5.27: The Sine Distribution

The *sine distribution* is a simple probability distribution based on a portion of the sine curve. It is also known as *Gilbert's sine distribution*, named for the American geologist Grove Karl (GK) Gilbert who used the distribution in 1892 to study craters on the moon.

The Standard Sine Distribution

Distribution Functions

The *standard sine distribution* is a continuous distribution on $[0,1]$ with probability density function g given by

$$g(z)=\frac{\pi}{2}\sin(\pi z), \quad z\in[0,1] \tag{5.27.1}$$

1. g is symmetric about $z=\frac{1}{2}$.
2. g increases and then decreases with mode at $z=\frac{1}{2}$.
3. g is concave downward.

Proof

From simple calculus, g is a probability density function: $\sin(\pi x)\geq 0$ for $x\in[0,1]$ and

$$\int_0^1 \sin(\pi z)dz=\frac{2}{\pi} \tag{5.27.2}$$

The properties follow from basic calculus since

$$g'(z)=\frac{\pi^2}{2}\cos(\pi z), \quad z\in[0,1] \tag{5.27.3}$$

$$g''(z)=-\frac{\pi^3}{2}\sin(\pi z), \quad z\in[0,1] \tag{5.27.4}$$

Open the Special Distribution Simulator and select the sine distribution. Run the simulation 1000 times and compare the emprical density function to the probability density function.

The distribution function G is given by $G(z)=\frac{1}{2}[1-\cos(\pi z)]$ for $z\in[0,1]$.

Proof

This follows from the PDF above and simple calculus.

The quantile function G^{-1} is given by $G^{-1}(p)=\frac{1}{\pi}\arccos(1-2p)$ for $p\in[0,1]$.

1. The first quartile is $q_1=\frac{1}{3}$.
2. The median is $\frac{1}{2}$.
3. The third quartile is $q_3=\frac{2}{3}$.

Proof

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

Open the Special Distribution Calculator and select the sine distribution. Compute a few quantiles.

Moments

Suppose that Z has the standard sine distribution. The moment generating function can be given in closed form.

The moment generating function m of Z is given by

$$m(t) = \mathbb{E}\left(e^{tZ}\right) = \frac{\pi^2 (1 + e^t)}{2(t^2 + \pi^2)}, \quad t \in \mathbb{R} \tag{5.27.5}$$

Proof

Note first that

$$m(t) = \frac{\pi}{2} \int_0^1 e^{tz} \sin(\pi z) \, dz \tag{5.27.6}$$

Integrating by parts with $u = e^{tz}$ and $dv = \sin(\pi z) dz$ gives

$$m(t) = \frac{t}{2}(1 + e^t) + \frac{t}{2} \int_0^1 e^{tz} \cos(\pi z) \, dz \tag{5.27.7}$$

Integrating by parts again with $u = e^{tz}$ and $dv = \cos(\pi z) dz$ gives

$$m(t) = \frac{t}{2}(1 + e^t) - \frac{t^2}{\pi^2} m(t) \tag{5.27.8}$$

Solving for $m(t)$ gives the result.

The moments of all orders exist, but a general formula is complicated and involves special functions. However, the mean and variance are easy to compute.

The mean and variance of Z are

1. $\mathbb{E}(Z) = 1/2$
2. $\text{var}(Z) = 1/4 - 2/\pi^2$

Proof

1. We know that the mean exists since the PDF is continuous on a bounded interval. By symmetry, the mean must be $1/2$.
2. Integration by parts (twice) gives

$$\mathbb{E}(Z^2) = \int_0^1 z^2 \frac{\pi}{2} \sin(\pi z) \, dz = \frac{1}{2} - \frac{2}{\pi^2} \tag{5.27.9}$$

   The variance then follows from the usual computational formula $\text{var}(Z) = \mathbb{E}(Z^2) - [\mathbb{E}(Z)]^2$.

Of course, the mean and variance could also be obtained by differentiating the MGF.

Numerically, $\text{sd}(Z) \approx 0.2176$.

Open the Special Distribution Simulator and select the sine distribution. Note the position and size of the mean $\pm$ standard deviation bar. Run the simulation 1000 times and compare the empirical mean and stadard deviation to the distribution mean and standard deviation.

The skewness and kurtosis of Z are

1. $\text{skew}(Z) = 0$
2. $\text{kurt}(Z) = (384 - 48\pi^2 + \pi^4)/(\pi^2 - 8)^2$

Proof

1. The skewness is 0 by the symmetry of the distribution.
2. The formula for the kurtosis follows from the usual computational formula and the first four moments: $\mathbb{E}(Z) = 1/2$, $\mathbb{E}(Z^2) = 1/2 - 2/\pi^2$, $\mathbb{E}(Z^3) = 1/2 - 3/\pi^2$, $\mathbb{E}(Z^4) = 1/2 + 24/\pi^4 - 6/\pi^2$.

Numerically, $\text{kurt}(Z) \approx 2.1938$.

Related Distributions

Since the distribution function and the quantile function have closed form representations, the standard sine distribution has the usual connection to the standard uniform distribution.

1. If U has the standard uniform distribution then $Z = G^{-1}(U) = \frac{1}{\pi}\arccos(1-2U)$ has the standard sine distribution.
2. If Z has the standard sine distribution then $U = G(Z) = \frac{1}{2}[1-\cos(\pi Z)]$ has the standard uniform distribution.

Part (a) of course leads to the random quantile method of simulation.

