AMS 7
26 May 17

read: LN pp. L-244 + L-268

this time: correlation & regression

next time: you can begin working on homework 4 now

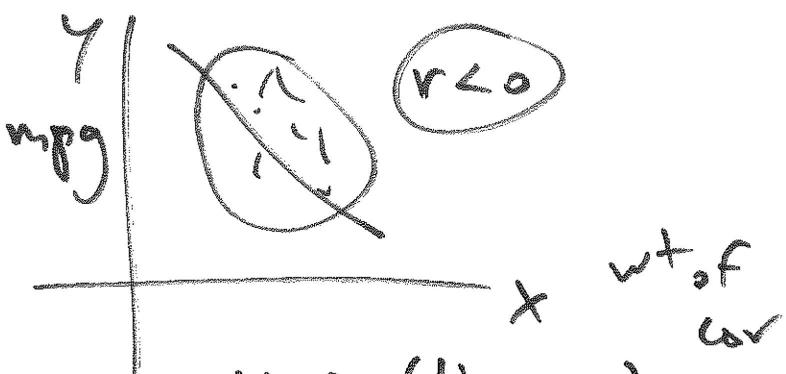
homework 3 due 11.59 pm tonight at canvas.ucsc.edu

1

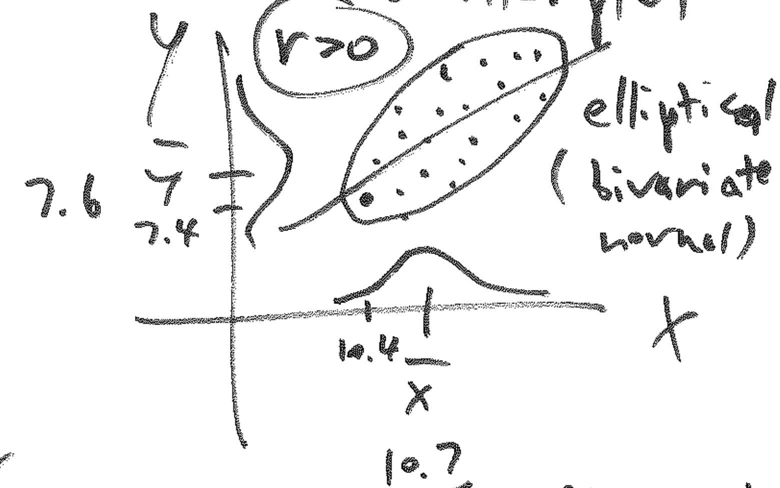
no class or disc. sec Mon; Mon disc. sec people need to go to section Tu-Fri next week

