العنوان:	STATISTICAL INFERENCE FOR FACTORIAL EXPERIMENTS UNDER AN INVERSE GAUSSIAN MODEL
المصدر:	المجلة العلمية للبحوث والدراسات التجارية
الناشر:	جامعة حلوان - كلية التجارة وإدارة الاعمال
المؤلف الرئيسي:	Saleh, Amina Ali
المجلد/العدد:	ع2
محكمة:	نعم
التاريخ الميلادي:	2007
الصفحات:	42 - 59
رقم MD:	661693
نوع المحتوى:	بحوث ومقالات
قواعد المعلومات:	EcoLink
مواضيع:	التحليل الإحصائي، قانون جاوس، التجارب العملية، العمليات الحسابية
رابط:	http://search.mandumah.com/Record/661693

Statistical Inference for Factorial Experiments under an Inverse Gaussian Model

Amina Ali Saleh

Department of Management Information System, Collage of Administrative Business, King Saud University, Saudi Arabia

Statistical Inference for Factorial Experiments under an Inverse Gaussian Model

Amina Ali Saleh

Department of Management Information System, Collage of Administrative Business, King Saud University, Saudi Arabia

Abstract

This article treats the statistical inference for factorial experiments under an inverse Gaussian distribution for the failure times. A reciprocal-linear model for the factor effects is motivated from the context of the underlying Wiener process. Simple and explicit form for the solution of the likelihood equations is given. Likelihood ratio tests for the main and interaction effects are derived. The test statistics for the main effects have an exact F distribution, while the test statistic for the interaction effects has an approximate F distribution. An analysis of reciprocals analogue of the usual normal theory analysis of variance, is investigated. An application of the procedures is illustrated with a data set of strength measurements of an insulating material. Fries and Bhattacharyya (1983) consider similar model, but they provide no explicit solution to the likelihood equations but via inverting some random matrix. The test statistics for main effects and interaction effect which provided by Fries and Bhattacharyya (1983), all have approximate F distributions. Also, a correction for their analysis of reciprocals is corrected in the present work.

Keywords: Inver Gaussian; Factorial experiment; Maximum likelihood; Analysis of reciprocals.

1. Introduction

The Inverse Gaussian family is a versatile one for modeling nonnegative right-skewed data such as the data obtained from reliability and life test studies. This family shares striking similarities with the Gaussian family. The Inverse family of distributions, denoted as $IG(\theta,\sigma)$ has probability density function given by

$$f(y;\theta,\sigma)=(2\pi\sigma)^{-1/2}\,y^{-3/2}e^{-(2\sigma y)^{-1}(y\theta^{-1}-1)^2};\ y>0,\theta>0,\sigma>0. \tag{1}$$

which belongs to the exponential families. The mean and variance of this distribution are θ and $\theta^3\sigma$ respectively.

The derivation of the IG distribution can be cast in the context of fatigue growth or accumulation of damage in a material over time. Specifically, if fatigue grows according to a Wiener process with a drift $\eta>0$ 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 to failure has the $IG(\theta,\sigma)$ distribution with $\theta=\omega/\eta$ and $\sigma=(\delta/\omega)^2$.

Statistical inference for one- and two-sample from $IG(\theta,\sigma)$ has been extensively studied (see Seshadri 1993 for a survey). Regarding the factorial experiment, few works are available. Tweedie (1957) considered the case of one-way classification where random samples from different IG populations $IG(\theta_i,\sigma_i), i=1,\ldots,a$ and that random samples of size n_i are drawn from population $i\,(i=1,\ldots,a)$, and considered the likelihood ratio (LR) test for testing the hypothesis $\theta_1=\theta_2=\cdots=\theta_a=\theta$ (say) under the assumption that all the populations have the same σ. He observed that the Inverse Gaussian family allows of "hierarchical analysis of variance" which is analogous to similar analysis for Gaussian family. Under the assumption of equal means, he showed that the LR test could be based on a test statistic which distributed as $F_{a-1,n_.-a}$ where $n_.=\sum n_i$. Tweedie called this procedure of testing "analysis of reciprocals". Shuster and Miura (1972) discuss a two-way layout and propose some heuristic tests of hypotheses assuming that θ is linear in the effects and $\theta^2\sigma$ remains constant (i.e. common ratios of mean to variance), this assumption is artificial when viewed in the context of an underlying Wiener process.

