

Research Article

Improved Inequalities for the Poisson and Binomial Distribution and Upper Tail Quantile Functions

Michael Short

Electronics & Control Group, Teesside University, Middlesbrough TS1 3BA, UK

Correspondence should be addressed to Michael Short; m.short@tees.ac.uk

Received 29 October 2013; Accepted 25 November 2013

Academic Editors: V. Makis and A. Volodin

Copyright © 2013 Michael Short. This is an open access article distributed under the Creative Commons Attribution License, which permits unrestricted use, distribution, and reproduction in any medium, provided the original work is properly cited.

The exact evaluation of the Poisson and Binomial cumulative distribution and inverse (quantile) functions may be too challenging or unnecessary for some applications, and simpler solutions (typically obtained by applying Normal approximations or exponential inequalities) may be desired in some situations. Although Normal distribution approximations are easy to apply and potentially very accurate, error signs are typically unknown; error signs are typically known for exponential inequalities at the expense of some pessimism. In this paper, recent work describing universal inequalities relating the Normal and Binomial distribution functions is extended to cover the Poisson distribution function; new quantile function inequalities are then obtained for both distributions. Exponential bounds—which improve upon the Chernoff-Hoeffding inequalities by a factor of at least two—are also obtained for both distributions.

1. Introduction

The Poisson and Binomial distributions are a good approximation for many random phenomena in areas such as telecommunications and reliability engineering, as well as the biological and managerial sciences [1, 2]. Let $Y \sim \text{Poi}(m)$ be a Poisson distributed random variable having mean $m > 0$, and let $P\{Y \leq k\}$ represent the cumulative distribution function (CDF) of Y with nonnegative integer support $k \in \{0, 1, \dots, \infty\}$:

$$P\{Y \leq k\} = e^{-m} \sum_{i=0}^k \frac{m^i}{i!}. \quad (1)$$

Similarly, let $X \sim \text{Bin}(n, p)$ be a Binomially distributed random variable with parameters $n \in \{1, 2, 3, 4, \dots\}$ and $p \in (0, 1)$, and let $P\{X \leq k\}$ represent the CDF of X for integer support $k \in \{0, 1, \dots, n\}$:

$$P\{X \leq k\} = \sum_{i=0}^k \binom{n}{i} p^i (1-p)^{n-i}. \quad (2)$$

Also, let the R th quantiles of Y and X for $R \in (0, 1)$ be obtained from the functions $Q_P(m, R)$ and $Q_B(n, p, R)$:

$$Q_P(m, R) = \{\min k \in N : P\{Y \leq k\} \geq R\}, \quad (3)$$

$$Q_B(n, p, R) = \{\min k \in N : P\{X \leq k\} \geq R\}. \quad (4)$$

Due to numerical and complexity issues, evaluation of the exponential and Binomial summations in (1) and (2) through recursive operations is only practical for small values of the input parameters (m or np and k). Instead, a better solution is to evaluate the CDFs directly through either their incomplete Beta/Gamma function representations which can be approximated to high precision by continued fractions or asymptotic expansions [3]. With respect to the quantiles of the distributions given by (3) and (4), no methods to exactly evaluate these functions without iterating the exponential/Binomial sums—or alternately employing a search until the required conditions are satisfied—seem to be known. Typically, a binary search to determine the smallest k satisfying (3) or (4) evaluating the respective CDF at each step would be a better general solution, given some initial upper bound for

k. Such methods (and related variants) are now employed very effectively in modern commercial and research-based statistical packages.

In some situations, one may desire simpler solutions to either approximate or bound these quantities. Typically an approximation can be obtained via the standard Normal distribution; the work of Mollenaar [4] contains a good description of several applicable variants. Although quickly applied and potentially very accurate for large input parameters (due to the central limit theorem), the sign of the approximation errors is typically unknown. Exceptions are the inequality of Bohman (see [1, page 169]), which always overestimates the true Poisson probability, and the expressions recently proposed by Zubkov and Serov for the Binomial distribution [5]. Methods to obtain provable bounds with known error signs (typically one would require to underestimate (1) and (2), whilst overestimating (3) and (4) in most engineering and computer science applications) principally include the Bernstein/Chernoff/Hoeffding-type exponential probability inequalities and their close variants [1, 2, 6, 7]. Although effective, one has to accept the unavoidable loss of accuracy with these bounds.

