Binary predictions

Types of research questions

So far, we have learned how to answer the following questions:

- What is the relationship between the explanatory variable(s) and the response?
- What is a "reasonable range" for a parameter in this relationship?
- Do we have strong evidence for a relationship between these variables?

> What other kinds of research questions might we ask?

Making predictions with the Titanic data

- For each passenger, we calculate $\hat{p}_i$ (estimated probability of survival)
- But, we want to predict *which* passengers actually survive

> How do we turn $\hat{p}_i$ into a binary prediction of survival / no survival?

Confusion matrix

		Actual	
		$Y = 0$	$Y = 1$
Predicted	$\widehat{Y} = 0$	344	70
	$\widehat{Y} = 1$	80	220

Did we do a good job predicting survival?

Why a threshold of 0.5?

Changing the threshold

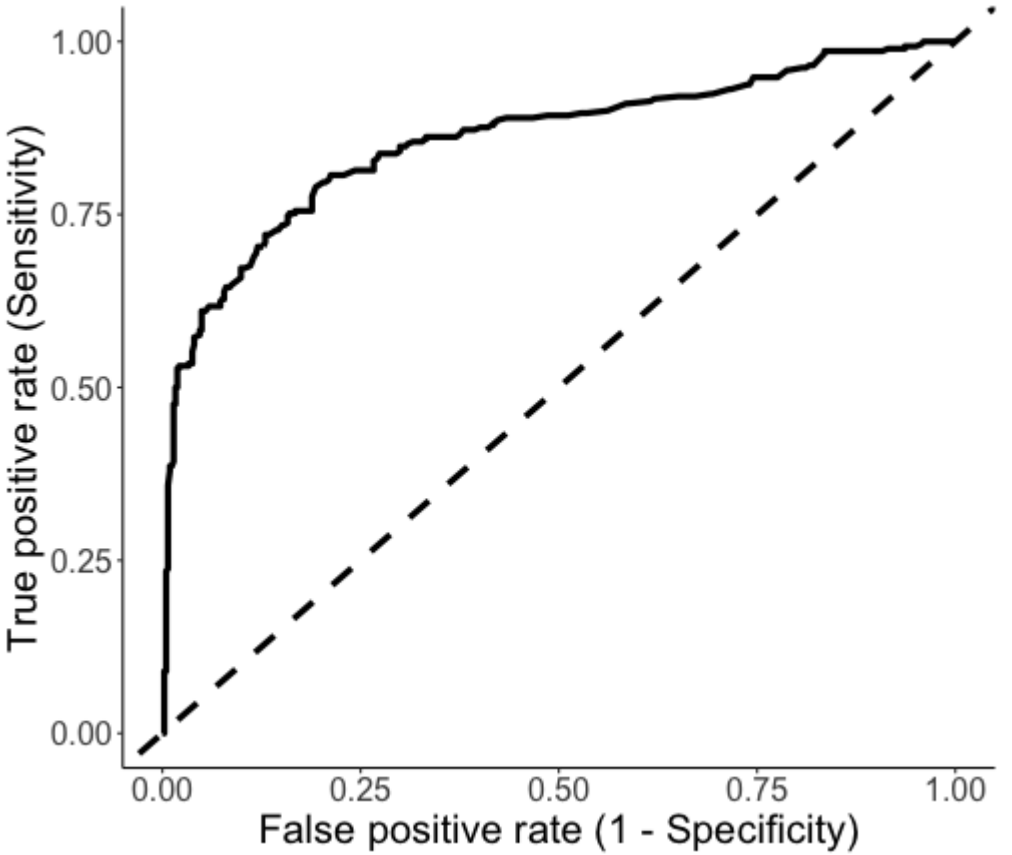
Using a threshold of 0.7:

		Actual	
		$Y = 0$	$Y = 1$
Predicted	$\widehat{Y} = 0$	412	136
	$\widehat{Y} = 1$	12	154

