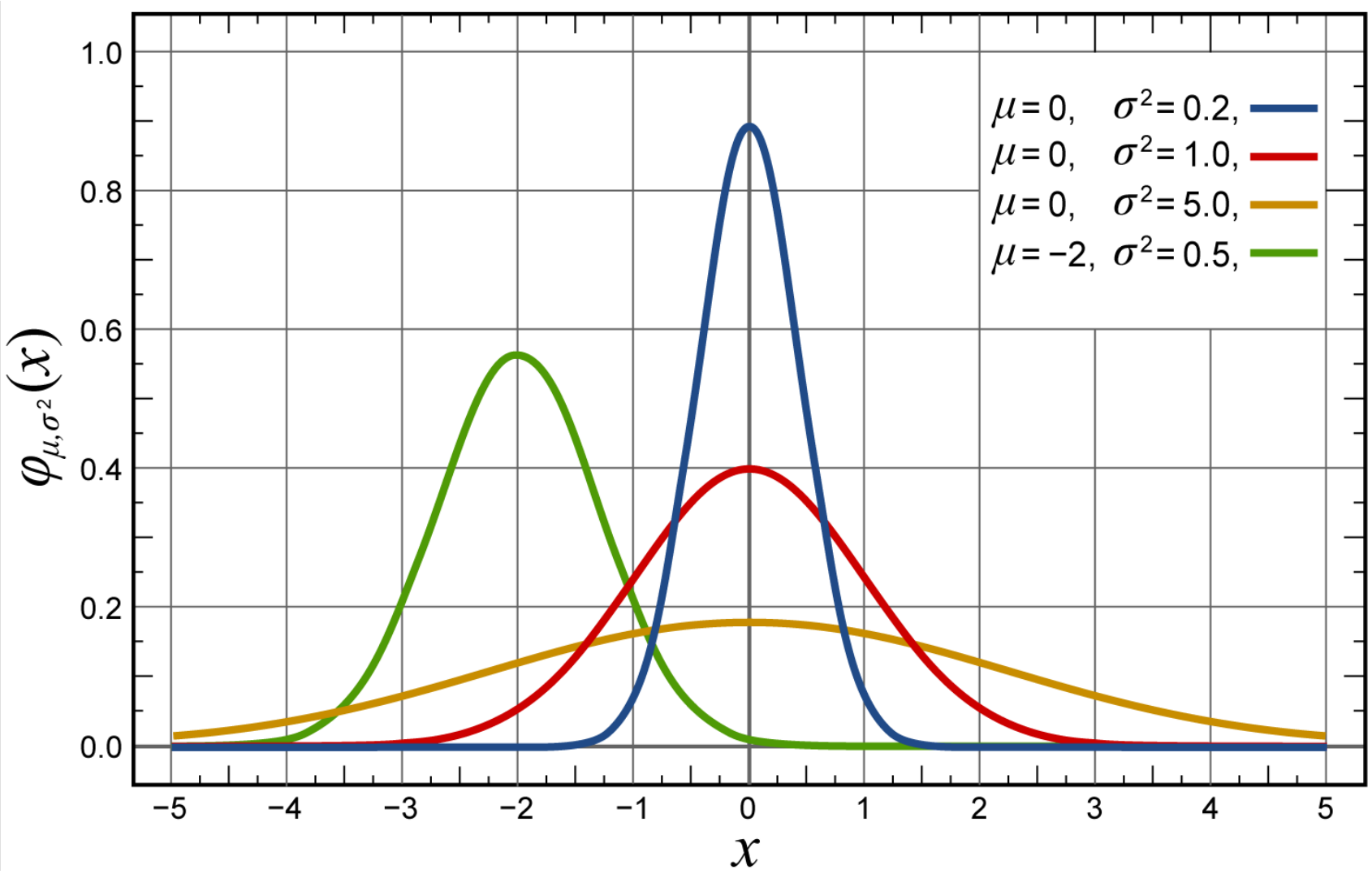
Introduction to Machine Learning

Brown University CSCI 1950-F, Spring 2011
Prof. Erik Sudderth

Lecture 5: Bayesian