The Bayesian clinician

Kulkarni UP

Seth GS Medical College and KEM Hospital, Parel, Mumbai - 400 012, India

Correspondence:
Uday P Kulkarni
E-mail: upk30@rediffmail.com

It is commonly believed that clinicians use statistics only when they are carrying out a research study. Most clinicians would be surprised if they are told that they (knowingly or unknowingly) use statistics at every step while reaching a diagnosis in each and every patient that they come across.

Basic Concepts and Terminologies

Before we see how clinicians depend upon statistical analysis in everyday practice let us first remind ourselves about the two types of statistical analyses: the frequentist and the Bayesian.[1] Bayesian statistics considers that our prior knowledge regarding any event constitutes the prior possibility of occurrence. Now if we have some new evidence regarding the event, then using this new knowledge, we can arrive at the 'posterior' possibility of occurrence for the event. Thus the Bayesian approach requires 'prior' and 'new knowledge' to arrive at the 'posterior' (or the possibility of occurrence obtained after our analysis). On the other hand the frequentist approach is not based on the concept of prior possibility of occurrence. Thus the possibility of occurrence is judged solely on the basis of 'new knowledge'.

It would at this point also not be out of place to differentiate between probability and odds. These terms are used to quantify the possibility of the occurrence of an event. The "odds" for any event is the ratio of the number of favorable outcomes to the number of unfavorable outcomes. Thus the odds of getting the digit 1 on rolling a fair dice would be 1:5. Probability is defined as chance of occurrence of the outcome. Thus the probability of getting the digit 1 on rolling a fair dice would be 1/6. Thus the relationship between odds (o) and probability (p) is

$$o = p: 1\text{-}p$$
$$\text{or } p = o/o+1$$

Frequentist or Bayesian

While managing a patient, the clinician's first task is to reach a probable clinical diagnosis that would enable him/ her to devise and implement a management plan consisting of investigations and therapeutic interventions. For example, when a patient presents with complaints of fever the clinician lists the common causes of fever in his mind considering the geographical location, season, current prevalence of various etiologies of fever, etc. This could be considered to constitute "prior knowledge". This subjective knowledge helps the clinician assign 'prior odds'. Then he asks the patient certain questions related to other complaints, performs clinical examination and orders for certain tests. The results of these maneuvers constitute 'new knowledge'.

Before we go further, let us remember that each point elicited in history, each clinical examination finding and each test has its own sensitivity, specificity and likelihood ratios. We will discuss these with the example provided in Table 1.

Table 1: Sample test result

Test result	Disease present	Disease absent	Total
Positive	160	10	170
Negative	40	90	130
Total	200	100	300

Figures indicate number of patients

Sensitivity = (Total number of persons with disease who had a positive test result × 100)/ (Total number of persons with disease who were tested)
= 160/ 200 or 80%.

This means that if 100 individuals having the disease are tested, the test will be positive in 80% of them.

False negative rate= 1- sensitivity = 20%
Specificity = (Total number of persons without disease who had a negative test result × 100)/ (Total number of persons without disease who were tested)
= 90/100 or 90%.

This means that if 100 individuals not having the disease are tested, the test will be negative in 90% of them.

False positive rate= 1- specificity = 10%

The likelihood ratio for a positive test result is given by[2]
Sensitivity/(100 - Specificity) = 80/(100-90) = 8

In other words, if the test is positive the patient is 8 times likely to have the disease.

The likelihood ratio for a negative test result is given by[2]
(100 - Sensitivity)/Specificity= (100-80)/90 or 2/9

In other words, if the test is negative, the patient is 2/9 times likely to have the disease.

As one would realize, the likelihood ratios require a concept of a prior possibility of occurrence for their interpretation. For example, after the test result is positive, the patient is 8 times likely to have the disease as compared to his/her prior possibility of occurrence. The likelihood ratios thus represent the Bayesian interpretation of a test.

Sensitivity and specificity are used for frequentist interpretation of the test. In the above example, the frequentist interpretation would be that if the patient has the disease then 80% of the times the test would be positive or if the patient does not have the disease, the test would be negative 90% of the times. Such an approach does not add anything towards answering the question "Does the patient actually have the disease?"

