Probability and Computation: Homework Assessment Solutions

Question 1. *Let A, B be independent uniformly random subsets of $[n] := \{1, \dots, n\}$.*

(i) Find $\mathbf{P}[\, A \subseteq B \,]$.

(ii) What is the distribution of $|A|$?

(iii) How about the distribution of $|A \backslash B|$?

(iv) How can you solve part (i) using part (iii)?

Solution: *For (i) we need that if $i \in A$ then $i \in B$ for all $i \in [n]$. For a specific i we have*

$$\begin{aligned}\mathbf{P}[\, i \in A \implies i \in B \,] &= \mathbf{P}\Big[\, i \in A \implies i \in B \Big| i \in B \,\Big] \mathbf{P}[\, i \in B \,] + \mathbf{P}\Big[\, i \in A \implies i \in B \Big| i \notin B \,\Big] \mathbf{P}[\, i \notin B \,] \\ &= (\mathbf{P}[\, i \notin A \,] + \mathbf{P}[\, i \in A \,])\, \mathbf{P}[\, i \in B \,] + (\mathbf{P}[\, i \notin A \,])\, \mathbf{P}[\, i \notin B \,] \\ &= 1 \cdot (1/2) + (1/2)(1/2) = 3/4.\end{aligned}$$

Thus $\mathbf{P}[\, A \subseteq B \,] = \prod_{i=1}^{n} \mathbf{P}[\, i \in A \implies i \in B \,] = (3/4)^n$.
For (ii) all elemets of $[n]$ are in A with probability $1/2$ independent of one and other so $|A|$ is distributed $Bin\,(n, 1/2)$.
For (iii) all elemets of $[n]$ are in $A \backslash B$ with probability $1/4$ independent of one and other so $|A \backslash B|$ is distributed $Bin\,(n, 1/4)$.
For (iv) observe $\{A \subseteq B\} = \{A \backslash B = \emptyset\}$. Thus $\mathbf{P}[\, A \subseteq B \,] = \mathbf{P}[\, Bin\,(n, 1/4) = 0 \,] = (1 - 1/4)^n$.

Question 2. *Let $X \geq 0$ be an integer valued random variable.*

(i) Show that $\mathbf{E}[\, X \,] = \sum_{i=0}^{\infty} \mathbf{P}[\, X > i \,]$.

(ii) Find a similar expression for $\mathbf{E}\big[\, X^2 \,\big]$.

Solution: *By disjointness $\mathbf{P}[\, X = x \,] = \mathbf{P}[\, X > x - 1 \,] - \mathbf{P}[\, X > x \,]$ thus*

$$\begin{aligned}\mathbf{E}[\, X \,] &= \sum_{x=1}^{\infty} x\mathbf{P}[\, X = x \,] = \sum_{x=1}^{\infty} x\mathbf{P}[\, X > x - 1 \,] - \sum_{x=1}^{\infty} x\mathbf{P}[\, X > x \,] \\ &= \sum_{x=0}^{\infty} (x+1)\mathbf{P}[\, X > x \,] - \sum_{x=0}^{\infty} x\mathbf{P}[\, X > x \,] = \sum_{x=0}^{\infty} \mathbf{P}[\, X > x \,]\end{aligned}$$

Item (ii) is similar.

Question 3. *Throw two fair dice and consider the following three events:*

$A := \{$*the sum of the dice is* $7\}$, $\quad B := \{$*the first dice rolled a* $3\}$, $\quad C := \{$*the second dice rolled a* $4\}$.

(i) Show that the events are pairwise independent.

(ii) Are the three events are independent?

Solution: *Item (i) follows from calculation. For Item (ii) notice $A \cap B \cap C$ is the event that a 3 is rolled followed by a 4 and thus $\mathbf{P}[\, A \cap B \cap C \,] = 1/6^2 \neq 1/6^3$.*

Question 4. *Suppose you are throwing an unbiased, 6-faced dice sequentially until a 6 turns up followed by a 5.*

(i) *What is the expected waiting time?*

(ii) *What happens if you are waiting for a 6 followed by a 6?*

(iii) *Explain the difference.*

Solution: *Item (i): Let $\tau_{65} = \inf\{t: \text{ roll a } 65\}$ and let $\mathcal{VI}$ be the event the last roll was a 6. Then*

$$\mathbf{E}[\,\tau_{65}\,] = \frac{1}{6}\left(1 + \mathbf{E}[\,\tau_{65} \mid \mathcal{VI}\,]\right) + \frac{5}{6}\left(1 + \mathbf{E}[\,\tau_{65}\,]\right). \tag{1}$$

