Journal of Statistical Theory and Applications, Vol. 15, No. 3 (September 2016), 207-213

Some Inferences on Semicircular Distribution

M. Ahsanullah

Rider University, 2083 Lawrenceville Road, Lawrenceville,
New Jersey, NJ 08648, USA
ahsan@rider.edu

Received 10 November 2015

Accepted 14 April 2016

Abstract

Several distributional properties of semicircular distribution are presented. The limiting distributions of the standardized extreme order statistics are given. Some characterizations of the distribution are shown.

1. Introduction

Wishart (1928) considered random matrix in connection to the statistical analysis of large data. A matrix is called a random matrix if the entries of the matrix are random numbers from a specified distribution. If the distribution is Gaussian, then we call it a Gaussian random matrix. Let $\lambda_i, i = 1,2,...,n$ be the eigenvalues of the matrix. The empirical spectral density (ESD) $m(\lambda)$ is defined as

$$m(\lambda) = \frac{1}{n}\sum_{i=1}^{n} \delta(\lambda - \lambda_i)$$

where δ is the Dirac delta function. The ESD is a probability measure over the complex plane. If the matrix is Hermitian, the eigenvalues are real and $m(\lambda)$ is a real probability density function. Wagner (1955) proved that the ESD of a Hermitian $n\times n$ Gaussian matrix when normalized by $\frac{1}{\sqrt{n}}$ tends to the semicircular distribution with the following pdf $f_1(x)$ where

$$f_1(x) = \frac{1}{2\pi}\sqrt{4-x^2}, -2 < x < 2.$$

This work has been extended by various authors. For details see Anderson et al. (2010), Bai and Yin (1988) and Tao (2012) .The pdf $f_R(x)$ of the generalized semicircular distribution is given by

$$\frac{2}{\pi R^2}\sqrt{R^2-x^2}\ , -R \le x \le R.$$

The pdf of the standardized SCD(–1, 1) $f_s(x)$ is given by

$$f_s(x) = \frac{2}{\pi}\sqrt{1-x^2}\ , -1 \le x \le 1.$$

Published by Atlantis Press

Suppose Y has a beta distribution with pdf $f(x)$ as

$$f_B(y) = \sqrt{y(1-y)}\,, \; 0 < y < 1,$$

then $W = 2y - 1$ is distributed as SCD(–1, 1). The standard semicircular distribution is symmetric around zero. The mean of SCD(–1, 1) is zero and variance = ¼.

2. Main Results

The cumulative distribution function (CDF) $F_s(x)$ of SCD(–1, 1) is given by

$$F_s(x) = \frac{1}{2\pi}(\pi + 2\arcsin x + 2x\sqrt{1-x^2})\,.$$

Figure 2.1 gives the pdf of SCD(–1, 1)-Black, SCD(–3, 3)-Red, SCD(–6, 6)-Green.

