The Annals of Statistics
1973, Vol. 1, No. 1, 114–126

CHARACTERIZATIONS OF POPULATIONS USING REGRESSION PROPERTIES[1]

BY F. S. GORDON[2]

C. W. Post College

In the present work, we study characterizations of populations obtained by using regression properties of one statistic on another. We extend some of the results of Lukacs and Laha by considering the cubic regression (polynomial regression of order 3) of a cubic statistic S on a linear one L. This assumption of cubic regression is used to derive a third order non-linear differential equation in the characteristic function $h(t)$ of a set of n independently and identically distributed random variables. The coefficients in this differential equation represent certain fixed relationships between the coefficients of the statistic S and the regression coefficients. For appropriate choices of the coefficients in the fundamental differential equation, the resulting equation can be solved to yield the characteristic function of a particular distribution. In this way, we are able to obtain a series of characterization theorems for each of a variety of populations including Normal, Gamma, Binomial, Poisson, Geometric and several others. Moreover, all of the results obtained by Lukacs and Laha for characterizing populations using quadratic and constant regression are shown to be special cases of the theorems obtained in the present work.

Finally, in the last section, we present an outline of a technique which can be used to study any general $r \leqq m$th order polynomial regression of any mth order statistic on a linear one. The approach used to generate the mth order differential equation is indicated and a method for determining the appropriate conditions on the coefficients is discussed.

1. Introduction. In the present work, we study characterizations of populations by using regression properties of one statistic on another. In particular, a series of results is obtained by considering the cubic regression (polynomial regression of order 3) of a cubic statistic S on a linear one L.

This assumption of cubic regression is used to derive a third order non-linear differential equation in the characteristic function $h(t)$ of a set of n independently and identically distributed random variables. The coefficients in this differential equation represent certain fixed relationships between the coefficients of the statistic S and the regression coefficients. For appropriate choices of the coefficients in the fundamental differential equation, the resulting equation can be solved to yield the characteristic function of a particular distribution. In this way, we are able to characterize a variety of populations including Normal, Gamma, Binomial, Poisson, Geometric, and several others. Moreover, all of the results

Received September 20, 1971; revised May 5, 1972.

[1] This paper is adapted from a portion of the author's doctoral dissertation written under the supervision of Professor A. M. Mathai and submitted to the Faculty of Graduate Studies and Research, McGill University, Montreal, Canada.

[2] The author wishes to acknowledge support of the National Research Council of Canada.

obtained by Lukacs and Laha for characterizing populations using quadratic and constant regression are shown to be special cases of the theorems obtained in the present work.

Finally, in the last section, we present an outline of a technique which can be used to study any general $r \leqq m$th order polynomial regression of any mth degree statistic on a linear one. A method for determining the appropriate conditions on the coefficients of the resulting mth order differential equation is also discussed.

2. Derivation of the fundamental differential equation. Let X and Y be two stochastic variables. For convenience, we denote the conditional expectation of X for a given value of $Y = y$, $E(X/Y = y)$, by $E(X/Y)$. We assume that these stochastic variables, as well as all others to be considered throughout the following discussions, are non-degenerate. If the conditional moment $E(X/Y)$ exists, we can define polynomial regression of X on Y in the following manner.

DEFINITION. X has polynomial regression of order m on Y if

$$E(X/Y) = \beta_0 + \beta_1 Y + \beta_2 Y^2 + \cdots + \beta_m Y^m, \quad \text{a.e.}$$

where the β_j $(j = 0, 1, \cdots, m)$ are real constants, known as the regression coefficients.

In particular, we consider a sample $X_1, \cdots, X_n$ of independently and identically distributed stochastic variables from a population with distribution function $F(x)$ and assume that all moments up to third order exist. We define

$$L = X_1 + \cdots + X_n,$$
$$S = \sum_{j,k,m} a_{jkm} X_j X_k X_m + \sum_{j,k} b_{jk} X_j X_k + \sum_j c_j X_j,$$

where a_{jkm}, b_{jk} and c_j (for all j, k, $m = 1, \cdots, n$) are real constants. We assume that S has cubic regression on L and study some problems where a suitable choice of the statistic S, given in terms of some relations between the coefficients a_{jkm}, b_{jk} and c_j of S and the regression coefficients β_0, β_1, β_2 and β_3, determines the population uniquely. These characterizations arise as solutions of a fundamental differential equation in the characteristic function of the population.

