Ph.D. MATHEMATICS

JONATHAN SHUSTER, M.Sc.

PROPERTIES OF THE INVERSE GAUSSIAN DISTRIBUTION

This dissertation is divided into two sections. The first, Chapters 4 to 9, concentrates on the distribution theory of the Inverse Gaussian distribution. The author develops a general limit theorem for convergence in law to this distribution. This result is of interest to nonparametric statisticians. Other results from this section include: a convenient method of obtaining percentage points; Bayes estimates of parameters; and four original characterizations of the distribution. A definition of the Multivariate Inverse Gaussian distribution (MVIG) is given. Necessary and sufficient conditions that all the marginal distributions of the MVIG be Inverse Gaussian are developed.

The second section, Chapters 10 to 12, deals with a general family of stochastic processes, of which the separable Inverse Gaussian and Poisson processes are members. The topic is treated from a measure theoretic point of view. Properties of the sample functions and Stieltjes stochastic integrals are treated in detail. Illustrative examples of how one's intuition, regarding path properties, can lead him astray, are included.

"PROPERTIES OF THE INVERSE GAUSSIAN DISTRIBUTION,"
by JONATHAN SHUSTER, MCGILL UNIVERSITY.

This manuscript is submitted to the Faculty of Graduate Studies and Research as partial fulfilment of the requirements of the degree of Doctor of Philosophy.

Signed,

Jon Shuster

Jonathan Shuster, M.Sc.
Dept. of Mathematics,
Feb. 20, 1969.

TABLE OF CONTENTS

TABLE OF CONTENTS, CONTINUED

THEOREM GUIDE

THEOREM GUIDE, CONTINUED

LEMMA GUIDE

EXAMPLE GUIDE

DEFINITION GUIDE

ACKNOWLEDGEMENTS

I would like to acknowledge several people for their generous assistance. First, I would like to thank my research director, Prof. V. Seshadri, who devoted countless hours to this project and its creation.

Secondly, I would like to express my gratitude to the National Research Council of Canada for their financial assistance in the form of a Postgraduate Scholarship.

Thirdly, My appreciation extends to the statistics seminar committee at McGill University for giving me the opportunity to give a survey seminar on the Inverse Gaussian distribution.

I would like to thank Prof. D. Dawson, for his suggestions with regard to notation of Chapters 10, 11, and 12; Prof M. Csorgo, for his suggestions in Chapter 4; and Prof. A. P. Dempster for his communications from Harvard.

Finally, I would like to thank my wife Sandra, who helped type this thesis, inspired this work, and proof read the prose sections of the dissertation.

Without the above assistance, completion of this thesis would not have been possible.

ORIGINAL RESULTS OF THIS THESIS

Apart from Chapters 1,2,3, and 10, to the best of the authors knowledge, each result is original.

Further, the author feels that:

(1) THEOREM 4.1 has great potential practical significance. Although, perhaps a shorter proof of the result <u>might</u> be possible, by bringing in Brownian Motion, the proof of Chapter 4 gives convergence rates.

(2) The author has shown a simple way to obtain percentage points of the Inverse Gaussian distribution, and investigated the structure of estimates of parameters.

(3) The author has also shown several sets of conditions which characterize the distribution. A workable definition of a multivariate Inverse Gaussian distribution is given, and investigated.

(4) Significant contributions to the field of Stochastic Processes appear in this work. The approach is perhaps different from the classical methods. The true essence of certain types of processes are indicated. Stochastic Integrals are defined, and their properties developed.

(5) Two research papers, by the author, solve thorny problems in the experimental area: Analysis of Reciprocals.

CHAPTER 1 : INTRODUCTION

Study of the Inverse Gaussian distribution was initiated by Tweedie [15], in 1957, although as early as 1915, Schrodinger showed that the distribution occurs as the first passage time of Brownian Motion with positive drift. It was Tweedie, who named the distribution, from the fact that the cumulant-generating functions of the Gaussian and Inverse Gaussian distributions are inverse functions to each other.

Khatri [7], in 1962, characterized the distribution with a rather remarkable theorem, based on the independence of two random variables. The author [12], in 1968, obtained a convenient method of obtaining the percentage points of the distribution. The most significant contributions to this field were made by Wasan [16], in 1966, in his work, "Monograph on Inverse Gaussian Distribution." Not only did he collect most of the original work mentioned above, but he developed a large quantity of new material, applicable in many areas of Statistics and Probability.

The first three chapters of this thesis serve as an introduction, and hence contain few original results. Chapter 2 is devoted to basic definitions and notation, while Chapter 3 includes results, quoted without proof, which can be found in the literature. In each case, a reference is given, as to where the proof may be found.

With the exception of a few results, given without proof, Chapters 4 through 9, contain, to the best of the author's knowledge, original research.

In Chapter 4, the author develops a situation, under very general conditions, in which Inverse Gaussian random variables occur.

Chapter 5 includes a simple proof of the author's result, [12], enabling one to find percentage points of the Inverse Gaussion distribution.

In Chapter 6, the author discusses confidence sets for the parameters of the distribution; while Chapter 7 treats Bayes estimates of parameters.

In Chapter 8, four characterizations of the Inverse Gaussian distribution are developed, three of which depend on a result of Prof. V. Seshadri, given in Chapter 5. Khatri's characterization is discussed in an example.

Chapter 9 is devoted to the definition of a multivariate Inverse Gaussian distribution. Lemma 9.1 gives necessary and sufficient conditions for sums of independently distributed, Inverse Gaussian random variables to be Inverse Gaussian distributed. This Lemma in turn gives necessary and sufficient conditions that the marginal distributions of a multivariate Inverse Gaussian random vector, all be univariate Inverse Gaussian distributions.

Chapters 10 to 12 deal with families of stochastic processes, of which the Inverse Gaussian Process is a member.

Chapter 10 is of an introductory nature, and as such, contains little original work. In Chapter 11, the author shows that for separable stochastic processes, such that $P(X_t \geqq X_s) = 1$ whenever $t > s$, almost every sample function is monotone non-decreasing. In Chapter 12, the author defines MISI stochastic processes. Stochastic integrals with respect to MISI processes are introduced. At the conclusion of the chapter, properties of the Inverse Gaussian stochastic process are listed.

Following Chapter 12, the author proposes certain conjectures and problems, which would make excellent research projects.

The author has also included two research papers, written after the typing of this manuscript had been completed. The results of these papers do not appear in the text of the thesis.

CHAPTER 2 : NOTATION AND DEFINITIONS

In this chapter, the reader will be introduced to the basic notation used throughout this thesis.

Notation Chart

1	$\{\ \}$	A set of points
2	$[a,b]$	$\{t : a \leqq t \leqq b\}$
3	(a,b)	$\{t : a < t < b\}$
4	Z^+	Set of positive integers
5	$[x]$	Largest integer $\leqq x$
6a	$X \sim F(x)$	Random variable X has distribution function F(x)
6b	$X \approx Y$	Random variables X and Y are identically distributed
7	$Y \sim \chi^2_n$	Y has the chi-square distribution with n degrees of freedom
8	$Y \sim F_{n,m}$	Y has the F-distribution with n and m degrees of freedom
9	$\chi^2_{n,\alpha}$	If $Y \sim \chi^2_n$, $P(Y > \chi^2_{n,\alpha}) = \alpha$
10	$F_{n,m,\alpha}$	If $Y \sim F_{n,m}$, $P(Y > F_{n,m,\alpha}) = \alpha$
11	$X_n \xrightarrow[L]{} F(x)$	As $n \longrightarrow \infty$, the distribution function of X_n converges to F(x), at every continuity point of F(x)

12	$\bar{X}$	Sample mean, a random variable
13	$\Gamma(\alpha)$	Gamma function
14	$\binom{n}{j}$	Binomial coefficient
15	iff	if and only if
16	$(\Omega, \mathcal{F}, P)$	Probability space
17	$E(X)$	Mathematical expectation of random variable X
18	a.e.	almost everywhere with respect to the measure P

Definitions

Definition 2.1:

A random variable X, has the Inverse Gaussian distribution with positive parameters μ and λ, if it has density function:

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

$$= 0 \qquad x < 0 .$$

The fact that X is an Inverse Gaussian random variable will be denoted by:

$X \sim IG(\cdot,\cdot)$.

The fact that X is an Inverse Gaussian random variable with parameters μ and λ, will be denoted by:

$X \sim IG(\mu,\lambda)$.

Definition 2.2:

If $X \sim IG(1,1)$, X is called a standard Inverse Gaussian random variable.

Definition 2.3:

The cumulative distribution function of an Inverse Gaussian random variable will be denoted by:

$$F(x;\mu,\lambda) = \int_0^x \left(\frac{\lambda}{2\pi t^3}\right)^{\frac{1}{2}} \exp\left[-\frac{\lambda(t-\mu)^2}{2\mu^2 t}\right] dt \qquad x > 0,$$

$$= \quad 0 \qquad x < 0.$$

Definition 2.4: (Infinitely Divisible Law)

A distribution function, $F(x)$, is an infinitely divisible law, if for every positive integer n, there exists a distribution function $F_n(x)$, such that if $X_1, X_2, \ldots, X_n$ is a random sample from $F_n(x)$, then

$$(X_1 + X_2 + \ldots + X_n) \sim F(x).$$

CHAPTER 3 : IMPORTANT THEOREMS FROM THE LITERATURE, CONCERNING INVERSE GAUSSIAN RANDOM VARIABLES

In this chapter, some results, whose applications will appear in later chapters, will be listed without proof. In each case, the author will give the reference where the proof may be found. Some results will be slightly generalized.

THEOREM 3.1: (Characteristic function of $F(x;\mu,\lambda)$)

If $X \sim IG(\mu,\lambda)$, and if θ is any real number, then

$$E(\exp(i\theta X)) = \exp\left[\frac{\lambda}{\mu}\left(1 - \left(1 - \frac{2i\theta\mu^2}{\lambda}\right)^{\frac{1}{2}}\right)\right]$$

Proof: See [16-1].

THEOREM 3.2:

If $X \sim IG(\mu,\lambda)$, then for every positive integer n, $E(X^n)$ and $E(X^{-n})$ exist, and in particular,

(a) $E((X/\mu)^n) = E((X/\mu)^{1-n})$

(b) $E(X) = \mu$

(c) $E(X^2) = \mu^2 + \mu^3/\lambda$

(d) $Var(X) = \mu^3/\lambda$

Proof: See [15-1] .

THEOREM 3.3:

Let $X \sim IG(\mu,\lambda)$. Let U,V be independent random variables with $\lambda U \sim \chi^2_1$ and $V \sim IG(1/\mu, \lambda/\mu^2)$. Then $1/X \overset{\sim}{\sim} U+V$.

Proof: See [15-2] .

THEOREM 3.4: (Corollary to THEOREM 3.3)

Let X,T be independent random variables with $X \sim IG(1,1)$ and $T \sim \chi^2_1$. Then $1/X \approx X+T$.

In this case, 1/X is the convolution of a standard Inverse Gaussian random variable and a chi-square-one random variable.

THEOREM 3.5:

If $X \sim IG(\mu,\lambda)$, and $c > 0$, then $cX \sim IG(c\mu, c\lambda)$.

Proof: By THEOREM 3.1,

$$E(\exp(i\theta cX) = \exp\left[\frac{c\lambda}{c\mu}\left(1-\left(1-\frac{2i\theta(c\mu)^2}{c\lambda}\right)^{\frac{1}{2}}\right)\right].$$

By uniqueness of Fourier Transform, the desired result follows. This completes the proof.

THEOREM 3.6:

Let $X_1, X_2, \ldots, X_n$ be a random sample from $F(x;\mu,\lambda)$.

Then (a) $(X_1+X_2+\ldots+X_n) \sim IG(n\mu, n^2\lambda)$

and (b) $\overline{X} \sim IG(\mu, n\lambda)$

Proof: See [16-2].

THEOREM 3.7:

Let $X_1, X_2, \ldots, X_n$ be a random sample from $F(x;\mu,\lambda)$.

Let $$Y = \lambda \sum_{j=1}^{n} \left(\frac{1}{X_j} - \frac{1}{\overline{X}}\right).$$

Then: (a) Y and $\overline{X}$ are stochastically independent

and (b) $Y \sim \chi^2_{n-1}$

Proof: See [16-3]

THEOREM 3.8 : (Khatri's Characterization)

Let $X_1, X_2, \ldots, X_n$ be a random sample of size $n \geq 2$, from an absolutely continuous distribution. Let $E(X)$, $E(X^2)$, $E(X^{-1})$, and $E(\overline{X}^{-1})$ be finite and not zero.

Then a necessary and sufficient condition that $X_j \sim IG(\cdot,\cdot)$, $j = 1, \ldots, n$, is that

$$\overline{X} \quad \text{and} \quad \sum_{j=1}^{n}\left(\frac{1}{X_j} - \frac{1}{\overline{X}}\right) \quad \text{are independent.}$$

Proof: See [7-1]

THEOREM 3.9 :

$F(X;\mu,\lambda)$ is an infinitely divisible law.

Proof: Let n be an arbitrary integer, and let $X_1, X_2, \ldots, X_n$ be a random sample from $F(X;\mu/n,\lambda/n^2)$. By THEOREM 3.6(a), $X_1 + X_2 + \cdots + X_n \sim IG(\mu,\lambda)$. Hence, by Definition 2.4, $F(X;\mu,\lambda)$ is an infinitely divisible law.

This completes the proof.

THEOREM 3.10: (Kolmogorov)

If $F(x)$ is an infinitely divisible law, with finite second moment, its characteristic function, $f(\theta)$, can be represented uniquely by:

$$\log(f(\theta)) = i\gamma\theta + \int_{-\infty}^{\infty} (\exp(i\theta u)-1-i\theta u)(1/u^2)dK(u)$$

where γ is a constant, and $K(u)$ is a non-decreasing function such that $K(-\infty) = 0$ and $K(\infty) < \infty$.

Proof: See [6-1].

Wasan [16-4], found the Kolmogorov representation of the characteristic function of $F(x;t,t^2)$. The author will obtain the representation for arbitrary Inverse Gaussian laws, by a different technique.

Lemma 3.1:

Let $F(x)$ be an infinitely divisible law with finite second moment, and characteristic function $f(\theta)$.

Then if $K(u)$ is the function given in the Kolmogorov representation of $f(\theta)$,

$$\frac{d^2\ln(f(\theta))}{d\theta^2} = -\int_{-\infty}^{\infty} (\exp(i\theta u)dK(u) .$$

Proof: Since $F(x)$ has finite second moment, $f(\theta)$, and hence $\ln(f(\theta))$, is twice differentiable. Differentiating twice under the integral sign gives the desired result.

This completes the proof.

THEOREM 3.11 :

The Kolmogorov representation of the characteristic function of $F(x;\mu,\lambda)$ is

$$\gamma = \mu \qquad dK(u) = au^{\frac{1}{2}}\exp(-bu)du \ , \ u > 0,$$
$$= 0 \qquad , \ u < 0,$$

where $a = (\lambda/2\pi)^{\frac{1}{2}}$ and $\frac{1}{2}b = \lambda/2\mu^2$.

Proof: By THEOREM 3.1, with $f(\theta)$ the characteristic function of $F(x;\mu,\lambda)$,

$$\log(f(\theta)) = (\lambda/\mu)(1 - (1 - \frac{2i\theta\mu^2}{\lambda})^{\frac{1}{2}}) \text{ , and hence}$$

$$- \frac{d^2\log(f(\theta))}{d\theta^2} = \frac{\mu^3}{\lambda}(1 - \frac{2i\theta\mu^2}{\lambda})^{-3/2} \qquad .. (1)$$

Using the elementary identity

$$\int_0^\infty (\beta^{-\alpha}/\Gamma(\alpha))y^{\alpha-1}\exp(-y/\beta)\exp(i\theta y)dy = (1 - i\beta\theta)^{-\alpha}$$

for positive constants α and β, one has

$$(\mu^3/\lambda)(1 - \frac{2i\theta\mu^2}{\lambda})^{-3/2} =$$

$$\int_0^\infty (\mu^3/\lambda)(\lambda/\mu^2)^{3/2}(2\pi)^{-\frac{1}{2}}y^{\frac{1}{2}}\exp(-\lambda y/\mu^2)\exp(i\theta y)dy =$$

$$\int_0^\infty (\lambda y/2\pi)^{\frac{1}{2}}\exp(-\lambda y/\mu^2)\exp(i\theta y)dy \ .$$

Setting

$$k(y) = (\lambda y/2\pi)^{\frac{1}{2}}\exp(-y\lambda/\mu^2) \qquad y > 0 \ ,$$
$$= 0 \qquad y < 0 \ ,$$

one has by .. (1):

$$- \frac{d^2\log(f(\theta))}{d\theta^2} = \int_{-\infty}^\infty \exp(i\theta u)k(u)du \ .$$

Hence, by Lemma 3.1:

$$\int_{-\infty}^{\infty} (\exp(i\theta u)dK(u) = \int_{-\infty}^{\infty} \exp(i\theta u)k(u)du \ .$$

From the definitions of k(u),a,and b, and the uniqueness of Fourier Transform, one has:

$$dK(u) = au^{\frac{1}{2}}\exp(-bu)du \qquad u > 0$$
$$= 0 \qquad u \leq 0 \ .$$

Thus, it remains only to prove: $\gamma = \mu$.

The Kolmogorov representation of $f(\theta)$ is:

$$f(\theta) = \exp\left[i\gamma\theta + \int_{-\infty}^{\infty} (\exp(i\theta u) - 1 - i\theta u)k(u)(1/u^2)du\right]$$

with k(u) as above.

$$\frac{df(\theta)}{d\theta} = \left[i\gamma + i\int_0^{\infty} (\exp(i\theta u)-1)\frac{k(u)}{u}du\right] f(\theta) \ .$$

Thus since $F(x;\mu,\lambda)$ has first moment μ,

$$\lim_{\theta\to 0} \frac{df(\theta)}{d\theta} = \lim_{\theta\to 0}\left[i\gamma + i\int_0^{\infty} (\exp(i\theta u) -1)\frac{k(u)}{u}du\right]$$

$$= i\mu \ .$$

<u>Remark</u>: In order to show that $\gamma = \mu$, it is therefore enough to show:

$$\lim_{\theta\to 0} \int_0^{\infty} (\exp(i\theta u)-1)\frac{k(u)}{u}du = 0 \ .$$

That is, it is sufficient to show:

$$\lim_{\theta\to 0} \int_0^{\infty} (\exp(i\theta u)-1)au^{-\frac{1}{2}}\exp(-bu)du = 0 \qquad .. (2)$$

Since $|(\exp(i\theta u)-1)| \leq 2$ for all real θ and u,

$$|(\exp(i\theta u)-1)au^{-\frac{1}{2}}\exp(-bu)| \leq 2au^{-\frac{1}{2}}\exp(-bu) \qquad u \geq 0 \ .$$

But since $\int_0^{\infty} 2au^{-\frac{1}{2}}\exp(-bu)du < \infty$, one has by the Lebesgue dominated convergence theorem [9-1]:

$$\lim_{\theta\to 0} \int_0^\infty (\exp(i\theta u)-1)au^{-\frac{1}{2}}\exp(-bu)du =$$

$$\int_0^\infty \lim_{\theta\to 0} (\exp(i\theta u)-1)au^{-\frac{1}{2}}\exp(-bu)du = 0.$$

Therefore, .. (2) holds, and hence by the remark,

$\gamma = \mu$.

This completes the proof.

CHAPTER 4 : A DERIVATION OF INVERSE GAUSSIAN RANDOM VARIABLES

To date, several derivations of Inverse Gaussian random variables have been obtained. Four such constructions can be found in [16]. In this chapter, the author will develop a new construction, with applications in nonparametric statistics.

Let $X_1,\ldots,X_n$ be an ordered sample from an absolutely continuous distribution $F(x)$. Label a partition of the real line as follows:

$$I_0 = (-\infty, X_1)$$

$$I_j = [X_j, X_{j+1}) \qquad j = 1,2,\ldots,n-1$$

$$I_n = [X_n, \infty)$$

Definition 4.1: The empirical distribution function, $F_n(x)$ is defined as follows:

$$F_n(x) = j/n \qquad x \in I_j \quad , \quad j = 0,1,\ldots,n.$$

Definition 4.2: (notation)

$$Y_n(d) = \min\left\{ j : \sup_{x\in I_j} (F(x)-F_n(x)) > d \right\}, \text{ if } \sup_{x\in R_1} (F(x)-F_n(x)) > d,$$

$$= n+1 \qquad \text{otherwise.}$$

Here, R_1 denotes the set of real numbers.

Definition 4.3: Let $Z_n(d)$ be the following conditional random variable:

$$Z_n(d) = \left(Y_n(d) \,\middle|\, \sup_{x \in R_1} (F(x) - F_n(x) > d \right) .$$

That is, $Z_n(d) = j$ is the event that the least x to violate the inequality $F(x)-F_n(x) \leq d$, given that the inequality is somewhere violated, occurs in the interval I_j, provided $j \leq n$.

THEOREM 4.1:

Let $\lambda > 0$, and $W_n = \dfrac{Z_n(\lambda/n^{\frac{1}{2}})}{n - Z_n(\lambda/n^{\frac{1}{2}})}$. Then

$$W_n \xrightarrow[L]{} IG(1, \lambda^2).$$

THEOREM 4.1 will be proved in a sequence of lemmas.

Lemma 4.1: The distribution of $Y_n(d)$.

$$P(Y_n(d) = r) = d\binom{n}{r}\left(d + \frac{r}{n}\right)^{r-1}\left(1 - d - \frac{r}{n}\right)^{n-r}$$

where $r = 0,1,\ldots,[n(1-d)]$.

Proof: Since $Y_n(d)$ is a function of a distribution free quantity, it is itself distribution free. Thus, without loss of generality, one can choose

$$\begin{aligned} F(x) &= 0 && x < 0 \\ &= x && 0 < x < 1 \\ &= 1 && 1 < x . \end{aligned}$$

By Definition 4.2, one readily sees:

$Y_n(d) = 0$ iff no observations in $[0,d)$.

$Y_n(d) = 1$ iff one observation in $[0,d)$ and no observation in $[d, d+\frac{1}{n})$

For $2 \leqq r \leqq n(1-d)$, $Y_n(d) = r$ iff

(a) at least $k+1$ observations in $[0, d+\frac{k}{n})$, $k = 0,1,\ldots,r-2$

(b) exactly r observations in $[0, d+\frac{r-1}{n})$, and

(c) no observation in $[d+\frac{r-1}{n}, d+\frac{r}{n})$.

By binomial probabilities, the desired result is easily verified for $r = 0$ or 1. For $2 \leqq r \leqq n(1-d)$,

$P(Y_n(d) = r) =$

$$(n!)\int_0^d \int_{Y_1}^{\frac{1}{n}+d}\int \cdots \int_{Y_{r-2}}^{\frac{r-2}{n}+d}\int_{Y_{r-1}}^{\frac{r-1}{n}+d}\int_{\frac{r}{n}+d}^{1}\int_{Y_{r+1}}^{1}\cdots\int_{Y_{n-1}}^{1} dY_n \ldots dY_1$$

$$= \quad d\binom{n}{r}\left(d+\frac{r}{n}\right)^{r-1}\left(1-d-\frac{r}{n}\right)^{n-r} \quad \text{by [2].}$$

This exact integral is evaluated in [2-1].

This completes the proof.

Lemma 4.2: For $\lambda \geqslant 0$ and $0 < y < 1$, let

$$f_n(\lambda, y) = \left(\frac{\lambda}{n^{\frac{1}{2}}y} + 1\right)^{ny}\left(1 - \frac{\lambda}{n^{\frac{1}{2}}(1-y)}\right)^{n(1-y)}$$

Let $f(\lambda,y) = \exp(-\frac{\lambda}{2y(1-y)})$, and A be any rectangle of finite Lebesgue measure, of the form:

$$A = \left\{ 0 \leqq \lambda \leqq b\ ,\ c \leqq y \leqq d \right\} \subset [0,\infty)X(0,1).$$

Then $f_n \longrightarrow f$, uniformly on A, as $n \longrightarrow \infty$.