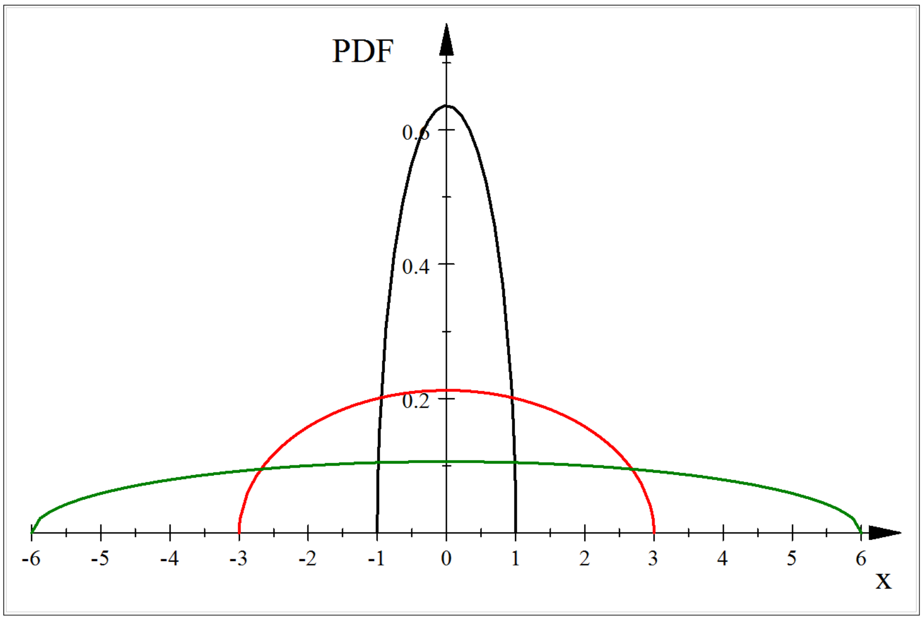

Figure 2.1.

The percentile points x_p of SCD (–1, 1) are given in Table 2.1.

Table 2.1. Percentile points of SCD(–1, 1).

p	x_p
0.1	–0.68705
0.2	–0.49186
0.3	–0.31969
0.4	–0.15774
0.5	0.0
0.6	0.15774
0.7	0.31969
0.8	0.49186
0.9	0.68705

The hazard rate $h(z)$ of SCD(–1, 1) is

$$h(z) = \frac{4\sqrt{1-x^2}}{\pi - 2\arcsin x - 2x\sqrt{1-x^2}}$$

The hazard rate is monotone increasing. The moment generating function $M(t)$ is

$$M(t) = \int_{-1}^{1} e^{tx} \frac{1}{2\pi}\sqrt{1-x^2}\,dx = \frac{2I_1(t)}{t}$$

where $I_1(t)$ is the modified Bessel function. Let $X_1, X_2, \ldots, X_n$ be an independent observation from SCD(–1, 1). Suppose $X_{1,n} \le X_{2,n} \le \ldots \le X_{n,n}$ is the associated order statistics. It can be shown that

$$\lim_{n\to\infty} \frac{F^{-1}(\frac{1}{n}) - F^{-1}(\frac{2}{n})}{F^{-1}(\frac{2}{n}) - F^{-1}(\frac{4}{n})} = 2^{-1./2}$$

Thus $X_{1,n}$ when normalized tends to the extreme Value type III (minimum) distribution (for details see Ahsanullah and Nevzorov (2001). Further

$$\lim_{n\to\infty} P(X_{1,n} \le a_n + b_n x) = 1 - e^{-\sqrt{x}}, x \ge 0,\ a_n = -1 \text{ and } b_n = n^{-2/3}$$

.It can also be shown that

$$\lim_{n\to\infty} \frac{F^{-1}(1-\frac{1}{n}) - F^{-1}(1-\frac{2}{n})}{F^{-1}(1-\frac{2}{n}) - F^{-1}(1-\frac{4}{n})} = 2^{-1./2}$$

Thus $X_{n,n}$ when normalized tends to the extreme Value type III (maximum) distribution. We have

$$\lim_{n\to\infty} P(X_{n,n} \le a_n + b_n x) = e^{-\sqrt{(-x)}}, x \le 0,\ a_n = -1 \text{ and } b_n = n^{-2/3}.$$

The following two lemmas will be used to prove the characterizing theorems.

Lemma 2.1. *Suppose the random variable X has an absolutely continuous cdf $F(x)$ and pdf $f(x)$. We assume $\alpha = \inf\{X \mid F(x) > 0\}$, $\beta = \sup\{x \mid F(x) < 1\}$, $E(X)$ exists and $f(x)$ is differentiable for all x in (α, β). If $E(X \mid X \le x) = g(x)\,\tau(x)$, $g(x)$ is a differentiable function for all x in (α, β) and $\tau(x) = \frac{f(x)}{F(x)}$, then*

