STAC63 Midterm 2024

Any results established in the class or in the Exercises, appropriately referenced, can be used as part of solving these questions.

Solutions

1(a). (10 marks) Suppose you need to approximate the following probability $P(X_1 > 0, \ldots, X_{10} > 0)$ where $\mathbf{X} = (X_1, \ldots, X_{10})' \sim N_{10}(\mu, \Sigma)$ where

$$\mu = (1, 2, 3, 1, 2, 3, 4, 5, 6, 7)'$$

and

$$\Sigma = \begin{pmatrix} 1.0 & 0.5 & \ldots & 0.5 \\ 0.5 & 1.0 & & \vdots \\ \vdots & & \ddots & 0.5 \\ 0.5 & \ldots & 0.5 & 1.0 \end{pmatrix}.$$

Discuss fully how you would carry out such an approximation including how you would assess the accuracy of the estimate. You do not have to specify code but make sure you describe each step necessary to implement your approximation in a language such as R (you can reference any of the capabilities of such a language).

Note that $\underset{\sim}{X} = \mu + \Sigma^{1/2} \underset{\sim}{Z}$ where $\underset{\sim}{Z} \sim N_k(\underset{\sim}{0}, I)$ and $\Sigma^{1/2}$ is the symmetric square root of Σ. So, using R we generate $\underset{\sim}{Z}_1, \ldots, \underset{\sim}{Z}_n \overset{iid}{\sim} N_k(\underset{\sim}{0}, I)$ and put $\underset{\sim}{X}_i = \mu + \Sigma^{1/2} \underset{\sim}{Z}_i$ for $i = 1, \ldots, n$. We then compute the average $\hat{p} = \frac{1}{n} \sum_{i=1}^{n} I_{(0,\infty)^k}(\underset{\sim}{X}_i)$ which records the proportion of times the generated values satisfy the inequalities where all the coordinates are positive. By the SLLN $\hat{p} \to p = P(X_1 > 0, \ldots, X_k > 0)$. Now $I_{(0,\infty)^k}(\underset{\sim}{X}) \sim$ Bernoulli(p) so $n\hat{p} \sim$ Binomial(n, p) and so by the CLT $\sqrt{n}(\hat{p} - p)/p(1-p) \xrightarrow{d} N(0,1)$ and since $\hat{p}(1-\hat{p}) \xrightarrow{wp1} p(1-p)$ Slutsky's Theorem says that $\sqrt{n}(\hat{p} - p)/\sqrt{\hat{p}(1-\hat{p})} \xrightarrow{d} N(0,1)$ and so the interval $[\hat{p} - 3\sqrt{\hat{p}(1-\hat{p})/n}, \hat{p} + 3\sqrt{\hat{p}(1-\hat{p})/n}]$ contains the true value of p with virtual certainty.

1(b). (15 marks) Suppose you want to estimate the integral $I = \int_{\mathbb{R}^k} h(x)\,dx$ and you will use importance sampling based on the importance sampler g. So, your estimate is of the form

$$I_n = \frac{1}{n}\sum_{i=1}^{n}\frac{h(x_i)}{g(x_i)}$$

where $x_1, \ldots, x_n \stackrel{iid}{\sim} g$. Use the Central Limit Theorem to determine a confidence interval for I based on I_n and state any conditions that need to be satisfied for the CLT to be valid.

Provided that the integral I is finite we have that $E(I_n) = I$. If, in addition, $\int \frac{h^2(x)}{g(x)}\,dx < \infty$ then $Var(I_n) = \frac{1}{n} Var_g\left(\frac{h(X)}{g(X)}\right) = \frac{\sigma_g^2}{n}$
$= \frac{1}{n}\left(E_g\left(\frac{h^2(X)}{g^2(X)}\right) - I^2\right) = \frac{1}{n}\left(\int \frac{h^2(x)}{g(x)}\,dx - I^2\right)$
and so $\sqrt{n}\,(I_n - I)/\sigma_g^2 \xrightarrow{d} N(0,1)$. Now we can estimate σ_g^2 by $\hat{\sigma}_g^2 = \frac{1}{n}\sum_{i=1}^{n}\frac{h^2(x_i)}{g^2(x_i)} - I_n^2$
and so by Slutsky's Thm we have $\sqrt{n}\,(I_n - I)/\hat{\sigma}_g \xrightarrow{d} N(0,1)$. Therefore, for large n $\left(I_n \pm \frac{3\hat{\sigma}_g}{\sqrt{n}}\right)$ contains the true value of I with virtual certainty.

2(a). (10 marks) Suppose $X_1, \ldots, X_n \stackrel{iid}{\sim} U(0,1)$, the uniform distribution on $(0,1)$. Determine the cdf of $X_{(n)} = \max\{X_1, \ldots, X_n\}$. Show how you could generate a value $X_{(n)}$ from its distribution based on generating a single value $U \sim U(0,1)$.

