

Classification Performance

Practical Machine Learning (with R)

UC Berkeley

CLASSIFICATION PERFORMANCE

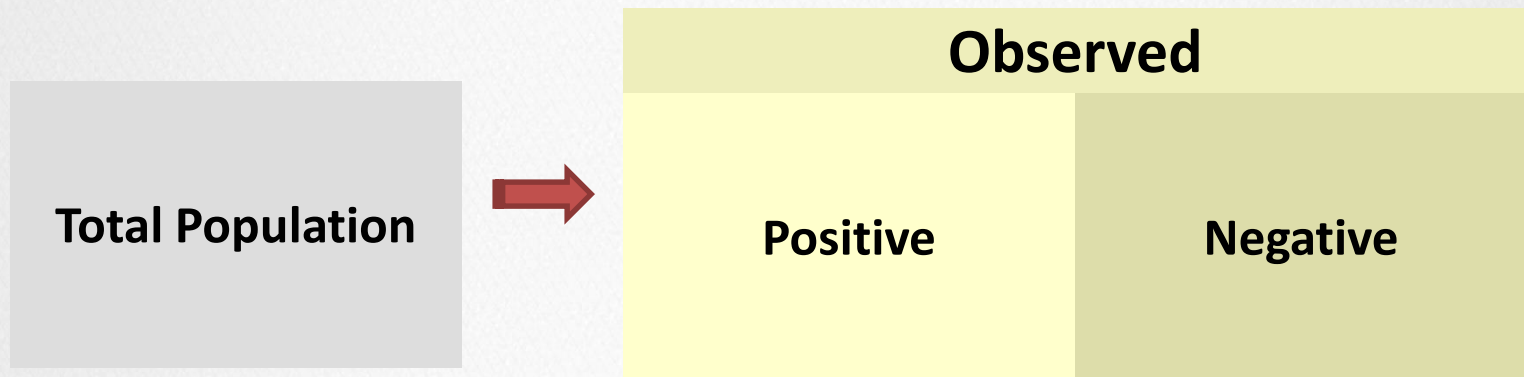


METRICS FOR BINOMIAL CLASSIFICATION



Total Population





Total Population



Predicted

Positive

Negative



Total Population



Observed	
Positive	Negative



Predicted	Positive
	Negative

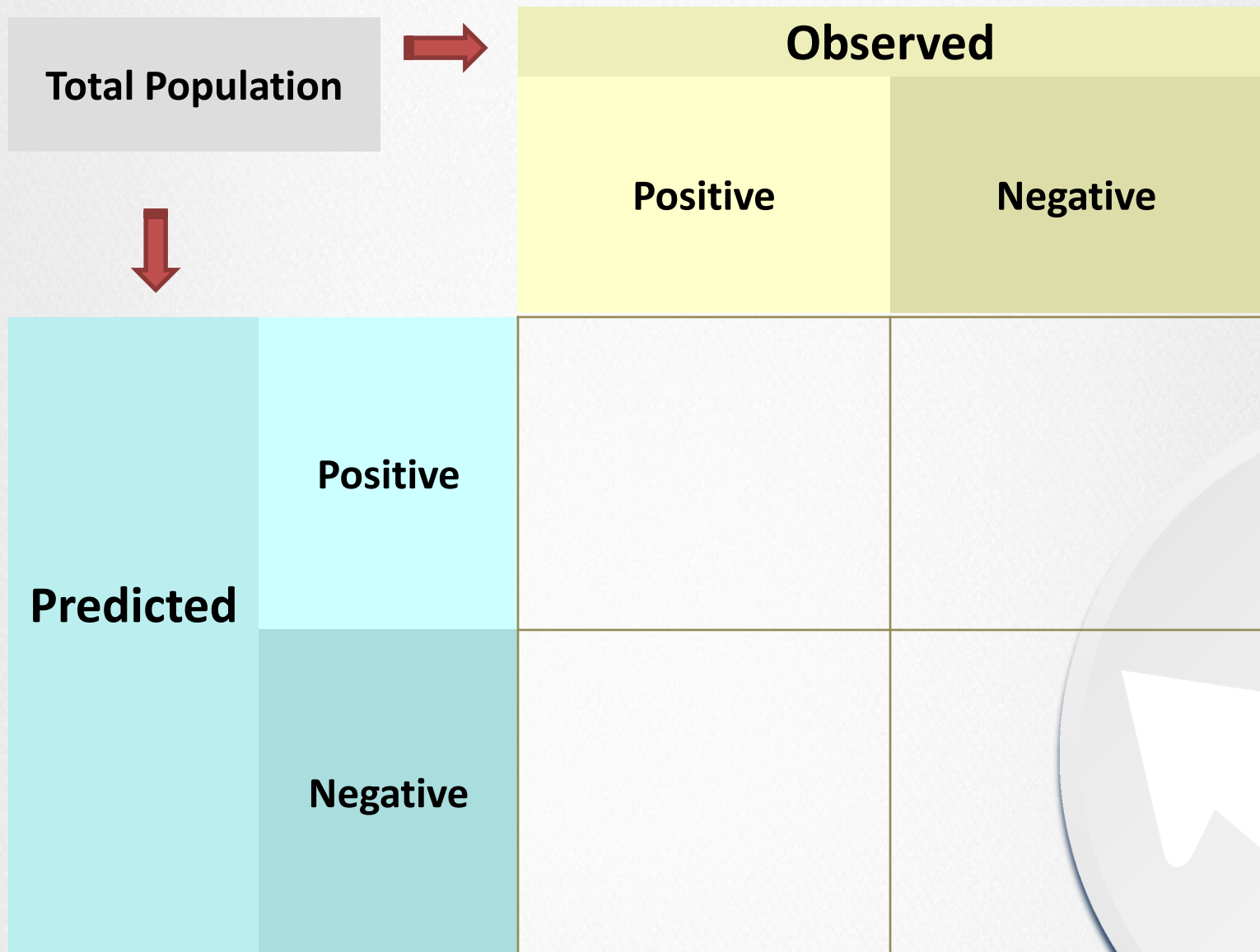
Accuracy

“How many did I get right”

