Analyzing human perceptions from survey data with Nonlinear CUB models

Wirtschafts Universität Wien, 13th May 2016

Marica Manisera and Paola Zuccolotto
University of Brescia, Italy

Wien - 13th May 2016

Credits

- Manisera M., Zuccolotto P. (2014) Modelling rating data with Nonlinear CUB models, *Computational Statistics and Data Analysis*, 78, 100–118.
- Manisera M., Zuccolotto P. (2014) Modelling "don't know" responses in rating scales. *Pattern Recognition Letters*, 45, 226-234
- Manisera M., Zuccolotto P. (2014). Nonlinear CUB models: the R code. *Statistica & Applicazioni*, XII, 205-223.
- Manisera M., Zuccolotto P. (2015). Identifiability of a model for discrete frequency distributions with a multidimensional parameter space, *Journal of Multivariate Analysis*, 140, 302-316.
- Manisera M., Zuccolotto P. (2015). Visualizing Multiple Results from Nonlinear CUB Models with R Grid Viewports. *Electronic Journal of Applied Statistical Analysis*, 8, 360-373.
- Manisera M., Zuccolotto P. (2016). Treatment of 'don't know' responses in a mixture model for rating data, *Metron*, 74, 99-115.
- Manisera M., Zuccolotto P. (2016). Estimation of Nonlinear CUB models via numerical optimization and EM algorithm, *Communications in Statistics - Simulation and Computation*, *forthcoming.*

Marica Manisera

Paola Zuccolotto

Credits

marica.manisera@unibs.it

paola.zuccolotto@unibs.it

Agenda

- Examples of rating data (real data case studies)
- The unconscious Decision Process (DP) driving individuals' responses on a rating scale
- **CUB models** (D'Elia&Piccolo 2005, *Computational Statistics and Data Analysis* – Iannario&Piccolo 2011, *Modern Analysis of Customer Surveys*)
- **NLCUB models** (Manisera&Zuccolotto 2014, *Computational Statistics and Data Analysis*)

SHAPE: "Statistical Modelling of Human Perception", STAR project - University of Naples Federico II - CUP: E68C13000020003
SYRTO: "SYstemic Risk TOmography: Signals, Measurements, Transmission Channels, and Policy Interventions", grant from the European Union Seventh Framework Programme - Project ID: 320270

Rating data

The analysis of human perception is often carried out by resorting to **surveys** and **questionnaires**, where respondents are asked to **express ratings about the objects being evaluated**.

The goal of the statistical tools proposed for this kind of data is to explicitly **characterize the respondents' perceptions about a latent trait**, by taking into account, at the same time, the **ordinal categorical scale of measurement** of the involved statistical variables.

Rating data – example 1 (superstition)

- A survey investigating confidence about assertions concerned with superstition in Romania
- dataset by Vlăsceanu et al. (2012), downloadable from the IQSS (Institute of Quantitative Social Science) Dataverse Network of the Harvard University
- Respondents (n = 1161) were asked to express a judgment about their degree of belief in some assertions, using a 4-point Likert scale (totally disagree, disagree, agree, totally agree)

Rating data – example 1 (superstition)

1. Evil has red eyes
2. Number 13 brings bad luck
3. If the palm of your left hand itches, you will receive money soon
4. Lucky at cards, unlucky in love
5. If a black cat crosses the street it is a sign of bad luck
6. Zodiacal signs influence nature and personality
7. Human civilization was created by aliens
8. There are some numbers that bring good luck to certain people

Rating data – example 2 (fraud management)

- A survey investigating the perceveid risk of being victim of frauds when using ICT
- dataset supplied by NetConsulting (2013)
- Respondents (n = 116 managers of small, mid-sized and large firms) were asked to express a judgment about their degree of perceived fraud risk when using some different IC Technologies, using a 4-point Likert scale (very low, low, high, very high)

Rating data – example 2 (fraud management)

SOCNET:	Web 2.0 and Social Networks
CLOUD:	Cloud storage and computing
BYOD:	Bring Your Own Device
LEG:	Legacy technologies

Rating data – example 3 (Standard Eurobarometer 81)

European Commission

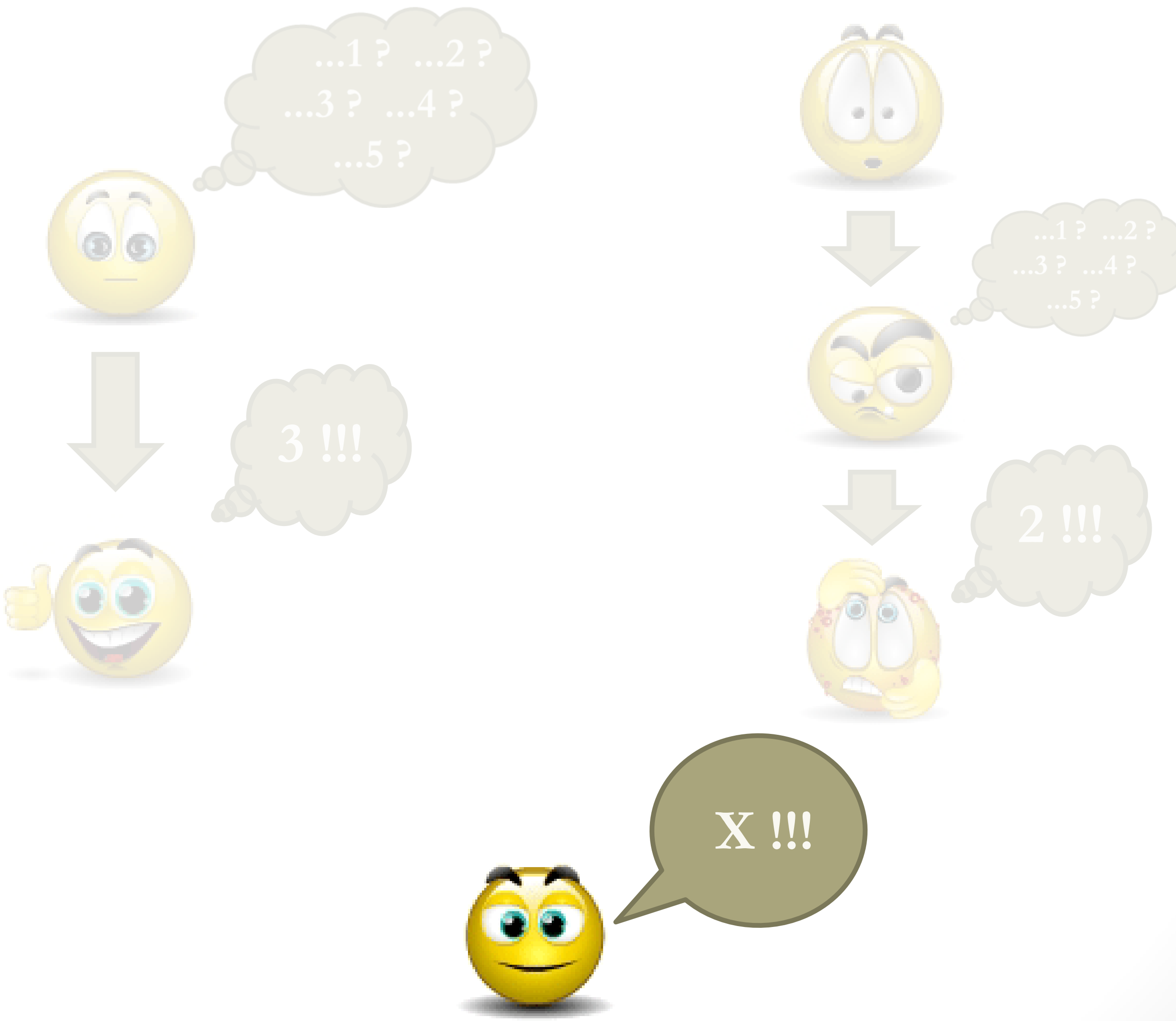
- A sample survey covering the national population of citizens of the 27 European Union Member States
- Questions asking respondents to rate their level of agreement with some statements using a 4-point Likert scale (totally disagree, tend to disagree, tend to agree, totally agree)
- “don’t know” option available

Rating data – example 3 (Standard Eurobarometer 81)

QA19.1:	I understand how the EU works
QA19.2:	Globalisation is an opportunity for economic growth
QA19.3:	(OUR COUNTRY) could better face the future outside the EU
QA19.4:	The EU should develop further into a federation of nation states
QA19.5:	More decisions should be taken at EU level
QA19.6:	We need a united Europe in today's world

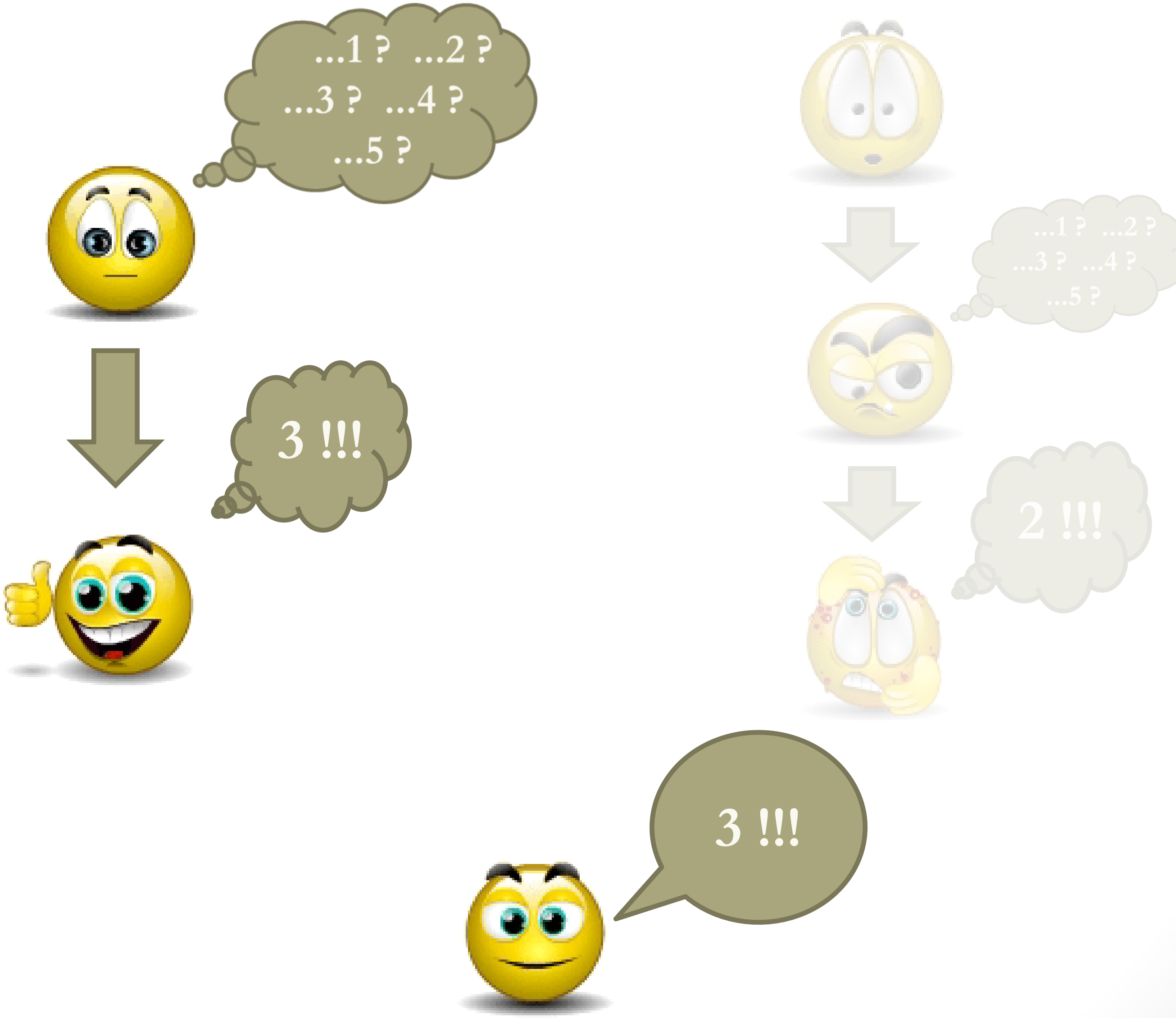
Are you satisfied with XYZ?

Express a rating from 1 (=completely unsatisfied) to 5 (=completely satisfied)



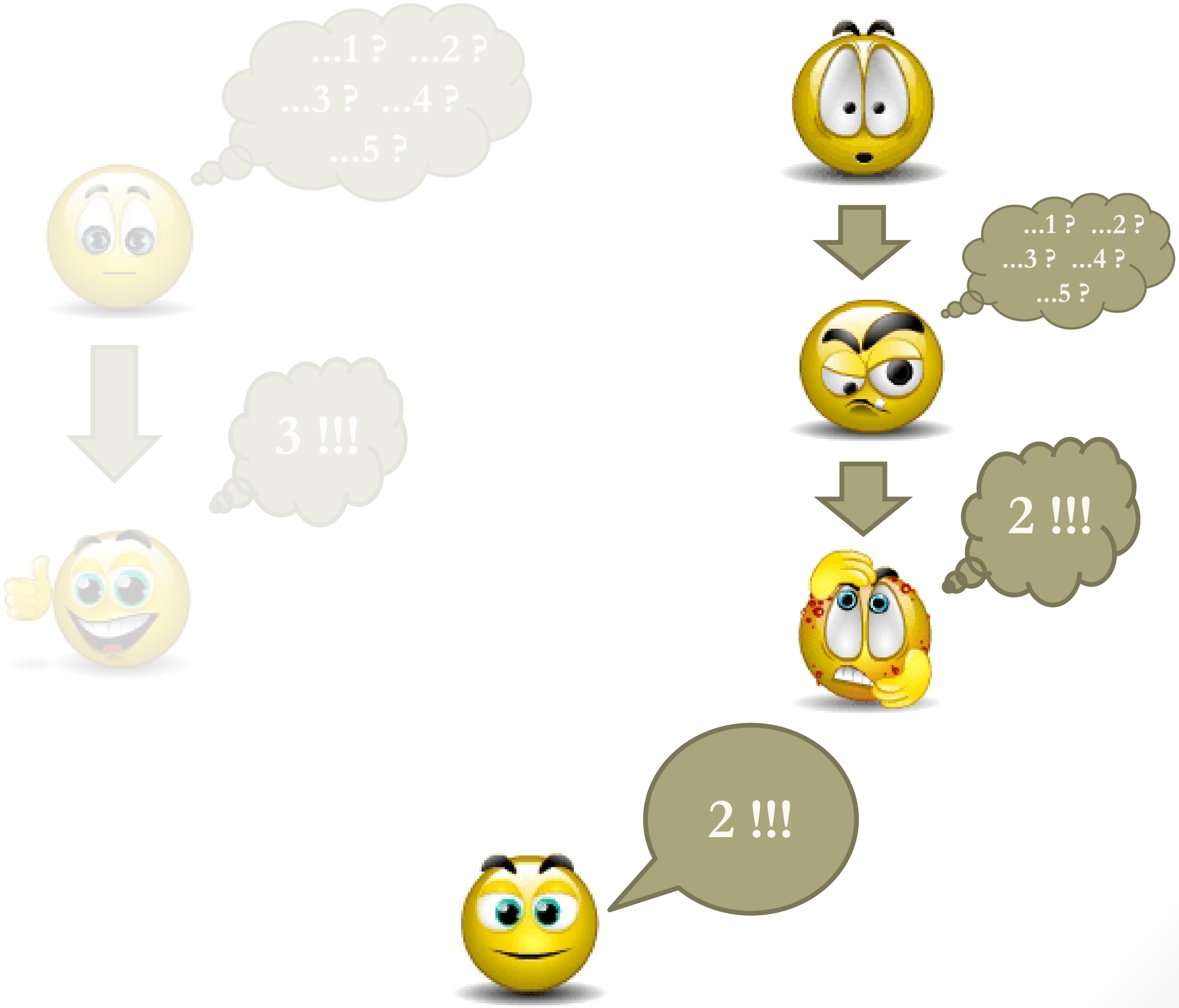
Are you satisfied with XYZ?

