

Lesson 21: Class Imbalance and Prevalence

MA376 Applied Statistics, AY27-1

Cadet: _____

Section: _____

Objectives

1. **Frame.** Say what a rare outcome does to each metric before any model is fit.
2. **Apply.** Handle imbalance three ways (threshold, class weights, undersampling) and compare what each one flags.
3. **Mean.** Each fix moves the cutoff and none improves the model. Precision moves with prevalence.
4. **Communicate.** Report a classifier's performance with the prevalence it was measured at.

1. The clinic model

4,000 clinic appointments. `lead_days` is how far ahead the appointment was booked. `no_show` is 1 if the patient missed it. The model is fit on the training set and flags a test appointment when $\hat{p} \geq 0.5$.

$$\text{logit } P(\text{no-show}) = \text{_____} + \text{_____} (\text{lead days})$$

	Truth: no-show	Truth: kept
Predicted no-show	TP =	FP =
Predicted kept	FN =	TN =

Accuracy	$\frac{TP + TN}{n} =$ _____	Sensitivity	$\frac{TP}{TP + FN} =$ _____
Specificity	$\frac{TN}{TN + FP} =$ _____	Precision	$\frac{TP}{TP + FP} =$ _____
F_1	$\frac{2TP}{2TP + FP + FN} =$ _____	Null classifier	_____

1. The model flags an appointment booked _____ days out or more. Is this a model worth having? Compare it with the null classifier.

2. What if the data looked like this?

Same clinic, but only _____ percent of appointments are no-shows. Same model form, same threshold of 0.5, scored on the new test set.

	Truth: no-show	Truth: kept
Predicted no-show	TP =	FP =
Predicted kept	FN =	TN =

	Accuracy	Null classifier	Sensitivity	Specificity	Precision	F_1
Before						
Now						

2. Which numbers are out of whack? Accuracy went up. Is the model better?

3. The cutoff moved from _____ days to _____ days. Why does a rare outcome push it that way?

3. Three fixes

Each fix is fit on the training set and scored on the same test appointments.

	Threshold	Flag at (days)	Caught	False alarms	Sensitivity	Precision	F_1
Plain	0.5						
Fix 1: tuned threshold							
Fix 2: class weights	0.5						
Fix 3: undersample	0.5						

4. Fix 1 changes the threshold and keeps the model. What do Fix 2 and Fix 3 change?

5. Did any fix make the model better at telling a no-show from a kept appointment? What did each one do?

6. A missed no-show wastes an appointment slot. A false alarm costs one reminder call. Which row would you give the clinic, and why? Would F_1 have picked it?

4. Precision moves with prevalence

Prevalence: _____

Precision: _____

$$\text{precision} = \frac{\text{sensitivity} \cdot \text{prevalence}}{\text{sensitivity} \cdot \text{prevalence} + (1 - \text{specificity})(1 - \text{prevalence})}$$

Precision when	Plain at 0.5	Fix 1	Fix 2	Fix 3
10% are no-shows				
2.6% are no-shows				

7. The model did not change between those two rows. Why did precision?

8. In one or two sentences, report Fix 1's performance to the clinic commander. Include the prevalence it was measured at.

Key Takeaway. A rare outcome breaks the 0.5 threshold, not the model. The threshold is yours to set, from what a miss and a false alarm cost. Report sensitivity and precision together, with the prevalence they were measured at.