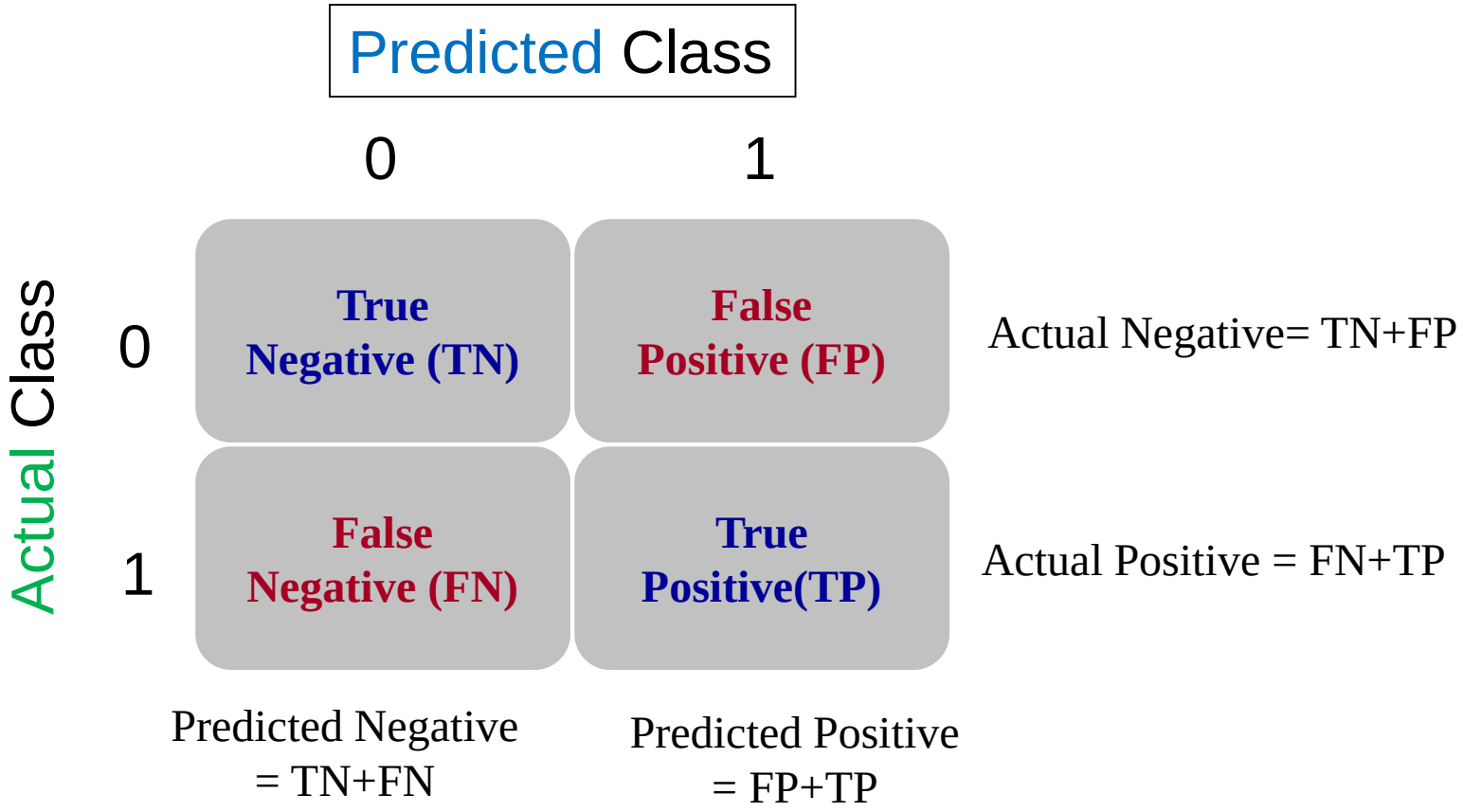
Overview of Statistical Metrics for Model Assessment