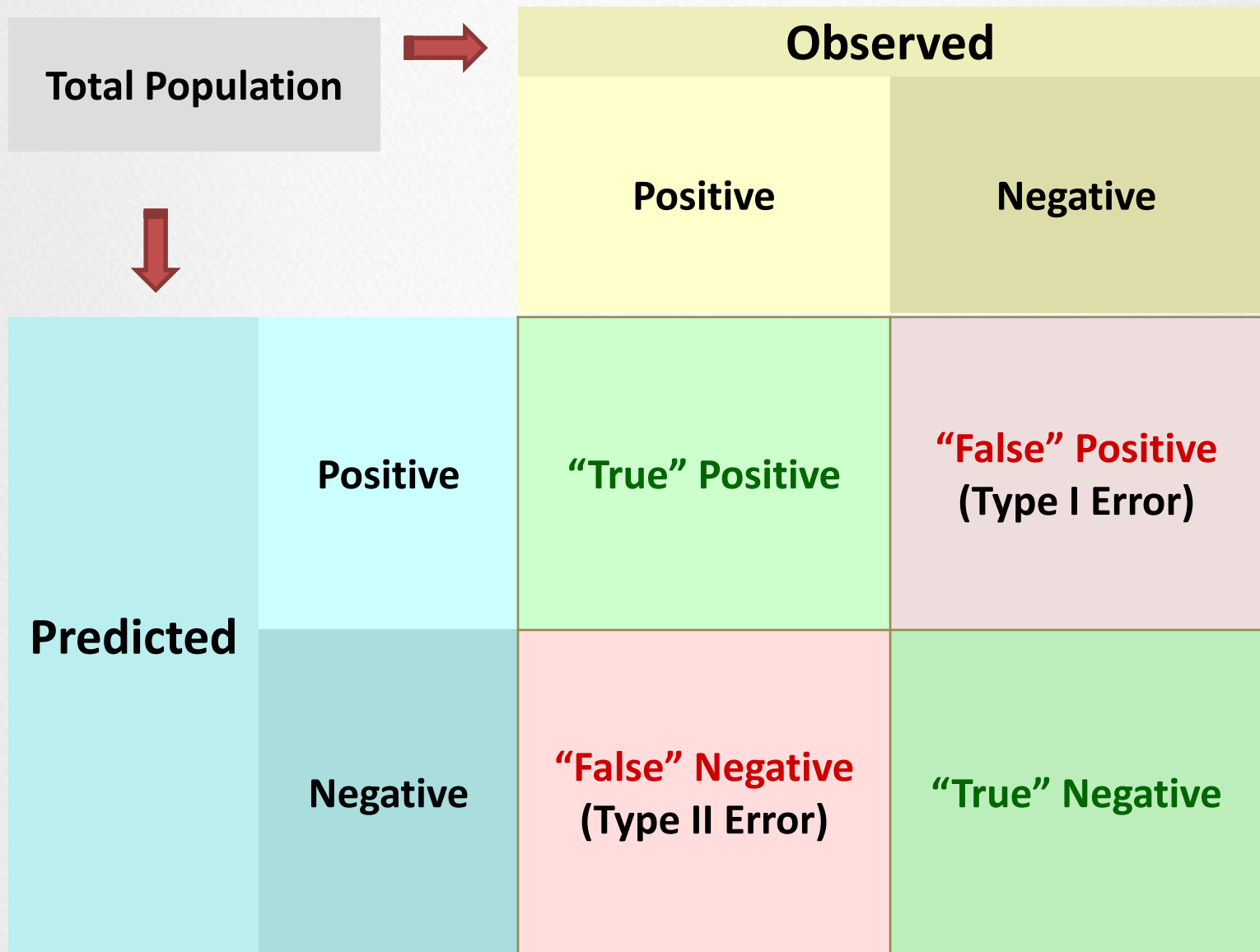
Error Rate

or Misclassification Rate
“How many did I get wrong”

CONFUSION MATRIX



- https://en.wikipedia.org/wiki/Sensitivity_and_specificity



Alternatives: Norm by Observed

		Observed	
		Positive	Negative
Predicted	Positive	True Positive Rate (TPR), Sensitivity , Recall $\frac{\text{True Positives}}{\text{Observed Positives}}$	False Positive Rate (FPR), Fall-Out $\frac{\text{False Positives}}{\text{Observed Negatives}}$
	Negative	False Neg. Rate (FNR), Miss rate $\frac{\text{False Negatives}}{\text{Observed Positives}}$	True Neg. Rate (TNR), Specificity (SPC) $\frac{\text{True Negatives}}{\text{Observed Negatives}}$

Alternatives: Norm by Predicted

		Observed	
		Positive	Negative
Predicted	Positive	Pos. Predictive Value (PPV), Precision $\frac{\text{True Positives}}{\text{Predicted Positives}}$	False Discovery Rate (FDR) $\frac{\text{False Positives}}{\text{Predicted Positives}}$
	Negative	False Omission Rate (FOR) $\frac{\text{False Negatives}}{\text{Predicted Negatives}}$	Negative Predictive Value (NPV) $\frac{\text{True Negatives}}{\text{Predicted Negatives}}$