Proof: Let

$$h_n(\lambda,y) = \frac{\frac{\partial}{\partial\lambda} f_n(\lambda,y)}{f_n(\lambda,y)} \quad . \text{ Hence,}$$

$$h_n(\lambda,y) = \left(-\frac{\lambda}{1-y} - \frac{\lambda}{y}\right)\left(\frac{\lambda}{n^{\frac{1}{2}}y} + 1\right)^{-1}\left(1 - \frac{\lambda}{n^{\frac{1}{2}}(1-y)}\right)^{-1} \quad .$$

For $(\lambda,y) \in A$, one has:

$$\frac{\frac{-\lambda}{y(1-y)}}{\left(1 - \frac{b}{n^{\frac{1}{2}}(1-d)}\right)} \leqq h_n(\lambda,y) \leqq \frac{\frac{-\lambda}{y(1-y)}}{\left(1 + \frac{b}{n^{\frac{1}{2}}c}\right)} \quad .$$

For $n > \frac{4b^2}{(1-d)^2}$, that is, $1/2 > \frac{b}{n^{\frac{1}{2}}(1-d)}$,

$$\left(1 - \frac{b}{n^{\frac{1}{2}}(1-d)}\right)^{-1} = \sum_{r=0}^{\infty} \left(\frac{b}{n^{\frac{1}{2}}(1-d)}\right)^r \leqq 1 + \frac{2b}{n^{\frac{1}{2}}(1-d)} \qquad ..(1)$$

and for $n > \frac{b^2}{c^2}$, that is $n^{\frac{1}{2}}c > b$,

$$\left(1 + \frac{b}{n^{\frac{1}{2}}c}\right)^{-1} = 1 + \sum_{r=1}^{\infty} \left(\frac{-b}{n^{\frac{1}{2}}c}\right)^r = 1 - \frac{bn^{\frac{1}{2}}c}{n^{\frac{1}{2}}c(n^{\frac{1}{2}}c+b)}$$

$$\geqq 1 - \frac{b}{n^{\frac{1}{2}}c} \qquad ..(2)$$

Substituting (1) and (2) in the double inequality for h_n, above, one obtains for $(\lambda, y) \in A$, and

$$n > M = \max\left[\frac{4b^2}{(1-d)^2}, \frac{b^2}{c^2}\right],$$

$$\frac{-\lambda}{y(1-y)}\left(1 + \frac{2b}{n^{\frac{1}{2}}(1-d)}\right) \leqq h_n(\lambda, y) \leqq \frac{-\lambda}{y(1-y)}\left(1 - \frac{b}{n^{\frac{1}{2}}c}\right) \qquad ..(3)$$

Integrating throughout (3) from 0 through λ, noting the definition of h_n, as well as the fact that $f_n(0,y) = 1$, for all n and y, one obtains for $n > M$, and $(\lambda, y) \in A$,

$$\frac{-\lambda^2}{2y(1-y)}\left(1 + \frac{2b}{n^{\frac{1}{2}}(1-d)}\right) \leqq \mathrm{Ln}\left(f_n(\lambda, y)\right) \leqq \frac{-\lambda^2}{2y(1-y)}\left(1 - \frac{b}{n^{\frac{1}{2}}c}\right)$$

Exponentiating the above, and substituting for $f(\lambda, y)$, one has:

$$f(\lambda, y)\exp\left[-\frac{\lambda^2 b}{y(1-y)n^{\frac{1}{2}}(1-d)}\right] \leqq f_n(\lambda, y) \leqq f(\lambda, y)\exp\left(\frac{\lambda^2 b}{2y(1-y)n^{\frac{1}{2}}c}\right)$$

Let $\in > 0$ be an arbitrary constant.

(a)

$$\exp\left[-\frac{\lambda^2 b}{y(1-y)n^{1/2}(1-d)}\right] \geqq \exp\left(\frac{-b^3}{Tn^{\frac{1}{2}}(1-d)}\right) \geqq 1 - \in ,$$

where $T = \min\left[c(1-c), d(1-d)\right]$, and n is sufficiently large, say $n > N_1(\in)$.

(b)

$$\exp\left(\frac{\lambda^2 b}{2y(1-y)n^{\frac{1}{2}}c}\right) \leqq \exp\left(\frac{b^3}{2Tn^{\frac{1}{2}}c}\right) \leqq 1 + \in ,$$

for n sufficiently large, say $n > N_2(\in)$.

Finally, noting that for all $(\lambda,y) \in A$, $0 \leqq f(\lambda,y) \leqq 1$, one obtains for $n > N(\epsilon) = \text{Max}\left(M, N_1(\epsilon), N_2(\epsilon)\right)$,

$$f(\lambda,y) - \epsilon f(\lambda,y) < f_n(\lambda,y) < f(\lambda,y) + \epsilon f(\lambda,y) .$$

That is, given $\epsilon > 0$, there exists an $N > 0$, independent of which point in A is chosen, but dependent on the choice of ϵ, such that for all $n > N$, and $(\lambda,y) \in A$,

$$\left|f(\lambda,y) - f_n(\lambda,y)\right| < \epsilon f(\lambda,y) \leqq \epsilon .$$

Therefore, $f_n \longrightarrow f$ uniformly on A.

This completes the proof.

Before proving the next lemma, the following definitions are required:

Definition 4.4:

Let λ be a non-negative constant, and n be a positive integer: We define for $j = 1,2,\ldots,[n-n^{\frac{1}{2}}\lambda]$,

$$h_n(j) = n^{\frac{1}{2}}\binom{n}{j}\left(\frac{\lambda}{n^{\frac{1}{2}}} + \frac{j}{n}\right)^{j-1}\left(1 - \frac{\lambda}{n^{\frac{1}{2}}} - \frac{j}{n}\right)^{n-j} \quad \text{, and}$$

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

Definition 4.5:

Let λ be a non-negative constant. We define

$$\Phi(y) = \left(2\pi y^3(1-y)\right)^{-\frac{1}{2}} \exp\left(\frac{-\lambda^2}{2y(1-y)}\right) \qquad 0 < y < 1,$$

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

Lemma 4.3 :

Let $B = \{y : c \leq y \leq d\} \subset (0,1)$, and let $\epsilon > 0$ be an arbitrary constant. Then there exists an integer $N = N(\epsilon)$, and a positive number $\delta = \delta(\epsilon)$ such that:

$$\left| h_n(j) - \phi(y) \right| < \epsilon \text{ , whenever}$$

(i) $n > N$, (ii) $\left| \frac{j}{n} - y \right| < \delta$, and (iii) j/n and $y \in B$.

Proof: Let $j/n \in B$.

$$h_n(j) = n^{\frac{1}{2}}\binom{n}{j}\left(\frac{\lambda}{n^{\frac{1}{2}}} + \frac{j}{n}\right)^{j-1}\left(1 - \frac{\lambda}{n^{\frac{1}{2}}} - \frac{j}{n}\right)^{n-j} =$$

$$\frac{j^{j+\frac{1}{2}}(n-j)^{n-j+\frac{1}{2}}}{n^{n+\frac{1}{2}}} \binom{n}{j} \left[(j/n)^3(1 - j/n)\right]^{-\frac{1}{2}} \left(\frac{\lambda}{n^{\frac{1}{2}}(j/n)} + 1\right)^{-1} f_n(\lambda, j/n)$$

where $f_n(\lambda, y)$ is as defined in Lemma 4.2.

<u>Step 1:</u>

Consider:

$$b_n(j) = \frac{j^{j+\frac{1}{2}}(n-j)^{n-j+\frac{1}{2}}}{n^{n+\frac{1}{2}}} \binom{n}{j}$$

$$= \frac{n!}{n^{n+\frac{1}{2}}e^{-n}} \quad \frac{j^{j+\frac{1}{2}}e^{-j}}{j!} \quad \frac{(n-j)^{n-j+\frac{1}{2}}e^{-(n-j)}}{(n-j)!} .$$

We shall use the following result found in [4-1] :

$$(2\pi)^{\frac{1}{2}}\exp\left(\frac{1}{12k+1}\right) < \frac{k!}{k^{k+\frac{1}{2}}e^{-k}} < (2\pi)^{\frac{1}{2}}\exp\left(\frac{1}{12k}\right) \qquad k \in Z^+.$$

Hence: one obtains the double inequality:

$$\sqrt{(2\pi)}^{-1}\exp\left(-\frac{1}{12j}-\frac{1}{12(n-j)}+\frac{1}{12n+1}\right) \leqq b_n(j)$$

$$\leqq \sqrt{(2\pi)}^{-1}\exp\left(-\frac{1}{12j+1}-\frac{1}{12(n-j)+1}+\frac{1}{12n}\right) .$$

But since $j/n \in B$, the above may be weakened to:

$$\sqrt{(2\pi)}^{-1}\exp\left(-\frac{1}{12nc}-\frac{1}{12n(1-d)}\right) \leqq b_n(j) \leqq \sqrt{(2\pi)}^{-1}\exp\left(\frac{1}{12n}\right) \quad ..(1)$$

Since the exponential terms converge to one independently of the choice of $j/n \in B$, for given $\epsilon_1 > 0$, one can find an integer $N_1 = N_1(\epsilon_1)$, such that whenever $n > N_1$, and $j/n \in B$,

$$\left|b_n(j) - (2\pi)^{-\frac{1}{2}}\right| < \epsilon_1 .$$

Step 2:

Consider:

$$g(y) = \left(y^3(1-y)\right)^{-\frac{1}{2}} .$$

Since $g(y)$ is continuous for $y \in (0,1)$, $g(y)$ is uniformly continuous for $y \in B$, a closed subinterval of $(0,1)$. Hence, given $\epsilon_2 > 0$, there exists a positive number $\delta = \delta_2(\epsilon_2)$ such that whenever $|j/n - y| < \delta_2$,

$$\left|g(y) - g(j/n)\right| < \epsilon_2 .$$

Step 3:

Consider $f_n(\lambda,y)$.

By Lemma 4.2, for any given $\epsilon_3 > 0$, there exists an integer $N_3 = N_3(\epsilon_3)$, such that whenever $n > N_3$, and $j/n \in B$,

$$|f_n(\lambda, j/n) - f(\lambda, j/n)| < \tfrac{1}{2}\epsilon_3 ,$$ where f_n, f are as in Lemma 4.2.

But $f(\lambda,y)$ is uniformly continuous in y, for $y \in B$. Hence, there exists a positive number $\delta_3 = \delta_3(\epsilon_3)$, such that whenever $|j/n - y| < \delta_3$, and $j/n \in B$,

$$|f(\lambda, j/n) - f(\lambda,y)| < \tfrac{1}{2}\epsilon_3 .$$

Thus, by the triangle inequality, whenever

(i) $n > N_3$, (ii) $|j/n - y| < \delta_3$, and (iii) $j/n, y \in B$:

$$|f_n(\lambda, j/n) - f(\lambda,y)| < \epsilon_3 .$$

Step 4:

For $j/n \in B$, one has:

$$1 > \left(1 + \frac{\lambda}{n^{\frac{1}{2}}(j/n)}\right)^{-1} \geqq 1 - \frac{\lambda}{n^{\frac{1}{2}}c} \longrightarrow 1 \qquad ..(2)$$

Thus, given $\epsilon_4 > 0$, there exists an integer $N_4 = N_4(\epsilon_4)$, such that whenever $n > N_4$, and $j/n \in B$,

$$\left| 1 - \left(1 + \frac{\lambda}{n^{\frac{1}{2}}(j/n)}\right)^{-1} \right| < \epsilon_4 .$$

Step 5:

The uniform boundedness on B, of the functions given in the first four steps will be shown.

(i) From (1), $b_n(j) \leq (2\pi)^{-\frac{1}{2}}\exp(1/12) < 2$

(ii) Since for every $y \in B$, $g(y) < \left(c^3(1-d)\right)^{-\frac{1}{2}} = M$, then for each $j/n \in B$,

$$g(j/n) < M.$$

(iii) By Lemma 4.2, choosing ϵ_5 between 0 and 1, there exists an $N_5 = N_5(\epsilon_5)$, such that whenever $n > N_5$, and $y \in B$,

$$\left|f_n(\lambda,y) - f(\lambda,y)\right| < \epsilon_5 < 1 \quad .$$

But since $f(\lambda,y) \leq 1$, one has for $n > N_5$,

$$f_n(\lambda,y) < 2.$$

(iv) From (2), one has:

$$\left(1 + \frac{\lambda}{n^{\frac{1}{2}}(j/n)}\right)^{-1} < 1 \quad .$$

<u>Step 6:</u> (Completion of proof) Let $\epsilon > 0$ be arbitrary.

Choose:

$$\epsilon_1 = \frac{\epsilon}{8M} \qquad \epsilon_2 = \frac{\epsilon(2\pi)^{\frac{1}{2}}}{8}$$

$$\epsilon_3 = \frac{\epsilon(2\pi)^{\frac{1}{2}}}{4M} \qquad \epsilon_4 = \frac{\epsilon(2\pi)^{\frac{1}{2}}}{4M}$$

$$\epsilon_5 = 1/2$$

Let $N = N(\epsilon) = \mathrm{Max}(\ N_1(\epsilon_1),\ N_3(\epsilon_3),\ N_4(\epsilon_4),\ N_5(\epsilon_5)\ )$

and $\delta = \delta(\epsilon) = \mathrm{Min}(\ \delta_2(\epsilon_2),\ \delta_3(\epsilon_3)\ )$

We shall now show that whenever :

(i) $n > N$, (ii) $|j/n - y| < \delta$, and (iii) $j/n, y \in B$,

$$|h_n(j) - \Phi(y)| < \epsilon \ .$$

We note the following:

$$\Phi(y) = (2\pi)^{-\frac{1}{2}} g(y) f(\lambda, y) \qquad \text{and}$$

$$h_n(j) = b_n(j) g(j/n) f_n(\lambda, j/n) \left(1 + \frac{\lambda}{n^{\frac{1}{2}}(j/n)}\right)^{-1}$$

Successively applying the triangle inequality, one has:

$$|h_n(j) - \Phi(y)| \leqq \left|b_n(j) - \sqrt{(2\pi)}^{-1}\right| g(j/n) f_n(\lambda, \tfrac{j}{n}) \left(1 + \frac{\lambda}{n^{\frac{1}{2}}(j/n)}\right)^{-1}$$

$$+ \sqrt{(2\pi)}^{-1} g(y) \left|f_n(\lambda, \tfrac{j}{n}) - f(\lambda, y)\right| \left(1 + \frac{\lambda}{n^{\frac{1}{2}}(j/n)}\right)^{-1}$$

$$+ \sqrt{(2\pi)}^{-1} \left|g(j/n) - g(y)\right| f_n(\lambda, \tfrac{j}{n}) \left(1 + \frac{\lambda}{n^{\frac{1}{2}}(j/n)}\right)^{-1}$$

$$+ \sqrt{(2\pi)}^{-1} g(y) f(\lambda, y) \left| \left(1 + \frac{\lambda}{n^{\frac{1}{2}}(j/n)}\right)^{-1} - 1 \right|$$

$$< \frac{\epsilon}{8M}(2M) + \frac{\epsilon(2\pi)^{\frac{1}{2}}}{8} \cdot \frac{2}{(2\pi)^{\frac{1}{2}}} + \frac{\epsilon(2\pi)^{\frac{1}{2}}M}{4M(2\pi)^{\frac{1}{2}}} + \frac{\epsilon(2\pi)^{\frac{1}{2}}M}{4(2\pi)^{\frac{1}{2}}M} = \epsilon$$

whenever: $n > N$, $|j/n - y| < \delta$, and $j/n, y \in B$.

This completes the proof.

Lemma 4.4:

Let $0 < c < d < 1$. Then

$$\lim_{n\to\infty} \sum_{j:c<\frac{j}{n}<d} h_n(j)/n = \int_c^d \phi(y)dy .$$

Proof: Partition $[c,d]$, in the following:

$P_n = \left\{c, c + 1/n, c + 2/n, \ldots, c + r/n, d\right\}$ with r satisfying:

$\left(d - c - r/n\right) < 1/n.$

Since ϕ is continuous on $[c,d]$, it is integrable there. Hence, given $\epsilon > 0$, there exists $N_1 = N_1(\epsilon)$, such that for $n > N_1$,

$$\left|\int_c^d \phi(y)dy - \sum_{j:c<\frac{j}{n}<d} \phi(j/n)/n\right| < \epsilon/2 \qquad ..(1)$$

But by Lemma 4.3, there exists $N_2 = N_2(\epsilon)$, such that for $n > N_2$, and $c < j/n < d$,

$\left|\phi(j/n) - h_n(j)\right| < \epsilon/2$. That is:

$$\left|\sum_{j:c<\frac{j}{n}<d} \phi(j/n)/n - \sum_{j:c<\frac{j}{n}<d} \frac{h_n(j)}{n}\right| \leqq \frac{1}{n} \sum_{j:c<\frac{j}{n}<d} \left|\phi(j/n) - h_n(j)\right|$$

$$< \frac{\epsilon r}{2n} < \epsilon/2 . \qquad ..(2)$$

Taking $N(\epsilon) = N = \text{Max}(N_1, N_2)$, one obtains, from (1) and (2) together with the triangle inequality, for $n > N$:

$$\left| \int_c^d \phi(y)dy - \sum_{j:c<\frac{j}{n}<d} h_n(j)/n \right| < \epsilon .$$

This completes the proof.

Lemma 4.5:

Let $0<d<1$. Then:

$$\int_0^d \phi(y)dy = \frac{\exp(-2\lambda^2)}{\lambda} F(d/(1-d);1,\lambda^2) ,$$

with $F(x;u,\lambda)$ as in Definition 2.3.

Proof:

$$\int_0^d \phi(y)dy = \int_0^d \left(2\pi y^3(1-y)\right)^{-\frac{1}{2}} \exp\left[\frac{-\lambda^2}{2y(1-y)}\right] dy .$$

Let $X = y/(1-y)$, that is $y = X/(X+1)$.

$$\int_0^d \phi(y)dy = \int_0^{d/(1-d)} \left(2\pi X^3\right)^{-\frac{1}{2}} \exp\left[\frac{-\lambda^2(X+1)^2}{2X}\right] dX$$

$$= \frac{\exp(-2\lambda^2)}{\lambda} \int_0^{d/(1-d)} \left(2\pi X^3\right)^{-\frac{1}{2}} \lambda \exp\left[\frac{-\lambda^2(X-1)^2}{2X}\right] dX$$

$$= \frac{\exp(-2\lambda^2)}{\lambda} F(d/(1-d);1,\lambda^2) .$$

This completes the proof.

The proof of the next lemma depends on the following two well known results concerning an absolutely continuous distribution, $F(x)$, and the empirical distribution function.

Result 1: (Smirnov, see [17-1])

$$\lim_{n\to\infty} P\left[\sup_{x\in R_1} (F(x)-F_n(x)) > \lambda/n^{\frac{1}{2}}\right] = 1 - \exp(-2\lambda^2)$$

Result 2: (Birnbaum and Tingey, see [2-2])

$$P\left[\sup_{x\in R_1} (F(x)-F_n(x)) > \lambda/n^{\frac{1}{2}}\right] = 1 - \sum_{j=1}^{n} (\lambda/n)h_n(j)$$

Lemma 4.6:

$$\lim_{n\to\infty} \sum_{j=1}^{n} h_n(j)/n = \int_0^1 \phi(y)dy$$

Proof:

Taking the limit as $n\to\infty$, in Result 2, and equating this with Result 1, one has:

$$\lim_{n\to\infty} \sum_{j=1}^{n} h_n(j)/n = \frac{\exp(-2\lambda^2)}{\lambda} \qquad *$$

But by Lemma 4.5 and the continuity of the integral,

$$\int_0^1 \phi(y)dy = \frac{\exp(-2\lambda^2)}{\lambda} \lim_{t\to\infty} F(t;1,\lambda^2) = \frac{\exp(-2\lambda^2)}{\lambda}$$

This completes the proof, by virtue of * .

Lemma 4.7 :

Let $0 < d < 1$. Then:

$$\lim_{n\to\infty} \sum_{j=1}^{[nd]} h_n(j)/n = \int_0^d \phi(y)dy .$$

Proof: By Lemma 4.6, one has:

$$\lim_{n\to\infty} \sum_{j=1}^{n} h_n(j)/n = \int_0^1 \phi(y)dy .$$

Hence, Lemma 4.4 and the above give, for $c < \frac{1}{2}$:

$$\lim_{n\to\infty} \sum_{j=1}^{[nc]} h_n(j)/n < \int_0^c \phi(y)dy + \int_{1-c}^1 \phi(y)dy .$$

By the absolute continuity of the integral, for given $\epsilon > 0$, one can find a number c, $0 < c < \frac{1}{2}$, such that $c < d$, and

$$\int_0^c \phi(y)dy + \int_{1-c}^1 \phi(y)dy < \epsilon .$$

Therefore: Lemma 4.4 and the above imply

$$\lim_{n\to\infty} \sum_{j=1}^{[nd]} h_n(j)/n > \lim_{n\to\infty} \sum_{j:c<\frac{j}{n}<d} h_n(j)/n = \int_c^d \phi(y)dy$$

$$> \int_0^d \phi(y)dy - \epsilon \qquad ..(1),$$

and:

$$\lim_{n\to\infty} \sum_{j=1}^{[nd]} h_n(j)/n = \lim_{n\to\infty} \sum_{j=1}^{[nc]} h_n(j)/n + \lim_{n\to\infty} \sum_{j:c<\frac{j}{n}<d} \left(h_n(j)/n\right)$$

$$< \int_c^d \phi(y)dy + \epsilon$$

$$< \int_0^d \phi(y)dy + \epsilon \qquad ..(2).$$

(1) and (2) together imply, that for given $\epsilon > 0$,

$$\left| \lim_{n\to\infty} \sum_{j=1}^{[nd]} h_n(j)/n - \int_0^d \phi(y)dy \right| < \epsilon .$$

By virtue of the fact that ϵ is arbitrarily chosen, this completes the proof.

Lemma 4.8:

$$\lim_{n\to\infty} P\left[Z_n(\lambda/n^{\frac{1}{2}}) < nd \right] = \lambda\exp(2\lambda^2)\int_0^d \phi(y)dy ,$$

where $Z_n(d)$ is defined in Definition 4.3.

Proof:

$$P\left(Z_n(\lambda/n^{\frac{1}{2}}) < nd\right) = \frac{P\left[Y_n(\lambda/n^{\frac{1}{2}}) < nd ,\ Y_n(\lambda/n^{\frac{1}{2}}) \in \left\{0,1,\ldots,[n-n^{\frac{1}{2}}\lambda]\right\}\right]}{P\left[Y_n(\lambda/n^{\frac{1}{2}}) \in \left\{0,1,\ldots,[n-n^{\frac{1}{2}}\lambda]\right\}\right]}$$

By Lemma 4.1 and definition 4.4, the above reduces to:

$$P\left[Z_n(\lambda/n^{\frac{1}{2}}) < nd\right] = \left(\sum_{j=0}^{[nd]} \lambda h_n(j)/n\right) \left(\sum_{j=0}^{n} \lambda h_n(j)/n\right)^{-1} .$$

Letting $n\to\infty$, and noting that the summands at $j = 0$ become negligeable, one obtains from Lemma 4.7,

$$\lim_{n\to\infty} P\left[Z_n(\lambda/n^{\frac{1}{2}}) < nd\right] = \left(\int_0^d \phi(y)dy\right)\left(\int_0^1 \phi(y)dy\right)^{-1} . \quad ..(1)$$

But as was seen in Lemma 4.6,

$$\int_0^1 \phi(y)dy = \frac{\exp(-2\lambda^2)}{\lambda} .$$

Substituting this result in (1), above, the desired result is immediate.

