RN 001-69

ON A GENERALIZATION OF THE LOGISTIC FUNCTION

Jan de Leeuw

June 1969

PSYCHOLOGICAL INSTITUTE - UNIVERSITY OF LEYDEN - NETHERLANDS

In probabilistic concept learning experiments the logistic function is a useful model. For quite a number of situations in which subjects give subjective probability estimates $\psi(H_1/D)$ that one of two mutually exclusive hypotheses is true the transformation

$$\log \frac{\psi(H_1/D)}{1 - \psi(H_1/D)} \tag{1}$$

produces values SLLR that are linearly related to the (objective) Baysian log likelihood ratios BLLR, with

$$BLLR = \frac{p(D/H_1)}{p(D/H_2)} \; . \tag{2}$$

(cf De Klerk 1968, De Leeuw 1968, 1969a, 1969 b).

The logistic law says, equivalently, that there are constants γ and δ such that

$$\psi(H_1/D) = \frac{1}{1 + \exp\{-\gamma BLLR - \delta\}} \; . \tag{3}$$

In other related experiments (Oppe, in preparation) subject's confidence in the hypotheses is measured by his choice reaction time (CRT). In this case the logit transformation cannot be applied. We propose the following substitute

$$\log \frac{CRT - \alpha}{\beta - CRT} = \gamma BLLR + \delta \; . \tag{4}$$

Solving for CRT gives

$$CRT = \frac{\alpha + \beta \exp(\gamma BLLR + \delta)}{1 + \exp(\gamma BLLR + \delta)} \; , \tag{5}$$

or, equivalently,

$$CRT = \frac{\beta + \alpha \exp(-\gamma BLLR - \delta)}{1 + \exp(-\gamma BLLR - \delta)} \; . \tag{6}$$

From the first expression

$$\lim_{BLLR \to -\infty} CRT = \alpha \; , \tag{7}$$

and from the second

$$\lim_{BLLR \to +\infty} CRT = \beta . \tag{8}$$

Of course if $\alpha = 0$ and $\beta = 1$ the functions defined by (3) and (6) are identical. Combining (3) and (4) gives

$$\log \frac{\psi(H_1/D)}{1 - \psi(H_1/D)} = \log \frac{CRT - \alpha}{\beta - CRT} = \gamma BLLR + \delta , \tag{9}$$

or

$$CRT = (\beta - \alpha)\, \psi(H_1/D) + \alpha . \tag{10}$$

In other words our proposal amounts to using the $\psi(H_1/D)$ of (3) as an underlying variable, which is linearly related to CRT. If $\beta > \alpha$ and $\gamma > 0$ then CRT is a strictly monotonic increasing function of BLLR with values $\alpha < CRT < \beta$, where the values approach the asymptotes $CRT = \beta$ and $CRT = \alpha$ as BLLR approaches $+\infty$ and $-\infty$ respectively. If $BLLR = -\delta/\gamma$ then $CRT = (\alpha + \beta)/2$, if $\alpha = \beta$ then $CRT = (\alpha + \beta)/2$ independent of the value of BLLR. Moreover CRT is symmetric around $BLLR = -\delta/\gamma$, in the sense that for $BLLR = \varepsilon - \delta/\gamma$ it is true that

$$CRT = \frac{\alpha + \beta \exp(\varepsilon\gamma)}{1 + \exp(\varepsilon\gamma)} , \tag{11}$$

and for $BLLR = -\varepsilon - \delta/\gamma$

$$CRT = \frac{\beta + \alpha \exp(\varepsilon\gamma)}{1 + \exp(\varepsilon\gamma)} . \tag{12}$$

Adding up (11) and (12) gives

$$CRT(\varepsilon - \delta/\gamma) + CRT(-\varepsilon - \delta/\gamma) = \alpha + \beta . \tag{13}$$

which is independent of ε.

-2-

Another natural generalisation that comes to mind is (writing x for BLLR and y for CRT)

$$\log \frac{y - \alpha}{\beta - y} = P_m(x), \tag{14}$$

where P_m is a polynomial of degree m. In other words

$$y = \frac{\beta + \alpha \exp(-P_m(x))}{1 + \exp(-P_m(x))} . \tag{15}$$

If m = 1 then this function is identical with the one discussed in the previous section. Of course (15) tends to become trivial if we take m large, both in the continuous case (Weierstrass' theorem) and in the finite case. If we have n data points (x_i, y_i) then, if only

$$\alpha < \min_{i=1}^{n}(y_i) \tag{16}$$

and

$$\beta > \max_{i=1}^{n}(y_i), \tag{17}$$

m = n - 1 garantees a perfect fit. We propose the more general form (15) because slight quadratic deviations from linearity, for example, may very well have a psychological meaning.

-3-

In fitting the function (15) to the data we proceed from equation (14). Let $\| z \|$ denote any vector norm. We want to minimize

$$\left\| \log \frac{y - \alpha}{\beta - y} - P_m(x) \right\| \tag{18}$$

for a fixed value of m. This formulation suggests, in a natural way, a two-stage process. Select any pair of values α^o and β^o in the convex region defined by (16) and (17). Transform y by the generalized logit transformation, and find the best fitting polynomial of degree m. Because in most cases we shall only be interested in values m $\leq$ 3, this can safely be done by the classical normal equation method (although orthogonal polynomials can, of course, also be used). Compute the values $P_m^o(x)$ and move α and β in the convex region in such a way that (18) is minimized for the current values of $P_m(x)$. Observe that the generalized logit transform posesses continuous derivatives everywhere in this region. Because fitting polynomials is (computationally) such an easy thing to do, and because the x-values (and thus the matrix of the normal equations) remain fixed, the most satisfactory way to minimize (18) may very well be to select a large set of permissible trial values (α^0, β^0) and to retain the best ones. The search can then be continued, if necessary, in a smaller region.

-4-

References:

De Klerk, L.F.W. — Probabilistic concept learning. Voorschoten, V.A.M., 1968.

De Leeuw, J. — Probabilistic concept learning. Department of Data Theory, Research Note RN 002-68 University of Leyden.

De Leeuw, J. — Probabilistic concept learning. Paper to be submitted for publication, 1969.

De Leeuw, J. — Likelihood ratio tests for PCL. Department of Data Theory, Research Note RN 002-69 University of Leyden.