

14.2: The Exponential Distribution

Basic Theory

The Memoryless Property

Recall that in the basic model of the Poisson process, we have “points” that occur randomly in time. The sequence of inter-arrival times is $\mathbf{X} = (X_1, X_2, \dots)$. The strong renewal assumption states that at each arrival time and at each fixed time, the process must probabilistically restart, independent of the past. The first part of that assumption implies that \mathbf{X} is a sequence of independent, identically distributed variables. The second part of the assumption implies that if the first arrival has not occurred by time s , then the time remaining until the arrival occurs must have the same distribution as the first arrival time itself. This is known as the *memoryless property* and can be stated in terms of a general random variable as follows:

Suppose that X takes values in $[0, \infty)$. Then X has the *memoryless property* if the conditional distribution of $X - s$ given $X > s$ is the same as the distribution of X for every $s \in [0, \infty)$. Equivalently,

$$\mathbb{P}(X > t + s \mid X > s) = \mathbb{P}(X > t), \quad s, t \in [0, \infty) \quad (14.2.1)$$

The memoryless property determines the distribution of X up to a positive parameter, as we will see now.

Distribution functions

Suppose that X takes values in $[0, \infty)$ and satisfies the memoryless property.

X has a continuous distribution and there exists $r \in (0, \infty)$ such that the distribution function F of X is

$$F(t) = 1 - e^{-rt}, \quad t \in [0, \infty) \quad (14.2.2)$$

Proof

Let $F^c = 1 - F$ denote the right-tail distribution function of X (also known as the *reliability function*), so that $F^c(t) = \mathbb{P}(X > t)$ for $t \geq 0$. From the definition of conditional probability, the memoryless property is equivalent to the *law of exponents*:

$$F^c(t + s) = F^c(s)F^c(t), \quad s, t \in [0, \infty) \quad (14.2.3)$$

Let $a = F^c(1)$. Implicit in the memoryless property is $\mathbb{P}(X > t) > 0$ for $t \in [0, \infty)$, so $a > 0$. If $n \in \mathbb{N}_+$ then

$$F^c(n) = F^c\left(\sum_{i=1}^n 1\right) = \prod_{i=1}^n F^c(1) = [F^c(1)]^n = a^n \quad (14.2.4)$$

Next, if $n \in \mathbb{N}_+$ then

$$a = F^c(1) = F^c\left(\frac{n}{n}\right) = F^c\left(\sum_{i=1}^n \frac{1}{n}\right) = \prod_{i=1}^n F^c\left(\frac{1}{n}\right) = \left[F^c\left(\frac{1}{n}\right)\right]^n \quad (14.2.5)$$

so $F^c\left(\frac{1}{n}\right) = a^{1/n}$. Now suppose that $m \in \mathbb{N}$ and $n \in \mathbb{N}_+$. Then

$$F^c\left(\frac{m}{n}\right) = F^c\left(\sum_{i=1}^m \frac{1}{n}\right) = \prod_{i=1}^m F^c\left(\frac{1}{n}\right) = \left[F^c\left(\frac{1}{n}\right)\right]^m = a^{m/n} \quad (14.2.6)$$

Thus we have $F^c(q) = a^q$ for rational $q \in [0, \infty)$. For $t \in [0, \infty)$, there exists a sequence of rational numbers (q_1, q_2, \dots) with $q_n \downarrow t$ as $n \uparrow \infty$. We have $F^c(q_n) = a^{q_n}$ for each $n \in \mathbb{N}_+$. But F^c is continuous from the right, so taking limits gives $a^t = F^c(t)$. Now let $r = -\ln(a)$. Then $F^c(t) = e^{-rt}$ for $t \in [0, \infty)$.

The probability density function of X is

$$f(t) = r e^{-rt}, \quad t \in [0, \infty) \quad (14.2.7)$$

1. f is decreasing on $[0, \infty)$.
2. f is concave upward on $[0, \infty)$.
3. $f(t) \rightarrow 0$ as $t \rightarrow \infty$.

Proof

This follows since $f = F'$. The properties in parts (a)–(c) are simple.

A random variable with the [distribution function](#) above or equivalently the [probability density function](#) in the last theorem is said to have the *exponential distribution* with *rate parameter* r . The reciprocal $\frac{1}{r}$ is known as the *scale parameter* (as will be justified [below](#)). Note that the mode of the distribution is 0, regardless of the parameter r , not very helpful as a measure of center.

