The software challenge in the Rasch model

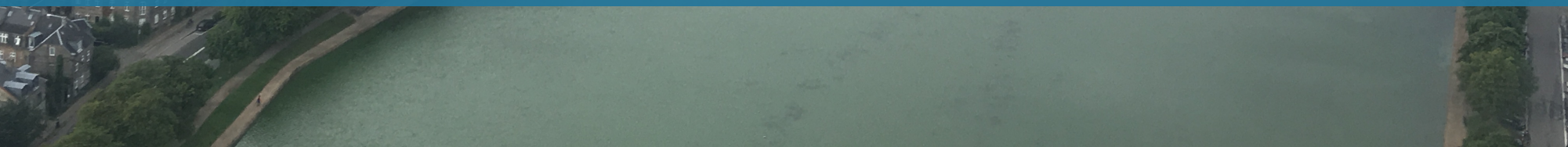
Karl Bang Christensen • Biostatistics, Univ. of Copenhagen
biostat.ku.dk/~kach/ • kach@sund.ku.dk


THE RASCH MODEL

- $\boldsymbol{X}_v = (X_{1v}, \ldots, X_{kv}) \in \{0,1\}^k$ item responses
  - $\boldsymbol{\eta} = (\eta_1, \ldots, \eta_k)$ item locations
  - $\theta_v \in \mathbb{R}$ latent variable, $\boldsymbol{\theta} = (\theta_1, \ldots, \theta_N)$

- $\theta \mapsto P(X_{iv} = 1 | \theta_v = \theta)$ functional relationship

- Test assumptions - get 'measurement' of θ



Rasch model part of general class of models[1]

$$P(\boldsymbol{X}_v = \boldsymbol{x} | \theta_v = \theta) = \frac{\exp(\boldsymbol{x}^T(\boldsymbol{B}\theta + \boldsymbol{A}\eta))}{\sum_z \exp(\boldsymbol{z}^T(\boldsymbol{B}\theta + \boldsymbol{A}\eta))}$$

choosing $\boldsymbol{A} = I$ and $\boldsymbol{B} = [1 \cdots 1]^T$ yields

$$R := \boldsymbol{x}^T\boldsymbol{B} = \sum_{i=1}^{k} x_i$$

seen to be sufficient for θ

[1] Adams, Wu (2007)
https://doi.org/10.1007/978-0-387-49839-3_4

Rasch model part of general class of models

$$P(\boldsymbol{X}_v = \boldsymbol{x}|\theta_v = \theta) = \frac{\exp(\boldsymbol{x}^T(\boldsymbol{B}\theta + \boldsymbol{A}\eta))}{\sum_z \exp(\boldsymbol{z}^T(\boldsymbol{B}\theta + \boldsymbol{A}\eta))}$$

choosing $\boldsymbol{A} = I$ and $\boldsymbol{B} = [1 \cdots 1]^T$ further yields

$$P(\boldsymbol{X} = \boldsymbol{x}|\theta) = \prod_i P(X_i = x_i|\theta)$$

local independence[2]

[2] Lazarsfeld, Henry (1968). Latent structure analysis. Boston, MA: Houghton Mifflin

Conditional probability independent of θ

$$P(\boldsymbol{X}_v = \boldsymbol{x} | R_v, \theta_v = \theta) = \frac{\exp(\boldsymbol{x}^T \boldsymbol{A} \boldsymbol{\eta})}{\sum_{z:\boldsymbol{z}^T \boldsymbol{B} = R} \exp(\boldsymbol{z}^T \boldsymbol{A} \boldsymbol{\eta})}$$

Conditional inference yields unbiased estimates $\hat{\eta}$ [3]
Historically a computational challenge. Not so anymore [4,5]

[3] Andersen (1970). JRSS B, 32, 283-301.
http://www.jstor.org/stable/2984535
[4] Mair, Hatzinger (2007). Journal of Statistical Software, 20(9), 9.
https://doi.org/10.18637/jss.v020.i09
[5] Christensen (2013). ISRN Computational Mathematics,
http://doi.org/10.1155/2013/617475


