

# Tutorial 7 Solutions

## Random Effects One-Way ANOVA and Two-Way ANOVA

STAT 292: Applied Statistics 2A

### R Functions Used in This Tutorial

In this tutorial we use output from the `mean_sq` column of the ANOVA table produced by the `aov` function to estimate the two **components of variation** in a one-way **random effects** model. We also use the `aov` function to fit a **two-way ANOVA** and we visualise the effects of the two factors in an **interaction graph**, produced using the `interaction.plot` function.

R functions that will be emphasised during this tutorial are:

Function	Description	Package
<code>aov</code>	Fit an analysis of variance model	<code>stats</code>
<code>interaction.plot</code>	Plot the mean (or other summary) of the response for two-way combinations of factors	<code>stats</code>

The actions of the functions in this tutorial are best illustrated by looking at the resulting output given below (and in the solutions), along with other examples given in Chapters 2 and 3 of the Part 2 Lecture Notes. Also refer back to the earlier tutorials.

Recall that the help file for any function can be produced by typing `?<FUNCTION_NAME>` or `help(FUNCTION_NAME)` (where `FUNCTION_NAME` is the name of the function) at the command line in the R console (e.g., `?aov`, `?interaction.plot`).

### Questions and Solutions

1. An education researcher wishes to know how much influence individual teachers have on students' understanding of the subject being taught. Five teachers are chosen randomly from a large pool of teachers. Fifty students are randomly chosen from a large group, and they are assigned randomly to the teachers, ten to each teacher.

The teachers teach new material to their group for four hours per week over three weeks, after which all 50 students are given a standard test for their understanding of the topic. The table below gives the students' scores on the test (as a percentage).

Teacher	Student test score (%)
1	69, 77, 59, 71, 63, 69, 62, 52, 71, 79
2	52, 58, 47, 38, 63, 58, 47, 49, 58, 49
3	72, 46, 56, 73, 53, 53, 60, 54, 71, 52
4	53, 84, 76, 81, 61, 74, 59, 53, 74, 70
5	83, 76, 81, 59, 80, 72, 69, 67, 78, 78

- a. Write the model for a one-way ANOVA with a random effect. What is the meaning of the different terms in the model?

The model equation is

$$Y_{ij} = \mu + A_i + E_{ij}$$

where:

$Y_{ij}$  is the random variable for the  $j$ -th score with Teacher  $i$ ,

$\mu$  is a parameter for an overall mean,

$A_i$  is a random variable for the random effect of Teacher  $i$ , with  $N(0, \sigma_A^2)$  distribution,

$E_{ij}$  is the error term random variable for score  $j$  with Teacher  $i$ , with  $N(0, \sigma^2)$  distribution.

- b. Do the ANOVA assumptions seem valid? To answer this question use the R commands and output given below and your knowledge of how the experiment was designed and run.

The output shows that it is reasonable to assume constant variance: there is a roughly constant overall spread in the boxplots, Levene's test has  $p = 0.3016$  (greater than 0.05, so do not reject the null hypothesis of equal variances), and the plot of Residuals versus Predicted Values, shows a level band across the page. The assumption of normality also seems valid: there is an approximately straight line in the Q-Q plot of Sample Quantiles versus Theoretical Quantiles, confirming normality.

We cannot check the independence of the errors from inspecting the data. However, the random selection of students supports an assumption of independence of the error terms, the random selection of five teachers supports the assumption of independence of the random variable  $A_i$ , and the independence of the  $A_i$  from  $E_{ij}$  is supported since the students were randomly assigned to the

teachers. We must trust that the experiment was run in such a way as to preserve that independence, e.g. students should not be allowed to communicate while taking the test. Students have very likely communicated with each other during the teaching, however, so some independence may have been lost.

