

5.17: The Beta Distribution

In this section, we will study the beta distribution, the most important distribution that has bounded support. But before we can study the beta *distribution* we must study the beta *function*.

The Beta Function

Definition

The beta function B is defined as follows:

$$B(a, b) = \int_0^1 u^{a-1} (1-u)^{b-1} du; \quad a, b \in (0, \infty) \quad (5.17.1)$$

Proof that B is well defined

We need to show that $B(a, b) < \infty$ for every $a, b \in (0, \infty)$. The integrand is positive on $(0, 1)$, so the integral exists, either as a real number or ∞ . If $a \geq 1$ and $b \geq 1$, the integrand is continuous on $[0, 1]$, so of course the integral is finite. Thus, the only cases of interest are when $0 < a < 1$ or $0 < b < 1$. Note that

$$\int_0^1 u^{a-1} (1-u)^{b-1} du = \int_0^{1/2} u^{a-1} (1-u)^{b-1} du + \int_{1/2}^1 u^{a-1} (1-u)^{b-1} du \quad (5.17.2)$$

If $0 < a < 1$, $(1-u)^{b-1}$ is bounded on $(0, \frac{1}{2}]$ and $\int_0^{1/2} u^{a-1} du = \frac{1}{a2^a}$. Hence the first integral on the right in the displayed equation is finite. Similarly, If $0 < b < 1$, u^{a-1} is bounded on $[\frac{1}{2}, 1)$ and $\int_{1/2}^1 (1-u)^{b-1} du = \frac{1}{b2^b}$. Hence the second integral on the right in the displayed equation is also finite.

The beta function was first introduced by Leonhard Euler.

Properties

The beta function satisfies the following properties:

1. $B(a, b) = B(b, a)$ for $a, b \in (0, \infty)$, so B is symmetric.
2. $B(a, 1) = \frac{1}{a}$ for $a \in (0, \infty)$
3. $B(1, b) = \frac{1}{b}$ for $b \in (0, \infty)$

Proof

1. Using the substitution $v = 1 - u$ we have

$$B(a, b) = \int_0^1 u^{a-1} (1-u)^{b-1} du = \int_0^1 (1-v)^{a-1} v^{b-1} dv = B(b, a) \quad (5.17.3)$$

2. $B(a, 1) = \int_0^1 u^{a-1} du = \frac{1}{a}$
3. This follows from (a) and (b).

The beta function has a simple expression in terms of the gamma function:

If $a, b \in (0, \infty)$ then

$$B(a, b) = \frac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)} \quad (5.17.4)$$

Proof

From the definitions, we can express $\Gamma(a+b)B(a, b)$ as a double integral:

$$\Gamma(a+b)B(a, b) = \int_0^\infty x^{a+b-1} e^{-x} dx \int_0^1 y^{a-1} (1-y)^{b-1} dy = \int_0^\infty \int_0^1 (xy)^{a-1} [x(1-y)]^{b-1} x e^{-x} dx dy \quad (5.17.5)$$

Next we use the transformation $w = xy, z = x(1 - y)$ which maps $(0, \infty) \times (0, 1)$ one-to-one onto $(0, \infty) \times (0, \infty)$. The inverse transformation is $x = w + z, y = w/(w + z)$ and the absolute value of the Jacobian is

$$\left| \det \frac{\partial(x, y)}{\partial(w, z)} \right| = \frac{1}{(w + z)} \quad (5.17.6)$$

Thus, using the change of variables theorem for multiple integrals, the integral above becomes

$$\int_0^\infty \int_0^\infty w^{a-1} z^{b-1} (w + z) e^{-(w+z)} \frac{1}{w + z} dw dz \quad (5.17.7)$$

which after simplifying is $\Gamma(a)\Gamma(b)$.

Recall that the gamma function is a generalization of the factorial function. Here is the corresponding result for the beta function:

If $j, k \in \mathbb{N}_+$ then

$$B(j, k) = \frac{(j-1)!(k-1)!}{(j+k-1)!} \quad (5.17.8)$$

Proof

Recall that $\Gamma(n) = (n-1)!$ for $n \in \mathbb{N}_+$, so this result follows from the previous one.

