5. Total Variation Distance for Discrete Random Variables

Quiz: Probability Mass Functions

0/1 point (graded)
Let X be a discrete random variable whose sample space is $\mathbb{Z}$, the set of integers. Let $p : \mathbb{Z} \to [0,1]$ denote the **probability mass function** (**pmf**) of X. What does $p(7) + p(10)$ represent?

- The probability that $X = 10$.
- The probability that $X = 7$.
- The probability that $X = 7$ **or** $X = 10$.
- (selected) The probability that $X = 7$ **and** $X = 10$.

Solution:

By definition, $p(7) + p(10) = P(X = 10) + P(X = 7)$. The events $X = 10$ and $X = 7$ are disjoint, so in fact $p(7) + p(10) = P(X = 10 \text{ or } X = 7)$.

提交 你已经尝试了1次（总共可以尝试1次）

Answers are displayed within the problem

Preparation: Probability of Complements

1/1 point (graded)
What is $\mathbf{P}_\theta(A^c) - \mathbf{P}_{\theta'}(A^c)$ in terms of $\mathbf{P}_\theta(A)$ and $\mathbf{P}_{\theta'}(A)$? (Recall A^c is the complement of A in the sample space.)

- (selected) $\mathbf{P}_{\theta'}(A) - \mathbf{P}_\theta(A)$
- $\mathbf{P}_\theta(A) - \mathbf{P}_{\theta'}(A)$

Solution:

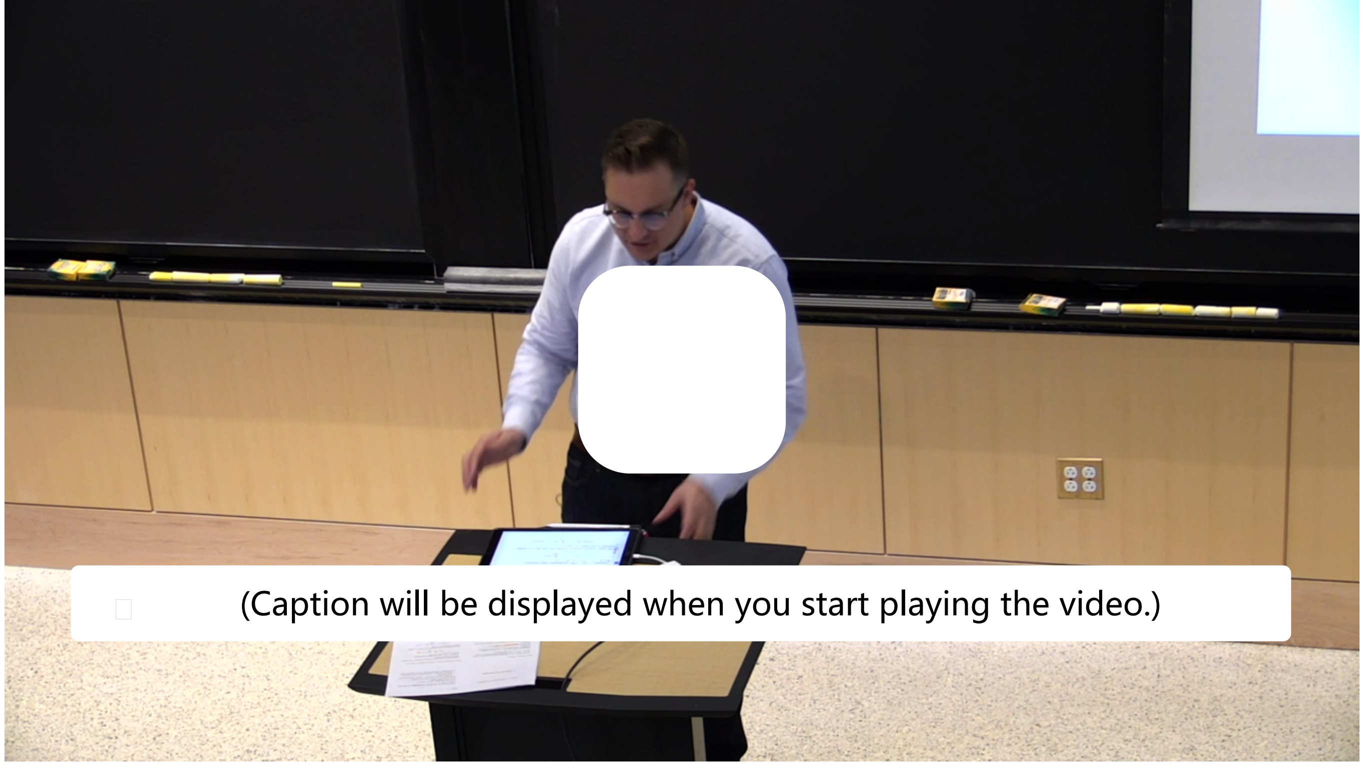
$$\mathbf{P}_\theta(A^c) - \mathbf{P}_{\theta'}(A^c) = (1 - \mathbf{P}_\theta(A)) - (1 - \mathbf{P}_{\theta'}(A)) = \mathbf{P}_{\theta'}(A) - \mathbf{P}_\theta(A).$$

