Ann. Inst. Statist. Math.
35 (1983), Part A, 49-61

ACCURATE CONFIDENCE INTERVALS FOR DISTRIBUTIONS WITH ONE PARAMETER

C. S. WITHERS

(Received Aug. 6, 1981; revised Feb. 18, 1982)

Summary

Let $\hat{\theta}_n$ be an estimate of a real parameter θ. Suppose that for some function $c(\cdot)$ and some random variable (r.v.) τ_n, the distribution of

$$Z_n=(c(\theta)-c(\hat{\theta}_n))/\tau_n$$

is continuous and depends only on θ and n and that the cumulants of Z_n can be expanded in the form

$$K_r(Z_n)\approx\sum_{i=r-1}^{\infty} a_{ri}(\theta)n^{-i}\ .$$

Then a confidence interval for θ can be constructed with level $1-\alpha+O(n^{-j/2})$ for any given value of α and j.

1. Introduction

This paper offers ways of improving the accuracy of approximate confidence intervals (C.I.'s) for one parameter problems when the cumulants of the parameter estimate have a very commonly occurring type of asymptotic expansion.

Section 2 summarises the usual first-order approximations to C.I.'s based on an asymptotically normal estimate, and indicates the magnitude of their error. Section 3 shows how to reduce this error, and Section 4 gives some examples.

2. First-order confidence intervals

Let θ be an unknown real parameter known to lie in an interval $[a, b]$, where $-\infty\leqq a<b\leqq\infty$. Let $c(\cdot)$ be a one to one increasing func-

Keywords and phrase: Accurate confidence intervals, Cornish-Fisher expansions, cumulants, percentiles.

tion on $[a, b]$. Let Φ, ϕ be the distribution function and density of $\mathcal{N}(0, 1)$. Suppose that $\tau_n>0$ is an r.v. bounded in probability away from 0 and ∞ and that the distribution of

$$Z_n=(c(\theta)-c(\hat{\theta}_n))/\tau_n$$

depends only on θ and n. When

(1) $$Z_n \text{ is asymptotically } \mathcal{N}(0, v(\theta)/n) ,$$

then a confidence interval for θ of level approximately $1-\alpha$ is

(2) $$V_{1n}(\hat{\theta}_n, x_2)\leqq\theta\leqq V_{1n}(\hat{\theta}_n, x_1)$$

where $V_{1n}(\theta, x)=c^{-1}(c(\theta)+n^{-1/2}x\tau_n v(\theta)^{1/2})$ and x_1, x_2 are chosen so that

(3) $$\Phi(x_1)-\Phi(x_2)=1-\alpha .$$

For a one-sided test one chooses $x_1=\infty$ or $x_2=-\infty$, and for a two-sided test, the usual choice is

(4) $$x_1=-x_2=\Phi^{-1}(1-\alpha/2) .$$

The C.I. (2) has level $1-\alpha+e_n$, where generally speaking *the error* e_n has magnitude $n^{-1/2}$ as $n\rightarrow\infty$, unless either (4) holds—i.e. the tails are equal—or the distribution of Z_n is symmetric; in either of these events the magnitude of the error reduces to n^{-1}. For more precise conditions see Withers [9].

Example 2.1. Suppose $\hat{\theta}_n\sim\mathcal{N}(\theta, V(\theta)/n)$. Then $c(\hat{\theta}_n)\sim\mathcal{N}(c(\theta), V_c(\theta)/n)$ where $V_c(\theta)=c^{(1)}(\theta)^2V(\theta)$. (We use $f^{(r)}(\theta)$ to denote the rth derivative of $f(\theta)$.) The choice $\tau_n^2=V_c(\hat{\theta}_n)$ implies $v(\theta)=1$. Generally $c(\cdot)$ is chosen either

(i) for simplicity—such as $c(\theta)=\theta$; or

(ii) to satisfy $c(a)=-\infty$, $c(b)=\infty$—so that the interval (2) contains no points outside $[a, b]$; or

(iii) so that $V_c(\theta)=1$—that is $c(\theta)=\int_{\theta_0}^{\theta} V(x)^{-1/2}dx$; or

(iv) to reduce the bias or skewness of $c(\hat{\theta}_n)$.

However *none* of these choices reduce the magnitude of the error of (2).

Example 2.1(a). Let $\hat{\theta}_n$ be the sample correlation of a sample of size $N=n$ from a bivariate normal population with correlation θ. Let $c(\theta)=\tanh^{-1}\theta$. This choice satisfies (ii), (iii), and (iv). But since its bias is still $O(n^{-1})$, the C.I. (2) still has error $e_n=O(n^{-1/2})$ or $O(n^{-1})$ if the tails are equal, the same as for the choice $c(\theta)=\theta$, $\tau_n=1-\hat{\theta}_n^2$. (The same is true with choices of n such as $N-1$ or $N-3$.)