This completes the proof.

Lemma 4.9: (Completion of THEOREM 4.1)

$$\lim_{n\to\infty} P\left[W_n(\lambda/n^{\frac{1}{2}}) < c\right] = F(c;1,\lambda^2) , \quad c > 0 ,$$

where $$W_n(\lambda/n^{\frac{1}{2}}) = \frac{Z_n(\lambda/n^{\frac{1}{2}})}{n - Z_n(\lambda/n^{\frac{1}{2}})} .$$

Proof: The following are immediate:

$$W_n(\lambda/n^{\frac{1}{2}}) = Z_n(\lambda/n^{\frac{1}{2}})/n \left(1 - Z_n(\lambda/n^{\frac{1}{2}})/n\right)^{-1} \quad ..(1)$$

and

$$W_n(\lambda/n^{\frac{1}{2}}) < c \quad \text{iff} \quad Z_n(\lambda/n^{\frac{1}{2}}) < nc/(1+c). \qquad ..(2)$$

Therefore,

$$\lim_{n\to\infty} P\left[W_n(\lambda/n^{\frac{1}{2}}) < c\right] = \lim_{n\to\infty} P\left[Z_n(\lambda/n^{\frac{1}{2}}) < nc/(1+c)\right]$$

$$= \lambda\exp(2\lambda^2)\int_0^{c/(1+c)} \phi(y)dy \qquad \text{(by Lemma 4.8)}$$

$$= F(c;1,\lambda^2) \qquad \text{(by Lemma 4.5).}$$

This completes the proof.

One has, therefore by Lemma 4.9,

$$W_n(\lambda/n^{\frac{1}{2}}) \xrightarrow[L]{} IG(1,\lambda^2).$$

This completes the proof of THEOREM 4.1.

We have thus shown that under quite general conditions, that the conditioned random variable is approximately Inverse Gaussian, for large n. Potential application of this result will be discussed at the end of this thesis.

CHAPTER 5 : EVALUATION OF THE INVERSE GAUSSIAN DISTRIBUTION FUNCTION

The author,[12-1], by means of a sequence of transformations, evaluated the distribution function of the Inverse Gaussian distribution. In this chapter, the result will be verified by a simpler, although less deductive method than appears in [12].

THEOREM 5.1:

$$F(c;\mu,\lambda) = H\left[(\lambda/c)^{\frac{1}{2}}\left(1 - \frac{c}{\mu}\right)\right] + \exp(2\lambda/\mu)H\left[(\lambda/c)^{\frac{1}{2}}\left(1 + \frac{c}{\mu}\right)\right],$$

for all $c > 0$, where

$$H(z) = \int_z^\infty (2\pi)^{-\frac{1}{2}}\exp(-t^2/2)dt \qquad -\infty < z < \infty,$$

and $F(c;\mu,\lambda)$ is as defined in Chapter 2.

Proof: Let

$$G(c;\mu,\lambda) = H\left[(\lambda/c)^{\frac{1}{2}}\left(1 - \frac{c}{\mu}\right)\right] + \exp(2\lambda/\mu)H\left[(\lambda/c)^{\frac{1}{2}}\left(1 + \frac{c}{\mu}\right)\right].$$

Then for $c > 0$,

$$\frac{\partial G(c;\mu,\lambda)}{\partial c} =$$

$$\left[\tfrac{1}{2}(\lambda/c^3)^{\frac{1}{2}}\left(1 - \frac{c}{\mu}\right) - \frac{1}{\mu}(\lambda/c)^{\frac{1}{2}} + \tfrac{1}{2}(\lambda/c^3)^{\frac{1}{2}}\left(1 + \frac{c}{\mu}\right) + \frac{1}{\mu}(\lambda/c)^{\frac{1}{2}}\right] k(c)$$

where $k(c) = (2\pi)^{-\frac{1}{2}}\exp\left(-\frac{\lambda(c-\mu)^2}{2\mu^2 c}\right)$.

Simplifying the above, one has:

$$\frac{\partial G(c;\mu,\lambda)}{\partial c} = \left(\frac{\lambda}{2\pi c^3}\right)^{\frac{1}{2}} \exp\left(-\frac{\lambda(c-\mu)^2}{2\mu^2 c}\right) = \frac{\partial F(c;\mu,\lambda)}{\partial c}.$$

Noting that as defined, $G(c;\mu,\lambda)$ is an absolutely continuous distribution function, one has by the Radon-Nikodym Theorem: $G(c;\mu,\lambda) = F(c;\mu,\lambda)$.
In view of the definition of $G(c;\mu,\lambda)$, this completes the proof.

The significance of this Theorem is that Inverse Gaussian probabilities may be easily obtained from Standard Normal and Exponential tables. ($H(z)$ is the upper tail of the Standard Normal distribution.)

Thus,with a minimum of computations, one can obtain the cumulative distribution function of random variables having the Inverse Gaussian distribution. Goodness of fit tests, for example, are quite easy using THEOREM 5.1 and an appropriate distribution free test.

At this time, we shall prove an important result, first obtained by V. Seshadri, using the characteristic function of the Inverse Gaussian distribution. The author [12-2], obtained this same result by a change of variable. Here, we shall prove it as a corollary of THEOREM 5.1.

THEOREM 5.2 :

If $X \sim IG(\mu,\lambda)$, then $Z = \frac{\lambda(X-\mu)^2}{\mu^2 X} \sim \chi^2_1$.

Proof: Let $a < \mu$, and $b = \frac{\lambda(a-\mu)^2}{\mu^2 a}$. Then clearly: if $a > 0$, $a < X < \mu^2/a$ iff $Z < b$.

Therefore, one has:

$$\begin{aligned} P(Z < b) &= P(a < X < \mu^2/a) \\ &= F(\mu^2/a;\mu,\lambda) - F(a;\mu,\lambda) \\ &= H\left[(\lambda/a)^{\frac{1}{2}}(\frac{a}{\mu} - 1)\right] - H\left[(\lambda/a)^{\frac{1}{2}}(1 - \frac{a}{\mu})\right] \\ &= H(-(b)^{\frac{1}{2}}) - H(b^{\frac{1}{2}}) \\ &= \int_{-b^{\frac{1}{2}}}^{b^{\frac{1}{2}}} (2\pi)^{-\frac{1}{2}}\exp(-t^2/2)dt \\ &= \int_{0}^{b} (2\pi y)^{-\frac{1}{2}}\exp(-y/2)dy, \end{aligned}$$

the integral of the χ^2_1 density from 0 to b. Since for suitable choice of a, b can be any positive number, one has $Z \sim \chi^2_1$. This completes the proof.

THEOREM 5.2 has great importance. The result will be used in several subsequent places in this thesis.

It is of some interest to note the following similarity between the Gaussian and Inverse Gaussian distributions:

In each case, minus twice the exponent, occurring in the density function, has the chi-square distribution with one degree of freedom.

CHAPTER 6 : CONFIDENCE SETS FOR THE PARAMETERS OF THE INVERSE GAUSSIAN DISTRIBUTION

Throughout this chapter, we shall assume that we have drawn a random sample: $X_1, X_2, \ldots, X_n$ from the Inverse Gaussian distribution $F(x;\mu,\lambda)$.

First, the required distribution theory will be developed.

THEOREM 6.1:

(a) $$\frac{n\lambda(\bar{X}-\mu)^2}{\mu^2\bar{X}} \sim \chi^2_1$$

(b) $$\frac{n(n-1)(\bar{X}-\mu)^2}{\mu^2\bar{X}\sum_{j=1}^{n}\left(\frac{1}{X_j}-\frac{1}{\bar{X}}\right)} \sim F_{1,n-1}.$$

Proof:

(a) By THEOREM 3.6, $\bar{X} \sim IG(\mu, n\lambda)$, and hence by THEOREM 5.2,

$$\frac{n\lambda(\bar{X}-\mu)^2}{\mu^2\bar{X}} \sim \chi^2_1.$$

(b) By THEOREM 3.7, $\lambda\sum_{j=1}^{n}\left(\frac{1}{X_j}-\frac{1}{\bar{X}}\right) \sim \chi^2_{n-1}$ and is independent of $\bar{X}$. Hence by part (a) of this THEOREM and elementary distribution theory, the assertion of part (b) of the THEOREM holds.

This completes the proof.

THEOREM 6.2: Confidence Interval for λ, μ known.

$$P\left(\frac{\mu^2\chi^2_{n,1-\alpha}}{\sum_{j=1}^{n}(X_j-\mu)^2/X_j} < \lambda < \frac{\mu^2\chi^2_{n,\beta}}{\sum_{j=1}^{n}(X_j-\mu)^2/X_j}\right) = 1-\alpha-\beta$$

where positive quantities α and β satisfy: $0 \leqq \alpha+\beta \leqq 1$.

Proof: By THEOREM 5.2, one has

$\frac{\lambda(X_j-\mu)^2}{\mu^2 X_j}$ for $j = 1,2,\ldots,n$ are n independent chi-square,

one degree of freedom, random variables.Therefore,

$$\frac{\lambda}{\mu^2}\sum_{j=1}^{n}(X_j-\mu)^2/X_j \sim \chi^2_n .$$

This fact in turn implies:

$$P\left(\chi^2_{n,1-\alpha} < \frac{\lambda}{\mu^2}\sum_{j=1}^{n}(X_j-\mu)^2/X_j < \chi^2_{n,\beta}\right) = 1-\alpha-\beta .$$

This completes the proof.

THEOREM 6.3: Alternate method to THEOREM 6.2.

$$P\left(\frac{\mu^2\bar{X}}{n(\bar{X}-\mu)^2}\chi^2_{1,1-\alpha} < \lambda < \frac{\mu^2\bar{X}}{n(\bar{X}-\mu)^2}\chi^2_{1,\beta}\right) = 1-\alpha-\beta ,$$

where α and β are as in THEOREM 6.2.

Proof: By THEOREM 3.6, $\bar{X} \sim IG(\mu, n\lambda)$. Hence, by Theorem 6.2 applied to $\bar{X}$, a sample of size one, the desired result is immediate.

This completes the proof.

THEOREM 6.4: Confidence interval for λ, μ unknown.

$$P\left(\frac{\chi^2_{n-1,1-\alpha}}{\sum_{j=1}^{n} (\frac{1}{X_j} - \frac{1}{\bar{X}})} < \lambda < \frac{\chi^2_{n-1,\beta}}{\sum_{j=1}^{n} (\frac{1}{X_j} - \frac{1}{\bar{X}})} \right) = 1-\alpha-\beta,$$

where α and β are as in THEOREM 6.2.

Proof: By THEOREM 3.7,

$$\lambda \sum_{j=1}^{n} (\frac{1}{X_j} - \frac{1}{\bar{X}}) \sim \chi^2_{n-1} .$$

Proceeding exactly as in THEOREM 6.2, the desired result follows.

This completes the proof.

THEOREM 6.5 : Confidence sets for μ, λ known.

(a) $$P\left(\mu^2 (\lambda \sum_{j=1}^{n} (X_j)^{-1} - C) - 2n\mu\lambda + \lambda \sum_{j=1}^{n} X_j < 0 \right) = 1-\alpha ,$$

(b) $$P\left(\mu^2 (\frac{n\lambda}{\bar{X}} - D) - 2n\mu\lambda + n\lambda\bar{X} < 0 \right) = 1-\alpha ,$$

where $0 < \alpha < 1$, $C = \chi^2_{n,\alpha}$, and $D = \chi^2_{1,\alpha}$.

Proof:

(a) As was noted in THEOREM 6.2,

$$\frac{\lambda}{\mu^2} \sum_{j=1}^{n} (X_j - \mu)^2 / X_j \sim \chi^2_n .$$

Therefore, $$P\left(\frac{\lambda}{\mu^2} \sum_{j=1}^{n} (X_j - \mu)^2 / X_j < C \right) = 1-\alpha .$$

By a simple rearrangement of terms in the above, one sees that equation (a) holds.

(b): In order to prove that equation (b) holds, one merely notes that $\bar{X} \sim IG(\mu, n\lambda)$, and applies equation (a) of this THEOREM to $\bar{X}$, a sample of size one from $F(x;\mu,n\lambda)$.

This completes the proof.

Definition 6.1:

The following notation will prove to be convenient throughout the balance of the chapter:

$$Y = \frac{\bar{X}}{n(n-1)} \sum_{j=1}^{n} \left(\frac{1}{X_j} - \frac{1}{\bar{X}} \right)$$

THEOREM 6.6 : Confidence set for μ, λ unknown.

$$P\left(\mu^2(1-EY) - 2\bar{X}\mu + \bar{X}^2 < 0\right) = 1-\alpha ,$$

where $E = F_{1,n-1,\alpha}$ and $0 < \alpha < 1$.

Proof: From the definition of Y, and THEOREM 6.1(b), one has:

$$\frac{(\bar{X}-\mu)^2}{\mu^2 Y} \sim F_{1,n-1} .$$

Proceeding exactly as in THEOREM 6.5, the desired result follows.

This completes the proof.

It is of some interest to investigate the confidence sets of THEOREMs 6.5 and 6.6 . Can these be reduced to intervals? THEOREMs 6.7 and 6.8 will answer this question.

Definition 6.2:

The following notation will prove useful in this investigation:

(a) $g(\mu) = \left(\frac{n\lambda}{\overline{X}} - D\right)\mu^2 - 2n\lambda\mu + n\lambda\overline{X}$,

(b) $h(\mu) = (1-EY)\mu^2 - 2\overline{X}\mu + \overline{X}^2$,

where D,E, and Y are defined as in THEOREM 6.5(b), THEOREM 6.6, and Definition 6.1, respectively.

Lemma 6.1:

With probability one, the roots of $g(\mu)$ and $h(\mu)$ are real.

Proof: Since $\overline{X}$ and Y are positive with probability one, and D and E are positive constants, the discriminants in each quadratic equation $g(\mu) = 0$, and $h(\mu) = 0$, are positive with probability one.

This completes the proof.

Remark: The degenerate case that g or h has a double root occurs with probability zero, since $D = 0$, and $EY = 0$ cannot occur.

Remark: The degenerate case that the coefficient of μ^2 vanishes in either quadratic forms are also events of zero probability, by virtue of the absolute continuity of random variables $\overline{X}$ and Y.

THEOREM 6.7:

Let $m < M$ be the roots of $g(\mu)$, computed after a sample has been drawn. Then, providing that $\mu > 0$,

$$\left\{\mu : g(\mu)<0\right\} = \left\{\mu : m<\mu<M\right\} \qquad \text{if} \quad \left(\frac{n\lambda}{\overline{X}} - D\right) > 0 ,$$

$$= \left\{\mu : M<\mu\right\} \qquad \text{if} \quad \left(\frac{n\lambda}{\overline{X}} - D\right) < 0.$$

Proof: If $(n\lambda - D\overline{X}) > 0$, one sees

$$g(\mu) = \left(\frac{n\lambda}{\overline{X}} - D\right)\mu^2 - 2n\lambda\mu + n\lambda\overline{X}$$

(i) has both its roots positive

(ii) has its minimum, by Rolle's Theorem, in the interval $[m,M]$.

Hence: $g(\mu) < 0$ iff $m < \mu < M$.

If $(n\lambda - D\overline{X}) < 0$, one sees that $g(\mu)$

(i) has one positive, and one negative root,

(ii) has its maximum in the interval $[m,M]$.

Hence: $g(\mu) < 0$ iff $M < \mu$.

This completes the proof.

THEOREM 6.8 :

Let $m < M$ be the roots of $h(\mu)$, after a sample has been drawn. Then, provided that $\mu > 0$,

$$\{\mu : h(\mu)<0\} = \{\mu : m<\mu<M\} \qquad \text{if} \quad (1-EY) > 0 ,$$
$$= \{\mu : M<\mu\} \qquad \text{if} \quad (1-EY) < 0.$$

Proof: The proof of this result is exactly the same as the previous one.

This completes the proof.

Example 6.1:

For a sample of size 20 from an Inverse Gaussian distribution: $Y = .189$ $\bar{X} = 1.69$, Find 95% and 99% confidence intervals for μ.

(a) 95% ; from the F-distribution table, $E = 4.38$.

$h(\mu) = .172\mu^2 - 3.38\mu + 2.86$.

$m = .89$ and $M = 18.9$.

Therefore, $.89 < \mu < 18.9$ is a 95% confidence interval for μ.

(b) 99% ; from the F-table, $E = 8.19$.

$h(\mu) = -.452\mu^2 - 3.38\mu + 2.86$.

$m < 0$ and $M = .78$.

Therefore, $.78 < \mu$ is a 99% confidence interval for μ.

Example 6.2: The analogous result does not apply to THEOREM 6.5(a). Let $\lambda=1, n=2$, $X_1 < .10$, and $X_2 > 5$.

What form has the 95% confidence set for μ, given in THEOREM 6.5(a)?

$$k(\mu) = (\lambda \sum_{j=1}^{n} (X_j)^{-1} - C)\mu^2 - 2n\lambda\mu + \lambda \sum_{j=1}^{n} X_j .$$

From the chi-square table, $C = 5.99$.

Substituting for known terms, one has

$$k(\mu) = (4+v)\mu^2 - 4\mu + (5+w) ,$$

where v and w are positive quantities depending on the values of X_1 and X_2 .

In this case, a case of positive probability, $k(\mu)$ has imaginary roots. Thus the analogous result to THEOREMs 6.7 and 6.8 do not hold.

CHAPTER 7 : BAYES ESTIMATES OF PARAMETERS

Wasan, [16-5], found a Bayes estimate of t^2, for a sample of size 1, from the Inverse Gaussian distribution $IG(t,t^2)$, for the apriori distribution:

$$g(t) = e^{-t} \qquad t > 0,$$
$$= 0 \qquad t < 0,$$ and loss function:

$$L(t^2,a) = (a-t^2)^2 .$$ (squared error loss)

In this section, two other restricted families of Inverse Gaussian distributions will be considered.

THEOREM 7.1: (Bayes estimate of λ)

Let the conditional distribution of X for each given λ, be $IG(1,\lambda)$. That is, the conditional density of X given λ is

$$f(x|\lambda) = \left(\frac{\lambda}{2\pi x^3}\right)^{\frac{1}{2}} \exp\left[-\frac{\lambda(x-1)^2}{2x}\right] \qquad x,\lambda > 0$$
$$= 0 \qquad \text{elsewhere.}$$

Let λ have apriori distribution:

$$g(\lambda) = \frac{\lambda^{r-1} e^{-\lambda}}{\Gamma(r)} \qquad \lambda > 0$$
$$= 0 \qquad \lambda < 0,$$

where r is a known positive constant.

For action a, let the loss function be

$$L(\lambda,a) = (\lambda-a)^2,$$ squared error loss.

Then $d(X) = \dfrac{2(r+\frac{1}{2})X}{X^2+1}$ is a Bayes estimate for λ, for apriori distribution $g(\lambda)$ and loss function $L(\lambda,a)$.

Proof: We shall use the well known result for squared error loss, (See Ferguson [5-1] , for example,) that $E(\lambda|X)$ is a Bayes estimate of λ. The first step, therefore, is to compute the density of λ given X, $f_1(\lambda|x)$.

The joint density of X and λ is clearly:

$$f_2(x,\lambda) = \frac{(2\pi x^3)^{-\frac{1}{2}}}{\Gamma(r)} \lambda^{r-\frac{1}{2}} \exp\left[-\tfrac{1}{2}\lambda(x + 1/x)\right] \qquad x,\lambda > 0$$

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

The unconditional density of X, for $x > 0$ is:

$$f_3(x) = \int_0^\infty f_2(x,\lambda)d\lambda$$

$$= \frac{(2\pi x^3)^{-\frac{1}{2}}}{\Gamma(r)} \int_0^\infty \lambda^{r-\frac{1}{2}} \exp\left[-\tfrac{1}{2}\lambda(x + 1/x)\right]d\lambda$$

$$= \frac{(2\pi x^3)^{-\frac{1}{2}}}{\Gamma(r)} 2^{r+\frac{1}{2}}\Gamma(r+\tfrac{1}{2}) \, (x + 1/x)^{-r-\frac{1}{2}} \quad ,$$

and the unconditional density of X, $f_3(x) = 0$, elsewhere.

For $x > 0$, we therefore obtain:

$$f_1(\lambda|x) = f_2(x,\lambda)/f_3(x) \quad , \qquad \lambda > 0,$$

$$= \frac{\lambda^{r-\frac{1}{2}}2^{-r-\frac{1}{2}}}{\Gamma(r+\frac{1}{2})} (x + 1/x)^{r+\frac{1}{2}} \exp\left(-\frac{\lambda}{2}(x + 1/x)\right) \qquad \lambda > 0,$$

and $f_1(\lambda|x) = 0 \qquad \lambda < 0.$

$$E(\lambda|x) = \int_0^\infty \lambda \, f_1(\lambda|x)d\lambda = \frac{2\,(x + 1/x)^{-1}\,\Gamma(r+3/2)}{\Gamma(r+\frac{1}{2})}$$

The above equation, by virtue of the fact that $\Gamma(\theta+1) = \theta\Gamma(\theta)$, reduces to:

$$E(\lambda|x) = \frac{2(r+\frac{1}{2})x}{x^2+1} \qquad x > 0.$$ That is

$$d(X) = \frac{2(r+\frac{1}{2})X}{X^2+1}$$ is a Bayes estimate of λ, for the apriori distribution and loss function given.

This completes the proof.

The following result is needed in the proof of THEOREM 7.2.

Lemma 7.1:

Let c,r be positive constants. Then

$$I = \int_0^\infty t^{r-1} \exp(-ct^2)dt = \tfrac{1}{2}\Gamma(r/2)\, c^{-r/2} \quad .$$

Proof: Let $y = ct^2$. Change of variable in the above gives:

$$I = \int_0^\infty \frac{y^{(r-2)/2}\, e^{-y}}{2c^{r/2}}\, dy = \tfrac{1}{2}\Gamma(r/2)\, c^{-r/2} \quad .$$

This completes the proof.

THEOREM 7.2 : (Bayes estimate of the reciprocal of the mean)

Let the conditional distribution of a random variable X, for given t be IG(1/t,1). That is, the conditional density of X given t, is:

$$f(x|t) = (2\pi x^3)^{-\frac{1}{2}} \exp\left[- \frac{t^2(x - 1/t)^2}{2x}\right] \qquad x,t > 0,$$

$$= 0 \qquad \text{elsewhere.}$$

Let t have the apriori distribution:

$$g(t) = \left(t^{r-1} e^{-t}\right)/\Gamma(r) \qquad t > 0$$

$$= 0 \qquad t < 0,$$

where r is a known positive constant.

Let the loss function be:

$L(t,a) = (t-a)^2$, for action a.

Then $d(X) = (2/X)^{\frac{1}{2}} \dfrac{\Gamma(\frac{1}{2}(r+1))}{\Gamma(\frac{1}{2}r)}$ is a Bayes estimate

of t, with respect to the apriori distribution g(t) and loss function L(t,a).

Proof: As in the case of the previous theorem, where the loss function was squared error, it suffices to show that

$E(t|X) = d(X)$.

The first step is therefore to compute $f_1(t|x)$, the conditional density of t given X.

