Item Response Theory - Final Essay

Marius Keute

August 29, 2022

Contents

submitted to:

Dr. Stefano Noventa

University of Tübingen

submitted by:

Marius Keute (QDS, 5991873)

Statutory Declaration: I hereby declare that I composed the present paper independently and that I have used no other resources than those indicated. The text passages which are taken from other works in wording or meaning have been identified as such. I also declare that this work has not been partly or completely used in another examination.

1 Introduction

2 Data preparation

The dataset consisted of 3376 observations, the variables being the ten items of the SCS, the sum score, gender and age. From the age variable, three cases where the reported age was 100 years or higher were set to missing values. The remaining cases had a mean age of 30.9 years (median 28 years, range [14, 85]). From the gender variable, 13 values were missing and 15 cases where the reported gender was “3” were set to missing values. Of the remaining cases, 2295 (68.5%) reported gender “1” and 1053 (31.4%) reported gender “2”. The SCS data contained at least one missing value for 133 cases.

The pattern of missing SCS items is shown in Figure 1. It can be seen that item Q9 was missing most often, though not by a large margin (Q9: 27 missing values, Q5: 13 missing values). It can be seen that the majority of cases with missing values (118 cases / 88.7%) had only a single item missing, while there were no prominent patterns of items that tended to be jointly missing. Eight cases where more than two SCS items were missing were excluded from all further analyses. For the remaining 3368 cases, the probability of missing values at each SCS variable was modeled as a function of the values in all other SCS variables using a logistic regression model: $P(M_{i,q} = 1|X_{i,q'}) = \sigma(X_{i,q'}\hat{\beta})$, where $M_{i,q}$ is 1 if the i^{th} person has a missing value at item $q \in \{Q1, Q2, ...Q10\}$, $X_{i,q'}$ denotes the item values of all other items, σ is the logistic function, and $\hat{\beta}$ are the estimated regression weights (Guan and Yusoff (2011)). Note that each variable’s pattern of missing values could only be predicted based on the observations without missing values in any other variable, since those cases were excluded by the logistic model by default of the implementation. Since the majority of cases had either no or only one variable missing, however, this should not bias the overall picture very much.

```
## Loading required package: ggplot2

## Loading required package: reshape2

## Loading required package: readxl

## Loading required package: VIM

## Loading required package: colorspace

## Loading required package: grid

## VIM is ready to use.

## Suggestions and bug-reports can be submitted at: https://github.com/statistikat/VIM/issues

##
## Attaching package: 'VIM'

## The following object is masked from 'package:datasets':
##
##     sleep

## Loading required package: mice

##
## Attaching package: 'mice'

## The following object is masked from 'package:stats':
##
##     filter

## The following objects are masked from 'package:base':
##
##     cbind, rbind

## Loading required package: dplyr
```

```
##
## Attaching package: 'dplyr'
```

```
## The following objects are masked from 'package:stats':
##
##     filter, lag
```

