A COMBINATORIAL CENTRAL LIMIT THEOREM.

by

Wassily Hoeffding

Special report to the Office of Naval Research
of work at Chapel Hill under Project NR 042 031
for research in probability and statistics

Institute of Statistics
Mimeograph Series No. 46

July 1, 1951

A COMBINATORIAL CENTRAL LIMIT THEOREM.[1]

By Wassily Hoeffding

Institute of Statistics, University of North Carolina.

1. Summary. Let $(Y_{n1}, \ldots, Y_{nn})$ be a random vector which takes on the n! permutations of $(1, \ldots, n)$ with equal probabilities. Let $c_n(i,j)$, $i,j = 1, \ldots, n$, be n^2 real numbers. Sufficient conditions for the asymptotic normality of

$$S_n = \sum_{i=1}^{n} c_n(i,Y_{ni})$$

are given (theorem 3). For the special case $c_n(i,j) = a_n(i)b_n(j)$ a stronger version of a theorem of Wald, Wolfowitz and Noether is obtained (theorem 4). A condition of Noether is simplified (theorem 1).

2. Introduction and statement of results.

An example of what is here meant by a combinatorial central limit theorem is a solution of the following problem. For every positive integer n there are given 2n real numbers $a_n(i)$, $b_n(i)$, $i = 1, \ldots, n$. It is assumed that the $a_n(i)$ are not all equal and the $b_n(i)$ are not all equal. Let $(Y_{n1}, \ldots, Y_{nn})$ be a random vector which takes on the n! permutations of $(1, \ldots, n)$ with equal probabilities $1/n!$. Under what conditions is

$$S_n = \sum_{i=1}^{n} a_n(i)b_n(Y_{ni}) \tag{1}$$

asymptotically normally distributed as $n \to \infty$?

Throughout this paper a random variable S_n will be called asymptotically normal or asymptotically normally distributed if

1. Work done under the sponsorship of the Office of Naval Research.

$$\lim_{n \to \infty} \Pr \left\{ S_n - ES_n \leq x \sqrt{\operatorname{var} S_n} \right\} = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{x} e^{-y^2/2} \, dy, \quad -\infty < x < \infty,$$

where ES_n and $\operatorname{var} S_n$ are the mean and the variance of S_n.

In the particular case $a_n(i) = b_n(i) = i$ the asymptotic normality of S_n was proved by Hotelling and Pabst [2]. The first general result is due to Wald and Wolfowitz [6] who showed that S_n is asymptotically normal if, as $n \to \infty$,

$$(2) \qquad \frac{\frac{1}{n}\sum_{i=1}^{n} (a_n(i) - \bar{a}_n)^r}{\left[\frac{1}{n}\sum_{i=1}^{n} (a_n(i) - \bar{a}_n)^2\right]^{r/2}} = O(1), \qquad r = 3, 4, \ldots$$

and

$$(3) \qquad \frac{\frac{1}{n}\sum_{i=1}^{n} (b_n(i) - \bar{b}_n)^r}{\left[\frac{1}{n}\sum_{i=1}^{n} (b_n(i) - \bar{b}_n)^2\right]^{r/2}} = O(1), \qquad r = 3, 4, \ldots$$

where

$$\bar{a}_n = \frac{1}{n}\sum_{i=1}^{n} a_n(i), \qquad \bar{b}_n = \frac{1}{n}\sum_{i=1}^{n} b_n(i).$$

Noether [5] proved that condition (3) can be replaced by the weaker condition

$$(4) \qquad \lim_{n \to \infty} \frac{\sum_{i=1}^{n} (b_n(i) - \bar{b}_n)^r}{\left[\sum_{i=1}^{n} (b_n(i) - \bar{b}_n)^2\right]^{r/2}} = 0, \qquad r = 3, 4, \ldots$$

This condition can be simplified as follows.

<u>Theorem 1</u>. <u>Condition (4) is equivalent to either of the following two conditions:</u>

(5) $$\lim_{n \to \infty} \frac{\sum_{i=1}^{n} \left| b_n(i) - \bar{b}_n \right|^r}{\left[\sum_{i=1}^{n} (b_n(i) - \bar{b}_n)^2 \right]^{r/2}} = 0 \qquad \underline{\text{for some}}\ r > 2;$$

(6) $$\lim_{n \to \infty} \frac{\max_{1 \le i \le n} (b_n(i) - \bar{b}_n)^2}{\sum_{i=1}^{n} (b_n(i) - \bar{b}_n)^2} = 0.$$

Hence conditions (2) and (5) or (2) and (6) are sufficient for the asymptotic normality of (1).

