U-Statistics and Imperfect Ranking in Ranked Set Sampling

Brett Presnell[1]

Centre for Mathematics and Its Applications

Australian National University

Lora L. Bohn

Department of Statistics

University of Florida

July 1997

Revised February 1998

[1]On leave from Department of Statistics, University of Florida

Key words and phrases. Ranked set sample, judgement ranking, U-statistic, asymptotic efficiency, majorization.

AMS 1991 subject classifications. Primary: 62G20. Secondary: 62E20, 62G30.

Abstract

Ranked set sampling has attracted considerable attention as an efficient sampling design, particularly for environmental and ecological studies. A number of authors have noted a gain in efficiency over ordinary random sampling when specific estimators and tests of hypotheses are applied to rank set sample data. We generalize such results by deriving the asymptotic distribution for random sample U-statistics when applied to ranked set sample data. Our results show that the ranked set sample procedure is asymptotically at least as efficient as the random sample procedure, regardless of the accuracy of judgement ranking. Some errors in the ranked set sampling literature are also revealed, and counterexamples provided. Finally, application of majorization theory to these results shows when perfect ranking can be expected to yield greater efficiency than imperfect ranking.

1 Introduction

In ranked set sampling, the items in each of k independent random samples of size k from some fixed population are separately subjected to judgement ranking. From the rth sample, $r = 1, \ldots, k$, the item judged to be the rth smallest is retained and measured. If the ranking is perfect, then the observation retained from the rth sample is the rth order statistic from that sample, but for imperfect judgement ranking the observation retained may be quite arbitrary. This entire process is repeated for m independent cycles. The data then consist of mk independent observations, with one item retained and measured from each of the mk independent k-samples. The observations retained from the ith cycle are denoted $\mathbf{X}_i = X_{[1]i}, \ldots, X_{[k]i}$, and the entire ranked set sample (RSS) is denoted $\mathbf{X}_1, \ldots, \mathbf{X}_m = X_{[1]1}, \ldots, X_{[k]1}, \ldots, X_{[1]m}, \ldots, X_{[k]m}$. The distribution function of the original population of X-measurements is denoted F. In the two-sample case, $\mathbf{Y}_1, \ldots, \mathbf{Y}_n = Y_{[1]1}, \ldots, Y_{[l]1}, \ldots, Y_{[1]n}, \ldots, Y_{[l]n}$ will denote an independent ranked set sample of size nl from a population with distribution G.

Judgement rankings may depend on the (unmeasured) values of the items in the block, as well as other possible covariates or randomness, but the process is assumed to be stochastically identical and independent from block to block. In practice, “ranker” and other biases which might otherwise violate this assumption can be minimized by blinding the ranker to a restricted randomization determining the judgement rank of the item to be measured within each block of k items [1].

Ranked set sampling is usually suggested in situations where ranking of items is easy relative to the cost or difficulty of measurement. A number of estimators and test statistics, including the sample mean [1], the sample variance [2], the empirical distribution function [3], the sign statistic [4], Mann-Whitney statistic [5, 6], and the signed-rank statistic [7], have been shown to gain in efficiency when applied to ranked set sample data, as opposed to the same number of i.i.d. observations.

When applied to i.i.d. data, each of the aforementioned statistics are U-statistics. Denote by U_{RSS} the value of a U-statistic when evaluated over ranked set data, and by U its value when evaluated over the same number of i.i.d. observations. In Section 2, we give the asymptotic distribution of U_{RSS} in the one- and two-sample (nondegenerate) cases, and explicit formulae for the asymptotic relative efficiency of U_{RSS} to U. These results are valid regardless of the accuracy of the judgement ranking process, and generalize previous asymptotic results for ranked set sampling.

As demonstrated by these theorems, powerful general results can be proven for quite arbitrary judgement ranking schemes. However, there are some errors in the literature relating to imperfect ranking, as shown in Section 3. In addition, remarks in the literature suggesting that perfect ranking leads to more efficient estimates than imperfect ranking do not hold in general. In Section 4, we use majorization to derive a result in the positive direction that holds for a broad class of "monotonic" U-statistics.

The construction of more efficient estimates tailored specifically to ranked set sampling is discussed by Kvam and Samaniego [8] and, in the more general setting of stratified random sampling, by Taga et al. [9] and Yanagawa [10]. Though we do not address this important issue, some connections between our work and that of Taga et al. [9] are discussed in Sections 2.2 and 4.

2 U-Statistics and Ranked Set Sample Data

2.1 A Key Lemma.

For $r = 1, \dots, k$, denote by $F_{[r]}(\cdot)$ the marginal distribution function of the rth judgement ordered statistic, $X_{[r]}$, and by $F_{(r)}(\cdot)$ the marginal distribution function of the rth order statistic, $X_{(r)}$, in a random sample of size k from F (continuous). The following lemma and its corollary Lemma 2 are crucial for proving both our results and those of most other papers concerned with ranked set sampling. The result is obvious if ranking is perfect, in which case $F_{[r]} = F_{(r)}$, but holds for *any* judgement ranking scheme, a fact first noted by Dell and Clutter [1]. Unfortunately their discussion of the result is somewhat unclear, and no explicit proof is available in the literature. The argument is straightforward, but because errors have arisen in the literature in connection with this lemma (see Section 3), we provide a detailed proof here.

