Section 5: Discrete Probability Distributions

Random Variable – A variable that cannot be predicted with certainty whose values occur according to a frequency distribution

Two Types of Random Variables

First, it is important to make sure that we distinguish between a discrete and continuous random variable. While the idea of a probability distribution is the same in both cases, the way we have to look at the probability distribution is different. We will focus on the idea of probability for a discrete random variable here.

Discrete Random Variable – potential values are separated points on the number line; there is a finite or countable number of values (outcomes are typically counted)

Examples
- The number of students in a statistics course
- The number of times you must flip a coin in order to get a tail

Continuous Random Variable – potential values fall in an interval on the number line; there is an infinite number of values (outcomes are typically measured)

Examples
- The height of a person
- The distance a person travels to get to work

A discrete random variable will include a finite or countable number of outcomes. Therefore, we assign probabilities to each value of the random variable.

Probability Distributions for Discrete Random Variables

Probability distribution – the collection of $P(x)$ values for all possible x values

For a discrete probability distribution to be valid, it must satisfy the following:
1. The variable, x, must be discrete (there is a finite or countable number of outcomes)
2. $0 \le P(x) \le 1$ for all x (all probabilities are between 0 and 1)
3. $\sum P(x) = 1$ (the sum of all the probabilities is 1)

Example
Flip a coin twice
x = number of tails
Identify the probability distribution for x

Answer
First we should identify the possible outcomes when flipping a coin twice. The four possible outcomes are {HH, HT, TH, TT} (could also do this with a tree diagram). Next we should think about the probability of each. In this case H and T have equal probability so the probability for the outcomes are the same, each is ¼. You can also figure this out using the multiplication rule. For example, getting HH is (½)(½) = ¼. If the probabilities are not equal, then the multiplication rule must be used to find the probabilities. Now that we understand the probabilities, we can deal with the random variable, *x*. The number of tails we will get when flipping a coin twice is either 0, 1, or 2. The probability of 0, P(*x* = 0), = ¼ because the only way to get 0 tails is to get HH. The probability of 1, P(*x* = 1), = ½ because there are two ways to get 1 tail, HT or TH. Combining these probabilities gives us ¼ + ¼ = ½. The probability of 2, P(*x* = 2), = ¼ because this would mean we got TT. The probability distribution follows:

x	P(*x*)
0	¼
1	½
2	¼

Notice that the probability distribution above satisfies the criteria needed for a valid discrete probability distribution. The random variable, *x*, is discrete (finite number of outcomes). Each probability is between 0 and 1 (any time you calculate a probability, it is important to ensure that it is between 0 and 1). The sum of the probabilities in the distribution is 1, ¼ + ½ + ¼ = 1 (any time you create a distribution you should check and make sure the sum is 1, if not there is an error)

The Mean and Standard Deviation for Discrete Random Variables

Expected Value – the mean of a probability distribution (This value represents the same quantity as the mean did in section 2. The difference in the calculation comes from the fact that outcomes in a discrete probability distribution are not equally likely in most cases. Thus, this calculation is a weighted average. If all probabilities are the same then the calculation from section 2 will give the same outcome. The other big difference here as compared to what we looked at in section 2 is that we are now dealing with a population mean and not a sample mean. In a probability distribution, we have all of the possible outcomes meaning that we are describing the population.)

The population mean for a discrete probability distribution is:
$$\mu = \sum xP(x)$$

μ is a lowercase Greek letter said as 'mu' (Can be any real number)

The difference in notation for the mean and standard deviation is important. We are now dealing with population values. It is crucial to remember the notation for both the sample and population because later in the course you will see population and sample values in the same formula. (in statistics, most population values are denoted by Greek letters and most sample values are denoted with our alphabet)

As always, in order to calculate standard deviation, we must start by calculating the variance. The population variance for a discrete probability distribution is:

$$\sigma^2 = \sum (x-\mu)^2 P(x) = \sum x^2 P(x) - \mu^2$$

The population standard deviation is the square root of the variance.

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

σ is a lowercase Greek letter said as 'sigma' (Can be any nonnegative real number)

As before, when dealing with variance, there are two different versions of the formula. The conceptual formula, $\sum (x-\mu)^2 \cdot P(x)$, and the computational formula, $\sum x^2 \cdot P(x) - \mu^2$. The conceptual formula comes directly from the definition. It still represents the average squared distance the data points are from the mean. Here, however, the variance is a weighted average. Comparing the conceptual formula for variance, $\sigma^2 = \sum (x-\mu)^2 \cdot P(x)$, to the formula for the mean, $\mu = \sum x \cdot P(x)$, you can see that the formulas are similar. The difference is for the mean we are averaging x, and for the variance we are averaging $(x-\mu)^2$, squared distance from the mean. The mean and variance formulas are the same except for these quantities. The computational formula, $\sum x^2 \cdot P(x) - \mu^2$, is an algebraic manipulation of the conceptual formula. This is the formula I would recommend using when doing these problems by hand. The algebra is presented in the example below.