The assumption of cubic regression for the statistics S on L leads to the relation

$$E(S/L) = \beta_0 + \beta_1 L + \beta_2 L^2 + \beta_3 L^3 \quad \text{a.e.}$$

By a result of Laha and Lukacs (1960), a necessary and sufficient condition for this to hold true is

$$E(Se^{itL}) = \beta_0 E(e^{itL}) + \beta_1 E(Le^{itL}) + \beta_2 E(L^2 e^{itL}) + \beta_3 E(L^3 e^{itL}). \tag{1}$$

If we substitute the assumed forms for S and L, the left-hand side is

$$E(Se^{itL}) = \sum_{j,k,m} a_{jkm} E(X_j X_k X_m e^{itL}) + \sum_{j,k} b_{jk} E(X_j X_k e^{itL}) + \sum_j c_j E(X_j e^{itL})$$
$$= E_1 + E_2 + E_3.$$

Since $X_1, \cdots, X_n$ are independently and identically distributed stochastic variables, we introduce their characteristic function

$$h(t) = E(e^{itX})$$

and its derivatives and thus obtain

$$\begin{aligned} E_1 &= \sum_j a_{jjj} ih'''h^{n-1} + \sum_{j\neq k}(a_{jjk} + a_{jkj} + a_{kjj})(-h'')(-ih')h^{n-2} \\ &\quad + \sum_{j\neq k\neq m} a_{jkm}(-ih')^3 h^{n-3} \\ &= iA_1 h'''h^{n-1} + iA_2 h''h'h^{n-2} + iA_3(h')^3 h^{n-3}, \end{aligned}$$

where

$$\begin{aligned} A_1 &= \sum_j a_{jjj} \\ A_2 &= \sum_{j\neq k} (a_{jjk} + a_{jkj} + a_{kjj}) \\ A_3 &= \sum_{j\neq k\neq m} a_{jkm} . \end{aligned}$$

In a similar manner, we find

$$\begin{aligned} E_2 &= -B_1 h''h^{n-1} - B_2(h')^2 h^{n-2} \\ E_3 &= -iCh'h^{n-1}, \end{aligned}$$

where

$$\begin{aligned} B_1 &= \sum_j b_{jj} \\ B_2 &= \sum_{j\neq k} b_{jk} \\ C &= \sum_j c_j . \end{aligned}$$

Proceeding in the same way, we evaluate the terms on the right-hand side of equation (1) to obtain

$$\begin{aligned} \beta_0 E(e^{itL}) &= \beta_0 h^n, \\ \beta_1 E(Le^{itL}) &= -in\beta_1 h'h^{n-1}, \\ \beta_2 E(L^2 e^{itL}) &= -n\beta_2 h''h^{n-1} - n(n-1)\beta_2(h')^2 h^{n-2}, \\ \beta_3 E(L^3 e^{itL}) &= in\beta_3 h'''h^{n-1} + 3in(n-1)\beta_3 h''h'h^{n-2} \\ &\quad + in(n-1)(n-2)\beta_3(h')^3 h^{n-3} . \end{aligned}$$

We substitute the above results into (1) and simultaneously divide each expression by h^n, for convenience. This is possible because the characteristic function $h(t)$ is nonzero in some neighborhood of the origin, by an application of the Intermediate Value Theorem. As a consequence, we obtain

$$\begin{aligned} i(n\beta_3 - A_1)(h'''/h) &+ i[3n(n-1)\beta_3 - A_2](h''/h)(h'/h) \\ &+ i[n(n-1)(n-2)\beta_3 - A_3](h'/h)^3 + (B_1 - n\beta_2)(h''/h) \\ &+ [B_2 - n(n-1)\beta_2](h'/h)^2 + i(C - n\beta_1)(h'/h) + \beta_0 = 0 . \end{aligned}$$

