A DISCUSSION OF SOME SUBJECTS RELATED TO THE INVERSE GAUSSIAN DISTRIBUTION

By

HODA MAHMOUD ALY MARZOUK

Bachelor of Science
Cairo University
Cairo, Egypt
1968

Master of Science
Cairo University
Cairo, Egypt
1974

Master of Arts
State University of New York at Buffalo
Buffalo, New York
1979

Submitted to the Faculty of the
Graduate College of the
Oklahoma State University
in partial fulfillment of
the requirements for
the Degree of
MASTER OF SCIENCE
July, 1982

A DISCUSSION OF SOME SUBJECTS

RELATED TO THE INVERSE

GAUSSIAN DISTRIBUTION

Thesis Approved:

R. Leroy Folks
Thesis Adviser

Lyle Broemeling

N. Mukhopadhyay

Norman N. Durham
Dean of the Graduate College

ACKNOWLEDGMENTS

The author wishes to express appreciation to Dr. J.L. Folks who served as her major thesis adviser. Appreciation is also expressed to Dr. L.D. Broemeling and Dr. N. Mukhopadhyay for serving on the author's advisory committee. Many thanks to all faculty members of the Department of Statistics at Oklahoma State University. Also, many thanks to Centre Associates who served as her typist.

TABLE OF CONTENTS

Chapter | Page

I. INTRODUCTION 1

Statement of the Problem 1

II. INVERSE GAUSSIAN PROBABILITY PLOTS 3

Basics of Probability Plot Paper 3
Construction of Inverse Gaussian Probability Plots 5
Wasan and Roy Tables 6
Inverse Gaussian Probability Plots . 6
Examples 8
Example (1) 8
Example (2) 9

III. THE POWER FUNCTION OF SOME TEST STATISTICS RELATED TO THE INVERSE GAUSSIAN DISTRIBUTION 15

Introduction 15
The Size of the Test Statistic U 17
The Power of the Test Statistic U 19
The Distribution of U 22

IV. SUMMARY 25

REFERENCES 28

LIST OF TABLES

Table | Page

I. The Values of the Ordered Values of the Sample Points $y_{(i)}$ (the Sampling Quantiles) with the Corresponding Theoretical Quantile Values $\tilde{y}_{(i)}$ for Example (1) 11

II. The Values of the Ordered Values of the Sample Points $y_{(i)}$ (the Sampling Quantiles) with the Corresponding Theoretical Quantile Values $\tilde{y}_{(i)}$ for Example (2) 13

LIST OF FIGURES

Figure | Page

1. The Values of the Ordered Values of the Sample Points $y_{(i)}$ (the Sampling Quantiles) with the Corresponding Theoretical Quantile Values $\tilde{y}_{(i)}$ for Example (1) 12

2. The Values of the Ordered Values of the Sample Points $y_{(i)}$ (the Sampling Quantiles) with the Corresponding Theoretical Quantile Values $\tilde{y}_{(i)}$ for Example (2) 14

CHAPTER I

INTRODUCTION

A random variable X is said to have an Inverse Gaussian Distribution if its probability density function is defined by,

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

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

where $\mu > 0$ and $\lambda > 0$

with

$E[X] = \mu$

$Var[x] = \mu^3/\lambda$

Many authors have studied the properties of the distribution and its applications; for example, Tweedie (1957a, 1957b), Khatri (1962), Chhikara and Folks (1974, 1975, 1976, 1977, 1978), Davis (1977) and many more. The object of this study is to investigate other subjects like the probability plots and the power function which is not investigated in the Inverse Gaussian Distribution theory.

Statement of the Problem

In this study I will discuss two different problems concerning the Inverse Gaussian Distribution. They are the probability plot procedure for the Inverse Gaussian

Distribution and the power function of some test statistics which appear in statistical inferences with the Inverse Gaussian Distribution.

First, Chapter II is concerned with the probability plot for the Inverse Gaussian Distribution. The subject is introduced for the first time into the study of the Inverse Gaussian Distribution theory and a method is suggested which describes how to make a probability plot for the Inverse Gaussian Distribution with some illustrated examples.

Secondly, Chapter III concerns itself with the derivation of the power function of some test statistics which arise in the study of linear statistical models which relate to the Inverse Gaussian Distribution. These test statistics arise in the study by Davis (1977).

CHAPTER II

INVERSE GAUSSIAN PROBABILITY PLOTS

Basics of Probability Plot Paper

Let $y_1, y_2, \ldots, y_n$ be a sample from a population with c.d.f. F(y) known theoretically. If $y_{(1)}, y_{(2)}, \ldots, y_{(n)}$ is the order statistic of $y_1, y_2, \ldots y_n$, then $F_n(y)$ is called the sample (or empirical) c.d.f. defined as

$$F_n(y) = \begin{cases} 0 & \text{if } y < y_{(1)} \\ i/n & \text{if } y_{(i)} \leq y < y_{(i+1)} \\ 1 & \text{if } y \geq y_{(n)} \end{cases}$$
$$i = 1, 2, \ldots n.$$

$F_n(y)$ is a good estimate for the theoretical function F(y), say the Inverse Gaussian Distribution, in the sense that:

1. $E[F_n(y)] = F(y)$

i.e., it is unbiased estimate for F(y) (Rohatgi, 1975, p. 300).

2. $F_n(y) \xrightarrow{P} F(y)$ as $n \longrightarrow \infty$

In words, $F_n(y)$ converges in probability to F(y) and this means that $F_n(y)$ is a consistent estimate for F(y).

> Hence, if the sample is large enough, the sample order statistics when plotted against the 'corresponding' quantities of the theoretical distribution will tend to yield a set of points falling on a straight line with slope 1 through the origin (Wilk and Gnanadesikan, 1962, p.1).

