

# Package ‘stabm’

March 23, 2020

**Title** Stability Measures for Feature Selection

**Version** 1.1.2

**Description** An implementation of many measures for the assessment of the stability of feature selection. Both simple measures and measures which take into account the similarities between features are available, see Bommert et al. (2017) <doi:10.1155/2017/7907163>.

**Depends** R (>= 3.5.0)

**Imports** BBmisc (>= 1.11), checkmate (>= 1.8.5), Matrix (>= 1.2-14), methods, stats, utils

**Suggests** cowplot (>= 0.9.2), data.table (>= 1.11.4), ggdendro (>= 0.1-20), ggplot2 (>= 3.0.0), igraph (>= 1.2.1), reshape2 (>= 1.4.3), testthat (>= 2.0.0)

**License** LGPL-3

**Encoding** UTF-8

**LazyData** true

**RoxygenNote** 6.1.1

**URL** <https://github.com/bommert/stabm>

**BugReports** <https://github.com/bommert/stabm/issues>

**NeedsCompilation** no

**Author** Andrea Bommert [aut, cre]

**Maintainer** Andrea Bommert <bommert@statistik.tu-dortmund.de>

**Repository** CRAN

**Date/Publication** 2020-03-23 18:00:02 UTC

## R topics documented:

stabm-package	2
listStabilityMeasures	3
plotFeatures	3
stabilityDavis	4

stabilityDice . . . . .	6
stabilityHamming . . . . .	8
stabilityIntersectionCount . . . . .	10
stabilityIntersectionGreedy . . . . .	12
stabilityIntersectionMBM . . . . .	14
stabilityIntersectionMean . . . . .	16
stabilityJaccard . . . . .	18
stabilityKappa . . . . .	20
stabilityLustgarten . . . . .	21
stabilityNogueira . . . . .	23
stabilityNovovicova . . . . .	24
stabilityOchiai . . . . .	26
stabilityPhi . . . . .	28
stabilitySomol . . . . .	29
stabilityUnadjusted . . . . .	31
stabilityWald . . . . .	32
stabilityYu . . . . .	34
stabilityZucknick . . . . .	36
<b>Index</b>	<b>39</b>

---

 stabm-package

*stabm: Stability Measures for Feature Selection*


---

## Description

An implementation of many measures for the assessment of the stability of feature selection. Both simple measures and measures which take into account the similarities between features are available, see Bommert et al. (2017) <doi:10.1155/2017/7907163>.

## Author(s)

**Maintainer:** Andrea Bommert <bommert@statistik.tu-dortmund.de>

## See Also

Useful links:

- <https://github.com/bommert/stabm>
- Report bugs at <https://github.com/bommert/stabm/issues>

---

listStabilityMeasures *List All Available Stability Measures*

---

**Description**

Lists all stability measures of package *stabm* and provides information about them.

**Usage**

```
listStabilityMeasures()
```

**Value**

data.frame

For each stability measure, its name, the information, whether it is corrected for chance by definition, the information, whether it is adjusted for similar features, its minimal value and its maximal value are displayed.

**Note**

The given minimal values might only be reachable in some scenarios, e.g. if the feature sets have a certain size. The measures which are not corrected for chance by definition can be corrected for chance with `correction.for.chance`. This however changes the minimal value. For the adjusted stability measures, the minimal value depends on the similarity structure.

---

plotFeatures *Plot Selected Features*

---

**Description**

Creates a heatmap of the features which are selected in at least one feature set. The sets are ordered according to average linkage hierarchical clustering based on the Manhattan distance. If `sim.mat` is given, the features are ordered according to average linkage hierarchical clustering based on `1 - sim.mat`. Otherwise, the features are ordered in the same way as the feature sets.

**Usage**

```
plotFeatures(features, sim.mat = NULL)
```

**Arguments**

features	list (length >= 2) Chosen features per dataset. Each element of the list contains the features for one dataset. The features must be given by their names (character) or indices (integerish).
sim.mat	numeric matrix Similarity matrix which contains the similarity structure of all features based on all datasets. The similarity values must be in the range of [0, 1] where 0 indicates very low similarity and 1 indicates very high similarity. If the list elements of features are integerish vectors, then the feature numbering must correspond to the ordering of sim.mat. If the list elements of features are character vectors, then sim.mat must be named and the names of sim.mat must correspond to the entries in features.

**Value**

Object of class ggplot.

**Examples**

```
feats = list(1:3, 1:4, 1:5)
mat = 0.92 ^ abs(outer(1:10, 1:10, "-"))
plotFeatures(features = feats)
plotFeatures(features = feats, sim.mat = mat)
```

---

stabilityDavis

*Stability Measure Davis*

---

**Description**

The stability of feature selection is defined as the robustness of the sets of selected features with respect to small variations in the data on which the feature selection is conducted. To quantify stability, several datasets from the same data generating process can be used. Alternatively, a single dataset can be split into parts by resampling. Either way, all datasets used for feature selection must contain exactly the same features. The feature selection method of interest is applied on all of the datasets and the sets of chosen features are recorded. The stability of the feature selection is assessed based on the sets of chosen features using stability measures.

**Usage**

```
stabilityDavis(features, p, correction.for.chance = "none", N = 10000,
  impute.na = NULL, penalty = 0)
```

**Arguments**

features	list (length >= 2) Chosen features per dataset. Each element of the list contains the features for one dataset. The features must be given by their names (character) or indices (integerish).
p	numeric(1) Total number of features in the datasets. Required, if <code>correction.for.chance</code> is set to "estimate" or "exact".
correction.for.chance	character(1) Should a correction for chance be applied? Correction for chance means that if features are chosen at random, the expected value must be independent of the number of chosen features. To correct for chance, the original score is transformed by $(score - expected)/(maximum - expected)$ . For stability measures whose score is the average value of pairwise scores, this transformation is done for all components individually. Options are "none", "estimate" and "exact". For "none", no correction is performed, i.e. the original score is used. For "estimate", N random feature sets of the same sizes as the input feature sets (features) are generated. For "exact", all possible combinations of feature sets of the same sizes as the input feature sets are used. Computation is only feasible for very small numbers of features (p) and numbers of considered datasets (length(features)).
N	numeric(1) Number of random feature sets to consider. Only relevant if <code>correction.for.chance</code> is set to "estimate".
impute.na	numeric(1) In some scenarios, the stability cannot be assessed based on all feature sets. E.g. if some of the feature sets are empty, the respective pairwise comparisons yield NA as result. With which value should these missing values be imputed? NULL means no imputation.
penalty	numeric(1) Penalty parameter, see Details.

**Details**

The stability measure is defined as (see Notation)

$$\max \left\{ 0, \frac{1}{|V|} \sum_{j=1}^p \frac{h_j}{m} - \frac{penalty}{p} \cdot median\{|V_1|, \dots, |V_m|\} \right\}.$$

**Value**

numeric(1) Stability value.

**Notation**

For the definition of all stability measures in this package, the following notation is used: Let  $V_1, \dots, V_m$  denote the sets of chosen features for the  $m$  datasets, i.e. `features` has length  $m$  and  $V_i$  is a set which contains the  $i$ -th entry of `features`. Furthermore, let  $h_j$  denote the number of sets that contain feature  $X_j$  so that  $h_j$  is the absolute frequency with which feature  $X_j$  is chosen. Also, let  $q = \sum_{j=1}^p h_j = \sum_{i=1}^m |V_i|$  and  $V = \bigcup_{i=1}^m V_i$ .

**References**

- C. A. Davis, F. Gerick, V. Hintermair, C. C. Friedel, K. Fundel, R. Küffner, and R. Zimmer, "Reliable gene signatures for microarray classification: assessment of stability and performance", *Bioinformatics*, vol. 22, no. 19, pp. 2356-2363, 2006.
- A. Bommert, J. Rahnenführer, and M. Lang, "A multi-criteria approach to find predictive and sparse models with stable feature selection for high-dimensional data", *Computational and mathematical methods in medicine*, 2017.

**See Also**

[listStabilityMeasures](#)

**Examples**

```
feats = list(1:3, 1:4, 1:5)
stabilityDavis(features = feats, p = 10)
```

---

stabilityDice

*Stability Measure Dice*

---

**Description**

The stability of feature selection is defined as the robustness of the sets of selected features with respect to small variations in the data on which the feature selection is conducted. To quantify stability, several datasets from the same data generating process can be used. Alternatively, a single dataset can be split into parts by resampling. Either way, all datasets used for feature selection must contain exactly the same features. The feature selection method of interest is applied on all of the datasets and the sets of chosen features are recorded. The stability of the feature selection is assessed based on the sets of chosen features using stability measures.

