Roll No.: ____________________

CS781 In-class Evaluation 2 (Autumn 2025)

Max marks: 10 **Duration: 10 mins**

- *You are required to answer each question only in the sheet provided to you.*
- *Only material written in the sheet will be graded.*
- *The exam is open book and notes. However, you are not allowed to consult others for your answers.*
- *Be brief, complete and stick to what has been asked.*
- *Unless asked for explicitly, you may cite results/proofs covered in class without reproducing them.*
- ***If you need to make any assumptions, state them clearly.***
- ***Do not copy solutions from others. Penalty for offenders: FR grade.***

The interaction of an environment with a reinforcement learning agent that has learnt a deterministic strategy, can often be described as a Markov chain. Consider such a Markov chain with 5 states $\{s_1, s_2, s_3, s_4, s_5\}$, but with an unknown transition matrix M. Suppose the interaction modeled by the Markov Chain is being monitored by a frequentist monitor. Every time the Markov chain reaches a state s_i, the monitor observes an event i.

Now consider the following sequence of observations of the monitor, where repeats are denoted in parentheses (e.g., 1(5) 2 3(2) denotes the sequence 1 1 1 1 1 2 3 3)

1 2(3) 1(2) 2(2) 3(2) 4 2 3(3) 2(3) 5 4(3) 2 5(2) 1(3) 5(2) 2 3 4(3) 2 5(3) 1(4) 2 3(3) 2(2) 3

Let $M_{i,j}$ denote the probability of the Markov chain transitioning to state s_j from state s_i.

For each of the following transition probability matrix entries or functions of such entries, give the 90% confidence interval from the sequence given above. Thus, when estimating $M_{i,j}$, you must give an interval $[l_{i,j}, u_{i,j}]$ such that $\mathsf{Pr}(l_{i,j} \leq M_{i,j} \leq u_{i,j}) \geq 0.9$

1. *[3 marks]* $M_{4,4}$
2. *[3 marks]* $M_{1,1} + M_{2,3}$
3. *[4 marks]* $M_{2,3} \times M_{2,5}$.

By Hoeffding's ineq., we know that.

$$Pr\left[\,|\mu - \hat{\mu}(\vec{x})| \leq \varepsilon\,\right] \geq 1 - 2.\exp\left(-\frac{2\varepsilon^2.n}{(b-a)^2}\right)$$

∴ If we want a confidence interval of $1-\delta$, where $0 < \delta \leq 1$, we need $\delta \geq 2\exp\left(-\frac{2\varepsilon^2.n}{(b-a)^2}\right)$.

$$\Rightarrow \varepsilon \geq (b-a)\sqrt{\ln\left(\frac{2}{\delta}\right) \times \frac{1}{2n}}$$

For notational convenience, let us use $E(b-a, n, \delta)$ as a shorthand for

$$(b-a).\sqrt{\frac{1}{2n} \times \ln\left(\frac{2}{\delta}\right)}$$

① Define $\vec{Y}_{4,4}$ as a sequence of i.i.d. random vars with mean $M_{4,4}$ as described in the "Monitoring Algorithmic Fairness" paper.

∴ we have the foll. sequence of observations (zooming on the relevant obsvns)

---- 3 3 4 2 ... 5 4 4 4 2 ---- 3 4 4 4 2 -----

$\vec{Y}_{4,4}$: 0 1 1 0 1 1 0

∴ $\hat{\mu}(\vec{Y}_{4,4}) = \frac{4}{7}$

Lower bound = a = 0 ; upper bound = b = 1 ∴ b − a = 1

for 90% confidence, we need $\delta \le 0.1$, and hence the reqd. interval is $\left[\frac{4}{7} - E(1,7,0.1), \ \frac{4}{7} + E(1,7,0.1)\right]$

② Define $\vec{Y}_{1,1}$, $\vec{Y}_{2,3}$ & $\vec{Y}_{1,1+2,3}$ as the relevant sequences of i.i.d. random vars with means $M_{1,1}$, $M_{2,3}$ & $M_{1,1} + M_{2,3}$ respectively.

We have the foll. seq. of relevant observations:

1 2 .... 1 1 2 ---- 5 5 1 1 1 5 ---- 5 5 5 1 1 1 1 2 -----

$\vec{Y}_{1,1}$: 0 1 0 1 1 0 1 1 1 0

1 2 2 2 1 1 2 2 3 ... 4 2 3 3 3 2 2 2 5 ---- 4 2 5 ---- 5 2 3 --- 4 2 5 .... 1 2 3 3 3 2 2 3

$\vec{Y}_{2,3}$: 0 0 0 0 1 1 0 0 0 0 1 0 1 0 1

∴ $\vec{Y}_{1,1+2,3}$: obtained by considering 10 elements of $\vec{Y}_{1,1}$ and $\vec{Y}_{2,3}$.

$\vec{Y}_{1,1}$: 0 1 0 1 1 0 1 1 1 0

$\vec{Y}_{2,3}$: 0 0 0 0 1 1 0 0 0 0

$\vec{Y}_{1,1+2,3}$: 0 1 0 1 2 1 1 1 1 0

∴ $\hat{\mu}(\vec{Y}_{1,1+2,3}) = \frac{8}{10}$

Also, we have lower bound = a = 0; upper bound = b = 2 ∴ b − a = 2

For 90% confidence, reqd. interval = $\left[\frac{8}{10} - E(2,10,0.1), \ \frac{8}{10} + E(2,10,0.1)\right]$

(3) For this part, note that we must use temporal shifting to ensure independence of random variables in the sequence $\vec{Y}_{2,3}$ and those in the sequence $\vec{Y}_{2,5}$

We have the foll. seq. of relevant observations:

1 222 11 22 3 ... 42333 2225 .... 425 .... 523 ...425 ....1 2333223

$\vec{Y}_{2,3}$: $\underline{0}$ 0 $\underline{0}$ 0 $\underline{1}$ 1 $\underline{0}$ 0 $\underline{0}$ 0 $\underline{1}$ 0 $\underline{1}$ 0 $\underline{1}$

$\vec{Y}_{2,5}$: 0 $\underline{0}$ 0 $\underline{0}$ 0 $\underline{0}$ 0 $\underline{0}$ 1 $\underline{1}$ 0 $\underline{1}$ 0 $\underline{0}$ 0

The underlined elements represent selection after temporal shift.

$\therefore$ Temporally shifted sequences of indep. random vars:

$\vec{Y}_{23}'$: 0 0 1 0 0 1 1 1

$\vec{Y}_{2,5}'$: 0 0 0 0 1 1 0

$\vec{Y}_{2,3\times 2,5}$: 0 0 0 0 0 1 0

$\therefore \hat{\mu}\left(\vec{Y}_{2,3\times 2,5}\right) = \dfrac{1}{7}$. Also, lower bound $= a = 0$

upper bound $= b = 1$

$\therefore b - a = 1$

For 90% confidence, reqd. interval $= \left[\frac{1}{7} - E(1, 7, 0.1), \frac{1}{7} + E(1, 7, 0.1)\right]$