If we introduce the substitution

$$g(t) = \log_e h(t)$$

and then set

$$f(t) = g'(t),$$

the above differential equation reduces to

$$\begin{aligned} i(n\beta_3 - A_1)f'' &+ i[3n^2\beta_3 - 3A_1 - A_2]f'f \\ &+ i[n^3\beta_3 - A_1 - A_2 - A_3]f^3 + (B_1 - n\beta_2)f' \\ &+ (B_1 + B_2 - n^2\beta_2)f^2 + i(C - n\beta_1)f + \beta_0 = 0\,. \end{aligned}$$

Alternatively, for convenience in the sequel, we will write this as

$$id_1 f'' + id_2 f'f + id_3 f^3 + d_4 f' + d_5 f^2 + id_6 f + d_7 = 0\,, \tag{2}$$

where we have put

$$\begin{aligned} d_1 &= n\beta_3 - A_1 \\ d_2 &= 3n^2\beta_3 - 3A_1 - A_2 \\ d_3 &= n^3\beta_3 - A_1 - A_2 - A_3 \\ d_4 &= B_1 - n\beta_2 \\ d_5 &= B_1 + B_2 - n^2\beta_2 \\ d_6 &= C - n\beta_1 \\ d_7 &= \beta_0\,. \end{aligned} \tag{3}$$

In the following section, we solve this fundamental differential equation (2) subject to particular sets of relations between the coefficients d_j and so characterize a number of different populations.

For future reference, we evaluate the initial conditions on the functions $h(t)$, $g(t)$ and $f(t)$. To begin,

$$\begin{aligned} h(0) &= E(e^{itX})|_{t=0} = 1 \\ h'(0) &= iE(Xe^{itX})|_{t=0} = i\mu \\ h''(0) &= -E(X^2e^{itX})|_{t=0} = -\sigma^2 - \mu^2\,, \end{aligned}$$

where μ and σ^2 are the mean and the variance of the population, respectively. Consequently,

$$\begin{aligned} g(0) = 0\,, \quad g'(0) = i\mu\,, \quad g''(0) = -\sigma^2 \\ f(0) = i\mu\,, \quad f'(0) = -\sigma^2\,. \end{aligned} \tag{4}$$

3. Characterizations of populations. In this section, we develop a series of characterizations for many well-known populations, beginning with the normal.

THEOREM 1. *Given a simple random sample $X_1, \cdots, X_n$ from a population with finite third moment, suppose*

$$\begin{aligned} d_3 &= d_5 = 0 \\ d_7 &= \sigma^2 d_4 \\ d_6 &= \sigma^2 d_2\,. \end{aligned}$$

Then, if any one of the following sets of conditions hold,

$$
\begin{array}{llll}
\text{(a)} & d_7 = 0\,, & d_6 = 0\,, & d_1 \neq 0\,, \\
\text{(b)} & d_7 = 0\,, & d_6 \neq 0\,, & d_1 = 0\,, \\
\text{(c)} & d_7 \neq 0\,, & d_6 = 0\,, & d_1 = 0\,, \\
\text{(d)} & d_7 = 0\,, & d_6 \neq 0\,, & d_1 \neq 0\,, \\
\text{(e)} & d_7 \neq 0\,, & d_6 = 0\,, & d_1 \neq 0\,, \\
\text{(f)} & d_7 \neq 0\,, & d_6 \neq 0\,, & d_1 = 0\,,
\end{array}
$$

a necessary and sufficient condition for S to have cubic regression on L is that the population be normal.

PROOF. The proofs of cases (a), (b), (c) and (e) follow readily as solutions of relatively simple differential equations and hence will be omitted.

(d) If we impose these conditions, the fundamental differential equation (2) reduces to

$$f'' + af'f + \sigma^2 af = 0\,,$$

where

$$a = d_2/d_1\,.$$

We can reduce this equation to a first order non-linear differential equation by letting $y = f' = df/dt$. Hence, we obtain

$$y\frac{dy}{df} = (-\sigma^2 af) + (-af)y\,,$$

an Abel differential equation of the second kind, as given by Murphy (1960). Let

$$y = u(f) + \int (-af)\,df = u(f) - \tfrac{1}{2}af^2\,,$$

so that the differential equation becomes

$$u\,du = \tfrac{1}{2}af^2\,du - \sigma^2 af\,df\,.$$

An integrating factor for this equation is $\exp(-u/\sigma^2)$, which yields

$$\tfrac{1}{2}af^2 = u + \sigma^2 + c_1 \exp(u/\sigma^2)\,.$$

We now apply the initial conditions (4) to determine $u(f) = y(f) + \frac{1}{2}af^2$ at $t = 0$. This gives

