The Semicircle Distribution

The *semicircle distribution* plays a very important role in the study of random matrices. It is also known as the *Wigner distribution* in honor of the physicist Eugene Wigner, who did pioneering work on random matrices.

The Standard Semicircle Distribution

Distribution Functions

The *standard semicircle distribution* is a continuous distribution on the interval $[-1, 1]$ with probability density function g given by

$$g(x) = \frac{2}{\pi}\sqrt{1-x^2}, \quad x \in [-1, 1] \tag{5.23.1}$$

Proof

The graph of $x \mapsto \sqrt{1-x^2}$ for $x \in [-1, 1]$ is the upper half of the circle of radius 1 centered at the origin. Hence the area under this graph is $\pi/2$ and therefore g is a valid PDF—the constant $2/\pi$ in g is the *normalizing constant*

As noted in the proof, $x \mapsto \sqrt{1-x^2}$ for $x \in [-1, 1]$ is the upper half of the circle of radius 1 centered at the origin, hence the name.

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

1. g is symmetric about $x = 0$.
2. g increases and then decreases with mode at $x = 0$.
3. g is concave downward.

Proof

As noted earlier, except for the normalizing constant, the graph of g is the upper half of the circle of radius 1 centered at the origin, and so these properties are obvious.

Open special distribution simulator and select the semicircle distribution. With the default parameter value, note the shape of the probability density function. Run the simulation 1000 times and compare the empirical density function to the probability density function.

The standard semicircle distribution function G is given by

$$G(x) = \frac{1}{2} + \frac{1}{\pi}x\sqrt{1-x^2} + \frac{1}{\pi}\arcsin x, \quad x \in [-1, 1] \tag{5.23.2}$$

Proof

Of course $G(x) = \int_{-1}^x g(t)\,dt$ for $-1 \le x \le 1$. The integral is evaluated by using the trigonometric substitution $t = \sin\theta$.

We cannot give the quantile function G^{-1} in closed form, but values of this function can be approximated. Clearly by symmetry, $G^{-1}\left(\frac{1}{2} - p\right) = -G^{-1}\left(\frac{1}{2} + p\right)$ for $0 \le p \le \frac{1}{2}$. In particular, the median is 0.

Open the special distribution simulator and select the semicircle distribution. With the default parameter value, note the shape of the distribution function. Compute the first and third quartiles.

Moments

Suppose that X has the standard semicircle distribution. The moments of X about 0 can be computed explicitly. In particular, the odd order moments are 0 by symmetry.

For $n \in \mathbb{N}$, the moment of order $2n+1$ is $\mathbb{E}\left(X^{2n+1}\right) = 0$ and the moment of order $2n$ is

$$\mathbb{E}\left(X^{2n}\right) = \left(\frac{1}{2}\right)^{2n} \frac{1}{n+1} \binom{2n}{n} \tag{5.23.3}$$

Proof

Clearly X has moments of all orders since the PDF g is bounded and the support interval is bounded. So by symmetry, the odd order moments are 0, and we just need to prove the result for the even order moments. Note that

$$\mathbb{E}\left(X^{2n}\right) = \int_{-1}^{1} x^{2n} \frac{2}{\pi} \sqrt{1-x^2}\, dx \tag{5.23.4}$$

We use the substitution $x = \sin\theta$ to get

$$\mathbb{E}\left(X^{2n}\right) = \int_{-\pi/2}^{\pi/2} \frac{2}{\pi} \sin^{2n}(\theta) \cos^2(\theta)\, d\theta \tag{5.23.5}$$

This integral can be evaluated by standard calculus methods to give the result above.

The numbers $C_n = \frac{1}{n+1}\binom{2n}{n}$ for $n \in \mathbb{N}$ are known as the *Catalan numbers*, and are named for the Belgian mathematician Eugene Catalan. In particular, we can compute the mean, variance, skewness, and kurtosis.

The mean and variance of X are

1. $\mathbb{E}(X) = 0$
2. $\text{var}(X) = \frac{1}{4}$

Open the special distribution simulator and select the semicircle distribution. With the default parameter value, note the size and location of the mean $\pm$ standard deviation bar. Run the simulation 1000 times and compare the empirical mean and standard deviation to the true mean and standard deviation.

The skewness and kurtosis of X are

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

Proof

The standard score of X is $2X$. Hence $\text{skew}(X) = E\left(2^3 X^3\right) = 0$. Of course, this is also clear from the symmetry of the distribution of X. Similarly, by the moment formula,

$$\text{kurt}(X) = \mathbb{E}\left(2^4 X^4\right) = 2^4 \left(\frac{1}{2}\right)^4 \frac{1}{3} \binom{4}{2} = 2 \tag{5.23.6}$$

It follows that the *excess kurtosis* is $\text{kurt}(X) - 3 = -1$.

