An Unusual Application of Cramér-Rao Inequality to Prove the Attainable Lower Bound for a Ratio of Complicated Gamma Functions

NITIS MUKHOPADHYAY
University of Connecticut
Department of Statistics
Storrs, Connecticut
USA
nitis.mukhopadhyay@uconn.edu

SRAWAN BISHNOI
University of Connecticut
Department of Statistics
Storrs, Connecticut
USA
srawan.bishnoi@uconn.edu

Abstract: A specific function $f(r)$ involving a ratio of complicated gamma functions depending upon a real variable $r(>0)$ is handled. Details are explained regarding how this function $f(r)$ appeared naturally for our investigation with regard to its behavior when r belongs to R^+. We determine explicitly where this function attains its unique minimum. In doing so, quite unexpectedly the customary Cramér-Rao inequality comes into play in order to nail down a valid proof of the required lower bound for $f(r)$ and locating where is that lower bound exactly attained.

Key–Words: Asymptotic distribution; CLT; Confidence interval; Cramér-Rao inequality; Gamma functions; Point estimation; Random CLT; Stopping time.

Received: January 14, 2019. Revised: April 21, 2019. Accepted: October 5, 2019.
Published: December 24, 2019.

1 Introduction

This investigation began with the following inquiry, Let

$$f(r) \equiv r^{-2}\left\{\pi^{1/2}\Gamma\left(r+\frac{1}{2}\right)\left(\Gamma\left(\frac{1}{2}r+\frac{1}{2}\right)\right)^{-2}-1\right\} \quad (1.1)$$

for $0<r<\infty$. Is it true that $f(r)\geq\frac{1}{2}$ for all $0<r<\infty$?

In a number of applications, to be highlighted shortly, we may be greatly interested in the behavior of $f(r)$ when r is smaller (than 10) rather than when r is larger (than 10). However, mathematically, we consider the behavior of $f(r)$ for all $r>0$.

Table 1 exhibits values of $f(r)$ for a selected set of r values between 0 and 6.0. While one may verify the following exact values:

$$f(1)=\tfrac{1}{2}\pi-1,\ f(2)=\tfrac{1}{2},\ f(3)=\tfrac{1}{9}(\tfrac{15\pi}{8}-1),$$
$$f(4)=\tfrac{2}{3},\ f(5)=\tfrac{1}{25}(\tfrac{945\pi}{128}-1),\text{ and } f(6)=\tfrac{113}{90},$$

a large majority of the exhibited values from Table 1 are clearly subject to reasonable numerical approximations built inside MAPLE.

Figure 1 tend to validate empirically our sentiment that $f(r)$ may attain its minimum value when $r=2$. Figure 2 additionally shows that $\frac{d}{dr}f(r)$ is negative (positive) when $0<r<2$ $(2<r<5)$ and thus $f(r)$ appears to be decreasing (increasing) when $0<r<2$ $(2<r<5)$.

The Empirical evidence from Figures 1-2 points in the direction of a much needed mathematical treatment in order to come up with a resolution of the query stated precisely in (1.1). The plots of $f(r)$ and and its derivative were both obtained directly using MAPLE.

The behavior of $f(r)$ for large r is rather straightforward and it may be treated as follows: Abramowitz and Stegun (1972, 6.1.38, p. 257) gave an approximate expression of the gamma function in the spirit of Stirling's approximation:

$$\Gamma(az+b)\sim\sqrt{2\pi}e^{-az}(az)^{az+b-\frac{1}{2}},\quad \text{for } a>0, z>0, \text{ and } z\to\infty. \quad (1.2)$$

The approximation (1.2) indicates that the ratio of the left-hand side and the right-hand side converges to 1 as $z\to\infty$. The approximation from (1.2) immediately leads to the following large-scale approximation:

$$f(r)\sim r^{-2}\left(2^{r-\frac{1}{2}}-1\right) \text{ as } r\to\infty.$$

While Stirling's approximation is well understood, exact monotonicity properties as well as minimization or maximization of complicated expressions involving ratios of gamma functions are almost never trivial. Indeed the mathematical analysis behind each

