EECS 461 PROBABILITY & STATISTICS
DISCRETE RANDOM VARIABLES

I. RANDOM VARIABLE

A RANDOM VARIABLE (RV) MAPS OUTCOMES OF A SAMPLE SPACE TO NUMBERS.

A RANDOM VARIABLE ASSIGNS NUMBERS TO OUTCOMES IN THE SAMPLE SPACE OF AN EXPERIMENT.

II. NOTATION

RV: Y

RANGE OF Y : S_Y : ALL POSSIBLE VALUES OF Y

PARTICULAR VALUE OF Y: y

DEF RANDOM VARIABLE

A RANDOM VARIABLE CONSISTS OF AN EXPERIMENT WITH A PROBABILITY MEASURE $P[\cdot]$ DEFINED ON A SAMPLE SPACE S & A FUNCTION THAT ASSIGNS A REAL NUMBER TO EACH OUTCOME IN THE SAMPLE SPACE OF THE EXPERIMENT.

S

IS THE SET OF ALL POSSIBLE OUTCOMES KNOWN AS A SAMPLE SPACE.
THE MATHMATICAL MODEL INCLUDES A RULE FOR ASSIGNING NUMBERS BETWEEN 0 & 1 TO SETS A IN S
THEREFORE FOR EVERY $A \subset S$ THE MODEL GIVES US A PROBABILITY $P[A]$ WHERE $0 \leq P[A] \leq 1$

PROBABILITY MODEL THAT ASSIGN NUMBERS TO THE OUTCOMES IN THE SAMPLE SPACE.
WHEN WE OBSERVE ONE OF THESE NUMBERS, WE REFER TO THE OBSERVATION AS A RANDOM VARIABLE.

THE SET OF POSSIBLE VALUES OF X IS THE RANGE OF X
WE DENOTE THE RANGE OF A RANDOM VARIABLE BY THE LETTER S WITH A SUBSCRIPT THAT IS THE NAME OF THE RANDOM VAR

S_X IS THE RANGE OF THE RANDOM VARIABLE X

S_Y IS THE RANGE OF THE RANDOM VARIABLE Y

A PROB. MODEL ALWAYS BEGINS WITH AN EXPERIMENT
EACH RANDOM VARIABLE IS DIRECTLY RELATED TO THE EXPERIMENT

THREE TYPES OF RELATIONSHIPS BETWEEN A RANDOM VARIABLE & AN EXPERIMENT

(1) THE RANDOM VARIABLE IS THE OBSERVATION
(2) THE RANDOM VARIABLE IS A FUNCTION OF THE OBSERVATION
(3) THE RANDOM VARIABLE IS A FUNCTION OF ANOTHER RANDOM VARIABLE.

THREE TYPES OF RELATIONSHIPS BETWEEN A RANDOM VARIABLE & AN EXPERIMENT

(1) THE RANDOM VARIABLE IS THE OBSERVATION
(2) THE RANDOM VARIABLE IS A FUNCTION OF THE OBSERVATION
(3) THE RANDOM VARIABLE IS A FUNCTION OF ANOTHER RANDOM VARIABLE.

(1) RV IS THE OBSERVATION

EX 3.1 THE EXPERIMENT IS TO ATTACH A PHOTO DETECTOR TO AN OPTICAL FIBER & COUNT THE THE NUMBER OF PHOTONS ARRIVING IN A ONE-MICROSECOND TIME INTERVAL.

EACH OBSERVATION IS A RANDOM VARIABLE X

THE RANGE OF X IS $S_X = \{0, 1, 2, \dots\}$

IN THIS CASE S_X IS THE RANGE OF X AND THE SAMPLE SPACE S ARE IDENTICAL.

(2) THE RV IS A FUNCTION OF THE OBSERVATION

EX 3.2

THE EXPERIMENT IS TO TEST SIX INTEGRATED CIRCUITS & AFTER EACH TEST OBSERVE WHETHER THE CIRCUIT IS ACCEPTED (A) OR REJECTED (R).

FOR EXAMPLE S_8 = AARAAA

THE SAMPLE SPACE S CONSISTS OF 64 POSSIBLE SEQUENCES.

$$\underset{1}{\underline{A/R}}\ \underset{2}{\underline{A/R}}\ \underset{3}{\underline{A/R}}\ \underset{4}{\underline{A/R}}\ \underset{5}{\underline{A/R}}\ \underset{6}{\underline{A/R}}$$

A RANDOM VARIABLE RELATED TO THIS EXPERIMENT IS N, THE NUMBER OF ACCEPTED CIRCUITS.

