Statistics & Decisions 11, 121-131 (1993)

SOME CHARACTERIZATIONS BY CONSTANT REGRESSION WITH RESPECT TO RESIDUALS

Jacek Wesołowski

Received: Revised version: September 16, 1992

Abstract. New characterizations of the normal, Poisson, binomial and negative binomial distributions are obtained. The results are based on a condition of the constant regression of some functions of independent random variables on residuals. Two applications for optimal estimators are also given.

1. INTRODUCTION

Consider the following problem: Let $X_1, \ldots, X_n$ be a sample from a population with a distribution function (d.f.) $F_\theta(x) = F(x-\theta)$. Suppose that $h(X_1, \ldots, X_n)$ is an unbiased estimator of a parametric function $g(\theta)$. Find such a family of d.f's for which h is an optimal estimator with respect to the quadratic loss. This problem may be reduced by application of a classical result from Rao (1952) to a particular case of the following characterizational scheme (for details see Kagan (1989)):

Let $X_1, \ldots, X_n$ be independent random variables (r.v's) and $h : \mathbb{R}^n \longrightarrow \mathbb{R}$ a Borel function such that $E|h(X_1, \ldots, X_n)| < \infty$. Denote by $Y = (X_1 - \bar{X}, \ldots, X_n - \bar{X})$ the vector of residuals, $\bar{X} = (X_1 + \ldots + X_n)/n$. What are the distributions of the r.v's $X_1, \ldots, X_n$ if

---

AMS (1980) subject classification: 62E10
Key words and phrases: normal law, Poisson law, binomial law, negative binomial law, constant regression, characterization of probability distributions, residuals.

$$E [h(X_1, \ldots, X_n) \mid Y] = \text{const} \quad ? \tag{1}$$

No general solution of this problem is known now. However in some cases the condition (1) determines the distributions of the r.v's - see Kagan et al. (1965), Khatri and Rao (1968), Bondesson (1974), Wesołowski (1987), Kagan and Rao (1988) and Wesołowski (1989).

In this note we give several new contributions taking into account new types of the function h. Two applications for optimal estimators are contained in Section 2.

2. A CHARACTERIZATION OF THE NORMAL LAW

It follows from the Darmois - Skitovitch theorem that independent r.v's $X_1, \ldots, X_n$ have normal distributions with a common variance if only the r.v. $S = n\bar{X}$ and the random vector Y are independent. A natural question arising here is if independence of S and Y may be replaced by the constancy of regression of a function of S on Y, i.e. by the condition

$$E [g(S) \mid Y] = \text{const} \quad , \tag{2}$$

where g is a Borel function such that $E|g(S)| < \infty$. This problem is positively solved for

(i) $g(x) = x$, in Kagan et al. (1965);

(ii) $g(x) = x^2$, in Wesołowski (1987) (with the assumption $n \geq 4$ lacking in the formulation of the result);

(iii) $g(x)$ = polynomial in x, in Kagan and Rao (1988) (with some additional assumptions).

In this note we investigate the case $g(x) = \exp(\alpha x)$, for a real $\alpha \neq 0$. It appears that even if we limit ourselves to absolutely continuous distributions the class of probability laws for which the condition (2) holds is wider than the normal distributions.

THEOREM 1. Let $X_1, \ldots, X_n$, $n \geq 3$, be independent r.v's such that $E[\exp(\alpha X_k)] < \infty$, $k = 1, \ldots, n$, for a real number $\alpha \neq 0$ and

$$E\,[\,\exp(\alpha S)\mid Y\,] = \text{const}\quad. \tag{3}$$

Then the density f_k of the r.v. X_k has a form

$$f_k = g_k h_k \quad, \tag{4}$$

where g_k is a normal density with a variance σ^2 not depending on k and h_k is a periodic function with a period $\alpha\sigma^2$ such that $h_k \geq 0$, $\int_{\mathbb{R}} g_k h_k = 1$, $k = 1, \ldots, n$. The factorization (4) is unique.

P r o o f : It is well known (see for example Lemma 1.1.1 in Kagan et al. (1973)) that (3) is equivalent to

$$E[\exp(\alpha \sum_{k=1}^{n} X_k)\exp(i \sum_{k=2}^{n} t_k(X_k - X_1))] =$$

$$= c\; E[\exp(i \sum_{k=2}^{n} t_k(X_k - X_1))] \tag{5}$$

for any real $t_2, \ldots, t_n$. Define

$$\phi_k(t) = \ln\{\, E\,[\exp(\alpha + itX_k)]\;/\;E\,[\exp(itX_k)]\,\}$$

$k = 1, \ldots, n$, for suitably small $|t|$. Then (5) yields

$$\phi_1(-\sum_{k=2}^{n} t_k) + \sum_{k=2}^{n} \phi_k(t_k) = c$$

for t_k, $k = 1, \ldots, n$, from a neighbourhood of the origin. Now by Lemma 1.5.1 in Kagan et al. (1973)) we obtain