Now, if we subject our variables to a linear transformation, i.e.

$$X = \frac{Y - \mu}{\beta}$$

then the basic linearity property will remain but the slope will no longer be 1 but β and the intercept will be μ / β. This new random variable X is called the "reduced variable".

When the distribution involved in this linear transformation is a parameter-free distribution function, i.e., does not involve unknown parameters, then special "probability paper" for that distribution may be prepared. On such paper one axis is scaled according to the quantiles of the "reduced" distribution and marked with the percentage values. Two outstanding examples are the normal distribution

$$F(y) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{y} e^{-t^2/2} \, dt$$

and the extreme value distribution of Type I

$$F(y) = \exp\,[-e^{-y}]$$

where μ is the mode and $\beta = \sqrt{\frac{6}{\pi}}\,\sigma$. There is probability paper for a few more distributions like the Binomial, lognormal, Weibull, Beta and the Uniform (Hahn and Shapiro, 1967).

Wilk and Gnaradesikan (1962) present a plotting procedure for the gamma distribution (Pearson's Type III) which includes the chi-squared and the exponential distributions as special cases. The distribution may be reduced or standardized, using a linear transformation EXCEPT for the shape parameter. Accordingly, no simple general probability paper can be prepared for this distribution. They constructed tables which give the percentage points for the reduced gamma distribution. Also, they suggested and employed an easy procedure for carrying the preparation of gamma probability plots by hand.

Here in this discussion I will adopt Wilk and Gnanadesikan's (1962) procedure for constructing Inverse Gaussian probability plots with some modification necessary for the new distribution, I will use the Wasan and Roy (1969) tables for constructing the Inverse Gaussian probability plots.

Construction of Inverse Gaussian Probability Plots

Suppose the random variable Y has Inverse Gaussian probability density function with parameters μ and λ. Now to construct general probability plot paper we have to be able to find a transformation from the random variable y to another random variable with probability density function independent of any of the parameters, say μ and λ, this kind of transformation is not possible for the Inverse Gaussian variables. So, for this reason there is no simple general probability plot paper which can be prepared for the Inverse Gaussian Distribution.

Wasan and Roy Tables

Now, the Inverse Gaussian probability density function has two parameteres μ and λ. We can reduce these two parameters to one parameter as was done by Wasan and Roy (1969). They described some of the properties of the Inverse Gaussian Distribution with $\mu = t$ and $\lambda = t2$. They showed that if Y is I. G. distributed with p.d.f. $f_Y(y;\mu,\lambda)$ then the random variable $X = \lambda Y/\mu^2$ is Inverse Gaussian distributed with p.d.f. $f_X(x;t,t2)$ where $t = \lambda/\mu$;

$$f_X(x;t,t2) = \begin{cases} \dfrac{t}{\sqrt{2\pi x^3}} \exp\left\{\dfrac{-(x-t)2}{2x}\right\} & , x>0 \\ & , t>0 \\ 0 & , \text{otherwise} \end{cases}$$

They present tables for the values of the percentage points of this Inverse Gaussian Distribution with parameter t for values of t ranging from 0.1 to 4000. The cases for t>4000 and t<.1 are also considered.

Inverse Gaussian Probability Plots

By using the Wasan and Roy (1969) Tables and the same steps which were suggested by Wilk and Gnanadesikan (1962) we can easily construct Inverse Gaussian probability plots. Let F(y) denote the c.d.f. of the

Inverse Gaussian variable Y, i.e.

$$F_Y(y) = \int_0^y f(z;t,t^2)dz \qquad , y>0.$$

Suppose that $y_{(1)} \leqslant y_{(2)} \leqslant \dots \leqslant y_{(n)}$ is an ordered random sample of n observations. Let $b_1, b_2 \dots b_n$ denote the appropriately chosen fractions of the Inverse Gaussian Distribution corresponding to these ordered statistics.

If $\tilde{y}_{(i)}$, i=1,... n satisfied $F_Y(\tilde{y}_{(i)}; t, t^2) = bi$ and if $y_{(1)}, y_{(2)}, \dots y_{(n)}$ is in fact an ordered random sample from the Inverse Gaussian distribution with parameter t, then the points $(y_{(i)}, \tilde{y}_{(i)})$ i=1, ... n will tend to fall along a straight line with slope 1 through the origin.

How should we choose the values of b_i for given sample size n? There are a lot of suggestions in this matter, but the usual choice which I will employ here is to take b_i as

$$b_i = (i - 1/2)/n.$$

So, the main thing to be done is to find $\tilde{y}_{(i)}$ for given b_i and t; this could be done by using the Wasan and Roy (1969) table. Then plot $(y_{(i)}, \tilde{y}_{(i)})$ and judge how near your plot to the straight line with slope 1 passed through the origin.

To use the Wasan and Roy (1969) table for preparing probability plot by hand, we follow a procedure similar to that suggested by Wilk and Gnanadesikan (1962, p.11) as follows:

1. Order the sample values to obtain

   $y_{(1)} \leq y_{(2)} \cdots \leq y_{(n)}$.

2. Compute $b_i = (i - 1/2)/n$, $i = 1, \ldots n$.
3. Determine $t = \lambda/\mu$ to be used. It it is known then we use the table directly, or if it is unknown, we can estimate it from the data,

$$\mu = \frac{1}{n}\sum_{i=1}^{n} y_i = \bar{y}$$

$$\frac{1}{\tilde{\lambda}} = \frac{1}{n}\sum_{i=1}^{n}\left(\frac{1}{y_i} - \frac{1}{\bar{y}}\right)$$

   then find the appropriate row in the table to be used to get $\tilde{y}_{(i)}$.

