Study on Poisson distribution and geometric distribution motivated by Chvátal’s conjecture

Kun Xu, Fu-Bo Li, Ze-Chun Hu*

Deyang Foreign Languages School, Deyang 618099, China
2483680155@qq.com

College of Mathematics, Sichuan University, Chengdu 610065, China
lifubo@scu.edu.cn

College of Mathematics, Sichuan University, Chengdu 610065, China
zchu@scu.edu.cn

Abstract

Let $B(n, p)$ denote a binomial random variable with parameters $n$ and $p$. Chvátal’s conjecture says that for any fixed $n \geq 2$, as $m$ ranges over $\{0, \ldots, n\}$, the probability $q_m := P(B(n, m/n) \leq m)$ is the smallest when $m$ is closest to $\frac{2n}{3}$. Motivated by this conjecture, in this note, we consider the corresponding minimum value problem on the probability that a random variable is not more than its expectation, when its distribution is Poisson distribution or geometric distribution. We give a complete answer for each case. Some related questions will be introduced.

MSC: 60C05, 60E15

Keywords: Poisson distribution, Geometric distribution, Chvátal’s conjecture.

1 Introduction

Let $B(n, p)$ denote a binomial random variable with parameters $n$ and $p$. Janson in [4] introduced the following conjecture suggested by Vašk Chvátal in a personal communication.

*Corresponding author
Conjecture 1 (Chvátal). For any fixed $n \geq 2$, as $m$ ranges over $\{0, \ldots, n\}$, the probability $q_m := P(B(n, m/n) \leq m)$ is the smallest when $m$ is closest to $\frac{2n}{3}$.

As to the probability of a binomial random variable exceeding its expectation, we refer to Greenberg and Mohri [3], Pelekis and Ramon [5] and Doerr [2].

Janson [4] proved that Conjecture 1 holds for large $n$. Barabesi, Pratelli and Rigo [1] and Sun [6] gave an affirmative answer to Conjecture 1 by using different methods for general $n \geq 2$.

Motivated by Conjecture 1, we will consider the corresponding minimum value problem on the probability that a random variable is not more than its expectation, when its distribution is Poisson distribution or geometric distribution.

Let $X$ be a random variable which has Poisson distribution with parameter $\lambda (\lambda > 0)$ or the geometric distribution with parameter $p (0 < p \leq 1)$. In Section 2, we consider the minimum value of the probability $P(X \leq \lambda)$ for $\lambda \in (0, \infty)$ when $X$ has Poisson distribution with parameter $\lambda$. In Section 3, we consider the minimum value of the probability $P(X \leq \frac{1}{p})$ for $p \in (0, 1]$ when $X$ has the geometric distribution with parameter $p$. In the final section, we propose some related questions.

2 Poisson distribution

Let $X$ be a random variable which has Poisson distribution with parameter $\lambda (\lambda > 0)$. Then its distribution can be expressed by

$$P(X = k) = \frac{\lambda^k e^{-\lambda}}{k!}, \quad k = 0, 1, 2, \ldots.$$ 

We know that its expectation $EX = \lambda$. Then

$$P(X \leq EX) = \sum_{k=0}^{[\lambda]} \frac{\lambda^k e^{-\lambda}}{k!} = \left(1 + \lambda + \frac{\lambda^2}{2!} + \cdots + \frac{\lambda^{[\lambda]}}{[\lambda]!}\right) e^{-\lambda}.$$ 

Hereafter, for a real number $a$, $[a]$ stands for the biggest integer which is not more than $a$.