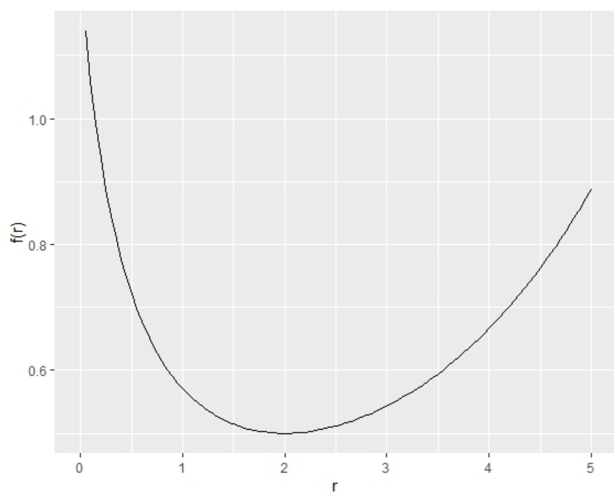

Figure 1: A plot of $f(r)$ coming from (1.1) when $0 < r < 5$.

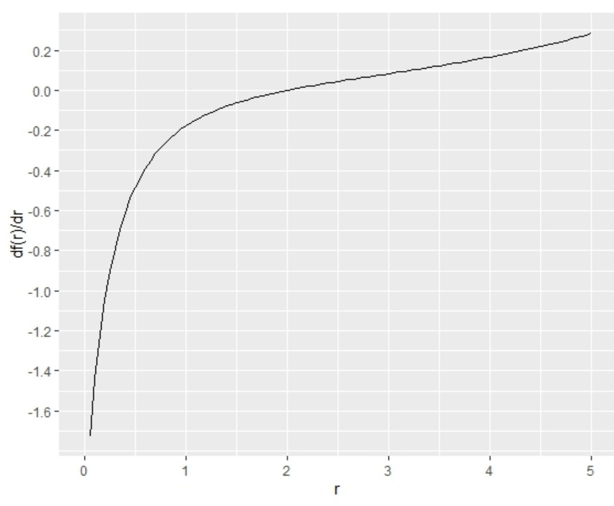

Figure 2: A plot of $df(r)/dr$ with $f(r)$ coming from (1.1) when $0 < r < 5$.

claim can become formidable and get out hand very quickly.

Instead of providing a full-blown review in this area, we provide a number of more recent references where inequalities and monotonicity properties have been addressed successfully in a number of interesting situations. One may look into the following resources: Mukhopadhyay and Bhattacharjee (2010) gave a number of extensions of Stirling's formula under heuristic approaches. In his 2010 paper of Mukhopadhyay developed a new sharper lower bound for a percentile of a Student's t distribution. Gut and Mukhopadhyay (2010) proved asymptotic and strict monotonicity of such a sharper lower bound for Student's t percentiles. In a 2011 paper, Mukhopadhyay developed a sharp Jensen's inequality along with some unusual applications. Mukhopadhyay and Son (2016) revisited Stirling's formula for gamma functions and bounds for ratios of gamma and beta functions and gave a synthesis with new results.

We must emphasize however that none of the results available in these cited sources or elsewhere lead to a successful resolution to the query stated in (1.1). In Section 2, we show explicitly how we came across to get hold of this complicated function, $f(r), r > 0$. In Section 3, we come up with a slightly generalized form of the function $f(r)$ and prove analytically (Theorem 3.1) where that generalized function is minimized exactly.

It is interesting to note that the query stated in (1.1) or an analogous query regarding a slightly generalized form of the function $f(r)$ as such should have absolutely nothing to do with any particular problem on statistical inference or statistical computation. Indeed the query from (1.1) is purely mathematical in nature. An answer should be very simple: "yes" or "no".

But, it so turns out that with some clever manipulation, we are able to connect the mathematical problem on hand indirectly with a suitably transformed problem in the context of the celebrated *Cramér-Rao inequality* (Cramér 1946; Rao 1945) and *minimum variance unbiased estimation* (MVUE). One may additionally refer to Mukhopadhyay (2000, pp. 365-371).

By doing precisely that, we thereby arrive at the following resolution of (1.1):

$$\boxed{\text{It is } \textbf{true} \text{ that } f(r) \geq \tfrac{1}{2} \text{ for all } 0 < r < \infty.} \quad (1.3)$$

Such a totally unexpected and outside-the-box proof may give this investigation a distinct edge in the minds of its readers. We close with brief concluding thoughts.

