Ann. Inst. Statist. Math.
Vol. 40, No. 4, 727-746 (1988)

NONPARAMETRIC CONFIDENCE INTERVALS FOR FUNCTIONS OF SEVERAL DISTRIBUTIONS

C. S. WITHERS

Applied Mathematics Division, Department of Scientific and Industrial Research, Wellington, New Zealand

(Received June 30, 1987; revised June 22, 1988)

Abstract. Let $F = (F_1 \cdots F_k)$ denote k unknown distribution functions and $\hat{F} = (\hat{F}_1 \cdots \hat{F}_k)$ their sample (empirical) functions based on random samples from them of sizes $n_1, \ldots, n_k$. Let $T(F)$ be a real functional of F. The cumulants of $T(\hat{F})$ are expanded in powers of the inverse of n, the minimum sample size. The Edgeworth and Cornish-Fisher expansions for both the standardized and Studentized forms of $T(\hat{F})$ are then given together with confidence intervals for $T(F)$ of level $1 - \alpha + O(n^{-j/2})$ for any given α in $(0, 1)$ and any given j. In particular, confidence intervals are given for linear combinations and ratios of the means and variances of different populations without assuming any parametric form for their distributions.

Key words and phrases: Confidence interval, nonparametric, Cornish-Fisher expansions, functional derivatives.

1. Introduction and summary

In a previous paper, Withers (1982*a*), the author gave Edgeworth and Cornish-Fisher expansions for the distribution and quantiles of a function of a number of unknown parameter estimates $t(\hat{\omega})$ say, in both its standardized and its Studentized forms. This was used in Withers (1983*b*) to obtain confidence interval expansions for $t(\omega)$.

In this paper we consider a real functional $T(F)$ of a number of unknown distributions $F = (F_1 \cdots F_k)$. These are estimated by their empirical distributions $\hat{F} = (\hat{F}_1 \cdots \hat{F}_k)$ based on independent random samples of sizes $n_1, \ldots, n_k$. By identifying $(\omega, \hat{\omega})$ with $(F, \hat{F})$ it is shown in Section 3 how to obtain from the previous parametric results cumulant, Edgeworth and Cornish-Fisher expansions for $T(\hat{F})$ in both standardized and Studentized forms as well as confidence interval expansions for $T(F)$. Putting $k = 1$ yields the one-sample results of Withers (1983*a*).

The one-sided intervals given are j-th order in the sense that the difference between their true level and nominal level has magnitude $n^{-j/2}$ where $j>0$ is any given integer and n is the sample size or minimum sample size. An alternative second order confidence interval is that obtainable by bootstrap methods (see Hall (1986, 1988)).

Section 4 gives as examples confidence intervals for linear combinations and ratios of the means and variances of a number of populations. Regularity conditions are discussed in Section 5.

The results are given in terms of the functional partial derivatives of $T(F)$, defined in Section 2. Example 2.1 allows these results to be stated in terms of ordinary partial derivatives when $T(F)$ has the form $H\left(\int f_1 dF_1, \ldots, \int f_k dF_k\right)$, where $H, f_1, \ldots, f_k$ are known functions. (All of the examples are of this form.) Section 2 also defines notation used in the other sections.

2. Functional partial derivatives and notation

Let $\mathscr{F}_s$ denote the space of distribution functions on R^s. Let x, y, $x_1, \ldots, x_r$ be points in R^s, $F \in \mathscr{F}_s$ and T: $\mathscr{F}_s \to R$. In Withers (1983*a*) the r-th order functional derivative of $T(F)$ at $(x_1 \cdots x_r)$, $T_{x_1 \cdots x_r} = T_F^{(r)}(x_1, \ldots, x_r)$, was defined. It is characterized by the formal functional Taylor series expansion: for $G \in \mathscr{F}_s$,

$$(2.1) \qquad T(G) - T(F) \approx \sum_{r=1}^{\infty} \int^r T_F^{(r)}(x_1, \ldots, x_r) \prod_{j=1}^{r} d(G(x_j) - F(x_j))/r! ,$$

where $\int^r$ denote r integral signs, and the constraints

$$(2.2) \qquad T_{x_1 \cdots x_r} \text{ is symmetric in its } r \text{ arguments} ,$$

and

$$(2.3) \qquad \int T_{x_1 \cdots x_r} \, dF(x_1) = 0 .$$

These imply $F(x_j)$ in (2.1) can be replaced by zero.

In particular, it was shown that for $0 \le \varepsilon \le 1$

$$(2.4) \qquad T_x = \partial T(F + \varepsilon(\delta_x - F))/\partial\varepsilon \quad \text{at} \quad \varepsilon = 0 ,$$

where δ_x is the distribution function putting mass 1 at x—that is $\delta_x(y) = 1\ (x \le y) = 1$ if $x_i \le y_i$ for $1 \le i \le s$ and 0 otherwise. For example, $T(F) = F(y)$ has first derivative

$$(2.5) \qquad T_x = T_F^{(1)}(x) = \delta_x(y) - F(y) = F(y)_x, \quad \text{say} .$$

Also, $T_{x_1 \cdots x_r} = 0$ if $T(F)$ is a 'polynomial in F' of degree less than r (for example a moment or cumulant of F of order less than r), so that the Taylor series in (2.1) consists of only $r-1$ terms. ($T(F)$ is a 'polynomial in F of degree m' if for any G in $\mathscr{F}_s$, $T(F+\varepsilon(G-F))$ is a polynomial in ε of degree m.)

Suppose now that $F=(F_1,\ldots,F_k)$ consists of k distributions on $R^{s_1},\ldots,R^{s_k}$, and that $T(F)$ is a real functional of F.

Then the *functional partial derivative* of $T(F)$ at $\begin{pmatrix} a_1 & \cdots & a_r \\ x_1 & \cdots & x_r \end{pmatrix}$,

$$T^{a_1 \cdots a_r}_{x_1 \cdots x_r} = T_F \begin{pmatrix} a_1 & \cdots & a_r \\ x_1 & \cdots & x_r \end{pmatrix}, \qquad (x_i \in R^{s_{a_i}}, \quad a_i \in \{1, 2, \ldots, k\}),$$

is obtained by treating the lower order functional partial derivatives and $T(F)$ as functionals of F_a alone for $a = a_1, \ldots, a_r$. For example, $T^{a \cdots a}_{x_1 \cdots x_r}$ is the ordinary functional derivative of $S(F_a) = T(F)$ at $(x_1 \cdots x_r)$, and $T^{a \cdots ab \cdots b}_{x_1 \cdots x_r y_1 \cdots y_s}$ is the ordinary functional derivative of $S(F_b) = T^{a \cdots a}_{x_1 \cdots x_r}$ at $(y_1 \cdots y_s)$.

Just as $\partial^2 f(x,y)/\partial x \partial y = \partial^2 f(x,y)/\partial y \partial x$ under mild conditions, swapping columns of $T^{a_1 \cdots a_r}_{x_1 \cdots x_r}$ (for example ${}^{a_1}_{x_1}$ and ${}^{a_2}_{x_2}$) will not alter its value.

The partial derivatives may also be characterized by the formal functional Taylor series expansion: for $G = (G_1, \ldots, G_k) \in \mathscr{F}_{s_1} \times \cdots \times \mathscr{F}_{s_k}$,

$$\text{(2.6)} \quad T(G) - T(F) \approx \sum_{r=1}^{\infty} \int^r T_F \begin{pmatrix} a_1 & \cdots & a_r \\ x_1 & \cdots & x_r \end{pmatrix} \prod_{j=1}^{r} d(G_{a_j}(x_j) - F_{a_j}(x_j))/r!\,,$$

with summation of the repeated subscripts $a_1 \cdots a_r$ over their range $1 \cdots p$ implicit, together with the constraints

$$\text{(2.7)} \qquad T^{a_1 \cdots a_r}_{x_1 \cdots x_r} \text{ is not altered by swapping columns},$$

and

$$\text{(2.8)} \qquad \int T^{a_1 \cdots a_r}_{x_1 \cdots x_r}\, dF_{a_1}(x_1) = 0\,.$$

These imply $F_{a_j}(x_j)$ in (2.6) can be replaced by zero.