Lemma 1.

$$\frac{1}{k}\sum_{j=1}^{k} F_{[j]}(x) = F(x) \quad \forall\, x\,. \tag{1}$$

Proof. Let $X_1, \ldots, X_k \sim$ i.i.d.F, let $X_{(1)} < \cdots < X_{(k)}$ be the order statistics, and let $X_{[1]}, \ldots, X_{[k]}$ be the judgement-ordered observations. Also, for $j = 1, \ldots, k$, let $J_j = i$ if the ith order statistic is judged to be the jth smallest observation, i.e., if $X_{[j]} = X_{(i)}$. Then clearly,

$$F_{[j]}(x) = \sum_{i=1}^{k} P(X_{[j]} \le x, J_j = i) = \sum_{i=1}^{k} P(X_{(i)} \le x, J_j = i)\,. \tag{2}$$

But for each fixed i, the events $\{J_j = i\}$, $j = 1, \ldots, k$, partition the sample space, because the ith order statistic receives one and only one judgement ranking. Thus,

$$F_{(i)}(x) = P(X_{(i)} \le x) = \sum_{j=1}^{k} P(X_{(i)} \le x, J_j = i)\,. \tag{3}$$

Combining (2) and (3) with an exchange in order of summation yields the result:

$$\sum_{j=1}^{k} F_{[j]}(x) = \sum_{j=1}^{k}\sum_{i=1}^{k} P(X_{(i)} \le x, J_j = i) = \sum_{i=1}^{k} F_{(i)}(x) = kF(x)\,. \quad \square$$

The following corollary is used extensively in proving Theorems 1 and 2.

Lemma 2. *For any function $g(x_1, \ldots, x_d)$, $d \le m$, for which the expectation on the left hand side exists,*

$$E[g(X_1, \ldots, X_d)] = k^{-d} \sum_{r_1=1}^{k} \cdots \sum_{r_d=1}^{k} E[g(X_{[r_1]1}, \ldots, X_{[r_d]d})]\,.$$

2.2 One-Sample U-Statistics.

Suppose that $h(x_1, \ldots, x_p)$ is a symmetric kernel of degree p for the parameter $\theta = \theta(F)$, and that U is the corresponding U-statistic. That is, for i.i.d. observations $X_1, \ldots, X_{mk}$

from F, $E[h(X_1, \ldots, X_p)] = \theta$, and

$$U = U(X_1, \ldots, X_{mk}) = \frac{1}{\binom{mk}{p}} \sum_{\boldsymbol{\alpha} \in \mathcal{A}} h(X_{\alpha_1}, \ldots, X_{\alpha_p}),$$

where $\mathcal{A}$ is the collection of all subsets of size p chosen from the integers $1, \ldots, mk$. Following closely the notation of Randles and Wolfe [11], let

$$h_1(x) = E[h(x, X_2, \ldots, X_p)] \tag{4}$$

and

$$\zeta_1 = \operatorname{Var}[h_1(X_1)]\,. \tag{5}$$

If $E[h^2(X_1, \ldots, X_p)] < \infty$ and $\zeta_1 > 0$, then U is a nondegenerate U-statistic and the One-Sample U-Statistic Theorem states that $\sqrt{mk}[U - \theta]$ is asymptotically normally distributed as $m \to \infty$, with mean 0 and variance $p^2\zeta_1$. In Theorem 1 we provide an analogous result when U is evaluated using the mk observations from a RSS, where the set size k is taken to be at least as large as the degree p.

Theorem 1. *Assume that $k \geq p$ and let*

$$U_{RSS} = U(\mathbf{X}_1, \ldots, \mathbf{X}_m) = U(X_{[1]1}, \ldots, X_{[k]1}, \ldots, X_{[1]m}, \ldots, X_{[k]m})\,.$$

For $r = 1, \ldots, k$, define

$$\gamma_r = E[h_1(X_{[r]1})] \qquad \textit{and} \qquad \xi_r = \operatorname{Var}[h_1(X_{[r]1})]\,.$$

If both $\operatorname{Var}[h(X_1, \ldots, X_p)] < \infty$ and $\sum_{r=1}^k \xi_r > 0$, then $\sqrt{mk}[U_{RSS} - \theta]$ is asymptotically normally distributed as $m \to \infty$ (with k and p fixed), with mean 0 and variance $p^2 k\zeta_1^*$, where

$$k\zeta_1^* = \frac{1}{k}\sum_{r=1}^{k} \xi_r = \zeta_1 - \frac{1}{k}\sum_{r=1}^{k}(\gamma_r - \theta)^2\,. \tag{6}$$

Corollary 1. *The asymptotic efficiency of U_{RSS} relative to U is*

$$\mathrm{ARE}(U_{RSS}, U) = 1 + \frac{\sum_{r=1}^{k} (\gamma_r - \theta)^2}{\sum_{r=1}^{k} \xi_r} .$$

These results generalize those found in the literature for the mean [1], the sign statistic [4], the empirical distribution function [3], the variance [2], and the signed-rank statistic [7]. Corollary 1 of course shows that $\mathrm{ARE}(U_{\mathrm{RSS}}, U) \geq 1$ regardless of the accuracy of the judgement ranking process.

