APSTA-GE 2094
APSY-GE 2524

Modern Approaches in Measurement: Lecture 5

Klint Kanopka

New York University

Check-In

- PollEv.com/klintkanopka

Table of Contents

Announcements

- PS2 is due a week from tomorrow (2.13 @ 11.59p)
- Today comes in two parts:
  - First we learn about modeling polytomous responses
  - Second we'll work on a consulting-style analysis

Polytomous Item Responses

Polytomous Responses

Polytomous Responses

- Up until now, we've mostly dealt with dichotomous responses

Polytomous Responses

- Up until now, we’ve mostly dealt with dichotomous responses
- Two response categories do a lot of work

Polytomous Responses

- Up until now, we’ve mostly dealt with dichotomous responses
- Two response categories do a lot of work
- Correct/incorrect; yes/no; agree/disagree; prefer A or B; etc.

Polytomous Responses

- Up until now, we’ve mostly dealt with dichotomous responses
- Two response categories do a lot of work
- Correct/incorrect; yes/no; agree/disagree; prefer A or B; etc.
- Polytomous responses have multiple categories

Polytomous Responses

- Up until now, we’ve mostly dealt with dichotomous responses
- Two response categories do a lot of work
- Correct/incorrect; yes/no; agree/disagree; prefer A or B; etc.
- Polytomous responses have multiple categories
- What are some cases where this might be better than dichotomous?

Polytomous Responses

- Up until now, we've mostly dealt with dichotomous responses
- Two response categories do a lot of work
- Correct/incorrect; yes/no; agree/disagree; prefer A or B; etc.
- Polytomous responses have multiple categories
- What are some cases where this might be better than dichotomous?
- Likert scales; partial credit; etc.

Polytomous Responses

- Up until now, we’ve mostly dealt with dichotomous responses
- Two response categories do a lot of work
- Correct/incorrect; yes/no; agree/disagree; prefer A or B; etc.
- Polytomous responses have multiple categories
- What are some cases where this might be better than dichotomous?
- Likert scales; partial credit; etc.
- Polytomous responses can be ordered or not

Building a Polytomous Item Response Model

As usual, with 3 ± 1:

1. What properties of IRT models do we want to preserve for polytomous models?
2. What complications do polytomous models provide?
3. What things won't work anymore?
4. How we could force an IRT model to work with these responses?

The Fundamental Idea of Dichotomization

The Fundamental Idea of Dichotomization

- Different polytomous IRT models work by turning polytomous responses into dichotomous responses

The Fundamental Idea of Dichotomization

- Different polytomous IRT models work by turning polytomous responses into dichotomous responses
- Different models imply different dichotomizations

The Fundamental Idea of Dichotomization

- Different polytomous IRT models work by turning polytomous responses into dichotomous responses
- Different models imply different dichotomizations
- These dichotomizations are the "step functions" from the NCME ITEMS paper

The Fundamental Idea of Dichotomization

- Different polytomous IRT models work by turning polytomous responses into dichotomous responses
- Different models imply different dichotomizations
- These dichotomizations are the "step functions" from the NCME ITEMS paper
- Before, we've modeled $P(X_{ij} = 1|\theta_j)$

The Fundamental Idea of Dichotomization

- Different polytomous IRT models work by turning polytomous responses into dichotomous responses
- Different models imply different dichotomizations
- These dichotomizations are the "step functions" from the NCME ITEMS paper
- Before, we've modeled $P(X_{ij} = 1|\theta_j)$
- Think of this as having a hidden condition because X_{ij} can only be 0 or 1
  - $P(X_{ij} = 1|\theta_j, X_{ij} \in \{0, 1\})$

The Fundamental Idea of Dichotomization

- Different polytomous IRT models work by turning polytomous responses into dichotomous responses
- Different models imply different dichotomizations
- These dichotomizations are the "step functions" from the NCME ITEMS paper
- Before, we've modeled $P(X_{ij} = 1|\theta_j)$
- Think of this as having a hidden condition because X_{ij} can only be 0 or 1
  - $P(X_{ij} = 1|\theta_j, X_{ij} \in \{0, 1\})$
- For polytomous models, we apply a dichotomous IRT model after making these conditions explicit

