The Exponential Distribution

Jonathan Wilson

November 25, 2019

Introduction to The Exponential Distribution

Applications of the exponential distribution are often used as a model for durations assuming a constant rate parameter. So predicting the waiting time until the next event such as a success, failure, or arrival of some event would be a general use case for this distribution. Some specific use cases for this distribution that statisticians, engineers, or others might find to be interesting are:

- the amount of time you need to wait for something such as a flight
- The time until someone finishes a phone call
- the time until you have your next car accident
- the time until a particular component needs to be replaced in an engine
- the time until a beam should be checked for safety
- the time until a radioactive particle decays, or the time between beeps of a Geiger counter

Historically this distribution has been used in areas such as queuing theory, reliability engineering, calculating hazard rates, barometrics, and hydrology.

In real life however, the constant rate may not always be assumed as constant so it could be said that this model can make approximate predictions in the case that the rate is roughly constant. To illustrate, the rate of traffic incidents for a particular area might decrease as we move the current time on a time interval. The rate of traffic incident occurrences during the times from 1am to 4am during the week days might be much lower than the rate during rush hour intervals.

The exponential distribution is considered a probability distribution of the time between events and is related to the Poisson distribution which measures the time between successes, or a process in which events occur continuously and independently at a constant average rate. The distribution represents time intervals which makes it continuous and not discrete as in the case of other distributions.[1]

Derivations for the Exponential Distribution

Probability Density Function

The probability density function (pdf) of an exponential distribution is:

$$f(x;\lambda) = \begin{cases} \lambda e^{-\lambda x} & x \geq 0, \\ 0 & x < 0. \end{cases}$$

The parameter λ is called the rate parameter which determines the rate at which events occur. And the support for the distribution itself is $[0, \infty)$

To verify the distribution satisfies the properties of a valid probability density function we need to check:

$$1. \int_0^\infty \lambda e^{-\lambda x} dx = 1$$

[1] https://en.wikipedia.org/wiki/Exponential_distribution

To solve this we integrate over the support and use u-substitution to integrate. Then we check and see if this is indeed equal to one.

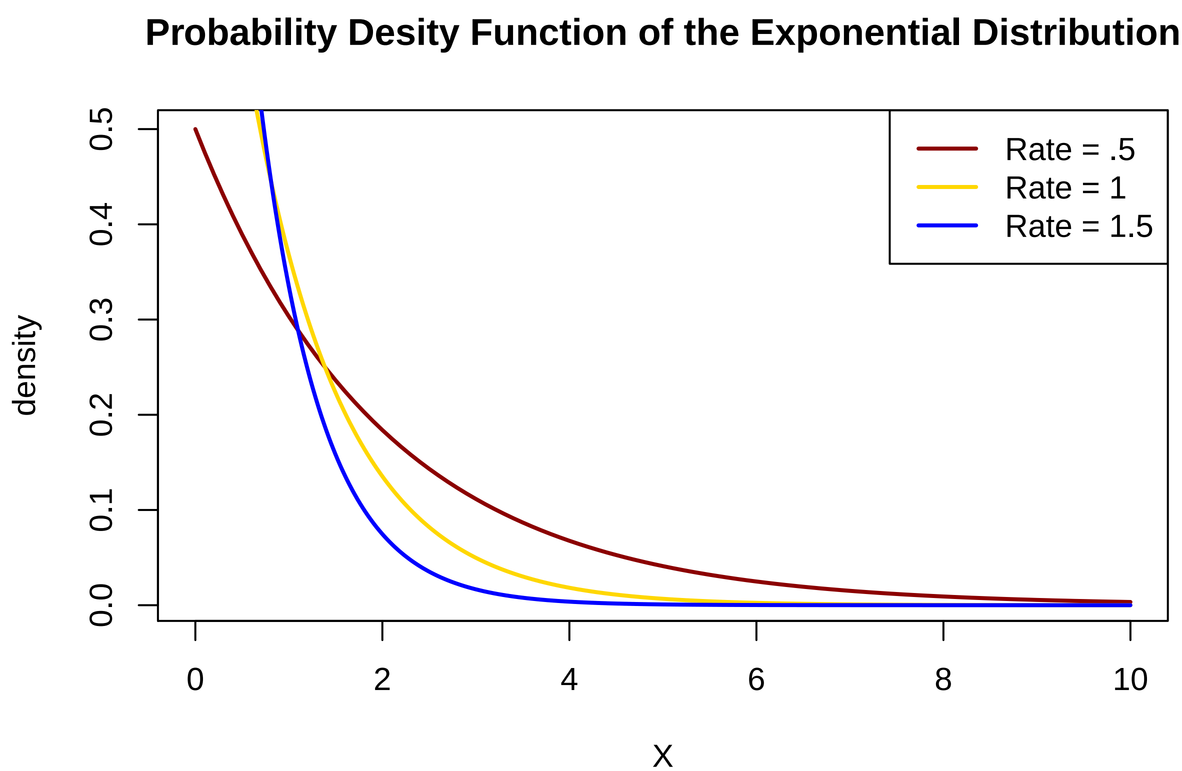
$$\int_0^\infty \lambda e^u(-\frac{1}{\lambda}du)$$
$$= \int_0^\infty e^u du$$
$$= -e^u|_{x=0}^{x=\infty} = -e^{-\lambda x}|_0^\infty = 0 - (-e^{-\lambda * 0}) = 1$$

