Plotting Non-linear & Complex Main Effects Models

- · Plotting single-predictor non-linear models
- Plotting 2-predictor non-linear models 2xQ, kxQ & QxQ
- · Plotting complex non-linear main effects models

Models with a single centered quantitative predictor

 $X_{cen} = X - X_{mean}$ y' = bX_{cen} + a

- $a \rightarrow$ regression constant
 - expected value of y when x = 0 (re-centered to mean=0)
 - height of y-x regression line
- $b \rightarrow$ regression weight
 - expected direction and extent of change in y for a 1-unit increase in x
 - slope of y-x regression line

X will be X_{cen} in all of the following models



$$\begin{split} X_{cen} &= X - X_{mean} \qquad X_{cen}{}^2 &= (X - X_{mean})^2 \\ y' &= b_1 X_{cen} + b_2 X_{cen}{}^2 + a \end{split}$$





Three kinds (for now, more later) ...

- centered quant variable & dummy-coded binary variable
- · centered quant variable & dummy-coded k-category variable
- 2 centered quant variables

What we're trying to do is to plot these models so that we can see how both of the predictors are related to the criterion.

Like when we're plotting data from a factorial design, we have to represent 3 variables -- the criterion & the 2 predictors X & Z -- in a 2-dimensional plot. We'll use the same solution..

We'll plot the relationship between one predictor and the criterion for different values of the other predictor

Non-linear Models with a centered quantitative predictor & a dummy coded binary predictor

 $y' = b_1 X + b_2 X^2 + b_3 Z + a$

This is called a main effects model \rightarrow there are no interaction terms.

a → regression constant

- expected value of y if X=0 (mean) and Z=0 (comparison group)
- · mean of the control group
- · height of control group quant-criterion regression line
- $b_1 \rightarrow$ regression weight for centered quant predictor linear main effect of X
 - expected direction and extent of change in y for a 1-unit increase in x , after controlling for the other variable(s) in the model
 - expected direction and extent of change in y for a 1-unit increase in x , for the comparison group (coded 0)
 - slope of quant-criterion regression line for the group coded 0 (comp)

Non-linear Models with a centered quantitative predictor & a dummy coded binary predictor

$$y' = b_1 X + b_2 X^2 + b_3 Z + a$$

This is called a main effects model \rightarrow there are no interaction terms.

- $b_2 \rightarrow$ weight for centered & squared quant predictor non-linear main effect of X
 - expected direction and extent of change in y-x slope for a 1-unit increase in x, after controlling for the other variable(s) in the model
 - expected direction and extent of change in y-x slope for a 1-unit increase in x, for the comparison group (coded 0)
 - curve of quant-criterion regression line for the group coded 0 (comp)

 $b_3 \rightarrow$ regression weight for dummy coded binary predictor – main effect of Z

- expected direction and extent of change in y for a 1-unit increase in x, after controlling for the other variable(s) in the model
- direction and extent of y mean difference between groups coded 0 & 1, after controlling for the other variable(s) in the model
- group mean/reg line height difference (when X = 0, the centered mean)

To plot the model we need to get separate regression formulas for each Z group. We start with the multiple regression model...

V' =

Model \rightarrow

$$b_1X + b_2X^2 + b_3Z + a$$

For the Comparison Group coded Z = 0

Substitute the 0 in for Z Simplify the formula

$$y' = b_1 X + b_2 X^2 + b_3 0 + a$$

$$y' = b_1 X + b_2 X^2 + a$$

slope curve height

For the Target Group coded Z = 1

Substitute the 1 in for Z Simplify the formula





Non-linear models with a centered quantitative predictor & a dummy coded k-category predictor

