NOTICE

The quality of this microform is heavily dependent upon the quality of the original thesis submitted for microfilming. Every effort has been made to ensure the highest quality of reproduction possible.

If pages are missing, contact the university which granted the degree.

Some pages may have indistinct print especially if the original pages were typed with a poor typewriter ribbon or if the university sent us an inferior photocopy.

Reproduction in full or in part of this microform is governed by the Canadian Copyright Act, R.S.C. 1970, c. C-30, and subsequent amendments.

AVIS

La qualité de cette microforme dépend grandement de la qualité de la thèse soumise au microfilmage. Nous avons tout fait pour assurer une qualité supérieure de reproduction.

S'il manque des pages, veuillez communiquer avec l'université qui a conféré le grade.

La qualité d'impression de certaines pages peut laisser a désirer, surtout si les pages originales ont été dactylographiées à l'aide d'un ruban usé ou si l'université nous a fait parvenir une photocopie de qualité inférieure

La reproduction, même partielle, de cette microforme est soumise à la Loi canadienne sur le droit d'auteur, SRC 1970, c. C-30, et ses amendements subséquents.

Inverse Gaussian Distribution and Analysis of Factorial Experiments

Nagalakshmi N. Tarikere

A Thesis
in
The Department
of
Mathematics and Statistics

Presented in Partial Fulfilment of the Requirements
for the Degree of Master of Science at
Concordia University
Montreal, Quebec, Canada

June 1991
Nagalakshmi N. Tarikere

ISBN 0-315-68770-3

Abstract

Inverse Gaussian Distribution and Analysis of Factorial Experiments.

Nagalakshmi Tarikere

This thesis deals with the analysis of nested factorial experiments under Inverse Gaussian Distribution. A reciprocal linear model for the factor effects is motivated from the context of the underlying Weiner process. Maximum likelihood estimates from the likelihood equations are derived and important properties such as strong consistency and limiting normality are stated.

A least squares approach and a weighted least squares approach using the reciprocals of the sample cell means are studied and compared with the maximum likelihood method.

Likelihood ratio tests and analysis of reciprocals analogous to the usual normal analysis of variance have been carried out.

Application of these procedures is illustrated with the help of two numerical examples.

Acknowledgements

I am grateful to my advisor Prof. Yogendra Chaubey for taking me as his student, and guiding me in my thesis. He has always been encouraging, patient, and available for discussion.

I am thankful to Prof. Garrido and Prof. A. K. Srivastava for agreeing to be on my thesis committee.

It is a pleasure to express my gratitude to Prof. T. N. Srivastava for providing many books and encouragement.

I thank my friend Ms. Ravinder Kaur, and the faculty and staff in the Department of Mathematics and Statistics.

Contents

Chapter 1

Inverse Gaussian Distribution

1.1 Introduction

This thesis provides a short review of some basic properties of the inverse Gaussian (IG) distribution and formulates the analysis of factorial experiments using inverse Gaussian distribution as the model. As a specific case, we consider the nested classification using the IG model.

Chapter 1 outlines the development of the inverse Gaussian distribution as a probability model and describes the important properties of this distribution. Also a few applications are considered. Many references are provided in the Bibliography which deal exhaustively with this distribution.

Chapter 2 provides a brief description of a *linear model* and the various methods of estimation of the parameters in such a model.

Chapter 3 describes the work done with regard to regression and factorial experiments using the IG distribution as a model by various researchers. And, in particular

the papers by Tweedie [43], Shuster and Miura [41], and Fries and Bhattacharyya [5] are discussed.

Chapter 4 follows the paper by Fries and Bhattacharyya [5] for nested classification. They have considered the crossed classification. Here as a specific case, the nested classification is used, but their paper has been followed to a large extent.

Chapter 5 deals with two numerical examples to illustrate the method of estimation and test of hypotheses discussed in Chapter 4.

1.2 Background of the Inverse Gaussian Distribution

The earliest known reference to inverse Gaussian distribution was made by Schrodinger [39]. He gave the probability distribution function of the first passage time in Brownian motion. No one else is known to have referred to it until the paper by Tweedie [43]. Tweedie proposed the name inverse Gaussian distribution to the first passage time distribution, since he found an inverse relationship between the cumulant generating function of this distribution and that of the normal distribution. The distribution was next given by Wald [44] who derived it as a limiting form for the distribution of the sample size in a sequential probability ratio test. Because of this derivation, the distribution is also called *Wald's distribution*, particularly in the Russian literature.

The other names for this distribution are *Gaussian first passage time distribution* [34] and the first passage time distribution for Brownian motion with positive drift [47]. However from the viewpoint of statistics it might more appropriately be

called Tweedie 's distribution. It had remained virtually unnoticed until Tweedie [43] investigated its characteristics, established some important statistical properties and depicted certain analogies between its statistical analysis and that of the normal distribution.

A characterization of the inverse Gaussian distribution by Khatri [30] paralleled the usual characterisation of the normal distribution by the independence of the sample mean and variance further reflecting the analogy. Wasan and his associates [46, 45] investigated some analytical and characteristic properties of this class of distributions particularly for the limiting forms. More recently Chhikara [17] and Chhikara and Folks [12, 13, 14, 15, 16] have developed further its statistical theory illuminating the analogy that Tweedie set in his paper [43] and have provided statistical methods based upon the inverse Gaussian particularly in the field of reliability. Useful applications of the inverse Gaussian distributions have been demonstrated in the works of Bannerjee and Bhattacharya [3], Lancaster [31], Hasofer [27], Shepperd [40] and Whitmore [49, 52, 53]. Ahmad, Chaubey and Sinha [1] have analysed the important problem of combining different estimators in the context of IG populations. Reference can here be made to the recent book by Chhikara and Folks [10] for various properties and applications of the inverse Gaussian distribution.

1.3 The PDF of the IG and its Properties:

The probability density function (pdf) of a random variable X distributed as inverse Gaussian with parameters μ and λ, denoted by $X \sim IG(\mu, \lambda)$ in Tweedie's notation

is given by

$$f(x;\mu,\lambda) = (\lambda/2\pi x^3)^{1/2} exp(-\lambda(x-\mu)^2)/(2\mu^2 x) \quad x > 0. \tag{1.1}$$

The inverse Gaussian is a two parameter family of distributions. The probability density function can be represented in several different forms, each of which would be convenient or suggestive for some purpose. Another important form can be obtained by supposing $W(t)$ as a Weiner process in one dimension with positive drift ν and $W(0) = 0$. Then T, the time required for $W(t)$ the Weiner process to reach the value a, an arbitrary real value, in unit time, is a random variable with density function

$$f(t) = a/\sigma\sqrt{2\pi t^3}\ exp\{-(a-\nu t)^2/(2\sigma^2 t)\} \quad t > 0, \quad \nu > 0. \tag{1.2}$$

This is a reparametrization of the density function given in (1.1) obtained by letting $\mu = a/\nu$ and $\lambda = a^2/\sigma^2$ where a is specified. νt and $\sigma/\sqrt{t}$ are in the distance scale and μ and λ are in the time scale. The inverse relationship between μ and ν can be viewed in the simple context of speed which can be measured in miles/hour or hours/mile.

There are various other reparametrizations of the density function, but we will denote the $IG(\mu,\lambda)$ as follows through out the thesis: A random variable X is distributed as inverse Gaussian, i.e. $X \sim IG(\mu,\lambda)$ if Y has the pdf given by

$$f(x;\mu,\lambda) = (\lambda/2\pi x^3)^{1/2} exp(-\lambda(x-\mu)^2)/(2\mu^2 x) \quad x > 0. \tag{1.3}$$

The mean and variance of this distribution are μ and μ^3/λ respectively.

The density function which can be seen to be a member of the exponential family

is unimodal and positively skewed, with its mode given as

$$x_{mode} = \mu\{(1 + \frac{9}{4\phi^2})^{1/2} - \frac{3}{2\phi}\} \tag{1.4}$$

where $\phi = \lambda/\mu$.

The cumulant generating function (cgf) can be easily found. The logarithm of the Laplace transform $E(e^{tx})$ of the probability density of the variate is in a sense, the cumulant generating function. We denote the relevant function operator by L with the variate symbol as a subscript and the other variables in parenthesis. Thus from the density function 1.1 we get

$$\begin{aligned} L_X(t;\mu,\lambda) &= ln[e^{\lambda(2\alpha)^{1/2}} \int_0^\infty e^{-(\alpha+t/\lambda)\lambda\alpha - \lambda/2x} dx\{\lambda/2\pi x^3\}^{1/2}] & (1.5) \\ &= \lambda(2\alpha)^{1/2} - \lambda 2^{1/2}(\alpha + t/\lambda)^{1/2} + ln \int_0^\infty f_1(x;\alpha + t/\lambda, \lambda) dx & (1.6) \end{aligned}$$

where $f_1(x;\alpha + t/\lambda, \lambda)$ is the IG density function with with $\alpha + t/\lambda$ and λ as its parameters. If t is imaginary, or if real or complex, if its real part exceeds $-\alpha\lambda$, the integral in the above expression is unity. Hence

$$\begin{aligned} L_X(t;\mu,\lambda) &= \lambda\{(2\alpha)^{1/2} - 2^{1/2}(\alpha + t/\lambda)^{1/2}\} & (1.7) \\ &= \lambda/\mu\{1 - (1 + 2\mu^2 t/\lambda)^{1/2}\} & (1.8) \\ &= \phi\{1 - (1 + 2\mu t/\phi)^{1/2}\} & (1.9) \\ &= \phi\{1 - (1 + 2\lambda t/\phi^2)^{1/2}\} & (1.10) \end{aligned}$$

The cgf (1.11) is unique to the density function 1.1. The cumulants can be obtained from the power series expansion of $L_x(t;\mu,\lambda)$. The first two cumulants are

$$\begin{aligned} k_1 &= \mu = \lambda\phi^{-1} = \lambda\lambda/\mu = \lambda^2\mu^{-1} & (1.11) \\ k_2 &= \mu^2\lambda{-}1 = \lambda^2\phi^{-2} & (1.12) \end{aligned}$$

The characteristic function is

$$C_X(t) = exp\{\phi[1 - (1 - \frac{2i\mu^2 t}{\lambda})^{1/2}]\} \tag{1.13}$$

All the positive and negative moments exist. The moments can be found from either the characteristic function or directly by integration.

The mean and variance are given by

$$E(X) = \mu. \tag{1.14}$$

$$Var(X) = \mu^3/\lambda. \tag{1.15}$$

Thus μ and λ are only partially interpretable as location and scale parameters. There is a remarkably simple relationship between positive and negative moments given by

$$E[X^{-r}] = \frac{E[X^{r+1}]}{\mu^{2r+1}}. \tag{1.16}$$

It is possible to express the distribution function of X , $F(x)$ in terms of the standard Normal distribution Φ

$$F(x) = \Phi((\lambda/x)^{1/2}(-1 + x/\mu)) + e^{2\lambda/\mu}\Phi(-\lambda/x)^{1/2}(1 + x/\mu)) \tag{1.17}$$

This result was obtained independently by Zigangirov [57] and Shuster [42]. Because of the many analogies between the Normal and the inverse Gaussian, it is natural to tabulate some "standard" inverse Gaussian distribution. If $Z = \lambda X/\mu^2$ then $Z \sim IG(\phi, \phi^2)$ where $\phi = \lambda/\mu$. Moreover $E(Z) = \phi$ and $var[Z] = \phi$. This transformation gives a single parameter family of distribution and Wasan and Roy [45] have tabulated the percentage points of Z for various values of ϕ.

The inverse Gaussian distribution shares with the gamma, and log normal and other skewed distributions, asymptotic convergence to normality. As $\lambda/\mu \to \infty$, the distribution of X is asymtotically normal with mean μ and variance μ^3/λ as shown by Wald [44].

The inverse Gaussian distribution has striking similarities with and provocative departures from a normal distribution. By the help of the characteristic function, it can be shown that if $X \sim IG(\mu, \lambda)$, then $cX \sim IG(c\mu, c\lambda)$ for $c > 0$. However unlike in the normal case, a linear combination of inverse Gaussian variates would not be inverse Gaussian. But they have a certain additive property. Tweedie showed that if X_i, $i = 1, 2, ...n$ are independent inverseGaussian variables with parameters μ_i and λ_i such that $\lambda_i/\mu_i^2 = \epsilon$ for all i, then $\sum X_i \sim IG(\sum \mu_i, \epsilon(\sum \mu_i)^2)$. The constancy of λ_i/μ_i was shown also to be a necessary condition for the sum to be IG. (see Chhikara [11] and Shuster and Miura [41]). Thus in order for the linear combination $\sum c_i X_i$ of independent inverse Gaussian variables where $X_i \sim IG(\mu_i, \lambda_i)$ $i = 1, 2, ..., n)$ to be inverse Gaussian, $\lambda_i/c_i\mu_i^2$ must be a constant for $i = 1,2$, ... n.

For a random sample $X_1, X_2....., X_n$ where $X_i \sim IG(\mu, \lambda)$ Schrodinger [39] showed that the maximum likelihood estimates of μ and λ are given by $\hat{\mu} = \bar{X}$ and $\lambda = n/\sum_{i=1}^{n}(1/X_i - 1/\bar{X})$ where $\bar{X} = \sum_{i=1}^{n} X_i/n$.

Tweedie proved that $\bar{X} \sim IG(\mu, n\lambda)$ and that $\lambda \sum_{i=1}^{n}(1/X_i - 1/\bar{X}) \sim \chi^2_{n-1}$ distribution. However the chief and most important result he proved was that the maximum likelihood estimators of μ and λ have stochastically independent distributions and are of a nature that permits the construction of the analogue of the analysis of variance for nested classification.

Independence of $\bar{X}$ and $\hat{\lambda}$ is fairly easily established by finding the conditional moment generating function of $\lambda \sum_{i=1}^{n}(1/X_i - 1/\bar{X})$ given $\bar{X} = \bar{x}$. It is that of a χ^2 variable with $(n-1)$ degrees of freedom for all $\bar{x}$. Therefore independence follows.

The statistics $\bar{X}$ and $\sum_{i=1}^{n}(1/X_i - 1/\bar{X})$ jointly form a complete sufficient statistic for (μ, λ).

The uniformly minimum variance unbiased estimators (UMVUE) of μ and λ^{-1} are

$$\hat{\mu} = \bar{X} \text{ and} \tag{1.18}$$

$$\hat{\lambda}^{-1} = \sum_{i=1}^{n} \frac{(1/X_i - 1/\bar{X})}{n-1}. \tag{1.19}$$

The UMVUE of the variance μ^3/λ has also recently been obtained which is of a very complicated form.

Thus summarising the properties if $X \sim IG(\mu, \lambda)$ where $\mu = E(X)$ and $\mu^3/\lambda = var(X)$, and $X_1, X_2 ... X_n$ are iid $\sim X$ then $\bar{X}$ and U where $\bar{X} = \sum_{i=1}^{n} X_i/n$ and $U = \sum_{i=1}^{n}(1/X_i - 1/\bar{X})$, are jointly minimal sufficient for μ and λ and complete, and are also independent. Moreover, $\bar{X} \sim IG(\mu, \lambda)$ and $\lambda U \sim \chi^2_{n-1}$. These results are analogous to those for a normal distribution. However, unlike in the normal case, in general an arbitrary linear combination $\sum c_i X_i$ does not follow an IG distribution, and while μ and λ admit UMVUE's which are easily found, the UMVUE of the variance μ^3/λ is complicated. This is a point of departure from the normal model. For some other similarities and departures from the normal case, the reader may see Ahmad etal [1] Letac, Seshadri and Whitmore [32], Pandey and Malik [38], Bravo and MacGibbon [9], Pal and Sinha [37], and Hsieh, Korwar and Rukhin [28].

Tweedie [43] also considered the distribution of the reciprocal of the IG variable, i.e., the distribution of $1/X$ and studied several interesting properties of the variate.

The pdf of $Y = 1/X$ is [43]

$$f(y) = (\lambda/2\pi y)^{1/2} exp\{-\lambda y/2 + \lambda/\mu - \lambda/2\mu^2 y\} \quad 0 < y < \infty. \tag{1.20}$$

The mode of Y is [43]

$$y_{mode} = 1/2\lambda[-1 + (1 + 4\phi^2)^{1/2}] \tag{1.21}$$

$$= 1/\mu\{(1 + 1/4\phi^2)^{1/2} - 1/2\phi\}. \tag{1.22}$$

The expressions for the mean and variance of $1/X$ are [43]

$$E(1/X) = 1/\mu + 1/\lambda. \tag{1.23}$$

$$var(1/X) = 1/\lambda\mu + 2/\lambda^2. \tag{1.24}$$

1.4 Applications

Although the inverse Gaussian distribution was long known in the literature of stochastic processes, its potential in statistical applications is increasingly recognised in recent years, especially in the area of engineering reliability. In choosing a stochastic model, one looks for such features as a sound theoretical basis, flexibility of empirical fit, and amenability to statistical inferences and model diagnostics. In addition to having a variety of shapes of the pdf and several convenient sampling distributions, the inverse Gaussian distribution derives from a stochastic formulation that is sufficiently general to allow wide applicability.

The physical aspect of Brownian motion which gives rise to inverse Gaussian as the first passage time distribution suggests its potential applications in modeling such phenemena as survival period, service duration and usage times. Moreover, a wide variety of shapes generated by the pdf makes it a prospective competitor to the Gamma, Weibull and Log-Normal models.

Inverse Gaussian distribution may be considered appropriate in any situation involving skewed positive data. It has the advantage over some other skewed distributions, like Weibull, and log-normal, that exact small sample theory is tractable and in some cases it parallels that of the normal distribution.

The inverse Gaussian distribution has found application in describing duration and failure phenomenon in the natural and social sciences. Areas of application have included lengths of strikes (Lancaster [31]), hospital stays (Eaton and Whitmore [23]), employee service times by (Whitmore [49]), equipment lives (Chhikara and Folks [15]), noise intensity (Marcus [33]), and tracer dilution curves (Wise [56]) to name a few.

In short, whenever a Brownian motion or random walk drift towards a threshold is a plausible mechanism, the first passage time can arise. It has been suggested by some authors that the analytical tractability of the inverse Gaussian is sufficient reason to use it for curve fitting, but probably is of limited scientific interest. In several applications, the simple Brownian motion model is only a starting point. It describes only the base line behavior and when the drift and the barrier have physical meaning, they can be manipulated in the attempt to replicate more complicated behavior.

Since the hazard function for the inverse Gaussian increases and then decreases, it serves as a good model for accelerated life tests. Some applications in marketing are

treated by Bannerjee and Bhattacharya [3] and Nadas [35], and Chhikara and Folks [14] consider applications of IG in life testing. Padgett [36] gives confidence bounds for the reliability function. Bhattacharya and Fries [8] argue that inverse Gaussian is more appropriate than the Birbaum Saunders fatigue distribution and Chhikara and Guttman [18] give sequential and Bayseian prediction limits. Thus the inverse Gaussian is becoming a widely used tool in reliability theory.

Many physiological processes evolve according to a diffusion process (in fact that is how Tweedie started his investigations). For instance, Gerstein and Mandelbred [26] modelled a neuron's membrane potential as a single number which moves towards or away from a threshold according to whether it receives an excitatory or inhibitory signal, respectively. If the signals are of roughly equal magnitude and come at high rates, then a Brownian motion approximation is appropriate. They showed that this model provides a good fit for the spontaneous activity of several neurons in the auditory cortex of a cat. They could interpret the parameters: the drift is a measure of the difference between the excitatory and inhibitory rates, and the barrier is a firing threshold. They also showed that, by introducing a time varying drift for the Brownian motion, they could replicate the observed behavior of one of the neurons subjected to periodic stimuli of various frequencies.

Other more sophisticated diffusion models are described in Fienberg [24]. Another important physiological application is the modelling of tracer dilution curves in cardiology (Wise [56]). A review of the various types of random walk models for physical systems and an extensive bibliography is given in Weiss [48].

One of the earliest applications of Brownian motion is due to Bachelier [2], who

modelled stock prices. Since then, investigators have applied it to a wide variety of phenemona (among others) in finance, marketing and labor. Recent examples in which the inverse Gaussian arises are the work of Bannerjee and Bhattacharya [7], who model interpurchase times as a renewal process with inverse Gaussian components, and Whitmore [49], who models labor turnover with this model. Whitmore has found that some of his data is very long-tailed, so he also uses a drift away from the barrier: thus, if T is the hitting time, then $P(T = \infty)$ is positive: this idea is similar to that of terminating renewal processes and gives rise to censored observations, which Whitmore [54] has studied in a regression context.

The maximum likelihood estimates of μ and λ the mean and the other parameter respectively of the inverse Gaussian distribution have stochastically independent distributions and are of a nature which permits the analysis of variance for nested and crossed classifications. This property was first investigated by Tweedie and this application to the analysis of variance is concentrated upon in this thesis.

1.5 IG Distribution as a Model for Failure Times

In this thesis we consider the inverse Gaussian distribution as a model for failure times. We are interested in the inverse Gaussian distribution as a model for failure times because of its several appealing features:

(i) : It accomodates a variety of shapes analogous to the Weibull and Lognormal distribution.

(ii) : It has the structure of an exponential family and many convenient properties

for the associated sampling distributions.

(iii) : Its derivation from a plausible stochastic formulation of the failure process often provides a physical support to its empirical fit.

The derivation of the inverse Gaussian distribution can be cast in the context of failure growth or accumulation of damage in a material over a period of time. Specifically if fatigue grows according to a Weiner process with drift η and diffusion constant δ^2 and if the material fails as soon as its accumulated fatigue exceeds a critical level $\omega > 0$, then the time for failure follows $IG(\theta, \sigma)$ distribution with

$$\theta = \omega/\eta, \quad \text{and} \tag{1.25}$$

$$\sigma = \frac{\delta^2}{\omega}. \tag{1.26}$$

The derivation is provided below. Some of the notations used in this derivation are

$g(t; \theta, \lambda)$: pdf of the inverse Gaussian distribution.

