Plotting Linear Main Effects Models

- Interpreting 1st order terms w/ & w/o interactions
- Coding & centering… gotta? oughta?
- Plotting single-predictor models – Q, 2 & k
- Plotting 2-predictor models – 2xQ, kxQ & QxQ

Coding & Transforming predictors for MR models

- Categorical predictors will be converted to dummy codes
- Quantitative predictors will be centered, usually to the mean

Is this absolutely necessary?

Not usually… Many sources, and nearly all older ones, used unit-coded & un-centered predictors and their multiplicative combinations. Much of the time it works just fine…

So, why is dummy-coding and centering a good idea?

Mathematically – 0s (as control group & mean) simplify the math & minimize collinearity complications

Interpretively – the "controlling for" included in multiple regression weight interpretations is really "controlling for all other variables in the model at the value 0"

– "0" as the comparison group & mean will make b interpretations simpler and more meaningful

Very important things to remember…

All 1st order predictor regression weights have the same interpretation…

The expected direction and extent of change in Y for a 1-unit increase in the X after controlling for the other variable(s) in the model at the value 0

If that 1st order predictor is not involved in a higher-order effect (interaction), then the regression weight is interpreted as a main or unconditional effect - not being part of an interaction, the effect of that variable is the same for all values of all other variables.

If that 1st order predictor is involved in a higher-order effect, then the regression weight is interpreted as a conditional effect when the other variable(s) involved in the interaction = 0 – being part of an interaction, the effect is different for different values of those other variable(s) & the interaction weight describes the direction and extent of those differences

Plotting Single-Predictor models

Let's start by getting, interpreting and plotting a model with each of the different kinds of predictor we will be working with…

Be sure to "get" how each regression weight represents a particular slope, height or height differences in the plot!!!

Models with a single quantitative predictor

$$y' = bX + a$$

a → regression constant

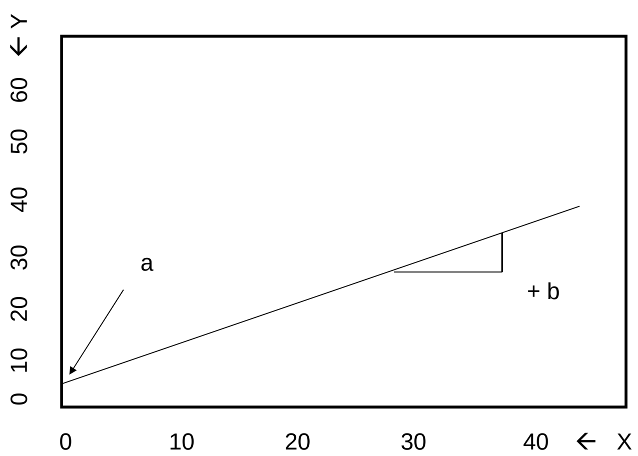
- expected value of y if x = 0
- height of predictor-criterion y-x regression line

b → regression weight

- expected direction and extent of change in y for a 1-unit increase in x
- slope of y-x regression line

Graphing & Interpreting
Models with a single quantitative predictor

$$y' = bX1 + a$$



a = height of line

b = slope of line

Models with a single centered quantitative predictor

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

a → regression constant

- expected value of y when x = 0 (re-centered to mean=0)
- height of y-x regression line

b → regression weight

- expected direction and extent of change in y for a 1-unit increase in x
- slope of y-x regression line

$\boxed{\text{X will be } X_{cen} \text{ in all of the following models}}$

Models with a single centered quantitative predictor

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

So, how do we plot this formula?

Simple → pick 2 values of x_{cen}, substitute them into the formula to get y' values and plot the line defined by those two x-y points

What x values?

- doesn't matter -- so keep it simple…
- 0 & 1, 0 & 10, 0 & 100, depending on the x-scale
- +/- 2 std isn't as simple, but tells you what x & y ranges are needed on the plot

Couple of things to remember…

- "0" is the center of the x-axis – X has been centered !!!
- the x-axis should extend about +/- 2 Std (include 96% of pop)

