SOME IMPORTANT DETAILS ON RICHARD GROWTH MODEL

by

Mehmet GURCAN [a*,**] ***and Arzu E. DEMIRELLI*** [b]

[a] Department of Statistics, Fırat University, Elazig, Turkey
[b] Department of Econometrics, Batman University, Batman, Turkey

Original scientific paper
https://doi.org/10.2298/TSCI190108350G

The distribution of the data is very important in all of the parametric methods used in the applied statistics. More clearly, if the experimental data fit well to the theoretical distribution, the results will be more efficient in parametric methods. The adaptability of experimental data to a theoretical distribution depends on the flexibility of the theoretical distribution used. If the flexibility of the theoretical distribution is sufficient, it can be used easily for experimental data. Most of the theoretical distributions have shape and location parameters. However, these two parameters are not always sufficient for the distribution adapt to the experimental data. Therefore, theoretical distributions with high flexibility in parametric methods are needed. Obtaining the new theoretical distributions that provide this feature is important for the literature. In this study, a new probability distribution has been obtained via Richard link function which has been high flexibility. In the introduction, important information is given related to growth models and Richard growth curve. Later, some details about the Richard distribution and wrapped distribution have been given.

Key words: *Richard growth curve, link function, wrapped distribution*

Introduction

The curves used in growth models are known as S-curves. In real applications, the basis of a growth model is examined in two forms. The first is the modelling of growth data observed associated with time and the second is the modelling of growth data observed as a result of the effect of a set of independent variables.

In the growth data observed associated with time, the organism observed or the size of the involved organism is taken into consideration. In the development of healthy children, variables such as height and weight may be seen as examples of this type of data [1-3]. The examination of increases occurring as the result of a set of independent variables is encountered more often in pharmacological research [4-7]. For example, the effective substance of a drug may produce changes observed in the patient related to the drug administered or to absorption in the blood which is dose-related.

In addition, growth curves are often used in the applications of logistic regression. In the logistic differentiation of two groups, the link functions are extremely important in the mapping of dependent variables which have taken two values such as zero and one. The flexibility of the curve is of great importance in the structuring of the curve appropriate to the data [8, 9]. Therefore, it is important to use more than one parameter which will increase the flexibility of the link functions used in both growth curves and logistic regression analysis.

[*] Corresponding author, e-mail: mgurcan@firat.edu.tr

[**] This article is produced from doctorate thesis

Natural growth cannot be expected to continue forever. However, it should be noted that associated with the structure, the growth of S-curves is greater at first then by slowing, becomes fixed. In this case, there must be an inflection point of the curve. Generally, in logistic differentiation problems, the differentiation is made based on this inflection point.

While the S-curve shows an increasing figure in some models, it stops at certain points. The definition of these points is of great importance in the examination of growth. These points, which cannot be clearly identified from the data, can be estimated by examining the appropriate growth model.

The most important method used in the estimation of model parameters is the least squares method. According to the Gauss-Markov theorem, the least squares estimate of model parameters is the estimate of minimum error. This theorem was proven by Johann Carl Friedrich Gauss and Andrey Markov. The German mathematician, Gauss was one of the most remarkable scientists of his age with many contributions to the field of mathematics. Andrey Markov also served in many areas of science throughout his life. The least squares method, which is the best estimate for linear models, is the currently most used most effective method.

The content of the method is based on the solution of the system of linear equations having non-quadratic coefficients matrix by converting the matrix of coefficients into quadratic form. In some applications, the matrix of coefficients converted to a quadratic form can be singular. In this case, the Ridge method can be used to solve the equation.

Least squares method is used in both linear models and non-linear models. When the model is non-linear, the first two terms that are linear are used by expanding the model equation the Taylor series. The relevant literature can be found in references [10, 11].

Richard distribution

The Richard link function can be written for $\alpha = 1$ as the following form:

$$F(x) = \left\{1 + (m-1)\exp(-kx)\right\}^{1/(1-m)}$$

where α is the upper limit of the growth curve. Detailed information on this function can be found in references [3, 12].

