Available online at www.sciencedirect.com
ScienceDirect
Journal of Statistical Planning and Inference 137 (2007) 3570–3590
journal of statistical planning and inference
www.elsevier.com/locate/jspi

Normal approximation to the hypergeometric distribution in nonstandard cases and a sub-Gaussian Berry–Esseen theorem☆

S.N. Lahiri*, A. Chatterjee, T. Maiti

Department of Statistics, Iowa State University, Ames, IA 50011, USA

Available online 21 April 2007

Abstract

In this paper, we consider simple random sampling without replacement from a dichotomous finite population. We investigate accuracy of the Normal approximation to the Hypergeometric probabilities for a wide range of parameter values, including the nonstandard cases where the sampling fraction tends to one and where the proportion of the objects of interest in the population tends to the boundary values, zero and one. We establish a non-uniform Berry–Esseen theorem for the Hypergeometric distribution which shows that in the nonstandard cases, the rate of Normal approximation to the Hypergeometric distribution can be considerably slower than the rate of Normal approximation to the Binomial distribution. We also report results from a moderately large numerical study and provide some guidelines for using the Normal approximation to the Hypergeometric distribution in finite samples.

MSC: primary 60F05; secondary 60G10; 62E20; 62D05

Keywords: Finite population; Sampling without replacement; Simple random sampling; Normal approximation; Berry–Esseen theorem

1. Introduction

Consider a dichotomous finite population of size N having M objects of 'type A' and $N - M$ objects of 'type B'. Suppose a sample of size n is drawn at random, without replacement from this population. Let X denote the number of 'type A'-individuals in the sample. Then, X is said to have the Hypergeometric distribution with parameters n, M, N, written as $X \sim \mathrm{Hyp}(n; M, N)$. The probability mass function (p.m.f) of X is given by

$$P(X = x) \equiv P(x; n, M, N) = \begin{cases} \dfrac{\binom{M}{x}\binom{N-M}{n-x}}{\binom{N}{n}} & \text{if } x = 0, 1, \ldots, n, \\ 0 & \text{otherwise,} \end{cases} \tag{1.1}$$

☆ Research partially supported by NSF Grant no. DMS 0306574.
* Corresponding author. Tel.: +1 515 294 2271; fax: +1 515 294 4040.
E-mail address: snlahiri@iastate.edu (S.N. Lahiri).

0378-3758/$ - see front matter
doi:10.1016/j.jspi.2007.03.033

Table 1
Error of Normal approximation

p	$f=0.5$	$f=0.6$	$f=0.7$	$f=0.8$	$f=0.9$
0.5	0.0562	0.0574	0.0613	0.0701	0.0929
0.6	0.0574	0.0593	0.0641	0.0743	0.0995
0.7	0.0613	0.0641	0.0702	0.0822	0.1112
0.8	0.0701	0.0743	0.0822	0.0972	0.1321
0.9	0.0929	0.0995	0.1112	0.1321	0.1787

Values of the absolute difference $|P((X-E(X))/Var(X)\leqslant x)-\Phi(x)|$ at $x=0$ for different values of the parameters p and f, where $X \sim \text{Hyp}(n;\, M,\, N)$ and $\Phi(\cdot)$ is the cdf of the N(0, 1) distribution. Here, $N=200$ and $M=Np$ and $n=Nf$.

where, for any two integers $r \geqslant 1$ and s,

$$\binom{r}{s} = \begin{cases} \dfrac{r!}{s!(r-s)!} & \text{if } 0 \leqslant s \leqslant r, \\ 0 & \text{otherwise,} \end{cases} \tag{1.2}$$

with $0!=1$ and $r! = 1\cdot 2 \ldots r$. Let $f=n/N$ denote the sampling fraction and let $p=M/N$ denote the proportion of the ‘type A’-objects in the population. The Hypergeometric distribution plays an important role in many areas of statistics, including sample surveys (Burstein, 1975; Wendell and Schmee, 1996), capture–recapture methods (Seber, 1970; Wittes, 1972), analysis of contingency tables (Blyth and Staudte, 1997), statistical quality control (von Collani, 1986; Patel and Samaranayake, 1991; Sohn, 1997), etc. Normal approximations to the Hypergeometric probabilities $P(.;\, n,\, M,\, N)$ of (1.1) are classical in the cases where the sampling fraction f and the proportion p are bounded away from 0 and 1; for example, see Feller (1971). The nonstandard cases correspond to the extremes where f or p take values near the boundary values 0 and 1. Although the nonstandard cases arise frequently in all these areas of applications, the validity and accuracy of the normal approximation in such situations are not well studied. The quality of Normal approximation deteriorates as the parameters f and p tend to their boundary values. For example, consider Table 1 where the cumulative distribution function (cdf) of the centered and scaled version $(X-E(X))/\sqrt{Var(X)}$ of X at zero is approximated by the Normal cdf at zero for different values of f and p.

Table 1 gives the values of the absolute difference of the cdfs of $(X-E(X))/\sqrt{Var(X)}$ and the standard Normal distribution at $x=0$. The population size is fixed at $N=200$ while the proportion p of ‘type A’-objects and the sampling fraction f are varied over a range of 0.5–0.9 in increments of 0.1 (the values of p and f between 0 and 0.5 are omitted due to the symmetry of the problem). In particular, for these choices of N and f, the sample size $n=Nf$ takes the values 100, 120, 140, 160 and 180. Note that even for such moderately large sample sizes, the error of approximation increases steadily as p approaches the boundary value 1. Indeed, for $p=0.9$ and $f=0.9$, the error of approximation is as high as 0.179 at the origin in the finite population sampling framework, which is significantly higher than 0.036, the error of Normal approximation to the Binomial cdf with parameters $n=180$ and $p=0.9$. This shows that the commonly known approximation results for the ‘with replacement sampling’ (or sampling from an infinite population) case do not give a representative picture in the finite population setting when the parameters p and f are close to their boundary values. For a better understanding, one needs to be able to quantify the accuracy of Normal approximation as a function of N, p and f in the finite population setting.

In this paper, we derive a non-uniform Berry–Esseen theorem on Normal approximation to the Hypergeometric distribution for a wide range of values of p and f, allowing these parameters go to the extreme points 0, 1. The non-uniform bound in the Berry–Esseen theorem shows that the difference of the cdfs of $(X-E(X))/\sqrt{Var(X)}$ and of a standard Normal variate at a point x is bounded above by

$$g(x)[Nf(1-f)p(1-p)]^{-1/2}, \tag{1.3}$$

where the function $g(x)=g(x;\, p,\, f)$ is bounded and decays at an exponential (sub-Gaussian) rate as a function of x. As a corollary, we also derive an exponential (sub-Gaussian) probability inequality for the tails of X, which may be of independent interest. Both the sub-Gaussian Berry–Esseen theorem and the exponential inequality seem to be new even in the standard case.

Note that the Hypergeometric variable X can be expressed as the sum of a collection of n *dependent* Bernoulli random variables and (1.3) yields the *uniform* bound $\mathrm{O}([Np(1-p)f(1-f)]^{-1/2})$ on the error of Normal approximation to the distribution of X. This rate is equivalent to the standard Berry–Esseen rate $\mathrm{O}(n^{-1/2})$ for the Binomial distribution only when p is bounded away from 0 and 1 and f bounded away from 1. However, for p and f close to these boundary points, the rate of approximation can be substantially slower. In such situations, the dependence of the Bernoulli random variables associated with X has a nontrivial effect on the accuracy of the Normal approximation. This provides the theoretical justification for the observed difference in the accuracy of Normal approximations to the Binomial probabilities and to the Hypergeometric probabilities in the nonstandard cases.

The theoretical findings of the paper and the numerical example in Table 1 also show that the existing guidelines for applying the Normal approximation to the Binomial distribution (based on *independent* Bernoulli random variables) are *not* appropriate for the Hypergeometric distribution in the nonstandard cases. To formulate a working guideline in such situations, we conduct a moderately large numerical study and investigate the effect of the dependence in finite samples. On the basis of the numerical study, in Section 3, we provide some ‘quick and easy’ guidelines for assessing the error of Normal approximation to the Hypergeometric distribution in practice.

The rest of the paper is organized as follows. We conclude Section 1 with a brief literature review. Section 2 introduces the asymptotic framework and states the Berry–Esseen theorem and the exponential inequality. Results from the numerical study are reported in Section 3. Proofs of all the results are given in Section 4.

For results on Normal approximations to Hypergeometric probabilities in the standard case where the sampling fraction f and the proportion p are bounded away from 0 and 1, see Feller (1971). For general p and f, Nicholson (1956) derived some very precise bounds for the point probabilities $P(.; n, M, N)$ (cf. (1.1)) using some special normalizations of the Hypergeometric random variable X. General methods for proving the central limit theorem (CLT) for sample means under sampling without replacement from finite populations are given by Madow (1948), Erdos and Renyi (1959), and Hajek (1960). In relation to the earlier work, the main contribution of our paper is to establish a non-uniform Berry–Esseen theorem for the Hypergeometric distribution for a wide range of parameter values, including the nonstandard case, and to provide some practical guidelines for using the Normal approximation in finite sample applications.

2. Main results

Let r be a positive integer valued variable and for each $r \in \mathbf{N}$ (where $\mathbf{N} = \{1, 2, \ldots\}$), let X_r be a random variable having the Hypergeometric distribution of (1.1) with parameters (n_r, M_r, N_r), where $n_r, M_r, N_r \in \mathbf{N}$. Thus we consider a sequence of dichotomous finite populations indexed by r, with the population of objects of type A and the sampling fraction, respectively, given by

$$p_r = \frac{M_r}{N_r} \quad \text{and} \quad f_r = \frac{n_r}{N_r} \quad \text{for all } r \in \mathbf{N}. \tag{2.1}$$

Let

$$\sigma_r^2 \equiv N_r p_r q_r f_r (1 - f_r), \tag{2.2}$$

where $q_r = 1 - p_r$. Also, let $\phi(\cdot)$ and $\Phi(\cdot)$, respectively, denote the density and the distribution function of a standard Normal random variable, i.e., $\phi(x) = (1/\sqrt{2\pi})\exp(-x^2/2)$, $x \in \mathbf{R}$ and $\Phi(x) = \int_{-\infty}^{x} \phi(t)\,\mathrm{d}t$, $x \in \mathbf{R}$. Let $I(\cdot)$ denote the indicator function. For $x, y \in \mathbf{R}$, write $x \wedge y = \min\{x, y\}$. Define

$$\delta_r = \tfrac{1}{10}(\max(a_{1r}, 2))^{-1}, \quad r \geqslant 1, \tag{2.3}$$

where $a_{1r} = (\bar{f}_r + 4)/4(1 - \bar{f}_r)$ and where $\bar{f}_r = \min\{f_r, 1 - f_r\}$. Then, we have the following result.