The proof is given in section 3. For a more general condition and a stronger but simpler condition see theorem 4 below.

One extension of this problem was considered by Daniels [1] who studied the asymptotic distribution of

$$\sum_{i=1}^{n} \sum_{j=1}^{n} a_n(i, j)\, b_n(Y_{ni}, Y_{nj}).$$

The present paper is concerned with an alternative extension. It considers the distribution of

(7) $$S_n = \sum_{i=1}^{n} c_n(i, Y_{ni}),$$

where $c_n(i, j)$, $i, j = 1, \ldots, n$, are n^2 real numbers, defined for every positive integer n. In the particular case $c_n(i, j) = a_n(i)\, b_n(j)$, (7) reduces to (1).

Let

(8) $$d_n(i,j) = c_n(i,j) - \frac{1}{n} \sum_{g=1}^{n} c_n(g,j) - \frac{1}{n} \sum_{h=1}^{n} c_n(i,h) + \frac{1}{n^2} \sum_{g=1}^{n} \sum_{h=1}^{n} c_n(g,h).$$

<u>Theorem 2</u>. <u>The mean and variance of</u>

$$S_n = \sum_{i=1}^{n} c_n(i, Y_{ni})$$

<u>are</u>

(9)
$$ES_n = \frac{1}{n} \sum_{i=1}^{n} \sum_{j=1}^{n} c_n(i,j),$$

(10)
$$\operatorname{var} S_n = \frac{1}{n-1} \sum_{i=1}^{n} \sum_{j=1}^{n} d_n^2(i,j).$$

Henceforth we assume that $d_n(i,j) \neq 0$ for some (i,j), so that $\operatorname{var} S_n > 0$. In the special case $c_n(i,j) = a_n(i)b_n(j)$ this corresponds to the assumption that the $a_n(i)$ are not all equal and the $b_n(j)$ are not all equal.

<u>Theorem 3</u>. <u>The distribution of</u> $S_n = \Sigma_{i=1}^n c_n(i, Y_{ni})$ <u>is asymptotically normal if</u>

(11)
$$\lim_{n \to \infty} \frac{\frac{1}{n} \sum_{i=1}^{n} \sum_{j=1}^{n} d_n^r(i,j)}{\left[\frac{1}{n} \sum_{i=1}^{n} \sum_{j=1}^{n} d_n^2(i,j)\right]^{r/2}} = 0, \qquad r = 3, 4, \ldots$$

<u>Condition</u> (11) <u>is satisfied if</u>

(12)
$$\lim_{n \to \infty} \frac{\max\limits_{1 \leq i,j \leq n} d_n^2(i,j)}{\frac{1}{n} \sum_{i=1}^{n} \sum_{j=1}^{n} d_n^2(i,j)} = 0.$$

Theorems 2 and 3 will be proved in sections 4 and 5.

For the special case $c_n(i,j) = a_n(i)b_n(j)$, theorem 3 immediately gives

<u>Theorem 4.</u> <u>The distribution of</u> $S_n = \Sigma_{i=1}^n a_n(i) b_n(Y_{ni})$ <u>is asymptotically normal if</u>

$$(13) \qquad \lim_{n \to \infty} n^{\frac{r}{2}-1} \frac{\sum_{i=1}^n (a_n(i) - \bar{a}_n)^r}{\left[\sum_{i=1}^n (a_n(i) - \bar{a}_n)^2\right]^{r/2}} \frac{\sum_{i=1}^n (b_n(i) - \bar{b}_n)^r}{\left[\sum_{i=1}^n (b_n(i) - \bar{b}_n)^2\right]^{r/2}} = 0,$$

$$r = 3, 4, \ldots$$

<u>Condition</u> (13) <u>is satisfied if</u>

$$(14) \qquad \lim_{n \to \infty} n \frac{\max_{1 \leq i \leq n} (a_n(i) - \bar{a}_n)^2}{\sum_{i=1}^n (a_n(i) - \bar{a}_n)^2} \frac{\max_{1 \leq i \leq n} (b_n(i) - \bar{b}_n)^2}{\sum_{i=1}^n (b_n(i) - \bar{b}_n)^2} = 0.$$

It will be observed that the symmetrical condition (13) contains Noether's condition (2) and (4) as a special case.

