ELSEVIER

Stud. Hist. Phil. Biol. & Biomed. Sci. 35 (2004) 613–626

Studies in History and Philosophy of Biological and Biomedical Sciences

www.elsevier.com/locate/shpsc

The early origins of the logit model

J. S. Cramer

University of Amsterdam and Tinbergen Institute, Amsterdam; Baambrugse Zuwe 194, 3645 AM Vinkeveen, Netherlands

Received 19 December 2003; received in revised form 16 April 2004

Abstract

This paper describes the origins of the logistic function and its history up to its adoption in bio-assay and the beginning of its wider acceptance in statistics, ca. 1950. The function was probably first invented in 1838 to describe population growth by the Belgian mathematician Verhulst, who gave it its name in 1845; but it was rediscovered independently several times over in the next eighty years, both for this purpose and for the description of autocatalytic chemical reactions. Its adoption in an altogether different role in bio-assay has been determined decisively by the individual actions and personal histories of a few scholars: the widespread acceptance of the growth function is due to Pearl and Reed, the general recognition of Verhulst's primacy to Yule, and the introduction of the function in bio-assay to Berkson.

Keywords: Growth curve; Autocatalytic function; Logistic model; Probit; Logit

1. Introduction

The sigmoid curve of Fig. 1 is traced by the *logistic function*

$$P(Z) = \frac{\exp Z}{1 + \exp Z} \tag{1}$$

As Z moves through the real number axis, P rises monotonically between the bounds of zero and one, and thus behaves like a distribution function; the corresponding density would be symmetrical around zero. The meaning of this function varies according to the definition of the variables. In the logit version of bio-assay,

E-mail address: mars.cram@worldonline.nl (J.S. Cramer).

1369-8486/$ - see front matter
doi:10.1016/j.shpsc.2004.09.003

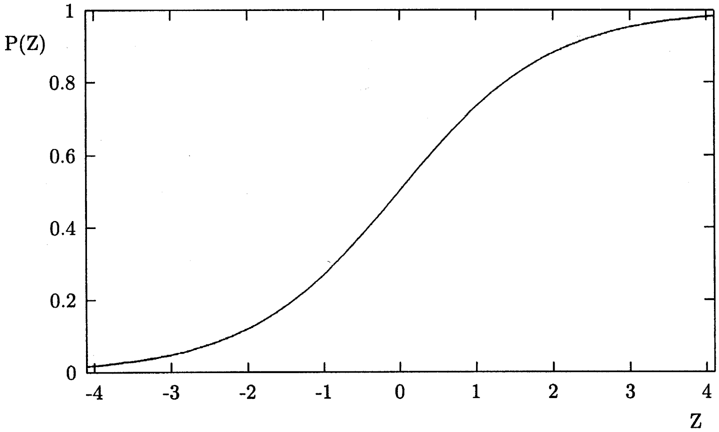

Fig. 1. The logistic curve.

P is the probability of a binary outcome (like the survival or death of an organism), and $Z = \alpha + \beta X$, with X a continuous stimulus or exposure variable (like the dosage of an insecticide). In logistic regression this is generalized to several determinants of P; $Z = x^T\beta$, with x a vector of covariates (including a unit constant) and β their coefficients. Originally, however, the logistic function was designed to describe the course of a *proportion* P over time, with $Z = \alpha + \beta\, t$; it is a *growth curve*, since $P(t)$ rises monotonically with time t.

Over a fairly wide central range, certainly for values of P from 0.3 to 0.7, the shape of the logistic curve closely resembles the normal probability distribution function. The two functions

$$P_l(X) = \frac{\exp\beta X}{1 + \exp\beta X} \tag{2}$$

and

$$P_n(X) = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{X} \exp\left(-\frac{u^2}{2\sigma^2}\right) \mathrm{d}u \tag{3}$$

both pass through the point (0, 0.5), and they can be made almost to coincide upon a suitable adjustment of β and σ. This is a sheer algebraic coincidence, for there appears to be no intrinsic relation between the two forms.

2. The origins of the logistic function: Verhulst

The logistic function was invented in the nineteenth century for the description of the growth of organisms and populations and for the course of autocatalytic chemical reactions. In either case we consider the time path of a quantity $W(t)$ and its growth rate:

$$\overset{o}{W}(t) = \mathrm{d}W(t)/\mathrm{d}t \tag{4}$$

If $W(t)$ is a living population, the simplest assumption is that $\overset{o}{W}(t)$ is proportional to $W(t)$:

$$\overset{o}{W}(t) = \beta W(t), \quad \beta = \overset{o}{W}(t)/W(t) \tag{5}$$

with β the constant rate of growth. This leads of course to exponential growth:

$$W(t) = A \exp \beta t$$

where A may be replaced by an initial value $W(0)$. This model of unopposed growth is familiar from Malthus's dictum that a human population, left to itself, will increase in geometric progression. It is a reasonable model for a young and empty country like the United States in its early years. Two hundred years later exponential growth played a major part in the *Report for the Club of Rome* (Meadows et al., 1972), and it is still implicit in many economic analyses. Like many others, Alphonse Quetelet (1795–1874), the Belgian astronomer who turned into a statistician and a social scientist, was well aware that the indiscriminate extrapolation of exponential growth must lead to impossible values. He experimented with several adjustments of (5) and also asked his pupil, the mathematician Pierre-François Verhulst (1804–1849), to look into the problem.