Example 2.2. Let $\{X_1, \cdots, X_n\}$ be a random sample from $F_0((x-\theta)/\sigma)$ where F_0 is a given distribution with variance 1. Choose $c(\theta)=\theta$ and $\hat{\theta}_n$, τ_n such that the distribution of $(\hat{\theta}_n-\theta)/\tau_n$ does not depend on (θ, σ). (This is true for a wide class of estimates $(\hat{\theta}_n, \tau_n)$.) Then in general the interval (2) has error $O(n^{-1/2})$ unless either the tails are equal or F_0 is symmetric, in which case the error is $O(n^{-1})$.

3. Improved approximations

We now give a method for obtaining a C.I. for θ with error $O(n^{-j/2})$ for any given j. We replace (1) by the stronger condition that the cumulants of Z_n have expansions of the form

$$(5) \qquad K_r(Z_n)\approx \sum_{i=r-1}^{\infty} a_{ri}n^{-i}, \qquad r\geqq 1,\ a_{10}=0$$

where $\{a_{ri}=a_{ri}(\theta)\}$ are functions of θ, and we assume that the distribution of Z_n is absolutely continuous. By Fisher and Cornish [3], (5) implies (1) with $v(\theta)=a_{21}(\theta)$.

The assumption (5) holds for a wide class of $(\hat{\theta}_n, \tau_n)$; see for example Withers [9] where formulas for $\{a_{ri}\}$ are given when $(\hat{\theta}_n, \tau_n)$ are regular functionals of the empirical distribution of a random sample of size n.

Let

$$\mathcal{P}_n(x)=\Pr(n^{1/2}a_{21}^{-1/2}Z_n\leqq x).$$

Upon substitution of (5) into the expansions of Cornish and Fisher, one obtains the asymptotic expansion

$$(6) \qquad \mathcal{P}_n^{-1}(\Phi(x))\approx x+\sum_{r=1}^{\infty} n^{-r/2}g_r(x)$$

where $g_r(x)$ is a polynomial of degree $r+1$ given in terms of

$$(7) \qquad A_{ri}=a_{21}^{-r/2}a_{ri},$$

in Appendix 1.

Set

$$g_r(x, \theta)=\begin{cases} x, & r=0 \\ g_r(x), & r\geqq 1 \end{cases}$$

and for a given value of x, set

$$p_i(\theta)=P_i(c(\theta))=-\tau_n a_{21}(\theta)^{1/2}g_{i-1}(x, \theta), \qquad i\geqq 1$$

and

$$R_{jnx}(\theta)=c(\theta)+\sum_{i=1}^{j} n^{-i/2}p_i(\theta) , \qquad j\geqq 1 .$$

THEOREM 3.1. *Suppose that* (6) *holds for* $x=x_1, x_2$ *satisfying* (3) *and that for some* $j\geqq 1$, $R_{jnx}(\theta)$ *is one to one increasing in a suitably large neighbourhood of* $\hat{\theta}_n$. *Then a confidence interval of level* $1-\alpha$ *with error* $O(n^{-j/2})$ *is given by*

$$V_{jn}(\hat{\theta}_n, x_2)\leqq\theta\leqq V_{jn}(\hat{\theta}_n, x_1) \tag{8}$$

where

$$V_{jn}(\theta, x)=c^{-1}\left(c(\theta)+\sum_{i=1}^{j} n^{-i/2}q_i(\theta)\right) , \tag{9}$$

$$q_1(\theta)=\tau_n x a_{21}(\theta)^{1/2} ,$$

$$q_2(\theta)=\tau_n a_{21}(\theta)^{1/2}g_1(x, \theta)+\tau_n^2x^2c^{(1)}(\theta)^{-1}a_{21}^{(1)}(\theta)/2 ,$$

$$\begin{aligned} q_3(\theta)=&\tau_n a_{21}(\theta)^{1/2}g_2(x, \theta)+\tau_n^2 x c^{(1)}(\theta)^{-1}\partial\{a_{21}(\theta)g_1(x, \theta)\}/\partial\theta \\ &+\tau_n^3x^3c^{(1)}(\theta)^{-2}a_{21}(\theta)^{1/2}(2a_{21}^{(2)}(\theta)+a_{21}^{(1)}(\theta)^2a_{21}(\theta)^{-1})/8 \\ &-\tau_n^3x^3c^{(1)}(\theta)^{-3}c^{(2)}(\theta)a_{21}(\theta)^{1/2}a_{21}^{(1)}(\theta)/4 , \end{aligned}$$

