Note Title 9/30/2009

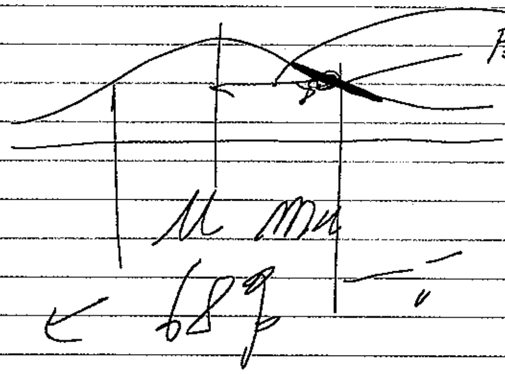
SST 200 9-30-09

NOTATION: TEXTBOOK N(MEAN, STD DEV) IQ N(100, 15)

MOST OF STAT N(MEAN, VARIANCE) IQ N(100, 225)

MORE VARIABLES — N((0, 2D), (Var 1st COV / COV Var 2nd))

RECALL BELL CURVE,



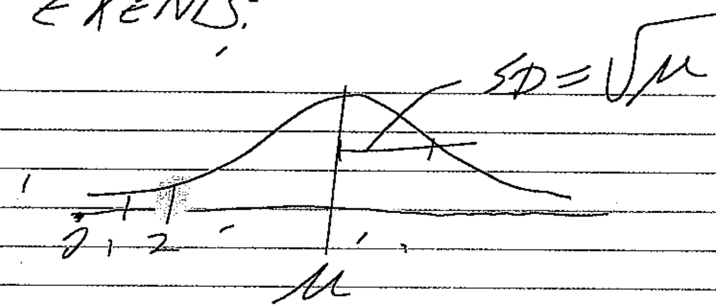
POINT OF σ SIGMA

95% 2σ

COUNTS OF RARE EVENTS.

$\approx$ NORMAL

POISSON



eg $\mu = 64$ FOR POISSON $\sigma = \sqrt{64} = 8$

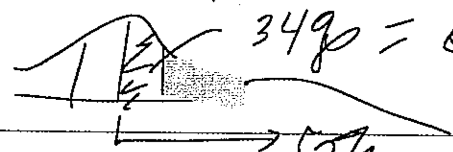
$\approx$ 68% CHANCE OF BETWEEN 64 ± 8

? $P(\text{GET MORE THAN } 72) \approx P\left(Z > \frac{72-64}{8}\right)$

STD SCORE

$z = \frac{72-64}{8} = 8/8 = 1.00$

RULE OF THUMB



34% = 68%

50%

$.5 - .34 = .16$

ANS.

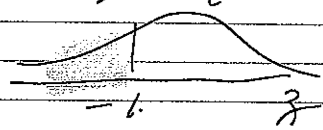
$P(Z > 1.00) = .16$

So $P(\text{GET} > 72) \approx P(Z > 1.00) = .16$

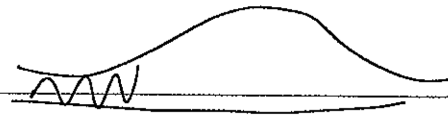
STD SCORE of 72

OR USE TABLE $P(Z > 1.00) = P(Z < -1.00)$

SAME AS



~~$z = -1.00$~~



$z = -1.00$

$.00$

$\boxed{0.1587}$ -1.0

KNOW z-TABLE

✓ KNOW POISSON $\approx$ NORMAL SD $\sigma = \sqrt{\mu}$

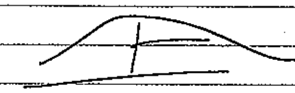
MEAN

PROVISO: $\mu \geq 3$ RULE OF THUMB FOR

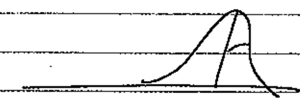
NORMAL APPROX OF POISSON.

