Statistical properties of negative binomial distribution under imprecise observation

Ajibola Akeem Adepoju*[1], Usman Mohammed[2], Safiya Sada Sani[3], Kabiru Adamu[4], Kabiru Tukur [1], Aliyu Ismail Ishaq[2]

[1]*Department of Statistics, Kano University of Technology, Wudil, Kano State, Nigeria*
[2]*Department of Statistics, Ahmadu Bello University, Zaria, Nigeria*
[3]*Department of Agronomy, Ahmadu Bello University, Zaria, Kaduna State, Nigeria*
[4]*Department of Statistics, Nuhu Bamali Polytechnic, Zaria, Kaduna State, Nigeria*

It is observed that some random variables are measured with uncertainty. In this paper, we intend to generate some properties of negative binomial distribution under imprecise measurement. These properties include fuzzy mean, fuzzy variance, fuzzy moment and fuzzy moment generating function. The uncertainty in the observations may not be addressed with the classical approach to probability distribution; therefore, fuzzy set theory helps to modify the classical approach.

Keyword: fuzzy probability density function; fuzzy mean; fuzzy variance; fuzzy moment; fuzzy moment generating function

1 Introduction

Negative binomial distribution is a discrete probability distribution which is generalization of geometric distribution. In an experiment of independent Bernoulli trials in which each trial has constant probability of success$p, 0 < p < 1$, are performed until a total of r success is obtained. This kind of experiment does occur in real-life on daily basis. Such occurrence could be found in clinical settings, games, engineering, etc. In the application of this technique, probability of success and failure is certain. However, there are situations whereby the probability may be fuzzy or uncertain. For instance, the probability of success of a particular student examination is about 0.3; meaning that such probability is uncertain but centered around 0.3.

Definition 1. Fuzzy sets: If $\tilde{A}$ denotes a fuzzy subset of universal Ω, $\tilde{A}$ is defined with its membership function $\tilde{A}_{(x)}$, which produces values within 0 and 1 for all x in Ω. $\tilde{A}_{(x)}$ is a function mapping Ω into [0,1]; x_0 belongs to $\tilde{A}$ if $\tilde{A}_{(x0)}$=1; meaning that the membership value of x_0 in $\tilde{A}$ is 1, and x_1 does not belong to $\tilde{A}$ if $\tilde{A}_{(x1)}$=0; meaning that the membership value of x_1 in $\tilde{A}$ is 0. If $\tilde{A}_{x2)}$=(0,1), we say the membership value of x_2 is in $\tilde{A}$ and thus defines fuzzy function. Also, whenever $\tilde{A}_{(x)}$ is equal to 1 or 0 the result is a crisp set (non-fuzzy) subset of Ω. Crisp is defined as not fuzzy or a regular set while a crisp number is a real number (Buckley, 2006).

Definition 2. Fuzzy numbers: A triangular fuzzy number $\widetilde{M}$ is defined by three numbers $a < b < c$ where the [a,c] are the base and its vertex is at $x = b$. For example, a triangular fuzzy number could be described as $\tilde{A} = (4.7,5,5.4)$. Similarly, a trapezoid fuzzy number $\widetilde{M}$ is

*Corresponding Author; E-mail: akeebola@gmail.com

described by four numbers $a < b < c < d$ such that [a,d] are the base and the top at [b,c] (Buckley, 2006).

Definition 3. Given a fuzzy set A on a universe X, then α-cuts are:

$$\tilde{A}_{\alpha} = \{x \in X \mid \tilde{A}(x) \geq \alpha\}, \alpha \in [0,1] \text{ (Zadeh, 1971} \quad (1)$$

The α-cuts are threshold through a fuzzy set producing regular crisp (non-fuzzy) sets. That is, by taking α as the degrees of the elements, we have different crisp subsets of X. Precisely α-cuts of a fuzzy set produces a fuzzy subset of any given set and the results are crisp. However, crispness simply means the exact value of a variable, while fuzzy simply means impreciseness, uncertainty or vagueness about the value of a variable. Role of α –cuts is to decompose a fuzzy set into a weighted combination of classical sets using the resolution identity. This principle is important in fuzzy set theory because it establishes a bridge between fuzzy sets and crisp sets. Suppose $\tilde{A}$ is a fuzzy number, then the $\tilde{A}_0$= all the real number in that set and it is called the base or support.