The proof of Theorem 1 given in the appendix shows that U_{RSS} is asymptotically equivalent to $\tilde{U}$, where $\tilde{U}$ is a U-statistic in the vectors $\mathbf{X}_1, \ldots, \mathbf{X}_m$ with expectation θ. If $p = 1$, as is the case for the mean, the sign statistic, and the empirical distribution function, then $U_{\mathrm{RSS}} = \tilde{U}$ and U_{RSS} is unbiased with exact variance $p^2 k \zeta_1^* / (mk)$.

Interestingly, the limiting distribution of U_{RSS} is identical to that of U^*, the generalized U-statistic of Taga et al. [9], though U^* is quite different in form and in the derivation of its asymptotic distribution from both U_{RSS} and $\tilde{U}$. Note that U^* is the UMVUE of θ over the class of all absolutely continuous distributions [9, Theorem 3.1].

2.3 Multi-Sample U-Statistics.

We present results only for the two-sample case, but the obvious generalization to multiple samples holds (for analogous results under ordinary random sampling see Theorems 3.4.13 and 3.6.9 of Randles and Wolfe [11]). Assume that $h(x_1, \ldots, x_p; y_1, \ldots, y_q)$ is a symmetric kernel of degree (p, q) for the parameter θ. Let k and l be fixed constants. For independent random samples $X_1, \ldots, X_{mk}$ from F and $Y_1, \ldots, Y_{nl}$ from G, the corresponding two-sample U-statistic for θ is

$$\begin{aligned} U &= U(X_1, \ldots, X_{mk}; Y_1, \ldots, Y_{nl}) \\ &= \frac{1}{\binom{mk}{p}\binom{nl}{q}} \sum_{\boldsymbol{\alpha} \in \mathcal{A}} \sum_{\boldsymbol{\beta} \in \mathcal{B}} h(X_{\alpha_1}, \ldots, X_{\alpha_p}; Y_{\beta_1}, \ldots, Y_{\beta_q}) , \end{aligned}$$

where $\mathcal{A}$ ($\mathcal{B}$) is the collection of all subsets of size p (q) chosen from the integers $1, \ldots, mk$ ($1, \ldots, nl$).

Again following the notation of Randles and Wolfe [11], let

$$\begin{aligned} h_{10}(x) &= E[h(x, X_2, \ldots, X_p; Y_1, \ldots, Y_q)]\,, \\ h_{01}(y) &= E[h(X_1, \ldots, X_p; y, Y_2, \ldots, Y_q)]\,, \end{aligned} \tag{7}$$

and

$$\zeta_{10} = \mathrm{Var}[h_{10}(X_1)]\,, \qquad \zeta_{01} = \mathrm{Var}[h_{01}(Y_1)]\,. \tag{8}$$

Further, let $N = mk + nl$, and assume that $m, n \to \infty$ in such a way that $mk/N \to \lambda$, where $0 < \lambda < 1$. Then if $E[h^2(X_1, \ldots, X_p; Y_1, \ldots, Y_q)] < \infty$ and $\zeta_{10} + \zeta_{01} > 0$, the Two-Sample U-Statistic Theorem implies that $\sqrt{N}[U - \theta]$ is asymptotically normally distributed with mean 0 and variance $[p^2\zeta_{10}/\lambda] + [q^2\zeta_{01}/(1-\lambda)]$. Theorem 2 provides the analogous result when U is evaluated using the observations from independent RSS's.

Theorem 2. *Assume that $k \geq p$ and $l \geq q$, and let $U_{RSS} = U(\mathbf{X}_1, \ldots, \mathbf{X}_m; \mathbf{Y}_1, \ldots, \mathbf{Y}_n)$. For $r = 1, \ldots, k$ and $s = 1, \ldots, l$, define*

$$\begin{aligned} \gamma_{r\cdot} &= E[h_{10}(X_{[r]1})]\,, &\qquad \xi_{r\cdot} &= \mathrm{Var}[h_{10}(X_{[r]1})]\,, \\ \gamma_{\cdot s} &= E[h_{01}(Y_{[s]1})]\,, &\qquad \xi_{\cdot s} &= \mathrm{Var}[h_{01}(Y_{[s]1})]\,, \end{aligned}$$

If $E[h^2(X_1, \ldots, X_p; Y_1, \ldots, Y_q)] < \infty$ and $\sum_{r=1}^{k} \xi_{r\cdot} + \sum_{s=1}^{l} \xi_{\cdot s} > 0$, then $\sqrt{N}[U_{RSS} - \theta]$ has a limiting normal distribution as $m, n \to \infty$ (with k, l, p, and q fixed), with mean 0 and variance $[p^2 k\zeta_{10}^*/\lambda] + [q^2 l\zeta_{01}^*/(1-\lambda)]$, where

$$\begin{aligned} k\zeta_{10}^* &= \frac{1}{k}\sum_{r=1}^{k} \xi_{r\cdot} = \zeta_{10} - \frac{1}{k}\sum_{r=1}^{k}(\gamma_{r\cdot} - \theta)^2 \\ l\zeta_{01}^* &= \frac{1}{l}\sum_{s=1}^{l} \xi_{\cdot s} = \zeta_{01} - \frac{1}{l}\sum_{s=1}^{l}(\gamma_{\cdot s} - \theta)^2\,. \end{aligned} \tag{9}$$