Example

Prove that $\sum (x-\mu)^2 \cdot P(x) = \sum x^2 \cdot P(x) - \mu^2$ (Show algebraic steps)

Answer

$$\begin{aligned}
&\sum (x-\mu)^2 \cdot P(x) \\
&= \sum (x^2 - 2\mu x + \mu^2) \cdot P(x) \\
&= \sum \left(x^2 \cdot P(x) - 2\mu x \cdot P(x) + \mu^2 \cdot P(x)\right) \\
&= \sum x^2 \cdot P(x) - \sum 2\mu x \cdot P(x) + \sum \mu^2 \cdot P(x) \\
&= \sum x^2 \cdot P(x) - 2\mu \sum x \cdot P(x) + \mu^2 \sum P(x) \\
&= \sum x^2 \cdot P(x) - 2\mu \cdot \mu + \mu^2 \cdot 1 \\
&= \sum x^2 \cdot P(x) - 2\mu^2 + \mu^2 \\
&= \sum x^2 \cdot P(x) - \mu^2
\end{aligned}$$

Example

What happens if we take the squared off the variance (think about our discussion of variance in section 3)

Simplify $\sum (x-\mu) \cdot P(x)$

Answer

$\sum(x-\mu)\cdot P(x)$
$=\sum\left(x\cdot P(x)-\mu\cdot P(x)\right)$
$=\sum x\cdot P(x)-\sum\mu\cdot P(x)$
$=\sum x\cdot P(x)-\mu\sum P(x)$
$=\mu-\mu\cdot 1$
$=\mu-\mu$
$=0$

Example

Roll a die and *x* is the number rolled.

a. Find the mean of *x*.
b. Find the variance of *x*.
c. Find the standard deviation of *x*.

Answers

In order to find the mean and standard deviation of a discrete probability distribution, I would recommend that you create a table like the one below. The first two columns of the table, x and $P(x)$, represent the probability distribution for the random variable. The other three columns, $x\cdot P(x)$, x^2, and $x^2\cdot P(x)$, are used to find the mean and variance.

x	$P(x)$	$xP(x)$	x^2	$x^2P(x)$
1	$1/6$	$1/6$	1	$1/6$
2	$1/6$	$2/6$	4	$4/6$
3	$1/6$	$3/6$	9	$9/6$
4	$1/6$	$4/6$	16	$16/6$
5	$1/6$	$5/6$	25	$25/6$
6	$1/6$	$6/6$	36	$36/6$
		$21/6=3.5$		$91/6=15.167$

a. $\mu=\sum x\cdot P(x)=3.5$
b. $\sigma^2=\sum x^2\cdot P(x)-\mu^2=15.167-(3.5)^2=15.167-12.25=2.917$
c. $\sigma=\sqrt{\sigma^2}=\sqrt{2.917}=1.708$

Example

This is a realistic example where at least the mean is an important value. Typically, standard deviation is not of much use in these situations. However, it is used as a unit of measure in more complex statistical analysis. It is important you understand the idea so that you will understand how standard deviation is used in more advanced analysis, which will be done in the coming material.

Suppose we sell life insurance policies. We will look at a $100,000 one year term policy. The random variable of interest in this case is x = payout of the policy. The payout is $100,000 if death occurs in the next year and $0 otherwise. When selling life insurance, a company figures out the probability of payout based on life expectancy. This is typically estimated based on a medical examination and/or a medical questionnaire. Suppose the policy has the following probability distribution.

x	$P(x)$
$0	0.995
$100,000	0.005

a. Find the mean of x.
b. What does the mean tell us in this example?
c. Find the variance of x.
d. Find the standard deviation of x.

Answers

x	$P(x)$	$xP(x)$	x^2	$x^2P(x)$
0	.995	0	0	0
100,000	.005	500	10,000,000,000	50,000,000
		500		50,000,000

a. $\mu = \sum x \cdot P(x) = \$500$
b. This is the average payout (break-even point on the policy not including overhead costs). In order to make money on these term life insurance policies you must charge more than $500.
c. $\sigma^2 = \sum x^2 \cdot P(x) - \mu^2 = 50{,}000{,}000 - (500)^2 = 49{,}750{,}000$
d. $\sigma = \sqrt{49{,}750{,}000} = 7053.37$

Thus far in this section we have introduced the general idea of a discrete probability distribution. It is important to note that all discrete probability distributions satisfy the characteristics discussed and the mean and standard deviation can always be calculated using the formulas presented. However, the calculations can sometimes become very cumbersome and difficult. Therefore, in statistics, it is common to classify discrete distributions based on their common properties. There are many different types of discrete probability distributions and we will just introduce one of them to illustrate this idea. If you know the properties of a distribution, there is often times shortcut formulas for the probability, mean, variance, and standard deviation.

