

Final exam review

final idea on regression: 2 June
p. 123 #4 Ans 5

Sat noon-2 (here)

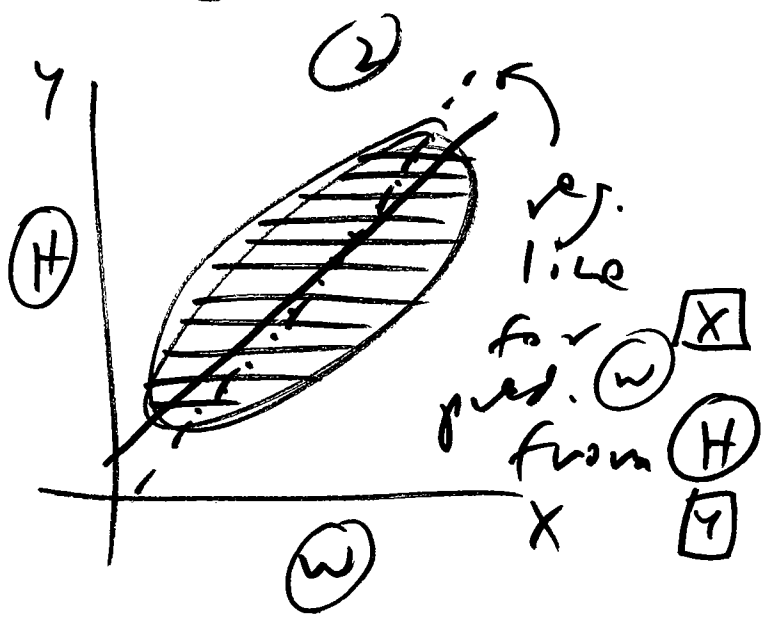
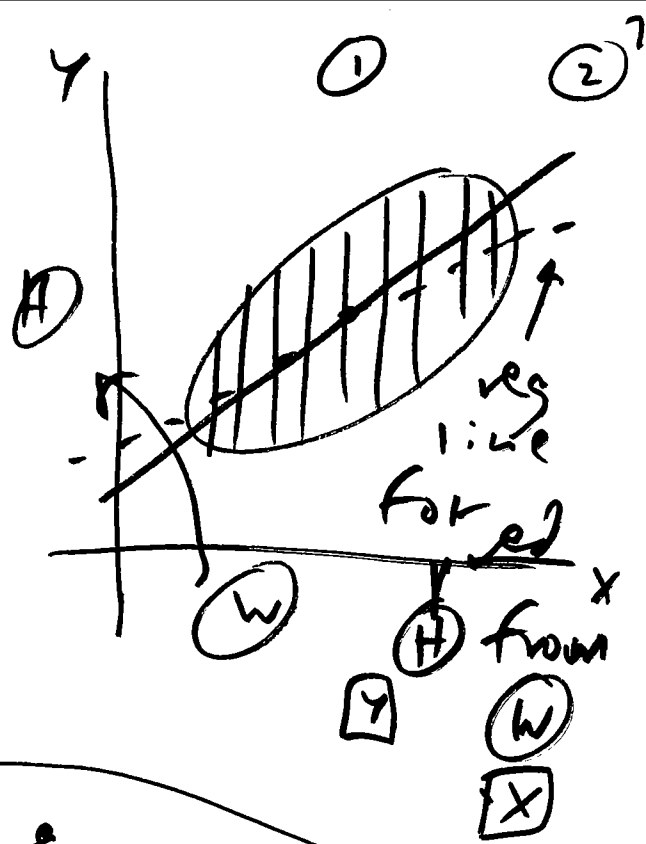
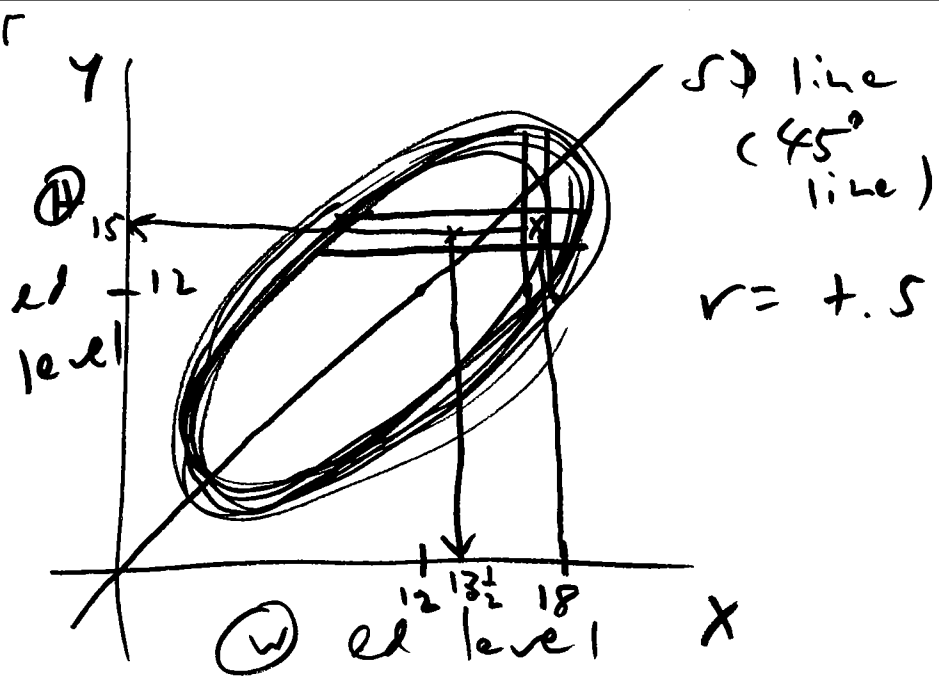
| variable | mean | SD |
|-------------------|--------|-------|
| (H) ed. level (Y) | 12 yrs | 3 yrs |
| (W) ed. level (X) | 12 yrs | 3 yrs |