$G(t)$: cdf of the inverse Gaussian distribution.

ω : critical level of failure.

$B(t)$: accumulated fatigue in time $[0, t]$.

$B^*(t) : sup_{s \in [0,t]} B(s)$.

η: diffusion parameter of the Weiner process.

δ^2: diffusion constant of Weiner process

$N[a, b]$: standard normal distribution with mean a and variance b.

$\Phi(*)$: cdf of standard Gaussian distribution.

We regard the accumulated fatigue in time $[0, t]$ to be governed by a Weiner

process $[B(t) : 0 < t < \infty]$ with drift $\eta > 0$ and a diffusion constant δ^2 . Thus for any $0 < s < t$ the distribution of $B(t) - B(s)$ is $N\{\eta[t-s], \delta^2[t-s]\}$

Considering that the material fails as soon as $B(t)$ exceeds ω, the time to failure T is a random variable. $T = inf\{t : B(t) > \omega\}$. Now define

$$B^*(t) = supB(s)_{s\in\{0,t\}} \tag{1.27}$$

The genesis of the distribution of T is founded upon the basic event relations

$$\begin{aligned} \{T > t\} &= \{B^*(t) \leq \omega\} \\ (B(t) \leq \omega) &= (B^*(t) \leq \omega) \bigcup (B(t) \leq \omega \text{ and } B^*(t) > \omega) \end{aligned} \tag{1.28}$$

which leads to the distribution function of T.

$$G(t) = \Phi(\eta\delta^{-1}t^{1/2} - \omega\delta^{-1}t^{-1/2}) +$$

$$exp(2\eta\omega\delta^{-2}) \quad t > 0.$$

Taking the derivative of $G(t)$ as can be seen in Cox and Miller [19] we obtain the density function of the inverse Gaussian distribution. Using the parameters

$$\theta = \omega/\eta \tag{1.29}$$

$$\lambda = \omega^2/\delta^2 \tag{1.30}$$

we get

$$g(t;\theta,\lambda) = (2\pi\lambda t^3)^{-1/2} exp[-\lambda\frac{(t\theta^{-1}-1)^2}{2t}] \tag{1.31}$$

$$t > 0, \quad \theta > 0 \quad \lambda > 0.$$

In this thesis we consider a reciprocal linear model for the mean of the inverse Gaussian distribution. We consider the following situation. Suppose that n similar

objects are subject to stress levels $x_1, x_2, ..., x_n$, until they break. Each object has a characteristic break threshold, which we take to be the same. The different stress levels however imply that the accumulation of fatigue proceeds at different objects at different levels. If the accumulation follows a Brownian motion, we may then model the drift function of the covariate x: and when this drift is linear in x, the time to break for the i^{th} object follows an inverse Gaussian with mean θ_i and the reciprocal of the mean linear in the factor effects. We assume a constant σ because the breaking threshold is the same for all objects. This appropriate generalisation of the 'reciprocal linear model' has several advantages. As stated earlier it arises out of a possible physical model. The hitting time distribution acquires a wide variety of shapes and the relevant sampling theory is in large part tractable.

The inverse Gaussian is thus a competitor to other well-known parametric families, such as the Gamma or the Weibull. However it does not suppliant either; for example, it does not adequately fit data from the Exponential distribution, and the theory for the inverse Gaussian with censoring is not well developed. Further more the strong resemblance between the inverse Gaussian and the Gaussian theories become somewhat weaker in the regression context. For example, in the Gaussian linear model, maximum likelihood estimates of parameters coincide with the least squares estimates. However in the inverse Gaussian reciprocal model, the two are not the same, and whereas (finite sample) sampling properties of the least squares estimates are easier to obtain than the maximum likelihood estimates, the former are less efficient for large samples.

Chapter 2

Some Preliminaries of Estimation in Linear Models

2.1 Introduction

One of the aims of science is to find, to describe and to predict relationships among the events in the world in which we live. One way of accomplishing it would be by finding a formula or equation that relates quantities in the real world. For example we may be interested in the relationship between temperature and pressure in a chemical process, or in the relationship between the number of apples in trees in an orchard and the amount of fertilizers the trees receive. Or we could be interested in the relationship of supply, demand and price of certain commodities, or in how a certain vaccine affects a disease or in the relationship of rainfall, temperature and humidity, or in the yields of various varities of wheat.

To study these types of problems, we investigate different kinds of relationships

through formulation of models. For example it could be a linear model or a quadratic model etc. In particular in this thesis we are interested in the linear model where we assume a linear relationship between different variables.

2.2 Linear Models

A general linear (statistical) model that is to be used to determine the value of a variable y from a knowledge of another variable x is usually written in the form

$$y = \mu(x) + e \tag{2.1}$$

where y and e are random variables, $\mu(x)$ is a function of x defined in some domain D and x is a non-random variable. The function $\mu(x)$ is defined to be the deterministic portion of the model. e is the random or stochastic portion. Also y is referred to as the dependent or response variable and x is referred to as the independent or predictor variable. If y represents say the blood pressure of an individual, and x represents the age of an individual, then $\mu(x)$ is the predicted value of the individual's blood pressure, except for a random error e from a knowledge of his age x. The random variable e is not observable, but something about the distribution of e is often stated as part of the model. In general the "functional form" of $\mu(x)$ is known, but it contains unknown parameters. The word "linear" indicates that $\mu(x)$ is a linear function of the independent variables.. The first step in our analysis would be to estimate the unknown parameters.

2.2.1 General Linear Model

: Consider the n equations

$$Y_i = \sum_{j=1}^{p} x_{ij}\beta_j + \epsilon_i. \tag{2.2}$$

$E(\epsilon_i) = 0; \quad i = 1, 2, 3, \ \ldots \ , n;$

where

1. the Y_i are observable independent variables;

2. the x_{ij} are observable random variables from the domain D;

3. the β_j are unknown parameters defined in a parameter space Ω_β ;

4. the ϵ_i are unobservable random variables such that $Cov[\epsilon_i, \epsilon_{i'}] = 0 \quad i \neq i'$

These specifications define a general linear model.

In matrix notation the equation would be

$$Y = X\beta + \epsilon. \tag{2.3}$$

where $Y = [Y_1, Y_2, Y_3, \cdots, Y_n]'$,

$\beta = [\beta_1, \beta_2, \beta_3, \cdots, \beta_p]'$,

$\epsilon = [\epsilon_1, \epsilon_2, \epsilon_3, \cdots, \epsilon_n']$,

$$X = \begin{bmatrix} x_{11} & x_{12} & \cdots & x_{1p} \\ x_{21} & x_{22} & \cdots & x_{2p} \\ \vdots & \vdots & \vdots & \vdots \\ x_{n1} & x_{n2} & \cdots & x_{np} \end{bmatrix}$$

Y is thus $(n \times 1)$ observable random vector, X is $(n \times p)$ matrix of fixed observable numbers, $\underline{\beta}$ is $(p \times 1)$ vector of unknown parameters with $E(\epsilon) = 0$ and $Cov = (\sigma_{i,j}) = \sum, Cov$ denoting the Covariance matrix. In many cases further assumptions about

the distributional properties of ϵ or Y or further assumptions about the structure of $\sum$ will be stated as part of the model.

2.2.2 An Example of a General Linear Model

Suppose that the distance s that a mass-particle moves from a point of reference in time t is given by the formula

$$s = \beta_1 + \beta_2 t. \tag{2.4}$$

In this problem t is a nonrandom variable and we assume it can be observed without error. However, suppose that s cannot be measured accurately, but instead of observing s, one can observe values of Y, where $Y = s + \epsilon$. That is, the observable variable y is equal to the true distance s plus a measurement error ϵ. If we substitute, we get

$$Y = \beta_1 + \beta_2 t + \epsilon, \quad E[\epsilon] = 0 \quad var[\epsilon] = \sigma^2. \tag{2.5}$$

Suppose that this model is assumed to hold only for t in the interval $0 \leq t \leq 100$. This is a functional relationship model with measurement error in the dependent variable. To obtain sample values, the investigator would preselect n values of time, say $t_1, t_2, ..., t_n$ and observe the corresponding distances $y_1, y_2, ..., y_n$. Thus the sample model is

$$Y_i = \beta_1 + \beta_2 t_i + \epsilon_i$$

$$E[\epsilon_i] = 0 \qquad i = 1, 2, ..., n \tag{2.6}$$

This, in vector notation, fits the equation of the general linear model (2.3) with $p = 2$.

Our analysis would consist of first estimating the unknown parameter $\underline{\beta}$,for which three methods are available in general, namely

1. the maximum likelihood method,

2. the method of least squares,

3. the method of weighted least squares.

These three methods are briefly described below.

2.3 Method of Maximum Likelihood

The method of maximum likelihood is one of the oldest and most important in estimation theory. It was used by Gauss in developing the theory of Least squares, which subsequently overshadowed maximum likelihood method until it was reintroduced by Fisher in 1912. This method, intimately connected with sufficiency and applicable to almost all problems of estimation has great practical appeal.

2.3.1 The Method

Let $x_1, x_2, \ldots\ldots, x_n$ be observed random variables from a population with probability distribution indexed by θ. The likelihood function L is defined, in this case where the x_i come from a continuous sample space with probability density $f(x;\theta)$, by

$$L(\theta; x_1, \ldots\ldots, x_n) = \prod_{i=1}^{n} f(x_i;\theta) \tag{2.7}$$

and for discrete sample spaces by

$$L(\theta; x_1, \ldots\ldots, x_n) = \prod_{i=1}^{n} P(x_i;\theta) \tag{2.8}$$

where $P(x_i,\theta)$ is the probability mass function. We should note that the likelihood function is a function of θ given $x_1, \ldots, x_n$. The method of maximum likelihood

consists of finding a value of θ which maximises the likelihood function. Assuming that the likelihood is a positive differentiable function of θ and the maximum does not occur on the boundary of the set of all admissible θ, we attempt to find a solution to the likelihood equation

$\frac{\partial L}{\partial \theta} = 0$ or to the equivalent , but usually simpler, equation $\frac{\partial logL}{\partial \theta} = 0$. since logarithm is a monotone function. We usually ignore any solution which is independent of the observations; i.e, any constant solutions.

2.3.2 Some Optimum Properties of the Estimates

. Let $f(x, \theta)$ be the probability density function of the population from which a sample of size n, $(x_1, ..., x_n)$, is observed. We shall denote the density $f(x_i, \theta)$ by f_i and the product by L. The following assumptions are made.

1. The derivates of $\frac{\partial logL}{\partial \theta}$ and $\frac{\partial^2 logL}{\partial \theta^2}$ exist and are continuous for every θ in a range R, including the true value, and for almost all x. For every θ in R, $\frac{\partial logL}{\partial \theta} < F_1(x)$ *and* $\frac{\partial^2 logL}{\partial \theta^2} < F_2(x)$ where $F_1(x)$ and $F_2(x)$ are integrable functions over $(-\infty, +\infty)$.

2.The derivative $\frac{\partial^3 logL}{\partial \theta^3} < M(x)$ and $E[M(x)] < K$ (a positive quantity).

3. For every θ in R,

$$\int_{-\infty}^{+\infty} -\frac{\partial^2 logL}{\partial \theta^2} L dx = I(\theta) \tag{2.9}$$

is finite and non-zero.

4. The range of integration is independent of θ.

Under these conditions the following theorems hold good.

Theorem 2.1. With probability approaching unity as $n \to \infty$, the likelihood equation $\frac{\partial logL}{\partial \theta} = 0$ has a solution which converges in probability to the true value θ_0. (Dugue, 1937) [22].

Theorem 2.2. Any consistent solution of the likelihood equation provides a maximum of the likelihood with probability tending to unity as the sample size tends to infinity. (Huzurbazar, 1948) [29].

Theorem 2.3. A consistent solution of the likelihood equation is asymptotically normally distributed about the true value θ. (Cramer, 1946) [20].

2.4 Method of Least Squares.

The method of Least Squares goes back to Gauss, who first developed this technique of estimation. Generally it is used for the estimation of parameters in a linear model.

Let $Y_1, Y_2, \ldots, Y_n$ be independent random variables such that $E(Y_i) = \lambda_1 + \lambda_2 X_i$, where X_i is a known constant and λ_1 and λ_2 are unknown constants. Let us consider the quadratic function

$$W(Y_1, Y_2, \ldots, Y_n, \lambda_1, \lambda_2) = \sum_{i=1}^{n} (Y_i - \lambda_1 - \lambda_2 X_i)^2 \qquad (2.10)$$

Let $\hat{\lambda_1}$ and $\hat{\lambda_2}$ be the estimates of λ_1 and λ_2 respectively, which minimize the above equation. Then $\hat{\lambda_1}$ and $\hat{\lambda_2}$ are called the least squares estimates of λ_1 and λ_2. They can be obtained by solving the so called normal equations.

$$\frac{\partial W(Y_1, Y_2, \ldots, Y_n, \lambda_1, \lambda_2)}{\partial \lambda_1} = -2 \sum_{i=1}^{n} (Y_i - \lambda_1 - \lambda_2 X_i) = 0$$

$$\text{or} \quad \sum_{i=1}^{n} Y_i = n\lambda_1 + \lambda_2 \sum X_i.$$

$$\text{and } \frac{\partial W(Y_1, Y_2, \ldots, Y_n, \lambda_1, \lambda_2)}{\partial \lambda_2} = -2\sum_{i=1}^{n} X_i(Y_i - \lambda_1 - \lambda_2 X_i) = 0$$

$$or \quad \sum_{i=1}^{n} Y_i X_i = \lambda_1 \sum X_i + \lambda_2 \sum X_i^2. \tag{2.11}$$

. By solving these equations one can obtain

$$\begin{aligned} \hat{\lambda_2} &= \frac{\sum_{i=1} n(Y_i - \bar{Y})(X_i - \bar{X})}{\sum_{i=1}^{n}(X_i - \bar{X})^2}. \\ \hat{\lambda_1} &= \bar{Y} - \hat{\lambda_2}\bar{X}. \end{aligned}$$

Consider the general linear model we defined earlier which in matrix notation is represented as

$$Y = X\beta + \epsilon \tag{2.12}$$

So now, the parameter estimates that minimize the sum of the squared deviations

$$S(\beta) = (Y - X\beta)'(Y - X\beta) \tag{2.13}$$

are called the Least Squares estimates.

We have

$$\begin{aligned} \epsilon'\epsilon &= (Y - X\beta)'(Y - X\beta). \\ &= YY' - 2\beta'X'R + \beta'X'X\beta. \end{aligned}$$

Using the fact that

$$(\beta'X'Y)' = Y'X\beta. \tag{2.14}$$

and differentiating $S(\beta)$ wrt β, we have

$$\frac{\partial(S(\beta)}{\partial\beta} = 0 \Rightarrow -2X'Y + 2X'X\beta = 0$$

$$\Rightarrow X'X\beta = X'Y$$

$$\Rightarrow \beta = (X'X)^{-1}X'Y. \qquad (2.15)$$

Here we assume that X is of full rank and hence $(X'X)^{-1}$ exists.

It can very easily be shown that any other value of β will lead to a larger value of $S(\beta)$.

A few lemmas and definitions with respect to the least squares principle of estimation are stated below:

1. An estimator is called a least-squares estimate if and only if it satisfies the set of normal equations.

Definition: Any linear function $\beta = b'\gamma$ of the unknown parameters $[\gamma_1, \gamma_2, ..., \gamma_k]$ with known constant coefficients $[b_1, b_2, ..., b_k]$ will be called a *parametric function.* If there exists an unbiased linear estimate of β, then β will be called an *estimable parametric function.*

2. A parametric function $\beta = b'\gamma$ is estimable if and only if $b' = \alpha'X$, that is if and only if b' is a linear combination of the columns of X.

3. Let U be the space spanned by the columns of X. Then, if β is estimable, there exists a unique unbiased linear estimator of β, say $\alpha^{*'}Y$, where α^* is in U. Furthermore, α^* is the projection of α on U, where $\alpha'Y$ is any other unbiased estimate of β.

Gauss-Markoff theorem: Let $E(Y) = X\gamma$ and $\sum = \sigma^2 I$. Then there exists a unique unbiased linear estimator $\hat{\beta}$ for every estimable function $\beta = b'\gamma$. Furthermore, $\hat{\beta}$ has minimum variance in the class of all unbiased estimates. It may be obtained

from β by replacing the $[\gamma_1, \gamma_2, ..., \gamma_k]$ by any least squares estimators $[\hat{\gamma}_1, \hat{\gamma}_2, ..., \hat{\gamma}_k]$.

In estimating β in the general linear model, our two principal assumptions, in the method of Least squares, about the error terms e_i were that

1. the errors are uncorrelated, and

2. the errors have equal variance, i.e. we assume the covariance matrix $\sum$ of ϵ to be $\sigma^2 I$.

Sometimes it so happens that all the observations do not have equal variance. In this case, the least squares method is not appropriate. We can resort to the weighted least squares method of estimation in this situation. This method is briefly described below.

2.5 Method of Weighted Least Squares

. The method of weighted least squares (which involves transforming the dependent variable by adding weights in order to stabilise the variance) is appropriate when we have the model 2.3 with

$$Cov(\epsilon) = \sigma^2[\Omega] \tag{2.16}$$

where Ω is the matrix

$$\Omega = W^{-1} = \begin{bmatrix} w_{11}^{-1} & 0 & 0 & 0 & \cdots & 0 \\ 0 & w_{12}^{-1} & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & w_{IJ}^{-1} \end{bmatrix} \tag{2.17}$$

We have assumed the covariance matrix as a diagonal matrix in which the diagonal elements (variances) are not equal.

The next step is to transform this model into a model in which the variances are equal. This can be achieved by multiplying the model

$$Y = X\beta + \epsilon. \tag{2.18}$$

by

$$W^{1/2} = \begin{bmatrix} \sqrt{w_{11}} & 0 & 0 & 0 & \cdots & 0 \\ 0 & \sqrt{w_{12}} & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & \sqrt{w_{IJ}} \end{bmatrix} \tag{2.19}$$

which leads to

$$W^{1/2}Y = W^{1/2}X\beta + \epsilon^* \tag{2.20}$$

where $\epsilon* = W^{1/2}\epsilon$. The error in the transformed model satisfies the usual assumptions

$$\begin{aligned} E(\epsilon^*) &= 0. \\ Cov(\epsilon^*) &= Cov(W^{1/2}\epsilon) \\ &= \sigma^2 W^{1/2}W^{-1}W^{1/2} \\ &= \sigma^2(I) \end{aligned}$$

The least squares estimates in the transformed model can now be obtained by minimizing the weighted sum of squares.

$$\begin{aligned} S(\beta) &= (Y - X\beta)'W(Y - X\beta). \\ &= Y'WY - 2\beta'X'WY + \beta'(X'WX)\beta. \end{aligned} \tag{2.21}$$

Differentiating wrt β and equating to zero, we get

$$(X'WX)\beta = X'WY. \tag{2.22}$$

The solution of the equation 2.22 is called the weighted least squares estimate of β denoted as $\hat{\beta}_w$

$$\hat{\beta}_w = (X'WX)^{-1}X'WY. \tag{2.23}$$

Since these estimates minimise the weighted sum of squares,

$$\begin{aligned} S(\beta) &= (Y - X\beta)'W(Y - X\beta). \\ &= \sum_i (w_i(Y_i - \sum_j x_{ij}\beta_j)^2 \end{aligned}$$

they are called the weighted least squares estimates. In this sum of squares, the observations are weighted in proportion to the reciprocal of their variances.

Further, since the weighted least squares is nothing but least squares on the transformed observations, analytical properties of $\hat{\beta}_w$ can be studied similar to those of $\hat{\beta}$. If the elements of W can be estimated consistently resulting in $\hat{W}$, the asymptotic distribution of $\hat{\beta}_w$ and $\hat{\beta}_{\hat{w}}$ are same.

Chapter 3

Testing Equality of Factor Effects in the context of IG Model

3.1 Introduction

A most frequent problem that occurs in statistical work is that of testing whether two samples differ significantly with respect to a characteristic. For example the population mean or the population variance. We may be interested in testing if the two population distributions are identical.

The reason this type of problem occurs so frequently is that experimentalists often design an experiment to compare a new technique or process with a standard one. For example an educator may believe that he has discovered a better way of teaching foreign languages than that being used in the institution; or a chemist may have discovered a new plastic that he believes may be superior to the one being manufactured in his plant, etc.

However situations often occur in which there are several methods or products rather than just two that are competing against each other. For example a manufacturer of cake mixes may vary the amount of a cake ingredient to obtain different mixes that he would like to compare for quality, or a business firm may use differnt calculating machines that it would like to compare with respect to performance.

It is very inefficient to compare several samples by comparing them two at a time, especially so in the case of agricultural experiments that are concerned with testing different types and amounts of fertilizers and different seed varieties. One of the methods that has been designed to solve problems of this type for continuous variables is known as *analysis of variance*. As the name indicates the method consists of decomposing the variability of the sample data into various components since it gives a measure of the effective variability.

The analysis of variance is essentially a method of separating the total variance of a response variable into its various components, corresponding to the sources of variations which can be identified. The data must clearly contain information on any given source of variation before its contribution can be estimated, and as a rule the components are best estimated from the experiments which have been designed for this purpose. The procedure to be used in the application of the analysis will depend on the number and the nature of the independent causes of variation which can be identified. It is possible to classify the data with respect to each such source of variation, and a complete classification is a necessary first step in the analysis.

There are two basic types of classifications:

1. Crossed classification, and

2. Nested or Hierarchal classification.

We will discuss these in more detail in the next chapter.

3.2 Fixed and Random Effects Model:

It is important to distiguish the kind of model we have chosen before we start analysing it. There are three kinds of models in the previous context:

1. model 1, the fixed effect or the systematic component model;

2. model 2, the random effects or the random-component model, and

3. the mixed effects model.

