CONDITIONS FOR UNIQUENESS IN THE PROBLEM OF MOMENTS

By M. G. Kendall

It was shown by Stieltjes [1] that in some circumstances it is possible for two different frequency distributions to have the same set of moments. For instance, the integral

$$\int z^{4n+3} e^{-z} e^{iz} \, dz$$

around a contour consisting of the positive x-axis, the infinite quadrant and the positive y-axis is seen to be zero and it follows that

$$\int_0^\infty x^n e^{-x^{\frac{1}{4}}} \sin x^{\frac{1}{4}} \, dx = 0.$$

Thus the frequency distribution

$$dF = \tfrac{1}{6} e^{-x^{\frac{1}{4}}}(1 - \lambda \sin x^{\frac{1}{4}}) \, dx \qquad \begin{matrix} 0 \leq x \leq \infty, \\ 0 \leq \lambda \leq 1 \end{matrix} \tag{1}$$

has moments which are independent of λ, and equation (1) may be regarded as defining a whole family of distributions each of which has the same moments. It is easy to see that moments of all orders exist, and in fact

$$\mu_r' \text{ (about the origin) } = \tfrac{1}{6}(4r + 3)!.$$

A second example of the same kind, also due to Stieltjes, is the distribution

$$dF = \frac{1}{e^{\frac{1}{4}}\sqrt{\pi}} x^{-\log x}\{1 - \lambda \sin (2\pi \log x)\} \, dx \qquad \begin{matrix} 0 \leq x \leq \infty, \\ 0 \leq \lambda \leq 1, \end{matrix} \tag{2}$$

for which

$$\mu_r' = e^{\frac{1}{4}r(r+2)}.$$

The question naturally arises, what are the conditions under which a given set of moments determines a frequency distribution uniquely? The question is of great interest to mathematicians, being closely linked with problems in the theory of asymptotic series, continued fractions and quasi-analytic functions; and it also has importance for statisticians since there is sometimes occasion to be satisfied that a problem of finding a frequency distribution has been uniquely solved by the ascertainment of its moments or semi-invariants. Stieltjes himself considered a more general problem: given a set of constants c_0,

Institute of Mathematical Statistics is collaborating with JSTOR to digitize, preserve, and extend access to *The Annals of Mathematical Statistics.* www.jstor.org

$c_1, \cdots c_r, \cdots$ does there exist a function F, non-decreasing and possessing an infinite number of points of increase, such that

$$\int_0^\infty x^r\,dF = c_r \tag{3}$$

and under what conditions is F unique, except for an additive constant? Stieltjes showed that if we express the series

$$\sum_{r=0}^{\infty} (-1)^r \frac{c_r}{z^r} \tag{4}$$

as a continued fraction of the form

$$\frac{1}{a_1 z +}\,\frac{1}{a_2 +}\,\frac{1}{a_3 z +}\,\frac{1}{a_4 +}\cdots\frac{1}{a_{2n-1} z +}\,\frac{1}{a_{2n} +}\cdots \tag{5}$$

it is a necessary and sufficient condition for the existence of at least one F that all the a's be positive; and that the function is unique or not according as the series $\sum_{r=0}^{\infty} (a_r)$ diverges or converges. (If the a's are positive it must do one or the other.) The integral of equation (3) is to be interpreted in the general Stieltjes sense, so that the result applies to discontinuous as well as to continuous distributions. This is also true of the results obtained below.

Hamburger [2] discussed the similar problem when the limits of the integral in equation (3) are $\pm\infty$, and showed that a function F exists if the expression of (4) as a continued fraction of the form

$$\frac{b_0}{a_0 + z +}\,\frac{b_1}{a_1 + z +}\,\frac{b_2}{a_2 + z +}\cdots$$

gives positive values of the b's. In order that F may be unique it is necessary and sufficient that the continued fraction be completely (vollständig) convergent in the sense defined by Hamburger.

