Discrete Probability Distributions

5CS037

Contents

Discrete Random Variable

- When the outcome of a random experiment is a numerical value, the outcome is called a **random variable**.
- A random variable can also be interpreted as a numerical value that's determined by chance.
- Random variables are commonly notated through the use of capital letters from the end of the alphabet.
- Most commonly a random variable will be denoted as **X** or **Y**.

- Many of the random experiments which have been previously introduced throughout the course can be represented as random variables

Example #1:

Given a random experiment which consists of rolling a fair six-sided die, define the variable as:

$$X = \text{The number of pips showing when the die lands.}$$

Then it can be said that X is a *random variable* whose possible values are 1, 2, 3, 4, 5, and 6.

Mean of a Discrete Random Variable

- The mean of a random variable is typically referred to as an **"expected value."**
- Expected values are typically denoted by:

$$\boxed{\begin{array}{l} 1.\ E(X) \\ 2.\ \mu \end{array}}$$

1. $E(X)$
2. μ

1. They are commonly interpreted as a long-run average. That is, if one were to run the experiment represented by our random variable repeatedly, and then average the outcomes of those experiments the result would be approximately equal to the ***expected value***.

2. The expected value can be interpreted as the population's mean when the probability distribution is representing a population.

3. The expected value can be interpreted as a weighted average of the possible values of a random variable.

Formal Definition for the Expected Value of a Discrete Random Variable

The mean of a discrete random variable, or expected value is formally defined as below:

$$\mu = \sum x * P(X = x)$$

where the summation covers all possible values of x of the random variable X.

Given a single roll of a six-sided die, find the expected value of a given die roll:

- Step 1. Determine the possible outcomes of the experiment
  - $S = 1, 2, 3, 4, 5, 6$
- Step 2. Determine the probabilities of each possible outcome

x	P(X = x)
1	$\frac{1}{6}$
2	$\frac{1}{6}$
3	$\frac{1}{6}$
4	$\frac{1}{6}$
5	$\frac{1}{6}$
6	$\frac{1}{6}$

- Step 3. Use the values to calculate the expected value
  - $\mu = \sum x * P(X = x)$
  - $\mu = x_1 * P(X = x_1) + x_2 * P(X = x_2) + \ldots + x_6 * P(X = x_6)$
  - $\mu = 1 * \frac{1}{6} + 2 * \frac{1}{6} + 3 * \frac{1}{6} + 4 * \frac{1}{6} + 5 * \frac{1}{6} + 6 * \frac{1}{6}$
  - $\mu = \frac{1}{6} + \frac{2}{6} + \frac{3}{6} + \frac{4}{6} + \frac{5}{6} + \frac{6}{6} = \frac{21}{6} = 3.5$

Another approach

x	P(X = x)	x * P(X =)
1	$\frac{1}{6}$	$\frac{1}{6}$
2	$\frac{1}{6}$	$\frac{2}{6}$
3	$\frac{1}{6}$	$\frac{3}{6}$
4	$\frac{1}{6}$	$\frac{4}{6}$
5	$\frac{1}{6}$	$\frac{5}{6}$
6	$\frac{1}{6}$	$\frac{6}{6}$
Sum	1	3.5

Standard Deviation of a Discrete Random Variable

Similar to finding the mean of a random variable, finding the variance and standard deviation of a random variable can help interpret and understand it's nature. Common notations for populations (σ, σ^2) are typically used for denoting the variance and standard deviation of a random variable.

The variance of a discrete random variable is formally defined as below:

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

The standard deviation of a discrete random variable is formally defined as below:

$$\sigma = \sqrt{\sum (x - \mu)^2 * P(X = x)}$$

where the summation covers all possible values of x of the random variable X.

Again, consider the experiment of a single six-sided die roll and find the variance:

x	P(X = x)	$(x-\mu)$	$(x-\mu)^2$	$(x-\mu)^2 * P(X=x)$
1	$\frac{1}{6}$	-2.5	6.25	1.042
2	$\frac{1}{6}$	-1.5	2.25	0.375
3	$\frac{1}{6}$	-0.5	0.25	0.0417
4	$\frac{1}{6}$	0.5	0.25	0.0417
5	$\frac{1}{6}$	1.5	2.25	0.375
6	$\frac{1}{6}$	2.5	6.25	1.042

- $\sigma^2 = 1.042 + 0.375 + 0.0417 + 0.0417 + 0.375 + 1.042 = 2.92$

Now, use the variance to calculate the standard deviation:

- $\sigma = \sqrt{\sigma^2} = \sqrt{2.92} = 1.71$

Let the discrete random variable X represent the number of rooms in a randomly selected rental housing unit in Denver. Given the following table representing our random variable, calculate the expected value:

x	P(X = x)
1	0.01
2	0.03
3	0.25
4	0.35
5	0.20
6	0.10
7	0.04
8	0.02

x	P(X = x)
1	0.01
2	0.03
3	0.25
4	0.35
5	0.20
6	0.10
7	0.04
8	0.02

First, begin by calculating the product of x and $P(X = x)$, then by summing the values find the expected value.

