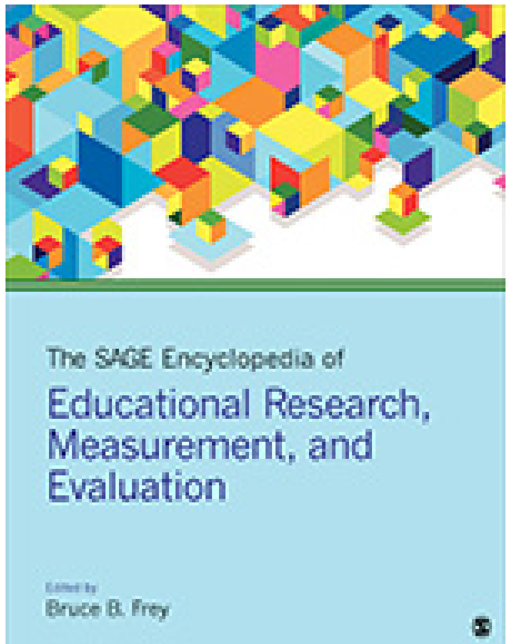


Sage Research Methods

The SAGE Encyclopedia of Educational Research, Measurement, and Evaluation

Author: Haiyan Liu, Zhiyong Zhang
Pub. Date: 2018
Product: Sage Research Methods
DOI: https://doi.org/10.4135/9781506326139
Methods: Educational research, Measurement

Disciplines: Education

Access Date: March 28, 2025

Publisher: SAGE Publications, Inc.

City: Thousand Oaks,

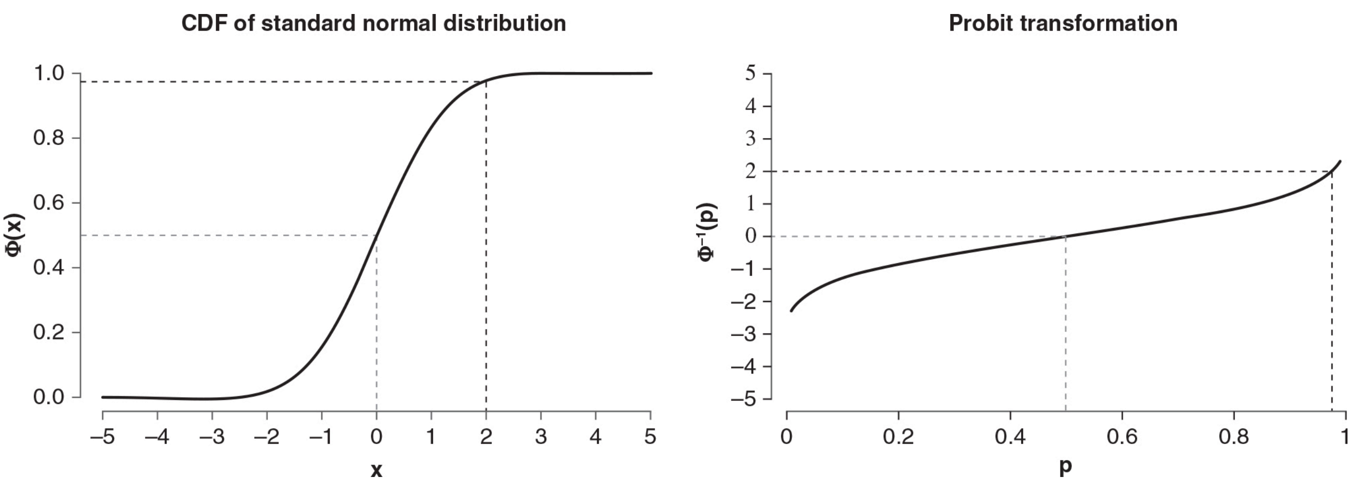
Online ISBN: 9781506326139

Probit transformation is widely used to transform a probability, percentage, or proportion to a value in the unconstrained interval $(-\infty,\infty)$, which is usually referred to as a *quantile* in probability theory. Strictly speaking, probit transformation is the inverse of the cumulative distribution function of the standard normal distribution. For any observed value $x \in (-\infty,\infty)$, the cumulative distribution function of the standard normal distribution, denoted by $\Phi(x)$, is defined as follows:

$$\Phi(x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} e^{-\frac{t^2}{2}} dt,$$

with t being a value that the standard normal distributed variable could take. It converts a value in the interval $(-\infty,\infty)$ to a value p in the interval (0,1) such that $p = \Phi(x)$. For a probability p, or more generally any value between 0 and 1, $\Phi^{-1}(p)$ is its probit transformation to transform p to the quantile x. For instance, $\Phi^{-1}(0) = -\infty$ and $\Phi^{-1}(1) = \infty$. It is true in general that $\Phi[\Phi^{-1}(p)] = p$. For example, when p is .975, $\Phi^{-1}(.975) = 1.96$ and $\Phi^{-1}(1.96) = .975$. An appealing feature of probit transformation is that it converts a sigmoid curve to a line that is almost linear (Figure 1). The linearization brings researchers great convenience because it allows them to model a linear line directly by a linear combination of other variables.