$F(x) = P(X_{(n)} \le x) = P(X_1 \le x, \ldots, X_n \le x)$

$= \prod_{i=1}^{n} P(X_i \le x)$ since the X_i are mutually statistically independent. Now when $X \sim U(0,1)$ we have $P(X \le x) = \begin{cases} 0, & x < 0 \\ x, & 0 \le x \le 1 \\ 1, & x > 1 \end{cases}$

and so $P(X_{(n)} \le x) = \begin{cases} 0, & x < 0 \\ x^n, & 0 \le x \le 1 \\ 1, & x > 1 \end{cases}$

since the X_i are identically distributed.

Therefore the quantile function associated with F is given by $F^{-1}(p) = p^{1/n}$ whenever $0 \le p \le 1$.

So we can use inversion to generate $X_{(n)}$ by generating $U \sim U(0,1)$ and putting $X_{(n)} = F^{-1}(U)$.

2(b). (10 marks) Suppose you want to generate a value from the probability distribution on $\mathbb{R}^1$ with density $f(x) = 3x^2/2 + 5x^4/2$ when $x \in (0,1)$ and $f(x) = 0$ elsewhere. Determine a rejection algorithm for this.

Since $0 < x^k < 1$ whenever $0 < x < 1$ and $k \geq 1$ we have that $\frac{3x^2}{2} + \frac{5x^4}{2} < \frac{3}{2} + \frac{5}{2} = 4g(x)$ where g is the pdf of the $U(0,1)$ distribution. So the rejection algorithm based on g is as follows.

1. generate $X \sim U(0,1)$ and an independent $U \sim U(0,1)$
2. if $4U \leq f(X) = \frac{3X^2}{2} + \frac{5X^4}{2}$ then return X
   else go to 1.

3(a). (10 marks) Suppose that $\mathbf{X}_n \xrightarrow{d} \mathbf{X} \sim N_k(\mathbf{0}, \Sigma)$ and $\mathbf{Y}_n = \mathbf{a} + \mathbf{B}\mathbf{X}_n$ for fixed $\mathbf{a} \in \mathbb{R}^l, \mathbf{B} \in \mathbb{R}^{l \times k}$. Determine the limiting distribution of $\mathbf{Y}_n$ and justify this.

The differential of the transformation $g(\underset{\sim}{x}) = \underset{\sim}{a} + B\underset{\sim}{x}$ is given by $G(\underset{\sim}{x}) = B$ and so $G(\underset{\sim}{\mu}) = G(\underset{\sim}{0}) = B$. Therefore by the delta theorem we have

$$g(\underset{\sim}{X}_n) - \underset{\sim}{a} \xrightarrow{d} G(\underset{\sim}{0})\underset{\sim}{X} \sim N_l(G(\underset{\sim}{0})\underset{\sim}{0}, G(\underset{\sim}{0})\Sigma G^t(\underset{\sim}{0}))$$
$$= N_l(\underset{\sim}{0}, B\Sigma B^t)$$

and so $g(\underset{\sim}{X}_n) \xrightarrow{d} N_l(\underset{\sim}{a}, B\Sigma B^t)$.

Or you could use the Continuous Mapping Theorem since g is a continuous function.

3(b). (10 marks) Suppose that $X_1, \ldots, X_n \stackrel{iid}{\sim} N(0,1)$. Determine an approximation to the asymptotic distribution of

$$Y_n = \frac{1}{n}\sum_{i=1}^{n} X_i^3$$

and recall that the k-th moment of the $N(0,1)$ is given by $\mu_k = 0$ when k is odd and $\mu_k = (k-1) \times (k-3) \times \cdots \times 3 \times 1$ when k is even.

If $X \sim N(0,1)$ then $E(X^3) = 0$ and $E(X^6) = 5 \cdot 3 = 15$

So $X_1^3, \ldots, X_n^3$ is a sample from a distribution

(5) with mean 0 and variance 15. Therefore,

by the CLT, $\sqrt{n}\left(\frac{1}{n}\sum_{i=1}^{n} X_i^3 - 0\right)/\sqrt{15} \xrightarrow{d} N(0,1)$

and we can approximate the distribution by

$\frac{1}{n}\sum_{i=1}^{n} X_i^3 \approx N(0, \frac{15}{n})$.

Now

3(c). (10 marks) Repeat 3(b) but for the statistic

$$W_n = \bar{X}^3 = \left(\frac{1}{n}\sum_{i=1}^{n} X_i\right)^3.$$

Now $\sqrt{n}\,\bar{X} \xrightarrow{d} N(0,1)$ and by the delta theorem $g(x)=x^3$ has $g'(x)=3x^2$ with $g'(0)=0$, then $\sqrt{n}\,\bar{X}^3 \xrightarrow{d} N(0,0)$ which is degenerate at 0 and this doesn't help to approximate the distribution of $\bar{X}^3$. We have, however, that

$x^3 = g(x) = g(0) + \frac{g'(0)}{1!}x + \frac{g''(0)}{2!}x^2 + \frac{g'''(0)}{3!}x^3$

where $g(0) = g'(0) = g''(0) = 0$ while $g'''(0) = 6$.