FOR OUTCOME S_8, $N=5$ CIRCUITS ARE ACCEPTED

THE RANGE OF N IS $S_N = \{0, 1, \dots, 6\}$

(3) THE RANDOM VARIABLE IS A FUNCTION OF ANOTHER RANDOM VARIABLE

NET REVENUE R OBTAINED FOR 6 INTEGRATED CIRCUITS AT \$5 FOR EACH CIRCUIT ACCEPTED MINUS \$7 FOR EACH CIRCUIT REJECTED

WHEN N CIRCUITS ARE ACCEPTED, $6-N$ CIRCUITS ARE REJECTED

THUS NET REVENUE R IS RELATED TO N BY THE FUNCTION

$$R = G(N) = 5N - 7(6-N) = 12N - 42 \text{ DOLLARS}$$

$S_N = \{0, \dots, 6\}$ THE RANGE OF R IS

$$S_R = \{-42, -30, -18, -6, 6, 18, 30\}$$

REVENUE ASSOCIATED WITH S_8 = AARAAA AND ALL OTHER OUTCOMES FOR WHICH $N=5$

THE RANDOM VARIABLE CAN ALWAYS BE DERIVED FROM THE OUTCOME OF THE UNDERLYING EXPERIMENT.

DEF 3.1 RANDOM VARIABLE

A RANDOM VARIABLE CONSISTS OF AN EXPERIMENT WITH A PROBABILITY MEASURE $P[\cdot]$ DEFINED ON A SAMPLE SPACE S & A FUNCTION THAT ASSIGNS A REAL NUMBER TO EACH OUTCOME IN THE SAMPLE SPACE OF THE EXPERIMENT.

DEF 3.2 DISCRETE RANDOM VARIABLE

X IS A DISCRETE RANDOM VARIABLE IF THE RANGE OF X IS A COUNTABLE SET

$$S_X = \{ x_1, x_2, \ldots \}$$

3.2 PROBABILITY MASS FUNCTION

THE PMF OF A RANDOM VARIABLE X EXPRESSES THE PROBABILITY MODEL OF AN EXPERIMENT AS A MATHMATICAL FUNCTION. THE FUNCTION IS THE PROBABILITY

$P[X = x]$ FOR EVERY NUMBER x

PROBABILITY MODEL OF A DISCRETE RANDOM VARIABLE ASSIGNS A NUMBER BETWEEN 0 & 1 TO EACH OUTCOME IN A SAMPLE SPACE. WHEN WE HAVE A DRV X WE EXPRESS THE PROBABILITY MODEL AS A PROBABILITY MASS FUNCTION (PMF) $P_X(x)$

$$P_X(x) = P[X = x]$$

$\underline{X = x}$ IS AN EVENT CONSISTING OF ALL OUTCOMES s OF THE UNDERLYING EXPERIMENT FOR WHICH $X(s) = x$

$P_X(x)$

X IS THE NAME OF THE RANDOM VARIABLE

x IS THE POSSIBLE VALUE OF THE RANDOM VARIABLE

THEOREM 3.1

FOR A DISCRETE RANDOM VARIABLE X WITH PMF $P_X(x)$ AND RANGE S_X

(1) FOR $\forall X$, $P_X(x) \geq 0$

(2) $\sum_{x \in S_X} P_X(x) = 1$

(3) FOR ANY EVENT $B \subset S_X$, THE PROBABILITY THAT X IS IN THE SET B IS

$$P[B] = \sum_{x \in B} P_X(x)$$

THEOREM 3.3

FOR ALL $b \geq a$,

$$F_X(b) - F_X(a) = P[a < X \leq b]$$

DEF 3.13 EXPECTED VALUE

THE EXPECTED VALUE OF X IS

$$E[X] = \mu_X = \sum_{x \in S_X} x P_X(x)$$

ALL OF THE MEMBERS OF THE RANGE OF X

DEFN 3.12 MEDIAN

A MEDIAN x_{MED} OF RANDOM VARIABLE X IS A MEMBER THAT SATISFIES

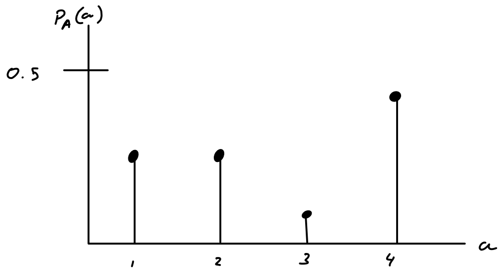
$$P[X \leq x_{MED}] \geq 1/2 \qquad P[X \geq x_{MED}] \geq 1/2$$

