Ann. Inst. Statist. Math.
Vol. 41, No. 1, 13–17 (1989)

CHARACTERIZATION BASED ON CONDITIONAL DISTRIBUTIONS

NORMAN L. JOHNSON[1] AND SAMUEL KOTZ[2]

[1]*University of North Carolina, Chapel Hill, NC 27514, U.S.A.*
[2]*Department of Management and Statistics, University of Maryland, College Park, MD 20742, U.S.A.*

(Received February 18, 1988; revised July 11, 1988)

Abstract. The field of application of a result given by Singh and Vasudeva (1984, *J. Indian Statist. Assoc.*, **22**, 93–96) which provides a way of characterizing the distribution of a random variable X, through conditional distributions of a second variable Z, given X, is extended.

Key words and phrases: Characterization, conditional distribution.

Singh and Vasudeva (1984) have proved the following result: *Let X, Y and Z be random variables such that X and Y are non-negative and*

$$\Pr[Z=k\,|\,X=t]=\Pr[Z=k\,|\,Y=t]=e^{-t}(1-e^{-t})^k \qquad (t\geq 0;\ k=0,1,2,\ldots)\,.$$

Then X and Y are identically distributed.

This result can be used to characterize the distribution of X from the conditional distribution of Z. Singh and Vasudeva use this fact to characterize the exponential distribution with density function

$$\alpha\, e^{-\alpha t} \qquad (t\geq 0,\ \alpha>0)\,,$$

by the fact that if $\Pr[Z=k\,|\,X=t]=e^{-t}(1-e^{-t})^k$ and Z has the Yule distribution

$$\Pr[Z=k]=\alpha B(\alpha+1,k+1) \qquad (k=0,1,\ldots)\,,$$

then X must have an exponential distribution.

Singh and Vasudeva's proof uses the extended Stone-Weierstrass theorem (Simmons (1963), p. 166). We present the following extension of this result, with a proof using somewhat simpler methods.

If X and Y have the same support, and

$$\Pr[Z = k \mid X = t] = \Pr[Z = k \mid Y = t] = g(t)\{h(t)\}^k , \tag{1}$$

($g(t)$, $h(t) > 0$) for all $k = 0, 1, 2, \ldots$ and all t in the common support of X and Y, and $h(t)$ is a strictly monotonic function of t, then X and Y have identical distributions.

PROOF. We note that since (1) holds for all $k = 0, 1, 2, \ldots$ we must have $h(t) < 1$ (otherwise $\Pr[Z = k \mid X = t] > 1$ for sufficiently large k). Since

$$\Pr[Z = k] = E_X[\Pr[Z = k \mid X]] = E_Y[\Pr[Z = k \mid Y]] , \tag{2}$$

$$\int_{-\infty}^{\infty} g(t)\{h(t)\}^k dF_X(t) = \int_{-\infty}^{\infty} g(t)\{h(t)\}^k dF_Y(t) .$$

In particular, putting $k = 0$,

$$\int_{-\infty}^{\infty} g(t) dF_X(t) = \int_{-\infty}^{\infty} g(t) dF_Y(t) ,$$

and (2) can be written

$$\int_{-\infty}^{\infty} \{h(t)\}^k dF_{X'}(t) = \int_{-\infty}^{\infty} \{h(t)\}^k dF_{Y'}(t) , \tag{3}$$

where

$$dF_{X'}(t) = \frac{g(t) dF_X(t)}{\int_{-\infty}^{\infty} g(t) dF_X(t)} , \qquad dF_{Y'}(t) = \frac{g(t) dF_Y(t)}{\int_{-\infty}^{\infty} g(t) dF_Y(t)} \tag{4}$$

correspond to cumulative distribution functions $F_{X'}(x')$ and $F_{Y'}(y')$ of random variables X', Y', respectively.

Equation (3) can also be written

$$E[\{h(X')\}^k] = E[\{h(Y')\}^k] \qquad (k = 0, 1, \ldots) .$$

That is, the random variables $h(X')$, $h(Y')$ have equal moments of all positive integer orders. Since $h(t)$ is bounded ($0 < h(t) < 1$), this means that $h(X')$ and $h(Y')$ have identical distributions. Since $h(t)$ is strictly monotonic, it follows that X' and Y' have identical distributions, that is, $dF_{X'}(t) = dF_{Y'}(t)$. From (4) it follows that X and Y have identical distributions (since we cannot have $F_X(t)/F_Y(t) = \text{constant} \neq 1$ for all t).

Remarks

1. Note that the condition that $h(t)$ is strictly monotonic excludes

the possibility that $g(t)\{h(t)\}^k$ does not depend on t, which would happen if Z were independent of both X and Y. In such a case, although we would have

$$\Pr[Z=k|X=t]=\Pr[Z=k|Y=t] \quad (=\Pr[Z=k])\,,$$

there would clearly be no restrictions on the distributions of X and Y.

2. It is in general necessary that (1) holds for more than a finite set of values of k (e.g., $k=0, 1, 2,\ldots, K$), because equality of a finite set of moments would not necessarily ensure identity of distributions.

