Package: TestDataImputation (via r-universe)

October 15, 2026

Type Package

Title Missing Item Responses Imputation for Test and Assessment Data

Version 2.3

Date 2021-10-18

Author Shenghai Dai [aut, cre], Xiaolin Wang [aut], Dubravka Svetina [aut]

Maintainer Shenghai Dai <s.dai@wsu.edu>

Description Functions for imputing missing item responses for dichotomous and polytomous test and assessment data. This package enables missing imputation methods that are suitable for test and assessment data, including: listwise (LW) deletion (see De Ayala et al. 2001 <doi:10.1111/j.1745-3984.2001.tb01124.x>), treating as incorrect (IN, see Lord, 1974 <doi:10.1111/j.1745-3984.1974.tb00996.x>; Mislevy & Wu, 1996 <doi:10.1002/j.2333-8504.1996.tb01708.x>; Pohl et al., 2014 <doi:10.1177/0013164413504926>), person mean imputation (PM), item mean imputation (IM), two-way (TW) and response function (RF) imputation, (see Sijtsma & van der Ark, 2003 <doi:10.1207/s15327906mbr3804_4>), logistic regression (LR) imputation, predictive mean matching (PMM), and expectation–maximization (EM) imputation (see Finch, 2008 <doi:10.1111/j.1745-3984.2008.00062.x>).

Depends R (>= 3.4.0), stats, mice, Amelia

Encoding UTF-8

NeedsCompilation no

LazyData true

License GPL (>= 2)

URL https://daiscode.github.io/TestDataImputation/

RoxygenNote 7.1.2

Config/pak/sysreqs cmake make libicu-dev libx11-dev zlib1g-dev

Repository https://daiscode.r-universe.dev

Date/Publication 2026-09-15 08:04:01 UTC

RemoteUrl https://github.com/daiscode/testdataimputation

RemoteRef HEAD

RemoteSha 0674e11f1f3836f75211084c260e53c74520a994

Contents

EMimpute 2
ImputeTestData 3
ItemMean 5
Listwise 6
LogisticReg 6
micePMM 7
PersonMean 8
ResponseFun 9
test.data 10
TreatIncorrect 11
Twoway 12

Index **13**

`EMimpute` *EM Imputation*

Description

This function imputes for all missing responses using EM imputation (see Finch, 2008) <doi: 10.1111/j.1745-3984.2008.00062.x>. The Amelia package (Honaker et al., 2011 <doi: 10.18637/jss.v045.i07>) is used for the imputation. Integrated scores are then obtained by rounding imputed values to the closest possible response value.

Usage

```
EMimpute(test.data, Mvalue = "NA", max.score = 1, round.decimal = 0)
```

Arguments

`test.data`	Test data set (a data frame or a matrix) containing missing responses. Missing values are coded as NA or other values (e.g., 8, 9).
`Mvalue`	Missing response indicators in the data (e.g. "NA", "8", "9", etc.). Mvalue="NA" by default.
`max.score`	The max possible response value in the test data. By default max.score=1 (i.e.,binary test data).
`round.decimal`	The number of digits or decimal places for the imputed value. The default value is 0.

Value

A data frame with all missing responses replaced by integrated imputed values.

References

Finch, H. (2008). "Estimation of Item Response Theory Parameters in the Presence of Missing Data." Journal of Educational Measurement, 45(3), 225-245. doi: 10.1111/j.1745-3984.2008.00062.x.

Honaker, J., King, G., & Blackwell, M. (2011). "Amelia II: A program for missing data." Journal of statistical software, 45(1), 1-47. doi: 10.18637/jss.v045.i07.

Examples

```
EMimpute(test.data, Mvalue="NA",max.score=1,round.decimal=0)
```

`ImputeTestData` *This main function imputes for missing responses using selected method*

Description

This function imputes for all missing responses using the selected imputation method. Integrated scores are obtained by rounding imputed values to the closest possible response value.