Unfortunately these criteria, though mathematically complete, are not very useful to statisticians because as a rule it is too difficult to express the coefficients a and b explicitly enough in terms of the given c's to enable questions of sign or of convergence to be decided. So far as I know, no more convenient criterion for the general Stieltjes problem has been found; but progress is possible if one considers the narrower question: given a set of moments, is the distribution which furnished them unique, that is to say, can any other distribution have furnished them? This is more limited than the Stieltjes problem because we know that at least one solution exists.

Contributions to this subject have been made by Lévy [3] and Carleman [4]. Lévy shows that if moments of all orders exist and are positive it is a sufficient condition for them to determine a distribution uniquely that $\mu_n^{1/n}/n$ remains finite as n tends to infinity. (Here and elsewhere in this paper μ_r refers to the moment of order r about any point, not necessarily the mean.) Carleman shows that, for the case of limits $-\infty$ to $+\infty$ the moments determine the distribution uniquely if

$$\sum_{r=0}^{\infty} \frac{1}{(\mu_{2r})^{1/(2r)}}$$

diverges. For the limits 0 to ∞ he gives the corresponding series

$$\sum_{r=0}^{\infty} \frac{1}{(\mu_r)^{1/(2r)}}$$

a criterion which can be improved upon, as will be shown below.

The purpose of this paper is to develop criteria of this kind more systematically and to give more general criteria suitable in cases where the moments are not known explicitly but the behavior of the frequency distribution at its terminals is known.

Three preliminary points necessary for the later argument may be noted.

(1) Define the absolute moment of order r by

$$\nu_r = \int_{-\infty}^{\infty} | x^r | \, dF$$

and recall that

$$\nu_1 \leq \nu_2^{\frac{1}{2}} \leq \nu_3^{\frac{1}{3}} \leq \cdots \leq \nu_r^{1/r} \leq \cdots$$

(cf. Hardy and others, [5]). In other words the quantities $\nu_r^{1/r}$ form an increasing positive sequence and their reciprocals a decreasing positive sequence.

(2) The quantity $\nu_n^{1/n}/n$ must either tend to a limit or diverge to infinity as $n \to \infty$. For suppose that

$$\overline{\lim}\, \nu_n^{1/n}/n = k,$$
$$\underline{\lim}\, \nu_n^{1/n}/n = l.$$

Writing temporarily $\nu_n^{1/n} = a_n$, we have that, given ϵ there is an N such that

$$a_n/n > k - \epsilon$$

for an infinity of values of n greater than N. Similarly there is an M such that

$$a_n/n < l + \epsilon$$

for an infinity of values of n greater than M. Now choose ρ such that a_ρ, $a_{\rho+1}$ are two consecutive values, one near the upper limit and one near the lower limit. This can always be done and we can take ρ as large as we please. We then have

$$a_\rho > \rho(k - \epsilon)$$
$$a_{\rho+1} < (\rho + 1)(l + \epsilon)$$

and hence, since $a_{\rho+1} \geq a_\rho$

$$(k - \epsilon)\rho < (\rho + 1)(l + \epsilon)$$

giving

$$(k - l) < \frac{l}{\rho} + 2\epsilon + \frac{\epsilon}{\rho}.$$

Thus $k - l$ can be made as small as we please and is thus zero.

The argument can be very simply adapted to the case in which k is infinite, and if l is not finite k, being not less than l, is infinite. Thus as $n \to \infty$ either $\lim a_n/n$ exists or $a_n/n \to \infty$.[1]

(3) If any moment fails to converge, so will all moments of higher order. It is evident that more than one distribution can exist having a limited number of finite moments given and the remainder infinite. Thus we need only consider the case when moments of all orders exist. Furthermore, if any even moment exists the absolute moment of next lowest order must exist; for if $\int_{-\infty}^{\infty} x^{2n}\,dF$ exists, then each of $\int_{-\infty}^{0} x^{2n}\,dF$ and $\int_{0}^{\infty} x^{2n}\,dF$ exist separately, each being positive. Hence $\int_{-\infty}^{0} x^{2n-1}\,dF$ and $\int_{0}^{\infty} x^{2n-1}\,dF$ exist separately and thus $\int_{-\infty}^{\infty} |x^{2n-1}|\,dF = -\int_{-\infty}^{0} x^{2n-1}\,dF + \int_{0}^{\infty} x^{2n-1}\,dF$ exists. Hence we need only consider the case in which absolute moments of all orders exist.

