

Introduction to Bayesian Methods for Medical Research

S B Tan,*^{PhD}

Abstract

Bayesian statistical methods provide a credible alternative to traditional statistical approaches used in medical research. In this article, the author gives a brief introduction to the key concepts involved in Bayesian methods and compares them with some commonly used non-Bayesian statistical techniques.

Ann Acad Med Singapore 2001; 30:444-6

Key words: *Bayesian approach, Probability distribution, Statistical methodology*

Introduction

The traditional statistical approach used in most areas of medical research involves what is commonly known as frequentist statistics. Bayesian methods provide an increasingly popular alternative to frequentist approaches.¹⁻³ The aim of this article is to give a brief introduction to Bayesian methods and to compare these with some commonly used frequentist procedures.

Examples of Some Commonly Used Frequentist Statistical Methods

A very brief overview of two popular frequentist statistical techniques used in medical research is given below. For more details on these and other frequentist methods, the reader is referred to Altman⁴ or any other introductory statistical text.

Hypothesis Tests

One of the most common problems that arise in medical research is that of wanting to compare two groups of subjects. For example, we may wish to compare whether there is a difference in systolic blood pressure between male and female subjects at age 30 within a certain population of interest. Or, in the context of a clinical trial, we may wish to compare whether patients given a new adjuvant therapy have a longer median survival than patients treated with surgery alone.

Subject to satisfying certain conditions such as the existence of a normally distributed population, a *t*-test will often be used to make the above comparisons. In essence,

a *t*-test involves the specification of a null hypothesis (typically this will be that there is no difference between the mean values of the two groups), evaluating the corresponding *t*-statistic and finding the '*P* value' corresponding to this *t*-statistic. A *P* value of (typically) less than 0.05 will allow us to conclude that there is a 'significant' difference between the two groups.

Estimation and Confidence Intervals

Hypothesis tests are only able to tell us if there is a difference between the two groups, but not how large this difference is. So, using a hypothesis test, we might be able to conclude, for example, that 30-year-old males have a higher systolic blood pressure than 30-year-old females within the population of interest. However, we would not be able to tell how much higher their blood pressure was.

Should we wish to determine this, we need to obtain an estimate of the population mean difference in blood pressures between the two groups (call this the true parameter of interest value). We then estimate this true parameter value by the sample mean or 'sample average'. However, this is only a point estimate, and a (usually 95%) confidence interval is often also evaluated. This is meant to reflect the uncertainty associated with the parameter estimate, and can be easily calculated using standard formulae combining the mean, standard deviation and sample size values.

A Bayesian Alternative

Many medical researchers often misinterpret the meaning of *P* values and frequentist confidence intervals. The

* Biostatistician

Division of Clinical Trials and Epidemiological Sciences
National Cancer Centre

Address for Reprints: Dr Tan Say Beng, Division of Clinical Trials and Epidemiological Sciences, National Cancer Centre, 11 Hospital Drive, Singapore 169610.

E-mail: ctetsb@nccs.com.sg

P value is often misunderstood to be the probability that the null hypothesis is true. This is not the case and the actual definition of the *P* value is that it is the probability of having observed our data (or more extreme data) when the null hypothesis is true. Similarly, the probability that the true parameter value is contained in a ‘95% confidence interval’ is not 0.95, as we would intuitively expect. Instead, the actual definition of a frequentist ‘95% confidence interval’ is as follows. Suppose the estimation process was to be repeated many, many times. Then 95% of the calculated ‘95% confidence intervals’ would be expected to contain the true parameter value.

These definitions are difficult for many medical researchers to fully understand, let alone correctly interpret. What would be more useful for researchers, would be if they could determine, for example, what is the probability that patients given the adjuvant therapy had a higher (or lower) median survival than patients treated with surgery alone. Or in the case of the estimation problem, what is the interval that contains the true parameter value with probability 0.95.

A Bayesian analysis is able to provide the answers to these and other questions of interest. Further, it can also provide a graphical representation of the actual probability distribution of the true parameter value. Such a representation will not only give information regarding the mean of the parameter value, but also its uncertainty and the shape of its distribution.

Another useful feature of a Bayesian approach is that it allows for the formal incorporation of relevant external information in the analysis. Thus, in the context of a clinical trial for example, the trial data can be explicitly combined with data from pilot studies, similar trials and even subjective clinical opinion as part of a Bayesian analysis. Any conclusions made will then have been formally based on all available sources of evidence known to the investigator.

The Bayesian Methodology

A brief overview of how a Bayesian analysis might be carried out as an alternative to conducting hypothesis tests or obtaining sample estimates and confidence intervals is given below.

The first thing to note is that, in a Bayesian framework, regardless of whether we have a frequentist hypothesis test problem or one of point estimation or any other standard statistical problem, the Bayesian approach is essentially the same. We start with a suitable prior distribution, combine this with the data to obtain the posterior distribution and finally analyse and draw conclusions based on the posterior distribution. Each of these steps is now addressed in turn.