Usage

```
ImputeTestData(
  test.data,
  Mvalue = "NA",
  max.score = 1,
  method = "LW",
  round.decimal = 0
)
```

Arguments

`test.data`	Test data set (a data frame or a matrix) containing missing responses. Missing values are coded as NA or other values (e.g., 8, 9).
`Mvalue`	Missing response indicators in the data (e.g. "NA", "8", "9", etc.). Mvalue="NA" by default.
`max.score`	The max possible response value in test data. By default max.score=1 (i.e.,binary test data). max.score = 2 if the response categories are (0, 1, 2), etc. Note: For IN and RF, the lowest response value should be zero (i.e., incorrect).
`method`	Missing response imputation methods. "LW" (by default) represents listwise that deletes all examinees who reported missing responses (see De Ayala et al. 2001 <doi:10.1111/j.1745-3984.2001.tb01124.x>) "IN" means treating all missing responses as incorrect (see Lord, 1974 <doi:

10.1111/j.1745-3984.1974.tb00996.x>; Mislevy & Wu, 1996 <doi: 10.1002/j.2333-8504.1996.tb01708.x>; Pohl et al., 2014 <doi: 10.1177/0013164413504926>).
"PM" imputes for all missing responses of an examinee by his/her mean on the available items.
"IM" imputes for all missing responses of an item by its mean on the available responses.
"TW" imputes for all missing responses using two-way imputation (if an examinee has no response to all items, the missing responses are replaced by item means first; see Sijtsma & van der Ark, 2003 <doi: 10.1207/s15327906mbr3804_4>).
"RF" imputes for all missing responses using response function imputation (Sijtsma & van der Ark, 2003 <doi: 10.1207/s15327906mbr3804_4>).
"LR" imputes for all missing responses using logistic regression (for binary responses) and polytomous regression (for polytmous responses) with mice package (Van Buuren & Groothuis-Oudshoorn, 2011 <doi: 10.18637/jss.v045.i03>).
"PMM" imputes for all missing responses using predictive mean matching with mice package (Van Buuren & Groothuis-Oudshoorn, 2011 <doi: 10.18637/jss.v045.i03>).
"EM" imputes for all missing responses using EM imputation with the Amelia package (Honaker et al., 2011 <doi: 10.18637/jss.v045.i07>). The imputed values are then rounded to the closest possible response value. (see Finch, 2008 <doi: 10.1111/j.1745-3984.2008.00062.x>).

`round.decimal` The number of digits or decimal places for the imputed value. The default value is 0.

Value

A data frame with all missing responses replaced by integrated imputed values.

References

De Ayala, R. J., Plake, B. S., & Impara, J. C. (2001). "The impact of omitted responses on the accuracy of ability estimation in item response theory." Journal of Educational Measurement, 38(3), 213-234. doi:10.1111/j.1745-3984.2001.tb01124.x.

Finch, H. (2008). "Estimation of Item Response Theory Parameters in the Presence of Missing Data." Journal of Educational Measurement, 45(3), 225-245. doi: 10.1111/j.1745-3984.2008.00062.x.

Honaker, J., King, G., & Blackwell, M. (2011). "Amelia II: A program for missing data." Journal of statistical software, 45(1), 1-47. doi: 10.18637/jss.v045.i07.

Lord, F. M. (1974). " Quick estimates of the relative efficiency of two tests as a function of ability level." Journal of Educational Measurement, 11(4), 247-254. doi: 10.1111/j.1745-3984.1974.tb00996.x.

Mislevy, R. J., & Wu, P. K. (1996). " Missing responses and IRT ability estimation: Omits, choice, time limits, and adaptive testing. " ETS Research Report Series, 1996(2), i-36. doi: 10.1002/j.2333-8504.1996.tb01708.x.