This can be interpretted that if one were to randomly select a rental unit in Denver, on average, it would have 4.26 bedrooms.

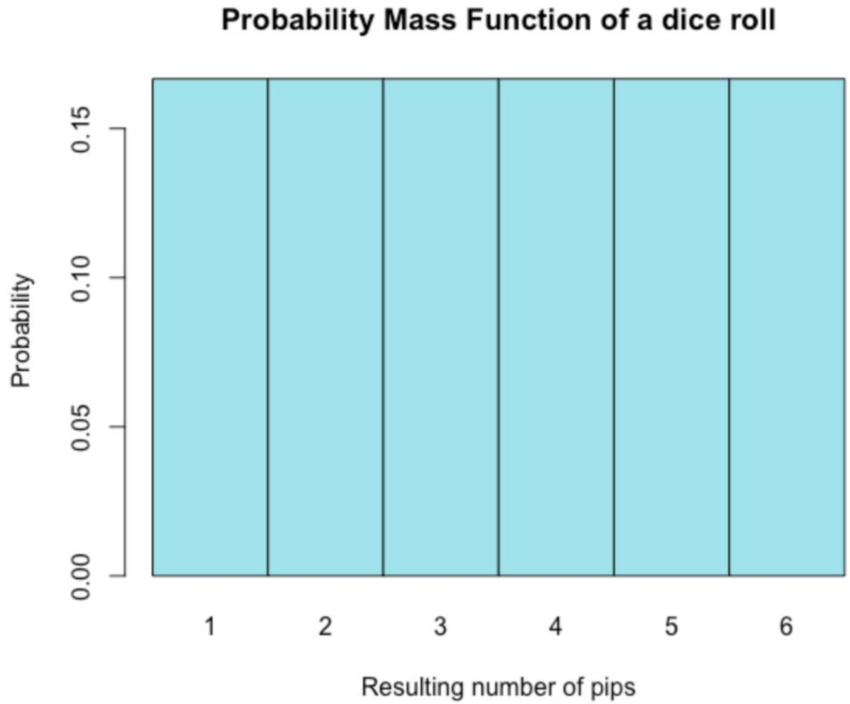
Probability Mass Function

- A natural thing to do with random variables is to consider how the probabilities differ across the different values that the random variable can take on.
- Since random variables can be defined in infinite ways, there are an infinite number of probability distributions.

- However, many random processes exhibit similarities that show up as commonalities in their probability distributions.

The distribution of outcomes of rolling a die or flipping a coin is uniform (assuming, of course the object is not "loaded" or unfair). When rolling a standard, 6-sided die, each value, 1 through 6, has a 1/6 chance of appearing. In this case the random variable would be "the number (or number of pips) on the upward face of a die." The fair die with this random variable results in this PMF:

- *Probability Mass Function* of a six sided die
  - $\frac{1}{6}$ for $1 \leq x \leq 6$
  - 0 otherwise



What's important to note is that, using the PMF, the exact probability of rolling exactly a 5 can be computed: $P(X = 5) = 1/6$. The PMF can also be used to calculate the probabilities for events which have multiple outcomes.

Calculate the probability of rolling a six-sided die and getting a result which shows three or fewer pips on the die.

- Step 1. Determine the summation necessary to represent the problem
    - $P(X \leq 3) = P(X=1) + P(X=2) + P(X=3)$

- Step 2. Use the PMF to calculate the individual probabilities
    - $P(X=1) = \frac{1}{6}$
    - $P(X=2) = \frac{1}{6}$
    - $P(X=3) = \frac{1}{6}$

- Complete the calculation
    - $P(X \leq 3) = P(X=1) + P(X=2) + P(X=3)$
    - $P(X \leq 3) = 1/6 + 1/6 + 1/6 = \frac{3}{6} = .5$

Cumulative Distribution Function

- Unlike the probability mass function or PMF, the cumulative distribution function (CDF) can be used to describe either continuous or discrete random variables. This lecture will focus specifically on the discrete instances of the CDF.

