The Annals of Statistics
1978, Vol. 6, No. 5, 1005-1011

AN ASYMPTOTIC EXPANSION FOR SAMPLES FROM A FINITE POPULATION

BY J. ROBINSON

Universities of Sydney and Waterloo

An asymptotic expansion is obtained for the distribution function of the standardized mean of a sample of s observations taken randomly without replacement from a finite population of n numbers. The expansion is given to order $1/n$ and agrees with the formal Edgeworth expansion. The proof of the result is obtained using an approximation to the characteristic function of the standardized sum.

Let $\{a_{ni}\}$ be a trianglar array of real numbers for $i = 1, \cdots, n$, $n = 2, 3, \cdots$ and suppose $\sum_i a_{ni} = 0$, $\sum_i a_{ni}^2 = 1$. Let

$$X_{ns} = \sum_{i=1}^{s} a_{nR_{ni}},$$

where $(R_{n1}, \cdots, R_{nn})$ is a uniform random permutation of $(1, \cdots, n)$. If $p = s/n$ and $q = 1 - p$, then it is easy to show that

$$EX_{ns} = 0, \qquad VX_{ns} = npq/(n-1).$$

Let $Y_{ns} = X_{ns}/(VX_{ns})^{\frac{1}{2}}$ and $F_{ns}(x) = P(Y_{ns} < x)$. Erdös and Rényi (1959) showed that $F_{ns}(x)$ converges to $\Phi(x)$, the distribution function of a standardized normal variate, if $b_n = \max_i |a_{ni}|$ tends to zero and Bikelis (1969) obtained an estimate of the remainder term for this approximation. Von Bahr (1972) considered a related problem, where the a_{ni} are themselves assumed to be random variables. We will obtain an approximation for $F_{ns}(x)$ by the asymptotic expansion

$$\begin{aligned} G_{ns}(x) = \Phi(x) &- H_2(x)\phi(x)\frac{q-p}{6(pq)^{\frac{1}{2}}}\sum_i a_{ni}^3 \\ &- H_3(x)\phi(x)\left[\frac{1-6pq}{24pq}\left(\sum_i a_{ni}^4 - 3n^{-1}\right) - \tfrac{1}{4}n^{-1}\right] \\ &- H_5(x)\phi(x)\frac{(q-p)^2}{72pq}\left(\sum_i a_{ni}^3\right)^2 \end{aligned}$$

where $\phi(x) = \Phi'(x) = (2\pi)^{-\frac{1}{2}}e^{-\frac{1}{2}x^2}$, $H_i(x)\phi(x) = (-1)^i(d^i/dx^i)\phi(x)$. The coefficients of $H_i(x)\phi(x)$ in this expansion differ from the cumulants of Y_{ns} by quantities of order n^{-1}.

Let $A_{rn} = \sum_k |a_{nk}|^r$. We will show that the expansion is a valid approximation accurate to the order of A_{5n}, subject to the condition:

(c) Given $C' > 0$, there exist $\varepsilon > 0$, $C > 0$ and $\delta > 0$ not depending on n,

Received December 1975; revised December 1976.

AMS 1970 *subject classifications.* Primary 60F05; Secondary 62E20.

Key words and phrases. Asymptotic expansions, Edgeworth series, sampling without replacement, permutation tests.

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

such that for any fixed x, the number of indices j, for which $|a_{nj}t - x - 2r\pi| > \varepsilon$, for all $t \in (C'b_n^{-1}, CA_{5n}^{-1})$ and all $r = 0, \pm 1, \pm 2, \cdots$, is greater than δn, for all n.

The condition is similar to that of Albers, Bickel and van Zwet (1976) who gave expansions for sampling with replacement. As with their condition, it ensures that the values of a_{ni} do not cluster around too few values. For example, if a_{ni} are the standardized values of $f(i/n)$, where f is a strictly increasing continuous function on [0, 1], such that $\int_0^1 f^2(x)\,dx < \infty$, then the condition is satisfied. In particular, if $f(x) = x$, Y_{ns} is the two-sample Wilcoxon statistic. In this case the expansion agrees with the first two terms of the expansion of Hodges and Fix (1955) based on a formal Edgeworth expansion. $F_{ns}(x)$ would have jumps of order $n^{-\frac{3}{2}}$ in this case, so further terms could not be used. On the other hand, the condition is not satisfied if the a_{ni} take only two distinct values. In this case $F_{ns}(x)$ has jumps of order $n^{-\frac{1}{2}}$.