**Usage**

```
stabilityDice(features, p = NULL, correction.for.chance = "none",
  N = 10000, impute.na = NULL)
```

**Arguments**

features	list (length >= 2) Chosen features per dataset. Each element of the list contains the features for one dataset. The features must be given by their names (character) or indices (integerish).
p	numeric(1) Total number of features in the datasets. Required, if correction.for.chance is set to "estimate" or "exact".
correction.for.chance	character(1) Should a correction for chance be applied? Correction for chance means that if features are chosen at random, the expected value must be independent of the number of chosen features. To correct for chance, the original score is transformed by $(score - expected)/(maximum - expected)$ . For stability measures whose score is the average value of pairwise scores, this transformation is done for all components individually. Options are "none", "estimate" and "exact". For "none", no correction is performed, i.e. the original score is used. For "estimate", N random feature sets of the same sizes as the input feature sets (features) are generated. For "exact", all possible combinations of feature sets of the same sizes as the input feature sets are used. Computation is only feasible for very small numbers of features (p) and numbers of considered datasets (length(features)).
N	numeric(1) Number of random feature sets to consider. Only relevant if correction.for.chance is set to "estimate".
impute.na	numeric(1) In some scenarios, the stability cannot be assessed based on all feature sets. E.g. if some of the feature sets are empty, the respective pairwise comparisons yield NA as result. With which value should these missing values be imputed? NULL means no imputation.

**Details**

The stability measure is defined as (see Notation)

$$\frac{2}{m(m-1)} \sum_{i=1}^{m-1} \sum_{j=i+1}^m \frac{2|V_i \cap V_j|}{|V_i| + |V_j|}$$

**Value**

numeric(1) Stability value.

**Notation**

For the definition of all stability measures in this package, the following notation is used: Let  $V_1, \dots, V_m$  denote the sets of chosen features for the  $m$  datasets, i.e. features has length  $m$  and  $V_i$  is a set which contains the  $i$ -th entry of features. Furthermore, let  $h_j$  denote the number of sets

that contain feature  $X_j$  so that  $h_j$  is the absolute frequency with which feature  $X_j$  is chosen. Also, let  $q = \sum_{j=1}^p h_j = \sum_{i=1}^m |V_i|$  and  $V = \bigcup_{i=1}^m V_i$ .

## References

- L. R. Dice, "Measures of the amount of ecologic association between species", Ecology, vol. 26, no. 3, pp. 297-302, 1945.
- A. Bommert, J. Rahnenführer, and M. Lang, "A multi-criteria approach to find predictive and sparse models with stable feature selection for high-dimensional data", Computational and mathematical methods in medicine, 2017.

## See Also

[listStabilityMeasures](#)

## Examples

```
feats = list(1:3, 1:4, 1:5)
stabilityDice(features = feats)
```

---

stabilityHamming

*Stability Measure Hamming*

---

## Description

The stability of feature selection is defined as the robustness of the sets of selected features with respect to small variations in the data on which the feature selection is conducted. To quantify stability, several datasets from the same data generating process can be used. Alternatively, a single dataset can be split into parts by resampling. Either way, all datasets used for feature selection must contain exactly the same features. The feature selection method of interest is applied on all of the datasets and the sets of chosen features are recorded. The stability of the feature selection is assessed based on the sets of chosen features using stability measures.

## Usage

```
stabilityHamming(features, p, correction.for.chance = "none",
  N = 10000, impute.na = NULL)
```

## Arguments

features	list (length >= 2) Chosen features per dataset. Each element of the list contains the features for one dataset. The features must be given by their names (character) or indices (integerish).
p	numeric(1) Total number of features in the datasets. Required, if correction.for.chance is set to "estimate" or "exact".



correction.for.chance	character(1) Should a correction for chance be applied? Correction for chance means that if features are chosen at random, the expected value must be independent of the number of chosen features. To correct for chance, the original score is transformed by $(score - expected)/(maximum - expected)$ . For stability measures whose score is the average value of pairwise scores, this transformation is done for all components individually. Options are "none", "estimate" and "exact". For "none", no correction is performed, i.e. the original score is used. For "estimate", N random feature sets of the same sizes as the input feature sets (features) are generated. For "exact", all possible combinations of feature sets of the same sizes as the input feature sets are used. Computation is only feasible for very small numbers of features (p) and numbers of considered datasets (length(features)).
N	numeric(1) Number of random feature sets to consider. Only relevant if correction.for.chance is set to "estimate".
impute.na	numeric(1) In some scenarios, the stability cannot be assessed based on all feature sets. E.g. if some of the feature sets are empty, the respective pairwise comparisons yield NA as result. With which value should these missing values be imputed? NULL means no imputation.

## Details

The stability measure is defined as (see Notation)

$$\frac{2}{m(m-1)} \sum_{i=1}^{m-1} \sum_{j=i+1}^m \frac{|V_i \cap V_j| + |V_i^c \cap V_j^c|}{p}$$

## Value

numeric(1) Stability value.

## Notation

For the definition of all stability measures in this package, the following notation is used: Let  $V_1, \dots, V_m$  denote the sets of chosen features for the  $m$  datasets, i.e. features has length  $m$  and  $V_i$  is a set which contains the  $i$ -th entry of features. Furthermore, let  $h_j$  denote the number of sets that contain feature  $X_j$  so that  $h_j$  is the absolute frequency with which feature  $X_j$  is chosen. Also, let  $q = \sum_{j=1}^p h_j = \sum_{i=1}^m |V_i|$  and  $V = \bigcup_{i=1}^m V_i$ .

## References

- K. Dunne, P. Cunningham, and F. Azuaje, "Solutions to instability problems with sequential wrapper-based approaches to feature selection", Technical Report, 2002.

## See Also

[listStabilityMeasures](#)

**Examples**

```
feats = list(1:3, 1:4, 1:5)
stabilityHamming(features = feats, p = 10)
```

---

```
stabilityIntersectionCount
```

*Stability Measure Adjusted Intersection Count*

---

**Description**

The stability of feature selection is defined as the robustness of the sets of selected features with respect to small variations in the data on which the feature selection is conducted. To quantify stability, several datasets from the same data generating process can be used. Alternatively, a single dataset can be split into parts by resampling. Either way, all datasets used for feature selection must contain exactly the same features. The feature selection method of interest is applied on all of the datasets and the sets of chosen features are recorded. The stability of the feature selection is assessed based on the sets of chosen features using stability measures.

**Usage**

```
stabilityIntersectionCount(features, sim.mat, threshold = 0.9,
  correction.for.chance = "estimate", N = 10000, impute.na = NULL)
```

**Arguments**

features	list (length >= 2) Chosen features per dataset. Each element of the list contains the features for one dataset. The features must be given by their names (character) or indices (integerish).
sim.mat	numeric matrix Similarity matrix which contains the similarity structure of all features based on all datasets. The similarity values must be in the range of [0, 1] where 0 indicates very low similarity and 1 indicates very high similarity. If the list elements of features are integerish vectors, then the feature numbering must correspond to the ordering of sim.mat. If the list elements of features are character vectors, then sim.mat must be named and the names of sim.mat must correspond to the entries in features.
threshold	numeric(1) Threshold for indicating which features are similar and which are not. Two features are considered as similar, if and only if the corresponding entry of sim.mat is greater than or equal to threshold.
correction.for.chance	character(1) How should the expected value of the stability score (see Details) be assessed? Options are "estimate", "exact" and "none". For "estimate", N random feature sets of the same sizes as the input feature sets (features) are generated. For

"exact", all possible combinations of feature sets of the same sizes as the input feature sets are used. Computation is only feasible for very small numbers of features and numbers of considered datasets (`length(features)`). For "none", the transformation  $(score - expected)/(maximum - expected)$  is not conducted, i.e. only `score` is used. This is not recommended.

N	numeric(1) Number of random feature sets to consider. Only relevant if <code>correction</code> for <code>chance</code> is set to "estimate".
impute.na	numeric(1) In some scenarios, the stability cannot be assessed based on all feature sets. E.g. if some of the feature sets are empty, the respective pairwise comparisons yield NA as result. With which value should these missing values be imputed? NULL means no imputation.

