Lecture 11: Polytomous Item Response Models (II)

(Baker & Kim (2004), Chapters 8 & 9, Embretson & Reise (2000), Chapter 5)

Partial Credit Model (PCM; Masters, 1982)

The partial credit model is based on the idea that the Rasch model be used to represent the conditional probability of scoring in category l of a k-category item given that the score is either in category $l-1$ or l:

$$\frac{P_{ijl}}{P_{ij,l-1}+P_{ijl}}=\frac{\exp(\theta_j-b_{il})}{1+\exp(\theta_j-b_{il})}$$

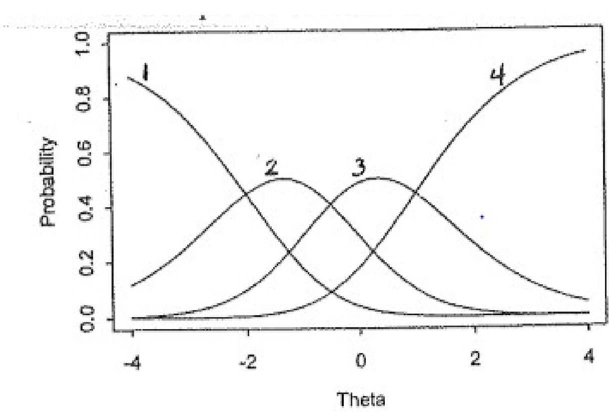
Because $\sum_{l=1}^{k}P_{il}(\theta_j)=1$ for all θ, the probability of scoring in score category l can be computed as:

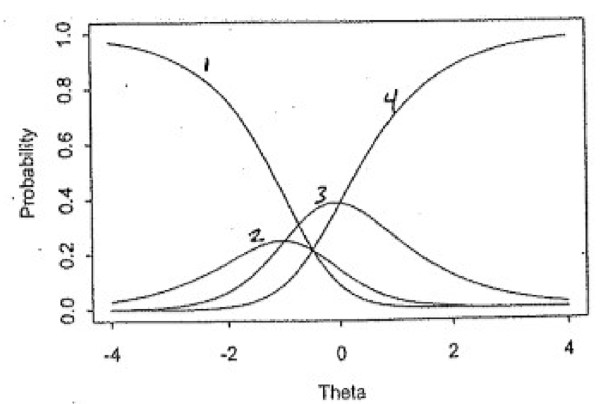
$$P_{ijl}=\frac{\exp\sum_{h=0}^{l}(\theta_j-b_{ih})}{\sum_{h=0}^{k}\exp\sum_{g=0}^{h}(\theta_j-b_{ig})},$$

where b_{i0}=0 for all θ_j.

Interpretation of Parameters

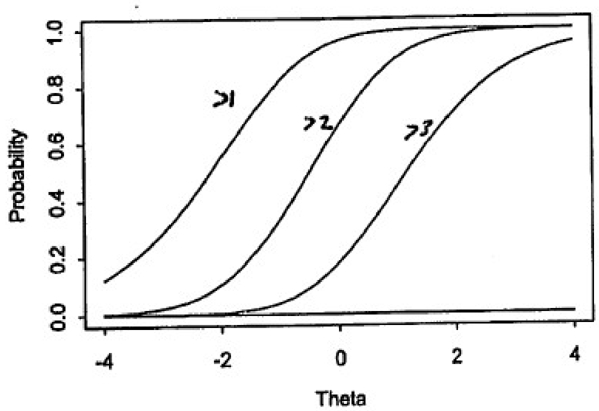
Because of PCM belongs to the Rasch family, there are only $k-1$ threshold parameters to be estimated for each item, and no slope parameters. An example of category characteristic curves for a couple of PCM items are shown below:

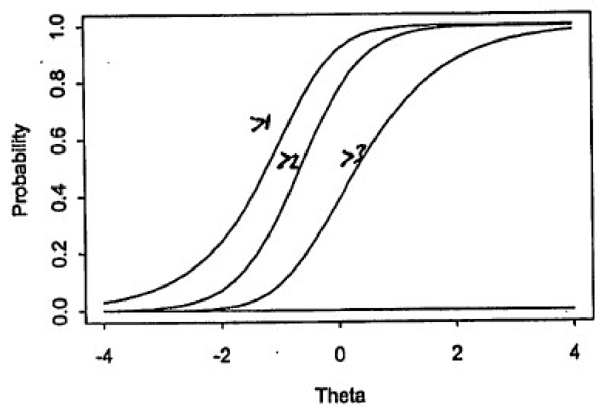




- Notice that the b_{il} parameters determine the locations at which the category characteristic curves for the score $l-1$ and l categories cross.

- Unlike the GRM, the threshold parameters do not have to be ordered. (They only determine the relative probabilities of scoring in adjacent categories, not all of the categories).
- The difference between the GRM and PCM is most noticeable with respect to their cumulative response functions. Notice that with the PCM, the boundary curves no longer conform to the same model (2PL with the same slope parameter in the GRM). Samejima distinguished polytomous models that are homogeneous (the boundary curves are based on the same model) and those that are heterogeneous (the boundary curves are not based on the same model).