We will consider an example to explain the three different cases. Let m be the mean tensile strength of items produced on a large group of machines. Consider the model

$$y_{ij} = m_i + e_{ij} \tag{3.1}$$

where y_{ij} is the measurment of the tensile strength of the j^{th} item produced by the i^{th} machine,

m_i is the mean tensile strength produced by the i^{th} machine and

e_{ij} is the deviation (error) due to the j^{th} item within the i^{th} machine.

If the group is relatively large, we may choose a random sample of say s machines out of the total and measure the tensile strength y_{ij} of n items from each chosen machine. In this case we may write $m_i = m + l_i$ where l_i and e_{ij} are random variables with variances σ_l^2 and σ_e^2 respectively. The variance σ_l^2 is a measure of the effect of variability *among* the machines in a group of machines on the tensile strength of items

produced on the machine, and σ_e^2 measures the variability within machines, i.e to say the effect on the tensile strength of items from an individual machine resulting from individual-machine variation.

The total variance of the observations y_{ij} equals $V(y_{ij}) = \sigma_l^2 + \sigma_e^2$. Consequently, we have decomposed the variance into two components. The quantities affecting y_{ij} are the overall mean m, l_i, and e_{ij}. Here m is a constant, while l_i and e_{ij}, because of the method of sampling, are random variables. This is an example of model 2 or the random components model.

Consider the situation where there are a limited number s of machines which are of interest. We may measure the tensile strengths of n items produced on each machine of interest. Again the measurement of the j^{th} item on machine i is denoted by y_{ij} and can be expressed as

$$y_{ij} = m + l_i + e_{ij}. \tag{3.2}$$

In this case our interest is in making inferences about the machines in particular. (In the previous case, we were interested in the characteristics of the population of machines, as reflected by the sample of s machines). Here $l_1, l_2, l_3, ...l_s$ are not random variables, but systematic components, each peculiar to a given machine. The e_{ij} are however random variables. This is an example of systematic- component model or the fixed effects model, also knowm as Model I.

Similarly in the case of two factor experiments, when one of the two factors consists of random components and the other one consists of fixed or systematic components, it is called a mixed model.

3.3 Review of analysis of Factorial Experiments

The literature on the analysis of variance in factorial experiments under the Gaussian model (normal distribution) is extensive. But literature on such analyses under the inverse Gaussian model is limited, though there have been recent developments. A brief review of the work done in this field will be outlined in this chapter.

3.3.1 Tweedie's Work

The earliest work can be dated back to Tweedie [43] who developed a one way analysis of variance for different means μ_i and constant σ. He gave an analogue of the analysis of variance for nested classification. Suppose that we have random samples from each of k classes denoted by $Y_{ij}, i = 1, 2, \ldots, k$ and $j = 1, 2, \ldots, n_i$. Denoting sample averages by dots in the proper positions, we can obtain the following algebraic identity:

$$\sum_{i=1}^{k}\sum_{j=1}^{n_i}\left(\frac{1}{Y_{ij}} - \frac{1}{Y_{..}}\right) = \sum_{i=1}^{k}\sum_{j=1}^{n_i}\left(\frac{1}{Y_{i.}} - \frac{1}{Y_{..}}\right) + \sum_{i=1}^{k}\sum_{j=1}^{n_i}\left(\frac{1}{Y_{ij}} - \frac{1}{Y_{i.}}\right). \tag{3.3}$$

Thus the total sum of residuals of reciprocals is partitioned into two components, a between class component and a within class component. Tweedie proved that if all the observations are from the same inverse Gaussian distribution, the three terms in the above equation are distributed as $1/\lambda$ times χ^2 variables with $\sum n_i - 1, k - 1$, and $\sum n_i - k$ degrees of freedom respectively. He also proved that the two terms on the right are independent. This gives an F test for testing the hypothesis of equal means.We simply divide the between mean reciprocal sum by the within mean reciprocal sum.

3.3.2 Work of Shuster and Miura

In 1972 Shuster and Miura [41] succeeded in giving tests for the two-way analysis of variance with equal cell sizes. Tests for main effects are obtained by applying the F-test for nested classifications to the totals obtained by summing over rows (or columns). A test for interaction was also devised. They considered the one-way modified analysis of reciprocals and also the randomised two-way lay out. Under the one-way modified analysis of variance they consider the following setup. Let $Y_{i1}, Y_{i2}, ..., Y_{in}$ $(i = 1, 2, ..., k)$ be k independently drawn random samples from $IG \sim (\mu_i, \lambda)$ respectively. They consider the test of $H_o : \mu_1 = \mu_2 ... = \mu_k$ against the alternative that H_o is false, at the significance level p. The proposed rejection criterion, which coincides with that of Tweedie [43] is: reject H_o if

$$T = \frac{(\sum_{i=1}^{k}(\bar{Y}_{i.}^{-1} - \bar{Y}_{..}^{-1}))}{\sum_{i=1}^{k}\sum_{j=1}^{n}(Y_{ij}^{-1} - \bar{Y}_{i.}^{-1})} > C_p \tag{3.4}$$

where $\bar{Y}_{i.}$ is the i^{th} sample mean, $\bar{Y}_{..}$ is the grand mean and

$C_p = \frac{(k-1)}{(n-1)k} F_{(k-1,(n-1)kn)}$.

Under H_o, the nk observations Y_{ij} all have the same inverse Gaussian distribution.

In the case of the randomised two way layout, they consider a model wherein the mean of the IG distribution is linear in the factor effects. They consider $Y_{ijk} \sim IG(\mu_i + \beta_j + \alpha_{ij}, \lambda)$ to be independent random variables, $(i = 1, ..., I; j = 1, ..., K)$ with $min(I, J, K) \geq 1$ and $\alpha_{i1} + ..., +\alpha_{iJ} = \alpha_{I1} + ..., +\alpha_{IJ} = 0$

This is the model that they assume and then test for the main effects and also for the interactions. Futher details can be obtained in their paper [41]. They illustrate the test with the help of a numerical example. However Fries and Bhattacharya [5] suggest

that their assumption that the mean is linear in the factor effects is inappropriate and that the tests they propose have the serious drawback that they involve arbitrary pairings of the observations from one cell to another.

Several regression models for the inverse Gaussian distribution have been studied by Davies [21], Whitmore and Yalovsky [55], Bhattacharya and Fries [4] and Whitmore [51]. The work of Davies concerns a single covariate x and is mainly confined to a zero intercept model $\theta = \beta x$ with σ or σx^2 constant. The model $\theta = \beta_0 + \beta_1 x$ is also considered but only some adhoc unbiased estimators are constructed. Whitmore and Yalovsky [55] use the formulation $\theta = exp(\underline{x}'\underline{\beta})$, $\sigma^{-1} = c\ exp(\underline{x}'\underline{\beta})$ where $\underline{x}$ denotes a column p-vector of covariates. However because they make a log-normal approximation by assuming that c is large, their procedure does not really provide a treatment of the inverse Gaussian regression problem. Whitmore [50, 51] considered the reciprocal linear structure of the mean for an IG multiple regression, and obtained solutions of the corresponding likelihood equations. In the setting of a simple regression, Bhattacharyya and Fries [4] investigated the conditions on the design points under which strong consistency and limiting normality of the estimators hold.

3.4 Work of Bhattacharya and Fries

Bhattacharya and Fries [6] considered the following regression problem. Suppose that at each design point $x_i \in R^p, i = 1, 2, \ldots k$, there are independent observations $[Y_{ij} : j = 1, 2..., n_j]$ and suppose that Y_{ij} is an $IG(\mu_i, \sigma)$ variate with $1/\mu_i = x_i'\beta$ or

$\mu^{-1} = X\beta$. Assume also that $k > p+1$. With the following notations

$$n = \sum_i n_i. \quad \bar{Y}_i = n_i^{-1} \sum_j Y_{ij}, \quad \bar{Y} = n^{-1} \sum_i \sum_j Y_{ij} \tag{3.5}$$

$$R = n^{-1} \sum_i \sum_j Y_{ij}^{-1}, \quad V = \sum_i \sum_j (Y_{ij}^{-1} - \bar{Y} i^{-1}), \quad \bar{x} = n^{-1} \sum_i n_i. \tag{3.6}$$

$$C = diag(n_1, ..., n_k), \quad D = diag(\bar{Y}_1, ..., \bar{Y}_k), \quad X' = (x_1, ...x_k). \tag{3.7}$$

The likelihood is proportional to

$\sigma^{-n/2} exp(-Q(\beta)/2\sigma)$.

where $Q(\underline{\beta}) = \sum_{i=1}^{k} \sum_{j=1}^{n_i} Y_{ij}^{-1} [Y_{ij} \underline{x}_i' \underline{\beta} - 1]^2$

The equations defining the maximum likelihood estimates are then

$$S\hat{\beta} = X'C\underline{1}$$

$$n\hat{\sigma} = Q(\beta) \tag{3.8}$$

which is similar to the Gaussian theory normal equations. When X has full rank, then S is non-singular and the above equation has unique roots (which provide a maximum of the likelihood) giving

$$\hat{\beta} = nS^{-1}\bar{x}\hat{\sigma} = R - \hat{\beta}'\bar{x}. \tag{3.9}$$

These roots, need not be the maximum likelihood estimates since $\hat{\beta}' x_i$ may be negative. However, unless n is very small, this is not a serious problem, because of the asymptotic results, sketched below.

Let $n \to \infty$ so that $n_i/n \to h_i > 0$. Define diagonal matrices M and H with $m_{ii} = \mu_i$ and $h_{ii} = h_i$, for $i = 1, 2, ..., k$. And let $\Delta = X'HMX$, then they show that the estimates $\hat{\beta}$ and $\hat{\sigma}$ are strongly consistent. Bhattacharya and Fries also show that

$$\sqrt{n}(\hat{\beta} - \beta) \to n_p(0, \sigma\Delta^{-1}) \sqrt{n}(\hat{\sigma}/\sigma - 1) \to n_1(0, 2). \tag{3.10}$$

and that $\hat{\beta}$ and $\hat{\sigma}$ are asymptotically independent.

Thus, appropriate testing procedures and confidence sets for parameters of this regression model are available for large n.

When there is a replicated design and the model includes a constant term, so that $n_i \geq 2$ for all i, $X = (1, X_1)$ and $\beta = (\beta_o, \beta_1')'$, an analysis of reciprocals is again developed by Bhattacharya and Fries [6] who use the identity

$$Y_{ij}^{-1} - \bar{Y}^{-1} = (x_i'\hat{\beta} - \bar{Y}^{-1}) + (\bar{Y}_i^{-1} - x_i'\hat{\beta}) + (Y_{ij}^{-1} - \bar{Y}_i^{-1}). \tag{3.11}$$

Summing over i and j we get the following equation which decomposes the sum of deviations from the grand mean into three components;

Q_{reg}, due to the regression of β_1 (given β_o),

Q_1, due to lack of fit of the model,

and Q_e, due to the error.

$$\sum_i \sum_j (Y_{ij}^{-1} - \bar{Y}^{-1}) = Q_{reg} + Q_1 + Q_e. \tag{3.12}$$

where

$$\begin{aligned} Q_{reg} &= n(\underline{\bar{x}}'\underline{\beta} - \bar{Y}^{-1}) \\ Q_1 &= \sum_{i=1}^{k} n_i(\bar{Y}_i^{-1} - (\underline{x}_i'\underline{\hat{\beta}} \\ Q_e &= \sum_{i=1}^{k} \sum_{j=1}^{n_i} (Y_{ij}^{-1} - \bar{Y}_i^{-1}). \end{aligned} \tag{3.13}$$

The hypothesis of interest are

(i) $\Theta_1 : \mu_i$ unrestricted,

(ii) $\Theta_2 : \mu_i^{-1} = x_i'\beta$, reciprocal linear model, and

(iii) $\Theta_3 : \mu_i^{-1} = \beta_o$, lack of regression.

Bhattacharya and Fries [6] show that the likelihood test for Θ_1 rejects for large values of

$$F_1 = \frac{Q_1/(k-p)}{Q_e/(n-k)} \tag{3.14}$$

and that F_1 has an approximate F distribution with $(k-p)$ and $(n-k)$ degrees of freedom. And for testing Θ_2, they justify a clever modification of the likelihood ratio test to get the the test statistic

$$F_{reg} = \frac{Q_{reg}/(p-1)}{Q_e/(n-k)} \tag{3.15}$$

which has an approximate F distribution with $(p-1)$ and $(n-k)$ degrees of freedom. Finally they provide inverse Gaussian regression diagnostics for checking model assumptions. By studying the one- sample case, they suggest the construction of a half-normal plot of the estimated standardised inverse Gaussian residuals. If the model is appropriate, then this plot will be approximately linear with slope one.

3.4.1 Factorial Experiments

Fries and Bhattacharya [5] have considered in detail the analysis of two factor experiments. They have considered the two way crossed classification whereas this thesis follows their paper for nested classification.

In their article, they develop inference procedures for balanced two factor experiment with the inverse Gaussian model. Their assumptions entail a linear model for the reciprocal mean $\theta^{-1}$and a constant σ for all levels of the factor.

They consider I levels of row factor A and J levels of the column factor B. At

each cell (i,j), n items are tested and failure times y_{ijk}, $k = 1, 2, ..., n$ are recorded. The observations are assumed to be independent with $y_{ijk} \sim IG(\theta_{ij}, \sigma)$.

They consider the following linear model for the reciprocal mean

$$(\theta_{ij})^{-1} = \mu + \alpha_i + \beta_j, \tag{3.16}$$

$i = 1,2,.....I$ $j = 1,2.....J$, $\sum_{i=1}^{I} \alpha_i = 0$; $\sum_{j=1}^{J} \beta_j = 0$

where μ is the general or the mean effect, α_i represents the the effect of the i^th level of the row factor A, β_j represents the effect of the j^{th} level of the column factor (B).

They find both the maximum likelihood and the least squares estimates of θ_{ij} and σ.

They differentiate the loglikelihood of the density function of the IG distribution

$$l = -(1/2)IJnln(\sigma) - (2\sigma)^{-1} \sum_{i=1}^{I} \sum_{j=1}^{J} \sum_{k=1}^{n} (y_{ijk})^{-1} [y_{ijk}(\mu + \alpha_i + \beta_{ij}) - 1]^2 \tag{3.17}$$

with respect to μ, α_i, β_j and σ and solve the equations in matrix notations to obtain the maximum likelihood estimates of Φ and σ where Φ is defined as $\Phi = \{\mu, \alpha_1, \alpha_2, ...\alpha_{I-1}, \beta_1, \beta_2, ... \beta_{J-1}\}$.

They show that, for every (i,j) there exists an $(I + J - 1)$ vector $\mathbf{x_{ij}}$ consisting of -1's, 0's and 1's such that

$\mu + \alpha_i + \beta_{ij} = \Phi' x_{ij}$ $(1 \leq i \leq I)$ and $(1 \leq j \leq J)$.

They define matrices X and M as

$X' = (x_{11}, x_{12}, ..., x_{IJ})$

$M = X'DX$

of order $IJ \times (I + J - 1)$ and $(I + J - 1) \times (I + J - 1)$ respectively where $D = diag(\bar{y}_{11}, \bar{y}_{12}, ..., \bar{y}_{IJ})$

They give the maximum likelihood estimate of Φ as

$$\hat{\Phi} = M^{-1}\delta \tag{3.18}$$

where $\delta = IJ(1, 0, ...0)' = IJu.$

And that of σ as

$$\hat{\sigma} = (IJn)^{-1}[R - n\delta' M^{-1}\delta]. \tag{3.19}$$

where $R = \sum_i \sum_j \sum_k y_{ijk}^{-1}$.

They find that the estimators $\hat{\Phi}$ and $\hat{\sigma}$ are strongly consistent and their limiting distributions are asymptotically normal.

They also consider the following hypotheses;

$\Omega_4 : \theta_{ij}'s$ unrestricted (general model),

$\Omega_3 : \theta_{ij}^{-1} = \mu + \alpha_i + \beta j. \quad \sum \alpha_i = \sum \beta_j = 0$ (additive model),

$\Omega_2 : \theta_{ij}^{-1} = \mu + \alpha_i, \ \sum \alpha_i = 0$ (no B effects),

$\Omega_1 : \theta_{ij}^{-1} = \mu + \beta_j, \ \sum \beta_j = 0$ (no A effects),

$\Omega_o : \theta_{ij}^{-1} = \mu$ (no factor effects).

The maximum likelihood estimates of σ , $\hat{\sigma}_s$ under the model $\Omega_s, s = 0, 1, 2, 3, 4$ respectively are found as

$$\begin{aligned} IJn\hat{\sigma}_4 &= R - n\sum\sum \bar{y}_{ij}^{-1} \\ IJn\hat{\sigma}_3 &= R - n\sum\sum \hat{\theta}_{ij}^{-1} = R - IJn\hat{\mu} \\ IJn\hat{\sigma}_2 &= R - nJ\sum \bar{y}_{i.}^{-1} \end{aligned}$$

$$IJn\hat{\sigma}_1 = R - nI\sum \bar{y}_{\cdot j}^{-1}$$

$$IJn\hat{\sigma}_0 = R - nIJ\bar{y}_{\cdot\cdot}^{-1}. \tag{3.20}$$

With the following quantities called sums of reciprocals, they test the different hypotheses:

$$R_A = IJn(\hat{\sigma}_1 - \hat{\sigma}_3) = n\sum\sum(\hat{\theta}_{ij}^{-1} - \bar{y}_{\cdot j}^{-1}).$$

$$R_B = IJn(\hat{\sigma}_2 - \hat{\sigma}_3) = n\sum\sum(\hat{\theta}_{ij}^{-1} - \bar{y}_{i\cdot}^{-1}).$$

$$R_{AB} = IJn(\hat{\sigma}_3 - \hat{\sigma}_4) = n\sum\sum(\bar{y}_{ij}^{-1} - \hat{\theta}_{ij}^{-1}).$$

$$R_E = IJn\hat{\sigma}_4 = R - n\sum\sum\bar{y}_{ij}^{-1}. \tag{3.21}$$

They show that the hypothesis Ω_3 or the test for additivity (or no interaction) rejects for large values of

$$F_{AB} = \frac{R_{AB}/(I-1)(J-1)}{R_E/IJ(n-1)} \tag{3.22}$$

where F_{AB} has approximate F distribution with $[(I-1)(J-1), IJ(n-1)]$ degrees of freedom.

Similarly, with some additional clever adjustments, they have devised F tests for testing the null hypotheses Ω_2 (no B effects) and Ω_1 (no A effects).

They also give the analysis of reciprocals, (ANOR) table very similar to the analysis of variance table (ANOVA). They also develop confidence intervals for the contrasts.

They discuss the least squares approach based on the initial reduction of the data by sufficiency. So given the model

$$\theta_{ij}^{-1} = \mu + \alpha_i + \beta_j \tag{3.23}$$

where μ α_i, β_j have the same meaning as before and assuming $y_{ijk} \sim IG(\theta_{ij}, \sigma)$, the moment results given by Tweedie [43] states that

$$E(\bar{y}_{ij}^{-1}) = \mu + \alpha_i + \beta_{ij} + (\sigma/n)$$

$$Var(\bar{y}_{ij}^{-1}) = (\mu + \alpha_i + \beta_{ij})\sigma/n + 2(\sigma/n)^2. \tag{3.24}$$

Let $\bar{y}_{ij}^{-1} = r_{ij}$. Then we have

$$E(r_{ij}) = \theta_{ij}^{-1} + \sigma/n. \tag{3.25}$$

Letting

$$\mu^* = \mu + \sigma/n \tag{3.26}$$

the expected value of r_{ij} is

$$E(r_{ij}) = \mu^* + \alpha_i + \beta_{ij}. \tag{3.27}$$

Now we may consider the linear model

$$R = X\Psi^* + e \tag{3.28}$$

where $R = (r_{11}, r_{12}, \quad \ldots \ r_{IJ})'$, X is the matrix of appropriate constants,

$\Psi^* = (\mu^*, \alpha', \beta')'$, $e = (e_{11}, e_{12}, \ldots . \ e_{IJ})'$, the error vector such that $E(e) = 0$.

The standard results for the two-way layout provide the simple linear (in reciprocals) unbiased estimators for $\hat{\mu}^*, \alpha_i, \beta_j$ as

$$\begin{aligned} \hat{\mu^*} &= \bar{r}_{..} \\ \hat{\alpha}_i &= \bar{r}_{i.} - \bar{r}_{..} \\ \hat{\beta}_i &= \bar{r}_{.j} - \bar{r}_{..} \end{aligned}$$

Furthermore the parameters σ and μ are estimated as

$$\hat{\sigma} = \frac{R - \sum\sum \bar{y}_{ij}}{IJ(n-1)}.$$

$$\hat{\mu} = \bar{r}_{..} - \hat{\sigma}/n.$$

Bhattacharyya and Fries [5] show that the least squares estimates are in general not efficient, but that they are consistent. By using the consistent estimators of the variances given in equation 3.24, as the weights, one can get efficient estimates using the weighted least squares method of estimation. It is easy to see that if we consider weights by estimating 3.24 upto the order $1/n$, the weighted least squares estimates coincides with that of maximum likelihood estimates.

In the next chapter, analogous results are obtained for the case of nested classification.

Chapter 4

Nested Classification under IG Model

4.1 Introduction

In this chapter we develop the analysis of nested factor experiments under the inverse Gaussian model. From equation 1.1 substituting $\mu = \theta$ and $\sigma = 1/\lambda$ we get the pdf of the IG distribution as

$$f(y;\theta,\sigma) = \frac{1}{(\sqrt{(2\pi\sigma)y^{3/2}})} exp - [\frac{((y/\theta)-1)^2}{2\sigma y}] \tag{4.1}$$

for $y > 0: \quad \theta > 0; \quad \sigma > 0.$

The derivation of the IG distribution can be cast in the context of fatigue growth or accumulation of fatigue or damage over a period of time according to a Weiner process as shown in Chapter 1.

