EECS 126 Probability and Random Processes — UC Berkeley, Spring 2024
Jiantao Jiao — Feb 27, 2024

Midterm 1

Last Name	First Name	SID

Left Neighbor First and Last Name	**Right Neighbor** First and Last Name

***Rules*.**

- **Unless otherwise stated, all your answers need to be justified and your work must be shown. Answers without sufficient justification will get no credit.**
- **All work you want graded should be on the fronts of the sheets in the space provided. Back sides may be used for scratch work, but will not be scanned/graded.**
- You have 10 minutes to read the exam and 70 minutes to complete the exam. (DSP students with $X\%$ time accommodation should spend $10 \cdot X\%$ time on reading and $70 \cdot X\%$ time on completing the exam).
- This exam is closed-book. You may reference one double-sided handwritten sheet of paper. No calculator or phones allowed.
- Collaboration with others is strictly prohibited. If you are caught cheating, you may fail the course and face disciplinary consequences.

Problem	out of
SID	1
Problem 1	72
Problem 2	30
Problem 3	15
Problem 4	8
Total	126

1 Similarity of Vectors [72 points]

Let $\mathcal{V}$ be an arbitrary set, and $N \in \mathbb{N}$ be a positive integer. Let $\mathbf{x} = (x_1, x_2, \ldots, x_N) \in \mathcal{V}^N$ and $\mathbf{y} = (y_1, y_2, \ldots, y_N) \in \mathcal{V}^N$ be two vectors where each element $x_i, y_i \in \mathcal{V}, \forall i \in [N] \triangleq \{1, 2, \ldots, N\}$. We are further given a function $f : \mathcal{V} \times \mathcal{V} \to \mathbb{R}$, and thus define the similarity of two vectors $\mathbf{x}$ and $\mathbf{y}$, denoted as $s(\mathbf{x}, \mathbf{y})$, as

$$s(\mathbf{x}, \mathbf{y}) \triangleq \sum_{i=1}^{N} f(x_i, y_i).$$

For each subpart, please carefully check the definition of $\mathcal{V}$ and $f(x, y)$ before you start solving problems.

1.1 Warm up: Bernoulli distribution [16 points]

Let $\mathcal{V} = \{0, 1\}$. Let $x_i \overset{\text{i.i.d.}}{\sim} \text{Bernoulli}(p_1)$ and $y_i \overset{\text{i.i.d.}}{\sim} \text{Bernoulli}(p_2)$ where $p_1, p_2 \in [0, 1]$ and $\mathbf{x}$ and $\mathbf{y}$ are independent. Let $f(x, y) = xy$.

(a) If $N = 1$, what is the distribution of $s(\mathbf{x}, \mathbf{y})$, and what is $\mathbb{E}[s(\mathbf{x}, \mathbf{y})]$ and $\text{Var}(s(\mathbf{x}, \mathbf{y}))$? [4 points]

(b) For a general fixed N and $p_2 = 1$, what is the distribution of $s(\mathbf{x}, \mathbf{y})$, and what is $\mathbb{E}[s(\mathbf{x}, \mathbf{y})]$ and $\text{Var}(s(\mathbf{x}, \mathbf{y}))$? [4 points]

(c) For a general fixed N and $p_1, p_2 \in (0, 1)$, what is the conditional probability $\mathbb{P}(x_1 = 1 | s(\mathbf{x}, \mathbf{y}) = 1)$? [8 points]

(a) When $N = 1$, $s(\mathbf{x}, \mathbf{y}) = x_1 y_1$ is the product of two Bernoulli random variables. By direct calculation, $s(\mathbf{x}, \mathbf{y}) \sim \text{Bernoulli}(p_1 p_2)$. [2 points] Therefore, $\mathbb{E}[s(\mathbf{x}, \mathbf{y})] = p_1 p_2$ [1 point] and $\text{Var}(s(\mathbf{x}, \mathbf{y})) = p_1 p_2 (1 - p_1 p_2)$. [1 point]

