AMS 7 24 Feb 09 ①

This time: 2-sample problems; next time: correlation & regression

lab 5 due by 5pm Fri 27 Feb

hwk 4 due Thu 5 Mar

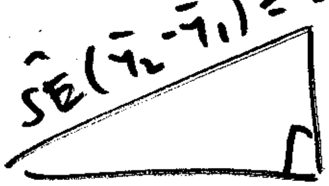
correction: lab 6 due by 5pm Fri 6 Mar

today: L-197 → L-220

long run hist of $(\bar{y}_2 - \bar{y}_1)$

$$\bar{y}_2 - \bar{y}_1 = \bar{y}_2 + (-\bar{y}_1)$$

main fact: $\hat{SE}(\bar{y}_1)$ & $\hat{SE}(\bar{y}_2)$ combine in calculating $\hat{SE}(\bar{y}_2 - \bar{y}_1)$ like edges of a right triangle:

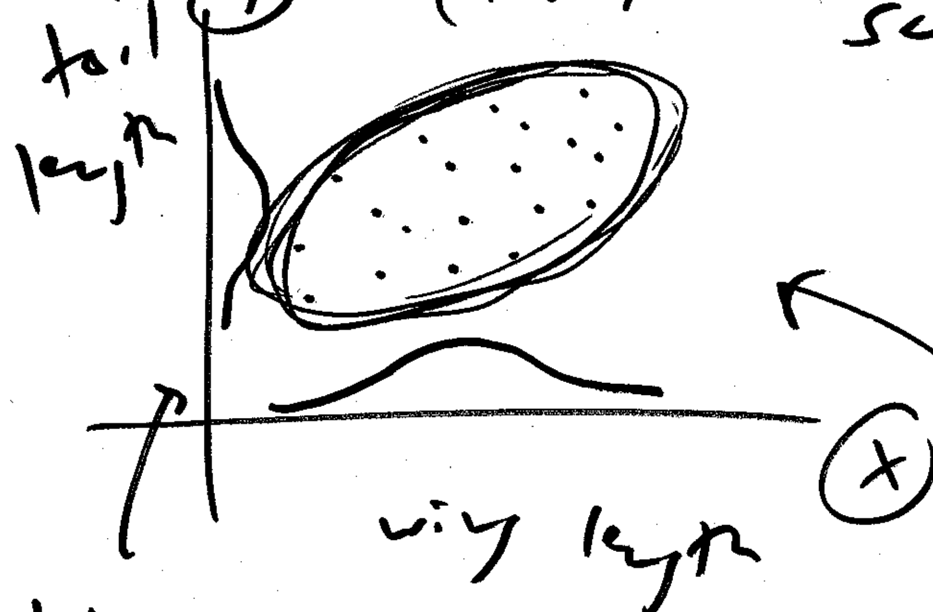


$$\hat{SE}(\bar{y}_2 - \bar{y}_1) = \sqrt{\left[\hat{SE}(\bar{y}_1)\right]^2 + \left[\hat{SE}(\bar{y}_2)\right]^2}$$

$$= \sqrt{\left(\frac{s_1}{\sqrt{n_1}}\right)^2 + \left(\frac{s_2}{\sqrt{n_2}}\right)^2}$$

$$= \sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}$$

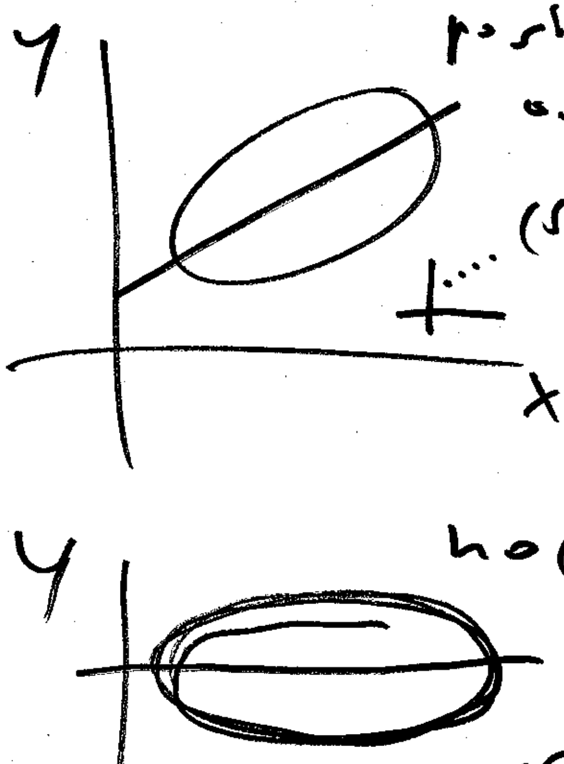
(1880)



