

## Discrete Random Variables

**CO-6:** Apply basic concepts of probability, random variation, and commonly used statistical probability distributions.

### Review:

- [Basic Probability Rules](#)
- [Conditional Probability and Independence](#)

### Video

**Video:** [Discrete Random Variables](#) (22:40 Total)

We begin with discrete random variables: variables whose possible values are a list of distinct values. In order to decide on some notation, let's look at the coin toss example again:

A fair coin is tossed twice.

- Let the random variable  $X$  be the number of tails we get in this random experiment.
- In this case, the possible values that  $X$  can assume are
  - 0 (if we get HH),
  - 1 (if get HT or TH),
  - and 2 (if we get TT).

### Notation

If we want to find the probability of the event “getting 1 tail,” we'll write:  $P(X = 1)$

If we want to find the probability of the event “getting 0 tails,” we'll write:  $P(X = 0)$

In general, we'll write:  $P(X = x)$  or  $P(X = k)$  to denote the probability that the **discrete** random variable  $X$  gets the value  $x$  or  $k$  respectively.

Many students prefer the second notation as keeping track of the difference between  $X$  and  $x$  can cause confusion.

- Here the  $X$  represents the random variable and  $x$  or  $k$  denote the value of interest in the current problem (0, 1, etc. ).
- Note that for the random variables we'll use a capital letter, and for the value we'll use a lowercase letter.

### Section Plan

The way this section on discrete random variables is organized is very similar to the way we organized our discussion about one quantitative variable in the Exploratory Data Analysis unit.

It will be separated into four sections.

1. We'll first discuss the probability **distribution** of a discrete random variable, ways to display it, and how to use it in order to find probabilities of interest.
2. We'll then move on to talk about the **mean and standard deviation** of a discrete random variable, which are measures of the center and spread of its distribution.
3. We'll conclude this part by discussing a special and very common class of discrete random variable: the **binomial** random variable.

### Probability Distributions

## Learning Objectives

**LO 6.12:** Use the probability distribution for a discrete random variable to find the probability of events of interest.

When we learned how to find probabilities by applying the basic principles, we generally focused on just one particular outcome or event, like the probability of getting exactly one tail when a coin is tossed twice, or the probability of getting a 5 when a die is rolled.

Now that we have mastered the solution of individual probability problems, we'll proceed to look at the big picture by considering all the possible values of a discrete random variable, along with their associated probabilities.

This list of possible values and probabilities is called the **probability distribution** of the random variable.

### Comments:

- In the Exploratory Data Analysis unit of this course, we often looked at the distribution of sample values in a quantitative data set. We would display the values with a histogram, and summarize them by reporting their mean.
- In this section, when we look at the probability distribution of a random variable, we consider all its possible values and their overall probabilities of occurrence.
- Thus, we have in mind an entire population of values for a variable. When we display them with a histogram or summarize them with a mean, these are representing a population of values, not a sample.
- The distinction between sample and population is an essential concept in statistics, because an ultimate goal is to draw conclusions about unknown values for a population, based on what is observed in the sample.

In the examples which follow we will sometimes illustrate how the probability distribution is created.

We do this to demonstrate the usefulness of the probability rules we previously discussed and to illustrate clearly how probability distributions can be created.

As we are more focused on data driven methods, you will often be given a probability distribution based upon data as opposed to constructing the theoretical probability distribution based upon flipping coins or similar classical probability experiments.

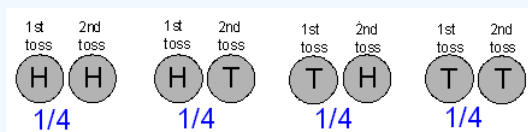
Recall our first example, when we introduced the idea of a random variable. In this example we tossed a coin twice.

### ✓ EXAMPLE: Flipping a Coin Twice

**What is the probability distribution of  $X$ , where the random variable  $X$  is the number of tails appearing in two tosses of a fair coin?**

We first note that since the coin is fair, each of the four outcomes HH, HT, TH, TT in the sample space  $S$  is equally likely, and so each has a probability of  $1/4$ .

