541B Final Solutions[1]

1. Question 1

Let U be a random variable uniform on the interval $(0,1)$.

- Describe in detail a function $g\colon (0,1)\to\{1,2,3\}$ such that $g(U)$ is uniformly distributed in $\{1,2,3\}$. Prove your assertion.
- Describe in detail a function $h\colon (0,1)\to\mathbf{R}$ such that $h(U)$ has a standard Gaussian distribution (i.e. $h(U)$ has PDF given by $f(x)=\frac{1}{\sqrt{2\pi}}e^{-x^2/2}$, $\forall\, x\in\mathbf{R}$.)
  Prove your assertion.

Solution. Define $g:=1$ if $U<1/3$, $g:=2$ if $1/3\leq U<2/3$ and $g:=3$ if $U\geq 2/3$. Then $\mathbf{P}(g(U)=1)=(1/3)-0=1/3$, $\mathbf{P}(g(U)=2)=2/3-1/3=1/3$ and $\mathbf{P}(g(U)=3)=1-2/3=1/3$, all by definition of g.

Let $\Phi(t):=\int_{-\infty}^{t}e^{-x^2/2}dx/\sqrt{2\pi}\ \forall\, t\in\mathbf{R}$. Then $\Phi\colon\mathbf{R}\to(0,1)$ and Φ is invertible since it is strictly increasing. Define $h:=\Phi^{-1}$. Then for any $t\in\mathbf{R}$

$$\mathbf{P}(h(U)\leq t)=\mathbf{P}(\Phi^{-1}(U)\leq t)=\mathbf{P}(U\leq\Phi(t))=\Phi(t).$$

The last equality follows by definition of U. We therefore conclude that $h(U)$ has PDF f.

2. Question 2

Suppose $\Theta=\{\theta_0,\theta_1\}$, $\Theta_0=\{\theta_0\}$, $\Theta_1=\{\theta_1\}$. Let H_0 be the hypothesis $\{\theta=\theta_0\}$ and let H_1 be the hypothesis $\{\theta=\theta_1\}$. Let $\{f_{\theta_0},f_{\theta_1}\}$ be two multivariable probability densities on $\mathbf{R}^n$. Fix $k\geq 0$. Define a **likelihood ratio test** $\phi\colon\mathbf{R}^n\to[0,1]$ to be

$$\phi(x):=\begin{cases}1 & \text{, if } f_{\theta_1}(x)>kf_{\theta_0}(x)\\ 0 & \text{, if } f_{\theta_1}(x)<kf_{\theta_0}(x)\\ \text{(unspecified)} & \text{, if } f_{\theta_1}(x)=kf_{\theta_0}(x).\end{cases}\qquad (*)$$

Define

$$\alpha:=\sup_{\theta\in\Theta_0}\beta(\theta)=\beta(\theta_0)=\mathbf{E}_{\theta_0}\phi(X).\qquad (**)$$

Let $\mathcal{T}$ be the class of all randomized hypothesis tests with significance level at most α.

Prove the following:

Any randomized hypothesis test satisfying $(*)$ is a UMP class $\mathcal{T}$ test.

Solution.

As we already noted in $(**)$, Θ_0 consists of a single point, so the supremum appearing in $(**)$ is just $\beta(\theta_0)$, and we will repeatedly use this fact below without further mention.

Let $\beta(\theta)$ be the power function of the test corresponding to ϕ. Let ϕ' another test in $\mathcal{T}$, and let $\beta'(\theta)$ be the power function of this test. By definition of ϕ, we have

$$[\phi(x)-\phi'(x)][f_{\theta_1}(x)-kf_{\theta_0}(x)]\geq 0,\qquad\forall\, x\in\mathbf{R}^n.$$

Therefore,

$$0\leq\int_{\mathbf{R}^n}[\phi(x)-\phi'(x)][f_{\theta_1}(x)-kf_{\theta_0}(x)]dx=\beta(\theta_1)-\beta'(\theta_1)-k[\beta(\theta_0)-\beta'(\theta_0)].\qquad (***)$$

[1]December 10, 2025,

Since ϕ has significance level α and ϕ' has significance level at most α, we have $\beta(\theta_0) - \beta'(\theta_0) \geq 0$. So, $k \geq 0$ and $(***)$ imply that $\beta(\theta_1) - \beta'(\theta_1) \geq 0$. That is, the ϕ test is UMP class $\mathcal{T}$.