Like Quetelet, Verhulst approached the problem by adding an extra term to (5) to represent the increasing resistance to further growth, as in

$$\overset{o}{W}(t) = \beta W(t) \text{-} \phi W(t) \tag{6}$$

and then experimenting with various forms of ϕ. The logistic appears when this is a simple quadratic, for in that case (6) may be rewritten with some minor adjustments as

$$\overset{o}{W}(t) = \beta W(t)\ \{\Omega \text{-} W(t)\} \tag{7}$$

where Ω denotes the upper limit or saturation *level* of W, its asymptote as $t \rightarrow \propto$. Growth is now proportional both to the population already attained $W(t)$ and to the remaining room for further expansion Ω - $W(t)$. If we express $W(t)$ as a proportion $P(t) = W(t)\ /\ \Omega$

$$\overset{o}{P}(t) = \beta P(t)\ \{1 \text{-} P(t)\} \tag{8}$$

and the solution of this differential equation is

$$P(t) = \frac{\exp(\alpha + \beta t)}{1 + \exp(\alpha + \beta t)} \tag{9}$$

which Verhulst named the *logistic* function, the population $W(t)$ then follows:

$$W(t) = \Omega \frac{\exp(\alpha + \beta t)}{1 + \exp(\alpha + \beta t)} \tag{10}$$

Verhulst published his ideas between 1838 and 1847 in three papers. The first is a brief note in the *Correspondance Mathématique et Physique*, edited by Quetelet, in 1838. It contains the essence of the argument in four small text pages, followed by a demonstration in tables that the curve agrees very well with the actual course of the population of France, Belgium, Essex and Russia for periods up to 1833. Verhulst explains that he did his research a couple of years before, that he did not have the time for a revision and that he publishes this note only at the insistence of Quetelet. He does not say how he fitted the curves. The second paper, in the *Mémoires* of the Belgian Royal Academy of 1845, is a much fuller account of the function and its properties. Here Verhulst names it the logistic, without further explanation: in a neat diagram, the *courbe logistique* is drawn alongside the *courbe logarithmique*, which we would nowadays call the exponential. Verhulst also determines the three parameters Ω, α and β of (10) by making the curve pass through three observed points. With data for some twenty or thirty years only this is a hazardous method, as is borne out by the resulting estimates of the limiting population Ω. Employing the known values of the Belgian population in 1815, 1830 and 1845 Verhulst finds a limiting population of 6.6 million for that country, and in a similar exercise 40 million for France. At present these populations number 10.2 and 58.7 million. In 1847 there followed a second paper in the same *Mémoires*, which is chiefly notable for an adjustment of the correction term that leads to a much better estimate of 9.5 millions for the Belgian Ω. But the major argument is set out in the article of 1845, and this has rightly been reprinted recently (if only in part) by David and Edwards in their rich collection of important but inaccessible papers from the past (David & Edwards, 2001, p. 65–77).

Verhulst was in poor health and died in 1849 at the age of forty-five. He was primarily a mathematician—professor of mathematics at the Belgian Military Academy—but sensitive to social and political issues. In his obituary of Verhulst, Quetelet attributes his early death to overwork and, rather curiously, to his great stature, as Verhulst was 1.89 meters or well over six feet tall (Quetelet, 1850). His discovery of the logistic curve was not taken up with much enthusiasm by Quetelet; as Vanpaemel has shown, the two men did not see eye to eye on the question of population growth (Vanpaemel, 1987). This may in part account for some curious elements in Quetelet's obituary of Verhulst; while ostensibly praising his pupil, Quetelet stresses his impulsive nature and depicts him as a somewhat silly man. Quetelet recounts at length Verhulst's adventures in Rome, where he was staying in the summer of 1830 when the news broke of the revolution in Paris and of the Belgian secession from the Netherlands. These events moved Verhulst strongly and set him drafting a democratic constitution for the Papal State. He submitted this document to some cardinals he had met, who expressed great interest; still the police were called in, and Verhulst banished from Rome. He left under somewhat dramatic circumstances, having at first barricaded his apartment with the intention of withstanding a siege by the forces of law and order. But then he was only twenty-six years old at the time.

