

Statistical Evaluation for Medical Screening & Diagnostic Tests

Events related to evaluation of a diagnostic test

Suppose we are interested in whether a subject is diseased or nondiseased, abbreviated D and $\bar{D}$, respectively. We consider a dichotomous diagnostic test with either positive (T) or negative ($\bar{T}$) results. We are interested in whether the test is predictive of disease status. In clinical application, such test can be a screening test, a biomedical test, or a status of presenting symptoms (T = present, $\bar{T}$ = absent). With two alternative disease states (D or $\bar{D}$) and two corresponding diagnostic alternatives (T or $\bar{T}$) there are four possible outcomes as summarized in Table 1.

Table 1 Classification of test results by disease status		
Test Result	Disease Status	
	Diseased (D)	Nondiseased ($\bar{D}$)
Positive (T)	True positive	False positive
Negative ($\bar{T}$)	False negative	True negative

In probability theory notations, the four outcomes are four joint events:

$$\begin{aligned} \text{True positive} &= T \cap D, & \text{False positive} &= T \cap \bar{D}, \\ \text{False negative} &= \bar{T} \cap D, & \text{True negative} &= \bar{T} \cap \bar{D}. \end{aligned}$$

But in clinical and biomedical application we are more interested in the conditional probabilities. The questions of interest would be

1. Given that a subject has the disease, what is the probability of a positive test result (or the presence of a symptom), *i.e.*, $P\{T | D\} = ?$
2. Given that a subject does not have the disease, what is the probability of a negative test result (or the absence of a symptom), *i.e.*, $P\{\bar{T} | \bar{D}\} = ?$
3. Given a positive screening test (or the presence of a symptom), what is the probability that the subject has the disease, *i.e.*, $P\{D | T\} = ?$
4. Given a negative screening test (or the absence of a symptom), what is the probability that the subject does not have the disease, *i.e.*, $P\{\bar{D} | \bar{T}\} = ?$

Sensitivity and specificity of a diagnostic test

The statistical measures for evaluating the diagnostic test based on questions 1 and 2 are summarized in Table 2

Table 2 Statistical measures for evaluating a diagnostic test (I)

Test Result	Disease Status	
	Diseased (D)	Nondiseased ($\bar{D}$)
Positive (T)	True positive fraction (TPF) $P\{T D\}$	False positive fraction (FPF) $P\{T \bar{D}\}$
Negative ($\bar{T}$)	False negative fraction (FNF) $P\{\bar{T} D\}$	True negative fraction (TNF) $P\{\bar{T} \bar{D}\}$
Sum	1	1

Since $TPF + FNF = 1$ and $FPF + TNF = 1$ only two fractions are needed to characterize the outcome. To evaluate a diagnostic test we can use

$$TPF = P\{T | D\} \text{ and } FPF = P\{T | \bar{D}\},$$

or use

$$\text{Sensitivity} = TPF = P\{T | D\},$$

$$\text{Specificity} = TNF = P\{\bar{T} | \bar{D}\}.$$

Estimations of Sensitivity and specificity of a diagnostic test

Suppose we have a sample of diseased subjects and a sample of nondiseased subjects. The test results for the two samples are recorded in Table 3

Table 3 Diseased samples classified according to test result

Test Result	Samples		
	Diseased (D)	Nondiseased ($\bar{D}$)	Total
Positive (T)	a	b	$a + b$
Negative ($\bar{T}$)	c	d	$c + d$
Total	$a + c$	$b + d$	n

$$\hat{TPF} = a/(a + c) \text{ (Sensitivity),}$$

$$\hat{FNF} = c/(a + c),$$

$$\hat{FPF} = b/(b + d),$$

$$\hat{TNF} = d/(b + d) \text{ (Specificity).}$$

Example 1 (EXAMPLE 3.5.1 on page 81, 10th Ed)

Table 4 Sample Classifications and Estimation for evaluation measures

Test Result	Samples (Diseased or not)			Estimations	
	Yes (D)	No ($\bar{D}$)	Total	diseased (D)	nondiseased ($\bar{D}$)
Positive (T)	436	5	441	$\hat{TPF} = 436/450$	$\hat{FPF} = 5/495$
Negative ($\bar{T}$)	14	495	509	$\hat{FNF} = 14/450$	$\hat{TNF} = 495/500$
Total	450	500	950	1	1

Sensitivity = $436/450 \approx 0.9689$, Specificity = $495/500 \approx 0.99$.