The Fundamental Idea of Dichotomization

- Different polytomous IRT models work by turning polytomous responses into dichotomous responses
- Different models imply different dichotomizations
- These dichotomizations are the "step functions" from the NCME ITEMS paper
- Before, we've modeled $P(X_{ij} = 1|\theta_j)$
- Think of this as having a hidden condition because X_{ij} can only be 0 or 1
  - $P(X_{ij} = 1|\theta_j, X_{ij} \in \{0, 1\})$
- For polytomous models, we apply a dichotomous IRT model after making these conditions explicit
- The dichotomization we pick defines the structure of the item parameters

Adjacent Category Models

Adjacent Category Models

Adjacent Category Models

- Consider an item with four response categories, $X_i \in \{0, 1, 2, 3\}$

Adjacent Category Models

- Consider an item with four response categories, $X_i \in \{0, 1, 2, 3\}$
- Model the probability that a response is in one category versus an adjacent category

Adjacent Category Models

- Consider an item with four response categories, $X_i \in \{0, 1, 2, 3\}$
- Model the probability that a response is in one category versus an adjacent category

$$P(X_i = 1 | X_i \in \{0, 1\})$$

Adjacent Category Models

- Consider an item with four response categories, $X_i \in \{0, 1, 2, 3\}$
- Model the probability that a response is in one category versus an adjacent category

$$P(X_i = 1 | X_i \in \{0, 1\})$$

$$P(X_i = 2 | X_i \in \{1, 2\})$$

Adjacent Category Models

- Consider an item with four response categories, $X_i \in \{0, 1, 2, 3\}$
- Model the probability that a response is in one category versus an adjacent category

$$P(X_i = 1 | X_i \in \{0, 1\})$$

$$P(X_i = 2 | X_i \in \{1, 2\})$$

$$P(X_i = 3 | X_i \in \{2, 3\})$$

Adjacent Category Models

- Consider an item with four response categories, $X_i \in \{0, 1, 2, 3\}$
- Model the probability that a response is in one category versus an adjacent category

$$P(X_i = 1 | X_i \in \{0, 1\})$$

$$P(X_i = 2 | X_i \in \{1, 2\})$$

$$P(X_i = 3 | X_i \in \{2, 3\})$$

- In general for K categories, model $P(X_i = k | X_i \in \{k-1, k\})$ for $k \in \{1, \ldots, K-1\}$

Adjacent Category Models

- Consider an item with four response categories, $X_i \in \{0, 1, 2, 3\}$
- Model the probability that a response is in one category versus an adjacent category

$$P(X_i = 1 | X_i \in \{0, 1\})$$

$$P(X_i = 2 | X_i \in \{1, 2\})$$

$$P(X_i = 3 | X_i \in \{2, 3\})$$

- In general for K categories, model $P(X_i = k | X_i \in \{k - 1, k\})$ for $k \in \{1, \ldots, K - 1\}$
- Most common version: Partial Credit Model (PCM; Masters, 1982)

Partial Credit Model

- Now we have to smash these together!

Partial Credit Model

- Now we have to smash these together!

$$P(X_i = j|\theta) = \frac{\exp\left[\sum_{k=1}^{j}(\theta - b_{ik})\right]}{1 + \sum_{r=1}^{K-1}\left[\exp\left[\sum_{k=1}^{r}(\theta - b_{ik})\right]\right]}$$

Partial Credit Model

- Now we have to smash these together!

$$P(X_i = j|\theta) = \frac{\exp\left[\sum_{k=1}^{j}(\theta - b_{ik})\right]}{1 + \sum_{r=1}^{K-1}\left[\exp\left[\sum_{k=1}^{r}(\theta - b_{ik})\right]\right]}$$

- With the caveat for $j = 0$:

$$P(X_i = 0|\theta) = \frac{1}{1 + \sum_{r=1}^{K-1}\left[\exp\left[\sum_{k=1}^{r}(\theta - b_{ik})\right]\right]}$$

Partial Credit Model

- Now we have to smash these together!

