CHARACTERIZATIONS OF SOME DISTRIBUTIONS BY CONDITIONAL MOMENTS

E. M. Bolger[1] and W. L. Harkness

Pennsylvania State University

1. Introduction and summary. Let X_1 and X_2 be independent random variables (r.v.'s) and assume that $Y = X_1 + X_2$ has finite second moment. We assume that the mean and variance of X_1, conditional on fixed values y of Y, satisfy the structural relations

$$\text{(i)}\ E(X_1 \mid Y = y) = \lambda_1 y/\lambda \qquad \text{and} \qquad \text{(ii)}\ V(X_1 \mid Y = y) = (\lambda_1\lambda_2/\lambda^2)u(y)$$

where λ_1 and λ_2 are positive constants, $\lambda = \lambda_1 + \lambda_2$, and $u(y)$ is non-negative. Laha [2] has given a simple necessary and sufficient condition for the regression $E(X_1 \mid Y = y)$ to be linear, as we assume in (i). We use the added condition (ii) to determine explicitly the distribution functions (d.f.'s) of X_1 and X_2 (and hence of Y) for various choices of $u(y)$. We prove in Section 2 a theorem on which our characterizations are based and illustrate the theorem in Section 3.

2. The theorem. Let X_1 and X_2 be independent r.v.'s with characteristic functions (ch.f.'s) φ_1 and φ_2, and let the ch.f. and d.f. of $Y = X_1 + X_2$ be denoted by g and G, respectively. If Y has finite second moment and $E(X_1 \mid Y = y)$ and $V(X_1 \mid Y = y)$ are given by (i) and (ii), then

$$\varphi_i(t) = [g(t)]^{\lambda_i/\lambda}, \qquad i = 1, 2 \tag{1}$$

and

$$g(t)(d^2/dt^2) \ln g(t) = -\int e^{ity}u(y)\, dG(y). \tag{2}$$

Proof. Let $\varphi(t, u)$ be the joint ch.f. of Y and X_1. Then $\varphi(t, u) = E[e^{itY + iuX_1}] = \varphi_1(t + u)\varphi_2(t)$ and $g(t) = \varphi_1(t)\varphi_2(t) = \varphi(t, 0)$. Also,

$$\varphi(t, u) = \int e^{ity}[\int e^{iux_1}\, dF_y(x_1)]\, dG(y) \tag{3}$$

where F_y is the conditional d.f. of X_1 for fixed $Y = y$. Thus,

$$\begin{aligned}(\partial/\partial u)\varphi(t, u)|_{u=0} &= i\int e^{ity}[\int x_1\, dF_y(x_1)]\, dG(y) \\ &= (i\lambda_1/\lambda)\int ye^{ity}\, dG(y) = (\lambda_1/)g'(t.\end{aligned} \tag{4}$$

But $(\partial/\partial u)\varphi(t, u)|_{u=0} = \varphi_1'(t)\varphi_2(t)$, so that

$$\varphi_1'(t)\varphi_2(t) = (\lambda_1/\lambda)g'(t). \tag{5}$$

Since g is a ch.f., does not vanish in a neighborhood of the orgin. Dividing both sides of (5) by g, it follows that $\lambda_2(d/dt) \ln \varphi_1(t) = \lambda_1(d/dt) \ln \varphi_2(t)$. Intergrating and using the boundary conditions $\varphi_1(0) = g(0) = 1$, we obtain

Received 16 March 1964; revised 29 October 1964.

[1] Now at Bucknell University.

