MTH 210A 2023: Midsem Exam

Instructions:

(a) Write your name and roll number on all supplemental sheets.

(b) Marks will not be given unless all work is shown. Show **all** mathematical details.

Questions:

1. (20 points) For $0 < \alpha < 1$, $0 < p < 1$ and $n \in \mathbb{N}$, consider the probability mass function

$$\Pr(X = i) = \begin{cases} \alpha + (1-\alpha)(1-p)^n & \text{if } i = 0 \\ (1-\alpha)\binom{n}{i} p^i (1-p)^{n-i} & \text{if } i \in \{1, 2, 3, \ldots, n\} \end{cases}$$

$Y \sim Bin(n, p)$
$P(Y=i) = {}^nc_i\, p^i (1-p)^{n-i}$, $i = 0, 1, \ldots, n$

10 (a) Write down all the steps to implement an accept-reject algorithm using a $\text{Bin}(n, p)$ proposal.

10 (b) Write down all the steps to sample from this distribution using the composition method.

(a) $Y \sim Bin(n, p)$ For A-R: then $c = \sup_{i\in\{0,1,\ldots,n\}} \frac{P(X=i)}{P(Y=i)} = \frac{\alpha + (1-\alpha)\ldots}{p^i(1-p)^{n-i}}$; $i = 0, 1, 2, \ldots, n$ for $i=0 = \frac{\alpha + (1-\alpha)(1-p)^n}{(1-p)^n}$

For A-R, we need to find

$$c = \sup_{i=\{0,1,\ldots,n\}} \frac{\Pr(X=i)}{\Pr(Y=i)} = \sup_i \frac{[\alpha + (1-\alpha)(1-p)^n]\,\mathbb{I}(i=0) + [(1-\alpha)\binom{n}{i} p^i (1-p)^{n-i}]\,\mathbb{I}[i\neq 0]}{\binom{n}{i} p^i (1-p)^{n-i}}$$

$i \neq 0 \Rightarrow (1-\alpha)$. $\frac{1}{(1-p)^n} + \ldots > 0$

$c = (1-\alpha) + \frac{\alpha}{(1-p)^n}$

for $i=0$: $\frac{\Pr(X=0)}{\Pr(Y=0)} = \frac{\alpha + (1-\alpha)(1-p)^n}{(1-p)^n} = (1-\alpha) + \underbrace{\frac{\alpha}{(1-p)^n}}_{>0}$

$i \neq 0$: $\frac{\Pr(X=i)}{\Pr(Y=i)} = \frac{(1-\alpha)\binom{n}{i} p^i (1-p)^{n-i}}{\binom{n}{i} p^i (1-p)^{n-i}} = (1-\alpha)$

(-4) or (-5) if not done calculation

$\Rightarrow c = \underbrace{(1-\alpha) + \frac{\alpha}{(1-p)^n}}$

(-3) or (-4) if didn't solve to get this

Algo: ① Draw $U \sim U(0,1)$ and $Y \sim Bin(n,p)$ {independently}. (-1) if not written

② If $U \leq \frac{\Pr(X=y)}{c.\Pr(Y=y)}$, then return $X = y$.

③ Else, go to step ①.

Algo: Draw $U \sim U[0,1]$ & $Y \sim Bin(n,p)$ independently. If $U \leq \frac{P(X=y)}{c.\Pr(Y=y)}$ then $X = Y$

(b) Note that:

$$\Pr(X=i) = [\alpha + (1-\alpha)(1-p)^n]\,\mathbb{I}(x=0) + [(1-\alpha)\binom{n}{i} p^i (1-p)^{n-i}]\,\mathbb{I}(x\neq 0)$$
$$= \alpha\,\mathbb{I}(x=0) + (1-\alpha)\left[\binom{n}{i} p^i (1-p)^{n-i}\right]\mathbb{I}(x\in\{0,1,2,\ldots,n\})$$

Define $X_{(1)}$: $\Pr(X_{(1)} = 0) = 1$ and $X_{(2)} \sim Bin(n, p)$.

(-5) if composition is not correct. or (-6)

Algo ① Draw $U \sim U(0,1)$

② If $U \leq \alpha$, then set $X = 0$

③ Else, draw $X \sim Bin(n,p)$ [independently of U]

Note: $P(X=i) = \alpha\,\mathbb{I}(x=0) + (1-\alpha)\{{}^nc_i\, p^i (1-p)^{n-i}\}\,\mathbb{I}(x\in\{0,1,\ldots,n\})$

