3. The Gamma Distribution

The Probability Density Function

We now know that the interarrival times $(X_1, X_2, \ldots)$ form a sequence of independent random variables, each having the exponential probability density function:

$$f(t) = r\, e^{-r\, t}, \quad t \ge 0$$

The k^{th} arrival time is simply the sum of the first k interarrival times:

$$T_k = \sum_{i=1}^{k} X_i$$

Therefore, the k^{th} arrival time has a continuous distribution and its probability density function is the convolution power of f of order k.

1. Show that the probability density function of the k^{th} arrival time is

$$f^{*k}\left(t\right) = r^k \frac{t^{k-1}}{(k-1)!} e^{-r\, t}, \quad t \ge 0$$

This distribution is the **gamma distribution** with **shape parameter** k and **rate parameter** r. Again, $\frac{1}{r}$ is knows as the scale parameter. A more general version of the gamma distribution, allowing non-integer shape parameters, is studied in the chapter on Special Distributions.

Note that since the arrival times are continuous, the probability of an arrival at any given instant of time is 0. Thus, we can also interpret N_t as the number of arrivals in $(0, t)$.

2. In the gamma experiment, vary r and k with the scroll bars and watch how the shape of the probability density function changes. Now set $r = 2$ and $k = 3$, run the experiment 1000 times with an update frequency of 10, and watch the apparent convergence of the empirical density function to the true density function.

3. Sketch the graph of the probability density function in Exercise 1. Show that the density function at first increases and then decreases, reaching its maximum value at the mode $\frac{k-1}{r}$.

4. Suppose that customers arrive at a service station according to the Poisson model, at a rate of $r = 3$ per hour. Relative to a given starting time, find the probability that the second customer arrives sometime after 1 hour.

5. Defects in a type of wire follow the Poisson model, with rate 1 per 100 meter. Find the probability that the 5^{th} defect is located between 450 and 550 meters.

Moments

The mean, variance, and moment generating function of T_k can be found using basic properties and the corresponding results for the exponential distribution

6. Show that $\mathbb{E}(T_k) = \frac{k}{r}$.

7. Show that $\operatorname{var}(T_k) = \frac{k}{r^2}$.

8. In the gamma experiment, vary r and k with the scroll bars and watch how the size and location of the mean/standard deviation bar changes. Now set $r = 2$ and $k = 3$, run the experiment 1000 times with an update frequency of 10, and watch the apparent convergence of the empirical moments to the true moments.

9. Show that $\mathbb{E}\left(e^{u\,T_k}\right) = \left(\frac{r}{r-u}\right)^k$ for $u < r$.

10. Suppose that requests to a web server follow the Poisson model with rate $r = 5$. Relative to a given starting time, compute the mean and standard deviation of the time of the 10^{th} request.

11. Suppose that Y has a gamma distribution with mean 40 and standard deviation 20. Find the shape parameter k and the rate parameter r.

Sums of Independent Gamma Variables

12. Suppose that V has the gamma distribution with shape parameter j and rate parameter r, that W has the gamma distribution with shape parameter k and rate parameter r, and that V and W are independent. Show that $V + W$ has the gamma distribution with shape parameter $j + k$ and rate parameter r.

a. Give an analytic proof, using moment generating functions.
b. Give an analytic proof, using probability density functions.
c. give a probabilistic proof, based on the Poisson process.

Normal Approximation

13. In the gamma experiment, vary r and k with the scroll bars and watch how the shape of the density function changes. Now set $r = 3$ and $k = 5$ run the experiment 1000 times with an update frequency of 10, and watch the apparent convergence of the empirical density function to the true density function.

Even though you are restricted to small values of k in the applet, note that the probability density function of the k^{th} arrival time becomes more bell shaped as k increases (for r fixed). This is yet another application of the central limit theorem, since the k^{th} arrival time is the sum of k independent, identically distributed random variables (the interarrival times).

14. Use the central limit theorem to show that the distribution of the standardized variable below converges to the standard normal distribution as $k \to \infty$:

$$Z_k = \frac{r\,T_k - k}{\sqrt{k}}$$

15. In the gamma experiment, set $k = 5$ and $r = 2$. Run the experiment 1000 times, updating after every run. Compute and compare the following:

a. $\mathbb{P}(1.5 \le T_5 \le 3)$.
b. The relative frequency of the event $\{1.5 \le T_5 \le 3\}$.
c. The normal approximation to $\mathbb{P}(1.5 \le T_5 \le 3)$.

16. Suppose that accidents at an intersection occur according to the Poisson model, at a rate of 8 per year. Compute the normal approximation to the event that the 10^{th} accident (relative to a given starting time) occurs within 2 years.

Estimating the Rate

In many practical situations, the rate r of the process in unknown and must be estimated based on observing the arrival times.

17. Show that $\mathbb{E}\left(\frac{T_k}{k}\right) = \frac{1}{r}$ and hence $\frac{T_k}{k}$ is an unbiased estimator of $\frac{1}{r}$.

Since the estimator is unbiased, the variance measures the mean square error of the estimator.

18. Show that $\operatorname{var}\left(\frac{T_k}{k}\right) = \frac{1}{k\,r^2}$ and hence $\operatorname{var}\left(\frac{T_k}{k}\right) \to 0$ as $k \to \infty$. This means that $\frac{T_k}{k}$ is a consistent estimator of $\frac{1}{r}$.

Note that

$$\frac{T_k}{k} = \frac{1}{k} \sum_{i=1}^{k} X_i$$

where X_i is the i^{th} interarrival time. Hence our estimator of $\frac{1}{r}$ can be interpreted as the sample mean of the interarrival times. In particular, by the law of large numbers, $\frac{T_k}{k} \to \frac{1}{r}$ as $k \to \infty$ with probability 1. A natural estimator of the rate itself is $\frac{k}{T_k}$. However, this estimator tends to overestimate r.

19. Use Jensen's inequality to show that $\mathbb{E}\left(\frac{k}{T_k}\right) \ge r$.

20. Suppose that requests to a web server follow the Poisson model. Starting at 12:00 noon on a certain day, the requests are logged. The 100^{th} request comes at 12:15. Estimate the rate of the process.