INVERSE GAUSSIAN REGRESSION AND ACCELERATED LIFE TESTS

Gouri K. Bhattacharyya and Arthur Fries

Department of Statistics, University of Wisconsin, Madison

1. Introduction

A parametric analysis of accelerated life test data largely depends on the model chosen for the distribution of the life time and the relation of the parameter(s) to the stress variable. In addition to the consideration of empirical fit, a life distribution derived from reasonable postulates of the underlying failure process adds credence to its statistical use. The exponential, Weibull and log-normal families have been the popular choices in the extensive literature of engineering applications of accelerated stress testing. The first two draw from the extreme value theory and have simple forms of the failure rate function while the third is capable of using the large resources of the normal theory inference results. In regard to the parameter-stress relation, some empirical engineering models, such as the Arrhenius, Eyring and inverse power law, are ordinarily used. These are cast in a common framework that makes the logarithm of the scale parameter of the life distribution linearly related to the stress. Consequently, the distribution of the log-life is in a location-scale form with a linear regression for the location. Inference procedures under these formulations are discussed in Mann, Schafer and Singpurwalla (1974), Nelson (1971) and others.

This article focuses on a versatile but not so well known life distribution, called the inverse Gaussian distribution $IG(\theta,\lambda)$, whose probability density function (pdf) is given by

$$(1)\qquad f(y;\theta,\lambda) = (2\pi\lambda^{-1}y^{3})^{-\frac{1}{2}} \exp\left\{-\frac{1}{2}\lambda y^{-1}(y\theta^{-1}-1)^{2}\right\},\ y>0,\ \theta>0,\ \lambda>0\ .$$

Its mean and variance are θ and θ^3/λ, respectively. Although it is not a location-scale family, it has the rare confluence of the three desirable features: a wide variety of shapes of the probability density curve, analytical tractability of many inferential results, and most important, its motivation from a plausible stochastic setting of the failure process. Tweedie (1957) studied the basic properties of this distribution, and an extensive literature has evolved in the last decade concerning mainly the one- and two-sample inference procedures and applications. A comprehensive survey is given by Folks and Chhikara (1978).

In the context of accelerated life tests, we refer to the physical basis of the inverse Gaussian distribution in Section 2 in order to formulate a plausible stochastic relation of the failure time to the intensity of stress x. The resulting regression model has the reciprocal-linear form $\theta^{-1} = \alpha + \beta x$ which is explored in the subsequent sections from the aspects of maximum likelihood and least squares estimation. The only previous work in this area is due to Davis (1977) who considers the traditional linear regression of the mean $\theta = \alpha + \beta x$ and assumes λ constant or proportional to θ^2. These formulations are somewhat artificial when viewed in the background of a Wiener process, and except for the uninteresting special case $\alpha = 0$, they lack the analytical advantage of our reciprocal-linear formulation.

2. The Reciprocal-Linear Regression Model

Along the ideas behind the development of the Birnbaum-Saunders (1969) fatigue life distribution, the genesis of the inverse Gaussian distribution can be cast in the context of fatigue growth in a material. Specifically, consider that a material fails when its accumulated fatigue, or depletion of strength exceeds a critical amount $\omega > 0$. Assume that the fatigue growth takes place over time according to a Wiener process with drift $\mu > 0$, and let δ^2 denote the diffusion constant of the process. Then the time to failure (y), alternatively called the first passage time through ω, has the inverse Gaussian distribution $IG(\theta,\lambda)$ with $\theta = \omega\mu^{-1}$ and $\lambda = \omega^2\delta^{-2}$ (cf. Cox and Miller 1965, p. 221).

For a stochastic relation of y to the intensity of stress x, we note that the parameter μ is the most obvious candidate to have a direct relation to x because it measures the mean fatigue growth per unit of time. A linear form $\mu = \alpha + \beta x$ with the natural positivity constraint $\beta \geq 0$ is a simple choice of the relation. On the other hand, the quantities δ and ω correspond, respectively, to the internal variability of the material and the critical damage that identifies a failure. It is therefore reasonable to assume that these are constants unrelated to x. Referring to the pdf (1) and absorbing ω into the parameters α and β, we then have the reciprocal-linear regression structure

$$\begin{gathered} \theta^{-1} = \alpha + \beta x \ , \ \lambda = \text{constant} > 0 \\ \alpha \geq 0, \ \beta \geq 0, \ \alpha + \beta > 0, \ x > 0 \ . \end{gathered} \tag{2}$$

The constancy of λ is analogous to the homoscedasticity assumption in the normal theory linear model. The positivity constraints in (2) are demanded by the fact that the pdf (1) is defined for $0 < \theta < \infty$. In a practical setting, we only require that $\alpha + \beta x > 0$ on a finite interval of x which corresponds to the admissible range of the stress. For a concrete discussion we assume that the origin is taken at the lower end point of this interval so $\alpha \geq 0$.