The partial derivatives may be calculated using

$$\text{(2.9a)} \qquad T^a_x = S_x \quad \text{for} \quad S(F_a) = T(F),$$

and

$$\text{(2.9b)} \qquad T^{a_1 \cdots a_{r+1}}_{x_1 \cdots x_{r+1}} = (T^{a_1 \cdots a_r}_{x_1 \cdots x_r})^{a_{r+1}}_{x_{r+1}} + \sum_{i=l}^{r} \delta_{a_i a_{r+1}} T \left\langle \begin{matrix} a_1 \cdots a_{r+1} \\ x_1 \cdots x_{r+1} \end{matrix} \right\rangle_i,$$

where $\delta_{ij} = 1$ or 0 for $i = j$ or $i \neq j$ and $\langle\,\rangle_i$ means 'drop the i-th column'. For example $T^{ab}_{xy} = (T^a_x)^b_y + \delta_{ab} T^b_y$. The proof of (2.9) is as for (2.6) of Withers (1983*a*).

The following notation is used: $n = \min_1^k n_i$, $\nu_a = n/n_a$,

$$[1, 12^3]_{ab} = \iint T^a_{x_1}(T^{ab}_{x_1x_2})^3 dF_a(x_1) dF_b(x_2) ,$$

$$[1, 2, 3^2, 123]_{abc} = \iiint T^a_{x_1} T^b_{x_2} (T^c_{x_3})^2 T^{abc}_{x_1x_2x_3} dF_a(x_1) dF_b(x_2) dF_c(x_3) ,$$

and so forth, and for S a number of sequences from and including 1, 2,…, r,

$$[S] = \sum_{a_1=1}^k \nu_{a_1}^{\lambda_1-1} \cdots \sum_{a_r=1}^k \nu_{a_r}^{\lambda_r-1} [S]_{a_1\cdots a_r} , \tag{2.10}$$

where λ_i is the number of times a_i occurs in S. For example,

$$[1^4] = \sum_{a=1}^k \nu_a^3 [1^4]_a ,$$

and

$$[12, 1233] = \sum_{a_1a_2a_3} \nu_{a_1}\nu_{a_2}\nu_{a_3} [12, 1233]_{a_1a_2a_3} .$$

Also we set

$$\langle 1^2 \rangle = \sum_1^k \nu_a^2 [1^2]_a , \tag{2.11}$$

and

$$\langle 1^2, 1^2 \rangle = \sum_1^k \nu_a^3 \left([1^2]_a \right)^2 . \tag{2.12}$$

Example 2.1. Let $f = (f_1,\ldots,f_k)$ be given functions from $R^s = (R^{s_1},\ldots, R^{s_k})$ to $R^t = (R^{t_1},\ldots, R^{t_k})$, and let H: $R^t \to R$ be a given function. Suppose $T(F) = H(\mu(F))$, where $\mu(F) = (\mu_1(F_1),\ldots,\mu_k(F_k))$ and $\mu_a(F_a) = (\mu_{a1}(F_a),\ldots,\mu_{as_a}(F_a)) = \int f_a\, dF_a$. For $1 \leq p_i \leq s_{a_i}$ set

$$H\begin{bmatrix} a_1 & \cdots & a_r \\ p_1 & \cdots & p_r \end{bmatrix} = \partial^r H(\mu)/(\partial\mu_{a_1p_1}\cdots\partial\mu_{a_rp_r}) \quad \text{at} \quad \mu = \mu(F) .$$

Then $\mu_{ap}(F_a)$ has first derivative

$$\mu_{apx_a} = f_{ap}(x_a) - \int f_{ap} dF_a \,, \tag{2.13}$$

for $1 \leq a \leq k$, $1 \leq p \leq s_a$, $x_a \in R^{s_a}$. Also

$$T^{a_1 \cdots a_r}_{x_1 \cdots x_r} = \Sigma \, H\begin{bmatrix} a_1 & \cdots & a_r \\ p_1 & \cdots & p_r \end{bmatrix} \mu_{a_1 p_1 x_1} \cdots \mu_{a_r p_r x_r} \,, \tag{2.14}$$

summed over $\{1 \leq p_1 \leq s_{a_1}, \ldots, 1 \leq p_k \leq s_{a_k}\}$, for $\{1 \leq a_i \leq k,\ x_i \in R^{s_{a_i}},\ 1 \leq i \leq r\}$. (This follows by induction from (3.15) of Withers (1983*a*).)

Hence setting $\mu_a^{p_1 \cdots p_r} = \int \mu_{ap_1x} \cdots \mu_{ap_rx} dF_a(x)$, we have

$$[11]_a = H\begin{bmatrix} a & a \\ p_1 & p_2 \end{bmatrix} \mu_a^{p_1 p_2} \,,$$

$$[1, 23, 123]_{abc} = H\begin{bmatrix} a \\ p_1 \end{bmatrix} H\begin{bmatrix} b & c \\ q_2 & q_3 \end{bmatrix} H\begin{bmatrix} a & b & c \\ r_1 & r_2 & r_3 \end{bmatrix} \mu_a^{p_1 r_1} \mu_b^{q_1 r_2} \mu_c^{q_3 r_3} \,,$$

$$[1, 12^2]_{ab} = H\begin{bmatrix} a \\ p_1 \end{bmatrix} H\begin{bmatrix} a & b \\ q_1 & q_2 \end{bmatrix} H\begin{bmatrix} a & b \\ r_1 & r_2 \end{bmatrix} \mu_a^{p_1 q_1 r_1} \mu_b^{q_2 r_2} \,,$$

and so forth, where summation over repeated suffixes (except for *abc*) is implicit.

3. The asymptotic expansions

Let $F_1, \ldots, F_k$ be unknown distribution functions on $R^{s_1}, \ldots, R^{s_k}$. In this section we show how to obtain nonparametric results from parametric results. This is applied to obtain cumulant, Edgeworth and Cornish-Fisher expansions for $T(\hat{F})$ in its standardized and Studentized forms where T is a real functional and $\hat{F} = (\hat{F}_1, \ldots, \hat{F}_k)$ is a set of empirical distributions from $F = (F_1, \ldots, F_k)$ based on random samples of sizes $n_1, \ldots, n_k$. Also given are expansions for confidence limits for $T(F)$.

In Withers (1982*a*, 1983*b*) we considered an estimate $\hat{\omega}$ of a parameter ω in R^p with cross-cumulants expandable in the form

$$\kappa(\hat{\omega}_{i_1}, \ldots, \hat{\omega}_{i_r}) \approx \sum_{j=r-1}^{\infty} k_j^{i_1 \cdots i_r} n^{-j}, \qquad r \geq 1 \,,$$

where n is known. We also assumed that $E\hat{\omega} \to \omega$ as $n \to \infty$. For $t\colon R^p \to R$ a function with partial derivatives

$$t_{i_1\cdots i_r} = \partial^r t(\omega)/(\partial\omega_{i_1}\cdots\partial\omega_{i_r}) ,$$

we obtained expansions for the cumulants of $t(\hat{\omega})$ of the form

$$\kappa_r(t(\hat{\omega})) \approx \sum_{r-1}^{\infty} a_{rj} n^{-j}, \qquad r \geq 1 ,$$

where $\{a_{rj}\}$, the cumulant coefficients of $t(\hat{\omega})$, are certain functions of the cumulant coefficients of $\hat{\omega}$ and the derivatives of $t(\omega)$—for instance, $a_{11} = I_1\binom{1}{0} + I_{01}\binom{2}{0}\Big/2$, where $I_1\binom{1}{0} = \sum_{i=1}^{p} t_i k_1^i$, and $I_{01}\binom{2}{0} = \sum_{ij} t_{ij} k_1^{ij}$.

