ON SUBSET SELECTION PROCEDURES FOR INVERSE GAUSSIAN POPULATIONS

by

Shanti S. Gupta
Purdue University

Hwa-Ming Yang
University of Toledo

Technical Report #82-7

Department of Statistics
Purdue University

March 1982

*This research was supported by the Office of Naval Research Contract N00014-75-C-0455 at Purdue University. Reproduction in whole or in part is permitted for any purpose of the United States Government.

Summary

The inverse Gaussian or the first passage time probability distribution for Brownian motion with a drift is particularly important for modeling and interpreting observed distributions of time intervals in many different fields of research. In this paper we deal with the problem of selecting a subset of k inverse Gaussian populations which includes the "best" population, i.e. the (unknown) population which is associated with the largest value of the unknown means. The shape parameters of the inverse Gaussian distributions are assumed to be equal for all the k populations. When the common shape parameter is known, a procedure R_1 is defined and studied which selects a subset which is nonempty, small in size and just large enough to guarantee that it includes the best population with a preassigned probability regardless of the true unknown values of the means. For the case when the common shape parameter is unknown a procedure R_2 is proposed. For the procedures R_1 and R_2, we obtain exact results for k=2 concerning the infimum of the probability of a correct selection. For $k \geq 3$ a lower bound on the probability of a correct selection is derived for each case. Formulas for the constants d_1 and d_2 which are necessary to carry out the procedures R_1 and R_2, respectively, are obtained. An upper bound on the expected number of populations retained in the selected subset is given.

If the best population is defined as the one associated with the largest shape parameter, it is shown that with a suitably chosen statistic, this problem concides with the problem of selecting a subset of k normal populations which includes the population with the smallest variance. Similarly, for the selection of a subset containing the smallest shape parameter, the problem reduces to selection in terms of the largest scale parameter of the gamma distributions.

On Subset Selection Procedures for Inverse Gaussian Populations*

by

Shanti S. Gupta
Purdue University

Hwa-Ming Yang
University of Toledo

1. Introduction, Basic Concepts and Notation

The inverse Gaussian or the first passage time probability density function (p.d.f.) for Brownian motion with a drift is particularly important for modeling and interpreting observed distributions of time intervals in many different fields of research. For example, Hasofer (1964) considered the inverse Gaussian model for the emptiness of dam, Marcus (1975, 1976) used it in communications noise and highway noise models, Banerjee and Bhattacharryya (1976) applied it in a study of purchase incidence models, Chhikara and Folks (1977) studied it in reliability and life testing, among others. Also the statistician often finds himself dealing with data of considerable skewness with no obvious choice of distribution suggested by physical consideration. In such cases the choice is always made upon the basis of goodness-of-fit and upon the ease of working with the chosen distribution. Because of the ease due to the exact sampling distribution theory of the inverse Gaussian it would appear to be a strong candidate in such cases and for this reason, Chhikara and Folks (1977) suggested that the use of the inverse Gaussian over the lognormal would be perferable.

The probability distribution of the first passage time in Brownian motion with a drift was first derived by Schrödinger (1915). Tweedie (1956, 1957a, 1957b) studied the properties of it and proposed the name inverse Gaussian distribution for it. This distribution is also known as Wald's distribution (cf. Wald (1947)).

---

* This research was supported by the Office of Naval Research contract N0014-75-C-0455 at Purdue University.

In this paper, we consider the problem of selecting a nonempty (small) subset of k different inverse Gaussian populations which contain the "best" i.e. the population which is associated with the largest unknown mean or the distribution which is associated with the largest shape parameter.

The inverse Gaussian distribution has two parameters with p.d.f. expressed, in two alternative parametizations, as

$$g(x;\ \nu,\ \sigma^2,\ a) = \frac{a}{\sigma\sqrt{2\pi x^3}}\ \exp\left\{-\frac{(a-\nu x)^2}{2\sigma^2 x}\right\},\ x,\nu,\sigma,\ a > 0 \qquad (1.1)$$

$$= 0, \text{ otherwise,}$$

and

$$f(x;\ \mu,\lambda) = \left(\frac{\lambda}{2\pi x^3}\right)^{\frac{1}{2}}\ \exp\left\{-\frac{\lambda(x-\mu)^2}{2\mu^2 x}\right\},\quad x,\mu,\lambda > 0 \qquad (1.2)$$

$$= 0, \text{ otherwise.}$$

Expression (1.1) is convenient for interpretation in terms of Brownian motion. Suppose W(x) is a Brownian motion (Wiener process, see Cox and Miller (1965)) with drift ν and variance parameter σ^2, i.e. a stochastic process with the following properties:

(a) W(0) = 0 a.e. and W(x) has independent increments;

(b) for any time interval (x_1,x_2), $W(x_2)-W(x_1)$ is normally distributed with mean $\nu(x_2-x_1)$ and variance $\sigma^2(x_2-x_1)$, then formula (1.1) gives the p.d.f. of the first passage time X of W(x) with positive drift ν to barrier $a > 0$.

Expression (1.2) is useful for deriving some results which are parallel to that of the usual normal distribution. It is known that the parameter μ is the mean and λ is a shape parameter.

From (1.1) and (1.2) it is easy to see that the relations $\mu = a/\nu$ and $\lambda = a^2/\sigma^2$ hold. Therefore, comparing k inverse Gaussian means μ_i's is equivalent in some sense to comparing the associated drifts of Brownian motion. Note that the inverse Gaussian distribution is a member of the exponential family. From now on we will use (1.2) to formulate our problem.

