International Journal of Statistics and Probability; Vol. 11, No. 3; May 2022
ISSN 1927-7032 E-ISSN 1927-7040
Published by Canadian Center of Science and Education

Approximation of the Binomial Probability Function Using the Discrete Normal Distribution

Mohammad Fraiwan Al-Saleh[1] & Doaa Suhail Obeidat[2]

[1] Department of Statistics, Yarmouk University, Jordan. Currently at Department of Mathematics and Statistics, Jordan University of Science and Technology, Jordan

[2] Department of Statistics, Yarmouk University, Jordan

Correspondence: Mohammad Fraiwan Al-Saleh, Department of Statistics, Yarmouk University, Jordan

Received: March 17, 2022 Accepted: April 19, 2022 Online Published: April 21, 2022

doi:10.5539/ijsp.v11n3p32 URL: https://doi.org/10.5539/ijsp.v11n3p32

Abstract

A new method of approximating the Binomial probability function is introduced. The method is based on the discrete normal distribution. In particular, the discrete normal probability function is used to approximate the binomial probability function. The new approximation is compared with the exact values and the approximation based on Central limit theorem. The maximum absolute error of the approximation is used to measure the accuracy of the method. It turned out that this method of approximation is useful and easy to use in practice. Also, the result can be an important theoretical statistical result that can be used in educational statistics.

Keywords: binomial probability function, discrete normal distribution, continuity correction; maximum absolute error, central limit theorem, maximum absolute error

1. Introduction and Some Closely Related Works

The Normal distribution is extremely important in statistics and is often used in practice as an approximate distribution for the distributions of real-valued random variables whose distributions are unknown. X is normally distributed with mean μ and variance σ^2, $X \sim N(\mu,\sigma^2)$, if it has the probability density function (pdf):

$$f(x;\mu,\sigma^2) = \frac{1}{\sigma\sqrt{2\pi}} e^{\frac{-(x-\mu)^2}{2\sigma^2}}, -\infty < x < \infty, \quad -\infty < \mu < \infty, \quad 0 < \sigma^2 < \infty.$$

The curve is bell-shaped, symmetric; mean=median= mode. If $\mu = 0$ & $\sigma = 1$, then X is a standard normal random variable; ϕ & Φ denotes pdf and the cumulative distribution function(cdf) of the N(0,1), respectively.

Using the normal distribution to approximate other distribution is a very old topic; it goes back to more than 300 years. In 1730, DeMoivre consider the problem of approximating the binomial probability function(pf) by the normal distribution. For more details, more explanation and references about the work on this topic see Govindarajulu (1965). Normal approximation of the main discrete distributions (binomial, Poisson, negative binomial, hypergeometric, etc.) was considered by Govindarajulu (1965).

The main comments that usually raised by users of statistics about the approximation of the binomial probability function(bbf) using the normal distribution are:

(1) We are approximating a discrete probability function using a continuous one;

(2) The bbf is skewed (except for p=0.5), while the normal pdf is symmetric.

(3) We need a correction factor to use this method of approximation.

Motivated by maximum entropy distribution, the "Discrete Normal Distribution" was first introduced by Kemp (1997). Consider the probability function(pf) of a discrete random variable X given by:

$$f(x;\lambda,q) = P(X = x) = \frac{\lambda^x q^{x(x-1)/2}}{\sum_{x=-\infty}^{\infty} \lambda^x q^{x(x-1)/2}}, \qquad x = 0, \pm 1, \pm 2, \ldots$$

where λ and q are parameters, $\lambda > 0$, $q \in (0,1)$. **Szablowski (2001)** took $q = e^{\frac{-1}{\sigma^2}}$ & $\lambda = e^{\left(\frac{-1}{2\sigma^2} + \frac{\alpha}{\sigma^2}\right)}$, $\alpha \in \Re$, which gives the discrete normal pdf. Thus, X is discrete normal, $X \sim DN(\mu, \sigma^2)$, if its pdf is

$$f(x) = P(X = x) = \frac{\frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-\mu)^2}{2\sigma^2}}}{\sum_{i=-\infty}^{\infty} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(i-\mu)^2}{2\sigma^2}}}, \quad x = 0, \pm 1, \pm 2, \ldots$$

For another discrete analogue of the Laplace distribution, see **Seidu et al. (2004).**

A Binomial r.v. X, $b(n, p)$, has the pf:

$$f(x; n, p) = P(X = x) = \binom{n}{x} p^x (1-p)^{n-x}, x = 0, 1, \ldots, n \quad \text{\& 0 O.w.}$$

$\mu = np$ & $\sigma^2 = np(1-p)$. Clearly, the variance is smaller than the mean with maximum value of $n/4$ occurs at $p = 0.5$.