Fries and Bhattacharyya (1983) consider the two-factor experiment under the IG model with the assumptions that critical level ω and the diffusion parameter δ^2 of the underlying Wiener processes are constants while the drift η (that is, the mean fatigue growth per unit time) is linear in the factor effects. These assumptions entail a linear model for the reciprocal mean θ^{-1} and a constant σ for all levels of the factors. The constancy of σ is parallel to the homoscedasticity assumption in the usual normal theory analysis of variance (ANOVA). Considering different IG populations; $IG\left((\mu+\alpha_i+\beta_j)^{-1},\sigma\right)$; $i=1,\ldots,a$; $j=1,\ldots,b$; $k=1,\ldots,n$ with the usual assumptions $\sum\alpha_i=\sum\beta_j=0$, they studied the maximum likelihood estimates (MLE's) of the parameters $\mu, \alpha's, \beta's$ and σ, but no explicit form for these MLE's were given. In order to obtain these estimates, they used different representation of the normal equation in order to obtain a unique restricted MLE's. This solution requires inverting some random matrix. Following this procedure, they propose LR test statistics for testing main effects and interaction; each of these test statistics has no explicit form in the observations and approximately follows an F distribution.

In this article we consider the same model of Fries and Bhattacharyya (1983). While they focus on the additive model, we consider interaction model, in a more general form by adding a term represent the interaction effects. We proved explicit form for all MLE's of the parameters $\mu, \alpha's, \beta's$ and σ. Then; LR test statistics for all different hypotheses are developed; each test statistic follows an exact F distribution. Generalization to k-factorial experiments with interactions is straightforward.

We organized the work as follows. In section 2, the case of one factor experiment is considered. As we mention before, Tweedie (1957) considers the case of one-way classification in the means parameterization but not on the effects parameterization as we consider here. This treatment of the one factor experiment is crucial as we will see later. Section 3 deals with the case of two-factor experiment with no interaction, where explicit solution to the likelihood equation is given and a numerical example is used to illustrate that our solution is identical to Fries and Bhattacharyya (1983). The case of two-factor experiment with interaction is considered in section 4. Explicit solution to the likelihood equation is also given. Section 5, discuss hypotheses testing for main and interaction an effect, also the analysis of reciprocals (ANOR) table is constricted as will as a decomposition of the reciprocal observations into components that can be ascribed to the factor effects.

2. The Case of One Factor Experiment

Consider a one factor life test with a levels of the factor. At each level, n items are tested and their failure times y_{ij}, $i = 1,\ldots,a$, and $j = 1,\ldots,n$ recorded. The observations are assumed to be independent with y_{ij} distributed as $IG(\theta_i, \sigma)$. Since the mean is inversely proportional to the drift, the usual parameterization suggests the model

$$\theta_i^{-1} = \mu + \alpha_i, \ \sum_{i=1}^{a} \alpha_i = 0 \tag{2}$$

where μ and α_i's represents the grand mean and the main factor effects respectively. For the IG distribution we must have $\theta_i \geq 0$ for all i and $\sigma > 0$. Thus the parameters μ, $\alpha' = (\alpha_1,\ldots,\alpha_a)$ and σ lie in the set

$$\Omega = \left\{(\mu, \alpha', \sigma) : \sum_i \alpha_i = 0;\ \mu + \alpha_i > 0, i = 1,\ldots,a;\ \sigma > 0\right\} \tag{3}$$

We introduced the basic notation for the totals and the means that will be used throughout the paper:

$$y_{i.} = \sum_j y_{ij} = n\bar{y}_{i.}, y_{..} = \sum_i \sum_j y_{ij} = na\bar{y}_{..}, R = \sum_i \sum_j y_{ij}^{-1} \tag{4}$$

Referring to (1) and (2), the log-likelihood function has the form

$$l = const. - (1/2)an\ \log\sigma - (2\sigma)^{-1} \sum \sum y_{ij}^{-1} \left[y_{ij}(\mu + \alpha_i) - 1\right]^2 \tag{5}$$

Expanding the squared term, we find that the set $(\bar{y}_{1.},\ldots,\bar{y}_{a.}, R)$ represent a set of $(a+1)$-dimensional sufficient statistics, with the parameter space Ω of dimension $(a+1)$ as well.

Equating to zero the first partial derivates of (5) with respect to (wrt) μ and α_i, we obtain

$$\begin{aligned} &\hat{\mu}\, y_{i.} + \sum_i \hat{\alpha}_i\, y_{i.} = na \\ &\hat{\mu}\, y_{i.} + \quad \hat{\alpha}_i\, y_{i.} = n, \qquad a \leq i \leq 1 \end{aligned} \tag{6}$$

and the derivative wrt σ leads to

$$\hat{\sigma} = \frac{1}{an} \sum \sum y_{ij}^{-1} \left[y_{ij}(\hat{\mu} + \hat{\alpha}_i) - 1\right]^2 \tag{7}$$

The system (6) of equations has the following unique solution

$$\hat{\mu} = \frac{1}{a}\sum_i \frac{1}{\bar{y}_{i.}}$$
$$\hat{\alpha}_i = \frac{1}{\bar{y}_{i.}} - \frac{1}{a}\sum_i \frac{1}{\bar{y}_{i.}}, \; i = 1,\ldots,a \tag{8}$$