$$P(X_i = j|\theta) = \frac{\exp\left[\sum_{k=1}^{j}(\theta - b_{ik})\right]}{1 + \sum_{r=1}^{K-1}\left[\exp\left[\sum_{k=1}^{r}(\theta - b_{ik})\right]\right]}$$

- With the caveat for $j = 0$:

$$P(X_i = 0|\theta) = \frac{1}{1 + \sum_{r=1}^{K-1}\left[\exp\left[\sum_{k=1}^{r}(\theta - b_{ik})\right]\right]}$$

- This is a *divide-by-total formulation* with $K - 1$ item parameters

Partial Credit Model

- Let's write this out explicitly for $j \in \{1, 2, 3\}$:

Partial Credit Model

- Let's write this out explicitly for $j \in \{1, 2, 3\}$:

$$P(X_i = 1|\theta) = \frac{\exp\left[\theta - b_{i1}\right]}{1 + \sum_{r=1}^{K-1}\left[\exp\left[\sum_{k=1}^{r}(\theta - b_{ik})\right]\right]}$$

Partial Credit Model

- Let's write this out explicitly for $j \in \{1, 2, 3\}$:

$$P(X_i = 1|\theta) = \frac{\exp\left[\theta - b_{i1}\right]}{1 + \sum_{r=1}^{K-1}\left[\exp\left[\sum_{k=1}^{r}(\theta - b_{ik})\right]\right]}$$

$$P(X_i = 2|\theta) = \frac{\exp\left[\theta - b_{i1} + \theta - b_{i2}\right]}{1 + \sum_{r=1}^{K-1}\left[\exp\left[\sum_{k=1}^{r}(\theta - b_{ik})\right]\right]} = \frac{\exp\left[2\theta - b_{i1} - b_{i2}\right]}{1 + \sum_{r=1}^{K-1}\left[\exp\left[\sum_{k=1}^{r}(\theta - b_{ik})\right]\right]}$$

Partial Credit Model

- Let's write this out explicitly for $j \in \{1, 2, 3\}$:

$$P(X_i = 1|\theta) = \frac{\exp\left[\theta - b_{i1}\right]}{1 + \sum_{r=1}^{K-1}\left[\exp\left[\sum_{k=1}^{r}(\theta - b_{ik})\right]\right]}$$

$$P(X_i = 2|\theta) = \frac{\exp\left[\theta - b_{i1} + \theta - b_{i2}\right]}{1 + \sum_{r=1}^{K-1}\left[\exp\left[\sum_{k=1}^{r}(\theta - b_{ik})\right]\right]} = \frac{\exp\left[2\theta - b_{i1} - b_{i2}\right]}{1 + \sum_{r=1}^{K-1}\left[\exp\left[\sum_{k=1}^{r}(\theta - b_{ik})\right]\right]}$$

$$P(X_i = 3|\theta) = \frac{\exp\left[\theta - b_{i1} + \theta - b_{i2} + \theta - b_{i3}\right]}{1 + \sum_{r=1}^{K-1}\left[\exp\left[\sum_{k=1}^{r}(\theta - b_{ik})\right]\right]} = \frac{\exp\left[3\theta - b_{i1} - b_{i2} - b_{i3}\right]}{1 + \sum_{r=1}^{K-1}\left[\exp\left[\sum_{k=1}^{r}(\theta - b_{ik})\right]\right]}$$

Generalized Partial Credit Model

Generalized Partial Credit Model

- Add item discrimination, one per item: a_i (Muraki, 1992)

$$P(X_i = j|\theta) = \frac{\exp\left[\sum_{k=1}^{j} a_i(\theta - b_{ik})\right]}{1 + \sum_{r=1}^{K-1}\left[\exp\left[\sum_{k=1}^{r} a_i(\theta - b_{ik})\right]\right]}$$

Generalized Partial Credit Model

- Add item discrimination, one per item: a_i (Muraki, 1992)

$$P(X_i = j|\theta) = \frac{\exp\left[\sum_{k=1}^{j} a_i(\theta - b_{ik})\right]}{1 + \sum_{r=1}^{K-1}\left[\exp\left[\sum_{k=1}^{r} a_i(\theta - b_{ik})\right]\right]}$$

