TI 2002-119/4
Tinbergen Institute Discussion Paper

The Origins of Logistic Regression

J.S. Cramer

Faculty of Economics and Econometrics, University of Amsterdam, and Tinbergen Institute.

The Origins of Logistic Regression

J.S. Cramer *

November 2002

Abstract

This paper describes the origins of the logistic function, its adoption in bio-assay, and its wider acceptance in statistics. Its roots spread far back to the early 19th century; the survival of the term *logistic* and the wide application of the device have been determined decisively by the personal histories and individual actions of a few scholars.

This is a much extended version of Chapter 9 of
Logit Models from Economics and Other Fields,
forthcoming at Cambridge University Press

*University of Amsterdam and Tinbergen Institute, Amsterdam; postal adress Baambrugse Zuwe 194, 3645 AM Vinkeveen, the Netherlands; e-mail address mars.cram@worldonline.nl. I had the benefit of comments of J. Sandee on an earlier version.

1 Introduction

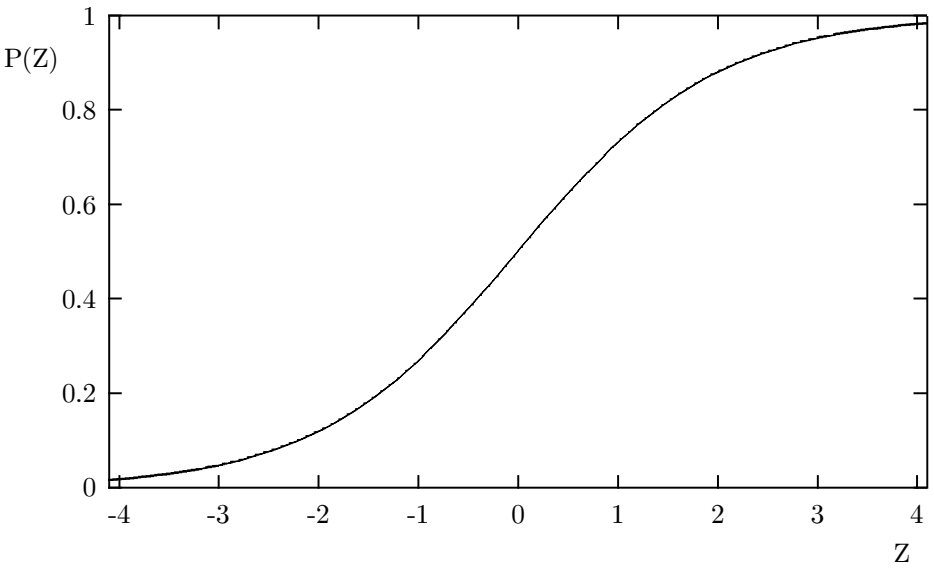

Figure 1. The logistic curve P(Z)

The sigmoid curve of Figure 1 is traced by the logistic function

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

P behaves like the distribution function of a symmetrical density, with midpoint zero; as Z moves through the real number axis, P rises monotonically between the bounds of zero and 1. The meaning of this function varies according to the the definition of the variables. In the logit version of bio-assay P is the probability of a binary outcome, and $Z = \alpha + \beta X$, with X a stimulus or exposure variable; α determines the location of the curve on the X-axis, and β its slope. In logistic regression there are several determinants of P, and $Z = x^T\beta$, with x a vector of covariates (including a unit constant) and β their coefficients. But the logistic function originally describes the course of a *proportion* P over *time* t, with $Z = \alpha + \beta t$; since $P(t)$ rises monotonically with t) this is a *growth curve*.

Over a fairly wide central range, for values of P from .3 to .7, the logistic curve closely resembles in shape as the normal probability distribution function. The two functions

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

and

$$P_n(x) = \frac{1}{\sigma\sqrt{2\pi}} \int_{-\inf}^{x} \exp\{-1/2(u/\sigma)^2\} du. \tag{3}$$

both pass through the point $(0, .5)$, 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

