# 6 Two-factor Experiments

Last week we considered a CRD (completely randomized design) for comparing insecticides where the levels of one factor (insecticide) vary while controlling other factors that influence survival time. The inferences from the one-way ANOVA apply to beetles with a given age from the selected strain that might be given the selected concentration of the insecticides. Any generalization of the conclusions to other situations must be justified scientifically, typically through further experimentation.

Recall the way we set up the model: $y_{ij}$ is the response for the $j^{th}$ experimental unit (replicate) in the $i^{th}$ treatment group, where $i = 1, 2, ..., I$;

$$y_{ij} = \mu_i + \epsilon_{ij},$$

where $\mu_i$ is the (unknown) population mean for all potential responses to the $i^{th}$ treatment, and $\epsilon_{ij}$ is the residual or deviation of the response from the population mean. The responses within and across treatments are assumed to be independent, normal random variables with constant variance. We further decomposed $\mu_i$ as $\mu_i = \mu + \alpha_i$.

There are several ways to broaden the scope of the study. For example, several strains of beetles or several concentrations of the insecticide might be used. For simplicity, consider a simple two-factor experiment where three concentrations (Low, Medium, and High) are applied with each of the four insecticides. This is a completely crossed **two-factor experiment** where each of the $4 \times 3 = 12$ combinations of the two factors (insecticide and dose) are included in the comparison of survival times. With this experiment, the scientist can compare insecticides, compare concentrations, and check for an interaction between dose and insecticide.

Assuming that 48 beetles are available, the scientist would randomly assign them to the 12 experimental groups, giving prespecified numbers of beetles to the 12 groups. For simplicity, assume that the experiment is **balanced**, that is, the same number of beetles (4) is assigned to each group ($12 \times 4 = 48$). This is a CRD with two factors.

### A Balanced Two-Factor Model

We will analyze survival times of groups of four beetles randomly allocated to twelve treatment groups obtained by crossing the levels of four insecticides (1,2,3,4) at each of three concentrations of the insecticides (1=Low, 2=Medium, 3=High). This is a balanced 4-by-3 **factorial** design (two-factor design) that is replicated four times. Three variables are needed to uniquely represent each response in the spreadsheet: dose (1-3, nominal), insecticide (1-4, nominal), and the survival time (called time). The unit of measure for the survival times is 10 hours. That is, .3 is a survival time of 3 hours. The data are given below, collected into 12 **cells** (4 rows and 3 columns):

| | Dose | | |
|---|---|---|---|
| Insecticide | 1 | 2 | 3 |
| 1 | .31, .45, .46, .43 | .36, .29, .40, .23 | .22, .21, .18, .23 |
| 2 | .82, 1.10, .88, .72 | .92, .61, .49, 1.24 | .30, .37, .38, .29 |
| 3 | .43, .45, .63, .76 | .44, .35, .31, .40 | .23, .25, .24, .22 |
| 4 | .45, .71, .66, .62 | .56, 1.02, .71, .38 | .30, .36, .31, .33 |

We model this in terms of population means, just as we did in the one-way ANOVA. Now, though, population means are indexed two ways, by insecticide and by dose, so we write

$$y_{ijk} = \mu_{ij} + \epsilon_{ijk}; \;\; i = 1, 2, \ldots, I; \;\; j = 1, 2, \ldots, J; \;\; k = 1, 2, \ldots, K$$

where i refers to insecticide (I=4), j refers to dose (J=3), and k refers to replicate (K=4). For instance $y_{314} = .76$. Since we have the same number (4) of replicates in each cell this is called balanced. More generally it happens that $k = 1, 2, \ldots, K_{ij}$, i.e. there can be different numbers of replicates (usually not designed that way, but things happen!), and the analysis gets somewhat more complicated. We will consider an unbalanced problem later.

In the one-way problem, the basic test of hypothesis is that all the means are equal. That is not very useful here. What we want to do is compare insecticides, compare doses, and see if the effect of dose varies with insecticide. We need to define some additional population averages to attack all these hypotheses. The population marginal mean for Insecticide i is $\bar{\mu}_{i.} = \frac{1}{J}\sum_{j=1}^{J}\mu_{ij} = \frac{1}{3}\sum_{j=1}^{3}\mu_{ij}$, the average of Insecticide i across Dose levels. The population marginal mean for Dose j is $\bar{\mu}_{.j} = \frac{1}{I}\sum_{i=1}^{I}\mu_{ij} = \frac{1}{4}\sum_{i=1}^{4}\mu_{ij}$, the average of Dose j across Insecticide levels. There also is an overall population average, $\bar{\mu}_{..} = \frac{1}{IJ}\sum_{i=1}^{I}\sum_{j=1}^{J}\mu_{ij} = \frac{1}{12}\sum_{i=1}^{4}\sum_{j=1}^{3}\mu_{ij}$, the average of all 12 population means. What we are interested in is the structure in the following table of population mean values:

| | Dose | | | |
|---|---|---|---|---|
| Insecticide | 1 | 2 | 3 | Insecticide marginal |
| 1 | $\mu_{11}$ | $\mu_{12}$ | $\mu_{13}$ | $\bar{\mu}_{1.}$ |
| 2 | $\mu_{21}$ | $\mu_{22}$ | $\mu_{23}$ | $\bar{\mu}_{2.}$ |
| 3 | $\mu_{31}$ | $\mu_{32}$ | $\mu_{33}$ | $\bar{\mu}_{3.}$ |
| 4 | $\mu_{41}$ | $\mu_{42}$ | $\mu_{43}$ | $\bar{\mu}_{4.}$ |
| Dose marginal | $\bar{\mu}_{.1}$ | $\bar{\mu}_{.2}$ | $\bar{\mu}_{.3}$ | $\bar{\mu}_{..}$ |

