Preface

Mean and variance are the pillars of classic statistics and statistical inference. These metrics are appropriate for symmetric distributions but are routinely used when distributions are not symmetric. Moreover, the mean is not coherent with human perception of centrality – we never sum but instead look for a typical value, the mode, in statistics language. Simply put, the mean is a suitable metric for computers, but the mode is for people. It is difficult to explain why the mean and median are parts of any statistical package but not the mode. Mean and variance penetrate the theory of statistical inference: the minimum variance unbiased estimator is the landmark of classic statistics. Unbiased estimators rarely exist outside of linear statistical models, but even when they do, they may produce incomprehensible values of the squared correlation coefficient or variance component. This work offers a new statistical theory for small samples with an emphasis on the exact optimal tests and unequal-tail confidence intervals (CI) using the cumulative distribution function (cdf) of a statistic as a pivotal quantity.

The organization of the book is as follows. The first chapter outlines the limitations of classic statistics and uses the normal variance statistical inference for a brief introduction to our theory. It contains a section on some extensions of the Neyman-Pearson lemma to be used in subsequent chapters. The second and third chapters introduce two competing approaches: maximum concentration and mode statistics with optimal confidence intervals and tests. An optimal confidence interval (CI) with coverage probability going to zero gives birth to a maximum concentration or mode estimator depending on the choice of the CI. Two tracks for statistical inference are offered combined under one umbrella of M-statistics: maximum concentration (MC) theory stems from the density level test, short-length CI and implied MC parameter estimator. Mode (MO) theory includes the test, unbiased CI, and implied parameter MO estimator. Unlike the classic approach, we suggest different CI optimality criteria and respective estimators depending on the parameter domain. For example, the CI with the minimum length on the log scale yields a new estimator of the binomial probability and Poisson rate, both positive, even when the number of successes is zero. New criteria for efficient estimation are developed as substitutes for the variance and the mean square error. Chapter 4 discusses definitions of the p-value for asymmetric distributions – a generally overlooked but an important statistical problem. M-statistics is demonstrated in action in Chapter 5. Novel exact optimal CIs and statistical tests are developed for major statistical parameters: effect size, binomial probability, Poisson rate, variance component in the meta-analysis model, correlation coefficient, squared correlation coefficient, and coefficient of determination in linear model. M-statistics is extended to the multidimensional parameter in Chapter 6. The exact confidence regions and unbiased tests for normal mean and variance, the shape parameters of the beta distribution, and nonlinear regression illustrate the theory.

The `R` codes can be freely downloaded from my website

`www.eugened.org`

stored at `GitHub`. By default, each code is saved in the folder `C:\Projects\Mode\` via `dump` command every time it is called. If you are using a Mac or do not want to save in this directory, remove or comment (`#`) this line before running the code (sometimes external `R` packages must be downloaded and installed beforehand).

I would like to hear comments, suggestions, and opinions from readers. Please e-mail me at `eugened@dartmouth.edu`.

Eugene Demidenko

Dartmouth College
Hanover, New Hampshire
April 2023