AMAT 362 – PROBABILITY FOR STATISTICS

LECTURE 9

3M's
- MEAN
- MEDIAN
- MODE

→ ALL DIFFERENT

Recall BINOMIAL RV

S_n BINOM.

RV that counts the # of successes out of n tries

$$P\begin{pmatrix}k \text{ successes} \\ \text{out of } n \text{ tries}\end{pmatrix} = \binom{n}{k} p^k q^{n-k}$$

// OR

$$P(S_n = k) = P_{S_n}(k)$$

where p = prob of success, $q = 1-p$

PMF notation (Probability Mass Function)

k	0	1	...	k	...	n
$P(k)$	q^n	$\binom{n}{1} p q^{n-1}$		$\binom{n}{k} p^k q^{n-k}$		p^n

Recall GEOMETRIC RV

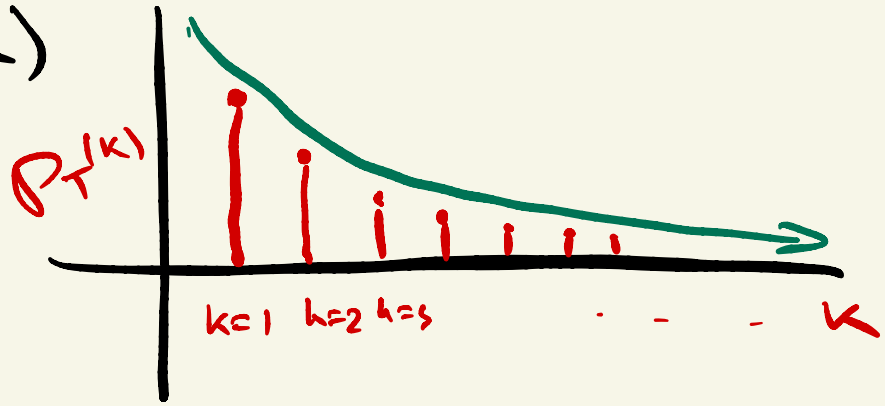
RV that counts # of tries needed for first success!

T = RV

$$P(\text{1st success is on kth try}) = q^{k-1}p$$

$$\| $$

$$P(T=k)$$



TODAY: 3 SUMMARY STATISTICS

↳ A single # that summarizes a probability distribution / RV

1) MEAN : AKA Average or Expectation

$$\mu = \bar{x} = \frac{x_1 + \cdots + x_n}{n} \quad \text{OR} \quad E(X)$$

μ = "mu"

$\bar{x}$ = Empirical Average

2) MEDIAN: A # where 50% of outcomes of X is below median & 50% above

⇒ NOT UNIQUE! (in general)

3) MODE: Most probable outcome or value for X

⇒ ALSO NOT UNIQUE (in general)

N.B. All of these are different P_0

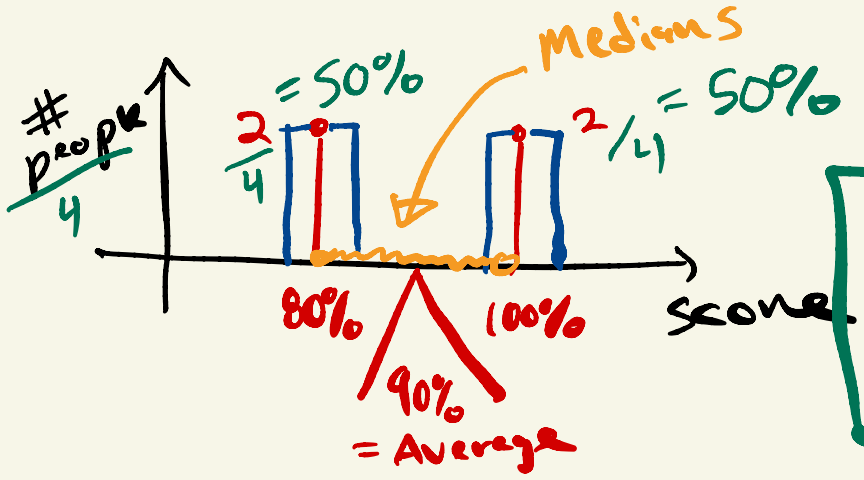
EX 4 people take a test

↳ 2 people get 80%

↳ 2 people get 100%

Mean = Empirical Average $\bar{x} = \frac{80+80+100+100}{4}$

$= \frac{360}{4} = 90\%$



EMPIRICAL PMF

$P(X=k) = \frac{\#\text{ times } k \text{ appears in a list of } n \text{ \#'s}}{n}$