Define a function

$$f(\lambda) := \left(1 + \lambda + \frac{\lambda^2}{2!} + \cdots + \frac{\lambda^{[\lambda]}}{[\lambda]!}\right) e^{-\lambda}, \quad \lambda > 0. \quad (2.1)$$

In the following, we will consider the minimum value of $f(\lambda)$ on the interval $(0, \infty)$, and the main result is

Proposition 2.1 The function $f(\lambda)$ has no minimum value on $(0, \infty)$, but

$$\inf_{\lambda \in (0, \infty)} f(\lambda) = \lim_{\lambda \uparrow 1} f(\lambda) = e^{-1}. \quad (2.2)$$
Proof. For $\lambda \in (0, 1)$, we have

$$f(\lambda) = e^{-\lambda}.$$ 

It follows that

$$\inf_{\lambda \in (0, 1)} f(\lambda) = \lim_{\lambda \uparrow 1} e^{-\lambda} = e^{-1}.$$ 

Let $x$ be a positive integer. For $\lambda \in [x, x+1)$, we have

$$f(\lambda) = \left(1 + \lambda + \frac{\lambda^2}{2!} + \cdots + \frac{\lambda^x}{x!}\right) e^{-\lambda}.$$ 

It follows that for any $\lambda \in (x, x+1)$,

$$f'(\lambda) = -\frac{\lambda^x}{x!} e^{-\lambda} < 0,$$

which implies that the function $f(\lambda)$ is strictly decreasing on the interval $[x, x+1)$. Hence we have

$$\inf_{\lambda \in [x,x+1)} f(\lambda) = \lim_{\lambda \uparrow x+1} f(\lambda) = \left(1 + (x+1) + \frac{(x+1)^2}{2!} + \cdots + \frac{(x+1)^x}{x!}\right) e^{-(x+1)}.$$ 

Define a sequence $\{b_n\}$ as follows:

$$b_n := \begin{cases} e^{-1}, & \text{if } n = 0, \\ \left(1 + (n+1) + \frac{(n+1)^2}{2!} + \cdots + \frac{(n+1)^n}{n!}\right) e^{-(n+1)}, & \text{if } n \geq 1. \end{cases} \quad (2.3)$$

In the following, we will analyze the minimum value of the sequence $\{b_n\}_{n \geq 0}$.

We know that the exponential function $e^x$ possesses the following Taylor’s formula:

$$e^x = 1 + x + \frac{x^2}{2!} + \cdots + \frac{x^n}{n!} + R_n(x), \quad \forall x \in (-\infty, +\infty),$$

where $R_n(x)$ is the remainder term.

By the Taylor’s formula above, we get that for any $n \geq 1$,

$$b_n = \frac{e^{n+1} - R_n(n+1)}{e^{n+1}} = 1 - \frac{R_n(n+1)}{e^{n+1}}. \quad (2.4)$$

By the integral remainder term, we have

$$R_n(n+1) = \frac{(-1)^n}{n!} \int_0^{n+1} (t-(n+1))^n e^t \, dt. \quad (2.5)$$

By the discussing for the parity of $n$ and change of variable, we get

$$R_n(n+1) = \frac{1}{n!} \int_0^{n+1} (n+1 - t)^n e^t \, dt = \frac{1}{n!} \int_0^{n+1} u^n e^{(n+1)-u} \, du. \quad (2.6)$$
It follows that
\[
\frac{R_n(n+1)}{e^{n+1}} = \frac{1}{n!} \int_0^{n+1} u^n e^{-u} du.
\]  

(2.7)

Define a sequence \( \{c_n\}_{n \geq 1} \) as follows:
\[
c_n := \frac{1}{n!} \int_0^{n+1} u^n e^{-u} du, \quad \forall n \geq 1.
\]  

(2.8)

Now, we come to analyze the maximum value of the sequence \( \{c_n\}_{n \geq 1} \).

For any \( n \geq 2 \), using integration by parts, we get
\[
c_n - c_{n-1} = \frac{1}{n!} \int_0^{n+1} u^n e^{-u} du - \frac{1}{(n-1)!} \int_0^n u^{n-1} e^{-u} du
\]
\[
= \frac{1}{n!} \left( -u^n e^{-u} \bigg|_0^{n+1} + \int_0^{n+1} nu^{n-1} e^{-u} du \right) - \frac{1}{(n-1)!} \int_0^n u^{n-1} e^{-u} du
\]
\[
= \frac{1}{(n-1)!} \int_n^{n+1} u^{n-1} e^{-u} du - \frac{(n+1)^n}{n! e^{n+1}}.
\]