Models with a single centered quantitative predictor

$y' = 1.5X + 30$

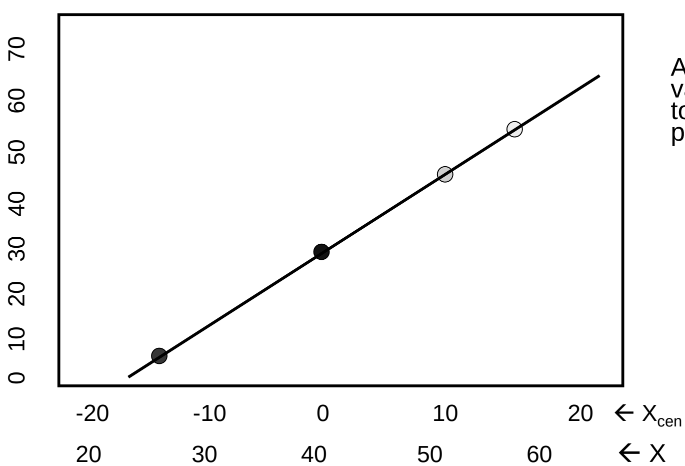
For $x_{cen} = 0$ → 1.6*0 + 30 y' = 30

For $x_{cen} = 10$ → 1.6*10 + 30 y' = 46

If X has mean = 42 & std = 7.5

For +2std → 1.6*15 + 30 y' = 54

For -2std → 1.6*-15 + 30 y' = 6



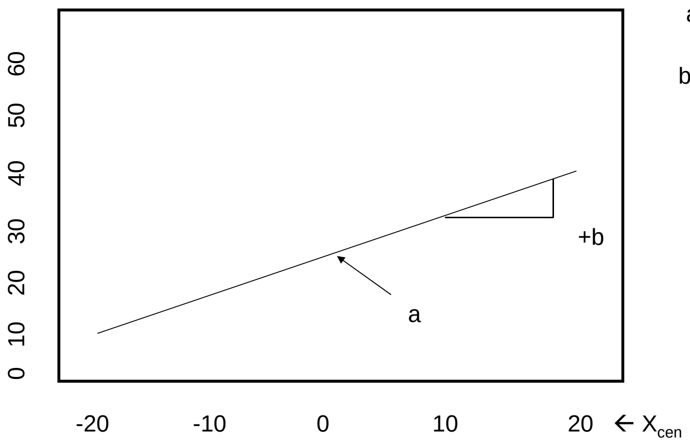
Any set of x values will lead to the same plotted line!

We can even substitute the original x scale back into the graph !!

Graphing & Interpreting
Models with a single centered quantitative predictor

$X_{cen} = X - X_{mean}$

$y' = bX_{cen} + a$



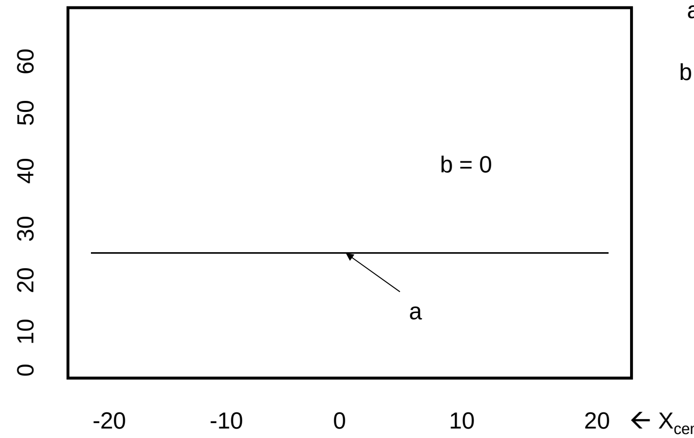
a = ht of line

b = slp of line

Graphing & Interpreting
Models with a single centered quantitative predictor

$X_{cen} = X - X_{mean}$

$y' = bX_{cen} + a$



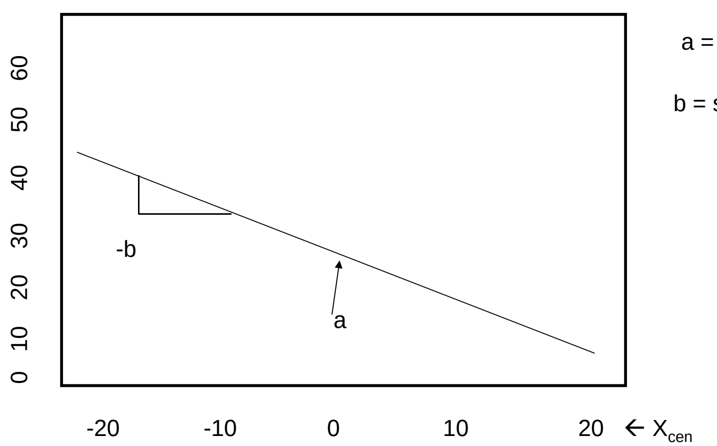
a = ht of line

b = slp of line

Graphing & Interpreting
Models with a single centered quantitative predictor

$X_{cen} = X - X_{mean}$

$y' = -bX_{cen} + a$



a = ht of line

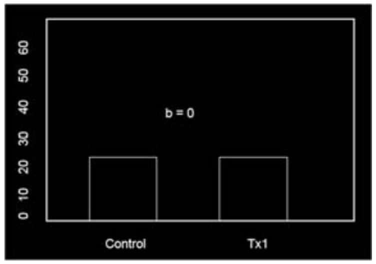
b = slp of line

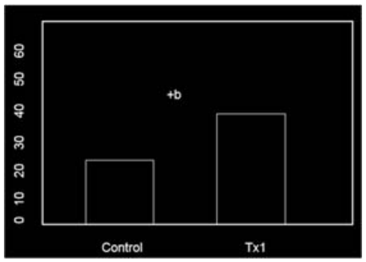
Plotting & Interpreting
models with a single binary predictor coded 1-2