A VERY IMPORTANT PROPERTY OF NORMAL DISTN

IF X IS NORMAL μ_x, σ_x



Y IS NORMAL μ_y, σ_y



SUPPOSE THESE ARE STATISTICALLY UNRELATED

("INDEPENDENT")

$\Rightarrow$ $X+Y$ IS NORMAL

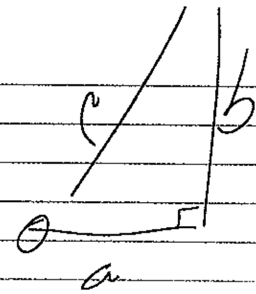
$$\mu_{x+y} = \mu_x + \mu_y$$

VARIANCE $\sigma^2_{x+y} = \sigma_x^2 + \sigma_y^2$

PYTHAGORAS' THEOREM

LIKEWISE

$$\sigma_{X+Y} = \sqrt{\sigma_X^2 + \sigma_Y^2}.$$



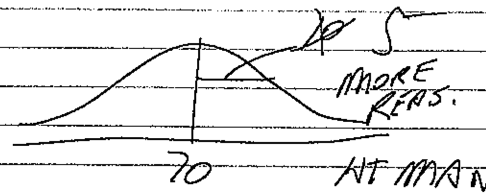
$$a^2+b^2=c^2$$

$$c=\sqrt{a^2+b^2}$$

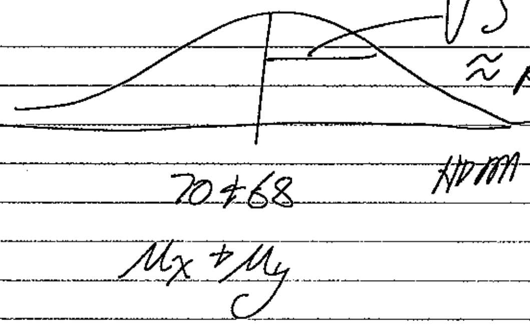
QUES. SUPPOSE HEIGHTS OF MEN

$N(70, 5)$ TEXT NOTATION

WOMEN $N(68, 4)$ TEXT NOTATION



PICK ONE OF EACH AT RANDOM + TAKE COMBINED HEIGHT



$\sqrt{5^2 + 4^2}$

$\approx$ ROUGHLY 6.5

HDMM + HTW

$\sqrt{41} \approx (6.5)^2$

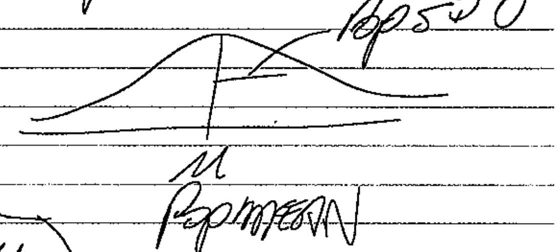
THE CONCLUSION IS

$\mu_{\text{TOTAL OF INDEP}} = \text{TOTAL OF } \mu_i$

$\sigma_{\text{TOTAL OF INDEP}} = \sqrt{\text{TOTAL OF } \sigma_i^2}$ GROWS SLOWLY WITH # OF VENTURES.

REVIEW OF IMPORTANT THINGS CH 6.

* NOTION OF SAMPLE S.D. + POP SD.
* NOTION OF NORMAL CURVE

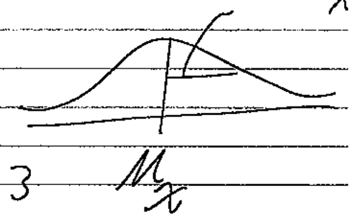


* ALL NORMAL DIST^{N}s ALIKE IN σ UNITS FROM μ
* $P(X < c) = P\left(Z < \frac{c - \mu_X}{\sigma_X}\right)$

