Section 2.6 – Discrete Random Variables

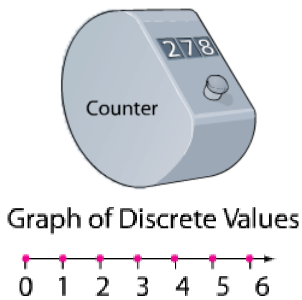
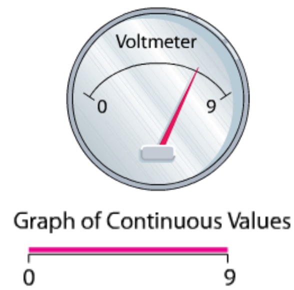
Defintions

A ***random variable*** is a variable (typically represented by *x*) that has a single numerical value, determined by chance, for each outcome of a procedure

A ***discrete random*** variable hass either a finite number of values or countable number of values, where "countable" refers to the fact that there might be infinitely many values, but they result from a counting process

A ***continuous random*** variable has infinitely many values, and those values can be associated with measurements on a continuous scale without gaps or interruptions

A ***Probability distribution*** is a description that gives the probability for each value of the random variable; often expressed in the format of a graph, table, or formula

<table>
<tr>
<td>
</td>
<td>
</td>
</tr>
<tr>
<td>Discrete Random Variable: Count of the number of movies patrons.</td>
<td>Continuous Random Variable: The measured voltage of a smoke detector battery</td>
</tr>
</table>

Example

The following are example of discrete and continuous radom variables.

1. ***Discrete***: Let x = the number of eggs that a hen lays in a day. This is a *discrete* random variable because its only possible values are 0, or 1, or 2, and so on. No han can lay 2.343115 eggs, which would have been possible if the data had come from a continuous scale.
2. ***Discrete***: The count of the number of statistics students present in class on a given day is a whole number and is therefore a discrete random variable. The counting device is capable of indicating only a finite number of values, so it is used to obtain values for a *discrete* random variable.
3. ***Continuous***: Let x = the amount of milk a cow produces in one day. This is a *continuous* random variable because it can have any value over a continuous span. During a single day, a cow might yield an amount of milk that can be any value between 0 gallon and 5 gallons. It owould be 4.123 gallons, because the coe is not restricted to the discrete amounts of 0, 1, 2, 3, 4, or 5 gallons.
4. ***Continuous***: The measure of voltage for a particular smoke detector battery can be any value between 0 volt and 9 volts. It is therefore a *continuous* random variable.

Requirements for Probability Distribution

- ✓ $\sum P(x) = 1$ where x assumes all possible values. The sum of all probailities must be 1. (but sunch 0.999 ot 1.001 are acceptable)
- ✓ $0 \le P(x) \le 1$ for every individual value of x

Example

Is the following a probability distribution?

x	$P(x)$
0	0.16
1	0.18
2	0.22
3	0.10
4	0.30
5	-0.01

No. $P(x = 5) = -0.01$

x	$P(x)$
0	0.001
1	0.015
2	0.088
3	0.264
4	0.396
5	0.237

Yes. $\sum P(x) = 1.001 \approx 1$

x	$P(x)$
0	0.16
1	0.18
2	0.22
3	0.10
4	0.30
5	0.01

No. $\sum P(x) = 0.97 < 1$

Example

Does $P(x) = \frac{x}{10}$ (where x can be 0, 1, 2, 3, or 4) determine a probability distribution?

Solution

$$P(0) = \frac{0}{10} = 0; \quad P(1) = \frac{1}{10}; \quad P(2) = \frac{2}{10}; \quad P(3) = \frac{3}{10}; \quad \text{and} \quad P(4) = \frac{4}{10}$$

$$\begin{aligned}\sum P(x) &= P(0) + P(1) + P(2) + P(3) + P(4) \\ &= \frac{0}{10} + \frac{1}{10} + \frac{2}{10} + \frac{3}{10} + \frac{4}{10} \\ &= \frac{10}{10} \\ &= 1\end{aligned}$$

Each value of the $P(x)$ is beteween 0 and 1.

Because both requirements are satisfied, the formula given is a probability distribution.

