

## ON THE MEDIANS OF GAMMA DISTRIBUTIONS AND AN EQUATION OF RAMANUJAN

K. P. CHOI

(Communicated by Wei Y. Loh)

**ABSTRACT.** For  $n \geq 0$ , let  $\lambda(n)$  denote the median of the  $\Gamma(n+1, 1)$  distribution. We prove that  $n + \frac{2}{3} < \lambda(n) \leq \min(n + \log 2, n + \frac{2}{3} + (2n+2)^{-1})$ . These bounds are sharp. There is an intimate relationship between  $\lambda(n)$  and an equation of Ramanujan. Based on this relationship, we derive the asymptotic expansion of  $\lambda(n)$  as follows:

$$\lambda(n) = n + \frac{2}{3} + \frac{8}{405n} - \frac{64}{5103n^2} + \frac{2^7 \cdot 23}{3^9 \cdot 5^2 n^3} + \cdots$$

Let  $\text{median}(Z_\mu)$  denote the median of a Poisson random variable with mean  $\mu$ , where the median is defined to be the least integer  $m$  such that  $P(Z_\mu \leq m) \geq \frac{1}{2}$ . We show that the bounds on  $\lambda(n)$  imply

$$\mu - \log 2 \leq \text{median}(Z_\mu) < \mu + \frac{1}{3}.$$

This proves a conjecture of Chen and Rubin. These inequalities are sharp.

### 1. INTRODUCTION AND STATEMENT OF MAIN RESULTS

For a continuous random variable  $X$ , the median is defined to be the least  $x \in (-\infty, \infty)$  such that  $P(X \leq x) = \frac{1}{2}$ ; and if  $X$  is an integer-valued random variable, the median of  $X$  is defined to be the least integer  $m$  such that  $P(X \leq m) \geq \frac{1}{2}$ . In either case, we denote the median of  $X$  by  $\text{median}(X)$ . For  $n \geq 0$ , let  $\lambda(n)$  denote the median of the  $\Gamma(n+1, 1)$  distribution, the Gamma distribution with parameters  $n+1$  and 1. In other words,

$$(1) \quad \frac{1}{n!} \int_0^{\lambda(n)} t^n e^{-t} dt = \frac{1}{n!} \int_{\lambda(n)}^\infty t^n e^{-t} dt = \frac{1}{2}.$$

One of the main results in this paper is

**Theorem 1.** For  $n \geq 0$ ,

$$(2) \quad n + \frac{2}{3} < \lambda(n) \leq \min(n + \log 2, n + \frac{2}{3} + (2n+2)^{-1}).$$

*These bounds are sharp.*

---

Received by the editors August 14, 1992.

1991 *Mathematics Subject Classification.* Primary 62E10, 41A58, 33B15.

*Key words and phrases.* Median, Gamma distribution, Poisson distribution, chi-square distribution, Poisson-Gamma relation, Ramanujan's equation.

In this paper, we also consider the median of the Poisson distribution. Let  $Z_\mu$  denote a Poisson random variable with mean  $\mu$ . Then it will be shown in the next section that Theorem 1 implies the following

**Theorem 2.** *Let  $\mu \in (0, \infty)$ . Then*

$$\mu - \log 2 \leq \text{median}(Z_\mu) < \mu + \frac{1}{3}.$$

*These bounds are best possible.*

This proves Conjecture 1a of Chen and Rubin [2]. Based on a numerical study, Chen and Rubin were led to a stronger conjecture, which states that  $\lambda(n) - n$  is decreasing in  $n$ . This still remains open. At the end of this section, we see that the coefficients of the first four terms of the asymptotic expansion of  $\lambda(n - 1) - (n - 1)$  are positive. This strongly suggests that their conjecture is very likely to hold.

