4.7: Poisson Distribution

The Poisson distribution is popular for modelling the number of times an event occurs in an interval of time or space. It is a discrete probability distribution that expresses the probability of a given number of events occurring in a fixed interval of time and/or space if these events occur with a known average rate and independently of the time since the last event.

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 and 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.
2. The Poisson distribution may be used to approximate the binomial if the probability of success is "small" (such as 0.01) and the number of trials is "large" (such as 1,000). You will verify the relationship in the homework exercises. n is the number of trials, and p is the probability of a "success."

The random variable $X =$ the number of occurrences in the interval of interest.

Example 4.7.1

The average number of loaves of bread put on a shelf in a bakery in a half-hour period is 12. Of interest is the number of loaves of bread put on the shelf in five minutes. The time interval of interest is five minutes. What is the probability that the number of loaves, selected randomly, put on the shelf in five minutes is three?

Solution

Let $X =$ the number of loaves of bread put on the shelf in five minutes. If the average number of loaves put on the shelf in 30 minutes (half-hour) is 12, then the average number of loaves put on the shelf in five minutes is $\left(\frac{5}{30}\right)(12) = 2$ loaves of bread.

The probability question asks you to find $P(x = 3)$.

Exercise 4.7.1

The average number of fish caught in an hour is eight. Of interest is the number of fish caught in 15 minutes. The time interval of interest is 15 minutes. What is the average number of fish caught in 15 minutes?

Answer

$\left(\frac{15}{60}\right)(8) = 2$ fish

Example 4.7.2

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

$P(x < 5)$

Exercise 4.7.2

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.

Answer

$P(x < 8)$

Example 4.7.3

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 $P($______$)$.

Solutions

a. one broadcast
b. 2
c. Let $X =$ the number of times the news reporter says "uh" during one broadcast.

$$x = 0, 1, 2, 3, \ldots \tag{4.7.1}$$

d. $P(x > 2)$

Exercise 4.7.3

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.

Answer

This problem wants to find the probability of events occurring in a fixed interval of time with a known average rate. The events are independent.

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

$$X \sim P(\mu) \tag{4.7.2}$$

Read this as "X is a random variable with a Poisson distribution." The parameter is μ (or λ); μ (or λ) $=$ the mean for the interval of interest.

Example 4.7.4

Leah's answering machine 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?

Solution

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

$$x = 0, 1, 2, 3, \ldots \tag{4.7.3}$$

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.

$X \sim P(0.75)$

Find $P(x > 1)$. $P(x > 1) = 0.1734$ (calculator or computer)

- Press 1 – and then press 2$^{\text{nd}}$ DISTR.
- Arrow down to poissoncdf. Press ENTER.
- Enter (.75,1).
- The result is $P(x > 1) = 0.1734$.

The TI calculators use λ (lambda) for the mean.

The probability that Leah receives more than one telephone call in the next 15 minutes is about 0.1734:

$P(x > 1) = 1 - \text{poissoncdf}(0.75, 1)$.

The graph of $X \sim P(0.75)$ is:

This graphs shows a poisson probability distribution. It has 5 bars that decrease in height from left to right. The x-axis shows values in increments of 1 starting with 0, representing the number of calls Leah receives within 15 minutes. The y-axis ranges from 0 to 0.5 in increments of 0.1.

Figure 4.7.1

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

Exercise 4.7.4

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? Use the TI-83+ or TI-84 calculator to find the answer.

Answer

$P(x > 4) = 0.0527$

Example 4.7.5

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.

a. What is the probability that an email user receives exactly 160 emails per day?
b. What is the probability that an email user receives at most 160 emails per day?
c. What is the standard deviation?

Solutions

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

Exercise 4.7.5

According to a recent poll by the Pew Internet Project, girls 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.

a. What is the probability that a teen girl sends exactly 175 texts per day?
b. What is the probability that a teen girl sends at most 150 texts per day?
c. What is the standard deviation?