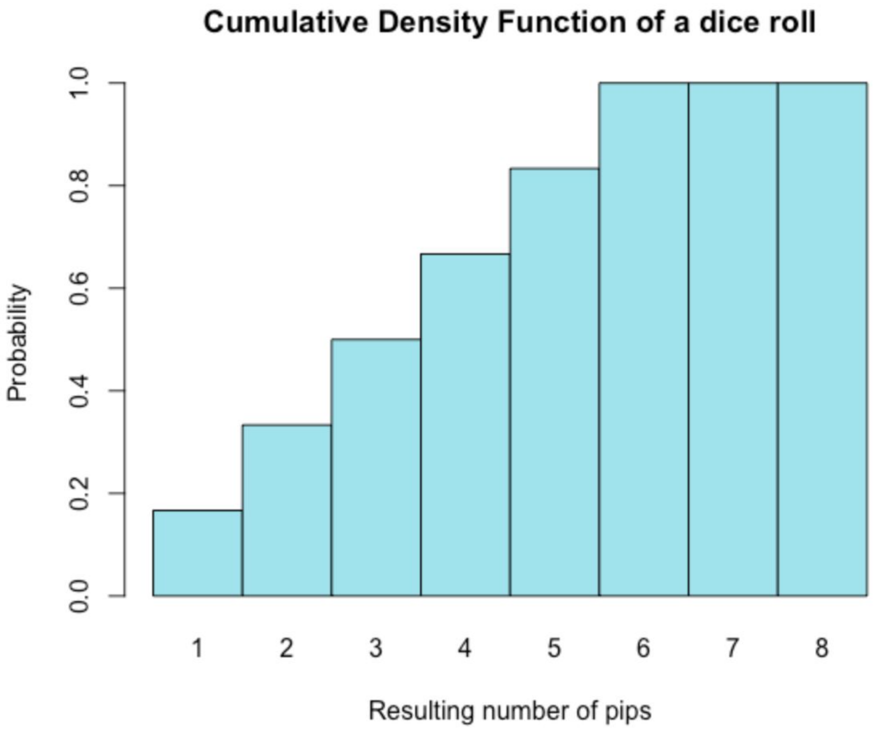
In the previous lesson, the ***PMF*** of a fair, six-sided die roll was given. From that, it was shown that we could calculate the probability for a ***range*** of values; from the PMF, one can generate a CDF which generalizes and simplifies this process. Where the PMF's purpose is to describe the probability of an exact outcome $P(X = x)$, the CDF represents ***cumulative*** probabilities from negative infinity to the given value $P(X \leq x)$. See an example below of the CDF for a six-sided die roll.

Whereas the PMF is typically notated as a lower-case $f_X(x)$ with the subscript of the random variable being represented, the CDF is notated similarly with a capital F as shown below.

$$\text{CDF} = F_X(x) = \begin{cases} 0 & : x < 1 \\ \frac{x}{6} & : 1 \leq x \leq 5 \\ 1 & : x > 5 \end{cases}$$

- Interpretation:
  - The *cumulative* probability for values less than 1 is zero
  - the *cumulative* probability for values between 1 and 5 is $\frac{x}{6}$
  - The *cumulative* probability for values greater than 5 is 1

Similar to the PMF, the CDF is often represented visually with the use of a histogram, see below for the histogram of the six-sided die roll:



An insurance company allows its policyholders to pay their premiums in a number of different ways. Let the random variable X represent the number of months between payments. The CDF of X is shown below:

$$F_X(x) = \begin{cases} 0.00 & : \ x < 1 \\ 0.30 & : \ 1 \leq x < 3 \\ 0.40 & : \ 3 \leq x < 4 \\ 0.45 & : \ 4 \leq x < 6 \\ 0.60 & : \ 6 \leq x < 12 \\ 1.00 & : \ x \geq 12 \end{cases}$$

Part #1:

Using only the CDF, calculate $P(3 \leq X \leq 6)$

Solution:

- Step 1. Determine the approach to solving the problem
  - $P(3 \leq X \leq 6) = P(X \leq 6) - P(X < 3)$
- Step 2. Complete the calculation
  - $P(3 \leq X \leq 6) = 0.60 - 0.30 = 0.30$

Part #2:

Using only the CDF, calulate $P(X \geq 4)$

Solution:

- Step 1. Determine the approach to solving the problem
    - Since the CDF produces values which are less than or equal to the value of x, the complement must be considered. Given the CDF, $P(X \leq 3)$ can be calculated.
        - *Note:* $P(X \leq 3)$ is the complement of $P(X \geq 4)$

- Step 2. Complete the calculation
    - Using the complement rule, $P(X \geq 4) = 1 - P(X \leq 3)$
    - $P(X \geq 4)$ = 1 - 0.40 = 0.60)

Bernoulli distribution

- The Bernoulli distribution is named after Swiss mathematician Jacob Bernoulli, and it represents random variables which consist of a single trial that has two possible outcomes.

