Journal of Statistics Applications & Probability
An International Journal

http://dx.doi.org/10.18576/jsap/060218

Maximum Entropy Formalism for Zero Truncated Poission and Binomial Distribution

Zahoor Ahmad and Adil Rashid*

P.G Department of Statistics, University of Kashmir, Srinagar, India

Received: 10 Sep. 2016, Revised: 9 Feb. 2017, Accepted: 12 Feb. 2017.
Published online: 1 Jul. 2017.

Abstract: The constraints, under which some Lagrangian probability distributions can be obtained as maximum-entropy distributions, have been found. The maximum entropy formalism is then used to study maximum entropy characterization of Zero Truncated Poisson and Binomial distribution.

Keywords: Maximum entropy formalism, Shanon entropy, Bayesian entropy, Zero Truncated Poisson Distribution, Zero Truncated Binomial Distribution.

1 Introduction

Most of the well known probability distributions for both discrete and continuous variates and for both univariate and multivariate cases can be obtained by maximizing the entropy of a probability distribution subject to certain constraints. The method of derivation also characteristics each probability distribution as the most likely or the most unprejudiced or the least biased distribution among a given class of probability distributions [Jaynes (1957, 1978), Kapur (1981 b,c,d,e,f,g), Tribus (1969)].

Consul and Shenton (1972) have obtained a class of Lagrangian distributions which include generalization of some discrete distributions. Recently Zahoor, Adil and Jan (2017) obtained a new discrete compound distribution which proved to be more flexible than classical distributions. The Maximum entropy formalism of some univariate and multivariate Lagrangian probability distributions has been studied by Kapur (1982). The object of the present paper is to study maximum entropy characterization of Zero Truncated Poisson and Binomial distribution.

2 Shanon and Bayesian Entropies

Shannon's [Shannon and Weaver (1949)] measure of entropy for a discrete probability distribution $(p_1, p_2, \dots, p_n)$ is given by

$$S = -\sum_{i=1}^{n} p_i \ln p_i \tag{2.1}$$

This is maximum when all the n outcomes are equally likely i.e. when

$$p_1 = p_2 = \cdots = p_n = \frac{1}{n} \tag{2.2}$$

This is consistent with Laplace's principal of insufficient reason that in the absence of any information to contrary, the uncertainty is maximum in the case of equally likely outcomes. However if, on the basis of experience or intuition, there is reason to believe that the most likely prior distribution is

$$p_1 = \alpha_1,\ p_2 = \alpha_2, \dots, p_n = \alpha_n\ ;\ \sum_{i=1}^{n} \alpha_i = 1 \tag{2.3}$$

Then we the Bayesian entropy [Kapur (1981 a)]

*Corresponding author e-mail: zahoor151@gmail.com

$$B = -\sum_{i=1}^{n} p_i \ln\left(\frac{p_i}{\alpha_i/[\alpha_i]_{min}}\right)$$

$$= -\sum_{i=1}^{n} p_i \ln\left(\frac{p_i}{\alpha_i}\right) - ln[\alpha_i]_{min}\sum_{i=1}^{n} p_i$$

$$= -\sum_{i=1}^{n} p_i \ln\left(\frac{p_i}{\alpha_i}\right) - ln[\alpha_i]_{min} \tag{2.4}$$

Where $[\alpha_i]_{min} = min[\alpha_1, \alpha_2, \ldots, \alpha_n]$ (2.5)

Thus maximizing B is equivalent to maximizing

$$-\sum_{i=1}^{n} p_i \ln\left\{\frac{p_i}{k\alpha_i}\right\} \tag{2.6}$$

where k is a constant.

When the number of outcomes is finite, Shannon's entropy is defined as

$$S = \sum_{i=1}^{\infty} p_i \ln p_i \tag{2.7}$$

In this case (2.2) does not make sense, yet in applications of maximum entropy principle, (2.7) is always used. This is equivalent to using (2.6) with α's equal. This implies the use of an 'improper' prior distribution, but this is taken as justified so long as the probability distribution obtained by maximizing S subject to $\sum_{i=1}^{\infty} p_i = 1$ and the given constraints is a 'proper' probability distribution. However, when possible, it is desirable to use (2.6) with $\sum_{i=1}^{\infty} \alpha_i$ as a convergent series.