AMTS

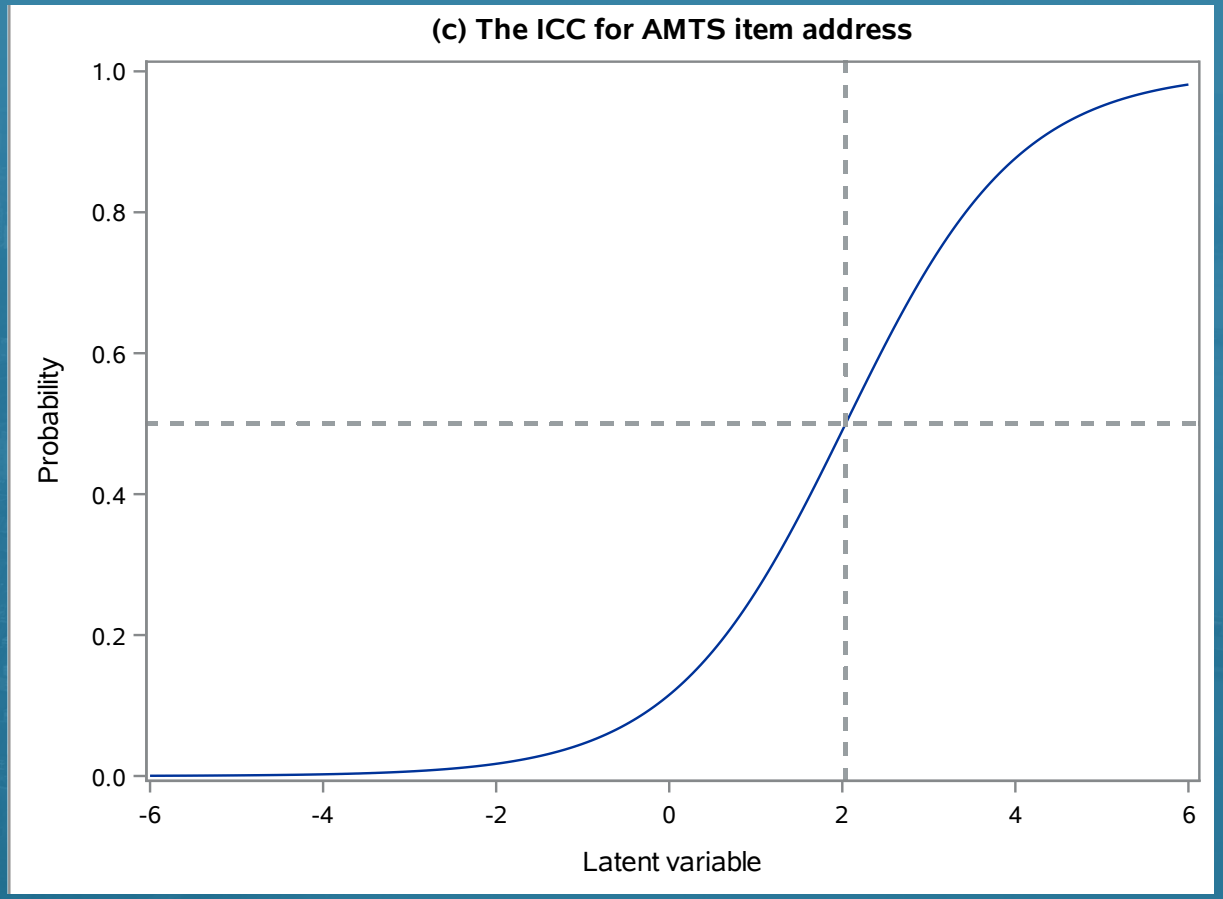
ABBREVIATED MENTAL TEST SCORE

Age // Time (to nearest hour) // Address
Name of Hospital (or area of town) // Current Year
Date of birth of patient // Month // Date of first world war
Name of Monarch // Count backwards

