

The Normal and Poisson Approximations to the Binomial: A Closer Look

Robert Decker

University of Hartford

West Hartford, CT 06117

Donna Jean Fitzgibbon

Trinity College, University of Cambridge

Cambridge, CB2 1TQ

England

### Author's Footnote

Bob Decker is an assistant professor of mathematics at the University of Hartford, West Hartford, Ct 06117. Donna Jean Fitzgibbon is a graduate student in statistics at Trinity College, University of Cambridge, Cambridge CB2 1TQ, England.

## Abstract

The normal and Poisson approximations to the binomial distribution are presented in all upper level undergraduate texts on probability and statistics, and in almost all introductory text on statistics. Though rules of thumb for when to use each approximation are often stated, uniform error bounds are never given to accompany these rules. Rules for determining when the normal or Poisson approximation is better are likewise not stated. Using a computer analysis and linear regression, we have determined improved, though simple, rules of thumb which give error bounds on each approximation, and which determine which of the two approximations is preferable for a given calculation.

## Key Words

DeMoivre-Laplace theorem, central limit theorem, statistical computing

## 1. INTRODUCTION

We will analyze the standard normal and Poisson approximations to the binomial distribution. Let  $X$  be a binomial random variable with parameters  $n$  and  $p$ , and let  $Y_N$  be a random variable with values  $-1, 0, 1, 2, \dots, n, n+1$  for

which

$$P(Y_N = a) = \begin{cases} \int_{-\infty}^{-0.5} f(x) dx & \text{if } a = -1 \\ \int_{a-0.5}^{a+0.5} f(x) dx & \text{if } 0 \leq a \leq n \\ \int_{n+0.5}^{\infty} f(x) dx & \text{if } a = n+1 \end{cases} \quad (1)$$

where  $f(x)$  is the normal density with mean  $np$  and variance  $np(1-p)$ . If we approximate  $P(X=a)$  with  $P(Y_N=a)$  we are using the normal approximation to the binomial.  $Y_N$  is a discrete version of a normal random variable; we have defined it as

above at  $a=-1$  and  $a=n+1$  so that we can treat it as a finite distribution as well.

Now define a random variable  $Y_P$  for which

$$P(Y_P = a) = \begin{cases} e^{-np} (np)^a / a! & \text{if } 0 \leq a \leq n \\ \sum_{k=n+1}^{\infty} e^{-np} (np)^k / k! & \text{if } a = n+1 \end{cases} \quad (2)$$

If we approximate  $P(X=a)$  with  $P(Y_P=a)$  we are using the Poisson approximation to the binomial. Again, the purpose of the definition at  $a=n+1$  is to be able to treat the Poisson as a finite distribution.

The normal and Poisson approximations enable approximate calculation of binomial probabilities in cases where binomial tables aren't easily obtained (in most text books the binomial tables only go up to  $n=20$ ) and direct calculations are impractical. Of course, with the aid of a computer, almost any binomial probability can be calculated; however, even then the calculation can be time consuming (especially for very large  $n$ ) and a quick approximation with error bounds preferable.

The Poisson approximation is recommended when  $p$  is small and the normal approximation when  $p$  is not small. The question of where the boundary between the two occurs is generally avoided; one textbook author (Snell (1988)) goes so far as to say that simple rules to determine this boundary do not exist. Furthermore, practical error bounds are rarely given for either approximation. In textbooks, the normal approximation generally comes equipped with a rule of thumb. Three common rules used for  $p \leq 1/2$  (for  $p > 1/2$  replace  $p$  with  $1-p$ ) are

- 1) "the approximation is good if  $np \geq C_1$ "
- 2) "the approximation is good if  $np(1-p) \geq C_2$ "
- 3) "the approximation is good if  $np \geq C_3\sigma$ , where  $\sigma = \sqrt{np(1-p)}$ "

where  $C_1$  and  $C_2$  are constants ranging from about 3 to 10, and  $C_3$  is usually either 2 or 3. Rules of the form of 3) are based on the idea that  $P(Y < 0)$  should be small, if  $Y$  is to approximate a binomial random variable. Rules of the form 1) or 2) are used for their simplicity. The chosen rule is generally followed by an example, where  $n$  and  $p$  are conveniently chosen so as minimize the error for the given rule (Feller (1945) criticizes textbook examples of this approximation). The Poisson approximation most often comes with a more ambiguous guideline such as "the approximation is good when  $p$  is small,  $n$  is large, and  $np$  is moderate." Some rules of thumb, as stated in various standard textbooks, are summarized in Table 1.

