Similarity and Substitutability

Paulo Natenzon
Washington University in Saint Louis

14th SAET Conference at Waseda University
August 21, 2014

Motivation

- Goal: useful parametrizations of stochastic data
- Important to understand the behavioral content of each parameter
- Two Thurstonian models
- Standard Probit and Bayesian Probit: $X \sim \mathcal{N}(\mu, \Sigma)$
- Correlation matters under low precision and close to indifference
- Substitutability in the Standard Probit
- Similarity in the Bayesian Probit

Louis Leon Thurstone (1887–1955)

'Law' of Comparative Judgement (Thurstone, 1927)

The 'law' is a **model** of binary comparisons:

Alternatives ordered in a **psychological continuum**

- gradations of gray, weight, excellence

The **discriminal process** for each alternative $X_i = \mu_i + \varepsilon_i$

ε_i discriminal deviation $\sim \mathcal{N}(0, \sigma_i^2)$

σ_i discriminal dispersion

$$\rho(1, \{1, 2\}) = \mathbb{P}\{X_1 > X_2\}$$

Example: choice of lotteries

Experimental data from Soltani, De Martino and Camerer (2012)

Each subject made 160 pairwise choices

In each trial, four seconds to evaluate the options:

And two seconds to choose

Standard Probit and Bayesian Probit

Same parameters for both models: $X \sim \mathcal{N}\left(\mu, \frac{1}{t}\Sigma\right)$

- Random variable X_i for each alternative i
- $X_1, \ldots, X_n$ joint normally distributed with
  - $\mu_i \in \mathbb{R}$ expectations
  - $\sigma_{ij} \in [0, 1]$ correlations
  - $1/t > 0$ equal variance

Standard probit: $\ddot{\rho}_t^{\mu\sigma}(j, B) = \mathbb{P}\{X_j \geq X_k, \forall k \in B\}$

Bayesian probit: $\rho_t^{\mu\sigma}(j, B) = \mathbb{P}\{m_j \geq m_k, \forall k \in B\}$

where $m = \left[I^{-1} + t\Sigma^{-1}\right]^{-1}\left[\Sigma^{-1}X\right]$ are mean posterior beliefs (prior is iid standard normal)

Binary Choice

Proposition

$$\rho_t^{\mu\sigma}(i,\{i,j\}) = \Phi\left(\frac{\sqrt{t}}{\sqrt{2}}\frac{(\mu_i-\mu_j)}{\sqrt{1-\sigma_{ij}}}\right) = \ddot{\rho}_t^{\mu\sigma}(i,\{i,j\})$$

Binary Choice

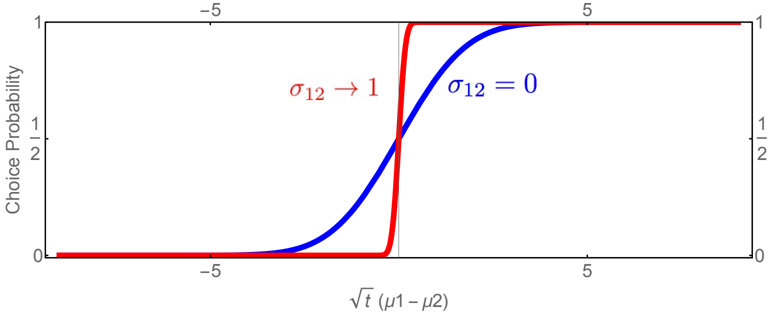
Proposition

$$\rho_t^{\mu\sigma}(i,\{i,j\}) = \Phi\left(\frac{\sqrt{t}}{\sqrt{2}}\frac{(\mu_i - \mu_j)}{\sqrt{1-\sigma_{ij}}}\right) = \ddot{\rho}_t^{\mu\sigma}(i,\{i,j\})$$

- Φ is standard normal cdf
- Equivalence for binary choice data
- Same estimation procedures
- Distinct interpretation for σ needs more alternatives

Effect of correlation

$$\rho_t^{\mu\sigma}(i,\{i,j\}) = \Phi\left(\frac{\sqrt{t}}{\sqrt{2}}\frac{(\mu_i-\mu_j)}{\sqrt{1-\sigma_{ij}}}\right) = \ddot{\rho}_t^{\mu\sigma}(i,\{i,j\})$$



- σ matters when $\sqrt{t}(\mu_i-\mu_j)$ is small

Substitutability in Standard Probit

Let $\mu_1 = \mu_2 = \mu_3$.

Proposition

$\ddot{\rho}_t^{\mu\sigma}(i, \{i, j, k\}) \geq \ddot{\rho}_t^{\mu\sigma}(j, \{i, j, k\})$ *if and only if* $\sigma_{ik} \leq \sigma_{jk}$.

Substitutability in Standard Probit

Let $\mu_1 = \mu_2 = \mu_3$.

Proposition

$\ddot{\rho}_t^{\mu\sigma}(i, \{i, j, k\}) \geq \ddot{\rho}_t^{\mu\sigma}(j, \{i, j, k\})$ *if and only if* $\sigma_{ik} \leq \sigma_{jk}$.

