KRUSKAL-WALLIS ONE-WAY ANALYSIS OF VARIANCE
BASED ON LINEAR PLACEMENTS

YICHCENG HONG AND SUNGCHUL LEE

Abstract. The limiting distribution for the linear placement statistics under the null hypotheses has been provided by Orban and Wolfe [9] and Kim [5] when one of the sample sizes goes to infinity, and by Kim, Lee and Wang [6] when the sample sizes of each group go to infinity simultaneously. In this paper we establish the generalized Kruskal-Wallis one-way analysis of variance for the linear placement statistics.

1. Introduction and statement of main results

To determine whether the mutually independent random samples come from a common underlying distribution, we use various kinds of two-sample tests. Two-sample tests that are valid with no assumptions (except the continuity of the underlying distribution), are called the distribution-free procedures. Some distribution-free procedures use the sample ranks; the rank-sum test of Wilcoxon [11] and the median test of Mood [8]. Other distribution-free tests utilize the sample placements; the linear placement test of Orban and Wolfe [9] and the fixed and updated linear placement test of Kim [5].

Asymptotic theories for the sample rank tests are well developed; see Chernoff and Savage [1], Govindarajulu, Le Cam and Raghavachari [3], Hajek [4], Pyke and Shorak [10], and Dupac and Hajek [2]. However, asymptotic theories for the sample linear placement tests are quite limited. Orban and Wolfe [9] provided the limiting theory when one of the sample sizes goes to infinity. Kim and Wolfe [7] investigated the iterative asymptotic distribution when the control group sample size goes to infinity then the treatment group sample size goes to infinity. Similarly, Kim [5] established the limiting theory of the fixed and updated linear placement test when the control group sample size goes to infinity.
infinity. Moreover, Kim, Lee, and Wang [6] further established the central limit theorem when the sample sizes of each group go to infinity simultaneously.

In this paper we develop the Kruskal-Wallis one-way analysis of variance based on linear placements. The usual Kruskal-Wallis one-way analysis of variance is based on ranks but our Kruskal-Wallis one-way analysis of variance is based on linear placements, which can be used for comparing more than two samples that are independent, or not related. Our main results are follows.

First we fix our notations. The data set consists of $k$ groups and the group $i$ has $n_i$ observations $X_{ij}, 1 \leq j \leq n_i$. Overall we have $N = \sum_{i=1}^{k} n_i$ observations. We assume that $X_{ij}, 1 \leq i \leq k, 1 \leq j \leq n_i$, are mutually independent samples from the continuous cumulative distribution functions $G_i(x) = F(x - \theta_i)$.

For each $X_{ij}$ let $V_{ij}$ be the “rank” of $X_{ij}$ among all observations excluding observations from the group that $X_{ij}$ belongs, and let $W_{ij}$ be its normalized rank:

$$V_{ij} = \sum_{s \neq i}^{n_s} \sum_{t=1}^{n_s} 1(X_{st} \leq X_{ij})$$  \hfill (1.1)

$$W_{ij} = \frac{V_{ij}}{N - n_i}.$$  \hfill (1.2)

We further let $\bar{V}_i$ be the sample mean rank of group $i$ observations and $\bar{V}_.$ be the sample mean rank of all observations:

$$\bar{V}_i = \frac{1}{n_i} \sum_{j=1}^{n_i} V_{ij},$$  \hfill (1.3)

$$\bar{V}_. = \frac{1}{N} \sum_{i=1}^{k} \sum_{j=1}^{n_i} V_{ij}.$$  \hfill (1.4)

When the sample sizes of all groups go to infinity simultaneously, we have the following Kruskal-Wallis one-way analysis of variance based on linear placements.

**Theorem 1.** If $\theta_1 = \cdots = \theta_k$ and if for each $i$, as $N \to \infty$

$$\frac{n_i}{N} \to \lambda_i, \quad 0 < \lambda_i < 1,$$

then as $N \to \infty$

$$H := \frac{12}{N(N+1)} \sum_{i=1}^{k} n_i (\bar{V}_i - E\bar{V}_i)^2 \Rightarrow \chi^2_{k-1},$$  \hfill (1.5)

where by (2.17), $E\bar{V}_i = (N - n_i)/2$.

In Figure 1 we create 200 samples of $H$ using iid standard normal $X_{ij}, i = 1, \ldots, 5, j = 1, \ldots, n_i$. More specifically, in Figure 1(a) we consider the equal group sizes and in Figure 1(b) we consider the unequal group sizes. In
KRUKAL-WALLIS ONE-WAY ANALYSIS

0.0 5 10 15 20
0 0.2 0.4 0.6 0.8 1
H values

(a) The equal group sizes; \( n_1 = 1010, n_2 = 1000, n_3 = 1020, n_4 = 995 \) and \( n_5 = 1015 \).