The joint density of X and t, $f_2(x,t)$, is readily seen to be:

$$f_2(x,t) = t^{r-1} (2\pi x^3)^{-\frac{1}{2}} \exp(-1/2x)\exp(-\tfrac{1}{2}xt^2)/\Gamma(r), \quad x,t > 0$$

$$= 0 \qquad \text{elsewhere.}$$

Hence, the unconditional density of X, for $x > 0$, is

$$f_3(x) = \int_0^\infty f_2(x,t)dt$$

$$= \frac{\exp(-1/2x)\ (2\pi x^3)^{-\frac{1}{2}}}{\Gamma(r)} \int_0^\infty t^{r-1}\exp(-\tfrac{1}{2}xt^2)$$

$$= \frac{\exp(-1/2x)\ (2\pi x^3)^{-\frac{1}{2}}\ \Gamma(r/2)\ (2/x)^{r/2}}{2\ \Gamma(r)},$$

the last equality by Lemma 7.1, and the unconditional density of X, for $x < 0$, is

$$f_3(x) = 0.$$

Therefore, for $x > 0$,

$$f_1(t|x) = \frac{f_2(x,t)}{f_3(x)}$$

$$= \frac{2(x/2)^{r/2}}{\Gamma(r/2)}\, t^{r-1} \exp(-\tfrac{1}{2}xt^2) \qquad t > 0,$$

$$f_1(t|x) = 0 \qquad t < 0.$$

$$E(t|x) = \int_0^\infty \frac{2(x/2)^{r/2}}{\Gamma(r/2)}\, t^r \exp(-\tfrac{1}{2}xt^2)dt$$

$$= (2/x)^{\frac{1}{2}} \frac{\Gamma(\frac{1}{2}(r+1))}{\Gamma(\frac{1}{2}r)}.$$

That is, $E(t|X) = d(X)$, as required.

This completes the proof.

CHAPTER 8 : CHARACTERISTIC PROPERTIES OF THE INVERSE GAUSSIAN DISTRIBUTION.

THEOREM 3.8, gives a characterization of the Inverse Gaussian distribution. In this chapter, further characterizations will be proved.

The reader might pose the following question: "Is the converse of THEOREM 5.2 true?"

That is to say, if

$$\frac{\lambda(X-\mu)^2}{\mu^2 X} \sim \chi^2_1 \quad , \text{ is it true that } X \sim IG(\mu,\lambda) \; ?$$

The following theorem will help negate this conjecture.

THEOREM 8.1 : (Characterization)

Let X be a non-negative random variable with density function $f(x)$.

Then in order that $X \sim IG(\mu,\lambda)$ it is necessary and sufficient that the following hold:

$$T(X) = \frac{\lambda(X-\mu)^2}{\mu^2 X} \sim \chi^2_1 \qquad ..(1)$$

and $$f(x) = \frac{\mu^3}{x^3} f(\mu^2/x) \qquad ..(2)$$

Proof: Necessity of conditions

If $X \sim IG(\mu,\lambda)$, (1) holds by THEOREM 5.2, and

$$f(x) = \left(\frac{\lambda}{2\pi x^3}\right)^{1/2} \exp\left[-\tfrac{1}{2}T(x)\right] .$$

Since $T(x) = T(\mu^2/x)$, (2) follows by a simple algebraic manipulation.

Sufficiency of conditions

Define a random variable $Y = X$ if $X < \mu$

$= \mu^2/X$ if $X \geqq \mu$.

Since for all X, T(Y) = T(X) , one has by (1),

$T(Y) \sim \chi^2_1$.

By elementary change of variable, Y has density g(y), defined by:

$$g(y) = f(y) + \frac{\mu^2}{y^2} f(\mu^2/y) \qquad 0 < y < \mu$$

$$= 0 \qquad \text{elsewhere.}$$

But by (2), this can be simplified to:

$$g(y) = \left(1 + \frac{y}{\mu}\right) f(y) \qquad 0 < y < \mu$$

$$0 \qquad \text{elsewhere.}$$

Now by elementary change of variable to the above, one has $T(Y) = \frac{\lambda(Y-\mu)^2}{\mu^2 Y}$ has density h(T) defined by:

$$h(T) = \frac{\mu^2 y^2}{\lambda(\mu^2 - y^2)} \left(1 + \frac{y}{\mu}\right) f(y) \qquad T > 0,\ 0 < y < \mu,$$

$$= 0 \qquad \text{elsewhere.}$$

This is readily simplified to:

$$h(T) = \frac{\mu y^2}{\lambda(\mu - y)} f(y) \qquad T > 0,\ 0 < y < \mu,$$

$$= 0 \qquad \text{elsewhere.}$$

But as was shown above, $T \sim \chi^2_1$, and hence:

$$h(T) = (2\pi T)^{-\frac{1}{2}}\exp(-\tfrac{1}{2}T) \qquad T > 0$$
$$= 0 \qquad T < 0.$$

Writing T in terms of y, a bijective function of T, one obtains:

$$h(T) = (2\pi\lambda)^{-\frac{1}{2}} \frac{\mu y^{\frac{1}{2}}}{\mu-y} \exp\left[-\frac{\lambda(y-\mu)^2}{2\mu^2 y}\right] \qquad T > 0,\ 0 < y < \mu$$
$$= 0 \qquad \text{elsewhere.}$$

Equating the two expressions for h(T) above, one obtains:

$$\frac{\mu y^2}{\lambda(\mu-y)} f(y) = (2\pi\lambda)^{-\frac{1}{2}} \frac{\mu y^{\frac{1}{2}}}{\mu-y} \exp\left[-\frac{\lambda(y-\mu)^2}{2\mu^2 y}\right] \qquad 0 < y < \mu$$

Solving for f(y), one has

$$f(y) = \left(\frac{\lambda}{2\pi y^3}\right)^{1/2} \exp\left[-\frac{\lambda(y-\mu)^2}{2\mu^2 y}\right] \qquad 0 < y < \mu.$$

By applying (2), one readily sees, that except possibly at $y = \mu$,

$$f(y) = \left(\frac{\lambda}{2\pi y^3}\right)^{1/2} \exp\left(-\frac{\lambda(y-\mu)^2}{2\mu^2 y}\right) \qquad 0 < y < \infty.$$

Therefore f, the density function of the random variable X, is the Inverse Gaussian density, required. That is, $X \sim IG(\mu,\lambda)$.

This completes the proof.

Example 8.1:

The random variable Y, defined in THEOREM 8.1, is clearly not Inverse Gaussian distributed, since it has density function $g(y)$ defined by:

$$g(y) = (1 + \frac{y}{\mu}) f(y) \qquad 0 < y < \mu$$

$$= 0 \qquad \text{elsewhere.}$$

One therefore sees, that whenever $0 < y < \mu$,

$g(y) \neq \frac{\mu^3}{y^3} g(\mu^2/y)$, since exactly one side of the equation, namely the right hand side, vanishes. That is, (2) does not hold.

But in the proof of THEOREM 8.1, T(Y) was shown to be χ_1^2. Therefore, condition (1) of THEOREM 8.1, is not sufficient by itself, to yield Inverse Gaussian random variables.

The following three lemmas will provide two further characterizations, based on THEOREM 5.2.

Lemma 8.1:

Let X be a non-negative random variable, and let $Y = \frac{(X-\mu)^2}{X}$, where μ is an arbitrary real constant.

Then if for a positive integer k, $E(Y^k)$ exists,

$$E(Y^k) = \sum_{j=0}^{2k} (-\mu)^j \binom{2k}{j} E(X^{j-k})$$

Proof: $E(Y^k) = E(X^{-k}(X-\mu)^{2k}) = E\left(\sum_{j=0}^{2k} (-\mu)^j \binom{2k}{j} X^{j-k}\right)$, the above following from the Binomial Theorem. By the linearity of the expectation operator, one obtains:

$$E(Y^k) = \sum_{j=0}^{2k} (-\mu)^j \binom{2k}{j} E(X^{j-k})$$

This completes the proof.

Lemma 8.2 :

Let X be a non-negative random variable, and $Y = \frac{(X-\mu)^2}{X}$, where μ is an arbitrary positive number.

Then if $E(Y^n)$ exists for every $n \in Z^+$, $E(X^n)$ and $E(X^{-n})$ also exist for every $n \in Z^+$.

Proof: Let $K = \text{INF}\{n \in Z^+ : \text{MAX}(E(X^n), E(X^{-n})) = +\infty\}$.

It is clearly sufficient to prove $K = +\infty$. We shall therefore assume that $K < \infty$, and obtain a contradiction.

By Lemma 8.1 and the fact that Y is clearly non-negative,

$$0 \leqq E(Y^K) = \sum_{j=0}^{2K} (-\mu)^j \binom{2K}{j} E(X^{j-K}) =$$

$$E(X^{-K}) + \mu^{2K} E(X^K) + \sum_{j=1}^{2K-1} (-\mu)^j \binom{2K}{j} E(X^{j-K}) < \infty . \quad ..(*)$$

But by the property that K is the smallest integer in Z^+ such that at least one of $E(X^K)$ or $E(X^{-K})$ does not exist, one sees: $E(X^j) < \infty$, $j = \pm1, \pm2, \ldots, \pm(K-1)$, and since K is finite, and μ is positive,

$$\left|\sum_{j=1}^{2K-1} (-\mu)^j \binom{2K}{j} E(X^{j-K})\right| < \infty \qquad (**)$$ and

hence: (*) gives

$$E(X^K) + \mu^{2K}E(X^{-K}) < \infty.$$

But since X is non-negative and $\mu > 0$,

$E(X^K) \leqq E(X^K) + \mu^{2K}E(X^{-K}) < \infty$, and similarly,

$E(X^{-K}) \leqq \mu^{-2K}E(X^K) + E(X^{-K}) < \infty$. But from the definition of K, one of $E(X^K)$ or $E(X^{-K})$ must be equal to $+\infty$, and therefore, the assumption that $K < \infty$, leads to the above contradiction.

Therefore $K = +\infty$. In view of the remark at the start of this proof:

This completes the proof.

Lemma 8.3:

Let X be a non-negative random variable with density function f(x), and let $E(X) = \mu$.

Then if $E(X^n)$ and $E(X^{-n})$ exist for all $n \in Z^+$, the following are equivalent: for all $x > 0$,

$$f(x) = \frac{\mu^3}{x^3} f(\mu^2/x) \qquad ..(2)$$

and $$E((X/\mu)^{-n}) = E((X/\mu)^{n+1}) \quad \text{for all } n \in Z^+. \qquad ..(3)$$

Proof: (2) implies (3):

$$E((X/\mu)^{-n}) = \int_0^\infty \frac{\mu^{n+3}}{x^{n+3}} f(\mu^2/x)dx \qquad \text{(By (2))}$$

By substituting $y = \mu^2/x$ in the above, elementary change of variable rules give

$$E((X/\mu)^{-n}) = \int_0^\infty \frac{y^{n+3}}{\mu^{n+3}} f(y)\frac{\mu^2}{y^2}\, dy = \int_0^\infty \frac{y^{n+1}}{\mu^{n+1}} f(y)dy$$

$$= E((X/\mu)^{n+1}) .$$

Therefore (2) implies (3).

Next, we shall show that (3) implies (2).
(3) is readily seen to be equivalent to

$$E((\mu^2/X)^n) = \frac{E(X^{n+1})}{\mu} \qquad \text{for all } n \in Z^+.$$

Let $Y = \mu^2/X$. The above equation, together with the fact that $E(X) = \mu$, give

$$E(Y^n) = \frac{E(X^{n+1})}{\mu} \qquad \text{for all non-negative integers, } n.$$

Let $M_y(t) = \int_0^\infty \exp(ty)dF(y)$, with $Y \sim F(y)$.

Since all the moments of Y exist,

$$M_y(t) = \sum_{j=0}^{\infty} \frac{t^r E(Y^r)}{r!} = \sum_{j=0}^{\infty} \frac{t^r E(X^{r+1})}{r!\mu}$$

Next, let $g(y) = \frac{y}{\mu}f(y)$, where f is the density of X.

Since for every $n \in Z^+ \cup \{0\}$,

$$\int_0^\infty y^n g(y)dy = \int_0^\infty \frac{y^{n+1}}{\mu} f(y)dy = \frac{E(X^{n+1})}{u} ,$$

$$\int_0^\infty \exp(ty)g(y)dy = \sum_{r=0}^{\infty} \frac{t^r E(X^{r+1})}{r!\mu} = M_y(t).$$

By uniqueness of Laplace Transform , we obtain:

Y has density $\frac{y}{\mu} f(y)$. (*)

But $X = \frac{\mu^2}{Y}$. Therefore, change of variable yields:

$$\frac{y}{\mu} f(y) = g(y) = \frac{\mu^2}{y^2} f(\mu^2/y) , \quad \text{for all } y > 0.$$

That is, for all $y > 0$, condition (2) holds.

This completes the proof.

The previous Lemma, gives rise to the following version of THEOREM 8.1:

THEOREM 8.2 : (Characterization)

Let X be a non-negative random variable, with density function f(x). Then in order that $X \sim IG(\mu, \lambda)$, it is necessary and sufficient that the following hold:

(1) $T(X) = \frac{\lambda(X-\mu)^2}{\mu^2 X} \sim \chi_1^2$, and

(3) $E((X/\mu)^{-n}) = E((X/\mu)^{n+1})$ for all non-negative integers n.

Proof: If $T(X) \sim \chi^2_1$, $E(T^n)$ exists for all $n \in Z^+ \cup \{0\}$.

That is, $E((X-\mu)^{2n} X^{-n})$ exists for every $n \in Z^+$. By virtue of the fact that X is a non-negative random variable, Lemma 8.2 gives: For every $n \in Z^+$, $E(X^n)$ and $E(X^{-n})$ exist.

Also, Inverse Gaussian random variables have moments of all positive and negative orders. Therefore, in both the Necessity and Sufficiency parts of the proof, moments of all positive and negative orders exist. Hence:

Conditions (1) and (3) are equivalent to (1) and (2), by Lemma 8.3, and conditions (1) and (2) are equivalent to $X \sim IG(\mu,\lambda)$. ((2) is defined in THEOREM 8.1).

This completes the proof.

The following theorem characterizes the Standard Inverse Gaussian distribution, IG(1,1).

THEOREM 8.3:

In order that $X \sim IG(1,1)$, it is necessary and sufficient that

(1)' $$\frac{(X-1)^2}{X} \sim \chi^2_1 \quad \text{and}$$

(4) $$X + t \approx 1/X, \quad \text{with } X, t \text{ independent, } t \sim \chi^2_1.$$

Proof: <u>Necessity of Conditions</u>

If $X \sim IG(1,1)$, THEOREMs 5.2 and 3.4 respectively, give (1)' and (4).

Sufficiency of Conditions

Conditions (1)' and (4) will give all the moments of X.

Let $Y = \frac{(X-1)^2}{X}$. By Lemma 8.1, (X is clearly non-negative)

$$E(Y^k) = \sum_{j=0}^{2k} (-1)^j \binom{2k}{j} E(X^{j-k}) \qquad \text{for all } k \in Z^+. \qquad ..(A)$$

Also, by (4), for all $k \in Z^+$,

$E((X+t)^k) = E(X^{-k})$. Applying the binomial expansion, one obtains:

$$\sum_{j=0}^{k} \binom{k}{j} E(X^j) E(t^{k-j}) = E(X^{-k}) \qquad \text{for all } k \quad Z^+ \qquad ..(B)$$

Since $t \approx Y \sim \chi_1^2$, a distribution with moments of all positive orders, the expectations given on (A) and (B), all exist. (Those of the form $E(X^j)$ exist by Lemma 8.2)

We shall evaluate $E(X^k)$ and $E(X^{-k})$ by induction. For k=1, equations (A) and (B) yield:

$$E(Y) = E(X^{-1}) - 2 + E(X) \quad ; \quad E(t) + E(X) = E(X^{-1}).$$

Since $Y \sim t \sim \chi_1^2$, $E(Y) = E(t) = 1$. Therefore, the solution to the above system is: $E(X) = 1$, $E(X^{-1}) = 2$.

Assume for all integers $k < n$, that $E(X^k)$ and $E(X^{-k})$ are uniquely determined quantities, from (1)' and (4), by equations (A) and (B). We shall then show that $E(X^n)$ and $E(X^{-n})$ are also uniquely determined from (1)' and (4) via (A) and (B).

Equations (A) and (B) yield:

$$E(X^{-n}) + E(X^{n}) = E(Y^{n}) - \sum_{j=1}^{2n-1} (-1)^{j}\binom{2n}{j}E(X^{j-n}) \qquad ..(A_n)$$

$$E(X^{-n}) - E(X^{n}) = \sum_{j=0}^{n-1} \binom{n}{j}E(X^{j})E(t^{n-j}) \qquad ..(B_n)$$

By the induction hypothesis, and the fact that the moments of Y and t, χ^2_1 random variables, are known, the right hand sides of equations (A_n) and (B_n) are known.

Since the equations (A_n) and (B_n) are non-singular in unknowns $E(X^n)$ and $E(X^{-n})$, the equations have a unique solution. Therefore, $E(X^n)$ and $E(X^{-n})$ are indeed uniquely determined from (1)' and (4). This completes the induction.

Therefore, for every $n \in Z^+$, conditions (1)' and (4) determine $E(X^n)$, uniquely.

Since as was remarked earlier, $E(X^n)$ exists for all $n \in Z^+$, conditions (1)' and (4) uniquely determine

$$E(\exp(\theta X)) = \sum_{j=0}^{\infty} \frac{\theta^j}{j!} E(X^j) \quad .$$

By uniqueness of Laplace Transform, there exists at most one probability distribution, satisfying (1)' and (4). But in the Necessity part of the proof, we showed: If $X \sim IG(1,1)$, then (1)' and (4) hold. Hence the only distribution such that (1)' and (4) hold is IG(1,1).

This completes the proof.

A particular case of a theorem of Patil and Seshadri [10-1] , can be stated as follows:

Let X_1, X_2 be non-negative independent random variables with common density function, $f(x)$; and Y_1, Y_2 be non-negative independent random variables with common density, $g(x)$.

Assume further that:

(1) $f(x),\ g(x) > 0$, for $x \geqq 0$.

(2) $d(X_1 | X_1+X_2) = d(Y_1 | Y_1+Y_2) =$ the conditional density of X_1 given (X_1+X_2) = the conditional density of Y_1 given (Y_1+Y_2), a continuous function of X_1, say, for all (X_1+X_2).

Then there exist constants a and b, such that

$$f(x) = ag(x)\exp(bx).$$

While this result will not be of direct use to us, to characterize the Inverse Gaussian distribution, as will be discussed below, the above theorem can be generalized to be of use.

We note, that assumption (1), demands that $f(0) = \lim_{h \searrow 0} f(h)$ must be positive. The quantity $f(0)$ appears in the denominator, in the proof of the above theorem, in one of the steps.

The Inverse Gaussian density vanishes at $x = 0^+$. The above theorem, must be modified, to deal with the Inverse Gaussian distribution.

The following lemmas will clear up this difficulty.

Lemma 8.4:

Let independent, non-negative random variables X_1, X_2, Y_1, Y_2 have respective density functions $f(x), f(x), g(x), g(x)$. Assume further that:

(1) $f(x), g(x) > 0$, whenever $X > 0$. ($f(0), g(0) = 0$, permitted.)

(2) For all values of X_1+X_2, and Y_1+Y_2, the conditional densities of X_1 given X_1+X_2 and Y_1 given Y_1+Y_2 are continuous and identical, $d(x|x+y)$.

Then $0 < \lim_{x \downarrow 0} \frac{f(x)}{g(x)} < \infty$, exists.

Proof: Let $k(x+y)$ be the density of X_1+X_2, and $h(x+y)$ be the density of Y_1+Y_2.

By (2), we obtain: for $x,y>0$,

$$d(x|x+y) = \frac{f(x)f(y)}{k(x+y)} = \frac{g(x)g(y)}{h(x+y)} \qquad (*) .$$

We immediately note the following:

(a) Since $d(x|x+y)$ is continuous, so must be $f(x)$ and $g(x)$ for all $x > 0$. ($k(x+y)$ and $h(x+y)$ cannot vanish for positive $x+y$, by (1), above.)

(b) $k(x+y)$ and $h(x+y)$ are also continuous for $x+y$ positive, by virtue of (a) and (*).

Rearranging in (*), we obtain:

$$\frac{f(x)}{g(x)} = \frac{k(x+y)g(y)}{h(x+y)f(y)}$$

Taking limits on both sides of the above, as $x \searrow 0$, one obtains by virtue of (a) and (b):

$$\lim_{x \searrow 0} \frac{f(x)}{g(x)} = \frac{k(y)g(y)}{h(y)f(y)}$$, a finite, positive quantity.

This completes the proof.

Lemma 8.5:

Under the conditions of Lemma 8.4, there exist constants a and b, such that

$f(x) = ag(x)\exp(bx)$.

Proof: Let the notation be as in the previous lemma. (*) of the previous lemma gives: for $x,y > 0$, and $0 < h < x+y$,

$$\frac{f(x)f(y)}{k(x+y)} = \frac{g(x)g(y)}{h(x+y)} \quad ; \quad \frac{f(h)f(x+y-h)}{k(x+y)} = \frac{g(h)g(x+y-h)}{h(x+y)} .$$

Therefore, for $x,y > 0$ and $0 < h < x+y$,

$$\frac{f(x)f(y)}{g(x)g(y)} = \frac{f(h)f(x+y-h)}{g(h)g(x+y-h)} .$$

Taking limits on both sides, as $h \searrow 0$, and setting $\lim_{h \searrow 0} \frac{f(h)}{g(h)} = a$, a being well defined, positive and finite, by Lemma 8.4; and using (a) of the previous lemma's proof, concerning the continuity of f and g, we obtain

$$\frac{f(x)f(y)}{g(x)g(y)} = \frac{a\, f(x+y)}{g(x+y)} \qquad \text{for } x,y > 0.$$

Let $\phi(x) = \frac{f(x)}{a\, g(x)}$ for all $x > 0$. The above yields: $\phi(x)\phi(y) = \phi(x+y)$.

This is the well known Cauchy Functional Equation, whose unique solution is:

$$\phi(x) = \exp(bx) \quad \text{for some constant } b, \text{ and } x > 0.$$

Therefore, for $x > 0$,

$$\phi(x) = \exp(bx) = \frac{f(x)}{a\, g(x)} \qquad x > 0.$$

That is, $f(x) = a\, g(x)\exp(bx) \quad x > 0$, for some constants a and b.

This completes the proof.

The machinery has now been established to prove a Patil-Seshadri type characterization of the Inverse Gaussian distribution.

THEOREM 8.4 : (Characterization)

Let X,Y be independent, identically distributed random variables, such that $E(X) < \infty$.

Then a necessary and sufficient condition that for some positive μ, that $X,Y \sim IG(\mu,\lambda)$, is that the conditional density of X given X+Y, is

$$d(x|x+y) = \frac{1}{2}\left(\frac{\lambda(x+y)^3}{2\pi x^3 y^3}\right)^{\frac{1}{2}} \exp\left[-\frac{\lambda}{2}\left(x^{-1} + y^{-1} - 4(x+y)^{-1}\right)\right],$$

$0 < x < x+y$.

Proof: Necessity of Condition

Let $X, Y \sim IG(\mu, \lambda)$ be independent random variables. By THEOREM 3.6(a), $X+Y \sim IG(2\mu, 4\lambda)$. Hence

$$d(x|x+y) = \frac{f(x)f(y)}{h(x+y)} , \qquad (*)$$

where f is the Inverse Gaussian density $\frac{d}{dz} F(z;\mu,\lambda)$ and

h is the Inverse Gaussian density $\frac{d}{dz} F(z;2\mu,4\lambda)$.

The required value of $d(x|x+y)$ is immediate from (*), upon substituting the functional forms of the Inverse Gaussian densities f and h.

