

An Alternative Approach of Binomial and Multinomial Distributions

Mian Arif Shams Adnan

Department of Mathematical Science, Ball State University, Muncie, IN 47304, USA
maadnan@bsu.edu

Md. Tareq Ferdous Khan

Department of Statistics, Jahangirnagar University, Dhaka-1342, Bangladesh

Md. Forhad Hossain

Department of Statistics, Jahangirnagar University, Dhaka-1342, Bangladesh

Abdullah Albalawi

Department of Mathematical Science, Ball State University, Muncie, IN 47304, USA

Received 4 April 2015

Accepted 18 September 2016

Abstract

In this paper we have tried to present an alternative approach for two discrete distributions such as Binomial and Multinomial with a new concept of sampling having a more general form apart from the traditional methods of sampling. The existing distributions may be obtained as a special case of our newly suggested technique. The basic distributional properties of the proposed distributions are also been examined including the limiting form of the proposed Binomial distribution to normal distribution.

Keywords: Generalized Binomial Distribution, Generalized Multinomial Distribution, Sampling Methods, Distributional Properties, Arithmetic Progression, Limiting Form.

Mathematics Subject Classification: 60E05, 62E10, 62E15, 62E20.

1. Introduction

The discrete distributions are widely used in the diversified field and among them the distribution like Binomial, Multinomial, Poisson and Geometric are the most commonly used discrete distributions. The other discrete distributions include uniform or rectangular, hypergeometric, negative binomial, power series etc. The truncated and censored forms of the different discrete distributions are used as probability distribution in statistics literature and in real life problems. The binomial distribution was first studied in connection with the games of pure chance but it is not limited within that area only, where the Multinomial distribution is considered as the generalization of the binomial distribution. The number of mutually exclusive outcomes from a single trial are k in multinomial distribution compared to two outcomes namely success or failure of Binomial distribution.

The usual binomial distribution is the discrete probability distribution of the number of successes 0 to n resulted from n independent Bernoulli trials each of which yields success with probability p and failure with

probability q . One of the important assumptions regarding binomial variate representing the number of success is that it can take only values in the sequence of $0, 1, 2, \dots, n$. But in real world, the successes of binomial may not occur in the usual way rather it may occur in a different sequence such as (i) $0, 2, 4, \dots, 2n$, (ii) $2, 4, \dots, 2n$, (iii) $0, 3, 6, \dots, 3n$, (iv) $3, 6, 9, \dots, 3n$ and so on.

In cases (ii) and (iv), truncated distribution is the better option to find the probability of number of success. In truncated distribution, it is assumed that the truncated values of the random variable have certain probabilities. If it is considered that there is no existence of the truncated values, the truncated distribution cannot provide the probability due to mathematical cumbersome. In the remaining cases, Binomial and truncated Binomial are completely helpless. In this context, we have suggested an alternative approach of Binomial and Multinomial distribution having generalized sequence of the values of the random variables. For convenience the distributions are defined as relatively generalized distribution. The number of successes of the proposed distributions is represented by the arithmetic progression $a + nd$, where, a is non-negative integer and termed as minimum number of success, d is positive integer representing the concentration of success and n is a non-negative integer indicating the total number of trials.

To justify the sequence with real life situation, let us consider an example of number of defective shoes. It is well known that the shoes are produced pair wise. That is, if we make n attempts to identify the number of defective shoes, then it is usual that the number of defective shoes would occur pair wise. In this context, the number success in the form $0, 2, 4, \dots, 2n$ is justified. If it is known that the minimum number of defective items in n attempts are a , then these occurrences are not chance outcome and may be regarded as constant and then we are interested to find the probability of chance outcome taking form $a, a + 2, a + 4, \dots, 2n$, where $a > 0$ and n is non-negative integer. That is, number of defective shoes larger than a . The other possible sequences are also justifiable in this way by real life examples.

One may confuse the proposed form of the distribution with the usual truncated distribution. The major difference is that in our form the probability exists only for the possible number of success shown in the sequence. In the example discussed above, the defective number of shoes take the values $0, 2, 4, \dots, 2n$. This indicates that there is no existence of the number of defective item(s) of the form $1, 3, 5, \dots$ and thus no probabilities. On the other hand, in truncated distribution, there is existence of the number of success which is truncated and also they have the probabilities.

A number of authors published their work under the heading of generalized binomial distribution but they were different in terms of the key concept of our present work. Altham¹ showed two generalizations of the binomial distribution when the random variables are identically distributed but not independent and assumed to have symmetric joint distribution with no second or higher order “interactions”. Two generalizations are obtained depending on whether the “interaction” for discrete variables is “multiplicative” or “additive”. The distribution has a new parameter $\theta > 0$ which controls the shape of the distribution and flexible to allow for both over or under-disperse than traditional Binomial distribution. Whereas, the beta-binomial distribution allows only for over-disperse distribution than the corresponding Binomial distribution (Johnson, Kemp and Kotz²). Dwass³ have provided a unified approach to a family of discrete distributions that includes the hypergeometric, Binomial, and Polya distributions by considering the simple sample scheme where after each drawing there is a “replacement” whose magnitude is a fixed real number. Paul⁴ derived a new three parameter distribution, a generalization of the binomial, the beta-binomial and correlated beta-binomial distribution. Further a modification on beta-correlated binomial distribution was proposed by Raul⁵. In the generalization of the probability distribution, Panaretos and Xekalaki⁶ developed cluster binomial and multinomial model and their probability distributions. In his study, Madsen⁷ discussed that in many cases binomial distribution fails to apply because of more variability in the data than that can be explained by the distribution. He pointed out a characterization of sequences of exchangeable Bernoulli random variables which can be used to develop models which is more fluxion than the traditional binomial distribution. His study exhibited sufficient conditions which will yield such models and show how existing models can be combined to generate further models.

