Foundations of Graphical Models: Homework 0

Due: Fr 2019-09-06

This homework will give you a sense of the background needed to take this course and the type of thinking and mathematics that you will use.

Please prepare your answers using LaTeX with the template provided on the course website. Submit the PDF of your completed assignment on Courseworks.

Problem 1 (via David Duvenaud)

Consider a probability density $p(x|\mu, \sigma^2) = \mathcal{N}(x|\mu, \sigma^2)$, where $\mu \in \mathbb{R}$ and $\sigma \in \mathbb{R}^+$.

a) For some $x \in \mathbb{R}$, can $p(x) < 0$?

b) For some $x \in \mathbb{R}$, can $p(x) > 1$?

Problem 2 (via Joe Blitzstein)

You have a jar of 1,000 coins. 999 are fair coins, and the remaining coin will always land heads. You take a single coin out of the jar and flip it 10 times in a row, all of which land heads. What is the probability your next toss with the same coin will land heads? Explain your answer.

Problem 3 (via David Duvenaud)

In the exponential family of distributions, $p(x|\theta) = h(x) \exp\{\eta(\theta)^\top t(x) - a(\theta)\}$, for $x \in \mathbb{R}^n, \theta \in \mathbb{R}^d, h(x) : \mathbb{R}^n \to \mathbb{R}, \eta(\theta) : \mathbb{R}^d \to \mathbb{R}^p, t(x) : \mathbb{R}^n \to \mathbb{R}^p$, and $a(\theta) : \mathbb{R}^d \to \mathbb{R}$. What must $a(\theta)$ be for $p(x|\theta)$ to be a valid probability distribution? Why?

Problem 4

Consider n i.i.d. random variables $X_1, \ldots, X_n$ with probability density

$$f(x|k) = kx^{k-1}e^{-x^k}.$$

The density is only defined for positive random variables, so assume all observations are positive. The parameter k is also positive.

a) Write down the log-likelihood of the data.

b) Write down the derivative of the log-likelihood with respect to k.

c) k is constrained to be positive, but often times it is convenient to work with parameters that live on an unconstrained space. To do this, we reparametrize the distribution. Let $\theta \in \mathbb{R}$ be such that $k = \exp(\theta)$. Write the log-likelihood as a function of θ, and the derivative of this log-likelihood with respect to θ.

d) Now consider the parameterization $k = \log(1 + \exp(\theta))$. Write the log-likelihood and derivative with respect to θ. At a high level, compare these two parameterizations. What are the benefits and drawbacks of each?