$$f(x) = ce^{\int \frac{x - g'(x)}{g(x)} dx},$$

where c is determined by the condition

$$\int_{\alpha}^{\beta} f(x)dx = 1.$$

Proof. We have

$$E(X \mid X \le x) = g(x)\,\tau(x)$$

and

$$g(x) = \frac{\int_\alpha^x uf(u)du}{f(x)}.$$

Thus

$$\int_\alpha^x uf(u)du = g(x)f(x).$$

Differentiating the above expression with respect to x, we obtain

$$xf(x) = f'(x)g(x) + f(x)g'(x).$$

Thus

$$\frac{f'(x)}{f(x)} = \frac{x - g'(x)}{g(x)}.$$

On integrating the above expression, we will have

$$f(x) = ce^{\int \frac{x-g'(x)}{g(x)}dx},$$

where c is determined by the condition

$$\int_\alpha^\beta f(x)dx = 1.$$

Lemma 2.2. *Suppose the random variable X has an absolutely continuous cdf $F(x)$ and pdf $f(x)$. We assume* $\alpha = \inf\{X|F(x) > 0\}$, $\beta = \sup\{x|F(x) < 1\}$, *$E(X)$ exists and $f(x)$ is differentiable for all x in (α, β). If $E(X|X \geq x)$* $= h(x)r(x)$, *$h(x)$ is a differentiable function for all x in (α, β) and $r(x) = \frac{f(x)}{1-F(x)}$, then*

$$f(x) = ce^{-\int \frac{x+h'(x)}{h(x)}dx},$$

where c is determined by the condition

$$\int_\alpha^\beta f(x)dx = 1.$$

Proof. We have

$$E(X \mid X \geq x) = h(x)r(x)$$

and

$$h(x) = \frac{\int_x^\beta uf(u)du}{f(x)}.$$

Thus

$$\int_x^\beta uf(u)du = h(x)f(x).$$

Differentiating the above expression with respect to x, we obtain

$$xf(x) = f'(x)h(x) + f(x)h'(x).$$

Thus

$$\frac{f'(x)}{f(x)} = -\frac{x + h'(x)}{h(x)}$$

On integrating the above expression, we will have

$$f(x) = ce^{-\int \frac{x-h'(x)}{h(x)}dx},$$

where c is determined by the condition

$$\int_{\alpha}^{\beta} f(x)dx = 1.$$

Theorem 2.1. *Suppose the random variable X has an absolutely continuous cdf $F(x)$ and pdf $f(x)$. We assume $\alpha = -1$ and $\beta = 1$, $f(x)$ is differentiable for all x in $(-1, 1)$ and $E(X)$ exists. Then*

$$E(X \mid X \leq x) = g(x)\frac{f(x)}{F(x)},$$

where

$$g(x) = \frac{x^2 - 1}{3} \text{ if and only if } f(x) = \frac{2}{\pi}\sqrt{1 - x^2}, -1 \leq x \leq 1.$$

Proof. If

$$f(x) = \frac{2}{\pi}\sqrt{1 - x^2},$$

then

$$g(x) = \frac{\int_{-1}^{x} u\sqrt{1-u^2}}{\sqrt{1-x^2}} = \frac{x^2 - 1}{3}.$$

Suppose

$$g(x) = \frac{\int_{-1}^{x} u\sqrt{1-u^2}}{\sqrt{1-x^2}} = \frac{x^2 - 1}{3},$$

then

$$g'(x) = 2x/3 \text{ and } \frac{x - g'(x)}{g(x)} = \frac{-x}{1 - x^2}.$$

We have by Lemma 2.1,

$$\frac{f'(x)}{f(x)} = -\frac{x}{1 - x^2}.$$

On integrating both sides of the above equation, we obtain

$$F(x) = c\sqrt{1 - x^2},$$

where c is a constant. Using the boundary condition $\int_{-1}^{1} f(x)dx = 1$, we obtain

$$f(x) = \frac{2}{\pi}\sqrt{1-x^2}, -1 \le x \le 1.$$

Theorem 2.2. *Suppose the random variable X has an absolutely continuous cdf F*(x) *and pdf f*(x). *We assume,* $\alpha = -1$ *and* $\beta = 1$, $f(x)$ *is differentiable for all x in* $(-1, 1)$ *and E*(X) *exists. Then*

$$E(X \mid X > x) = h(x)\frac{f(x)}{F(x)},$$

where

$$h(x) = \frac{x^2-1}{3} \text{ if and only if } f(x) = \frac{2}{\pi}\sqrt{1-x^2}, -1 \le x \le 1.$$

