1 The beta distribution

The beta distribution is useful in a number of applications. Specifically, it is often used in Bayesian inference.

1.1 Definition

The pdf of the beta distribution has the interval $[0, 1]$ as closed support. The pdf f depends on two parameters, α and β:

$$f(x) \propto x^{\alpha-1}(1-x)^{\beta-1} \tag{1}$$

Here, the alpha indicates proportionality; there will also be a normalization constant that will depend on α and β:

$$f(x) = \frac{1}{B(\alpha,\beta)} x^{\alpha-1}(1-x)^{\beta-1} \tag{2}$$

Here, $B(\alpha, \beta)$ is known as the *beta function*.

1.2 The beta and gamma function relationship

It turns out that the beta function can be expressed in terms of the gamma function, which is defined as

$$\Gamma(x) = \int_0^\infty e^{-u} u^{x-1}\, du \tag{3}$$

To see this connection, consider the following expression:

$$\begin{aligned}
\Gamma(x)\Gamma(y) &= \int_0^\infty e^{-u} u^{x-1}\, du \int_0^\infty e^{-v} v^{y-1}\, dv && (4)\\
&= \int_0^\infty \int_0^\infty e^{-(u+v)} u^{x-1} v^{y-1}\, du\, dv && (5)
\end{aligned}$$

Now, consider the following change of variables:

$$(u, v) \mapsto (z, t), \quad u = zt, \quad v = z(1-t) \tag{6}$$

Adding the definitions of the two new variables, we get $z = u + v$. From the first definition, we now get $t = \frac{u}{z} = \frac{u}{u+v}$. So z can take on any value, but t must be between 0 and 1. The corresponding Jacobian matrix is:

$$J = \begin{pmatrix} \frac{\partial u}{\partial z} & \frac{\partial v}{\partial z} \\ \frac{\partial u}{\partial t} & \frac{\partial v}{\partial t} \end{pmatrix} = \begin{pmatrix} t & 1-t \\ z & -z \end{pmatrix} \tag{7}$$

The determinant of J is $t(-z) - z(1-t) = -z$, so $|\det J| = z$. Now the integral reads:

$$\int_0^1 \int_0^\infty e^{-z}(zt)^{x-1}(z(1-t))^{y-1} \cdot \underbrace{z}_{|\det J|} dz\ dt \tag{8}$$

Now rearrange to get:

$$\int_0^1 t^{x-1}(1-y)^{y-1}\ dt \int_0^\infty e^{-z} z^{x+y-1}\ dz = B(x,y)\Gamma(x+y) \tag{9}$$

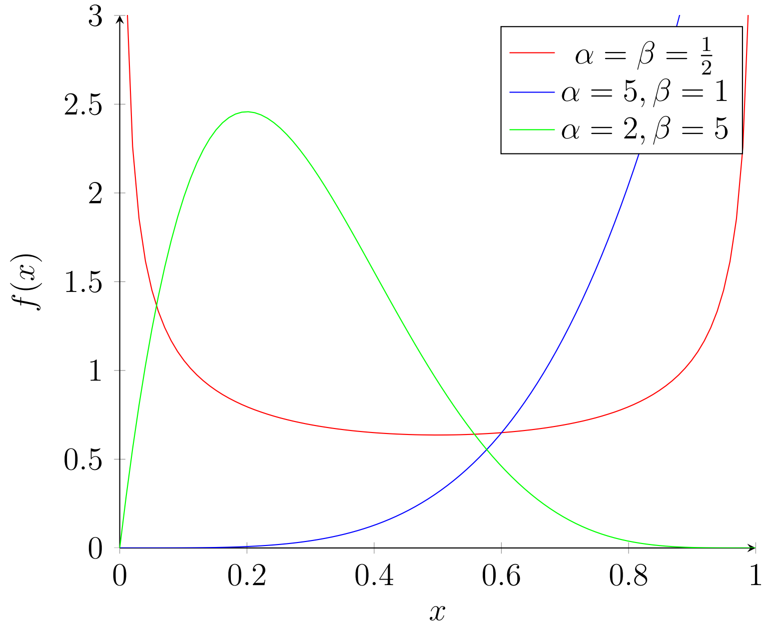
This means that:

$$\Gamma(x)\Gamma(y) = B(x,y)\Gamma(x+y) \Leftrightarrow B(x,y) = \frac{\Gamma(x)\Gamma(y)}{\Gamma(x+y)} \tag{10}$$