Dr. Goutam Chakraborty

Outline

- Metrics used by statisticians/data scientists
  - Misclassification Rate, Hit-ratio, Sensitivity, Specificity, Recall, Precision, F1 Score..
  - ROC curve and area under this curve (AUC), Gains curve, Lift curve (Lorenz curve, Concentration curve), K-S statistic …

Confusion Matrix



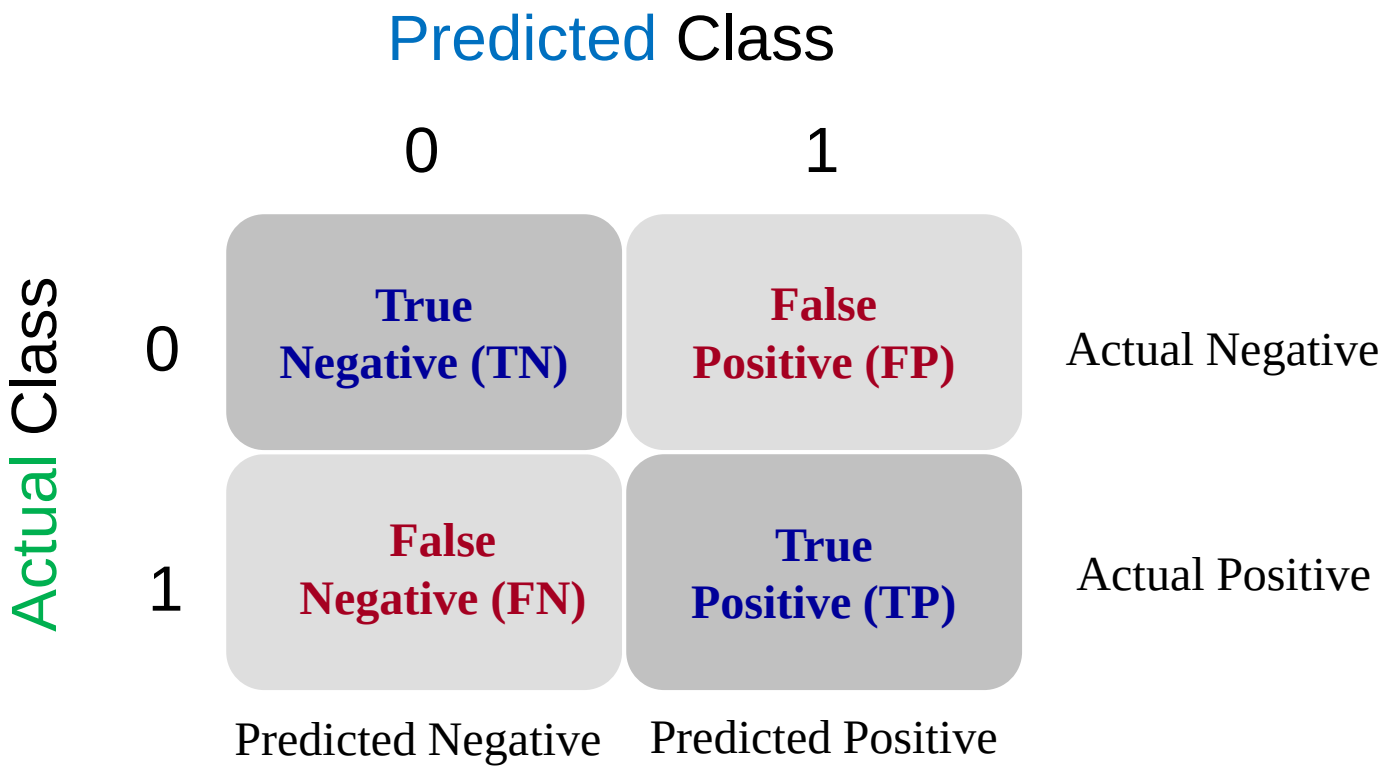
Misclassification Rate = (FP+FN)/(FP+FN+TN+TP)
Hit-ratio (overall accuracy) = 1- Misclassification Rate