4. Interpolate b_i in the list of percentages and obtain the corresponding quantile $\tilde{y}_{(i)}$. A rough graph of the tabulated percentages versus quantiles may be used for interpolation purposes to yield quick and sufficiently accurate answers.
5. Plot $y_{(i)}$ versus $\tilde{y}_{(i)}$ on ordinary linear by linear graph paper.

Examples

<u>Example (1)</u>

The following sets give runoff amount at Jug Bridge, Maryland:

0.17, 0.19, 0.23, 0.33, 0.39, 0.40, 0.45, 0.52, 0.56,

0.64, 0.66, 0.70, 0.76, 0.77, 0.78, 0.95, 0.97, 1.02, 1.12, 1.24, 1.59, 1.74, 2.92. (Folks and Chhikara, 1978, Example 4)

We want to test if these data represent a random sample from a population which has the Inverse Gaussian Distribution. We will use the procedure which was developed here:

1. Estimate μ and λ from the data.
   $\tilde{\mu}$ = .80 and $\tilde{\lambda}$ = 1.44
2. Construct Table I from the data and Wason and Roy (1969) tables by using $t = \tilde{\lambda} / \tilde{\mu} = 1.8$.
3. Plot $(y_{(i)}, \tilde{y}_{(i)})$ as done in Figure 1.
4. Conclusion: From the plots in Figure 1 we see that the fit is very satisfactory, i.e., the data are very well described by the Inverse Gaussian and this is the same conclusion as in Folks and Chhikara (1978).

Example 2

The following set gives data of precipitation (inches) from Jug Bridge, Maryland:

1.01, 1.11, 1.13, 1.15, 1.16, 1.17, 1.17, 1.20, 1.52, 1.54, 1.54, 1.57, 1.64, 1.73, 1.79, 1.79, 2.09, 2.09, 2.57, 2.75, 2.93, 3.29, 2.54, 3.57, 5.11, 5.62. (Folks and Chhikara, 1978, Example 3)

Steps:

1. Estimate μ and λ from the data
   $\tilde{\mu}$ = 2.16 and $\tilde{\lambda}$ = 8.08.
2. Construct the Table II from the data and Wason and Ray (1969)

table by using $t = \tilde{\lambda} / \tilde{\mu} = 3.7$.

3. Plot $(y_{(i)}, \tilde{y}_{(i)})$ as done in Figure 2.
4. Conclusion: From the plots in Figure 2 we see that the fit is not satisfactory, i.e., the data are not described by the Inverse Gaussian and this is the same conclusion as in Folks and Chhikara (1978).

TABLE I

THE VALUES OF THE ORDERED VALUES OF THE SAMPLE POINT $y(i)$ (THE SAMPLING QUANTILES) WITH THE CORRESPONDING THEORETICAL QUANTILE VALUE $\tilde{y}_{(i)}$

order i	$y_{(i)}$	$b_i = (i - 1/2)/n$	$\tilde{y}_{(i)}$
1	0.17	0.02	0.4093
2	0.19	0.06	0.4893
3	0.23	0.10	0.6079
4	0.33	0.14	0.7000
5	0.39	0.18	0.7500
6	0.39	0.22	0.8500
7	0.40	0.26	0.8955
8	0.45	0.30	1.0000
9	0.52	0.34	1.0750
10	0.56	0.38	1.1500
11	0.59	0.42	1.2500
12	0.64	0.46	1.3000
13	0.66	0.50	1.4140
14	0.70	0.54	1.5500
15	0.76	0.58	1.6500
16	0.77	0.62	1.8000
17	0.78	0.66	1.9500
18	0.95	0.70	2.0750
19	0.97	0.74	2.2596
20	1.02	0.78	2.4000
21	1.12	0.82	2.5356
22	1.24	0.86	3.0500
23	1.59	0.90	3.4100
24	1.74	0.94	4.2992
25	1.92	0.98	6.2556