While

$$\hat{\sigma} = \frac{1}{an}\left[R - \frac{an}{a}\sum_i \frac{1}{\bar{y}_{i.}}\right] \text{ or } = \frac{1}{an}\left[R - an\hat{\mu}\right] \tag{9}$$

To test the hypothesis of no main effects, i.e. to test $H_0 : \alpha_1 = \alpha_2 = \ldots = \alpha_a = 0$, against $H_1 : \alpha_1 \neq \alpha_2 \neq \ldots \neq \alpha_a$, the LR test statistic is given by

$$\Lambda = 2\left[l_{\max(\Omega_1)} - l_{\max(\Omega_0)}\right] = a\,n\log\left(\frac{\hat{\sigma}_0}{\hat{\sigma}_1}\right) \tag{10}$$

The last expression in (10) obtains from the general result that the maximized log-likelihood, under each model Ω_i, $i = 0,1$, has the value $-\left(\frac{1}{2}\right)a\,n\left(\log\hat{\sigma}_i + 1\right)$, ignoring the constant term,. Expression for $\hat{\sigma}_1$ is given by (9) while $\hat{\sigma}_0$ is given by

$$\hat{\sigma}_0 = \frac{1}{an}\left[R - a\,n\frac{1}{\bar{y}_{..}}\right] \tag{11}$$

The rejection region consists of the large values of the statistic in (10).

We note that;

$$\Lambda = an\log\left(1 + \frac{\hat{\sigma}_0 - \hat{\sigma}_1}{\hat{\sigma}_1}\right) \tag{12}$$

which is strictly increasing function of $R_{01} = \frac{\hat{\sigma}_0 - \hat{\sigma}_1}{\hat{\sigma}_1}$. Consequently, the LR test can equivalently be based on R_{01} with large values in the rejection region.

Let $Q_1^* = \hat{\sigma}_1$ and

$$Q_2^* = \hat{\sigma}_0 - \hat{\sigma}_1 = \sum_i\sum_j\left[\frac{1}{y_{ij}} - \frac{1}{\bar{y}_{..}}\right] - \sum_i\sum_j\left[\frac{1}{y_{ij}} - \frac{1}{\bar{y}_{i.}}\right]$$

$$= \sum_i \left[\frac{n}{\bar{y}_{i.}} - \frac{n}{\bar{y}_{..}} \right] \tag{13}$$

hence, $R_{01} = \dfrac{\hat{\sigma}_0 - \hat{\sigma}_1}{\hat{\sigma}_1} = \dfrac{Q_2^*}{Q_1^*}$ (14)

The statistic $Q_1^* = \sum_i \sum_j \left[\frac{1}{Y_{ij}} - \frac{1}{\bar{Y}_{i.}} \right]$ divided by σ has χ^2 distribution with $a(n-1)$. While, under the assumption of no main effects, the statistics $Q_2^* = \sum_i \left[\frac{n}{\bar{Y}_{i.}} - \frac{n}{\bar{Y}_{..}} \right]$ divided by σ has χ^2 distribution with $a-1$ degrees of freedom. The two statistics are independent (see Seshadri (1983) and Datta (2005) for properties of inverse Gaussian distribution); hence one can use the F test based on the statistic

$$T_{01} = \frac{a(n-1)Q_2^*}{(a-1)Q_1^*} \tag{15}$$

with $(a-1)$ and $a(n-1)$ degrees of freedom.

3. The Case of Two-factor Experiment With no Interaction

In this section, we consider the same model as Fries and Bhattacharyya (1983). While Fries and Bhattacharyya (1983) could not obtain an explicit solution to the set of normal equation, we will provide a solution to these equations and hence obtain an explicit form of the ML estimates of the model's parameters. A numerical example used before by Fries and Bhattacharyya (1983), is used here to show that the two sets of solutions leads to identical estimates of the means.

The two-factor life test consists of a levels of factor **A** and b levels of factor **B**. At each factor setting or cell (i, j), n items are tested and their failure times y_{ijk}, $i = 1,\ldots,a$, $j = 1,\ldots,b$ and $k = 1,\ldots,n$ recorded. The observations are independent with y_{ijk} distributed as IG. We focus in this section on the additive or no-interaction model, which assumes that the drift of the Wiener process corresponding to each cell is the sum of the factor effects. Assuming that the mean is inversely proportional to the drift, the usual parameterization of the model is

$$\theta_{ij}^{-1} = \mu + \alpha_i + \beta_j, \qquad \sum_{i=1}^{a} \alpha_i = \sum_{j=1}^{b} \beta_j = 0 \tag{16}$$

where μ, α_i's, and β_j's represent the grand mean, the main effects of factor **A**, and the main effects of factor **B**, respectively. We must have $\theta_{ij} \geq 0$ for all i, j and $\sigma > 0$. Thus the parameters μ, $\alpha' = (\alpha_1, \ldots, \alpha_a)$, $\beta' = (\beta_1, \ldots, \beta_b)$, and σ lie in the set

$$\begin{aligned} \Omega = \{ (\mu, \alpha', \beta', \sigma) &: \sum_i \alpha_i = \sum_j \beta_j = 0; \\ &\mu + \alpha_i + \beta_j > 0, \quad i = 1, \ldots, a; \ j = 1, \ldots, b; \sigma > 0 \} \end{aligned} \tag{17}$$