```
## The following objects are masked from 'package:base':
##
##     intersect, setdiff, setequal, union
```

```
## Loading required package: tidyr
```

```
##
## Attaching package: 'tidyr'
```

```
## The following object is masked from 'package:reshape2':
##
##     smiths
```

```
## Loading required package: psych
```

```
##
## Attaching package: 'psych'
```

```
## The following objects are masked from 'package:ggplot2':
##
##     %+%, alpha
```

```
## Loading required package: ggcorrplot
```

```
##
##    0    1    2    3
##   13 2295 1053   15
##  [1]  41  50  23  42  36  29  24  35  26  43  21  39  37  64  28  46  34  31  47
## [20]  22  61  16  40  33  30  56  49  51  18  20  45  32  15  27  25  59  58  19
## [39]  14  38  48  44  55 100  65  17  77  57  60  52  53  62  71  78  54  63  67
## [58]  68  72 999  85  69  70  66  84 123  73
```

```
## Warning in plot.aggr(res, ...): not enough vertical space to display frequencies
## (too many combinations)
```

```
##
##  Variables sorted by number of missings:
##  Variable Count
##        Q9    27
##        Q3    22
##        Q4    22
##        Q7    22
##        Q8    21
##        Q6    20
##       Q10    17
##        Q2    16
##        Q1    15
##        Q5    13
##
##  iter imp variable
##   1   1  Q1  Q2  Q3  Q4  Q5  Q6  Q7  Q8  Q9  Q10  gender  age
##   1   2  Q1  Q2  Q3  Q4  Q5  Q6  Q7  Q8  Q9  Q10  gender  age
##   1   3  Q1  Q2  Q3  Q4  Q5  Q6  Q7  Q8  Q9  Q10  gender  age
##   1   4  Q1  Q2  Q3  Q4  Q5  Q6  Q7  Q8  Q9  Q10  gender  age
##   1   5  Q1  Q2  Q3  Q4  Q5  Q6  Q7  Q8  Q9  Q10  gender  age
##   2   1  Q1  Q2  Q3  Q4  Q5  Q6  Q7  Q8  Q9  Q10  gender  age
##   2   2  Q1  Q2  Q3  Q4  Q5  Q6  Q7  Q8  Q9  Q10  gender  age
##   2   3  Q1  Q2  Q3  Q4  Q5  Q6  Q7  Q8  Q9  Q10  gender  age
##   2   4  Q1  Q2  Q3  Q4  Q5  Q6  Q7  Q8  Q9  Q10  gender  age
##   2   5  Q1  Q2  Q3  Q4  Q5  Q6  Q7  Q8  Q9  Q10  gender  age
##   3   1  Q1  Q2  Q3  Q4  Q5  Q6  Q7  Q8  Q9  Q10  gender  age
##   3   2  Q1  Q2  Q3  Q4  Q5  Q6  Q7  Q8  Q9  Q10  gender  age
##   3   3  Q1  Q2  Q3  Q4  Q5  Q6  Q7  Q8  Q9  Q10  gender  age
##   3   4  Q1  Q2  Q3  Q4  Q5  Q6  Q7  Q8  Q9  Q10  gender  age
##   3   5  Q1  Q2  Q3  Q4  Q5  Q6  Q7  Q8  Q9  Q10  gender  age
##   4   1  Q1  Q2  Q3  Q4  Q5  Q6  Q7  Q8  Q9  Q10  gender  age
##   4   2  Q1  Q2  Q3  Q4  Q5  Q6  Q7  Q8  Q9  Q10  gender  age
##   4   3  Q1  Q2  Q3  Q4  Q5  Q6  Q7  Q8  Q9  Q10  gender  age
##   4   4  Q1  Q2  Q3  Q4  Q5  Q6  Q7  Q8  Q9  Q10  gender  age
##   4   5  Q1  Q2  Q3  Q4  Q5  Q6  Q7  Q8  Q9  Q10  gender  age
##   5   1  Q1  Q2  Q3  Q4  Q5  Q6  Q7  Q8  Q9  Q10  gender  age
##   5   2  Q1  Q2  Q3  Q4  Q5  Q6  Q7  Q8  Q9  Q10  gender  age
##   5   3  Q1  Q2  Q3  Q4  Q5  Q6  Q7  Q8  Q9  Q10  gender  age
##   5   4  Q1  Q2  Q3  Q4  Q5  Q6  Q7  Q8  Q9  Q10  gender  age
##   5   5  Q1  Q2  Q3  Q4  Q5  Q6  Q7  Q8  Q9  Q10  gender  age
```

```
## Warning in biserialc(x[, j], y[, i], j, i): For x = 1 y = 1 x seems to be
## dichotomous, not continuous
```

```
## Warning in biserialc(x[, j], y[, i], j, i): For x = 2 y = 2 x seems to be
## dichotomous, not continuous
```

```
## Warning in biserialc(x[, j], y[, i], j, i): For x = 3 y = 3 x seems to be
## dichotomous, not continuous
```

```
