Equation 1 shows the formula for the discriminant score for class k, δ_k. x_i^* represents the expression level of gene i in the array to be classified. x^* represents a vector of expression levels for all p genes to be used for classification drawn from the array to be classified, i.e. $x^* = (x_1^*, x_2^*, x_3^*, \ldots, x_p^*)$. $\bar{x}'_k$ is the shrunken centroid calculated from the training data and shrinkage factor, and $\bar{x}'_{ik}$ is the component of $\bar{x}'_k$ corresponding to gene i. $s_i$is the pooled within-class standard deviation for gene i in the training data, and is also calculated during the training of the classifier. s_0 is a specified positive constant. It is typically set to the median of s_i across all genes on the array. Lastly, π_k represents the prior probability of a sample belonging to class k. It should be set to the frequency of class k in the population being smapled from, or if such information is not available, π_k for all classes can be set identically to $1/K$, where K is the number of possible classes.

$$\delta_k(x^*) = \sum_{i=1}^{p} \frac{(x_i^* - \bar{x}'_{ik})^2}{(s_i + s_0)^2} - 2 \log \pi_k \tag{1}$$

In summary, $\bar{x}'_{ik}$, s_i, and s_0 are calculated from the training data, and the π_k are specified *a priori* (though s_0 may also be specified *a priori*). Together, these parameters comprise the trained classifier. The x_i^* represent the sample that is to be classified by this trained classifier.

Based on the discriminant scores, class probabilities can be calculated. Equation 2 shows the formula for the probability that the array x^* belongs to class k. Note that the denominator is simply the sum of the numerator across all possible classes.

$$\hat{p}_k(x^*) = \frac{e^{-\frac{1}{2}\delta_k(x^*)}}{\sum_{l=1}^{K} e^{-\frac{1}{2}\delta_l(x^*)}} \tag{2}$$

Equation 3 expresses the classification rule. Simply, the array is assigned to the class that it has the highest probability of belonging to.

$$C(x^*) = \underset{k \in \{1,K\}}{\arg\max}\, \hat{p}_k(x^*) \tag{3}$$

Depending on the application, the class prediction $C(x^*)$ or the class probabilities $\hat{p}_k(x^*)$ may be used as the final output.