STT 200 4-1-09a

RANDOM VARIABLES

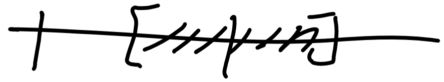
CH17: LAST SECTION (COVARIANCE & CORRELATION FOR r.v.)

CH 18: ALL IT IS ABOUT THE NOTIONS OF

(i) SAMPLING DISTRIBUTION (OF $\hat{p}$, OF $\bar{x}$)

(ii) CENTRAL LIMIT THEOREM (CLT)

CONSIDER THE ABOVE AS THEORY BEHIND CI (TESTS (TO BE SEEN))



$\bar{x}$ CHOL MSU U GRADS

SUPPOSE THIS IS AVG CHOL LEVEL AT ADMISSION

$$\bar{x} \pm 1.96 \frac{\sigma}{\sqrt{n}} \quad [\quad]$$

BIG NEWS!

~~CH 7~~ (END) CORRELATION FOR R.V.

RECALL LIST (x_i, y_i)

i	x_i	y_i
1	3	7
2	6	11
3	2	5

SAMPLE CORRELATION

$$r = (\rho \text{ RHO}) = \frac{\overline{xy} - \bar{x}\bar{y}}{\sqrt{\overline{x^2} - \bar{x}^2}\sqrt{\overline{y^2} - \bar{y}^2}}$$

DATA

RANDOM VARIABLES

X, Y X = YOUR RET^N ON ROULETTE

Y = YOUR FRIEND'S RETURN ON SAME PLAY

POSS OUTCOMES

x	y	$p(x,y)$	$xy\,p(x,y)$
2	7	.2 (JAY)	2.8
6	4	.7	16.8
3	8	.1	2.4
		1	$E(XY) = 22$

CORRELATION(X,Y)

$$\overset{\text{DEF}}{=} \frac{E(XY) - (EX)(EY)}{\sqrt{EX^2 - (EX)^2}\sqrt{EY^2 - (EY)^2}}$$

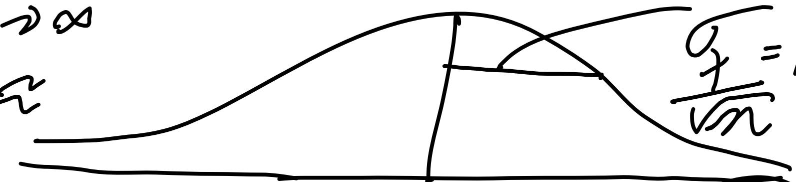
IMPLICATIONS OF $\rho_{x,y}$, σ_x, σ_y FOR r.v.'s

(FOR DATA WE'D USE ESTIMATES $\rho_{x,y}$, $\hat{\sigma}_x = \sqrt{\overline{x^2} - \bar{x}^2}$
$\hat{\sigma}_y = \sqrt{\overline{y^2} - \bar{y}^2}$)

RECALL $\bar{X}$ (FROM WITH-REPL SAMPLE OF n)

