Communications for Statistical Applications and Methods
2014, Vol. 21, No. 2, 183–191

DOI: http://dx.doi.org/10.5351/CSAM.2014.21.2.183
ISSN 2287-7843

A Clarification of the Cauchy Distribution

Hwi-Young Lee[a], Hyoung-Jin Park[a], Hyoung-Moon Kim[1,a]

[a]Department of Applied Statistics, Konkuk University, Korea

Abstract

We define a multivariate Cauchy distribution using a probability density function; subsequently, a Ferguson's definition of a multivariate Cauchy distribution can be viewed as a characterization theorem using the characteristic function approach. To clarify this characterization theorem, we construct two dependent Cauchy random variables, but their sum is not Cauchy distributed. In doing so the proofs depend on the characteristic function, but we use the cumulative distribution function to obtain the exact density of their sum. The derivation methods are relatively straightforward and appropriate for graduate level statistics theory courses.

Keywords: Cauchy distribution, dependency, linear combination, characteristic function, distribution function.

1. Introduction

The Cauchy distribution is an example of a distribution which has no mean, variance or higher moments defined. Hence it has no moment generating function (mgf). The sample mean will have the same standard Cauchy distribution if $X_1, \ldots, X_n$ are independent and identically distributed random variables with a standard Cauchy distribution. This example serves to show that the hypothesis of finite variance in the central limit theorem cannot be dropped. For the applications of the multivariate Cauchy distribution, see Kotz and Nadarajah (2004) and Crovella *et al.* (1998). Recently Park and Bera (2009) well summarized the maximum entropy probability distributions. Specifically the standard Cauchy distribution is the maximum entropy probability distribution for a random variable X for which $E(\ln(1 + X^2)) = \ln(4)$. Zhang (2010) proposed a new unbiased L-estimator based on order statistics. Originally, Ferguson (1962) defined a multivariate Cauchy distribution as follows:

"A random vector X is said to have a multivariate Cauchy distribution if, and only if, for every real vector b, the random variable $b^\top X$ has a Cauchy distribution. The distribution is said to be symmetric if the mass is distributed symmetrically with respect to some point in p-dimensional space."

In addition, characteristic function (cf) and the cumulative distribution function (cdf) approaches for the Cauchy distribution are also possible. Presumably, the easiest approach is using the probability density function (pdf) which is commonly adopted in modern mathematical statistics because the multivariate Cauchy density is equal to that of the multivariate t distribution when the degrees of freedom is one.

Thus, a multivariate Cauchy distribution is defined as follows (Kotz and Nadarajah, 2004). Suppose a p-dimensional random vector $X = (X_1, \ldots, X_p)^\top$ follows a multivariate Cauchy distribution,

The paper was supported by Konkuk University in 2013.
[1] Corresponding author: Professor, Department of Applied Statistics, Konkuk University, Seoul 143-701, Korea.
E-mail: hmkim@konkuk.ac.kr

Published online 31 March 2014 / journal homepage: http://csam.or.kr

that is, $X \sim \text{Cauchy}_p(\mu, \Sigma)$; then, its density function is given by

$$f_X(x;\mu,\Sigma) = \frac{\Gamma\left(\frac{1+p}{2}\right)}{\Gamma\left(\frac{1}{2}\right)\pi^{\frac{p}{2}}|\Sigma|^{\frac{1}{2}}\left[1+(x-\mu)^\top\Sigma^{-1}(x-\mu)\right]^{\frac{1+p}{2}}}, \quad x \in \mathbb{R}^p, \tag{1.1}$$

where $\mu \in \mathbb{R}^p$ is a location vector and Σ is a positive-definite scale matrix. When $p = 1$, its density function is

$$f_X\left(x;\mu,\sigma^2\right) = \frac{1}{\pi}\left[\frac{\sigma}{(x-\mu)^2+\sigma^2}\right], \quad x \in \mathbb{R}. \tag{1.2}$$

Furthermore, the special case when $\mu = 0$ and $\sigma = 1$ is called the standard Cauchy distribution.