Quetelet did not pay much attention to the logistic curve in his writings; it is mentioned only in an aside, in his *Système social* (Quetelet, 1848). But Verhulst's work was quoted with approval in the statistics textbook of Liagre, who was his colleague at the Military Academy, and in the second edition of this textbook of 1879, its editor Camille Peney repeats the estimation of Ω for Belgium on the basis of more recent population figures, arriving at a value of 13.7 millions. This is however an exception, and the work of Verhulst and his naming of the logistic curve lay largely dormant until the 1920's when they were revived through the publications of Pearl and Reed.

3. Independent discoveries of the function and the revival of the logistic

Apart from Verhulst, the function (1) has been applied to the course of chemical reactions, to the growth of individual organisms and to the growth of populations, in that order. Until 1920 it was commonly known as the autocatalyic function since it arises naturally from a diferential equation like (7) if one considers the course of autocatalytic chemical reactions. Autocatalysis (and indeed catalysis) were discovered in the 1880s by Wilhelm Ostwald (1853–1932); the autocatalytic function originates in a two-part paper on the dynamics of chemical reactions of 1883, where Ostwald considers simple differential equations that describe the course of various types of chemical reactions over time (Ostwald, 1883). Ostwald, who was a great chemical experimentalist, soon became the grand old man of German chemistry. At that time Germany was among the first nations in chemistry, and American students would often complete their training in that country. Ostwald's work (and the man himself) were thus quite well known in the United States; moreover, he visited the United States three times between 1903 and 1905, when he lectured at Harvard as one of the first visiting professors. He received the 1909 Nobel prize for his work on catalysis.

Lloyd has given a very interesting survey of the early history of applications of the function to the growth of organisms and of living populations (Lloyd, 1967). He cites a dozen authors from around 1900 until the 1930s, although some are only quoted by proxy. The later authors show remarkable ignorance of what went on before, and the logistic is rediscovered independently over and again, which leads Lloyd to deplore the 'utter lack of communication between the bulk of physiologists and biological scientists and the human demographers' (Lloyd, 1967, p. 103).

Among the early explorers, Lloyd pays special attention to Thorburn Bradford Robertson (1884–1930). Robertson read biology at Adelaide, and published his first papers about the growth of individual organisms even before obtaining his D.Sc. in 1909. Immediately afterwards he went to Berkeley and continued this line of research. He published tables for the graduation of the autocatalytic function, and used it in many articles about the growth of organisms, initially published in German scientific journals (until the World War intervened); two are quoted in Pearl & Reed (1920). A full account is given in Robertson's book of 1923, which lists thirty-seven papers by the author among its 500 references. Robertson uses the autocatalytic function freely, but he is more interested in the physiology of growth (and in the factors that determine it) than in curve-fitting. He studies the growth of such diverse organisms as man (unborn and born), cows (both Holstein and Jersey), white rats, Rhode Island fowl, sunflowers, and oats, the regeneration of tadpoles' tails and the growth of tumors. But by the time the book was published Robertson had returned to Adelaide (via Toronto), where he became Professor of Physiology (later Biochemistry) in 1919, and his attention had turned to other matters, notably the introduction of insulin in Australia (for which his Toronto connection must have been quite helpful). As far as I know he never returned to growth curves.

The best known work about the autocatalytic function as a model of human population growth is the paper by Pearl and Reed of 1920. They fit the curve to Census figures of the United States by making it pass through three points, find a good fit for the period from 1790 to 1910, and then boldly present an estimate of Ω of 197 millions. This again in retrospect compares badly with the present value of about 270 million, but at the time it attracted a good deal of interest. Pearl and Reed wrote two more articles about this piece of work (Pearl & Reed, 1922, 1923), and Pearl included it in several of his popular books (Pearl, 1922, 1924, 1925). Along with the pursuit of many other interests, Pearl and his collaborators in the next twenty years went on to apply the logistic growth curve to almost any living population from fruit flies to the human population of the French colonies in North Africa as well as to the growth of cantaloupes. Among these publications is an article by Reed, together with Berkson, on the application of the logistic curve to a number of autocatalytic chemical reactions (Reed & Berkson, 1929). We shall hear more about this co-author in the next section.

In 1920, Raymond Pearl (1879–1940) had just been appointed Director of the Department of Biometry and Vital Statistics at Johns Hopkins University, and Lowell J. Reed (1886–1966) was his deputy (and his successor when a few years later Pearl was promoted to Professor of Biology). Pearl was trained as a biologist and acquired his statistics as a young man by spending the year 1905–1906 in London with Karl Pearson (and later quarrelling with him). He became a prodigious investigator and a prolific writer on a wide variety of phenomena like longevity, fertility, contraception, and the effects of alcohol and tobacco consumption on health, all subsumed under the heading of human biology. During World War I Pearl worked in the US Food Administration, and this may account for his preoccupation with the food needs of a growing population in the 1920 paper. Reed, who was trained as a mathematician, made a quiet career in biostatistics; he excelled as a teacher and as an administrator, and was brought back in 1953 from retirement to serve as President of Johns Hopkins.