Pohl, S., Gräfe, L., & Rose, N. (2014). "Dealing with omitted and not-reached items in competence tests evaluating approaches accounting for missing responses in item response theory models. " Educational and Psychological Measurement, 74(3), 423-452. doi: 10.1177/0013164413504926.

Sijtsma, K., & Van der Ark, L. A. (2003). "Investigation and treatment of missing item scores in test and questionnaire data." Multivariate Behavioral Research, 38(4), 505-528. doi: 10.1207/s15327906mbr3804_4.

Van Buuren, S., & Groothuis-Oudshoorn, K. (2011). "mice: Multivariate imputation by chained equations in R." Journal of statistical software, 45(1), 1-67. DOI: 10.18637/jss.v045.i03.

Examples

```
ImputeTestData(test.data, Mvalue="NA",max.score=1, method ="TW",round.decimal=0)
```

`ItemMean` *Item Mean (IM) Imputation*

Description

This function imputes for all missing responses of an item by its mean (i.e., IM) on the available responses. Integrated scores for items are obtained by rounding their means to the closest possible response value.

Usage

```
ItemMean(test.data, Mvalue = "NA", max.score = 1, round.decimal = 0)
```

Arguments

`test.data`	Test data set (a data frame or a matrix) containing missing responses. Missing values are coded as NA or other values (e.g., 8, 9).
`Mvalue`	Missing response indicators in the data (e.g. "NA", "8", "9", etc.). Mvalue="NA" by default.
`max.score`	The max possible response value in test data. By default max.score=1 (i.e.,binary test data).
`round.decimal`	The number of digits or decimal places for the imputed value. The default value is 0.

Value

A data frame with all missing responses replaced by Integrated item means.

Examples

```
ItemMean(test.data, Mvalue="NA",max.score=1,round.decimal=0)
```

Listwise *Listwise Deletion (LW)*

Description

This function deletes examinees who report missing responses (see De Ayala et al. 2001) <doi:10.1111/j.1745-3984.2001.tb01124.x>.

Usage

```
Listwise(test.data, Mvalue = "NA")
```

Arguments

test.data	Test data set (a data frame or a matrix) containing missing responses. Missing values are coded as NA or other values (e.g., 8, 9).
Mvalue	Missing response indicators in the data (e.g. "NA", "8", "9", etc.). Mvalue="NA" by default.

Value

A data frame with no missing responses.

References

De Ayala, R. J., Plake, B. S., & Impara, J. C. (2001). "The impact of omitted responses on the accuracy of ability estimation in item response theory." Journal of Educational Measurement, 38(3), 213–234. doi:10.1111/j.1745-3984.2001.tb01124.x.

Examples

```
Listwise(test.data, Mvalue="NA")
```

LogisticReg *Logistic Regression (LR) Imputation*

Description

This function imputes for all missing responses using logistic regression (for binary responses) or polytomous regression (for polytomous responses). The mice () function with default settings from the mice package (Van Buuren & Groothuis-Oudshoorn, 2011 <doi: 10.18637/jss.v045.i03>) is used to impute for the missing responses.

Usage

```
LogisticReg(test.data, Mvalue = "NA", max.score = 1)
```

Arguments

test.data	Test data set (a data frame or a matrix) containing missing responses. Missing values are coded as NA or other values (e.g., 8, 9).
Mvalue	Missing response indicators in the data (e.g. "NA", "8", "9", etc.). Mvalue="NA" by default.
max.score	The max possible response value in test data. By default max.score=1 (i.e.,binary test data).

Value

A data frame with all missing responses replaced by integrated imputed values.

References

Van Buuren, S., & Groothuis-Oudshoorn, K. (2011). "mice: Multivariate imputation by chained equations in R." Journal of statistical software, 45(1), 1-67. DOI: 10.18637/jss.v045.i03.