In this chapter, we develop inference procedures for a nested two factor experiment

with the IG model. We assume that the critical level ω and the diffusion parameter δ^2 of the underlying Weiner processes are constants, while the drift η (the drift η measures the mean fatigue growth per unit time) is linear in its factor effects. So the two important assumptions here are

(i) ω, the critical level after which point the object fails and δ^2, the diffusion parameter are both constants,

(ii) drift η is linear in factor effects.

But

$$\theta = \omega/\eta. \text{ and} \tag{4.2}$$

$$\sigma = [\delta/\omega]^2. \tag{4.3}$$

These lead us to a linear model for the reciprocal mean θ^{-1} and a constant σ for all levels of the factor. The constancy of σ parallels the homoscedasticity assumption in the normal theory analysis of variance.

4.2 Nested Classification

In certain multifactor experiments the levels of one factor (for e.g., say factor B) are similar but not identical for different levels of another factor (e.g., A). Such an arrangement is called a *nested* or *hierarchial* design, with the levels of factor B nested under the levels of factor A. For example, consider a company that purchases its raw material from three different suppliers. The company wishes to determine if the purity of the raw material is the same from each supplier. There are four batches

of raw material available from each supplier, and three determinations of purity are to be taken from each batch.

This is a *two-stage nested* or *hierarchial design*, with batches nested within suppliers. One may think as to why the two factors suppliers and batches are not crossed. If the factors are crossed, then batch 1 would always refer to the same batch, batch 2 would always refer to the same batch and so on. This is clearly not the case, as the batches from each supplier are unique for that particular supplier. That is batch 1 from supplier 1 has no connection with batch 1 from any other supplier, batch 2 from supplier 1 has no connection with batch 2 from any other supplier, etc. To emphasize the fact that batches from each supplier are different batches, we may remember the batches as 1,2,3 and 4 from supplier 1, then 5,6,7 8 from supplier 2, and finally 9,10,11 and 12 from supplier 3.

Sometimes we may be uncertain as to whether a factor is crossed or nested. If the levels of the factor can be renumbered arbitrarily as stated, then the factor is nested. In nested classification, the subordinate classification is nested within the higher level of classification where as in crossed classification, the two sources of variation are of equal rank.

4.3 The Model

We now consider a two factor nested life test with J levels of the secondary or the column factor B nested within I levels of the primary or row factor A.

We consider the fixed effects or the systematic components model.

At each factor setting or cell (i,j), n items are tested and their failure times y_{ijk}, $k = 1, 2.....n$ are recorded.

These observations are assumed to be independent with y_{ijk} distributed as inverse Gaussian with mean θ_{ij} and the other parameter being σ, ie $y_{ijk} \sim IG(\theta_{ij}, \sigma)$.

We have assumed that the drift η is linear in the nested factor effects. But since the mean life time is inversely proportional to the drift, we now assume a linear model for the reciprocal of the mean i.e. the reparametrisation of the nested factor effects suggests the model

$$(\theta_{ij})^{-1} = \mu + \alpha_i + \beta_{ij}. \tag{4.4}$$

$i = 1, 2, ...I$, $j = 1, 2.....J$, $\sum_{i=1}^{I} \alpha_i = 0$, and $\sum_{j=1}^{J} \beta_{ij} = 0$

where μ is the general or the mean effect, α_i represents the the effect of the i^th level of the primary factor A, and β_{ij} represents the effect of the j^{th} level of the secondary factor (B) nested in the i^{th} level of the primary factor (A).

The assumptions of $\sum \alpha_i = 0$, $\sum \beta_{ij} = 0$ assures that there are IJ unknowns. For the IG distribution we must have $(\theta_{ij})^{-1} > 0$ for all (i,j) and $\sigma > 0$

Thus the parameters are

μ, $\alpha' = (\alpha_1, \alpha_2, ...\alpha_I)$, $\beta' = (\beta_{11}, \beta_{12}, ...\beta_{1J}, \beta_{21}, \beta_{22}, ...\beta_{IJ})$ and σ. And they lie in the set $\Omega = \{\mu, \alpha', \beta', \sigma\}$ with $\sum_{i=1}^{I} \alpha_i = 0$ $\quad , \sum_{i=1}^{I} \sum_{j=1}^{J} \beta_{ij} = 0$; and $\mu + \alpha_i + \beta_{ij} > 0$; $\quad \forall i, j, \quad \sigma > 0$;.

4.3.1 An Illustration

If we assume 2 levels for the primary factor A and 3 levels for the secondary factor B, we would have the following model with $I = 2$; $J = 3$;

$$(\theta_{ij})^{-1} = \mu + \alpha_i + \beta_{ij} \tag{4.5}$$

$i = 1, 2$ and $j = 1, 2, 3$.

We would have the following system of equations:

$$\begin{aligned}
\theta_{11}^{-1} &= \mu + \alpha_1 + \beta_{11} \\
\theta_{12}^{-1} &= \mu + \alpha_1 + \beta_{12} \\
\theta_{13}^{-1} &= \mu + \alpha_1 + \beta_{13} \\
\theta_{21}^{-1} &= \mu + \alpha_2 + \beta_{21} \\
\theta_{22}^{-1} &= \mu + \alpha_2 + \beta_{22} \\
\theta_{23}^{-1} &= \mu + \alpha_2 + \beta_{23}
\end{aligned}$$

But since we have $\sum_{i=1}^{2} \alpha_i = 0; \alpha_2 = -\alpha_1$ and

$\sum_{i=j}^{3} \beta_{ij} = 0 \Rightarrow \beta_{13} = -(\beta_{11} + \beta_{12})$ *and* $\beta_{23} = -(\beta_{21} + \beta_{22})$.

So the new set of equations is:

$$\begin{aligned}
\theta_{11}^{-1} &= \mu + \alpha_1 + \beta_{11} \\
\theta_{12}^{-1} &= \mu + \alpha_1 + \beta_{12} \\
\theta_{13}^{-1} &= \mu + \alpha_1 - (\beta_{11} + \beta_{12}) \\
\theta_{21}^{-1} &= \mu - \alpha_1 + \beta_{21} \\
\theta_{22}^{-1} &= \mu - \alpha_1 + \beta_{22}
\end{aligned}$$

$$\theta_{23}^{-1} = \mu - \alpha_1 - (\beta_{21} + \beta_{22})$$

and the parameters to be estimated are $\{\mu, \alpha_1, \beta_{11}, \beta_{12}, \beta_{21}, \beta_{22}\}$. We will discuss this example in more detail in the next section.

4.3.2 Basic Notations

We now introduce some basic notations for the totals and the means:

1. y_{ij} is the total of the observations y_{ijk} in the $(i,j)^{th}$ cell summed over k from 1 to n, ie

$$y_{ij} = \sum_{k=1}^{n} y_{ijk} \tag{4.6}$$

$$\bar{y}_{ij} = \frac{\sum_{k=1}^{n} y_{ijk}}{n} \tag{4.7}$$

the mean of the $(i,j)^{th}$ cell.

2. $y_{i.}$ is the total of the observations y_{ij} summed over j.

$$y_{i.} = \sum_{j=1}^{J} y_{ij}, \tag{4.8}$$

$$\bar{y}_{i.} = \frac{\sum_{j=1}^{J} y_{ij}}{nJ}. \tag{4.9}$$

3. $y_{..}$ is the grand total of all the observations summed over i, j, k

$$y_{..} = \sum_{i=1}^{I}\sum_{j=1}^{J} y_{ij}, \tag{4.10}$$

$$\bar{y}_{..} = \frac{\sum_{i=1}^{I}\sum_{j=1}^{J} y_{ij}}{nIJ}. \tag{4.11}$$

4. $y_{.j}$ is the total of the observations y_{ij} summed over i from 1 to I.

$$y_{.j} = \sum_{i=1}^{I} y_{ij}, \tag{4.12}$$

$$\bar{y}_{.j} = \frac{\sum_{i=1}^{I} y_{ij}}{nI}. \tag{4.13}$$

5. D is a matrix with its diagonal elements as the mean of the observations y_{ijk} summed over k. $D = diag(\bar{y}_{11}, \bar{y}_{12}, ...\bar{y}_{IJ})$.

$$D = \begin{pmatrix} \bar{y}_{11} & 0 & 0 & 0 & \cdots & 0 \\ 0 & \bar{y}_{12} & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & \bar{y}_{IJ} \end{pmatrix} \tag{4.14}$$

6.R represents the grand sum of the reciprocals of the observations y_{ijk} summed over i, j and$k$

ie $R = \sum_{i=1}^{I} \sum_{j=1}^{J} \sum_{k=1}^{n} y_{ijk}^{-1}$.

The unit vector (1,0 0 0....0 0 0) will be denoted by u' and its dimensionality would be indicated and clear from the context. The symbols $N_p(\mu, \Sigma)$ denotes a p variate normal distribution with mean μ and covariance matrix Σ, and χ^2_ν denotes a chi squared distribution with ν degrees of freedom.

Some important properties of the inverse Gaussian distribution proved in Chapter 1 will be used here, some of which are:

(i) y_{ij} and $\sum_{k=1}^{n}(y_{ijk}^{-1} - \bar{y}_{ij}^{-1})$ are independent

(ii) $\bar{y}_{ij} \sim IG(\theta_{ij}, \sigma/n)$ and

(iii) $\sum_{k=1}^{n}(y_{ijk}^{-1} - y_{ij}^{-1}) \sim \chi^2_{(n-1)}$.

Moreover

$$E(\bar{y_{ij}}^{-1}) = \theta_{ij}^{-1} + \sigma/n. \tag{4.15}$$

$$Var(\bar{y_{ij}}^{-1}) = \theta_{ij}^{-1}\sigma/n + 2(\sigma/n)^2. \tag{4.16}$$

4.4 Maximum Likelihood Estimation

We have $y_{ijk} \sim IG(\theta_{ij}, \sigma)$. From equation 4.1 the loglikelihood function is given by

$$\begin{aligned} l &= ln\{\prod_{i=1}^{I}\prod_{j=1}^{J}\prod_{k=1}^{n}[(2\pi\sigma)^{1/2}(y_{ijk})^{-3/2} \\ &\quad exp[-(2\sigma y_{ijk})^{-1}(y_{ijk}\theta_{ij}^{-1} - 1)^2]]\} \\ &= (-1/2)IJn\ ln\ 2\pi - (1/2)IJn\sigma - (3/2)\sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{k=1}^{n} ln\ y_{ijk} \\ &\quad -(2\sigma)^{-1}\sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{k=1}^{n}(y_{ijk})^{-1}[y_{ijk}(\theta_{ij})^{-1} - 1]^2 \end{aligned} \tag{4.17}$$

Substituting for θ_{ij}^{-1} as $\mu+\alpha_i+\beta_{ij}$ and ignoring the first and the third terms which do not involve any parameters, the loglikelihood apart from an additive constant is

$$l = -(1/2)IJn\ ln\ (\sigma) - (2\sigma)^{-1}\sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{k=1}^{n}(y_{ijk})^{-1}[y_{ijk}(\mu + \alpha_i + \beta_{ij}) - 1]^2 \tag{4.18}$$

We now obtain a set of linear equations by finding the partial derivatives of this loglikelihood function with respect to μ, α_i, β_{ij} and σ and equating them to zero.

Differentiating the above equation with respect to μ we obtain

$$y_{..}\hat{\mu} + \sum_i \hat{\alpha}_i y_{i.} + \sum_i\sum_j \beta_{ij}y_{ij} = IJn. \tag{4.19}$$

Differentiating wrt α_i gives

$$y_{i.}\hat{\mu} + \hat{\alpha}_i y_{i.} + \sum_j \beta_{ij}y_{ij} = Jn. \tag{4.20}$$

Differentiating wrt β_{ij} gives

$$y_{ij}\hat{\mu} + \hat{\alpha}_i y_{ij} + \beta_{ij}y_{ij} = n. \tag{4.21}$$

And finally differentiating wrt σ gives

$$\hat{\sigma} = (IJn)^{-1}\sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{k=1}^{n}(y_{ijk})^{-1}[y_{ijk}(\hat{\mu} + \hat{\alpha}_i + \hat{\beta}_{ij}) - 1]^2 \tag{4.22}$$

We can observe here that these equations are linear in the parameters just like the likelihood equations under the usual normal model (analysis of variance setting). However there is a remarkable difference in the positions of the random and constant coeffients.Their positions are reversed here. The right sides of these equations are constants where as in the normal case they are random variables. And the coefficients in the left sides are random variables whereas in the normal case they are constants.

Using the constraints $\sum_i \alpha_i = \sum_j \beta_{ij} = 0$ we can reduce the problem to a system of IJ equations in IJ unknowns.

We now try to use a matrix representation of these equations and use it to prove that the root is unique and that it maximises the likelihood.

We delete the last component of α, i.e. α_I and I components from β using the conditions $\sum_i \alpha_i = \sum_j \beta_{ij} = 0$ and hence the parameters to be estimated are

$\{\mu, \alpha_1, \alpha_2, ...\alpha_{I-1}, \beta_{11}, \beta_{12}, ... \ \beta_{1J-1}, \beta_{21}, \beta_{22}, ... \ \beta_{2J-1}, ... \ \beta_{IJ-1}\}$.

Let $\Phi = \{\mu, \alpha_1, \alpha_2, ...\alpha_{I-1}, \beta_{11}, \beta_{12}, ... \ \beta_{1J-1}, \beta_{21}, \beta_{22}, ... \ \beta_{2J-1}, ... \ \beta_{IJ-1}\}'$

and $\Psi = (\Phi', \sigma)'$

Now for every (i, j) we can identify the IJ vector $\mathbf{x_{ij}}$ consisting of -1 0 and 1 such that

$\mu + \alpha_i + \beta_{ij} = \Phi' x_{ij}$ $(1 \leq i \leq I)$ and $(1 \leq j \leq J)$

For example in the previous section we considered an illustration where we assumed $I = 2$ and $J = 3$. Our equations were

$$\theta_{11}^{-1} = \mu + \alpha_1 + \beta_{11}. \tag{4.23}$$

$$\theta_{12}^{-1} = \mu + \alpha_1 + \beta_{12}. \tag{4.24}$$

$$\theta_{13}^{-1} = \mu + \alpha_1 - (\beta_{11} + \beta_{12}). \tag{4.25}$$

$$\theta_{21}^{-1} = \mu - \alpha_1 + \beta_{21}. \tag{4.26}$$

$$\theta_{22}^{-1} = \mu - \alpha_1 + \beta_{22}. \tag{4.27}$$

$$\theta_{23}^{-1} = \mu + \alpha_1 - (\beta_{21} + \beta_{22}). \tag{4.28}$$

So

$\Phi = (\mu, \alpha_1, \beta_{11}, \beta_{12}, \beta_{21}, \beta_{22})$

$\Theta^{-1} = X\phi$

where

$$X = \begin{pmatrix} 1 & 1 & 1 & 0 & 0 & 0 \\ 1 & 1 & 0 & 1 & 0 & 0 \\ 1 & 1 & -1 & -1 & 0 & 0 \\ 1 & -1 & 0 & 0 & 1 & 0 \\ 1 & -1 & 0 & 0 & 0 & 1 \\ 1 & -1 & 0 & 0 & -1 & -1 \end{pmatrix} \tag{4.29}$$

$\Theta^{-1} = (\theta_{11}^{-1}, \theta_{12}^{-1}, \theta_{13}^{-1}, \theta_{21}^{-1}, \theta_{22}^{-1}, \theta_{23}^{-1})'$

$\Phi = (\mu, \alpha_1, \beta_{11}, \beta_{12}, \beta_{21}, \beta_{22})'$

In this case x_{IJ} or x_{23} = [1 -1 0 0 -1 -1] since

$\theta_{23}^{-1} = \Phi' \mathbf{x_{23}} \Rightarrow$

$$\begin{bmatrix} \mu \\ \alpha_1 \\ \beta_{11} \\ \beta_{12} \\ \beta_{21} \\ \beta_{22} \end{bmatrix} [1\ -1\ 0\ 0\ -1\ -1] = \theta_{23}^{-1} = \mu - \alpha_1 - (\beta_{21} + \beta_{22}).$$

We now define the $(IJ \times IJ)$ and $(IJ \times IJ)$ matrices X and M as

$X' = (x_{11}\ \ x_{12}\ \ x_{13}, \ldots\ \ x_{IJ})$;

$M = X'DX$ where $D = diag(\bar{y}_{11}, \bar{y}_{12}, \ldots . \bar{y}_{IJ})$.

$$D = \begin{bmatrix} \bar{y}_{11} & 0 & 0 & 0 & \cdots & 0 \\ 0 & \bar{y}_{12} & 0 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & \cdots & \bar{y}_{IJ} \end{bmatrix} \tag{4.30}$$

We define $X'\mathbf{1} = IJ(1\ \ 0\ \ 0\ \ 0\ \ \ldots 0\ \ 0) = IJu = \delta$.

For example in the previous case

$$X'1 = \begin{bmatrix} 1 & 1 & 1 & 1 & 1 & 1 \\ 1 & 1 & 1 & -1 & -1 & -1 \\ 1 & 0 & -1 & 0 & 0 & 0 \\ 0 & 1 & -1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & -1 \\ 0 & 0 & 0 & 0 & 1 & -1 \end{bmatrix} \begin{bmatrix} 1 \\ 1 \\ 1 \\ 1 \\ 1 \\ 1 \end{bmatrix} = \begin{bmatrix} 6 \\ 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} \tag{4.31}$$

$= 6[1\ \ 0\ \ 0\ \ 0\ \ 0\ \ 0]' = IJ[1\ \ 0\ \ 0\ \ 0\ \ 0\ \ 0] = IJu$ which we define as δ

To see that the matrix M is positive definite, we can observe that the diagonal matrix D is positive definite, since the $y'_{ij}s$ are positive. Also the standard results for the two way layout entail that rank(X') = rank $(X'X) = IJ$. So the only vector ω satisfying $(X\omega) = 0$ is the null vector. Consequently, for an arbitrary $\omega \neq 0$, we have $\omega' M\omega = (X\omega)'D'(X\omega) > 0$. Hence M is positive definite.

We now try to change the form of the loglikelihood equation into the aforementioned notation.

We have the loglikelihood function derived previously as

$$l = -(1/2)(IJn)\ ln\ (\sigma) - (2\sigma)^{-1}\sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{k=1}^{n}(y_{ijk})^{-1}[y_{ijk}(\mu + \alpha_i + \beta_{ij}) - 1]^2 \quad (4.32)$$

Now substituting for

$$\mu + \alpha_i + \beta_{ij} = \Phi' x_{ij} \quad (4.33)$$

we have

$$\begin{aligned} l &= -(1/2)(IJn)\ ln\ (\sigma) - (2\sigma)^{-1}\sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{k=1}^{n}(y_{ijk})^{-1}[y_{ijk}(\Phi' x_{ij}) - 1]^2 \quad (4.34)\\ &= -(1/2)IJn\ ln\ (\sigma) - (2\sigma)^{-1}\sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{k=1}^{n}(y_{ijk})^{-1} \\ &\quad [y_{ijk}^2(\Phi' x_{ij})x'_{ij}\Phi + 1 - 2y_{ijk}\Phi' x_{ij}]. \quad (4.35)\\ &= -(1/2)IJn\ ln\ (\sigma) - (2\sigma)^{-1}\sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{k=1}^{n}[y_{ijk}(\Phi' x_{ij})x'_{ij}\Phi + y_{ijk}^{-1} - 2\Phi' x_{ij}] \quad (4.36) \end{aligned}$$

But we have denoted $\sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{k=1}^{n} y_{ijk}^{-1} = R$ so we have

$$\begin{aligned} l &= -(1/2)IJn\ ln\ (\sigma) - (2\sigma)^{-1}[\Phi'\sum_{i=1}^{I}\sum_{j=1}^{J} x_{ij}x'_{ij}\sum_{k=1}^{n} \\ &\quad y_{ijk}\Phi + R - 2n\Phi'\sum_{i=1}^{I}\sum_{j=1}^{J} x_{ij}]. \quad (4.37)\\ &= -(1/2)IJnln(\sigma) - (2\sigma)^{-1}[\Phi'\sum_{i=1}^{I}\sum_{j=1}^{J} x_{ij}d_{ij}x'_{ij} + R - 2n\Phi'\delta]. \quad (4.38) \end{aligned}$$

Since $\sum i \sum j x_{ij} = IJ = IJ[1 \quad 0 \quad 0...0 \quad 0 \quad 0] = IJu$ in vector notation.

So

$$\begin{aligned} l &= -(1/2)IJn\ ln\ (\sigma) - (2\sigma)^{-1}[R - 2n\Phi'\delta + n\Phi' M\Phi]. \\ &= -(1/2)IJn\ ln\ (\sigma) - (2\sigma)^{-1}[R - 2n\Phi'\delta + n\Phi' X' DX\Phi]. \\ &= -(1/2)IJn\ ln\ (\sigma) - (2\sigma)^{-1}[R - 2n\Phi'\delta + n\Phi' M\Phi]. \end{aligned} \tag{4.39}$$

The unknown parameters to be estimated here are Φ and σ. So differentiating the above equation wrt Φ we obtain

$$\frac{\partial l}{\partial \Phi} = \frac{n}{\sigma}[\delta - M\Phi]. \tag{4.40}$$

And differentiating wrt σ we obtain

$$\frac{\partial l}{\partial \sigma} = \frac{IJn}{2\sigma} + 1/2\sigma^2[R - 2n\Phi'\delta + n\Phi' M\Phi] \tag{4.41}$$

Since M is positive definite, the likelihood equations corresponding to the above equations have unique solution, i.e.

