STA221

Neil Montgomery

Last edited: 2017-03-15 15:07

experiments with two factors

fire retardant example

There aren't many datasets that come with Chapter 26, so I found another to use as a motivating example.

In this dataset a new fire retardant treatment of cotton fabric is being tested. The name of this factor is `Bath`. There are two levels names `I` and `II`, which stand for "received the treatment" and "did not receive the new treatment" respectively.

(It's called "Bath" because the fabric is bathed in the treatment solution.)

fire retardant example

There aren't many datasets that come with Chapter 26, so I found another to use as a motivating example.

In this dataset a new fire retardant treatment of cotton fabric is being tested. The name of this factor is `Bath`. There are two levels names `I` and `II`, which stand for "received the treatment" and "did not receive the new treatment" respectively.

(It's called "Bath" because the fabric is bathed in the treatment solution.)

The amount of time it takes each cloth sample to start to burn is recorded.

Here is a numerical summary of the results:

Bath	n	Means	SD
I	24	16.875	5.921
II	24	9.154	5.670

but there is also another variable

The new treatment works. But there is also the matter of the efficiency with which the treatment can be applied.

There is another factor variable in this dataset: the number of “launderings”, which is the way a retardant treatment is applied.

This variable has two levels named 5 and 10, corresponding to the actual number of launderings.

Here is a summary of the results with respect to this variable:

Launderings	n	Means	SD
10	24	15.067	5.609
5	24	10.963	7.620

the effect of both variables

The new treatment is better than the old treatment (all else being equal). More laundering is better than less laundering (all else being equal).

But what about the combination of both variables? There are a few different ways they might combine to affect the results. We'll be concerned with these possibilities:

the effect of both variables

The new treatment is better than the old treatment (all else being equal). More laundering is better than less laundering (all else being equal).

But what about the combination of both variables? There are a few different ways they might combine to affect the results. We'll be concerned with these possibilities:

Their effects could simply be *additive*.

the effect of both variables

The new treatment is better than the old treatment (all else being equal). More laundering is better than less laundering (all else being equal).

But what about the combination of both variables? There are a few different ways they might combine to affect the results. We'll be concerned with these possibilities:

Their effects could simply be *additive*.

The effect of one factor could depend on the level of the other factor. In this case we say there is an *interaction* between the two factor variables.

the effect of both variables

The new treatment is better than the old treatment (all else being equal). More laundering is better than less laundering (all else being equal).

But what about the combination of both variables? There are a few different ways they might combine to affect the results. We'll be concerned with these possibilities:

Their effects could simply be *additive*.

The effect of one factor could depend on the level of the other factor. In this case we say there is an *interaction* between the two factor variables.

A good graphical method to evaluate the relationship is called an *interaction plot*.

the effect of both variables

The new treatment is better than the old treatment (all else being equal). More laundering is better than less laundering (all else being equal).

But what about the combination of both variables? There are a few different ways they might combine to affect the results. We'll be concerned with these possibilities:

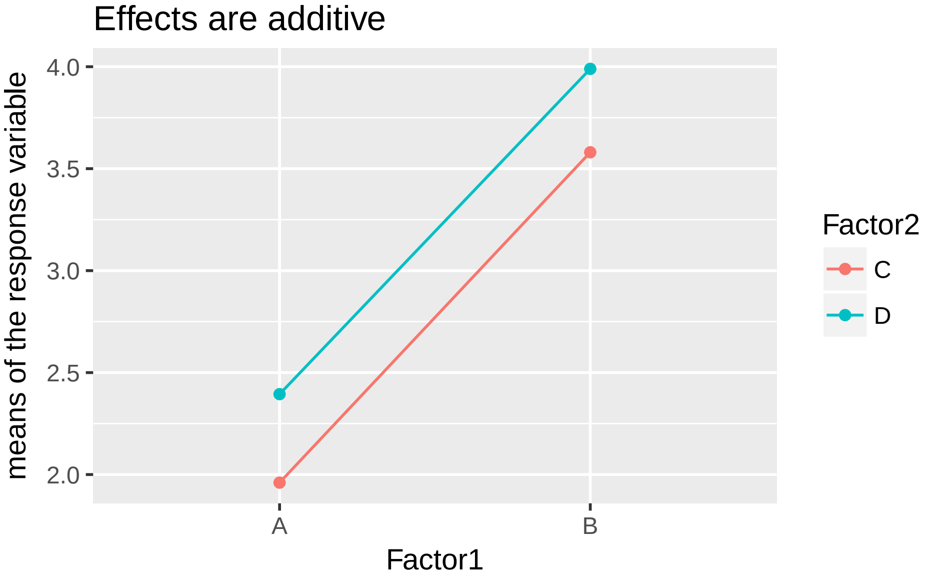
Their effects could simply be *additive*.

