. The main effects in the sequential tables (using `anova`) depend on the order. The main effects in the method of fitting constants tables `car::Anova(type=2)` do not depend on order and are the same as the second position of the sequential tables. The main effects in the weighted squares of means tables `car::Anova(type=3)` do not depend on order; they violate the principal of marginality because they are conditional on taking out the interaction effect first. The `Residuals` and interaction `comb:strain` sums of squares are identical in all six tables.

```
> data(rhiz.clover)

> ## drop two observation to illustrate Type II and III sums of squares
> ## I am dropping the non-outlier observations in 3DOk5
> cloverD <- rhiz.clover[-c(7,9,10),]

> old.opt <- options(show.signif.stars=FALSE, digits=3)
```

```
> cloverDsc.aov <-
+   aov(Npg ~ strain * comb,
+       data=cloverD,
+       contrasts=
+         list(strain=contr.sum,
+              comb=contr.sum))

> anova(cloverDsc.aov)[,c(2,1,4,5)]
            Sum Sq Df F value  Pr(>F)
strain         657  5   29.28 4.2e-13
comb            20  1    4.46    0.04
strain:comb    594  5   26.47 2.3e-12
Residuals      202 45

> Anova(cloverDsc.aov, type=2)
Anova Table (Type II tests)

Response: Npg
            Sum Sq Df F value  Pr(>F)
strain         675  5   30.07 2.7e-13
comb            20  1    4.46    0.04
strain:comb    594  5   26.47 2.3e-12
Residuals      202 45

> Anova(cloverDsc.aov, type=3)
Anova Table (Type III tests)

Response: Npg
            Sum Sq Df F value  Pr(>F)
(Intercept)  38711  1  8621.1 < 2e-16
strain        1049  5    46.7 < 2e-16
comb            86  1    19.3 6.8e-05
strain:comb    594  5    26.5 2.3e-12
Residuals      202 45
```

```
> cloverDcs.aov <-
+   aov(Npg ~ comb * strain,
+       data=cloverD,
+       contrasts=
+         list(strain=contr.sum,
+              comb=contr.sum))

> anova(cloverDcs.aov)[,c(2,1,4,5)]
            Sum Sq Df F value  Pr(>F)
comb             2  1    0.52    0.48
strain         675  5   30.07 2.7e-13
comb:strain    594  5   26.47 2.3e-12
Residuals      202 45

> Anova(cloverDcs.aov, type=2)
Anova Table (Type II tests)

Response: Npg
            Sum Sq Df F value  Pr(>F)
comb            20  1    4.46    0.04
strain         675  5   30.07 2.7e-13
comb:strain    594  5   26.47 2.3e-12
Residuals      202 45

> Anova(cloverDcs.aov, type=3)
Anova Table (Type III tests)

Response: Npg
            Sum Sq Df F value  Pr(>F)
(Intercept)  38711  1  8621.1 < 2e-16
comb            86  1    19.3 6.8e-05
strain        1049  5    46.7 < 2e-16
comb:strain    594  5    26.5 2.3e-12
Residuals      202 45

> options(old.opt)
```

### 13.6.3 Example—Application to Body Fat Data

We revisit in Table 13.27 the analysis begun in Section 9.2 of a portion of the body fat data `data(fat)` using the two predictors `abdomin` and `biceps` of the response `bodyfat`.

The $F$-value 54.92 applies to the composite hypothesis $H_0$:$\beta_1 = \beta_2 = 0$ against the alternative that at least one $\beta_i$ is nonzero, where the $\beta_i$ are the coefficients of the two predictors. The corresponding small $p$-value indicates that either `abdomin` or `biceps` or both are linearly related to `bodyfat`. The $R^2 = 0.714$ tells us that 71.4% of the variability in these subjects' `bodyfat` is accounted for by their `abdomin` and `biceps` measurements. The remaining 28.6% of `bodyfat` variability is explained by other measurable variables not presently in the model as well as the random error component of the model in Equation (9.1).

The Type I sums of squares are sequential in that the sum of squares 2440.5 for the first listed predictor, `abdomin`, is calculated assuming that this is the only predictor in the model, while the sum of squares 209.32 for the second listed predictor, `biceps`, is calculated assuming that the first listed predictor is already in the model. In general, the top to bottom ordering of sources of variation in the Type I sum of squares table is the same as the ordering of these sources in the `Model` statement.