(b) The unequal group sizes; \( n_1 = 1010, n_2 = 500, n_3 = 1020, n_4 = 995 \) and \( n_5 = 1015 \).

**Figure 1.** Comparison of the empirical distribution of \( H \) and the Chi-square distribution. We create 200 samples of \( H \) using iid standard normal \( X_{ij}, i = 1, \ldots, 5, j = 1, \ldots, n_i \), from both the equal group sizes (a) and the unequal group sizes (b). In both cases we see that the distribution of \( H \) is very close to the Chi-square distribution.

In both cases we see that the distribution of the Kruskal-Wallis test statistic \( H \) based on linear placements is very close to the Chi-square distribution as we present in Theorem 1.

If we have the equal group sizes, we replace \( E\overline{V}_{/i} \) in \( H \) with \( \overline{V} \) and we still get the Chi-square limit.

**Theorem 2.** If \( \theta_1 = \cdots = \theta_k \) and if for each \( i \), as \( N \to \infty \)

\[(1.6) \quad n_i = \frac{N}{k} + f_i(N), \quad f_i(N) = O(N^{\beta_i}), \quad 0 \leq \beta_i < \frac{1}{2}, \]

then as \( N \to \infty \)

\[(1.7) \quad \hat{H} := \frac{12}{N(N+1)} \sum_{i=1}^{k} n_i(\overline{V}_{/i} - \overline{V})^2 \Rightarrow \chi^2_{k-1}. \]

However, if we don’t have the same group sizes, we cannot replace \( E\overline{V}_{/i} \) in \( H \) with \( \overline{V} \) to get the Chi-square limit.

**Proposition 1.** If \( \theta_1 = \cdots = \theta_k \), if for each \( i \), as \( N \to \infty \)

\[(1.8) \quad n_i = \frac{N}{k} + f_i(N), \quad f_i(N) = O(N^{\beta_i}), \quad 0 \leq \beta_i < 1, \]

and if for some \( i \), as \( N \to \infty \)

\[(1.9) \quad n_i = \frac{N}{k} + f_i(N), \quad f_i(N) = \Omega(N^{\beta_i}), \quad \frac{1}{2} < \beta_i < 1, \]

then as \( N \to \infty \)

\[ E(H - \hat{H}) \to \infty \quad \text{and} \quad \text{Var}(H - \hat{H}) \to \infty. \]
Remark. If in Proposition 1, instead of (1.8) and (1.9) we assume for each \(i\), as \(N \to \infty\)

\[
(1.10) \quad n_i = \frac{N}{k} + f_i(N), \quad f_i(N) = O(N^{\beta_i}), \quad 0 \leq \beta_i \leq \frac{1}{2},
\]

and if for some \(i\), as \(N \to \infty\)

\[
(1.11) \quad n_i = \frac{N}{k} + f_i(N), \quad f_i(N) = \Omega(N^{\beta_i}), \quad \beta_i = \frac{1}{2},
\]

we cannot draw any meaningful conclusion for this borderline case.

**(a)** The equal group sizes; \(n_1 = 1010, n_2 = 1000, n_3 = 1020, n_4 = 995\) and \(n_5 = 1015\).

**(b)** The unequal group sizes; \(n_1 = 1010, n_2 = 500, n_3 = 1020, n_4 = 995\) and \(n_5 = 1015\).

**Figure 2.** Comparison of the empirical distribution of \(\hat{H}\) and the Chi-square distribution. We create 200 samples of \(\hat{H}\) using iid standard normal \(X_{ij}, i = 1, \ldots, 5, j = 1, \ldots, n_i\), from both the equal group sizes (a) and the unequal group sizes (b). In the equal group sizes case (a) we see that the distribution of \(\hat{H}\) is very close to the Chi-square distribution. However, in the unequal group sizes case (b) the distribution of \(\hat{H}\) is far away from the Chi-square distribution.

**Proposition 2.** If \(\theta_1 = \cdots = \theta_k\) and if for each \(i\), as \(N \to \infty\)

\[
(1.12) \quad \frac{n_i}{N} \to \lambda_i, \quad 0 < \lambda_i < 1,
\]

and

\[
(1.13) \quad \lambda_i \neq \frac{1}{k} \quad \text{for some } i,
\]

then as \(N \to \infty\)

\[
E(H - \hat{H}) \to \infty \quad \text{and} \quad Var(H - \hat{H}) \to \infty.
\]
In Figure 2 we create 200 samples of $\hat{H}$ using iid standard normal $X_{ij}$, $i = 1, \ldots, 5$, $j = 1, \ldots, n_i$. More specifically, in Figure 2(a) we consider the equal group sizes and in Figure 2(b) we consider the unequal group sizes. In the equal group sizes case we see that the distribution of the Kruskal-Wallis test statistic $\hat{H}$ based on linear placements is very close to the Chi-square distribution as we present in Theorem 2. However, in the unequal group sizes case the distribution of $\hat{H}$ is far away from the Chi-square distribution as we present in Propositions 1 and 2.