提交 你已经尝试了1次（总共可以尝试1次）

Answers are displayed within the problem

Total Variation Distance for Discrete Distributions

Start of transcript. Skip to the end.



So actually let's just do a little bit of algebra.

[INAUDIBLE]

And so, yeah, we'll just do a little bit of algebra to see--

I'm not going prove you this result, but I'm actually--

so remember the total variation between say p--

let's just write it p--

well, p theta and theta prime.

视频
[下载视频文件]

字幕
[下载 SubRip (.srt) file]
[下载 Text (.txt) file]

Let $\mathbf{P}$ and $\mathbf{Q}$ be probability measures with a discrete sample space E and probability mass functions f and g. Then, the total variation distance between $\mathbf{P}$ and $\mathbf{Q}$:

$$\text{TV}\left(\mathbf{P},\mathbf{Q}\right) = \max_{A\subset E}|\mathbf{P}\left(A\right) - \mathbf{Q}\left(A\right)|,$$

can be computed as

$$\text{TV}\left(\mathbf{P},\mathbf{Q}\right) = \frac{1}{2}\sum_{x\in E}|f\left(x\right) - g\left(x\right)|.$$

Equivalence of Formulas

4/4 points (graded)
Let $E = \{1,2,3,4\}$ be a discrete sample space. Let $\mathbf{P}$ and $\mathbf{Q}$ be probability measures with probability mass functions f and g as follows:

$$f\left(1\right) = 1/4,\ f\left(2\right) = 1/4,\ f\left(3\right) = 1/8,\ f\left(4\right) = 3/8$$

$$g\left(1\right) = g\left(2\right) = g\left(3\right) = g\left(4\right) = 1/4$$

Find the value of $|\mathbf{P}\left(A\right) - \mathbf{Q}\left(A\right)|$ for the following choices of A.

For $A = \{3\}$:

$|\mathbf{P}\left(A\right) - \mathbf{Q}\left(A\right)| =$ 1/8 **Answer:** 1/8

For $A = \{4\}$:

$|\mathbf{P}\left(A\right) - \mathbf{Q}\left(A\right)| =$ 1/8 **Answer:** 1/8

For $A = \{3,4\}$?

$|\mathbf{P}(A) - \mathbf{Q}(A)| =$ 0 Answer: 0

What is the value of $\max_{A \subset E} |\mathbf{P}(A) - \mathbf{Q}(A)|$?

1/8 Answer: 1/8

STANDARD NOTATION

Solution:

First, compute $|\mathbf{P}(A) - \mathbf{Q}(A)|$ for the different choices of A:

- When $A = \{3\}$, $\mathbf{P}(A) = f(3) = 1/8$ and $\mathbf{Q}(A) = g(3) = 1/4$. Therefore, $|\mathbf{P}(A) - \mathbf{Q}(A)| = 1/8$.
- When $A = \{4\}$, $\mathbf{P}(A) = f(4) = 3/8$ and $\mathbf{Q}(A) = g(4) = 1/4$. Therefore, $|\mathbf{P}(A) - \mathbf{Q}(A)| = 1/8$.
- When $A = \{3,4\}$, $\mathbf{P}(A) = f(3) + f(4) = 1/2$ and $\mathbf{Q}(A) = g(3) + g(4) = 1/2$. Therefore, $|\mathbf{P}(A) - \mathbf{Q}(A)| = 0$.