So requirement one is met.

$$2.\lambda e^{-\lambda x} \geq 0, \forall x$$

For the second requirement we need to make sure the f(x) is greater than zero. We can check this via inspection. Looking at the first part we see that $\lambda > 0$ by what λ was defined on being greater than zero. Looking at $e^{-\lambda x}$ we know that $e^t \forall t$ is a positive number. Therefore f(x) is positive for all x.

Changes to the parameters change the center, spread, and shape of the curve. As the rate (λ) increases the curve decays faster. That is, the curve starts off very steep and levels off to zero quickly. The means tend to be smaller for rates less than one. The spread is not as great for larger rates. As the rate decreases the curve spreads our/stretches out the events or in other words the events happen more slowly. The means tend to be bigger than one for rates smaller than 1. These behaviors can be seen in the graph below using 3 different parameters for λ.



Cumulative Distribution Function

The cumulative distribution function can be derived using the the probability density function from above by integrating the PDF from zero over some variable k. Below is the derivation:

$$
\begin{aligned}
F(k) &= \int_0^k \lambda e^{-\lambda x} dx \\
&= -e^{-\lambda x}|_0^k \\
&= 1 - e^{-\lambda x}
\end{aligned}
$$

Which can be written as:

$$
F(x;\lambda) = \begin{cases} 1 - e^{-\lambda x} & x \geq 0, \\ 0 & x < 0. \end{cases}
$$

To verify this is a valid cumulative distribution function we must satisfy the follwing requirements:

1. $\lim_{x\to-\infty} F(x) = 0$ by the definition of the Cumulative Distribution Function we just defined above we already know $\lim_{x\to-\infty} 1 - e^{-\lambda x} = 0$
2. $\lim_{x\to\infty} F(x) = 1 : \lim_{x\to\infty} 1 - e^{-\lambda x} = 1$ This is because $e^{-\lambda x}$ goes to zero as x goes to infinity.
3. Nondecreasing? If $x_1 \geq x_2$ then we need to show $F(x_1) \geq F(x_2)$. Assume $x_1 \geq x_2 \iff \lambda x_1 \geq \lambda x_2 \iff -\lambda x_1 \leq -\lambda x_2 \iff e^{-\lambda x} \leq e^{-\lambda x} \Rightarrow -e^{-\lambda x} \geq -e^{-\lambda x} \iff 1 - e^{-\lambda x} \geq 1 - e^{-\lambda x}$ Therefore this cumulative distribution function is nondecreasing.
4. Right Coninuous? Since we already know this is a continuous distribution this is something we are not too concerned with as in the case with discrete random variables.

Thus we can conclude that this distribution is a valid cumulative distribution function.

Changes to the parameters change the center, spread, and shape of the cumulative distribution. As the rate increases we would ecpect to see the shape of the curve to reach 1 quickly as compared to when the rate is small we would see the curve shape take longer to get to 1. The center for a small rate tends to be larger than the center for larger rates.