Let $S_N$ be the total sum of the normalized ranks $W_{ij}$, i.e.,

$$S_N = \sum_{i=1}^{k} \sum_{j=1}^{n_i} W_{ij}.$$ \hfill (1.14)

Then we also have the following variant of the central limit theorem of the linear placement statistics of Kim, Lee, and Wang [6].

Theorem 3. If $\theta_1 = \cdots = \theta_k$ and if for each $i$, as $N \to \infty$

$$\frac{n_i}{N} \to \lambda_i, \quad 0 < \lambda_i < 1,$$ \hfill (1.15)

and

$$\lambda_i \neq \frac{1}{k} \quad \text{for some } i,$$ \hfill (1.16)

then as $N \to \infty$

$$\frac{S_N - E(S_N)}{\sqrt{\text{Var}(S_N)}} \Rightarrow N(0,1) \quad \text{in distribution.}$$ \hfill (1.17)

Remark. If in Theorem 3, instead of (1.16) we assume for each $i$

$$\lambda_i = \frac{1}{k},$$ \hfill (1.18)

we cannot prove the CLT. However, even in this case based on the simulation we think the CLT should hold under certain mild conditions.

A little caution is needed here. As we see in (3.1), since in the extremely balanced case $n_1 = \cdots = n_k = N/k$, the sum $\sum_{i=1}^{k} \sum_{j=1}^{n_i} R_{ij}$ is a deterministic constant $N(N+1)/2$, $S_N$ is not random and hence in this case the CLT cannot be true.

In Figure 3 we create 200 samples of $\hat{S}$ using iid standard normal $X_{ij}$, $i = 1, \ldots, 5$, $j = 1, \ldots, n_i$. More specifically, in Figure 3(a) we consider the equal group sizes and in Figure 3(b) we consider the unequal group sizes. In both cases we see that the distribution of $\hat{S}$ is very close to the standard normal distribution as we present in Theorem 3.
We can still consider the generalization of Theorem 3 in line with Kim, Lee, and Wang [6]. For a non-constant $C^1$ function $\varphi : [0, 1] \to \mathbb{R}$ we let

$$S_N^\varphi = \sum_{i=1}^{k} \sum_{j=1}^{n_i} \varphi(W_{ij}).$$

If $\theta_1 = \cdots = \theta_k$ and if for each $i$, as $N \to \infty$

$$\frac{n_i}{N} \to \lambda_i, \quad 0 < \lambda_i < 1,$$

then as $N \to \infty$ we conjecture that

$$\hat{S}^\varphi := \frac{S_N^\varphi - E(S_N^\varphi)}{\sqrt{\text{Var}(S_N^\varphi)}} \Rightarrow \mathcal{N}(0, 1) \text{ in distribution.}$$

In Figure 4 with $\varphi_1(x) = x^2, \varphi_2(x) = \sqrt{x}, \varphi_3(x) = \log(1 + x), \varphi_4(x) = e^x$, we create 200 samples of $\hat{S}^{\varphi_i}$ using iid standard normal $X_{ij}, i = 1, \ldots, 5, j = 1, \ldots, n_i$. More specifically, on the left we consider the equal group sizes and on the right we consider the unequal group sizes. In these equal and unequal cases we see that the distribution of $\hat{S}^{\varphi_i}$ is close to the standard normal distribution as we conjecture. More interestingly, we notice that the distribution of $\hat{S}^{\varphi_i}$ for the unequal group sizes is closer to the standard normal distribution than that for the equal group sizes. This agrees with the fact that in the case $\varphi(x) = x$ of Theorem 3, $S_N$ in the extremely balanced case $n_1 = \cdots = n_k = N/k$, is not random and hence in this case the CLT does not hold. So, for general $\varphi$ we conjecture even further that in the extremely balanced case $n_1 = \cdots = n_k =$
Figure 4. Comparison of the empirical distribution of $\hat{S}_{\varphi_i}$ and the standard normal distribution. With $\varphi_1(x) = x^2$, $\varphi_2(x) = \sqrt{x}$, $\varphi_3(x) = \log(1 + x)$, $\varphi_4(x) = e^x$, we create 200 samples of $\hat{S}_{\varphi_i}$ using iid standard normal $X_{ij}$, $i = 1, \ldots, 5$, $j = 1, \ldots, n_i$, from both the equal group sizes $n_1 = 1010$, $n_2 = 1000$, $n_3 = 1020$, $n_4 = 995$, $n_5 = 1015$ on the left and the unequal group sizes $n_1 = 1010$, $n_2 = 500$, $n_3 = 1020$, $n_4 = 995$, $n_5 = 1015$ on the right. In these equal and unequal cases we see that the distribution of $\hat{S}_{\varphi_i}$ is close to the standard normal distribution. We also note that the distribution of $\hat{S}_{\varphi_i}$ for the unequal group sizes on the right is closer to the standard normal distribution than that for the equal group sizes on the left.