When error bounds are given (see for example Golberg (1984)) they are usually too complicated to be of practical use. Smith (1953) has published a set of tables and graphs which can be used with the normal and Poisson approximations to find cumulative binomial probabilities to one decimal place; this work, however, likewise

fails a simplicity requirement. Also, approximations with error bounds are most often stated in terms of the cumulative distribution function, where as for many probability calculations error bounds given in terms of the mass function are more appropriate, as will be seen below. Schader and Schmid (1989) showed empirically that the maximum absolute distance between cumulative distribution functions (which they call  $MABS(n,p)$ ) for the normal approximation is a roughly linear function of  $p$  for rules of thumb of the form 1) or 2) given above; a good rule of thumb, however, should leave the error measure constant.

We have developed simple empirically-based rules of thumb which keep  $MABS(n,p)$  as well as two other measures of "closeness" for probability distributions (described below) relatively constant over large ranges of both  $n$  and  $p$  for both approximations. All rules are of the form  $n^k p \geq \text{constant}$  or  $n^k p \leq \text{constant}$ . For each rule and each measure of error, the error along the curve given by  $n^k p = \text{constant}$  is given. Keeping the rule simple enables the user of the approximation to solve for either  $n$  or  $p$  to determine the necessary sample size or the probability required to achieve the desired accuracy. Finally, we present a rule for determining

when the normal or Poisson is the better approximation.

Computer programs were developed to determine  $(n,p)$  pairs which kept the measures of error constant. Linear regression was then used on the pairs  $(\ln n, \ln p)$  to determine the rules of thumb.

## 2. THE THREE ERROR MEASURES

Suppose we are to approximate the random variable  $X$  with the random variable  $Y$ . In order to estimate the error in the calculation of  $P(X \in S)$  where  $S$  is some subset of the range of  $X$ , we need a global measure of the distance between the distribution of  $X$  and that of  $Y$ . For any two discrete random variables  $X$  and  $Y$ , we introduce three measures of distance (and hence error if one is approximating the other):

$$M_1 = \max_{-\infty < z < \infty} | F_X(z) - F_Y(z) | \quad (3)$$

$$M_2 = \sum_{z=-\infty}^{\infty} | P_X(z) - P_Y(z) | \quad (4)$$

$$M_3 = \max_{-\infty < z < \infty} | P_X(z) - P_Y(z) | \quad (5)$$

$F_X$  and  $F_Y$  are the respective cumulative distributions for  $X$  and  $Y$  and  $P_X$  and

$P_Y$  the respective mass functions. When writing  $M_k(n,p)$  for  $k=1, 2, 3$ , we are referring to either the normal or Poisson approximation to the binomial (i.e.  $X$  is binomial with parameters  $n,p$  and  $Y$  is normal or Poisson as defined in (1) or (2));  $M_1(n,p)$  is referred to as MABS( $n,p$ ) above.

For general discrete random variables  $X$  and  $Y$  where  $S$  is a subset of the combined ranges of  $X$  and  $Y$  we will show that

$$| P( X \in S ) - P( Y \in S ) | \leq 1/2 M_2 \quad (6)$$

and

$$| P( X \in S ) - P( Y \in S ) | \leq \#S M_3 \quad (7)$$

where  $\#S$  is the number of elements in  $S$ . It will also be shown that

$$| P( a \leq X \leq b ) - P( a \leq Y \leq b ) | \leq 2 M_1 \quad (8)$$

and

$$| P( X \leq b ) - P( Y \leq b ) | \leq M_1 \quad (9)$$

for any  $a$  and  $b$ . Thus, depending on the calculation, each of the three measures could be useful. It will be shown that

$$M_1 \leq 1/2 M_2 \quad (10)$$

although, for the distributions investigated, we have found empirically that

$$M_1(n,p) \approx 1/4 M_2(n,p) \quad (11)$$

over the range of  $n,p$  of interest for both the normal and Poisson approximations.

