Characterization of distributions based on certain powers of random variables

Wen-Jang Huang[1], **Yen-Ning Huang**[2] **and Chia-Ling Lai**[1]

[1]Department of Applied Mathematics, National University of Kaohsiung

[2]Department of Mathematics, National Tsing-Hua University

Abstract

It is well known that if W is $\mathcal{N}(0,1)$ distributed, then W^2 has the χ_1^2 distribution. Roberts and Geisser(1966) generalized this result and gave a necessary and sufficient condition for the square of a random variable to be gamma distributed. In this note, first the class of random variables is characterized when the distribution of its nth power is given, where n is a positive integer. Next, some characterization results based on certain quadratic statistics are also provided.

Keywords and phrases: characterization, exponential family.
AMS 2000 subject classifications: Primary 62E10; secondary 60E05.

1. Introduction

It is well known that if W is $\mathcal{N}(0,1)$ distributed, then W^2 has the χ_1^2 distribution. Roberts and Geisser(1966) generalized this result and gave a necessary and sufficient condition for the square of a random variable to be gamma distributed. In Roberts (1971), it was further explored that a necessary and sufficient condition was given for the nth, $n \geq 0$, power of a random variable to be gamma distributed. On the other hand, from elementary algebra, it is well-known that the equation $x^n = b$, where n is a positive integer, $b > 0$, has a unique real number solution $x = b^{1/n}$, if n is odd; and has two real number solutions $x = b^{1/n}$ and $x = -b^{1/n}$, if n is even. From this, it is natural to ask that when the distribution of the nth power of a random variable X is given, what can we say about the distribution of X? The answer of this question will be provided in Section 2.

Next, some characterization results based on certain quadratic statistics are presented, which are generalizations of Roberts and Geisser(1966), Roberts(1971), and Gupta et al.(2004).

2. Characterization of distributions of random variables whose nth power is given

Let $f_X(x)$, $x \in A$, be the probability density function (p.d.f.)of a continuous random variable X, and the corresponding distribution function of X be

$$F_X(x) = \int_{-\infty}^{x} f_X(t)dt, \quad \forall x \in R.$$

Also let the probability mass function(p.m.f.) $f_X(x)$ of a discrete random variable X be given by

$$f_X(x) = P(X = x), \quad x \in A,$$

where A is the support of the distribution of X.

We investigate the problem discussed above and present the main result of this section in the following .

Theorem 1 Let n be a positive integer, $g(y)$, $y \in A$, a continuous p.d.f. Also assume $A \subset [0, \infty)$, when n is even. Then X^n has g as its p.d.f., if and only if the p.d.f. of X is

$$f_X(x) = \begin{cases} nx^{n-1}g(x^n) & , n \text{ is odd,} \\ s(x)|x|^{n-1}g(x^n) & , n \text{ is even,} \end{cases} \tag{1}$$

where $x \in B = \{x | x \in R, x^n \in A\}$, $s(x) \geq 0$, and $s(x) + s(-x) = n$, $\forall x \in B$.

Proof. Let $Y = X^n$. When n is odd, the transformation between X and Y is one to one, the result is obvious. We now prove the case that n is even.

First we show the sufficiency. Observe that $f_Y(y)$, the p.d.f. of Y, is given by

$$f_Y(y) = (ny^{1-1/n})^{-1}(f_X(y^{1/n}) + f_X(-y^{1/n})), \quad y \in A.$$

It follows that

$$\begin{aligned} f_Y(y) &= (ny^{1-1/n})^{-1}(s(y^{1/n}) + s(-y^{1/n}))y^{(n-1)/n}g(y) \\ &= g(y), \quad y \in A, \end{aligned}$$

using the fact that $s(x) + s(-x) = n$, $\forall x \in B$.