- https://en.wikipedia.org/wiki/Sensitivity_and_specificity

MORE FUN ...

https://en.wikipedia.org/wiki/Sensitivity_and_specificity

https://en.wikipedia.org/wiki/precision_and_recall



EXERCISE: CLASSIFICATION METRICS



MUTLI-NOMIAL CLASSIFICATION



CLASSIFICATION PERFORMANCE

⇒ `predict` methods can provide

- Classes
- Class probabilities

Better!

⇒ Class probs → Classes?

- Apply **softmax** function

$$\hat{p}_\ell^* = \frac{e^{\hat{y}_\ell}}{\sum_{l=1}^C e^{\hat{y}_l}}$$

⇒ Probabilities often need post predict → calibrations (talk about this with deployment)

CLASSIFICATION PERFORMANCE

- ➔ Accuracy ... problems?
- ➔ Confusion Matrix
 - `table`
 - `caret::confusionMatrix`
- ➔ Cohen's Kappa: $\kappa = \frac{O-E}{1-E}$
 - Kappa values within (0.30-0.50)+ → good fit
- ➔ ROC Curves / Lift Charts



EVEN MORE COMPLICATION ...

- ⇒ Not all errors need count “equivocal zone” or “intermediate zone”
- ⇒ *Prevalent when the model has three choices, e.g. A or B or Nothing.*



TERMS

- ⇒ Kappa Statistic,
 - ⇒ S-Statistic, F-Statistic
-