Thus (6) and (8) give roughly the same error estimate for intervals  $S = [a, b]$ . Also, each of the inequalities (6) through (10) is sharp; we can find random variables  $X$  and  $Y$  and a set  $S$  or interval  $[a, b]$  which makes each into an equality.

Let  $X$  and  $Y$  be finite, integer valued random variables. We now prove (6) - (10).

Proofs of (6) and (7): We first note that

$$\sum_{P_X > P_Y} P_X(z) + \sum_{P_X < P_Y} P_X(z) = 1 \quad (11)$$

and

$$\sum_{P_X > P_Y} P_Y(z) + \sum_{P_X < P_Y} P_Y(z) = 1. \quad (12)$$

Subtracting (11) from (12) and rearranging terms gives us

$$\sum_{P_X > P_Y} (P_X(z) - P_Y(z)) = \sum_{P_X < P_Y} (P_Y(z) - P_X(z)). \quad (13)$$

Now, since

$$M_2 = \sum_z |P_X(z) - P_Y(z)|$$

$$= \sum_{P_X > P_Y} (P_X(z) - P_Y(z)) + \sum_{P_X < P_Y} (P_Y(z) - P_X(z))$$

we get using (13) that

$$M_2 = 2 \sum_{P_X > P_Y} (P_X(z) - P_Y(z)) = 2 \sum_{P_X < P_Y} (P_Y(z) - P_X(z)). \quad (14)$$

Thus, if  $S$  is any set we get

$$|P(X \in S) - P(Y \in S)| = \left| \sum_{z \in S} (P_X(z) - P_Y(z)) \right|$$

$$= \left| \sum_{z \in S, P_X > P_Y} (P_X(z) - P_Y(z)) + \sum_{z \in S, P_X < P_Y} (P_X(z) - P_Y(z)) \right|$$

$$= \left| \sum_{z \in S, P_X > P_Y} |P_X(z) - P_Y(z)| - \sum_{z \in S, P_X < P_Y} |P_X(z) - P_Y(z)| \right|$$

$$\leq \max \left( \sum_{z \in S, P_X > P_Y} |P_X(z) - P_Y(z)|, \sum_{z \in S, P_X < P_Y} |P_X(z) - P_Y(z)| \right)$$

$$\leq 1/2 M_2 \quad \text{by virtue of (14) which proves (6).}$$

Finally, since

$$|P(X \in S) - P(Y \in S)| = \left| \sum_{z \in S} (P_X(z) - P_Y(z)) \right| \leq \sum_{z \in S} |P_X(z) - P_Y(z)|$$

$$\leq \sum_{z \in S} M_3 = \#S M_3 \quad \text{we have shown (7).}$$

Proofs of (8) and (9): We have

$$\begin{aligned} | P( a \leq X \leq b ) - P( a \leq Y \leq b ) | &= | F_X(b) - F_X(a-1) - F_Y(b) + F_Y(a-1) | \\ &\leq | F_X(b) - F_Y(b) | + | F_X(a-1) - F_Y(a-1) | \leq M_1 + M_1 = 2M_1 \end{aligned}$$

which proves (8). Similarly

$$| P(X \leq b) - P(Y \leq b) | = | F_X(b) - F_Y(b) | \leq M_1$$

which proves (9).

Proof of (10): Since

$$| F_X(z) - F_Y(z) | = | P( X \in (-\infty, z] ) - P( Y \in (-\infty, z] ) | \leq 1/2 M_2$$

by (6), we get (10) from the definition of  $M_1$ .