Bernoulli distribution

- The Bernoulli distribution is named after Swiss mathematician Jacob Bernoulli, and it represents random variables which consist of a single trial that has two possible outcomes.
- These outcomes are labeled as either a "success," or a "failure." This single trial is referred to as a "Bernoulli Trial".
- Successes are denoted with the value 1, and failures are assigned the value zero.
- The canonical example of a Bernoulli trial is the flip of a coin. (H or T - 1 or 0 - Success or Failure)

For example, if a heads result is considered to be a success, and we represent the outcome of our trial with the random variable X; $X = 1$ denotes a flip resulting in heads, and $X = 0$ denotes a flip resulting in tails.

Given that the coin is fair, it can also be stated that:

$$P(heads) = P(success) = P(X = 1) = 0.5$$

$$P(tails) = P(failure) = P(X = 0) = 0.5$$

- **Note:** It's not required that the two outcomes be equally likely, that is only the case with this given example. However, the two probabilities must always sum to 1, because they are collectively exhaustive & mutually exclusive.
- Another example: The probability that a website visitor will purchase an item might be:

$$\boxed{P(\text{makes a purchase}) = P(X=1) = p = 0.02}$$

$$\boxed{P(\text{doesn't make a purchase}) = P(X=0) = 1 - p = 1.0 - 0.02 = 0.98}$$

Notation

When a random variable (In this case X) follows a Bernoulli Distribution, the common notation is:

$$X \sim Bernoulli(p)$$

Much like our examples, there is one input value, or ***parameter*** p.

Definition: Expected Value

$$E(X) = p \times 1 + (1 - p) \times 0 = p$$

Applications

- The Bernoulli distribution is commonly used to model binary outcomes or success/failure events in various fields, such as medical diagnosis (e.g., presence or absence of a disease), quality control (defective or non-defective items), and in surveys (yes/no responses).

- The Bernoulli distribution serves as the building block for more complex distributions, such as the Binomial distribution (a series of independent Bernoulli trials) and the Geometric distribution (the number of trials needed to get the first success in a sequence of independent Bernoulli trials).

The Binomial distribution

- It is used to model a random variable which represents the number of success in a sequence of Bernoulli Trials.
- The PMF of the binomial distribution can be used to model probabilities which represent the number of heads in ten successive coin flips.

Definition: Probability Mass Function

$$P(X = k) = \binom{n}{k} p^k (1-p)^{n-k}$$

Where n represents the number of trials, and k represents the number of successes for which we're calculating the probability of.

When the random variable X follows a binomial distribution it is typically denoted as below:

$$X \sim binomial(n, p) \text{ or } X \sim Bin(n, p)$$

Where the Bernoulli distribution only has one parameter p, which represents the probability of success, the binomial distribution takes two parameters.

- $p =$ the probability of success in a given trial
- $n =$ the number of trials

NBC News reported in May of 2013, that 1 in 20 children in the USA have some sort of food allergy. Consider selecting a random sample of 25 children and let the random variable X represent the number of children from our sample who have an allergy.

Determine the probability that in the given sample there will be three children with a food allergy.

- Step 1. Identify the given information in the problem
  - $n =$ number of Bernoulli Trials $= 25$
  - $k =$ number of success $= 3$
  - $p =$ probability of success in a given trial $= 0.05$

- Step 2. Calculate the probabilities using the binomial PDF
  - $P(X = k) = \binom{n}{k} * p^k * (1-p)^{n-k}$
  - $P(X = 3) = \binom{25}{3} * 0.05^3 * (1-0.05)^{25-3} = 0.093$

- Step 3. Interpret the result
  - The PMF has given the probability 0.093, or 9.3%; given 25 randomly chosen children from the USA, there is a 9.3% chance that three of them will have some sort of food allergy.

