33-765 Homework 1

Nathaniel D. Hoffman

February 3, 2020

8. The Gamma Function

One of the first "special functions" one learns about in college—after having covered exponentials, logarithms, and trigonometric functions in high school—is probably the Gamma function, $\Gamma(x)$, which is usually defined as the following integral:

$$\Gamma(x) = \int_0^\infty \mathrm{d}t\, t^{x-1} e^{-t} \quad x \in \mathbb{C}, \mathrm{Re}(x) > 0$$

Prove the recurrence relation $\Gamma(x+1) = x\Gamma(x)$ and deduce from it that $\Gamma(N+1) = N!$ for $N \in \mathbb{N}_0$. Also, calculate $\Gamma\left(\frac{1}{2}\right)$.

We prove the recurrence relation by integrating by parts ($\int u\,\mathrm{d}v = uv - \int v\,\mathrm{d}u$), with

$$\Gamma(x+1) = \int_0^\infty \mathrm{d}t\, t^x e^{-t}$$

$$\begin{cases} u = t^x \\ \mathrm{d}u = xt^{x-1} \\ \mathrm{d}v = \mathrm{d}t\, e^{-t} \\ v = -e^{-t} \end{cases}$$

$$\Gamma(x+1) = -t^x e^{-t}\Big|_{t=0}^\infty + \int_0^\infty xt^{x-1} e^{-t}$$

$$= x\Gamma(x)$$

Next, we want to show the relation between the Gamma function and the factorial.

$$\Gamma(N+1) = N\Gamma(N) = N(N-1)\Gamma(N-1) = N(N-1)(N-2)\Gamma(N-2) = \cdots$$
$$= N(N-1)(N-2)\cdots(N-(N-1))\Gamma(1)$$

where

$$\Gamma(1) = \int_0^\infty \mathrm{d}t\, te^{-t} = -\frac{t+1}{e^t}\Big|_{t=0}^\infty = 0 - (-1) = 1$$

This is equivalent to the definition of the factorial:

$$N! = N(N-1)(N-2)\cdots(N-(N-1))^{\nearrow 1}$$

Finally, we need to calculate $\Gamma\left(\frac{1}{2}\right)$:

$$\Gamma\left(\frac{1}{2}\right) = \int_0^\infty \mathrm{d}t\, \frac{1}{\sqrt{t}} e^{-t}$$

Next, make the substitution $u^2 = t$, $u = \sqrt{t}$, $\mathrm{d}u = \frac{\mathrm{d}t}{2\sqrt{t}}$:

$$\Gamma\left(\frac{1}{2}\right) = \int_0^\infty 2\,\mathrm{d}u\,e^{-u^2}$$

This function is even, so we can change the bounds of integration and divide by two:

$$\Gamma\left(\frac{1}{2}\right) = \int_{-\infty}^\infty e^{-u^2}\,\mathrm{d}u$$

This is now the form of an un-normalized Gaussian centered at $\mu = 0$ with a standard deviation of $\sigma = \frac{1}{\sqrt{2}}$. If it was normalized, there would be an additional factor of $\frac{1}{\sqrt{\pi}}$ multiplying it, and that would make the integral equal to 1:

$$\int_{-\infty}^\infty \frac{1}{\sqrt{\pi}}e^{-u^2}\,\mathrm{d}u = 1$$

so

$$\Gamma\left(\frac{1}{2}\right) = \sqrt{\pi}$$

9. The Cauchy-Lorentz Distribution

The p-density of the Cauchy-Lorentz distribution $p_{x_0,a}(x)$ with location parameter x_0 and scale parameter a is defined as

$$p_{x_0,a}(x) = \frac{a/\pi}{a^2 + (x - x_0)^2}.$$

1. Show that $p_{x_0,a}(x)$ is normalized. Also show that none of its higher moments exist.

To show that the distribution is normalized, we integrate over the domain and show that we get 1. I will assume it is okay if I know that the integral of $\frac{1}{1-x^2} = \arctan(x)$ is well known, so the a and x_0 just scale and shift the integral while the $\frac{1}{\pi}$ factor can be brought outside:

$$\int_{-\infty}^\infty p_{x_0,a}(x)\,\mathrm{d}x = \frac{1}{\pi}\arctan\left(\frac{x - x_0}{a}\right)\Big|_{-\infty}^\infty = \frac{1}{\pi}\left(\frac{\pi}{2} - \left(-\frac{\pi}{2}\right)\right) = 1$$