### 3. EMPIRICAL FINDINGS AND RULES OF THUMB

For  $p$  in the range  $0.0001 \leq p \leq 0.5$  and for  $n$  in the range  $20 \leq n \leq 1,000,000$  computer programs were developed and used to measure  $M_1$ ,  $M_2$ , and  $M_3$  for various combinations of  $n$  and  $p$ . Pairs  $(n, p)$  of constant error for each of the three measures and the two approximations were then determined for selected error values. Due to (11) only the data corresponding to  $M_2$  and  $M_3$  was used.

These pairs determined the rules of thumb using linear regression on  $(\ln n, \ln p)$ . It was found that the graph of  $\ln p$  vs  $\ln n$  was not sufficiently linear over the entire range of  $n$  and  $p$  values; thus as many as three piece-wise linear equations were needed to approximate the  $(\ln n, \ln p)$  curves of constant error. The only case where regression could not be used was in the Poisson approximation using  $M_2$ . For this case the  $\ln n$  vs  $\ln p$  curves are vertical, i.e., the  $M_2$  (and hence  $M_1$ ) error for the Poisson approximation depends only on  $p$ .

The regression results were simplified and adjusted so that the approximately constant error became an error bound. For example, the result of the regression using the normal approximation for points  $(n,p)$  which satisfy  $M_3(n,p) = 0.01$  and for  $0.0001 \leq p \leq 0.14$  was the curve  $n^{0.985} p = 8.59$  with a correlation of  $r = 0.9999$ ; this was simplified and adjusted to the rule  $np \geq 10$  to ensure that  $M_3 \leq 0.01$  when  $0.0001 \leq p \leq 0.14$ . The simplified rules are given in Table 2 (normal) and Table 3 (Poisson).

Finally, points  $(n,p)$  which gave equal error for the Poisson and normal approximations were determined, and regression used to find a curve which represents a

boundary between the two approximations for  $M_2$  and  $M_3$ . The simplified rule along with some data are presented in Table 4.

### 3.1 The Normal Approximation to the Binomial

Looking at Table 2 we see that the exponent of the simplified  $n^k p$  rule depends strongly on  $p$  but weakly on  $n$  for both the  $M_2$  and  $M_3$  errors. In the range  $0.0001 \leq p \leq 0.14$  we see that the form of the rule is  $np \geq \text{constant}$  in agreement with the simplest of the standard rules of thumb. However, we find that the value of  $np$  needed to achieve a reasonable degree of accuracy is generally much higher than one might infer from the common rules; for 0.05  $M_2$  error we need  $np \geq 26$  and for 0.01  $M_2$  error we need  $np \geq 633$ . Recall that due to (6) these provide bounds of 2.5% and 0.5% for the calculation of any probability. In the middle  $p$  ranges ( $0.14 \leq p \leq 0.34$  for  $M_2$  and  $0.14 \leq p \leq 0.39$  for  $M_3$ ) the  $np$  rule should be replaced by an  $n^{0.4} p$  ( $M_2$  error) or  $n^{0.6} p$  rule ( $M_3$  error). For the highest  $p$  values one should look at  $n^{0.1} p$  ( $M_2$  error) or  $n^{0.15} p$  ( $M_3$  error). Most textbooks give examples of the normal approximation to the binomial for  $p$  near 0.5, precisely where the  $np$  rule is least

applicable.

### 3.2 The Poisson Approximation to the Binomial

The  $M_2$  and  $M_3$  results (Table 3) for the Poisson approximation are less consistent with each other than for the normal approximation. For  $M_2$ , the errors are roughly constant for constant  $p$ ; thus we get an  $n^0 p$  rule. We note that the approximation

$$M_2(n,p) \approx 1/2 p \text{ for } np \geq 1 \quad (15)$$

which was determined empirically is quite good. For  $M_3$  we find that the  $\ln p$  vs  $\ln n$  graphs of constant error are roughly linear both for small  $n$  and for large  $n$  but that the slopes in these two regions are negatives of each other. For each error level we get an  $n^{k_1} p$  rule for small  $n$  and an  $n^{-k_2} p$  rule for large  $n$  (where  $k_1$  and  $k_2$  are positive). Thus the  $M_2$  results are somewhat consistent with textbook rules of thumb (which emphasize the dependence on  $p$ ), whereas, if we are using the  $M_3$  bound, we can use increasingly larger  $p$ 's as  $n$  gets large and maintain the error level.

