Extended Sequential Item Response Model for Multiple-Choice, Multiple-Attempt Test Items

Yikai Lu[1], Ying Cheng[1]

[1]University of Notre Dame

Friday 4th November, 2022

Outline

Motivation of our research

- Can we further improve person estimation ($\hat{\theta}$s) for multiple-choice test items?
- E.g.) A geography question.

Which is a European country?

A. Mexico

B. India

C. Austria

D. Australia

Can we get more information?

- **C** is the right answer!

Which is a European country?

A. Mexico

B. India

C. Austria

D. Australia

Can we get more information?

- **C** is the right answer!
- What if we allow them to rank the options in terms of plausibility?
  - e.g.) **C**DAB, D**C**AB.

Which is a European country?

A. Mexico

B. India

C. Austria

D. Australia

Can we get more information?

- **C** is the right answer!
- What if we allow them to rank the options in terms of plausibility?
  - e.g.) **C**DAB, D**C**AB.
- What if we allow them to have chances until they get the correct answer (i.e., Answer-Until-Correct procedure)?

Which is a European country?

A. Mexico

B. India

C. Austria

D. Australia

How about this problem?

Which is a European country?

A. Mexico

B. Brazil

C. Austria

D. Chile

A Multiple-Attempt Model

- **Analogy**: we dont treat all the test items equally in IRT
  - People who have the same total sum score of 8/10 could still have individual differences.
- Likewise, we don’t treat (first) wrong responses equally by allowing multiple attempts!

A Multiple-Attempt Model

- **Analogy**: we dont treat all the test items equally in IRT
  - People who have the same total sum score of 8/10 could still have individual differences.
- Likewise, we don’t treat (first) wrong responses equally by allowing multiple attempts!
- Some people have partial knowledge to identity some distractors but not all.
- *Partial information on a multiple-choice test item is defined as the ability to eliminate some, but not all, the incorrect choices, thus restricting guessing to a proper subset of choices that includes the correct choice. (Frary, 1980)*

Multiple-Attempt/Answer-Until-Correct (AUC)

- Scoring scheme (partial credits) in classical test theory
  - $s = K - k$
  - where K is the number of answer options and k is the number of attempts needed to get the correct answer option.
- Gilman and Ferry (1972) reported higher reliability than zero/one scoring, but Frary (1980) found that it failed to yield consistent improvements in reliability because of guessing and item differences.

Sequential Item Response Theory

Tutz (1990) proposed sequential item response models including the sequential Rasch model and the sequential rating scale model.

Sequential Item Response Theory

Tutz (1990) proposed sequential item response models including the sequential Rasch model and the sequential rating scale model.

- Let Y_{ik} be the response of the kth attempt to the item i.
- $Y_{ik} = 1$ if correct at the kth attempt. 0 otherwise.
- We let $P(Y_{ik} = 1 | Y_{ik-1} = 0, ..., Y_{i1} = 0, \theta) = H_{ik}(\theta)$.
- Then the unconditional probability of reaching the correct answer at the kth attempt is:

$$P(X_i = k|\theta) = P(Y_{i1} = 0, ..., Y_{ik-1} = 0, Y_{ik} = 1|\theta) \tag{1}$$

$$= \prod_{h=1}^{k-1} [1 - H_{ih}(\theta)] H_{ik}(\theta) \tag{2}$$

Sequential Item Response Theory

Tutz (1990) proposed sequential item response models including the sequential Rasch model and the sequential rating scale model.

- The sequential Rasch model is:

$$H_{ik}(\theta) = \frac{\exp(\theta - b_{ik})}{1 + \exp(\theta - b_{ik})} \tag{3}$$

$$P(X_i = k|\theta) = \prod_{h=1}^{k-1}[1 - H_{ih}(\theta)]H_{ik}(\theta) \tag{4}$$

$$= \frac{\exp(\theta - b_{ik})}{\prod_{h=1}^{k}(1 + \exp(\theta - b_{ih}))} \tag{5}$$

Sequential Item Response Theory

Table 1
Family of Sequential Models

Constraint	Model Name	Abbreviation
α_{jk}, β_{jk} unconstrained	2p(jk) sequential model	SM-2p(jk)
$\alpha_{jk} = \alpha_j$ for all k	2p(j) sequential model	SM-2p(j)
$\alpha_{jk} = \alpha_k$ for all j	2p(k) sequential model	SM-2p(k)
$\alpha_{jk} = 1$	Sequential Rasch model Tutz (1990)	SM-Rasch
$\alpha_{jk} = 1$ and $\beta_{jk} = \beta_j - \gamma_k$	Sequential rating scale model Tutz (1990)	SRSM

