Computational Methods for Differential Equations
http://cmde.tabrizu.ac.ir
Vol. 3, No. 1, 2015, pp. 45-50

A continuous approximation fitting to the discrete distributions using ODE

Hossein Bevrani
Department of Statistics,
Faculty of Mathematical Sciences,
University of Tabriz, Tabriz, 5166615648, Iran.
E-mail: bevrani@gmail.com, bevrani@tabrizu.ac.ir

Abstract The probability density functions fitting to the discrete probability functions has always been needed, and very important. This paper is fitting the continuous curves which are probability density functions to the binomial probability functions, negative binomial geometrics, poisson and hypergeometric. The main key in these fittings is the use of the derivative concept and common differential equations.

Keywords. Ordinary differential equations, Probability density functions, Pearson's family distribution.

2010 Mathematics Subject Classification. 65E17, 97K60.

1. Introduction

The discrete distributions approximation with continuous distributions have always been used and taken into consideration, for example, the normal distribution has been used for related approximation probabilities to poisson or binomial distribution [1].

Karl Pearson for the first time, investigated the continuous curves to the discrete probability functions. He suggested Pearsons's family that included twelve members of probability density functions. The details related to this family have been presented in Elderton (1968) and Pollard (1977) essays [2,3].

Bevrani and Sharififar (2013), using common differential equations for binomial distribution, produced a continuous approximation which was probability density function. This approach has been used for other discrete distributions such as geometric, negative binomial, poisson and hypergeometric, therefore, the aforementioned approach will be described briefly [1].

The binomial probability distribution is:

$$P(X=m)=C_n^m p^m (1-p)^{n-m}; \quad m=1,2,\cdots,n, \quad 0 \le p \le 1 \tag{1.1}$$

Received: 17 August 2015 ; Accepted: 1 September 2015.

The computation of variations ratio are as follows:

$$\frac{P(m+1)-P(m)}{P(m)}=\frac{np-(1-p)-m}{1-p+(1-p)m}, \tag{1.2}$$

and partition m up to $m+1$ to k compartment, and limited k to ∞, the following differential equation will be obtained.

$$\frac{1}{y}\frac{dy}{dx}=\frac{r+x}{b_0+b_1x}, \tag{1.3}$$

In which the right portion of equation (1.3) has obtained from the right portion of equation (1.2) which x, r, b_0 and b_1 in equation (1.3) are, $-m$, $np-(1-p)$, $1-p$, $p-1$.

Solving the differential equation (1.3) we will get:

$$y=C(b_0+b_1x)e^{(r-\frac{b_0}{b_1})/b_1}e^{x/b_1}, \tag{1.4}$$

which is the same desired continuous curve, in which C is normalizing function provided that $\int_{-\infty}^{\infty} ydx=1$ holds and b_0, b_1 are the obtained density function parameters. The follow-up, the proper density function, using the equation (1.2), to geometric distributions, negative binomial, poisson and hypergeometric and also equation (1.3) which is the same differential equation will be fitting.

2. The curve fitting to geometric distribution

The geometric probability function will be used in reaching the first success of probability computation. This function has a lot of applications in discussion of reliability and economy [4]. The relation of probability function is as follows:

$$P(X=m)=pq^{m-1}; \quad m=1,2,\cdots \tag{2.1}$$

The ratio of variations can be easily obtained as follows:

$$\begin{aligned}\frac{P(m+1)-P(m)}{P(m)}&=\frac{pq^m-pq^{m-1}}{pq^{m-1}}\\&=q-1,\end{aligned} \tag{2.2}$$

consequently, its related differential equation will be as follows:

$$\frac{1}{y}\frac{dy}{dn}=a_0. \tag{2.3}$$

Solving the differential equation (2.3) we will get:

$$y=Ce^{ax}, \quad x>0, \tag{2.4}$$

which C is normalizing parameter and a_0 is parameter function.

The equation (2.4) actually is exponential probability density function of a parameter which is the proof for the closeness of geometric and exponential distribution [5].

3. The curve fitting to negative binomial distribution

The negative binomial probability is used for computation of getting to r success and it has many application in probability, statistical inference, medical sciences, and pharmaceutical [4]. The equation as follows:

$$P(X=m)=\binom{m-1}{r-1}p^r q^{m-r+1}; \quad m=r,r+1,\cdots \tag{3.1}$$

The variation ration is in the form of:

$$\begin{aligned}\frac{P(m+1)-P(m)}{P(m)} &= \frac{\binom{m}{r-1}p^r q^{m-r+2}-\binom{m-1}{r-1}p^r q^{m-r+1}}{\binom{m-1}{r-1}p^r q^{m-r+1}}\\ &= \frac{q\binom{m}{r-1}-\binom{m-1}{r-1}}{\binom{m-1}{r-1}}\\ &= \frac{m(q-1)+(r-1)}{m+1-r},\end{aligned} \tag{3.2}$$

consequently, the differential equation

$$\frac{1}{y}\frac{dy}{dn}=\frac{a_0+b_0x}{x+a_1}, \tag{3.3}$$

which we will get

$$\begin{aligned}\ln y &= \int \frac{a_0+b_0x}{x+a_1}dx\\ &= \int \frac{b_0(x+a_1)+a_0-a_1b_0}{x+a_1}dx\\ &= \int (b_0+\frac{a_0-a_0b_0}{x+a_1})dx,\end{aligned} \tag{3.4}$$

consequently

$$\ln y = b_0x - a_0(1-b_0)\ln(x+a_1)+\ln c. \tag{3.5}$$

Therefore, the probability density function concerned will be obtained

$$y=Ce^{b_0x}(x+a_0)^{-a_0(1-b_0)}, \tag{3.6}$$

in which C is the value of normalizing and a_0 and b_0 are the probability density function parameters. The probability density function (3.6) is similar to gamma distribution which Bevrani et al. in 2015 have computed the convergence rate of this two distributions [6]. Meanwhile, we know that negative binomial distribution and gamma distribution are generalized state of geometric and exponential distributions, respectively.

