Statistical proof

Statistical proof is the rational demonstration of degree of certainty for a proposition, hypothesis or theory that is used to convince others subsequent to a statistical test of the supporting evidence and the types of inferences that can be drawn from the test scores. Statistical methods are used to increase the understanding of the facts and the proof demonstrates the validity and logic of inference with explicit reference to a hypothesis, the experimental data, the facts, the test, and the odds. Proof has two essential aims: the first is to convince and the second is to explain the proposition through peer and public review.[1]

The burden of proof rests on the demonstrable application of the statistical method, the disclosure of the assumptions, and the relevance that the test has with respect to a genuine understanding of the data relative to the external world. There are adherents to several different statistical philosophies of inference, such as Bayes theorem versus the likelihood function, or positivism versus critical rationalism. These methods of reason have direct bearing on statistical proof and its interpretations in the broader philosophy of science.[1][2]

A common demarcation between science and non-science is the hypothetico-deductive proof of falsification developed by Karl Popper, which is a well-established practice in the tradition of statistics. Other modes of inference, however, may include the inductive and abductive modes of proof.[3] Scientists do not use statistical proof as a means to attain certainty, but to falsify claims and explain theory. Science cannot achieve absolute certainty nor is it a continuous march toward an objective truth as the vernacular as opposed to the scientific meaning of the term "proof" might imply. Statistical proof offers a kind of proof of a theory's falsity and the means to learn heuristically through repeated statistical trials and experimental error.[2] Statistical proof also has applications in legal matters with implications for the legal burden of proof.[4]

Contents

Axioms

Test and proof

Bayes theorem

In legal proceedings

See also

References

Notes

Axioms

There are two kinds of axioms, 1) conventions that are taken as true that should be avoided because they cannot be tested, and 2) hypotheses.[5] Proof in the theory of probability was built on four axioms developed in the late 17th century:

1. The probability of a hypothesis is a non-negative real number: $\left\{ \mathbf{Pr}(h) \geqq 0 \right\}$;

2. The probability of necessary truth equals one: $\left\{ \mathbf{Pr}(t) = 1 \right\}$;

3. If two hypotheses h_1 and h_2 are mutually exclusive, then the sum of their probabilities is equal to the probability of their disjunction: $\left\{ \mathbf{Pr}(h_1) + \mathbf{Pr}(h_2) = \mathbf{Pr}(h_1 \, or \, h_2) \right\}$;

4. The conditional probability of h_1 given h_2 $\left\{\mathbf{Pr}(h_1|h_2)\right\}$ is equal to the unconditional probability $\left\{\mathbf{Pr}(h_1 \& h_2)\right\}$ of the conjunction h_1 and h_2, divided by the unconditional probability $\left\{\mathbf{Pr}(h_2)\right\}$ of h_2 where that probability is positive $\left\{\dfrac{\mathbf{Pr}(h_1|h_2) = \mathbf{Pr}(h_1 \& h_2)}{\mathbf{Pr}(h_2)}\right\}$, where $\left\{\mathbf{Pr}(h_2) > 0\right\}$.

The preceding axioms provide the statistical proof and basis for the laws of randomness, or objective chance from where modern statistical theory has advanced. Experimental data, however, can never prove that the hypotheses (h) is true, but relies on an inductive inference by measuring the probability of the hypotheses relative to the empirical data. The proof is in the rational demonstration of using the logic of inference, math, testing, and deductive reasoning of significance.[1][2][6]

Test and proof

The term *proof* descended from its Latin roots (provable, probable, *probare* L.) meaning *to test*.[7][8] Hence, proof is a form of inference by means of a statistical test. Statistical tests are formulated on models that generate probability distributions. Examples of probability distributions might include the binary, normal, or poisson distribution that give exact descriptions of variables that behave according to natural laws of random chance. When a statistical test is applied to samples of a population, the test determines if the sample statistics are significantly different from the assumed null-model. True values of a population, which are unknowable in practice, are called parameters of the population. Researchers sample from populations, which provide estimates of the parameters, to calculate the mean or standard deviation. If the entire population is sampled, then the sample statistic mean and distribution will converge with the parametric distribution.[9]