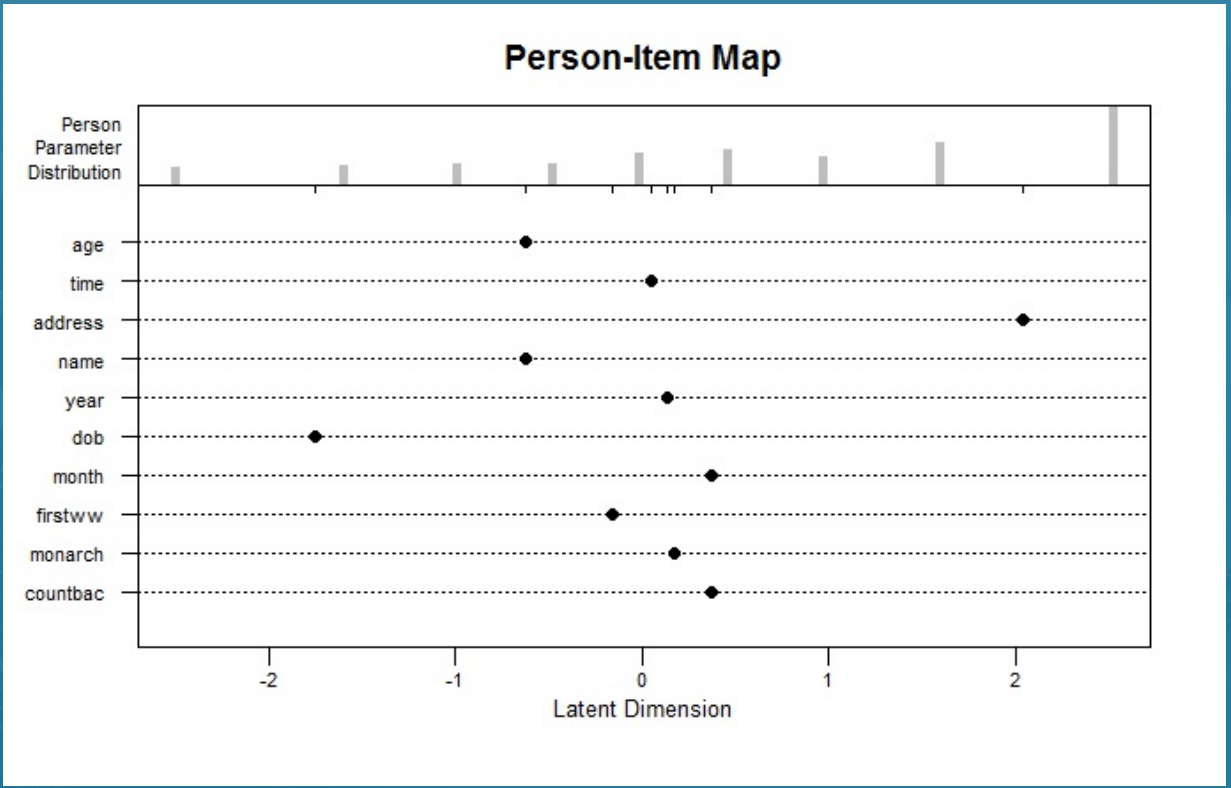
6/10 or less abnormal in elderly.[6]

[6] Hodkinson HM (1972). Age Ageing. 1(4):233-8
https://doi.org/10.1093/ageing/afs148

SAS


(c) The ICC for AMTS item address
Probability
1.0
0.8
0.6
0.4
0.2
0.0
-6
-4
-2
0
2
4
6
Latent variable

R


Person-Item Map
Person
Parameter
Distribution
age
time
address
name
year
dob
month
firstww
monarch
countbac
-2
-1
0
1
2
Latent Dimension


RASCH SOFTWARE

- RUMM[7] (pairwise) conditional estimation of $\hat{\eta}$, ML estimation of $\hat{\theta}$
- WINSTEPS[8] joint likelihood estimation of $(\hat{\eta}, \hat{\theta})$

Model fit evaluated by plugging in $\hat{\theta}_v$

[7] Andrich, Sheridan, Luo (2010). RUMM2030 [Computer software and manual]. Perth, Australia.

[8] Linacre (2017). Winsteps® Rasch measurement computer program. Beaverton, Oregon: Winsteps.com


INDIVIDUAL ITEM FIT

Standard evaluation of individual item fit using

$$Z_{iv} = \frac{X_{iv} - E(X_{iv}|\theta_v = \hat{\theta}_v)}{\sqrt{V(X_{iv}|\theta_v = \hat{\theta}_v)}}$$

that rely on (biased) estimates $\hat{\theta}_v$.[9] Fit statistics like

$$OUTFIT_i = \frac{1}{N}\sum_{v=1}^{N} Z_{iv}^2$$

with no established null distribution (early Rasch literature claims of χ^2 distribution.).