A generalization of the binomial distribution is introduced by Drezner and Farnum⁸ that allow the dependence between trials, non-constant probabilities of success from trial to trial and which contains usual binomial distribution as special case. A new departure in the generalization was carried out by Fu and Sproule⁹ by adopting the assumption that the underlying Bernoulli trials take on the values α or β where $\alpha < \beta$, rather than the conventional values 0 or 1. This rendered a four parameter binomial distribution of the form $B(n, p, \alpha, \beta)$. In a recent work, Altham and Hankin¹⁰ introduced two generalizations of multinomial and Binomial distributions which arise from the multiplicative Binomial distribution of Altham¹. The forms of the generalized distributions are of exponential family form and termed as “multivariate multinomial distribution” and “multivariate multiplicative binomial distribution”. Like the Altham’s generalized distribution, both the distribution has an additional shape parameter θ which corresponds usual distribution if it takes value 1 and over and under-disperse for greater and less than 1 respectively.

2. New Approach of Binomial Distribution

In the traditional Binomial sampling each time we draw a sample having either a success or a failure, we continue this method up to n trials and we may have number of successes starting from 0 and may ended up to n , but in real world, the success of Binomial distribution may not occur in the usual sequences rather it may happen as

- i) $0, 2, 4, \dots, 2n$
- ii) $2, 4, \dots, 2n$
- iii) $0, 3, 6, \dots, 3n$
- iv) $3, 6, 9, \dots, 3n$

and so on.

These indicate that the number of success may follow an arithmetic progression $a + nd$, where, a is a non-negative integer representing the minimum number of success, d is a positive integer representing the concentration of success occurring and n is a non-negative integer indicating the total number of trails.

Definition 2.1. A random variable X is said to have a relatively general binomial distribution if it has the following probability mass function

$$P(x; a, n, d, p) = \frac{\binom{a+nd}{x} p^x q^{a+nd-x}}{\sum_{x=a}^{a+nd} \binom{a+nd}{x} p^x q^{a+nd-x}} \quad ; \quad x = a, a+d, a+2d, \dots, a+nd \quad (2.1)$$

$$= \frac{1}{k} \binom{a+nd}{x} p^x q^{a+nd-x}$$

where, $a \geq 0$ is the minimum number of success, $d > 0$ is the concentration of occurrence of success, n is a predefined finite number of non-negative integer represents the number of trials and p is the probability of success such that $p + q = 1$ and $k = \sum_{x=a}^{a+nd} \binom{a+nd}{x} p^x q^{a+nd-x}$ is a constant. The probability mass function $P(x; a, n, d, p)$ stands for probability of getting x success out of maximum of $a + nd$ successes in n trials.

Theorem 2.1. Show that the probability mass function of generalized binomial distribution is a probability function.

Theorem 2.2. Show the relationship between generalized binomial distribution with parameter $a \geq 0, d > 0, n \geq 0$ and p and usual binomial distribution with parameter $n \geq 0$ and p .

It can be easily shown that the probability mass function of generalized binomial distribution reduces to the probability mass function usual binomial distribution when $a = 0$.

Theorem 2.3 Show that the moment generating function of generalized binomial distribution is

$$M_X(t) = \frac{\sum_{x=a}^{a+nd} \binom{a+nd}{x} (pe^t)^x q^{a+nd-x}}{\sum_{x=a}^{a+nd} \binom{a+nd}{x} p^x q^{a+nd-x}} \quad (2.2)$$

Theorem 2.4. The mean and variance of generalized binomial distribution are $(a + nd)p$ and $(a + nd)pq$ respectively.

Theorem 2.5. If X follows the generalized binomial distribution then find the 3rd and 4th raw and central moments of X respectively are

$$\mu'_3 = (a + nd)^3 p^3 - 3(a + nd)^2 p^3 + 2(a + nd)p^3 + 3(a + nd)^2 p^2 - 3(a + nd)p^2 + (a + nd)p \quad (2.4)$$

$$\mu'_4 = (a + nd)^4 p^4 - 6(a + nd)^3 p^4 + 11(a + nd)^2 p^4 - 6(a + nd)p^4 + 6(a + nd)^3 p^3 - 18(a + nd)^2 p^3 + 12(a + nd)p^3 + 7(a + nd)^2 p^2 - 7(a + nd)p^2 + (a + nd)p \quad (2.5)$$

$$\mu_3 = (a + nd)pq(1 - 2p) \quad (2.6)$$

$$\mu_4 = (a + nd)pq[1 + 3((a + nd) - 2)pq] \quad (2.7)$$

Still the special case for 3rd and 4th raw and central moments holds for $a = 0$ and $d = 1$ and turned to the form of usual Binomial distributions.

Theorem 2.6. The measures and coefficient of skewness and kurtosis of generalized binomial distribution are:

Skewness

$$\text{Measures of Skewness } (\beta_1) = \frac{(1 - 2p)^2}{(a + nd)pq} \quad (2.8)$$

$$\text{Coefficient of Skewness } (\gamma_1) = \sqrt{\beta_1} = \frac{1 - 2p}{\sqrt{(a + nd)pq}} \quad (2.9)$$

From coefficient of skewness, the following conclusion can be drawn and still surprising that the nature of skew depends only on p only and which is similar to the usual binomial distribution as:

- i) The distribution is positively skewed if $p < \frac{1}{2}$.
- ii) On the other hand, the distribution is negatively skewed if $p > \frac{1}{2}$.
- iii) And the distribution is symmetric if $p = \frac{1}{2}$.