The basic unit of analysis is **sample cell means**, which are the direct estimators of the above population averages. We have a sample of K observations in cell ij - the natural estimator of $\mu_{ij}$ is $\bar{y}_{ij.} = \frac{1}{K}\sum_{k=1}^{K} y_{ijk} = \frac{1}{4}\sum_{k=1}^{4} y_{ijk}$ . We define sample marginal means as we did for population values above (row averages and column averages), $\bar{y}_{i..} = \frac{1}{J}\sum_{j=1}^{J}\bar{y}_{ij.}$, $\bar{y}_{.j.} = \frac{1}{I}\sum_{i=1}^{I}\bar{y}_{ij.}$, and $\bar{y}_{...} = \frac{1}{IJ}\sum_{i=1}^{I}\sum_{j=1}^{J}\bar{y}_{ij.}$ . this gives us natural estimators of the above population values as the sample values:

| | Dose | | | |
|---|---|---|---|---|
| Insecticide | 1 | 2 | 3 | Insecticide marginal |
| 1 | $\bar{y}_{11.}$ | $\bar{y}_{12.}$ | $\bar{y}_{13.}$ | $\bar{y}_{1..}$ |
| 2 | $\bar{y}_{21.}$ | $\bar{y}_{22.}$ | $\bar{y}_{23.}$ | $\bar{y}_{2..}$ |
| 3 | $\bar{y}_{31.}$ | $\bar{y}_{32.}$ | $\bar{y}_{33.}$ | $\bar{y}_{3..}$ |
| 4 | $\bar{y}_{41.}$ | $\bar{y}_{42.}$ | $\bar{y}_{43.}$ | $\bar{y}_{4..}$ |
| Dose marginal | $\bar{y}_{.1.}$ | $\bar{y}_{.2.}$ | $\bar{y}_{.3.}$ | $\bar{y}_{...}$ |

Values calculated with these data are as follows:

| | Dose | | | |
|---|---|---|---|---|
| Insecticide | 1 | 2 | 3 | Insect marg |
| 1 | .413 | .320 | .210 | .314 |
| 2 | .880 | .815 | .335 | .677 |
| 3 | .568 | .375 | .235 | .393 |
| 4 | .610 | .667 | .325 | .534 |
| Dose marg | .618 | .544 | .276 | .479 |

Because the experiment is balanced, a marginal mean also is the average of all observations that receive a given treatment. For example, the marginal mean for insecticide 1 is the average survival time for the 12 beetles given insecticide 1.

The basic model for the two-factor design, as applied to this experiment is that the

$$\text{Response} = \text{Grand Mean} + \text{Insect Effect} + \text{Dose Effect} + \text{Insect} * \text{Dose Interaction} + \text{Residual}.$$

or

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

i.e. $\mu_{ij} = \mu + \alpha_i + \beta_j + (\alpha\beta)_{ij}$ . The assumptions for the analysis of the model are identical to those for a one-way ANOVA on the 12 treatment combinations (insecticide and dose), i.e. all 48 residual effects $\epsilon_{ijk}$ are independent and variances are all the same, $\sigma^2_{ij} = \sigma^2$. There are constraints put on the terms above, since there are too many, but those depend upon the software used. We will discuss this at some length.

The ANOVA table for this experimental design decomposes the total variation in the data, as measured by the Total SS, into components that measure the variation of marginal sample means of insecticide and dose individually (the Insecticide SS and Dose SS), a component that measures the degree to which the factors interact (the insecticide by dose SS), and a component that pools the sample variances across the 12 samples (the Error SS). Each SS has a df, given in the following ANOVA table. As usual, the MS for each source of variation is the corresponding SS divided by the df. The MS Error estimates the common population variance for the 12 treatments.

| Source | df | SS | MS = SS/df |
|---|---|---|---|
| Insecticide | $I-1=4-1=3$ | $JK\sum_{i=i}^{I}(\bar{y}_{i..}-\bar{y}_{...})^2$ | |
| Dose | $J-1=3-1=2$ | $IK\sum_{j=i}^{J}(\bar{y}_{.j.}-\bar{y}_{...})^2$ | |
| Interaction | $(I-1)(J-1)=(3)(2)=6$ | $K\sum_{i=1}^{I}\sum_{j=1}^{J}(\bar{y}_{ij.}-\bar{y}_{i.}-\bar{y}_{.j.}+\bar{y}_{...})^2$ | |
| Error | $IJ(K-1)=(4)(3)(3)=36$ | $\sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{k=1}^{K}(y_{ijk}-\bar{y}_{ij.})^2$ | |
| Total | $IJK-1=48-1=47$ | $\sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{k=1}^{K}(y_{ijk}-\bar{y}_{...})^2$ | |

Believe it or not, these formulas actually make sense! SS for Insecticide just compares the Insecticide marginal means to each other by computing their variance (up to a constant), and similarly with the SS for Dose. SS Total is the usual sum of all squared deviations from the overall mean. MS Error is just $\frac{1}{IJ}\sum_{i=1}^{I}\sum_{j=1}^{J}s^2_{ij} = s^2_{pooled}$, the average of the sample variances from each of the cells, a natural way to estimate $\sigma^2$ (this is the within cell variability). The only term that is not easily understood is SS for Interaction. We will turn to that in a little while (after we have decided what interaction is).

