Unit 4 - Week 3: Randomness and entropy

Weekly Assignment 3

The due date for submitting this assignment has passed.
As per our records you have not submitted this assignment.

Due on 2020-10-07, 23:59 IST.

Note: All the questions below may have multiple correct answers. There is no negative marking for wrong choices. However, partial marking is considered only when none of the marked choices is wrong.

Questions 1 and 2 are based on the following:

Let $\mathcal{X}$ be a finite alphabet, and P be a probability distribution over $\mathcal{X}$. Define $A_1, A_2 \subset \mathcal{X}$ as follows:
$A_1 = \{x \in \mathcal{X} : -\log P(x) \geq 4\}$
$A_2 = \{x \in \mathcal{X} : -\log P(x) \leq 4\}$.
Moreover, it is known that $P(A_1) \geq 7/8$ and $P(A_2) \geq 7/8$.

1) Which of the following can be deduced about $|A_1|$ and $|A_2|$? *1 point*

- $|A_1| \leq 16$
- $|A_1| \geq 14$
- $|A_2| \leq 16$
- $|A_2| \geq 14$

No, the answer is incorrect.
Score: 0
Accepted Answers:
$|A_1| \geq 14$
$|A_2| \leq 16$

2) Let $A = A_1 \cap A_2$. Which of the following can be deduced about A? *1 point*

- $P(A) \geq 6/8$
- $P(A) \leq 1/2$
- $P(x) = 1/16$ for every $x \in A$
- $|A| \geq 12$

No, the answer is incorrect.
Score: 0
Accepted Answers:
$P(A) \geq 6/8$
$P(x) = 1/16$ *for every* $x \in A$
$|A| \geq 12$

3) Let $P = (1/2, 1/2)$ and $Q = (1/4, 3/4)$ be two probability distributions over some alphabet $\mathcal{X}$, where $|\mathcal{X}| = 2$. Which of the following are true? (d_{TV} stands for the total variation distance.) *1 point*

- $d_{TV}(P,Q) = 1/4$
- $d_{TV}(P,Q) = 1/2$
- Can't compute $d_{TV}(P,Q)$ unless we have more information about $\mathcal{X}$
- $d_{TV} = 1/4$ if $\mathcal{X} = \{0,1\}$, whereas $d_{TV} = 0$ if $\mathcal{X} = \{-1,1\}$

No, the answer is incorrect.
Score: 0
Accepted Answers:
$d_{TV}(P,Q) = 1/4$

4) Let P, Q be two probability distributions on $\mathcal{X}$, such that $d_{TV}(P,Q) \leq 0.2$. Let $A \subset \mathcal{X}$ with $P(A) \geq 0.7$. What can we say about $Q(A)$? *1 point*

- $Q(A) \geq 0$
- $Q(A) \geq 0.5$
- $Q(A) \leq 1$
- $Q(A) \leq 0.9$

No, the answer is incorrect.
Score: 0
Accepted Answers:
$Q(A) \geq 0$
$Q(A) \geq 0.5$
$Q(A) \leq 1$

5) Let P be a probability distribution on $\mathcal{X}$. Let $A \subset \mathcal{X}$ with $P(A) \geq 7/8$. Define a new distribution $\bar{P}$ on $\mathcal{X}$ as follows: *1 point*
$\bar{P}(x) = \begin{cases} P(x)/P(A), & \text{if } x \in A \\ 0, & \text{if } x \notin A \end{cases}$.
Which of the following are true?

- $d_{TV}(P,\bar{P}) = \sum_{x \notin A} P(x)$
- $d_{TV}(P,\bar{P}) = \sum_{x \in A} P(x) \cdot \left(\frac{1}{P(A)} - 1\right)$
- $d_{TV}(P,\bar{P}) \geq 7/8$
- $d_{TV}(P,\bar{P}) \leq 1/8$

No, the answer is incorrect.
Score: 0
Accepted Answers:
$d_{TV}(P,\bar{P}) = \sum_{x \notin A} P(x)$
$d_{TV}(P,\bar{P}) = \sum_{x \in A} P(x) \cdot \left(\frac{1}{P(A)} - 1\right)$
$d_{TV}(P,\bar{P}) \leq 1/8$

6) You are stuck in a desert island with an unfair coin whose probability of Heads is 2/3. To keep yourself occupied, you organize a cricket match between a team of crocodiles and a team of alligators. You are forced to use the unfair coin to decide which team bats first. How do you make a fair decision? This is the problem of simulating a fair coin from an unfair coin. von Neumann proposed the following scheme: *1 point*

Toss the unfair coin twice:
- *If the first toss is Heads and the second toss is Tails: Declare Heads.*
- *If the first toss is Tails and the second toss is Heads: Declare Tails.*
- *If either both the tosses are Heads or both the tosses are Tails: Repeat (i.e. toss the coin twice again).*

Let *N* be the number of rounds we need to declare a result. (Here, each round involves tossing the unfair coin twice.) Clearly, *N* is not a fixed deterministic number. Which of the following are true?

