Ethical Guidelines for Statistical Practice

Prepared by the Committee
on Professional Ethics of the
American Statistical Association

Approved by the ASA Board in April 2016

ASA AMERICAN STATISTICAL ASSOCIATION
Promoting the Practice and Profession of Statistics®

Contents

Purpose of the Guidelines

The American Statistical Association's Ethical Guidelines for Statistical Practice are intended to help statistics practitioners make decisions ethically. Additionally, the Ethical Guidelines aim to promote accountability by informing those who rely on statistical analysis of the standards that they should expect. The discipline of statistics links the capacity to observe with the ability to gather evidence and make decisions, providing a foundation for building a more informed society. Because society depends on informed judgments supported by statistical methods, all practitioners of statistics, regardless of training and occupation or job title, have an obligation to work in a professional, competent, and ethical manner and to discourage any type of professional and scientific misconduct.[1]

Good statistical practice is fundamentally based on transparent assumptions, reproducible results, and valid interpretations. In some situations, Guideline principles may conflict, requiring individuals to prioritize principles according to context. However, in all cases, stakeholders have an obligation to act in good faith, to act in a manner that is consistent with these Guidelines, and to encourage others to do the same. Above all, professionalism in statistical practice presumes the goal of advancing knowledge while avoiding harm; using statistics in pursuit of unethical ends is inherently unethical.[2]

Good statistical practice is fundamentally based on transparent assumptions, reproducible results, and valid interpretations.

The principles expressed here should guide both those whose primary occupation is statistics and those in all other disciplines who use statistical methods in their professional work. Therefore, throughout these Guidelines, the term "statistician" includes all practitioners of statistics and quantitative sciences, regardless of job title or field of degree, comprising statisticians at all levels of the profession and members of other professions who utilize and report statistical analyses and their implications.

A. Professional Integrity and Accountability

The ethical statistician uses methodology and data that are relevant and appropriate, without favoritism or prejudice, and in a manner intended to produce valid, interpretable, and reproducible results.[3] The ethical statistician does not knowingly accept work for which he/she is not sufficiently qualified, is honest with the client about any limitation of expertise, and consults other statisticians when necessary or in doubt.

The ethical statistician:

1. Identifies and mitigates any preferences on the part of the investigators or data providers that might predetermine or influence the analyses/results.[4]
2. Employs selection or sampling methods and analytic approaches appropriate and valid for the specific question to be addressed, so that results extend beyond the sample to a population relevant to the objectives with minimal error under reasonable assumptions.[5]
3. Respects and acknowledges the contributions and intellectual property of others.
4. When establishing authorship order for posters, papers, and other scholarship, strives to make clear the basis for this order, if determined on grounds other than intellectual contribution.[6]
5. Discloses conflicts of interest, financial and otherwise, and manages or resolves them according to established (institutional/regional/local) rules and laws.[7]
6. Accepts full responsibility for his/her professional performance. Provides only expert testimony, written work, and oral presentations that he/she would be willing to have peer reviewed.

B. Integrity of Data and Methods

The ethical statistician is candid about any known or suspected limitations, defects, or biases in the data that may impact the integrity or reliability of the statistical analysis. Objective and valid interpretation of the results requires that the underlying analysis recognizes and acknowledges the degree of reliability and integrity of the data.[8]

The ethical statistician:

1. Acknowledges statistical and substantive assumptions made in the execution and interpretation of any analysis. When reporting on the validity of data used, acknowledges data editing procedures, including any imputation and missing data mechanisms.
2. Reports the limitations of statistical inference and possible sources of error.[9]
3. In publications, reports, or testimony, identifies who is responsible for the statistical work if it would not otherwise be apparent.
4. Reports the sources and assessed adequacy of the data; accounts for all data considered in a study and explains the sample(s) actually used.
5. Clearly and fully reports the steps taken to preserve data integrity and valid results.
6. Where appropriate, addresses potential confounding variables not included in the study.
7. In publications and reports, conveys the findings in ways that are both honest and meaningful to the user/reader. This includes tables, models, and graphics.
8. In publications or testimony, identifies the ultimate financial sponsor of the study, the stated purpose, and the intended use of the study results.
9. When reporting analyses of volunteer data or other data that may not be representative of a defined population, includes appropriate disclaimers and, if used, appropriate weighting.
10. To aid peer review and replication, shares the data used in the analyses whenever possible/allowable, and exercises due caution to protect proprietary and confidential data, including all data that might inappropriately reveal respondent identities.
11. Strives to promptly correct any errors discovered while producing the final report or after publication. As appropriate, disseminates the correction publicly or to others relying on the results.