Answer

a. $P(x = 175) = \text{poissonpdf}(187, 175) \approx 0.0203$
b. $P(x \leq 150) = \text{poissoncdf}(187, 150) \approx 0.0030$
c. Standard Deviation $= \sigma = \sqrt{\mu} = \sqrt{187} \approx 13.6748$

Example 4.7.6

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

a. How many text messages does a text message user receive or send per hour?

b. What is the probability that a text message user receives or sends two messages per hour?
c. What is the probability that a text message user receives or sends more than two messages per hour?

Solutions

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. $X \sim P(1.7292)$, 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 4.7.6

Atlanta's Hartsfield-Jackson International Airport is the busiest airport in the world. On average there are 2,500 arrivals and departures each day.

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

Answer

a. Let $X =$ the number of airplanes arriving and departing from Hartsfield-Jackson in one hour. The average number of arrivals and departures per hour is $\frac{2{,}500}{24} \approx 104.1667$.
b. $X \sim P(104.1667)$, so $P(x=100) = \text{poissonpdf}(104.1667, 100) \approx 0.0366$
c. $P(x \leq 100) = \text{poissoncdf}(104.1667, 100) \approx 0.3651$

The Poisson distribution can be used to approximate probabilities for a binomial distribution. This next example demonstrates the relationship between the Poisson and the binomial distributions. Let n represent the number of binomial trials and let p represent the probability of a success for each trial. If n is large enough and p is small enough then the Poisson approximates the binomial very well. In general, n is considered "large enough" if it is greater than or equal to 20. The probability p from the binomial distribution should be less than or equal to 0.05. When the Poisson is used to approximate the binomial, we use the binomial mean $\mu = np$. The variance of X is $\sigma^2 = \sqrt{\mu}$ and the standard deviation is $\sigma = \sqrt{\mu}$. The Poisson approximation to a binomial distribution was commonly used in the days before technology made both values very easy to calculate.

✓ Example 4.7.7

On May 13, 2013, 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:

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

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.

> **Exercise 4.7.7**
>
> On May 13, 2013, 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?
>
> **Answer**
>
> Let $X =$ the number of days with moderate seismic activity.
>
> Using the binomial distribution: $P(x=5) = \text{binompdf}(100, 0.0143, 5) \approx 0.0115$
>
> Using the Poisson distribution:
>
> - Calculate $\mu = np = 100(0.0143) = 1.43$
> - $P(x=5) = \text{poissonpdf}(1.43, 5) = 0.0119$
>
> 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—the difference between the values is 0.0004.

References

1. “ATL Fact Sheet,” Department of Aviation at the Hartsfield-Jackson Atlanta International Airport, 2013. Available online at www.atl.com/about-atl/atl-factsheet/ (accessed May 15, 2013).
2. Center for Disease Control and Prevention. “Teen Drivers: Fact Sheet,” Injury Prevention & Control: Motor Vehicle Safety, October 2, 2012. Available online at http://www.cdc.gov/Motorvehiclesafet...factsheet.html (accessed May 15, 2013).
3. “Children and Childrearing,” Ministry of Health, Labour, and Welfare. Available online at http://www.mhlw.go.jp/english/policy...ing/index.html (accessed May 15, 2013).
4. “Eating Disorder Statistics,” South Carolina Department of Mental Health, 2006. Available online at http://www.state.sc.us/dmh/anorexia/statistics.htm (accessed May 15, 2013).
5. “Giving Birth in Manila: The maternity ward at the Dr Jose Fabella Memorial Hospital in Manila, the busiest in the Philippines, where there is an average of 60 births a day,” theguardian, 2013. Available online at www.theguardian.com/world/gal...471900&index=2 (accessed May 15, 2013).
6. “How Americans Use Text Messaging,” Pew Internet, 2013. Available online at pewinternet.org/Reports/2011/...in-Report.aspx (accessed May 15, 2013).
7. Lenhart, Amanda. “Teens, Smartphones & Testing: Texting volum is up while the frequency of voice calling is down. About one in four teens say they own smartphones,” Pew Internet, 2012. Available online at www.pewinternet.org/~/media/F...nd_Texting.pdf (accessed May 15, 2013).
8. “One born every minute: the maternity unit where mothers are THREE to a bed,” MailOnline. Available online at http://www.dailymail.co.uk/news/arti...thers-bed.html (accessed May 15, 2013).
9. Vanderkam, Laura. “Stop Checking Your Email, Now.” CNNMoney, 2013. Available online at management.fortune.cnn.com/20...our-email-now/ (accessed May 15, 2013).
10. “World Earthquakes: Live Earthquake News and Highlights,” World Earthquakes, 2012. www.world-earthquakes.com/ind...thq_prediction (accessed May 15, 2013).