The logistic function was invented in the 19th century for the description of the growth of populations and the course of autocatalytic chemical reactions. In either case we consider the time path of a quantity $W(t)$ and its growth rate

$$\dot{W}(t) = \mathrm{d}W(t)/\mathrm{d}t. \tag{4}$$

The simplest assumption is that $\dot{W}(t)$ is proportional to $W(t)$

$$\dot{W}(t) = \beta W(t),\ \beta = \dot{W}(t)/W(t), \tag{5}$$

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

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

where A is sometimes replaced by the initial value $W(0)$. This is a reasonable model for unopposed population growth in a young country like the United States in its early years; as Malthus (1789) put it, a human population, left to itself, will increase *in geometric progression*.[1] But Alphonse Quetelet (1795–1874), the Belgian astronomer turned statistician, 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 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

$$\dot{W}(t) = \beta W(t) - \phi(W(t)). \tag{6}$$

[1] Two hundred years later exponential growth played a major part in the *Report to the Club of Rome* of Meadows et al (1972), and it still lies at the basis of many economic analyses.

and then experimenting with various forms of ϕ. The logistic appears when this is a quadratic, for then we may rewrite (6) as

$$\dot{W}(t) = \beta W(t)(\Omega - W(t)) \tag{7}$$

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

$$P(t) = \beta P(t)\{1 - P(t)\}, \tag{8}$$

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 suggestions 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 pages, followed by a demonstration 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 an update and that he publishes these notes only at the insistence of Quetelet. He does not say how he fitted the curves. The second paper, in the *Proceedings* 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 of 6.6 millions for Belgium and 40 million for France: at present these populations number 10.2 and 58.7 million. In 1847 there

followed a second paper in the *Proceedings*, 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.

Verhulst was primarily a mathematician, but sensitive to social and political issues; he was in poor health and died young. Quetelet attributes his early death to overwork and, rather curiously, to his great stature, as Verhulst was 1.89 meters or six feet tall. His impulsive nature was demonstrated in the summer of 1830. Verhulst had gone to Italy for his health and was staying in Rome when the news of the revolution in Paris and of the Belgian secession from the Netherlands broke. These events moved him 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.

The logistic function was discovered anew in 1920 by Pearl and Reed in a study of the population growth of the United States. They were unaware of Verhulst's work (though not of the curves for autocatalytic reactions discussed presently), and they arrived independently at the logistic curve of (10). When this was fitted to Census figures, again by making the curve pass through three points, it gave a good fit for the period from 1790 to 1910. But the estimate of Ω of 197 millions again compares badly with the present value of about 270 millions. In spite 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; we list a few of these studies in the bibliography.

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 in 1905–1906 by spending a year 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 U.S. 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. Among his publications in the aftermath of the 1920 paper with Pearl is an application of the logistic curve to autocatalytic reactions, Reed and Berkson (1929). We shall hear more about this co-author in the next section.

The term *logistic* was of course not used until Verhulst's work was rediscovered, which was soon after Pearl and Reed's first paper of 1920. The immediate sequel, Pearl and Reed (1922), does not mention Verhulst; his priority is first acknowledged in a footnote in Pearl (1922). In Pearl and Reed (1923) Verhulst is again named and references are given to his two papers of 1845 and 1847, but his terminology is not adopted. Pearl and Reed 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 harsh judgment. 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 linked to Halley, de Moivre, Euler and Verhulst, and these authors are briefly introduced; Halley, for example, as *"the famous astronomer"*, and Verhulst as *"a Belgian who died in 1847"*. No references are given. It is not clear how Du Pasquier knew about Verhulst, nor how Pearl and Reed knew about Du Pasquier. On the first point, Du Pasquier (1876 – ?), who was a Professor of Mathematics at the University of Neuchatel, in Switzerland, was educated in mathematics and economics in Zürich and Paris, and he may have read about Verhulst in the French literature. On the second point, Du Pasquier may well have taken the initiative in establishing contact with Pearl or Reed; after all, Pearl and Reed had published their paper in the *Proceedings of the National Academy of Sciences*, which would reach Switzerland, and Du Pasquier in the *Vierteljahrsschrift der Naturforschenden Gesellschaft in Zürich* which was unlikely to reach Baltimore. But all this is pure speculation.

The next important publication is Yule's Presidential Address to the Royal Statistical Society of 1925. Yule, who says he owes the reference to Pearl (1922), treats Verhulst much more handsomely than Pearl and Reed did, devoting an appendix to his work and reviving the name *logistic*. It would take until 1933 for Miner (a collaborator of Pearl) to pay tribute to

Verhulst, if in an indirect way: instead of reproducing at least one of Verhulst's papers, he published a translation of Verhulst's obituary by Quetelet, with the addition of an extract from the memoirs of Queen Hortense de Beauharnais, who records the episode of Verhulst's stay in Rome in 1830.