(b) When $p_2 = 1$, we have $y_i = 1$ almost surely. Therefore, $s(\mathbf{x}, \mathbf{y}) = \sum_{i=1}^{N} x_i$, which is the summation of N i.i.d. Bernoulli random variables. Therefore, $s(\mathbf{x}, \mathbf{y}) \sim \text{Binomial}(N, p_1)$. [2 points] Therefore, $\mathbb{E}[s(\mathbf{x}, \mathbf{y})] = N p_1$ [1 point] and $\text{Var}(s(\mathbf{x}, \mathbf{y})) = N p_1 (1 - p_1)$. [1 point]

(c) Applying Bayes' rule, we have

$$\mathbb{P}(x_1 = 1 | s(\mathbf{x}, \mathbf{y}) = 1) = \frac{\mathbb{P}(x_1 = 1)\mathbb{P}(s(\mathbf{x}, \mathbf{y}) = 1 | x_1 = 1)}{\mathbb{P}(s(\mathbf{x}, \mathbf{y}) = 1)} \text{[2 points]}.$$

Since x_1 is Bernoulli and $s(\mathbf{x}, \mathbf{y})$ is Binomial by the extension of (b), we can directly obtain that $\mathbb{P}(x_1 = 1) = p_1$ [1 point] and $\mathbb{P}(s(\mathbf{x}, \mathbf{y}) = 1) = N p_1 p_2 (1 - p_1 p_2)^{N-1}$. [1 point] To calculate $\mathbb{P}(s(\mathbf{x}, \mathbf{y}) = 1 | x_1 = 1)$, note that if $s(\mathbf{x}, \mathbf{y}) = 1$, either $x_1 y_1 = 1$ and $\sum_{i=2}^{N} x_i y_i = 0$, or $x_1 y_1 = 0$ and $\sum_{i=2}^{N} x_i y_i = 1$. Let $u = \sum_{i=2}^{N} x_i y_i \sim \text{Binomial}(N-1, p_1 p_2)$. Therefore,

$$\begin{aligned}
\mathbb{P}(s(\mathbf{x},\mathbf{y}) = 1 | x_1 = 1) =& \mathbb{P}(x_1 y_1 = 1, u = 0 | x_1 = 1) + \mathbb{P}(x_1 y_1 = 0, u = 1 | x_1 = 1) \text{ [1 point]} \\
=& \mathbb{P}(y_1 = 1, u = 0 | x_1 = 1) + \mathbb{P}(y_1 = 0, u = 1 | x_1 = 1) \\
=& \mathbb{P}(y_1 = 1, u = 0) + \mathbb{P}(y_1 = 0, u = 1) \\
=& \mathbb{P}(y_1 = 1)\mathbb{P}(u = 0) + \mathbb{P}(y_1 = 0)\mathbb{P}(u = 1) \text{ [1 point]} \\
=& p_2 (1 - p_1 p_2)^{N-1} + (1 - p_2)(N-1) p_1 p_2 (1 - p_1 p_2)^{N-2}. \text{ [1 point]}
\end{aligned}$$

Plugging all three probabilities in, we can finally get

$$\mathbb{P}(x_1 = 1 | s(\mathbf{x},\mathbf{y}) = 1) = \frac{p_1 p_2 (1 - p_1 p_2)^{N-1} + (1 - p_2)(N-1) p_1^2 p_2 (1 - p_1 p_2)^{N-2}}{N p_1 p_2 (1 - p_1 p_2)^{N-1}}. \text{ [1 point]}$$

1.2 Text similarity: Uniform distribution [20 points]

Let $\mathcal{V} = \{\text{'a'}, \text{'b'}, \dots, \text{'z'}, \text{'A'}, \text{'B'}, \dots, \text{'Z'}\}$ with $|\mathcal{V}| = 52$ which contains all 26 upper case letters and all 26 lower case letters. Let $f(x, y) = \mathbb{1}\{x = y\}, \forall x, y \in \mathcal{V}$, i.e., $f(x, y) = 1$ if and only if x and y are the same letter (case-sensitive). For example, $f(\text{'a'}, \text{'C'}) = 0$, $f(\text{'B'}, \text{'B'}) = 1$, $f(\text{'Z'}, \text{'z'}) = 0$. For a vector $\mathbf{x} \in \mathcal{V}^N$, it is essentially a text, and we will equivalently represent it using a string. For example, if $N = 5$, and $\mathbf{x} = (x_1, x_2, \dots, x_5) = (\text{'A'}, \text{'p'}, \text{'p'}, \text{'l'}, \text{'e'})$, we will also write $\mathbf{x} = \text{'Apple'}$ for convenience.