→ $n \to \infty$

~



$\frac{\sigma_x}{\sqrt{n}} = \frac{\sqrt{EX^2 - (EX)^2}}{\sqrt{n}}$

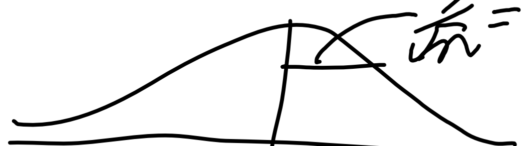
EX $\bar{x}$ POSS AVG YOU MIGHT GET

PR QUANTITY

LIKEWISE FOR Y

~

DIST OF $\bar{Y}$



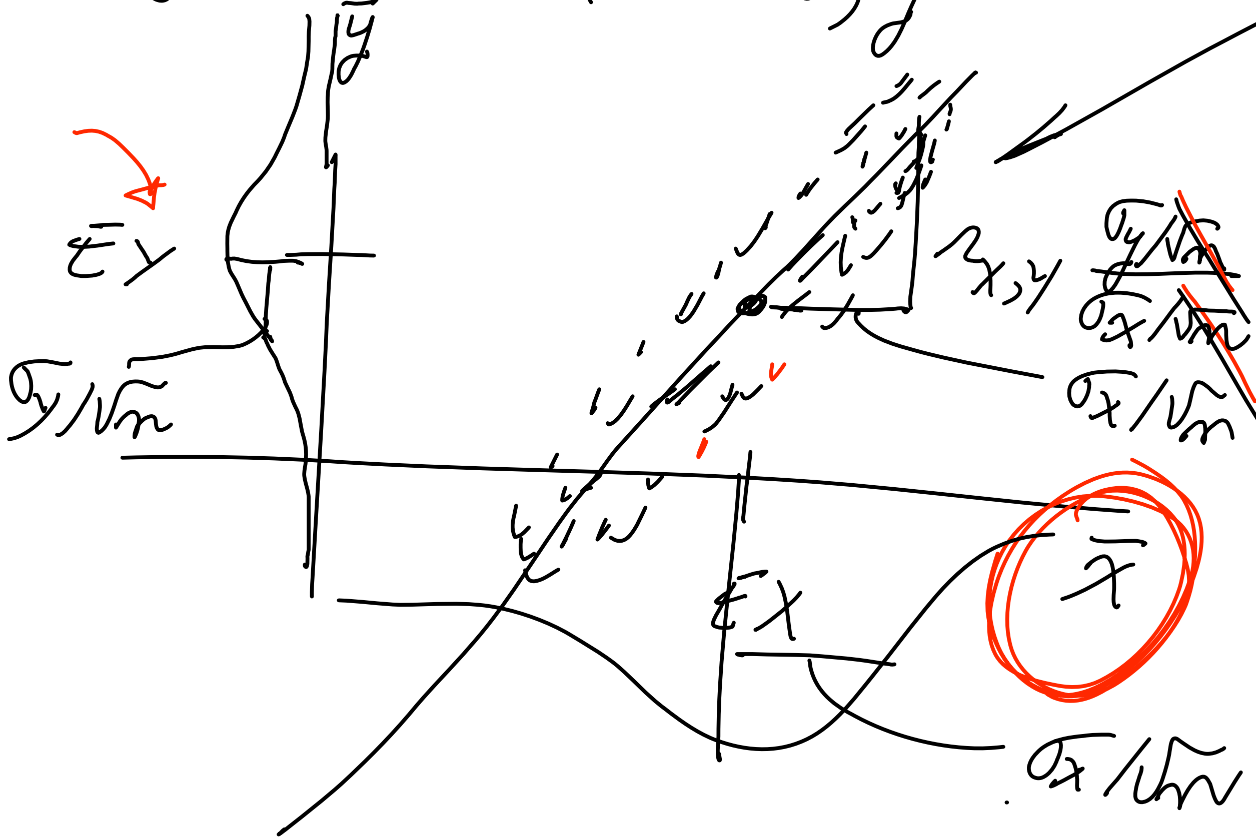
$\frac{\sigma_y}{\sqrt{n}} = \frac{\sqrt{EY^2 - (EY)^2}}{\sqrt{n}}$

AVG (NOT A SINGLE PLAY)

EY $\bar{y}$ POSS $\bar{y}$ COMPANION MAY GET

JOINT BEHAVIOR OF $\bar{X}, \bar{Y}$?

CASE OF $r_{X,Y} > 0$



WHERE $(\bar{X}, \bar{Y})$ ARE LIKELY TO FALL

SAME REGR LINE AS FOR $n=1$

OUR $\bar{X}, \bar{Y}$ MAY TAKE ANY OF A VERY LARGE SUITE OF POSSIBLE VALUES $(\bar{x}, \bar{y})$.

BACK

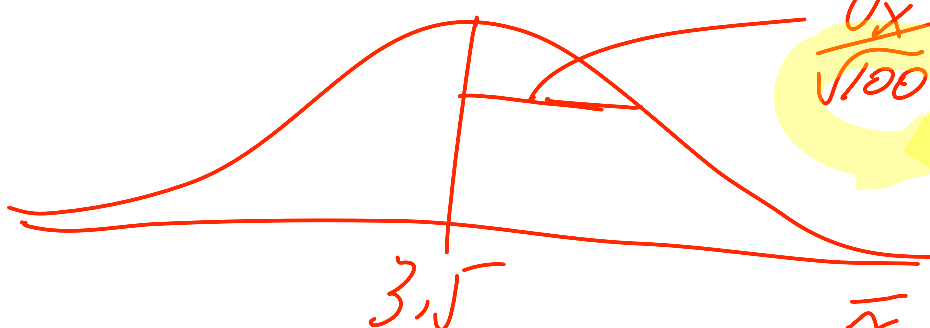
DICE 1 2 3 4 5 6

$EX = 3.5$

$$\sqrt{EX^2 - (EX)^2} = \sigma_X$$

PROSPECT OF $\bar{X}$ FROM $n = 100$ TOSSES.