$N/k$, the CLT does not hold and that the convergence rate for the unequal group sizes is greater than that for the equal group sizes.
In Section 2, we prove Kruskal-Wallis one-way analysis of variance based on linear placements, Theorems 1, 2, and Proposition 1 using the tools developed for Kruskal-Wallis one-way analysis of variance based on ranks. In Section 3, we prove the variant of the central limit theorem of the linear placement statistics of Kim, Lee, and Wang [6] using Lyapunov CLT.

2. Kruskal-Wallis one-way analysis of variance based on linear placements

In this section, we prove Kruskal-Wallis one-way analysis of variance based on linear placements, Theorems 1, 2, and Proposition 1 using the tools developed for Kruskal-Wallis one-way analysis of variance based on ranks. To prove this, we start with the following lemma that characterizes the Chi-square distribution.

**Lemma 1.** Let \( X = (X_1, \ldots, X_n)^T \) be an \( n \) dimensional multivariate normal random vector with mean vector \( \mathbf{0} \) and covariance matrix \( \mathbf{B} \). If the covariance matrix \( \mathbf{B} \) satisfies the following two conditions
\[
\mathbf{B}^2 = \mathbf{B} \quad \text{and} \quad \text{rank}(\mathbf{B}) = k,
\]
then \( \sum_{i=1}^{n} X_i^2 \) comes from the Chi-square distribution with the degree of freedom \( k \):
\[
\sum_{i=1}^{n} X_i^2 \sim \chi^2_k.
\]

**Proof.** Since \( \mathbf{B} \) is symmetric, there is an orthogonal matrix \( \mathbf{P} \) such that \( \mathbf{P}^{-1} \mathbf{B} \mathbf{P} = \mathbf{A} \), where \( \mathbf{A} \) is a diagonal matrix. Since \( \text{rank}(\mathbf{B}) = k \), \( \mathbf{A} \) has \( k \) non-zero diagonal elements. By rearranging the coordinates if needed, we may assume that the first \( k \) diagonal elements of \( \mathbf{A} \) are non-zero. Since \( \mathbf{B}^2 = \mathbf{B} \), all the non-zero diagonal elements of \( \mathbf{A} \) are in fact 1.

Since \( X \sim \mathcal{N}_n(\mathbf{0}, \mathbf{B}) \) and since \( \mathbf{P}^{-1} \mathbf{B} \mathbf{P} = \mathbf{A} \), we see that \( \mathbf{Z} := \mathbf{P}^{-1} \mathbf{X} = (Z_1, \ldots, Z_n)^T \) is an \( n \) dimensional multivariate normal random vector with mean vector \( \mathbf{0} \) and covariance matrix \( \mathbf{A} \), i.e., \( \mathbf{Z} \sim \mathcal{N}_n(\mathbf{0}, \mathbf{A}) \). Since the covariance matrix \( \mathbf{A} \) is diagonal and since the first \( k \) diagonal elements of \( \mathbf{A} \) are 1 and all other elements of \( \mathbf{A} \) are 0, \( Z_1, \ldots, Z_k \) are iid standard normal and \( Z_{k+1} = \cdots = Z_n = 0 \). Therefore,
\[
\sum_{i=1}^{n} X_i^2 = \mathbf{X}^T \mathbf{X} = \mathbf{Z}^T \mathbf{P}^{-1} \mathbf{P} \mathbf{Z} = \mathbf{Z}^T \mathbf{Z} = \sum_{i=1}^{k} Z_i^2 \sim \chi^2_k.
\]

**Proof of Theorem 1.** Let \( \mathbf{T} := (T_1, \ldots, T_k)^T \) be given by
\[
(2.1) \quad T_i := \sqrt{\frac{12m_i}{N(N+1)}} (\mathbf{V}_i - E\mathbf{V}_i), \ i = 1, \ldots, k.
\]
By Lemma 1, to prove Theorem 1 it suffices to prove the following three:
\[
(2.2) \quad \mathbf{T} \sim \mathcal{N}_k(\mathbf{0}, \mathbf{B}),
\]
(2.3) \( B = B^2 \),
(2.4) \( \text{rank}(B) = k - 1 \).