| $y' = b_1 X + b_2 X^2 + b_3 Z_1 + b_4 Z_2 + a_3$ | This is called a main effects model → there are no interaction terms. | $y' = b_1 X + b_2 X^2 + b_3 Z_1 + b_4 Z_2 + a_1$ | This is called a main effects model \rightarrow there are no interaction terms. |
|---|--|---|--|
| a → regression constant expected value of y if Z₁ & Z₂ = 0 (the control group mean of the control group height of control group quant-criterion regressio | $\begin{array}{cccc} & \text{Group} & Z_1 & Z_2 \\ \text{Dup}) & 1 & 1 & 0 \\ 1 & 2 & 0 & 1 \\ n \text{ line} & 3^* & 0 & 0 \end{array}$ | | $\begin{array}{ccccc} \text{Group} & \text{Z}_1 & \text{Z}_2 \\ 1 & 1 & 0 \\ 2 & 0 & 1 \\ 3^* & 0 & 0 \end{array}$ |
| b₁ → regression weight for centered quant predictor – main effect of X expected direction and extent of change in y for a 1-unit increase in x slope of quant-criterion regression line (for both groups) | | b₃ → regression weight for dummy coded comparison of G1 vs G3 – main effect expected direction and extent of change in y for a 1-unit increase in x direction and extent of y mean difference between groups 1 & 3 group height difference between comparison & target groups (3 & 1) (X = 0) | |
| $b_2 \rightarrow$ weight for centered & squared quant predictor – r | non-linear main effect of X | | |
| expected direction and extent of change in y-x slope for a 1-unit increase in x, after controlling for the other variable(s) in the model | | $b_4 \rightarrow$ regression weight for dummy coded comparison of G2 vs. G3 – main effect | |
| expected direction and extent of change in y-x slope for a 1-unit increase in x , for the comparison group (coded 0) curve of quant-criterion regression line for the group coded 0 (comp) | | direction and extent of y mean difference between groups coded 0 & 1 group height difference between comparison & target groups (3 & 2) (X = 0) | |
| | | | |
| | | | |

Non-linear models with a centered quantitative predictor &

a dummy coded k-category predictor

To plot the model we need to get separate regression formulas for each Z group. We start with the multiple regression model...

| Model \rightarrow y' = b ₁ X + b ₂ X ² | $+ b_3 Z_1 + b_4 Z_2 + a$ | $\begin{array}{cccc} \text{Group} & \text{Z}_1 & \text{Z}_2 \\ 1 & 1 & 0 \\ 2 & 0 & 1 \\ 3^* & 0 & 0 \end{array}$ | | |
|---|--|---|--|--|
| For the Comparison Group - coded $Z_1 = 0 \& Z_2 = 0$ | | | | |
| Substitute the Z code values Simplify the formula | y' = $b_1X + b_2X^2 + b_1X + b_2X^2 + a_2X^2 + $ | $b_2^{*0} + b_3^{*0} + a$ a ve height | | |
| For the Target Group 1 - coded $Z_1 = 1 \& Z_2 = 0$ | | | | |
| Substitute the Z code values Simplify the formula | $y' = b1X + b_2X^2 + 1$ $y' = b_1X + b_2X^2 + 1$ slope \checkmark curve | $b_2^*1 + b_3^*0 + a$ ($b_2 + a$) | | |
| For the Target Group 2- coded $Z_1 = 0 \& Z_2 = 1$ | | | | |
| Substitute the Z code values Simplify the formula | $y' = b_1X + b_2X^2 + b_2$ $y' = b_1X + b_2X^2 + (b_3)$ | *0 + b ₃ *1 + a ₃ + a) | | |
| | slope / curve | height | | |



weights for non-linear main effects models w/ centered quantitative predictor &

a dummy coded binary predictor **Of** a dummy coded k-category predictor

Constant "a"

- the expected value of y when the value of all predictors = 0
- height of the Y-X regression line for the comparison group

b for a centered quantitative variable - main effect

- the direction and extent of the expected change in the value of y for a 1-unit increase in that predictor, holding the value of all other predictors constant at 0
- slope of the Y-X regression line for the comparison group

b for a centered & squared quantitative variable - main effect

- the direction and extent of the expected change in the y-x for a 1-unit increase in x, holding the value of all other predictors constant at 0
- curve of the Y-X regression line for the comparison group

b for a dummy coded binary variable -- main effect

- the direction and extent of expected mean difference of the Target group from the Comparison group, holding the value of all other predictors constant at 0
- difference in height of the Y-X regression lines for the comparison & target groups
 comparison & target groups Y-X regression lines have same slope no interaction

b for a dummy coded k-group variable – main effect

· direction and extent of expected mean difference of Target group for that dummy code from Comparison group, holding the value of all other predictors constant at 0
difference in height of the Y-X regression lines for the comparison & target groups
comparison & target groups Y-X regression lines have same slope – no interaction and the same slope – no interaction and th

Non-linear Models with 2 centered quantitative predictors

$$y' = b_1 X_1 + b_2 X_1^2 + b_3 X_2 + b_4 X_2^2 + a$$

This is called a main effects model \rightarrow there are no interaction terms.