To show that none of its higher moments exist, let's look at the first moment. For simplicity (and without loss of generality), I will take the location parameter to be $x_0 = 0$ and the scale parameter to be $a = 1$ (If the moments did exist, I could just reposition and scale them after this calculation):

$$\frac{1}{\pi}\int_{-\infty}^\infty \frac{x}{1 + x^2}\,\mathrm{d}x = \frac{1}{2\pi}\ln\left(1 + x^2\right)\Big|_{x=-\infty}^\infty$$

The x^2 term in the logarithm makes this tricky. Both ends of the integral will evaluate to the same limit (∞) at the same speed (logarithmically), so if we take the integral as a limit:

$$\int_{-\infty}^\infty xp(x) \to \lim_{a\to\infty}\int_{-a}^a = \lim_{a\to\infty}\frac{1}{2\pi}\ln\left(\frac{1 + a^2}{1 + a^2}\right) = \frac{1}{2\pi}\ln 1 = 0$$

However we could change the bounds of integration to be

$$\lim_{a\to\infty}\int_{-a}^{2a} xp(x) = \lim_{a\to\infty}\frac{1}{2\pi}\ln\left(\frac{1 + 4a^2}{1 + a^2}\right) = \frac{\ln 4}{2\pi} \neq 0$$

Since these limits should give the same result (and don't), the limit does not exist. For the next expectation value, we can factor out the x^2:

$$\frac{1}{\pi}\int_{-\infty}^{\infty}\frac{x^2}{1+x^2}\,\mathrm{d}x = \int_{-\infty}^{\infty}\frac{1}{\pi} - \int_{-\infty}^{\infty}p(x)\,\mathrm{d}x = \infty$$

For odd moments after this, we can factor out an even power of x^2 to get the first result where the limits don't match, and for even moments, we will have results similar to the second example, where the integral diverges to infinity.

2. For the special case $x_0 = 0$, calculate the characteristic function $\tilde{p}_{0,a}(k)$.

In the first homework, we were given

$$\tilde{p}(k) = \int \mathrm{d}x\, p(x)e^{\imath k x}$$

Using this definition,

$$\tilde{p}_{0,a}(k) = \frac{1}{\pi}\int \mathrm{d}x\, \frac{ae^{\imath k x}}{a^2 + x^2}$$

We can integrate over a closed contour in complex space and use residue theorem to simplify this problem. Since the integrand is holomorphic in the upper half-plane and continuous except at $z = +\imath a$, Jordan's lemma says that if k is positive, we can close the contour in the upper half-plane. If k is negative, we have to close in the lower half-plane, where the function is discontinuous at $z = -\imath a$. Let us first consider positive k:

$$\tilde{p}(k) = \frac{1}{\pi}\oint_C \mathrm{d}z\, \frac{ae^{\imath k z}}{a^2 + z^2} = \frac{1}{\pi}\int_{C_1} \mathrm{d}z\, \frac{ae^{\imath k z}}{a^2 + z^2} + \frac{1}{\pi}\int_{-R}^{R} \mathrm{d}x\, \frac{ae^{\imath k x}}{a^2 + x^2}$$

where C_1 is a CCW-oriented semicircle in the upper half-plane and we want the limit as $R \to \infty \in \mathbb{R}$. Residue theorem tells us that the integral around C is proportional to the residues enclosed, and Jordan's lemma tells us the integral over C_1 vanishes as we take $R \to \infty$:

$$\tilde{p}(k) = \frac{2\pi\imath}{\pi}\mathrm{Res}(f, \imath a) \quad f = \frac{ae^{\imath k z}}{a^2 + z^2}$$

The residue is

$$\mathrm{Res}(f, \imath a) = \left.\frac{ae^{\imath k z}}{z + \imath a}\right|_{z = \imath a} = -\frac{1}{2}\imath e^{-ak}$$

so, for $k > 0$:

$$\tilde{p}_{0,a}(k) = 2\imath\left(-\frac{1}{2}\imath e^{-ak}\right) = e^{-ak} \quad k > 0$$

For $k < 0$, the sign in the exponential changes, while the sign from residue theorem also changes because we are now integrating in the lower half-plane:

$$\tilde{p}_{0,a}(k) = e^{ak} \quad k < 0$$

All together,

$$\tilde{p}_{0,a}(k) = e^{-a|k|}$$

3. Let the random variables X_i be the outcomes of independent measurements of some observable that is Cauchy-Lorentz distributed with $p_{0,a}(x)$. We want to accurately measure that observable, so we define the average $Y_N = \frac{1}{N}\sum_{i=1}^{N} X_i$ and aim for a large N. What is the p-density of Y_N? Discuss your findings in the context of the Central Limit Theorem!