When the a_{ni} are random variables, it is necessary to assume that the condition (c) holds except in a set E with probability of order A_{5n}. In particular, if the a_{ni} are assumed to be independent, identically distributed continuous random variables then this condition is satisfied. Under this condition we can obtain an expansion for the conditional distribution of Y_{ns} given the values of the order statistic in this set and the expansion for the marginal distribution is obtained by taking expectations.

The expansion is in a form which may be applied to obtain an approximation to the level of significance of a two-sample permutation test or rank test. It should provide considerably better accuracy than a simple normal approximation and it is quite simple to calculate.

THEOREM. *If condition* (c) *holds, then*

$$|F_{ns}(x) - G_{ns}(x)| < BA_{5n}\,, \tag{1}$$

for all x, where B is a function of p only.

PROOF. The characteristic function of X_{ns} can be put in a form obtained by Erdös and Rényi (1959) as

$$\begin{aligned} f_{ns}^*(u) &= \binom{n}{s}^{-1} \textstyle\sum^* \exp\{iu(a_{ni_1} + \cdots + a_{ni_s})\} \\ &= [2\pi B_{ns}(p)]^{-1} \int_{-\pi}^{\pi} \prod_{k=1}^n [q + pe^{i(ua_{nk}+\theta)}]e^{-i\theta s}\,d\theta\,, \end{aligned}$$

where $\sum^*$ denotes summation over all choices of $i_1, \cdots, i_s$, with $1 \leqq i_1 < i_2 < \cdots < i_s \leqq n$, from $1, \cdots, n$, and

$$B_{ns}(p) = \binom{n}{s}p^s q^{n-s}\,.$$

So the characteristic function of Y_{ns} is

$$\begin{aligned} f_{ns}(t) &= f_{ns}^*[t\{(n-1)/npq\}^{\frac{1}{2}}] \\ &= [(npq)^{\frac{1}{2}}2\pi B_{ns}(p)]^{-1} \int \prod_{k=1}^n \rho_k(\psi, t)\,d\psi\,, \end{aligned}$$

where the integral is over the range $-\pi(npq)^{\frac{1}{2}} < \psi < \pi(npq)^{\frac{1}{2}}$ and

$$\rho_k(\psi, t) = qe^{-ip\xi_{nk}(pq)^{-\frac{1}{2}}} + pe^{iq\xi_{nk}(pq)^{-\frac{1}{2}}}$$

where

$$\xi_{nk} = n^{-\frac{1}{2}}\psi + n^{-\frac{1}{2}}(n-1)^{\frac{1}{2}}ta_{nk}\,.$$

We will use $\theta_1, \theta_2, \cdots$ to denote quantities which are bounded by numbers depending only on p and $B_1, B_2, \cdots$ to denote positive quantities depending only on p.

For $|\xi_{nk}| < (pq)^{\frac{1}{2}}$, we have

$$\begin{aligned}\prod_{k=1}^n \rho_k(\psi,t) &= \exp[\textstyle\sum_{k=1}^n \log(qe^{-ip\xi_{nk}(pq)^{-\frac{1}{2}}} + pe^{iq\xi_{nk}(pq)^{-\frac{1}{2}}})] \\ &= \exp\left[-\tfrac{1}{2}(\psi^2+t^2) + \frac{t^2}{2n} + \textstyle\sum_{j=3}^r \frac{\gamma_j}{j!}\sum_{k=1}^n (i\xi_{nk})^j + \theta_1 \sum_{k=1}^n |\xi_{nk}|^{r+1}\right],\end{aligned}$$

where γ_j are the standardized cumulants of a binomial distribution arising from a single trial, so $\gamma_3 = (pq)^{-\frac{1}{2}}(q-p)$, $\gamma_4 = (pq)^{-1}(1-6pq)$. Let

$$V(z) = \frac{t^2z^2}{2n} + \sum_{j=3}^r \frac{\gamma_j}{j!}\sum_{k=1}^n (i\xi_{nk})^j z^{j-2} + \theta_1 \sum_{k=1}^n |\xi_{nk}|^{r+1}z^{r-1}, \tag{2}$$

and consider the power series expansion in z, $|z| \leqq 1$, for

$$e^{V(z)} = 1 + \sum_{j=1}^{r-2} P_j z^j + R(z) \tag{3}$$

where $R(z) = 0(z^{r-1})$ as z tends to zero and P_j are polynomials in ψ and t of degree $3j$ to be considered explicitly later.