Although the need for provable, accurate bounds has been well documented in computer science, information engineering, and reliability analysis applications ([1, 2] provide such discussions), the motivation for the current work arose from a recent application in probabilistic schedulability analysis for real-time systems described by Short and Proenza [8]. In this work, the authors consider efficient admission controls for providing probabilistic schedulability guarantees for real-time messages traversing communication channels with error arrival characteristics that can be approximated by Poisson or Binomial distributions. Ultimately, it is required in this application to evaluate many upper tail quantiles (directly corresponding to (3) or (4)) in a very short space of time by a (possibly resource-constrained) embedded microcontroller or microcomputer. Clearly the use of a commercial statistical package is not possible; several logarithmic inequalities were instead developed for these purposes. Although the bounds were shown to be tight in terms of relative errors (which become vanishingly small as the input magnitude becomes large), the absolute errors on the other hand become increasingly large as the input parameters increase. Therefore, one of the motivations for the current work was to tighten these quantile inequalities, with the goal of making them asymptotically near-exact in the input parameters. In this paper such tighter bounds are obtained, along with several other inequalities which may have a more general interest.

The remainder of the article is organized as follows. In Section 2, the recent work on the categorization of the Binomial with respect to the Normal distribution in [5] is first extended to obtain universal inequalities (with known error signs) relating the Poisson and Normal distribution functions. Section 3 obtains asymptotically near-exact analytic inequalities relating both the Poisson and Binomial upper tail quantiles to the Normal quantiles. Improved Chernoff/Hoeffding-type exponential inequalities are then obtained for both distributions in Section 4. A brief summary is then given in Section 5.

2. Distribution Function Inequalities

Consider the following recently proven universal inequality on the distribution function of a Binomially distributed random variable.

Theorem 1. *Let $X \sim \text{Bin}(n, p)$ be a random variable with parameters $n \in \{1, 2, 3, 4, \dots\}$ and $p \in (0, 1)$, where the integer n represents the number of trials and p the probability of success in each trial. Denoting the CDF of X as $P\{X \leq k\}$ as per (2), then, for $k = 0$ and $k = n$, one has the exact equalities $P\{X \leq 0\} = (1 - p)^n$ and $P\{X \leq n\} = 1$, and for all $n > 0$, $p \in (0, 1)$, and $k \in \{1, 2, \dots, n - 2\}$ the following inequalities hold:*

$$\begin{aligned} & \Phi \left(\text{sign}(k - np) \cdot \sqrt{2nD \left(p, \frac{k}{n} \right)} \right) \\ & < P\{X \leq k\} \\ & < \Phi \left(\text{sign}(k + 1 - np) \cdot \sqrt{2nD \left(p, \frac{(k + 1)}{n} \right)} \right). \end{aligned} \quad (5)$$

And for $k = n - 1$ one also has:

$$\Phi \left(\text{sign}(k - np) \cdot \sqrt{2nD \left(p, \frac{k}{n} \right)} \right) < P\{X \leq n - 1\}, \quad (6)$$

where $\text{sign}(x)$ is the usual signum function with argument x , $\Phi(y)$ is the distribution function of a standard normal variable with argument y , and the function $D(p, c) = c \cdot \ln(c/p) + (1 - c) \cdot \ln((1 - c)/(1 - p))$.

Proof. Zubkov and Serov [5]. □

Although it was not explicitly denoted as such in [5], it is easy to see that $D(p, c)$ represents the Kullback-Leibler (KL) divergence between two Bernoulli variables with respective probabilities of success p and c ; hence, $nD(p, c)$ represents the KL divergence of n summed pairs of such variables. This observation allows the relatively straightforward extension of the above result to the case of a Poisson distributed random variable, which is given in Theorem 2.