2 Motivation behind this investigation

In this section, we provide our key motivation behind raising the specific query stated in (1.1). Often, a stopping time N associated with a sequential sampling strategy can be generally expressed as:

$$N \equiv N_\nu = \inf\{n \geq m : \ n \geq \psi_\nu T_n\}, \qquad (2.1)$$

where m is the pilot sample size, $\{\psi_\nu;\ \nu \geq 1\}$ is a sequence of positive numbers such that $\psi_\nu \to \infty$ as $\nu \to \infty$, and $\{T_n;\ n \geq m\}$ is a sequence of positive statistics *with probability one* (w.p.1). We may cite two specific examples.

The stopping times from Anscombe (1952,1953), Ray (1957), and Chow and Robbins (1965) resemble exactly like that in (2.1) with

$$T_n \equiv (n-1)^{-1}\Sigma_{i=1}^{n-1} Z_i^2, \qquad (2.2)$$

where the Z_i's are *independent and identically distributed* (i.i.d.) $N(0,1)$ random variables under the auspices of a sequential fixed-width confidence interval problem for the mean in normal population with variance unknown.

Additionally, the stopping times from Robbins (1959), Starr (1966), and Ghosh and Mukhopadhyay (1976) also resemble exactly like that in (2.1) with

$$T_n \equiv \{(n-1)^{-1}\Sigma_{i=1}^{n-1} Z_i^2\}^{1/2} \qquad (2.3)$$

where the Z_i's are again i.i.d. $N(0,1)$ random variables under the auspices of a sequential minimum risk point estimation problem for the mean in normal population with variance unknown.

In (2.2)-(2.3), we customarily relate T_n to be the sample variance and sample standard deviation respectively after expressing the sample variance as the sample mean of $n-1$ i.i.d. χ_1^2 random variables by exploiting Helmert's orthogonal transformation (Mukhopadhyay 2000, pp. 197-201). In the context of both (2.2)-(2.3), it is understood that ψ_ν will coincide with the optimal fixed-sample-size n^* had σ^2 been known.

Under these two situations (2.2)-(2.3), however, the stopping rule (2.1) may be alternatively expressed in an unified fashion as follows:

$$\begin{gathered} N_r \equiv N_{\nu,r} = \inf\{n \geq m : \ n \geq \psi_\nu T_{n,r}\} \\ \text{with } T_{n,r} = \{(n-1)^{-1}\Sigma_{i=1}^{n-1} |Z_i|^r\, a_r^{-1}\}^{u/r}, \\ \psi_\nu \equiv n^*, a_s = \pi^{-1/2} 2^{s/2} \Gamma\left(\frac{1}{2}s + \frac{1}{2}\right), s = r, 2r \end{gathered} \qquad (2.4)$$

where (i) $u=2, r=2$ under (2.2) and (ii) $u=1, r=2$ under (2.3).

However, we allow $r(>0)$ to be *arbitrary* but it is held *fixed*. That way, we have a class of purely sequential stopping times indexed by $r(>0)$.

The exact distribution of $N_{\nu,r}$ is generally hard to obtain analytically in a closed form since the event $[N_{\nu,r} = n]$ depends on the $(n-m)$-dimensional statistic $\{T_{m,r}, T_{m+1,r}, \ldots, T_{n,r}\}$ for any fixed but otherwise arbitrary $n \geq m$. In some specific circumstances, an exact distribution of a stopping time may be determined with significant effort if the parameter(s) involved in the joint distribution of $\{T_{m,r}, T_{m+1,r}, \ldots, T_{n,r}\}$ could be assumed known for all $n \geq m$. Robbins (1959) gave an algorithm to determine the exact distribution of his associated stopping time. Some recent references include Zacks (2005,2009), Zacks and Mukhopadhyay (2009), Mukhopadhyay and Zacks (2018), and Mukhopadhyay and Zhang (2018).