$$u(i\mu) = f'(0) + \tfrac{1}{2}a(i\mu)^2 = -\sigma^2 - \tfrac{1}{2}a\mu^2\,.$$

Evaluating the integrated expression above at $t = 0$, we find

$$-\tfrac{1}{2}a\mu^2 = (-\sigma^2 - \tfrac{1}{2}a\mu^2) + \sigma^2 + c_1 \exp(u(i\mu)/\sigma^2)\,,$$

which implies that $c_1 = 0$ and hence

$$u = \tfrac{1}{2}af^2 - \sigma^2\,.$$

As a result,

$$y = \frac{df}{dt} = \tfrac{1}{2}af^2 - \sigma^2 + (-\tfrac{1}{2}af^2) = -\sigma^2 .$$

Therefore, if we integrate with respect to t and apply the initial conditions (4), we obtain

$$f = -\sigma^2 t + i\mu ,$$

which therefore leads to

$$h(t) = \exp(i\mu t - \tfrac{1}{2}\sigma^2 t^2) ,$$

the characteristic function of the normal population.

(f) In this case, the fundamental differential equation is

$$(f' + \sigma^2)(id_2 f + d_4) = 0 .$$

Consequently, either

$$f' + \sigma^2 \equiv 0 \qquad \text{or} \qquad id_2 f + d_4 \equiv 0 .$$

The latter possibility implies

$$f = id_4/d_2 = i\mu ,$$

using the initial conditions (4). This leads to

$$h(t) = \exp(i\mu t) ,$$

the characteristic function of a degenerate distribution. However, since the stochastic variables studied are assumed to be non-degenerate, this case is impossible. Hence,

$$f' = -\sigma^2 ,$$

which easily leads to the characteristic function of the normal distribution.

Moreover, in each of the six cases (a)—(f), the converse is obtained by direct substitution.

If, in the statistic S originally considered in Section 2, we put $a_{jkm} = 0$ (for all $j, k, m = 1, \cdots, n$) then the cubic statistic S reduces to a quadratic statistic

$$Q = \textstyle\sum_{j,k} b_{jk} X_j X_k + \sum_j c_j X_j .$$

Furthermore, putting $\beta_3 = 0$, the regression of Q on L assumes the quadratic form

$$E(Q/L) = \beta_0 + \beta_1 L + \beta_2 L^2 .$$

Moreover, the fundamental differential equation (2) becomes

$$(B_1 - n\beta_2)f' + (B_1 + B_2 - n^2\beta_2)f^2 + i(C - n\beta_1)f + \beta_0 = 0 .$$

Then, Theorem 1 given by Laha and Lukacs (1960) in their study of quadratic regression is a special case of part (c) of the above theorem.

In addition, if we let $\beta_1 = \beta_2 = 0$, we find that Theorem 6.2.1 of Lukacs and Laha (1964) is also a special case of the above.

Obviously, it is also possible to obtain additional results by considering any statistic S of degree $s \leqq 3$ which has polynomial regression on L of order $m \leqq s$.

We next consider the Gamma distribution and obtain characterizations for it in terms of cubic regression for the statistic S on L.

THEOREM 2. *Given a simple random sample $X_1, \cdots, X_n$ from a population with finite third moment and nonzero mean, suppose*

$$d_6 = d_7 = 0$$
$$d_4 = -(\mu^2/\sigma^2)d_5$$
$$2\sigma^4 d_1 + \sigma^2\mu^2 d_2 + \mu^4 d_3 = 0 .$$

Then if any one of the following sets of conditions holds,

(a) $d_1 = 0$, $d_2 = 0$, $d_3 = 0$, $d_4 \neq 0$,
(b) $d_1 \neq 0$, $d_2 = 0$, $d_3 \neq 0$, $d_4 = 0$,
(c) $d_1 = 0$, $d_2 \neq 0$, $d_3 \neq 0$, $d_4 = 0$,
(d) $d_1 \neq 0$, $d_2 \neq 0$, $d_3 = 0$, $d_4 = 0$,
(e) $d_1 = -(\mu^2/\sigma^2)d_2 = -(\mu^4/\sigma^4)d_3 \neq 0$, $d_4 = 0$,

a necessary and sufficient condition for S to have cubic regression on L is that the population be Gamma.