Theorem 2. *Let $Y \sim \text{Poi}(m)$ be a Poisson distributed random variable with mean m . Let the distribution function $P\{Y \leq k\}$ be defined as in (1), with integer support $k \in \{0, 1, \dots, \infty\}$. For $k = 0$ and $k = \infty$, one has $P\{Y \leq 0\} = e^{-m}$ and $P\{Y \leq \infty\} = 1$. For every other $k = \{1, 2, 3, \dots\}$ one has the following inequalities:*

$$\begin{aligned} & \Phi \left(\text{sign}(k - m) \cdot \sqrt{2H(m, k)} \right) \\ & < P\{Y \leq k\} \\ & < \Phi \left(\text{sign}(k + 1 - m) \cdot \sqrt{2H(m, k + 1)} \right), \end{aligned} \quad (7)$$

where $H(m, k)$ is the KL divergence between two Poisson distributed random variables with respective means m and k :

$$H(m, k) = m - k + k \cdot \ln \left(\frac{k}{m} \right). \quad (8)$$

Proof. The cases $k = 0$ and $k = \infty$ are exact equalities which are easily derived from the distribution function (1). For the other cases, first we form the variable $X \sim \text{Bin}(n, p)$ with some finite integer n and $p \in (0, 1)$, such that $np = m$. Clearly, this variable satisfies inequality (5) for any choice of $n - 2 > k > 0$. Now, suppose we increase n by one and reduce p such that the constraint $np = m$ still holds. Again, this variable still satisfies inequality (5). Now, incrementally repeat this procedure for increasing n under the constraint that $np = m$ indefinitely; in the limit as $n \rightarrow \infty$ one has the following:

$$\lim_{\substack{n \rightarrow \infty \\ np=m}} \left[P\{X \leq k\} = \sum_{i=0}^k \binom{n}{i} p^i (1-p)^{n-i} \right] = e^{-m} \sum_{i=0}^k \frac{m^i}{i!}. \quad (9)$$

With the identity above being the famous limit theorem of Poisson, a contemporary description of which may be found in [1]. Now consider the limit of the argument to $\Phi(x)$; observe that as $n \rightarrow \infty$ as the Bernoulli variables with success probability $p = m/n$ become Poisson distributed with mean $np = nm/n = m$, then it must also follow that as $n \rightarrow \infty$ the infinite number of Bernoulli variables with success probability k/n must also become Poisson distributed with mean $nk/n = k$. Thus, the KL divergence of this infinite number of Bernoulli variable pairs with respective success probabilities m/n and k/n becomes the KL divergence of two Poisson distributions with respective means m and k , which is given by the function $H(m, k) = m - k + k \cdot \ln(k/m)$ [9]. Thus, the following limit must hold:

$$\lim_{\substack{n \rightarrow \infty \\ np=m}} \left[nD\left(p, \frac{k}{n}\right) \right] = m - k + k \cdot \ln\left(\frac{k}{m}\right). \quad (10)$$

Noting that, as (5) holds at each step as n is increased to infinity, it must also hold in this limit for any $k = \{1, 2, 3, \dots\}$, which implies the inequalities stated in (7). \square

As in the Binomial case, these relationships may be used to bound the quantile of a Poisson random variable to a pair of adjacent integers. However, given a desired R , no analytical expression for a corresponding k may be obtained due to the presence of the term $k \ln(k/m)$ in the expression for $H(m, k)$. A similar restriction occurs with $D(p, c)$ due to the presence of multiple such terms involving c and p and their natural logarithm. Slightly weaker inequalities having an analytical form for the upper tail quantiles are thus obtained in the next section.

3. Upper Tail Quantile Inequalities

The first step in the Poisson quantile inequality is to obtain an expressive bound on the function $H(m, k)$ defined by (8).

Lemma 3. For any $m > 0$ and $0 < k < \infty$, one has.

$$H(m, k) \geq H_b(m, k) = 6m + 3k - 3\sqrt{3m^2 + 6mk}. \quad (11)$$

Proof. Observe that, for $k = m$, $H(m, m) = H_b(m, m) = 0$. Elementary calculations yield the first partial derivatives of both functions with respect to k :

$$\begin{aligned} H(m, k)' &= \ln\left(\frac{k}{m}\right), \\ H_b(m, k)' &= 3 - \sqrt{\frac{27m^2}{m^2 + 2mk}}. \end{aligned} \quad (12)$$

And observe again that, when $k = m$, $H(m, m)' = H_b(m, m)' = 0$. Now considering the sign of the derivatives, when $k > m$ we have that $H(m, k)' > 0$ since $k/m > 1$; we also have that $H_b(m, k)' > 0$ as the quantity $m^2 + 2mk > 3m^2$. Similarly, for $k < m$, $H(m, k)' < 0$ since $k/m < 1$ and $H_b(m, k)' < 0$ as $m^2 + 2mk < 3m^2$. Thus, both functions are monotonically decreasing in k over the interval $(0, m)$ and monotonically increasing in k over the interval (m, ∞) .