It is easy nowadays to find the value of the pf at any value of X using calculators. However, there is a theoretical interest, when teaching statistical courses, to approximate the binomial using some well-known distribution such as normal, Poisson, etc. These approximations are mentioned in almost any text book in statistics and probability. Several ways have been used to approximate the binomial pf. The most popular way is by using the normal distribution with $\mu = np$ & $\sigma^2 = np(1-p)$. This is justified by the central limit theorem (CLT). The following is the simplest form of CLT:

If $X_1, X_2, \ldots, X_n$ are iid with $E(X_i) = \mu$ & $Var(X_i) = \sigma^2$, $-\infty < \mu < \infty$, $0 < \sigma^2 < \infty$, then as $n \to \infty$, $\frac{\sqrt{n}(\bar{X} - \mu)}{\sigma} \xrightarrow{d} N(0,1)$.

Now, if $X \sim b(n, p)$, then $X = \sum_{i=1}^{n} X_i$, where X_i's are iid $b(1, p)$. Thus, for large n we have

$$\frac{X - E(X)}{\sqrt{\text{var}(X)}} = \frac{X - np}{\sqrt{np(1-p)}} \approx N(0,1),$$

i.e. for large n,

$$X \approx N(np, np(1-p)).$$

As mentioned above, one problem of this approximation is that a continuous distribution is being used to approximate a discrete one. To get around this problem, what is called "continuity correction" has been used:

$$\Pr(X = k) = \Pr(k - 0.5 \le X \le k + 0.5), k = 0, 1, \ldots, n.$$

There are some restrictions on the use of CLT approximation. The two rules of thumb for the approximation to be accurate were introduced and investigated by **Schader & Schmid (1989),** these two rules are:

- $np(1-p) > 9$.
- $np > 5$ for $0 < p \le 0.5$ and $n(1-p) > 5$ for $0.5 < p < 1$.

The shape of the normal pdf is bell shaped. Thus, normal approximation of the binomial pf may not be suitable to use if the shape of the binomial distribution is very skewed; i.e. p is closed to zero or to one. However, it easier to use the standard normal distribution tables especially by beginner students than to use the exact formula of the binomial pf.

Another way to approximate the Binomial distribution was proposed by **Chang et al. (2008)** based on the skew normal distribution:

$$P(X \le k) = F_{n,p}(k) \approx \Psi_\lambda\left(\frac{k + 0.5 - \mu}{\sigma}\right),$$

where, $\Psi_{\lambda}(c)=\int_{-c}^{c}2\varphi(z)\Phi(\lambda z)dz$. The values of the parameters σ^2, μ and λ are chosen suitably based on n & p.

A discrete random variable X has a Poisson distribution if its pf is:

$$f(x;\lambda)=p(X=x)=\frac{\lambda^x e^{-\lambda}}{x!},\quad x=0,1,2,\dots\ ,\ \lambda>0.$$

$\mu=\sigma^2=\lambda$. If n is large and p is small, with $np=\lambda$ fixed, the terms of $b(n,p)$ are found to be near the Poisson distribution, $P(\lambda)$;

$$\binom{n}{x}p^x(1-p)^{n-x}\approx\frac{(np)^x e^{-np}}{x!}\quad \textbf{(See Filler (1968)).}$$

(For more details see Hogg and Tanis (1996)).

In this paper, the discrete normal is used instead of the normal. In section 2, the approximation of the binomial probability function, bpf, using discrete normal is investigated. Conclusions and some suggested future works are the content of section 3.

