# Chapter 15

## *The Gamma Function and Related Distributions*

In this chapter we'll explore some of the strange and wonderful properties of the Gamma function $\Gamma(s)$, defined by

For $s > 0$ (or actually $\Re(s) > 0$), the **Gamma function** $\Gamma(s)$ is

$$\Gamma(s) \;=\; \int_0^\infty e^{-x} x^{s-1} dx \;=\; \int_0^\infty e^{-x} x^s \frac{dx}{x}.$$

There are countless integrals or functions we can define. Just looking at it, there's nothing that suggests it's one of the most important functions in all of mathematics, appearing throughout probability and statistics (and many other fields), but it does. We'll see where it occurs and why, and discuss many of its most important properties. If you can't wait, before reading on evaluate the integral for $s = 1, 2, 3$ and 4, and try and figure out the pattern.

### 15.1 Existence of $\Gamma(s)$

Looking at the definition of $\Gamma(s)$, it's natural to ask: *Why do we have restrictions on $s$?* Whenever you're given an integrand, you must make sure it's well-behaved before you can conclude the integral exists. The purpose of this section is to highlight some useful techniques to investigate integrals. Frequently there are two trouble points to check, near $x = 0$ and near $x = \pm\infty$ (okay, three points).

For example, consider the function $f(x) = x^{-1/2}$ on the interval $[0, \infty)$. This function blows up at the origin, but only mildly. Its integral is $2x^{1/2}$, and this is integrable near the origin. This just means that

$$\lim_{\epsilon \to 0} \int_\epsilon^1 x^{-1/2} dx$$

exists and is finite. Unfortunately, even though this function is tending to zero, it approaches zero so slowly for large $x$ that it's not integrable on $[0, \infty)$. The problem is that integrals such as

$$\lim_{B\to\infty} \int_1^B x^{-1/2} dx$$

are infinite. Can the reverse problem happen, namely our function decays fast enough for large $x$ but blows up too rapidly for small $x$? Sure – consider $g(x) = 1/x^2$. Note $g$ has a nice integral:

$$G(x) \;=\; \int g(x)dx \;=\; \int \frac{dx}{x^2} \;=\; -\frac{1}{x}.$$

Now the integral over large $x$ is fine and finite, being just

$$\lim_{B\to\infty} \int_1^B g(x)dx \;=\; \lim_{B\to\infty} -\frac{1}{x}\Big|_1^B \;=\; \lim_{B\to\infty} \left[1 - \frac{1}{B}\right] \;<\; \infty;$$

however, the integral over small $x$ blows up:

$$\lim_{\epsilon\to\infty} \int_\epsilon^1 g(x)dx \;=\; \lim_{\epsilon\to 0} -\frac{1}{x}\Big|_\epsilon^1 \;=\; \lim_{\epsilon\to 0} \left[\frac{1}{\epsilon} - 1\right] \;=\; \infty.$$

So it's possible for a positive function to fail to be integrable because it decays too slowly for large $x$, or it blows up too rapidly for small $x$. As a rule of thumb, if as $x \to \infty$ a function is decaying faster than $1/x^{1+\epsilon}$ for any epsilon, then the integral at infinity will be finite. For small $x$, if as $x \to 0$ the function is blowing up slower than $x^{-1+\epsilon}$ then the integral at 0 will be okay near zero. You should always do tests like this, and get a sense for when things will exist and be well-defined.

Returning to the Gamma function, let's make sure it's well-defined for any $s > 0$. The integrand is $e^{-x}x^{s-1}$. As $x \to \infty$, the factor $x^{s-1}$ is growing polynomially but the term $e^{-x}$ is decaying exponentially, and thus their product decays rapidly. If we want to be a bit more careful and rigorous, we can argue as follows: choose some integer $M > s + 1701$ (we put in a large number to alert you to the fact that the actual value of our number does not matter). We clearly have $e^x > x^M/M!$, as this is just one term in the Taylor series expansion of $e^x$ (all terms have a positive contribution as $x > 0$). Thus $e^{-x} < M!/x^M$, and the integral for large $x$ is finite and well-behaved, as it's bounded by

$$\begin{aligned}
\int_1^B e^{-x}x^{s-1}dx \;&\leq\; \int_1^B M!x^{-M}x^{s-1}dx \\
&=\; \int M! \int_1^B x^{s-M-1} \\
&=\; M!\frac{x^{s-M}}{s-M}\Big|_1^B \\
&=\; \frac{M!}{s-M}\left[\frac{1}{B^{M-s}} - 1\right].
\end{aligned}$$

Remember, our goal is not just to understand the Gamma function, but to understand functions in general. Thus, it's important to get a sense of what techniques are available, and when a method has a chance of succeeding. Our approach above was a very good choice. We know $e^x$ grows very rapidly, so $e^{-x}$ decays quickly. We're **borrowing some of the decay** from $e^{-x}$ to handle the $x^{s-1}$ piece; **borrowing decay** is a great technique to bound the behavior of integrals.

What about the other issue, near $x = 0$? Well, near $x = 0$ the function $e^{-x}$ is bounded; it's largest value is when $x = 0$ so it's at most 1. Thus

$$\int_0^1 e^{-x}x^{s-1}dx \ \le \ \int_0^1 1 \cdot x^{s-1}dx \ = \ \frac{x^s}{s}\Big|_0^1 \ = \ \frac{1}{s}.$$

We've shown everything is fine for $s > 0$; what if $s \le 0$? Could these values be permissible as well? The same type of argument as above shows that there are no problems when $x$ is large. Unfortunately, it's a different story for small $x$. For $x \le 1$ we clearly have $e^{-x} \ge 1/e$; before we had an upper bound to show the integral was okay, now we need a lower bound to show it blows up. Thus our integrand is at least as large as $x^{s-1}/e$. If $s \le 0$, this is no longer integrable on $[0, 1]$. For definiteness, let's do $s = -2$. Then we have

$$\int_0^\infty e^{-x}x^{-3}dx \ \ge \ \int_0^\infty \frac{1}{e}\,x^{-3}dx \ = \ -\frac{1}{e}x^{-2}\Big|_0^1 \ = \ \infty,$$

and this blows up.

The arguments above can (and should!) be used every time you meet an integral. Even though our analysis hasn't suggested a reason why anyone would *care* about the Gamma function, we at least know that it's well-defined and exists for all $s > 0$. In the next section we'll show how to make sense of Gamma for all values of $s$. This should be a bit alarming – we've just spent this section talking about being careful and making sure we only use integrals where they are well-defined, and now we want to talk about putting in values such as $s = -1/2$? Obviously, whatever we do, it won't be anything as simple as just plugging $s = -1/2$ into the formula.

If you're interested, $\Gamma(-1/2) = \sqrt{\pi}$ – we'll prove this soon!

If you're looking for a fun integral, explore whether or not $\int_0^\infty f(x)dx$ exists, where

$$f(x) \ = \ \begin{cases} \frac{1}{(x+1)\log^2(x+1)} & \text{if } x > 0 \\ 0 & \text{otherwise.} \end{cases}$$

Is the integral fine at infinity? At zero?

## 15.2 The Functional Equation of $\Gamma(s)$

We turn to *the* most important property of $\Gamma(s)$. This property allows us to make sense of *any* value of $s$ as input, such as the $s = -1/2$ of the last section. Obviously this can't mean just naively throwing in any $s$ in the definition, though many good mathematicians have accidentally done so. What we're going to see is the **Analytic (or Meromorphic) Continuation** . The gist of this is that we can take a function $f$ that makes sense in one region and extend its definition to a function $g$ defined on a larger region in such a way that our new function $g$ agrees with $f$ where they are both defined, but $g$ is defined for more points.

The following absurdity is a great example. What is

$$1 + 2 + 4 + 8 + 16 + 32 + 64 + \cdots ?$$

Well, we're adding all the powers of 2, thus it's clearly infinity, right? Wrong – the "natural" meaning for this sum is $-1$! A sum of infinitely many positive terms is negative? What's going on here?

This example comes from something you've probably seen many times, a geometric series. If we take the sum

$$1 + r + r^2 + r^3 + r^4 + r^5 + r^6 + \cdots$$

then, *so long as* $|r| < 1$, the sum is just $\frac{1}{1-r}$. There are many ways to see this. The most common, as well as one of the most boring, is to let

$$S_n = 1 + r + \cdots + r^n.$$

If we look at $S_n - rS_n$, almost all the terms cancels; we're left with

$$S_n - rS_n = 1 - r^{n+1}.$$

We factor the left hand side as $(1-r)S_n$, and then dividing both sides by $1-r$ gives

$$S_n = \frac{1 - r^{n+1}}{1 - r}.$$

If $|r| < 1$ then $\lim_{n\to\infty} r^n = 0$, and thus taking limits gives

$$\sum_{m=0}^{\infty} r^m = \lim_{n\to\infty} S_n = \lim_{n\to\infty} \frac{1 - r^{n+1}}{1 - r} = \frac{1}{1-r}.$$

This is known as the **geometric series formula**, and is used in a variety of problems. See §1.2 for a more entertaining derivation.