Examples

```
LogisticReg(test.data, Mvalue="NA",max.score=1)
```

---

micePMM *Predictive mean matching (PMM)*

Description

This function imputes for all missing responses using predictive mean matching. The mice () function with default settings from the mice package (Van Buuren & Groothuis-Oudshoorn, 2011 <doi: 10.18637/jss.v045.i03>) is used to impute for the missing responses.

Usage

```
micePMM(test.data, Mvalue = "NA")
```

Arguments

test.data	Test data set (a data frame or a matrix) containing missing responses. Missing values are coded as NA or other values (e.g., 8, 9).
Mvalue	Missing response indicators in the data (e.g. "NA", "8", "9", etc.). Mvalue="NA" by default.

Value

A data frame with all missing responses replaced by integrated imputed values.

References

Van Buuren, S., & Groothuis-Oudshoorn, K. (2011). "mice: Multivariate imputation by chained equations in R." Journal of statistical software, 45(1), 1-67. DOI: 10.18637/jss.v045.i03.

Examples

```
micePMM(test.data, Mvalue="NA")
```

`PersonMean` *Person Mean Imputation (PM)*

Description

This function imputes for all missing responses of an examinee by his/her mean (i.e., PM) on the available items. Integrated scores for examinees are obtained by rounding their means to the closest possible response value.

Usage

```
PersonMean(test.data, Mvalue = "NA", max.score = 1, round.decimal = 0)
```

Arguments

`test.data`	Test data set (a data frame or a matrix) containing missing responses. Missing values are coded as NA or other values (e.g., 8, 9).
`Mvalue`	Missing response indicators in the data (e.g. "NA", "8", "9", etc.). Mvalue="NA" by default.
`max.score`	The max possible response value in test data. By default max.score=1 (i.e.,binary test data).
`round.decimal`	The number of digits or decimal places for the imputed value. The default value is 0.

Value

A data frame with all missing responses replaced by person means.

References

Sijtsma, K., & Van der Ark, L. A. (2003). "Investigation and treatment of missing item scores in test and questionnaire data." Multivariate Behavioral Research, 38(4), 505-528.DOI: 10.1207/s15327906mbr3804_4.

Examples

```
PersonMean(test.data, Mvalue="NA",max.score=1,round.decimal=0)
```

ResponseFun *Response Function Imputation (RF)*

Description

This function imputes for all missing responses using the response function imputation (Sijtsma and van der Ark, 2003 <doi: 10.1207/s15327906mbr3804_4>).

Usage

```
ResponseFun(test.data, Mvalue = "NA", max.score = 1, round.decimal = 0)
```

Arguments

test.data	Test data set (a data frame or a matrix) containing missing responses. Missing values are coded as NA or other values (e.g., 8, 9).#'
Mvalue	Missing response indicators in the data (e.g. "NA", "8", "9", etc.). Mvalue="NA" by default.
max.score	The max possible response value in test data (the lowest response value should be 0). By default max.score=1 (i.e.,binary test data). max.score = 2 if the response categories are (0, 1, 2), etc.
round.decimal	The number of digits or decimal places for the imputed value. The default value is 0.

Value

A data frame with all missing responses imputed with response function.

References

Sijtsma, K., & Van der Ark, L. A. (2003). "Investigation and treatment of missing item scores in test and questionnaire data." Multivariate Behavioral Research, 38(4), 505-528. DOI: 10.1207/s15327906mbr3804_4.

Examples

```
ResponseFun(test.data, Mvalue="NA",max.score=1,round.decimal=0)
```

`test.data` *Example test data*

Description

This dataset contains binary responses of 775 participants to 20 items.Missing responses are coded as NA.

Usage

```
data("test.data")
```

Format

A data frame with 775 observations on the following 20 items.

