5.19: The Arcsine Distribution

The *arcsine distribution* is important in the study of Brownian motion and prime numbers, among other applications.

The Standard Arcsine Distribution

Distribution Functions

The *standard arcsine distribution* is a continuous distribution on the interval $(0, 1)$ with probability density function g given by

$$g(x) = \frac{1}{\pi \sqrt{x(1-x)}}, \quad x \in (0, 1) \tag{5.19.1}$$

Proof

There are a couple of ways to see that g is a valid PDF. First, it's the beta PDF with parameters $a = b = \frac{1}{2}$:

$$g(x) = \frac{1}{B(1/2, 1/2)} x^{-1/2}(1-x)^{-1/2}, \quad x \in (0, 1) \tag{5.19.2}$$

since we recall that $B\left(\frac{1}{2}, \frac{1}{2}\right) = \pi$. A direct proof is also easy: The substitution $u = \sqrt{x}$, $x = u^2$, $dx = 2u\,du$ gives

$$\int_0^1 \frac{1}{\pi\sqrt{x(1-x)}} dx = \int_0^1 \frac{2}{\pi\sqrt{1-u^2}} du = \frac{2}{\pi} \arcsin u \Big|_0^1 = \frac{2}{\pi}\left(\frac{\pi}{2} - 0\right) = 1 \tag{5.19.3}$$

The occurrence of the arcsine function in the proof that g is a probability density function explains the name.

The standard arcsine probability density function g satisfies the following properties:

1. g is symmetric about $x = \frac{1}{2}$.
2. g decreases and then increases with minimum value at $x = \frac{1}{2}$.
3. g is concave upward
4. $g(x) \to \infty$ as $x \downarrow 0$ and as $x \uparrow 1$.

Proof

1. Note that g is a function of x only through $x(1-x)$.
2. This follows from standard calculus:

$$g'(x) = \frac{2x - 1}{2\pi[x(1-x)]^{3/2}} \tag{5.19.4}$$

3. This also follows from standard calculus:

$$g''(x) = \frac{3 - 8x + 8x^2}{4\pi[x(1-x)]^{5/2}} \tag{5.19.5}$$

4. The limits are clear.

In particular, the standard arcsine distribution is U-shaped and has no mode.

Open the Special Distribution Simulator and select the arcsine distribution. Keep the default parameter values and note the shape of the probability density function. Run the simulation 1000 times and compare the emprical density function to the probability density function.

The distribution function has a simple expression in terms of the arcsine function, again justifying the name of the distribution.

The standard arcsine distribution function G is given by $G(x) = \frac{2}{\pi}\arcsin(\sqrt{x})$ for $x \in [0, 1]$.

Proof

Again, using the substitution $u = \sqrt{t}$, $t = u^2$, $dt = 2u\,du$:

$$G(x) = \int_0^x \frac{1}{\pi\sqrt{t(1-t)}}dt = \int_0^{\sqrt{x}} \frac{2}{\pi\sqrt{1-u^2}}du = \frac{2}{\pi}\arcsin(t)\Big|_0^{\sqrt{x}} = \frac{2}{\pi}\arcsin(\sqrt{x}) \tag{5.19.6}$$

Not surprisingly, the quantile function has a simple expression in terms of the sine function.

The standard arcinse quantile function G^{-1} is given by $G^{-1}(p) = \sin^2\left(\frac{\pi}{2}p\right)$ for $p \in [0,1]$. In particular, the quartiles are

1. $q_1 = \sin^2\left(\frac{\pi}{8}\right) = \frac{1}{4}(2-\sqrt{2}) \approx 0.1464$, the first quartile
2. $q_2 = \frac{1}{2}$, the median
3. $q_3 = \sin^2\left(\frac{3\pi}{8}\right) = \frac{1}{4}(2+\sqrt{2}) \approx 0.8536$, the third quartile

Proof

The formula for the quantile function follows from the distribution function by solving $p = G(x)$ for x in terms of $p \in [0,1]$.

Open the Special Distribution Calculator and select the arcsine distribution. Keep the default parameter values and note the shape of the distribution function. Compute selected values of the distribution function and the quantile function.

Moments

Suppose that random variable Z has the standard arcsine distribution. First we give the mean and variance.

The mean and variance of Z are

1. $\mathbb{E}(Z) = \frac{1}{2}$
2. $\text{var}(Z) = \frac{1}{8}$

Proof

1. The mean is $\frac{1}{2}$ by symmetry.
2. Using the usual substitution $u = \sqrt{x}$, $x = u^2$ $dx = 2u\,du$ and then the substitution $u = \sin\theta$, $du = \cos\theta\,d\theta$ gives

$$\mathbb{E}\left(Z^2\right) = \int_0^1 \frac{1}{\pi\sqrt{x(1-x)}}dx = \int_0^1 \frac{2u^4}{\pi\sqrt{1-u^2}} = \int_0^{\pi/2} \frac{2}{\pi}\sin^4(\theta)d\theta = \frac{2}{\pi}\frac{3\pi}{16} = \frac{3}{8} \tag{5.19.7}$$

Open the Special Distribution Simulator and select the arcsine distribution. Keep the default parameter values. Run the simulation 1000 times and compare the empirical mean and stadard deviation to the true mean and standard deviation.