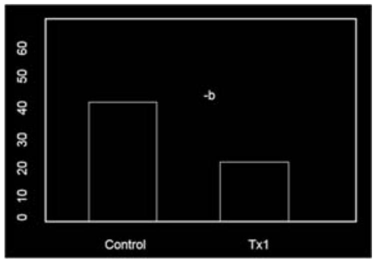
y' = bZ + a

Z = Tx1 vs. Cx

Cx = 1 Tx1 = 2







Models with a single dummy coded binary predictor

y' = bZ + a

a → regression constant

- expected value of y if z = 0 (the control group)
- mean of the control group
- height of control group

b → regression weight

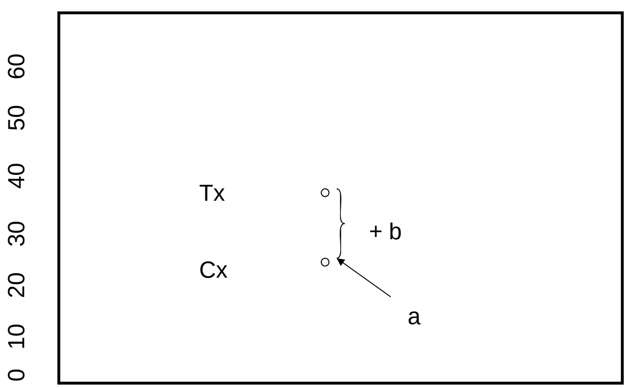
- expected direction and extent of change in y for a 1-unit increase in x
- direction and extent of y mean difference between groups coded 0 & 1
- group height difference

The way we're going to graph this model looks strange, but will provide simplicity & clarity for more complex models.

Plotting & Interpreting
Models with a single dummy coded binary predictor

y' = bZ + a

Z = Tx1 vs. Cx Cx = 0 Tx = 1



a = ht Cx

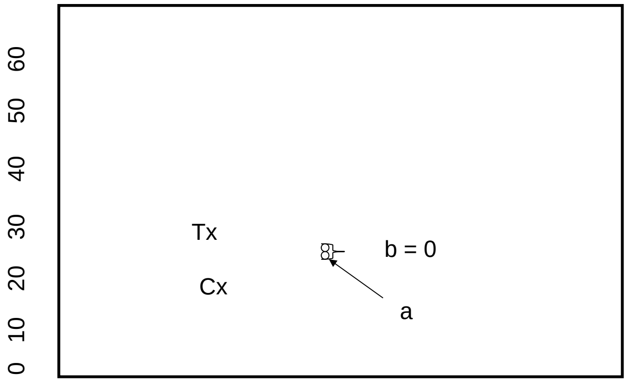
b = htdif Cx & Tx

Plotting & Interpreting

Models with a single dummy coded binary predictor

$y' = bZ + a$

Z = Tx1 vs. Cx   Cx = 0   Tx = 1



a = ht Cx

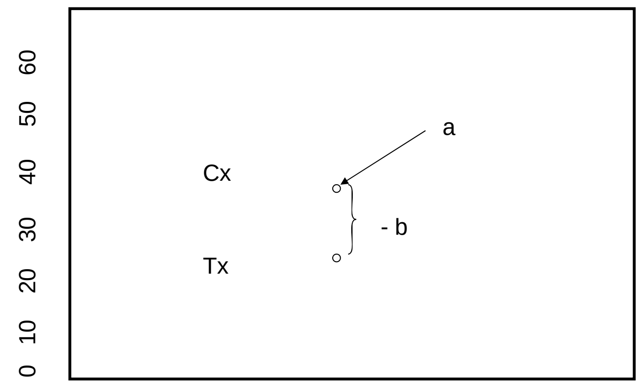
b = htdif Cx & Tx

Plotting & Interpreting

Models with a single dummy coded binary predictor

$y' = -bZ + a$

Z = Tx1 vs. Cx   Cx = 0   Tx = 1



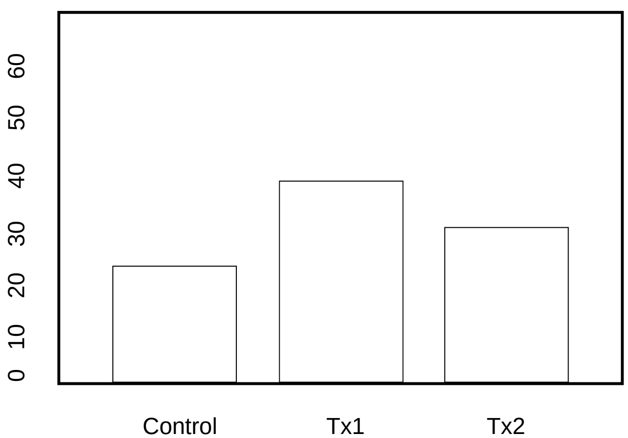
a = ht Cx

b = htdif Cx & Tx

Models with a single k-category predictor coded 1-3

~~y' = bX + a~~

Can't put the 1-3 coded variable into a regression – not a quant/interval variable



