Thresholds In Multiple Testing - False Positive Discovery And Non Discovery Rates

Dr. Sampoornam. W

The Tamilnadu Dr.M.G.R. Medical University, Associate Professor, Mental health nursing department, Dhanvantri College of Nursing Pallakkapalayam, Namakkal. Tamilnadu, India, PN: 9865677976
sampoornamwebster@yahoo.in

Abstract: The false discovery rate (FDR) is a method of conceptualizing the rate of type I errors in null hypothesis testing when conducting multiple comparisons. FDR controlling procedures are designed to control the expected proportion of "discoveries" (rejected null hypotheses) that are false (incorrect rejections). In other words FDR is designed to control the proportion of false positives among the set of rejected hypotheses. This is more sensitive than traditional methods simply because of using a more lenient metric for false positives. False discoveries infiltrate the science world. The probability of a false positive finding increases with the numbers of statistical analytical tests. With recently expanding possibilities for piling data, concerns about the effects of multiplicities on false positive discoveries in the scientific endeavor have increased. However, awareness has not steep evenly throughout all branches of science. The role of new statistical approaches such as the false discovery rate for controlling false positive findings as well as the impact of false positive findings on science shall be highlighted in this critical review.

Key Words: False discovery rate, False non discovery rate, Multiple testing, Nurse scientist

Introduction
It is a fact that 80-95% of cases made based on observational data which fails to replicate when rigorously re tested [7]. Although there are a number of explanations for the failure of apparent scientific findings to replicate one of the major ones are undoubtedly failure to allow the effects of multiplicity on the probability of false results. Although statistical methods for controlling error rates in multiple testing have been available for some time, they have been ignored in many scientific areas. When the number of hypothesis is relatively limited or the tests relate to different aspects of the study, it is easy although unwise to ignore the problem. With the increase in the simultaneous testing of substantial numbers of hypotheses in recent years, the problems have become impossible to ignore [5].

1.1 False Positive Discovery Rate
Benjamini and Hochberg [1] introduced the concept of false discovery rate (FDR) as a way to allow inference when many tests are being conducted. The False Discovery Rate approach is a more recent development. This approach also determines adjusted p-values for each test. However, it controls the number of false discoveries in those tests that result in a discovery (significant result). FDR allows the researcher to tolerate a certain number of tests to be incorrectly discovered. The word rate in the FDR is the proportion of discoveries that are false among all discoveries i.e. proportion of incorrect rejections among all rejections of the null hypothesis. The False Discovery Proportion (FDP) to be the (unobserved) proportion of false discoveries among total rejections [2]. A compound decision theoretic framework for spatial multiple testing and propose a class of asymptotically optimal data driven procedures that control the FDR, false discovery exceedance (FDX) and false cluster rate (FCR) respectively. The control of the FDX and FCR is quite challenging from the classical perspective [8]. As in standard multiple testing problems, controlling the fraction of errors is an alternative to the traditional approach of controlling the overall probability of type I errors, also denoted as the family wise error etc. False discovery control (FDC) under arbitrary covariance dependence is a very challenging and important open problem in the modern research [4].

1.2 False Non Discovery Rate
The FNR, called false non-discovery rate was developed by Genovese and Wasserman [3]. False Non discovery Proportion (FNP) is the proportion of missed discoveries among those tests for which the null is retained. False non-discovery rate (FNR) are developed for dependent test statistics under a model where the number of true null hypotheses is assumed fixed and a mixture model where different configurations of true and false null hypotheses are assumed to have certain probabilities. This result verify some desirable properties of FNR as measures of error rates and extend some previously known results, providing further insights into the notions of FNR and related measures under dependence. A simulation study is also conducted investigating how the modified Bonferroni or Sidak procedure performs compared to its unmodified version in terms of a measure of power involving both FDR and FNR.

2. Nursing Literatures and Existing Gap
Nurse scientist do not test false discoveries in research work due paucity of statistical knowledge. Only few critical reviews denote multiple comparison analytical testing. There are no nursing research studies to report novelty pathways in the false discovery rate (FDR). Few nursing literatures on genomics and nuclear magnetic resonance metabolomics have tested the false discoveries [6].

References
[1] Y. Benjamini, Y. Hochberg, "Controlling the False Discovery Rate: A Practical and Powerful Approach to Multiple Testing", J R Statist Soc B, 57(1), pp.289-300, 1995.

[2] R. Christopher, Genovese, "A Tutorial on False Discovery Control - CMU Statistics", Carnegie, NIH Grant, 2003.

[3] C. Genovese, L. Wasserman, "Operating characteristics and extensions of the false dscovery rate procedure", Journal of Royal Statistical Society, Ser B 64, pp.499–517, 2002.

[4] Jianqing Fan, Xu Han & Weijie Gu, "Estimating False Discovery Proportion Under Arbitrary Covariance Dependence", Journal of the American Statistical Association, 107:499, pp.1019-1035, 2012.

[5] P. Juliet Shaffer, "Multiplicities and False-Positive Rates in Science: Overview", Annual Meeting, Advancing Science Serving Society, 2010.

[6] S. Marcia , Monteiro, " Nuclear Magnetic Resonance metabolomics reveals an excretory metabolic signature of renal cell carcinoma", Scientific Reports, 2016.

[7] Stanley Young, "National Institute of Statistical Sciences. Health Findings and False Discoveries", Annual Meeting, Advancing Science Serving Society, 2010.

[8] Wenguang Sun etal, " False Discovery Control in Large-Scale Spatial Multiple Testing", pp. 1-26