may 31, 2017

This time: regression
Next time: ANOVA

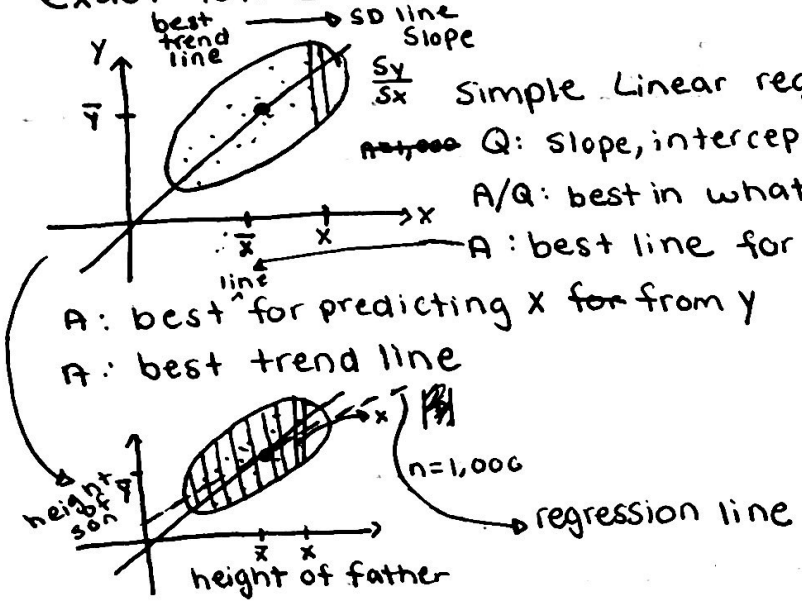
read: LN pp. L-245 →

HW #4 due Friday June 9, 2017 @ 11:59 pm
↳ Final will be passed out that day a due a week later on Friday June 16, 2017 at 11:59pm

*There will be no HW #5 *

★ There will be either 1 or 2 makeup lectures in finals week, on Monday (or Tuesday)
↳ more details soon

exact 95% CI method: LN pp. L-231 → L-244



Simple Linear regression
Q: slope, intercept of best line?
A/Q: best in what sense?
A: best line for predicting y from x
A: best ^ for predicting x ~~for~~ from y
A: best trend line

eq. of regression line

$\hat{y} = \hat{\beta}_0 + \beta_1 x$

"beta"

predicted y value

y-intercept

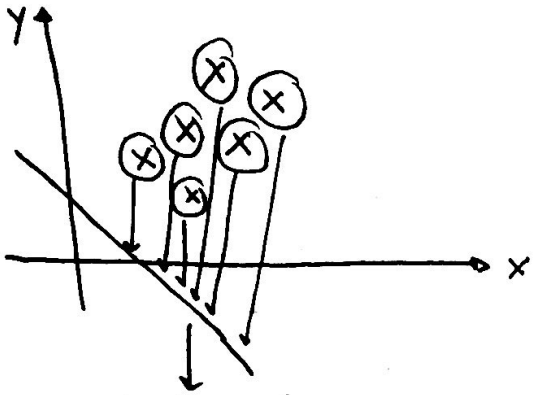
slope of regression line

$$\hat{\beta}_1 = r \cdot \frac{S_y}{S_x}$$ equation #17 p. R-25

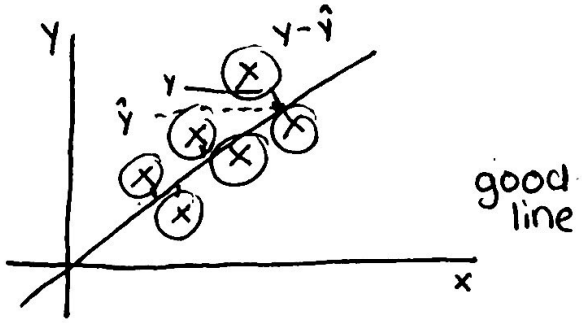
regression line goes through $(\bar{x}, \bar{y})$

$\bar{y} = \hat{\beta}_0 + \beta_1 x$ so $\hat{\beta}_0 = \bar{y} - \beta_1 x$ eq. # 21 p. R-25

regression to the mean



bad line (all predictions too low)



good line

find $(\hat{\beta_0}, \hat{\beta_1})$ to

best line solves this problem: minimize

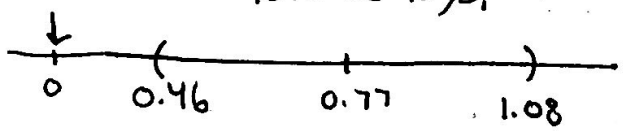
$$\frac{1}{n}\sum_{i=1}^{n}\left[y_i - (\hat{\beta_0} + \hat{\beta_1}x)\right]^2$$

$$\begin{bmatrix} Y_1 & X_1 \\ \vdots & \vdots \\ Y_2 & X_2 \\ \vdots & \vdots \\ Y_n & X_n \end{bmatrix}$$

$\hat{y} = \hat{\beta_0} + \hat{\beta_1}x$

math fact: least squares lines = regression line

95% CI for β_1



the difference between 0.77 and 0 (is) statsig