3 Lagrangian Distributions

Lagrangian distributions are based on the use of Lagrange's theorem

$$f(t) = f(0) + \sum_{s=1}^{\infty} \frac{u^s}{s!}\frac{d^{s-1}}{dt^{s-1}}[\{g(t)\}^s f'(t)]_{t=0}\ ; u = \frac{t}{g(t)} \tag{3.1}$$

where f(t) and g(t) are probability generating functions. This gives {see Consul and Shenton (1972)}

$$P[X = x] = \frac{1}{x!}\frac{d^{x-1}}{dt^{x-1}}[\{g(t)\}^x f'(t)]_{t=0}$$

$$P[X = x] = 0 \tag{3.2}$$

By using different f(t) and g(t), we get different Lagrangian distributions.

4 Maximum-Entropy Formalism:

Maximizing

$$S = -\sum_{i=1}^{n} P(x) \ln\frac{P(x)}{\alpha(x)} \tag{4.1}$$

Subject to

$$\sum_{x=1}^{n} P(x) = 1\ , \quad \sum_{x=1}^{n} P(x) g_r(x) = \bar{g}_r \ ; \ r = 1, 2, \ldots, m \tag{4.2}$$

We get, by using Lagrange's method

$$\left[-1 - \ln\frac{P(x)}{\alpha(x)}\right] - \lambda_0 - \sum_{r=1}^{m} \lambda_r g_r(x) = 0 \tag{4.3}$$

so that

$$P(x) = A\,\alpha(x)\exp[-\lambda_1 g_1(x) - \lambda_2 g_2(x)\ \ldots\ \lambda_m g_m(x)]\ , \tag{4.4}$$

Where $A,\ \lambda_1, \lambda_2, \ldots, \lambda_m$ are obtained by using the constraints (4.2)

5 Maximum Entropy Characterization of Zero Truncated Poisson and Binomial distribution

a) Zero Truncated Poisson Distribution

$$Let\ \alpha(x) \propto \left(\frac{1}{x!}\right) \quad ; x = 1, 2, \ldots \tag{5.1}$$

This is an 'improper' prior. Let the mean be prescribed as $(\lambda/1-e^{-\lambda})$, then we have

$$P(x) = A\,\frac{1}{x!}b^x \qquad ;x = 1,2,\ldots \tag{5.2}$$

where

$$\sum_{x=1}^{n} P(x) = 1 \qquad and \qquad \sum_{x=1}^{n} xP(x) = \frac{\lambda}{1-e^{-\lambda}} \tag{5.3}$$

From (5.3), we have

$$\sum_{x=1}^{\infty} A\,\frac{1}{x!}b^x = 1$$

$$\Rightarrow\ A\left(\frac{b}{1!}+\frac{b^2}{2!}+\frac{b^3}{3!}\ldots\right) = 1$$

$$\Rightarrow\ A = \frac{1}{(e^b-1)} \tag{5.4}$$

Again from (5.3), we have

$$\sum_{x=1}^{\infty} x\,\frac{A\,b^x}{x!} = \frac{\lambda}{1-e^{-\lambda}}$$

$$\Rightarrow\ Ab\left(1+\frac{b}{1!}+\frac{b^2}{2!}+\ \ldots\right) = \frac{\lambda}{1-e^{-\lambda}}$$

$$\Rightarrow\ \frac{be^b}{e^b-1} = \frac{\lambda}{1-e^{-\lambda}} \qquad \{\text{using (5.4)}\}$$

$$\Rightarrow \frac{b}{1-e^{-b}} = \frac{\lambda}{1-e^{-\lambda}}$$

$$\Rightarrow b = \lambda \tag{5.5}$$

Using (5.5) in (5.4), we have

$$A = \frac{1}{e^{\lambda}-1} = \frac{e^{-\lambda}}{1-e^{-\lambda}} \tag{5.6}$$

Giving $$P(x) = \frac{Ab^x}{x!} = \frac{e^{-\lambda}\lambda^x}{x!(1-e^{-\lambda})} \qquad ;x = 1,2,3,\ \ldots \tag{5.7}$$