$\hat{\Psi} = (\hat{\Phi}' \quad \hat{\sigma})$

From (4.40) equating it to 0 we obtain

$$\begin{aligned} \frac{\sigma}{n}[\delta - M\Phi] &= 0 \\ \delta - M\Phi &= 0 \Rightarrow \delta = M\Phi \\ or \quad \hat{\Phi} &= M^{-1}\delta \end{aligned} \tag{4.42}$$

From (4.41) we obtain

$$-\frac{IJn}{2\sigma} + 1/2\sigma^2[R - 2n\Phi'\delta + n\Phi' M\Phi] = 0$$

$$1/\sigma[R - 2n\Phi'\delta + n\Phi' M\Phi] = IJ.$$

$$\hat{\sigma} = IJn^{-1}[R - 2n\hat{\Phi}'\delta + n\hat{\Phi}' M\hat{\Phi}] \tag{4.43}$$

Substituting for $\hat{\Phi} = M^{-1}\delta$, we have

$$\begin{aligned}
\hat{\sigma} &= IJn^{-1}[R - 2n\hat{\Phi}'\delta + n\hat{\Phi}' MM^{-1}\delta] \\
&= IJn^{-1}[R - 2n\hat{\Phi}'\delta + n\hat{\Phi}'\delta] \\
&= IJn^{-1}[R - n\hat{\Phi}'\delta] \quad \text{Since } \hat{\Phi}' = \delta' M^{-1} \\
&= IJn^{-1}[R - n\delta' M^{-1}\delta] \qquad (4.44)
\end{aligned}$$

So the Maximum likelihood estimators in vector notations for $\hat{\Phi}$ and $\hat{\sigma}$ are

$$\begin{aligned}
\hat{\Phi} &= M^{-1}\delta \\
\hat{\sigma} &= IJn^{-1}[R - n\delta' M^{-1}\delta]
\end{aligned}$$

Taking the second partial derivatives of l with respect to Φ and σ we have

$$\begin{aligned}
\frac{\partial^2 l}{\partial\Phi\partial\Phi'} &= \partial/\partial\Phi \quad [\partial l/\partial\Phi'] \\
&= \partial/\partial\Phi[n/\sigma(\delta' - \Phi' M)] \\
&= (n/\sigma)M. \qquad (4.45)
\end{aligned}$$

Similarly

$$\begin{aligned}
\frac{\partial^2 l}{\partial^2\sigma} &= \partial/\partial\sigma[-\frac{IJn}{2\sigma} + 1/2\sigma^2[R - 2n\Phi'\delta + n\Phi' M\Phi]] \\
&= \partial/\partial\sigma[R - n\delta' M^{-1}\delta] \\
&= 1/2[IJ\hat{\sigma^{-2}}] \qquad (4.46)
\end{aligned}$$

And $\frac{\partial^2 l}{\partial\sigma\partial\Phi'} = 0$, $\frac{\partial^2 l}{\partial\Phi'\partial\sigma} = 0$

Therefore

$$\frac{\partial^2 l}{\partial \Psi \partial \Psi'} = \begin{bmatrix} \frac{\partial^2 l}{\partial \Phi \partial \Phi'} & \frac{\partial^2 l}{\partial \sigma \partial \Phi'} \\ \frac{\partial^2 l}{\partial \Phi' \partial \sigma} & \frac{\partial^2 l}{\partial^2 \sigma} \end{bmatrix} \tag{4.47}$$

$$= -n \begin{bmatrix} \hat{\sigma}^{-1} M & 0 \\ 0 & 1/2[IJ\hat{\sigma}^{-2}] \end{bmatrix} \tag{4.48}$$

This matrix is positive definite since M is positive definite and $\hat{\sigma} > 0$ and hence $\hat{\Psi}$ indeed maximises the likelihood.

To claim that $\hat{\Psi}$ is the maximum likelihood estimator, we need to show that it obtains within the restricted parameter space Ω ie $\hat{\theta_{ij}}^{-1} = \hat{\Phi}' x_{ij} > 0$ for all i, j. This result holds for the special case $I = J = 2$, which has been proved for by Fries [25]. But it has not been proved generally. However the theorem proved by Fries and Bhattacharya [5] given below shows that Ψ serves the primary goal of maximum likelihood estimation, which is to provide an efficient likelihood estimator.

Theorem 1. The estimators $\hat{\Phi}$ and $\hat{\sigma}$ are strongly consistent. and the limiting distributions of $n^{1/2}(\hat{\Phi} - \Phi)$ and $n^{1/2}(\hat{\sigma} - \sigma)$ are $N_{IJ}(0, \sigma\Gamma_{-1})$ and $N_1(0, 2\sigma^2(IJ)^{-1})$ respectively where

$$\Gamma = X'\Theta X, \quad \Theta = diag(\theta_{11}, \theta_{12}, ..., \theta IJ) \tag{4.49}$$

Moreover, $\hat{\Phi}$ and $\hat{\sigma}$ are asymptotically independent.

For proof see Fries and Bhattacharyya [5].

4.5 Test of Hypotheses

In the preceeding section, we have estimated the parameters using the maximum likelihood method of estimation. We now use these estimates and other results to develop likelihood ratio tests and tests of hypotheses. Instead of resorting entirely to the asymptotic results for the likelihood ratio statistics some variants of these tests will be constructed by intermixing the asymptotic and exact sampling distribution properties of the component statistics.

We now consider the following relevant models for testing the hypotheses.

1. Ω_1 : θ_{ij} 's unrestricted. (general model)

Under this general model we have the estimate of $\hat{\theta_{ij}} = y_{ij}$. This can also be shown from the equations we obtained by differentiating the loglikelihood function. Consider the equation we obtained by differentiating l wrt β_{ij}

$$y_{ij}\hat{\mu} + \hat{\alpha}_i y_{ij} + \beta_{ij} y_{ij} = n. \tag{4.50}$$

From this equation we obtain

$$\begin{aligned} y_{ij}[\hat{\mu} + \hat{\alpha}_i + \beta_{ij}] &= n, \\ \text{i.e. } y_{ij}\theta_{ij}^{-1} &= n. \\ \hat{\theta_{ij}}^{-1} &= \bar{y_{ij}}^{-1}. \end{aligned} \tag{4.51}$$

We would like an estimate of $\hat{\sigma}$ for the various models. From the previous section we have

$$\hat{\sigma} = (IJn)^{-1} \sum_{i=1}^{I} \sum_{j=1}^{J} \sum_{k=1}^{n} (y_{ijk})^{-1} [y_{ijk}(\theta_{ij}^{-1}) - 1]^2$$

$$
\begin{aligned}
IJn\hat{\sigma} &= \sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{k=1}^{n}(y_{ijk})^{-1}[y_{ijk}^2(\theta_{ij}^{-2}) + 1 - 2\theta_{ij}^{-1}y_{ijk}] \\
&= \sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{k=1}^{n}[y_{ijk}(\theta_{ij}^{-2}) + y_{ijk}^{-1} - 2\theta_{ij}^{-1}] \\
&= \sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{k=1}^{n}[y_{ijk}(\theta_{ij}^{-2}) + \sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{k=1}^{n}y_{ijk}^{-1} - 2\sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{k=1}^{n}\theta_{ij}^{-1}] \\
&= \sum_{i=1}^{I}\sum_{j=1}^{J}ny_{ijk}(\theta_{ij}^{-2}) + R - 2n\sum_{i=1}^{I}\sum_{j=1}^{J}\theta_{ij}^{-1}]
\end{aligned}
$$

Substituting for $\hat{\theta_{ij}}^{-1}$ as

$$\hat{\theta}_{ij}^{-1} = \bar{y_{ij}}^{-1} \tag{4.52}$$

we get

$$
\begin{aligned}
IJn\hat{\sigma}_1 &= \sum_{i=1}^{I}\sum_{j=1}^{J}n\bar{y_{ij}}^{-1} + R - 2n\sum_{i=1}^{I}\sum_{j=1}^{J}\bar{y_{ij}}^{-1} \\
&= R - n\sum_{i=1}^{I}\sum_{j=1}^{J}\bar{y_{ij}}^{-1}.
\end{aligned} \tag{4.53}
$$

2. Consider the nested model $\Omega_2 : \theta_{ij}^{-1} = \mu + \alpha_i + \beta_{ij}$

$\sum_{i=1}^{I}\alpha_i = 0; \quad \sum_{j=1}^{J}\beta_{ij} = 0$

Here in this case we have obtained the maximum likelihood estimate of $\hat{\theta_{ij}}^{-1}$ as $\Phi' x_{ij}$. So we have

$$
\begin{aligned}
IJn\hat{\sigma}_2 &= \sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{k=1}^{n}y_{ijk}^{-1} - n\sum_{i=1}^{I}\sum_{j=1}^{J}\hat{\theta_{ij}^{-1}} \\
IJn\hat{\sigma}_2 &= R - n\sum\sum\hat{\Phi}' x_{ij} \\
&= R - nIJ\hat{\mu}
\end{aligned} \tag{4.54}
$$

3. Consider the hypothesis

$\Omega_3 : \theta_{ij}^{-1} = \mu + \alpha_i$

Here again μ is the general mean, α_i, the effect on θ_{ij}^{-1} due to the i^{th} level of the primary factor and the secondary factor is absent. This is equivalent to one-way

classification. The loglikelihood equation in this case is

$$l = -(1/2)IJn\ ln\ (\sigma) - (2\sigma)^{-1}\sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{k=1}^{n}(y_{ijk})^{-1}[y_{ijk}(\theta_{ij}^{-1}) - 1]^2$$

$$l = -(1/2)IJn\ ln\ (\sigma) - (2\sigma)^{-1}\sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{k=1}^{n}(y_{ijk})^{-1}[y_{ijk}(\mu + \alpha_i) - 1]^2$$

Equating to zero the first partial derivatives wrt μ and α_i of this equation respectively, we obtain

$$\begin{aligned} \hat{\mu}y_{..} + \sum_{i=1}^{I}\alpha_i y_{i.} &= IJn. \\ \hat{\mu}y_{i.} + \alpha_i y_{i.} &= Jn \end{aligned} \tag{4.55}$$

From (4.55) we obtain

$$\begin{aligned} y_{i.}[\hat{\mu} + \alpha_i] &= Jn \\ \hat{\mu} + \alpha_i &= Jn/y_{i.} = Jny_{i.}^{-1} = \bar{y_{i.}}^{-1} \end{aligned} \tag{4.56}$$

Thus

$\hat{\theta_{ij}}^{-1} = \bar{y_{i.}}^{-1}$

Substituting for $\hat{\theta_{ij}}^{-1}$ in the equation

$$\begin{aligned} IJn\hat{\sigma_3} &= R - n\sum_i\sum_j \hat{\theta_{ij}}^{-1} \\ IJn\hat{\sigma_3} &= R - n\sum_i\sum_j \bar{y_{i.}}^{-1} \\ &= R - nJ\sum_{i=1}^{I}\bar{y_{i.}}^{-1} \end{aligned} \tag{4.57}$$

4, Consider the following hypothesis Ω_4 where we consider the model

$\Omega_4 : \theta_{ij}^{-1} = \mu + \beta_{ij}$

Here again μ is the general effect or the general mean, β_{ij} is the effect due to the j^{th} level of the secondary factor nested in the i^{th} level of the primary factor. The loglikelihood equation in this case is

$$l = -(1/2)(IJn)\ ln\ (\sigma) - (2\sigma)^{-1} \sum_{i=1}^{I} \sum_{j=1}^{J} \sum_{k=1}^{n} (y_{ijk})^{-1} [y_{ijk}(\theta_{ij}^{-1}) - 1]^2 \qquad (4.58)$$

$$l = -(1/2)(IJn)\ ln\ (\sigma) - (2\sigma)^{-1} \sum_{i=1}^{I} \sum_{j=1}^{J} \sum_{k=1}^{n} (y_{ijk})^{-1} [y_{ijk}(\mu + \beta_{ij}) - 1]^2 \qquad (4.59)$$

Equating to zero the first partial derivatives wrt μ and β_{ij} we obtain the following equations

$$\hat{\mu} y_{ij} + \hat{\beta}_{ij} y_{ij} = n \qquad (4.60)$$

$$\hat{\mu} y_{..} + \sum_{i} \sum_{j} \beta_{ij} y_{ij} = IJn \qquad (4.61)$$

From (4.60) we obtain

$$\begin{aligned} \hat{\mu} + \hat{\beta}_{ij} y_{ij} &= n \\ \hat{\mu} + \hat{\beta}_{ij} &= n y_{ij}^{-1} = \bar{y_{ij}}^{-1} \end{aligned}$$

or

$$\hat{\theta_{ij}^{-1}} = \mu + \beta_{ij} = \bar{y_{ij}}^{-1} \qquad (4.62)$$

We should here note that the estimate of $\hat{\theta}_{ij}^{-1}$ under this model is the same as that under the model Ω_2. Thus the parameters of the models Ω_2 and Ω_4 cannot be distinguished. In practice it is thus of importance to test Ω_4 first. If Ω_3 is accepted, we can use one way approach for testing the main factor effects; otherwise analysis of individual cell means is of interest.

Thus we will consider testing Ω_2 and Ω_3 only. Let $\hat{\sigma}_s$ denote the MLE of σ and $l_{max}(\Omega_s)$, the maximised loglikelihood under the model Ω_s where $s = 1, 2, 3$.

In general it has been proved that for large n, $-2log\lambda$ has a χ^2 distribution. Here λ is the likelihood ratio statistic for testing a null hypothesis Ω_s nested within a model Ω_t.

Let $l_{max}(\Omega_s)$ represent the likelihood function under the hypothesis Ω_s, and let $l_{max}(\Omega_t)$ represent the likelihood function in the unrestricted case Ω_t.

We have

$$l_{max}(\Omega_s) = -(1/2)(IJn)ln(\hat{\sigma}_s) - (2\hat{\sigma}_s)^{-1}\sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{k=1}^{n}(y_{ijk})^{-1}[y_{ijk}(\mu + \alpha_i + \beta_{ij}) - 1]^2 \tag{4.63}$$

Substituting for $\hat{\sigma}_s$

$$\hat{\sigma}_s = (IJn)^{-1}\sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{k=1}^{n}(y_{ijk})^{-1}[y_{ijk}(\hat{\mu} + \hat{\alpha}_i + \hat{\beta}_{ij}) - 1]^2 \tag{4.64}$$

we obtain

$$\begin{aligned} l_{max}(\Omega_s) &= -(1/2)IJnln(\hat{\sigma}_s) - -1/2(IJn)\frac{\sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{k=1}^{n}(y_{ijk})^{-1}[y_{ijk}(\hat{\theta_{ij}}^{-1}) - 1]^2}{\sum_{i=1}^{I}\sum_{j=1}^{J}\sum_{k=1}^{n}(y_{ijk})^{-1}[y_{ijk}(\hat{\theta_{ij}}^{-1}) - 1]^2} \\ l_{max}(\Omega_s) &= -(1/2)IJnlog\hat{\sigma}_s - (1/2)IJn. \\ &= -IJn/2(log\hat{\sigma}_s + 1). \end{aligned} \tag{4.65}$$

Thus the maximised loglikelihood under each model Ω_s has the value

$-(1/2)IJn(log\hat{\sigma}_s + 1)$.

So we have

$$-2log\lambda = -2(l_{max}(\Omega_s) - l_{max}(\Omega_t))$$

$$= -2 - [(1/2)IJn[log\hat{\sigma}_s + 1] - (1/2)IJn(log\hat{\sigma}_t + 1)]$$

$$= 2 \times 1/2[IJn(log\hat{\sigma}_s + 1 - log\hat{\sigma}_t - 1]$$

$$= IJnlog\frac{\hat{\sigma}_s}{\hat{\sigma}_t}. \tag{4.66}$$

As calculated before we have the expressions for $\hat{\sigma}_s$ for $s = 1,2,3$, under the relevant model hypotheses Ω_1 Ω_2, Ω_3, as

$$IJn\hat{\sigma}_1 = R - n\sum_{i=1}^{I}\sum_{j=1}^{J}\bar{y_{ij}}^{-1}.$$

$$IJn\hat{\sigma}_2 = R - n\sum_{i=1}^{I}\sum_{j=1}^{J}\hat{\theta_{ij}}^{-1} = R - IJn\hat{\mu}.$$

$$IJn\hat{\sigma}_3 = R - nJ\sum_{i=1}^{I}\bar{y_{i.}}^{-1}.$$

where $\hat{\theta_{ij}}$ and μ are the efficient likelihood estimators developed in the previous section. When n is large one can perform the likelihood ratio test using the asymptotic distribution of the test statistic $-2log\lambda$. The rejection region of a level α test would be set as $-2log\lambda \geq \chi^2_\alpha$ where χ^2_α is the upper α point of χ^2 with degrees of freedom equating the number of parametric constants imposed on Ω_s. The results of the lemmas along with some exact sampling distributions stated at the end of the chapter establish the limiting χ^2 distribution of the LR statistic.

Now instead of resorting entirely to the asymptotic distribution of the test statistic, we can consider an alternate representation of the statistic. This would lead us to an analogue of the usual normal theory analysis of variance table. Also some modifications would provide a better approximation of the nominal significance level for moderate sample sizes.

Consider a new test statistic which we define as

$$T_{st} = IJn\frac{\hat{\sigma}_s - \hat{\sigma}_t}{\hat{\sigma}_t} \tag{4.67}$$

We now represent $2log\lambda = \Lambda_{st}$ in terms of T_{st}

$$T_{st} = IJn[\frac{\hat{\sigma}_s}{\hat{\sigma}_t} - 1]. \tag{4.68}$$

Taking log on both sides

$$\begin{aligned} logT_{st} &= logIJn[\frac{\hat{\sigma}_s}{\hat{\sigma}_t} - 1] \\ IJnlog[\frac{T_{st}}{IJn}] &= IJnlog((IJn/IJn)[\frac{\hat{\sigma}_s}{\hat{\sigma}_t} - 1]) \\ IJnlog[\frac{T_{st}}{IJn}] &= IJnlog[\frac{\hat{\sigma}_s}{\hat{\sigma}_t} - 1]) \\ IJnlog[\frac{T_{st}}{IJn} + 1] &= IJnlog[\frac{\hat{\sigma}_s}{\hat{\sigma}_t}] \\ \Lambda_{st} &= IJnlog[1 + T_{st}/IJn]. \end{aligned} \tag{4.69}$$

We can observe that Λ is a strictly increasing function of T_{st}. Consequently the LR statistic can be equivalently based on the test statistic T_{st} with large value in the rejection region. We will try to examine the individual test statistic by introducing some basic quantities which will be referred to as sum of reciprocals.

To construct an analogue of the usual analysis of variance like in the normal case, we now introduce certain basic quantities that will henceforth be referred as sum of reciprocals.

In order to test the effect of the j^{th} level of factor β or B nested within the i^{th} level of the primary factor A We consider the null hypothesis H_o that β_{ij} is not significant

against the alternative H_A that it is. To this end we define $R_{B(A)}$

$$\begin{aligned} R_{B(A)} &= IJn\hat{\sigma}_3 - IJn\hat{\sigma}_2 \\ &= R - nJ\sum_i \bar{y}_{i.}^{-1} - R + n\sum_i\sum_j \hat{\theta}_{ij}^{-1} \\ &= n\sum_i\sum_j \hat{\theta}_{ij}^{-1} - J\sum_i \bar{y}_{i.}^{-1} \\ &= n\sum_i\sum_j (\hat{\theta}_{ij}^{-1} - \bar{y}_{i.}^{-1}) \end{aligned} \tag{4.70}$$

$$\begin{aligned} R_E &= IJn\hat{\sigma}_4 \\ &= R - n\sum_i\sum_j \bar{y}_{ij}^{-1}. \end{aligned} \tag{4.71}$$

So we have

$$R_{B(A)} = n\sum_i\sum_j (\hat{\theta}_{ij}^{-1} - \bar{y}_{i.}^{-1}).$$

$$R_E = R - n\sum_i\sum_j \bar{y}_{ij}^{-1}$$

From the form of $l_{max}(\Omega_s)$ it is clear that $\hat{\sigma}_s > \hat{\sigma}_t$ for $s < t$. So the statistics given above are all non-negative.

We devise an approximate F test for the problem of testing the effect due to the nested factor β_{ij} i.e. the effect of the secondary factor j nested within the ith level of the primary factor α_i. Consider the test statistic T_{32} based on the likelihood ratio statistic Ω_{32}. We define

$$F_{B(A)} = \frac{(R_{B(A)}/IJ - I)}{(R_E/IJ(n-1))} \tag{4.72}$$

We have the following results on the null distribution of $F_{B(A)}$.

(a) R_E/σ has χ^2 distribution with $IJ(n-1)$ degrees of freedom ie is $\chi^2_{IJ(n-1)}$,

(b) R_E and $R_(B(A))$ are independent,

(c) $R_{B(A)}/\sigma$ is asymptotically distributed as χ^2_{I-1}.

The first property follows from the basic sampling property of the IG distribution stated in Chapter 1. However we need an additional adjustment to attain the exact independence of the components, stated in (b). Referring to T_{42} we define

$$F_{B(A)} = \frac{(R_{B(A)}/IJ-I)}{(R_E/IJ(n-1))} = \frac{n-1}{n(IJ-I)} T_{32}(\hat{\sigma}_2/\hat{\sigma}_1). \tag{4.73}$$

The lemma proved by Fries and Bhattacharyya [5] stated below shows that the adjustment factor $\hat{\sigma}_2/\hat{\sigma}_1$ rapidly approaches 1 in the sense that $n^r(\hat{\sigma}_2/\hat{\sigma}_1 - 1) \to 0$ in probability for any $r < 1$. Thus the use of this adjustment factor assures that $F_{B(A)}$ has properties (a)-(c).

Lemma 1. Under the nested model Ω_2, the asymptotic distribution of $IJn(\hat{\sigma}_3 - \hat{\sigma}_4)/\sigma$ is $\chi^2_(IJ-I)$. For proof see Fries and Bhattacharyya [5]

As stated earlier, the effect of the primary factor or factor A cannot be separated out in our model. This is because the models $\mu + \alpha_i + \beta_{ij}$ and $\mu + \beta_{ij}$ cannot be identified to be distinct. Hence we test for the effect due to the secondary factor β_{ij} and if it is found insignificant, then test for the primary factor effect α_i through the one way analysis of variance.