For a random variable X distributed according to (1.2), we denote $X \sim I(\mu,\lambda)$. For this distribution λ is a shape parameter, the mean is μ and the variance is μ^3/λ. If $X_1,\ldots,X_n$ is a random sample from $I(\mu,\lambda)$, Schrödinger (1915) showed that the maximum likelihood estimates of μ and λ are given by

$$\hat{\mu} = \bar{X} \text{ and } \hat{\lambda} = n/\sum_{i=1}^{n}(1/X_i - 1/\bar{X}),$$

where

$$\bar{X} = \sum_{i=1}^{n} X_i/n.$$

Tweedie (1957a) proved that $\bar{X} \sim I(\mu,\lambda)$, $\lambda\sum_{i=1}^{n}(1/X_i - 1/\bar{X}) \sim \chi^2_{n-1}$, the chi-square distribution with n-1 degrees of freedom and that they are stochastically independent. The statistics $\bar{X}$ and $\Sigma(1/X_i - 1/\bar{X})$ jointly are sufficient and complete for (μ,λ), and $\bar{X}$ is a complete sufficient statistic for μ if λ is known.

Let $\pi_1,\ldots,\pi_k$ be k independent inverse Gaussian populations with means $\mu_1,\ldots,\mu_k$ and shape parameters $\lambda_1,\ldots,\lambda_k$, respectively. Let $\mu_{[1]} \leq \ldots \leq \mu_{[k]}$ be the ordered μ_i's. It is assumed that there is no prior knowledge of the correct pairing of the ordered and the unordered μ_i's. Let X_{ij}, $j=1,\ldots,n_i$, $i=1,\ldots,k$ be independent samples for $\pi_1,\ldots,\pi_k$, respectively, and let

$\bar{X}_i = \sum_{j=1}^{n_i} X_{ij}/n_i$, $i=1,\ldots,k$ denote the sample means. Let $\bar{X}_{(i)}$ and $n_{(i)}$ denote

the sample mean and sample size associated with the unknown population $\pi_{(i)}$ with mean $\mu_{[i]}$, i=1,...,k.

Given any P*, 1/k < P* < 1, our goal is to select a subset of these k populations such that the subset contains the best population with probability at least P*, no matter what the true configuration of μ_i's. Selection of a subset which contains the best population is called a correct selection and is denoted by CS. Therefore, we are interested in defining (and studying) a selection procedure R such that

$$\inf_{\underline{\mu}\in\Omega} P_{\underline{\mu}}(CS|R) \geq P^* \tag{1.3}$$

where Ω is the set of all k-tuples $(\mu_1,\ldots,\mu_k)$, $\mu_i > 0$, i=1,...,k. This requirement will be referred to as the P*-condition. In Sections 2 and 3, we discuss the cases of known and unknown common shape parameter λ, respectively. For each case, a conditional selection procedure is proposed and studied. In Section 4, the problem of selecting a subset which contains the larges shape parameter is considered. It is shown that with a suitably chosen statistic this problem is equivalent to the problem of selecting a subset of k normal populations which includes the population with the smallest variance. In other words, the problem of selecting the inverse Gaussian population with the largest (smallest) shape parameter reduces to the problem of selecting the gamma population with the smallest (largest) scale parameter.

2. Selection of the Inverse Gaussian Population with the Largest Mean When $\lambda_i = \lambda$, i=1,...,k is Known

2.1. A Conditional Selection Procedure R_1

When the common shape parameter is known, we propose the following conditional selection procedure R_1:

R_1: Select the population π_i if and only if

$$\bar{X}_i \geq \max_{1\leq j\leq k} \bar{X}_j - d_1(t), \text{ given } T = \sum_{i,j} X_{ij} = t ,$$

where $t > 0$ and $d_1(t)$ is the smallest positive value to satisfy the P*-condition.

It is known that for two independent random samples $X_{11},\ldots,X_{1n}$, from $I(\mu_1,\lambda)$ and $X_{21},\ldots,X_{2n_2}$ from $I(\mu_2,\lambda)$, the joint p.d.f. constitutes a three-parameter exponential family and may be written in the form

$$\exp(\psi t + \theta u + \eta v),$$

where

$$\begin{aligned} \psi &= -\lambda(n_1\mu_1^{-2} + n_2\mu_2^{-2})/2(n_1+n_2), \\ \theta &= -\lambda(\mu_1^{-2} - \mu_2^{-2})\, n_1n_2/2(n_1+n_2), \\ \eta &= -\lambda/2, \end{aligned} \tag{2.1}$$

and t, u, v denote the values of the statistics

$$T = \sum_{i=1}^{n_1} X_{1i} + \sum_{j=1}^{n_2} X_{1j} ,$$

$$U = \bar{X}_1 - \bar{X}_2, \quad \text{where } \bar{X}_i = \sum_{\ell=1}^{n_i} X_{i\ell}/n_i, \quad i = 1,2,$$

and

$$V = \sum_{i=1}^{n_1} X_{1i}^{-1} + \sum_{j=1}^{n_2} X_{2j}^{-1} ,$$

respectively.

For $k = 2$, the following theorem gives us an exact result.

Theorem 2.1. For a given P^*, $1/k < P^* < 1$, $k = 2$, let $d_1(t)$ be the smallest value such that

$$P_{\underline{\mu} \in \Omega_0} (\bar{X}_1 - \bar{X}_2 \leq d_1(t) | T=t) = P^*$$

where $\Omega_0 = \{\underline{\mu} \in \Omega | \mu_1 = \ldots = \mu_k > 0\}$.