Let's rewrite the above. The summation notation is nice and compact, but that's not what we want right now – we want to really see what's going on. We have

$$1 + r + r^2 + r^3 + r^4 + r^5 + r^6 + \cdots = \frac{1}{1-r}, \quad |r| < 1.$$

Note the left hand side makes sense only for $|r| < 1$, but the right hand side makes sense for *all* values of $r$ other than 1! We say the right hand side is an **analytic continuation** of the left, with a **pole** at $s = 1$ (poles are where our functions blow-up).

Let's define the function

$$f(x) = 1 + x + x^2 + x^3 + x^4 + x^5 + x^6 + \cdots.$$

For $|x| < 1$ we also have

$$f(x) = \frac{1}{1-x}.$$

We're now ready for the big question: what's $f(2)$? If we use the second definition, it's just $\frac{1}{1-2} = -1$, while if we use the first definition it's that strange sum of all the powers of 2. *THIS* is the sense in which we mean the sum of all the powers of 2 is -1. We don't mean plugging in 2 for the series expansion; instead, we evaluate the extended function at 2.

It's now time to apply these techniques to the Gamma function. We'll show, using integration by parts, that $\Gamma(s)$ can be extended for all $s$ (or at least for all $s$ except the negative integers and zero). Before doing the general case, let's do a few representative examples to see why integration by parts is such a good thing to do, and to get a feeling for the Gamma function's behavior. Recall

$$\Gamma(s) \;=\; \int_0^\infty e^{-x}x^{s-1}dx, \quad s > 0.$$

The easiest value of $s$ to take is $s = 1$, as then the $x^{s-1}$ term becomes the harmless $x^0 = 1$. In this case, we have

$$\Gamma(1) \;=\; \int_0^\infty e^{-x}dx \;=\; -e^{-x}\Big|_0^\infty \;=\; -0+1 \;=\; 1.$$

Building on our success, what's the next easiest value of $s$ to take? A little experimentation suggests we try $s = 2$. This makes $x^{s-1}$ equal $x$, a nice, integer power. We find

$$\Gamma(2) \;=\; \int_0^\infty e^{-x}xdx.$$

Now we can begin to see why integration by parts will play such an important role. If we let $u = x$ and $dv = e^{-x}dx$, then $du = dx$ and $v = -e^{-x}$, then we'll see great progress – we start with needing to integrate $xe^{-x}$ and after integration by parts we're left with having to do $e^{-x}$, a wonderful savings. Putting in the details, we find

$$\Gamma(2) \;=\; uv\Big|_0^\infty - \int_0^\infty vdu \;=\; -xe^{-x}\Big|_0^\infty + \int_0^\infty e^{-x}dx.$$

The boundary term vanishes (it's clearly zero at zero; use L'Hopital's Rule to evaluate it at $\infty$, giving $\lim_{x\to\infty}\frac{x}{e^x} = \lim_{x\to\infty}\frac{1}{e^x} = 0$), while the other integral is just $\Gamma(1)$. We've thus shown that

$$\Gamma(2) \;=\; \Gamma(1);$$

however, it's more enlightening to write this in a slightly different way. We took $u = x$ and then said $du = dx$; let's write it as $u = x^{\mathbf{1}}$ and $du = \mathbf{1}dx$. This leads us to

$$\Gamma(2) \;=\; \mathbf{1}\cdot\Gamma(1).$$

At this point you should be skeptical – does it really matter? Anything times 1 is just itself! It does matter, and should remind you of our work with binomial coefficients and combinatorics. If we were to calculate $\Gamma(3)$, we would find it equals $\mathbf{2}\cdot\Gamma(2)$, and if we then progressed to $\Gamma(4)$ we would see it's just $\mathbf{3}\cdot\Gamma(3)$. This pattern suggests $\Gamma(s+1) = s\Gamma(s)$, which we now prove.

*Proof that* $\Gamma(s+1) = s\Gamma(s)$ *for* $\Re(s) > 0$. We have

$$\Gamma(s+1) \;=\; \int_0^\infty e^{-x}x^{s+1-1}dx \;=\; \int_0^\infty e^{-x}x^s dx.$$

We now integrate by parts. Let $u = x^s$ and $dv = e^{-x}dx$; we're basically forced to do it this way as $e^{-x}$ has a nice integral, and by setting $u = x^s$ when we differentiate the power of our polynomial goes down, leading to a simpler integral. We thus have

$$u \;=\; x^s, \quad du \;=\; sx^{s-1}dx, \quad dv \;=\; e^{-x}dx, \quad v \;=\; -e^{-x},$$

which gives

$$\begin{aligned}\Gamma(s+1) &= -x^s e^{-x}\Big|_0^\infty + \int_0^\infty e^{-x}sx^{s-1}dx \\ &= 0 + s\int_0^\infty e^{-x}x^{s-1}dx \;=\; s\Gamma(s),\end{aligned}$$

completing the proof. □

This relation is so important its worth isolating it, and giving it a name.

**Functional equation of $\Gamma(s)$:** The Gamma function satisfies

$$\Gamma(s+1) \;=\; s\Gamma(s).$$

This allows us to extend the Gamma function to all $s$. We call the extension the Gamma function as well, and it's well-defined and finite for all $s$ save the negative integers and zero.

Let's return to the example from the previous section. Later we'll prove that $\Gamma(1/2) = \sqrt{\pi}$. For now we assume we know this, and show how we can figure out what $\Gamma(-3/2)$ should be. From the functional equation, $\Gamma(s+1) = s\Gamma(s)$. We can rewrite this as $\Gamma(s) = s^{-1}\Gamma(s+1)$, and we can now use this to 'walk up' from $s = -3/2$, where we don't know the value, to $s = 1/2$, where we assume we do. We have

$$\Gamma\left(-\frac{3}{2}\right) \;=\; -\frac{2}{3}\Gamma\left(-\frac{1}{2}\right) \;=\; -\frac{2}{3}\cdot(-2)\,\Gamma\left(\frac{1}{2}\right) \;=\; \frac{4\sqrt{\pi}}{3}.$$

This is the power of the functional equation – it allows us to define the Gamma function essentially everywhere, so long as we know its values for $s > 0$ (or more generally for $\Re(s) > 0$). Why are zero and the negative integers special? Well, let's look at $\Gamma(0)$:

$$\Gamma(0) \;=\; \int_0^\infty e^{-x}x^{0-1}dx \;=\; \int_0^\infty e^{-x}x^{-1}dx.$$

The problem is that this isn't integrable. While it decays very rapidly for large $x$, for small $x$ it looks like $1/x$. The details are:

$$\lim_{\epsilon\to 0}\int_\epsilon^1 e^{-x}x^{-1}dx \;\geq\; \frac{1}{e}\lim_{\epsilon\to 0}\int_\epsilon^1 \frac{dx}{x} \;=\; \frac{1}{e}\lim_{\epsilon\to 0}\log x\Big|_\epsilon^1 \;=\; \lim_{\epsilon\to 0} -\log\epsilon \;=\; \infty.$$

Thus $\Gamma(0)$ is undefined, and hence by the functional equation it's also undefined for all the negative integers.

## 15.3 The Factorial Function and $\Gamma(s)$

In the last section we showed that $\Gamma(s)$ satisfies the functional equation $\Gamma(s+1) = s\Gamma(s)$. This is reminiscent of a relation obeyed by a better known function, the factorial function. Remember

$$n! \ = \ n \cdot (n-1) \cdot (n-2) \cdots 3 \cdot 2 \cdot 1;$$

we write this in a more suggestive way as

$$n! \ = \ n \cdot (n-1)!.$$

Note how similar this looks to the relationship satisfied by $\Gamma(s)$. It's not a coincidence – the Gamma function is a generalization of the factorial function!

> **$\Gamma(s)$ and the Factorial Function.** If $n$ is a non-negative integer, then $\Gamma(n+1) = n!$. Thus the Gamma function is an extension of the factorial function.

We've shown that $\Gamma(1) = 1$, $\Gamma(2) = 1$, $\Gamma(3) = 2$, and so on. We can interpret this as $\Gamma(n) = (n-1)!$ for $n \in \{1, 2, 3\}$; however, applying the functional equation allows us to extend this equality to all $n$. We proceed by induction. Proofs by induction have two steps, the base case (where you show it holds in some special instance) and the inductive step (where you assume it holds for $n$ and then show that it holds for $n+1$). See §A.2 for a review and additional examples of this technique.

We've already done the base case, as we've checked $\Gamma(1) = 0!$. (This is probably one of the few times in your life when you are grammatically correct to end a sentence with an exclamation point *and* a period. It's a good idea not to use another exclamation point for excitement, as the !!, called the double factorial, has a meaning in probability too!) We checked a few more cases than we needed. Typically that's a good strategy when doing inductive proofs. By getting your hands dirty and working out a few cases in detail, you often get a better sense of what's going on, and you can see the pattern. Remember, we initially wrote $\Gamma(2) = \Gamma(1)$, but after some thought (as well as years of experience) we rewrote it as $\Gamma(2) = \mathbf{1} \cdot \Gamma(1)$.