Express a rating from 1 (=completely unsatisfied) to 5 (=completely satisfied)



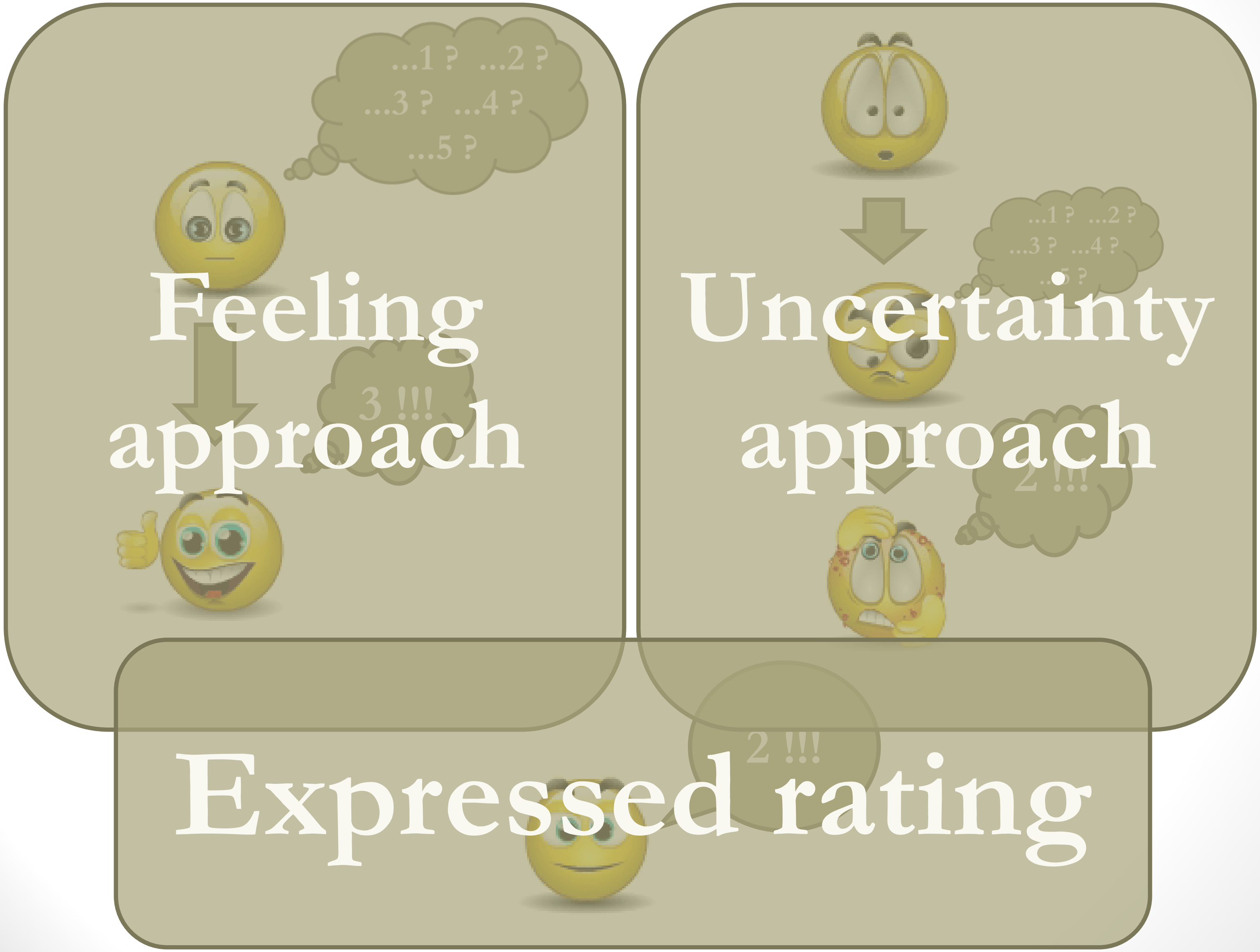
Are you satisfied with XYZ?

Express a rating from 1 (=completely unsatisfied) to 5 (=completely satisfied)



Are you satisfied with XYZ?

Express a rating from 1 (=completely unsatisfied) to 5 (=completely satisfied)



Are you satisfied with XYZ?

Express a rating from 1 (=completely unsatisfied) to 5 (=completely satisfied)

reasoned and logical thinking, the set of emotions, perceptions, subjective evaluations that individuals have with regard to the latent trait being evaluated

indecision inherently present in any human choice, not depending on the individuals' position on the latent variable

Expressed rating

Are you satisfied with XYZ?

How do CUB models fit into this framework?

Express a rating from 1 (=completely unsatisfied) to 5 (=completely satisfied)

Feeling approach CUB: (shifted) Binomial random variable (V)

$$b_r(\xi) = P(V = r) = \binom{m-1}{r-1} \xi^{m-r} (1-\xi)^{r-1}$$

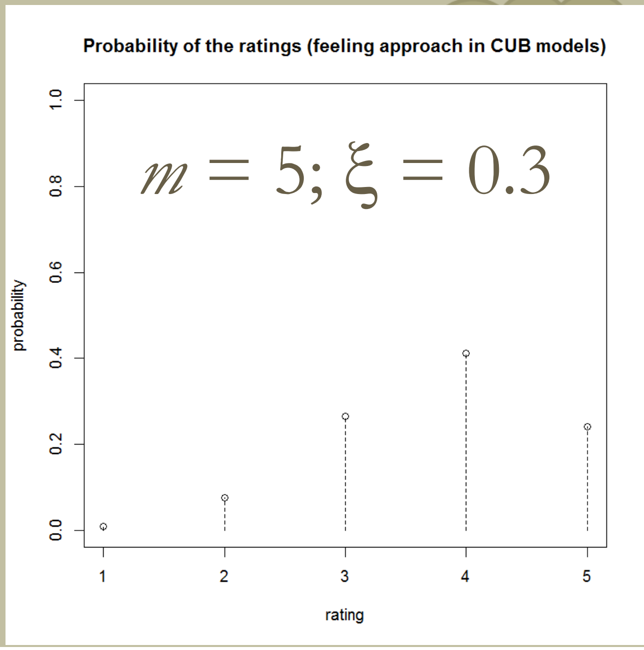
indecision inherently present in any human choice, not depending on the individuals' position on the latent variable

Expressed rating

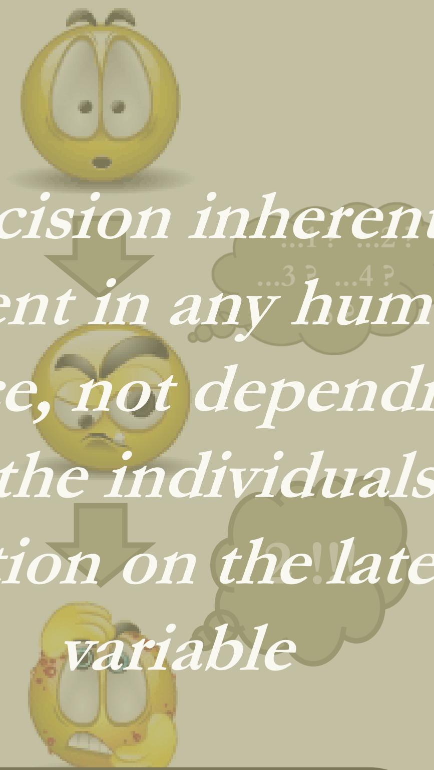
Are you satisfied with XYZ?

How do CUB models fit into this framework?

Express a rating from 1 (=completely unsatisfied) to 5 (=completely satisfied)



$$b_r(\xi) = P(V = r) = \binom{m-1}{r-1} \xi^{m-r} (1-\xi)^{r-1}$$



indecision inherently present in any human choice, not depending on the individuals' position on the latent variable

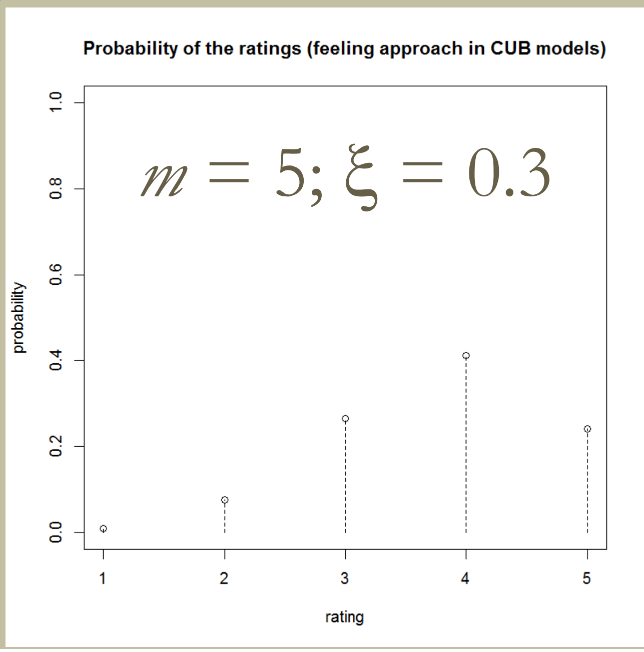
Expressed rating

CUB models

Are you satisfied with XYZ?

How do CUB models fit into this framework?

Express a rating from 1 (=completely unsatisfied) to 5 (=completely satisfied)



$$b_r(\xi) = P(V = r) = \binom{m-1}{r-1} \xi^{m-r} (1-\xi)^{r-1}$$

Uncertainty approach CUB: Uniform random variable (U)

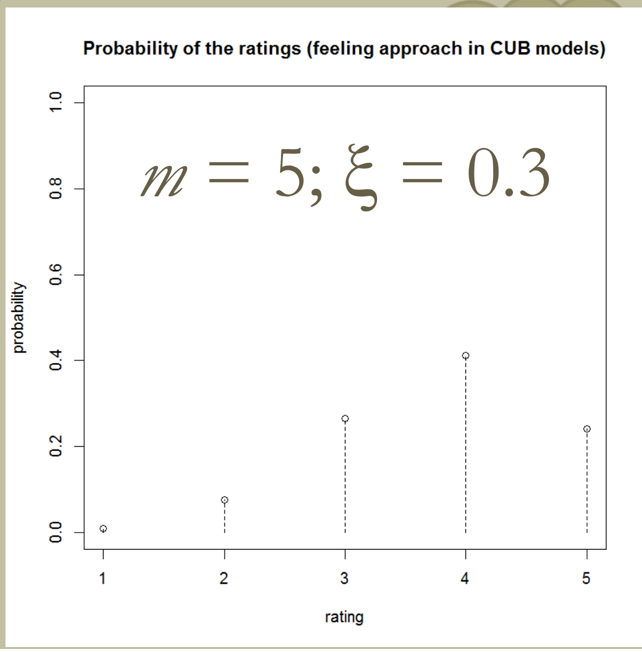
$$P(U = r) = 1/m$$

Expressed rating

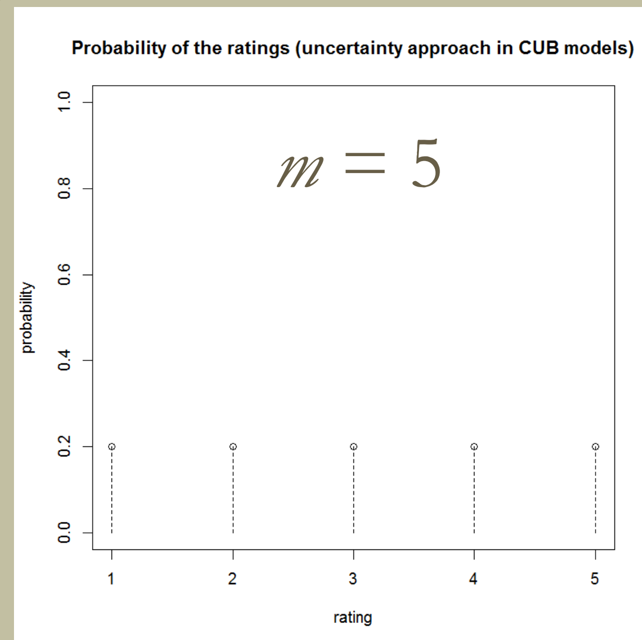
Are you satisfied with XYZ?

How do CUB models fit into this framework?

Express a rating from 1 (=completely unsatisfied) to 5 (=completely satisfied)



$$b_r(\xi) = P(V = r) = \binom{m-1}{r-1} \xi^{m-r} (1-\xi)^{r-1}$$



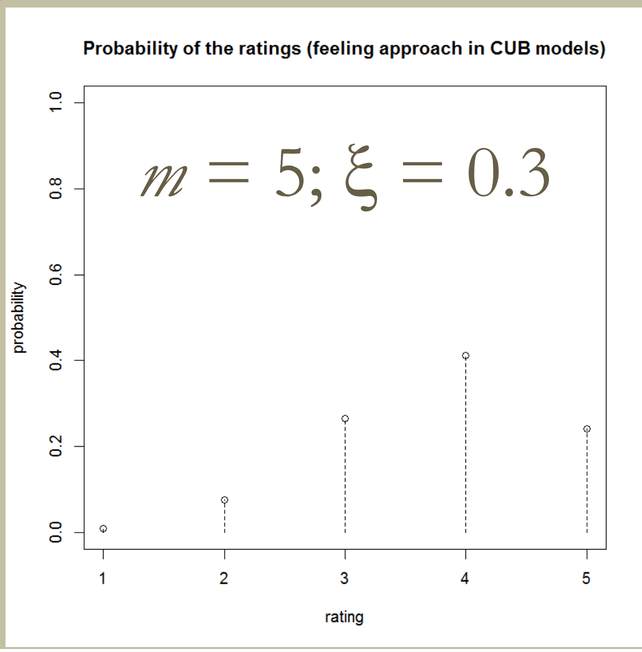
$$P(U = r) = 1/m$$

Expressed rating

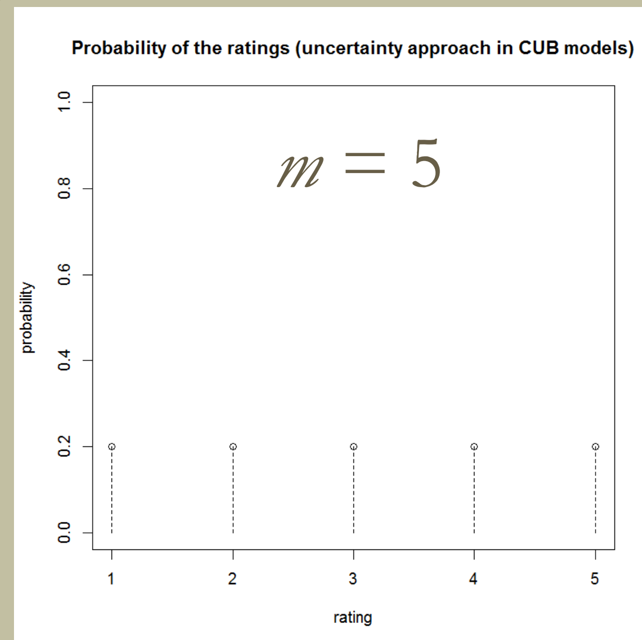
Are you satisfied with XYZ?