Corollary 2. *The asymptotic relative efficiency of* U_{RSS} *to* U *is*

$$\mathrm{ARE}(U_{RSS}, U) = 1 + \frac{[(1-\lambda)p^2/k]\sum_{r=1}^{k}(\gamma_{r\cdot} - \theta)^2 + [\lambda q^2/l]\sum_{s=1}^{l}(\gamma_{\cdot s} - \theta)^2}{[(1-\lambda)p^2/k]\sum_{r=1}^{k}\xi_{r\cdot} + [\lambda q^2/l]\sum_{s=1}^{l}\xi_{\cdot s}}.$$

These results apply for example to the Mann-Whitney statistic [5, 6]. Clearly, as in the one-sample case, $\mathrm{ARE}(U_{\mathrm{RSS}}, U) \geq 1$.

3 Examples and Counterexamples.

As noted in Section 2.1, confusion surrounding the proof of Lemma 1 has lead to some errors in the literature. In both Bohn and Wolfe [6] and Hettmansperger [4], it is claimed that

$$F_{[j]}(x) = \sum_{i=1}^{k} p_{ij} F_{(i)}(x), \tag{10}$$

where $p_{ij} = P(J_j = i)$. Comparing with (2), we see that this equation arises from the assumption that $P(X_{(i)} \leq x, J_j = i) = P(X_{(i)} \leq x)P(J_j = i)$, i.e., that the value of the ith order statistic and the event that it receives judgement rank j are independent. However this is not generally true for models which properly incorporate the notion that observations whose values are nearly the same are more difficult to rank correctly than those whose values are clearly separated.

In such a model the probability of correct ranking must depend on the configuration of the data at hand, since it is the actual realized values of the order statistics (and not just their expected values for example) which directly determine the accuracy of judgement ranking. Thus realistic models for judgement ranking may be described most naturally by conditioning on the unobserved (or indirectly observed) measurements, $X_1, \ldots, X_k$. Note that this does not in any way imply that the observer knows the true measured values of the observations and then perversely insists on using a random mechanism to rank them. Instead conditioning is merely a probabilistic device for modeling the dependence of the judgement ranks on the actual unmeasured data.

We present two counterexamples to (10). In both examples the conditional probability of correctly ranking two observations is close to 1/2 when the observations are very close together, while the conditional probability of correctly ranking widely separated observations is close to 1. Details of the calculations are available from the authors.

Example 1. Suppose that $k = 2$, and that $X_1, X_2 \sim$ i.i.d. $U(0,1)$. We model the judgement ranking scheme by assuming that conditional on X_1, X_2,

$$J_1 = 1 \quad \text{with probability} \quad (1/2)\big[1 + (X_{(2)} - X_{(1)})\big] \,. \tag{11}$$

For this model, $F_{[1]}(1/2) = 5/8$, while

$$\begin{bmatrix} p_{11} & p_{12} \\ p_{21} & p_{22} \end{bmatrix} = \begin{bmatrix} 2/3 & 1/3 \\ 1/3 & 2/3 \end{bmatrix} , \tag{12}$$

so that $p_{11}F_{(1)}(1/2) + p_{21}F_{(2)}(1/2) = 7/12$, disproving (10).

The scheme can be extended in a natural way to nonuniformly distributed random variables by replacing $X_{(1)}$ and $X_{(2)}$ in (11) by $F(X_{(1)})$ and $F(X_{(2)})$. A further extension replaces $\big(F(X_{(2)}) - F(X_{(1)})\big)$ in (11) by $\big(F(X_{(2)}) - F(X_{(1)})\big)^\rho$ for some $\rho > 0$. In this case, smaller values of ρ imply more accurate judgement ranking.

Example 2. Suppose again that $k = 2$ and that $X_1, X_2 \sim$ i.i.d. $N(0,1)$. We model the judgement ranking scheme by assuming that the observer's perception of X_1, X_2 is clouded by noise. Specifically we assume that the judgement ranks assigned to X_1 and X_2 are actually the joint ranks of T_1 and T_2, where $T_i = X_i + \epsilon_i$, $i = 1, 2$, with $\epsilon_1, \epsilon_2 \sim$ i.i.d. $N(0, \sigma^2)$, independently of X_1, X_2. Thus smaller values of σ imply more accurate judgement ranking. This model was used in simulation studies by Dell and Clutter [1], and could also be defined by stating that conditional on X_1, X_2,

$$J_1 = 1 \quad \text{with probability} \quad \Phi\left(\frac{X_{(2)} - X_{(1)}}{\sqrt{2}\,\sigma}\right),$$

where Φ is the standard normal distribution function.