4. The curve fitting to Poisson distribution

Poisson probability function is used in computation probability of a number of input, output or occurrences in time or location intervals. This function has a lot of applications in probability, queueing theory, reliability and simulation [4].

Its equation is as follows:

$$P(X=m)=\frac{e^{-\lambda}\lambda^m}{m!};\quad m=0,1,2,\cdots \tag{4.1}$$

variation ratio will be obtained as follows:

$$\begin{aligned}\frac{P(m+1)-P(m)}{P(m)}&=\frac{\dfrac{e^{-\lambda}\lambda^{m+1}}{(m+1)!}-\dfrac{e^{-\lambda}\lambda^m}{m!}}{\dfrac{e^{-\lambda}\lambda^m}{m!}}\\&=\frac{\lambda-m-1}{m+1}\\&=\frac{a_0-x}{a_1+x},\end{aligned} \tag{4.2}$$

consequently, the differential equation will be as follows:

$$\frac{1}{y}\frac{dy}{dx}=\frac{a_0-x}{a_1+x}, \tag{4.3}$$

In order to solve this, we have:

$$\begin{aligned}\ln y&=\int\frac{a_0-x_0}{a_1+x}dx\\&=\int\frac{a_0-x-a_1+a_1}{a_1+x}dx\\&=\int(\frac{a_0+a_1}{a_1+x}-1)dx\\&=(a_0+a_1)\ln(a_1+x)-x+\ln c\\&=\ln\left[(a_1+x)^{a_0+a_1}c/e^{-x}\right],\end{aligned} \tag{4.4}$$

consequently

$$y=C(a_1+x)^{a_0+a_1}e^{-x}. \tag{4.5}$$

5. The curve fitting to Hypergeometric distribution

This function is used in probability computation related to the sampling without replacement and it has applications in sampling diagnostics, probability, and game theory [4]. Its equation is as follows:

$$P(X=m)=\frac{\binom{k}{m}\binom{N-k}{n-m}}{\binom{N}{n}};\quad m=0,1,\cdots,\min\{k,n\}, \tag{5.1}$$

variation ratio wil be obtained as follows: Due to voluminous equation results, the final result has been presented.

$$\frac{P(X=m+1)-P(X=m)}{P(X=m)}=\frac{-m(N+2)-(N-k-n-kn+1)}{(m+1)(m+N-k-n+1)}. \tag{5.2}$$

And, the differential equation which will be obtained with the help of equation (5.2) consist of

$$\frac{1}{y}\frac{dy}{dn}=\frac{a_0+a_1x}{b_0+b_1x+b_2x^2}, \tag{5.3}$$

which is the same differential equation presented by Karl Pearson. Which through solving it, twelve Pearson's family members will be obtained. This family has a lot of applications which some of them have been mentioned in the references [7], [8], [9], [10].

6. Conclusion

The approximation for discrete functions such as binomial, geometric, negative binomial, poisson and hypergeometric, using differential equations, proper probability density functions were obtained. Through, the estimate of the obtained function parameters, the probabilities related to discrete probability functions can be approximated. This paper will properly show the application of differential equations is solving the statistic and probability problems.

References

[1] Bevrani, H., Sharififar, S. An approximation to binomial distribution, Journal of Statistical theory and Practic, 14(2), (2014), 1–8.

[2] Elderton, V., Johnson, N. System of frequency curves, Cambridge University Press, 1962.

[3] Pollard, J. H. A handbook of numerical and statistical thechniques, Cambridge University Press, 1977.

[4] Johnson, N. L., Kamp, A. W., Kotz, S. Univariate Discrete Distributions, 3rd Edtion, Wiley, 2005.

[5] Selivanova, D. O. Estimates of the rate of convergence of random sums, ph.D. Thesis, Mascow State University, 1955.

[6] Bevrani, H., Bening, V. E., and Korolev, V. YU. On the occuracy of the negative binomial distribution by the gamma distribution and convergency ratio, Journal of Mathematical Sciences, 205(1), (2015).

[7] Andreev, A., Kanto, A., Malo, P. Computational examples of a new method for distribution selection in the Pearson system. Journal of applied Statistics, 34(4), (2007), 487–506.

[8] Derrode, S., Mercier, G., Lecaillec, J., Garello, R. Estimation of Sec-Ice SAR clutter Statistics from Pearsons system of distributions, JGARSS, IEEE, (2001), 9–13.

[9] Griffs, V. W., Stedinger, J. R. Log-Pearson type III distribution and its application in flood frequency analysis, I: Distribution characteristics. Journal of Hydrologic Engineering, 12(5), (2007), 482–491.

[10] Karvanen, J., Koivunen, V. Blind Separation methods based on Pearson system and its extensions, Signal Processing, 82, (2002), 663–673.