probable due date: Fri 2 Jun

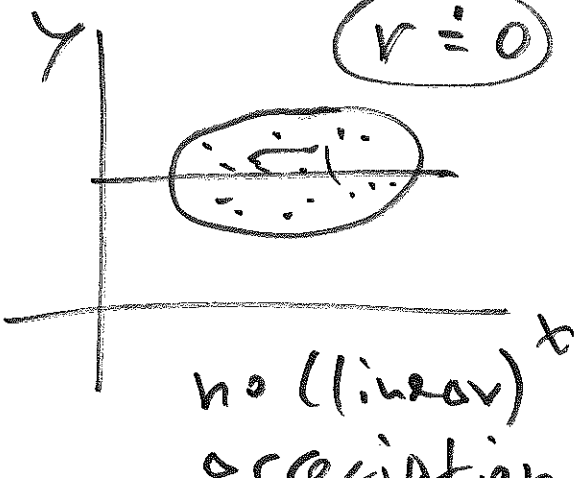
scatterplot



negative (linear) association



positive (linear) association

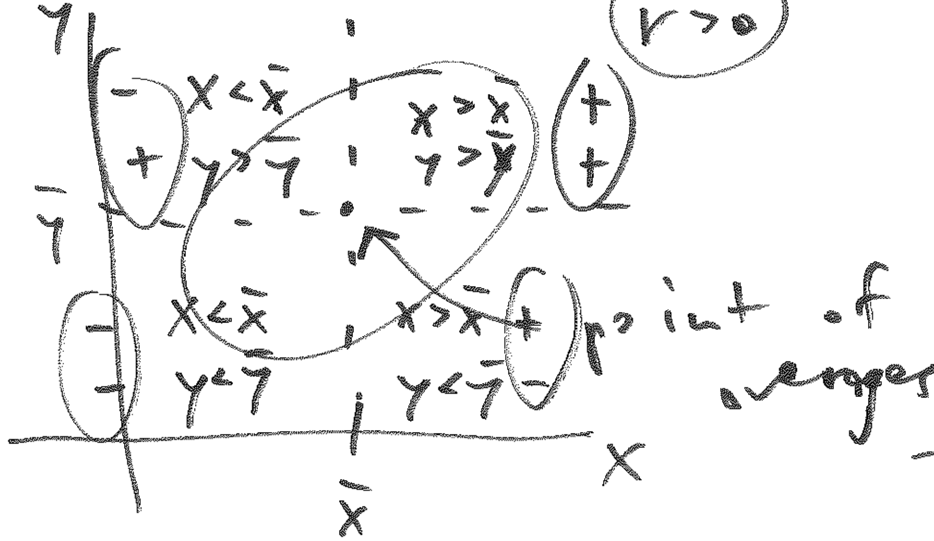


no (linear) association

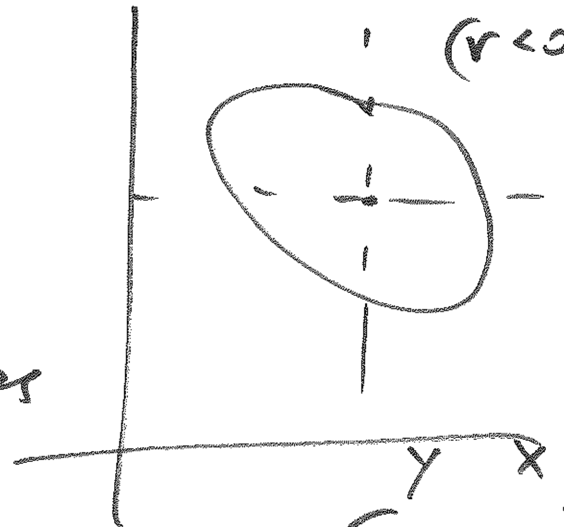
Karl Pearson & Francis Galton uk (1885)

r = (Pearson's product-moment) correlation coefficient

$(r > 0)$



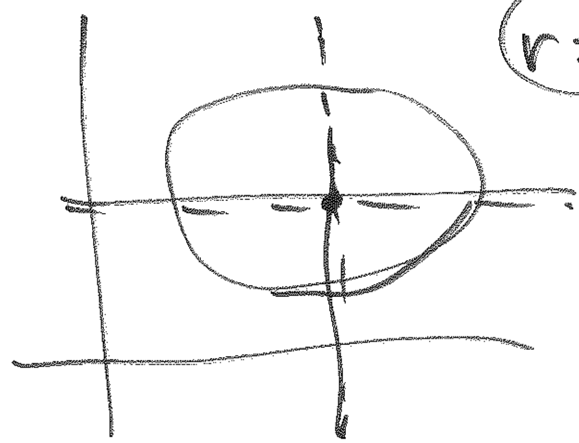
$(r < 0)$



	Y	X
	Y_1	X_1
	Y_i	X_i
	Y_n	X_n
mean	$\bar{Y}$	$\bar{X}$
SD	s_Y	s_Y

n

$(r \doteq 0)$



$$\boxed{r = \frac{1}{n}\sum_{i=1}^{n}\left(\frac{X_i-\bar{X}}{s_X^*}\right)\cdot\left(\frac{Y_i-\bar{Y}}{s_Y^*}\right)}$$