Example

$B = \{1, 2, 3\}$

$$\left.\begin{array}{l} \ddot{\rho}_t^{\mu\sigma}(1, B) = 0.4 \\ \ddot{\rho}_t^{\mu\sigma}(2, B) = 0.3 \\ \ddot{\rho}_t^{\mu\sigma}(3, B) = 0.3 \end{array}\right\} \Longrightarrow \sigma_{12} = \sigma_{13} < \sigma_{23}$$

Alternatives 2 and 3 have a higher degree of substitutability.

Proof

Let $A = \{1, 2, 3\}$.

$$\begin{bmatrix} X_1 \\ X_2 \\ X_3 \end{bmatrix} \sim \mathcal{N}\left(\begin{bmatrix} \mu_1 \\ \mu_2 \\ \mu_3 \end{bmatrix}, \frac{1}{t}\begin{bmatrix} 1 & \sigma_{12} & \sigma_{13} \\ \sigma_{12} & 1 & \sigma_{23} \\ \sigma_{13} & \sigma_{23} & 1 \end{bmatrix}\right)$$

$$\begin{aligned} \ddot{\rho}_t^{\mu\sigma}(1, A) &= \mathbb{P}\left(\{X_1 > X_2\} \cap \{X_1 > X_3\}\right) \\ &= \int_{-\infty}^{\infty}\int_{-\infty}^{x}\int_{-\infty}^{x} \varphi(x, y, z)\; dz\, dy\, dx \end{aligned}$$

closed form expression?

Proof

Let $L_1 = \begin{bmatrix} -1 & 1 & 0 \\ -1 & 0 & 1 \end{bmatrix}$

Let $B = (B_1, B_2)$ be standard normal

Let $MM' = L_1 \Sigma L_1 = \mathrm{Var}(L_1 X)$

Then

$$\begin{aligned}
\ddot{\rho}_t^{\mu\sigma}(1, A) &= \mathbb{P}\left(\{X_2 - X_1 < 0\} \cap \{X_3 - X_1 < 0\}\right) \\
&= \mathbb{P}\{L_1 X < 0\} \\
&= \mathbb{P}\{MB < 0\} \\
&= \mathbb{P}\left\{ \begin{array}{l} B_1 \leq 0 \\ \text{and} \\ B_2 \leq -B_1 \dfrac{(1+\sigma_{23}-\sigma_{12}-\sigma_{13})}{\sqrt{4(1-\sigma_{12})(1-\sigma_{13})-(1+\sigma_{23}-\sigma_{12}-\sigma_{13})^2}} \end{array} \right. \\
&= \frac{1}{4} + \frac{1}{2\pi} \arctan\left(\frac{(1+\sigma_{23}-\sigma_{12}-\sigma_{13})}{\sqrt{4(1-\sigma_{12})(1-\sigma_{13})-(1+\sigma_{23}-\sigma_{12}-\sigma_{13})^2}} \right)
\end{aligned}$$

Proof

Since arctan is strictly increasing,

$$\ddot{\rho}_t^{\mu\sigma}(1, \{1, 2, 3\}) > \ddot{\rho}_t^{\mu\sigma}(2, \{1, 2, 3\})$$

if and only if

$$\frac{(1+\sigma_{23}-\sigma_{12}-\sigma_{13})}{\sqrt{4(1-\sigma_{12})(1-\sigma_{13})-(1+\sigma_{23}-\sigma_{12}-\sigma_{13})^2}} > \frac{(1+\sigma_{13}-\sigma_{12}-\sigma_{23})}{\sqrt{4(1-\sigma_{12})(1-\sigma_{23})-(1+\sigma_{13}-\sigma_{12}-\sigma_{23})^2}}$$

if and only if

$$1 + \sigma_{23} - \sigma_{12} - \sigma_{13} > 1 + \sigma_{13} - \sigma_{12} - \sigma_{23}$$

if and only if

$$\sigma_{23} > \sigma_{13}$$

Q.E.D.

Similarity in Bayesian Probit

Let $\rho_0^{\mu\sigma}(i, B) = \lim_{t \to 0+} \rho_t^{\mu\sigma}(i, B)$

Proposition

$\rho_0^{\mu\sigma}(i, \{i, j, k\}) \geq \rho_0^{\mu\sigma}(j, \{i, j, k\})$ *if and only if* $\sigma_{ik} \geq \sigma_{jk}$.

Example

$B = \{1, 2, 3\}$

$$\left.\begin{array}{l} \rho_0^{\mu\sigma}(1, B) = 0.4 \\ \rho_0^{\mu\sigma}(2, B) = 0.3 \\ \rho_0^{\mu\sigma}(3, B) = 0.3 \end{array}\right\} \Longrightarrow \sigma_{23} < \sigma_{12} = \sigma_{13}$$

Alternatives 2 and 3 have a lower degree of similarity.

Similarity in Bayesian Probit: intutition

Take any symmetric, absolutely continuous prior

1	1	2	2	3	3
2	3	1	3	1	2
3	2	3	1	2	1

If you only learn that

$(2 \succ 3)$

1	2	2
2	1	3
3	3	1

or

$(3 \succ 2)$

1	3	3
3	1	2
2	2	1

then you never choose alternative 1.

Conclusion

- Goal: useful parametrizations of stochastic data
- Important to understand the behavioral content of each parameter
- Two Thurstonian models
- Standard Probit and Bayesian Probit: $X \sim \mathcal{N}(\mu, \Sigma)$
- Correlation matters under low precision and close to indifference
- Substitutability in the Standard Probit
- Similarity in the Bayesian Probit