Note, with only 10 observations per teacher, the appearance of the boxplots is very much dependent on how close together the individual scores are. There are only five observations either side of the median in each boxplot (the median is half way between the fifth and sixth ranked values for each teacher). Then there are only two observations between each of: the minimum and the LQ (which is the third ranked value for each teacher), the LQ and the median, the median and the UQ (which is the eighth ranked value for each teacher), and the UQ and the maximum. So, for example, if any scores are repeated or only differ by one or two marks, the 'hinges' of the boxes (and/or the 'whiskers') can appear very short. This is because of the small number of observations and the recording of the scores as whole numbers—it does not mean, necessarily, that the underlying data are not normally distributed, at least approximately.

```
## Teacher effects

score <- c(69, 77, 59, 71, 63, 69, 62, 52, 71, 79,
          52, 58, 47, 38, 63, 58, 47, 49, 58, 49,
          72, 46, 56, 73, 53, 53, 60, 54, 71, 52,
          53, 84, 76, 81, 61, 74, 59, 53, 74, 70,
          83, 76, 81, 59, 80, 72, 69, 67, 78, 78)

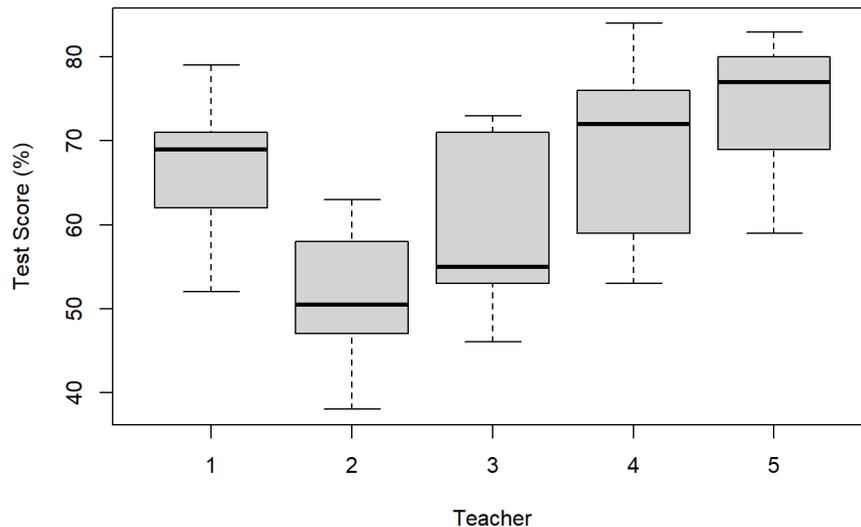
teacher <- rep(c(1,2,3,4,5), each = 10)

# Display mean score for the different teachers using tapply.
tapply(score, teacher, mean)
```

```
1 2 3 4 5
67.2 51.9 59.0 68.5 74.3
```

```
# Produce a boxplot of score by teacher.
# Note the effect that similar scores have in the plots,
# on the appearance of the boxes and the whiskers.
boxplot(score ~ factor(teacher),
        main = "Distribution of Student Score by Teacher",
        xlab = "Teacher", ylab = "Test Score (%)")
```