[9] Warm (1989). Psychometrika, 54, 427–450.

It has been suggested that the Wilson-Hilferty cube-root transformation

$$t_i = (OUTFIT_i^{1/3} - 1)\frac{3}{V(OUTFIT_i)} + \frac{V(OUTFIT_i)}{3}$$

has an approximate t distribution and that

$$FitResid_i = \frac{f(\log(N \cdot OUTFIT_i) - \log(f))}{\sqrt{V(N \cdot OUTFIT_i)}}$$

with $f = (Nk - N - k + 1)/k$, has a symmetrical distribution with mean zero and variance one.

Item χ^2 fit statistic in RUMM groups respondents in G 'class intervals' based on the estimated person locations $\hat{\theta}_v$. Computed as

$$\chi^2(X_i) = \sum_g \frac{\left(\sum_{v \in V_g} X_{vi} - \sum_{v \in V_g} E(X_{vi})\right)^2}{\sum_{v \in V_g} V(X_{vi})}$$

where V_g denotes the set of respondents in class interval g

still relies on $\hat{\theta}_1, \hat{\theta}_2, \ldots$

In SAS:

- Re-write $\sum_{v=1}^{N} Z_{iv}^2 = \sum_{r=0}^{k} \sum_{v:R_v=r} Z_{iv}^2$
- replace $X_{iv}|(\theta_v = \hat{\theta}_v)$ by $X_{iv}|(R_v = r)$:

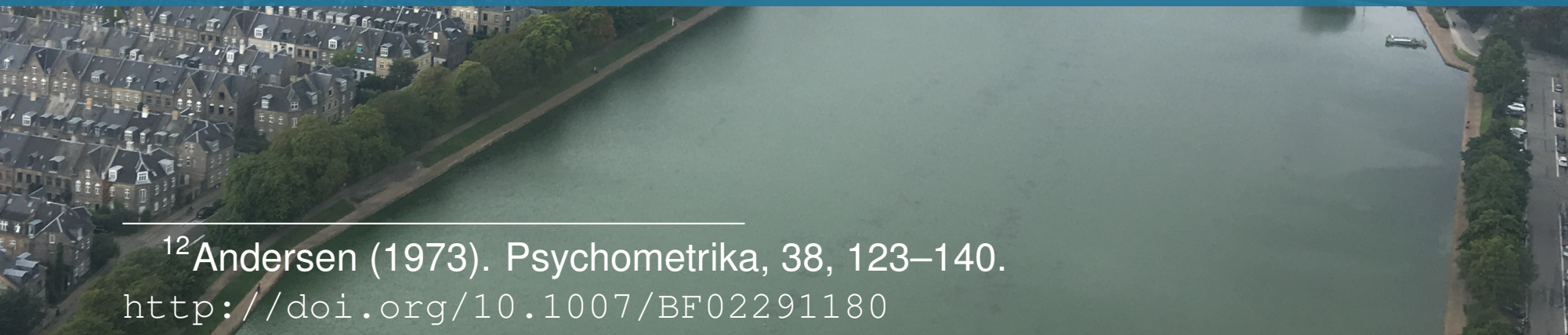
$$\sum_{v=1}^{N} Z_{iv}^2 = \sum_{r=0}^{k} \underbrace{\sum_{v:R_v=r} Z_{i|r}^2}_{iid}$$

has known null distribution[10] (abandoned[11])

[10] Christensen, Kreiner (2013). Rasch Models in Health (chap. 5, pp. 83–104), Wiley https://doi.org/10.1002/9781118574454.ch5

[11] Wright, Panchapakesan (1969). Educational and Psychological Measurement, 29, 23-48. http://doi.org/10.1177/001316446902900102