The following table presents the sums of reciprocals components associated with the various factor effects and the approximate F tests discussed earlier. The mean sum of reciprocals MR is defined as a sum of reciprocals divided by the corresponding degrees of freedom. The ANOR table has a striking resemblance with the normal

theory ANOVA table with the sum of reciprocals playing the role of sums of squares.

ANOR TABLE

source	sums of reciprcls	degrees of fdm	MR	appr F ratio
Factor B in A	$R_{B(A)}$	$IJ-I$	$MR_{B(A)}$	$MR_{B(A)}/MR_E$
residual	R_E	$IJ(n-1)$	MR_E	

4.6 Least Squares Estimation.

In the maximum likelihood method of estimation, we were not able to investigate the biases and the exact variances of the estimators since it involves the inverse of a random matrix. So here we try to develop the least squares approach. Our model is

$$\theta_{ij}^{-1} = \mu + \alpha_i + \beta_{ij} \tag{4.74}$$

where μ , α_i, β_{ij} have the same interpretations as before.

θ_{ij} is the mean of the observations y_{ijk} and we assume that $y_{ijk} \sim IG(\theta_{ij}, \sigma)$

The moment results stated in the beginning of the chapter give

$$E(\bar{y_{ij}}) = \mu + \alpha_i + \beta_{ij} + (\sigma/n).$$

$$Var(\bar{y_{ij}}) = (\mu + \alpha_i + \beta_{ij})\sigma/n + 2(\sigma/n)^2. \tag{4.75}$$

Let $y_{ij}^{-1} = r_{ij}$, then we have

$$E(r_{ij}) = \theta_{ij}^{-1} + \sigma/n. \tag{4.76}$$

Letting

$$\mu^* = \mu + \sigma/n \tag{4.77}$$

we have

$$E(r_{ij}) = \mu^* + \alpha_i + \beta_{ij}. \tag{4.78}$$

The following linear model results:

$$R = X\Psi^* + e \tag{4.79}$$

where $R = (r_{11}, r_{12}, \quad \ldots \ r_{IJ})'$

X is the matrix of appropriate constants, $\Psi^* = (\mu^*, \alpha', \beta')'$,

$e = (e_{11}, e_{12}, \ldots . \ e_{IJ})'$ is the error vector, $\mu^* = \mu + \sigma/n$, $\alpha' = (\alpha_1, \alpha_2, \ldots . \ \alpha_{I-1})$ and $\beta' = (\beta_{11}, \beta_{12}, \ldots . \ \beta_{IJ-1})$.

Given the model

$$r_{ij} = \mu^* + \alpha_i + \beta_{ij} + e_{ij} \tag{4.80}$$

where $E(e_{ij}) = 0$, we find the least squares estimators of $\mu^*, \alpha_i, \beta_{ij}$. . First we form the sum of squares of the errors: $e_{ij} = r_{ij} - \mu^* - \alpha_i - \beta_{ij}$ i.e.

$$\sum_i \sum_j e_{ij}^2 = \sum_i \sum_j (r_{ij} - \mu^* - \alpha_i - \beta_{ij})^2 \tag{4.81}$$

Differentiating with respect to μ^* we have

$$-2 \sum_i \sum_j (r_{ij} - \mu^* - \alpha_i - \beta_{ij}) = 0. \tag{4.82}$$

Differentiating with respect to α_i we have

$$-2 \sum_j (r_{ij} - \mu^* - \alpha_i - \beta_{ij}) = 0. \tag{4.83}$$

Differentiating with respect to β_{ij} we have

$$-2(r_{ij} - \mu^* - \alpha_i - \beta_{ij}) = 0. \tag{4.84}$$

From (4.82) we have

$$IJ\mu^* + J\sum_i \alpha_i + \sum_i \sum_j \beta_{ij} = r_{..}. \tag{4.85}$$

But since in our model we have assumed

$$\sum_i \alpha_i = \sum_i \sum_j \beta_{ij} = 0, \tag{4.86}$$

we have

$$\begin{aligned} IJ\mu^* &= r_{..}, \\ \mu^* &= r_{..}/IJ = \bar{r}_{..}. \end{aligned} \tag{4.87}$$

From (4.83) we have

$$\begin{aligned} J\mu^* + J\alpha_i &= r_{i.}, \\ \hat{\alpha}_i &= r_{i.}/J - \mu^*, \\ &= \bar{r}_{i.} - \bar{r}_{..}. \end{aligned} \tag{4.88}$$

From (4.84) we have

$$\begin{aligned} \mu^* + \alpha_i + \beta_{ij} &= r_{ij}, \\ \beta_{ij} &= r_{ij} - \alpha_i - \mu^*, \\ \hat{\beta}_{ij} &= r_{ij} - \bar{r}_{..} - \bar{r}_{i.} + \bar{r}_{..}, \\ &= (r_{ij} - \bar{r}_{i.}). \end{aligned} \tag{4.89}$$

So we have the least squares estimates of $\mu^*, \alpha_i, \beta ij$ respectively as

$$\hat{\mu}^* = \bar{r}_{..},$$

$$\hat{\alpha}_i = \bar{r}_{i.} - r_{..},$$

$$\hat{\beta}_{ij} = (r_{ij} - r_{i..}$$

$\hat{\sigma}$ is obtained from the previous section as MR_E i.e.

$$\hat{\sigma} = \frac{\sum_i \sum_j \sum_k y_{ijk}^{-1} - n \sum_i \sum_j y_{ij}^{-1}}{IJ(n-1)}. \tag{4.90}$$

So we obtain

$$\hat{\mu} = \bar{r}_{..} - (\hat{\sigma}/n). \tag{4.91}$$

Theorem 2. The asympototic distributions of $n^{1/2}(\hat{\Phi} - \Phi)$ and $n^{1/2}(\hat{\sigma} - \sigma)$ under least squares estimation are $N_{IJ}(0, \sigma\Gamma_{-1})$ and $N_1(0, 2\sigma^2(IJ)^{-1})$ respectively where

$$\Gamma_1 = V'\Theta^{-1}V, \quad V' = (X'X)^{-1}X. \tag{4.92}$$

Also $\hat{\Phi}$ and $\hat{\sigma}$ are asymptotically independent.

The least squares estimates cannot be used for testing the effect of the various factors and this is where the simplicity of the least squares ends. This approach is based on initial reduction of the data by sufficiency. It provides unbiased estimators as well as exact expressions for their variances and covariances. The weighted least squares with estimated weights offers a convenient and useful modification. (see sec 4.7)

4.6.1 Estimation of Variances and Covariances

Since the $r'_{ij}s$ are independent of each other and σ, the exact variances and covariances of these estimators are readily calculated.

1. Variance of $\hat{\mu}$.

$$\begin{aligned} Var(\hat{\mu}) &= Var(\bar{r}_{..} - \sigma/n), \\ &= Var(\bar{r}_{..}) + Var(\sigma/n), \\ &= Var\frac{\sum_i \sum_j r_{ij}}{IJ} + (1/n^2)Var\hat{\sigma}, \\ &= \frac{1}{I^2J^2}Var\sum_i \sum_j r_{ij} + \frac{1}{n^2}Var\hat{\sigma}, \end{aligned} \tag{4.93}$$

But we know that $Var(r_{ij}) = (\mu + \alpha_i + \beta_{ij})\sigma/n + 2(\sigma/n)^2$, Therefore we can write

$$\begin{aligned} Var(\hat{\mu}) &= \frac{1}{I^2J^2}\sum_i \sum_j [(\mu + \alpha_i + \beta_{ij})(\sigma/n) + 2(\sigma/n)^2] + 1/n^2(\sigma^2/IJ(n-1)), \\ &= \frac{1}{I^2J^2}[(IJ\mu + J0 + 0)\sigma/n + 2IJ(\sigma/n)^2] + 1/n^2(\sigma^2/IJ(n-1)). \\ &= \frac{\mu\sigma}{IJn} + 2/IJ(\sigma/n)^2 + 1/n^2(\sigma^2/IJ(n-1)), \\ &+ \frac{\mu\sigma}{IJn} + \frac{2\sigma^2}{n^2IJ}(n-1+1)/n - 1, \\ &= \frac{\mu\sigma}{IJn} + \frac{2\sigma^2}{n^2IJ(n-1)}, \\ &= \sigma(IJn)^{-1}[\mu + 2\sigma/n - 1]. \end{aligned} \tag{4.94}$$

Thus $Var(\mu) = \sigma/IJn[\mu + 2\sigma/n - 1]$.

2. Variance of $\hat{\alpha}_i$

$$\begin{aligned} Var(\hat{\alpha}_i) &= Var(\bar{r}_{i.} - \bar{r}_{..}), \\ &= Var(\frac{\sum_j r_{ij}}{J} - \frac{\sum_i \sum_j r_{ij}}{IJ}), \\ &= Var(\frac{\sum_j r_{ij}}{J}) - Var(\frac{\sum_i \sum_j r_{ij}}{IJ}) \\ &\quad -2Cov(\frac{\sum_j r_{ij}}{J}\frac{\sum_i \sum_j r_{ij}}{IJ}), \\ &= \frac{\mu\sigma}{IJn} + \frac{2\sigma^2}{n^2IJ} + 1/J^2[\sum_j(\mu + \alpha_i + \beta_{ij})(\sigma/n) \end{aligned}$$

$$
\begin{aligned}
& +2(\sigma/n)^2] - \frac{2\mu\sigma}{IJn} - \frac{2\alpha_i\sigma}{IJn} - \frac{4\sigma^2}{IJn^2}, \\
= \; & \frac{\mu\sigma}{IJn} + \frac{2\sigma^2}{IJn^2} + \frac{\mu\sigma}{Jn} + \frac{\alpha_i\sigma}{Jn} \\
& + \frac{2\sigma^2}{Jn^2} - \frac{2\mu\sigma}{IJn} - \frac{2\alpha_i\sigma}{IJn} - \frac{4\sigma^2}{IJn^2}, \\
= \; & \frac{\mu\sigma}{Jn} + \frac{\alpha_i\sigma}{Jn} + \frac{2\sigma^2}{Jn^2} - \frac{\mu\sigma}{IJn}, \\
& - \frac{2\alpha_i\sigma}{IJn} - \frac{4\sigma^2}{IJn^2}, \\
= \; & \frac{\sigma}{IJn}[(I-1)\mu + (I-2)\alpha_i + (I-1)\sigma/n]. \qquad (4.95)
\end{aligned}
$$

3. Variance of $\hat{\sigma}$

$$
Var(\hat{\sigma}) = Var[\frac{\sum_i \sum_j \sum_k y_{ijk}^{-1} - n \sum_i \sum_j y_{ij}^{-1}}{IJ(n-1)}]. \qquad (4.96)
$$

Multiply and divide by σ.

$$
\begin{aligned}
Var(\hat{\sigma}) &= Var\sigma \left[\frac{\sum_i \sum_j \sum_k y_{ijk}^{-1} - n \sum_i \sum_j y_{ij}^{-1}}{IJ(n-1)\sigma} \right] \\
&= \frac{\sigma^2}{I^2J^2(n-1)^2} Var(X) \text{ where } X \sim \chi^2_{IJ(n-1)}
\end{aligned}
$$

We know that $Var(X) = 2IJ(n-1)$. Hence

$$
\begin{aligned}
Var(\hat{\sigma}) &= \frac{\sigma^2}{I^2J^2(n-1)^2} 2IJ(n-1), \\
&= \frac{2\sigma^2}{IJ(n-1)}. \qquad (4.97)
\end{aligned}
$$

Thus the variance of $\hat{\sigma}$ is $2\sigma^2/IJ(n-1)$.

4. Covariance between $\hat{\mu}$ and $\hat{\sigma}$

$$
\begin{aligned}
Cov(\hat{\mu}, \hat{\sigma}) &= Cov(\bar{r}_{..} - \hat{\sigma}/n)(\hat{\sigma}) \\
&= -1/nVar(\hat{\sigma})
\end{aligned}
$$

Since r_{ij} and σ are independent. From the property proved above we know that $Var(\hat{\sigma}) = \frac{2\sigma^2}{IJ(n-1)}$.

Substituting we have

$$Cov(\hat{\mu}, \hat{\sigma}) = -1/n[\frac{2\sigma^2}{IJ(n-1)}] \tag{4.98}$$

Thus we have

$Cov(\hat{\mu}, \hat{\sigma}) = -[\frac{2\sigma^2}{IJn(n-1)}]$

5. Covariance between $\hat{\mu}$ and $\hat{\alpha}_i$

$$\begin{aligned}
Cov(\hat{\mu}, \hat{\alpha}_i) &= Cov\,[(\bar{r}_{..} - \hat{\sigma}/n)(\bar{r}_{i.} - \bar{r}_{..})] \\
&= -Var(\bar{r}_{..}) + Cov[(\bar{r}_{..})(\bar{r}_{i.})], \\
&= -[\frac{1}{I^2J^2}(IJ\mu)\sigma/n + \frac{2}{IJ}(\sigma/n)^2] + Cov[\sum\sum r_{ij}/IJ,\ \sum_j r_{ij}/J], \\
&= -\frac{\mu\sigma}{IJn} - \frac{2\sigma^2}{IJn^2} + J/IJ^2[\mu\sigma/n + \alpha - i\sigma/n] + \frac{2\sigma^2}{IJn^2}, \\
&= \frac{\alpha_i\sigma}{IJn}, \\
&= \alpha_i\sigma(IJn)^{-1}.
\end{aligned} \tag{4.99}$$

Thus the covariance between $\hat{\mu}$ and $\hat{\alpha}_i$ is $\alpha_i\sigma(IJn)^{-1}$

6. Covariance between $\hat{\alpha}_i$ and $\hat{\alpha}_{i'}$

$$\begin{aligned}
Cov(\hat{\alpha}_i, \hat{\alpha}_{i'}) &= Cov[(\bar{r}_{i.} - \bar{r}_{..}), (\bar{r}_{i'.} - \bar{r}_{..})] \\
&= Cov[(\bar{r}_{i.})(\bar{r}_{i'.})] + Var(\bar{r}_{..}) - Cov[(\bar{r}_{i.})(\bar{r}_{..})] \\
&\quad -Cov[(\bar{r}_{i'.})(\bar{r}_{..})]
\end{aligned}$$

The first term vanishes since it is zero. Hence

$$Cov(\hat{\alpha}_i, \hat{\alpha}_{i'}) = \frac{\mu\sigma}{IJn} + \frac{2\sigma^2}{IJn^2} - \frac{\mu\sigma}{IJn}$$

$$-\frac{\alpha_i\sigma}{IJn} \quad -\frac{2\sigma^2}{IJn} - \frac{\mu\sigma}{IJn} - \frac{\alpha_{i'}\sigma}{IJn} - \frac{2\sigma^2}{IJn}$$

$$= \frac{\sigma}{IJn}[\mu + \alpha_i + \alpha_{i'} + 2\sigma/n] \tag{4.100}$$

Thus covariance between $\hat{\alpha}_i$ and $\hat{\alpha}_{i'}$ is $\frac{\sigma}{IJn}[\mu + \alpha_i + \alpha_{i'} + 2\sigma/n]$.

7. Variance of $\hat{\beta}_{ij}$

$$\begin{aligned} Var(\hat{\beta}_{ij}) &= Var(r_{ij} - \bar{r}_{..}), \\ &= Varr_{ij} + Var(\bar{r}_{..}) - 2\left[Covr_{ij}\sum_i\sum_j r_{ij}/IJ\right], \\ &= (\mu + \alpha_i + \beta_{ij})\sigma/n + 2\sigma^2/n^2 + \frac{\mu\sigma}{IJn} + \frac{2\sigma^2}{n^2IJ} \\ &\quad -\frac{2}{IJ}[(\mu + \alpha_i + \beta_{ij})\sigma/n + \frac{2\sigma^2}{n^2}], \\ &= \frac{\sigma}{IJn}\left[(IJ-1)\mu + (IJ-2)\alpha_i + (IJ-2)\beta_{ij} - (IJ-1)\frac{2\sigma}{n}\right] \end{aligned} \tag{4.101}$$

Thus variance of $\hat{\beta}_{ij}$ is $-\frac{\sigma}{IJn}[(IJ-1)\mu + (IJ-2)\alpha_i + (IJ-2)\beta_{ij} - (IJ-1)\frac{2\sigma}{n}]$.

4.7 Weighted Least Squares

We have the model (4.79) where the elements of R are independent, but the variances are not constants, especially if α_i and β_{ij} are significant, since

$$Var(\bar{y_{ij}}) = (\mu + \alpha_i + \beta_{ij})\sigma/n + 2(\sigma/n)^2. \tag{4.102}$$

Hence we can apply the weighted least squares method of estimation in this situation. Here we assume the weights w_{ij}^{-1} as

$$w_{ij}^{-1} = \hat{\theta}_{ij}^{-1} + 2\hat{\sigma}/n. \tag{4.103}$$

where $\hat{\theta}_{ij}$ and $\hat{\sigma}$ are the estimated values from either the maximum likelihood or the least squares method of estimation.

Thus applying the weighted least squares method dicussed earlier in section(2.5) of chapter(2) to the equation

$$R = X\Psi + \epsilon. \tag{4.104}$$

with the weights (4.103) we obtain the weighted least squares estimate of Ψ as

$$\hat{\Psi}_w = (X'WX)^{-1}X'WR. \tag{4.105}$$

We can observe here that if we use weights only upto the order $1/n$ i.e.

$w_{ij} = \theta_{ij}$

we get the same estimate of Ψ as that in the case of maximum likelihood estimation. This will be illustrated in the numerical examples solved in the next chapter.

Chapter 5

Numerical Examples

5.1 Example 1

As an example of nested classification, we consider a study of three different sprays used on trees. Each of the three sprays was applied to four trees. After one week the concentration of nitrogen was measured in each of six leaves picked in a random way from each tree. Here the experimental units are the four trees; we consider them to have been chosen at random from a large (∞) population of trees. The leaves thus form four subsamples, each consisting of six leaves from a large (∞) population of leaves on the particular tree. The following table gives the data . Each measurement of nitrogen concentration is denoted by y_{ijk}

	TREE 1	TREE 2	TREE 3	TREE 4
SPRAY 1	4.50	5.78	13.32	11.59
	15.32	14.53	10.89	15.12
	7.18	6.70	5.94	4.00
	7.04	7.69	15.05	8.96
	14.97	14.51	10.27	13.79
	7.98	8.28	5.78	5.46
SPRAY 2	4.98	12.68	12.67	10.95
	14.81	12.61	12.21	15.32
	5.51	6.99	7.59	5.40
	5.48	5.89	12.42	9.87
	14.26	16.13	12.77	11.95
	7.48	6.40	7.21	6.85
SPRAY 3	6.54	4.07	10.03	10.48
	15.58	13.65	10.45	12.56
	7.55	4.96	6.12	7.74
	7.20	4.08	13.50	12.79
	16.01	14.78	11.44	15.31
	5.64	7.03	7.13	6.81

The cell means $\bar{y}_{ij}$ are given in the following table:

	tree 1	tree 2	tree 3	tree 4
spray1	5.956672	6.698328	12.831701	10.773310
spray 2	15.208433	14.368437	11.338383	14.008350
spray 3	6.889995	6.726668	6.628355	6.056679

where i is the number of the treatment and runs from 1 to I or 1 to 3 here.

j is the number of the experimental unit within the ith treatment and runs from 1 to J or 1 to 4 in this case.

k is the sample number of the unit (leaf number) within the $(i,j)^{th}$ cell, which runs from 1 to n or 1 to 6 in this case.

We assume that $y_{ijk} \sim IG(\theta_{ij}, \sigma)$ and with the reciprocal linear model for the mean as

$$\theta_{ij}^{-1} = \mu + \alpha_i + \beta_{ij}. \tag{5.1}$$

where i = 1,2,3 and j = 1,2,3,4 ie I=3 and J=4 In the next section we estimate the parameters $\hat{\theta}_{ij}$ and $\hat{\sigma}$ perform the analysis using the three methods of estimation.