A further application of the calculus yields the second partial derivatives of $H(m, k)$ and $H_b(m, k)$ with respect to k :

$$\begin{aligned} H(m, k)'' &= k^{-1}, \\ H_b(m, k)'' &= \frac{27m^2}{(3m^2 + 6mk)^{1.5}}. \end{aligned} \quad (13)$$

And it is easy to verify that both functions are positive for all positive nonzero k and m . Now, the objective is to show that $H(m, k) \geq H_b(m, k)$ for the specified ranges of m and k . Form the function $f(m, k) = (H(m, k)'' - H_b(m, k)'')''$. Standard analytic techniques yield the two roots of f as $k = m$ and $k = -m/8$. As only the former root lies in the interval of positive k and m , $H(m, k)''$ and $H_b(m, k)''$ can intersect only at this root which implies that the sign of the function f can only (potentially) change once at this location. Verification that $H(m, k)''$ dominates $H_b(m, k)''$ reduces to demonstrating that, for arbitrary positive non-zero m , $f > 0$ for some k satisfying $0 < k < m$ and also that $f > 0$ for some $k > m$. For simplicity, let us choose $k = m/2$ and $k = 2m$:

$$\begin{aligned} f\left(m, \frac{m}{2}\right) &= \left(\frac{2}{m} - \frac{27}{(6\sqrt{6})m}\right) = \frac{0.1628\dots}{m} > 0, \\ f(m, 2m) &= \left(\frac{1}{2m} - \frac{27}{(15\sqrt{15})m}\right) = \frac{0.0352\dots}{m} > 0. \end{aligned} \quad (14)$$

Therefore, $H(m, k)'' \geq H_b(m, k)''$ for positive k and m , with equality occurring if and only if $k = m$ whence both functions equal k^{-1} . The lemma follows by observing that, when $k = m$, $H(m, m) = H_b(m, m) = H(m, m)' = H_b(m, m)' = 0$, and moving k away from m either towards zero or infinity causes a smaller corresponding increase in $H_b(m, k)$ than in $H(m, k)$ due to the dominance of $H(m, k)''$ over $H_b(m, k)''$. \square

Remarks. $H_b(m, k)$ provides a good quality bound over all ranges of k and m , but it is at its tightest when $k \approx m$; as the

mean m increases and the quantile clusters around the mean ($k \approx m$), the bound can be expected to be asymptotically very tight. Although the right-hand side of (11) seems to have a strange form, it allows one to obtain a tail quantile bound with a simple structure. Let $\Phi^{-1}(R)$ for $R \in (0, 1)$ be the inverse of the standard normal CDF $\Phi(x)$ (i.e., the “probit” function), such that $\Phi(\Phi^{-1}(R)) = R$. The inequality can be stated.

Theorem 4. *Let $Y \sim \text{Poi}(m)$ be a Poisson distributed random variable with mean m , and let $1 > R > 0.5$. Then, one has the following bound on the R th quantile of Y :*

$$Q_P(m, R) \leq \left\lceil m + \Phi^{-1}(R) \sqrt{m} + \frac{\Phi^{-1}(R)^2}{6} \right\rceil. \quad (15)$$

Proof. Consider the lower inequality in (7), and suppose that the left-hand side evaluates to R ; this clearly implies $R > P\{Y \leq k\}$. Since $R > 0.5$, we are working in the upper tail ($k > m$) and $\text{sign}(k - m) = 1$; thus, we seek an integer k such that the following holds true:

$$\Phi(\sqrt{2H(m, k)}) \geq R. \quad (16)$$

Applying the probit function to both sides of the equation, then squaring both sides and dividing through by 2 isolates H on the left hand side:

$$\begin{aligned} \sqrt{2H(m, k)} &\geq \Phi^{-1}(R), \\ H(m, k) &\geq \frac{(\Phi^{-1}(R))^2}{2}. \end{aligned} \quad (17)$$