We now turn to the inductive step. We assume $\Gamma(n) = (n-1)!$, and we must show $\Gamma(n+1) = n!$. From the functional equation, $\Gamma(n+1) = n\Gamma(n)$; but by the inductive step $\Gamma(n) = (n-1)!$. Combining gives $\Gamma(n+1) = n(n-1)!$, which is just $n!$, or what we needed to show. This completes the proof. □

We now have two different ways to calculate say $1020!$. The first is to do the multiplications out: $1020 \cdot 1019 \cdot 1018 \cdots$. The second is to look at the corresponding integral:

$$1020! \ = \ \Gamma(1021) \ = \ \int_0^\infty e^{-x} x^{1020} dx.$$

There are advantages to both methods; I want to discuss some of the benefits of the integral approach, as this is definitely not what most people have seen. Integration is hard; most students don't see it until late in high school or college. We all know how to multiply numbers – we'be been doing this since grade school. Thus, why make our lives difficult by converting a simple multiplication problem to an integral?

The reason is a general principle of mathematics – often by looking at things in a different way, from a higher level, new features emerge that you can exploit. Also, once we write it as an integral we have a lot more tools in our arsenal; we can use results from integration theory and from analysis to study this. We do this in Chapter 18, and see just how much we can learn about the factorial function by recasting it as an integral.

*Remark: The relation in this section is so important it's worth one last look before moving on. In the early chapters of the book we did a lot with combinatorics and probability. The factorial function was almost always lurking in the background, either directly through multiplicative trees of probabilities, or indirectly through binomial coefficients (recall $\binom{n}{k}$, the number of ways of choosing $k$ objects from $n$ when order doesn't matter, is $n!/k!(n-k)!$). This section connects the factorial function to the Gamma function, and suggests the possibility of a greater understanding through calculus and real analysis.*

## 15.4 Special Values of $\Gamma(s)$

We know that $\Gamma(s+1) = s!$ whenever $s$ is a non-negative integer. Are there other choices of $s$ that are important, and if so, what are they? In other words, we've just generalized the factorial function. What was the point? It may be that the non-integral values are just curiosities that don't really matter, and the entire point might be to have the tools of calculus and analysis available to study $n!$. This, however, is most emphatically *not* the case. Some of these other values are very important in probability; in a bit of foreshadowing, we'll say they play a *central* role in the subject.

So, what are the important values for $s$? Because of the functional equation, once we know $\Gamma(1)$ we know the Gamma function at all non-negative integers, which gives us all the factorials. So 1 is an important choice of $s$. What should we look at next? The simplest number after the integers are the half-integers, those of the form $n/2$ where $n$ is an integer. The simplest one which isn't an integer is 1/2. We'll now see that $s = 1/2$ is also very important.

One of the most important, if not the most important, distribution is the normal distribution (see Chapter 14 for a detailed tour). We say $X$ is normally distributed with mean $\mu$ and variance $\sigma^2$, written $X \sim N(\mu, \sigma^2)$, if the density function is

$$f_{\mu,\sigma}(x) \;=\; \frac{1}{\sqrt{2\pi\sigma^2}}\, e^{-(x-\mu)^2/2\sigma^2}.$$

Looking at this density, we see there are two parts. There's the exponential part, and the constant factor of $1/\sqrt{2\pi\sigma^2}$. Because the exponential function decays so rapidly, the integral is finite and thus, if appropriately normalized, we will have a probability density. The hard part is determining just what this integral is. Let $g(x) = e^{-(x-\mu)^2/2\sigma^2}$. As $g$ decays rapidly and is never negative, it can be rescaled to integrate to one and hence become a probability density. That scale factor is just

$1/c$, where

$$c \;=\; \int_{-\infty}^{\infty} e^{-(x-\mu)^2/2\sigma^2}.$$

In Chapter 14 we'll see numerous applications and uses of the normal distribution. It's not hard to make an argument that it's important, and thus we *need* to know the value of this integral. That said, why is this in the Gamma function chapter?

The reason is that, with a little bit of algebra and some change of variables, we'll see that this integral is just $\sqrt{2}\Gamma(1/2)\sigma^2$. We might as well assume $\mu = 0$ and $\sigma = 1$ (if not, then step 1 is just to change variables and let $t = \frac{x-\mu}{\sigma}$). So let's look at

$$I \;:=\; \int_{-\infty}^{\infty} e^{-x^2/2}dx \;=\; 2\int_0^{\infty} e^{-x^2/2}dx,$$

where we **exploited the symmetry** to reduce the integration to be from 0 to infinity (see §A.4). This only vaguely looks related to the Gamma function. The Gamma function is the integral of $e^{-x}$ times a polynomial in $x$, while here we have the exponential of $-x^2/2$. Looking at this, we see that there's a natural change of variable to try to make our integral look like the Gamma function at some special point. We try $u = x^2/2$, as this is the only way we'll end up with the exponential of the negative of our variable. We want to find $dx$ in terms of $u$ and $du$ for the change of variables, thus we rewrite $u = x^2/2$ as $x = (2u)^{1/2}$, which gives $dx = (2u)^{-1/2}du$. Plugging all of these in, we see

$$I \;=\; 2\int_0^{\infty} e^{-u}(2u)^{-1/2}du \;=\; \sqrt{2}\int_0^{\infty} e^{-u}u^{-1/2}du.$$

We're almost done – this does look very close to the Gamma function. There are just two issues: one trivial and one minor. The first is that we're using the letter $u$ instead of $x$, but that's fine as we can use whatever letter we want for our variable. The second is that $\Gamma(s)$ involves a factor of $u^{s-1}$ and we have $u^{-1/2}$. This is easily fixed; we just write

$$u^{-\frac{1}{2}} \;=\; u^{\frac{1}{2}-\frac{1}{2}-\frac{1}{2}} \;=\; u^{\frac{1}{2}-1};$$

we just **added zero**, one of the most useful things to do in mathematics. (It takes awhile to learn how to 'do nothing' well, which is why we keep pointing this out.) Thus

$$I \;=\; \sqrt{2}\int_0^{\infty} e^{-u}u^{\frac{1}{2}-1}du \;=\; \sqrt{2}\Gamma(1/2).$$

We did it – we've found another value of $s$ that's important. Now we just need a way to find out what $\Gamma(1/2)$ equals! We could of course just go back to the standard normal's density and do the polar coordinate trick (see §14.1); however, it's possible to evaluate this directly by using the cosecant identity:

**The cosecant identity.** If $s$ is not an integer, then

$$\Gamma(s)\Gamma(1-s) \;=\; \pi\csc(\pi s) \;=\; \frac{\pi}{\sin(\pi s)}.$$

We'll give a few different proofs in Section 15.8; note it implies that $\Gamma(1/2) = \sqrt{\pi}$.

*Remark: It's worth remarking again* why *we chose to study $s = 1/2$. We had already mastered the Gamma function at the positive integers, and we needed to figure out what to study next. It's best to walk before running. Before running to $s = \sqrt{2}$ or $s = \pi$, it's a good idea to try the simplest numbers remaining. So, what's the simplest numbers that aren't positive integers? Those that are* almost *positive integers. This thought process led us to the half-integers, and I hope you see that these are natural items to investigate.*

## 15.5 The Beta Function and the Gamma Function

The **Beta function** is defined by

$$B(a,b) \;=\; \int_0^1 t^{a-1}(1-t)^{b-1}dt, \quad a,b>0.$$

Note the similarities with the Gamma function; both involve the integration variable raised to a parameter minus 1. It turns out this isn't just a coincidence or a stretch of the imagination, but rather these two functions are intimately connected by

**Fundamental Relation of the Beta Function:** For $a, b > 0$ we have

$$B(a,b) \;:=\; \int_0^1 t^{a-1}(1-t)^{b-1}dt \;=\; \frac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)}.$$

With a little bit of algebra, we can rearrange the above and find

$$\frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)}\int_0^1 t^{a-1}(1-t)^{b-1}dt \;=\; 1;$$

this means that we've discovered a new density, the density of the **Beta distribution**.

**Beta distribution**: Let $a, b > 0$. If $X$ is a random variable with the **Beta distribution** with parameters $a$ and $b$, then its density is

$$f_{a,b} \;=\; \begin{cases} \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)}t^{a-1}(1-t)^{b-1}dt & \text{if } 0 \le t \le 1 \\ 0 & \text{otherwise.} \end{cases}$$

We write $X \sim B(a,b)$.

We'll discuss this distribution in a bit more detail in §15.7. For now we'll just say briefly that it's an important family of densities as often our input is between 0 and 1, and the two parameters $a$ and $b$ give us a lot of freedom in creating 'one-hump' distributions (namely densities that go up and then go down). We plot several of these densities in Figure 15.1.

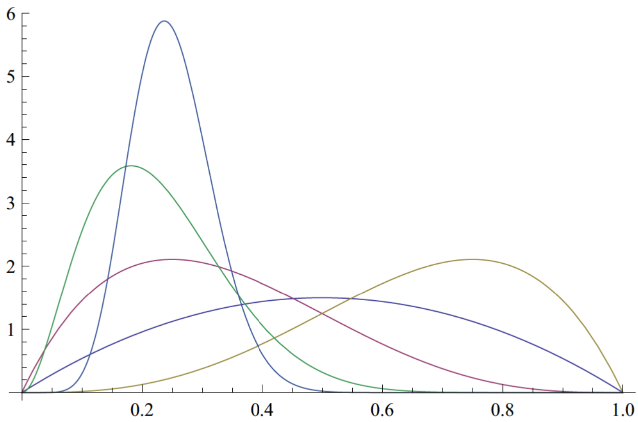


Figure 15.1: Plots of Beta densities for $(a, b)$ equal to (2, 2), (2, 4), (4, 2), (3, 10), and (10, 30).