$\lambda_2 \ln \varphi_1(t) = \lambda_1 \ln \varphi_2(t)$, from which (1) follows immediately. On differentiating $\varphi(t, u)$ twice, we see from (3) that

$$
\begin{aligned}
(\partial^2/\partial u^2)\varphi(t, u)|_{u=0} &= -\int e^{ity} E(X_1^2 \mid Y = y)\, dG(y) \\
&= -\int e^{ity}[(\lambda_1\lambda_2/\lambda^2)u(y) + (\lambda_1 y/\lambda)^2]\, dG(y) \\
&= -(\lambda_1\lambda_2/\lambda^2) \int e^{ity} u(y)\, dG(y) + (\lambda_1/\lambda)^2 g''(t).
\end{aligned}
\tag{6}
$$

Differentiating each side of (5) and noting that $\lambda_1 \ln \varphi_2(t) = \lambda_2 \ln \varphi_1(t)$, we find that

$$
\lambda_1^2 g''(t) = \lambda[\varphi_1(t)]^{\lambda_2/\lambda_1}\{\lambda_2[\varphi_1'(t)]^2/\varphi_1(t) + \lambda_1\varphi_1''(t)\}. \tag{7}
$$

Also,

$$
(\partial^2/\partial u^2)\varphi(t, u)|_{u=0} = (\partial^2/\partial u^2)\varphi_1(t + u)\varphi_2(t) = \varphi_1''(t)[\varphi_1(t)]^{\lambda_2/\lambda_1}. \tag{8}
$$

Equation (2) now follows directly from (6), (7), and (8).

REMARK. The referee pointed out that if X_1 and X_2 are independent r.v.'s, $Y = X_1 + X_2$, $E(|Y|) < \infty$, and $\lambda_2 \ln \varphi_{X_1}(t) = \lambda_1 \ln \varphi_{X_2}(t)$, then $E(X_1|Y = y) = (\lambda_1 y/\lambda)$ a.e. For the assumptions immediately imply Equation (4), which in turn implies that $\int e^{ity}[E(X_1|Y = y) - \lambda_1 y/\lambda]\, dG(y) = 0$. Then, from the uniqueness theorem of Fourier transforms of functions of bounded variation, it follows that $E(X_1|Y = y) = \lambda_1 y/\lambda$ a.e. If we further assume that $E(Y^2) < \infty$, then proceeding as in the proof of the theorem, one obtains Equation (2), with $(\lambda^2/\lambda_1\lambda_2)V(X_1|Y = y)$ replacing $u(y)$. It then follows that $(\lambda^2/\lambda_1\lambda_2)V(X_1|Y = y)$ is a function of y and an unbiased estimate of $V(Y)$.

3. An example. We consider the case when $u(y)$ is at most quadratic in y, say $u(y) = \alpha y^2 + \beta y + \gamma$. Under this assumption, Equation (2) can be written in the simpler form

$$
(1 - \alpha)(d^2/dt^2) \ln g(t) = \alpha[(d/dt) \ln g(t)]^2 + i\beta[(d/dt) \ln g(t)] - \gamma, \tag{9}
$$

which can be solved in general by first solving for $(d/dt) \ln g(t)$ by separation of variables and then finding $g(t)$ by integration. In the ensuing discussion of the solutions of the differential equation (9), it is convenient to distinguish three cases, corresponding to (i) $\Delta = 0$, (ii) $\Delta > 0$, and (iii) $\Delta < 0$, where $\Delta = \beta^2 - 4\alpha\gamma$. In what follows, we put $i\theta = g'(0)$.

(i) $\Delta = 0$. We must distinguish three cases corresponding to various values of α. If $\alpha = 0$ and $\beta = 0$, with $\gamma > 0$, we find that $\ln g(t) = i\theta t - \gamma t^2/2$, which is the cumulant generating function of the normal distribution with mean θ and variance γ. Then, X_i, $i = 1, 2$, is normally distributed with mean $\theta\lambda_i/\lambda$ and variance $\gamma\lambda_1/\lambda$. If γ is zero, g is degenerate, while the case $\gamma < 0$ is impossible. For $\alpha = 1$, g is again a degenerate ch.f. Finally, if $\alpha \neq 0$ or 1, we find that $g(t) = [1 - ict/\mu]^{-\mu} e^{-it(\gamma/\alpha)}$, $c = \theta + (\gamma/\alpha)^{\frac{1}{2}}$, $\mu = (1 - \alpha)/\alpha$. In order that $g(t)$ be a ch.f., we must have $\mu > 0$ (implying that $0 < \alpha < 1$ and γ is non-negative), in which case $g(t)$ is the ch.f. of a gamma distribution.