Kurtosis

$$\text{Measures of kurtosis } (\beta_2) = 3 + \frac{(1 - 6pq)}{(a + nd)pq} \quad (2.10)$$

$$\text{Coefficient of kurtosis } (\gamma_2) = \beta_2 - 3 = \frac{(1 - 6pq)}{(a + nd)pq} \quad (2.11)$$

These equations tell us that the generalized distribution is

- i) Mesokurtic if $= \frac{1}{6}$.
- ii) Platykurtic $pq > \frac{1}{6}$ and
- iii) Leptokurtic if $pq < \frac{1}{6}$.

Theorem 2.7 Show that the maximum likelihood estimate of the parameter p is $\frac{x}{a+nd}$, where x is the total number of success from maximum of $a + nd$ success in n trials.

Theorem 2.8. The MLE estimator of number of trial of the distribution is $\hat{n} = \frac{1}{d} \left(\frac{\sum_{i=1}^n x_i}{\hat{p}} - a \right)$.

Theorem 2.9. Necessary conditions to derive the generalized Poisson distribution from generalized binomial distribution.

Generalized Poisson distribution can be derived from the generalized binomial distribution under the following assumptions:

- i) p , the probability of success in a Bernoulli trial is very small. i.e. $p \rightarrow 0$.
- ii) n , the number of trials is very large. i.e. $n \rightarrow \infty$.
- iii) $(a + nd)p = \lambda$ is finite constant, that is average number of success is finite. Under this condition, we have

$$(a + nd)p = \lambda \quad \therefore p = \frac{\lambda}{(a+nd)} \text{ and } q = 1 - \frac{\lambda}{(a+nd)}$$

Theorem 2.10 Normal distribution as a limiting form of generalized binomial distribution.

3. New Approach of Multinomial Distribution

Under the sampling scheme described in Section 2, consider the situation where one of the k mutually exclusive outcomes is possible from a single trial other than only success or failure. More specifically, if the k outcomes are denoted by $e_1, e_2, e_3, \dots, e_k$ and the number of occurrences of the respective outcomes are denoted by x_1, x_2, \dots, x_k such that $\sum_{i=1}^k x_i = a + nd$, where a is non-negative integer and termed as minimum number of success, d is positive integer representing the concentration of success and n is a non-negative integer indicating the total number of trials, then our suggested general form of Binomial as well as traditional Multinomial distribution cannot provide the probability that the event e_1 occurred x_1 times, the event e_2 occurred x_2 times and so on the event e_k occurred x_k times. To overcome this situation, we have suggested the new approach of Multinomial distribution and termed as relatively more general Multinomial or generalized Multinomial distribution. In this section we would present only the definition of the suggested distribution and statement of the theorems that we have derived for our present work to keep the paper size standard and other reason is that the derivations are very much similar to the suggested Binomial distribution.

Definition 3.1. k discrete random variable X_1, X_2, \dots, X_k is said to have a *generalized multinomial distribution* if it has the following probability mass function

$$\begin{aligned}
 P(x_1, x_2, \dots, x_k; a, n, d, p_1, p_2, \dots, p_k) &= \frac{\frac{1}{x_1! x_2! \dots x_k!} p_1^{x_1} p_2^{x_2} \dots p_k^{x_k}}{\sum_{x_1, x_2, \dots, x_k=a}^{a+nd} \frac{1}{x_1! x_2! \dots x_k!} p_1^{x_1} p_2^{x_2} \dots p_k^{x_k}} \\
 &= \frac{\frac{(a+nd)!}{x_1! x_2! \dots x_k!} p_1^{x_1} p_2^{x_2} \dots p_k^{x_k}}{\sum_{x_1, x_2, \dots, x_k=a}^{a+nd} \frac{(a+nd)!}{x_1! x_2! \dots x_k!} p_1^{x_1} p_2^{x_2} \dots p_k^{x_k}} \quad (3.1)
 \end{aligned}$$

where, $a \geq 0, d > 0, n \geq 0$ and p_1, p_2, \dots, p_k such that $\sum_{i=1}^k p_i = 1$, are the parameters of the distribution. and $\sum_{i=1}^k x_i = a + nd$.

The defined function is a probability mass function as it holds the conditions of the probability function and easily reduces to traditional Multinomial distribution if $a = 0$ and $d = 1$ and thus may be said that the traditional Multinomial distribution is a special case of the suggested Multinomial distribution.

Theorem 3.1. *Generalized Binomial distribution is a special case of generalized Multinomial distribution.*

Theorem 3.2. *The moment generating function of generalized Multinomial distribution is*

$$M_{X_1, X_2, \dots, X_k}(t_1, t_2, \dots, t_k) = \frac{\sum_{x_1, x_2, \dots, x_k=a}^{a+nd} \frac{(a+nd)!}{\prod_{i=1}^k x_i!} \prod_{i=1}^k (p_i e^{t_i})^{x_i}}{\sum_{x_1, x_2, \dots, x_k=a}^{a+nd} \frac{(a+nd)!}{\prod_{i=1}^k x_i!} \prod_{i=1}^k p_i^{x_i}} \quad (3.3)$$

and the mean and the variance of X_i ($i = 1, 2, \dots, k$) respectively are

$$E(X_i) = \mu'_{1_i} = (a + nd)p_i \quad (3.4)$$

$$V(X_i) = (a + nd)p_i(1 - p_i) \quad (3.5)$$

Theorem 3.3. *The 3rd and 4th raw and central moments of generalized Multinomial distribution are*

$$\mu'_{3_i} = (a + nd)^3 p_i^3 - 3(a + nd)^2 p_i^3 + 2(a + nd)p_i^3 + 3(a + nd)^2 p_i^2 - 3(a + nd)p_i^2 + (a + nd)p_i \quad (3.6)$$