- As before, for $j = 0$:

$$P(X_i = 0|\theta) = \frac{1}{1 + \sum_{r=1}^{K-1}\left[\exp\left[\sum_{k=1}^{r} a_i(\theta - b_{ik})\right]\right]}$$

Generalized Partial Credit Model

- Add item discrimination, one per item: a_i (Muraki, 1992)

$$P(X_i = j|\theta) = \frac{\exp\left[\sum_{k=1}^{j} a_i(\theta - b_{ik})\right]}{1 + \sum_{r=1}^{K-1}\left[\exp\left[\sum_{k=1}^{r} a_i(\theta - b_{ik})\right]\right]}$$

- As before, for $j = 0$:

$$P(X_i = 0|\theta) = \frac{1}{1 + \sum_{r=1}^{K-1}\left[\exp\left[\sum_{k=1}^{r} a_i(\theta - b_{ik})\right]\right]}$$

- Let's build a GPCM IRF in Desmos

Threshold Models

Threshold Models

Threshold Models

- Build a model that checks whether a response is in a category or higher

Threshold Models

- Build a model that checks whether a response is in a category or higher
- Model $P(X_i \geq k)$ for $k \in \{1, \ldots, K-1\}$

Threshold Models

- Build a model that checks whether a response is in a category or higher
- Model $P(X_i \geq k)$ for $k \in \{1, \ldots, K-1\}$
- Most common: Graded Response Model (GRM; Samejima, 1969)

Threshold Models

- Build a model that checks whether a response is in a category or higher
- Model $P(X_i \geq k)$ for $k \in \{1, \ldots, K-1\}$
- Most common: Graded Response Model (GRM; Samejima, 1969)

$$P(X_i \geq 0) = 1$$

Threshold Models

- Build a model that checks whether a response is in a category or higher
- Model $P(X_i \geq k)$ for $k \in \{1, \ldots, K-1\}$
- Most common: Graded Response Model (GRM; Samejima, 1969)

$$P(X_i \geq 0) = 1$$

$$P(X_i \geq 1) = \frac{1}{1 + \exp\big(-a_i(\theta - b_{i1})\big)}$$

Threshold Models

- Build a model that checks whether a response is in a category or higher
- Model $P(X_i \geq k)$ for $k \in \{1, \ldots, K-1\}$
- Most common: Graded Response Model (GRM; Samejima, 1969)

$$P(X_i \geq 0) = 1$$

$$P(X_i \geq 1) = \frac{1}{1 + \exp\big(-a_i(\theta - b_{i1})\big)}$$

$$P(X_i \geq k) = \frac{1}{1 + \exp\big(-a_i(\theta - b_{ik})\big)}$$

Graded Response Model

Graded Response Model

- Model threshold probabilities, then recover category probabilities

Graded Response Model

- Model threshold probabilities, then recover category probabilities

$$P(X_i = 0) = 1 - P(X_i \geq 1)$$

Graded Response Model

- Model threshold probabilities, then recover category probabilities

$$P(X_i = 0) = 1 - P(X_i \geq 1)$$

$$P(X_i = 1) = P(X_i \geq 1) - P(X_i \geq 2)$$

Graded Response Model

- Model threshold probabilities, then recover category probabilities

$$P(X_i = 0) = 1 - P(X_i \geq 1)$$

$$P(X_i = 1) = P(X_i \geq 1) - P(X_i \geq 2)$$

$$P(X_i = 2) = P(X_i \geq 2) - P(X_i \geq 3)$$

Graded Response Model

- Model threshold probabilities, then recover category probabilities

$$P(X_i = 0) = 1 - P(X_i \geq 1)$$

$$P(X_i = 1) = P(X_i \geq 1) - P(X_i \geq 2)$$

$$P(X_i = 2) = P(X_i \geq 2) - P(X_i \geq 3)$$

- For this to work, we require $b_{ik} \leq b_{i,k+1}$

Graded Response Model

- Model threshold probabilities, then recover category probabilities

$$P(X_i = 0) = 1 - P(X_i \geq 1)$$

$$P(X_i = 1) = P(X_i \geq 1) - P(X_i \geq 2)$$

$$P(X_i = 2) = P(X_i \geq 2) - P(X_i \geq 3)$$

- For this to work, we require $b_{ik} \leq b_{i,k+1}$
- So GRM enforces category ordering more strictly than GPCM