$$\begin{aligned} q_4(\theta)=&-p_4(\theta)-\sum_{i=1}^{3}\bar{P}_{4-i}^{(1)}q_i(\theta)-\bar{P}_2^{(2)}q_1(\theta)^2/2-\bar{P}_1^{(2)}q_1(\theta)q_2(\theta) \\ &-\bar{P}_1^{(3)}q_1(\theta)^3/6 , \end{aligned}$$

and

$$\begin{aligned} q_5(\theta)=&-p_5(\theta)-\sum_{i=1}^{4}P_{5-i}^{(1)}q_i(\theta)-\bar{P}_3^{(2)}q_1(\theta)^2/2-\bar{P}_2^{(2)}q_1(\theta)q_2(\theta) \\ &-\bar{P}_1^{(2)}(q_2(\theta)^2/2+q_1(\theta)q_3(\theta))-\bar{P}_2^{(3)}q_1(\theta)^3/6-\bar{P}_1^{(3)}q_1(\theta)^2q_2(\theta)/2 \\ &-\bar{P}_1^{(4)}q_1(\theta)^4/24 , \end{aligned}$$

where $\{\bar{P}_i^{(r)}=P_i^{(r)}(c(\theta)),\ 1\leqq r\leqq 4\}$ *are given by*

$$\bar{P}_i^{(1)}=c^{(1)}(\theta)^{-1}p_i^{(1)}(\theta)$$

$$\bar{P}_i^{(2)}=c^{(1)}(\theta)^{-2}p_i^{(2)}(\theta)-c^{(1)}(\theta)^{-3}c^{(2)}(\theta)p_i^{(1)}(\theta) ,$$

$$\begin{aligned} \bar{P}_i^{(3)}=&c^{(1)}(\theta)^{-3}p_i^{(3)}(\theta)-3c^{(1)}(\theta)^{-4}c^{(2)}(\theta)p_i^{(2)}(\theta) \\ &+\{3c^{(1)}(\theta)^{-5}c^{(2)}(\theta)^2-c^{(1)}(\theta)^{-4}c^{(3)}(\theta)\}p_i^{(1)}(\theta) , \end{aligned}$$

$$\begin{aligned} \bar{P}_i^{(4)}=&c^{(1)}(\theta)^{-4}p_i^{(4)}(\theta)-6c^{(1)}(\theta)^{-5}c^{(2)}(\theta)p_i^{(3)}(\theta) \\ &+\{15c^{(1)}(\theta)^{-6}c^{(2)}(\theta)^2-4c^{(1)}(\theta)^{-5}c^{(3)}(\theta)\}p_i^{(2)}(\theta) \\ &+\{-15c^{(1)}(\theta)^{-6}c^{(2)}(\theta)^3+10c^{(1)}(\theta)^{-5}c^{(2)}(\theta)c^{(3)}(\theta) \\ &\quad -c^{(1)}(\theta)^{-4}c^{(4)}(\theta)\}p_i^{(1)}(\theta) . \end{aligned}$$

PROOF. By (6) with probability $\Phi(x)+O(n^{-j/2})$, $c(\hat{\theta}_n)\geqq R_{jnx}(\theta)$, which for R_{jnx} one to one is equivalent to

$$c(V_{jn}(\hat{\theta}_n, x))+O_p(n^{-j/2})\geqq c(\theta) ,$$

where $q_i(\theta)=Q_j(c(\theta))$, and $g(x)\simeq x+\sum_1^\infty n^{-r/2}Q_r(x)$ is the inverse of $x(g)\simeq g+\sum_1^\infty n^{-r/2}P_r(g)$. Now apply Appendix 2.

Note 1. For Example 2.1 for a given $c(\cdot)$, $\{q_j(\hat{\theta}_n)\}$ are independent of the choice of τ_n, provided τ_n is a function of $\hat{\theta}_n$ independent of n; thus $\{q_j\}$ are most easily computed choosing $\tau_n=1$.