In the gamma experiment, set $n = 1$ so that the simulated random variable has an exponential distribution. Vary r with the scroll bar and watch how the shape of the probability density function changes. For selected values of r , run the experiment 1000 times and compare the empirical density function to the probability density function.

The quantile function of X is

$$F^{-1}(p) = \frac{-\ln(1-p)}{r}, \quad p \in [0, 1) \quad (14.2.8)$$

1. The median of X is $\frac{1}{r} \ln(2) \approx 0.6931 \frac{1}{r}$
2. The first quartile of X is $\frac{1}{r} [\ln(4) - \ln(3)] \approx 0.2877 \frac{1}{r}$
3. The third quartile X is $\frac{1}{r} \ln(4) \approx 1.3863 \frac{1}{r}$
4. The interquartile range is $\frac{1}{r} \ln(3) \approx 1.0986 \frac{1}{r}$

Proof

The formula for F^{-1} follows easily from solving $p = F^{-1}(t)$ for t in terms of p .

In the special distribution calculator, select the exponential distribution. Vary the scale parameter (which is $1/r$) and note the shape of the distribution/quantile function. For selected values of the parameter, compute a few values of the distribution function and the quantile function.

Returning to the Poisson model, we have our first formal definition:

A process of random points in time is a Poisson process with rate $r \in (0, \infty)$ if and only the interarrival times are independent, and each has the exponential distribution with rate r .

Constant Failure Rate

Suppose now that X has a continuous distribution on $[0, \infty)$ and is interpreted as the lifetime of a device. If F denotes the distribution function of X , then $F^c = 1 - F$ is the reliability function of X . If f denotes the probability density function of X then the failure rate function h is given by

$$h(t) = \frac{f(t)}{F^c(t)}, \quad t \in [0, \infty) \quad (14.2.9)$$

If X has the exponential distribution with rate $r > 0$, then from the results above, the reliability function is $F^c(t) = e^{-rt}$ and the probability density function is $f(t) = re^{-rt}$, so trivially X has constant rate r . The converse is also true.

If X has constant failure rate $r > 0$ then X has the exponential distribution with parameter r .

Proof

Recall that in general, the distribution of a lifetime variable X is determined by the failure rate function h . Specifically, if $F^c = 1 - F$ denotes the reliability function, then $(F^c)' = -f$, so $-h = (F^c)' / F^c$. Integrating and then taking exponentials gives

$$F^c(t) = \exp\left(-\int_0^t h(s) ds\right), \quad t \in [0, \infty) \quad (14.2.10)$$

In particular, if $h(t) = r$ for $t \in [0, \infty)$, then $F^c(t) = e^{-rt}$ for $t \in [0, \infty)$.

The memoryless and constant failure rate properties are the most famous characterizations of the exponential distribution, but are by no means the only ones. Indeed, entire books have been written on characterizations of this distribution.

Moments

Suppose again that X has the exponential distribution with rate parameter $r > 0$. Naturally, we want to know the mean, variance, and various other moments of X .

If $n \in \mathbb{N}$ then $\mathbb{E}(X^n) = n! / r^n$.

Proof

By the change of variables theorem for expected value,

$$\mathbb{E}(X^n) = \int_0^\infty t^n r e^{-rt} dt \quad (14.2.11)$$

Integrating by parts gives $\mathbb{E}(X^n) = \frac{n}{r} \mathbb{E}(X^{n-1})$ for $n \in \mathbb{N}^+$. Of course $\mathbb{E}(X^0) = 1$ so the result now follows by induction.

More generally, $\mathbb{E}(X^a) = \Gamma(a+1) / r^a$ for every $a \in [0, \infty)$, where Γ is the gamma function.

In particular.

1. $\mathbb{E}(X) = \frac{1}{r}$
2. $\text{var}(X) = \frac{1}{r^2}$
3. $\text{skew}(X) = 2$
4. $\text{kurt}(X) = 9$

In the context of the Poisson process, the parameter r is known as the *rate* of the process. On average, there are $1/r$ time units between arrivals, so the arrivals come at an average rate of r per unit time. The Poisson process is completely determined by the sequence of inter-arrival times, and hence is completely determined by the rate r .

Note also that the mean and standard deviation are equal for an exponential distribution, and that the median is always smaller than the mean. Recall also that skewness and kurtosis are standardized measures, and so do not depend on the parameter r (which is the reciprocal of the scale parameter).

