Measurement, Scaling, & Dimensional Analysis
Prof. Adam Enders
ICPSR, 2019

Assignment 2:
Cumulative Scaling

We begin by submitting all 18 items to the automated item selection procedure (results of which are depicted below), which reveals that all items are cumulatively scalable to the weak (necessary, but not sufficient) criterion of an H coefficient of 0.30 or greater.

```
> aisp(toldata)
            0.3
atheists      1
atheistt      1
atheistb      1
racists       1
racistt       1
racistb       1
communists    1
communistt    1
communistb    1
militarists   1
militaristt   1
militaristb   1
homos         1
homot         1
homob         1
muslims       1
muslimt       1
muslimb       1
```

In fact, the scale H coefficient is 0.54, which is suggests an "excellent" fit according to rules of thumb.

```
> coefH(toldata)

[Lengthy output omitted]

$H
 Scale H se
   0.540 (0.011)
```

Next, we move on to checking the assumptions of the Monotone Homogeneity model, which can help more firmly establish the scalability of subjects (or, row objects) along the latent

dimension. This can be accomplished simply by checking whether the estimated item response functions for the individual items (column objects) are monotonically nondecreasing.

```
> summary(check.monotonicity(toldata))
            ItemH #ac #vi #vi/#ac maxvi sum sum/#ac zmax #zsig crit
atheists     0.54  21   0       0     0   0       0    0     0    0
atheistt     0.46  21   0       0     0   0       0    0     0    0
atheistb     0.53  21   0       0     0   0       0    0     0    0
racists      0.49  21   0       0     0   0       0    0     0    0
racistt      0.49  21   0       0     0   0       0    0     0    0
racistb      0.48  21   0       0     0   0       0    0     0    0
communists   0.56  21   0       0     0   0       0    0     0    0
communistt   0.45  21   0       0     0   0       0    0     0    0
communistb   0.57  21   0       0     0   0       0    0     0    0
militarists  0.52  21   0       0     0   0       0    0     0    0
militaristt  0.53  21   0       0     0   0       0    0     0    0
militaristb  0.54  21   0       0     0   0       0    0     0    0
homos        0.61  10   0       0     0   0       0    0     0    0
homot        0.53  21   0       0     0   0       0    0     0    0
homob        0.55  21   0       0     0   0       0    0     0    0
muslims      0.64  21   0       0     0   0       0    0     0    0
muslimt      0.72  21   0       0     0   0       0    0     0    0
muslimb      0.62  21   0       0     0   0       0    0     0    0
```

This set of statistical tests reveal no violations of monotonicity, whatsoever. We could also visually inspect estimated item response functions to diagnose severity in the case of statistically significant model violations using `plot(check.monotonicity(toldata))`.

With a high H coefficient and no violations of the Monotone Homogeneity model, we can be very comfortable asserting the cumulative scalability of subjects (row objects) along the latent dimension. Next, we can consider whether a scale of these items (the full set, in this case) also conform to the Double Monotonicity model, which would allow us to order the items (column objects) along the latent dimension, as well. As a first step in the examination of DM model violations, we could check for violations of item response function non-intersection using restscores.

```
> summary(check.restscore(toldata))
            ItemH #ac #vi #vi/#ac maxvi  sum sum/#ac zmax #zsig crit
atheists     0.54 100   0    0.00  0.00 0.00  0.0000 0.00     0    0
atheistt     0.46 100   7    0.07  0.10 0.45  0.0045 2.88     3   57
atheistb     0.53  98   0    0.00  0.00 0.00  0.0000 0.00     0    0
racists      0.49 101   4    0.04  0.09 0.26  0.0026 2.72     2   42
racistt      0.49  99   9    0.09  0.11 0.68  0.0068 4.45     6   81
racistb      0.48 101   5    0.05  0.11 0.34  0.0033 3.42     3   55
communists   0.56 102   8    0.08  0.14 0.67  0.0066 4.65     6   80
```

```
communistt   0.45 101   6    0.06  0.14 0.56  0.0055 4.65     4   75
communistb   0.57 102   9    0.09  0.11 0.57  0.0056 3.61     4   64
militarists  0.52 102   4    0.04  0.06 0.17  0.0017 1.58     0   14
militaristt  0.53 101   4    0.04  0.09 0.22  0.0022 2.72     1   34
militaristb  0.54 102   2    0.02  0.06 0.10  0.0010 1.60     0    9
homos        0.61 102   0    0.00  0.00 0.00  0.0000 0.00     0    0
homot        0.53 102   0    0.00  0.00 0.00  0.0000 0.00     0    0
homob        0.55 100   0    0.00  0.00 0.00  0.0000 0.00     0    0
muslims      0.64  93   3    0.03  0.11 0.24  0.0026 4.21     2   44
muslimt      0.72  98   1    0.01  0.04 0.04  0.0004 2.09     1    8
muslimb      0.62 102   4    0.04  0.11 0.34  0.0034 4.45     2   48
```

