A Beta Unfolding Model for Continuous Bounded Responses[1]

Yvonnick Noel

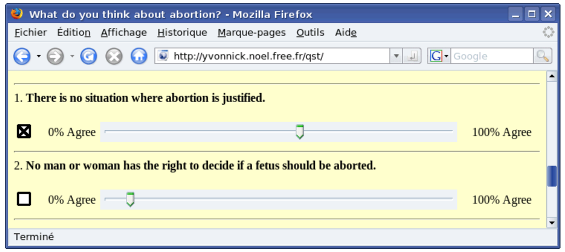
University of Brittany at Rennes (France)

Durham, NH, June 30th 2008

[1]These slides are available for download at: http://yvonnick.noel.free.fr/papiers

Outline

The continuous response format



Some advantages of continuous responses

- Some **interaction effects** may appear with continuous responses, that did not appear with Likert data (Russel & Bobko, 1992).
- Continuous responses seem to be more sensitive to the **mesurement of change** (Pfennings, Cohen & Van Der Ploeg, 1995).
- Grant et al. (1999) report higher **sensitivity** and **reliability** of continuous scores.
- Albaum, Best & Hawkins (1981) argue that continuous scores allow for better between-subject **discrimination**.
- McKelvie (1978) report that subjects tend to find continuous scales **more pleasing** to use.

The interpolation mechanism

- The only reference points in a visual analogue scale are the **labels** at the segment boudaries.
- It is assumed that some **latent values**, v_0 et v_1, are granted to both extreme responses, graphically located at $\lambda_0 = 0$ and $\lambda_1 = 1$ (in arbitrary units).
- It is then assumed (Noël & Dauvier, 2007) that subjects **interpolate** their agreement response x following :

$$x = \frac{\lambda_0 v_0 + \lambda_1 v_1}{v_0 + v_1} = \frac{v_1}{v_0 + v_1}$$

Justification

- **Remark**: This construction garantees that the response lies within $[0; 1]$.
- This hypothetical mechanism is very similar to classical **choice models** in psychology (Luce, 1959), or **discrete choice models** in econometrics (McFadden, 1974), or to the **matching law** in reinforcement learning (Herrnstein, 1961).

A random model

- The "psychological values" are assumed to be **non negative quantities**.
- It is assumed that subjects sample the values from two **Gamma densities** with different shape parameters but a common scale parameter:

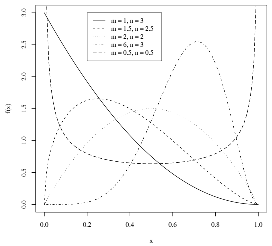
$$v_0 \sim \Gamma(n, s)$$
$$v_1 \sim \Gamma(m, s)$$

- It is known that (Kotz & Johnson, 1982, p.229):

$$X = \frac{v_1}{v_0 + v_1} \sim \beta(m, n)$$

The beta density

$$f(x; m, n) = \frac{\Gamma(m+n)}{\Gamma(m)\Gamma(n)} x^{m-1}(1-x)^{n-1} \text{ for } x \in [0;1], \ m, n > 0$$



A Beta Response Model

- A **beta response model** on the response X_{ij} from subject i ($i = 1, ..., N$) to item j ($j = 1, ..., p$) may be written as:

$$f(x_{ij}; m_{ij}, n_{ij}) = \frac{\Gamma(m_{ij}+n_{ij})}{\Gamma(m_{ij})\Gamma(n_{ij})} x_{ij}^{m_{ij}-1}(1-x_{ij})^{n_{ij}-1} \text{ with } x_{ij} \in [0;1]$$

- m_{ij} et n_{ij} may be interpreted as **acceptance** and **refusal** parameters, respectively.

A structural model

- Distribution parameters are linked to **subject** and **item-specific** parameters by posing:

$$m_{ij} = \exp\left(\frac{\theta_i - \delta_j}{2}\right)$$
$$n_{ij} = \exp\left(-\frac{\theta_i - \delta_j}{2}\right)$$

where θ_i et δ_j are attitude and difficulty parameters, respectively.

- The acceptance parameter varies as a monotone increasing function of attitude and a monotone decreasing function of item difficulty, and conversely for the refusal parameter.