There are three usual tests of interest.

1. The test of no insecticide effect. The absence of an insecticide effect implies that each level of insecticide has the same **population mean** response **when the means are averaged over levels of** dose. The test for no insecticide effect is based on the p-value for the $F-$statistic: $F_{obs}$ = MS Insecticide/MS Error. This hypothesis is rejected when the insecticide marginal means vary significantly relative to the within cell variation. Formally, this is a test of $H_0 : \bar{\mu}_{1.} = \bar{\mu}_{2.} = \bar{\mu}_{3.} = \bar{\mu}_{4.}(= \bar{\mu}_{..})$. The form of the SS certainly matches this hypothesis.

2. The test of no dose effect. The absence of a dose effect implies that each dose level has the same **population mean** response **when the means are averaged over levels of** insecticide. The test for no dose effect is based on the p-value for the $F-$statistic: $F_{obs}$ = MS Dose/MS Error. This hypothesis is rejected when the marginal means for dose vary significantly relative to the within cell variation. Formally, this is a test of $H_0 : \bar{\mu}_{.1} = \bar{\mu}_{.2} = \bar{\mu}_{.3}(= \bar{\mu}_{..})$. The form of the SS certainly matches this hypothesis.

3. The test of no interaction between dose and insecticide is based on the p-value for the $F-$statistic: $F_{obs}$ = MS Interaction/MS Error. This is a test of a hypothesis that the structure is simple. Let's explore what that means.

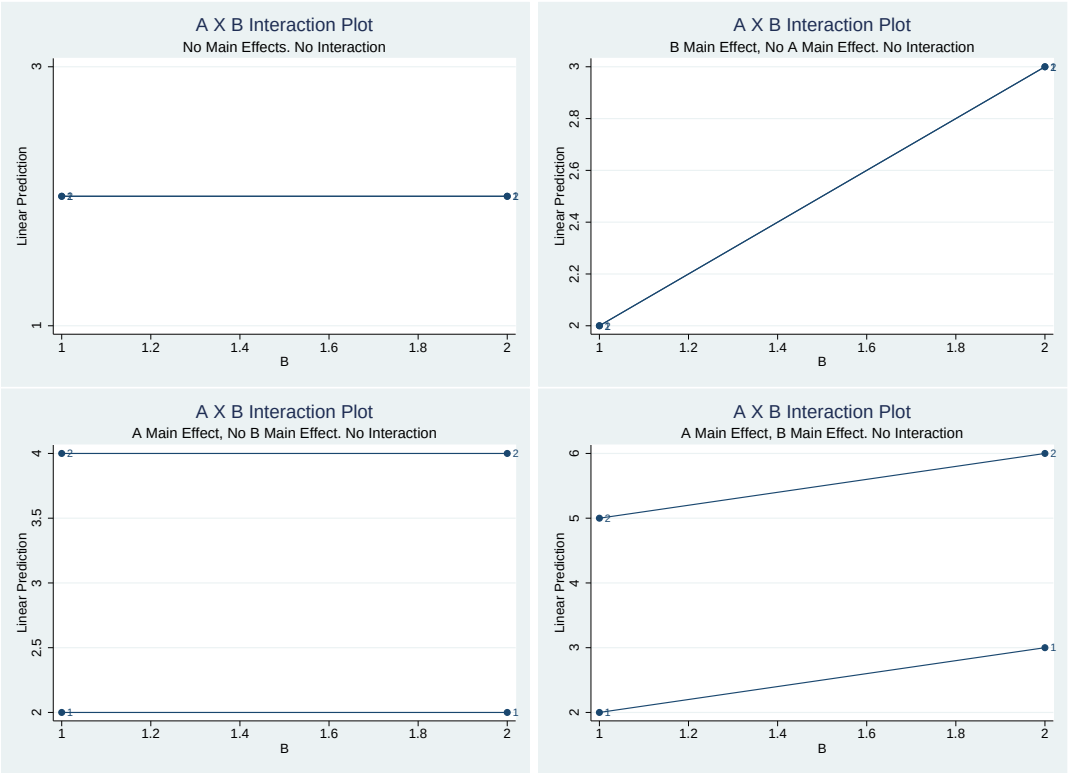
## Interpretation of Interaction

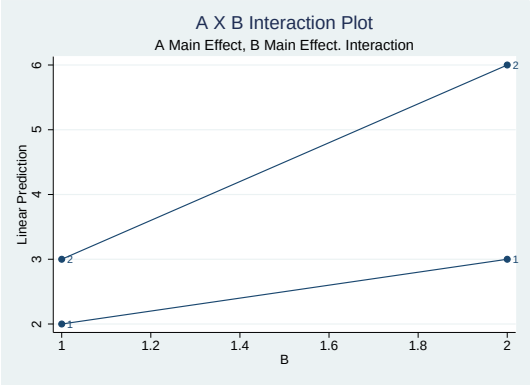
The idea of **no** interaction is that the margins of the table tell you what the structure is. If row 2, for instance, is "large" for one column, it is similarly large for all the other columns. This gets awkward to quantify, though, and we need a better approach. We have already seen the decomposition $\mu_{ij} = \mu + \alpha_i + \beta_j + (\alpha\beta)_{ij}$ . This imposes no restrictions on the cell means. We *can* impose a restriction if we force $(\alpha\beta)_{ij} \equiv 0$ so that $\mu_{ij} = \mu + \alpha_i + \beta_j$ . This **additive** model is how we force the margins of the table to tell us everything. The hypothesis of no interaction is thus formally $H_0 : (\alpha\beta)_{ij} = 0$ for all i and j.