Cumulative Distribution Function

Unlike some distributions, there is not a powerful general equation for the CDF of a binomial distribution. Instead, to derive the CDF, simply take an additive approach. See the formula below:

Definition: Cumulative Distribution Function

$$P(X \leq k) = \sum_{i=0}^{k} \binom{n}{i} p^i (1-p)^{n-i}$$

- Step 1. Identify the given information in the problem
  - $n =$ number of Bernoulli Trials $= 25$
  - $k =$ number of success $= 3$
  - $p =$ probability of success in a given trial $= 0.05$

- Step 2. Determine an approach to solving the problem
  - $P(X \leq 3) = P(X = 3) + P(X = 2) + P(X = 1) + P(X = 0)$

- Step 3. Calculate the probabilities using the binomial PDF
  - $P(X = 3) = \binom{25}{3} \cdot 0.05^3 * (1 - 0.05)^{25-3} = 0.093$
  - $P(X = 2) = \binom{25}{2} \cdot 0.05^2 * (1 - 0.05)^{25-2} = 0.231$
  - $P(X = 1) = \binom{25}{1} \cdot 0.05^1 * (1 - 0.05)^{25-1} = 0.365$
  - $P(X = 0) = \binom{25}{0} \cdot 0.05^0 * (1 - 0.05)^{25-0} = 0.277$

- Step 4. Apply the summation
    - $P(X \leq 3) = P(X = 3) + P(X = 2) + P(X = 1) + P(X = 0)$
    - $P(X \leq 3) = 0.093 + 0.231 + 0.365 + 0.277 = 0.966$
- Step 5. Interpret the result
    - Our summation shows that there is a 0.966 probability, or 96.6% chance that a given sample of 25 will have three or fewer children with a food allergy.

Expected Value

The expected value for the binomial distribution is relatively easy to compute. See below for the definition:

> ⓘ **Definition: Expected Value**
>
> $$E(X) = \bar{X} = \mu_X = np$$

Note: The notation μ is often used to denote the mean or expected value of a random variable, because in general the binomial distribution represents a population or a "long-run average".

Definitions

Variance:

$$Var(X) = \sigma^2 = np(1-p)$$

Standard Deviation:

$$\sigma = \sqrt{\sigma^2} = \sqrt{np(1-p)}$$

Note:

The notation σ is often used to denote the variance or standard deviation of a random variable, because in general the binomial distribution represents a population or in this case a "long-run variation".

Geometric Distribution

- The Geometric Distribution is used to model some sequence of Bernoulli Trials. Whereas the Binomial Distribution models **k** number of successes from **n** trials, the Geometric Distribution models ***how many trials until the first success is encountered.***
- Once the first success is encountered, the random experiment is considered complete.
- The **k** represents the number of trials up to and including the first success, i.e. the first success happens on the **kth** trial.

The other common interpretation that is used models k to represent how many failures occurred BEFORE the first success.

Notation

When the random variable X follows a Geometric Distribution, it is typically denoted as below:

$$X \sim geometric(p) \text{ or } X \sim G(p)$$

Like the Bernoulli Distribution, the Geometric Distribution only has one parameter p, which represents the probability of success on a single Bernoulli Trial.

- $p =$ The probability of success on a single Bernoulli Trial

Geometric Distribution Probability Mass Function

The formal definition for the Geometric Distribution's *PMF*, where k represents the trial in which the first success is encountered is as follows below:

$$P(X = k) = (1 - p)^{k-1} * p; \qquad \text{for } 1 \leq k \leq \infty$$

Note: Note that the value(s) which are possible for k begin at one, and continue to infinity. This is because k represents the trial in which the first success is encountered, and a success cannot occur until the first trial. Theoretically, it is also possible that one could conduct a Bernoulli try infinitely before encountering their first success, though for most trials, the probability of this occuring is very low.

Example #1:

Let the random variable X represent the outcome of a random experiment consisting of sequential rolls of a fair, six-sided die. A "successful" roll of the die is defined as rolling the die resulting in six pips showing (reminder: a pip is the name of the dot on the dice). The experiment is ended after the first success is encountered. **What is the probability that the first roll resulting in a six will occur on the tenth roll?**

- Step 1. Identify the given information in the problem
  - $P(X = 6) = 1/6$
  - $k = 10$

- Step 2. Calculate the probability using the binomial *PMF*
  - $P(X = k) = (1 - p)^{k-1} * p$
  - $P(X = 10) = (1 - \frac{1}{6})^{10-1} * \frac{1}{6} = 0.0323$

- Step 3. Interpret the result
  - Shown above, the probability that the first six will be rolled on the tenth roll is $0.0323 \text{ or } 3.23$

Cumulative Density Function

The Geometric Distribution has a generalized function that can be used for the *CDF*. Unlike some other discrete probability distributions like the Poisson or the Binomial, there is no need to compute a summation in order to find the cumulative probability for some given value. See below for the definition of the *CDF*:

$$P(X \leq k) = 1 - (1-p)^k; \qquad \text{for } 1 \leq k \leq \infty$$

Example #2:

Given the same random experiment from ***Example #1***, find the probability that the first die roll resulting in six pips showing will occur on or before the fourth die roll.

Example #2:

Given the same random experiment from ***Example #1***, find the probability that the first die roll resulting in six pips showing will occur on or before the fourth die roll.