An Example with Very Unbalanced Classes

		Predicted/Classified	
		Negative	Positive
Actual	Negative	998	0
	Positive	1	1

Accuracy = 99.9%+, but what if the positive here represents a fraud case that may cost
a bank millions of $ or, represents a terrorist or, represents a person carrying deadly virus that can spread quickly

Sensitivity, Specificity, PV+ and PV-



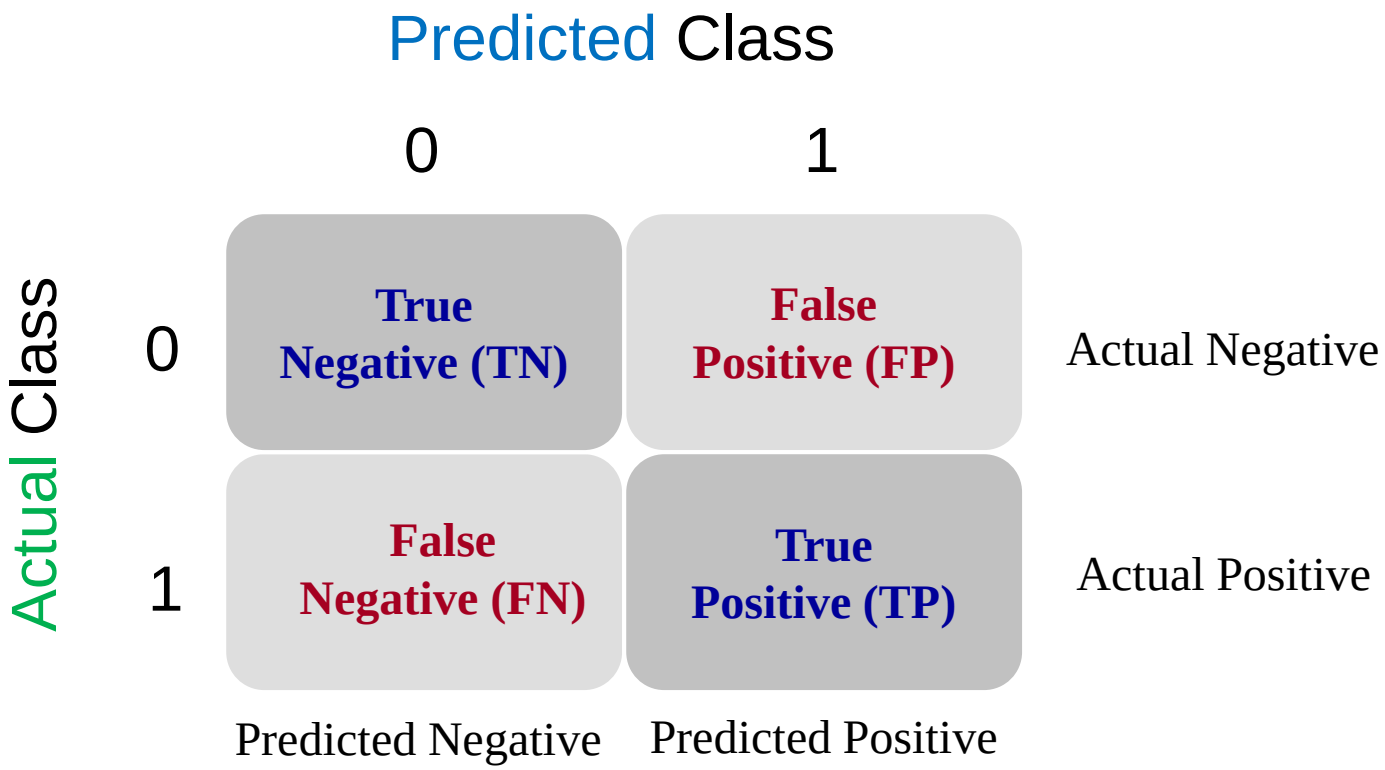
Sensitivity = TP/Actual Positive = TP/(FN+TP)