Pearl and Reed discovered Verhulst's work soon after their first paper of 1920. The immediate sequel, another joint paper of 1922, does not mention it, but

Verhulst's priority is acknowledged in a footnote in Pearl's book (Pearl, 1922), and, at greater length, in the third paper by Pearl and Reed of 1923. It is not clear how these authors came to know about Verhulst. In the 1923 paper, they call Verhulst's papers 'long since forgotten', except for a single article by Du Pasquier (1918), and they then go out of their way to criticize that author for an 'entirely unjustified and in practice usually incorrect modification' of Verhulst's formula, without substantiating this judgment (Pearl & Reed, 1923, p. 10). In fact Du Pasquier's paper is a harmless reflection on four mathematical theories of population, of a very formal and abstract character to the point of inanity. The four theories are ascribed to Halley, de Moivre, Euler and Verhulst, and these authors are briefly introduced; Halley, for example, as 'the famous astronomer', and Verhulst, rather oddly, as 'a Belgian who died in 1847' (Du Pasquier, 1918, p. 241, p. 242). No references are given.

Louis-Gustave Du Pasquier (1876–1957), Professor of Mathematics at the University of Neuchâtel, took his degrees in mathematics in Zürich, but followed courses in the social sciences as well when he spent the year 1900–1901 in Paris at a variety of academic institutions. He may well have read about Verhulst in Liagre or elsewhere in the French literature, but I have been unable to find any traces of this in his textbook of probability (Du Pasquier, 1926). As with Pearl and Reed, it is unclear how Du Pasquier learned about Verhulst, or, for that matter, how Pearl and Reed learned about Du Pasquier.[1]

The next important publication is Yule's Presidential Address to the Royal Statistical Society of 1925. Yule states expressly that he owes the reference to Pearl's book of 1922 (the two were lifelong friends), but he treats Verhulst much more handsomely than Pearl and Reed did, devoting an appendix to his work (Yule, 1925, pp. 4, 41–45). It would take until 1933 for Miner (a collaborator of Pearl) to pay tribute to Verhulst, if in an oblique way. Instead of reproducing at least one of Verhulst's papers, he gives a translation of Quetelet's obituary, and emphasizes Verhulst's Roman imbroglio by adding an extract from the memoirs of Queen Hortense de Beauharnais recording this episode.

After 1922, the rediscovered name *logistic* quickly caught on, and it was adopted by influential authors like Yule and Wilson, a mathematician and statistician at Harvard and life-long managing editor of the *Proceedings of the National Academy of Sciences* (see Wilson, 1925, for an example). The basic idea of the logistic growth curve is simple and effective, and it is used to this day to model population growth and market penetration of new products and technologies. The introduction of mobile telephones is an autocatalytic process, and so is the spread of many new products and techniques in industry. A number of these (and other) applications from the past as well as from more recent times are listed in Tsokos and DiCroce (1991). A remarkable generalization is given in the article of Emmens

[1] The Pearl Archives at the American Philosophical Society in Philadelphia contain Pearl's correspondence with several hundreds of individuals, but Du Pasquier is not among them; nor are Robertson and Ostwald.

(1941); the author uses the logistic to relate the growth of certain reproductive organs not to time, but to various doses of stimulantia. But while this demonstrates the use of the logistic as a dose-response curve, the response is still continuous, not discrete.

4. The introduction of the probit in bio-assay

The invention of the probit model is usually credited to Gaddum's report of 1933 and the papers of Bliss (1934a,b), but one look at the historical section of Finney's textbook or indeed at Gaddum's paper and his references will show that this is too simple (Finney, 1971, Ch. 3.6). The roots of the method and in particular the transformation of frequencies to equivalent normal deviates can be traced to the German scholar Fechner (1801–1887); Stigler recounts how Fechner was drawn to study human responses to external stimuli by experimental test of the ability to distinguish differences in weight (Stigler, 1986). The issue of the variability of human responses had been raised by astronomers, who relied on human observers of celestial phenomena and found that their readings showed much unaccountable variation. Fechner recognized that human response to an identical stimulus is not uniform, and he was the first to transform observed differences to equivalent normal deviates. The historical sketches of Finney (1971, Ch. 3.6), and of Aitchison & Brown (1957, Ch. 1.2), record a long line of largely independent rediscoveries of this approach that spans the seventy years from Fechner to the publications of Gaddum and Bliss in the early 1930s. Both the later authors regard the assumption of a normal distribution as commonplace, and attach more importance to the logarithmic transformation of the stimulus. Their papers contain no major innovations, but they mark the emergence of a standard paradigm of bio-assay and of a new terminology. Gaddum wrote a comprehensive and authoritative report with the emphasis on practical aspects of the experiments and on the statistical interpretation of bio-assay, giving several worked examples from the medical and pharmaceutical literature. Bliss published two brief notes in *Science*, introducing the term *probit*; he followed this up with a series of articles setting out the maximum likelihood estimation of the probit curve, in one instance with assistance from R. A. Fisher (see Bliss, 1935). Both Gaddum and Bliss set standards of estimation; until the 1930s this was largely a matter of ad hoc numerical and graphical adjustment of curves to categorical data.