Prior Distribution

This is a probability distribution that represents the prior information associated with the parameter of interest. There are many types of prior distributions and many ways to derive them (see Berry and Stangl² and Spiegelhalter et al³ for some examples). They can be based on information available in the literature, clinical databases, expert opinion or information from any other appropriate source. A prior distribution can also be non-informative, that is representing the state of very little or no relevant prior information. A Bayesian analysis with a non-informative prior will usually give very similar conclusions to what is obtained using a frequentist analysis. However, the output and presentation of the analysis will still be very different (and often more intuitive and helpful).

Combining the Prior with the Data

Once a prior distribution has been obtained, it needs to be combined with the data to give a posterior distribution. This process is fundamental to the Bayesian approach and is best explained graphically. The top two plots of Figure 1 show a prior distribution and likelihood (distribution of the data) corresponding to the mean difference in systolic blood pressure between 30-year-old males and 30-year old females in a certain population. The plots are based on hypothetical data and are meant to just illustrate the Bayesian procedure. The third plot at the bottom of the figure is the corresponding posterior distribution that is derived from updating the prior distribution with the likelihood. Formally, this is done via the application of what is known as Bayes theorem. Conceptually, the idea involves ‘merging’ the prior distribution with the likelihood to give a new combined, posterior distribution.

Thus, we see from Figure 1 that the prior distribution is ‘merged’ with the likelihood to give a final posterior distribution which is, in a sense, an ‘average’ of the two distributions. Note, however, that the posterior distribution has a smaller variance (that is, it has a tighter distribution) than either the prior distribution or likelihood. This reflects the greater amount of information present in the posterior distribution since it combines information from both the prior and the likelihood.

Analysing the Posterior Distribution

Various analyses can now be performed on the posterior distribution. Among these are computing the mean of the posterior distribution and the 95% Bayesian confidence interval (given by: posterior mean $\pm 1.96 \times$ posterior SD). Note that, unlike in the case of the frequentist confidence interval, the true parameter value has a probability of 0.95 of being contained in the 95% Bayesian confidence interval.

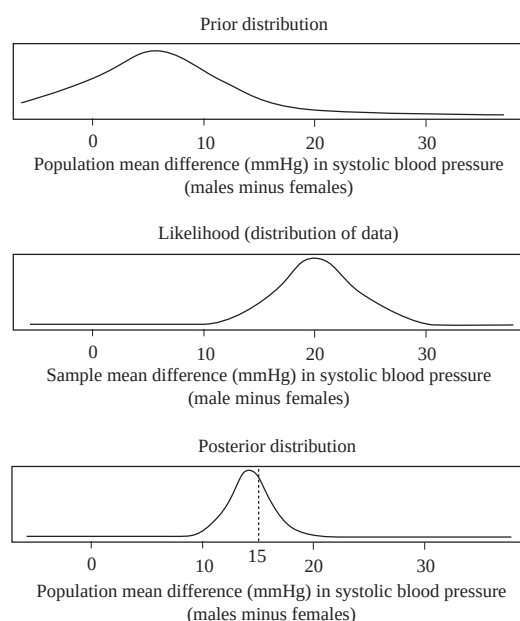


Fig.1. Hypothetical prior distribution, likelihood and posterior distribution of the difference in systolic blood pressure between 30-year-old males and females

Other statistics can also be computed, such as the probability that the mean male systolic blood pressure is at least 15 mmHg higher than females. This is just the area under the curve to the right of the dotted line in the posterior distribution plot given in Figure 1. Other similar probabilities can also be computed, depending on what is of clinical interest and relevance.

Conclusion

Bayesian methods are becoming increasingly popular, not just in medical research but also in fields as diverse as

software design, financial prediction and aeronautical engineering. A key reason for this is the advent of fast computer processors, which have allowed many previously very difficult Bayesian computations to now be easily performed. Freely available Bayesian software packages such as BUGS⁵ and an increase in the number of statisticians trained in Bayesian methodology have also encouraged its adoption.

There is now a need for clinicians and other medical researchers to become familiar with the basics of Bayesian methodology. Frequentist statistics is and will continue to be the dominant approach to statistical analysis for many years to come. However, medical researchers need to be aware of the potential of alternative Bayesian approaches, and to ensure that they are in a position to understand and discuss reports of Bayesian analyses given in the literature.

It is the author's hope that this article, together with others like it, would facilitate this process of familiarising medical researchers with some of the key concepts of Bayesian statistics.

REFERENCES

1. Spiegelhalter D J, Myles P, Jones D R, Abrams K R. An introduction to bayesian methods in health technology assessment. *BMJ* 1999; 319:508-12.
2. Berry D A, Stangl D K. Bayesian biostatistics. New York: Marcel Dekker Inc, 1996.
3. Spiegelhalter D J, Freedman L S, Parmar M K B. Bayesian approaches to randomized trials. *J Roy Stat Soc A* 1994; 157:357-87.
4. Altman D G. Practical statistics for medical research. Boca Raton: Chapman and Hall, 1991.
5. Spiegelhalter D, Thomas A, Best N G. BUGS: Bayesian inference using Gibbs sampling. Medical Research Council Biostatistics Unit, 1996.