The characteristic function (Kotz and Nadarajah, 2004) of (1.1) is obtained as follows:

$$\Psi_X(t) = \exp\left(it^\top\mu - \left\|\Sigma^{\frac{1}{2}}t\right\|\right), \quad \forall\, t \in \mathbb{R}^p, \tag{1.3}$$

where $\| t \| = \sqrt{t^\top t}$ and $i = \sqrt{-1}$. Therefore the cf of (1.2) is $\Psi_X(t) = \exp(i\mu t - \sigma|t|)$, $\forall\, t \in \mathbb{R}$.

We first argue that Ferguson's definition of a multivariate Cauchy distribution can be viewed as a characterization theorem using a characteristic function approach as in the case of the multivariate normal distribution. Ferguson's definition of a multivariate Cauchy distribution may give a wrong impression. Every component of a random vector is Cauchy since each component can be expressed as a linear combination; therefore, the sum of every component is itself Cauchy. This fact may give an illusion that a sum of Cauchy random variables should be Cauchy such as in the case of independent Cauchy random variables. Similarly, this argument can be applied to the normal distribution. Even for the normal distribution, this is a problem as in Melnick and Tenenbein (1982).

However, we could not find any clear example of it: as far as we know, examples are abundant in the multivariate normal case (Romano and Siegel, 1986; Novosyolov, 2006, among others). Even though for the normal case, many students in statistics courses obtain a mistaken impression concerning the property of the normal distribution. Therefore the goal of this paper is to present an informative and intriguing example, where the sum of two marginal univariate Cauchy random variables is not Cauchy for clarifying a mistaken impression about the Cauchy distribution. In doing so, some derivation methods are relatively straightforward using the cf and cdf of the Cauchy distribution. Hence this material is appropriate for graduate level statistics theory courses.

The rest of this paper is organized as follows. Section 2 describes the problem setup to construct random variables that are marginally univariate Cauchy distributed, but whose sum is not Cauchy distributed. In Section 3, we show that the random variables given in Section 2 fulfill our assertion. Furthermore some interesting results are mentioned as remarks. Finally concluding remarks are given in Section 4.

2. Problem Setup

Suppose that a p-dimensional random vector $X = (X_1, \ldots, X_p)^\top$ follows a multivariate Cauchy distribution, that is, $X \sim \text{Cauchy}_p(\mu, \Sigma)$; then, its density function is given by (1.1). We characterize a multivariate Cauchy random vector X using every linear combination of the components of X in the following theorem which is usually done in the case of the multivariate normal distribution. The proof depend on the cf since there is no mgf for Cauchy distribution.

Theorem 1. (Characterization) *$X \sim Cauchy_p(\mu, \Sigma)$ if and only if $a^\top X \sim Cauchy(a^\top\mu, a^\top\Sigma a)$, $\forall a \neq 0 \in \mathbb{R}^p$.*

Proof: Suppose that $X \sim \text{Cauchy}_p(\mu, \Sigma)$, then its cf is given by (1.3). Thus the cf of $a^\top X$ is $\Psi_{a^\top X}(t) = \Psi_X(at) = \exp(ita^\top\mu - \sqrt{a^\top\Sigma a}|t|)$. For the other part, observe that if $a^\top X \sim \text{Cauchy}(a^\top\mu, a^\top\Sigma a)$, $\forall a \neq 0 \in \mathbb{R}^p$, then we have $\Psi_{a^\top X}(t) = \exp(ita^\top\mu - \sqrt{a^\top\Sigma a}|t|)$, $\forall t \in \mathbb{R}$. Now given that $\Psi_{a^\top X}(t) = \Psi_X(at)$, taking $t = 1$ and comparing with $\Psi_X(t)$ in (1.3), we have $X \sim \text{Cauchy}_p(\mu, \Sigma)$ since a is arbitrary. □

