# ON THE STRUCTURAL PROPERTIES OF THE CONDITIONAL DISTRIBUTIONS

A.M. Mathai

(received March 7, 1966)

Summary. If $x, x_1, \ldots, x_n$ are independent stochastic variables and if the conditional distribution of $x$ given $x_1 + \ldots + x_n$ is known, what can be said about the marginal distributions of $x, x_1, \ldots, x_n$ ? In this paper we will show that if the conditional distribution of $x$ given a subset of $x_1, x_2, \ldots, x_{n-1}$, $x + x_1 + \ldots + x_n$ has a certain structural form then $x, x_1, \ldots, x_n$ are distributed as members of the linear exponential family of distributions and further $x_1, \ldots, x_n$ are identically distributed. A few interesting corollaries are obtained and it is pointed out that with the help of the characterization properties parametric inference problems can be transformed into non-parametric inference problems.

1. Introduction. Mauldon (1961) has studied the problem of determining the parent distribution through the knowledge of the distribution of a sample statistic. Moran (1951) dealt with the characterization of the Poisson distribution from the conditional distribution of $x$ given $x+y$ in a bivariate case. Patil and Seshadri (1964) considered the problem of characterizing the Binomial, Poisson, Exponential, Normal and Power series from the conditional distribution of $x$ given $x+y$, in a bivariate case.

For convenience, the stochastic variables as well as the values assumed by them will be denoted by the same letters. A stochastic variable $x$ is said to have a linear exponential distribution if its probability function has the form

$$f(x) = a(x)\, e^{wx} / J(w) ,$$

where $a(x) > 0$, $w \in \Omega$ (parameter space) and $J(w) > 0$ is a

Canad. Math. Bull. vol. 10, no. 2, 1967

normalizer such that

$$J(w) = \int_T a(x)\, e^{wx} ,$$

where $T$ is a subset of the set of real numbers, $T$ may be discrete or continuous and the integration or summation is denoted by $\int_T$ .

## 2. A Characterization Theorem.

THEOREM. Let $x, x_1, \ldots, x_n$ be independent non-degenerate stochastic variables whose probability functions do not vanish at the origin. Let the conditional distribution of $x$ given $x_1, x_2, \ldots, x_{n-1}, x+x_1+\ldots+x_n$ have the structural form $C(x, z)$. Let the conditional distributions of $x_i$ given subsets of $x, x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_{n-1}, x+x_1+\ldots+x_n$ have the structural forms $C_i(x_i, z)$ for all $i$ and for every subset, where $z = x+x_1+\ldots+x_n$. If $C(x, z)$ is such that

$$\frac{C(x, z)\, C(x_1, z) \ldots C(x_n, z)\, C(0, z)}{C(0, z)\, C(0, z) \ldots C(0, z)\, C(z, z)} = \frac{h(x)\, h(x_1) \ldots h(x_n)}{h(x+x_1+\ldots+x_n)}$$

for some non-negative measurable function $h(x)$ then $x, x_1, \ldots, x_n$ all belong to the linear exponential family and further $x_1, \ldots, x_n$ are identically distributed.

Proof. $f(x)\, f_1(x_1) \ldots f_n(x_n) = g(x, x_1, \ldots, x_{n-1}, x+x_1+\ldots+x_n)$

$$\begin{aligned} (1) \qquad &= g_1(x_1 | x, x_2, \ldots, x_{n-1}, z)\, g_2(x_2 | x, x_3, \ldots) \ldots g_n(x|z)\, g_{n+1}(z) \\ &= C_1(x_1, z)\, C_2(x_2, z) \ldots C_{n-1}(x_{n-1}, z)\, C(x, z)\, C_0(z) \end{aligned}$$

where $f(.)$, $f_i(.)$, $g(.)$ denote the corresponding probability functions of $(.)$, $g_i(.|.)$ denotes the conditional distributions, and $g_{n+1}(z)$ is the marginal distribution of $Z$.
In (1) put $x = 0$ and replace $x_n$ by $x_n + x$ to get

$$(2) \qquad f(0)\, f_1(x_1) \ldots f_n(x_n+x) = C_1(x_1, z) \ldots C_{n-1}(x_{n-1}, z)\, C(0, z)\, C_0(z) .$$

Dividing (1) by (2) we get

$$(3) \qquad \frac{f(x)\ f_n(x_n)}{f(0)\ f_n(x_n+x)} = \frac{C(x, z)}{C(0, z)} .$$