The work of this paper is also motivated by its connection with a well-known equation of Ramanujan. Ramanujan [7] and also his first letter to Hardy on January 16, 1913 contained the following statement: For  $n \geq 1$ , let  $\theta(n)$  be defined as

$$(3) \quad \frac{e^n}{2} = \sum_{k=0}^{n-1} \frac{n^k}{k!} + \theta(n) \frac{n^n}{n!};$$

then  $\frac{1}{3} \leq \theta(n) \leq \frac{1}{2}$ . He also gave the first four terms of the asymptotic expansion of  $\theta(n)$ :

$$(4) \quad \theta(n) = \frac{1}{3} + \frac{4}{135n} - \frac{8}{2835n^2} - \frac{16}{8505n^3} + O\left(\frac{1}{n^4}\right).$$

This problem has attracted a lot of attention. For example, Szegő [9] and Watson [10] proved that  $\theta(n)$  lies between  $\frac{1}{3}$  and  $\frac{1}{2}$ , whereas further coefficients of the asymptotic expansion have been found by Bowman, Shenton, and Szekeres [1] and Marsaglia [6]. Knuth considered the asymptotic expansion of  $P(Z_n \leq n - 1)/P(Z_n = n)$  from which he gave another derivation of the asymptotic expansion of  $\theta(n)$ , though with fewer terms (see Knuth [5, pp. 112–118]). We find that there is an intimate relationship between  $\lambda(n)$  and  $\theta(n)$  (see the first paragraph of §3); namely,

$$(5) \quad 1 - \theta(n) = \int_n^{\lambda(n)} t^{n-1} e^{-t} / n^{n-1} e^{-n} dt.$$

This will enable us to deduce bounds on  $\theta(n)$  based on Theorem 1 and, conversely, asymptotic expansion of  $\lambda(n)$  based on that of  $\theta(n)$ .

**Theorem 3.** *With  $\theta(n)$  defined as above and  $n \geq 1$ , we have*

$$\frac{1}{3} - \frac{73}{162n} + \frac{19}{243n^2} < \theta(n) < \frac{1}{3} + \frac{4}{81n} - \frac{4}{243n^3}.$$

The upper bound in Theorem 3 is a decreasing function in  $n$ , and hence  $\theta(n) < \frac{1}{3} + \frac{8}{243} = 0.36625\dots$ , which is an improvement of the upper bound  $\frac{1}{2}$ .

**Theorem 4.** *With  $\lambda(n)$  defined as above, we have*

$$(6) \quad \lambda(n) = n + \frac{2}{3} + \frac{8}{405n} - \frac{64}{5103n^2} + \frac{2^7 \cdot 23}{3^9 \cdot 5^2 n^3} + O\left(\frac{1}{n^4}\right).$$

The first three terms of the asymptotic expansion can also be derived from Dinges [3, Proposition 3.6]. Our method here is different, and with more work many more terms can be derived. Doodson [4] considered

$$\frac{\text{median}(X_n) - \text{mode}(X_n)}{\text{mean}(X_n) - \text{mode}(X_n)} = \text{median}(X_n) - (n - 1),$$

where  $X_n$  denotes a Gamma random variable with parameters  $n$  and 1. Using our notation and results, this ratio is

$$\lambda(n - 1) - (n - 1) = \frac{2}{3} + \frac{8}{405n} + \frac{144}{3^6 \cdot 5 \cdot 7n^2} + \frac{2^3 \cdot 281}{3^9 \cdot 5^2 \cdot 7n^3} + \dots$$

## 2. PROOFS OF THEOREMS 1 AND 2

*Proof of Theorem 1.* Let  $X_{n+1}$  be  $\Gamma(n + 1, 1)$  distributed. For  $x \in [\frac{2}{3}, 1]$ ,  $n \geq 0$ , let  $a_n(x) = P(X_{n+1} \geq n + x)$  and  $A_n(x) = (n + 1)![a_{n+1}(x) - a_n(x)]$ . For  $n \geq 0$  and  $t > 0$ , let  $\psi_n(t) = t^{n+1}e^{-t}$ . Therefore,

$$\begin{aligned} A_n(x) &= \int_{n+x+1}^{\infty} t^{n+1}e^{-t} dt - (n + 1) \int_{n+x}^{\infty} t^n e^{-t} dt \\ &= (n + x)^{n+1}e^{-(n+x)} - \int_{n+x}^{n+x+1} t^{n+1}e^{-t} dt \\ &= \int_{n+x}^{n+x+1} [\psi_n(n + x) - \psi_n(t)] dt. \end{aligned}$$