Models with a dummy coded k-category predictor – 2 dummy codes

$$y' = b_1Z_1 + b_2Z_2 + a$$

Group	Z_1	Z_2
1	1	0
2	0	1
3*	0	0

a → regression constant
- expected value of y if Z1 & Z2 = 0 (the control group)
- mean of the control group
- height of control group

b_1 → regression weight related to group 1 vs. target group
- expected direction and extent of change in y for a 1-unit increase in x
- direction and extent of y mean difference between comparison group and target group coded 1 on this variable
- group height difference between comparison & target groups (3 & 1)

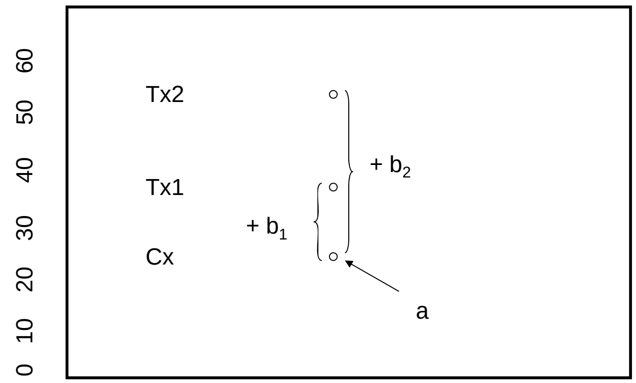
b_2 → regression weight related to group 2 vs. target group
- expected direction and extent of change in y for a 1-unit increase in x
- direction and extent of y mean difference between comparison group and target group coded 1 on this variable
- group height difference between comparison & target groups (3 & 2)

Again, the way we're going to graph this model looks strange, but will provide simplicity & clarity for more complex models.

Plotting & Interpreting
Models with a dummy coded k-category predictor – 2 dummy codes

$$y' = b_1Z_1 + b_2Z_2 + a$$

Z_1 = Tx1 vs. Cx(0) Z_2 = Tx2 vs. Cx(0)

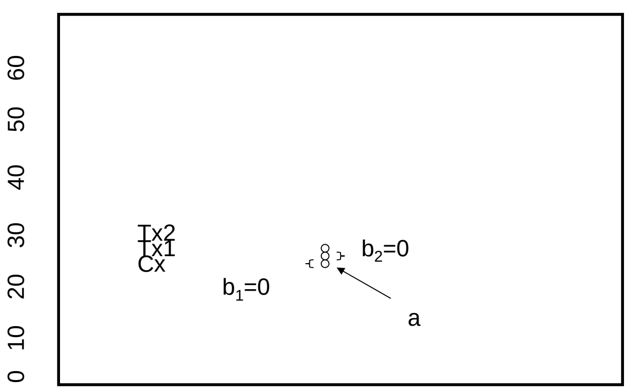


a = ht Cx
b_1 = htdif Cx & Tx_1
b_2 = htdif Cx & Tx_2

Plotting & Interpreting
Models with a dummy coded k-category predictor – 2 dummy codes

$$y' = b_1Z_1 + b_2Z_2 + a$$

Z_1 = Tx1 vs. Cx(0) Z_2 = Tx2 vs. Cx(0)



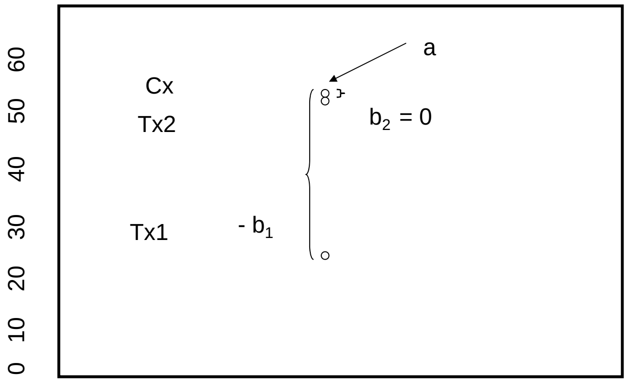
a = ht Cx
b_1 = htdif Cx & Tx_1
b_2 = htdif Cx & Tx_2

The dots should overlap – but that makes it hard to see…

Plotting & Interpreting
Models with a dummy coded k-category predictor – 2 dummy codes

$$y' = -b_1Z_1 + b_2Z_2 + a$$

Z_1 = Tx1 vs. Cx(0) Z_2 = Tx2 vs. Cx(0)



a = ht Cx

b_1 = htdif Cx & Tx_1

b_2 = htdif Cx & Tx_2

The top 2 dots should overlap – but that makes it hard to see…

Plotting 2-Predictor Linear Main Effects Models

Now we get to the fun part – plotting multiple regression equations involving multiple variables.