Sections 2.1-2.3 summarize (i) Anscombe's (1952) *random central limit theorem* (Random CLT: Theorem 2.1), (ii) Ghosh and Mukhopadhyay's (1975) theorem (Theorem 2.2) and asymptotic distribution of N_ν, and (iii) the asymptotic distribution of $N_{\nu,r}$

2.1. Asymptotic Distribution of T_{N_ν} from (2.1)

Under the setup from (2.1), we first summarize (i) Anscombe's (1952) random *central limit theorem* (Random CLT) and (ii) Ghosh and Mukhopadhyay's (1975) theorem. At the outset, a customary CLT is assumed to hold for T_n, that is, we suppose:

$$n^{1/2}(T_n - g)/h \xrightarrow{\pounds} N(0,1), \text{ as } n \to \infty, \qquad (2.5)$$

with some real numbers $g(>0)$ and $h(>0)$. Anscombe's (1952) formulation of his Random CLT specified a set of sufficient conditions under which the following result:

$$N_\nu^{1/2}(T_{N_\nu} - g)/h \xrightarrow{\pounds} N(0,1), \text{ as } \nu \to \infty, \qquad (2.6)$$

would hold. One may review from Mukhopadhyay and Solanky (1994, pp. 42-43) and Ghosh et al. (1997, Section 2.7), among other sources. More recent treatments can be found in Mukhopadhyay and Chattopadhyay (2012) and Mukhopadhyay and Zhang (2018).

Theorem 2.1 (Anscombe's Random CLT, 1952). *Let $\{T_n;\ n \geq 1\}$ be a sequence of random variables. Let $\{n_\nu;\ \nu \geq 1\}$ be an increasing sequence of positive integers such that $n_\nu \to \infty$ as $\nu \to \infty$, and let $\{N_\nu; \nu \geq 1\}$ be a general sequence of positive* (*w.p.1*) *integer-valued random variables. Suppose that the following conditions hold*:

(*i*) $N_\nu / n_\nu \xrightarrow{P} 1$ *as* $\nu \to \infty$,

(*ii*) (*2.6*) *holds, that is,* $n_\nu^{1/2}(T_{n_\nu} - g)/h \xrightarrow{\pounds} N(0,1)$ *as* $\nu \to \infty$ *for some* $g(>0)$ *and* $h(>0)$, *and*

(*iii*) *for any given* $\varepsilon > 0, \eta > 0$, *there exist* $\delta > 0$ *and* n_0 *such that for* $n_\nu \geq n_0$, *we have*:

$\mathrm{P}\left\{ \sup_{|n'_\nu - n_\nu| \leq \delta n_\nu} n_\nu^{1/2} |T_{n'_\nu} - T_{n_\nu}| > \varepsilon \right\} < \eta$
[*Anscombe's u.c.i.p./tightness condition*

Under these sufficient conditions (*i*)-(*iii*), *we have*:

$$N_\nu^{1/2}(T_{N_\nu} - g)/h \xrightarrow{\pounds} N(0,1), \text{ as } \nu \to \infty,$$

that is, (*2.6*) *would hold.*

Condition (iii) is known as Anscombe's *uniform continuity in probability* (u.c.i.p.) condition. It is also referred to as the *tightness* condition. The sufficient conditions (i)-(iii) are satisfied when T_n is a sample mean, a U-statistic, or a *maximum likelihood estimator* (MLE), among others, obtained from a sequence of *independent and identically distributed* (i.i.d.) observations.

In the context of a variety of sequential inference strategies, n_ν is interpreted as an appropriate fixed sample size, had the (nuisance) parameter(s) been known. In such cases, N_ν from (2.1) would estimate n_ν, and $\{T_n; n \geq 1\}$ would be a sequence of observable positive (w.p.1) random variables.

2.2. Asymptotic Distribution of N_ν from (2.1)

Again, under the setup from (2.1), Ghosh and Mukhopadhyay (1975) developed a technique to "transfer" an asymptotic distribution of T_{N_ν} from Theorem 2.1 further along to conclude an asymptotic distribution of N_ν. That theorem comes from an unpublished manuscript which formed a part of Mukhopadhyay's (1975, Chapter 2) thesis and it played a key role in Carroll's (1977) derivation of the asymptotic normality result of stopping times based on robust estimators.

One should realize that the usual sufficient conditions to ensure (2.6) are the ones stated in Theorem 2.1 with $n_\nu \equiv g\psi_\nu$. A proof of (2.4) is rather brief and hence we omit it. One may review from Mukhopadhyay and Solanky (1994, Section 2.4), Ghosh et al. (1997, Exercise 2.7.4), Mukhopadhyay and de Silva (2009), and Zacks (2017), among other sources.