The moment generating function of X is

$$M(s) = \mathbb{E}(e^{sX}) = \frac{r}{r-s}, \quad s \in (-\infty, r) \quad (14.2.12)$$

Proof

By the change of variables theorem

$$M(s) = \int_0^{\infty} e^{st} r e^{-rt} dt = \int_0^{\infty} r e^{(s-r)t} dt \quad (14.2.13)$$

The integral evaluates to $\frac{r}{r-s}$ if $s < r$ and to ∞ if $s \geq r$.

In the gamma experiment, set $n = 1$ so that the simulated random variable has an exponential distribution. Vary r with the scroll bar and watch how the mean \pm standard deviation bar changes. For various values of r , run the experiment 1000 times and compare the empirical mean and standard deviation to the distribution mean and standard deviation, respectively.

Additional Properties

The exponential distribution has a number of interesting and important mathematical properties. First, and not surprisingly, it's a member of the general exponential family.

Suppose that X has the exponential distribution with rate parameter $r \in (0, \infty)$. Then X has a one parameter general exponential distribution, with natural parameter $-r$ and natural statistic X .

Proof

This follows directly from the form of the PDF, $f(x) = r e^{-rx}$ for $x \in [0, \infty)$, and the definition of the general exponential family.

The Scaling Property

As suggested earlier, the exponential distribution is a scale family, and $1/r$ is the scale parameter.

Suppose that X has the exponential distribution with rate parameter $r > 0$ and that $c > 0$. Then cX has the exponential distribution with rate parameter r/c .

Proof

For $t \geq 0$, $\mathbb{P}(cX > t) = \mathbb{P}(X > t/c) = e^{-r(t/c)} = e^{-(r/c)t}$.

Recall that multiplying a random variable by a positive constant frequently corresponds to a change of units (minutes into hours for a lifetime variable, for example). Thus, the exponential distribution is preserved under such changes of units. In the context of the Poisson process, this has to be the case, since the memoryless property, which led to the exponential distribution in the first place, clearly does not depend on the time units.

In fact, the exponential distribution with rate parameter 1 is referred to as the *standard exponential distribution*. From the previous result, if Z has the standard exponential distribution and $r > 0$, then $X = \frac{1}{r}Z$ has the exponential distribution with rate parameter r . Conversely, if X has the exponential distribution with rate $r > 0$ then $Z = rX$ has the standard exponential distribution.

Similarly, the Poisson process with rate parameter 1 is referred to as the *standard Poisson process*. If Z_i is the i th inter-arrival time for the standard Poisson process for $i \in \mathbb{N}_+$, then letting $X_i = \frac{1}{r}Z_i$ for $i \in \mathbb{N}_+$ gives the inter-arrival times for the Poisson process with rate r . Conversely if X_i is the i th inter-arrival time of the Poisson process with rate $r > 0$ for $i \in \mathbb{N}_+$, then $Z_i = rX_i$ for $i \in \mathbb{N}_+$ gives the inter-arrival times for the standard Poisson process.

Relation to the Geometric Distribution

In many respects, the geometric distribution is a discrete version of the exponential distribution. In particular, recall that the geometric distribution on \mathbb{N}_+ is the only distribution on \mathbb{N}_+ with the memoryless and constant rate properties. So it is not surprising that the two distributions are also connected through various transformations and limits.

Suppose that X has the exponential distribution with rate parameter $r > 0$. Then

1. $\lfloor X \rfloor$ has the geometric distributions on \mathbb{N} with success parameter $1 - e^{-r}$.
2. $\lceil X \rceil$ has the geometric distributions on \mathbb{N}_+ with success parameter $1 - e^{-r}$.

Proof

1. For $n \in \mathbb{N}$ note that $\mathbb{P}(\lfloor X \rfloor = n) = \mathbb{P}(n \leq X < n+1) = F(n+1) - F(n)$. Substituting into the distribution function and simplifying gives $\mathbb{P}(\lfloor X \rfloor = n) = (e^{-r})^n (1 - e^{-r})$.
2. For $n \in \mathbb{N}_+$ note that $\mathbb{P}(\lceil X \rceil = n) = \mathbb{P}(n-1 < X \leq n) = F(n) - F(n-1)$. Substituting into the distribution function and simplifying gives $\mathbb{P}(\lceil X \rceil = n) = (e^{-r})^{n-1} (1 - e^{-r})$.

The following connection between the two distributions is interesting by itself, but will also be very important in the section on splitting Poisson processes. In words, a random, geometrically distributed sum of independent, identically distributed exponential variables is itself exponential.

