The Annals of Mathematical Statistics
1970, Vol. 41, No. 4, 1363–1368

GENERAL FORMULAE FOR THE CENTRAL MOMENTS OF CERTAIN SERIAL CORRELATION COEFFICIENT APPROXIMATIONS

BY ADRIENNE W. KEMP

The Queen's University of Belfast

1. Introduction and summary. The distributions considered in this paper have the probability density functions (pdf's)

$$\frac{(1-x^2)^{\frac{1}{2}(n-1)}(1+r^2-2rx)^{-\frac{1}{2}n}}{B[\frac{1}{2}(n+1),\frac{1}{2}]}, \qquad -1 \leqq x \leqq 1, \tag{1.1}$$

and

$$\frac{(n+1)(1-x^2)^{\frac{1}{2}n-1}(1+r^2-2rx)^{\frac{1}{2}(1-n)}(1-x)}{(n-nr+1+r)B(\frac{1}{2}n,\frac{1}{2})}, \qquad -1 \leqq x \leqq 1, \tag{1.2}$$

and are respectively the Madow–Leipnik approximation distribution for the serial correlation coefficient, circularly defined with known non-null mean, and Daniels' (1956) modified approximation distribution for the coefficient, circularly defined with fitted mean. (Note correction of misprints in (1.2).)

General expressions for the uncorrected moments (u.m's) of (1.1) were given by Kendall (1957) and by White (1957); these may also be obtained from Leipnik's (1958) Neumann-type series for the characteristic function (ch.f.). They have the form of polynomials in r, with coefficients involving the Hermite polynomial coefficients. The first four central moments (c.m's) of (1.1) have been derived from the u.m's by Jenkins (1956), Kendall (1957) and White (1957). General expressions for the u.m's and c.m's of (1.2) do not seem to appear in the literature.

In this paper we consider firstly the u.m's of (1.1) about $x = 1$; these are shown to be proportional to Gaussian hypergeometric functions, and to lead to representations of the c.m's and of the ch.f. by hypergeometric functions in two variables. The u.m's of (1.2) about $x = 1$ are closely related to those of (1.1), yielding corresponding formulas for its c.m's and ch.f.

2. Notation and terminology.

(i) Let ${}_n\mu_j'$ (1) and ${}_n\mu_j$ denote the jth moment about $x = 1$ and the jth central moment respectively, for distribution (1.1).

(ii) Let ${}_n\bar{\mu}_j'$ (1) and ${}_n\bar{\mu}_j$ denote the jth moment about $x = 1$ and the jth central moment respectively, for distribution (1.2).

(iii) The normalizing coefficient for the pdf (1.2) occurs repeatedly; put $(n+1)/(n-nr+1+r) = K$, say.

(iv) A number of results concerning hypergeometric functions will be used in

Received September 9, 1969; revised January 22, 1970.

subsequent sections; these are summarized here together with relevant references:

$$\tag{2.1} \begin{gathered} {}_2F_1[a,b;c;z] = \frac{\Gamma(c)}{\Gamma(b)\Gamma(c-b)}\int_0^1 y^{b-1}(1-y)^{c-b-1}(1-yz)^{-a}\,dy, \\ c > b > 0, \qquad |\arg(1-z)| < \pi, \end{gathered}$$

Erdélyi (1953) 2.12 (1).

$$\tag{2.2} \begin{aligned} c[c-1-(2c-a-b-1)z]\,{}_2F_1[a,b;c;z] + (c-a)(c-b)z\,{}_2F_1[a,b;c+1;z] \\ -c(c-1)(1-z)\,{}_2F_1[a,b;c-1;z] = 0, \end{aligned}$$

Erdélyi (1953), 2.8 (45).

$$\tag{2.3} {}_2F_1[a,b;a-b+1;z] = (1+z)^{-a}\,{}_2F_1\left[\frac{a}{2},\frac{a+1}{2};a-b+1;\frac{4z}{(1+z)^2}\right],$$

Erdélyi (1953), 2.11 (34), a Goursat transformation.

$$\tag{2.4} {}_2F_1[a,b;c;z] = (1-z)^{-b}\,{}_2F_1[c-a,b;c;z/(z-1)],$$

Erdélyi (1953), 2.9 (4), a Kummer transformation.

