Bayesian isn't frequentist

Guangji Zou (Ray)
Student ID: 301216239
Email: guangjiz@sfu.ca
Instructor: Jack Davis

August 6, 2018

1 Introduction

We have now learned about two schools of statistical inference: Bayesian and frequentist. The essential difference between Bayesian and frequentist statisticians is in how probability is used. Frequentists use probability only to model certain processes broadly described as "sampling." Bayesians use probability more widely to model both sampling and other kinds of uncertainty. Frequentist inference, however, tends to be less computationally oriented as it depends more on likelihood than it does on the probability of the hypothesis notice.

2 Content

Bayes' formula, a perfectly abstract statement about conditional probabilities of events:

$$P(H|D) = \frac{P(D|H) \times P(H)}{P(D)}$$

H is a hypothesis and D is data which may give evidence for or against H. Each term in Bayes' formula has a name and a role. The prior P(H) is the probability that H is true before the data is considered. The posterior $P(H|D)$ is the probability that H is true after the data is considered. The likelihood $P(D|H)$ is the evidence about H provided by the data D. P(D) is the total probability of the data taking into account all possible hypotheses.

However, the frequentist method only uses conditional distributions of data given specific hypotheses. The presumption is that some hypothesis (parameter specifying the conditional distribution of the data) is true and that the observed data is sampled from that distribution. The frequentist approach does not depend on a subjective prior that may vary from one investigator to another.

Frequentist measures like p-values and confidence intervals continue to dominate research, especially in the life sciences. However, in the current era of powerful computers and big data, Bayesian methods have undergone an enormous renaissance in fields like machine learning and genetics. There are now several large, ongoing clinical trials using Bayesian protocols, something that would have been hard to imagine a generation ago. While professional divisions remain, the consensus forming among top statisticians is that the most effective approaches to complex problems often draw on the best insights from both schools working in concert.

3 Ideas

Here are some different ideas between the frequentist school and the Bayesian school:

1) The frequentist school thinks that the sampling population is infinite. In the infinite sampling population, the rules of the decision can be very accurate; while the Bayes' school thinks that the world is constantly changing, the unknown variables and events have a certain probability. This probability will change the state of the world at any time (the posterior probabilities mentioned above are corrections to prior probabilities).

2) The frequentist school believes that the parameters of the model are fixed. After a model is sampled at numerous times, all the parameters should be the same; and the Bayes' school believes that the data should be fixed. Our rules come from our observation and understanding of the world. What we see is true and correct. The point of observation of things is to estimate the parameters.

3) The frequentist school holds that no model has a prior, but prior plays an important role in the Bayesian school.

4) The frequentist school advocates an evaluation paradigm. It has no prior and is more objective. The Bayes' school advocates a model method. By establishing an unknown parameter model. Before the sample is observed, all parameters are uncertain. The parameters are estimated using the observed sample values. The obtained parameters are brought into the model to make the model. The posterior model best fits the observed data.

The controversy between the Bayes' school and the frequentist school is the prior distribution. The Bayes' school thinks that the prior distribution can be subjective, and it does not need frequentist explanation. The frequentist school believes that only a prior distribution does not depend on the subjective meaning and can be allowed to use a prior distribution in statistical inference when it is determined according to the appropriate theory or previous experience, otherwise the objectivity will be lost.

First, the theory of frequentist first establishes an null model, and then calculates the possibility of the parameters obtained from the actual data under the premise of the invalid model. If this possibility is very small, we think that the null model is not established so as to choose the alternative model; and Bayesian theory is concerned with a certain model on the premise of the current data. The probability of classification is obtained by a specific probability value, which is not used to judge a hypothesis.

Second, the interpretation of probability in the theory of frequentist is the frequency of an event occurring in a longer period of time; the interpretation of probability by Bayes is the degree of recognition of the occurrence of an event.

Third, Bayes' theory is good at making use of past knowledge and sampling data, while the frequentist theory only uses sampling data. Therefore, the posterior probability distribution obtained from Bayesian inference can be used as the prior probability.

Fourth, the different interpretation of the confidence interval: the 95 percent confidence interval in the frequentist theory is explained as: 95 of the 100 confidence intervals obtained by 100 samples are included in the total parameters and 5 are not, but can not be interpreted as the possibility that there are 95 percent of the probability in one sampling. This is because the general parameters in classical statistics are regarded as a constant value and can not be interpreted from the perspective of probability; the confidence interval of Bayes' theory can be interpreted as a form of probability, because in Bayes' analysis, the overall parameter is a random variable, but not a constant value.

If all variables are random variables, then many concepts of the frequentist school will be meaningless. Unbiased estimation, for example,

If E (T) =t, the statistics T is unbiased estimator of unknown parameter t. If the parameter t is a random variable, that equal sign is meaningless, because the expected E (T) of the statistics T is a number, and it can't be equal to a random variable, except in the case of degenerate distribution.

4 Conclusion

In addition, when explaining the meaning of confidence interval, it does not have to be as effortless as the frequentist school. Which unknown parameters are unknown and which are fixed values, and that the interval is a random interval, because the endpoints of the interval are statistics and are also random variables. Each time the interval estimation is different with the observed samples, the interval estimates are different. When the number of experiments is large enough, about 95 percent of the interval contains the fixed unknown parameter.

References

[Cir14] Prajwal Ciryam. *The Management of Proteome Metastability in Health and Disease*. PhD thesis, Northwestern University, 2014.

[DWGTS10] Sarah Donegan, Paula Williamson, Carrol Gamble, and Catrin Tudur-Smith. Indirect comparisons: a review of reporting and methodological quality. *PloS one*, 5(11):e11054, 2010.

[OB] Jeremy Orloff and Jonathan Bloom. Comparison of frequentist and bayesian inference. class 20, 18.05, spring 2014.

[Ste12] Jacob Steinhardt. Beyond bayesians and frequentists. 2012.

[Val] Jordi Vallverdu. A view on statistical schools.

[WSC+09] GA Wells, SA Sultan, L Chen, M Khan, and D Coyle. Indirect evidence: indirect treatment comparisons in meta-analysis. *Ottawa: Canadian Agency for Drugs and Technologies in Health*, pages 1–94, 2009.