Set $n = \max n_a$ and $v_a = n/n_a$. Then $\hat{F}$ has cumulants

$$\kappa(\hat{F}_{a_1}(x_1),\ldots,\hat{F}_{a_r}(x_r)) = \begin{cases} n_a^{1-r}\kappa_a = n^{1-r}v_a^{r-1}\kappa_a & \text{if} \quad a_1 = \cdots = a_r = a \quad \text{say} , \\ 0, & \text{otherwise,} \end{cases}$$

where $\kappa_a = \kappa(1(X_a \leq x_1),\ldots, 1(X_a \leq x_r)) = \kappa(x_1\cdots x_r F_a)$ say, and X_a is a random variable from F_a.

Hence we may identify ω with F and $\hat{\omega}$ with $\hat{F}$, to obtain

$$\kappa_r(T(\hat{F})) \approx \sum_{r-1}^{\infty} a_{rj} n^{-j}, \qquad r \geq 1 , \tag{3.1}$$

with $a_{10},\ldots, a_{43}$ given by p. 59 of Withers (1982*a*) in terms of $t(\omega) = T(F)$ and $I_2\binom{2}{0},\ldots, I_{301}\binom{222}{000}$ appropriately interpreted. This is done by setting $k_j^{i_1\cdots i_r} = 0$ if $j \neq r-1$ and replacing Σ by $\int$, $t_{ij\cdots}$ by $T_{x_i x_j\cdots}^{a_i a_j\cdots}$ and $k_{r-1}^{i_1\cdots i_r}$ by $\sum_{a_1=1}^{k} v_{a_1}^{r-1} 1(a_1 = \cdots = a_r) d\kappa(x_1\cdots x_r F_{a_1})$.

Thus one obtains $I_1\binom{1}{0} = 0$ and

$$\begin{aligned} I_{01}\binom{2}{0} &= \sum_{a_1 a_2} \iint T_{x_1 x_2}^{a_1 a_2} v_{a_1}^{2-1} 1(a_1 = a_2) d\kappa(x_1 x_2 F_{a_1}) \\ &= \sum_{a=1}^{k} v_a \iint T_{x_1 x_2}^{a\,a} d\kappa(x_1 x_2 F_a) . \end{aligned}$$

For X a random variable with distribution G on R^s, *the cumulant function* $\kappa_{x_1\cdots x_r} = \kappa(x_1\cdots x_r G)$ may be expressed in terms of the *central moment functions*

$$\mu_{x_1\cdots x_r} = \mu(x_1\cdots x_r G) = E \prod_{j=1}^{r} (1(X \leq x_j) - G(x_j)) , \tag{3.2}$$

according to the usual rule for expressing cross-cumulants in terms of cross-moments, namely (II.b) of Leonov and Shiryaev (1959). In particular,

$$\kappa_{x_1x_2} = \mu_{x_1x_2} = G(x_1 \wedge x_2) - G(x_1)G(x_2)\,,$$

$$\kappa_{x_1x_2x_3} = \mu_{x_1x_2x_3} = G(x_1 \wedge x_2 \wedge x_3) - \overset{3}{\Sigma}\, G(x_1 \wedge x_2)G(x_3) + 2\prod_1^3 G(x_i)\,,$$

and

$$\kappa_{x_1\cdots x_4} = \mu_{x_1\cdots x_4} - \overset{3}{\Sigma}\, \mu_{x_1x_2}\mu_{x_3x_4}\,,$$

where $x_1 \wedge \cdots \wedge x_r = \left(\min_{i=1}^{r} x_{i1},\ldots, \min_{i=1}^{r} x_{is}\right)$ for $x_1,\ldots,x_r$ in R^s,

$$\overset{3}{\Sigma} f_{123} = f_{123} + f_{231} + f_{312}\,,$$

and

$$\overset{3}{\Sigma} f_{12}f_{34} = f_{12}f_{34} + f_{13}f_{24} + f_{14}f_{23}\,.$$

Because of (2.8) in calculating expressions such as $\iint T^{a\,a}_{x_1x_2}d\kappa(x_1x_2F_a)$ above, we may replace $\mu_{x_1\cdots x_r}$ by $G(x_1 \wedge \cdots \wedge x_r)$. Thus $I_{01}\binom{2}{0} = \Sigma\, \nu_a \cdot \iint T^{aa}_{xy}dF_a(x \wedge y) = \Sigma\, \nu_a \iint T^{aa}_{xx}dF_a(x) = [11]$. In this way one obtains $a_{10} = T(F)$ and $a_{21},\ldots,a_{43}$ in terms of

(3.3) $I_2\binom{2}{0} = [1^2]$, $\quad I_1\binom{1}{0} = 0\,,$

$I_{01}\binom{2}{0} = [11]$, $\quad I_3\binom{3}{0} = [1^3]\,,$

$I_{21}\binom{22}{00} = [1,\, 12,\, 2]$, $\quad \bar{I}_2\binom{2}{0} = 0\,,$

$I_{11}\binom{3}{0} = [1,\, 11]$, $\quad I_{02}\binom{22}{00} = [12^2]\,,$

$I_{101}\binom{22}{00} = [1,\, 122]$, $\quad I_4\binom{4}{0} = [1^4] - 3\langle 1^2, 1^2\rangle\,,$

$I_{31}\binom{23}{00} = [1,\, 12,\, 2^2]$, $\quad I_{22}\binom{222}{000} = [1,\, 2,\, 23,\, 31]$ and

$$I_{301}\binom{222}{000} = [1, 2, 3, 123]\,,$$

and hence

$$(3.4)\qquad \begin{aligned} &a_{10} = T(F), \quad a_{21} = [1^2], \quad a_{11} = [11]/2\,,\\ &a_{32} = [1^3] + 3[1, 12, 2]\,,\\ &a_{22} = [1, 11] + [12^2]/2 + [1, 122] \quad \text{and}\\ &a_{43} = [1^4] - 3\langle 1^2, 1^2\rangle + 12[1, 12, 2^2] + 12[1, 2, 23, 31]\\ &\qquad + 4[1, 2, 3, 123]\,. \end{aligned}$$

In particular, $n^{1/2}(T(\hat{F}) - T(F))$ has asymptotic variance

$$\sigma(F)^2 = a_{21} = [1^2]\,.$$

Hence the standardized form of $T(\hat{F})$, $Y_{n1} = n^{1/2}(T(\hat{F}) - T(F))/\sigma(F)$ satisfies (with $Y_n = Y_{n1}$)

$$(3.5)\qquad \kappa_r(Y_n) \approx n^{r/2} \sum_{r-1}^{\infty} A_{ri} n^{-i}\,,$$

with $A_{10} = 0$, $A_{21} = 1$ and $A_{ri} = a_{21}^{-r/2} a_{ri}$ for $(ri) \neq (10)$.

In particular, setting $[11](F) = [11]$, $T(\hat{F})$ as an estimator of $T(F)$ has bias $[11]n^{-1}/2 + O(n^{-2})$, while

$$(3.6)\qquad T(\hat{F}) - [11](\hat{F})n^{-1}/2\,,$$

the 'infinitesimal jackknife estimator', has bias $O(n^{-2})$.