We also have

$$\mathbf{E}[\,\tau_{65} \mid \mathcal{VI}\,] = \frac{1}{6} + \frac{1}{6}\left(1 + \mathbf{E}[\,\tau_{65} \mid \mathcal{VI}\,]\right) + \frac{4}{6}\left(1 + \mathbf{E}[\,\tau_{65}\,]\right). \tag{2}$$

Thus $5\mathbf{E}[\,\tau_{65} \mid \mathcal{VI}\,] = 6 + 4\mathbf{E}[\,\tau_{65}\,]$ *by* (1) *and substituting this into* (2) *gives* $\mathbf{E}[\,\tau_{65}\,] = 36$.

Item (ii): Let $\tau_{66} = \inf\{t: \text{ roll a } 66\}$, this calculation is actually easier:

$$\mathbf{E}[\,\tau_{66}\,] = \frac{1}{6}\left(1 + \frac{1}{6} + \frac{5}{6}\left(1 + \mathbf{E}[\,\tau_{66}\,]\right)\right) + \frac{5}{6}\left(1 + \mathbf{E}[\,\tau_{66}\,]\right). \tag{3}$$

Solving (3) *gives* $\mathbf{E}[\,\tau_{66}\,] = 42$.

Item (iii): The probability we roll x followed by y is $1/36$, if we think of this as an independent trial then we expect to fail 36 times before success. This is the first experiment. In the second experiment there is no overlap between trials and a full trial takes time 2 total, thus we must add the expected number of times we roll 6 then something else before rolling 66, this is what gives the additional 6 expected time.

Question 5. *For an event $\mathcal{E}_n$ say that $\mathcal{E}_n$ occurs with high probability (w.h.p.) if $\mathbf{P}[\,\mathcal{E}_n\,] = 1 - o(1)$. Let $\{X_n\}_{n\geq 0}$ be a sequence of non-negative integer random variables. Show that if $\lim_{n\to\infty} \mathbf{E}[\,X_n\,] = 0$ then $X_n = 0$ w.h.p..*

Solution: *By Markov's inequality,*

$$\mathbf{P}[\,X_n \geq 1\,] \leq \mathbf{E}[\,X_n\,].$$

Thus if $\mathbf{E}[\,X_n\,] = o(1)$, then $\mathbf{P}[\,X_n \geq 1\,] = o(1)$. Since X_n is non-negative integer valued, $\mathbf{P}[\,X_n = 0\,] = 1 - o(1)$.

Question 6. *Let X be a random variable with expected value $\mu < \infty$ and variance $0 < \sigma^2 < \infty$.*

(i) *Show that for any real number $k > 0$ we have $\mathbf{P}[\,|X - \mu| \geq k\sigma\,] \leq 1/k^2$.*

(ii) *Deduce that if $X \geq 0$ then $\mathbf{P}[\,X = 0\,] \leq \sigma^2/\mu^2$.*

(iii) *Let $X = X_1 + \cdots + X_n$ where X_i are indicator random variables with $\mathbf{P}[\,X_i = 1\,] = p_i$ and $\mathbf{P}[\,X_i = 0\,] = 1 - p_i$. Show that*

$$\mathbf{Var}\,[\,X\,] \leq \mathbf{E}[\,X\,] + \sum_{1\leq i\neq j\leq n} \mathbf{Cov}\,[\,X_i, X_j\,], \quad \textit{where} \quad \mathbf{Cov}\,[\,X_i, X_j\,] = \mathbf{E}[\,X_iX_j\,] - \mathbf{E}[\,X_i\,]\,\mathbf{E}[\,X_j\,].$$

Solution: *Item (i) is in any standard text, even Wikipedia*

Item (ii) follows since $\mathbf{P}[\,X = 0\,] \leq \mathbf{P}[\,|X - \mu| \geq \mu\,]$.

Item (iii): To begin from the handout we have

$$\mathbf{Var}\left[\sum_{i=1}^{n} X_i\right] = \sum_{i=1}^{n} \mathbf{Var}\,[\,X_i\,] + \sum_{i\neq j} \mathbf{Cov}\,[\,X_i, X_j\,].$$

For X_i an indicator we have $\mathbf{Var}\,[\,X_i\,] = p_i(1 - p_i) \leq p_i = \mathbf{E}[\,X_i\,]$.

