

CUT IT SHORT: A NEW ITEM RESPONSE THEORY-BASED APPROACH FOR SHORTENING TESTS

Ottavia M. Epifania^{1,2}, Pasquale Anselmi¹, Egidio Robusto¹

`ottavia.epifania@unipd.it`

¹University of Padova

²Catholic University of the Sacred Heart

September 7th, Bologna (IT)

Statistics, Technology and Data Science for Economic and Social
Development, 2023



- 1 **Introduction**
- 2 Item Response Theory and information functions
- 3 IRT procedures for shortening tests
 - Benchmark procedure
 - Procedures based on θ targets
- 4 Practical Implications
- 5 Some final remarks

Many items/questions in a questionnaire

Good

High assessment precision

High information/reliability

But

Respondent's fatigue

Response quality might be compromised

A viable solution

A short test form (STF) with few items but high reliability

European Social Surveys

Cross-national survey carried on every two years since 2001

Assessment of attitudes, beliefs, and behavior patterns of diverse populations in different countries. Main focus → change/stability of:

- Living conditions
- Social structure
- Public opinion

Round 10:

Socio-demographic information

+

Well being, social exclusion, human values

European Social Surveys

Cross-national survey carried on every two years since 2001

Assessment of attitudes, beliefs, and behavior patterns of diverse populations in different countries. Main focus → change/stability of:

- Living conditions
- Social structure
- Public opinion

Round 10:

Socio-demographic information

+

Well being, social exclusion, human values

Focus on the item

The information at the item level is crucial → Each item taps on a specific location of the latent trait

Item Response Theory

1 Introduction

2 Item Response Theory and information functions

3 IRT procedures for shortening tests

- Benchmark procedure
- Procedures based on θ targets

4 Practical Implications

5 Some final remarks

Item Response Theory

2-PL Model

$$P(x_{pj} = 1 | \theta_p, b_j, a_j) = \frac{\exp[a_j(\theta_p - b_j)]}{1 + \exp[a_j(\theta_p - b_j)]}$$

where:

$P(x_{pj} = 1)$: Probability of endorsing item j by respondent p

θ_p : Ability of respondent p

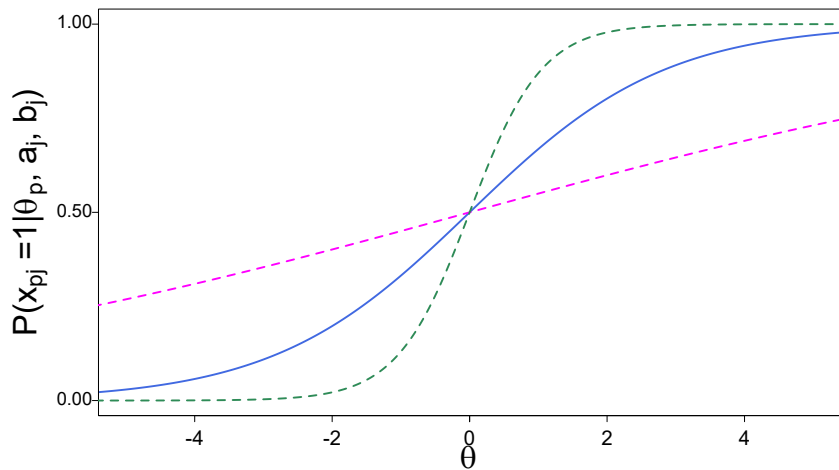
b_j : Difficulty (location on the latent trait) of item j