Note 2. If $j=2$ and $\tau_n=1$ or $V_c(\hat{\theta}_n)^{1/2}$, then for Example 2.1 the interval (8) is just the same as that given by Withers [6] for

$$Y_n(\theta)=n^{1/2}(c(\theta)-c(\hat{\theta}_n))V_c(\hat{\theta}_n)^{-1/2} .$$

4. Some examples

Example 4.1. Returning to Example 2.1(a) we have

Case 1. $c(\theta)=\theta$ and $\tau_n=1$: then $a_{21}(\theta)=(1-\theta^2)^2$ and by (4.12) of Withers [6], $g_1(x)=\theta(x^2-1/2)$, $g_2(x)=(3x-x^3)/4+\theta^2(-5x+4x^3)/4$, so that

$$q_1(\theta)=x(1-\theta^2) , \qquad q_2(\theta)=-(\theta-\theta^3)(1/2+x^2) ,$$

and

$$q_3(\theta)=(1-\theta^2)\{x-x^3+\theta^2(5x+4x^3)\}/4 .$$

Case 2. $c(\theta)=\tanh^{-1}\theta$ and $\tau_n=1$: then $a_{21}(\theta)=1$ and by (4.13) of Withers [6], $g_1(x)=-\theta/2$, $g_2(x)=(9x+x^3)/12-\theta^2x/4$, so that

Table 1. Error in exact probability of one-sided nominally 95% confidence interval given by (8) with $\tau_n=1$ for Example 2.1(a)

	θ	$j=1$	$j=2$	$j=3$
$n=5$				
	−.9	−.202	−.09(6)	−.05(3)[2]
	−.5	−.154	−.06(6)	−.03(2)
Case 1	0	−.09(0)	−.02(8)	−.007
	.5	−.01(7)	.00(9)	.010
	.9	.043	−.19(2)	.01(4)
	−.9	−.07(1)	−.04(1)[2]	−.01(5)[2]
	−.5	−.05(8)	−.03(3)	−.01(0)
Case 2	0	−.04(1)	−.024	−.005
	.5	−.021	−.017	−.001
	.9	−.00(1)	−.01(5)	.00(2)

Table 1. (Continued)

	θ	$j=1$	$j=2$	$j=3$
$n=10$				
	$-.9$	$-.12(8)$	$-.04(5)$	$-.01(9)$
	$-.5$	$-.09(3)$	$-.02(7)$	$-.00(9)$
Case 1	0	$-.04(5)$	$-.00(5)$	$.00(1)$
	$.5$	$.00(6)$	$.00(8)$	$.00(0)$
	$.9$	$.04(6)$	$-.07(1)$	$.01(3)^{2}$
	$-.9$	$-.03(5)$	$-.01(6)$	$-.00(4)$
	$-.5$	$-.028$	$-.013$	$-.002$
Case 2	0	$-.018$	$-.01(0)$	$-.00(1)$
	$.5$	$-.00(7)$	$-.00(8)$	$-.00(0)$
	$.9$	$.00(2)^{2}$	$-.00(8)^{3}$	$.00(1)$
$n=20$				
	$-.9$	$-.08(4)$	$-.02(2)$	$-.00(7)$
	$-.5$	$-.05(7)$	$-.01(1)$	$-.00(2)$
Case 1	0	$-.02(3)$	$.00(0)$	$.00(0)^{3}$
	$.5$	$.01(3)$	$.00(2)^{2}$	$-.00(0)^{2}$
	$.9$	$.04(2)$	$-.03(0)$	$.00(6)$
	$-.9$	$-.01(9)$	$-.00(5)$	$-.00(1)$
	$-.5$	$-.015$	$-.00(5)$	$-.00(0)$
Case 2	0	$-.00(8)$	$-.00(4)^{2}$	$-.00(0)$
	$.5$	$-.00(1)$	$-.00(3)^{3}$	$.00(0)^{2}$
	$.9$	$.00(4)$	$-.00(4)$	$.00(0)^{5}$

Tables 1 and 2 were calculated by quadratic interpolation on the nearest three points in David's '*Tables of the Correlation Coefficient*' (1938). The error of this formula was found using the tables on a point outside this range and is indicated by the brackets and superscripts:

$.00(8)$ means $.008 \pm .001$, $\quad .00(6)^{3}$ means $.006 \pm .003$.

Table 2. Error in exact probability of two-sided nominally 90% confidence interval given by (8) with $\tau_n=1$ for Example 2.1(a)

	θ	$j=1$	$j=2$	$j=3$
$n=5$				
	0	$-.18(0)$	$-.05(6)$	$-.01(5)$
Case 1	$\pm .5$	$-.17(1)$	$-.05(7)$	$-.02(2)$
	$\pm .9$	$-.15(9)$	$-.28(9)$	$-.03(9)^{3}$
	0	$-.08(2)^{2}$	$-.04(9)$	$.01(0)$
Case 2	$\pm .5$	$-.07(9)^{2}$	$-.05(0)$	$-.01(2)^{3}$
	$\pm .9$	$-.07(2)^{2}$	$-.05(6)^{3}$	$-.01(2)^{4}$

