Basic ideas of a latent variable in the framework of Item response theory exemplified by the Rasch model

Rolf Steyer

Friedrich Schiller University Jena
Institute of Psychology
Chair of Methodology and Evaluation Research

Motivation and Theory of Science Background

Many theories in psychology involve theoretical variables that cannot directly be observed. Obviously, these theories can only be tested in a strict sense if we succeed in translating these concepts of colloquial language into a language that is directly related to our statistical data analysis. So what is the language that is directly related to our statistical data analysis? The answer is obvious: it is the language in which probability, conditional probability, expectation, conditional expectation, variance, correlation, distribution, etc. are defined, i.e., probability theory.

- How can we define latent variables in terms of probability theory?
- What is the difference between a manifest random variable and a latent random variable?

In this presentation I introduce the basic ideas using a simple example.

Logit transformation

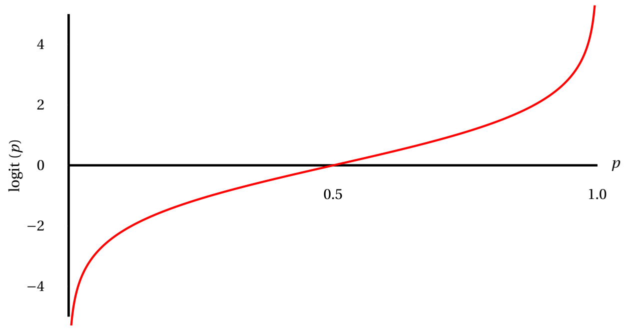

Abbildung 1: Graph of the logit transformation of p

The function $\text{logit}: \,]0,1[\, \rightarrow \mathbb{R}$ is defined by: $\text{logit}(p) = \ln \frac{p}{1-p}, \quad \forall p \in \,]0,1[$. The logit transformation is bijective. Its inverse function is the logistic transformation $\text{logit}^{-1}: \mathbb{R} \rightarrow \,]0,1[$ defined by

$$\text{logit}^{-1}(x) = \frac{\exp(x)}{1+\exp(x)}, \quad \forall x \in \mathbb{R}.$$

Hence,

$$\text{logit}^{-1}\left[\text{logit}(p)\right] = \frac{\exp\left[\text{logit}(p)\right]}{1+\exp\left[\text{logit}(p)\right]} = p.$$

The simplest example for the Rasch model: Joe and Ann – compressed

Tabelle 1: Joe and Ann with two items satisfying the Rasch model

U	$P(U=u)$	$P(Y_1=1 \mid U)$	$P(Y_2=1 \mid U)$	$P(Y_1=1, Y_2=1 \mid U)$
Joe	1/2	.60	.40	.24
Ann	1/2	.80	.64	.512

Rasch model: Joe and Ann with latent variable – compressed

Tabelle 2: Joe and Ann with two items satisfying the Rasch model

U	$P(U=u)$	$P(Y_1=1 \mid U)$	$P(Y_2=1 \mid U)$	$P(Y_1=1, Y_2=1 \mid U)$	ξ
Joe	1/2	.60	.40	.24	0.4054651
Ann	1/2	.80	.64	.512	1.3862943