Define X_1: $P(X_1 = 0) = 1$, $P(X_1 \neq 0) = 0$ & $X_2 \sim Bin(n,p)$

2. (20 points) Consider target distribution F with supposrt $\mathcal{X}$ and density $f(x) = m\tilde{f}(x)$, where $m > 0$ is **unknown**, and $\tilde{f}$ is known. Suppose G is a proposal distribution with density $g(x)$ with support $\mathcal{Y}$ such that $\mathcal{X} \subseteq \mathcal{Y}$. Suppose there exists $M < \infty$ such that

$$\sup_{x \in \mathcal{X}} \frac{\tilde{f}(x)}{g(x)} \leq M.$$

$f(x) = g_1(x) + g_2(x)$

$p \int_x g_1(x) = p$

so, $\tilde{g}_1(x) = \frac{1}{p} g_1(x)$

$f(x) = p\tilde{g}_1(x) + (1-p)\tilde{g}_2(x)$

$g_2(x) \leq h(x) \ \forall x$

Consider the following accept-reject algorithm:

Step 1. Generate $Y \sim G$ and independently generate $U \sim U[0,1]$

Step 2. If $U \leq \dfrac{\tilde{f}(Y)}{Mg(Y)}$, then set $X = Y$

Step 3. Else, go to step 1. $U \leq \tilde{g}$

15 (a) *Prove* that the algorithm returns $X \sim F$. Provide <u>all</u> steps of the proof.

5 (b) What is the expected number of loops the algorithm takes before returning an output?

For $B \subseteq \mathcal{X}$, consider

$Pr(X \in B) = Pr(Y \in B \mid accept)$

$= Pr\left(Y \in B, U \leq \frac{\tilde{f}(Y)}{Mg(Y)}\right) \cdot \frac{1}{P(accept)}$

$= mM \cdot E\left[E\left[\mathbb{I}\left(Y \in B, U \leq \frac{\tilde{f}(Y)}{Mg(Y)}\right) \Big| Y\right]\right]$

$= mM \cdot E\left[\mathbb{I}(Y \in B)\, Pr\left(U \leq \frac{\tilde{f}(Y)}{Mg(Y)} \Big| Y\right)\right]$

$= mM \cdot E\left[\mathbb{I}(Y \in B) \cdot \frac{\tilde{f}(Y)}{Mg(Y)}\right]$

$= \frac{m\cancel{M}}{\cancel{M}} \int_{\mathcal{Y}} \mathbb{I}(y \in B) \cdot \frac{\tilde{f}(y)}{\cancel{g(y)}} \cancel{g(y)}\, dy$

$= \int_B m\tilde{f}(y)\, dy$

$= \int_B f(y)\, dy = F(B)$

$Pr(accept) = Pr\left(U \leq \frac{\tilde{f}(Y)}{Mg(Y)}\right)$

$P_r(accept) = E\left(\right.$

$= E\left[E\left[\mathbb{I}\left(U \leq \frac{\tilde{f}(Y)}{Mg(Y)}\right) \Big| Y\right]\right]$

$= E_g\left[Pr\left(U \leq \frac{\tilde{f}(Y)}{Mg(Y)} \Big| Y\right)\right]$

$= E_G\left[\frac{\tilde{f}(Y)}{Mg(Y)}\right] = \int_{\mathcal{Y}} \frac{\tilde{f}(y)}{Mg(y)} g(y)\, dy$

$\frac{1}{M}\left[\int_{\mathcal{X}} \tilde{f}(y)\, dy + \underbrace{\int_{\mathcal{Y} \setminus \mathcal{X}} \tilde{f}(y)\, dy}_{=0}\right]$

$(1-p)$

$= \frac{1}{mM} \int_{\mathcal{X}} f(y)\, dy$

$= \frac{1}{mM}$

Expected # loops:

loops $\sim$ Geomertic ($Pr(accept)$)

$\therefore$ Expected # loops is $\frac{1}{\frac{1}{mM}} = mM$

(-10) if proof not given only or (-12) reasoning is given.

(-4) or (-5) if using discrete random variable notation.

(-3) or (-4) if steps are incorrect or missing

For RoU: $f^{\frac{1}{2}}(x)$ $\quad xf^{\frac{1}{2}}(x)$ $\quad -\frac{1}{x^\alpha} \to \infty$ $\quad x = \left(\frac{-1}{\log U}\right)^{1/\alpha}$

3. (25 points) Consider the Fréchet distribution which for $\alpha > 0$ has cumulative distribution function and density function

$$F(x) = e^{-\frac{1}{x^\alpha}} \quad \text{and} \quad f(x) = \alpha x^{-\alpha-1} e^{-\frac{1}{x^\alpha}} \quad ; x > 0 .$$

(5) (a) Write the steps of the Inverse Transform method to sample from this distribution.

(20) (b) Write steps for the Ratio-of-Uniforms algorithm to sample from this distribution. Does the RoU method work for all α? If not, then for what values of α does it fail, and why?

