Predictive Modeling

Lecture II

Michael C Sachs

Measures of accuracy

Binary predictor, binary outcome

	Outcome (Y)	
Predictor (X)	0	1
———	–	–
0	True negative	False negative
1	False positive	True positive

- *Sensitivity* = Predict positive / outcome positive
- *Specificity* = Predict negative / outcome negative
- *Positive predictive value* = Outcome positive / predict positive
- *Negative predictive value* = Outcome negative / predict negative

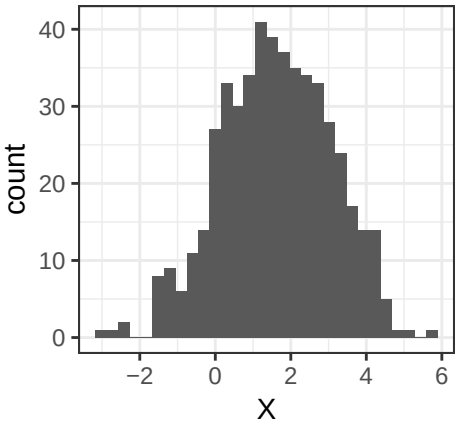
Covered in Biostat 1, slide 30

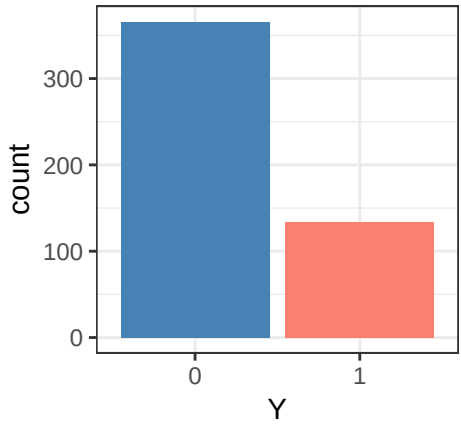
Cautionary notes

1. Base rate fallacy, aka, why one number is not enough:
   - the TPF of my great new test for Zika is 99%
   - A randomly select person tests positive. What is the probability that they have Zika?
   - Need to know the FPF and prevalence
2. The uninformative diagnostic test:
   - In the absence of any information, say every test is positive.
   - Has 100% sensitivity!
   - Has 0% specificity
   - Remember test positives have costs, too

Continuous predictor, binary outcome

We still have a binary outcome $Y \in \{0, 1\}$, now our test X is continuous, and can take any value.

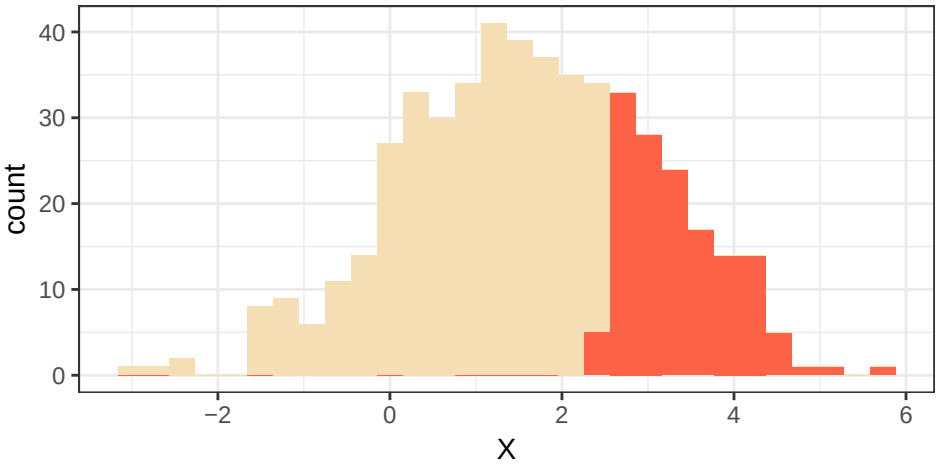




what to do?

Dichotomize the predictor

Choose a cutoff c, and say that we call the event $X > c$ a "test positive", otherwise it is a "test negative".



Now we can define

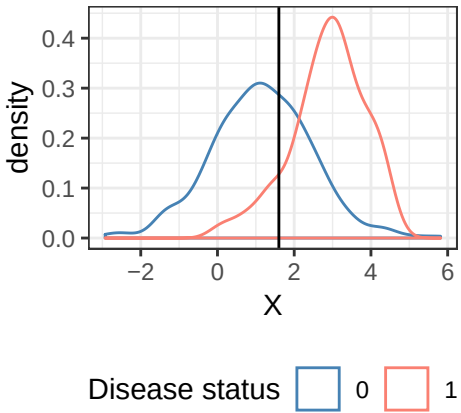
$$TPF(c) = P(X > c|Y = 1) \text{ and } FPF(c) = P(X > c|Y = 0)$$

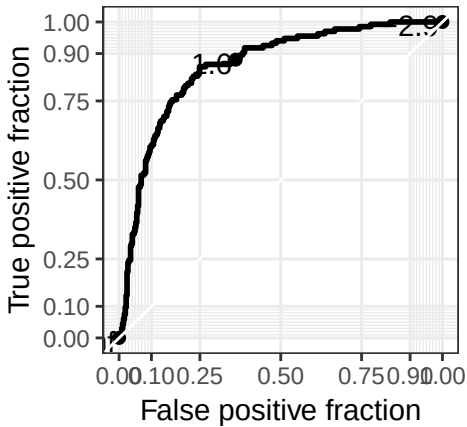
and

$$PPV(c) = P(Y = 1|X > c) \text{ and } NPV(c) = P(Y = 0|X \leq c).$$

What about the cutoff?