Let's generalize this result. First, recall from our study of combinatorial structures that for $a \in \mathbb{R}$ and $j \in \mathbb{N}$, the *ascending power* of base a and order j is

$$a^{[j]} = a(a+1) \cdots [a+(j-1)] \quad (5.17.9)$$

If $a, b \in (0, \infty)$, and $j, k \in \mathbb{N}$, then

$$\frac{B(a+j, b+k)}{B(a, b)} = \frac{a^{[j]} b^{[k]}}{(a+b)^{[j+k]}} \quad (5.17.10)$$

Proof

Recall that $\Gamma(a+j) = a^{[j]} \Gamma(a)$, so the result follows from the representation [above](#) for the beta function in terms of the gamma function.

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

Proof

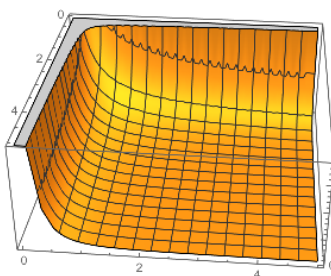


Figure 5.17.1: The graph of $B(a, b)$ on the square $0 < a < 5, 0 < b < 5$

The Incomplete Beta Function

The integral that defines the beta function can be generalized by changing the interval of integration from $(0, 1)$ to $(0, x)$ where $x \in [0, 1]$.

The *incomplete beta function* is defined as follows

$$B(x; a, b) = \int_0^x u^{a-1} (1-u)^{b-1} du, \quad x \in (0, 1); a, b \in (0, \infty) \quad (5.17.11)$$

Of course, the ordinary (complete) beta function is $B(a, b) = B(1; a, b)$ for $a, b \in (0, \infty)$.

The Standard Beta Distribution

Distribution Functions

The beta distributions are a family of continuous distributions on the interval $(0, 1)$.

The (standard) *beta distribution* with *left parameter* $a \in (0, \infty)$ and *right parameter* $b \in (0, \infty)$ has probability density function f given by

$$f(x) = \frac{1}{B(a, b)} x^{a-1} (1-x)^{b-1}, \quad x \in (0, 1) \quad (5.17.12)$$

Of course, the beta function is simply the normalizing constant, so it's clear that f is a valid probability density function. If $a \geq 1$, f is defined at 0, and if $b \geq 1$, f is defined at 1. In these cases, it's customary to extend the domain of f to these endpoints. The beta distribution is useful for modeling random probabilities and proportions, particularly in the context of Bayesian analysis. The distribution has just two parameters and yet a rich variety of shapes (so in particular, both parameters are *shape parameters*). Qualitatively, the first order properties of f depend on whether each parameter is less than, equal to, or greater than 1.

For $a, b \in (0, \infty)$ with $a + b \neq 2$, define

$$x_0 = \frac{a-1}{a+b-2} \quad (5.17.13)$$

1. If $0 < a < 1$ and $0 < b < 1$, f decreases and then increases with minimum value at x_0 and with $f(x) \rightarrow \infty$ as $x \downarrow 0$ and as $x \uparrow 1$.
2. If $a = 1$ and $b = 1$, f is constant.
3. If $0 < a < 1$ and $b \geq 1$, f is decreasing with $f(x) \rightarrow \infty$ as $x \downarrow 0$.
4. If $a \geq 1$ and $0 < b < 1$, f is increasing with $f(x) \rightarrow \infty$ as $x \uparrow 1$.
5. If $a = 1$ and $b > 1$, f is decreasing with mode at $x = 0$.
6. If $a > 1$ and $b = 1$, f is increasing with mode at $x = 1$.
7. If $a > 1$ and $b > 1$, f increases and then decreases with mode at x_0 .

