

## 4.5: Poisson Distribution

Another useful probability distribution is the Poisson distribution, or waiting time distribution. This distribution is used to determine how many checkout clerks are needed to keep the waiting time in line to specified levels, how many telephone lines are needed to keep the system from overloading, and many other practical applications. A modification of the Poisson, the Pascal, invented nearly two centuries ago, is used today by telecommunications companies worldwide for load factors, satellite hookup levels and Internet capacity problems. The distribution gets its name from Simeon Poisson who presented it in 1837 as an extension of the binomial distribution which we will see can be estimated with the Poisson.

There are two main characteristics of a Poisson experiment.

1. The **Poisson probability distribution** gives the probability of a number of events occurring in a **fixed interval** of time or space if these events happen with a known average rate.
2. The events are independently of the time since the last event. For example, a book editor might be interested in the number of words spelled incorrectly in a particular book. It might be that, on the average, there are five words spelled incorrectly in 100 pages. The interval is the 100 pages and it is assumed that there is no relationship between when misspellings occur.
3. The random variable  $X$  = the number of occurrences in the interval of interest.

### ? Exercise 4.5.1

A bank expects to receive six bad checks per day, on average. What is the probability of the bank getting fewer than five bad checks on any given day? Of interest is the number of checks the bank receives in one day, so the time interval of interest is one day. Let  $X$  = the number of bad checks the bank receives in one day. If the bank expects to receive six bad checks per day then the average is six checks per day. Write a mathematical statement for the probability question.

**Answer**

$$P(x < 5)$$

### 🔗 Try It 4.5.1

An electronics store expects to have ten returns per day on average. The manager wants to know the probability of the store getting fewer than eight returns on any given day. State the probability question mathematically.

### ? Exercise 4.5.2

You notice that a news reporter says "uh," on average, two times per broadcast. What is the probability that the news reporter says "uh" more than two times per broadcast.

This is a Poisson problem because you are interested in knowing the number of times the news reporter says "uh" during a broadcast.

- a. What is the interval of interest?
- b. What is the average number of times the news reporter says "uh" during one broadcast?
- c. Let  $X =$  \_\_\_\_\_. What values does  $X$  take on?
- d. The probability question is  $P$  (\_\_\_\_\_).

**Answer**

- a. one broadcast measured in minutes
- b. 2
- c. Let  $X$  = the number of times the news reporter says "uh" during one broadcast.  $x = 0, 1, 2, 3, \dots$
- d.  $P(x > 2)$

### Try It 4.5.2

An emergency room at a particular hospital gets an average of five patients per hour. A doctor wants to know the probability that the ER gets more than five patients per hour. Give the reason why this would be a Poisson distribution.

## Notation for the Poisson: $P =$ Poisson Probability Distribution Function

$$X \sim P(\mu)$$

Read this as "  $X$  is a random variable with a Poisson distribution." The parameter is  $\mu$  ( or  $\lambda$ );  $\mu$  ( or  $\lambda$ ) = the mean for the interval of our interest. The mean is the number of occurrences that occur on average during the interval period.

The formula for computing probabilities that are from a Poisson process is:

$$P(x) = \frac{\mu^x e^{-\mu}}{x!} \quad (4.5.1)$$

where  $P(X)$  is the probability of  $X$  successes,  $\mu$  is the expected number of successes based upon historical data,  $e$  is the natural logarithm approximately equal to 2.718, and  $X$  is the number of successes per unit, usually per unit of time.

Remember your algebra class:  $x^{-n} = \frac{1}{x^n}$ . The Poisson distribution has both a mean,  $\mu$ , the average number of occurrences per unit of time, and also a standard deviation,  $\sigma = \sqrt{\mu}$ .

In order to use the Poisson distribution, certain assumptions must hold. These are: the probability of a success,  $\mu$ , is unchanged within the interval, there cannot be simultaneous successes within the interval, and finally, that the probability of a success among intervals is independent, the same assumption of the binomial distribution.

In a way, the Poisson distribution can be thought of as a clever way to convert a continuous random variable, usually time, into a discrete random variable by breaking up time into discrete independent intervals. This way of thinking about the Poisson helps us understand why it can be used to estimate the probability for the discrete random variable from the binomial distribution. The Poisson is asking for the probability of a number of successes during a period of time while the binomial is asking for the probability of a certain number of successes for a given number of trials.