**Distribution of Student Score by Teacher**



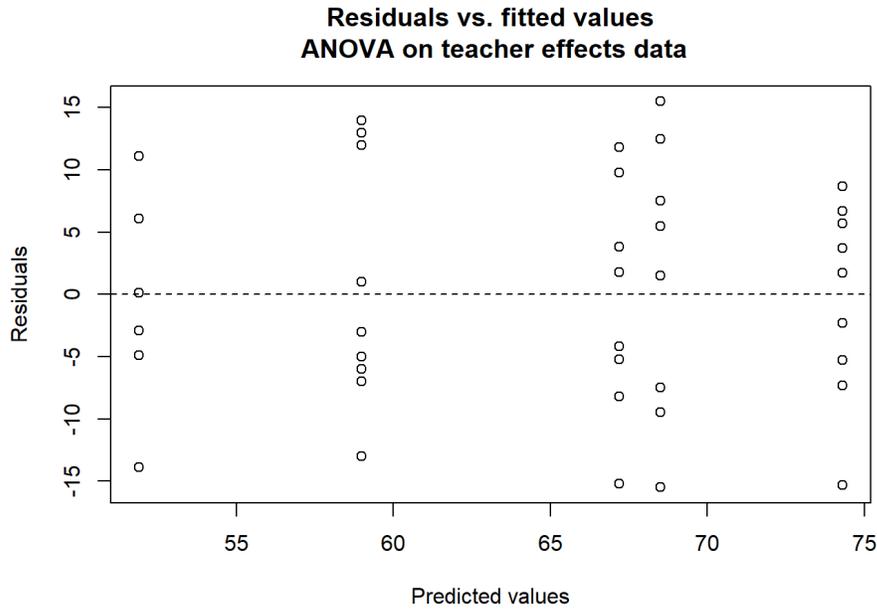
```
# Carry out a one-way ANOVA of score on teacher.
teacher.ANOVA <- aov(score ~ factor(teacher))
summary(teacher.ANOVA)
```

```
          Df Sum Sq Mean Sq F value    Pr(>F)
factor(teacher) 4   3078   769.6   9.654 1.01e-05 ***
Residuals      45   3587    79.7
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

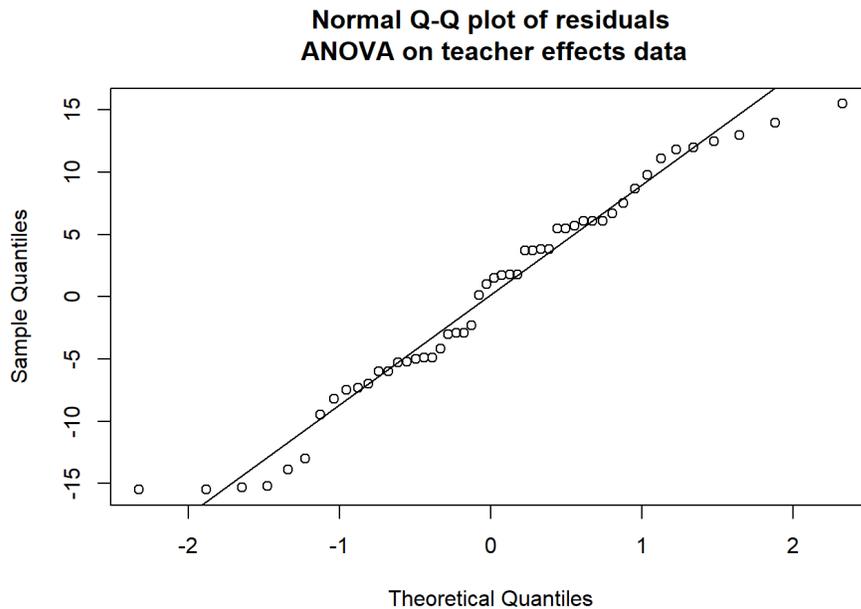
```
# Carry out Levene's test for score on teacher.
# Needs the "car" package to have been loaded.
leveneTest(score ~ factor(teacher), center = "mean")
```

```
Levene's Test for Homogeneity of Variance (center = "mean")
  Df F value Pr(>F)
group 4  1.2552 0.3016
  45
```

```
# Scatterplot of residuals vs. fitted values for score on teacher.
plot(x = teacher.ANOVA$fitted.values, y = teacher.ANOVA$residuals,
     main = "Residuals vs. fitted values\n ANOVA on teacher effects data",
     xlab = "Predicted values", ylab = "Residuals")
abline(h = 0, lty = 2)
```



