

Editor's note: The uncorrupted statistical power

Jean Descôteaux

Université de Sherbrooke

In 1999, Wilkinson and the Task Force on Statistical Inference published a number of recommendations concerning testing – related issues including, most importantly, statistical power. These recommendations are discussed prior to the presentation of the structure and the various articles of this special issue on statistical power. The contents of these articles will most certainly prove quite useful to those wishing to follow the Task Force's recommendations.

In the late nineties, a select group of specialists (the so-called Task Force on Statistical Inference; TFSI) tackled the task of elucidating controversial issues related to statistical testing, its underlying models and recent methods of analysis. Their effort followed in the wake of Cohen's influential article "The earth is round ($p > .05$)", in which he describes a number of ways statistically significant findings are frequently being misinterpreted (Cohen, 1994). After a few years of work, the TFSI came up with a series of recommendations which were to be implemented in a revision of the *American Psychological Association Publication Manual* (the revised edition: APA, 2001). However, prior to that revision, the members of the Task Force felt they should publish an article in *American Psychologist* in order to initiate discussion about their recommendations.

This paper came out in 1999 (Wilkinson & the TFSI, 1999) and has had a large influence, if judged by the number of references made to it. Most importantly, one of the recommendations pertained directly to power and sample size. It read as follows:

"Provide information on sample size and the process that led to sample size decisions. Document the effect sizes, sampling and measurement assumptions, as well as analytic

procedures used in power calculations. Because power computations are most meaningful when done before data are collected and examined, it is important to show how effect-size estimates have been derived from previous research and theory in order to dispel suspicions that they might have been taken from data used in the study or, even worse, constructed to justify a particular sample size. Once the study is analyzed, confidence intervals replace calculated power in describing results" (p. 596).

Within this paragraph, Wilkinson and the TFSI describe the most common use of power, that is, to decide on a sample size suitable for the purpose of a study. This application is frequently called *a priori power analysis*, whereas the last sentence indirectly refers to what is sometimes called *a posteriori power analysis*.

Moreover, and most interestingly, the paper's section on power ends with the following little paragraph:

"Many of us encounter power issues when applying for grants. Even when not asking for money, think about power. Statistical power does not corrupt." (p. 597).

This formulation suggests that: 1) more and more funding agencies are making a priori power analysis a requirement, 2) it is not always easy to come up with satisfactory estimates of sample sizes based on past research results in order to compute sample size through power analysis, and 3) some people are or were under the impression that statistical power somehow undermines or invalidates the research process.

The present special issue on statistical power, for which I had the pleasure of acting as guest editor, offers extremely

Jean Descôteaux, Département de psychologie, Faculté des lettres et sciences humaines, Université de Sherbrooke, 2500, boul. de l'Université, Sherbrooke (Qc), J1K 2R1, Phone: (819)821-8000, ext. 63366, Email: jean.descoteaux@usherbrooke.ca

valuable information on the topics and issues brought up by the Task Force, as illustrated above. By providing detailed procedures, methods and tools, the articles of this special issue will most certainly help readers follow the recommendations issued by the Task Force (and adopted by more and more editorial boards). Moreover, by providing some historical background as well as provocative discussions about the limits of power-related computations, they also present readers with the necessary background to understand why the Task Force chose to end its recommendation by writing “Statistical power does not corrupt”.

By content, the articles of this issue split into two broad categories: conceptual and methodological. The more conceptual articles are presented first and are followed by the papers of the methodological category. Descôteaux’s (2007) article opens the issue with an historical presentation of the statistical power concept, which partly explains the “corruption” described by the Task Force, as well as a short introduction of what is described as *a priori* power analysis. The article by Laurencelle (2007) then proceeds to describe some problems and limitations linked with the use of statistical power in addition to providing thought-provoking propositions. This article directly deals with point 2) about the difficulties related to obtaining satisfactory estimates of sample sizes.

The work of Wilson VanVoorhis and Morgan (2007) begins the methodological section by offering useful rules-of-thumb guiding the selection of sample sizes (*a priori* power analysis). It is followed by the article of Mayr, Erdfelder, Buchner, and Faul (2007), which presents some guidelines for the use of a very handy program called “GPower” that can be downloaded for free from their website. The program, among other applications, can be used both for *a priori* and *a posteriori* power computations. The remaining three articles are more mathematical in nature and deal with the “mechanics” of the power computations (either *a priori* or *a posteriori*) associated with various statistical tests. Cousineau’s (2007) work deals with the *t*-test, while Hélie’s (2007) presents an interesting method based on the use of noncentral distributions which can be used to understand and compute power when more than two groups are involved in the analysis. The final article by Chartier and Allaire (2007) does a great job reducing the complexities associated with power estimation for Multivariate Analysis of Variance. In homage to Jacob Cohen, who pioneered the field of statistical power and deceased not so long ago, we chose to end the issue on the summary of his landmark article “A power primer” (Cohen, 1992).

As can be gathered from this brief presentation, the

special issue covers many aspects associated with statistical power so that it represents, at least in my opinion, a very good introduction to the subject. Although wide in scope, it is also very precise in its presentation, providing for instance many numerical examples which can be transferred and adapted to many research contexts. I feel assured that readers, both novices and more experienced, will find something valuable to them in the following pages. It has been a pleasure for me to act as guest editor for this special issue and I wish to express my gratitude to all the authors who have made its publication possible by contributing a piece on statistical power.

References

- American Psychological Association (2001). *Publication Manual of the American Psychological Association* (5th ed.). Washington, DC.
- Chartier, S., & Allaire, J.-F. (2007). Power estimation in Multivariate analysis of variance. *Tutorials in Quantitative Methods for Psychology*, 3, 70-78.
- Cohen, J. (1994). The earth is round ($p < .05$). *American Psychologist*, 49, 997-1003.
- Cohen, J. (1992). A power primer. *Psychological Bulletin*, 112, 155-159.
- Cousineau, D. (2007). Computing the power of a *t* test. *Tutorials in Quantitative Methods for Psychology*, 3, 60-62.
- Descôteaux, J. (2007). Statistical power: An historical introduction. *Tutorials in Quantitative Methods for Psychology*, 3, 28-34.
- Hélie, S. (2007). Understanding statistical power using noncentral probability distributions: Chi-squared, G-Squared, and ANOVA. *Tutorials in Quantitative Methods for Psychology*, 3, 63-69.
- Laurencelle, L. (2007). Inventer ou estimer la puissance statistique? Quelques considérations utiles pour le chercheur. *Tutorials in Quantitative Methods for Psychology*, 3, 35-42.
- Mayr, S., Erdfelder, E., Buchner, A., & Faul, F. (2007). A short tutorial of GPower. *Tutorials in Quantitative Methods for Psychology*, 3, 51-59.
- Wilkinson, L., & the Task Force on Statistical Inference (1999). Statistical methods in psychology journals. *American Psychologist*, 54, 594-604.
- Wilson Van Voorhis, C.R., & Morgan, B.L. (2007). Understanding power and rules of thumb for determining sample sizes. *Tutorials in Quantitative Methods for Psychology*, 3, 43-50.