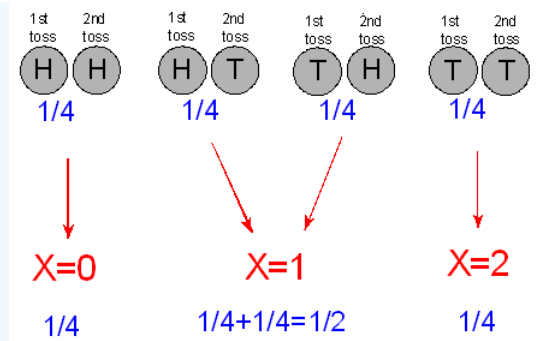
(Alternatively, the multiplication principle can be applied to find the probability of each outcome to be  $1/2 * 1/2 = 1/4$ .)



$X$  takes the value 0 only for the outcome **HH**, so the probability that  $X = 0$  is  $1/4$ .

$X$  takes the value 1 for outcomes **HT** or **TH**. By the addition principle, the probability that  $X = 1$  is  $1/4 + 1/4 = 1/2$ .

Finally,  $X$  takes the value 2 only for the outcome **TT**, so the probability that  $X = 2$  is  $1/4$ .



The **probability distribution of the random variable X** is easily summarized in a table:

List of possible values	x	0	1	2
Probability of each value	P(X=x)	1/4	1/2	1/4

As mentioned before, we write “ $P(X = x)$ ” to denote “the probability that the random variable X takes the value x.”

The way to interpret this table is:

- X takes the values 0, 1, 2 and  $P(X = 0) = 1/4$ ,  $P(X = 1) = 1/2$ ,  $P(X = 2) = 1/4$ .

Note that events of the type  $(X = x)$  are subject to the principles of probability established earlier, and will provide us with a way of systematically exploring the behavior of random variables.

In particular, the first two principles in the context of probability distributions of random variables will now be stated.

The ideal gas law is easy to remember and apply in solving problems, as long as you get the **proper values** aAny **probability distribution** of a **discrete random variable** must satisfy:

1.  $0 \leq P(X = x) \leq 1$
2.  $\sum_x P(X = x) = 1$

The probability distribution for two flips of a coin was simple enough to construct at once.

For more complicated random experiments, it is common to first construct a table of all the outcomes and their probabilities, then use the addition principle to condense that information into the actual probability distribution table.

### ✓ EXAMPLE: Flipping a Coin Three Times

A coin is tossed three times. Let the random variable X be the number of tails.

**Find the probability distribution of X.**

We’ll follow the same reasoning we used in the previous example:

First, we specify the 8 possible outcomes in S, along with the number and the probability of that outcome.

- Because they are all equally likely, each has probability 1/8.
- Alternatively, by the multiplication principle, each particular sequence of three coin faces has probability  $1/2 * 1/2 * 1/2 = 1/8$ .

Then we figure out what the value of X is (number of tails) for each possible outcome.

Outcome	Probability	X
HHH	$1/2 \cdot 1/2 \cdot 1/2 = 1/8$	0
HHT	$1/8$	1
HTH	$1/8$	1
THH	$1/8$	1
HTT	$1/8$	2
THT	$1/8$	2
TTH	$1/8$	2
TTT	$1/8$	3

Next, we use the addition principle to assert that

- $P(X = 1) = P(\text{HHT or HTH or THH}) = P(\text{HHT}) + P(\text{HTH}) + P(\text{THH}) = 1/8 + 1/8 + 1/8 = 3/8$ .
- Similarly,  $P(X = 2) = P(\text{HTT or THT or TTH}) = 3/8$ .

Outcome	Probability	X	
HHH	$1/8$	0	→ $1/8$
HHT	$1/8$	1	→
HTH	$1/8$	1	→ $1/8 + 1/8 + 1/8 = 3/8$
THH	$1/8$	1	→
HTT	$1/8$	2	→
THT	$1/8$	2	→ $1/8 + 1/8 + 1/8 = 3/8$
TTH	$1/8$	2	→
TTT	$1/8$	3	→ $1/8$

The resulting probability distribution is:

X	0	1	2	3
$P(X=x)$	$1/8$	$3/8$	$3/8$	$1/8$

In the previous two examples, we needed to specify the probability distributions ourselves, based on the physical circumstances of the situation.