We shall show at the end of the proof that

$$(7) \quad A_n(\log 2) > 0, \quad n \geq 2;$$

$$(8) \quad A_n\left(\frac{2}{3} + (2n + 2)^{-1}\right) > 0, \quad n \geq 1;$$

$$(9) \quad A_n\left(\frac{2}{3}\right) < 0, \quad n \geq 1.$$

By the Central Limit Theorem, we have  $a_n(x) \rightarrow \frac{1}{2}$  as  $n \rightarrow \infty$ . From (9) we have  $a_n(\frac{2}{3})$  is a strictly decreasing sequence; therefore,  $a_n(\frac{2}{3}) > \frac{1}{2}$ . Hence,  $\lambda(n) > n + \frac{2}{3}$ ,  $n \geq 1$ . Direct computation shows that  $\lambda(0) = \log 2 > \frac{2}{3}$ , establishing inequality (2) in the case  $n = 0$ . The second inequality is proved in the same way. To prove (7)–(9), we write  $A_n(x) = B_n(x) + E_n(x)$ , where  $B_n(x)$  is due to the Taylor expansion of  $\psi_n(x)$  at  $n + 1$  up to the third derivative and

with  $E_n(x)$  the error term. It is not difficult to derive that

$$B_n(x) = \left\{ \frac{(3x - 2)}{6(n + 1)} - \frac{6x^2 - 8x + 3}{12(n + 1)^2} \right\} \psi_n(n + 1),$$

and there exist  $\xi \in [n + x, n + 1]$  and  $\zeta(t)$  lying between  $n + 1$  and  $t$  such that

$$E_n(x) = \frac{(1 - x)^4}{4!} \psi_n^{(4)}(\xi) - \int_{n+x}^{n+x+1} \frac{(t - n - 1)^4}{4!} \psi_n^{(4)}(\zeta(t)) dt.$$

Write  $\psi_n^{(k)}(t) = w_{k,n}(t)\psi_n(t)/t^k$  and  $u_{k,n}(s) = w_{k,n}(n + 1 + s)$  for  $s \in [-(1 - x), x] \subseteq [-1, 1]$ . It is straightforward to verify that  $u_{4,n}(s) = s^4 + (n + 1)[3(n - 1) - 8s - 6s^2]$ , it is decreasing in  $s$ , and  $u_{4,n}(s) \geq u_{4,n}(1) > 0$  when  $n \geq 6$ . For  $n \geq 6$ , this implies that  $\psi_n^{(4)}(\zeta(t)) > 0$ ; hence,  $E_n(x)$  is bounded above by

$$\frac{(1 - x)^4 u_{4,n}(-1)}{24(n + x)^4} \psi_n(n + 1).$$

Therefore,  $A_n(\frac{2}{3}) < B_n(\frac{2}{3}) + \psi_n(n + 1)/[648(n + 1)^2] < 0$ , proving (9) for  $n \geq 6$ . Remaining cases can be verified by direct calculation.

Next, for  $n \geq 6$ ,  $E_n(x)$  is bounded below by

$$\frac{-[x^5 + (1 - x)^5]u_{4,n}(-1)}{120(n + x)^4} \psi_n(n + 1).$$

Since the leading term of  $A_n(\log 2)$  is  $(3 \log 2 - 2)/[6(n + 1)]$ , it is apparent that  $A_n(\log 2) > 0$  for  $n$  sufficiently large. Indeed, it can be verified to be positive for  $n \geq 6$ . Also, we have  $B_n(\frac{2}{3} + \frac{1}{2(n+1)}) = (\frac{2}{9(n+1)^2} - \frac{1}{8(n+1)^4})\psi_n(n + 1)$  and  $E_n(\frac{2}{3} + \frac{1}{2(n+1)}) > \frac{-3}{40(n+1)^2} \psi_n(n + 1)$ . Hence,  $A_n(\frac{2}{3} + \frac{1}{2(n+1)}) > 0$ . Remaining cases can be verified by direct calculation. Since  $\lambda(0) = \log 2$  and  $\lambda(n) - n \rightarrow \frac{2}{3}$  as  $n \rightarrow \infty$ , the bounds are sharp. This completes the proof of Theorem 1.  $\square$

*Remark 1.* The lower bound in (2) and hence the upper bound in Theorem 2 were proved by an ingenious and yet elementary method in Chen and Rubin [2].