Estimation, Decision Theory, & ROC Curves

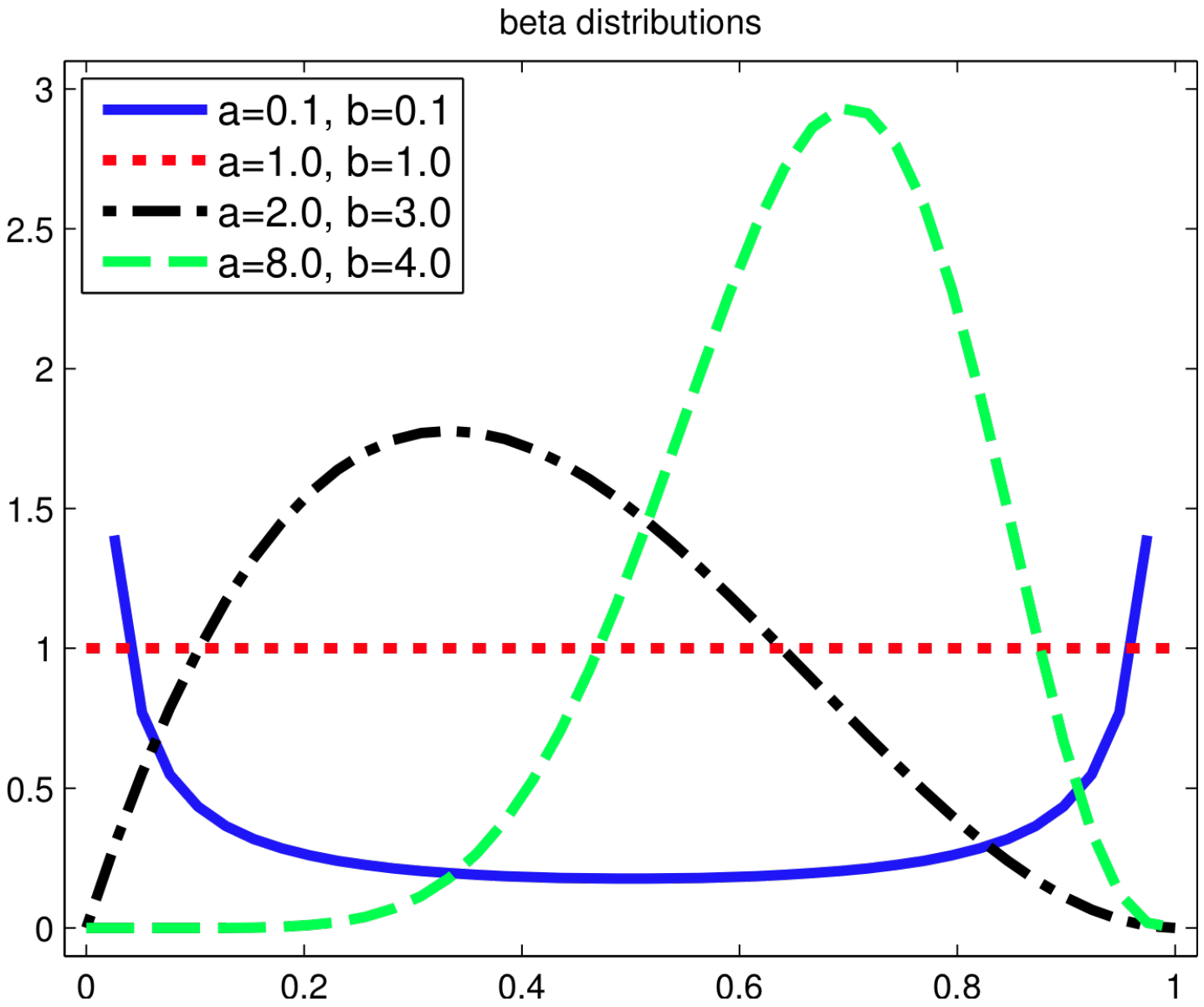
Many figures courtesy Kevin Murphy's textbook,
Machine Learning: A Probabilistic Perspective