$$\xi := \text{logit}[P(Y_1=1 \mid U)]$$

Rasch model: Joe and Ann – full table

Tabelle 3: Joe and Ann With Two Items Satisfying the Rasch Model

Poss. outcomes (Unit, Item 1, Item 2)	$P(\{\omega_i\})$	U	Y_1	Y_2	$P(Y_1{=}1\|U)$	$P(Y_2{=}1\|U)$	$\text{logit}_1 =: \xi$	logit_2
$\omega_1 = (\textit{Joe}, -, -)$	.12	*Joe*	0	0	.6	.4	0.4054651	−0.4054651
$\omega_2 = (\textit{Joe}, -, +)$	.08	*Joe*	0	1	.6	.4	0.4054651	−0.4054651
$\omega_3 = (\textit{Joe}, +, -)$	.18	*Joe*	1	0	.6	.4	0.4054651	−0.4054651
$\omega_4 = (\textit{Joe}, +, +)$	.12	*Joe*	1	1	.6	.4	0.4054651	−0.4054651
$\omega_5 = (\textit{Ann}, -, -)$	.036	*Ann*	0	0	.8	.64	1.3862943	0.5753641
$\omega_6 = (\textit{Ann}, -, +)$	.064	*Ann*	0	1	.8	.64	1.3862943	0.5753641
$\omega_7 = (\textit{Ann}, +, -)$	.144	*Ann*	1	0	.8	.64	1.3862943	0.5753641
$\omega_8 = (\textit{Ann}, +, +)$	.256	*Ann*	1	1	.8	.64	1.3862943	0.5753641

The two random variables logit_1 and logit_2 are defined by

$$\text{logit}_i := \text{logit}\,[P(Y_i{=}1|U)] := \ln \frac{P(Y_i{=}1|U)}{1 - P(Y_i{=}1|U)}, \quad i = 1,2,$$

where $P(Y_i{=}1|U) := E(1_{Y_i=1}|U)$ denotes the conditional expectation of $1_{Y_i=1}$.

$$\begin{aligned}\text{logit}_2 &= \xi - \beta_2 \qquad \text{(Rasch homogeneity)}\\ &= \xi - (\text{logit}_1 - \text{logit}_2) = \xi - 0.8109302.\end{aligned}$$

Rasch model: Homogeneity

We consider the random experiment of sampling a person and observing the responses to m binary items Y_i with values 0 or 1. For all $i \in I := \{1,\dots,m\}$, we assume $0 < P(Y_i = 1|U) < 1$ and define the random variable

$$\text{logit}_i := \ln\left(\frac{P(Y_i = 1|U)}{1 - P(Y_i = 1|U)}\right), \tag{1}$$

where U denotes the person variable, the value of which is the person drawn. Furthermore, we define

$$\xi := \text{logit}_1. \tag{2}$$

Assumption 1: Rasch homogeneity

$$\forall i \in I\ \exists \beta_i \in \mathbb{R}: \quad \text{logit}_i = \xi - \beta_i. \tag{3}$$

Inserting (1) and solving Equation (3) for $P(Y_i = 1|U)$ yields

$$\forall i \in I: \quad P(Y_i = 1|U) = \frac{\exp(\xi - \beta_i)}{1 + \exp(\xi - \beta_i)}. \tag{4}$$

Assumption 1 and the definition of ξ also imply ξ-conditional independence of U and each item Y_i, i.e.,

$$\forall i \in I: \quad P(Y_i = 1|U) = P(Y_i = 1|\xi). \tag{5}$$

Rasch model: U-conditional independence of the items

Assumption 2: U-conditional independence of the items $Y_1, \ldots, Y_m$

$$\forall i \in I: \quad P(Y_i = 1 \mid U, Y_1, \ldots, Y_{i-1}, Y_{i+1}, \ldots, Y_m) = P(Y_i = 1 \mid U). \tag{6}$$

Assumptions 1, 2, and the definition of ξ imply ξ-conditional independence of the items, i.e.,

$$\forall i \in I: \quad P(Y_i = 1 \mid \xi, Y_1, \ldots, Y_{i-1}, Y_{i+1}, \ldots, Y_m) = P(Y_i = 1 \mid \xi), \tag{7}$$

which is equivalent to

$$\forall J \subset I: \quad P\Big(\bigcap_{i \in J} \{Y_i = y_i\} \Big| \xi\Big) = \prod_{i \in J} P(Y_i = y_i \mid \xi), \quad \text{where } y_i \in \{0, 1\}. \tag{8}$$