[1] Bergner et al. (2019)

Sequential Item Response Theory

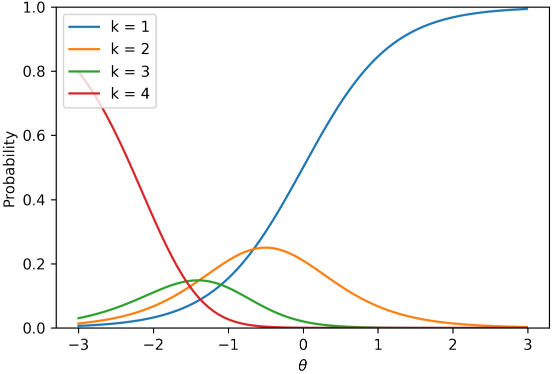


Figure: $a = 1.7, \mathbf{b} = (0, -0.5, -1, -1.5)$

Model Misspecification?

- Suppose K is the number of response choices/the maximum number of attempts.
- Then $P(X_i = K|\theta) \to 1$ as $\theta \to -\infty$.
- This means that when people have almost no ability, they always need K attempts to reach the correct choice.
- Is this natural?

SIRT models for Multiple-Choice, Multiple-Attempt Test Items

SIRT models for Multiple-Choice, Multiple-Attempt Test Items

Some thought experiments...

- What's $P(X_i = k|\theta)$ when $\theta \to -\infty$? assuming all the options look equally uncertain to them (homegenious).

SIRT models for Multiple-Choice, Multiple-Attempt Test Items

Some thought experiments...

- What's $P(X_i = k|\theta)$ when $\theta \to -\infty$? assuming all the options look equally uncertain to them (homegenious).
- It's $\frac{1}{K}$. Why?

SIRT models for Multiple-Choice, Multiple-Attempt Test Items

Some thought experiments...

- What's $P(X_i = k|\theta)$ when $\theta \to -\infty$? assuming all the options look equally uncertain to them (homegenious).
- It's $\frac{1}{K}$. Why?

$$P(X_i = 1|\theta) = \frac{1}{K} \tag{6}$$

$$P(X_i = 2|\theta) = \frac{K-1}{K} \cdot \frac{1}{K-1} = \frac{1}{K} \tag{7}$$

$$\tag{8}$$

and so on...

SIRT models for Multiple-Choice, Multiple-Attempt Test Items

Some thought experiments...

- What's $P(X_i = k|\theta)$ when $\theta \to -\infty$? assuming all the options look equally uncertain to them (homegenious).
- It's $\frac{1}{K}$. Why?

$$P(X_i = 1|\theta) = \frac{1}{K} \tag{6}$$

$$P(X_i = 2|\theta) = \frac{K-1}{K} \cdot \frac{1}{K-1} = \frac{1}{K} \tag{7}$$

$$\tag{8}$$

  and so on...
- Generating a random permutation of ABCD. The probability of C being at the kth position is the same.

SIRT models for Multiple-Choice, Multiple-Attempt Test Items

Some thought experiments...

- What's $P(X_i = k|\theta)$ when $\theta \to -\infty$? assuming all the options look equally uncertain to them (homegenious).
- It's $\frac{1}{K}$. Why?

$$P(X_i = 1|\theta) = \frac{1}{K} \tag{6}$$

$$P(X_i = 2|\theta) = \frac{K-1}{K} \cdot \frac{1}{K-1} = \frac{1}{K} \tag{7}$$

$$\tag{8}$$

  and so on...

- Generating a random permutation of ABCD. The probability of C being at the kth position is the same.
- Therefore, $P(X_i = k|\theta)$ should converge to $\frac{1}{K}$ when $\theta \to -\infty$.

SIRT models for Multiple-Choice, Multiple-Attempt Test Items

Some thought experiments...

- How about $P(X_i = k|\theta)$ when $\theta \neq -\infty$ assuming all the **distractors** are homegenious.
- Let p_T be the probability of considering the correct choice as TRUE.
- Let p_D be the probability of considering one distractor as TRUE.

SIRT models for Multiple-Choice, Multiple-Attempt Test Items

Some thought experiments...

- How about $P(X_i = k|\theta)$ when $\theta \neq -\infty$ assuming all the **distractors** are homegenious.
- Let p_T be the probability of considering the correct choice as TRUE.
- Let p_D be the probability of considering one distractor as TRUE.
- The probability of selecting the correct choice at the 1st attempt is: $\frac{p_T}{p_T+(K-1)p_D}$.

SIRT models for Multiple-Choice, Multiple-Attempt Test Items

Some thought experiments...

