

# Package ‘TestDataImputation’

August 11, 2016

**Type** Package

**Title** Missing Item Responses Imputation for Test and Assessment Data

**Version** 1.0

**Date** 2016-08-10

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

**Maintainer** Shenghai Dai <dais@indiana.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, treating as incorrect (IN), person mean imputation (PM), item mean imputation (IM), two-way imputation (TW), logistic regression imputation (LR), and EM imputation.

**Depends** R (>= 3.2.5), mice, Amelia, stats

**NeedsCompilation** no

**LazyData** true

**License** GPL (>= 2)

**RoxygenNote** 5.0.1

**Repository** CRAN

**Date/Publication** 2016-08-11 18:40:22

## R topics documented:

EMimpute . . . . .	2
ImputeTestData . . . . .	2
ItemMean . . . . .	3
Listwise . . . . .	4
LogsticReg . . . . .	5
PersonMean . . . . .	5
test.data . . . . .	6
TreatIncorrect . . . . .	7
Twoway . . . . .	8

**Index** **9**


---

<code>EMimpute</code>	<i>EM Imputation</i>
-----------------------	----------------------

---

**Description**

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

**Usage**

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

**Arguments**

<code>test.data</code>	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).
<code>Mvalue</code>	Missing response indicators in the data (e.g. "NA", "8", "9", etc.). <code>Mvalue="NA"</code> by default.
<code>max.score</code>	The max possible response value in test data. By default <code>max.score=1</code> (i.e., binary test data).

**Value**

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

**Examples**

```
EMimpute(test.data, Mvalue="8",max.score=1)
```

---

<code>ImputeTestData</code>	<i>This main function imputes for missing responses using selected method</i>
-----------------------------	---

---

**Description**

This function imputes for all missing responses using selected miputation methods. 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")
```

**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).#’ @param Mvalue Missing response indicators in the data (e.g. "NA", "8", "9", etc.). Mvalue="NA" by default.
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).
method	Missing response imputation methods. "LW" (by default) represents listwise that deletes all examinees who reported missing responses; "IN" means treating all missing responses as incorrect; "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.; "LR" imputes for all missing responses using logistic regression; "EM" imputes for all missing responses using EM imputation.

**Value**

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

**Examples**

```
ImputeTestData(test.data, Mvalue="8",max.score=1, method ="TW")
```

---

ItemMean	<i>Item Mean (IM) Imputation</i>
----------	----------------------------------

---

**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)
```

**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 item means.

**Examples**

```
ItemMean(test.data, Mvalue="8",max.score=1)
```

---

Listwise

*Listwise Deletion (LW)*

---

**Description**

This function deletes examinees who report missing responses.

**Usage**

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

**Arguments**

<code>test.data</code>	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).#’ @param Mvalue Missing response indicators in the data (e.g. "NA", "8", "9", etc.). Mvalue="NA" by default.
<code>Mvalue</code>	Missing response indicators in the data (e.g. "NA", "8", "9", etc.). Mvalue="NA" by default.

**Value**

A data frame with no missing responses.

**Examples**

```
Listwise(test.data, Mvalue="8")
```

---

LogsticReg

*Logistic Regression (LR) Imputation*

---

### Description

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

### Usage

```
LogsticReg(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.

### Examples

```
LogsticReg(test.data, Mvalue="8",max.score=1)
```

---

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)
```

**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).#’ @param Mvalue Missing response indicators in the data (e.g. "NA", "8", "9", etc.). Mvalue="NA" by default.
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 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.

**Examples**

```
PersonMean(test.data, Mvalue="8",max.score=1)
```

---

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	<i>Treat missing responses as incorrect (IN)</i>
----------------	--

---

### Description

This function replaces all missing responses by zero.

### 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).# @param Mvalue Missing response indicators in the data (e.g. "NA", "8", "9", etc.). Mvalue="NA" by default.
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 imputed.

**Examples**

```
TreatIncorrect(test.data, Mvalue="8")
```

---

Tway

*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.

**Usage**

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

**Arguments**

<code>test.data</code>	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).# @param Mvalue Missing response indicators in the data (e.g. "NA", "8", "9", etc.). Mvalue="NA" by default.
<code>Mvalue</code>	Missing response indicators in the data (e.g. "NA", "8", "9", etc.). Mvalue="NA" by default.
<code>max.score</code>	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 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.

**Examples**

```
Tway(test.data, Mvalue="8",max.score=1)
```



# Index

\*Topic **datasets**

test.data, [6](#)

EMimpute, [2](#)

ImputeTestData, [2](#)

ItemMean, [3](#)

Listwise, [4](#)

LogsticReg, [5](#)

PersonMean, [5](#)

test.data, [6](#)

TreatIncorrect, [7](#)

TwoWay, [8](#)