### ? Exercise 4.5.3

Leah receives about six telephone calls between 8 a.m. and 10 a.m. What is the probability that Leah receives more than one call in the next 15 minutes?

Let  $X =$  the number of calls Leah receives in 15 minutes. (The interval of interest is 15 minutes or  $\frac{1}{4}$  hour.)  $x = 0, 1, 2, 3, \dots$

If Leah receives, on the average, six telephone calls in two hours, and there are eight 15 minute intervals in two hours, then Leah receives

$\left(\frac{1}{8}\right)(6) = 0.75$  calls in 15 minutes, on average. So,  $\mu = 0.75$  for this problem.

Find  $P(x > 1)$ .

The Poisson distribution is discrete, and thus  $> 1$  includes all whole numbers through infinity. The solution is to subtract the probability less than 1 thus  $1 - [P(x = 0) + P(x = 1)]$

$$\begin{aligned} P(x) &= \frac{\mu^x e^{-\mu}}{x!} \\ P(x > 1) &= 1 - P(x \leq 1) \\ &= 1 - [P(x = 0) + P(x = 1)] \\ &= 1 - \left[ \frac{0.75^0 e^{-0.75}}{0!} + \frac{0.75^1 e^{-0.75}}{1!} \right] \\ &= 1 - \left[ \frac{(1)(0.4724)}{1} + \frac{(0.75)(0.4724)}{1} \right] \\ &= 1 - [0.4724 + 0.3543] = 0.1733 \end{aligned} \quad (4.5.2)$$

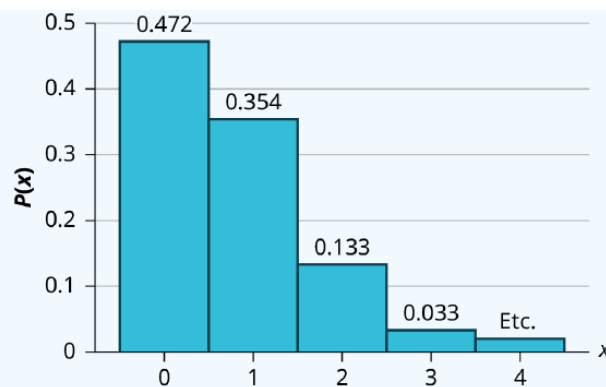


Figure 4.5.1:

The y-axis in Figure 4.5.1 contains the probability of  $x$  where  $X$  = the number of calls in 15 minutes

### Try It 4.5.3

A customer service center receives about ten emails every half-hour. What is the probability that the customer service center receives more than four emails in the next six minutes?

### ? Exercise 4.5.4

According to Baydin, an email management company, an email user gets, on average, 147 emails per day. Let  $X$  = the number of emails an email user receives per day. The discrete random variable  $X$  takes on the values  $x = 0, 1, 2, \dots$ . The random variable  $X$  has a Poisson distribution:  $X \sim P(147)$ . The mean is 147 emails.

- What is the probability that an email user receives exactly 160 emails per day?
- What is the probability that an email user receives at most 160 emails per day?
- What is the standard deviation?

### Answer

- $P(x = 160) = \text{poissonpdf}(147, 160) \approx 0.0180$
- $P(x \leq 160) = \text{poissoncdf}(147, 160) \approx 0.8666$
- Standard Deviation =  $\sigma = \sqrt{\mu} = \sqrt{147} \approx 12.1244$

### Try It 4.5.4

According to a recent poll by the Pew Internet Project, people between the ages of 14 and 17 send an average of 187 text messages each day. Let  $X$  = the number of texts that a girl aged 14 to 17 sends per day. The discrete random variable  $X$  takes on the values  $x = 0, 1, 2, \dots$ . The random variable  $X$  has a Poisson distribution:  $X \sim P(187)$ . The mean is 187 text messages.

- What is the probability that a person sends exactly 175 texts per day?
- What is the probability that a person sends at most 150 texts per day?
- What is the standard deviation?

### ? Exercise 4.5.5

Text message users receive or send an average of 41.5 text messages per day.

- How many text messages does a text message user receive or send per hour?
- What is the probability that a text message user receives or sends two messages per hour?
- What is the probability that a text message user receives or sends more than two messages per hour?