Several items show more than violation – that is, there are several points in a plot of one pair of IRFs, and/or across several plots of various pairs of IRFs, where IRFs significantly intersect. Before taking any action, we might want to look at the results of the "p matrix" method of testing IRF nonintersection to look for similarities and differences.

```
> summary(check.pmatrix(toldata))
             ItemH #ac #vi #vi/#ac maxvi  sum sum/#ac zmax #zsig crit
atheists      0.54 272   2    0.01  0.03 0.07  0.0002 5.60     2   39
atheistt      0.46 272   7    0.03  0.04 0.27  0.0010 6.55     7   69
atheistb      0.53 272   1    0.00  0.04 0.04  0.0002 6.15     1   37
racists       0.49 272   8    0.03  0.04 0.27  0.0010 5.82     8   66
racistt       0.49 272  13    0.05  0.08 0.57  0.0021 8.74    13  101
racistb       0.48 272   7    0.03  0.04 0.26  0.0010 5.82     7   64
communists    0.56 272   9    0.03  0.05 0.33  0.0012 8.33     9   79
communistt    0.45 272   8    0.03  0.05 0.30  0.0011 8.33     8   82
communistb    0.57 272   9    0.03  0.05 0.34  0.0013 5.88     9   67
militarists   0.52 272  11    0.04  0.05 0.41  0.0015 6.15    11   77
militaristt   0.53 272   6    0.02  0.04 0.21  0.0008 5.69     6   57
militaristb   0.54 272   4    0.01  0.03 0.13  0.0005 5.23     4   46
homos         0.61 272   0    0.00  0.00 0.00  0.0000 0.00     0    0
homot         0.53 272   0    0.00  0.00 0.00  0.0000 0.00     0    0
homob         0.55 272   1    0.00  0.03 0.03  0.0001 5.36     1   31
muslims       0.64 272   7    0.03  0.08 0.32  0.0012 8.35     7   74
muslimt       0.72 272   0    0.00  0.00 0.00  0.0000 0.00     0    0
muslimb       0.62 272   5    0.02  0.05 0.22  0.0008 8.74     5   66
```

The p matrix method also reveals several statistically significant violations of the DM model for several items.

This is the point at which the distinction between art and science blurs. The best starting point for removing items (assuming that you are interested in scaling items along the latent continuum) is to find those which are involved in the most statistically significant violations across methods. Once an item has been removed, one should repeat the analysis from the bottom up: re-examine the reduced-scale H coefficient, check the MH model, and

re-run DM model diagnostics. Stop and proceed with substantive interpretation and subsequent statistical analyses (e.g., a regression model) once you're satisfied that no problematic violations of either the MH or DM model exist.

In this particular case, the `racistt` item is most problematic according to both methods of checking IRF non-intersection. In the end, I removed all of the items that had to do with teaching (variable names ending in a `t`), as well as the `militarists` item. The resultant scale has an H coefficient of 0.602 and no violations of the MH model.

```
> toldata2 <- cbind(toldata[,c(1,3,4,6,7,9,12,13,15,16,18)])
> head(toldata2)

[Lengthy output omitted]

>
> coefH(toldata2)

[Lengthy output omitted]

$H
 Scale H se
   0.602 (0.012)

>
> summary(check.monotonicity(toldata2))
            ItemH #ac #vi #vi/#ac maxvi sum sum/#ac zmax #zsig crit
atheists     0.56  15   0       0     0   0       0    0     0    0
atheistb     0.58  15   0       0     0   0       0    0     0    0
racists      0.53  15   0       0     0   0       0    0     0    0
racistb      0.56  15   0       0     0   0       0    0     0    0
communists   0.59  15   0       0     0   0       0    0     0    0
communistb   0.63  15   0       0     0   0       0    0     0    0
militaristb  0.57  15   0       0     0   0       0    0     0    0
homos        0.62  15   0       0     0   0       0    0     0    0
homob        0.59  15   0       0     0   0       0    0     0    0
muslims      0.72  15   0       0     0   0       0    0     0    0
muslimb      0.71  15   0       0     0   0       0    0     0    0
```