Next we show the necessity. Since the p.d.f. of the random variable Y, $g(y)$, $y \in A$, is continuous, it follows that the distribution $F_X(x) = P(X \leq x)$ is absolutely continuous with respect to Lebesgue measure, and write $f_X(x)$ as the p.d.f. of X. Define

$$s(x) = f_X(x)|x|^{1-n}(g(x^n))^{-1}, \quad x \in B. \tag{2}$$

We will show that this A satisfies $s(x) + s(-x) = n$, $\forall x \in B$. Given the p.d.f. of X

$$f_X(x) = s(x)|x|^{n-1}g(x^n), \quad x \in B, \tag{3}$$

the p.d.f. of Y is

$$\begin{aligned} f_Y(y) &= (ny^{1-1/n})^{-1}(s(y^{1/n}) + s(-y^{1/n}))y^{(n-1)/n}g(y) \\ &= n^{-1}(s(y^{1/n}) + s(-y^{1/n}))g(y), \quad y \in A. \end{aligned}$$

By the assumption, the p.d.f. of Y is $g(y)$, hence almost everywhere in A, $s(y^{1/n}) + s(-y^{1/n}) = n$, or equivalently, $s(x) + s(-x) = n$, almost everywhere in B, which completes the proof.

¿From the above theorem, for every even integer n, unlike the case in real equation that there are only two real solutions for the equation $x^n = b$, where $b > 0$, there are infinitely many nonnegative functions $s(x)$ satisfy $s(x) + s(-x) = n$. Consequently, there are infinitely many distributions for X, satisfying $X^n \stackrel{d}{=} Y$, when the distribution of Y is given.

We have the following immediate consequence.

Corollary 1 Let n be a positive integer. The distribution of X^n belongs to the exponential family, with p.d.f.

$$g(y) = c(\theta)h(y)e^{\sum_{j=1}^{k} w_j(\theta)t_j(y)}, \quad y \in A, \tag{4}$$

if and only if the p.d.f. of X is

$$f_X(x) = \begin{cases} nx^{n-1}c(\theta)h(x^n)e^{\sum_{j=1}^{k} w_j(\theta)t_j(x^n)} & , n \text{ is odd}, \\ c(\theta)s(x)|x|^{n-1}h(x^n)e^{\sum_{j=1}^{k} w_j(\theta)t_j(x^n)} & , n \text{ is even}, \end{cases} \tag{5}$$

where $x \in B = \{x | x^n \in A\}$, $s(x) \geq 0$, and $s(x) + s(-x) = n$, $\forall x \in B$.

As gamma distribution belongs to the exponential family, Corollary 1 is a generalization of Roberts and Geisser(1966).

Example 1 Note that by Corollary 1, X^2 has a χ^2_1 distribution with the p.d.f.

$$g(y) = \frac{1}{\sqrt{2\pi}} y^{-1/2} e^{-y/2}, \quad y > 0,$$

if and only if the p.d.f. of X is

$$\begin{aligned} f_X(x) &= \frac{1}{\sqrt{2\pi}} |x||x|^{-1} e^{-x^2/2} s(x) \\ &= \frac{1}{\sqrt{2\pi}} e^{-x^2/2} s(x), \quad x \in R, \end{aligned}$$

where

$$s(x) + s(-x) = 2, \quad \forall x > 0. \tag{6}$$

Obviously $s(x) = 2F(x)$ satisfies (6), where F is a symmetric distribution, that is

$$F(x) + F(-x) = 1, \quad \forall x \in R. \tag{7}$$

This result is exactly the Corollary 1 of Roberts and Geisser(1966). Hence the random variable Z with a $S\mathcal{N}(\lambda)$ distribution or $SG\mathcal{N}(\lambda_1, \lambda_2)$ distribution, satisfies $Z^2 \sim \chi^2_1$. Here Z has a skew-normal distribution with parameter $\lambda > 0$, denoted by $Z \sim S\mathcal{N}(\lambda)$, if its p.d.f. is given by

$$f(z|\lambda) = \phi(z) 2\Phi(\lambda z), \quad z \in R,$$

where ϕ and Φ are the p.d.f. of $\mathcal{N}(0,1)$ and distribution function of $\mathcal{N}(0,1)$, respectively (see Azzalini (1985)). Moreover U has a skew-generalized normal distribution with parameters $\lambda_1 \in R$, $\lambda_2 \geq 0$, denoted by $U \sim SG\mathcal{N}(\lambda_1, \lambda_2)$, if its p.d.f. is given by

$$f(u|\lambda_1, \lambda_2) = 2\phi(u)\Phi(\frac{\lambda_1 u}{\sqrt{1 + \lambda_2 u^2}}), \quad u \in R$$

(see Arellano-Valle et al. (2003)). Note that $S\mathcal{N}(0) = \mathcal{N}(0,1)$, $SG\mathcal{N}(0, \lambda_2) = \mathcal{N}(0,1)$. So both skew-normal and skew-generalized normal classes contain the $\mathcal{N}(0,1)$ distribution.