```
# Normal Q-Q plot of residuals for teacher effects data.
qqnorm(teacher.ANOVA$residuals,
       main = "Normal Q-Q plot of residuals\n ANOVA on teacher effects data")
qqline(teacher.ANOVA$residuals)
```



c. Do the random effects analysis, *i.e.*, estimate:

- the two components of variance,
- the total variance,
- the percentage of the variability in the data due to individual teacher differences and the percentage due to unexplained variance.

Using the ANOVA output,

- $\hat{\sigma}^2 = 79.7,$

$$\hat{\sigma}^2 + r\hat{\sigma}_A^2 = 769.6.$$

Hence, using  $r = 10$ ,  $\hat{\sigma}_A^2 = (769.6 - 79.7)/10 = 68.99$

- Total variance =  $\hat{\sigma}_A^2 + \hat{\sigma}^2 = 68.99 + 79.7 = 148.69$
- Percentage due to teacher differences =  $100 \times 68.99/148.69 = 46.4\%$ ,  
and percentage unexplained =  $53.6\%$ .

**d. Comment on your results—are teacher differences important here?**

The individual teacher is important here, as teacher differences account for 46.4% of the variability in the data.

*Note: How much is classed as important? There is no objective cut-off. It depends on the context. In some situations, such as explaining weather patterns, researchers are pleased to have a predictor which accounts for even 15–20% of the variation. Animal and plant breeders are also happy to use sources of variation which may only account for 10%; it gives them a way of improving the herd of cows, flock of sheep, stable of racehorses, or variety of vegetable.*

**2. Two methods of delivering a drug into a patient's body are tested: capsules (taken internally) or a skin patch taped to the wrist. For each method, a low, medium and high dosage is tried.**

**Twelve patients were available, two were randomly allocated to each treatment. The response variable is the concentration of the drug in the blood after two hours.**

	Low dose	Medium dose	High dose
Patch delivery	12, 14	19, 17	21, 24
Capsule delivery	11, 14	25, 23	32, 37

**a. How would you enter the data in R for a two-factor ANOVA?**

Some R commands to read in the data are given below.

```
# Store the drug concentration data in separate variables for
# drug concentration, drug dose, and drug delivery method.

conc <- c(12, 14, 19, 17, 21, 24, 11, 14, 25, 23, 32, 37)
dose <- rep(rep(c("Low", "Medium", "High"), each = 2), times = 2)
delivery <- rep(c("Patch", "Capsule"), each = 6)
```

**b. Some R commands and output are given below. Explain why the degrees of freedom (df) take the values that you can see in the R output.**

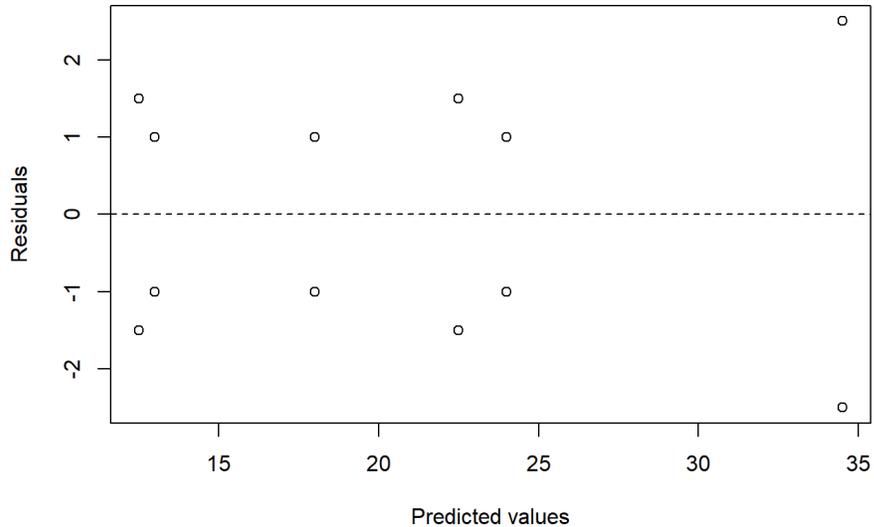
The two factors have three levels (for `dose`) and two levels (for `delivery`) respectively, so there are  $3 \times 2 = 6$  treatments and hence 6 parameters that need to be estimated overall—i.e., the six mean values for the six treatments. Therefore the residuals from the complete model will have  $n - 6 = 12 - 6 = 6$  degrees of freedom. In the complete model, an overall mean needs to be estimated, using up one parameter. Then each factor has one less independent parameters than its number of levels: so  $3 - 1 = 2$  parameters (= df used) for `dose` and  $2 - 1 = 1$  parameter (= df used) for `delivery`. The interaction between the factors uses  $(3 - 1) \times (2 - 1)$  or  $2 \times 1 = 2$  independent parameters (= df used). So, as expected, there are  $1 + 2 + 1 + 2 = 6$  parameters estimated in total (= df used).

