Report on AOS1910-054

This paper studies the phase transition phenomenon in large scale multiple testing and its implications to large scale genome-wide association studies (GWAS). As a phenomenon, phase transition in large scale multiple testing may no longer be surprising. Nevertheless, the paper provides detailed descriptions of the phenomenon in four different scenarios of multiple independent chi-squared tests and two scenarios of multiple independent two-sided normal test. These results significantly broaden the scope of the phenomenon. As far as I can see, the proofs of the results are all correct. Quite importantly, the paper makes a good connection between these results and large scale GWAS. Basically, the results can be read as saying that a large scale GWAS could be either a complete success or a complete failure, depending on the signal size per sample and the sample size, given the sparsity of significant SNPs. This makes a strong case for careful selection of the sample size, whose importance has been recognized in the field of GWAS for quite a while [2]. I therefore recommend acceptance for publication of the paper after some minor changes.

Comments on the Main Text

1. It would be helpful to emphasize somewhere that in the sub-threshold case, all thresholding procedures fail to get support recovery, whether they control the FWER, FDR, or some other error rate.
2. Section 4.1: It seems that in GWAS on some disorders, a difficulty is that the odds ratio of most disease-associated SNPs is close to one. Therefore, it would be useful to report the asymptotic of $w^2(R)$ as $R \to 1$.
3. Section 4.2: the relevance of this part needs to be made more clear. The optimal ϕ_1 is for one SNP. Since there can be a large number of disease-associated SNPs, and it may be unknown beforehand which are adversarial, which are protective, and how strong factors they are, how can an optimal ϕ_1 be found / estimated for the entire GWAS?
4. Section 4.2: perhaps it is useful to consider the following. With fixed ϕ_1, among all SNPs with the same R, which ones are the most significant? This should be easy to answer based on Proposition 4.1, as $w^2(R)$ is a function only in θ_1.
5. Example 4.1: in Fig. 3, the total number of cases and controls in each panel is half of n. The explanation of the halving is given only later, in Example 4.2. It should be presented in a paragraph on experiment setup before the examples.
6. I would prefer to move Section 6 to the supplement, while moving the current content in the supplement to Section 6.

Comments on the Supplement

1. Proof of Lemma A.1: the proof can and should be simplified by using L'Hôpital's rule,

$$\frac{\int_x^\infty t^{a-1}e^{-t/b}\,\mathrm{d}t}{bx^{a-1}e^{-x/b}} \to 1, \quad x \to \infty.$$

2. Eq. (A.4): $F^{\leftarrow}$ has not been defined.
3. Lemma A.2: this is a classical result; see for example, Example 5.3.4 in [3]. No proof is needed.
4. Lemma A.3: the proof can be simplified by using the aforementioned asymptotic based on L'p̂ital's rule.

5. Eq. (B.5): it would be helpful to mention below the display that t_p is the cut-off of an arbitrary thresholding procedure, not just Bonferroni.
6. P. 5: you have all the necessary ingredients, why not just use

$$\min_{i\in S}\frac{(Z_\nu(i)+\sqrt{\bar{\Delta}})^2}{u_p} = [1+o(1)]\min_{i\in S}\frac{(Z_\nu(i)+\sqrt{\bar{\Delta}})^2}{2\log p}$$
$$\leq [1+o(1)]\left(\min_{i\in S}\frac{Z_\nu(i)}{\sqrt{2\log p}}+\sqrt{\bar{r}}\right)^2 \xrightarrow{P} (-\sqrt{1-\beta}+\sqrt{\bar{r}})^2 < 1$$

   to directly go from (B.7) to the paragraph starting with "Finally, ..." on p.6. The last limit within the square perhaps should be stated as a lemma as it is also used in the proofs of Theorems 3.4 and 5.2.
7. Eq. (B.19): this is again by L'Hôpital's rule, not some original result.
8. Eq. (B.23): what does "$i \in [p]$" in the sum stand for?
9. Lemma B.2: this is Theorem 2 in [1] and should be cited as Jaeschke's theorem; also see [4], p. 600–601.
10. P. 13–14: the proof again can and should be streamlined. I don't think Lemma B.3 is needed. The lemma itself is a known result (cf. [4], p. 424) combined with the fact that the quantile of the normal distribution is slowly varying at $\pm\infty$.

References

[1] Eicker, F. (1979). The asymptotic distribution of the suprema of the standardized empirical processes. *Ann. Statist.* **7**, 116–138.

[2] Nishino, J., Ochi, H., Kochi, Y., Tsunoda, T., and Matsui, S. (2018). Sample size for successful genome-wide association study of major depressive disorder. *Front. Genet.* **9**:227. doi: 10.3389/fgene.2018.00227.

[3] Ravishanker, N. and Dipak, K. D. (2002). *A First Course in Linear Model Theory*. Texts in Statistical Science. Chapman & Hall/CRC.

[4] Shorack, G. R. and Wellner, J. A. (1986). *Empirical processes with applications to statistics*. John Wiley & Sons, Inc., New York.