Theorem 2.2 (Ghosh-Mukhopadhyay Theorem, 1975). *Let* N_ν *be defined as in* (*2.1*). *Suppose that as* $\nu \to \infty$, *the following holds*:

$$N_\nu^{1/2}(T_{N_\nu} - g)/h \xrightarrow{\pounds} N(0,1) \quad \text{and } N_\nu^{1/2}(T_{N_\nu - 1} - g)/h \xrightarrow{\pounds} N(0,1). \quad (2.7)$$

Then we have,

$$g^{1/2}(N_\nu - g\psi_\nu)/(h\psi_\nu^{1/2}) \xrightarrow{\pounds} N(0,1) \text{ as } \nu \to \infty. \quad (2.8)$$

2.3. Asymptotic Distribution of $N_{\nu,r}$ from (2.4)

Now, we are in a position to precisely state and prove an asymptotic distribution of $N_{\nu,r}$ under the setup from (2.4) where u and r are both held *fixed*. Let us denote $W_{n,r} = (n-1)^{-1}\Sigma_{i=1}^{n-1} |Z_i|^r a_r^{-1}$ so that $T_{n,r} = W_{n,r}^{u/r}$. The CLT gives:

$$n^{1/2}(W_{n,r} - 1) \xrightarrow{\pounds} N\left(0, a_{2r}a_r^{-2} - 1\right) \text{ as } n \to \infty. \quad (2.9)$$

From (2.4), we recall that

$$a_s = \pi^{-1/2} 2^{s/2} \Gamma\left(\frac{1}{2}s + \frac{1}{2}\right), s = r, 2r.$$

Then, Mann-Wald theorem (Rao 1973, pp. 385-386; Mukhopadhyay 2000, pp. 261-262) and (2.9) leads to (as $n \to \infty$):

$$n^{1/2}(T_{n,r} - 1)/h_r \xrightarrow{\pounds} N(0,1) \text{ where } h_r^2 = u^2 f(r), \quad (2.10)$$

in the spirit of (2.5) where $f(r)$ was defined in (1.1).

Next, since the conditions of Theorem 2.1 are satisfied (with $n_\nu \equiv \psi_\nu$), we can immediately conclude (as $\nu \to \infty$):

$$N_{\nu,r}^{1/2}(T_{N_{\nu,r}} - 1)/h_r \xrightarrow{\pounds} N(0,1) \text{ and } N_{\nu,r}^{1/2}(T_{N_{\nu,r}-1} - 1)/h_r \xrightarrow{\pounds} N(0,1). \quad (2.11)$$

Then, in view of (2.11), Theorem 2.2 implies:

$$\left(N_{\nu,r} - \psi_\nu\right)/\psi_\nu^{1/2} \xrightarrow{\pounds} N(0, h_r^2) \text{ as } \nu \to \infty, \quad (2.12)$$

with h_r coming from (2.10).

2.4. More Insight from Combining (2.4) with (2.12) and Motivation Behind (1.1)

We may exploit the tools from nonlinear renewal theory of Woodroofe (1977,1982), Lai and Siegmund (1977,1979), Mukhopadhyay (1988), and Mukhopadhyay and Solanky (1994, pp. 48,50, Theorem 2.4.8, part iv) to claim (as $\nu \to \infty$):

$$\left(N_{\nu,r} - \psi_\nu\right)^2/\psi_\nu \text{ is uniformly integrable if } m > 1 + u, \quad (2.13)$$

that is, we can express:

$$Var[N_{\nu,r}] = u^2 f(r) + o(1) \text{ if } m > 1 + u. \tag{2.14}$$

In the contexts of the two specific sequential estimation problems addressed via (2.2) and (2.3), we recall that $\psi_\nu \equiv n^*$, the optimal fixed-sample-size and we can identify $(u = 1, r = 2)$ and $(u = 2, r = 2)$ under (2.2)-(2.3) respectively. Noting that $f(2) = \frac{1}{2}$, We may summarize as follows:

(i) Interval Estimation (2.2): $u = 2, r = 2$: $Var[N_{\nu,r}] = 2 + o(1)$ if $m \geq 4$;

(ii) Point Estimation (2.3): $u = 1, r = 2$: $Var[N_{\nu,r}] = \frac{1}{2} + o(1)$ if $m \geq 3$. (2.15)

Now, since $N_{\nu,r}$ estimates $\psi_\nu \equiv n^*$, the optimal fixed sample size, it should make sense to identify $r = r^*(> 0)$ analytically such that $f(r)$ is minimized at this $r = r^*$. Such N_{ν,r^*} may be referred to as the optimal (or most tight around ψ_ν) stopping time in the class of $\{N_{\nu,r}; r > 0\}$.