We extended the basic notation for the totals and the means to the two-factors experiments as follows

$$\begin{aligned} & y_{ij.} = \sum_k y_{ijk} = n\bar{y}_{ij.} & & y_{i..} = \sum_j \sum_k y_{ijk} = nb\bar{y}_{i..} \\ & y_{.j.} = \sum_i \sum_k y_{ijk} = na\bar{y}_{.j.} & & y_{...} = \sum_i \sum_j \sum_k y_{ijk} = nab\bar{y}_{...} \\ & R = \sum_i \sum_j \sum_k y_{ijk}^{-1} & & \end{aligned} \tag{18}$$

The log-likelihood function has the form

$$\begin{aligned} l &= const. - (1/2)an \log \sigma \\ &\quad - (2\sigma)^{-1} \sum_i \sum_j \sum_k y_{ijk}^{-1} \left[y_{ijk} (\mu + \alpha_i + \beta_j) - 1 \right]^2 \end{aligned} \tag{19}$$

Equating to zero the first partial derivates of (18) wrt μ and α_i, we obtain

$$\left. \begin{aligned} & \hat{\mu} y_{...} + \sum_i \hat{\alpha}_i y_{i..} + \sum_j \hat{\beta}_j y_{.j.} = nab \\ & \hat{\mu} y_{i..} + \hat{\alpha}_i y_{i..} + \sum_j \hat{\beta}_j y_{ij.} = bn, \quad 1 \leq i \leq a \\ & \hat{\mu} y_{.j.} + \sum_i \hat{\alpha}_i y_{ij.} + \hat{\beta}_j y_{.j.} = an, \quad 1 \leq j \leq b \end{aligned} \right\} \tag{20}$$

and the derivative wrt σ leads to

$$\hat{\sigma} = \frac{1}{abn} \sum_i \sum_j \sum_k y_{ijk}^{-1} \left[y_{ijk} \left(\hat{\mu} + \hat{\alpha}_i + \hat{\beta}_j \right) - 1 \right]^2 \tag{21}$$

Note that the summation of the a equations associated with the α_i's yields the first equation; also, the summation of the b equations associated with the β_j's yields the first equation.

It seems that there is no explicit solution to the system (20) of normal equations. In fact, Fries and Bhattacharyya (1983) used the conditions $\sum_i \alpha_i = \sum_j \beta_j = 0$ to delete the last components of α and β, and hence, obtain restricted ML estimates of the model's parameters by inverting some random matrix. The following theorem provide explicit expressions for the ML estimates of the model's parameters, which is a natural generalization of the solution obtained before in the case of the one factor experiment.

Theorem 1

The solution to the system of linear equation (20) is given by

$$\left.\begin{aligned} \hat{\mu} &= \frac{1}{a}\sum_i \frac{1}{\bar{y}_{i..}} + \frac{1}{b}\sum_j \frac{1}{\bar{y}_{.j.}} - \frac{1}{\bar{y}_{...}} \\ \hat{\alpha}_i &= \frac{1}{\bar{y}_{i..}} - \frac{1}{a}\sum_i \frac{1}{\bar{y}_{i..}}, \quad i = 1,\ldots,a \\ \hat{\beta}_j &= \frac{1}{\bar{y}_{.j.}} - \frac{1}{b}\sum_i \frac{1}{\bar{y}_{.j.}}, \quad j = 1,\ldots,b \end{aligned}\right\} \tag{22}$$

and the ML estimate of σ is

$$\hat{\sigma} = \frac{1}{abn}\left[R - abn\hat{\mu}\right] \tag{23}$$

Proof

First notice that $\sum_i \hat{\alpha}_i = \sum_j \hat{\beta}_j = 0$ and that $\hat{\theta}_{ij}^{-1} = \hat{\mu} + \hat{\alpha}_i + \hat{\beta}_j$ $= \frac{1}{\bar{y}_{i..}} + \frac{1}{\bar{y}_{.j.}} - \frac{1}{\bar{y}_{...}} \geq 0$ for all i, j; i.e. they obtained within the parameter space Ω. Second, It is easy to verify that these estimators satisfy the first equation of the system. Finally, substituting these estimators into the a equations associated with the α_i's, and summing over i yields the first equation. The same thing is obtained regarding the system of the b equations associated with the β_j's.