3. Question 3

Suppose X is a binomial distributed random variable with parameters 2 and $\theta \in \{1/4, 3/4\}$. (That is, X is the number of heads that result from flipping two coins, where each coin has probability θ of landing heads.)
We want to test the hypothesis H_0 that $\theta = 1/4$ versus the hypothesis H_1 that $\theta = 3/4$.

Let $\mathcal{T}$ be the set of hypothesis tests with significance level at most 1/80.
(Recall that the significance level of a hypothesis test $\phi\colon \mathbf{R} \to [0,1]$ is $\sup_{\theta\in\Theta_0} \mathbf{E}_\theta \phi(X)$.)

You may take it as given that the hypothesis test

$$\phi(x) := \begin{cases} 0 & ,\ \text{if}\, x \neq 2 \\ 1/5 & ,\ \text{if}\, x = 2. \end{cases}$$

is UMP class $\mathcal{T}$.

Is ϕ unique? That is, is there another hypothesis test ψ with $\psi \neq \phi$ such that ψ is UMP class $\mathcal{T}$? Explain in detail.

Solution. The Neyman-Pearson Lemma says that, since ϕ is UMP class $\mathcal{T}$, ϕ is uniquely defined, up to changes on a set $D \subseteq \mathbf{R}$ where $\mathbf{P}_{\theta_0}(X \in D) = \mathbf{P}_{\theta_1}(X \in D) = 0$. For each $i \in \{0,1,2\}$ and for each $\theta \in \{\theta_0, \theta_1\}$, we have $\mathbf{P}_\theta(X = i) > 0$. So, ϕ is uniquely defined on the set $\{0,1,2\}$. However, if $D = \mathbf{R} \setminus \{0,1,2\}$, then $\mathbf{P}_\theta(X \in D) = 0$ for each $\theta \in \{\theta_0, \theta_1\}$. So, ϕ is not uniquely defined on the set D.

4. Question 4

- Give an example of a Markov Chain that is not reversible. Prove your assertion.
- Give an example of a Markov Chain where every state has period 3. Prove your assertion.

Solution.
Consider

$$P := \begin{pmatrix} 0 & 1 & 0 \\ 0 & 0 & 1 \\ 1 & 0 & 0 \end{pmatrix}$$

Suppose π is stationary for P, i.e. $\pi = \pi P$. Denote $\pi = (a,b,c)$ with $0 \leq a,b,c \leq 1$, $a+b+c = 1$. Then $\pi = \pi P$ says $a = c$, $a = b$ and $b = c$, i.e. the unique stationary distribution for P is $\pi = (1/3, 1/3, 1/3)$. However, π does not satisfy the detailed balance condition since, if we denote $\Omega = \{1,2,3\}$, we have

$$\pi(1)P(1,2) = (1/3)(1) \neq 0 = (1/3)\cdot 0 = \pi(2)P(2,1)$$

If P were reversible, then P would have a stationary distribution satisfying the detailed balance condition. However, we just showed the unique stationary distribution for P does not satisfy the detailed balance condition. We conclude that P is not reversible.

5. Question 5

Let $f\colon [0,5]\to[0,3]$ be the PDF of a real-valued random variable X with maximum value $m:=3$. Let $(X_1,Y_1),(X_2,Y_2),\ldots$ be i.i.d random variables uniformly distributed in the rectangle $[0,5]\times[0,3]$.

Using accept/reject sampling, describe how to sample a random variable Z such that the PDF of Z is f.

Prove that the random variable Z has PDF f.

Solution. Define $I:=\inf\{n\geq 1\colon Y_n\leq f(X_n)\}$. Output $Z:=X_I$. We show that Z has PDF f.

Let $z\in\mathbf{R}$. Define $\varepsilon:=\frac{1}{m(b-a)}\int_a^b f(x)dx=\frac{1}{m(b-a)}$. Then

$$\begin{aligned}\mathbf{P}(Z\leq z)&=\sum_{i=1}^{\infty}\mathbf{P}(Z\leq z,I=i)=\sum_{i=1}^{\infty}\mathbf{P}(X_i\leq z,\,Y_i\leq f(X_i))\prod_{j=1}^{i-1}\mathbf{P}(Y_j>f(X_j))\\&=\sum_{i=1}^{\infty}\frac{\int_{-\infty}^{z}f(x)dx}{m(b-a)}(1-\varepsilon)^{i-1}=\int_{-\infty}^{z}f(x)dx\sum_{i=1}^{\infty}\varepsilon(1-\varepsilon)^{i-1}=\int_{-\infty}^{z}f(x)dx.\end{aligned}$$