$r = +0.5$

(1) (W) has 18 yrs of schooling
 $\rightarrow \frac{18-12}{3} = 2$
 SDs above ave for (W) \rightarrow we

expect low (H) only ~~2~~ $r \cdot 2$ SDs above ave
 in ed. $\rightarrow (0.5)(2 \text{ SDs}) = 1 \text{ SD above ave} \rightarrow$
 predict (H) ed level = 15 yrs

(2) (H) has 15 yrs of schooling $\rightarrow \frac{15-12}{3} = 1 \text{ SD above ave}$
 for (H) \rightarrow we expect his (W) to be only
 $r \cdot 1 = (0.5)(1 \text{ SD above ave}) \rightarrow \frac{1}{2} \text{ yrs above ave}$
 $\rightarrow 13.5 \text{ yrs of schooling predicted for (W)}$



There are actually 2 different reg. lines in any scatter plot: one for predicting

Y from X, & a different line for pred.

X from Y

note: if somebody

switches the role of X & Y in the middle of a problem, watch out.

final is open-book, open notes; 5 problems ⁽³⁾
in no particular order:

- correlation & regression
- probability models for ^{mid-term} sums (escalator problem)
- 2 indep samples (continuous data)
- 1-sample problem with 0/1 data (Spock)
- 2 sample paired comparison

some where in 1 of these problems:
you will be asked either to do inference
or to say why doing inference would
be inappropriate

S. (a) \swarrow 2d. level

$$\hat{y} = 66.75_{in} + \left(0.25 \frac{in}{yr}\right) (x)$$

ht. for a guy with high school ed,

$$x=12 \rightarrow \hat{y} = 66.75_{in} + \left(0.25 \frac{in}{yr}\right) (12_{yr}) \\ = 69.75_{in}$$

$$\vec{y} = 66.75 \text{ in} + (0.25 \frac{\text{in}}{\text{yr}})(16 \text{ yrs})$$

$$= 70.75 \text{ in} \quad (\text{1 inch taller})$$

(4)

Q: What does this mean?
 (going back to college, seeing if the guy gets taller) longitudinal question but data cross-sectional

$$\vec{y} = \beta \frac{\text{in}}{\text{yr}}$$

↑
0.25

A₁: slope came out pos → corr. is pos → PCF: income (income ↑ height ↑ (nutrition) income ↑ ed ↑ \$)

A₂: If we look at 2 groups of men who differ on age by 1 yr of age we can expect them to differ on age by 0.25 in in ht.

(b) + corr due to better nutrition as child (might get a raise with elevator shares, but simpler explanation of this corr. is PCF) (nutrition)