$n=10$					
		0	$-.09(0)^2$	$-.01(1)$	$.00(2)$
	Case 1	$\pm.5$	$-.08(7)^2$	$-.02(1)^2$	$-.00(8)$
		$\pm.9$	$-.08(2)^2$	$-.11(7)^2$	$-.00(6)^3$
		0	$-.03(3)^3$	$-.02(0)$	$-.00(2)^2$
	Case 2	$\pm.5$	$-.03(5)$	$-.02(1)$	$-.00(3)$
		$\pm.9$	$-.03(3)^2$	$-.02(5)^3$	$-.00(3)^3$
$n=20$					
		0	$-.04(6)^2$	$-.00(0)^2$	$.00(1)^4$
	Case 1	$\pm.5$	$-.04(4)^2$	$-.00(8)^3$	$-.00(2)^3$
		$\pm.9$	$-.04(2)^2$	$-.05(2)$	$-.00(1)^2$
		0	$-.01(6)^2$	$-.00(8)^3$	$-.00(0)^3$
	Case 2	$\pm.5$	$-.01(6)$	$-.00(9)^4$	$-.00(0)^3$
		$\pm.9$	$.01(5)^2$	$-.01(0)$	$-.00(0)^6$

$$q_1(\theta)=x\ , \qquad q_2(\theta)=-\theta/2\ , \qquad q_3(\theta)=(3x+x^3)/12+\theta^2x/4\ . \tag{10}$$

When $p_n(\hat{\theta}_n;\theta)$, the density of $\hat{\theta}_n$, satisfies $p_n(\hat{\theta}_n;-\theta)=p_n(-\hat{\theta}_n;\theta)$, as is true for Example 2.1(a), then the error of the two-sided confidence interval given by (4), (8) is symmetric in θ.

Example 4.2. Let $X_1,\cdots,X_n$ be i.i.d. χ^2_θ, $S=\sum_1^n X_i$, $\hat{\theta}_n=S/n$. Taking $c(\theta)=\theta$ and $\tau_n=1$ gives $a_{21}(\theta)=2\theta$ and

$$(\theta-\hat{\theta}_n)(n/2\theta)^{1/2}=(m-S)(2m)^{-1/2}\ ,$$

where $m=n\theta$. Since $S\sim\chi^2_m$,

$$g_r(x,\theta)=\theta^{-r/2}(-1)^r g_{r0}(x)$$

where $g_{r0}(x)$ denotes $g_r(x)$ for $Z_m=(\chi^2_m-m)/m$, given, for example, for $1\leqq r\leqq 6$, by (3a) of Fisher and Cornish [3].

Hence

$$q_1(\theta)=(2\theta)^{1/2}x\ , \qquad q_2(\theta)=(2+x^2)/3\ ,$$

$$q_3(\theta)=-(2\theta)^{-1/2}(2x+x^3)/18\ ,$$

Table 3. Error in exact probability of one-sided nominally 95% confidence interval given by (8) with $\tau_n=1$ for Example 4.2

$m=n\theta$	$j=1$	$j=2$	$j=3$	$j=4$
5	$-.0795$	.0085	$-.0034$	.0012
10	$-.0502$	.0034	$-.0007$	.00012
15	$-.0388$	.0021	$-.00032$	.00004
20	$-.0324$	.0015	$-.00019$	.00002
100	$-.0128$	.00026	$-.00007$	.00000(5)

$$q_4(\theta)=2\theta^{-1}(-16+7x^2+3x^4)/405 .$$

Example 4.3. Returning to Example 2.2, $a_{21}(\theta)=1$ and $g_r(x, \theta)=g_r(x)$ do not depend on θ, so that

$$q_i(\theta)=\tau_n g_{i-1}(x)$$

and

$$V_{jn}(\hat{\theta}_n, x)=\hat{\theta}_n+\tau_n n^{-1/2}\left\{x+\sum_{r=1}^{j-1} n^{-r/2}g_r(x)\right\} .$$

Fisher and Cornish [3] give $\{g_r(x)\}$ for the case of Student's t-statistic.

Acknowledgements

The first draft of this paper was written in 1978 while an Exchange Visitor at the Division of Mathematics and Statistics, C.S.I.R.O., Canberra. I wish to thank D.M.S. for its support, and also David Harte for assistance with Tables 1 and 2.

Addendum

Since this paper was written, Winterbottom [5] has published formulae (A.1)–(A.3) equivalent to ours for the case when $c(\theta)=\theta$, $\tau_n=1$, and $j=5$.