a_j : Discrimination of item j

Cut it short

└ IRT and Information Functions

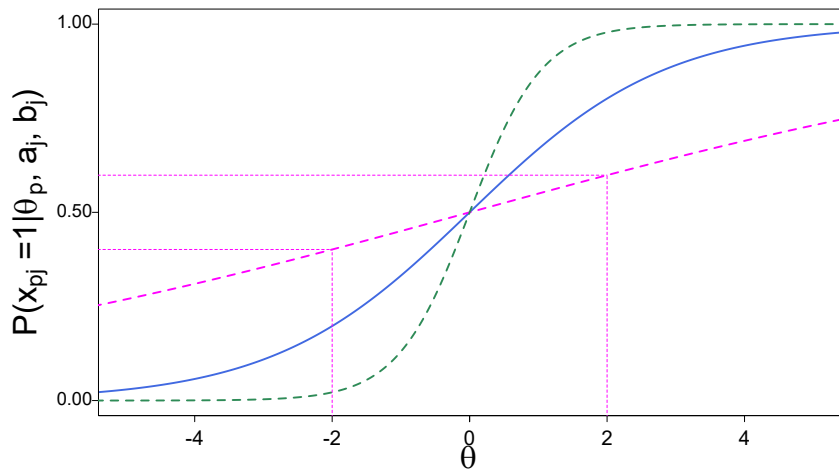
└ 2-PL Model



Cut it short

└ IRT and Information Functions

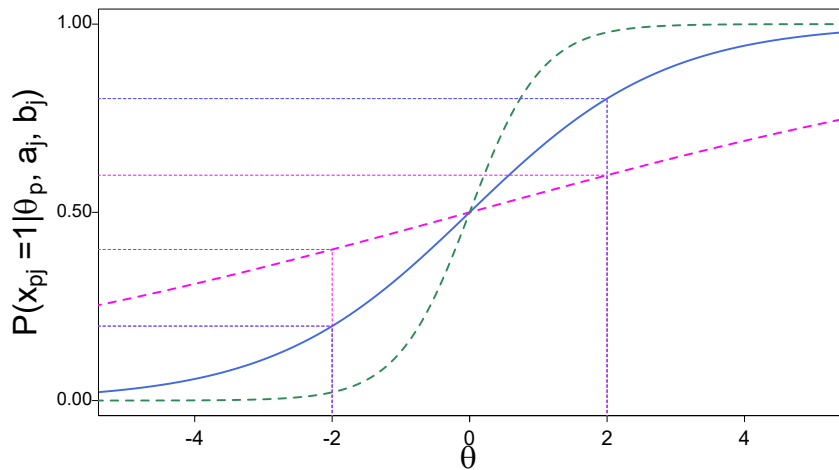
└ 2-PL Model



Cut it short

└ IRT and Information Functions

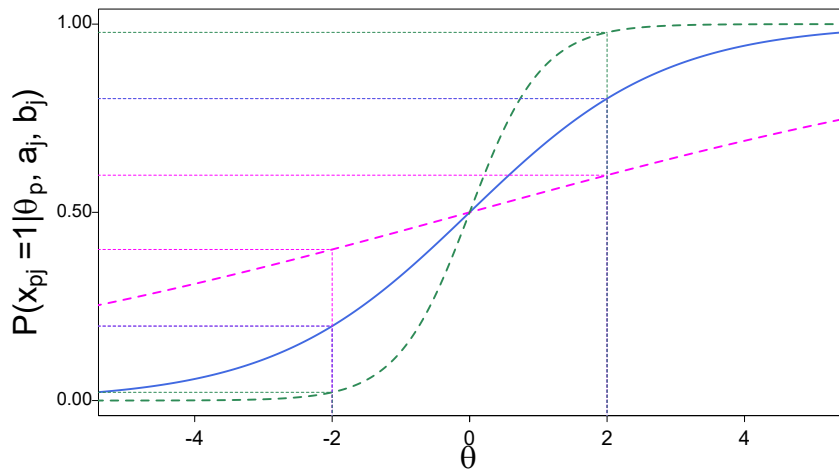
└ 2-PL Model



Cut it short

└ IRT and Information Functions

└ 2-PL Model



Information functions

Item Information Function

$$IIF_j = a_j^2 [P(\theta)(1 - P(\theta))]$$

Information functions

Item Information Function

$$IIF_j = a_j^2 [P(\theta)(1 - P(\theta))]$$

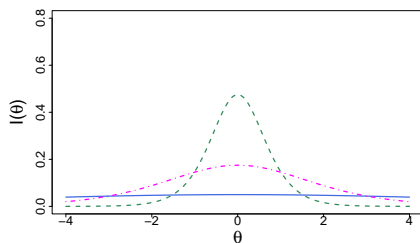


Figure 1: $a = 0.20$, $a = 0.70$, $a = 1.90$,
 $b = 0$

Information functions

Item Information Function

$$IIF_j = a_j^2 [P(\theta)(1 - P(\theta))]$$

Test Information Function

$$TIF = \sum_{j=1}^J IIF_j$$

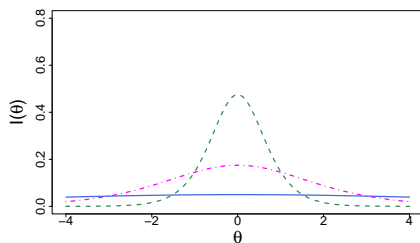


Figure 1: $a = 0.20$, $a = 0.70$, $a = 1.90$,
 $b = 0$

Information functions

Item Information Function

$$IIF_j = a_j^2 [P(\theta)(1 - P(\theta))]$$

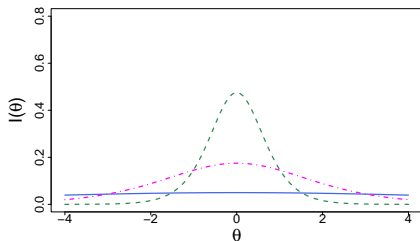


Figure 1: $a = 0.20$, $a = 0.70$, $a = 1.90$,
 $b = 0$

Test Information Function