Suppose that $\mathbf{X} = (X_1, X_2, \dots)$ is a sequence of independent variables, each with the exponential distribution with rate r . Suppose that U has the geometric distribution on \mathbb{N}_+ with success parameter p and is independent of \mathbf{X} . Then $Y = \sum_{i=1}^U X_i$ has the exponential distribution with rate rp .

Proof

Recall that the moment generating function of Y is $P \circ M$ where M is the common moment generating function of the terms in the sum, and P is the probability generating function of the number of terms U . But $M(s) = r/(r-s)$ for $s < r$ and $P(s) = ps/[1 - (1-p)s]$ for $s < 1/(1-p)$. Thus,

$$(P \circ M)(s) = \frac{pr/(r-s)}{1 - (1-p)r/(r-s)} = \frac{pr}{pr-s}, \quad s < pr \quad (14.2.14)$$

It follows that Y has the exponential distribution with parameter pr

The next result explores the connection between the Bernoulli trials process and the Poisson process that was begun in the Introduction.

For $n \in \mathbb{N}_+$, suppose that U_n has the geometric distribution on \mathbb{N}_+ with success parameter p_n , where $np_n \rightarrow r > 0$ as $n \rightarrow \infty$. Then the distribution of U_n/n converges to the exponential distribution with parameter r as $n \rightarrow \infty$.

Proof

Let F_n denote the CDF of U_n/n . Then for $x \in [0, \infty)$

$$F_n(x) = \mathbb{P}\left(\frac{U_n}{n} \leq x\right) = \mathbb{P}(U_n \leq nx) = \mathbb{P}(U_n \leq \lfloor nx \rfloor) = 1 - (1 - p_n)^{\lfloor nx \rfloor} \quad (14.2.15)$$

But by a famous limit from calculus, $(1 - p_n)^n = \left(1 - \frac{np_n}{n}\right)^n \rightarrow e^{-r}$ as $n \rightarrow \infty$, and hence $(1 - p_n)^{nx} \rightarrow e^{-rx}$ as $n \rightarrow \infty$. But by definition, $\lfloor nx \rfloor \leq nx < \lfloor nx \rfloor + 1$ or equivalently, $nx - 1 < \lfloor nx \rfloor \leq nx$ so it follows that $(1 - p_n)^{\lfloor nx \rfloor} \rightarrow e^{-rx}$ as $n \rightarrow \infty$. Hence $F_n(x) \rightarrow 1 - e^{-rx}$ as $n \rightarrow \infty$, which is the CDF of the exponential distribution.

To understand this result more clearly, suppose that we have a sequence of Bernoulli trials processes. In process n , we run the trials at a rate of n per unit time, with probability of success p_n . Thus, the actual time of the first success in process n is U_n/n . The last result shows that if $np_n \rightarrow r > 0$ as $n \rightarrow \infty$, then the sequence of Bernoulli trials processes converges to the Poisson process with rate parameter r as $n \rightarrow \infty$. We will return to this point in subsequent sections.

Orderings and Order Statistics

Suppose that X and Y have exponential distributions with parameters a and b , respectively, and are independent. Then

$$\mathbb{P}(X < Y) = \frac{a}{a+b} \quad (14.2.16)$$

Proof

This result can be proved in a straightforward way by integrating the joint PDF of (X, Y) over $\{(x, y) : 0 < x < y < \infty\}$. A more elegant proof uses conditioning and the [moment generating function](#) above:

$$\mathbb{P}(Y > X) = \mathbb{E}[\mathbb{P}(Y > X | X)] = \mathbb{E}(e^{-bX}) = \frac{a}{a+b} \quad (14.2.17)$$

The following theorem gives an important random version of the memoryless property.

Suppose that X and Y are independent variables taking values in $[0, \infty)$ and that Y has the exponential distribution with rate parameter $r > 0$. Then X and $Y - X$ are conditionally independent given $X < Y$, and the conditional distribution of $Y - X$ is also exponential with parameter r .

Proof

Suppose that $A \subseteq [0, \infty)$ (measurable of course) and $t \geq 0$. Then

$$\mathbb{P}(X \in A, Y - X \geq t | X < Y) = \frac{\mathbb{P}(X \in A, Y - X \geq t)}{\mathbb{P}(X < Y)} \quad (14.2.18)$$

But conditioning on X we can write the numerator as

$$\mathbb{P}(X \in A, Y - X > t) = \mathbb{E}[\mathbb{P}(X \in A, Y - X > t | X)] = \mathbb{E}[\mathbb{P}(Y > X + t | X), X \in A] = \mathbb{E}[e^{-r(t+X)}, X \in A] = e^{-rt} \mathbb{E}(e^{-rX}, X \in A) \quad (14.2.19)$$