$$\begin{aligned}
 \mu'_{4_i} &= (a + nd)^4 p_i^4 - 6(a + nd)^3 p_i^4 + 11(a + nd)^2 p_i^4 - 6(a + nd)p_i^4 + 6(a + nd)^3 p_i^3 \\
 &\quad - 18(a + nd)^2 p_i^3 + 12(a + nd)p_i^3 + 7(a + nd)^2 p_i^2 - 7(a + nd)p_i^2 \\
 &\quad + (a + nd)p_i \quad (3.7)
 \end{aligned}$$

$$\mu_{3_i} = (a + nd)p_i(1 - p_i)(1 - 2p_i) \quad (3.8)$$

and

$$\mu_4 = (a + nd)p_i(1 - p_i)[1 + 3((a + nd) - 2)p_i(1 - p_i)] \quad (3.9)$$

In all the moments if we put $a = 0$ and $d = 1$ then these reduces to the form of moments of traditional Multinomial distribution which again justifies that the traditional Multinomial distribution is a special case of our suggested generalized Multinomial distribution.

Theorem 3.4. *Karl Pearson's measures and coefficients of skewness and kurtosis of generalized multinomial distribution.*

Measure of Skewness:

$$\beta_{1i} = \frac{(1 - 2p_i)^2}{(a + nd)p_i(1 - p_i)} \quad (3.10)$$

Coefficient of Skewness:

$$\gamma_{1i} = \sqrt{\beta_{1i}} = \frac{1 - 2p_i}{\sqrt{(a + nd)p_i(1 - p_i)}} \quad (3.11)$$

Measure of Kurtosis:

$$\beta_{2i} = 3 + \frac{\{1 - 6p_i(1 - p_i)\}}{(a + nd)p_i(1 - p_i)} \quad (3.12)$$

Coefficient of Kurtosis

$$\gamma_{2i} = \beta_{2i} - 3 = \frac{\{1 - 6p_i(1 - p_i)\}}{(a + nd)p_i(1 - p_i)} \quad (3.13)$$

The conclusion from the measures and coefficient of skewness and kurtosis of the generalized Multinomial distribution can be drawn in the same way as drawn in the generalized Binomial distribution depending on the values of p_i .

Theorem 3.5. *The maximum likelihood estimator of the parameters of generalized multinomial distribution is $\hat{p}_i = \frac{x_i}{(a+nd)}$, where x_i is the number of success comprising $(a + nd - x_i)$ is the total number of failure subject to condition that maximum number of success is $a + nd$.*

4. Discussion and Conclusion

We have suggested relatively more general form of two discrete distributions such as Binomial and Multinomial for the different sampling scheme which is described above and termed as generalized distribution. It is evident from the generalized distribution that if sampling is drawn in the usual manner, then our suggested distributions reduces to the traditional form and thus it may conclude that the traditional Binomial and Multinomial distribution are the special cases of our proposed generalized Binomial and Multinomial distribution. Like the traditional distributions, all of the distributional properties including limiting theorems of the suggested distributions have derived. The truncated cases of the traditional distribution can be address accurately by our new approach of the distributions. In general, the new approach of the distributions are providing more access and broaden the scope from the theoretical point of view as well as from the standpoint of real world problem solving. Generalized sequence of the number of success of the proposed distributions may be considered as an added advantage in the distribution theory.

References

- [1] P. M. E. Altham, Two Generalizations of the Binomial Distribution, *J. R. Stat. Soc. Ser. C. Appl. Stat.* **27 (2)** 1978 162-167.
- [2] N. L. Johnson, A. W. Kemp and S. Kotz, *Univariate Discrete Distributions*. 3rd edn. (John Wiley & Sons, New Jersey, 2005).
- [3] M. Dwass, A Generalized Binomial Distribution, *Amer. Statist.* **33(2)** 1979 86-87.

- [4] S. R. Paul, A Three-Parameter Generalization of the Binomial Distribution, *Comm. Statist. Theory Methods* **14(6)** 1985 1497-1506.
- [5] S. R. Raul, On the beta-correlated binomial (bcb) distribution-a three parameter generalization of the binomial distribution, *Comm. Statist. Theory and Methods* **16(5)** 1987 1473-1478.
- [6] J. Panaretos. and E. Xekalaki, On Generalized Binomial and Multinomial Distributions and Their Relation to Generalized Poisson Distribution, *Ann. Inst. Statist. Math.* **38(2)** 1986 223-231.
- [7] R. W. Madsen, Generalized binomial distributions, *Comm. Statist. Theory and Methods* **22(11)** 1993 3065-3086.
- [8] Z. Drezner and N. Farnum, A generalized binomial distribution, *Comm. Statist. Theory and Methods* **22(11)** 1993 3051-3063.
- [9] J. Fu and R. Sproule, A generalization of the binomial distribution, *Comm. Statist. Theory and Methods* **24(10)** 1995 2645-2658.
- [10] P. M. E. Altham and R. K. S. Hankin, Multivariate Generalizations of the Multiplicative Binomial Distribution: Introducing the MM Package, *J. Statist. Software.* **46(12)** 2012 1-23.

Appendix

Proof of Theorem 2.1. We know, the probability mass function of the suggested generalized binomial distribution is

$$P(x; a, n, d, p) = \frac{\binom{a+nd}{x} p^x q^{a+nd-x}}{\sum_{x=a}^{a+nd} \binom{a+nd}{x} p^x q^{a+nd-x}} ; x = a, a+d, a+2d, \dots, a+nd \quad (2.1)$$

It is clear from the function that $P(x; a, n, d, p) \geq 0$ for all values of X with different values of a, d, n and p .