In this example $F_{(1)}(x) = 1 - [1 - \Phi(x)]^2$ and $F_{(2)}(x) = [\Phi(x)]^2$, and some calculus shows that

$$F_{[1]}(x) = 2\int_{-\infty}^{x} \left[1 - \Phi\left(z/\sqrt{1+2\sigma^2}\right)\right]\phi(z)\,dz\,,$$
$$F_{[2]}(x) = 2\int_{-\infty}^{x} \Phi\left(z/\sqrt{1+2\sigma^2}\right)\phi(z)\,dz\,,$$

and $p_{11} = P(J_1 = 1) = 1 - \arctan(\sigma)/\pi$. With these results it is apparent that $F_{[1]}(x) \neq p_{11}F_{(1)}(x) + p_{21}F_{(2)}(x)$. For example, $F_{[1]}(0) = 1 - \arctan(\sqrt{1+2\sigma^2})/\pi$, while $p_{11}F_{(1)}(0) + p_{21}F_{(2)}(0) = 3/4 - \arctan(\sigma)/(2\pi)$. While these expressions are equal when $\sigma = 0$ (perfect ranking) and in the limit as $\sigma \to \infty$ (completely random ranking), they do not agree for any $0 < \sigma < \infty$.

Example 3 (The Sign Test). In Hettmansperger [4], the null hypothesis variance of the sign statistic under ranked set sampling with judgement ranking is given by $nk\delta_p^2/4$, where

$$\delta_p^2 = 1 - \frac{4}{k}\sum_{j=1}^{k}\left\{\sum_{i=1}^{k} p_{ij}\left(F_{(i)}(0) - \frac{1}{2}\right)\right\}^2 . \tag{H5}$$

Because of the error described above, this expression and its corollary, Theorem 3, are false for general judgement ranking schemes. (Theorem 2 of Hettmansperger [4] is correct however; see Section 4.) The correct expression is

$$\delta_p^2 = 1 - \frac{4}{k}\sum_{j=1}^{k}\left\{F_{[j]}(0) - \frac{1}{2}\right\}^2 . \tag{H5$'$}$$

In Example 1 (after recentering the uniform distribution at 0), (H5) yields $\delta_p^2 = 35/36$, while the correct value from (H5′) is $\delta_p^2 = 15/16$. Similarly, in Example 2, expression (H5) yields $\delta_p^2 = 1 - [1/2 - \arctan(\sigma)/\pi]^2$, while the correct expression from (H5′) is $\delta_p^2 = 1 - [1 - 2\arctan(\sqrt{1+2\sigma^2})/\pi]^2$. Again these agree at $\sigma = 0$ and in the limit as $\sigma \to \infty$, but for all $0 < \sigma < \infty$ the value of δ_p^2 from (H5′) is less than that provided by (H5). Thus (H5) overstates the variance of the sign statistic ($nk\delta_p^2/4$) in both examples.

Example 4 (The Mann-Whitney Test). Bohn and Wolfe [6] give expressions for the null hypothesis variance of the Mann-Whitney statistic under imperfect judgement ranking. Since their derivation depends crucially on (10), these results are not valid for general judgement ranking schemes. In fact, we will see that for both of our examples, as for the sign statistic, this assumption leads us to overstate the true asymptotic variance of the statistic.

Recall that for the Mann-Whitney statistic, $h(x, y) = I(x < y)$, so that (7) yields $h_{10}(x) = 1 - G(x)$ and $h_{01}(y) = F(y)$, where we assume F and G continuous. Thus, under ranked set sampling the asymptotic variance of U is given by Theorem 2, with $\gamma_{r\cdot} = 1 - E[G(X_{[r]})]$ and $\gamma_{\cdot s} = E[F(Y_{[s]})]$. When $F = G$, the usual null hypothesis, $\theta = 1/2$, and $\zeta_{10} = \zeta_{01} = 1/12$.

If (10) holds, then it is straightforward to show that as a consequence $\gamma_{r\cdot} = 1 - \sum_{i=1}^{k} p_{ir} E[G(X_{(i)})]$ and $\gamma_{\cdot s} = \sum_{i=1}^{l} q_{is} E[F(Y_{(i)})]$, where the q_{ij} are defined in the same way as the p_{ij} but for the second sample. Under the null hypothesis $F = G$, these simplify to $\gamma_{r\cdot} = 1 - (1/(k+1)) \sum_{i=1}^{k} i p_{ir}$ and $\gamma_{\cdot s} = (1/(l+1)) \sum_{i=1}^{l} i q_{is}$.

Suppose now that both populations are uniformly distributed, and that ranked set sampling is carried out with $k = l = 2$, and under the model for judgement ranking given in Example 1. Then $\gamma_{1\cdot} = \gamma_{\cdot 2} = 7/12$ and $\gamma_{2\cdot} = \gamma_{\cdot 1} = 5/12$, so that (9) yields $2\zeta_{10}^* = 2\zeta_{01}^* = 11/144$. However, in this case $p_{11} = q_{11} = 2/3$, so that naively assuming (10) to hold and applying the simple formula of the last paragraph yields $2\zeta_{10}^* = 2\zeta_{01}^* = 13/162$, which is about five percent larger than the correct value.