Each predictor's Type III sum of squares is calculated assuming that all other predictors are already in the model. Thus the Type III sum of squares for `abdomin`, 1823.02, is *conditional* in the sense that it is calculated under the assumption that the model already contains the other predictor `biceps`. In general, any entry in a Type III sum of squares table is conditioned on the existence in the model of all sources above it in this table.

The parameter estimates in both Table 9.1 and Table 13.27 are based on this same "last-in" rule, corresponding to the Type III sums of squares. The Type III $F$-value for `abdomin`, 75.57, is the square of the $t$-value for `abdomin`, 8.69, in the `Parameter` section of Table 13.27 and in Table 9.1. This corresponds to the interpretation of the $t$-tests for the regression coefficients, each of which measures the marginal contribution of its predictor variable conditional on all the other predictor variables already being in the model.

In this context, it is preferable to work with the Type I analysis if the investigator believes that a model containing `biceps` makes no sense unless `abdomin` is already in the model. Otherwise, with continuous predictor variables, the Type III approach is preferred. In this example, each predictor has a statistically significant impact on `bodyfat` after the other predictor has already been included in the model. In general, it is possible for one predictor to have an insignificant additional impact on the response when other more prominent predictors are already in the model. See Section 9.11 on collinearity for a discussion of this issue.

**Table 13.27** `car::Anova` from the **car** package display for two-$X$ regression of `bodyfat`. The sequential sums of squares (Type I sums of squares) correspond to the display in Table 9.1. The sequential sums of squares are an orthogonal partitioning of the model sum of squares. The weighted squares of means (Type III sums of squares) is not an orthogonal partitioning. The sum of the two values for `abdomin` and `biceps` is not equal to the model sum of squares.

```
> library(car)

> data(fat)

> fat.lm <- lm(bodyfat ~ abdomin + biceps, data=fat)

> ## regression coefficients
> coef(summary(fat.lm))
               Estimate Std. Error   t value     Pr(>|t|)
(Intercept) -14.5937363 6.69222199 -2.180701 3.459548e-02
abdomin       0.6829379 0.07855885  8.693329 4.168613e-11
biceps       -0.9221544 0.31304822 -2.945726 5.133920e-03

> ## sequential sums of squares (Type I)
> anova(fat.lm)
Analysis of Variance Table

Response: bodyfat
          Df  Sum Sq Mean Sq  F value    Pr(>F)
abdomin    1 2440.50 2440.50 101.1718 5.581e-13 ***
biceps     1  209.32  209.32   8.6773  0.005134 **
Residuals 44 1061.38   24.12
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

> ## weighted squares of means (Type III)
> Anova(fat.lm, type="III")
Anova Table (Type III tests)

Response: bodyfat
             Sum Sq Df F value    Pr(>F)
(Intercept)  114.71  1  4.7555  0.034595 *
abdomin     1823.02  1 75.5740 4.169e-11 ***
biceps       209.32  1  8.6773  0.005134 **
Residuals   1061.38 44
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

> ## model sum of squares
> var(fat$bodyfat) * (nrow(fat)-1) - sum(fat.lm$residuals^2)
[1] 2649.817
```

## 13.7 Exercises

**13.1.** Consider an experiment to determine which of the four types of `valve` used in an artificial heart maximizes blood pressure control as measured by maximum `flow` gradient (mm Hg). `Flow` was maintained at each of the same six `pulse` rates for each valve type. Two `runs` were made for each valve type. The order of the eight runs at the four valve types was randomized. Note that `run` is a random factor, nested within `valve`. The dataset `data(heartvalve)` comes from Anderson and McLean (1974). Perform a thorough analysis including plots of the data.

**13.2.** An experiment reported in Lewin and Shakun (1976) investigated whether an Octel filter (`type=2`) or a standard filter (`type=1`) provided superior suppression of `noise` produced by automobile exhaust systems. The experiment considered three vehicle `sizes` coded 1 small, 2 medium, 3 large; and both the right 1 and left 2 `side` of cars. The dataset is available as `data(filter)`. Perform a thorough analysis leading to a recommendation of which filter to use under the various experimental conditions.

**13.3.** An experiment explored the abilities of six commercial laboratories to accurately measure the percentage `fat` content in samples of powdered eggs. A pair of samples from a single can was sent to each lab. The labs were told that the samples were of two `types`, but in fact they were from the same can. Each lab assigned two `technicians` to analyze each type. The dataset from Bliss (1967) is `data(eggs)`. Analyze the data in order to recommend which lab(s) have superior or inferior abilities to ascertain the fat content of powdered eggs.