Again,

$$\begin{aligned} & \sum_{x=a}^{a+nd} P(x; a, n, d, p) \\ &= \sum_{x=a}^{a+nd} \frac{\binom{a+nd}{x} p^x q^{a+nd-x}}{\sum_{x=a}^{a+nd} \binom{a+nd}{x} p^x q^{a+nd-x}} \\ &= \frac{1}{\sum_{x=a}^{a+nd} \binom{a+nd}{x} p^x q^{a+nd-x}} \sum_{x=a}^{a+nd} \binom{a+nd}{x} p^x q^{a+nd-x} \\ &= 1 \end{aligned}$$

As $P(x; a, n, d, p) \geq 0$ and $\sum_{x=a}^{a+nd} P(x; a, n, d, p) = 1$, so we may conclude that $P(x; a, n, d, p)$ is a probability function.

Proof of Theorem 2.3. According to the definition of moment generating function,

$$\begin{aligned} M_X(t) &= E(e^{tX}) = \sum_{x=a}^{a+nd} e^{tx} P(x; a, n, d, p) \\ \Rightarrow M_X(t) &= \sum_{x=a}^{a+nd} e^{tx} \frac{1}{k} \binom{a+nd}{x} p^x q^{a+nd-x} \end{aligned}$$

$$\begin{aligned} \Rightarrow M_X(t) &= \frac{1}{k} \sum_{x=a}^{a+nd} \binom{a+nd}{x} (pe^t)^x q^{a+nd-x} \\ \therefore M_X(t) &= \frac{\sum_{x=a}^{a+nd} \binom{a+nd}{x} (pe^t)^x q^{a+nd-x}}{\sum_{x=a}^{a+nd} \binom{a+nd}{x} p^x q^{a+nd-x}} \end{aligned} \quad (2.2)$$

This is the moment generating function of generalized binomial distribution.

Proof of Theorem 2.4. The mean of the distribution can be calculated from its moment generating function by differentiating it with respect to t and equating $t = 0$. That is

$$\begin{aligned} E(X) = \mu'_1 &= \left. \frac{d}{dt} [M_X(t)] \right|_{t=0} \\ &= \left[\frac{\sum_{x=a}^{a+nd} \binom{a+nd}{x} x (pe^t)^{x-1} (pe^t) q^{a+nd-x}}{\sum_{x=a}^{a+nd} \binom{a+nd}{x} p^x q^{a+nd-x}} \right]_{t=0} \\ &= \left[\frac{\sum_{x=a}^{a+nd} x \binom{a+nd}{x} (pe^t)^x q^{a+nd-x}}{\sum_{x=a}^{a+nd} \binom{a+nd}{x} p^x q^{a+nd-x}} \right]_{t=0} \\ &= \frac{\sum_{x=a}^{a+nd} x \binom{a+nd}{x} p^x q^{a+nd-x}}{\sum_{x=a}^{a+nd} \binom{a+nd}{x} p^x q^{a+nd-x}} \\ &= \frac{\sum_{x=a}^{a+nd} x \frac{(a+nd)(a+nd-1)!}{x(x-1)!(a+nd-1-(x-1))!} p^{x-1} q^{a+nd-1-(x-1)}}{\sum_{x=a}^{a+nd} \binom{a+nd}{x} p^x q^{a+nd-x}} \\ &= \frac{(a+nd)p \left\{ \sum_{x-1=a-1}^{a+nd-1} \frac{(a+nd-1)!}{(x-1)!(a+nd-1-(x-1))!} p^{x-1} q^{a+nd-1-(x-1)} \right\}}{\sum_{x=a}^{a+nd} \binom{a+nd}{x} p^x q^{a+nd-x}} \\ &= \frac{(a+nd)p \left\{ \sum_{x-1=a-1}^{a+nd-1} \binom{a+nd-1}{x-1} p^{x-1} q^{a+nd-1-(x-1)} \right\}}{\sum_{x=a}^{a+nd} \binom{a+nd}{x} p^x q^{a+nd-x}} \\ &= \frac{(a+nd)p \left\{ \sum_{x=a}^{a+nd} \binom{a+nd}{x} p^x q^{a+nd-x} \right\}}{\sum_{x=a}^{a+nd} \binom{a+nd}{x} p^x q^{a+nd-x}} \\ &= (a+nd)p \end{aligned} \quad (2.3)$$

which is the mean of the distribution.

Similarly, it can be found that the variance of the distribution is $(a+nd)pq$. If we substitute $a = 0$ and $d = 1$ in mean and variance of the generalized distribution then it reduces to mean and variance of usual distribution which holds the property that usual distribution is special case of the generalized distribution.

Proof of Theorem 2.7. We know, the probability mass function of binomial distribution itself is a likelihood function. Therefore, the like function of generalized binomial distribution is

$$L = \frac{\binom{a+nd}{x} p^x q^{a+nd-x}}{\sum_{x=a}^{a+nd} \binom{a+nd}{x} p^x q^{a+nd-x}} = \frac{1}{k} \binom{a+nd}{x} p^x q^{a+nd-x} \quad (2.12)$$

where, x is the total number of success from maximum of $a + nd$ success in n trials.