Median Any number m w/

$P(X \geq m) = \frac{1}{2} = 50\%$

AND $P(X \leq m) = \frac{1}{2} = 50\%$

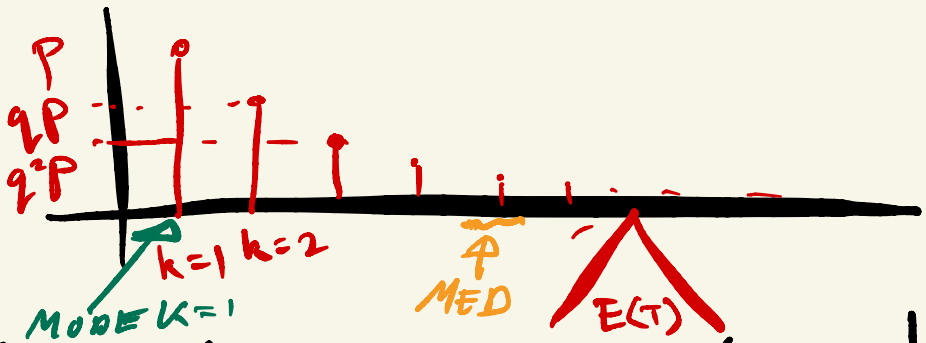
Ex Here any $\#\, m \in [80, 100)$ is the median

MODE Most probable outcome

i.e. k such that $P(X=k)$ is maximum possible value

Ex Both 80 & 100 are modes

Ex Geometric RV T



MEAN Intuitively... if $p = \frac{1}{10}$ ... $E(T) = 10$ takes 10 trials

MEDIAN Tough calculation ~ $p = 1/10$

69% × (10)
" $(\log_e 2) \times (\frac{1}{p})$

MODE MOST LIKELY is $k = 1$!

Ex 1 revisited

Turn grades (w/ repetition) into a RV

k	80	100
$p(X=k)$	1/2	1/2

$$\bar{X} = 80 \cdot \frac{1}{2} + 100 \cdot \frac{1}{2}$$
$$= 40 + 50$$
$$= 90$$

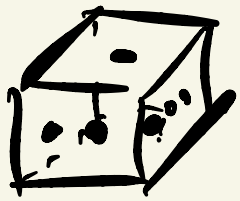
GENERAL THEORETICAL DEF of MEAN

Def If X is a discrete RV, then the mean or expectation of X is

$$E(X) := \sum_{\text{all possible values } k} k \cdot P(X=k)$$

Ex Roll a 6-sided die

X = # that is facing up



PMF Table

k	1	2	3	4	5	6
$p(X=k)$	$1/6$	$1/6$	$1/6$	$1/6$	$1/6$	$1/6$

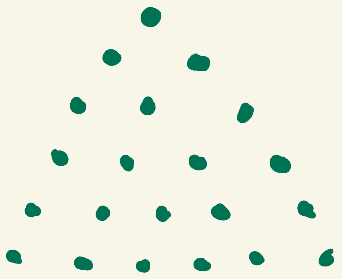
← uniform distribution on $\{1,2,\dots,6\}$

$$E(X) = \sum_{k=1}^{6} k \cdot p(X=k)$$

$$= 1\cdot\left(\frac{1}{6}\right) + 2\left(\frac{1}{6}\right) + 3\left(\frac{1}{6}\right) + 4\left(\frac{1}{6}\right) + 5\left(\frac{1}{6}\right) + 6\left(\frac{1}{6}\right)$$

$$= \frac{1}{6}(1+2+3+4+5+6) = \frac{1}{\cancel{6}} \cdot \frac{\cancel{6}\cdot 7}{2} = 3.5$$

ASIDE ON TRIANGULAR NUMBERS



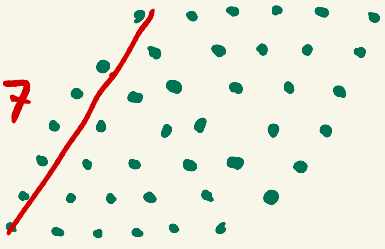
T_n = # dots in a triangle w/ n dots on base