To prove these we use the tools developed for Kruskal-Wallis one-way analysis of variance based on ranks. For each \( X_{ij} \) let \( R_{ij} \) be the rank of \( X_{ij} \) among all the observations from all groups, and let \( L_{ij} \) be the rank of \( X_{ij} \) among all the observations from group \( i \):

\[
R_{ij} = \sum_{s=1}^{k} \sum_{t=1}^{n_s} 1(X_{st} \leq X_{ij}),
\]
\[
L_{ij} = \sum_{t=1}^{n_i} 1(X_{it} \leq X_{ij}).
\]

We further let \( \overline{R}_i \) be the sample mean rank of group \( i \) observations and \( \overline{R}_. \) be the sample mean rank of all observations:

\[
(2.5) \quad \overline{R}_i = \frac{1}{n_i} \sum_{j=1}^{n_i} R_{ij} \equiv \frac{1}{n_i} R_i,
\]
\[
(2.6) \quad \overline{R}. = \frac{1}{N} \sum_{i=1}^{k} \sum_{j=1}^{n_i} R_{ij} \equiv \frac{1}{N} R_.,
\]

By the definitions of \( V_{ij}, R_{ij}, \) and \( L_{ij} \), we can rewrite \( V_{ij} \) as

\[
(2.7) \quad V_{ij} = R_{ij} - L_{ij}.
\]

Since

\[
(2.8) \quad \sum_{j=1}^{n_i} L_{ij} = \sum_{j=1}^{n_i} j = \frac{n_i(n_i + 1)}{2},
\]

by (2.7) we see \( V_i - EV_i = R_i - ER_i \), and hence

\[
(2.9) \quad V_i - EV_i = \overline{R}_i - E\overline{R}_i.
\]

So, we can rewrite \( T_i \) as

\[
(2.10) \quad T_i = \sqrt{\frac{12n_i}{N(N+1)}(\overline{R}_i - E\overline{R}_i)}, \quad i = 1, \ldots, k.
\]

Now the standard results on the Kruskal-Wallis one-way analysis of variance based on ranks say that \( T := (T_1, \ldots, T_k)^T \) with \( T_i \) given by (2.10) satisfies (2.2)-(2.4). Therefore, Theorem 1 follows from Lemma 1. \( \square \)

**Lemma 2.** For \( R_{ij}, L_{ij}, \) and \( R_i \), we have

\[
ER_{ij} = \frac{n_i + 1}{2}, \quad \text{Var}(R_{ij}) = \frac{(N-1)(N+1)}{12}, \quad \text{Cov}(R_{ij}, R_{ik}) = -\frac{N+1}{12},
\]
\[
EL_{ij} = \frac{n_i + 1}{2}, \quad \text{Var}(L_{ij}) = \frac{(n_i-1)(n_i+1)}{12}, \quad \text{Cov}(L_{ij}, L_{ik}) = -\frac{N+1}{12},
\]
\[
ER_i = \frac{n_i(N+1)}{2}, \quad \text{Var}(R_i) = \frac{n_i(N-n_i)(N+1)}{12}, \quad \text{Cov}(R_i, R_j) = -\frac{n_i n_j(N+1)}{12}.
\]
Proof. By symmetry $R_{ij}$ is uniform on \{1, 2, \ldots, N\}, i.e., for any observation $X_{ij}$

$$P(R_{ij} = r) = \frac{1}{N} \quad \text{for } 1 \leq r \leq N.$$ 

Also, by symmetry for any two different observations $X_{ij}$ and $X_{st}$

$$P(R_{ij} = r, R_{st} = l) = \frac{1}{N(N - 1)} \quad \text{for } 1 \leq r \neq l \leq N.$$ 

The lemma for $R_{ij}$ and $R_{i.}$ follows from the direct calculation using the above joint probability mass function. By the same reasoning one can get the lemma for $L_{ij}$.

By the definitions of $\hat{H}$ and $H$, we can decompose $\hat{H}$ into $H$ and $C$:

\begin{equation}
\hat{H} = \frac{12}{N(N + 1)} \sum_{i=1}^{k} n_i (\bar{V}_i - \bar{V})^2
\end{equation}

\begin{equation}
= \frac{12}{N(N + 1)} \sum_{i=1}^{k} n_i ((\bar{V}_i - E\bar{V}_i) + (E\bar{V}_i - \bar{V}))(\bar{V}_i - E\bar{V}_i) = H + C,
\end{equation}

where, by (2.9)

\begin{equation}
C := \frac{12}{N(N + 1)} \left[ \sum_{i=1}^{k} n_i (E\bar{V}_i - \bar{V})^2 + \sum_{i=1}^{k} 2n_i (E\bar{V}_i - \bar{V})(\bar{V}_i - E\bar{V}_i) \right]
\end{equation}

\begin{equation}
= \frac{12}{N(N + 1)} \left[ \sum_{i=1}^{k} n_i (E\bar{V}_i - \bar{V})^2 + \sum_{i=1}^{k} 2n_i (E\bar{V}_i - \bar{V})(\bar{R}_i - E\bar{R}_i) \right].
\end{equation}