Similarly, conditioning on X gives $\mathbb{P}(X < Y) = \mathbb{E}(e^{-rX})$. Thus

$$\mathbb{P}(X \in A, Y - X > t | X < Y) = e^{-rt} \frac{\mathbb{E}(e^{-rX}, X \in A)}{\mathbb{E}(e^{-rX})} \quad (14.2.20)$$

Letting $A = [0, \infty)$ we have $\mathbb{P}(Y > t) = e^{-rt}$ so given $X < Y$, the variable $Y - X$ has the exponential distribution with parameter r . Letting $t = 0$, we see that given $X < Y$, variable X has the distribution

$$A \mapsto \frac{\mathbb{E}(e^{-rX}, X \in A)}{\mathbb{E}(e^{-rX})} \quad (14.2.21)$$

Finally, because of the factoring, X and $Y - X$ are conditionally independent given $X < Y$.

For our next discussion, suppose that $\mathbf{X} = (X_1, X_2, \dots, X_n)$ is a sequence of independent random variables, and that X_i has the exponential distribution with rate parameter $r_i > 0$ for each $i \in \{1, 2, \dots, n\}$.

Let $U = \min\{X_1, X_2, \dots, X_n\}$. Then U has the exponential distribution with parameter $\sum_{i=1}^n r_i$.

Proof

Recall that in general, $\{U > t\} = \{X_1 > t, X_2 > t, \dots, X_n > t\}$ and therefore by independence, $F^c(t) = F_1^c(t)F_2^c(t) \cdots F_n^c(t)$ for $t \geq 0$, where F^c is the reliability function of U and F_i^c is the reliability function of X_i for each i . When X_i has the exponential distribution with rate r_i for each i , we have $F^c(t) = \exp[-(\sum_{i=1}^n r_i)t]$ for $t \geq 0$.

In the context of reliability, if a series system has independent components, each with an exponentially distributed lifetime, then the lifetime of the system is also exponentially distributed, and the failure rate of the system is the sum of the component failure rates. In the context of random processes, if we have n independent Poisson processes, then the new process obtained by combining the random points in time is also Poisson, and the rate of the new process is the sum of the rates of the individual processes (we will return to this point later).

Let $V = \max\{X_1, X_2, \dots, X_n\}$. Then V has distribution function F given by

$$F(t) = \prod_{i=1}^n (1 - e^{-r_i t}), \quad t \in [0, \infty) \quad (14.2.22)$$

Proof

Recall that in general, $\{V \leq t\} = \{X_1 \leq t, X_2 \leq t, \dots, X_n \leq t\}$ and therefore by independence, $F(t) = F_1(t)F_2(t) \cdots F_n(t)$ for $t \geq 0$, where F is the distribution function of V and F_i is the distribution function of X_i for each i .

Consider the special case where $r_i = r \in (0, \infty)$ for each $i \in \mathbb{N}_+$. In statistical terms, \mathbf{X} is a random sample of size n from the exponential distribution with parameter r . From the last couple of theorems, the minimum U has the exponential distribution with rate nr while the maximum V has distribution function $F(t) = (1 - e^{-rt})^n$ for $t \in [0, \infty)$. Recall that U and V are the first and last order statistics, respectively.

In the order statistic experiment, select the exponential distribution.

1. Set $k = 1$ (this gives the minimum U). Vary n with the scroll bar and note the shape of the probability density function. For selected values of n , run the simulation 1000 times and compare the empirical density function to the true probability density function.
2. Vary n with the scroll bar, set $k = n$ each time (this gives the maximum V), and note the shape of the probability density function. For selected values of n , run the simulation 1000 times and compare the empirical density function to the true probability density function.

Curiously, the distribution of the maximum of independent, identically distributed exponential variables is also the distribution of the sum of independent exponential variables, with rates that grow linearly with the index.