Sufficiency of Condition

Let $g(x) = \frac{d}{dx} F(x;\mu,\lambda)$. Suppose that densities $f(x)$ and $g(x)$ yield:

$$d(x|x+y) = \frac{1}{2}\left(\frac{\lambda(x+y)^3}{2\pi x^3 y^3}\right)^{\frac{1}{2}} \exp\left[-\frac{\lambda}{2}\left(x^{-1} + y^{-1} - 4(x+y)^{-1}\right)\right]$$

whenever $0 < x < x+y$. By Lemma 8.5, there exist constants a and b, such that whenever $0 < x < \infty$,

$$\begin{aligned} f(x) &= ag(x)\exp(bx) \\ &= a\left(\frac{\lambda}{2\pi x^3}\right)^{\frac{1}{2}} \exp\left[-\frac{\lambda}{2\mu^2}\left[\left(1 - \frac{2b\mu^2}{\lambda}\right)x - 2\mu + \frac{\mu^2}{x}\right]\right] , \end{aligned}$$

and $f(x) = 0$, $x < 0$.

Let $K = a\left(\frac{\lambda}{2\pi}\right)^{\frac{1}{2}} \exp(\lambda/\mu)$ and $c = 1 - \frac{2b\mu^2}{\lambda}$

Therefore,

$$f(x) = K x^{-3/2} \exp\left[-\frac{\lambda}{2\mu^2}\left(cx + \frac{\mu^2}{x}\right)\right] .$$

(i) If $c > 0$, by completing the square of the exponent, one obtains the fact that $f(x)$ is an Inverse Gaussian density. The fact that f has the correct value of λ, follows from the necessity part of this proof, for if f had some other value of λ, say λ', one would have

$$\frac{d(x|x+y)}{d(x|x+y)} = 1$$

$$= (\lambda/\lambda')^{\frac{1}{2}} \exp\left[-\tfrac{1}{2}(\lambda-\lambda')(x^{-1} + y^{-1} - 4(x+y)^{-1})\right],$$

whenever $0 < x < x+y$. This is an obvious contradiction.

(ii) If $c = 0$, one has

$$f(x) = Kx^{-3/2}\exp(-\lambda/2x) \quad , \quad x > 0$$

$$= 0 \quad , \quad x < 0.$$

$$E(X) = \int_0^\infty K x^{-\frac{1}{2}} \exp(-\lambda/2x)dx = +\infty.$$

This contradicts the fact that $E(X)$ exists. Therefore, $c = 0$ cannot occur under the assumptions of the theorem.

(iii) If $c < 0$, we obtain:

$$\lim_{x\to\infty} f(x) = \lim_{x\to\infty} K x^{-3/2} \exp\left(-\frac{\lambda}{2u^2}\left(cx + \frac{u^2}{x}\right)\right) = +\infty .$$

That is, $f(x)$ cannot be a probability density.

In view of the fact that (i) is the only possible case that can occur,

This completes the proof.

Example 8.2:

We shall show, by example, that in Khatri's characterization of the Inverse Gaussian distribution, THEOREM 3.8, assumptions on the moments cannot be dropped.

Let X_1, X_2 be independent random variables, with common density function

$$f(x) = (2\pi x^3)^{-\frac{1}{2}} \exp(-1/2x) \qquad x > 0$$
$$= 0 \qquad x < 0.$$

Result (i): f is a density function.

This is easily seen, from the fact that $Y \sim \chi_1^2$ iff $1/Y$ has density $f(x)$, above.

Result (ii): $U = X_1 + X_2$ and $V = 1/X_1 + 1/X_2 - 4/(X_1+X_2)$ are independent.

Argument : Let $Y_1 = \min(X_1, X_2)$; $Y_2 = \max(X_1, X_2)$. Then $U = Y_1 + Y_2$; $V = 1/Y_1 + 1/Y_2 - 4/(Y_1+Y_2)$.

Also, the joint density of Y_1, Y_2 wherever positive, is

$$f_1(Y_1, Y_2) = \frac{1}{\pi} (Y_1 Y_2)^{-3/2} \exp\left[-1/2Y_1 - 1/2Y_2\right] \qquad 0 < Y_1 < Y_2.$$

and

$$\frac{\partial(U,V)}{\partial(Y_1,Y_2)} = (Y_2)^{-2} - (Y_1)^{-2} < 0 .$$

The joint density of U and V is therefore,

$$f_2(U,V) = \frac{1}{\pi} (Y_1 Y_2)^{-3/2} \left(\frac{1}{Y_1^2} - \frac{1}{Y_2^2} \right) \exp\left[-\tfrac{1}{2}(Y_1^{-1} + Y_2^{-1})\right]$$

$$= \frac{1}{\pi} \frac{(Y_1 Y_2)^{\frac{1}{2}}}{Y_2^2 - Y_1^2} \exp(-2/U)\exp(-\tfrac{1}{2}V).$$

Next, we note that:

$$V = Y_1^{-1} + Y_2^{-1} - 4(Y_1+Y_2)^{-1} = \frac{(Y_2-Y_1)^2}{Y_1Y_2(Y_1+Y_2)} .$$ Hence,

$$U^{3/2}V^{1/2} = \frac{Y_2^2 - Y_1^2}{(Y_1Y_2)^{\frac{1}{2}}} .$$

Therefore, by substituting the above in the expression for f_2, the joint density of U and V is seen to be

$$f_2(U,V) = \frac{1}{\pi} U^{-3/2}\exp(-2/U)V^{-1/2}\exp(-V/2) , \quad U,V > 0 ,$$

$$= 0 \qquad \text{elsewhere.}$$

Since f_2 can be factored into a function of U times a function of V, U and V are indeed independent random variables.

We have shown, that f(x), a non Inverse Gaussian density, enjoys the property of Khatri's characterization. Assumptions on the moments cannot be dropped, therefore.

CHAPTER 9 : THE MULTIVARIATE INVERSE GAUSSIAN DISTRIBUTION

Definition 9.1:

A random vector $X = (x_1,\ldots,x_n)^T$ is "Multivariate Inverse Gaussian distributed," if there exists a non-singular matrix P, of n rows and n columns, such that

$Z = PX$ is a vector of independently distributed Inverse Gaussian random variables.

In symbols, one writes: $X \sim$ MVIG.

The reader may ask: "Under what conditions are all the univariate marginal distributions of an MVIG random vector, Inverse Gaussian distributions?"

The next important lemma will provide the key to answering this question.

Lemma 9.1 :

Let $X_1,\ldots,X_n$ be independent random variables with $X_j \sim IG(\mu_j,\lambda_j)$, $j = 1,\ldots,n$.

Then a necessary and sufficient condition that

$X_0 = \sum_{j=1}^{n} X_j$ be an Inverse Gaussian random variable

is that: $(\mu_j)^2/\lambda_j$ does not depend on j, $j = 1,\ldots,n$.

Proof: <u>Necessity of condition</u>

Let $X_0 \sim IG(\mu_0, \lambda_0)$, and let $f_j(\theta)$ be the characteristic function of X_j, $j = 0,1,\ldots,n$.

By THEOREM 3.11, for $j = 0,1,\ldots,n$,

$$\log f_j(\theta) = i\mu_j\theta + \int_0^\infty \left(\exp(iu\theta)-1-iu\theta\right)u^{-3/2}a_j\exp(-b_ju)du ,$$

where $a_j = (\lambda_j/2\pi)^{\frac{1}{2}}$ and $b_j = (\lambda_j/\mu_j^2)$.

But since $X_1\ldots X_n$ are independent random variables, one obtains from the definition of X_0 ,

$$\log f_0(\theta) = \sum_{j=1}^{n} \log f_j(\theta) =$$

$$i\theta\sum_{j=1}^{n}\mu_j + \int_0^\infty \left(\exp(iu\theta)-1-iu\theta\right)u^{-3/2}\left(\sum_{j=1}^{n} a_j\exp(-b_ju)\right) du .$$

By the uniqueness of the Kolmogorov representation, the following is clear:

$$a_0\exp(-b_0u) = \sum_{j=1}^{n} a_j\exp(-b_ju) \qquad \text{for positive } u. \qquad ..(1)$$

By taking limits on both sides of (1), as $u \searrow 0$, one obtains:

$$a_0 = \sum_{j=1}^{n} a_j .$$

By differentiating both sides of (1), and letting u tend to 0, one also obtains:

$$a_0 b_0^m = \sum_{j=1}^{n} a_j b_j^m \qquad m = 1,2,\ldots$$

That is, since $a_0 > 0$,

$$(b_0)^m = \sum_{j=1}^{n} (a_j/a_0)(b_j)^m \qquad m = 0,1,\ldots \qquad ..(2).$$

Next, consider the following discrete distribution, for a random variable Y:

$P(Y=b_j) = a_j/a_0 \qquad j = 1,\ldots,n$ *

$P(Y=t) = 0 \qquad t \neq a_j$ for any j.

The above is clearly a probability function, by (2) above, with m = 0. Also by (2), for $m = 1,2,\ldots$

$$E(Y^m) = \sum_{j=1}^{n}(a_j/a_0)(b_j)^m = (b_0)^m \ .$$

Hence, the moment generating function of Y is:

$E(\exp(ty)) = \exp(b_0 t)$.

By uniqueness of Laplace Transform, one obtains:

$P(Y=b_0) = 1.$ **

By the definition of a_j and the fact that Inverse Gaussian random variables have $\mu > 0$, $a_j > 0$, $j = 1,\ldots,n$. * and ** therefore imply, for $j = 1,\ldots,n$, that

$b_j = b_0$. That is, by definition of b_j,

$(\mu_j^2/\lambda_j) = (b_0)^{-1} \qquad j = 1,\ldots,n$, and is therefore independent of j.

Sufficiency of condition

Setting $\mu_j^2/\lambda_j = C$, $j = 1,\ldots,n$, THEOREM 3.1 gives:

$$f_j(\theta) = \exp\left(\frac{\lambda_j}{\mu_j}\left[1 - (1 - 2i\theta C)^{\frac{1}{2}}\right]\right) \qquad j = 1,\ldots,n.$$

By the independence of $X_1,\ldots,X_n$, one obtains

$$f_0(\theta) = \prod_{j=1}^{n} f_j(\theta) = \exp\left(\left(\sum_{j=1}^{n}\frac{\lambda_j}{\mu_j}\right)\left[1 - (1 - 2i\theta C)^{\frac{1}{2}}\right]\right).$$

Hence, by THEOREM 3.1, and uniqueness of Fourier Transform, one obtains

$$X_0 \sim IG\left(\sum_{j=1}^{n}\mu_j , \frac{1}{C}\left(\sum_{j=1}^{n}\mu_j\right)^2\right)$$

This completes the proof.

The following theorem gives necessary and sufficient conditions that a Multivariate Inverse Gaussian random vector have all its univariate components Inverse Gaussian distributed.

THEOREM 9.1:

Let $Z = (z_1,\ldots,z_n)^T$ be a vector of independent random variables with $z_j \sim IG(\mu_j,\lambda_j)$, $j = 1,\ldots,n$.

Let P be an nxn non-singular matrix, and let $X = (x_1,\ldots,x_n)^T$ be the random vector satisfying:

$Z = PX$.

Denote P^{-1} by $Q = \begin{pmatrix} q_{11} & q_{12} & \cdots & q_{1n} \\ \vdots & & & \vdots \\ q_{n1} & \cdot & \cdot & q_{nn} \end{pmatrix}$

Then necessary and sufficient conditions that for a fixed $i, i = 1, \ldots, n$, that x_i be Inverse Gaussian distributed, is that for each j such that $q_{ij} \neq 0$, the following hold:

(1) $q_{ij} > 0$

(2) $\dfrac{\mu_j^2 q_{ij}}{\lambda_j}$ does not depend on j.

Proof: Necessity of (1):

We assume $x_i \sim IG(\cdot, \cdot)$ and shall conclude $q_{ij} \geqq 0$. Suppose it were possible that $x_i \sim IG(\cdot, \cdot)$, and $q_{ij} < 0$. But $x_i = \sum_{j=1}^{n} q_{ij} z_j$. Were some $q_{ij} < 0$, the independence and non-negativity of the z_j give:

$P(x_i < 0) > 0$. This cannot happen for Inverse Gaussian random variables, which are positive with probability 1. The assumption that some q_{ij} could be negative, with x_i Inverse Gaussian distributed cannot be compatible. That is, (1) is indeed necessary.

Necessity and sufficiency of (2):

Let $\{q_{ij_1}, \ldots, q_{ij_r}\}$ be the set of positive q_{ij} for a fixed i. This set is non-empty, since P is non-singular.

$$x_i = \sum_{k=1}^{r} q_{ij_k} z_{j_k} = \sum_{k=1}^{r} w_{j_k} \quad , \text{ where}$$

$w_{j_k} = q_{ij_k} z_{j_k}$, $k = 1,\ldots,r$, are independent random variables, such that; by THEOREM 3.5,

$$w_{j_k} \sim IG(q_{ij_k}\mu_{j_k},\ q_{ij_k}\lambda_{j_k}) \qquad k = 1,\ldots,r$$

Hence, by Lemma 9.1,

$x_i \sim IG(\cdot,\cdot)$ iff $\dfrac{q_{ij_k}\mu_{j_k}^2}{\lambda_{j_k}}$ $k = 1,\ldots,r$, does not depend on k. In view of the definition of q_{ij_k} , $k = 1,\ldots,r$,

This completes the proof.

<u>Remark:</u>

If the properties enjoyed by the fixed i, is enjoyed by all the i, $i=1,\ldots,n$, all univariate marginals are Inverse Gaussian distributions.

THEOREM 9.2 :

In the statement of THEOREM 9.1, if $z_1,\ldots,z_n$, is a random sample with $z_j \sim IG(\mu,\lambda)$, then a necessary and sufficient condition that all the marginals of X, be Inverse Gaussian distributions, is that all the row vectors of Q, $(q_{i1},\ldots q_{in})$ consist only of zeros and a positive quantity c_i, depending only on i.

Proof: (Special case of THEOREM 9.1)

Example 9.1: (Sequential sampling)

Let $z_1,\dots,z_n$ be a random sample, with $z_j \sim IG(\mu,\lambda)$.

Let:

$$X = \begin{pmatrix} 1 & 0 & 0 & 0 & \cdots & 0 \\ \frac{1}{2} & \frac{1}{2} & 0 & 0 & \cdots & 0 \\ 1/3 & 1/3 & 1/3 & 0 & \cdots & 0 \\ \cdot & \cdot & \cdot & \cdot & \cdots & \cdot \\ \cdot & \cdot & \cdot & \cdot & \cdots & \cdot \\ 1/n & 1/n & 1/n & \cdot & \cdots & 1/n \end{pmatrix} \begin{pmatrix} z_1 \\ z_2 \\ \cdot \\ \cdot \\ \cdot \\ z_n \end{pmatrix}$$

$X \sim$ MVIG since the matrix above is non-singular. (The determinant is the product of the terms of the main diagonal).

X has all its marginals Inverse Gaussian, by THEOREM 9.2, with $c_i = 1/i \quad i = 1,\dots,n$.

CHAPTER 10: MEASURE THEORETIC INTRODUCTION TO THE INVERSE GAUSSIAN PROCESS

In this chapter, the author will introduce the notation required in Chapters 11 and 12. Illustrative examples will be given to show how ones intuition can lead him to false conclusions, with regard to properties of the sample functions of a stochastic process.

IMPORTANT DEFINITIONS

Definition 10.1: (Stochastic Process)

A Stochastic Process is a 4-tuple $(\Omega, \mathcal{F}, P, X_t : t \in T)$ such that:

(1) T is a subset of the non-negative real numbers,

(2) $\Omega = \prod_{t \in T} \Omega_t$, where each Ω_t is the sample space of the random variable X_t,

(3) $\mathcal{F}$ is the sigma-field generated by $\prod_{t \in T} \mathcal{F}_t$, where each $\mathcal{F}_t$ is the Borel field generated by sets of the form: $\{w \in \Omega_t : X_t(w) < a\}$, and

(4) P is the probability measure induced on $\mathcal{F}$, by the finite dimensional distributions of $(X_{t_1}, \ldots, X_{t_n})$, for every finite subset $\{t_1, \ldots, t_n\} \subset T$.

Definition 10.2: (Notation)

Let S be a finite subset of T. We define $\mathcal{F}^*_S$ by:

$$\mathcal{F}^*_S = \prod_{t \in T} \mathcal{G}_t \quad , \text{ where } \quad \mathcal{G}_t = \mathcal{F}_t \ , \ t \in S$$
$$\qquad = \{\Omega_t\} \ , \ t \notin S.$$

Definition 10.3: (Notation)

Let $S = \{t_1, \ldots, t_n\}$ be a finite subset of T. We shall define P_S as the measure induced on $(\Omega, \mathcal{F}^*_S)$ by the finite dimensional distribution of $(X_{t_1}, \ldots, X_{t_n})$.

Definition 10.4: (Inverse Gaussian Process)

An Inverse Gaussian Stochastic Process, $X_t : t \geqq 0$, is one with the following properties:

(1) $(X_t - X_s) \sim IG(\alpha(t-s), \beta(t-s)^2)$ for $t > s$, with α and β positive constants.

(2) Process has independent increments. That is, the set of random variables $\{(X_{t_{j+1}} - X_{t_j}) : j = 1, \ldots, n\}$ are mutually independent whenever $0 \leqq t_1 < t_2 < \ldots < t_{n+1} < \infty$.

(3) $X_0 = 0$ a.e.

Definition 10.5:

A Stochastic Process is "Monotonic increasing in probability" if whenever $t > s$,

$P(X_t \geqq X_s) = 1.$

Definition 10.6:

A Stochastic Process $(\Omega, \mathcal{F}, P, X_t : t \geq 0)$ is " a.e. monotonic increasing" if there exists a set $N \in \mathcal{F}$, such that:

(1) $P(N) = 0$ and

(2) For every fixed $w \in (\Omega - N)$, $X_t(w)$ is a monotonic non-decreasing function of t.

Definition 10.7:

A Stochastic Process is "continuous in probability" if for every $\epsilon > 0$, and $t \geq 0$:

$$\lim_{h \to 0} P(|X_{t+h} - X_t| > \epsilon) = 0 \qquad (*)$$

If (*) holds subject to the condition that $h \to 0$ from above, the process is "right-continuous in probability".

Definition 10.8:

A Stochastic Process $(\Omega, \mathcal{F}, P, X_t : t \geq 0)$ is " a.e. continuous (right-continuous), if there exists a set $N \in \mathcal{F}$, such that:

(1) $P(N) = 0$ and

(2) For every fixed $w \in (\Omega - N)$, $X_t(w)$ is a continuous (right-continuous) function of t.

Definition 10.9: (Separable process)

A Stochastic Process $(\Omega, \mathcal{F}, P, X_t : t \geqq 0)$ is called "Separable", if there exists a countable set $S \subset [0, \infty)$, with the following property:

For every open interval I, of finite length,

$$\sup_{t \in I} X_t(w) = \sup_{t \in I \cap S} X_t(w) \quad \text{a.e.} \quad \text{and}$$

$$\inf_{t \in I} X_t(w) = \inf_{t \in I \cap S} X_t(w) \quad \text{a.e.}$$

Definition 10.10: (Universal Separating Set, USS)

A set S with the property required in Definition 10.9 is called a universal separating set.

Definition 10.11: (Modification of a process)

A Stochastic Process, $\{Z_t : t \in T\}$, is a modification of a process $\{X_t : t \in T\}$, if for every $t \in T$,

$P(Z_t = X_t) = 1.$

Definition 10.12:

A Stochastic Process is finite a.e., if for all finite T, $P(\sup_{t<T} X_t(w) = \infty) = 0$.

Definition 10.13:

A Stochastic Process is unbounded in probability if for every $K > 0$, $t > 0$:

$P(X_t > K) > 0.$

Before one can continue with the introduction to the measure theoretic properties of the Inverse Gaussian Process, one logical question must be answered: "How does one know that there exists a measure P, such that P extends the measures induced by each finite dimensional distribution?"

The answer lies in the famous Kolmogorov Theorem, which we state without proof.

THEOREM 10.1: (Kolmogorov)

Let Ω be defined by (2) of definition 10.1, with each Ω_t a copy of the real line. Then if there exists a probability measure P_S on $(\Omega, \mathcal{F}^*_S)$ for every finite $S \subset T$, such that whenever $S_1 \subset S_2$, P_{S_2} is an extension of P_{S_1}, it follows that there exists a probability measure P which extends every P_S. ($\mathcal{F}^*_S$ and P_S are as in Definitions 10.2 and 10.3 respectively)

Proof: See [8-1] or [14-1] .

THEOREM 10.2: (Restatement of above)

Each Stochastic Process whose finite dimensional distributions are specified, corresponds to at least one Stochastic Process. (i.e. the Process, given by Definition 10.4, corresponds to at least one process in the sense of Definition 10.1, provided the distributional conditions are consistent with each other.)

THEOREM 10.3:

Inverse Gaussian Processes as in the sense of Definition 10.4, exist in the sense of Definition 10.1.

Proof: Wasan [16-6], showed that conditions (1),(2), and (3) of Definition 10.4 are consistent, and uniquely determine the finite dimensional distributions of $(X_{t_1},\ldots,X_{t_n})$. (In fact, these are clearly Multivariate Inverse Gaussian random vectors. See Chapter 9.) Hence, by Theorem 10.2, the process exists in the sense of Definition 10.1.

This completes the proof.

We shall discuss the Inverse Gaussian Process in the context of Definition 10.1. However, we shall see by example, that little can be said about many aspects of the process, unless a suitable modification is chosen.

THEOREM 10.4:

Every Inverse Gaussian Process is monotonic increasing in probability.

Proof: $P(X_t \geqq X_s) = P(X_t - X_s \geqq 0) = 1$ whenever $t > s$, since Inverse Gaussian random variables are positive with probability 1, and P extends $P_{\{s,t\}}$.

This completes the proof.

Intuitively, one might feel that a Stochastic Process, monotonic increasing in probability is a.e. monotonic increasing. This need not be the case. This fact is dramatically demonstrated in the following example.

Example 10.1:

Let $(\Omega, \mathcal{F}, P, X_t : t \geq 0)$ be an Inverse Gaussian Process, whose measure P is complete. (This can be done by adjunction to $\mathcal{F}$, all subsets of sets of measure zero, and assigning P-measure zero to each of these.)

Let Y be the cumulative distribution function of $(X_2 - X_1)$, (i.e. of $X_{t=2} - X_{t=1}$). Y is uniformly distributed over the interval (0,1).

Define the following process: $(\Omega, \mathcal{F}, P, Z_t : t \geq 0)$, where

$$Z_t = X_t \quad \text{if} \quad Y \neq t$$
$$= 0 \quad \text{if} \quad Y = t$$

Since the event $\{w : Y(w) = t\}$ has probability zero for every t, with respect to the measure $P_{\{t\}}$ (and hence P), the finite dimensional distributions of the Z_t process are identical with those of the X_t process. The Z_t process is therefore Inverse Gaussian, and as such, by Theorem 10.4, is monotonic increasing in probability.

We shall now show that the Z_t process is not a.e. monotonic increasing. In fact, it will be shown that the w-set for which $X_t(w)$ is monotonic increasing, is a set of P-measure zero.

Argument :

Let c be an arbitrary constant such that $0 < c < 1$. Since Z_c is an Inverse Gaussian random variable, and P is an extension of $P_{\{c\}}$,

$$P(Z_c > 0) = 1.$$

Under the assumption that the Z_t process is a.e. monotonic increasing, one has by the completeness of P,

$$P(Z_t > 0, \text{ for all } t: c<t<1) = 1 .$$

Hence, taking the complimentary event to the above,

$$P(Z_t \leqq 0, \text{ for some } t: c<t<1) = 0. \qquad ..(1)$$

But from the definition of Y, $Z_t = 0$ if $t=Y$. Hence

$$P(Z_t \leqq 0, \text{ for some } t: c<t<1) \geqq P(c<Y<1) = 1 - c \qquad ..(2)$$

Since $c < 1$, the assumption that the Z_t process is a.e. monotonic increasing, leads to contradictory results (1) and (2). Therefore, the Z_t process is not a.e. monotonic increasing.