Theorem 1. *Suppose that* $X_r \sim \mathrm{Hyp}(n_r, M_r, N_r)$, $r \in \mathbf{N}$. *Assume that r is such that*:

$$\delta_r \sigma_r > 1. \tag{2.4}$$

Then there exists universal constants $C_1, C_2 \in (0, \infty)$ (*not depending on* r, n_r, M_r *and* N_r) *such that*:

$$\left| P\left(\frac{X_r - n_r p_r}{\sigma_r} \leqslant x\right) - \Phi(x) \right| \leqslant \frac{C_1}{\sigma_r} \frac{1+|x|^2}{\lambda_r(x)} \exp(-C_2 x^2 \lambda_r^2(x)) \tag{2.5}$$

for all $x \in \mathbf{R}$, *where* $\lambda_r(x) = q_r I(x \leqslant 0) + p_r I(x \geqslant 0)$.

Theorem 1 is a non-uniform Berry–Esseen theorem for the Hypergeometric distribution. It shows that the error of Normal approximation to the Hypergeometric distribution dies at a sub-Gaussian rate in the tails. The only condition needed for the validity of this bound is (2.4). It is easy to check that

$$\delta_r \in \left(\tfrac{1}{25}, \tfrac{1}{20}\right] \tag{2.6}$$

for all r satisfying (2.4). Hence, the bound in (2.5) is available for all r such that $\sigma_r \geqslant 25$.

As pointed out in Section 1, when both the sequences $\{p_r\}_{\{r\geqslant 1\}}$ and $\{f_r\}_{\{r\geqslant 1\}}$ are bounded away from 0 and 1, the rate of approximation in Theorem 1 matches the standard rate $O(n_r^{-1/2})$ of Normal approximation for the sum of n_r iid random variables with a finite third moment. Although the Hypergeometric random variable X_r can be written as a sum of n_r *dependent* Bernoulli (p_r) variables, the lack of independence of the summands does not affect the rate of Normal approximation as long as the sequence $\{p_r\}_{r\geqslant 1}$ is bounded away from 0 and 1 and $\{f_r\}_{r\geqslant 1}$ is bounded away from 1. On the other hand, if either of the sequences $\{p_r\}_{\{r\geqslant 1\}}$ and $\{f_r\}_{\{r\geqslant 1\}}$ converge to one of the extreme values 0 and 1, then

$$\sigma_r = o(n_r^{1/2}) \quad \text{as } r \to \infty$$

and the rate of Normal approximation to the Hypergeometric distribution is indeed *worse* than the standard rate $O(n_r^{-1/2})$ in such *nonstandard cases*.

An immediate consequence of Theorem 1 is the following exponential (sub-Gaussian) probability bound on the tails of X_r.

Theorem 2. *Suppose that* $X_r \sim \mathrm{Hyp}(n_r, M_r, N_r)$, $r \in \mathbf{N}$. *Then, there exist universal constants* $C_3, C_4 \in (0, \infty)$ (*not depending on* r, n_r, M_r, N_r) *such that for all r satisfying* (2.4),

$$P\left(\left|\frac{X_r - n_r p_r}{\sigma_r}\right| \geqslant x\right) \leqslant \frac{C_3}{(p_r \wedge q_r)^3} \exp(-C_4 x^2 [p_r \wedge q_r]^2) \quad \textit{for all } x > 0.$$

3. Numerical results

To gain some insight into the quality of Normal approximation to the Hypergeometric distribution in finite samples and to compare it with the accuracy in the case of the Binomial distribution, first we consider some joint plots of the cdfs of normalized Hypergeometric and Binomial random variables against the standard Normal cdf. Figs. 1–5 show these plots for different values of the parameters n and p for $N = 60, 200$.

From the figures, it follows that the quality of Normal approximation to the Hypergeometric distribution is comparable to that for the Binomial distribution for values of f and p close to 0.5, but there is a stark loss of accuracy for high values of f and p.

Next, to get a quantitative picture of the error of Normal approximation, we conducted a moderately large numerical study with different values of the population size N and with different values of the parameters p and f. The population sizes considered were $N = 60, 200, 500, 2000$. For a given value of N, the set of values of p and f considered was $\{0.5, 0.6, 0.7, 0.8, 0.9\}$. We considered the Kolmogorov distance, i.e., the maximal distance between the cdfs of the normalized Hypergeometric variable and a standard Normal variable as a measure of accuracy. More specifically, the measure of accuracy for the Hypergeometric case is defined as

$$\Delta(N, p, f) = \left| P\left(\frac{X - np}{\sigma} \leqslant x\right) - \Phi(x) \right|, \tag{3.1}$$

where $X \sim \mathrm{Hyp}(n, M, N)$, $f = n/N$, $p = M/N$, $\sigma^2 = Nf(1-f)p(1-p)$ and $\Phi(\cdot)$ denotes the cdf of the N(0,1) distribution.

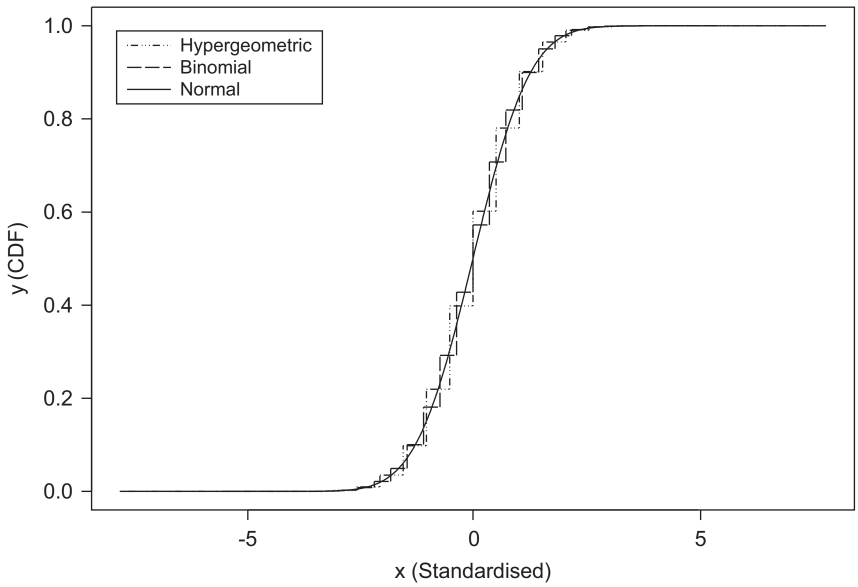

Fig. 1. A plot of the cdfs of normalized Hypergeometric and Binomial random variables against the standard Normal cdf for the parameter values $N = 60$, $p = 0.5$ and $f = 0.5$.

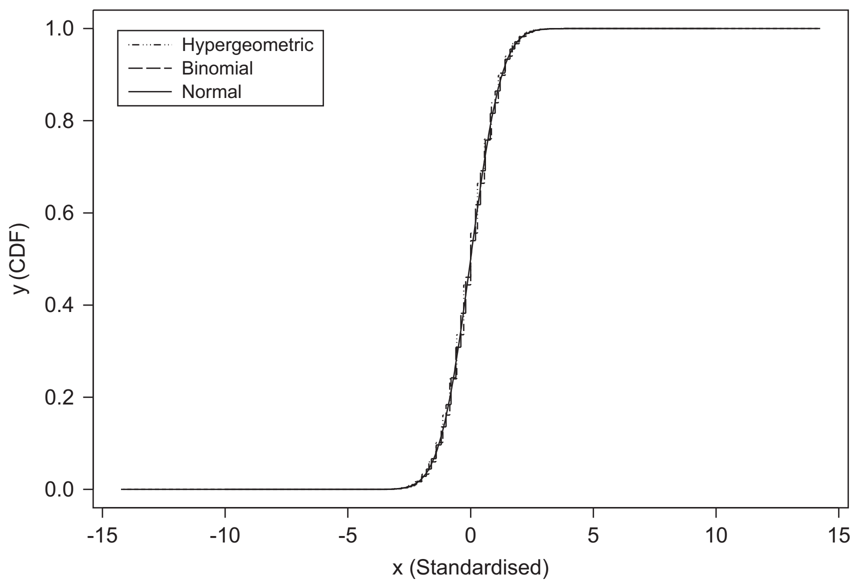

Fig. 2. A plot of the cdfs of normalized Hypergeometric and Binomial random variables against the standard Normal cdf for the parameter values $N = 200$, $p = 0.5$ and $f = 0.5$.

Tables 2–5 give the values of $\Delta(N, p, f)$ for different combinations of the parameter values as indicated above. For comparison, we also included the values of the maximal distance of the cdfs of normalized Binomial (n, p) and N(0,1) random variables.

From Tables 2–5, it follows that the accuracy of the Normal approximation to the Hypergeometric distribution deteriorates as p or f tend to the boundary value 1. Unlike the case of Normal approximation to the Binomial distribution, the values of the sample size n and p alone are not a good indicator of the level of accuracy attainable in this case. For example, for an iid sample of size $n = 54$ with $p = 0.8$, one may expect a reasonable accuracy of the Normal

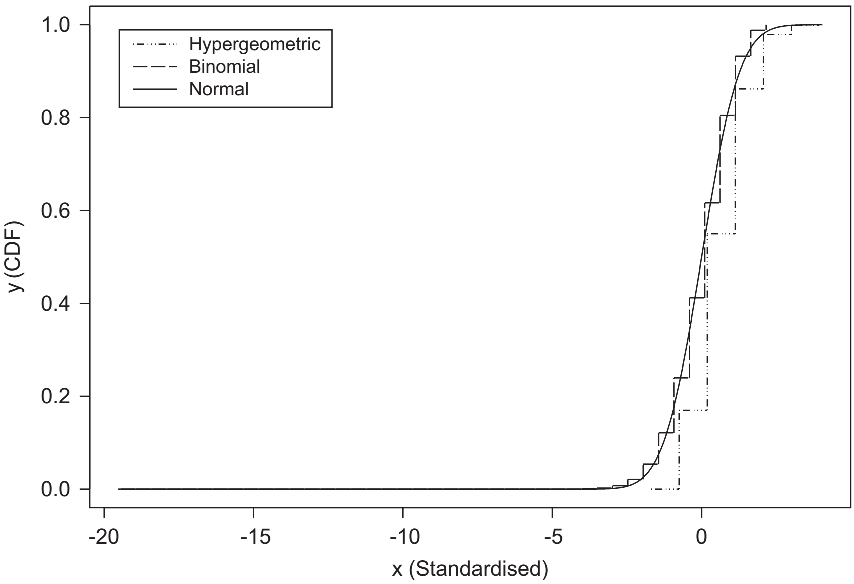

Fig. 3. A plot of the cdfs of normalized Hypergeometric and Binomial random variables against the standard Normal cdf for the parameter values $N = 60$, $p = 0.9$ and $f = 0.7$.

