Unit 5 - Week 4: Information and statistical inference-1

Weekly Assignment 4

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

Due on 2020-10-14, 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.

Recall some notations from Week 4 lectures. Let $\mathcal{X}, \mathcal{Y}$ be alphabets. W denotes a channel $W : \mathcal{X} \to \mathcal{Y}$, where for each $x \in \mathcal{X}, W_x(\cdot) = P_{Y|X}(\cdot|X = x)$. In the questions on hypothesis testing, it is assumed that the observation is from the alphabet $\mathcal{Y}$ and we want to decide which element in alphabet $\mathcal{X}$ "caused" the observation.

1) Let $\mathcal{X} = \{0,1\}$ and assume that $P_X(0) = P_X(1) = 1/2$. Consider the binary hypothesis testing problem *1 point*
$\mathcal{H}_0 : Y \sim W_0$
$\mathcal{H}_1 : Y \sim W_1$.
Suppose it is known that there is a set $A \subset \mathcal{Y}$ such that $W_0(A) \geq 0.7$ and $W_1(A) \leq 0.5$. Let $P_e^*(\text{unif})$ be the probability of error of the Bayes optimal test. What can we say about $P_e^*(\text{unif})$?

- $P_e^*(\text{unif}) < 0.5$
- $P_e^*(\text{unif}) \leq 0.4$
- $P_e^*(\text{unif}) > 0.6$
- $P_e^*(\text{unif}) \geq 0.8$

No, the answer is incorrect.
Score: 0
Accepted Answers:
$P_e^*(\text{unif}) < 0.5$
$P_e^*(\text{unif}) \leq 0.4$

2) Let $\mathcal{X} = \{1,2,\ldots,M\}$ and let P_X be the uniform distribution over $\mathcal{X}$. *1 point*
Consider an M-ary hypothesis testing problem, where $\mathcal{H}_i : Y \sim W_i$ for $i \in \{1,2,\ldots,M\}$.
Let $A_1,\ldots,A_M$ form a partition of $\mathcal{Y}$, i.e., $\cup_{i=1}^M A_i = \mathcal{Y}$ and $A_i \cap A_j = \phi$ for $i \neq j$.
Consider a test $g : \mathcal{Y} \to \mathcal{X}$ defined as $g(y) = i$ if $y \in A_i$. Let $\hat{X} = g(Y)$. Let $P_e(g|P_X) \triangleq \mathbb{P}(\hat{X} \neq X)$ be the probability of error of the test g. Which of the following are true?

- $P_e(g|P_X) = \sum_{i=1}^M \mathbb{P}(\hat{X} \neq i|X = i)$
- $P_e(g|P_X) = \sum_{i=1}^M \mathbb{P}(\hat{X} \neq X, X = i)$
- $P_e(g|P_X) = \frac{1}{M}\sum_{i=1}^M \mathbb{P}(\hat{X} \neq i|X = i)$
- $P_e(g|P_X) = \frac{1}{M}\sum_{i=1}^M \sum_{j\neq i} W_i(A_j)$

No, the answer is incorrect.
Score: 0
Accepted Answers:
$P_e(g|P_X) = \sum_{i=1}^M \mathbb{P}(\hat{X} \neq X, X = i)$
$P_e(g|P_X) = \frac{1}{M}\sum_{i=1}^M \mathbb{P}(\hat{X} \neq i|X = i)$
$P_e(g|P_X) = \frac{1}{M}\sum_{i=1}^M \sum_{j\neq i} W_i(A_j)$

3) Let $Y \sim \mathcal{N}(\mu, \sigma^2 I_d)$, where $\mathcal{N}(\mu, \sigma^2 I_d)$ denotes the d-dimensional Gaussian distribution with mean vector μ and covariance matrix $\sigma^2 I_d$. Here, μ is a $d \times 1$ column vector with entries $\mu_1,\ldots,\mu_d$, and I_d is the $d \times d$ identity matrix. Let Y_i be the i-th entry of the (random) vector Y. Which of the following are true? ($||\cdot||_2$ denotes the Euclidean norm.) *1 point*

- $Y_i \sim \mathcal{N}(\mu_i, \sigma^2)$
- $Y_i \sim \mathcal{N}(\mu_i, d\sigma^2)$
- $\mathbb{E}[||Y-\mu||_2^2] = \sigma^2$
- $\mathbb{E}[||Y-\mu||_2^2] = d\sigma^2$

No, the answer is incorrect.
Score: 0
Accepted Answers:
$Y_i \sim \mathcal{N}(\mu_i, \sigma^2)$
$\mathbb{E}[||Y-\mu||_2^2] = d\sigma^2$

4) Let $\mathcal{X} = \{0,1\}$ and let $P = (0.5+\epsilon, 0.5-\epsilon)$ be a distribution on $\mathcal{X}$. Using notation from Week 4 lectures, let $P^2 \triangleq P \times P$ be the corresponding product distribution. Which of the following are true? *1 point*

- P^2 is a distribution on $\mathcal{X}$.
- $P^2 = ((0.5+\epsilon)^2, (0.5-\epsilon)^2)$.
- P^2 is a distribution on $\mathcal{X} \times \mathcal{X}$.
- $P^2 = ((0.5+\epsilon)^2, 0.25-\epsilon^2, 0.25-\epsilon^2, (0.5-\epsilon)^2)$.