Predictive values of a test

The statistical measures for evaluating the diagnostic test based on questions 3 and 4 are summarized in Table 5

Table 5 Statistical measures for evaluating a diagnostic test (II)			
	Disease Status		Sum
	Diseased (D)	Nondiseased ($\bar{D}$)	
Positive (T)	Positive predictive value = $P\{D T\}$	$P\{\bar{D} T\}$	1
Negative ($\bar{T}$)	$P\{D \bar{T}\}$	Negative predictive value = $P\{\bar{D} \bar{T}\}$	1

The positive predictive value and negative predictive value are called predictive value positive and predictive value negative in the text book. They are referred to as the predictive values of a positive result or a negative result. In the next part we will see that given the prevalence of the disease, $P\{D\}$, these two measure can be calculated from sensitivity and specificity.

Bayes' theorem

$$P\{D|T\} = \frac{P\{T|D\}P\{D\}}{P\{T|D\}P\{D\} + P\{T|\bar{D}\}P\{\bar{D}\}} \quad (1)$$

$$P\{\bar{D}|\bar{T}\} = \frac{P\{\bar{T}|\bar{D}\}P\{\bar{D}\}}{P\{\bar{T}|\bar{D}\}P\{\bar{D}\} + P\{\bar{T}|D\}P\{D\}} \quad (2)$$

Example 1 (continued) Assuming that the prevalence of Alzheimer's disease of the population (people with age being 65 years or older) is 0.113. Find the positive predictive value of test.

Sensitivity and specificity are $P\{T | D\} = 0.9689$ and $P\{\bar{T} | \bar{D}\} = 0.99$. The prevalence is $P\{D\} = 0.113$. Hence the positive predictive value is

$$P\{D | T\} = \frac{0.9686 \times 0.113}{0.9686 \times 0.113 + (1 - 0.99) \times (1 - 0.113)} = 0.925.$$

Diagram illustrating the components of the Positive Predictive Value (PPV) formula:

- Sensitivity points to 0.9686
- Prevalence points to 0.113
- $1 - \text{Prevalence}$ points to $(1 - 0.113)$
- $\text{FPF} = 1 - \text{specificity}$ points to $(1 - 0.99)$

Note that before the test is taken an individual who is randomly selected from the population has an 11.3% chance of having Alzheimer's disease. This is called *prior probability*. After the test is taken and the result is positive, the same individual has a 92.5% chance of having the disease. This is called *posterior probability*. The posterior probability takes into account a new piece of information—the positive test result.

We can also calculate the negative predictive value of the test as follows

$$P\{\bar{D} | \bar{T}\} = \frac{0.99 \times (1 - 0.113)}{0.99 \times (1 - 0.113) + (1 - 0.9689) \times 0.113} = 0.996$$

Diagram illustrating the components of the Negative Predictive Value (NPV) formula:

- Specificity points to 0.99
- $1 - \text{Prevalence}$ points to $(1 - 0.113)$
- Prevalence points to 0.113
- $\text{FNF} = 1 - \text{sensitivity}$ points to $(1 - 0.9689)$

Example 2 Alpha-fetoprotein Blood (AFP) Test is widely used in diagnosis of liver cancer. A study indicates that the sensitivity and specificity of AFP are 0.95 and 0.90. It was reported that the prevalence of liver cancer in a population is 0.0004. What are the predictive values of the test?

Solution

$P\{T | D\} = 0.95$, $P\{\bar{T} | \bar{D}\} = 0.90$ and $P\{D\} = 0.0004$. Hence the positive predictive value is

$$P\{D | T\} = \frac{0.95 \times 0.0004}{0.95 \times 0.0004 + (1 - 0.90) \times (1 - 0.0004)} = 0.00379,$$

and negative predictive value is

$$P\{\bar{D} | \bar{T}\} = \frac{0.90 \times (1 - 0.0004)}{0.90 \times (1 - 0.0004) + (1 - 0.95) \times 0.0004} = 0.99998.$$

Note that the predictive values of a test depend upon the prevalence. In evaluating the performance we prefer to use sensitivity and specificity rather than predictive values.

General form of Bayes' theorem

Suppose that $A_1, A_2, \dots, A_n$ are n mutually exclusive and exhaustive events, i.e.

$$P\{A_1 \cup A_2 \cup \dots \cup A_n\} = P\{A_1\} + P\{A_2\} + \dots + P\{A_n\} = 1.$$

Then

$$P\{A_i | B\} = \frac{P\{A_i\}P\{B | A_i\}}{P\{A_1\}P\{B | A_1\} + P\{A_2\}P\{B | A_2\} + \dots + P\{A_n\}P\{B | A_n\}}, \quad (i = 1, 2, \dots, n)$$

Homework (due Aug 3)

Exercises 3.5.1 and 3.5.2 on page 83 (10th Ed)

Review exercises 9 (page 88) and 22 (page 90)