Specificity = TN/Actual Negative = TN/(FP+TN)

Positive Predicted Value (PV+) = TP/Predicted Positive = TP/(FP+TP)

Negative Predicted Value (PV-) = TN/Predicted Negative = TN/(FN+TN)

Recall and Precision



Sensitivity = TP/Actual Positive = TP/(FN+TP) = **Recall**

Positive Predicted Value (PV+) = TP/Predicted Positive = TP/(FP+TP) = **Precision**

F1 Score : Balancing Between Precision and Recall

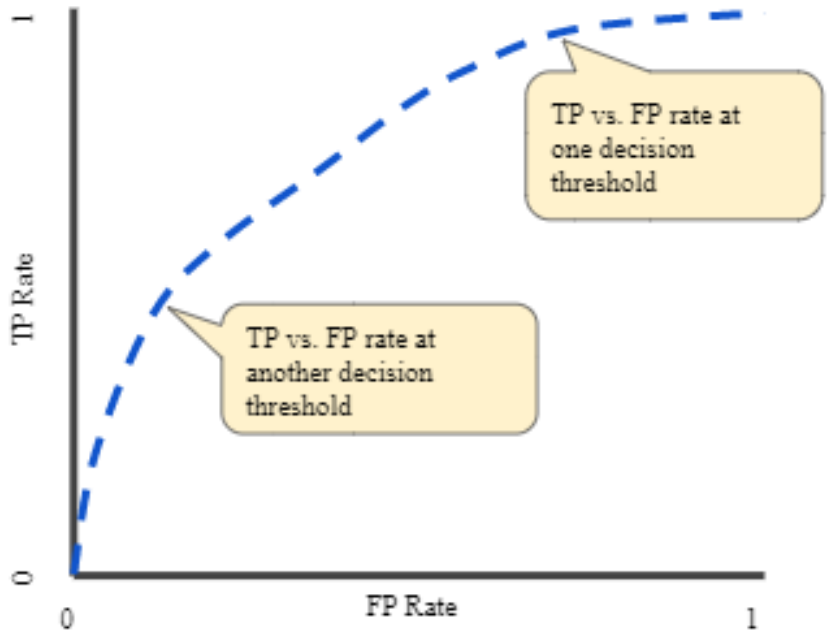
F1 Score is the **harmonic mean** of precision and recall

$$F1 = 2 \times \frac{Precision * Recall}{Precision + Recall}$$

How to Select Which Metric?

- **First Question:** Does both True Positive and True Negatives matters to the business or just True Positives? If both is important, *Accuracy* is may be the you go.
- **Second Question:** If True Positive is what you *are concerned with more*, ask yourself, which one has a higher costs to business, False Positives or False Negatives?
  - If having large number of False Negatives has a higher cost to business, choose Recall.
  - If having large number of False Positives has a higher cost to business, choose Precision.
  - If you cannot decide or think that its best to reduce both, False Negatives and False Positives then choose F1.