Suppose that $r_i = ir$ for each $i \in \{1, 2, \dots, n\}$ where $r \in (0, \infty)$. Then $Y = \sum_{i=1}^n X_i$ has distribution function F given by

$$F(t) = (1 - e^{-rt})^n, \quad t \in [0, \infty) \quad (14.2.23)$$

Proof

By assumption, X_k has PDF f_k given by $f_k(t) = kre^{-krt}$ for $t \in [0, \infty)$. We want to show that $Y_n = \sum_{i=1}^n X_i$ has PDF g_n given by

$$g_n(t) = nre^{-rt}(1 - e^{-rt})^{n-1}, \quad t \in [0, \infty) \quad (14.2.24)$$

The PDF of a sum of independent variables is the convolution of the individual PDFs, so we want to show that

$$f_1 * f_2 * \cdots * f_n = g_n, \quad n \in \mathbb{N}_+ \quad (14.2.25)$$

The proof is by induction on n . Trivially $f_1 = g_1$, so suppose the result holds for a given $n \in \mathbb{N}_+$. Then

$$\begin{aligned} g_n * f_{n+1}(t) &= \int_0^t g_n(s)f_{n+1}(t-s)ds = \int_0^t nre^{-rs}(1 - e^{-rs})^{n-1}(n+1)re^{-(n+1)(t-s)}ds \\ &= r(n+1)e^{-r(n+1)t} \int_0^t n(1 - e^{-rs})^{n-1}re^{rs}ds \end{aligned}$$

Now substitute $u = e^{rs}$ so that $du = re^{rs}ds$ or equivalently $rd s = du/u$. After some algebra,

$$\begin{aligned} g_n * f_{n+1}(t) &= r(n+1)e^{-r(n+1)t} \int_1^{e^{rt}} n(u-1)^{n-1}du \\ &= r(n+1)e^{-r(n+1)t}(e^{rt} - 1)^n = r(n+1)e^{-rt}(1 - e^{-rt})^n = g_{n+1}(t) \end{aligned}$$

This result has an application to the *Yule process*, named for George Yule. The Yule process, which has some parallels with the Poisson process, is studied in the chapter on Markov processes. We can now generalize the [order probability](#) above:

For $i \in \{1, 2, \dots, n\}$,

$$\mathbb{P}(X_i < X_j \text{ for all } j \neq i) = \frac{r_i}{\sum_{j=1}^n r_j} \quad (14.2.26)$$

Proof

First, note that $X_i < X_j$ for all $i \neq j$ if and only if $X_i < \min\{X_j : j \neq i\}$. But the minimum on the right is independent of X_i and, by [result on minimums](#) above, has the exponential distribution with parameter $\sum_{j \neq i} r_j$. The result now follows from [order probability](#) for two events above.

Suppose that for each i , X_i is the time until an event of interest occurs (the arrival of a customer, the failure of a device, etc.) and that these times are independent and exponentially distributed. Then the *first* time U that one of the events occurs is also exponentially distributed, and the probability that the first event to occur is event i is proportional to the rate r_i .

The probability of a total ordering is

$$\mathbb{P}(X_1 < X_2 < \cdots < X_n) = \prod_{i=1}^n \frac{r_i}{\sum_{j=i}^n r_j} \quad (14.2.27)$$

Proof

Let $A = \{X_1 < X_j \text{ for all } j \in \{2, 3, \dots, n\}\}$. then

$$\mathbb{P}(X_1 < X_2 < \cdots < X_n) = \mathbb{P}(A, X_2 < X_3 < \cdots < X_n) = \mathbb{P}(A)\mathbb{P}(X_2 < X_3 < \cdots < X_n | A) \quad (14.2.28)$$

But $\mathbb{P}(A) = \frac{r_1}{\sum_{i=1}^n r_i}$ from the previous result, and $\{X_2 < X_3 < \cdots < X_n\}$ is independent of A . Thus we have

$$\mathbb{P}(X_1 < X_2 < \cdots < X_n) = \frac{r_1}{\sum_{i=1}^n r_i} \mathbb{P}(X_2 < X_3 < \cdots < X_n) \quad (14.2.29)$$

so the result follows by induction.

Of course, the probabilities of other orderings can be computed by permuting the parameters appropriately in the formula on the right.

The result on [minimums](#) and the [order probability](#) result above are very important in the theory of continuous-time Markov chains. But for that application and others, it's convenient to extend the exponential distribution to two degenerate cases: point mass at 0 and point mass at ∞ (so the first is the distribution of a random variable that takes the value 0 with probability 1, and the second the distribution of a random variable that takes the value ∞ with probability 1). In terms of the rate parameter r and the distribution function F , point mass at 0 corresponds to $r = \infty$ so that $F(t) = 1$ for $0 < t < \infty$. Point mass at ∞ corresponds to $r = 0$ so that $F(t) = 0$ for $0 < t < \infty$. The memoryless property, as expressed in terms of the reliability function F^c , still holds for these degenerate cases on $(0, \infty)$:

$$F^c(s)F^c(t) = F^c(s+t), \quad s, t \in (0, \infty) \quad (14.2.30)$$

We also need to extend some of results above for a finite number of variables to a countably infinite number of variables. So for the remainder of this discussion, suppose that $\{X_i : i \in I\}$ is a countable collection of independent random variables, and that X_i has the exponential distribution with parameter $r_i \in (0, \infty)$ for each $i \in I$.