```
# Fit a two-way ANOVA to the drug concentration data.
drug.ANOVA <- aov(conc ~ factor(dose) * factor(delivery))
summary(drug.ANOVA)
```

```
          Df Sum Sq Mean Sq F value    Pr(>F)
factor(dose)      2  496.5   248.25  54.164 0.000145 ***
factor(delivery)  1  102.1   102.08  22.273 0.003260 **
factor(dose):factor(delivery)  2    78.2    39.08   8.527 0.017627 *
Residuals        6    27.5     4.58
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

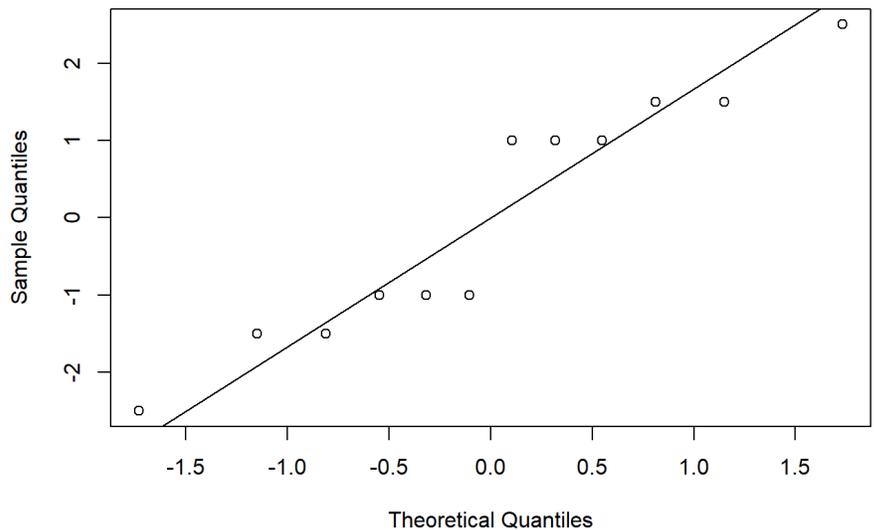
```
# Scatterplot of residuals vs. fitted values for drug concentration data.
plot(x = drug.ANOVA$fitted.values, y = drug.ANOVA$residuals,
     main = "Residuals vs. fitted values\n ANOVA for drug concentration data",
     xlab = "Predicted values", ylab = "Residuals")
abline(h = 0, lty = 2)
```

