STAT 7
26 Nov 19

①

this time: correlation & next time: regression; ANOVA

read: LN pp. 1-269 + 282

today: LN p. 221 → (216)

hwk 4 due Sun night 8 Dec

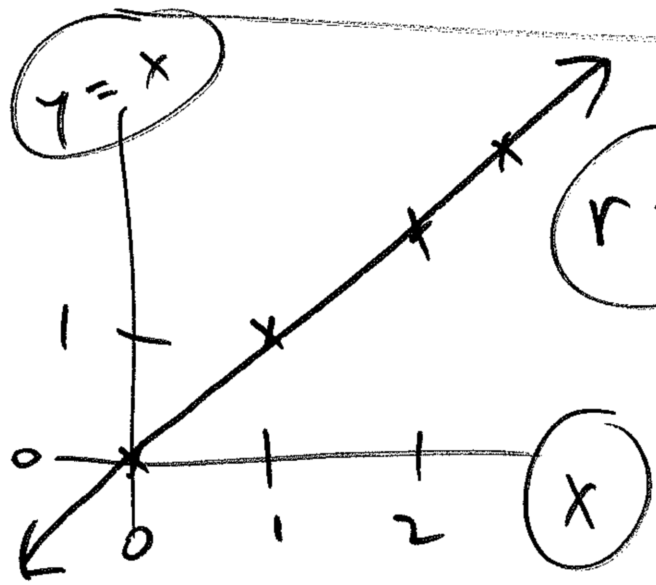
quiz 8 due tonight

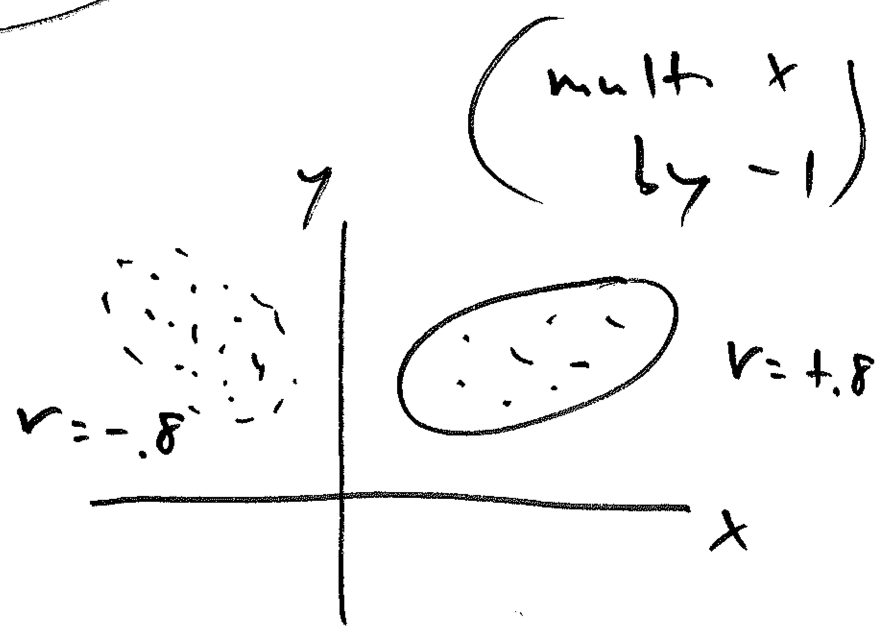
take-home final due Sun night 15 dec

quiz 9 ...

make-up lecture? Mon night 2 Dec 7.10-8.45p Baskin Auditorium

disc. sections (will) occur next week





Q: is $r = +.87$ large in practical terms?

A: L-216 ① start with the smallest spans:

$x \doteq 10$ cm $\rightarrow$ $y \doteq 7$ cm ② then consider the largest spans: $x \doteq 11.5$ cm

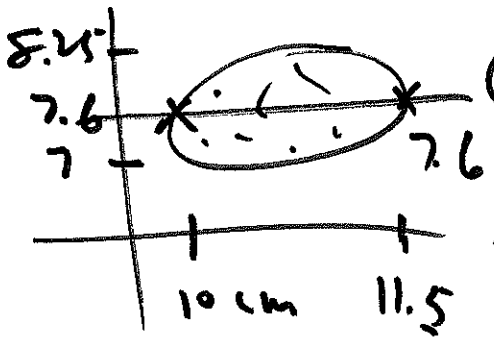
$\rightarrow y \doteq 8.25$ cm

or x goes from 10 to 11.5

y goes from 7 to 8.25

$\frac{8.25 - 7}{7} = +1.25$

A: Yes



$r = 0$

$\frac{7.6 - 7.6}{7.6} = 0\%$

not practically

$= 18\%$

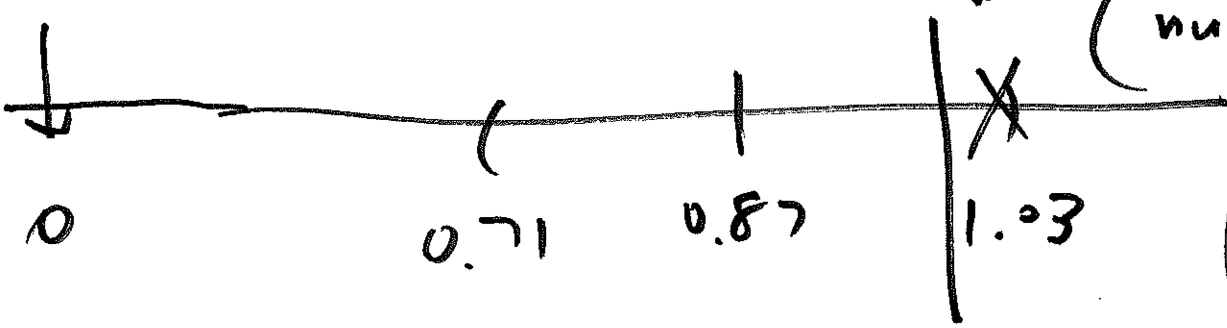
for lack of time skip "exact"