Let $X_n = (X_{n1}, \ldots, X_{nn})$ be independent of and have the same distribution as $Y_n = (Y_{n1}, \ldots, Y_{nn})$.

<u>Theorem 5.</u> <u>The random variable</u>

$$(15) \qquad S_n' = \sum_{i=1}^n c_n(X_{ni}, Y_{ni})$$

<u>has the same distribution as</u> S_n <u>in</u> (7).

In fact, the conditional distribution of S_n' given that $X_n = p$, a fixed permutation of $(1, \ldots, n)$, is independent of p because the distribution of Y_n is invariant under permutations of its components.

The distribution of sums of the form (1) has attracted the attention of statisticians in connection with non-parametric tests (see, for example, [2], [6], [3]) and sampling from a finite population (which leads to the case $a_n(i) = 0$ for $i > m$; cf. also Madow [4]). More general sums of the form (7) or (15) are likewise of interest in non-parametric theory. Thus it follows from results of Lehmann and Stein [3] that a test of the hypothesis that

$U_1, \ldots, U_n$ are independent and identically distributed, which is most powerful similar against the alternative that the joint frequency function is $f_1(u_1) \ldots f_n(u_n)$ is based on a statistic of the form (7) with

$$c_n(i,j) = \log f_i(u_j),$$

where the u_j are the observed sample values. If the n pairs $(U_1,V_1), \ldots, (U_n,V_n)$ are independent and identically distributed, a test of the hypothesis that U_i and V_i are independent which is most powerful similar against the alternative that their joint frequency function is $f(u,v)$ is based on a statistic of the form (15) with $c_n(i,j) = \log f(u_i,v_j)$, where $(u_1,v_1), \ldots, (u_n,v_n)$ are the observed values.

In these examples the numbers $c_n(i,j)$ are random variables. An application of some of the present results to such cases will be considered by the author in a forthcoming paper.

3. Proof of theorem 1. Let

$$g_i = \frac{b_n(i) - \bar{b}_n}{\left[\sum_{i=1}^{n} (b_n(i) - \bar{b}_n)^2\right]^{1/2}},$$

$$G_n = \max(|g_1|, \ldots, |g_n|).$$

Theorem 1 asserts the equivalence of the three relations

(16) $$\lim_{n \to \infty} \sum_{i=1}^{n} g_i^r = 0, \qquad r = 3, 4, \ldots;$$

(17) $$\lim_{n \to \infty} \sum_{i=1}^{n} |g_i|^r = 0 \qquad \text{for some } r > 2;$$

(18) $$\lim_{n \to \infty} G_n = 0.$$

We have

(19) $$\sum_{i=1}^{n} g_i^2 = 1,$$

and hence

(20) $$G_n \leq 1.$$

If $0 \leq r < s < t$, Hölder's inequality gives

(21) $$\sum_{i=1}^{n} |g_i|^{s} = \sum_{i=1}^{n} |g_i|^{\frac{r(t-s)}{t-r}} |g_i|^{\frac{t(s-r)}{t-r}}$$

$$\leq (\sum_{i=1}^{n} |g_i|^{r})^{\frac{t-s}{t-r}} (\sum_{i=1}^{n} |g_i|^{t})^{\frac{s-r}{t-r}}.$$

If we let $r = 2k$, $t = 2k + 2$, $k = 1, 2, \ldots$, it follows that relation (16) is equivalent to

(22) $$\lim_{n \to \infty} \sum_{i=1}^{n} |g_i|^{r} = 0, \qquad 2 < r < \infty$$

From (21) with $r = 2$ and (19) we have

(23) $$\sum_{i=1}^{n} |g_i|^{s} \leq (\sum_{i=1}^{n} |g_i|^{t})^{\frac{s-2}{t-2}} \qquad \text{if } 2 < s < t.$$

By (20), for $r < s$

(24) $$\sum_{i=1}^{n} |g_i|^{s} \leq G_n^{s-r} \sum_{i=1}^{n} |g_i|^{r} \leq \sum_{i=1}^{n} |g_i|^{r}.$$

Inequalities (23) and (24) show that (17) implies (22). Since obviously (22) implies (17), these relations are equivalent.

