# Intro to Stats

Correlations

## Correlation Coefficients

- Use it when:
  - test the relationship between variables (not differences in means)
  - Only two variables at a time

# Correlation coefficient

- Captures how the value of one variable changes when the value of the other changes
- Ranges from −1 to +1
- A Pearson correlation is based on continuous variables

- Important to remember this is a relationship for a group, not each person
- Reflects the amount of variability shared by two variables

# Computations

$$r_{xy} = \frac{n\,\Sigma XY - \Sigma X\,\Sigma Y}{\sqrt{[n\,\Sigma X^2 - (\Sigma X)^2][n\,\Sigma Y^2 - (\Sigma Y)^2]}}$$

- $r_{xy}$ = correlation coefficient between x & y
- n = size of sample
- X = score on X variable
- Y = score on Y variable

# Why you can't say "x caused y"

- Even though you may suspect there's a causal relationship you can only make causal statements if:
  - X definitely preceded Y
  - X was manipulated so that it was the only probable factor that could cause changes in Y
- When talking correlations, you can use "relationship", "relate", "associated"

# Table: Correlation Matrix

| | | yrsed | income |
|---|---|---|---|
| yrsed | Pearson Correlation | 1 | .95(***) |
| | Sig. (2-tailed) | | .000 |
| | N | 6 | 6 |
| income | Pearson Correlation | .95(***) | 1 |
| | Sig. (2-tailed) | .000 | |
| | N | 6 | 6 |

## Interpretations

| | |
|---|---|
| .80 to 1.0 | Very strong |
| .60 to .80 | Strong |
| .40 to .60 | Moderate |
| .20 to .40 | Weak |
| .00 to .20 | Weak/ None |

## Interpretations

- Coefficient of Determination
  - Percentage of variance in one variable that is accounted for by variance in the other
- Square the correlation coefficient

- If r = .70
  $r^2$ = .49
  49% of variance in one is explained by variance in other