Please provide complete and well-written solutions to the following exercises.

Due November 17, 9AM, to be submitted in blackboard, under the Assignments tab.

Homework 6

Exercise 1. Let $f: \mathbf{R} \to [0,\infty)$ be a PDF. Suppose you can sample (on a computer) any number of i.i.d. real valued random variables $X_1, X_2, \ldots$, each with PDF X. Let $g: \mathbf{R} \to [0,\infty)$ be another PDF. Assume there exists some $m > 0$ such that $g(x) \leq mf(x)$ $\forall\, x \in \mathbf{R}$.

The goal of this exercise is to sample from a random variable with PDF g.

State a version of accept/reject sampling with these assumptions and goal.

Prove that your version of accept/reject sampling outputs some random variable Z with PDF g.

Exercise 2. Let $X_1, \ldots, X_n$ be i.i.d. Gaussian random variables with unknown mean $\mu \in \mathbf{R}$ and variance 1. Define $Y_i := \max(X_i, 0)$, for all $1 \leq i \leq n$. Without loss of generality (i.e. by re-ordering the random variables), assume that $Y_1, \ldots, Y_m > 0$ and $Y_{m+1} = \cdots = Y_n = 0$.

In this problem, we assume that we cannot access $X_1, \ldots, X_n$, but we can access $Y_1, \ldots, Y_n$.

- Explain how you could use the EM algorithm to estimate μ from $Y_1, \ldots, Y_m$. Give details about the E and M steps. Let μ_k denote the estimate of μ from the k^{th} iteration of the EM algorithm. Show that
$$\mu_{k+1} = \frac{1}{n}\sum_{i=1}^{m} Y_i \quad + \frac{n-m}{m}\mu_k - \frac{n-m}{m}\frac{\phi(\mu_k)}{\Phi(-\mu_k)}, \qquad \forall\, k \geq 1. \qquad (*)$$
Here ϕ is the PDF of a standard Gaussian, and Φ is the CDF of a standard Gaussian.
- Find the log-likelihood function $\log \ell(\mu)$ based on the observed data $Y_1, \ldots, Y_n$, and use it to write down a (nonlinear) equation that the MLE Z_n satisfies.
- Use the equation in the previous part to verify that Z_n is a fixed point of the recursion $(*)$.
- Prove that μ_k converges in distribution to μ as $k \to \infty$, for any starting value of $\mu_0 \in \mathbf{R}$, assuming that $m \geq 1$. Hint: Show that $|\mu_k - Y_k|$ decreases as $k \to \infty$. Hint: Use the Mean Value Theorem, and you can also freely use the following inequality
$$0 < \frac{\phi(x)[\phi(x) - x\Phi(-x)]}{[\Phi(-x)]^2} < 1, \qquad \forall\, x \in \mathbf{R}.$$

Exercise 3. Let $\theta \in (0, 3/4)$ be unknown. Define
$$(p_1, p_2, p_3, p_4) := ((1+\theta)/2,\ (1-\theta)/4,\ 1/4 - \theta/3,\ \theta/12).$$

Let $X=(X_1,X_2,X_3,X_4)$ be a multinomial distribution, so that

$$\mathbf{P}(X=x)=\binom{200}{x_1,x_2,x_3,x_4}\prod_{i=1}^{4}p_i^{x_i},$$

for any $x=(x_1,x_2,x_3,x_4)$ that are nonnegative integers with $x_1+x_2+x_3+x_4=200$.

- Write an equation that would need to be solved in order to obtain an MLE for θ.
- Assume now that instead you have data from a multinomial distribution $Y=(Y_1,\ldots,Y_6)$, but $Y_1,\ldots,Y_4$ are not observed, and $X_1=Y_1+Y_2$, $X_2=Y_3+Y_4$ are both observed, along with $Y_5=X_3$ and $Y_6=X_4$. (Choose convenient probabilities $q_1,\ldots,q_6$ for the multinomial $Y_1,\ldots,Y_6$. For example, consider $q_1=1/2$, $q_2=\theta/2$, $q_3=1/4-\theta/3$, $q_4=\theta/12$, $q_5=1/4-\theta/3$, $q_6=\theta/12$.) Write down the E and M steps of the EM algorithm for estimating θ, under the above assumptions.
- Is the EM algorithm simpler than directly finding the MLE of θ?

Exercise 4. Let $0<p,q<1$. Let $P=\begin{pmatrix}1-p & p\\ q & 1-q\end{pmatrix}$. Find the (left) eigenvectors of P, and find the eigenvalues of P. By writing any row vector $x\in\mathbf{R}^2$ as a linear combination of eigenvectors of P (whenever possible), find an expression for xP^n for any $n\geq 1$. What is $\lim_{n\to\infty}xP^n$? Is it related to the vector $\pi=(q/(p+q),p/(p+q))$?

Exercise 5. Suppose we have a Markov chain $X_0,X_1,\ldots$ with finite state space Ω. Let $y\in\Omega$. Define $L_y:=\max\{n\geq 0\colon X_n=y\}$. Is L_y a stopping time? Prove your assertion.

Exercise 6 (Knight Moves)**.** Consider a standard 8×8 chess board. Let V be a set of vertices corresponding to each square on the board (so V has 64 elements). Any two vertices $x,y\in V$ are connected by an edge if and only if a knight can move from x to y. (The knight chess piece moves in an L-shape, so that a single move constitutes two spaces moved along the horizontal axis followed by one move along the vertical axis (or two spaces moved along the vertical axis, followed by one move along the horizontal axis.) Consider the simple random walk on this graph. This Markov chain then represents a knight randomly moving around a chess board. For every space x on the chessboard, compute the expected return time $\mathbf{E}_xT_x$ for that space. (It might be convenient to just draw the expected values on the chessboard itself.)