His notation	$\tilde{\theta}(\xi)$	ξ	θ	T	$v(\theta)$	$(-1)^r\kappa_{r,s}$
Our notation	$V_{5n}(\hat{\theta}_n, x)$	x	θ	$\hat{\theta}_n$	$a_{21}(\theta)$	$A_{r,s}$

(Equivalently, taking $c(\theta)=-\theta$, his $\kappa_{r,s}$ is our $A_{r,s}$.)

He applied it to $\hat{\theta}_n=(\chi_n^2(\lambda)-n)/n$ with $\theta=\lambda^2/n$ (NOT λ/n as stated), also to the maximum likelihood estimate, and to Case 2 of Example 4.1, for which he obtains $q_4(\theta)$ and $q_5(\theta)$.

APPENDIX 1

From Corollary 3.1 of Withers [6] we have

LEMMA 1. (5), (7) *implies* (6) *with*

$$g_1(x)=A_{11}+A_{32}(x^2-1)/6 ,$$

$$g_2(x)=A_{22}x/2+A_{43}(x^3-3x)/24+A_{32}^2(-2x^3+5x)/36 ,$$

$$g_3(x)=A_{12}+A_{33}(x^2-1)/6+A_{22}A_{32}(-x^2+1)/6$$

$$+A_{54}(x^4-6x^2+3)/120+A_{32}A_{43}(-x^4+5x^2-2)/24$$
$$+A_{32}^3(12x^4-53x^2+17)/324\ ,$$

and

$$g_r(x)=\sum_{0\leq k\leq (r-1)/2} G'_{r-2k,2k}\, g^*_{r-2k}(x)$$

where for $1\leqq r\leqq 6$, $G_{r,0}$ *and* $g_r^*(x)$ *are given on pages* 214, 215 *of Fisher and Cornish* [3] *with* $a=A_{11}$, $b=A_{22}$, $c=A_{32}$, $d=A_{43}$, $e=A_{54}$, $f=A_{65}$, $g=A_{76}$, $h=A_{87}$: *for example,*

$$(G_{4,0})_3=\textit{coefficient in line 3 of } IV=A_{22}A_{32}^2\ ,$$

and

$$(g_4^*(x))_3=\textit{polynomial/divisor in line 3 of } IV$$
$$=5(2x^3-5x)/72\ ,$$

while the other $\{G_{r,k}\}$ *needed for* $4\leqq r\leqq 6$ *are as follows.*

For $r=4$: $G'_{22}=(A_{23},\ A_{44},\ 2A_{32}A_{33})$.

For $r=5$: $G'_{32}=(A_{13},\ A_{34})$,

$$G'_{14}=(A_{22}A_{33}+A_{32}A_{23},\ A_{55},\ A_{43}A_{33}+A_{32}A_{44},\ 3A_{32}^2A_{33})\ .$$

For $r=6$: $G'_{42}=(A_{42},\ A_{45},\ A_{33}^2+2A_{32}A_{34})$,

$$G'_{24}=(2A_{22}A_{23},\ A_{22}A_{44}+A_{43}A_{23},\ A_{32}^2A_{23}+2A_{32}A_{22}A_{33},\ A_{66},\ A_{33}A_{54}+A_{32}A_{55},\ 2A_{43}A_{44},\ 2A_{32}^3A_{33}A_{43},\ 4A_{32}^3A_{33})\ .$$

APPENDIX 2

A1. *Summary*

This contains some formulas for inverting series. Section 2 gives the inverse of

$$y(\varepsilon)=\sum_1^\infty \varepsilon^r P_r \tag{A1.1}$$

as a power series in ε. Section 3 gives the inverse of

$$x(g)=g+\sum_1^\infty \varepsilon^r P_r(g) \tag{A1.2}$$

as a power series in ε, and gives some statistical applications.

A2. *The first inversion series*

Expressions for the inverse of (A1.1) are well known, e.g. §3.6.25

of Abramowitz and Stegun [1]. However these expressions only give the first few terms of the inverse, as a series in ε. The general term may easily be expressed using the notation of the following lemma.