2. Approximation of bpf Using Discrete Normal pdf

In this section, we investigate the appropriateness of using the general discrete normal distribution to approximate the bpf. The approximate value is compared with the exact value and with normal approximation.

Let X be a random variable with discrete normal distribution, $DN(\mu,\sigma^2)$. The pdf of X is:

$$f(x;\mu,\sigma^2)=\frac{\frac{1}{\sigma\sqrt{2\pi}}e^{-\frac{(x-\mu)^2}{2\sigma^2}}}{\sum_{i=-\infty}^{\infty}\frac{1}{\sigma\sqrt{2\pi}}e^{-\frac{(i-\mu)^2}{2\sigma^2}}},\quad x=0,\pm1,\pm2,\dots;\quad \sigma>0,\ -\infty<\mu<\infty.$$

Clearly, the most inconvenient part of this pf is its denominator.

Zsablowski (2001) introduced the following general formula:

$$\text{For }\ 0<q<1,\ \sum_{k=-\infty}^{\infty}q^{k^2/2}=\sqrt{\frac{2\pi}{\ln(1/q)}}\left(1+2\sum_{k=1}^{\infty}\exp\left(-\frac{2\pi^2k^2}{\ln(1/q)}\right)\right).$$

In particular, for $q\ge\exp(-2\pi^2/(7\ln10-\ln5))=0.25653$, he showed that

$2\exp\left(\frac{-2\pi^2}{\ln(1/q)}\right)\le10^{-6}$. So, with high accuracy level (up to the fifth digit), we have

$$\sum_{k=-\infty}^{\infty}q^{k^2/2}=\sqrt{\frac{2\pi}{\ln(1/q)}}.$$

By taking $q=\exp(-1)\cong0.36788>0.25658$, he got with accuracy up to 10^{-6}:

$$\sum_{k=-\infty}^{\infty}\frac{\exp(-k^2/2)}{\sqrt{2\pi}}\approx1,\quad \&\quad \sum_{k=-\infty}^{\infty}k^2\,\frac{\exp(-k^2/2)}{\sqrt{2\pi}}\approx1.$$

Thus, based on these results, if X is $DN(\mu,\sigma^2)$, then for $\sigma^2>0.73$, we have:

$$f(x;\mu,\sigma^2)=\frac{\frac{1}{\sigma\sqrt{2\pi}}e^{-\frac{(x-\mu)^2}{2\sigma^2}}}{\sum_{i=-\infty}^{\infty}\frac{1}{\sigma\sqrt{2\pi}}e^{-\frac{(i-\mu)^2}{2\sigma^2}}},\approx\frac{1}{\sigma\sqrt{2\pi}}e^{\frac{-(x-\mu)^2}{2\sigma^2}},\quad x=0,\pm1,\pm2,\dots,$$

$$E(X) \approx \mu,\ Var(X) \approx \sigma^2 .$$

Let $Y \sim b(n,p)$, as we mentioned in the previous section, by CLT and under some conditions on n and p, the distribution of Y is approximately normal with mean np and variance $np(1-p)$. If $X \sim DN(\mu,\sigma^2)$, by using the methods of moment similar to the work of **Chang et al. (2008),** we have $E(Y)=np$, $Var(Y)=np(1-p)$ equating them to μ and σ^2, we obtained an approximate value of (μ,σ^2) as $\mu^*=np,\ \sigma^{*2}=np(1-p)$. Thus, we can use the $DN(np,np(1-p))$ to approximate the bpf. Thus, for $k=0,1,2,...,n$, we have

$$P(Y=k)=\binom{n}{k}p^k(1-p)^{n-k} \approx f(k;np,np(1-p))=\frac{1}{\sqrt{2\pi np(1-p)}}e^{\frac{-(k-np)^2}{2np(1-p)}}.$$

For different values for n & p, we computed the approximate value for bpf using discrete normal probability function at $X=k$, and compared it with the exact value. The maximum absolute error, $MABE(n,p)$, is used to measure the accuracy of the approximation:

$$MABE(n,p)=\max_{k\in\{0,1,2,...,n\}}\left|P(Y=k)-f(k;np,np(1-p))\right|.$$

The smaller the $MABE(n,p)$, the better the approximation. Also, this approximation is compared with the CLT approximation.