How do CUB models fit into this framework?

Express a rating from 1 (=completely unsatisfied) to 5 (=completely satisfied)



$$b_r(\xi) = P(V = r) = \binom{m-1}{r-1} \xi^{m-r} (1-\xi)^{r-1}$$



$$P(U = r) = 1/m$$

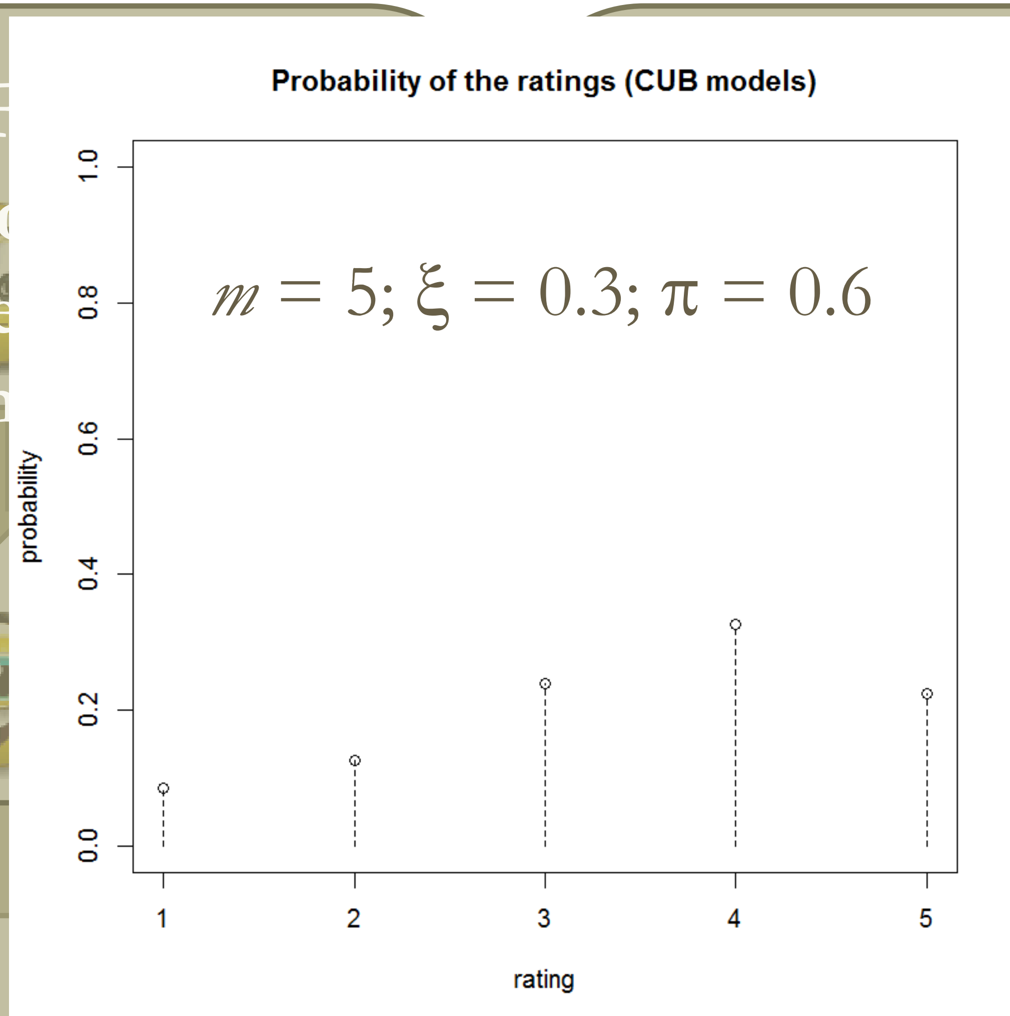
Expressed rating CUB: mixture of V and U (R)

$$P(R = r \mid \boldsymbol{\theta}) = \pi\, b_r(\xi) + (1-\pi)\, P(U = r)$$

Are you satisfied with XYZ?

How do CUB models fit into this framework?

Express a rating from 1 (=completely unsatisfied) to 5 (=completely satisfied)



$$P(R = r \mid \boldsymbol{\theta}) = \pi\, b_r(\xi) + (1-\pi)\, P(U = r)$$

Are you satisfied with XYZ?

How do CUB models fit into this framework?

Express a rating from 1 (=completely unsatisfied) to 5 (=completely satisfied)

Feeling approach CUB: (shifted) Binomial random variable (V)

Feeling parameter:
1 - ξ

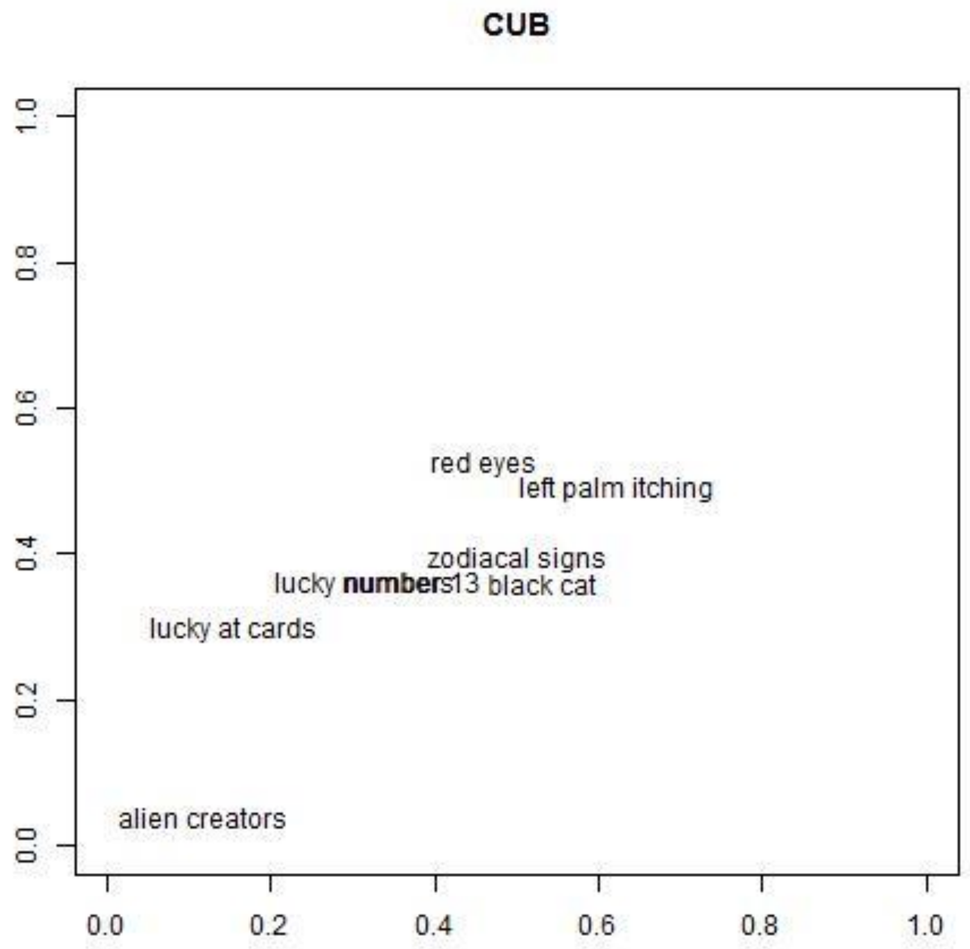
Uncertainty approach CUB: Uniform random variable (U)

Uncertainty parameter:
1 - π

Expressed rating CUB: mixture of V and U (R)

$$P(R = r \mid \boldsymbol{\theta}) = \pi\, b_r(\xi) \;+\; (1-\pi)\, P(U = r)$$

Example 1 (superstition)



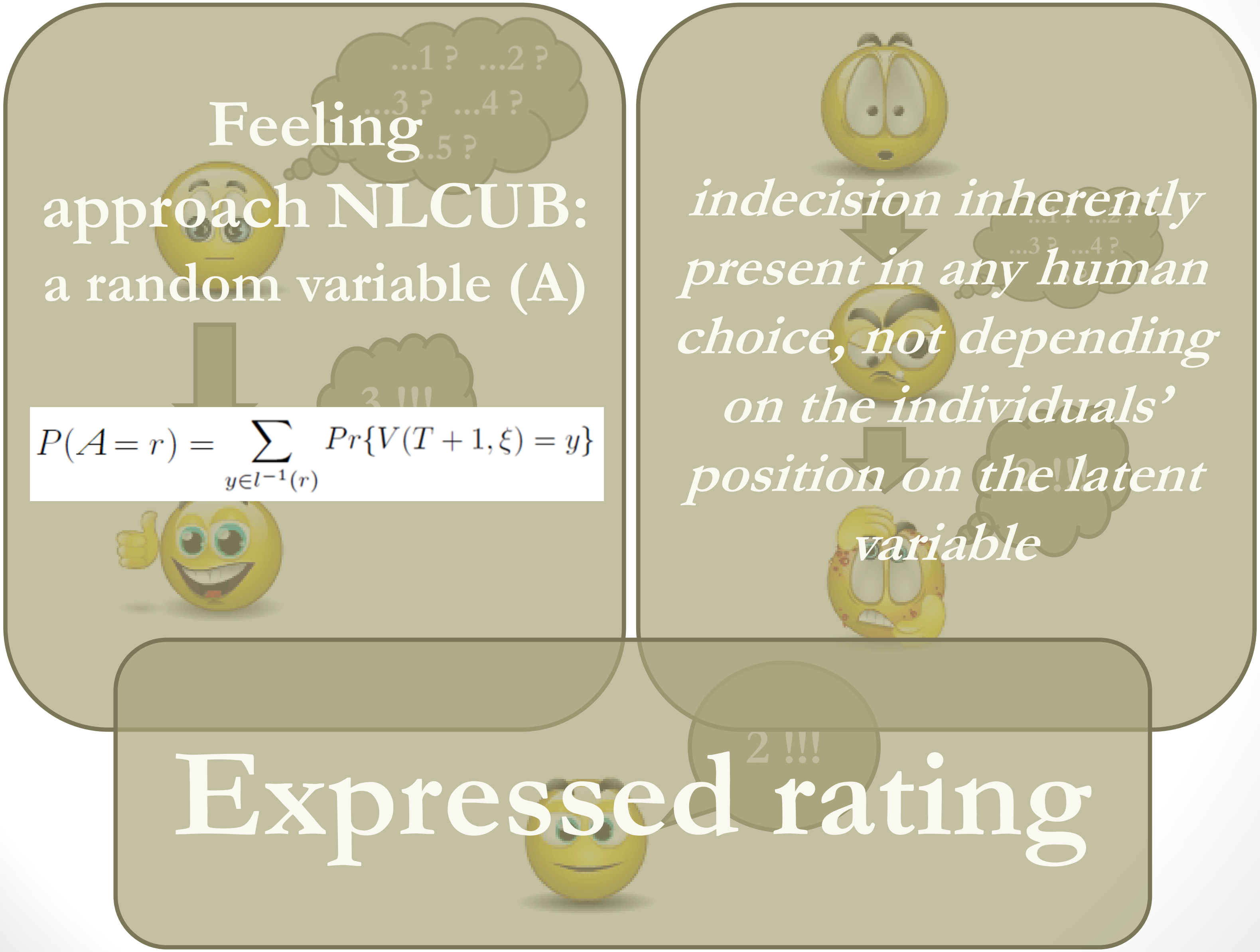
Feeling parameter:
$1 - \xi$

Uncertainty parameter:
$1 - \pi$

Are you satisfied with XYZ?

How do NLCUB models fit into this framework?

Express a rating from 1 (=completely unsatisfied) to 5 (=completely satisfied)



Feeling approach NLCUB:
a random variable (A)

$$P(A = r) = \sum_{y \in l^{-1}(r)} Pr\{V(T+1, \xi) = y\}$$

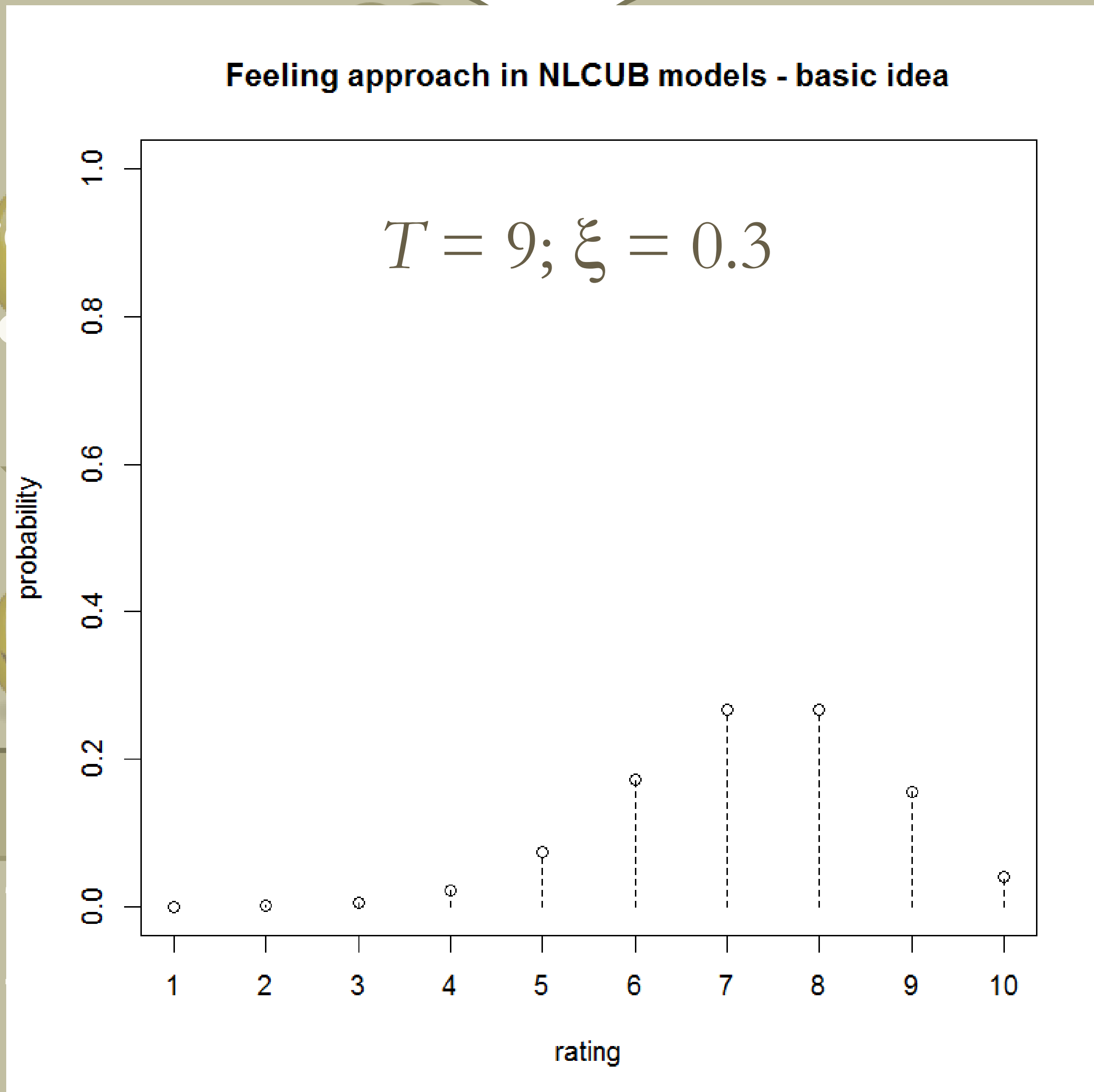
indecision inherently present in any human choice, not depending on the individuals' position on the latent variable

Expressed rating

Are you satisfied with XYZ?