  ↑ TABLE — STD SCORE OF C

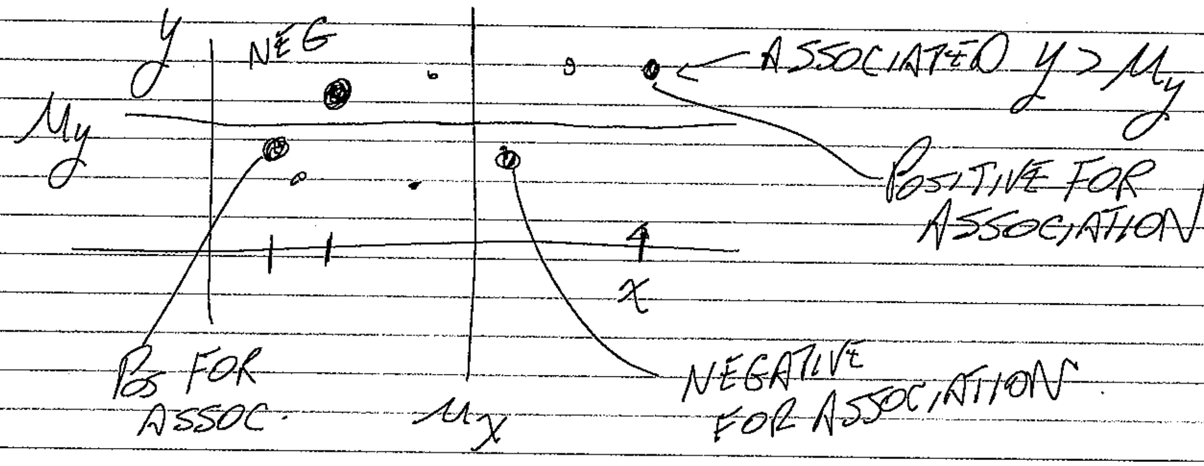
* SPECIAL CASE DIST OF COUNTS OF RARE EVENTS — POISSON ≈ $\mu_X \geq 3$

$\sigma_X = \sqrt{\mu_X}$



Ch. 7. TREATMENT OF TWO VARIABLES

"CO-RELATION" → CORRELATION



WANT DIMENSIONLESS MEASURE OF POS ASSOCIATION.

$\cancel{(x, y)}$ $\left(\frac{x-\mu_x}{\sigma_x}, \frac{y-\mu_y}{\sigma_y}\right)$

z_x z_y

PRODUCT OF THESE IS

$-$ $+$

0 $-$

$+$

$\frac{\sum z_x z_y}{n}$ CORRELATION

(x_i, y_i)

OR (SAME) TEXT $\sum \left(\frac{x_i-\bar{x}}{s_x}\right)\left(\frac{y_i-\bar{y}}{s_y}\right) / (n-1)$