DEFN 3.11 MODE

FOR DISCRETE RANDOM VARIABLE X IS A NUMBER x_{MOD} SATISFYING

$P_X(x_{MOD}) \geq P_X(x)$ FOR ALL X

EXAMPLE $P_A(a)$



$E[A] = \mu_A = 1(0.25) + 2(0.25) + 3(0.1) + 4(0.4) = 2.65 \quad \notin S_A$

a_{MED} IS $\forall\, a$ S.T. $2 \leq a \leq 3$

$a_{MED} = 2$ SINCE $P[A \leq 2] = 0.5$ & $P[A \geq 2] = 0.75$

$a_{MED} = 3$ SINCE $P[A \leq 3] = 0.6$ & $P[A \geq 3] = 0.5$

FOR ALL $2 \leq a_{MOD} \leq 3$ $\quad P[A \leq a_{MOD}] = 0.5$ & $P[A \geq a_{MOD}] = 0.5$

$a_{MOD} = 4$ SINCE $P_A(4) = 0.4 \geq P_A(a) \quad a = 1, 2, 3, 4$

EECS 461 PROBABILITY & STATISTICS

DISCRETE RANDOM VARIABLES

3.5 AVERAGES & EXPECTED VALUES

MEANS OF COMMON DRV FAMILIES

THM 3.4 BERNOULLI (P) RANDOM VARIABLE X HAS CREATED VALUE $E[X] = p$

PROOF $E[X] = 0 \cdot P_X(0) + 1 P_X(1) = 0(1-p) + 1(p) = p$

THM 3.5 GEOMETRIC (p) RANDOM X HAS EXPECTED VALUE $E[X] = 1/p$

PROOF LET $q = 1-p$, THEN PMF OF X BECOMES

$$P_X(x) = \begin{cases} pq^{x-1} & x = 1, 2, \ldots \\ 0 & \text{OTHERWISE} \end{cases}$$

THE EXPECTED VALUE $E[X]$ IS THE INFINITE SUM

THM 3.7

BINOMIAL (N, p) RANDOM VARIABLE X OF DEFN 3.6

$$E[X] = NP$$

PASCAL (K, p) RANDOM VARIABLE X OF DEFN 3.7

$E[$

DISCRETE UNIFORM (K, ℓ) RANDOM VARIABLE X OF DEFN 3.8

3.6 FUNCTION OF A RANDOM VARIABLE

A FUNCTION $Y = g(X)$ OF RANDOM VARIABLE X IS ANOTHER RANDOM VARIABLE. THE PMF $P_Y(y)$ CAN BE DERIVED $P_X(x)$ & $g(X)$

SAMPLE VALUE OF $Y = g(X)$

PREFORM AN EXPERIMENT & OBSERVE AN OUTCOME S
FROM s FIND x, THE CORRESPONDING VALUE OF RANDOM VARIABLE X
OBSERVE y BY CALCULATING $y = g(x)$

$$Y = g(X)$$

EX AN INSTRUMENT 1 OF THE 4 POSSIBLE VALUES EACH TIME IT IS READ CALL THE SAMPLE VALUE DRV A & LET A HAVE THE FOLLOWING PMF

$$P_A(-1) = 0.35$$

$$P_A(1) = 0.35$$

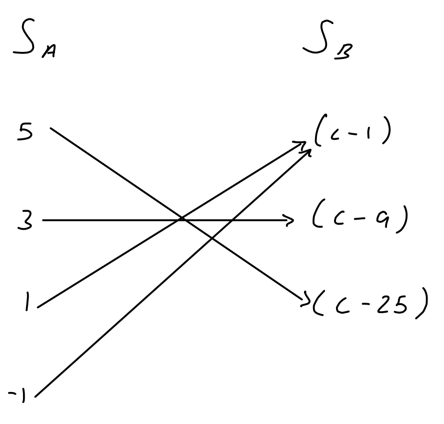
$$P_A(3) = 0.2$$

$$P_A(5) = 0.1$$

NOW ASSUME THE VALUES -1 & 1 WILL REQUIRE THE SAME TREATMENT SO COLLAPSE THEM INTO ONE VALUE BY SQUARING ALL VALUES OF a, (a^2)
ALSO ASSUME THAT WE WANT TO REVERSE THE ORDER OF THE VALUES $(-a^2)$
& ENSURE THAT THE DERIVED RV HAS MEAN (EXPECTED VALUE) = 0