How do NLCUB models fit into this framework?

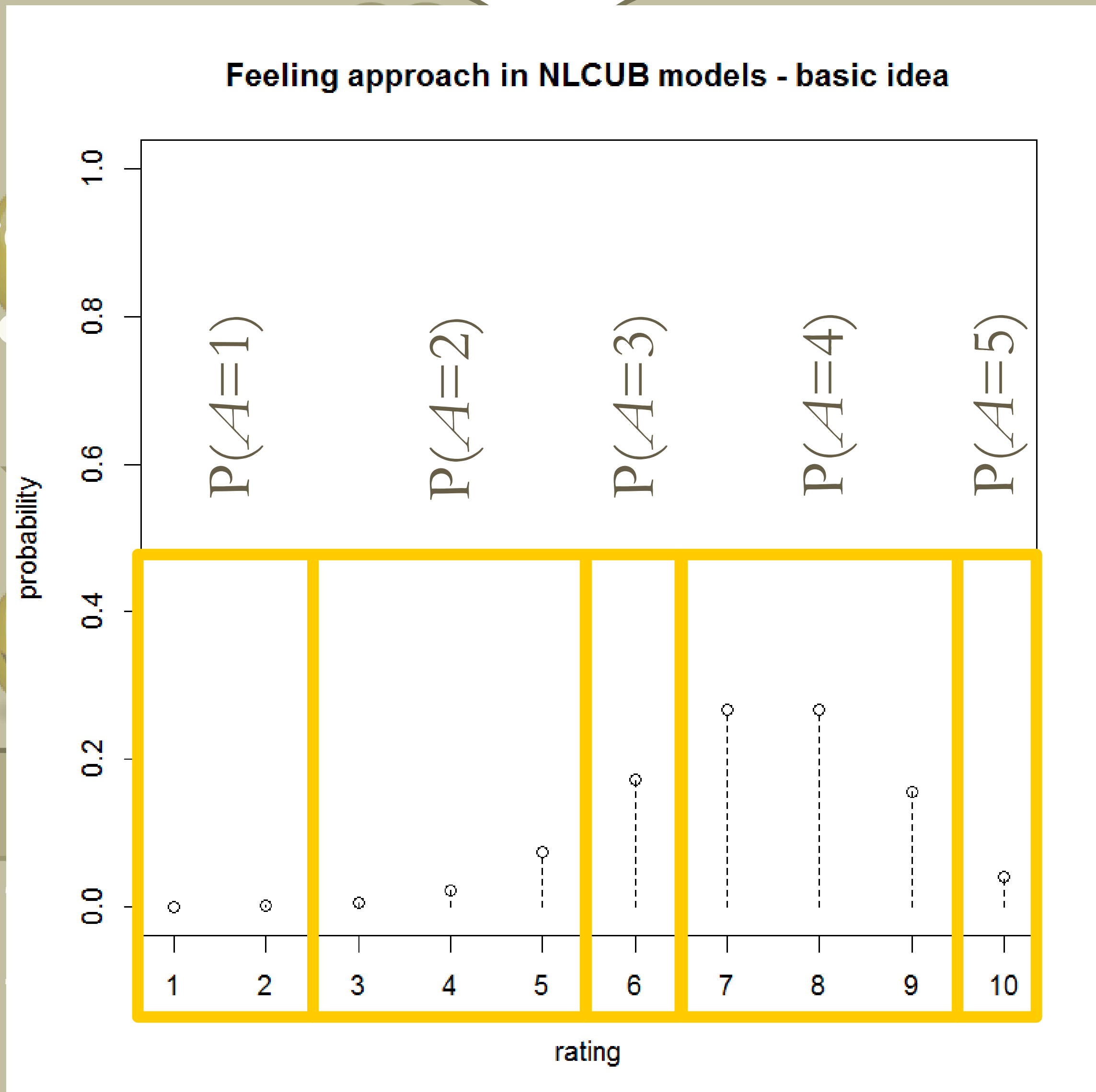
Express a rating from 1 (=completely unsatisfied) to 5 (=completely satisfied)



Are you satisfied with XYZ?

How do NLCUB models fit into this framework?

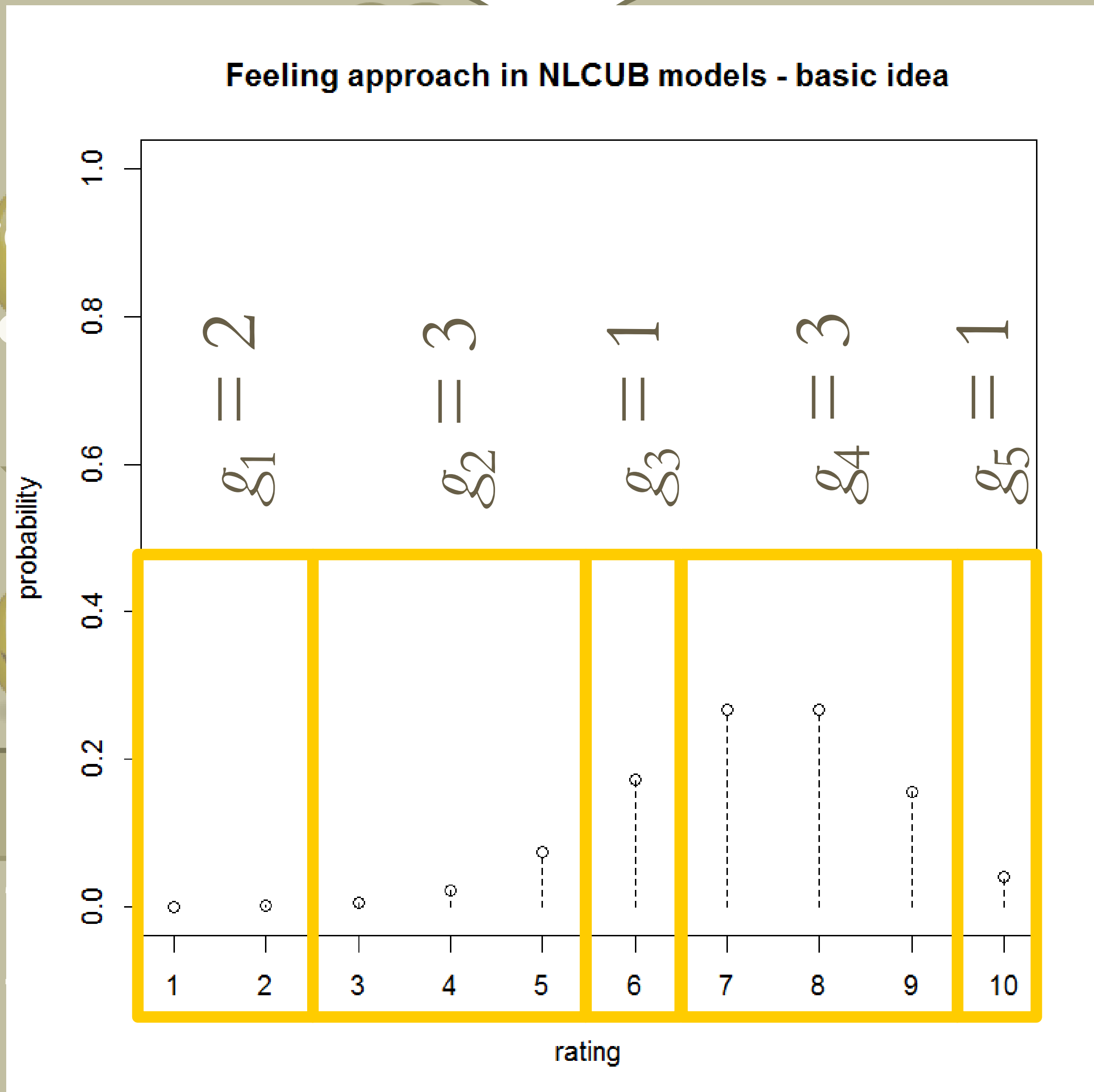
Express a rating from 1 (=completely unsatisfied) to 5 (=completely satisfied)



Are you satisfied with XYZ?

How do NLCUB models fit into this framework?

Express a rating from 1 (=completely unsatisfied) to 5 (=completely satisfied)

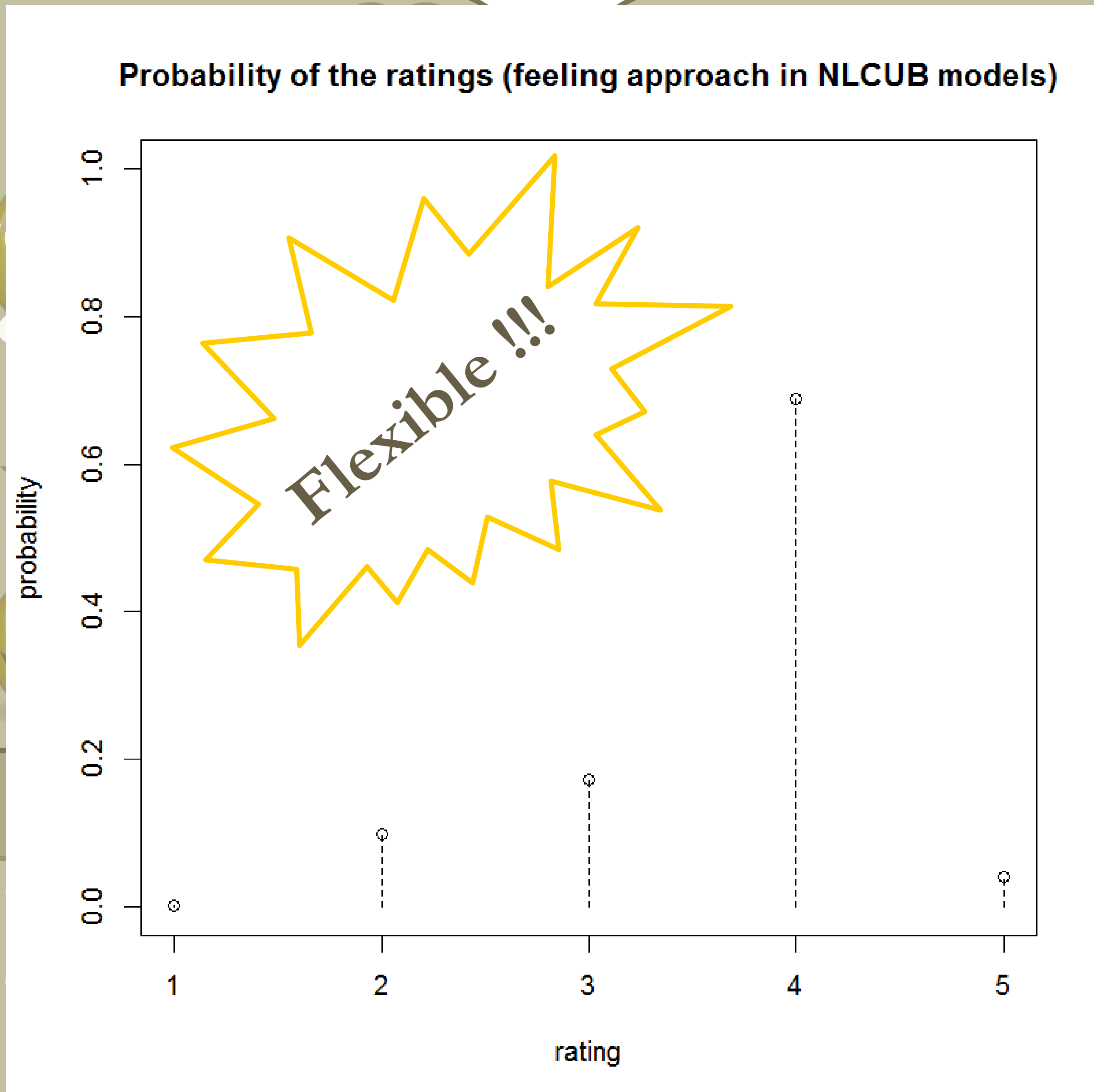


NLCUB models

Are you satisfied with XYZ?

How do NLCUB models fit into this framework?

