Linear Models: Interactions

BIOE 498/598 PJ

Spring 2021

What is an interaction?

Imagine we're modeling the response (y) from two input variables, x_1 and x_2. The simplest model is

$$y = \beta_1 x_1 + \beta_2 x_2 + \epsilon$$

What is an interaction?

Imagine we're modeling the response (y) from two input variables, x_1 and x_2. The simplest model is

$$y = \beta_1 x_1 + \beta_2 x_2 + \epsilon$$

The coefficient β_1 measures the effect of x_1 and β_2 measures the effect of x_2. These effects are **independent**.

What is an interaction?

Imagine we're modeling the response (y) from two input variables, x_1 and x_2. The simplest model is

$$y = \beta_1 x_1 + \beta_2 x_2 + \epsilon$$

The coefficient β_1 measures the effect of x_1 and β_2 measures the effect of x_2. These effects are **independent**.

What is there is another effect that depends on both x_1 and x_2? This is an **interaction** between x_1 and x_2.

How do we model interactions?

We model the interaction of x_1 and x_2 using the product of these variables.

$$y = \beta_1 x_1 + \beta_2 x_2 + \beta_{12} x_1 x_2 + \epsilon$$

The coefficient β_{12} is the effect size of the interaction.

How do we model interactions?

We model the interaction of x_1 and x_2 using the product of these variables.

$$y = \beta_1 x_1 + \beta_2 x_2 + \beta_{12} x_1 x_2 + \epsilon$$

The coefficient β_{12} is the effect size of the interaction.

Why do we multiply x_1 and x_2? There are at least two ways to interpret this term.

The coded factor interpretation

Often we set up design matrices using **coded variables**. If we're testing the variable at two levels, we code the variable as "on/off" ($\{0, 1\}$) or "low/high" ($\{-1, +1\}$).

The coded factor interpretation

Often we set up design matrices using **coded variables**. If we're testing the variable at two levels, we code the variable as "on/off" ($\{0, 1\}$) or "low/high" ($\{-1, +1\}$).

on/off $\rightarrow$ interaction when both "on"

x_1	x_2	$x_1 x_2$
0	0	0
0	1	0
1	0	0
1	1	1

The coded factor interpretation

Often we set up design matrices using **coded variables**. If we're testing the variable at two levels, we code the variable as "on/off" ($\{0, 1\}$) or "low/high" ($\{-1, +1\}$).

on/off $\rightarrow$ interaction when both "on"

x_1	x_2	$x_1 x_2$
0	0	0
0	1	0
1	0	0
1	1	1

high/low $\rightarrow$ interaction when both "high" or both "low"

x_1	x_2	$x_1 x_2$
-1	-1	$+1$
-1	$+1$	-1
$+1$	-1	-1
$+1$	$+1$	$+1$

The augmented slope interpretation

We can also interpret the interaction as one variable changing the effect of the other variable.

$$
\begin{aligned}
y &= \beta_1 x_1 + \beta_2(x_1) x_2 + \epsilon \\
&= \beta_1 x_1 + (\beta_2 + \beta_{12} x_1) x_2 + \epsilon \\
&= \beta_1 x_1 + \beta_2 x_2 + \beta_{12} x_1 x_2 + \epsilon
\end{aligned}
$$

Interactions with `lm`

Recall the data frame from out blood pressure clinical trial:

```
## # A tibble: 6 x 3
##   BPchange treated male
##      <dbl> <lgl>   <lgl>
## 1   -0.525 TRUE    FALSE
## 2    4.17  TRUE    FALSE
## 3    6.03  TRUE    TRUE
## 4   -1.40  TRUE    FALSE
## 5    0.493 TRUE    FALSE
## 6   12.9   FALSE   TRUE
```

Adding an interaction term to our model

```
##
## Call:
## lm(formula = BPchange ~ treated + male + treated:male, data =
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -20.465  -4.407   2.309   5.887  18.738
##
## Coefficients:
##                       Estimate Std. Error t value Pr(>|t|)
## (Intercept)             12.319      7.895   1.560    0.145
## treatedTRUE            -10.090      9.116  -1.107    0.290
## maleTRUE                14.898      9.116   1.634    0.128
## treatedTRUE:maleTRUE     1.049     12.893   0.081    0.936
##
## Residual standard error: 11.17 on 12 degrees of freedom
## Multiple R-squared:  0.5606, Adjusted R-squared:  0.4507
## F-statistic: 5.102 on 3 and 12 DF,  p-value: 0.01666
```