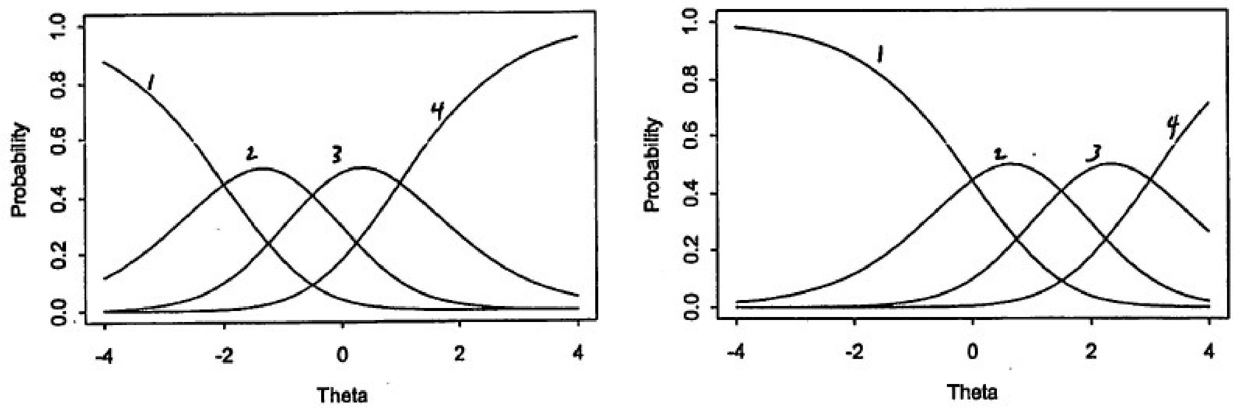


<u>Some additional features of the PCM</u>

- Conditional maximum likelihood estimation is possible. Because test score is a sufficient statistic for θ, it ultimately can be used to take the place of θ in finding maximum likelihood estimates of the threshold parameters.
- Evaluating items based on locations of thresholds. Although there is not a slope parameter, the spacing and location of threshold parameters can be used to evaluate the usefulness of the item at various points along the θ scale.
- Related models - Rating scale model (Andrich, 1978). The idea behind the rating scale model is that the spacing of the item categories stays the same across items. But the items can differ in terms of the locations of the thresholds categories. In effect, each item has its own unique location parameter, but a common difference exists between category thresholds across items. Sometimes applications of this model refer to a decomposition of the item category thresholds:

$$b_{im} = b_i + d_m .$$

Examples of 2 items from the rating scale model:



Generalized Partial Credit Model (GPCM; Muraki, 1992)

The GPCM is essentially an extension of the PCM that incorporates a slope parameter that can vary across items.

The conditional probability of scoring in category l given that the score is either in category $l-1$ or l is now given by

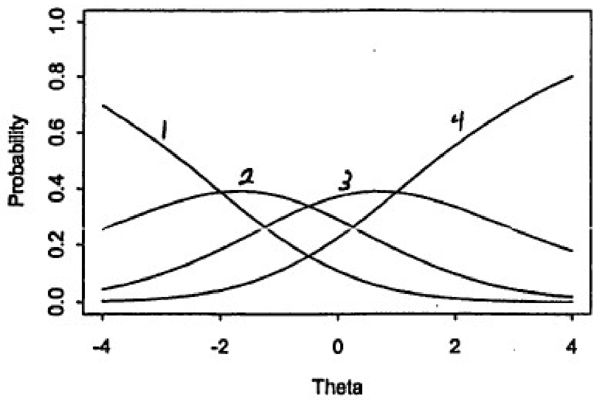
$$\frac{P_{ijl}}{P_{ij,l-1}+P_{ijl}}=\frac{\exp[a_i(\theta_j-b_{il})]}{1+\exp[a_i(\theta_j-b_{il})]}$$

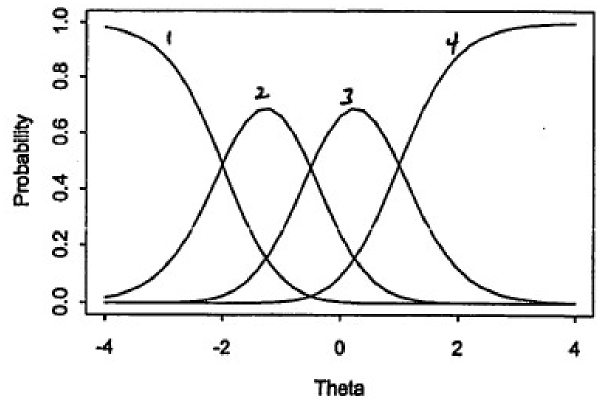
and the probability of scoring in category l is given by

$$P_{ijl}=\frac{\exp\sum_{h=0}^{l}a_i(\theta_j-b_{ih})}{\sum_{h=0}^{k}\exp\sum_{g=0}^{h}a_i(\theta_j-b_{ig})},$$

where b_{i0}=0 for all θ_j.