The equivalence of (17) and (18) is seen from the inequalities

$$G_n^r \leq \sum_{i=1}^{n} |g_i|^r \leq G_n^{r-2} \sum_{i=1}^{n} g_i^2 = G_n^{r-2}, \qquad r > 2.$$

This completes the proof.

4. Proof of theorem 2. The subscript n in Y_{ni}, $c_n(i,j)$ etc. will henceforth be omitted. We note that if the subscripts $i_1, \ldots, i_m$ are distinct, the expected value of a function $f(Y_{i_1}, \ldots, Y_{i_m})$ is equal to

$$\frac{1}{n(n-1)\ldots(n-m+1)} \sum_{j_1, \ldots, j_m}' f(j_1, \ldots, j_m),$$

where the sum Σ' is extended over all m-tuples $(j_1, \ldots, j_m)$ of distinct integers from 1 to n.

Relation (9) follows immediately.

Let

(25) $$T_n = \sum_{i=1}^{n} d(i,Y_i),$$

where $d(i,j) = d_n(i,j)$ is defined by (8). Using (9), we get

(26) $$T_n = S_n - ES_n.$$

Also

(27) $$\sum_{i=1}^{n} d(i,j) = 0 \text{ for all } j, \qquad \sum_{j=1}^{n} d(i,j) = 0 \text{ for all } i.$$

Hence

$$Ed(i,Y_i) = 0,$$

$$Ed^2(i,Y_i) = \frac{1}{n} \sum_{j=1}^{n} d^2(i,j),$$

and if $i \neq j$,

$$Ed(i,Y_i)d(j,Y_j) = \frac{1}{n(n-1)} \sum_{g,h}' d(i,g)d(j,h)$$

$$= \frac{-1}{n(n-1)} \sum_{g=1}^{n} d(i,g)d(j,g).$$

Hence

$$\text{var } S_n = \text{var } T_n = \sum_{i=1}^{n} Ed^2(i,Y_i) + \sum_{i,j}' Ed(i,Y_i)d(j,Y_j)$$

$$= \frac{1}{n} \sum_{i=1}^{n} \sum_{j=1}^{n} d^2(i,j) - \frac{1}{n(n-1)} \sum_{g=1}^{n} \sum_{i,j}' d(i,g)d(j,g)$$

$$= \frac{1}{n} \sum_{i=1}^{n} \sum_{j=1}^{n} d^2(i,j) + \frac{1}{n(n-1)} \sum_{g=1}^{n} \sum_{i=1}^{n} d^2(i,g),$$

which gives relation (10).

5. Proof of theorem 3. Let

(28) $$M_{r,n} = \frac{1}{n} \sum_{i=1}^{n} \sum_{j=1}^{n} d^r(i,j),$$

(29) $$\overline{M}_{r,n} = \frac{1}{n} \sum_{i=1}^{n} \sum_{j=1}^{n} \left| d(i,j) \right|^r,$$

(30) $$D_n = \max_{1 \leq i,j \leq n} \left| d(i,j) \right| .$$

Then $\text{var } S_n = \frac{n}{n-1} M_{2,n}$. Since, by hypothesis, $\text{var } S_n > 0$, we may and shall assume that

(31) $$M_{2,n} = 1.$$

Conditions (11) and (12) can now be written as

(32) $$\lim_{n \to \infty} M_{r,n} = 0, \qquad r = 3, 4, \ldots$$

and

$$\lim_{n \to \infty} D_n = 0. \tag{33}$$

That (33) implies (32) is seen from the inequalities

$$\left| M_{r,n} \right| \leq \overline{M}_{r,n} \leq D_n^{r-2} M_{2,n} = D_n^{r-2} \qquad \text{for } r > 2.$$

Since (cf. inequality (21))

$$\overline{M}^2_{2k+1,n} \leq M_{2k,n} M_{2k+2,n}, \qquad k = 1, 2, \ldots,$$

condition (32) implies

$$\lim_{n \to \infty} \overline{M}_{r,n} = 0, \qquad r = 3, 4, \ldots \tag{34}$$

As var $S_n \to 1$, it is now sufficient to demonstrate that under conditions (31) and (34), $T_n = S_n - ES_n$ has a normal limiting distribution with mean 0 and variance 1. This will be proved by showing that

$$\lim_{n \to \infty} ET_n^r = \begin{cases} 1\cdot 3 \cdots (r-1) & \text{if } r \text{ is even} \\ 0 & \text{if } r \text{ is odd.} \end{cases} \tag{35}$$