**13.4.** Box and Cox (1964), reprinted in Hand et al. (1994), reported the results of a $3^3$ factorial experiment. The dataset is `data(wool)`. The response is the `cycles` under tension to failure of worsted yarn. The three factors are `length` of test specimen (250, 300, 350 mm), `amplitude` of loading cycle (8, 9, 10 mm), and `load` (40, 45, 50 g). The levels of all three factors are coded as $-1, 0, 1$ in the data file. The authors recommend a preliminary log transformation of the response. Perform an analysis to determine the influences of the factors on the response. This dataset is from the paper defining the Box–Cox transformation.

**13.5.** A 5×3×4 factorial experiment is designed to compare the wear resistance of vulcanized rubber Davies (1954, p. 192). The three factors are

filler: 5 qualities

pretreatment: 3 methods

raw: 4 qualities

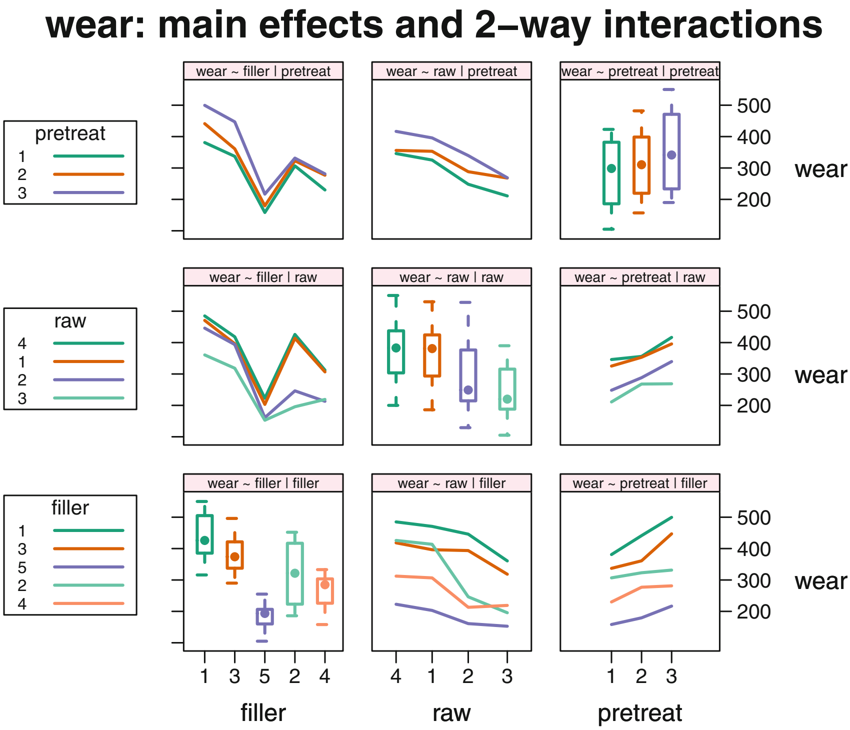



**Fig. 13.15** Main effects and two-way interactions for wear resistance of vulcanized rubber.

There is only one replicate; thus the assumption must be made that the three-factor interaction is negligible and the three-factor sum of squares can be used for the error term.

The data are available as `data(vulcan)`. The graph of the main effects and two-way interactions is in Figure 13.15. The simple effects of `filler` and `raw` are in Figure 13.16.

a. Determine from the ANOVA table whether any of the main effects or two-way interactions are significant.
b. Why can't we test the three-way interaction?
c. From the figures and tables of means, determine if any levels of any factors can be eliminated from further consideration. Assume that we are looking for big numbers for the best wear resistance.
d. Of the treatment combinations that are left for consideration, are any clearly dominant? Would we need to make a conditional recommendation to the client?