$$b = g(a) = c - a^2 \quad \text{FOR SOME } c$$

PMF OF DERIVED RV
IN EXAMPLE, B HAS ONLY 3 VALUES & THE FUNCTION (MAPPING) CAN BE REPRESENTED



$P_B(b)$

$P_A(1) + P_A(-1) = 0.7$ $\quad B = C-1$ IF $A = 1$ OR $A = -1$

$P_A(3) = 0.2$ $\quad P_B(c-1) = P(A=1 \cup A=1)$

$P_A(5) = 0.1$ $\quad = P_A(1) + P_A(1)$

THM 3.9 GENERALIZES THIS

$$P_B(b) = \sum_{a:g(a)=b} P_A(a)$$

EXPECTED VALUE OF DERIVED RANDOM VARIABLE

GENERAL RESULT

WE CAN GET $E[B]$ OF DRV B (DERIVED FROM A) FROM $P_B(b)$, THE PMF OF B, BUT WE CAN ALSO GET IT WITHOUT COMPUTING THE PMF OF B

EX $$E[B] = \sum_b b\, P_B(b) = (c-1)(0.7) + (c-9)(0.2) + (c-25)(0.1)$$

WANT $E[B] = 0$

$$E[B] = 0 = c - 0.7 - 1.8 - 2.5 \Rightarrow c = 5$$

$$E[B] = 4(0.7) + (-4)(0.2) + (-20)(0.1) = 0$$

BUT NOTICE THAT WE CAN ALSO COMPUTE $E[B]$ AS:

$$E[B] = (5-(-1)^2)(0.35) + (5-(1)^2)(0.35) + (5-3^2)(0.2) + (5-5^2)(0.1) = 0$$

THM 3.10: $$E[B] = \mu_Y = \sum_{x \in S_X} g(x) P_X(x)$$

3.7 EXPECTED VALUE OF A DERIVED RANDOM VARIABLE

AFFINE TRANSFORMATION

THEOREM FOR ANY RANDOM VARIABLE X

$$E[aX + b] = aE[X] + b$$

EECS 461 PROBABILITY & STATISTICS

DISCRETE RANDOM VARIABLES

3.5 AVERAGES & EXPECTED VALUES

QUIZ 3.6

MONITOR THREE CUSTOMERS PURCHASING SMARTPHONES AT PHONESMART STORE & OBSERVE WHETHER EACH BUYS AN APRICOT PHONE FOR \$450 OR A BANANA PHONE FOR \$300. THE RANDOM VARIABLE N IS THE NUMBER OF CUSTOMERS PURCHASING AN APRICOT PHONE ASSUME N HAS A PMF

$$P_N(n) = \begin{cases} 0.4 & n = 0 & \leftarrow 450N \\ 0.2 & n = 1, 2, 3 & \leftarrow 300(3-N) \\ 0 & \text{OTHERWISE} \end{cases}$$

m DOLLARS IS THE AMOUNT OF MONEY PAID BY THREE CUSTOMERS

(A) EXPRESS M AS A FUNCTION OF N

AS A FUNCTION OF N, THE MONEY SPENT BY THE THREE CUSTOMERS IS,

$$M = \underset{\text{APRICOT}}{450N} + \underset{\text{BANANA}}{300(3-N)} \qquad m = 450N + 300(3-N)$$

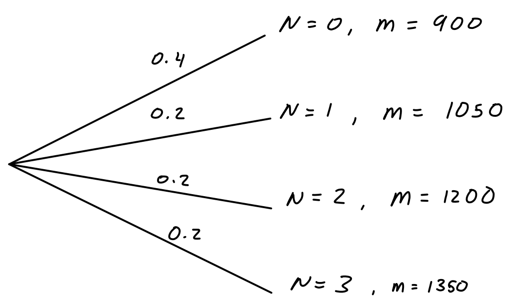
$$= 900 - 300N + 450N$$

$$= 900 + 150N$$

(B) FIND $P_M(m)$ & $E(M)$

TO FIND THE PMF OF M, WE CAN DRAW THE FOLLOWING TREE & MAP THE OUTCOMES TO THE VALUES OF M:

$$m = 450(N) + 300(3-N)$$



$$m = 450(0) + 300(3-0) = 900$$

$$m = 450(1) + 300(3-1) = 450 + 600 = 1050$$

$$m = 450(2) + 300(3-2) = 900 + 300 = 1200$$

$$m = 450(3) + 300(3-3) = 1350$$

REPRESENTED BY THE PMF

$$P_M(m) = \begin{cases} 0.4 & m = 900 \\ 0.2 & m = 1050, 1200, 1250 \\ 0 & \text{OTHERWISE} \end{cases}$$

$$E[M] = 900\, P_M(900) + 1050\, P_M(1050) + 1200\, P_M(1200) + 1350\, P_M(1350$$