IP: Index of Performance (IP) for the Rugby Attack Assessment Instrument (RAAI)

Jose Barrera-Gómez[a,b]
jbarrera@creal.cat

September 4, 2013

[a]Centre for Research in Environmental Epidemiology (CREAL), Barcelona, Spain
[b]Departament de Matemàtiques, Universitat Autònoma de Barcelona, Spain

Contents

1 Introduction

This document is a guide for the R package IP usage, which provides the Index of Performance (IP) [1]. Both point and confidence interval are computed. In addition, the package also can perform a test for the significance of the difference between two trials. Confidence intervals and p-value for the test are obtained using parametric bootstrap sampling assuming the Poisson distribution [2]. The IP package can be downloaded at `http://www.mat.uab.cat/~jbarrera/Software.html`.

2 Getting started

First, ensure you have R installed on your computer[1]. Then, download and install the package as follows:

- Windows users:
  1. Download the file `IP_0.6.zip` from `http://www.mat.uab.cat/~jbarrera/Software.html` to the desktop of your computer.
  2. Open R.
  3. Go to the "Packages" menu and select the option "Install package(s) from local zip files...".
  4. Then, browse your desktop and select the downloaded file `IP_0.6.zip`.
- Mac OS X users:
  1. Download the file `IP_0.6.tar.gz` from `http://www.mat.uab.cat/~jbarrera/Software.html` to the desktop of your computer.
  2. Open Terminal.
  3. Write and execute (⏎)

     ```
     cd Desktop
     ```
  4. Write and execute (⏎)

     ```
     R CMD INSTALL IP_0.6.tar.gz
     ```
  5. Close Terminal.
  6. Open R.

  We can now start the R session loading the package as follows:

```
> library(IP)
```

  If needed, this document can be recovered executing:

```
> vignette("IP")
```

2.1 The function `computeIP`

We can get information about the function `computeIP`, which calculates the IP value and a confidence interval:

```
> ?computeIP
```

The input parameters for the function `computeIP` are:

[1] R is a free and open source software and it is available at `http://cran.r-project.org/`.

- `w`: a numerical vector with the weights of the actions.
- `f`: a numerical vector with the frequencies of the actions.
- `conf`: the percentage confidence level for the confidence interval. Default is 95.
- `nsim`: the number of simulations for computing the parametric bootstrap confidence interval, based on the Poisson distribution. Default is 5000.
- `seed`: the seed for reproducibility of the confidence interval. Default is 4321.
- `digits`: number of decimal places to round printed results. Default is 3.

As a result, the function provides the IP score and a confidence interval.

2.2 The function `compareIP`

We can get information about the function `compareIP`, which calculates the significance of the IP difference between two trials (both confidence interval and p-value):

```
> ?compareIP
```

The input parameters for the function `compareIP` are:

- `w`: a numerical vector with the weights of the actions.
- `f1`: a numerical vector with the frequencies of the actions from trial 1.
- `f2`: a numerical vector with the frequencies of the actions from trial 2.
- `conf`: the percentage confidence level for the confidence interval. Default is 95.
- `nsim`: the number of simulations for computing the parametric bootstrap confidence interval, based on the Poisson distribution. Default is 5000.
- `seed`: the seed for reproducibility of the confidence interval. Default is 4321.
- `digits`: number of decimal places to round printed results. Default is 3.

As a result, the function provides the IP difference, a confidence interval and the p-value when testing equality of IP scores.

3 Some usage examples

Next examples are based on real data [1].

3.1 Example 1. Computing IP

Suppose we have defined 14 actions, $A1, \ldots, A14$, with weights 8, 7, 5, 3, 2, 1, 0, -1, -2, -3, -4, -5, -6 and -7, respectively [1]. Suppose we have performed a trial, say N1, whose observed frequencies for these actions were: 9, 0, 3, 5, 10, 63, 7, 9, 0, 6, 1, 1, 0 and 0, respectively. Then the IP value and a 95% confidence interval can be obtained as follows:

```
> myWeights <- c(8, 7, 5, 3, 2, 1, 0, -1, -2, -3, -4, -5, -6, -7)
> N1 <- c(9, 0, 3, 5, 10, 63, 7, 9, 0, 6, 1, 1, 0, 0)
> N1IP <- computeIP(w = myWeights, f = N1)
> N1IP

IP: 1.307
95% confidence interval based on 5000 parametric bootstrap samples: (0.858, 1.771)
```