Now

$$|\textstyle\sum_k \xi_{nk}^j| \leqq \max_k|\xi_{nk}|^{j-2}\sum_k \xi_{nk}^2 \leqq (\psi^2+t^2)[|\psi|n^{-\frac{1}{2}} + tb_n]^{j-2}\,.$$

So we can find C' depending only on p and r, such that for $|\psi| < 2C'n^{\frac{1}{2}}$, $|t| < C'b_n^{-1}$ and $n > 2$, $|\xi_{nk}| < (pq)^{\frac{1}{2}}$ and

$$V(1) < \tfrac{1}{4}(\psi^2+t^2)\,.$$

For $n = 2$ the theorem follows immediately by choosing B in (1) large enough. Also for $j > 2$,

$$\begin{aligned}\textstyle\sum_k |\xi_{nk}|^j &\leqq 2^{j-1}[|\psi|^j n^{-\frac{1}{2}(j-2)} + |t|^j \textstyle\sum_k |a_{nk}|^j] \\ &\leqq 2^{j-1}A_{jn}[|\psi| + |t|]^j\end{aligned}$$

since from the Hölder inequality, for $j > 2$,

$$1 = \textstyle\sum_k a_{nk}^2 \leqq n^{(j-2)/j}(\sum_k |a_{nk}|^j)^{2/j}$$

and so

$$A_{jn} \geqq n^{-\frac{1}{2}(j-2)}\,. \tag{4}$$

Also from the Hölder inequality, for $2 < j \leqq r$,

$$\begin{aligned}A_{jn} &= \textstyle\sum_k |a_{nk}|^j \leqq (\sum_k a_{nk}^2)^{1-(j-2)/(r-1)}(\sum_k |a_{nk}|^{r+1})^{(j-2)/(r-1)} \\ &= A_{r+1,n}^{(j-2)/(r-1)}\,.\end{aligned}$$

So for $i = 1, \cdots, r-1$, $|\psi| < 2C'n^{\frac{1}{2}}$ and $|t| < C'b_n^{-1}$;

(5) $$|V^{(i)}(1)| < B_1 A_{r+1,n}^{i/(r-1)}(|\psi| + |t|)^{i+2}$$

where $V^{(i)}(1)$ is the ith derivative of $V(z)$ with respect to z, evaluated at $z = 1$. Now

$$|R(1)| \leqq \left|\left[\frac{d^{r-1}}{dz^{r-1}} e^{V(z)}\right]\right|$$

for some $|z| \leqq 1$, so

(6) $$|R(1)| < |\sum a_{j_1 \cdots j_\nu} V^{(j_1)}(1) \cdots V^{(j_\nu)}(1)|e^{|V(1)|}| ,$$

where the summation is over all choices of $j_1, \cdots, j_\nu$ with $j_1 + \cdots + j_\nu = r-1$ and $a_{j_1 \cdots j_\nu}$ are quantities depending only on r. So using (5) and (6) we have

$$|R(1)| < A_{r+1,n} P_1(|\psi| + |t|) e^{\frac{1}{4}(\psi^2+t^2)} ,$$

where $P_1(x), \cdots$ are polynomials in x of degree $3(r-1)$ with coefficients depending only on p and r. Thus for $|\psi| < 2C'n^{\frac{1}{2}}$ and $|t| < C'b_n^{-1}$,

(7) $$|\prod_{k=1}^n \rho_k(\psi, t) - e^{-\frac{1}{2}(\psi^2+t^2)}(1 + \sum_{j=1}^{r-2} P_j)| < A_{r+1,n} P_1(|\psi| + |t|) e^{-\frac{1}{4}(\psi^2+t^2)} .$$

Let

(8) $$(2\pi)^{-\frac{1}{2}} \int_{-\infty}^{\infty} e^{-\frac{1}{2}(\psi^2+t^2)}(1 + \sum_{j=1}^{r-2} P_j)\, d\psi = e^{-\frac{1}{2}t^2}[1 + \sum_{j=1}^{r-2} Q_j^*(t)] ,$$

where $Q_j^*(t)$ are polynomials in t whose explicit value will be considered later. The difference between this integral and the integral of the same function over the range $(-2C'n^{\frac{1}{2}}, 2C'n^{\frac{1}{2}})$ is less than

(9) $$B_2 n^{\frac{1}{2}(3r-5)} e^{-2C'^2 n} < B_3 A_{r+1,n} .$$

So using (7), (8) and (9), we have

(10) $$|(2\pi)^{-\frac{1}{2}} \int_{-2C'n^{\frac{1}{2}}}^{2C'n^{\frac{1}{2}}} \prod_{k=1}^n \rho_k(\psi, t)\, d\psi - e^{-\frac{1}{2}t^2}[1 + \sum_{j=1}^{r-2} Q_j^*(t)]| < P_2(|t|) A_{r+1,n} e^{-\frac{1}{4}t^2}$$

for $|t| < C'b_n^{-1}$.