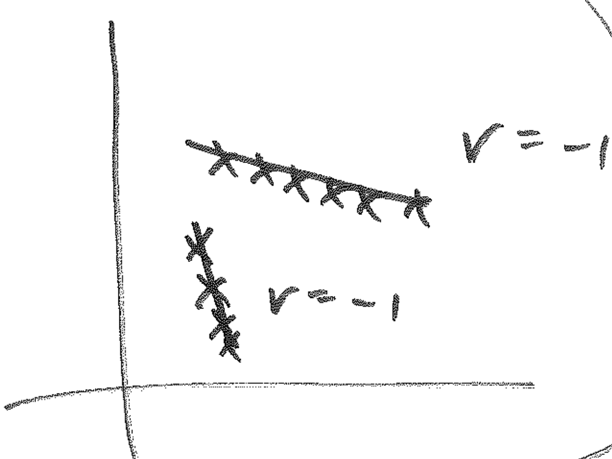
R-24 formula (15)

$$s_X^* = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(X_i-\bar{X})^2} \qquad s_Y^* = \sqrt{\frac{1}{n}\sum_{i=1}^{n}(Y_i-\bar{Y})^2}$$

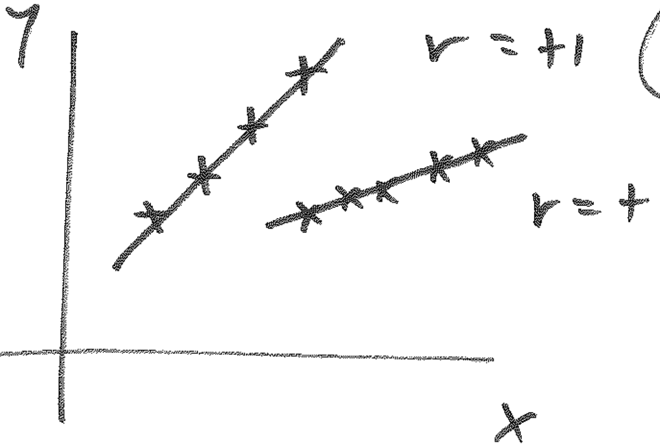
Facts About r

1. r is a pure number without units

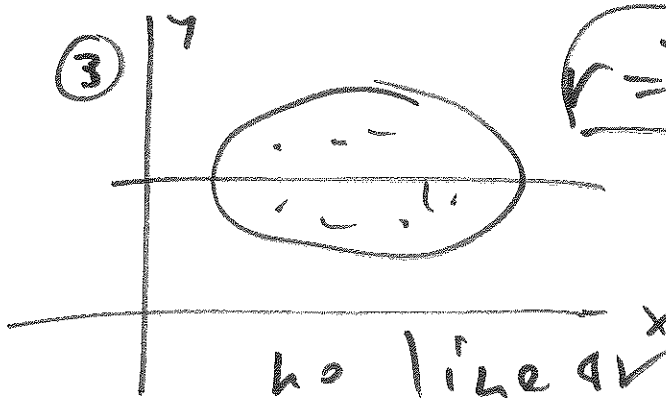
② $-1 \le r \le +1$



③



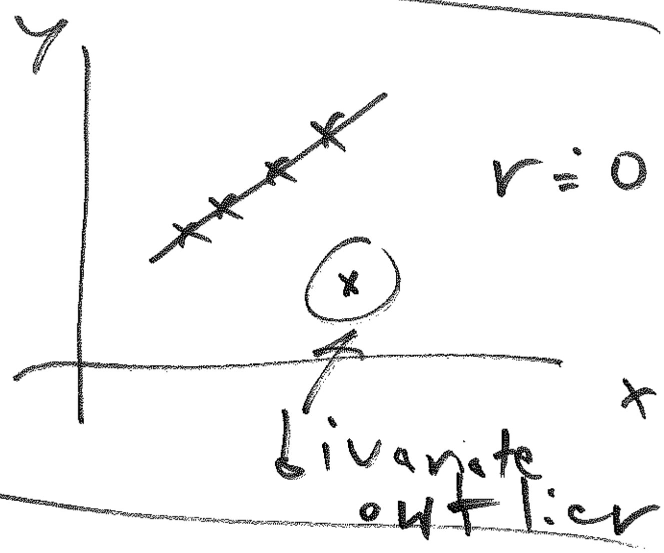
③

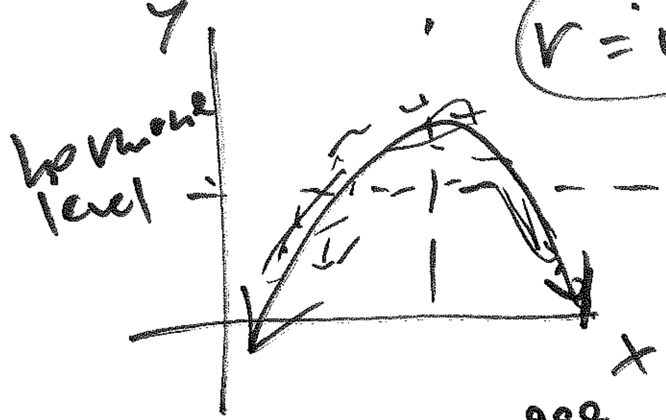