In a real life situation, vagueness and uncertainty do occur as a result of human error usually due to judgmental decision based on qualitative measures, such as 'the distance is far', but if one would have asked 'how far is the distance?' different people will have different responses as in: 'the distance is very far', 'not too far', 'too far', 'not too very far'. All these responses are based on qualitative measurement which might not be presented with exact value. Zadeh (1965) and (1975) proposed the idea of the fuzzy sets theory and the type-1 fuzzy sets, with a degree of membership called crisp membership value, whose values are over the range, 0 to 1. He also proposed the Type-2 fuzzy sets, which is an extension of the type-1 fuzzy sets, the type-2 fuzzy set is three dimensional, that is, it comprises of two membership functions; the upper membership function and lower membership function and the representative value (Zadeh, 1975). On fuzzy binomial distribution, Buckley (2006) intensively studied the fuzzy binomial distribution and its applications. Likewise, Buckley (2006) exclusively explored Fuzzy Poisson Distribution, its properties and application. Kareema and Abdul (2012) studied the properties of fuzzy geometric distribution and the results reveal some properties of the fuzzy geometric distribution.

Considerate number of authors which include but not limited to Wang, *et al.* (2019), Adepoju (2018), Kahraman and Kabak (2016), Poongodi and Muthulakshmi (2015), Sheldon (2010) and Mendel, *et al.* (2006) have made various contributions to the extensional development in fuzzy studies. However, this paper aimed at addressing the challenges of fuzziness in the estimate of some of the parameters of negative binomial distribution. Such parameters include fuzzy mean, fuzzy variance, fuzzy moment and the fuzzy moment generating function.

2 Methodology

2.1 Fuzzy negative binomial distribution

In a classical experiment, the probability p of occurrence of an event if known from previous experiment is expressed with certainty about the parameter of the random experiment. Assuming, the probability of an event is expressed with some degree of uncertainty, and the estimate is needed or collected from opinion of professionals. Therefore, the probability values tend to be fuzzy or imprecise. Let $\tilde{p}$ and $\tilde{q}$ be the fuzzy probability of success and failure, respectively. Then;

$$p \in \tilde{p}$$

$$q \in \tilde{q}$$

$$p + q = 1$$

$\tilde{p} + \tilde{q}$ may not necessarily be equal to one,

where p and q are the classical probability of success and failure, while $\tilde{p}$ and $\tilde{q}$ are the fuzzy probability of success and failure. For instance, if the probability of success in a trial is 0.3, then $p = 0.3$, that probability is certain and without ambiguity, but if the probability is described with uncertainty or fuzziness, say, the probability of a success is about 0.3. This implies that the probability is not far from 0.3. It denotes that it is a bit less or a bit greater than 0.3. Therefore, p is a subset of $\tilde{p}$.

In a given experiment that follows a negative binomial distribution, let $\tilde{p}(x)$ be the fuzzy probability of success from an independent trials. Then the α-cuts are slices through a fuzzy set resulting in crisp sets (non-fuzzy). Therefore, the α-cuts fuzzy negative binomial probability will be:

$$\tilde{p}(x)[\alpha] = \left[\binom{x-1}{r-1} p^r q^{x-r} / S\right] x = r, r+1 \qquad (2.1)$$

$$\text{for} 0 \leq \binom{x-1}{r-1} p^r q^{x-r} \leq 1$$

where $S = p \in \tilde{p}[\alpha], q \in \tilde{q}[\alpha], p + q = 1$

$$\text{but} \tilde{p}(x) \neq \left[\binom{x-1}{r-1} \tilde{p}^r \tilde{q}^{x-r} / S\right] \qquad (2.2)$$