So by Theorem 1, to prove Theorem 2 it suffices to prove that under the conditions of Theorem 2, $C \to 0$ in probability, or

\begin{equation}
EC \to 0,
\end{equation}

\begin{equation}
Var(C) \to 0,
\end{equation}

and to prove Propositions 1 and 2 we just need to show that under the conditions of Propositions 1 and 2,

\begin{equation}
EC \to \infty,
\end{equation}

\begin{equation}
Var(C) \to \infty.
\end{equation}

By (2.7), (2.8) and Lemma 2,

\begin{equation}
E\bar{V}_i = \frac{1}{n_i} \left[ E \left( \sum_{j=1}^{n_i} R_{ij} \right) - E \left( \sum_{j=1}^{n_i} L_{ij} \right) \right] = \frac{N - n_i}{2}.
\end{equation}
By (2.17), (2.7) and (2.8), we see that $E\overline{V}_i - \overline{V}_\cdot$ is actually a deterministic constant:

$$E\overline{V}_i - \overline{V}_\cdot = \frac{N - n_i}{2} - \frac{1}{N} \sum_{s=1}^{k} \sum_{t=1}^{n_s} V_{st} = \frac{1}{N} \sum_{s=1}^{k} \frac{n_s^2}{2} - \frac{n_i}{2}, \quad (2.18)$$

By (2.12) and (2.18), the random part of $C$ is just $\frac{12}{N(N+1)} \sum_{i=1}^{k} 2n_i (E\overline{V}_i - \overline{V}_\cdot)R_i$. Again, by (2.18) we can rewrite the major part $\sum_{i=1}^{k} 2n_i (E\overline{V}_i - \overline{V}_\cdot)R_i$ of the random part of $C$ as

$$\sum_{i=1}^{k} 2n_i (E\overline{V}_i - \overline{V}_\cdot)R_i = \sum_{i=1}^{k} \frac{1}{N} \sum_{s=1}^{k} n_s^2 - n_i \frac{R_i}{N}, \quad (2.19)$$

Take the expectation and variance on the both sides of (2.19). Then, by Lemma 2 we have

$$E \left( \sum_{i=1}^{k} 2n_i (E\overline{V}_i - \overline{V}_\cdot)R_i \right) = \sum_{i=1}^{k} \frac{1}{N} \sum_{s=1}^{k} n_s^2 - n_i \cdot E(R_i) = 0, \quad (2.20)$$

and

$$\text{Var} \left( \sum_{i=1}^{k} 2n_i (E\overline{V}_i - \overline{V}_\cdot)R_i \right) = \text{Var} \left( \sum_{i=1}^{k} \left( \frac{1}{N} \sum_{s=1}^{k} n_s^2 - n_i \right) R_i \right) = \frac{N+1}{12} \left[ N \sum_{i=1}^{k} n_i^3 - \left( \sum_{s=1}^{k} n_s^2 \right)^2 \right]. \quad (2.21)$$

So, by (2.12), (2.18), (2.20), and (2.21) we have

$$E(C) = 3f(N), \quad (2.22)$$

$$\text{Var}(C) = 12f(N), \quad (2.23)$$

where

$$f(N) := \frac{1}{N^2(N+1)} \left[ N \sum_{i=1}^{k} n_i^3 - \left( \sum_{s=1}^{k} n_s^2 \right)^2 \right]. \quad (2.24)$$

Since $n_i = \frac{N}{k} + f_i(N)$ and since $\sum_{i=1}^{k} f_i(N) = 0$, we have

$$f(N) = \frac{1}{N^2(N+1)} \left[ N \sum_{i=1}^{k} \left( \frac{N}{k} + f_i(N) \right)^3 - \left( \sum_{i=1}^{k} \left( \frac{N}{k} + f_i(N) \right)^2 \right)^2 \right]. \quad (2.25)$$
\[
\frac{1}{N(N+1)} \sum_{i=1}^{k} f_i^3(N) + \frac{1}{k(N+1)} \sum_{i=1}^{k} f_i^2(N) - \frac{1}{N^2(N+1)} \left( \sum_{i=1}^{k} f_i^2(N) \right)^2.
\]

**Proof of Theorem 2.** By dropping the last negative term in the right hand side of (2.25), we have

(2.26) \[ f(N) \leq \frac{1}{N(N+1)} \sum_{i=1}^{k} f_i^3(N) + \frac{1}{k(N+1)} \sum_{i=1}^{k} f_i^2(N). \]

If by (1.6) for all \( i, f_i(N) = O(N^\beta), 0 \leq \beta < \frac{1}{2} \), by (2.26) as \( N \to \infty \)

\[ f(N) \to 0, \]

and hence by (2.22)-(2.24) both the mean and the variance of \( C \) go to 0. Therefore, by (2.13) and (2.14) Theorem 2 follows. \( \square \)

**Proof of Proposition 1.** By (1.12) the first term in the right hand side of (2.25) is bounded by \( \frac{1}{N} \sum_{i=1}^{k} f_i^2(N) \):