(a) If $N = 7$, and $\mathbf{x}$ is the first name of Prof. Jiantao Jiao, and $\mathbf{y}$ is the first name of head TA Tianhao Wu, what is x_2 and y_4? [4 points]

(b) If $N = 7$, $\mathbf{x} = \text{'Tianhao'}$ and $\mathbf{y} = \text{'Jiantao'}$, what is the similarity between $\mathbf{x}$ and $\mathbf{y}$, i.e., $s(\mathbf{x}, \mathbf{y})$? If $N = 11$, $\mathbf{x} = \text{'TATianhaoWu'}$ and $\mathbf{y} = \text{'probability'}$, what is $s(\mathbf{x}, \mathbf{y})$? [4 points]

(c) Assume each x_i and y_i are i.i.d. uniformly sampled from $\mathcal{V}$. For a general fixed N, please calculate $\mathbb{E}[s(\mathbf{x}, \mathbf{y})]$ and $\text{Var}(s(\mathbf{x}, \mathbf{y}))$. [4 points]

(d) Assume each x_i and y_i are still i.i.d. uniformly sampled from $\mathcal{V}$. Additionally, N follows a uniform distribution over $\{1, 2, 3, 4, 5\}$ and N, $\mathbf{x}$, $\mathbf{y}$ are mutually independent. Please calculate $\mathbb{E}[s(\mathbf{x}, \mathbf{y})]$ and $\text{Var}(s(\mathbf{x}, \mathbf{y}))$. You may denote $q = 1/52$ and leave q in your final answer. [8 points]

(a) Since $\mathbf{x} = \text{'Jiantao'}$ and $\mathbf{y} = \text{'Tianhao'}$, we can obtain that $x_2 = \text{'i'}$ [2 points] and $y_4 = \text{'n'}$ [2 points].

(b) When $\mathbf{x} = \text{'Tianhao'}$ and $\mathbf{y} = \text{'Jiantao'}$, by comparing letters one by one, we have $s(\mathbf{x}, \mathbf{y}) = 5$. [2 points] When $\mathbf{x} = \text{'TATianhaoWu'}$ and $\mathbf{y} = \text{'probability'}$, by careful comparison, we have $s(\mathbf{x}, \mathbf{y}) = 1$. [2 points]

(c) First, observe that $\mathbb{P}(f(x_i, y_i) = 1) = \mathbb{P}(x_i = y_i) = 1/52$. We denote $q = 1/52$. Then $s(\mathbf{x}, \mathbf{y}) \sim \text{Binomial}(N, q)$, from which we can conclude $\mathbb{E}[s(\mathbf{x}, \mathbf{y})] = Nq = \frac{N}{52}$ [2 points] and $\text{Var}(s(\mathbf{x}, \mathbf{y})) = Nq(1 - q) = \frac{51N}{52^2}$. [2 points]

(d) By the law of iterated expectation,

$$\begin{aligned}\mathbb{E}[s(\mathbf{x}, \mathbf{y})] =& \mathbb{E}[\mathbb{E}[s(\mathbf{x}, \mathbf{y})|N]] \text{ [1 point]}\\ =& \mathbb{E}[Nq] \text{ [1 point]}\\ =& q\mathbb{E}[N] = 3q. \text{ [1 point]}\end{aligned}$$