Taking logarithm in both sides, we have

$$\begin{aligned} \ln L &= \ln \left[\frac{1}{k} \binom{a+nd}{x} p^x q^{a+nd-x} \right] \\ \Rightarrow \ln L &= -\ln k + \ln \binom{a+nd}{x} + x \log p + (a+nd-x) \ln(1-p) \end{aligned} \quad (2.13)$$

Differentiating equation (2.13) with respect to p and equating to zero, we get

$$\begin{aligned} \frac{\partial}{\partial p} (\ln L) &= \frac{x}{\hat{p}} + \frac{a+nd-x}{1-\hat{p}} (-1) = 0 \\ \Rightarrow \frac{x}{\hat{p}} - \frac{a+nd-x}{1-\hat{p}} &= 0 \\ \Rightarrow \frac{x - \hat{p}x - (a+nd)\hat{p} + \hat{p}x}{\hat{p}(1-\hat{p})} &= 0 \\ \Rightarrow x - (a+nd)\hat{p} &= 0 \\ \Rightarrow (a+nd)\hat{p} &= x \\ \Rightarrow \hat{p} &= \frac{x}{a+nd} \end{aligned} \quad (2.14)$$

Hence, the maximum likelihood estimate of the parameter p is $\frac{x}{a+nd}$, where x is the total number of success from maximum of $a + nd$ success in n trials.

Alternative Proof. The likelihood function of generalized binomial distribution can be written as

$$\begin{aligned} L &= \left\{ \binom{a+d}{x_1} p^{x_1} q^{a+d-x_1} \right\} \left\{ \binom{d}{x_2} p^{x_2} q^{d-x_2} \right\} \left\{ \binom{d}{x_3} p^{x_3} q^{d-x_3} \right\} \dots \left\{ \binom{d}{x_n} p^{x_n} q^{d-x_n} \right\} \\ \Rightarrow L &= p^{\sum_{i=1}^n x_i} q^{a+nd-\sum_{i=1}^n x_i} \left\{ \binom{a+d}{x_1} \binom{d}{x_2} \binom{d}{x_3} \dots \binom{d}{x_n} \right\} \\ \Rightarrow L &= C \cdot p^{\sum_{i=1}^n x_i} (1-p)^{a+nd-\sum_{i=1}^n x_i} \end{aligned} \quad (2.15)$$

where, $C = \binom{a+d}{x_1} \binom{d}{x_2} \binom{d}{x_3} \dots \binom{d}{x_n}$ is a constant.

Taking logarithm in both sides, we have

$$\ln L = \ln [C \cdot p^{\sum_{i=1}^n x_i} (1-p)^{a+nd-\sum_{i=1}^n x_i}]$$

$$= \ln C + \sum_{i=1}^n x_i \ln p + \left(a + nd - \sum_{i=1}^n x_i \right) \ln(1 - p) \quad (2.16)$$

Differentiating equation (2.16) with respect to p and equating to zero, we get

$$\begin{aligned} \frac{\delta}{\delta p}(\ln L) &= \frac{\sum_{i=1}^n x_i}{\hat{p}} + \frac{(a + nd - \sum_{i=1}^n x_i)}{1 - \hat{p}} (-1) = 0 \\ \Rightarrow \frac{\sum_{i=1}^n x_i - \hat{p} \sum_{i=1}^n x_i - (a + nd)\hat{p} + \hat{p} \sum_{i=1}^n x_i}{\hat{p}(1 - \hat{p})} &= 0 \\ \Rightarrow \sum_{i=1}^n x_i - (a + nd)\hat{p} &= 0 \\ \Rightarrow (a + nd)\hat{p} &= \sum_{i=1}^n x_i \\ \Rightarrow \hat{p} &= \frac{\sum_{i=1}^n x_i}{a + nd} \end{aligned} \quad (2.17)$$

where, $\sum_{i=1}^n x_i = x$ is the total number of success out of maximum of $a + nd$ success in n trials.

Proof of Theorem 2.8. If the probability of success p is known from prior knowledge or estimated on the basis of observation then we can also estimate the number of trial n . Let us define \tilde{p} is the probability of success which is either known or estimated. Therefore,

$$\begin{aligned} \tilde{p} &= \frac{\sum_{i=1}^n x_i}{a + \hat{n}d} \\ \Rightarrow a + \hat{n}d &= \frac{\sum_{i=1}^n x_i}{\tilde{p}} \\ \Rightarrow \hat{n}d &= \frac{\sum_{i=1}^n x_i}{\tilde{p}} - a \\ \Rightarrow \hat{n} &= \frac{1}{d} \left(\frac{\sum_{i=1}^n x_i}{\tilde{p}} - a \right) \end{aligned} \quad (2.18)$$

Again, the value of a is the initial number of success or in other words, the number success in first trial. Whereas, d is the difference between the occurred number of success in successive trials.

Proof of Theorem 2.10. Normal distribution can be derived from the generalized binomial distribution under the following conditions:

- i) The probability of success p or the probability of failure q are not so small and
- ii) n , the number of trails is very large i.e. n tends to infinity.

We know, the probability mass function of generalized binomial variate X with parameter a , n , d , and p is

$$P(x; a, n, d, p) = \frac{\binom{a + nd}{x} p^x q^{a + nd - x}}{\sum_{x=a}^{a + nd} \binom{a + nd}{x} p^x q^{a + nd - x}} ; x = a, a + d, a + 2d, \dots, a + nd$$

$$= \frac{\frac{(a+nd)!}{x!(a+nd-x)!} p^x q^{a+nd-x}}{\sum_{x=a}^{a+nd} \frac{(a+nd)!}{x!(a+nd-x)!} p^x q^{a+nd-x}} \quad (2.19)$$

Now, let us consider the standard normal variate as

$$z = \frac{x - E(X)}{\sqrt{\text{var}(X)}} = \frac{x - (a+nd)p}{\sqrt{(a+nd)pq}} \quad (2.20)$$

When $x = a$, $z = \frac{a - (a+nd)p}{\sqrt{(a+nd)pq}} = -\sqrt{\frac{(a+nd)p}{q}} + \frac{a}{\sqrt{(a+nd)pq}}$
 and when $x = a + nd$, $z = \frac{a+nd - (a+nd)p}{\sqrt{(a+nd)pq}} = \frac{(a+nd)q}{\sqrt{(a+nd)pq}} = \sqrt{\frac{(a+nd)q}{p}}$

Thus, in the limit as $n \rightarrow \infty$ Z takes the values $-\infty$ to ∞ Hence the distribution of Z will be a continuous distribution over the range $-\infty$ to ∞ with mean zero and variance unity.

