Behavior of Binomial Distribution near Its Median

N. A. Volkov, D. I. Dmitriev, and M. E. Zhukovskii

Presented by Academician of the RAS V.V. Kozlov December 27, 2021

Received January 6, 2022; revised January 6, 2022; accepted February 10, 2022

Abstract—We study the behavior of the cumulative distribution function of a binomial random variable with parameters $n$ and $b/(n+c)$ at the point $b-1$ for positive integers $b \leq n$ and real $c \in [0,1]$. Our results can be applied directly to the well-known problem about small deviations of sums of independents random variables from their expectations. Moreover, we answer the question about the monotonicity of the Ramanujan function for the binomial distribution posed by Jogdeo and Samuels in 1968.

Keywords: binomial distribution, median, Ramanujan function, small deviations of sums of independent random variables

DOI: 10.1134/S1064562422020193

1. POISSON DISTRIBUTION AND RAMANUJAN’S PROBLEM

Let $b \in \mathbb{N}$ be a positive integer. We expand $e^b$ in a Taylor series: $e^b = \sum_{j=0}^{\infty} \frac{b^j}{j!}$. What is the smallest positive integer $\mu$ such that $\sum_{j=0}^{\mu} \frac{b^j}{j!} \geq \frac{1}{2} e^b$? The answer to this question is well known.

Let $\xi$ be a nonnegative integer random variable. The median of $\xi$ is the smallest nonnegative integer $\mu := \mu(\xi)$ such that $P(\xi \leq m) \geq \frac{1}{2}$. For a Poisson random variable $\eta_b$ with a positive integer parameter $b$, it is known that $\mu(\eta_b) = b$ [1], which answers the above question. How close is the probability $P(\eta_b \leq b)$ to $\frac{1}{2}$?

Ramanujan conjectured [2] that

$$y_b = \frac{1}{3} + \frac{4}{135(b + \alpha_b)},$$

where $8/45 \geq \alpha_b \geq 2/21$. This conjecture was proved in 1995 by Flajolet et al. [5]. In 2003, Alm [6] showed that $\alpha_b$ decreases, and, in 2004, Alzer [7] strengthened Ramanujan’s conjecture:

$$y_b = \frac{1}{3} + \frac{4}{135b} - \frac{8}{2835(b^2 + \beta_b)},$$

where

$$-\frac{1}{3} < \beta_b \leq -1 + \frac{4}{\sqrt{21(368 - 135e)}};$$

moreover, the indicated bounds are sharp.

2. BINOMIAL DISTRIBUTION AND SAMUELS’ PROBLEM

Let $\xi_{b,n}$ be a binomial random variable with parameters $n$ and $b/n$, where $b \leq n$ are positive integers. It is well known that $\xi_{b,n}$ converges in distribution to a random variable $\eta_b$ as $n \to \infty$. Accordingly, it is natural to expect that the properties of the Poisson distribution described in Section 1 hold for the binomial distribution for sufficiently large $n$. However, can the same questions be answered for all $n$?

Since $\mu(\xi_{b,n}) - b \leq \ln 2$ [8], the median of the binomial random variable is $\mu(\xi_{b,n}) = b$. In 1968, Jogdeo and Samuels [9] considered a quantity similar to

---

a Moscow Institute of Physics and Technology (National Research University), Dolgoprudny, Moscow oblast, Russia

b ETH Zurich and ETH AI Center, Zurich, Switzerland

e-mail: zhukmax@gmail.com
the one introduced by Ramanujan for the Poisson distribution, namely,
\[ z_{b,n} := \frac{1/2 - P(\xi_{b,n} < b)}{P(\xi_{b,n} = b)}. \]

**Theorem 1** (Jogdeo, Samuels, 1968 [9]). The quantity \( z_{b,n} \) decreases for \( n \geq 2b \), \( z_{b,n} \to y_n \) as \( n \to \infty \).

Moreover, \( \frac{1}{3} < z_{b,n} < \frac{2}{3} \) for all \( n > 2b \), \( \frac{1}{2} < z_{b,n} < \frac{2}{3} \) for \( b < n < 2b \), and \( z_{b,2b} = \frac{1}{2} = z_{b,b} \).