Now, we find $\max_{A \subset E} |\mathbf{P}(A) - \mathbf{Q}(A)|$. We have already considered $A = \{3\}$, $A = \{4\}$, and $A = \{3,4\}$. For any other non-empty set A, $|\mathbf{P}(A) - \mathbf{Q}(A)|$ takes on one of the values that we have already computed because $f(1) = f(2) = g(1) = g(2) = 1/4$.

In particular, for any set that includes 3 but does not include 4, $|\mathbf{P}(A) - \mathbf{Q}(A)| = |-1/8| = 1/8$. For any set that includes 4 but does not include 3, $|\mathbf{P}(A) - \mathbf{Q}(A)| = |1/8| = 1/8$. And finally, for any set that includes both 3 and 4, $|\mathbf{P}(A) - \mathbf{Q}(A)| = 0$.

Therefore, $\max_{A \subset E} |\mathbf{P}(A) - \mathbf{Q}(A)| = 1/8$, with the maximum achieved with numerous sets as discussed above.

提交 你已经尝试了1次（总共可以尝试2次）

Answers are displayed within the problem

Equivalence of Formulas (cont.)

1/1 point (graded)

Setup as above:

Let $E = \{1,2,3,4\}$ be a discrete sample space. Let $\mathbf{P}$ and $\mathbf{Q}$ be probability measures with probability mass functions f and g as follows:

$$f(1) = 1/4, \; f(2) = 1/4, \; f(3) = 1/8, \; f(4) = 3/8$$

$$g(1) = g(2) = g(3) = g(4) = 1/4$$

Question: What is the value of $\frac{1}{2} \sum_{x \in E} |f(x) - g(x)|$?

1/8 Answer: 1/8

Solution:

$$\frac{1}{2} \sum_{x \in E} |f(x) - g(x)| = \frac{1}{2}\left(0 + 0 + \frac{1}{8} + \frac{1}{8}\right) = \frac{1}{8}.$$

This is the same result as in the previous problem.

提交 你已经尝试了1次（总共可以尝试2次）

Answers are displayed within the problem

Computing Total Variation Distance I

1/1 point (graded)
Let $X \sim \mathbf{P} = \text{Ber}(1/2)$ and $Y \sim \mathbf{Q} = \text{Ber}(1/2)$. What is $\text{TV}(\mathbf{P}, \mathbf{Q})$, the total variation distance between the distributions of the Bernoulli random variables X and Y?

Note that we make no assumptions about X and Y being independent.

0

Answer: 0.0

Solution:

Intuitively, since X and Y have the same distribution, we expect the (total variation) distance between their distributions to be 0. And indeed this is the case. Observe that for any event, $\mathbf{P}(A) = \mathbf{Q}(A)$ since $\mathbf{P}$ and $\mathbf{Q}$ are both $\text{Ber}(1/2)$.

$$\text{TV}(\mathbf{P}, \mathbf{Q}) = \max_{A \subset E} |\mathbf{P}(A) - \mathbf{Q}(A)| = 0.$$

Note that the distance between two distributions only depends on the distributions themselves and *not* their relation to each other (the joint distribution). This is why assuming X and Y are independent (or not) does not affect the total variation distance.

提交 你已经尝试了1次（总共可以尝试3次）

Answers are displayed within the problem

Computing Total Variation II

1/1 point (graded)
Let $X \sim \mathbf{P} = \text{Ber}(1/2)$ and $Y \sim \mathbf{Q} = \text{Ber}(1/3)$. What is $\text{TV}(\mathbf{P}, \mathbf{Q})$, the total variation distance between the distributions of the Bernoulli random variables X and Y?

1/6

Answer: 1/6

Solution:

For this problem, the sample space of X and Y is $\{0, 1\}$. Let f be the pmf of X and let g be the pmf of Y. Note that $f(1) = f(0) = 1/2$ and $g(1) = 1/3, g(0) = 2/3$. Hence, we can apply the given formula:

$$\begin{aligned} \text{TV}(\mathbf{P}, \mathbf{Q}) &= \frac{1}{2} \sum_{x \in E} |f(x) - g(x)| \\ &= \frac{1}{2}(|f(0) - g(0)| + |f(1) - g(1)|) \\ &= \frac{1}{2}(1/6 + 1/6) = 1/6 \approx 0.16667. \end{aligned}$$

Remark: In general, we have the formula

$$\text{TV}(\text{Ber}(p), \text{Ber}(q)) = |p - q|.$$

提交 你已经尝试了1次（总共可以尝试10次）

Answers are displayed within the problem

讨论

显示讨论