## Warning in biserialc(x[, j], y[, i], j, i): For x = 4 y = 4 x seems to be
## dichotomous, not continuous
```

```
## Warning in biserialc(x[, j], y[, i], j, i): For x = 5 y = 5 x seems to be
## dichotomous, not continuous
```

```
## Warning in biserialc(x[, j], y[, i], j, i): For x = 6 y = 6 x seems to be
## dichotomous, not continuous
```

```
## Warning in biserialc(x[, j], y[, i], j, i): For x = 7 y = 7 x seems to be
## dichotomous, not continuous
```

```
## Warning in biserialc(x[, j], y[, i], j, i): For x = 8 y = 8 x seems to be
## dichotomous, not continuous
```

```
## Warning in biserialc(x[, j], y[, i], j, i): For x = 9 y = 9 x seems to be
## dichotomous, not continuous
```

```
## Warning in biserialc(x[, j], y[, i], j, i): For x = 10 y = 10 x seems to be
## dichotomous, not continuous
```

```
## Warning in biserialc(x[, j], y[, i], j, i): For x = 1 y = 11 y is not
## dichotomous
```

```
## Warning in biserialc(x[, j], y[, i], j, i): For x = 2 y = 11 y is not
## dichotomous
```

```
## Warning in biserialc(x[, j], y[, i], j, i): For x = 3 y = 11 y is not
## dichotomous

## Warning in biserialc(x[, j], y[, i], j, i): For x = 4 y = 11 y is not
## dichotomous

## Warning in biserialc(x[, j], y[, i], j, i): For x = 5 y = 11 y is not
## dichotomous

## Warning in biserialc(x[, j], y[, i], j, i): For x = 6 y = 11 y is not
## dichotomous

## Warning in biserialc(x[, j], y[, i], j, i): For x = 7 y = 11 y is not
## dichotomous

## Warning in biserialc(x[, j], y[, i], j, i): For x = 8 y = 11 y is not
## dichotomous

## Warning in biserialc(x[, j], y[, i], j, i): For x = 9 y = 11 y is not
## dichotomous

## Warning in biserialc(x[, j], y[, i], j, i): For x = 10 y = 11 y is not
## dichotomous

## Warning in biserialc(x[, j], y[, i], j, i): For x = 11 y = 11 y is not
## dichotomous

## Warning in biserialc(x[, j], y[, i], j, i): For x = 12 y = 11 y is not
## dichotomous

## Warning in biserialc(x[, j], y[, i], j, i): For x = 13 y = 11 y is not
## dichotomous

## Warning in biserialc(x[, j], y[, i], j, i): For x = 12 y = 12 x seems to be
## dichotomous, not continuous

## Warning in biserialc(x[, j], y[, i], j, i): For x = 1 y = 13 y is not
## dichotomous

## Warning in biserialc(x[, j], y[, i], j, i): For x = 2 y = 13 y is not
## dichotomous

## Warning in biserialc(x[, j], y[, i], j, i): For x = 3 y = 13 y is not
## dichotomous

## Warning in biserialc(x[, j], y[, i], j, i): For x = 4 y = 13 y is not
## dichotomous

## Warning in biserialc(x[, j], y[, i], j, i): For x = 5 y = 13 y is not
## dichotomous

## Warning in biserialc(x[, j], y[, i], j, i): For x = 6 y = 13 y is not
## dichotomous

## Warning in biserialc(x[, j], y[, i], j, i): For x = 7 y = 13 y is not
## dichotomous

## Warning in biserialc(x[, j], y[, i], j, i): For x = 8 y = 13 y is not
## dichotomous

## Warning in biserialc(x[, j], y[, i], j, i): For x = 9 y = 13 y is not
## dichotomous

## Warning in biserialc(x[, j], y[, i], j, i): For x = 10 y = 13 y is not
## dichotomous
## Warning in biserialc(x[, j], y[, i], j, i): For x = 11 y = 13 y is not
## dichotomous

## Warning in biserialc(x[, j], y[, i], j, i): For x = 12 y = 13 y is not
## dichotomous

## Warning in biserialc(x[, j], y[, i], j, i): For x = 13 y = 13 y is not
## dichotomous

## Warning in base::rbind(...): number of columns of result is not a multiple of
## vector length (arg 2)
```