In this case, the first derivative of the link function will be:

$$\frac{\mathrm{d}F(x)}{\mathrm{d}x} = k\left\{1 + (m-1)\exp(-kx)\right\}^{m/(1-m)}\exp(-kx)$$

Now, to be the density function on the positive half-axis of this function, the following condition will provide the form by $C = k\{1 - m^{1/(1-m)}\}^{-1}$:

$$\int_0^\infty C\,k\left\{1 + (m-1)\exp(-kx)\right\}^{m/(1-m)}\exp(-kx)\mathrm{d}x = 1$$

In this case, the Richard density function can be written:

$$f(x) = k\left\{1 - m^{1/(1-m)}\right\}^{-1}\left\{1 + (m-1)\exp(-kx)\right\}^{m/(1-m)}\exp(-kx)$$

Consequently, X~Richard(k, m, x) Richard distribution function can be written:

$$F(x) = \left[1 - m^{1/(1-m)}\right]^{-1}\left\{\left[1 + (m-1)\exp(-kx)\right]^{1/(1-m)} - m^{1/(1-m)}\right\}$$

In particular cases of this distribution function, for $m = 0$, it will become an exponential distribution function with parameter k.

The expected value of Richard distribution:

$$EX = \int_0^{\infty} \left[1 - F(x)\right] \mathrm{d}x$$

However, the aforementioned integral cannot be calculated analytically. Finding the mean of this distribution be able to apply statistical analysis requires great care. This problem is handled with the definition of Richard distribution in sub-interval.

Now, with $\alpha_1 T \le X \le \alpha_2 T$, the density and the distribution function of X~Richard ($\alpha_1 T$, $\alpha_2 T$, k, m, x) distribution is obtained. For this, the value of the integral is calculated:

$$\int_{\alpha_1 T}^{\alpha_2 T} k\exp(-kx)\left\{1+(m-1)\exp(-kx)\right\}^{m/(1-m)} \mathrm{d}x = c_2^{1/(1-m)} - c_1^{1/(1-m)}$$

Here, c_1 and c_2 are reel constant:

$$c_1 = 1+(m-1)\exp(-k\alpha_1 T)$$

$$c_2 = 1+(m-1)\exp(-k\alpha_2 T)$$

In this case, the density function is written:

$$f(x) = \frac{k\exp(-kx)}{c_2^{1/(1-m)} - c_1^{1/(1-m)}}\left\{1+(m-1)\exp(-kx)\right\}^{m/(1-m)}$$

From this, the distribution function can be written:

$$F(x) = \frac{1}{c_2^{1/(1-m)} - c_1^{1/(1-m)}}\left\{\left[1+(m-1)\mathrm{e}^{-kx}\right]^{1/(1-m)} - \left[1+(m-1)\mathrm{e}^{-k\alpha_1 T}\right]^{1/(1-m)}\right\}$$

Under the conditions $\alpha_1 = 0$ and $\alpha_2 \longrightarrow \infty$, variable X has Richard (k, m, x) distribution. The expected value of the Richard distribution defined in a sub-interval:

$$EX = \frac{k}{c_2^{1/(1-m)} - c_1^{1/(1-m)}} \int_{\alpha_1 T}^{\alpha_2 T} x\mathrm{e}^{-kx}\left\{1+(m-1)\exp(-kx)\right\}^{m/(1-m)} \mathrm{d}x$$

If partial integration is used, the integral is in the form with $C = c_2^{1/(1-m)} - c_1^{1/(1-m)}$,

$$EX = \frac{1}{C}\left\{\alpha_2 T c_2^{1/(1-m)} - \alpha_1 T c_1^{1/(1-m)} - \int_{\alpha_1 T}^{\alpha_2 T}\left\{1+(m-1)\mathrm{e}^{-kx}\right\}^{1/(1-m)} \mathrm{d}x\right\}$$

If the integral of the expected value is calculated numerically, it can be obtained:

$$I = \frac{1}{n}\sum_{j=1}^{n}\left\{1+(m-1)\exp(-k)\left[\alpha_1 T + \frac{jT}{n}(\alpha_2 - \alpha_1)\right]\right\}^{1/(1-m)}, \; n \to \infty$$

Now, using the relationship between the density and distribution function of the radom variable X~Richard(k, m, x), the following theorem can be given.