### Details

The stability measure is defined as (see Notation)

$$\frac{2}{m(m-1)} \sum_{i=1}^{m-1} \sum_{j=i+1}^m \frac{I(V_i, V_j) - E(I(V_i, V_j))}{\sqrt{|V_i| \cdot |V_j|} - E(I(V_i, V_j))}$$

with

$$I(V_i, V_j) = |V_i \cap V_j| + \min(C(V_i, V_j), C(V_j, V_i))$$

and

$$C(V_k, V_l) = |\{x \in V_k \setminus V_l : \exists y \in V_l \setminus V_k \text{ with } \text{Similarity}(x, y) \geq \text{threshold}\}|.$$

### Value

numeric(1) Stability value.

### Notation

For the definition of all stability measures in this package, the following notation is used: Let  $V_1, \dots, V_m$  denote the sets of chosen features for the  $m$  datasets, i.e. `features` has length  $m$  and  $V_i$  is a set which contains the  $i$ -th entry of `features`. Furthermore, let  $h_j$  denote the number of sets that contain feature  $X_j$  so that  $h_j$  is the absolute frequency with which feature  $X_j$  is chosen. Also, let  $q = \sum_{j=1}^p h_j = \sum_{i=1}^m |V_i|$  and  $V = \bigcup_{i=1}^m V_i$ .

### See Also

[listStabilityMeasures](#)

### Examples

```
feats = list(1:3, 1:4, 1:5)
mat = 0.92 ^ abs(outer(1:10, 1:10, "-"))
stabilityIntersectionCount(features = feats, sim.mat = mat, N = 1000)
```

---

 stabilityIntersectionGreedy

*Stability Measure Adjusted Intersection Greedy*


---

## Description

The stability of feature selection is defined as the robustness of the sets of selected features with respect to small variations in the data on which the feature selection is conducted. To quantify stability, several datasets from the same data generating process can be used. Alternatively, a single dataset can be split into parts by resampling. Either way, all datasets used for feature selection must contain exactly the same features. The feature selection method of interest is applied on all of the datasets and the sets of chosen features are recorded. The stability of the feature selection is assessed based on the sets of chosen features using stability measures.

## Usage

```
stabilityIntersectionGreedy(features, sim.mat, threshold = 0.9,
  correction.for.chance = "estimate", N = 10000, impute.na = NULL)
```

## Arguments

features	list (length >= 2) Chosen features per dataset. Each element of the list contains the features for one dataset. The features must be given by their names (character) or indices (integerish).
sim.mat	numeric matrix Similarity matrix which contains the similarity structure of all features based on all datasets. The similarity values must be in the range of [0, 1] where 0 indicates very low similarity and 1 indicates very high similarity. If the list elements of features are integerish vectors, then the feature numbering must correspond to the ordering of sim.mat. If the list elements of features are character vectors, then sim.mat must be named and the names of sim.mat must correspond to the entries in features.
threshold	numeric(1) Threshold for indicating which features are similar and which are not. Two features are considered as similar, if and only if the corresponding entry of sim.mat is greater than or equal to threshold.
correction.for.chance	character(1) How should the expected value of the stability score (see Details) be assessed? Options are "estimate", "exact" and "none". For "estimate", N random feature sets of the same sizes as the input feature sets (features) are generated. For "exact", all possible combinations of feature sets of the same sizes as the input feature sets are used. Computation is only feasible for very small numbers of features and numbers of considered datasets (length(features)). For "none", the transformation $(score - expected)/(maximum - expected)$ is not conducted, i.e. only score is used. This is not recommended.

N	numeric(1) Number of random feature sets to consider. Only relevant if correction.for.chance is set to "estimate".
impute.na	numeric(1) In some scenarios, the stability cannot be assessed based on all feature sets. E.g. if some of the feature sets are empty, the respective pairwise comparisons yield NA as result. With which value should these missing values be imputed? NULL means no imputation.

## Details

The stability measure is defined as (see Notation)

$$\frac{2}{m(m-1)} \sum_{i=1}^{m-1} \sum_{j=i+1}^m \frac{I(V_i, V_j) - E(I(V_i, V_j))}{\sqrt{|V_i| \cdot |V_j|} - E(I(V_i, V_j))}$$

with

$$I(V_i, V_j) = |V_i \cap V_j| + GMBM(V_i \setminus V_j, V_j \setminus V_i).$$

$GMBM(V_i \setminus V_j, V_j \setminus V_i)$  denotes a greedy approximation of  $MBM(V_i \setminus V_j, V_j \setminus V_i)$ , see [stabilityIntersectionMBM](#).

## Value

numeric(1) Stability value.

## Notation

For the definition of all stability measures in this package, the following notation is used: Let  $V_1, \dots, V_m$  denote the sets of chosen features for the  $m$  datasets, i.e. `features` has length  $m$  and  $V_i$  is a set which contains the  $i$ -th entry of `features`. Furthermore, let  $h_j$  denote the number of sets that contain feature  $X_j$  so that  $h_j$  is the absolute frequency with which feature  $X_j$  is chosen. Also, let  $q = \sum_{j=1}^p h_j = \sum_{i=1}^m |V_i|$  and  $V = \bigcup_{i=1}^m V_i$ .

## See Also

[listStabilityMeasures](#)

## Examples

```
feats = list(1:3, 1:4, 1:5)
mat = 0.92 ^ abs(outer(1:10, 1:10, "-"))
stabilityIntersectionGreedy(features = feats, sim.mat = mat, N = 1000)
```

---

 stabilityIntersectionMBM

*Stability Measure Adjusted Intersection MBM*


---

## Description

The stability of feature selection is defined as the robustness of the sets of selected features with respect to small variations in the data on which the feature selection is conducted. To quantify stability, several datasets from the same data generating process can be used. Alternatively, a single dataset can be split into parts by resampling. Either way, all datasets used for feature selection must contain exactly the same features. The feature selection method of interest is applied on all of the datasets and the sets of chosen features are recorded. The stability of the feature selection is assessed based on the sets of chosen features using stability measures.

## Usage

```
stabilityIntersectionMBM(features, sim.mat, threshold = 0.9,
  correction.for.chance = "estimate", N = 10000, impute.na = NULL)
```

## Arguments

features	list (length >= 2) Chosen features per dataset. Each element of the list contains the features for one dataset. The features must be given by their names (character) or indices (integerish).
sim.mat	numeric matrix Similarity matrix which contains the similarity structure of all features based on all datasets. The similarity values must be in the range of [0, 1] where 0 indicates very low similarity and 1 indicates very high similarity. If the list elements of features are integerish vectors, then the feature numbering must correspond to the ordering of sim.mat. If the list elements of features are character vectors, then sim.mat must be named and the names of sim.mat must correspond to the entries in features.
threshold	numeric(1) Threshold for indicating which features are similar and which are not. Two features are considered as similar, if and only if the corresponding entry of sim.mat is greater than or equal to threshold.
correction.for.chance	character(1) How should the expected value of the stability score (see Details) be assessed? Options are "estimate", "exact" and "none". For "estimate", N random feature sets of the same sizes as the input feature sets (features) are generated. For "exact", all possible combinations of feature sets of the same sizes as the input feature sets are used. Computation is only feasible for very small numbers of features and numbers of considered datasets (length(features)). For "none", the transformation $(score - expected)/(maximum - expected)$ is not conducted, i.e. only score is used. This is not recommended.

N	numeric(1) Number of random feature sets to consider. Only relevant if correction.for.chance is set to "estimate".
impute.na	numeric(1) In some scenarios, the stability cannot be assessed based on all feature sets. E.g. if some of the feature sets are empty, the respective pairwise comparisons yield NA as result. With which value should these missing values be imputed? NULL means no imputation.

## Details

The stability measure is defined as (see Notation)

$$\frac{2}{m(m-1)} \sum_{i=1}^{m-1} \sum_{j=i+1}^m \frac{I(V_i, V_j) - E(I(V_i, V_j))}{\sqrt{|V_i| \cdot |V_j|} - E(I(V_i, V_j))}$$

with

$$I(V_i, V_j) = |V_i \cap V_j| + MBM(V_i \setminus V_j, V_j \setminus V_i).$$

$MBM(V_i \setminus V_j, V_j \setminus V_i)$  denotes the size of the maximum bipartite matching based on the graph whose vertices are the features of  $V_i \setminus V_j$  on the one side and the features of  $V_j \setminus V_i$  on the other side. Vertices  $x$  and  $y$  are connected if and only if  $Similarity(x, y) \geq threshold$ .

## Value

numeric(1) Stability value.