THEOREM 1. *A set of moments determines a distribution uniquely if the series* $\sum_{r=0}^{\infty} \frac{\nu_r t^r}{r!}$ *converges for some real non-zero* t.

Consider the characteristic function

$$\phi(t) = \int_{-\infty}^{\infty} e^{ixt}\,dF.$$

This is uniformly continuous in t, and so are its derivatives of all orders. Thus we have, in the neighborhood of $t = 0$ the Maclaurin expansion

$$\phi(t) = \sum_{r=0}^{r} \left\{\frac{t^r}{r!}\left[\frac{d^r\phi}{dt^r}\right]_{t=0}\right\} + R$$

$$= \sum_{r=0}^{r} \frac{(it)^r}{r!}\mu_r + R.$$

[1] This proof is necessary to the use of limits in the following theorems, but Theorems 2 and 3 are equally valid if $\overline{\lim}$ is substituted for lim therein. It is not generally true that if a_n and b_n are increasing monotonic sequences either $\lim a_n/b_n$ exists or $a_n/b_n \to \infty$ as $n \to \infty$.

Consequently, under the condition of the theorem, which implies that $\Sigma \frac{(it)^r}{r!} \mu_r$ is absolutely convergent for some radius ρ, $\phi(t)$ has a Taylor expansion in the neighborhood of the origin and is thus uniquely determined by the moments for $t < \rho$. Furthermore, in the neighborhood of $t = t_0$ we have

$$\phi(t) = \sum_{r=0}^{r} \left\{ \frac{i^r (t - t_0)^r}{r!} \int_{-\infty}^{\infty} x^r e^{it_0 x} dF \right\} + R.$$

The modulus of the coefficient of $\frac{(t - t_0)^r}{r!}$ is not greater than ν_r. Therefore $\phi(t)$ can be expanded in the neighborhood of $t = t_0$ in a Taylor series with a radius of convergence at least equal to ρ. Hence the function defining $\phi(t)$ in the neighborhood of the origin can be continued analytically throughout the range $-\infty$ to $+\infty$ and $\phi(t)$ is uniquely determined in that range.

But the characteristic function unqiuely determines the distribution; and hence the theorem follows.

As a result of Theorem 1 we have the following generalization of the criterion given by Lévy.

THEOREM 2. *A set of moments completely determines a distribution if* $\lim_{n\to\infty} \nu_n^{1/n}/n$ *is finite.*

It has already been seen that unless $\nu_n^{1/n}/n$ becomes infinite the limit exists. By the Cauchy test for convergence the series $\Sigma \frac{\nu_r t^r}{r!}$ converges if

$$\lim_{n\to\infty} \left(\frac{\nu_n t^n}{n} \right)^{1/n} < 1. \tag{7}$$

As $n \to \infty$, $(n!)^{1/n}$ tends, in accordance with Stirling's theorem, to $(\sqrt{2\pi n}\, e^{-n} n^n)^{1/n}$ i.e. to n/e. Consequently the condition (7) becomes

$$\lim\, [\nu_n^{1/n}/n]\, et < 1.$$

Thus if $\lim \nu_n^{1/n}/n = k$, say, the inequality (7) is satisfied for $t < 1/(ek)$ and the theorem follows.

An important corollary, which enables us to disregard the absolute moments (which may not be given if part of the range is negative) is

THEOREM 3. *A set of moments uniquely determines a distribution if* $\lim_{n\to\infty} \mu_{2n}^{1/(2n)}/n$ *is finite.*

For
$$\nu_{2n-1}^{1/(2n-1)} \le \nu_{2n}^{1/(2n)} = \mu_{2n}^{1/(2n)}.$$

Thus,
$$\lim \frac{1}{2n-1} \cdot \nu_{2n-1}^{1/(2n-1)} \le \lim \frac{2n}{2n-1} \cdot \frac{1}{2n} \mu_{2n}^{1/(2n)}$$
$$\le \lim \frac{1}{2n} \mu_{2n}^{1/(2n)}$$

and is therefore finite if the limit on the right is finite. Thus lim $\nu_n^{1/n}/n$, which cannot be greater than the greater of the two limits of $\nu_{2n-1}^{1/(2n-1)}/(2n-1)$ and $\nu_{2n}^{1/(2n)}/(2n)$, must be finite; and the theorem follows from Theorem 2.

