PROBABILITY CLASS 22
M 4/4/22

REVIEW CLASS 19 3/24

3 COIN TOSS. FAIR

RANDOM EXPERIMENT: TOSS 3 COINS.

OBSERVE 2 RANDOM VARIABLES FROM EACH EXPERIMENT.

OUTCOMES: 8 WAYS 3 COINS COME UP.

2 RV: X: # HEADS
Y: POSITION OF 1st H

OUTCOME	X	Y
TTT	0	0
TTH	1	3
THT	1	2
THH	2	2
HTT	1	1
HTH	2	1
HHT	2	1
HHH	3	1

PMF MASS FUNCTIONS

PMF X :

0	1
1	3
2	3
3	1

PMF Y

0	1
1	4
2	2
3	1

IF 1st HEAD WAS AT POSITION 1,

WHAT'S PROB FOR # HEADS.

$P(X|Y)$

$X : \{0, 1, 2, 3\}$

$Y : \{0, 1, 2, 3\}$

x	y	#(x, y)	P
0	0	1	1/8
0	1	0	
0	2	6	
0	3	0	
1	0	6	
1	1	1	1/8
1	2	1	
1	3	1	
2	0	0	
2	1	2	
2	2	1	
2	3	0	
3	0	6	
3	1	1	
3	2	0	
3	3	0	

8 ✓

CONDITIONALS

$$P(X|Y)\,P(Y) = P(X \,\&\, Y)$$

$$P(X|Y) = \frac{P(X \,\&\, Y)}{P(Y)}$$

$$P(X=1|Y=1) = \frac{P(X=1 \,\&\, Y=1)}{P(Y=1)}$$

$$= \frac{1/8}{1/2} = 1/4$$

$E[X]$

$E[X^2]$

$$VAR[X] = E[(X - E[X])^2]$$

$$= E[X^2] - (E[X])^2$$

2 RV

$\mathrm{COV}(X, Y) =$

$E[(X - E[X])(Y - E[Y])]$

$= E[XY] - E[X]E[Y]$

E.G. 2 LINKED COINS.

T → 0 H → 1

	H	T	
H	.3	.2	.5
T	.2	.3	.5
	.5	.5	

X: 1st

Y: 2nd

$E[X] = .5 \times 0 + .5 \times 1 = .5 = 1/2$

$E[Y] = 1/2$

$E[XY] = .3 \times 1 + .2 \times 0 + .2 \times 0 + .3 \times 0 = .3$

$\mathrm{COV}(X, Y) = E[XY] - E[X]E[Y] = .3 - .25 = .05$

POSITIVE: X, Y CORRELATED.

FOR ANYTHING $E[?] =$

$\sum$ (POSSIBLE VALUES FOR ?) $\times$ (PROBABILITY)

TOSS COIN

X	P
0	1/2
1	1/2

$E[X] = 0 \cdot \frac{1}{2} + 1 \cdot \frac{1}{2} = \frac{1}{2}$

FOR $E[XY]$ "X TIMES Y"

X	Y	XY	P
0	0	0	.3
0	1	0	.2
1	0	0	.2
1	1	1	.3

$E(XY) = 0 + 0 + 0 + .3 = .3$

$$\mathrm{VAR}[X] = \underbrace{E[X^2]}_{1/2} - (\underbrace{E(X)}_{1/2})^2$$

$$= \frac{1}{2} - \frac{1}{2}\cdot\frac{1}{2} = \frac{1}{4}$$

$$\mathrm{STD}[X] = \sqrt{\mathrm{VAR}(X)} = 1/2$$

$$\mathrm{COV}(X,Y) = \underbrace{E[XY]}_{.3} - \underbrace{E[X]}_{.5}\underbrace{E[Y]}_{.5}$$

$$= .05$$

CORREC COEF

$$\rho = \frac{\mathrm{COV}(X,Y)}{\mathrm{STD}(X)\,\mathrm{STD}(Y)} = \frac{.05}{.25}$$

$$= .2$$

.2 IS SMALL.

$$|\rho| \le 1$$

REDO FOR STRONGLY CORRELATED COINS

X

Y	H	T	
H	.4	.1	.5
T	.1	.4	.5
	.5	.5	

$E[Y] = 1/2$

$VAR[Y] = \frac{1}{4}$

$E[XY] = .4$ $\quad COV(X,Y) = .4 - .25$

$= .15$

CORREL $= \frac{.15}{.25} = .6$

<u>PROBABLY</u>

CHAPTER 7 P359

N RV $X_1 \cdots X_n$

$$S_n = \sum_{i=1}^{N} X_i$$

$$E[S_n] = \sum E[X_i]$$

$$VAR(S_n) = E\left(\left(S_n - E[S_n]\right)^2\right)$$

$$= E\left[\left(\sum X_i - \sum E[X_i]\right)^2\right]$$

$$= E\left[\left(\sum X_i\right)^2 - 2\sum X_i\, E[X_i] + E[X_i]^2\right)$$

$$=$$

$$= \sum VAR(X_i) + \sum_{i \neq j}\sum COV(X_i X_j)$$

GENERALLY

$$VAR\left(\sum X_i\right) \neq \sum VAR(X_i)$$

IF ALL $COV(X_i, X_j) = 0$

$$VAR(\Sigma X_i) = \Sigma VAR(X_i)$$

EX 7.1 p360

IF ALL X_i IID

INDEP
IDENT
DIST.

$$\sigma = STD(X_i)$$

$$VAR(\Sigma X_i) = n\sigma^2$$

§2.2 SAMPLE MEAN

LAWS OF LARGE NUMBERS.