## Notation

For the definition of all stability measures in this package, the following notation is used: Let  $V_1, \dots, V_m$  denote the sets of chosen features for the  $m$  datasets, i.e. `features` has length  $m$  and  $V_i$  is a set which contains the  $i$ -th entry of `features`. Furthermore, let  $h_j$  denote the number of sets that contain feature  $X_j$  so that  $h_j$  is the absolute frequency with which feature  $X_j$  is chosen. Also, let  $q = \sum_{j=1}^p h_j = \sum_{i=1}^m |V_i|$  and  $V = \bigcup_{i=1}^m V_i$ .

## See Also

[listStabilityMeasures](#)

## Examples

```
feats = list(1:3, 1:4, 1:5)
mat = 0.92 ^ abs(outer(1:10, 1:10, "-"))
stabilityIntersectionMBM(features = feats, sim.mat = mat, N = 1000)
```

---

 stabilityIntersectionMean

*Stability Measure Adjusted Intersection Mean*


---

### Description

The stability of feature selection is defined as the robustness of the sets of selected features with respect to small variations in the data on which the feature selection is conducted. To quantify stability, several datasets from the same data generating process can be used. Alternatively, a single dataset can be split into parts by resampling. Either way, all datasets used for feature selection must contain exactly the same features. The feature selection method of interest is applied on all of the datasets and the sets of chosen features are recorded. The stability of the feature selection is assessed based on the sets of chosen features using stability measures.

### Usage

```
stabilityIntersectionMean(features, sim.mat, threshold = 0.9,
  correction.for.chance = "estimate", N = 10000, impute.na = NULL)
```

### Arguments

features	list (length >= 2) Chosen features per dataset. Each element of the list contains the features for one dataset. The features must be given by their names (character) or indices (integerish).
sim.mat	numeric matrix Similarity matrix which contains the similarity structure of all features based on all datasets. The similarity values must be in the range of [0, 1] where 0 indicates very low similarity and 1 indicates very high similarity. If the list elements of features are integerish vectors, then the feature numbering must correspond to the ordering of sim.mat. If the list elements of features are character vectors, then sim.mat must be named and the names of sim.mat must correspond to the entries in features.
threshold	numeric(1) Threshold for indicating which features are similar and which are not. Two features are considered as similar, if and only if the corresponding entry of sim.mat is greater than or equal to threshold.
correction.for.chance	character(1) How should the expected value of the stability score (see Details) be assessed? Options are "estimate", "exact" and "none". For "estimate", N random feature sets of the same sizes as the input feature sets (features) are generated. For "exact", all possible combinations of feature sets of the same sizes as the input feature sets are used. Computation is only feasible for very small numbers of features and numbers of considered datasets (length(features)). For "none", the transformation $(score - expected)/(maximum - expected)$ is not conducted, i.e. only score is used. This is not recommended.



N	numeric(1) Number of random feature sets to consider. Only relevant if correction.for.chance is set to "estimate".
impute.na	numeric(1) In some scenarios, the stability cannot be assessed based on all feature sets. E.g. if some of the feature sets are empty, the respective pairwise comparisons yield NA as result. With which value should these missing values be imputed? NULL means no imputation.

### Details

The stability measure is defined as (see Notation)

$$\frac{2}{m(m-1)} \sum_{i=1}^{m-1} \sum_{j=i+1}^m \frac{I(V_i, V_j) - E(I(V_i, V_j))}{\sqrt{|V_i| \cdot |V_j|} - E(I(V_i, V_j))}$$

with

$$I(V_i, V_j) = |V_i \cap V_j| + \min(C(V_i, V_j), C(V_j, V_i)),$$

$$C(V_k, V_l) = \sum_{x \in V_k \setminus V_l : |G_x^{kl}| > 0} \frac{1}{|G_x^{kl}|} \sum_{y \in G_x^{kl}} \text{Similarity}(x, y)$$

and

$$G_x^{kl} = \{y \in V_l \setminus V_k : \text{Similarity}(x, y) \geq \text{threshold}\}.$$

### Value

numeric(1) Stability value.

### Notation

For the definition of all stability measures in this package, the following notation is used: Let  $V_1, \dots, V_m$  denote the sets of chosen features for the  $m$  datasets, i.e. features has length  $m$  and  $V_i$  is a set which contains the  $i$ -th entry of features. Furthermore, let  $h_j$  denote the number of sets that contain feature  $X_j$  so that  $h_j$  is the absolute frequency with which feature  $X_j$  is chosen. Also, let  $q = \sum_{j=1}^p h_j = \sum_{i=1}^m |V_i|$  and  $V = \bigcup_{i=1}^m V_i$ .

### See Also

[listStabilityMeasures](#)

### Examples

```
feats = list(1:3, 1:4, 1:5)
mat = 0.92 ^ abs(outer(1:10, 1:10, "-"))
stabilityIntersectionMean(features = feats, sim.mat = mat, N = 1000)
```

---

stabilityJaccard      *Stability Measure Jaccard*

---

### Description

The stability of feature selection is defined as the robustness of the sets of selected features with respect to small variations in the data on which the feature selection is conducted. To quantify stability, several datasets from the same data generating process can be used. Alternatively, a single dataset can be split into parts by resampling. Either way, all datasets used for feature selection must contain exactly the same features. The feature selection method of interest is applied on all of the datasets and the sets of chosen features are recorded. The stability of the feature selection is assessed based on the sets of chosen features using stability measures.

### Usage

```
stabilityJaccard(features, p = NULL, correction.for.chance = "none",
  N = 10000, impute.na = NULL)
```

### Arguments

features	list (length >= 2) Chosen features per dataset. Each element of the list contains the features for one dataset. The features must be given by their names (character) or indices (integerish).
p	numeric(1) Total number of features in the datasets. Required, if correction.for.chance is set to "estimate" or "exact".
correction.for.chance	character(1) Should a correction for chance be applied? Correction for chance means that if features are chosen at random, the expected value must be independent of the number of chosen features. To correct for chance, the original score is transformed by $(score - expected)/(maximum - expected)$ . For stability measures whose score is the average value of pairwise scores, this transformation is done for all components individually. Options are "none", "estimate" and "exact". For "none", no correction is performed, i.e. the original score is used. For "estimate", N random feature sets of the same sizes as the input feature sets (features) are generated. For "exact", all possible combinations of feature sets of the same sizes as the input feature sets are used. Computation is only feasible for very small numbers of features (p) and numbers of considered datasets (length(features)).
N	numeric(1) Number of random feature sets to consider. Only relevant if correction.for.chance is set to "estimate".
impute.na	numeric(1) In some scenarios, the stability cannot be assessed based on all feature sets. E.g.

if some of the feature sets are empty, the respective pairwise comparisons yield NA as result. With which value should these missing values be imputed? NULL means no imputation.

### Details

The stability measure is defined as (see Notation)

$$\frac{2}{m(m-1)} \sum_{i=1}^{m-1} \sum_{j=i+1}^m \frac{|V_i \cap V_j|}{|V_i \cup V_j|}.$$

### Value

numeric(1) Stability value.

### Notation

For the definition of all stability measures in this package, the following notation is used: Let  $V_1, \dots, V_m$  denote the sets of chosen features for the  $m$  datasets, i.e. features has length  $m$  and  $V_i$  is a set which contains the  $i$ -th entry of features. Furthermore, let  $h_j$  denote the number of sets that contain feature  $X_j$  so that  $h_j$  is the absolute frequency with which feature  $X_j$  is chosen. Also, let  $q = \sum_{j=1}^p h_j = \sum_{i=1}^m |V_i|$  and  $V = \bigcup_{i=1}^m V_i$ .

### References

- P. Jaccard, "Étude comparative de la distribution florale dans une portion des alpes et du jura", Bulletin de la Société Vaudoise des Sciences Naturelles, vol. 37, pp. 547-579, 1901.
- A. Bommert, J. Rahnenführer, and M. Lang, "A multi-criteria approach to find predictive and sparse models with stable feature selection for high-dimensional data", Computational and mathematical methods in medicine, 2017.

### See Also

[listStabilityMeasures](#)

### Examples

```
feats = list(1:3, 1:4, 1:5)
stabilityJaccard(features = feats)
```

---

 stabilityKappa

*Stability Measure Kappa*


---

### Description

The stability of feature selection is defined as the robustness of the sets of selected features with respect to small variations in the data on which the feature selection is conducted. To quantify stability, several datasets from the same data generating process can be used. Alternatively, a single dataset can be split into parts by resampling. Either way, all datasets used for feature selection must contain exactly the same features. The feature selection method of interest is applied on all of the datasets and the sets of chosen features are recorded. The stability of the feature selection is assessed based on the sets of chosen features using stability measures.