Now

$$|\rho_k(\psi, t)|^2 = 1 - 2pq[1 - \cos \xi_{nk}(pq)^{-\frac{1}{2}}] .$$

For $2C'n^{\frac{1}{2}} < |\psi| < \pi(npq)^{\frac{1}{2}}$ and $|t| < C'b_n^{-1}$, $|\xi_{nk}| > C'$. Then $C' < |\xi_{nk}| < 2\pi(pq)^{\frac{1}{2}} - C'$, and so

$$1 - \cos \xi_{nk}(pq)^{-\frac{1}{2}} \geqq 1 - \cos C'(pq)^{-\frac{1}{2}} \geqq \frac{C'^2}{2pq} - \frac{C'^4}{24p^2q^2} \geqq \frac{C'^2}{3pq}$$

for $C' < 2(pq)^{\frac{1}{2}}$, so

$$|\rho_k(\psi, t)|^2 = 1 - 2pq[1 - \cos \xi_{nk}(pq)^{-\frac{1}{2}}] < e^{-\frac{2}{3}C'^2}$$

for $C' < |\xi_{nk}| < 2\pi(pq)^{\frac{1}{2}} - C'$ and $C' < 2(pq)^{\frac{1}{2}}$. Thus

(11) $$\prod_{k=1}^n |\rho_k(\psi, t)| = e^{-\frac{1}{3}C'^2 n} \leqq \exp[-\tfrac{1}{4}t^2 - \tfrac{1}{12}C'^2 n]$$

for $2C'n^{\frac{1}{2}} < |\psi| < \pi(npq)^{\frac{1}{2}}$ and $|t| < C'b_n^{-1}$. So using the two estimates (10) and

(11), we have for $|t| < C'b_n^{-1}$,

$$|(2\pi)^{-\frac{1}{2}} \int_{-\pi(npq)^{\frac{1}{2}}}^{\pi(npq)^{\frac{1}{2}}} \prod_{k=1}^{n} \rho_k(\psi, t)\, d\psi - e^{-\frac{1}{2}t^2}[1 + \sum_{j=1}^{r-2} Q_j^*(t)]| < A_{r+1,n} P_3(|t|) e^{-\frac{1}{4}t^2}. \tag{12}$$

As a particular case of this result, we have

$$\begin{aligned}(2\pi npq)^{\frac{1}{2}} B_{ns}(p) &= (2\pi)^{-\frac{1}{2}} \int_{-\pi(npq)^{\frac{1}{2}}}^{\pi(npq)^{\frac{1}{2}}} \prod_{k=1}^{n} \rho_k(\psi, 0)\, d\psi \\ &= 1 + \sum_{j=1}^{r-2} Q_j^*(0) + \theta_2 A_{r+1,n}.\end{aligned}$$

So for $|t| < C'b_n^{-1}$, we have

$$|f_{ns}(t) - e^{-\frac{1}{2}t^2}[1 + \sum_{j=1}^{r-2} Q_j^*(t)][1 + \sum_{j=1}^{r-2} Q_j^*(0)]^{-1}| < A_{r+1,n} P_4(|t|) e^{-\frac{1}{4}t^2}. \tag{13}$$

Restricting attention to the case $r = 5$, we will calculate $Q_1^*(t)$ and $Q_2^*(t)$ explicitly and show that $Q_3^*(t)$ is a polynomial with zero constant term and all coefficients bounded by by A_{5n} times some constant depending on p only. From (2) and (3) we have

$$\begin{aligned}P_1 &= \frac{\gamma_3}{3!} \sum_k (i\xi_{nk})^3 \\ P_2 &= \frac{\gamma_4}{4!} \sum_k (i\xi_{nk})^4 + \frac{t^2}{2n} + \tfrac{1}{2}\left[\frac{\gamma_3}{3!} \sum_k (i\xi_{nk})^3\right]^2 \\ P_3 &= \frac{\gamma_5}{5!} \sum_k (i\xi_{nk})^5 + \left[\frac{\gamma_3}{3!} \sum_k (i\xi_{nk})^3\right]\left[\frac{\gamma_4}{4!} \sum_k (i\xi_{nk})^4 + \frac{t^2}{2n}\right] \\ &\quad + \frac{1}{3!}\left[\frac{\gamma_3}{3!} \sum_k (i\xi_{nk})^3\right]^3.\end{aligned}$$

