

How can I calculate KL Divergence in R, and can you provide an example?

Authored by
stats writer

June 26, 2024

RECOMMENDED CITATION

stats writer (2024). *How can I calculate KL Divergence in R, and can you provide an example?*. PSYCHOLOGICAL SCALES. Retrieved from <https://scales.arabpsychology.com/?p=153617>

KL Divergence, also known as Kullback-Leibler Divergence, is a measure of how different two probability distributions are from each other. In R, the KL Divergence can be calculated using the "KLdiv" function from the "entropy" package. This function takes two probability distributions as input and returns the KL Divergence value. An example of calculating KL Divergence in R would be:

```
# Load the "entropy" package
library(entropy)

# Define two probability distributions
P
```

Calculate KL Divergence in R (With Example)

In statistics, the Kullback-Leibler (KL) divergence is a distance metric that quantifies the difference between two probability distributions.

If we have two probability distributions, P and Q, we typically write the KL divergence using the notation $KL(P \parallel Q)$, which means "P's divergence from Q."

We calculate it using the following formula:

$$KL(P \parallel Q) = \sum P(x) \ln(P(x) / Q(x))$$

If the KL divergence between two distributions is zero, then it indicates that the distributions are identical.

The easiest way to calculate the KL divergence between

two probability distributions in R is to use the `KL()` function from the `philentropy` package.

The following example shows how to use this function in practice.

Example: Calculating KL Divergence in R

Suppose we have the following two probability distributions in R:

```
#define two probability distributions  
P <- c(.05, .1, .2, .05, .15, .25, .08, .12)  
Q <- c(.3, .1, .2, .1, .1, .02, .08, .1)
```

Note: It's important that the probabilities for each distribution sum to one.

We can use the following code to calculate the KL divergence between the two distributions:

```
library(philentropy)
```

```
#rbind distributions into one matrix  
x <- rbind(P,Q)
```

```
#calculate KL divergence
```

```
KL(x, unit='log')
```

Metric: 'kullback-leibler' using unit: 'log'; comparing: 2 vectors.

```
kullback-leibler
```

```
0.5898852
```

The KL divergence of distribution P from distribution Q is about 0.589.

Note that the units used in this calculation are known as , which is short for *natural unit of information*.

Thus, we would say that the KL divergence is 0.589 nats.

Also note that the KL divergence is not a symmetric metric. This means that if we calculate the KL divergence of distribution Q from distribution P, we will likely get a different value:

```
library(philtropy)
```

```
#rbind distributions into one matrix
```

```
x <- rbind(Q,P)
```

```
#calculate KL divergence
```

```
KL(x, unit='log')
```

Metric: 'kullback-leibler' using unit: 'log'; comparing: 2 vectors.

```
kullback-leibler
```

```
0.4975493
```

The KL divergence of distribution Q from distribution P is about 0.497 nats.

Also note that some formulas use log base-2 to calculate the KL divergence. In this case, we refer to the divergence in terms of instead of nats.

To calculate the KL divergence in terms of bits, you can instead use log2 in the unit argument:

```
library(philterropy)
```

```
#rbind distributions into one matrix
```

```
x <- rbind(P,Q)
```

```
#calculate KL divergence (in bits)
```

KL(x, unit='log2')

Metric: 'kullback-leibler' using unit: 'log2'; comparing: 2 vectors.

kullback-leibler

0.7178119

The KL divergence of distribution P from distribution Q is about 0.7178 bits.

ARABPSYCHOLOGY.COM