### Usage

```
stabilityKappa(features, p, impute.na = NULL)
```

### Arguments

features	list (length >= 2) Chosen features per dataset. Each element of the list contains the features for one dataset. The features must be given by their names (character) or indices (integerish).
p	numeric(1) Total number of features in the datasets.
impute.na	numeric(1) In some scenarios, the stability cannot be assessed based on all feature sets. E.g. if some of the feature sets are empty, the respective pairwise comparisons yield NA as result. With which value should these missing values be imputed? NULL means no imputation.

### Details

The stability measure is defined as the average kappa coefficient between all pairs of feature sets. It can be rewritten as (see Notation)

$$\frac{2}{m(m-1)} \sum_{i=1}^{m-1} \sum_{j=i+1}^m \frac{|V_i \cap V_j| - \frac{|V_i| \cdot |V_j|}{p}}{\frac{|V_i| + |V_j|}{2} - \frac{|V_i| \cdot |V_j|}{p}}$$

### Value

numeric(1) Stability value.

**Notation**

For the definition of all stability measures in this package, the following notation is used: Let  $V_1, \dots, V_m$  denote the sets of chosen features for the  $m$  datasets, i.e. `features` has length  $m$  and  $V_i$  is a set which contains the  $i$ -th entry of `features`. Furthermore, let  $h_j$  denote the number of sets that contain feature  $X_j$  so that  $h_j$  is the absolute frequency with which feature  $X_j$  is chosen. Also, let  $q = \sum_{j=1}^p h_j = \sum_{i=1}^m |V_i|$  and  $V = \bigcup_{i=1}^m V_i$ .

**References**

J. Cohen, "A coefficient of agreement for nominal scales", Educational and psychological measurement, vol. 20, no. 1, pp. 37-46, 1960.

**See Also**

[listStabilityMeasures](#)

**Examples**

```
feats = list(1:3, 1:4, 1:5)
stabilityKappa(features = feats, p = 10)
```

---

stabilityLustgarten     *Stability Measure Lustgarten*

---

**Description**

The stability of feature selection is defined as the robustness of the sets of selected features with respect to small variations in the data on which the feature selection is conducted. To quantify stability, several datasets from the same data generating process can be used. Alternatively, a single dataset can be split into parts by resampling. Either way, all datasets used for feature selection must contain exactly the same features. The feature selection method of interest is applied on all of the datasets and the sets of chosen features are recorded. The stability of the feature selection is assessed based on the sets of chosen features using stability measures.

**Usage**

```
stabilityLustgarten(features, p, impute.na = NULL)
```

**Arguments**

<code>features</code>	<code>list</code> (length $\geq 2$ ) Chosen features per dataset. Each element of the list contains the features for one dataset. The features must be given by their names (character) or indices (integerish).
<code>p</code>	<code>numeric</code> (1) Total number of features in the datasets.

impute.na      numeric(1)  
 In some scenarios, the stability cannot be assessed based on all feature sets. E.g. if some of the feature sets are empty, the respective pairwise comparisons yield NA as result. With which value should these missing values be imputed? NULL means no imputation.

### Details

The stability measure is defined as (see Notation)

$$\frac{2}{m(m-1)} \sum_{i=1}^{m-1} \sum_{j=i+1}^m \frac{|V_i \cap V_j| - \frac{|V_i| \cdot |V_j|}{p}}{\min\{|V_i|, |V_j|\} - \max\{0, |V_i| + |V_j| - p\}}$$

### Value

numeric(1) Stability value.

### Notation

For the definition of all stability measures in this package, the following notation is used: Let  $V_1, \dots, V_m$  denote the sets of chosen features for the  $m$  datasets, i.e. `features` has length  $m$  and  $V_i$  is a set which contains the  $i$ -th entry of `features`. Furthermore, let  $h_j$  denote the number of sets that contain feature  $X_j$  so that  $h_j$  is the absolute frequency with which feature  $X_j$  is chosen. Also, let  $q = \sum_{j=1}^p h_j = \sum_{i=1}^m |V_i|$  and  $V = \bigcup_{i=1}^m V_i$ .

### References

- J. L. Lustgarten, V. Gopalakrishnan, and S. Visweswaran, "Measuring stability of feature selection in biomedical datasets", AMIA Annual Symposium proceedings/AMIA Symposium. AMIA Symposium, vol. 2009, pp. 406-410, 2009.
- A. Bommert, J. Rahnenführer, and M. Lang, "A multi-criteria approach to find predictive and sparse models with stable feature selection for high-dimensional data", Computational and mathematical methods in medicine, 2017.

### See Also

[listStabilityMeasures](#)

### Examples

```
feats = list(1:3, 1:4, 1:5)
stabilityLustgarten(features = feats, p = 10)
```

---

 stabilityNogueira      *Stability Measure Nogueira*


---

## Description

The stability of feature selection is defined as the robustness of the sets of selected features with respect to small variations in the data on which the feature selection is conducted. To quantify stability, several datasets from the same data generating process can be used. Alternatively, a single dataset can be split into parts by resampling. Either way, all datasets used for feature selection must contain exactly the same features. The feature selection method of interest is applied on all of the datasets and the sets of chosen features are recorded. The stability of the feature selection is assessed based on the sets of chosen features using stability measures.

## Usage

```
stabilityNogueira(features, p, impute.na = NULL)
```

## Arguments

features	list (length >= 2) Chosen features per dataset. Each element of the list contains the features for one dataset. The features must be given by their names (character) or indices (integerish).
p	numeric(1) Total number of features in the datasets.
impute.na	numeric(1) In some scenarios, the stability cannot be assessed based on all feature sets. E.g. if some of the feature sets are empty, the respective pairwise comparisons yield NA as result. With which value should these missing values be imputed? NULL means no imputation.

## Details

The stability measure is defined as (see Notation)

$$1 - \frac{\frac{1}{p} \sum_{j=1}^p \frac{m}{m-1} \frac{h_j}{m} \left(1 - \frac{h_j}{m}\right)}{\frac{q}{mp} \left(1 - \frac{q}{mp}\right)}.$$

## Value

numeric(1) Stability value.

**Notation**

For the definition of all stability measures in this package, the following notation is used: Let  $V_1, \dots, V_m$  denote the sets of chosen features for the  $m$  datasets, i.e. features has length  $m$  and  $V_i$  is a set which contains the  $i$ -th entry of features. Furthermore, let  $h_j$  denote the number of sets that contain feature  $X_j$  so that  $h_j$  is the absolute frequency with which feature  $X_j$  is chosen. Also, let  $q = \sum_{j=1}^p h_j = \sum_{i=1}^m |V_i|$  and  $V = \bigcup_{i=1}^m V_i$ .

**References**

S. Nogueira, "Quantifying the Stability of Feature Selection", Diss. PhD thesis, University of Manchester, 2018.

**See Also**

[listStabilityMeasures](#)

**Examples**

```
feats = list(1:3, 1:4, 1:5)
stabilityNogueira(features = feats, p = 10)
```

---

stabilityNovovicova    *Stability Measure Novovičová*

---

**Description**

The stability of feature selection is defined as the robustness of the sets of selected features with respect to small variations in the data on which the feature selection is conducted. To quantify stability, several datasets from the same data generating process can be used. Alternatively, a single dataset can be split into parts by resampling. Either way, all datasets used for feature selection must contain exactly the same features. The feature selection method of interest is applied on all of the datasets and the sets of chosen features are recorded. The stability of the feature selection is assessed based on the sets of chosen features using stability measures.

**Usage**

```
stabilityNovovicova(features, p = NULL, correction.for.chance = "none",
  N = 10000, impute.na = NULL)
```

**Arguments**

features	list (length >= 2) Chosen features per dataset. Each element of the list contains the features for one dataset. The features must be given by their names (character) or indices (integerish).
p	numeric(1) Total number of features in the datasets. Required, if correction.for.chance is set to "estimate" or "exact".



correction.for.chance	character(1) Should a correction for chance be applied? Correction for chance means that if features are chosen at random, the expected value must be independent of the number of chosen features. To correct for chance, the original score is transformed by $(score - expected)/(maximum - expected)$ . For stability measures whose score is the average value of pairwise scores, this transformation is done for all components individually. Options are "none", "estimate" and "exact". For "none", no correction is performed, i.e. the original score is used. For "estimate", N random feature sets of the same sizes as the input feature sets (features) are generated. For "exact", all possible combinations of feature sets of the same sizes as the input feature sets are used. Computation is only feasible for very small numbers of features (p) and numbers of considered datasets (length(features)).
N	numeric(1) Number of random feature sets to consider. Only relevant if correction.for.chance is set to "estimate".
impute.na	numeric(1) In some scenarios, the stability cannot be assessed based on all feature sets. E.g. if some of the feature sets are empty, the respective pairwise comparisons yield NA as result. With which value should these missing values be imputed? NULL means no imputation.