It follows that
\[
n!(c_n - c_{n-1}) = n \int_n^{n+1} u^{n-1} e^{-u} du - (n+1)^n e^{-(n+1)}
\]
\[
:= I - J,
\]

where
\[
I = n \int_n^{n+1} u^{n-1} e^{-u} du = n \int_0^1 (n+1-x)^{n-1} e^{-(n+1-x)} dx,
\]  

(2.9)

\[
J = (n+1)^n e^{-(n+1)}.
\]  

(2.10)

In the following, we will show that \( \frac{I}{J} < 1 \) for \( n \geq 5 \).

By (2.9) and (2.10), we have
\[
\frac{I}{J} = \frac{n}{n+1} \int_0^1 \left( 1 - \frac{x}{n+1} \right)^{n-1} e^x dx.
\]

Define a function
\[
g(x) := \frac{n}{n+1} \left( 1 - \frac{x}{n+1} \right)^{n-1} e^x, \quad x \in [0, 1].
\]  

(2.11)

Then \( \frac{I}{J} = \int_0^1 g(x) dx \). We will show that for \( n \geq 5 \),
\[
\int_0^1 g(x) dx < 1.
\]  

(2.12)
By (2.11), we have
\[ g'(x) = \frac{n}{n+1} \left( 1 - \frac{x}{n+1} \right)^{n-1} \cdot e^x \cdot \frac{2-x}{n+1-x} > 0, \quad \forall x \in [0,1], \]
\[ g''(x) = g'(x) \cdot \frac{x^2 - 4x + 5 - n}{(n+1-x)(2-x)} < 0, \quad \forall x \in (0,1), \quad \forall n \geq 5. \]

It follows that \( g(x) \) is a strictly increasing concave function on \([0,1]\) for \( n \geq 5 \).

By (2.11), we have
\[ g(0) = \frac{n}{n+1} < 1, \]
\[ g(1) = \frac{e}{(1+\frac{1}{n})^n} > 1, \]
\[ g \left( \frac{1}{2} \right) = \left( \frac{e}{(1+\frac{1}{n})^2(1+\frac{1}{2n+1})^{2(n-1)}} \right)^{1/2} < 1, \]

where in the third inequality, we used that for any \( n \geq 2, \)
\[ \left( 1 + \frac{1}{n} \right)^2 \left( 1 + \frac{1}{2n+1} \right)^{2(n-1)} > \left( 1 + \frac{1}{2n+1} \right)^{2n+2} > e. \]

Then we know that there is a unique \( a \in (0,1) \) such that \( g(a) = 1 \) and \( a > 1/2 \). As to the curve associated with the function \( g(x) \) on \([0,1]\), we draw a tangent line at \((a,g(a))\). The corresponding equation is
\[ y = g'(a)x + (g(a) - g'(a)a). \]

Then we have Figure 1 on the sketch map of the function \( g(x) \) for \( n \geq 5 \) on next page.

Denote by \( D_1 \) the domain bounded by \( x = 0, y = 1, y = g(x) \), by \( D_2 \) the domain bounded by \( x = 0, y = 1, y = g(x) \), by \( D_3 \) the domain bounded by \( y = 1, x = 1 \) and the tangent line \( y = g'(a)x + (g(a) - g'(a)a) \), by \( D_4 \) the domain bounded by \( y = 1, x = 1, y = g(x) \), and by \( D_5 \) the domain bounded by \( x = 0, x = 1, y = 0, y = g(x) \).

For \( i = 1, \ldots, 5, \) we denote by \( S(D_i) \) the area of the domain \( D_i \). Then by Figure 1 below, we know that
\[ S(D_1) > S(D_2) > S(D_3) > S(D_4), \]
which implies that
\[ S(D_5) = 1 - S(D_1) + S(D_4) < 1, \]
i.e.
\[ \int_0^1 g(x)dx < 1. \]

Hence \( I < J \) and thus \( c_n - c_{n-1} < 0 \) for \( n \geq 5 \). So the sequence \( \{c_n\}_{n \geq 4} \) is strictly decreasing.
Figure 1: Sketch map of the function $g(x)$ on $[0, 1]$ for $n \geq 5$.

