Gamma and Poisson Distribution

The gamma and Poisson distributions are extensions of the exponential distribution. The exponential distribution, as a reminder, is the time until the next event where the probability of that time is

$$p(t) = ae^{-at}$$

A gamma distribution is the time until k events occur. Examples of this would be

- The time until k customers arrive,
- The time until k atoms decay,
- The time until you've been invited to k parties,
- The time until you've been exposed to Corona virus k times,
- etc.

A Poisson distribution is the probability of N events occurring over a time interval of M seconds. For example,

- The number of pieces of mail you receive each day (the sending time is exponential)
- The number of cars through in intersection in one minute
- The number of atoms decaying over a one minute interval.
- The number of customers arriving at a restaurant in one hour

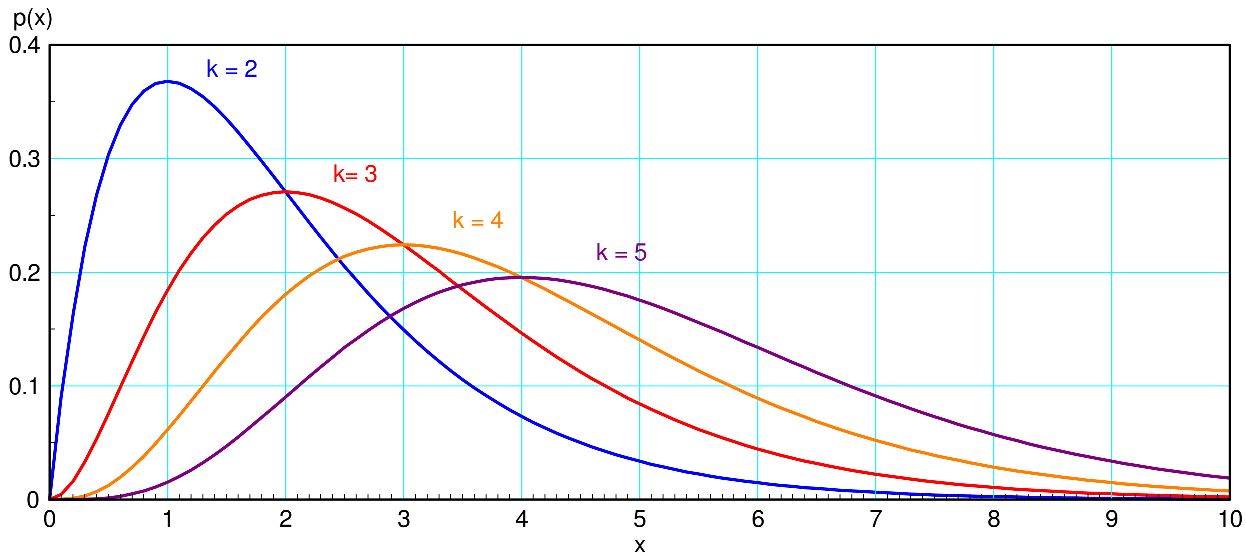
A Poisson distribution can also be used to approximate binomial distributions where n is large.

Gamma Distribution

The moment generating function is an extension of the exponential distribution (time until k events vs. 1 event). Likewise, the mean, variance, moment generating functions are all very similar

	Exponential	Gamma
pdf	$f_x = a\ e^{-ax}$	$f_x = \left(\frac{a^k}{(k-1)!}\right) x^{k-1}\ e^{-ax}$
cdf	$F_x = 1 - e^{-ax}$	complicated
mgf	$\left(\frac{a}{s+a}\right)$	$\left(\frac{a}{s+a}\right)^k$
mean	$\left(\frac{1}{a}\right)$	$\left(\frac{k}{a}\right)$
variance	$\left(\frac{1}{a^2}\right)$	$\left(\frac{k}{a^2}\right)$

The shape of a gamma distribution looks like the following for k=2, 3, 4, and 5:



pdf for a Gamma distribution with an average arrival time of 1

These can all be derived from the moment generating function.

Example 1: Determine the pdf and cdf for a Gamma distribution with

- k = 3, and
- a = 0.2

Solution: The moment generating function (i.e. LaPlace transform) is

$$\psi(s) = \left(\frac{0.2}{s+0.2}\right)^3$$

From a table of LaPlace transforms (CRC handbook of Mathematics, 1964 edition)

$$\left(\frac{1}{s-a}\right)^3 \rightarrow \frac{1}{2!}\, t^2\, e^{at}\, u(t)$$

Substituting

$$f_x = (0.2)^3\, \frac{1}{2}\, t^2\, e^{-0.2t}\, u(t)$$

The cumulative density function (cdf) is the integral of the pdf

$$F_X(s) = \left(\frac{0.2}{s+0.2}\right)^3 \left(\frac{1}{s}\right)$$

Doing partial fraction expansion

$$F_X(s) = \left(\frac{1}{s}\right) + \left(\frac{0.04}{(s+0.2)^3}\right) + \left(\frac{-0.2}{(s+0.2)^2}\right) + \left(\frac{-1}{s+0.2}\right)$$

Taking the inverse LaPlace transform gives you the cdf

$$F_x = (1 + (0.04\, t^2 - 0.2\, t - 1)\, e^{-0.2t})\, u(t)$$

Poisson Distribution.