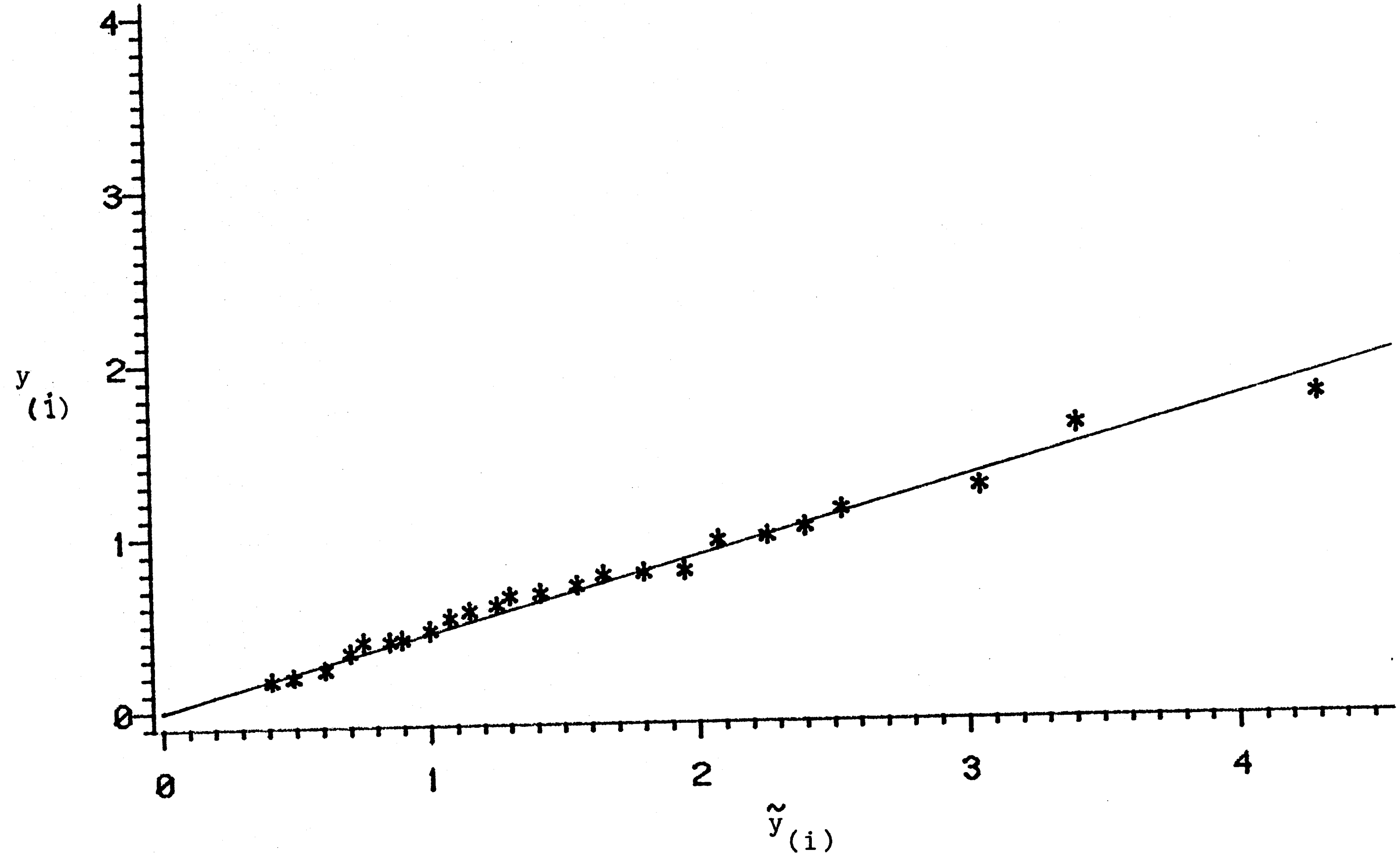

Figure 1. The Values of the Ordered Values of the Sample Points $y_{(i)}$ (The Sampling Quantiles) with the Corresponding Theoretical Quantiles Values $\tilde{y}_{(i)}$

TABLE II

THE VALUES OF THE ORDERED VALUES OF THE SAMPLE POINT $y(i)$ (THE SAMPLING QUANTILES) WITH THE CORRESPONDING THEORETICAL QUANTILE VALUE $\tilde{y}_{(i)}$

order i	$y_{(i)}$	$b_i = (i - 1/2)/n$	$\tilde{y}_{(i)}$
1	1.01	0.02	0.2784
2	1.11	0.06	0.4719
3	1.13	0.10	1.7409
4	1.15	0.14	1.9000
5	1.16	0.18	2.0500
6	1.17	0.22	2.2000
7	1.17	0.26	2.3289
8	1.20	0.30	2.5000
9	1.52	0.34	2.6500
10	1.54	0.38	2.8000
11	1.54	0.42	2.9500
12	1.57	0.46	3.1000
13	1.64	0.50	3.2518
14	1.73	0.54	3.4500
15	1.79	0.58	3.6500
16	2.09	0.62	3.9000
17	2.09	0.66	4.1000
18	2.57	0.70	4.3000
19	2.75	0.74	4.5505
20	2.93	0.78	4.8000
21	3.19	0.82	4.9406
22	3.54	0.86	5.7000
23	3.57	0.90	6.1054
24	5.11	0.94	7.2079
25	5.62	0.98	9.4004