"Gauss, sum $1+2+3+4+\cdots+100$"

$$\begin{array}{r} 100+99+98+97+\cdots+1 \\ \hline 101+101+101+101+\cdots+101 \end{array}$$

100 times

$$\frac{100\cdot 101}{2} = T_{100}$$



$\Rightarrow 42/2$

$$\frac{6\cdot 7}{2} = 21 \qquad n = 6 \text{ (die)}$$

$$T_n = \frac{n(n+1)}{2}$$

3 Interpretations of $E(X)$

I. Long Run Average

If you perform an experiment enough times & record the outcome X then

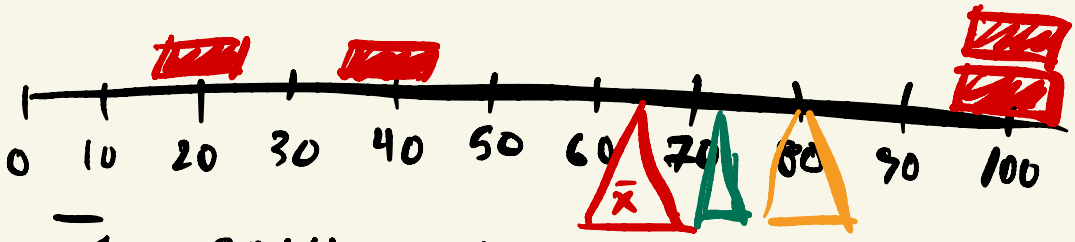
$$\text{Avg Value} = \text{Empirical Mean} \overset{\uparrow}{\sim} E(X)$$

very close to

$\Rightarrow$ LAW OF AVERAGES

II. CENTER OF MASS / GRAVITY / OF PMF X

Ex) 4 people take exam



$$\bar{X} = \frac{20+40+100+100}{4} = \frac{260}{4} = 65$$

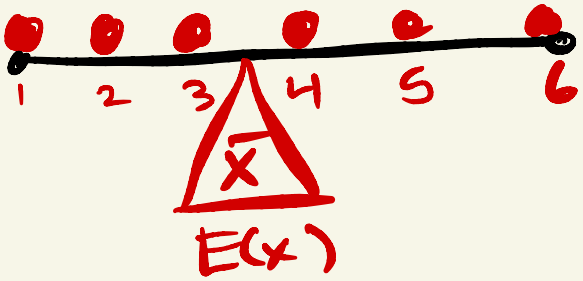
Remove 40 $\quad \bar{x} = \frac{220}{3} = 73.\overline{3}$

Remove 20 $\quad \bar{x} = \frac{240}{3} = 80$

Different effect

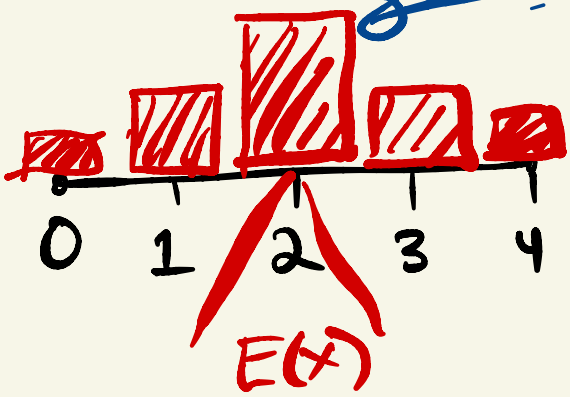
N.B. For uniform or symmetric distributions
the mean ~ middle

For 6-sided die



Every outcome is equally likely so Mode $\in \{1, \ldots, 6\}$

Flip a fair coin 4 times count #heads



Unique Mode!

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

III. FAIR PRICE / GAMBLER'S WINNINGS

Ex Sps you rent a car
They offer you extra insurance for $200.

⟹ **Should you buy it???**

Two Scenarios

No insurance X → $2000 accident w/ prob p
→ $0 if no accident $q = 1-p$

Insurance Y → $200 regardless

The insurance is worth it if

$$E(Y) \leq E(X)$$

$$\$200 \cdot 1 \qquad \$0 \cdot q + \$2000\, p$$

$$\Rightarrow 200 \leq 2000\, p$$

$$\Rightarrow p \geq 10\%$$

MORAL: If you are less than 10% likely to get into an accident then don't buy it!