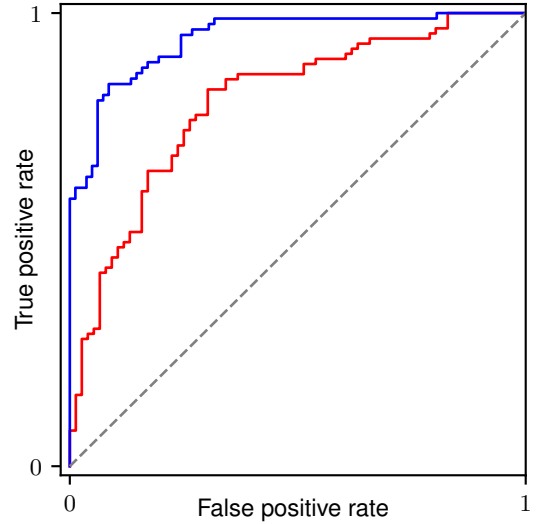


Figure 5.11 The receiver operator characteristic (ROC) curve is a plot of true positive rate against false positive rate, and it characterizes the trade-off between type 1 and type 2 errors in a classification problem. The upper blue curve represents a better classifier than the lower red curve. Here the dashed curve represents the performance of a simple random classifier.



recall, and is therefore defined by

$$F = \frac{2 \times \text{precision} \times \text{recall}}{\text{precision} + \text{recall}} \quad (5.38)$$

$$= \frac{2N_{\text{TP}}}{2N_{\text{TP}} + N_{\text{FP}} + N_{\text{FN}}}. \quad (5.39)$$

Of course, we can also combine the confusion matrix in [Figure 5.9](#) with the loss matrix in [Figure 5.6](#) to compute the expected loss by multiplying the elements pointwise and summing the resulting products.

Although the ROC curve can be extended to more than two classes, it rapidly becomes cumbersome as the number of classes increases.

5.3. Generative Classifiers

Section 5.2.4

We turn next to a probabilistic view of classification and show how models with linear decision boundaries arise from simple assumptions about the distribution of the data. We have already discussed the distinction between the discriminative and the generative approaches to classification. Here we will adopt a generative approach in which we model the class-conditional densities $p(\mathbf{x}|\mathcal{C}_k)$ as well as the class priors $p(\mathcal{C}_k)$ and then use these to compute posterior probabilities $p(\mathcal{C}_k|\mathbf{x})$ through Bayes' theorem.

First, consider problems having two classes. The posterior probability for class