### Answer

- a. Let  $X$  = the number of texts that a user sends or receives in one hour. The average number of texts received per hour is  $\frac{41.5}{24} \approx 1.7292$
- b.  $P(x = 2) = \frac{\mu^x e^{-\mu}}{x!} = \frac{1.7292^2 e^{-1.7292}}{2!} = 0.265$
- c.  $P(x > 2) = 1 - P(x \leq 2) = 1 - \left[ \frac{7^0 e^{-7}}{0!} + \frac{7^1 e^{-7}}{1!} + \frac{7^2 e^{-7}}{2!} \right] = 0.250$

#### Try It 4.5.5

Atlanta's Hartsfield-Jackson International Airport is the busiest airport in the world. On average there are 2,700 arrivals and departures each day.

- How many airplanes arrive and depart the airport per hour?
- What is the probability that there are exactly 100 arrivals and departures in one hour?
- What is the probability that there are at most 100 arrivals and departures in one hour?



#### Solution

- a. Let  $X$  = the number of texts that a user sends or receives in one hour. The average number of texts received per hour is  $\frac{41.52441.524}{24} \approx 1.7292$ .
- b.  $X \sim P(1.7292)$ , so  $P(x = 2) = \text{poissonpdf}(1.7292, 2) \approx 0.2653$
- c.  $P(x > 2) = 1 - P(x \leq 2) = 1 - \text{poissoncdf}(1.7292, 2) \approx 1 - 0.7495 = 0.2505$

#### Exercise 4.5.6

On a specific day in May starting at 4:30 PM, the probability of low seismic activity for the next 48 hours in Alaska was reported as about 1.02%. Use this information for the next 200 days to find the probability that there will be low seismic activity in ten of the next 200 days. Use both the binomial and Poisson distributions to calculate the probabilities. Are they close?

#### Answer

Let  $X$  = the number of days with low seismic activity.

Using the binomial distribution:

$$P(x = 10) = \frac{200!}{10!(200 - 10)!} \times .0102^{10} \times .9898^{190} = 0.000039 \quad (4.5.3)$$

Using the Poisson distribution:

Calculate  $\mu = np = 200(0.0102) \approx 2.04$

$$P(x = 10) = \frac{\mu^x e^{-\mu}}{x!} = \frac{2.04^{10} e^{-2.04}}{10!} = 0.000045 \quad (4.5.4)$$

We expect the approximation to be good because  $n$  is large (greater than 20) and  $p$  is small (less than 0.05). The results are close-both probabilities reported are almost 0.



#### Solution

Let  $X$  = the number of days with low seismic activity.

Using the binomial distribution:

- $P(x = 10) = \text{binompdf}(200, .0102, 10) \approx 0.000039$

Using the Poisson distribution:

- Calculate  $\mu = np = 200(0.0102) \approx 2.04$
- $P(x = 10) = \text{poissonpdf}(2.04, 10) \approx 0.000045$

We expect the approximation to be good because  $n$  is large (greater than 20) and  $p$  is small (less than 0.05). The results are close—both probabilities reported are almost 0.

#### Try It 4.5.6

On a specific day in May starting at 4:30 PM, the probability of moderate seismic activity for the next 48 hours in the Kuril Islands off the coast of Japan was reported at about 1.43%. Use this information for the next 100 days to find the probability that there will be low seismic activity in five of the next 100 days. Use both the binomial and Poisson distributions to calculate the probabilities. Are they close?

#### Exercise 4.5.7

The Poisson distribution is often referred to as a “waiting time” distribution. In a sense this is tied to the link between the discrete distribution measuring the number of occurrences and the treatment of the continuous random variable time. The Poisson takes the continuous random variable time and breaks it into a discrete random variable by measuring the discrete number of occurrences. In Example 4.16 we were measuring the number of messages sent, an occurrence, a discrete random variable, per day, a continuous random variable.

The Poisson distribution can be used for many other applications. Any continuous random variable that can be broken into discrete measures can use the Poisson distribution to calculate probabilities of the number of a given discrete value of  $x$ , the number of occurrences in which we have an interest. We are interested in the number of potholes in a highway to evaluate the quality of the paving job. In this case the pothole is the occurrence and can be counted per square miles of highway. Mile is a continuous random variable, and potholes is the discrete random variable. We may want to know the probability that a particular paving technique results in more than 20 potholes in 10 miles of paving. If that is the case, we may discard that paving technique. Alternatively, our chocolate cookies claim that each cookie has more than 5 chocolate pieces in each cookie. Sampling the cookie dough, we can calculate the probability that the resulting cookies will meet the claim of 5 chocolate pieces per cookie. Perhaps the production of glass for automobile windshields suffers from a flaw known as the “dot.” In the production process, air infrequently becomes lodged in the hot glass and upon cooling leaves a dot that distracts the driver. This would also be a case where use of the Poisson probability distribution would be the appropriate tool. To calculate the probability that the results have fewer than 3 “dots,” the number acceptable to the manufacturing process, the Poisson distribution would be useful.