Proof

These results follow from standard calculus. The first derivative is

$$f'(x) = \frac{1}{B(a, b)} x^{a-2} (1-x)^{b-2} [(a-1) - (a+b-2)x], \quad 0 < x < 1 \quad (5.17.14)$$

From part (b), note that the special case $a = 1$ and $b = 1$ gives the continuous uniform distribution on the interval $(0, 1)$ (the *standard uniform distribution*). Note also that when $a < 1$ or $b < 1$, the probability density function is unbounded, and hence the distribution has no mode. On the other hand, if $a \geq 1$, $b \geq 1$, and one of the inequalities is strict, the distribution has a unique mode at x_0 . The second order properties are more complicated.

For $a, b \in (0, \infty)$ with $a + b \notin \{2, 3\}$ and $(a-1)(b-1)(a+b-3) \geq 0$, define

$$x_1 = \frac{(a-1)(a+b-3) - \sqrt{(a-1)(b-1)(a+b-3)}}{(a+b-3)(a+b-2)} \quad (5.17.15)$$

$$x_2 = \frac{(a-1)(a+b-3) + \sqrt{(a-1)(b-1)(a+b-3)}}{(a+b-3)(a+b-2)} \quad (5.17.16)$$

For $a < 1$ and $a + b = 2$ or for $b < 1$ and $a + b = 2$, define $x_1 = x_2 = 1 - a/2$.

1. If $a \leq 1$ and $b \leq 1$, or if $a \leq 1$ and $b \geq 2$, or if $a \geq 2$ and $b \leq 1$, f is concave upward.
2. If $a \leq 1$ and $1 < b < 2$, f is concave upward and then downward with inflection point at x_1 .
3. If $1 < a < 2$ and $b \leq 1$, f is concave downward and then upward with inflection point at x_2 .
4. If $1 < a \leq 2$ and $1 < b \leq 2$, f is concave downward.
5. If $1 < a \leq 2$ and $b > 2$, f is concave downward and then upward with inflection point at x_2 .

6. If $a > 2$ and $1 < b \leq 2$, f is concave upward and then downward with inflection point at x_1 .
7. If $a > 2$ and $b > 2$, f is concave upward, then downward, then upward again, with inflection points at x_1 and x_2 .

Proof

These results follow from standard (but very tedious) calculus. The second derivative is

$$f''(x) = \frac{1}{B(a, b)} x^{a-3} (1-x)^{b-3} [(a+b-2)(a+b-3)x^2 - 2(a-1)(a+b-3)x + (a-1)(a-2)] \quad (5.17.17)$$

In the special distribution simulator, select the beta distribution. Vary the parameters and note the shape of the beta density function. For selected values of the parameters, run the simulation 1000 times and compare the empirical density function to the true density function.

The special case $a = \frac{1}{2}$, $b = \frac{1}{2}$ is the *arcsine distribution*, with probability density function given by

$$f(x) = \frac{1}{\pi \sqrt{x(1-x)}}, \quad x \in (0, 1) \quad (5.17.18)$$

This distribution is important in a number of applications, and so the arcsine distribution is studied in a separate section.

The beta distribution function F can be easily expressed in terms of the [incomplete beta function](#). As usual a denotes the left parameter and b the right parameter.

The beta distribution function F with parameters $a, b \in (0, \infty)$ is given by

$$F(x) = \frac{B(x; a, b)}{B(a, b)}, \quad x \in (0, 1) \quad (5.17.19)$$

The distribution function F is sometimes known as the *regularized incomplete beta function*. In some special cases, the distribution function F and its inverse, the quantile function F^{-1} , can be computed in closed form, without resorting to special functions.

If $a \in (0, \infty)$ and $b = 1$ then

1. $F(x) = x^a$ for $x \in (0, 1)$
2. $F^{-1}(p) = p^{1/a}$ for $p \in (0, 1)$

If $a = 1$ and $b \in (0, \infty)$ then

1. $F(x) = 1 - (1-x)^b$ for $x \in (0, 1)$
2. $F^{-1}(p) = 1 - (1-p)^{1/b}$ for $p \in (0, 1)$

If $a = b = \frac{1}{2}$ (the arcsine distribution) then

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

There is an interesting relationship between the distribution functions of the beta distribution and the binomial distribution, when the beta parameters are positive integers. To state the relationship we need to embellish our notation to indicate the dependence on the parameters. Thus, let $F_{a,b}$ denote the beta distribution function with left parameter $a \in (0, \infty)$ and right parameter $b \in (0, \infty)$, and let $G_{n,p}$ denote the binomial distribution function with trial parameter $n \in \mathbb{N}_+$ and success parameter $p \in (0, 1)$.