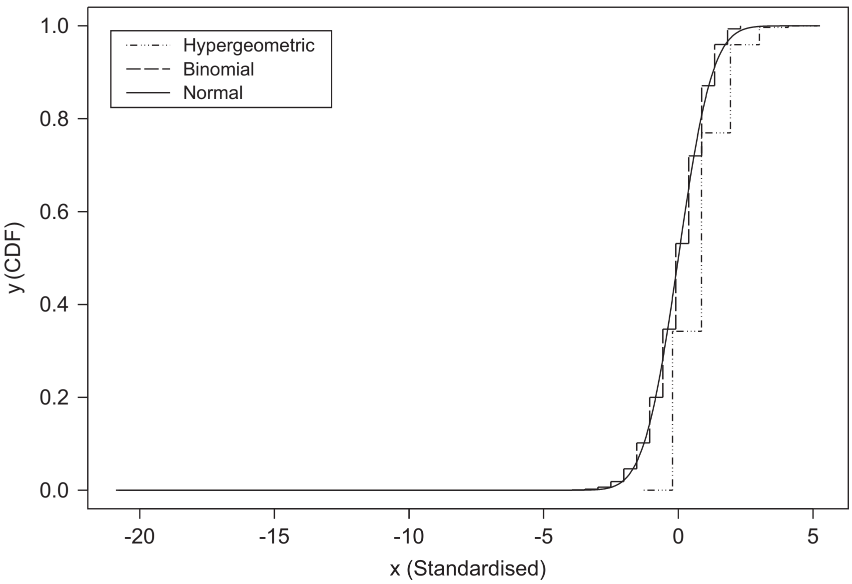

Fig. 4. A plot of the cdfs of normalized Hypergeometric and Binomial random variables against the standard Normal cdf for the parameter values $N = 60$, $p = 0.9$ and $f = 0.8$.

approximation; the maximal error of approximation to the Binomial (54, 0.8) distribution is 0.0803. However, with $N = 60$, $n = 54$ and $p = 0.9$, the maximal error of Normal approximation to the Hypergeometric distribution is as high as 0.4633, making the approximation practically useless. With about a 9-fold increase in the sample size, at $n = 450$, the accuracy of the approximation in the Hypergeometric case only improves to 0.1683 for the same values of f and p. The corresponding maximal error for the Normal approximation to the Binomial distribution with parameters $n = 450$ and $p = 0.8$ is only 0.0282. Thus, the loss in accuracy in this case is an astounding 600% compared to the Binomial

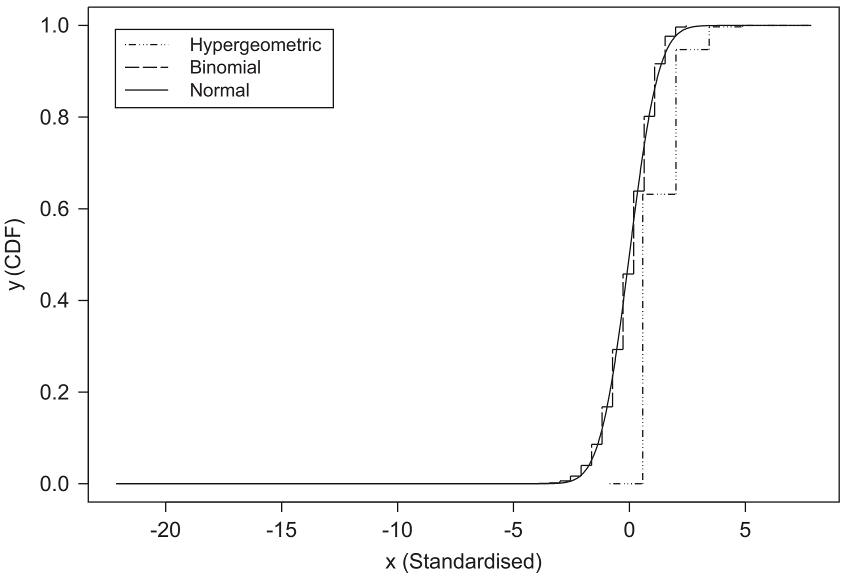

Fig. 5. A plot of the cdfs of normalized Hypergeometric and Binomial random variables against the standard Normal cdf for the parameter values $N = 60$, $p = 0.9$ and $f = 0.9$.

Table 2
Values of the maximal error of Normal approximation to Hypergeometric distribution (viz., $\Delta(N, p, f)$ of (3.1)) at $N = 60$ and the corresponding values of the maximal error for the Binomial (n, p) distribution where $n = Nf$ and $p, f \in \{0.5, 0.6, 0.7, 0.8, 0.9\}$

	Hypergeometric					Binomial				
	$n = 30$	36	42	48	54	$n = 30$	36	42	48	54
$p = 0.5$	0.1017	0.1038	0.1106	0.1260	0.1646	0.0722	0.0660	0.0612	0.0573	0.0540
$p = 0.6$	0.1038	0.1066	0.1171	0.1284	0.1817	0.0785	0.0722	0.0661	0.0626	0.0588
$p = 0.7$	0.1106	0.1171	0.1283	0.1476	0.1844	0.0888	0.0808	0.0757	0.0711	0.0670
$p = 0.8$	0.2268	0.2528	0.2896	0.3480	0.4633	0.1070	0.0992	0.0922	0.0859	0.0803
$p = 0.9$	0.3128	0.3550	0.4046	0.4633	0.7169	0.1474	0.1391	0.1289	0.1184	0.1145

Table 3
Values of the maximal error of Normal approximation to Hypergeometric distribution (viz., $\Delta(N, p, f)$ of (3.1)) at $N = 200$ and the corresponding values of the maximal error for the Binomial (n, p) distribution where $n = Nf$ and $p, f \in \{0.5, 0.6, 0.7, 0.8, 0.9\}$

	Hypergeometric					Binomial				
	$n = 100$	120	140	160	180	$n = 100$	120	140	160	180
p = 0.5	0.0562	0.0574	0.0613	0.0701	0.0929	0.0398	0.0363	0.0337	0.0315	0.0297
$p = 0.6$	0.0574	0.0593	0.0641	0.0743	0.0995	0.0433	0.0395	0.0366	0.0343	0.0323
$p = 0.7$	0.0613	0.0641	0.0702	0.0822	0.1112	0.0491	0.0449	0.0416	0.0389	0.0367
$p = 0.8$	0.1261	0.1373	0.1559	0.1887	0.2634	0.0595	0.0543	0.0503	0.0471	0.0444
$p = 0.9$	0.1764	0.1922	0.2181	0.2634	0.3652	0.0832	0.0761	0.0705	0.0661	0.0623

case. Indeed, similar high levels of loss in accuracy occur for values of f and p near 1 even when the population size N is increased to 2000 and beyond. As a consequence, the commonly used guidelines for the accuracy in the Binomial case can be misleading for assessing accuracy of the Normal approximation to the Hypergeometric distribution in the extreme cases.

Table 4
Values of the maximal error of Normal approximation to Hypergeometric distribution (viz., $\Delta(N, p, f)$ of (3.1)) at $N = 500$ and the corresponding values of the maximal error for the Binomial (n, p) distribution where $n = Nf$ and $p, f \in \{0.5, 0.6, 0.7, 0.8, 0.9\}$

	Hypergeometric					Binomial				
	$n = 250$	300	350	400	450	$n = 250$	300	350	400	450
$p = 0.5$	0.0356	0.0364	0.0389	0.0445	0.0592	0.0252	0.0230	0.0213	0.0199	0.0188
$p = 0.6$	0.0364	0.0376	0.0407	0.0472	0.0635	0.0274	0.0250	0.0232	0.0217	0.0205
$p = 0.7$	0.0389	0.0407	0.0446	0.0523	0.0712	0.0311	0.0284	0.0263	0.0246	0.0232
$p = 0.8$	0.0801	0.0872	0.0990	0.1200	0.1683	0.0378	0.0345	0.0319	0.0299	0.0282
$p = 0.9$	0.1124	0.1224	0.1390	0.1683	0.2355	0.0530	0.0484	0.0449	0.0420	0.0396

Table 5
Values of the maximal error of Normal approximation to Hypergeometric distribution (viz., $\Delta(N, p, f)$ of (3.1)) at $N = 2000$ and the corresponding values of the maximal error for the Binomial (n, p) distribution where $n = Nf$ and $p, f \in \{0.5, 0.6, 0.7, 0.8, 0.9\}$

	Hypergeometric					Binomial				
	$n = 1000$	1200	1400	1600	1800	$n = 1000$	1200	1400	1600	1800
$p = 0.5$	0.0178	0.0182	0.0195	0.0223	0.0297	0.0126	0.0115	0.0107	0.0010	0.0094
$p = 0.6$	0.0182	0.0188	0.0204	0.0237	0.0319	0.0137	0.0125	0.0116	0.0109	0.0102
$p = 0.7$	0.0195	0.0204	0.0224	0.0263	0.0358	0.0156	0.0142	0.0132	0.0123	0.0116
$p = 0.8$	0.0401	0.0437	0.0496	0.0602	0.0846	0.0189	0.0173	0.0160	0.0150	0.0141
$p = 0.9$	0.0564	0.0614	0.0698	0.0846	0.1189	0.0266	0.0243	0.0225	0.0210	0.0198

From Tables 2–5, it also follows that for a given value of N, if the parameter p is held fixed at a given level, the maximal error of approximation to the Hypergeometric distribution increases monotonically as the value of f (i.e., n) increases, and *vice versa*. However, the sample size n and the value of p alone do not give a true indication of the accuracy of the Normal approximation to the Hypergeometric distribution. To get a better estimate of the level of accuracy, one must consider the combined effect of all three parameters N, f and p. Theorem 1 implies that the combined effect of all three parameters on the maximal error of approximation can be expressed in terms of $1/\sigma(N, p, f)$, where $[\sigma(N, p, f)]^2 = Nfp(1-p)(1-f)$. To that effect, we define the co-efficient $c(N, p, f) \in (0, \infty)$ by the relation

$$c(N, p, f) = \Delta(N, p, f)\sigma(N, p, f). \tag{3.2}$$

Thus,

$$\Delta(N, p, f) \equiv c(N, p, f)[\sigma(N, p, f)]^{-1}.$$

Using Theorem 1 and the lattice property of $(X - E(X))/\sqrt{Var(X)}$, it can be shown that there exist two constants $C_5, C_6 \in (0, \infty)$ such that:

$$C_5 \leqslant c(N, p, f) \leqslant C_6$$

for all N, p, f, whenever $\sigma(N, p, f) > 0$. If we knew the approximate value of C_6 or of the co-efficient $c(N, p, f)$, we could use $C_6[\sigma(N, p, f)]^{-1}$ or $c(N, p, f)[\sigma(N, p, f)]^{-1}$ as a guideline for assessing the accuracy of Normal approximation for the Hypergeometric distribution. Tables 6–9 below give the values of the co-efficient $c(N, p, f)$ for different values of N, p, f.