(ii) $\Delta > 0$. We consider various cases. If $\alpha = 1$, we find that $g(t)$ is the ch.f. of a degenerate distribution. If $\alpha = 0$, so that $\beta \neq 0$, then $u(y) = \beta y + \gamma$, and we find that $\ln g(t) = -i\gamma t/\beta + \nu(e^{i\beta t} - 1)$, $\nu = \beta^{-2}(\theta\beta + \gamma)$, so that $g(t)$ is the ch.f. of a Poisson-type distribution. The p.d.f. of Y is given by $P(Y = y) = e^{-\nu}\nu^j/j!$, for $y = \beta j - (\gamma/\beta)$, $j = 0, 1, \cdots$. In particular, if $\beta = 1$ and $\gamma = 0$, then Y has p.d.f. $e^{-\theta}\theta^y/y!$, $y = 0, 1, \cdots$. Chatterji [1] has recently given a characterization of the Poisson distribution under a set of assumptions equivalent to assuming that the conditional density of X, given $Y = y$, is binomial.

It remains to discuss the case when $\alpha \neq 0$ or 1. In this case we find that

$$g(t) = e^{-i\beta t/2\alpha}[pe^{\frac{1}{2}\rho t} + qe^{-\frac{1}{2}\rho t}]^{-\mu} \tag{10}$$

with $\mu = (1 - \alpha)/\alpha$, $\rho = (4\alpha\gamma - \beta^2)^{\frac{1}{2}}/(1 - \alpha)$, $p = \delta(1 - \delta)^{-1}$, $q = 1 - p$, and $\delta = [(2\alpha\theta + \beta)i - (1 - \alpha)\rho]/[(2\alpha\theta + \beta)i + (1 - \alpha)\rho]$, so that ρ is purely imaginary and δ is real. Following the discussion of Lukacs [3], we conclude that if p and q are both positive, then (10) can be a ch.f. only if μ is a negative integer. The corresponding distribution is in this case a binomial distribution. If p and q are not both positive, then one of them must be positive and the other negative (since $p + q = 1$). In either of these cases, (10) is the ch.f. of a negative binomial distribution, provided μ is positive.

(iii) $\Delta < 0$. It is not hard to show that if $\alpha < 0$ or $\alpha > 1$, then the solutions to (9) are not ch.f.'s, so that we need only examine the two cases $\alpha = 1$ and $0 < \alpha < 1$ (note that α cannot be zero if $\Delta < 0$). We consider the situation when $0 < \alpha < 1$ first. Then the solution to (9) is again given by (10). Since $\Delta < 0$, we see that ρ is now real and δ is complex. Lukacs [3] has shown that in order for $g(t)$ to be ch.f. in this case, we must have $\mu > 0$ and $p = q = \frac{1}{2}$, so that (10) reduces to $g(t) = [\cosh \frac{1}{2} \rho t]^{-\mu}$.

Finally, if $\alpha = 1$ with $\Delta < 0$, we obtain the ch.f. of the Cauchy distribution if the further assumption is made that the resulting differential equation (9) is valid at all points except the origin. This last case shows that it is not really necessary to assume the existence and finiteness of the first moments if this additional assumption is made.

Acknowledgment. The authors thank the referee for his many helpful suggestions.

REFERENCES

[1] Chatterji, S. D. (1963). Some elementary characterizations of the Poisson distribution. *Amer. Math. Monthly* **70** 958–964.

[2] Laha, R. G. (1956). On a characterization of the stable law with finite expectation. *Ann. Math. Statist.* **27** 187–195.

[3] Lukacs, E. (1952). The stochastic independence of symmetric and homogeneous linear and quadratic statistics. *Ann. Math. Statist.* **23** 442–449.