A shortcut for adding interactions and main effects

```
##
## Call:
## lm(formula = BPchange ~ treated * male, data = bp_data)
##
## Residuals:
##     Min      1Q  Median      3Q     Max
## -20.465  -4.407   2.309   5.887  18.738
##
## Coefficients:
##                       Estimate Std. Error t value Pr(>|t|)
## (Intercept)             12.319      7.895   1.560    0.145
## treatedTRUE            -10.090      9.116  -1.107    0.290
## maleTRUE                14.898      9.116   1.634    0.128
## treatedTRUE:maleTRUE     1.049     12.893   0.081    0.936
##
## Residual standard error: 11.17 on 12 degrees of freedom
## Multiple R-squared:  0.5606, Adjusted R-squared:  0.4507
## F-statistic: 5.102 on 3 and 12 DF,  p-value: 0.01666
```

A shortcut for adding interactions and main effects

```
##
## Call:
## lm(formula = y ~ x1 * x2 * x3)
##
## Residuals:
##         1         2         3         4         5         6
## -0.053712  0.130828  0.062348 -0.331228  0.004367  0.059872
##         8         9        10
##  0.184778 -0.123954  0.023741
##
## Coefficients:
##             Estimate Std. Error t value Pr(>|t|)
## (Intercept)   -11.02       9.52  -1.158    0.367
## x1             12.04      10.54   1.142    0.372
## x2             17.03      17.79   0.957    0.440
## x3             20.78      21.16   0.982    0.430
## x1:x2         -17.16      20.71  -0.828    0.495
## x1:x3         -20.57      23.86  -0.862    0.479
## x2:x3         -31.54      36.32  -0.868    0.477
## x1:x2:x3       30.02      42.31   0.710    0.551
##
```

How many interactions are there?

term	x_1	x_2	x_3
β_0	0	0	0
$\beta_1 x_1$	1	0	0
$\beta_2 x_2$	0	1	0
$\beta_3 x_3$	0	0	1
$\beta_{12} x_1$	1	1	0
$\beta_{13} x_2$	1	0	1
$\beta_{23} x_2 x_3$	0	1	1
$\beta_{123} x_1 x_2 x_3$	1	1	1

How many interactions are there?

term	x_1	x_2	x_3
β_0	0	0	0
$\beta_1 x_1$	1	0	0
$\beta_2 x_2$	0	1	0
$\beta_3 x_3$	0	0	1
$\beta_{12} x_1$	1	1	0
$\beta_{13} x_2$	1	0	1
$\beta_{23} x_2 x_3$	0	1	1
$\beta_{123} x_1 x_2 x_3$	1	1	1

A model with n factors has 2^n possible terms; $2^n - n - 1$ of these are interactions.

Hierarchical ordering to the rescue

Hierarchical Ordering principle

- Lower order effects are more likely to be important than higher order effects.
- Effects of the same order are equally likely to be important.

How many interactions are there?

n	intercept	main effects	TWI	higher-order
1	1	1	0	0
2	1	2	1	0
3	1	3	3	1
4	1	4	6	5
5	1	5	10	16
6	1	6	15	42
7	1	7	21	99
8	1	8	28	219
9	1	9	36	466
10	1	10	45	968

How many interactions are there?

n	intercept	main effects	TWI	higher-order
1	1	1	0	0
2	1	2	1	0
3	1	3	3	1
4	1	4	6	5
5	1	5	10	16
6	1	6	15	42
7	1	7	21	99
8	1	8	28	219
9	1	9	36	466
10	1	10	45	968

We will design experiments that focus on main effects and two-way interactions.

Hierarchical ordering to the rescue

Hierarchical Ordering Principle

- Lower-order effects are more likely to be important than higher-order effects.
- Effects of the same order are equally likely to be important.

If we neglect an important higher-order term, the effects can appear anywhere in our model!

We can design the experiment to constrain where higher-order effects appear.

Things to remember about interactions

- Interaction are modeled as the product of variables.
- The interaction effect is “above and beyond" the independent effects (synergy/super-additivity, antagonism/sub-additivity).
- Higher-order interactions are possible (e.g. $x_1 x_2 x_3$), but these are rare.
- Proper experiment design is needed when "ignoring" higher-order interactions.