$\approx$



$$\frac{\sigma_X}{\sqrt{100}} = \frac{\sqrt{(1-3.5)^2 + \cdots + (6-3.5)^2 / 6}}{10}$$

AVG — n

$$\sigma_X = \sqrt{\operatorname{Var} X} = \sqrt{\frac{(1-3.5)^2 + (2-3.5)^2 + \cdots + (6-3.5)^2}{6}}$$

← $n-1$ FOR DATA

STT 200 4-1-09b

READ (a) CORRECATION OF R.V. = $\dfrac{E(XY) - EX\,EY}{\sqrt{EX^2-(EX)^2}\sqrt{EY^2-(EY)^2}}$

END OF CH 8? 17?

(b) CH 18 (i) SAMPLING DISTRIBUTION ($\hat{p}$, $\bar{x}$).

(ii) CENTRAL LIMIT THEOREM.

RECALL FIRST OF CLASS — NOTION OF CI FOR A POPULATION MEAN. EXAMPLE: SAMPLE MSU UNDERGRADS SAMPLE $n = 100$ MEAS THEIR CHOLESTEROL LEVELS. $X_1 \ldots X_{100}$. THEN FORM $\bar{X}$

"WE CLAIMED" $P(\bar{X} \pm 1.96\,\sigma/\sqrt{100}$ COVERS POP AVG C.L.$) \sim .95$

POINT EST OF AVG (POP) CHOL LEVEL

WIDEN $\pm 1.96 \frac{\sigma}{\sqrt{100}}$?

WHAT IS THE THEORY BEHIND THIS?

CENTRAL LIMIT THEOREM — ASSERTS THAT DIST OF $\bar{X} \sim$ NORMAL

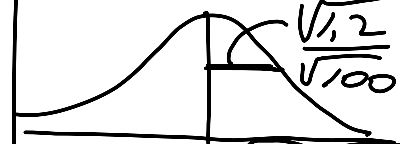
FIRSD, GO BACK TO 1·N· RECALL EX, $\sigma_X = \sqrt{EX^2-(EX)^2}$

eg.

x	$p(x)$	$x\,p(x)$	$x^2 p(x)$
0	.3	0	$0^2 \cdot .3 = 0$
4	.6	2.4	$4^2 \cdot .6 = 9.6$
6	.1	.6	$6^2 \cdot .1 = 3.6$
	1	$EX = 3$	$EX^2 = 10.2$

PROBY DISTRIBUTION (ONE PLAY)

INSTEAD, LOOK AT AVG $\bar X$ OF 100



$3 = EX$ $\boxed{\frac{t}{100} = \bar X}$

SO $Var\, X = \sqrt{10.2 - 9} = \sqrt{1.2}$

SO (ONE PLAY) $\mu_X = 3$, $\sigma_X = \sqrt{1.2}$

mu

DO NOT SAY ANYTHING USEFUL ABOUT 1 PLAY.

