5.1 Discrete Random Variable

Definition. A **random variable** is a variable which values are determined by the outcome of a random experiment.

Notation. Capital Latin letters at the end of the alphabet X, Y, Z, T, W, V.

Example. A coin is flipped twice. The sample space is $S = \{HH, HT, TH, TT\}$. Let X be the number of heads that appear. $X(HH)=2$, $X(HT)=X(TH)=1$, *and* $X(TT)=0$.

Note that X assumes a numeric value for each outcome, and thus is a random variable.

Definition. A **discrete random variable** assumes a finite or countably infinite number of values.

Definition. A **continuous random variable** assumes values in an interval.

5.2 Probability Distribution of a Discrete Random Variable

Definition. The **probability distribution** of a **discrete** random variable is the list of all possible values for that variable and their corresponding probabilities.

Notation. Values of a random variable X are denoted by x, and the probability distribution of X is denoted by $P(x)=P(X=x)$.

Properties of Probability Distribution

1. $0 \le P(x) \le 1$, for each x,

and

2. $\Sigma \mathrm{P}(x) = 1.$

Example. A fair coin is flipped twice. Find the probability distribution of the number of heads.

Solution. $P(0)=P(X=0)=P(TT)=P(T)P(T)$ $=(1/2)(1/2)=1/4$, $P(1)= P(X=1)=P(HT \text{ or } TH)=P(HT)+P(TH)=1/4+1/4=1/2$, $P(2)=P(X=2)=P(HH)=1/4$.

Definition. A **probability distribution table** contains a column of x's and a column of $P(x)$'s.

Example. In our example,

x	$P(x)$
0	1/4
1	1/2
2	1/4

Example. Check that the table below is the **probability distribution table**.

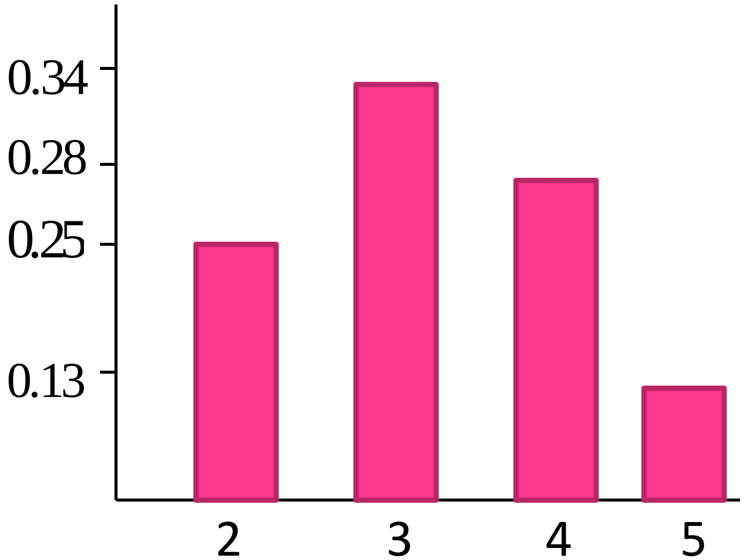
x	***P*(*x*)**
2	0.25
3	0.34
4	0.28
5	0.13

We check that (i) $0 \le P(x) \le 1$, for each x,

and (ii) $\Sigma P(x) = 1$

Definition. A **graphical presentation of a probability distribution** is a bar graph with bar heights corresponding to probabilities.

Example. In the previous example,



5.3 Mean of a Discrete Random Variable

Definition. The **mean** (or **expected value** or **expectation**) of a discrete random variable X is

$$\mu = E(X) = \sum_{x} x \cdot P(X = x)$$

The interpretation of the mean is the average value of X when an experiment is repeated many times.

Example 1. The probability distribution table is

x	$P(x)$
2	0.25
3	0.34
4	0.28
5	0.13

$$\mu = E(X) = (2)(0.25) + (3)(0.34) + (4)(0.28) + (5)(0.13) = 3.29.$$

Example 2. Bob is offered to play the following game. A fair coin is flipped three times. If exactly two heads appear, Bob is paid $10. Otherwise, he has to pay $7. Should he agree to play this game?

Solution. The sample space for this random experiment is

$$S = \{HHH, HHT, HTH, THH, HTT, THT, TTH, TTT\}$$

Let X be Bob's gain. The values of X are

$X(HHT)=X(HTH)=X(THH)=$ \$10, and

$X(HHH)=X(HTT)=X(THT)=X(TTH)=X(TTT)=$ - \$7.