- How about $P(X_i = k|\theta)$ when $\theta \neq -\infty$ assuming all the **distractors** are homegenious.
- Let p_T be the probability of considering the correct choice as TRUE.
- Let p_D be the probability of considering one distractor as TRUE.
- The probability of selecting the correct choice at the 1st attempt is: $\frac{p_T}{p_T+(K-1)p_D}$.
- That is, the conditional probability $H_{ik}(\theta) = \frac{p_T}{p_T+(K-k)p_D}$.
- We want the 1st attempt probability the same as the 3PL model with a fixed guessing parameter. That is:

$$H_{i1}(\theta) = \frac{1}{K} + (1 - \frac{1}{K})\frac{\exp(a_i(\theta - b_i))}{1 + \exp(a_i(\theta - b_i))} \quad (9)$$

How can we derive $P(X_i = K|\theta)$

Some thought experiments...

- Let's consider the reciprocal!
- $\frac{1}{H_{ik}(\theta)} = \frac{p_T + (K-k)p_D}{p_T} = 1 + (K-k)\frac{p_D}{p_T}$
- Solve $\alpha = \frac{p_D}{p_T}$ by

$$\frac{1}{H_{i1}(\theta)} = \{\frac{1}{K} + (1 - \frac{1}{K})\frac{\exp(a_i(\theta - b_i))}{1 + \exp(a_i(\theta - b_i))}\}^{-1} \tag{10}$$

$$= 1 + (K-1)\alpha \tag{11}$$

- After we solve this...

$$\frac{1}{H_{ik}(\theta)} = \frac{K-k}{1 + K\exp(a(\theta - b))} + 1 \tag{12}$$

Finally

Let $f(k) = \frac{1}{H_{ik}(\theta)}$.

$$P(X_i = k|\theta) = [\prod_{h=1}^{k-1}(1 - H_{ih}(\theta))] \cdot H_{ik}(\theta) \tag{13}$$

$$= [\prod_{h=1}^{k-1}(\frac{f(h)-1}{f(h)})] \cdot \frac{1}{f(k)} \tag{14}$$

$$\ldots \tag{15}$$

$$= \frac{[1 + K\exp(a_i(\theta - b_i))]\prod_{h=1}^{k-1}(K-h)}{\prod_{h=1}^{k}[K - h + 1 + K\exp(a_i(\theta - b_i))]} \tag{16}$$

$$\tag{17}$$

This is the final form. Only parameters are a_i and b_i.

Item Category Response Function

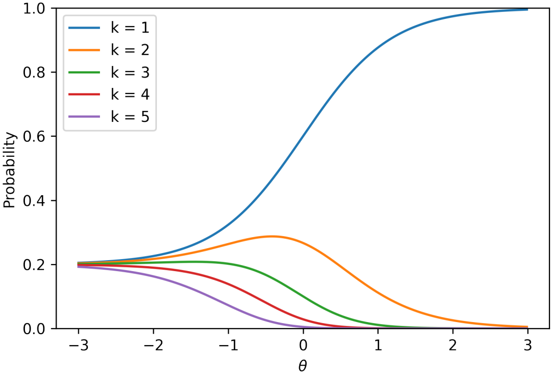


Figure: $a = 1.7, b = 0$

Information Function

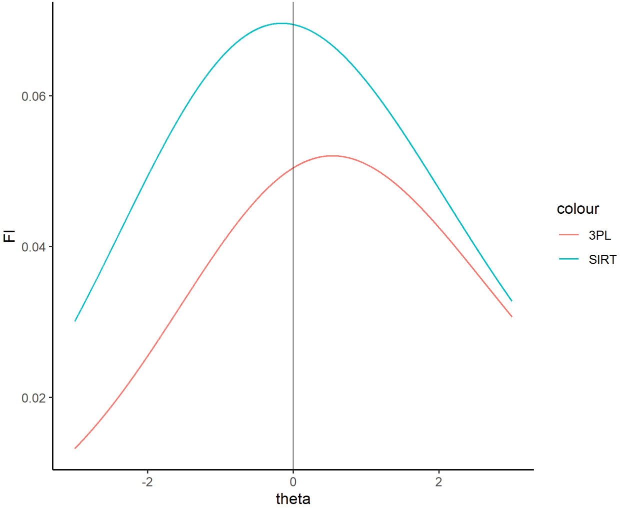


Figure: Fisher Information of the model with $a = 0.58, b = 0$ and corresponding 3PL.

Let's think back about the assumption we made...

Remember...?

- How about $P(X_i = k|\theta)$ when $\theta \neq -\infty$ assuming all the **distractors** are homegenious.