LEMMA A2. *For* $j=0, 1, 2, \cdots$

$$\left(\sum_{i=1}^{\infty} \varepsilon^i Q_i\right)^j = \sum_{r=j}^{\infty} \varepsilon^r C_{rj}(\{Q_i\})$$

where

$$C_{r0}(\{Q_i\}) = \begin{cases} 1, & r=0 \\ 0, & r>0, \end{cases} \tag{A2.1}$$

and for $r \geqq j \geqq 0$, $C_{rj}(\{Q_i\}) = \sum Q_{k_1} \cdots Q_{k_j}$ *summed over* $\{k_1 + \cdots + k_j = r,\ k_1 \geqq 1, \cdots, k_j \geqq 1\}$, *or equivalently*

$$C_{rj}(\{Q_i\}) = \sum (k_1, \overset{j}{\cdots}, k_r) Q_1^{k_1} \cdots Q_r^{k_r} \tag{A2.2}$$

summed over $\{k_1 + \cdots + k_r = j,\ k_1 + 2k_2 + \cdots + rk_r = r,\ k_1 \geqq 0, \cdots, k_r \geqq 0\}$, *where* $(k_1, \overset{j}{\cdots}, k_r)$ *is the multinomial coefficient* $j!/(k_1! \cdots k_r!)$.

For example $C_{r1}(\{Q_i\}) = Q_r$, $C_{jj}(\{Q_i\}) = Q_1^j$, $C_{j+1,j}(\{Q_i\}) = jQ_1^{j-1}Q_2$,

$$C_{j+2,j}(\{Q_i\}) = \binom{j}{2} Q_1^{j-2} Q_2^2 + \binom{j}{1} Q_1^{j-1} Q_3,$$

$$C_{j+3,j}(\{Q_i\}) = \binom{j}{1} Q_1^{j-1} Q_4 + \binom{j}{3} Q_1^{j-3} Q_2^3 + j(j-1) Q_1^{j-2} Q_2 Q_3.$$

THEOREM A1. *When both series converge, the inverse of*

$$y(\varepsilon) = \sum_{r=1}^{\infty} \varepsilon^r P_r$$

is given for $P_1 \neq 0$ *by*

$$\varepsilon(y) = \sum_{r=1}^{\infty} y^r Q_r$$

where Q_r *is defined recursively by* $Q_1 = P_1^{-1}$,

$$Q_r = -P_1^{-1} \sum_{s=2}^{r} P_s C_{rs}(\{Q_i\}), \qquad r>1.$$

PROOF. Set $y = y(\varepsilon)$, $\varepsilon = \varepsilon(y)$. Then

$$P_1 \varepsilon = y - \sum_{2}^{\infty} P_s \varepsilon^s.$$

But

$$\varepsilon^s = \sum_{r=s}^{\infty} y^r C_{rs}(\{Q_i\}).$$

An alternative formula for P_r was given by McMahon in 1894. An extension of his result to the problem of expressing a power of $\varepsilon(y)$ as a series in $\{y^r\}$ is given in Part IX of David, et al. [2]. Their Table 9 may be used as an alternative to Theorem 1 to obtain Q_r for $r\leqq 11$.

A3. *The second inversion series*

Let $\{P_r\}$ be functions on R with derivatives $\{P_r^{(j)}\}$.

THEOREM A2. *When both series converge, the inverse of*

$$x(g)=g+\sum_{r=1}^{\infty} \varepsilon^r P_r(g)$$

is given by

$$g(x)=x+\sum_{r=1}^{\infty} \varepsilon^r Q_r(x)$$

where $Q_r(x)$ is defined recursively by

$$Q_r(x)=-\sum_{j=0}^{r-1}\sum_{k=j}^{r-1} P_{r-k}^{(j)}(x)C_{kj}(\{Q_i(x)\})/j! \tag{A2.3}$$

and

$$P_r^{(j)}(x)=(d/dx)^j P_r(x) .$$

PROOF. $Q_r(x)$ is the coefficient of ε^r in the Taylor series expansion for

$$g=x-\sum_{r=1}^{\infty} \varepsilon^r P_r(g) .$$

The first five Q_r are as follows.

$$Q_1=-P_1 , \qquad Q_2=-P_2+P_1^{(1)}P_1 ,$$

$$\begin{aligned} Q_3&=-P_3-P_2^{(1)}Q_1-P_1^{(1)}Q_2-P_1^{(2)}Q_1^2/2 \\ &=-P_3+P_1P_2^{(1)}+P_1^{(1)}P_2-P_1^{(1)2}P_1-P_1^{(2)}P_1^2/2 , \end{aligned}$$

$$Q_4=-P_4-P_3^{(1)}Q_1-P_2^{(1)}Q_2-P_1^{(1)}Q_3-P_2^{(2)}Q_1^2/2-P_1^{(2)}Q_1Q_2-P_1^{(3)}Q_1^3/6 ,$$

$$\begin{aligned} Q_5=&-P_5-P_4^{(1)}Q_1-P_3^{(1)}Q_2-P_2^{(1)}Q_3-P_1^{(1)}Q_4-P_3^{(2)}Q_1^2/2-P_2^{(2)}Q_1Q_2 \\ &-P_1^{(2)}(Q_2^2/2+Q_1Q_3)-P_2^{(3)}Q_1^3/6-P_1^{(3)}Q_1^2Q_2/2-P_1^{(4)}Q_1^4/24 . \end{aligned}$$

An alternative formula for $Q_r(x)$ involving multivariate Bell polynomials is given by (3) of Riordan [4]. His formula seems more difficult for algebraic manipulation.