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

Models with a centered quantitative predictor & a dummy coded binary predictor

$$y' = b_1X + b_2Z + a$$

This is called a main effects model → 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 → 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 , 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)

b_2 → 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…

Model → $y' = b_1X + b_2Z + a$

For the Comparison Group coded Z = 0

Substitute the 0 in for Z $y' = b_1X + b_2{*}0 + a$

Simplify the formula $y' = b_1X + a$

slope height

For the Target Group coded Z = 1

Substitute the 1 in for Z $y' = b_1X + b_2{*}1 + a$

Simplify the formula $y' = b_1X + \underbrace{(b_2 + a)}$

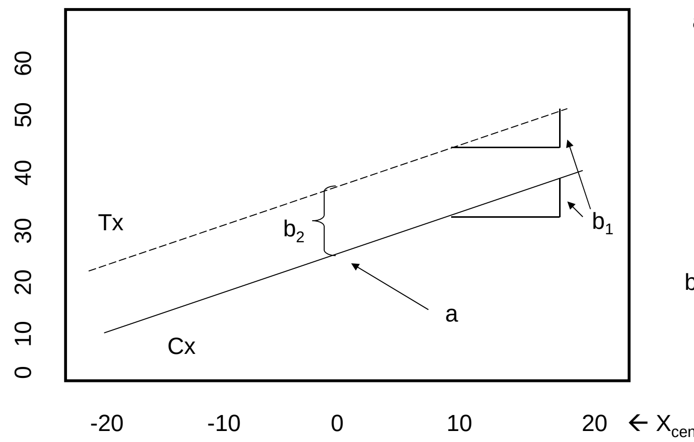
slope height

Plotting & Interpreting Models
with a centered quantitative predictor & a dummy coded binary predictor

$y' = b_1X + b_2 Z + a$

This is called a main effects model → no interaction → the regression lines are parallel.

$X_{cen} = X – X_{mean}$ Z = Tx(1) vs. Cx(0)



a = ht of Cx line
→ mean of Cx

b_1 = slp of Cx line

Cx slp = Tx slp
No interaction

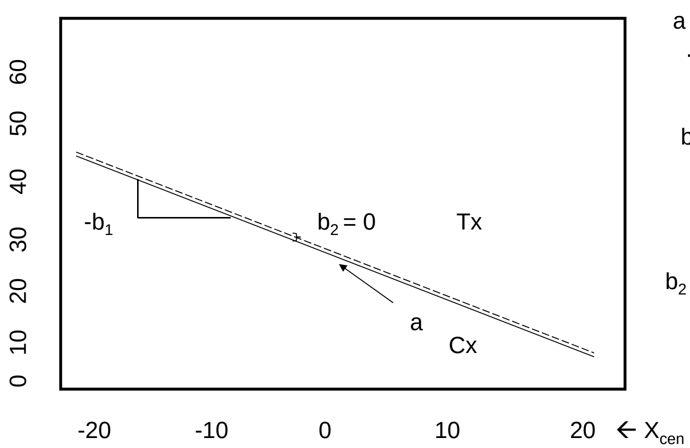
b_2 = htdif Cx & Tx
→ Cx & Tx mean dif

Plotting & Interpreting Models
with a centered quantitative predictor & a dummy coded binary predictor

$y' = -b_1X + -b_2 Z + a$

This is called a main effects model → no interaction → the regression lines are parallel.

$X_{cen} = X – X_{mean}$ Z = Tx(1) vs. Cx(0)



a = ht of Cx line
→ mean of Cx

b_1 = slp of Cx line

Cx slp = Tx slp
No interaction

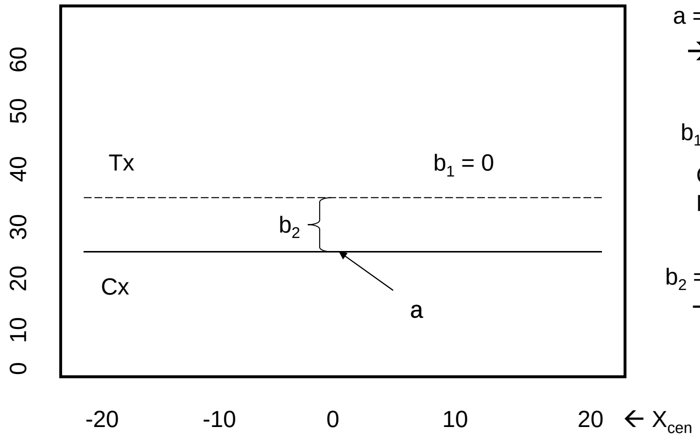
b_2 = htdif Cx & Tx
→ Cx & Tx mean dif

Plotting & Interpreting Models
with a centered quantitative predictor & a dummy coded binary predictor