Tables (1-3) contain the numerical calculations for the exact value and the values of the discrete normal and normal approximations for selected values of (p,n); p =0.1, 0.3, 0.5, $n=$ 10, 30, 50. Since $MABE(n,p)$ is symmetric around $p=0.5$; i.e. $MABE(n,p)=MABE(n,1-p)$, for $DN(np,np(1-p))$ & $N(np,np(1-p))$, there is no need to consider the values of p that are larger than 0.5.

Table (1a). The exact and the approximate values of $p(X=k)$ when $n=10, p=0.1$

k	$B(n,p)$	$DN(np,np(1-p))$	$N(np,np(1-p))$
0	0.3487	0.2413	0.2422
1	0.3874	0.4205	0.4018
2	0.1937	0.2413	0.2422
3	0.0575	0.0456	0.0527
4	0.0112	0.0029	0.0041
≥ 5	0.0016	0.0001	0.0001
$MABE(n,p)$		**0.1074**	**0.1065**

Table (1b). The exact and the approximate values of $p(X=k)$ when $n=30, p=0.1$

k	$B(n,p)$	$DN(np,np(1-p))$	$N(np,np(1-p))$
0	0.0425	0.0458	0.0475
1	0.1413	0.1158	0.1166
2	0.2277	0.2018	0.1998
3	0.2361	0.2428	0.2391
4	0.1771	0.2018	0.1998
5	0.1023	0.1158	0.1166
6	0.0474	0.0459	0.0475
≥ 7	0.0258	0.0202	0.0166
$MABE(n,p)$		**2.5911×10^{-2}**	**2.7858×10^{-2}**

Table (1c). The exact and the approximate values of $p(X=k)$ when $n=50, p=0.1$

k	$B(n,p)$	$DN(np,np(1-p))$	$N(np,np(1-p))$
≤ 2	0.1117	0.1232	0.1145
3	0.1386	0.1206	0.1204
4	0.1809	0.1683	0.1671
5	0.1849	0.1881	0.1863
6	0.1541	0.1683	0.1671
7	0.1076	0.1206	0.1204
≥ 8	0.1222	0.1298	0.1193
$MABE(n,p)$		$\mathbf{1.7983\times10^{-2}}$	$\mathbf{1.8112\times10^{-2}}$

Table (2a). The exact and the approximate values of $p(X=k)$ when $n=10, p=0.3$

k	$B(n,p)$	$DN(np,np(1-p))$	$N(np,np(1-p))$
10	0.1398	0.1411	0.1403
11	0.1271	0.1325	0.1319
12	0.1033	0.1098	0.1096
13	0.0755	0.0804	0.0804
14	0.0499	0.0519	0.05216
≥ 15	0.0607	0.0549	0.0558
$MABE(n,p)$		$\mathbf{7.0740\times10^{-3}}$	$\mathbf{7.3598\times10^{-3}}$

Table (2b). The exact and the approximate values of $p(X=k)$ when $n=30, p=0.3$

k	$B(n,p)$	$DN(np,np(1-p))$	$N(np,np(1-p))$
≤ 6	0.1595	0.1579	0.1595
7	0.1218	0.1157	0.1154
8	0.1501	0.1468	0.1460
9	0.1573	0.1589	0.1579
10	0.1416	0.1468	0.1460
11	0.1103	0.1157	0.1154
12	0.0748	0.0778	0.0780
≥ 13	0.0845	0.0802	0.0802
$MABE(n,p)$		$\mathbf{6.1444\times10^{-3}}$	$\mathbf{6.4243\times10^{-3}}$

Table (2c). The exact and the approximate values of $p(X=k)$ when $n=50, p=0.3$

k	$B(n,p)$	$DN(np,np(1-p))$	$N(np,np(1-p))$
≤ 12	0.2229	0.2193	0.2202
13	0.1050	0.1018	0.1015
14	0.1190	0.1174	0.1170
15	0.1224	0.1231	0.1226
16	0.1147	0.1174	0.1170
17	0.0983	0.1018	0.1015
18	0.0772	0.0802	0.0801
≥ 19	0.1406	0.1391	0.1400
$MABE(n,p)$		$\mathbf{3.6270\times10^{-3}}$	$\mathbf{3.6730\times10^{-3}}$