### 15.5.1 Proof of the Fundamental Relation

We prove the fundamental relation of the Beta function. While this is an important result, remember that our purpose in doing so is to help you see how to attack problems like this. Multiplying both sides by $\Gamma(a+b)$, we see that we must prove

$$\Gamma(a)\Gamma(b) \quad \text{and} \quad \Gamma(a+b)\int_0^1 t^{a-1}(1-t)^{b-1}dt$$

are equal. There are two ways to do this; we can either work with the product of the Gamma functions, or expand the $\Gamma(a+b)$ term and combine it with the other integral.

Let's try working with the product of the Gamma functions. Note that we can use the integral representation freely, as we've assumed $a, b > 0$. We'll argue along the lines of our first proof of the cosecant identity (see §15.8.1), and we find

$$\begin{aligned}\Gamma(a)\Gamma(b) &= \int_0^\infty e^{-x}x^{a-1}dx \int_0^\infty e^{-y}y^{b-1}dy \\ &= \int_{y=0}^\infty \int_{x=0}^\infty e^{-(x+y)}x^{a-1}y^{b-1}dxdy.\end{aligned}$$

Remember, we can't change the order of integration, as that won't gain us anything as the two variables are not mixed. Our only remaining option is to change variables. We've fixed $y$ and are integrating with respect to $x$. Let's try $x = yu$ so $dx = ydu$; this at least mixes things up, and turns out to be a good choice for many problems.

We find

$$\begin{aligned}
\Gamma(a)\Gamma(b) &= \int_{y=0}^{\infty}\left[\int_{u=0}^{\infty} e^{-(1+y)u}(yu)^{a-1}y^{b-1}ydu\right]dy \\
&= \int_{y=0}^{\infty}\int_{u=0}^{\infty} y^{a+b-1}u^{a-1}e^{-(1+u)y}dudy \\
&= \int_{u=0}^{\infty}\int_{y=0}^{\infty} y^{a+b-1}u^{a-1}e^{-(1+u)y}dydu.
\end{aligned}$$

We've changed variables and then switched the order of integration. So right now we're fixing $u$ and then integrating with respect to $y$. For $u$ fixed, consider the change of variables $t = (1+u)y$. This is a good choice, and a somewhat reasonable one to try. We need to get a $\Gamma(a+b)$ somehow. For that, we want something like the exponential of the negative of one of our variables. Right now we have $e^{-(1+u)y}$, which isn't of the desired form. By letting $t = (1+u)y$, however, it now becomes $e^{-t}$. Again, what drives this change of variables is trying to get something looking like $\Gamma(a+b)$; *note how useful it is to have a sense of what the answer is!*

Anyway, if $t = (1+u)y$ then $dy = dt/(1+u)$ and our integral becomes

$$\begin{aligned}
\Gamma(a)\Gamma(b) &= \int_{u=0}^{\infty}\int_{t=0}^{\infty}\left(\frac{t}{1+u}\right)^{a+b-1} u^{a-1}e^{-t}\frac{1}{1+u}\,dtdu \\
&= \int_{u=0}^{\infty}\left(\frac{u}{1+u}\right)^{a-1}\left(\frac{1}{1+u}\right)^{b+1}\left[\int_{t=0}^{\infty} e^{-t}t^{a+b-1}dt\right]du \\
&= \Gamma(a+b)\int_{u=0}^{\infty}\left(\frac{u}{1+u}\right)^{a-1}\left(\frac{1}{1+u}\right)^{b+1}du,
\end{aligned}$$

where we used the definition of the Gamma function to replace the $t$-integral with $\Gamma(a+b)$. We're definitely making progress – we've found the $\Gamma(a+b)$ factor.

We should also comment on how we wrote the algebra above. We combined everything that was to the $a-1$ power together, and what was left was to the $b+1$ power. Again, this is a promising sign; we're trying to show that this equals $\Gamma(a+b)$ times an integral involving $x^{a-1}$ and $(1-x)^{b-1}$; it's not exactly this, but it's close. (You might be a bit worried that we have a $b+1$ and not a $b-1$ – it'll work out after yet another change of variables.) So, looking at what we have and again comparing it with where we want to go, what's the next change of variables? Let's try $\tau = \frac{u}{1+u}$, so $1-\tau = \frac{1}{1+u}$ and $d\tau = \frac{du}{(1+u)^2}$ (by the quotient rule), or $du = (1+u)^2 d\tau = \frac{d\tau}{(1-\tau)^2}$. Since $u : 0 \to \infty$, we have $\tau : 0 \to 1$,

$$\begin{aligned}
\Gamma(a)\Gamma(b) &= \Gamma(a+b)\int_0^1 \tau^{a-1}(1-\tau)^{b+1}\frac{d\tau}{(1-\tau)^2} \\
&= \Gamma(a+b)\int_0^1 \tau^{a-1}(1-\tau)^{b-1}d\tau,
\end{aligned}$$

which is what we needed to show! Why did we set $\tau$ equal to $\frac{u}{1+u}$? Remember we're trying to get the beta integral, which involves integrating the product of $\tau$ (which is less than one) to a power times one minus $\tau$ to another power. As $u$ ranges from 0 to $\infty$, $\frac{u}{1+u}$ runs from 0 to 1. This suggests that $\tau = \frac{u}{1+u}$ could be a useful change of variables.

*Remark: As always, after going through a long proof we should stop, pause, and think about what we did and why. There were several change of variables and an interchange of orders of integration. As we've already discussed why these changes of variables are reasonable, we won't rehash that here. Instead, we'll talk one more time about how useful it is to know the answer. If you can guess the answer somehow, that can provide great insight as to what to do. For this problem, knowing we wanted to find a factor of $\Gamma(a+b)$ helped us make the change of variables to fix the exponential. And knowing we wanted factors like a variable to the $a-1$ power suggested the change of variables $\tau = \frac{u}{1+u}$.*

## 15.5.2 The Fundamental Relation and $\Gamma(1/2)$

We give yet another derivation of $\Gamma(1/2)$, this time using properties of the Beta function. Taking $a = b = 1/2$ gives

$$\begin{aligned}\Gamma\left(\frac{1}{2}\right)\Gamma\left(\frac{1}{2}\right) &= \Gamma\left(\frac{1}{2}+\frac{1}{2}\right)\int_0^1 t^{1/2-1}(1-t)^{1/2-1}dt \\ &= \Gamma(1)\int_0^1 t^{-1/2}(1-t)^{-1/2}dt.\end{aligned}$$

As always, the question becomes: what's the right change of variables? If we think back to our studies of the Gamma function and the cosecant identity, we have $\Gamma(1/2)^2$ was supposed to be $\pi/\sin(\pi/2)$. This is telling us that trig functions should play a big role, so perhaps we want to do something to facilitate using trig functions or trig substitution. If so, one possibility is to take $t = u^2$. This makes the factor $(1-t)^{-1/2}$ equal to $(1-u^2)^{-1/2}$, which is ideally suited for a trig substitution.

Now for the details. We set $t = u^2$ or $u = t^{1/2}$, so $du = dt/2t^{1/2}$ or $t^{-1/2}dt = 2du$; we write it like this as we have a $t^{-1/2}dt$ already! The bounds of integration are still 0 to 1, and we find

$$\Gamma\left(\frac{1}{2}\right)^2 = \int_0^1 (1-u^2)^{-1/2}2du.$$

We now use trig substitution. Take $u = \sin\theta$, $du = \cos\theta d\theta$, so $u : 0 \to 1$ becomes $\theta : 0 \to \pi/2$ (we chose $u = \sin\theta$ over $u = \cos\theta$ as this way the bounds of integration become 0 to $\pi/2$ and not $\pi/2$ to 0, though of course either approach is fine). We now have

$$\begin{aligned}\Gamma\left(\frac{1}{2}\right)^2 &= 2\int_0^{\pi/2}(1-\sin^2\theta)^{-1/2}\cos\theta d\theta \\ &= 2\int_0^{\pi/2}\frac{\cos\theta d\theta}{(\cos^2\theta)^{1/2}} \\ &= 2\int_0^{\pi/2} d\theta \;=\; 2\cdot\frac{\pi}{2} \;=\; \pi,\end{aligned}$$

which gives us yet another way to see $\Gamma(1/2) = \sqrt{\pi}$.

## 15.6 The Normal Distribution and the Gamma Function

It would be irresponsible to cover the Gamma function without mentioning some of the other connections with the normal distribution. The three most important integrals related to the standard normal are

$$\begin{aligned} 1 &= \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}}\, e^{-x^2/2} dx \\ 0 &= \int_{-\infty}^{\infty} x \cdot \frac{1}{\sqrt{2\pi}}\, e^{-x^2/2} dx \\ 1 &= \int_{-\infty}^{\infty} (x-0)^2 \cdot \frac{1}{\sqrt{2\pi}}\, e^{-x^2/2} dx; \end{aligned}$$

a fourth useful one is the $2m^{\text{th}}$ moment,

$$\mu_{2m} = \int_{-\infty}^{\infty} x^{2m} \cdot \frac{1}{\sqrt{2\pi}}\, e^{-x^2/2} dx = (2m-1)!!,$$

where the **double factorial** means we take every other term until we reach 2 or 1 (so $5!! = 5 \cdot 3 \cdot 1$ while $6!! = 6 \cdot 4 \cdot 2$); we don't bother recording the odd moments as these are trivially zero.