This is called a main effects model → no interaction → the regression lines are parallel.

$$y' = b_1X + b_2Z + a$$

$X_{cen} = X – X_{mean}$ Z = Tx(1) vs. Cx(0)



a = ht of Cx line
→ mean of Cx

b_1 = slp of Cx line

Cx slp = Tx slp
No interaction

b_2 = htdif Cx & Tx
→ Cx & Tx mean dif

Models with a centered quantitative predictor & a dummy coded k-category predictor

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

$$y' = b_1X + b_2Z_1 + b_3Z_2 + a$$

Group	Z_1	Z_2
1	1	0
2	0	1
3*	0	0

a → regression constant
- expected value of y if Z_1 & $Z_2 = 0$ (the control group)
- mean of the control group
- height of control group quant-criterion regression line

b_1 → 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)

b_2 → 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_3 → regression weight for dummy coded comparison of G2 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 coded 0 & 1
- group height difference between comparison & target groups (3 & 2) (X = 0)

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

Model → $y' = b_1X + b_2Z_1 + b_3Z_2 + a$

Group	Z_1	Z_2
1	1	0
2	0	1
3*	0	0

For the Comparison Group coded $Z_1 = 0$ & $Z_2 = 0$

Substitute the Z code values $y' = b1X + b_2*0 + b_3*0 + a$

Simplify the formula $y' = \underbrace{b_1}_{\text{slope}}X + \underbrace{a}_{\text{height}}$

For the Target Group coded $Z_1 = 1$ & $Z_2 = 0$

Substitute the Z code values $y' = b1X + b_2*1 + b_3*0 + a$

Simplify the formula $y' = \underbrace{b_1}_{\text{slope}}X + \underbrace{(b_2 + a)}_{\text{height}}$

For the Target Group coded $Z_1 = 0$ & $Z_2 = 1$

Substitute the Z code values $y' = b1X + b_2*0 + b_3*1 + a$

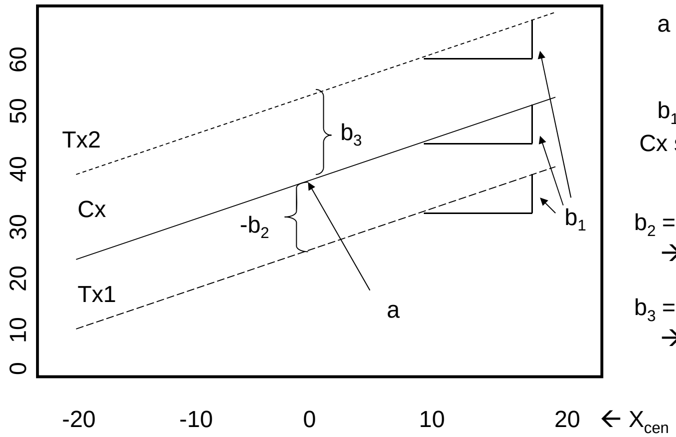
Simplify the formula $y' = \underbrace{b_1}_{\text{slope}}X + \underbrace{(b_3 + a)}_{\text{height}}$

Plotting & Interpreting Models
with a centered quantitative predictor & a dummy coded k-category predictor

$y' = b_1X + -b_2 Z_1 + b_3 Z_2 + a$

This is called a main effects model → no interaction → the regression lines are parallel.

$X_{cen} = X – X_{mean}$ Z_1 = Tx1 vs. Cx(0) Z_2 = Tx2 vs. Cx (0)



a = ht of Cx line
→ mean of Cx

b_1 = slp of Cx line
Cx slp = Tx_1 slp = Tx_2 slp
No interaction

b_2 = htdif Cx & Tx_1
→ Cx & Tx_1 mean dif

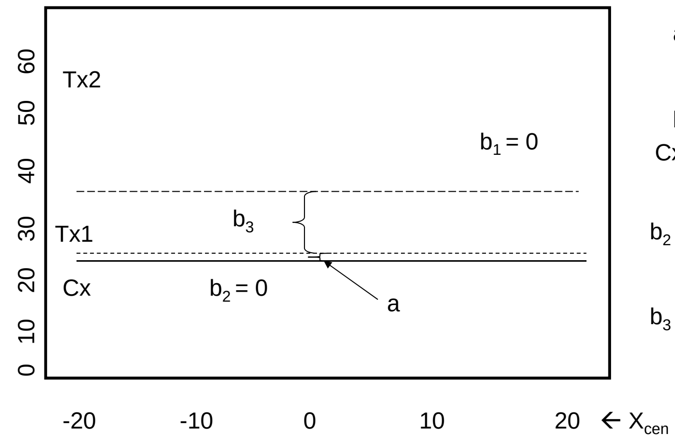
b_3 = htdif Cx & Tx_2
→ Cx & Tx_2 mean dif