This is well-known in the case of the normal distribution, but we were unable to find any direct proof that uses the cf of the Cauchy distribution. To clarify this characterization theorem, we start with two simple examples to show that the sum of two marginal univariate Cauchy random variables is not Cauchy. Let X be a univariate standard Cauchy random variable. Then $-X$ is also a univariate standard Cauchy random variable but $X - X = 0$. This is a simple example to roughly illustrate that linear combinations of Cauchy random variables need not be Cauchy distribution if the degenerate distribution is not included in the Cauchy distribution. To be more general, let $X \sim \text{Cauchy}(\mu, \sigma^2)$, then $Y = 2\mu - X \sim \text{Cauchy}(\mu, \sigma^2)$ and $X + Y = X + 2\mu - X = 2\mu$ is not Cauchy distributed by the same reason as above.

However these examples are not so informative since the degenerate distribution should not be included in the Cauchy distribution. An example about the normal distribution from Romano and Siegel (1986) can be extended to the Cauchy distribution. Let $X \sim \text{Cauchy}(0, 1)$. Observe X; then toss a fair coin and define Y as follows:

$$Y = \begin{cases} X, & \text{if the toss is "heads"}, \\ -X, & \text{if the toss is "tails"}. \end{cases}$$

By symmetry of X about 0, $Y \sim \text{Cauchy}(0, 1)$. However the sum $X + Y$ is a mixture of a discrete and a continuous distribution, so it cannot have a Cauchy distribution. This is not degenerate because it is continuously distributed whenever the toss is heads, but too tricky. Therefore we develop an informative and intriguing example to clarify a mistaken impression about Cauchy distribution.

Now we construct random variables such that X_1 and X_2 are univariate Cauchy, but their sum, $X_1 + X_2$, is not a Cauchy random variable. Suppose that X_1 and Z are independent univariate standard Cauchy random variables. This means that the joint distribution of X_1 and Z is standard bivariate Cauchy. Consider a new random variable X_2 such that

$$X_2 = \begin{cases} |Z|, & X_1 \geq 0, \\ -|Z|, & X_1 < 0. \end{cases} \tag{2.1}$$

Then, the random variable X_2 has a standard univariate Cauchy distribution, but the joint distribution of (X_1, X_2) is not bivariate Cauchy. The latter can be shown using the definition of X_2. Since X_1 and X_2 always have the same sign by construction, they cannot be bivariate Cauchy. Moreover the distribution of $X_1 + X_2$ is not Cauchy.

In the following section we will prove the following two statements:

- The random variable X_2 is distributed as a Cauchy distribution.

- The distribution of $X_1 + X_2$ is not a Cauchy distribution.

3. Completing the Example

3.1. X_2 is Cauchy

We show that the random variable X_2 defined by (2.1) has the same distribution, that of the standard Cauchy random variable Z.

Theorem 2. *The cf of X_2 defined by* (2.1) *is equal to that of Z.*

Proof: Denote $\Psi_{|Z|}(t) = a(t)+ib(t)$ as the cf of $|Z|$, where $a(t) = E\cos(t|Z|)$ and $b(t) = E\sin(t|Z|)$. Then the cf of Z becomes $\Psi_Z(t) = a(t)$ since the property of trigonometric functions. Furthermore, $\Psi_{X_2}(t) = E_{X_2}(e^{itX_2}) = E_{X_1}[E_{X_2}(e^{itX_2}|X_1)]$, which becomes $\Psi_{|Z|}(t)/2 + \Psi_{|Z|}(-t)/2$ by the construction (2.1). The result follows using the property of cf and $a(t)$ and $b(t)$ are even and odd functions, respectively. □

3.2. $X_1 + X_2$ is not Cauchy

We show that the random variable $X_1 + X_2$, where X_2 defined by (2.1), is not Cauchy distributed using the cf of it.