Gaussian (Normal) Distributions

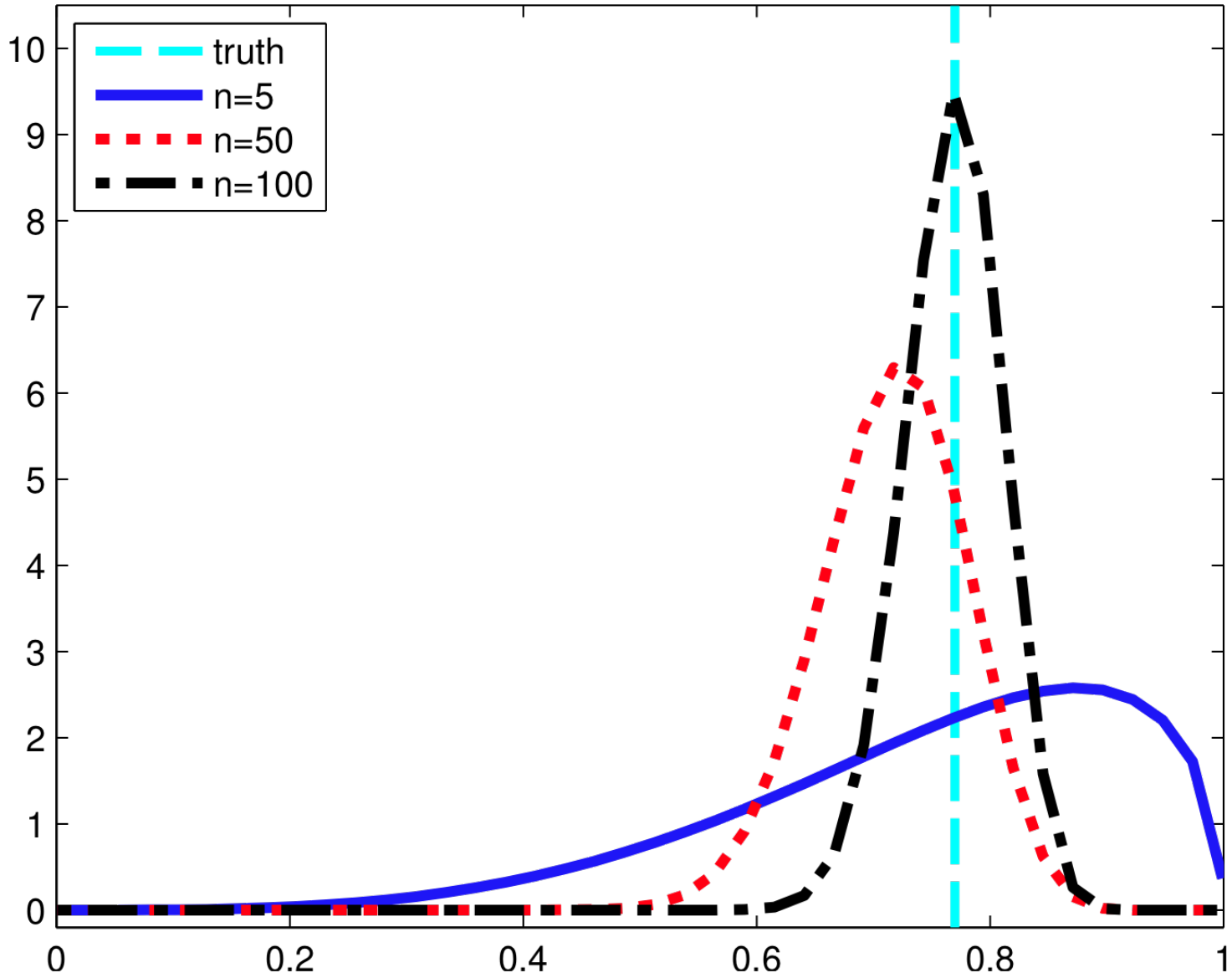


Summaries: Mean, median, mode, variance, standard deviation

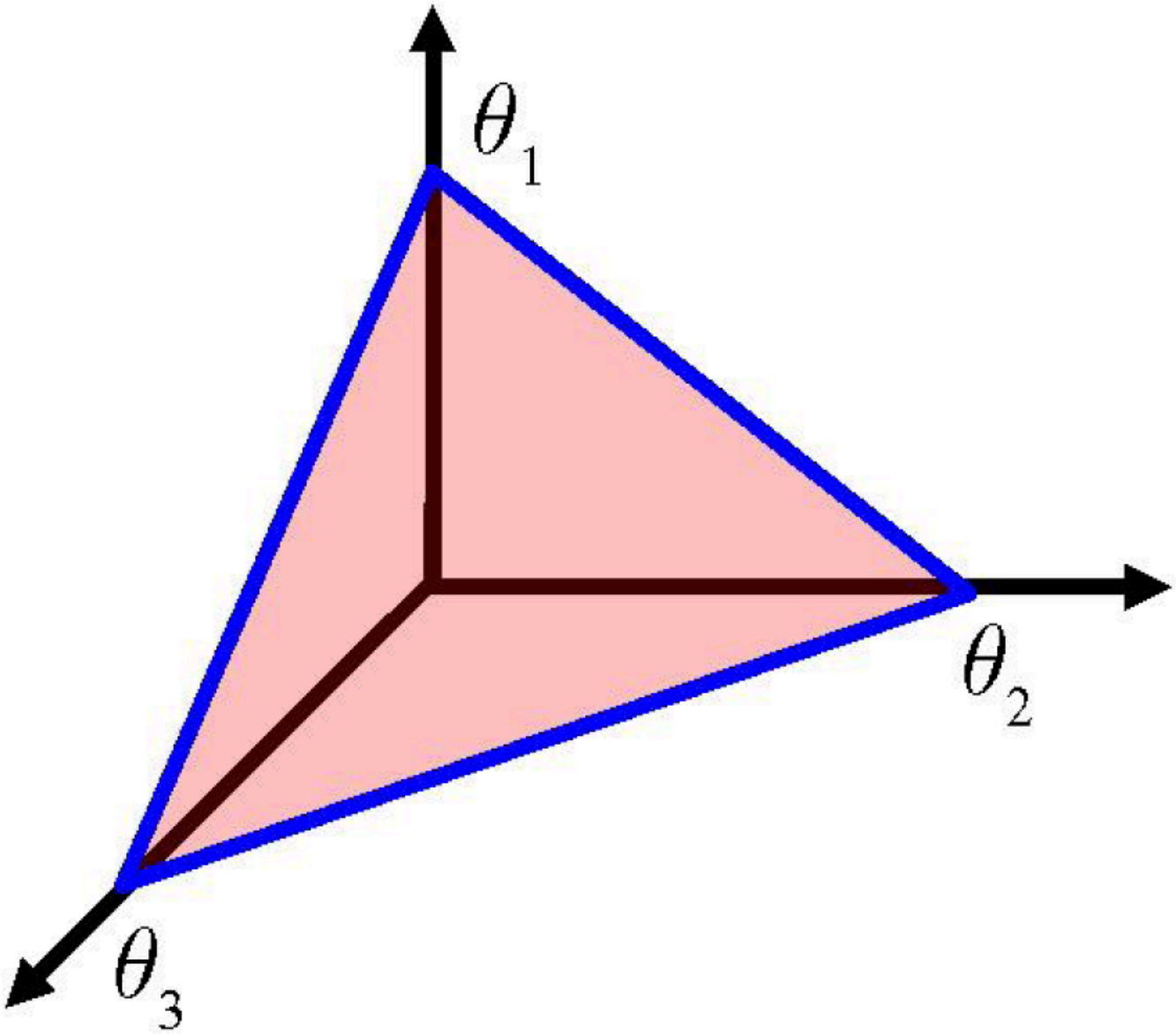
Beta Distributions



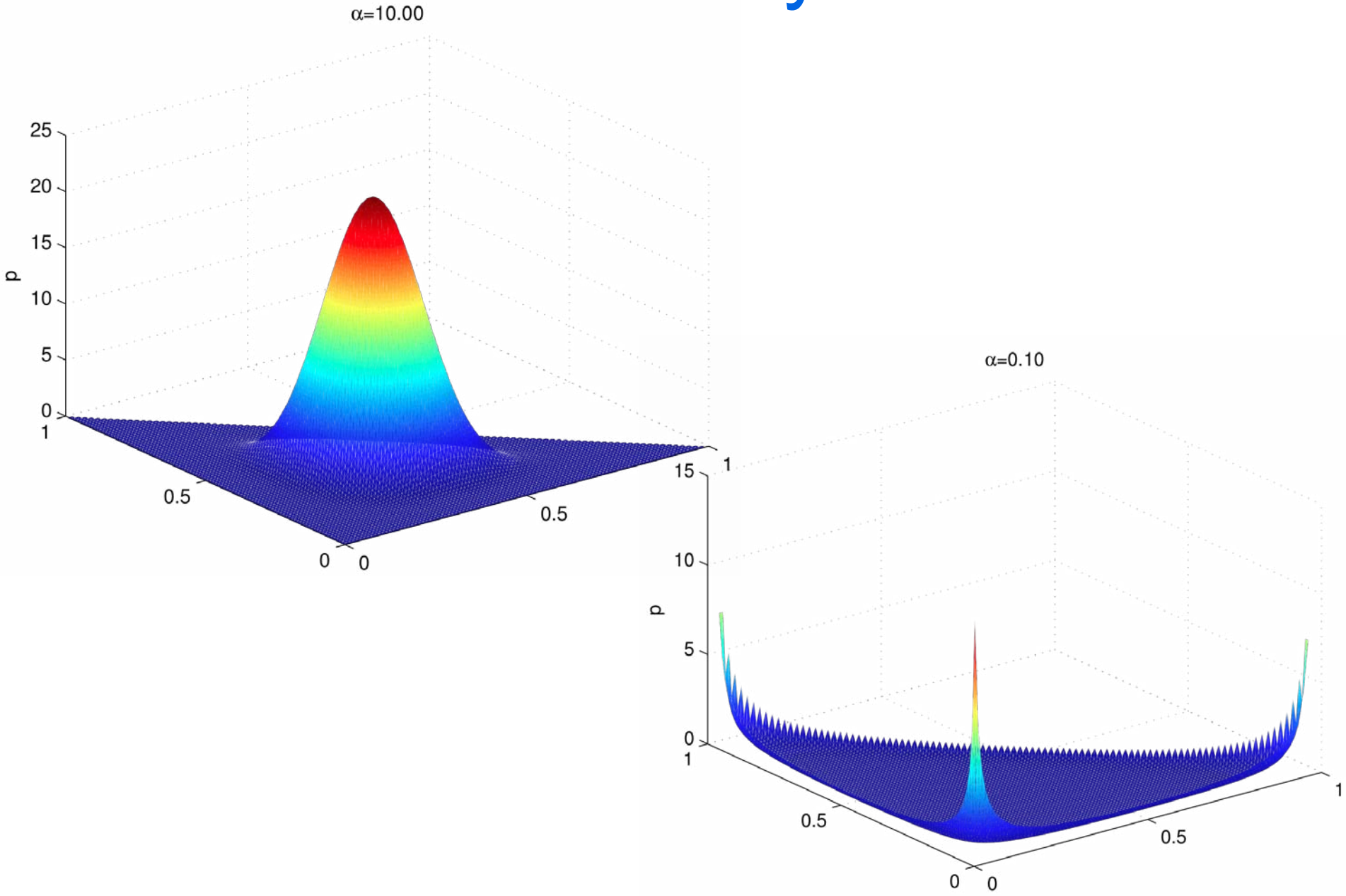
Sequence of Beta Posteriors