If $\tilde{p}(x)[\alpha] = p_{n1}(\alpha), p_{n2}(\alpha)$

then;

$$p_{n1}(\alpha) = min\left[\binom{x-1}{r-1} p^r q^{x-r} / S\right], \; p_{n2}(\alpha) = max\left[\binom{x-1}{r-1} p^r q^{x-r} / S\right], \qquad (2.3)$$

where p_{n1} and p_{n2} are the probabilities of occurrence of event at the lower value and upper value, respectively. While $p_{n1}(\alpha)$ and $p_{n2}(\alpha)$ are the decompositions or slices of the probabilities of occurrence of the event both at the lower (minimum) value and upper (maximum) value.

2.2 Fuzzy mean and variance of negative binomial distribution

From the classical negative binomial distribution, the mean and variance are expressed as:

$\mu = \frac{r}{p}$ and $\sigma^2 = \frac{qr}{p^2}$.

We can now obtain the fuzzy mean and fuzzy variance from the classical property thus:

$$\mu[\alpha] = \left\{\sum_{x=1}^{n} x \binom{x-1}{r-1} p^r q^{x-r} / S\right\} = \left\{\frac{r}{p} / S\right\}. \quad (2.4)$$

$$\tilde{\mu} = \frac{r}{\tilde{p}} \quad (2.5)$$

The fuzzy variance

$$\tilde{\sigma}^2[\alpha] = \left\{E(x^2) - \left(E(x)\right)^2 / S\right\} = \left\{\sum_{x=1}^{n} x^2 \binom{x-1}{r-1} p^r q^{x-r} - \left(\sum_{x=1}^{n} x \binom{x-1}{r-1} p^r q^{x-r}\right)^2 / S\right\} = \left\{\frac{qr}{p^2} / S\right\} \quad (2.6)$$

$$\tilde{\sigma}^2 = \frac{\tilde{q}r}{\tilde{p}^2} \quad (2.7)$$

Moments

Let X be a random variable. Let k be a positive integer and c be a constant. Then the moment of order k or k^{th} moment of X about the origin is denoted by:

$$\mu_k^1 = E(x^k) = \sum_i x_i^k f(x_i). \quad (2.8)$$

The fuzzy moment of negative binomial is defined as:

$$E(x^k)[\alpha] = \sum_{x=1}^{n} x^k \binom{x-1}{r-1} p^r q^{x-r} / S. \quad (2.9)$$

Moment generating function

If $M_x(t, \tilde{p})$ is the fuzzy moment generating function of a negative binomial random variable, then its α-cuts are determined as:

$$M_x(t, \tilde{p})[\alpha] = E(e^{tx})[\alpha] = \sum_{x=r}^{\infty} e^{tx} \binom{x-1}{r-1} p^r q^{x-r} / S \quad (2.10)$$

$$M_x(t, \tilde{p})[\alpha] = E(e^{tx})[\alpha] = p^r \sum_{x=r}^{\infty} (e^t)^{x-r} (e^t)^r \binom{x-1}{r-1} q^{x-r} / S$$

$$= (pe^t)^r \sum_{x=r}^{\infty} \binom{x-1}{r-1} qe^{t(x-r)} / S$$

$$= (pe^t)^r (1 - qe^t)^{-r} / S. \quad (2.11)$$

$$M_x(t,\tilde{p})[\alpha] = \left[\frac{pe^t}{1-qe^t}\right]^r /S. \quad (2.12)$$

$$M_x(t,\tilde{p}) = \left[\frac{\tilde{p}e^t}{1-\tilde{q}e^t}\right]^r . \quad (2.13)$$

