

Choose k , a positive integer which is large but small compared to the sample sizes. Specify a metric in the sample space, for example ordinary Euclidean distance. Pool the two samples and find, of the k values in the pooled samples which are nearest to z , the number M which are X 's. Let $N = k - M$ be the number which are Y 's. Proceed with the likelihood ratio discrimination, using however $\frac{M}{k}$ in place of $f(z)$ and $\frac{N}{k}$ in place of $g(z)$. That is, assign Z to F if and only if

$$\frac{M}{k} < c \frac{N}{k}.$$