R Tutorial

An R Introduction to Statistics

HOME   DOWNLOAD   SALES   EBOOK   SITE MAP

Kruskal-Wallis Test

A collection of data samples are **independent** if they come from unrelated populations and the samples do not affect each other. Using the **Kruskal-Wallis Test**, we can decide whether the population distributions are identical *without* assuming them to follow the normal distribution.

Example

In the built-in data set named **airquality**, the daily air quality measurements in New York, May to September 1973, are recorded. The ozone density are presented in the data frame column Ozone.

```
> head(airquality)
  Ozone Solar.R Wind Temp Month Day
1    41     190  7.4   67     5   1
2    36     118  8.0   72     5   2
    .....
```

Problem

Without assuming the data to have normal distribution, test at .05 significance level if the monthly ozone density in New York has identical data distributions from May to September 1973.

Solution

The null hypothesis is that the monthly ozone density are identical populations. To test the hypothesis, we apply the kruskal.test function to compare the independent monthly data. The p-value turns out to be nearly zero (6.901e-06). Hence we reject the null hypothesis.

```
> kruskal.test(Ozone ~ Month, data = airquality)

        Kruskal-Wallis rank sum test

data:  Ozone by Month
Kruskal-Wallis chi-squared = 29.267, df = 4, p-value = 6.901e-06
```
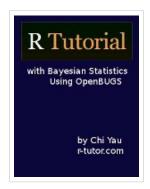
Answer

At .05 significance level, we conclude that the monthly ozone density in New York from May to September 1973 are *nonidentical* populations.

‹ Mann-Whitney-Wilcoxon Test   up   Simple Linear Regression ›

Tags: Elementary Statistics with R   Kruskal-Wallis test   non-parametric   head   kruskal.test
airquality

Search

R Tutorial eBook

R Tutorial

with Bayesian Statistics Using OpenBUGS

by Chi Yau
r-tutor.com

R Tutorials

Recent Articles

- Installing CUDA Toolkit 7.5 on Fedora 21 Linux

  September 10, 2015

- Installing CUDA Toolkit 7.5 on Ubuntu 14.04 Linux

  September 10, 2015

- Hierarchical Linear Model

  July 22, 2013

- Bayesian Classification with Gaussian Process

  January 6, 2013