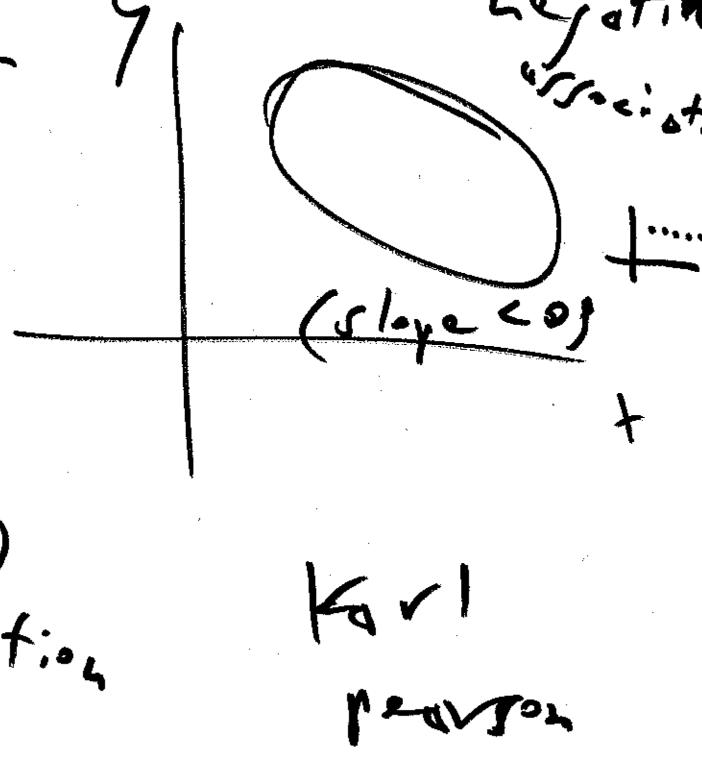
scatter plot (scatter diagram)

elliptical scatterplot (football-shaped)

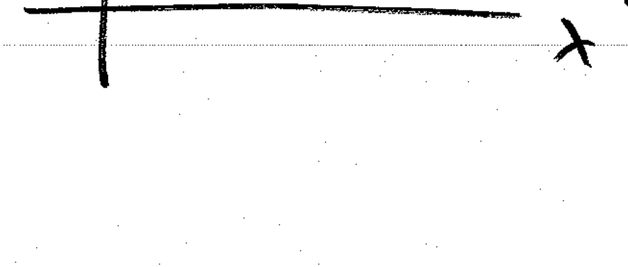
bivariate normal distribution



positive association (slope > 0)

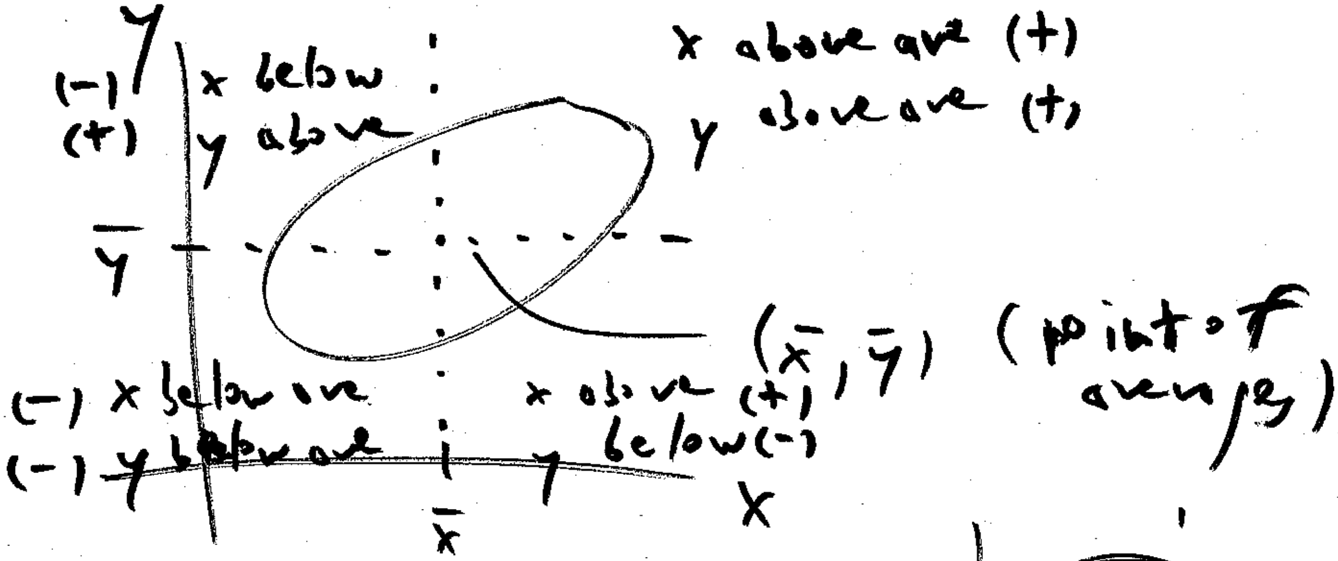


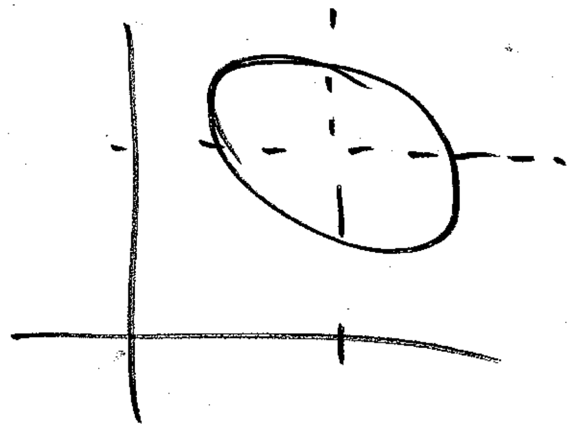
negative association (slope < 0)



no (linear) association (slope = 0)

Karl Pearson (1880s)





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

1	y_1	x_1
2	y_2	x_2
i	y_i	x_i
n	y_n	x_n

$= r$ (~~product-moment~~)

correlation coefficient