If both samples are from standard normal populations with $k = l = 2$ and with the judgement ranking model of Example 2 (with σ the same for both samples), we find that $2\zeta_{10}^* = 2\zeta_{01}^* = 1/12 - \left\{2E\left[\Phi(Z)\Phi\left(Z/\sqrt{1+2\sigma^2}\right)\right] - 1/2\right\}^2$, where Z is a standard normal random variable. However, in this case $p_{11} = q_{11} = 1 - \arctan(\sigma)/\pi$, so that incorrectly assuming (10) to hold and applying the simple formulas given above yields $2\zeta_{10}^* = 2\zeta_{01}^* = 1/12 - (1/9)\left[1/2 - \arctan(\sigma)/\pi\right]^2$. These expressions agree when $\sigma = 0$ and in the limit as $\sigma \to \infty$, but for any $0 < \sigma < \infty$ numerical evaluation reveals that the first yields a smaller value than the second. For example, if $\sigma = 0.5$, the incorrect formula exaggerates the asymptotic variance by thirteen percent.

4 When is Perfect Ranking Best?

Stokes and Sager [3] remark that "the two extremes in usefulness of the RSS occur when judgement is perfect and when it is no better than random." Though this statement is true for the distribution function estimator they considered, it is possible for imperfect judgement ranking to yield more efficient estimates than perfect ranking in ordinary ranked set sampling.

Example 5 (The Sample Variance). Stokes [2] examined estimation of the population variance, σ^2, under ranked set sampling using the usual sample variance, s^2. This is of course a U-statistic with symmetric kernel $h(x_1, x_2) = (1/2)(x_1 - x_2)^2$, for which $h_1(x) = (1/2)[(x-\mu)^2 + \sigma^2]$. Suppose that $k = 2$, and that X's follow a $U(0,1)$ distribution. Then, in the notation of Theorem 1, $\theta = 1/12$, $\zeta_1 = 1/720$, and for perfect ranking it is easily seen that $\gamma_1 = \gamma_2 = 1/12$. Thus, in this example, perfect ranking provides no improvement over an ordinary random sample of size $2m$, and the asymptotic variance of the sample variance is $(2m)^{-1}p^2k\zeta_1^* = (2m)^{-1}(1/180)$.

Now suppose that judgement ranking is such that, conditional on X_1, X_2,

$$J_1 = \begin{cases} 1, & \text{if } |X_1 - \frac{1}{2}| > |X_2 - \frac{1}{2}|, \\ 2, & \text{otherwise.} \end{cases}$$

so that the observation furthest from the mean in either direction is always judged "smallest." Then straightforward calculations show that $\gamma_1 = 5/48$ and $\gamma_2 = 1/16$, and the asymptotic variance of s^2 is $(2m)^{-1}p^2k\zeta_1^* = (2m)^{-1}(1/480)$. Thus in this example, imperfect ranking has asymptotic efficiency $8/3$ relative to perfect ranking.

Though the judgement ranking scheme of the last example is artificial, it demonstrates that no unqualified statement can be made concerning the efficacy of perfect versus imperfect judgement ranking. The next theorem gives a result in the positive direction for the class of U-statistic estimators considered in Section 2.2.

Theorem 3. *Under the conditions of Theorem 1, if $h_1(\cdot)$ is a monotone function, then* *the asymptotic variance of U_{RSS} is minimized (over all ranking schemes) by perfect ranking.*

Proof. Let $\gamma_{(r)} = E[h_1(X_{(r)})]$, and $\gamma_{[r]} = E[h_1(X_{[r]})]$, for $r = 1, \ldots, k$. Assume that h is nondecreasing. Then the ordered values of $h(X_{[1]}), \ldots, h(X_{[k]})$ are $h(X_{(1)}) \leq \ldots \leq h(X_{(k)})$, and by Proposition 12.D.1 of Marshall and Olkin [12], it follows that $(\gamma_{(1)}, \ldots, \gamma_{(k)})$ majorizes $(\gamma_{[1]}, \ldots, \gamma_{[k]})$. If h is nonincreasing exactly the same majorization result holds. By a classical theorem of Hardy, Littlewood, and Polya [12, Proposition 3.C.1], $\sum_{r=1}^{k} g(\gamma_{[r]}) \leq \sum_{r=1}^{k} g(\gamma_{(r)})$ for any convex function g. Thus $\sum_{r=1}^{k}(\gamma_{[r]} - \theta)^2 \leq \sum_{r=1}^{k}(\gamma_{(r)} - \theta)^2$, and the result follows from Theorem 1. □

Theorem 3 applies to the sample mean, the sign statistic, the empirical distribution function, and the signed rank statistic. As noted in Section 2.2, with the exception of the signed-rank statistic, the variance formula given by Theorem 1 is exact for these examples, so that the result of Theorem 3 actually holds for all sample sizes. Thus this also provides a corrected proof for Theorem 2 of Hettmansperger [4]. The previous counterexample exploits the fact that for the sample variance h_1 is not monotone.

Because both $\tilde{U}$, used in our proof of Theorem 1, and U^*, the generalized U-statistic of Taga et al. [9], have the same limiting distribution as U_{RSS}, Theorem 3 applies to both. The obvious generalization to the two-sample case, requiring h_{10} and h_{01} monotone, also holds and applies for example to the Mann-Whitney statistic.

Acknowledgements

The first author is grateful to Takashi Yanagawa for a helpful discussion and for providing a reprint of Taga et al. [9].