Now consider the series $\sum_{r=0}^{\infty} \frac{1}{\nu_r^{1/r}}$. Since the successive terms form a monotonic sequence it is a sufficient as well as a necessary condition for convergence that $n/\nu_n^{1/n}$ tend to zero. Thus, if the series is divergent $n/\nu_n^{1/n}$ cannot tend to zero and so $\nu_n^{1/n}/n$ cannot become infinite. Hence it must tend to a finite limit, which may in particular be zero. Hence from Theorem 3 we get

THEOREM 4. *A frequency distribution is uniquely determined by its moments if* $\sum_{r=0}^{\infty} \frac{1}{\nu_r^{1/r}}$ *diverges.*

Since $1/\nu_r^{1/r}$ is a decreasing sequence the series $\Sigma\, 1/\nu_r^{1/r}$ converges or diverges with $\Sigma\, 1/\mu_{2r}^{1/(2r)}$. The Carleman criterion, given by him for the case of limits $\pm\infty$, follows. For the case of limits 0 to ∞ the absolute moments are the same as the moments and the criterion can be the divergence of either $\Sigma\, 1/\mu_r^{1/r}$ or $\Sigma\, 1/\mu_{2r}^{1/(2r)}$. Since μ_r is greater than unity in the type of case under consideration the former series provides a more stringent test than that given by Carleman.

At first sight it is rather surprising that the uniqueness of the distribution depends only on the behavior of the even moments, particularly when, by a simple extension of the above result, it is seen that a sufficient condition for uniqueness is the divergence of $\Sigma\, 1/\mu_{4n}^{1/(4n)}$ or $\Sigma\, 1/\mu_{mn}^{1/(mn)}$ or any infinite subset chosen from the moments. It will, however, be remembered that the odd moments are conditioned to some extent by the even moments, and that uniqueness is really determined by the limiting form of ν_n as n tends to infinity.

It is evident that other tests may be derived from Theorem 1 by using the various tests for the convergence of an infinite series. For instance it is a sufficient condition for a set of moments to determine uniquely a distribution with positive range that

$$\frac{\mu_n}{n!}\bigg/\frac{\mu_{n+1}}{(n+1)!} = 1 + \frac{\alpha}{n} + O\left(\frac{1}{n^{1+\beta}}\right), \quad \text{where } \begin{matrix}\alpha > 1\\ \beta > 0\end{matrix}$$

i.e. that

$$\frac{\mu_n}{\mu_{n+1}} = 1 + \frac{\gamma}{n} + O\left(\frac{1}{n^{1+\beta}}\right), \qquad \gamma > 0. \tag{8}$$

It may be noted in passing that the distribution

$$dF = e^{-x}\,dx \qquad 0 \leq x \leq \infty,$$

for which

$$\mu_r \text{ (about origin)} = r!$$

is completely determined by its moments. In fact, by direct reference to Theorem 1 we see that the series $\Sigma\,(it)^r$ converges for $t < 1$.

A frequency distribution of finite range is uniquely determined by its moments. For if the range is 0 to A we have

$$\mu_r = \int_0^A x^r \, dF \leq A^r$$

and hence $1/\mu_r^{1/r} \geq 1/A$ so that the series $\Sigma \, 1/\mu_r^{1/r}$ is divergent.

A proof for the case when the frequency distribution is continuous has been given by Lévy, though on entirely different lines from the above.

THEOREM 5. *A frequency distribution of infinite range is uniquely determined by its moments if it tends to zero at the infinite terminals faster than* e^{-x}.