Then, $\inf_{\underline{\mu} \in \Omega} P_{\underline{\mu}} (CS|R) = P_{\underline{\mu} \in \Omega_0} (CS|R_1) = P^*$.

Note that the infimum of $P(CS|R)$ does not depend on the common value of $\mu_1 = \mu_2 = \mu$.

Proof: Since λ is known, the joint p.d.f. of $X_{11}, \ldots, X_{1n_1}$ and $X_{21}, \ldots, X_{2n_2}$ belongs to a two-parameter exponential family. It follows from an argument similar to that in Lehmann (1959, p. 136) that

$$\begin{aligned} P_{\underline{\mu}}(CS|R) &= P_{\theta \leq 0} (\bar{X}_{(1)} - \bar{X}_{(2)} \leq d_1(t) | T=t) \\ &\geq P_{\theta=0} (\bar{X}_{(1)} - \bar{X}_{(2)} \leq d_1(t) | T=t) \\ &= P_{\underline{\mu} \in \Omega_0} (CS|R_1) . \end{aligned}$$

Hence $\inf_{\underline{\mu} \in \Omega} P_{\underline{\mu}} (CS|R_1) = P_{\underline{\mu} \in \Omega_0} (CS|R_1) = P^*$.

Lemma 2.1. If two random variables Z and X are independent of another random variable Y, and if the just p.d.f.'s $f_{Z,X}$ and $f_{Z,X+Y}$ exist, then

$$f_{Z,X+Y}(z,t) = \int_{-\infty}^{\infty} f_{Z,X}(z,x) f_Y(t-x) dx \; ; \tag{2.2}$$

$$= \int_{0}^{t} f_{Z,X}(z,x) f_Y(t-x) dx, \tag{2.3}$$

if both random variables X and Y take only positive values.

Proof: The proof is straight forward and hence is omitted.

For $k \geq 3$, based on the Bonferroni inequalities and Lemma 2.1, we drive a lower bound on the probability of a correct selection in Theorem 2.2.

Theorem 2.2. For $k \geq 3$, and given P^*, $1/k < P^* < 1$ and $T = \sum_{i,j} X_{ij} = t$, let $P_1^* = 1 - \frac{1-P^*}{k-1}$ and let $d_{ij}^{(1)}(r)$ be the smallest value such that for any $\underline{\mu} \in \Omega_0 = \{\underline{\mu} | \mu_1 = \ldots = \mu_k = \mu > 0\}$, and any $i,j,j \neq i$,

$$P_{\underline{\mu}}(U_{ij} < d_{ij}^{(1)}(r) \mid T_{ij} = r) = P_1^* \tag{2.4}$$

where

$$U_{ij} = \bar{X}_{(i)} - \bar{X}_{(j)}$$

$$T_{ij} = n_{(i)} \bar{X}_{(i)} + n_{(j)} \bar{X}_{(j)}, \quad 1 \leq i \neq j \leq k.$$

Let

$$d_1(t) = \max \{d_{ij}^{(1)}(r) \mid 1 \leq i \neq j \leq k, 0 < r \leq t\} \tag{2.5}$$

then

$$\inf_{\underline{\mu} \in \Omega} P_{\underline{\mu}}(CS|R_1) \geq P^* .$$

Proof: For all $\underline{\mu} \in \Omega$

$$\begin{aligned} & P_{\underline{\mu}}(CS|R_1) \\ &= P_{\underline{\mu}}(\bar{X}_{(k)} \geq \max_{1 \leq j \leq k-1} \bar{X}_{(j)} - d_1(t)|T=t) \\ &= 1 - P_{\underline{\mu}}(\bar{X}_{(k)} < \max_{1 \leq j \leq k-1} \bar{X}_{(j)} - d_1(t)|T=t) \\ &\geq 2 - k + \sum_{j=1}^{k-1} P_{\underline{\mu}}(U_{jk} \leq d_1(t)|T=t) \end{aligned} \tag{2.6}$$

For any j, $1 \leq j \leq k-1$, using Lemma 2.1 we have

$$P_{\underline{\mu}}\,(U_{jk} \le d_1(t) \mid T = t)$$

$$= \int_0^t P(U_{jk} \le d_1(t) \mid T_{jk}=r)\; f_{T_{jk}}(r)\cdot f_{T-T_{jk}}(t-r)dr/f_T(t)$$

$$\ge \int_0^t P(U_{jk} \le d_{jk}^{(1)}(r) \mid T_{jk} = r)\; f_{T_{jk}}(r)\cdot f_{T-T_{jk}}(t-r)dr/f_T(t)$$

$$\ge P_1^* \,.$$

Hence $\inf_{\underline{\mu}\in\Omega} P_{\underline{\mu}}\,(CS|R_1) \ge 2-k + (k-1)\, P_1^* = P^* \,.$