- Step 1. Identify the given information in the problem
  - $p = \frac{1}{6}$
  - Trying to find $P(X \leq 4)$

- Step 2. Apply the CDF
  - $P(X \leq k) = 1 - (1-p)^k$
  - $P(X \leq 4) = 1 - (1 - \frac{1}{6})^4 = 0.518$

- Step 3. Interpret the solution
  - The probability that the first six in a sequence of die rolls happening on or before the fourth die roll is 0.518, or 51.8

Expected Value

The *Expected Value* for the Geometric Distribution is relatively easy to compute. See below for the definition:

$$E(X) = \mu = \frac{1}{p}$$

Note: The notation μ is often used to denote the mean or expected value of a random variable, because in general the geometric distribution represents a population or a "long-run average".

Variance and Standard Deviation

There is also a special formula for the variance of a Geometric Distribution which is much less computationally expensive than the calculation of variance for a generic random variable. See below for the definition:

$$Var(X) = \sigma^2 = \frac{1-p}{p^2}$$

As usual, to derive the standard deviation, take the square root of the variance, as shown below:

$$\bar{X} = \sqrt{\sigma^2} = \sqrt{\frac{1-p}{p^2}}$$

Note: The notation σ is often used to denote the variance or standard deviation of a random variable, because, in general, the geometric distribution represents a population or, in this case, a "long-run variation".

Example #3:

Determine the mean (expected value) and standard deviation of the previous problem.

- Step 1. Identify the given information in the problem
    - $p = \frac{1}{6}$

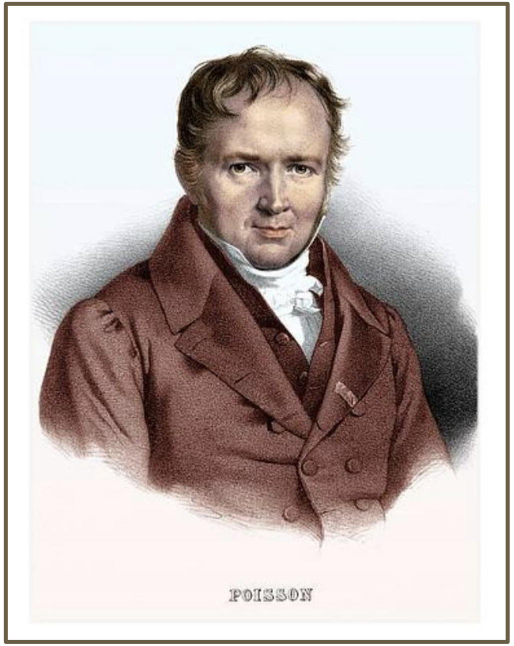
- Step 2. Apply the formula for expected value
    - $E(X) = \frac{1}{p}$
    - $E(X) = \frac{1}{\frac{1}{6}} = 6$

- Step 3. Apply the formulas for variance and standard deviation
  - $\sigma^2 = \frac{1-p}{p^2}$
  - $\sigma^2 = \frac{1-\frac{1}{6}}{\frac{1}{6}^2} = 30$
  - $\sigma = \sqrt{\sigma^2}$
  - $\sigma = \sqrt{30} = 5.478$

- Step 4. Interpret the solution
  - On average, you would expect to get the die roll resulting with six pips showing on the sixth sequential die roll
  - The standard deviation is 5.478. One could say they would expect to get the first six on the sixth roll, give or take 5 rolls.

Poisson distribution

- The poisson random variable, or poisson distribution is named after the French mathematician **Simeon Denis Poisson.**



Poisson distribution

- The poisson random variable, or poisson distribution is named after the French mathematician **Simeon Denis Poisson.**
- This distribution allows one to model the probability of a given number of events occuring in some fixed interval of time or space.
- The poisson distribution makes a few assumptions; first, the events must occur at a previously known (relatively) constant rate. Second, we must assume that each occurence of an event must occur independently.

An example of a Poisson distribution is used to model watching for shooting stars in the night sky. Given that an individual has historically seen shooting stars in the night sky at a rate of 20 shooting stars per hour, the Poisson distribution's PMF gives the ability to calculate the probability that someone may see 30 shooting stars in some arbitrarily chosen hour.

Probability Mass Function Unlike the Bernoulli or binomial distributions previously discussed, the sample space for a poisson distribution is infinite in nature.

One calculates the probability of events occurring between zero and infinity.

Extremely large values may produce extremely small probabilities.

However, these tiny probabilities are not impossible.

$$P(X = k) = \frac{\lambda^k e^{-\lambda}}{k!}$$

Where k = the number of events occuring and λ = the known rate of events occuring.