Expected Response Function

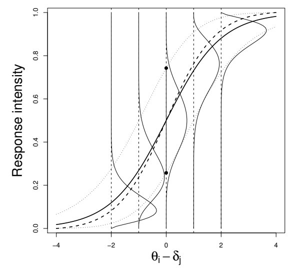
- In a beta model $\beta(m_{ij}, n_{ij})$ the **expectation** reads:

$$\mu_{ij} = E(X_{ij}; m_{ij}, n_{ij}) = \frac{m_{ij}}{m_{ij} + n_{ij}}$$

- The **expected response function** has a familiar form:

$$E(X_{ij}; \theta_i, \delta_j) = \frac{\exp\left(\frac{\theta_i - \delta_j}{2}\right)}{\exp\left(\frac{\theta_i - \delta_j}{2}\right) + \exp\left(-\frac{\theta_i - \delta_j}{2}\right)}$$
$$= \frac{\exp(\theta_i - \delta_j)}{1 + \exp(\theta_i - \delta_j)}$$

Response function



Noel, Y., & Dauvier, B. (2007). A beta response model for continuous bounded responses, *Applied Psychological Measurement*, 31, 47-73.

Ambivalent items

- A sample item from the **attitude towards abortion** questionnaire (Roberts, Donoghue and Laughlin, 2000): "I cannot whole-heartedly support either side of the abortion debate".
- This item may be **refused** for two different reasons (Andrich & Luo, 1993):
  - because you are **in favor** of an unconditional right to abortion,
  - because your are a convinced **opponent** of abortion.

A latent random model

- For such items, we consider **three latent evaluations**, for: Agree (A), Disagree for the first reason (D_1), Disagree for the second reason (D_2).
- The corresponding latent values $v_{ij}^{(A)}$, $v_{ij}^{(D_1)}$ and $v_{ij}^{(D_2)}$, for subject i and item j, are assumed to follow:

$$\begin{aligned} v_{ij}^{(A)} &\sim \Gamma(m_{ij}, s) \\ v_{ij}^{(D_1)} &\sim \Gamma(p_{ij}, s) \\ v_{ij}^{(D_2)} &\sim \Gamma(q_{ij}, s) \end{aligned}$$

A beta model

- A **global disagree value** may be defined as:

$$v_{ij}^{(D)} = v_{ij}^{(D_1)} + v_{ij}^{(D_2)}$$

- From the **properties of the Gamma**, and assuming conditional independence of the latent values, we have:

$$v_{ij}^{(D)} \sim \Gamma(n_{ij}, s) \text{ with } n_{ij} = p_{ij} + q_{ij}$$

- Then from the **interpolation response mechanism**:

$$X_{ij} = \frac{v_{ij}^{(A)}}{v_{ij}^{(D)} + v_{ij}^{(A)}} = \frac{v_{ij}^{(A)}}{v_{ij}^{(D_1)} + v_{ij}^{(D_2)} + v_{ij}^{(A)}} \sim \beta(m_{ij}, p_{ij} + q_{ij})$$

A structural model

- Distribution parameters are connected to **attitude** (θ_i) and **item location** (δ_j) parameters by:

$$\begin{cases} m_j &= \exp \lambda_j \\ n_{ij} &= p_{ij} + q_{ij} \\ &= \exp(\theta_i - \delta_j) + \exp(\delta_j - \theta_i) \end{cases}$$

with λ_j an item-specific acceptance parameter.

- The refusal parameter n_{ij} is the sum of two refusal sources operating **in opposite directions**.

Expected response function

- For this basic model, the **expected response function** is:

$$\begin{aligned} E(X_{ij}|\theta_i) &= \mu_{ij} \\ &= \frac{m_j}{m_j + n_{ij}} \\ &= \frac{m_j}{m_j + p_{ij} + q_{ij}} \\ &= \frac{\exp \lambda_j}{\exp \lambda_j + \exp(\theta_i - \delta_j) + \exp(\delta_j - \theta_i)} \\ &= \frac{\exp \lambda_j}{\exp \lambda_j + 2\cosh(\theta_i - \delta_j)} \end{aligned}$$

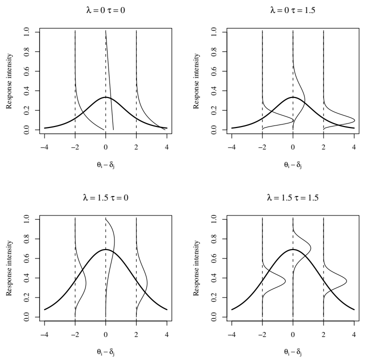
A generalized model

- A more flexible 4-parameter model is obtained by posing:

$$\begin{cases} m_j &= \exp(\lambda_j + \tau_j) \\ n_{ij} &= \exp\{\alpha_j(\theta_i - \delta_j) + \tau_j\} + \exp\{-\alpha_j(\theta_i - \delta_j) + \tau_j\} \end{cases}$$

- Parameter interpretation:
  - λ_j affects the global level of **item acceptation**,
  - α_j is a **discrimination** parameter,
  - τ_j affects response **variance**.