If $j, k \in \mathbb{N}_+$ and $x \in (0, 1)$ then

$$F_{j,k}(x) = G_{j+k-1, 1-x}(k-1) \quad (5.17.20)$$

Proof

By definition

$$F_{j,k}(x) = \frac{1}{B(j,k)} \int_0^x t^{j-1} (1-t)^{k-1} dt \quad (5.17.21)$$

Integrate by parts with $u = (1-t)^{k-1}$ and $dv = t^{j-1} dt$, so that $du = -(k-1)(1-t)^{k-2}$ and $v = t^j/j$. The result is

$$F_{j,k}(x) = \frac{1}{jB(j,k)} (1-x)^{k-1} x^j + \frac{k-1}{jB(j,k)} \int_0^x t^j (1-t)^{k-2} dt \quad (5.17.22)$$

But by the [property of the beta function](#) above, $B(j,k) = (j-1)!(k-1)!/(j+k-1)!$. Hence $1/jB(j,k) = \binom{j+k-1}{k-1}$ and $(k-1)/jB(j,k) = 1/B(j+1,k-1)$. Thus, the last displayed equation can be rewritten as

$$F_{j,k}(x) = \binom{j+k-1}{k-1} (1-x)^{k-1} x^j + F_{j+1,k-1}(x) \quad (5.17.23)$$

Recall from the [special case](#) above that $F_{j+k-1,1}(x) = x^{j+k-1}$. Iterating the last displayed equation gives the result.

In the special distribution calculator, 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.

1. $a = 1, b = 1$
2. $a = 1, b = 3$
3. $a = 3, b = 1$
4. $a = 2, b = 4$
5. $a = 4, b = 2$
6. $a = 4, b = 4$

Moments

The moments of the beta distribution are easy to express in terms of the beta function. As before, suppose that X has the beta distribution with left parameter $a \in (0, \infty)$ and right parameter $b \in (0, \infty)$.

If $k \in [0, \infty)$ then

$$\mathbb{E}(X^k) = \frac{B(a+k, b)}{B(a, b)} \quad (5.17.24)$$

In particular, if $k \in \mathbb{N}$ then

$$\mathbb{E}(X^k) = \frac{a^{[k]}}{(a+b)^{[k]}} \quad (5.17.25)$$

Proof

Note that

$$\mathbb{E}(X^k) = \int_0^1 x^k \frac{1}{B(a,b)} x^{a-1} (1-x)^{b-1} dx = \frac{1}{B(a,b)} \int_0^1 x^{a+k-1} (1-x)^{b-1} dx = \frac{B(a+k, b)}{B(a, b)} \quad (5.17.26)$$

If $k \in \mathbb{N}$, the formula simplifies by the [property of the beta function](#) above.

From the general formula for the moments, it's straightforward to compute the mean, variance, skewness, and kurtosis.

The mean and variance of X are

$$\mathbb{E}(X) = \frac{a}{a+b} \quad (5.17.27)$$

$$\text{var}(X) = \frac{ab}{(a+b)^2(a+b+1)} \quad (5.17.28)$$

Proof

The formula for the mean and variance follow from the [formula for the moments](#) and the computational formula $\text{var}(X) = \mathbb{E}(X^2) - [\mathbb{E}(X)]^2$

Note that the variance depends on the parameters a and b only through the product ab and the sum $a + b$.

Open the special distribution simulator and select the beta distribution. Vary the parameters and note the size and location of the mean \pm standard deviation bar. For selected values of the parameters, run the simulation 1000 times and compare the sample mean and standard deviation to the distribution mean and standard deviation.

