this time: regression

next time: ANOVA

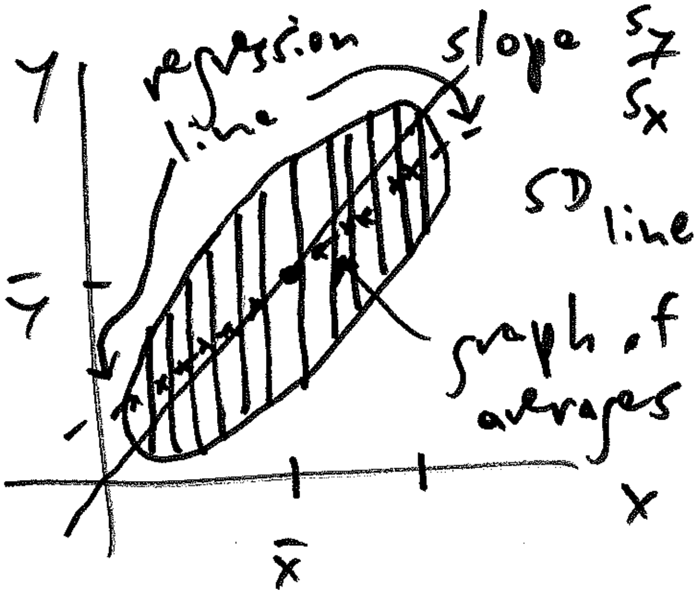
read: LN pp. L-269 + L-301 | AMS 7 27 Nov 18

today: L-245 → L-261 (1)

hwk 3 due by 11.59 pm tomorrow @ canvas

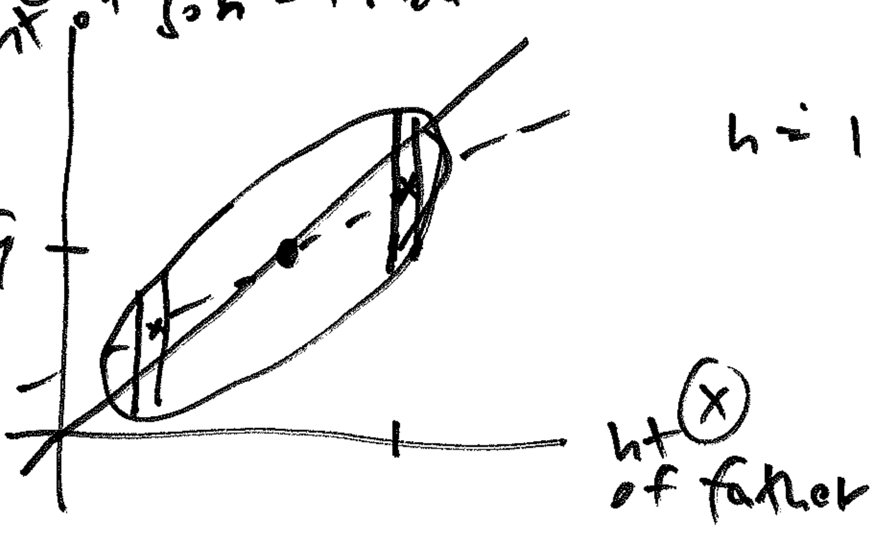
D) extra office hour wed (tomorrow) 4-5 pm

L-216 | y = tail length | x = wing length



best line for predicting y from x?

Gauss (~1800)



Francis Galton (1890)

$n \doteq 1050$

ht of son	ht of father
y	x
$\bar{y}$ s_y	$\bar{x}$ s_x

$n \doteq 1{,}050$

2

R - 25 (17)

slope of regression line for predicting y from x $= \hat{\beta}_1 = r \cdot \frac{s_y}{s_x}$

equation of regression line:

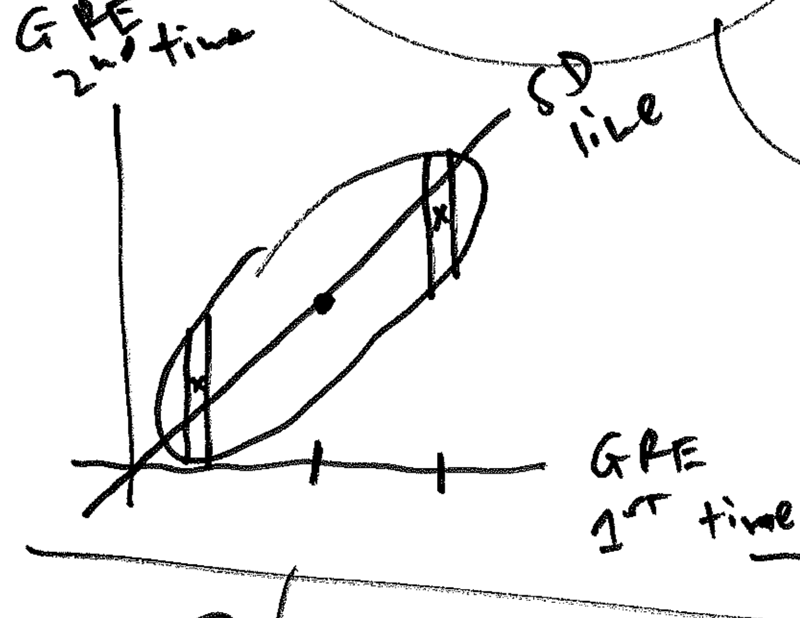
$$\hat{y} = \hat{\beta}_0 + \hat{\beta}_1 x$$

