

# Package: oii (via r-universe)

August 25, 2024

**Title** Crosstab and Statistical Tests for OII MSc Stats Course

**Version** 1.0.2.1

**Description** Provides simple crosstab output with optional statistics (e.g., Goodman-Kruskal Gamma, Somers' d, and Kendall's tau-b) as well as two-way and one-way tables. The package is used within the statistics component of the Masters of Science (MSc) in Social Science of the Internet at the Oxford Internet Institute (OII), University of Oxford, but the functions should be useful for general data analysis and especially for analysis of categorical and ordinal data.

**Author** Scott Hale [aut, cre], Jon Bright [aut], Grant Blank [aut]

**Maintainer** Scott Hale <scott.hale@oii.ox.ac.uk>

**Depends** R (>= 3.2.1)

**License** MIT + file LICENSE

**LazyData** true

**Imports** rapportools, gmodels, Deducer

**RoxygenNote** 6.0.1

**NeedsCompilation** no

**Date/Publication** 2017-10-22 20:34:17 UTC

**Repository** <https://computermacgyver.r-universe.dev>

**RemoteUrl** <https://github.com/cran/oii>

**RemoteRef** HEAD

**RemoteSha** 43d91f691b526801f223acae0cdc4e0c871b75b1

## Contents

association.measures . . . . .	2
concordant.pairs . . . . .	3
discordant.pairs . . . . .	4
oii . . . . .	4
oii.freq . . . . .	5

oii.summary . . . . .	6
oii.xtab . . . . .	7
tied.pairs . . . . .	9

<b>Index</b>	<b>10</b>
--------------	-----------

---

association.measures *Measures of association*

---

## Description

This function calculates basic measures of association

## Usage

```
association.measures(x, y = NULL, warnings = FALSE)
```

## Arguments

x	a table or matrix if y is NULL, or a numeric vector for the row variable
y	the column variable, a numeric vector used only when x is not a table or matrix.
warnings	a logical value indicating whether warnings should be shown (defaults to FALSE, no warnings).

## Value

A list with the following elements is returned:

phi	Phi, a chi-square-based measures of association.
contingency_coefficient	Contingency coefficient, a chi-square-based measures of association.
cramersv	Cramer's V, a chi-square-based measures of association.
pairs_total	Total number of pairs
pairs_concordant	Number of concordant pairs
pairs_discordant	Number of discordant pairs
pairs_tied_first	The number of pairs tied on the first variable (but not both variables)
pairs_tied_second	The number of pairs tied on the second variable (but not both variables)
pairs_tied_both	The number of pairs tied on both the first and second variables
minimum_dim	Minimum dimension of x and y
n	Number of cases

gamma	Goodman-Kruskal Gamma
somersd	Somers' d (assuming the column variable is the dependent variable)
taub	Kendall's tau-b
tauc	Stuart's tau-c

**See Also**

[oii.xtab](#), [likelihood.test](#), [lambda.test](#), [concordant.pairs](#), [discordant.pairs](#), [tied.pairs](#)

**Examples**

```
#Create var1 as 200 A's, B's, and C's
var1<-sample(LETTERS[1:3],size=200,replace=TRUE)
#Create var2 as 200 numbers in the range 1 to 4
var2<-sample(1:4,size=200,replace=TRUE)

#Print a simple cross tab of var1 and var2
association.measures(var1,var2)
```

---

concordant.pairs      *The number of concordant pairs in a table or matrix*

---

**Description**

The number of concordant pairs in a table or matrix

**Usage**

```
concordant.pairs(x, y = NULL)
```

**Arguments**

x                    a table or matrix if y is NULL, or a numeric vector for the row variable  
y                    the column variable, a numeric vector used only when x is not a table or matrix.

**Value**

The number of concordant pairs

**See Also**

[association.measures](#), [discordant.pairs](#), [tied.pairs](#)

---

<code>discordant.pairs</code>	<i>The number of discordant pairs in a table or matrix</i>
-------------------------------	------------------------------------------------------------

---

**Description**

The number of discordant pairs in a table or matrix

**Usage**

```
discordant.pairs(x, y = NULL)
```

**Arguments**

<code>x</code>	a table or matrix if <code>y</code> is <code>NULL</code> , or a numeric vector for the row variable
<code>y</code>	the column variable, a numeric vector used only when <code>x</code> is not a table or matrix.