Now let X_1 and X_2 be two random variables satisfying $X_1^n \stackrel{d}{=} X_2^n$, where n is an even integer. Also let ξ be an even function. Then $\xi(x_1) \stackrel{d}{=} \xi(x_2)$. In particular let $X_1 \sim \mathcal{N}(0,1)$, $X_2 \sim SN(\lambda)$, $X_3 \sim SG\mathcal{N}(\lambda_1, \lambda_2)$. Then $\xi(X_1) \stackrel{d}{=} \xi(X_2) \stackrel{d}{=} \xi(X_3)$, for every even function ξ.

We also state a generalization of Roberts(1971), the proof is exactly the same as that of Theorem 1 hence is omitted.

Corollary 2 Let $n \in R \setminus \{0\}$, $g(y)$, $y \in A$, be a continuous p.d.f. Then $|X|^n$ has g as its p.d.f., if and only if the p.d.f. of X is

$$f_X(x) = s(x)|x|^{n-1} g(|x|^n), \tag{8}$$

where $x \in B = \{x | x \in R, x^n \in A\}$, and $s(x) + s(-x) = n$, $\forall x \in B$.

To this end, we state without proof a theorem for the discrete case.

Theorem 2 Let n be a positive integer, $g(y)$, $y \in A$, a p.m.f., where A is assumed to be a finite or countable set. Also assume $A \subset [0, \infty)$, when n is even. Then X^n has g as its p.m.f., if and only if the p.m.f. of X is

$$f_X(x) = \begin{cases} g(x^n) & , n \text{ is odd}, \\ s(x) g(x^n) & , n \text{ is even}, \end{cases}$$

where $x \in B = \{x | x \in R, x^n \in A\}$, $s(x) \geq 0$, and $s(x) + s(-x) = n$, $\forall x \in B$.

3. Characterizations based on quadratic statistics

Apparently other than $\mathcal{N}(0,1)$ distribution, there are infinitely many distributions for X, such that X^2 is χ_1^2 distributed. With some other types of conditions, characterizations of the $\mathcal{N}(0,1)$ distribution can also be obtained. Roberts and Geisser(1966) showed that if X_1 and X_2 are independent and identically distributed(i.i.d.) random variables, then X_1^2, X_2^2, and $\frac{1}{2}(X_1 + X_2)^2$ are all χ_1^2 distributed, if and only if X_1 and X_2 are $\mathcal{N}(0,1)$ distributed. A generalization of Roberts and Geisser(1966) was given in Theorem 3 of Gupta et al.(2004), where $\frac{1}{2}(X_1+X_2)^2$ was replaced by $(AX_1+BX_2)^2$, $A, B \neq 0$ and $A^2+B^2 = 1$. Roberts(1971) gave a sufficient condition for a quadratic form to be χ^2_k distributed, and stated a qualified necessary condition.

Based on some quadratic statistics, and under the assumption that the distribution is uniquely determined by its sequence of moments, Gupta et al.(2004) gave two characterization results for the $S\mathcal{N}(\lambda)$ distribution. Their results are stated in the following with a more general form.

Theorem 3 (Gupta et al.(2004) Theorem 1). Let X and Y be i.i.d. F_0, a given distribution that is uniquely determined by its sequence of moments where all of them do exist. Denote by G_0 the distribution of X^2 (and Y^2), and by H_0 the distribution of $\frac{1}{2}(X+Y)^2$. Let X_1 and X_2 be i.i.d. F, an unspecified distribution with sequence of moments which all exist. Then $X_1^2 \sim G_0$, $X_2^2 \sim G_0$, and $\frac{1}{2}(X_1 + X_2)^2 \sim H_0$, if and only if $F(x) = F_0(x)$ or $F(x) = 1 - F_0(-x)$.

Theorem 4 (Gupta et al.(2004) Theorem 2). Let F_0 be a given distribution uniquely determined by its sequence of moments where all of them exist. Let $Y \sim F_0$. Let G_0 be the distribution of Y^2, and H_0 be the distribution of $(Y+a)^2$, where $a \neq 0$ is a constant. Let $X \sim F$, an unspecified distribution which admits moments of all order. Then $X^2 \sim G_0$, and $(X+a)^2 \sim H_0$, if and only if $F = F_0$.