$$E[\exp(\alpha X_k)\exp(itX_k)] = a_k \exp(ibt) E[\exp(itX_k)] \tag{6}$$

for some real numbers $a_k > 0$, $b \neq 0$ depending on α and $|t| < \varepsilon$, where $\varepsilon > 0$ is suitably small, $k = 1, \ldots, n$. Choose any k. Assume that $\alpha > 0$. Define $w_k(z) = a_k \exp(ibz) E\exp(izX_k)$ for $\bar{D} = \{z: \mathrm{Re}(z) \in [-\varepsilon, \varepsilon]$ and $\mathrm{Im}(z) \in [-\alpha, 0]\}$ and $s_k(t) = \int_{\mathbb{R}} \exp(ity)\, dS_k(y)$, where $dS_k(y) = \exp(\alpha)\, dF_k(y)$ (F_k is a distribution function of X_k), $t \in \mathbb{R}$. Then w_k is continuous in $\bar{D}$ and analytic in $D = \{z: \mathrm{Re}(z) \in (-\varepsilon, \varepsilon)$ and $\mathrm{Im}(z) \in (-\alpha, 0)\}$. Additionally w_k and s_k are equal on $(-\varepsilon, \varepsilon)$. Consequently by Theorem 3.2.1 from Ramachandran (1967) we can extend the equation (6) to the whole real line. Similar argumentation is valid in the case $\alpha < 0$.

Hence

$$\int_{\mathbb{R}} e^{itx} e^{\alpha x} f_k(x)\, dx = a_k \int_{\mathbb{R}} e^{itx} f_k(x - b)\, dx$$

for any $t \in \mathbb{R}$. The equality of the Fourier transforms yields

$$e^{\alpha x} f_k(x) = a_k f_k(x - b) \quad , \quad x \in \mathbb{R}. \tag{7}$$

Define $H = f_k/G$, where

$$G(x) = \exp[-\alpha(x - \ln(a_k)/\alpha + b/2)^2/(2b)] \quad , \quad x \in \mathbb{R}.$$

Substituting then $f_k = GH$ in (7), after easy computations, we get

$$H(x) = H(x - b) \quad , \quad x \in \mathbb{R}.$$

Consequently we have (4) with g_k being a normal density with the mean $\ln(a_k)/\alpha - b/2$ and the variance $\sigma^2 = |b/\alpha|$.

Assume now that $f_k = \underline{g}_k\underline{h}_k$ is another factorization of the type (4) with $\underline{g}_k$ being another normal density and suitable periodic function $\underline{h}_k$. From the equation $g_k h_k = \underline{g}_k\underline{h}_k$ we immediately obtain the uniqueness of (4). ∎

Now let us consider again a problem in the theory of optimal estimators: It is well known that $h(\bar{X}) = \exp(\bar{X} - \frac{\sigma^2}{2n})$ is an optimal unbiased estimator of the parameter function e^m in the normal population with the variance σ^2 and the unknown mean m. Is this a characteristic property of the normal distribution? An answer gives the following

COROLLARY 1. Let $X_1, \ldots, X_n$ be i.i.d. observations from an absolutely continuous distribution with a d.f. $F_\theta(x) = F(x-\theta)$ depending on a location parameter $\theta \in \mathbb{R}$. Suppose that $E_\theta[\exp(2X_1/n)] < \infty$. If $K_n\exp(\bar{X})$ (K_n being non-random) is the best unbiased estimator of the parameter function e^θ with respect to the quadratic loss then the density of F has the form (4).

P r o o f : Since $K_n\exp(\bar{X})$ is square integrable then by the Rao theorem mentioned in Section 1 (see also Lemma 7.2.1 in Kagan et al. (1973)) we have

$$E_\theta[K_n \exp(\bar{X})\ H] = 0 \quad , \quad \theta \in \mathbb{R}$$

for any measurable bounded H such that $E_\theta[H] = 0$, $\theta \in \mathbb{R}$. It suffices to take all the H's of the form H = H(Y). Consequently for any such H's

$$E_0[E_0(\exp(\bar{X})|Y)\ H(Y)] = 0$$

and thus the assumptions of Th. 1 with $\alpha = 1/n$ hold. ■

We observe that condition (3) fulfilled for a single value α is not sufficient for characterizing the normal law. On the other hand it is quite easy to show that $X_1, \ldots, X_n$ are normal r.v's if only (3) holds for any real α. It appears that to obtain a characterization of the normal distribution it suffices to assume (3) for two suitably chosen values of α.