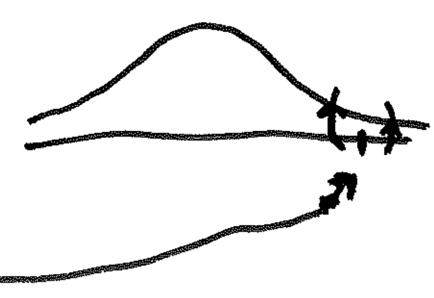
predicted y value

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

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

R - 25 (21)





L - 217 | L - 248 inference about

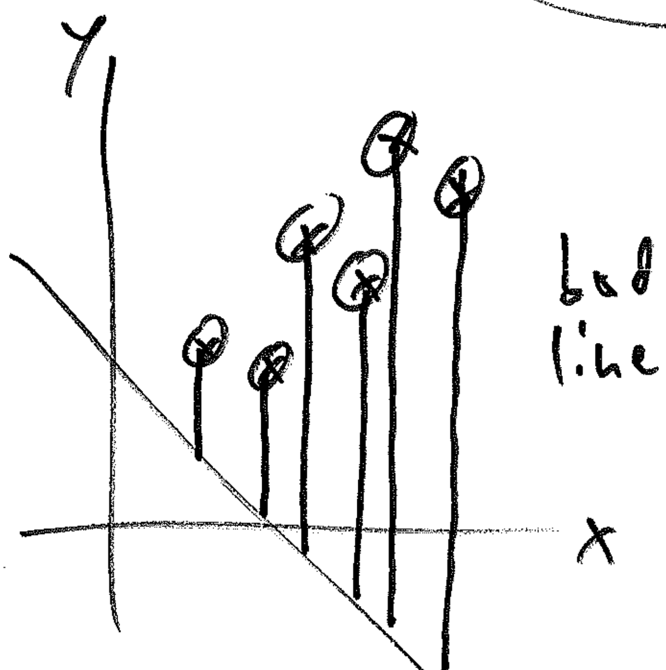
R - 25 (18), (19) / (12.44) slope

5050

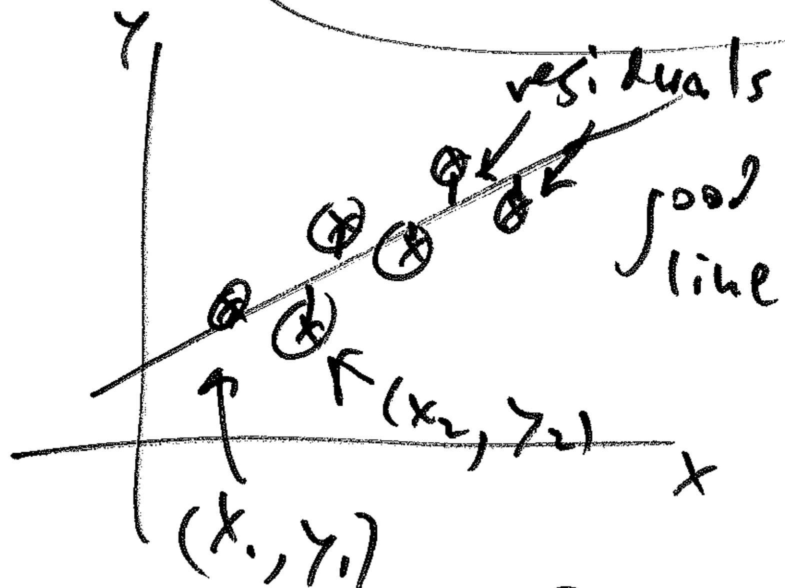
Gauss

1	100	101
2	99	101
3	98	101
⋮	⋮	
50	51	101

$\} 50$



Laplace 1.1



$$\begin{pmatrix} y_1 & x_1 \\ \vdots & \vdots \\ y_i & x_i \\ \vdots & \vdots \\ y_n & x_n \end{pmatrix}$$

Gauss

$$\frac{1}{n} \sum_{i=1}^{n} \Big[y_i - \underbrace{(\hat{\beta}_0 + \hat{\beta}_1 x_i)}_{\hat{y}_i} \Big]^2$$

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

result is the <u>least-squares estimates</u>