$$TIF = \sum_{j=1}^J IIF_j$$

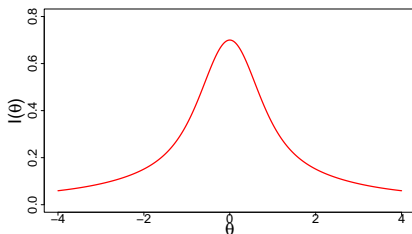


Figure 2: $TIF = IIF_1 + IIF_2 + IIF_3$

1 Introduction

2 Item Response Theory and information functions

3 IRT procedures for shortening tests

- Benchmark procedure
- Procedures based on θ targets

4 Practical Implications

5 Some final remarks

1 Introduction

2 Item Response Theory and information functions

3 IRT procedures for shortening tests

- Benchmark procedure
- Procedures based on θ targets

4 Practical Implications

5 Some final remarks

Benchmark procedure

Selected items → items with the highest *IIFs*

e.g.: 3-item short form from 10-item full-length test

item	<i>b</i>	<i>a</i>	<i>IIF</i>
1	−0.67	0.71	0.08
2	0.50	1.19	0.15
3	−2.43	0.25	0.01
4	2.12	1.98	0.24
5	1.72	0.39	0.03
6	−2.28	1.62	0.19
7	0.64	0.50	0.05
8	−2.51	1.68	0.19
9	−0.66	0.44	0.04
10	0.72	0.33	0.02

Benchmark procedure

Selected items → items with the highest *IIFs*

e.g.: 3-item short form from 10-item full-length test

item	<i>b</i>	<i>a</i>	<i>IIF</i>
4	2.12	1.98	0.24
8	−2.51	1.68	0.19
6	−2.28	1.62	0.19
2	0.50	1.19	0.15
1	−0.67	0.71	0.08
7	0.64	0.50	0.05
9	−0.66	0.44	0.04
5	1.72	0.39	0.03
10	0.72	0.33	0.02
3	−2.43	0.25	0.01