Now (3.5) implies that the distribution of Y_n has Edgeworth and Cornish-Fisher expansions in powers of $n^{-1/2}$ given by (1.5) of Withers (1984). In particular,

$$(3.7)\qquad P(Y_n \leq x) \approx \Phi(x) - \phi(x) \sum_{1}^{\infty} h_r(x, F) n^{-r/2}\,,$$

where $h_1(\cdot, F) = A_{11} + A_{32} He_2$,

$$h_2(\cdot, F) = (A_{22} + A_{11}^2)He_1/2 + (A_{43} + 4A_{11}A_{32})He_3/24 + A_{32}^2 He_5/72\,,$$

He_r is the r-th Hermite polynomial, and Φ and ϕ are the distribution and density of a unit normal random variable.

Now consider the Studentized statistic

$$T_0(\hat{F}) = (T(\hat{F}) - T(F))/\sigma(\hat{F}) \,. \tag{3.8}$$

This can be identified with $t_0(\hat{\omega})$ of Withers (1982*a*), the Studentized form of $t(\hat{\omega})$. It follows that

$$\kappa_r(T_0(\hat{F})) \approx \sum_{r-1}^{\infty} (a_{ri})_r n^{-i}, \qquad r \geq 1 \,, \tag{3.9}$$

where $(a_{10})_1 \cdots (a_{43})_4$ are given on p. 60 of Withers (1982*a*) in terms of the expressions given by (3.3) and

$$I_3\binom{22}{01} = [1^3], \quad I_2\binom{12}{01} = 0, \quad I_2\binom{22}{02} = -2\langle 1^2 \rangle \,, \tag{3.10}$$

$$I_{11}\binom{22}{01} = [1, 11], \quad I_4\binom{23}{10} = I_4\binom{222}{101} = [1^4] - \langle 1^2, 1^2 \rangle \,,$$

$$I_4\binom{222}{020} = -2\langle 1^2, 1^2 \rangle \quad \text{and}$$

$$I_{31}\binom{222}{010} = I_{31}\binom{222}{001} = [1, 12, 2^2] \,.$$

These are obtained as was demonstrated for $I_{01}\binom{2}{0}$. Where a derivative of a cumulant function occurs—as indicated by a 1 or 2 on the bottom row of an $I_{\cdot}\binom{\cdots}{\cdots}$ function—one may no longer replace the corresponding $\mu_{x_1x_2}$ by $G(x_1 \wedge x_2)$. For example, $I_2\binom{22}{02}$ involves $k^{kl}_{1,ij}$, which becomes $dS_F\binom{a_i\, a_j}{x_i x_j}$ where $S(F) = \sum_{a_k} v_{a_k}^{2-1} 1(a_k = a_l)\kappa(x_k x_l F_{a_k})$. Set

$$W(F_a) = \kappa(x_k x_l F_a) = F_a(x_k \wedge x_l) - F_a(x_k)F_a(x_l) \,.$$

Since $V(F_a) = F_a(y)$ has second derivative 0, it follows setting $U(F_a) = F_a(x_k)F_a(x_l)$ that

$$W^{(2)}_{F_a}(x_i, x_j) = -\, U^{(2)}_{F_a}(x_i, x_j) = -\sum_{ij} F_a(x_k)_{x_i} F_a(x_l)_{x_j} \,,$$

defined by (2.5) where $\sum_{ij} f_{ij} = f_{ij} + f_{ji}$. Thus

$$I_2\binom{22}{02} = \int^4 \sum_a v_a d\kappa(x_i x_j F_a) \sum_b v_b dW^{(2)}_{F_a}(x_i, x_j) 1(a = b) T^b_{x_k} T^b_{x_l}$$

$$= - \sum_a v_a^2 \int^4 d\kappa(x_i x_j F_a) \sum_{ij} dF_a(x_k)_{x_i} dF_a(x_l)_{x_j} T^a_{x_k} T^b_{x_l}$$

$$= - 2 \sum v_a^2 \int^2 d\kappa(x_i x_j F_a) T^a_{x_i} T^a_{x_i} \quad \text{since} \quad \int T^a_x dF_a(x)_y = T^a_y$$

$$= - 2 \sum v_a^2 [1^2]_a = - 2\langle 1^2 \rangle \quad \text{(as claimed)} .$$

Hence one obtains $(a_{10})_1 = 0$, $(a_{21})_2 = 1$ and

(3.11)
$$(a_{11})_1 = [1^2]^{-1/2}[11]/2 - [1^2]^{-3/2}\big([1^3]/2 + [1, 12, 2]\big) ,$$

$$(a_{32})_3 = - [1^2]^{-3/2}\big(2[1^3] + 3[1, 12, 2]\big) ,$$

$$\begin{aligned}(a_{22})_2 = {} & [1^2]^{-1}\big(\langle 1^2\rangle - [1, 11] - [12^2]/2\big) \\ & - [1^2]^{-2}\big\{[11]\big([1^3]/2 + [1, 12, 2]\big) - 2\langle 1^2, 1^2\rangle + 4[1, 12, 2^2] \\ & \qquad + 2[1, 2, 23, 31] + 2[1, 2, 3, 123]\big\} \\ & + 7[1^2]^{-3}\big([1^3]/2 + [1, 12, 2]\big)^2 ,\end{aligned}$$

$$\begin{aligned}(a_{43})_4 = {} & 2[1^2]^{-2}\big(6\langle 1^2, 1^2\rangle - [1^4] - 6[1, 2, 23, 31] - 4[1, 2, 3, 123] \\ & \qquad - 12[1, 12, 2^2]\big) \\ & + 6[1^2]^{-3}\big([1^3] + 2[1, 12, 2]\big)\big(2[1^3] + 3[1, 12, 2]\big) .\end{aligned}$$

Since $(a_{10})_1 = 0$ and $(a_{21})_2 = 1$, $Y_{n2} = n^{1/2} T_0(\hat{F})$ satisfies (3.5) with $A_{ri} = (a_{ri})_r$ and so has Edgeworth and Cornish-Fisher expansions given by (1.5) of Withers (1984) in terms of these $\{A_{ri}\}$.

Let $H_r(x, F)$ denote $h_r(x, F)$ of (3.7) for $Y_n = Y_{n2}$, that is with $A_{ri} \equiv (a_{ri})_r$. Thus H_1 and H_2 are specified by (3.7) and (3.11). In particular,

$$H_1(x, F) = [1^2]^{-1/2}[11]/2 - [1^2]^{-3/2}\big\{[1^3](2x^2 + 1)/6 + [1, 12, 2](x^2 + 1)/6\big\} .$$

By (3.7) for $Y_n = Y_{n2}$, confidence intervals (C.I.s) based on the approximation $Y_{n2} \sim \mathcal{N}(0, 1)$ have error ε_n of magnitude $n^{-1/2}$ for one-sided intervals or n^{-1} for a two-sided equal-tailed interval. This error can be reduced to magnitude ε_n^2 by the method of second order inference: applying Withers (1982*b*) (with $\theta = T(F)$, $Y_n(\theta) = - Y_{n2}$, $h_r(x) = - H_r(- x, F)$) one-sided C.I.s for $T(F)$ of level $\Phi(x) + O(n^{-1})$ are given by

(3.12)
$$T(\hat{F}) - \sigma(\hat{F}) n^{-1/2}(x + H_1(x, \hat{F}) n^{-1/2}) < T(F)$$

and

$$T(F) < T(\hat{F}) + \sigma(\hat{F})n^{-1/2}(x - H_1(x, \hat{F})n^{-1/2}), \tag{3.13}$$

a two-sided C.I. of level $2\Phi(x) - 1 + O(n^{-2})$ is given by

$$T(\hat{F}) - \hat{\zeta} < T(F) < T(\hat{F}) + \hat{\zeta}, \tag{3.14}$$

where $\hat{\zeta} = \sigma(\hat{F})n^{-1/2}(x + H_2(x, \hat{F})n^{-1})$, and the p-value of the hypothesis H_0: $T(F) = T_0$, is given by

$$\Phi(Y_{n0} + H_1(Y_{n0}, \hat{F})n^{-1/2}) + O_p(n^{-1}), \tag{3.15}$$

where $Y_{n0} = n^{1/2}(T_0 - T(\hat{F}))/\sigma(\hat{F})$.