The general moments about 0 can be expressed as products.

For $n \in \mathbb{N}$,

$$\mathbb{E}\left(Z^n\right) = \prod_{j=0}^{n-1} \frac{2j+1}{2j+2} \tag{5.19.8}$$

Proof

The same integral substitutions as before gives

$$\mathbb{E}(Z^n) = \int_0^{\pi/2} \frac{2}{\pi}\sin^{2n}(\theta)d\theta = \prod_{j=0}^{n-1} \frac{2j+1}{2j+2} \tag{5.19.9}$$

Of course, the moments can be used to give a formula for the moment generating function, but this formula is not particularly helpful since it is not in closed form.

Z has moment generating function m given by

$$m(t) = \mathbb{E}\left(e^{tZ}\right) = \sum_{n=0}^\infty \left(\prod_{j=0}^{n-1} \frac{2j+1}{2j+2}\right) \frac{t^n}{n!}, \quad t \in \mathbb{R} \tag{5.19.10}$$

Finally we give the skewness and kurtosis.

The skewness and kurtosis of Z are

1. $\text{skew}(Z) = 0$
2. $\text{kurt}(Z) = \frac{3}{2}$

Proof

1. The skewness is 0 by the symmetry of the distribution.
2. The result for the kurtosis follows from the standard formula for kurtosis in terms of the moments: $\mathbb{E}(Z) = \frac{1}{2}$, $\mathbb{E}\left(Z^2\right) = \frac{3}{8}$, $\mathbb{E}\left(Z^3\right) = \frac{5}{16}$, and $\mathbb{E}\left(Z^4\right) = \frac{35}{128}$.

Related Distributions

As noted earlier, the standard arcsine distribution is a special case of the beta distribution.

The standard arcsine distribution is the beta distribution with left parameter $\frac{1}{2}$ and right parameter $\frac{1}{2}$.

Proof

The beta distribution with parameters $a = b = \frac{1}{2}$ has PDF

$$x \mapsto \frac{1}{B(1/2, 1/2)} x^{-1/2}(1-x)^{-1/2}, \quad x \in (0,1) \tag{5.19.11}$$

But $B\left(\frac{1}{2}, \frac{1}{2}\right) = \pi$, so this is the standard arcsine PDF.

Since the quantile function is in closed form, the standard arcsine distribution can be simulated by the random quantile method.

Connections with the standard uniform distribution.

1. If U has the standard uniform distribution (a random number) then $X = \sin^2\left(\frac{\pi}{2}U\right)$ has the standard arcsine distribution.
2. If X has the standard arcsine distribution then $U = \frac{2}{\pi}\arcsin\left(\sqrt{X}\right)$ has the standard uniform distribution.

Open the random quantile simulator and select the arcsine distribution. Keep the default parameters. Run the experiment 1000 times and compare the empirical probability density function, mean, and standard deviation to their distributional counterparts. Note how the random quantiles simulate the distribution.

The following exercise illustrates the connection between the Brownian motion process and the standard arcsine distribution.

Open the Brownian motion simulator. Keep the default time parameter and select the last zero random variable. Note that this random variable has the standard arcsine distribution. Run the experiment 1000 time and compare the empirical probability density function, mean, and standard deviation to their distributional counterparts. Note how the last zero simulates the distribution.

The General Arcsine Distribution

The standard arcsine distribution is generalized by adding location and scale parameters.

Definition

If Z has the standard arcsine distribution, and if $a \in \mathbb{R}$ and $b \in (0, \infty)$, then $X = a + bZ$ has the *arcsine distribution* with location parameter a and scale parameter b.

So X has a continuous distribution on the interval $(a, a+b)$.

Distribution Functions

Suppose that X has the arcsine distribution with location parameter $a \in \mathbb{R}$ and scale parameter $w \in (0, \infty)$.

X has probability density function f given by

$$f(x) = \frac{1}{\pi\sqrt{(x-a)(a+w-x)}}, \quad x \in (a, a+w) \tag{5.19.12}$$

1. f is symmetric about $a + \frac{1}{2}w$.
2. f decreases and then increases with minimum value at $x = a + \frac{1}{2}w$.
3. f is concave upward.
4. $f(x) \to \infty$ as $x \downarrow a$ and as $x \uparrow a + w$.

Proof

Recall that $f(x) = \frac{1}{w} g\left(\frac{x-a}{w}\right)$ where g is the PDF of the standard arcsine distribution.

An alternate parameterization of the general arcsine distribution is by the endpoints of the support interval: the left endpoint (location parameter) a and the right endpoint $b = a + w$.

Open the Special Distribution Simulator and select the arcsine distribution. Vary the location and scale parameters and note the shape and location of the probability density function. For selected values of the parameters, run the simulation 1000 times and compare the emprical density function to the probability density function.

Once again, the distribution function has a simple representation in terms of the arcsine function.