The Binomial Probability Distribution

A binomial probability distribution is a common type of discrete probability distribution.

A binomial experiment has the following properties:

- Experiment consists of n identical and independent trials
- Each trial results in one of two outcomes "success" = S or "failure" = F
- $P(S) = p$, $P(F) = 1 - p = q$ for all trials
- The random variable of interest x is the number of successes in the n trials

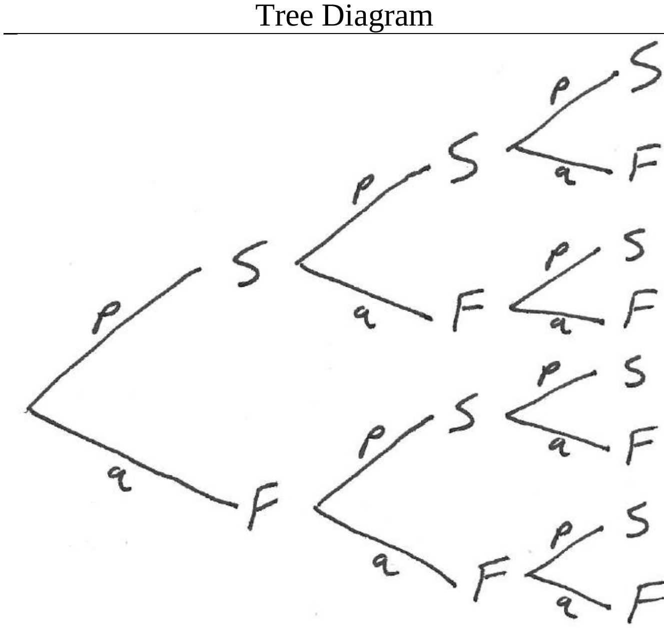
Example

Flip a fair coin 10 times and x = number of tails

$n = 10$
"success" → S = tails (you must define a success based on the experiment)
"failure" → F = heads (here we are interested in getting tails so that is a success)
$p = 1/2$
$q = 1/2$

Illustration of the Binomial Probability Formula

For this illustration, we will assume $n = 3$. We will look at the sample space in terms of successes and failures utilizing a tree diagram. On the branches of the tree diagram we will include p and q to represent the probability of a success and failure, respectively. Now with the sample space, we can identify probability. Keep in mind that we do not know p and q so we cannot assume the outcomes are equally likely. This would only happen if p = ½. Therefore, we must use the multiplication rule as seen below.

Tree Diagram	Sample Space	Probability
	SSS	$p \cdot p \cdot p = p^3$
	SSF	$p \cdot p \cdot q = p^2 q$
	SFS	$p \cdot q \cdot p = p^2 q$
	SFF	$p \cdot q \cdot q = pq^2$
	FSS	$q \cdot p \cdot p = p^2 q$
	FSF	$q \cdot p \cdot q = pq^2$
	FFS	$q \cdot q \cdot p = pq^2$
	FFF	$q \cdot q \cdot q = q^3$

The next step is to look at the probability distribution of x. With $n = 3$, the possible outcomes for x (number of successes) would be 0, 1, 2, or 3. The probability for each outcome is included below.

$$P(x=0) = 1 \cdot p^0 q^3$$
$$P(x=1) = 3 \cdot p^1 q^2$$
$$P(x=2) = 3 \cdot p^2 q^1$$
$$P(x=3) = 1 \cdot p^3 q^0$$

Notice that there is a pattern to how these probabilities work. Each probability starts with a constant representing the number of outcomes from the tree diagram which satisfy the value of x (for example when x is 1, there are 3 ways to get exactly 1 success). Thus, the first part of the binomial probability formula gives the number of possible arrangements of S and F. The next thing in each probability is p to some power. Notice the power on p is always equal to x. So, the next part of the binomial probability formula is p^x. Finally, each probability has a q to some power. The power on q is always $n - x$. Therefore, the last part of the binomial probability formula is q^{n-x}. Knowing this pattern gives us a shortcut way to find probability when we know an experiment follows the properties of a binomial experiment. The formula is below.