The mean is easily understood – we're integrating an odd function over a symmetric region, and as our integrand decays very fast the integral converges and is zero. The other ones are a bit harder, and we had to do a lot of work to show the Gaussian's density did in fact integrate to 1, and that the variance was 1.

*If*, and this is a big if, we know the Gamma function very well, then we can immediately get any even moment. All we have to do is a little change of variables. We have

$$\mu_{2m} = \int_{-\infty}^{\infty} x^{2m} \cdot \frac{1}{\sqrt{2\pi}}\, e^{-x^2/2} dx = 2\int_{0}^{\infty} x^{2m} \cdot \frac{1}{\sqrt{2\pi}}\, e^{-x^2/2} dx.$$

How should we change variables? Looking at the definition of the Gamma function, we see that it has a term $e^{-u}$; our exponential term is $e^{-x^2/2}$. This suggests that we set $u = x^2/2$, which implies $x = (2u)^{1/2}$, so

$$x^{2m} = 2^m u^m, \quad dx = \frac{du}{\sqrt{2u}}.$$

Doing this gives

$$\mu_{2m} = \frac{2}{\sqrt{2\pi}} \int_0^{\infty} 2^m u^m e^{-u} \frac{du}{\sqrt{2u}} = \frac{2^m}{\sqrt{\pi}} \int_0^{\infty} u^{m-\frac{1}{2}} e^{-u} du.$$

We now do a nice trick: we add zero. Remember **adding zero** is one of the most powerful tools in our arsenal. We *almost* have the definition of the Gamma function, but we need to have $u^{s-1}$ and we have $u^{m-\frac{1}{2}}$. Thus we'll write

$$u^{m-\frac{1}{2}} = u^{m+\frac{1}{2}-1},$$

implying that $s = m + \frac{1}{2}$. The integral above is now just $\Gamma\left(m + \frac{1}{2}\right)$, and we end up with

$$\mu_{2m} = \frac{2^m}{\sqrt{\pi}} \Gamma\left(m + \frac{1}{2}\right).$$

We see now why there was a big *if* before; we have answers for the moments, but unless you really know a lot about the Gamma function, the answers don't look that useful. For example, if we take $m = 0$ we get the area under the curve. This is supposed to be 1; our formula tells us it's $2^0\Gamma(1/2)/\sqrt{\pi}$. It just so happens that $\Gamma(1/2) = \sqrt{\pi}$. What about the variance? That requires us to take $m = 1$ (remember we're looking at the $2m^{\text{th}}$ moment). In this case, we find $2^1\Gamma(3/2)/\sqrt{\pi}$, and (as you surely have guessed) we do have $\Gamma(3/2) = \sqrt{\pi}/2$.

The Gamma function satisfies a lot of beautiful properties. We showed in §15.2 from integrating by parts that $\Gamma(s+1) = s\Gamma(s)$, at least if $s > 0$. We gave several proofs that $\Gamma(1/2) = \sqrt{\pi}$. Using these twofacts, it's a nice exercise to show that $\Gamma(m + 1/2) = \frac{(2m-1)!!}{2^m}\Gamma(1/2)$, and thus $\mu_{2m} = (2m-1)!!$.

## 15.7 Families of Random Variables

We could easily fill up many more chapters by going through all the different, important distributions in general. Even if we restricted ourselves to distributions related to the Gamma function we would still have many more chapters to write. Instead of doing that, what we'll do instead is discuss one such distribution in greater detail (the chi-square distribution, the subject of Chapter 16), and briefly comment on a few here.

We've already talked about the Beta distribution in §15.5. We give two other **families** of densities (we'll explain the terminology in a bit).

**The Gamma and Weibull Distributions.** A random variable $X$ has the **Gamma distribution** with (positive) parameters $k$ and $\sigma$ if its density is

$$f_{k,\sigma}(x) = \begin{cases} \frac{1}{\Gamma(k)\sigma^k} x^{k-1} e^{-x/\sigma} & \text{if } x \geq 0 \\ 0 & \text{otherwise.} \end{cases}$$

We call $k$ the **shape** parameter and $\sigma$ the **scale** parameter, and write $X \sim \Gamma(k, \sigma)$ or $X \sim \text{Gamma}(k, \sigma)$.

A random variable $X$ has the **Weibull** distribution with (positive) parameters $k$ and $\sigma$ if its density is

$$f_{k,\sigma}(x) = \begin{cases} (k/\sigma)(x/\sigma)^{k-1} e^{-(x/\sigma)^k} & \text{if } x \geq 0 \\ 0 & \text{otherwise.} \end{cases}$$

We call $k$ the **shape** parameter and $\sigma$ the **scale** parameter, and write $X \sim W(k, \sigma)$.

Note these two distributions are fundamentally different from each other and from the Beta distribution. All three densities have a polynomial factor, but the Gamma and Weibull have (different) exponential factors if $k \neq 1$ and are non-zero outside $[0,1]$. What's particularly nice about these distributions is that we can vary the parameters and get different, but related, distributions. This leads us to the notion of a **family of distributions**. These are densities that are different specializations of parameters. In practice, we frequently have reason to believe a natural or mathematical phenomenon is modeled by some distribution, but with unknown values of the parameters. We then try to figure out the value of these parameters, either through mathematical analysis or through statistical inference.

One of my favorite examples of this is some work I did with the Weibull distribution to provide a theoretical justification for a formula used to predict a baseball team's winning percentage knowing just its average runs scored and allowed per game (see [Mil], or `http://www.youtube.com/watch?v=gFDly_6qOn4` for a lecture on the subject). It turns out that, for appropriate choices of parameters, a Weibull distribution does an excellent job fitting the runs scored and allowed data.

The more distributions you know, the more likely you are to make such a connection. I strongly urge you to read the paper; it's a nice application of basic probability and mathematical modeling (and some elementary statistics). I started by exploring what would happen if the runs scored or allowed were drawn from an exponential distribution (the density is proportional to $e^{-x/\sigma}$) and a Rayleigh distribution (the density is proportional to $xe^{-x^2/2\sigma^2}$). I knew about these distributions from physics, and saw I could get a nice answer, but not a perfect one. Then, inspiration hit: I noticed these two densities were of the form $x^{k-1}e^{-x^k/\lambda}$. They sat inside a family, and by choosing 'good' values for $k$ and $\lambda$ I could get both good fits with the real world data, as well as have mathematically tractable integrals. This is how I learned about the Weibull distribution.

The Weibull distribution is used in many problems involving survival analysis; there are similarly applications of the Beta and Gamma distributions (Wikipedia and a Google search will quickly yield many examples). Again, the point is to have families of distributions on your radar. The bigger your tool chest, the better job you can do modeling.

## 15.8 Appendix: Cosecant Identity Proofs

Books have entire chapters on the various identities satisfied by the Gamma function. In this section we'll concentrate on one that's particularly well-suited to our investigation of $\Gamma(1/2)$, namely the cosecant identity.

**The cosecant identity.** If $s$ is not an integer, then

$$\Gamma(s)\Gamma(1-s) \;=\; \pi\csc(\pi s) \;=\; \frac{\pi}{\sin(\pi s)}.$$

Before proving this, let's take a moment to use this to finish our study. For almost all $s$ the cosecant identity relates two values, Gamma at $s$ and Gamma at $1-s$; if you know one of these values, you know the other. Unfortunately, this means that in order for this identity to be useful, we have to know at least one of the two values. Unless, of course, we make the *very* special choice of taking $s = 1/2$. As $1/2 = 1 - 1/2$, the two values are the same, and we find

$$\Gamma(1/2)^2 \;=\; \Gamma(1/2)\Gamma(1/2) \;=\; \frac{\pi}{\sin(\pi/2)} \;=\; \pi;$$

taking square-roots gives $\Gamma(1/2) = \sqrt{\pi}$. We're quite fortunate that the very special value happens to be the value we wanted earlier!

In the following subsections I'll give various proofs of the cosecant identity. If all you care about is using it, you can of course skip this; however, if you read on you'll get some insight as to how people come up with formulas like this, and how they prove them. The arguments will become involved in places, but I'll try to point out why we're doing what we're doing, so that if you come across a situation like this in the future, a new situation where you are the first one looking at a problem and there's no handy guidebook available, you'll have some tools for your studies.

### 15.8.1 The Cosecant Identity: First Proof

*Proof of the cosecant identity.* We've seen the cosecant identity is useful; now let's see a proof. How should we try to prove this? Well, one side is $\Gamma(s)\Gamma(1-s)$. Both of these numbers can be represented as integrals. So this quantity is really a double integral. Whenever you have a double integral, you should start thinking about changing variables or changing the order of integration, or maybe even both! The point is using the integral formulations give us a starting point. This argument might not work, but it's something to try (and, for many math problems, one of the hardest things is just figuring out where to begin).