ROC Curve

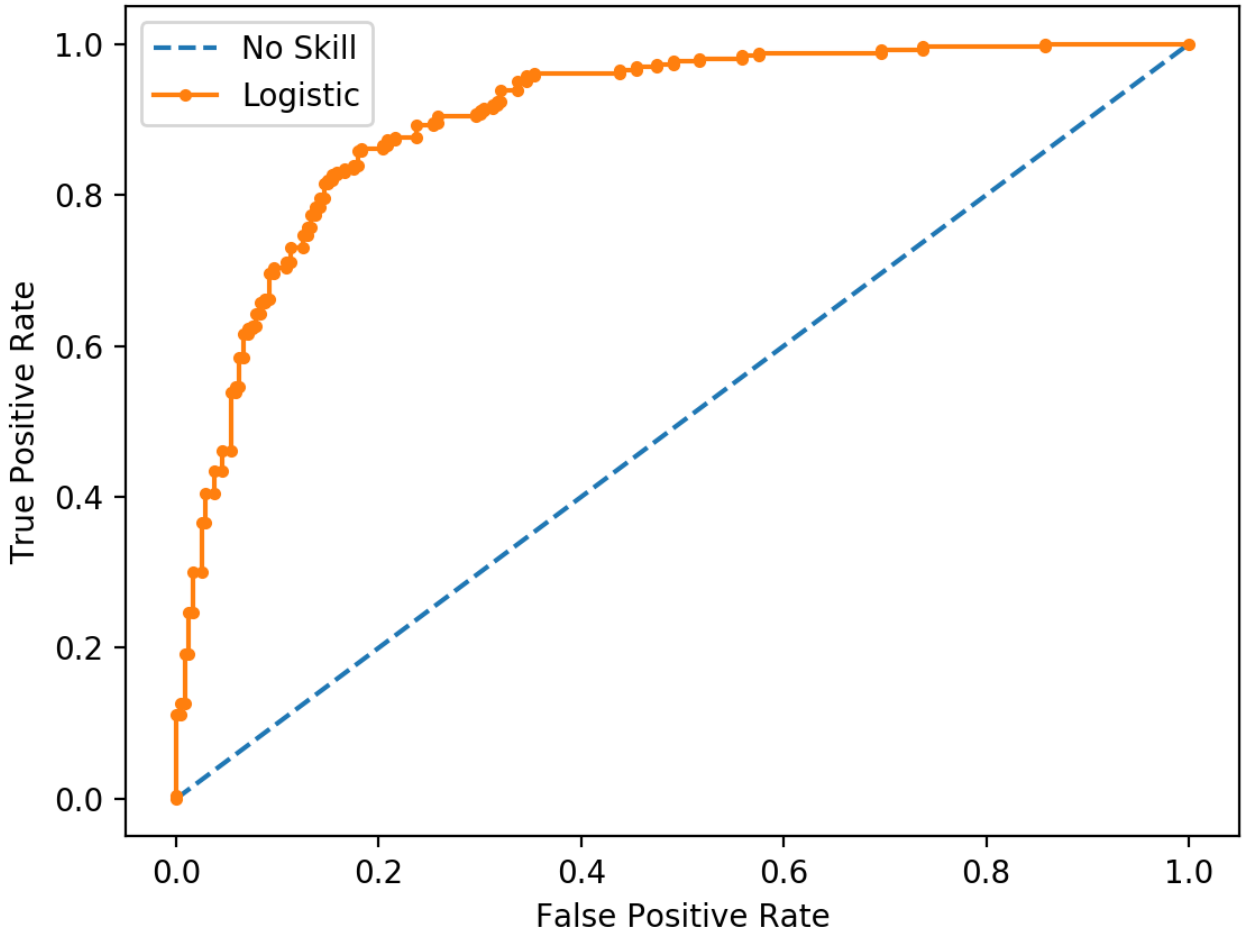


AUC: Area Under the ROC Curve

AUC stands for "Area under the ROC Curve." That is, AUC measures the entire two-dimensional area underneath the entire ROC curve (think integral calculus) from (0,0) to (1,1).

