Stats 545: Midterm exam 2

This is a 75-minute exam for 30 points. Write your name and PUID on each sheet, and also include the number of answer sheets. Attempt all questions.

1 Root-finding and optimization [5 pts]

1. Let $f(x)$ be a continuous 1-dimensional function. Let x^* be a root of f, i.e. $f(x^*) = 0$. Explain the bracketing condition to verify that an interval $[a, b]$ contains x^*. Give an example f and x^*, where the root x^* does not satisfy this. [2pts]

2. Briefly explain the secant and false position methods to find the root x^*, giving an advantage and disadvantage of the former. [2pts]

3. Briefly explain the advantage of conjugate gradient descent over standard gradient descent for minimization. [1pts]

2 Monte Carlo sampling [9 pts]

1. To sample from $p(x) \propto f(x) \leq Mq(x)$, rejection sampling proposes x^* from q and accepts with probability $\frac{f(x^*)}{Mq(x^*)}$. Briefly explain why this scheme produces samples from $p(x)$. What is the overall acceptance probability? [2pts]

2. You can randomly generate uniform numbers between 0 and 1. You want to simulate from the exponential distribution truncated to $[1, 2]$: $p(x) \propto \exp(-x)\mathbb{1}_{[1,2]}(x)$. How will you do this using a) the inverse-cdf method b) rejection sampling. For the latter, give M and q. [3pts]

3. Consider the Laplace distribution: $p(x) = \frac{1}{2}\exp(-|x|)$ where x lies on the real line. You want to calculate $\mathbb{E}_p[X^2]$ under this distribution using importance sampling. You can only generate samples from a standard normal $N(0, 1)$. Explain how you will do this, giving the importance weights. Explain what a problem with this estimate might be. [2pts]

4. Give an two advantages and two disadvantages of importance sampling vs regular Monte Carlo. [2pts]

3 MCMC [11 pts]

1. What is meant by the stationary distribution of a Markov chain? What are *all* the stationary distributions of the Markov chain $x_{i+1} = x_i$ (i.e. that just stays in its initial state). Given another example of a Markov chain that does not have a unique stationary distribution. [3pts]

2. Derive the transition kernel of Metropolis-Hastings (MH) chain that targets the stationary distribution $\pi(x)$ with proposal distribution $q(x^*|x)$, explaining all terms. [2pts]

3. What is detailed balance in the context of Markov chains? Show that a Markov chain satifying detailed balance w.r.t. a distribution π has π as its stationary distribution. The converse is not true: give an example of a Markov chain that has a stationary distribution, but does not satisfy detailed balance with respect to its stationary distribution. [3pts]

4. What is effective sample size (ESS) in the context of MCMC? Explain how running an MCMC chain for *more* iterations can result in a *decrease* in ESS. Briefly explain the Gelman-Rubin diagnostic and how it is useful. [3pts]

4 MCMC 2 [5 pts]

You want to sample from $p(x, y) \propto \exp\left(-x^2 - (y - x^2)^2\right)$ where both x and y are real-valued. The only random-number generators you have are a Gaussian and a Uniform(0,1) random number generator.

1. Describe a Metropolis-Hastings sampler to do this, giving the proposal distribution and the acceptance probability. [2pts]

Consider a Gibbs sampler to do this:

2. What is the conditional distribution $p(y|x)$? In particular, what family does it belong to? [1pts]

3. Write down the conditional distribution $p(x|y)$ That is not a standard distribution. If you were to sample from this by rejection sampling, write down the proposal distribution and acceptance probability. [2pts]