C. Responsibilities to Science/Public/Funder/Client

The ethical statistician supports valid inferences, transparency, and good science in general, keeping the interests of the public, funder, client, or customer in mind (as well as professional colleagues, patients, the public, and the scientific community).[10]

The ethical statistician:

1. To the extent possible, presents a client or employer with choices among valid alternative statistical approaches that may vary in scope, cost, or precision.
2. Strives to explain any expected adverse consequences of failure to follow through on an agreed-upon sampling or analytic plan.
3. Applies statistical sampling and analysis procedures scientifically, without predetermining the outcome.
4. Strives to make new statistical knowledge widely available to provide benefits to society at large and beyond his/her own scope of applications.[11]
5. Understands and conforms to confidentiality requirements of data collection, release, and dissemination and any restrictions on its use established by the data provider (to the extent legally required), and protects use and disclosure of data accordingly. Guards privileged information of the employer, client, or funder.

D. Responsibilities to Research Subjects

The ethical statistician protects and respects the rights and interests of human and animal subjects at all stages of their involvement in a project. This includes respondents to the census or to surveys, those whose data are contained in administrative records, and subjects of physically or psychologically invasive research.

The ethical statistician:

1. Keeps informed about and adheres to applicable rules, approvals, and guidelines for the protection and welfare of human and animal subjects.[12]
2. Strives to avoid the use of excessive or inadequate numbers of research subjects, and excessive risk to research subjects (in terms of health, welfare, privacy, and ownership of their own data), by making informed recommendations for study size.[13]
3. Protects the privacy and confidentiality of research subjects and data concerning them, whether obtained from the subjects directly, other persons, or existing records. Anticipates and solicits approval for secondary and indirect uses of the data, including linkage to other data sets, when obtaining approvals from research subjects, and obtains approvals appropriate to allow for peer review and independent replication of analyses.
4. Knows the legal limitations on privacy and confidentiality assurances and does not over-promise or assume legal privacy and confidentiality protections where they may not apply.
5. Considers whether appropriate research-subject approvals were obtained before participating in a study involving human beings or organizations, before analyzing data from such a study, and while reviewing manuscripts for publication or internal use. The statistician considers the treatment of research subjects (e.g., confidentiality agreements, expectations of privacy, notification, consent, etc.) when evaluating the appropriateness of the data source(s).[14]
6. In contemplating whether to participate in an analysis of data from a particular source, refuses to do so if participating in the analysis could reasonably be interpreted by individuals who provided information as sanctioning a violation of their rights.
7. Recognizes that any statistical descriptions of groups may carry risks of stereotypes and stigmatization. Statisticians should contemplate, and be sensitive to, the manner in which information is framed so as to avoid disproportionate harms to vulnerable groups.

E. Responsibilities to Research Team Colleagues

Science and statistical practice are often conducted in teams made up of professionals with different professional standards. The statistician must know how to work ethically in this environment.

The ethical statistician:

1. Recognizes that other professions have standards and obligations, that research practices and standards can differ across disciplines, and that statisticians do not have obligations to standards of other professions that conflict with these Guidelines.
2. Ensures that all discussion and reporting of statistical design and analysis is consistent with these Guidelines.
3. Avoids compromising scientific validity for expediency.[15]
4. Strives to promote transparency in design, execution, and reporting or presenting of all analyses.

F. Responsibilities to Other Statisticians or Statistics Practitioners

The practice of statistics requires consideration of the entire range of possible explanations for observed phenomena, and distinct observers drawing on their own unique sets of experiences can arrive at different and potentially diverging judgments about the plausibility of different explanations. Even in adversarial settings, discourse tends to be most successful when statisticians treat one another with mutual respect and focus on scientific principles, methodology and the substance of data interpretations.