Review

A Poisson probability distribution of a discrete random variable gives the probability of a number of events occurring in a fixed interval of time or space, if these events happen at a known average rate and independently of the time since the last event. The Poisson distribution may be used to approximate the binomial, if the probability of success is "small" (less than or equal to 0.05) and the number of trials is "large" (greater than or equal to 20).

Formula Review

$X \sim P(\mu)$ means that X has a Poisson probability distribution where $X =$ the number of occurrences in the interval of interest.

X takes on the values $x = 0, 1, 2, 3, \ldots$

The mean μ is typically given.

The variance is $\sigma = \mu$, and the standard deviation is

$$\sigma = \sqrt{\mu} \tag{4.7.4}$$

.

When $P(\mu)$ is used to approximate a binomial distribution, $\mu = np$ where n represents the number of independent trials and p represents the probability of success in a single trial.

Use the following information to answer the next six exercises: On average, a clothing store gets 120 customers per day.

? Exercise 4.7.8

Assume the event occurs independently in any given day. Define the random variable X.

? Exercise 4.7.9

What values does X take on?

Answer

0, 1, 2, 3, 4, …

? Exercise 4.7.10

What is the probability of getting 150 customers in one day?

? Exercise 4.7.11

What is the probability of getting 35 customers in the first four hours? Assume the store is open 12 hours each day.

Answer

0.0485

? Exercise 4.7.12

What is the probability that the store will have more than 12 customers in the first hour?

? Exercise 4.7.13

What is the probability that the store will have fewer than 12 customers in the first two hours?

Answer

0.0214

? Exercise 4.7.14

Which type of distribution can the Poisson model be used to approximate? When would you do this?

Use the following information to answer the next six exercises: On average, eight teens in the U.S. die from motor vehicle injuries per day. As a result, states across the country are debating raising the driving age.

? Exercise 4.7.15

Assume the event occurs independently in any given day. In words, define the random variable X.

Answer

$X =$ the number of U.S. teens who die from motor vehicle injuries per day.

Exercise 4.7.16

$X \sim$ _____(_____,_____)

Exercise 4.7.17

What values does X take on?

Answer

$0, 1, 2, 3, 4, \ldots$

Exercise 4.7.18

For the given values of the random variable X, fill in the corresponding probabilities.

Exercise 4.7.19

Is it likely that there will be no teens killed from motor vehicle injuries on any given day in the U.S? Justify your answer numerically.

Answer

No

Exercise 4.7.20

Is it likely that there will be more than 20 teens killed from motor vehicle injuries on any given day in the U.S.? Justify your answer numerically.

Glossary

Poisson Probability Distribution

a discrete random variable (RV) that counts the number of times a certain event will occur in a specific interval; characteristics of the variable:

- The probability that the event occurs in a given interval is the same for all intervals.
- The events occur with a known mean and independently of the time since the last event.

The distribution is defined by the mean μ of the event in the interval. Notation: $X \sim P(\mu)$. The mean is $\mu = np$. The standard deviation is $\sigma = \sqrt{\mu}$. The probability of having exactly x successes in r trials is $P(X = x) =$

$$\left(e^{-\mu}\right) \frac{\mu^x}{x!} \tag{4.7.5}$$

. The Poisson distribution is often used to approximate the binomial distribution, when n is "large" and p is "small" (a general rule is that n should be greater than or equal to 20 and p should be less than or equal to 0.05).