As an application in statistics, consider the problem investigated by Fisher and Cornish [3]. Many standardized asymptotically normal random variables Y_n have rth cumulant of the form

$$\begin{aligned} &l_{1n}=O(n^{-1/2}), && r=1 \\ &1+l_{2n}=1+O(n^{-1}), && r=2 \\ &l_{rn}=O(n^{1-r/2}), && r>2 \text{ as } n\to\infty, \end{aligned}$$

(Here n is usually the sample size or associated degrees of freedom.) Under this assumption they showed that $P_n(x)=\Pr(Y_n\leqq x)$ satisfies expansions of the form

$$\Phi^{-1}(P_n(x))=x-\sum_1^\infty f_r(x, \mathcal{L}_n)$$

and

$$P_n^{-1}(\Phi(x))=x+\sum_1^\infty g_r(x, \mathcal{L}_n)$$

where

$$\Phi(x)=(2\pi)^{-1/2}\int^x \exp(-y^2/2)dy$$

and f_r, g_r are polynomials of degree $r+1$ involving $\mathcal{L}_n=\{l_{rn}\}$ and having magnitude $O(n^{-r/2})$.

They gave the first four f_r and the first six g_r, but no expression for the general term. Expressions for f_5 and f_6 may be obtained from the following application of Theorem A2.

COROLLARY A1. *Let* $Q_r(x, \{P_i\})$ *denote* $Q_r(x)$ *of* (A2.3). *Then*

$$g_r(x, \mathcal{L})=Q_r(x, \{-f_i(x, \mathcal{L})\})$$

and

$$f_r(x, \mathcal{L})=-Q_r(x, \{g_i(x, \mathcal{L})\}).$$

An expression for $f_r(x, \mathcal{L})$ for general r was given by (2.8) of Withers [6]. This may be used in Corollary 1 to obtain any desired $g_r(x, \mathcal{L})$.

In most instances $\sum_1^\infty f_r(x, \mathcal{L}_n)$ and $\sum_1^\infty g_r(x, \mathcal{L}_n)$ can be rewritten in the form $\sum_1^\infty n^{-r/2}f_r(x)$ and $\sum_1^\infty n^{-r/2}g_r(x)$. In this case $\{f_r(x)\}$ and $\{g_r(x)\}$ have the same relationship to each other as do $\{f_r(x, \mathcal{L})\}$ and $\{g_r(x, \mathcal{L})\}$, so that $\{f_r(x), 1\leqq r\leqq 6\}$ are obtainable using Appendix 1 when (5) holds.

D.S.I.R.

REFERENCES

[1] Abramowitz, M. and Stegun, I. A. (1964). *Handbook of Mathematical Functions*, National Bureau of Standards, Washington, D.C.

[2] David, F. N., Kendall, M. G. and Barton, D. E. (1966). *Symmetric Function and Allied Tables*, University Press, Cambridge.

[3] Fisher, R. A. and Cornish, E. A. (1960). The percentile points of distributions having known cumulants, *Technometrics*, **2**, 209-225.

[4] Riordan, John (1949). Inversion formulas in normal variable mapping, *Ann. Math. Statist.*, **20**, 417-425.

[5] Winterbottom, Alan (1979). Cornish-Fisher expansions for confidence limits, *J. R. Statist. Soc.*, B, **41**, 69-75.

[6] Withers, C. S. (1980a). Expansions for asymptotically normal random variables, *Tech. Report* 94, A.M.D., D.S.I.R., Wellington, Submitted to *J. R. Statist. Soc.*, B.

[7] Withers, C. S. (1980b). The distribution and quantiles of a regular functional of the empirical distribution, *Tech. Report* 96, A.M.D., D.S.I.R., Wellington, Submitted to *Ann. Statist.*

[8] Withers, C. S. (1980c). Accurate non-parametric inference—the one sample case, *Tech. Report* 97, A.M.D., D.S.I.R., Wellington, Submitted to *Ann. Statist.*

[9] Withers, C. S. (1983). Expansions for the distribution and quantiles of a regular functional of the empirical distribution with applications to nonparametric confidence intervals, *Ann. Statist.*, **11**, June issue.