6. Question 6

Let $G=(V,E)$ be a finite graph. (So V is a finite vertex set, and $E\subseteq\{\{x,y\}\colon x,y\in V,\,x\neq y\}$.) Let A denote the set of all elements of $\xi\in\{0,1\}^V$ such that, if $\{v,w\}\in E$ then $\xi(v),\xi(w)$ are not both equal to 1. The **hard core model** μ is a probability measure on the set $\{0,1\}^V$ that is uniform on the set A. That is, $\mu(\xi)=1/|A|$ for all $\xi\in A$, and $\mu(\xi)=0$ for all $\xi\in\{0,1\}^V\setminus A$.

Define a Markov chain $X_0,X_1,\ldots$ with state space A as follows. Initialize X_0 to be the zero function on V. For any $n\geq 0$, we will define X_{n+1} using X_n. For each integer $n\geq 0$, repeat the following procedure.

- Select one $v\in V$ uniformly at random.
- Let Y_n be uniformly distributed in $\{0,1\}$ and independent of all previously defined random variables.
- If $Y_n=1$, and if all vertices $w\in V$ adjacent to v satisfy $X_n(w)=0$, then set $X_{n+1}(v):=1$. Otherwise, set $X_{n+1}(v):=0$.
- For all $w\in V$ with $w\neq v$, define $X_{n+1}(w):=X_n(w)$.

You may take it as given that this stochastic process $X_0,X_1,\ldots$ is a Markov Chain with state space A that is irreducible and aperiodic.

Show: $X_0,X_1,\ldots$ has unique stationary distribution μ.

Solution. It suffices to show that μ satisfies the detailed balance condition. We know the Markov chain is irreducible, so we know the stationary distribution exists and is unique. If we show that μ satisfies the detailed balance condition, then μ is the (unique) stationary distribution for the Markov chain, by a proposition in the notes. So let us verify the detailed balance condition holds. That is, it remains to show that

$$\mu(\xi)P(\xi,\zeta)=\mu(\zeta)P(\zeta,\xi),\qquad\forall\,\xi,\zeta\in A,\qquad(*)$$

where P is the transition matrix of the Markov chain P. Let d be the number of $v\in V$ such that $\xi(v)\neq\zeta(v)$. If $d=0$, then $\xi=\zeta$, so both sides of $(*)$ are equal. If $d\geq 2$, then $P(\xi,\zeta) = P(\zeta,\xi) = 0$ by definition of the Markov chain, so both sides of $(*)$ are zero. To prove $(*)$ holds, it therefore remains to consider the case $d = 1$. If $d = 1$, then there exists exactly one vertex $v \in V$ such that $\xi(v) \neq \zeta(z)$, i.e. $\xi(w) = \zeta(w)$ for all other $w \in V$. Without loss of generality, assume $\xi(v) = 1$. Since $\xi \in A$, each neighbor w of v satisfies $\xi(w) = 0$, so that $\mu(\xi) = \mu(\zeta) > 0$ (recalling $\xi, \zeta \in A$), and $P(\xi,\zeta) = P(\zeta,\xi)$ since $\xi(w) = \zeta(w) = 0$ for all neighbors w of v, so $(*)$ holds.

7. Question 7

Let $X_0, X_1, \ldots, Y_0, Y_1, \ldots$ be a Hidden Markov Model.

Prove: for any $y_0, \ldots, y_{n-1} \in T$, and for any $x, z \in \Omega$,

$$\mathbf{P}(X_n = x \,|\, X_{n-1} = z, Y_0 = y_0, \ldots, Y_{n-1} = y_{n-1}) = P(z,x).$$