Consider first of all the case when only one end of the range is infinite, so that we may take the range to be 0 to ∞.

If $(\mu_n/n!)^{1/n}$ has a finite limit the distribution is unique, by Theorem 2. We have then only to consider the cases (if any) in which $(\mu_n/n!)^{1/n}$ tends to infinity. It will be shown that in fact such cases do not occur.

Given any (small) ϵ there exists an X such that

$$\frac{f(x)}{e^{-x}} < \epsilon, \qquad x > X$$

where $f(x)$ is the distribution. Thus

$$\int_X^\infty f(x)x^n \, dx < \epsilon \int_X^\infty e^{-x}x^n \, dx < \epsilon n!. \tag{9}$$

This is true for all n and X is independent of n. Now,

$$\int_0^\infty f(x)x^n \, dx = \int_0^X f(x)x^n \, dx + \int_X^\infty f(x)x^n \, dx.$$

The first integral on the right is not greater than X^n. The integral on the left tends, for large n, to something of greater order than $n!$, by our hypothesis, and hence to something of greater order than n^n. This is of greater order than X^n (since X, however large, is independent of n) and consequently the second integral on the right is also of greater order than $n!$. But this is contrary to equation (9).

The case for the range which is infinite in both directions may be dealt with similarly.

It is easily seen that the two examples of equations (1) and (2) do not tend to infinity faster than e^{-x}.

Except for the general result of Stieltjes, all the above criteria provide sufficient conditions, but whether the condition of Theorem 1 is also necessary is not certain. An inquiry into the circumstances in which the moment-series of Theorem 1 does not converge throws some light on the question.

It will be remembered that the characteristic function always exists and is uniformly continuous in t. Since the moments of all orders are assumed to exist we always have

$$\left[\frac{d^r}{dt^r}\phi(t)\right]_{t=0} = (i)^r \mu_r.$$

Thus, if $\phi(t)$ can be expanded in an infinite Taylor series that series must be $\Sigma \frac{(it)^r}{r!} \mu_r$. And if this series does not converge then $\phi(t)$ cannot be expanded as an infinite Taylor series. But it can always be expanded in the finite form with remainder

$$\phi(t) = \sum_{r=0}^{r} \frac{(it)^r}{r!} \mu_r + R.$$

Thus, when the series does not converge, $\phi(t)$ can be expanded in powers of t only asymptotically.

Now it is known that there exist an infinite number of functions which have a given set of coefficients in an asymptotic expansion; for instance, if $\psi(t)$ has an asymptotic expansion in t the functions $\psi(t) + \lambda t^{-\log t}$ all have the same expansion. It is therefore hardly surprising that when the conditions of Theorem 1 break down there can be more than one frequency distribution with the same set of moments.

But it does not follow from what has been said that there *must* be more than one frequency distribution. There must be more than one function, but those functions may not qualify as frequency distributions, e.g. they may be negative in part of the range. In the example just given $t^{-\log t}$ cannot be a characteristic function, for it does not obey the well-known condition that $\phi(t)$ and $\phi(-t)$ should be conjugate.

However, the question is more of mathematical than of statistical interest since the criteria provided above are likely to be adequate for the distributions encountered in practice. For example they establish the uniqueness of the Pearson curves (including the normal curve), the Poisson and the binomial. It would seem that distributions like those of equations (1) and (2) will appear only as statistical curiosities.

REFERENCES

[1] J. Stieltjes, *Recherches sur les fractions continues*, Oeuvres, Groningen, 1918.
[2] H. Hamburger, "Über eine Erweiterung des Stietjesschen Momentproblems," *Math. Annalen*, Vol. 81 (1920), p. 235, and Vol. 82 (1921), pp. 120 and 168.
[3] P. Lévy, *Calcul des Probabilités*, 1925, Paris.
[4] T. Carleman, *Les Fonctions Quasi-analytiques*, 1925, Paris.
[5] G. H. Hardy, J. E. Littlewood and G. Pólya, *Inequalities*, 1934, Cambridge, England.

London, England.