(2.27) \[ \frac{1}{N(N+1)} \sum_{i=1}^{k} f_i^3(N) = o \left( \frac{1}{N} \sum_{i=1}^{k} f_i^2(N) \right). \]

Again by (1.12)

(2.28) \[ \frac{1}{k} - \frac{1}{N^2} \sum_{i=1}^{k} f_i^2(N) \to \frac{1}{k}. \]

So, by (2.28) among the second and third terms in the right hand side of (2.25) the second is the major term;

(2.29) \[
\frac{1}{k(N+1)} \sum_{i=1}^{k} f_i^2(N) - \frac{1}{N^2(N+1)} \left( \sum_{i=1}^{k} f_i^2(N) \right)^2 = \left( \frac{1}{N+1} \sum_{i=1}^{k} f_i^2(N) \right) \left( \frac{1}{k} - \frac{1}{N^2} \sum_{i=1}^{k} f_i^2(N) \right).
\]

Combining (2.27), (2.28), and (2.29) we see that \( f(N) \) is of order \( \frac{1}{N} \sum_{i=1}^{k} f_i^2(N) \):

(2.30) \[ f(N) = \Omega \left( \frac{1}{N} \sum_{i=1}^{k} f_i^2(N) \right). \]

If by (1.9) for some \( i, f_i(N) = \Omega(N^{\beta_i}), \frac{1}{2} < \beta_i < 1 \), by (2.30) as \( N \to \infty \)

\[ f(N) \to \infty, \]
and hence by (2.22)-(2.24) both the mean and the variance of $C$ go to $\infty$. Therefore, by (2.15) and (2.16) Proposition 1 follows.

Proof of Proposition 2. Under the conditions (1.12) and (1.13) we have unequal $0 < \lambda_i < 1$ with $\sum_{i=1}^k \lambda_i = 1$. For these $\lambda_i$ we consider a random variable $X$ with

$$P(X = \lambda_i) = \lambda_i \quad \text{for } 1 \leq i \leq k.$$ 

Let’s calculate its variance. Since $EX = \sum_{i=1}^k \lambda_i^2$ and since $EX^2 = \sum_{i=1}^k \lambda_i^3$, we have

$$\text{Var}(X) = EX^2 - (EX)^2 = \sum_{i=1}^k \lambda_i^3 - \left(\sum_{i=1}^k \lambda_i^2\right)^2. \quad (2.31)$$ 

Since $X$ is not a constant but finite random variable, it has a non-zero but finite variance. Therefore, by (2.31) we see that, for unequal $0 < \lambda_i < 1$ with $\sum_{i=1}^k \lambda_i = 1$

$$0 < \sum_{i=1}^k \lambda_i^3 - \left(\sum_{i=1}^k \lambda_i^2\right)^2 < \infty. \quad (2.32)$$

By (2.24) and (2.32), as $N \to \infty$

$$f(N) = N \left[ \sum_{i=1}^k \lambda_i^3 - \left(\sum_{i=1}^k \lambda_i^2\right)^2 + o(1) \right] \to \infty,$$

and hence by (2.22)-(2.24) both the mean and the variance of $C$ go to $\infty$. Therefore, by (2.15) and (2.16) Proposition 2 follows. \qed

3. Proof of Theorem 3

In this section we use Lyapunov’s CLT and prove Theorem 3. By (2.7) and (2.8) we can rewrite $S_N$ as

$$S_N = \sum_{i=1}^k \sum_{j=1}^{n_i} \frac{R_{ij}}{N - n_i + 1} - \sum_{i=1}^k \frac{n_i(n_i + 1)}{2(N - n_i)}. \quad (3.1)$$

Since $\{U_{ij} := \frac{R_{ij}}{N - n_i + 1}\}$ have the following (joint) probability mass function

$$P \left( U_{ij} = \frac{k}{N + 1} \right) = \frac{1}{N} \quad \text{for } 1 \leq k \leq N,$$

$$P \left( U_{ij} = \frac{k}{N + 1}, U_{ij'} = \frac{k'}{N + 1} \right) = \frac{1}{N(N - 1)} \quad \text{for } 1 \leq k \neq k' \leq N,$$

$\{U_{ij}\}$ have the following mean, variance, and covariance

$$EU_{ij} = \frac{1}{2}. \quad (3.2)$$
(3.3) \[ \text{Var}(U_{ij}) = \frac{2N + 1}{6N + 6} - \frac{1}{4} - \frac{1}{6}N^{-1} + O(N^{-2}), \]

(3.4) \[ \text{Cov}(U_{ij}, U_{i'j'}) = \frac{3N^2 - N - 2}{12(N + 1)(N - 1)} - \frac{1}{4} = -\frac{1}{12}N^{-1} + O(N^{-2}). \]