$r \doteq 0$

no <u>linear</u> association

r can be fooled by outliers and/or nonlinearity





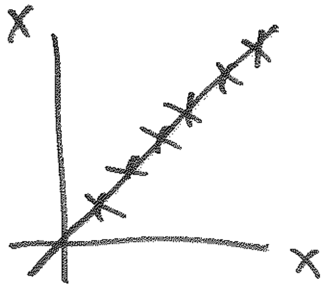
$r \doteq 0$

nonlinear relationship between

x & y

④ $r(x, x) = +1$ for any variable x ④

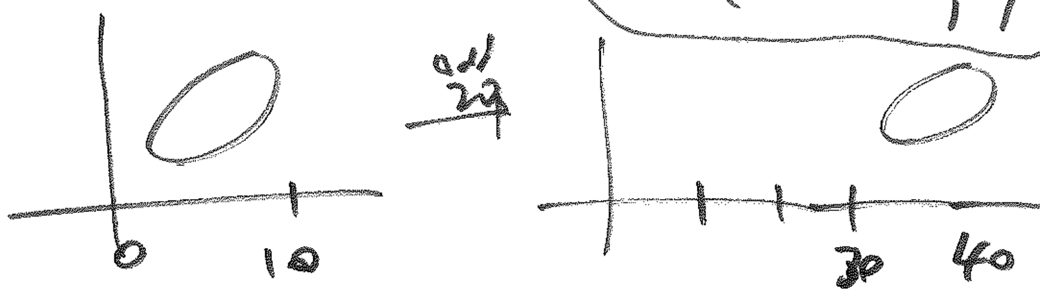
correlation between x and x



⑤ $r(x, y) = r(y, x)$

for all var. x & y

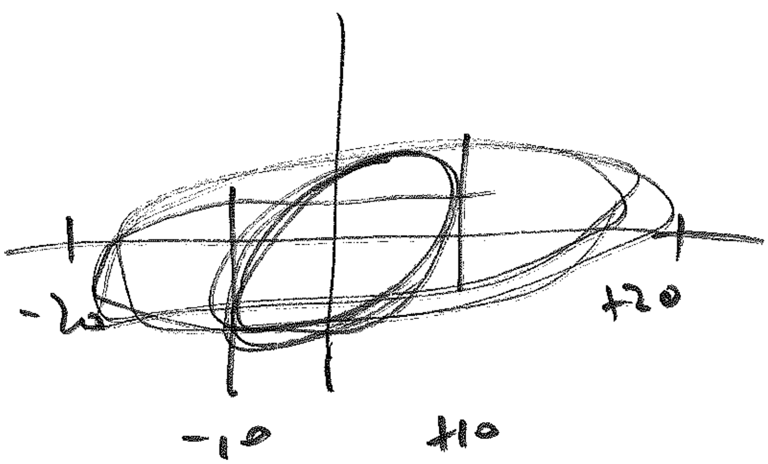
⑥ add c to all x values → $r(x, y) = r(x + c, y)$