Tables 6–9 show that the co-efficients $c(N, p, f)$ are surprisingly stable as a function of N, i.e., it exhibits very minor variation as a function of N for a given value of the pair (p, f). In a specific application with a given set of values of p and f, one can use $c(N, p, f)[\sigma(N, p, f)]^{-1}$ from Tables 6–9 to decide on the suitability of normal approximation to the Hypergeometric probabilities.

Table 6
Values of the co-efficients $c(N, p, f)$ at $N = 60$ for $p, f \in \{0.5, 0.6, 0.7, 0.8, 0.9\}$

	$f = 0.5$	0.6	0.7	0.8	0.9
$p = 0.5$	0.1970	0.1969	0.1964	0.1952	0.1913
$p = 0.6$	0.1969	0.1982	0.2036	0.1949	0.2069
$p = 0.7$	0.1964	0.2036	0.2086	0.2095	0.1963
$p = 0.8$	0.3513	0.3837	0.4112	0.4313	0.4306
$p = 0.9$	0.3634	0.4041	0.4308	0.4306	0.4998

Table 7
Values of the co-efficients $c(N, p, f)$ at $N = 200$ for $p, f \in \{0.5, 0.6, 0.7, 0.8, 0.9\}$

	$f = 0.5$	0.6	0.7	0.8	0.9
$p = 0.5$	0.1987	0.1987	0.1985	0.1982	0.1970
$p = 0.6$	0.1987	0.2012	0.2037	0.2058	0.2068
$p = 0.7$	0.1985	0.2037	0.2086	0.2132	0.2162
$p = 0.8$	0.3567	0.3806	0.4041	0.4269	0.4470
$p = 0.9$	0.3742	0.3994	0.4239	0.4470	0.46478

Table 8
Values of the co-efficients $c(N, p, f)$ at $N = 500$ for $p, f \in \{0.5, 0.6, 0.7, 0.8, 0.9\}$

	$f = 0.5$	0.6	0.7	0.8	0.9
$p = 0.5$	0.1992	0.1992	0.1991	0.1989	0.1985
$p = 0.6$	0.1992	0.2018	0.2043	0.2068	0.2088
$p = 0.7$	0.1991	0.2043	0.2095	0.2145	0.2189
$p = 0.8$	0.3581	0.3820	0.4058	0.4293	0.4517
$p = 0.9$	0.3771	0.4023	0.4273	0.4517	0.4740

Table 9
Values of the co-efficients $c(N, p, f)$ at $N = 2000$ for $p, f \in \{0.5, 0.6, 0.7, 0.8, 0.9\}$

	$f = 0.5$	0.6	0.7	0.8	0.9
$p = 0.5$	0.1994	0.1994	0.1994	0.1993	0.1992
$p = 0.6$	0.1994	0.2020	0.2047	0.2073	0.2098
$p = 0.7$	0.1994	0.2047	0.2010	0.2152	0.2203
$p = 0.8$	0.3588	0.3827	0.4066	0.4305	0.4540
$p = 0.9$	0.3785	0.4038	0.4290	0.4540	0.4785

4. Proofs

We now introduce some notation and notational convention to be used in this section. For real numbers x, y, let $x \wedge y = \min\{x, y\}$ and $x \vee y = \max\{x, y\}$. Let $\lfloor x \rfloor$ denote the largest integer not exceeding x, $x \in \mathbf{R}$. For $a \in (0, \infty)$, write $\phi_a(x) = (1/a)\phi(x/a)$ and $\Phi_a(x) = \Phi(x/a)$, $x \in \mathbf{R}$, for the density and distribution functions of a $\mathrm{N}(0, a^2)$ variable. Write $\phi_a = \phi$ and $\Phi_a = \Phi$ for $a = 1$. Let

$$\Delta_r^*(x) = P\left(\frac{X_r - n_r p_r}{\sigma_r} \leqslant x\right) - \Phi(x), \quad x \in \mathbf{R}. \tag{4.1}$$

Let $\mathbf{N} = \{1, 2, \ldots\}$, $\mathbf{Z}_+ = \{0, 1, \ldots\}$ and $\mathbf{Z} = \{\ldots, -1, 0, 1, \ldots\}$.

For notational simplicity, we shall drop the suffix r from notation, except when it is important to highlight the dependence on r. Thus, we write n, M, N for n_r, M_r, N_r, respectively, and set $p = M/N$, $q = 1 - p$ and $f = n/N$. We shall use C to denote a generic positive constant that does not depend on r. Unless otherwise stated, limits in order symbols are taken by letting $r \to \infty$.

For proving the result, we shall frequently make use of Stirling's approximation (cf. Feller, 1971)

$$m! = \sqrt{2\pi} e^{-m+\varepsilon_m} m^{m+1/2} \quad \text{for all } m \in \mathbf{N}, \tag{4.2}$$

where the error term ε_m admits the bound

$$\frac{1}{12m+1} \leqslant \varepsilon_m \leqslant \frac{1}{12m} \quad \text{for all } m \in \mathbf{N}.$$

Also note that for $g(y) = \log y$, $y \in (0, \infty)$, the kth derivative of g is given by $g^{(k)}(y) = ((-1)^{k-1}(k-1)!)/y^k$, $y \in (0, \infty)$, $k \in \mathbf{N}$. Hence, for any $k \in \mathbf{N}$ and $\delta \in (0, 1)$,

$$|g^{(k)}(1+x)| \leqslant \frac{(k-1)!}{(1-\delta)^k} \quad \text{for all } 0 \leqslant |x| < \delta. \tag{4.3}$$

For Lemma 1 below, let $X \sim \mathrm{Hyp}(n; M, N)$ for a *given* set of integers $n, M, N \in \mathbf{N}$ with $1 \leqslant n \leqslant (N-1)$, $1 \leqslant M \leqslant (N-1)$. Let

$$x_{k,n} = \frac{k - np}{\sqrt{npq}} \quad \text{and} \quad a_{k,n} = \frac{x_{k,n}}{(1-f)\sqrt{npq}}, \quad 0 \leqslant k \leqslant n, \tag{4.4}$$

where $f = n/N$, $p = M/N$ and $q = 1 - p$. Lemma 1 gives a basic approximation to Hypergeometric probabilities solely under condition (4.5) stated below.

Lemma 1. *Suppose that* $X \sim \mathrm{Hyp}(n; M, N)$ *for a* given *set of integers* $n, M, N \in \mathbf{N}$ *such that*

$$0 < f < 1, \quad 0 < p < 1 \quad \text{and} \quad 6(np \wedge nq) \geqslant 1, \tag{4.5}$$

where $f = n/N$, $p = M/N$ *and* $q = 1 - p$. *Then, for any given* $\delta \in (0, \frac{1}{2}]$,

$$\log P(k; n, M, N) = -\frac{x_{k,n}^2}{2(1-f)} - \frac{1}{2} \log(2\pi npq(1-f)) + r_n^*(k) \tag{4.6}$$

for all $k \in \{0, \ldots, n\}$ *with* $|a_{k,n}| \leqslant \delta$, *where* $P(k; n, M, N) = P(X = k)$ (*cf.* (1.1)) *and where the remainder term* $r_n^*(k)$ *admits the bound*

$$\begin{aligned} |r_n^*(k)| \leqslant{} & \frac{1}{6npq(1-\delta)(1-f)} + \left[\frac{1}{2}|a_{k,n}| + a_{k,n}^2 \left\{\frac{1}{4} + \frac{2\delta}{(1-\delta)^3}\right\}\right] \\ & + |a_{k,n}|^3 npq \left(\frac{f}{4} + 1\right) \left\{\frac{1}{2} + \frac{2(1+\delta)}{(1-\delta)^3}\right\}, \end{aligned} \tag{4.7}$$

provided $|a_{k,n}| \leqslant \delta$.

Proof. For $k \in \{0, 1, \ldots, n\}$,

$$\begin{aligned} P(k, n, M, N) &= \binom{n}{k} p^k q^{n-k} \frac{\prod_{j=1}^{k-1}(1 - j/Np) \prod_{j=1}^{n-k-1}(1 - j/Nq)}{\prod_{j=1}^{n-1}(1 - j/N)} \\ &= \binom{n}{k} p^k q^{n-k} \; R(k, n, M, N) \quad \text{say.} \end{aligned} \tag{4.8}$$

First consider $R(k; n, M, N)$. By (4.2),

$$\prod_{j=1}^{n-1}\left(1-\frac{j}{N}\right) = \frac{e^{(-N+\varepsilon_N)}N^{N+1/2}}{e^{(-(N-n)+\varepsilon_{N-n})}(N-n)^{N-n+1/2}}\,\frac{1}{N^n}$$

$$= \frac{e^{(\varepsilon_N-\varepsilon_{N-n})}e^{-n}}{(1-f)^{N(1-f)+1/2}},$$

$$\prod_{j=1}^{k-1}\left(1-\frac{j}{Np}\right)\prod_{j=1}^{n-k-1}\left(1-\frac{j}{Nq}\right) = \frac{e^{-n}e^{\varepsilon_{Np}-\varepsilon_{Np-k}+\varepsilon_{Nq}-\varepsilon_{Nq-n+k}}}{(1-k/Np)^{Np-k+1/2}(1-(n-k)/Nq)^{Nq-n+k+1/2}}.$$

Note that by (4.4),

$$\frac{k}{Np} = f + x_{k,n}\sqrt{\frac{fq}{Np}} \quad \text{and} \quad \frac{n-k}{Nq} = f - x_{k,n}\sqrt{\frac{fp}{Nq}}. \tag{4.9}$$

Next write

$$z_{k,n} = \frac{x_{k,n}\sqrt{fp/Nq}}{1-f}, \quad y_{k,n} = \frac{x_{k,n}\sqrt{fq/Np}}{1-f}$$

and

$$\varepsilon^* = \varepsilon_{Np} - \varepsilon_{Np-k} + \varepsilon_{Nq} - \varepsilon_{Nq-n+k} + \varepsilon_{N-n} - \varepsilon_N. \tag{4.10}$$

