Working on Medicinal Frontiers: Possibilities of Statistical Support?

Kurt Neumann*

Department of Medicine, Hungary

*****Corresponding author:** Kurt Neumann, Department of Medicine, Hungary

ARTICLE INFO

ABSTRACT

Citation: Kurt Neumann. Working on Medicinal Frontiers: Possibilities of Statistical Support?. Biomed J Sci & Tech Res 22(1)-2019. BJSTR. MS.ID.003689.

Introduction

My professional experience with my clients indicates that all types of physicians generally think there must be some general scientific information available prior seeking detailed statistical advice. I wrote this article to show the reader that this is a fundamental scientific error from my view, under the assumption that your problem originates based on human patients living on our planet earth, and your key data are at least approximately continuous and quantitative. This information can be transformed as evidence that all human patients represent a big – but finite – sampling population. A report of United Nations displays for 2019 an estimated world population of 7713 million of people. Other statistical sources report an average population increase of 27 humans worldwide as difference of total number of births minus total number of deaths in every ten seconds. There are substantial dynamics in the world population figures contained! The number of statistical methods has increased from about some twenty thousand in the years around 2000 by about a factor of two or even more until today based on an internal assessment due to all the many big data and other developments. The concept of parameter free tolerance limits based on a theorem from Wilks [1] and its application in the frontiers of science will be discussed herein.

Methods

The theorem of Wilks describes the functional connection between the percentage share of the true population data in a random sample of size n with desired confidence levels in the interval between the minimum and maximum of any continuous data distribution sampled from an infinite population. In my experience over some five decades of professional statistical work, I would judge the finite world population is just causing negligible error in the requirement of an infinite sampling population.

Some Numerical Examples

1) In case you plan a pilot study with six patients. You want to know at the routine 95% confidence level in statistics which percentage of the data of the unknown and otherwise unspecified distribution are contained between the smallest and largest data point. The answer is 42%.

2) In case you reduce the level to 90% then the answer will be 49% or about half of the true distribution.

3) In case you liked to increase your confidence level to 99.9% then your sample interval between the extreme values will cover only about 18% of the true distribution.

4) In case you plan a study with given confidence levels and given percentage share of the unknown true distribution then you can solve the Wilks equation to find the necessary sample size n.

5) In case you were satisfied with a 90% confidence level for a pilot study and you would like a safety level of coverage of 90% of the unknown distribution then your sample size must be 38.

6) In case you want to change the 90% to 95% as statistical safety level in the above example then your sample size should be 100. An increase to 99% in the previous example would require a sample of 600.

7) I think in the early stages of medical research these considerations have tremendous implications: In case your priority is safety then you want to expose only six patients to a new treatment scheme, but the price in lack of statistical power for your primary goals seems to me too high.

8)   In case you needed a sponsor's support for your research you can benefit from scratch with the ethical review board for your project: You could sketch a plan that after a very safety oriented first phase with a sample size in the range from 6 to 38, the second step should be planned with sample sizes about 100 to 600. In this second development time period you have then at least a solid sample size basis for the solid application of sample size calculations for subsequent marketing or other required authorizations for your research in case of success in every project step.

9)   In my professional activities I could very frequently observe that projected phase III studies failed due to an insufficient quality of the estimates of the standard deviations. I think this proposed approach of tolerance intervals can assist to prevent those expensive experiences with a high level of safety from a statistical perspective.

Conclusion/Discussion

My experience indicates that even at university level educated professional statisticians have seldom the concept of tolerance limits in their minds when they are consulted in the study planning phase. A clear limitation of this article is the omission of implications of multiple testing or calculations of tolerance limits for the actually planned treatment group(s) in a future study. In my view the gain in safety of future decisions is worth the relatively benign, consequential sample size increases. Another limitation is that the formulae used here do not apply for categorical yes/no variables but are restricted to continuous data. It should be noted that additional techniques based on the same mathematical principles are available in the statistical literature. Ordinal data with a small number of gradings might be used as very crude approximations only. I think that the most important impact on the usage of tolerance intervals is the availability of reliable estimates for the treatment and control group's sample size and the subsequent estimation of standard deviations and other distributional parameters prior big decisive studies are envisaged minimizing financial risks for any type of sponsor or researcher budgets. The costs of medical interventions expressed in currency units, despite being exactly a discrete variable, for health technology applications in the context of economic evaluations induce in the medical context only negligible errors and can be safely categorized as continuous. In my view these techniques of tolerance limits help to improve the quality of research projects. Modern information technology infrastructures offer today very economical tools for complex calculations with unprecedented end user comfort. Overall, the question in the title could get in my view a clear YES, if you knew this theorem of Wilks and used it in your scientific work.

References

1. Wilks SS (1941) Determinations of sample sizes for setting tolerance limits; Ann Math Statis 12(1): 91-96.

Submission Link: https://biomedres.us/submit-manuscript.php

Assets of Publishing with us

- Global archiving of articles
- Immediate, unrestricted online access
- Rigorous Peer Review Process
- Authors Retain Copyrights
- Unique DOI for all articles

https://biomedres.us/