What we are about to write looks like it does what we have decided to do, but there's ***two*** subtle mistakes:

$$\begin{aligned}\Gamma(s)\Gamma(1-s) &= \int_0^\infty e^{-x}x^{s-1}dx \cdot \int_0^\infty e^{-x}x^{1-s-1}dx \\ &= \int_0^\infty e^{-x}x^{s-1} \cdot e^{-x}x^{1-s-1}dx. \end{aligned} \tag{15.1}$$

Why is this wrong? The first expression is the integral representation of $\Gamma(s)$, the second expression is the integral representation of $\Gamma(1-s)$, so their product is $\Gamma(s)\Gamma(1-s)$ and then just collect terms.... Unfortunately, ***NO!*** The problem is that we used the same dummy variable for both integrations. We can't write it as one integral – we had two integrations, each with a $dx$, and then ended up with just one $dx$. This is one of the most common mistakes students make. By not using a different letter for the variables in each integration, we accidentally combined them and went from a double integral to a single integral.

We should use two different letters, which in a fit of creativity we'll take to be $x$ and $y$. Then

$$\begin{aligned}\Gamma(s)\Gamma(1-s) &= \int_0^\infty e^{-x}x^{s-1}dx \cdot \int_0^\infty e^{-y}y^{1-s-1}dy \\ &= \int_{y=0}^\infty \int_{x=0}^\infty e^{-x}x^{s-1}e^{-y}y^{-s}dxdy.\end{aligned}$$

While the result we're gunning for, the cosecant formula, is beautiful and important, even more important (and far more useful!) is to learn how to attack problems like this. There aren't that many options for dealing with a double integral. You can integrate as given, but in this case that would be a bad idea as we would just get back the product of the Gamma functions. What else can we do? We can switch the orders of integration. Unfortunately, that too isn't any help; switching orders can only help us if the two variables are mingled in the integral, and that isn't the case now. Here, the two variables aren't seeing each other; if we switch the order of integration, we haven't really changed anything. Only one option remains: we need to change variables.

*This* is the hardest part of the proof. We have to figure out a good change of variables. Let's look at the first possible choice. We have $x^{s-1}y^{-s} = (x/y)^{s-1}y^{-1}$ (we could have written it as $(x/y)^s x^{-1}$, but since the definition of the Gamma function involves a variable to the $s-1$ power, we try this first). Perhaps a good change of variables would be to let $u = x/y$? If we do this, we fix $y$, and then for fixed $y$ we set $u = x/y$, giving $du = dx/y$. The $1/y$ is encouraging, as we had an extra $y$ earlier. This leads to

$$\Gamma(s)\Gamma(1-s) \quad = \quad \int_{y=0}^{\infty} e^{-y}\left[\int_{u=0}^{\infty} e^{-uy}u^{s-1}du\right]dy.$$

Now switching orders of integration is non-trivial, as $u$ and $y$ appear together. That gives

$$\begin{aligned}\Gamma(s)\Gamma(1-s) &= \int_{u=0}^{\infty} u^{s-1}\left[\int_{y=0}^{\infty} e^{-(u+1)y}dy\right]du \\ &= \int_{u=0}^{\infty} u^{s-1}\left[-\frac{e^{-(u+1)y}}{u+1}\Bigg|_0^{\infty}\right]dy \\ &= \int_{u=0}^{\infty} u^{s-1}\frac{1}{u+1}\,dy \;=\; \int_{u=0}^{\infty}\frac{u^{s-1}}{u+1}\,du.\end{aligned}$$

*Warning: we have to be very careful above, and make sure the interchange is justified.* Remember earlier in the chapter when we had a long discussion about the importance of making sure an integral makes sense? The integrand above is $\frac{u^{s-1}}{u+1}$. It has to decay sufficiently rapidly as $u \to \infty$ and it cannot blow up too quickly as $u \to 0$ if the integral is to be finite. If you work out what this entails, it forces $s \in (0,1)$; if $s \le 0$ then it blows up too rapidly near 0, while if $s \ge 1$ it doesn't decay fast enough at infinity.

In hindsight, this restriction isn't surprising, and in fact we should have expected it. Why? Remember earlier in the proof we remarked that there were ***two*** mistakes in (15.1); if you were really alert, you would have noticed we only mentioned ***one*** mistake! What is the missing mistake? We used the integral representation of the Gamma function. That is only valid when the argument is positive. Thus we need $s > 0$ and $1 - s > 0$; these two inequalities force $s \in (0,1)$. If you didn't catch this mistake this time, don't worry about it; just be aware of the danger in the future. This is one of the most common errors made (by both students and researchers). It's so easy to take a formula that works in some cases and accidentally use it in a place where it's not valid.

Alright. For now, let's restrict ourselves to taking $s \in (0,1)$. We leave it as an exercise to show that if the relationship holds for $s \in (0,1)$ then it holds for all $s$. *Hint:* keep using the functional equation of the Gamma function. It's easy to see how $\csc(\pi s)$ or $\sin(\pi s)$ changes if we increase $s$ by 1; the Gamma pieces follow with a bit more work.

Now we really can say

$$\Gamma(s)\Gamma(1-s) \;=\; \int_0^\infty \frac{u^{s-1}}{u+1}\,du. \tag{15.2}$$

What next? Well, we have two factors, $u^{s-1}$ and $\frac{1}{u+1}$. Note the second looks like the sum of a geometric series with ratio $-u$. To see that, we can write $\frac{1}{u+1}$ as $\frac{1}{1-(-u)}$, which is the sum of a geometric series with ratio 1 (so long as $|u| < 1$). Admittedly, this isn't going to be an obvious identification at first, but the more math you do, the more experience you gain and the easier it's to recognize patterns. We know $\sum_{n=0}^{\infty} r^n = \frac{1}{1-r}$, so all we have to do is take $r = -u$.

We must be careful – we're about to make the same mistake again, namely using a formula where it isn't applicable. It's very easy to fall into this trap. Fortunately, there's a way around it. We split the integral into two parts, the first part is when $u \in [0,1]$ and the second when $u \in [1,\infty]$. In the second part we'll then change variables by setting $v = 1/u$ and do a geometric series expansion there. **Splitting an integral** is another useful technique to master. It allows us to break a complicated problem up into simpler ones, ones where we have more results at our disposal to attack it. We need to do something like this as we're searching for a Taylor series expansion. We want to get rid of an infinity and replace it with something we know.

For the second integral, we'll make the change $v = 1/u$. This gives $dv = -du/u^2$ or $du = -v^2 dv$ (since $1/u^2 = v^2$), and the bounds of integration go from being $u : 1 \to \infty$ to $v : 1 \to 0$ (we'll then use the negative sign to switch the order of integration to the more common $v : 0 \to 1$). Continuing onward, we have

$$\begin{aligned}
\Gamma(s)\Gamma(1-s) &= \int_0^1 \frac{u^{s-1}}{u+1}\,du + \int_1^\infty \frac{u^{s-1}}{u+1}\,du \\
&= \int_0^1 \frac{u^{s-1}}{u+1}\,du - \int_1^0 \frac{(1/v)^{s-1}}{(1/v)+1}\,v^2 dv \\
&= \int_0^1 \frac{u^{s-1}}{u+1}\,du + \int_0^1 \frac{v^{-s}}{v+1}\,dv.
\end{aligned}$$

Note how similar the two expressions are (and are the same at the very special value of $s = 1/2$). We now use the geometric series formula, and then we'll interchange the integral and the sum. Everything can be justified (see §B.2) because $s \in (0,1)$, so all the integrals exist and are well behaved, giving

$$\begin{aligned}
\Gamma(s)\Gamma(1-s) &= \int_0^1 u^{s-1}\sum_{n=0}^{\infty}(-1)^n u^n du + \int_0^1 v^{-s}\sum_{m=0}^{\infty}(-1)^m v^m dv \\
&= \sum_{n=0}^{\infty}(-1)^n \int_0^1 u^{s-1+n} du + \sum_{m=0}^{\infty}(-1)^m \int_0^1 v^{m-s} dv \\
&= \sum_{n=0}^{\infty}(-1)^n \frac{u^{s+n}}{n+s}\Big|_0^1 + \sum_{m=0}^{\infty}(-1)^m \frac{v^{m+1-s}}{m+1-s}\Big|_0^1 \\
&= \sum_{n=0}^{\infty}(-1)^n \frac{1}{n+s} + \sum_{m=0}^{\infty}(-1)^m \frac{1}{m+1-s}.
\end{aligned}$$

Note we used two different letters for the different sums. While we could have used the letter $n$ twice, it's a good habit to use different letters. What happens now is that we'll adjust the counting a bit to easily combine them.

The two sums look very similar. They both look like a power of negative one divided by either $k+s$ or $k-s$. Let's rewrite both sums in terms of $k$. The first sum has one extra term, which we'll pull out. In the first sum we'll set $k=n$, while in the second we'll set $k=m+1$ (so $(-1)^m$ becomes $(-1)^{k-1}=(-1)^{k+1}$). We get

$$\begin{aligned}
\Gamma(s)\Gamma(1-s) &= \frac{1}{s} + \sum_{k=1}^{\infty}(-1)^k \frac{1}{k+s} + \sum_{k=1}^{\infty}(-1)^{k+1}\frac{1}{k-s} \\
&= \frac{1}{s} + \sum_{k=1}^{\infty}(-1)^k \left[\frac{1}{k+s} - \frac{1}{k-s}\right] \\
&= \frac{1}{s} + \sum_{k=1}^{\infty}(-1)^k \frac{2s}{k^2-s^2} \\
&= \frac{1}{s} - \sum_{k=1}^{\infty}(-1)^k \frac{-2s}{k^2-s^2}.
\end{aligned}$$

It may not look like it, but we've just finished the proof. The problem is recognizing the above is $\pi\csc(\pi s) = \pi/\sin(\pi s)$. This is typically proved in a complex analysis course; see for instance [SS2].