The terms involving powers of ψ only in P_3 are all of odd power so they do not appear on the right-hand side of (8), where $Q_3^*(t)$ is defined. It is readily seen that

$$\begin{aligned}Q_1^*(t) &= \frac{q-p}{6(pq)^{\frac{1}{2}}} (it)^3 \sum_k a_{nk}^3 (1 + n^{-1}\theta_3) \\ Q_2^*(t) &= \left\{\frac{1-6pq}{24pq}\left[\frac{3}{n} + \frac{6t^2}{n} + t^4 \sum_k a_{nk}^4\right] + \frac{t^2}{2n}\right. \\ &\quad \left. - \frac{(p-q)^2}{72pq}\left[\frac{15}{n} + \frac{18t^2}{n} + \frac{9t^4}{n} + t^6(\sum_k a_{nk}^3)^2\right]\right\}(1 + n^{-1}\theta_4).\end{aligned}$$

Now

$$[1 + Q_1^*(t) + Q_2^*(t)][1 + Q_1^*(0) + Q_2^*(0)]^{-1} = 1 + [Q_1(t) + Q_2(t)](1 + n^{-1}\theta_5)$$

where

$$Q_1(t) = \frac{q-p}{6(pq)^{\frac{1}{2}}} (it)^3 \sum_k a_{nk}^3$$

and

$$Q_2(t) = (it)^4\left[\frac{1-6pq}{24pq}(\sum_k a_{nk}^4 - 3n^{-1}) - \tfrac{1}{4}n^{-1}\right] + (it)^6 \frac{(q-p)^2}{72pq} (\sum_k a_{nk}^3)^2.$$

From Hölder's inequality if $r > s$,

$$\sum_k |a_{nk}|^s \leqq n^{(r-s)/r}(\sum_k |a_{nk}|^r)^{s/r} .$$

So using (4) we have for $r > s$

$$A_{rn} \geqq n^{-(r-s)/s} A_{sn}^{r/s} \geqq (n^{-(r-s)/2} A_{sn}) A_{sn}^{(r-s)/s} n^{(r-s)(s-2)/2s}$$
$$\geqq A_{sn}^{(r-s)/s} n^{(r-s)(s-2)/2s} .$$

Applying this to the terms in $Q_3^*(t)$ it is seen that all coefficients are bounded by A_{5n} times some constant depending on p only. Then we have for $|t| < C'b_n^{-1}$,

$$(14) \qquad |f_{ns}(t) - g_{ns}(t)| < P_5(|t|)e^{-\frac{1}{4}t^2}(A_{5n}|t| + A_{6n}) ,$$

since the characteristic function corresponding to $G_{ns}(x)$ is

$$g_{ns}(t) = e^{-\frac{1}{2}t^2}[1 + Q_1(t) + Q_2(t)] .$$

The number of indices k for which, for any fixed ψ, $|ta_{nk} + \psi n^{-\frac{1}{2}} - 2r\pi| > \varepsilon$, for all $r = 0, \pm 1, \pm 2, \cdots$ and all $C'b_n^{-1} < |t| < CA_{5n}^{-1}$, is greater than δn, for each n, so

$$(15) \qquad \prod_{k=1}^n |\rho_k(\psi, t)|^2 = \prod_{k=1}^n [1 - 2pq(1\text{-}\cos[\{\psi n^{-\frac{1}{2}} + ta_{nk}n^{-\frac{1}{2}}(n-1)^{\frac{1}{2}}\}(qp)^{-\frac{1}{2}}])]$$
$$< e^{-\frac{1}{2}\delta\varepsilon^2 n} .$$

Also for $|t| > C'b_n^{-1}$,

$$(16) \qquad |g_{ns}(t)| < B_4 b_n^{-6} e^{-\frac{1}{2}C'^2 b_n^{-2}} .$$

We will use the well-known inequality (see, for example, Feller (1966), page 510)

$$|F_{ns}(x) - G_{ns}(x)| < \int_{-T}^{T} |t|^{-1}|f_{ns}(t) - g_{ns}(t)|dt + 12m(\pi T)^{-1} ,$$

where $m = \sup_x G'_{ns}(x)$ and we will take $T = CA_{5n}^{-1}$. From (14), (15) and (16), we have

$$(17) \qquad \int_{-T}^{-T^{-1}} + \int_{T^{-1}}^{T} |t|^{-1}|f_{ns}(t) - g_{ns}(t)|\, dt \leqq B_5 A_{5n} + B_6 A_{6n} \log T \leqq B_7 A_{5n} .$$