On the contrary, the Bayesian interpretation of the test which would be as follows: "if the test is positive, the patient is 8 times more likely to have the disease" does take us towards the answer to this question.

In general, whether the test result is positive or negative, the Bayesian interpretation of the test result tells us that the patient is some 'x' times likely to have the disease. This likelihood ratio can be directly multiplied with the prior odds to get the posterior odds as both odds and likelihood ratio are ratios.

Posterior or post-test odds = prior odds x likelihood ratio

However, when we are using the prior probability instead of prior odds to express the prior possibility of occurrence, then we need to use a formula to calculate the posterior probability.

Post-test probability
= [(Pre-test probability × test sensitivity)]/
[(Pre-test probability× test sensitivity+ (1- disease prevalence)× test false-positive rate)][3]

We can also use a nomogram to arrive at the posterior probability.[4] Although both odds and probability can be used to express the prior possibility of occurrence, the calculations seem easier when we use odds instead of probability.

Prior knowledge from prevalence of disease
As stated above, prior odds are generally based on subjective information. However, there are situations when prior odds could be derived from known prevalence of the disease. For example, consider that the prevalence of HIV infection in the general population is 0.1% (P= 0.1/100 or 0.001) while the prevalence of HIV infection among intravenous (IV) drug users is 10% (P= 10/100 or 0.1). Consider that we use a diagnostic

test for HIV infection with specificity and sensitivity of 99%. The likelihood ratio of a positive test result in this case is
Sensitivity/(100 - Specificity) = 99/(100-99) = 99.

If the test is positive in two individuals, one who does not have any risk factors for HIV and whose prior odds for HIV are given by:
Prior odds for the person without risk factors
= p:1-p
= 0.001:1-0.001 = 0.001: 0.999 = 1: 999

And another who is an intravenous drug user and whose prior odds for HIV are given by
Prior odds for the IV drug user
= p:1-p
= 0.1:1-0.1 = 0.1:0.9 = 1:9

The posterior odds for these individuals after the test provides a positive result can be determined as follows:
1:999 x 99 = 11:111 for the person with no risk factors
1:9 x 99 = 11:1 for the IV drug user

The posterior probability of having HIV infection for the individual without a risk factor would be
(11/111)/ [(11/111) +1]
= (11/111)/[122/111] = 11/122 = 0.09 or 9%

The posterior probability of having HIV infection for the individual with history of IV drug abuse would be
(11/1)/[(11/1) +1]
= 11/12 = 0.917 or 91.7% for the IV drug user

Even with a positive test result, the possibility of having HIV infection differs in these individuals. In the individual without a risk factor, the diagnosis of HIV infection still seems unlikely. But in the patient with IV drug abuse, HIV infection seems to be the most probable diagnosis. The test modifies the prior odds by a fixed multiplication factor in the case of both these individuals. But the prior odds affect the posterior odds and hence interpretation of the result is different in the two individuals. In contrast, the frequentist approach does not appreciate any difference between a person without risk factors and the intravenous drug user in the interpretation of the test result. The clinician invariably understands and appreciates this difference (although subjectively in terms of risk factors) and thus he uses the Bayesian approach in principle, though not in the same way as a statistician would.

Prior knowledge from the clinical scenario
Let us take another example. Consider a scenario wherein a patient presents with hemiparesis. The clinician makes a list of all possible causes of hemiparesis in his mind. He, then, enlists the prior probabilities or odds for the presence of each of these etiologies. Assigning of prior odds can be based on the prevalence of each cause. It can also be based on the clinical scenario.

For example, consider two patients presenting to the emergency department with hemiparesis. One of them is a 60-year-old male with history of hypertension while the other is a 35-year-old male

with history of heart disease and irregular heart beats. Clearly, the prior odds of an embolic event, which is one of the causes for hemiparesis, are higher for the second patient than for the first one with hypertension.

Adding new knowledge from history, physical examination and investigations

Continuing with the above example, the questions asked while taking the history (for example, whether the patient had loss of consciousness, whether the weakness was maximal at onset, whether the speech is affected, etc) are like 'diagnostic tests' for at least one of the causes for hemiparesis. They have their own sensitivity, specificity and likelihood ratios for determining etiology for hemiparesis. Similarly, each of the findings of clinical examination and investigations has its own sensitivity, specificity and likelihood ratios for the various causes for hemiparesis. Assuming that all these diagnostic tests are independent (for example, all the peripheral signs of aortic regurgitation, which are because of a wide pulse pressure and hence are interdependent, will be equivalent only to one diagnostic test), we can multiply the likelihood ratios obtained from multiple diagnostic tests for a particular cause to get a final likelihood ratio for a particular cause.

