

8. The Beta Distribution

In this section, we will study a two-parameter family of distributions that has special importance in probability and statistics.

The Beta Function

The **beta function**, first introduced by **Leonhard Euler**, is defined as follows:

$$B(a, b) = \int_0^1 u^{a-1} (1-u)^{b-1} du; \quad a > 0, b > 0$$

1. Show that the beta function is well-defined, that is, $B(a, b) < \infty$ for any $a > 0$ and $b > 0$, using these steps:

- Break the integral into two parts, from 0 to $\frac{1}{2}$ and from $\frac{1}{2}$ to 1.
- If $0 < a < 1$, the integral is improper at $u = 0$, but $(1-u)^{b-1}$ is bounded on $(0, \frac{1}{2}]$.
- If $0 < b < 1$, the integral is improper at $u = 1$, but u^{a-1} is bounded on $[\frac{1}{2}, 1)$.

2. Show that

- $B(a, b) = B(b, a)$
- $B(a, 1) = \frac{1}{a}$

3. Show that the beta function can be written in terms of the **gamma function** as follows:

$$B(a, b) = \frac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)}, \quad a > 0, b > 0$$

- Express $\Gamma(a+b)B(a, b)$ as a double integral with respect to x and y where $x > 0$ and $0 < y < 1$.
- Use the transformation $w = xy$, $z = x - xy$ and the change of variables theorem for multiple integrals.
- The transformation maps the (x, y) region in (a) one-to-one and onto the region $z > 0$, $w > 0$.
- The Jacobian of the inverse transformation has magnitude $\frac{1}{z+w}$.
- Show that the transformed integral is $\Gamma(a)\Gamma(b)$.

4. Show that if j and k are positive integers, then

$$B(j, k) = \frac{(j-1)!(k-1)!}{((j+k)-1)!}$$

Let's generalize this result. First, recall the generalized permutation formula from our study of [combinatorial structures](#): for $a \in \mathbb{R}$, $s \in \mathbb{R}$, and $j \in \mathbb{N}$, we defined

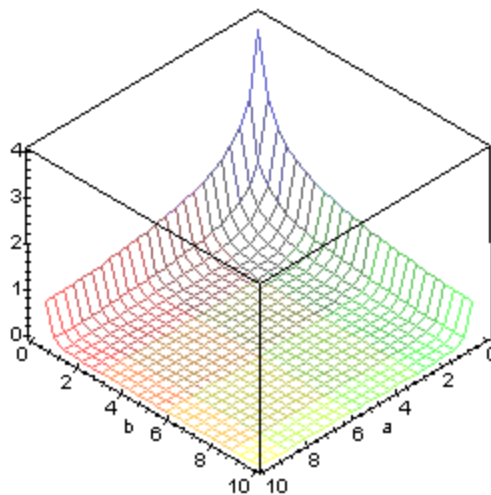
$$a^{(s,j)} = a(a+s)(a+2s)\cdots(a+(j-1)s)$$

5. Suppose that $a > 0$, $b > 0$, $j \in \mathbb{N}$, and $k \in \mathbb{N}$. Show that

$$\frac{B(a+j, b+k)}{B(a, b)} = \frac{a^{(1,j)} b^{(1,k)}}{(a+b)^{(1,j+k)}}$$

6. Show that $B\left(\frac{1}{2}, \frac{1}{2}\right) = \pi$

A graph of $B(a, b)$ on the square $0 < a < 10$, $0 < b < 10$ is shown below.



The Density Function

7. Show that the function f below is a [probability density function](#) for any $a > 0$, $b > 0$

$$f(u) = \frac{1}{B(a, b)} u^{a-1} (1-u)^{b-1}, \quad 0 < u < 1$$

The distribution with the density in the previous exercise is called the **beta distribution** with **left parameter** a and **right parameter** b . The beta distribution is useful for modeling random probabilities and proportions, particularly in the context of Bayesian analysis. The distribution has two parameters and yet a rich variety of shapes:

8. Sketch the graph of the beta probability density function. Note the qualitative differences in the shape of the density for the following parameter ranges:

- $0 < a < 1$, $0 < b < 1$
- $a = 1$, $b = 1$. This special case corresponds to the uniform distribution.
- $a = 1$, $0 < b < 1$

- d. $0 < a < 1, b = 1$
- e. $0 < a < 1, b > 1$
- f. $a > 1, 0 < b < 1$
- g. $a = 1, b > 1$
- h. $a > 1, b = 1$
- i. $a > 1, b > 1$. Show that in this special case, the mode occurs at $u = \frac{a-1}{(a+b)-2}$.

9. In the **random variable experiment**, select the beta distribution. Set the parameters to values in each of the ranges of the previous exercise. In each case, note the shape of the beta density function. Run the simulation 1000 times with an update frequency of 10 and note the apparent convergence of the empirical density function to the true density function.