The expression of $\hat{\sigma}$ is obtained in a straightforward way. □

We illustrate the above result with the following example which was used before by Fries and Bhattacharyya (1983).

Example 1

Shuster and Miura (1972) analyzed a data set from Ostel (1963), which is in the form of a randomized 2×5 layout with 10 replicates per cell. The data consist of the impact strength, in foot-pounds, from tests on 5 lots of the same type of insulating material that are cut either lengthwise or crosswise. The use of an IG distribution is plausible since the impact strength is determined by building up stresses until failure occurs. The assumption of constant diffusion parameter is also appropriate since the same type of insulating material is being tested under a fixed specification of the failure criterion.

Fries and Bhattacharyya (1983) used this data and obtain the following restricted ML estimates for σ and $\phi = (\mu,\ \alpha_1,\ \beta_1,\ \beta_2,\ \beta_3,\ \beta_4)'$

$$\hat{\sigma} = .02261$$

$$\hat{\phi} = (1.342,\ -.039,\ -.134,\ -.352,\ .182,\ -.212)'$$

wile, $\hat{\alpha}_1 = -\hat{\alpha}_2 = .039$ and $\hat{\beta}_5 = -\sum_{j=1}^{4}\beta_j = -.516$

The ML estimates for σ and $\mu,\ \alpha = (\alpha_1, \alpha_2)'$, $\beta = (\beta_1, \beta_2, \beta_3, \beta_4, \beta_5)'$ using the solution given by (22) are

$$\hat{\sigma} = .02261$$

$$\hat{\mu} = 1.342,\ \ \hat{\alpha} = (-.035, .035)'\ \hat{\beta} = (-.137, -.350, .181, -.210, .516)'$$

The corresponding estimates for the cell means are calculated by using the relation $\hat{\theta}_{ij}^{-1} = \hat{\mu} + \hat{\alpha}_i + \hat{\beta}_j$. These estimated cell means are identical to those obtained by Fries and Bhattacharyya (F&B) and are given in table1.

Table1. MLE of Mean Impact Strengths (10 Replicates per Cell)

Type of cut		*Lot Number*				
		I	*II*	*III*	*IV*	*V*
Lengthwise	$\bar{y}$	.919	.997	.690	.870	.551
	RMLE (F&B)	.855	1.051	.673	.916	.550
	MLE	.855	1.050	.672	.912	.552
Crosswise	$\bar{y}$	.743	1.022	.624	.899	.526
	RMLE (F&B)	.803	.972	.640	.856	.527
	MLE	.807	.974	.642	.857	.528

Inspection of Table1 shows that both sets of estimators, the restricted ML estimators given by Fries and Bhattacharyya (1983) and the ML estimators given by (21) give the same estimates of the cell means $\hat{\theta}_{ij}^{-1}, i=1,\dots,a;\ j=1,\dots,b$. That is to say that the parameters θ_{ij}'s are estimable. However, our estimates, besides being explicit, are easy to calculate and accurate, since no rounded errors due to inverting any kind of matrices are involved. □

4. The Case of Two-factor Experiment With Interaction (The full model)

In this section, we consider the same model as in the previous section and impose the term of interaction to the means. As before, assume that the mean is inversely proportional to the drift, the usual parameterization of the model is

$$\theta_{ij}^{-1} = \mu + \alpha_i + \beta_j + \delta_{ij}, \quad \sum_{i=1}^{a} \alpha_i = \sum_{j=1}^{b} \beta_j = \sum_{i} \delta_{ij} = \sum_{j} \delta_{ij} = 0 \qquad (24)$$

where μ, α_i's , β_j's and δ_{ij}'s represent the grand mean, the main effects of factor **A**, the main effects of factor **B** and the interaction effects at the *i*th level of factor A and the *j*th level of factor B, respectively. We must have $\theta_{ij} \geq 0$ for all i, j and $\sigma > 0$. Thus the parameters μ, $\alpha' = (\alpha_1,\dots,\alpha_a)$, $\beta' = (\beta_1,\dots,\beta_b)$, $\delta' = (\delta_{11},\dots,\delta_{ab})$ and σ lie in the set

$$\Omega = \left\{(\mu, \alpha', \beta', \delta', \sigma): \sum_i \alpha_i = \sum_j \beta_j = \sum_i \delta_{ij} = \sum_j \delta_{ij} = 0; \right.$$
$$\left. \mu + \alpha_i + \beta_j + \delta_{ij} > 0, \quad i = 1,\ldots,a; \; j = 1,\ldots,b; \sigma > 0 \right\} \qquad (25)$$

