2. Compute Total Variation Distance

(a)

3/3 points (graded)

Compute the total variation distance between

$$\mathbf{P} = X \quad \text{and} \quad \mathbf{Q} = X + c, \quad \text{where } X \sim \mathsf{Ber}(p), p \in (0,1), \text{and } c \in \mathbb{R}.$$

(If applicable, enter **abs(x)** for $|x|$. Simplify your answer to have the minimum number of absolute signs possible.)

For $c \notin \{-1, 0, 1\}$:

$\mathsf{TV}(\mathbf{P}, \mathbf{Q}) =$ 1

1

Answer: 1

For $c = 0$:

$\mathsf{TV}(\mathbf{P}, \mathbf{Q}) =$ 0

0

Answer: 0

For $c = 1$ or $c = -1$:

$\mathsf{TV}(\mathbf{P}, \mathbf{Q}) =$ (abs(2*p-1)+1)/2

$\frac{\text{abs}(2\cdot p-1)+1}{2}$

Answer: 1/2*(1+abs(1-2*p))

STANDARD NOTATION

Solution:

- For $c \notin \{-1, 0, 1\}$, the support of X and $X + c$ are disjoint, hence $\mathsf{TV}(X, X + c) = 1$.
- For $c = 0$, by the definiteness property $\mathsf{TV}(X, X) = 0$.
- For $c = 1$ (*resp.* $c = -1$), the support of X and $X + c$ intersect at $X = 1$ (*resp.* at $X = 0$). Hence

$$\begin{aligned} \mathsf{TV}(X, X+c) &= \frac{1}{2}(|1-p| + |p - (1-p)| + |p|) \\ &= \frac{1}{2}(1 + |1-2p|) \qquad \text{where } c = 1, \text{ or } -1. \end{aligned}$$

提交 你已经尝试了2次（总共可以尝试2次）

Answers are displayed within the problem

(b)

2/2 points (graded)

Compute the total variation distance between

$$\mathbf{P} = \text{Ber}(p) \quad \text{and} \quad \mathbf{Q} = \text{Ber}(q), \quad \text{where } p, q \in [0,1].$$

(If applicable, enter **abs(x)** for $|x|$.)

$\text{TV}(\mathbf{P}, \mathbf{Q}) =$ abs(p-q) Answer: abs(p-q)

$\text{abs}(p-q)$

Let $X_1, \ldots, X_n$ be n i.i.d. Bernoulli random variables with some parameter $p \in [0,1]$, and $\bar{X}_n$ be their empirical average. Consider the total variation distance $\text{TV}(\text{Ber}(\bar{X}_n), \text{Ber}(p))$ between $\text{Ber}(\bar{X}_n)$ and $\text{Ber}(p)$ as a function of the random variable $\bar{X}_n$, and hence a random variable itself. Does $\text{TV}(\text{Ber}(\bar{X}_n), \text{Ber}(p))$ necessarily converge in probability to a constant? If yes, enter the constant below; if not; enter DNE.

$\text{TV}(\text{Ber}(\bar{X}_n), \text{Ber}(p)) \xrightarrow[n\to\infty]{(\mathbf{P})}$ 0 Answer: 0

0

STANDARD NOTATION

Solution:

To compute the total variation distance between two Bernoulli variables, again use the formula relating TV to the pmfs:

$$\begin{aligned}\text{TV}(\text{Ber}(p), \text{Ber}(q)) &= \frac{1}{2}[|p-q| + |(1-p)-(1-q)|] \\ &= |p-q|.\end{aligned}$$

Let $X_1, \ldots, X_n$ be n i.i.d. Bernoulli random variables with some parameter $p \in [0,1]$, and $\bar{X}_n$ be their empirical average. By the Law of Large Numbers, we know that $\bar{X}_n$ will converge to p. Now, imagine another Bernoulli distribution with parameter $\bar{X}_n$, say $\text{Ber}(\bar{X}_n)$. What can we infer about the two distributions? What is the total variation distance between $\text{Ber}(\bar{X}_n)$ and $\text{Ber}(p)$? Intuitively, the two distribution should behave similarly since

$$\bar{X}_n \xrightarrow[n\to\infty]{i.p.} p.$$

Recall that by definition, the convergence in probability means

$$P(|\bar{X}_n - p| > \epsilon) \xrightarrow[n\to\infty]{} 0.$$

Remember that the total variation distance between $\text{Ber}(q)$ and $\text{Ber}(p)$ is $|q-p|$. We want to calculate the total variation distance between $\text{Ber}(\bar{X}_n)$ and $\text{Ber}(p)$, that is, $|\bar{X}_n - p|$. This is the same as what we've seen above! By the Law of Large Numbers, we can say that the total variation distance will converge in probability to 0 as n goes to infinity.

$$\text{TV}(\text{Ber}(\bar{X}_n), \text{Ber}(p)) = |\bar{X}_n - p| \xrightarrow[n\to\infty]{p.} 0.$$

提交 你已经尝试了2次（总共可以尝试2次）

Answers are displayed within the problem

(c)

1/1 point (graded)
Compute the total variation distance between

$$P = \text{Unif}([0, s]) \quad \text{and} \quad Q = \text{Unif}([0, t]), \quad \text{where } 0 < s < t.$$

$\text{TV}(\mathbf{P}, \mathbf{Q}) =$ 1/2*(1/s-1/t)*s+1/2*1/t* Answer: 1 - s/t

$\frac{1}{2} \cdot \left(\frac{1}{s} - \frac{1}{t}\right) \cdot s + \frac{1}{2} \cdot \frac{1}{t} \cdot (t - s)$

STANDARD NOTATION

Solution:

To compute the total variation distance between two uniform distributions, denote the densities of the two distributions by

$$f_s(x) = \frac{1}{s}\mathbf{1}\{0 \le x \le s\}, \quad f_t(x) = \frac{1}{t}\mathbf{1}\{0 \le x \le t\}.$$

With this, we have

$$\begin{aligned} \text{TV}(\text{Unif}([0, s]), \text{Unif}([0, t])) &= \frac{1}{2}\int_{\mathbb{R}} |f_s(x) - f_t(x)|\, dx \\ &= \frac{1}{2}\left[\int_0^s \left|\frac{1}{s} - \frac{1}{t}\right| dx + \int_s^t \left|\frac{1}{t}\right| dx\right] \\ &= \frac{1}{2}\left[\left(1 - \frac{s}{t}\right) + \left(1 - \frac{s}{t}\right)\right] \\ &= 1 - \frac{s}{t}. \end{aligned}$$

Hence, $\text{TV}(\text{Unif}([0, s]), \text{Unif}([0, t]))$ is a continuous function in t that decreases to 0 as t approaches s.

提交 你已经尝试了1次（总共可以尝试2次）

Answers are displayed within the problem

(d)

1.0/1 point (graded)

Let $X \sim N(\mu, \sigma^2)$ and $Y \sim \text{Ber}(p)$. Compute the total variation distance between the distributions of $\text{sign}(X)$ and $Y - 1$. Note that $\text{sign}(X)$ is a function of the random variable with

$$\text{sign}(X) = \begin{cases} 1 & \text{if } X > 0 \\ 0 & \text{if } X = 0 \\ -1 & \text{if } X < 0. \end{cases}$$

(If applicable, enter **abs(x)** for $|x|$, **Phi(x)** for $\Phi(x) = \mathbf{P}(Z \le x)$ where $Z \sim \mathcal{N}(0, 1)$, and **q(alpha)** for q_α, the $1 - \alpha$-quantile of a standard normal distribution, e.g. enter **q(0.01)** for $q_{0.01}$.)

$\text{TV}(\text{sign}(X), Y - 1) =$ 1/2*(abs(1-p-Phi(-mu/sigma))+p+1-Phi(-mu/sigma))

Answer: 0.5*(Phi(mu/sigma)+p+abs(1-p-Phi(-mu/sigma)))

STANDARD NOTATION

Solution:

Observe that $\frac{X - \mu}{\sigma} \sim \mathcal{N}(0, 1)$. Hence

$$\text{sign}(X) = \begin{cases} -1 & \text{with probability } \Phi\left(-\frac{\mu}{\sigma}\right) \\ 1 & \text{with probability } 1 - \Phi\left(-\frac{\mu}{\sigma}\right) = \Phi\left(\frac{\mu}{\sigma}\right) \end{cases}$$

Hence,

$$\begin{aligned} 2\mathsf{TV}\left(\mathrm{sign}\left(X\right), Y-1\right) &= \left|\Phi\left(\frac{\mu}{\sigma}\right)\right| + |p| + \left|(1-p) - \Phi\left(-\frac{\mu}{\sigma}\right)\right| \\ &= \Phi\left(\frac{\mu}{\sigma}\right) + p + \left|1-p-\Phi\left(-\frac{\mu}{\sigma}\right)\right|. \end{aligned}$$

提交 你已经尝试了1次（总共可以尝试3次）

Answers are displayed within the problem

(e)

0/1 point (graded)

Compute the total variation distance between

$$\mathbf{P} = \mathsf{Ber}\left(p\right) \quad \text{and} \quad \mathbf{Q} = \mathsf{Poiss}\left(p\right), \quad \text{where } p \in (0,1)\,.$$

$\mathsf{TV}\left(\mathbf{P},\mathbf{Q}\right) =$ 1/2*(p*(1-exp(-p))+abs(exp(-p)-1+p)) Answer: p*(1-e^(-p))

$\frac{1}{2}\cdot\left(p\cdot\left(1-\exp\left(-p\right)\right)+\mathrm{abs}\left(\exp\left(-p\right)-1+p\right)\right)$

没考虑到support漏了一项；而且没化简：e^(-p) > 1 - p when p > 0

STANDARD NOTATION

Solution:

Recall the pmf $f_X\left(x\right)$ for $X \sim \mathsf{Poiss}\left(p\right)$ is

$$f_X\left(x\right) = e^{-p}\frac{p^x}{x!} \qquad \text{for } x = 0, 1, 2 \ldots.$$

Hence,

$$\begin{aligned} 2\mathsf{TV}\left(\mathsf{Ber}\left(p\right), \mathsf{Poiss}\left(p\right)\right) &= \left|e^{-p} - (1-p)\right| + \left|pe^{-p} - p\right| + e^{-p}\left(\frac{p^2}{2!} + \frac{p^3}{3!} + \cdots\right) \\ &= \left(e^{-p} - (1-p)\right) + \left(p\left(1-e^{-p}\right)\right) + e^{-p}\left(e^p - (1+p)\right) \quad \text{since } e^{-p} > (1-p) \text{ for } p > 0 \\ &= 2\left(p\left(1-e^{-p}\right)\right). \\ \iff \mathsf{TV}\left(\mathsf{Ber}\left(p\right), \mathsf{Poiss}\left(p\right)\right) &= p\left(1-e^{-p}\right). \end{aligned}$$

提交 你已经尝试了2次（总共可以尝试2次）

Answers are displayed within the problem

讨论

显示讨论

主题: Unit 3 Methods of Estimation:Homework 4: TV distance, KL-Divergence, and Introduction to MLE / 2. Compute Total Variation Distance

认证证书是什么?