As we have indicated there is another early root of the logistic function in chemistry, where it was employed (again with some variations) to describe the course of autocatalytic reactions. These are chemical chain reactions where the product itself acts as a catalyst for the process while the supply of raw material is fixed. This leads naturally to a differential equation like (8) and hence to the logistic function for the time path of the amount of the reaction product. The review of the application of logistic curves to a number of such processes by Reed and Berkson (1929) quotes work of the German professor of chemistry Wilhelm Ostwald of 1883. Authors like Yule (1925) and Wilson (1925) were well aware of this strand of the literature.

The basic idea of logistic growth 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.

3 The invention of the probit and the advent of the logit

The invention of the probit model is usually credited to Gaddum (1933) and Bliss (1934a), (1934b), but one look at the historical section of Finney (1971) or indeed at Gaddum's paper and his references will show that this is too simple. 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 (1986) recounts how Fechner was drawn to study human responses to external stimuli by experimental test of the ability to distinguish differences in weight. 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 and Brown (1957), Ch. 1.2, record a long line of largely independent rediscoveries of this approach that spans the seventy years from Fechner (1860) to the early 1930's when Gaddum and Bliss published their

contributions. Both 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. 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 (Bliss (1935)). Both Gaddum and Bliss set standards of estimation; until the 1930's 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 he embarked 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 U.S. 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 in Leningrad where he lived from 1936 and 1938. The political conditions were not propitious for serious work. 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.

In their early writings on bio-assay both authors 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 was generally accepted. For any (relative) frequency f there is an equivalent normal deviate $\tilde{Z}$ such that the cumulative

normal distribution at $\tilde{Z}$ equals f; $\tilde{Z}$ is the solution of

$$f = \frac{1}{\sqrt{2\pi}} \int_{-\inf}^{\tilde{Z}} \exp\{-1/2u^2\} du,$$

and can be read of from a table of the normal distribution. The probit is the equivalent normal deviate increased by 5. This ensures that the probit is almost always positive, which facilitates calculation; at the time such additive constants were a common device.

The acceptance of the probit method was aided by the articles of Bliss, who published regularly in this field until the 1950's, and by Finney and others (Gaddum returned to pharmacology). The full flowering of this school in bio-assay probably coincides with the first edition of Finney's monograph in 1947. Applications in other fields like economics and market research appear already in the 1950's: Farrell (1954) employed a probit model for the ownership of cars of different vintage as a function of household income, and Adam (1958) fitted lognormal demand curves to survey data of the willingness to buy cigarette lighters and the like at various prices. The classic monograph on the lognormal distribution of Aitchison and Brown (1957) brought probit analysis to the notice of a wider audience of economists.

As far as I can see the introduction of the logistic as an alternative to the normal probability function is the work of a single person, namely Joseph Berkson (1899–1982), Reed's co-author of the paper on autocatalytic functions of 1929. Berkson 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 collaborated with Reed on autocatalytic functions. Berkson then moved to the Mayo Clinic where he remained for the rest of his working life as chief statistician. In the 1930's he published numerous papers on medical and public health matters, but in 1944 he turned his attention to the statistical methodology of bio-assay and proposed the use of the logistic, coining the term 'logit' by analogy to the 'probit' of Bliss (for which he was initially much derided). 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) and Berkson (1980). He often adopted a somewhat provocative style, and much controversy ensued.

The close resemblance of the logistic to the normal distribution function must have been common knowledge among those who were familiar with

the logistic; it had been demonstrated by Wilson (1925) and written up by Winsor (1932) (another collaborator of Pearl). Wilson was probably the first to publish an application of the logistic in bio-assay in Wilson and Worcester (1943), just before Berkson (1944). But it was Berkson who persisted and fought a long and spirited campaign which lasted for several decades.