Therefore, by the Continuity Theorem for Convergence in Distribution $(\sqrt{n}\,\bar{X})^3 \xrightarrow{d} Z^3$ where $Z \sim N(0,1)$. Note that by the Change of Variable theorem $Y = Z^3$ has density given by $\frac{1}{3}y^{-2/3}\varphi(y^{1/3})$.

4(a) (5 marks) Suppose a sequence of random variables $X_0, X_1, X_2 \ldots$ is generated as follows: $X_0 = 0$ and $X_i = X_{i-1} + Z_i$ where $Z_i \sim \text{Poisson}(\lambda)$. Is this a time homogeneous Markov chain and, if so, what is the state space S, the initial distribution ν and the transition probability p_{ij}? Justify all your claims.

Note that $X_n = \sum_{i=1}^{n} Z_i$ where $Z_1, \ldots, Z_n \overset{iid}{\sim} \text{Poisson}(\lambda)$

Therefore $\{X_n : n \in \mathbb{N}_0\}$ is a random walk and it was proved in class that this is a time homogeneous MC. We have that $p_{ij} = P(X_2 = j \mid X_1 = i)$ by TH

$= P(Z_1 + Z_2 = j \mid X_1 = i) = P(Z_2 = j - i)$

$$= \begin{cases} \dfrac{\lambda^{j-i}}{(j-i)!} e^{-\lambda} & j \geq i \\ 0 & j < i \end{cases}$$

The state space is $S = \mathbb{N}_0$ and the initial distribution is $\nu_0 = 1$.

4(b) (10 marks) Consider a MC with state space $S = \{1, 2, 3, 4\}$, initial probability distribution

$$\nu = (1/2, 1/6, 1/6, 1/6)$$

and transition probabilities

$$P = \begin{pmatrix} 1/2 & 1/4 & 1/4 & 0 \\ 0 & 2/3 & 1/3 & 0 \\ 0 & 1/5 & 4/5 & 0 \\ 0 & 0 & 0 & 1 \end{pmatrix}.$$

Compute $P(X_1 = 1, X_2 = 3, X_4 = 2)$.

$$P(X_1=1, X_2=3, X_4=2) = P(X_4=2 \mid X_1=1, X_2=3)\, P(X_1=1, X_2=3)$$

$$= P(X_4=2 \mid X_2=3)\, P(X_2=3 \mid X_1=1)\, P(X_1=1)$$

$$= \left(\sum_{i=1}^{4} P(X_3=i, X_4=2 \mid X_2=3)\right) \left(\tfrac{1}{4}\right) \left(\tfrac{1}{2}\right)$$

$$= \left(\tfrac{1}{4}\right)\left(\tfrac{1}{2}\right)\left(\sum_{i=1}^{4} P(X_4=2 \mid X_2=3, X_3=i)\, P(X_3=i \mid X_2=3)\right)$$

$$= \left(\tfrac{1}{4}\right)\left(\tfrac{1}{2}\right) \sum_{i=1}^{4} P(X_4=2 \mid X_3=i)\, P(X_3=i \mid X_2=3)$$

$$\overset{TH}{=} \left(\tfrac{1}{4}\right)\left(\tfrac{1}{2}\right) \sum_{i=1}^{4} P(X_2=2 \mid X_1=i)\, P(X_2=i \mid X_1=3)$$

$$= \left(\tfrac{1}{4}\right)\left(\tfrac{1}{2}\right)\left(\left(\tfrac{1}{4}\right)0 + \left(\tfrac{2}{3}\right)\left(\tfrac{1}{5}\right) + \left(\tfrac{1}{5}\right)\left(\tfrac{4}{5}\right) + 0 \cdot 0\right)$$

$$= \left(\tfrac{1}{4}\right)\left(\tfrac{1}{2}\right)\left(\left(\tfrac{2}{3}\right)\left(\tfrac{1}{5}\right) + \left(\tfrac{1}{5}\right)\left(\tfrac{4}{5}\right)\right) = \frac{11}{600}$$

4(c) (5 marks) Is the chain in 4(b) irreducible? Identify all the recurrent and transient states,

The chain is not irreducible because $1 \nrightarrow 4$.

We have that 4 is recurrent as are 2 and 3 but 1 is transient since $f_{11} = 0$ while $f_{22} = f_{33} = f_{44} = 1$, 4(d) (5 marks) Prove that for a time homogeneous Markov chain with state space $S = \mathbb{N}$, we have that

$$E_i(N_i) = \sum_{k=1}^{\infty} p_{ii}^{(k)} = \sum_{k=1}^{\infty} f_{ii}^{k}.$$

Hint: Use the result proved in class that $E_i(N_i) = \sum_{k=1}^{\infty} p_{ii}^{(k)}$ and the Choquet formula for expectation of nonnegative random variables.

By Prop. III.6 $E_i(N_i) = \sum_{k=1}^{\infty} p_{ii}^{(k)}$.

But by Choquet's Formula we also have $E_i(N_i) = \sum_{k=1}^{\infty} P_i(N_i \geq k)$ and we worked out in class that $P_i(N_i \geq k) = f_{ii}^{k}$.