2.2. Evaluation of Values of $d_1(t)$ for the Procedure R_1

For two independent random samples $X_{11},\ldots,X_{1n_1}$ from $I(\mu_1,\lambda)$ and $X_{21},\ldots,X_{2n_2}$ from $I(\mu_2,\lambda)$, let $T = \sum_i X_{1i} + \sum_j X_{2j}$, then it follows from Tweedie (1957a) that $T \sim I((n_1+n_2)\mu,\ (n_1+n_2)^2\,\lambda)$ if $\mu_1=\mu_2=\mu$. Chhikara (1975) derived the conditional p.d.f. $g(u|t)$ of $U \equiv \bar{X}_1 - \bar{X}_2$, given $T = t$, $\mu_1=\mu_2$, as

$$g(u|t) = \left[\frac{n_1n_2(n_1+n_2)^2\,\lambda t^3}{2\pi(t+n_2\,u)^3(t-n_1u)^3}\right]^{1/2} \exp\left[-\frac{n_1n_2(n_1+n_2)^2\,\lambda u^2}{2t(t+n_2u)(t-n_1u)}\right], \qquad (2.7)$$

$$-\frac{t}{n_2} < u < \frac{t}{n_1} \quad .$$

By using the one - to-one transformation

$$y = \frac{(n_1n_2\lambda)^{\frac{1}{2}}(n_1+n_2)u}{[t(t+n_2u)(t-n_1u)]^{\frac{1}{2}}}\,, \qquad (2.8)$$

it can be shown that the P*-percentile point $i(P^*) \equiv i(P^*, n_1, n_2, t)$ of U, given $T=t$ i.e., the solution of the equation

$$\int_{-\infty}^{i(P^*)} g(u|t)du = P^*$$

is given by the following equation

$$\Phi(d_0(t)) + \frac{n_2-n_1}{n_1+n_2} \exp\left(\frac{2n_1n_2\lambda}{t}\right)\left\{1-\Phi\left[\left(d_0^2(t) + \frac{4n_1n_2\lambda}{t}\right)^{\frac{1}{2}}\right]\right\} = P^* , \qquad (2.9)$$

where

$$d_0(t) = i(P^*)(n_1+n_2)\ [n_1n_2\ \lambda/t(t+n_2\ i(P^*))(t-n_1\ i(P^*)]^{\frac{1}{2}} , \qquad (2.10)$$

and Φ is the cumulative distribution function (c.d.f.) of a standard normal distribution.

When $n_1 = n_2 = n$, the equation (2.8) will simplify to $d_0(t) = z(P^*)$, the P*-percentile point of the standard normal distribution. Hence we have

$$i(P^*)=\left[\frac{z^2(P^*)t^3}{4n^4\lambda+z^2(P^*)tn^2}\right]^{\frac{1}{2}} \qquad (2.11)$$

which is increasing in t, if n is fixed. Note that $i(p^*) \to 0$ as $n \to \infty$ if $t = 0(n)$.

Corollary 2.1. For k=2, the constant $d_1(t)$ associated with the procedure R_1 is given by

$$d_1(t) = i(P^*), \qquad (2.12)$$

where $i(P^*)$ is given by (2.9) or (2.11).

Corollary 2.2. For $k \geq 3$, the constant $d_1(t)$ associated with the procedure R_1 is given by

$$d_1(t) = \max\{i(P_i^*, n_i, n_j, r) \mid 1 \leq i \neq j \leq k, 0 < r \leq t\}$$

$$= i(P_1^*, n, n, t), \quad \text{if } n_1 = \ldots = n_k = n. \qquad (2.13)$$

2.3 An Upper Bound on the Expected Subset Size and Other Properties of Procedure R_1

For any given values of k and P*, the size of the selected subset S by using the procedure R_1 is a function of the true configuration $\underline{\mu} = (\mu_1,\ldots,\mu_k)$ and it also depends on $n_1,\ldots,n_k$. Note that S is an interger-valued random variable which takes values 1 to k inclusive. Hence, (in alalogy with power of the hypothesis testing problem) $E_{\underline{\mu}}(S|R_1)$ can be looked upon as a measure of the efficiency of the procedure R_1. We now discuss how to evaluate it. We consider the space of all slippage configurations of the type $\mu_{[1]}=\ldots=\mu_{[k-1]}=\mu$ and $\mu_{[k]}=\delta\mu$, $\delta > 1$; and we denote this space by $\Omega(\delta)$. We also assume that $n_1 = n_2 =\ldots=n_k = n$.
Then, for any $\underline{\mu}\in\Omega(\delta)$, the expected size of the selected subset is

$$
\begin{aligned}
E_{\underline{\mu}}(S|R_1) &= P_{\underline{\mu}}\left(\bar{X}_{(k)} \geq \max_{1\leq j\leq k-1} \bar{X}_{(j)} - d_1(t) \,|\, T=t\right) \\
&\quad + (k-1)\, P_{\underline{\mu}}\left(\bar{X}_{(1)} \geq \max_{2\leq j< k} \bar{X}_{(j)} - d_1(t) \,|\, T=t\right) \\
&\leq P_{\underline{\mu}}\left(\bar{X}_{(1)} - \bar{X}_{(k)} \leq d_1(t) \,|\, T=t\right) \\
&\quad + (k-1)\, P_{\underline{\mu}}\left(\bar{X}_{(k)} - \bar{X}_{(1)} \leq d_1(t) \,|\, T=t\right) \\
&\leq 1 + (k-1) \int_0^{t/n} \int_0^{x+d_1(t)} f_{\bar{X}_{(k)}|T/n}(y|t/n)\, f_{\bar{X}_{(1)}|T/n}(x|t/n)\,dy\,dx
\end{aligned}
\tag{2.14}
$$

where

$$
f_{\bar{X}_{(1)}|T/n}(x|t/n) = \int_0^{t/n} f_{\bar{X}_{(1)},T/n-\bar{X}_{(k)}}(x,r) f_{\bar{X}_{(k)}}(t-r)\,dr / f_{T/n}(t/n),
$$

$$
f_{\bar{X}_{(k)}|T/n}(x|t/n) = f_{\bar{X}_{(k)}}(x) f_{T/n-\bar{X}_{(k)}} / f_{T/n}\ (t/n),
$$

and

$$
f_{T/n}(t/n) = \int_0^{t/n} f_{T/n-\bar{X}_{(k)}}(r) f_{\bar{X}_{(k)}}(t/n-r)\,dr.
$$

Note that $f_{T/n-\bar{X}_{(k)}}(\cdot)$ is the p.d.f. of $I((k-1)\mu, n(k-1)^2\lambda)$.