The skewness and kurtosis of X are

$$\text{skew}(X) = \frac{2(b-a)\sqrt{a+b+1}}{(a+b+2)\sqrt{ab}} \quad (5.17.29)$$

$$\text{kurt}(X) = \frac{3a^3b + 3ab^3 + 6a^2b^2 + a^3 + b^3 + 13a^2b + 13ab^2 + a^2 + b^2 + 14ab}{ab(a+b+2)(a+b+3)} \quad (5.17.30)$$

Proof

These results follow from the computational formulas that give the skewness and kurtosis in terms of $\mathbb{E}(X^k)$ for $k \in \{1, 2, 3, 4\}$ and the [formula for the moments](#) above.

In particular, note that the distribution is positively skewed if $a < b$, unskewed if $a = b$ (the distribution is symmetric about $x = \frac{1}{2}$ in this case) and negatively skewed if $a > b$.

Open the special distribution simulator and select the beta distribution. Vary the parameters and note the shape of the probability density function in light of the previous result on skewness. For various values of the parameters, run the simulation 1000 times and compare the empirical density function to the true probability density function.

Related Distributions

The beta distribution is related to a number of other special distributions.

If X has the beta distribution with left parameter $a \in (0, \infty)$ and right parameter $b \in (0, \infty)$ then $Y = 1 - X$ has the beta distribution with left parameter b and right parameter a .

Proof

This follows from the standard change of variables formula. If f and g denote the PDFs of X and Y respectively, then

$$g(y) = f(1-y) = \frac{1}{B(a,b)}(1-y)^{a-1}y^{b-1} = \frac{1}{B(b,a)}y^{b-1}(1-y)^{a-1}, \quad y \in (0,1) \quad (5.17.31)$$

The beta distribution with right parameter 1 has a reciprocal relationship with the Pareto distribution.

Suppose that $a \in (0, \infty)$.

1. If X has the beta distribution with left parameter a and right parameter 1 then $Y = 1/X$ has the Pareto distribution with shape parameter a .
2. If Y has the Pareto distribution with shape parameter a then $X = 1/Y$ has the beta distribution with left parameter a and right parameter 1.

Proof

The two results are equivalent. In (a), suppose that X has the beta distribution with parameters a and 1. The transformation $y = 1/x$ maps $(0,1)$ one-to-one onto $(0,\infty)$. The inverse is $x = 1/y$ with $dx/dy = -1/y^2$. Recall also that $B(a,1) = 1/a$. By the change of variables formula, the PDF g of $Y = 1/X$ is given by

$$g(y) = f\left(\frac{1}{y}\right) \frac{1}{y^2} = a\left(\frac{1}{y}\right)^{a-1} \frac{1}{y^2} = \frac{a}{y^{a+1}}, \quad y \in (0,\infty) \quad (5.17.32)$$

We recognize g as the PDF of the Pareto distribution with shape parameter a .

The following result gives a connection between the beta distribution and the gamma distribution.

Suppose that X has the gamma distribution with shape parameter $a \in (0, \infty)$ and rate parameter $r \in (0, \infty)$, Y has the gamma distribution with shape parameter $b \in (0, \infty)$ and rate parameter r , and that X and Y are independent. Then $V = X/(X + Y)$ has the beta distribution with left parameter a and right parameter b .

Proof

Let $U = X + Y$ and $V = X/(X + Y)$. We will actually prove stronger results: U and V are independent, U has the gamma distribution with shape parameter $a + b$ and rate parameter r , and V has the beta distribution with parameters a and b . First note that (X, Y) has joint PDF f given by

$$f(x, y) = \frac{r^a}{\Gamma(a)} x^{a-1} e^{-rx} \frac{r^b}{\Gamma(b)} y^{b-1} e^{-ry} = \frac{r^{a+b}}{\Gamma(a)\Gamma(b)} x^{a-1} y^{b-1} e^{-r(x+y)}; \quad x, y \in (0, \infty) \quad (5.17.33)$$

The transformation $u = x + y$ and $v = x/(x + y)$ maps $(0, \infty) \times (0, \infty)$ one-to-one onto $(0, \infty) \times (0, 1)$. The inverse is $x = uv$, $y = u(1 - v)$ and the absolute value of the Jacobian is

$$\left| \det \frac{\partial(x, y)}{\partial(u, v)} \right| = u \quad (5.17.34)$$