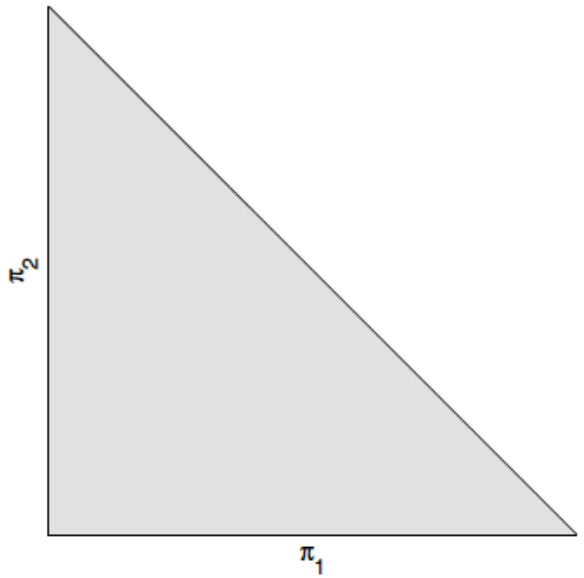
Multinomial Simplex



Dirichlet Probability Densities

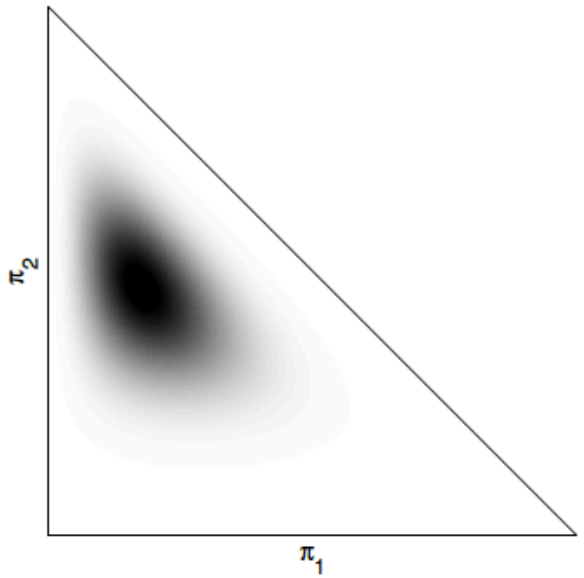


Dirichlet Probability Densities

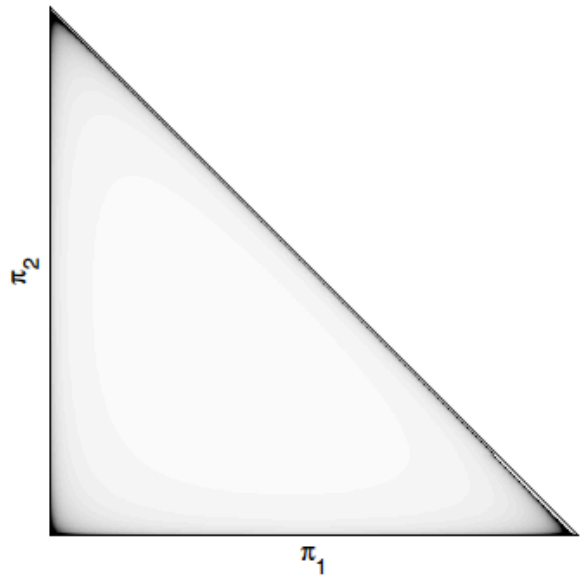


$\pi \sim \mathrm{Dir}(1, 1, 1)$