Recalling that $H(m, k) \geq H_b(m, k)$, an integer k which satisfies $H_b(m, k) \geq \Phi^{-1}(R)^2/2$ also clearly guarantees that $P\{Y \leq k\} > R$. At this point, let us substitute into (17) the label $C = \Phi^{-1}(R)^2/2$, let k be real valued, assume an equality, and replace $H(m, k)$ with the definition of $H_b(m, k)$ from (11):

$$6m + 3k - 3\sqrt{3m^2 + 6mk} = C. \quad (18)$$

Simple rearrangement and squaring eliminate the square root:

$$\begin{aligned} \sqrt{27m^2 + 54mk} &= 6m + 3k - C, \\ 27m^2 + 54mk &= (6m + 3k - C)^2. \end{aligned} \quad (19)$$

Expanding the right hand side and then gathering terms leads to a quadratic in k :

$$(9)k^2 - (6C + 18m)k - (C^2 + 9m^2 + 12Cm) = 0. \quad (20)$$

Taking the principle root (to ensure $k > m$) gives:

$$k = m + \sqrt{2Cm} + \frac{C}{3}, \quad (21)$$

from which (15) is recovered by employing the ceiling function to make k integer and then substituting $C = \Phi^{-1}(R)^2/2$ back into the resulting expression. \square

For the Binomial quantile, below is presented a corresponding bound for the Binomial upper tail quantile $Q_B(n, p, R)$. The inequality is very similar to that of Theorem 4, but skewness correction terms now also appear.

Theorem 5. *Let $X \sim \text{Bin}(n, p)$ be a Binomial distributed random variable with parameters n and p , and let $1 > R > 0.5$. The following bound holds on the R th quantile of X :*

$$\begin{aligned} Q_B(n, p, R) &\leq \left\lceil np + \sqrt{\Phi^{-1}(R) \cdot \left[np(1-p) + (1-2p) \frac{\Phi^{-1}(R)^2}{6} \right]} \right. \\ &\quad \left. + (1-2p) \frac{\Phi^{-1}(R)^2}{6} \right\rceil. \end{aligned} \quad (22)$$

Proof. The proof proceeds upon almost identical lines to that of Theorem 4, except that the lower inequality in (5) provides the starting point. The inequality on $D(p, c)$ given in (23) below, which as shown by Janson [10] is valid for $c \geq p$ provides the analytical relaxation allowing the quantile to be explicitly solved for:

$$D(p, c) \geq \frac{(c-p)^2}{2(np(1-p) + (1-2p) \cdot (c-p)/3)}. \quad (23)$$

Simplifying the resulting expression leads to (22). \square

Observe that the Poisson upper tail quantile inequality (15) is virtually identical to the expression obtained from applying a Cornish-Fisher expansion with a continuity correction to the Poisson quantile [4]:

$$\begin{aligned} Q_P(m, R) &\approx m + \Phi^{-1}(R) \sqrt{m} \\ &\quad + \frac{\Phi^{-1}(R)^2}{6} - \frac{4}{6} + O\left(\frac{1}{\sqrt{m}}\right). \end{aligned} \quad (24)$$

It is immediately seen from Theorem 4 that, if overestimation of the true quantile by a small factor is always desired, it suffices to simply neglect the $4/6$ correction and the contribution of the asymptotic term in the above expansion. The sharpness of (15) is evident as for sufficiently large m the contribution of the term $O(m^{-0.5}) \rightarrow 0$ and the gap between (15) and (29) is always ≤ 1 . A similar relationship holds between expression (22) and the asymptotic expansion of the Binomial quantile [4]. Practical experience with both quantile bounds indicates that equalities can be achieved even for very small values of standard deviation: Figure 1 gives an illustrative comparison between the exact and bounded R th quantiles for a Poisson process having a mean of $m(t) = 0.1t$ for increasing values of t .

The sharpness of the bound is evident, and it may be observed that the gap between the bound and the exact quantile quickly reduces as t (and hence the mean and standard deviation) increases in both cases. This illustrates the quick convergence to the observation that the asymptotic gap is always ≤ 1 .