As usual, the fuzzy moment generating function would be used to determine the first moment and second moment so as to obtain the mean and variance of x. Hence, we will calculate the first two derivatives of $M_x(t,\tilde{p})[\alpha]$ in (2.12)and set $t = 0$, herein the mean and variance of the fuzzy negative binomial distribution can be obtained thus:

$$M_x(t,\tilde{p})[\alpha] = \left[\frac{pe^t}{1-qe^t}\right]^r /S$$

$$\frac{dM_x(t,\tilde{p})}{dt}[\alpha] = r\left(\frac{pe^t}{1-qe^t}\right)^{r-1} \frac{(1-qe^t)pe^t+pe^tqe^t}{(1-qe^t)^2} /S$$

$$= r\frac{(pe^t)^r}{(1-qe^t)^{r+1}} /S \quad (2.14)$$

$$\frac{dM_x(0,\tilde{p})}{dt}[\alpha] = \frac{r}{p}/S$$

$$\tilde{\mu} = \frac{r}{\tilde{p}}$$

Also, the second derivative:

$$\frac{d^2M_x(t,\tilde{p})}{dt^2}[\alpha] = \left\{r\left[\frac{(1-qe^t)r(pe^t)^{r-1}pe^t-(pe^t)^r(r+1)(1-qe^t)^r(-qe^t)}{(-qe^t)^{2r+2}}\right]/S\right\} = \left\{r\frac{(pe^t)^r(r+qe^t)}{(1-qe^t)^{r+2}}S\right\} \quad (2.15)$$

$$\frac{d^2M_x(0,\tilde{p})}{dt^2}[\alpha] = \left\{\frac{r^2+rq}{p^2}/S\right\} \quad (2.16)$$

$$\tilde{\sigma}^2 = \frac{d^2M_x(0,\tilde{p})}{dt^2}[\alpha] - \left[\frac{dM_x(0,\tilde{p})}{dt}\right]^2[\alpha] = \left[\frac{r^2+rq}{p^2}/S\right] - \left[\frac{r}{p}/S\right]^2. \quad (2.17)$$

Therefore, the variance is:

$$\tilde{\sigma}^2 = \frac{\tilde{q}r}{\tilde{p}^2} \quad (2.13)$$

Parameters with such as $\tilde{\mu}$, $\tilde{\sigma}^2$, $\tilde{p}$, and $\tilde{q}$ are the fuzzy mean, fuzzy variance, fuzzy probability of success and fuzzy probability of failure respectively. The relationship between them is that p is a subset of $\tilde{p}$. Mapple 18 software was used to estimates the values of the probability problem illustrated in this work.

3. Illustration

If the probability is about 0.4 that a child exposed to a certain contagious disease will catch it, what is the probability that tenth child exposed to the disease will be third to catch it? This is a typical fuzzy negative binomial problem. Here, the probability of success is not certain. The only guidance about the probability of success is that it centers on 0.4, therefore, it could be 0.3, 0.4 and/or 0.5. Thus:

$\tilde{p} = 0.3, 0.4, 0.5$ and $\tilde{q} = 0.5, 0.6, 0.7$ and r = 3.

If $p \in \tilde{p}[\alpha]$

then, $q = 1 - \tilde{p} \in \tilde{q}[\alpha]$

$p_{n1}[\alpha] = \min(36p^3q^7/S)$

$p_{n2}[\alpha] = \max(36p^3q^7/S)$

$\tilde{p}(3)[\alpha] = [36p_1^3(\alpha)q_1^7, 36p_2^3(\alpha)q_2^7]$,

where $\tilde{p}[\alpha] = [p_1(\alpha), p_2(\alpha)] = [0.3 + 0.1\alpha, 0.5 - 0.1\alpha]$.

Table 1: Alpha-cuts of the fuzzy probability

α	$p_1(\alpha)$	$p_2(\alpha)$
0	0.080048380	0.035156250
0.1	0.079860352	0.038008609
0.2	0.079307603	0.040930742
0.3	0.078409365	0.043906332
0.4	0.077187469	0.046917225
0.5	0.075665887	0.049943488
0.6	0.073870293	0.052963505
0.7	0.071827627	0.055954093
0.8	0.069565693	0.058890647
0.9	0.067112770	0.061747326
1	0.064497254	0.064497254