We can at least see it's reasonable. We're claiming

$$\frac{\pi}{\sin(\pi s)} = \frac{1}{s} - \sum_{k=1}^{\infty}(-1)^k \frac{2s}{k^2-s^2}.$$

If $s$ is an integer then $\sin(\pi s)=0$ and thus the left hand side is infinite, while exactly one of the terms on the right hand side blows up. This at least shows our answer is reasonable. Or mostly reasonable. It seems likely that our sum is $c/\sin(\pi s)$ for some $c$, but it isn't clear that $c$ equals $\pi$. Fortunately, there's even a way to get that, but it involves knowing a bit more about certain special sums. If we take $s=1/2$

then the sum becomes

$$\begin{aligned}
\frac{1}{1/2} - \sum_{k=1}^{\infty} (-1)^k \frac{1}{k^2 - (1/2)^2} &= 2 - \sum_{k=1}^{\infty} \frac{(-1)^k}{k^2 - 1/4} \\
&= 2 - \sum_{k=1}^{\infty} \frac{(-1)^k 4}{4k^2 - 1} \\
&= 2 - 4\sum_{k=1}^{\infty} \frac{(-1)^k}{(2k-1)(2k+1)} \\
&= 2 - 4\sum_{k=1}^{\infty} \frac{(-1)^k}{2}\left(\frac{1}{2k-1} - \frac{1}{2k+1}\right) \\
&= 2 + 2\left(\frac{1}{1} - \frac{1}{3}\right) - 2\left(\frac{1}{3} - \frac{1}{5}\right) + \cdots \\
&= 4\left(1 - \frac{1}{3} + \frac{1}{5} - \cdots\right).
\end{aligned}$$

As the alternating sum of the reciprocals of the odd numbers is $\pi/4$, this proves our constant $c$ is $\pi$ as claimed. This is the **Gregory-Leibniz formula** for $\pi$, and of course it's a bit of work to prove it.

*Remark: This was a long proof, but there were a lot of good ideas in it. At the end, we tried to check the reasonableness of our formula by looking at special values. This is a great idea, but it's only as useful as our ability to find special values. Knowing the Gregory-Leibniz formula allowed us to verify the claim at $s = 1/2$, which fortunately is the value we care about most!*

Here's a sketch of the proof of the Gregory-Leibniz formula. Use the derivative of $\arctan(x)$ is $\frac{1}{1+x^2}$ to get $\int_0^1 \frac{1}{1+x^2}dx = \frac{\pi}{4}$. Write $1+x^2$ as $1-(-x^2)$, and expand $\frac{1}{1+x^2}$ using the geometric series formula with $r = -x^2$. Justify interchanging the sum and the integral, integrate term by term, and smile!

## 15.8.2 The Cosecant Identity: Second Proof

We already have a proof of the cosecant identity for the Gamma function – why do we need another? For us, the main reason is educational. The goal of this book is not to teach you how to answer one specific problem at one moment in your life, but rather to give you the tools to solve a variety of new problems whenever you encounter them. Because of that, it's worth seeing multiple proofs as different approaches emphasize different aspects of the problem, or generalize better for other questions.

Let's go back to the set-up. We had $s \in (0,1)$ and

$$\begin{aligned}
\Gamma(s)\Gamma(1-s) &= \int_0^{\infty} e^{-x}x^{s-1}dx \cdot \int_0^{\infty} e^{-y}y^{1-s-1}dy \\
&= \int_{y=0}^{\infty}\int_{x=0}^{\infty} e^{-x}x^{s-1}e^{-y}y^{-s}dxdy.
\end{aligned}$$

We've already talked about what our options are. We can't integrate it as is, or we'll just get back the two Gamma functions. We can't change the order of integration, as the $x$ and $y$ variables are not mingled and thus changing the order of integration won't really change the problem. The only thing left to do is change variables.

Before we set $u = x/y$. We were led to this because we saw $x^{s-1}y^{-s} = (x/y)^{s-1}y^{-1}$, and thus it's not unreasonable to set $u = x/y$. Are there any other 'good' choices for a change of variable? There is, but it's not surprising if you don't see it. It's our old friend, polar coordinates.

It should seem a little strange to use polar coordinates here. After all, we use those for problems with radial and angular symmetry. We use them for integrating over circular regions. **NONE** of this is happening here! That said, we think a good case can be made for trying the **polar coordinate trick**.

- First, we don't know that many change of variables; we do know polar coordinates, so we might as well try it.

- Second, we're trying to show the answer is $\pi \csc(\pi s) = \pi/\sin(\pi s)$. The answer involves the sine function, so perhaps this suggests we should try polar coordinates.

At the end of the day, a method either works or it doesn't. We hope the above at least motivates why we're trying this here, and can provide guidance for you in the future.

Recall for polar coordinates we have the following relations:

$$x = r\cos\theta, \quad y = r\sin\theta, \quad dxdy = rdrd\theta.$$

What are the bounds of integration? We're integrating over the upper right quadrant, $x, y : 0 \to \infty$. In polar coordinates it becomes $r : 0 \to \infty$ and $\theta : 0 \to \pi/2$. Our integral now becomes

$$\begin{aligned}
\Gamma(s)\Gamma(1-s) &= \int_{\theta=0}^{\pi/2}\int_{r=0}^{\infty} e^{-r\cos\theta}(r\cos\theta)^{s-1}e^{-r\sin\theta}(r\sin\theta)^{-s}rdrd\theta \\
&= \int_{\theta=0}^{\pi/2}\int_{r=0}^{\infty} e^{-r(\cos\theta+\sin\theta)}\left(\frac{\cos\theta}{\sin\theta}\right)^{s-1}\frac{1}{\sin\theta}drd\theta \\
&= \int_{\theta=0}^{\pi/2}\left(\frac{\cos\theta}{\sin\theta}\right)^{s-1}\frac{1}{\sin\theta}\left[\int_{r=0}^{\infty} e^{-r(\cos\theta+\sin\theta)}dr\right]d\theta \\
&= \int_{\theta=0}^{\pi/2}\left(\frac{\cos\theta}{\sin\theta}\right)^{s-1}\frac{1}{\sin\theta}\left[-\frac{e^{-r(\cos\theta+\sin\theta)}}{\cos\theta+\sin\theta}\right]_0^{\infty}d\theta \\
&= \int_{\theta=0}^{\pi/2}\left(\frac{\cos\theta}{\sin\theta}\right)^{s-1}\frac{1}{\sin\theta}\frac{1}{\cos\theta+\sin\theta}\,d\theta.
\end{aligned}$$

It doesn't look like we've made much progress, but we're just one little change of variables away from a great simplification. Note that a lot of the integrand only depends on $\cos\theta/\sin\theta = \mathrm{ctan}\theta$ (the cotangent of $\theta$). If we do make the change of variables $u = \mathrm{ctan}\theta$ then $du = -\csc^2\theta = -1/\sin^2\theta$; if you don't remember this formula, you can get it by the quotient rule:

$$\begin{aligned}\text{ctan}'(\theta) = \left(\frac{\cos\theta}{\sin\theta}\right)' &= \frac{\cos'\theta\sin\theta - \sin'\theta\cos\theta}{\sin^2\theta} \\ &= \frac{-\sin^2\theta - \cos^2\theta}{\sin^2\theta} = -\frac{1}{\sin^2\theta}.\end{aligned}$$

Now things are looking really promising; our proposed change of variables needs a $1/\sin^2\theta$, and we already have a $1/\sin\theta$ in the integrand. We get the other by writing

$$\frac{1}{\cos\theta+\sin\theta} = \frac{1}{\sin\theta}\frac{1}{(\cos\theta/\sin\theta)+1} = \frac{1}{\sin\theta}\frac{1}{\text{ctan}\theta+1}.$$

All that remains is to find the bounds of integration. If $u = \text{ctan}\theta = \cos\theta/\sin\theta$, then $\theta : 0 \to \pi/2$ corresponds to $u : \infty \to 0$ (don't worry that we're integrating from infinity to zero – we have a minus sign floating around, and that will flip the order of integration).

Putting all the pieces together, we find

$$\begin{aligned}\Gamma(s)\Gamma(1-s) &= \int_{\theta=0}^{\pi/2} \frac{\text{ctan}^{s-1}\theta}{\text{ctan}\theta+1}\frac{d\theta}{\sin^2\theta} \\ &= \int_{u=\infty}^{0} \frac{u^{s-1}}{u+1}(-du) = \int_0^\infty \frac{u^{s-1}}{u+1}du.\end{aligned}$$

This integral should look familiar – it's exactly the integral we saw in the previous section, in equation (15.2). Thus from here onward we can just follow the steps in that section.

*Remark: A lot of students freeze when they first see a difficult math problem. Why varies from student to student, but a common refrain is:* "I didn't know where to start." *For those who feel that way, this should be comforting. There are (at least!) two different change of variables we can do, both leading to a solution for the problem. As you continue in math you'll see again and again that there are many different approaches you can take. Don't be afraid to try something. Work with it for awhile and see how it goes. If it isn't promising you can always backtrack and try something else.*

## 15.8.3 The Cosecant Identity: Special Case $s = 1/2$

While obviously we want to prove the cosecant formula for arbitrary $s$, *the* most important choice of $s$ is clearly $s = 1/2$. We need $\Gamma(1/2)$ in order to write down the density functions for normal distributions, and to compute its moments. Thus, while it would be nice to have a formula for any $s$, it's still cause for celebration if we can handle just $s = 1/2$.