$r(x, y) = r(x, y + c)$

mult. by $c > 0$

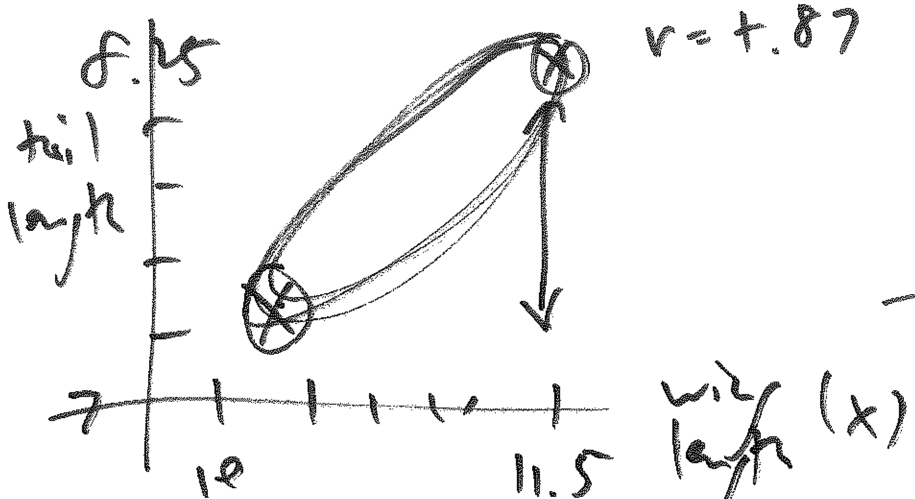
$r(x, y) = r(cx, y) = r(x, cy)$

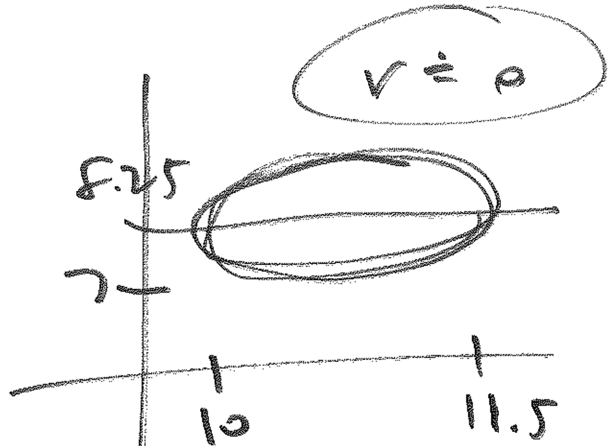


Q: Is r(wing length, tail length) = +0.87 large in practical terms?

A: Does 0.87 differ from 0 by an amount that matters (real-world)

null, boring value $r = 0$

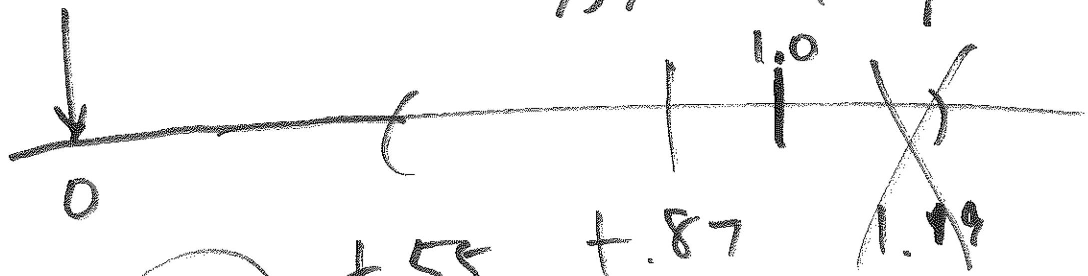




as we go from one sparrow with ($x = 10$, $y \doteq 7$) to ($x = 11.5$, $y \doteq 8.25$); 8.25 is much larger practically than 7; but if $r = 0$, ($x = 10$, $y = 7.6$), ($x = 11.5$, $y = 7.6$) ← no diff approx (large n) 95% CI for ρ:

$$r \pm 2\,\widehat{SE}(r) = r \pm 2\underbrace{\sqrt{\frac{1-r^2}{n-2}}}_{0.16}$$

95% CI for ρ



$(+.55, +1.0)$

highly shaky