What Joe and Ann teach us about latent variables

- When we talk about a latent variable **we refer to a random experiment**. In the Joe-Ann-example, this random experiment consists of drawing a person from a set of persons and observing the responses to the two items. In this example the set of persons is $\Omega_U = \{Joe, Ann\}$.
- **The random experiment is represented by a probability space** $(\Omega, \mathscr{A}, P)$. In this example, the set of (possible) outcomes is $\Omega = \{\omega_1, \ldots, \omega_8\}$, the σ-algebra $\mathscr{A}$, i.e., the set of (possible) events, is the power set of Ω, and the probability measure P is determined by the probabilities $P(\{\omega_i\})$ of the elementary events $\{\omega_i\}$, $i = 1, \ldots, 8$.
- The **person variable** U is an ordinary random variable. Therefore, it has a distribution and a joint distribution with the other random variables on this probability space, such as Y_1 and Y_2.

What Joe and Ann teach us about latent variables—continued 1

- **The latent variable** is defined exclusively using the joint distribution of the random variables U, Y_1, and Y_2, by

$$\xi := \text{logit}_1 := \text{logit}\,[P(Y_1{=}1\,|\,U)]\,,$$

  where $P(Y_1{=}1|U)$ denotes the U-conditional probability of the event $\{Y_1{=}1\}$.The two assumptions defining the Rasch model, Rasch homogeneity and U-conditional independence of the items, are not involved in the definition of the latent variable. They only broaden its meaning. [Compare the 'definitions' of a latent variable in Bollen (2002).]
- **Without the person variable U, the latent variable ξ cannot be defined.**
- The definition $\xi := \text{logit}\,[P(Y_1{=}1|U)]$ is arbitrary to some degree. ξ is just an arbitrarily chosen member of the family $(\xi_k)_{k\in K}$ of random variables satisfying $\xi_k = \xi + \alpha$, $\alpha \in \mathbb{R}$. Hence, each ξ_k is another version of the 'same' latent variable. We say that the latent variable is *uniquely defined up to translations*. In other words, the latent variable has a *difference scale*.

What Joe and Ann teach us about latent variables (continued 2)

- The **substantive meaning of the latent variable** is exclusively determined by (a) the formal properties of ξ that follow from its definition, by (b) the concrete random experiment considered, and (c) the concrete two items referred to in the table.
- One of these formal properties implied by the definition of ξ is that it is the composition $g \circ U = g(U)$ of the random variable U and a function $g\colon \Omega_U \to \mathbb{R}$ that assigns to each $u \in \Omega_U$ its value $g(u) = \text{logit}[P(Y_1{=}1|U{=}u)]$. This makes clear that the values of ξ are attributes of the persons considered in the set Ω_U of persons (the 'population') and that it makes sense to estimate the values of ξ. (Compare the discussion on the indeterminacy of factor scores.)

What Joe and Ann teach us about latent variables (continued 3)

- Assumptions 1 and 2 of the Rasch model and the definition $\xi := \text{logit}[P(Y_1{=}1\,|\,U)]$ imply

$$P(Y_1{=}1\,|\,U) \;=\; P(Y_1{=}1\,|\,\xi) \tag{9}$$

  and

$$P(Y_2{=}1\,|\,U, Y_1) \;=\; P(Y_2{=}1\,|\,\xi) \tag{10}$$

  This in turn implies that, for $i = 1, 2$, the 'items' Y_i and all functions of U are ξ-conditionally independent, i.e.,

$$P[Y_i{=}1\,|\,U, f(U)] \;=\; P(Y_i{=}1\,|\,\xi), \quad i = 1, 2. \tag{11}$$

  Examples for such functions $f(U)$ are gender, race, and any other attributes of the persons. In psychometrics we refer to this property saying that 'there is no differential item functioning'.
- Equation (9) is a causality condition, called the *complete cause condition*. In terms of colloquial English it says that ξ is 'the only cause of Y_i'. Each item Y_i is U-conditionally independent from all variables that are prior or simultaneous to ξ. In this case, all functions of U such as sex, race, educ. status, ...
- The relationship between a latent variable — if properly defined — and 'indicators' is a causal one. Noteworthy, it is a cause that cannot be manipulated. (In contrast, see Pearl's do-operator approach as well as Rubin's approach (e.g., Rubin, 2005) and the discussion of Holland, 1986).