In (3) put $x_n = 0$ and replace $x_{n-1}$ by $x_{n-1} + x_n$ ; we then have,

$$(4) \qquad \frac{f(x)\ f_n(0)}{f(0)\ f_n(x)} = \frac{C(x, z)}{C(0, z)} .$$

Since $f(x)\, f_1(x_1) \ldots f_n(x_n)$ can be written as a joint distribution of $x, x_1, \ldots, x_{i-1}, x_{i+1}, \ldots, x_n, z,$ we have,

$$(5) \qquad \frac{f(x)\ f_i(0)}{f(0)\ f_i(x)} = \frac{C(x, z)}{C(0, z)} .$$

Equation (5) indicates that $x_1, \ldots, x_n$ are identically distributed. In (3) put $x_n = 0$ , replace $x$ by $x + x_n$ , and thus obtain

$$(6) \qquad \frac{f(x+x_n)\ f_n(0)}{f(0)\ f_n(x+x_n)} = \frac{C(x+x_n, z)}{C(0, z)} .$$

Now in (6) put $x = 0$ and replace $x_{n-1}$ by $x+x_{n-1}$ , obtaining

$$(7) \qquad \frac{f(x_n)\ f_n(0)}{f(0)\ f_n(x_n)} = \frac{C(x_n, z)}{C(0, z)} .$$

Multiplying (3) by (7),

$$(8) \qquad \frac{f(x)\ f(x_n)\ f_n(0)}{f(0)\ f(0)\ f_n(x_n+x)} = \frac{C(x, z)\ C(x_n, z)}{C(0, z)\ C(0, z)} .$$

In (3) put $x_n = 0$ and replace $x$ by $x+x_n$ , so that

$$(9) \qquad \frac{f(x+x_n)\ f_n(0)}{f(0)\ f_n(x+x_n)} = \frac{C(x+x_n, z)}{C(0, z)} .$$

Dividing (8) by (9), we get

$$(10)\qquad \frac{f(x)\ f(x_n)\ f(0)}{f(0)\ f(0)\ f(x+x_n)} = \frac{C(x,z)\ C(x_n,z)\ C(0,z)}{C(0,z)\ C(0,z)\ C(x+x_n,z)} .$$

Now in (10) put $x_1 = 0$ and replace $x$ by $x+x_1$, obtaining

$$(11)\qquad \frac{f(x+x_1)\ f(x_n)\ f(0)}{f(0)\ f(0)\ f(x+x_1+x_n)} = \frac{C(x+x_1,z)\ C(x_n,z)\ C(0,z)}{C(0,z)\ C(0,z)\ C(x+x_1+x_n,z)} .$$

In (11) putting $x = 0$ and replacing $x_n$ by $x_n + x$ we see that

$$(12)\qquad \frac{f(x_1)\ f(x+x_n)}{f(0)\ f(x+x_1+x_n)} = \frac{C(x_1,z)\ C(x+x_n,z)}{C(0,z)\ C(x+x_1+x_n,z)} .$$

Multiplying (10) and (12) we get

$$(13)\qquad \frac{f(x)\ f(x_1)\ f(x_n)\ f(0)}{f(0)\ f(0)\ f(0)\ f(x+x_1+x_n)} = \frac{C(x,z)\ C(x_1,z)\ C(x_n,z)\ C(0,z)}{C(0,z)\ C(0,z)\ C(0,z)\ C(x+x_1+x_n,z)} .$$

Proceeding in a similar way, we arrive at the result

$$\frac{f(x)\ f(x_1)\ \ldots\ f(x_n)\ f(0)}{f(0)\ f(0)\ \ldots\ f(0)\ f(z)} = \frac{C(x,z)\ C(x_1,z)\ldots C(x_n,z)\ C(0,z)}{C(0,z)\ C(0,z)\ \ldots\ C(0,z)\ C(z,z)} ,$$

that is

$$\frac{f(x)\ f(x_1)\ \ldots\ f(x_n)\ f(0)}{f(0)\ f(0)\ \ldots\ f(0)\ f(z)} = \frac{h(x)\ h(x_1)\ldots h(x_n)}{h(x+x_1+\ldots+x_n)} .$$

By putting $\phi(x) = f(x)/f(0)\ h(x)$ we get a Cauchy functional equation

$$\phi(x)\ \phi(x_1)\ \ldots\ \phi(x_n) = \phi(x+x_1+\ldots+x_n) ,$$

whose solution is

$$\phi(x) = e^{ax}$$

for some constant $a$, since $f(x)$ is a probability function, $f(0) \neq 0$ and $h(x)$ is measurable (Sierpinski, 1920), where $f(0)$



https://doi.org/10.4153/CMB-1967-022-2 Published online by Cambridge University Press

is an appropriate normalizer which will make $f(x)$ a probability function. Hence

$$f(x) = f(0)\ h(x)\ e^{ax}$$

From (5),

$$f_i(x) = f_i(0)\ f(x)\ C(0, z)/f(0)\ C(x, z)$$

$$= f_i(0)\ h(x)\ e^{ax}\ C(0, z)/C(x, z)\ .$$

This completes the proof.