Remark: The density function of statistic T in $\Omega(\delta)$ or in any other non-homogeneous space is difficult to evaluate in an exact form. One of the reasons for such difficulty is that the inverse Gaussian random variables have only restrictive additive property as explained below (see Chhikara and Folks (1975)). We know that if $X_1, X_2, \ldots, X_k$ are independent inverse Gaussian variables with parameter μ_i and λ_i, then $\Sigma X_i \sim I(\Sigma\mu_i, \xi(\Sigma\mu_i)^2)$ if and only if $\lambda_i/\mu_i^2 = \xi$ for all i. The sufficient part was shown by Tweedie and the necessary part was given by Chhikara (1972) and Shuster and Miura (1972).

Let $\underline{X}=(X_{11},\ldots,X_{1n_1},X_{21},\ldots,X_{2n_2},\ldots,X_{k1},\ldots,X_{kn_k})$ be a random vector, then $\underline{X}=\underline{x}\in\mathbb{R}^N$, where $N=n_1+\ldots+n_k$. A selection rule can be denoted by $\varphi(\underline{x})=(\varphi_1(\underline{x}),\ldots,\varphi_k(\underline{x}))$, where $\varphi_i(\underline{x})$: $\mathbb{R}^N \to [0,1]$ is the probability that π_i is included in the selected subset when $\underline{X}=\underline{x}$ is observed. Similarly, a conditional selection rule can be denoted by $\varphi^{\underline{T}}(\underline{X})=(\varphi_1^{\underline{T}}(x),\ldots,\varphi_k^{\underline{T}}(x))$, where $\varphi_i^{\underline{T}}(x)$ is the conditional probability that π_i is included in the selected subset, given $\underline{T}=\underline{t}$, when $\underline{X}=\underline{x}$ is observed. It is easy to see that $\underline{T}=T$, $\varphi_i^T(\underline{x})=0$ or 1 and $\varphi^T(\underline{x})\neq\underline{0}$ when rule R_1 is used.

Definition: A selection rule is scale invariant if for every $\underline{x}\in\mathbb{R}^N$, for every real number $c>0$ and for every $i=1,\ldots,k$, $\varphi_i(c\underline{x})=\varphi_i(\underline{x})$.

We may define scale invariance for conditional selection procedures

in a similar way, then we have the following theorem:

Theorem 2.3. With equal sample size, the procedure with constant $d_1(t)$ given by Theorem 2.1. or Theorem 2.2. is scale invariant.

Proof: From Corollary 2.1. and Corollary 2.2. we see that the constants $d_1(t)$ given by Theorem 2.1. or Theorem 2.2. have the property

$$d_1(ct)=cd_1(t) \text{ for all } c>0,$$

so we have

$$\varphi_i^{T(c\underline{x})}(c\underline{x})=\varphi_i^{T}(\underline{x}) \text{ for all } c>0, \text{ and } i=1,\ldots,k,$$

hence the theorem is proved.

2.4 Applications to a Test of Homogeneity for $\mu_1=\ldots=\mu_k$

When the common shape parameter λ is known, for the problem of the test of homogeneity of k inverse Gaussian populations, i.e. the test of hypothesis:

$$\text{null hypothesis } H_0:\mu_1=\mu_2=\ldots=\mu_k$$

$$\text{versus } H_1:\mu_i\text{'s are not all equal,}$$

we propose the following conditional procedure (at level α) $\Psi(T)$.
The procedure $\Psi(T)$ is:

The null hypothesis H_0 is rejected if and only if $\bar{X}_{[k]}-\bar{X}_{[1]}>d_1(t)$, given $T=t$, where $d_1(t)$ is given by (2.12) or (2.13) with $P^*=1-\alpha/k$.

It is easy to see that the procedure $\Psi(t)$ has the probability of

of type-one error less than α, since under null hypothesis

$$
\begin{aligned}
& P_{\underline{\mu}}(\bar{X}_{[k]}-\bar{X}_{[1]}>d_1(t)|T=t) \\
&= P(\bar{X}_{[1]}<\bar{X}_{[k]}-d_1(t)|T=t) \\
&= P(\bar{X}_j<\bar{X}_{[k]}-d_1(t) \text{ for some } j|T=t) \\
&\leq \sum_{j=1}^{k} P(\bar{X}_j<\bar{X}_{[k]}-d_1(t)|T=t) \\
&= k-kP(\bar{X}_1>\bar{X}_{[k]}-d_1(t)|T=t) \\
&= k-kP^* \\
&= \alpha.
\end{aligned}
$$

For k=2, it has been shown that the procedure $\Psi(t)$ is an UMP unbiased test, see Chhikara (1975) (also Lehmann (1959)).

3. Selection of the Inverse Gaussian Population with the Largest Mean when the Common Shape Parameter λ is unknown

With the same notations as before and let $V=\sum_{i=1}^{k}\sum_{j=1}^{n_i} X_{ij}^{-1}$.