Now, under conditions (i) and (ii), we shall find the limiting form of the (12.1) by applying the Stirling's approximation for factorials. For large n , equation (12.1) can be written as

$$\lim_{n \rightarrow \infty} P(x; a, n, d, p) = \lim_{n \rightarrow \infty} \frac{\frac{(a+nd)!}{x!(a+nd-x)!} p^x q^{a+nd-x}}{\sum_{x=a}^{a+nd} \frac{(a+nd)!}{x!(a+nd-x)!} p^x q^{a+nd-x}} \quad (2.21)$$

Let us first work with numerator only

$$\begin{aligned} & \lim_{n \rightarrow \infty} \frac{(a+nd)!}{x!(a+nd-x)!} p^x q^{a+nd-x} \\ &= \lim_{n \rightarrow \infty} \frac{\sqrt{2\pi} e^{-(a+nd)} (a+nd)^{a+nd+\frac{1}{2}}}{\sqrt{2\pi} e^{-x} x^{x+\frac{1}{2}} \sqrt{2\pi} e^{-(a+nd-x)} (a+nd-x)^{a+nd-x+\frac{1}{2}}} p^x q^{a+nd-x} \\ &= \lim_{n \rightarrow \infty} \frac{(a+nd)^{a+nd+\frac{1}{2}}}{\sqrt{2\pi} x^{x+\frac{1}{2}} (a+nd-x)^{a+nd-x+\frac{1}{2}}} p^x q^{a+nd-x} \\ &= \lim_{n \rightarrow \infty} \frac{\{(a+nd)p\}^{x+\frac{1}{2}} \{(a+nd)q\}^{a+nd-x+\frac{1}{2}}}{\sqrt{2\pi} \sqrt{(a+nd)pq} x^{x+\frac{1}{2}} (a+nd-x)^{a+nd-x+\frac{1}{2}}} \\ &= \lim_{n \rightarrow \infty} \frac{1}{\sqrt{2\pi} \sqrt{(a+nd)pq}} \left\{ \frac{(a+nd)p}{x} \right\}^{x+\frac{1}{2}} \left\{ \frac{(a+nd)q}{(a+nd-x)} \right\}^{a+nd-x+\frac{1}{2}} \end{aligned} \quad (2.22)$$

We have from equation (2.20)

$$x = (a+nd)p + z\sqrt{(a+nd)pq}$$

$$\Rightarrow \frac{x}{(a+nd)p} = 1 + z \sqrt{\frac{q}{(a+nd)p}}$$

and

$$\frac{a+nd-x}{(a+nd)p} = 1 - z \sqrt{\frac{p}{(a+nd)q}}$$

Again,

$$\frac{dx}{dz} = \sqrt{(a+nd)pq}$$

From equation (2.22) we have

$$\begin{aligned} & \lim_{n \rightarrow \infty} \frac{(a+nd)!}{x! (a+nd-x)!} p^x q^{a+nd-x} \\ &= \lim_{n \rightarrow \infty} \frac{1}{\sqrt{2\pi} \sqrt{(a+nd)pq}} \left\{ \frac{(a+nd)p}{x} \right\}^{x+\frac{1}{2}} \left\{ \frac{(a+nd)q}{(a+nd-x)} \right\}^{a+nd-x+\frac{1}{2}} \\ &= \lim_{n \rightarrow \infty} \frac{\left\{ \frac{1}{1+z\sqrt{\frac{q}{(a+nd)p}}} \right\}^{(a+nd)p+z\sqrt{(a+nd)pq}+\frac{1}{2}} \left\{ \frac{1}{1-z\sqrt{\frac{p}{(a+nd)q}}} \right\}^{a+nd-(a+nd)p-z\sqrt{(a+nd)pq}+\frac{1}{2}}}{\sqrt{2\pi}} \\ &= \lim_{n \rightarrow \infty} \frac{\left\{ \frac{1}{1+z\sqrt{\frac{q}{(a+nd)p}}} \right\}^{(a+nd)p+z\sqrt{(a+nd)pq}+\frac{1}{2}} \left\{ \frac{1}{1-z\sqrt{\frac{p}{(a+nd)q}}} \right\}^{(a+nd)q-z\sqrt{(a+nd)pq}+\frac{1}{2}}}{\sqrt{2\pi}} \\ &= \lim_{n \rightarrow \infty} \frac{1}{\sqrt{2\pi} N} \end{aligned} \tag{2.23}$$

where,

$$\begin{aligned} N &= \left\{ 1 + z \sqrt{\frac{q}{(a+nd)p}} \right\}^{(a+nd)p+z\sqrt{(a+nd)pq}+\frac{1}{2}} \left\{ 1 - z \sqrt{\frac{p}{(a+nd)q}} \right\}^{(a+nd)q-z\sqrt{(a+nd)pq}+\frac{1}{2}} \\ &\Rightarrow \log N = \left\{ (a+nd)p + z\sqrt{(a+nd)pq} + \frac{1}{2} \right\} \log \left\{ 1 + z \sqrt{\frac{q}{(a+nd)p}} \right\} \\ &\quad + \left\{ (a+nd)q - z\sqrt{(a+nd)pq} + \frac{1}{2} \right\} \log \left\{ 1 - z \sqrt{\frac{p}{(a+nd)q}} \right\} \end{aligned}$$