The special case $a = \frac{1}{2}, b = \frac{1}{2}$ is the **arcsine distribution** (the name will be explained below). Thus, the probability density function of the arcsine distribution is

$$f(x) = \frac{1}{\pi \sqrt{x(1-x)}}, \quad 0 < x < 1$$

Distribution Function

In some special cases, the beta **distribution function** F and its inverse, the **quantile function** F^{-1} , can be computed in closed form. In the following exercises, a denotes the left parameter and b the right parameter, as usual.

10. Suppose that $a > 0$ and $b = 1$. Show that

- a. $F(x) = x^a, 0 < x < 1$
- b. $F^{-1}(p) = p^{1/a}, 0 < p < 1$

11. Suppose that $a = 1$ and that $b > 0$. Show that

- a. $F(x) = 1 - (1-x)^b, 0 < x < 1$
- b. $F^{-1}(p) = 1 - (1-p)^{1/b}, 0 < p < 1$

12. Suppose that $a = b = \frac{1}{2}$ (the arcsine distribution). Show that

- a. $F(x) = \frac{2}{\pi} \arcsin(\sqrt{x})$ for $0 < x < 1$
- b. $F^{-1}(p) = \sin^2\left(\frac{\pi}{2} p\right)$ for $0 < p < 1$

In general, there is an interesting relationship between the distribution functions of the beta distribution and the [binomial distribution](#).

13. Fix $n \in \mathbb{N}_+$. Let F_p denote the binomial distribution function with trial parameter n and success parameter $p \in (0, 1)$, and let G_k denote the beta distribution function with left parameter $n - k + 1$ and right parameter k , where $k \in \{1, 2, \dots, n + 1\}$. Show that

$$F_p(k - 1) = G_k(1 - p)$$

Hint: Express $G_k(1 - p)$ as an integral of the beta probability density function, and then integrate by parts.

14. In the [quantile applet](#), select the beta distribution. Vary the parameters and note the shape of the density function and the distribution function. In each of the following cases, find the median, the first and third quartiles, and the interquartile range. Sketch the boxplot.

- $a = 1, b = 1$
- $a = 1, b = 3$
- $a = 3, b = 1$
- $a = 2, b = 4$
- $a = 4, b = 2$
- $a = 4, b = 4$



Moments

The [moments](#) of the beta distribution are easy to express in terms of the beta function.

15. Suppose that U has the beta distribution with left parameter a and right parameter b . Show that

$$\mathbb{E}(U^k) = \frac{B(a + k, b)}{B(a, b)}$$

16. Suppose that U has the beta distribution with left parameter a and right parameter b . Show that

- $\mathbb{E}(U) = \frac{a}{a + b}$
- $\text{var}(U) = \frac{ab}{(a + b)^2 (a + b + 1)}$

In particular, if $a = b = \frac{1}{2}$, so that U has the arcsine distribution, then $\mathbb{E}(U) = \frac{1}{2}$ and $\text{var}(U) = \frac{1}{8}$

17. In the simulation of the [random variable experiment](#), select the beta distribution. Set the parameters to values in each of the ranges of [Exercise 8](#). In each case, note the size and location of the mean/standard deviation bar. In each case, run the simulation 1000 times with an update frequency of 10. Note the apparent convergence of the sample moments to the distribution moments.

Transformations

18. Suppose that X has the [gamma distribution](#) with shape parameter a and scale parameter r , that Y has the gamma distribution with shape parameters b and scale parameter r , and that X and Y are [independent](#). Show that $U = \frac{X}{X+Y}$ has the beta distribution with left parameter a and right parameter b .

19. Suppose that U has the beta distribution with left parameter a and right parameter b . Show that $V = 1 - U$ has the beta distribution with parameters left parameter b and right parameter a .

20. Suppose that X has the [F distribution](#) with m degrees of freedom in the numerator and n degrees of freedom in the denominator. Show that

$$U = \frac{(m/n)X}{1 + (m/n)X}$$

has the beta distribution with left parameter $a = \frac{m}{2}$ and right parameter $b = \frac{n}{2}$.

21. Suppose that X has the beta distribution with left parameter a and right parameter b . Show that the distribution is a [two-parameter exponential family](#) with natural parameters $a - 1$ and $b - 1$, and natural statistics $\ln(X)$ and $\ln(1 - X)$.

The beta distribution is also the distribution of the [order statistics](#) of a [random sample](#) from the [uniform distribution](#).

[Virtual Laboratories](#) > [5. Special Distributions](#) > [1](#) [2](#) [3](#) [4](#) [5](#) [6](#) [7](#) **[8](#)** [9](#) [10](#) [11](#) [12](#) [13](#) [14](#) [15](#)

[Contents](#) | [Applets](#) | [Data Sets](#) | [Biographies](#) | [External Resources](#) | [Key words](#) | [Feedback](#) | [©](#)