Remember in (15.2) that we showed

$$\Gamma(s)\Gamma(1-s) = \int_0^\infty \frac{u^{s-1}}{u+1}\,du.$$

Taking $s = 1/2$ gives

$$\Gamma(1/2)^2 \;=\; \int_0^\infty \frac{u^{-1/2}}{1+u}\, du.$$

We're going to solve this with a highly non-obvious change of variable. Let's state it first, see how it works, and then discuss why this is a reasonable thing to try. Here it is: take $u = z^2$, so $z = u^{1/2}$ and $dz = du/2\sqrt{u}$. Note how beautifully this fits with our integral. We have a $u^{-1/2}du$ term already, which becomes $2dz$. Substituting gives

$$\Gamma(1/2)^2 \;=\; \int_0^\infty \frac{2dz}{1+z^2} \;=\; 2\int_0^\infty \frac{dz}{1+z^2}.$$

Looking at this integral, you should think of the trigonometric substitutions from calculus. Whenever you see $1-z^2$ you should try $z = \sin\theta$ or $z = \cos\theta$; when you see $1+z^2$ you should try $z = \tan\theta$. Let's make this change of variables. The reason it's so useful is the Pythagorean formula

$$\sin^2\theta + \cos^2\theta \;=\; 1$$

becomes, on dividing both sides by $\cos^2\theta$,

$$\tan^2\theta + 1 \;=\; \frac{1}{\cos^2\theta} \;=\; \sec^2\theta.$$

Letting $z = \tan\theta$ means we replace $1+z^2$ with $\sec^2\theta$. Further, $dz = \sec^2\theta d\theta$ (if you don't remember this, just use the quotient rule applied to $\tan\theta = \sin\theta/\cos\theta$). As $z : 0 \to \infty$, we have $\theta : 0 \to \pi/2$. Collecting everything gives

$$\begin{aligned}\Gamma(1/2)^2 \;&=\; 2\int_0^{\pi/2} \frac{1}{\sec^2\theta}\sec^2\theta d\theta \\ &=\; 2\int_0^{\pi/2} d\theta \;=\; 2\frac{\pi}{2} \;=\; \pi,\end{aligned}$$

and if $\Gamma(1/2)^2 = \pi$ then $\Gamma(1/2) = \sqrt{\pi}$ as claimed!

And there we have it: a correct, elementary proof that $\Gamma(1/2) = \sqrt{\pi}$. You should be able to follow the proof line by line, but that's not the point of mathematics. The point is to *see* why the author is choosing to do these steps so that you too could create a proof like this.

There were two changes of variables. The first was replacing $u$ with $z^2$, and the second was replacing $z$ with $\tan\theta$. The two changes are related. How can anyone be expected to think of these? To be honest, when writing this chapter I had to consult my notes from teaching a similar course several years ago. I remembered that somehow tangents came into the problem, but couldn't remember the exact trick I used so long ago. It's not easy. It takes time, but the more you do, the more patterns you can detect. We have a $1+u$ in the denominator; we know how to handle terms such as $1+z^2$ through trig substitution. As the cosecant identity involves trig functions, that suggests this could be a fruitful avenue to explore. It's not a guarantee, but we might as well try it and see where it leads.

Flush with our success, the most natural thing to try next are these substitutions for general $s$. If we do this, we would find

$$\begin{aligned}\Gamma(s)\Gamma(1-s) &= \int_0^\infty \frac{z^{2s-2}}{1+z^2}\, 2z dz\\ &= 2\int_0^\infty \frac{z^{2s-1}}{1+z^2}\, dz\\ &= 2\int_0^{\pi/2} \frac{\tan^{2s-1}\theta}{\sec^2\theta}\, \sec^2\theta d\theta\\ &= 2\int_0^{\pi/2} \tan^{2s-1}\theta d\theta.\end{aligned}$$

We now see how special $s = 1/2$ is. For this, and only for this, value does the integrand collapse to just being the constant function 1, which is easily integrated. Any other choice of $s$ forces us to have to find integrals of powers of the tangent function, which is no easy task! Formulas do exist; for example,

$$\begin{aligned}&\int \tan^{1/2}\theta d\theta\\ &= \frac{1}{2\sqrt{2}}\Big[-2\arctan\left(1-\sqrt{2}\sqrt{\tan\theta}\right) + 2\arctan\left(1+\sqrt{2}\sqrt{\tan\theta}\right)\\ &\quad + \log\left(1-\sqrt{2}\sqrt{\tan\theta}+\tan\theta\right) - \log\left(1+\sqrt{2}\sqrt{\tan\theta}+\tan\theta\right)\Big].\end{aligned}$$

*Remark: If we remember that the derivative of* $\arctan(z)$ *is* $\frac{1}{1+z^2}$*, we can avoid the* $z = \tan\theta$ *substitution and directly evaluate* $\int_0^\infty \frac{1}{1+z^2}dz$ *as* $\arctan(\infty) - \arctan(0) = \pi/2$*. One of the best ways to see this is to note that if* $f(g(x)) = x$*, then by the chain rule* $f'(g(x))g'(x) = 1$*, or* $g'(x) = 1/f'(g(x))$*. Use this relation with* $g(x) = \arctan(x)$ *and* $f(x) = \tan(x)$ *to find the derivative of* $\arctan(x)$*. The difficult part is drawing the correct right triangle to get the nice expression for* $f'(g(x))$*.*

## 15.9 Additional Problems

**Problem 15.9.1** *Find* $\Gamma(3/2)$

**Problem 15.9.2** *Find* $\Gamma(-1/2)$

**Problem 15.9.3** *Prove* $\Gamma(m+1/2) = \frac{(2m-1)!!}{2^m}\Gamma(1/2)$*, and thus* $\mu_{2m} = (2m-1)!!$.

**Problem 15.9.4** *Find* $\Gamma(1/2-m)$ *for positive interger* $m$*.*

**Problem 15.9.5** *Prove that if the relationship* $\Gamma(s)\Gamma(1-s) = \pi\csc(\pi s)$ *holds for* $s \in (0,1)$ *then it holds for all* $s$ *(or at least all* $s$ *that are not integers, as if* $s$ *is an integer then we have to interpret the equality among two infinities).*

**Problem 15.9.6** *For what values of* $a$ *and* $b$ *is the Beta distribution symmetric about its mean?*

**Problem 15.9.7** *Find the mean of the Beta distribution.*

**Problem 15.9.8** *Find the mean of the Weibull distribution.*

**Problem 15.9.9** *Find the mean of the Gamma distribution with parameters $k$ and $\sigma$.*

**Problem 15.9.10** *Find the variance of the Gamma distribution with parameters $k$ and $\sigma$.*

**Problem 15.9.11** *Find $E[X^n]$ for $X$ a Gamma variable with parameters $k$ and $\sigma$.*

**Problem 15.9.12** *Comment on the relationship between the Gamma distribution and the Erlang distribution.*

**Problem 15.9.13** *For what $x$ does the Gamma distribution take on its maximum value given parameters $k$ and $\theta$?*

**Problem 15.9.14** *Justify that the kth smallest variable in a list of $n$ identical, independent random variables $X_1, X_2, \ldots X_k$ each with density $f$ and cdf $F$ is*

$$nf(x)\binom{n-1}{k-1}(F(x))^{k-1}(1-F(x))^{n-k}.$$

**Problem 15.9.15** *Use the formula from the previous problem to show that the $k^{th}$ smallest of a set of $n$ uniform variables on (0,1) can be modeled with a Beta distribution. Find the appropriate parameters for the Beta distribution.*

**Problem 15.9.16** *The incomplete lower gamma function is defined by*

$$\gamma(s,x) \;=\; \int_0^x e^{-x}x^{s-1}dx.$$

*Find a recurrence relation relating $\gamma(s,x)$to $\gamma(s-1,x)$.*

**Problem 15.9.17** *Prove the* ***Gregory-Leibniz formula*** *for $\pi$ by evaluating $\int_0^1 \frac{dx}{1+x^2}$ two ways: (1) use the geometric series formula to expand, interchange the summation and the integral, and integrate term by term (all this must be justified); (2) use the derivative of $\arctan x$ is $1/(1+x^2)$.*

**Problem 15.9.18** *Due to the Gamma distributions relationship to the exponential distribution and the exponential distributions 'memorylessness" Gamma distributions are useful in measuring web server traffic. Particularly, the time at which the $k^{th}$ person connects can be modeled by a Gamma distribution with shape parameter $k$ and some scale parameter $\sigma$. If $\sigma = 1/10$, find the probability that the 100th person connects within an hour of the start time.*

**Problem 15.9.19** *Wind speed is well approximated by a Weibull distribution. Find the probability of the wind speed being over 20 if the wind speed in a given area is a Weibull variable with shape parameter 2 and scale parameter 10.*

**Problem 15.9.20** *Use Mathematica to plot the Beta distribution with parameters $a = 2, b = 3$. Shade in the area under the curve corresponding to $P(.2 < X < .6)$. (There are several ways to do this, some look better than others. Play around with it some.)*