In some situations, the probability distribution may be specified with a formula.

Such a formula must be consistent with the constraints imposed by the laws of probability, so that the probability of each outcome must be between 0 and 1, and the probabilities of all possible outcomes together must sum to 1.

We will see this with the binomial distribution.

## Probability Histograms

We learned to display the distribution of sample values for a quantitative variable with a histogram in which the horizontal axis represented the range of values in the sample.

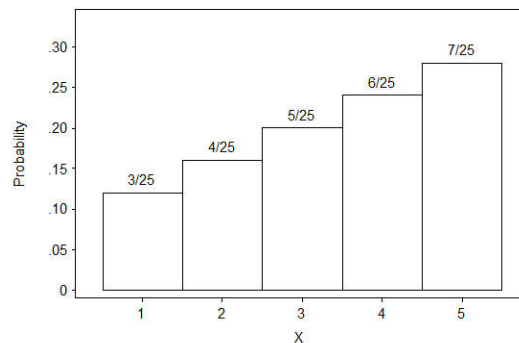
- The vertical axis represented the frequency or relative frequency (sometimes given as a percentage) of sample values occurring in that interval.
- The width of each rectangle in the histogram was an interval, or part of the possible values for the quantitative variable.
- The height of each rectangle was the frequency (or relative frequency) for that interval.

Similarly, we can display the probability distribution of a random variable with a probability histogram.

- The horizontal axis represents the range of all possible values of the random variable
- The vertical axis represents the probabilities of those values.

Here an example of a probability histogram.

(Such probabilities are not always increasing; they just happen to be so in this example).



## Area of a Probability Histogram

Notice that each rectangle in the histogram has a width of 1 unit. The height of each rectangle is the probability that it will occur. Thus, the area of each rectangle is base times height, which for these rectangles is 1 times its probability for each value of  $X$ .

This means that for **probability distributions of discrete random variables**, the sum of the areas of all of the rectangles is the same as the sum of all of the probabilities. **The total area = 1.**

For probability distributions of discrete random variables, this is equivalent to the property that the sum of all of the probabilities must equal 1.

**Learn by Doing:** [Probability Distributions](#)

## Finding Probabilities

We've seen how probability distributions are created. Now it's time to use them to find probabilities.

### ✓ EXAMPLE: Changing Majors

A random sample of graduating seniors was surveyed just before graduation. One question that was asked is: How many times did you change majors?  
The results are displayed in a probability distribution.

x	0	1	2	3	4	5
P(X = x)	.28	.37	.23	.09	.02	.01

Using this probability distribution, we can answer probability questions such as:

**What is the probability that a randomly selected senior has changed majors more than once?**

This can be written as  $P(X > 1)$ .

We can find this probability by adding the appropriate individual probabilities in the probability distribution.

- $P(X > 1)$
- $= P(X = 2) + P(X = 3) + P(X = 4) + P(X = 5)$
- $= 0.23 + 0.09 + 0.02 + 0.01$
- $= 0.35$

As you just saw in this example, we need to pay attention to the wording of the probability question.

The key words that told us which values to use for  $X$  are **more than**.

The following will clarify and reinforce the **key words** and their meanings.

## Key Words

Let's begin with some everyday situations using **at least** and **at most**.

Suppose someone said to you, "I need you to write **at least 10 pages** for a term paper."

- What does this mean?
- It means that 10 pages is the smallest amount you are going to write.
- In other words, you will write **10 or more** pages for the term paper.
- This would be the same as saying, "**not less than** 10 pages."
- So, for example, writing 9 pages would be unacceptable.

On the other hand, suppose you are considering the number of children you will have. You want **at most 3 children**.

- This means that 3 children is the most that you wish to have.
- In other words, you will have **3 or fewer**
- This would be the same as saying, "**not more than** 3 children."
- So, for example, you would not want to have 4 children.

The following table gives a list of some key words to know.

Suppose a random variable  $X$  had possible values of 0 through 5.