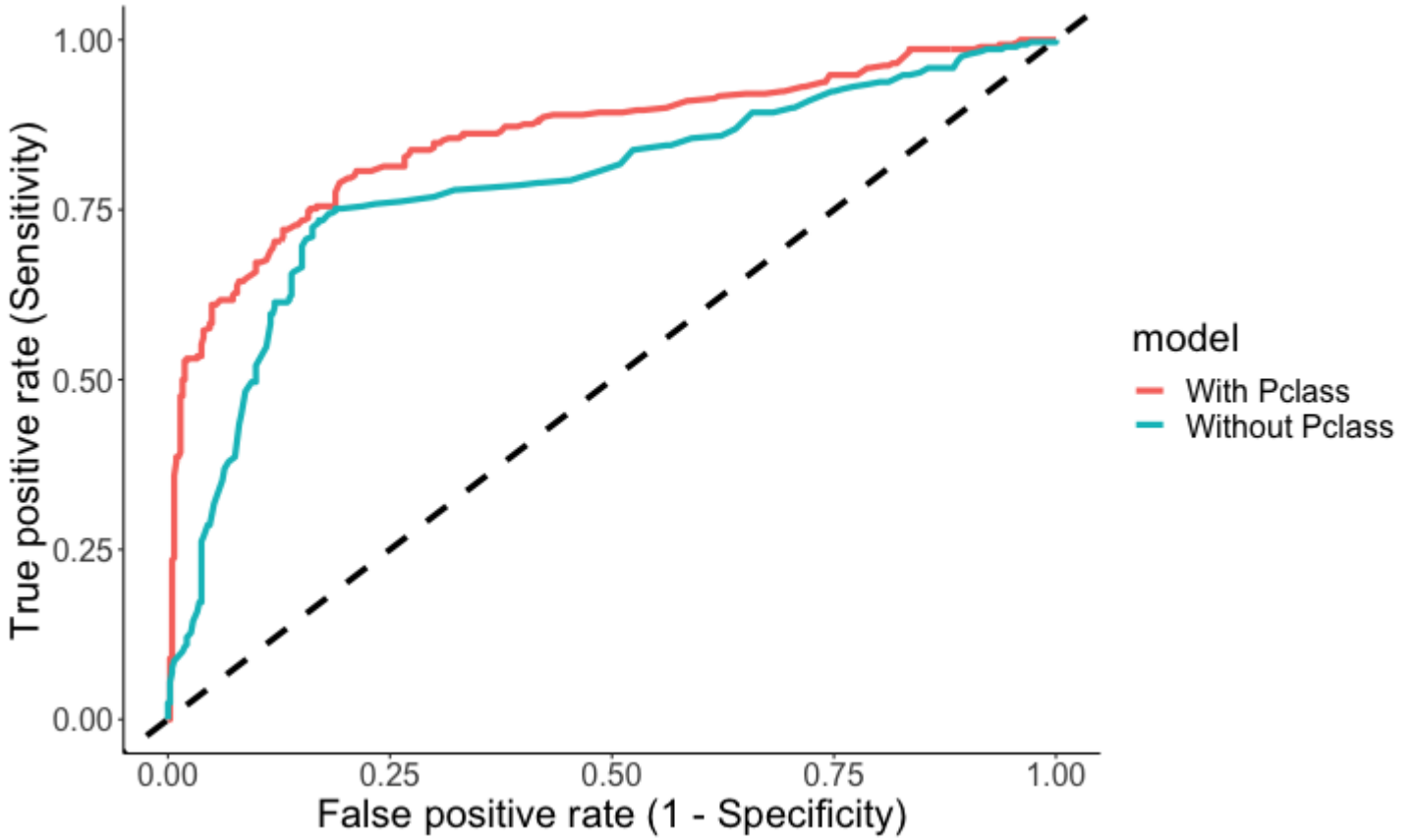
Using a threshold of 0.3:

		Actual	
		$Y = 0$	$Y = 1$
Predicted	$\widehat{Y} = 0$	309	49
	$\widehat{Y} = 1$	115	241

ROC curve: consider all thresholds



Comparing models with ROC curves



Problem: reusing data...

> It is generally a bad idea to assess performance of a model on the same data we used to train it. This can lead to overfitting.
>
> What can we do instead?