Plotting & Interpreting Models
with a centered quantitative predictor & a dummy coded k-category predictor

$y' = b_1X + b_2 Z_1 + b_3 Z_2 + a$

This is called a main effects model → no interaction → the regression lines are parallel.

$X_{cen} = X – X_{mean}$ Z_1 = Tx1 vs. Cx(0) Z_2 = Tx2 vs. Cx (0)



a = ht of Cx line
→ mean of Cx

b_1 = slp of Cx line
Cx slp = Tx_1 slp = Tx_2 slp
No interaction

b_2 = htdif Cx & Tx_1
→ Cx & Tx_1 mean dif

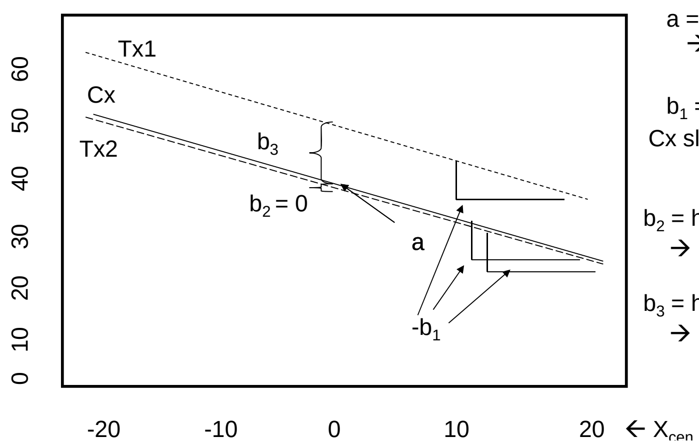
b_3 = htdif Cx & Tx_2
→ Cx & Tx_2 mean dif

Plotting & Interpreting Models
with a centered quantitative predictor & a dummy coded k-category predictor

$y' = -b_1X + b_2 Z_1 + b_3 Z_2 + a$

This is called a main effects model → no interaction → the regression lines are parallel.

$X_{cen} = X – X_{mean}$ Z_1 = Tx1 vs. Cx(0) Z_2 = Tx2 vs. Cx (0)



a = ht of Cx line
→ mean of Cx

b_1 = slp of Cx line
Cx slp = Tx_1 slp = Tx_2 slp
No interaction

b_2 = htdif Cx & Tx_1
→ Cx & Tx_1 mean dif

b_3 = htdif Cx & Tx_2
→ Cx & Tx_2 mean dif

weights for main effects models w/ a centered quantitative predictor **&** a dummy coded binary predictor **or** 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 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

- the direction and extent of the expected mean difference of the Target group for that dummy code 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

Models with 2 centered quantitative predictors

$$y' = b_1X_1 + b_2X_2 + a$$

This is called a main effects model → 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 X & Z variables are centered quantitative variables
- what we'll do is to plot the Y-X_1 regression line for different values of X_2
- the most common approach is to plot the Y-X regression line for…
  - the mean of X_2
  - +1 std above the mean of X_2
  - -1 std below the mean of X_2

We'll plot 3 lines

Models with 2 centered quantitative predictors

$$y' = b_1X_1 + b_2X_2 + a$$

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

a → regression constant

- expected value of y if X=0 (mean) Z=0 (mean)
- mean criterion score of those having Z=0 (mean)
- height of quant-criterion regression line for those with Z=0 (mean)

b_1 → regression weight for X_1 centered quant predictor

- expected direction and extent of change in y for a 1-unit increase in X_1, 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 X_2=0 (mean)

b_2 → regression weight for X_2 centered quant predictor

- expected direction and extent of change in y 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 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

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

Model ➔ $y' = b_1X + b_2X_1 + a$

For $X_2 = 0$ (the mean of centered Z)

