Journal of Statistical Research *https://doi.org/10.3329/jsr.v59i1.83680*
2025, Vol. 59, No. 1, pp. 5-8 *ISSN 0256 - 422 X*

A CONCENTRATION INEQUALITY FOR THE INVERSE GAUSSIAN DISTRIBUTION: A TRIBUTE TO THE LATE PROFESSOR A.K. MD. E. SALEH

MALAY GHOSH
Department of Statistics, University of Florida, Box 118545, Gainesville, FL 32611-8545
Email: ghoshm@stat.ufl.edu

SUMMARY

Concenration inequalities involving tail behavior of distributions have currently become very popular owing to their applications in the growing machine leaning literature. The paper offers one such inequality for the less known, but equally important inverse Gaussian distribution.

1 Introduction

It is my pleasure and priviledge to offer this short note as a humble tribute to the memory of Professor A.K.Md. E. Saleh, a man whom I have revered for more than five decades. Professor Saleh is indeed a unique individual who has contributed eminently to many facets of statistics. He started his career working on nonparametric statistics, but diversified himself over the years to several other areas of the subject, all in a very timely manner.

Of particular interest, in addition to his research, is the number of books written by Professor Saleh. It started with his sole author book entitled "Theory of Preliminary Test and Stein Type Estimation with Applications (2006)." This book is highly valuable in the sense that it compares and contrasts very nicely preliminary test estimators initiated by Bancroft and his associates with Stein type empirical Bayes estimators. This was followed later in his coauthored book with Rohatgi entitled "Introduction to Probability and Statistics" (2015). This is an ideal Masters level text with a very comprehensive treatment of both probability and inference at an appropriate level. This book has received more than 2,500 citations, truly a great feat, attained very rarely by authors of statistics texts.

Next comes "Theory of Ridge Regression Estimation with Applications" joint with Arashi and Kibria (2019). This was also very timely because of statisticians' increased interest in regularized estimation. Finally, even towards the end of his life, when he was suffering from multiple health issues, Professor Saleh got interested in machine learning, tying it up with his earlier work on non-parametics and shrinkage estimation. This culminated evetually in the book "Rank Based methods for Shrinkage and Selection with Application to Machine Learning" coauthored with Arashi, RA Saleh and Norouzirad (2022).

The publication list of Professor Saleh is enormous, many in *The Annals of Mathematical Statistics* and later in *The Annals of Statistics*. This is indeed an enviable achievement since these journals have always been included in the list of top four journals of statistics. More importantly, his publications in these journals spanned for nealy three decades, an enviable feat for any statistician.

What I find remarkable of Professor Saleh is that in spite of all his research and teaching accomplishments, he was always a modest and down to earth person, friendly to all his colleagues, in and outside his own intitution. I myself had research collaboration with him for a very short period, and was able to write an influential paper entitled "Empirical Bayes Subset Estimation in Regression Models" joint with him and Professor P.K. Sen. The paper offered for the first time an empirical Bayes approach in contrast to a preliminary test approach for subset selection in regression models, with the attractive features of minimaxity. It was a pleasant and fruitful collaboration which I still cherish.

With this eulogy to a great scientist, teacher and scholar, let me now move into the technical part of the paper. I want to prove here a concentration inequality for the Inverse Gaussian distribution. Inverse Gaussian distribution belongs to the exponential family with a single mode and long tail. This distribution, often suitable for modeling nonnegative positively skewed data, has been found very appropriate for a wide range of applications such as survival analysis, finance, medicine and others. Moreover, this distribution, also often referred to as the Wald distribution, arose quite naturally as the limiting distribution of the sample size in the Sequential Probability Ratio Test as developed by Abraham Wald.

Despite its many uses, Inverse Gaussian distribution has remained obscure for most statistics researchers, and has seldom appeared in statistics textbooks. I want to adddress in this note the tail bahavior of the sample mean for an inverse Gaussian distribution, which to my knowledge has not appeared elsewhere.

General techniques for finding such tail behavior have appeared in the articles of Chernoff (1952, 1956) and Bahadur and Ranga Rao (1960). Their work led to very important large deviation results, very succintly reproduced in the monograph of Bahadur (1971), and subsequently generalized in several directions by a large number of authors. My aim is to point out also an explicit large deviation index for the inverse Gaussian distibution.

The following section states and proves the main results followed by several comments.

2 The Main Result

Let $X_1, \ldots, X_n$ be iid with a common inverse Gaussian pdf

$$f_\mu(x) = (2\pi x^3)^{-1/2} \exp[-(1/(2x))(x/\mu - 1)^2], \tag{2.1}$$

$x > 0$ and $\mu > 0$. Then it is well-known that $E(X) = \mu$ and the moment generating function is given by (Chhikara and Folks, 1989: Sedasdri, 1999)

$$M_X(t) = \exp[(1/\mu)\{1 - (1 - 2\mu^2 t)^{1/2}\}],\ t < (2\mu^2)^{-1}. \tag{2.2}$$

The following theorem provides tail behavior of the sample mean for an inverse Gaussin distribution.