Since the set of all sample functions which are not monotonic increasing, contain

$$\{w: Z_t(w) \leqq 0, \text{ for some } t: c<t<1\} \cap N',$$

where $N' = \{w: Z_c(w) > 0\}$, a set of P-measure 1,

therefore, the set of all sample functions which are not monotonic increasing, contain a set of measure 1-c, (above by (2)), for arbitrary $c: 0<c<1$, and as such, is a set of measure 1.

We shall see in Chapter 11, that Definitions 10.5 and 10.6 are equivalent provided P is a complete probability measure, and the process is separable. The next theorem, given without proof, is therefore of great importance.

THEOREM 10.5 :

Every Stochastic Process, whose measure is complete, has a separable modification.

Proof: See [14-2] .

Example 10.2:

The Z_t modification of X_t of Example 10.1 is not separable.

<u>Argument</u> : Let $0<c<1$. Then for Y as in Example 10.1:

$$\left\{w: \inf_{t\in[c,1]} Z_t(w) \leq 0\right\} \supset \left\{w: c<Y<1\right\} .$$

That is,

$$\left\{w: \inf_{t\in[c,1]} Z_t(w) \leq 0\right\} \quad \text{contains a set of P-measure } 1-c. \quad (*)$$

But for any countable set S:

$$P\left(\left\{w: \inf_{t\in[c,1]\cap S} Z_t(w) \leq 0\right\}\right) = P\left(\bigcup_{t\in[c,1]\cap S}\left\{w: Z_t(w) \leq 0\right\}\right)$$

$$\leq \sum_{t\in[c,1]\cap S}\left(P\left(\left\{w: Z_t(w) \leq 0\right\}\right)\right) = 0 .$$

Therefore, $\inf_{t\in[c,1]\cap S} Z_t > 0$ a.e.

Now were the Z_t process separable, one would have:

$$\underset{t \in [c,1]}{\mathrm{INF}} Z_t > 0 \quad \text{a.e., and hence}$$

$$P\left(w: \underset{t \in [c,1]}{\mathrm{INF}} Z_t(w) \leqq 0 \right) = 0 \qquad (**)$$

Since $c<1$, the assumption that the Z_t process is separable leads to contradictory equations (*) and (**). The Z_t process is therefore not separable.

CHAPTER 11 : SEPARABLE STOCHASTIC PROCESSES, MONOTONIC INCREASING IN PROBABILITY

The first goal of this chapter will be to show that for separable processes whose measures are complete, Definitions 10.5 and 10.6 are equivalent. The reader will also be shown that for separable processes monotonic in probability, with complete measures, there exist right continuous modifications.

We shall now set up the machinery that will be used in our proofs.

Definition 11.1: (Upcrossings and Downcrossings)

Let $f: R_1 \longrightarrow R_1$ be an arbitrary function, where R_1 is the real line. Let $A = \{x_1, x_2, \ldots, x_n\} \subset R_1$ be an arbitrary finite set, with $x_1 < x_2 < \ldots < x_n$. Further, let $[a,b] \subset R_1$ be an arbitrary closed interval of positive length.

We shall first define $y_1, \ldots, y_n$ as follows:

$$\begin{aligned} y_1 &= 0 && \text{if} \quad f(x_1) < a \\ &= \tfrac{1}{2} && \text{if} \quad a \leqq f(x_1) \leqq b \\ &= 1 && \text{if} \quad b < f(x_1) \end{aligned}$$

and for $j = 2, 3, \ldots, n$:

$$\begin{aligned} y_j &= 0 && \text{if} \quad f(x_j) < a \\ &= y_{j-1} && \text{if} \quad a \leqq f(x_j) \leqq b \\ &= 1 && \text{if} \quad b < f(x_j) \; . \end{aligned}$$

Consider the sequence $y_1, \dots y_n$:

U = the number of times that 0 is immediately followed by 1

= the "number of upcrossings of $[a,b]$ by $\{f(x): x \in A\}$."

We shall denote this by:

$$U = U\left(f, A, [a,b]\right) .$$

Also,

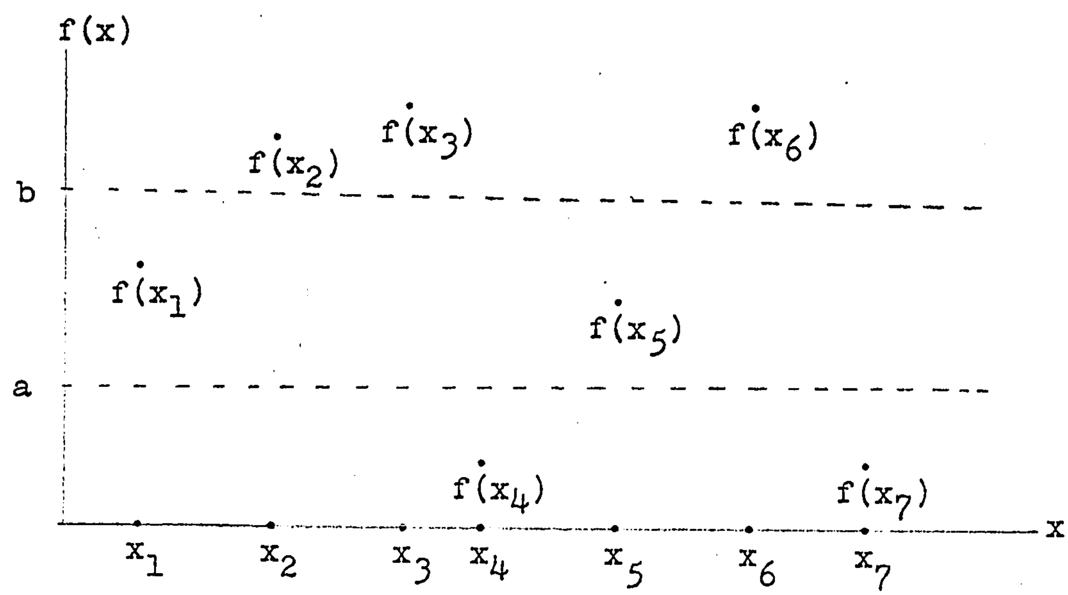
D = the number of times that 1 is immediately followed by 0

= the "number of downcrossings of $[a,b]$ by $\{f(x): x \in A\}$."

We shall denote this by:

$$D = D\left(f, A, [a,b]\right) .$$

Example 11.1:



Let $A = \{x_1, \dots, x_7\}$. The y-sequence by definition is: $\frac{1}{2}, 1, 1, 0, 0, 1, 0$. Therefore:

$$U\left(f,A,[a,b]\right) = 1 \quad \text{and} \quad D\left(f,A,[a,b]\right) = 2.$$

Definition 11.2: (Upcrossings, Downcrossings: infinite case)

Let $f: R_1 \longrightarrow R_1$ be an arbitrary function; $B \subset R_1$ be an arbitrary set; and $[a,b]$ be an arbitrary closed interval of finite, positive length.

We define the number of upcrossings and downcrossings of $[a,b]$ by $\{f(x): x \in B\}$ respectively by:

$$U\left(f,B,[a,b]\right) = \sup_{\substack{A \subset B \\ A \text{ finite}}} U\left(f,A,[a,b]\right), \text{ and}$$

$$D\left(f,B,[a,b]\right) = \sup_{\substack{A \subset B \\ A \text{ finite}}} D\left(f,A,[a,b]\right).$$

Lemma 11.1:

Let $f: [0,\infty) \longrightarrow R_1$ be an arbitrary function, and S be a dense countable subset of $[0,\infty)$ satisfying:

for every interval I, with endpoints in S,

$$\sup_{x \in I} f(x) = \sup_{x \in I \cap S} f(x) \quad \text{and} \quad \inf_{x \in I} f(x) = \inf_{x \in I \cap S} f(x).$$

Then for every finite set $A \subset [0,\infty)$, and every interval I with endpoints in S,

$$D\left(f, I\ A, [a,b]\right) \leqq D\left(f, I\ S, [a,b]\right) \qquad ..(1).$$

Proof: We shall assume that for some finite A, and interval I_o with endpoints in S, that (1) is false, and obtain a contradiction. That is, we assume

$$n_o = D\left(f, I_o \cap A, [a,b]\right) > D\left(f, I_o \cap S, [a,b]\right).$$

It is clear that n_o is finite, since n_o must be smaller than the number of elements of the finite set, A.

Since S is dense, it is easily seen that I_o can be partitioned into n_o intervals, each of which has its endpoints in S, and each, when intersected with A, has exactly one downcrossing of [a,b] by f.

Since $D\left(f, I_o \cap S, [a,b]\right) < n_o$, at least one of the n_o intervals, above, when intersected with S, has no downcrossing of [a,b] by f. Let one such interval be denoted by I^*.

We have thus constructed an interval I^*, with endpoints in S, such that

$$D\left(f, I^* \cap S, [a,b]\right) = 0 \quad \text{and} \quad D\left(f, I^* \cap A, [a,b]\right) = 1 \quad ..(2).$$

Since it is clear that for any B, such that $D\left(f, B, [a,b]\right)$ is finite: $\left| D\left(f, B, [a,b]\right) - U\left(f, B, [a,b]\right) \right| \leqq 1$, (see Definitions 11.1 and 11.2), two cases can exist. These are: Case 1: $U\left(f, I^* \cap S, [a,b]\right) = 0$; Case 2: $U\left(f, I^* \cap S, [a,b]\right) = 1$. We shall in turn treat each, and produce an appropriate contradiction.

Case 1: $U\left(f, I^* \cap S, [a,b]\right) = 0$

Since we also have $D\left(f, I^* \cap S, [a,b]\right) = 0$, then either:

(a) $\sup_{x \in I^* \cap S} f(x) \leqq b$ or (b) $\inf_{x \in I^* \cap S} f(x) \geqq a$.

(Otherwise, an upcrossing or a downcrossing would exist.)

But in view of the given properties of f, this implies that either

(a) $b \geqq \sup_{x \in I^*} f(x) \geqq \sup_{x \in I^* \cap A} f(x)$ or

(b) $a \leqq \inf_{x \in I^*} f(x) \leqq \inf_{x \in I^* \cap A} f(x)$

This in turn implies:

$D\left(f, I^* \cap A, [a,b]\right) = 0$, contrary to equation (2), which gives $D\left(f, I^* \cap A, [a,b]\right) = 1$. Hence Case 1 cannot occur.

Case 2: $U\left(f, I^* \cap A, [a,b]\right) = 1$

Partition I^* into intervals J_1 and J_2 , with endpoints in S , such that $\sup\left\{x : x \in J_1\right\} = \inf\left\{x : x \in J_2\right\}$ and

$\sup_{x \in J_1 \cap S} f(x) \leqq b$ and $\inf_{x \in J_2 \cap S} f(x) \geqq a$.

(This is possible, since S is dense, and since $D\left(f, I^* \cap S, [a,b]\right) = 0$.)

But by the given properties of f, we obtain:

$b \geqq \sup_{x \in J_1} f(x) \geqq \sup_{x \in J_1 \cap A} f(x)$ and

$$a \leq \inf_{x \in J_2} f(x) \leq \inf_{x \in J_2 \cap A} f(x) \quad .$$

That is, $f(x)$ is never above b for $x \in J_1 \cap A$, and $f(x)$ is never below a for $x \in J_2 \cap A$. Therefore,

$D\left(f, I^* \cap A, [a,b]\right) = 0$, since $J_1 \cup J_2 = I^*$, and J_1 is "to the left of " J_2. Again, this gives a contradiction of equation (2), which gives $D\left(f, I^* \cap A, [a,b]\right) = 1$. Therefore, Case 2 cannot occur.

Hence, for every finite set A, and every interval I, with endpoints in S,

$D\ f, I \cap A, [a,b] \leq D\left(f, I \cap S, [a,b]\right)$, since contrary assumptions led us to the contradiction:

"There exists an interval I^*, with endpoints in S, such that $0 = D\left(f, I^* \cap A, [a,b]\right) = 1$."

This completes the proof.

Lemma 11.2:

Under the conditions of Lemma 11.1, for any interval I with endpoints in S,

$$D\left(f, I, [a,b]\right) = D\left(f, I \cap S, [a,b]\right) \quad .$$

Proof: By Lemma 11.1, one has

$$D(f, I, [a,b]) = \sup_{\substack{A \subset I \\ A \text{ finite}}} D\left(f, A, [a,b]\right) \leq D\left(f, I \cap S, [a,b]\right) \quad (*).$$

On the other hand, Definition 11.2 gives:

$$D\left(f, I \cap S, [a,b]\right) = \sup_{\substack{A \subset I \cap S \\ A \text{ finite}}} D\left(f, A, [a,b]\right) \leqq \sup_{\substack{A \subset I \\ A \text{ finite}}} D\left(f, A, [a,b]\right)$$

$$\leqq D\left(f, I, [a,b]\right) . \qquad (**)$$

Combining the inequalities (*) and (**), one has:

$$D\left(f, I, [a,b]\right) = D\left(f, I \cap S, [a,b]\right) .$$

This completes the proof.

Lemma 11.2 will be instrumental in proving THEOREM 11.1. The result will be used with S as the universal separating set of a separable Stochastic Process, and for almost every w, by the definition of separability, $X_t(w)$ will have the same property that f had in Lemmas 11.1 and 11.2.

THEOREM 11.1 :

Let $(\Omega, \mathfrak{F}, P, X_t, t \geqq 0)$ be a separable Stochastic Process, monotonic increasing in probability, and complete measure P.

Then the process is a.e. monotonic increasing.

Proof: Let S be a universal separating set for the process. Without loss of generality, we may select S as a dense subset of the positive real numbers. (The rational numbers together with an arbitrary universal separating set form a dense universal separating set, for example.) Also, by definition S is countable.

One has therefore,

$$P\left\{w: X_t(w) \geqq X_s(w) \quad \text{for all } t,s \in S, \text{with } t > s\right\}$$

$$= P\left(\bigcap_{\substack{t,s\in S\\ t>s}} \left\{w: X_t(w) \geqq X_s(w)\right\}\right) = 1.$$

The above follows from the facts that since the process is monotonic increasing in probability, $\{w: X_t(w) \geqq X_s(w)\}$ is an event of probability 1, whenever $t > s$, and countable intersections of events of probability 1, have probability 1.

Therefore, noting that for arbitrary finite intervals I, $[a,b]$:

$$\left\{w: D\left(X_t(w), I\cap S, [a,b]\right) = 0\right\} \supset$$

$$\left\{w: X_t(w) \geqq X_s(w) \quad \text{for all } t,s \quad S \quad \text{with } t > s\right\},$$

one has by the completeness of the measure P,

$$P\left\{w: D\left(X_t(w), I\cap S, [a,b]\right) = 0\right\} = 1 \qquad ..(1).$$

Next, let $\mathcal{J} = \left\{[\alpha,\beta]: \alpha,\beta \in S \text{ with } \beta > \alpha\right\}$

= the set of all closed intervals with endpoints in S.

Let $E_{a,b} = \left\{w: D\left(X_t(w), I\cap S, [a,b]\right) = D\left(X_t(w), I, [a,b]\right), \forall\ I \in \mathcal{J}\right\}$

Therefore, by Lemma 11.2,

$$E_{a,b} \supset \bigcap_{I\in\mathcal{J}} \left\{w: \sup_{t\in I} X_t(w) = \sup_{t\in I\cap S} X_t(w), \ \inf_{t\in I} X_t(w) = \inf_{t\in I\cap S} X_t(w)\right\}$$

But since the process is separable, one has:

$$P\left(\left\{w: \sup_{t\in I} X_t(w) = \sup_{t\in I\cap S} X_t(w),\ \inf_{t\in I} X_t(w) = \inf_{t\in I\cap S} X_t(w)\right\}\right) = 1$$

holds for every interval I.

Also, it is clear that since there are only countably many choices for the endpoints of the intervals in $\mathcal{J}$, $\mathcal{J}$ is itself a countable set.

Therefore $E_{a,b}$ contains a countable intersection of events of measure 1, and hence it contains an event of measure 1. By the completeness of P, one hence obtains:

$$P(E_{a,b}) = 1 \qquad ..(2).$$

By (1), and the fact that $\mathcal{J}$ is countable, one has:

$$P\left(\left\{w: D\left(X_t(w), I\cap S, [a,b]\right) = 0,\ \forall\ I\in\mathcal{J}\right\}\right) = 1,$$

and rewriting (2), using the definition of $E_{a,b}$ above,

$$P\left(\left\{w: D\left(X_t(w), I\cap S, [a,b]\right) = D\left(X_t(w), I, [a,b]\right),\ \forall\ I\in\mathcal{J}\right\}\right) = 1.$$

Combining the above, one obtains:

$$P\left(\left\{w: D\left(X_t(w), I, [a,b]\right) = 0,\ \forall\ I\in\mathcal{J}\right\}\right) = 1 \qquad ..(3)$$

Now if a path $X_t(w_o)$ is not monotonic increasing, there will exist an $I\in\mathcal{J}$, and rational numbers a,b with a<b, such that

$$D\left(X_t(w_o), I, [a,b]\right) \geqq 1.$$

Hence, if $E = \left\{w: X_t(w) \text{ is monotonic, nondecreasing in } t\right\}$,

$$E = \bigcap_{\substack{a<b \\ a,b \text{ rational}}} \left\{w: D\left(X_t(w), I, [a,b]\right) = 0,\ \forall\ I\in\mathcal{J}\right\}$$

That is, E is a countable intersection of events of probability 1,and is therefore itself an event of probability 1. The process is therefore a.e. monotonic increasing.

This completes the proof.

The next Theorem is a converse to THEOREM 11.1, and under restricted conditions, characterizes separability.

THEOREM 11.2:

Let $(\Omega, \mathcal{F}, P, X_t : t \geqq 0)$ be a Stochastic Process, a.e. monotonic increasing, with complete measure P.

Then:

(a) The process is monotonic increasing in probability, and

(b) If the process is also continuous in probability, Definition 10.7, it is separable.

Proof:

(a) Let s,t be arbitrary real numbers with $0<s<t$.

$\{w : X_t(w) \geqq X_s(w)\} \supset \{w : X_t(w)$ is monotonic, non-decreasing in $t\}$.

Therefore, $\{w : X_t(w) \geqq X_s(w)\}$ contains an event of probability 1, and by the completeness of P, it must itself be an event of probability 1. (Note: For (a) completeness of P is not really needed. P, being a monotone set function, and X_s, X_t being measurable functions give this result.)

(b) Let S = the set of positive rational numbers.

Since the process is a.e. monotonic increasing,

$$P\left(\left\{w: \sup_{t\in[a,b]} X_t = X_b \ , \ \inf_{t\in[a,b]} X_t = X_a\right\}\right) = 1 \qquad ..(1).$$

Let $t_1, t_2, \ldots$ be an increasing sequence of rationals in $[a,b]$ such that $\lim_{n\to\infty} t_n = b$.

By the continuity in probability, for every $\epsilon > 0$,

$$\lim_{n\to\infty} P\left(\left\{w: \left|X_{t_n} - X_b\right| > \epsilon\right\}\right) = 0 .$$

Hence, for every $\epsilon > 0$, since P is a monotone set function,

$$P\left(\left\{w: \sup_{t\in[a,b]\cap S} X_t \geqq X_b - \epsilon\right\}\right) = 1.$$

That is, since the choice of ϵ is arbitrary,

$$P\left(\left\{w: \sup_{t\in[a,b]\cap S} X_t \geqq X_b\right\}\right) = 1.$$

Also, by equation (1), we have

$$P\left(\left\{w: \sup_{t\in[a,b]\cap S} X_t \leqq X_b\right\}\right) = 1.$$

Combining the above equations, we readily obtain:

$$P\left(\left\{w: \sup_{t\in[a,b]\cap S} X_t = X_b\right\}\right) = 1.$$

Similarly, by taking a sequence of rationals in $[a,b]$, decreasing to a, we obtain

$$P\left(\left\{w: \inf_{t\in[a,b]\cap S} X_t = X_a\right\}\right) = 1.$$

Combining the last two equations with (1), it is readily seen that:

$$\sup_{t\in[a,b]} X_t = \sup_{t\in[a,b]\cap S} X_t = X_b \quad \text{a.e.} \quad \text{and}$$

$$\inf_{t\in[a,b]} X_t = \inf_{t\in[a,b]\cap S} X_t = X_a \quad \text{a.e.,}$$

for any closed interval $[a,b] \subset [0,\infty)$.

This is clearly equivalent to the conditions for separability, since open intervals can be represented as countable unions of closed intervals, and closed intervals can be represented as countable intersections of open intervals, and countable intersections or unions of events of probability 1, are events of probability 1.

This completes the proof.

We shall now develop an a.e. right continuous modification of certain types of Stochastic Processes.

Definition 11.3:

Let $(\Omega, \mathcal{F}, P, X_t : t \geqq 0)$ be a Stochastic Process, which is a.e. monotonic increasing, with $X_0 = 0$.

We shall define X_t^+ for $t \geqq 0$, in the following way:

$$X_t^+(w) = 0 \quad \text{if } w \in N$$
$$= \lim_{h \downarrow 0} X_{t+h}(w) \quad \text{if } w \in (\Omega - N),$$

where $N = \{w : X_t(w) \text{ is not monotonic, non-decreasing in } t\}$.

Lemma 11.3:

For fixed $w \in \Omega$, $X_t^+(w)$ is well defined for all $t \geqq 0$.

Proof: By Definition 11.3, the result is evident for $w \in N$.

Let $w \in (\Omega - N)$, and let $\{h_n\}$ and $\{h'_n\}$ be two sequences of positive numbers decreasing to 0, as n tends to infinity.

Since both sequences converge to 0 from above, for every integer $n \geqq 0$, there exists an integer $m \geqq 0$, such that

$$h_n > h'_m .$$

By the monotonicity of X_t, for all $w \in (\Omega - N)$, we have

$$\lim_{n\to\infty} X_{t+h_n}(w) \geqq \lim_{n\to\infty} X_{t+h'_n}(w)$$

Reversing the roles of the sequences h_n and h'_n, the same argument gives the above inequality in the opposite sense. Therefore, we have shown:

$$\lim_{n\to\infty} X_{t+h_n}(w) = \lim_{n\to\infty} X_{t+h'_n}(w) .$$

Hence, the limit exists independently of the choice of the sequence converging to zero from above. That is, for $w \in \Omega - N$,

$$X_t^+(w) = \lim_{h \searrow 0} X_{t+h}(w) \quad \text{is also well defined.}$$

This completes the proof.

The next Lemma will further assert that X_t^+ is a Stochastic Process.

Lemma 11.4:

$(\Omega, \mathcal{F}, P, X_t^+ : t \geqq 0)$ is a Stochastic Process, where X_t^+ is defined in Definition 11.3.

Proof: By Lemma 11.3, for $a \geqq 0$, and $b < 0$:

$$\left\{w : X_t^+(w) \leqq a\right\} = \bigcap_{n=1}^{\infty}\left\{w : X_{t+1/n}(w) \leqq a\right\} \cup N \qquad \text{and}$$

$$\left\{w : X_t^+(w) \leqq b\right\} = \phi \qquad \text{(Since } X_o = 0).$$

Since the right hand sides of both equations above are members of $\mathcal{F}$, we have for each t, X_t^+ is a random variable. In view of Definition 10.1,

This completes the proof.

THEOREM 11.3 :

Let $(\Omega, \mathcal{F}, P, X_t : t \geqq 0)$ be a Stochastic Process with complete measure P, monotonic increasing in probability and contiuous in probability, with $X_o = 0$.

Then there exists a separable modification of the process, which is a.e. monotonic increasing, and a.e. right continuous. (i.e. there exists a separable modification with almost every sample function being right continuous and monotonic, non-decreasing.)

