UNIVERSITY OF TORONTO
Faculty of Arts and Science

STA2311H: Advanced Computational Methods for Statistics I

Final Exam

December 5, 2023

3 hours

Name: ______________________________

Student Number: ________________________

- Do not open this test until you are told to begin.
- This final exam is closed-book; a one-sided handwritten cheat sheet is allowed.
- There are four questions on the final, each worth 25 marks. Take a quick scan through the questions first and prioritize your time accordingly.
- Show all of your work for full marks, and ensure your notation is legible, correct, and consistent with that used in the course.
- If you need to use a result from lecture or the practice problems, briefly describe it.
- If you run out of space, use the back of the page.
- **PhD students:** All four questions will be counted, for a total mark out of 100.
- **Masters students:** Your best three questions will be counted, for a total mark out of 75.

Good luck!

1. (a) (12.5 points) Suppose we are interested in estimating the following integral using Monte Carlo:

$$I = \int_0^\infty (2 - e^{-e^x}) \cdot \lambda e^{-\lambda x} \, dx,$$

where $\lambda > 0$. Describe (in detail) a Monte Carlo variance reduction strategy that would be effective in this case.

(b) (12.5 points) Derive a slice sampler to sample from the density $f(x) \propto e^{-(x+\mu)^2/2}$, where $x \in [0, 1]$ and $\mu > 0$.

2. Let π, g be two densities, and suppose there exists a constant $M > 1$ such that $\pi(x) < M \cdot g(x)$ for all $x \in \text{Supp}(\pi)$.

(a) (10 points) Show that a rejection sampler for π with proposal g has an acceptance probability of $1/M$.

(b) (15 points) Show that an independence sampler for π with proposal g has an acceptance probability of *at least* $1/M$.

3. Suppose we are in the importance sampling setup, where the goal is to estimate $I = \mathbb{E}_f[h(X)]$ and f is a hard density to sample from, g is an easy density to sample from, and $\text{Supp}(f) \subseteq \text{Supp}(g)$. Recall that the importance weights are defined by

$$w_i = \frac{f(Y_i)/g(Y_i)}{\sum_{j=1}^n f(Y_j)/g(Y_j)}, \quad Y_1, \dots, Y_n \overset{iid}{\sim} g.$$

Consider a two-stage algorithm in which we first draw $Y_1, \dots, Y_n \overset{iid}{\sim} g$, compute the weights as above, and then for each $i = 1, \dots, n$, we take $Z_i = Y_j$ with probability w_j.

(a) (15 points) Show that, assuming a law of large numbers holds, each Z_i is asymptotically distributed according to f as $n \to \infty$ (equivalently, $\mathbb{P}(Z_n \leq z) \xrightarrow{n\to\infty} \mathbb{P}(X \leq z)$, where $X \sim f$[1]).

(b) (10 points) Show that $\hat{I} = \sum_{i=1}^n w_i \cdot h(Y_i)$ is a consistent estimator of I.

[1]Technically $\mathbb{P}(Z_n \leq z)$ is a random variable because the Y_n are defined on a different probability space, but the mode of convergence is not important here.

4. (a) (15 points) Consider a Gibbs sampler to sample (X_1, X_2, X_3) from some 3-dimensional target distribution π. Suppose that at each iteration t, we sample at random (without replacement) indices i_1, i_2 from $\{1, 2, 3\}$ and then update only (X_{i_1}, X_{i_2}); that is, we draw

$$(X_{i_1}^{(t)}, X_{i_2}^{(t)}) \sim \pi(\cdot, \cdot \mid X_{i_3}^{(t-1)}), \quad t \geq 1$$

where $i_3 \in \{1, 2, 3\} \setminus \{i_1, i_2\}$. Show that the Markov chain underpinning this algorithm has π as a stationary distribution.

(b) (10 points) Suppose that of n nuclear reactors, the number of pump failures y_i for plant i follows a Poisson distribution with parameter $\lambda_i t_i$, where t_i is the duration of observation (which is known). If we impose independent Gamma(α, β) priors on each λ_i, a Gamma(γ, δ) hyperprior on β, and α, γ, and δ are all known, derive a Gibbs sampler for sampling from the posterior distribution of $(\lambda_1, \ldots, \lambda_n, \beta)$ given the data $\mathbf{y} = (y_1, \ldots, y_n)$.[2]

[2] Recall that the Poisson(λ) distribution has pmf $p(k) = \lambda^k e^{-\lambda}/k!$ for $k \in \mathbb{N}$ and the Gamma(α, β) distribution has pdf $f(x) = (\beta^\alpha/\Gamma(\alpha))x^{\alpha-1}e^{-\beta x}$ for $x > 0$.