STT 200 12th 7-14-10

Note Title 7/14/2010

(1) RANDOM VARIABLE X (values denoted x)

IS NUMERICAL FUNCTION ON THE OUTCOMES OF A PROBABILITY EXPERIMENT.

(2) DISTRIBUTION OF A RANDOM VARIABLE.

(1) TOSS OF A DIE

	1	2	3	4	5	6
PR	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$

TOSS OF RED + GREEN DICE

LET $X = R+G$

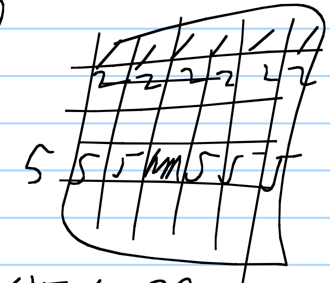
x	2	3	... 7	3	... 8	7	... 12
$p(x)$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$

R \ G	1	2	3	4	5	6
1	2	3	4	5	6	7
2	3	4	5	6	7	8
3	4	5	6	7	8	9
4	5	6	7	8	9	10
5	6	7	8	9	10	11
6	7	8	9	10	11	12

DEFINE "EXPECTED VALUE OF RANDOM VARIABLE (r.v.) X" $=7$

$$EX = \sum x\,p(x) = 2\left(\tfrac{1}{36}\right) + \dots + 7\left(\tfrac{1}{36}\right) + \dots + 7\left(\tfrac{1}{36}\right) + \dots + 12\left(\tfrac{1}{36}\right)$$

EXPECT EX <u>ON AVERAGE</u> (RELEVANT TO MANY PLAYS)



HAVE TWO r.v. R, G $\quad X = R+G$

RULE: $E(R+G) = ER + EG$ ← EACH 3.5

$$E(aR+bG+c) = aER + bEG + c$$ <u>IN GENERAL</u>

<u>DISTRIBUTION OF r.v. X</u>: DISTINCT VALUES v AND THEIR PROBABILITIES

v	2	3	4	5	6	7	8	9	10	11	12
$p(v)$	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{3}{36}$	. .		$\frac{6}{36}$	. .		$\frac{3}{36}$	$\frac{2}{36}$	$\frac{1}{36}$

$$EX \overset{\text{ALSO}}{=} \sum v\,p(v) = 2\left(\tfrac{1}{36}\right) + 3\left(\tfrac{2}{36}\right) + \dots + 7\left(\tfrac{6}{36}\right) + \dots + 12\left(\tfrac{1}{36}\right)$$

GROUPING TERMS $=7$

ALSO $EX = E(R+G) = ER + EG = 3.5 + 3.5 = 7$

WHAT ABOUT VARIANCE?

$$\sigma_X^2 = E(X^2) - (EX)^2 \overset{\text{DEF.}}{=} E(X - EX)^2$$

SO TOO, THIS CAN BE CALCULATED IF DISTRIBUTION IS KNOWN FOR X

ONLY DEPENDS OF DISTRIBUTION OF r.v. X

$$EX = \sum v\, p(v)$$ v DISTINCT VALUES OF X

$$\sigma^2 = \sum (v - EX)^2 p(v)$$

or $$EX = \sum x\, p(x)$$

OUTCOMES FOR X (SOME POSSIBLY DUPLICATED))

PROBABILITY OF THE PARTICULAR INSTANCE OF X.

EXAMPLE $(R-3)^2$

R	1	2	3	4	5	6
$(R-3)^2$	4	1	0	1	4	9

$$E(R-3)^2 = \sum x\,p(x) = 4\left(\tfrac{1}{6}\right) + 1\left(\tfrac{1}{6}\right) + 0\left(\tfrac{1}{6}\right) + 1\left(\tfrac{1}{6}\right) + 4\left(\tfrac{1}{6}\right) + 9\left(\tfrac{1}{6}\right)$$

HAPPENS TO BE SAME $p(x)$ AS

$$E(R-3)^2 = \sum_{v} v\,p(v)$$

DISTINCT VALUE

$$= 0\left(\tfrac{1}{6}\right) + 1\underbrace{\left(\tfrac{1}{6}+\tfrac{1}{6}\right)}_{p(v=1)} + 4\underbrace{\left(\tfrac{1}{6}+\tfrac{1}{6}\right)}_{p(v=4)} + 9\left(\tfrac{1}{6}\right)$$

AGAIN EX AND $\sigma_X^2 = \sum (v - EX)^2 p(v) = E(X-EX)^2$

IMPORTANT INSIGHT:

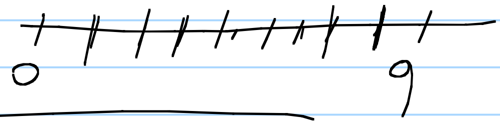
X = \$ ON YOUR PERSON

Y = LAST DIGIT STUDENT #

XY SUPPOSE X, Y ARE INDEP

$4.5 = EY$

WHAT IS $E(XY)$? $= (EX)(EY)$



SO (AVG \$ HELD BY MSU STUDENTS)(LAST DIGIT OF ST. #)

CLAIM AVG PRODUCT OF STATISTICALLY UNRELATED R.V.'s IS EQUAL TO PRODUCT OF THEIR RESPECTIVE AVGS.