As we know, this assumption does not always hold!

To break the homogenity assumption...

$$P(X_i = k|\theta) = \frac{[1 + K\exp(a_i(\theta - b_i + \gamma_{ik}))]\prod_{i=1}^{k-1}(K-i)}{\prod_{i=1}^{k}[K - i + 1 + K\exp(a_i(\theta - b_i + \gamma_{ik}))]} \tag{18}$$

(19)

where $\gamma_{i1} = \gamma_{iK} = 0$ always.

Item Category Response Function

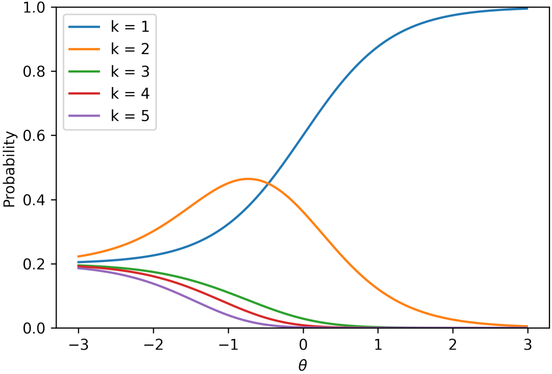


Figure: $a = 1.7, b = 0, \gamma_2 = 1$

Sequential Item Response Theory

Table 1
Family of Sequential Models

Constraint	Model Name	Abbreviation
α_{jk}, β_{jk} unconstrained	2p(jk) sequential model	SM-2p(jk)
$\alpha_{jk} = \alpha_j$ for all k	2p(j) sequential model	SM-2p(j)
$\alpha_{jk} = \alpha_k$ for all j	2p(k) sequential model	SM-2p(k)
$\alpha_{jk} = 1$	Sequential Rasch model Tutz (1990)	SM-Rasch
$\alpha_{jk} = 1$ and $\beta_{jk} = \beta_j - \gamma_k$	Sequential rating scale model Tutz (1990)	SRSM

[1] Bergner et al. (2019)

Sequential Item Response Theory

Table 1. Family of SIRT-MM models

Constraint	Description
a_{jk}, b_j, γ_{jk} unconstrained	The SIRT-MM model with the highest degrees of freedom
$a_{jk} = a_k$	
$a_{jk} = a_j$	The second SIRT-MM model we formulated (eq. 13)
$a_{jk} = a_j, \gamma_{jk} = 0$ for all $3 < k < K$	
$a_{jk} = a_j, \gamma_{jk} = 0$ for all $2 < k < K$	A reduced version of the second SIRT-MM model
$a_{jk} = a_j, \gamma_{jk} = 0$ for all $1 < k < K$	The first SIRT-MM model we formulated (eq. 10)
$\vdots$	
$a_{jk} = 1, \gamma_{jk} = 0$	The simplest SIRT-MM model

[1] Bergner et al. (2019)

Extended SIRT-MM models (SIRT-MMe)

- In SIRT-MM models, $P(X_i = k|\theta) = \frac{1}{K}$ when $\theta \to -\infty$.
- Can we break this assumption? That is, can we incorporate a parameter similar to the pseudo-guessing parameter in the 3PL model?

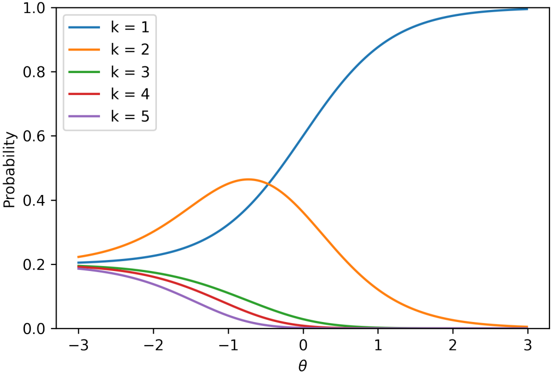


Figure: $a = 1.7, b = 0, \gamma_2 = 1$

Extended SIRT-MM models (SIRT-MMe)

Remember?

- We let $P(Y_{ik}=1|Y_{ik-1}=0,...,Y_{i1}=0,\theta)=H_{ik}(\theta)$.
- The unconditional probability of getting the correct answer at the *k*th attempt solely depends on the conditional probability of it:

$$P(X_i=k|\theta)=P(Y_{i1}=0,...,Y_{ik-1}=0,Y_{ik}=1|\theta) \quad (20)$$

$$=\prod_{h=1}^{k-1}[1-H_{ih}(\theta)]H_{ik}(\theta) \quad (21)$$

Extended SIRT-MM models (SIRT-MMe)

Remember?