There are serious implications of this hypothesis. One is that $\mu_{ij} = \bar{\mu}_{i.} + \bar{\mu}_{.j} - \bar{\mu}_{..}$ . If you look back at the SS for Interaction, this is exactly what is being tested. That is not nearly as useful or interesting, though, as this: If $i, i', j, j'$ are legal indexes, then $\mu_{ij} - \mu_{ij'} = \mu_{i'j} - \mu_{i'j'}$, which is to say the difference between doses j and j' is the same for insecticide i as for insecticide i'; and $\mu_{ij} - \mu_{i'j} = \mu_{ij'} - \mu_{i'j'}$, which is to say the difference between insecticides i and i' is the same for dose j as it is for dose j'. These differences in cell means are slopes of line segments in interaction plots. What no interaction tells you is that the slopes of the line segments (connecting) sample cell means should be approximately parallel, and the formal test for no interaction is a check on whether the profile plots of the population means are perfectly parallel.

## Prototype Interaction Plots

These profile plots are extremely important tools for understanding our analysis, so let us examine various possible patterns. Consider the simplest example with two factors A and B each at two levels, and let the population cell means $\mu_{ij}$ be broken down as $\mu_{ij} = \mu+\alpha_i+\beta_j+(\alpha\beta)_{ij}$. Following are plots of those population cell means (those based on data have noise in them so will not be so perfect) for various combinations of effects present. Make sure you understand why each appears as it does and match it with the appropriate model.






## Stata Analysis of Insecticide Data

The data need to go into three columns in the spreadsheet with full information for each observation as follows (note that I folded output to save space – this should go on for 48 rows):

| poison | dose | time | poison | dose | time | poison | dose | time |
|---|---|---|---|---|---|---|---|---|
| 1 | 1 | .43 | 2 | 2 | 1.24 | 3 | 3 | .23 |
| 1 | 1 | .46 | 2 | 2 | .61 | 3 | 3 | .25 |
| 1 | 1 | .31 | 2 | 2 | .49 | 3 | 3 | .24 |
| 1 | 1 | .45 | 2 | 2 | .92 | 3 | 3 | .22 |
| 1 | 2 | .4 | 2 | 3 | .29 | 4 | 1 | .66 |
| 1 | 2 | .29 | 2 | 3 | .3 | 4 | 1 | .45 |
| 1 | 2 | .36 | 2 | 3 | .37 | 4 | 1 | .62 |
| 1 | 2 | .23 | 2 | 3 | .38 | 4 | 1 | .71 |
| 1 | 3 | .23 | 3 | 1 | .63 | 4 | 2 | .56 |
| 1 | 3 | .21 | 3 | 1 | .43 | 4 | 2 | .71 |
| 1 | 3 | .18 | 3 | 1 | .76 | 4 | 2 | 1.02 |
| 1 | 3 | .22 | 3 | 1 | .45 | 4 | 2 | .38 |
| 2 | 1 | .88 | 3 | 2 | .4 | 4 | 3 | .3 |
| 2 | 1 | 1.1 | 3 | 2 | .35 | 4 | 3 | .31 |
| 2 | 1 | .72 | 3 | 2 | .31 | 4 | 3 | .36 |
| 2 | 1 | .82 | 3 | 2 | .44 | 4 | 3 | .33 |

The table of cell and marginal means

| | Dose | | | |
|---|---|---|---|---|
| Insecticide | 1 | 2 | 3 | Insect marg |
| 1 | .413 | .320 | .210 | .314 |
| 2 | .880 | .815 | .335 | .677 |
| 3 | .568 | .375 | .235 | .393 |
| 4 | .610 | .667 | .325 | .534 |
| Dose marg | .618 | .544 | .276 | .479 |

is produced within **Stata** thus

```
. tabulate poison dose,summarize(time) means
                           Means of time

           |                  dose
    poison |         1          2          3 |     Total
-----------+---------------------------------+----------
         1 |     .4125        .32        .21 | .31416667
         2 | .88000001  .81500001       .335 | .67666667
         3 |     .5675       .375       .235 |     .3925
         4 |       .61  .66749999  .32500001 | .53416667
-----------+---------------------------------+----------
     Total |     .6175   .544375    .27625 |   .479375
```

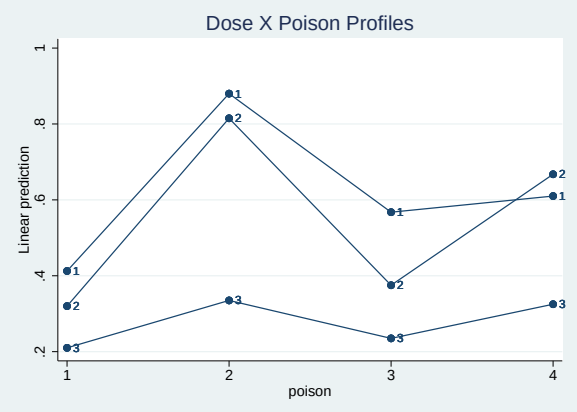
The ANOVA table is produced using the `anova` command forcing the software to fit the model $\mu_{ij} = \mu + \alpha_i + \beta_j + (\alpha\beta)_{ij}$

```
. anova time poison dose poison*dose
                          Number of obs =      48     R-squared     =  0.7335
                          Root MSE      = .149139     Adj R-squared =  0.6521
                  Source |  Partial SS    df       MS           F     Prob > F
              -----------+----------------------------------------------------
                   Model |  2.20435628    11  .200396025       9.01     0.0000
                         |
                  poison |  .921206282     3  .307068761      13.81     0.0000
                    dose |  1.03301249     2  .516506246      23.22     0.0000
             poison*dose |  .250137502     6  .041689584       1.87     0.1123
                         |
                Residual |  .800724989    36  .022242361
              -----------+----------------------------------------------------
                   Total |  3.00508126    47  .063937899
```

