

CS-498 HW 5: Due 11 Nov 2013. Do this individually.

In this exercise, you will compare Maximum Likelihood and Maximum A Posteriori inference methods using a simulation.

- 1) Build a simple simulation of a Bernoulli random variable X . Write p for $P(X=1)$. Your simulation should accept a value of p , then obtain N IID samples of X (use `rand` in Matlab).
- 2) Now draw 10 independent samples from a beta distribution with parameters $\alpha=2$, $\beta=2$. These will serve as values of p . For each N value in the set $\{1, 2, 5, 10, 15, 100\}$, draw 10 sets of N IID samples from your simulation, one for each of your values of p .
- 3) For each of these sixty datasets, infer the value of p using (a) maximum likelihood; (b) MAP estimation with a beta prior with $\alpha=2$, $\beta=2$; and (c) MAP estimation with a beta prior with $\alpha=8$, $\beta=8$. Compare the results of (a), (b), and (c) with one another, and with the right answer. You should plot a scatter plot of the results of each method against the right answer. Under which circumstances is which method better?
- 4) The beta prior with $\alpha=8$, $\beta=8$ is not a good model of the true prior, but it's not awful either (the peak is in the right place). What happens if you use a beta prior with $\alpha=16$, $\beta=1$?