Variance and Expected Value

The expected value of X can be derived using the law of the unconscious statistician and integration by parts. The steps for achieving this are below:

$$
E(X) = \int_0^\infty x\lambda e^{-\lambda x} dx
$$

$$
\begin{aligned}
&= \lambda[\frac{-xe^{-\lambda x}}{\lambda}]_0^\infty + \frac{1}{\lambda}\int_0^\infty e^{-\lambda x} dx \\
&= \lambda[0 + \frac{1}{\lambda}\frac{-e^{-\lambda x}}{\lambda}|_0^\infty] \\
&= \lambda\frac{1}{\lambda^2} = \frac{1}{\lambda}
\end{aligned}
$$

So

$$
E(x) = \frac{1}{\lambda}
$$

So as λ gets larger, the thing we are waiting for to happen tends to happen more quickly hence the nickname of the parameter as the rate. The variance of X can be found by using the short cut formula

$$
Var(X) = E[X^2] - E[X]^2
$$

. In this case we already have $E[X]$ we just need to square it. What we still need to find is E[X^2]. The process for solving this is exactly the same as we did when solving for E[X] we just need to plug in X^2 this time using the law of the unconscious statistician and integration by parts. The integration is not that much harder than we had done before and trival to even write out. The derivation is written below:

$$E(X^2) = \int_0^\infty x^2 \lambda e^{-\lambda x} dx = ... = \frac{2}{\lambda^2}$$

So we have,

$$Var(X) = E[X^2] - E[X]^2 = \frac{2}{\lambda^2} - \frac{1}{\lambda^2} = \frac{1}{\lambda^2}$$

Thus $Var(X) = \frac{1}{\lambda^2}$

Moments and Moment Generating Function

- If the parameters have natural interpretations as moments, describe those here.

The moment generating function is derived below:

$$E[e^{tx}] = \int_{-\infty}^{\infty} e^{tx} f(x) dx$$

$$= \int_0^\infty e^{tx} e^{-\lambda x} dx = \lambda \int_0^\infty e^{(t-\lambda)x} dx$$

$$= \lambda [\frac{1}{t-\lambda} e^{(t-\lambda)x}]_0^\infty = \lambda [0 - \frac{1}{t-\lambda}]$$

So the moment generating function for the exponential distribution is:

$$M_{X(t)} = \frac{\lambda}{\lambda - t}, t < \lambda$$

Using the moment generating function we can derive E(X):

$$E[X] = \frac{dM_{X(t)}}{dt}|_{t=0} = \frac{\lambda}{(\lambda - 0)^2} = \frac{1}{\lambda}$$

This is what we had derived earlier when we were deriving the expected value. The natural interpretation of $\frac{1}{\lambda}$ is a reciprocal of the rate λ in Poisson. As an example, if you have 4 failures happen per hour you would expect to get one failure every 1/4 of an hour. With this moment generating function we can genererate $E[X^n]$ for any number n $\in$ N.