To calculate $\text{Var}(s(\mathbf{x}, \mathbf{y}))$, we need to calculate $\mathbb{E}[\text{Var}(s(\mathbf{x}, \mathbf{y})|N)]$ and $\text{Var}(\mathbb{E}[s(\mathbf{x}, \mathbf{y})|N])$. For the first term,

$$\mathbb{E}[\text{Var}(s(\mathbf{x}, \mathbf{y})|N)] = \mathbb{E}[Nq(1 - q)] = q(1 - q)\mathbb{E}[N] = 3q(1 - q). \text{ [2 points]}$$

For the second term,

$$\begin{aligned}\mathrm{Var}(\mathbb{E}[s(\mathbf{x},\mathbf{y})|N]) =& \mathbb{E}[\mathbb{E}^2[s(\mathbf{x},\mathbf{y})|N]] - \mathbb{E}^2[\mathbb{E}[s(\mathbf{x},\mathbf{y})|N]] \\ =& \mathbb{E}[N^2q^2] - \mathbb{E}^2[Nq] \text{ [1 point]} \\ =& 11q^2 - 9q^2 = 2q^2. \text{ [1 point]}\end{aligned}$$

Therefore, by the law of total variance, we can obtain that

$\mathrm{Var}(s(\mathbf{x},\mathbf{y})) = \mathbb{E}[\mathrm{Var}(s(\mathbf{x},\mathbf{y})|N)] + \mathrm{Var}(\mathbb{E}[s(\mathbf{x},\mathbf{y})|N]) = 3q(1-q) + 2q^2 = q(3-q)$. [1 point]

1.3 Exponential fun: Exponential distribution [20 points]

Let $\mathcal{V} = \mathbb{R}$ be the set of all real numbers. Let $f(x, y) = xy$ be the product of x and y. In this case, $s(\mathbf{x}, \mathbf{y})$ coincides with the inner product of two real-valued vectors. Also, we assume $\mathbf{x}$ is independent of N and $\mathbf{y}$. Moreover, $x_i \overset{\text{i.i.d.}}{\sim} \exp(\lambda)$ for all $i \in [N]$ for some fixed $\lambda > 0$. Therefore, the p.d.f. of x_i is $f_{x_i}(t) = \frac{1}{\lambda} e^{-\lambda t}$ for $t \geq 0$.

(a) What is the MGF (moment generating function) of x_i? You only need to provide the answer without justification. Be careful about the domain of the MGF. [4 points]

(b) Let $y_i = 1$ for all $i = 1, 2, \ldots$ For a fixed N, what is the MGF of $s(\mathbf{x}, \mathbf{y})$? [4 points]

(c) Let $y_i = 1$ for all $i = 1, 2, \ldots$ Let $N \sim \text{Geo}(p)$ which is a geometric distribution with head probability $p \in (0, 1)$. What is the MGF of $s(\mathbf{x}, \mathbf{y})$? From the MGF of $s(\mathbf{x}, \mathbf{y})$, what is the distribution of $s(\mathbf{x}, \mathbf{y})$? *Hint: Consider applying the law of iterated expectation. Again be careful about the domain of the MGF.* [8 points]

(d) Let $y_i \overset{\text{i.i.d.}}{\sim} \text{Bernoulli}(p)$ for some fixed $p \in (0, 1)$ for $i = 1, 2, \ldots$. Let $N = \min\{n \geq 1 \mid y_n = 0\}$. What is the distribution of N and $s(\mathbf{x}, \mathbf{y})$? You only need to provide the answer without justification. [4 points]

(a) $M_{x_i}(t) = \frac{\lambda}{\lambda - t}$ [2 points] for $t < \lambda$ [1 point].

(b) Since $s(\mathbf{x}, \mathbf{y}) = \sum_{i=1}^{N} x_i$ where x_i are i.i.d. exponential random variable, [1 point] we have

$$M_{s(\mathbf{x},\mathbf{y})}(t) = \prod_{i=1}^{N} M_{x_i}(t) = \left(\frac{\lambda}{\lambda - t}\right)^N. \text{[2 points]}$$

(c) By law of iterated expectation, we have

$$M_{s(\mathbf{x},\mathbf{y})}(t) = \mathbb{E}[e^{s(\mathbf{x},\mathbf{y})t}] = \mathbb{E}[\mathbb{E}[e^{s(\mathbf{x},\mathbf{y})t}|N]] = \mathbb{E}\left[\left(\frac{\lambda}{\lambda - t}\right)^N\right]. \text{ [2 points]}$$