Model of essentially τ-equivalent variables

Linear measurement models are analog to logistic measurement models. In linear models we consider a similar random experiment as before. However, quantitative manifest random variables $Y_1, \ldots, Y_m$ are observed and we define

$$\eta \;:=\; \tau_1, \tag{12}$$

where $\tau_1 := E(Y_1\,|\,U)$ denotes the *true-score variable of* Y_1, i.e., the U-conditional expectation of Y_1.

The two assumptions of the *model of essentially τ-equivalent variables* are:

$$\forall i \in I := \{1, \ldots, m\}\; \exists \lambda_i \in \mathbb{R}: \quad E(Y_i\,|\,U) \;=\; \eta - \lambda_i, \tag{13}$$

and

$$\forall i \in I: \quad E(Y_i\,|\,U, Y_1, \ldots, Y_{i-1}, Y_{i+1}, \ldots, Y_m) \;=\; E(Y_i\,|\,U). \tag{14}$$

Compare these two assumptions to the two assumptions defining the Rasch model [see Eqs. (3) and (6)]!

More Details

More details on latent variables including latent variables in (ordinary) structural equation models and in longitudinal designs are found in

Steyer, Mayer, Geiser und Cole (2015). A Theory of States and Traits — Revised, *Annual Review of Clinical Psychology, 11,* 71-98

and its supplements. (The revisions concern the concept of a person as a dynamic entity and its amazing implications, also for the properties of latent variables in longitudinal designs.)

Important mile stones, in particular in defining and understanding *method factors* (= specific latent variables) are the papers by Michael Eid (e.g., Eid, 2000) and his colleagues and by Steffi Pohl (e.g., Pohl, Steyer & Kraus, 2008) .

Concepts and notation used in this presentation are introduced in

Steyer und Nagel (in Druck). *Probability and Conditional Expectation: Fundamentals of the Empirical Sciences,* Chichester: Wiley.

Conclusion

- Both, latent variables and causal effects, can be defined exclusively using well-known and well-defined concepts of probability theory. In this case, the implications of our theories, assumptions, and hypotheses can be tested directly in appropriate samples. Statistical inference aims at aspects of the (joint) distributions of random variables. Only if theories refer to the same (joint) distributions can we learn from empirical findings about the validity of our theories.
- In the definitions of latent variables and of causal effects we do not refer to substantive or philosophical theories. It is straight-forward mathematics and we can use logical instead of plausibility arguments in the derivation of further properties and implications.

References

Bollen, K. A. (2002). Latent variables in psychology and the social sciences. *Annual Review of Psychology*, *53*, 605–634.

Eid, M. (2000). A multitrait-multimethod model with minimal assumptions. *Psychometrika*, *65*, 241–261.

Holland, P. W. (1986). Statistics and causal inference. *Journal of the American Statistical Association*, *81*, 945–960.

Pohl, S., Steyer, R. & Kraus, K. (2008). Modelling method effects as individual causal effects. *Journal of the Royal Statistical Society: Series A*, *171*, 41–63.

Rubin, D. B. (2005). Causal inference using potential outcomes: Design, modeling, decisions. *Journal of the American Statistical Association*, *100*, 322–331. doi: 10.1198/016214504000001880

Steyer, R., Mayer, A., Geiser, C. & Cole, D. (2015). A theory of states and traits — revised. *Annual Review of Clinical Psychology*, *11*, 71–98.

Steyer, R. & Nagel, W. (in Druck). *Probabilistic foundations of empirical research: Probability and conditional expectation.* Chichester, England: Wiley.

Book