Referee Report for AOS2501-026 "Berry-Esseen bounds for design-based causal inference with possibly diverging treatment levels and varying group sizes"

1 Summary and evaluation

This paper addresses the statistical inference of causal effects in the context of potentially diverging treatment levels and varying group sizes, within the randomization or design-based inference framework. The authors formulate the causal inference problem in terms of linear permutational statistics and derive several Berry-Esseen bounds for both the linear and quadratic functions of the point causal effect estimator. These bounds can handle general experimental designs with diverging treatment levels, varying sample sizes across treatment groups, and potentially differing numbers of causal effects of interest.

This paper makes a significant contribution to the literature on design-based causal inference in general settings, particularly those that allow treatment levels to diverge and group sizes to vary. The asymptotic normality of causal effect estimators has been well-studied in contexts where the number of treatment levels is fixed and sample sizes across treatment groups tend to infinity. However, the Berry-Esseen bounds and asymptotic normality in regimes where treatment levels diverge and group sizes vary remain underexplored in the existing literature. This paper addresses this important theoretical gap.

Moreover, the paper proposes conservative variance estimators for unreplicated and mixture designs, where sample sizes may be one in certain treatment levels. In such cases, deriving a variance estimator is challenging because the sample variance is not well-defined when the sample size is one. To address this, the paper proposes new variance estimators for design-based causal inference and establish the asymptotic properties of these estimators, representing another novel contribution.

Overall, the paper is well-written, the results are novel, and the proofs appear to be correct.

2 Comments and questions

My main comments are as follows:

- In Definition 3, the treatment arms are divided into $\mathcal{Q}_L$ and $\mathcal{Q}_S = \mathcal{Q}_U \cup \mathcal{Q}_R$. In practice, how can one distinguish between $\mathcal{Q}_L$ and $\mathcal{Q}_R$? That is, what constitutes a 'large' sample size? Moreover, for $q \in \mathcal{Q}_L$, the sample size N_q is of the same order as N_0. Is it possible for the sample sizes of some treatment arms to converge at a faster rate than others?
- It is fine to assume the bounded fourth-moment condition for the potential outcomes, as stated in Condition 3. However, I am curious whether it is possible to relax this condition to $\max_{q\in[Q]} N^{-1} \max_{i=1,\ldots,N} |Y_i(q) - \bar{Y}(q)| = o(1)$, as typically required by the finite-population central limit theorem?
- In Theorem 2, a constant of 19/4 appears in relation to H. Is this constant sharp? I would appreciate a discussion on this aspect.
- The proposed variance estimators for unreplicated and mixture designs share similarities with those used in finely stratified survey sampling and experiments. Could the authors provide a more detailed discussion of the similarities and differences between them after the formal introduction of the variance estimators?
- In the simulation, the total sample size is $N = 1780$. Do the proposed methods perform well with smaller sample sizes?

My minor comments are as follows:

- On page 4, the authors state "In the main theoretical results developed later, we will further assume that $N_0 \to \infty$ whereas $\bar{n} = O(1)$ asymptotically." Do all the theoretical results require this assumption? If not, it would be helpful to specify which theorems rely on this condition.
- The first line on page 6, $[Q]$, appears to be undefined; it seems that it should be $[Q-1]$.
- The first line above Example 6 on page 12, "We first apply Corollary 2 to the general factorial design".
- The first line under equation (17) on page 13, there are two extra spaces.
- In Theorem 4 on Page 16, "Validty" should be "Validity".

- The first line below equation (31) on page 21, “corresponds to” should be “correspond to”.
- In remark 2 on page 21, “based on the following ~~the~~ partition: ...”
- Page 24, “We use the sandwich variance estimators based on ~~based on~~ WLS with the target factors and their two-way interactions”.
- In the second paragraph in Section 4.3 on Page 24, “asymptotic Normality” should be “asymptotic normality”.