Introduction to Machine Learning

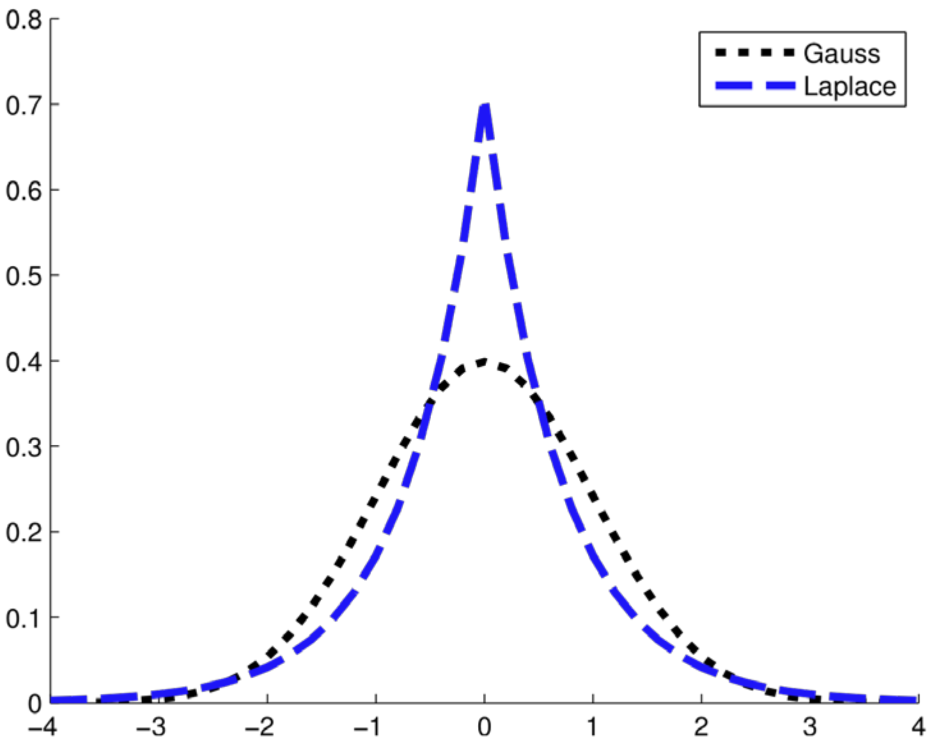
Brown University CSCI 1950-F, Spring 2011
Prof. Erik Sudderth

Lecture 12: Exponential Families, Generalized Linear Models, Robust Regression

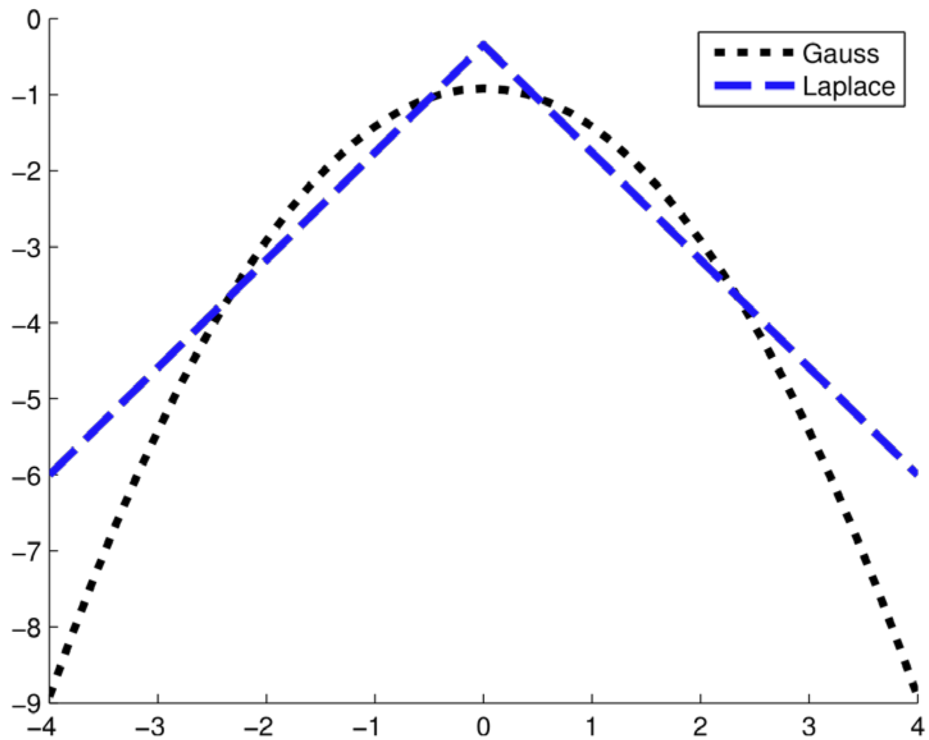
Many figures courtesy Kevin Murphy's textbook, *Machine Learning: A Probabilistic Perspective*

