# §12.1: Inferences for Regression

Skills

- Construct a confidence interval for the slope of the regression line
- Conduct a test of significance for the slope of the regression line

# The Idea

It begins with a question:

- what line describes the true relationship between these variables?

If we could gather all the data, we'd be able to construct this line...i.e., find the two numbers (true slope, true intercept) needed to write the equation.

Since we can't measure the population, we'll have to settle for a sample.

Every sample is going to produce a different line.
Every sample is going to produce two statistics (sample slope, sample intercept).
...so there are really two sampling distributions involved in this!

What's more important: slope, or intercept?

The Parameters: α, β
The Statistics: *a*, *b*

The Sampling Distribution of *b*:

Mean: $\mu_\beta = \beta$
Standard Error: $SE_b = \sqrt{\dfrac{\sum (y_i - \hat{y}_i)^2}{(n-2)\sum (x_i - \overline{x})^2}}$

(*let's not get into its shape just yet*)

Under certain circumstances,

$$t = \frac{b - \beta}{SE_b}$$

has a *t* distribution with *n* - 2 degrees of freedom.

About those conditions...

# The Conditions

- The variables must have a linear relationship
- The data were collected randomly (sample or experiment)
- *10% Rule applies if this is a sample
- The standard deviation of $y$ does not vary with $x$
- For any particular value of $x$, $y$-values should be normally distributed

# Linear Relationship

There should be no curvature in the residual plot.

# Random Sample/Experiment

This is as it has always been.

Remember that the 10% rule only applies if the data were collected via a random sample.

# Standard Deviation of $y$

Check the residual plot for a triangle pattern.

# Normality of $y$-values

Check a histogram of the residuals for normality.

- If the histogram is approximately normal, all is well.
- If the histogram isn't approximately normal, then we need at least 30 data.

# The Interval

statistic ± (critical value)(measure of variation)

$b \pm t^* SE_b$

*...and all of the usual stuff from there.*

# The Test

$H_0$: β = 0

$H_a$: β ?? 0

This calls for a *t*-test for slope.
*...and then all the usual stuff.*

# More about $SE_b$...

- Finding it in computer output
- Tricking the calculator into giving it to you
- The meaning of $s$

# Example

A study examined the age (years) of young mothers and the birth mass (grams) of their children...

| | | | | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| x | 15 | 17 | 18 | 15 | 16 | 19 | 17 | 16 | 18 | 19 |
| y | 2289 | 3393 | 3271 | 2648 | 2897 | 3327 | 2970 | 2535 | 3138 | 3573 |