Since $N \sim \text{Geo}(p)$, we have

$$\begin{aligned}
\mathbb{E}\left[\left(\frac{\lambda}{\lambda - t}\right)^N\right] &= \sum_{i=1}^{\infty} \mathbb{P}(N = i) \left(\frac{\lambda}{\lambda - t}\right)^i \\
&= \sum_{i=1}^{\infty} (1 - p)^{i-1} p \cdot \left(\frac{\lambda}{\lambda - t}\right)^i \\
&= \frac{p\lambda}{\lambda - t} \sum_{i=0}^{\infty} \left(\frac{(1 - p)\lambda}{\lambda - t}\right)^i. \text{ [2 points]}
\end{aligned}$$

When $\frac{(1-p)\lambda}{\lambda-t} \in (0,1)$, which is $t < p\lambda$ [1 point], we have

$$\sum_{i=0}^{\infty} \left(\frac{(1-p)\lambda}{\lambda - t}\right)^i = \frac{1}{1 - \frac{(1-p)\lambda}{\lambda-t}} = \frac{\lambda - t}{p\lambda - t}.$$ [1 point]

Therefore, we have $M_{s(\mathbf{x},\mathbf{y})}(t) = \frac{p\lambda}{p\lambda - t}$ where $t < p\lambda$. [1 point]

From the MGF of $s(\mathbf{x},\mathbf{y})$, we have $s(\mathbf{x},\mathbf{y}) \sim \exp(p\lambda)$. [1 point]

(d) $N \sim \text{Geo}(1-p)$. [1 point]. $s(\mathbf{x},\mathbf{y})$ follows a mixed distribution, where with probability $1-p$ it is 0, and with probability p it follows $\exp((1-p)\lambda)$. [2 points]

1.4 Concentration: Gaussian distribution [16 points]

Let $\mathcal{V} = \mathbb{R}$ be the set of all real numbers, and let $f(x, y) = xy$ be the product of x and y. Also, we assume $\mathbf{x}$, $\mathbf{y}$ and N are jointly independent. Moreover, $x_i \overset{\text{i.i.d.}}{\sim} \mathcal{N}(0, \sigma^2)$ for all $i \in [N]$ for some fixed $\sigma > 0$.

(a) Let $N = 126$ and $y_i = 2^{\frac{i-1}{2}}$ for all $i \in [N]$. What is the distribution of $s(\mathbf{x}, \mathbf{y})$? You only need to provide your answer without justification. [4 points]

(b) Let $N = 1$ and $y_1 = 1$. Please use Chebyshev's inequality to prove that $\mathbb{P}(s(\mathbf{x}, \mathbf{y}) \geq k\sigma) \leq \frac{1}{2k^2}$ for $k \geq 1$. [4 points]

(c) Let $N = 1$ and $y_1 = 1$. Now, we want to get a tighter bound for large k. We could apply Markov's inequality to $e^{t \cdot s(\mathbf{x}, \mathbf{y})}$ (which is known as Chernoff bound), to obtain that

$$\mathbb{P}(s(\mathbf{x}, \mathbf{y}) \geq k\sigma) = \mathbb{P}(e^{t \cdot s(\mathbf{x}, \mathbf{y})} \geq e^{t \cdot k\sigma}) \leq \frac{\mathbb{E}[e^{t \cdot s(\mathbf{x}, \mathbf{y})}]}{e^{t \cdot k\sigma}}, \forall t > 0.$$

By choosing an appropriate t, one can get a tighter bound than we have in (b) for large k. Now please prove

$$\mathbb{P}(s(\mathbf{x}, \mathbf{y}) \geq k\sigma) \leq e^{-\frac{k^2}{2}}.$$

by choosing an appropriate t. You need to solve t explicitly. *Hint: You may use without proof that the MGF of* $Z \sim \mathcal{N}(\mu, \sigma^2)$ *is* $M_Z(t) = \exp\left(\mu t + \frac{1}{2}\sigma^2 t^2\right)$. [8 points]