<u>Binomial probability formula</u>

$$P(x) = \frac{n!}{x!(n-x)!} p^x q^{n-x}$$

<u>Example</u>
Flip a fair coin 10 times and x = number of tails
Find the probability that you get exactly 7 tails

<u>Answer</u>

$$n = 10$$
$$p = \tfrac{1}{2}$$

$$P(x=7) = \frac{n!}{x!(n-x)!} p^x q^{n-x}$$
$$= \frac{10!}{7!(10-7)!} \left(\tfrac{1}{2}\right)^7 \left(\tfrac{1}{2}\right)^{10-7}$$
$$= \frac{10!}{7! \cdot 3!} \left(\tfrac{1}{2}\right)^7 \left(\tfrac{1}{2}\right)^3$$
$$= 120 \left(\tfrac{1}{2}\right)^{10}$$
$$= 0.117$$

About 11.7% of the time you will get 7 tails out of 10 flips

Example
Roll a fair die 5 times and x = # of sixes
Find the probability of getting exactly 2 sixes

Answer
$n = 5$
$p = 1/6$

$$P(x=2) = \frac{n!}{x!(n-x)!} p^x q^{n-x}$$
$$= \frac{5!}{2!(5-2)!} \left(1/6\right)^2 \left(5/6\right)^{5-2}$$
$$= \frac{5!}{2! \cdot 3!} \left(1/6\right)^2 \left(5/6\right)^3$$
$$= 10 \cdot \left(1/6\right)^2 \left(5/6\right)^3$$
$$= 0.161$$

Approximately 16.1% of the time we will get exactly 2 sixes out of 5 rolls

Example
Suppose a cancer treatment produces remission in ¼ of patients. If three patients are currently using the treatment, what is the probability exactly 2 go into Remission?

Answer
$n = 3$
$p = 1/4$

$$P(x=2) = \frac{n!}{x!(n-x)!} p^x q^{n-x}$$
$$= \frac{3!}{2!(3-2)!} \left(1/4\right)^2 \left(3/4\right)^{3-2}$$
$$= \frac{3!}{2! \cdot 1!} \left(1/4\right)^2 \left(3/4\right)$$
$$= 3\left(1/16\right)\left(3/4\right)$$
$$= .1406$$

About 14.1% of the time exactly 2 of the 3 patients would go into remission.

Example

Suppose x is binomial with $n = 8$ and $p = .4$

Find $P(x \leq 6)$

Answer

First, notice this is different than the rest of the probability questions that were asked because this one is an inequality. We want anything less than or equal to 6 in this case. There are two ways to do this problem. You could calculate it directly but 0, 1, 2, 3, 4, 5, and 6 all satisfy the inequality. Therefore, to do this problem directly you would have to calculate all these probabilities and add them together. That would be a lot of work. There is an easier way to do this one. Any time you are calculating probability and it seems difficult, you should think about the complement as it could be easier. With $n = 8$ and the complement being $P(x > 6)$, only two probabilities would have to be calculated to get the complement, 7 and 8. Therefore, it is much easier here to calculate the complement and then subtract from 1 (complement rule) to get the probability. Any time you are dealing with these inequalities, consider the direct route and the complement. This will allow you to choose the easier process.

$$P(x = 7) = \frac{n!}{x!(n-x)!} p^x q^{n-x} = \frac{8!}{7!(8-7)!} (.4)^7 (.6)^1 = 8(.4)^7(.6) = .007864$$
$$P(x = 8) = \frac{n!}{x!(n-x)!} p^x q^{n-x} = \frac{8!}{8!(8-8)!} (.4)^8 (.6)^0 = (.4)^8 = .000655$$

$$P(x > 6) = P(x = 7) + P(x = 8) = .007864 + .000655 = .008519$$

$$P(x \leq 6) = 1 - P(x > 6) = 1 - .008519 = .991481$$

Mean and Standard Deviation of the Binomial

As with the probability, the mean, variance, and standard deviation for the binomial have short-cut formulas. These formulas are very straight forward and much easier than using the general formulas for a discrete probability distribution. Remember these formulas only work for the binomial. The formulas are as follows.

$$\text{Mean (Expected Value)} = \mu = np$$
$$\text{Variance} = \sigma^2 = npq$$
$$\text{Standard Deviation} = \sigma = \sqrt{\sigma^2} = \sqrt{npq}$$

Example

You take a 100 question multiple-choice exam and use random guessing as your strategy. Each question has 4 choices. The random variable *x* is the number of correct answers on the exam.

a. Find the mean of *x*.
b. Find the variance of *x*.
c. Find the standard deviation of *x*.

Answers

$n = 100$

$p = \frac{1}{4}$

$q = \frac{3}{4}$

a. $\mu = n \cdot p = 100(1/4) = 25$
   On average you would get 25 questions right (25%). This should make since given the scenario.
b. $\sigma^2 = n \cdot p \cdot q = 100(1/4)(3/4) = 18.75$
c. $\sigma = \sqrt{\sigma^2} = \sqrt{n \cdot p \cdot q} = \sqrt{18.75} = 4.330$