Then $R(k; n, M, N)$ can be expressed as

$$\begin{aligned}\log R(k; n, M, N) &= \varepsilon^* - \frac{\log(1-f)}{2} - \left(Np(1-f)(1-y_{k,n}) + \frac{1}{2}\right)\log(1-y_{k,n})\\ &\quad - \left(Nq(1-f)(1+z_{k,n}) + \frac{1}{2}\right)\log(1+z_{k,n})\\ &\equiv \varepsilon^* - \frac{\log(1-f)}{2} - A_1 - A_2 \quad \text{say.}\end{aligned} \tag{4.11}$$

Fix $\delta \in (0, \frac{1}{2})$. By Taylor's expansion and (4.3),

$$\begin{aligned}A_1 &= \left(Np(1-f)(1-y_{k,n}) + \frac{1}{2}\right)\log(1-y_{k,n})\\ &= \left(Np(1-f)(1-y_{k,n}) + \frac{1}{2}\right)\left(-y_{k,n} - \frac{y_{k,n}^2}{2} + r_{1n}(k)\right)\\ &= -\,y_{k,n}\left(Np(1-f) + \frac{1}{2}\right) - \frac{y_{k,n}^2}{2}\left(\frac{1}{2} - Np(1-f)\right) + r_{2n}(k),\end{aligned} \tag{4.12}$$

where $r_{1n}(k)$ and $r_{2n}(k)$ are remainder terms, defined by the equality of the successive expressions. By (4.3), for all n, k satisfying $|y_{k,n}| \leqslant \delta$,

$$|r_{1n}(k)| \leqslant \frac{2}{(1-\delta)^3}\,\frac{|y_{k,n}|^3}{3!}$$

and

$$|r_{2n}(k)| \leqslant \frac{Np}{2}(1-f)|y_{k,n}|^3 + \left|Np(1-f)(1-y_{k,n}) + \frac{1}{2}\right| \cdot |r_{1n}(k)|. \tag{4.13}$$

By similar arguments,

$$A_2 = \left[Nq(1-f)(1+z_{k,n}) + \frac{1}{2}\right]\log(1+z_{k,n})$$
$$= \left(Nq(1-f) + \frac{1}{2}\right) z_{k,n} + \frac{z_{k,n}^2}{2}\left[Nq(1-f) - \frac{1}{2}\right] + r_{3n}(k), \tag{4.14}$$

where for all n, k, satisfying $|z_{k,n}| \leqslant \delta$,

$$|r_{3n}(k)| \leqslant Nq(1-f)\frac{|z_{k,n}|^3}{2} + \left|Nq(1-f)(1+z_{k,n}) + \frac{1}{2}\right| \cdot \frac{|z_{k,n}|^3}{3(1-\delta)^3}. \tag{4.15}$$

From, (4.11), (4.12) and (4.14), we have

$$\log R(k; n, M, N) = \varepsilon^* - \frac{\log(1-f)}{2} - \left[\frac{(z_{k,n} - y_{k,n})}{2} + \frac{z_{k,n}^2}{2}\left\{Nq(1-f) - \frac{1}{2}\right\}\right.$$
$$\left. + \frac{y_{k,n}^2}{2}\left\{Np(1-f) - \frac{1}{2}\right\} + r_{2n}(k) + r_{3n}(k)\right]$$
$$= \varepsilon^* - \frac{1}{2}\log(1-f) - \frac{x_{k,n}^2 f}{2(1-f)} + r_{4n}(k), \tag{4.16}$$

where for all n, k satisfying $(|y_{k,n}| \vee |z_{k,n}|) \leqslant \delta$,

$$|r_{4n}(k)| \leqslant |r_{2n}(k)| + |r_{3n}(k)| + \tfrac{1}{2}|y_{k,n} - z_{k,n}| + \tfrac{1}{4}(y_{k,n}^2 + z_{k,n}^2).$$

Next using Stirling's formula on the binomial term, we have

$$\log\left\{\binom{n}{k} p^k q^{n-k}\right\}$$
$$= \log\left\{\frac{e^{(\varepsilon_n - \varepsilon_k - \varepsilon_{n-k})}}{\sqrt{2\pi npq}}\right\} - \left(nq - x_{k,n}\sqrt{npq} + \frac{1}{2}\right)\log\left\{1 - x_{k,n}\sqrt{\frac{p}{nq}}\right\}$$
$$- \left(np + x_{k,n}\sqrt{npq} + \frac{1}{2}\right)\log\left\{1 + x_{k,n}\sqrt{\frac{q}{np}}\right\}$$
$$\equiv \varepsilon^{**} - \log\sqrt{2\pi npq} - A_3 - A_4 \quad \text{say}, \tag{4.17}$$

where $\varepsilon^{**} = \varepsilon_n - \varepsilon_k - \varepsilon_{n-k}$. Next write $\tilde{y}_{k,n} = x_{k,n}\sqrt{p/nq}$ and $\tilde{z}_{k,n} = x_{k,n}\sqrt{q/np}$. Then, by arguments similar to (4.12) and (4.14),

$$A_3 = \left(nq - x_{k,n}\sqrt{npq} + \frac{1}{2}\right)\log\left(1 - x_{k,n}\sqrt{\frac{p}{nq}}\right)$$
$$= -\tilde{y}_{k,n}\left(nq + \frac{1}{2}\right) + \frac{\tilde{y}_{k,n}^2}{2}\left(nq - \frac{1}{2}\right) + r_{5n}(k)$$

and

$$A_4 = \left(np + x_{k,n}\sqrt{npq} + \frac{1}{2}\right)\log\left(1 + x_{k,n}\sqrt{\frac{q}{np}}\right)$$
$$= \tilde{z}_{k,n}\left(np + \frac{1}{2}\right) + \frac{\tilde{z}_{k,n}^2}{2}\left(np - \frac{1}{2}\right) + r_{6n}(k),$$

where for all k and n satisfying $(|\tilde{y}_{k,n}| \vee |\tilde{z}_{k,n}|) \leqslant \delta$,

$$|r_{5n}(k)| + |r_{6n}(k)| \leqslant \frac{n}{2}[q|\tilde{y}_{k,n}|^3 + p|\tilde{z}_{k,n}|^3] + \frac{2}{(1-\delta)^3}\left[\left(nq + \frac{1}{2} + nq|\tilde{y}_{k,n}|\right)|\tilde{y}_{k,n}|^3 + \left(np + \frac{1}{2} + np|\tilde{z}_{k,n}|\right)|\tilde{z}_{k,n}|^3\right]. \tag{4.18}$$

Hence, as in (4.16), it follows that

$$\log\left\{\binom{n}{k} p^k q^{n-k}\right\} = \varepsilon^{**} - \log\sqrt{2\pi npq} - \frac{1}{2}x_{k,n}^2 + r_{7n}(k), \tag{4.19}$$

where for all n, k satisfying $(|\tilde{y}_{k,n}| \vee |\tilde{z}_{k,n}|) \leqslant \delta$,

$$|r_{7n}(k)| \leqslant |\tfrac{1}{2}(\tilde{z}_{k,n} + \tilde{y}_{k,n}) - \tfrac{1}{4}(\tilde{y}_{k,n}^2 + \tilde{z}_{k,n}^2)| + |r_{5n}(k)| + |r_{6n}(k)|.$$

Note that

$$fq + fp + (1-f)p + (1-f)q = 1,$$

$$(fq)^2 + (fp)^2 + ((1-f)p)^2 + ((1-f)q)^2 = (1-2pq)(1-2(1-f)) < 1,$$

and by (4.4), $y_{k,n} = fqa_{k,n}$, $z_{k,n} = fpa_{k,n}$, $\tilde{y}_{k,n} = (1-f)pa_{k,n}$, and $\tilde{z}_{k,n} = (1-f)qa_{k,n}$. Hence, it follows that,

$$\tfrac{1}{2}|a_{k,n}| + \tfrac{1}{4}a_{k,n}^2 \geqslant \tfrac{1}{2}(|y_{k,n}| + |\tilde{y}_{k,n}| + |z_{k,n}| + |\tilde{z}_{k,n}|) + \tfrac{1}{4}(y_{k,n}^2 + \tilde{y}_{k,n}^2 + z_{k,n}^2 + \tilde{z}_{k,n}^2). \tag{4.20}$$

Now, combining (4.8), (4.16) and (4.18) and using (4.20) and the above identities, after some algebra, we get

$$\log P(k; n, M, N) = -\frac{x_{k,n}^2}{2(1-f)} - \frac{1}{2}\log(2\pi npq(1-f)) + r_n^*(k),$$

where for all k, n satisfying $|a_{k,n}| \leqslant \delta$,

$$\begin{aligned}
|r_n^*(k) - \varepsilon^* - \varepsilon^{**}| &\leqslant |r_{4n}(k)| + |r_{7n}(k)| \\
&\leqslant \frac{npq}{2}|a_{k,n}|^3[(1-f)(fq)^2 + (1-f)(fp)^2 + p^2 + q^2] \\
&\quad + \frac{2npq}{(1-\delta)^3}|a_{k,n}|^3[(1-f)f^2\{(1+\delta fq)q^2 + (1+\delta fp)p^2\} \\
&\quad + (1+\delta p)p^2 + (1+\delta q)q^2] + \frac{2}{(1-\delta)^3}|a_{k,n}|^3\frac{1}{2}[(1+f^3)(p^3+q^3)] \\
&\quad + \frac{1}{2}|a_{k,n}| + \frac{1}{4}a_{k,n}^2 \\
&\leqslant \frac{1}{2}|a_{k,n}| + a_{k,n}^2\left\{\frac{1}{4} + \frac{2\delta}{(1-\delta)^3}\right\} \\
&\quad + |a_{k,n}|^3 npq\left(\frac{f}{4} + 1\right)\left\{\frac{1}{2} + \frac{2(1+\delta)}{(1-\delta)^3}\right\}.
\end{aligned} \tag{4.21}$$

Note that for all k, n satisfying $|a_{k,n}| \leqslant \delta$,

$$Np - k \geqslant Np - (np + \delta(1-f)npq) > np\frac{(1-f)}{2} > 0$$

and

$$Nq - (n-k) > nq\frac{(1-f)}{2} > 0.$$

Hence, by the error bound in Stirling's approximation, for all k, n with $|a_{k,n}| \leqslant \delta$ and $6(np \wedge nq) \geqslant 1$,

$$\begin{aligned}
\varepsilon^* &\geqslant \frac{1}{12Np+1} - \frac{1}{12(Np-k)} + \frac{1}{12Nq+1} - \frac{1}{12(Nq-(n-k))} + \frac{1}{12(N-n)+1} - \frac{1}{12N} \\
&\geqslant -\frac{12k+1}{(12Np+1)(12(Np-k))} - \frac{12(n-k)+1}{(12Nq+1)(12(Nq-n+k))} \\
&\geqslant -\frac{1}{6Np(1-\delta)(1-f)} - \frac{1}{6Nq(1-\delta)(1-f)} \\
&= -\frac{f}{6npq(1-\delta)(1-f)},
\end{aligned}$$

$$\varepsilon^* \leqslant 0 + 0 + \left[\frac{1}{12(N-n)+1} - \frac{1}{12N}\right] \leqslant \frac{f}{6npq(1-\delta)(1-f)},$$

$$\varepsilon^{**} \leqslant \frac{1}{12n} - \frac{1}{12k+1} - \frac{1}{12(n-k)+1} \leqslant 0,$$

$$\varepsilon^{**} \geqslant \frac{1}{12n+1} - \frac{1}{12k} - \frac{1}{12(n-k)} \geqslant -\frac{n}{12k(n-k)} \geqslant -\frac{1}{6npq(1-\delta)}.$$

Hence, the lemma follows from (4.21) and the above inequalities. □

Lemma 2. *Let $g : \mathbf{R} \longrightarrow [0, \infty)$ be such that g is $\uparrow$ on $(-\infty, a)$ and g is $\downarrow$ on (a, ∞) for some $a \in \mathbf{R}$. Then, for any $k \in \mathbf{N}$, $b \in \mathbf{R}$ and $h \in (0, \infty)$,*

$$\sum_{i=o}^{k} g(b+ih) \leqslant \int_b^{b+hk} g(x)\,\mathrm{d}x + 2hg(x_0), \tag{4.22}$$

where $g(x_0) = \max\{g(b+ih) : i = 0, 1, \ldots, k\}$.