In other words, the pdf for a beta distribution with parameters α and β is:

$$f(x) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} x^{\alpha-1}(1-x)^{\beta-1} \tag{11}$$

A few of the possible pdf's are graphed below:



1.3 Moments

Consider the raw n'th moment of a beta distribution with parameters α and β:

$$\mu_n = \int_0^1 x^n f(x)\ dx = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} \int_0^1 x^n x^{\alpha-1}(1-x)^{\beta-1} \tag{12}$$

But this integral is simply the reciprocal of the normalization constant for a beta distribution with parameters $\alpha' = \alpha + n$ and β. In other words, the moment is:

$$\mu_n = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} \frac{\Gamma(\alpha+n)\Gamma(\beta)}{\Gamma(\alpha+\beta+n)} = \frac{\Gamma(\alpha+\beta)\Gamma(\alpha+n)}{\Gamma(\alpha)\Gamma(\alpha+\beta+n)} \tag{13}$$

1.3.1 Mean

When $n = 1$, this becomes:

$$\mu = \frac{\Gamma(\alpha+\beta)\Gamma(\alpha+1)}{\Gamma(\alpha)\Gamma(\alpha+\beta+1)} \tag{14}$$

But we know that $\Gamma(x+1) = x\Gamma(x)$, so:

$$\mu = \frac{\Gamma(\alpha+\beta)\alpha\Gamma(\alpha)}{\Gamma(\alpha)(\alpha+\beta)\Gamma(\alpha+\beta)} = \frac{\alpha}{\alpha+\beta} \tag{15}$$

1.3.2 Second moment and variance

When $n = 2$, we get:

$$\mu_2 = \frac{\Gamma(\alpha+\beta)\Gamma(\alpha+2)}{\Gamma(\alpha)\Gamma(\alpha+\beta+2)} \tag{16}$$

Now use the functional equation twice:

$$\Gamma(x+2) = (x+1)\Gamma(x+1) = x(x+1)\Gamma(x) \tag{17}$$

So the second, raw moment can be written:

$$\mu_2 = \frac{\Gamma(\alpha+\beta)\alpha(\alpha+1)\Gamma(\alpha)}{\Gamma(\alpha)(\alpha+\beta)(\alpha+\beta+1)\Gamma(\alpha+\beta)} = \frac{\alpha(\alpha+1)}{(\alpha+\beta)(\alpha+\beta+1)} \tag{18}$$

The variance is equal to $\mu_2 - \mu^2$:

$$\sigma^2 = \frac{\alpha(\alpha+1)}{(\alpha+\beta)(\alpha+\beta+1)} - \left(\frac{\alpha}{\alpha+\beta}\right)^2 \tag{19}$$

Getting a common denominator:

$$\sigma^2 = \frac{\alpha(\alpha+1)(\alpha+\beta) - \alpha^2(\alpha+\beta+1)}{(\alpha+\beta)^2(\alpha+\beta+1)} \tag{20}$$

Expand the numerator:

$$\alpha^3 + \alpha^2\beta + \alpha^2 + \alpha\beta - \alpha^3 - \alpha^2\beta - \alpha^2 = \alpha\beta \tag{21}$$

So the variance is:

$$\sigma^2 = \frac{\alpha\beta}{(\alpha+\beta)^2(\alpha+\beta+1)} \tag{22}$$