*Proof of Theorem 2.* Making use of the Poisson-Gamma relation, which states that for  $n \geq 0, \mu > 0$

$$(10) \quad P(Z_\mu \leq n) = P(X_{n+1} \geq \mu),$$

we see that  $P(Z_{\lambda(n)} \leq n) = \frac{1}{2}$  and, for  $\lambda(n) < \mu \leq \lambda(n + 1)$ ,  $\text{median}(Z_\mu) = n + 1$ . Since  $\lambda(n + 1) - \lambda(n) > n + 1 + \frac{2}{3} - (n + \log 2) > 0.97$ , for any  $\mu \in (\log 2, \infty)$ , there exists  $n \geq 0$  such that  $\lambda(n) < \mu \leq \lambda(n + 1)$ . Therefore,

$$\text{median}(Z_\mu) - \mu = n + 1 - \mu \geq n + 1 - \lambda(n + 1) \geq -\log 2$$

and

$$\text{median}(Z_\mu) - \mu = n + 1 - \mu < n + 1 - \lambda(n) < \frac{1}{3}.$$

For  $\mu \in (0, \log 2]$ ,  $\text{median}(Z_\mu) = 0$ , so  $-\log 2 \leq \text{median}(Z_\mu)$  with equality holds when  $\mu = \log 2$ . Let  $\mu = \lambda(n) + \epsilon$ , where  $\epsilon$  is an arbitrarily small positive number. Then  $\text{median}(Z_\mu) - \mu = n + 1 - \lambda(n) - \epsilon = \frac{1}{3} - \epsilon + O(n^{-1})$ .

This shows that the second inequality is sharp, and this completes the proof of Theorem 2.  $\square$

### 3. PROOFS OF THEOREMS 3 AND 4

It is obvious that equation (4) is equivalent to

$$P(Z_n \leq n) - [1 - \theta(n)]P(Z_n = n) = \frac{1}{2}.$$

Recalling the Poisson-Gamma relation (see equation (10)) and the definitions of  $\psi_n$  and  $\lambda(n)$ , we obtain (5).

*Proof of Theorem 3.* Recall the definition of  $\psi_n(t)$ . Then

$$1 - \theta(n) < \int_n^{n+2/3+(2n)^{-1}} \psi_{n-1}(t)/\psi_{n-1}(n) dt.$$

Expanding  $\psi_{n-1}(t)$  at  $t = n$  up to the second derivative and estimating the error term as in the proof of Theorem 1, we can show that  $\psi_{n-1}(t)/\psi_{n-1}(n)$  is bounded above by

$$\begin{aligned} 1 - \frac{(t-n)^2}{2n} + \frac{(t-n)^3}{6n^3} \max_{0 \leq s \leq 2/3+(2n)^{-1}} ((2+3s)n - s^3) \\ \leq 1 - \frac{(t-n)^2}{2n} + \frac{(t-n)^3}{6n^3} \left(4n + \frac{3}{2}\right), \end{aligned}$$

from which we obtain the lower bound for  $\theta(n)$ . As shown in the proof of Theorem 1, for  $n \geq 6$ ,  $\psi_{n-1}^{(4)}(t) > 0$  on  $[n, n + \frac{2}{3}]$ . Therefore,  $\psi_{n-1}'''(\xi) \geq \psi_{n-1}'''(n)$ . It now follows that

$$\frac{\psi_{n-1}(t)}{\psi_{n-1}(n)} \geq 1 - \frac{(t-n)^2}{2n} + \frac{(t-n)^3}{3n^2}.$$

For  $n \geq 6$ , this inequality and Theorem 1 imply

$$1 - \theta(n) \geq \int_n^{n+2/3} \frac{\psi_{n-1}(t)}{\psi_{n-1}(n)} dt \geq \frac{2}{3} - \frac{4}{81n} + \frac{4}{243n^2},$$

and the upper bound for  $\theta(n)$  follows. It can be verified that the upper bound still holds for  $1 \leq n \leq 5$ . This completes the proof of Theorem 3.  $\square$

*Proof of Theorem 4.* Using the Taylor expansion of  $\psi_{n-1}$  at  $t = n$ , we have

$$1 - \theta(n) = \sum_{k=0}^{N-1} \frac{\psi_{n-1}^{(k)}(n)}{\psi_{n-1}(n)} \frac{(\lambda(n) - n)^{k+1}}{(k+1)!} + R_N.$$