Let $U = \inf\{X_i : i \in I\}$. Then U has the exponential distribution with parameter $\sum_{i \in I} r_i$

Proof

The proof is almost the same as the one [above](#) for a finite collection. Note that $\{U \geq t\} = \{X_i \geq t \text{ for all } i \in I\}$ and so

$$\mathbb{P}(U \geq t) = \prod_{i \in I} \mathbb{P}(X_i \geq t) = \prod_{i \in I} e^{-r_i t} = \exp\left[-\left(\sum_{i \in I} r_i\right)t\right] \quad (14.2.31)$$

If $\sum_{i \in I} r_i < \infty$ then U has a proper exponential distribution with the sum as the parameter. If $\sum_{i \in I} r_i = \infty$ then $\mathbb{P}(U \geq t) = 0$ for all $t \in (0, \infty)$ so $\mathbb{P}(U = 0) = 1$.

For $i \in \mathbb{N}_+$,

$$\mathbb{P}(X_i < X_j \text{ for all } j \in I - \{i\}) = \frac{r_i}{\sum_{j \in I} r_j} \quad (14.2.32)$$

Proof

First note that since the variables have continuous distributions and I is countable,

$$\mathbb{P}(X_i < X_j \text{ for all } j \in I - \{i\}) = \mathbb{P}(X_i \leq X_j \text{ for all } j \in I - \{i\}) \quad (14.2.33)$$

Next note that $X_i \leq X_j$ for all $j \in I - \{i\}$ if and only if $X_i \leq U_i$ where $U_i = \inf\{X_j : j \in I - \{i\}\}$. But U_i is independent of X_i and, by [previous result](#), has the exponential distribution with parameter $s_i = \sum_{j \in I - \{i\}} r_j$. If $s_i = \infty$, then U_i is 0 with probability 1, and so $\mathbb{P}(X_i \leq U_i) = 0 = r_i/s_i$. If $s_i < \infty$, then X_i and U_i have proper exponential distributions, and so the result now follows from [order probability](#) for two variables above.

We need one last result in this setting: a condition that ensures that the sum of an infinite collection of exponential variables is finite with probability one.

Let $Y = \sum_{i \in I} X_i$ and $\mu = \sum_{i \in I} 1/r_i$. Then $\mu = \mathbb{E}(Y)$ and $\mathbb{P}(Y < \infty) = 1$ if and only if $\mu < \infty$.

Proof

The result is trivial if I is finite, so assume that $I = \mathbb{N}_+$. Recall that $\mathbb{E}(X_i) = 1/r_i$ and hence $\mu = \mathbb{E}(Y)$. Trivially if $\mu < \infty$ then $\mathbb{P}(Y < \infty) = 1$. Conversely, suppose that $\mathbb{P}(Y < \infty) = 1$. Then $\mathbb{P}(e^{-Y} > 0) = 1$ and hence $\mathbb{E}(e^{-Y}) > 0$. Using independence and the [moment generating function](#) above,

$$\mathbb{E}(e^{-Y}) = \mathbb{E}\left(\prod_{i=1}^{\infty} e^{-X_i}\right) = \prod_{i=1}^{\infty} \mathbb{E}(e^{-X_i}) = \prod_{i=1}^{\infty} \frac{r_i}{r_i + 1} > 0 \quad (14.2.34)$$

Next recall that if $p_i \in (0, 1)$ for $i \in \mathbb{N}_+$ then

$$\prod_{i=1}^{\infty} p_i > 0 \text{ if and only if } \sum_{i=1}^{\infty} (1 - p_i) < \infty \quad (14.2.35)$$

Hence it follows that

$$\sum_{i=1}^{\infty} \left(1 - \frac{r_i}{r_i + 1}\right) = \sum_{i=1}^{\infty} \frac{1}{r_i + 1} < \infty \quad (14.2.36)$$

In particular, this means that $1/(r_i + 1) \rightarrow 0$ as $i \rightarrow \infty$ and hence $r_i \rightarrow \infty$ as $i \rightarrow \infty$. But then

$$\frac{1/(r_i + 1)}{1/r_i} = \frac{r_i}{r_i + 1} \rightarrow 1 \text{ as } i \rightarrow \infty \quad (14.2.37)$$

By the comparison test for infinite series, it follows that

$$\mu = \sum_{i=1}^{\infty} \frac{1}{r_i} < \infty \quad (14.2.38)$$

Computational Exercises

Show directly that the exponential probability density function is a valid probability density function.