Related Distributions

The semicircle distribution has simple connections to the continuous uniform distribution.

If (X, Y) is uniformly distributed on the circular region in $\mathbb{R}^2$ centered at the orgin with radius 1, then X and Y each have the standard semicircular distribution.

Proof

(X, Y) has joint PDF $(x, y) \mapsto 1/\pi$ on $C = \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 \le 1\}$. Hence X has PDF

$$g(x) = \int_{-\sqrt{1-x^2}}^{\sqrt{1-x^2}} \frac{1}{\pi}\, dy = \frac{2}{\pi}\sqrt{1-x^2}, \quad x \in [-1, 1] \tag{5.23.7}$$

It's easy to simulate a random point that is uniformly distributed on circular region in the previous theorem, and this provides a way of simulating a standard semicircle distribution. This is important since we can't use the random quantile method of simulation.

Suppose that U, V, and W are independent random variables, each with the standard uniform distribution (random numbers). Let $R = \max\{U, V\}$ and $\Theta = 2\pi W$, and then let $X = R\cos\Theta$, $Y = R\sin\Theta$. Then (X, Y) is uniformly distributed on the circular region of radius 1 centered at the origin, and hence X and Y each have the standard semicircle distribution.

Proof

U and V have CDF $u \mapsto u$ for $u \in [0, 1]$ and therefore R has CDF $r \mapsto r^2$ for $r \in [0, 1]$. Hence R has PDF $r \mapsto 2r$ for $r \in [0, 1]$. On the other hand, Θ is uniformly distributed on $[0, 2\pi)$ and hence has density $\theta \mapsto 1/2\pi$ on $[0, 2\pi)$. By independence, the Joint PDF of (R, Θ) is $(r, \theta) \mapsto (2r)(1/2\pi) = r/\pi$ on $\{(r, \theta) : 0 \le r \le 1, 0 \le \theta \le 2\pi\}$. For the polar coordinate transformation $(x, y) \mapsto (r\cos\theta, r\sin\theta)$, the Jacobian is r. Hence by the change of variables theorem, (X, Y) has PDF

$$(x, y) \mapsto \frac{r}{\pi}\frac{1}{r} = \frac{1}{\pi} \text{ on } \{(x, y) \in \mathbb{R}^2 : x^2 + y^2 \le 1\} \tag{5.23.8}$$

Of course, note that X and Y in the previous theorem are not independent. Another method of simulation is to use the rejection method. This method works well since the semicircle distribution has a bounded support interval and a bounded probability density function.

Open the rejection method app and select the semicircle distribution. Keep the default parameters to get the standard semicirle distribution. Run the simulation 1000 times and note the points in the scatterplot. Compare the empirical density function, mean, and standard deviation to their distributional counterparts.

The General Semicircle Distribution

Like so many *standard distributions*, the standard semicircle distribution is usually generalized by adding location and scale parameters.

Definition

Suppose that Z has the standard semicircle distribution. For $a \in \mathbb{R}$ and $r \in (0, \infty)$, $X = a + rZ$ has the *semicircle distribution* with *center* (location parameter) a and *radius* (scale parameter) r.

Distribution Functions

Suppose that X has the semicircle distribution with center $a \in \mathbb{R}$ and radius $r \in (0, \infty)$.

X has probability density function f given by

$$f(x) = \frac{2}{\pi r^2}\sqrt{r^2 - (x - a)^2}, \quad x \in [a - r, a + r] \tag{5.23.9}$$

Proof

This follows from a standard result for location-scale families. Recall that

$$f(x) = \frac{1}{r} g\left(\frac{x - a}{r}\right), \quad \frac{x - a}{r} \in [-1, 1] \tag{5.23.10}$$

where g is the standard semicircle PDF.

The graph of $x \mapsto \sqrt{r^2 - (x - a)^2}$ for $x \in [a - r, a + r]$ is the upper half of the circle of radius r centered at a. The area under this semicircle is $\pi r^2/2$ so as a check on our work, we see that f is a valid probability density function.

The probability density function f of X satisfies the following properties:

1. f is symmetric about $x = a$.
2. f increases and then decreases with mode at $x = a$.

3. f is concave downward.

Open special distribution simulator and select the semicircle distribution. Vary the center a and the radius r, and note the shape of the probability density function. For selected values of a and r, run the simulation 1000 times and compare the empirical density function to the probability density function.