3. On the other hand, it is not necessary that Z takes only values 0, 1, 2,.... In fact,

$$\Pr[\bigcup_{k=0}^{\infty}(Z=k)|X=t]=g(t)\{1-h(t)\}^{-1}$$

can be as small as desired. The remainder of the distribution of Z can be quite arbitrary (of course, it is necessary that $g(t)\leq 1-h(t)$).

4. The result still holds, even if (1) is true only for $k=0, r, 2r,\ldots$ where r is a positive integer. The proof is exactly the same, except that $h(t)$ is replaced by $\{h(t)\}^r$.

5. The range of values of t (i.e., the support of X and Y) need not be restricted to $t\geq 0$.

6. The result will still hold if the $(\{h(t)\}^k)$ are replaced by some other set of functions $\{h_k(t)\}$, such that the expected values of $h_k(X)$ determine the distribution of X uniquely.

7. If (1) is valid, and X has a mixture distribution of form

$$F_X(t)=\sum_{j=1}^{m} w_jF_j(t) \quad (0\leq w_j;\ \sum_{j=1}^{m} w_j=1)\,,$$

where the $F_j(\cdot)$'s are proper cumulative distribution functions, then the overall distribution of Z is a mixture of the corresponding distributions in the same proportions. From our result it follows that, conversely, if Z has a mixture distribution over 0, 1, 2,... and (1) is valid, then X has a unique corresponding mixture distribution.

8. If Z takes only the values 0, 1, 2,..., then $g(t)=1-h(t)$. In this case the conditional distribution of $(Z+1)$, given $X=t$, is that of the number of independent trials needed to observe an event which has probability $\{1-h(t)\}$ of occuring at any one trial.

9. In the situation just described, if $h(t)$ is a strictly increasing proper cumulative distribution function over the relevant range of values of t, the conditions of (1) are satisfied and the overall distribution of $(Z+1)$ is that of the number of observed values of independent random variables

W_1, $W_2, \ldots,$ each having cumulative distribution function $h(t)$ needed to obtain one exceeding a random observed value of X.

In particular, if X has the same distribution as each of the W's, then

$$\begin{aligned}
\Pr[Z=k] &= \int_{-\infty}^{\infty} \{1-h(t)\}\{h(t)\}^k dh(t) \\
&= [(k+1)^{-1}\{h(t)\}^{k+1} - (k+2)^{-1}\{h(t)\}^{k+2}]_{t=-\infty}^{t=\infty} \\
&= (k+1)^{-1} - (k+2)^{-1} .
\end{aligned} \tag{5}$$

Using (1), we see that if (5) holds, then X must have the same distribution as each of the W's.

10. Taking $h(t) = t/(1+t)$, and the density function of X as

$$f_X(t) = \frac{1}{B(\alpha,\beta)} \cdot \frac{t^{\beta-1}}{(1+t)^{\alpha+\beta}} \qquad (0 < t;\ \alpha, \beta > 0) , \tag{6}$$

we obtain

$$\begin{aligned}
\Pr[Z=k] &= \frac{1}{B(\alpha,\beta)} \int_0^{\infty} (1+t)^{-1} t^k (1+t)^{-k} t^{\beta-1} (1+t)^{-\alpha-\beta} dt \\
&= \frac{1}{B(\alpha,\beta)} \int_0^{\infty} t^{(\beta+k)-1} (1+t)^{-(\alpha+1)-(\beta+k)} dt \\
&= \frac{B(\alpha+1, \beta+k)}{B(\alpha,\beta)} \qquad (k = 0, 1, \ldots) .
\end{aligned}$$

If $\beta = 1$ we obtain

$$\Pr[Z=k] = B(\alpha+1, k+1)/B(\alpha, 1) = \alpha B(\alpha+1, k+1) ,$$

as Singh and Vasudeva (1984) obtained with $h(t) = 1 - e^{-t}$ and $f_X(t) = \alpha e^{-\alpha t}$ ($t > 0$; $\alpha > 0$) (this shows, incidentally, that the distribution of Z, and conditional geometric distributions given $X = t$, do not determine $h(t)$ and the distribution of X).

If we take $h(t) = [t/(1+t)]^{\gamma}$ ($\gamma > 0$; $t > 0$) with X still having the same density function, we obtain

$$\Pr[Z=k] = \{B(\alpha, k\gamma+\beta) - B(\alpha, (k+1)\gamma+\beta)\}/B(\alpha,\beta) \qquad (k = 0, 1, \ldots) .$$

This may be regarded as a "generalized" Yule distribution.

11. The result also applies if X has a discrete distribution. For example if

$$\Pr[Z = k \mid X = t] = e^{-t}(1 - e^{-t})^k \qquad (t \geq 0;\ k = 0, 1, 2, \ldots)$$

and

$$\Pr[Z = k] = e^{-\theta} \sum_{j=0}^{k} (-1)^j \binom{k}{j} \exp(\theta e^{-j-1}),$$

then X must have a Poisson distribution with expected value θ.

REFERENCES

Simmons, G. F. (1963). *Introduction to Topology and Modern Analysis*, McGraw-Hill, New York.

Singh, H. and Vasudeva, H. L. (1984). A characterization of exponential distribution by Yule distribution, *J. Indian Statist. Assoc.*, **22**, 93–96.