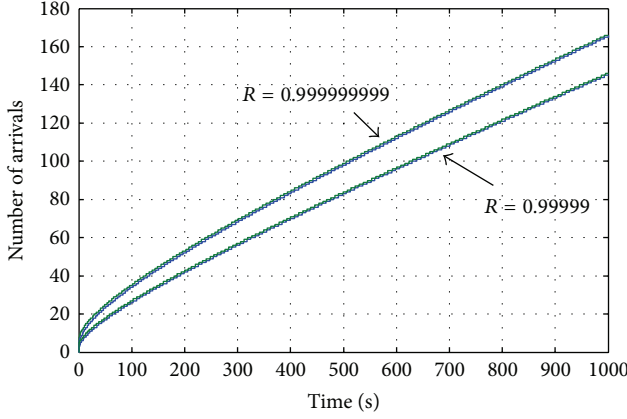


FIGURE 1: Comparison of the exact (blue) and bounded (green) number of arrivals versus time t , for a Poisson process having intensity $\lambda = 0.1$ for confidence probabilities $R = (1 - 10^{-5})$ and $R = (1 - 10^{-9})$.

4. Exponential and Logarithmic Inequalities

It is easy to verify the sharpness of the Poisson CDF universal inequality presented in Section 2 (the sharpness of the inequality for the Binomial distribution follows from the discussions in [5]), and its form makes it relatively easy to compute and implement; both the standard Normal CDF $\Phi(x)$ and its inverse $\Phi^{-1}(R)$ for the quantile inequalities can be quickly calculated to machine precision by simple rational approximations [3, 11]. It may still, however, be desired to have simpler bounds that have closed forms (e.g., it may be needed to algebraically manipulate a probability expression for the probabilistic analysis of an algorithm). Firstly, observe that the trivial inequality $1 - \exp(-x^2/2) \leq \Phi(x)$ for $x > 0$ can be used to recover the Chernoff exponential bound for the Poisson upper tail CDF by substitution of x with $\sqrt{2H(m, k)}$. The same inequality may be used to recover the bound of Hoeffding [7] by replacing x with $\sqrt{2nD(p, k/n)}$. This leads to consider the possibility of recovering tighter exponential bounds which retain simple forms for these quantities, employing known inequalities on the standard normal CDF. To achieve this, consider first the quantity known as Mills' ratio $M(x)$, which is defined for real arguments x in the usual way:

$$M(x) = \frac{1 - \Phi(x)}{\phi(x)}, \quad (25)$$

where $\phi(x) = e^{-0.5x^2}/\sqrt{2\pi}$ is the standard normal density function. For nonnegative x , it is known that $M(x)$ is monotonically decreasing in x and achieves a maximum of $\sqrt{0.5\pi}$ at $x = 0$. Various simple inequalities are known for this function; consider the simple upper bound $M(x) \leq x^{-1}$ for $x \geq 0$ proved by Gordon [12]. For small x , x^{-1} can become larger than $\sqrt{0.5\pi}$ and the basic bound can be improved by selecting the smaller of these two values:

$$M(x) \leq \min \left\{ \frac{1}{x}, \sqrt{\frac{\pi}{2}} \right\}. \quad (26)$$

Expression (26) may be employed to sharpen the Chernoff/Hoeffding bounds considerably via the inequalities of Theorems 1 and 2.

Corollary 6. *Let $Y \sim \text{Poi}(m)$ be a Poisson distributed random variable with mean m . For $k \geq m$, the following inequality holds:*

$$P(Y \leq k) > 1 - \frac{e^{-H(m, k)}}{\max \{2, \sqrt{4\pi H(m, k)}\}}. \quad (27)$$

Proof. Set $x = \sqrt{2H(m, k)}$. Then from Theorem 2 and (25) and (26) we can write the following:

$$\begin{aligned} P\{Y \leq k\} &> \Phi(x) \\ &= 1 - M(x)\phi(x) \\ &= 1 - M(x) \frac{e^{-0.5x^2}}{\sqrt{2\pi}} \\ &\geq 1 - \min \left\{ \frac{1}{x}, \sqrt{\frac{\pi}{2}} \right\} \cdot \frac{e^{-0.5x^2}}{\sqrt{2\pi}}. \end{aligned} \quad (28)$$