The cutoff c is not fixed a priori, so we make plots:



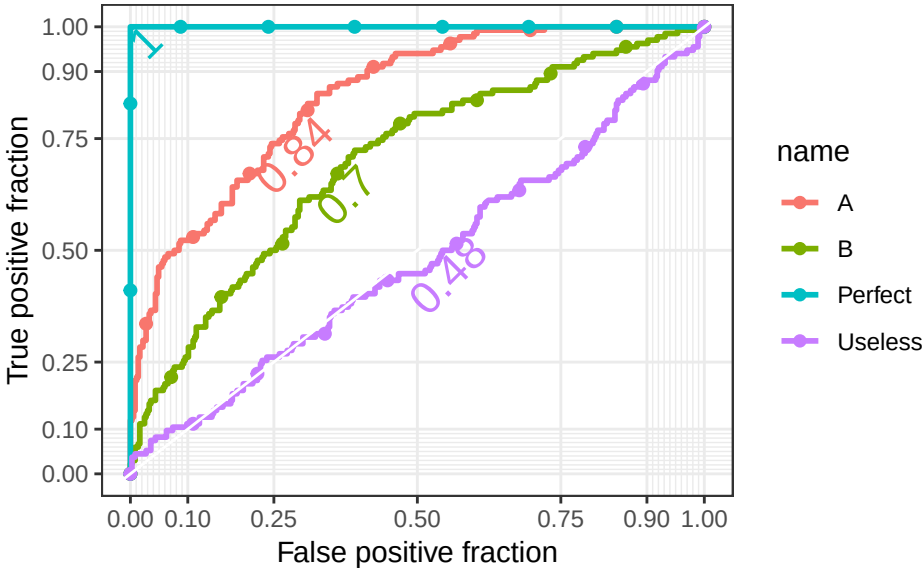


This is the ROC curve

Receiver **O**perating **C**haracteristic curve.

- Plot of $FPF(c)$ versus $TPF(c)$ for all possible c
- Perfect test: $FPF = 0$, $TPF(c) = 1$ for all c
- Useless test: $FPF(c) = TPF(c)$ for all c

The area under the ROC curve (AUC)



Relationship to c-index

Concordance index:

Estimates the concordance between predicted and true responses

- Observations with $Y = 1$ should have higher predicted values than those with $Y = 0$
- $P(X_1 > X_0)$, where X_1 is a randomly selected observation with $Y = 1$ and X_0 is a randomly selected observation with $Y = 0$
- This is equal to the AUC for binary Y.
- To estimate, generate all possible pairs of discordant on Y and count how many pairs have the correct ordering on X
- For survival data, there are some extensions of the c-index

Prospective measures

Again, we can define $PPV(c) = P(Y = 1|X > c)$, but also consider $P(Y = 1|X = c)$:

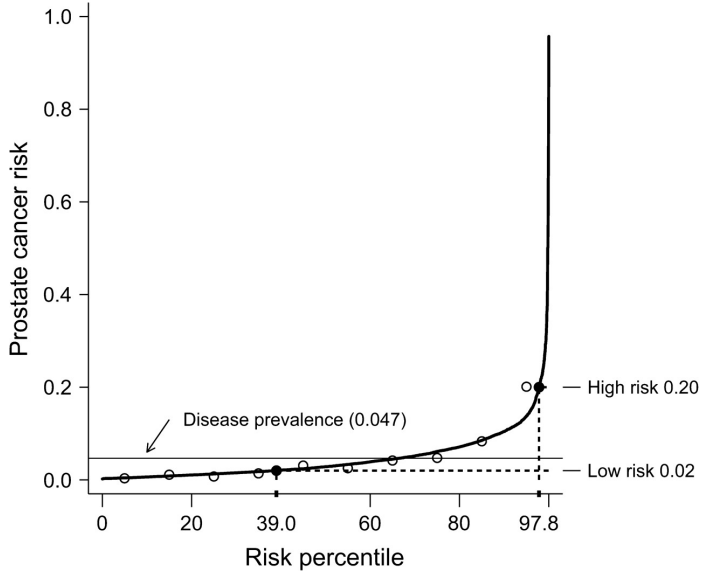
- These measures are more relevant for risk prediction
- But, they are harder to estimate

Logistic regression models

$$\text{logit}\{P(Y = 1|X = c)\} = \beta_0 + \beta_1 \cdot c.$$

From this, or a similar model, we get fitted values $\hat{p}$ which are individual estimates of probabilities of $Y = 1$ (risk estimates)

Assessing the operating characterisitics of a risk estimate



Sources of bias

Type of bias	Description
Verification bias	Non-random selection of patients for definitive diagnosis
Error in gold standard	Imperfect gold standard diagnosis/classification
Spectrum bias	Cases and controls come from difference populations
Interpretation bias	Information is available in the study that can distort the classification
Overdiagnosis	Subclinical disease may be detected by screening that would have never progressed

Study designs

Name	Description
Prospective	Measure biomarker, then follow up people for the true outcome
Retrospective	Take sample with known outcome, measure biomarker
Case-control	Retrospective where sampling depends on the outcome
Clinical trial	Secondary study as part of clinical trial
Prospective-retrospective	Study pre-registered with protocol, biomarker measured in stored specimens with known outcome

Phases of prediction model development

- Anecdotal Observation
- Designed case-control or cohort studies - associations, preliminary predictive models
- Pre-clinical experiments - Laboratory development of clinical assay, *analytic validity*: Can we measure what we want to?
- Signature finding studies - Retrospective or prospective
- Prospective evaluation of fixed signature - *clinical validity*: Does the signature accurately predict the clinical outcome?
- Randomized efficacy studies - *clinical utility*: Does the use of the prediction in the clinic lead to benefit to patients?

Note the parallels to the phases of drug development