- └ Short form procedures
- └ Procedures based on θ targets

1 Introduction

2 Item Response Theory and information functions

3 IRT procedures for shortening tests

- Benchmark procedure
- Procedures based on θ targets

4 Practical Implications

5 Some final remarks

θ -target procedures

Selected items \rightarrow items with highest $IIFs$ in respect to θ targets (θ')

e.g.: 3-item short form from 10-item full-length test

	θ'_1	θ'_2	θ'_3
item	−2.67	0.01	2.67
1			
2			
3			
4			
5			
6			
7			
8			
9			
10			

θ -target procedures

Selected items \rightarrow items with highest $IIFs$ in respect to θ targets (θ')

e.g.: 3-item short form from 10-item full-length test

	θ'_1	θ'_2	θ'_3
item	-2.67	0.01	2.67
1	0.04	0.12	0.08
2	0.09	0.33	0.03
3	0.01	0.01	0.02
4	0.73	0.06	0.01
5	0.04	0.03	0.02
6	0.01	0.06	0.59
7	0.05	0.06	0.03
8	0.01	0.04	0.69
9	0.03	0.05	0.04
10	0.02	0.03	0.02

θ -target procedures

Selected items \rightarrow items with highest $IIFs$ in respect to θ targets (θ')

e.g.: 3-item short form from 10-item full-length test

item	θ'_1 -2.67	θ'_2 0.01	θ'_3 2.67
1	0.04	0.12	0.08
2	0.09	0.33	0.03
3	0.01	0.01	0.02
4	0.73	0.06	0.01
5	0.04	0.03	0.02
6	0.01	0.06	0.59
7	0.05	0.06	0.03
8	0.01	0.04	0.69
9	0.03	0.05	0.04
10	0.02	0.03	0.02

θ -target procedures

Selected items \rightarrow items with highest $IIFs$ in respect to θ targets (θ')

e.g.: 3-item short form from 10-item full-length test

item	θ'_1 -2.67	θ'_2 0.01	θ'_3 2.67
1	0.04	0.12	0.08
2	0.09	0.33	0.03
3	0.01	0.01	0.02
4	0.73	0.06	0.01
5	0.04	0.03	0.02
6	0.01	0.06	0.59
7	0.05	0.06	0.03
8	0.01	0.04	0.69
9	0.03	0.05	0.04
10	0.02	0.03	0.02

θ -target procedures

Selected items \rightarrow items with highest $IIFs$ in respect to θ targets (θ')

e.g.: 3-item short form from 10-item full-length test

item	θ'_1 -2.67	θ'_2 0.01	θ'_3 2.67
1	0.04	0.12	0.08
2	0.09	0.33	0.03
3	0.01	0.01	0.02
4	0.73	0.06	0.01
5	0.04	0.03	0.02
6	0.01	0.06	0.59
7	0.05	0.06	0.03
8	0.01	0.04	0.69
9	0.03	0.05	0.04
10	0.02	0.03	0.02

θ -target procedures

Selected items \rightarrow items with highest $IIFs$ in respect to θ targets (θ')

e.g.: 3-item short form from 10-item full-length test

item	θ'_1 -2.67	θ'_2 0.01	θ'_3 2.67
1	0.04	0.12	0.08
2	0.09	0.33	0.03
3	0.01	0.01	0.02
4	0.73	0.06	0.01
5	0.04	0.03	0.02
6	0.01	0.06	0.59
7	0.05	0.06	0.03
8	0.01	0.04	0.69
9	0.03	0.05	0.04
10	0.02	0.03	0.02

θ -target procedures

Selected items \rightarrow items with highest $IIFs$ in respect to θ targets (θ')

e.g.: 3-item short form from 10-item full-length test

item	θ'_1 -2.67	θ'_2 0.01	θ'_3 2.67
1	0.04	0.12	0.08
2	0.09	0.33	0.03
3	0.01	0.01	0.02
4	0.73	0.06	0.01
5	0.04	0.03	0.02
6	0.01	0.06	0.59
7	0.05	0.06	0.03
8	0.01	0.04	0.69
9	0.03	0.05	0.04
10	0.02	0.03	0.02