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 (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"* (p.72). Berkson was aware of this defect and tried to remedy it by adapting the autocatalytic argument, in Berkson (1951), but this did not convince as the autocatalytic argument essentially deals with a process over time. In retrospect it is surprising that so much importance was attached to these somehwat 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 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 Cochran (from his comments on Fisher (1954), p.147.) *".. 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 .."*. On this count the logit spread much more quickly in workfloor practice than in the academic discourse. Until the advent of the computer and the pocket calculator, some trwenty years later, all numerical work was done by hand, that is with pencil and paper, sometimes aided by graphical inspection of 'freehand curves', 'fitted by eye'. For probit and logit analyses of grouped data or class frequencies there was graph paper with a special grid on which a probit or logit curve would appear as a straight line. Wilson (1925) had introduced the logistic

(or 'autocatalytic') grid, and examples of lognormal paper can be found in Aitchison and Brown (1957) and Adam (1958);[2] Berkson himself had designed logistic graph paper as well as several nomograms.[3] Numerical work was 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; punched card equipment was helpful if numerous data had to be analysed. Values of the normal distribution (and of exponentials and logarithms) were obtained from printed tables like Pearson's *Biometrika* Tables or the *Statistical Tables* of Fisher and Yates (1938). From the first edition the latter carried specially designed tables for probit analysis (with auxiliary tables contributed by Bliss and by Finney), but from the fifth edition of 1957 onwards they also included special tables for logit analysis.

In time, the ideological conflict over bio-assay abated. Finney, who had ignored the logit in the second edition of his textbook of 1952, made amends in the third edition of 1970, 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"*. Between 1960 and 1970 the logit indeed gradually achieved an equal footing with the probit. By then it was also slowly recognized that its analytical properties permit much wider statistical applications, beyond bio-assay: it can be linked to discriminant analysis, it leads to loglinear models, it can be used with retrospective samples as in case-control studies, and so on. One of the first to recognize and exploit these avenues was Cox, in a series of articles in the 1960's, and in Cox (1969). This general development is illustrated in Table 1, which is drawn from the JSTOR electronic repertory of twelve major statistical journals in the english language. The table show the number of articles which contain the word "probit" or "logit". The number of statistical journals included in JSTOR increases over time, as does the number of articles in each journal; from 1935 to 1985 the total number of articles covered annually increases about eightfold. It is therefore the *relative* position of "probit" and "logit" that counts. By 1970 logit reaches parity, and thereafter soars ahead.

[2] Finney (1947) traces the invention of the probability grid to a French artilleryman of the late 1890's.

[3] A nomogram is a graph from which one can read off a transformations, as from a table; sophisticated nomograms may permit the quick solution of more complicated equations.

Table 1. Number of articles in statistical journals containing the word 'probit' or 'logit'.

	probit	logit
1935 – 39	6	-
1940 – 44	3	1
1945 – 49	22	6
1950 – 54	50	15
1955 – 59	53	23
1960 – 64	41	27
1965 – 69	43	41
1970 – 74	48	61
1975 – 79	45	72
1980 – 84	93	147
1985 – 89	98	215
1990 – 94	127	311

Both probit and logit were also adopted beyond bio-assay, in economics, in epidemiology and in the social sciences. The close link to tolerance levels or threshold values was dissolved and less stringent interpretations were admitted; the elegant but quite abstract model of the latent regression equation was probably first explicitly formulated by McKelvey and Zavoina (1975) for an ordered probit model of the voting behaviour of U.S. Congressmen, far removed from bio-assay. Analyses linking binary discrete responses to several covariates became known as logistic regression. This wider acceptance was greatly helped by the advent of the computer and by the introduction of package routines for the maximum likelihood estimation of both logit and probit models from individual data. The BMDP or BIOMEDICAL DATA PROCESSING computer package of 1977 was probably the first to offer this facility, which soon became a standard feature of most statistical packages. By the time the first comprehensive textbook of Hosmer and Lemeshow (1989) appeared the use of such routines was taken for granted.

Of the two causes Berkson advocated, minimum chi-squared was thus overtaken by the computer revolution, but the logit was there to stay. Its multinomial generalization was first mooted by Cox (1966) and then, independently, by Theil (1969) who immediately saw its potential as a general approach to the modelling of shares. The simple algebra of this generalisation opened up a very wide field of applications in economics and other social sciences, and interest in an interpretation in terms of an underlying process

waned. But in 1973 McFadden, working as a consultant for a Californian public transportation project, first linked the multinomial logit to the theory of discrete choice from mathematical psychology. This provided a theoretical foundation of the logit model that is much more profound than any theory put forward for the use of the probit in bio-assay. It earned McFadden a Nobel prize in 2000.

References

Adam, Daniel (1958), *Les réactions du consommateur devant les prix*. Paris: Sedes.

