What are Exponential Distributions

Edris Safari

Bellevue University, Nebraska U.S.A.

Abstract

Exponential distributions are distributions in which events takes place before or after a period of time. This paper describes it in more detail.

What are Exponential Distributions

We've learned that a distribution is a set of finite observations (i.e. head or tail) that occur a number of times after a finite number of observations are made (i.e. tossing a coin 10 times). We've used histograms, simplified them with PMFs and CDFs to make better sense of the data by looking at its shapes. With PDM, we calculated the probability of an observation rather than the number of observations. The value of the probability being between 0 and 1 gave us a better view of the distribution. We evaluated mean, and variance of the PMF whose functions are bit different than the distributions that we can plot histograms for. The mean of a PMF distribution is

$$Mean(x) = \sum PMF(xi) * xi$$

, and the variance is

$$S^2 = \sum PMF(xi) * (xi - mean(xi)$$

While PMF works well for distribution with a small sample space, it does not for distributions with non-discrete and thus larger sample space. Cumulative Distribution Functions (CDFs) can be used to mitigate this issue. The CDF is a function that maps a value to its percentile rank. Percentile rank is a fraction of scores or values less than or equal to a given value. The equation for percentile rank is

Percentile_rank = 100*Count/number of values in sample set

Count is the number of values in the distribution that are less than or equal to a given value. So if x is in the 90th percentile rank, x is greater than equal to 90% of the values in the distribution. The CDF of a value x in a distribution is percentile_rank/100 making the value of CDF between 0 and 1-as in PMF but with larger sample sets.

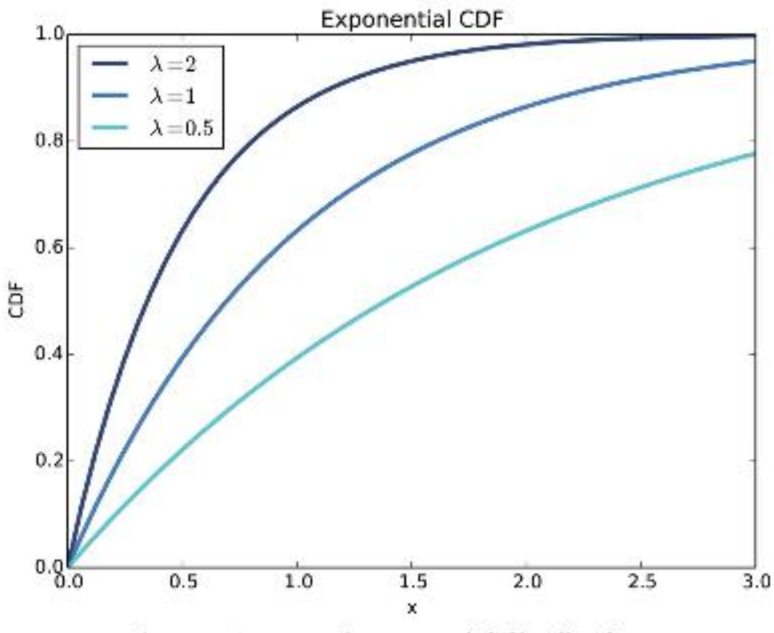
The distributions in PMF and CDF are based on empirical or verifiable distributions with finite samples. When faced with larger distributions and infinite samples, we resort to computing the properties of the distribution using a CDF function. As we know, CDF is the probability that an outcome is less than or equal to a given value.

When a distribution is composed of a series of events and there are measured times between events, then that distribution is called an exponential distribution. For example, the amount of time it takes for a car to pass through a tollgate or the amount of time it takes before the next customer walks in or visits a web site have an exponential distribution. The CDF function for exponential distribution is:

$$CDF(x) = 1 - e^{-\lambda x}$$

Where the parameter

λ is the mean of the distribution and it value shapes the distribution as shown below:



Exponential distribution is the inverse of Poisson distribution which has the property of the number of events in a given time period (i.e. number of cars passing a tollgate in an hour). Usually we have a PMF and/or a CDF distribution for the Poisson distribution with a given mean value called lambda. For example, mean of 3 cars passing a tollgate in one hour.

Notice that the values in the Poisson distribution are discrete. To make them continuous so we consider the amount of time before the occurrence of an event and use the reciprocal of the mean of PMF λ . So instead of saying 3 cars per hour, we can say 20 minutes for 3 cars or $\mu = \frac{1}{\lambda}$. This turns the equation above to:

$$CDF(x) = 1 - e^{-x/\mu}$$

The resulting exponential distribution allows us to calculate how many cars within 25 minutes or after 12 minutes. However, we cannot calculate the number of events at an exact time such as at 15 minutes.

One usage of exponential distribution is in reliability. For example, in a factory the machines that perform various tasks have parts that need regular maintenance. Parameters such a mean time before failure and mean time to repair contribute to the overall equipment effectivity (OEE). With these measures, maintenance can be predicted and scheduled rather than reacting to a failure which is costly in terms of equipment down time.

References

- The Difference Between Poisson and Exponential Distributions From <https://www.youtube.com/watch?v=Z-8FtjZNlb4>
- The Exponential Distribution From https://courses.lumenlearning.com/introstats1/chapter/the-exponential-distribution/
- Probability Exponential Distribution Problems From <https://www.youtube.com/watch?v=J3KSjZFVbis>
- Exponential Distribution! Definition | Calculations | Why is it called "Exponential"? From Exponential Distribution! Definition | Calculations | Why is it called "Exponential"?

References

1. Rahul Gupta,Michael Koffie,Brandon May,Tushar Muley,Edris Safari –"As The World Churns: Customer Data, Business Models, and Predicting Customer Trends and Behaviors". Final project DSC500 Bellevue University
2. https://www.kdnuggets.com/2017/04/value-exploratory-data-analysis.html
3. Kelleher, John D. Data Science (MIT Press Essential Knowledge series) (p. 1). The MIT Press.