- $P(X_i = k|\theta) = \prod_{h=1}^{k-1}[1 - H_{ih}(\theta)]H_{ik}(\theta)$
- The reciprocal of $H_{ik}(\theta)$ is easier to handle:

$$\frac{1}{H_{ik}(\theta)} = \frac{p_T + (K - k)p_D}{p_T} = 1 + (K - k)\frac{p_D}{p_T} \tag{22}$$

- Solve $\alpha = \frac{p_D}{p_T}$ by

$$\frac{1}{H_{i1}(\theta)} = \{c + (1 - c)\frac{1}{\exp(-a(\theta - b))}\}^{-1} = 1 + (K - 1)\alpha \tag{23}$$

- After we solve this...

$$\frac{1}{H_{ik}(\theta)} = \frac{(K - k)(1 - c)}{(K - 1)(c + \exp(a(\theta - b)))} + 1 \tag{24}$$

Extended SIRT-MM models (SIRT-MMe)

Let $f(k) = \frac{1}{H_{ik}(\theta)}$.

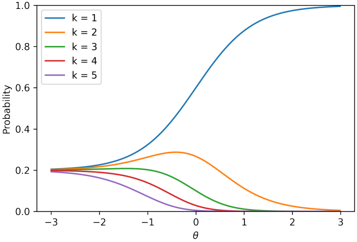
$$
\begin{aligned}
P(X_i = k|\theta) &= [\prod_{h=1}^{k-1}(1 - H_{ih}(\theta))] \cdot H_{ik}(\theta) \\
&= [\prod_{h=1}^{k-1}(\frac{f(h)-1}{f(h)})] \cdot \frac{1}{f(k)} \\
&= \frac{\prod_{h=1}^{k-1} \frac{(K-h)(1-c_i)}{(K-1)(c_i+\exp(a_i(\theta-b_i)))}}{\prod_{h=1}^{k}[\frac{(K-h)(1-c_i)}{(K-1)(c_i+\exp(a_i(\theta-b_i)))} + 1]} \\
&= \frac{[c_i + \exp(a_i(\theta - b_i))] \prod_{h=1}^{k-1} \frac{1-c_i}{K-1}(K-h)}{\prod_{i=h}^{k}[\frac{1-c_i}{K-1}(K-h) + c_i + \exp(a_i(\theta - b_i))]}
\end{aligned}
$$

To break the homogenity assumption...

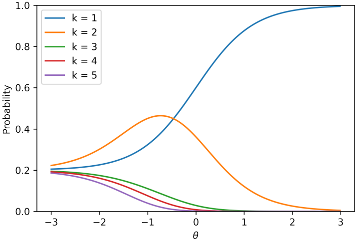
$$P(X_i = k|\theta) = \frac{[c_i + \exp(a_i(\theta - b_i + \gamma_{ik}))]\prod_{h=1}^{k-1} \frac{1-c_i}{K-1}(K-h)}{\prod_{i=h}^{k}[\frac{1-c_i}{K-1}(K-h) + c_i + \exp(a_i(\theta - b_i + \gamma_{ik}))]} \tag{25}$$

(26)

where $\gamma_{i1} = \gamma_{iK} = 0$.



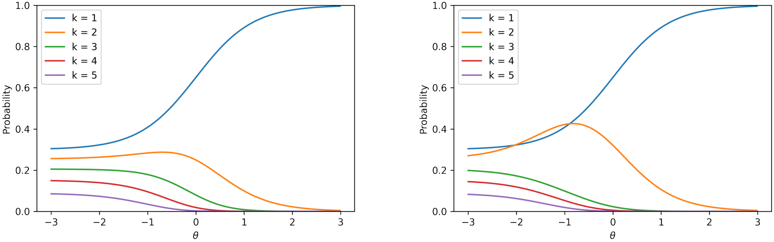
(a) $\gamma_2 = 0$



(b) $\gamma_2 = 1$

Figure: Item category response functions of SIRT-MMe models with $a = 1.7, b = 0.0, c = 0.2, K = 5$ and different γ_2. It is equivalent to a SIRT-MM model when $c = \frac{1}{K}$.

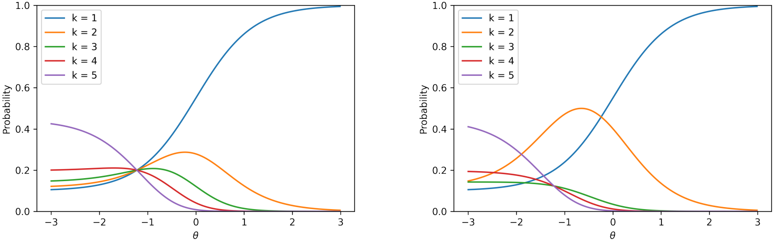
ICRFs of SIRT-MMe Models



(a) $\gamma_2 = 0$ (b) $\gamma_2 = 1$

Figure: Item category response functions of SIRT-MMe models with $a = 1.7, b = 0.0, c = 0.3, K = 5$ and different γ_2.

