The Gamma Distribution and all its Friends

Sean Wu (slwu89berkeley.edu)

2/11/2020

The Story of the Gamma

The Density Function

The Gamma distribution is a distribution that describes how long one must wait until a certain number of things happen. We'll use the shape, rate (α, β) parameterization in this document, following Wikipedia. It has a density function that looks like this:

$$f(x; \alpha, \beta) = \frac{\beta^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\beta x}$$

Just like we mentioned in class, the "kernel" of the density function is the part that depends on x; the rest of the function is just to make it integrate to one. Let's rewrite the density but instead of saying it is equal to $(=)$ the mathematical function on the right hand side, we'll just say its "proportional to" $(\propto)$:

$$f(x; \alpha, \beta) \propto x^{\alpha-1} e^{-\beta x}$$

Why can we get away with this? Well consider how a probability density function works. You plug in some known values of α, β, and you get a curve over all possible values that x could take on, such that the integral (area under) that curve is one, because the summation/integral of probabilities of all possible events must equal one. We can be even more formal in a way familiar to statistical journals. Call $\mathcal{X}$ the total space that the probability distribution is defined on. In this case $\mathcal{X} = (0, \infty)$, because it is a distribution over positive real numbers. Then $x \in \mathcal{X}$ says that x is any *particular* real number in the set of all positive real numbers $\mathcal{X}$. In notation, saying it integrates to one looks like this:

$$\int\limits_{x \in \mathcal{X}} f(x; \alpha, \beta) dx = \int\limits_{x \in \mathcal{X}} \frac{\beta^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\beta x} dx = 1$$

Basically, we go over all the possible real numbers x, take the density at that point $f(x; \alpha, \beta)$, and multiply by the infintesimal bit dx, and sum it together. Thanks to the presence of $\frac{\beta^\alpha}{\Gamma(\alpha)}$, it will integrate to one.

Interpretation

What does the Gamma distribution mean however? Cutting to the chase, it gives the waiting time until one observes α (shape) number of events occur, where each event takes $\frac{1}{\beta}$ units of time to occur. Incidentally, each event takes one over β time to occur because β is a rate, which is in units of $time^{-1}$. So let's say one clears out their email inbox. Let's assume an email arrives on average once every 10 minutes, so $\beta = \frac{1}{10}$. Let's further assume that emails arrive *independently*, that is, if one email arrives at time t, other emails are no more or less likely to arrive due to the arrival of that one. Now let's say we want to know about how long we'd have to wait to see 20 emails show up in our empty inbox $(\alpha = 20)$.

Intuition would say that if each email takes 10 minutes on average to arrive and we want to know how long 20 of them take to show up, we'd wait on average 200 = 10*20 minutes. Somewhat uncommonly in probability theory, in this case intuition is right! We can test this in a simple Monte Carlo simulation. Let's simulate 1,000,000 Gamma distributed random numbers with those parameters.

```r
sample <- rgamma(n = 1e6,shape = 20,rate = 1/10)
mean(sample)
```

```
## [1] 200.0326
```

As expected, we get an average of about 200 (the simulation replicates deleting our inbox 1,000,000 times, waiting until 20 emails arrive, and jotting down the time of the 20^{th} arrival).

Connection to the Exponential Distribution and the Poisson Process

The astute reader may notice that the assumption of a constant arrival rate of independent events sounds suspiciously like the *Poisson Process*. And they'd be right. As a reminder, a Poisson Process is a simple stochastic process that counts the number of things that happen over some time interval $[0, t)$. In a Poisson Process, each event occurs after an Exponentially distributed delay (the "mechanistic" interpretation is that each event takes that long to "happen"). For the reader familiar with stochastic processes, this is the time-homogeneous Poisson Process, and is an example of a Markov Chain in continuous time.

Let's observe how long it takes a Poisson Process started from time 0 to accumulate 20 events, where each event takes $\frac{1}{\beta}$ units of time to happen. Again, we'll do a Monte Carlo simulation where we simulate 1,000,000 Poisson Processes until the 20th email arrives.

```r
poisson_process <- function(){
  # time of last arrival
  t <- 0
  # counter for arrivals
  i <- 0
  # quit when we have 20
  while(i < 20){
    t <- t + rexp(n = 1,rate = 1/10)
    i <- i + 1
  }
  return(t)
}

sample_pp <- replicate(n = 1e6,expr = poisson_process())
mean(sample_pp)
```

```
## [1] 199.8999
```

As expected, we also get 200 minutes, just like we got from directly sampling a Gamma distributed random number with the right parameters!

At this point you might suspect that a Gamma distribution with an integer α and β is equal to summing α Exponential random variables, each with rate parameter β. Again you'd be right!

```r
samp_exp <- replicate(n = 1e6,expr = {
  sum(rexp(n = 20,rate = 1/10))
})
mean(samp_exp)
```

```
## [1] 200.064
```

Right around 200. Just like we expected.