John Henry Gaddum (1900–1965) studied medicine at Cambridge but failed in his final examinations. He turned to pharmacology and worked under Trevan at the Wellcome Laboratories, then transferred to the National Institute for Medical Research (where he wrote the 1933 report) before embarking on an academic career of professorships in pharmacology in Cairo, London and Edinburgh. He was elected to the Royal Society in 1945 and knighted in 1964. To this day the British Pharmacological Society awards an annual Gaddum Memorial Prize for pharmaceutical research. Charles Ittner Bliss (1899–1979) studied as an entomologist at Ohio State University and was a field worker with the US Department of

Agriculture until this employment was terminated in 1933. He then spent two years in London studying statistics with R. A. Fisher, and Fisher found him a job as a statistician in Leningrad where he lived from 1936 and 1938. The political conditions were not propitious for serious research. Bliss returned to the Connecticut Agricultural Experiment Station, combining his work as a practising statistician with a Lecturership at Yale from 1942 until his retirement. He played an important role in the founding of the Biometric Society.

Both Gaddum and Bliss adhere firmly to the classical model of bio-assay, where the stimulus is determinate and responses are random because of the variability of individual tolerance levels. Bliss introduced the term *probit* (short for *probability unit*) originally as a convenient *scale* for normal deviates, but abandoned this within a year in favour of a different definition which has since been generally accepted. For any (relative) frequency f there is an equivalent normal deviate Z such that the cumulative normal distribution at Z equals f; Z is the solution of

$$f = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\tilde{Z}} \exp\left(-\frac{u^2}{2\sigma^2}\right) \mathrm{d}u \tag{11}$$

and this can be read off from a table of the normal distribution. The probit of f is this equivalent normal deviate Z, or initially Z increased by 5; this ensures that the probit is almost always positive, which facilitates calculation. In the 1930s such additive constants were a common device. In the analysis, probits of relative frequencies or of probabilities f are linearly related to (the logarithm of the) stimulus.

The acceptance of the probit method was aided by the articles of Bliss, who published regularly in this field until the 1950s, and by Finney and others (Gaddum returned to pharmacology). The full flowering of this school probably coincides with the first edition of Finney's monograph in 1947. Without the underlying theory of bio-assay, probit analysis was quickly used for any relation of a discrete binary outcome to one or more determinants. In economics and market research, for example, the first applications appear already in the 1950s: In 1954, Farrell used a probit model for the ownership of cars of different vintage as a function of household income, and in 1958 Adam fitted lognormal demand curves to survey data of the willingness to buy cigarette lighters and the like at various prices (Farrel, 1954; Adam, 1958). The classic monograph of 1957 on the lognormal distribution of Aitchison and Brown brought probit analysis to the notice of a wider audience of economists.

5. The advent of the logit

As we have indicated earlier the shape of the logistic curve closely resembles the normal probability distribution function. In the 1920s this was common knowledge among practitioners; it was demonstrated in 1925 by Wilson, and later written up by Winsor (another collaborator of Pearl) (Winsor, 1932). Wilson was probably the first to publish an application of the logistic curve as a substitute for the normal in bio-assay in 1943, just before the seminal paper of Berkson (Wilson &

Worcester, 1943; Berkson, 1944). Still the introduction of the logit as an alternative for the probit is very much the work of Berkson, and of Berkson alone, for it was Berkson who persisted and fought a long and spirited campaign which lasted for several decades.

In his 1944 paper, Berkson proposed the use of the logistic instead of the normal probability function in bio-assay, and coined the term *logit* by analogy to the *probit* of Bliss (for which he was initially much derided). By the inverse of the logistic function (1) we immediately have

$$\text{logit }(P) = \log\{P/(1\text{-}P)\} = Z \tag{12}$$

and this is of course much simpler than the definition of the probit of (11). At the time the simplicity of the computations was still of some importance. The issue of logit versus probit was tangled by Berkson's simultaneous attacks on the method of maximum likelihood, and his advocacy of minimum chi-squared estimation instead. Between 1944 and 1980 he wrote a large number of papers on both issues; examples are Berkson (1951, 1980). He often adopted a somewhat provocative style, and much controversy ensued.

Joseph Berkson (1899–1982) read physics at Columbia, then went to Johns Hopkins for his M.D. and a doctorate in statistics in 1928. He stayed on as an assistant for three years, and this is when he was co-author of Reed of a paper on autocatalytic functions (Reed & Berkson, 1929). Berkson then moved to the Mayo Clinic where he remained for the rest of his working life as chief statistician. In the 1930s he published numerous papers on medical and public health matters, but from 1944 on he turned his attention to the statistical methodology of bio-assay.

