

Find solutions for your homework

Search

home / study / math / statistics and probability / statistics and probability questions and answers / 2. consider that  $p = N(\mu, \Sigma)$  and  $q = N(0, I)$ . here  $I = \text{diag}(1, 1, \dots, 1)$

**Question: 2. Consider that  $p = N(\mu, \Sigma)$  and  $q = N(0, I)$ . Here  $I = \text{diag}(1, 1, \dots, 1)$**

2. Consider that  $p = N(\mu, \Sigma)$  and  $q = N(0, I)$ . Here  $\Sigma = \text{diag}(\sigma_1, \sigma_2, \dots, \sigma_k)$ . Then, shown that the KL divergence between  $p$  and  $q$  is defined as below.

$$D_{KL}(p||q) = \frac{1}{2} \sum_{i=1}^k (\sigma_i^2 + \mu_i^2 - 1 - \log_e(\sigma_i^2)).$$

Show transcribed image text

## Expert Answer



Anonymous answered this  
58 answers

Was this answer helpful?

 1

 0

Ans.

KL divergence between two distribution  $p$  &  $q$  of a continuous random variable is given by.

$$D_{KL}(p||q) = \int_{\mathcal{X}} p(x) \log \frac{p(x)}{q(x)} dx$$

$p = N(\mu, \Sigma)$  and  $q = N(0, I)$

$$p(x) = \frac{1}{(2\pi)^{k/2} |\Sigma|^{1/2}} \exp\left(-\frac{1}{2}(x-\mu)^T \Sigma^{-1}(x-\mu)\right)$$

$$D_{KL}(p||q) = E_p[\log(p) - \log(q)]$$

$$= E_p\left[\frac{1}{2} \log \frac{|q|}{|\Sigma|} - \frac{1}{2}(x-\mu)^T \Sigma^{-1}(x-\mu) + \frac{1}{2}(x-\mu)^T \Sigma^{-1}(x-\mu)\right]$$

$$= \frac{1}{2} \log \frac{|q|}{|\Sigma|} - \frac{1}{2} E_p[(x-\mu)^T \Sigma^{-1}(x-\mu)] + \frac{1}{2} E_p[(x-\mu)^T \Sigma^{-1}(x-\mu)]$$

## Post a question

Answers from our experts for your toughest homework questions

Enter question

Continue to post

20 questions remaining



**Snap a photo from your phone to post a question**

We'll send you a one-time download link

888-888-8888

Text n

By providing your phone number, you agree to receive a automated text message with a link to get the app. Standard messaging rates may apply.

## My Textbook Solutions



Finite...

9th Edition



Calculus for...

1st Edition



Introduc

10th Editi

[View all solutions](#)

Now,

$$\begin{aligned} (\underline{x} - \underline{\mu}_p)^T \underline{\Sigma}_p^{-1} (\underline{x} - \underline{\mu}_p) &= \text{tr} \left\{ (\underline{x} - \underline{\mu}_p)^T \underline{\Sigma}_p^{-1} (\underline{x} - \underline{\mu}_p) \right\} \\ &= \text{tr} \left\{ (\underline{x} - \underline{\mu}_p) (\underline{x} - \underline{\mu}_p)^T \underline{\Sigma}_p^{-1} \right\} \end{aligned}$$

Also

$$\begin{aligned} \frac{1}{2} E_p \left[ \text{tr} \left\{ (\underline{x} - \underline{\mu}_p) (\underline{x} - \underline{\mu}_p)^T \underline{\Sigma}_p^{-1} \right\} \right] \\ &= \frac{1}{2} \text{tr} \left\{ E_p \left[ (\underline{x} - \underline{\mu}_p) (\underline{x} - \underline{\mu}_p)^T \underline{\Sigma}_p^{-1} \right] \right\} \\ &= \frac{1}{2} \text{tr} \left\{ E_p \left[ (\underline{x} - \underline{\mu}_p) (\underline{x} - \underline{\mu}_p)^T \underline{\Sigma}_p^{-1} \right] \right\} \end{aligned}$$

We know,

$$\begin{aligned} E_p \left[ (\underline{x} - \underline{\mu}_p) (\underline{x} - \underline{\mu}_p)^T \right] &= \underline{\Sigma}_p \\ &= \frac{1}{2} \text{tr} \left\{ \underline{\Sigma}_p \underline{\Sigma}_p^{-1} \right\} \\ &= \frac{1}{2} \text{tr} \left\{ \underline{I}_k \right\} \\ &= k/2. \end{aligned}$$

Similarly, on simplifying.

$$E_p \left[ (\underline{x} - \underline{\mu}_p)^T \underline{\Sigma}_q^{-1} (\underline{x} - \underline{\mu}_q) \right] = (\underline{\mu}_p - \underline{\mu}_q)^T \underline{\Sigma}_q^{-1} (\underline{\mu}_p - \underline{\mu}_q) + \text{tr} \{ \underline{\Sigma}_q^{-1} \underline{\Sigma}_p \}.$$

Now combining all we get.

$$D_{KL}(p||q) = \frac{1}{2} \left[ \log \frac{|\underline{\Sigma}_q|}{|\underline{\Sigma}_p|} - k + (\underline{\mu}_p - \underline{\mu}_q)^T \underline{\Sigma}_q^{-1} (\underline{\mu}_p - \underline{\mu}_q) + \text{tr} \{ \underline{\Sigma}_q^{-1} \underline{\Sigma}_p \} \right].$$

When  $q$  is  $N(\underline{0}, \underline{I})$  we get.

$$D_{KL}(p||q) = \frac{1}{2} \left[ \underline{\mu}_p^T \underline{\mu}_p + \text{tr} \{ \underline{\Sigma}_p \} - k - \log_e |\underline{\Sigma}_p| \right]$$

We have  $\underline{\Sigma} = \text{diag}(\sigma_1^2, \sigma_2^2, \dots, \sigma_k^2)$ .

$$\underline{\mu}_p^T \underline{\mu}_p = \sum_{i=1}^k \mu_i^2$$

$$\log_e |\underline{\Sigma}_p| = \sum_{i=1}^k \log_e (\sigma_i^2)$$

Putting these value in expression of  $D_{KL}$  we get.

$$D_{KL}(p||q) = \frac{1}{2} \sum_{i=1}^k (\sigma_i^2 + \mu_i^2 - 1 - \log_e (\sigma_i^2)).$$

Proved.

Comment ➤

### Questions viewed by other students


Q: QUESTION 4 "in the trp operon, if there are high levels of tryptophan available" the trp operon is being transcribed at relatively high levels tryptophan is inactivating the repressor protein the trp operon is being transcribed at low levels lactose is inactivating the repressor protein QUESTIONS In the lac operon the lacI gene codes for: the lac inducer, responsible for negative...

A: [See answer](#) 100% (1 rating)

Q: Definition of Kullback-Leibler (KL) Divergence Let  $P$  and  $Q$  be discrete probability distributions with pmfs  $p$  and  $q$  respectively. Let's also assume  $P$  and  $Q$  have a common sample space  $E$ . Then the KL divergence (also known as

A: [See answer](#)

92% (13 ratings)

[Show more](#) 

COMPANY

LEGAL & POLICIES

CHEGG PRODUCTS AND SERVICES

CHEGG NETWORK

CUSTOMER SERVICE