Theorem 3. *The cf of $X_1 + X_2$, where X_2 defined by* (2.1)*, can't be the form of cf in Cauchy.*

Proof: Denote $\Psi_{|Z|}(t) = a(t) + ib(t)$ as the cf of $|Z|$, where $a(t) = E\cos(t|Z|)$ and $b(t) = E\sin(t|Z|)$. Then the cf of X_1 becomes $\Psi_{X_1}(t) = a(t)$ since the property of trigonometric functions, $E[e^{itX_1}I(X_1 \geq 0)] = \Psi_{|Z|}(t)/2$ and $E[e^{itX_1}I(X_1 < 0)] = \Psi_{|Z|}(-t)/2$. Note that

$$\begin{aligned}\Psi_{X_1+X_2}(t) &= E\left[e^{it(X_1+X_2)}\right] = E_{X_1}\left[E_{X_2}\left(e^{it(X_1+X_2)}\middle|X_1\right)\right]\\ &= \Psi_{|Z|}(t)E_{X_1}\left[e^{itX_1}I(X_1 \geq 0)\right] + \Psi_{|Z|}(-t)E_{X_1}\left[e^{itX_1}I(X_1 < 0)\right]\\ &= a(t)^2 - b(t)^2,\end{aligned}$$

which cannot be the form of cf in Cauchy. □

We showed that the random variable $X_1 + X_2$ is not Cauchy distributed using the cf of it, but we still do not know the exact pdf of it. We employ the cdf approach to resolve this; consequently, some informative results appear and become useful presentation materials for graduate level statistics theory courses.

Define $\upsilon(x)$ as the pdf of the univariate standard Cauchy distribution given by (1.2) with $\mu = 0$ and $\sigma = 1$ and define $\Upsilon(x)$ as its cdf; that is, $\Upsilon(x) = P(X \leq x) = (1/\pi)\arctan(x)+1/2,\ x \in \mathbb{R}$. Furthermore $g(x, y)$ denotes the pdf of the bivariate standard Cauchy distribution as given by

$$g(x, y) = \upsilon(x)\upsilon(y), \quad x, y \in \mathbb{R}.$$

This is the pdf of (X_1, Z) as defined in the problem setup. Since, by construction, (2.1), X_1 and X_2 have the same sign, the pdf of (X_1, X_2) is twice as large as that of (X_1, Z) when $xy > 0$ and is equal to 0 when $xy < 0$. Hence,

$$f_{X_1,X_2}(x, y) = \begin{cases} 2\upsilon(x)\upsilon(y), & xy > 0,\\ 0, & \text{otherwise}.\end{cases} \tag{3.1}$$

Based on (3.1), the cdf of $X_1 + X_2$ can be obtained as

$$F_{X_1+X_2}(z) = P(X_1 + X_2 \leq z) = \int_{x+y\leq z} 2\upsilon(x)\upsilon(y)\,dxdy, \quad z \in \mathbb{R}. \tag{3.2}$$

Note that the joint density of (3.1) is symmetric about the origin and the distribution of $X_1 + X_2$ is symmetric with respect to 0; that is, $F_{X_1+X_2}(z) = 1 - F_{X_1+X_2}(-z)$. Therefore, it is sufficient to evaluate it for a positive z. For $z > 0$, (3.2) becomes

$$F_{X_1+X_2}(z) = \frac{1}{2} + \int_{x\geq 0,\, y\geq 0,\, x+y\leq z} 2\upsilon(x)\upsilon(y)\,dxdy. \tag{3.3}$$

To find the integral in (3.3), we develop the following Lemma 1.