Key Words	Meaning	Symbols	Values for $X$
more than 2	strictly larger than 2	$X > 2$	3, 4, 5
no more than 2	2 or fewer	$X \leq 2$	0, 1, 2
fewer than 2	strictly smaller than 2	$X < 2$	0, 1
no less than 2	2 or more	$X \geq 2$	2, 3, 4, 5
at least 2	2 or more	$X \geq 2$	2, 3, 4, 5
at most 2	2 or fewer	$X \leq 2$	0, 1, 2
exactly 2	2, no more or no less, only 2	$X = 2$	2

Before we move on to the next section on the means and variances of a probability distribution, let's revisit the changing majors example:

### ✓ EXAMPLE: Changing Major

$x$	0	1	2	3	4	5
$P(X = x)$	.28	.37	.23	.09	.02	.01

**Question:** Based upon this distribution, do you think it would be unusual to change majors 2 or more times?

**Answer:**

- $P(X \geq 2) = 0.35$ .
- So, 35% of the time a student changes majors 2 or more times.
- This means that it is not unusual to do so.

**Question:** Do you think it would be unusual to change majors 4 or more times?

**Answer:**

- $P(X \geq 4) = 0.03$ .
- So, 3% of the time a student changes majors 4 or more times.
- This means that it is fairly unusual to do so.

We can even answer more difficult questions using our probability rules!

**Question:** What is the probability of changing majors only once given at least one change in major.

**Answer:**

- $P(X = 1 | X \geq 1) = P(X = 1 \text{ AND } X \geq 1) / P(X \geq 1)$  [using Probability Rule 7]
- $= P(X = 1) / P(X \geq 1)$  [since the only outcome that satisfies both  $X = 1$  and  $X \geq 1$  is  $X = 1$ ]
- $= (0.37) / (0.37 + 0.23 + 0.09 + 0.02 + 0.01) = 0.37 / 0.72 = 0.5139$ .
- So, among students who change majors, 51% of these students will only change majors one time.

After we learn about means and standard deviations, we will have another way to answer these types of questions.

## Mean of a Discrete Random Variable

### Learning Objectives

**LO 6.13:** Find the mean, variance, and standard deviation of a discrete random variable.

In the Exploratory Data Analysis (EDA) section, we displayed the distribution of one quantitative variable with a histogram, and supplemented it with numerical measures of center and spread.

We are doing the same thing here.

- We display the probability distribution of a discrete random variable with a table, formula or histogram.
- And supplement it with numerical measures of the center and spread of the probability distribution.

These measures are the **mean** and **standard deviation** of the **random variable**.

This section will be devoted to introducing these measures. As before, we'll start with the numerical measure of center, the mean. Let's begin by revisiting an example we saw in EDA.

### ✓ EXAMPLE: World Cup Soccer

Recall that we used the following data from 3 World Cup tournaments (a total of 192 games) to introduce the idea of a **weighted average**.

We've added a third column to our table that gives us relative frequencies.

total # goals/game	frequency	relative frequency
0	17	$17 / 192 = 0.089$
1	45	$45 / 192 = 0.234$
2	51	$51 / 192 = 0.266$
3	37	$37 / 192 = 0.193$
4	25	$25 / 192 = 0.130$
5	11	$11 / 192 = 0.057$
6	3	$3 / 192 = 0.016$
7	2	$2 / 192 = 0.010$
8	1	$1 / 192 = 0.005$

The mean for this data is:

$$\frac{0(17) + 1(45) + 2(51) + 3(37) + 4(25) + 5(11) + 6(3) + 7(2) + 8(1)}{192}$$

Distributing the division by 192 we get:

$$0 \left( \frac{17}{192} \right) + 1 \left( \frac{45}{192} \right) + 2 \left( \frac{51}{192} \right) + \cdots + 8 \left( \frac{1}{192} \right)$$

Notice that the mean is each number of goals per game multiplied by its relative frequency.

Since we usually write the relative frequencies as decimals, we can see that:

**Mean** number of goals per game =

$$\bullet \quad 0(0.089) + 1(0.234) + 2(0.266) + 3(0.193) + 4(0.130) + 5(0.057) + 6(0.016) + 7(0.010) + 8(0.005)$$

$$= 2.36, \text{ rounded to two decimal places.}$$

In Exploratory Data Analysis, we used the **mean** of a sample of quantitative values—their arithmetic average—to tell the **center** of their distribution. We also saw how a weighted mean was used when we had a frequency table. These frequencies can be changed to relative frequencies.