Figure 1 The left panel is plot of the cumulative distribution function (CDF) of the standard normal distribution, and the right panel contains the plot of probit transformation



Probit transformation is often used in modeling categorical, especially binary, outcome data. The well-known probit regression analysis exemplifies its most notable application. In binary data analysis, one is often interested in predicting the binary outcome variable Y. It usually assumes that there is an underlying normally distributed variable Y^* and a threshold τ such that $Y = 0$ when $Y^* \leq \tau$, and $Y = 1$ when $Y^* > \tau$. Therefore,

$$p = Pr(Y = 1) = Pr(Y^* > \tau).$$

The underlying continuous variable Y^* can be analyzed by a regression model with given predictors X,

$$Y^* = X\beta + \varepsilon.$$

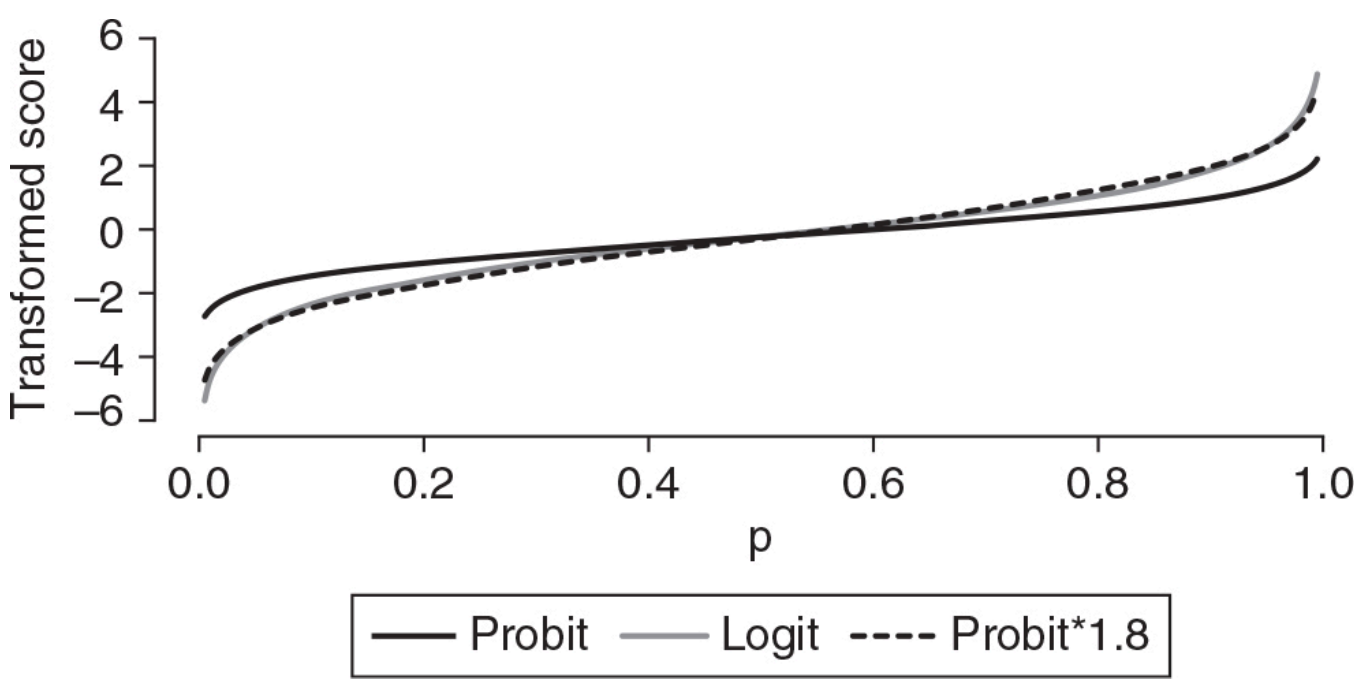
To identity the model, it is usually assumed that $\varepsilon \sim N(0,1)$. Consequently, the probit model has the following form:

$$p = Pr(Y = 1).$$
$$\Phi^{-1}(p) = -\tau + X\beta.$$

There are other transformation methods to convert the interval (0,1) to the unconstrained interval $(-\infty,\infty)$ such as the logit transformation. For a value p in the interval (0,1), its logit transformation is $\log\frac{p}{1-p}$.

Although both probit and logit transformations linearize a sigmoid curve, the slopes of the two linear lines are different, as shown in Figure 2. The slope from the logit transformation is around 1.8 times as large as the one from the probit transformation. For researchers, both probit and logistic transformations have their own appealing features. In probit transformation, the underlying Y^* is assumed to be normally distributed, which is consistent with the normal assumption on the latent constructs in the social and educational sciences, while in the logit transformation, one assumes the underlying continuous variable Y^* follows a logistic distribution. The results from the logit transformation are more interpretable in terms of the odds ratio. The curves under these two transformations are hardly distinguishable when the probit transformation is scaled by 1.8.

Figure 2 Probit versus logit transformation



See also Normal Distribution

Haiyan Liu Zhiyong Zhang

Further Readings

Bliss, C. I. (1934). The method of probits. *Science*, 79(2037), 38–39. doi:

Bliss, C. I. (1935). The calculation of the dosage-mortality curve. *Annals of Applied Biology*, 22(1), 134–167. doi:

Camilli, G. (1994). Origin of the scaling constant *d* = 1.7 in item response theory. *Journal of Educational and Behavioral Statistics*, 19(3), 293–295. doi:

Finney, D. J. (Ed.). (1971). *Probit analysis* (3rd ed.). Cambridge, UK: Cambridge University Press.

https://doi.org/10.4135/9781506326139