this regression time:

next time: ANOVA

read: LN pp. L-245 + 274

AMS 7
2 June 17

(1)

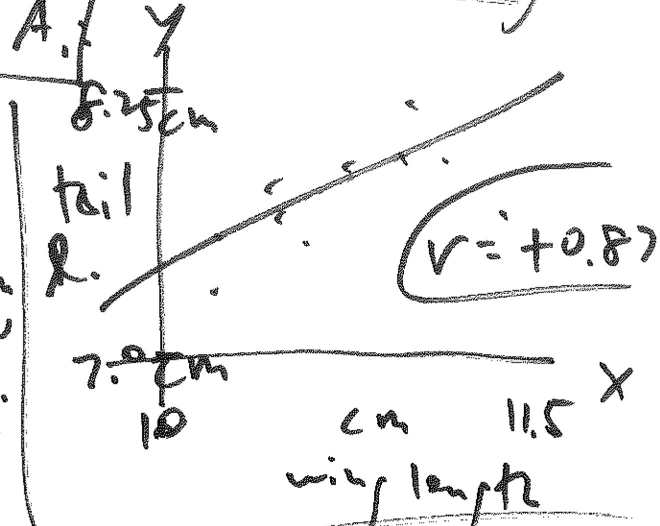
$\boxed{Q_1:}$ When is a regression slope $\hat{\beta}_1$ large in practical terms (practsig)?

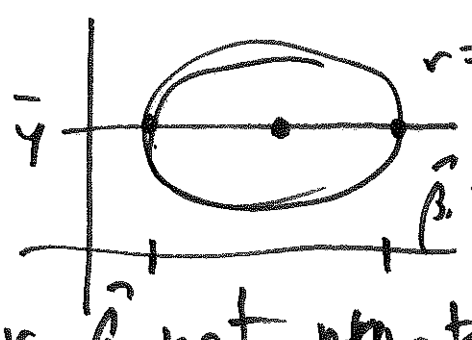
$\boxed{A_1:}$ Same method as in answering the question

$\boxed{Q_2:}$ when is a correlation coefficient (r) large in practical terms?

$\boxed{A_2 = A_1}$

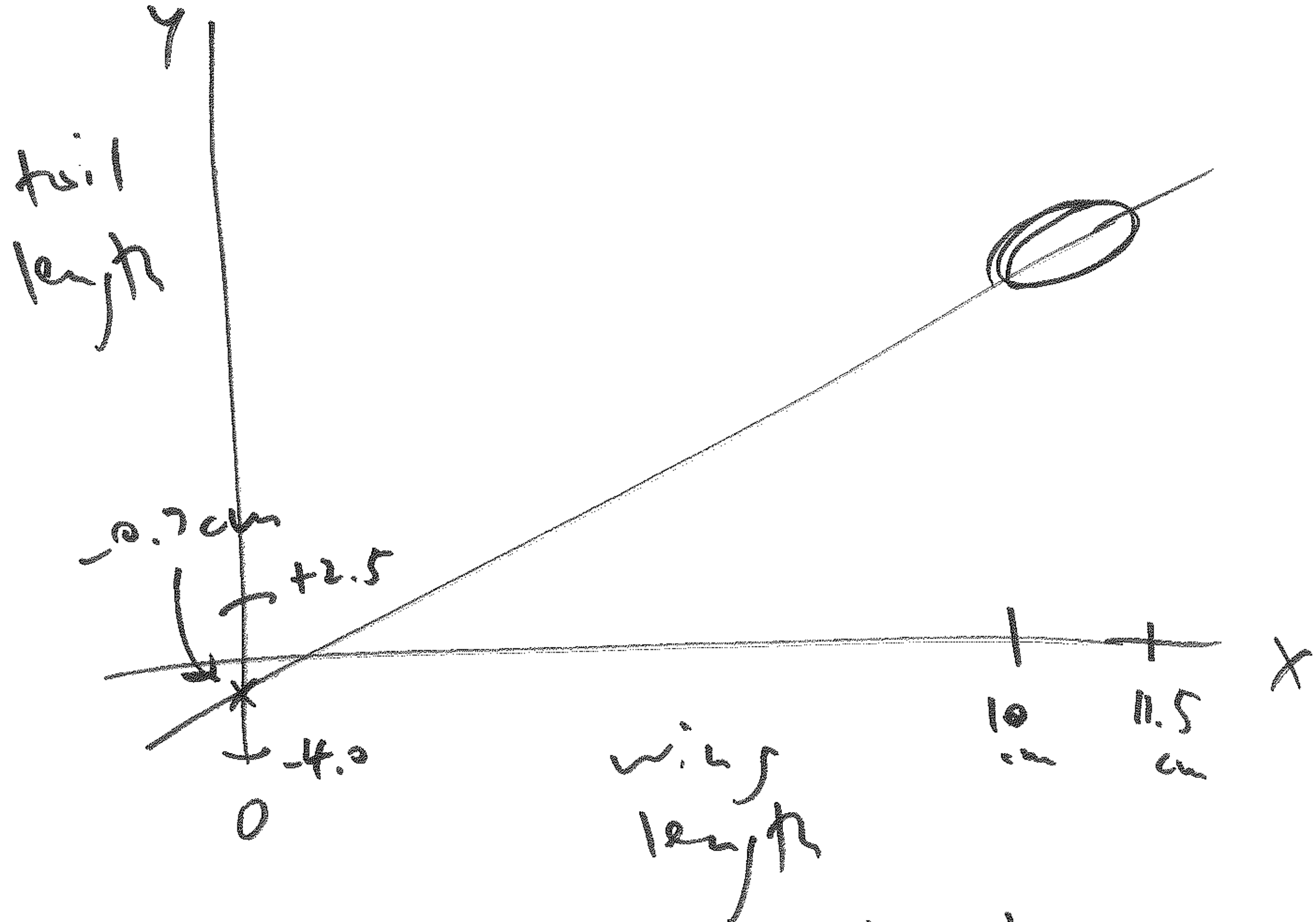
sage sparrows: smallest birds have x = wing length $\doteq$ 10cm, $\hat{y}$ = predicted tail length $\doteq$ 7.0 cm; largest have x = 11.5 cm, $\hat{y} \doteq$ 8.25cm; since 8.25cm differs from 7.0cm by an amount that's practsig, the corr. (r) & slope $\hat{\beta}_1$ associated with predicting y from x are also practsig