2.4.1. Motivation Behind the Query Stated in (1.1)

When one fixes $r = 2$ in defining (2.4), we note that the associated stopping time $T_{\nu,2}$ clearly reduces to one whose boundary crossing condition becomes an appropriate exponent of the sample variance.

Intuitively, then, since $u(> 0)$ is held fixed, it is appropriate to expect that in the class of all stopping rules $\{T_{\nu,r}; r > 0$ arbitrary, but $u > 0$ fixed$\}$, the stopping time $T_{\nu,2}$ with $r = 2$ should have the associated smallest asymptotic variance, $u^2 f(2)$. But, is this sentiment theoretically sound?

We will prove (1.3) by appropriately invoking the Cramér-Rao inequality, even though the query in itself has nothing to do with any kind of inference problem in statistics.

3 Proof of generalized version of (1.3)

Indeed, we first state and prove a result (Theorem 3.1) that is slightly more general than (1.3). Our proof involves a totally unexpected reliance on the customary Cramér-Rao inequality. In the end, we briefly return to (1.1) and address its resolution as stated in (1.3).

Theorem 3.1. *For all fixed* $\alpha > 0, s > 0$, *we have*:

$$\Gamma(\alpha + 2s)\,\Gamma(\alpha)\,\{\Gamma(\alpha + s)\}^{-2} \geq 1 + s^2\alpha^{-1}.$$

The equality holds if and only if $s = 1$.

Proof: We begin with a random variable Y Gamma(α, β), where $\beta(> 0)$ is an unknown parameter, but $\alpha(> 0)$ is known so that the *probability density function* (p.d.f.) of Y is given by:

$$a(y; \beta) = \{\beta^\alpha \Gamma(\alpha)\}^{-1} y^{\alpha-1} \exp(-y/\beta) I(y > 0), \tag{3.1}$$

where $I(.)$ stands for the indicator function of (.).

Then, in view of (3.1), we immediately have:

$$\begin{aligned} E_\beta[Y^s] &= \textstyle\int_0^\infty y^s a(y;\beta)dy = b(\beta;\alpha,s), \text{ say; and} \\ V_\beta[Y^s] &= b(\beta;\alpha,2s) - b^2(\beta;\alpha,s). \end{aligned} \tag{3.2}$$

Now, $U \equiv Y^s$ is an unbiased estimator of a parametric function (of β) which we have defined as $b(\beta; \alpha, s)$ in (3.2). Observe that $V_\beta[Y^s]/E^2_\beta[Y^s]$ does not involve β. Also, we obviously have:

$$\tfrac{\partial}{\partial\beta} b(\beta;\alpha,s) = s\beta^{-1} b(\beta;\alpha,s). \tag{3.3}$$

Next, let us evaluate Fisher-information about β in a single observation Y following the p.d.f. from (3.1) which belongs to a one-parameter exponential family as follows (in view of (3.3)):

$$\begin{aligned} \boldsymbol{I}_Y(\beta) &= E_\beta\left[\left\{\tfrac{\partial}{\partial\beta}\ln(a(Y;\beta))\right\}^2\right] \\ &= \mathrm{E}_\beta\left[\tfrac{(Y-\alpha\beta)^2}{\beta^4}\right] = \alpha\beta^{-2}. \end{aligned} \tag{3.4}$$