In this section some characterization of distributions based on certain quadratic statistics are investigated. To begin with, a lemma by Gupta et al. (2004) is presented.

Lemma 1 (Gupta et al.(2004) Lemma 6). The skew-normal distribution is uniquely determined by its sequence of moments.

Now we have the following generalized result.

Theorem 5 Let X and Y be i.i.d. F_0, a given distribution that is uniquely determined by its sequence of moments where all of them exist. Denote by G_0 the distribution of X^2(and Y^2), and by H_0 the distribution of $(X+aY^{2k})^2$ for some fixed nonnegative integer k and $a \neq 0$. Let X_1 and X_2 be i.i.d. with distribution function F, an unspecified distribution with sequence of moments where all of them exist.

Then $X_1^2 \sim G_0$, $X_2^2 \sim G_0$ and $(X_1 + aX_2^{2k})^2 \sim H_0$, if and only if $F(x) = F_0(x)$.

Proof: The sufficiency follows directly from the definition of F_0, G_0 and H_0. To prove the necessity, first note that since all moments of F_0 exist and X and Y are independent, it follows that all moments of G_0 and H_0 do exist. Define the following for $i = 1, 2, \cdots$.

μ_i is the ith moment of F,

μ_i^0 is the ith moment of F_0.

Since both $(X_1 + aX_2^{2k})^2$ and $(X + aY^{2k})^2 \sim H_0$, we have

$$E[(X_1 + aX_2^{2k})^2]^l = E[(X + aY^{2k})^2]^l, \quad \forall l = 1, 2, \cdots. \tag{9}$$

The even moments of F coincide with the even moments of F_0, i.e.,

$$\mu_{2i} = \mu_{2i}^0 \quad \forall i = 1, 2, \cdots, \tag{10}$$

since $X_1^2 \sim G_0$ and $X^2 \sim G_0$. So, we will only proceed by induction to show that all the odd moments of F coincide with the odd moments of F_0, i.e.

$$\mu_{2i-1} = \mu_{2i-1}^0 \quad \forall i = 1, 2, \cdots. \tag{11}$$

First, taking $l = 1$ in (9) and using (10), we get $\mu_1 = \mu_1^0$. Hence, the induction statement (11) is true when $i = 1$. Now suppose $\mu_{2i-1} = \mu_{2i-1}^0$, $i = 1, 2, \cdots, n$. Again from (9) we have

$$\sum_{m=0}^{2l} \binom{2l}{m} a^{2l-m} \mu_m \mu_{2k(2l-m)} = \sum_{m=0}^{2l} \binom{2l}{m} a^{2l-m} \mu_m^0 \mu_{2k(2l-m)}^0 \tag{12}$$

holds $\forall l = 1, 2, \cdots$. Take $l = n + 1$, then (12) becomes

$$\begin{aligned} &\sum_{m=0}^{2n} \binom{2n+2}{m} a^{2(n+1)-m} \mu_m \mu_{2k(2n+2-m)} + a\mu_{2n+1}\mu_{2k} + \mu_{2n+2} \\ &= \sum_{m=0}^{2n} \binom{2n+2}{m} a^{2(n+1)-m} \mu_m^0 \mu_{2k(2l-m)}^0 + a\mu_{2n+1}^0 \mu_{2k} + \mu_{2n+2}^0. \end{aligned} \tag{13}$$

By (10), all terms with even moments in the left–hand side of (13) cancel with the corresponding terms in the right–hand side of (13). Also by the induction hypothesis, $\mu_{2i-1} = \mu_{2i-1}^0$ $\forall i = 1, 2, \cdots, n$. It follows that (13) would give $\mu_{2n+1} = \mu_{2n+1}^0$. Hence, the induction argument is complete which proves that $\mu_{2i-1} = \mu^0_{2i-1}$, $\forall i = 1, 2, \cdots$. Consequently, we have obtained $\mu_i = \mu^0_i$, $\forall i = 1, 2, \cdots$, and it follows that $F = F_0$.