BIG $\Longrightarrow$ IMPLICATION

$$\mathrm{Var}(X+Y) \underset{\text{INDEP}}{=} \mathrm{Var}\,X + \mathrm{Var}\,Y.$$

(IF)

BIG DEAL B-CAUSE

$$X_1 + X_2 + \cdots + X_n \quad \text{INDEP PLAYS.}$$

$$E(\underbrace{X_1 + \cdots + X_n}_{\text{RANDOMLY EVOLVING TOTAL}}) = EX_1 + EX_2 + \cdots + EX_n$$

INDEP PLAYS? IF SO

$$\mathrm{Var}(X_1 + \cdots + X_n) = \mathrm{Var}\,X_1 + \cdots + \mathrm{Var}\,X_n.$$

SO

$$\sigma_{X_1 + \cdots + X_n} = \sqrt{\mathrm{Var}\,X_1 + \cdots + \mathrm{Var}\,X_n}$$

EXPECTATION GROWS (OR DECAYS) WITH n = # PLAYS

BE $\sigma_{X_1+\cdots+X_n}$ ONLY LIKE $\sqrt{\operatorname{Var} X_1 + \cdots + \operatorname{Var} X_n}$.

$$X_1+\cdots+X_n = (EX_1+\cdots+EX_n) + \left[(X_1+\cdots+X_n) - ((EX_1)+\cdots+EX_n)\right]$$

TAKES LESS THAN $\sqrt{\operatorname{Var} X_1 + \cdots + \operatorname{Var} X_n}$ TO KILL THIS.

CONCLUSION: IN MOST ACTIVITIES OF THIS KIND YOUR EARNINGS $X_1+\cdots+X_n$ WILL NOT DEPART TOO IMPORTENTLY FROM WHAT IS EXPECTED, NAMELY $EX_1+\cdots+EX_n$.

EXAMPLE: $R, G \quad X = R+G$

CLAIM: $EX = ER + EG = 3.5 + 3.5 = 7$ *

[DIST^N OF X:

2	3	4	5	6	7	8	9	10	11	12
$1/36$	$2/36$				$6/36$				$2/36$	$1/36$

]

"SEE" $EX = 7$. CONFIRMING *

^ BALANCE POINT = MEAN

WHAT ABOUT $\mathrm{Var}\, X$? R, G ARE INDEPENDENT.

SO SUPPOSED TO BE TRUE THAT $\mathrm{Var}\, X = \mathrm{Var}\, R + \mathrm{Var}\, G$

$= 2 \mathrm{Var}\, R = 2(8.47)$?

$$\mathrm{Var}\, X = \sum_i (x - EX)^2 p(x)$$

$$= (2-7)^2 \frac{1}{36} + (3-7)^2 \frac{2}{36} + \cdots + (7-7)^2 \frac{6}{36} + (8-7)^2 \frac{5}{36} + \cdots + (12-7)^2 \frac{1}{36}.$$

EXAMPLE TOSS COIN 100 TIMES

1 H $\quad pr. = \frac{1}{2}$

0 T $\quad pr = \frac{1}{2}$

ONE TOSS $\quad E X_1$ — 1 or 0 FIRST TOSS

$= \sum$ VALUE TIMES PROBABILITY

$= 0\left(\frac{1}{2}\right) + 1\left(\frac{1}{2}\right) = \frac{1}{2}$ ON AVE GET $\frac{1}{2}$ HEAD IN ONE TOSS OF A COIN

LET $X = X_1 + X_2 + \cdots + X_{100}$

$\overbrace{\text{H H H T T H T H} \ldots \text{H T H}}^{\text{100 TOSSES}}$

VALUE OF X_i $\quad$ 1 1 1 0 0 1 0 1 $\quad$ 1

$X_1 = 1$ B.CAUSE GOT "H" ON TOSS 1

$$EX = E(X_1 + \cdots + X_{100}) = EX_1 + \cdots + EX_{100} = 100\left(\tfrac{1}{2}\right) = 50$$

