this: regression

time: ↓

next time: ANOVA

read: LN pp. L-245 → 274

AMS7 31 May 17

I am cancelling homework ⑤

homework 4 due 11.59pm Fri 9 Jun 17

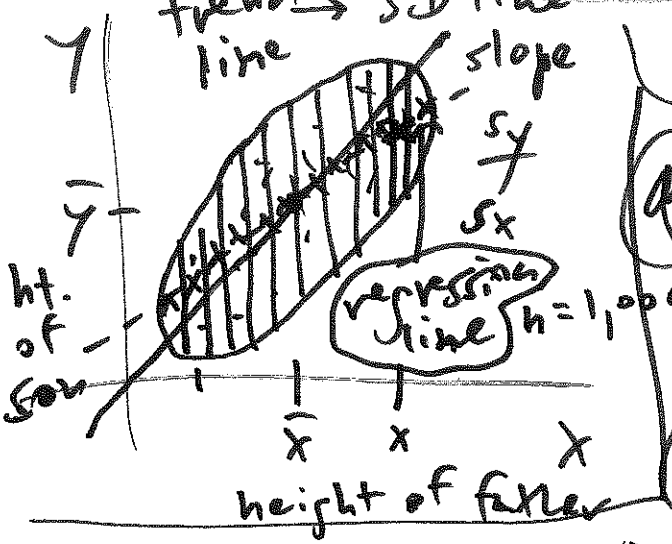
there will be either 1 or 2 makeup lectures in finals week, on Mon (& Tues). more details soon

take home final handed out in class Fri 9 Jun, due 1 week later

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?

Gauss (1800)

Galton (1890)

A: ← now best line for predicting y from x

(best line for pred. x from y) (best trend line)

slope of regression line

$$\hat{\beta}_1 = r \cdot \frac{s_y}{s_x}$$

equation (1) p. R-25

equation of regression line

$$\hat{y} = \hat{\beta}_0 + \hat{\beta}_1 x$$

("beta")

$\hat{y}$: predicted y value

$\hat{\beta}_0$: y-intercept

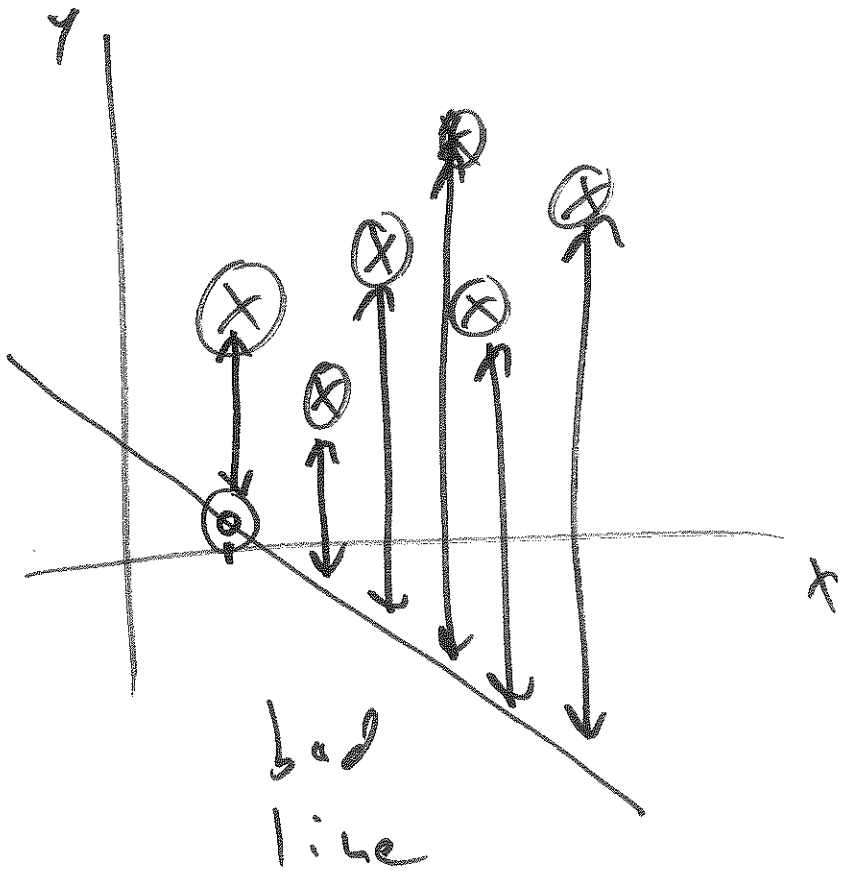
reg. line goes through $(\bar{x}, \bar{y})$

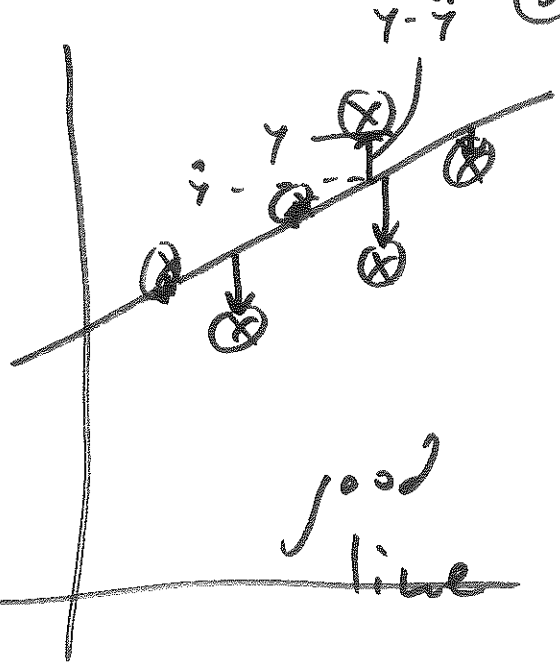
$$\bar{y} = \hat{\beta}_0 + \hat{\beta}_1 \bar{x}$$

so $$\hat{\beta}_0 = \bar{y} - \hat{\beta}_1 \bar{x}$$

eqn. (21) p. R-25

regression to the mean





Gauss

best line solves fit problem: find $(\hat{\beta}_0, \hat{\beta}_1)$ to minimize $\frac{1}{n}\sum_{i=1}^{n}\left[y_i - \overbrace{(\hat{\beta}_0 + \hat{\beta}_1 x_i)}^{\hat{y}_i}\right]^2$

Laplace

y_i	x_i
y_i	x_i
y_n	x_n

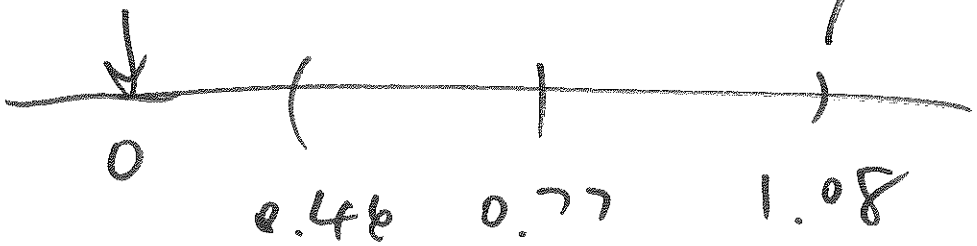
$$\hat{y} = \hat{\beta}_0 + \hat{\beta}_1 x$$

least squares = lines

math fact

regression line

95% CI for β_1



(4)

the diff between 0.77 & 0 is statistically (15) significant