ICRFs of SIRT-MMe Models



(a) $\gamma_2 = 0$ (b) $\gamma_2 = 1$

Figure: Item category response functions of SIRT-MMe models with $a = 1.7, b = 0.0, c = 0.1, K = 5$ and different γ_2.

Fisher Information of SIRT-MMe Models

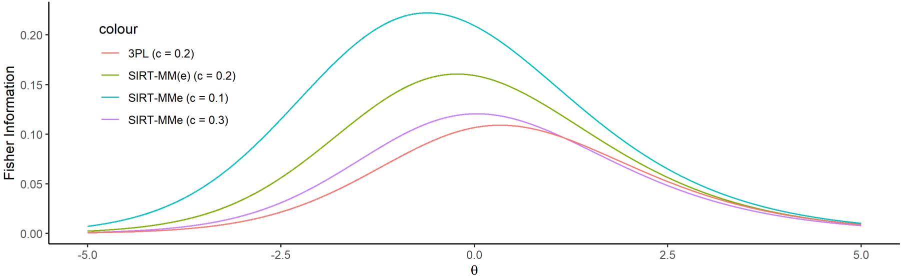


Figure: Fisher information of θ with $a = 0.8, b = 0.0, \gamma_k = 0, K = 5$.

Marginal MLE with an EM algorithm for Item Parameters

We want to maximize the marginal probability of the observed response patterns u_j:

$$P_j(\mathbf{u}_j) = \int_{-\infty}^{\infty} \prod_{i}^{M} P(X_i = u_{ji}|\theta)\phi(\theta)d\theta \tag{27}$$

The likelihood function will be:

$$L = C\prod_{j=1}^{S} P_j(\mathbf{u}_j)^{r_j} \tag{28}$$

$$\log L = \sum_{j=1}^{S} r_j \log P_j(\mathbf{u}_j) + C \tag{29}$$

where $S = min(K^M, N)$ is the number of kinds of response patterns and r_j is the number of observed response patterns j.

Integration is hard

Remember the marginal probability

$$P_j(\mathbf{u}_j) = \int_{-\infty}^{\infty} \prod_{i}^{M} P(X_i = u_{ji}|\theta)\phi(\theta)d\theta \tag{30}$$

Gauss-Hermite quadratures allow you to approximate this kind of integrals well!

$$\int_{-\infty}^{\infty} \exp(-x^2)f(x)dx \approx \sum_{i=1}^{n} w_i f(x_i) \tag{31}$$

If you want to approximate a normal distribution, just transform x_i and scale the sum a bit.

Integration is hard

We want to maximize L but it involves integration.
Solution: Use Gauss-Hermite quadratures by assuming $\phi(\theta)$ is the standard normal distribution.

$$\log L = \sum_{f=1}^{F} \sum_{i=1}^{M} \sum_{k=1}^{K} \hat{r}_{fik} \log P(X_i = k|\theta_f) + C \tag{32}$$

- Now, we only need to find $\hat{r}_{fik}$ and maximize $\log L$!
- $\hat{r}_{fik}$ is a provisional expected number of people who made k attempts for item i in θ_f.

$\hat{N}_f$ can be calculated by $\sum_{i=1}^{M} \sum_{k=1}^{K} \hat{r}_{fik}$.

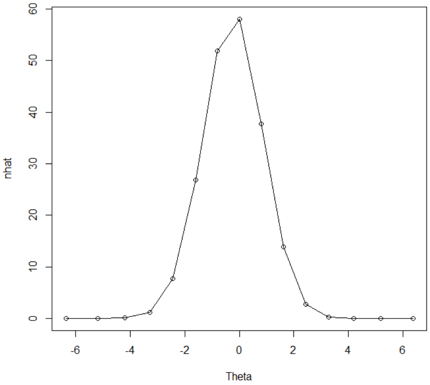


Figure: $\hat{N}_f$ from a population simulated by an uniform distribution.

Integration is hard

We want to maximize L but it involves integration.
Solution: Use Gauss-Hermite quadratures by assuming $\phi(\theta)$ is the standard normal distribution.

$$\log L = \sum_{f=1}^{F} \sum_{i=1}^{M} \sum_{k=1}^{K} \hat{r}_{fik} \log P(X_i = k|\theta_f) + C \tag{33}$$

Now, we only need to find $\hat{r}_{fik}$ and maximize $\log L$!