Lemma 1. *For $z > 0$,*

$$\int_{x\geq 0,\, y\geq 0,\, x+y\leq z} 2\upsilon(x)\upsilon(y)\,dxdy = \frac{1}{2} - \Upsilon(z) + 2\int_0^z \upsilon(x)\Upsilon(z-x)\,dx.$$

Proof:

$$\int_{x\geq 0,\, y\geq 0,\, x+y\leq z} 2\upsilon(x)\upsilon(y)\,dxdy = \frac{2}{\pi^2}\int_0^z \frac{1}{(1+x^2)}\arctan(z-x)\,dx$$
$$= 2\int_0^z \upsilon(x)\left(\Upsilon(z-x) - \frac{1}{2}\right)dx = \frac{1}{2} - \Upsilon(z) + 2\int_0^z \upsilon(x)\Upsilon(z-x)\,dx.$$

Note the definitions of pdf and cdf of the univariate standard Cauchy distribution. □

By the direct application of Lemma 1, we have

$$F_{X_1+X_2}(z) = 1 - \Upsilon(z) + 2\int_0^z \upsilon(x)\Upsilon(z-x)\,dx.$$

For illustration, this cdf is plotted in Figure 1 with the standard Cauchy cdf, the cdf of $X_1 + Z$, which is the sum of two independent standard Cauchy random variables, and the one of Cauchy(1, 1) as references. Note the peculiar shape around $z = 0$.

Differentiating this cdf with respect to z, we have the pdf of $X_1 + X_2$; that is,

$$f_{X_1+X_2}(z) = -\upsilon(z) + 2\frac{d}{dz}\int_0^z \upsilon(x)\Upsilon(z-x)\,dx.$$

For differentiating the integral $\int_0^z \upsilon(x)\Upsilon(z-x)dx$, we use Leibnitz's rule (Casella and Berger, 2002), which results in

$$\frac{d}{dz}\int_0^z \upsilon(x)\Upsilon(z-x)dx = \frac{1}{2}\upsilon(z) + \int_0^z \upsilon(x)\upsilon(x-z)\,dx. \tag{3.4}$$

To complete the calculation, we need evaluate the integral in (3.4). Blyth (1986) used a similar approach in Lemma 2 to find the density of the sum of two independent Cauchy random variables.

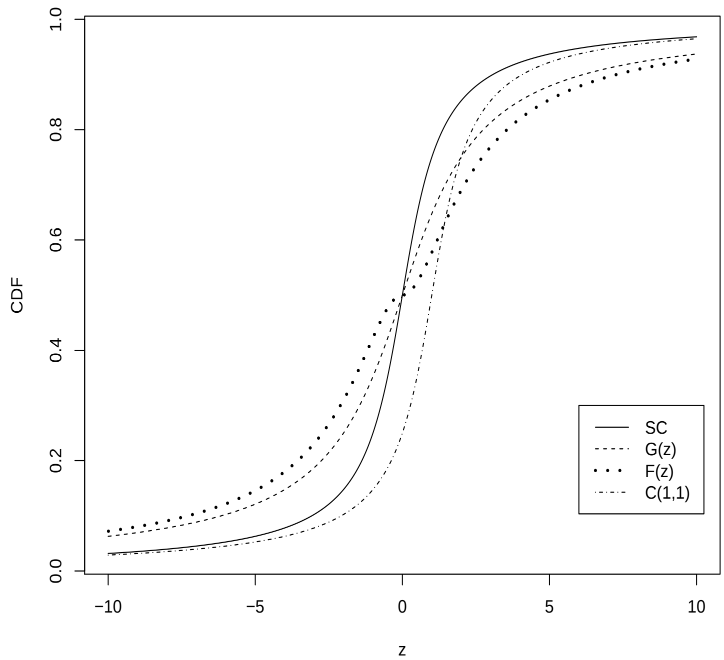

Figure 1: *The cdfs of the standard Cauchy distribution (denoted as SC), X_1+Z (denoted as $G(z)$), X_1+X_2 (denoted as $F(z)$), and Cauchy$(1,1)$ (denoted as $C(1,1)$).*