Laplace Distribution

Probability Densities



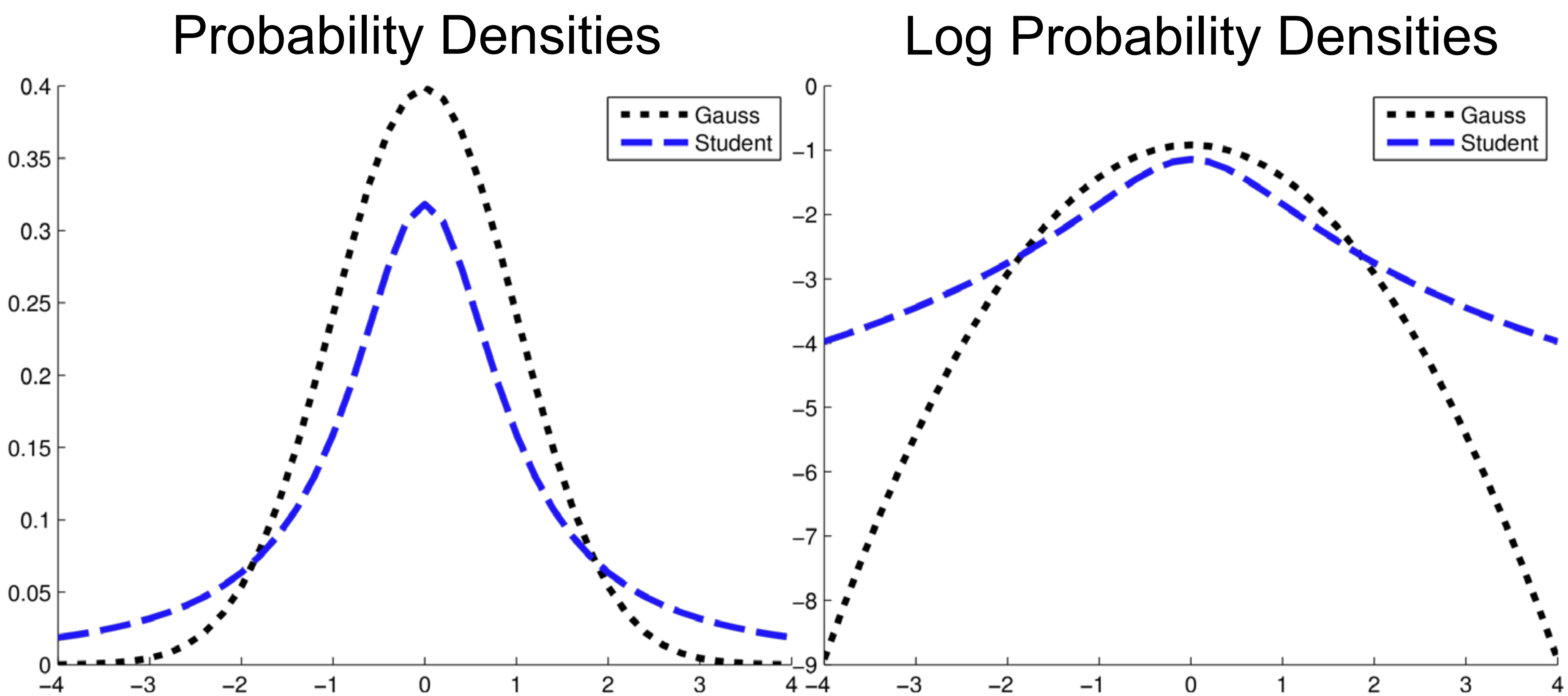
Log Probability Densities



Relative to Gaussian distributions with equal variance:

- Many samples are near zero
- Occasional large-magnitude samples are far more likely
- Negative log probability density is *convex but not smooth*

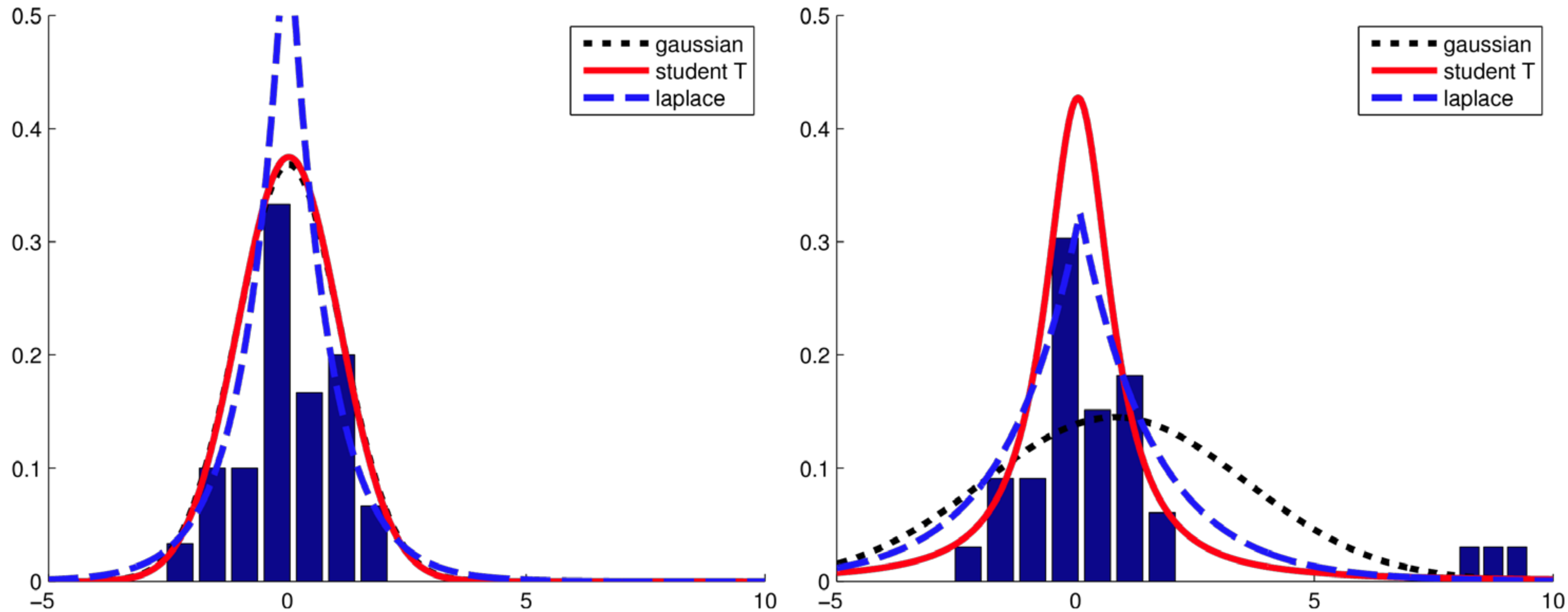
Student T Distribution



Relative to Gaussian distributions with equal variance:

- Approaches Gaussian as DOF parameter approaches infinity
- For small DOF, large-magnitude samples are far more likely
- Negative log probability density is *smooth but not convex*

