Unit 9 - Week 8: Information theoretic lower bounds and Data Compression-1

Weekly Assignment 8

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

Due on 2020-11-11, 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.

```
Unless stated otherwise, assume that probability distributions are defined over a finite alphabet, an
d that log is to the base 2. In addition to the lectures in Week 8, also review the lecture on `Lower
bound for hypothesis testing' from Week 7 for this assignment.
```

1) Let X be a random variable taking values in alphabet $\mathcal{X}$. Consider the binary hypothesis testing problem $H_0 : X \sim P, H_1 : X \sim Q$. Recall that $\beta_\epsilon(P,Q)$ is defined as $\beta_\epsilon(P,Q) = \min_{T:\mathcal{X}\to\{0,1\}} \{\sum_{x\in\mathcal{X}} Q(x)T(0|x) : \sum_{x\in\mathcal{X}} P(x)T(1|x) \le \epsilon\}$. Define another quantity $\hat{\beta}_\epsilon(P,Q)$ as follows: $\hat{\beta}_\epsilon(P,Q) \triangleq \min_{A\subset\mathcal{X}} \{Q(A) : P(A) \ge 1-\epsilon\}$. Which of the following are true? *1 point*

- $\hat{\beta}_\epsilon(P,Q) \ge \beta_\epsilon(P,Q)$
- $\hat{\beta}_\epsilon(P,Q) \le \beta_\epsilon(P,Q)$
- $\hat{\beta}_\epsilon(P,Q) = \beta_\epsilon(P,Q)$
- The two quantities are incomparable.

No, the answer is incorrect.
Score: 0
Accepted Answers:
$\hat{\beta}_\epsilon(P,Q) \ge \beta_\epsilon(P,Q)$

2) Let X be a random variable taking values in alphabet $\mathcal{X}$. Consider the binary hypothesis testing problem $H_0 : X \sim P, H_1 : X \sim Q$. Let $T : \mathcal{X} \to \{0,1\}$ be a test. Let $P_{FA} = P(T(X) = 1)$ be the probability of false alarm, and let $P_D = Q(T(X) = 1)$ be the probability of detection. Which of the following are true? *1 point*
($\mathrm{Ber}(p)$ denotes Bernoulli distribution with parameter p.)

- $D(P||Q) \ge D(\mathrm{Ber}(P_{FA})||\mathrm{Ber}(P_D))$
- $D(P||Q) \ge D(\mathrm{Ber}(P_{FA})||\mathrm{Ber}(1-P_D))$
- $D(P||Q) \le D(\mathrm{Ber}(P_{FA})||\mathrm{Ber}(P_D))$
- $D(P||Q) \le D(\mathrm{Ber}(1-P_{FA})||\mathrm{Ber}(1-P_D))$

No, the answer is incorrect.
Score: 0
Accepted Answers:
$D(P||Q) \ge D(\mathrm{Ber}(P_{FA})||\mathrm{Ber}(P_D))$

3) Let X be a random variable taking values in alphabet $\mathcal{X}$ with distribution P. Let $f : \mathcal{X} \to \{0,1\}^\ell$ and let Q be a distribution over $\{0,1\}^\ell$ defined as $Q(y) = P(f(X) = y)$ for every $y \in \{0,1\}^\ell$. Suppose it is given that $Q(y) \in [(1-\epsilon)2^{-\ell}, (1+\epsilon)2^{-\ell}]$ for every $y \in \{0,1\}^\ell$. Which of the following are true? *1 point*

- $H(P) \le H(Q)$
- $H(P) \ge H(Q)$
- $H(P) \ge \ell(1-\epsilon) - 1$
- $H(Q) \ge \ell(1-\epsilon) - 1$

No, the answer is incorrect.
Score: 0
Accepted Answers:
$H(P) \ge H(Q)$
$H(P) \ge \ell(1-\epsilon) - 1$
$H(Q) \ge \ell(1-\epsilon) - 1$

4) Let $X_1, \ldots, X_n$ be positive-valued random variables. Which of the following are true? *1 point*