Cumulant, Edgeworth and Cornish-Fisher expansions for

$$T_n(\hat{F}) \approx \sum_0^\infty T_{(i)}(\hat{F})n^{-i/2}$$

may be obtained similarly from Section 4 of Withers (1983*a*). Likewise, C.I.s for $T(F)$ may be obtained from Sections 1 and 2 of Withers (1983*a*) by inserting the expressions given in (3.3), (3.10) and

$$\begin{aligned} &I_2\begin{pmatrix}12\\10\end{pmatrix} = 0, \quad I_4\begin{pmatrix}222\\011\end{pmatrix} = [1^4] - \langle 1^2, 1^2\rangle, \\ &I_4\begin{pmatrix}23\\01\end{pmatrix} = [1^4] - 3\langle 1^2, 1^2\rangle, \\ &I_{11}\begin{pmatrix}12\\00\end{pmatrix} = 0 \quad \text{and} \quad \bar{I}_{11}\begin{pmatrix}22\\01\end{pmatrix} = [1, 11]. \end{aligned} \tag{3.16}$$

Thus for $j \geq 1$ as $n \to \infty$ one-sided C.I.s of level $\Phi(x) + O(n^{-j/2})$ are given by

$$V_{jn}(\hat{F}, x) \leq T(F), \tag{3.17}$$

and

$$T(F) \leq V_{jn}(\hat{F}, -x), \tag{3.18}$$

while a two-sided C.I. of level $2\Phi(x) - 1 + O(n^{-J/2})$, ($J = j + 1$ for j odd, $J = j$ for j even, $x > 0$) is given by

$$V_{jn}(\hat{F}, x) \leq T(F) \leq V_{jn}(\hat{F}, -x), \tag{3.19}$$

where

$$(3.20)\quad V_{jn}(F,x) = T(F) + \sum_{r=1}^{j} q_r(F,x)n^{-r/2},$$

$$q_1(F,x) = -[1^2]^{1/2}x,$$

$$q_2(F,x) = -[11]/2 + [1^2]^{-1}\big\{[1^3](2x^2+1)/6 + [1,\,12,\,2](x^2+1)/2\big\},$$

and

$$\begin{aligned} q_3(F,x) = {} & [1^2]^{-1/2}\big\{-\langle 1^2\rangle/2 + [1,\,11] + [12^2]/4 + [1,\,122]/2\big\}x \\ & + [1^2]^{-3/2}\big\{\langle 1^2,\,1^2\rangle(x^3/2 + x) - [1^4](3x^3+5x)/12 \\ & \qquad - [1,\,12,\,2^2](2x^3+5x)/2 - \big(3[1,\,2,\,23,\,31] \\ & \qquad + [1,\,2,\,3,\,123]\big)(x^3+3x)/6\big\} \\ & + [1^2]^{-5/2}\big\{[1^3]^2(16x^3+23x) + 48[1^3][1,\,12,\,2](x^3+2x) \\ & \qquad + 18[1,\,12,\,2]^2(2x^3+5x)\big\}\big/72 . \end{aligned}$$

Note that since $q_2(F,x) = -[1^2]^{1/2}H_1(x,F)$, (3.12) and (3.13) are just (3.17) and (3.18) with $j = 1$. However, (3.14) is a simpler C.I. than (3.19) with $j = 3$, though both have error of magnitude n^{-2}.

In the one sample case $\langle 1^2\rangle = [1^2]$ and $\langle 1^2, 1^2\rangle = [1^2]^2$, so that except for (3.6), (3.14) and (3.15), the results of this section specialize to those of Withers (1983*a*). The errors in these approximate C.I.s satisfy Theorem 3 of Withers (1983*a*) and Theorem 3.2 of Withers (1980) with $[1^2]$ and $q_{j+1}(F,x)$ as redefined here.

The results of this section were initially found by generalizing the technique of Withers (1983*a*). However, the present method is more instructive and probably less prone to error. It may be useful—for example to apply Theorem 3.2 of Withers (1980)—to include the following version of $(a_{21})_{101}$ of Section 5 of Withers (1983*a*):

$$\begin{aligned} (a_{21})_{101} = {} & -[1^2]^{-1}\big([1,\,11] + [1,\,122]\big) \\ & + [1^2]^{-2}\big\{[11]K/2 + \big([1^4]/3 - \langle 1^2,\,1^2\rangle + [1,\,12,\,2^2]\big)(2x^2+1) \\ & \qquad + \big(2[1,\,2,\,23,\,31] + 2[1,\,12,\,2^2] + [1,\,2,\,3,\,123]\big)(x^2+1)\big\} \\ & - [1^2]^{-3}\big\{[1^3](2x^2+1) + 3[1,\,12,\,2](x^2+1)\big\}K/2 , \end{aligned}$$

where $K = [1^3] + 2[1,\,12,\,2]$.

4. Examples

The expansions of Section 3 were given in terms of the $[\cdot]$ and $\langle\cdot\rangle$ functions defined in Section 2. Here we obtain their values for some simple cases. We use $\mu_a = \mu_a(F_a) = \int f_a dF_a$ to denote the mean of the a-th distribution after a given transformation f_a: $R^{s_a} \to R$ has been applied, and $\mu_{aj} = \mu_{aj}(F_a) = \int (f_a - \mu_a)^j dF_a$ to denote the j-th central moment of the transformed distribution. Typically $s_a \equiv 1$ and $f_a(x) \equiv x$; however, log and power transformations are also commonly used.

For S a series of sequences from 1, 2,... set

$$[S](T_{(a)}) = [S]_a = [S]_{aa\cdots}\,,$$

of Section 2 with $(T, F) = (T_{(a)}, F_a)$.

Example 4.1. Linear combinations: $T(F) = \sum_1^k T_{(a)}(F_a)$:

$$\text{since} \qquad T^{ab\cdots}_{xy\cdots} = 1(a = b = \cdots)T^{aa\cdots}_{(a)xy\cdots}\,,$$

we have

$$[S]_{ab\cdots} = 1(a = b = \cdots)[S]_a\,,$$

so that

$$\begin{aligned}
&[1^i, 11^j] = \Sigma\, v_a^{i+2j-1}[1^i, 11^j]_a\,,\\
&[1, 12, 2^j] = \Sigma\, v_a^{j+1}[1, 12, 2^j]_a\,,\\
&[12^2] = \Sigma\, v_a^2[12^2]_a\,,\\
&[1, 122] = \Sigma\, v_a^2[1, 122]_a\,,\\
&[1, 2, 23, 31] = \Sigma\, v_a^3[1, 2, 23, 31]_a\,,\\
&[1, 2, 3, 123] = \Sigma\, v_a^3[1, 2, 3, 123]_a\,,\\
&\langle 1^2\rangle = \Sigma\, v_a^2[1^2]_a \quad \text{and} \quad \langle 1^2, 1^2\rangle = \Sigma\, v_a^3[1^2]_a^2\,.
\end{aligned}$$

Example 4.1.1. A linear combination of means: $T(F) = \Sigma\, c_a\mu_a$: $\mu_a(F_a)$ has first derivative $\mu_{ax} = f_a(x) - \mu_a$ and higher derivatives vanish; the non-zero terms are

$$[1^j] = \Sigma\, v_a^{j-1}c_a^j\mu_{aj}, \quad \langle 1^2\rangle = \Sigma\, v_a^2c_a^2\mu_{a2} \quad \text{and} \quad \langle 1^2, 1^2\rangle = \Sigma\, v_a^3c_a^4\mu_{a2}^2\,.$$

C.I.s are given by (3.12)–(3.14) with

$$H_1(x,F) = -[1^2]^{-3/2}[1^3](2x^2+1)/6 \quad \text{and}$$
$$H_2(x,F) = [1^2]^{-1}\langle 1^2\rangle x/2 + [1^2]^{-2}\{\langle 1^2, 1^2\rangle(x^3-x)/2 - [1^4](x^3-3x)/12\}$$
$$+ [1^2]^{-3}[1^3]^2(x^5+2x^3-3x)/18\,,$$

and by (3.17)–(3.20) with

$$q_1(F,x) = -[1^2]^{1/2}x, \qquad q_2(F,x) = [1^2]^{-1}[1^3](2x^2+1)/6 \quad \text{and}$$
$$q_3(F,x) = -[1^2]^{-1/2}\langle 1^2\rangle x/2 + [1^2]^{-3/2}\{\langle 1^2, 1^2\rangle(x^3/2+x) - [1^4](3x^3+5x)/12\}$$
$$+ [1^2]^{-5/2}[1^3]^2(16x^3+23x)/72\,.$$

If one knew *a priori* that the populations were not skew, i.e., that $[1^3]=0$, then instead of (3.12), (3.13) one would use (2.3) of Withers (1982*b*).