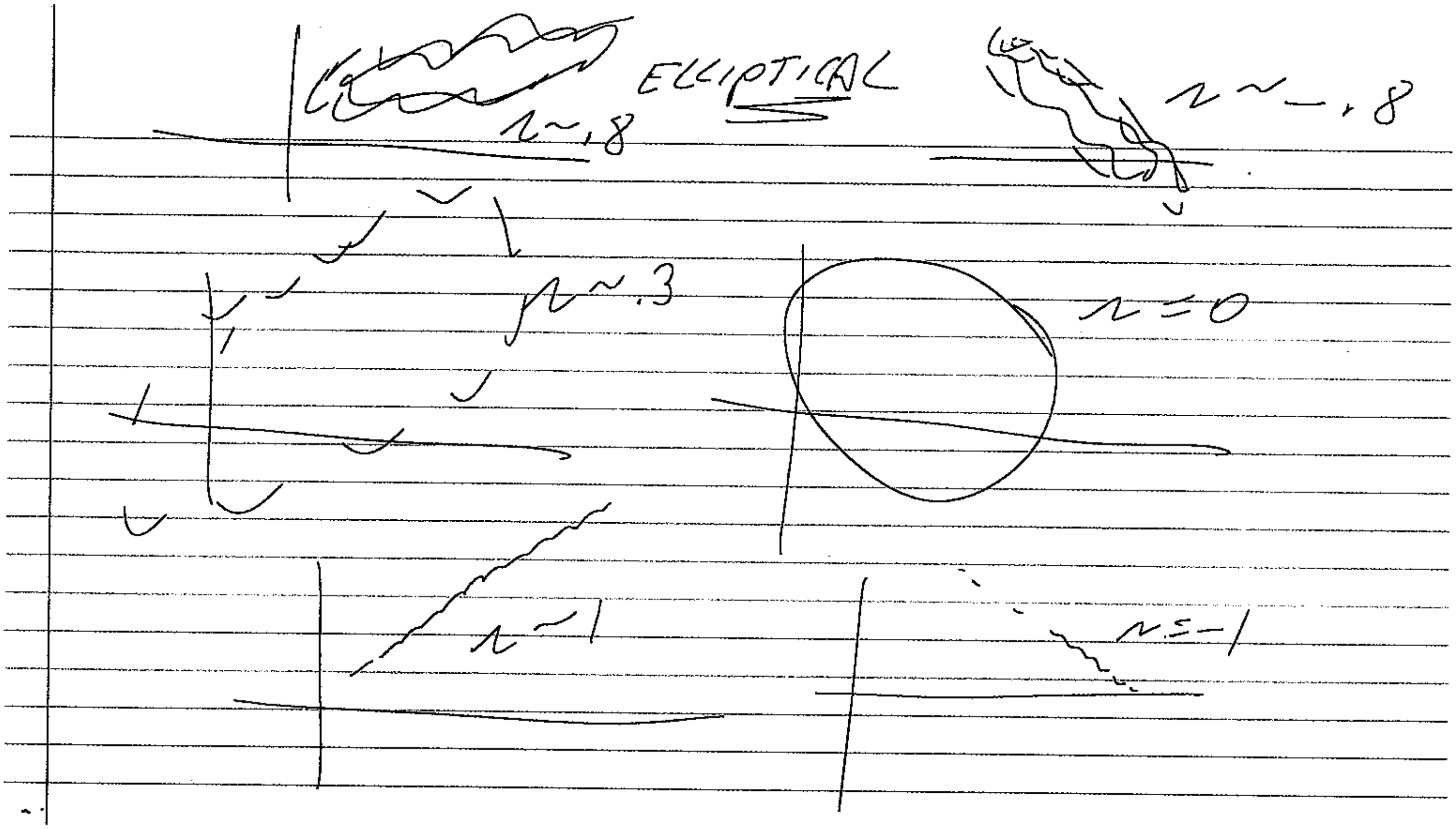
THE CORRELATION DENOTED r, ρ RHO

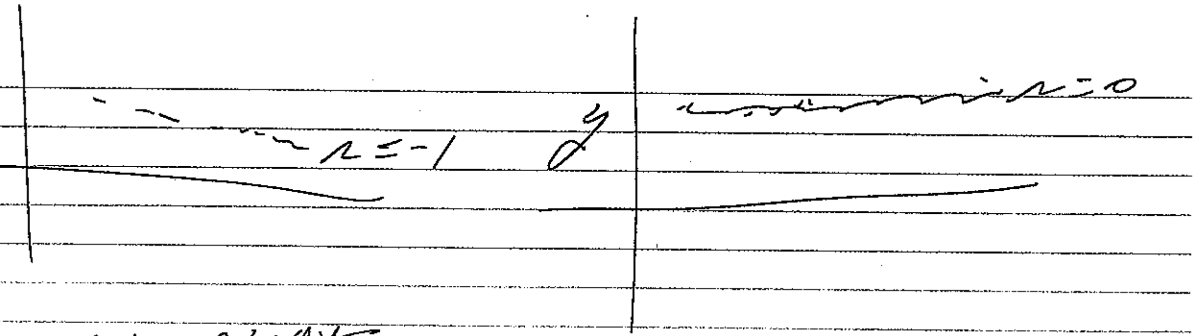
$$|r| \leq 1$$

r IS A MEASURE OF LINEAR ASSOCIATION OF SCORES (x, y).

WHAT CAN YOU DO WITH CORRELATION?

