Block 2

Discrete Probability

- **random variable**: fxn that assigns one and only one number to each outcome in a sample space S
  X or Y
  - **support/domain for a random variable**: all of the #s that can be assigned to the function
    - can be discrete or continuous depending on the problem

EX: Roll 2 4-sided die (red, green)

$S = \{11, 12, 13, 14, 21, 22, 23, 24, 31, 32, 33, 34, 41, 42, 43, 44\}$

..., red# green #, ...

can have 2 diff. random variables:

(1) $X \equiv$ the sum of the numbers that are on the face
i.e. red + green

$X(13) = 4$ support $X = [2, 8]$

(2) $Y \equiv$ the difference b/t the red and green die
i.e. red − green

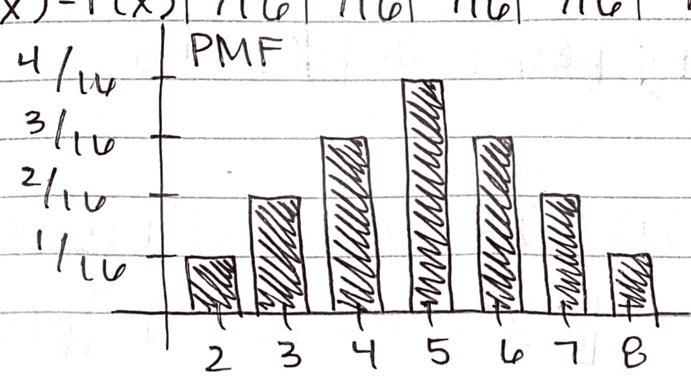
$Y(13) = -2$ support $Y = [-3, 3]$

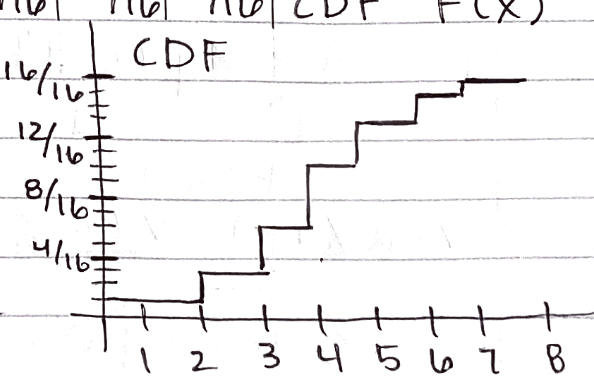
Assumption: Fair die

- **probability mass function**

support

x	2	3	4	5	6	7	8	
$P(X=x)=p(x)$	1/16	2/16	3/16	4/16	3/16	2/16	1/16	PMF $p(x)$ ↳ sum to 1
$P(X \le x)=F(x)$	1/16	3/16	6/16	10/16	13/16	15/16	16/16	CDF $F(x)$



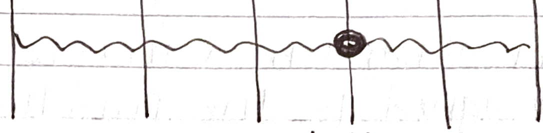


SEE Rcode

EX: We know the dist. of flash drives sold in store

X	1	2	4	8	16
p(X)	.05	.1	.35	.4	.1

$X \equiv$ amt. of memory for purchased flash drive (GB)



6.45

$E(X^2) = 1^2(.05) + 2^2(.1) + (4^2).35 + 8^2(.4) + 16^2(.1) =$

SEE R

$-6.45 = 15.65$

- What is the long run avg. size of purchased flash drive?

mean $= E(X) = \mu = (.05)1 + (.1)2 + (.35)4 + (.4)8 + (.1)16 = 6.45$ GB

is the weighted average

Expected Value $= \sum_x x \cdot p(x) = \mu$

True Average

SEE R

variance $= (1-6.45)^2(.05) + (2-6.45)^2(.1) + (4-6.45)^2(.35) + (8-6.45)^2(.4) + (16-6.45)^2(.1) = 15.647$