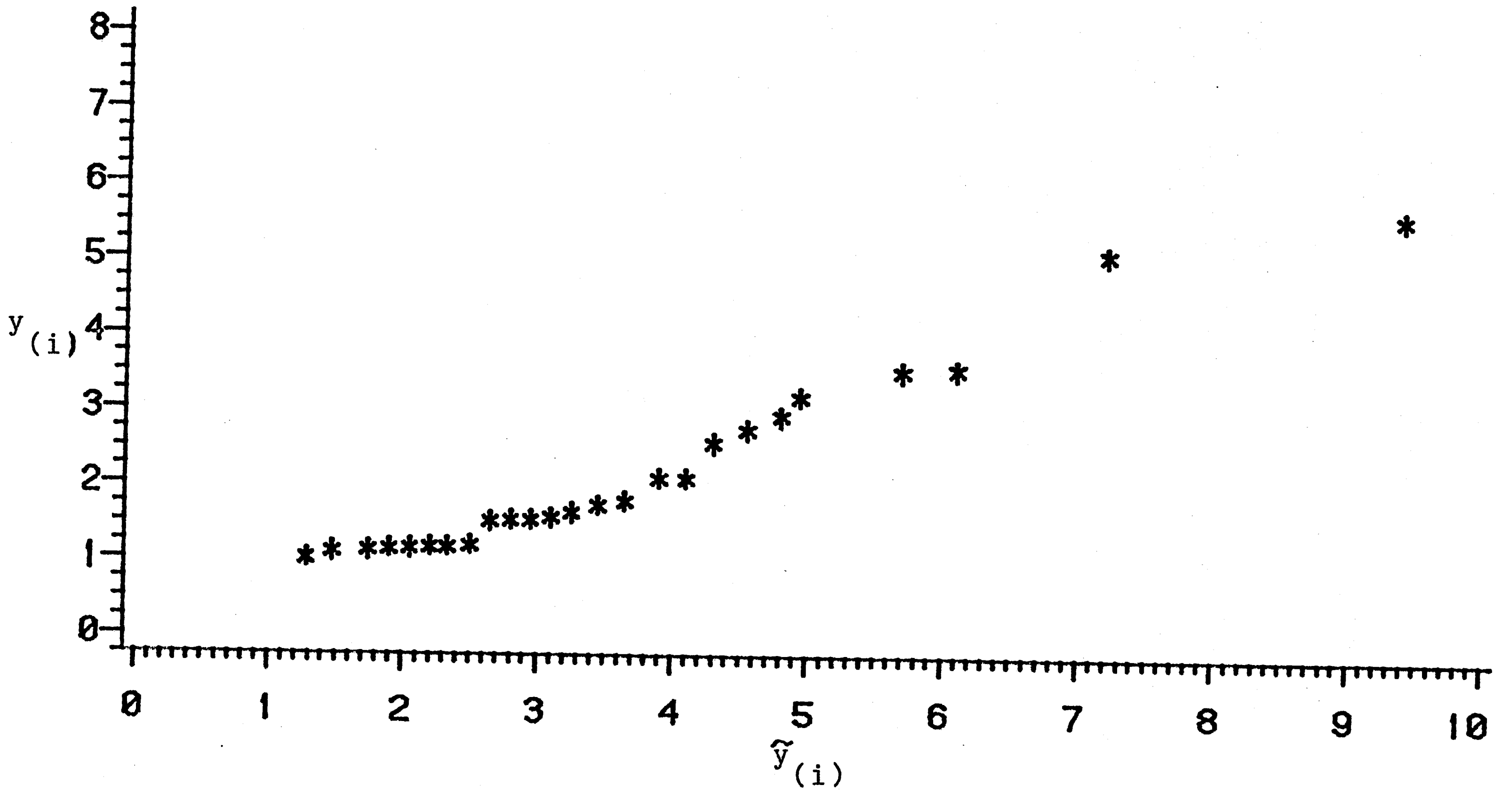

Figure 2. The Values of the Ordered Values of the Sample Points $y_{(i)}$ (The Sampling Quantiles) with the Corresponding Theoretical Quantiles Values $\tilde{y}_{(i)}$

CHAPTER III

THE POWER FUNCTION OF SOME TEST STATISTICS RELATED TO THE INVERSE GAUSSIAN DISTRIBUTION

Introduction

In this chapter I will derive the power function of some test statistics which arise in the the study of linear statistical models. More specifically these tests statistics arise in the study of Davis (1977)

Anne Davis (1977) studied four linear models in detail. I will derive the power function of some test statistics used in one of the models which is Model A (Davis, 1977), but the same idea and similar approach can be applied to the other models.

Now, let $y_1, y_2, \ldots y_n$ be a random sample from an Inverse Gaussian distribution with parameters βx_i and λ such that $y_i = \beta x_i + \epsilon_i$ where is an unknown scalar constant, x_i is a known quantity and ϵ_i is an error term with zero mean and independent of ϵ_j $(i = j)$. The maximum likelihood estimates of the parameters are

$$\text{(i)} \quad \tilde{\beta} = \frac{\sum_{i=1}^{n} \frac{y_i}{y_i^2}}{\sum_{i=1}^{n} \frac{1}{x_i}}$$

with Inverse Gaussian Distribution with parameters β and $\lambda \sum_{i=1}^{n} 1/x_i$

$$\text{(ii)} \quad \tilde{\lambda} = \frac{n}{\sum_{i=1}^{n} \left(\frac{1}{y_i} - \frac{1}{\tilde{\beta} x_i}\right)}$$

$$= \frac{n \sum_{i=1}^{n} \frac{y_i}{x_i^2}}{\sum_{i=1}^{n} \frac{1}{y_i} \sum_{i=1}^{n} \frac{y_i}{x_i} - \left(\sum_{i=1}^{n} \frac{1}{x_i}\right)^2}$$

where $\frac{n\lambda}{\tilde{\lambda}}$ has χ^2- distributiiion with (n - 1) d.f. and is independent of $\tilde{\beta}$. To study the power function of test statistics, we have to study the theory of testing hypotheses in which these test statistics appear. In our situation, we want to test some hypothsis about the parameter β . Suppose that

$$H_0: \quad \beta = \beta^*$$
$$H_1: \quad \beta = \beta_0 \neq \beta^*$$

For this model if we use the likelihood ratio test to test the null hypothesis, then our test statistic will be

$$U=\left\{\frac{\lambda\sum_{i=1}^{n}\frac{1}{x_i}(\tilde{\beta}-\beta)^2}{\tilde{\beta}\ \beta^2}\right\}\Bigg/\left\{\frac{\lambda\left[\sum_{i=1}^{n}\frac{1}{y_i}\sum_{i=1}^{n}\frac{y_i}{x_i^2}-\left(\sum_{i=1}^{n}\frac{1}{x_i}\right)^2\right]}{\sum_{i=1}^{n}\frac{y_i}{x_i^2}}\right\}$$

Now we want to answer the following questions:

(1) What is the size of the test statistic U, say α ?

(2) What is the power of the test statistic U, say P?

Davis (1977) answered the first question and I will answer the second question. To answer these two questions we have to investigate the distribution of the test statistic U under the null hypothesis H_0 to get the size of the test statistic U and also to investigate the distribution of the test statistic U under alternative hypothesis H_1 to get the power of the test statistic U.