**Residuals vs. fitted values  
ANOVA for drug concentration data**



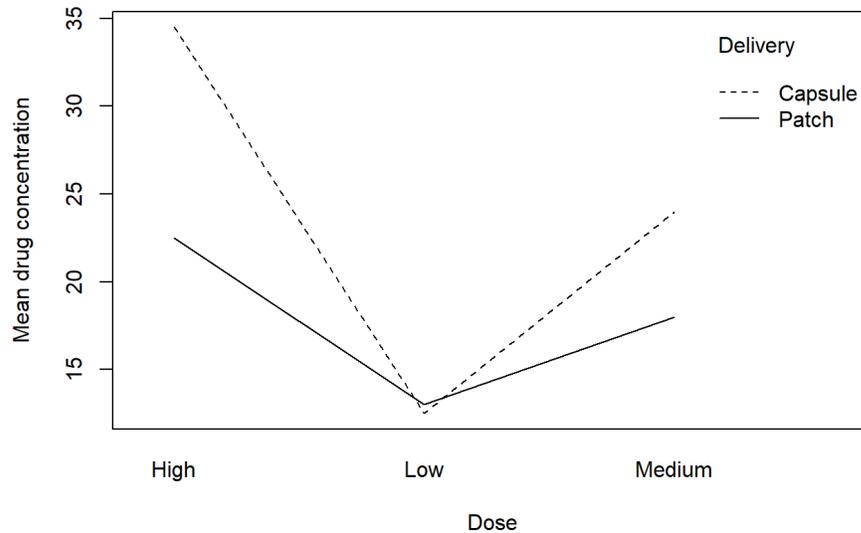
```
# Normal Q-Q plot of residuals for drug concentration data.
qqnorm(drug.ANOVA$residuals,
        main = "Normal Q-Q plot of residuals\n ANOVA for drug concentration data")
qqline(drug.ANOVA$residuals)
```

**Normal Q-Q plot of residuals  
ANOVA for drug concentration data**



```
# Interaction graph for drug concentration data.
# Terms used in the function are generally self-explanatory.
interaction.plot(x.factor = dose,
                 trace.factor = delivery,
                 response = conc,
                 fun = mean,
                 xlab = "Dose",
                 ylab = "Mean drug concentration",
                 main = "Interaction graph for drug delivery",
                 legend = TRUE, xpd=TRUE,
                 trace.label="Delivery")
```

### Interaction graph for drug delivery



**c. From the diagnostic graphs, do you think the assumptions are satisfied? Discuss why the Q-Q plot has 'steps' in it.**

The plot of Residuals versus Predicted Values shows a fairly level band; it's not perfect, as the two points at the right make it look a bit like funnelling. Possibly a log transformation could be tried, but we will not do it here. In this case Levene's test can't be applied, since the size of all (two) residuals within each treatment are identical (since the mean is half way between the two observations), so there is no variation on which to base the test. Boxplots for each cell are useless with only two observations per cell, so we have no other way of checking for constant variance.

Checking normality, the Q-Q plot of sample quantiles versus theoretical quantiles shows little steps and stairs, which is typical of data that have been rounded a lot and hence include repeated values. If the data had been given to one decimal place, this plot would have looked smoother. If we ignore the steps, it looks like a fairly straight line, confirming approximate normality. Again, a boxplot (a 5-number summary) for each cell would be useless, with only two observations per cell.

Without any special knowledge of how the experiment was run, other than the random allocation of patients to treatments (which is good!), we cannot check for independence.

**d. State any relevant hypotheses, and give your statistical conclusions (using a 5% level of significance).**

The effects model for two-way ANOVA has the model equation

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

where  $\mu$ ,  $\alpha_i$ ,  $\beta_j$  and  $(\alpha\beta)_{ij}$  are parameters with the following constraints:

- $\sum_{i=1}^a \alpha_i = 0$ ,
- $\sum_{j=1}^b \beta_j = 0$ ,
- for each  $j = 1, 2, \dots, b$ ,  $\sum_{i=1}^a (\alpha\beta)_{ij} = 0$ ,
- for each  $i = 1, 2, \dots, a$ ,  $\sum_{j=1}^b (\alpha\beta)_{ij} = 0$ .

The model assumptions are the constraints above, together with assuming that the error term,  $E_{ijk}$ , has a  $N(0, \sigma^2)$  distribution and the errors are independent. See the Part 2 Lecture Notes (e.g. Section 3.4) for further details.

We start with the interaction test:

$\mathcal{H}_0$  : There is no interaction—all  $(\alpha\beta)_{ij} = 0$ , versus

$\mathcal{H}_1$  : There is interaction—at least one  $(\alpha\beta)_{ij} \neq 0$ .

As seen in the ANOVA table from the given R output, the interaction is significant at the 5% level ( $p = 0.0176 < 0.05$ ), so we reject  $\mathcal{H}_0$  at the 5% significance level.