The effect of one factor could depend on the level of the other factor. In this case we say there is an *interaction* between the two factor variables.

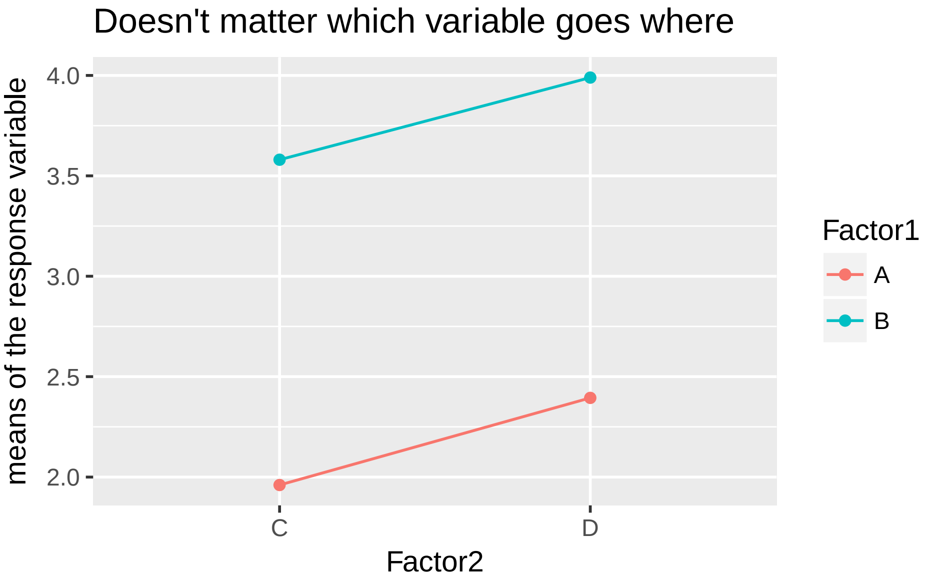
A good graphical method to evaluate the relationship is called an *interaction plot*.

To illustrate, I have simulated a dataset with two variables. `Factor1` has levels `A` and `B` while `Factor2` has levels `C` and `D`.