LOOK AT PICTURES — APPEARANCE OF BI-VARIATE NORMAL DATA


ELLIPTICAL
r~.8
r~-.8
r~.3
r=0
r~1
r=-1

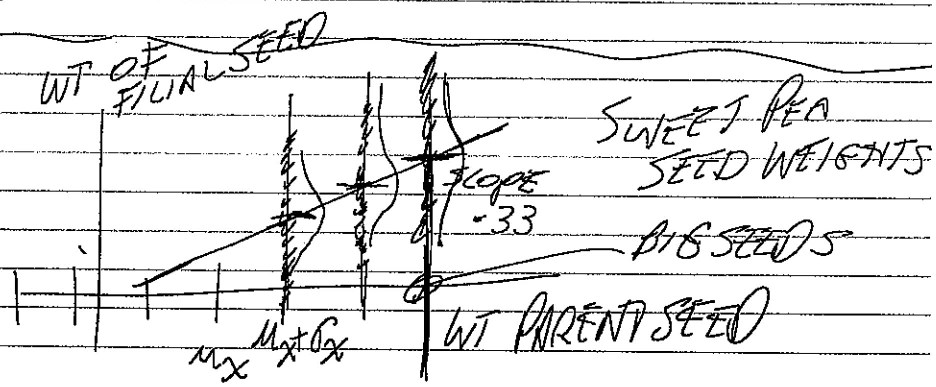


$-1 \le r \le 1$ **ALWAYS**

GALTON 1877

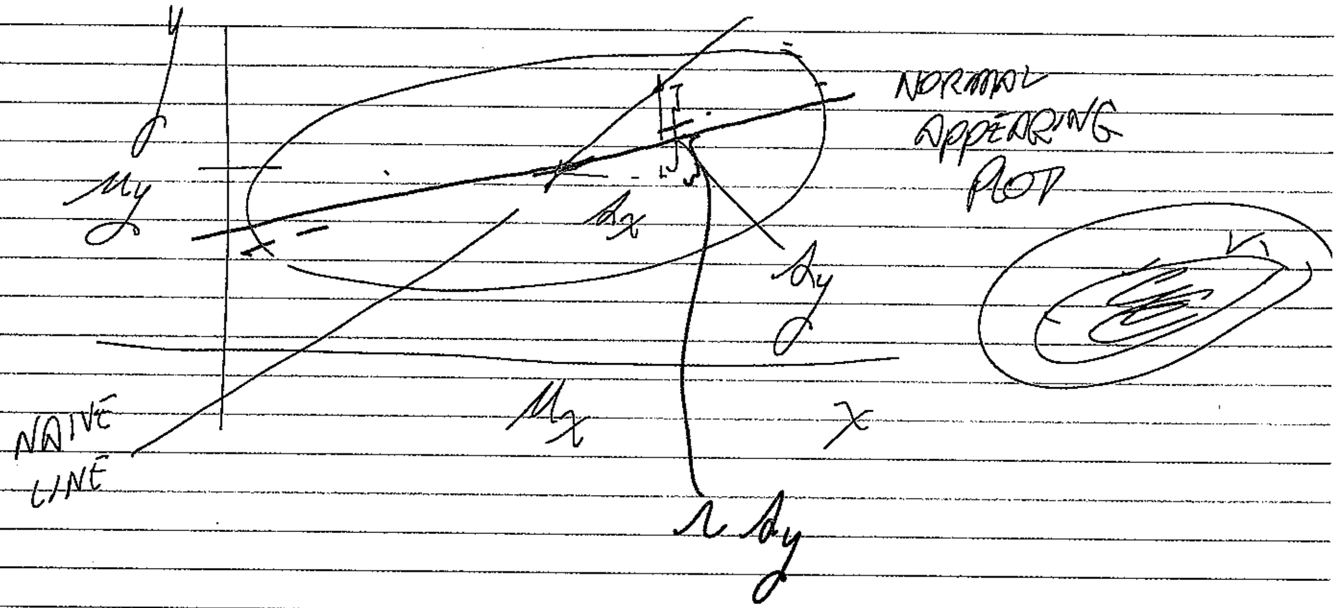
CO-RELATION $r = .33$

LATER CORRELATION

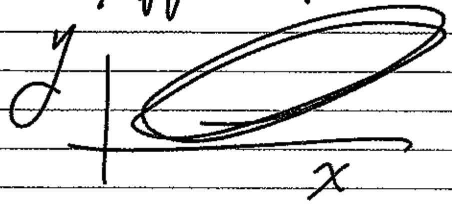


HUGE - SAYS THAT FOR BIVARIATE NORMAL

~~DATA (AS GALTON HAD) THE FOLLOWING ARE TRUE~~



APPLICATION: x = EXAM 1 SCORE



y = " 2 "

SUPPOSE $r = 0.9$ (SAY)

THEN IF I SCORED $1\,\sigma_x$ ABOVE μ_x (EXAM 1)

"PREDICT" $r\,1\,\sigma_y = .9\,\sigma_y$ ABOVE μ_y (EXAM 2)