CI method for ρ: L-23 ~~→ L-24~~

(crude approx.) 95% CI for ρ: $r \pm 2\,\hat{SE}(r)$

$0.87 \pm 2\,(0.8)$

(null: $\rho = 0$)

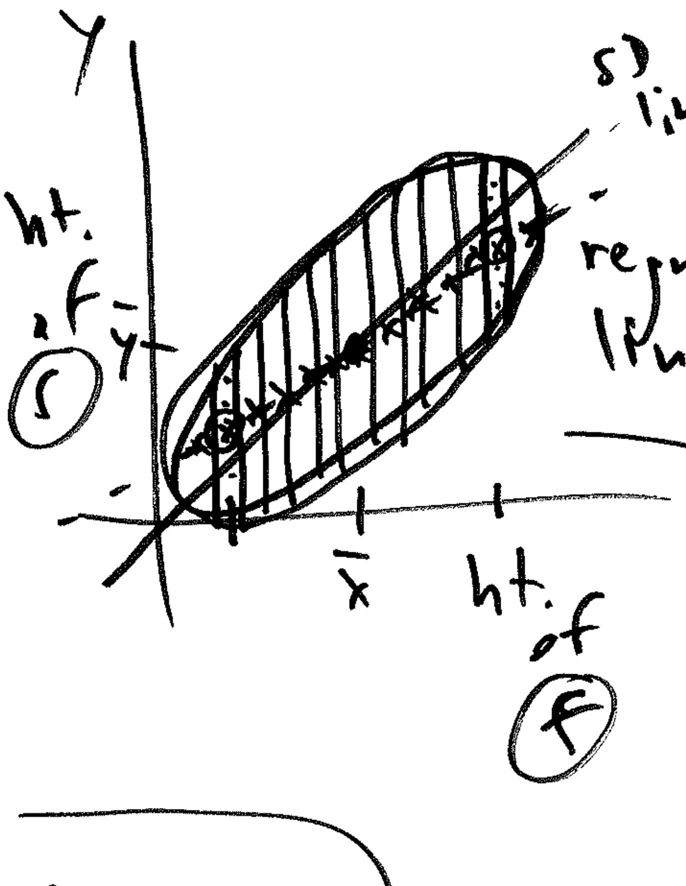


(highly statis.)

1.0 best

how predict y from x? (eugenics)

Galton (1895)



ht. of s (sons)	ht. of f
y	x
mean ~~68~~ 67 in	67 in
SD 2.5 in	2.5 in

($r = +.65$)

$n = 1050$

1 row for each family

Mendel (1865)

1885

1890

regression line for pred. y from x : $\hat{y} = \hat{\beta}_0 + \hat{\beta}_1 x$

$\hat{y}$: pred. y value; $\hat{\beta}_0$: y intercept; $\hat{\beta}_1$: slope

$$\hat{\beta}_1 = r \cdot \left(\frac{s_y}{s_x}\right)$$

↑ SD line

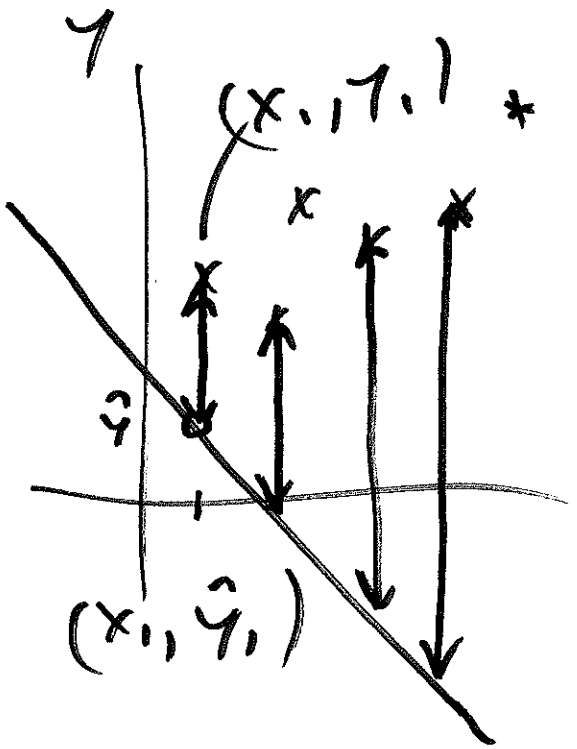
$$\bar{y} = \hat{\beta}_0 + \hat{\beta}_1 \bar{x}$$

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

(Legendre) (Gauss)

|—|

(Laplace)



regression line = least squares line

$$\frac{1}{n}\sum_{i=1}^{n} (y_i - \hat{y}_i)^2 =$$

find $(\hat{\beta}_0, \hat{\beta}_1)$ to

$$\frac{1}{n}\sum_{i=1}^{n} \left(y_i - (\hat{\beta}_0 + \hat{\beta}_1 x_i)\right)^2$$

minimize ↗