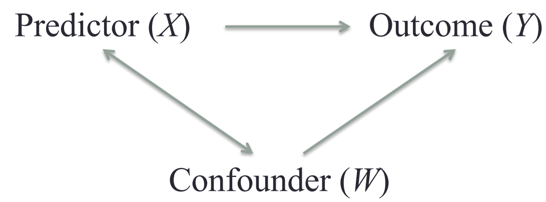
Three-Way Tables: Confounders and Effect Modifiers

Confounding

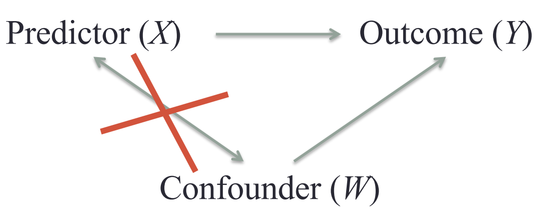
One definition: A **confounder** is a variable that is associated with the predictor of interest (X) and causally related to the outcome of interest (Y).



→ You cannot tell if the effect seen on Y is really due to X or is due to W. X and W are said to be **confounded**.

Random Assignment

The goal of **randomly assigning** values of X to experimental units is to (on average) remove the association between X and W:



Confounding: Example

X = Race (Afr. American or white)

Y = Attend college (yes or no)

Confounders??

- Example: In the U.S., household income is likely to confound the relationship between race (Afr. American vs. white) and college attendance.
  - Children from higher income families are more likely to attend college (affects Y)
  - On average, household income is lower among African Americans than whites (related to X).

How to Deal with Confounding

Need to "adjust" for the confounder → Look at the relationship between X and Y at fixed level of W.

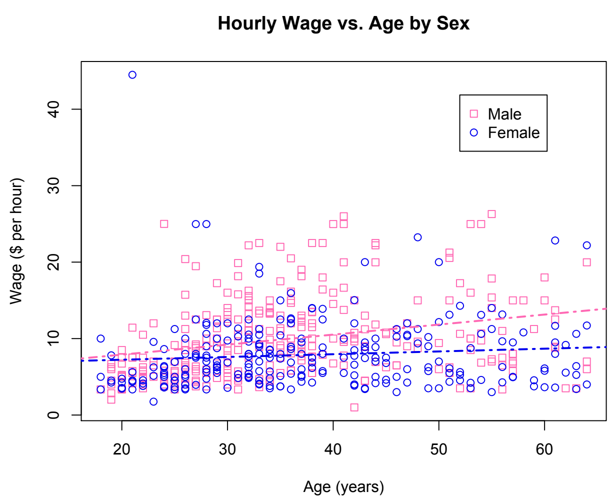
- If W is categorical, we can stratify on W.
- If W is continuous, adjust for it in a regression context.

Effect Modification

If the association between a predictor of interest X and outcome Y varies depending upon the level of W, then W is said to be an **effect modifer**.

- We also say that an **interaction** exists between X and W (on Y).

Regression Example



- X = Age
- Y = Hourly Wage ($)
- W = Gender (M/F)

How to Deal with Effect Modification (Interaction)

- Again, consider the association between X and Y separately at fixed levels of W.
- Adjust for the interaction in a regression context.

I x J x K Tables

- *X* has *I* levels; *Y* has *J* levels; *W* has *K* levels.
- A slice of the three-way table classifying *X* and *Y* at a given level of *W* is called a **partial table**.
  - Associations in partial tables are called **conditional associations**.
- The two-way table resulting from combining the partial tables is called the ***XY* marginal table**.
  - Associations in marginal tables are called **marginal associations**.

Association in 3-way Tables

- **Conditional independence** of *X* and *Y* *given W* → *X* and *Y* are independent in each partial table (for each value of *W*)
- **Marginal independence** of *X* and *Y* → *X* and *Y* are independent in the marginal table (summed across all values of *W*)

Example: 2 x 2 x 2 Table

Partial tables of X and Y given W:

	W = 1		W = 2	
	Y = 1	Y = 2	Y = 1	Y = 2
X = 1	4	6	16	4
X = 2	6	9	4	1

Are X and Y conditionally independent given W?

Are X and Y marginally independent?

Is W an effect modifier? a confounding variable?

Special case: 2 x 2 x K Tables

X and *Y* are binary; *W* has *K* levels. Notation:

- (conditional) odds ratio of *X* and *Y* given $W = k$: $\theta_{XY(k)}$
- (marginal) odds ratio of *X* and *Y*: θ_{XY}

We say there is ***homogenous association*** between *X* and *Y* if

$$\theta_{XY(1)} = \theta_{XY(2)} = \ldots = \theta_{XY(K)}$$

i.e., there is *no interaction* between *X* and *W* on their effects on *Y*.

If X and Y are conditionally independent given W, then

$$\theta_{XY(1)} = \theta_{XY(2)} = \ldots = \theta_{XY(K)} = \mathbf{1}$$

Inference for Three-Way Tables

- Test whether X and Y are independent at each level of W, where W takes on K levels → **Mantel-Haenszel Test**
  - Will see for $2 \times 2 \times K$ tables in Ch. 4.
- Or test whether odds ratios between X and Y are the same across all levels of W → **Breslow-Day Test**
  - Will see $2 \times 2 \times K$ tables in Ch. 4.
- Use a generalized linear model (Chs. 3-6) with Y as the response variable, and X and W as predictors in the model.