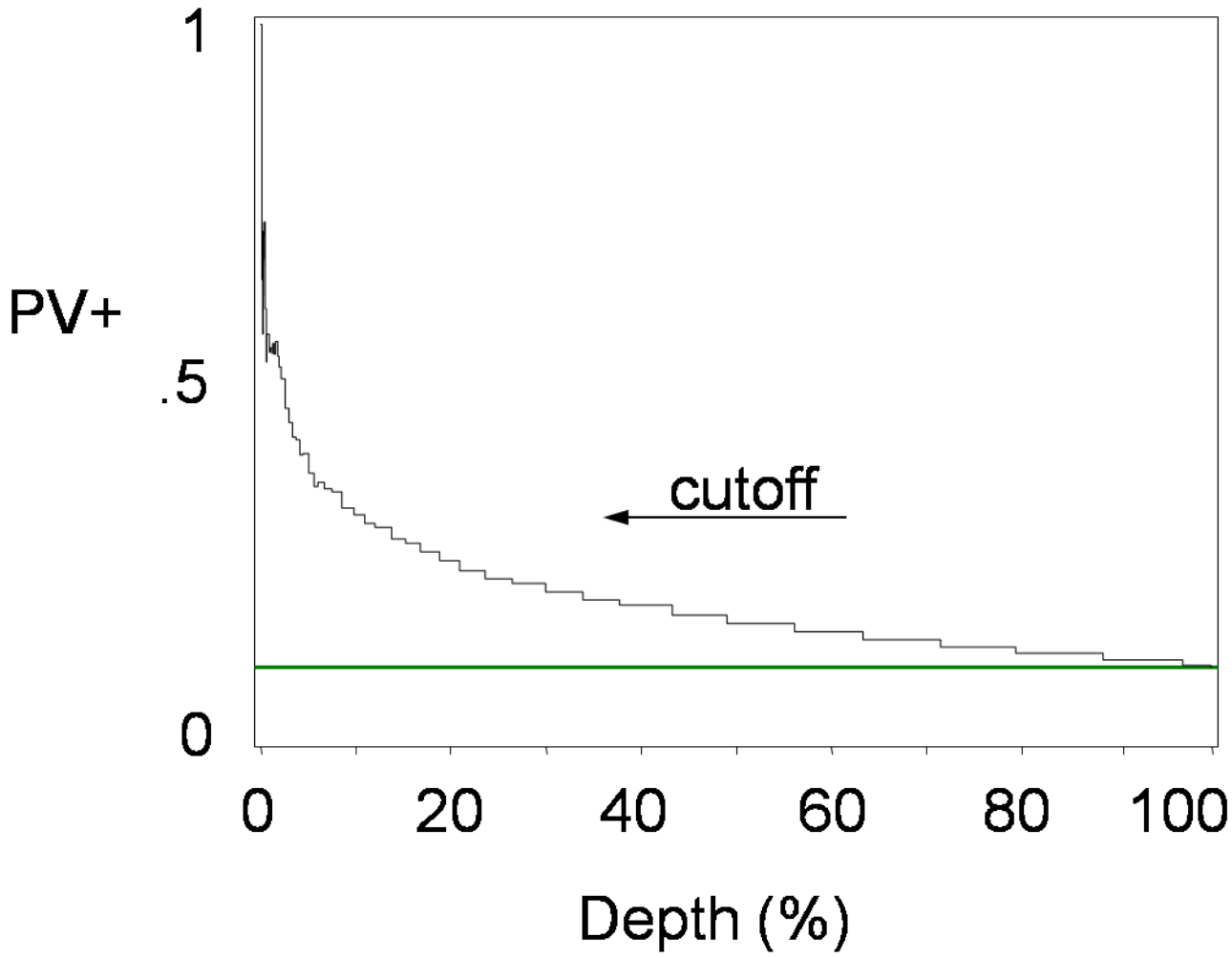
An ROC curve plots TPR vs. FPR at different classification thresholds. Lowering the classification threshold classifies more items as positive, thus increasing both False Positives and True Positives.

