Youden’s Index and the Weight of Evidence

G. Hughes

Keywords
Youden’s index, log-likelihood ratio, Bayes’ theorem, diagnostic test

Summary
By means of Shannon’s relationship between information and probability, Youden’s index for rating diagnostic tests is shown to be a probability-scale analogue of the log-likelihood ratio of a positive test outcome.

Correspondence to:
Dr. Gareth Hughes
SRUC
The King’s Buildings
West Mains Road
Edinburgh EH9 3JG
UK
E-mail: gareth.hughes@sruc.ac.uk

Methods Inf Med 2015; 54: 198–199
http://dx.doi.org/10.3414/ME14-04-0003
received: January 20, 2015
accepted: January 22, 2015
epub ahead of print: February 6, 2015

Dear Editor,

Consider a diagnostic test resulting in a binary classification of subjects. The proportions correctly classified among subjects actually positive (sensitivity) and subjects actually negative (specificity) are properties of the test in question. The quantity sensitivity + specificity − 1 is Youden’s index for rating diagnostic tests [1]. By 1950, when Youden’s index was first described [2], the use of Bayes’ theorem by Turing and colleagues had proved crucial in cryptographical work during World War II [3, 4] (incidentally initiating a renaissance for statistical applications of Bayesian methods), and Shannon’s groundbreaking work providing the basis for information theory had recently been published [5]. In order to describe how these analyses are linked, a little notation and terminology is first required. To describe a test, two groups of subjects (cases (D+)) and controls (D−)) are identified by means of a gold standard, independent of the test in question. Then, all subjects in both groups are tested and the outcomes (positive (T+) or negative (T−)) recorded (Youden’s analysis in its original form uses a binary classification of subjects and of test outcomes). Although diagnostic tests are imperfect, a large proportion of subjects who are cases should provide a positive test outcome (these results are true positives); while a small proportion of controls will also provide a positive test outcome (these results are false positives). The proportion of cases testing positive (sensitivity, true positive proportion, denoted TPP) is an estimate of the conditional probability Pr(T+|D+), and the proportion of controls testing positive (1−specificity, false positive proportion, denoted FPP) is an estimate of Pr(T+|D−). Then Youden’s index is J = TPP − FPP = Pr(T+|D+) − Pr(T+|D−).

Now consider Shannon’s analysis, and in particular the information content of a message (such as the outcome of a test). Briefly, if a positive test outcome were considered to be very unlikely before the result of the test became known, the information content of a positive outcome – when realized – would be large. And conversely, if a positive test outcome were considered likely, the realized information content of that outcome would be small. Shannon’s analysis defines the information content of a message h(p) as a function of the probability of that message (p): h(p) = −log(p) (the base of the logarithm can be chosen to suit the application). Then h(p) declines monotonically from ∞ to 0 as p increases from 0 to 1, so meeting the requirements as stated. In Figure 1, Shannon’s information graph is shown, with Pr(T+|D+) = TPP and Pr(T+|D−) = FPP marked on the probability axis (TPP > FPP is assumed, but otherwise the numerical values are not important here). The corresponding values on the information axis are −log(Pr(T+|D+)) and −log(Pr(T+|D−)), respectively. On the horizontal (probability) axis, the interval TPP − FPP = Pr(T+|D+) − Pr(T+|D−) = J. The corresponding interval on the vertical (information) axis is −log(Pr(T+|D+)) − (−log(Pr(T+|D+))) = log(Pr(T+|D+))/Pr(T+|D−)). This quantity is the log-likelihood ratio of a positive test outcome, denoted here log[LR(+)]. Now note that Bayes’ theorem may be stated as follows: the log-likelihood ratio is added to the initial log-odds to obtain the final log-odds. Turing and colleagues, who referred to the log-likelihood ratio as the weight of evidence, adopted this format; as did Van den Ende and colleagues in the

Figure 1
Shannon’s information graph, showing Youden’s index J as an interval on the horizontal (probability, p) axis and the log-likelihood ratio for a positive test outcome log[LR(+)] as the corresponding interval on the vertical (information, −log(p)) axis.
context of diagnostic clinical epidemiology [6, 7].

The correspondence between \( J \) and \( \log[LR+] \) is such that there is a monotonic increasing trend in \( \log[LR+] \) from \( J = 0 \ (\log[LR+] = 0) \) towards \( J = 1 \). This brings within the Bayesian paradigm the observation that \( \Pr(D+|T+) \) (the posterior probability that a subject is a case given a positive test outcome) > \( \Pr(D+) \) (the prior probability that a subject is a case) if and only if \( J > 0 \) (a “process qualification” requirement for a test used for diagnosis) [8]. Between the extremes, there is not a one-to-one correspondence between \( J \) and \( \log[LR+] \) because diagnostic tests with equal values of \( J \) will not necessarily have equal values of \( \log[LR+] \). Notwithstanding, the correspondence is likely to be of interest, for example in studies concerning the evaluation of diagnostic tests where both Youden's index and likelihood ratios are calculated [9, 10]. Previously, Zhou et al. noted informally that Youden's index “reflects the likelihood of a positive result among patients with versus without the condition” [1]. The present analysis now provides an analytical underpinning for that description. In words: Youden's index is a probability-scale analogue of the log-likelihood ratio of a positive test outcome.

Acknowledgment
SRUC receives grant-in-aid from the Scottish Government.

References