Inequality (27) results after some further simplification. \square

Corollary 7. *Let $X \sim \text{Bin}(n, p)$ be a Binomially distributed random variable with parameters p and n . For $k \geq np$, the following inequality holds:*

$$P(X \leq k) > 1 - \frac{e^{-nD(p, k/n)}}{\max \{2, \sqrt{4\pi nD(p, k/n)}\}}. \quad (29)$$

Proof. Set $x = \sqrt{2nD(p, k/n)}$. The result follows from Theorem 1 using the method of Corollary 6. \square

Expressions (28) and (29) are tighter than the corresponding Chernoff/Hoeffding bounds by a factor of at least 2, due to the presence of the additional denominator terms; in fact, they are much tighter for larger deviations as the denominators under the exponential become $\gg 2$. An illustration of the improvement that is obtained by adopting expression (29) over Hoeffding's original bound is given in Figure 2. The figure illustrates the improvement in the lower bound on the distribution function that is obtained for increasing values of the normalized argument $nD(p, k/n)$ in the range from 0 to 6.

Moving deeper into the tail, it is easy to verify that the improvement is continually increasing; this can also be illustrated by way of a simple example. Suppose $n = 1000$, $p = 0.1$, and $k = 150$. Then the true Binomial distribution function has a value of 0.999999723 and Hoeffding's inequality gives a lower bound of 0.999995143. Application of (29) gives a value of 0.999999608 which is an order of magnitude closer to the true Binomial function; the complimentary probability has been reduced by a factor of 12.399... , which is in fact the value of the function $\sqrt{4\pi nD(p, k/n)}$ at this point. Extension of (29) to cover the Poisson-Binomial distribution (i.e., a sum of n nonidentically distributed Bernoulli variables having a mean

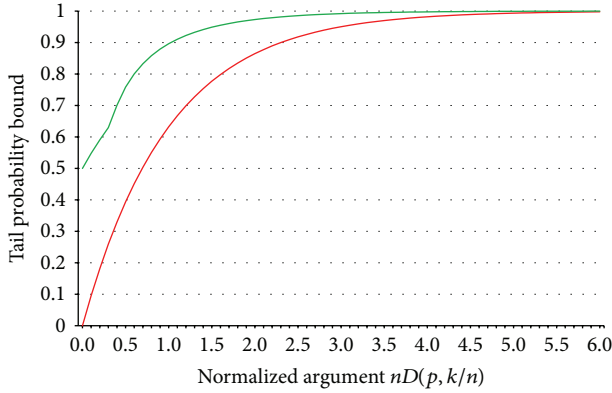


FIGURE 2: Comparison of the improvement of the distribution function lower bound in terms of the normalized argument $nD(p, k/n)$ between expression (29) (green) and Hoeffding's inequality (red).

$n\mu$) is obtained by employing (29) with $x = \sqrt{2nD(\mu, k/n)}$, as per Hoeffding [7]. Sharper inequalities for $M(x)$ can be used to improve these exponential bounds further, at the expense of increased complexity.

Finally, observe that simple Chernoff-style logarithmic quantile inequalities are obtained from (15) and (22) by using the known relationships $\Phi^{-1}(R) \leq \sqrt{-2 \cdot \ln(2 \cdot (1 - R))}$ for $R \in [0.5, 1)$ or the slightly sharper result of Chiani et al. [13]:

$$\Phi^{-1}(R) \leq \sqrt{2 \cdot \ln\left(\frac{1 + \sqrt{1 + 16 \cdot (1 - R)}}{8 \cdot (1 - R)}\right)}. \quad (30)$$

5. Conclusions

In this paper, some improved inequalities with a relatively simple form have been developed for the Poisson and Binomial distribution and quantile functions. Analysis and observations have helped to illustrate some improvements over previous work and related bounds. The obtained expressions should prove to be most useful in situations where provable and accurate bounds having analytic or closed forms are required and/or situations in which the use of commercial statistical software packages is not possible (see [8] for an example application arising in probabilistic real-time analysis). As a final remark, it seems that an interesting bound on the natural logarithm which to the authors' knowledge does not seem to have been previously described can also be obtained as a direct corollary of Lemma 3, which slightly sharpens similar bounds such as $\ln(x) \geq (2x - 2)/(x + 1)$ (see [14, page 160]).