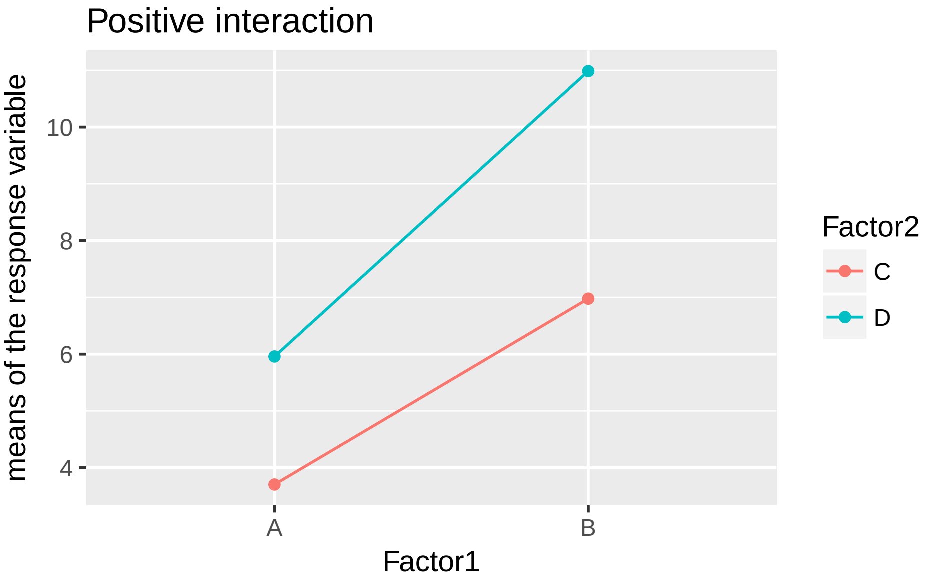
interaction plot example 1a - additive



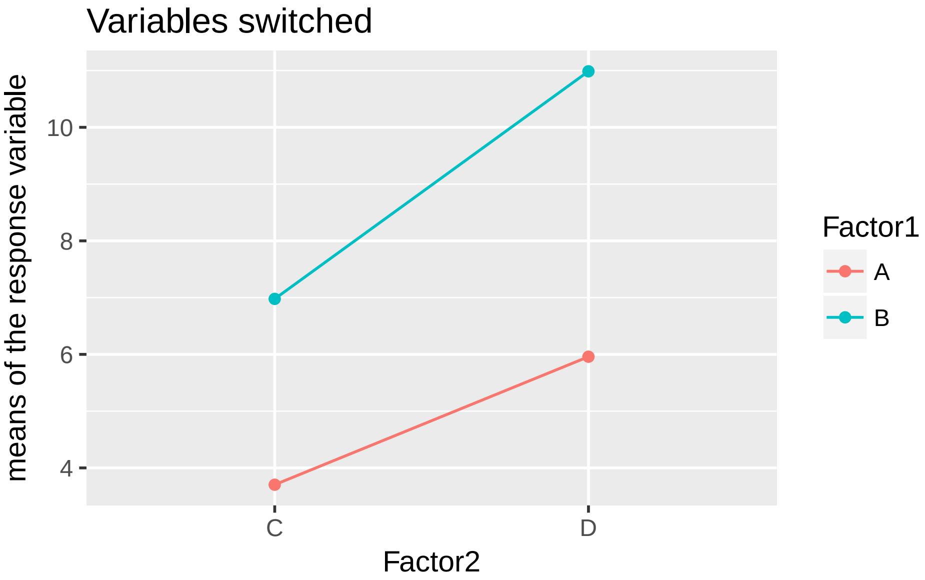
interaction plot example 1b - additive



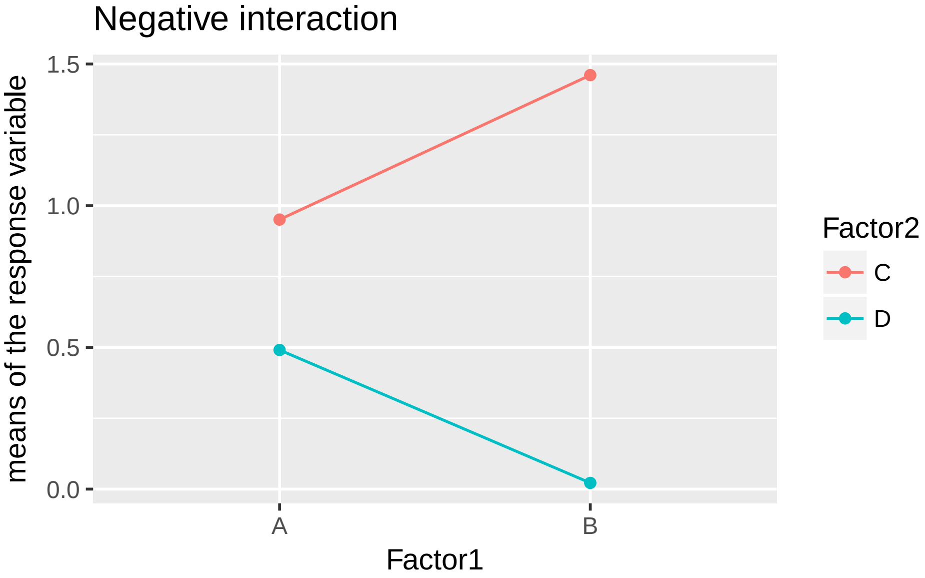
interaction plot example 2a - positive interaction