### 3.3 Comparison of the Normal and Poisson Approximations

Due to the qualitative difference between  $M_2$  and  $M_3$  results for the Poisson case, we might expect that the curves of equal error for the normal and Poisson approximations would be different for the two error measures. Surprisingly, this is not the case; for both measures the curves of equal error are very close to straight lines (in the  $\ln p$  vs  $\ln n$  plane) and these lines are nearly identical for both measures. For  $M_2$  the curve of equal error using regression was  $pn^{0.3018} = 0.4476$ ; for  $M_3$  the curve of equal error was  $pn^{0.3159} = 0.4862$ . Thus, we combine these to get a robust and simple measure dictating when to use the normal approximation as opposed to the Poisson approximation: if  $pn^{0.31} \geq 0.47$ , use the normal approximation; otherwise, use the Poisson approximation.

Included in Table 4 are a comparison the values of  $M_2$  and  $M_3$  for the normal and Poisson approximations along the curve  $pn^{0.31} = 0.47$  so that the reader can judge how well the rule works.

#### 4. SOME TYPICAL CALCULATIONS

We conclude with an application of the rules of thumb to some probability calculations, showing how each of the three measures  $M_1$ ,  $M_2$  and  $M_3$  can come into play.

In  $n = 10,000$  trials 3 dice are rolled; triple 1's come up 30 times. If the dice are fair, triple 1's should come up with probability  $p = 1/216$ . Thus the expected number in 10,000 trials is about  $np \approx 46$ . We let  $X = \#$  of triple ones, and calculate the binomial probability  $P(30 \leq X \leq 62)$  to see if 30 is reasonable (roughly a two tailed hypothesis test). Since  $n^{0.31}p = 0.08$  we use the Poisson approximation (Table 4). Since  $p = 0.000372$ , from Table 3 we get  $M_2 \leq 0.005$  and hence from (11)  $M_1 \leq 0.00125$ . Since  $n^{-0.9}p = 1.2 \times 10^{-6}$ , and  $n^{-0.85}p = 1.8 \times 10^{-6}$  we have  $M_3 \leq 0.0005$  (again from Table 3). Thus, letting the actual error =  $E_1$ , we get

$$E_1 \leq 0.0025 \quad \text{by (6) and}$$

$$E_1 \leq 33 (0.0005) = 0.017 \quad \text{by (7)}$$

and so in this case,  $M_2$  provides the better bound. If we wanted to calculate  $P(X = 30)$ , then letting the actual error =  $E_2$  we get

$$E_2 \leq 0.0025 \quad \text{by (6) and}$$

$$E_2 \leq 0.0005 \quad \text{by (7).}$$

In this case,  $M_3$  provides the better bound. Finally, if we wanted to find  $P( X \leq 30 )$

(a one tailed hypothesis test) with the error given by  $E_3$  we get

$$E_3 \leq 0.0025 \quad \text{by (6) and}$$

$$E_3 \leq 31 (0.0005) = 0.016 \quad \text{by (7) and}$$

$$E_3 \leq M_1 \approx 1/4 M_2 = 0.00125 \quad \text{by (9) and (11)}$$

and so for this case,  $M_1$  provides the best bound. To summarize: for large intervals

or general sets use  $M_2$  ; for small intervals or sets use  $M_3$  ; for tails use  $M_1$ .