X has distribution function F given by

$$F(x) = \frac{2}{\pi}\arcsin\left(\sqrt{\frac{x-a}{w}}\right), \quad x \in [a, a+w] \tag{5.19.13}$$

Proof

Recall that $F(x) = G[(x-a)/w]$ where G is the CDF of the standard arcsine distribution.

As before, the quantile function has a simple representation in terms of the sine functioon

X has quantile function F^{-1} given by $F^{-1}(p) = a + w\sin^2\left(\frac{\pi}{2}p\right)$ for $p \in [0, 1]$ In particular, the quartiles of X are

1. $q_1 = a + w\sin^2\left(\frac{\pi}{8}\right) = a + \frac{1}{4}\left(2 - \sqrt{2}\right) w$, the first quartile
2. $q_2 = a + \frac{1}{2}w$, the median
3. $q_3 = a + w\sin^2\left(\frac{3\pi}{8}\right) = a + \frac{1}{4}\left(2 + \sqrt{2}\right) w$, the third quartile

Proof

Recall that $F^{-1}(p) = a + wG^{-1}(p)$ where G^{-1} is the quantile function of the standard arcsine distribution.

Open the Special Distribution Calculator and select the arcsine distribution. Vary the parameters and note the shape and location of the distribution function. For various values of the parameters, compute selected values of the distribution function and the quantile function.

Moments

Again, we assume that X has the arcsine distribution with location parameter $a \in \mathbb{R}$ and scale parameter $w \in (0, \infty)$. First we give the mean and variance.

The mean and variance of X are

1. $\mathbb{E}(X) = a + \frac{1}{2}w$
2. $\text{var}(X) = \frac{1}{8}w^2$

Proof

These results from the representation $X = a + wZ$ and the results for the mean and variance of Z.

Open the Special Distribution Simulator and select the arcsine distribution. Vary the parameters and note the size and location of the mean±standard deviation bar. For various values of the parameters, run the simulation 1000 times and compare the empirical mean and stadard deviation to the true mean and standar deviation.

The moments of X can be obtained from the moments of Z, but the results are messy, except when the location parameter is 0.

Suppose the location parameter $a = 0$. For $n \in \mathbb{N}$,

$$\mathbb{E}(X^n) = w^n \prod_{j=0}^{n-1} \frac{2j+1}{2j+2} \tag{5.19.14}$$

Proof

This follows from the representation $X = wZ$ and the results for the moments of Z.

The moment generating function can be expressed as a series with product coefficients, and so is not particularly helpful.

X has moment generating function M given by

$$M(t) = \mathbb{E}\left(e^{tX}\right) = e^{at} \sum_{n=0}^{\infty} \left(\prod_{j=0}^{n-1} \frac{2j+1}{2j+2}\right) \frac{w^n t^n}{n!}, \quad t \in \mathbb{R} \tag{5.19.15}$$

Proof

Recall that $M(t) = e^{at} m(wt)$ where m is the moment generating function of Z.

Finally, the skewness and kurtosis are unchanged.

The skewness and kurtosis of X are

1. $\text{skew}(X) = 0$
2. $\text{kurt}(X) = \frac{3}{2}$

Proof

Recall that the skewness and kurtosis are defined in terms of the standard score of X and hence are invariant under a location-scale transformation.

Related Distributions

By construction, the general arcsine distribution is a location-scale family, and so is closed under location-scale transformations.

If X has the arcsine distribution with location parameter $a \in \mathbb{R}$ and scale parameter $w \in (0, \infty)$ and if $c \in \mathbb{R}$ and $d \in (0, \infty)$ then $c + dX$ has the arcsine distribution with location parameter $c + ad$ scale parameter dw.

Proof

By definition we can take $X = a + wZ$ where Z has the standard arcsine distribution. Hence $c + dX = (c + da) + (dw)Z$.

Since the quantile function is in closed form, the arcsine distribution can be simulated by the random quantile method.

Suppose that $a \in \mathbb{R}$ and $w \in (0, \infty)$.

1. If U has the standard uniform distribution (a random number) then $X = a + w\sin^2\left(\frac{\pi}{2}U\right)$ has the arcsine distribution with location parameter a and scale parameter b.
2. If X has the arcsine distribution with location parameter a and scale parameter b then $U = \frac{2}{\pi}\arcsin\left(\sqrt{\frac{X-a}{w}}\right)$ has the standard uniform distribution.

Open the random quantile simulator and select the arcsine distribution. Vary the parameters and note the location and shape of the probability density function. For selected parameter values, run the experiment 1000 times and compare the empirical probability density function, mean, and standard deviation to their distributional counterparts. Note how the random quantiles simulate the distribution.

The following exercise illustrates the connection between the Brownian motion process and the arcsine distribution.

Open the Brownian motion simulator and select the last zero random variable. Vary the time parameter t and note that the last zero has the arcsine distribution on the interval $(0, t)$. Run the experiment 1000 time and compare the empirical probability density function, mean, and standard deviation to their distributional counterparts. Note how the last zero simulates the distribution.