## Details

The stability measure is defined as (see Notation)

$$\frac{1}{q \log_2(m)} \sum_{j: X_j \in V} h_j \log_2(h_j).$$

## Value

numeric(1) Stability value.

## Notation

For the definition of all stability measures in this package, the following notation is used: Let  $V_1, \dots, V_m$  denote the sets of chosen features for the  $m$  datasets, i.e. features has length  $m$  and  $V_i$  is a set which contains the  $i$ -th entry of features. Furthermore, let  $h_j$  denote the number of sets that contain feature  $X_j$  so that  $h_j$  is the absolute frequency with which feature  $X_j$  is chosen. Also, let  $q = \sum_{j=1}^p h_j = \sum_{i=1}^m |V_i|$  and  $V = \bigcup_{i=1}^m V_i$ .

## References

- J. Novovičová, P. Somol, and P. Pudil, "A new measure of feature selection algorithms' stability", in Proceedings of the 2009 IEEE International Conference on Data Mining Workshops, ICDMW 2009, pp. 382–387, December 2009.
- A. Bommert, J. Rahnenführer, and M. Lang, "A multi-criteria approach to find predictive and sparse models with stable feature selection for high-dimensional data", Computational and mathematical methods in medicine, 2017.

**See Also**[listStabilityMeasures](#)**Examples**

```
feats = list(1:3, 1:4, 1:5)
stabilityNovovicova(features = feats)
```

---

stabilityOchiai	<i>Stability Measure Ochiai</i>
-----------------	---------------------------------

---

**Description**

The stability of feature selection is defined as the robustness of the sets of selected features with respect to small variations in the data on which the feature selection is conducted. To quantify stability, several datasets from the same data generating process can be used. Alternatively, a single dataset can be split into parts by resampling. Either way, all datasets used for feature selection must contain exactly the same features. The feature selection method of interest is applied on all of the datasets and the sets of chosen features are recorded. The stability of the feature selection is assessed based on the sets of chosen features using stability measures.

**Usage**

```
stabilityOchiai(features, p = NULL, correction.for.chance = "none",
  N = 10000, impute.na = NULL)
```

**Arguments**

features	list (length >= 2) Chosen features per dataset. Each element of the list contains the features for one dataset. The features must be given by their names (character) or indices (integerish).
p	numeric(1) Total number of features in the datasets. Required, if <code>correction.for.chance</code> is set to "estimate" or "exact".
correction.for.chance	character(1) Should a correction for chance be applied? Correction for chance means that if features are chosen at random, the expected value must be independent of the number of chosen features. To correct for chance, the original score is transformed by $(score - expected) / (maximum - expected)$ . For stability measures whose score is the average value of pairwise scores, this transformation is done for all components individually. Options are "none", "estimate" and "exact". For "none", no correction is performed, i.e. the original score is used. For "estimate", N random feature sets of the same sizes as the input feature sets (features) are generated. For "exact", all possible combinations of feature sets

of the same sizes as the input feature sets are used. Computation is only feasible for very small numbers of features ( $p$ ) and numbers of considered datasets ( $\text{length}(\text{features})$ ).

N	numeric(1) Number of random feature sets to consider. Only relevant if correction.for.chance is set to "estimate".
impute.na	numeric(1) In some scenarios, the stability cannot be assessed based on all feature sets. E.g. if some of the feature sets are empty, the respective pairwise comparisons yield NA as result. With which value should these missing values be imputed? NULL means no imputation.

### Details

The stability measure is defined as (see Notation)

$$\frac{2}{m(m-1)} \sum_{i=1}^{m-1} \sum_{j=i+1}^m \frac{|V_i \cap V_j|}{\sqrt{|V_i| \cdot |V_j|}}.$$

### Value

numeric(1) Stability value.

### Notation

For the definition of all stability measures in this package, the following notation is used: Let  $V_1, \dots, V_m$  denote the sets of chosen features for the  $m$  datasets, i.e.  $\text{features}$  has length  $m$  and  $V_i$  is a set which contains the  $i$ -th entry of  $\text{features}$ . Furthermore, let  $h_j$  denote the number of sets that contain feature  $X_j$  so that  $h_j$  is the absolute frequency with which feature  $X_j$  is chosen. Also, let  $q = \sum_{j=1}^p h_j = \sum_{i=1}^m |V_i|$  and  $V = \bigcup_{i=1}^m V_i$ .

### References

- A. Ochiai, "Zoogeographic studies on the soleoid fishes found in Japan and its neighbouring regions", Bulletin of the Japanese Society for the Science of Fish, vol. 22, no. 9, pp. 526-530, 1957.
- A. Bommert, J. Rahnenführer, and M. Lang, "A multi-criteria approach to find predictive and sparse models with stable feature selection for high-dimensional data", Computational and mathematical methods in medicine, 2017.

### See Also

[listStabilityMeasures](#)

### Examples

```
feats = list(1:3, 1:4, 1:5)
stabilityOchiai(features = feats)
```

---

 stabilityPhi

*Stability Measure Phi*


---

### Description

The stability of feature selection is defined as the robustness of the sets of selected features with respect to small variations in the data on which the feature selection is conducted. To quantify stability, several datasets from the same data generating process can be used. Alternatively, a single dataset can be split into parts by resampling. Either way, all datasets used for feature selection must contain exactly the same features. The feature selection method of interest is applied on all of the datasets and the sets of chosen features are recorded. The stability of the feature selection is assessed based on the sets of chosen features using stability measures.

### Usage

```
stabilityPhi(features, p, impute.na = NULL)
```

### Arguments

features	list (length >= 2) Chosen features per dataset. Each element of the list contains the features for one dataset. The features must be given by their names (character) or indices (integerish).
p	numeric(1) Total number of features in the datasets.
impute.na	numeric(1) In some scenarios, the stability cannot be assessed based on all feature sets. E.g. if some of the feature sets are empty, the respective pairwise comparisons yield NA as result. With which value should these missing values be imputed? NULL means no imputation.

### Details

The stability measure is defined as the average phi coefficient between all pairs of feature sets. It can be rewritten as (see Notation)

$$\frac{2}{m(m-1)} \sum_{i=1}^{m-1} \sum_{j=i+1}^m \frac{|V_i \cap V_j| - \frac{|V_i| \cdot |V_j|}{p}}{\sqrt{|V_i| \left(1 - \frac{|V_i|}{p}\right) \cdot |V_j| \left(1 - \frac{|V_j|}{p}\right)}}$$

### Value

numeric(1) Stability value.

**Notation**

For the definition of all stability measures in this package, the following notation is used: Let  $V_1, \dots, V_m$  denote the sets of chosen features for the  $m$  datasets, i.e. features has length  $m$  and  $V_i$  is a set which contains the  $i$ -th entry of features. Furthermore, let  $h_j$  denote the number of sets that contain feature  $X_j$  so that  $h_j$  is the absolute frequency with which feature  $X_j$  is chosen. Also, let  $q = \sum_{j=1}^p h_j = \sum_{i=1}^m |V_i|$  and  $V = \bigcup_{i=1}^m V_i$ .

**References**

- S. Nogueira and G. Brown, "Measuring the stability of feature selection", in Machine Learning and Knowledge Discovery in Databases, vol. 9852 of Lecture Notes in Computer Science, pp. 442-457, Springer International Publishing, Cham, 2016.
- A. Bommert, J. Rahnenführer, and M. Lang, "A multi-criteria approach to find predictive and sparse models with stable feature selection for high-dimensional data", Computational and mathematical methods in medicine, 2017.

**See Also**

[listStabilityMeasures](#)

**Examples**

```
feats = list(1:3, 1:4, 1:5)
stabilityPhi(features = feats, p = 10)
```

---

stabilitySomol

*Stability Measure Somol*

---

**Description**

The stability of feature selection is defined as the robustness of the sets of selected features with respect to small variations in the data on which the feature selection is conducted. To quantify stability, several datasets from the same data generating process can be used. Alternatively, a single dataset can be split into parts by resampling. Either way, all datasets used for feature selection must contain exactly the same features. The feature selection method of interest is applied on all of the datasets and the sets of chosen features are recorded. The stability of the feature selection is assessed based on the sets of chosen features using stability measures.