The distribution function F of X is

$$F(x) = \frac{1}{2} + \frac{x-a}{\pi r^2}\sqrt{r^2-(x-a)^2} + \frac{1}{\pi}\arcsin\left(\frac{x-a}{r}\right), \quad x \in [a-r, a+r] \tag{5.23.11}$$

Proof

This follows from a standard result for location-scale families:

$$F(x) = G\left(\frac{x-a}{r}\right), \quad \frac{x-a}{r} \in [-1, 1] \tag{5.23.12}$$

where G is the standard semicircle CDF.

As in the standard case, we cannot give the quantile function F^{-1} in closed form, but values of this function can be approximated. Recall that $F^{-1}(p) = a + rG^{-1}(p)$ where G^{-1} is the standard semicircle quantile function. In particular, $F^{-1}\left(\frac{1}{2}-p\right) = 2a - F^{-1}\left(\frac{1}{2}+p\right)$ for $0 \le p \le \frac{1}{2}$. The median is a.

Open the special distribution simulator and select the semicircle distribution. Vary the center a and the radius r, and note the shape of the distribution function. For selected values of a and r, compute the first and third quartiles.

Moments

Suppose again that X has the semicircle distribution with center $a \in \mathbb{R}$ and radius $r \in (0, \infty)$, so by definition we can assume $X = a + rZ$ where Z has the standard semicircle distribution. The moments of X can be computed from the moments of Z. Using the binomial theorem and the linearity of expected value we have

$$\mathbb{E}(X^n) = \sum_{k=0}^n \binom{n}{k} r^k a^{n-k} \mathbb{E}(Z^k), \quad n \in \mathbb{N} \tag{5.23.13}$$

In particular,

The mean and variance of X are

1. $\mathbb{E}(X) = a$
2. $\text{var}(X) = r^2/4$

When the center is 0, the general moments have a simple form:

Suppose that $a = 0$. For $n \in \mathbb{N}$ the moment of order $2n+1$ is $\mathbb{E}\left(X^{2n+1}\right) = 0$ and the moment of order $2n$ is

$$\mathbb{E}\left(X^{2n}\right) = \left(\frac{r}{2}\right)^{2n} \frac{1}{n+1}\binom{2n}{n} \tag{5.23.14}$$

Proof

This follows from the moment results for Z since $X^m = r^m Z^m$ for $m \in \mathbb{N}$.

Open the special distribution simulator and select the semicircle distribution. Vary the center a and the radius r, and note the size and location of the mean $\pm$ standard deviation bar. For selected values of a and r, 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. $\text{kurt}(X) = 2$

Proof

These results follow immediately from the skewness and kurtosis of the standard distribution. Recall that skewness and kurtosis are defined in terms of the standard score, which is independent of the location and scale parameters..

Once again, the excess kurtosis is $\text{kurt}(X) - 3 = -1$.

Related Distributions

Since the semicircle distribution is a location-scale family, it's invariant under location-scale transformations.

Suppose that X has the semicircle distribution with center $a \in \mathbb{R}$ and radius $r \in (0, \infty)$. If $b \in \mathbb{R}$ and $c \in (0, \infty)$ then $b + cX$ has the semicircle distribution with center $b + ca$ and radius cr.

Proof

Again from the definition we can take $X = a + rZ$ where Z has the standard semicircle distribution. Then $b + cX = (b + ca) + (cr)Z$.

One member of the beta family of distributions is a semicircle distribution:

The beta distribution with left parameter $3/2$ and right parameter $3/2$ is the semicircle distribution with center $1/2$ and radius $1/2$.

Proof

By definition, the beta distribution with left and right parameters $3/2$ has PDF

$$f(x) = \frac{1}{B(3/2, 3/2)} x^{1/2}(1-x)^{1/2}, \quad x \in [0, 1] \tag{5.23.15}$$

But $B(3/2, 3/2) = \pi/8$ and $x^{1/2}(1-x)^{1/2} = \sqrt{x - x^2}$. Completing the square gives

$$f(x) = \frac{8}{\pi} \sqrt{\frac{1}{4} - \left(x - \frac{1}{2}\right)^2}, \quad x \in [0, 1] \tag{5.23.16}$$

which is the PDF of the semicircle distribution with center $1/2$ and radius $1/2$.

Since we can simulate a variable Z with the standard semicircle distribution by the method above, we can simulate a variable with the semicircle distribution with center $a \in \mathbb{R}$ and radius $r \in (0, \infty)$ by our very definition: $X = a + rZ$. Once again, the rejection method also works well since the support and probability density fucntion of X are bounded.

Open the rejection method app and select the semicircle distribution. For selected values of a and r, run the simulation 1000 times and note the points in the scatterplot. Compare the empirical density function, mean and standard deviation to their distributional counterparts.