5.1.1 Estimates of the Parameters

The computer program developed to estimate the parameters attached to the appendix was used for the purpose. The algorithm is simple and involves multiplication of matrices and finding their inverses. The estimates of μ, α, β using maximum likelihood estimation was found to be

$\hat{\mu} = 0.116053$

$\hat{\alpha}_1 = 0.005928, \hat{\alpha}_2 = -0.042319$

$\hat{\beta}_{11} = 0.045898, \hat{\beta}_{12} = 0.027310, \hat{\beta}_{13} = -0.044049$

$\hat{\beta}_{21} = -0.007980, \hat{\beta}_{22} = -0.004136, \hat{\beta}_{23} = 0.014463$

$\hat{\beta}_{31} = -0.007306, \hat{\beta}_{32} = -0.003782, \hat{\beta}_{33} = -0.001576$

The estimates of θ_{ij}^{-1} was found to be

$\hat{\theta}_{11}^{-1} = 0.167879, \hat{\theta}_{12}^{-1} = 0.149291 \hat{\theta}_{13}^{-1} = 0.077932 \hat{\theta}_{14}^{-1} = 0.092822$

$\hat{\theta}_{21}^{-1} = 0.065753, \hat{\theta}_{22}^{-1} = 0.069597, \hat{\theta}_{23}^{-1} = 0.088196 \hat{\theta}_{24}^{-1} = 0.071386$

$\hat{\theta}_{31}^{-1} = 0.145138, \hat{\theta}_{32}^{-1} = 0.148662, \hat{\theta}_{33}^{-1} = 0.150867, \hat{\theta}_{34}^{-1} = 0.165107.$

The estimate of $\hat{\sigma}$ is found to be 0.004979

Least squares method of estimation gives us the following estimates for $\hat{\mu}, \hat{\alpha}, \beta$

$\hat{\mu} = 0.116053$

$\hat{\alpha}_1 = 0.005928, \hat{\alpha}_2 = -0.042319, \hat{\alpha}_3 == 0.036391$

$\hat{\beta}_{11} = 0.045898, \ \hat{\beta}_{12} = 0.027310, \hat{\beta}_{13} = -0.044049, \hat{\beta}_{14} = -0.029159.$

$\hat{\beta}_{21} = -0.007980, \hat{\beta}_{23} = -0.004136, \hat{\beta}_{23} = 0.014463, \hat{\beta}_{24} = -0.002347.$

$\hat{\beta}_{31} = -0.007306, \hat{\beta}_{32} = -0.003782, \hat{\beta}_{33} = -0.001576, \hat{\beta}_{34} = 0.012664.$

The estimates of θ_{ij}^{-1} was found to be

$\hat{\theta}_{11}^{-1} = 0.167049, \hat{\theta}_{12}^{-1} = 0.148461, \hat{\theta}_{13}^{-1} = 0.077102, \hat{\theta}_{14}^{-1} = 0.091992$

$\hat{\theta}_{21}^{-1} = 0.064924, \hat{\theta}_{22} =^{-1} 0.068768, \hat{\theta}_{23}^{-1} = 0.087367, \hat{\theta}_{24}^{-1} = 0.070556.$

$\hat{\theta}_{31}^{-1} = 0.144308, \hat{\theta}_{32}^{-1} = 0.147832, \hat{\theta}_{33}^{-1} = 0.150038, \hat{\theta}_{34}^{-1} = 0.164277.$

In the weighted least squares we use weights

(i) $wij = \hat{\theta_{ij}} = \bar{y}_{ij}$

(2) $wij^{-1} = \theta_{ij}^{-1} + 2\sigma/n$

In the first case we get exactly the same estimates as in the case of maximum likelihood. In the second case, they differ slightly. The estimates of μ, α, β using

weighted least squares method of estimation was found to be

$\hat{\mu} = 0.116052$

$\hat{\alpha}_1 = 0.005929, \hat{\alpha}_2 = -0.042320$

$\hat{\beta}_{11} = 0.045898, \hat{\beta}_{12} = 0.027310, \hat{\beta}_{13} = -0.044049$

$\hat{\beta}_{21} = -0.007980, \hat{\beta}_{22} = -0.004136, \hat{\beta}_{23} = 0.014463$

$\hat{\beta}_{31} = -0.007306, \hat{\beta}_{32} = -0.003781, \hat{\beta}_{33} = -0.001576$

The estimates of θ_{ij}^{-1} was found to be

$\hat{\theta}_{11}^{-1} = 0.1679, \hat{\theta}_{12}^{-1} = 0.1493, \hat{\theta}_{13}^{-1} = 0.0779, \hat{\theta}_{14}^{-1} = 0.0928$

$\hat{\theta}_{21}^{-1} = 0.0658, \hat{\theta}_{22}^{-1} = 0.06956, \hat{\theta}_{23}^{-1} = 0.0882, \hat{\theta}_{24}^{-1} = 0.0714$

$\hat{\theta}_{31}^{-1} = 0.1451, \hat{\theta}_{32}^{-1} = 0.1487, \hat{\theta}_{33}^{-1} = 0.1509, \hat{\theta}_{34}^{-1} = 0.1651.$

5.1.2 MLE, LSE and WLSE of the mean

$\hat{\theta}_{ij}$	MLE	LSE	WLSE1	WLSE2
$\hat{\theta}_{11}$	5.956667	5.9862575	5.596667	5.9559261
$\hat{\theta}_{12}$	6.698335	6.7357757	6.698335	6.6889632
$\hat{\theta}_{13}$	12.831668	12.969832	12.831668	12.83697
$\hat{\theta}_{14}$	10.773335	10.87051	10.773335	10.775962
$\hat{\theta}_{21}$	15.208334	15.402625	15.208334	15.197568
$\hat{\theta}_{22}$	14.368334	14.541647	14.368334	14.376078
$\hat{\theta}_{23}$	11.338334	11.445969	11.338334	11.337868
$\hat{\theta}_{24}$	14.008333	14.173139	14.008333	14.005602
$\hat{\theta}_{31}$	6.890000	6.9296228	6.890000	6.68917988
$\hat{\theta}_{32}$	6.726666	6.7644353	6.726666	6.7249496
$\hat{\theta}_{33}$	6.628334	6.6649782	6.628334	6.6269052
$\hat{\theta}_{34}$	6.056666	6.0872794	6.056666	6.60569352

It can be observed above that WLSE1 for which weights are $w_{ij} = \theta_{ij}$, the estimates are the same as that of the MLE's.

5.1.3 Test of Hypotheses

The analysis of reciprocals using maximum likelihood estimation is summarised in the following table:

ANOR TABLE

source of variation	sums of reciprcls	deg of fdm	MR	appr F ratio
var of sprays within trees	0.309	9	0.0343333	6.8956
residual	0.298740	60	0.004979	

It can be seen from the above table that since

$F = 6.89 > F_{.05(9,60)} = 3.01$

we reject the hypothesis $H_o : (\beta_{ij} = 0 \quad i = 1,2,3 \quad j = 1,2,3,4)$ that the effect of the j^{th} level of the secondary factor nested within the i^{th} level of primary factor β_{ij} is insignificant. ie. there is a significant effect on the mean nitrogen concentration in the leaves due to the variation among the trees within the sprays.

5.2 Example 2

The next example we consider to illustrate the above method is an experiment that was conducted to compare three varieties of wheat. In each of three villages A,B and C, three plots were sown with the same variety. The following table gives the yield per acre y_{ijk} of wheat in appropriate units.

	VILLAGE 1	VILLAGE 2	VILLAGE 3
WHEAT 1	46	41	35
	19	11	31
	23	17	27
WHEAT 2	37	49	45
	18	35	66
	18	35	61
	20	43	75
WHEAT 3	32	65	34
	43	67	66
	32	58	58
	18	74	67

The cell means $\bar{y}_{ij}$ is given in the following table

	1	2	3
1	30.75	23.5	28.5
2	23.25	41	61.5
3	31.25	66	56.25

We assume that $y_{ijk} \sim IG(\theta_{ij}, \sigma)$ and analyse the data for nested factors. The three different varieties of wheat form the primary factor or the primary effect. And the secondary factor (village factor) is nested within each level of the primary factor. So here we have

i from 1 to 3,

j from 1 to 3, k from 1 to 4

and n the number of plots sown in each village is 4. $I = 3$, $J = 3$

5.2.1 Parameter Estimation

Again in this case also the data has been run on the programs listed in the appendix to get the following results. The estimates of μ, α, β using maximum likelihood estimation was found to be

$\hat{\mu} = 0..028750$

$\hat{\alpha}_1 = 0.007970$, $\hat{\alpha}_2 = -0.000863$

$\hat{\beta}_{11} = -0.004200$, $\hat{\beta}_{12} = 0.005833$,

$\hat{\beta}_{21} = 0.015124$, $\hat{\beta}_{22} = -0.003497$

$\hat{\beta}_{31} = 0.010357$, $\hat{\beta}_{32} = -0.006492$

The estimates of θ_{ij}^{-1} was found to be

$\hat{\theta}_{11}^{-1} = 0.032520, \hat{\theta}_{12}^{-1} = 0.042553$, $\hat{\theta}_{13}^{-1} = 0.035088$

$\hat{\theta}_{21}^{-1} = 0.043011$, $\hat{\theta}_{22}^{-1} = 0.024390$, $\hat{\theta}_{23}^{-1} = 0.016260$

$\hat{\theta}_{31}^{-1} = 0.03200$, $\hat{\theta}_{32}^{-1} = 0.015152$, $\hat{\theta}_{33}^{-1} = 0.017778$

The estimate of $\hat{\sigma}$ is found to be 0.00

Least squares method of estimation gives us the following estimates for μ, α, β

$mu = 0.028750$

$\hat{\alpha}_1 = 0.07970$, $\hat{\alpha}_2 = -0.000863$, $\hat{\alpha}_3 = -0.007107$

$\hat{\beta}_{11} = -0.00420$, $\hat{\beta}_{12} = 0.005833$, $\hat{\beta}_{13} = -0.001633$

$\hat{\beta}_{21} = 0.015124$, $\hat{\beta}_{23} = -0.003497$, $\hat{\beta}_{23} = -0.010357$

$\hat{\beta}_{31} = 0.010357$, $\hat{\beta}_{32} = -0.006492$, $\hat{\beta}_{33} = -0.003865$

The estimates of θ_{ij}^{-1} was found to be

$\hat{\theta}_{11}^{-1} = 0.031690, \quad \hat{\theta}_{12}^{-1} = 0.041723, \quad \hat{\theta}_{13} = 0.034250$

$\hat{\theta}_{21}^{-1} = 0.042181, \quad \hat{\theta}_{22}^{-1} = 0.023560, \quad \hat{\theta}_{23}^{-1} = 0.015430$

$\hat{\theta}_{31}^{-1} = 0.031170, \quad \hat{\theta}_{32}^{-1} = 0.014322, \quad \hat{\theta}_{33}^{-1} = 0.016948$

The estimates of μ, α, β using weighted Least squares method of estimation was found to be

$\hat{\mu} = 0.028750$

$\hat{\alpha}_1 = 0.007970, \quad \hat{\alpha}_2 = -0.000863$

$\hat{\beta}_{11} = -0.00420, \quad \hat{\beta}_{12} = 0.005833$

$\hat{\beta}_{21} = 0.015124, \quad \hat{\beta}_{22} = -0.003497$

$\hat{\beta}_{31} = 0.010357, \quad \hat{\beta}_{32} = -0.006491$

The estimates of θ_{ij}^{-1} was found to be

$\hat{\theta}_{11}^{-1} = 00.0325, \quad \hat{\theta}_{12}^{-1} = .0426, \quad \hat{\theta}_{13}^{-1} = 0.0351$

$\hat{\theta}_{21}^{-1} = 0.0430, \quad \hat{\theta}_{22}^{-1} = 0.0244, \quad \hat{\theta}_{23}^{-1} = 0.0163$

$\hat{\theta}_{31}^{-1} = 0.0320, \quad \hat{\theta}_{32}^{-1} = 0.0152, \quad \hat{\theta}_{33}^{-1} = 0.0178$

5.2.2 Comparison of MLE, LSE and WLSE of the mean

The MLE, LSE and WLSE of the mean θ_{ij} is given in the following table.

$\hat{\theta}_{ij}$	MLE	LSE	WLSE1	WLSE2
$\hat{\theta}_{11}$	30.749998	31.55569	30.749998	30.769230
$\hat{\theta}_{12}$	23.49999	23.96759	23.49999	23.4741784
$\hat{\theta}_{13}$	28.50000	29.19708	28.5000	28.4900249
$\hat{\theta}_{21}$	23.250002	23.707356	23.250002	23.255813
$\hat{\theta}_{22}$	40.9999992	42.44482	40.9999992	40.9836063
$\hat{\theta}_{23}$	61.499999	64.806814	61.499999	61.34969325
$\hat{\theta}_{31}$	31.24999	32.08213	31.24999	31.25
$\hat{\theta}_{32}$	65.99998	69.82265	65.99998	65.78947368
$\hat{\theta}_{33}$	56.249999	59.00401	56.249999	56.1797752

5.2.3 Test of Hypotheses

The analysis of reciprocals using maximum likelihood estimation is summarised in the following table:

ANOR TABLE.

source of variation	sums of reciprcls	deg of fdm	MR	appr F ratio
nested village factor	0.169161	6	0.0281935	7.190385
residual	0.105867	27	0.003921	

It can be seen from the above table that since

$F = 7.19 > F_{.05(6,24)} = 4.10$

we reject the hypothesis $H_o : (\beta_{ij} = 0 \quad i = 1,2,3 \quad j = 1,2,3,4)$ that the effect of the j^{th} level of the secondary factor nested within the i^{th} level of primary factor β_{ij} is insignificant. ie. there is a significant effect on the mean yield of wheat/acre

beacause of the village the variety of wheat is grown in.

The Appendix gives the program code using C for estimating the parameters using the three methods of estimation.

Bibliography

[1] Ahmad.M, Chaubey, Y.P and Sinha, B.K. (1991) Estimation of a common mean of several Univariate inverse Gaussian Populations, To appear in *Annals of Inst of Statistical Mathematics*, **31**.

[2] Bachelier, L (1900) Theorie de la Speculation, *Ann. Sci. Ec. Norm. Super*, 17. 21-86

[3] Banarjee, A. K. and Bhattacharyya, G. K. (1976). A Purchase Incidence Model with Inverse Gaussian Interpurchase Times, *Journal of American Statist. Association.*, **71**, 823-829.

[4] Bhattacharyya, G.K. and Fries, A. (1982) Inverse Gaussian Regression and accelerated life Tests, in *Proceedings of the special Topics Meeting on Survival Analysis,* 138 th meeting of the institute of Mathematical Statistics, Columbus, Ohio, Oct. 1981, 101-117.

[5] Fries, A. and Bhattacharyya, G. K. (1983) Analysis of two factor experiments under an Inverse Gaussian Model, *J. of the American Statistical Association* **78(384)**, 820-826.

[6] Bhattacharyya, G.K. and Fries, A. (1986) On the Inverse Gaussian Multiple Regression and model checking procedures, *Reliability and Quality Control*, 87-100.

[7] Banerjee, A. et al (1983) Bayesian results for the inverse Gaussian distribution with application, *Technometrics*, **21**, 823-829.

[8] Bhattacharya, G. et al (1982), Fatigue failure models: Birbaum-Saunders vs. Inverse Gaussian, *IEEE Trans. Rel.* , **31**, 439-441.

[9] Bravo, G. and MacGibbon, B. (1988), Improved Estimation for the parameters of in Inverse Gaussian Distribution, *Comm. Statis., Theory and Methods*, **17 (12)**, 4285-4299.

[10] Chikkara, R. S. and Folks, J.L. (1989), *The Inverse Gaussian Distribution*, Marcel Dekker, New York.

[11] Chikkara, R. S. (1972), Statistival inference related to the Inverse Gaussian distribution, Ph.D, Dissertation, Oklahama State University.

[12] Chhikara. R. S. and Folks. J. L. (1974) Estimation of the Inverse Gaussian distribution, *J Amer. Statist. Ass.*, **69**, 250-254.

[13] Chhikara. R. S. and Folks. J. L. (1975) Statistical distributions related to the inverse Gaussian distribution, function. *Communications in Statistics*, **4**, 1081-1091.

[14] Chhikara. R. S. and Folks. J. L. (1976) Optimum test procedures for the mean of the first passage time distributions in Brownian motion with positive drift, *Technometrics,* **18**, 189-193.

[15] Chhikara. R. S. and Folks. J. L. (1974) The Inverse Gaussian Distribution as a lifetime model, *Technometrics*, **19**, 461-468.

[16] Chhikara. R. S. and Folks. J. L. (1978) The Inverse Gaussian Distribution and its statistical application- A Review, *J. R. Statist. Soc. B*, **40**, no 5, pp 263-289.

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

[18] Chhikara. R. et al (1982), Prediction limits for the Inverse Gaussian Distribution, *Technometrics*, **21**, 319-324.

[19] Cox, D. R. and Miller, H. D. (1965), *The Theory of Stochastic Processes*, London, Methuen.

[20] Cramer, H. (1946) *Mathematical methods in Statistics*, Princeton University Press.

[21] Davis, A. S. (1977) Linear Statistical Inference as related to Inverse Gaussian distribution, Ph.D. thesis, Dept of Statistic, Oklahoma State University.

[22] Dugue, D. (1937), Application des propprietes de la limite au sens du calcul des probabilites a l etude des diverses questions d' estimation. *J. Ecol. Poly.*, **3**, no. 4, 305.

[23] Eaton, W. W and Whitmore, G. A.(1977), Length of stay as a stochastic process: a general approach and application to hospitalisation for schizophrema, *J. Math. Sociol*, **5**, 273-292.

[24] Frienberg, S. (1974) Single neuron in the brain: A Survey, *Biometrika*, **30**, 399-427.

[25] Fries, A. (1982) Analysis of factorial experiments and accelerated life testsunder an Inverse Gaussian model, Unpublished Ph.D thesis, University of Wisconsin, Department of Statistics.

[26] Gerstein. G. et al. (1964), Random walk models for the spike activity of a neuron, *Biophys. Jour.*, **4**, 41-68.

[27] Hasofer, A. M. (1964), A Dam with Inverse Gaussian Input, *Proc. Camb. Phil. Soc.*, **60**, 931-933.

[28] Hsiech, H. K.. Korwar, R.M. and Rukhin, A. L (1990) Inadmissibility of the Maximum Likelihood Estimate of the Inverse Gaussian mean, *Stat. Prob. Lett.*, **9**, no1, 83-90.

[29] Huzurbazar, V. S. (1948) The likelihood equation, consistency, and maxima of the likelihood function. *Ann. Eugen. Lomdon*, **14**, 185.

[30] Khatri C. G. (1962) A characterisation of the Inverse Gaussian Distribution, *Annals of Math. Statist.*, **33**, 800-803.

[31] Lancaster, A. A stochastic model for the duration of a strike, *J. Roy. Statist. Soc. Ser. A.*, **135**, 257-271, 1972.

[32] Letac, G. Seshadri, V. and Whitmore, G. A. (1985). An exact Chi-squared Decomposition Theorem for Inverse Gaussian Variates, *J. R. Statist. Soc. B.*, **47**, 476-481.

[33] Marcus, Allan H. (1975) Some exact distributions in traffic noise theory, *Adv. in Appl. Probab.*, **7**, 593-606.

[34] Moran, P. A. P. (1968), *An introduction to Probability Theory*, Oxford Clarendon Press.

[35] Nadas, A. (1973) Best tests for zero drift based on first passage times in Brownian motion, *Technometrics*, **18**, 125-132.

[36] Padgett, W. (1981) Bayes estimation of reliability for the inverse Gaussian model, *IEEE Trans. Relia.*, **30**, 84-85.

[37] Pal, N. and Sinha, B.K (1989) Improved Estimators of Dispersion of an Inverse Gaussian Distribution, *Statistical Data Analysis and Inference*, (editor Y. Dodge), North Holland.

[38] Pandey, B. N. and Malik, H.J. (1988) Some improved estimators for a measure of dispersion of an Inverse Gaussian Distribution, *Comm. Statis., Theory and methods*, **17(11)**, 3935-3949.

[39] Schrodinger, E. (1915) Zur Theorie der fall- und Steigversuche an Tellchen mit Brownscher Bewegung, *Physikalische Zeitschrift*, **16**, 289-295.

[40] Sheppard, C. W. (1962) *Basic Principles of the Tracer Method*, New York, Wiley.

[41] Shuster, J. J. and Muira, C. (1972) Two way analysis of reciprocals, *Biometrika*, **59**, 478-481.

[42] Shuster, J.J (1968) On the Inverse Gaussian distribution function, *J. Amer. Statist. Assoc.*, **63**, 1514-1516.

[43] Tweedie, M. C. K. (1957a) Statistical properties of inverse gaussian distribution, *Ann. Math. Statist.*, **28**, 362-377.

[44] Wald, A. (1947) *Sequential Analysis*, New York, Wiley.

[45] Wasan, M. T. (1969), On the Inverse Gaussian process, *Skand. Actuar.*, **51**, 69-96.

[46] Wasan, M. T. and Roy, L. K. (1969) Tables on Inverse Gaussian percentage points, *Technometrics*, **11**, 591-604.

[47] Wasan, M. T. (1969) First passage time distribution of Brownian motion with positive drift, (inverse Gaussian distribution), *Queen's Papers in Pure and Applied Mathematics*, **no 19**, Queen's University, Kingston, Ontario.

[48] Weiss, G. (1983) Random Walks and their applications, *Amer. Scientist*, **71**, 65-71.

[49] Whitmore, G. A.(1979) An Inverse Gaussian model for labour turnover, *J. R. Statist. Soc.* series A, 142(4), 468-479.

[50] Whitmore, G. A.(1979) An exact Multiple Regression modelfor Inverse Gaussian Data, working paper, McGill University, Faculty of Management,

[51] Whitmore, G. A.(1983) Aregression method for Censored Inverse Gaussian Data, *Canadian Journal of Statistics,* 11, 305-315.

[52] Whitmore, G. A.(1986) Normal Gamma Mixtures of Inverse Gaussian distributions, *Scand. J. Statist.*, **13**, 211-220.

[53] Whitmore, G. A.(1986) Inverse Gaussian Ratio Estimation, *Applied Statistics*, **35**, 8-15.

[54] Whitmore, G. A. (1976) Management applications of the inverse Gaussian distribution, *Internat. J. Acta. Phy. Pharmacol. Neerlandica.*, **14**, 175-204.

[55] Whitmore, G. A. and Yalovsky, M. (1978) A normalising logarithmic transformation for Inverse Gaussian random variables, *Technometrics*, **20**, 207-208.

[56] Wise, M. E. (1966). Tracer dilution curves in cardiology and random walk and lognormal distri butions, *Acia Physiologica Pharmacologica Neeriandica.*, **14**, 175-204.

[57] Zigangirov, K. S. (1962) Expression for the Wald distribution in terms of normal distribution, *Radiotekhnika Electronika*, **7**, 164-168.