Example ROC Curve

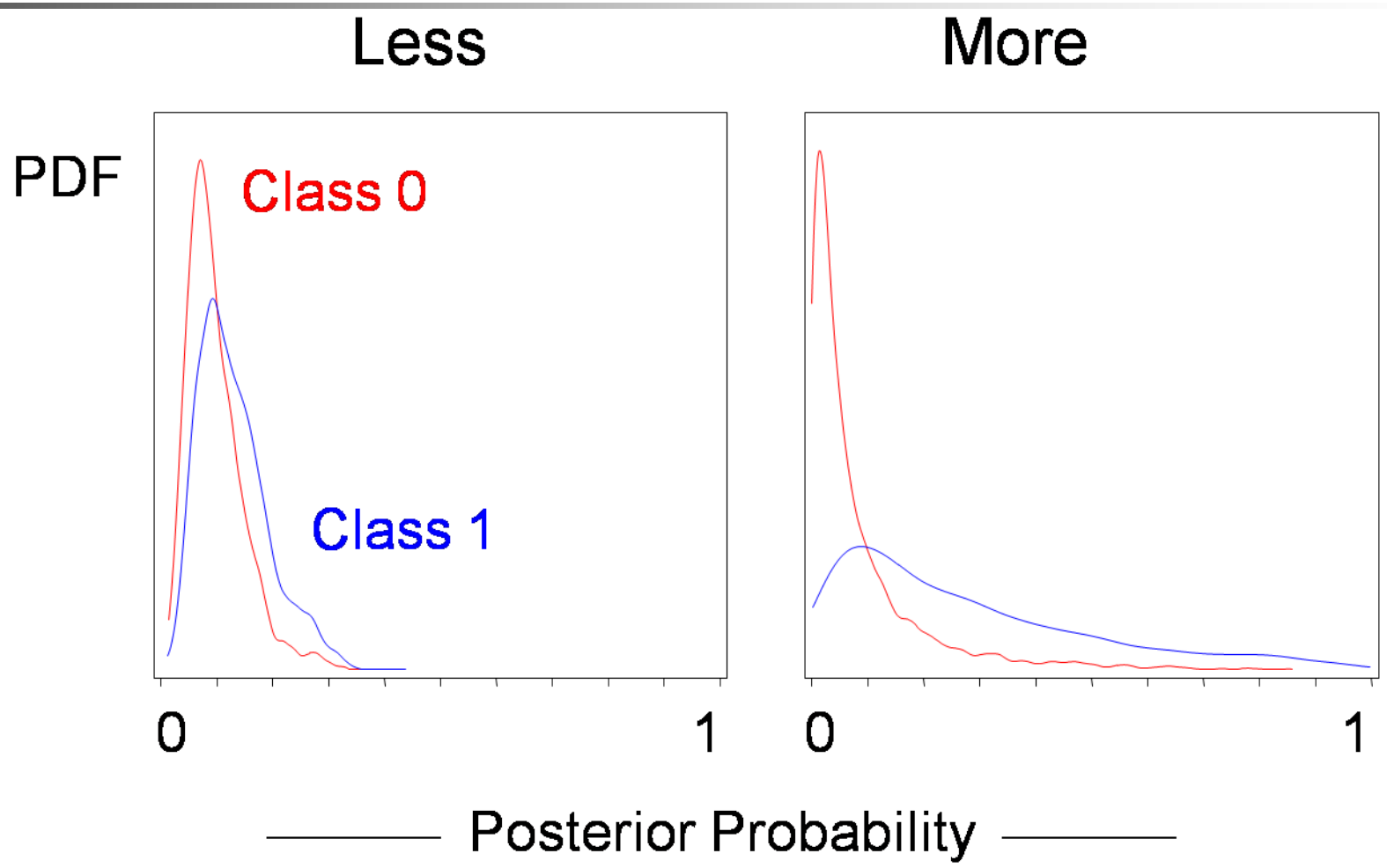


No Skill (baseline): ROC AUC=0.500
Logistic: ROC AUC=0.903