THEOREM 2. Let α_1 and α_2 be real non-zero numbers such that their quotient is irrational. Assume that $X_1, \ldots, X_n$, $n \geq 3$, are independent r.v's with continuous densities and $E[\exp(\alpha_j X_k)] < \infty$, $j = 1, 2$, $k = 1, \ldots, n$. If the condition (3) holds for $\alpha = \alpha_j$, $j = 1, 2$, then $X_1, \ldots, X_n$ have normal distributions with a common variance.

P r o o f : From Th. 1 for any $k = 1, \ldots, n$,

$$f_k = g_{1k}h_{1k} = g_{2k}h_{2k} ,$$

where two factorizations (4) are associated with two values of α. By uniqueness of the factorization we have $g_k = g_{1k} = g_{2k}$ and $h_k = h_{1k} = h_{2k}$ with g_k being a normal density and h_k - a periodic function with periods $\beta_j = \alpha_j\sigma^2$, $j = 1, 2$. Define a set of periods of h_k:

$$A = \{ y \in \mathbb{R} : h_k(x) = h_k(x + y) \ \forall \ x \in \mathbb{R} \}.$$

The set A is closed with respect to addition and substraction and $\beta_j \in A$, $j = 1, 2$. Since the quotient β_1/β_2 is irrational then A is dense in $\mathbb{R}$. Obviously $h_k|_A$ is constant. Consequently the regularity assumption imposed on f_k implies $h_k = 1$ on $\mathbb{R}$. Hence $f_k = g_k$. ■

Observe that similar argument as in the proof of Corollary 1 leads to the following application of the above Theorem.

COROLLARY 2. Let $X_1, \ldots, X_n$ be i.i.d. observations with an absolutely continuous d.f. $F_\theta(x) = F(x - \theta)$ having a continuous density. Suppose that $E_\theta[\exp(\alpha_i X_1/n)] < \infty$, $i = 1, 2$, where α_1, α_2 are real numbers such that α_1/α_2 is irrational. If the statistic $K_i(n)\exp(2\alpha_i \bar{X})$ ($K_i(n)$ being non-random) is the best unbiased estimator of the parametric function $\exp(\alpha_i\theta)$, $i = 1, 2$ with respect to quadratic loss then F is a normal d.f. ∎

Remark: Conversely, if $X_1, \ldots, X_n$ are i.i.d. observations from $F(x - \theta)$, $\theta \in \mathbb{R}$, where F is normal, then $\bar{X}$ is a complete sufficient statistic and thus $g(\bar{X})$ is the best unbiased estimator of its expectation for any Borel function g such that $E\, g^2(\bar{X}) < \infty$.

3. A CHARACTERIZATION OF THE POISSON DISTRIBUTION

In Wesołowski (1989) the Poisson distribution was characterized by considering (1) with $h(X_1, \ldots, X_n) = X_1 \cdot \ldots \cdot X_n$. Now we investigate the product of inverses: $X_1^{-1} \cdot \ldots \cdot X_n^{-1}$. We recall that the condition (1) with h being the sum of inverses was used in Khatri and Rao (1968) to characterize the gamma law.

THEOREM 3. Let $X_1, \ldots, X_n$, $n \geq 3$, be independent r.v's such that $E|X_k^{-1}| < \infty$, $k = 1, \ldots, n$. If

$$E\left[\prod_{k=1}^{n} X_k^{-1} \mid Y\right] = \text{const} \tag{8}$$

then $X_1, \ldots, X_n$ have Poisson type distributions with the same common location and scale parameter.

P r o o f : Similarly as in the proof of Th. 1 from (8) we have for any $k = 1, \ldots, n$,

$$E[X_k^{-1}\exp(itX_k)] = a_k\exp(-ibt)E[\exp(itX_k)]$$

for suitably small $|t| < \varepsilon$ and real $a_k \neq 0$, b. We differentiate the above equation and get

$$i\phi_k(t) = -ia_k b\exp(-ibt)\,\phi_k(t) + a_k\exp(-ibt)\,\phi_k'(t) \quad ,$$

$|t| < \varepsilon$, where ϕ_k is a characteristic function of X_k. Hence

$$\phi_k'(t)/\phi_k(t) = i(\exp(ibt) + a_k b)/a_k \quad , \quad |t| < \varepsilon. \tag{9}$$

For $b = 0$ we have $\phi_k(t) = \exp(i\alpha t)$ and since in this case $(EX_k)^{-1} = EX_k^{-1}$ then $\alpha = \pm 1$. Now assume $b \neq 0$. Then the solution of (9) has a form

$$\phi_k(t) = \exp(ibt)\exp(c_k(e^{ibt} - 1) \quad , \quad |t| < \varepsilon,$$

where $c_k = (a_k b)^{-1}$ since $Var(X_k) = b/a_k$. The final result is a consequence of the fact that the Poisson characteristic function is uniquely determined by its form in a neighbourhood of the origin. ■

4. A CHARACTERIZATION OF THE BINOMIAL AND NEGATIVE BINOMIAL DISTRIBUTIONS

Our last result is a joint characterization of the binomial and negative binomial laws by the condition (1) with a suitable function h.