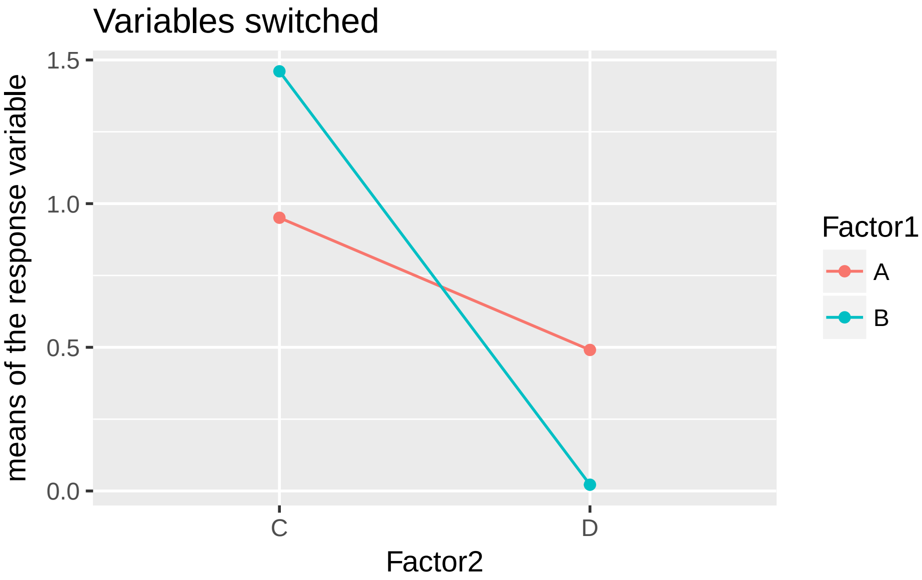
interaction plot example 2b - positive interaction



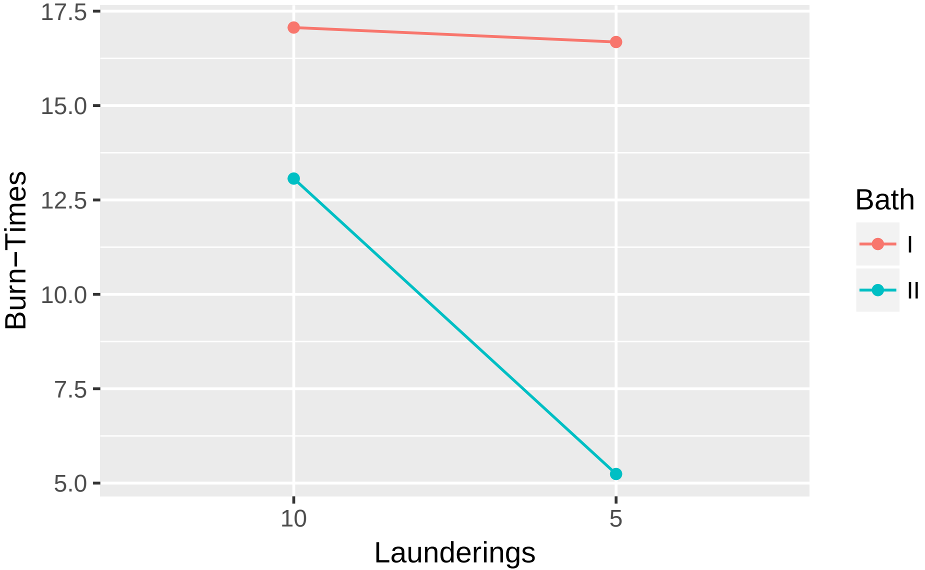
interaction plot example 3a - negative interaction