- ***Poisson Distribution is a limiting case of Binomial Distribution*** under the following conditions:
- n, the number of trials is indefinitely large i.e. n□ ∞
- P, the constant probability of success for each trial is indefinitely small i.e p□ 0
- np = λ (constant and finite)

Definition: Probability Mass Function

$$P(X = k) = \binom{n}{k} p^k (1 - p)^{n-k}$$

Where n represents the number of trials, and k represents the number of successes for which we're calculating the probability of.

Now, we obtain probability mass function as a limiting case as n□∞, p□0 and np = λ

$$P(X=k) = \lim_{n\to\infty} (nC_k)\left(\frac{\lambda}{n}\right)^k \left(1-\frac{\lambda}{n}\right)^{n-k} = \lim_{n\to\infty} \frac{n!}{(n-k)!\;k!} \frac{\lambda^k}{n^k} \left(1-\frac{\lambda}{n}\right)^n \left(1-\frac{\lambda}{n}\right)^{-k}$$

$$P(X=k) = \lim_{n\to\infty} \frac{n(n-1)(n-2)\ldots(n-k+1)}{n^k} \frac{\lambda^k}{k!} \left(1-\frac{\lambda}{n}\right)^n \left(1-\frac{\lambda}{n}\right)^{-k}$$

$$P(X=k) = \frac{\lambda^k}{k!} \times \lim_{n\to\infty} \frac{n(n-1)(n-2)\ldots(n-k+1)}{n^k} \times \lim_{n\to\infty} \left(1-\frac{\lambda}{n}\right)^n \left(1-\frac{\lambda}{n}\right)^{-k}$$

$$P(X=k) = \frac{\lambda^k}{k!} \times \lim_{n\to\infty} \frac{n^k + n^{k-1} + n^{k-2} + \ldots}{n^k} \times \lim_{n\to\infty} \left(1-\frac{\lambda}{n}\right)^n \left(1-\frac{\lambda}{n}\right)^{-k}$$

$$P(X=k) = \frac{\lambda^k}{k!} \times 1 \times \lim_{n\to\infty} \left(1-\frac{\lambda}{n}\right)^n \times \lim_{n\to\infty} \left(1-\frac{\lambda}{n}\right)^{-k}$$

$$P(X=k) = \frac{\lambda^k}{k!} \times 1 \times e^{-\lambda} \times \lim_{n\to\infty} \left(1-\frac{\lambda}{n}\right)^{-k}$$

$$P(X=k) = \frac{\lambda^k}{k!} \times 1 \times e^{-\lambda} \times 1$$

$$P(X=k) = \frac{\lambda^k}{k!} e^{-\lambda}$$

Probability Mass Function(PMF) of Poisson's Distribution.

Example #1:

You are the data scientist for a prominent web company which receives a lot of traffic on their website, however they are struggling to attract new users to their website.

A new urgent request just came in! The ops team needs to push a new crucial bug fix, but will disable the site and it will go down for 5 minutes while the changes are deployed. Assume that on average, ten users visit the site every 5 minutes.

What is the probability that during the 5 minute period where the site is down, that no one will visit the site?

(Assume that each visitor visits the site independently, and that the rate given is constant and accurate)

- Step 1. Write down all of the given information
  - $\lambda =$ the historical rate of visitors per 5 minutes $= 10$
  - $k =$ the # of visitors for which the probability is to be calculated $= 0$

- Step 2. Apply the PMF of the Poisson distribution
  - $P(X = k) = \frac{\lambda^k e^{-\lambda}}{k!}$
  - $P(X = 0) = \frac{10^0 e^{-10}}{0!} = e^{-10} = 0.000045$

- Step 3. Interpret the solution
  - Shown here, there is an incredibly low probability that the site will not see any visitors in the given time frame.

Notation

When a random variable follows a Poisson distribution, it is typically notated as follows:

$$X \sim \mathrm{Poisson}(\lambda\backslash) \text{ or } X \sim Po(\lambda\backslash)$$

Note: Note that there is only a single parameter, λ for a Poisson distribution.

The Poisson Process

The poisson process allows the use of the Poisson distribution to calculate probabilities for different variants of time or space than the given rate. For example, in our previous problem defines a rate of an event's occurence, to apply the process, we will assign this value to the variable α. Second, to apply the Poisson process, we are given a second variable t which in this case represents the amount of time in the interval one would like to compute a probability for. See the definition below to calculate a "new" lambda to use in the Poisson PMF:

$$\lambda = \alpha t$$

After seeing projections and probabilities for the number of website visitors who will be affected by the outage, the engineering team has gone to work, and has come up with a fix which will only take 2 minutes to implement. Given this new fix, determine the probability at least one visitor will visit the site during the outage.