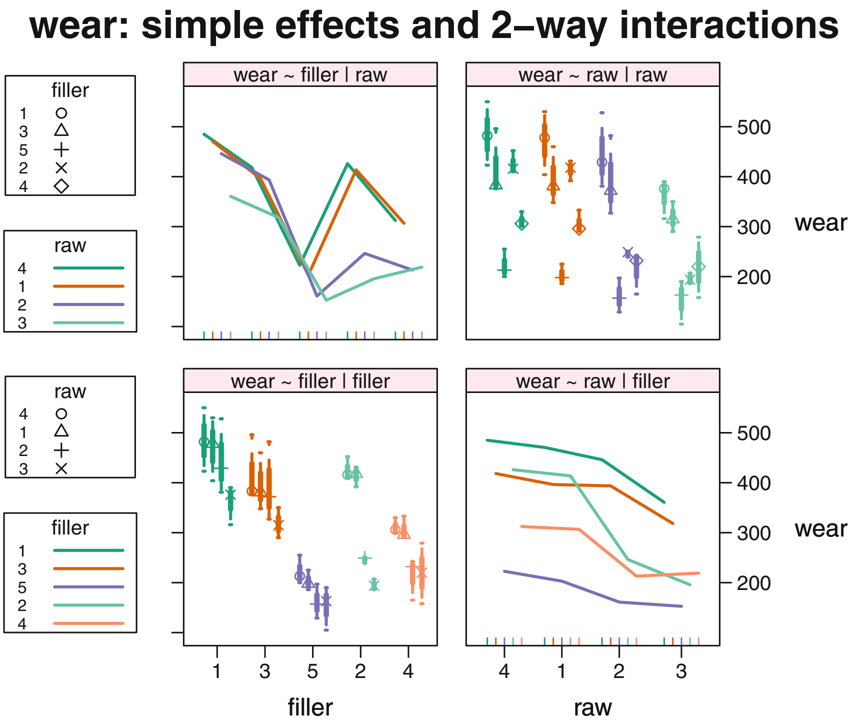



**Fig. 13.16** Main effects and simple effects of `filler` and `raw` for wear resistance of vulcanized rubber.

**13.6.** Continue the Latin square example using the dataset `data(tires)` in Section 13.2. The treatment sum of squares with 3 degrees of freedom is linearly dependent on the Row×Column interaction with $(r-1) \times (c-1)$ degrees of freedom. Demonstrate the dependency by showing that each of the dummy variables for `brand` has a zero residual when regressed on the dummy variables for `car*position`.

You may use chunk 39 of file `HHscriptnames(13)`, reproduced in Table 13.28, as a starting point. Explain why the residual sum of squares in `tr1.lm` (and the analogous `tr2.lm` and `tr3.lm`) is 0.

**13.7.** Peterson (1985) discusses an experiment to assess the effects on strengths of spot welds (psi $\times 10^{-3}$) created by robots on automobile assembly lines. On each of two assembly `lines` (blocks) there were three fixed treatment factors: `maker` at two levels, rod `diameters` at three levels (30 mm, 60 mm, 90 mm), and `chromium` content at three levels (1%, 1.5%, 2%). The 18 treatment combinations were randomly assigned to 18 robots on each assembly line. The dataset is `data(weld)`. Analyze the data, including

**Table 13.28** Chunk 39 of file HHscriptnames(13) to be used as starting point for Exercise 13.6.

```
## R defaults to treatment contrasts for factors.
## We need an orthogonal set of factors for this exercise.
##
data(tires)
contrasts(tires$car)      <- contr.helmert(4)
contrasts(tires$position) <- contr.helmert(4)
contrasts(tires$brand)    <- contr.helmert(4)

tires.aov <- aov(wear ~ car + position + brand, data=tires, x=TRUE)
anova(tires.aov)
tires.rc.aov <- aov(wear ~ car * position, data=tires, x=TRUE)
anova(tires.rc.aov)

t(tires.aov$x[,8:10])
t(tires.rc.aov$x[,8:16])
tr1.lm <- lm(tires.aov$x[,8] ~ tires.rc.aov$x[,8:16])
anova(tr1.lm)
```

a. a discussion of interaction among the treatment factors.

b. a recommendation of the combination of the treatment factors for maximizing `strength`. Explain how you know that your recommendation for `diameter` is distinctly better than the next-best choice of `diameter`.

**13.8.** Peterson (1985) describes an investigation to compare the abilities of seven washday `products` to remove dirt in cloth:

- liquid detergent
- granular detergent
- detergent flakes
- liquid detergent plus phosphate
- granular detergent plus phosphate
- detergent flakes with phosphate
- soap

Each of these seven `products` was assigned to three bedsheets soiled in a standard way, and the amount of `dirt` removed (mg) from each bedsheet was recorded. The data is in `data(washday)`. Analyze these data.

a. Perform a one-way analysis of variance to assess whether the mean amount of `dirt` removed is the same for all seven `products`.

b. Partition the 6 degrees of freedom sum of squares for `product` into 6 mutually orthogonal 1 degree-of-freedom sums of squares, each of which has an interpretation based on the similarities and differences among the `products`.

c. Estimate each of the six corresponding contrasts.

d. The six levels (hence five contrasts) within detergent can be specified as the crossing of three levels of form (liquid, granular, flakes) and two levels of ingredient (none, phosphate). Rewrite the model as a crossing of form and phosphate nested within soap.vs.detergent.

e. Assuming that the costs per wash are roughly the same for all seven `products`, provide recommendations for consumers.