Express a rating from 1 (=completely unsatisfied) to 5 (=completely satisfied)

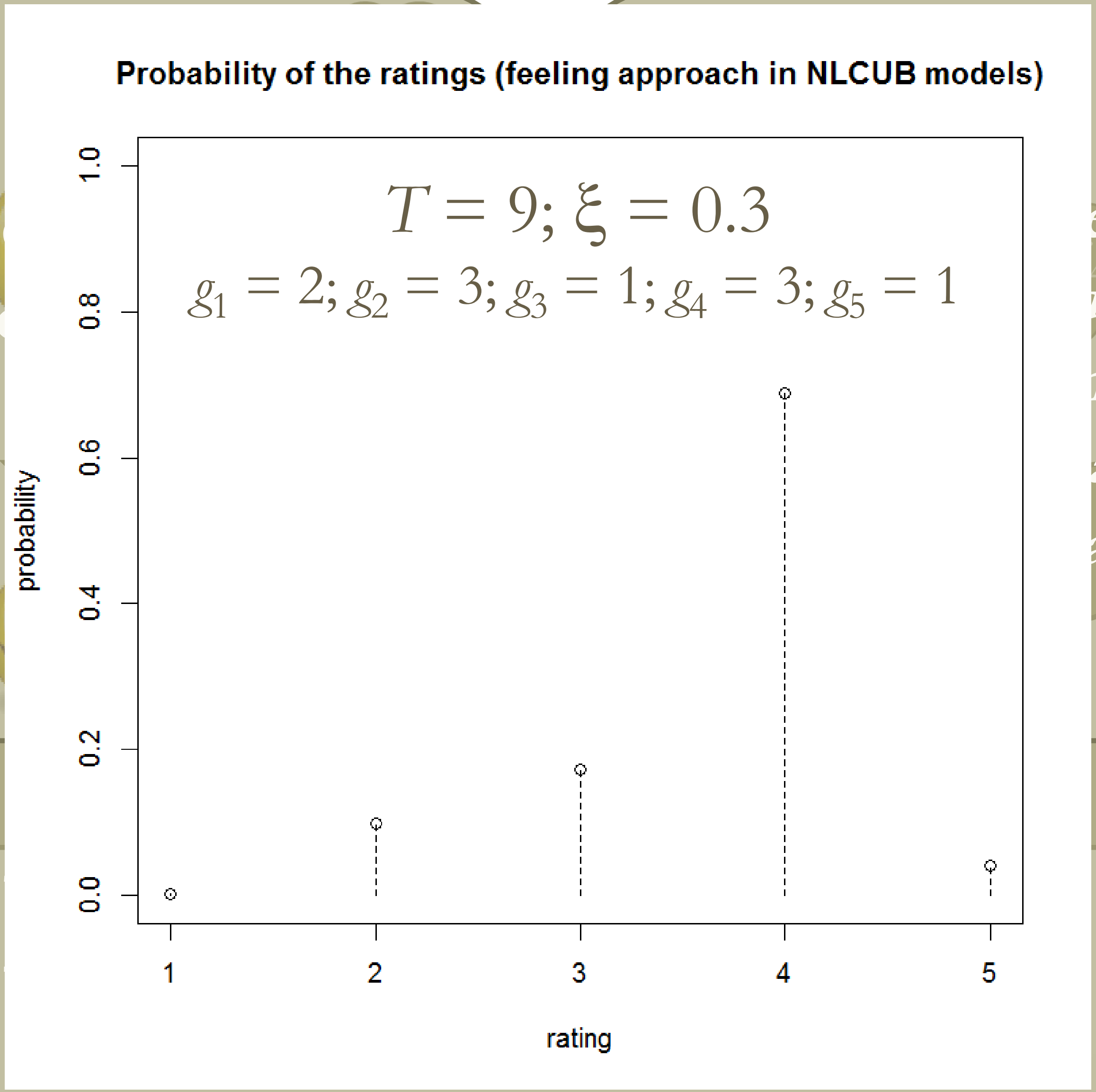


NLCUB models

Are you satisfied with XYZ?

How do NLCUB models fit into this framework?

Express a rating from 1 (=completely unsatisfied) to 5 (=completely satisfied)

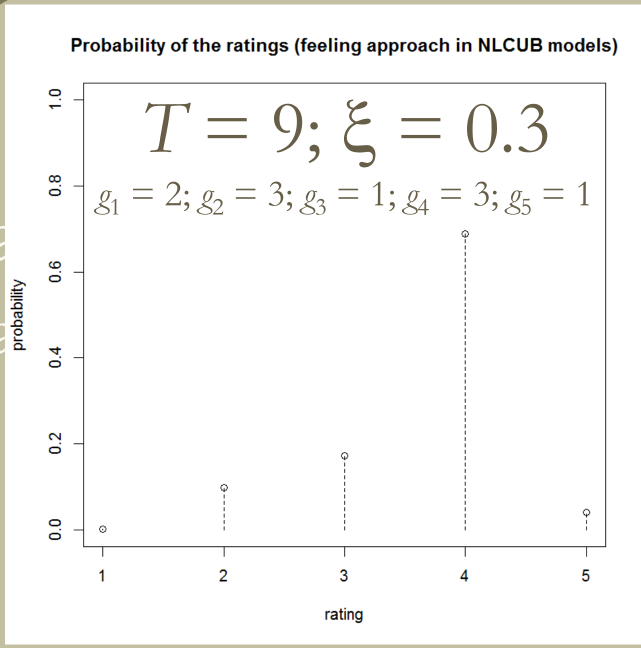


NLCUB models

Are you satisfied with XYZ?

How do NLCUB models fit into this framework?

Express a rating from 1 (=completely unsatisfied) to 5 (=completely satisfied)



$$P(A=r) = \sum_{y \in l^{-1}(r)} Pr\{V(T+1,\xi)=y\}$$

Uncertainty approach NLCUB: Uniform random variable (U)

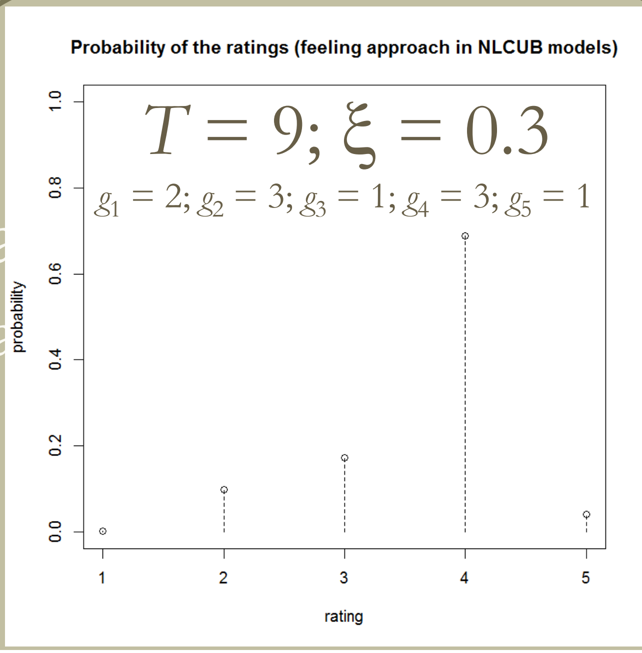
$$P(U=r) = 1/m$$

Expressed rating

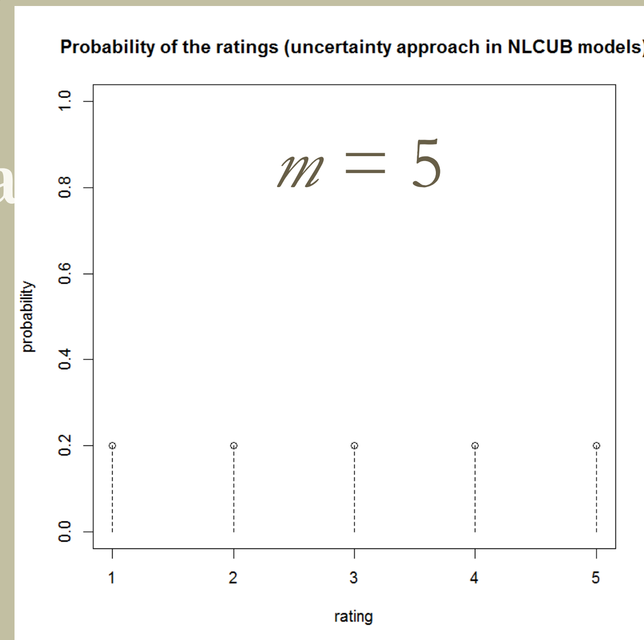
Are you satisfied with XYZ?

How do NLCUB models fit into this framework?

Express a rating from 1 (=completely unsatisfied) to 5 (=completely satisfied)



$$P(A = r) = \sum_{y \in l^{-1}(r)} Pr\{V(T+1, \xi) = y\}$$



$$P(U = r) = 1/m$$

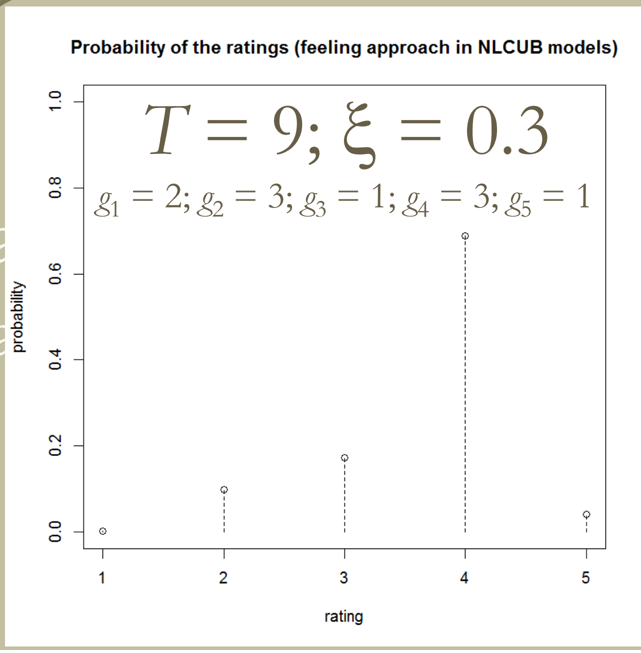
Expressed rating

NLCUB *models*

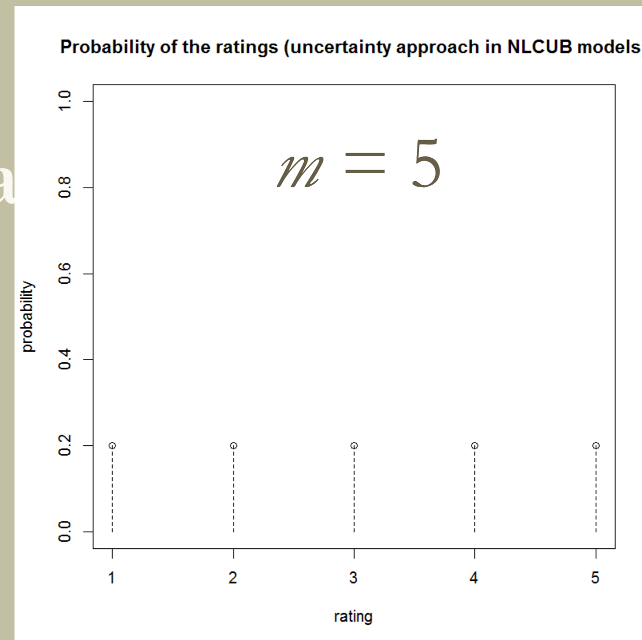
Are you satisfied with XYZ?

How do NLCUB models fit into this framework?

Express a rating from 1 (=completely unsatisfied) to 5 (=completely satisfied)



$$P(A=r) = \sum_{y \in l^{-1}(r)} Pr\{V(T+1,\xi) = y\}$$



$$P(U=r) = 1/m$$

Expressed rating NLCUB: mixture of A and U (R)

$$P(R=r \mid \boldsymbol{\theta}) = \pi\, P(A=r) + (1-\pi)\, P(U=r)$$

NLCUB models

Are you satisfied with XYZ?

How do NLCUB models fit into this framework?

Express a rating from 1 (=completely unsatisfied) to 5 (=completely satisfied)

F

approa

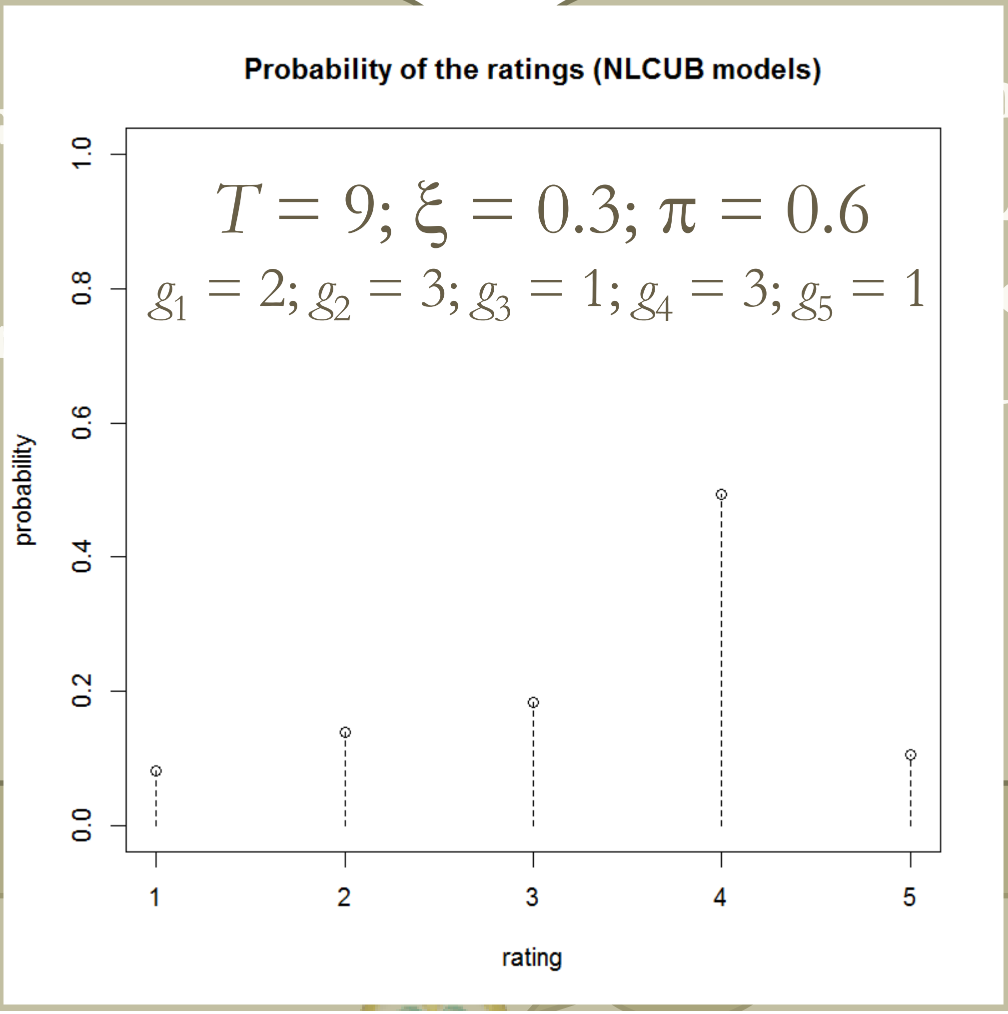
a rando

nty

LCUB:

dom

U)



$$P(R = r \mid \boldsymbol{\theta}) = \pi\, P(A = r) + (1 - \pi)\, P(U = r)$$

NLCUB models

Are you satisfied with XYZ?

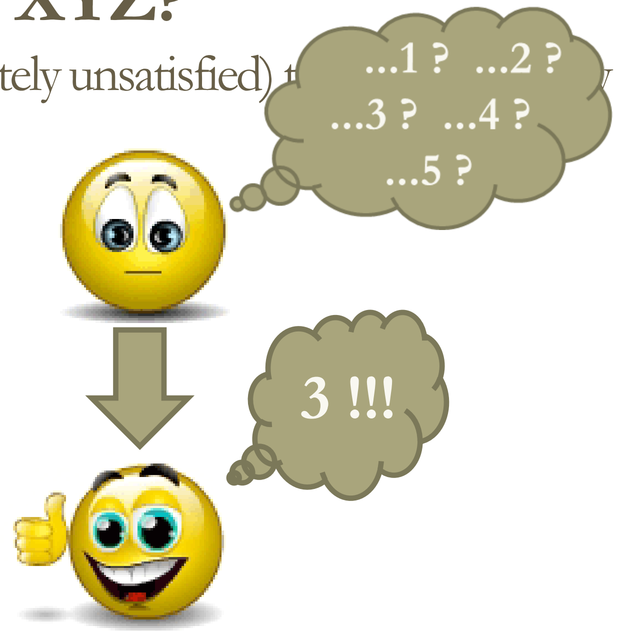
Express a rating from 1 (=completely unsatisfied) to 5 (=completely satisfied)

Let us focus on the Feeling approach

The Feeling approach

Are you satisfied with XYZ?