Proof: (five steps)

(i) By THEOREM 10.5, the process has a separable modification. Since the properties, described in probability, are preserved under modifications, this modification is monotonic increasing in probability. Hence by THEOREM 11.1, the modified process $(\Omega, \mathcal{F}, P, X_t : t \geqq 0)$ is a.e. monotonic increasing.

(ii) Next, X_t^+ will be shown to be a modification of X_t, the separable version of the process.

By part (i) above, and Definition 11.3,

$$P(X_t \leqq X_t^+ \leqq X_{t+1/n}) = 1 \qquad \text{for all} \quad t \geqq 0,\ n \in Z^+.$$

But by continuity in probability we have, for all $\epsilon > 0$,

$$\lim_{n \to \infty} P(|X_{t+1/n} - X_t| > \epsilon) = 0 .$$

Combining the two equations above, we obtain, for all $\epsilon > 0$,

$$P(X_t \leqq X_t^+ \leqq X_t + \epsilon) = 1.$$

Since the choice of ϵ is arbitrary, we have

$$P(X_t = X_t^+) = 1.$$

Therefore, X_t^+ is indeed a modification of X_t.

(iii) Next, X_t^+ will be shown to be a.e. monotonic increasing.

Let $A = \{w : X_t(w)$ is a monotonic non-decreasing function of $t\}$.

By part (i) of this proof, again assuming X_t is separable,

$$P(A) = 1.$$

But for every $w \in A$, and whenever $0 \leqq t < s$, we have, by the monotonicity of $X_t(w)$,

$$X_s^+(w) = \lim_{h \downarrow 0} X_{s+h}(w) \geqq \lim_{h \downarrow 0} X_{t+h}(w) = X_t^+(w) .$$

Therefore for every $w \in A$, $X_t^+(w)$ is a monotonic non-decreasing function of t. Hence,

$B = \{w: X_t^+(w) \text{ is monotonic non-decreasing in } t\} \supset A$,

a set of P-measure 1. Therefore, by the completeness of P, B is a set of P-measure 1.

X_t^+ is thus indeed an a.e. monotonic increasing process.

(iv) The $X_t^+(w)$ process is a separable modification, since by hypothesis, $X_t(w)$ is continuous in probability, and so must be its modification (by (ii)) X_t^+ ; X_t^+ is a.e. monotonic increasing; and thus, the conditions of THEOREM 11.2(b), being satisfied, imply the separability of X_t^+.

(v) It remains only to show, that X_t^+ is a.e. right continuous. For $w \in A = \{w: X_t(w) \text{ is monotonic non-decreasing in } t\}$, with X_t taken as a separable modification, we have for every $h > 0$, that since the X_t process is a.e. monotonic increasing,

$$X_{t+h}^+(w) \leqq X_{t+2h}(w).$$

Taking limits on both sides, and applying Lemma 11.3,

$$\lim_{h \downarrow 0} X_{t+h}^+(w) \leqq \lim_{h \downarrow 0} X_{t+2h}(w) = X_t^+(w)$$

But since by (iii), X_t^+ is a.e. monotonic increasing, we have

$$\lim_{h \downarrow 0} X_{t+h}^+(w) \geqq X_t^+$$, on a set containing A. See (iii).

Therefore, for every $w \in A$, we have

$$\lim_{h \downarrow 0} X^+_{t+h}(w) = X^+_t \quad , \text{ for all } t \geqq 0 .$$

But by (i), the set A, of monotonic non-decreasing $X_t(w)$ sample functions, is an event of probability 1. In view of the completeness of P, the set of all right continuous sample functions, is an event of probability 1. Therefore X^+_t is indeed a.e. right continuous.

We have thus constructed a modification X^+_t of X_t, such that

(1) X^+_t is separable

(2) X^+_t is a.e. monotonic increasing

(3) X^+_t is a.e. right continuous.

This completes the proof.

CHAPTER 12 : STOCHASTIC INTEGRATION WITH RESPECT TO MISI PROCESSES

In this chapter, we shall further restrict the processes occurring in the discussions to the following:

Definition 12.1:

A MISI Stochastic Process $(\Omega, \mathcal{F}, P, X_t : t \geqq 0)$ is a separable process, with complete measure P, satisfying:

(1) process is monotone increasing in probability,

(2) is an independent increment process, that is, if $0 = t_o \leqq t_1 \leqq t_2 \leqq \dots \leqq t_n$, then

$$\left\{ \left(X_{t_{j+1}} - X_{t_j} \right) : j = 0, 1, \dots, n-1 \right\}$$ are a set of mutually independent random variables,

(3) is a stationary increment process, that is, if h is an arbitrary non-negative number and $0 = t_o \leqq t_1 \leqq t_2 \leqq \dots \leqq t_n$, the joint distribution of

$$(X_{t_1+h} - X_{t_o+h} , X_{t_2+h} - X_{t_1+h} , \dots , X_{t_n+h} - X_{t_{n-1}+h})$$

does not depend on h, and

(4) $X_o = 0$.

MISI means Monotonic, with Independent and Stationary Increments. The separable Inverse Gaussian and Poisson processes are examples of MISI processes.

Definition 12.2: (Stieltjes Stochastic Integral)

Let $(\Omega, \mathcal{F}, P, X_t : t \geqq 0)$ be a MISI process, and $f: [a,b] \longrightarrow R_1$ be a continuous function with $0 \leqq a < b < \infty$. We define

$$I_f(w) = \int_a^b f(t)\,dX_t(w) \qquad w \in \Omega - N$$

$$= 0 \qquad w \in N,$$

with $N = \{w: X_t(w) \text{ is not monotonic non-decreasing in } t\}$, and the integral is the ordinary Stieltjes integral.

$I_f(w)$ is called the Stieltjes Stochastic Integral of f with respect to the MISI process X_t over the interval $[a,b]$.

THEOREM 12.1 :

If X_t is a MISI process, which is finite a.e., then I_f exists for almost every w.

Proof: Let $N_1 = \{w: X_b(w) = \infty\}$, and N be as in Definition 12.1. Since a MISI process satisfies the conditions of THEOREM 11.1, N is therefore a set of measure 0. Also, since the process is finite a.e., N_1 is also a set of measure 0.

Hence for $w \in (\Omega - (N \cup N_1))$, a set of measure 1, we have: $f(t)$ is continuous and $X_t(w)$ is a monotonic, non-decreasing function of bounded variation. By the elementary theory of Riemann-Stieltjes integration, [13-1], I_f exists on $\{\Omega - (N \cup N_1)\}$.

This completes the proof.

THEOREM 12.2:

For MISI processes, finite a.e., $I_f(w)$ is a random variable.

Proof: Clearly, it is sufficient to prove $I_{f^+}(w)$ is a random variable, where $f^+ = MAX(f,0)$, for if this were the case, $I_{f^-}(w) = I_{(-f)^+}(w)$ would likewise be a random variable, and since I_{f^+} and I_{f^-} are finite a.e., by THEOREM 12.1, $I_f = I_{f^+} - I_{f^-}$ would be a random variable.

Let P_n be the following sequence of partitions of $[a,b]$:

$$P_n = \left\{ t_{k,n} = a + \frac{k}{n}(b-a) : k = 0,1,\ldots,n \right\} \text{ and } n = 1,2,\ldots$$

Let $$C_{k,n} = \text{INF}\left\{ f^+(t) : t \in [t_{k-1,n}, t_{k,n}] \right\} \quad k = 1,\ldots n .$$

Finally, let

$$S_n^+(w) = \sum_{k=1}^{n} C_{k,n}\left(X_{t_{k,n}}(w) - X_{t_{k-1,n}}(w) \right) .$$

Since f is continuous, it is clear that f^+ is continuous. Hence, by THEOREM 12.1, we have for almost every w,

$$\lim_{r\to\infty} \sup_{n>r} S_n^+(w) = I_{f^+}(w) , \qquad ..(1)$$

(since the $S_n^+(w)$ are merely the Riemann-Stieltjes sums, and these converge a.e. to the Riemann-Stieltjes integral.)

But for each n, $S_n^+(w)$ is a linear combination of random variables, and hence is a random variable. Since the lim SUP of a family of random variables, which is countable, is itself a random variable, the left hand side of equation (1) is a random variable. But (1) holds for almost every w. Hence by the completeness of the measure P, $I_{f^+}(w)$ is a random variable.

In view of the remark at the start of the proof, we obtain that $I_f(w)$ is a random variable.

This completes the proof.

THEOREM 12.3 :

For a MISI process, finite a.e., the following hold:

If f and g are continuous functions from $[a,b]$ to R_1, with $0 \leqq a < b < \infty$, and c, any real constant, then

(1) $I_{f+g}(w) = I_f(w) + I_g(w)$ a.e.

(2) $I_{cf}(w) = cI_f(w)$

That is, the integral is linear with respect to the functions f and g.

Proof: The proof of the above follow immediately from the corresponding results for ordinary Riemann-Stieltjes integrals.

THEOREM 12.4 :

For a MISI process X_t, and $f:[a,b] \longrightarrow R_1$, a continuous function, with $0 \leqq a < b < \infty$, $I_{f^+}(w)$ and $I_{f^-}(w)$ are stochastically independent, provided the process is a.e finite.

Proof: Let P_n be the following partition of $[a,b]$:

$$P_n = \left\{ t_{k,n} = a + \frac{k}{n}(b-a) \ , \ k = 0,1,\dots,n \right\} \quad , \ n = 1,2,\dots$$

Let
$$C_{k,n} = \text{INF}\left\{ f^+(t) : t \in \left[t_{k-1,n} , t_{k,n}\right]\right\}, \ k = 1,\dots,n,$$
$$D_{k,n} = \text{INF}\left\{ f^-(t) : t \in \left[t_{k-1,n} , t_{k,n}\right]\right\}, \ k = 1,\dots,n.$$

Next, let
$$S_n^+(w) = \sum_{k=1}^{n} C_{k,n}\left(X_{t_{k,n}}(w) - X_{t_{k-1,n}}(w)\right) \quad \text{and}$$

$$S_n^-(w) = \sum_{k=1}^{n} D_{k,n}\left(X_{t_{k,n}}(w) - X_{t_{k-1,n}}(w)\right) .$$

We now note that P_{2^q} is a refinement of P_{2^r}, whenever $q > r$, $q,r \in Z^+$. Hence by the fundemental results of lower Riemann-Stieltjes sums: THEOREM 12.1 gives

$$\begin{aligned} S^+_{2^r}(w) &\nearrow I_{f^+}(w) \quad \text{a.e., and} \\ S^-_{2^r}(w) &\nearrow I_{f^-}(w) \quad \text{a.e. as } r \to \infty. \end{aligned} \qquad (*)$$

But from the definition of $C_{k,n}$ and $D_{k,n}$, it is clear that $C_{k,n}$ and $D_{k,n}$ cannot both be positive for fixed k and n, (for if this were the case, there would exist an interval such that, f is always positive and always negative.) Hence at least one of $C_{k,n}$ and $D_{k,n}$ must vanish.

Therefore $S_n^+(w)$ is a linear combination of the increments $\left(X_{t_{k,n}} - X_{t_{k-1,n}}\right)$ such that $f(t) > 0$, for all t in the interval $[t_{k-1,n}, t_{k,n}]$, (i.e. $C_{k,n} > 0$).

Also, $S_n^-(w)$ is a linear combination of the increments $\left(X_{t_{k,n}} - X_{t_{k-1,n}}\right)$ such that $f(t) < 0$, for all t in the interval $[t_{k-1,n}, t_{k,n}]$, (i.e. $D_{k,n} > 0$).

Since a MISI process has independent increments, and $S_n^+(w)$ is a function of increments of intervals disjoint of those upon which $S_n^-(w)$ depend, we see that for arbitrary n, $S_n^+(w)$ and $S_n^-(w)$ are independent random variables.

By (*), and the fact that P is a measure, and therefore continuous from below, we obtain for any $c > 0$, $d > 0$:

$$P[I_{f+}(w) > c,\ I_{f-} > d] = \lim_{r\to\infty} P[S_{2^r}^+(w) > c,\ S_{2^r}^-(w) > d]$$

$$= \lim_{r\to\infty}\left(P[S_{2^r}^+(w) > c]\ P[S_{2^r}^-(w) > d]\right) \qquad (\ S_n^+,\ S_n^-\ \text{Independent}))$$

$$= P[I_{f^+}(w) > c]\ P[I_{f^-}(w) > d]$$

That is, $I_{f^+}(w)$ and $I_{f^-}(w)$ are indeed independent random variables.

This completes the proof.

Remark: The assumption that the process be a.e. finite was never used in the proof. In fact, neither was the fact that $I_{f^+}(w)$ is a random variable dependent, on this fact. However, the result is of no value to our discussion unless $I_f(w)$ is a random variable, and this requires that the process be finite a.e..

The next theorem gives a rather remarkable result, which is proved using THEOREM 12.4.

THEOREM 12.5: (Characterization)

Let X_t be a MISI process, finite a.e., and unbounded in probability. (See Definition 10.13). Let f be a continuous function from $[a,b]$ to R_1, with $0 \leqq a < b < \infty$.

Then a necessary and sufficient condition that

(a) $I_f(w) = 0$ a.e. is that

(b) $f(t) = 0$ for all $t \in [a,b]$.

Proof: <u>Necessity of condition</u>

If $I_f(w) = 0$ a.e., then

$$I_{f^+}(w) = I_{f^-}(w) \quad \text{a.e.} \qquad ..(1)$$

Hence for any $c \geqq 0$, THEOREM 12.4 gives

$$P[I_{f^+}(w) < c] = P[I_{f^+}(w) < c,\ I_{f^-}(w) < c]$$

$$= P[I_{f^+}(w) < c]P[I_{f^-}(w) < c] = P[I_{f^+}(w) < c]^2 .$$

The above equation has solutions

$$P[I_{f^+}(w) < c] = 0 \text{ or } 1.$$

Therefore, there exists a constant $d \geqq 0$, such that

$$I_{f^+}(w) = I_{f^-}(w) = d \quad \text{a.e.}$$

(Since the process is finite a.e., d must be finite, by THEOREM 12.1)

We shall assume, for some $t_o \in [a,b]$, that $f(t_o) > 0$, and obtain a contradiction.

Since $f(t)$ is continuous, and $f(t_o) > 0$, there exists a positive number δ, and an interval of finite length, $[t_1,t_2] \subset [a,b]$, with $t_o \in [t_1,t_2]$, such that

$$f(t) \geqq \delta \text{ , for all } t \in [t_1,t_2].$$

By the monotonicity of almost every sample function, and by dominated convergence, we obtain:

$$I_{f^+}(w) \geqq \delta(X_{t_2} - X_{t_1}) \quad \text{a.e.}$$

Hence by the above, and the fact that a MISI process has stationary increments, as well as the fact that the process is unbounded in probability, we obtain:

$$P[I_{f^+}(w) > d] \geqq P[X_{t_2} - X_{t_1} > d/\delta] = P[X_{t_2-t_1} > d/\delta] > 0.$$

This is a contradiction of (2), which states $I_{f^+} = d$ a.e.

The assumption that $f(t)$ is positive for some $t \in [a,b]$, gives rise to a contradiction. Hence, for all $t \in [a,b]$,

$f(t) \leqq 0$.

The same arguement applied to $I_{f^-}(w)$ readily gives:

$f(t) \geqq 0$, for all $t \in [a,b]$.

Therefore, $f(t) = 0$ for all $t \in [a,b]$

<u>Sufficiency of condition</u>

If $f(t) = 0$ for all $t \in [a,b]$, THEOREM 12.1 gives

$$I_f(w) = \int_a^b 0 dX_t(w) \qquad \text{a.e.}$$

$$= 0 \qquad \text{a.e.}$$

This completes the proof.

THEOREM 12.6 : (Characterization)

Let X_t be a MISI process, finite a.e., and unbounded in probability. Let f and g be continuous functions from $[a,b]$ to R_1 , with $0 \leqq a < b < \infty$.

Then a necessary and sufficient condition that

(a) $I_f(w) = I_g(w)$ a.e. is that

(b) $f(t) = g(t)$ for all $t \in [a,b]$.

Proof: $I_f(w) = I_g(w)$ a.e. , by THEOREM 12.3, is equivalent to $I_{f-g}(w) = 0$ a.e., which in turn, by THEOREM 12.5, is equivalent to $f(t) - g(t) = 0$ for all $t \in [a,b]$.

This completes the proof.

A Note on the Integral:

For MISI processes, finite a.e., and unbounded in probability, the integral I_f defines a 1-1 mapping from the continuous functions with domain, $[a,b] \subset [0,\infty)$, to the set of random variables on $(\Omega, \mathcal{F}, P)$.

The separable Inverse Gaussian and Poisson Processes are examples of MISI processes, finite a.e., and unbounded in probability. The results of Chapters 11 and 12 therefore hold for these processes.

Table 12.1 lists some properties of the separable Inverse Gaussian Process, in particular.

Table 12.1: (For complete measures)

Properties of the Separable Inverse Gaussian Process

(1) Continuous in probability

(2) Monotone increasing in probability

(3) Unbounded in probability

(4) a.e. monotonic increasing

(5) MISI process

(6) If a and b are the parameters given in Definition 10.4, for any interval $I \subset [0,\infty)$, such that I is closed, Separability, (1) and (4) give:

$$P[\sup_{t \in I} X_t < c] = P[X_M < c] = F(c;aM,bM^2),$$

where M is the right hand endpoint of I, $\sup\{t : t \in I\}$. This result was obtained first in [16-7], by means of double Laplace Transform. The result is in general false, if the process is not separable. (In example 10.1, let Y_t be the following modification of X_t.

$$\begin{aligned} Y_t &= X_t \qquad Y \neq t \\ &= \infty \qquad Y = t. \end{aligned}$$
)

(7) By (6), the process is finite a.e.

(8) By (3),(5), and (7) all the results in Chapter 12, with regard to Stochastic Integrals hold.

(9) The process has a separable modification, a.e. monotonic increasing, and a.e. right continuous. (THEOREM 11.3)

APPENDIX 1: CONCLUSIONS AND CONJECTURES

(1) Testing for the mean of an Inverse Gaussian population when λ is unknown.

Let $X_1, \ldots, X_n$ be a random sample from $F(x;\mu,\lambda)$, where u and λ are unknown. Assuming that our decision is based on:

$$Z = n(n-1)(\bar{X}-\mu_o)^2 \mu_o^{-2} \left(\sum_{j=1}^{n} \left(\frac{\bar{X}}{X_j} - 1 \right) \right)^{-1} ,$$

how would one test:

$H_0 : \mu = \mu_o$ against

$H_A : \mu = \mu_a$ at significance level α?

Under H_0, THEOREM 6.1 gives: $Z \sim F_{1,n-1}$. Therefore the distribution of Z, under the null hypothesis does not depend on λ. One might conjecture that the best test based on Z, would be to reject if Z were too large. The following computation supports this conjecture:

Since $\bar{X}$ and $\sum_{j=1}^{n} (1/X_j - 1/\bar{X})$ are independent random variables, and the distribution of the latter does not depend on μ, in order to minimize E(Z), it is equivalent to minimize :

$E\left(\frac{(\bar{X} - \mu_o)^2}{\bar{X}\mu_o^2} \right)$ with respect to μ, the true mean of the population.

By THEOREM 6.1, $\bar{X} \sim IG(\mu, n\lambda)$. THEOREM 3.3 gives

$$E(\bar{X}^{-1}) = \frac{1}{\mu} + \frac{1}{n\lambda} .$$

Using the above, simple algebraic manipulation gives:

$$E\left[\frac{(\bar{X} - \mu_o)^2}{\mu_o^2 \bar{X}}\right] = \frac{1}{n\lambda} + \frac{(\mu-\mu_o)^2}{\mu_o^2 \mu} ,$$

where μ is the true mean. The expectation, above, is clearly minimum at $\mu = \mu_o$. Hence, intuitively, the smaller Z is, the closer we would expect the true mean to be to μ_o. A challenging problem would be to obtain the distribution of Z under H_A, and apply the Neyman-Pearson lemma, to check this conjecture.

(2) Analysis of Reciprocals.

Tweedie, [15-3], suggests a method of comparing the means of several Inverse Gaussian random samples, having the same λ. The method is known as "Analysis of Reciprocals". Let $X_{i1}, \ldots, X_{in}$ be a random sample of size n from $F(x, \mu_i, \lambda)$, $i = 1, \ldots, k$. Let $\bar{X}_i$ be the ith sample mean, and $\bar{X}$, be the mean of the nk observations, X_{ij}. Let

$$S = \sum_{i=1}^{k} (1/\bar{X}_i - 1/\bar{X}) \quad ; \quad T = \sum_{i=1}^{k} \sum_{j=1}^{n} (1/X_{ij} - 1/\bar{X}_i) .$$

The suggested method of testing :

$H_0: \mu_1 = \ldots = \mu_k$ versus H_A: H_0 false, is: reject if

$$\frac{n(n-1)kS}{(k-1)T} > F_{k-1,k(n-1),\alpha} .$$

Conjecture: This rejection region is UMP.

(3) For MISI processes, unbounded in probability, it was noted that the stochastic integral, defined in Chapter 12, can be regarded as a 1-1 mapping from the continuous functions on a closed interval to the random variables on the measure space of the process.

The definition of the integral is easily extended to funtions with at most countable discontinuities on the positive real numbers. Interesting problems to be investigated in further work are:

(1) What is the image space of the integral of an Inverse Gaussian MISI process? That is, what distributions can the integral have?

(2) Find a class of functions which will produce a canonical form for all infinitely divisible laws. Here the integration will be with respect to a Poisson MISI process. The extended definition would have to be used, and allowance would have to be made for addition of an independent Gaussian term.

The fact that the mapping is 1-1, of course, does not imply, that there is a unique way to obtain a given distribution. The class of functions selected, therefore, must bring about this uniqueness.

(4) Potential application of THEOREM 4.1 might occur in the following test of goodness of fit, for large samples:

H_0: $F = F_0$ versus H_1: $F = F_1$, where F is absolutely continuous, and

$$F_0(x) > F_1(x) \qquad x < X_o \ ,$$
$$< F_1(x) \qquad x > X_o \ , \text{ with } F_0(X_o) = F_1(X_o) = \tfrac{1}{2} \text{ , say.}$$

Here, the fact that the inequality, given in Definition 4.3, is violated, is not in itself a significant fact, but the fact that it was violated early, is significant.

The rejection of H_0 in favor af H_1 would be made on the joint facts that the inequality was violated, and that it was first violated before an appropriate value of x, determined from $F_0(x)$.

Note that are procedure is vastly different from locating the sup of the difference $F_0(x) - F_n(x)$, and determining where this occurs. The first violation of the inequality, is not necessarily the sup.

APPENDIX 2 : THE INVERSE GAUSSIAN PROCESS AS A CENTERED PROCESS

Definition A.2.1: (Centered Stochastic Process)

A stochastic process, $\{X_t : a \leqq t \leqq b\}$, is a centered stochastic process provided

(1) It is a stationary, independent increment process;

(2) For each $t \in (a,b)$,

$$X_{t^-} = \lim_{s \uparrow t} X_s \quad \text{and} \quad X_{t^+} = \lim_{s \downarrow t} X_s \quad , \text{ where}$$

the limits are sequential, exist with probability 1 and are independent of the choice of the sequence with probability 1;

(3) If (a) $P(X_t - X_s = C) = 1 \quad t \neq s$, or

(b) $P(X_{t^-} - X_{s^-} = C) = 1 \quad t \neq s$, or

(c) $P(X_{t^+} - X_{s^+} = C) = 1 \quad t \neq s$, then

$C = 0$;

(4) For all but countably many fixed $t \in (a,b)$,

$P(X_{t^-} = X_t = X_{t^+}) = 1$.

This definition occurs in Doob, [3-1].