(a) $F(x) = e^{-\frac{1}{x^\alpha}} \Rightarrow -\frac{1}{x^\alpha} = \log F(x) \Rightarrow x = \left(\frac{-1}{\log F(x)}\right)^{1/\alpha}$

Algo ① Draw $U \sim U(0,1)$ $\quad$ (-1) or (-2) for small errors.

② Set $X = \left(\frac{-1}{\log U}\right)^{1/\alpha}$

(b) $f(x) = \alpha x^{-\alpha-1} e^{-1/x^\alpha}$; $x > 0$; $\alpha > 0$.

We need to find $a = \sup_{x \in \mathcal{X}} \sqrt{f(x)}$; $b = \inf_{x \le 0} x\sqrt{f(x)}$; $c = \sup_{x > 0} x\sqrt{f(x)}$

$g(x) = \log\sqrt{f(x)} = \frac{1}{2}\log f(x) = \frac{1}{2}\left[\log\alpha - (\alpha+1)\log x - \frac{1}{x^\alpha}\right]$; taking derivative

$g'(x) = \frac{1}{2}\left[-\frac{(\alpha+1)}{x} + \frac{\alpha}{x^{\alpha+1}}\right] \overset{set}{=} 0 \Rightarrow -(\alpha+1) + \frac{\alpha}{x^\alpha} = 0 \Rightarrow x^\alpha = \frac{\alpha}{1+\alpha} \Rightarrow x_a = \left(\frac{\alpha}{\alpha+1}\right)^{1/\alpha}$

$\boxed{a = \sqrt{f\left(\left(\frac{\alpha}{1+\alpha}\right)^{1/\alpha}\right)}}$ $\quad$ (-2) for small mistakes each. (-4) for bigger mistakes.

$b = \inf_{x \le 0} x\sqrt{f(x)}$; $f(x) = 0$ for $x \le 0$, $\Rightarrow \boxed{b = 0}$

$\tilde{g}(x) = \log\left[x\sqrt{f(x)}\right] = \log x + \frac{\log\alpha - (\alpha+1)\log x}{2} - \frac{1}{2x^\alpha}$

$\tilde{g}'(x) = \frac{1}{x} - \frac{(\alpha+1)}{2x} + \frac{\alpha}{2x^{\alpha+1}} \overset{set}{=} 0 \Rightarrow -\frac{(\alpha-1)}{2} + \frac{\alpha}{2x^{\alpha+1}} = 0 \Rightarrow x_c = \left(\frac{\alpha}{\alpha-1}\right)^{1/\alpha}$

if $\alpha \geq 1$

$\boxed{c = \sqrt{x_c^2 f(x_c)}}$

Algorithm ① Draw $(U,V) \sim U[0,a] \times U[0,c]$ — (-3) or (-4) if algorithm is wrong.

② If $U \le \sqrt{f\left(\frac{V}{U}\right)}$, then set $X = \frac{V}{U}$

③ Else, go to step ①.

RoU doesn't work for $\alpha < 1$ since c doesn't exist as region C won't be bdd.

(-4) if not explained.

4. For $X \sim N(0,1)$, consider estimating $\theta = \mathrm{E}\left[e^{-X^4}\right]$ using simple importance sampling.

(a) (~~10~~ 5 points) Consider a standard Cauchy importance proposal with density

$$g(x) = \frac{1}{\pi}\frac{1}{1+x^2}.$$

Write the steps to obtain the simple importance sampling estimate of θ. Call this $\hat{\theta}_g$.

$$\theta = \int_{-\infty}^{\infty} e^{-x^4} \frac{1}{\sqrt{2\pi}} e^{-x^2/2}\, dx \qquad \text{where } h(x) = e^{-x^4}, \quad f(x) = \frac{1}{\sqrt{2\pi}} e^{-x^2/2}$$

$$\hat{\theta}_g = \frac{1}{N}\sum_{t=1}^{N} \frac{h(Z_t) f(Z_t)}{g(Z_t)} \qquad Z_1, Z_2, \ldots, Z_N \overset{iid}{\sim} \text{Cauchy}$$

(-1) if not iid

$$= \frac{1}{N}\sum_{t=1}^{N} \frac{\frac{1}{\sqrt{2\pi}} e^{-Z_t^4 - Z_t^2/2}}{\frac{1}{\pi}\frac{1}{(1+Z_t^2)}}$$

$$= \frac{1}{N}\sum_{t=1}^{N} \sqrt{\frac{\pi}{2}}\,(1+Z_t^2)\, e^{-Z_t^4 - Z_t^2/2}$$

$$h(x) = e^{-x^4}$$

$$f(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$$

$$\hat{\theta}_g = \frac{1}{N}\sum_{t=1}^{N} \frac{h(Z_t) f(Z_t)}{g(Z_t)}$$

$$= \frac{1}{N}\sum \frac{\frac{1}{\sqrt{2\pi}} e^{-}}{}$$

(b) i) 0.6291959

ii) $\frac{0.353516}{1000}$ (-3) if not divided by 1000.