- $\frac{1}{n}\sum_{i=1}^n \mathbb{E}[X_i \log X_i] \ge \frac{1}{n}\sum_{i=1}^n \mathbb{E}[X_i] \log \mathbb{E}[X_i]$
- $\frac{1}{n}\sum_{i=1}^n \mathbb{E}[X_i \log X_i] \le \frac{1}{n}\sum_{i=1}^n \mathbb{E}[X_i] \log \mathbb{E}[X_i]$
- $\frac{1}{n}\sum_{i=1}^n \mathbb{E}[X_i \log X_i] \ge \left(\frac{1}{n}\sum_{i=1}^n \mathbb{E}[X_i]\right) \log\left(\frac{1}{n}\sum_{i=1}^n \mathbb{E}[X_i]\right)$
- $\frac{1}{n}\sum_{i=1}^n \mathbb{E}[X_i \log X_i] = \left(\frac{1}{n}\sum_{i=1}^n \mathbb{E}[X_i]\right) \log\left(\frac{1}{n}\sum_{i=1}^n \mathbb{E}[X_i]\right)$

No, the answer is incorrect.
Score: 0
Accepted Answers:
$\frac{1}{n}\sum_{i=1}^n \mathbb{E}[X_i \log X_i] \ge \frac{1}{n}\sum_{i=1}^n \mathbb{E}[X_i] \log \mathbb{E}[X_i]$
$\frac{1}{n}\sum_{i=1}^n \mathbb{E}[X_i \log X_i] \ge \left(\frac{1}{n}\sum_{i=1}^n \mathbb{E}[X_i]\right) \log\left(\frac{1}{n}\sum_{i=1}^n \mathbb{E}[X_i]\right)$

5) Let X, Y, Z be random variables. Let P_{XYZ} and Q_{XYZ} be two joint distributions of these random variables. Which of the following are true? *1 point*

- $D(P_{XYZ}||Q_{XYZ}) = D(P_X||Q_X) + \mathbb{E}_X[D(P_{Y|X}||Q_{Y|X})] + \mathbb{E}_{X,Y}[D(P_{Z|X,Y}||Q_{Z|X,Y})]$
- $D(P_{XYZ}||Q_{XYZ}) = D(P_X||Q_X) + \mathbb{E}_Y[D(P_{X|Y}||Q_{X|Y})] + \mathbb{E}_Z[D(P_{XY|Z}||Q_{XY|Z})]$
- $D(P_{XYZ}||Q_{XYZ}) = D(P_X||Q_X) + \mathbb{E}_X[D(P_{YZ|X}||Q_{YZ|X})]$
- $D(P_{XYZ}||Q_{XYZ}) = D(P_Z||Q_Z) + \mathbb{E}_Z[D(P_{XY|Z}||Q_{XY|Z})]$

No, the answer is incorrect.
Score: 0
Accepted Answers:
$D(P_{XYZ}||Q_{XYZ}) = D(P_X||Q_X) + \mathbb{E}_X[D(P_{Y|X}||Q_{Y|X})] + \mathbb{E}_{X,Y}[D(P_{Z|X,Y}||Q_{Z|X,Y})]$
$D(P_{XYZ}||Q_{XYZ}) = D(P_X||Q_X) + \mathbb{E}_X[D(P_{YZ|X}||Q_{YZ|X})]$
$D(P_{XYZ}||Q_{XYZ}) = D(P_Z||Q_Z) + \mathbb{E}_Z[D(P_{XY|Z}||Q_{XY|Z})]$

6) Suppose X is a random variable such that $\mathbb{P}(|X| \ge 3) \ge 1/2$. Which of the following are true? *1 point*

- $\mathbb{E}[|X|] \ge 3/2$
- $\mathbb{E}[|X|] \le 3/2$
- $\mathbb{P}(|X| \ge 1) \ge 2/3$
- $\mathbb{P}(|X| \ge 4) \ge 1/2$

No, the answer is incorrect.
Score: 0
Accepted Answers:
$\mathbb{E}[|X|] \ge 3/2$

7) Consider the Gaussian mean estimation problem, where, given i.i.d. samples $X_1, \ldots, X_n \sim \mathcal{N}(\mu, 1)$, the task is to estimate $\mu \in \mathbb{R}$. Suppose we have an estimator $\hat{\mu} : (X_1, \ldots, X_n) \to \mathbb{R}$ with the guarantee that $\mathbb{E}_\mu[|\mu - \hat{\mu}(X_1, \ldots, X_n)|] \le 0.1$ for every $\mu \in \mathbb{R}$. *1 point*