Out of respect for fellow statistical practitioners, the ethical statistician:

1. Promotes sharing of data and methods as much as possible and as appropriate without compromising propriety. Makes documentation suitable for replicate analyses, metadata studies, and other research by qualified investigators.
2. Helps strengthen the work of others through appropriate peer review; in peer review, respects differences of opinion and assesses methods, not individuals. Strives to complete review assignments thoroughly, thoughtfully, and promptly.
3. Instills in students and non-statisticians an appreciation for the practical value of the concepts and methods they are learning or using.
4. Uses professional qualifications and contributions as the basis for decisions regarding statistical practitioners' hiring, firing, promotion, work assignments, publications and presentations, candidacy for offices and awards, funding or approval of research, and other professional matters.
5. Does not harass or discriminate.

G. Responsibilities Regarding Allegations of Misconduct

The ethical statistician understands the difference between questionable scientific practices and practices that constitute misconduct, avoids both, but knows how each should be handled.

The ethical statistician:

1. Avoids condoning or appearing to condone incompetent or unethical practices in statistical analysis.
2. Recognizes that differences of opinion and honest error do not constitute misconduct; they warrant discussion, but not accusation.
3. Knows the definitions of, and procedures relating to, misconduct. If involved in a misconduct investigation, follows prescribed procedures.
4. Maintains confidentiality during an investigation, but discloses the investigation results honestly to appropriate parties and stakeholders once they are available.
5. Following an investigation of misconduct, supports the appropriate efforts of all involved, including those reporting the possible scientific error or misconduct, to resume their careers in as normal a manner as possible.
6. Avoids, and acts to discourage, retaliation against or damage to the employability of those who responsibly call attention to possible scientific error or misconduct.

H. Responsibilities of Employers, Including Organizations, Individuals, Attorneys, or Other Clients Employing Statistical Practitioners

Those employing any person to analyze data are implicitly relying on the profession's reputation for objectivity. However, this creates an obligation on the part of the employer to understand and respect statisticians' obligation of objectivity.
Those employing statisticians are expected to:

1. Recognize that the Ethical Guidelines exist, and were instituted, for the protection and support of the statistician and the consumer alike.[16]
2. Recognize that valid findings result from competent work in a moral environment. Employers, funders, or those who commission statistical analysis have an obligation to rely on the expertise and judgment of qualified statisticians for any data analysis. This obligation may be especially relevant in analyses that are known or anticipated to have tangible physical, financial, or psychological impacts.[17]
3. Recognize that the results of valid statistical studies cannot be guaranteed to conform to the expectations or desires of those commissioning the study or the statistical practitioner(s).
4. Recognize that it is contrary to these Guidelines to report or follow only those results that conform to expectations without explicitly acknowledging competing findings and the basis for choices regarding which results to report, use, and/or cite.[18]
5. Recognize that the inclusion of statistical practitioners as authors, or acknowledgement of their contributions to projects or publications, requires their explicit permission because it implies endorsement of the work.
6. Support sound statistical analysis and expose incompetent or corrupt statistical practice.
7. Strive to protect the professional freedom and responsibility of statistical practitioners who comply with these Guidelines.[19]

Discussion

[1] The use of statistics in medical diagnoses and biomedical research may affect whether individuals live or die, whether their health is protected or jeopardized, and whether medical science advances or becomes sidetracked. Life, death, and health may be at stake in statistical analyses of occupational, environmental, or transportation safety. Early detection and control of new or recurrent infectious diseases depend on sound epidemiological statistics. Mental and social health may be at stake in psychological and sociological applications of statistical analysis.

Effective functioning of the economy depends on the availability of reliable, timely, and properly interpreted economic data. The profitability of individual firms depends in part on their quality control and market research, both of which should rely on statistical methods. Agricultural productivity benefits greatly from statistically sound applications to research and output reporting. Governmental policy decisions regarding public health, criminal justice, social equity, education, the environment, national defense, and security depend in part on sound statistics.

[2] Misconduct is not limited to instances of plagiarism and data fabrication or falsification. More broadly, misconduct includes all professional dishonesty, by commission or omission, and, within the realm of professional activities and expression, all harmful disrespect for people, unauthorized use of their intellectual and physical property, and unjustified detraction from their reputations.

[3] Typically, each study should be based on a competent understanding of the subject-matter issues and statistical protocols that are clearly defined for the stage (exploratory, intermediate, or final) of analysis before looking at those data that will be decisive for that stage and technical criteria to justify both the practical relevance of the study and the amount of data to be used.