Using the scientific method of falsification, the probability value that the sample statistic is sufficiently different from the null-model than can be explained by chance alone is given prior to the test. Most statisticians set the prior probability value at 0.05 or 0.1, which means if the sample statistics diverge from the parametric model more than 5 (or 10) times out of 100, then the discrepancy is unlikely to be explained by chance alone and the null-hypothesis is rejected. Statistical models provide exact outcomes of the parametric and estimates of the sample statistics. Hence, the burden of proof rests in the sample statistics that provide estimates of a statistical model. Statistical models contain the mathematical proof of the parametric values and their probability distributions.[10][11]

Bayes theorem

Bayesian statistics are based on a different philosophical approach for proof of inference. The mathematical formula for Bayes's theorem is:

$$Pr[Parameter|Data] = \frac{Pr[Data|Parameter] \times Pr[Parameter]}{Pr[Data]}$$

The formula is read as the probability of the parameter (or hypothesis =h, as used in the notation on axioms) "given" the data (or empirical observation), where the horizontal bar refers to "given". The right hand side of the formula calculates the prior probability of a statistical model (Pr [Parameter]) with the likelihood (Pr [Data | Parameter]) to produce a posterior probability distribution of the parameter (Pr [Parameter | Data]). The posterior probability is the likelihood that the parameter is correct given the observed data or samples statistics.[12] Hypotheses can be compared using Bayesian inference by means of the Bayes factor, which is the ratio of the posterior odds to the prior odds. It provides a measure of the data and if it has increased or decreased the likelihood of one hypothese relative to another.[13]

The statistical proof is the Bayesian demonstration that one hypothesis has a higher (weak, strong, positive) likelihood.[13] There is considerable debate if the Bayesian method aligns with Karl Poppers method of proof of falsification, where some have suggested that "...there is no such thing as "accepting" hypotheses at all. All that one does in science is assign degrees of belief..."[14]:180 According to Popper, hypotheses that have withstood testing and have yet to be falsified are not verified but corroborated. Some

researches have suggested that Popper's quest to define corroboration on the premise of probability put his philosophy in line with the Bayesian approach. In this context, the likelihood of one hypothesis relative to another may be an index of corroboration, not confirmation, and thus statistically proven through rigorous objective standing.[6][15]

In legal proceedings

Statistical proof in a legal proceeding can be sorted into three categories of evidence:

1. The occurrence of an event, act, or type of conduct,
2. The identity of the individual(s) responsible
3. The intent or psychological responsibility[16]

> "Where gross statistical disparities can be shown, they alone may in a proper case constitute *prima facie* proof of a pattern or practice of discrimination."[nb 1]:271

Statistical proof was not regularly applied in decisions concerning United States legal proceedings until the mid 1970's following a landmark jury discrimination case in *Castaneda v. Partida*. The US Supreme Court ruled that gross statistical disparities constitutes "*prima facie* proof*" of discrimination, resulting in a shift of the burden of proof from plaintiff to defendant. Since that ruling, statistical proof has been used in many other cases on inequality, discrimination, and DNA evidence.[4][17][18] However, there is not a one-to-one correspondence between statistical proof and the legal burden of proof. "The Supreme Court has stated that the degrees of rigor required in the fact finding processes of law and science do not necessarily correspond."[18]:1533

In an example of a death row sentence (*McCleskey v. Kemp*[nb 2]) concerning racial discrimination, the petitioner, a black man named McCleskey was charged with the murder of a white police officer during a robbery. Expert testimony for McClesky introduced a statistical proof showing that "defendants charged with killing white victims were 4.3 times as likely to receive a death sentence as charged with killing blacks.".[19]:595 Nonetheless, the statistics was insufficient "to prove that the decisionmakers in his case acted with discriminatory purpose."[19]:596 It was further argued that there were "inherent limitations of the statistical proof",[19]:596 because it did not refer to the specifics of the individual. Despite the statistical demonstration of an increased probability of discrimination, the legal burden of proof (it was argued) had to be examined on a case by case basis.[19]

See also

- Mathematical proof
- Data analysis

References

1. Gold, B.; Simons, R. A. (2008). *Proof and other dilemmas: Mathematics and philosophy* (https://books.google.com/books?id=wPhwJdjI-dIC&printsec=frontcover&dq=Proof+and+other+dilemmas:+mathematics+and+philosophy&hl=en&ei=SVq7TvQ4g-KIAteutIkC&sa=X&oi=book_result&ct=book-thumbnail&resnum=1&ved=0CDEQ6wEwAA#v=onepage&q&f=false). Mathematics Association of America Inc. ISBN 0-88385-567-4.