By (2.8), we get that

$$C_1 = \int_0^2 u e^{-u} du = 1 - \frac{3}{e^2},$$

$$C_2 = \frac{1}{2!} \int_0^3 u^2 e^{-u} du = 1 - \frac{17}{2e^3},$$

$$C_3 = \frac{1}{3!} \int_0^4 u^3 e^{-u} du = 1 - \frac{71}{3e^4},$$

$$C_4 = \frac{1}{4!} \int_0^5 u^4 e^{-u} du = 1 - \frac{523}{8e^5}.$$

By the fact that $e = 2.71 \cdots < 2.72$, we can easily check that

$$\frac{3}{e^2} < \frac{17}{2e^3} < \frac{71}{3e^4} < \frac{523}{8e^5},$$

which implies that $C_1 > C_2 > C_3 > C_4$. Hence the sequence $\{c_n\}_{n \geq 1}$ is strictly decreasing. By (2.4), (2.7) and (2.8), we know that the sequence $\{b_n\}_{n \geq 1}$ is strictly increasing. It follows that

$$\min_{n \geq 1} b_n = b_1 = 3e^{-2}.$$

By the fact that $e = 2.71 \cdots < 2.72$ again, we get that $3e^{-2} > e^{-1} = b_0$.

By the analysis above, we know that the function $f(\lambda)$ has no minimum value on $(0, \infty)$, but

$$\inf_{\lambda \in (0, \infty)} f(\lambda) = b_0 = e^{-1} = \lim_{\lambda \uparrow 1} f(\lambda).$$

The proof is complete. \qed
3 Geometric distribution

Let $X$ be a random variable which has the geometric distribution with parameter $p(0 < p \leq 1)$. Now its distribution can be expressed by

$$P(X = k) = (1 - p)^{k-1}p, \ k = 1, 2, 3, \ldots.$$  

We know that the expectation of $X$ is $EX = 1/p$. Then we have

$$P(X \leq EX) = \sum_{k=1}^{\lfloor 1/p \rfloor} p(1 - p)^{k-1} = (p + p(1 - p) + \cdots + p(1 - p)^{\lfloor 1/p \rfloor - 1}) = 1 - (1 - p)^{\lfloor 1/p \rfloor}.$$  

Define

$$f(p) := 1 - (1 - p)^{\lfloor 1/p \rfloor}, \quad 0 < p \leq 1. \quad \text{(3.1)}$$

In the following, we will analyze the minimum value of $f(p)$ on the interval $(0, 1]$ and the main result is

**Proposition 3.1** The function $f(p)$ has no minimum value on $(0, 1]$, but

$$\inf_{p \in (0, 1]} f(p) = \lim_{p \downarrow \frac{1}{x+1}} f(p) = \frac{1}{2}. \quad \text{(3.2)}$$

**Proof.** Let $x$ be a positive integer. For any $1/p \in [x, x + 1)$, we have

$$f(p) = 1 - (1 - p)^x.$$  

Then for $p \in (\frac{1}{x+1}, \frac{1}{x})$,

$$f'(p) = x (1 - p)^{x-1} > 0,$$

which implies that the function $f(p)$ is strictly increasing on the interval $(\frac{1}{x+1}, \frac{1}{x}]$. Thus we have

$$\inf_{p \in (\frac{1}{x+1}, \frac{1}{x}]} f(p) = \lim_{p \downarrow \frac{1}{x+1}} f(p) = 1 - \left(1 - \frac{1}{x+1}\right)^x.$$  

Define a sequence $\{a_n\}$ as follows:

$$a_n := 1 - \left(1 - \frac{1}{n+1}\right)^n, \quad \forall n \geq 1. \quad \text{(3.3)}$$

In the following, we will consider the minimum value of the sequence $\{a_n\}_{n \geq 1}$.
Define a function
\[ g(x) := 1 - \left( 1 - \frac{1}{x+1} \right)^x, \quad \forall x \geq 1. \] (3.4)

