Stats 545: Midterm exam (75 minutes for 30 points)

Write your name and PUID on each sheet, and also include the number of answer sheets. Attempt all questions. The difficulty of a question need not be correlated with its score.

1 Miscellaneous [9 pts]

1. Briefly explain the bisection method, and Newton's method for root-finding, explaining the intuition behind the latter. Give an advantage and disadvantage of Newton's method over bisection. [4pts]

2. Given a function $f(x)$, you want to find an x^* satisfying $f(x^*) = 0$. Recast this root-finding problem as an optimization problem. Explain why gradient descent is not enough to solve this. [2pts]

3. Briefly explain the difference between gradient descent and stochastic gradient descent. [1pts]

4. a) What is a prior distribution? b) What is a conjugate prior? c) Give an example of a conjugate prior, specifying the associated likelihood. [2pts]

2 Monte Carlo methods [7 pts]

1. Let X and Y be Poisson distributed random variables with parameters λ_X and λ_Y (recall $\text{Poiss}(x|\lambda) = \lambda^x \exp(-\lambda)/x!$. Assume you can simulate from and evaluate the Poisson distribution for any parameter. Give a simple Monte Carlo algorithm to estimate the probability $p(X = Y)$. Explain whether or not your algorithm is inefficient if λ_Y is much larger than λ_X. If so, outline an algorithm that will work for this case too. [3pts].

2. For importance sampling with proposal and target distributions $q(x)$ and $p(x)$, a sample x has weight $w(x) = p(x)/q(x)$. a) What is the mean and variance of $w(x)$? b) Define effective sample size (ESS) for importance sampling, and explain its intuition. c) Suppose $p(x) = f(x)/Z$. How will you estimate Z? [4pts]

3 MCMC [14 pts]

1. What are the pros and cons of MCMC versus Monte Carlo sampling? [1pt]

2. a) What is the stationary distribution of a Markov chain? b) Give a Markov chain without a stationary distribution. c) What does irreducibility mean? d) Give a *non-irreducible* MCMC chain involving (i) Metropolis-Hastings (MH), and (ii) Gibbs sampling. [4pts]

3. a) Write the transition kernel of MH chain with stationary distrib. $\pi(x)$ and proposal distrib. $q(x^*|x)$? b) Explain why for MH your acceptance probability should in general not be too small or too close to 1? [3pts]

4. Let X_1, X_2, and X_3 be three Poisson random variables with parameter λ. You want to calculate the probability that all three are less than 4, given that their sum is less than 10. Describe a) a rejection sampling approach, and b) a Gibbs sampling approach to estimate this. Explain when you might prefer one over the other. [4pts]

5. You want to sample from $p(x) \propto \frac{1}{1+x^2}$ using MH. If your current sample is x, you propose $x^* \sim N(x, 1 + x^2)$. Thus your proposal distribution $q(x^*|x)$ is a Gaussian with mean x and variance $(1 + x^2)$. What is the acceptance probability? How will you use this to estimate the mean of $\log(x)$ under $p(x)$? [2pts]