PROOF. (a) The fundamental differential equation now becomes

$$f' = (\sigma^2/\mu^2)f^2 . \tag{5}$$

Since f is nonzero (otherwise, $h(t)$ would be a constant and hence not a characteristic function),

$$-1/f = (\sigma^2/\mu^2)t + c_1 .$$

From the initial conditions (4) and a further integration, we find

$$g(t) = (\mu^2/\sigma^2)\log(i\mu/[\sigma^2 t + i\mu])$$

and accordingly,

$$h(t) = (1 - i\theta t)^{-\lambda} ,$$

where

$$\theta = \sigma^2/\mu , \qquad \lambda = \mu^2/\sigma^2 .$$

This is the characteristic function of the Gamma distribution.

(b) The fundamental differential equation becomes

$$f'' = 2(\sigma^4/\mu^4)f^3 .$$

Multiplying both sides of this equation by f' and integrating, we apply the initial conditions (4) to find

$$(f')^2 = (\sigma^4/\mu^4)f^4 .$$

Taking the positive square root, we obtain

$$f' = (\sigma^2/\mu^2)f^2 ,$$

which is the same differential equation solved in part (a), equation (5).

It is worth noting that the positive square root is essential here, since the alternate equation $f' = -(\sigma^2/\mu^2)f^2$ does not satisfy the initial conditions.

(e) The fundamental differential equation becomes

$$f'' = (\sigma^2/\mu^2)f'f + (\sigma^4/\mu^4)f^3 .$$

As in Theorem 1d, we let $y = f'$ and so obtain

$$y\frac{dy}{df} = (\sigma^4/\mu^4)f^3 + (\sigma^2/\mu^2)f \cdot y .$$

An integrating factor for this differential equation is

$$M(f, y) = y - (\sigma^2/\mu^2)f^2 .$$

Integrating the resulting equation and using the initial conditions (4), we find

$$y^3/3 - (\sigma^2/\mu^2)f^2y^2/2 + (\sigma^6/\mu^6)f^6/6 = [y - (\sigma^2/\mu^2)f^2]^2[y + (\sigma^2/\mu^2)f^2/2] = 0 .$$

Since the second factor cannot be zero (otherwise the initial conditions are not satisfied), we are once more led to equation (5) whose solution is the characteristic function of the Gamma population.

The proofs of cases (c) and (d) follow easily, as do the converses in each of the five cases.

As in the discussion following Theorem 1, it is possible to obtain a series of corollaries by considering any statistic S of degree $s \leqq 3$ which has polynomial regression on L of order $m \leqq s$. In particular, Theorem 4 of Laha and Lukacs (1960) and Theorem 6.2.2 of Lukacs and Laha (1964) are special cases of part (a).

Some characterizations for the Binomial and Negative Binomial distributions are also possible using the property of cubic regression for S on L.

THEOREM 3. *Given a simple random sample $X_1, \cdots, X_n$ from a population with finite third moment, suppose*

$$d_7 = 0 .$$

Then, if any one of the following sets of conditions holds for some $p \in (0, 1)$,

(a) $d_1 = d_2 = d_3 = 0$, $\quad d_4 \doteq (\mu/p)d_5 = -d_6 \neq 0$,

(b) $d_1 = d_4 = d_6 = 0$, $\quad d_2 = (\mu/p)d_3 = d_5 \neq 0$,

(c) $d_3 = d_5 = d_6 = 0$, $\quad d_1 = \frac{1}{2}\mu^2 d_2/(\mu - \sigma^2) = d_4 \neq 0$,

a necessary and sufficient condition for S to have cubic regression on L is that the population be Binomial.

PROOF. In each case, the proof follows readily from the solution of a relatively simple differential equation. Moreover, the converses follow easily.

As a special case of the above, we have Laha and Lukacs' (1960) Theorem 3.