Hence by the multivariate change of variables theorem, the PDF g of (U, V) is given by

$$g(u, v) = f[uv, u(1 - v)]u = \frac{r^{a+b}}{\Gamma(a)\Gamma(b)} (uv)^{a-1} [u(1 - v)]^{b-1} e^{-ru} u \quad (5.17.35)$$

$$= \frac{r^{a+b}}{\Gamma(a)\Gamma(b)} u^{a+b-1} e^{-ru} v^{a-1} (1 - v)^{b-1} \quad (5.17.36)$$

$$= \frac{r^{a+b}}{\Gamma(a+b)} u^{a+b-1} e^{-ru} \frac{\Gamma(a+b)}{\Gamma(a)\Gamma(b)} v^{a-1} (1 - v)^{b-1}; \quad u \in (0, \infty), v \in (0, 1) \quad (5.17.37)$$

The results now follow from the factorization theorem. The factor in u is the gamma PDF with shape parameter $a + b$ and rate parameter r while the factor in v is the beta PDF with parameters a and b .

The following result gives a connection between the beta distribution and the F distribution. This connection is a minor variation of the previous result.

If X has the F distribution with $n \in (0, \infty)$ degrees of freedom in the numerator and $d \in (0, \infty)$ degrees of freedom in the denominator then

$$Y = \frac{(n/d)X}{1 + (n/d)X} \quad (5.17.38)$$

has the beta distribution with left parameter $a = n/2$ and right parameter $b = d/2$.

Proof

If X has the F distribution with $n > 0$ degrees of freedom in the numerator and $d > 0$ degrees of freedom in the denominator then X can be written as

$$X = \frac{U/n}{V/d} \quad (5.17.39)$$

where U has the chi-square distribution with n degrees of freedom, V has the chi-square distribution with d degrees of freedom, and U and V are independent. Hence

$$Y = \frac{(n/d)X}{1 + (n/d)X} = \frac{U/V}{1 + U/V} = \frac{U}{U + V} \quad (5.17.40)$$

But the chi-square distribution is a special case of the gamma distribution. Specifically, U has the gamma distribution with shape parameter $n/2$ and rate parameter $1/2$, V has the gamma distribution with shape parameter $d/2$ and rate parameter $1/2$ and

again U and V are independent. Hence by the [previous result](#), Y has the beta distribution with left parameter $n/2$ and right parameter $d/2$.

Our next result is that the beta distribution is a member of the general exponential family of distributions.

Suppose that X has the beta distribution with left parameter $a \in (0, \infty)$ and right parameter $b \in (0, \infty)$. Then the distribution is a two-parameter exponential family with natural parameters $a - 1$ and $b - 1$, and natural statistics $\ln(X)$ and $\ln(1 - X)$.

Proof

This follows from the definition of the general exponential distribution, since the PDF f of X can be written as

$$f(x) = \frac{1}{B(a, b)} \exp[(a - 1) \ln(x) + (b - 1) \ln(1 - x)], \quad x \in (0, 1) \quad (5.17.41)$$

The beta distribution is also the distribution of the order statistics of a random sample from the standard uniform distribution.

Suppose $n \in \mathbb{N}_+$ and that (X_1, X_2, \dots, X_n) is a sequence of independent variables, each with the standard uniform distribution. For $k \in \{1, 2, \dots, n\}$, the k th order statistics $X_{(k)}$ has the beta distribution with left parameter $a = k$ and right parameter $b = n - k + 1$.

Proof

See the section on order statistics.

One of the most important properties of the beta distribution, and one of the main reasons for its wide use in statistics, is that it forms a conjugate family for the success probability in the binomial and negative binomial distributions.

Suppose that P is a random probability having the beta distribution with left parameter $a \in (0, \infty)$ and right parameter $b \in (0, \infty)$. Suppose also that X is a random variable such that the conditional distribution of X given $P = p \in (0, 1)$ is binomial with trial parameter $n \in \mathbb{N}_+$ and success parameter p . Then the conditional distribution of P given $X = k$ is beta with left parameter $a + k$ and right parameter $b + n - k$.