- $a \rightarrow$ regression constant
 - expected value of y if X₁=0 (mean) X₂=0 (mean)
 - mean criterion score of those having $X_1=0$ (mean) $X_2=0$ (mean)
 - height of quant-criterion regression line for those with $X_2=0$ (mean)
- $b_1 \rightarrow$ regression weight for X₁ centered quant predictor
 - expected direction and extent of change in y for a 1-unit increase in X₁ after controlling for the other variable(s) in the model
 - expected direction and extent of change in y for a 1-unit increase in X_2 , for those with $X_2=0$ (mean)
 - slope of quant-criterion regression line when X₂=0 (mean)
- $b_2 \rightarrow$ weight for centered & squared quant predictor non-linear main effect of X_1
 - expected direction and extent of change in y-x slope for a 1-unit increase in X₁, after controlling for the other variable(s) in the model
 - expected direction and extent of change in y-x slope for a 1-unit increase in X1, for those with $X_2=0$ (mean)
 - curve of quant-criterion regression line when $X_2 = 0$ (mean)

Non-linear models with 2 centered quantitative predictors

$$y' = b_1X_1 + b_2X_1^2 + b_3X_2 + b_4X_2^2 + a$$

This is called a main effects model \rightarrow there are no interaction terms.

Same idea ...

 we want to plot these models so that we can see how both of the predictors are related to the criterion

Different approach ...

- when the second predictor was binary or had k-categories, we plotted the Y-X regression line for each Z group
- now, however, we don't have any groups both the X₁ & X₂ variables are centered quantitative variables
- what we'll do is to plot the Y-X₁ regression line for different values of X_2
- the most common approach is to plot the Y-X₁ regression line for...
 - the mean of X_2
 - +1 std above the mean of X_2
 - -1 std below the mean of X2

Ye'll plot 3 lines

Non-linear Models with 2 centered quantitative predictors

 $y' = b_1 X_1 + b_2 X_1^2 + b_3 X_2 + b_4 X_2^2 + a$

This is called a main effects model \rightarrow there are no interaction terms.

- $b_3 \rightarrow$ regression weight for X₂ centered quant predictor
 - expected direction and extent of change in y for a 1-unit increase in X₂, after controlling for the other variable(s) in the model
 - expected direction and extent of change in y for a 1-unit increase in X_2 , for those with $X_1=0$ (mean)
 - expected direction and extent of change of the height of Y-X regression line for a 1-unit change in X_2 when $X_1{=}0$ (mean)
- $b_4 \rightarrow$ weight for centered & squared quant predictor non-linear main effect of X_2
 - expected direction and extent of change in y-x slope for a 1-unit increase in X_2 , after controlling for the other variable(s) in the model
 - expected direction and extent of change in y-x slope for a 1-unit increase in $X_2,$ for those with $X_1{=}0~(mean)$
 - curve of quant-criterion regression line when $X_1 = 0$ (mean)

To plot the model we need to get separate regression formulas for each chosen value of Z. Start with the multiple regression model..