Theorem 1. Let the density and distribution functions of the Richard distribution defined in the positive half-axis be $f(x)$ and $F(x)$, respectively:

$$f(x) = \frac{k}{(m-1)\left(1-m^{1/(1-m)}\right)}\left\{u(x) - u^m(x)\right\}$$

$$u(x)=\left[1-m^{1/(1-m)}\right]F(x)+m^{1/(1-m)}$$

The expected value of Richard distribution defined in the positive half-axis can be expressed:

$$EX=\frac{(1-m)}{k}+\frac{1}{1-m^{1/(1-m)}}\int_0^\infty\left[1-u^m(x)\right]\mathrm{d}x$$

Proof. Since the integral of the density function makes 1, the following equation can be written:

$$\int_0^\infty\left[u(x)-u^m(x)\right]\mathrm{d}x=\frac{(m-1)\left[1-m^{1/(1-m)}\right]}{k}$$

The distribution function $F(x)$ can be written in terms of $u(x)$:

$$F(x)=\frac{u(x)-m^{1/(1-m)}}{1-m^{1/(1-m)}}$$

According to this, the $1-F(x)$ expression can be written:

$$1-F(x)=\frac{1-u(x)}{1-m^{1/(1-m)}}$$

In this case, the expected value of the distribution is written:

$$EX=\frac{1}{1-m^{1/(1-m)}}\int_0^\infty\left[1-u(x)\right]\mathrm{d}x$$

The equation below can be written with the help of the integral of the differences of the $u(x)$ and $u^m(x)$ functions:

$$\left[1-m^{1/(1-m)}\right]EX-\int_0^\infty\left[1-u^m(x)\right]\mathrm{d}x=\frac{(1-m)\left(1-m^{1/(1-m)}\right)}{k}$$

According to this, the desired result will be obtained:

$$EX=\frac{1}{1-m^{1/(1-m)}}\left\{\frac{(1-m)\left(1-m^{1/(1-m)}\right)}{k}+\int_0^\infty\left[1-u^m(x)\right]\mathrm{d}x\right\}$$

From the obtained equation, for $m = 0$, the expected value of the exponential distribution with Exp(k) parameter can be easily obtained as $1/k$. The same procedures can be applied for Richard distributions defined in sub-intervals.

Division according to the stopping times of the growth curve

As the Richard link function is simultaneously a growth curve model, there is no inflection at any point. However, there may be inflection at the Richard distribution starting point defined in the positive half-axis. Due to this feature, Richard distribution defined in sub-intervals may cover the whole positive axis by subsequent additions. Naturally, by adding subsequent distributions in sub-intervals, a new growth model is formed and the structure of the distribution function will be damaged but the probability will be provided to be able to place stopping points in the growth curve model. Similar studies can be found in references [13, 14].

Let the stopping points determined in the positive half-axis of the growth curve be t_1, t_2,... Let the defined sub-intervals of Richard distribution be taken as $A_j = [\alpha_j T, \alpha_{j+1} T]$, $j = 0, 1...$

In this case, the first stopping point and subsequent stopping points will be equivalent to the three points of the sub-intervals:

$$\alpha\ T = 0\ ,\ \alpha_j T = t_j\ ,\ \ j = 1, 2, ...$$

Let F_j be the distribution function of Richard distribution defined in the sub-interval A_j, $j = 0, 1,...$ The growth curve model can be written:

$$F(x) = \sum_j \left[j + F_j(x) \right] I(A_j)$$

where $I(A_j)$ is the characteristic function of the sub-interval A_j.

This growth curve model obtained in a similar way can also be used for the examination of periodic data associated with time. However, in this case, at the determined breaking points of the series, the generated kernel model should be used as $1 - F$ in decreasing trend intervals. When increasing and decreasing trends are considered together, the kernel model can be given:

$$1 - \alpha + (2\alpha - 1)F$$

where

$$\begin{cases} \alpha = 0 & \text{, decreasing trend} \\ \alpha = 1 & \text{, increasing trend} \\ \alpha = 1/2 & \text{, stationary trend} \end{cases}$$

The stationary of the series can be tested when the distribution of α is known.