To examine interaction, consider the Dose*Insecticide profile plot given below. For each dose, we have a plot of the mean survival times across insecticides, giving 3 profiles. There is no interaction in the data if these profiles are parallel. The formal test for no interaction is a check on whether the profile plots of the population means are perfectly parallel. Every statistical package requires some special means to obtain these plots. In earlier versions of Stata we downloaded a command named `cmeans`, but that appears no longer available. The method now is easy enough if a little obscure looking:

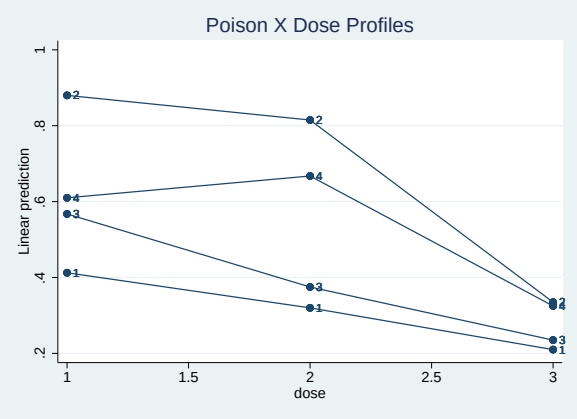
```
anova time poison dose poison*dose
predict yhat,xb
sort dose poison
scatter yhat poison,c(L) mlabel(dose) title(Dose X Poison Profiles)
```

The points being plotted by this method are those actually fit by the `anova` model – with the `poison*dose` term we are imposing no restrictions so these are the averages we calculated earlier. Had we left that term off we would have forced the profiles to be parallel since we would have imposed an additive model. The order of the sort is very important here. The `c(L)` option is a very special connected line version suited to this application (`c(l)` does not work). `mlabel` lets us label dose levels.




The last two lines above could be modified to produce four profiles, one for each poison. The plots are equally useful – sometimes both are worth examining.

```
sort poison dose
scatter yhat dose,c(L) mlabel(poison) title(Poison X Dose Profiles)
```




These lines are not parallel, but they have the same general trend. If we accept that the interaction is not significant, we may fit the additive model instead and base inferences on additive structure. To fit the model with the interaction term we type `anova time poison dose poison*dose`; the additive model is fit using `anova time poison dose`.

### Interpretation of the ANOVA

The "Model" row in the ANOVA table gives a p-value for testing no differences among the population mean survival times for the 12 dose and insecticide combinations (or whatever model we fit – in this case we allow all 12 means to vary with no restriction because we fit the interaction term). The p-value of .0000 strongly suggests that the population mean survival times are not equal across all 12 groups.

The ANOVA table gives a breakdown of the Model SS into the SS for insecticide, dose, and the insecticide by dose interaction. The Mean Squares, F-statistics and p-values for testing these effects are given. The p-values indicate that the dose and insecticide effects are significant at the .0001 level. However, the F-test for no dose by insecticide interaction is not significant at the .10 level (p-value=.1123).

The cell means give us an idea about the nature of the differences among doses and insecticides (the F-tests only tell us if *some* difference appears to be there). In particular, the insecticides have noticeably different mean survival times averaged over doses, with insecticide 1 having the lowest mean survival time averaged over doses. Similarly, higher doses tend to produce lower survival times. More formal comparisons of the doses and insecticides are possible using the output from the Tukey comparisons of LS MEANS in **JMP-IN** or **SAS**, or from Fisher comparisons in **Stata**. For example, using **JMP-IN** output for the Tukey comparisons (not shown here), the high dose is significantly different from the low and medium doses, which are not significantly different from each other. We obtain Fisher's method in **Stata** using the commands

```
test _b[dose[2]]=_b[dose[1]]
test _b[dose[3]]=_b[dose[1]]
test _b[dose[3]]=_b[dose[2]]
```

These test the hypotheses $H_0 : \beta_1 = \beta_2$, $H_0 : \beta_1 = \beta_3$, and $H_0 : \beta_2 = \beta_3$ respectively. This is the same as testing $H_0 : \bar{\mu}_{.1} = \bar{\mu}_{.2}$, $H_0 : \bar{\mu}_{.1} = \bar{\mu}_{.3}$, and $H_0 : \bar{\mu}_{.2} = \bar{\mu}_{.3}$ We obtain:

```
 ( 1) - dose[1] + dose[2] = 0
       F(  1,    36) =    0.30
            Prob > F =    0.5889
 ( 1) - dose[1] + dose[3] = 0
       F(  1,    36) =    7.30
```

```
           Prob > F =    0.0104
 ( 1) - dose[2] + dose[3] = 0
       F(  1,    36) =   10.55
            Prob > F =    0.0025
```

We see that doses 1 and 2 are not significantly different from each other, but dose 3 is significantly different from doses 1 or 2, *averaged over the poison effects*. Bonferroni comparisons simply multiply the above p-values by 3 (the number of comparisons), so Bonferroni, Tukey, and Fisher all agree here.