The DM model fit statistics, while technically imperfect, reveal very few model violations, none of which are particularly obvious upon visual inspection of estimated IRF pairs using `plot(check.restscore(toldata2))`. Furthermore only one "crit" value is above the recommend 40 "troublesome" cutoff, and only by 4 (`racists` is at 44, according to the p matrix method).

```
> summary(check.restscore(toldata2))
            ItemH #ac #vi #vi/#ac maxvi  sum sum/#ac zmax #zsig crit
```

```
atheists     0.56  49   0    0.00  0.00 0.00  0.0000 0.00     0    0
atheistb     0.58  49   0    0.00  0.00 0.00  0.0000 0.00     0    0
racists      0.53  49   3    0.06  0.06 0.19  0.0039 1.97     2   39
racistb      0.56  49   3    0.06  0.07 0.19  0.0039 1.93     2   37
communists   0.59  49   3    0.06  0.07 0.19  0.0039 1.97     2   36
communistb   0.63  49   4    0.08  0.06 0.22  0.0044 1.85     2   36
militaristb  0.57  49   1    0.02  0.04 0.04  0.0007 1.19     0    1
homos        0.62  50   0    0.00  0.00 0.00  0.0000 0.00     0    0
homob        0.59  49   0    0.00  0.00 0.00  0.0000 0.00     0    0
muslims      0.72  42   1    0.02  0.04 0.04  0.0010 1.55     0   -3
muslimb      0.71  50   1    0.02  0.04 0.04  0.0008 1.55     0   -3
> summary(check.pmatrix(toldata2))
            ItemH #ac #vi #vi/#ac maxvi  sum sum/#ac zmax #zsig crit
atheists     0.56  90   1    0.01  0.03 0.03   4e-04 5.36     1   32
atheistb     0.58  90   0    0.00  0.00 0.00   0e+00 0.00     0    0
racists      0.53  90   2    0.02  0.03 0.06   7e-04 5.82     2   44
racistb      0.56  90   1    0.01  0.03 0.03   4e-04 5.82     1   34
communists   0.59  90   1    0.01  0.03 0.03   3e-04 4.64     1   26
communistb   0.63  90   1    0.01  0.03 0.03   4e-04 5.12     1   27
militaristb  0.57  90   1    0.01  0.03 0.03   4e-04 5.12     1   30
homos        0.62  90   0    0.00  0.00 0.00   0e+00 0.00     0    0
homob        0.59  90   1    0.01  0.03 0.03   4e-04 5.36     1   30
muslims      0.72  90   0    0.00  0.00 0.00   0e+00 0.00     0    0
muslimb      0.71  90   0    0.00  0.00 0.00   0e+00 0.00     0    0
```

This scale is also highly reliable, by any measure of reliability.

```
> check.reliability(toldata2)
$MS
[1] 0.9043729

$alpha
[1] 0.8865142

$lambda.2
[1] 0.8906687
```

I end my analysis being able to comfortably array 11 of the original 18 items along a latent (in)tolerance dimension in a fixed order, along with respondents. The order of the items appears below. Remember that these items are coded so that a "1" denotes an intolerant response. In 2016, Americans are much more tolerant toward homosexuals than muslims, regardless of the civil liberty in question. One would need to be very high on the scale (10, 11) to be intolerant of homosexuals, which would also imply an intolerant response to all other groups. However, a scale score of 1 or 2 suggests intolerance toward muslims, specifically, since the items do have a fixed order along the latent dimension. A scale score of 2 means that the given individual "dominated" both the `muslims` and `muslimb` items.

```
> rank(1-colMeans(toldata2))
   atheists    atheistb     racists     racistb  communists  communistb
          9           8           3           4           5           6
militaristb       homos       homob     muslims     muslimb
          7          11          10           1           2
```

The distribution of the (in)tolerance scale reveals that less than 30% of Americans are completely tolerant of all groups. The mean of the scale is 3.37, suggesting that, on average, people are intolerant of muslims and racists speaking in their community, or of their local library containing texts authored by muslims.

```
> mokkenscale <- rowSums(toldata2)
>
> summary(mokkenscale)
   Min. 1st Qu.  Median    Mean 3rd Qu.    Max.
   0.00    0.00    2.00    3.37    6.00   11.00
>
> histogram(~mokkenscale,
+           aspect = 1,
+           xlab = "Mokken Scale Scores",
+           )
```