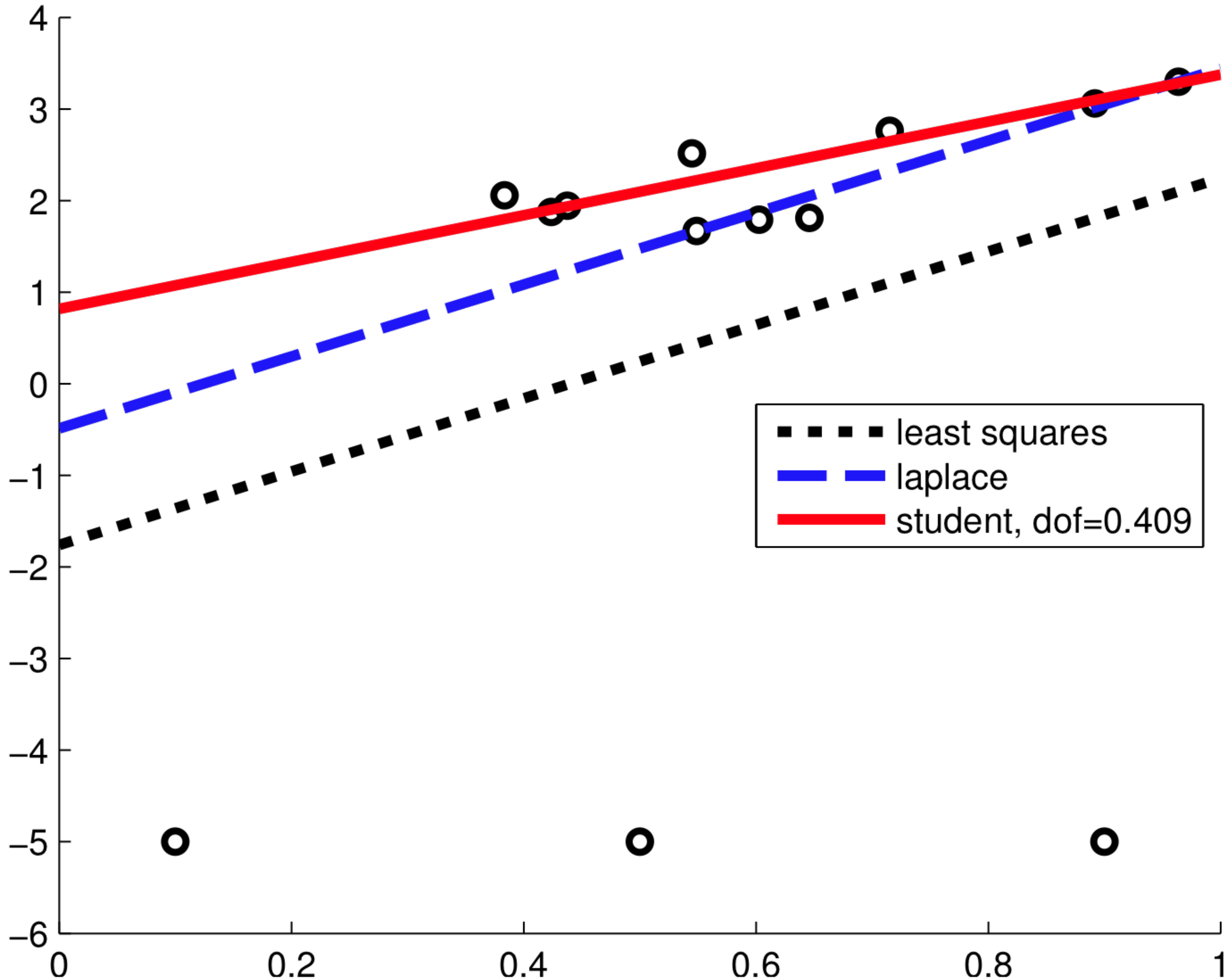
Outliers & ML Estimation



Maximum likelihood estimates of mean parameters:

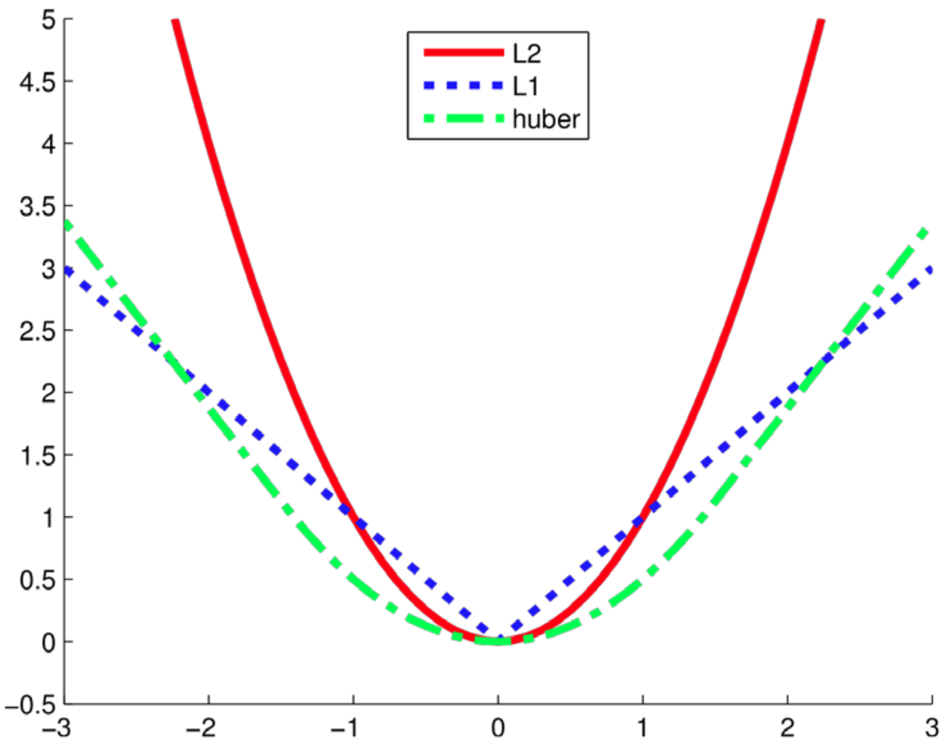
- Gaussian: Sample mean of data
- Laplacian: Sample median of data
- Student T: No closed form, optimize via gradient methods

Outliers & Linear Regression

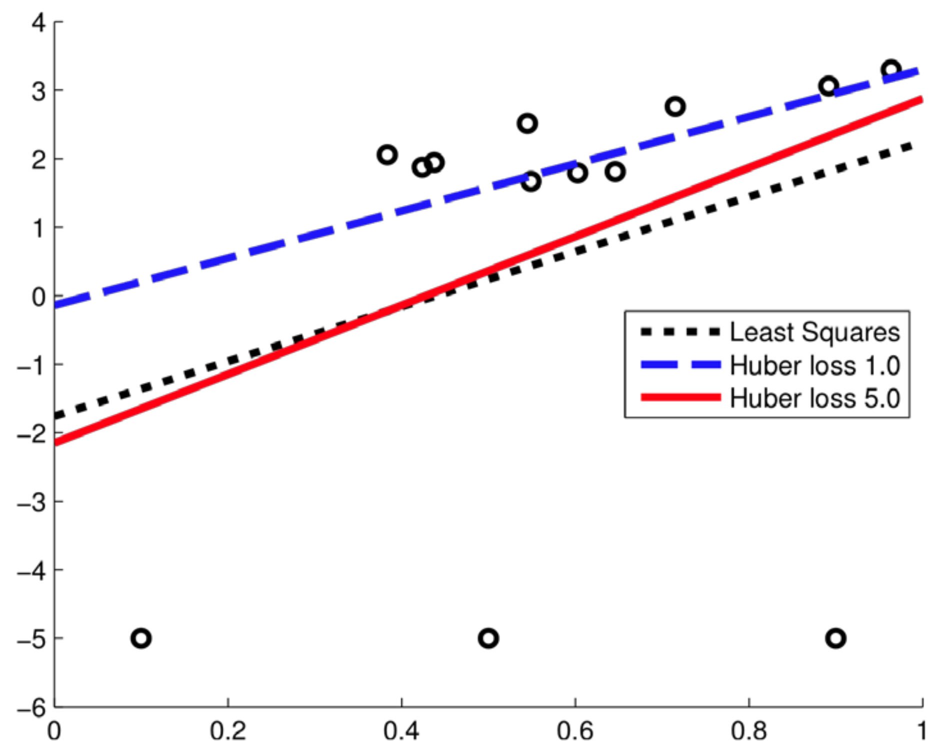


Huber Loss Function

Negative Log Probabilities



Robust Linear Regression



Relative to Gaussian distributions with equal variance:

- Behaves like Gaussian near origin ("non-outliers")
- Behaves like Laplacian far from origin (robustness)
- Negative log probability density is *smooth and convex*