Proof. For $b \geqslant a$, by monotonicity,

$$h\sum_{i=0}^{k} g(b+ih) \leqslant hg(b) + \int_b^{b+hk} g(x)\,\mathrm{d}x.$$

For $b < a$, let $k_1 = \sup\{i : b + ih < a\}$ and $b_1 = b + k_1 h$. Then,

$$\begin{aligned}
h\sum_{i=0}^{k_1} g(b+ih) &\leqslant \sum_{i=0}^{k_1-1} \int_{b+ih}^{b+(i+1)h} g(x)\,\mathrm{d}x + hg(b+k_1 h) \\
&\leqslant \int_b^{b_1} g(x)\,\mathrm{d}x + hg(b_1).
\end{aligned}$$

Hence, for $b < a$ and $k > k_1$,

$$\begin{aligned}
h\sum_{i=0}^{k} g(b+ih) &= h\sum_{i=0}^{k_1} g(b+ih) + h\sum_{i=k_1+1}^{k} g(b+ih)\\
&= h\sum_{i=0}^{k_1} g(b+ih) + h\sum_{j=0}^{k-k_1-1} g(b_1+h+jh)\\
&\leqslant \int_b^{b_1} g(x)\,\mathrm{d}x + hg(b_1) + hg(b_1+h) + \int_{b_1+h}^{b_1+h+(k-k_1-1)h} g(x)\,\mathrm{d}x - hg(b_1)\\
&\leqslant \int_b^{b+hk} g(x)\,\mathrm{d}x + 2hg(x_0).
\end{aligned}$$

For $b < a$ and $k < k_1$, it is easy to check (using the arguments above) that bound (4.22) trivially holds. This completes the proof of the lemma. □

Lemma 3. *Let* $\phi(x) = (1/\sqrt{2\pi})\exp(-x^2/2)$, $x \in \mathbf{R}$. *Then, for any* $h \in (0,\infty)$, $b \in [0,\infty)$, $j_0 \in \mathbf{N}$,

$$\left| h\sum_{i=0}^{j_0} \phi(b+ih) - \int_{b-h/2}^{b+(j_0+1/2)h} \phi(x)\,\mathrm{d}x \right| \leqslant \frac{h^2}{12}\left[\int_{b-h/2}^{b+j_0h+h/2} |\phi''(x)|\,\mathrm{d}x \right.$$
$$\left. +(4+h)\max\left\{|\phi''(x)| : b-\frac{h}{2} < x < b+j_0h+\frac{h}{2}\right\}\right]. \quad (4.23)$$

Proof. Note that the function $|\phi''(x)| = |x^2-1|\phi(x)$ is even, and on $[0,\infty)$, it is increasing on $[1, 3^{1/2}]$ and decreasing on each of the intervals $[0, 1)$ and $(3^{1/2}, \infty)$, with the maximum value $1/\sqrt{2\pi}$ at $x = 0$ and the minimum value 0 at $x = 1$. First suppose that $(b-h/2, b+(j_0+\frac{1}{2})h) \cap \{0, \sqrt{3}\} = \emptyset$. Then, writing $b_i = b+ih$, $i \geqslant 0$, and using Taylor's expansion, one can show that the left side of (4.23) is bounded above by

$$\begin{aligned}
&\sum_{i=0}^{j_0} \left| \int_{b_i-h/2}^{b_i+h/2} (\phi(x)-\phi(b_i))\,\mathrm{d}x \right|\\
&\quad\leqslant \frac{1}{2}\sum_{i=0}^{j_0} \int_{b_i-h/2}^{b_i+h/2} (x-b_i)^2 \left\{\sup_{y\in(b_i-h/2,b_i+h/2)} |\phi''(y)|\right\}\mathrm{d}x\\
&\quad\leqslant \frac{1}{2}\sum_{i=0}^{j_0} \left(2\int_0^{h/2} y^2\,\mathrm{d}y\right) \times \left\{\left|\phi''\left(b_i-\frac{h}{2}\right)\right| \vee \left|\phi''\left(b_i+\frac{h}{2}\right)\right|\right\}\\
&\quad\leqslant \frac{h^3}{24}\sum_{i=0}^{j_0} \left\{\left|\phi''\left(b_i-\frac{h}{2}\right)\right| + \left|\phi''\left(b_i+\frac{h}{2}\right)\right|\right\}\\
&\quad\leqslant \frac{h^3}{12}\sum_{i=0}^{j_0+1} \left|\phi''\left(b_i-\frac{h}{2}\right)\right|.
\end{aligned}$$

Hence by two applications of Lemma 2, one can show that

$$h\sum_{i=0}^{j_0+1} \left|\phi''\left(b_i-\frac{h}{2}\right)\right| \leqslant \int_{b-h/2}^{b+j_0h+h/2} |\phi''(x)|\,\mathrm{d}x + 4\max\left\{|\phi''(x)| : b-\frac{h}{2} \leqslant x \leqslant b+j_0h+\frac{h}{2}\right\}.$$

Next consider the case where $0 \in [b-h/2, b+h/2)$. Then, by Taylor's expansion,

$$\left| h\phi(b) - \int_{b-h/2}^{b-h/2} \phi(x)\,\mathrm{d}x \right| \leqslant h^3 |\phi''(0)|/24.$$

Now using similar arguments for the case '$\sqrt{3} \in (b-h/2, b+(j_0+\frac{1}{2})h)$' and using the above bounds, one can complete the proof of the lemma. □

Proof of Theorem 1. Let $r \in \mathbf{N}$ be an integer such that (2.4) holds. Since r will be held *fixed* all through the proof, we shall drop r from the notation for simplicity, and write $f_r = f$, $\sigma_r = \sigma$, $p_r = p$, $q_r = q$, $n_r - n$, etc. First, suppose that $f \leqslant \frac{1}{2}$. Consider the case $x \leqslant 0$. Let $\tilde{x}_k = x_k/\sqrt{1-f} = (k-np)/\sigma$, $k = 0, 1, \ldots, n$. Define

$$K_0 = \sup\{k \in \mathbf{Z}_+ : \tilde{x}_k \leqslant 0\},$$

$$K_1 = \inf\{k \in \mathbf{Z}_+ : \tilde{x}_k \geqslant -1\},$$

$$K_2 = \inf\{k \in \mathbf{Z}_+ : \tilde{x}_k \geqslant -\delta\sigma\}$$

and

$$J_x = \lfloor np + x\sigma \rfloor, \quad x \in \mathbf{R},$$

where $\delta \equiv \delta_r \in (0, \frac{1}{2}]$ is as in (2.3). Note that by definition,

$$K_1 - 1 < np - \sigma \leqslant K_1, \quad K_2 - 1 < np - \delta\sigma^2 \leqslant K_2,$$

$$\tilde{x}_j \in [-1, 0] \quad \text{for all } K_1 \leqslant j \leqslant K_0 \quad \text{and} \quad \tilde{x}_j \in [-\delta\sigma, -1) \quad \text{for all } K_2 \leqslant j < K_1.$$

Hence, for any $x \in [-\delta\sigma, 0]$,

$$\begin{aligned} \left| P\left(\frac{X-np}{\sigma} \leqslant x\right) - \Phi(x) \right| &= |P(X \leqslant J_x) - \Phi(x)| \\ &\leqslant P(X < K_2) + \sum_{j=K_2}^{J_x} \left| P(X=j) - \frac{\phi(\tilde{x}_j)}{\sigma} \right| + \left| \sum_{j=K_2}^{J_x} \frac{\phi(\tilde{x}_j)}{\sigma} - \Phi(x) \right| \\ &= I_1 + I_2 + I_3 \quad \text{say.} \end{aligned} \tag{4.24}$$

Consider I_2 for $x \in [-\delta\sigma, -1)$. Note that for $x < -1$, $(J_x - np)/\sigma \leqslant x < -1$. Hence $J_x < K_1$ and $\tilde{x}_j < -1$ for all $j < J_x$. From Lemma 1,

$$\begin{aligned} |r^*(j)| &\leqslant \frac{1}{6\sigma^2(1-\delta)} + \left[\frac{|\tilde{x}_j|^2}{2\sigma} + \frac{|\tilde{x}_j|^2}{\sigma^2} \left\{ \frac{1}{4} + \frac{2\delta}{(1-\delta)^3} \right\} + \frac{|\tilde{x}_j|^3}{2\sigma} A \right], \\ &\equiv r^{**}(j), \end{aligned} \tag{4.25}$$

where $A = a_1(1 + (4(1+\delta))/((1-\delta)^3))$ and $a_1 \equiv a_{1r} = (f+4)/4(1-f)$ (cf. (2.3)). For the given choice of δ, it is easy to verify that $\delta \leqslant \frac{1}{20}$ and $\delta A < 0.59$. Hence

$$\begin{aligned} |r^*(j)| &\leqslant (0.2)\sigma^{-2} + \frac{\tilde{x}_j^2}{2} \left[\frac{1}{\sigma} + \frac{2}{\sigma^2}(0.3667) + \delta A \right] \\ &\leqslant (0.2)\sigma^{-2} + \frac{\tilde{x}_j^2}{2} \left[\min\{0.86, \frac{6}{5\sigma} + 0.59\} \right]. \end{aligned} \tag{4.26}$$