Berkson's suggestion was not well received by the biometric establishment. In the first place, the logit was regarded as somewhat inferior and disreputable because unlike the probit it can not be related to an underlying (normal) distribution of tolerance levels. Aitchison and Brown in 1957 dismiss the logit in a single sentence, because it 'lacks a well-recognized and manageable frequency distribution of tolerances which the probit curve does possess in a natural way' (Aitschison & Brown, 1957, p. 72). Berkson had tried to remedy this defect by adapting the autocatalytic argument, but this did not convince as the argument essentially deals with a process over time (see Berkson, 1951). In retrospect it is surprising that so much importance was attached to these somewhat ideological points of interpretation; at the time no one, not even Berkson, seems to have recognized the formidable power of the logistic's analytical properties. In the second place, Berkson's case for the logit was not helped by his simultaneous attacks on the established wisdom of maximum likelihood estimation and his advocacy of minimum chi-squared. The unpleasant atmosphere in which this discussion was conducted can be gauged from the acrimonious exchanges between R. A. Fisher and Berkson in the discussion following a paper by the former (Fisher, 1954).

In the practical aspect of ease of computation the logit had a clear advantage over the probit, even with maximum likelihood estimation. To quote from Cochran's contribution to the same discussion, 'the speed with which a new technique becomes widely used is considerably influenced by the simplicity or otherwise of the calculations that it requires. Next door to the lecture room in which the probit method is expounded one may still find the laboratory in which the workers compute their LD 50s by the [much less sophisticated] Behrens (Reed-Muench) method' (Fisher, 1954, p.147). On this count the logit spread much more quickly in workfloor practice than in the academic discourse. At the time, all empirical work was done by hand, that is with pencil and paper, supported rather feebly by the slide rule and by mechanical calculating machines, driven by hand or powered by a small electric motor, which were capable of addition and multiplication. Many analyses relied on freehand curves, drawn by hand; for bio-assay, grouped data or class frequencies could be plotted on graph paper with a special grid, on which a probit or logit curve would appear as a straight line: Wilson had introduced the logistic (or autocatalytic) grid in 1925, examples of lognormal paper can be found in Aitchison and Brown's book and Adam's study (Aitchison & Brown, 1958; Adam, 1958)[2], and Berkson himself designed logistic graph paper as well as several nomograms.[3] For numerical work, the values of the normal distribution (and of exponentials and logarithms) were obtained from printed tables like Pearson's *Biometrika tables* (Pearson & Hartley, 1954) or the *Statistical tables* of Fisher and Yates (1938). The latter carried specially designed tables for probit analysis (with auxiliary tables contributed by Bliss and by Finney) from their first edition of 1938, and from the fifth edition of 1957 onwards also included special tables for logit analysis.

6. The sequel

Between 1950 and 1970 the logit gradually gained ground on the probit in bio-assay, and in time the ideological conflict abated. Finney, who had ignored the logit in the second edition of his textbook of 1952, made amends in the third edition of 1971, recognizing that 'what matters is the dependence of P on dose and the unknown parameters, and the tolerance distribution is merely a substructure leading to this' (Finney, 1971, p. 47). But by 1971 the narrow conflict between probit and logit in bio-assay had long been overtaken by independent developments in other fields, where the superior analytical properties of the logit had soon been recognized. From 1960 onwards the logit terminology and the logit transformation (12) were widely adopted outside biology, and their origins forgotten. These developments took place in mathematical statistics, in epidemiology, in medicine and in the social sciences and economics. Once more there was not much communication between the various disciplines, and many statistical properties and generalizations have been

[2] Finney (1971), p. 42, traces the invention of the probability grid to Henry, a French artilleryman of the 1890s.

[3] A nomogram is a graph from which one may read off a transformation. More sophisticated nomograns may permit the quick solution of more complicated equations.

discovered independently and separately by practitioners of these different disciplines. But our account stops here, and we shall not attempt to trace the history of the logit transform beyond its adoption in bio-assay.

Acknowledgements

This is an expanded version of 'The origins and development of the logit model', previously published as Chapter 9 in J. S. Cramer, *Logit models from economics and other fields* (Cambridge University Press, 2003), reproduced with permission.

For useful comments on earlier drafts and ready help in obtaining valuable information about Verhulst, Pearl and Du Pasquier, I thank John Glaus (Maine), Jan Sandee (Boskoop), Ida Stamhuis (Amsterdam), Professor G. Vanpaemel (Gent) and Professor J. Aghion (Liège). I am greatly indebted to the American Philosophical Society, Philadelphia, for permission to consult the Pearl Archives, which contain the correspondence and pocket diaries of Raymond Pearl, and to Robert Cox for his help; to Professor Remy Scheurer, who provided invaluable material about Du Pasquier from the Archives of the University of Neuchâtel, and to François Sigrist (also from Neuchâtel) for consulting the University Library on my behalf; to Michel Guillot (Paris) and Anton Barten (Leuven) who kindly assisted me in consulting Du Pasquier's books at the Bibliothèque Nationale and the University Library, Leuven. I am also grateful to Mary Morgan, July Klein and an unknown referee for helpful suggestions.