At this point, we invoke the Cramér-Rao inequality (Cramér 1946; Rao 1945). One may refer to numerous other sources including Mukhopadhyay (2000, pp. 366-371). Using (3.4), we can claim:

$$\begin{aligned} V_\beta[Y^s] &\geq \left\{\tfrac{\partial}{\partial\beta} b(\beta;\alpha,s)\right\}^2 \boldsymbol{I}_Y^{-1}(\beta) \\ &= \left\{s\beta^{-1} b(\beta;\alpha,s)\right\}^2 (\alpha\beta^{-2})^{-1} \\ &= s^2\alpha^{-1}\{b(\beta;\alpha,s)\}^2. \end{aligned} \tag{3.5}$$

The last expression seen in (3.5) is the *Cramér-Rao lower bound* (CRLB) for the variance of an unbiased estimator of a parametric function $b(\beta; \alpha, s)$, that is, we have:

$$\text{CRLB} = s^2\alpha^{-1}\{b(\beta;\alpha,s)\}^2. \tag{3.6}$$

But, using (3.2), we can alternatively express:

$$V_\beta[Y^s]/E^2_\beta[Y^s] = b(\beta;\alpha,2s)\{b(\beta;\alpha,s)\}^{-2} - 1, \tag{3.7}$$

so that the required inequality follows by combining (3.5) and (3.7).

In the context of the Cramér-Rao inequality, the Remark 7.5.1 in Mukhopadhyay (2000, p. 368) clearly leads us to conclude that the CRLB from (3.6) will be attained by the variance of an unbiased estimator of $b(\beta;\alpha,s)$ when $s=1$. Hence, equality all across (3.5) will hold when $s=1$. Now, the proof is complete.

3.1. A Resolution of the Query Stated in (1.1)

We immediately arrive at the resolution stated in (1.3) once we plug in $s=\frac{1}{2}r, \alpha=\frac{1}{2}$ in Theorem 3.1, by noting that $\Gamma\left(\frac{1}{2}\right)=\sqrt{\pi}$. Also, in (1.2), the equality will hold, that is $f(r)$ will be $\frac{1}{2}$ when $s=1$ or equivalently when $r=2$.

References:

[1] Abramowitz, M. and Stegun, I. (1972). *Handbook of Mathematical Functions*, ninth printing, New York: Dover.

[2] Anscombe, F. J. (1952). Large-Sample Theory of Sequential Estimation, *Proceedings of Cambridge Philosophical Society* 48: 600-607.

[3] Anscombe, F. J. (1953). Sequential Estimation, *Journal of Royal Statistical Society, Series B* 15: 1-29.

[4] Carroll, R. J. (1977). On the Asymptotic Normality of Stopping Times Based on Robust Estimators, *Sankhyā, Series A* 39: 355-377.

[5] Chow, Y. S. and Robbins, H. (1965). On the Asymptotic Theory of Fixed Width Confidence Intervals for the Mean, *Annals of Mathematical Statistics* 36: 457-462.

[6] Cramér, H. (1946). A Contribution to the Theory of Statistical Estimation, *Skandnavian Acturial Journal*: 85-94.

[7] Ghosh, M. and Mukhopadhyay, N. (1975). Asymptotic Normality of Stopping Times in Sequential Analysis, unpublished manuscript, Indian Statistical Institute, Calcutta, India.

[8] Ghosh, M. and Mukhopadhyay, N. (1976). On Two Fundamental Problems of Sequential Estimation, *Sankhyā, Series B* 38: 203-218.

[9] Ghosh, M., Mukhopadhyay, N., and Sen, P. K. (1997). *Sequential Estimation*, New York: Wiley.

[10] Gut, A. and Mukhopadhyay, N. (2010). On Asymptotic and Strict Monotonicity of a Sharper Lower Bound for Student's *t* Percentiles, *Methodology and Computing in Applied Probability* 12: 647-657.

[11] Lai, T. L. and Siegmund, D. (1977). A Nonlinear Renewal Theory with Applications to Sequential Analysis I, *Annals of Statistics* 5: 946-954.

[12] Lai, T. L. and Siegmund, D. (1979). A Nonlinear Renewal Theory with Applications to Sequential Analysis II, *Annals of Statistics* 7: 60-76.

[13] Mukhopadhyay, N. (1975). *Sequential Methods in Estimation and Prediction*, Ph.D. dissertation, May 1975, Indian Statistical Institute, Calcutta, India.

[14] Mukhopadhyay, N. (1988). Sequential Estimation Problems for Negative Exponential Populations, *Communications in Statistics - Theory & Methods* (*Reviews Section*) 17: 2471-2506.