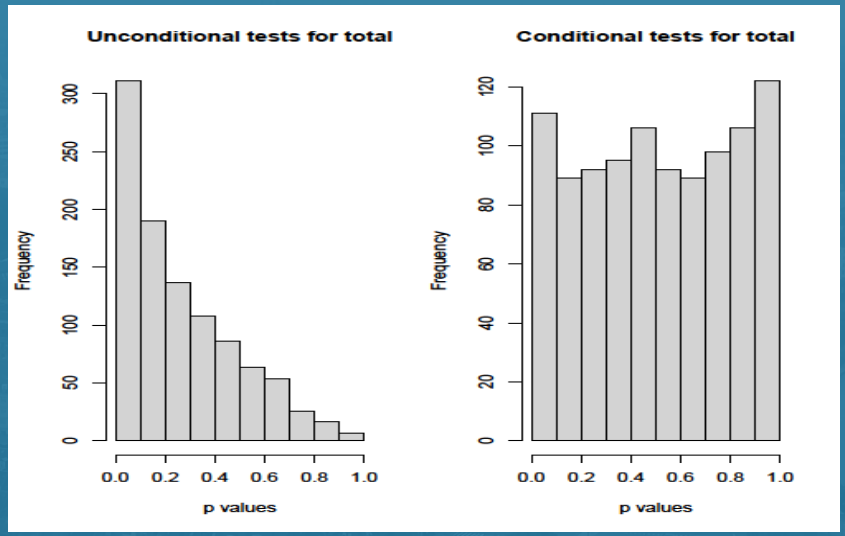
	RUMM/WINSTEPS	R/SAS
Overall model fit	total χ^2	CLR[12]
Individual item fit	OUTFIT item χ^2 :	cond. OUTFIT cond. item χ^2 :



[12] Andersen (1973). Psychometrika, 38, 123–140.
http://doi.org/10.1007/BF02291180

OVER-ALL MODEL FIT

Over-all fit[13]



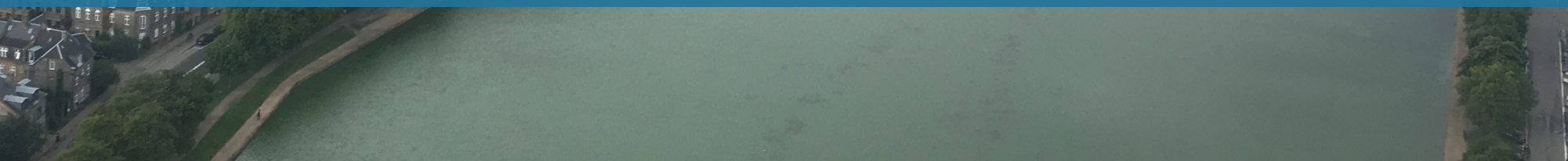
[13] Müller, Kreiner (2015). Res. Report 6, Department of Biostatistics, University of Copenhagen. https://ifsv.sund.ku.dk/biostat/annualreport/images/2/2f/Research_Report_15-06.pdf