Theorem 1. *Let $\bar{X}$ denote the sample mean. Then for every $c > 0$,*

$$P[\bar{X} - \mu > c] \leq \exp\left[-\frac{nc^2}{2\mu^2(\mu + c)}\right].$$

Proof. Following the standard approach, we apply Bernstein's inequality, which is an exponential version of Markov's inequality. This gives

$$P[\bar{X} - \mu > c] \leq \inf_{0<t<(2\mu^2)^{-1}} \exp[-nt(\mu + c) + (n/\mu)\{1 - (1 - 2\mu^2 t)^{1/2}\}]. \tag{2.3}$$

Let $g(t) = -nt(\mu + c) + (n/\mu)\{1 - (1 - 2\mu^2 t)^{1/2}\}$. Then

$$g'(t) = n[\mu(1 - 2t\mu^2)^{-1/2} - (\mu + c)]. \tag{2.4}$$

Also, $g^{''}(t) = n\mu^3(1 - 2t\mu^2)^{-3/2} > 0$. Thus it follows from (4) that $g(t)$ is minimized at $t = t_0$, where t_0 satisfies $1 - 2t_0\mu^2 = \mu^2/(\mu + c)^2$, or equivalently $t_0 = (1/2\mu^2)(2\mu + c)/(\mu + c)^2$. Now it follows from (3) that $\inf_{0<t<(2\mu^2)^{-1}} \exp[g(t)] = \exp[g(t_0)]$, where

$$\begin{aligned}
\exp[g(t_0)] &= \exp[-nt_0(\mu + c) + (n/\mu)\{1 - (1 - 2t_0\mu^2)^{1/2}\}] \\
&= \exp[-nc(2\mu + c)/(2\mu^2(\mu + c)) + (n/\mu)(1 - \mu/(\mu + c))] \\
&= \exp[-n\{c(2\mu + c)/(2\mu^2(\mu + c) - c/(\mu(\mu + c))\}] \\
&= \exp[-nc^2/(2\mu^2(\mu + c))]. \qquad (2.5)
\end{aligned}$$

□

Remark 1. There are some examples of curved exponential family distributions. One well known example is a N(μ, μ) distribution, namely a normal distribution with mean and variance both equal to $\mu(> 0)$. Let $X_1, \ldots, X_n$ be iid with the above distribution. Then applying the standard Bernstein inequality for a N$(0, 1)$ distribution, one gets for $c > 0$,

$$P(\bar{X} - \mu > c) = P[N(0, 1) > c/\mu^{1/2}] \leq \exp(-nc^2/(2\mu)).$$

An analogous concentration inequality for an Inverse Gaussian distribution is given below.

Consider the Inverse Gaussian distribution with pdf given by $f_\mu(x) = (\mu/2\pi x^3)^{1/2} \exp[-(x - \mu)^2/(2x)$. Then $E(X) = V(X) = \mu$. The corresponding mgf is given by $M_X(t) = \exp[\mu\{1 - (1 - 2t)^{1/2}\}]$, $0 < t < 1/2$. Then we have the following concentration inequality.

Theorem 2. $P(\bar{X} - \mu > c) \leq \exp[-nc^2/(2(\mu + c))]$, $c > 0$.

Proof. Once again, an application of Bernstein's inequality provides

$$P(\bar{X} - \mu > c) \leq \inf_{0<t,1/2} \exp[-nt(\mu + c) + n\mu\{1 - (1 - 2t)^{1/2}\}]. \tag{2.6}$$

Let $g(t) = -nt(\mu + c) + n\mu\{1 - (1 - 2t)^{1/2}\}$. Then $g'(t) = -n(\mu + c) + n\mu(1 - 2t)^{-1/2}$. Also, $g^{''}(t) = n\mu(1-2t)^{-3/2} > 0$. Hence, $g(t)$ is minimized at $t = t_0$, where $(1-2t_0)^{-1/2} = (\mu+c)/\mu$, or equivalently $t_0 = (c\mu + c^2/2)/(\mu + c)^2$. Hence, from (6), one is led to the inequality

$$\begin{aligned} P(\bar{X} - \mu > c) &\leq \exp[-nt_0(\mu + c) + n\mu\{1 - (1 - 2t_0)^{1/2}\}] \\ &= \exp[-n(c\mu + c^2/2)/(\mu + c) + nc\mu/(\mu + c)] \\ &= \exp[-nc^2/(2(\mu + c))]. \end{aligned}$$

This proves the theorem. □

Remark 2. Following the work of Chernoff (1956) and Bahadur and Ranga Rao (1960), we define the large deviation index $\rho = \lim_{n\to\infty}(-1/n)\log P(\bar{X}_n - \mu > c)$ for iid samples $X_1, \ldots, X_n$. From Theorem 1, it follows that $\rho = c^2/[2\mu^2(\mu + c)]$. Also, as a consequence of Theorem 2, $\rho = c^2/[2(\mu + c)]$.

References

Bahadur, R. R., and Ranga Rao, R. (1960), "On deviations of the sample mean," *Annals of Mathematical Statistics*, 31, 1015–1027.

Bahadur, R. R. (1971), *Some Limit Theorems in Statistics*, SIAM, Philadelphia, Pennsylvania.

Chernoff, H. (1952), "A measure of asymptotic efficiency for tests of a hypothesis based on the sums of observations," *Annals of Mathematical Statistics*, 23, 493–507.

Chernoff, H. (1956), "Large sample theory—parametric case," *Annals of Mathematical Statistics*, 27, 1–22.

Chhikara, R. S., and Folks, J. L. (1989), *The Inverse Gaussian Distribution: Theory, Methodology and Applications*, Marcel Dekker, New York.

Seshadri, V. (1999), *The Inverse Gaussian Distribution*, Springer Verlag.

Received: *November 25, 2024*

Accepted: *February 26, 2025*