THEOREM 4. Let $X_1, \ldots, X_n$, $n \geq 3$, be independent r. v's with $E|X_k^{-1}| < \infty$, $k = 1, \ldots, n$. If

$$E\left[\prod_{k=1}^{n} (1 + X_k^{-1}) \mid Y\right] = \text{const} \tag{10}$$

then for any $k = 1, \ldots, n$, X_k has a

(i) binomial distribution if only $EX_k^{-1} < -1$;

(ii) negative binomial distribution if only $EX_k^{-1} > -1$;

(iii) degenerate distribution if only $EX_k^{-1} = -1$.

P r o o f : As earlier from (10) we conclude that

$$E[(1 + X_k^{-1})\exp(itX_k)] = a_k \exp(ibt) E[\exp(itX_k)] \quad ,$$

$|t|<\varepsilon$. Hence $a_k = 1 + EX_k^{-1}$ and thus for $a_k = 0$ from the above equation we have $\phi_k'(t) = -i\phi_k(t)$, $|t| < \varepsilon$. Consequently X_k is concentrated in -1. Now assume $a_k \neq 0$. Then we have

$$\frac{\phi_k'(t)}{\phi_k(t)} = i\,\frac{a_k \ b\exp(ibt) - 1}{1 - a_k \ \exp(ibt)} \quad , \qquad |t| < \varepsilon. \tag{11}$$

Observe that $b \neq 1$ since if $b = 1$ then $a_k = 0$. The solution of (11) for $b \neq 0$ has the following form

$$\phi_k(t) = [p_k \ \exp(it\mu_1) + q_k \ \exp(it\mu_2)]^{\rho} \quad , \qquad |t| < \varepsilon,$$

where $p_k = (1 - a_k)^{-1} = 1 - q_k$, $\rho = 1/b - 1$, $\mu_1 = b/(b - 1)$, $\mu_2 = b^2/(b - 1)$. Since $\mathrm{Var}(X_k) = a_k b(b - 1)(1 - a_k)^{-2}$ then

1° $a_k < 0$ implies $\rho > 0$, $q_k > 0$, $p_k > 0$ (binomial law);

2° $a_k > 0$ implies $\rho < 0$, $q_k p_k < 0$ (negative binomial law).

The final result is a consequence of the fact that both the distributions are uniquely determined by their moments.

For $b = 0$ the formula for the variance implies degeneracy of X_k. ■

Remark. The results of this and the previous section may be combined together and slightly generalized (with application of the same method) by considering the condition

$$E\left[\prod_{k=1}^{n} (\alpha_k + \beta_k X_k^{-1}) \mid Y\right] = \text{const} \quad .$$

References

[1] Bondesson, L.: Characterizations of probability laws through constant regression. Z. Wahrschein. verw. Geb. 29, 95-115 (1974).

[2] Kagan, A.: Constancy of regression of a polynomial of sample average of positive random variables on their ratios characterizes gamma distribution. In: Statistical Data Analysis and Inference, Y. Dodge (Ed.), 309-316, Elsevier Science Publishers B. V. (North Holland) (1989).

[3] Kagan, A.M., Linnik Yu.V., Rao C.R., On a characterization of the normal law based on a property of the sample average. Sankhya, Ser. A 27, 405-406 (1965).

[4] Kagan, A.M., Linnik, Yu.V., Rao, C.R.: Characterization Problems in Mathematical Statistics, New York, Wiley (1973).

[5] Kagan, A. and Rao, C.R.: Constancy of regression of a polynomial of sample average on residuals characterizes normal distribution. Technical Report No. 88-30, Center for Multivariate Analysis, Univ. of Pittsburgh (1988).

[6] Khatri, C.G. and Rao, C.R.: Some characterizations of the gamma distribution, Sankhya, Ser. A, 30, 157-166 (1968).

[7] Ramachandran, B.: Advanced Theory of Characteristic Functions, Calcutta, Statist. Publ. Soc. (1967).

[8] Rao, C.R.: Some theorems on minimum variance estimation. Sankhya, Ser. A 12, 27-42 (1952).

[9] Wesołowski, J.: A regressional characterization of the normal law, Statist. Probab. Letters, 6, 11-12 (1987).

[10] Wesołowski, J.: A regressional characterization of the Poisson distribution. In: Stability Problems for Stochastic Models (Proc., Sukhumi, 1987), Lect. Notes Math. 1412, 349-351, New York - Berlin - Heidelberg, Springer (1989).

Jacek Wesołowski
Mathematical Institute
Warsaw University of Technology
Plac Politechniki 1
00-661 Warsaw
Poland