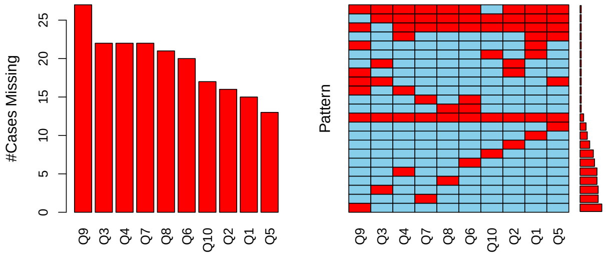

Figure 1: Pattern of missing SCS values.

For dichotomization of the item data, I considered two options, namely, thresholding each of the 10 items at its own median, to ensure an even distribution of observations into both categories for each item, or finding a common threshold for all items. Since the items have only four levels each, a median split would not necessarily lead to a very balanced dichotomization. Furthermore, the item levels are designed to have the same meaning across all items, therefore I decided to dichotomize at a common threshold of 2, i.e., the dichotomous items $D_q \in \{D_1, D_2, ..., D_{10}\}$ were defined such that

$$D_{i,q} = \begin{cases} 0 \text{ if } Q_{i,q} \in \{1, 2\}, \\ 1 \text{ if } Q_{i,q} \in \{3, 4\}, \end{cases}$$

The distribution of the dichotomized items is shown in 2 Since most variables' median was 2, this was not much different from an item-wise median threshold (see 1). Subsequently, I calculated biserial correlations between all pairs of dichotomized items. Moreover, I calculated item discrimination, i.e., each items ability to discriminate between high- and low-scoring individuals, using the adjusted item-total correlation method (Reynolds and Livingston (2021)), i.e., by calculating biserial correlation coefficients between each (dichotomized) item's scores and the sum of all other (dichotomized) items.

#TODO make tables smaller (too wide) #TODO make captions and table referencing work

3 Results

Descriptive characteristics of the 10 SCS items are shown in 1, the proportions of 'correct' responses, i.e., responses greater than 2, are shown in 2.

Item intercorrelations are shown in Figure 2 #TODO discuss

Table 1: Descriptive item statistics (mean, median and range before dichotomization)

stat	Q1	Q10	Q2	Q3	Q4	Q5	Q6	Q7	Q8	Q9
max	4.0	4.0	4.0	4.0	4.0	4.0	4.0	4.0	4.0	4.0
mean	2.3	2.5	2.2	2.2	1.9	2.2	3.1	2.2	2.3	2.5
median	2.0	3.0	2.0	2.0	2.0	2.0	3.0	2.0	2.0	2.0
min	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0	1.0

Table 2: Distribution and discrimination of dichotomized items

percent in category 1	40.6	37.8	37.3	27	38.2	71.9	37.8	41.4	49.3	50.7	432	131.4
number of cases in category 1	1366	1273	1255	911	1286	2423	1272	1393	1662	1709	14550	4426
discrimination	0.45	0.45	0.44	0.34	0.29	0.26	0.42	0.37	0.31	0.35	0.45	0.45

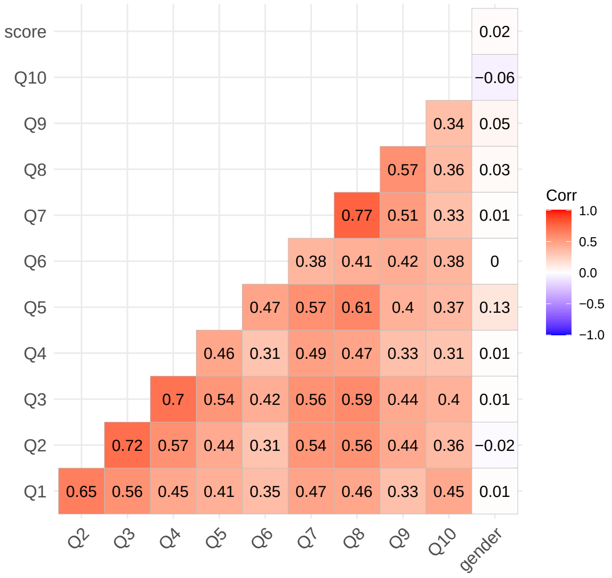

Figure 2: Pattern of missing SCS values.

4 References

Guan, Ng Chong, and Muhamad Saiful Bahri Yusoff. 2011. "Missing Values in Data Analysis: Ignore or Impute?" *Education in Medicine Journal* 3 (1).

Reynolds, Cecil R, and RA Livingston. 2021. *Mastering Modern Psychological Testing*. Springer.