(v) The following hypergeometric functions in two variables will also be required:

Appell's hypergeometric function of the first kind

$$\tag{2.5} F_1(a;b,c;d;x,y) = \sum_{m,n} \frac{(a)_{m+n}(b)_m(c)_n\,x^m y^n}{(d)_{m+n}\,m!\,n!}$$

$$\tag{2.6} = \frac{\Gamma(d)}{\Gamma(a)\Gamma(d-a)}\int_0^1 t^{a-1}(1-t)^{d-a-1}(1-tx)^{-b}(1-ty)^{-c}\,dt,$$

$d > a > 0$, Erdélyi (1953), 5.8(5).

Humbert's corresponding confluent hypergeometric function

$$\tag{2.7} \phi_1(a;b;d;x,y) = \sum_{m,n} \frac{(a)_{m+n}(b)_m\,x^m y^n}{(d)_{m+n}\,m!\,n!}.$$

Appell's hypergeometric function of the third kind

$$\tag{2.8} F_3(a,b,c,d;e;x,y) = \sum_{m,n} \frac{(a)_m(b)_n(c)_m(d)_n\,x^m y^n}{(e)_{m+n}\,m!\,n!}.$$

Humbert's corresponding confluent hypergeometric function

$$\tag{2.9} Z_1(a,b,c;e;x,y) = \sum_{m,n} \frac{(a)_m(b)_n(c)_m\,x^m y^n}{(e)_{m+n}\,m!\,n!}.$$

N.B. $(u)_i$ is Pochhammer's symbol for the rising factorial $u(u+1)\cdots(u+i-1)$.

3. Circular coefficient with known non-null mean. The u.m's about the upper end-point of the distribution (1.1) may be obtained by using the transformation $y = (1+x)/2$ and the basic hypergeometric integral (2.1):

$$
\begin{aligned}
{}_n\mu_j'(1) &= \int_{-1}^{1} \frac{(x-1)^j(1-x^2)^{\frac{1}{2}(n-1)}(1+r^2-2rx)^{-\frac{1}{2}n}}{B[\frac{1}{2}(n+1),\frac{1}{2}]}\,dx \\
&= \frac{(-1)^j 2^{j+n}}{(1+r)^n}\int_0^1 \frac{y^{\frac{1}{2}(n-1)}(1-y)^{j+\frac{1}{2}(n-1)}\{1-4ry/(1+r)^2\}^{-\frac{1}{2}n}}{B[\frac{1}{2}(n+1),\frac{1}{2}]}\,dy \qquad (3.1)\\
&= \frac{(-2)^j[\frac{1}{2}(n+1)]_j}{(1+r)^n(n+1)_j}\,{}_2F_1\left[\frac{n}{2},\frac{n+1}{2};j+n+1;4r/(1+r)^2\right].
\end{aligned}
$$

A three-term recurrence formula for ${}_n\mu'_{j+1}$ (1) follows, using the relationship for contiguous hypergeometric functions (2.2):

$$
\begin{aligned}
(n+2+2j)r\,{}_n\mu'_{j+1}(1) &= \{(n+j)(1-r)^2-2(1+2j)r\}\,{}_n\mu_j'(1) \qquad (3.2)\\
&\quad +(n-1+2j)(1-r)^2\,{}_n\mu'_{j-1}(1).
\end{aligned}
$$

Application of the Goursat and Kummer transformations (2.3) and (2.4) to (3.1) yields

$$
{}_n\mu_j'(1) = \frac{(-2)^j[\frac{1}{2}(n+1)]_j}{(n+1)_j}\,{}_2F_1[n,-j;j+n+1;r] \qquad (3.3)
$$

$$
= \frac{2^j(r-1)^j[\frac{1}{2}(n+1)]_j}{(n+1)_j}\,{}_2F_1[j+1,-j;j+n+1;r/(r-1)]. \qquad (3.4)
$$

Other representations (e.g. as a Legendre function or as a Jacobi polynomial) are obtainable. Note that (3.3) is a polynomial in r (and will yield the usual formulas for the u.m's about $x = 0$), whereas (3.4) involves a terminating inverse factorial series in n (and so will give more readily asymptotic formulas for large n).