[15] Mukhopadhyay, N. (2000). *Probability and Statistical Inference*, New York: Dekker.

[16] Mukhopadhyay, N. (2010). On a New Sharper Lower Bound for a Percentile of a Student's *t* Distribution with an Application, *Methodology and Computing in Applied Probability* 12: 609-622.

[17] Mukhopadhyay, N. (2011). On Sharp Jensen's Inequality and Some Unusual Applications, *Communications in Statistics, Theory & Methods* 40: 1283-1297.

[18] Mukhopadhyay, N. and Bhattacharjee, D. (2010). Stirling's Formula and Its Extensions: Heuristic Approaches, *Communications in Statistics, Theory & Methods* 39: 1046-1053.

[19] Mukhopadhyay, N. and Chattopadhyay, B. (2012). A Tribute to Frank Anscombe and Random Central Limit Theorem from 1952, *Sequential Analysis* 31: 265-277.

[20] Mukhopadhyay, N. and de Silva, B. M. (2009). *Sequential Methods and Their Applications*, Boca Raton: CRC.

[21] Mukhopadhyay, N. and Solanky, T. K. S. (1994). *Multistage Selection and Ranking Procedures*, New York: Dekker.

[22] Mukhopadhyay, N. and Son, M. (2016). Stirling's Formula for Gamma Functions and Bounds for Ratios of Gamma and Beta Functions: A Synthesis with New Results, *Methodology and Computing in Applied Probability* 18: 1117-1127.

[23] Mukhopadhyay, N. and Zacks, S. (2018). Modified Linex Two-Stage and Purely Sequential Estimation of the Variance in a Normal Distribution with Illustrations Using Horticultural Data, in *Advances in Statistical Planning and Inference*, 70th Birthday Festschrift for Professor Bimal K. Sinha and Professor Bikas K. Sinha, Carlos A. Coelho, Thomas Mathew, and Nitis Mukhopadhyay, eds., *Journal of Statistical Theory and Practice* 12: 111-135.

[24] Mukhopadhyay, N. and Zhang, C. (2018). EDA on the Asymptotic Normality of the Standardized Sequential Stopping Times, Part-I: Parametric Models, *Sequential Analysis* 37: 342-374.

[25] Rao, C. R. (1945). Information and Accuracy Attainable in the Estimation of Statistical Parameters, *Bulletin of Calcutta Mathematical Society* 37: 81-91.

[26] Rao, C. R. (1973). *Linear Statistical Inference and its Applications,* 2nd edition, New York: Wiley.

[27] Ray, W. D. (1957). Sequential Confidence Intervals for the Mean of a Normal Population with Unknown Variance, *Journal of Royal Statistical Society, Series B* 10: 133-143.

[28] Robbins, H. (1959). Sequential Estimation of the Mean of a Normal Population, in *Probability and Statistics*, H. Cramér volume, U. Grenander, ed., pp. 235-245, Uppsala: Almquist and Wiksell.

[29] Starr, N. (1966). On the Asymptotic Efficiency of a Sequential Procedure for Estimating the Mean, *Annals of Mathematical Statistics* 37: 1173-1185.

[30] Woodroofe, M. (1977). Second Order Approximations for Sequential Point and Interval Estimation, *Annals of Statistics* 5: 984-995.

[31] Woodroofe, M. (1982). *Nonlinear Renewal Theory in Sequential Analysis*, CBMS Lecture Notes #39, Philadelphia: SIAM.

[32] Zacks, S. (2005). Some Recent Results on the Distributions of Stopping Times of Compound Poisson Processes with Linear Boundaries, in *Special Issue in Honor of Herman Chernoff, Journal of Statistical Planning and Inference* 130: 95-109.

[33] Zacks, S. (2009). *Stage-Wise Adaptive Designs*, New York: Wiley.

[34] Zacks, S. (2017). *Sample Path Analysis and Distributions of Boundary Crossing Times*, Lecture Notes in Mathematics, New York: Springer.

[35] Zacks, S. and Mukhopadhyay, N. (2009). On Exact and Asymptotic Properties of Two-Stage and Sequential Estimation of the Normal Mean Under Linex Loss, *Communications in Statistics - Theory & Methods* 38: 2992-3014.