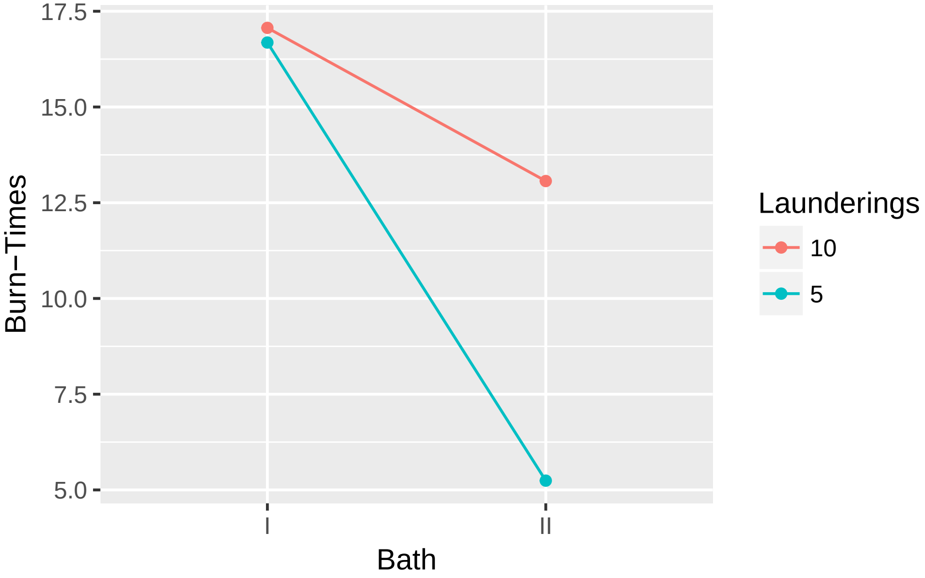
interaction plot example 3a - negative interaction



interaction plot of the fire retardant data



fire data - variables switched



tentative conclusion from fire data | new models to consider

It seems that the new treatment is just as good, even after only 5 launderings, so it is also more efficient.

tentative conclusion from fire data | new models to consider

It seems that the new treatment is just as good, even after only 5 launderings, so it is also more efficient.

Here are the models we will consider:

$$y_{ijk} = \mu + \tau_i + \gamma_j + \varepsilon_{ijk} \qquad \text{no interaction assumed}$$

$$y_{ijk} = \mu + \tau_i + \gamma_j + (\tau\gamma)_{ij} + \varepsilon_{ijk} \qquad \text{general model}$$

tentative conclusion from fire data | new models to consider

It seems that the new treatment is just as good, even after only 5 launderings, so it is also more efficient.

Here are the models we will consider:

$$y_{ijk} = \mu + \tau_i + \gamma_j + \varepsilon_{ijk} \qquad \text{no interaction assumed}$$

$$y_{ijk} = \mu + \tau_i + \gamma_j + (\tau\gamma)_{ij} + \varepsilon_{ijk} \qquad \text{general model}$$

μ is the grand, overall average.

tentative conclusion from fire data | new models to consider

It seems that the new treatment is just as good, even after only 5 launderings, so it is also more efficient.

Here are the models we will consider:

$$y_{ijk} = \mu + \tau_i + \gamma_j + \varepsilon_{ijk} \qquad \text{no interaction assumed}$$

$$y_{ijk} = \mu + \tau_i + \gamma_j + (\tau\gamma)_{ij} + \varepsilon_{ijk} \qquad \text{general model}$$

μ is the grand, overall average.