THEOREM 4. *Given a simple random sample $X_1, \cdots, X_n$ from a population with* *finite third moment, suppose*

$$d_7 = 0\,.$$

Then, if either one of the following sets of conditions holds, for some $p \in (0, 1)$,

$$\text{(a)}\quad d_1 = d_2 = d_3 = 0\,, \qquad d_4 = -p\mu d_5/(1-p) = -d_6 \neq 0\,,$$

$$\text{(b)}\quad d_1 = d_4 = d_6 = 0\,, \qquad d_2 = -p\mu d_3/(1-p) = d_5 \neq 0\,,$$

a necessary and sufficient condition for S to have cubic regression on L is that the population be Negative Binomial.

PROOF. The proofs of both cases, as well as that of the converses, are immediate and hence are omitted.

As before, the above result contains Laha and Lukacs' (1960) Theorem 3, as a special case.

Several characterizations for the Geometric Distribution are easily obtained as special cases of those of the Negative Binomial Distribution.

THEOREM 5. *Given a sample random sample* $X_1, \cdots, X_n$ *from a population with finite third moment and nonzero mean, suppose*

$$d_7 = 0\,.$$

Then, if either one of the following sets of conditions holds,

$$\text{(a)}\quad d_1 = d_2 = d_3 = 0\,, \qquad d_4 = -d_5 = -d_6 \neq 0\,,$$

$$\text{(b)}\quad d_1 = d_4 = d_6 = 0\,, \qquad d_2 = -d_3 = d_5 \neq 0\,,$$

a necessary and sufficient condition for S to have cubic regression on L is that the population be Geometric.

Our next result will involve a characterization for the Poisson-type distribution whose characteristic function is given by

$$h(t) = \exp[\lambda(e^{i\rho t} - 1) + i\nu t]\,, \tag{6}$$

where $\lambda > 0$, $\rho \neq 0$, and ν are three real constants. The mean and variance for this distribution are given by

$$\mu = \rho\lambda + \nu\,, \qquad \sigma^2 = \rho^2\lambda\,.$$

In particular, if $\nu = 0$, we have $\lambda = \mu/\rho = \sigma^2/\rho^2$ and the above characteristic function reduces to

$$h(t) = \exp[(\mu/\rho)(e^{i\rho t} - 1)]\,. \tag{7}$$

In addition, in the special case where $\rho = 1$, we have $\lambda = \mu = \sigma^2$ and the characteristic function reduces to the characteristic function of the usual Poisson distribution.

THEOREM 6. *Given a simple random sample* $X_1, \cdots, X_n$ *from a population with finite third moment, suppose*

$$d_2 = d_3 = d_5 = d_6 = d_7 = 0\,.$$

Then, if the following condition holds for some nonzero real constant ρ,

$$d_4 = \rho d_1 \neq 0 ,$$

a necessary and sufficient condition for S to have cubic regression on L is that the population be Poisson-type with characteristic function given by equation (6).

PROOF. Under the hypotheses, the fundamental differential equation (2) becomes

$$f'' - i\rho f' = 0 .$$

If we integrate this and apply the initial conditions, we find

$$f' - i\rho f = \rho\mu - \sigma^2 ,$$

which in turn may be solved to yield

$$f = i(\sigma^2/\rho)e^{i\rho t} + i(\mu - \sigma^2/\rho) .$$

Consequently, we are led to

$$h = \exp[(\sigma^2/\rho^2)(e^{i\rho t} - 1) + i(\mu - \sigma^2/\rho)t] .$$

Thus, for $\lambda = \sigma^2/\rho^2$ and $\nu = \mu - \sigma^2/\rho = \mu - \rho\lambda$, we have the characteristic function of the Poisson-type distribution given in (6).

As usual, the converse follows immediately.

The following theorem involves characterizations of the Poisson-type distribution whose characteristic function is given in (7).

THEOREM 7. *Given a simple random sample* $X_1, \cdots, X_n$ *from a population with finite third moment and nonzero mean, suppose*

$$d_3 = d_7 = 0 .$$

Then, if any one of the following sets of conditions holds for some nonzero constant ρ,

(a) $d_1 = d_2 = d_5 = 0 , \qquad d_6 = -\rho d_4 \neq 0 ,$

(b) $d_1 = d_4 = d_6 = 0 , \qquad d_5 = \rho d_2 \neq 0 ,$

(c) $d_2 = d_5 = 0 , \qquad (\sigma^4/\mu^2)d_1 = (\sigma^2/\mu)d_4 + d_6 \neq 0 ,$

a necessary and sufficient condition for S to have cubic regression on L is that the population be Poisson-type with characteristic function given by equation (7).