Aitchison, J. and J.A. C. Brown (1957) *The Lognormal Distribution*. Cambridge: Cambridge University Press.

Berkson, J. (1944) Application of the Logistic Function to Bio-assay. *Journal of the American Statistical Association*, **9**, 357–365.

Berkson, J. (1951) Why I Prefer Logits to Probits. *Biometrics*, **9**, 357–365.

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.

Cox, David R. (1966) Some Procedures Connected with the Logistic Qualitative Response Curve. In: F.N. David (ed) *Research Papers in Statistics: Festschrift for J. Neyman*. London: Wiley.

Cox, David R. (1969) *Analysis of Binary Data*. London: Chapman and Hall.

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

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

Finney, D. (1971) *Probit Analysis*. Cambridge: Cambridge University Press. (third edition; first edition 1947)

Fisher, Sir Ronald (1954) The Analysis of Variance With Various Binomial Transformations (With Comments by M. S. Bartlett, F. J. Anscombe, W. G. Cochran and J. Berkson), *Biometrics*, **10**, 130–151.

Gaddum, J.H. (1933) *Report on Biological Standards III: Methods of Biological Assay Depending on Quantal Response*. Special Report Series of the Medical Research Council, no.183. London: Medical Research Council.

Hosmer, David W. and Stanley Lemeshow (1989) *Applied Logistic Regression* New York: Wiley.

Malthus, T.R. (1798) *An Essay on the Principle of Population.* London.

McFadden, Daniel (2001) Economic Choices. (Nobel prize acceptance speech.) *American Economic Review*, **91**, 352–370.

Meadows, D.H., Meadows, D.L., Randers J. and W.W. Behren (1972) *The Limits to Growth* New York: Universe Books.

Miner, J. R. (1933) Pierre-François Verhulst, the Discoverer of the Logistic Function. *Human Biology*, **5**, 673–689.

Pearl, Raymond (1922) *The Biology of Death.* Philadelphia: Lippincott.

Pearl, Raymond (1927) The Indigenous Population of Algeria in 1926. *Science* **66**, 593–594.

Pearl, Raymond (1939) *The Natural History of Population* Oxford: Oxford University Press.

Pearl, Raymond and Lowell J. Reed (1920) On the Rate of Growth of the Population of the United States and its Mathematical Representation. *Proceedings of the National Academy of Sciences*, **6**, 275–288.

Pearl, Raymond and Lowell J. Reed (1922) A Further Note on the Mathematical Theory of Population Growth. *Proceedings of the National Academy of Sciences*, **8**, 365–368.

Pearl, R., C.P. Winsor and F.B. White (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.

Pearl, R., L.J. Reed and J.F. Kish (1940) The Logistic Curve and the Census Count of 1940. *Science*, **92**, 486–488.

Reed, L.J., and J. Berkson (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, Mass.: Harvard University Press.

Theil, H. (1969) A multinomial extension of the linear logit model. *International Economic Review*, **10**, 251–259.

Verhulst, Pierre-François (1838) Notice sur la loi que la population suit dans son accroissement. *Correspondance mathématique et Physique, publiée par A. Quetelet*, **10**, 113–120.

Verhulst, Pierre-François (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**, 1–38.

Verhulst, Pierre-François (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.

Wilson, Edwin B. (1925) The Logistic or Autocatalytic Grid. *Proceedings of the national Academy of Sciences*, **11**, 431–456.

Wilson, E.B. and Jane Worcester (1943) The Determination of L.D.50 and Its Sampling Error in Bio-assay. *Proceedings of the National Academy of Sciences*, **29**, 79–85.

Winsor, C.P. (1932) A Comparison of Certain Symmetrical Growth Curves. *Proceedings of the Washington Academy of Sciences*, **22**, 73–84.

Yule, G. Udney (1925) The Growth of Population and the Factors which Control It. *Journal of the Royal Statistical Society*, **38**, 1–59.

Other Biographical Sources

On Pearl:

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

Miner, John R. and Joseph Berkson (1940) Raymond Pearl, 1879–1940. *The Scientific Monthly*, **52**, 1092–194.

On Reed:

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

On Gaddum:

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

On Berkson:

Armitage, P., and T. Colton (eds) (1998) Joseph Berkson 1899–1982. *Encyclopedia of Biostatistics*, volume I, 290–300. New York: Wiley.

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