So it only remains to show that

$$(18) \qquad \int_{-T^{-1}}^{T^{-1}} |t|^{-1}|f_{ns}(t) - g_{ns}(t)|\, dt < B_8 A_{5n} .$$

Now

$$\int_{-T^{-1}}^{T^{-1}} |t|^{-1}e^{-\frac{1}{2}t^2}|Q_1(t) + Q_2(t)|\, dt < B_9 A_{5n}$$

so we need only consider $|t|^{-1}|f_{ns}(t) - e^{-\frac{1}{2}t^2}|$ in the range $(-T^{-1}, T^{-1})$. Now

$$|t|^{-1}|f_{ns}(t) - e^{-\frac{1}{2}t^2}| \leqq |t|^{-1} \int_0^t |f'_{ns}(\eta) + \eta e^{-\frac{1}{2}\eta^2}|d\eta$$
$$\leqq \sup_{0 \leqq \eta \leqq t} |f'_{ns}(\eta) + \eta e^{-\frac{1}{2}\eta^2}| .$$

Also

$$\frac{d}{dt} \prod_{k=1}^n \rho_k(\psi, t) = \sum_{k=1}^n ia_{nk}(pq)^{\frac{1}{2}}(-e^{-ip\zeta_k} + e^{iq\zeta_k})(qe^{-ip\zeta_k} + pe^{iq\zeta_k})^{-1}$$
$$\times n^{-\frac{1}{2}}(n-1)^{\frac{1}{2}} \prod_{j=1}^n \rho_j(\psi, t) ,$$

where $\zeta_k = (\psi + (n-1)^{\frac{1}{2}} t a_{nk})(npq)^{-\frac{1}{2}}$. But for $|\psi| < 2C'n^{\frac{1}{2}}$, $|\zeta_k| < 3C'(pq)^{-\frac{1}{2}}$; so

$$(-e^{-ip\zeta_k} + e^{iq\zeta_k})(qe^{-ip\zeta_k} + pe^{iq\zeta_k})^{-1} = i\zeta_k + \theta_5|\zeta_k|^2 ,$$

and so

$$\frac{d}{dt} \prod_{k=1}^n \rho_k(\psi, t) = \prod_{k=1}^n \rho_k(\psi, t)[-t + \theta_6(\psi^2 n^{-\frac{1}{2}} + 2\psi t n^{-\frac{1}{2}} + t^2 \textstyle\sum_k |a_{nk}|^3)] .$$

Integrating this with respect to ψ and using (7) and then using (11) for $2C'n^{\frac{1}{2}} < |\psi| < \pi(npq)^{\frac{1}{2}}$, as before, we have

$$\sup_{0 \leq \eta \leq t} |f'_{ns}(\eta) + \eta e^{-\frac{1}{2}\eta^2}| < B_{10} ,$$

for $-T^{-1} < t < T^{-1}$. So (18) holds and combining (17) and (18) we have (1).

REFERENCES

ALBERS, W., BICKEL, P. J. and VAN ZWET, W. R., (1976). Asymptotic expansions for the power of distribution free tests in the one sample problem. *Ann. Statist.* **4** 108–156.

VON BAHR, B. (1972). On sampling from a finite set of independent random variables. *Z. Wahrscheinlichkeitstheorie und Verw. Gebiete* **24** 279–286.

BIKELIS, A. (1969). On the estimation of the remainder term in the central limit theorem for samples from finite populations (in Russian). *Studia Sci. Math. Hungar.* **4** 345–354.

ERDÖS, P. and RÉNYI, A. (1959). On the central limit theorem for samples from a finite population. *Publ. Math. Inst. Hungar. Acad. Sci.* **4** 49–61.

FELLER, W. (1966). *An Introduction to Probability Theory and its Applications*, **2**. Wiley, New York.

HODGES, J. L. and FIX, E. (1955). Significance probabilities of the Wilcoxon test. *Ann. Math. Statist.* **26** 301–312.

DEPARTMENT OF MATHEMATICAL STATISTICS
UNIVERSITY OF SYDNEY
2006, NEW SOUTH WALES
AUSTRALIA