So we are essentially using the relative frequency approach to find probabilities. We can use this to find the **mean**, or **center**, of a **probability distribution for a discrete random variable**, which will be a weighted average of its values; the more probable a value is the more weight it gets.

As always, it is important to distinguish between a concrete sample of observed values for a variable versus an abstract population of all values taken by a random variable in the long run.

Whereas we denoted the mean of a sample as  $\bar{x}$ , we now denote the mean of a random variable using the **Greek letter mu** with a subscript for the random variable we are using.

Let's see how this is done by looking at a specific example.

#### ✓ EXAMPLE: Xavier's Production Line

Xavier's production line produces a variable number of defective parts in an hour, with probabilities shown in this table:

X	0	1	2	3	4
P(X=x)	.15	.30	.25	.20	.10

How many defective parts are typically produced in an hour on Xavier's production line? If we sum up the possible values of X, each weighted with its probability, we have

$$\mu_X = 0(0.15) + 1(0.30) + 2(0.25) + 3(0.20) + 4(0.10) = 1.8$$

Here is the general definition of the mean of a discrete random variable:

In general, for any discrete random variable X with probability distribution

X	$x_1$	$x_2$	$x_3$	$\dots$	$x_n$
P(X=x)	$p_1$	$p_2$	$p_3$	$\dots$	$p_n$

The **mean** of X is defined to be

$$\mu_X = x_1p_1 + x_2p_2 + \dots + x_np_n = \sum_{i=1}^n x_ip_i$$

- In general, the mean of a random variable tells us its "long-run" average value.
- It is sometimes referred to as the **expected value** of the random variable.

Although "**expected value**" is a common, and even preferred term in the field of statistics, this expression may be somewhat misleading, because in many cases it is impossible for a random variable to actually equal its expected value.

For example, the mean number of goals for a World Cup soccer game is 2.36. But we can never expect any single game to result in 2.36 goals, since it is not possible to score a fraction of a goal. Rather, 2.36 is the long-run average of all World Cup soccer games.

In the case of Xavier's production line, the mean number of defective parts produced in an hour is 1.8. But the actual number of defective parts produced in any given hour can never equal 1.8, since it must take whole number values.

To get a better feel for the mean of a random variable, let's extend the defective parts example:

#### ✓ EXAMPLE: Xavier's and Yves' Production Lines

Recall the probability distribution of the random variable  $X$ , representing the number of defective parts in an hour produced by Xavier's production line.

$X$	0	1	2	3	4
$P(X=x)$	.15	.30	.25	.20	.10

The number of defective parts produced each hour by Yves' production line is a random variable  $Y$  with the following probability distribution:

$Y$	0	1	2	3	4
$P(Y=y)$	.05	.05	.10	.75	.05

Look at both probability distributions. Both  $X$  and  $Y$  take the same possible values (0, 1, 2, 3, 4).

However, they are very different in the way the probability is distributed among these values.

**Learn by Doing:** [Comparing Probability Distributions #1](#)

**Did I Get This?:** [Mean of Discrete Random Variable](#)

## Variance and Standard Deviation of a Discrete Random Variable

### Learning Objectives

**LO 6.13:** Find the mean, variance, and standard deviation of a discrete random variable.

In Exploratory Data Analysis, we used the mean of a sample of quantitative values (their arithmetic average,  $\bar{x}$ ) to tell the center of their distribution, and the standard deviation ( $s$ ) to tell the typical distance of sample values from their mean.

We described the center of a probability distribution for a random variable by reporting its mean which we denoted by the Greek letter  $\mu$ .

Now we would like to establish an accompanying measure of **spread**.

Our measure of spread will still report the typical distance of values from their means, but in order to distinguish the spread of a population of all of a random variable's values from the spread ( $s$ ) of sample values, we will denote the standard deviation of the random variable  $X$  with the Greek lower case "**sigma**," and use a subscript to remind us what is the variable of interest (there may be more than one in later problems):

We will also focus more frequently than before on the squared standard deviation, called the **variance**, because some important rules we need to invoke are in terms of variance rather than standard deviation.