The τ_i are the effects of the levels of the first factor.

tentative conclusion from fire data | new models to consider

It seems that the new treatment is just as good, even after only 5 launderings, so it is also more efficient.

Here are the models we will consider:

$$y_{ijk} = \mu + \tau_i + \gamma_j + \varepsilon_{ijk} \qquad \text{no interaction assumed}$$

$$y_{ijk} = \mu + \tau_i + \gamma_j + (\tau\gamma)_{ij} + \varepsilon_{ijk} \qquad \text{general model}$$

μ is the grand, overall average.

The τ_i are the effects of the levels of the first factor.

The γ_j are the effects of the levels of the second factor.

tentative conclusion from fire data | new models to consider

It seems that the new treatment is just as good, even after only 5 launderings, so it is also more efficient.

Here are the models we will consider:

$$y_{ijk} = \mu + \tau_i + \gamma_j + \varepsilon_{ijk} \qquad \text{no interaction assumed}$$

$$y_{ijk} = \mu + \tau_i + \gamma_j + (\tau\gamma)_{ij} + \varepsilon_{ijk} \qquad \text{general model}$$

μ is the grand, overall average.

The τ_i are the effects of the levels of the first factor.

The γ_j are the effects of the levels of the second factor.

The $(\tau\gamma)_{ij}$ are the effects of all combinations of the levels of hte two factors.

tentative conclusion from fire data | new models to consider

It seems that the new treatment is just as good, even after only 5 launderings, so it is also more efficient.

Here are the models we will consider:

$$y_{ijk} = \mu + \tau_i + \gamma_j + \varepsilon_{ijk} \qquad \text{no interaction assumed}$$

$$y_{ijk} = \mu + \tau_i + \gamma_j + (\tau\gamma)_{ij} + \varepsilon_{ijk} \qquad \text{general model}$$

μ is the grand, overall average.

The τ_i are the effects of the levels of the first factor.

The γ_j are the effects of the levels of the second factor.

The $(\tau\gamma)_{ij}$ are the effects of all combinations of the levels of hte two factors.

ε_{ijk} is random noise, assumed to be $N(0, \sigma)$.

the analysis

We have new sums of squares decompositions. The details can get brutal - so we'll stay symbolic.

We'll call the factor variables `A` and `B`, with I and J levels respectively.

In the additive case we get, assuming a common sample size of n for each combination of factor levels (strongly encouraged in practice!):

$$SS_{Total} = SS_A + SS_B + SS_{Error}$$

the analysis

We have new sums of squares decompositions. The details can get brutal - so we'll stay symbolic.

We'll call the factor variables `A` and `B`, with I and J levels respectively.

In the additive case we get, assuming a common sample size of n for each combination of factor levels (strongly encouraged in practice!):

$$SS_{Total} = SS_A + SS_B + SS_{Error}$$

Everything has χ^2 distributions. The degrees of freedom add up (N is the grand sample size):

$$N - 1 = (I - 1) + (J - 1) + (N - I - J + 1)$$

the analysis

We have new sums of squares decompositions. The details can get brutal - so we'll stay symbolic.

We'll call the factor variables `A` and `B`, with I and J levels respectively.

In the additive case we get, assuming a common sample size of n for each combination of factor levels (strongly encouraged in practice!):

$$SS_{Total} = SS_A + SS_B + SS_{Error}$$

Everything has χ^2 distributions. The degrees of freedom add up (N is the grand sample size):

$$N - 1 = (I - 1) + (J - 1) + (N - I - J + 1)$$

When the experiment is *balanced* (strongly encouraged!) with n in each combination of levels, the error degrees of freedom simplifies to $IJ(n - 1)$.

fire example - no interaction (?!)

```
##             Df Sum Sq Mean Sq F value    Pr(>F)
## Bath         1  715.3   715.3  23.954 0.0000131
## Launderings  1  202.1   202.1   6.769    0.0125
## Residuals   45 1343.8    29.9
```