Wrapped Richard distribution

When the observed values are periodic, the distribution be used for the data is the circular distribution. Circular distributions are one of the most important subjects in recent years. The first definition related to this subject is the definition of the wrapped density function. Let $X > 0$ be a random variable with a period length of $[0.2\pi]$. The integral of the density function must be one on the positive half-axis. According to this, if the positive half-axis is separated into parts of 2π length, the equation can be written:

$$[0.\infty) = [0.2\pi) + [2\pi, 4\pi) + \cdots = \sum_{k=0}^{\infty} \left[2k\pi, (2k+2)\pi \right)$$

In this case, the integral of density function in the positive half-axis can be written as the total of the integrals of the sub-intervals:

$$\int_0^{\infty} f(x)\mathrm{d}x = \sum_{k=0}^{\infty} \int_{k\pi}^{(2k+2)\pi} f(x)\mathrm{d}x = \sum_{k=0}^{\infty} \int_0^{2\pi} f(\varphi + 2k\pi)\mathrm{d}\varphi = 1$$

The wrapped density function obtained from this result can be defined:

$$f_w(\varphi)=\sum_{k=0}^{\infty} f(\varphi+2k\pi)$$

When it is desired to define this density function in all real axis, the following equation can be written:

$$f_w(\varphi)=\sum_{k=-\infty}^{\infty} f(\varphi+2k\pi)$$

Adapting the density function the periodic data and restricting it to the [–π, π] interval are different processes. In the first, density function is fragmented according to the length of the period while in the second, by using a link function, the interval is restricted. The most important link function is written:

$$X=\frac{\cos\varphi}{1-\sin\varphi}$$

The previous transformation is a topological homeomorphism.

Even if the applied method seems appropriate for each density function, it should conform to the wrapped distribution of the observed values. Therefore, distributions such as uniform, normal, Cauchy which are frequently used in applications have wrapped form.

The density functions of wrapped uniform, normal and Cauchy distributions, respectively are given:

$$WU(\varphi)=\frac{1}{2\pi}$$

$$WN(\varphi,\mu,\sigma)=\frac{1}{\sigma\sqrt{2\pi}}\sum_{k=-\infty}^{\infty}\exp\left[\frac{-(\varphi-\mu-2k\pi)^2}{2\sigma^2}\right]$$

$$WC(\varphi,\varphi_0,\gamma)=\sum_{k=-\infty}^{\infty}\frac{\gamma}{\pi\left[\gamma^2+(\varphi+2k\pi-\varphi_0)^2\right]}=$$

$$=\frac{1}{2\pi}\frac{\sinh\gamma}{\cosh\gamma-\cos(\varphi-\varphi_0)}$$

In addition, one of the most widely used of the circular distributions is the von Mises distribution. This distribution is obtained with the help of the feature of the Bessel function. The density function of the distribution:

$$Wf(x)=\frac{\exp\left[K\cos(x-\mu)\right]}{2\pi I_0(K)}$$

where $I_0(K)$ is the modified Bessel function with zero degrees.

In this section, we will obtain the wrapped form of Richard distribution. Let $X>0$ be a random variable with period length of $[0.2\pi]$. When Richard density function is $f(x)$, wrapped density function will be written:

$$Wf(\theta)=\sum_{k=0}^{\infty}\frac{k\mathrm{e}^{-k(\theta+2n\pi)}}{1-m^{1/(1-m)}}\left[1+(m-1)\mathrm{e}^{-k(\theta+2n\pi)}\right]^{m/(1-m)}$$

$$F(x)=\sum_{j}\left[j+F_j(x)\right]I\left(A_j\right)$$

and the integral in the range [0.2π]. of the wrapped Richard density function can be obtained:

$$I=\int_{0}^{2\pi}Wf(\theta)\,\mathrm{d}\theta=$$

$$=\sum_{n=0}^{\infty}\frac{\left\{\left[1+(m-1)\mathrm{e}^{-(n+1)2k\pi}\right]^{1/(1-m)}-\left[1+(m-1)\mathrm{e}^{-2nk\pi}\right]^{1/(1-m)}\right\}}{1-m^{1/(1-m)}}=$$

$$=\frac{1}{1-m^{1/(1-m)}}\lim_{n\to\infty}\left\{\left[1+(m-1)\mathrm{e}^{-(n+1)2k\pi}\right]^{1/(1-m)}-m^{1/(1-m)}\right\}=1$$

Wrapped Richard distribution shows structural similarities to Richard distributions defined in sub-intervals. When each sub-interval is taken as having the same period length and when distributions are continued consecutively, wrapped Richard distribution will be obtained in a similar way to that defined in the positive half-axis. However, in this process, the coefficients must be changed in an appropriate manner.