$$y' = b_1X_1 + b_2X_1^2 + b_3X_2 + b_4X_2^2 + a_4X_2^2$$

For X₂ = 0 (the mean of centered X₂) Substitute the 0 in for X₂ $y' = b_1X_1 + b_2X_1^2 + b_3 0 + b_4 0 + a$ Simplify the formula For X₂ = +1 std Substitute the std for X₂ $y' = b_1X + b_2X_1^2 + a_2$ height Substitute the std for X₂ $y' = b_1X + b_2X_1^2 + b_2$ *std + b_4 std² + a Simplify the formula For X₂ = -1 std Substitute the std for X₂ $y' = b_1X + b_2X_1^2 - b_2$ *std - b_4 std² + a Simplify the formula Substitute the std for X₂ $y' = b_1X + b_2X_1^2 - b_2$ *std - b_4 std² + a Substitute the std for X₂ $y' = b_1X + b_2X_1^2 - b_2$ *std - b_4 std² + a Simplify the formula Substitute the std for X₂ $y' = b_1X + b_2X_1^2 - b_2$ *std - b_4 std² + a Simplify the formula Substitute the std for X₂ $y' = b_1X + b_2X_1^2 - b_2$ *std - b_4 std² + a Simplify the formula Substitute the std for X₂ $y' = b_1X + b_2X_1^2 + (-b_2$ *std - b_4 std² + a) Substitute the std for X₂ $y' = b_1X + b_2X_1^2 + (-b_2$ *std - b_4 std² + a) Substitute the std for X₂ $y' = b_1X + b_2X_1^2 + (-b_2$ *std - b_4 std² + a) Substitute the std for X₂ $y' = b_1X + b_2X_1^2 + (-b_2$ *std - b_4 std² + a) Substitute the std for X₂ $y' = b_1X + b_2X_1^2 + (-b_2$ *std - b_4 std² + a) Substitute the std for X₂ $y' = b_1X + b_2X_1^2 + (-b_2$ *std - b_4 std² + a)

It can be a challenge to "see" the shape of the variable represented with the different lines for different values...



X1





| Plotting Complex Linear Main Effects Models | So, we have this model | |
|---|--|--|
| Sometimes we have a model with more than 2 predictors, and | pert' = $b_1 age + b_2 mar1 + b_3 mar2 + b_4 era + b_5 exp + a$ | |
| sometimes we want to plot these more complex models. | age & exp(erience) are centered quantitative variables era is dummy coded 0 – 1980s 1 – 2010s | |
| There is only so much that can be put into a single plot and | mar(ital status) is dummy coded | |
| | mar1 single = 1 mar2 divorced = 1 married = 0 | |
| Remember, with main effects models (with no interactions), all | To plot this we have to decide what we want to show, say | |
| | How are experience & marital status related to performance for 25-year-olds | |
| So, our plots are a series of lines, each for a different | | |
| combination of variable groups/values. | What we do is fill in the values of the "selection variables" and simplify the | |
| The most common version of these plots is to show how 2 | tormula to end up with a plotting function for each regression line | |
| variables relate to the criterion, for a given set of values for the other variables (usually their mean) | For this plot: | |
| | performance (the criterion) is on the Y axis experience (a quantitative predictor) is on the X axis | |
| Here's an example! | we'll have 3 regression lines, one each for single, married & divorced | |
| | | |
| | | |
| | | |
| How are experience & marital status related to perf for 25-year-olds from 2010s? | | |
| perf' = $b_1 aqe + b_2 mar1 + b_2 mar2 + b_4 era + b_e exp + a$ | | |
| • era $0 = 1980s$ $1 = 2010s$ | | |
| • mar mar1 single = 1 mar2 divorced = 1 married = 0 | | |
| | | |
| For married \rightarrow perf' = $b_1^*25 + b_2^*0 + b_3^*0 + b_4^*1 + b_5 exp + a$ | | |
| we plot $perf' = b_5 exp + (b_4^{*1} + b_1^{*25} + a)$ height | | |
| `slope | | |
| For singles \rightarrow perf' = $b_1^* 25 + b_2^* 1 + b_3^* 0 + b_4^* 1 + b_5 exp + a$ | | |
| we plot perf' = $b_5 exp + (b_2 + b_4^*1 + b_1^*25 + a)$ height | | |
| slope | | |
| For divorged -> perf' = b *25 + b *0 + b *1 + b *1 + b even + a | | |
| For alvorced \rightarrow peri = $D_1^{-1}23 + D_2^{-1}0 + D_3^{-1}1 + D_4^{-1}1 + D_5^{-1}exp + a$ | | |
| we plot $perf = b_5 exp + (b_3 + b_4 + b_1 + b_1 + 25 + a)$ height | | |
| | | |
| | | |

How are experience & marital status related to perf for 25-year-olds from 2010s?

perf' = $b_1age + b_2mar1 + b_3mar2 + b_4era + b_5exp + a$

• era 0 = 1980s 1 = 2010s • mar mar1 single = 1 mar2 divorced = 1 married = 0

Let's look at the plotting formulas again

For married \rightarrow perf' = b₅exp + (b₄*1 + b₁*25 + a)

For singles \rightarrow perf' = b₅exp + (b₂ + b₄*1 + b₁*25 + a)

For divorced \rightarrow perf' = b₅exp + (b₃ + b₄*1 + b₁*25 + a)

Notice that each group's regression line has the same slope (b_5) , because there is no interaction.