`Item_1` a numeric vector

`Item_2` a numeric vector

`Item_3` a numeric vector

`Item_4` a numeric vector

`Item_5` a numeric vector

`Item_6` a numeric vector

`Item_7` a numeric vector

`Item_8` a numeric vector

`Item_9` a numeric vector

`Item_10` a numeric vector

`Item_11` a numeric vector

`Item_12` a numeric vector

`Item_13` a numeric vector

`Item_14` a numeric vector

`Item_15` a numeric vector

`Item_16` a numeric vector

`Item_17` a numeric vector

`Item_18` a numeric vector

`Item_19` a numeric vector

`Item_20` a numeric vector

Details

A test data that contain binary responses of 775 participants to 20 items.Missing responses are coded as NA.

Examples

```
data(test.data)
## maybe str(test.data) ; plot(test.data) ...
```

`TreatIncorrect` *Treat missing responses as incorrect (IN)*

Description

This function replaces all missing responses by zero (see Lord, 1974 <doi: 10.1111/j.1745-3984.1974.tb00996.x>; Mislevy & Wu, 1996 <doi: 10.1002/j.2333-8504.1996.tb01708.x>; Pohl et al., 2014 <doi: 10.1177/0013164413504926>);).

Usage

```
TreatIncorrect(test.data, Mvalue = "NA")
```

Arguments

`test.data`	Test data set (a data frame or a matrix) containing missing responses. Missing values are coded as NA or other values (e.g., 8, 9).
`Mvalue`	Missing response indicators in the data (e.g. "NA", "8", "9", etc.). Mvalue="NA" by default.

Value

A data frame with all missing responses replaced by '0'.

References

Lord, F. M. (1974). " Quick estimates of the relative efficiency of two tests as a function of ability level." Journal of Educational Measurement, 11(4), 247-254. doi: 10.1111/j.1745-3984.1974.tb00996.x.

Mislevy, R. J., & Wu, P. K. (1996). " Missing responses and IRT ability estimation: Omits, choice, time limits, and adaptive testing. " ETS Research Report Series, 1996(2), i-36. doi: 10.1002/j.2333-8504.1996.tb01708.x.

Pohl, S., Gräfe, L., & Rose, N. (2014). "Dealing with omitted and not-reached items in competence tests evaluating approaches accounting for missing responses in item response theory models. " Educational and Psychological Measurement, 74(3), 423-452. doi: 10.1177/0013164413504926.

Examples

```
TreatIncorrect(test.data, Mvalue="NA")
```

Twoway *Two-Way Imputation (TW)*

Description

This function imputes for all missing responses using two-way imputation. Integrated responses are obtained by rounding imputed values to the closest possible response value. If a case showed missingness on all the variables (i.e., empty record), the missing values are replaced by item means first. see Sijtsma and van der Ark (2003) <doi: 10.1207/s15327906mbr3804_4>;

Usage

```
Twoway(test.data, Mvalue = "NA", max.score = 1, round.decimal = 0)
```

Arguments

test.data	Test data set (a data frame or a matrix) containing missing responses. Missing values are coded as NA or other values (e.g., 8, 9).#'
Mvalue	Missing response indicators in the data (e.g. "NA", "8", "9", etc.). Mvalue="NA" by default.
max.score	The max possible response value in test data. By default max.score=1 (i.e.,binary test data).
round.decimal	The number of digits or decimal places for the imputed value. The default value is 0.

Value

A data frame with all missing responses replaced by integrated two-way imputed values.

References

Bernaards, C. A., & Sijtsma, K. (2000). " Influence of imputation and EM methods on factor analysis when item nonresponse in questionnaire data is nonignorable." Multivariate Behavioral Research, 35(3), 321-364.DOI: 10.1207/S15327906MBR3503_03.

Examples

```
Twoway(test.data, Mvalue="NA",max.score=1,round.decimal=0)
```

Index

∗ **datasets**
    test.data, 10

EMimpute, 2

ImputeTestData, 3
ItemMean, 5

Listwise, 6
LogisticReg, 6

micePMM, 7

PersonMean, 8

ResponseFun, 9

test.data, 10
TreatIncorrect, 11
Twoway, 12