The Poisson distribution is slightly different than the gamma distribution. Instead of the pdf being

- The time until the kth customer arrives, (Gamma)

it is

- The probability that k customers will arrive in a fixed interval (Poisson)

Likewise, the Poisson distribution is actually a discrete probability function.

The Poisson distribution is useful if you want to know

- How many cars will go through an intersection in one hour,
- How many customers will arrive in one hour,
- How many patients will go to the emergency room in one day, or
- The number of times your boss will notice you over the course of one week.
- The probability of a binomial distribution when n is large

The basic assumptions behind a Poisson distribution (Wikipedia) are:

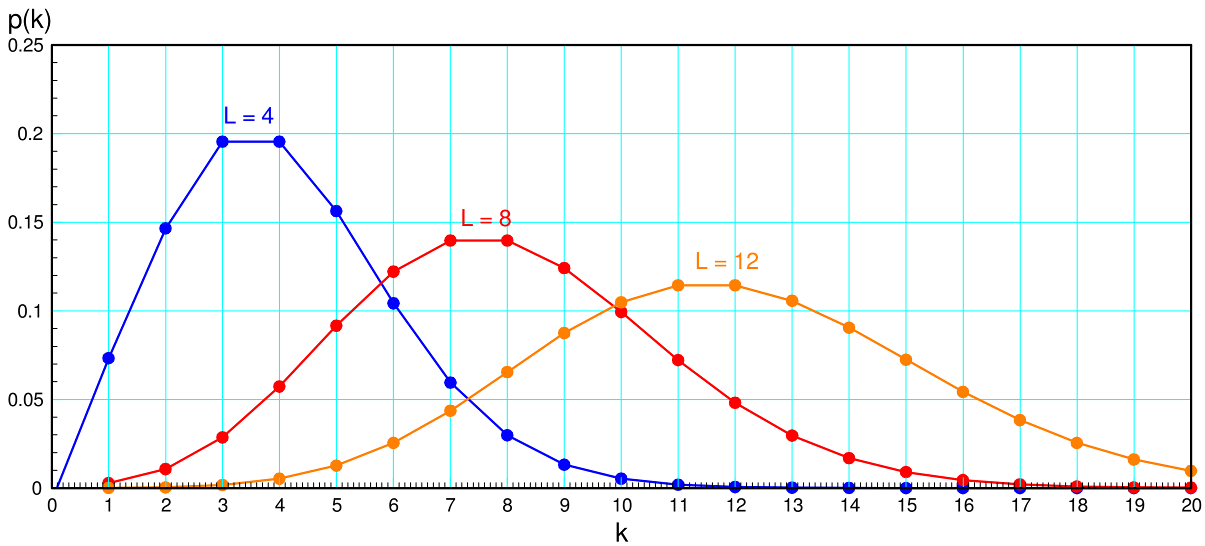
- k is the number of times an event occurs in an interval and k can take values 0, 1, 2, ...
- The occurrence of one event does not affect the probability that a second event will occur. That is, events occur independently.
- The average rate at which events occur is independent of any occurrences. For simplicity, this is usually assumed to be constant, but may in practice vary with time.
- Two events cannot occur at exactly the same instant; instead, at each very small subinterval exactly one event either occurs or does not occur.

The pdf for a Poisson distribution is (wikipedia)

	Exponential	Poisson
pdf	$f_x = a\ e^{-ax}$	$f(k, \lambda) = \frac{\lambda^k e^{-\lambda}}{k!}$
cdf	$F_x = 1 - e^{-ax}$	complicated
mgf	$\left(\frac{a}{s+a}\right)$	$\exp(\lambda(z-1))$
mean	$\left(\frac{1}{a}\right)$	λ
variance	$\left(\frac{1}{a^2}\right)$	λ

Here, $\lambda = 1/a$ for equating exponential and Poisson processes.