$\pi \sim \mathrm{Dir}(4, 4, 4)$



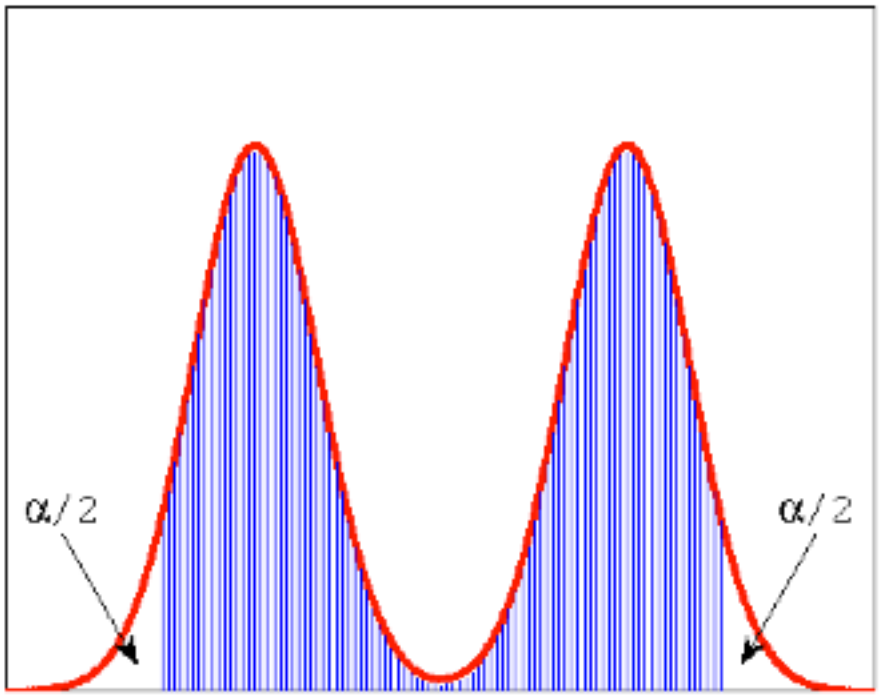
$\pi \sim \mathrm{Dir}(4, 9, 7)$



$\pi \sim \mathrm{Dir}(0.2, 0.2, 0.2)$

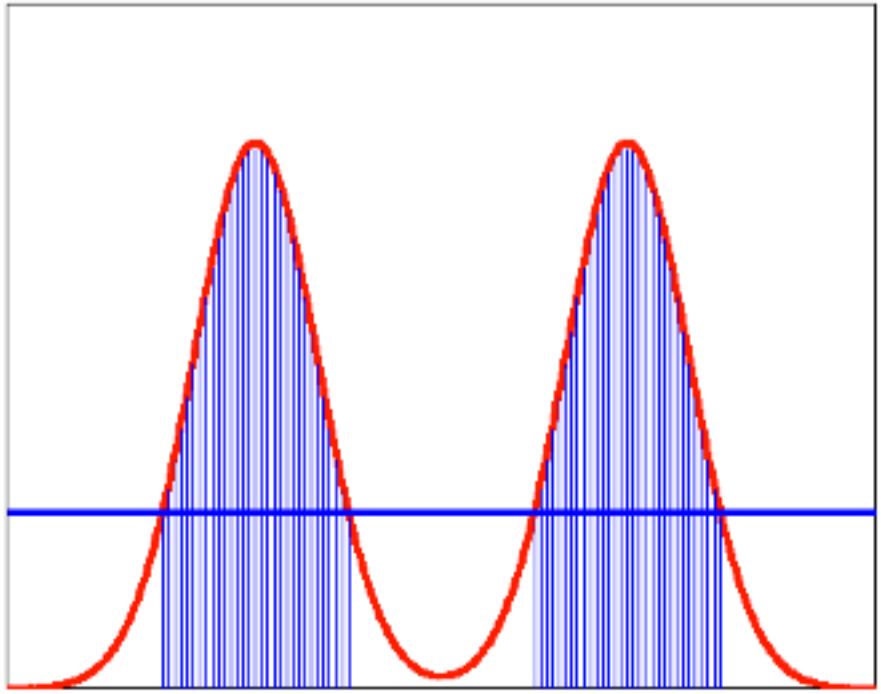
Summaries of Posterior Distributions

Central Credible Interval



$$C_\alpha(\mathcal{D}) = (\ell, u) : P(\ell \leq \theta \leq u|\mathcal{D}) = 1 - \alpha$$

Highest Posterior Density Region



$$C_\alpha(\mathcal{D}) = \{\theta : p(\theta|\mathcal{D}) \geq p^*\}$$

$$1 - \alpha = \int_{\theta:p(\theta|\mathcal{D})>p^*} p(\theta|\mathcal{D})d\theta$$