2. Gattei, S. (2008). *Thomas Kuhn's "Linguistic Turn" and the Legacy of Logical Empiricism: Incommensurability, Rationality and the Search for Truth* (https://books.google.com/books?id=HDOuKjMc0AEC&printsec=frontcover&dq=Thomas+Kuhn's+%22linguistic+turn%22+and+the+legacy+of+logical+empiricism&hl=en&ei=Xz-8TpznC6aPiALPn6mdAw&sa=X&oi=book_result&ct=result&resnum=1&ved=0CC4Q6AEwAA#v=onepage&q&f=false). Ashgate Pub Co. p. 277. ISBN 0-7546-6160-1.

3. Pedemont, B. (2007). "How can the relationship between argumentation and proof be analysed?" (http://www.springerlink.com/content/424p8r420820316l/). *Educational Studies in Mathematics* **66** (1): 23–41. doi:10.1007/s10649-006-9057-x (https://doi.org/10.1007%2Fs10649-006-9057-x).

4. Meier, P. (1986). "Damned Liars and Expert Witnesses" (http://users.stat.umn.edu/~sandy/courses/8801/articles/Law/meierlaw.pdf) (PDF). *Journal of the American Statistical Association* **81** (394): 269–276. doi:10.1080/01621459.1986.10478270 (https://doi.org/10.1080%2F01621459.1986.10478270).

5. Wiley, E. O. (1975). "Karl R. Popper, Systematics, and Classification: A Reply to Walter Bock and Other Evolutionary Taxonomists". *Systematic Zoology*. Society of Systematic Biologists **24** (2): 233–43. doi:10.2307/2412764 (https://doi.org/10.2307%2F2412764). ISSN 0039-7989 (https://www.worldcat.org/issn/0039-7989). JSTOR 2412764 (https://www.jstor.org/stable/2412764) – via JSTOR. (Registration required (help)).

6. Howson, Colin; Urbach, Peter (1991). "Bayesian reasoning in science". *Nature*. **350** (6317): 371–4. doi:10.1038/350371a0 (https://doi.org/10.1038%2F350371a0). ISSN 1476-4687 (https://www.worldcat.org/issn/1476-4687). (Registration required (help)).

7. Sundholm, G. "Proof-Theoretical Semantics and Fregean Identity Criteria for Propositions" (https://openaccess.leidenuniv.nl/bitstream/handle/1887/11990/9_054_019.pdf?sequence=1) (PDF). *The Monist*. **77** (3): 294–314. doi:10.5840/monist199477315 (https://doi.org/10.5840%2Fmonist199477315).

8. Bissell, D. (1996). "Statisticians have a Word for it" (http://www.rsscse-edu.org.uk/tsj/wp-content/uploads/2011/03/bissell1.pdf) (PDF). *Teaching Statistics*. **18** (3): 87–89. doi:10.1111/j.1467-9639.1996.tb00300.x (https://doi.org/10.1111%2Fj.1467-9639.1996.tb00300.x).

9. Sokal, R. R.; Rohlf, F. J. (1995). *Biometry* (https://books.google.com/books?id=N6KCNw5NHNkC&dq=biometry&hl=en&ei=_ka7TuLtCObmiAKfuKSDAg&sa=X&oi=book_result&ct=book-thumbnail&resnum=1&ved=0CDUQ6wEwAA) (3rd ed.). W.H. Freeman & Company. p. 887. ISBN 0-7167-2411-1.

10. Heath, David (1995). *An introduction to experimental design and statistics for biology* (https://books.google.com/books?id=sJVHoJ2ML40C&printsec=frontcover&dq=statistics+for+biologists&hl=en&ei=9e-6TGzMq3MiQKrkfiODA&sa=X&oi=book_result&ct=book-thumbnail&resnum=7&ved=0CEgQ6wEwBg#v=onepage&q=parametric&f=false). CRC Press. ISBN 1-85728-132-2.