Substitute the 0 in for X_2 $\qquad y' = b_1X + b_2{*}0 + a$

Simplify the formula $\qquad y' = b_1X + a$

slope height

For $X_2 = +1$ std

Substitute the std value in for X_2 $\qquad y' = b_1X + b_2{*}\text{std} + a$

Simplify the formula $\qquad y' = b_1X + (b_2{*}\text{std} + a)$

slope

For $X_2 = -1$ std

Substitute the std value in for X_2 $\qquad y' = b_1X + -b_2{*}\text{std} + a$

Simplify the formula $\qquad y' = b_1X + (-b_2{*}\text{std} + a)$

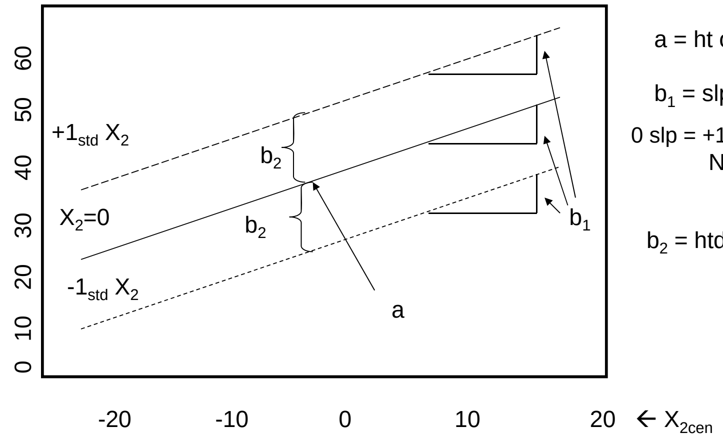
slope height

Plotting & Interpreting Models with 2 centered quantitative predictors

$y' = b_1X_{1cen} + b_2X_{2cen} + a$

This is called a main effects model → no interaction → the regression lines are parallel.

$X_{1cen} = X_1 - X_{1mean}$ $\qquad X_{2cen} = X_2 - X_{2mean}$



a = ht of X_{2mean} line

b_1 = slp of X_{2mean} line

0 slp = +1std slp = -1std slp
No interaction

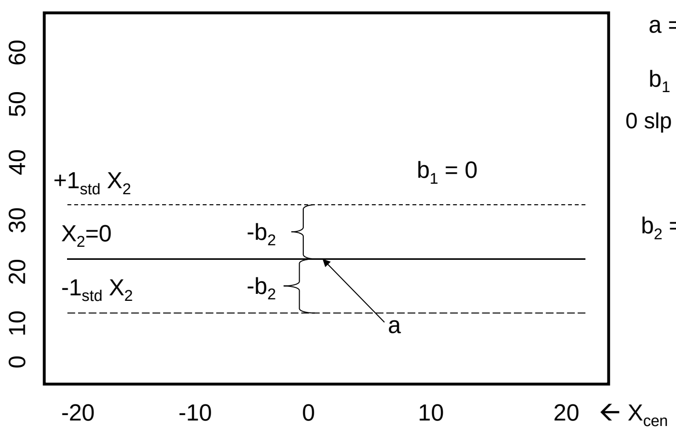
b_2 = htdifs among X_2-lines

Plotting & Interpreting Models with 2 centered quantitative predictors

$y' = b_1X_{1cen} + b_2X_{2cen} + a$

This is called a main effects model → no interaction → the regression lines are parallel.

$X_{1cen} = X_1 - X_{1mean}$ $\qquad X_{2cen} = X_2 - X_{2mean}$



a = ht of X_{2mean} line

b_1 = slp of X_{2mean} line

0 slp = +1std slp = -1std slp
No interaction

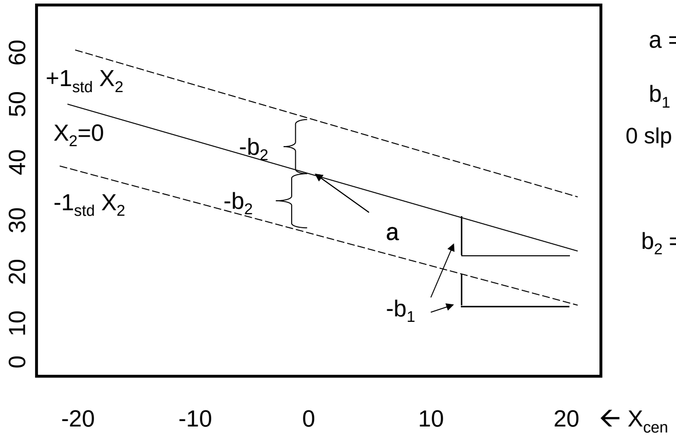
b_2 = htdifs among X_2-lines

Plotting & Interpreting Models
with 2 centered quantitative predictors

This is called a main effects model → no interaction → the regression lines are parallel.

$y' = b_1 X_{1cen} + b_2 X_{2cen} + a$

$X_{1cen} = X_1 – X_{1mean}$ $X_{2cen} = X_2 – X_{2mean}$



a = ht of X_{2mean} line

b_1 = slp of X_{2mean} line

0 slp = +1std slp = -1std slp
No interaction

b_2 = htdifs among X_2-lines