Now, from (4.25), for all $K_2 \leqslant j < K_1$,

$$|r^*(j)| \leqslant (0.2)\sigma^{-2} + |\tilde{x}_j|^3 \left| \left[\frac{1}{2\sigma} + \frac{1}{\sigma^2}(0.3667) + \frac{3a_1}{\sigma} \right] \leqslant 4|\tilde{x}_j|^3 \frac{a_1}{\sigma}. \tag{4.27}$$

Next note that $(J_x - np)/\sigma \leqslant x \in \mathbf{R}$, and

$$\int_a^\infty y^3 \exp\left(-\frac{by^2}{2}\right) \mathrm{d}y = \frac{1}{2b^2}(1 + ba^2)\mathrm{e}^{-ba^2} \quad \text{for all } a, b \in (0, \infty),$$

and that for any $a \in (0, \infty)$, the function $g(y; a) = y^3 \exp(-ay)$, $y \in [0, \infty)$, is increasing on $[0, \sqrt{3/2a}]$, and decreasing on $(\sqrt{3/2a}, \infty)$. Hence, by Lemmas 1 and 2, (4.26) and (4.27), with $c = 0.07$, we have

$$\begin{aligned}
I_2 &\leqslant \sum_{j=K_2}^{J_x} \left| \frac{\phi(\tilde{x}_j)}{\sigma} \exp(r^*(j)) - \frac{\phi(\tilde{x}_j)}{\sigma} \right| \\
&\leqslant \frac{1}{\sigma} \sum_{j=K_2}^{J_x} \phi(\tilde{x}_j)|r^*(j)| \exp(|r^*(j)|) \\
&\leqslant \frac{4a_1}{\sqrt{2\pi}\sigma^2} \exp(\sigma^{-2}) \sum_{j=K_2}^{J_x} |\tilde{x}_j|^3 \exp(-c\tilde{x}_j^2) \\
&\leqslant \frac{4a_1 \exp(\sigma^{-2})}{\sqrt{2\pi}\sigma} \left[\int_{(K_2-np)/\sigma}^{(J_x-np)/\sigma} |y|^3 \exp(-c|y|)\,\mathrm{d}y + \frac{2}{\sigma} \max\{|y|^3 \exp(-c|y|) : K_2 \leqslant np + \sigma y \leqslant J_x\} \right] \\
&\leqslant \frac{C}{\sigma(1-f)} [(1 + x^2) \exp(-cx^2)].
\end{aligned} \tag{4.28}$$

Also, for $-1 \leqslant x \leqslant 0$, by Lemma 1,

$$\begin{aligned}
\Delta_1(x) \equiv &\left| P\left(-1 \leqslant \frac{X - np}{\sigma} \leqslant x\right) - \sum_{j=K_1}^{K_0} \frac{1}{\sigma}\phi(\tilde{x}_j) \right| \\
&\leqslant \sum_{j=K_1}^{K_0} \left| P(X = j) - \frac{1}{\sigma}\phi(\tilde{x}_j) \right| \\
&\leqslant \sum_{j=K_1}^{K_0} \exp\left(-\frac{\tilde{x}_j^2}{2}\right) |r^*(j)| \frac{\exp(|r^*(j)|)}{\sqrt{2\pi}\sigma}.
\end{aligned}$$

For $K_1 \leqslant j \leqslant K_0$, from (4.25) and (4.26),

$$\begin{aligned}
|r^*(j)| &\leqslant \left[\frac{1}{2\sigma}|\tilde{x}_j| + r^{**}(j) \right] \wedge \left[\frac{1}{5\sigma^2} + \frac{1}{2\sigma} + \frac{1}{2\sigma^2}(0.3667) + \frac{A}{2\sigma} \right] \\
&\leqslant \left[\frac{1}{2\sigma} + \frac{1}{5\sigma^2} + (0.43)\tilde{x}_j^2 \right] \wedge \left[\frac{1}{2\sigma} + \frac{1}{5\sigma^2} + \frac{0.3667}{\sigma^2} + \frac{3a_1}{\sigma} \right] \\
&\leqslant \frac{1}{\sigma} + \left[(0.43)\tilde{x}_j^2 \right] \wedge \left[\frac{4a_1}{\sigma} \right].
\end{aligned}$$

Hence, for $-1 \leqslant x \leqslant 0$, noting that $K_0 - K_1 \leqslant \sigma$,

$$
\begin{aligned}
|\Delta_1(x)| &\leqslant \sum_{j=K_1}^{K_0} \exp(-\tilde{x}_j^2(0.07)) \exp(\sigma^{-1}) \frac{5a_1}{\sqrt{2\pi}\sigma^2} \\
&\leqslant (K_0 - K_1) \exp(\sigma^{-1}) \frac{5a_1}{\sqrt{2\pi}\sigma^2} \\
&\leqslant \frac{C}{\sigma}.
\end{aligned} \tag{4.29}
$$

Thus, the bound (4.28) on I_2 holds for all $x \in [-\delta\sigma, 0]$.

Next consider I_1. Note that for $j \in \{0, 1, \ldots, n\}$,

$$
\begin{aligned}
&P(X = j+1) \gtreqless P(X = j) \\
&\Leftrightarrow \frac{Np - j}{j+1} \cdot \frac{n - j}{Nq - n + j + 1} \gtreqless 1 \\
&\Leftrightarrow j \lesseqgtr \frac{n(Np+1)}{N+2} - \frac{Nq+1}{N+2}.
\end{aligned} \tag{4.30}
$$

Thus, $P(X = j) < P(X = j+1)$ for all $0 \leqslant j \leqslant np - 1$. Hence, by (4.26) and Lemma 1,

$$
\begin{aligned}
I_1 &= \sum_{j=0}^{K_2 - 1} P(X = j) \\
&< K_2 P(X = K_2) \\
&\leqslant K_2 \frac{1}{\sigma} \phi\left(\tilde{x}_{K_2}\right) \exp(r^*(K_2)) \\
&\leqslant \frac{K_2}{\sqrt{2\pi}\sigma} \exp\left(\frac{1}{5\sigma^2}\right) \exp(-\tilde{x}_{K_2}^2(.07)) \\
&\leqslant \frac{K_2}{\sqrt{2\pi}\sigma} \exp\left(\frac{1}{5\sigma^2}\right) \exp\left(-\left(\delta\sigma - \frac{1}{\sigma}\right)^2 (0.07)\right) \\
&\leqslant \frac{K_2}{\sqrt{2\pi}\sigma} \exp(-\delta^2\sigma^2(0.07) + 2\delta(0.07) + 0.13\sigma^{-2}) \\
&\leqslant \frac{np}{\sqrt{2\pi}\sigma} \exp(-\delta^2\sigma^2(0.07)) \exp(0.014) \\
&\leqslant (q(1-f))^{-1} \sigma \exp(-\delta^2\sigma^2(0.07)).
\end{aligned}
$$

It is easy to check that,

$$
\frac{\sigma \exp(-\delta^2\sigma^2(0.07))}{(1+x^2)\exp(-x^2(0.07))} \leqslant
\begin{cases}
\dfrac{2}{(0.07)\delta^2\sigma} & \text{if } x \in \left[0, \dfrac{\delta\sigma}{\sqrt{2}}\right], \\[2ex]
\dfrac{2}{\delta^2\sigma} & \text{if } x \in \left[\dfrac{\delta\sigma}{\sqrt{2}}, \delta\sigma\right].
\end{cases}
$$

Hence, it follows that for all $x \in [-\delta a, 0]$,

$$
I_1 \leqslant \frac{C}{\delta^2 q\sigma(1-f)} (1+x^2) \exp(-x^2(0.07)). \tag{4.31}
$$

Since $\tilde{x}_{J_x} \leqslant x$ and $\tilde{x}_{K_2} \leqslant -\delta\sigma + \sigma^{-1}$, by Lemma 3, for $x \in [-\delta\sigma, 0]$, one gets

$$\begin{aligned}
I_3 &\leqslant \left| \frac{1}{\sigma} \sum_{j=K_2}^{J_x} \phi(\tilde{x}_j) - \int_{\tilde{x}_{K_2}-(2\sigma)^{-1}}^{\tilde{x}_{J_x}+(2\sigma)^{-1}} \phi(y)\,\mathrm{d}y \right| + |\Phi(x) - \Phi(\tilde{x}_{J_x} + (2\sigma)^{-1})| + \Phi(\tilde{x}_{K_2} - (2\sigma)^{-1}) \\
&\leqslant \frac{1}{12\sigma^2} \left[\int_{-\infty}^{x+1/2\sigma} |\phi''(y)|\,\mathrm{d}y + 5 \max\left\{ |\phi''(y)| : -\infty < y < x + \frac{1}{2\sigma} \right\} \right] \\
&\quad + \Phi\left(x + \frac{1}{2\sigma}\right) - \Phi\left(x - \frac{1}{2\sigma}\right) + \Phi\left(-\delta\sigma + \frac{1}{2\sigma}\right).
\end{aligned}$$

Note that for any $a \in (0, \infty)$,

$$\begin{aligned}
&\int_a^\infty y^2 \mathrm{e}^{(-y^2/2)}\mathrm{d}y \leqslant \frac{1}{a} \int_a^\infty y^3 \mathrm{e}^{(-y^2/2)}\mathrm{d}y = \frac{2}{a} \int_{a^2/2}^\infty t\,\mathrm{e}^{-t}\,\mathrm{d}t = \frac{a^2+2}{a}\mathrm{e}^{-a^2/2}, \\
&\int_a^\infty y^2 \mathrm{e}^{-(y^2/2)}\,\mathrm{d}y \leqslant \int_0^\infty y^2 \mathrm{e}^{-(y^2/2)}\,\mathrm{d}y \leqslant \sqrt{\frac{\pi}{2}}, \\
&\max\{|\phi''(y)| : a < y < \infty\} \leqslant \frac{1}{\sqrt{2\pi}} I(0 < a < \sqrt{3}) + |\phi''(a)| I(a \geqslant \sqrt{3}).
\end{aligned}$$

And, for all $a \in (0, \delta\sigma)$,

$$\exp\left(-\frac{(a-(2\sigma)^{-1})^2}{2}\right) \leqslant \exp\left(-\frac{a^2}{2} + \frac{a}{2\sigma}\right) \leqslant \exp\left(-\frac{a^2}{2} + \frac{\delta}{2}\right).$$

Also note that, for $0 < a \leqslant 1$, $b \in (0, \infty)$,

$$\begin{aligned}
&1 - \Phi(b) \leqslant \frac{1}{b}\phi(b), \\
&1 - \Phi(a) \leqslant \int_a^1 \phi(x)\,\mathrm{d}x + \phi(1) \leqslant \phi(a)(1-a) + \phi(a) = (2-a)\phi(a).
\end{aligned}$$