The log-likelihood function has the form

$$l = const. - (1/2) an \log \sigma$$
$$- (2\sigma)^{-1} \sum_i \sum_j \sum_k y_{ijk}^{-1} \left[y_{ijk} \left(\mu + \alpha_i + \beta_j + \delta_{ij} \right) - 1 \right]^2 \qquad (26)$$

Equating to zero the first partial derivates of (26) wrt μ and α_i, we obtain

$$\hat{\mu} y_{...} + \sum_i \hat{\alpha}_i y_{i..} + \sum_j \hat{\beta}_j y_{.j.} + \sum_i \sum_j \delta_{ij} y_{ij.} = nab$$
$$\hat{\mu} y_{i..} + \hat{\alpha}_i y_{i..} + \sum_j \hat{\beta}_j y_{ij.} + \sum_j \delta_{ij} y_{ij.} = bn, \qquad 1 \le i \le a$$
$$\hat{\mu} y_{.j.} + \sum_i \hat{\alpha}_i y_{ij.} + \hat{\beta}_j y_{.j.} + \sum_j \delta_{ij} y_{ij.} = an, \qquad 1 \le j \le b \qquad (27)$$
$$\hat{\mu} y_{ij.} + \hat{\alpha}_i y_{ij.} + \hat{\beta}_j y_{ij.} + \delta_{ij} y_{ij.} = n, \qquad 1 \le i \le a, 1 \le j \le b$$

and the derivative wrt σ leads to

$$\hat{\sigma} = \frac{1}{abn} \sum_i \sum_j \sum_k y_{ijk}^{-1} \left[y_{ijk} \left(\hat{\mu} + \hat{\alpha}_i + \hat{\beta}_j + \delta_{ij} \right) - 1 \right]^2 \qquad (28)$$

It can be seen easily that the solution to the normal equations given in (27) is

$$\left.\begin{aligned}
\hat{\mu} &= \frac{1}{ab} \sum_i \sum_j \frac{1}{\bar{y}_{ij.}} \\
\hat{\alpha}_i &= \frac{1}{ab} \sum_i \sum_j \frac{1}{\bar{y}_{ij.}} - \frac{1}{b} \sum_j \frac{1}{\bar{y}_{ij.}}, \quad i = 1,\ldots,a \\
\hat{\beta}_j &= \frac{1}{ab} \sum_i \sum_j \frac{1}{\bar{y}_{ij.}} - \frac{1}{a} \sum_i \frac{1}{\bar{y}_{ij.}}, \quad j = 1,\ldots,b \\
\hat{\delta}_{ij} &= \frac{1}{\bar{y}_{ij.}} - \hat{\mu} - \hat{\alpha}_i - \hat{\beta}_j, \quad i = 1,\ldots,a, \quad j = 1,\ldots,b
\end{aligned}\right\} \qquad (29)$$

with

$$\hat{\theta}_{ij}^{-1} = \frac{1}{\bar{y}_{ij.}} \tag{30}$$

while

$$\hat{\sigma} = \frac{1}{abn}\left[R - nab\,\hat{\mu}\right] = \frac{1}{abn}\left[R - n\sum_i\sum_j \frac{1}{\bar{y}_{ij.}}\right] \tag{31}$$

5. Hypotheses Tests

The results of the preceding sections will now be used to develop LR tests for ANOVA-type hypotheses.

For testing the hypotheses of additivity or absence of main factor effects, the relevant models (hypotheses) are

$$\left.\begin{array}{ll}
\Omega_4 : \theta_{ij}^{-1} = \mu + \alpha_i + \beta + \delta_{ij} \ , & \\
\quad \sum_{i=1}^{a}\alpha_i = \sum_{j=1}^{b}\beta_j = \sum_i \delta_{ij} = \sum_j \delta_{ij} = 0 & \text{(Full model)} \\
\Omega_3 : \theta_{ij}^{-1} = \mu + \alpha_i + \beta_j & \\
\quad \sum_{i=1}^{a}\alpha_i = \sum_{j=1}^{b}\beta_j = 0 & \text{(additive model)} \\
\Omega_2 : \theta_{ij}^{-1} = \mu + \alpha_i \ , \ \sum_{i=1}^{a}\alpha_i = 0 & \text{(no } \mathbf{B} \text{ effects)} \\
\Omega_1 : \theta_{ij}^{-1} = \mu + \beta_j \ , \ \sum_{j=1}^{b}\beta_j = 0 & \text{(no } \mathbf{A} \text{ effects)} \\
\Omega_0 : \theta_{ij}^{-1} = \mu & \text{(no factor effects)}
\end{array}\right\} \tag{32}$$

It is understood that each model also has the nuisance parameter σ and that all θ_{ij}'s are all constrained to be positive.