(a) $s(\mathbf{x}, \mathbf{y}) \sim \mathcal{N}(0, (2^{126} - 1)\sigma^2)$. [3 points]

(b) Since $s(\mathbf{x}, \mathbf{y}) = x_1 \sim \mathcal{N}(0, \sigma^2)$, we have $\mathbb{E}[s(\mathbf{x}, \mathbf{y})] = 0$ and $\text{Var}(s(\mathbf{x}, \mathbf{y})) = \sigma^2$. [1 point] By Chebyshev's inequality, we have

$$\mathbb{P}(|s(\mathbf{x}, \mathbf{y})| \geq k\sigma) \leq \frac{1}{k^2}. \text{ [1 point]}$$

By symmetry of the Gaussian distribution, we have $\mathbb{P}(s(\mathbf{x}, \mathbf{y}) \geq k\sigma) \leq \frac{1}{2k^2}$. [1 point]

(c) Since $\mathbb{E}[e^{t \cdot s(\mathbf{x}, \mathbf{y})}] = \exp(\frac{1}{2}\sigma^2 t^2)$, [1 point] we have

$$\mathbb{P}(s(\mathbf{x}, \mathbf{y}) \geq k\sigma) \leq \exp(\frac{1}{2}\sigma^2 t^2 - k\sigma t). \text{[1 point]}$$

By minimizing $\frac{1}{2}\sigma^2 t^2 - k\sigma t$, we have can solve $t = \frac{k}{\sigma}$. [2 points]

2 Card Game [30 points]

x club ♣ cards and y heart ♡ cards are stacked face down in a uniformly random order. Two players Alice and Bob, knowing x and y but not knowing the order, take turns to choose to either deal the top card to their opponent or themselves, after which the distributed card is immediately revealed. Dealing a ♡ to oneself gives one another turn (i.e. skips one's opponent's next turn). Whenever one player is distributed a ♣, the game ends and the other player wins. Alice deals first.

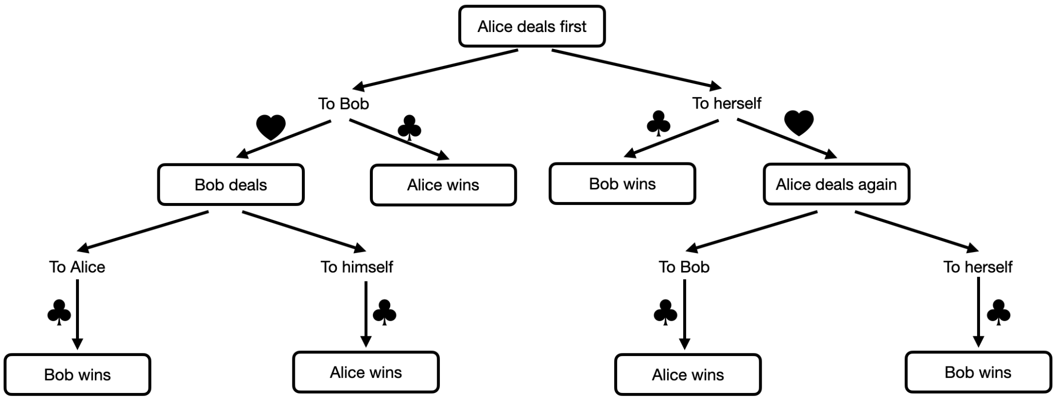

Figure 1: Example flow of the game when $x = y = 1$

(a) (15 points) Consider the setting where $x = 2, y = 3$. If both player always deal the card to their opponent, find:

$$\mathbb{P}(\text{ the third card is } \clubsuit \mid \text{Alice wins }).$$

(b) (15 points) Consider the setting where $x = 1, y = 2$. A strategy is defined as a mapping from the numbers of remaining ♣ and ♡ cards in the pile to the choice of dealing (possibly random). Describe Alice's optimal strategy and find her probability to win under this strategy.