Lemma 2. *For $z > 0$,*

$$\int_0^z \upsilon(x)\upsilon(x-z)dx = \frac{2}{\pi^2 z(4+z^2)}\left[z\arctan(z) + \ln\left(1+z^2\right)\right].$$

Proof: We write down the product of Cauchy densities as

$$\upsilon(x)\upsilon(z-x) = \frac{1}{\pi^2}\left(\frac{a+bx}{1+x^2} + \frac{c+d(x-z)}{1+(x-z)^2}\right). \tag{3.5}$$

The integration of the right-hand side of (3.5) is straightforward and we only need to find the constants a, b, c, and d. Brought to a common denominator, the right-hand side has for its numerator a cubic in x whose constant term must be 1 and whose other coefficients must all be 0. Solving these equations, we get

$$a = \frac{1}{4+z^2} = c \quad \text{and} \quad b = \frac{2}{z(4+z^2)} = -d.$$

Plugging in these values in (3.5), we have the result. □

We evaluated the density of X_1+X_2 for positive values. Due to symmetry of the density of X_1+X_2, we have $f_{X_1+X_2}(z) = f_{X_1+X_2}(-z)$. Finally, using Lemma 2 and after some calculations, the pdf of X_1+X_2

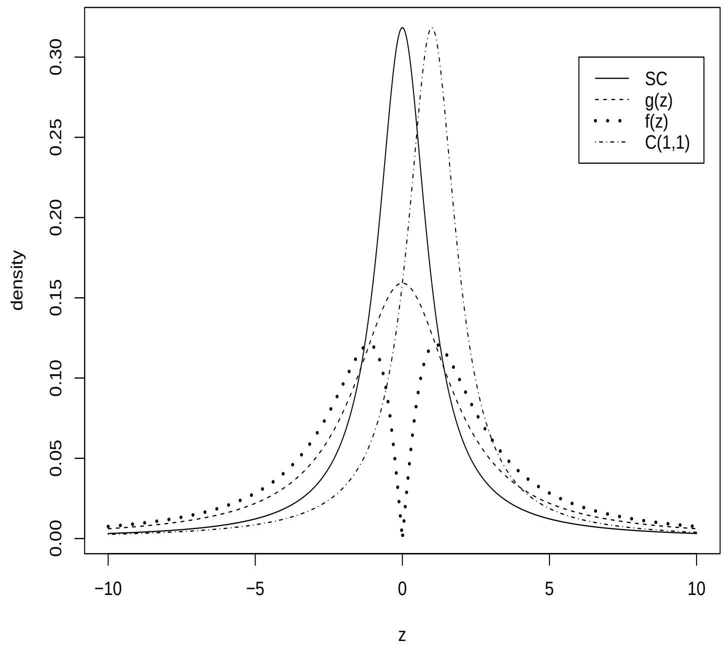

Figure 2: *Densities of the standard Cauchy distribution (denoted as SC),* X_1+Z *(denoted as* $g(z)$*),* X_1+X_2 *(denoted as* $f(z)$*), and Cauchy*$(1,1)$ *(denoted as* $C(1,1)$*).*

is given by

$$f_{X_1+X_2}(z) = \begin{cases} \dfrac{4}{\pi^2|z|\,(4+z^2)}\left[z\arctan(z) + \ln\left(1+z^2\right)\right], & |z|>0, \\ 0, & \text{otherwise.} \end{cases} \tag{3.6}$$

For illustration, this density is plotted in Figure 2 with the standard Cauchy density, the density of $X_1 + Z$, which is the sum of two independent standard Cauchy random variables, and the density of Cauchy$(1, 1)$ as references. Obviously, from Figure 2, its bimodal form is not Cauchy.

Some interesting results are mentioned as remarks.