In the case of $T(F)=\mu_1$, the mean of a single population with variance $\sigma^2=\mu_{12}$, these simplify: putting $\lambda_j=\sigma^{-j}\mu_{1j}$ the standardized j-th moment, we have

$$H_1(x,F) = -\sigma^{-1}q_2(F,x) = -\lambda_3(2x^2+1)/6\,,$$
$$H_2(x,F) = x^3/2 - \lambda_4(x^3-3x)/12 + \lambda_3^2(x^5+2x^3-3x)/18\,,$$
$$\sigma^{-1}q_3(F,x) = (x^3+x)/2 - \lambda_4(3x^3+5x)/12 + \lambda_3^2(16x^2+23x)/72\,,$$

while from Section 5 of Withers (1983*a*), $\sigma^{-1}q_4(F,x) = \lambda_3(-19-19x^2+36x^4)/12 + \lambda_5(27+86x^2+24x^4)/120 - \lambda_3\lambda_4(14+55x^2+18x^4)/36 + \lambda_3^3(110+529x^2+192x^4)/648$.

These C.I.s provide nonparametric alternatives to the C.I. derived from the k-sample t-statistic

$$t_f = f^{1/2}(N(\hat F)-N(F))/\sigma(\hat F) \quad \text{where} \quad f=\Sigma\,(n_a-1)\,,$$
$$N(F)=\Sigma\, c_a\mu_a \quad \text{and} \quad \sigma^2(F)=(\Sigma\, c_a^2 n_a^{-1})\,\Sigma\, n_a\mu_{a2}\,.$$

The latter C.I. is only exact for normal populations with equal variances; if the variances are not equal, its α-level is not even asymptotically correct unless

$$(\Sigma\, c_a^2 v_a)\,\Sigma\, v_a^{-1}\mu_{a2} = (\Sigma\, v_a^{-1})\,\Sigma\, c_a^2 v_a\mu_{a2}\,, \tag{4.1}$$

which is true if $|c_a|/n_a$ does not depend on a. (In the general case, an expansion for the distribution of t_f is obtainable by applying Section 3 to

$T(\hat{F}) = f^{-1/2}t_f$; c.f. Subsections 2(b) and (c) of Geary (1947) as corrected on p. 403 of Gayen (1950), and Example 2 of Withers (1983*a*); an alternative expansion when $k = 2$ and cumulants beyond the fourth are zero is given by Tan (1982), but the magnitude of the error of this approximation when cumulants are non-zero is not given.)

Example 4.1.2. A linear combination of variances: $T(F) = \Sigma\, c_a\mu_{a2}$: μ_{a2} has first and second derivatives $\mu_{a2x} = \mu_{ax}^2 - \mu_{a2}$ and $\mu_{a2xy} = -2\mu_{ax}\mu_{ay}$, where $\mu_{ax} = f_a(x) - \mu_a$, while higher derivatives vanish. Hence

$$[1^2] = \Sigma\, \nu_a c_a^2(\mu_4 - \mu_2^2)_a\,,$$

where $(\mu_4 - \mu_2^2)_a = \mu_{a4} - \mu_{a2}^2$. Similarly,

$$\begin{aligned}
&[1^3] = \Sigma\, \nu_a^2 c_a^3(\mu_6 - 3\mu_4\mu_2 + 2\mu_2^3)_a\,,\\
&[1^4] = \Sigma\, \nu_a^3 c_a^4(\mu_8 - 4\mu_6\mu_2 + 6\mu_4\mu_2^2 - 3\mu_2^3)_a\,,\\
&\langle 1^2\rangle = \Sigma\, \nu_a^2 c_a^2(\mu_4 - \mu_2^2)_a\,,\\
&\langle 1^2, 1^2\rangle = \Sigma\, \nu_a^3 c_a^4\{(\mu_4 - \mu_2^2)_a\}^2\,,\\
&[11] = -2\,\Sigma\, \nu_a c_a\mu_{a2}\,,\\
&[1, 12, 2] = -2\,\Sigma\, \nu_a^2 c_a^3\mu_{a3}^2\,,\\
&[1, 11] = -2\,\Sigma\, \nu_a^2 c_a^2(\mu_4 - \mu_2^2)_a\,,\\
&[12^2] = 4\,\Sigma\, \nu_a^2 c_a^2\mu_{a2}^2\,,\\
&[1, 12, 2^2] = -2\,\Sigma\, \nu_a^3 c_a^4(\mu_5\mu_3 - 2\mu_3^2\mu_2)_a\,,\\
&[1, 2, 23, 31] = 4\,\Sigma\, \nu_a^3 c_a^4\mu_{a3}^2\mu_{a2} \quad \text{and}\\
&[1, 122] = [1, 2, 3, 123] = 0\,.
\end{aligned}$$

Example 4.2. Products and quotients: $T(F) = T_{(1)}(F_1)T_{(2)}(F_2) = T_1T_2$, say: $T^{1\cdots 1\,2\cdots 2}_{x_1\cdots x_i\,y_1\cdots y_j} = T_{(1)x_1\cdots x_i}T_{(2)y_1\cdots y_j}$, so that

$$[1^j] = \nu_1^{j-1}[1^j]_1 T_2^j + \nu_2^{j-1}[1^j]_2 T_1^j = \overset{2}{\Sigma}\, \nu_1^{j-1}[1^j]T_2^j\,,$$

where

$$\overset{2}{\Sigma} f(T_{(1)}, T_{(2)}, \nu_1, \nu_2) = f(T_{(1)}, T_{(2)}, \nu_1, \nu_2) + f(T_{(2)}, T_{(1)}, \nu_2, \nu_1)\,,$$

$$[11] = \overset{2}{\Sigma}\nu_1[11]_1T_2, \qquad [1, 11] = \overset{2}{\Sigma}\nu_1^2[1, 11]_1T_2^2\,,$$

$$[1, 12, 2^j] = \overset{2}{\Sigma}\left\{\nu_1^{j+1}[1, 12, 2^j]_1T_2^{j+2} + \nu_1\nu_2^j T_1^j[1^2]_1T_2[1^{j+1}]_2\right\},$$

$$[12^2] = \sum^2 \nu_i^2[12^2]_i T_2^2 + 2\nu_1\nu_2[1^2]_1[1^2]_2 ,$$

$$[1,\ 122] = \sum^2 \left\{\nu_i^2[1,\ 122]_i T_2^2 + \nu_1\nu_2[1^2]_i[11]_2 T_2\right\} ,$$

$$[1,\ 2,\ 23,\ 31] = \sum^2 \left\{\nu_i^3[1,\ 2,\ 23,\ 31]_i T_2^4 + \nu_1^2\nu_2[1^2]_2 T_2^2\left([1^2]_i^2 + 2T_i[1,\ 12,\ 2]_i\right)\right\} ,$$

$$[1,\ 2,\ 3,\ 123] = \sum^2 \left\{\nu_i^3[1,\ 2,\ 3,\ 123]_i T_2^4 + 3\nu_i^2\nu_2 T_1 T_2^2[1,\ 12,\ 2]_i[1^2]_2\right\} ,$$

$$\langle 1^2\rangle = \sum^2 \nu_i^2[1^2]_i T_2^2 \quad \text{and} \quad \langle 1^2, 1^2\rangle = \sum^2 \nu_i^3[1^2]_i^2 T_2^4 .$$