AMTS INDIVIDUAL ITEM FIT

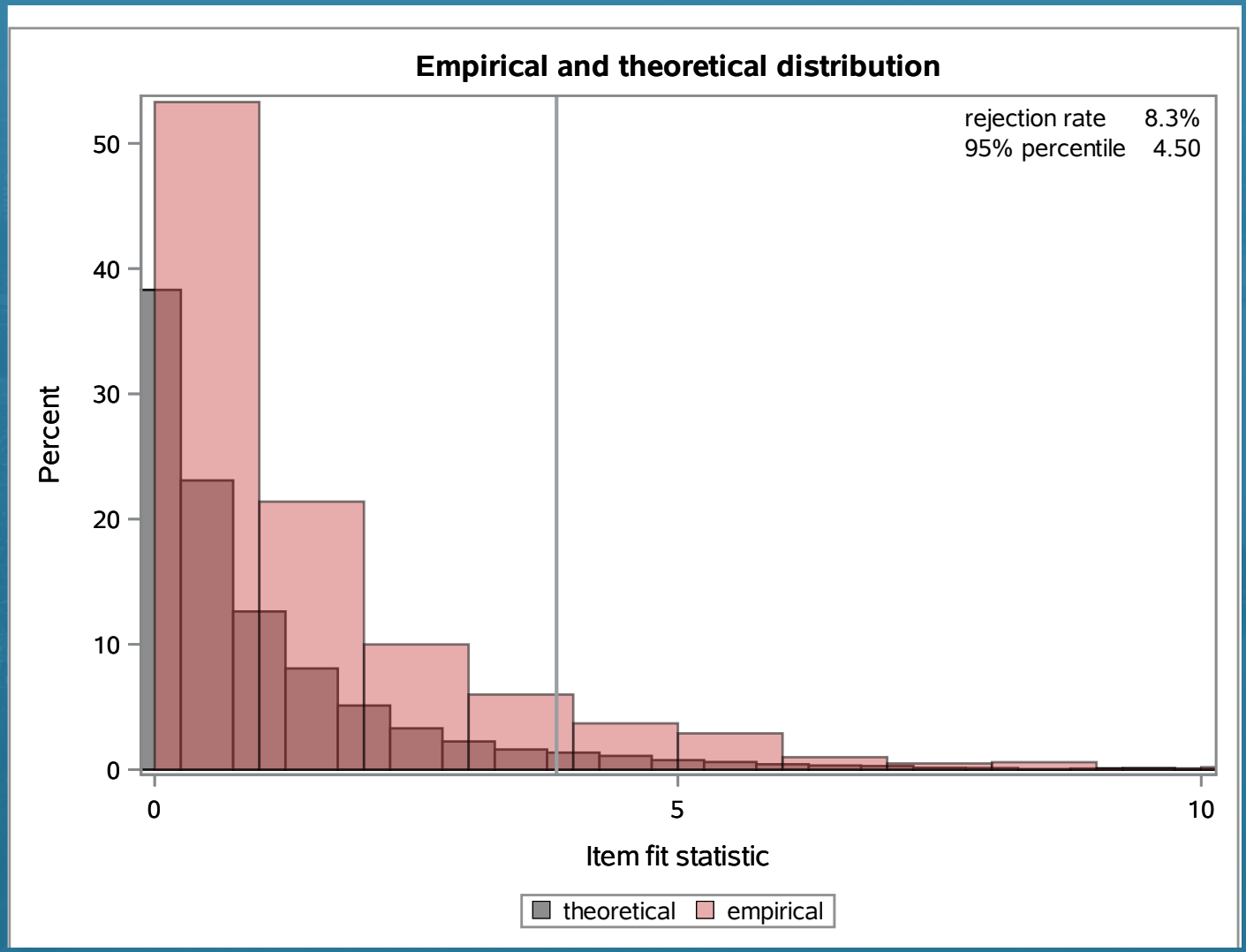
item	chisq	df	p
ADDRESS	1.73	1	0.1889
AGE	2.74	1	0.0976
COUNTBAC	0.98	1	0.3226
DOB	0.60	1	0.4382
FIRSTWW	2.25	1	0.1337
MONARCH	3.39	1	0.0656
MONTH	12.99	1	0.0003
NAME	2.74	1	0.0976
TIME	2.21	1	0.1369
YEAR	6.25	1	0.0124

- Look at the item 'MONTH'
- Simulate 1000 data sets based on $\hat{\eta}$ and $\hat{\theta}$
- Look at empirical distribution of $\chi^2(X_{MONTH})$