Table 1 reveals the probabilities for different values of α-cut ranging from 0 to 1. It contains all the possible probabilities in this experiment. The $\boldsymbol{p_1(\alpha)}$ and $\boldsymbol{p_2(\alpha)}$ are `probabilities values of the assumed values of 0.3 and 0.5 that centered around the given probability of success 0.4 at different α levels. At $\alpha=1$, $\boldsymbol{p_1(\alpha)}$ and $\boldsymbol{p_2(\alpha)}$ concise at 0.064497254 and this is the exact probability that tenth child exposed to the disease will be third to catch it, knowing that probability that a child exposes to such contagious diseases will catch it is 0.4. This can be confirmed by the classical method too. However, at other α levels $\boldsymbol{p_1(\alpha)}$ and $\boldsymbol{p_2(\alpha)}$ give different probability.

Figure 1 is the graphical presentation of the result obtained in the table. This indicates a typical triangular fuzzy membership function.

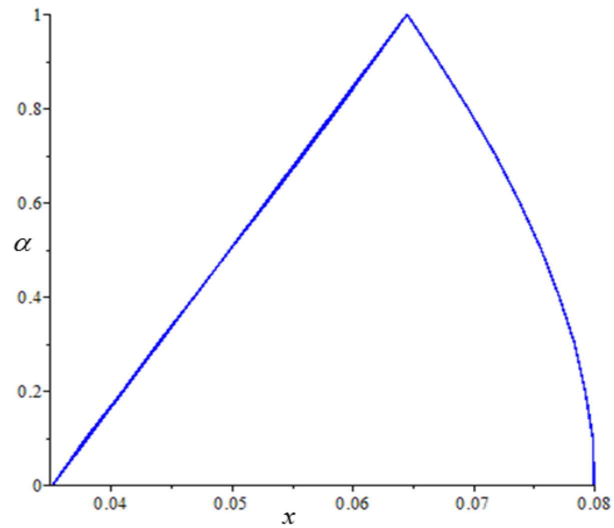

Figure 1. Triangular fuzzy membership function

4 Conclusion

The research shows that while dealing with uncertainty about the parameter of negative binomial distribution, fuzzy logic could be employed to obtain the fuzzy number over the parameter domain due to its flexibility in dealing with uncertainty. With this approach, other estimates, such as mean, variance, moment and moment generating function were obtained.

References

Adepoju, A.A. (2018). Performance of Fuzzy Control Chart Over the Traditional Control Chart, *Benin Journal of Statistics,* 1(1), 101-112.

Buckley, J.J. (2006). *Fuzzy Probability and Statistics*, Springer, New York.

Kahraman, C. and Kabak, O. (2016). *Fuzzy Statistical Decision-Making: Theory and Application,* Springer, Switzerland.

Kareema, A.A. and Abdul, H.A. (2012). Fuzzy Geometric Distribution with Properties, *Achieves Des Sciences*, 65(2), 13-19.

Mendel, J.M., John, R.I. and Liu, F. (2006). Interval Type-2 Fuzzy Logic Systems made Simple, *IEEE Transactions on Fuzzy Systems*, 14(6), 808–821.

Poongodi, T. and Muthulakshmi, S. (2015). Fuzzy Control Chart for Number of Customers of E/M/1 Queuing Model, *International Journal of Advanced Scientific and Technical Research*, 3(5), 9–22.

Sheldon, R. (2010). *A First Course in Probability* (8^{th} Ed.), Pearson Education Inc., Waterloo, Canada.

Wang, G., Xu, Y. and Qin, C. (2019). Basic Fuzzy Event Space and Probability Distribution of Probability Fuzzy Space, *Mathematics*, 7(542), 1-15.

Zadeh, L.A. (1971). Similarity Relations and Fuzzy Orderings, *Information Sciences*, 3(2), 177–200.

Zadeh, L.A. (1975). The concept of a Linguistic Variable and Its Application to Approximate Reasoning, *Information Sciences*, 8, 199–249.

Zadeh, L.A. (1965). Fuzzy Sets, *Information and control*, 8(3), 338–353