The pdf for a Poisson distribution looks like the following ($\lambda = 10$ (a = 1/10) for illustration purposes).



pdf for a Poisson distribution with λ = {4, 8, 12}.
Note that this is a discrete pdf - so only the integer values of k matter

Poisson Approximation for a Binomial Distribution

A Poisson distribution is also a good approximation for a binomial distribution where the number of rolls is large. For this approximation, match the means:

$$\lambda = np$$

Example 1: Plot the probability density function for a binomial distribution with

n = 100

p = 0.05

np = 20

$$f_1(x) = \binom{100}{x}(0.05)^x(0.95)^{100-x}$$

Compare this to a Poisson approximation with

$$\lambda = np$$

$$f_2(x) = \frac{1}{x!} \cdot 5^x \cdot e^{-5}$$

Matlab Code:

First, generate the binomial pdf. Use a gamma function rather than factorials so that the resulting pdf is continuous (it's easier to compare graphs this way.)

```
x = [0:0.1:20]';

f1 = gamma(100) ./ (gamma(x) .* gamma(100-x)) .* (0.05 .^ x) .* (0.95 .^
(100-x) );
```

Next, generate the Poisson approximation:

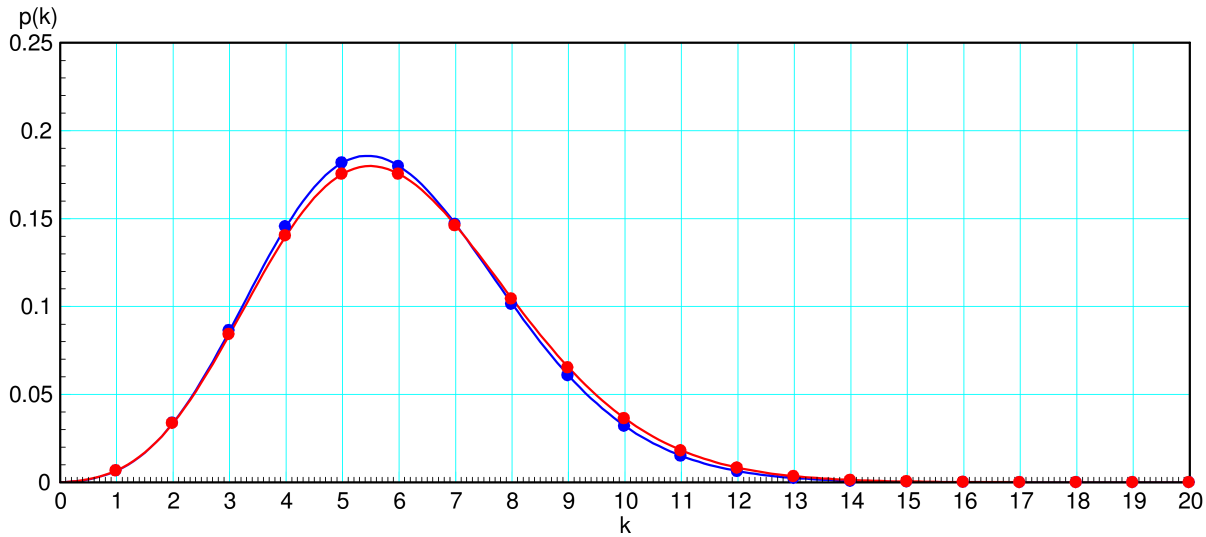
```
f2 = ( 1 ./ gamma(x) ) .* (5 .^ x) .* exp(-5);
```

Throw in a fudge factor so that the area is one. I'm not sure why it isn't, but I do know the area has to be one.

```
f1 = f1 / ( sum(f1) * 0.01 );
f2 = f2 / ( sum(f2) * 0.01 );
```

Plot the resulting pdf's:

```
plot(x,f1,x,f2)
```



Binomial (blue) vs. Poisson (red) with np = 5

A Poisson approximation is a slightly more complicated approximation for a binomial distribution than a Normal approximation. It's more accurate however.

- A normal distribution goes from $-\infty$ to $+\infty$
- A Poisson distribution is zero for $k < 0$

In the case of a binomial distribution, you'll never get a negative total. Hence, the Poisson approximation is slightly more accurate than a Normal approximation.

Example 2: Plot the probability density function for a binomial distribution with

$n = 10{,}000$

$p = 0.000\,5$

$np = 5$

This doesn't work really well using a Binomial pdf:

$$f(x) = \binom{10{,}000}{x} (0.0005)^x (0.9995)^{10{,}000-x}$$

10,000! is a really big number.

Instead, use a Poisson distribution with $\lambda = np = 5$. This gives you the same results as we got before.

$$f(x) \approx \tfrac{1}{x!} \cdot \lambda^x e^{-\lambda} \tfrac{1}{x!} \cdot 5^x \cdot e^{-5}$$