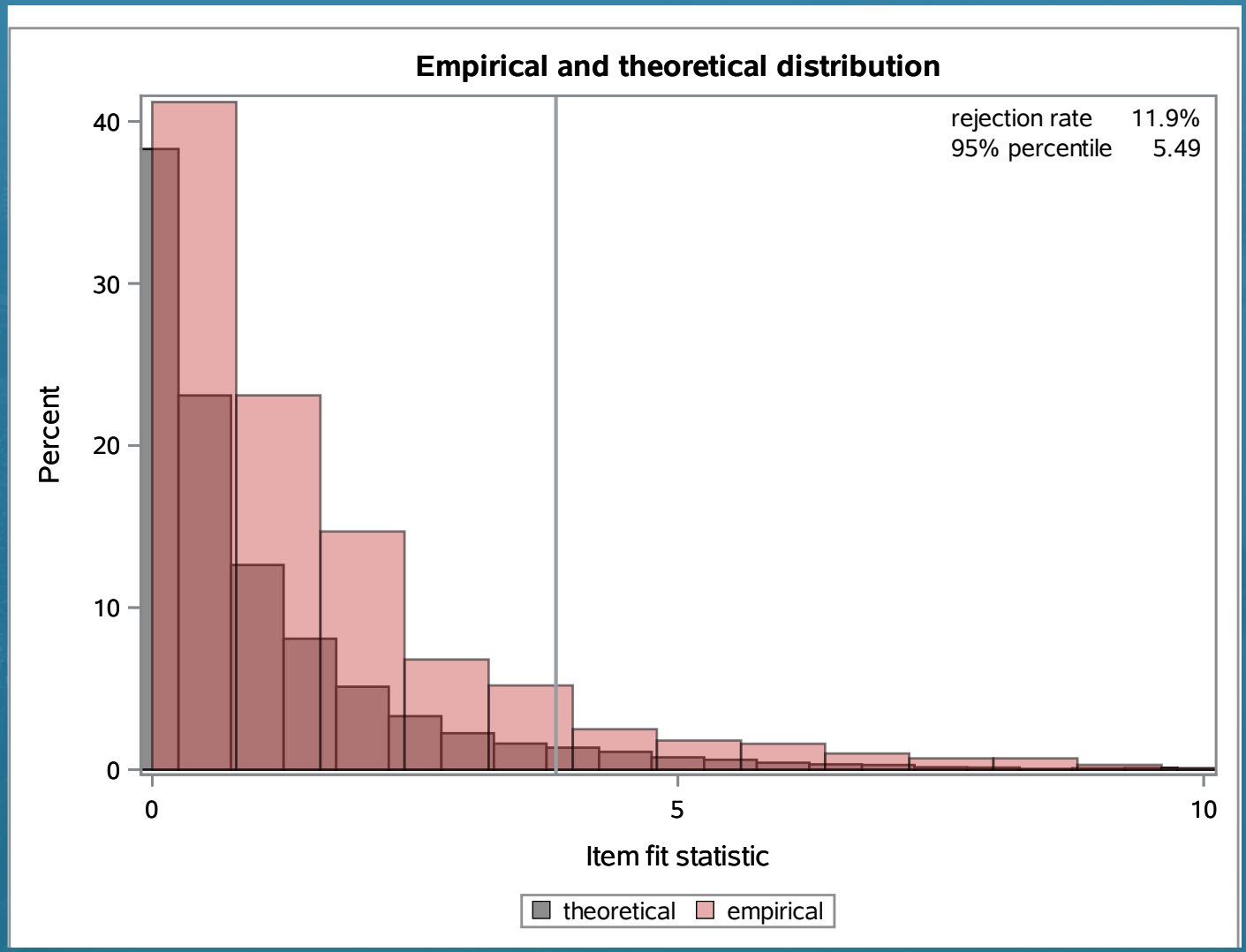
- in data set of original size
- in data set with twice the number of respondents



item 'MONTH' (sample size *N*)



item 'MONTH' (sample size $2 \cdot N$)



- Large observed χ^2 values indicate misfit

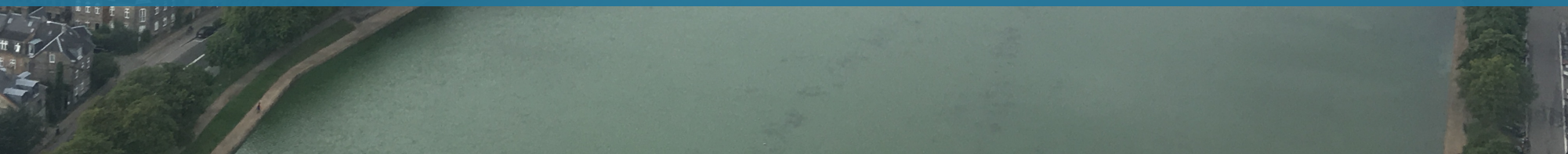
- Simulations: *p*-values cannot be trusted
- Makes it difficult to address multiple testing

- Simulations provide tables of critical values

- Fit statistics with known asymptotic distributions are known - are being implemented in SAS/R.

SUMMARY

RUMM/WINSTEPS	R/SAS
Rasch model only	Rasch part of toolbox
stand-alone software (sell-by-date)	implementations not perfect
faulty stat. methods	valid methods not readily applicable
	time to meet challenge of teaching Rasch




PERSPECTIVES

Perspectives in using simulation

- Appropriate critical value for $OUTFIT_i$?
- Appropriate critical value for $\max_i OUTFIT_i$?
- is $FitResid_i$ better then χ^2 ?

- Tabulate critical values
- Implement parametric bootstrap


THANK YOU