Some examples of GPCM items that differ only in the slope parameters:





Nominal Response Model (NRM; Bock, 1975)

The nominal response model, unlike the PCM and GPCM, does not assume an ordering to the score categories for an item. Its basis, like the PCM and GPCM, however, is in the conditional probabilities of item score categories.

The choice model:

$$P_A = \frac{\pi_A}{\pi_A + \pi_B}$$

$$P_B = \frac{\pi_B}{\pi_A + \pi_B}$$

For a *k*-category item, this looks like

$$P_l = \frac{\pi_l}{\pi_1 + \pi_2 + \ldots + \pi_m}$$

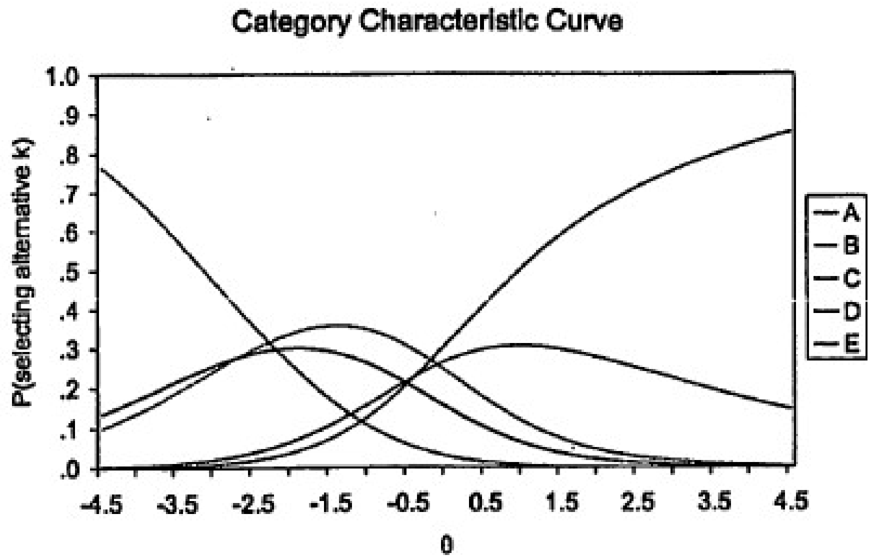
In the nominal model,

$$\pi_l = \exp(a_{il}\theta_j + c_{il}),$$

so the probability of scoring in a particular score category is:

$$P_{il}(\theta) = \frac{\exp(a_{il}\theta_j + c_{il})}{\sum_{l=1}^{m} \exp(a_{il}\theta_j + c_{il})}$$

Example of category characteristic curves for a nominal response item:



Some properties of the item parameters of the nominal response model:

- For each item, one category should be monotonically increasing in θ and one monotonically decreasing in θ. The category with the largest a parameter is monotonically increasing, the one with the smallest a is monotonically decreasing.
- Categories with the same a have the same maximum, regardless of the value of c.
- There is an indeterminacy to the a and c parameters that requires fixing one a and one c. It is common to resolve this indeterminacy by making the mean of the a s and the mean of the c s equal to zero across all categories within an item.

<u>A common formulation for these four models</u>

Thissen and Steinberg (1986) discusses an alternative parameterization of the nominal model that can also be used to illustrate the similarities of each. This nominal model is the most general model. Each of the other three can be shown to be a special case of the nominal model. In MULTILOG, we fit the PCM, rating scale, and GPCM models by essentially imposing constraints on the nominal model.

```
ITEM   1:      5 NOMINAL CATEGORIES,  5 HIGH
 CATEGORY(K): 1     2     3     4     5
  A(K)      0.11  -0.98  -0.04  1.05  -0.14
  C(K)      0.39  -1.79   0.42  2.14  -1.16

@THETA:      INFORMATION:   (Theta values increase in steps of 0.2)
 -3.0 - -1.6  0.349  0.370  0.388  0.404  0.415  0.423  0.427  0.427
 -1.4 -  0.0  0.421  0.410  0.393  0.372  0.346  0.318  0.287  0.256
  0.2 -  1.6  0.225  0.196  0.168  0.144  0.121  0.102  0.085  0.071
  1.8 -  3.0  0.059  0.049  0.040  0.033  0.027  0.022  0.018
```

Theta -3.0: 0.349
Theta -2.8: 0.370
Theta -2.6: 0.388
⋮

Theta 1.2: 0.102
Theta 1.4: 0.085
Theta 1.6: 0.071

```
OBSERVED AND EXPECTED COUNTS/PROPORTIONS IN
CATEGORY(K):  1      2      3      4      5
OBS. FREQ.    133    34     147    697    32
OBS. PROP.  0.1275 0.0326 0.1409 0.6683 0.0307
EXP. PROP.  0.1254 0.0362 0.1390 0.6690 0.0304
```