Let $\hat{\sigma}_s$ denote the ML estimate of σ. Expressions for $\hat{\sigma}_s$'s are given by

$$\hat{\sigma}_4 = \frac{1}{abn}\left[R - n\sum_i\sum_j \bar{y}_{ij.}^{\,-1}\right]$$

$$
\begin{aligned}
\hat{\sigma}_3 &= \frac{1}{abn}\left[R - bn\sum_i \bar{y}_{i..}^{\;-1} - an\sum_j \bar{y}_{.j.}^{\;-1} + abn\,\bar{y}_{...}^{\;-1}\right] \\
\hat{\sigma}_2 &= \frac{1}{abn}\left[R - bn\sum_i \bar{y}_{i..}^{\;-1}\right] \\
\hat{\sigma}_1 &= \frac{1}{abn}\left[R - an\sum_j \bar{y}_{.j.}^{\;-1}\right] \\
\hat{\sigma}_0 &= \frac{1}{abn}\left[R - abn\,\bar{y}_{...}^{\;-1}\right]
\end{aligned} \tag{33}
$$

Let $l_{\max}(\Omega_s)$ denote the maximized log-likelihood under Ω_s, $s = 0,1,\ldots,4$. In general terms, the LR statistic, for testing a null hypothesis Ω_s, nested within the full model Ω_4, is given by

$$
\Lambda_{s4} = 2\left[l_{\max(\Omega_4)} - l_{\max(\Omega_s)}\right] = abn\log\left(\frac{\hat{\sigma}_s}{\hat{\sigma}_4}\right) \tag{34}
$$

When n is large, one can perform a LR test using the asymptotic distribution of the test statistic (34). The rejection region of a level α test would then be set as $\Lambda_{st} \geq \chi^2_\alpha$ where χ^2_α is the upper α point of χ^2 with the degrees of freedom equating the number of parametric constraints imposed by Ω_s. However exact test statistics are available as illustrated in the following.

As in (12),

$$
\Lambda_{s4} = abn\log\left(1 + \frac{\hat{\sigma}_s - \hat{\sigma}_4}{\hat{\sigma}_4}\right) \tag{35}
$$

which is strictly increasing function of $R_{s4} = \dfrac{\hat{\sigma}_s - \hat{\sigma}_4}{\hat{\sigma}_4}$. Consequently, each LR test can equivalently be based on R_{s4}. To examine the individual test statistics, we introduce the following statistics

We can see that

$$
abc(\hat{\sigma}_0 - \hat{\sigma}_4) = \sum_i \sum_j \left(\frac{n}{\bar{y}_{ij.}} - \frac{n}{y_{...}}\right) = Q_0 \text{ (say)}, \tag{36}
$$

Under the assumption of no factor effects, Q_0/σ has χ^2_{ab-1} distribution.

$$abc(\hat{\sigma}_1 - \hat{\sigma}_4) = \sum_j \left(\frac{na}{\bar{y}_{ij.}} - \frac{na}{\bar{y}_{.j.}} \right) = Q_1 \text{ (say)}, \tag{37}$$

Under the assumption of no **A** effects, Q_1/σ has χ^2_{b-1} distribution.

$$abc(\hat{\sigma}_2 - \hat{\sigma}_4) = \sum_i \left(\frac{nb}{\bar{y}_{ij.}} - \frac{nb}{\bar{y}_{i..}} \right) = Q_2 \text{ (say)}, \tag{38}$$

Under the assumption of no **B** effects, Q_2/σ has χ^2_{a-1} distribution.

$$abc(\hat{\sigma}_3 - \hat{\sigma}_4) = \sum_i \sum_j \left(\frac{n}{\bar{y}_{ij.}} - \frac{n}{\bar{y}_{i..}} - \frac{n}{\bar{y}_{.j.}} - \frac{n}{y_{...}} \right) = Q_3 \text{ (say)}, \tag{39}$$

Under the assumption of no **AB** effects, i.e. no interaction, proved that Q_3/σ is asymptotically distributed as $\chi^2_{(a-1)(b-1)}$.

$$abn\,\hat{\sigma}_4 = \sum_i \sum_j \sum_k \left[\frac{n}{y_{ijk}} - \frac{n}{\bar{y}_{ij.}} \right] = Q_4 \text{ (say)}, \tag{40}$$

Q_4/σ has $\chi^2_{ab(n-1)}$ distribution, and is independent of the other Q's.

LR test for no factor effects

The LR test is based on F_{04} where

$$F_{04} = \frac{Q_0/(ab-1)}{Q_4/ab(n-1)}, \tag{41}$$

under the null hypothesis of no factor effects, $F_{04} \sim F_{ab-1,ab(n-1)}$.