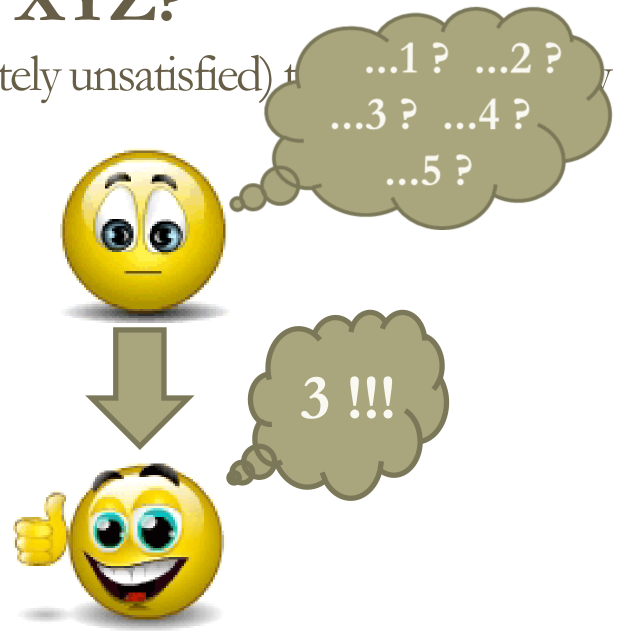
Express a rating from 1 (=completely unsatisfied) t… satisfied)



- We assume that the Feeling approach proceeds through T consecutive steps.
- At each step a basic judgment is formulated.
- Step-by-step, the basic judgments are accumulated and transformed into provisional ratings.
- The rating at the end of the Feeling approach is given by the last provisional rating.

Are you satisfied with XYZ?

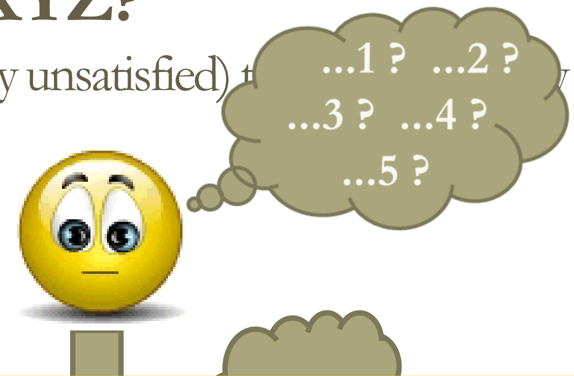
Express a rating from 1 (=completely unsatisfied) t… satisfied)



- We assume that the Feeling approach proceeds through T consecutive steps.
- At each step a basic judgment is formulated.
- Step-by-step, the basic judgments are accumulated and transformed into provisional ratings.
- The rating at the end of the Feeling approach is given by the last provisional rating.

Are you satisfied with XYZ?

Express a rating from 1 (=completely unsatisfied) t… y satisfied)

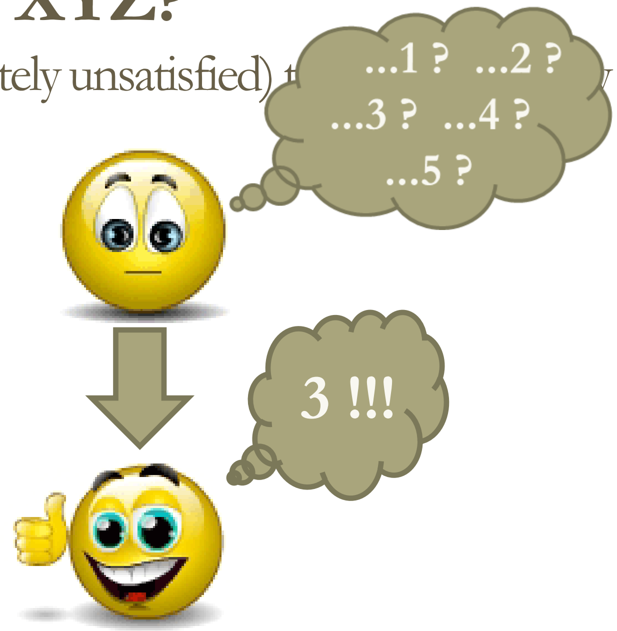


An evaluation about the latent trait, but **a simpler task** than the rating expression

- We assume th… eeling approach proceeds through T consecu…e steps.
- At each step a basic judgment is formulated.
- Step-by-step, the basic judgments are accumulated and transformed into provisional ratings.
- The rating at the end of the Feeling approach is given by the last provisional rating.

Are you satisfied with XYZ?

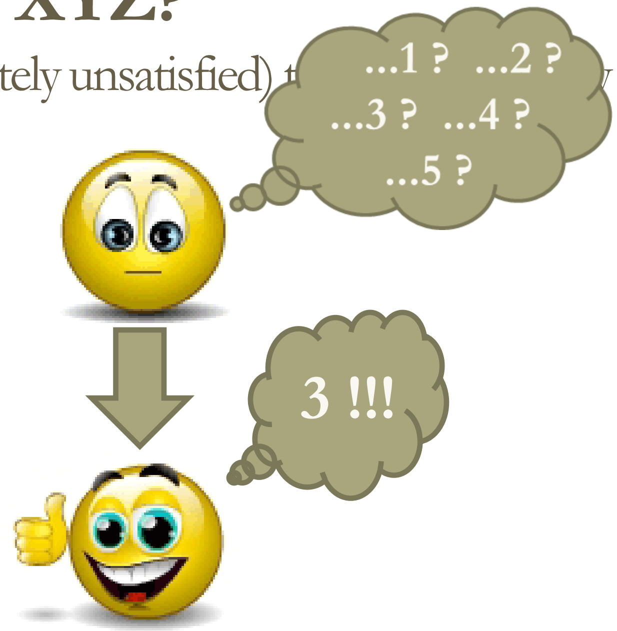
Express a rating from 1 (=completely unsatisfied) t ... y satisfied)



- We assume that the Feeling approach proceeds through T consecutive steps.
- At each step a basic judgment is formulated.
- Step-by-step, the basic judgments are accumulated and transformed into provisional ratings.
- The rating at the end of the Feeling approach is given by the last provisional rating.

Are you satisfied with XYZ?

Express a rating from 1 (=completely unsatisfied) t y satisfied)



- We assume that the Feeling approach proceeds through T consecutive steps.
- At each step a basic judgment is formulated.
- Step-by-step, the basic judgments are accumulated and transformed into provisional ratings.
- The rating at the end of the Feeling approach is given by the last provisional rating.

Are you satisfied with XYZ?

Express a rating from 1 (=completely unsatisfied) to 5 (=completely satisfied)

We can obtain **several different models**, depending on the assumptions we make about:

- the **distribution** of the basic judgments
- the **accumulation** function
- the **transformation** function

- We assume that the Feeling approach proceeds through T consecutive steps.
- At each step a basic judgment is formulated.
- Step-by-step, the basic judgments are accumulated and transformed into provisional ratings.
- The rating at the end of the Feeling approach is given by the last provisional rating.

Are you satisfied with XYZ?

Express a rating from 1 (=completely unsatisfied) to 5 (=completely satisfied)

We can obtain **several different models**, depending on the assumptions we make about:

- the **distribution** of the basic judgments
- the **accumulation** function
- the **transformation** function

Both CUB and NLCUB models can be derived following this paradigm, when some **specific assumptions (... ... ...) are made about these three points**

A) FEELING APPROACH

1. *Elementary judgments:* An *iid* sequence of random variables $X_1, \cdots, X_T$ with domains $\mathcal{D}_{X_1}, \cdots, \mathcal{D}_{X_T}$ generates T elementary judgments $x_1, \cdots, x_T$ progressively expressed along T steps.

2. *Accumulating function:* At each step t, a function $f : \mathcal{D}_{X_1} \times \cdots \times \mathcal{D}_{X_t} \to \Psi_t \subseteq \mathbb{R}$ summarizes the t past elementary judgments (for example, by summation). We say that f is an accumulating function, i.e. we require it obeys the following property: $\Psi_t \subseteq \Psi_{t+1}, \forall t$.

3. *Accumulated judgments:* A sequence of random variables $W_1, \cdots, W_T$, $W_t = f(X_1, \cdots, X_t)$, with domains $\mathcal{D}_{W_1} \equiv \Psi_1, \cdots, \mathcal{D}_{W_T} \equiv \Psi_T$ is then originated along the T steps of the DP with T corresponding realizations $w_1, \cdots, w_T$, $w_t = f(x_1, \cdots, x_t)$, called accumulated judgments.

4. *'Likertization' function:* At each step t, a non-decreasing function $d : \mathcal{D}_{W_T} \to (1, \cdots, m)$ transforms w_t into a provisional rating. Note that from the definition of accumulating function derives $\mathcal{D}_{W_1} \subseteq \cdots \subseteq \mathcal{D}_{W_T}$, so that d can always be computed on the domain of W_t, for all t.

5. *Provisional ratings:* A sequence of random variables $R_1, \cdots, R_T$, $R_t = d(W_t)$, with domains the space $(1, \cdots, m)$ is then originated along the T steps of the feeling path with T corresponding realizations $r_1, \cdots, r_T$, $r_t = d(w_t)$, called provisional ratings.

Which is the advantage of fitting CUB and NLCUB models into this paradigm?

TRANSITION PROBABILITIES

The probability of increasing one (provisional) rating point in the next step of the decision process

$$\phi_t(s) = Pr(R_{t+1} = s+1 | R_t = s)$$

$$\phi_t(s) = \frac{\sum_{w_t \in d^{-1}(s)} Pr(\underline{x}(s) < X_{t+1} \leq \overline{x}(s) | W_t = w_t) Pr(W_t = w_t)}{\sum_{w_t \in d^{-1}(s)} Pr(W_t = w_t)}$$

with $t : \mathcal{D}_{W_t} \cap d^{-1}(s) \neq \emptyset$, $t < T$, where $\underline{x}(s) = \max\{d^{-1}(s)\} - w_t$ and $\overline{x}(s) = \max\{d^{-1}(s+1)\} - w_t$. In order to consider also what happens during the first step of the DP, we define $w_0 := 0$ and $\phi_0 = \phi_0(s) := Pr(\underline{x}(s) < X_1 \leq \overline{x}(s))$ with $s = d(w_0) = d(0)$.

Which is the advantage of fitting CUB and NLCUB models into this paradigm?

TRANSITION PROBABILITIES

The probability of increasing one (provisional) rating point in the next step of the decision process

$$\phi_t(s) = Pr(R_{t+1} = s+1 | R_t = s)$$

In CUB models:

$$\phi_t(s) = 1 - \xi$$

for all *t* and *s*

Which is the advantage of fitting CUB and NLCUB models into this paradigm?

TRANSITION PROBABILITIES

The probability of increasing one (provisional) rating point in the next step of the decision process

$$\phi_t(s) = Pr(R_{t+1} = s+1 | R_t = s)$$

In NLCUB models:

$$\phi_t(s) = (1-\xi)\frac{\binom{t}{w_{g_s s}}(1-\xi)^{w_{g_s s}}\xi^{t-w_{g_s s}}}{\sum_{h=1}^{g_s}\binom{t}{w_{hs}}(1-\xi)^{w_{hs}}\xi^{t-w_{hs}}}$$

Which is the advantage of fitting CUB and NLCUB models into this paradigm?

TRANSITION PROBABILITIES

The probability of increasing one (provisional) rating point in the next step of the decision process

$$\phi_t(s) = Pr(R_{t+1} = s+1 | R_t = s)$$

In NLCUB models:

Different values for different *t* and *s*

TRANSITION PROBABILITIES

$$\phi_t(s) = Pr(R_{t+1} = s+1 | R_t = s)$$

$$\boxed{\phi(s) = av_t(\phi_t(s))}$$

"Perceived closeness" between rating *s* and *s* + 1

$$\delta_s = h(\phi(s)) \quad \text{for example} \quad \delta_s = -\log(\phi(s))$$

"Perceived distance" between rating *s* and *s* + 1

TRANSITION PLOT

A broken line joining points $(s, \tilde{\phi}(s))$, where

$$s = 0, \cdots, m-1$$

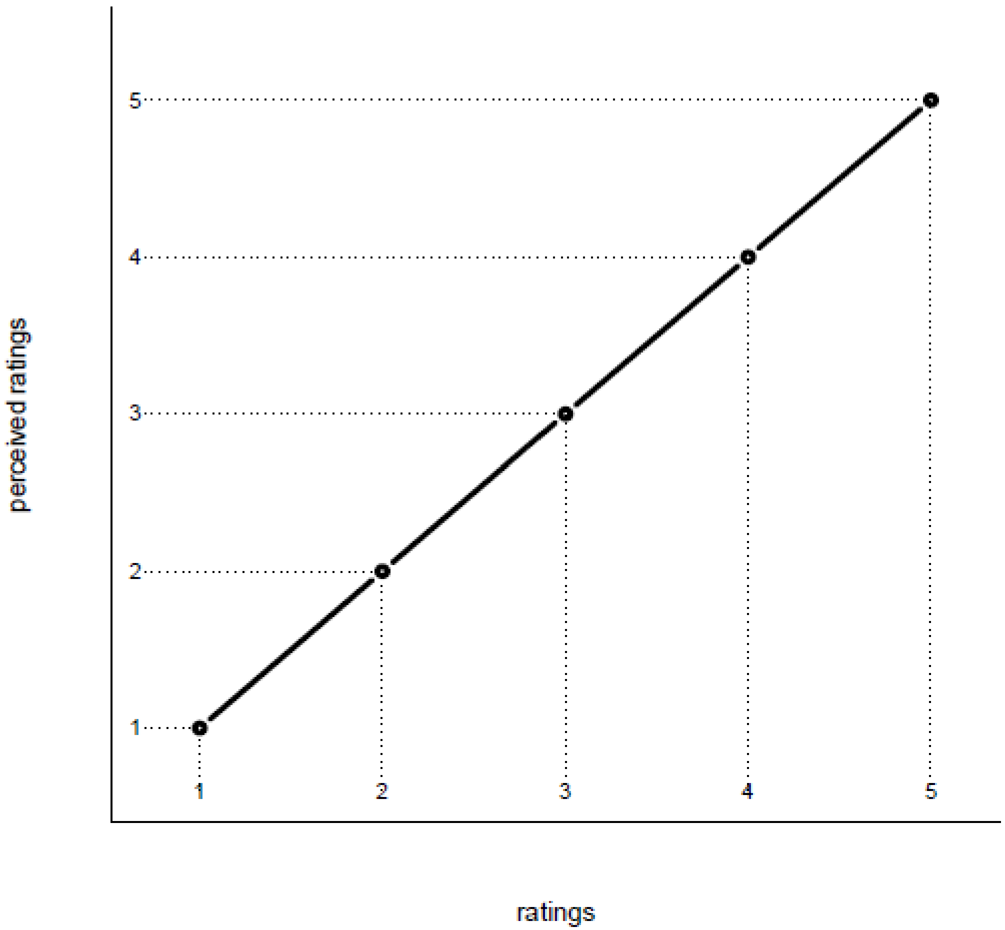
$$\tilde{\phi}(s) = (\delta_1 + \cdots + \delta_s)/(\delta_1 + \cdots + \delta_{m-1})$$

i.e.: the **cumulated** "perceived distances".