The r-th moment of T_n,

$$ET_n^r = E \sum_{i_1=1}^{n} \cdots \sum_{i_r=1}^{n} d(i_1, Y_{i_1}) \ldots d(i_r, Y_{i_r}), \tag{36}$$

can be written as a sum of terms of the form

$$I(r, e_1, \ldots, e_m) = \sum_{i_1, \ldots, i_m}{}' Ed^{e_1}(i_1, Y_{i_1}) \ldots d^{e_m}(i_m, Y_{i_m}), \tag{37}$$

where $e_i \geq 1$, $e_1 + \ldots + e_m = r$. The number of terms (37) is independent of n.

It will be shown that

$$(38) \qquad \lim_{n \to \infty} I(r, e_1, \ldots, e_m) = 0 \quad \text{unless } r \text{ even, } m = r/2,\ e_1 = \ldots = e_m = 2,$$

$$(39) \qquad \lim_{n \to \infty} I(r, 2, \ldots, 2) = 1 \qquad \text{if } r \text{ even,}$$

and that the number of terms $I(r,2, \ldots, 2)$ in (36) with r even equals $1 \cdot 3 \ldots (r-1)$. Then (35) holds, and the theorem will be proved.

We have for $n \to \infty$

$$(40) \qquad I(r, e_1, \ldots, e_m) \sim n^{-m} \sum_{i_1, \ldots, i_m}{}' \sum_{j_1, \ldots, j_m}{}' d^{e_1}(i_1, j_1) \ldots d^{e_m}(i_m, j_m).$$

The right-hand side can be written as a sum of terms which, apart from the sign, are of the form

$$(41) \qquad n^{-m} J(r,p,q,e_1, \ldots, e_m) = n^{-m} \sum_{i_1=1}^{n} \ldots \sum_{i_p=1}^{n} \sum_{j_1=1}^{n} \ldots \sum_{j_q=1}^{n} d^{e_1}(i_{c_1}, j_{d_1}) \ldots d^{e_m}(i_{c_m}, j_{d_m})$$

where

$$1 \leq p \leq m, \qquad 1 \leq q \leq m,$$

$$1 \leq c_g \leq p, \qquad 1 \leq d_h \leq q, \quad (g,h = 1, \ldots, m),$$

and for every integer u, $1 \leq u \leq p$ $(1 \leq u \leq q)$ at least one $c_g (d_h)$ is equal to u. The number of terms (41) is independent of n.

The sum J in (41) can be written as a product of $s \geq 1$ sums of a similar form,

$$(42) \qquad J(r,p,q,e_1, \ldots, e_m) = \prod_{k=1}^{s} J(r_k, p_k, q_k, e_{k1}, \ldots, e_{km_k}),$$

where

$$(e_{k1}, \ldots, e_{km_k}), \qquad k = 1, \ldots, s,$$

are s disjoint subsets of $(e_1, \ldots, e_m)$,

$$(43)\quad \begin{cases} e_{k1} + \dots + e_{km_k} = r_k, \\ r_1 + \dots + r_s = r, \\ p_1 + \dots + p_s = p, \qquad q_1 + \dots + q_s = q, \\ m_1 + \dots + m_s = m. \end{cases}$$

We observe that

$$(44)\qquad 1 \le p_k \le m_k, \qquad 1 \le q_k \le m_k, \qquad m_k \le r_k.$$

It will be assumed that s is the greatest possible number of factors into which $J(r,p,q,e_1,\dots,e_m)$ can be decomposed in the form (42). If $s = 1$, the number of equalities between the subscripts c or between the subscripts d in (41) must be at least $m-1$. The total number of subscripts c,d being 2m, there are at most $m + 1$ distinct subscripts, so that $p + q \le m + 1$. If

$$(45)\qquad (c_g,d_g) = (c_h,d_h) \quad \text{for some } (g,h),\ g \neq h,$$

we have strict inequality. For an arbitrary s we have in a similar way

$$(46)\qquad p_k + q_k \le m_k + 1, \qquad k = 1, \dots, s,$$

and hence

$$(47)\qquad p + q \le m + s,$$

with strict inequality in the case (45).