Example

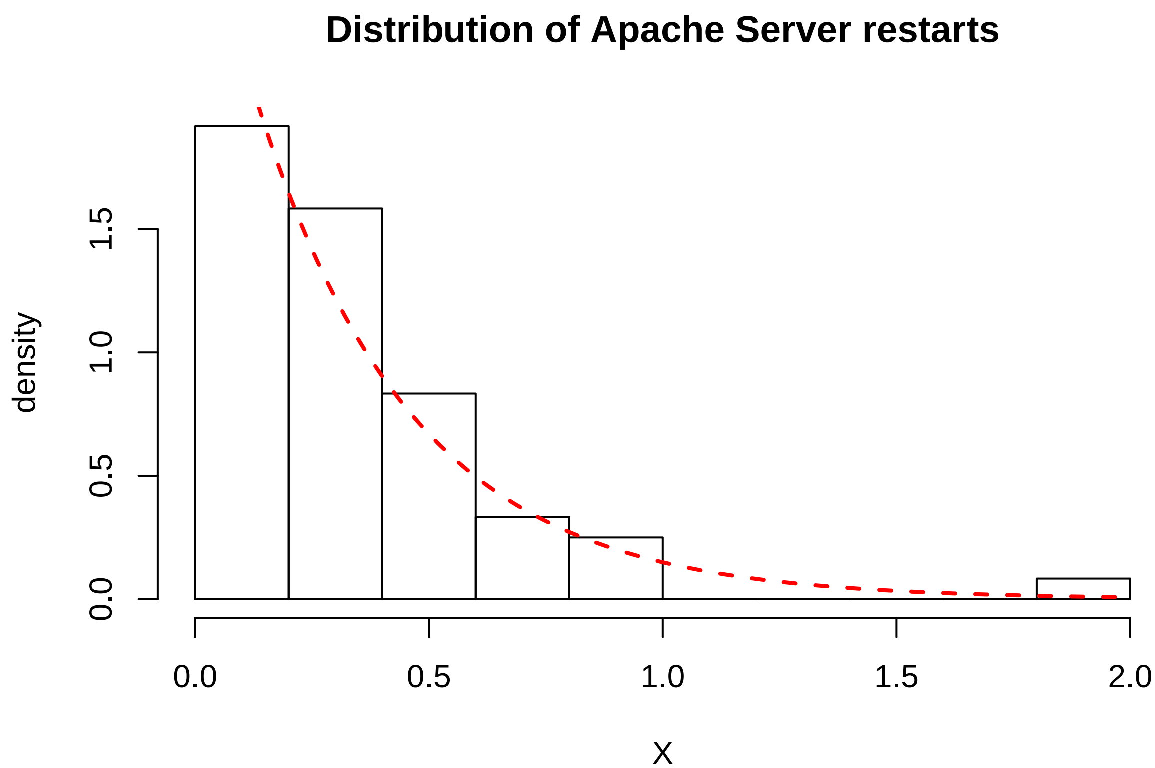
The Exponential Distribution used as a probability model

As a software developer I would like to understand certain behaviors about the Apache server that runs our application for Network Monitoring and Information. The Apache server is said to automatically recycle processes after 10,000 requests to the server according to apache.org. Looking at our server logs I was able to gather data and do a few calculations. From this information I have found we are getting on avearge 500 requests per day over the span of 60 days. This means that on average the server will restart 3 times every 60 days. This will give us a lambda of 3 since there is a restart 3 times per 60 days or a restart every 20 days. I am also making the assumption that these numbers stay roughly consistant which in the case of web applications there can certainly be spikes and dips depeding on the day. I would like to know the probability of the next restart is less than 15 days. Using an exponetial distribution would be an excelant choice for this kind of analysis since the number of restarts in 60 days are not many and we are wanting to know the number of occurances between a time interavel (60 days).

By hand we could solve this using the CDF:

$$P(T < 15days) = 1 - e^{-(3*.25)} = .5276$$

or 52.76% Below is a graph of the histogram of the data with an overlaid probability density function.



Code Appendix

```
#Plot for the Probability Desity Function of the Exponential Distribution
curve(dexp(x, rate=.5),
      xlab="X",
      ylab="density",
      main="Probability Desity Function of the Exponential Distribution",
      col="darkred", lwd=2,
      xlim=c(0,10))
curve(dexp(x, rate=1),
      n=1000,
      add=T,
      col="gold", lwd=2,
      xlim=c(0,10))
```

```r
curve(dexp(x, rate=1.5),
      n=1000,
      add=T,
      col="blue", lwd=2,
      xlim=c(0,10))

legend("topright", legend=c("Rate = .5", "Rate = 1", "Rate = 1.5"),
       col=c("darkred", "gold", "blue"), lty=1, lwd=2)

#Example: Graph of the histogram of the data with an overlaid probability density function
data <- c(432, 530, 454, 434, 479, 487, 510, 583, 597, 470, 468, 404,
          494, 519, 513, 578, 569, 575, 555, 554, 413, 567, 458, 547,
          545, 486, 549, 508, 546, 485, 511, 497, 528, 598, 505, 579,
          506, 565, 405, 472, 572, 557, 495, 529, 535, 582, 480, 568,
          409, 574, 406, 521, 539, 449, 469, 431, 586, 594, 591, 561)
z <- rexp(60,rate = 3)
hist(z, prob = TRUE,
      main="Distribution of Apache Server restarts",
      xlab="X",
      ylab="density")
curve(dexp(x, rate = 3), col = 2, lty = 2, lwd = 2, add = TRUE)
```