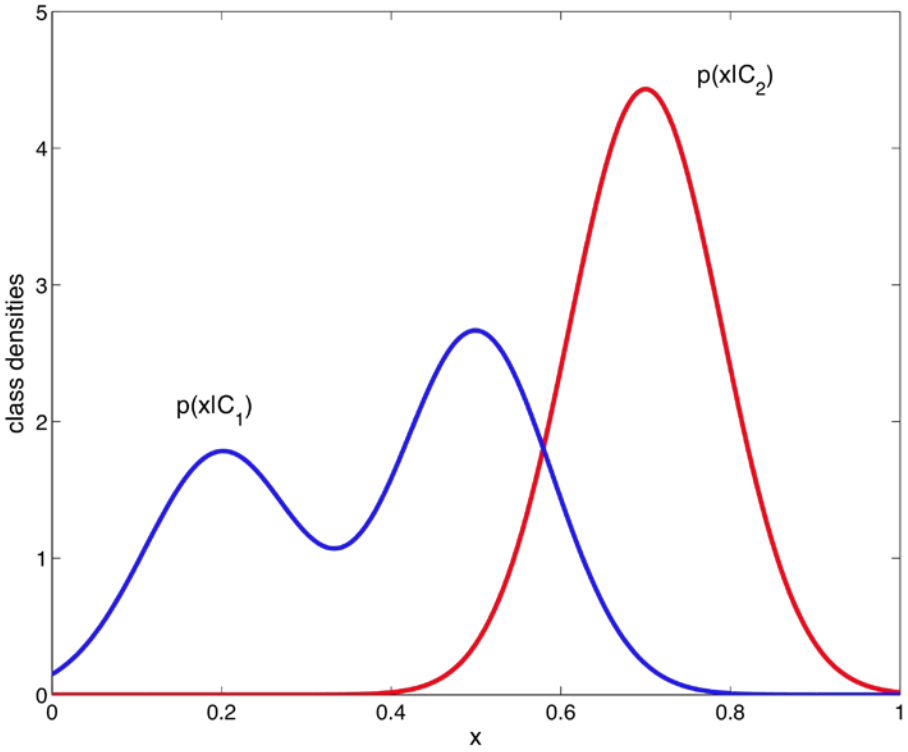
Model Selection: Bayes' Factors

$$BF_{1,0} := \frac{p(\mathcal{D}|M_1)}{p(\mathcal{D}|M_0)}$$

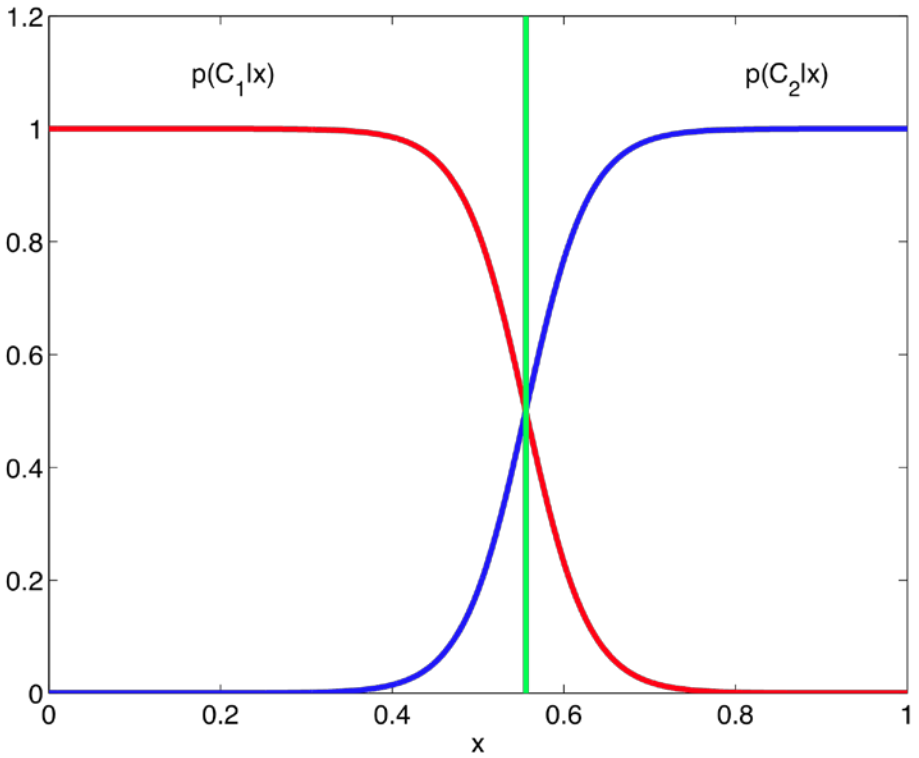
Bayes factor $BF(1,0)$	Interpretation
$B < \frac{1}{100}$	Decisive evidence for H_0
$B < \frac{1}{10}$	Strong evidence for H_0
$\frac{1}{10} < B < \frac{1}{3}$	Moderate evidence for H_0
$\frac{1}{3} < B < 1$	Weak evidence for H_0
$1 < B < 3$	Weak evidence for H_1
$3 < B < 10$	Moderate evidence for H_1
$B > 10$	Strong evidence for H_1
$B > 100$	Decisive evidence for H_1

As suggested by Jeffreys. Caveats: Can exhibit sensitivity to choice of priors for each model's parameters.
Most reliable when comparing pairs of "similar" models.