$$\begin{aligned}
 \Rightarrow \log N &= \left[(a+nd)p + z\sqrt{(a+nd)pq} \right. \\
 &\quad \left. + \frac{1}{2} \right] \log \left[z\sqrt{\frac{q}{(a+nd)p}} - \frac{z^2}{2} \frac{q}{(a+nd)p} + \frac{z^3}{3} \left\{ \frac{q}{(a+nd)p} \right\}^{3/2} + \dots \right] \\
 &+ \left[(a+nd)q - z\sqrt{(a+nd)pq} + \frac{1}{2} \right] \log \left[-z\sqrt{\frac{p}{(a+nd)q}} - \frac{z^2}{2} \frac{p}{(a+nd)q} - \frac{z^3}{3} \left\{ \frac{p}{(a+nd)q} \right\}^{3/2} + \dots \right] \\
 \Rightarrow \log N &= \left[z\sqrt{(a+nd)pq} - \frac{z^2}{2} q + z^2 q + \frac{z}{2} \sqrt{\frac{q}{(a+nd)p}} - \frac{z^2}{4} \frac{q}{(a+nd)p} + \dots \right] \\
 &+ \left[-z\sqrt{(a+nd)pq} - \frac{z^2}{2} p + z^2 p - \frac{z}{2} \sqrt{\frac{p}{(a+nd)q}} - \frac{z^2}{4} \frac{p}{(a+nd)q} + \dots \right] \\
 \Rightarrow \log N &= \frac{z}{2\sqrt{(a+nd)}} \left(\sqrt{\frac{q}{p}} - \sqrt{\frac{p}{q}} \right) + \frac{z^2}{2} - \frac{z^2}{4(a+nd)} \left(\frac{q}{p} - \frac{p}{q} \right) + \dots
 \end{aligned}$$

Hence, as limit $n \rightarrow \infty$, we get $\log N = \frac{z^2}{2} \Rightarrow N = e^{\frac{z^2}{2}}$.

Now, substituting the value of N in equation (12.5) we have

$$\lim_{n \rightarrow \infty} \frac{(a+nd)!}{x!(a+nd-x)!} p^x q^{a+nd-x} = \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}}$$

Putting these value in equation (12.3) we have

$$\lim_{n \rightarrow \infty} P(x; a, n, d, p) = f(z) = \frac{\frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}}}{\int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}}} = \frac{1}{\sqrt{2\pi}} e^{-\frac{z^2}{2}} \quad ; -\infty < z < \infty$$

which is the probability density function of standard normal variate Z .

If we consider the transformation $z = \frac{x-(a+nd)p}{\sqrt{(a+nd)pq}}$ then the probability density function of X will be

$$f(x) = \frac{1}{\sqrt{2\pi} \sqrt{(a+nd)pq}} e^{-\frac{1}{2} \left(\frac{x-(a+nd)p}{\sqrt{(a+nd)pq}} \right)^2} \quad ; -\infty < x < \infty \quad (2.24)$$

If we denote the mean of X by μ and variance of X by σ^2 then the probability density function become

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2} \left(\frac{x-\mu}{\sigma} \right)^2} \quad ; -\infty < x < \infty$$

Hence, the proof of the theorem.

Proof of Theorem 3.1. The form of the probability mass function of generalized Multinomial distribution is

$$P(x_1, x_2, \dots, x_k; a, n, d, p_1, p_2, \dots, p_k) = \frac{\frac{1}{x_1! x_2! \dots x_k!} p_1^{x_1} p_2^{x_2} \dots p_k^{x_k}}{\sum_{x_1, x_2, \dots, x_k=a}^{a+nd} \frac{1}{x_1! x_2! \dots x_k!} p_1^{x_1} p_2^{x_2} \dots p_k^{x_k}}$$

Considering $k = 2$ such that $x_1 + x_2 = a + nd$ and $p_1 + p_2 = 1$, we obtain

$$\begin{aligned} P(x_1, x_2; a, n, d, p_1, p_2) &= \frac{\frac{(a+nd)!}{x_1! x_2!} p_1^{x_1} p_2^{x_2}}{\sum_{x_1, x_2=a}^{a+nd} \frac{(a+nd)!}{x_1! x_2!} p_1^{x_1} p_2^{x_2}} \\ \Rightarrow P(x_1; a, n, d, p_1, p_2) &= \frac{\frac{(a+nd)!}{x_1! (a+nd-x_1)!} p_1^{x_1} p_2^{a+nd-x_1}}{\sum_{x_1=a}^{a+nd} \frac{(a+nd)!}{x_1! (a+nd-x_1)!} p_1^{x_1} p_2^{a+nd-x_1}} \end{aligned}$$

Letting $x_1 = x$, $p_1 = p$ and $p_2 = 1 - p_1 = 1 - p = q$, we have

$$\begin{aligned} P(x; a, n, d, p) &= \frac{\frac{(a+nd)!}{x! (a+nd-x)!} p^x q^{a+nd-x}}{\sum_{x=a}^{a+nd} \frac{(a+nd)!}{x! (a+nd-x)!} p^x q^{a+nd-x}} \\ \Rightarrow P(x; a, n, d, p) &= \frac{\binom{a+nd}{x} p^x q^{a+nd-x}}{\sum_{x=a}^{a+nd} \binom{a+nd}{x} p^x q^{a+nd-x}} \end{aligned}$$

Hence,

$$P(x; a, n, d, p) = \frac{\binom{a+nd}{x} p^x q^{a+nd-x}}{\sum_{x=a}^{a+nd} \binom{a+nd}{x} p^x q^{a+nd-x}} ; x = a, a+d, a+2d, \dots, a+nd \quad (3.2)$$

which is the probability mass function of generalized Binomial distribution.