(YOU ARE EXPECTING TO EARN 50 OVER COURSE OF 100 PLAYS IF COIN IS FAIR)

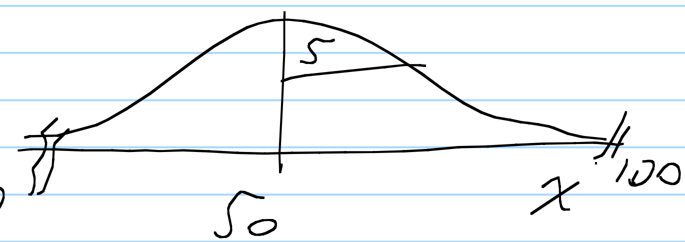
X_1	
1	$\frac{1}{2}$
0	$\frac{1}{2}$

$$\mathrm{Var}\, X_1 = EX_1^2 - (EX_1)^2 = \frac{1}{2} - \left(\frac{1}{2}\right)^2 = \frac{1}{4}$$

$$\text{SO } \mathrm{Var}\, X = \mathrm{Var}(X_1 + \cdots + X_{100}) \underset{\text{BY INDEP}}{=} 100\left(\tfrac{1}{4}\right) = 25$$

$$\text{SO } \sigma_X = \sqrt{25} = 5.$$

CENTRAL LIMIT THEOREM: $\sim$ DIST OF X $\sim$



EXAMPLE

54% OF VOTERS ARE D.
46% R

$$X_1 = \begin{cases} 1 & D & .54 \\ 0 & R & .46 \end{cases} \qquad EX_1 = 0(.46) + 1(.54) = .54$$

$$\text{Var } X_1 = \underbrace{E X_1^2}_{X_1^2 = X_1} - (E X_1)^2 = .54 - (.54)^2$$

$$= .54(1 - .54) = pq$$

$$p = .54$$

EX BINOMIAL $= np$

Var X " $= np(1-p)$

n = # VOTERS SAMPLED.

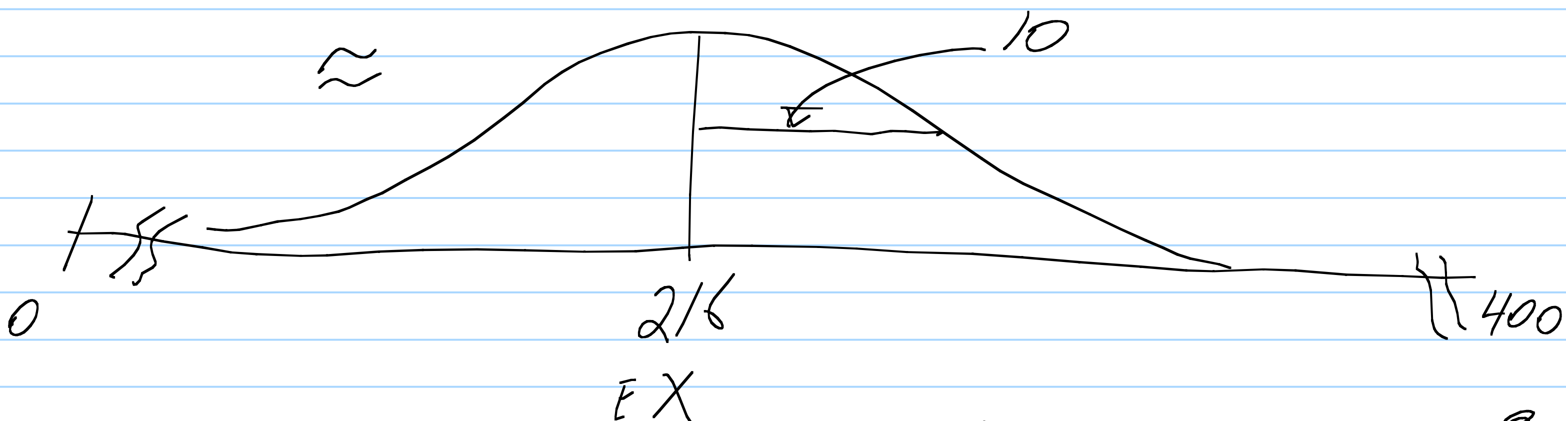
TRY $n = 400$ VOTERS. INDEP ASSUMED

X = TOTAL # D IN $n = 400$

$$EX = np = 400(.54) = 216$$

$$\operatorname{Var} X = np(1-p) = 400(.54)(.46) \approx 99.36$$

$$\sigma_X = \sqrt{99.36} \approx 10$$



CHANCE GET BETWEEN 216 ± 10 AROUND 68%

CHANCE " " $2/6 \pm 1.96/10$ " 95%