**13.9.** Neter et al. (1996) describe an experiment to compare the work of market research firms. The dataset is `data(market)`. It was desired to evaluate the effects on `quality` of work performed by 48 firms of the factors of the three crossed factors fee level (`feelevel`), `scope`, and `supervision`. Fee level has three levels (1 = high, 2 = average, 3 = low), scope has two levels (1 = all performed in-house; 2 = some contracted out), and supervision has two levels (1 = local supervisors, 2 = traveling supervisors). Construct an analysis of variance table. Produce and interpret interaction plots for any interaction found significant in the table. Compare the means of the levels of any factors not involved in a significant interaction.

**13.10.** In Table 13.25 we use the R `match` function to identify which of the implied dummy variables in a nested design are actually used. The complete command using the `match` function is

```
match(dimnames(coef(summary.lm(turkey3.aov)))[[1]],
      dimnames(turkey3.aov$x)[[2]])
```

Study the command by picking up pieces of it and dropping them into the Console window. For example, assuming you have already defined all the variables by running the R statements leading up to Table 13.25, open file `HHscriptnames(13)` in your editor and highlight and run the pieces of code corresponding to the lines in Table 13.29.

**13.11.** It is desired to compare a response variable `dimvar`, dimensional variability, of a component produced by each of three `machines`. Each machine is comprised of two `spindles`, and four components are selected at random from each spindle. This example is attributable to Montgomery (2001), and the dataset is `data(spindle)`. Perform an analysis to determine the effects of `spindle` and `machine` on `dimvar`, assuming that both factors are fixed.

**Table 13.29** Isolated code fragments to be run one line at a time to help learn what the complete statement is doing. See Exercise 13.10 for more detail.

```
                    summary.lm(turkey3.aov)
               coef(summary.lm(turkey3.aov))
      dimnames(coef(summary.lm(turkey3.aov)))
      dimnames(coef(summary.lm(turkey3.aov)))[[1]]

               turkey3.aov$x
      dimnames(turkey3.aov$x)
      dimnames(turkey3.aov$x)[[2]]

match(dimnames(coef(summary.lm(turkey3.aov)))[[1]],
      dimnames(turkey3.aov$x)[[2]])
```

**13.12.** In an experiment reported by Montgomery (2001), the response variable is a measure of `surface` finish of a metal part. Each part is produced by one of four `machines`, a fixed factor. Three `operators` are assigned to produce parts on each `machine`. The operators are selected at random and a total of 12 different operators are chosen for the 4 machines. Analyze the data in `data(surface)` to determine the effects on `surface` of `machine` and `operator`.

## 13.A Appendix: Orientation for Boxplots

We display the boxplots for the turkey data in two orientations in Figures 13.13 and 13.17. We prefer the vertical orientation for the values of the response variable because it accords with how we have been trained to think of functions—levels of the independent variable along the abscissa and the response variable along the ordinate. Most of the graphs in this book are oriented with the response variable in the vertical direction.

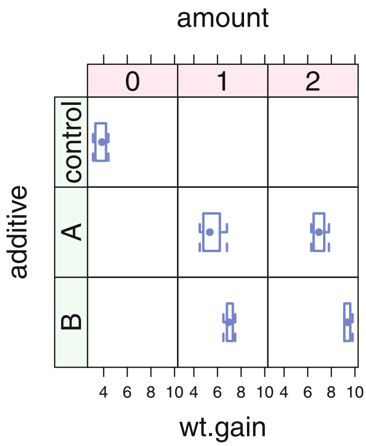



**Fig. 13.17** Turkey data with factor structure. This figure is essentially the same as Figure 13.13, but with horizontal boxplots instead of vertical boxplots.

We chose to display Figures 12.7, and 12.8 in vertical orientation. In Section 12.14.5 we discuss three other options for varying the horizontal and vertical orientation and varying the conditioning variable. One of our concerns was legibility of the labels when there are too many long labels on the abscissa. Code chunks for viewing the options are included in file `HHscriptnames(12)`.