Corollary 8. For real $x \geq 1$,

$$\ln(x) \geq 3 - \sqrt{\frac{27}{1 + 2x}}, \quad (31)$$

with equality occurring only for $x = 1$ and the sign of inequality reversed if $0 < x < 1$.

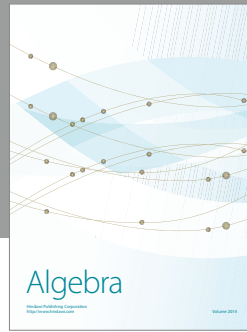
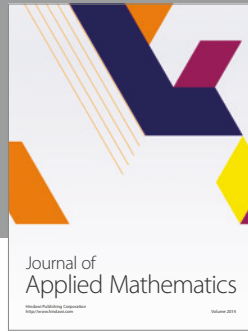
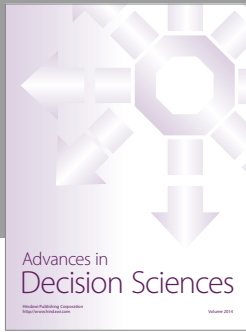
Proof. Replace k/m by x in (12); the corollary follows directly from Lemma 3 and some algebraic simplifications. \square

Conflict of Interests

The author declares that there is no conflict of interests regarding the publication of this paper.

References

- [1] N. L. Johnson, A. W. Kemp, and S. Kotz, *Univariate Discrete Distributions*, Wiley-Interscience, New York, NY, USA, 3rd edition, 2005.
- [2] N. Alon and A. J. Spencer, *Probabilistic Method*, John Wiley & Sons, New York, NY, USA, 2nd edition, 2000.
- [3] W. H. Press, S. A. Teukolsky, W. T. Vetterling, and B. P. Flannery, *Numerical Recipes in C: The Art of Scientific Computing*, Cambridge University Press, 1992.
- [4] W. Mollenaar, *Normal Approximations to the Poisson, Binomial and Hypergeometric Functions*, Tract 31, Amsterdam Mathematics Center, 1973.
- [5] A. M. Zubkov and A. A. Serov, "A complete proof of universal inequalities for distribution function of binomial law," *Teoriya Veroyatnostei i ee Primeneniya*, vol. 57, no. 3, pp. 597–602, 2012.
- [6] T. Hagerup and C. Rüb, "A guided tour of chernoff bounds," *Information Processing Letters*, vol. 33, no. 6, pp. 305–308, 1990.
- [7] W. Hoeffding, "Probability inequalities for sums of bounded random variables," *Journal of the American Statistical Association*, vol. 58, no. 301, pp. 13–30, 1963.
- [8] M. Short and J. Proenza, "Towards efficient probabilistic scheduling guarantees for real-time systems subject to random errors and random bursts of errors," in *Proceedings of the 25th Euromicro Conference on Real-Time Systems (ECRTS '13)*, pp. 259–268, Paris, France, July 2013.
- [9] Y. Li and L. Wang, "Testing for homogeneity in mixture using weighted relative entropy," *Communications in Statistics*, vol. 37, no. 10, pp. 1981–1995, 2008.
- [10] S. Janson, "Large deviation inequalities for sums of indicator variables," Tech. Rep., Department of Mathematics, Uppsala University, Uppsala, Sweden, 1994, <http://www.math.uu.se/~svante/papers/index.html>.
- [11] M. J. Wichura, "Algorithm AS 241: the percentage points of the normal distribution," *Applied Statistics*, vol. 37, no. 3, pp. 477–484, 1988.
- [12] R. D. Gordon, "Values of Mills' ratio of area bounding ordinate and of the normal probability integral for large values of the argument," *Annals of Mathematical Statistics*, vol. 12, no. 3, pp. 364–366, 1941.
- [13] M. Chiani, D. Dardari, and M. K. Simon, "New exponential bounds and approximations for the computation of error probability in fading channels," *IEEE Transactions on Wireless Communications*, vol. 2, no. 4, pp. 840–845, 2003.
- [14] P. S. Bullen, *A Dictionary of Inequalities*, Addison Wesley/Longman, 1998.



Hindawi

Submit your manuscripts at
<http://www.hindawi.com>