- However, $\hat{r}_{fik}$ is unknown. Thus, we estimate it by taking the expectation of r_{fik}, which is the E step. And then, maximize $\log L$, which is the M step.
- In the M step, we can estimate the item parameters by the Fisher's scoring (NewtonRaphson) method.

$$\mathbf{v}_q = \mathbf{v}_{q-1} + \mathbf{I}^{-1}\mathbf{t} \tag{34}$$

where $\mathbf{v}$ is the parameter estimates, $\mathbf{t}$ is the first derivative of $\log L$, and $\mathbf{I}$ is the Fisher information matrix.

Simulation Study: Item Parameter Recovery

- $M = 25$ items are simulated by
  - $K = 4$
  - $\theta \sim N(0,1)$
  - $a \sim Unif(0.75, 1.33)$
  - $c \sim Unif(0.15, 0.35)$
  - $\gamma_2 \sim Unif(0,1)$
- The c parameter needs to be regularized/penalized. In this simulation study, we used a Ridge penalty.
- We have two conditions for the b parameter: we included the $b \sim Unif(-2,2)$ and the $b \sim Unif(0,2)$ condition to see if the new SIRT model can recover the partial information, which will manifest when subjects respond to difficult items with multiple attempts.
- We take the averages of 30 replications.

Results: Item Parameter Recovery

b parameter	N	SE				BIAS				RMSE			
		b	a	c	γ_2	b	a	c	γ_2	b	a	c	γ_2
$Unif(-2,2)$	500	0.71	0.32	0.23	0.32	-0.02	0.15	-0.01	-0.03	0.69	0.47	0.21	0.32
$Unif(-2,2)$	1000	0.46	0.21	0.15	0.22	-0.06	0.05	-0.03	-0.02	0.54	0.28	0.19	0.22
$Unif(-2,2)$	2000	0.31	0.14	0.10	0.16	-0.04	0.01	-0.02	-0.01	0.44	0.21	0.15	0.18
$Unif(-2,2)$	4000	0.23	0.10	0.08	0.12	-0.05	-0.01	-0.02	0.01	0.37	0.15	0.12	0.12
$Unif(-2,2)$	8000	0.15	0.07	0.05	0.08	-0.01	0.00	-0.01	0.00	0.24	0.10	0.09	0.08
$Unif(0,2)$	500	0.47	0.34	0.11	0.26	-0.17	0.18	-0.04	-0.05	0.56	0.63	0.15	0.28
$Unif(0,2)$	1000	0.29	0.21	0.07	0.19	-0.11	0.05	-0.02	-0.03	0.40	0.32	0.11	0.19
$Unif(0,2)$	2000	0.19	0.15	0.05	0.13	-0.05	-0.00	-0.02	-0.01	0.31	0.22	0.08	0.13
$Unif(0,2)$	4000	0.12	0.10	0.03	0.10	-0.02	-0.00	-0.01	-0.00	0.18	0.15	0.05	0.10
$Unif(0,2)$	8000	0.09	0.07	0.02	0.07	-0.02	-0.01	-0.01	0.00	0.13	0.11	0.04	0.07

Table: Item Recovery Statistics for 10 conditions

Simulation Study: Person Parameter Recovery

- Item parameters are estimated by simulated response matrices.
- θ is estimated by Expected A Posteriori (i.e. $E(\theta|\mathbf{X})$).
- Our model is compared to a 3PL model.

Results: Person Parameter Recovery

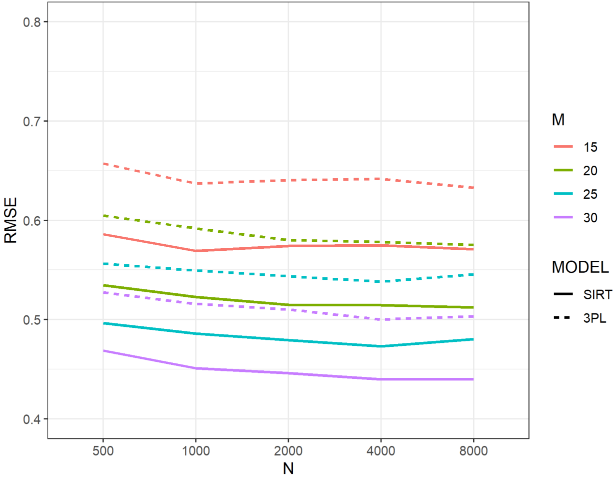


Figure: RMSE for θ estimates in $b \sim Unif(-2, 2)$ condtion.