3.1. A Conditional Selection R_2

R_2: Select the population π_i if and only if

$$\bar{X}_i \geq \max_{1\leq j\leq k} \bar{X}_j - d_2(t,v), \text{ given } T=t,\ V=v$$

where $t>0$, $v>0$; $d_2(t,v)$ is the smallest positive values chosen to satisfy the P*-condition.

For k=2, we have the following theorem:

Theorem 3.1. Given $\frac{1}{2} < P^* < 1$, $k = 2$, $T = t$ and $V = v$, let

$$h(u) = \frac{[n_1 n_2 (n_1+n_2-2)]^{\frac{1}{2}} (n_1+n_2)u}{\{[tv-(n_1+n_2)^2](T+n_2u)(T-n_1u)\}^{\frac{1}{2}}}$$

$$\times \left[1 - \frac{n_1 n_2 (n_1+n_2)^2 u^2}{[tv-(n_1+n_2)^2](t+n_2u)(t-n_1u)} \right]^{-\frac{1}{2}} , \qquad (3.1)$$

where

$$-1 \leq \frac{(n_1 n_2)^{\frac{1}{2}}(n_1+n_2)u}{\{[tv-(n_1+n_2)^2](t+n_2u)(t-n_1u)\}^{\frac{1}{2}}}$$

and let

$$d_2(t,v) = h^{-1}(c(t,v)), \qquad (3.2)$$

where the constant $c \equiv c(t,v)$ is determined by

$$H_{t;n_1+n_2-2}(c) + \frac{n_2-n_1}{n_1+n_2} \left[\frac{tv-(n_1-n_2)^2}{tv-(n_1+n_2)^2} \right]^{(n_1+n_2-3)/2} \left(1-H_{t;n_1+n_2-2}(c')\right) = P^* , \qquad (3.3)$$

where

$$c' = \{c^2 + 4n_1n_2(n_1+n_2-2)/[tv-(n_1-n_2)^2]\}^{\frac{1}{2}}$$

and $H_{t;n_1+n_2-2}$ denotes the c.d.f. of students's t-distribution with n_1+n_2-2 degrees of freedom. Then $\inf_{\Omega} P(CS|R_2) = \inf_{\Omega_0} P(CS|R_2) = P^*$.

Proof: For fixed t and v, h(u) is a monotone nondecreasing function in u, hence h^{-1} exists and

$$h^{-1}(w) = \{(n_2-n_1)\, ty^2 + ty[(n_1+n_2)^2 y^2 + 4]^{\frac{1}{2}}\}/2(1+n_1n_2y^2) \qquad (3.4)$$

where $y = w[tv-(n_1+n_2)^2]^{\frac{1}{2}}/(n_1n_2)^{\frac{1}{2}}(n_1+n_2)[n_1+n_2-2+w^2]^{\frac{1}{2}}$.

With the same notations as that in Section 2, it follows from an argument similar to that as in Lehmann (1959) [see P.136] that

$$\begin{aligned}
&\inf_{\Omega} P_{\underline{\mu}} (CS|R_2) \\
&= \inf_{\Omega} P_{\underline{\mu}} (\bar{X}_{(1)}-\bar{X}_{(2)} \leq d_2(t,v)|T=t, V=v) \\
&= P_{\theta=0} (\bar{X}_{(1)}-\bar{X}_{(2)} \leq d_2(t,v)|T=t, V=v) \\
&= P_{\theta=0} (h(U) \leq h(d_2(t,v))|T=t, V=v) \\
&= P_{\theta=0} (h(U) \leq c(t,v)|T=t, V=v) \\
&= P^*,
\end{aligned}$$

by the definition of c(t,v) (see Chhikara (1975), p.81).

Corollary 3.1. In Theorem 3.1, if we have a common sample size, say $n_1=n_2=n$, then the constant c is determined by

$$H_{t;2n-2}(c) = P^* \quad \text{i.e.} \quad c = H^{-1}_{t;2n-2}(P^*). \tag{3.5}$$

Thus c is given by the P*-percentile of a t-distribution with 2n-2 degrees of freedom. Consequently,

$$d_2(t,v) = \frac{tc(tv-4n^2)^{\frac{1}{2}}}{n[c^2tv+4n^2(2n-2)]^{\frac{1}{2}}} \tag{3.6}$$

which is increasing in t and v, if n is fixed. Note that $d_2(t,v) \to 0$ as $n \to \infty$ if both $t = O(n)$ and $v = O(n)$.

Similar to Theorem 2.2., the following theorem gives a lower bound on the probability of a correct selection in case of $k \geq 3$ where the common shape parameter λ is unknown.

Theorem 3.2. For $k \geq 3$, given P^*, $1/k < P^* < 1$, $T=t$, $V=v$, suppose the common shape parameter λ is unknown. Let $P_1^* = 1 - \frac{1-P^*}{k-1}$ and let $d_{ij}^{(2)}(t,v)$ be the smallest value such that for any $\underline{\mu} \in \Omega_0 = \{\underline{\mu} | \mu_1 = \ldots = \mu_k > 0\}$

$$P_{\underline{\mu}}(U_{ij} < d_{ij}^{(2)}(r,s) | T_{ij}=r, V_{ij}=s) = P^*$$

where

$$U_{ij} = \bar{X}_{(i)} - \bar{X}_{(j)} ,$$

$$T_{ij} = n_{(i)}\bar{X}_{(i)} + n_{(j)}\bar{X}_{(j)} ,$$

and

$$V_{ij} = \sum_{\ell=1}^{n_{(i)}} X_{(i\ell)}^{-1} + \sum_{\ell=1}^{n_{(j)}} X_{(j\ell)}^{-1} , \quad 1 \leq i \neq j \leq k.$$

Let $d_2(t,v) = \max\{d_{ij}^{(2)}(r,s) | 1 \leq i \neq j \leq k,\ 0 < r \leq t,\ 0 < s \leq v\}$.