THEOREM A.2.1 : (Characterization)

Let X_t , $a \leqq t \leqq b$, be a centered process, with no fixed points of discontinuity, (that is, continuous in probability). Then the following are equivalent:

(1) If the process is separable, almost every sample function is continuous; and

(2) For each $s,t \in [a,b]$, the random variable $X_t - X_s$ has a normal distribution. Process is separable.

Proof: See [3-2] .

THEOREM A.2.2:

A separable Inverse Gaussian process is a centered process.

Proof:

By definition, the process has stationary, independent increments. By Lemmas 11.3 and 11.4, and Definition 11.3, $X_{t+} = X_t^{+}$ satisfies (2) of Definition A.2.1. By THEOREM 11.3, X_{t+} is a modification of X_t . Hence $P(X_t = X_{t+}) = 1$.

The same arguments in Chapter 11, are easily seen to carry through for X_{t-} , giving analogues of Lemmas 11.3 and 11.4 , and THEOREM 11.3. Hence (1),(2), and (4) of Definition A.2.1 hold.

(3) clearly holds, since in each case, we have an Inverse Gaussian random variable, which, of course has an absolutely continuous distribution function.

X_t is therefore a centered process.

This completes the proof.

Remark: If a stochastic process is a.e. continuous, it is clearly separable. (See Definition 10.9, with S = the set of rational numbers.)

THEOREM A.2.3 :

There does not exist a modification of the Inverse Gaussian process, almost all of whose sample functions are continuous.

Proof:

We shall asume such a process exists, and obtain a contradiction.

By the above remark, the process must be separable. Since it does not have any fixed discontinuities, and is a centered process, THEOREM A.2.1 gives: $X_t - X_s$ is a Gaussian random variable, for $t \neq s$. This cannot happen, since Gaussian random variables are negative, with positive probability, while Inverse Gaussian random variables are negative with probability 0. (Contradiction)

This completes the proof.

Remark: a.e. right continuous and a.e. left continuous modifications of the Inverse Gaussian process exist. However, a.e. continuity cannot be attained. This is intuitively logical, since the Poisson process has all of the properties used above, except absolute continuity. The Poisson process is highly discontinuous.

APPENDIX 3: THE CAUCHY FUNCTIONAL EQUATION

In this section, we shall derive the solution of the Cauchy functional equation, with a minimum of assumptions.

THEOREM A.3.1:

Let $f(x)$ be an extended real valued function, satisfying

(1) $f(x) + f(y) = f(x+y)$ for all $x,y \in R_1$,

(2) $f(x)$ is bounded in an interval $[0,C]$.

Then $f(x) = f(1)x$ for all $x \in R_1$.

Proof:

<u>Step 1:</u> f is right continuous at 0.

Let $t \in [0,C]$. Since f is bounded on $[0,C]$, there exists a positive number K, such that

$f(t) < K$.

But by (1), for arbitrary $n \in Z^+$,

$$f(t/n) = \frac{f(t)}{n} .$$

Therefore, for all $t \in [0,C]$, $f(t/n) < K/n$.

That is, whenever $t < C/n$, $f(t) < K/n$. Hence:

$$\lim_{t \downarrow 0} f(t) = 0 .$$

By (1), $f(0) + f(y) = f(y)$, and therefore, $f(0) = 0$.

f is thus right continuous at 0.

<u>Step 2:</u> f is continuous at all $x \in R_1$.

(a) Right continuity:

By (1), $f(x) + f(t) = f(x+t)$. Therefore, for $x \in R_1$,

$$f(x) + \lim_{t \searrow 0} f(t) = \lim_{t \searrow 0} f(x+t) \quad \text{, and hence:}$$

$$f(x) = \lim_{t \searrow 0} f(x+t) .$$

f is right continuous at x.

(b) Left continuity:

By (1), $f(x) = f(x-t) + f(t)$. Therefore, for $x \in R_1$,

$$f(x) = \lim_{t \searrow 0} f(x-t) + \lim_{t \searrow 0} f(t)$$

$$= \lim_{t \searrow 0} f(x-t) .$$

f is left continuous, and hence continuous at x.

<u>Step 3:</u> The theorem holds for the rationals.

Let $x = p/q$ be a positive rational number, with p,q positive integers.

$f(x) = p\, f(1/q)$ (By (1)).

$f(1) = q\, f(1/q)$ (Put x = 1 in above)

Combination of the above yields:

$$f(x) = f(1) \frac{p}{q} = f(1)\, x .$$

This gives the theorem for negative rationals:

If y is a negative rational, ane obtains:

$f(y) + f(-y) = f(0) = 0$. Therefore,

$f(y) = -f(-y) = f(1)\, y$.

Therefore, the theorem indeed holds for the rationals.

Step 4: Completion of proof.

Let $x \in R_1$, and r_n , $n = 1,2,\ldots$ be a sequence of rationals converging to x. (Such sequences exist by the density of the rational numbers)

$$\text{By Step 2,} \quad f(x) = \lim_{n\to\infty} f(r_n)$$

$$= \lim_{n\to\infty} [f(1)\, r_n] \qquad (r_n \text{ rational})$$

$$= f(1)x .$$

This completes the proof.

Remark: If the domain of f is restricted to an interval containing 0, it clearly has a unique extension from (1) to the domain R_1 .

Remark: If there does not exist an interval containing 0 such that f is bounded, the only solutions of the required form are $f = \infty$, except possibly at $x = 0$, or $f = -\infty$, except possibly at $x=0$.

Remark: By letting $\phi(x) = \exp(f(x))$, we obtain:

$\phi(x)\phi(y) = \phi(x+y)$ has its unique solution

$\phi(x) = \exp(cx) \quad c \in R_1$, provided ϕ is bounded in an interval containing 0.

BIBLIOGRAPHY

[1] Arnold, B.C. (1967). "A Note on Multivariate Distributions with Specific Marginals." Journal of the American Statistical Association, 62, pp 1460-62.

[2] Birnbaum, Z.W. and Tingey, F.H. (1951). "One Sided Contours for Probability Distributions." Ann. Math. Stats., 22, pp 592-96.

[2-1] from formula (3.4) page 594, J(d,n,k) is obtained. Our integral is J(d,n,k) - J(d,n,k+1).

[2-2] Main result of paper.

[3] Doob, J.L. (1953). "Stochastic Processes". John Wiley, New York.

[3-1] Page 407

[3-2] Page 420

[4] Feller, W. (1957). "Introduction to Probability Theory and its Applications, Vol. I, 2nd Edition." John Wiley, New York.

[4-1] Page 52

[5] Ferguson, T. (1967). "Mathematical Statistics, A Decision Theoretic Approach." Academic Press, New York.

[5-1] Page 46, Rule 1

[6] Gnedenko, B.V. (1967). "Probability Theory, 4th Edition". Chelsea.

[6-1] Page 323

[7] Khatri, C.G. (1962)."A Characterization of the Inverse Gaussian Distribution." Ann. Math. Stats.,33,pp 800-803.

[7-1] Theorem.

[8] Kolmogorov, A. (1932). "Foundations of the Theory of Probability." Chelsea. (English version 1950)

[8-1]

[9] Munroe, M.E. (1953)."Introduction to Measure and Integration." Addison-Wesley.

[9-1] Page 233.

[10] Patil, G.P. and Seshadri, V. (1964)."Characterization Theorems for some Univariate Probability Distributions." The Journal of the Royal Statistical Society, Series B, 26,pp 286-92.

[10-1] Page 287, Theorem 1.

[11] Shachtman, R. (1968). "Sample Paths of Bounded Variation." Notices of the American Mathematical Society,15,p 936.

[12] Shuster, J. (1968). "A Note on the Inverse Gaussian Distribution Function." Journal Amer. Stat. Assoc.,63, to appear December, 1968.

[12-1] Theorem 2

[12-2] Theorem 1

[13] Taylor, A.E. (1965). "General Theory of Functions and Integration." Blaisdell Publishing Co.

[13-1] Page 394, Theorem 9-5. II

[14] Tucker, H. (1967). "A Graduate Course in Probability." Academic Press, New York.

[14-1] Page 30, Theorem 2

[14-2] Page 251, Theorem 2

[15] Tweedie, M.C.K. (1957). "Statistical Properties of Inverse Gaussian Distribution I and II". Ann. Math. Stats., 28,pp 362-77 and 696-705.

[15-1] Page 366 and Page 372.

[15-2] Page 373

[15-3] Page 369

[16] Wasan, M.T. (1966). "Monograph on Inverse Gaussian Distribution". Queen's University Press, Kingston, Ont.

[16-1] Page 2.6

[16-2] Page 2.12

[16-3] Page 2.18 and Page 2.21

[16-4] Page 9.22 (Section 9.7)

[16-5] Page 4.28 (Section 4.9)

[16-6] Page 3.2, last two paragraphs

[16-7] Page 3.10

[17] Wilks, S. (1962). "Mathematical Statistics." John Wiley, New York. (Corrected Edition)

[17-1] Page 339, 11.6.2.

TWO RESEARCH PAPERS, WRITTEN BY JONATHAN SHUSTER.

The results of these papers supplement those of this thesis. They were discovered after the typing of the manuscript had been completed. Both articles center around problems in testing hypotheses. As of Feb. 1969, they are in the hands of referees for the Annals of Mathematical Statistics.

LIKELIHOOD RATIO TESTS FOR INVERSE GAUSSIAN POPULATIONS,

by JONATHAN SHUSTER, McGILL UNIVERSITY. [1]

1. Introduction. In this paper, we shall develop the likelihood ratio test for the mean of an inverse Gaussian population, and further, we shall show that the "Analysis of Reciprocal" procedure, developed by Tweedie, [3], is in fact, the likelihood ratio test in its setting.

Definition: A random variable, X, is said to follow the inverse Gaussian distribution, with positive parameters μ and λ, if it has density function:

$$f(x;\mu,\lambda) = \begin{cases} \lambda^{\frac{1}{2}}(2\pi x^3)^{-\frac{1}{2}}\exp\left[-\tfrac{1}{2}\lambda\mu^{-2}(x-\mu)^2x^{-1}\right] & x > 0 \\ 0 & x < 0. \end{cases} \qquad ..(1)$$

2. The Fundamental Lemma. Let $X_{i_1}, \ldots, X_{i_{n_i}}$, $i=1,\ldots,k$, respectively, be k independently drawn random samples from the inverse Gaussian density $f(x;\mu_i,\lambda)$. Let $\bar{X}_i$ be the ith sample mean, and $\bar{X}$ be the mean of the totality of observations.

The joint maximum likelihood estimators of $\mu_1,\ldots,\mu_k$ and λ, respectively, are:

$$\hat{\mu}_i = \bar{X}_i\ , \ i = 1,\ldots,k, \text{ and}$$

$$\hat{\lambda} = \left(\sum_{i=1}^{k} n_i\right)\left[\sum_{i=1}^{k}\ \sum_{j=1}^{n_i}\left(X_{ij}^{-1} - \bar{X}_i^{-1}\right)\right]^{-1}.$$

[1] Work supported by National Research Council of Canada.

Proof: The proof is very elementary, and therefore will merely be outlined. As usual, the log of the joint density of the $(n_1 + \cdot \cdot + n_k)$ random variables is maximized, the first derivatives of the function giving the estimates, and the matrix of second derivatives, diagonal at the point in $(1+k)$-dimensional space, given by the estimate, checking that the function indeed has been maximized.

3. Likelihood Ratio Test for μ. We are now prepared to determine the likelihood ratio test for μ, the population mean of an inverse Gaussian random sample:

H_o: $\mu = \mu_o$ against H_a: $u \neq u_o$, at level α.

Let $X_1, \ldots, X_n$ be a random sample from a population whose density is $f(x;\mu,\lambda)$. Then the rejection region of the likelihood ratio test is: Reject if $T =$

$$(n-1)\mu_o^{-2}\,\overline{X}^{-1}\,(\overline{X}-\mu_o)^2 n \left[\sum_{j=1}^{n}\left(X_j^{-1} - \overline{X}^{-1}\right)\right]^{-1} > K_\alpha ,$$

where K_α, determined in section 5, is a constant depending only on α.

Proof: Let $L(\mu,\lambda)$ be the joint density of $X_1, \ldots, X_n$, for given μ and λ. Let $A = (0,\infty)X(0,\infty)$; and $B = \{\mu_o\}X(0,\infty)$.

The likely ratio test is therefore: Reject if

$$\left(\text{Sup}\left\{L(\mu,\lambda):(\mu,\lambda)\in A\right\}\right)\left(\text{Sup}\left\{L(\mu,\lambda):(\mu,\lambda)\in B\right\}\right)^{-1} > C_\alpha .$$

Applying the fundamental lemma, and substituting the functional form given by equation (1), we obtain:

$$\left(\operatorname{Sup}\left\{L(\mu,\lambda):(\mu,\lambda)\in A\right\}\right)\left(\operatorname{Sup}\left\{L(\mu,\lambda):(\mu,\lambda)\in B\right\}\right)^{-1}$$

$$= \left[u_o^{-2}\sum_{j=1}^{n} X_j^{-1}(X_j-u_o)^2 \Big/ \sum_{j=1}^{n}\left(X_j^{-1}-\overline{X}^{-1}\right)\right]^{n/2} ,$$

$$= \left(1 + (n-1)^{-1}T\right)^{n/2} , \text{ where}$$

$$T = \left\{n(n-1)u_o^{-2}\,\overline{X}^{-1}(\overline{X}-u_o)^2\right\}\left[\sum_{j=1}^{n}\left(X_j^{-1}-\overline{X}^{-1}\right)\right]^{-1} .$$

The rejection region is therefore: Reject if

$$T > K_\alpha = \left(C_\alpha^{2/n} - 1\right)(n-1) .$$

This completes the proof.

4. Likelihood Ratio Test for the Equality of Several Inverse Gaussian Populations. Under the conditions of the fundamental lemma, we can determine the likelihood ratio test for:

H_o: $\mu_1=\mu_2=\dots=\mu_k$ against H_a: H_o is false, at the level α.

The rejection region of this test is given by: Reject if

$$S = \left[\sum_{i=1}^{k} n_i\left(\overline{X}_i^{-1}-\overline{X}^{-1}\right)\right]\left[\sum_{i=1}^{k}\sum_{j=1}^{n_i}\left(X_{ij}^{-1}-\overline{X}_i^{-1}\right)\right]^{-1} > M_\alpha,$$

where M_α, determined in section 5, is a constant, depending only on α.

Proof: Let $L(\mu_1,\dots,\mu_k,\lambda)$ be the joint density of the totality of random variables X_{ij}. Let $A' = (0,\infty)^{k+1}$; and $B' = \left\{(\mu_1,\dots,\mu_k,\lambda)\in A' : \mu_1=\mu_2=\dots=\mu_k\right\}$. In B', the fundamental lemma in applied to the totality of observations X_{ij} as a single random sample from $f(x;\mu,\lambda)$. Computation as in 3, yields the desired result.

5. Evaluation of K_α and M_α .

(a) $K_\alpha = F_{1,n-1,\alpha}$.

Proof: If $X_1, \ldots, X_n$ is a random sample from the inverse Gaussian distribution, with density $f(x,\mu,\lambda)$, Wasan, [4], page 2.26, showed that $\bar{X}$ has the inverse Gaussian density, $f(x;\mu,n\lambda)$. The author, [2], shows that for an inverse Gaussian random variable, with density $f(x;\mu,\lambda)$, that

$\lambda\mu^{-2}X^{-1}(X-u)^2 \sim \chi^2_1$. Tweedie [3], showed that

$\lambda \sum_{j=1}^{n}\left(X_j^{-1} - \bar{X}^{-1}\right) \sim \chi^2_{n-1}$, and is independent of $\bar{X}$.

Hence $T \sim F_{1,n-1}$, where T is defined in 3, gives the required value of K_α .

(b) $M_\alpha = (N-k)^{-1}(k-1) \quad F_{k-1,n-k,\alpha}$, where

$N = \sum_{i=1}^{k} n_i$.

This is Tweedie's distributional result, the basis of Analysis of Reciprocals, namely, that

$(N-k)(k-1)^{-1} S \sim F_{k-1,N-k}$ under H_o.

6. Summary and Conclusions. We have produced the likelihood test for the mean of an inverse Gaussian population, a test remarkably similar to the two-sided t-test. We have also shown that Tweedie's Analysis of Reciprocal procedure is in fact the likelihood ratio test.

REFERENCES

[1] E. Lehmann, "Testing Statistical Hypotheses," John Wiley and Sons, 1959.

[2] J. Shuster, "A Note on the Inverse Gaussian Distribution Function," Jour. Amer. Stat. Assoc., Vol. 63, pp. , 1968.

[3] M.C.K. Tweedie, "Statistical Properties of Inverse Gaussian Distribution I," Ann. Math. Stat., Vol. 28, pp. 362-77, 1957.

[4] M.T. Wasan, "Monograph on Inverse Gaussian Distribution," Queen's University Press, Kingston, Ont., 1966.

BARTLETT'S TEST FOR INVERSE GAUSSIAN POPULATIONS

by JONATHAN SHUSTER, McGILL & UNIV. of FLORIDA [1]

0. Introduction. Tweedie [3] and the author [2], have investigated a method for testing the equality of the means of several Inverse Gaussian populations. The major assumption of this procedure is that the secondary parameter λ_i, does not vary from sample to sample. This is the analogous assumption to that of constant variance in regression analysis. In this paper, the author will develop a modification of Bartlett's Test, to test the validity of the assumption on the secondary parameter λ_i.

1. Construction of the Test.

Definition: A random variable X, follows the Inverse Gaussian distribution, if it has density function

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

$$= 0 \qquad x < 0$$

Lemma: Let $X_{i_1}, \ldots, X_{i_{n_i}}$, $i = 1,\ldots,K$, be K independently drawn samples respectively, from $f(x;\mu_i,\lambda_i)$.

Let $A = (0,\infty)^{2K}$, and $B = \left\{(\mu_1,\ldots,\mu_K,\lambda_1,\ldots,\lambda_K) \in A : \lambda_1 = \ldots = \lambda_K\right\}$.

The Likelihood Ratio, M, for the test of hypothesis:

H_o: $(\mu_1,\ldots,\mu_K,\lambda_1,\ldots,\lambda_K) \in B$ versus

H_a: $(\mu_1,\ldots,\mu_K,\lambda_1,\ldots,\lambda_K) \in A \setminus B$, is

[1] Work supported by National Research Council of Canada

$$M = \hat{\lambda}^{\frac{1}{2}N} (\hat{\lambda}_1^{n_1} \hat{\lambda}_2^{n_2} \cdots \hat{\lambda}_K^{n_K})^{-\frac{1}{2}} \text{, where}$$

$$N = \sum_{i=1}^{K} n_i \,, \quad \hat{\lambda}_i = n_i \left[\sum_{j=1}^{n_i} (X_{ij}^{-1} - \bar{X}_i^{-1}) \right]^{-1} \,, \quad \hat{\lambda} = N \left[\sum_{i=1}^{K} (n_i/\hat{\lambda}_i) \right]^{-1},$$

and $\bar{X}_i$ is the i-th sample mean.

Proof: Under H_0, the joint density of the N random variables is readily seen to be maximized with respect to $(\mu_1,\ldots,\mu_K,\lambda_1,\ldots,\lambda_K)$ at the point $(\bar{X}_1,\ldots,\bar{X}_K,\hat{\lambda},\ldots,\hat{\lambda})$; while under H_a, the joint density of the N random variables is readily seen to be maximized with respect to $(\mu_1,\ldots,\mu_K,\lambda_1,\ldots,\lambda_K)$ at the point $(\bar{X}_1,\ldots,\bar{X}_K,\hat{\lambda}_1,\ldots,\hat{\lambda}_K)$.

The above value of M is obtained by taking the ratio of the joint density evaluated at $(\bar{X}_1,\ldots,\bar{X}_K,\hat{\lambda},\ldots,\hat{\lambda})$ to the joint density evaluated at $(\bar{X}_1,\ldots,\bar{X}_K,\hat{\lambda}_1,\ldots,\hat{\lambda}_K)$.

Definition: A two parameter family of probability densities is "Regular", provided that it satisfies:

(1) $$E\left[\frac{\partial \log g(\mu,\lambda)}{\partial \mu}\right] = E\left[\frac{\partial \log g(\mu,\lambda)}{\partial \lambda}\right] = 0 \quad \text{and}$$

(2) $$E\left[\frac{\partial^2 \log g(\mu,\lambda)}{\partial \mu \partial \lambda}\right] + E\left[\left(\frac{\partial \log g(\mu,\lambda)}{\partial \mu}\right)\left(\frac{\partial \log g(\mu,\lambda)}{\partial \lambda}\right)\right] = 0 ,$$

where $g(\mu,\lambda)$ is a probability density of a random variable X, for (μ,λ) in an appropriate parameter space.

Lemma: The Inverse Gaussian densities, $f(x;\mu,\lambda)$, form a regular family.

Proof: The proof of this statement uses the following:

(a) Tweedie [3], showed $E(X) = \mu$, $Var(X) = \mu^3/\lambda$, and

(b) the author [1], showed $\left[\lambda(X-\mu)^2/\mu^2 X\right] \sim \chi^2_1$, hence $E\left[(X-\mu)^2/\mu^2 X\right] = 1/\lambda$, whenever X has the Inverse Gaussian distribution, defined above.

$$E\left[\frac{\partial \log f(X;\mu,\lambda)}{\partial \mu}\right] = E\left[\lambda X/\mu^3\right] - \lambda/\mu^2 = \lambda E\left[\frac{\partial^2 \log f(X;\mu,\lambda)}{\partial \mu \, \partial \lambda}\right] = 0$$

$$E\left[\frac{\partial \log f(X;\mu,\lambda)}{\partial \lambda}\right] = 1/2\lambda - E\left[\tfrac{1}{2}(X-\mu)^2/\mu^2 X\right] = 0$$

$$E\left[\left(\frac{\partial \log f(X;\mu,\lambda)}{\partial \mu}\right)\left(\frac{\partial \log f(X;\mu,\lambda)}{\partial \lambda}\right)\right] =$$

$$\left(-\lambda/\mu^2\right)\left\{E\left(\frac{\partial \log f(X;\mu,\lambda)}{\lambda}\right)\right\} + \left(\lambda/\mu^3\right) E\left[X/2\lambda - \tfrac{1}{2}(X-\mu)^2/\mu^2\right] = 0 .$$

This completes the proof.

THEOREM: (Bartlett's Test for the equality of the λ_i)

The critical region for the test in the first lemma, is a symptotically: $\left\{M : -2 \log M > \chi^2_{K-1,p}\right\}$, where p is the significance level of the test, and M is as given in the first lemma.

Proof: This resultS immediately from : (a) the second lemma, (b) 13.8.1, page 419 of Wilks [4], on the limiting distribution of the likelihood ratio of regular densities, (c) H_o's equivlence to H_o^* : $\lambda_i - \lambda_1 = 0$, $i = 2, \ldots, K$, (K-1 constraints), and (d) $-2 \log M$ is a monotonic decreasing function of M.

2. Conclusion. Analysis of Reciprocals, depends on the λ_i remaining constant, from sample to sample, of Inverse Gaussian random variables. The above modification of Bartlett's Test gives an approximate method of checking out this homogeneity.

REFERENCES

[1] J. Shuster, "A Note on the Inverse Gaussian Distribution Function," Jour. Amer. Stat. Assoc., Vol. 63, pp. , 1968.

[2] J. Shuster, "Likelihood Ratio Tests for Inverse Gaussian Populations," Submitted to Ann. Math. Stat., 1969.

[3] M.C.K. Tweedie, "Statistical Properties of Inverse Gaussian Distribution I," Ann. Math. Stat., Vol. 28, pp. 362-77, 1957.

[4] S.S. Wilks, "Mathematical Statistics," John Wiley and Sons, 1962.