Again, from the first homework, we showed two important things. First, summing random distributions multiplies their characteristic functions, and second, scaling the distribution in x also scales it in k. Therefore, we can write down the characteristic function of the distribution:

$$\tilde{p}_Y(k) = \prod_{i=1}^{N} e^{-\frac{a}{N}|k|} = e^{-a|k|}$$

However, this is just the characteristic function for the single distribution! The p-density of Y_N is $p_{0,a}(x)$. In terms of the Central Limit Theorem, this original distribution does not have any higher order moments (even the mean does not exist, as we showed above). The Gaussian that comes from the Central Limit Theorem has mean and variance which depend on the random variables that are input, so it makes sense that we don't gain any information about these moments by summing the distributions.

10. Another Frequently Occurring Distribution Function

Let X be a random variable which has the p-density $P_{\mu,\sigma^2}(x)$ given by

$$P_{\mu,\sigma^2}(x) = \frac{1}{x\sqrt{2\pi\sigma^2}} \exp\{-\frac{(\log x - \mu)^2}{2\sigma^2}\} \quad x \in \mathbb{R}^+$$

1. Show that $P_{\mu,\sigma^2}(x)$ is normalized and that its rth moment is $\langle X^r \rangle = e^{r\mu + \frac{1}{2}r^2\sigma^2}$ for all $r \in \mathbb{R}$.

To show that it is normalized, all we have to do is make the following substitution in typical normalization integral:

$$u = \ln x \qquad \mathrm{d}u = \frac{\mathrm{d}x}{x}$$

such that the normalization is now

$$\int_{-\infty}^{\infty} \frac{\mathrm{d}u}{\sqrt{2\pi\sigma^2}} e^{-\frac{(u-\mu)^2}{2\sigma^2}} = 1$$

since this is just an integral over a normalized Gaussian.
Next, we want to find the moments. If we do the same substitution as before, we can use the fact that $x = e^u$ to make

$$\langle X^r \rangle = \int e^{ru} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(u-\mu)^2}{2\sigma^2}} \, \mathrm{d}u$$

If we just look at the exponentials, we have

$$-\frac{1}{2\sigma^2}\left[u^2 + 2u\mu - \mu^2 + 2ru\sigma^2\right] = -\frac{1}{2\sigma^2}\left[(u - (\mu + r\sigma^2))^2 - 2r\mu\sigma^2 - r^2\sigma^4\right]$$

We can put this back into the equation to get

$$\langle X^r \rangle = \int e^{-r\mu - \frac{1}{2}r^2\sigma^2} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(u-(\mu+r\sigma^2))^2}{2\sigma^2}} \, \mathrm{d}u$$

$$= e^{-r\mu - \frac{1}{2}r^2\sigma^2} \int \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(u-(\mu+r\sigma^2))^2}{2\sigma^2}} \, \mathrm{d}u \overset{1}{\nearrow}$$

Since the integral is just the normalization of a shifted Gaussian

2. If $P(x)$ is a p-density, a number m (not necessarily unique!) such that $\int_{-\infty}^{m} \mathrm{d}x\, P(x) = \frac{1}{2}$ is called a *median* of that distribution. Show that for $P_{\mu,\sigma^2}(x)$ the median is unique and given by $m = e^\mu$.

Let's now make the substitution

$$u = \ln x - \mu \qquad u + \mu = \ln x \qquad \mathrm{d}u = \frac{1}{x}\,\mathrm{d}x$$

so

$$\frac{1}{2} = \int_{-\infty}^{\ln m - \mu} \frac{1}{2\pi\sigma^2} e^{-\frac{u^2}{2\sigma^2}}\,\mathrm{d}u$$

This is just a Gaussian centered at 0. The median of a Gaussian is the mean, so

$$0 = \ln m - \mu \implies m = e^{\mu}$$

The only case where the median would not be unique is if there are portions (intervals) of the distribution function which are zero (or for some reason, negative). However, the only place this distribution can be zero is at $x = \infty$ and $x = 0$, so the integral will be monotonically increasing over the whole distribution.

3. If we consider the new random variable $Y = \ln X$, show that its p-density is a Gaussian with mean μ and variance σ^2.

By the transformation theorem,

$$P_Y(y) = \int \mathrm{d}x\, \delta(y - \ln x) \frac{1}{x\sqrt{2\pi\sigma^2}} e^{-\frac{(\ln x - \mu)^2}{2\sigma^2}}$$

The δ-function is just the substitution I made before, and it takes $\frac{\mathrm{d}x}{x} \to \mathrm{d}y$ and $\ln x \to y$, so the result is

$$P_Y(y) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{(y - \mu)^2}{2\sigma^2}}$$

which is a Gaussian with mean μ and standard deviation σ.