Then $\inf_{\Omega} P(CS|R_2) \geq P^*$.

Proof: Proof is almost the same as the proof of Theorem 2.2., hence it is omitted.

Corollary 3.2. In Theorem 3.2., if $n_1=n_2=\ldots=n_k=n$ then $d_2(t,v)$ is given by (3.7) with $c = H_{t;2n-2}^{-1}(P_1^*)$; note that the procedure R_2 is scale invariant.

4. Selection From Inverse Gaussian Populations in Terms of the Shape Parameters

In ranking inverse Gaussian populations in terms of their shape parameters, we defined the best population as the one associated with $\lambda_{[k]}$. With the same assumptions as given in Section 1, for all $i = 1,\ldots,k$ let

$$S_i^2 = \frac{1}{\mu_i^2} \sum_{j=1}^{n_i} \frac{(X_{ij}-\mu_i)^2}{X_{ij}} , \quad \text{if } \mu_i \text{ is known,} \tag{4.1}$$

$$= \sum_{j=1}^{n_i} \left(\frac{1}{X_{ij}} - \frac{1}{\bar{X}_i} \right) , \quad \text{if } \mu_i \text{ is unknown,} \tag{4.2}$$

then $\lambda_i S_i^2$ has a chi-square distribution $\chi^2_{\nu_i}$ with ν_i degrees of freedom where $\nu_i = n_i$ or $n_i - 1$ depending on the case whether μ_i is known or unknown. Therefore, there is no need to deal with the cases of known or unknown means separately.

Using statistics S_i^2 , i=1,...k, the problem of selecting from inverse Gaussian populations in terms of shape parameter is equivalent to the problem of selection from normal populations in terms of variances (see Gupta and Panchapakesan (1979)).

For an equal sample size case, parallel to the rule of Gupta and Sobel (1962a), we propose a rule R_3.

R_3: Select population π_i if and only if

$$S_i^2 < C^{-1} S_{[1]}^2 \quad ,$$

where $0 < C \equiv C(\nu,k,P^*) \leq 1$ is determined so that the P*-condition is satisfied. Here $\Omega = \{\underline{\lambda} | \lambda_i > 0\}$. It is easy to see that the infimum of $P(CS|R_3)$ occurs when $\lambda_{[1]} = \lambda_{[2]} = \ldots = \lambda_{[k]}$ and is independent of the common value. Thus we have

$$\inf_{\Omega} P_{\underline{\lambda}}(CS|R_3) = \int_0^\infty [1-\chi^2_\nu(cx)]^{k-1} \, d\chi^2_\nu(x) \qquad (4.3)$$

and also we have $\sup_{\Omega} E_{\underline{\lambda}}(S|R_3) = kP^*$.

The c-values can be found in Gupta and Sobel (1962b) for k=2(1) 11, ν=2(2) 50 and P*=0.75, 0.9, 0.95 and 0.99.

For an unequal sample size case, some results are available in Gupta and Huang (1976) [see also Gupta and Panchapakesan (1979)].

Remark 4.1. Let $\bar{X}_{iH} \equiv \left[\left(\sum_{j=1}^{n_i} X_{ij}^{-1}\right)/n_i\right]^{-1}$ be the harmonic sample mean of π_i and let

$$\tilde{S}_i^2 = \bar{X}_{iH}^{-1} + \frac{2}{\mu_i^2}(\bar{X} - 2\mu_i), \text{ if } \mu_i \text{ is known,}$$

$$= \bar{X}_{iH}^{-1} - \bar{X}_i^{-1} \text{ , if } \mu_i \text{ is unknown,}$$

then using the statistic $\tilde{S}_i^2$ is equivalent to using the statistics S_i^2 , i=1,...k, since $S_i^2 = n_i \tilde{S}_i^2$ for all i.

Remark 4.2. It should be pointed out that the problem of selecting the inverse Gaussian populations in terms of $\lambda_{[1]}$ is equivalent to the problem of selecting from gamma populations with densities $\frac{1}{\Gamma(\nu)} \frac{1}{\theta_i} e^{-x/\theta_i} (\frac{x}{\theta_i})^{\nu-1}$, those that have large values of θ_i. This problem has been solved in Gupta (1963), where appropriate tables are also provided.

References

Banerjee, A. K. and Bhattacharyya, G. K. (1976). A purchase incidence model with inverse Gaussian interpurchase times. J. Amer. Statist. Assn., 71, 823-829.

Chhikara, R. S. (1972). Statistical inference related to the inverse Gaussian distribution. Ph.D. Dissertation, Oklahoma State University.

Chhikara, R. S. (1975). Optimum tests for the comparison of two inverse Gaussian means. Austral J. Statist., 17, 77-83.

Chhikara, R. S. and Folks, J. L. (1974). Estimation of the inverse Gaussian distribution function. J. Amer. Statist. Assn., 69, 250-254.

Chhikara, R. S. and Folks, J. L. (1975). Statistical distributions related to the inverse Gaussian. Commun. in Statist., 4, 1081-1091.

