Binary classification

Kristian Wichmann

January 21, 2017

1 Basic definitions

Binary classification is a situation where there's two distinct outcomes which we are trying to predict. Associated predictors are known as binary predictors.

1.1 Positives and negatives

The two outcomes are denoted *positives* and *negatives*, respectively. Given a data point, and the corresponding prediction of a binary predictor, there's four different possibilities:

- True positive: Both the actual outcome and the prediction is positive.
- False positive: The prediction is positive, but the actual outcome is negative. This is also known as a type I error.
- False negative: The prediction is negative, but the actual outcome is positive. This is also known as a type II error.
- True negative: Both the actual outcome and the prediction is negative.

1.2 The confusion matrix

The frequency of the four above-mentioned events are usually presented in matrix form, in what is known as the *confusion matrix*. It is shown in table 1.2, where TP means 'True Positive' and so on.

The entries can either be specified absolutely or relatively.

Actual/Predicted	Positive	Negative
Positive	TP	FN
Negative	FP	TN

Table 1: The confusion matrix.

2 Sensitivity, specificity and prevalence

A common way of describing a predictor is through *sensitivity* and *specificity*:

• Sensitivity - also known as *recall* or *true positive rate* (TPR) - is the rate of actual positives that are classified as such. It can be expressed as:

$$TPR = \frac{TP}{TP + FN} \tag{2.1}$$

• Specificity - also known as *true negative rate* (TNR) - is the rate of actual negatives that are classified as such. It can be expressed as:

$$TNR = \frac{TN}{TN + FP}$$
 (2.2)

One might think, that if a predictor has a high specificity and sensitivity, then it is useful. It turns out that this is not automatically true: prevalence, i.e. the overall occurrence rate of positives also plays a role. The prevalence is TP + FN.

2.1 Example

Assume that a disease has a prevalence of 0.1% in a given population. A screening procedure with a specificity of 99% and a sensitivity of 95%. At first glance, this looks like a good predictor. But in practice, it is less impressive:

Since only 0.1% = 0.001 of the population actually has the disease (positive), 99.9% has not (negative). Of that 0.1%, 99% are true positives:

$$TP = 0.001 \cdot 0.99 = 0.00099 \tag{2.3}$$

The final 1% of the actual positives are false negatives:

$$TP = 0.001 \cdot 0.01 = 0.00001 \tag{2.4}$$

Out of the actual positives, 95% are true negatives:

$$TN = 0.999 \cdot 0.95 = 0.94905 \tag{2.5}$$

Actual/Predicted	Positive	Negative
Positive	0.00099	0.00001
Negative	0.04995	0.94905

Table 2: Example confusion matrix.

The 5% of the actual negatives are false positives:

$$FP = 0.999 \cdot 0.05 = 0.04995 \tag{2.6}$$

The corresponding confusion matrix is shown in table 2.1.

3 Positive and negative predictive value

In practice, the following statistics are often of great importance:

• The *positive predictive value* (PPV) is the conditional probability that a sample is actually positive given that the prediction is positive. It can be expressed:

$$PPV = \frac{TP}{TP + FP} \tag{3.1}$$

• The negative predictive value (NPV) is the conditional probability that a sample is actually negative given that the prediction is negative. It can be expressed:

$$NPV = \frac{TN}{TN + FN}$$
 (3.2)

3.1 Example

In the example above, the positive predictive value is:

$$PPV = \frac{0.00099}{0.00099 + 0.04995} \approx 0.019 \tag{3.3}$$

The negative predictive value is:

$$NPV = \frac{0.94905}{0.94905 + 0.00001} \approx 1.000 \tag{3.4}$$

So, the screening is very good at actually predicting negatives: If the test is negative, you're almost certain not to have the disease. However, if the test is positive, there's only a 1.9% chance that you actually have the disease! This is not very reassuring! The deeper reason for this low number, is that since the prevalence is low, there ends up being a comparatively large number of false positives, even if the specificity is high.

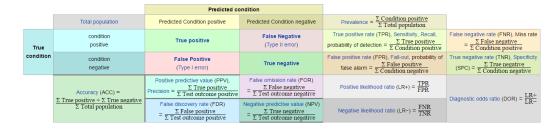


Figure 1: Confusion matrix and derived statistics. Source: Wikipedia.