**Value**

The number of discordant pairs

**See Also**

[association.measures](#), [concordant.pairs](#), [tied.pairs](#)

---

<code>oii</code>	<i>Commands for the OII MSc Stats course</i>
------------------	----------------------------------------------

---

**Description**

This package provides a few commands that are used within the MSc course at the Oxford Internet Institute, University of Oxford

**Details**

The only functions you're likely to need from `oii` are [oii.summary](#), [oii.freq](#), and [oii.xtab](#).

---

`oii.freq`*Frequency tables*

---

**Description**

This function prints a simple frequency table with totals and percentages

**Usage**

```
oii.freq(x)
```

**Arguments**

`x` input variable, (usually of class `factor`)

**Value**

A `data.frame` with one row per each unique value of `x`. These values of `x` are assigned to the `row.names` of the `data.frame`. The `data.frame` also has rows for:

Valid Total      The total number of non-missing cases (i.e., `sum(!is.na(x))`)

Missing          The total number of missing/NA cases (i.e., `sum(is.na(x))`)

Total            The total number of cases (i.e., `length(x)`)

The `data.frame` has the following columns:

`freq`            The number of cases with this value

`percent`        The percentage of all cases that this value represents

`valid_percent` The percentage of all valid (i.e., not missing) cases that this value represents

`cum_percent`    The cumulative percentage of valid cases

**See Also**

[data.frame](#), [row.names](#) [is.na](#), [length](#), [summary](#), [table](#)

**Examples**

```
#Create var as 200 A's, B's, and C's
var<-sample(LETTERS[1:3],size=200,replace=TRUE)

#Generate a frequency table for the counts of A's, B's, and C's
oii.freq(var)
```

---

oii.summary

---

*Print summary statistics for a numeric variable*


---

### Description

This function is designed to be like the built-in [summary](#) function but include a few additional values. If the input is not numeric, the built-in summary command is executed.

### Usage

```
oii.summary(x, extended = FALSE, warnings = FALSE)
```

### Arguments

x	a numeric vector for which summary statistics should be generated.
extended	a logical value indicating whether additional statistics should be printed (see Value section). Defaults to FALSE stripped before the computation proceeds (defaults to TRUE).
warnings	a logical value indicating whether warnings should be shown (defaults to FALSE, no warnings).

### Value

If x is not numeric, the built-in summary command is executed. If x is numeric (that is, `is.numeric(x)` returns TRUE), then a list with the following elements is returned:

cases	The number of non-missing values in x (Valid N)
na	The number of missing values in x (Missing N).
mean	The mean value of x after missing values are removed. See <a href="#">mean</a>
sd	The standard deviation for values in x. See <a href="#">sd</a>
min	The minimum/smallest value in x. See <a href="#">min</a>
max	The maximum/largest value in x. See <a href="#">max</a>

This function also calculates the following statistics, but these are not printed by default unless `extended` is set to TRUE

var	The variance of x after missing values are removed. See <a href="#">var</a>
median	The median value of x after missing values are removed. See <a href="#">median</a>
p25	The 25th percentile of x after missing values are removed
p75	The 75th percentile of x after missing values are removed
skewness	The skewness coefficient for x after missing values are removed. See <a href="#">skewness</a>
kurtosis	The kurtosis coefficient for x after missing values are removed. See <a href="#">kurtosis</a>

**See Also**

[summary](#), [min](#), [median](#), [mean](#), [max](#), [sd](#), [is.na](#), [is.numeric](#), [skewness](#), [kurtosis](#), [var](#)

**Examples**

```
#Generate data from a normal distribution with mean 0 and sd 1
#store the result in a variable called tmp
tmp<-rnorm(500,mean=0,sd=1)

#Print the summary statistics about tmp
oii.summary(tmp)
#Print even more summary statistics about tmp
oii.summary(tmp,extended=TRUE)
```

---

oii.xtab

*A cross-tabulation with measures of association*


---

**Description**

This function prints a 2-way table with optional cell statistics and measures of association

**Usage**

```
oii.xtab(r, c = NULL, s = NULL, row = FALSE, col = FALSE,
  pctcell = FALSE, stats = FALSE, rescell = FALSE, chistd = FALSE,
  expcell = FALSE, chicell = FALSE, warnings = FALSE, varnames = NULL,
  include.missing = FALSE, ...)
```