$$\sigma_g^2 = \frac{1}{N} \mathrm{Var}\left(\frac{1}{N}\sum \_\_\right)$$

$$= \frac{1}{N^2} \sum \mathrm{Var}(\_)$$

$$\frac{1}{N}\mathrm{Var}(\ldots)$$

(c) (20 points) Recall that if $\int_{\mathcal{X}} |h(x)|f(x)dx \neq 0$, then for estimating $\theta = \int_{\mathcal{X}} h(x)f(x)dx$, the optimal proposal density for simple importance sampling is

$$g^*(x) = \frac{|h(x)|f(x)}{\int_{\mathcal{X}} |h(x)|f(x)dx}.$$

This proposal is difficult to use in practice since $\int_{\mathcal{X}} |h(x)|f(x)dx$ is often unknown. For $\theta = \mathrm{E}[e^{-X^4}]$, $X \sim N(0,1)$, answer the following:

10 (i) Use the algorithm in Question 2 of this exam to obtain samples from g^* setting $G = N(0,1)$ in the algorithm. Explain clearly what m, M, f, g are.

5 (ii) Why is the simple importance sampling estimator not practically usable for this proposal density?

5 (iii) Using the samples obtained in part (i), construct a weighted importance sampling estimator of θ.

i) Algo says. $f(x) = m\tilde{f}(x)$ where m is unknown. Here our target is the proposal density g^*, so that we can sample the values we want to propose.

$$g^*(x) = \frac{e^{-x^4}\frac{1}{\sqrt{2\pi}}e^{-x^2/2}}{\int_{-\infty}^{\infty} e^{-x^4}\frac{1}{\sqrt{2\pi}}e^{-x^2/2}} = m.\, e^{-x^4}\frac{1}{\sqrt{2\pi}}e^{-x^2/2} \;;\; \boxed{m = \frac{1}{\theta}}$$

In the notation of Q2: $\tilde{f}(x) = \frac{1}{\sqrt{2\pi}}e^{-x^4}e^{-x^2/2}$; $g(x) = \frac{1}{\sqrt{2\pi}}e^{-x^2/2}$.

$h(x) = e^{-x^4}$

$f(x) = \frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}$

$$\sup_{x\in\mathcal{X}} \frac{\tilde{f}(x)}{g(x)} = \sup e^{-x^4} = \boxed{1 = M}$$

(-6) if setup not explaind

Algorithm: ① Draw $U \sim U(0,1)$ and $Y \sim N(0,1)$ (independently)

② If $U \le \frac{\tilde{f}(Y)}{Mg(Y)} = e^{-Y^4}$, then $X = Y$

(-2) if algorithm not given

③ Else, go to ①.

$g^*(x) = \frac{1}{\theta}\frac{1}{\sqrt{2\pi}}e^{-\frac{x^2}{2}}e^{-x^4}$

ii) If we were to use simple importance sampling estimator; for $Z_1, Z_2, \dots, Z_N \overset{iid}{\sim} G^*$

$$\hat{\theta}_{g^*} = \frac{1}{N}\sum_{t=1}^{N}\frac{h(Z_t)f(Z_t)}{g(Z_t)} = \frac{1}{N}\sum_{t=1}^{N}\frac{e^{-Z_t^4}f(Z_t)}{e^{-Z_t^4}f(Z_t)/\theta} = \theta!$$

So the estimator is θ itself which is unknown.

(-2) or (-3) if not described

iii) Given $Z_1, Z_2, \dots, Z_n \overset{iid}{\sim} G^*$ which has density $g^*(x) = \frac{e^{-x^4}e^{-x^2/2}}{\theta} = \frac{1}{\theta}\tilde{g}(x)$

$$\hat{\theta}_W = \frac{\frac{1}{N}\sum_{t=1}^{N}\frac{h(Z_t)f(Z_t)}{\tilde{g}(Z_t)}}{\frac{1}{N}\sum_{t=1}^{N}\frac{f(Z_t)}{\tilde{g}(Z_t)}} = \frac{\sum_{t=1}^{N}\frac{e^{-Z_t^4}e^{-Z_t^2/2}\frac{1}{\sqrt{2\pi}}}{e^{-Z_t^4}e^{-Z_t^2/2}}}{\sum_{t=1}^{N}\frac{e^{-Z_t^2/2}\frac{1}{\sqrt{2\pi}}}{e^{-Z_t^4}e^{-Z_t^2/2}}} = \frac{N}{\sum_{t=1}^{N} e^{Z_t^4}}$$

$\sum e^{x_t^4}$

(-3) or (-4) if details not provided.