Solution. We use the definition of conditional probability, then the definition of an HMM (for n and $n-1$ separately), then sum over y and use that Q is a stochastic matrix,

$$\begin{aligned}
&\mathbf{P}(X_n = x \,|\, X_{n-1} = z, Y_0 = y_0, \ldots, Y_{n-1} = y_{n-1}) \\
&= \frac{\mathbf{P}(X_n = x, X_{n-1} = z, Y_0 = y_0, \ldots, Y_{n-1} = y_{n-1})}{\mathbf{P}(X_{n-1} = z, Y_0 = y_0, \ldots, Y_{n-1} = y_{n-1})} \\
&= \frac{\sum_{x_0,\ldots,x_{n-2}\in\Omega, y\in T} \mu(x_0)Q(x_0,y_0)\prod_{i=1}^{n-2} Q(x_i,y_i)P(x_{i-1},x_i)\cdot Q(z,y_{n-1})P(x_{n-2},z)Q(x,y)P(z,x)}{\sum_{x_0,\ldots,x_{n-2}\in\Omega} \mu(x_0)Q(x_0,y_0)\prod_{i=1}^{n-2} Q(x_i,y_i)P(x_{i-1},x_i)\cdot Q(z,y_{n-1})P(x_{n-2},z)} \\
&= \frac{\sum_{x_0,\ldots,x_{n-2}\in\Omega} \mu(x_0)Q(x_0,y_0)\prod_{i=1}^{n-2} Q(x_i,y_i)P(x_{i-1},x_i)\cdot Q(z,y_{n-1})P(x_{n-2},z)P(z,x)}{\sum_{x_0,\ldots,x_{n-2}\in\Omega} \mu(x_0)Q(x_0,y_0)\prod_{i=1}^{n-2} Q(x_i,y_i)P(x_{i-1},x_i)\cdot Q(z,y_{n-1})P(x_{n-2},z)} \\
&= P(z,x).
\end{aligned}$$

8. Question 8

Let $X_0, X_1, \ldots, Y_0, Y_1, \ldots$ be a Hidden Markov Model.

Fix $y_0, \ldots, y_n \in T$.

For all $x \in \Omega$, define $v_0(x) := \mathbf{P}(X_0 = x, Y_0 = y_0) = \mu(x)Q(x,y_0)$.

For any $1 \leq j \leq n$, define iteratively

$$v_j(x) = \Big(\max_{z\in\Omega} v_{j-1}(z)P(z,x)\Big)\cdot Q(x,y_j), \qquad \forall\, x \in \Omega.$$

Define also $w_0(x) := v_0(x)$ for all $x \in \Omega$, and for any $1 \leq j \leq n$, define

$$w_j(x) := \max_{x_0,\ldots,x_{j-1}\in\Omega} \mathbf{P}(X_0 = x_0, \ldots, X_{j-1} = x_{j-1}, X_j = x, Y_0 = y_0, \ldots, Y_j = y_j).$$

Prove by induction that: for all $1 \leq j \leq n$, and for all $x \in \Omega$,

$$w_j(x) = v_j(x).$$

Solution. The base case $j = 0$ follows by definition of v_0, w_0. We then prove the inductive step. Suppose the assertion holds for all $x \in \Omega$ and for all $1 \leq j \leq n-1$. We then consider the case $j = n$. By the HMM definition,

$$\begin{aligned}
&\mathbf{P}(X_0 = x_0, \ldots, X_n = x_n, Y_0 = y_0, \ldots, Y_n = y_n) \\
&= \mu(x_0) Q(x_0, y_0) \prod_{i=1}^{n} Q(x_i, y_i) P(x_{i-1}, x_i) \\
&= \mathbf{P}(X_0 = x_0, \ldots, X_j = x_{n-1}, Y_0 = y_0, \ldots, Y_{n-1} = y_{n-1}) \cdot Q(x_n, y_n) P(x_{n-1}, x_n)
\end{aligned}$$

Taking the maximum over $x_0, \ldots, x_{n-1} \in \Omega$ and applying the inductive hypothesis, we get

$$\begin{aligned}
&w_n(x_n) \\
&= \max_{x_{n-1} \in \Omega} \max_{x_0, \ldots, x_{n-2} \in \Omega} \mathbf{P}(X_0 = x_0, \ldots, X_j = x_{n-1}, Y_0 = y_0, \ldots, Y_{n-1} = y_{n-1}) Q(x_n, y_n) P(x_{n-1}, x_n) \\
&= \max_{x_{n-1} \in \Omega} w_{n-1}(x_{n-1}) \cdot Q(x_n, y_n) P(x_{n-1}, x_n) \\
&= \Big(\max_{x_{n-1} \in \Omega} v_{n-1}(x_{n-1}) P(x_{n-1}, x_n) \Big) Q(x_n, y_n).
\end{aligned}$$

The right side is $v_n(x_n)$, by definition of v_n. Having completed the inductive step, the proof is done.