The Size of the Test Statistic U

Under the null hypothses H_0, the statistic U will be,

$$U=\left\{\frac{\lambda\sum_{i=1}^{n}\frac{1}{x_i}(\tilde{\beta}-\beta^*)^2}{\tilde{\beta}\ \beta^{*2}}\right\}\Bigg/\left\{\frac{\lambda\left[\sum_{i=1}^{n}\frac{1}{y_i}\sum_{i=1}^{n}\frac{y_i}{x_i^2}-\left(\sum_{i=1}^{n}\frac{1}{x_i}\right)^2\right]}{\sum_{i=1}^{n}\frac{y_i}{x_i^2}}\right\}$$

but

$$\frac{\lambda \sum_{i=1}^{n} \frac{1}{x_i}(\tilde{\beta} - \beta^*)^2}{\tilde{\beta}\ \beta^{*2}}$$

has χ^2 - distribution with 1 d.f. (Davis, 1977; p. 40).
Also

$$\frac{n\lambda}{\tilde{\lambda}} = \frac{\lambda\left[\sum_{i=1}^{n} \frac{1}{y_i} \sum_{i=1}^{n} \frac{y_i}{x_i^2} - \left(\sum_{i=1}^{n} \frac{1}{x_i}\right)^2\right]}{\sum_{i=1}^{n} \frac{y_i}{x_i}}$$

has χ^2 - distribution with (n - 1) d.f. and is independent of $\tilde{\beta}$. Therefore

$$(n-1)U = \frac{(n-1) \sum_{i=1}^{n} \frac{1}{x_i}\left(\tilde{\beta} - \beta^*\right)^2 \sum_{i=1}^{n} \frac{y_i}{x_i^2}}{\tilde{\beta}\ \beta^{*2}\left[\sum_{i=1}^{n} \frac{1}{y_i} \sum_{i=1}^{n} \frac{y_i}{x_i^2} - \left(\sum_{i=1}^{n} \frac{1}{x_i}\right)^2\right]}$$

has F - distribution with 1 and (n - 1) d.f. respectively. Hence, the size of the test statistic U, say α , is

$$\alpha = Pr\left\{(n-1)U > F_{1,n-1;1-\alpha} \mid H_0\right\}$$

$$= \int_{F_{1,n-1;1-\alpha}}^{\infty} F \, dF$$

The Power of the Test Statistic U

The power of the test statistic U, say P, is

$$P = \int_{F_{1,n-1;1-\alpha}}^{\infty} (n-1)\, U \quad dU$$

this probability evaluated in the case where the alternative hypothsis H_1 is true, i.e.,

$$P = Pr\left\{(n-1)U > F_{1,n-1;1-\alpha} \mid H_1\right\}$$

So, we have to investigate the distribution of

$$(n-1)U = \left\{\frac{\lambda \sum_{i=1}^{n} \frac{1}{x_i}\left(\tilde{\beta} - \beta^*\right)^2}{\tilde{\beta}\ \beta^{*2}}\right\} \Bigg/ \left\{\frac{\lambda\left[\sum_{i=1}^{n} \frac{1}{y_i} \sum_{i=1}^{n} \frac{y_i}{x_i^2} - \left(\sum_{i=1}^{n} \frac{1}{x_i}\right)^2\right]}{\sum_{i=1}^{n} \frac{y_i}{x_i^2}}\right\}$$

when

$$\beta = \beta_0 \neq \beta^*$$

Now I will examine the distribution of each of the numerator and the denominator of U separately, then I will get the distribution of U and this is done in 3 steps as follows:

1. $\tilde{\beta}$ has the Inverse Gaussian Distribution with parameters β_0 and $\lambda \sum_{1=i}^{n} \frac{1}{x_i}$.

   So,

$$\frac{\lambda \sum_{i=1}^{n} \frac{1}{x_i} \left(\tilde{\beta} - \beta^*\right)^2}{\tilde{\beta}\ \beta^{*2}}$$

   does not has χ^2 - distribution, but the distribution of this quantity can be found easily by using the work done by Chhikara and Folks (1975, theorem 2).

2.

$$\frac{\lambda\left[\sum_{i=1}^{n} \frac{1}{y_i} \sum_{i=1}^{n} \frac{y_i}{x_i^2} - \left(\sum_{i=1}^{n} \frac{1}{x_i}\right)^2\right]}{\sum_{i=1}^{n} \frac{y_i}{x_i^2}}$$

   has χ^2 - distribution with (n - 1) d.f.. Note that the distribution of this quantity is the same under both H_0 and H_1.

3. The distribution of (n - 1) U can be obtained by two ways: (a) Using theorem (2) of Chhikara and Folks (1975) which gives the distribution of

$$t = \sqrt{\frac{\lambda \sum_{i=1}^{n} \frac{1}{x_i} (\tilde{\beta} - \beta^*)^2}{\tilde{\beta} \ \beta^{*2}}}$$

and then by using transformation of variables we get the distribution of t^2. Then the distribution of $((n-1)t^2)/\chi^2 = (n-1)U$ can be obtained, where in that theorem, the parameter λ is replaced by $\lambda \sum_{i=1}^{n} 1/x_i$, the parameter μ_0 is replaced by β^* and n is replaced by (n-1). (b) Using theorem (4) of Chhikara and Folks (1975) directly, theorem (4) gives us the distribution of (n-1)U = W and then we can get the distribution of (n-1)U by driving the distribution of W^2, where in theorem (4) the parameter λ is replaced by $\lambda \sum_{i=1}^{n} \frac{1}{x_i}$, the parameter μ_0 is replaced by β^* and n is replaced by (n-1). The derivation is carried out in the next section.