Chhikara, R. S. and Folks, J. L. (1977). The inverse Gaussian distribution as a lifetime model. Technometrics, 19, 461-468.

Cox, D. R. and Miller, H. D. (1965). The Theory of Stochastic Processes. London: Methuen.

Folks, J. L. and Chhikara, R. S. (1978). The inverse Gaussian distribution and its statistical application -- a review. J. R. Statist. Soc. B, 40, No. 3, 263-289,

Gupta, S. S. (1963). On a selection and ranking procedure for gamma populations. Ann. Inst. Statist. Math., 14, 199-216.

Gupta, S. S. and Huang, D. Y. (1976). Selection procedures for the means and variances of normal populations: unequal sample size case. Samkhyā Ser. B, 38, 112-128.

Gupta, S. S. and Panchapakesan, S. (1979). Multiple Decision Procedures: Theory and Methodology of Selection and Ranking Populations. John Wiley, New York.

Gupta, S. S. and Sobel, M. (1962a). On selecting a subset containing the population with the smallest variance. Biometrika, 49, 495-507.

Gupta, S. S. and Sobel, M. (1962b). On the smallest of several correlated F-statistics. Biometrika, 49, 509-523.

Hasofer, A. M. (1964). A dam with inverse Gaussian input. Proc. Camb. Phil. Soc., 60, 931-933.

Lehmann, E. L. (1959). Testing Statistical Hypotheses. Wiley, New York.

Marcus, A. H. (1975). Some exact distributions in traffic noise theory. Adv. Appl. Prob., 7, 593-606.

Marcus, A. H. (1976). Power sum distributions: an easier approach using the Wald distribution. J. Amer. Statist. Assn., 71, 237-238.

Schrödinger, E. (1915). Zur Theorie der Fall-und Steigversuche an Teilchen mit Brownscher Bewegung. Physikalische Zeitschrift, 16, 289-295.

Shuster, J. J. (1968). On the inverse Gaussian distribution function. J. Amer. Statist. Assn., 63, 1514-1516.

Shuster, J. J. and Miura, C. (1972). Two way analysis of reciprocals. Biometrika, 59, 478-481.

Tweedie, M. C. K. (1956). Some statistical properties of inverse Gaussian distributions. Virginia J. Sci., 7, 160-165.

Tweedie, M. C. K. (1957a). Statistical properties of inverse Gaussian distributions I. Ann. Math. Statist., 28, 362-377.

Tweedie, M. C. K. (1957b). Statistical properties of inverse Gaussian distributions II. Ann. Math. Statist., 28, 696-705.

Wald, A. (1947). Sequential Analysis. Wiley, New York.

REPORT DOCUMENTATION PAGE

READ INSTRUCTIONS BEFORE COMPLETING FORM

1. REPORT NUMBER: Technical Report #82-7
2. GOVT ACCESSION NO.:
3. RECIPIENT'S CATALOG NUMBER:

4. TITLE (and Subtitle): ON SUBSET SELECTION PROCEDURES FOR INVERSE GAUSSIAN POPULATIONS

5. TYPE OF REPORT & PERIOD COVERED: Technical

6. PERFORMING ORG. REPORT NUMBER: Technical Report #82-7

7. AUTHOR(s): Shanti S. Gupta and Hwa-Ming Yang

8. CONTRACT OR GRANT NUMBER(s): N00014-75-C-0455

9. PERFORMING ORGANIZATION NAME AND ADDRESS
Purdue University
Department of Statistics
West Lafayette, IN 47907

10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS:

11. CONTROLLING OFFICE NAME AND ADDRESS
Office of Naval Research
Washington, DC

12. REPORT DATE: March 1982

13. NUMBER OF PAGES: 20

14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office):

15. SECURITY CLASS. (of this report): Unclassified

15a. DECLASSIFICATION/DOWNGRADING SCHEDULE:

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release, distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

Brownian Motion, Wiener Processes, First passage time, Positive drift Inverse Gaussian distributions, Subset selection procedures, Conditional selection procedure, Exponential family, Gamma distributions.

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

The inverse Gaussian or the first passage time probability distribution for Brownian motion with a drift is particularly important for modeling and interpreting observed distributions of time intervals in many different fields of research. In this paper we deal with the problem of selecting a subset of k inverse Gaussian populations which includes the "best" population, i.e. the (unknown) population which is associated with the largest value of the unknown means. The shape parameters of the inverse Gaussian distributions are assumed to be equal for all the k populations. When the common shape parameter is known, a procedure R_1 is defined

and studied which selects a subset which is nonempty, small in size and just large enough to guarantee that it includes the best population with a preassigned probability regardless of the true unknown values of the means. For the case when the common shape parameter is unknown a procedure R_2 is proposed. For the procedures R_1 and R_2, we obtain exact results for $k = 2$ concerning the infimum of the probability of a correct selection. For $k \geq 3$ a lower bound on the probability of a correct selection is derived for each case. Formulas for the constants d_1 and d_2 which are necessary to carry out the procedures R_1 and R_2, respectively, are obtained. An upper bound on the expected number of populations retained in the selected subset is given

If the best population is defined as the one associated with the largest shape parameter, it is shown that with a suitably chosen statistic, this problem coincides with the problem of selecting a subset of k normal populations which includes the population with the smallest variance. Similarly, for the selection of a subset containing the smallest shape parameter, the problem reduces to selection in terms of the largest scale parameter of the gamma distributions.