It gives an idea of the state of mind of respondents toward the rating scale.

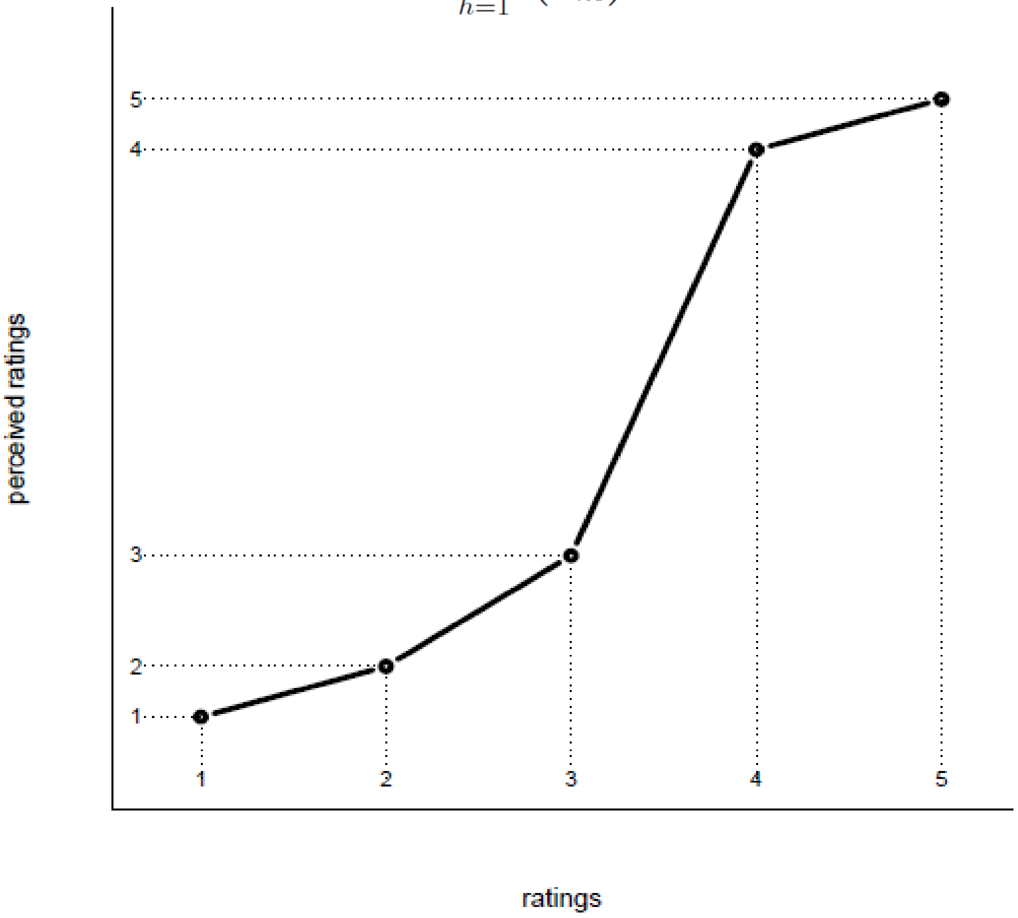
TRANSITION PLOT – CUB model

$$\phi_t(s) = 1 - \xi$$



TRANSITION PLOT – NLCUB model

$$\phi_t(s) = (1-\xi)\frac{\binom{t}{w_{g_s s}}(1-\xi)^{w_{g_s s}}\xi^{t-w_{g_s s}}}{\sum_{h=1}^{g_s}\binom{t}{w_{hs}}(1-\xi)^{w_{hs}}\xi^{t-w_{hs}}}$$



The transition plot

NonLinear CUB models

- Derive from a different assumed mechanism in the Feeling approach (the Uncertainty approach is unchanged)
- Allow us to gain insight about the state of mind toward the rating scale
- Include traditional CUB models as a special case

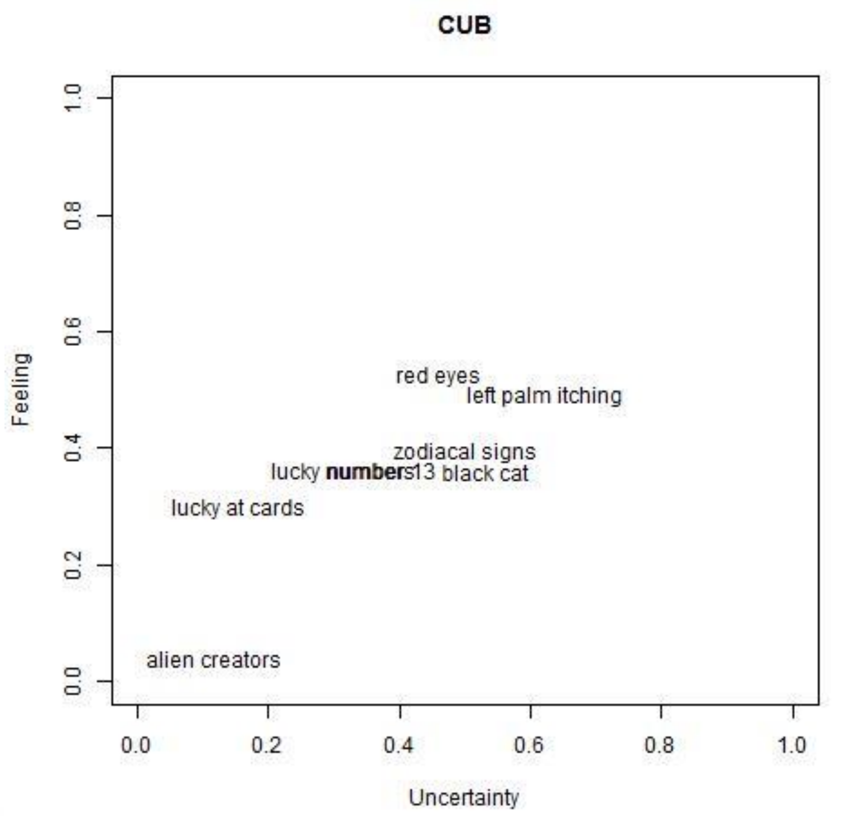
NonLinear CUB models

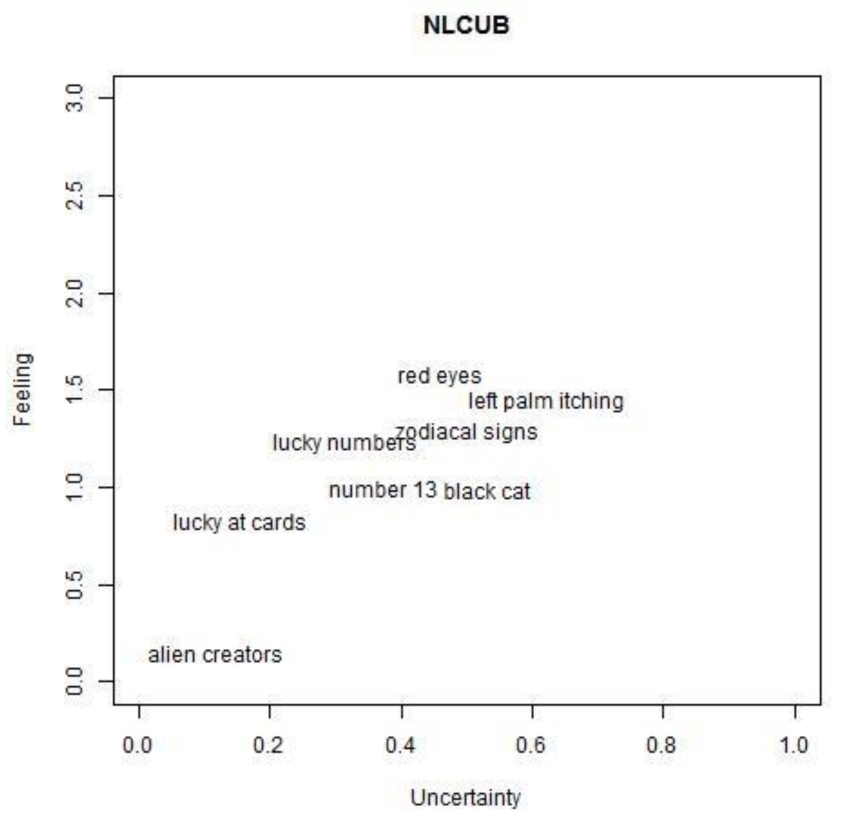
- Derive from a different assumed mechanism in the Feeling approach (the Uncertainty approach is unchanged)
- Allow us to gain insight about the state of mind toward the rating scale
- Include traditional CUB models as a special case

NonLinear CUB models

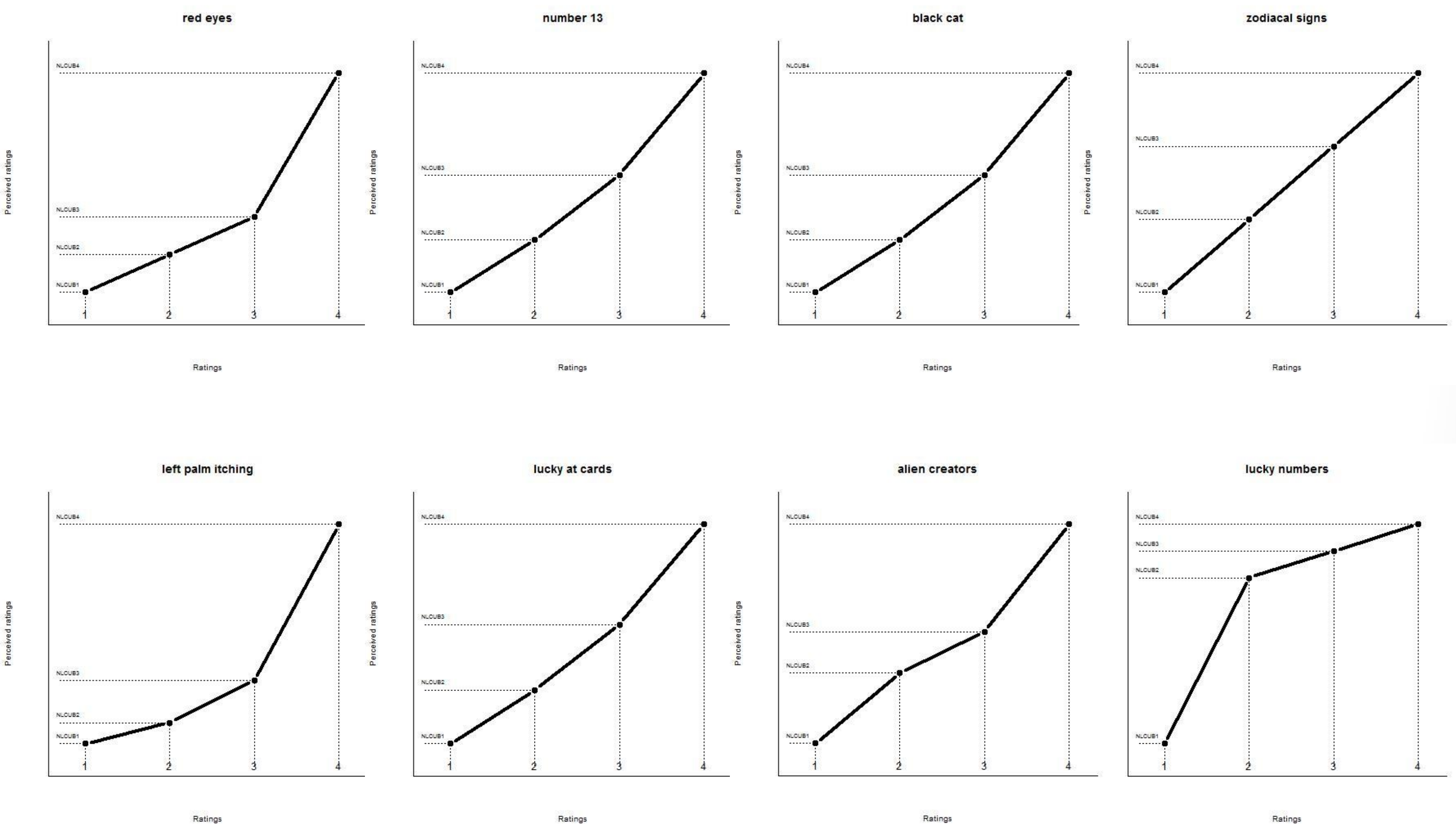
- Derive from a different assumed mechanism in the Feeling approach (the Uncertainty approach is unchanged)
- Allow us to model nonlinear DPs, gaining insight about the state of mind toward the rating scale
- Include traditional CUB models as a special case

Example 1 (superstition)



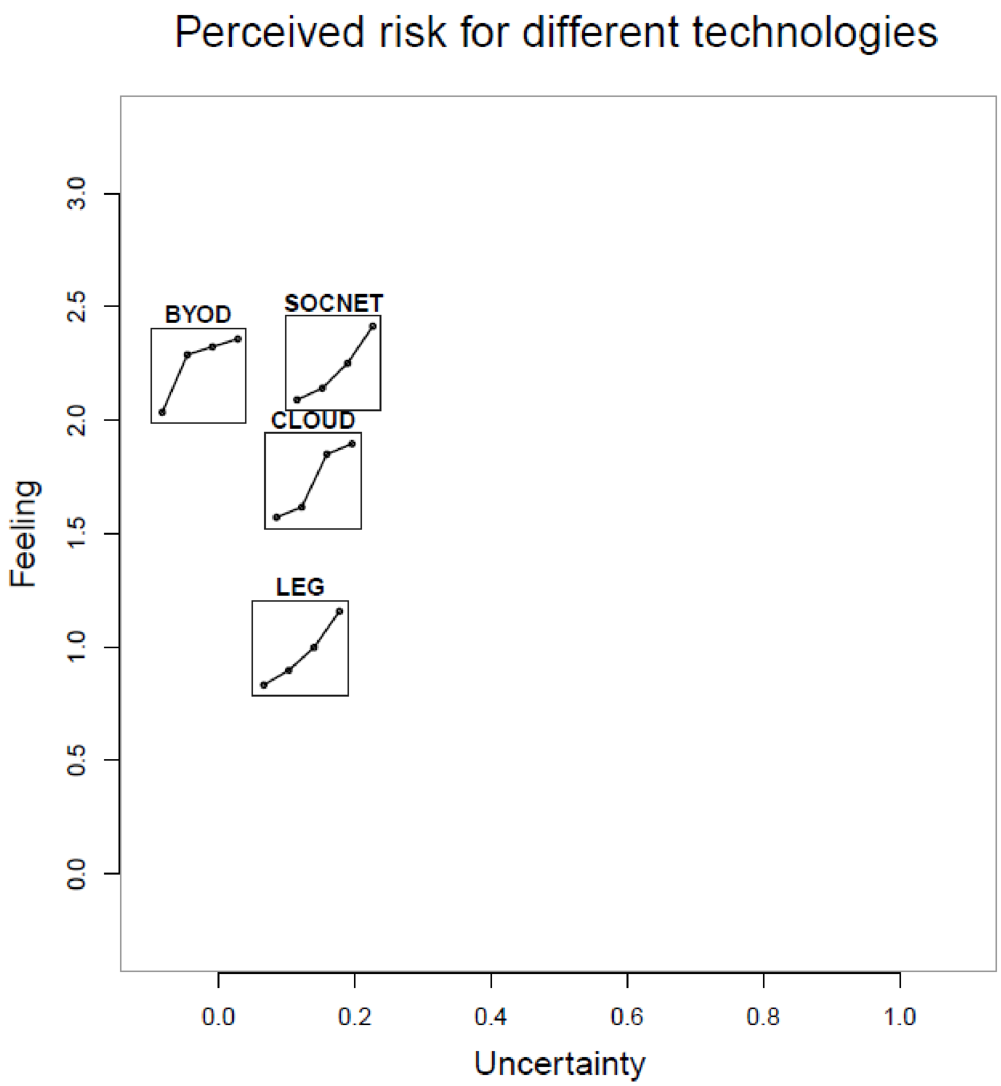


Example 1 (superstition)