Now, suppose we denote the probability density function of S = (n-1)U by f(s), then the power of the test statistic U is,

$$P = Pr\left\{(n-1)U > F_{1,\ n-1;1-\alpha} \mid H_1\right\}$$
$$= \int_{F_{1,n-1;1-\alpha}}^{\infty} f(s)ds$$

The Distribution of U

First, by using the same notation as in theorem (4) (Chhikara and Folks, 1975), we can find the distribution of W^2 as follows: Let X have Inverse Gaussian Distribution with parameters μ and λ be independent of the random variable $Z = V/\lambda$ where V has χ^2 - distribution with n d.f.. Also, let

$$W = \frac{\sqrt{n}(X - \mu_0)}{\mu_0 \sqrt{XZ}} \qquad , \mu_0 > 0$$

then the probability density function of $Y = W^2$ is as follows:

$$f_Y(y) = \frac{\exp\{-\lambda(\mu - \mu_0)^2/2\mu^2\mu_0\}}{\sqrt{2\pi}\,\Gamma(n/2)}\left[2 + \frac{(\mu^2+\mu_0^2)}{\mu^2}\frac{y}{n}\right]^{-\frac{n+1}{2}}$$

$$\left\{\sum_{k=0}^{\infty}\left(\frac{\mu^2-\mu_0^2}{2\mu^3\mu_0}\right)^{2k}\frac{1}{\sqrt{y}\,(2k)!}\left(\frac{y/n}{\left[2+\frac{\mu^2+\mu_0^2}{\mu^2}\frac{y}{n}\right]^2}\right)^{k} G_{n+4k-1,2k}(y) + \right.$$

$$\left.\sum_{k=0}^{\infty}\left(\frac{\mu^2-\mu_0}{2\mu^2\mu_0}\right)^{k}\frac{1}{(2k+1)!}\left(\frac{-\sqrt{y}/\sqrt{n}}{\left[2+\frac{\mu^2+\mu_0^2}{\mu^2}\frac{y}{n}\right]}\right)^{2k+1}\frac{\mu_0}{\sqrt{n}}G_{n+4k,2k}(y)\right\}, \quad y>0$$

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

where

$$G_{ij}(y) = \int_0^{\infty}\left[2\lambda\mu_0\left(2+\frac{\mu^2+\mu_0^2}{\mu^2}\frac{y}{n}\right)+\frac{\mu^2 y}{n}u\right]^{j/2} u^{(i-j)/2}\, e^{-u/2}\, du$$

Proof:

Now, $y = W^2$ then $W = \pm\sqrt{y}$

therefore

$$f_Y(y) = \left[g_W(\sqrt{y}) + g_W(-\sqrt{y}) \right] \left\| \frac{dw}{dy} \right\|_{W \to y}$$

where g_W (.) denote to the p.d.f. of the random variable W.

$$f_Y(y) = \frac{C}{2\sqrt{y}} \left[2 + \frac{\mu^2 + \mu_0^2}{\mu^2} \frac{y}{n}\right]^{-\frac{n+1}{2}} \left\{ \sum_{k=0}^{\infty} \frac{1}{k!} \left(\frac{\sqrt{y}/\sqrt{n}}{\left[2 + \frac{\mu^2+\mu_0^2}{\mu^2}\frac{y}{n}\right]} \right)^k \left(\frac{\mu^2 - \mu_0^2}{2\mu^2\mu_0}\right)^k \left[G_{n+2k-1,k}(y) - \frac{\mu_0\sqrt{y}}{\sqrt{n}} G_{n+2k-1,k-1}(y) \right] \right.$$

$$+ \sum_{k=0}^{\infty} \frac{1}{k!} \left(\frac{-\sqrt{y}/\sqrt{n}}{\left[2 + \frac{\mu^2+\mu_0^2}{\mu^2}\frac{y}{n}\right]} \right)^k \left(\frac{\mu^2 - \mu_0}{2\mu^2\mu_0}\right)^k \left[G_{n+2k-1,k}(y) + \frac{\mu_0\sqrt{y}}{\sqrt{n}} G_{n+2k-1,k-1}(y) \right]$$

$$= \frac{C}{\sqrt{n}} \left[2 + \frac{\mu^2 + \mu_0^2}{\mu^2} \frac{y}{n}\right]^{-\frac{n+1}{2}} \left\{ \sum_{k=0}^{\infty} \frac{C_1}{2k!} \left(\frac{y/n}{\left[2 + \frac{\mu^2+\mu_0^2}{\mu^2}\frac{y}{n}\right]} \right)^k G_{n+4k-1,2k}(y) + \right.$$

$$\left. \sum_{k=0}^{\infty} \frac{C_2}{2k+1!} \left(\frac{-\sqrt{y}/\sqrt{n}}{\left[2 + \frac{\mu^2+\mu_0^2}{\mu^2}\frac{y}{n}\right]} \right)^{2k+1} \frac{\mu_0\sqrt{y}}{\sqrt{n}} G_{n+4k,2k}(y) \right\} \quad , y > 0$$

where:

$$C = \frac{\exp\left[-\lambda(\mu - \mu_0)^2 / 2\mu^2\mu_0\right]}{\sqrt{2\pi}\ \Gamma(n/2)} \quad ,$$

$$C_1 = \left(\frac{\mu^2 - \mu_0^2}{2\mu^2\mu_0}\right)^k \quad \text{and} \quad C_2 = \left(\frac{\mu^2 - \mu_0^2}{2\mu^2\mu_0}\right)^{2k}$$

Second by using our notation, i.e. replacing λ by $\lambda \sum_{i=1}^{n} 1/x_i$, the parameter μ_0 by β^* and n by (n - 1), we can get the probability density function of S = (n - 1)U, i.e., f(s) as follows:

$$f_S(s) = \frac{\exp\left[-\lambda \sum x_i (\beta_0 - \beta^*)^2 / 2\beta_0^2 \beta^*\right]}{\sqrt{(n-1)\pi}\;\Gamma(n-1/2)} \left[2 + \frac{\beta_0^2 + \beta^{*2}}{\beta_0^2} \frac{s}{n-1}\right]^{-\frac{n}{2}}$$