We can rewrite \( g(x) \) by
\[ g(x) = 1 - e^{x\ln(1 - \frac{1}{x+1})}, \quad \forall x \geq 1. \]

It follows that
\[ g'(x) = -e^{x\ln(1 - \frac{1}{x+1})} \left[ \ln \left( 1 - \frac{1}{x+1} \right) + \frac{1}{x+1} \right]. \] (3.5)

Define
\[ h(y) := \ln(1 - y) + y, \quad \forall y \in [0, 1/2]. \] (3.6)

Then
\[ h'(y) = 1 - \frac{1}{1-y} = -\frac{y}{1-y} < 0, \quad \forall y \in (0, 1/2]. \] (3.7)

It follows that for any \( y \in (0, 1/2] \),
\[ h(y) < h(0) = 0, \]
which implies that for any \( x \geq 1 \),
\[ g'(x) > 0, \]
and thus \( g(x) \) is strictly increasing on \([1, +\infty)\). And so \( g(x) \) has a unique minimum value point \( x = 1 \) on \([1, \infty)\). Hence the sequence \( \{a_n\}_{n \geq 1} \) reaches the minimum value at \( n = 1 \).

By the analysis above, we know that \( f(p) \) has no minimum value on \((0, 1]\), but its infimum satisfies
\[ \inf_{p \in (0, 1]} f(p) = \lim_{p \downarrow \frac{1}{2}} f(p) = a_1 = \frac{1}{2}. \]

Hence (3.2) holds. The proof is complete. \( \square \)

## 4 Questions

We know that Pascal distribution includes geometric distribution as a special case, and negative binomial distribution includes Pascal distribution as a special case. Two natural questions arise:

**Question 1.** If \( X \) is a random variable satisfying Pascal distribution with parameter \( r \) (\( r \) is a positive integer) and \( p(p \in (0, 1]) \), how about the minimum value of the probability \( P(X \leq EX) \)?

**Question 2.** If \( X \) is a random variable satisfying the negative binomial distribution with parameter \( r \) (\( r > 0 \)) and \( p(p \in (0, 1]) \), how about the minimum value of the probability \( P(X \leq EX) \)?
Of course, one can consider the corresponding minimum value problems on other distributions, such as hypergeometric distribution, $\chi^2$ distribution, Gamma distribution, student $t$ distribution, Pareto distribution, $F$ distribution, $\beta$ distribution, log-normal distribution etc.

In addition, Chvátal’s conjecture only concerns the minimum value of the probability $P(B(n, m/n) \leq m)$ as $m$ ranges over $\{0, \ldots, n\}$. We can ask the following question:

**Question 3.** For any fixed integer $n \geq 2$, how about the minimum value of the probability $P(B(n, p) \leq np)$ for $p \in (0, 1]$?

**Acknowledgments** We’d like to thank Xuesong Li and Jingzi Yan for the discussing on the proof of Proposition 2.1. This work was supported by the National Natural Science Foundation of China (12171335) and the Science Development Project of Sichuan University (2020SCUNL201).

**References**

[1] L. Bababesi, L. Pratelli, P. Rigo, On the Chvátal-Janson conjecture, arXiv: 2104.11971v1, 2021.

[2] B. Doerr, An elementary analysis of the probability that a binomial random variable exceeds its expectation, Statis. Probab. Lett. 139 (2018) 67-74.

[3] S. Greenberg, M. Mohri, Tight lower bound on the probability of a binomial exceeding its expectation, Statis. Probab. Lett. 86 (2014) 91-98.

[4] S. Janson, On the probability that a binomial variable is at most its expectation, Statis. Probab. Lett. 171 (2021) 109020.

[5] C. Pelekis, J. Ramon, A lower bound on the probability that a binomial random variable is exceeding its mean, Statis. Probab. Lett. 119 (2016) 305-309.

[6] P. Sun, Strictly unimodality of the probability that the binomial distribution is more than its expectation, Discrete Appl. Math. 301 (2021) 1-5.