Table (3a). The exact and the approximate values of $p(X = k)$ when $n = 10,\ p = 0.5$

k	$B(n, p)$	$DN(np, np(1-p))$	$N(np, np(1-p))$
≤ 2	0.0547	0.0537	0.0567
3	0.1172	0.1134	0.1145
4	0.2051	0.2066	0.2045
5	0.2461	0.2523	0.2482
6	0.2051	0.2066	0.2045
7	0.1172	0.1134	0.1145
≥ 8	0.0547	0.0537	0.0567
$MABE(n, p)$		$\mathbf{6.2195 \times 10^{-3}}$	$\mathbf{2.7198 \times 10^{-3}}$

Table (3b). The exact and the approximate values of $p(X = k)$ when $n = 30,\ p = 0.5$

k	$B(n, p)$	$DN(np, np(1-p))$	$N(np, np(1-p))$
≤ 11	0.1002	0.0994	0.1006
12	0.0805	0.0799	0.0800
13	0.1115	0.1116	0.1113
14	0.1354	0.1363	0.1356
15	0.1445	0.1457	0.1449
16	0.1354	0.1363	0.1356
17	0.1115	0.1116	0.1113
18	0.0806	0.0800	0.0800
≥ 19	0.1002	0.0994	0.1006
$MABE(n, p)$		$\mathbf{1.2087 \times 10^{-3}}$	$\mathbf{5.1919 \times 10^{-4}}$

Table (3c). The exact and the approximate values of $p(X = k)$ when $n = 50,\ p = 0.5$

k	$B(n, p)$	$DN(np, np(1-p))$	$N(np, np(1-p))$
≤ 22	0.2399	0.2390	0.2398
23	0.0960	0.1084	0.0959
24	0.1080	0.1128	0.1081
25	0.1123	0.1084	0.1125
26	0.1080	0.0962	0.1081
27	0.0960	0.0787	0.0959
≥ 28	0.2399	0.2390	0.3357
$MABE(n, p)$		$\mathbf{5.6274 \times 10^{-4}}$	$\mathbf{2.4528 \times 10^{-4}}$

Comparison Exact value and Discrete Normal Approximation:

Based on Table 1(a), ($n = 10, p = 0.1$), that *the exact*, $B(n, p)$, and approximate value, $DN(np, np(1-p))$, are close for almost all values of k. The maximum absolute error is 0.1074. The approximation is not so good for small values of k. Similar comments can be said based on Table 1(b), $(n = 30,\ p = 0.1)$, Table 1(c), $(n = 50,\ p = 0.1)$; $MABE(50, 0.1)$ is 0.02591 and $MABE(30, 0.1)$ is 0.01798. It can be seen that $MABE(n, p)$ is decreasing in n for fixed p.

Based on Table 2 (a,b,c), it can be seen that the $MABE(n, 0.2)$ is decreasing in n: $MABE(10, 0.3) = 0.01685$, $MABE(30, 0.3) = 0.00614$, $MABE(50, 0.3) = 0.00363$. Clearly the approximation is very accurate. Based on Table 3(a,b,c); *MABE*(10,0.5)=0.00622, *MAB(*30,0.5)=0.00121, *MAB(*50,0.5)=0.000562. Clearly the approximation is extremely accurate.

Comparison of Discrete Normal and Central Limit Theorem Approximations:

Several ways have been used to approximate the Binomial distribution. The most popular way is the one using the Normal distribution with $\mu = np$ and $\sigma^2 = np(1-p)$. This approximation is justified by the central limit theorem, one of the most important theorems in probability and statistics:

$$\frac{X - E(X)}{\sqrt{\mathrm{var}(X)}} = \frac{X - np}{\sqrt{np(1-p)}} \approx N(0,1),$$