Bibliographical Sources

I have consulted a number of biographies, obituaries and other sources about Ostwald, Robertson, Pearl, Reed, Wilson, Bliss, Gaddum and Berkson. Apart from the website of the University of Adelaide these are the following:

Armitage, P., & Colton, T. (Eds.) (1998). Joseph Berkson 1899–1982. In *Encyclopedia of Biostatistics, Vol. I* (pp. 290-300). New York: Wiley.

Cochran, W. G. (1967). Lowell Jacob Reed. *Journal of the Royal Statistical Society, Series A*, *130*, 279-281.

Cochran, W. G. (1979). Chester Ittner Bliss, 1899–1979. *Biometrics*, *35*, 715-717.

Feldberg, W. (1967). John Henry Gaddum, 1900–1965. *Biographical Memoirs of Fellows of the Royal Society*, *13*, 57-77.

Jennings, H. S. (1941). Raymond Pearl, 1879–1940. *Biographical Memoirs of the National Academy of Sciences*, *22(*14), 295-347.

Miner, J. R., & Berkson, J. (1940). Raymond Pearl, 1879–1940. *The Scientific Monthly*, *52*, 1092-1094.

Ostwald, W. (1933). *Lebenslinien. Eine Selbstbiographie*. Berlin: Klasing.

Salsburg, D. (2001). *The lady tasting tea*. New York: Holt. (About Bliss)

Taylor, W. F. (1983) Joseph Berkson, 1899–1982. *Journal of the Royal Statistical Society, Series A*, *146*, 413-419.

Worcester, J. (1965). Edwin Budwell Wilson, 1879–1865. *The American Statistician*, *19*, 52-53.

References

Adam, D. (1958). *Les réactions du consommateur devant les prix*. Observation Economique, no. 15. Paris: Sedes.

Aitchison, J., & Brown, J. A. C. (1957). *The lognormal distribution*. University of Cambridge, Department of Applied Economics Monographs, no. 5. Cambridge: Cambridge University Press.

Berkson, J. (1944). Application of the logistic function to bio-assay. *Journal of the American Statistical Association*, *39*, 357–365.

Berkson, J. (1951). Why I prefer logits to probits. *Biometrics*, *7*, 327–339.

Berkson, J. (1980). Minimum chi-square, not maximum likelihood!. *Annals of Mathematical Statistics*, *8*, 457–487.

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

Bliss, C. I. (1934b). The method of probits. *Science*, *79*, 409–410.

Bliss, C. I. (1935). The calculation of the dosage-mortality curve (with an appendix by R. A. Fisher). *Annals of Applied Biology*, *22*, 134–167.

David, H. A., & Edwards, A. W. F. (2001). *Annotated readings in the history of statistics*. New York: Springer.

Du Pasquier, L. G. (1918). Esquisse d'une nouvelle théorie de la population. *Vierteljahrsschrift der Naturforschenden Gesellschaft in Zürich*, *63*, 236–249.

Du Pasquier, L. G. (1926). *Le calcul des probabilités*. Paris: Hermann.

Emmens, C. W. (1941). The dose/response relation for certain principles of the pituitary gland, and the serum and urine of pregnancy. *Journal of Endocrinology*, *2*, 194–225.

Farrell, M. J. (1954). The demand for motorcars in the United States. *Journal of the Royal Statistical Society, Series A*, *117*, 171–200.

Finney, D. (1971). *Probit analysis*. (3rd ed.). Cambridge: Cambridge University Press. (First published 1947)

Fisher, R. A., & Yates, F. (1938). *Statistical tables for biological, agricultural and medical research*. (1st ed.). Edinburgh: Oliver and Boyd. (6th ed., 1963)

Fisher, R. A. (1954). The analysis of variance with various binomial transformations (with comments by M. S. Bartlett, F. J. Anscombe, W. G. Cochran, & J. Berkson). *Biometrics*, *10*, 130–151.

Gaddum, J. H. (1933). *Reports on Biological Standard III. Methods of biological assay depending on a quantal response*. Special Report Series of the Medical Research Council, no. 183. London: Medical Research Council.

Lloyd, P. J. (1967). American, German and British antecedents to Pearl and Reed's logistic curve. *Population Studies*, *21*, 99–108.

Meadows, D. H., Meadows, D. L., Randers, J., & Behrens, W. W. (1972). *The limits to growth: A report for the Club of Rome's project on the predicament of mankind*. New York: Universe Books.

Ostwald, W. (1883). Studien zur chemischen Dynamik. *Journal fuer Praktische Chemie*, *136*(n.s. 27, 28), 1–39, 449–495.

