AMS 7 - Lecture 5/22/18

1

THIS TIME: 2-independent samples

NEXT TIME: Correlation

→ Read LN pg L-214 → L-268

Today LN pg L-281

- Dichotomous Outcomes (redwood example)
  - 2 samples. CA & OR (2 diff designs)
    Independent: 281 vs. 265 sample sizes

* like daphnia case study, except dichotomous

* The OR rate of infection is more than twice as high ($\hat{p}_1 = 3.4\%$ for CA & $\hat{p}_2 = 7.1\%$ for OR)

$$SE_{IID}(\hat{p}_1) = \frac{\sigma_1}{\sqrt{n_1}} = \frac{\sqrt{p_1(1-p_1)}}{\sqrt{n_1}} = \sqrt{\frac{p_1(1-p_1)}{n_1}}$$

$$\hat{SE}_{IID}(\hat{p}_1) = \sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1}} \doteq 1.1\%$$

Larger than we might like b/c dichotomous is causing more uncertainty

$$\hat{SE}_{IID}(\hat{p}_2) = \sqrt{\frac{\hat{p}_2(1-\hat{p}_2)}{n_2}} \doteq 1.5\%$$

$$\hat{SE}_{IID\ indep.}(\hat{p}_1 - \hat{p}_2) = ?$$

$1.5 \le \text{Hypotenus} \le 3.0$

$\hat{SE}(\hat{p}_2) = 1.5$

Shortcut formula

$$\hat{SE}_{IID\ indep.}(\hat{p}_1 - \hat{p}_2) = \sqrt{\left(\sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1}}\right)^2 + \left(\sqrt{\frac{\hat{p}_2(1-\hat{p}_2)}{n_2}}\right)^2}$$

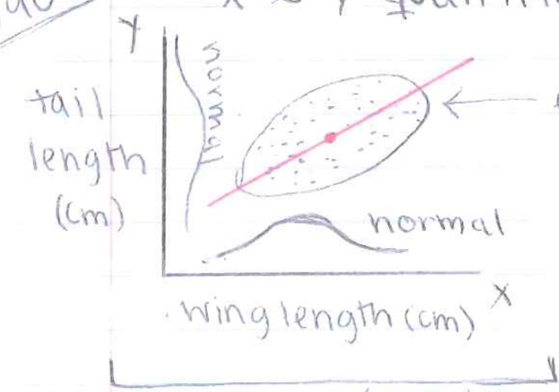
$$= \sqrt{\frac{\hat{p}_1(1-\hat{p}_1)}{n_1} + \frac{\hat{p}_2(1-\hat{p}_2)}{n_2}}$$

5/22/18

(2)

L-214 • Simple Correlation & simple Linear Regression
Pg. L-221 (Bird's Tail & Wing length)
* Studying the relationship b/w 2 variables
X & Y quantitative continous

1890's



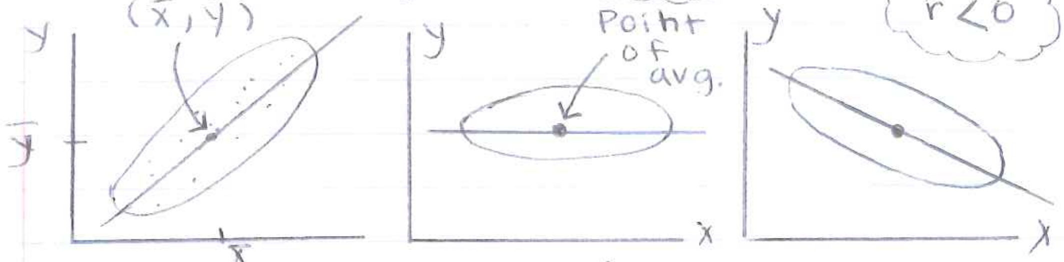
ellipse

"two variables"

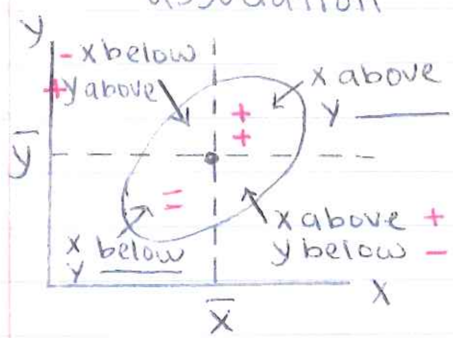
summary of bivariate normal distribution

* works of Karl pearson & Francis Galton

Scatterplot / scatter diagram



positive association | no (linear) association | negative association



$\frac{x-\bar{x}}{S_x} > 0$ & $\frac{y-\bar{y}}{S_y} > 0$

$\frac{x-\bar{x}}{S_x} < 0$ & $\frac{y-\bar{y}}{S_y} < 0$

$$n \begin{array}{cc} y & x \end{array} \left[\begin{array}{c|c} y_1 & x_1 \\ y_2 & x_2 \\ \vdots & \vdots \end{array}\right]$$

mean $\bar{y}$ & $\bar{x}$

SD S_y & S_x

$$\frac{1}{n}\sum_{i=1}^{n}\left(\frac{x_i-\bar{x}}{S_x *}\right)\cdot\left(\frac{y_i-\bar{y}}{S_y *}\right)$$

$\hookrightarrow = r$

$S_x^* \sqrt{\frac{1}{n}\sum_{i=1}^{n}(x_i-\bar{x})^2}$

$S_y^* \sqrt{\frac{1}{n}\sum_{i=1}^{n}(y_i-\bar{y})^2}$

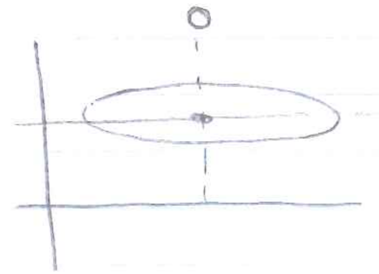
5/22/18

3

* r = Sample correlation (coefficient) b/w x & y

→ FACTS ABOUT (r)

$-1 \le r \le +1$



-1 ↑ perfect linearity with (-) slope

+1 ↑ perfect linearity with (+) slope

R-73