Proof. If

$$f(x) = \frac{2}{\pi}\sqrt{1-x^2},$$

then

$$h(x) = \frac{\int_x^1 u\sqrt{1-u^2}}{\sqrt{1-x^2}} = \frac{1-x^2}{3}.$$

Suppose

$$h(x) = \frac{1-x^2}{3},$$

then

$$h'(x) = -2x/3 \text{ and } -\frac{x+h'(x)}{h(x)} = \frac{-x}{1-x^2}.$$

We have by Lemma 2.2,

$$\frac{f'(x)}{f(x)} = -\frac{x}{1-x^2}.$$

On integrating both sides of the above equation, we obtain

$$F(x) = c\sqrt{1-x^2},$$

where c is a constant. Using the boundary condition $\int_{-1}^{1} f(x)dx = 1$, we obtain

$$f(x) = \frac{2}{\pi}\sqrt{1-x^2}, -1 \le x \le 1.$$

We define

$$S_{k,n} = \frac{1}{k}(X_{1,n} + X_{2,n} + \cdots + X_{k,n})$$

and

$$T_{k,n} = \frac{1}{k-1}(X_{k+1,n} + X_{k+2,n} + \cdots + X_{n,n}).$$

The following two characterization theorems are based on the conditional expectation of $S_{k,n}$ and $T_{k,n}$.

Theorem 2.3. *Suppose the random variable X has an absolutely continuous cdf $F(x)$ and pdf $f(x)$. We assume $\alpha = -1$ and $\beta = 1$, $f(x)$ is differentiable for all x in $(-1, 1)$ and $E(X)$ exists. Then*

$$E(S_{k,n} \mid X_{k,n} \le x) = g(x)\frac{f(x)}{F(x)},$$

where

$$g(x) = \frac{1-x^2}{3} \text{ if and only if } f(x) = \frac{2}{\pi}\sqrt{1-x^2}, -1 \le x \le 1.$$

Proof. We have

$$E(S_{k,n} \mid X_{k,n} = x) = E(X \mid X \le x),$$

thus the proof follows from Theorem 2.1.

Theorem 2.4. *Suppose the random variable X has an absolutely continuous cdf $F(x)$ and pdf $f(x)$. We assume $\alpha = -1$ and $\beta = 1$, $f(x)$ is differentiable for all x in $(-1, 1)$ and $E(X)$ exists. Then*

$$E(T_{k,n} \mid X_{k,n} = x) = h(x)\frac{f(x)}{F(x)},$$

where

$$h(x) = \frac{1-x^2}{3} \text{ if and only if } f(x) = \frac{2}{\pi}\sqrt{1-x^2}, -1 \le x \le 1.$$

Proof. We have

$$E(T_{k,n} \mid X_{k,n} = x) = E(X \mid X \ge x),$$

thus the proof follows from Theorem 2.2.

References

[1] M. Ahsanullah, V.B. Nevzorov, and M. Shakil, *An Introduction to Order Statistics* (Atlantis-Press, Paris , France, 2015).
[2] M. Ahsanullah and V.B. Nevzorov, *Ordered Random Variables* (Nova Science Publishers Inc., New York, USA, 2001).
[3] T.W. Anderson, A. Guiomet and O. Zeitoumi, *An Introduction to Random Matrix* (Cambridge University Press, Cambridge, U.K., 2010).
[4] Z.D. Bai and V. Yin, Converegence to Semi Circular Law, in *Annals of Probability* **16** (1988) 1729-1741.
[5] T. Tao, *Topics in Random Matrix Theory* (American Mathematical Society, Providence, R.I., USA, 2012).
[6] E.P. Wagner, Characteristic vectors of borthered matrices with infinite dimensions, *Annals of Mathematics* **52**(2) (1955) 545-564.
[7] J. Wishart, The generalized product moment distribution in samples from a normal distribution, *Biometrika* **28A** (1928) 32-52.