The interpretation is that the factors of dose and method of delivery interact in their effect on the concentration of the drug in the blood after two hours. We cannot discuss the effect of dose on concentration without specifying the method of delivery, and we cannot discuss the effect of the delivery method on concentration without specifying the dose.

Since there is interaction, we do not try to interpret the main effects terms in the ANOVA table. Instead we focus on more detailed interpretation of the interaction, using an interaction graph.

- e. The interaction graph produced by the `interaction.plot` command in R could be better. Explain why. What happened to cause this poor graph? How could it be fixed?

It is good to have dose on the horizontal axis, since it is an ordinal (quantitative) factor. This was ensured in the `interaction.plot` command in R by specifying `dose` as the `x.factor`. Since the levels of the factor are ordered categories, however, we would like them in the correct order: Low, Medium, High. The `interaction.plot` command in R sorts them alphanumerically though, which is not what we want. We can override this by giving the levels of factor `dose` new names; e.g. a for "low", b for "medium" and c for "high". Or, perhaps more clearly, 1.Low, 2.Medium, 3.High —which still ensures we know what the original names of the levels were. Labels can be given to the levels of the factor if desired, as in the R code given in part (f) below, which can then be used in the `interaction.plot` command.

- f. Work out the cell means from the data table above, and use them to sketch your own better interaction graph. Use your sketch to interpret your statistical results. How could you get those cell means from the fitted model in R?

The means of the response variable, **Drug concentration**, go on the vertical axis. It is better to put the factor `dose` on the horizontal axis, as it is ordinal, and to draw one line each for the two levels of `delivery`, i.e., **Patch** and **Capsule**. Make sure to put the `dose` levels in ascending order.

Cell means for the interaction graph are given below. Simply average the two values in each cell or, equivalently, go half way between them.

	Low dose	Medium dose	High dose
Patch delivery	13	18	22.5
Capsule delivery	12.5	24	34.5

Note that the cell means are the predicted values for each cell from the complete model, so these are available and we have already used them earlier in our diagnostic graph to check for constant variance. Here's some R code (below) to give the cell means, and also to draw an interaction graph with the desired ordering for the 'dose' levels.

**Comments:** The significant interaction shows up as non-parallel lines in the interaction graph. Note that at low doses, the two delivery methods have similar effectiveness, but at high doses capsule delivery is more effective than patch delivery.

*What could cause this? It's a matter for the scientist to consider. Perhaps there's a maximum rate of transfer through the skin, so the high dose is limited in its transfer, while no such limiting mechanism operates with the capsules.*

```
# List all 12 predicted values from the fitted model.
drug.ANOVA$fitted.values
```

```
 1  2  3  4  5  6  7  8  9 10 11 12
13.0 13.0 18.0 18.0 22.5 22.5 12.5 12.5 24.0 24.0 34.5 34.5
```

```
# List just the cell means (these are all different here).
# Need to round to 1dp though, since "equal" floating point numbers are
# not necessarily equal when identified using the unique() function.
unique(round(drug.ANOVA$fitted.values,1))
```

```
[1] 13.0 18.0 22.5 12.5 24.0 34.5
```

```
# Create a new dose factor, with the levels in the
# same order as originally entered (to match the concentration data),
# but now ordered numerically, so that low comes first.
newdose <- rep(rep(c("1.Low", "2.Medium", "3.High"), each = 2), times = 2)
```

```
# Give the desired labels to the newly created dose factor
dose.new <- factor(newdose, labels = c("Low", "Medium", "High"))
```

```
# Interaction graph for drug concentration data.
# Now has a new dose factor with levels in the desired order.
interaction.plot(x.factor = dose.new,
                trace.factor = delivery,
                response = conc,
                fun = mean,
                xlab = "Dose",
                ylab = "Mean drug concentration",
                main = "Interaction graph for drug delivery",
                legend = T, xpd=T,
                trace.label="Delivery")
```

Interaction graph for drug delivery