Hence, the Zero Truncated Poisson Distribution is Maximum Bayesian Entropy Distribution (MBED) when the prior probability distribution is giving by (5.1) and the mean of the distribution is prescribed.

b) Zero Truncated Binomial Distribution:

$$Let\ \alpha(x) \propto \binom{n}{x} \qquad ;x = 1,2,\ldots n \tag{5.8}$$

and let the mean be prescribed as $(np/1-q^n)$, then we get

$$P(x) = A\binom{n}{x}b^x \qquad ;x = 1,2,\ldots n \tag{5.9}$$

Where

$$\sum_{x=1}^{n} P(x) = 1 \qquad and \qquad \sum_{x=1}^{n} xP(x) = \frac{np}{1-q^n} \tag{5.10}$$

From (5.10), we have

$$\sum_{x=1}^{\infty} A\binom{n}{x}b^x = 1$$

$$\Rightarrow\ A\{(1+b)^n-1\} = 1$$

$$\Rightarrow\ A = \frac{1}{(1+b)^n-1} \tag{5.11}$$

Again, from (5.10), we have

$$\sum_{x=1}^{\infty} x\, A \binom{n}{x} b^{x} = \frac{np}{1-q^{n}}$$

$$\Rightarrow\ Ab(1+b)^{n-1} = \frac{p}{1-q^{n}}$$

$$\Rightarrow\ \frac{b(1+b)^{n-1}}{(1+b)^{n}-1} = \frac{p}{1-q^{n}}\ \{\text{using } (5.11)\}$$

$$\Rightarrow b = \frac{p}{q} \tag{5.12}$$

Using (5.12) in (5.11), we have

$$A = \frac{1}{\left(1+\frac{p}{q}\right)^{n}-1} = \frac{q^{n}}{1-q^{n}} \tag{5.13}$$

Giving $$P(x) = A\binom{n}{x} b^{x} = \frac{\binom{n}{x} p^{x} q^{n-x}}{1-q^{n}} \quad ; x = 1, 2, 3,\ \dots, n \tag{5.14}$$

Thus, the Zero Truncated Binomial Distribution can be characterized as the Maximum Entropy Bayesian Distribution (MBED) for which the prior probability distribution is giving by (5.8) and for which the mean is prescribed.

References

[1] P.C. Consul and L. R. Shenton (1972) Use of Lagrange's expansion for generating discrete generalized probability distributions, SIAM J. Appl. Math. 23.239-248.

[2] E. T. Jaynes (1957) Information theory and statistical mechanics, Phys. Rev. 106, 620-630.

[3] C. E. Shannon and W. Weaver (1949) The mathematical theory of communication, University of lllinois Press, Urbana.

[4] M. Tribus (1969) Rational descriptions, decisions and designs, Pergamon, Oxford.

[5] E. T. Jaynes (1978) Where do we stand on maximum entropy? In the Maximum Entropy Formalism edited by R. D. Levine and M. Tribus, MIT Press, pp. 15-118.

[6] J. N. Kapur (1981a) Bayesian entropy and its application to entropy-maximization models, Univ. Manitoba Res. Report No. 15.

[7] J. N. Kapur (1981b) Characterization of probability distributions through entropy-maximization principle, Univ. Mnitoba Res. Report No. 16.

[8] Zahoor, A., Adil, R and Jan, T.R (2017) A new discrete compound distribution with application, J. Stat. Appl. Pro. 6, No. 1, 185-201(2017)

Zahoor Ahmad is Lecturer of Statistics in Department of School Education, Jammu and Kashmir. He holds Master degree in Statistics from the University of Kashmir and has been a gold medalist. Presently he is doing P.hd in the Department of Statistics, University of Kashmir. He has published work in international and national journals of repute. His research interests include Distribution theory, Estimation theory and Bio-Statistics.

Adil Rashid was born in 1984. He holds M.Sc. and M. Phil. degrees in Statistics from University of Kashmir, Srinagar, India in 2009 and 2013 respectively. At present, he is pursuing his Ph. D degree in Statistics in the Post Graduate Department of Statistics, University of Kashmir, Srinagar, India. He has to his credit several research publications both in national and international journals of repute. His current research interests are in Compound Probability Distributions, Bio-Statistics and Models for Complex Systems.