We estimate the error term  $R_N$  as follows. Recall that

$$\psi_{n-1}^{(N)}(t) = w_{N, n-1}(t)\psi_{n-1}(t)/t^N.$$

It is not difficult to show by induction that  $\|w_{N, n-1}\| = \max\{|w_{N, n-1}(t)| : n \leq t \leq n + \log 2\} \leq N!n^{N/2}$ . Therefore,

$$\begin{aligned} |R_N| &= \left| \int_n^{\lambda(n)} \frac{(t-n)^N}{N!} \frac{\psi_{n-1}^{(N)}(\xi(t))}{\psi_{n-1}(n)} dt \right| \leq \frac{(\lambda(n)-n)^{N+1}}{(N+1)!} \frac{\|w_{N, n-1}\|}{n^N} \\ &\leq \frac{(\lambda(n)-n)^{N+1}}{(N+1)n^{N/2}} \rightarrow 0 \text{ as } N \rightarrow \infty. \end{aligned}$$

Let  $a_k = \psi_{n-1}^{(k)}(n)/\psi_{n-1}(n)$  for  $k \geq 0$ . It is easy to verify that  $a_0 = 1, a_1 = 0$ , and, for  $k \geq 2, a_k = -(k-1)[a_{k-1} + a_{k-2}]/n$ . Then we obtain

$$1 - \theta(n) = \sum_{k=0}^{\infty} a_k \frac{(\lambda(n)-n)^{k+1}}{(k+1)!}.$$

From this we can derive the asymptotic expansion of  $\lambda(n)$ . This finishes the proof.  $\square$

#### 4. REMARKS

It is not difficult to see that  $\lambda \text{ median}(\Gamma(\alpha, \lambda)) = \text{median}(\Gamma(\alpha, 1))$ . Therefore, we have the following corollary from Theorem 1.

**Corollary 5.** For  $n \geq 1$  and  $\lambda > 0$ , we have

$$\frac{-1}{3\lambda} \leq \text{median}(\Gamma(n, \lambda)) - \text{mean}(\Gamma(n, \lambda)) \leq \frac{-1 + \log 2}{\lambda}.$$

The bounds are best possible.

Using the fact that  $\chi_n^2 = \Gamma(\frac{n}{2}, \frac{1}{2})$ , we have

**Corollary 6.** For  $m \geq 1$ ,

$$-\frac{2}{3} \leq \text{median}(\chi_{2m}^2) - \text{mean}(\chi_{2m}^2) \leq 2(-1 + \log 2).$$

The bounds are best possible.

#### ACKNOWLEDGMENT

I would like to thank the referee for many helpful suggestions which improved the paper.

#### REFERENCES

1. K. O. Bowman, L. R. Shenton, and G. Szekeres, *A Poisson sum up to the mean and a Ramanujan Problem*, *J. Statist. Comput. Simulation* **20** (1984), 167–173.
2. J. Chen and H. Rubin, *Bounds for the difference between median and mean of Gamma and Poisson distributions*, *Statist. Probab. Lett.* **4** (1986), 281–283.
3. H. Dinges, *Special cases of second order Wiener Germ approximations*, *Probab. Theory Related Fields* **83** (1989), 5–57.
4. A. T. Doodson, *Relation of the mode, median and mean in frequency curves*, *Biometrika* **11** (1971), 425–429.
5. D. E. Knuth, *The art of computer programming*, Vol. 1, 2nd ed., Addison-Wesley, Reading, MA, 1973.

6. J. C. W. Marsaglia, *The incomplete Gamma function and Ramanujan's rational approximation to  $e^x$* , J. Statist. Comput. Simulation **24** (1986), 163–168.
7. S. Ramanujan, J. Indian Math. Soc. **3** (1911), 151–152.
8. ———, *Collected Papers*, Chelsea, New York, 1927.
9. G. Szegő, *Über einige von S. Ramanujan gestellte Aufgaben*, J. London Math. Soc. **3** (1928), 225–232.
10. G. N. Watson, *Theorems stated by Ramanujan (V): Approximations connected with  $e^x$* , Proc. London Math. Soc. (2) **29** (1927), 293–308.

DEPARTMENT OF MATHEMATICS, NATIONAL UNIVERSITY OF SINGAPORE, SINGAPORE 0511, REPUBLIC OF SINGAPORE

*E-mail address:* matckp@nusvm.bitnet