(a) Define:

$$S = \{\text{the third card is } \clubsuit\}$$
$$W = \{\text{Alice wins}\}.$$

There are only 4 orders out of $\binom{5}{2} = 10$ total orderings in which Alice losses:

(a) ♡, ♣, ♣, ♡, ♡;
(b) ♡, ♣, ♡, ♣, ♡;
(c) ♡, ♣, ♡, ♡, ♣;
(d) ♡, ♡, ♡, ♣, ♣;

It follows that $\mathbb{P}(W) = 1 - \frac{4}{10} = 6/10$ (3 points). Notice that $\mathbb{P}(S) = 2/5$ (3 points), $\mathbb{P}(W|S) = 3/4$ (3 points) since Alice losses only when the second card is ♣ given S. By

Bayes rule,

$$\begin{aligned}\mathbb{P}(S|W) &= \frac{\mathbb{P}(W|S)\mathbb{P}(S)}{\mathbb{P}(W)} \text{ (4 points)} \\ &= \frac{\frac{3}{4}\cdot\frac{2}{5}}{\frac{6}{10}} \\ &= 1/2 \text{ (2 points)}.\end{aligned}$$

(b) Alice's best strategy is to always deal to Bob (5 points). Let

$$\begin{aligned}W &= \{\text{Alice wins}\} \\ T &= \{\text{the first card is } \clubsuit\}.\end{aligned}$$

From Figure 1, it is obvious that the winning probability is always 1/2 for both player when $x = y = 1$, whatever strategy they use (3 points). By law of total probability,

$$\begin{aligned}\mathbb{P}(\text{Alice wins}) &= \mathbb{P}(T)\cdot\mathbb{P}(W|T) + \mathbb{P}(T^c)\cdot\mathbb{P}(W|T^c) \text{ (3 points)} \\ &= \frac{1}{3}\cdot 1 + \frac{2}{3}\cdot\frac{1}{2} \text{ (3 points)} \\ &= 2/3 \text{ (1 points)}.\end{aligned}$$

3 Independence [15 points]

Random variable (ξ, η) has joint density

$$p(x,y) = \begin{cases} \frac{1+xy}{4}, & |x|<1, |y|<1 \\ 0, & \text{otherwise} \end{cases}.$$

Prove that ξ and η are not independent but ξ^2 and η^2 are independent.

We have

$$f_\xi(x) = \begin{cases} \frac{1}{2}, & |x|<1 \\ 0, & \text{otherwise} \end{cases}$$
$$g_\eta(y) = \begin{cases} \frac{1}{2}, & |y|<1 \\ 0, & \text{otherwise} \end{cases}.$$

(3 points)
Since $p(x,y) \neq f_\eta(x) g_\eta(y)$ (3 points), ξ and η are not independent.
Notice for any $u \in [0,1]$,

$$\mathbb{P}(\xi^2 < u) = \int_{x^2<u} f_\xi(x)dx = \sqrt{u}$$
$$\mathbb{P}(\eta^2 < v) = \sqrt{v}.$$

(3 points)
Furthermore,

$$\mathbb{P}(\xi^2 < u, \eta^2 < v) = \int\int_{x^2<u, y^2<v} p(x,y)dxdy = \sqrt{uv} \text{ (3 points)}.$$

Since $\mathbb{P}(\xi^2 < u, \eta^2 < v) = \mathbb{P}(\xi^2 < u)\mathbb{P}(\eta^2 < v)$ (3 points), it follows that ξ^2 and η^2 are independent.

4 Concentration Inequalities [8 points]

Prove that for any random variable X,

$$\mathbb{P}(X = 0) \leq \frac{\mathrm{Var}(X)}{\mathbb{E}[X]^2}.$$

By Chebyshev's inequality,

$$\begin{aligned}\mathbb{P}(X = 0) &\leq \mathbb{P}(|X - \mathbb{E}[X]| \geq |\mathbb{E}[X]|) \text{ (4 points)} \\ &\leq \frac{\mathrm{Var}(X)}{\mathbb{E}[X]^2}. \text{ (4 points)}\end{aligned}$$