Thus, for any $x \in (0, \infty)$,

$$\Phi(x) \leqslant \mathrm{e}^{-x^2/2}.$$

Since $(2\sigma)^{-1} < \frac{1}{8}$ and $|y + (2\sigma)^{-1}| \leqslant |y|$ for $y < -\frac{1}{8}$, we have, for all $x \in [-\delta a, 0]$,

$$\begin{aligned}
I_3 &\leqslant \frac{1}{12\sigma^2} \left[2I(-2 \leqslant x \leqslant 0) + 5|x|\phi\left(x + \frac{1}{2\sigma}\right) I(-\delta\sigma \leqslant x \leqslant -2) \right. \\
&\quad \left. +5\left\{ \frac{1}{\sqrt{2\pi}} I(-2 \leqslant x \leqslant 0) + (x^2+1)\phi\left(x + \frac{1}{2\sigma}\right) I(-\delta\sigma \leqslant x \leqslant -2) \right\} \right] \\
&\quad + \frac{1}{\sqrt{2\pi}\sigma} I(-2 \leqslant x \leqslant 0) + \frac{1}{\sigma}\phi\left(x + \frac{1}{2\sigma}\right) I(-\delta a \leqslant x < -2) \\
&\quad + \Phi\left(-\delta\sigma + \frac{1}{2\sigma}\right) \\
&\leqslant \frac{1}{2\sigma} I(-2 \leqslant x \leqslant 0) + 2\left\{ \frac{x^2+1}{2\sigma^2} + \frac{1}{\sigma} \right\} \phi\left(x + \frac{1}{2\sigma}\right) I(-\delta\sigma \leqslant x \leqslant -2) \\
&\quad + \exp\left(-\frac{(\delta\sigma - 1/2\sigma)^2}{2}\right) \\
&\leqslant \frac{C}{\sigma}(1 + |x|)\exp\left(-\frac{x^2}{2}\right). \qquad (4.32)
\end{aligned}$$

Next note that

$$P\left(\frac{X-np}{\sigma}\leqslant x\right)=0 \quad \text{for all } x<-\frac{np}{\sigma}$$

and for $-np/\sigma\leqslant x\leqslant -\delta\sigma$,

$$\begin{aligned}P\left(\frac{X-np}{\sigma}\leqslant x\right) &\leqslant I_1\leqslant (q(1-f))^{-1}\sigma\exp(-\delta^2\sigma^2(0.07))\\ &=(q(1-f))^{-1}\frac{(\delta\sigma)^2}{\delta^2\sigma}\exp\left(-\delta^2q^2(1-f)^2\left[\frac{-np}{\sigma}\right]^2(0.07)\right)\\ &\leqslant(\delta^2q(1-f)\sigma)^{-1}|x|^2\exp(-\delta^2q^2(1-f)^2x^2(0.07)).\end{aligned}$$

Hence, for all $x\leqslant -\delta\sigma$,

$$\begin{aligned}\left|P\left(\frac{X-np}{\sigma}\leqslant x\right)-\Phi(x)\right| &\leqslant\frac{|x|^2\exp(-\delta^2q^2(1-f)^2x^2(0.07))+\exp(-x^2/2)}{\delta q(1-f)\sigma}\\ &\leqslant\frac{2}{\delta q(1-f)\sigma}x^2\exp(-\delta^2q^2(1-f)^2x^2(0.07)).\end{aligned}\tag{4.33}$$

Now using the fact that $\delta\in[\frac{2}{45},\frac{1}{20}]$ for all $f\in(0,\frac{1}{2}]$, from (4.28),(4.29) and (4.31)–(4.33), it follows that there exist numerical constants C_1 and C_2, not depending on n, M, N, such that for all $x\in(-\infty,0]$,

$$\left|P\left(\frac{X-np}{\sigma}\leqslant x\right)-\Phi(x)\right|\leqslant\frac{C_1}{\sigma q}(1+x^2)\exp(-C_2qx^2),$$

provided $\delta\sigma>1$. This proves (2.5) for $x\in(-\infty,0]$ and $f\leqslant\frac{1}{2}$. To prove the theorem for $x\geqslant 0$ and $f\leqslant\frac{1}{2}$, define

$$V_r=n_r-X_r,\quad r\in\mathbf{N}.$$

Note that V_r has a Hypergeometric distribution with parameters (n_r, N_r-M_r, N_r). Further,

$$\frac{X_r-n_rp_r}{\sigma_r}=-\frac{V_r-n_rq_r}{\sigma_r}\quad\text{for all } r\in\mathbf{N}.$$

Hence, the derived bound on the right tails of $(X_r-n_rp_r)/\sigma_r$, can be obtained by repeating the arguments above with X_r replaced by V_r and p_r replaced by q_r for any r such that $\delta\sigma_r>1$. This proves (2.5) for $x\in[0,\infty)$ and $f\leqslant\frac{1}{2}$.

Next suppose that $f_r>\frac{1}{2}$. Consider the collection of N_r-n_r objects that are left after the sample of size n_r has been selected from the population of size N_r. Let Y_r is the number of 'type A'-objects in this collection. Then, for all $r\in\mathbf{N}$ and $j\in\mathbf{Z}$,

$$Y_r\sim \mathrm{Hyp}(N_r-n_r;M_r,N_r)\quad\text{and } P(X_r=j)=P(Y_r=M_r-j).\tag{4.34}$$

Hence,

$$P(X_r\leqslant k)=\sum_{j=0}^{k}P(X_r=j)=\sum_{j=0}^{k}P(Y_r=M_r-j)=P(Y_r\geqslant M_r-k).$$

Further, note that $Var(Y_r)=(N_r-n_r)p_rq_r(1-(N_r-n_r)/N_r)=\sigma_r^2$. Hence, for each $x\in\mathbf{R}$,

$$\begin{aligned}P\left(\frac{X_r-n_rp_r}{\sigma_r}\leqslant x\right) &=P\left(X_r\leqslant n_rp_r+x\sigma_r\right)\\ &=P(X_r\leqslant\lfloor n_rp_r+x\sigma_r\rfloor)\\ &=P(Y_r\geqslant M_r-\lfloor n_rp_r+x\sigma_r\rfloor)\\ &=P(\tilde{Y}_r\geqslant\check{x}_r)\quad\text{(say)},\end{aligned}$$

where $\tilde{Y}_r = (Y_r - (N_r - n_r)p_r)/\sigma_r$ and $\check{x}_r = (M_r - \lfloor n_r p_r + x\sigma_r \rfloor - (N_r - n_r)p_r)/\sigma_r$. Note that,

$$\check{x}_r < \frac{1}{\sigma_r}[N_r p_r - (n_r p_r + x\sigma_r - 1) - N_r p_r + n_r p_r] = -x + \sigma_r^{-1}$$

and similarly, $\check{x}_r \geqslant -x$. Hence, this implies,

$$P(\tilde{Y}_r < \check{x}_r) \leqslant P(\tilde{Y}_r \leqslant \check{x}_r) \leqslant P(\tilde{Y}_r \leqslant -x + \sigma_r^{-1})$$

and

$$P(\tilde{Y}_r < \check{x}_r) \geqslant P(\tilde{Y}_r < -x) \geqslant P(\tilde{Y}_r \leqslant -x - \sigma_r^{-1}).$$

Now using the above identity and inequalities, we have

$$\begin{aligned}\left| P\left(\frac{X_r - n_r p_r}{\sigma_r} \leqslant x\right) - \Phi(x) \right| &= |P(\tilde{Y}_r \geqslant \check{x}_r) - (1 - \Phi(-x))| \\ &= |\Phi(-x) - P(\tilde{Y}_r < \check{x}_r)| \\ &\leqslant \max\{|P(\tilde{Y}_r \leqslant -x - \sigma_r^{-1}) - \Phi(-x - \sigma_r^{-1})|, \\ &\quad |P(\tilde{Y}_r \leqslant -x + \sigma_r^{-1}) - \Phi(-x + \sigma_r^{-1})|\} \\ &\quad + \max\{|\Phi(-x) - \Phi(-x - \sigma_r^{-1})|, |\Phi(-x) - \Phi(-x + \sigma_r^{-1})|\}. \qquad (4.35)\end{aligned}$$

The proof of (2.5) for '$f \in [\frac{1}{2}, 1]$ and $x \in \mathbf{R}$' follows by replacing the above arguments with X_r, f_r replaced by Y_r, $1 - f_r$, respectively, and using the bound (4.34) and (4.35). This completes the proof of the theorem. □

Proof of Theorem 2. Use (2.5) and the inequality '$\exp(x) \geqslant (1 + x)$ for all $x \in (0, \infty)$'. □

References

Blyth, C.R., Staudte, R.G., 1997. Hypothesis estimates and acceptability profiles for 2 × 2 contingency tables. J. Amer. Statist. Assoc. 92, 694–699.
Burstein, H., 1975. Finite population correction for binomial confidence limits. J. Amer. Statist. Assoc. 70, 67–69.
Erdos, P., Renyi, A., 1959. On the central limit theorem for samples from a finite population. Magyar Tudoanyos Akademia Budapest Matematikai Kutato Intezet Koezlemenyei. Trudy Publications, vol. 4, pp. 49–57.
Feller, W., 1971. An Introduction to Probability Theory and its Applications, vol. I. Wiley, New York, NY.
Hajek, J., 1960. Limiting distributions in simple random sampling from a finite population. Magyar Tudoanyos Akademia Budapest Matematikai Kutato Intezet Koezlemenyei. Trudy Publications, vol. 5, pp. 361–374.
Nicholson, W.L., 1956. On the normal approximation to the hypergeometric distribution. Ann. Math. Statist. 27, 471–483.
Madow, W.G., 1948. On the limiting distributions of estimates based on samples from finite universes. Ann. Math. Statist. 19, 535–545.
Patel, J.K., Samaranayake, V.A., 1991. Predictions intervals for some discrete distributions. J. Qual. Technol. 23, 270–278.
Seber, G.A.F., 1970. The effects of trap response on tag recapture estimates. Biometrics 26, 13–22.
Sohn, S.Y., 1997. Accelerated life-tests for intermittent destructive inspection, with logistic failure-distribution. IEEE Trans. Reliab. 46, 122–1295.
von Collani, E., 1986. The α-optimal acceptance sampling scheme. J. Qual. Technol. 18, 63–66.
Wendell, J.P., Schmee, J., 1996. Exact inference for proportions from a stratified finite population. J. Amer. Statist. Assoc. 91, 825–830.
Wittes, J.T., 1972. On the bias and estimated variance of Chapman's two-sample capture-recapture population estimate. Biometrics 28, 592–597.