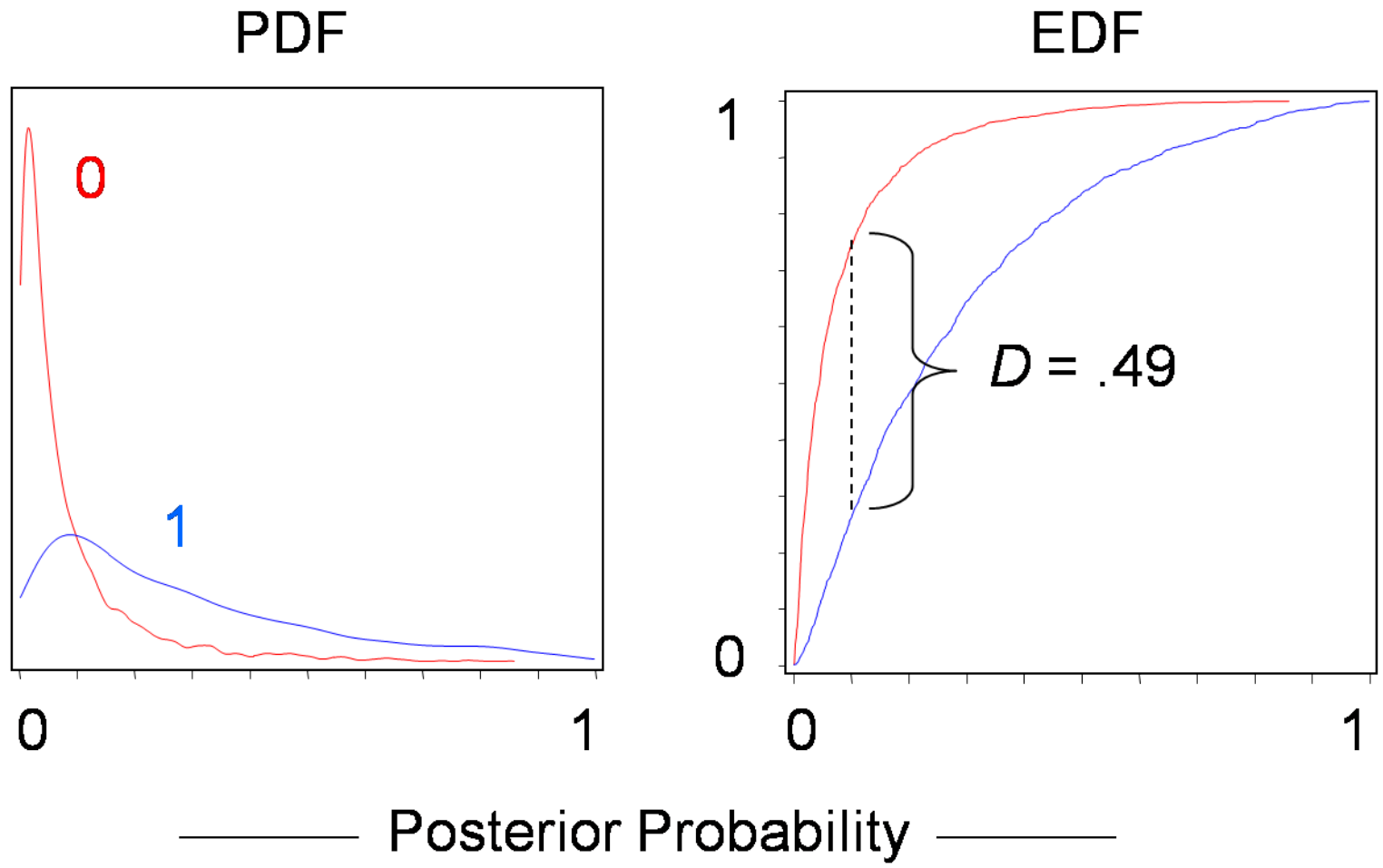
Gains Chart



Class Separation (K-S Statistic)



K-S Statistic



Area under the ROC Curve