Graded Response Model

- Model threshold probabilities, then recover category probabilities

$$P(X_i = 0) = 1 - P(X_i \geq 1)$$

$$P(X_i = 1) = P(X_i \geq 1) - P(X_i \geq 2)$$

$$P(X_i = 2) = P(X_i \geq 2) - P(X_i \geq 3)$$

- For this to work, we require $b_{ik} \leq b_{i,k+1}$
- So GRM enforces category ordering more strictly than GPCM
- Let's build a GRM IRF in Desmos

Nominal Response Model

Nominal Response Model

Nominal Response Model

- Here we do not enforce category ordering

Nominal Response Model

- Here we do not enforce category ordering
- Model each response category relative to a baseline category $k = 0$

Nominal Response Model

- Here we do not enforce category ordering
- Model each response category relative to a baseline category $k = 0$
- $P(X_i = k | X_i \in \{0, k\})$ for $k \in \{1, \ldots, K-1\}$

Nominal Response Model

- Here we do not enforce category ordering
- Model each response category relative to a baseline category $k = 0$
- $P(X_i = k | X_i \in \{0, k\})$ for $k \in \{1, \ldots, K-1\}$

$$P(X_i = k | \theta, X_i \in \{0, k\}) = \frac{1}{1 + \exp(a_{ik}\theta + b_{ik})}$$

Nominal Response Model

- Here we do not enforce category ordering
- Model each response category relative to a baseline category $k = 0$
- $P(X_i = k | X_i \in \{0, k\})$ for $k \in \{1, \ldots, K-1\}$

$$P(X_i = k | \theta, X_i \in \{0, k\}) = \frac{1}{1 + \exp(a_{ik}\theta + b_{ik})}$$

$$P(X_i = 0 | \theta) = \frac{1}{1 + \sum_{k=1}^{K-1} \exp(a_{ik}\theta + b_{ik})}$$

Nominal Response Model

- Here we do not enforce category ordering
- Model each response category relative to a baseline category $k = 0$
- $P(X_i = k | X_i \in \{0, k\})$ for $k \in \{1, \ldots, K-1\}$

$$P(X_i = k | \theta, X_i \in \{0, k\}) = \frac{1}{1 + \exp(a_{ik}\theta + b_{ik})}$$

$$P(X_i = 0 | \theta) = \frac{1}{1 + \sum_{k=1}^{K-1} \exp(a_{ik}\theta + b_{ik})}$$

$$P(X_i = k | \theta) = \frac{\exp(a_{ik}\theta + b_{ik})}{1 + \sum_{k=1}^{K-1} \exp(a_{ik}\theta + b_{ik})}$$

Nominal Response Model

- Here we do not enforce category ordering
- Model each response category relative to a baseline category $k = 0$
- $P(X_i = k | X_i \in \{0, k\})$ for $k \in \{1, \ldots, K-1\}$

$$P(X_i = k|\theta, X_i \in \{0, k\}) = \frac{1}{1 + \exp(a_{ik}\theta + b_{ik})}$$

$$P(X_i = 0|\theta) = \frac{1}{1 + \sum_{k=1}^{K-1} \exp(a_{ik}\theta + b_{ik})}$$

$$P(X_i = k|\theta) = \frac{\exp(a_{ik}\theta + b_{ik})}{1 + \sum_{k=1}^{K-1} \exp(a_{ik}\theta + b_{ik})}$$

- This is *much* more flexible than other formulations!