Mean, Variance and Standard Deviation of a Probability Distribution

$\mu = \sum [x \cdot P(x)]$	Mean of Discrete random variable
$\sigma^2 = \sum \left[(x-\mu)^2 \cdot P(x) \right]$	Variance for a probability distribution (easier to understand)
$\sigma^2 = \left[\sum x^2 \cdot P(x) \right] - \mu^2$	Variance for a probability distribution (easier computations)
$\sigma = \sqrt{\left[\sum x^2 \cdot P(x) \right] - \mu^2}$	Standard deviation for a probability distribution

Example

Find the mean, variance, and standard deviation for the probaiblity distribution described in the table.

x (# of peas)	$P(x)$
0	0.001
1	0.015
2	0.088
3	0.264
4	0.396
5	0.237

Solution

Mean: $\mu = \sum [x \cdot P(x)]$

$$= 0(0.001) + 1(0.015) + 2(0.088) + 3(0.264) + 4(0.396) + 5(0.237)$$

$$\underline{= 3.752}$$

Variance: $\sigma^2 = \sum \left[(x-\mu)^2 \cdot P(x) \right]$

$$= (0-3.8)^2(0.001) + (1-3.8)^2(0.015) + (2-3.8)^2(0.088) + (3-3.8)^2(0.264)$$
$$+ (4-3.8)^2(0.396) + (5-3.8)^2(0.237)$$

$$\underline{= 0.940574}$$

Standard deviation: $\sigma = \sqrt{0.940574}\ \underline{= 0.9698}$

Expected Value

The expected value of a discrete random variable is denoted by E, and it represents the mean value of the outcomes. It is obtained by finding the value of $\sum\left[x\cdot P(x)\right]$

$$E=\sum\left[x\cdot P(x)\right]$$

Example

You areconsidering placing a bet either on the number 7 in roulette or on the "pass line" in the dice game of craps at the casino.

a) If you bet \$5 on the number 7 in roulette, the probability of losing \$5 is $\frac{37}{38}$ and the probability of making a net gain of \$175 is $\frac{1}{38}$. (The prize id \$180, including your \$5 bet, so the net gain is \$175.) Find your expected value if you bet \$5 on the number 7 in roulette.

b) If you bet \$5 on the pass line in the dice game of craps, the probability of losing \$5 is $\frac{251}{495}$ and the probability of making a net gain of \$5 is $\frac{244}{495}$. (If you bet \$5 on the Pass Line and win, you are given \$10 that includes your bet, so the net gain is \$5.) Find your expected value if you bet \$5 on the pass line in the dice game? Why?.

Solution

a)

Event	x	$P(x)$	$x\cdot P(x)$
Lose	–\$5	$\frac{37}{38}$	–\$4.87
Gain (net)	\$175	$\frac{1}{38}$	\$4.61
Total			–\$0.26 ***Or*** –26¢

You can expect to lose an average of **26¢**.

b)

Event	x	$P(x)$	$x\cdot P(x)$
Lose	–\$5	$\frac{251}{495}$	–\$2.54
Gain (net)	\$5	$\frac{244}{495}$	\$2.46
Total			–\$0.08 ***Or*** –8¢

You can expect to lose an average of 8¢.

Exercises *Section* 2.6 – Discrete Random Variables

1. Determine whether or not a probability distribution is given. If a probability is given, find its mean and standard deviation. If the probability is not given, identify the requirements that are not satisfied.

a)

x	$P(x)$
0	0.125
1	0.375
2	0.375
3	0.125

b)

x	$P(x)$
0	0.22
1	0.16
2	0.21
3	0.16

c)

x	$P(x)$
0	0.528
1	0.360
2	0.098
3	0.013
4	0.001
5	0^+

0^+ *denotes a positive probability value that is very small.*

d)

x	$P(x)$
0	0.02
1	0.15
2	0.29
3	0.26
4	0.16
5	0.12

2. Based on past results found in the *Information Please Almanac*, there is a 0.1919 probability that a baseball World Series context will last 4 games, is a 0.2121 probability that it will last 5 games, a 0.2222 probability that it will last 6 games, a 0.3737 probability that us will last 7 games.

a) Does the given information describe a probability distribution?

b) Assuming that the given information describes a probability distribution, find the mean and standard deviation for the numbers of games in World Series contests.

c) Is it unusual for a team to "sweep" by winning in four games? Why or why not?