**Arguments**

<code>r</code>	the row variable. If <code>r</code> is a <a href="#">table</a> , <a href="#">data.frame</a> , or <a href="#">matrix</a> , then <code>c</code> and <code>s</code> are ignored.
<code>c</code>	the column variable.
<code>s</code>	the split variable. The <code>r</code> and <code>c</code> will be separately tabulated for each unique value of <code>s</code> .
<code>row</code>	Show row percentages? Defaults to FALSE.
<code>col</code>	Show column percentages? Defaults to FALSE.
<code>pctcell</code>	Print cell percentages? Defaults to FALSE.
<code>stats</code>	Print measures of association? Defaults to FALSE. This parameter is ignored either <code>r</code> or <code>c</code> has only one value. See <a href="#">association.measures</a> .
<code>rescell</code>	Print residual cell count under the null hypothesis? Defaults to FALSE.
<code>chistd</code>	Print cell standardized residuals to pearson chi-square? Defaults to FALSE.
<code>expcell</code>	Print expected cell count under the null hypothesis? Defaults to FALSE.
<code>chicell</code>	Print cell contribution to pearson chi-square? Defaults to FALSE.

warnings	a logical value indicating whether warnings should be shown (defaults to FALSE, no warnings).
varnames	Names used to refer to r, c, and s in the printed output.
include.missing	Set to TRUE to include factor levels with no instances in the output. Default (FALSE) excludes them.
...	Additional parameters to be passed to <a href="#">CrossTable</a> .

### See Also

[association.measures](#), [CrossTable](#), [likelihood.test](#), [lambda.test](#)

### Examples

```
#Create var1 as 200 A's, B's, and C's
var1<-sample(LETTERS[1:3],size=200,replace=TRUE)
#Create var2 as 200 numbers in the range 1 to 4
var2<-sample(1:4,size=200,replace=TRUE)

#Print a simple 2-way table of var1 and var2
oii.xtab(var1,var2)

#Print the row and column percents
oii.xtab(var1,var2,row=TRUE,col=TRUE)

#Print measures of association statistics
oii.xtab(var1,var2,stats=TRUE)

#If the variables are part of a data.frame
my.data.frame<-data.frame(var1,var2)
#We can use the $ to get the variables
oii.xtab(my.data.frame$var1,my.data.frame$var2)
#or use the with(...) command to save some typing
with(my.data.frame,oii.xtab(var1,var2))

#Three-way tables are also possible
#Create var3 as 200 "I"'s, "II"'s, and "III"'s
var3<-sample(c("I","II","III"),size=200,replace=TRUE)
oii.xtab(var1,var2,var3)

#We can also pass in a data.frame directly as the first argument
my.data.frame<-data.frame(var1,var2,var3)
oii.xtab(my.data.frame,stats=TRUE)
#The variables in the data.frame are used in order;
#so, sometimes it is useful to re-order them. For example,
oii.xtab(my.data.frame[,c("var3","var1","var2")],stats=TRUE)
#Of course, it is also possible to pass in the variables one
#at a time or use with(...) as shown above.
```

---

tied.pairs	<i>The number of tied pairs, a measure of association</i>
------------	-----------------------------------------------------------

---

**Description**

The number of tied pairs, a measure of association

**Usage**

```
tied.pairs(x, y = NULL)
```

**Arguments**

x	a table or matrix if y is NULL, or a numeric vector for the first variable
y	the second variable, a numeric vector used only when x is not a table or matrix.

**Value**

A list with the following values:

first	The number of pairs tied on the first variable, but not both variables
second	The number of pairs tied on the second variable, but not both variables
both	The number of pairs tied on both the first and second variables

**See Also**

[association.measures](#), [concordant.pairs](#), [discordant.pairs](#)

# Index

association.measures, 2, 3, 4, 7–9

concordant.pairs, 3, 3, 4, 9

CrossTable, 8

data.frame, 5, 7

discordant.pairs, 3, 4, 9

factor, 5

is.na, 5, 7

is.numeric, 7

kurtosis, 6, 7

lambda.test, 3, 8

length, 5

likelihood.test, 3, 8

matrix, 7

max, 6, 7

mean, 6, 7

median, 6, 7

min, 6, 7

oii, 4

oii-package (oii), 4

oii.freq, 4, 5

oii.summary, 4, 6

oii.xtab, 3, 4, 7

row.names, 5

sd, 6, 7

skewness, 6, 7

summary, 5–7

table, 5, 7

tied.pairs, 3, 4, 9

var, 6, 7