The approximation of the binomial by normal probability function is suitable if $np > 5$, **(See Schader and Schmid (1989)).** Since the normal distribution is continuous and the binomial is discrete and takes only nonnegative integers, the correction factor for continuity is being used:

$$P(X = k) = P(k - 0.5 \le X \le k + 0.5).$$

For example, if X is $b(10,0.1)$, then the exact value of $P(X = 1)$ is

$$P(X = 1) = \binom{10}{1}(0.1)^1(0.9)^9 = 0.3874;$$

while the approximate value using CLT is

$$P(X = 1) = P(0.5 \le X \le 1.5) \cong \int_{0.5}^{1.5} \frac{1}{\sqrt{1.8\pi}} e^{\frac{-1}{1.8}(x-1)^2} = 0.40184.$$

Based on the above tables, we can say that except for very few cases of fluctuations, the approximation using the discrete normal distribution is satisfactory and recommended for use. Furthermore, using the discrete normal probability function is more reasonable and easier.

3. Conclusions and Some Suggested Future Works

In this paper, we considered the discrete normal pdf to approximate the binomial probability function. The choice of this pdf is motivated by CLT and the fact that the binomial random variable is a discrete one. Taking into account the accuracy and easiness as well, the discrete normal strongly recommended for use in practice.

It might of interest to go for some theoretical results. The approximate discrete probability function can be also used to make inference about *p*, when it is unknown. For example, when the distribution of X is approximated by $N(np, np(1-p))$. Based on this approximated distribution an approximate a ($1-\alpha$)% C.I. can be obtained for *p*. Similar things can be done using the discrete pdf.

Other methods of discretizing the normal distribution can be used. For example:

$$p(Y = k) = p(X \le k) - p(X \le k-1) = \Phi\left(\frac{k-\mu}{\sigma}\right) - \Phi\left(\frac{k-1-\mu}{\sigma}\right)$$

is an approach **(See Gomes-Deniz, Vazquez-Polo, Garcia-Garcia (2012)).** Another approach is by using the greatest integer of the continuous random variable; $Y = \left[X\right]$. These approaches can be used to obtain discrete normal distributions, which can be investigated for the approximation of the binomial function. Truncated discrete normal is another choice motivated by the binomial r.v. being nonnegative. Since the Poisson pf is the limit of the binomial pf as $n \to \infty, p \to \infty$ and $np = \lambda$, it can be used when p is very small.

References

Chang, C. H., Lin, J. J, Pal, N., & Chiang, M. C. (2008). A Note on Improved Approximation of the Binomial Distribution by the Skew-Normal Distribution. *The American Statistician, 62*(2), 167-170. https://doi.org/10.1198/000313008X305359

Filler, W. (1968). *An Introduction to Probability Theory and its Applications.* (3rd ed.) Wiley Eastern Limited.

Gomiz-Deniz, E., Vazquez-Polo, F., & Garcia-Garcia, V. (2012). A Discrete Version of the Half-Normal Distribution and its Generalization with applications. *Statistical Papers, 55,* 497-511. https://doi.org/10.1007/s00362-012-0494-6

Govindarajulu, Z. (1965). Normal Approximations to the Classical Discrete Distributions. *Sankhyā: The Indian Journal of Statistics, Series A, 27,* 143–72. http://www.jstor.org/stable/25049363

Hogg, R., & Tanis, E. (1996). Probability and Statistical Inference. (5th ed.). Prentice Hall.

Kemp, A. W. (1997). Characterizations of a Discrete Normal Distribution. *Journal of Statistical Planning and Inference, 63*, 223 -229. https://doi.org/10.1016/S0378-3758(97)00020-7

Schader, M., & Schmid, F. (1989). Tow Rules of Thumb for the Approximation of the Binomial Distribution by the Normal Distribution. *The American Statistician*, *43*(1), 23-24. https://doi.org/10.1080/00031305.1989.10475601

Seidu, I., & Kozubowski, T. J. (2004). A Discrete analogue of the Laplace distribution. *Journal of Statistical Planning and Inference, 136,* 1090 -1102. https://doi.org/10.1016/j.jspi.2004.08.014

Szablowski, P. J. (2001). Discrete Normal Distribution and its Relationship with Jacobi Theta Function. *Statistics & Probability Letters, 52,* 289-299. https://doi.org/10.1016/S0167-7152(00)00223-6

Copyrights