#### ✓ EXAMPLE: Xavier's Production Line

Recall that the number of defective parts produced each hour by Xavier's production line is a random variable  $X$  with the following probability distribution:

X	0	1	2	3	4
P(X=x)	.15	.30	.25	.20	.10

We found the mean number of defective parts produced per hour to be 1.8.

Obviously, there is variation about this mean: some hours as few as 0 defective parts are produced, whereas in other hours as many as 4 are produced.

### Typically, how far does the number of defective parts fall from the mean of 1.8?

As we did for the spread of sample values, we measure the spread of a random variable by calculating the square root of the average squared deviation from the mean.

Now “average” is a weighted average, where more probable values of the random variable are accordingly given more weight.

Let’s begin with the variance, or average squared deviation from the mean, and then take its square root to find the standard deviation:

Values of X	0	1	2	3	4
Dev. from mean	(0-1.8)	(1-1.8)	(2-1.8)	(3-1.8)	(4-1.8)
Sq. deviations	(0-1.8) <sup>2</sup>	(1-1.8) <sup>2</sup>	(2-1.8) <sup>2</sup>	(3-1.8) <sup>2</sup>	(4-1.8) <sup>2</sup>
Probabilities	.15	.30	.25	.20	.10

$$\begin{aligned}\text{Variance} = \sigma_X^2 &= (0 - 1.8)^2(0.15) + (1 - 1.8)^2(0.30) + (2 - 1.8)^2(0.25) \\ &\quad + (3 - 1.8)^2(0.20) + (4 - 1.8)^2(0.1) \\ &= 1.46\end{aligned}$$

$$\text{standard deviation} = \sigma_X = \sqrt{1.46} = 1.21$$

How do we interpret the standard deviation of X?

- Xavier’s production line produces an average of 1.80 defective parts per hour.
- **The number of defective parts varies from hour to hour; typically (or, on average), it is about 1.21 away from the mean 1.80.**

Here is the formal definition:

In general, for any discrete random variable X with probability distribution

X	x <sub>1</sub>	x <sub>2</sub>	x <sub>3</sub>	...	x <sub>n</sub>
P(X=x)	p <sub>1</sub>	p <sub>2</sub>	p <sub>3</sub>	...	p <sub>n</sub>

The **variance** of X is defined to be

$$\begin{aligned}\sigma_X^2 &= (x_1 - \mu_X)^2 p_1 + (x_2 - \mu_X)^2 p_2 + \dots + (x_n - \mu_X)^2 p_n \\ &= \sum_{i=1}^n (x_i - \mu_X)^2 p_i\end{aligned}$$

There is also a “short-cut” formula which is faster for by-hand calculation. In the formula below we have dropped the subscript for the variable in the notation. In this short-cut, we simply need to

- square each X,
- multiply by the probability of that X,
- then sum those values.
- From that result we subtract the square of the mean to find the variance.

$$\text{Var}(X) = \sigma^2 = \sum_{i=1}^n [x_i^2 P(X = x_i)] - \mu^2$$

The **standard deviation** is the square root of the variance

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

### Did I Get This?: [Standard Deviation of a Discrete Random Variable](#)

The purpose of the next activity is to give you better intuition about the mean and standard deviation of a random variable.

### Learn by Doing: [Comparing Probability Distributions #2](#)

#### ✓ EXAMPLE: Xavier's and Yves' Production Lines

Recall the probability distribution of the random variable  $X$ , representing the number of defective parts per hour produced by Xavier's production line, and the probability distribution of the random variable  $Y$ , representing the number of defective parts per hour produced by Yves' production line:

$X$	0	1	2	3	4
$P(X=x)$	.15	.30	.25	.20	.10

$Y$	0	1	2	3	4
$P(Y=y)$	.05	.05	.10	.75	.05

Look carefully at both probability distributions. Both  $X$  and  $Y$  take the same possible values (0, 1, 2, 3, 4).