NLCUB models - example 1

Example 2 (fraud management)



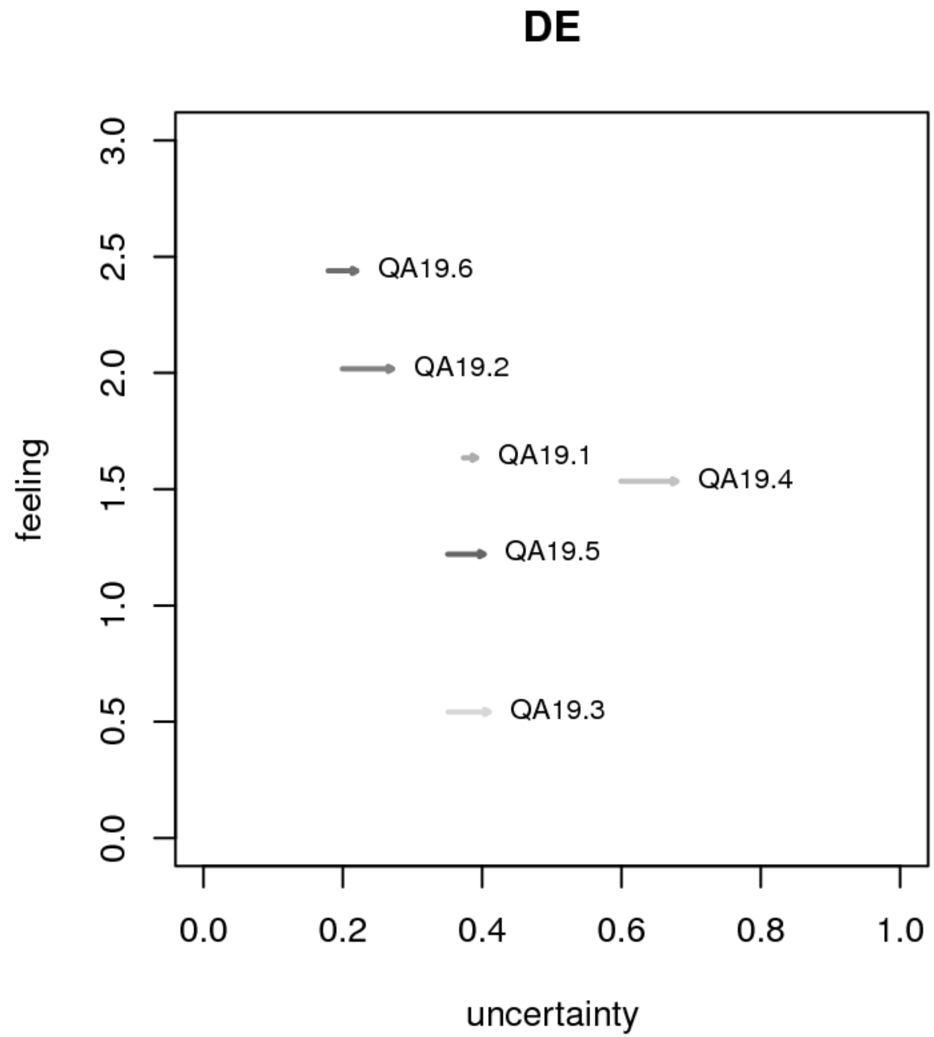
Example 3 (Standard Eurobarometer 81)

- Manisera & Zuccolotto (*Pattern Recognition Letters*, 2014) have proposed a procedure to take into account the presence of "don't know" responses (DK)
- The idea is that DKs inform about the uncertainty of the respondents, so they can be introduced in the CUB framework
- DKs determine an adjustment of the uncertainty parameter

Example 3 (Standard Eurobarometer 81)

The arrows show the shift in uncertainty due to the presence of DK responses.

The arrows are coloured in a gray-level scale. The darker the colour, the higher the degree of nonlinearity of the transition plot, according to a nonlinearity index λ proposed by Manisera&Zuccolotto (*QdS - Journal of Methodological and Applied Statistics,* 2013)



Parameter estimation (two-steps procedure)

$$L(\xi, \pi | \mathbf{g}; \mathbf{s}) = \sum_{i=1}^{n} \log \left\{ \pi \left[\sum_{h=1}^{g_{s_i}} \binom{T}{w_{hs_i}} (1-\xi)^{w_{hs_i}} \xi^{T - w_{hs_i}} \right] + (1-\pi)\frac{1}{m} \right\}$$

Step 1: Fix a maximum value T_{max} for T, and maximize ( ) with respect to ξ and π, for all the possible configurations of $g_1, \cdots, g_m$ such that $g_1 + \cdots + g_m \leq T_{max} + 1$. At the end of this step, we have one NLCUB model for each configuration of $g_1, \cdots, g_m$, along with the corresponding ML estimates of the parameters ξ and π.

Likelihood function for fixed g (step 1)

$$L(\xi, \pi | \mathbf{g}; \mathbf{s}) = \sum_{i=1}^{n} \log \left\{ \pi \left[\sum_{h=1}^{g_{s_i}} \binom{T}{w_{hs_i}} (1-\xi)^{w_{hs_i}} \xi^{T - w_{hs_i}} \right] + (1-\pi) \frac{1}{m} \right\}$$

A good choice may be
$T_{max} = 2m\text{-}1$
(stylized facts from a
large exploratory study
+ identifiability issues)

OPTIMIZATION:
(1) numerical methods
(2) EM algorithm

Model selection (step 2)

Step 2: Among the models defined in Step 1, select the 'best one' according to a given criterion. Let $\hat{\mathbf{g}}$ be the configuration corresponding to the 'best' model, the NLCUB model parameters are finally estimated by $\hat{\boldsymbol{\theta}} = (\hat{\xi}, \hat{\pi}, \hat{\mathbf{g}})'$, where

$$\hat{\xi}, \hat{\pi} = \arg\max_{\xi,\pi} L(\xi, \pi | \hat{\mathbf{g}}; \mathbf{s})$$

Model selection (step 2)

Step 2: Among the models defined in Step 1, select the 'best one' according to a given criterion. Let $\hat{\mathbf{g}}$ be the configuration corresponding to the 'best' model, the NLCUB model parameters are finally estimated by $\hat{\boldsymbol{\theta}} = (\hat{\xi}, \hat{\pi}, \hat{\mathbf{g}})'$, where

Maximum Likelihood

Information criteria

Out-of-sample predictive measures

NLCUB: R functions available!

Description

Generic code for Nonlinear CUB estimation, graphical representations, fit evaluation, data simulation

Usage

```
NLCUB(r,g = c(), m = c(), maxT = c(), param0 = c(0.5,0.5), freq.table
= TRUE, method = "EM", draw.plot = TRUE, dk = c() )
```

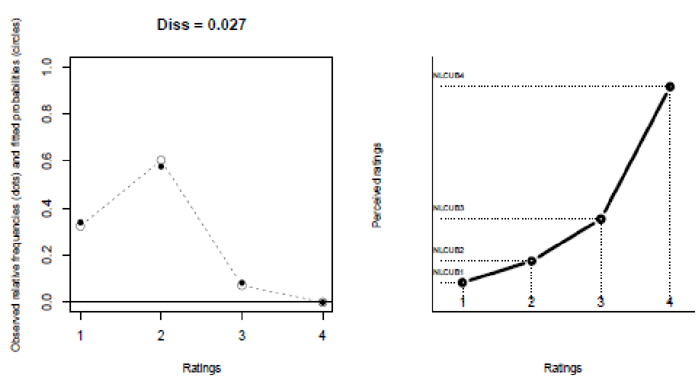
Arguments

`r`	a vector of observed ratings (either microdata or the m observed frequencies - frequency table); see `freq.table`
`m`	integer: number of categories of the response scale (active only when `g` is not declared)
`g`	a vector of the 'latent' categories assigned to each rating point; if `g` is declared, Nonlinear CUB parameters are estimated for fixed `g`, else model selection is performed in order to determine the optimal `g`
`maxT`	integer: maximum value for T (must be `maxT`$> m-1$, default is $2m-1$) (active only when `g` is not declared)
`param0`	starting values for π and ξ
`freq.table`	logical: if `TRUE`, the data in `r` is the vector of the m observed frequencies (frequency table)
`method`	character: method to use for likelihood maximization; `method="NM"` for likelihood based - Melder-Mead maximization - `method="EM"` for likelihood based - EM algorithm
`draw.plot`	logical: if `TRUE`, two graphs are plotted: observed vs fitted frequencies and transition plot
`dk`	proportion of 'don't know' responses; if declared, in addition to the estimate of π, the estimated of π adjusted for the presence of dk responses is provided

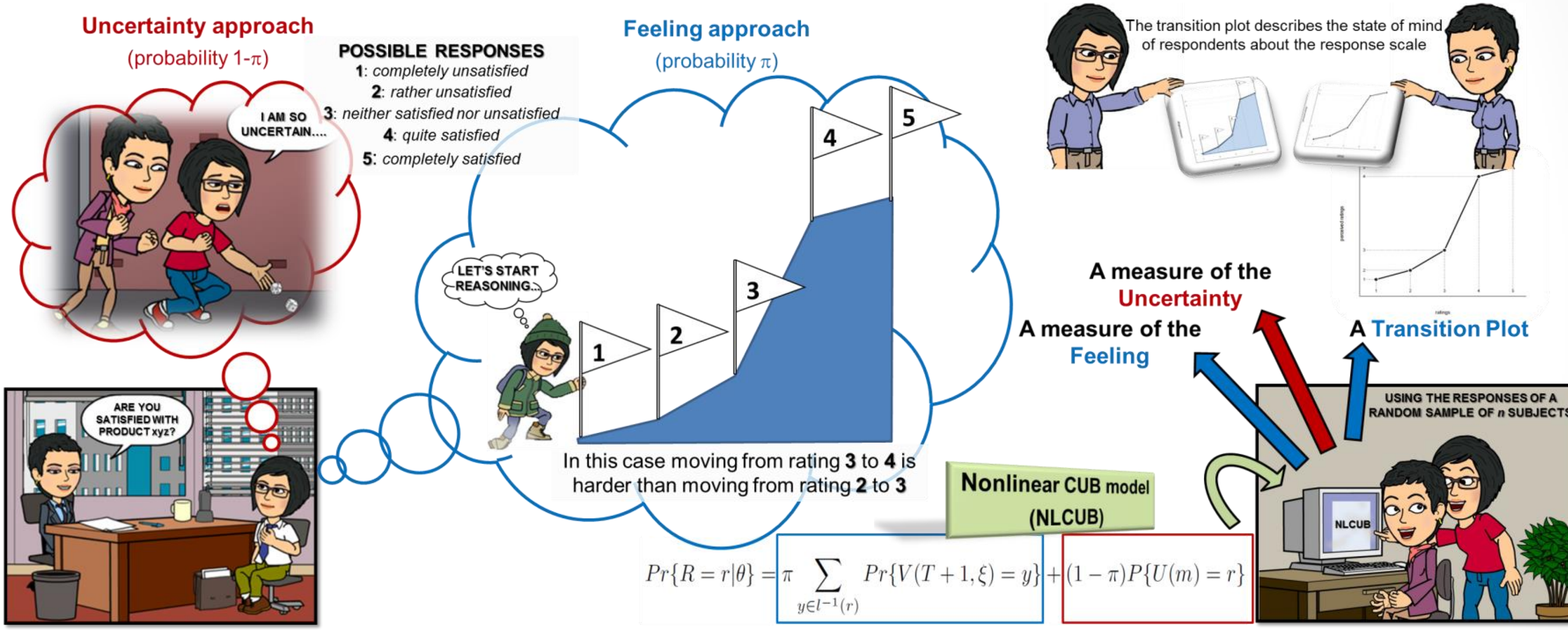
NLCUB: R functions available!

Value

`pai`	parameter estimate for π
`csi`	parameter estimate for ξ
`g`	optimal value for $\mathbf{g} == [g_1, \cdots, g_m]$ (if $\mathbf{g}$ is not declared as input)
`Varmat`	estimated asymptotic variance-covariance matrix of the ML estimator for (π, ξ) for fixed $\mathbf{g}$
`Infmat`	estimated Information matrix
`Fit`	m fitted frequencies, obtained according to the estimated NLCUB model
`diss`	the dissimilarity index value
`transprob_mat`	transition probability matrix containing $\phi_t(s)$
`transprob`	$m-1$ transition probabilities $\phi(s)$
`uncondtransprob`	unconditioned transition ϕ probability
`mu`	estimate of μ
`NL_index`	the nonlinearity index value
`pai_adj`	estimate of the uncertainty parameter adjusted for the presence of 'don't know' (dk) responses



Summarizing…



Basic References

- D'Elia A., Piccolo D. (2005) A mixture model for preference data analysis, *Computational Statistics and Data Analysis,* 49, 917-934.
- Iannario M., Piccolo D. (2011) CUB Models: Statistical Methods and Empirical Evidence, in: Kenett, R. S. and Salini, S. (eds.), *Modern Analysis of Customer Surveys,* Wiley, NY, 231–254.
- Manisera M., Zuccolotto P. (2013) Nonlinear CUB models: some stylized facts. QdS - Journal of Methodological and Applied Statistics, 15, 1-20
- Manisera M., Zuccolotto P. (2014) Modelling rating data with Nonlinear CUB models, *Computational Statistics and Data Analysis*, 78, 100–118.
- Manisera M., Zuccolotto P. (2014) Modelling "don't know" responses in rating scales. Pattern Recognition Letters, 45, 226-234

Basic References

- Manisera M., Zuccolotto P. (2014). Nonlinear CUB models: the R code. *Statistica & Applicazioni*, XII, 205-223.
- Manisera M., Zuccolotto P. (2015). Identifiability of a model for discrete frequency distributions with a multidimensional parameter space, *Journal of Multivariate Analysis*, 140, 302-316.
- Manisera M., Zuccolotto P. (2015). Visualizing Multiple Results from Nonlinear CUB Models with R Grid Viewports. *Electronic Journal of Applied Statistical Analysis*, 8, 360-373.
- Manisera M., Zuccolotto P. (2016). Treatment of 'don't know' responses in a mixture model for rating data, *Metron*, 74, 99-115.
- Manisera M., Zuccolotto P. (2016). Estimation of Nonlinear CUB models via numerical optimization and EM algorithm, *Communications in Statistics - Simulation and Computation*, *forthcoming*.

Thank you

Thank you for your attention