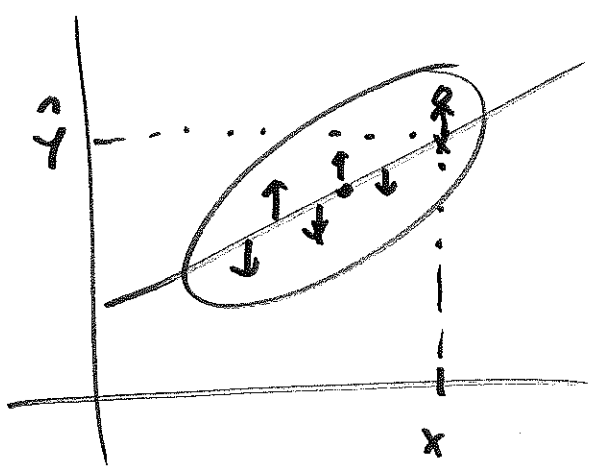
r, $\hat{\beta}_1$ not practsig

How to judge whether $\hat{\beta}_1$ is statsig?



often, y-intercept involves huge extrapolation away from data; it may therefore be meaningless

---



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

If x available,

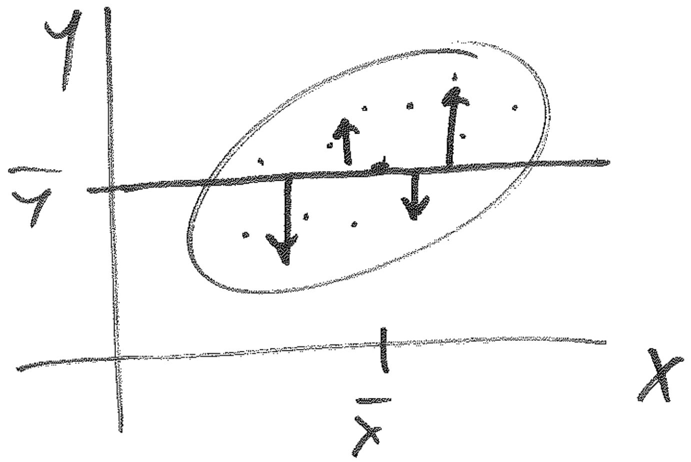
$\hat{y}_x = \hat{\beta}_0 + \hat{\beta}_1 x$

and $\hat{SE}(\hat{y}_x) \doteq s_{y|x}$

SD(residuals)
= RMSE

method (2)

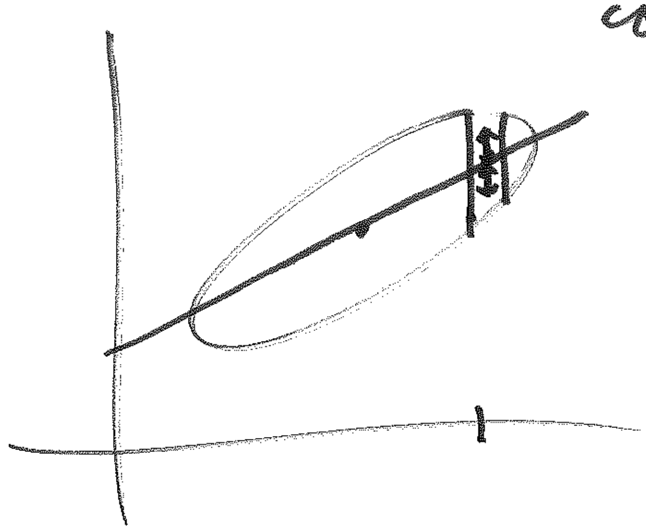
(3)



case (A) Don't have x or choose to ignore x:

best $\hat{y} = \bar{y}$

& $SE(\hat{y}) = S_y$



case (B) use x to predict y:

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

& $\widehat{SE}(\hat{y}) = S_{y|x}$

$\doteq S_y \sqrt{1-r^2}$

Regression is successful if $S_y\sqrt{1-r^2}$ is smaller than S_y by an amount that's meaningful

One way to quantify this is

$$\frac{\overset{\text{ignore } x}{s_y} \;-\; \overset{\text{use } x}{s_y\sqrt{1-r^2}}}{s_y \;\; \text{ignore } x} = 1 - \sqrt{1-r^2}$$

better measure of success of regression than r^2