$Var(X_1 + \cdots + X_{100}) = 100\, Var X = 100(1.2)$

$E\bar X = E\frac{X_1 + \cdots + X_{100}}{100} = \mu_X = 3$

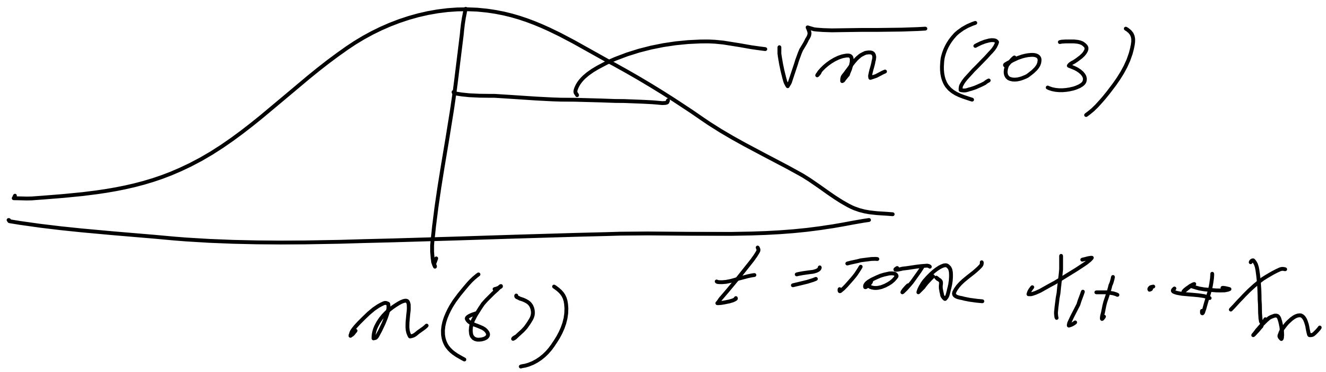
TOTAL OF 100 INDEP PLAYS $\approx$

$300 = 100\mu_X$ t = TOTAL

SEE ABOVE "BOX" /100

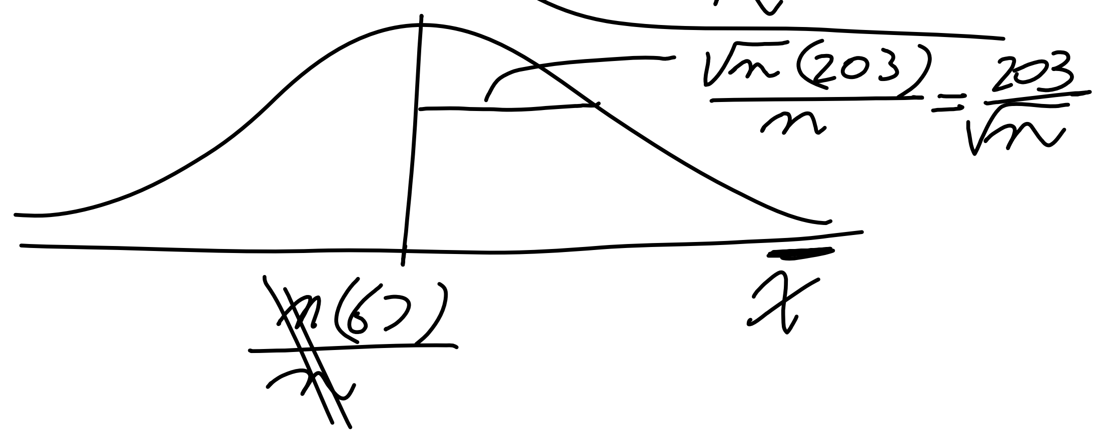
GEN'L: LOTTERY X $\quad EX = 67$ (SAY) $\quad \sigma_X = (203)$

$\approx$ DIST OF TOTAL $X_1 + \cdots + X_n$ (n LARGE)



BUT, IF INSTEAD YOU LOOK AT $\bar{X} = \dfrac{X_1 + \cdots + X_n}{n} = \dfrac{\text{TOTAL}}{n}$

SAME PICTURE AS ABOVE BUT WITH MEAN & SD (OF PIC) DIVIDED BY n

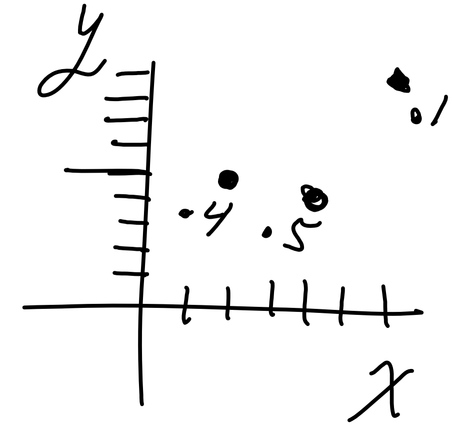


CORRELATION BETWEEN TWO r.v. X, Y.

LET'S SAY WE'RE PLAYING THE SAME ROULETTE TABLE.

SIMPLIFY TO (X = MY RETURN, Y = YOUR RETURN).

x	y	$p(x,y)$	$x\,y\,p(x,y)$
2	5	.4	$2 \cdot 5 \cdot .4 = 4$
4	4	.5	$4 \cdot 4 \cdot .5 = 8$
6	9	.1	$6 \cdot 9 \cdot .1 = 5.4$
		1	$E(XY) = 17.4$



APPEARS CORRELATION IS POSITIVE.

(FOR r.v.) CORRELATION(X,Y): $\dfrac{E(XY)-(EX)(EY)}{\sqrt{EX^2-(EX)^2}\sqrt{EY^2-(EY)^2}}$

$E(XY) = 17.4$

" "

SUPPOSE WE FIND $EX =$ σ_X

$EY =$ σ_Y

CORR (X,Y)

WHAT DOES THIS SAY ABOUT (X, Y)

PIC DISPLAYING WHAT OUR RESPECTIVE $(\underbrace{\bar{X}}_{EX}, \underbrace{\bar{Y}}_{EY})$

SHOULD LOOK LIKE.

EMPIRICAL (RANDOM) AVG (PER PLAY) RETURN. $\bar{X} \sim EX$

THEOR. EXPECTED RETURN FOR 1 PLAY.