Therefore, $P(X=\$10)=3/8$, $P(X=-\$7)=5/8$.

Bob's expected gain is

$$\boldsymbol{E}(\boldsymbol{X}) = (\$10)(3/8) + (-\$7)(5/8) = \$(30-35)/8$$

$$= -\$5/8 = -\$0.625.$$

So, on average, Bob would be losing 62.5 cents. He should **NOT** agree to play.

Question. When should Bob agree to play?

Answer. When his expected gain is non-negative (he doesn't expect to lose money).

BUT

If his mean gain is strictly positive, no one would want to play with him, because they would lose money, on average.

SO…….

The only way both sides would agree to play is when the expected gain of both sides is zero.

Definition. A **fair game** is a game in which the expected gain of both sides is zero.

Question. In our setting, Bob will be paid \$10, but how much should he pay to make it a fair game?

Answer. Bob's expected gain should be zero, that is, $E(X) = (\$10)(3/8) + (\$?)(5/8) = \$0$

$\$? = -\$6$

So, Bob should be paying \$6 instead of \$7.

Example 3. A state lottery awards at random, for each 100,000 one-dollar tickets sold,

1 – $10,000 prize,

18 – $200 prizes,

120 – $25 prizes,

270 – $20 prizes.

What is the expected value of the winnings of one ticket in this lottery? Do we want to play?

Solution. Let W be the winning of one ticket. The probability distribution of W is

$P(W=\$10,000)=1/100,000$, $P(W=\$200)=18/100,000$, $P(W=\$25) = 120/100,000$, $P(W=\$20)=270/100,000$, $P(W=\$0)$ = everything else.

Then the expected value of W is

$$E(W)=\frac{1}{100,000}\left[(\$10,000)(1)+(\$200)(18)+(\$25)(120)+(\$20)(270)\right]$$

$$=\$0.22=22\text{cents.}$$

So, we pay \$1 to buy a ticket but expect to win only 22 cents. We don’t want to play.

Example 4. A grab-bag contains 6 packages worth \$2 each, 11 packages worth \$3 each, and 8 packages worth \$4 each. Find the expected winnings of one package.

Solution. Let W be the winnings of one package. Then $P(W=\$2)=6/(6+11+8)=6/25$, $P(W=\$3)=11/25$, $P(W=\$4)=8/25$. Hence,

$$E(W) = (\$2)(6/25) + (\$3)(11/25) + (\$4)(8/25) = \$3.08.$$

Is it reasonable to pay $3.50 for the option of selecting one of these packages at random?

It is unfair to pay $3.50. To make it a fair game, the cost should be $3.08.

5.3 Standard Deviation of a Discrete Random Variable

Definition. The **variance** of a discrete random variable *X* is

$$\sigma^2 = \sum_x (x-\mu)^2 P(X = x).$$

The computational formula is

$$\sigma^2 = \sum_x x^2 P(X = x) - \mu^2.$$

Definition. The **standard deviation** of a discrete random variable X is

$$\sigma = \sqrt{\sigma^2}.$$

Example 1 (continued…). The mean $\mu = 3.29$, hence, the variance and standard deviation are

$$\sigma^2 = (2)^2(0.25) + (3)^2(0.34) + (4)^2(0.28) + (5)^2(0.13)$$
$$-3.29^2 = 0.97, \text{ and } \sigma = \sqrt{0.97} = 0.98.$$

Example 2 (continued…). In a fair game, the mean $E(X)$=\$0. Therefore,

$$\sigma^2 = (\$10)^2(3/8) + (-\$6)^2(5/8) - (\$0)^2 = \$^2 60, \text{ and}$$

$$\sigma = \sqrt{\$^2 60} = \$7.75.$$

Example 3 (continued…). $E(W)$=\$0.22,

$$\sigma^2 = \frac{1}{100{,}000}[(\$10{,}000)^2(1) + (\$200)^2(18) + (\$25)^2(120)$$

$$+ (\$20)^2(270)] - (\$0.22)^2 = \$^2 1008.98, \text{ and}$$

$$\sigma = \sqrt{\$^2 1008.98} = \$31.76.$$

<u>Example 4</u> (continued…). The variance and **standard deviation** of winnings of one package are

$$\sigma^2 = (\$2)^2(6/25) + (\$3)^2(11/25) + (\$4)^2(8/25)$$

$$-(\$3.08)^2 = (\$)^2 0.5536, \text{ and } \sigma = \sqrt{(\$)^2 0.5536} = \$0.74.$$