By Hölder's inequality, from (41),

$$\left| J(r,p,q,e_1,\dots,e_m) \right| \le \prod_{g=1}^{m} \Big(\sum_{i_1} \dots \sum_{i_p} \sum_{j_1} \dots \sum_{j_q} \left| d(i_{c_g}, j_{d_g}) \right|^{r} \Big)^{e_g/r}$$

$$= \prod_{g=1}^{m} \left(n^{p+q-1}\, \bar{M}_{r,n} \right)^{e_g/r} = n^{p+q-1}\, \bar{M}_{r,n}.$$

Similarly,

$$\left|J(r_k,p_k,q_k,e_{k1},\ldots,e_{km_k})\right| \leq n^{p_k+q_k-1}\,\bar{M}_{r_k,n}.$$

Hence, by (42),

$$(48) \quad n^{-m}\left|J(r,p,q,e_1,\ldots,e_m)\right| \leq n^{p+q-s-m}\,\bar{M}_{r_1,n}\cdots\bar{M}_{r_s,n}.$$

If, for some k, $r_k = 1$, then, by (44) and (43), $p_k=q_k=m_k=e_{k1}=1$, and hence $J = 0$ by (27). Thus we may assume $r_k \geq 2$, $k = 1, \ldots, s$. Then, by (34), $\bar{M}_{r_1n}\cdots\bar{M}_{r_sn} \to 0$ unless $r_1 = \ldots = r_s = 2$. It now follows from (48) and (47) that

$$(49) \qquad \lim_{n\to\infty} n^{-m}\,J(r,p,q,e_1,\ldots,e_m) = 0$$

except perhaps when $r_1 = \ldots = r_s = 2$.

If $r_1 = \ldots = r_s = 2$, we have

$$(50) \qquad n^{-m}\,J(r,p,q,e_1,\ldots,e_m) = 0\,(n^{p+q-s-m}).$$

By (44), $r_k = 2$ implies $m_k = 1$ or 2.

If $m_k = 2$, then $e_{k1} = e_{k2} = 1$ and $p_k + q_k \leq 3$ by (46). If $p_k + q_k = 3$, the corresponding J-factor is of the form

$$\sum_i\sum_j\sum_k d(i,j)d(i,k) \qquad \text{or} \qquad \sum_i\sum_j\sum_k d(i,k)d(j,k),$$

both of which vanish by (27). If $m_k = 2$ and $p_k + q_k = 2$, we have case (45) and hence, by the remark following (47), $p + q - s - m < 0$. By (50), this implies (49).

Thus the only case where (49) need not hold is $r_k = 2$, $m_k = 1$ for $k = 1, \ldots, s$. Then $p_k = q_k = 1$, $e_{k1} = 2$, hence

$$r = 2s = 2m, \qquad p = q = r/2$$

$$e_1 = \ldots = e_m = 2.$$

This proves relation (38), and (39) follows from

$$I(r,2,\ldots,2) \sim n^{-\frac{r}{2}} J(r, \tfrac{r}{2}, \tfrac{r}{2}, 2, \ldots, 2)$$
$$= n^{-\frac{r}{2}} [J(2,1,1,2)]^{r/2}$$
$$= M_{2,n}^{r/2} = 1.$$

It remains to determine the number of terms $I(r,2,\ldots,2)$ in (36) when r is even. This is the number of ways the subscripts $i_1,\ldots,i_r$ can be tied in $r/2$ groups of two, which is $(r-1)(r-3)\ldots3\cdot1$. The proof is complete.

REFERENCES

[1] H. E. Daniels. The relation between measures of correlation in the universe of sample permutations. Biometrika, Vol.33, pp.129-135(1944).

[2] H. Hotelling and M. Pabst. Rank correlation and tests of significance involving no assumption of normality. Annals of Math.Stat., Vol. 7, pp. 29-43 (1936).

[3] E. L. Lehmann and C. Stein. On the theory of some non-parametric hypotheses. Annals of Math. Stat., Vol. 20, pp. 28-45 (1949).

[4] W. G. Madow. On the limiting distributions of estimates based on samples from finite universes. Annals of Math.Stat., Vol.19, pp.535-545 (1948).

[5] G. E. Noether. On a theorem by Wald and Wolfowitz. Annals of Math. Stat., Vol. 20, pp. 455-458 (1949).

[6] A. Wald and J. Wolfowitz. Statistical tests based on permutations of the observations. Annals of Math.Stat., Vol. 15, pp. 358-372 (1944).