(Assume that each visitor visits the site independently, and that the rate given is constant and accurate)

- Step 1. Write down all of the given information
    - $\alpha =$ the historical rate of visitors per 5 minutes $= 10$
    - $k =$ the # of visitors for which the probability is to be calculated $= 0$

- Step 2. Determine the "new" lambda value
    - $\lambda = \alpha t$
    - $t = \frac{2}{5}$
        - *Note:* use algebra to compute t
        - In this case, to convert our given time period of 5 minutes to a time period of two minutes t must be equal to $\frac{2}{5}$
    - $\lambda = 10 * \frac{2}{5} = \frac{20}{5} = 4$

- Step 3. Apply the Poisson PMF with the new λ value
  - $P(X = k) = \frac{\lambda^k e^{-\lambda}}{k!}$
  - $P(X = 0) = \frac{4^0 e^{-4}}{0!} = e^{-4} = 0.018$

- Step 4. Interpret the solution:
  - There is still only a (approximately) 2 percent chance that the site will see zero visitors in this shortened time-frame. That is, there is (approximately) a 98% chance there will be at least one visitor affected by the outage.

Cumulative Distribution Formula for the Poisson Distribution

Much like the binomial distribution, there is not a unique or special formula to represent the CDF of a Poisson distribution, so again, a summation approach works best. See below for the definition and an example implementation:

$$P(X \leq k) = \sum_{i=0}^{k} \frac{\lambda^k e^{-\lambda}}{k!}$$

Example #3:

Given the previous examples, the ops team has determined that they are willing to accept that some users will be affected by the outage. They have found that 4 or fewer affected visitors during the 2 minute outage is an acceptable number of users to inconvenience, and they would like to know that the probability of having more than 4 visitors affected will be less than 35%. Is the two-minute deployment short enough to meet their demands?

- Step 1. Write down all of the given information
  - $\alpha =$ the historical rate of visitors per 5 minutes $= 10$
  - $k =$ the # of visitors for which the probability is to be calculated $= 0$
  - $\lambda = 10 * \frac{2}{5} = \frac{20}{5} = 4$

- Step 2. Determine the approach to calculating that 4 or more visitors will be affected
    - $P(X \geq 5) = 1 - P(X \leq 4)$
    - $P(X \leq 4) = P(X = 4) + P(X = 3) + P(X = 2) + P(X = 1) + P(X = 0)$
- Step 3. Apply the Poisson PMF to calculate the values
    - $P(X = k) = \frac{\lambda^k e^{-\lambda}}{k!}$
    - $P(X = 0) = \frac{4^0 e^{-4}}{0!} = 1 * e^{-4} = 0.018$
    - $P(X = 1) = \frac{4^1 e^{-4}}{1!} = 4 * e^{-4} = 0.073$
    - $P(X = 2) = \frac{4^2 e^{-4}}{2!} = 16 * e^{-4} = 0.147$
    - $P(X = 3) = \frac{4^3 e^{-4}}{3!} = 64 * e^{-4} = 0.195$
    - $P(X = 4) = \frac{4^4 e^{-4}}{4!} = 256 * e^{-4} = 0.195$

- Step 4. Apply calculations to the previously determined approach
    - $P(X \leq 4) = P(X = 4) + P(X = 3) + P(X = 2) + P(X = 1) + P(X = 0)$
    - $P(X \leq 4) = 0.195 + 0.195 + 0.147 + 0.073 + 0.018 = 0.628$
    - $P(X \geq 5) = 1 - P(X \leq 4)$
    - $P(X \geq 5) = 1 - 0.628 = 0.372$
- Step 5. Interpret the solution
    - Though this solution comes very close to meeting the desires of the operations team, the 2-minute outage still produces a greater than 35% probability of there being more than 4 visitors during the outage.

Expected value and standard deviation of a Poisson Distribution

Here, we have possibly the best/easiest shortcuts to determine both the mean (or expected value), and the variance of a poisson random variable, they are defined as follows:

$$E(X) = \bar{X} = \mu_X = \lambda$$

$$Var(X) = \lambda = \sigma_X^2$$

$$sd(X) = \sigma_X = \sigma = \sqrt{\sigma_X^2} = \sqrt{\lambda}$$

Note: The notation μ is often used to denote the mean or expected value of a random variable, and σ is used to denote the standard variation, because in general the Poisson distribution represents a population or "long-run average," and "long-run variation," respectively.

Thank You