Assuming the interaction is not important, we can obtain the three estimated pairwise differences in the three doses *using any poison* with the commands

```
lincom _b[dose[2]]-_b[dose[1]]
lincom _b[dose[3]]-_b[dose[1]]
lincom _b[dose[3]]-_b[dose[2]]
```

we obtain

```
. lincom _b[dose[2]]-_b[dose[1]]
------------------------------------------------------------------------------
        time |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
         (1) |      .0575    .105457     0.55   0.589    -.1563767    .2713767
------------------------------------------------------------------------------
. lincom _b[dose[3]]-_b[dose[1]]
------------------------------------------------------------------------------
        time |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
         (1) |      -.285    .105457    -2.70   0.010    -.4988767   -.0711233
------------------------------------------------------------------------------
. lincom _b[dose[3]]-_b[dose[2]]
------------------------------------------------------------------------------
        time |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
         (1) |     -.3425    .105457    -3.25   0.003    -.5563767   -.1286233
------------------------------------------------------------------------------
```

We see, for instance, that we are 95% confident that the mean difference in survival time between dose 2 and dose 3 is between 0.13 and 0.56. Put another way, beetles given dose 2 last between 1.3 and 5.6 hours longer *on average* than those given dose 3, *regardless of the insecticide used*. The last part of this statement would not hold if we had an important interaction and a more detailed analysis of *how* the difference changed with the insecticide used would be warranted.

Results from fitting the *additive model* are similar, although the confidence intervals are tighter:

```
. lincom _b[dose[2]]-_b[dose[1]]
------------------------------------------------------------------------------
        time |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
         (1) |   -.073125   .0559247    -1.31   0.198    -.1859855    .0397355
------------------------------------------------------------------------------
. lincom _b[dose[3]]-_b[dose[1]]
------------------------------------------------------------------------------
        time |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
         (1) |    -.34125   .0559247    -6.10   0.000    -.4541105   -.2283895
------------------------------------------------------------------------------
. lincom _b[dose[3]]-_b[dose[2]]
------------------------------------------------------------------------------
        time |      Coef.   Std. Err.      t    P>|t|     [95% Conf. Interval]
-------------+----------------------------------------------------------------
         (1) |   -.268125   .0559247    -4.79   0.000    -.3809855   -.1552645
------------------------------------------------------------------------------
```

### More Interpretation of the Dose Effect

The interpretation of the dose and insecticide effects (called the **main-effects**) depends on whether interaction is present. The distinction is important, so I will give both interpretations to emphasize the differences. Given that the test for interaction was not significant, I would likely summarize the main effects assuming no interaction. For simplicity, I will restrict attention to the dose effect.

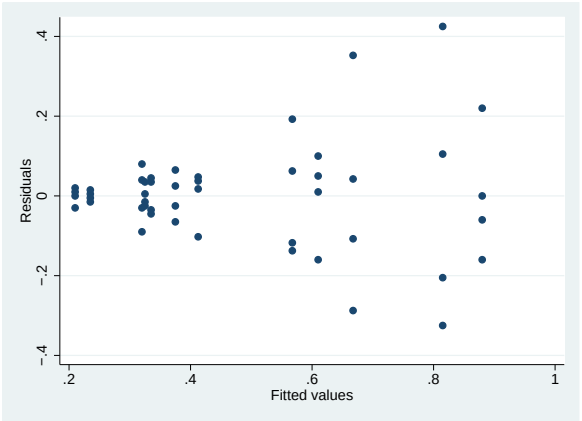
The average survival time decreases as the dose increases, with estimated mean survival times of .618, .544, and .276, respectively. If dose and insecticide **interact**, you can conclude that beetles given a high dose of the insecticide typically survive for shorter periods of time **averaged over insecticides**. You can not, in general, conclude that the highest dose yields the lowest survival time **regardless** of insecticide. For example, the difference in the medium and high dose marginal means of .544 - .276 = .268 estimates the typical decrease in survival time achieved by using the high dose instead of the medium dose, averaged over insecticides.

If the two factors interact, then the difference in mean times between the medium and high doses on a given insecticide may be significantly greater than .268, significantly less than .268, or even negative. In the latter case the medium dose would be **better** than the high dose for the given insecticide, even though the high dose gives better performance averaged over insecticides. An interaction forces you to use the cell means to decide which combination of dose and insecticide gives the best results.

If dose and insecticide **do not interact**, then the difference in marginal dose means averaged over insecticides also estimates the difference in population mean survival times between two doses, **regardless of the insecticide**. This follows from the parallel profiles definition of no interaction. Thus, the difference in the medium and high dose marginal means (.544 - .276 = .268) estimates the expected decrease in survival time anticipated from using the high dose instead of the medium dose, **regardless of the insecticide** (and hence also when averaged over insecticides).

A practical implication of no interaction is that you can conclude that the high dose is best, regardless of the insecticide used. The difference in marginal means for two doses estimates the difference in average survival expected, regardless of the insecticide.

As a final note, I will mention that the residual plot suggests that the variability in the survival times increases with increasing mean (obtained using `rvfplot`). A transformation to the reciprocal scale (which turns the response into a rate) is often suggested with these data. You should repeat the analysis on that scale to see the improvement.