No, the answer is incorrect.
Score: 0
Accepted Answers:
P^2 is a distribution on $\mathcal{X} \times \mathcal{X}$.
$P^2 = ((0.5+\epsilon)^2, 0.25-\epsilon^2, 0.25-\epsilon^2, (0.5-\epsilon)^2)$.

5) Let $\mathcal{X} = \{0,1\}$, and let $P = (0.5+\epsilon, 0.5-\epsilon)$ and $Q = (0.5-\epsilon, 0.5+\epsilon)$ be distributions on $\mathcal{X}$. Which of the following are true? *1 point*
(d denotes the total variation distance.)

- $d(P,Q) = 2\epsilon$
- $d(P^2,Q^2) = 0$.
- $d(P^2,Q^2) = 2\epsilon$
- $d(P^2,Q^2) \leq 2\, d(P,Q)$

No, the answer is incorrect.
Score: 0
Accepted Answers:
$d(P,Q) = 2\epsilon$
$d(P^2,Q^2) = 2\epsilon$
$d(P^2,Q^2) \leq 2\, d(P,Q)$

6) Let $\mathcal{X} = \{0,1\}$. Consider a binary hypothesis testing problem $\mathcal{H}_0 : Y \sim W_0, \mathcal{H}_1 : Y \sim W_1$. Let $g : \mathcal{Y} \to \mathcal{X}$ be the test defined as follows: *1 point*
$g(y) = 0$ for every $y \in \mathcal{Y}$. Which of the following are true for the test g?

- Probability of Type-I error is 0.
- Probability of Type-I error is 1.
- Probability of Type-II error is 0.
- Probability of Type-II error is 1.

No, the answer is incorrect.
Score: 0
Accepted Answers:
Probability of Type-I error is 0.
Probability of Type-II error is 1.

7) Let $\mathcal{X} = \{0,1\}$. Consider a binary hypothesis testing problem $\mathcal{H}_0 : Y \sim W_0, \mathcal{H}_1 : Y \sim W_1$. Let $\tau > 0$ be such that *1 point*

$$W_0\left(\log \frac{W_0(Y)}{W_1(Y)} \geq \tau\right) \geq 1-\epsilon.$$

Let $g : \mathcal{Y} \to \mathcal{X}$ be the test defined as follows: $g(y) = 0$ if $\log \frac{W_0(y)}{W_1(y)} \geq \tau, g(y) = 1$ otherwise. What can we say about the test g?

- Probability of Type-I error is at most ϵ.
- Probability of Type-II error is at most ϵ.
- Probability of Type-I error is at most $2^{-\tau}$.
- Probability of Type-II error is at most $2^{-\tau}$.

No, the answer is incorrect.
Score: 0
Accepted Answers:
Probability of Type-I error is at most ϵ.
Probability of Type-II error is at most $2^{-\tau}$.

8) Which of the following are true for every $x > -1$? *1 point*
($\ln$ denotes the natural logarithm, $\log$ denotes logarithm to the base 2.)

- $\ln(1+x) \leq x$.
- $\ln(1+x) \leq (\log e)x$
- $\log(1+x) \leq x$
- $\log(1+x) \leq (\log e)x$

No, the answer is incorrect.
Score: 0
Accepted Answers:
$\ln(1+x) \leq x$.
$\ln(1+x) \leq (\log e)x$
$\log(1+x) \leq (\log e)x$

9) Consider two coins -- a fair coin (probability of Heads 0.5) and a biased coin (probability of Heads 0.6). One coin is picked uniformly at random and you have to decide whether the coin is fair or biased. You design a test which outputs the correct answer with probability at least 95/100. You are unwilling to reveal anything about how your test works; in particular, you don't want to tell how many coin tosses your test needs to make its decision. Let r be the number of coin tosses your test needs to make its decision. What can still be said about r? *1 point*
(Assume following fact is already known: For $P = \text{Ber}(0.5), Q = \text{Ber}(0.6)$, $d(P^n, Q^n) \leq 0.15\sqrt{n}$.)

- $r \geq 70$
- $r \geq 35$
- $r \leq 105$
- Nothing can be claimed as such, unless the test is revealed

No, the answer is incorrect.
Score: 0
Accepted Answers:
$r \geq 35$

10) Let $P \equiv \mathcal{N}(-1,1)$ and $Q \equiv \mathcal{N}(1,1)$, where $\mathcal{N}(\mu,\sigma^2)$ denotes the Gaussian distribution on $\mathbb{R}$ with mean μ and variance σ^2. Which of the following are true? *1 point*

- $d(P,Q) = 1$
- $d(P,Q) = \int_{-\infty}^{1} e^{-x^2/2}dx - \int_1^{\infty} e^{-x^2/2}dx$
- $D(P||Q) = 2$
- $D(P||Q) = 4$

No, the answer is incorrect.
Score: 0
Accepted Answers:
$d(P,Q) = \int_{-\infty}^{1} e^{-x^2/2}dx - \int_1^{\infty} e^{-x^2/2}dx$
$D(P||Q) = 2$