General formulas for the c.m's of the Madow–Leipnik distribution can now be found, either by summation of the u.m's using expression (3.1),

$$
{}_n\mu_j = (1-\mu)^j(1+r)^{-n}F_3\left(\tfrac{1}{2}(n+1),\tfrac{1}{2}(n+1),\tfrac{1}{2}n,-j;n+1;\frac{4r}{(1+r)^2},\frac{2}{1-\mu}\right), \qquad (3.5)
$$

or by integration of $(x-\mu)^j$ over the pdf using (2.6),

$$
{}_n\mu_j = (-1-\mu)^j(1+r)^{-n}F_1\left(\tfrac{1}{2}(n+1);\tfrac{1}{2}n,-j;n+1;\frac{4r}{(1+r)^2},\frac{2}{1+\mu}\right), \qquad (3.6)
$$

where $\mu = 1+{}_n\mu_1'$ (1) $= rn/(n+2)$ (the mean). The equivalence of (3.5) and (3.6) can be shown using a well-known transformation for such series, see e.g. Erdélyi (1953), 5.11 (11).

Similar results are obtainable for the ch.f. Summation of the u.m's using expression (3.1) gives

$$\text{(3.7)} \qquad \begin{aligned} \text{ch.f.} &= e^{it}\textstyle\sum_j {}_n\mu_j'(1)(it)^j/j! \\ &= e^{it}(1+r)^{-n}Z_1\left(\tfrac{1}{2}(n+1), \tfrac{1}{2}(n+1), \tfrac{1}{2}n; n+1; \frac{4r}{(1+r)^2}, -2it\right), \end{aligned}$$

whereas integration of e^{itx} over the pdf using the confluent form of (2.6) yields

$$\text{(3.8)} \qquad \text{ch.f.} = e^{-it}(1+r)^{-n}\phi_1\left(\tfrac{1}{2}(n+1); \tfrac{1}{2}n; n+1; \frac{4r}{(1+r)^2}, 2it\right).$$

The confluent form of Erdélyi's formula 5.11 (11) demonstrates the equivalence of (3.7) and (3.8). Note that Leipnik's Neumann-type series is also a double power series.

4. Circular coefficient with fitted (non-null) mean. The u.m's about the upper end-point of the distribution (1.2) are readily expressible in terms of those of distribution (1.1),

$$\text{(4.1)} \qquad \begin{aligned} {}_n\bar{\mu}_j'(1) &= \int_{-1}^{1} \frac{(n+1)(x-1)^j(1-x^2)^{\frac{1}{2}n-1}(1+r^2-2rx)^{\frac{1}{2}(1-n)}(1-x)}{(n-nr+1+r)B(\frac{1}{2}n, \frac{1}{2})}\,dx \\ &= -K\cdot {}_{n-1}\mu_{j+1}'(1), \end{aligned}$$

whence expressions corresponding to (3.1), (3.2), (3.3) and (3.4) may be obtained.

General formulas for the central moments and for the ch.f. follow as before; these are

$$\text{(4.2)} \qquad {}_n\bar{\mu}_j = K(1-\bar{\mu})^j(1+r)^{1-n}F_3\left(\tfrac{1}{2}n, \tfrac{1}{2}n+1, \tfrac{1}{2}(n-1), -j; n+1; \frac{4r}{(1+r)^2}, \frac{2}{1-\bar{\mu}}\right)$$

$$\text{(4.3)} \qquad = K(-1-\bar{\mu})^j(1+r)^{1-n}F_1\left(\tfrac{1}{2}n; \tfrac{1}{2}(n-1), -j; n+1; \frac{4r}{(1+r)^2}, \frac{2}{1+\bar{\mu}}\right),$$

and

$$\text{(4.4)} \qquad \text{ch.f.} = K\,e^{it}(1+r)^{1-n}Z_1\left(\tfrac{1}{2}n, \tfrac{1}{2}n+1, \tfrac{1}{2}(n-1); n+1; \frac{4r}{(1+r)^2}, -2it\right)$$

$$\text{(4.5)} \qquad = K\,e^{-it}(1+r)^{1-n}\phi_1\left(\tfrac{1}{2}n; \tfrac{1}{2}(n-1); n+1; \frac{4r}{(1+r)^2}, 2it\right),$$

where $\bar{\mu} = 1 + {}_n\bar{\mu}_1'(1)$

$$\text{(4.6)} \qquad = \left[-1+(n-1)r-\frac{(n-1)nr^2}{n+3}\right]\Big/(n-nr+1+r)$$

(the mean).