References

[1] T. R. Dell and J. L. Clutter, *Ranked set sampling theory with order statstics background*, Biometrics **28** (1972), 545–555.

[2] S. L. Stokes, *Estimation of variance using judgment ordered ranked set samples*, Biometrics **36** (1980), 35–42.

[3] S. L. Stokes and T. W. Sager, *Characterization of a ranked-set sample with application to estimating distribution functions*, J. Amer. Statist. Assoc. **83** (1988), 374–381.

[4] T. P. Hettmansperger, *The ranked-set sample sign test*, J. Nonparametr. Statist. **4** (1995), 263–270.

[5] L. L. Bohn and D. A. Wolfe, *Nonparametric two-sample procedures for ranked-set samples data*, J. Amer. Statist. Assoc. **87** (1992), 552–561.

[6] L. L. Bohn and D. A. Wolfe, *The effect of imperfect judgment rankings on properties of procedures based on the ranked-set samples analog of the Mann-Whitney-Wilcoxon statistic*, J. Amer. Statist. Assoc. **89** (1994), 168–176.

[7] L. L. Bohn, *A ranked-set sample signed-rank statistic*, J. Nonparametr. Statist. **9** (1998), 295–306.

[8] P. H. Kvam and F. J. Samaniego, *On the inadmissibility of empirical averages as estimators in ranked set sampling*, J. Statist. Plann. Inference **36** (1993), 39–55.

[9] Y. Taga, K. Wakimoto, and T. Yanagawa, *Generalized U-statistics for stratified random samples*, Skandinavisk Aktuarietidskrift **56** (1973), 30–51.

[10] T. Yanagawa, *Stratified random sampling; Gain in precision due to stratification in the case of proportional allocation*, Journal of the Institute of Mathematical Statistics **27** (1975), 33–44.

[11] R. H. Randles and D. A. Wolfe, *Introduction to the Theory of Nonparametric Statistics*, Krieger Publishing Company, Malabar, Florida, 1991.

[12] A. W. Marshall and I. Olkin, *Inequalities: Theory of Majorization and Its Applications*, Academic Press, San Diego, 1979.

Appendix: Proof of Theorem 1

To set notation, for $\boldsymbol{\alpha} \in \mathcal{A}$, write α_i uniquely as $\alpha_i = (c_i - 1)k + r_i$, with $1 \leq r_i \leq k$ and $1 \leq c_i \leq m$, and identify with $\boldsymbol{\alpha}$ the subset $X_{[r_1]c_1}, \ldots, X_{[r_p]c_p}$ of the ranked set sample. Let $\mathcal{A}^*$ consist of all $\boldsymbol{\alpha} \in \mathcal{A}$ for which $c_1 < \cdots < c_p$, i.e., those $\boldsymbol{\alpha}$ for which all of the observations selected are from different cycles. Then

$$
\begin{aligned}
U_{\mathrm{RSS}} &= U(X_{[1]1}, \ldots, X_{[k]1}, \ldots, X_{[1]m}, \ldots, X_{[k]m}) \\
&= \frac{1}{\binom{mk}{p}} \sum_{\boldsymbol{\alpha} \in \mathcal{A}} h(X_{[r_1]c_1}, \ldots, X_{[r_p]c_p}) \\
&= \frac{1}{\binom{mk}{p}} \sum_{\boldsymbol{\alpha} \in \mathcal{A}^*} h(X_{[r_1]c_1}, \ldots, X_{[r_p]c_p}) + \frac{1}{\binom{mk}{p}} \sum_{\boldsymbol{\alpha} \in \mathcal{A}/\mathcal{A}^*} h(X_{[r_1]c_1}, \ldots, X_{[r_p]c_p}) \\
&= \left[\binom{m}{p} k^p / \binom{mk}{p}\right] \tilde{U} + R\,,
\end{aligned}
$$

where

$$\tilde{U} = \frac{k^{-p}}{\binom{m}{p}} \sum_{\boldsymbol{\alpha} \in \mathcal{A}^*} h(X_{[r_1]c_1}, \ldots, X_{[r_p]c_p})$$

and

$$R = \frac{1}{\binom{mk}{p}} \sum_{\boldsymbol{\alpha} \in \mathcal{A}/\mathcal{A}^*} h(X_{[r_1]c_1}, \ldots, X_{[r_p]c_p})\,. \tag{13}$$

The proof proceeds by showing that $\sqrt{mk}(\tilde{U} - \theta)$ is asymptotically normal with mean 0 and variance $p^2 k \zeta_1^*$, while $\sqrt{mk}R \to 0$ in probability as $m \to \infty$. The result follows immediately from these facts, since $[\binom{m}{p} k^p / \binom{mk}{p}] = 1 + O(m^{-1})$.