$= \sum_x (x-\mu)^2 p(x)$

SEE R

$P(X=x)$ $\quad$ $P(X \leq x)$

$P(X=4)$ $\quad$ $P(X \leq 4)$

E.V. and Var Rules for $aX+b$ where a,b are constants and X is the r.v.:

$E(aX+b) = \sum_x (ax+b) \cdot p(x) = \sum_x (ax \cdot p(x)) + \sum_x (b \cdot p(x))$

$= \underbrace{a \sum_x x \cdot p(x)}_{E(X)} + b \xrightarrow{1} \cancel{\sum_x p(x)} = aE(X)+b$

$V(aX+b) = \underline{V(aX) + V(b)} = \underline{\sum_x (ax-a\mu)^2 p(x)} + \cancel{\sum_x (b-b)^2 p(x)} \rightarrow 0 = \sum_x [a^2x^2 - 2a^2x\mu + a^2\mu^2] p(x) = a^2 \cdot V(X)$

Shortcut formula for V(X)

$$V(X) = \sum_x (x-\mu)^2 \cdot p(x) = \sum_x (x^2 - 2x\mu + \mu^2) \cdot p(x) =$$

$$\underbrace{\sum_x x^2 p(x)}_{} - 2\mu \overbrace{\cancel{\sum_x x p(x)}}^{\mu} + \mu^2 \overbrace{\cancel{\sum_x p(x)}}^{1}$$

$$E(X^2) - 2\mu^2 + \mu^2 = E(X^2) - \mu^2 = E(X^2) - E(X)^2$$

- Bernoulli R.V.

: any r.v. whose possible values are 0 or 1
(failure or success)

$X \sim$ Bernoulli (p) p is probability of success

PMF X	0	1
$P(X=x)=p(x)$	$1-p$	p

$$E(X) = \sum_x x \cdot p(x) = 0(1-p) + 1(p) = p$$

$$E(X^2) = \sum_x x^2 \cdot p(x) = 0^2(1-p) + 1^2(p) = p$$

$$V(X) = p - p^2 = p(1-p) \text{ or } pq$$

- Binomial Dist. (n,p) : sum of n Bernoulli trials

must meet following:

1) must have n trials
2) each trial must be S or F
3) independent trials
4) P(success) = p is constant

- answers Q's about the prob. of getting x successes out of n trials

EX. Lebron makes 75% free throws, takes 4 foul shots, what is the FT% over the next 4?

$X \equiv$ the # out of next 4 Lebron makes

$X \sim bin(4, .75)$ support - 0, 1, 2, 3, 4

All prob of next 4: SSSS $P(X=4) = \binom{4}{4} .75^4 \, 0.25^0$

SSSF SSFS SFSS FSSS	makes 3 throws	$P(X=3)=\binom{4}{3}.75^3.25^1$
SSFF SFSF SFFS FSSF FSFS FFSS	makes 2 throws	$P(X=2)=\binom{4}{2}.75^2.25^2$
FFFS FFSF FSFF SFFF	makes 1 throw	$P(X=1)=\binom{4}{1}.75^1.25^3$
FFFF	makes NONE	$P(X=0)=\binom{4}{0}.75^0.25^4$

x	0	1	2	3	4
$P(X=x)$ $=p(x)$	$\binom{4}{0}.75^0.25^4$	$\binom{4}{1}.75^1.25^3$	$\binom{4}{2}.75^2.25^2$	$\binom{4}{3}.75^3.25^1$	$\binom{4}{4}.75^4.25^4$

In general $X \sim bin(n,p)$

$$P(X=x)=\binom{n}{x}p^x(1-p)^{n-x}$$

$E(X) = E(Y_1+Y_2+...+Y_n)$ $\quad X=Y_1+Y_2+...+Y_n$

$= E(Y_1)+E(Y_2)+...+E(Y_n)$ $\quad Y_i \sim bernoulli(p)$

$V(X) = V(Y_1)+V(Y_2)+...+V(Y_n)$

$= p(1-p)+p(1-p)+p(1-p) = np(1-p)$