In particular

$$(4.7)\qquad {}_n\bar{\mu}_2 = K\left[\frac{1+2\bar{\mu}}{n+1}+\bar{\mu}^2\right]-K\frac{(n-1)r}{n+1}\left[\frac{3}{n+3}+2\bar{\mu}+\bar{\mu}^2\right]$$

$$+K\frac{(n-1)nr^2}{(n+1)(n+3)}[1+2\bar{\mu}]-K\frac{(n-1)nr^3}{(n+3)(n+5)}$$

$$(4.8)\qquad = \frac{n-\dfrac{2(n-1)nr}{n+3}+\dfrac{2(n-1)(n-3)nr^3}{(n+3)(n+5)}-\dfrac{(n-1)^2(n-3)nr^4}{(n+3)^2(n+5)}}{(n-nr+1+r)^2},$$

$$(4.9)\qquad \frac{{}_n\bar{\mu}_3}{K} = -\left[\frac{3}{(n+1)(n+3)}+\frac{3\bar{\mu}+3\bar{\mu}^2}{n+1}+\bar{\mu}^3\right]$$

$$+\frac{(n-1)r}{n+1}\left[\frac{3(1+3\bar{\mu})}{n+3}+3\bar{\mu}^2+\bar{\mu}^3\right]$$

$$-\frac{(n-1)nr^2}{(n+1)(n+3)}\left[\frac{6}{n+5}+3\bar{\mu}+3\bar{\mu}^2\right]$$

$$+\frac{(n-1)nr^3}{(n+3)(n+5)}[1+3\bar{\mu}]-\frac{(n-1)n(n+2)r^4}{(n+3)(n+5)(n+7)},$$

$$(4.10)\qquad \frac{{}_n\bar{\mu}_4}{K} = \left[\frac{3(1+4\bar{\mu})}{(n+1)(n+3)}+\frac{6\bar{\mu}^2+4\bar{\mu}^3}{n+1}+\bar{\mu}^4\right]$$

$$-\frac{(n-1)r}{n+1}\left[\frac{15}{(n+3)(n+5)}+\frac{3(4\bar{\mu}+6\bar{\mu}^2)}{n+3}+4\bar{\mu}^3+\bar{\mu}^4\right]$$

$$+\frac{(n-1)nr^2}{(n+1)(n+3)}\left[\frac{6(1+4\bar{\mu})}{n+5}+6\bar{\mu}^2+4\bar{\mu}^3\right]$$

$$-\frac{(n-1)nr^3}{(n+3)(n+5)}\left[\frac{10}{n+7}+4\bar{\mu}+6\bar{\mu}^2\right]$$

$$+\frac{(n-1)n(n+2)r^4}{(n+3)(n+5)(n+7)}[1+4\bar{\mu}]-\frac{(n-1)n(n+2)r^5}{(n+5)(n+7)(n+9)}.$$

REFERENCES

[1] Daniels, H. E. (1956). The approximate distribution of serial correlation coefficients. *Biometrika* **43** 169–185.

[2] Erdélyi, A. (1953). *Higher Transcendental Functions* **1**. McGraw-Hill, New York.

[3] Jenkins, G. M. (1956). Tests of hypotheses in the linear auto-regressive model. II Null distributions for higher order schemes: non-null distributions. *Biometrika* **43** 186–199.

[4] Kendall, M. G. (1957). The moments of the Leipnik distribution. *Biometrika* **44** 270–272.

[5] Leipnik, R. (1958). Note on the characteristic function of a serial-correlation distribution. *Biometrika* **45** 559–562.

[6] White, J. S. (1957). Approximate moments for the serial correlation coefficient. *Ann. Math. Statist.* **28** 798–802.