Nominal Response Model

- Here we do not enforce category ordering
- Model each response category relative to a baseline category $k = 0$
- $P(X_i = k | X_i \in \{0, k\})$ for $k \in \{1, ..., K-1\}$

$$P(X_i = k|\theta, X_i \in \{0, k\}) = \frac{1}{1 + \exp(a_{ik}\theta + b_{ik})}$$

$$P(X_i = 0|\theta) = \frac{1}{1 + \sum_{k=1}^{K-1} \exp(a_{ik}\theta + b_{ik})}$$

$$P(X_i = k|\theta) = \frac{\exp(a_{ik}\theta + b_{ik})}{1 + \sum_{k=1}^{K-1} \exp(a_{ik}\theta + b_{ik})}$$

- This is *much* more flexible than other formulations!
- We have one discrimination per comparison category and uses "slope-intercept" form

Nominal Response Model

- Here we do not enforce category ordering
- Model each response category relative to a baseline category $k = 0$
- $P(X_i = k | X_i \in \{0, k\})$ for $k \in \{1, \ldots, K-1\}$

$$P(X_i = k|\theta, X_i \in \{0, k\}) = \frac{1}{1 + \exp(a_{ik}\theta + b_{ik})}$$

$$P(X_i = 0|\theta) = \frac{1}{1 + \sum_{k=1}^{K-1} \exp(a_{ik}\theta + b_{ik})}$$

$$P(X_i = k|\theta) = \frac{\exp(a_{ik}\theta + b_{ik})}{1 + \sum_{k=1}^{K-1} \exp(a_{ik}\theta + b_{ik})}$$

- This is *much* more flexible than other formulations!
- We have one discrimination per comparison category and uses "slope-intercept" form
- Let's build a NRM IRF in Desmos

Fitting Polytomous Models with `mirt`

Fitting Polytomous Models with `mirt`

- Data should be an item response matrix
  - Individual responses are numeric
  - If response categories are ordered, use ordered integers
  - It does not matter whether the lowest category starts at 0 or 1

Fitting Polytomous Models with `mirt`

- Data should be an item response matrix
  - Individual responses are numeric
  - If response categories are ordered, use ordered integers
  - It does not matter whether the lowest category starts at 0 or 1
- Choose model type using `itemtype` in `mirt()`
  - `'Rasch'` fits a partial credit model with discriminations fixed to 1
  - `'gpcm'` fits a generalized partial credit model
  - `'graded'` fits a graded response model
  - `'nominal'` fits a nominal response model

Break

Consulting Query

Consulting Query

- Many of us are engaged in our own research or work
- If you have a question related to class content and can share context or data, we can discuss it as a group
- If you can share data, it's possible that we even can do your work for you!
- The goal is for this class to be responsive and useful for your own needs

The Problem

Someone approaches you with the following request seeking your newfound consulting expertise:

I am conducting a study and want to measure grit. I found some questions online and asked them to a bunch of people without doing any real pilot or preparation. I spent all of my research budget doing this, so I can't recollect my data. I need grit scores for each person and want to see how they relate to age. Can you help me make sense of this?

With 3 ± 1, you'll figure out how screwed they are! Download the data here.

- Take a look at the data
  - What looks good?
  - What looks bad?
  - What information do you have?
- Come up with a (vague) plan!
  - We'll check in before I turn you loose to execute your plan

The Items

Items are rated on a 5-point scale from 1 being *not at all like me* to 5 being *very much like me*. The item names and their text are listed below:

- (`grit_diligent`) I am diligent.
- (`grit_finish`) I finish whatever I begin.
- (`grit_focus`) I have difficulty maintaining my focus on projects that take more than a few months to complete.
- (`grit_goals`) I often set a goal but later choose to pursue a different one.
- (`grit_hardworking`) I am a hard worker.
- (`grit_interest`) I have been obsessed with a certain idea or project for a short time but later lost interest.
- (`grit_new_projects`) New ideas and projects sometimes distract me from previous ones.
- (`grit_setbacks`) Setbacks don’t discourage me.

Wrap Up

Recap

- If you have questions tied to your own research/work, please reach out
- Core idea: polytomous models imply fundamental dichotomizations
  - Adjacent categories
  - Thresholding
  - One-vs-all
- These are all estimable in `mirt`
  - The model you select implies how strongly you believe in response ordering
  - For polytomous items, parameters need to be interpreted together

Check-Out

- PollEv.com/klintkanopka