In Theorem 5, we give a characterization result based on the distribution X_1^2, X_2^2 and $(X_1 + aX_2^{2k})^2$ for some fixed nonnegative integer k and $a \neq 0$. Obviously, unlike the situation in Theorem 1, we cannot replace the power 2 by an arbitrary integer power $n \geq 2$ in Theorem 5. On the other hand, it is interesting to know whether the result of Theorem 5 still holds if the quadratic statistic $(X_1 + aX_2^{2k})^2$ is replaced by the quadratic statistic $(X_1 + aX_2^{2k+1})^2$ for some fixed positive integer k and $a \neq 0$. We are unable to prove the "only if "part in this case. In addition , we also consider to replace $(X_1 + aX_2^{2k})^2$ by the form such as $(X_1^p + aX_2^{2k})^2$ where $p \geq 3$ is odd. Yet even for the simplest case $(X_1^3 + X_2^2)^2$, we can only obtain $\mu_{2i} = \mu^0_{2i}$ and $\mu_{6i-3} = \mu^0_{6i-3}$, $\forall i = 1, 2, \cdots$. In other words we cannot arrive at all of the moments of X and X_1 are equal.

Again, it is easy to see Theorem 4 is an immediate consequence of Theorem 5. We give two more simple corollaries in the following.

Corollary 3 Let X_1 and X_2 be i.i.d. F, an unspecified distribution which admits moments of all order. Then $X_1^2 \sim \chi_1^2$, $X_2^2 \sim \chi_1^2$ and $(X_1 + aX_2^{2k})^2 \sim H_0(\lambda)$, for some fixed nonnegative integer k and $a \neq 0$, if and only if $F = S\mathcal{N}(\lambda)$, where $H_0(\lambda)$ is the distribution of $(X + aY^{2k})^2$when X and Y are i.i.d. $S\mathcal{N}(\lambda)$ random variables.
Proof: Take $F_0 = S\mathcal{N}(\lambda)$, so that $G_0 = \chi_1^2$ and $H_0 = H_0(\lambda)$. Apply Theorem 5 and note that $S\mathcal{N}(\lambda)$ distribution is uniquely determined by its moments by Lemma 1.

Corollary 4 Let X_1 and X_2 be i.i.d. F, an unspecified distribution which admits moments of all order. Then $X_1^2 \sim \chi_1^2$, $X_2^2 \sim \chi_1^2$ and $(X_1 + aX_2^{2k})^2 \sim H_0$, for some fixed nonnegative integer k and $a \neq 0$, if and only if $F = \mathcal{N}(0, 1)$, where H_0 is the distribution of $(X + aY^{2k})^2$ when X and Y are i.i.d. $\mathcal{N}(0, 1)$ random variables.
Proof: The result is obtained from Corollary 3 by taking $\lambda = 0$. Alternatively, take $F_0 = \mathcal{N}(0, 1)$, so that $G_0 = \chi_1^2$ and apply Theorem 5.

4. Conclusion

The characterizations we give in Section 3, need the assumption that the distributions are determined by their moments, respectively. Although this assumption is not needed in Roberts and Geisser(1966). For X_1 and X_2 being i.i.d., along the lines of Roberts and Geisser(1966), it is easy to obtain X_1^2, X_2^2 and $(AX_1 + BX_2)^2$, where $A = 1/\sqrt{2}$, $B = -1/\sqrt{2}$, are all χ_1^2 distributed, if and only if X_1 and X_2 are $\mathcal{N}(0, 1)$ distributed. Yet without the moments assumption, we cannot find other similar characterizations for different pairs of A and B.

Acknowledgements. We are grateful to Professor Gupta for sending us the preprints of Gupta et al.(2004).

References

1. Arellano-Valle, R.B., Gómez, H.W. and Quintana, F.A.(2003). A new class of skew-normal distributions. Commun. *Statist.-Theory and Meth.*, 33, 1465-1480.
2. Azzalini, A.(1985). A class of distributions which includes the normal ones. *Scand.J. Statist.*, 12, 171-178.
3. Gupta, A.K., Nguyen, T.T. and Sanqui, J.A.T.(2004). Characterization of the skew-normal distribution. *Ann. Inst. Statist. Math.*, 56, 351-360
4. Robers, C. and Geisser, S.(1966). A necessary and sufficient condition for the square of a random variable to be gamma. *Biometrika*, 53, 275-279.
5. Robers, C.(1971). On the distribution of random variables whose mth absolute power is gamma. *Sankhyā*, Ser.A, 33, 229-232.