The coefficients for a quotient may be obtained by applying this to

$$T_2 = T_{(2)}(F_2) = T_{(3)}(F_2)^{-1} = T_3^{-1}, \quad \text{say:}$$

from

$$T_{2x} = -\ T_3^{-2}T_{3x}, \qquad T_{2xy} = -\ T_3^{-2}T_{3xy} + 2T_3^{-3}T_{3x}T_{3y} ,$$

and

$$T_{2xyz} = -\ T_3^{-2}T_{3xyz} + 2T_3^{-3}\sum^3 T_{3x}T_{3yz} - 6T_3^{-4}T_{3x}T_{3y}T_{3z} ,$$

where

$$\sum^3 f_{xyz} = f_{xyz} + f_{yzx} + f_{zxy} ,$$

we have

$$[1^j]_2 = \left(-\ T_3^{-2}\right)^j[1^j]_3, \qquad [11]_2 = -\ T_3^{-2}[11]_3 + 2T_3^{-3}[1^2]_3 ,$$

$$[1,\ 11]_2 = T_3^{-4}[1,\ 11]_3 - 2T_3^{-5}[1^3]_3 ,$$

$$[1,\ 12,\ 2^j]_2 = \left(-\ T_3^{-2}\right)^{j+2}\left\{[1,\ 12,\ 2^j]_3 - 2T_3^{-1}[1^2]_3[1^{j+1}]_3\right\} ,$$

$$[12^2]_2 = T_3^{-4}[12^2]_3 - 4T_3^{-5}[1,\ 12,\ 2]_3 + 4T_3^{-6}[1^2]_3^2 ,$$

$$[1,\ 122]_2 = T_3^{-4}[1,\ 122]_3 - 2T_3^{-5}\left\{[1^2]_3[11]_3 + 2[1,\ 12,\ 1]_3\right\} + 6T_3^{-6}[1^2]_3^2 ,$$

$$[1,\ 2,\ 23,\ 31]_2 = T_3^{-8}[1,\ 2,\ 23,\ 31]_3 - 4T_3^{-9}[1^2]_3[1,\ 12,\ 1]_3 + 4T_3^{-10}[1^2]_3^3 ,$$

and

$$[1,\ 2,\ 3,\ 123]_2 = T_3^{-8}[1,\ 2,\ 3,\ 123]_3 - 6T_3^{-9}[1^2]_3[1,\ 12,\ 1]_3 + 6T_3^{-10}[1^2]_3^3 .$$

Example 4.2.1. The ratio of two means: $T(F) = \mu_1/\mu_2 = T$, say:

setting $\overline{G}_1 = \nu_1\mu_2^{-2}\mu_{12}$, $\overline{G}_2 = \nu_2\mu_2^{-2}\mu_{22}T^2$, $\overline{H}_1 = \nu_1^2\mu_2^{-3}\mu_{13}$, $\overline{H}_2 = \nu_2^2\mu_2^{-3}\mu_{23}T^3$, the coefficients needed for H_1 and q_2 are

$$[1^2] = \overline{G}_1 + \overline{G}_2, \quad [11] = 2\overline{G}_2T^{-1},$$
$$[1^3] = \overline{H}_1 - \overline{H}_2 \quad \text{and} \quad [1, 12, 2] = [1^2][11];$$

the coefficients needed for H_2 and q_3 are

$$\langle 1^2\rangle = \nu_1\overline{G}_1 - \nu_2\overline{G}_2, \quad \langle 1^2, 1^2\rangle = \nu_1\overline{G}_1^2 + \nu_2\overline{G}_2^2, \quad [1, 11] = -2\overline{H}_2T^{-1},$$
$$[1^4] = \nu_1^3\mu_2^{-4}\mu_{14} + \nu_2^3\mu_2^{-4}\mu_{24}T^4, \quad [12^2] = [1, 12, 2]T^{-1},$$
$$[1, 12, 2^2] = -\overline{H}_2(\overline{G}_1 + 2\overline{G}_2)T^{-1}, \quad [1, 2, 23, 31] = \overline{G}_2(\overline{G}_1 + 4\overline{G}_2)[1^2]T^{-2},$$
$$[1, 2, 3, 123] = 6\overline{G}_2^2[1^2]T^{-2} \quad \text{and} \quad [1, 122] = 2\overline{G}_2(\overline{G}_1 + 3\overline{G}_2)T^{-2}.$$

In particular,

$$(a_{11})_1 = -[1^2]^{-1/2}\overline{G}_2T^{-1} - [1^2]^{-3/2}[1^3]/2,$$
$$(a_{32})_3 = -6[1^2]^{-1/2}\overline{G}_2T^{-1} - 2[1^2]^{-3/2}[1^3],$$
$$(a_{22})_2 = -3\overline{G}_2T^{-2} + [1^2]^{-1}\left(\langle 1^2\rangle + 6\overline{H}_2T^{-1} + 6\overline{G}_2^2T^{-2}\right)$$
$$+ [1^2]^{-2}\left\{\overline{G}_2T^{-1}(15\overline{H}_1 - 11\overline{H}_2) + 2\langle 1^2, 1^2\rangle\right\} + 7[1^2]^{-3}[1^3]/4,$$
$$(a_{43})_4 = -12\overline{G}_2T^{-2} + 12[1^2]^{-1}(7\overline{G}_2^2T^{-2} + 2\overline{H}_2T^{-1})$$
$$+ 12[1^2]^{-2}\left\{\langle 1^2, 1^2\rangle - [1^4]/6 - 3\overline{G}_2T^{-2} + \overline{G}_2T^{-1}(7\overline{H}_1 - 5\overline{H}_2)\right\}$$
$$+ 12[1^2]^{-3}[1^3]^2,$$

and

$$H_1(x, F) = -[1^2]^{-1/2}\overline{G}_2T^{-1}x^2 - [1^2]^{-3/2}[1^3](2x^2 + 1)/6.$$

Note that if it is known that $\mu_2 \neq 0$, then the test of $\mu_1 = \mu_2$ derived from the confidence interval for μ_1/μ_2 provides an alternative to the test of $\mu_1 = \mu_2$ derived from the confidence interval for $\mu_1 - \mu_2$ given in Example 4.1.1; a third alternative for a test that $\mu_1 = \mu_2$ is that derived from the confidence interval for μ_2/μ_1. Their AREs are one. A similar remark applies to the tests of $\sigma_1^2 = \sigma_2^2$ derivable from Example 4.1.2 and the next example.