Open the random quantile simulator and select the sine distribution. Note the shape of the distribution and density functions. Run the simulation 1000 times and note the random quantiles. Compare the empirical density function to the probability density function.

Since the probability density function is continuous and is defined on a closed, bounded interval, the standard sine distribution can also be simulated using the rejection method.

Open the rejection method app and select the sine distribution. Run the simulation 1000 times and compare the empirical density function to the probability density function.

The General Sine Distribution

As with so many other "standard distributions", the standard sine distribution is generalized by adding location and scale parameters.

Suppose that Z has the standard sine distribution. For $a \in \mathbb{R}$ and $b \in (0, \infty)$, random variable $X = a + bZ$ has the *sine distribution* with *location parameter* a and *scale parameter* h.

Distribution Functions

Analogies of the results above for the standard sine distribution follow easily from basic properties of the location-scale transformation. Suppose that X has the sine distribution with location parameter $a \in \mathbb{R}$ and scale parameter $b \in (0, \infty)$. So X has a continuous distribution on the interval $[a, a+b]$.

The probability density function f of X is given by

$$f(x) = \frac{\pi}{2b}\sin\left(\pi\frac{x-a}{b}\right), \quad x \in [a, a+b] \tag{5.27.10}$$

1. f is symmetric about $x = a + b/2$.
2. f increases and then decreases, with mode $x = a + b/2$.
3. f is concave downward.

Proof

Recall that

$$f(x) = \frac{1}{b} g\left(\frac{x-a}{b}\right), \quad x \in \mathbb{R} \tag{5.27.11}$$

where g is the standard PDF.

Pure scale transformations ($a = 0$ and $b > 0$) are particularly common, since X often represents a *random angle*. The scale transformation with $b = \pi$ gives the angle in radians. In this case the probability density function is $f(x) = \frac{1}{2}\sin(x)$ for $x \in [0, \pi]$. Since the radian is the standard angle unit, this distribution could also be considered the "standard one". The scale transformation with $b = 90$ gives the angle in degrees. In this case, the probability density function is $f(x) = \frac{\pi}{180}\sin\left(\frac{\pi}{90}x\right)$ for $x \in [0, 90]$. This was Gilbert's original formulation.

In the special distribution simulator, select the sine distribution. Vary the 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 empirical density function to the probability density function.

The distribution function F of X is given by

$$F(x) = \frac{1}{2}\left[1 - \cos\left(\pi \frac{x-a}{b}\right)\right], \quad x \in [a, a+b] \tag{5.27.12}$$

Proof

Recall that

$$F(x) = G\left(\frac{x-a}{b}\right), \quad x \in \mathbb{R} \tag{5.27.13}$$

where G is the standard CDF.

The quantile function F^{-1} of X is given by

$$F^{-1}(p) = a + \frac{b}{\pi}\arccos(1-2p), \quad p \in (0,1) \tag{5.27.14}$$

1. The first quartile is $a + b/3$.
2. The median is $a + b/2$.
3. The third quartile is $a + 2b/3$

Proof

Recall that $F^{-1}(p) = a + bG^{-1}(p)$ for $p \in (0,1)$, where G^{-1} is the standard quantile function.

In the special distribution calculator, select the sine distribution. Vary the parameters and note the shape and location of the probability density function and the distribution function. For selected values of the parameters, find the quantiles of order 0.1 and 0.9.

Moments

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

The moment generating function M of X is given by

$$M(t) = \frac{\pi^2\left(e^{at} + e^{(a+b)t}\right)}{2\left(b^2t^2 + \pi^2\right)}, \quad t \in \mathbb{R} \tag{5.27.15}$$

Proof

Recall that $M(t) = e^{at}m(bt)$ where m is the standard MGF.

The mean and variance of X are

1. $\mathbb{E}(X) = a + b/2$
2. $\operatorname{var}(X) = b^2(1/4 - 2/\pi^2)$

Proof

By definition we can assume $X = a + bZ$ where Z has the standard sine distribution. Using the mean and variance of Z we have

1. $\mathbb{E}(X) = a + b\mathbb{E}(Z) = a + b/2$
2. $\operatorname{var}(X) = b^2\operatorname{var}(Z) = b^2(1/4 - 2/\pi^2)$

In the special distribution simulator, select the sine distribution. Vary the parameters and note the shape and location of the mean $\pm$ standard deviation bar. For selected values of the parameters, run the simulation 1000 times and compare the empirical mean and standard deviation to the distribution mean and standard deviation.

The skewness and kurtosis of X are

1. $\text{skew}(X) = 0$
2. $\text{kurt}(X) = (384 - 48\pi^2 + \pi^4)/(\pi^2 - 8)^2$

Proof

Recall that skewness and kurtosis are defined in terms of the standard score, and hence are invariant under location-scale transformations. So the skewness and kurtosis of X are the same as the skewness and kurtosis of Z.

Related Distributions

The general sine distribution is a location-scale family, so it is trivially closed under location-scale transformations.

Suppose that X has the sine distribution with location parameter $a \in \mathbb{R}$ and scale parameter $b \in (0, \infty)$, and that $c \in \mathbb{R}$ and $d \in (0, \infty)$. Then $Y = c + dX$ has the sine distribution with location parameter $c + ad$ and scale parameter bd.

Proof

Again by definition we can take $X = a + bZ$ where Z has the standard sine distribution. Then $Y = c + dX = (c + ad) + (bd)Z$.