PROOF. The proofs of cases (a) and (b) follow readily from relatively simple differential equations whose solutions lead to the Poisson-type characteristic function in (7).

(c) The fundamental differential equation (2) may now be written as

$$f'' - i(d_4/d_1)f' + [(\sigma^4/\mu^2) - (\sigma^2/\mu)(d_4/d_1)]f = 0 .$$

If we put $a = d_4/d_1$, the solution to this equation can readily be found to be

$$f = Ae^{i\sigma^2 t/\mu} + Be^{i(a-\sigma^2/\mu)t} .$$

From the initial conditions on f and f' we determine

$$A = i\mu , \qquad B = 0 .$$

Consequently, we are led to

$$h = \exp[(\mu/\rho)(e^{i\rho t} - 1)] ,$$

the characteristic function of the Poisson-type distribution given in (7) with $\rho = \sigma^2/\mu$.

In each case, the converse follows immediately.

As in the previous theorems, the above result contains, as special cases, Laha and Lukacs' (1960) Theorem 2 and Lukacs and Laha's (1964) Theorem 6.2.4.

Our final result involves characterizations for a distribution whose characteristic function has the form

$$h(t) = e^{imt}[\cosh at + ib \sinh at]^{-r} , \tag{8}$$

where a, b, m and r are real constants and $r > 0$, $a \neq 0$. This distribution was originally studied by Laha and Lukacs (1960).

THEOREM 8. *Given a simple random sample $X_1, \cdots, X_n$ from a population with finite third moment, suppose that any one of the following sets of conditions holds,*

$$\begin{aligned}
&\text{(a)}\quad d_1 = d_2 = d_3 = 0 , \quad d_4 \neq 0 , \quad d_5 \neq 0 , \quad d_7 \neq 0 , \\
&\qquad d_4 \cdot d_5 < 0 , \quad d_6^2 + 4d_5 \cdot d_7 < 0 , \\
&\text{(b)}\quad d_1 = d_4 = d_7 = 0 , \quad d_2 \neq 0 , \quad d_3 \neq 0 , \quad d_5 \neq 0 , \\
&\qquad d_2 \cdot d_3 < 0 , \quad d_5^2 + 4d_3 \cdot d_6 < 0 , \\
&\text{(c)}\quad d_3 = d_5 = d_6 = d_7 = 0 , \quad d_1 \neq 0 , \quad d_2 \neq 0 , \\
&\qquad \sigma^2 d_1 + \tfrac{1}{2}\mu^2 d_2 - \mu d_4 \neq 0 , \quad d_1 \cdot d_2 < 0 , \\
&\qquad d_4^2 + 4d_2(\sigma^2 d_1 + \tfrac{1}{2}\mu^2 d_2 - \mu d_4) < 0 .
\end{aligned}$$

Then a necessary and sufficient condition for S to have cubic regression on L is that the population have the characteristic function of the form in equation (8).

PROOF. (a) Under the hypotheses, the fundamental differential equation (2) now becomes

$$d_4 f' + d_5 f^2 + i d_6 f + d_7 = 0 .$$

That is,

$$f' = (-d_5/d_4)[(f + \tfrac{1}{2} i d_6/d_5)^2 - K^2] ,$$

where

$$K^2 = -(d_6^2 + 4d_5 d_7)/4d_5^2 \geqq 0 .$$

We introduce the substitution $f = iy$ so that the above differential equation may be rewritten as

$$\frac{dy}{(y + \frac{1}{2}d_6/d_5)^2 + K^2} = -i(d_5/d_4)\, dt .$$

We now let

$$y + \tfrac{1}{2}d_6/d_5 = K \tan\theta\,,$$

so that the above equation becomes

$$d\theta/K = -(id_5/d_4)\,dt\,.$$

Integrating this equation with respect to t, we obtain

$$(1/K)\tan^{-1}[(1/K)(y + \tfrac{1}{2}d_6d_5)] = -(id_5/d_4)t + c_1\,, \tag{9}$$