Results: Person Parameter Recovery

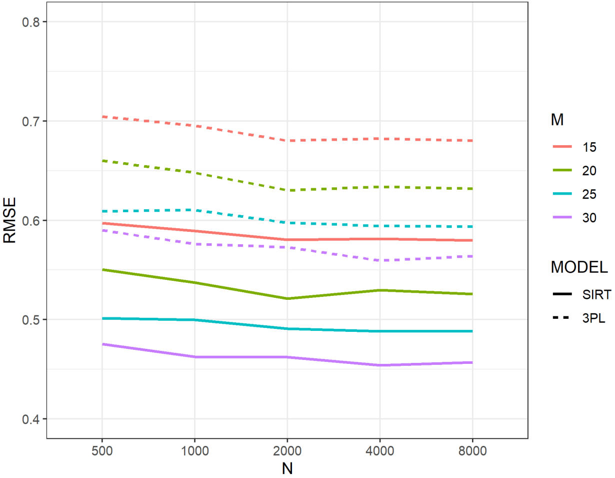
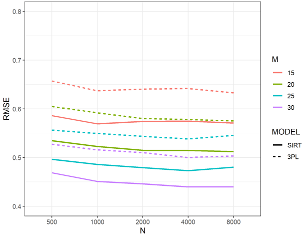
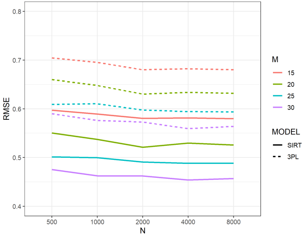


Figure: RMSE for θ estimates in $b \sim \mathit{Unif}(0, 2)$ condtion.

ICRFs of SIRT-MMe Models



(a) RMSE for θ estimates in $b \sim Unif(-2, 2)$ condtion.



(b) RMSE for θ estimates in $b \sim Unif(0, 2)$ condtion.

Results: Person Parameter Recovery

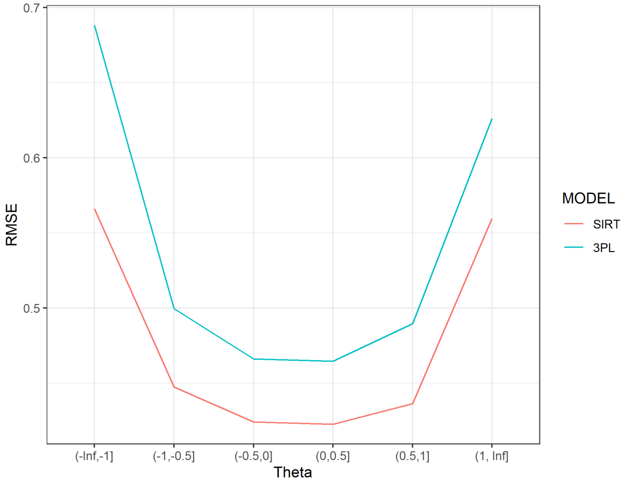


Figure: RMSE conditioning on θ ranges when $N = 4000, M = 25$.

Conclusion & Next Steps

- We have demonstrated that the new SIRT model parameters could be estimated reliably with $N = 2000$.
- It offers more accurate person parameter estimates than the 3PL model.
- Next steps include:
  - Getting reliable real data to try our model.
  - Allowing non-immediate (i.e., possibly forgetting/remembering previous answer choices) and/or longitudinal responses.
  - Jointly modeling response time.
  - Applications to Cognitive Diagnostic Models.
  - Fit statistics and model selection.
  - Detailed comparison between the SIRT and CTT in AUC.
  - Application to Computerized Adaptive Testing.

Bergner, Y., Choi, I., & Castellano, K. E. (2019). Item response models for multiple attempts with incomplete data. *Journal of Educational Measurement*, *56*(2), 415–436. https://doi.org/10.1111/jedm.12214

Frary, R. B. (1980). The effect of misinformation, partial information, and guessing on expected multiple-choice test item scores. *Applied Psychological Measurement*, *4*(1), 79–90. https://doi.org/10.1177/014662168000400109

Gilman, D. A., & Ferry, P. (1972). Increasing test reliability through self-scoring procedures. *Journal of Educational Measurement*, *9*(3), 205–207. Retrieved August 2, 2021, from https://www.jstor.org/stable/1434166

Tutz, G. (1990). Sequential item response models with an ordered response. *British Journal of Mathematical and Statistical Psychology*, *43*(1), 39–55. https://doi.org/10.1111/j.2044-8317.1990.tb00925.x

Questions?