3.2 Example 2. Comparing two IP values

Following with the example above, suppose now that we are interested in comparing N1 and N2 trials, corresponding both trials to the same team, having performed N2 two weeks after N1. N2 frequencies were: 12, 0, 7, 4, 12, 51, 6, 5, 0, 7, 1, 0, 0 and 0, respectively. Then the difference of IP values and its significances can be obtained as follows:

```
> myWeights <- c(8, 7, 5, 3, 2, 1, 0, -1, -2, -3, -4, -5, -6, -7)
> N1 <- c(9, 0, 3, 5, 10, 63, 7, 9, 0, 6, 1, 1, 0, 0)
> N2 <- c(12, 0, 7, 4, 12, 51, 6, 5, 0, 7, 1, 0, 0, 0)
> N1N2IP <- compareIP(w = myWeights, f1 = N1, f2 = N2)
> N1N2IP

IP scores and 95% confidence interval based on 5000 parametric bootstrap samples:
  Trial 1: 1.307 (0.858, 1.771)
  Trial 2: 1.79 (1.275, 2.327)
  Trial 1 - Trial 2: -0.483 (-1.182, 0.218)

Hyphotesis test for the significance of the difference:
  Null hyphotesis: performance profiles are the same in both trials.
  Alternative hyphotesis: performance profiles are not equal.
  p-value: 0.1814
```

3.3 Example 3. Computing and comparing IP values from data stored in a `xls` file

R can perform IP calculations with data stored in a `xls` file (amongst other file types). For instance, we provide the `IPdata.xls` file[2]. You can use this file as a template

[2] `http://www.mat.uab.cat/~jbarrera/Software.html`

replacing or adding in your own data. Once you have stored your data in the `IPdata.xls` file, using a spreadsheet application (e.g, Excel), you should save the file as a *comma separated values* file (`csv`) from the "Save as.." menu option. We also provide the `IPdata.csv` file. These files contain data about the K1, N1, N2 and E1 trials [1]. We next illustrate how to use the `IPdata.csv` file to perform the calculations.

First, download the file `IPdata.csv` to your desktop. Then, ensure your working directory in the R session is also your desktop (Menu "File", "Change dir..."). Then, we can reproduce results in Section 3.1 as follows.

We load the data and store it in an object, say `myData`:

```
> myData <- read.csv2("IPdata.csv", header = TRUE)
```

We now can see the data:

```
> myData
```

```
   action weight K1 N1 N2 E1
1      A1      8  0  9 12 13
2      A2      7  0  0  0  0
3      A3      5  4  3  7  4
4      A4      3  1  5  4  9
5      A5      2  9 10 12 15
6      A6      1 27 63 51 68
7      A7      0  5  7  6  3
8      A8     -1  4  9  5  2
9      A9     -2  0  0  0  0
10    A10     -3  8  6  7  3
11    A11     -4  1  1  1  0
12    A12     -5  4  1  0  2
13    A13     -6  0  0  0  0
14    A14     -7  1  0  0  0
```

Variables in `myData` are available using the $ symbol. For instance, the weights are

```
> myData$weight
```

```
 [1]  8  7  5  3  2  1  0 -1 -2 -3 -4 -5 -6 -7
```

and the frequencies from trial N1 are

```
> myData$N1
```

```
 [1]  9  0  3  5 10 63  7  9  0  6  1  1  0  0
```

Then, we can reproduce results in Section 3.1:

```
> N1IPnew <- computeIP(w = myData$weight, f = myData$N1)
> N1IPnew

IP: 1.307
95% confidence interval based on 5000 parametric bootstrap samples: (0.858, 1.771)
```

References

[1] Llobet B, López-Ros V, Barrera-Gómez J, Comino J. How to assess game performance in rugby union initiation: the Rugby Attack Assessment Instrument. *Journal of Teaching in Physical Education* [Submitted].

[2] Wasserman L. *All of Nonparametric Statistics*. Springer Texts in Statistics. Springer Science + Business: New York, 2006.