θ -target procedures

Selected items \rightarrow items with highest $IIFs$ in respect to θ targets (θ')

e.g.: 3-item short form from 10-item full-length test

item	θ'_1 -2.67	θ'_2 0.01	θ'_3 2.67
1	0.04	0.12	0.08
2	0.09	0.33	0.03
3	0.01	0.01	0.02
4	0.73	0.06	0.01
5	0.04	0.03	0.02
6	0.01	0.06	0.59
7	0.05	0.06	0.03
8	0.01	0.04	0.69
9	0.03	0.05	0.04
10	0.02	0.03	0.02

- 1 Introduction
- 2 Item Response Theory and information functions
- 3 IRT procedures for shortening tests
 - Benchmark procedure
 - Procedures based on θ targets
- 4 Practical Implications**
- 5 Some final remarks

An overall look

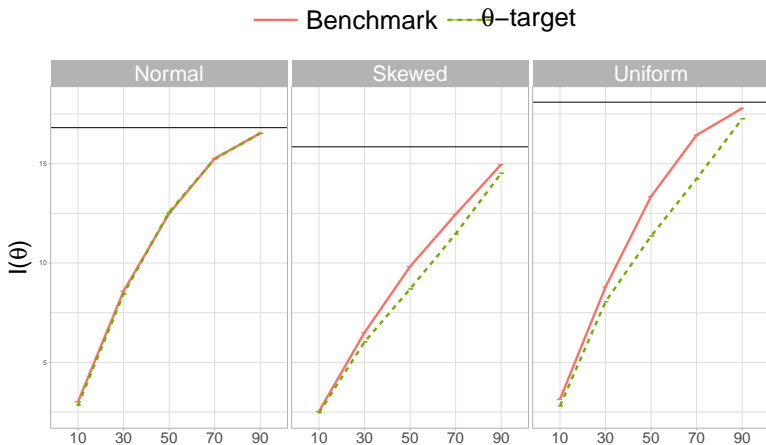


Figure 3: Overall Information of the short test forms

A closer look

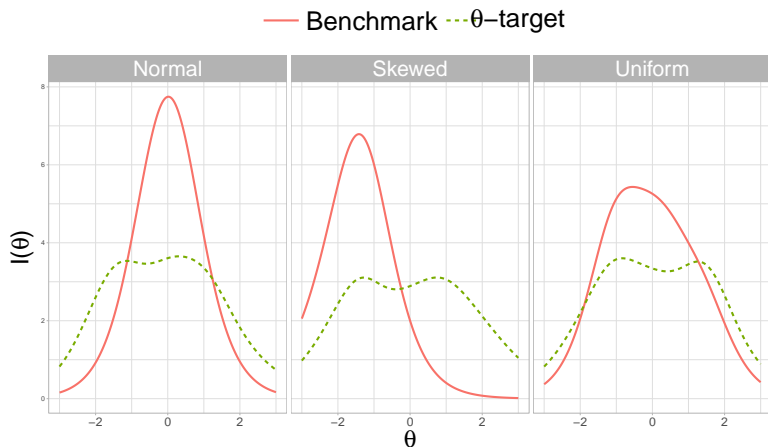


Figure 4: TIF of the 10-item short test form

- 1 Introduction
- 2 Item Response Theory and information functions
- 3 IRT procedures for shortening tests
 - Benchmark procedure
 - Procedures based on θ targets
- 4 Practical Implications
- 5 Some final remarks**

- Item response theory provides a valid framework for shortening tests without losing information and reliability
- Targeting vs. ordering: There is no “one-fits-all” solution
- In the future → Which is the ideal number of item?

Thank you!

ottavia.epifania@unipd.it