($\mathbb{P}$ denotes probability, $\mathbb{E}$ denotes expectation.)

- $\mathbb{P}(N = 1) = 2/9$
- $\mathbb{P}(N = 1) = 4/9$
- *N* is a Geometric random variable
- $\mathbb{E}[N] = 9/4$

No, the answer is incorrect.
Score: 0
Accepted Answers:
$\mathbb{P}(N = 1) = 4/9$
N is a Geometric random variable
$\mathbb{E}[N] = 9/4$

7) In a reversal of situation, suppose that you only had a fair coin with you (i.e. with probability of Heads 1/2). But you feel a special tenderness towards crocodiles and want them to win the toss with probability 2/3. During the toss, crocodiles choose Heads. Using the fair coin, how do you ensure that you declare Heads with probability 2/3? Consider the following scheme: *1 point*

Toss the fair coin (Toss 1) -- if it shows Heads, declare Heads.
If it shows Tails, toss the coin again (Toss 2) -- if it shows Heads, declare Tails.
If it shows Tails, toss the coin again (Toss 3) -- if it shows Heads, declare Heads.
If it shows Tails, toss the coin again (Toss 4) -- if it shows Heads, declare Tails ...and so on.

That is
- Whenever the coin shows Tails: Toss it again.
- Whenever the coin shows Heads: Declare Heads if it was an `odd' toss (Toss 1, Toss 3,...); Declare Tails if it was an `even' toss (Toss 2, Toss 4,...).

Which of the following are true?

- Probability of declaring Heads is $2\sum_{i=1}^{\infty} 2^{-2i}$
- Probability of declaring Heads is $\sum_{i=1}^{\infty} 2^{-2i}$
- Probability of declaring Heads is 2/3
- Expected number of tosses required to declare a result (either Heads or Tails) is 2

No, the answer is incorrect.
Score: 0
Accepted Answers:
Probability of declaring Heads is $2\sum_{i=1}^{\infty} 2^{-2i}$
Probability of declaring Heads is 2/3
Expected number of tosses required to declare a result (either Heads or Tails) is 2

8) A coin with probability of Heads p is tossed n times independently. Let $Y = (X_1, \ldots, X_n)$ be the result of the n tosses, where each X_i takes values in an alphabet of cardinality 2 (Heads or Tails). Thus, Y can be thought of as a random variable that takes values in an alphabet of cardinality 2^n. Which of the following are true? *1 point*

- $H(Y) = n$
- $H(Y) = np$
- $H(Y) = -np\log(1-p)$
- $H(Y) = -np\log p - n(1-p)\log(1-p)$

No, the answer is incorrect.
Score: 0
Accepted Answers:
$H(Y) = -np\log p - n(1-p)\log(1-p)$

9) $P = (\frac{1}{4}, \frac{1}{4}, \frac{1}{2})$ is a probability distribution over an alphabet $\mathcal{X}$ of cardinality 3. Let $X_1, \ldots, X_n$ be independent and identically distributed (i.i.d.) random variables with distribution P. For each $i \in \{1, \ldots, n\},$ let $Z_i = -\log P(X_i)$. Note that $Z_1, \ldots, Z_n$ are also i.i.d. random variables. *1 point*

Which of the following are true about Z_1?
($\mathbb{E}$ denotes expectation, $\mathbf{var}$ denotes variance.)

- $\mathbb{E}[Z_1] = 1$
- $\mathbb{E}[Z_1] = 3/2$
- $\mathbf{var}(Z_1) = 1/4$
- $\mathbf{var}(Z_1) = 1/2$

No, the answer is incorrect.
Score: 0
Accepted Answers:
$\mathbb{E}[Z_1] = 3/2$
$\mathbf{var}(Z_1) = 1/4$

10) Consider the same setting as in Question 9. Let $Y = (X_1, \ldots, X_n)$. Which of the following are true? *1 point*
($\mathbb{P}$ denotes probability)

- $H(Y) = \sum_{i=1}^{n} \mathbb{E}[Z_i]$
- $H(Y) = 3n/2$
- $\mathbb{P}(\sum_{i=1}^{n} Z_i \geq \frac{3n}{2} - \sqrt{\frac{n}{4\epsilon}}) \geq 1 - \epsilon$
- $\mathbb{P}(\sum_{i=1}^{n} Z_i \leq \frac{3n}{2} + \sqrt{\frac{n}{4\epsilon}}) \geq 1 - \epsilon$

No, the answer is incorrect.
Score: 0
Accepted Answers:
$H(Y) = \sum_{i=1}^{n} \mathbb{E}[Z_i]$
$H(Y) = 3n/2$
$\mathbb{P}(\sum_{i=1}^{n} Z_i \geq \frac{3n}{2} - \sqrt{\frac{n}{4\epsilon}}) \geq 1 - \epsilon$
$\mathbb{P}(\sum_{i=1}^{n} Z_i \leq \frac{3n}{2} + \sqrt{\frac{n}{4\epsilon}}) \geq 1 - \epsilon$