However, they are very different in the way the probability is distributed among these values. We saw before that this makes a difference in means:

$$\mu_X = 1.8$$

$$\mu_Y = 2.7$$

We now want to get a sense about how the different probability distributions impact their standard deviations.

Recall that the **standard deviation** of a **random variable** can be **interpreted** as a **typical** (or the **long-run average**) **distance** between the **value of  $X$  and its mean**.

### Learn by Doing: [Comparing Probability Distributions #3](#)

So, 75% of the time  $Y$  will assume a value (3) that is very close to its mean (2.7), while  $X$  will assume a value (2) that is close to its mean (1.8) much less often—only 25% of the time.

The long-run average, then, of the distance between the values of  $Y$  and their mean will be much smaller than the long-run average of the distance between the values of  $X$  and their mean.

Therefore

$$\sigma_Y < \sigma_X = 1.21$$

Actually we have

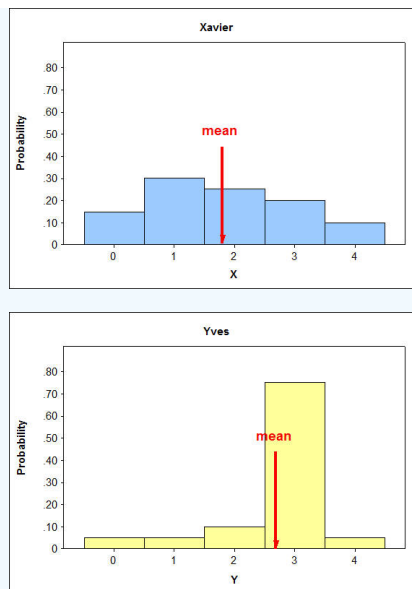
$$\sigma_Y = 0.85$$

So we can draw the following conclusion:

Yves' production line produces an average of 2.70 defective parts per hour.

**The number of defective parts varies from hour to hour; typically (or, on average), it is about 0.85 away from 2.70.**

Here are the histograms for the production lines:



When we compare distributions, the distribution in which it is **more likely** to find values that are further from the mean will have a **larger** standard deviation.

Likewise, the distribution in which it is **less likely** to find values that are further from the mean will have the **smaller** standard deviation.

### Did I Get This?: Standard Deviation of a Discrete Random Variable #2

#### Comment:

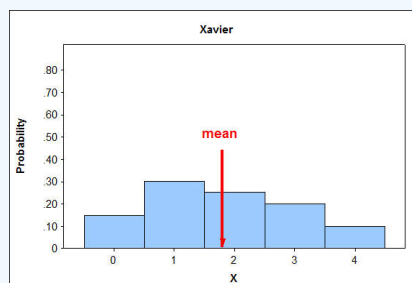
As we have stated before, using the mean and standard deviation gives us another way to assess which values of a random variable are unusual.

For reasonably symmetric distributions, any values of a random variable that fall within 2 or 3 standard deviations of the mean would be considered ordinary (not unusual).

For any distribution, it is unusual for values to fall outside of 3 or 4 standard deviations – depending on your definition of “unusual.”

#### ✓ EXAMPLE: Xavier's Production Line-Unusual or Not?

Looking once again at the probability distribution for Xavier's production line:



#### Would it be considered unusual to have 4 defective parts per hour?

We know that the mean is 1.8 and the standard deviation is 1.21.

Ordinary values are within 2 (or 3) standard deviations of the mean.

- $1.8 - 2(1.21) = -0.62$  and

- $1.8 + 2(1.21) = 4.22$ .

This gives us an interval from -0.62 to 4.22.

Since we cannot have a negative number of defective parts, the interval is essentially from 0 to 4.22.

Because 4 is within this interval, it would be considered ordinary. Therefore, it is **not unusual**.

**Would it be considered unusual to have no defective parts?**

Zero is within 2 standard deviations of the mean, so it would not be considered unusual to have no defective parts.

The following activity will reinforce this idea.

**Learn by Doing:** [Unusual or Not?](#)

---

Discrete Random Variables is shared under a [CC BY-NC-SA 4.0](#) license and was authored, remixed, and/or curated by LibreTexts.