4. Plot the distribution function $P_{\mu,\sigma^2}(x)$ for the three cases $\mu = 0$ and $\sigma^2 \in \left\{\frac{1}{3}, 1, 3\right\}$ in the same diagram (you might want to choose $x \in [0,5]$). Mark the mean and the median in each case. Comment on what you find noteworthy.

The plot can be seen in the attached figure. A few things that I found noteworthy were that first, the median doesn't change, since $\mu = 0$ and we showed that the median is $e^{\mu} = 1$ for these distributions. Additionally, the means are all larger than the median, and they are increasingly larger with the values of σ^2. While I didn't expect the mean to be μ, I really didn't expect it to be completely independent of the median in this case. However, the equations for the median and the mean explain why these values seem to be independent for $\mu = 0$, since the median doesn't depend on σ^2 while the mean depends only on σ^2.

11. Jensen's Inequality

Let $f \colon \mathbb{R} \supset G \to \mathbb{R}$, $x \mapsto f(x)$ be a function which is convex and differentiable on the connected domain G. As a consequence, for all $x, x_0 \in G$ we have

$$f(x) \geq f(x_0) + f'(x_0)(x - x_0).$$

1. Draw an educationally pristine graph to illustrate that the inequality makes sense as a definition of convexity.
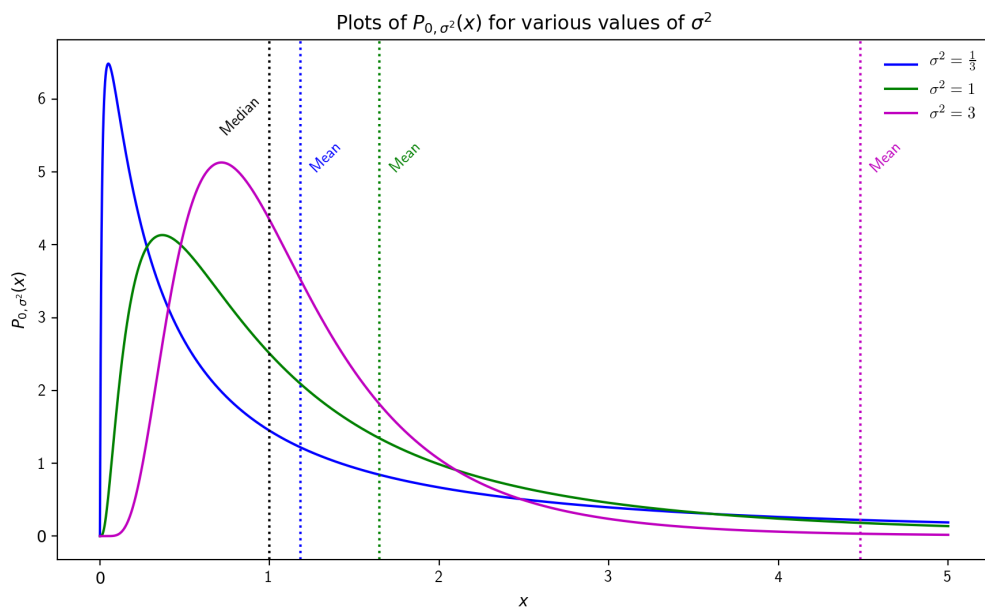
Figure 0.1: Plot for 10.4

> I am not dedicated enough to draw this with LaTeX and even though my hand-drawn art isn't that great, it's a lot faster.

2. Let X be a continuous random variable and $f(x)$ be a convex (differentiable) function. Prove Jensen's inequality:

$$\langle f(X) \rangle \geq f(\langle X \rangle),$$

where the average is taken over whatever p-density underlies the randomness of X.

Let's work with a line tangent to $f(X)$ at $\langle X \rangle = \mu$. This line can be written as

$$L(X) = f(\mu) + f'(\mu)(X - \mu) = (f(\mu) - f'(\mu)(\mu)) + f'(\mu)(X) = a + bX$$

If f is convex, we know from the above definition that

$$\langle f(X) \rangle \geq \langle L(X) \rangle$$

Since the expectation value is a linear functional, we can perform the following operations:

$$\langle f(X) \rangle \geq \langle L(X) \rangle = \langle a + bX \rangle = a + b \langle X \rangle = L(\langle X \rangle) = f(\langle X \rangle)$$

The last step works because we defined the line to be tangent to f at $\langle X \rangle$, so it will have the same value there.