LR test for no A effects

The LR test is based on F_{14} where

$$F_{14} = \frac{Q_1/(a-1)}{Q_4/ab(n-1)}, \tag{42}$$

under the null hypothesis of no A effects, $F_{14} \sim F_{a-1,ab(n-1)}$.

LR test for no B effects

The LR test is based on F_{24} where

$$F_{24} = \frac{Q_2 /(b-1)}{Q_4 / ab(n-1)}, \tag{43}$$

under the null hypothesis of no **B** effects, $F_{24} \sim F_{b-1,ab(n-1)}$.

LR test for no interaction effects

The LR test is based on F_{34} where

$$F_{34} = \frac{Q_3 /(a-1)(b-1)}{Q_4 / ab(n-1)} \tag{44}$$

under the null hypothesis of no interaction effect, F_{34} is approximately follows $F_{(a-1)(b-1),ab(n-1)}$.

Table 2, called the analysis of reciprocals (ANOR) table, presents the sums of reciprocals components associated with the various factor effects, and the 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 similarity with the normal theory ANOVA table with sum of reciprocals playing the role of sum of squares.

Table 2 Analysis of Reciprocals (ANOR) Table

Source	*Sum of Reciprocals*	*Degrees of freedom*	*MR*	*F Ratio*
Factor A	Q_1	$a-1$	MR_A	MR_A / MR_E
Factor B	Q_2	$b-1$	MR_B	MR_B / MR_E
Interaction AB	Q_3	$(a-1)(b-1)$	MR_{AB}	MR_{AB} / MR_E *
Residual	Q_4	$ab(n-1)$	MR_E	

* this statistic is distributed approximately as an *F* distribution.

The resemblance between the ANOR for inverse Gaussian model and the normal theory ANOVA is further enhanced by a decomposition of the reciprocal observations $1/y_{ijk}$ into components that can be ascribed to the factor effects. To this end, we formally write out the identities

$$\frac{1}{y_{ijk}}=\frac{1}{\bar{y}_{...}}+\left(\frac{1}{\bar{y}_{ij.}}-\frac{1}{\bar{y}_{.j.}}\right)+\left(\frac{1}{\bar{y}_{ij.}}-\frac{1}{\bar{y}_{i..}}\right)-\left(\frac{1}{\bar{y}_{ij.}}-\frac{1}{\bar{y}_{i..}}-\frac{1}{\bar{y}_{.j.}}+\frac{1}{\bar{y}_{...}}\right)+\left(\frac{1}{y_{ijk}}-\frac{1}{\bar{y}_{ij.}}\right) \quad (45)$$

Summing (44) over i, j and k, we obtain

$$R=abn\,\bar{y}_{...}^{-1}+Q_1+Q_2-Q_3+Q_4 \quad (46)$$

The first term in (44), can be called the general reciprocal mean. The second term represent an A effect, similarly, the third term represent an B effect. The fourth term called an interaction effect.. The term $(y_{ijk}^{-1}-\bar{y}_{ij.}^{-1})$, has the obvious interpretation as a residual provided we bear in mind that the linear model is really on a reciprocal scale. However, unlike the normal theory decomposition, the interaction effect has a negative sign, this is due to the fact that the interaction may be negative. Fries and Bhattacharyya (1983), could not give a complete decomposition of the reciprocal as we demonstrate here, they attribute that to a nonorthogonality component.

Example 1 (continue)

To illustrate the hypothesis testing procedures developed above, we use the same data of example 1 to test for no interactions and for no main effects. The relevant calculations are presented in table 3.

Table 3 The ANOR Table

Source	*Sum of Reciprocals*	*Degrees of freedom*	*MR*	*F Ratio*
Cut	0.379074	1	0.3791	17.0754
Lot	6.822565	4	1.7056	76.8307
Interaction	0.284883	4	0.0712	3.20815
Error	2.0015	90	0.0222	

References

Datta, G. S. (2005), *An alternative derivation of the distributions of the maximum Likelihood estimators of the parameters in an inverse Gaussian distribution,* Biometrika, 94, 4, pp 975-977.

Fries, A. and Bhattacharyya, G. K. (1983), *Analysis of two- factor Experiments Under an Inverse Gaussian Model,* Journal of the American Statistical Association, 78, 384,pp 820-826

Ostel, B. (1963), Statistics in research, Ames: Iowa State University Press.

Seshadri, V. (1993). *The Inverse Gaussian Distribution: A case Study in Exponential Families.* Oxford Science Publications.

Shuster, J.J. and Miura, C. (1972). *Two way analysis of reciprocals.* Biometrika, **59,** 478-81.

Tweedie, M.C.K. (1957), *Statistical properties of inverse Gaussian distributions I.* Annals of Mathematical Statistics, 28, 362-77.

Tweedie, M.C.K. (1957), *Statistical properties of inverse Gaussian distributions II.* Annals of Mathematical Statistics, 28, 696-705.