Table 1. Rules of Thumb for the Normal and Poisson Approximations  
from Selected Books

<u>Normal Rules</u> (for $p \leq 1/2$ )	<u>Author</u>
$np \geq 5$	Hogg and Tanis (1988); Wolf (1962); Lindgren (1969)
$np(1-p) \geq 10$	Parzen (1960); Ross (1984)
$np(1-p) \geq 5$	Solomon (1987)
$np(1-p) \geq 3$	Golberg (1984)
$np \geq 3\sqrt{npq}$	Drake (1967); Mosteller, Rourke, Thomas(1970)
$np \geq 2\sqrt{npq}$	Mendenhall (1967)
<u>Poisson Rules</u>	<u>Author</u>
n large, p small, np moderate	Ross (1984); Solomon (1987); Alexander (1961)
n large, p small	Golberg (1984); Drake (1967)
$p \leq 0.1$	Parzen (1960)
$n \geq 20, p \leq 0.05$ (fairly succesful)	Hogg and Tanis (1988)
$n \geq 100, np \leq 10$ (very succesful)	

Table 2. Summary of Rules of Thumb and Errors for the Normal Approximation.  
Valid for  $0.0001 \leq p \leq 0.5$  and for  $20 \leq n \leq 1,000,000$

<u><math>M_2</math> Error Bound</u>	<u>Rule of Thumb</u>		
	<u><math>0.0001 &lt; p &lt; 0.14</math></u>	<u><math>0.14 &lt; p &lt; 0.34</math></u>	<u><math>0.34 &lt; p &lt; 0.5</math></u>
$M_2 \leq 0.10$	$np \geq 7$	$n^{0.4} p \geq 0.6$	
$M_2 \leq 0.05$	$np \geq 26$	$n^{0.4} p \geq 1.05$	
$M_2 \leq 0.01$	$np \geq 633$	$n^{0.4} p \geq 4.0$	$n^{0.1} p \geq 0.65$
$M_2 \leq 0.005$	$np \geq 2533$	$n^{0.4} p \geq 7.0$	$n^{0.1} p \geq 0.75$

  

<u><math>M_3</math> Error Bound</u>	<u>Rule of Thumb</u>		
	<u><math>0.0001 &lt; p &lt; 0.14</math></u>	<u><math>0.14 &lt; p &lt; 0.39</math></u>	<u><math>0.39 &lt; p &lt; 0.5</math></u>
$M_3 \leq 0.01$	$np \geq 10$	$n^{0.6} p \geq 1.85$	
$M_3 \leq 0.005$	$np \geq 20$	$n^{0.6} p \geq 2.70$	
$M_3 \leq 0.001$	$np \geq 95$	$n^{0.6} p \geq 6.95$	$n^{0.15} p \geq .80$
$M_3 \leq 0.0005$	$np \geq 188$	$n^{0.6} p \geq 11.00$	$n^{0.15} p \geq .88$
$M_3 \leq 0.0001$	$np \geq 930$	$n^{0.6} p \geq 28.00$	$n^{0.15} p \geq 1.11$

Table 3. Summary of Rules of Thumb and Errors for the Poisson Approximation.  
Valid for  $0.0001 \leq p \leq 0.5$  and for  $20 \leq n \leq 1,000,000$

<u><math>M_2</math> Error Bound</u>	<u>Rule of Thumb</u>
$M_2 \leq 0.10$	$p \leq 0.183$
$M_2 \leq 0.05$	$p \leq 0.095$
$M_2 \leq 0.01$	$p \leq 0.0175$
$M_2 \leq 0.005$	$p \leq 0.009$
<u><math>M_3</math> Error Bound</u>	<u>Rule of Thumb</u>
$M_3 \leq 0.01$	$n^{-0.70} p \leq 6.2 \times 10^{-3}$
$M_3 \leq 0.005$	$n^{-0.75} p \leq 1.5 \times 10^{-3}$
$M_3 \leq 0.001$	$n^{0.4} p \leq 0.026$ for $n < 100$ ; $n^{-0.85} p \leq 4.7 \times 10^{-5}$ for $n \geq 100$
$M_3 \leq 0.0005$	$n^{0.4} p \leq 0.017$ for $n < 285$ ; $n^{-0.85} p \leq 1.5 \times 10^{-5}$ for $n \geq 285$
$M_3 \leq 0.0001$	$n^{0.4} p \leq 0.0065$ for $n < 992$ ; $n^{-0.90} p \leq 5.3 \times 10^{-7}$ for $n \geq 992$

Table 4. Simplified Rule for Normal vs Poisson Approximation

Simplified Rule of Thumb

Use the normal approximation if  $n^{0.31}p \geq 0.47$ , otherwise use the Poisson approximation.