- HAVE n R.V. X_i
- DON'T KNOW $E[X_i]$
- WANT IT.

DEFINE NEW R.V. $M_n = \frac{1}{n} \sum_{1}^{n} X_i$

"SAMPLE MEAN"

THE SAMPLE MEAN HAS A MEAN.

$$E[M_n] = \frac{1}{n} \sum E[X_i] = \mu$$

μ: UNKNOWN POPULATION MEAN.

HOW MUCH DOES $E[M_n]$ BOUNCE AROUND?

		HAPPEN	
		Y	N
$3\sigma =$ HERE	Y	✓	✓
	N	✗	✗

X_i NOT CORRELATED

$$\mathrm{VAR}[\mu_n] = \mathrm{VAR}\left[\frac{1}{n}\sum X_i\right]$$

$$= \frac{1}{n^2}\,\mathrm{VAR}\left[\sum X_i\right]$$

NOTE !!

$$= \frac{1}{n^2}\sum \mathrm{VAR}(X_i) = \frac{1}{n^2}\,n\sigma^2$$

$$= \frac{\sigma^2}{n}$$

LARGER SAMPLE (LARGER n) $\Rightarrow$ SMALLER VARIANCE OF SAMPLE MEAN.

OBSERVE ROBINS (BIRDS)

X_n = # MINUTES ~~BIRD~~ #n SINGS TODAY.

X_n = 30, 28, 35, 20, 40, 32, ...

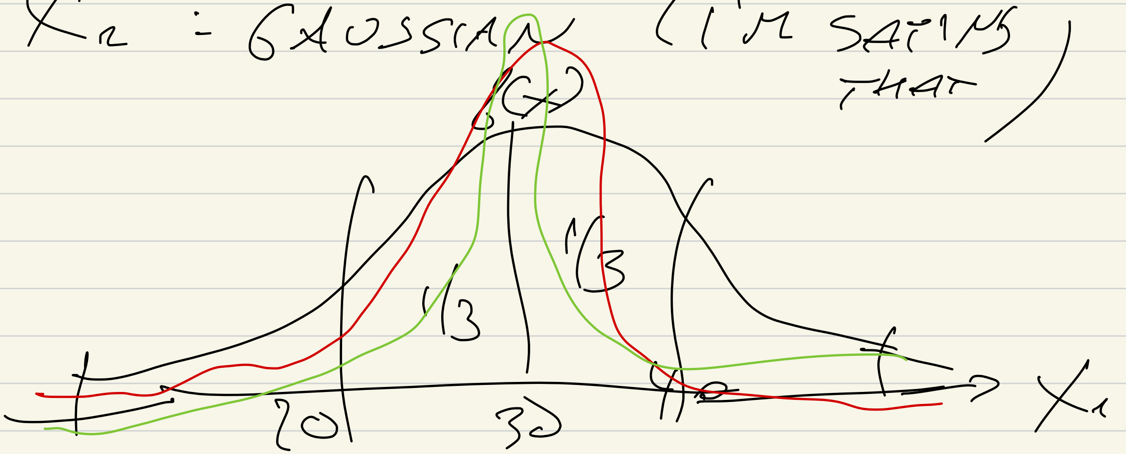
WHOLE POPULATION IS 10000 ROBINS.

$X_n \quad n = 1 \dots 10000$

POP MEAN $\mu = 30$

$\sigma = 10$

X_n : GAUSSIAN (I'M SAYING THAT)



DISTNOE SAME MEAN TIGHTER

$P(X > 40) = 1/6$

$\mu + \sigma$

TAKE SAMPLES OF 4 BIRDS.

S_1 32, 25, 33, 29

S_2 40, 30, 25, 31

S_3 29, 29, 32, 30

SAMPLE MEAN M_i

ALL SAMPLE MEANS HAVE A MEAN $E[M] = \mu = 30$

$VAR[M] = \frac{\sigma^2}{n}$

$STD[M] = \frac{\sigma}{\sqrt{n}}$

$n = 4$ SAMPLE $\sigma = 5$

HOW CLOSE IS UNKNOWN TRUE POP MEAN μ TO SAMPLE MEAN?

USE CHEBYSHEV. EQ 7.19 368

$$P\left[|M_n - \mu| > \varepsilon\right] \le \frac{\sigma^2}{n\varepsilon^2}$$

THIS IS A VERY LOOSE BOUND.

WEAK LAW OF LARGE NUMBERS

$$\lim_{n\to\infty} P\left[|M_n - \mu| < \varepsilon\right] = 1$$

ASSUMES POP VARIANCE IS FINITE.

7.3 CENTRAL LIMIT THM. P 369

M_n BECOMES GAUSSIAN AS $n \to \infty$.

REGARDLESS OF PDF OF X_i