Solution

Clearly $f(t) = re^{-rt} > 0$ for $t \in [0, \infty)$. Simple integration that

$$\int_0^{\infty} re^{-rt} dt = 1 \quad (14.2.39)$$

Suppose that the length of a telephone call (in minutes) is exponentially distributed with rate parameter $r = 0.2$. Find each of the following:

1. The probability that the call lasts between 2 and 7 minutes.
2. The median, the first and third quartiles, and the interquartile range of the call length.

Answer

Let X denote the call length.

1. $\mathbb{P}(2 < X < 7) = 0.4237$
2. $q_1 = 1.4384$ $q_2 = 3.4657$ $q_3 = 6.9315$ $q_3 - q_1 = 5.4931$

Suppose that the lifetime of a certain electronic component (in hours) is exponentially distributed with rate parameter $r = 0.001$. Find each of the following:

1. The probability that the component lasts at least 2000 hours.
2. The median, the first and third quartiles, and the interquartile range of the lifetime.

Answer

Let T denote the lifetime

1. $\mathbb{P}(T \geq 2000) = 0.1353$
2. $q_1 = 287.682$ $q_2 = 693.147$ $q_3 = 1386.294$ $q_3 - q_1 = 1098.612$

Suppose that the time between requests to a web server (in seconds) is exponentially distributed with rate parameter $r = 2$. Find each of the following:

1. The mean and standard deviation of the time between requests.
2. The probability that the time between requests is less than 0.5 seconds.
3. The median, the first and third quartiles, and the interquartile range of the time between requests.

Answer

Let T denote the time between requests.

1. $\mathbb{E}(T) = 0.5$, $\text{sd}(T) = 0.5$
2. $\mathbb{P}(T < 0.5) = 0.6321$
3. $q_1 = 0.1438$ $q_2 = 0.3466$ $q_3 = 0.6931$ $q_3 - q_1 = 0.5493$

Suppose that the lifetime X of a fuse (in 100 hour units) is exponentially distributed with $\mathbb{P}(X > 10) = 0.8$. Find each of the following:

1. The rate parameter.
2. The mean and standard deviation.
3. The median, the first and third quartiles, and the interquartile range of the lifetime.

Answer

Let X denote the lifetime.

1. $r = 0.02231$
2. $\mathbb{E}(X) = 44.814$, $\text{sd}(X) = 44.814$
3. $q_1 = 12.8922$ $q_2 = 31.0628$ $q_3 = 62.1257$ $q_3 - q_1 = 49.2334$

The position X of the first defect on a digital tape (in cm) has the exponential distribution with mean 100. Find each of the following:

1. The rate parameter.
2. The probability that $X < 200$ given $X > 150$.
3. The standard deviation.
4. The median, the first and third quartiles, and the interquartile range of the position.

Answer

Let X denote the position of the first defect.

1. $r = 0.01$
2. $\mathbb{P}(X < 200 \mid X > 150) = 0.3935$
3. $\text{sd}(X) = 100$
4. $q_1 = 28.7682$ $q_2 = 69.3147$ $q_3 = 138.6294$ $q_3 - q_1 = 109.8612$

Suppose that X , Y , Z are independent, exponentially distributed random variables with respective parameters a , b , $c \in (0, \infty)$. Find the probability of each of the 6 orderings of the variables.

Proof

1. $\mathbb{P}(X < Y < Z) = \frac{a}{a+b+c} \frac{b}{b+c}$
2. $\mathbb{P}(X < Z < Y) = \frac{a}{a+b+c} \frac{c}{b+c}$
3. $\mathbb{P}(Y < X < Z) = \frac{b}{a+b+c} \frac{a}{a+c}$
4. $\mathbb{P}(Y < Z < X) = \frac{b}{a+b+c} \frac{c}{a+c}$
5. $\mathbb{P}(Z < X < Y) = \frac{c}{a+b+c} \frac{a}{a+b}$
6. $\mathbb{P}(Z < Y < X) = \frac{c}{a+b+c} \frac{b}{a+b}$

This page titled [14.2: The Exponential Distribution](#) is shared under a [CC BY 2.0](#) license and was authored, remixed, and/or curated by [Kyle Siegrist \(Random Services\)](#) via [source content](#) that was edited to the style and standards of the LibreTexts platform.