**Usage**

```
stabilitySomol(features, p, impute.na = NULL)
```

**Arguments**

features	list (length >= 2) Chosen features per dataset. Each element of the list contains the features for one dataset. The features must be given by their names (character) or indices (integerish).
p	numeric(1) Total number of features in the datasets.
impute.na	numeric(1) In some scenarios, the stability cannot be assessed based on all feature sets. E.g. if some of the feature sets are empty, the respective pairwise comparisons yield NA as result. With which value should these missing values be imputed? NULL means no imputation.

**Details**

The stability measure is defined as (see Notation)

$$\frac{\left( \sum_{j=1}^p \frac{h_j}{q} \frac{h_{j-1}}{m-1} \right) - c_{\min}}{c_{\max} - c_{\min}}$$

with

$$c_{\min} = \frac{q^2 - p(q - q \bmod p) - (q \bmod p)^2}{pq(m-1)},$$

$$c_{\max} = \frac{(q \bmod m)^2 + q(m-1) - (q \bmod m)m}{q(m-1)}.$$

**Value**

numeric(1) Stability value.

**Notation**

For the definition of all stability measures in this package, the following notation is used: Let  $V_1, \dots, V_m$  denote the sets of chosen features for the  $m$  datasets, i.e. features has length  $m$  and  $V_i$  is a set which contains the  $i$ -th entry of features. Furthermore, let  $h_j$  denote the number of sets that contain feature  $X_j$  so that  $h_j$  is the absolute frequency with which feature  $X_j$  is chosen. Also, let  $q = \sum_{j=1}^p h_j = \sum_{i=1}^m |V_i|$  and  $V = \bigcup_{i=1}^m V_i$ .

**References**

- P. Somol and J. Novovičová, "Evaluating stability and comparing output of feature selectors that optimize feature subset cardinality", IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. 32, no. 11, pp. 1921-1939, 2010.
- A. Bommert, J. Rahnenführer, and M. Lang, "A multi-criteria approach to find predictive and sparse models with stable feature selection for high-dimensional data", Computational and mathematical methods in medicine, 2017.

**See Also**[listStabilityMeasures](#)**Examples**

```
feats = list(1:3, 1:4, 1:5)
stabilitySomol(features = feats, p = 10)
```

---

stabilityUnadjusted     *Stability Measure Unadjusted*

---

**Description**

The stability of feature selection is defined as the robustness of the sets of selected features with respect to small variations in the data on which the feature selection is conducted. To quantify stability, several datasets from the same data generating process can be used. Alternatively, a single dataset can be split into parts by resampling. Either way, all datasets used for feature selection must contain exactly the same features. The feature selection method of interest is applied on all of the datasets and the sets of chosen features are recorded. The stability of the feature selection is assessed based on the sets of chosen features using stability measures.

**Usage**

```
stabilityUnadjusted(features, p, impute.na = NULL)
```

**Arguments**

features	list (length >= 2) Chosen features per dataset. Each element of the list contains the features for one dataset. The features must be given by their names (character) or indices (integerish).
p	numeric(1) Total number of features in the datasets.
impute.na	numeric(1) In some scenarios, the stability cannot be assessed based on all feature sets. E.g. if some of the feature sets are empty, the respective pairwise comparisons yield NA as result. With which value should these missing values be imputed? NULL means no imputation.

**Details**

The stability measure is defined as (see Notation)

$$\frac{2}{m(m-1)} \sum_{i=1}^{m-1} \sum_{j=i+1}^m \frac{|V_i \cap V_j| - \frac{|V_i| \cdot |V_j|}{p}}{\sqrt{|V_i| \cdot |V_j| - \frac{|V_i| \cdot |V_j|}{p}}}$$

This is what [stabilityIntersectionMBM](#), [stabilityIntersectionGreedy](#), [stabilityIntersectionCount](#) and [stabilityIntersectionMean](#) become, when there are no similar features.

**Value**

numeric(1) Stability value.

**Notation**

For the definition of all stability measures in this package, the following notation is used: Let  $V_1, \dots, V_m$  denote the sets of chosen features for the  $m$  datasets, i.e. `features` has length  $m$  and  $V_i$  is a set which contains the  $i$ -th entry of `features`. Furthermore, let  $h_j$  denote the number of sets that contain feature  $X_j$  so that  $h_j$  is the absolute frequency with which feature  $X_j$  is chosen. Also, let  $q = \sum_{j=1}^p h_j = \sum_{i=1}^m |V_i|$  and  $V = \bigcup_{i=1}^m V_i$ .

**See Also**

[listStabilityMeasures](#)

**Examples**

```
feats = list(1:3, 1:4, 1:5)
stabilityUnadjusted(features = feats, p = 10)
```

---

stabilityWald

*Stability Measure Wald*

---

**Description**

The stability of feature selection is defined as the robustness of the sets of selected features with respect to small variations in the data on which the feature selection is conducted. To quantify stability, several datasets from the same data generating process can be used. Alternatively, a single dataset can be split into parts by resampling. Either way, all datasets used for feature selection must contain exactly the same features. The feature selection method of interest is applied on all of the datasets and the sets of chosen features are recorded. The stability of the feature selection is assessed based on the sets of chosen features using stability measures.

**Usage**

```
stabilityWald(features, p, impute.na = NULL)
```

**Arguments**

<code>features</code>	<code>list</code> (length $\geq 2$ ) Chosen features per dataset. Each element of the list contains the features for one dataset. The features must be given by their names (character) or indices (integerish).
<code>p</code>	numeric(1) Total number of features in the datasets.



impute.na      numeric(1)  
 In some scenarios, the stability cannot be assessed based on all feature sets. E.g. if some of the feature sets are empty, the respective pairwise comparisons yield NA as result. With which value should these missing values be imputed? NULL means no imputation.

## Details

The stability measure is defined as (see Notation)

$$\frac{2}{m(m-1)} \sum_{i=1}^{m-1} \sum_{j=i+1}^m \frac{|V_i \cap V_j| - \frac{|V_i| \cdot |V_j|}{p}}{\min\{|V_i|, |V_j|\} - \frac{|V_i| \cdot |V_j|}{p}}.$$

## Value

numeric(1) Stability value.

## Notation

For the definition of all stability measures in this package, the following notation is used: Let  $V_1, \dots, V_m$  denote the sets of chosen features for the  $m$  datasets, i.e. `features` has length  $m$  and  $V_i$  is a set which contains the  $i$ -th entry of `features`. Furthermore, let  $h_j$  denote the number of sets that contain feature  $X_j$  so that  $h_j$  is the absolute frequency with which feature  $X_j$  is chosen. Also, let  $q = \sum_{j=1}^p h_j = \sum_{i=1}^m |V_i|$  and  $V = \bigcup_{i=1}^m V_i$ .

## References

- Wald, R., T. M. Khoshgoftaar, and A. Napolitano (2013). “Stability of filter-and wrapper-based feature subset selection.” In: 2013 IEEE 25th International Conference on Tools with Artificial Intelligence, pp. 374–380.

## See Also

[listStabilityMeasures](#)

## Examples

```
feats = list(1:3, 1:4, 1:5)
stabilityWald(features = feats, p = 10)
```

---

 stabilityYu

*Stability Measure Yu*


---

### Description

The stability of feature selection is defined as the robustness of the sets of selected features with respect to small variations in the data on which the feature selection is conducted. To quantify stability, several datasets from the same data generating process can be used. Alternatively, a single dataset can be split into parts by resampling. Either way, all datasets used for feature selection must contain exactly the same features. The feature selection method of interest is applied on all of the datasets and the sets of chosen features are recorded. The stability of the feature selection is assessed based on the sets of chosen features using stability measures.

### Usage

```
stabilityYu(features, sim.mat, threshold = 0.9,
  correction.for.chance = "estimate", N = 10000, impute.na = NULL)
```

### Arguments

features	list (length >= 2) Chosen features per dataset. Each element of the list contains the features for one dataset. The features must be given by their names (character) or indices (integerish).
sim.mat	numeric matrix Similarity matrix which contains the similarity structure of all features based on all datasets. The similarity values must be in the range of [0, 1] where 0 indicates very low similarity and 1 indicates very high similarity. If the list elements of features are integerish vectors, then the feature numbering must correspond to the ordering of sim.mat. If the list elements of features are character vectors, then sim.mat must be named and the names of sim.mat must correspond to the entries in features.
threshold	numeric(1) Threshold for indicating which features are similar and which are not. Two features are considered as similar, if and only if the corresponding entry of sim.mat is greater than or equal to threshold.
correction.for.chance	character(1) How should the expected value of the stability score (see Details) be assessed? Options are "estimate", "exact" and "none". For "estimate", N random feature sets of the same sizes as the input feature sets (features) are generated. For "exact", all possible combinations of feature sets of the same sizes as the input feature sets are used. Computation is only feasible for very small numbers of features and numbers of considered datasets (length(features)). For "none", the transformation $(score - expected)/(maximum - expected)$ is not conducted, i.e. only score is used. This is not recommended.