[4] Running multiple tests on the same data set at the same stage of an analysis increases the chance of obtaining at least one invalid result. Selecting the one "significant" result from a multiplicity of parallel tests poses a grave risk of an incorrect conclusion. Failure to disclose the full extent of tests and their results in such a case would be highly misleading.

[5] For example, address the multiple potentially confounding factors in observational studies and use due caution in drawing causal inferences. The fact that a procedure is automated does not ensure its correctness or appropriateness; it is also necessary to understand the theory, data, and methods used in each statistical study.

[6] Preferably, authorship order in statistical publications should be by degree of intellectual contribution to the study and material to be published, to the extent that such ordering can feasibly be determined. When some other rule of authorship order is used in a statistical publication, the rule should be disclosed in a footnote or endnote. Where authorship order by contribution is assumed by those making decisions about hiring, promotion, or tenure, for example, failure to disclose an alternative rule may improperly damage or advance careers.

[7] This may sometimes require divestiture of the conflicting personal interest or withdrawal from the professional activity.

Examples where conflict of interest may be problematic include grant reviews, other peer reviews, and tensions between scholarship and personal or family financial interests.

[8] For the general public, convey the scope, relevance, and conclusions of a study without technical distractions. For the professional literature, strive to answer the questions likely to occur to your peers.

[9] For example, disclose any significant failure to follow through fully on an agreed sampling or analytic plan and explain any resulting adverse consequences. Address the suitability of the analytic methods and their inherent assumptions relative to the circumstances of the specific study. Identify the computer routines used to implement the analytic methods.

[10] Statisticians are encouraged to participate in professional activities contributing to the improvement of the community and to work that elevates the statistical profession in the United States and the world. It is recognized that the ability to do pro bono work may be limited by the conditions of the statistician's employment and personal situations, but statisticians should be open to opportunities for pro bono and other work, including service to the local community or to international organizations. Service to the profession—including service on ASA committees, sections, and chapters—is also encouraged.

[11] Statistical methods may be broadly applicable to many classes of problem or application. Statistical innovators may well be entitled to monetary or other rewards for their writings, software, or research results.

[12] Ensure adequate planning to support the practical value of the research, validity of expected results, ability to provide the protection promised, and consideration of all other ethical issues involved.

[13] These recommendations may be based on prospective power analysis, the planned precision of the study endpoint(s), or other methods to ensure appropriate scope to either frequentist or Bayesian approaches. Study scope also should take into consideration the feasibility of obtaining research subjects and the value of the data elements to be collected.

[14] Appropriate approval is not necessarily given by the research subjects themselves. Approval may be required from a parent or legal guardian who must consent, and the actual participant may be required to assent in some cases. This includes legal guardians for minors and adults who are unable to give consent. In animal studies, statisticians should enquire as to proper forms of owner or institutional consent for animals that are part of a research project.

In all cases, avoid or minimize the use of deception. Where it is necessary and provides significant knowledge—as in some psychological, sociological, and other research—ensure prior independent ethical review of the protocol and continued monitoring of the research. Where full disclosure of study parameters to subjects or other investigators is not advisable, as in some randomized clinical trials, generally inform them of the nature of the information withheld and the reason for withholding it. As with deception, ensure independent ethical review of the protocol and continued monitoring of the research.

[15] All research is constrained by resources, including the available funding, timing, and personnel. Thus few, if any, studies can achieve the highest standards imaginable. However, the statistician should always strive to achieve the most valid results with the resources available and avoid working on studies likely to produce misleading or meaningless results.

[16] This safeguard will lower your risk of learning only after the fact that you have collaborated on an unethical study.

[17] In cases of conflict, statistical practitioners and those employing them are encouraged to resolve issues of ethical practice privately. If private resolution is not possible, recognize that statistical practitioners have an ethical obligation to expose incompetent or corrupt practice before it can cause harm to research subjects or society at large.

[18] Any measures taken to ensure a particular outcome will lessen the validity of the analysis. Pressure on a statistical practitioner to deviate from these guidelines is likely to damage both the validity of study results and the professional credibility of the practitioner.

[19] Within organizations and within professions using statistical methods generally, statistics practitioners with greater prestige, power, or status have a responsibility to protect the professional freedom and responsibility of more subordinate statistical practitioners who comply with these guidelines.