

## How we Deal with Small Data?

09 Novembre 2019

Manuela Ghica, Irina B?ncescu, Denisa Ioana Udeanu

### Abstract

Most commonly used statistical methods in analyzing and interpreting clinical trial data are based on a specific distribution of the data, such as normal, binomial and exponential models. The alternative is to choose nonparametric methods based on rank, but with the disadvantage that for small size samples, they may give erroneous results. In this article, we refer an accurate statistical method based on computational simulations in order to characterize and compare as efficiently as possible clinical trial studies having few values. The bootstrap technique is a statistical method that does not depend on any data distribution hypothesis and can be considered adequate to estimate some essential parameters in data analysis by different types of bootstrap confidence intervals. The effect-size statistics and its confidence intervals are also presented, being useful when we compare group means.

### Table of contents

1. Introduction
2. Confidence intervals
3. Effect size
4. Cohen Effect size for mean difference
5. Effect size for non normal - data
6. Conclusions

### 1. Introduction

Computational statistics along with statistical software packages can describe and shape any biomedical phenomena using bootstrap replication methods [1]. With the help of these bootstrap techniques, it is possible to estimate using confidence intervals any statistical parameter belonging to a statistical population which does not respect the hypotheses of classical inferential statistics.

Challenges occur when data distribution is uncertain and fails to be approximated with a normal one and, last but not least, when we deal with insufficient number of data, thus not being able to obtain a known statistical distribution of the data at a reasonable risk. The hypotheses of inferential statistics require a selection volume of at least 30 values, which disagrees with the ethical code and the economic conditions of animal experiment studies [2]. Thus, the research on the number of animals necessary to obtain credible results is pervasive [3] but, also, the statistical approach must be modern, based on computational simulations [4].

The bootstrapping technique, also called resampling, is based on random replications having sample extracts returns (that is, the same individual may be encountered several times in the bootstrap set and also in any other bootstrap sample), from the dataset under analysis, replications having the same size as the original dataset. This method assumes that the original dataset is randomly extracted from an infinite population and is representative of it.

This bootstrap resampling technique is also used in bioequivalence studies of formulations to obtain a confidence interval for the difference of two treatments, one tested and one reference [5].

This article is divided into four sections. In the first section, we analyze the need to use computational techniques when the hypotheses of inferential parametric statistics are not accomplished. The second section describes several models of bootstrap confidence intervals and the assumptions they imply. In the next section, we amplify the new statistics approach by exposing the effect size measures both under normality conditions and when this assumption is not met. The last section is reserved for conclusions.

Click [here](#) see the full version.

Contributo selezionato da Filodiritto tra quelli pubblicati nei Proceedings “17th Romanian National Congress of Pharmacy - 2018”

Per acquistare i Proceedings clicca qui:

<https://www.filodiritto.com/proceedings>

Contribution selected by Filodiritto among those published in the Proceedings “17th Romanian National Congress of Pharmacy - 2018”

To buy the Proceedings click here:

<https://www.filodiritto.com/proceedings>

## References

1. Efron, B., Tibshirani, R.J. (1993). An Introduction to the Bootstrap. Chapman and Hall: London.
2. Fitts, D. A. (2011). Ethics and Animal Numbers: Informal Analyses, Uncertain Sample Sizes, Inefficient Replications, and Type I Errors. *Journal of the American Association for Laboratory Animal Science: JAA LAS*, 50(4), pp. 445-453.
3. Charan, J., & Kantharia, N. D. (2013). How to calculate sample size in animal studies? *Journal of Pharmacology & Pharmacotherapeutics*, 4(4), pp. 303-306.
4. Bolton, S., Bon, C. (2004). *Pharmaceutical Statistic. Practical and Clinical Applications*, Marcel Dekker, Inc., New York Basel.
5. Patterson, S., Jones B. (2006). *Bioequivalence and Statistics in Clinical Pharmacology*, Published in Chapman & Hall/CRC Taylor & Francis Group.
6. Li, J.C.H. (2016). Effect size measures in a two-independent-samples case with nonnormal and nonhomogeneous data, *Behavior Research Methods* 48(4), pp. 1560-1574.
7. Cohen, J. (1988). *Statistical power analysis for the behavioral sciences* (2nd ed.). Hillsdale, NJ: Erlbaum.
8. Lee, D. K. (2016). Alternatives to P-value: confidence interval and effect size. *Korean Journal of Anesthesiology*, 69(6), pp. 555-562.
9. Hedge LV, Olkin I. (2014). *Statistical methods for meta-analysis*. Orlando: Academic Press Inc.

10. Hogarty, K. Y., & Kromrey, J. D. (2001). We've been reporting some effect sizes: Can you guess what they mean? Paper presented at the annual meeting of the American Educational Research Association, Seattle, WA.
11. Algina, J., Keselman, H. J., & Penfield, R. D. P. (2005). An alternative to Cohen's standardized mean difference effect size: A robust parameter and confidence interval in the two independent groups case. *Psychological Methods*, 10, pp. 317-328.

**TAG:** *simulation, small data, effect size, proceedings*

---

### **Avvertenza**

*La pubblicazione di contributi, approfondimenti, articoli e in genere di tutte le opere dottrinarie e di commento (ivi comprese le news) presenti su Filodiritto è stata concessa (e richiesta) dai rispettivi autori, titolari di tutti i diritti morali e patrimoniali ai sensi della legge sul diritto d'autore e sui diritti connessi (Legge 633/1941). La riproduzione ed ogni altra forma di diffusione al pubblico delle predette opere (anche in parte), in difetto di autorizzazione dell'autore, è punita a norma degli articoli 171, 171-bis, 171-ter, 174-bis e 174-ter della menzionata Legge 633/1941. È consentito scaricare, prendere visione, estrarre copia o stampare i documenti pubblicati su Filodiritto nella sezione Dottrina per ragioni esclusivamente personali, a scopo informativo-culturale e non commerciale, esclusa ogni modifica o alterazione. Sono parimenti consentite le citazioni a titolo di cronaca, studio, critica o recensione, purché accompagnate dal nome dell'autore dell'articolo e dall'indicazione della fonte, ad esempio: Luca Martini, La discrezionalità del sanitario nella qualificazione di reato perseguibile d'ufficio ai fini dell'obbligo di referto ex. art 365 cod. pen., in "Filodiritto" (<https://www.filodiritto.com>), con relativo collegamento ipertestuale. Se l'autore non è altrimenti indicato i diritti sono di Inforomatica S.r.l. e la riproduzione è vietata senza il consenso esplicito della stessa. È sempre gradita la comunicazione del testo, telematico o cartaceo, ove è avvenuta la citazione.*

---