where, from the intial conditions,

$$c_1 = (1/K)\tan^{-1}[(1/K)(\mu + \tfrac{1}{2}d_6/d_5)]\,. \tag{10}$$

We now write (9) as

$$\begin{aligned}(1/K)(y + \tfrac{1}{2}d_6/d_5) &= \tan K[(-id_5/d_4)t + c_1]\\ &= \frac{\tan(-iKd_5t/d_4) + \tan(Kc_1)}{1 - \tan(-iKd_5t/d_4)\cdot\tan(Kc_1)}\,.\end{aligned}$$

However, if we introduce the value of c_1 from (10) and use the fact that $\tan(iz) = i\tanh(z)$, we find

$$(1/K)(y + \tfrac{1}{2}d_6/d_5) = \frac{-i\tanh(Kd_5t/d_4) + (1/K)(\mu + \tfrac{1}{2}d_6/d_5)}{1 + i(1/K)(\mu + \tfrac{1}{2}d_6/d_5)\tanh(Kd_5t/d_4)}\,.$$

For simplicity, we introduce

$$b = (\mu + \tfrac{1}{2}d_6/d_5)/K$$

and use the relation $f = iy$ to write

$$f = -\tfrac{1}{2}id_6/d_5 + K\frac{\sinh(Kd_5t/d_4) + ib\cosh(Kd_5t/d_4)}{\cosh(Kd_5t/d_4) + ib\sinh(Kd_5t/d_4)}\,.$$

This latter equation may easily be integrated with respect to t and hence leads to

$$h = \exp[-\tfrac{1}{2}i(d_6/d_5)t]\cdot[\cosh(Kd_5t/d_4) + ib\sinh(Kd_5t/d_4)]^{(d_4/d_5)}\,,$$

which is a characteristic function of the form given in (8).

(b) and (c). The proofs of these two parts follow readily from the above proof.

In each of the three cases, the converse is immediate.

Moreover, as before, the above theorem contains as a special case Laha and Lukacs' (1960) Theorem 5.

4. General polynomial regression. We note in conclusion that the method utilized throughout the present work can be generalized to a study of polynomial regression of order m for an mth order statistic on a linear one. This assumption leads to an mth order non-linear differential equation in the characteristic function, $h(t)$. The solutions of this differential equation subject to various conditions on its coefficients yield a series of characterizations for the different distributions. We now outline the method for obtaining these conditions. Given the characteristic function of the particular distribution of interest, we substitute it into the fundamental differential equation. For some distributions, the resulting expression will be a polynomial in t (for example, for the Normal and the Gamma distributions); for some others, (such as the Binomial, Negative Binomial, Poisson and Geometric distributions), the expression will be a polynomial in e^{iat}; and so forth. Since t, and accordingly e^{iat}, is arbitrary, these polynomials are identically zero if, and only if, all of their coefficients are zero. It is precisely these zero conditions which yield the required relationships between the coefficients of the fundamental differential equation.

Finally, as consequences of the results thus obtained, it is possible, as in the cubic case, to derive characterizations based on polynomial regression of any order $r \leqq m$.

5. Acknowledgment. The author wishes to express her sincere gratitude to Professor A. M. Mathai for many helpful discussions and suggestions during the preparation of the doctoral dissertation upon which this work is based in part.

REFERENCES

LAHA, R. G. (1954). On a characterization of the Gamma distribution. *Ann. Math. Statist.* **25** 784–787.

LAHA, R. G. and LUKACS, E. (1960). On a problem connected with quadratic regression. *Biometrika* **47** 335–343.

LUKACS, E. (1960). *Characteristic Functions.* Griffin, London.

LUKACS, E. and LAHA, R. G. (1964). *Applications of Characteristic Functions.* Griffin, London.

MURPHY, G. M. (1960). *Ordinary Differential Equations and Their Solutions.* Van Nostrand, Princeton.

DEPARTMENT OF MATHEMATICS
C. W. POST COLLEGE
LONG ISLAND UNIVERSITY
GREENVALE, NEW YORK 11548