Pearl, R. (1922). *The biology of death*. Phildelphia: Lippincott.

Pearl, R. (1924). *Studies in human biology*. Baltimore: Williams and Wilkins.

Pearl, R. (1925). *The biology of population growth*. New York: Knopf.

Pearl, R., & Reed, L. J. (1920). On the rate of growth of the population of the United States since 1870 and its mathematical representation. *Proceedings of the National Academy of Sciences*, *6*, 275–288.

Pearl, R., & Reed, L. J. (1922). A further note on the mathematical theory of population growth. *Proceedings of the National Academy of Sciences*, *8*, 365–368.

Pearl, R., & Reed, L. J. (1923). On the mathematical theory of population growth. *Metron*, *5*, 6–19.

Pearson, E. S., & Hartley, H. O. (Eds.). (1954). *Biometrika tables for statisticians*. Cambridge: published for the Biometrika Trustees at the University Press. (Rev. ed. of Pearson, 1914)

Quetelet, A. (1848). *Le système social et les lois qui le régissent*. Paris: Guillaume.

Quetelet, A. (1850). Notice sur Pierre-François Verhulst. *Annuaire de l'Académie Royale des Sciences, Lettres et des Beaux-arts*, *16*, 97–124.

Reed, L. J., & Berkson, J. (1929). The application of the logistic function to experimental data. *Journal of Physical Chemistry*, *33*, 760–779.

Stigler, S. M. (1986). *The history of statistics*. Cambridge, MA: Harvard University Press.

Tsokos, C. P., & DiCroce, P. S. (1991). Applications in health and social sciences. In N. Balakrishnan (Ed.), *Handbook of the logistic distribution* (pp. 449–494). New York: Marcel Dekker.

Vanpaemel, G. (1987). Quetelet en Verhulst over de mathematische wetten van de bevolkings-groei. *Academiae Analecta, Mededelingen van de Koninklijke Academie voor Wetenschappen, Letteren en Schoone Kunsten van België*, *49*, 99–114.

Wilson, E. B. (1925). The logistic or autocatalytic grid. *Proceedings of the National Academy of Sciences*, *11*, 451–456.

Wilson, E. B., & Worcester, J. (1943). The determination of L.D.50 and its sampling error in bio-assay. *Proceedings of the National Academy of Sciences*, *29*, 79–85. (First of a series of three articles)

Winsor, C. P. (1932). A comparison of certain symmetrical growth curves. *Journal of the Washington Academy of Sciences*, *22*, 73–84.

Yule, G. U. (1925). The growth of population and the factors which control it. *Journal of the Royal Statistical Society*, *138*, 1–59.

Further reading

De Beauharnais, H. (1927). *Mémoires de la Reine Hortense, publiées par le prince Napoléon* (3 vols.). Paris: Plon.

Fechner, G. T. (1860). *Elemente der Psychophysik*. Leipzig: Breitkopf und Härtel.

Liagre, J. B. J. (1852). *Calcul des probabilités et théorie des erreurs*. Bruxelles: Société pour l'emancipation intellectuelle A. Jamard.

Liagre, J. B. J. (1879). *Calcul des probabilités et théorie des erreurs* (2e éd., revue par Camille Peney). Bruxelles, Paris: Muquardt, Gauthier-Villars.

Malthus, T. E. (1798). *An essay on the principle of population*. London.

Miner, J. R. (1933). Pierre-François Verhulst, the discoverer of the logistic curve. *Human Biology*, *5*, 673–689.

Pearl, R. (1927). The indigenous population of Algeria in 1926. *Science*, *66*, 593–594.

Pearl, R., Reed, L. J., & Kish, J. F. (1940). The logistic curve and the census count of 1940. *Science*, *92*, 486–488.

Pearl, R., Winsor, C. P., & White, F. B. (1928). The form of the growth curve of the cantaloupe (Cucumis melo) under field conditions. *Proceedings of the National Academy of Sciences*, *14*, 895–901.

Pearson, K. (1914). *Tables for statisticians and biometricians*. Cambridge: Cambridge University Press.

Robertson, T. B. (1923). *The chemical basis of growth and senescence*. Philadelphia: Lippincott.

Verhulst, P.-F. (1838). Notice sur la loi que la population suit dans son accroissement. *Correspondance Mathématique et Physique publiée par A. Quetelet*, *10*, 113.

Verhulst, P.-F. (1845). Recherches mathématiques sur la loi d'accroissement de la population. *Nouveaux Mémoires de l'Académie Royale des Sciences, des Lettres et des Beaux-Arts de Belgique*, *18*, 1–32.

Verhulst, P.-F. (1847). Deuxième mémoire sur la loi d'accroissement de la population. *Nouveaux Mémoires de l'Académie Royale des Sciences, des Lettres et des Beaux-Arts de Belgique*, *20*, 1–32.