Values of  $M_2$  for the Normal and Poisson Approximations  
along the Curve  $n^{0.31}p = 0.47$

<u>n</u>	<u>p</u>	<u>Normal</u>	<u>Poisson</u>
51	0.139	0.0724	0.0718
201	0.0907	0.0510	0.0462
794	0.0593	0.0333	0.0296
3128	0.0388	0.0215	0.0191
12323	0.0254	0.0137	0.0124
48545	0.0166	0.00865	0.00809
191227	0.0108	0.00544	0.00527
753274	0.00708	0.00341	0.00344

Values of  $M_3$  for the Normal and Poisson Approximations  
along the Curve  $n^{0.31}p = 0.47$

<u>n</u>	<u>p</u>	<u>Normal</u>	<u>Poisson</u>
51	0.139	0.0121	0.0114
201	0.0907	0.00485	0.00447
794	0.0593	0.00191	0.00180
3128	0.0388	0.000746	0.000721
12323	0.0254	0.000291	0.000291
48545	0.0166	0.000113	0.000118
191227	0.0108	0.0000441	0.0000479
753274	0.00708	0.0000172	0.0000195

## REFERENCES

- Alexander, H. W. (1961), Elements of Mathematical Statistics, New York: John Wiley and Sons, Inc., p. 218.
- Drake, A. W. (1967), Fundamentals of Applied Probability Theory, New York: McGraw-Hill Book Co., p. 219.
- Feller, W. (1945), "On the Normal Approximation to the Binomial Distribution," Annals of Mathematical Statistics, 16, pp. 319-329.
- Goldberg, M. A. (1984), An Introduction to Probability Theory with Statistical Applications, New York: Plenum Press, pp. 219, 377.
- Hogg, R. V., Tanis, E. A. (1988), Probability and Statistical Inference (3rd ed.), New York: Macmillan Publishing Co., pp. 315, 322.
- Lindgren, B. W., McElrath, G. W. (1969), Introduction to Probability and Statistics (3rd ed.), New York: Macmillan Publishing Co., p. 154.
- Mendenhall, W. (1967), Introduction to Probability and Statistics (2nd ed.), Belmont, CA.: Wadsworth Publishing Co., Inc., p. 132.
- Mosteller, F., Rourke, R. E. K., and Thomas Jr., G. B. (1970), Probability with Statistical Applications (2nd ed.), Reading, MA.: Addison-Wesley Publishing Co., p. 290.

- Parzen, E. (1960), Modern Probability Theory and Its Applications, New York: John Wiley and Sons, Inc., p. 245.
- Ross, S. (1984), A First Course in Probability (2nd ed.), New York: Macmillan Publishing Co., pp. 162, 163.
- Schader, M., Schmid, F. (1989), "Two Rules of Thumb for the Approximation of the Binomial Distribution by the Normal Distribution," The American Statistician, 43, 23-24.
- Smith, E. S. (1953), Binomial, Normal and Poisson Probabilities, Bel Air, MD.: Author, pp. 51-66.
- Snell, J. L. (1988), Introduction to Probability, New York: Random House, Inc., p. 334.
- Solomon, F. (1987), Probability and Stochastic Processes, Englewood Cliffs, N.J.: Prentice-Hall, Inc., pp. 157, 302.
- Wolf, F. L. (1962), Elements of Probability and Statistics, New York: McGraw-Hill Book Co., Inc., p. 213.