Question 7. *Let $G = (V, E)$ be an undirected graph of $|V| = n$ vertices. Assume that G has been constructed at random according to the following procedure: for any pair of vertices $\{u, v\} \in V \times V$, we put an undirected edge between u and v with probability p (with probability $1 - p$ there is no edge between u and v). We call a vertex u isolated if there is no edge incident to u.*

(i) Let X be the number of isolated vertices in G. Find an expression (which might depend on n and p) for $\mathbf{E}[\,X\,]$.

(ii) Use Question 5 to show that if $p > \frac{\ln n}{n}$ then with high probability there are no isolated vertices in G. You might need to use the fact that $\lim_{n\to\infty}\left(1 - \frac{x}{n}\right)^n = \mathrm{e}^{-x}$.

(iii) Use Question 6 (ii) and (iii) to argue that $\lim_{n\to\infty} \mathbf{P}[\,X > 0\,] = 1$ whenever $p < \frac{\ln n}{n}$. That is, G contains at least one isolated vertex with high probability.

Solution: *For Item (i) let X_i is the indicator random variable for the event that vertex i is isolated (not connected to any other vertex). A vertex is isolated with probability $(1-p)^{n-1}$ and so*

$$\mathbf{E}[\,X\,] = \sum_{i\in V}\mathbf{E}[\,X_i\,] = \binom{n}{1}(1-p)^{n-1}.$$

For Item (ii) if $p = \frac{(\ln n)+x}{n}$ for x not too large (or small), i.e. $|x| = o(n)$ then we have

$$\mathbf{E}[\,X\,] = n(1-p)^{n-1} = n\cdot\frac{e^{-np+O(np^2)}}{1-p} = n(1+o(1))e^{-\ln n - x} = (1+o(1))e^{-x}. \tag{4}$$

Thus provided $x \to \infty$ then by Q5 we have $\mathbf{P}[\,X > 0\,] \le (1+o(1))e^{-x} \to 0$.

For Item (iii) if $i \ne j$ then

$$\begin{aligned}\mathbf{Cov}\,[\,X_i, X_j\,] &= \mathbf{P}[\,X_i = 1, X_j = 1\,] - \mathbf{P}[\,X_i = 1\,]\,\mathbf{P}[\,X_j = 1\,]\\ &= (1-p)\cdot\left((1-p)^{n-2}\right)^2 - \left((1-p)^{n-1}\right)^2\\ &= p(1-p)^{2n-3}.\end{aligned}$$

Now notice that $\sum_{1\le i\ne j\le n}\mathbf{Cov}\,[\,X_i, X_j\,] = \binom{n}{2}p(1-p)^{2n-3} \le p(\mathbf{E}[\,X\,])^2$ for large n. Thus by Q6(iii) we have

$$\mathbf{Var}\,[\,X\,] \le \mathbf{E}[\,X\,] + \sum_{1\le i\ne j\le n}\mathbf{Cov}\,[\,X_i, X_j\,] \le \mathbf{E}[\,X\,] + p(\mathbf{E}[\,X\,])^2$$

and so by Q6(ii) and (4) we have

$$\mathbf{P}[\,X = 0\,] = \frac{\mathbf{Var}\,[\,X\,]}{\mathbf{E}[\,X\,]^2} \le \frac{\mathbf{E}[\,X\,] + p(\mathbf{E}[\,X\,])^2}{\mathbf{E}[\,X\,]^2} = \frac{1}{\mathbf{E}[\,X\,]} + p = (1+o(1))e^{x} + \frac{\ln n + x}{n}.$$

Thus if we take $x \to -\infty$, $x > -\ln n$ we obtain $\mathbf{P}[\,X = 0\,] = o(1)$ and so there are isolated vertices w.h.p.

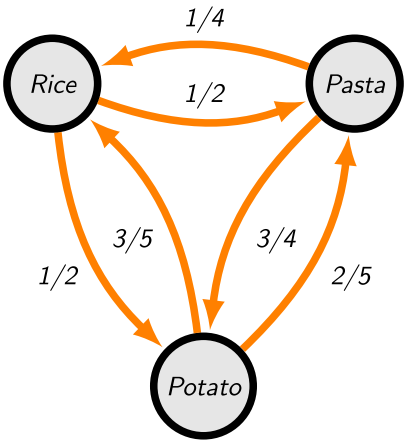
Question 8. *Let $\mathcal{F}$ be a finite collection of binary strings of finite lengths and assume no member of $\mathcal{F}$ is a prefix of another one. Let N_i denote the number of strings of length i in F. Prove that*