APPENDIX

```
/* PROGRAM TO FIND MAXIMUM LIKELIHOOD ESTIMATE */

#include <math.h>
#include </home/f1/murali/laks/nrutil.h>
#include </home/f1/murali/laks/nrutil.c>


main() {

  int I,J,IJ,i,j;
  float **x, **m, **minv, **mm, **d, **xp, **phi, **delta, **xpd, **tij;
  void prt(),mul(),gaussj();

  printf("Enter I , J : \n");
  scanf("%d %d",&I,&J);
  IJ = I*J;

  printf("I: %d , J: %d \n\n",I,J);

  x = matrix(1,IJ,1,IJ);
  xp = matrix(1,IJ,1,IJ);

  printf("Enter matrix x : \n");
  for(i=1;i<= IJ;i++) {
     for(j=1;j<=IJ;j++) {
         scanf("%f",&x[i][j]);
         xp[j][i] = x[i][j];
     }
  }

  printf("\nMatrix x : \n");
  prt(x,IJ,IJ);

  printf("\nMatrix xp : \n");
  prt(xp,IJ,IJ);


  d = matrix(1,IJ,1,IJ);
  for(i=1;i<= IJ;i++) {
     for(j=1;j<=IJ;j++) {
          d[i][j] = 0.0;
     }
```

```
}
printf("Enter diagonal elements of d : \n");
for(i=1;i<= IJ;i++) {
    scanf("%f",&d[i][i]);
}

printf("\nMatrix d : \n");
prt(d,IJ,IJ);

xpd = matrix(1,IJ,1,IJ);
mul(xp,IJ,IJ,d,IJ,xpd);

printf("\nMatrix xpd : \n");
prt(xpd,IJ,IJ);

m = matrix(1,IJ,1,IJ);
mul(xpd,IJ,IJ,x,IJ,m);
printf("Matrix m: \n");
prt(m,IJ,IJ);

minv = matrix(1,IJ,1,IJ);
for(i=1;i<=IJ;i++) {
   for(j=1;j<=IJ;j++) {
      minv[i][j]=m[i][j];
   }
}

delta = matrix(1,IJ,1,1);
for(i=1;i<=IJ;i++) {
   for(j=1;j<=1;j++) {
      delta[i][j] = 0.0;
   }
}
delta[1][1] = IJ;

gaussj(minv,IJ,delta,1);

printf("\nMatrix minv: \n");
prt(minv,IJ,IJ);


mm = matrix(1,IJ,1,IJ);
mul(m,IJ,IJ,minv,IJ,mm);
printf("\nMatrix mm: \n");
prt(mm,IJ,IJ);
```

```
  phi = matrix(1,IJ,1,1);
  for(i=1;i<=IJ;i++) {
     for(j=1;j<=1;j++) {
        phi[i][j] = delta[i][j];
     }
  }

  tij = matrix(1,IJ,1,1);
  mul(x,IJ,IJ,phi,1,tij);

  printf("\n\n");
  for(i=1;i<=IJ;i++)
     printf("tij[%d]:  %f    1/tij:  %f  phi[%d]:  %f \n",i,tij[1][1],1.0/tij[1

}

void prt(a,m,n)
  float **a;
  int m,n;
{ int i,j;

  for(i=1;i<= m;i++) {
     for(j=1;j<=n;j++) {
         printf("%6.4f  ",a[i][j]);
     }
     printf("\n");
  }
}

void mul(a,m,n,b,k,c)
  float **a;              /* first matrix */
  int m,n;                /* dimension of a */
  float **b;              /* second matrix */
  int k;                  /* no. of coulumns in b */
  float **c;              /* product matrix */
{ int i,j,l;

  for(i=1;i<=m;i++) {
     for(j=1;j<=k;j++) {
         c[i][j] = 0.0;
         for(l=1;l<=n;l++) {
            c[i][j] += a[i][l] * b[l][j];
         }
     }
```

```
  }
}


#define SWAP(a,b) {float temp=(a);(a)=(b);(b)=temp;}

void gaussj(a,n,b,m)
float **a,**b;
int n,m;
{
int *indxc,*indxr,*ipiv;
int i,icol,irow,j,k,l,ll,*ivector();
float big,dum,pivinv;
void nrerror(),free_ivector();

indxc=ivector(1,n);
indxr=ivector(1,n);
ipiv=ivector(1,n);
for (j=1;j<=n;j++) ipiv[j]=0;
for (i=1;i<=n;i++) {
big=0.0;
for (j=1;j<=n;j++)
if (ipiv[j] != 1)
for (k=1;k<=n;k++) {
if (ipiv[k] == 0) {
if (fabs(a[j][k]) >= big) {
big=fabs(a[j][k]);
irow=j;
icol=k;
}
} else if (ipiv[k] > 1) nrerror("GAUSSJ: Singular Matrix-1");
}
++(ipiv[icol]);
if (irow != icol) {
for (l=1;l<=n;l++) SWAP(a[irow][l],a[icol][l])
for (l=1;l<=m;l++) SWAP(b[irow][l],b[icol][l])
}
indxr[i]=irow;
indxc[i]=icol;
if (a[icol][icol] == 0.0) nrerror("GAUSSJ: Singular Matrix-2");
pivinv=1.0/a[icol][icol];
a[icol][icol]=1.0;
for (l=1;l<=n;l++) a[icol][l] *= pivinv;
for (l=1;l<=m;l++) b[icol][l] *= pivinv;
for (ll=1;ll<=n;ll++)
```

```
if (ll != icol) {
dum=a[ll][icol];
a[ll][icol]=0.0;
for (l=1;l<=n;l++) a[ll][l] -= a[icol][l]*dum;
for (l=1;l<=m;l++) b[ll][l] -= b[icol][l]*dum;
}
}
for (l=n;l>=1;l--) {
if (indxr[l] != indxc[l])
for (k=1;k<=n;k++)
SWAP(a[k][indxr[l]],a[k][indxc[l]]);
}
free_ivector(ipiv,1,n);
free_ivector(indxr,1,n);
free_ivector(indxc,1,n);
}

#undef SWAP
```

```
/* PROGRAM TO ESTIMATE SIGMA */


#include <math.h>
#include </home/f1/murali/laks/nrutil.h>
#include </home/f1/murali/laks/nrutil.c>


main ()
{

int i,j,k,m,n;
float yij[250],rij[250],sum,mu,**ybar,sum2,sigma;

sum = sum2= 0.0;
printf ("enter m,n,k\n");
scanf ("%d  %d  %d", &m,&n,&k);
for (i=1; i<= m*n*k; i++)
{
scanf ("%f", &yij[i]);
rij[i] = 1/yij[i];
printf ("rij[%d] is: %f \n", i, rij[i]);
sum = sum + rij[i];
}
printf ("The sum of the reciprocals over i,j,k is: %f \n", sum);

ybar=matrix(1,m,1,n);

for (i =1; i<= m; i++)
{
  for (j=1; j<=n; j++)
     {
     scanf ("%f", &ybar[i][j]);
     sum2 = sum2 + (1/( ybar[i][j]));
     }
}
sigma = (sum -( k*sum2))/(m*n*(k-1));
printf ("The vaue of sigma is: %f \n", sigma);
}
```

```
/* PROGRAM FOR LEAST SQUARES ESTIMATION */


#include <math.h>
#include </home/f1/murali/laks/nrutil.h>
#include </home/f1/murali/laks/nrutil.c>

main ()
{
int i,j,k m,n;
float **yij,**rij,**bij,sum,ai[10],sum1[10],sum2[10],mu,a,**tij,**eij;
float suma,sumb;

printf ("enter m,n : \n");
scanf ("%d %d",&m,&n);

yij=matrix(1,m,1,n);
rij=matrix(1,m,1,n);
tij=matrix(1,m,1,n);
eij=matrix(1,m,1,n);

bij=matrix(1,m,1,n);
sum = suma = sumb = 0.0;
for(i=1;i<=m;i++) sum2[i]=0.0;
for(j=1;j<=n;j++) sum1[j]=0.0;

printf ("enter matrix yij: \n");

for (i=1; i<=m; i++) {
    for (j=1; j<=n; j++) {
         scanf ("%f", &yij[i][j]);
         rij[i][j]= 1/yij[i][j];
         sum = sum + rij[i][j];
         sum2[i] += rij[i][j];
         sum1[j] += rij[i][j];
    }
}
mu = sum/(m*n);
printf ("The overall mean, mu is: %f \n", mu);
    printf ("The mean summed over j is: \n");
for(i=1;i<=m;i++)
{
    a = sum2[i]/n;
    ai[i] = a - mu;
    suma = suma + ai[i];
```

```
    printf ("sum2[%d]:  %f, ai[%d]: %f\n", i,a,i, ai[i]);
    }
printf ("The value of the R(A) is: %f \n", m * suma);
printf ("The value of the nested factors bij is: \n");
for (i= 1; i<=m; i++)
{
     for (j= 1; j<= n; j++)
{
          bij[i][j] = (rij[i][j] - (sum2[i]/n));
          printf ("bij[%d][%d]: %f\n", i,j, bij[i][j]);
  sumb = sumb + bij[i][j];
}
}
printf ("The value of the R(B(A)) is: %f \n ", n * sumb);
for (i = 1; i<= m; i++)
{
     for (j=1; j <= n; j++)
      {
         tij[i][j] = mu - ((0.004979)/6) + ai[i] + bij[i][j];
          eij[i][j] = rij[i][j] - tij[i][j];
          printf ("tij[%d][%d]: %f\n\n", i,j, tij[i][j]);
          printf ("eij[%d][%d]: %f\n\n", i,j, eij[i][j]);
      }
}
}
```

```
/* PROGRAM TO FIND WEIGHTED LEAST SQUARES ESTIMATE */


#include <math.h>
#include </home/f1/murali/laks/nrutil.h>
#include </home/f1/murali/laks/nrutil.c>


main() {

  int I,J,IJ,i,j;
  float **x, **m, **minv, **mm, **d, **xp, **phi, **delta;
  float  **xpd, **tij,**s,**w,**the;

  void prt(),mul(),gaussj();

  printf("Enter I , J : \n");
  scanf("%d %d",&I,&J);
  IJ = I*J;

  printf("I: %d , J: %d \n\n",I,J);

  x = matrix(1,IJ,1,IJ);
  xp = matrix(1,IJ,1,IJ);

  printf("Enter matrix x : \n");
  for(i=1;i<= IJ;i++) {
     for(j=1;j<=IJ;j++) {
         scanf("%f",&x[i][j]);
         xp[j][i] = x[i][j];
     }
  }

  printf("\nMatrix x : \n");
  prt(x,IJ,IJ);

  printf("\nMatrix xp : \n");
  prt(xp,IJ,IJ);


  d = matrix(1,IJ,1,IJ);
  for(i=1;i<= IJ;i++) {
     for(j=1;j<=IJ;j++) {
          d[i][j] = 0.0;
     }
```

```
}
printf("Enter diagonal elements of d : \n");
for(i=1;i<= IJ;i++) {
    scanf("%f",&d[i][i]);
}

printf("\nMatrix d : \n");
prt(d,IJ,IJ);

xpd = matrix(1,IJ,1,IJ);
mul(xp,IJ,IJ,d,IJ,xpd);

printf("\nMatrix xpd : \n");
prt(xpd,IJ,IJ);

m = matrix(1,IJ,1,IJ);
mul(xpd,IJ,IJ,x,IJ,m);
printf("Matrix m: \n");
prt(m,IJ,IJ);

minv = matrix(1,IJ,1,IJ);
for(i=1;i<=IJ;i++) {
   for(j=1;j<=IJ;j++) {
      minv[i][j]=m[i][j];
   }
}

delta = matrix(1,IJ,1,1);
for(i=1;i<=IJ;i++) {
   for(j=1;j<=1;j++) {
      delta[i][j] = 0.0;
   }
}
delta[1][1] = IJ;

gaussj(minv,IJ,delta,1);

printf("\nMatrix minv: \n");
prt(minv,IJ,IJ);


mm = matrix(1,IJ,1,IJ);
mul(m,IJ,IJ,minv,IJ,mm);
printf("\nMatrix mm: \n");
prt(mm,IJ,IJ);
```

```
  phi = matrix(1,IJ,1,1);
  for(i=1;i<= IJ;i++) {
      scanf("%f",&phi[i][1]);
  }
  printf("\n,Matrix phi:\n");
  prt (phi,IJ,1);
  s = matrix(1,IJ,1,1);

  w = matrix(1,IJ,1,1);

  mul(d,IJ,IJ,phi,1,w);
  printf ("\nMatrix w: \n");
  prt (w,IJ,1);

  mul(xp,IJ,IJ,w,1,s);
  printf ("\nMatrix s: \n");
  prt (s,IJ,1);

  tij = matrix(1,IJ,1,1);
  mul(minv,IJ,IJ,s,1,tij);

  printf("\n\n");
  for(i=1;i<=IJ;i++)
    {
      printf("tij[%d]:  %f  \n",i,tij[i][1]);
    }

  the = matrix(1,IJ,1,1);
  mul (x,IJ,IJ,tij,1,the);
  printf ("\nMatrix the:\n");
  prt(the,IJ,1);
}
void prt(a,m,n)
  float **a;
  int m,n;
{ int i,j;

  for(i=1;i<= m;i++) {
     for(j=1;j<=n;j++) {
         printf("%6.4f  ",a[i][j]);
     }
     printf("\n");
  }
}
```

```
void mul(a,m,n,b,k,c)
  float **a;              /* first matrix */
  int m,n;                /* dimension of a */
  float **b;              /* second matrix */
  int k;                  /* no. of coulumns in b */
  float **c;              /* product matrix */
{ int i,j,l;

  for(i=1;i<=m;i++) {
     for(j=1;j<=k;j++) {
         c[i][j] = 0.0;
         for(l=1;l<=n;l++) {
            c[i][j] += a[i][l] * b[l][j];
         }
     }
  }
}


#define SWAP(a,b) {float temp=(a);(a)=(b);(b)=temp;}

void gaussj(a,n,b,m)
float **a,**b;
int n,m;
{
int *indxc,*indxr,*ipiv;
int i,icol,irow,j,k,l,ll,*ivector();
float big,dum,pivinv;
void nrerror(),free_ivector();

indxc=ivector(1,n);
indxr=ivector(1,n);
ipiv=ivector(1,n);
for (j=1;j<=n;j++) ipiv[j]=0;
for (i=1;i<=n;i++) {
big=0.0;
for (j=1;j<=n;j++)
if (ipiv[j] != 1)
for (k=1;k<=n;k++) {
if (ipiv[k] == 0) {
if (fabs(a[j][k]) >= big) {
big=fabs(a[j][k]);
irow=j;
icol=k;
```

```
}
} else if (ipiv[k] > 1) nrerror("GAUSSJ: Singular Matrix-1");
}
++(ipiv[icol]);
if (irow != icol) {
for (l=1;l<=n;l++) SWAP(a[irow][l],a[icol][l])
for (l=1;l<=m;l++) SWAP(b[irow][l],b[icol][l])
}
indxr[i]=irow;
indxc[i]=icol;
if (a[icol][icol] == 0.0) nrerror("GAUSSJ: Singular Matrix-2");
pivinv=1.0/a[icol][icol];
a[icol][icol]=1.0;
for (l=1;l<=n;l++) a[icol][l] *= pivinv;
for (l=1;l<=m;l++) b[icol][l] *= pivinv;
for (ll=1;ll<=n;ll++)
if (ll != icol) {
dum=a[ll][icol];
a[ll][icol]=0.0;
for (l=1;l<=n;l++) a[ll][l] -= a[icol][l]*dum;
for (l=1;l<=m;l++) b[ll][l] -= b[icol][l]*dum;
}
}
for (l=n;l>=1;l--) {
if (indxr[l] != indxc[l])
for (k=1;k<=n;k++)
SWAP(a[k][indxr[l]],a[k][indxc[l]]);
}
free_ivector(ipiv,1,n);
free_ivector(indxr,1,n);
free_ivector(indxc,1,n);
}

#undef SWAP
```

```
/* PROGRAM TO FIND SUM OF RECIPROCALS */


#include <math.h>
#include </home/f1/murali/laks/nrutil.h>
#include </home/f1/murali/laks/nrutil.c>

main()
{
 int i,j,k,m,n,l;
 float **yij,**rij, **tij,**ti, rba, rbaa,mea[30],mu[30],**rm;
  scanf ("%d %d", &m,&n);

  yij=matrix(1,m,1,n);
  rij=matrix(1,m,1,n);
  tij=matrix(1,m,1,n);
  ti=matrix(1,m,1,n);
  rm=matrix(1,m,1,n);
  for(i=1;i<=m;i++) mea[i]=0.0;

  for (i=1; i<=m; i++)
    {
       for (j=1; j<=n; j++)
         {
           scanf ("%f", &yij[i][j]);
           mea[i] = mea[i] + yij[i][j];
           scanf ("%f", &ti[i][j]);
           tij[i][j] = 1/(ti[i][j]);
 }
    mu[i] = n/mea[i];
    }

  rba=0.0;
  for (i=1; i<=m; i++)
    {
       for (j=1; j<=n; j++)
         {
             rm[i][j] = tij[i][j] - mu[i];
             rba = rba + rm[i][j];
 }
     }
   rbaa = 6 * rba;
   printf ("The value of rbaa is %f \n", rbaa);
}
```

```
/*  NRUTIL.C */

#include <malloc.h>
#include <stdio.h>

void nrerror(error_text)
char error_text[];
{
void exit();

fprintf(stderr,"Numerical Recipes run-time error...\n");
fprintf(stderr,"%s\n",error_text);
fprintf(stderr,"...now exiting to system...\n");
exit(1);
}



float *vector(nl,nh)
int nl,nh;
{
float *v;

v=(float *)malloc((unsigned) (nh-nl+1)*sizeof(float));
if (!v) nrerror("allocation failure in vector()");
return v-nl;
}

int *ivector(nl,nh)
int nl,nh;
{
int *v;

v=(int *)malloc((unsigned) (nh-nl+1)*sizeof(int));
if (!v) nrerror("allocation failure in ivector()");
return v-nl;
}

double *dvector(nl,nh)
int nl,nh;
{
double *v;

v=(double *)malloc((unsigned) (nh-nl+1)*sizeof(double));
if (!v) nrerror("allocation failure in dvector()");
```

```
return v-nl;
}



float **matrix(nrl,nrh,ncl,nch)
int nrl,nrh,ncl,nch;
{
int i;
float **m;

m=(float **) malloc((unsigned) (nrh-nrl+1)*sizeof(float*));
if (!m) nrerror("allocation failure 1 in matrix()");
m -= nrl;

for(i=nrl;i<=nrh;i++) {
m[i]=(float *) malloc((unsigned) (nch-ncl+1)*sizeof(float));
if (!m[i]) nrerror("allocation failure 2 in matrix()");
m[i] -= ncl;
}
return m;
}

double **dmatrix(nrl,nrh,ncl,nch)
int nrl,nrh,ncl,nch;
{
int i;
double **m;

m=(double **) malloc((unsigned) (nrh-nrl+1)*sizeof(double*));
if (!m) nrerror("allocation failure 1 in dmatrix()");
m -= nrl;

for(i=nrl;i<=nrh;i++) {
m[i]=(double *) malloc((unsigned) (nch-ncl+1)*sizeof(double));
if (!m[i]) nrerror("allocation failure 2 in dmatrix()");
m[i] -= ncl;
}
return m;
}

int **imatrix(nrl,nrh,ncl,nch)
int nrl,nrh,ncl,nch;
{
int i,**m;
```

```
m=(int **)malloc((unsigned) (nrh-nrl+1)*sizeof(int*));
if (!m) nrerror("allocation failure 1 in imatrix()");
m -= nrl;

for(i=nrl;i<=nrh;i++) {
m[i]=(int *)malloc((unsigned) (nch-ncl+1)*sizeof(int));
if (!m[i]) nrerror("allocation failure 2 in imatrix()");
m[i] -= ncl;
}
return m;
}



float **submatrix(a,oldrl,oldrh,oldcl,oldch,newrl,newcl)
float **a;
int oldrl,oldrh,oldcl,oldch,newrl,newcl;
{
int i,j;
float **m;

m=(float **) malloc((unsigned) (oldrh-oldrl+1)*sizeof(float*));
if (!m) nrerror("allocation failure in submatrix()");
m -= newrl;

for(i=oldrl,j=newrl;i<=oldrh;i++,j++) m[j]=a[i]+oldcl-newcl;

return m;
}



void free_vector(v,nl,nh)
float *v;
int nl,nh;
{
free((char*) (v+nl));
}

void free_ivector(v,nl,nh)
int *v,nl,nh;
{
free((char*) (v+nl));
}
```

```
void free_dvector(v,nl,nh)
double *v;
int nl,nh;
{
free((char*) (v+nl));
}



void free_matrix(m,nrl,nrh,ncl,nch)
float **m;
int nrl,nrh,ncl,nch;
{
int i;

for(i=nrh;i>=nrl;i--) free((char*) (m[i]+ncl));
free((char*) (m+nrl));
}

void free_dmatrix(m,nrl,nrh,ncl,nch)
double **m;
int nrl,nrh,ncl,nch;
{
int i;

for(i=nrh;i>=nrl;i--) free((char*) (m[i]+ncl));
free((char*) (m+nrl));
}

void free_imatrix(m,nrl,nrh,ncl,nch)
int **m;
int nrl,nrh,ncl,nch;
{
int i;

for(i=nrh;i>=nrl;i--) free((char*) (m[i]+ncl));
free((char*) (m+nrl));
}



void free_submatrix(b,nrl,nrh,ncl,nch)
float **b;
int nrl,nrh,ncl,nch;
```

```
{
free((char*) (b+nrl));
}



float **convert_matrix(a,nrl,nrh,ncl,nch)
float *a;
int nrl,nrh,ncl,nch;
{
int i,j,nrow,ncol;
float **m;

nrow=nrh-nrl+1;
ncol=nch-ncl+1;
m = (float **) malloc((unsigned) (nrow)*sizeof(float*));
if (!m) nrerror("allocation failure in convert_matrix()");
m -= nrl;
for(i=0,j=nrl;i<=nrow-1;i++,j++) m[j]=a+ncol*i-ncl;
return m;
}



void free_convert_matrix(b,nrl,nrh,ncl,nch)
float **b;
int nrl,nrh,ncl,nch;
{
free((char*) (b+nrl));
}
```

```
/* NRUTIL.H */

float *vector();
float **matrix();
float **convert_matrix();
double *dvector();
double **dmatrix();
int *ivector();
int **imatrix();
float **submatrix();
void free_vector();
void free_dvector();
void free_ivector();
void free_matrix();
void free_dmatrix();
void free_imatrix();
void free_submatrix();
void free_convert_matrix();
void nrerror();
```