$$\left\{ \sum_{k=0}^{\infty} \left(\frac{\beta_0^2 - \beta^{*2}}{2\beta_0^2 \beta^*}\right)^{2k} \frac{1}{\sqrt{s}\; 2k!} \left(\frac{s/n-1}{\left[2 + \frac{\beta_0^2 + \beta^{*2}}{\beta_0^2} \frac{s}{n-1}\right]^2}\right)^{k} G_{n+4k-2,2k}(s) \; + \right.$$

$$\left. \sum_{k=0}^{\infty} \left(\frac{\beta_0^2 - \beta^{*}}{2\beta_0^2 \beta^*}\right)^{k} \frac{1}{(2k+1)!} \left(\frac{-\sqrt{s}/\sqrt{n-1}}{\left[2 + \frac{\beta_0^2 + \beta^{*2}}{\beta_0^2} \frac{s}{n-1}\right]}\right)^{2k+1} \frac{\beta^*}{\sqrt{n-1}} G_{n+4k-1,2k}(s) \right\}$$

$$, s > 0$$

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

CHAPTER IV

SUMMARY

This study discusses two different subjects related to the Inverse Gaussian Distribution. They are the probability plot procedure for the Inverse Gaussian Distribution and the power function of some test statistics which appears in statistical inference with the Inverse Gaussian Distribution.

Probability Plot Procedure

In Chapter II a new procedure is carried out describing how to make a probability plot. In other words, if we have a sample of size n, taken from some population, and we want to test if this data represents a sample taken from a population with probability density function of the Inverse Gaussian Distribution, we just apply the procedure which is suggested here in this study. The procedure is as follows:

1. Order the sample values to obtain

   $y_{(1)} \leqslant y_{(2)} \leqslant \cdots\cdot \leqslant y_{(n)}$
2. Compute $b_i = (i-0.5)/n$, $i = 1, 2, \ldots, n$
3. Determine $t = \lambda / \mu$ to be used. If it is known then we use the table directly, or if it is

unknown, we can estimate it from the data

$$\tilde{\mu} = 1/n \sum_{i=1}^{n} y_i = y$$

$$1/\tilde{\lambda} = 1/n \sum_{i=1}^{n} (1/y_i - 1/y)$$

then find the appropriate row in the table to be used to get $\tilde{y}_{(i)}$.

4. Interpolate b_i in the list of percentages and obtain the corresponding quantities $\tilde{y}_{(i)}$. A rough graph of the tabulated percentages versus quantiles may be used for interpolation purposes to yield quick and sufficiently accurate answers.

5. Plot $y_{(i)}$ versus $\tilde{y}_{(i)}$ on ordinary linear by linear graph paper. If your data comes from a population with probability density function as the Inverse Gaussian Distribution, then the Graph will represent a straight line with slope 1 and passing through the origin and as we go away from that straight line with slope 1 passing through the origin we go away from the conclusion that the data represent a sample from a population with Inverse Gaussian Distribution. Also, Chapter II includes two illustrated examples.

The Power Function

Chapter III concerns the derivation of the power function of some test statiatics which arise in the study of linear statistical models which are related to the Inverse

Gaussian Distribution. These test statistics arise in the study by Davis (1977, p. 38).

Davis (1977) studied four linear models in detail. In Chapter III the derivation of the power function of some test statistics used in one of her models (model A) is given.

REFERENCES

1. Davis, A.S. (1977). "Linear Statistical Inference as Related to the Inverse Gaussian Distribution." (Ph.D. Thesis, Department of Statistics, Oklahoma State University.)

2. Folks, J.L. and Chhikara, R.S. (1974). "Estimation of the Inverse Gaussian Distribution Function." Journal of the American Statistical Association, 69, 250-254.

3. Folks, J.L. and Chhikara, R.S. (1975). "Statistical Distribution Related to the Inverse Gaussian." Communications in Statistics, 4, 1081-1091.

4. Folks, J.L. and Chhikara, R.S. (1978). "The Inverse Gaussian Distribution and Its Statistical Application - A Review." Journal of the Royal Statistical Society B, 40, 263-239

5. Hahn, G.J. and Shapiro, S.S. Statistical Models in Engineering. New York: John Wiley and Sons, Inc., 1967.

6. Khatri, C.G. (1962). "A Characterization of the Inverse Gaussian Distribution." Annals of Mathematical Statistics, 33, 800-803.

7. Tweedie, M.C.K. (1957a). "Statistical Properties of Inverse Gaussian Distributions I." Annal of Mathematical Statistics, 28, 696-377.

8. Tweedie, M.C.K. (1957b). "Statistical Properties of Inverse Gaussian Distributions II." Annal of Mathematical Statistics, 28, 696-705.

9. Wasan, M.T. and Roy, L.K. (1969). "Tables of Inverse Gaussian Percentage Points." Technometrics, 11, 591-604.

10. Wilk, M.B. and Gnanadezikan, R. and Huyett, M.J. (1962). "Probability Plots for the Gamma Distribution." Technometrics, 4, 1-20.

VITA

Hoda Mahmoud Aly Marzouk

Candidate for the Degree of

Master of Science

Thesis: A DISCUSSION OF SOME SUBJECTS RELATED TO THE INVERSE GAUSSIAN DISTRIBUTION

Major Field: Statistics

Biographical:

Personal Data: Born in Port-Said, Egypt, July 1, 1946, the daughter of Mahmoud Aly Marzouk.

Education: Graduated from Port-Said High School in May, 1964; received Bachelor of Science degree in statistics in May, 1968, Cairo University; received Master of Science degree in statistics from Cairo University, June, 1974; completed requirements for the Master of Arts in statistics at State University of New York at Buffalo; completed requirements for the Master of Science Degree in statistics at Oklahoma State University inJuly, 1982.

Professional Experience: Teacher in statistics at El-Azher University, Egypt, 1969-1974; Assistant lecturer in statistics at El-Azher University, Egypt, 1974-1976.