Example 4.2.2. The ratio of two variances: $T(F) = \mu_{12}/\mu_{22}$: the required coefficients are given by the first part of Example 4.2 in terms of $T_1 = \mu_{12}$, $T_2 = \mu_{22}^{-1}$,

$$[1^2]_1 = \mu_{14} - \mu_{12}^2, \quad [1^3]_1 = \mu_{16} - 3\mu_{14}\mu_{12} + 2\mu_{12}^2 ,$$

$$[1^4]_1 = \mu_{18} - 4\mu_{16}\mu_{12} + 6\mu_{14}\mu_{12}^2 - 3\mu_{12}^3 ,$$

$$[11]_1 = -2\mu_{12}, \quad [1, 12, 2]_1 = -2\mu_{13}^2, \quad [1, 11]_1 = -2(\mu_{14} - \mu_{12}^2) ,$$

$$[12^2]_1 = 4\mu_{12}^2, \quad [1, 12, 2^2]_1 = -2(\mu_{15}\mu_{13} - 2\mu_{13}^2\mu_{12}) ,$$

$$[1, 2, 23, 31]_1 = 4\mu_{13}^2\mu_{12}, \quad [1, 122]_1 = [1, 2, 3, 123]_1 = 0 ,$$

and the second part of Example 4.2 with $T_3 = \mu_2$ and $[\cdot]_3 = [\cdot]_1$ with $\{\mu_{1i}\}$ replaced by $\{\mu_{2i}\}$. In particular, $T(\hat{F})$ has asymptotic variance

$$\begin{aligned}[1^2]/n &= (\mu_{12}/\mu_{22})^2 \sum_1^2 (\mu_{a4}/\mu_{a2}^2 - 1)/n_a \\ &= 2(n_1^{-1} + n_2^{-1})T(F)^2 \qquad \text{for } (F_1, F_2) \text{ normal} .\end{aligned}$$

The associated confidence interval for a variance ratio and the associated test for equality of variances are important as they provide alternatives to the confidence interval and test based on the usual assumption that

$$n_1(n_2 - 1)n_2^{-1}(n_1 - 1)^{-1}T(\hat{F})/T(F) \sim F_{n_1-1,n_2-1} .$$

The latter are not even consistent unless $\mu_{a4}/\mu_{a2}^2 \equiv 3$, as for normal populations. In short, procedures based on the F-distribution are notoriously inconsistent for non-normal populations. The 'correction terms' $q_2(F, x)$, etc., are NOT zero even for normal populations: in fact, if (F_1, F_2) are normal, then

$$q_2(F, x) = \{5(\nu_1 - \nu_2) + 2x^2(2\nu_1 + \nu_2)\}T(F)/3 ,$$

while $T(\hat{F})$ has bias $\approx a_{11}n^{-1} = (-n_1^{-1} + 3n_2^{-1})T(F)$ (negligible only if $n_2 \approx 3n_1$) and skewness $(\mu_3) \approx a_{32}n^{-2} = 8(n_1^{-1} + n_2^{-1})(n_1^{-1} + 2n^{-1})T(F)^3$.

Note. If c is a one to one increasing function on the range of $T(F)$, an alternative to a C.I. for $q_0 = T(F)$ based on $V_{jn} = \sum_0^j n^{-r/2}q_r$ is to use a C.I. for $q_0^c = c(T(F))$ based on $V_{jn}^c = \sum_0^j n^{-r/2}q_r^c$, where q_r^c is the coefficient of $n^{-r/2}$ in $c(V_{jn})$ for $j \geq r$:

$$q_1^c = c_1q_1, \quad q_2^c = c_1q_2 + c_2q_1^2/2, \quad q_3^c = c_1q_3 + c_2q_1q_2 + c_3q_1^3/6 ,$$

and

$$q_4^c = c_1 q_4 + c_2(q_1 q_3 + q_2^2/2) + c_3 q_1^2 q_2/2 + c_4 q_4/24\ ,$$

where $c_i = c^{(i)}(T(F))$. In particular, the C.I.s for μ_1/μ_2 (if known to be positive) and μ_{12}/μ_{22} derived in this way from Examples 4.2.1 and 4.2.2 with $c(T) = \log T$ are likely to be more robust than the original C.I.s. (These V_{jn}^c could also be derived directly from Example 4.1.)

5. Regularity conditions

For $k = 1$ these are given in Withers (1983*a*), based on the results of Bhattacharya and Ghosh (1978) for $T(F) = H\left(\int f dF\right)$. Extension of their results to k-samples—that is, to

$$(5.1)\qquad \begin{aligned} &F = (F_1, \ldots, F_k)\ , \\ &T(F) = H\left(\int f_1 dF_1, \ldots, \int f_k dF_k\right) = H(\mu_1, \ldots, \mu_k) \quad \text{say}\ , \end{aligned}$$

should be straightforward. We note here an alternative approach which is available for (5.1) when

$$(5.2)\qquad n_a \equiv m_a N \text{ for } m_1, \ldots, m_k \text{ bounded integers and } N \text{ an integer}\ .$$

Set $\hat{\mu}_{a,n_a} = \int f_a d\hat{F}_a$. Then $\hat{\mu}_{a,n_a}$ may be considered as the mean of a random sample of size N, say $\hat{\mu}_{a,n_a} = N^{-1} \sum_{i=1}^{N} \tilde{X}_{ia}$, where $\tilde{X}_{ia}$ has the same distribution as $\hat{\mu}_{a,m_a}$, say $F_{am_a}(x)$, the m_a-fold convolution of $P(f_a(X_a) \leq x)$, where $X_a \sim F_a$.

Thus writing $\mu = (\mu_1, \ldots, \mu_k)$ and $\hat{\mu} = (\hat{\mu}_1, \ldots, \hat{\mu}_k)$ we have

$$\hat{\mu} = N^{-1} \sum_{i=1}^{N} \tilde{X}_a \quad \text{where} \quad \tilde{X}_a = (\tilde{X}_{1a}, \ldots, \tilde{X}_{ka})\ .$$

Hence the results of Bhattacharya and Ghosh (1978) may be applied or the results of Withers (1983*a*), where now

$$(5.3)\qquad F(x) = \prod_{1}^{k} F_{am_a}(x_a)\ ,$$

with empirical distribution $F_N(x) = N^{-1} \sum_{1}^{N} 1(\tilde{X}_a \leq x)$.

REFERENCES

Bhattacharya, R. N. and Ghosh, J. K. (1978). On the validity of the formal Edgeworth expansion, *Ann. Statist.*, **6**, 434–451.

Gayen, A. K. (1950). Significance of difference between the means of two non-normal samples, *Biometrika*, **37**, 399–408.

Geary, R. C. (1947). Testing for normality, *Biometrika*, **34**, 109–242.

Hall, P. (1986). On the bootstrap and confidence intervals, *Ann. Statist.*, **14**, 1431–1452.

Hall, P. (1988). Theoretical comparisons of bootstrap confidence intervals, *Ann. Statist.*, (to appear).

Leonov, V. P. and Shiryaev, A. N. (1959). On a method for calculating semi-invariants, *Theory Probab. Appl.*, **4**, 319–329.

Tan, W. Y. (1982). Sampling distributions and robustness of t, F and variance-ratio in two samples and ANOVA models with respect to departure from normality, *Comm. Statist. A—Theory Methods*, **11** (22), 2485–2511.

Welch, B. L. (1947). The generalisation of 'Student's' problem when several different population variances are involved, *Biometrika*, **34**, 28–35.

Withers, C. S. (1980). Accurate nonparametric inference—the one sample case, Tech. Report 97, Applied Mathematics Division, DSIR, Wellington, New Zealand.

Withers, C. S. (1982*a*). The distribution and quantiles of a function of parameter estimates, *Ann. Inst. Statist. Math.*, **34**, 55–68.

Withers, C. S. (1982*b*). Second order inference for asymptotically normal random variables, *Sankhyā Ser. B*, **44**, 19–27.

Withers, C. S. (1983*a*). Expansions for the distribution and quantiles of a regular functional of the empirical distribution with applications to nonparametric confidence intervals, *Ann. Statist.*, **11**, 577–587.

Withers, C. S. (1983*b*). Accurate confidence intervals when nuisance parameters are present, (submitted).

Withers, C. S. (1984). Asymptotic expansions for distributions and quantiles with power series cumulants, *J. Roy. Statist. Soc. Ser. B*, **46**, 389–396.