N	numeric(1) Number of random feature sets to consider. Only relevant if correction.for.chance is set to "estimate".
impute.na	numeric(1) In some scenarios, the stability cannot be assessed based on all feature sets. E.g. if some of the feature sets are empty, the respective pairwise comparisons yield NA as result. With which value should these missing values be imputed? NULL means no imputation.

### Details

Let  $O_{ij}$  denote the number of features in  $V_i$  that are not shared with  $V_j$  but that have a highly similar feature in  $V_j$ :

$$O_{ij} = |\{x \in (V_i \setminus V_j) : \exists y \in (V_j \setminus V_i) \text{ with } \text{Similarity}(x, y) \geq \text{threshold}\}|.$$

Then the stability measure is defined as (see Notation)

$$\frac{2}{m(m-1)} \sum_{i=1}^{m-1} \sum_{j=i+1}^m \frac{I(V_i, V_j) - E(I(V_i, V_j))}{\frac{|V_i|+|V_j|}{2} - E(I(V_i, V_j))}$$

with

$$I(V_i, V_j) = |V_i \cap V_j| + \frac{O_{ij} + O_{ji}}{2}.$$

Note that this definition slightly differs from its original in order to make it suitable for arbitrary datasets and similarity measures and applicable in situations with  $|V_i| \neq |V_j|$ .

### Value

numeric(1) Stability value.

### Notation

For the definition of all stability measures in this package, the following notation is used: Let  $V_1, \dots, V_m$  denote the sets of chosen features for the  $m$  datasets, i.e. features has length  $m$  and  $V_i$  is a set which contains the  $i$ -th entry of features. Furthermore, let  $h_j$  denote the number of sets that contain feature  $X_j$  so that  $h_j$  is the absolute frequency with which feature  $X_j$  is chosen. Also, let  $q = \sum_{j=1}^p h_j = \sum_{i=1}^m |V_i|$  and  $V = \bigcup_{i=1}^m V_i$ .

### References

- L. Yu, Y. Han, and M. E. Berens, "Stable gene selection from microarray data via sample weighting", IEEE/ACM Transactions on Computational Biology and Bioinformatics, vol. 9, no. 1, pp. 262-272, 2012.
- M. Zhang, L. Zhang, J. Zou, C. Yao, H. Xiao, Q. Liu, J. Wang, D. Wang, C. Wang, and Z. Guo, "Evaluating reproducibility of differential expression discoveries in microarray studies by considering correlated molecular changes", Bioinformatics, vol. 25, no. 13, pp. 1662-1668, 2009.

**See Also**

[listStabilityMeasures](#)

**Examples**

```
feats = list(1:3, 1:4, 1:5)
mat = 0.92 ^ abs(outer(1:10, 1:10, "-"))
stabilityYu(features = feats, sim.mat = mat, N = 1000)
```

---

 stabilityZucknick

*Stability Measure Zucknick*


---

**Description**

The stability of feature selection is defined as the robustness of the sets of selected features with respect to small variations in the data on which the feature selection is conducted. To quantify stability, several datasets from the same data generating process can be used. Alternatively, a single dataset can be split into parts by resampling. Either way, all datasets used for feature selection must contain exactly the same features. The feature selection method of interest is applied on all of the datasets and the sets of chosen features are recorded. The stability of the feature selection is assessed based on the sets of chosen features using stability measures.

**Usage**

```
stabilityZucknick(features, sim.mat, threshold = 0.9,
  correction.for.chance = "none", N = 10000, impute.na = NULL)
```

**Arguments**

features	list (length >= 2) Chosen features per dataset. Each element of the list contains the features for one dataset. The features must be given by their names (character) or indices (integerish).
sim.mat	numeric matrix Similarity matrix which contains the similarity structure of all features based on all datasets. The similarity values must be in the range of [0, 1] where 0 indicates very low similarity and 1 indicates very high similarity. If the list elements of features are integerish vectors, then the feature numbering must correspond to the ordering of sim.mat. If the list elements of features are character vectors, then sim.mat must be named and the names of sim.mat must correspond to the entries in features.
threshold	numeric(1) Threshold for indicating which features are similar and which are not. Two features are considered as similar, if and only if the corresponding entry of sim.mat is greater than or equal to threshold.

correction.for.chance	character(1) Should a correction for chance be applied? Correction for chance means that if features are chosen at random, the expected value must be independent of the number of chosen features. To correct for chance, the original score is transformed by $(score - expected)/(maximum - expected)$ . For stability measures whose score is the average value of pairwise scores, this transformation is done for all components individually. Options are "none", "estimate" and "exact". For "none", no correction is performed, i.e. the original score is used. For "estimate", N random feature sets of the same sizes as the input feature sets (features) are generated. For "exact", all possible combinations of feature sets of the same sizes as the input feature sets are used. Computation is only feasible for very small numbers of features (p) and numbers of considered datasets (length(features)).
N	numeric(1) Number of random feature sets to consider. Only relevant if correction.for.chance is set to "estimate".
impute.na	numeric(1) In some scenarios, the stability cannot be assessed based on all feature sets. E.g. if some of the feature sets are empty, the respective pairwise comparisons yield NA as result. With which value should these missing values be imputed? NULL means no imputation.

## Details

The stability measure is defined as

$$\frac{2}{m(m-1)} \sum_{i=1}^{m-1} \sum_{j=i+1}^m \frac{|V_i \cap V_j| + C(V_i, V_j) + C(V_j, V_i)}{|V_i \cup V_j|}$$

with

$$C(V_k, V_l) = \frac{1}{|V_l|} \sum_{(x,y) \in V_k \times (V_l \setminus V_k) \text{ with } Similarity(x,y) \geq threshold} Similarity(x,y).$$

Note that this definition slightly differs from its original in order to make it suitable for arbitrary similarity measures.

## Value

numeric(1) Stability value.

## Notation

For the definition of all stability measures in this package, the following notation is used: Let  $V_1, \dots, V_m$  denote the sets of chosen features for the  $m$  datasets, i.e. `features` has length  $m$  and  $V_i$  is a set which contains the  $i$ -th entry of `features`. Furthermore, let  $h_j$  denote the number of sets that contain feature  $X_j$  so that  $h_j$  is the absolute frequency with which feature  $X_j$  is chosen. Also, let  $q = \sum_{j=1}^p h_j = \sum_{i=1}^m |V_i|$  and  $V = \bigcup_{i=1}^m V_i$ .

**References**

- M. Zucknick, S. Richardson, and E. Stronach, "Comparing the characteristics of gene expression profiles derived by univariate and multivariate classification methods", *Statistical Applications in Genetics and Molecular Biology*, vol. 7, no. 1, pp. 1-34, 2008.
- A. Bommert, J. Rahnenführer, and M. Lang, "A multi-criteria approach to find predictive and sparse models with stable feature selection for high-dimensional data", *Computational and mathematical methods in medicine*, 2017.

**See Also**

[listStabilityMeasures](#)

**Examples**

```
feats = list(1:3, 1:4, 1:5)
mat = 0.92 ^ abs(outer(1:10, 1:10, "-"))
stabilityZucknick(features = feats, sim.mat = mat)
```

# Index

`listStabilityMeasures`, 3, 6, 8, 9, 11, 13,  
15, 17, 19, 21, 22, 24, 26, 27, 29,  
31–33, 36, 38

`plotFeatures`, 3

`stabilityDavis`, 4

`stabilityDice`, 6

`stabilityHamming`, 8

`stabilityIntersectionCount`, 10, 31

`stabilityIntersectionGreedy`, 12, 31

`stabilityIntersectionMBM`, 13, 14, 31

`stabilityIntersectionMean`, 16, 31

`stabilityJaccard`, 18

`stabilityKappa`, 20

`stabilityLustgarten`, 21

`stabilityNogueira`, 23

`stabilityNovovicova`, 24

`stabilityOchiai`, 26

`stabilityPhi`, 28

`stabilitySomol`, 29

`stabilityUnadjusted`, 31

`stabilityWald`, 32

`stabilityYu`, 34

`stabilityZucknick`, 36

`stabm` (`stabm`-package), 2

`stabm`-package, 2