Response expectation and variance

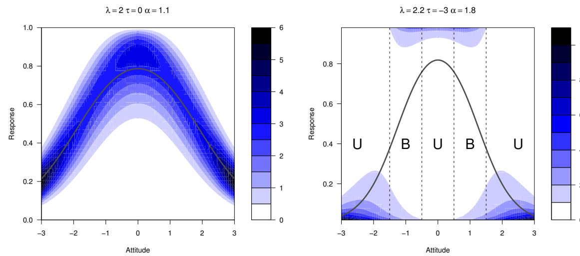


Conditions of bimodality

- The $\beta(m_j, n_{ij})$ density is **bimodal** when $m_j < 1$ et $n_{ij} < 1$.
- Response density is bimodal when **little value is granted to both acceptation and refusal**.
- ... that is, when we have simultaneoulsy $(\lambda_j^* = \lambda_j + \tau_j)$:

$$\begin{cases} \lambda_j^* & < & 0 \\ \tau_j & < & -\ln\left[2\cosh\left(\alpha_j(\theta_i - \delta_j)\right)\right] \end{cases}$$

Unimodal and bimodal items



'U': Unimodal density, 'B': Bimodal density

EM algorithm for finite mixtures

- The attitude variable is **discretized** to a set Θ of fixed θ_k $(k = 1, ..., K)$ values, with probabilities π_k (Woodruf & Hanson, 1996).
- The density of an **oberved response vector** $\mathbf{x}_i$, given the whole set Δ of item parameters, is:

$$f(\mathbf{x}_i|\Delta, \pi) = \sum_k f(\mathbf{x}_i|\theta_k, \Delta)\pi_k$$

- At each step s of the **EM algorithm**, the expectation:

$$Q(\Delta, \pi|\Delta^{(s)}, \pi^{(s)}) = E_\theta \left\{ \ln \left[\prod_{i=1}^{N} f(\mathbf{x}_i, \theta_i|\Delta, \pi) \right] |\mathbf{X}, \Delta^{(s)}, \pi^{(s)} \right\}$$

is computed and maximized with respect to the π_k and δ_j.

Attitude toward abortion

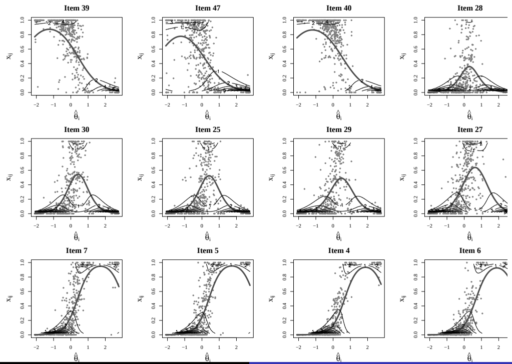
- A 50-item questionnaire on attitude toward abortion (Roberts, Donoghue and Laughlin, 2000; Roberts, Lin & Laughlin, 2001) was submitted to 443 subjects, on a web interface.
- Sample items:
  1. Abortion is a threat to our society (6)
  2. I find myself agreeing with arguments both for and against abortion (27)
  3. I cannot whole-heartedly support either side of the abortion debate (30)
  4. Abortion should be an accepted mechanism for family planning (48)

Concurrent BUM models

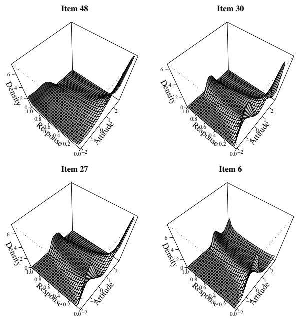
Model	*Parameters*	*Loglik.*	*#Param.*	*#Constr.*	*AIC*
CBUM-2	λ_j, δ_j	-8384.559	128	1	17025.12
UBUM-2	λ_j, δ_j	-8384.559	128	1	17025.12
CBUM-3	$\lambda_j, \delta_j, \tau_j$	14642.29	158	21	-28968.59
UBUM-3	$\lambda_j, \delta_j, \tau_j$	20046.07	178	1	-39736.15
CBUM-4	$\lambda_j, \delta_j, \tau_j, \alpha_j$	16092.28	217	12	-31750.56
UBUM-4	$\lambda_j, \delta_j, \tau_j, \alpha_j$	20436.73	227	2	-40419.47*

'C': Constrained to unimodality, 'U': Unconstrained

Data and model plots



Response surfaces



Concluding remarks

1. Continuous responses have some **natural advantages** over categorical data:
   1. Discriminant,
   2. Easier to plot,
   3. Fewer parameters needed.
2. An unexpected benefit is the **emergence of bimodality** which proves to be psychologically meaningful.
3. More work is needed to test this unusual feature in **other fields** (behavior change, developmental processes...).