1. Kullback-Leibler divergence

Instructions:

For the following pairs of distributions $(\mathbf{P}, \mathbf{Q})$, compute the Kullback-Leibler divergence $\text{KL}(\mathbf{P}, \mathbf{Q})$.

If the KL divergence is $+\infty$ or $-\infty$, enter **+inf** or **-inf**.

(a)

1/1 point (graded)

$$\mathbf{P} = \mathcal{N}(a, \sigma^2), \quad \mathbf{Q} = \mathcal{N}(b, \sigma^2), \quad a, b \in \mathbb{R}, \, \sigma^2 > 0.$$

(If applicable, enter **ln(x)** for $\ln(x)$. Do NOT enter "log".)

$\text{KL}(\mathbf{P}, \mathbf{Q}) =$ (a-b)^2/(2*sigma^2)

$\frac{(a-b)^2}{2\cdot\sigma^2}$

Answer: (a - b)^2/(2*sigma^2)

STANDARD NOTATION

Solution:

If we write $X \sim \mathbf{P}$, we can compute:

$$\begin{aligned}\text{KL}(\mathbf{P}, \mathbf{Q}) &= \mathbb{E}_{\mathbf{P}}\left[\ln\left(\frac{\frac{1}{\sqrt{2\pi}}\exp\left(-\frac{-(X-a)^2}{2\sigma^2}\right)}{\frac{1}{\sqrt{2\pi}}\exp\left(-\frac{(X-b)^2}{2\sigma^2}\right)}\right)\right] \\ &= \mathbb{E}_{\mathbf{P}}\left[-\frac{(X-a)^2}{2\sigma^2} + \frac{(X-b)^2}{2\sigma^2}\right] \\ &= \frac{1}{2\sigma^2}\mathbb{E}_{\mathbf{P}}\left[2(a-b)(X-a) + (a-b)^2\right] \\ &= \frac{(a-b)^2}{2\sigma^2},\end{aligned}$$

because $\mathbb{E}_{\mathbf{P}}[(X-a)] = 0$.

提交 你已经尝试了2次（总共可以尝试2次）

Answers are displayed within the problem

(b)

1/1 point (graded)

$$\mathbf{P} = \text{Ber}(a), \quad \mathbf{Q} = \text{Ber}(b), \quad a, b \in (0, 1)$$

(If applicable, enter **ln(x)** for $\ln(x)$. Do NOT enter "log".)

$\textsf{KL}(\mathbf{P}, \mathbf{Q}) =$ (1-a)*ln((1-a)/(1-b))+a*ln(a/b)

$(1-a)\cdot\ln\left(\frac{1-a}{1-b}\right) + a\cdot\ln\left(\frac{a}{b}\right)$

Answer: a * ln(a/b) + (1-a) *ln((1-a)/(1-b))

STANDARD NOTATION

Solution:

If we write $X \sim \mathbf{P}$, $Y \sim \mathbf{Q}$, we have

$$\begin{aligned}\textsf{KL}(\mathbf{P}, \mathbf{Q}) &= \mathbf{P}(X=0)\ln\frac{\mathbf{P}(X=0)}{\mathbf{P}(Y=0)} + \mathbf{P}(X=1)\frac{\mathbf{P}(X=1)}{\mathbf{P}(Y=1)}\\ &= a\ln\frac{a}{b} + (1-a)\ln\frac{1-a}{1-b}.\end{aligned}$$

提交 你已经尝试了1次（总共可以尝试3次）

Answers are displayed within the problem

(c)

1/2 points (graded)

$$P = \textsf{Unif}([0,\theta_1]), \quad Q = \textsf{Unif}([0,\theta_2]), \quad 0 < \theta_1 < \theta_2.$$

Hint: Note the support of each distribution when computing the expectation.

(If applicable, enter **ln(x)** for $\ln(x)$. Do NOT enter "log". If applicable, enter **theta_1** for θ_1 and **theta_2** for θ_2.)

$\textsf{KL}(\mathbf{P}, \mathbf{Q}) =$ ln(theta_2/theta_1)/theta_1*theta_1

$\frac{\ln\left(\frac{\theta_2}{\theta_1}\right)}{\theta_1}\cdot\theta_1$

Answer: ln(theta_2/theta_1)

$\textsf{KL}(\mathbf{Q}, \mathbf{P}) =$ ln(theta_1/theta_2)/theta_2*theta_1

$\frac{\ln\left(\frac{\theta_1}{\theta_2}\right)}{\theta_2}\cdot\theta_1$

Answer: inf

没考虑到support，漏了一项

STANDARD NOTATION

Solution:

We compute

$$\begin{aligned}\textsf{KL}(\mathbf{P}, \mathbf{Q}) &= \mathbb{E}_{\mathbf{P}}\left[\ln\frac{\frac{1}{\theta_1}}{\frac{1}{\theta_2}}\right]\\ &= \ln\left(\frac{\theta_2}{\theta_1}\right).\end{aligned}$$

If we try to compute the $\textsf{KL}$ divergence the other way round, we notice that $\mathbf{P}$ is not supported between for $\theta_1 < X < \theta_2$. We compute the expectation by integrating explicitly:

$$\textsf{KL}(\mathbf{Q}, \mathbf{P}) = \mathbb{E}_{\mathbf{Q}}\left[\ln\frac{q}{p}\right] \quad \text{where } p, q, \text{are the pdfs of } \mathbf{P}, \mathbf{Q} \text{ respectively}$$

$$= \int_0^{\theta_1} \frac{1}{\theta_2} \ln \frac{1/\theta_2}{1/\theta_1} dx + \int_{\theta_1}^{\theta_2} \frac{1}{\theta_2} \ln \frac{1/\theta_2}{0} dx$$

$$= +\infty$$

because the second term diverges to $+\infty$. **Remark:** In general, $\text{KL}(\mathbf{P}, \mathbf{Q}) \neq \text{KL}(\mathbf{Q}, \mathbf{P})$.

提交 你已经尝试了3次（总共可以尝试3次）

Answers are displayed within the problem

(d)

0/1 point (graded)

$$P = \text{Exp}(\lambda), \quad Q = \text{Exp}(\mu), \quad \lambda, \mu \in (0, \infty).$$

(If applicable, enter **ln(x)** for $\ln(x)$. Do NOT enter "log".)

$\text{KL}(\mathbf{P}, \mathbf{Q}) =$ +inf

$+inf$

Answer: ln(lambda/mu) + mu/lambda - 1

没用期望做

STANDARD NOTATION

Solution:

If $X \sim P$, then

$$\begin{aligned}
\text{KL}(\mathbf{P}, \mathbf{Q}) = & \; \mathbb{E}_P \left[\ln \frac{\lambda e^{-\lambda x}}{\mu e^{-\mu x}} \right] \\
= & \; \mathbb{E}_P \left[\ln \frac{\lambda}{\mu} + (\mu - \lambda) X \right] \\
= & \; \ln \frac{\lambda}{\mu} + (\mu - \lambda) \frac{1}{\lambda} \\
= & \; \ln \frac{\lambda}{\mu} + \frac{\mu}{\lambda} - 1,
\end{aligned}$$

because $\mathbb{E}_P[X] = \frac{1}{\lambda}$.

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 / 1. Kullback-Leibler divergence

认证证书是什么?