For example, if the positive and negative likelihood ratios for each diagnostic test for a particular cause are 3 and 1/3 respectively and we perform 10 such tests of which 8 turn out to be positive and 2 are negative, then the final likelihood ratio for that particular cause after our entire workup is $3^8/3^2 = 3^6 = 729$. Now if the prior probability of that particular cause was 10% ($P = 10/100 = 0.1$), then the posterior probability after our workup will be calculated as follows:

$$\text{Prior odds} = p{:}1{-}p = 0.1{:}1{-}0.1 = 1{:}9$$
$$\text{Posterior odds} = 1{:}9 \times 729 = 81{:}1$$
$$\text{Posterior probability} = (81/1)/[(81/1) +1] = 81/82 = 0.988 = 98.8\%$$

This posterior probability or posterior odds are to be evaluated considering similarly derived posterior probabilities or odds for the other causes.

Posterior probability or odds and differential diagnosis

After such an exercise, the clinician has a list of causes for hemiparesis with their posterior probabilities or odds, which in fact is the differential diagnosis with one or two diagnoses that are most probable. Although the clinician does not perform all these calculations while arriving at such a list, in principle he does use a method that parallels the above-mentioned complicated process.

Clinical acumen

In actual practice the likelihood ratios of most findings of clinical examination and medical history are unknown; so they have to be assigned by the clinician. Clinicians who have 'good clinical acumen' are the ones who are able to assign the appropriate prior probabilities or odds and the likelihood ratios consistently although they may refer to them as 'judgment of

clinical scenario' or 'clinical experience'. Clinical experience in fact helps the clinician to arrive at appropriate prior probabilities or odds and likelihood ratios in different scenarios consistently. Thus although it is amazing to know the thought process of a clinician, one needs vast clinical experience to use the process. We may know how to use the Bayesian approach but still we should be able to assign appropriate prior odds and likelihood ratios to become better clinicians. And in order to achieve this, we need to keep gaining clinical experience and reading medical literature. Nonetheless, knowledge of the thought process may help identify and appreciate the important goals and components of medical history and clinical examination of every patient. Thus although clinical experience is indispensable, knowledge of the Bayesian approach may help us use our experience optimally for a particular patient.

Limitations

Bayesian interpretation is not without its critics. It is stated that the subjectivity of the prior limits the utility of the Bayesian approach.[5] Another argument is that the Bayesian thought does not exactly reflect the clinicians' thought process because the clinicians have a certain concept of a threshold, thus if the posterior odds of two or more life-threatening diseases cross that threshold, clinicians would treat the patient for all those diseases irrespective of what diagnosis they think as more likely.[6]

Thus the Bayesian approach may not be a perfect simulation of the clinicians' thought process. But even then it is worth knowing this approach because it is probably the only approach that represents the clinicians' thought process and also can be explicitly described. Other ways like the artificial neural networks that closely match clinicians' decisions are known, but these approaches are probably known explicitly only to computers or sometimes not even to them!

Acknowledgments

References

1. Bland JM, Altman DG. Bayesians and frequentists. BMJ 1998;317:1151-60.
2. Gill CJ, Sabin L, Schmid CH. Why clinicians are natural Bayesians? BMJ 2005;330:1080-3.
3. Mark DB. Decision-making in clinical medicine. *In*: Kasper DL, Braunwald E, Fauci AS, Hauser SL, Longo DL, Jameson JL, editors. Harrison's Principles of Internal Medicine. 16th ed. McGraw-Hill: New York; 2005. p. 6-13.
4. Fagan TJ. Letter: Nomogram for Bayes theorem. N Eng J Med 1975;293:257.
5. Petrie A, Bulman JS, Osborn JF. Bayesian statistics. Br Dent J 2003;194:129-34.
6. Pauker SG, Kassirer JP. The threshold approach to clinical decision making. N Engl J Med 1980;302:1109-17.