Remark 1. The conditional distributions of X_i, $i = 1, 2$ given $X_1 + X_2 = z$ are as follows:

$$f_{X_i|X_1+X_2=z}(x_i) = \begin{cases} \dfrac{\pi^2|z|\left(4+z^2\right)\upsilon(x_i)\upsilon(z-x_i)}{2\left[z\arctan(z)+\ln\left(1+z^2\right)\right]}, & x_i(z-x_i)>0,\ |z|>0, \\ 0, & \text{otherwise.} \end{cases}$$

by (3.1) and (3.6).

Remark 2. Suppose that X_1 and Z are independent univariate standard Cauchy random variables. If we make a new random variable X_2 instead of (2.1) such that

$$X_2 = \begin{cases} Z, & X_1 \le \lambda z, \\ -Z, & X_1 > \lambda z, \end{cases} \quad \lambda \in \mathbb{R}.$$

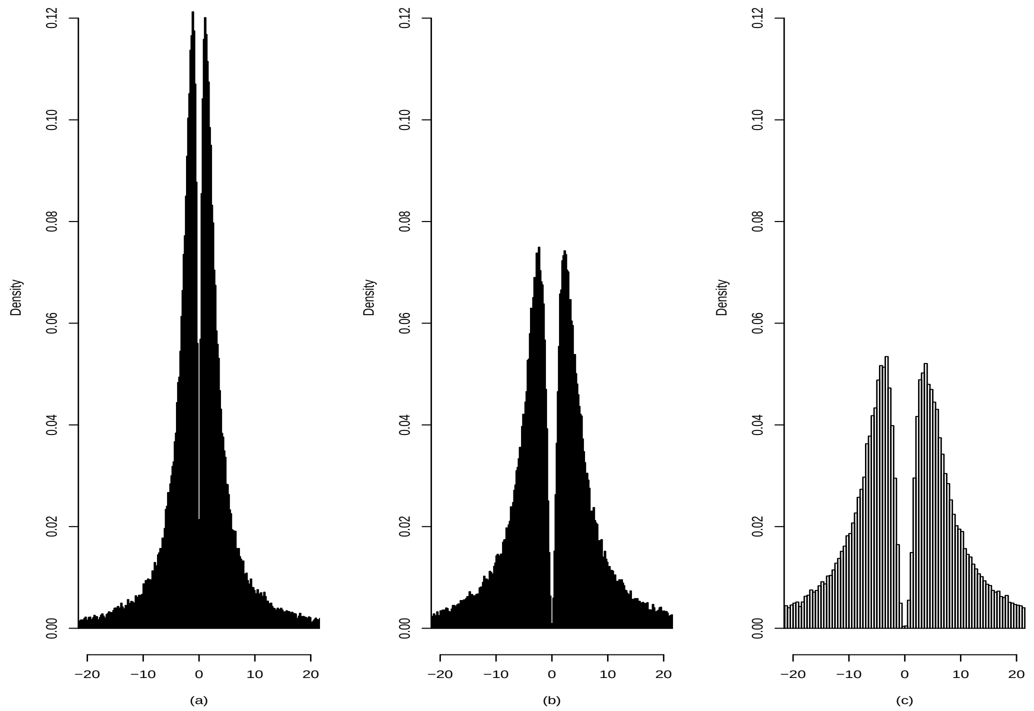

Figure 3: *(a) Histogram of* $X_1 + X_2$*; (b) Histogram of* $X_1 + X_2 + X_3$*; (c) Histogram of* $X_1 + X_2 + X_3 + X_4$ *.*

Then, the distribution of X_2 follows a skew-Cauchy distribution having a pdf $2\upsilon(x_2)\Upsilon(\lambda x_2)$, $x_2 \in \mathbb{R}$ (Azzalini and Capitanio, 1999).