By way of warning, the Poisson is a discrete random variable. We are counting occurrences. It occurred or it did not during the time, the miles of road, or the batch of cookie dough. In this sense the Poisson is like the binomial in that the occurrence is binary, happened or did not. Because of this link to the binomial, we can use the Poisson to estimate a binomial distribution, and this is discussed below.

#### Try It 4.5.7

In a small city, a survey 500 vehicles reveals that there is a 2% chance of an accident occurring in a week. Calculate the probability that at most 2 accidents occur in any given week.

## Estimating the Binomial Distribution with the Poisson Distribution

We found before that the binomial distribution provided an approximation for the hypergeometric distribution. Now we find that the Poisson distribution can provide an approximation for the binomial. We say that the binomial distribution approaches the Poisson. The binomial distribution approaches the Poisson distribution is as  $n$  gets larger and  $p$  is small such that  $np$  becomes a constant value. There are several rules of thumb for when one can say they will use a Poisson to estimate a binomial. One suggests that  $np$ , the mean of the binomial, should be less than 25. Another author suggests that it should be less than 7. And another, noting that the mean and variance of the Poisson are both the same, suggests that  $np$  and  $npq$ , the mean and variance of the binomial, should be greater than 5. There is no one broadly accepted rule of thumb for when one can use the Poisson to estimate the binomial.

As we move through these probability distributions we are getting to more sophisticated distributions that, in a sense, contain the less sophisticated distributions within them. This proposition has been proven by mathematicians. This gets us to the highest level of sophistication in the next probability distribution which can be used as an approximation to all of those that we have discussed so far. This is the normal distribution.

### ? Exercise 4.5.8

A survey of 500 seniors in the Price Business School yields the following information. 75% go straight to work after graduation. 15% go on to work on their MBA. 9% stay to get a minor in another program. 1% go on to get a Master's in Finance.

What is the probability that more than 2 seniors go to graduate school for their Master's in finance?

#### Answer

This is clearly a binomial probability distribution problem. The choices are binary when we define the results as "Graduate School in Finance" versus "all other options." The random variable is discrete, and the events are, we could assume, independent. Solving as a binomial problem, we have:

Binomial Solution

$$\begin{aligned}
 n \cdot p &= 500 \cdot 0.01 = 5 = \mu \\
 P(0) &= \frac{500!}{0!(500-0)!} 0.01^0 (1-0.01)^{500-0} = 0.00657 \\
 P(1) &= \frac{500!}{1!(500-1)!} 0.01^1 (1-0.01)^{500-1} = 0.03318 \\
 P(2) &= \frac{500!}{2!(500-2)!} 0.01^2 (1-0.01)^{500-2} = 0.08363
 \end{aligned} \tag{4.5.5}$$

Adding all 3 together = 0.12339

$$1 - 0.12339 = 0.87661 \tag{4.5.6}$$

Poisson approximation

$$\begin{aligned}
 n \cdot p &= 500 \cdot 0.01 = 5 = \mu \\
 n \cdot p \cdot (1-p) &= 500 \cdot 0.01 \cdot (0.99) \approx 5 = \sigma^2 = \mu \\
 P(X) &= \frac{e^{-np} (np)^x}{x!} = \left\{ P(0) = \frac{e^{-5} \cdot 5^0}{0!} \right\} + \left\{ P(1) = \frac{e^{-5} \cdot 5^1}{1!} \right\} + \left\{ P(2) = \frac{e^{-5} \cdot 5^2}{2!} \right\} \\
 &0.0067 + 0.0337 + 0.0842 = 0.1247 \\
 &1 - 0.1247 = 0.8753
 \end{aligned} \tag{4.5.7}$$

An approximation that is off by 1 one thousandth is certainly an acceptable approximation.

### 🔗 Try It 4.5.8

In Example 4.5.8, what is the probability that less than 4 seniors will go to graduate school for their master's in finance?

This page titled [4.5: Poisson Distribution](#) is shared under a [CC BY 4.0](#) license and was authored, remixed, and/or curated by [OpenStax](#) via [source content](#) that was edited to the style and standards of the LibreTexts platform.