Suppose, now, that someone tells us that the samples we are observing are coming either from $\mathcal{N}(0,1)$ or from $\mathcal{N}(1,1)$. Our task is to determine which of these hypotheses is true. Let $H_0 : X_1, \ldots, X_n \sim \mathcal{N}(0,1)$ and $H_1 : X_1, \ldots, X_n \sim \mathcal{N}(1,1)$. Since we already have an estimator $\hat{\mu}$ with us, we design a test $T : (X_1, \ldots, X_n) \to \{0,1\}$ as follows: $T(X_1, \ldots, X_n) = 0$, if $|\hat{\mu}(X_1, \ldots, X_n)| \le 0.5$; $T(X_1, \ldots, X_n) = 1$, if $|\hat{\mu}(X_1, \ldots, X_n)| > 0.5$. Let the probability of error of this test be $P_e(T) = \frac{1}{2}(\mathbb{P}(T = 1|H_0) + \mathbb{P}(T = 0|H_1))$. Which of the following are true?
(You may use the fact that $D(\mathcal{N}(0,1)||\mathcal{N}(1,1)) = \frac{1}{2\ln 2}$, where the $D(\cdot||\cdot)$ is computed using $\log_2$. Furthermore, $d(P,Q) \le \sqrt{\frac{\ln 2}{2} D(P||Q)}$ for distributions P, Q.)

- $P_e(T) \ge 1/3$
- $P_e(T) \le 1/5$
- $P_e(T) \ge \frac{1}{2}\left(1 - \frac{\sqrt{n}}{2}\right)$
- $P_e(T) \le \frac{1}{2}\left(1 - \frac{\sqrt{n}}{2}\right)$

No, the answer is incorrect.
Score: 0
Accepted Answers:
$P_e(T) \le 1/5$
$P_e(T) \ge \frac{1}{2}\left(1 - \frac{\sqrt{n}}{2}\right)$

8) Let X, Y be random variables with joint distribution P_{XY}, where $X \in \mathcal{X}, Y \in \{0,1\}^*$, and X is distributed uniformly over $\mathcal{X}$. Suppose there is a function $f : \{0,1\}^* \to \mathcal{X}$ such that $P_{XY}(X \ne f(Y)) \le 1/4$. Which of the following are true? *1 point*

- $I(X \wedge Y) \le \frac{3}{4}\log|\mathcal{X}| - 1$
- $I(X \wedge Y) \le \frac{3}{4}\log|\mathcal{X}| - 1$
- $I(X \wedge Y) \le H(Y)$
- $I(X \wedge Y) \le \log|\mathcal{X}|$

No, the answer is incorrect.
Score: 0
Accepted Answers:
$I(X \wedge Y) \le \frac{3}{4}\log|\mathcal{X}| - 1$
$I(X \wedge Y) \le H(Y)$
$I(X \wedge Y) \le \log|\mathcal{X}|$

9) Let X, Y be random variables with joint distribution P_{XY}, where $X \in \mathcal{X}, Y \in \{0,1\}^*$. Let $Z = \mathrm{len}(Y)$, where $\mathrm{len}(y)$ denotes the length of the binary string y. Let $\ell = \mathbb{E}[Z]$. Which of the following are true? *1 point*

- $H(Z|Y) = \ell$
- $H(Z|Y) = 0$
- $H(Z|Y) \le \ell$
- $H(Y|Z) \le \sum_{z=0}^\infty z P_Z(z)$

No, the answer is incorrect.
Score: 0
Accepted Answers:
$H(Z|Y) = 0$
$H(Z|Y) \le \ell$
$H(Y|Z) \le \sum_{z=0}^\infty z P_Z(z)$

10) Let $f(x) = x\log x$ for $x > 0$. Which of the following are true? *0 points*
($\ln$ denotes the natural log.)

- $f''(x) \ge 0$ for every $x > 0$
- $f''(x) \le 0$ for every $x > 0$
- $f(x) - f(x-1) \ge \frac{1}{\ln 2}(1 + \ln(x-1))$ for every $x > 1$
- $f(x) - f(x-1) \le \frac{1}{\ln 2}(1 + \ln(x-1))$ for every $x > 1$

No, the answer is incorrect.
Score: 0
Accepted Answers:
$f''(x) \ge 0$ *for every* $x > 0$
$f(x) - f(x-1) \ge \frac{1}{\ln 2}(1 + \ln(x-1))$ *for every* $x > 1$
$f(x) - f(x-1) \le \frac{1}{\ln 2}(1 + \ln(x-1))$ *for every* $x > 1$