Remark 3. One possible extension of constructing more than a new random variable can be done similar to (2.1). Suppose that X_1, Z_1, and Z_2 are independent standard Cauchy random variables. The random variable X_2 is the same as (2.1) replacing Z with Z_1. X_3 is constructed as

$$X_3 = \begin{cases} |Z_2|, & X_1 \geq 0, \\ -|Z_2|, & X_1 < 0. \end{cases} \tag{3.7}$$

Then what is the distribution of $X_1 + X_2 + X_3$? In a similar manner as (3.7), if we have another standard Cauchy random variable Z_3 independent of X_1, Z_1, and Z_2, then a new random variable X_4 can be constructed as follows:

$$X_4 = \begin{cases} |Z_3|, & X_1 \geq 0, \\ -|Z_3|, & X_1 < 0. \end{cases} \tag{3.8}$$

We found the density of $X_1 + X_2$ which is given in (3.6). For a partial answer of the distributions of $X_1 + X_2 + X_3$, and $X_1 + X_2 + X_3 + X_4$, we simulated $N = 100{,}000$ samples from each independent standard Cauchy distribution. By the constructions (2.1), (3.7), and (3.8), we obtained X_2, X_3, and X_4 consecutively. The following Figure 3 shows the histograms of each random variable, that is, those of $X_1 + X_2$, $X_1 + X_2 + X_3$, and $X_1 + X_2 + X_3 + X_4$. Based on the Figure, we conjecture that if we increase the number of components in the sum, then the density becomes heavier at the tails and less mass is concentrated around 0.

4. Conclusions

In this paper, we define a multivariate Cauchy distribution using a probability density function commonly adopted in modern mathematical statistics. Then Ferguson's definition of a multivariate Cauchy distribution can be viewed as a characterization theorem using the characteristic function approach as in the multivariate normal distribution. We also illustrated that linear combinations of Cauchy random variables need not themselves be Cauchy. The correct statement is that any linear combination of random variables from a multivariate Cauchy distribution is Cauchy distributed; therefore, multivariate Cauchy implies univariate Cauchy but not vice versa. The derivation is quite simple and is suitable for presentation in statistics graduate theory courses. For future work, the exact densities of $X_1 + X_2 + X_3$ and $X_1 + X_2 + X_3 + X_4$ need to be developed. Furthermore, it is desirable to develop the approximate distribution of $\sum_{i=1}^{n} X_i$ when $n \to \infty$.

References

Azzalini, A. and Capitanio, A. (1999). Statistical applications of the multivariate skew-normal distribution, *Journal of the Royal Statistical Society, Series B*, **61**, 579–602.

Blyth, C. R. (1986). Convolutions of Cauchy distributions, *The American Mathematical Monthly*, **93**, 645–647.

Casella, G. and Berger, R. L. (2002). *Statistical Inference*, 2nd, Duxbury, CA.

Crovella, M. E., Taqqu, M. S. and Bestavros, A. (1998). *A Practical Guide to Heavy Tails: Statistical Techniques and Applications*, Adler, R., Feldman, R. and Taqqu, M., editors, Birkhauser, Boston.

Ferguson, T. S. (1962). A representation of the symmetric bivariate Cauchy distribution, *The Annals of Mathematical Statistics*, **33**, 1256–1266.

Kotz, S. and Nadarajah, S. (2004). *Multivariate t Distributions and Their Application*, Cambridge University Press, Cambridge.

Melnick, E. L. and Tenenbein, A. (1982). Misspecifications of the normal distribution, *The American Statistician*, **36**, 372–373.

Novosyolov, A. (2006). *The Sum of Dependent Normal Variables may be not Normal*, Technical report, Institute of computational modeling SB RAS.

Park, S. Y. and Bera, A. K. (2009). Maximum entropy autoregressive conditional heteroskedasticity model, *Journal of Econometrics*, **150**, 219–230.

Romano, J. P. and Siegel, A. F. (1986). *Counterexamples in Probability and Statistics*, Wadsworth & Brooks/Cole Advanced Books & Software, CA.

Zhang, J. (2010). A highly efficient *L*-estimator for the location parameter of the Cauchy distribution, *Computational Statistics*, **25**, 97–105.

Received January 16, 2014; Revised February 17, 2014; Accepted February 17, 2014