Proof

The joint PDF f of (P, X) on $(0, 1) \times \{0, 1, \dots, n\}$ is given by

$$f(p, k) = \frac{1}{B(a, b)} p^{a-1} (1-p)^{b-1} \binom{n}{k} p^k (1-p)^{n-k} = \frac{1}{B(a, b)} \binom{n}{k} p^{a+k-1} (1-p)^{b+n-k-1} \quad (5.17.42)$$

The conditional PDF of P given $X = k$ is simply the normalized version of the function $p \mapsto f(p, k)$. We can tell from the functional form that this distribution is beta with the parameters given in the theorem.

Suppose again that P is a random probability having the beta distribution with left parameter $a \in (0, \infty)$ and right parameter $b \in (0, \infty)$. Suppose also that N is a random variable such that the conditional distribution of N given $P = p \in (0, 1)$ is negative binomial with stopping parameter $k \in \mathbb{N}_+$ and success parameter p . Then the conditional distribution of P given $N = n$ is beta with left parameter $a + k$ and right parameter $b + n - k$.

Proof

The joint PDF f of (P, N) on $(0, 1) \times \{k, k + 1, \dots\}$ is given by

$$f(p, n) = \frac{1}{B(a, b)} p^{a-1} (1-p)^{b-1} \binom{n-1}{k-1} p^k (1-p)^{n-k} = \frac{1}{B(a, b)} \binom{n-1}{k-1} p^{a+k-1} (1-p)^{b+n-k-1} \quad (5.17.43)$$

The conditional PDF of P given $N = n$ is simply the normalized version of the function $p \mapsto f(p, n)$. We can tell from the functional form that this distribution is beta with the parameters given in the theorem.

In both cases, note that in the posterior distribution of P , the left parameter is increased by the number of successes and the right parameter by the number of failures. For more on this, see the section on Bayesian estimation in the chapter on point estimation.

The General Beta Distribution

The beta distribution can be easily generalized from the support interval $(0, 1)$ to an arbitrary bounded interval using a linear transformation. Thus, this generalization is simply the location-scale family associated with the standard beta distribution.

Suppose that Z has the standard beta distribution with left parameter $a \in (0, \infty)$ and right parameter $b \in (0, \infty)$. For $c \in \mathbb{R}$ and $d \in (0, \infty)$ random variable $X = c + dZ$ has the beta distribution with left parameter a , right parameter b , location parameter c and scale parameter d .

For the remainder of this discussion, suppose that X has the distribution in the definition above.

X has probability density function

$$f(x) = \frac{1}{B(a, b)d^{a+b-1}}(x - c)^{a-1}(c + d - x)^{b-1}, \quad x \in (c, c + d) \quad (5.17.44)$$

Proof

This follows from a standard result for location-scale families. If g denotes the [standard beta PDF](#) of Z , then X has PDF f given by

$$f(x) = \frac{1}{d}g\left(\frac{x - c}{d}\right), \quad x \in (c, c + d) \quad (5.17.45)$$

Most of the results in the previous sections have simple extensions to the general beta distribution.

The mean and variance of X are

1. $\mathbb{E}(X) = c + d\frac{a}{a+b}$
2. $\text{var}(X) = d^2 \frac{ab}{(a+b)^2(a+b+1)}$

Proof

This follows from the [standard mean and variance](#) and basic properties of expected value and variance.

1. $\mathbb{E}(X) = c + d\mathbb{E}(Z)$
2. $\text{var}(X) = d^2\text{var}(Z)$

Recall that skewness and variance are defined in terms of standard scores, and hence are unchanged under location-scale transformations. Hence the skewness and kurtosis of X are just as for the [standard beta distribution](#).

This page titled [5.17: The Beta Distribution](#) is shared under a [CC BY 2.0](#) license and was authored, remixed, and/or curated by [Kyle Siegrist \(Random Services\)](#) via [source content](#) that was edited to the style and standards of the LibreTexts platform.