$$\sum_i \frac{N_i}{2^i} \le 1.$$

Solution: Since $\mathcal{F}$ is a collection of binary strings of finite length we can assume no string in $\mathcal{F}$ is longer than L. Let A_i be the event that a string of length i in $\mathcal{F}$ is the prefix of a random binary string of length L. Since no string in $\mathcal{F}$ is a prefix of any other the events A_i are disjoint. Thus

$$1 \ge \mathbf{P}[\,\cup_i A_i\,] = \sum_i \mathbf{P}[\,A_i\,] = \sum_i \frac{N_i}{2^i}.$$

Question 9. *The "College Carbs" Markov chain below makes an appearance in Lecture 2:*



This has transition matrix:

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

*(i) If I had Pasta on Monday what is the probability that I have Pasta on Thursday **and** Potato on Friday?*

(ii) If I have Pasta today then how many days should I expect to wait until I have Rice?

(iii) If I have Pasta today then how many days should I expect to wait until I next have Pasta again?

Solution: Let Pasta, Potato and Rice be states $1, 2$ and 3 respectively. For part (i),

$\mathbf{P}[\, X_3 = 1, X_4 = 2 \mid X_0 = 1\,] = \mathbf{P}[\, X_4 = 2 \mid X_3 = 1, X_0 = 1\,] \cdot \mathbf{P}[\, X_3 = 1 \mid X_0 = 1\,] = 3\mathbf{P}[\, X_3 = 1 \mid X_0 = 1\,]/4.$

Also observe that

$$\mathbf{P}[\, X_3 = 1 \mid X_0 = 1\,] = \sum_{1 \le i,j \le 3} P_{1,i}P_{i,j}P_{j,1} = P_{1,3}P_{3,2}P_{2,1} + P_{1,2}P_{2,3}P_{3,1} = \frac{11}{40},$$

thus the result for part (i) is $(3/4) \cdot (11/40) = 33/160$.

For parts (ii) and (iii) we have the following pair of linear equations by the Markov property

$$h_{1,3} = 1 + P_{1,2}h_{2,3} + P_{1,3} \cdot 0 \qquad \text{and} \qquad h_{2,3} = 1 + P_{2,1}h_{1,3} + P_{2,3} \cdot 0.$$

Thus we have $h_{1,3} = 1 + P_{1,2}(1 + P_{2,1}h_{1,3})$ and so $h_{1,3} = (1 + P_{1,2})/(1 + P_{1,2}P_{2,1}) = 35/26$.

More generally for $i \neq j \neq k \in \{1, 2, 3\}$ we have

$$h_{i,j} = \frac{1 + P_{i,k}}{1 + P_{i,k}P_{k,j}}$$

Thus we have

$$\mathbf{E}_1\left[\tau_1^+\right] = 1 + P_{1,2}h_{2,1} + P_{1,3}h_{3,1} = P_{1,2}\frac{1 + P_{2,3}}{1 + P_{2,3}P_{3,1}} + P_{1,3}\frac{1 + P_{3,2}}{1 + P_{3,2}P_{2,1}} = 13/4.$$

This also follows as we know $\pi_1 = 4/13$.

Question 10. *The Gamblers ruin chain appears in Lecture 2. This is a Markov Chain on $\{0, \dots, n\}$ with transition matrix P is given by $P_{i,i+1} = a$, $P_{i,i-1} = b = 1 - a$ for $1 \le i \le n - 1$ and $P_{0,0} = 1$, $P_{n,n} = 1$.*

(i) Describe the vectors v such that $vP = v$.

(ii) Why does this not contradict the results on the stationary distribution given in Lecture 2?

Solution: *Part (i):* Since $P_{0,0} = 1$, $P_{n,n} = 1$ the standard basis vectors $e_0 = (1, 0, \dots, 0)$ and $e_n = (0, \dots, 0, 1)$ and also any convex combination of the two satisfy $vP = v$, where v is a convex combination of e_0, e_n if there exists some $\alpha \in (0, 1)$ such that $v = \alpha e_0 + (1 - \alpha)e_n$.

Part (ii): The above says that there are infinitely many stationary vectors (they are a subspace of $\mathbb{R}^n$ of dimension 2). This does not contradict the Theorem on uniqueness of stationary distribution for finite irreducible chains as this chain is not irreducible.