Conclusion

In general, the first and second derivative properties of the probability distribution and growth curve are very similar. In the references [15-18], the first and second derivative properties of the growth curves were clearly examined. This similarity provides the use of a growth curve as a probability distribution. In the research, Richard curve is adapted to provide probability distribution characteristics. The obtained probability distribution is more flexible than others. This strengthens the compliance of the distribution the data. Also, the method of dividing the distribution into sub-intervals provides the convenience of use. The finally wrapped Richard distribution ensures that the distribution is consistent with the circular data.

References

[1] Gregorczyk, A., Richards Plant Growth Model, *Journal Agronomy and Crop Science*, *181* (1998), 4, pp. 243-247
[2] Bertalanffy, Von L., Untersuchungen über die Gesetzlichkeit des Wachstums. I. Allgemeine Grundlagen der Theorie mathematische und physiologische Gesetzlichkeiten des Wachstums bei Wassertieren (in German), *Arch. Entwicklungsmech*, *131* (1934), pp. 613-652
[3] Richard F. J., A Flexible Growth Function for Emprical Use, *Journal Exp. Bot*., *10* (1959), 29, pp. 290-300
[4] Gurcan M., Oner Y., On The Existence Problem in Logistic Regression Models By Alternative Form, *Adv. and Appl. in Stat*., *1* (2001), 2, pp. 165-174
[5] Gurcan, M., Colak, C., The Polynomial Approach in Growth Curves, *Turkiye Klinikleri Bioistatistik Dergisi*, *1-2* (2009), pp. 54-62
[6] Gurcan, M., *et al*., Bernstein Polynomial Approach Against to Some Frequently Used Growth Curve Models on Animal Data, *Pak. J. Statist*, *26* (2010), 3, pp. 509-516
[7] Gurcan, M., Colak, C., Generalization of Korovkin Type Approximation by Appropriate Random Variables & Moments and an Application in Medicine, *Pak. J. Statist*., *27* (2011), 3, pp. 283-297
[8] Tachev, G., Voronovskaja's Theorem Revisited, *Journal of Mathematical Analysis and Applications*, *343* (2008), 1, pp. 399-404
[9] Gonska, H., Tachev, G., A Quantitative Variant of Voronovskaja's Theorem, *Results in Mathematics*, *53* (2009), 3, pp. 287-294

[10] Albert, A., Anderson, J. A., On the Existence of Maximum Likelihood Estimates in Logistic Regression Models, *Biometrica*, *71* (1984), 1, pp. 1-10
[11] Delgado, J., Pena, J. M., A Linear Complexity Algorithm for the Bernstein Basis, *Proceedings*, IEEE Int. Conf. on Geometric Modelling and Graphics (GMAG), London, UK, 2003, Vol. 3, pp. 162-167
[12] Oner, Y., *et al*., On Continuous Deformation of Richards Family, *International Journal of Pure and Applied Mathematics*, *18* (2005), 3, pp. 375-377
[13] Tachev, G., Approximation of Bounded Continuous Functions by Linear Combinations of Phillips Operators, *Demonstratio Mathematica*, *47* (2014), 3, pp. 662-671
[14] Horowitz, J. L., Mammen, E., Non-Parametric Estimation of an Additive Model with a Link Function, *The Annals of Statistics*, *32* (2004), 6, pp. 2412-2443
[15] Tachev, G., The Distance between Bezier Curve and its Control Polygon, Jubilee Collection of Papers Dedicated to the 60th Anniversary of Prof. Mihail M. Konstantinov, 2013
[16] Tachev, G., On the Second Moment of Rational Bernstein Functions, *Journal of Computational Analysis and Applications*, *12* (2010), 2, pp. 471-479
[17] Tachev, G., Voronovskaja Theorem for Schoenberg Operator, *Mathematical Inequalities and Applications (MIA)*, *15* (2012), 1, pp. 49-59
[18] Tachev, G., New Estimates in Voronovskaja's Theorem, *Numerical Algorithms*, *59* (2012), 1, pp. 119-129

Paper Submitted: January 08, 2019
Paper revised: June 19, 2019
Paper accepted: July 25, 2019