Teaching the Likelihood Paradigm in Biostatistics

Charles A. Rohde
Johns Hopkins University
Baltimore, Maryland

In this paper it is argued that most of the conventional topics in basic biostatistics and statistics can be adequately discussed using the likelihood paradigm. The advantages of this approach without reference to repeated sampling properties or the addition of prior distribution are emphasized.

Most introductory biostatistics and statistics courses focus on the analysis and interpretation of data. In particular the traditional courses approach to statistical inference focuses on the interpretation of data in the context of a probability model.

A variety of "definitions" of statistical inference have been used to delineate the subject matter typically discussed in basic courses. These definitions range from Fisherian to decision theoretic to Bayesian. It is argued that basic statistics courses should be concerned with emphasizing that statistics is the discipline concerned with statistical evidence: producing it, interpreting it and using it. Loosely defined "statistical evidence consists of observations in the context of a probability model where (i) a model is a collection of probability distributions indexed by a parameter and (ii) observations are conceptualized as having been generated according to one of the distributions in the model.

As emphasized by Royall (1997) observations can lead to answering one of three questions (i) what do I believe, (ii) what do I do and (iii) what does the data say. The first requires input of prior beliefs, the second requires both prior beliefs and statements of consequences of actions (losses). The third which is called interpretation of statistical evidence provided by data has not received the attention it deserves in beginning courses in biostatistics or statistics even though it is the fundamental reason for studying statistics.

It is argued that representing the evidence in data with respect to the probability model which generated the data is best achieved by using the Law of Likelihood and likelihood (support) intervals).

In general we have a family of probability distributions $\boldsymbol{\mathcal{F}}$ indexed by a parameter θ i.e.

$$\boldsymbol{\mathcal{F}} = \{f(\cdot;\theta) \ : \ \theta \in \Theta\}$$

where Θ is the parameter space. $f(\mathbf{y};\theta)$ thus represents the pdf of $\mathbf{Y}$ at $\mathbf{y}$ when θ is the value of the parameter. For a fixed value of $\mathbf{y}$, $f(\mathbf{y};\theta)$ is a function of θ called the **likelihood function** and is denoted by $\mathrm{L}\,(\theta;\mathbf{y})$.

In terms of the likelihood function the Law of Likelihood simply states that θ_2 is better supported by $\mathbf{y}$ than θ_1 if

$$\frac{\mathrm{L}\left(\theta_2;\mathbf{y}\right)}{\mathrm{L}\left(\theta_1;\mathbf{y}\right)} > 1$$

and the magnitude of the ratio gives the strength of the evidence favoring θ_2 vis a vis θ_1. For convenience the likelihood function is normalized so as to have maximum value 1. Any two observations which generate the same likelihood function are equivalent as evidence.

Given the likelihood for a parameter θ we define the $1/k$ unit likelihood interval for θ as

$$\{\theta : \ \mathrm{Lik}_p(\theta;\mathbf{y}) \geq 1/k\}$$

where typical choices of k are 8 and 32. These choices roughly correpsond to the conventional 95 and 99 percent confidence intervals, respectively. The interpretation is however, entirely different. Loosely speaking the $1/k$ likelihood interval consists of those values of θ supported by the data or consistent with the data. Profile likelihoods for parameters of interest can be used when nuisance parameters are present.

Examples are provided to show that the presentation of statistical evidence graphically as a plot of the likelihood or profile likelihood function provides a useful and informative display of the evidence provided by the data. These displays provide the same information as customary frequency based analyses or Bayesian analyses and are useful evidential counterparts to exploratory plots.

Resume

En cet article on lui discute que la plupart des matires conventionnelles dans la biostatistique et les statistiques de base peuvent tre en juste proportion discutes en utilisant le paradigme de probabilit. Les avantages de ceci s'approchent sans rfrence aux properties d'chantillonnage rpts ou l'addition de la distribution antrieure sont soulignes

References

(1) Royall, R.M. (1997) Statistical Evidence: A Likelihood Paradigm. Chapman and Hall

(2) Royall, R.M. (2000) On the Probability of Observing Misleading Statistical Evidence. Journal of the American Statistical Association; 95:760-767