11. Hald, Anders (2006). *A History of Parametric Statistical Inference from Bernoulli to Fisher, 1713-1935* (https://books.google.com/books?id=Uc9C90KKW_UC&printsec=frontcover&dq=A+history+of+parametric+statistical+inference+from+Bernoulli+to+Fisher&hl=en&ei=_fe6Tui3MliuiQKhw5yMDA&sa=X&oi=book_result&ct=book-thumbnail&resnum=1&ved=0CDAQ6wEwAA#v=onepage&q&f=false). Springer. p. 260. ISBN 0-387-46408-5.

12. Huelsenbeck, J. P.; Ronquist, F.; Bollback, J. P. (2001). "Bayesian Inference of Phylogeny and Its Impact on Evolutionary Biology" (http://bama.ua.edu/~molsyst/page2/assets/Huelsenbeck2001.pdf) (PDF). *Science*. **294** (5550): 2310–2314. doi:10.1126/science.1065889 (https://doi.org/10.1126%2Fscience.1065889).

13. Wade, P. R. (2000). "Bayesian methods in conservation biology" (http://www.utm.utoronto.ca/~collinsn/1310/Nov%203/optional/Bayesian%20methods%20in%20conserv%20biol.pdf) (PDF). *Conservation Biology*. **14** (5): 1308–1316. doi:10.1046/j.1523-1739.2000.99415.x (https://doi.org/10.1046%2Fj.1523-1739.2000.99415.x).

14. Sober, E. (1991). *Reconstructing the Past: Parsimony, Evolution, and Inference* (https://books.google.com/books?id=WlUL_r06GnEC&pg=PA180&lpg=PA180&dq=%22All+that+one+does+in+science+is+assign+degrees+of+belief%22&source=bl&ots=ii2pXy2BzM&sig=FYFilXaMbM-9rRNrJLDs0rhBrD0&hl=en&ei=s3287ZHKiZiQLA9sGIAw&sa=X&oi=book_result&ct=result&resnum=2&sqi=2&ved=0CCIQ6AEwAQ#v=onepage&q=%22All%20that%20one%20does%20in%20science%20is%20assign%20degrees%20of%20belief%22&f=false). A Bradford Book. p. 284. ISBN 0-262-69144-2.

15. Helfenbein, K. G.; DeSalle, R. (2005). "Falsifications and corroborations: Karl Popper's influence on systematics" (http://research.amnh.org/users/desalle/pdf/Helfenbein.2005.MPE.pdf) (PDF). *Molecular Phylogenetics and Evolution*. **35**: 271–280. doi:10.1016/j.ympev.2005.01.003 (https://doi.org/10.1016%2Fj.ympev.2005.01.003).

16. Fienberg, S. E.; Kadane, J. B. "The presentation of Bayesian statistical analyses in legal proceedings". *Journal of the Royal Statistical Society Series D*. **32** (1/2): 88–98. doi:10.2307/2987595 (https://doi.org/10.2307%2F2987595). JSTOR 2987595 (https://www.jstor.org/stable/2987595).

17. Garaud, M. C. (1990). "Legal Standards and Statistical Proof in Title VII Litigation: In Search of a Coherent Disparate Impact Model". *University of Pennsylvania Law Review*. **139** (2): 455–503. JSTOR 3312286 (https://www.jstor.org/stable/3312286).

18. The Harvard Law Review Association (1995). "Developments in the Law: Confronting the New Challenges of Scientific Evidence". *Harvard Law Review*. **108** (7): 1481–1605. doi:10.2307/1341808 (https://doi.org/10.2307%2F1341808). JSTOR 1341808 (https://www.jstor.org/stable/1341808).

19. Faigman, D. L. (1991). "Normative Constitutional Fact-Finding": Exploring the Empirical Component of Constitutional Interpretation". *University of Pennsylvania Law Review*. **139** (3): 541–613. JSTOR 3312337 (https://www.jstor.org/stable/3312337).

Notes

1. Supreme Court of the United States *Castaneda v. Partida*, 1977 [1] (https://www.law.cornell.edu/supct/html/historics/USSC_CR_0430_0482_ZS.html) cited in Meier (1986) Ibid. who states "Thus, in the space of less than half a year the Supreme Court had moved from the traditional legal disdain for statistical proof to a strong endorsement of it as being capable, on its own, of establishing a prima facie case against a defendant."[4]
2. 481 U.S. 279 (1987).[19]

Retrieved from "https://en.wikipedia.org/w/index.php?title=Statistical_proof&oldid=841096266"