Additionally, it was noted in [9] that \( z_{b+1,n} < z_{b,n} \) for all sufficiently large \( n \), but the authors failed to improve this result.

Now we consider a binomial random variable \( \xi_{b,n,c} \) with parameters \( n \) and \( \frac{b}{n + c} \), where \( b < n \) are positive integers and \( c \in [0, 1] \). Define \( p_{b,n,c} := P(\xi_{b,n,c} < b) \).

The study of the monotonicity of \( p_{b,n,c} \) with respect to \( b \) is motivated by the well-known problem of small deviation inequality posed by Samuels [10], which can be formulated as follows: find the minimum of \( P(\xi_1 + \xi_2 + \ldots + \xi_n < n + c) \) over all sets of independent nonnegative random variables \( \{\xi_1, \ldots, \xi_n\} \) with an identical expectation equal to 1. This problem is still unsolved. Nevertheless, it is known that optimal random variables are quantities taking two values with probability 1 (i.e., with two atoms). If the consideration is restricted to identically distributed random variables with two atoms, then the original problem is reduced to analyzing the monotonicity of \( p_{b,n,c} \) with respect to \( b \).

The above-mentioned result of Szegö and Watson implies that \( P(\eta_b < b) \) increases. Since \( \lim_{n \to \infty} p_{b,n,c} = P(\eta_b < b) \), it follows that \( p_{b+1,n,c} > p_{b,n,c} \) for sufficiently large \( n \). On the other hand, for example, \( n = b + 1 \) and \( c = 0 \), we have \( 0 = p_{b+1,n,c} < p_{b,n,c} \). Thus, the monotonicity of \( p_{b,n,c} \) (regarded as a function of \( b \)) changes with increasing \( n \).

**3. NEW RESULTS**

We have been able to solve the problem posed by Jogdeo and Samuels concerning the monotonicity of \( z_{b,n} \) with respect to \( b \).

**Theorem 2.** Let \( \varepsilon > 0 \). Then there exists \( n_0 \) such that, for all \( n \geq n_0 \), the following assertions are true:

1. If \( n - (1 + \varepsilon) \frac{77}{360} n > b > (1 + \varepsilon) \frac{77}{360} n \), then \( z_{b+1,n} > z_{b,n} \).

2. If either \( b > n - (1 - \varepsilon) \frac{77}{360} n \) or \( b < (1 - \varepsilon) \frac{77}{360} n \), then \( z_{b+1,n} < z_{b,n} \).

Additionally, we have examined the function \( p_{b,n,c} \) for monotonicity with respect to \( b \).

**Theorem 3.** The following assertions hold:

1. If \( n \geq 3b + 2 \), then \( p_{b+1,n,0} > p_{b,n,0} \). If \( n \leq 3b + 1 \), then \( p_{b+1,n,0} < p_{b,n,0} \).

2. For all \( 1 \leq b < n \), it is true that \( p_{b+1,n,1} > p_{b,n,1} \).

3. If \( n \geq 3b + 2 \) and \( c \in (0, 1) \), then \( p_{b+1,n,c} > p_{b,n,c} \). Unfortunately, a complete result was obtained only for \( c = 0 \) and \( c = 1 \). Nevertheless, we found an asymptotic threshold after which monotonicity changes.

**Theorem 4.** For all positive \( \delta \) and \( \varepsilon \), sufficiently large \( n \), and integer \( b \in (en, n) \), the following assertions hold:

1. If \( b < \frac{n(1 - \delta)}{3(1 - c)} \), then \( p_{b+1,n} > p_{b,n} \).

2. If \( b > \frac{n(1 + \delta)}{3(1 - c)} \), then \( p_{b+1,n} < p_{b,n} \).

Note that Theorem 3 implies the following result for the small deviation problem with \( c = 1 \). Suppose that \( \alpha \in (0, 1) \) and \( \beta > 1 \). Let \( b \) be an integer such that \( b < n + 1 - \frac{n\alpha}{\beta} \leq b + 1 \). Then \( P(\xi_1 + \ldots + \xi_n < n + 1) \geq p_{b,n,1} \), where \( \xi_1, \ldots, \xi_n \) are independent identically distributed random variables with mean 1, and the equality holds if and only if \( \alpha = 0 \) and \( n + 1 = b + 1 \). Similar assertions might be stated for any \( c \in (0, 1) \) if the asymptotic result in Theorem 4 were proved to be sharp.