3. Based on information from MRI Network, some job applicants are required to have several interviews before a decision is made. The number of required interviews and the corresponding probabilities are: 1 (0.09); 2 (0.31); 3 (0.37); 4 (0.12); 5 (0.05); 6 (0.05).

a) Does the given information describe a probability distribution?

b) Assuming that a probability distribution is described, find its mean and standard deviation.

c) Use the range rule of thumb to identify the range of values for usual numbers of interviews.

d) Is it unusual to have a decision after just one interview? Explain?

4. Based on information from Car dealer, when a car is randomly selected the number of bumper stickers and the corresponding probabilities are: 0 (0.824); 1 (0.083); 2 (0.039); 3 (0.014); 4 (0.012); 5 (0.008); 6 (0.008); 7 (0.004); 8 (0.004); 9 (0.004).

a) Does the given information describe a probability distribution?

b) Assuming that a probability distribution is described, find its mean and standard deviation.

c) Use the range rule of thumb to identify the range of values for usual numbers of bumper stickers.

d) Is it unusual for a car to have more than one bumper sticker? Explain?

5. A Company hired 8 employees from a large pool of applicants with an equal numbers of males and females. If the hiring is done without regard to sex, the numbers of females hired and the corresponding probabilities are: 0 (0.004); 1 (0.0313); 2 (0.109); 3 (0.219); 4 (0.273); 5 (0.219); 6 (0.109); 7 (0.031); 8 (0.004).

a) Does the given information describe a probability distribution?

b) Assuming that a probability distribution is described, find its mean and standard deviation.

c) Use the range rule of thumb to identify the range of values for usual numbers of females hired in such groups of 8.

d) If the most recent group of 8 newly hired employees does not include any females, does there appear to be discrimination based on sex? Explain?

6. Let the random variable x represent the number of girls in a family of 4 children. Construct a table describing the probability distribution; then find the mean and the standard deviation. (Hint: List the different possible outcomes.) Is it unusual for a family of 3 children to consist of 3 girls?

7. In 4 lottery game, you pay 50¢ to select a sequence of 4 digits, such 1332. If you select the same sequence of 4 digits that are drawn, you win and collect \$2788.

a) How many different selections are possible?

b) What is the probability of winning?

c) If you win, what is your net profit?

d) Find the expected value.

8. When playing roulette at casino, a gambler is trying to decide whether to bet \$5 on the number 13 or bet \$5 that the outcomes any one of these 5 possibilities: 0 or 00 or 1 or 2 or 3.the expected value of the \$5 bet for a single number is –26¢. For the \$5 bet that the outcome 0 or 00 or 1 or 2 or 3, there is a probability of $\frac{5}{38}$ of making a net profit of \$30 and a $\frac{33}{38}$ probability of losing \$5.

a) Find the expected value for the \$5 bet that the outcome is 0 or 00 or 1 or 2 or 3.

b) Which bet is better: A \$5 bet on the number 13 or a \$5 bet the outcome is 0 or 00 or 1 or 2 or 3? Why?

9. There is a 0.9986 probability that a randomly selected 30-year-old male lives through the year. As insurance company charges \$161 for insuring that the male will live through the year. If the male does not survive the year, the policy pays out \$100,000 as a death benefit.

a) From the perspective of the 30-year-old male, what are the values corresponding to the 2 events of surviving the year and not surviving?

b) If a 30-year-old male purchases the policy, what is his expected value?

c) Can the insurance company expect to make a profit from many such policies? Why?

10. An insurance company charges a 21-year-old male a premium of \$500 for a one-year \$100,000 life insurance policy. A 21-year-old male has a 0.9985 probability of living for a year.

a) From the perspective of a 21-year-old male (or estate), what are the values of the two different outcomes?

b) What is the expected value for a 21-year-old male who buys the insurance?

c) What would be the cost of the insurance if the company just breaks even (in the long run with many such policies), instead of making a profit?

d) Given that the expected value is negative (so the insurance company can make a profit), why