COROLLARY 1. Let $x, x_1, \ldots, x_n$ be as defined in the theorem and let $C(x, z) = 1/z$, $0 < x < z$ for every given $z = x+x_1+\ldots+x_n$. Then $x, x_1, \ldots, x_n$ are identically distributed as negative exponential distributions.

Here it is easily seen that $h(x) = 1$ and therefore

$$f(x) = f(0)\ e^{ax} \quad \text{and} \quad f_i(x) = f_i(0)\ e^{ax}\ ,$$

where the normalizing factor $f(0)$ is easily seen to be $a$.

This corollary helps us to transform problems of statistical inference regarding the negative exponential population to inference problems regarding the rectangular distribution. Furthermore, this helps us to reduce a parametric hypothesis to a non-parametric hypothesis, since $z$ is known.

COROLLARY 2. Let $x, x_1, \ldots, x_n$ be as defined in the theorem. Let $C(x, z) = \text{Const. Exp.} -(x-z/2)^2/2\sigma^2$. Then $x, x_1, \ldots, x_n$ are identically normally distributed.

In this case it can be easily seen that $h(x) = e^{-x^2/2\sigma^2}$ and hence $f(x) = f(0)\ e^{-x^2/2\sigma^2} \cdot e^{ax}$ and $f_i(x) = f_i(0)\ e^{-x^2/2\sigma^2} \cdot e^{ax}$ where $a$ is a constant. The advantage of this result is that it enables us to go from a normal with unknown parameter $a$ to a normal with the known parameter $z$ when $\sigma$ is known.

COROLLARY 3. Let $x, x_1, \ldots, x_n$ be as defined in the theorem and let $C(x, z) = \text{Const. } z^{x+1}/(x+1)$ for all given $z$.

Then $x_1, \ldots, x_n$ are identically distributed as the geometric series distribution. In this case it can be easily seen that $h(x) = 1/(x+1)$ and hence

$$f_i(x) = f_i(0)\, e^{ax}\, (x+1)/(x+1)z^x = f_i(0)\, (e^a/z)^x \text{ for } x = 1, 2, \ldots,$$

which is a geometric distribution for $e^a < z$.

In statistical problems of testing hypotheses, in the non-distribution free procedures, usually a parent distribution is to be assumed. In order to get a test criterion having some desirable properties, the consideration of conditional distribution given an ancillary statistic arises. Also in other similar statistical problems, if the complete or partial knowledge of the conditional distribution determines the parent distribution, our assumptions regarding the underlying distribution in an experimental problem can be justified to a great extent. In many cases this procedure helps us to reduce a parametric problem to a non-parametric problem. This paper throws some light on these aspects of statistical analysis.

## REFERENCES


1. T.S. Ferguson, Location and scale parameters in exponential families of distributions. Ann. Math. Statist. 33 (1962a), pages 986-1001.

2. T.S. Ferguson, A characterization of the geometric distribution. Ann. Math. Statist. 33 (1962b), page 1207.

3. R.G. Laha, On the laws of Cauchy and Gauss. Ann. Math. Statist. 30 (1959), pages 1165-1174.

4. A.M. Mathai, Some characterizations of the one parameter family of distributions. Canadian Math. Bull. 9 (1966), pages 95-102.

5. J.G. Mauldon, Characterization properties of statistical distributions. Quart. J. Math. 7 (1961), pages 155-160.

6. P.A.P. Moran, A characteristic property of the Poisson distribution. Proc. Cam. Phil. Soc. 48 (1951), pages 206-207.

7. G.P. Patil and V. Seshadri, A characterization of a bivariate distribution by the marginal and the conditional distributions of the same component. Ann. Inst. Statist. Math. Tokyo, 15 (1963), pages 215-221.

8. G.P. Patil and V. Seshadri, Characterization theorems for some univariate probability distributions. Roy. Statist. Soc. Series, B. 26 (1964), pages 286-292.

9. W. Sierpinski, Sur l'equation fonctionnelle $f(x+y) = f(x) + f(y)$. Fund. Math. 1 (1920), pages 116-122.

McGill University