## An Unbalanced Two-Factor Experiment and Analysis

The sample sizes are rarely equal for the different treatments in an experiment. This has no consequence on the specification of a model, and we proceed as in the balanced case.

## Example: Insulin Levels in Rats

The data below are the insulin levels in rats a certain length of time after a fixed dose of insulin was injected into their jugular or portal veins. This is a two-factor study with two vein types (jugular=1, portal=2) and three time levels (0 minutes = 1, 30 minutes = 2, and 60 minutes = 3). A feature of this experiment is that the rats used in the six vein and time combinations are distinct. I will fit a two-factor interaction model, which assumes that the responses are independent within and across treatments. The design is unbalanced, with sample sizes varying from 3 to 12.

| Vein | Time | Insulin Levels | | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|---|---|---|
| jugular | 0 | 18 | 36 | 12 | 24 | 43 | | | | | | | |
| jugular | 30 | 61 | 116 | 63 | 132 | 68 | 37 | | | | | | |
| jugular | 60 | 18 | 133 | 33 | | | | | | | | | |
| portal | 0 | 96 | 72 | 34 | 41 | 98 | 77 | 120 | 49 | 92 | 111 | 99 | 94 |
| portal | 30 | 146 | 193 | 78 | 127 | 136 | 144 | 115 | 199 | 253 | 338 | | |
| portal | 60 | 132 | 110 | 141 | 204 | 69 | 152 | 196 | 195 | 84 | 105 | 71 | 83 |

An alternative experimental design might randomly assign rats to the two vein groups, and then measure the insulin levels of each rat at the three time points. Depending on the questions of interest, you could compare veins using a one-way MANOVA, or a repeated measures design that allows correlated responses within rats.

The model written abstractly is

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

Here, $i = 1, 2$ denote the two vein types, $j = 1, 2, 3$ denote the three times, and $k = 1, 2, \ldots, K_{ij}$ denotes the $k^{th}$ rat out of $K_{ij}$ in the group with vein $i$ and time $j$. You should verify that $K_{11} = 5$, $K_{12} = 6$, $K_{13} = 3$, $K_{21} = 12$, $K_{22} = 10$, and $K_{13} = 12$. The `tabulate` command makes that fairly easy

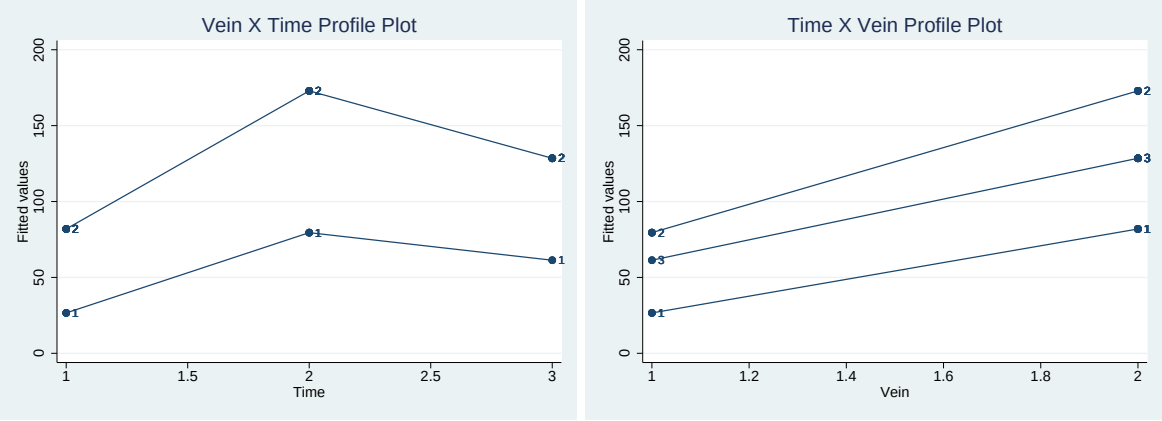
```
. tabulate vein time,summarize(insulin)
           Means, Standard Deviations and Frequencies of Insulin
           |              Time
      Vein |         1          2          3 |     Total
-----------+---------------------------------+----------
         1 |      26.6       79.5  61.333333 | 56.714286
           |  12.75931   36.44585  62.516664 | 41.899933
           |         5          6          3 |        14
-----------+---------------------------------+----------
         2 | 81.916667      172.9      128.5 | 125.11765
           | 27.747099  76.117526  49.718297 | 63.525115
           |        12         10         12 |        34
-----------+---------------------------------+----------
     Total | 65.647059    137.875  115.06667 | 105.16667
           | 35.284453   78.10239  57.218212 | 65.621848
           |        17         16         15 |        48
```

In order to get the interaction (profile) plots, we need to fit the ANOVA with interaction present.

```
. anova insulin vein time vein*time
                          Number of obs =      48     R-squared     =  0.4915
                          Root MSE      = 49.5009     Adj R-squared =  0.4310
                  Source |  Partial SS    df       MS           F     Prob > F
              -----------+----------------------------------------------------
                   Model |  99478.4833     5  19895.6967       8.12     0.0000
                         |
                    vein |  48212.7037     1  48212.7037      19.68     0.0001
                    time |   37734.188     2   18867.094       7.70     0.0014
                vein*time |   2745.9139     2  1372.95695       0.56     0.5752
                         |
                Residual |  102914.183    42   2450.3377
              -----------+----------------------------------------------------
                   Total |  202392.667    47  4306.22695
```

It probably makes sense to look at both profile plots:

```
predict yhat
sort vein time
scatter yhat time, c(L) ml(vein) title(Vein X Time Profile Plot)
sort time vein
scatter yhat time, c(L) ml(vein) title(Vein X Time Profile Plot)
```




The profile or *interaction* plots show roughly parallel profiles indicating that the interaction term may not be important. The profile plots, along with the table of means , indicate that the insulin level is at its highest (of the three times considered) at 30 minutes for either vein considered alone, or averaged over veins. The portal vein yields a higher insulin level at any of the three time periods and averaged over the three time periods.

