

# Package ‘tidyinftheo’

November 30, 2017

**Type** Package

**Title** Some Information-Theoretic Functions in the 'Tidy' Style

**Version** 0.2.1

**Description** A frontend to the Shannon entropy, conditional Shannon entropy, and mutual information calculations provided by the 'infotheo' package. See Cover and Thomas (2001) <doi:10.1002/0471200611> for an explanation of these measures. Also provides a convenient heatmap to compare more than two columns in a pairwise fashion.

**License** GPL-3

**Encoding** UTF-8

**LazyData** true

**Depends** R (>= 3.1.0)

**Imports** dplyr, infotheo, ggplot2, magrittr, purrr, rlang, tibble, tidyselect, tidyverse

**Suggests** knitr, rmarkdown (>= 0.2.65), testthat, covr

**VignetteBuilder** knitr

**RoxygenNote** 6.0.1

**URL** <https://github.com/pohlio/tidyinftheo>

**BugReports** <https://github.com/pohlio/tidyinftheo/issues>

**NeedsCompilation** no

**Author** Andy Pohl [aut, cre],  
Pohl Data, LLC [cph]

**Maintainer** Andy Pohl <andy@pohl.io>

**Repository** CRAN

**Date/Publication** 2017-11-30 18:23:07 UTC

## R topics documented:

tidyinftheo-package . . . . .	2
mutual_info . . . . .	3
mutual_info_heatmap . . . . .	3
mutual_info_matrix . . . . .	4
shannon_cond_entropy . . . . .	5
shannon_entropy . . . . .	5

<b>Index</b>	<b>7</b>
--------------	----------

---

tidyinftheo-package    *tidyinftheo: tidy-style information theoretic routines*

---

## Description

tidyinftheo provides an easy way to calculate shannon entropy, mutual information, etc. from fields in a 'tibble'

## Details

To learn more about tidyinftheo, start with the vignettes: `browseVignettes(package = "tidyinftheo")`

## Author(s)

**Maintainer:** Andy Pohl <andy@pohl.io>

Other contributors:

- Pohl Data, LLC [copyright holder]

## See Also

Useful links:

- <https://github.com/pohlio/tidyinftheo>
- Report bugs at <https://github.com/pohlio/tidyinftheo/issues>

---

mutual_info	<i>Mutual information</i> $MI(X;Y) = H(X) - H(X Y) = H(Y) - H(Y X)$
-------------	---

---

**Description**

This calculates the mutual information between two variables in a tibble. (if normalized). It's assumed these columns are character typed with no NAs.

**Usage**

```
mutual_info(.data, ..., normalized = FALSE, na.rm = FALSE)
```

**Arguments**

.data	A tibble with the column of interest
...	two columns (variables) selected
normalized	if TRUE, scale from 0 to 1
na.rm	remove all rows with NA values in at least one of the columns

**Value**

a double with the calculated value

**Examples**

```
# make an all-character version of mtchars
mt_tib <- as_tibble(mtcars) %>% mutate_all(as.character)
mutual_info(mt_tib, vs, am)
mutual_info(mt_tib, 'am', 'vs')
mutual_info(mt_tib, vs, am, normalized=TRUE)
mutual_info(mt_tib, starts_with('c'))
```

---

mutual_info_heatmap	<i>Plot heatmap of mutual infos</i>
---------------------	-------------------------------------

---

**Description**

Given a matrix produced from mutual\_info\_matrix(), plot a heatmap with ggplot2

**Usage**

```
mutual_info_heatmap(mi_matrix, title = NULL, font_sizes = c(12, 12))
```

**Arguments**

mi\_matrix      a table produced from [mutual\\_info\\_matrix](#)  
 title          title of plot  
 font\_sizes     A length-2 vector of x-axis and y-axis variable font sizes

**Value**

a double with the calculated value

**See Also**

[mutual\\_info\\_matrix](#)

**Examples**

```
p <- mtcars %>%
  mutual_info_matrix(cyl, vs, am, gear, carb, normalized=TRUE) %>%
  mutual_info_heatmap()
p
```

---

mutual\_info\_matrix      *Mutual information Matrix*

---

**Description**

To simplify the task of comparing variables, this calculates a matrix of mutual information values from each pairwise combination of the variables selected. If 6 variables are selected, that would yield a table with 15 rows (choose(6,2)), and 3 columns.

**Usage**

```
mutual_info_matrix(.data, ..., normalized = FALSE, na.rm = FALSE)
```

**Arguments**

.data            A tibble with the columns of interest  
 ...             a selection of columns, selected in the same way as [select](#)  
 normalized     if TRUE, scale from 0 to 1  
 na.rm          remove all rows with NA values in at least one of the columns

**Value**

a 3 column tibble with each pairwise combination and its calculated mutual information

**Examples**

```
# make an all-character version of mtchars
mt_tib <- as_tibble(mtcars) %>% mutate_all(as.character)
mutual_info_matrix(mt_tib, 8:11)
```

---

shannon\_cond\_entropy    *Conditional Shannon Entropy  $H(X|Y)$  i.e. " $H(X \text{ given } Y)$ "*

---

### Description

This calculates conditional shannon entropy of two columns in a tibble. It's assumed these columns are character typed with no NAs.

### Usage

```
shannon_cond_entropy(.data, ..., na.rm = FALSE)
```

### Arguments

.data	A tibble with the columns of interest
...	two columns (variables) selected (order is important)
na.rm	remove all rows with NA values in at least one of the columns

### Value

a double with the calculated value

### See Also

[shannon\\_entropy](#)

### Examples

```
# make an all-character version of mtcars
mt_tib <- as_tibble(mtcars) %>% mutate_all(as.character)
shannon_cond_entropy(mt_tib, vs, am)
shannon_cond_entropy(mt_tib, 'vs', 'am')
shannon_cond_entropy(mt_tib, starts_with('c'))
shannon_cond_entropy(mt_tib, 9:8)
```

---

shannon\_entropy    *Shannon Entropy  $H(X)$*

---

### Description

This calculates shannon entropy of a variable in a tibble. It's assumed these columns are character typed with no NAs.

### Usage

```
shannon_entropy(.data, X, na.rm = FALSE)
```

**Arguments**

<code>.data</code>	A tibble with the column of interest
<code>x</code>	Name of the column
<code>na.rm</code>	remove all rows with NA values in at least one of the columns

**Value**

a double with the calculated value

**See Also**

[shannon\\_cond\\_entropy](#)

**Examples**

```
shannon_entropy(iris, Species)
iris %>% as_tibble() %>% shannon_entropy(Species)
shannon_entropy(iris, 'Species')
shannon_entropy(iris, 5)
```

# Index

`mutual_info`, [3](#)

`mutual_info_heatmap`, [3](#)

`mutual_info_matrix`, [4](#), [4](#)

`select`, [4](#)

`shannon_cond_entropy`, [5](#), [6](#)

`shannon_entropy`, [5](#), [5](#)

`tidyinftheo` (`tidyinftheo`-package), [2](#)

`tidyinftheo`-package, [2](#)