The Beta Distribution

We say that a continuous random variable $X$ has a Beta distribution with parameters $\alpha > 0$ and $\beta > 0$ if it has pdf:

$$f_X(x) = \begin{cases} \frac{1}{B(\alpha, \beta)} x^{\alpha-1} (1-x)^{\beta-1} & \text{for } 0 < x < 1 \\ 0 & \text{otherwise} \end{cases}$$

where $B(\alpha, \beta) = \int_0^1 x^{\alpha-1} (1-x)^{\beta-1} \, dx$ is called the Beta function and is defined for $\alpha > 0$ and $\beta > 0$.

On the homework you show that

$$\Gamma(\alpha) \Gamma(\beta) = \Gamma(\alpha + \beta) B(\alpha, \beta).$$

This gives $B(\alpha, \beta) = \frac{\Gamma(\alpha) \Gamma(\beta)}{\Gamma(\alpha + \beta)}$. Thus, the Beta distribution with parameters $\alpha$ and $\beta$ (written $\text{Beta}(\alpha, \beta)$) has pdf that can be written as

$$f_X(x) = \begin{cases} \frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha) \Gamma(\beta)} x^{\alpha-1} (1-x)^{\beta-1} & \text{for } 0 < x < 1 \\ 0 & \text{otherwise} \end{cases}$$

Moments of the $\text{Beta}(\alpha, \beta)$ distribution

For $k$ a positive integer, the $k^{th}$ moment is

$$E[X^k] = \int_0^1 x^k \frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha) \Gamma(\beta)} x^{\alpha-1} (1-x)^{\beta-1} \, dx$$

$$= \frac{\Gamma(\alpha + \beta)}{\Gamma(\alpha) \Gamma(\beta)} \Gamma(k+\alpha) \Gamma(k+\beta) \int_0^1 x^{k+\alpha-1} (1-x)^{\beta-1} \, dx$$

proportional to a $\text{Beta}(k+\alpha, \beta)$ density

$$= \frac{\Gamma(\alpha + \beta) \Gamma(k+\alpha) \Gamma(\beta)}{\Gamma(\alpha) \Gamma(\beta) \Gamma(k+\alpha+\beta)} \int_0^1 \frac{\Gamma(k+\alpha+\beta)}{\Gamma(k+\alpha+\beta) \Gamma(\beta)} x^{k+\alpha-1} (1-x)^{\beta-1} \, dx$$

$$= \frac{\Gamma(\alpha + \beta) \Gamma(k+\alpha)}{\Gamma(k+\alpha+\beta) \Gamma(\alpha)} \frac{\Gamma(k+\alpha+\beta)}{\Gamma(k+\alpha+\beta) \Gamma(\alpha+\beta)} \frac{\Gamma(k+\alpha)}{\Gamma(k+\alpha-1) \ldots \Gamma(\alpha+\beta)}$$

$$= \frac{(\alpha+\beta-1) \ldots (\alpha)}{(k+\alpha+\beta-1) \ldots (\alpha+\beta)}$$
For $k = 1$
\[
E[X] = \frac{\alpha}{\alpha + \beta}
\]

For $k = 2$
\[
E[X^2] = \frac{(\alpha + 1)\alpha}{(\alpha + \beta + 1)(\alpha + \beta)}
\]

Then
\[
\text{Var}(X) = E[X^2] - E[X]^2 = \frac{(\alpha + 1)\alpha}{(\alpha + \beta + 1)(\alpha + \beta)} - \frac{\alpha^2}{(\alpha + \beta)^2}
\]
\[
= \frac{(\alpha^2 + \alpha)(\alpha + \beta) - (\alpha^3 + \alpha^2 \beta + \alpha^2)}{(\alpha + \beta + 1)(\alpha + \beta)}
\]
\[
= \frac{\alpha^3 + \alpha^2 \beta + \alpha^2 + \alpha \beta - \alpha^3 - \alpha^2 \beta - \alpha^2}{(\alpha + \beta + 1)(\alpha + \beta)^2}
\]
\[
\text{Var}(X) = \frac{\alpha \beta}{(\alpha + \beta + 1)(\alpha + \beta)^2}
\]

**Special Case**

If $\alpha = 1$ and $\beta = 1$, then $f_x(x) = \begin{cases} 1 & \text{for } 0 < x < 1 \\ 0 & \text{otherwise} \end{cases}$

I.e., The Beta$(1, 1)$ distribution is the Uniform$(0, 1)$ distribution.

**Different shapes of the density**

1. If $\alpha > 1$ and $\beta > 1$,

   $f_x(x)$ goes to $0$ or $x \to 0$ or $x \to 1$.
2. If $\alpha < 1$ and $\beta < 1$

If $\alpha < 1$ and $\beta < 1$

3. If $\alpha < 1$ and $\beta > 1$

If $\alpha < 1$ and $\beta > 1$

4. If $\alpha > 1$ and $\beta < 1$

If $\alpha > 1$ and $\beta < 1$

5. If $\alpha = \beta$ then $f_x(x)$ is symmetric about $\frac{1}{2}$.

Linear transformations of $X$: Let $a < b$ be constants, and let $Y = a + (b-a)X$. The inverse is $X = \frac{Y-a}{b-a}$ and the derivative of the inverse transformation is $\frac{1}{b-a}$. Then the pdf of $Y$ is $f_Y(y) = f_x\left(\frac{Y-a}{b-a}\right) \frac{1}{b-a}$

$$=egin{cases} \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} \left(\frac{y-a}{b-a}\right)^{\alpha-1} \left(\frac{b-y}{b-a}\right)^{\beta-1} \frac{1}{b-a} & 0 < Y < b \\ 0 & \text{otherwise} \end{cases}$$

So the Beta distribution (which is on $(0,1)$) is the basis for a large family of distributions on arbitrary finite intervals.
Example: Suppose the arrival time of a friend is between 1pm and 3pm and has density proportional to \((y-1)(3-y)^4\). What is the expected arrival time of the friend?

Sol: Comparing to the translated Beta distribution we can see that the density of \(Y\) (here \(Y\) denotes the arrival time) is:

\[
f_Y(y) = \begin{cases} \frac{\Gamma(7)}{\Gamma(2)\Gamma(5)} \frac{(y-1)(3-y)^4}{2^6} & 1 \leq y \leq 3 \\ 0 & \text{otherwise} \end{cases}
\]

and we know that \(Y = 1 + 2X\), where \(X\) has a \(\text{Beta}(2,5)\) distribution. Then \(E[Y] = 1 + 2E[X] = 1 + \frac{2}{2+5} = 1 + \frac{4}{7} = \frac{11}{7}\)