4. **B- FUNCTION**

Recall that \( B(x, y) = \int_0^1 t^{x-1}(1 - t)^{y-1} \). The proofs of the results stated in Section 3 are based on the following convenient expressions we derived for \( p_{b+1,n,c} - p_{b,n,c} \) and \( z_{b,n} \).

**Proposition 1.** For all positive integer \( b \leq n \) and \( c \in [0, 1] \),

\[
\begin{align*}
p_{b,n,c} &= \frac{1 - b}{\beta^{n+1}} \int_0^1 (1 - z)^{b-1} z^{n-b} dz \\
z_{b,n} &= \frac{1}{2} b \left[ \frac{1}{1-b/n} - \int_0^1 (1-z)^{b-1} z^{n-b} dz \right] \left( \frac{b}{n} \right)^b \left( 1 - \frac{b}{n} \right)^{n-b}.
\end{align*}
\]

DOKLADY MATHEMATICS Vol. 105 No. 2 2022
The analysis of these expressions is reduced to examining the behavior of the function $g(z) = (1 - z)^{b-1}z^{n-b}$ on $[1 - \frac{b+1}{n+c}, 1 - \frac{b}{n+c}]$. Its behavior can be studied using the Taylor formula (up to the fifth term) with a Lagrange remainder and the following convenient representation of the derivatives of $g$:

$$
\frac{\partial^\ell f}{\partial z^\ell} = (1 - z)^{b-\ell}z^{n-b-\ell} \\
\times \sum_{i=0}^{\ell} \binom{\ell}{i} (-1)^{\ell-i} z^{i-\ell} \frac{(n-1-i)!}{(n-1-\ell)!} \frac{(n-b)!}{(n-b-i)!}, \\
\ell \in \{1, ..., \min\{b-1, n-b\}\}.
$$

FUNDING

This work was supported by the Russian Science Foundation, project no. 21-71-10092.

CONFLICT OF INTEREST

The authors declare that they have no conflicts of interest.

REFERENCES

1. K. P. Choi, “On the medians of Gamma distributions and an equation of Ramanujan,” Proc. Am. Math. Soc. 121 (1), 245–251 (1994).
2. S. Ramanujan, “Question 294,” Indian Math. Soc. 3, 151–152 (1911).
3. G. Szegő, “Über einige von S. Ramanujan gestellte Aufgaben,” J. London Math. Soc. 3, 225–232 (1928).
4. G. N. Watson, “Theorems stated by Ramanujan (V): Approximations connected with $e^x$,” Proc. London Math. Soc. 29, 293–308 (1929).
5. P. Flajolet, P. J. Grabner, P. Kirschenhofer, and H. Prodinger, “On Ramanujan’s $Q$-function,” J. Comput. Appl. Math. 58, 103–116 (1995).
6. S. E. Alm, “Monotonicity of the difference between median and mean of gamma distributions and of a related Ramanujan sequence,” Bernoulli 9, 351–371 (2003).
7. H. Alzer, “On Ramanujan’s inequalities for $\exp(k)$,” J. London Math. Soc. 69, 639–656 (2004).
8. K. Hamza, “The smallest uniform upper bound on the distance between the mean and the median of the binomial and Poisson distributions,” Stat. Probab. Lett. 23, 21–25 (1995).
9. K. Jogdeo and S. M. Samuels, “Monotone convergence of binomial probabilities and a generalization of Ramanujan’s equation,” Ann. Math. Stat. 39 (4), 1191–1195 (1968).
10. S. M. Samuels, “On a Chebyshev-type inequality for sums of independent random variables,” Ann. Math. Stat. 37, 248–259 (1966).

Translated by I. Ruzanova