The ANOVA table indicates that the vein and time effects are significant, with p-values of .0001 and .0014, respectively, but that the interaction is not significant (p-value=.575). Recall that the profiles are reasonably parallel, which is consistent with a lack of interaction.

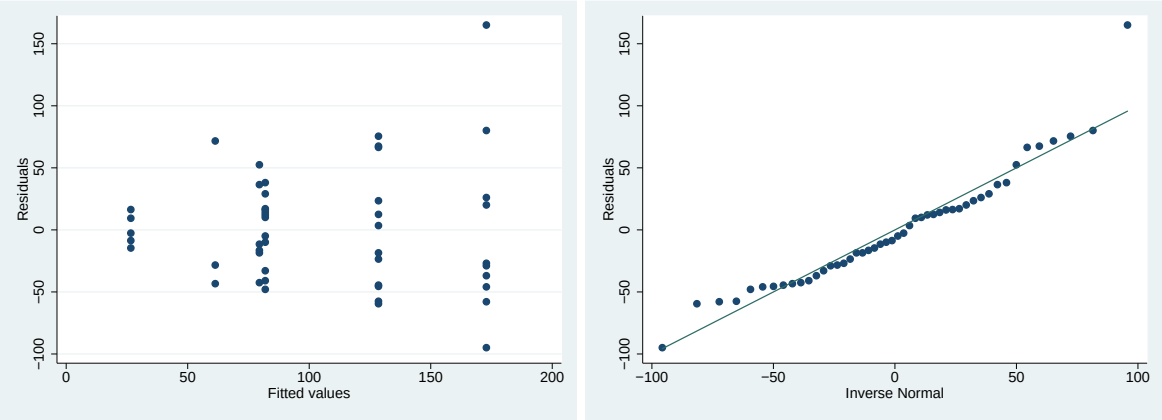
The means table above shows that the mean insulin level in the portal vein is significantly greater than the mean insulin level in the jugular vein. Because of the lack of interaction, the difference in mean levels for the portal veins is reasonably consistent across times.

Since we accept that there is no interaction here, it makes sense to compare the overall main effects *vein* and *time* using pairwise comparisons. A test that there is no difference in vein types (`test _b[vein[2]] = _b[vein[1]]`) yields a $p$-value of 0.0416; we reject that there is no difference. We estimate the difference in insulin levels from the portal versus the jugular veins in **Stata** using the command `lincom _b[vein[2]]-_b[vein[1]]` and find the estimate to be $\hat{\alpha}_2 - \hat{\alpha}_1 = 67.2$ with a 95% CI of (2.7, 132), *independent of time*. This is close to the estimate obtained from the marginal

means as 125 - 57 = 68, *but need not be in unbalanced designs!!* For this reason, one should always use the model estimates from **Stata** rather than estimates obtained from looking at a raw means table.

Since there is no interaction here, the difference in insulin levels is the same at time = 0, time = 30, and time = 60 minutes. Recall that when no interaction is present we say the model is *additive*. Similarly we may look at *differences* in insulin levels at the three times *independent of vein type*. The $p$-values for testing that there is no difference between (1) 30 minutes and 0 minutes, (2) 60 minutes and 0 minutes, and (3) 60 minutes and 30 minutes are (1) 0.0001, (2) 0.0262, and (3) 0.0423. These are Fisher values. The corresponding Bonferroni-adjusted values are obtained by multiplying each by $\begin{pmatrix} 3 \\ 2 \end{pmatrix} = 3$: (1) 0.0003 (2) 0.078, and (3) 0.13; the only significant difference is between 30 minutes and 0 minutes. Using `lincom` we would estimate this difference in insulin levels to be about 72, *independent of vein type*.

What are your thoughts on the residual plots?




Finally, note that we accept $H_0 : (\alpha\beta)_{ij} = 0$ at any reasonable significance level. We can thus fit and base inferences on the additive model if we choose to do so. The ANOVA table is

```
      Number of obs =      48     R-squared     =  0.4779
      Root MSE      = 49.0037     Adj R-squared =  0.4424
------------------------------------------------------------------------------
      Source |  Partial SS    df       MS           F     Prob > F
-------------+----------------------------------------------------------------
       Model |  96732.5694     3  32244.1898      13.43     0.0000
             |
        vein |  51594.4685     1  51594.4685      21.49     0.0000
        time |  50332.2893     2  25166.1447      10.48     0.0002
             |
    Residual |  105660.097    44  2401.36585
-------------+----------------------------------------------------------------
       Total |  202392.667    47  4306.22695
```

Now, using estimates from the additive model, we obtain the estimated mean difference in vein effects to be $\hat{\alpha}_2 - \hat{\alpha}_1 = 73.0$ with a 95% CI of (41, 105). The CI is now *smaller* than when calculated with a model that includes an interaction term!! This is good news as it provides a tighter range of plausible values and thus more powerful inference.