Since \( \frac{N+1}{N-n_i} \to \frac{1}{1-\lambda_i} \), and since \( \sum_{i=1}^{k} \frac{n_i(n_i+1)}{2N(N-n_i)} \) is non-random, we see that the asymptotics of \( S_N \) is the same as that of \( T_N \), where \( T_N \) is defined as

(3.5) \[ T_N = \sum_{i=1}^{k} \sum_{j=1}^{n_i} A_i U_{ij}, \]

where

(3.6) \[ A_i = \frac{1}{1 - \lambda_i}. \]

By (3.3), (3.4), and (3.6),

(3.7) \[ \text{Var}(T_N) = \text{Var} \left( \sum_{i=1}^{k} \sum_{j=1}^{n_i} A_i U_{ij} \right) \]

\[ = \sum_{ij} A_i^2 \left( \frac{1}{12} - \frac{1}{6}N^{-1} + O(N^{-2}) \right) \]

\[ - \sum_{ij} \sum_{i'j' \neq ij} A_i A_{i'} \left( \frac{1}{12}N^{-1} + O(N^{-2}) \right) \]

\[ = \frac{N}{12} \left[ \sum_{i=1}^{k} \frac{\lambda_i}{(1 - \lambda_i)^2} - \sum_{i=1}^{k} \sum_{i' \neq i} \frac{\lambda_i \lambda_{i'}}{(1 - \lambda_i)(1 - \lambda_{i'})} + O(N^{-1}) \right] \]

\[ = \frac{N}{12} \left[ \sum_{i=1}^{k} \frac{\lambda_i}{(1 - \lambda_i)^2} - \left( \sum_{i=1}^{k} \frac{\lambda_i}{1 - \lambda_i} \right)^2 + O(N^{-1}) \right]. \]

Under the conditions (1.15) and (1.16) we have unequal \( 0 < \lambda_i < 1 \) with \( \sum_{i=1}^{k} \lambda_i = 1 \). For these \( \lambda_i \), we consider a random variable \( Y \) with

\[ P \left( Y = \frac{1}{1 - \lambda_i} \right) = \lambda_i \quad \text{for} \quad 1 \leq i \leq k. \]

Let's calculate its variance. Since \( EY = \sum_{i=1}^{k} \lambda_i/(1 - \lambda_i) \) and since \( EY^2 = \sum_{i=1}^{k} \lambda_i/(1 - \lambda_i)^2 \), we have

(3.8) \[ \text{Var}(Y) = EY^2 - (EY)^2 = \sum_{i=1}^{k} \frac{\lambda_i}{(1 - \lambda_i)^2} - \left( \sum_{i=1}^{k} \frac{\lambda_i}{1 - \lambda_i} \right)^2. \]

Since by (1.16), \( Y \) is not a constant but finite random variable, it has a non-zero but finite variance. Therefore, by (3.8) we see that, for unequal \( 0 < \lambda_i < 1 \)
with \( \sum_{i=1}^{k} \lambda_i = 1 \)
\[
0 < \sum_{i=1}^{k} \frac{\lambda_i}{(1 - \lambda_i)^2} - \left( \sum_{i=1}^{k} \frac{\lambda_i}{1 - \lambda_i} \right)^2 < \infty.
\]

Therefore, by (3.7) and (3.9) we have
\[
(3.10) \quad \text{Var}(T_N) = \Omega(N).
\]

Since \( 0 \leq U_{ij} \leq 1 \), for any fixed \( \delta > 0 \), \( E |U_{ij} - EU_{ij}|^{2+\delta} \) is bounded by 1. So,
\[
(3.11) \quad \sum_{i=1}^{k} \sum_{j=1}^{n_i} E |A_iU_{ij} - EA_iU_{ij}|^{2+\delta}
= \sum_{i=1}^{k} \sum_{j=1}^{n_i} A_i^{2+\delta} E |U_{ij} - EU_{ij}|^{2+\delta}
\leq \sum_{i=1}^{k} n_i A_i^{2+\delta}
\leq N \left[ \sum_{i=1}^{k} \frac{\lambda_i}{(1 - \lambda_i)^{2+\delta}} + o(1) \right].
\]

By (3.10) and (3.11) one can easily check the Lyapunov condition;
\[
\lim_{n \to \infty} \left( \text{Var}(T_N) \right)^{-\frac{1}{2}} \sum_{i=1}^{k} \sum_{j=1}^{n_i} E |A_iU_{ij} - EA_iU_{ij}|^{2+\delta} \to \infty.
\]

Hence, by Lyapunov’s CLT Theorem 3 follows.

**References**


Yicheng Hong
Department of Mathematics
Yanbian University
Yanj 133002, P. R. China
and
Department of Mathematics
Yonsei University
Seoul 120-749, Korea
E-mail address: tophych@gmail.com

Sungchul Lee
Department of Mathematics
Yonsei University
Seoul 120-749, Korea
E-mail address: sungchul@yonsei.ac.kr