Binary MAP Estimation



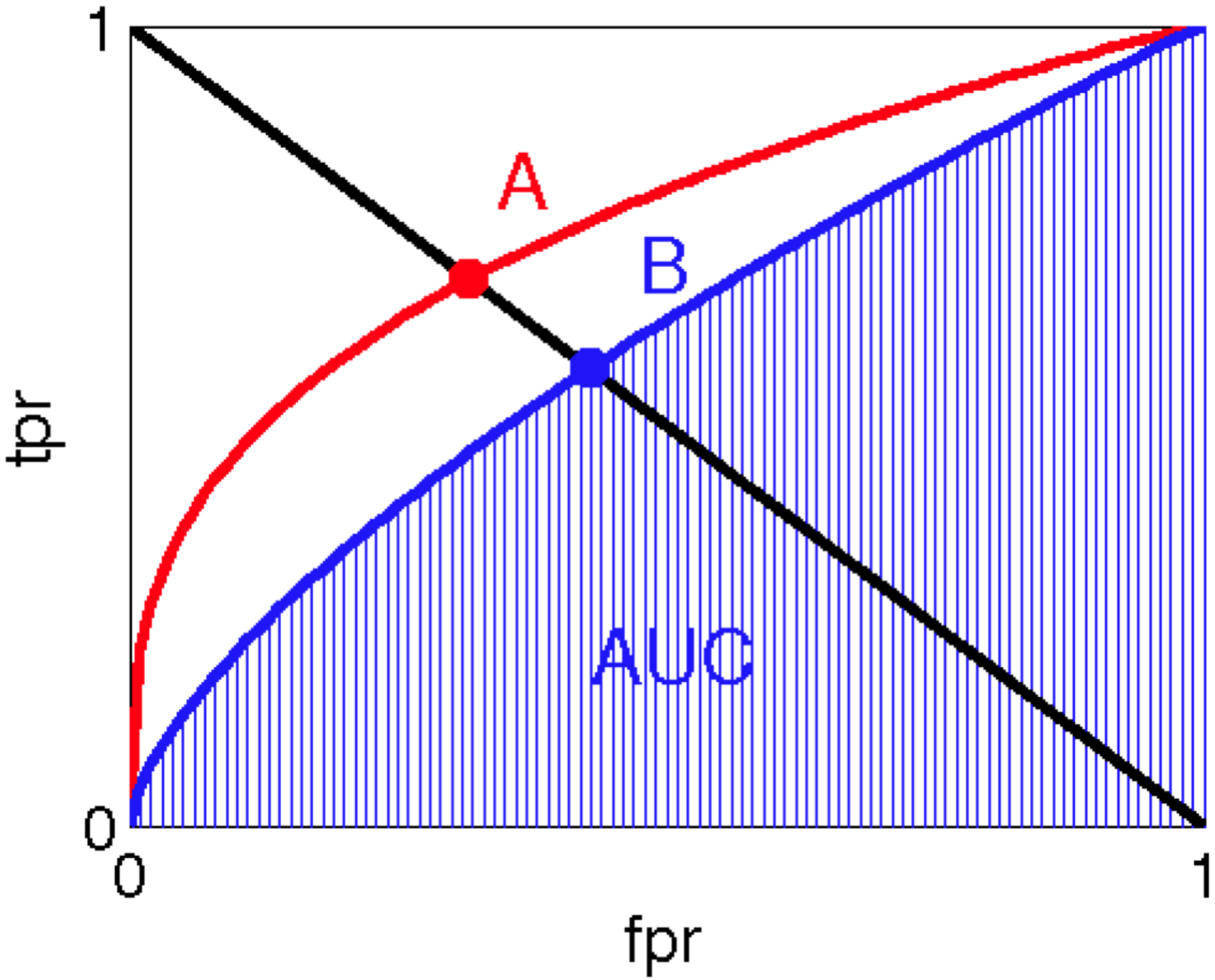
Likelihood Functions



Posterior Probabilities

False Positives vs. False Negatives

		Truth		
		1	0	Σ
Estimate	1	TP	FP	$\hat{N}_+ = TP + FP$
	0	FN	TN	$\hat{N}_- = FN + TN$
	Σ	$N_+ = TP + FN$	$N_- = FP + TN$	$N = TP + FP + FN + TN$

	$y = 1$	$y = 0$	$y = 1$	$y = 0$
$\hat{y} = 1$	$TP/\hat{N}_+$=precision	$FP/\hat{N}_+$=FDP	TP/N_+=TPR	FP/N_-=FPR
$\hat{y} = 0$	$FN/\hat{N}_-$	$TN/\hat{N}_-$=NPV	FN/N_+=FNR	TN/N_-=TNR

ROC Curves