The difference between the group's regression lines are their height – which differ based on the relationship between marital status and perf !!!

How are experience & marital status related to perf for 25-year-olds from 2010s? perf' = $b_125 + b_2 mar1 + b_3 mar2 + b_4 2010s + b_5 exp + a$



The selection of the value of Age (b_1) & Era (b_4) determine the height of the set of lines!

Plotting Complex Non-Linear Main Effects Models

Sometimes we have a non-linear model with more than 2 predictors, and sometimes we want to plot these more complex models.

There is only so much that can be put into a single plot and reasonably hope the reader will be able to understand it.

Remember, with main effects models (with no interactions), all lines are parallel!!!

So, our plots are a series of lines, each for a different combination of variable groups/values.

The most common version of these plots is to show how 2 variables relate to the criterion, for a given set of values for the other variables (usually their mean)

Here's an example!

So, we have this model... How are experience & marital status related to perf for 25-year-olds from 2010s? perf' = $b_1age + b_2mar1 + b_3mar2 + b_4era + b_5exp + a$ perf' = $b_1age + b_2age^2 + b_3mar1 + b_4mar2 + b_5era + b_6exp + b_7exp^2 + a$ • age & exp(erience) are centered quantitative variables • era 0 = 1980s 1 = 2010s • era is dummy coded 0 = 1980s 1 = 2010s • mar mar1 single = 1 mar2 divorced = 1 married = 0 mar(ital status) is dummy coded mar1 single = 1mar2 divorced = 1 married = 0perf' = $b_1 25 + b_2 25^2 + b_3 0 + b_4 0 + b_5 1 + b_6 exp + b_7 exp^2 + a$ For married \rightarrow perf' = $b_6 exp + b_7 exp^2 + (b_5^{*1} + b_1^{*25} + b_2^{25^2} + a)$ we plot To plot this we have to decide what we want to show, say... Curve slope 1 How are experience & marital status related to performance for 25-year-olds 🔶 height from 1980s? For singles \rightarrow perf' = $b_1 25 + b_2 25^2 + b_3 1 + b_4 0 + b_5 1 + b_6 exp + b_7 exp^2 + a$ All we do is fill in the values of the "selection variables" and simplify the perf' = $b_{e}exp + b_{7}exp^{2} + (b_{3} + b_{5}^{*}1 + b_{1}^{*}25 + b_{2}25^{2} + a)$ we plot formula to end up with a plotting function for each regression line... curve height slope For this plot: • performance (the criterion) is on the Y axis For divorced \rightarrow perf' = $b_1 25 + b_7 25^2 + b_3 0 + b_4 1 + b_5 1 + b_6 exp + b_7 exp^2 + a$ experience (a quantitative predictor) is on the X axis perf' = $b_6 exp + b_7 exp^2 + (b_4 + b_5^{*1} + b_1^{*25} + b_2^{25^2} + a)$ slope \checkmark curve we plot we'll have 3 regression lines, one each for single, married & divorced height

How are experience & marital status related to perf for 25-year-olds from 2010s?
perf' = b₁age + b₂age² + b₃ mar1 + b₄mar2 + b₅era + b₆exp + b₇exp² + a

era coded 0 = 1980s 1 = 2010s
mar mar1 single = 1 mar2 divorced = 1 married = 0

Let's look at the plotting formulas again
For married → perf' = b₆exp + b₇exp² + (b₅*1 + b₁*25 + b₂25² + a)
For singles → perf' = b₆exp + b₇exp² + (b₃ + b₅*1 + b₁*25 + b₂25² + a)
For divorced → perf' = b₆exp + b₇exp² + (b₄ + b₅*1 + b₁*25 + b₂25² + a)
For divorced → perf' = b₆exp + b₇exp² + (b₄ + b₅*1 + b₁*25 + b₂25² + a)
Notice that each group's regression line has the same slope (b₆), because there is no interaction.
Notice that each group's regression line has the same curve (b₇), because there is no interaction.
The difference between the group's regression lines are their height – which differ based on the relationship between marital status and perf !!!