We begin by showing that $\sqrt{mk}R \to 0$ in probability as $m \to \infty$. By Lemma 2,

$$E[|h(X_1, \ldots, X_p)|] = k^{-p} \sum_{r_1=1}^{k} \cdots \sum_{r_p=1}^{k} E[|h(X_{[r_1]1}, \ldots, X_{[r_p]p})|]\,. \tag{14}$$

But for any $\alpha \in \mathcal{A}$, $E[|h(X_{[r_1]c_1}, \ldots, X_{[r_p]c_p})|] = E[|h(X_{[r_1]1}, \ldots, X_{[r_p]p})|]$, and thus (14) implies that

$$E[|h(X_{[r_1]c_1}, \ldots, X_{[r_p]c_p})|] \leq k^p E[|h(X_1, \ldots, X_p)|] \quad \forall\, \boldsymbol{\alpha} \in \mathcal{A},$$

so that the expectations of the terms in (13) are uniformly bounded by a constant not depending on m. Since the number of elements in $\mathcal{A}/\mathcal{A}^*$ is $\binom{mk}{p} - \binom{m}{p}k^p$, and since $[\binom{mk}{p} - \binom{m}{p}k^p]/\binom{mk}{p} = O(m^{-1})$, we see that $E[\sqrt{m}R] \to 0$.

Similarly, Lemma 2 implies that

$$E[h^2(X_{[r_1]c_1}, \ldots, X_{[r_p]c_p})] \le k^p E[h^2(X_1, \ldots, X_p)] \quad \forall\, \boldsymbol{\alpha} \in \mathcal{A},$$

so that the variances and covariances of the terms in (13) are also uniformly bounded by a constant not depending on m. Since

$$\mathrm{Var}(R) = \frac{1}{\binom{mk}{p}^2} \sum_{\boldsymbol{\alpha} \in \mathcal{A}/\mathcal{A}^*} \sum_{\boldsymbol{\alpha}' \in \mathcal{A}/\mathcal{A}^*} \mathrm{Cov}[h(X_{[r_1]c_1}, \ldots, X_{[r_p]c_p}), h(X_{[r_1']c_1'}, \ldots, X_{[r_p']c_p'})]$$

has $[\binom{mk}{p} - \binom{m}{p}k^p]^2$ bounded terms, and since $[m/\binom{mk}{p}^2][\binom{mk}{p} - \binom{m}{p}k^p]^2 = O(m^{-1})$, we have $\mathrm{Var}[\sqrt{m}R] \to 0$. Thus $\sqrt{m}R \to 0$ in probability as $m \to \infty$, and since k is fixed, $\sqrt{mk}R \to 0$ in probability.

Turning to $\tilde{U}$, we note that

$$\tilde{U} = \tilde{U}(\mathbf{X}_1, \ldots, \mathbf{X}_m) = \frac{1}{\binom{m}{p}} \sum_{1 \le c_1 < \cdots < c_p \le m} \cdots \sum h^*(\mathbf{X}_{c_1}, \ldots, \mathbf{X}_{c_p})$$

is a U-statistic in the ranked sets $\mathbf{X}_1, \ldots, \mathbf{X}_m$, with symmetric kernel

$$h^*(\mathbf{X}_1, \ldots, \mathbf{X}_p) = k^{-p} \sum_{r_1=1}^{k} \cdots \sum_{r_p=1}^{k} h(X_{[r_1]1}, \ldots, X_{[r_p]p})\,.$$

Applying Lemma 2, we see that

$$E[h^*(\mathbf{X}_1, \ldots, \mathbf{X}_p)] = E[h(X_1, \ldots, X_p)] = \theta\,.$$

Now for $\mathbf{x} = (x_1, \ldots, x_k)$, let

$$h_1^*(\mathbf{x}) = E[h^*(\mathbf{x}, \mathbf{X}_2, \ldots, \mathbf{X}_p)] = k^{-1} \sum_{r=1}^{k} h_1(x_r)\,,$$

where the second equality follows from Lemma 2 and the definition of h_1 in (4). By the One-Sample U-Statistic Theorem, $\sqrt{m}(\tilde{U} - \theta)$ is asymptotically normal with mean 0 and variance $p^2 \zeta_1^* = p^2 \operatorname{Var}[h_1^*(\mathbf{X}_1)]$, as long as $\zeta_1^* > 0$. But since $X_{[1]1}, \ldots, X_{[k]1}$ are independent,

$$\zeta_1^* = k^{-2} \sum_{r=1}^{k} \operatorname{Var}[h_1(X_{[r]1})] = k^{-2} \sum_{r=1}^{k} \xi_r^2\,.$$

It is immediate that $\sqrt{mk}(\tilde{U} - \theta)$ is asymptotically normal with mean 0 and variance $p^2 k \zeta_1^*$, establishing the results of Theorem 1 through the first equality in (6).

To establish the second equality in (6), note that from Lemma 2 we have

$$\theta = E[h_1(X_1)] = \frac{1}{k} \sum_{r=1}^{k} E[h_1(X_{[r]1})] = \frac{1}{k} \sum_{r=1}^{k} \gamma_r\,,$$

and

$$E[h_1^2(X_1)] = \frac{1}{k} \sum_{r=1}^{k} E[h_1^2(X_{[r]1})] = \frac{1}{k} \sum_{r=1}^{k} \xi_r^2 + \frac{1}{k} \sum_{r=1}^{k} \gamma_r^2\,,$$

from which it follows easily that

$$\zeta_1 = \operatorname{Var}[h_1(X_1)] = \frac{1}{k} \sum_{r=1}^{k} \xi_r^2 + \frac{1}{k} \sum_{r=1}^{k} (\gamma_r - \theta)^2\,. \quad \square$$