We now consider the observations (x_i, y_i), $i = 1,\ldots,n$ from n runs of an accelerated life test experiment where y_i denotes the failure time corresponding to the stress setting x_i. The random variables $y_1,\ldots,y_n$ are independent and y_i is distributed as $IG(\theta_i,\lambda)$, with $\theta_i^{-1} = \alpha + \beta x_i$.

Referring to (1), the log-likelihood function is given by

$$
\begin{aligned}
\ell &= -\frac{n}{2}\log(2\pi\lambda^{-1}) - \frac{\lambda}{2}\sum_{i=1}^{n} y_i^{-1}\left\{y_i(\alpha+\beta x_i)-1\right\}^2 - \frac{3}{2}\sum_{i=1}^{n}\log y_i \\
&= c(\alpha,\beta,\lambda) - \left\{\frac{\lambda\alpha^2}{2}\sum y_i + \lambda\alpha\beta\sum x_i y_i + \frac{\lambda\beta^2}{2}\sum x_i^2 y_i + \frac{\lambda}{2}\sum y_i^{-1}\right\} - \frac{3}{2}\sum \log y_i \ ,
\end{aligned} \tag{3}
$$

which is defined on the restricted parameter space

$$\Omega = \{(\alpha,\beta,\lambda):\ \alpha \geq 0,\ \beta \geq 0,\ \alpha+\beta > 0,\ \lambda > 0\} \ .$$

The second form of (3) shows that we have an exponential family of distributions with four-dimensional sufficient statistics. However, the natural parameter space being only three-dimensional, the standard theories of inference for the exponential families do not readily apply.

3. Maximum Likelihood Estimation

For the moment we disregard the restricted form of the parameter space and consider maximization of ℓ with respect to $\underset{\sim}{\psi} = (\alpha,\beta,\lambda)'$. The first expression in (3) yields the likelihood equations

$$
\begin{aligned}
\partial\ell/\partial\alpha &= \lambda\sum_{i=1}^{n}\{1-(\alpha+\beta x_i)y_i\} = 0 \\
\partial\ell/\partial\beta &= \lambda\sum_{i=1}^{n} x_i\{1-(\alpha+\beta x_i)y_i\} = 0 \\
\partial\ell/\partial\lambda &= \frac{1}{2}n\lambda^{-1} - \frac{1}{2}\sum_{i=1}^{n} y_i^{-1}\{1-(\alpha+\beta x_i)y_i\}^2 = 0 \ ,
\end{aligned} \tag{4}
$$

We introduce convenient notation for the basic statistics:

$$V_j = n^{-1} \sum_{i=1}^{n} y_i x_i^j \ , \ j = 0,1,2$$

(5)

$$\bar{x} = n^{-1} \sum_{i=1}^{n} x_i \ , \quad \bar{y} = n^{-1} \sum_{i=1}^{n} y_i \ , \quad R = n^{-1} \sum_{i=1}^{n} y_i^{-1} \ .$$

The first two equations in (4) simplify to

$$\alpha V_0 + \beta V_1 = 1 \ , \quad \alpha V_1 + \beta V_2 = \bar{x} \quad ,$$

which are linear in the parameters as are the corresponding likelihood equations under the usual normal theory linear regression model. Interestingly however, the coefficients on the left side are random variables, each a linear function of $\underset{\sim}{y}$, while the terms on the right side are nonrandom. In the normal regression case, the situation is reversed. The likelihood equations yield the unique root $\hat{\underset{\sim}{\psi}}_L = (\hat{\alpha}_L, \hat{\beta}_L, \hat{\lambda}_L)'$ given by

$$\hat{\alpha}_L = (V_2 - \bar{x}V_1)\, D^{-1} = \bar{y}^{-1}(1 - V_1\hat{\beta}_L)$$

(6)

$$\hat{\beta}_L = (\bar{x}V_0 - V_1)\, D^{-1} = -(nD)^{-1} \sum_{i=1}^{n} (x_i - \bar{x})(y_i - \bar{y})$$

$$\hat{\lambda}_L^{-1} = n^{-1} \sum_{i=1}^{n} (y_i^{-1} - \hat{\alpha}_L - \hat{\beta}_L x_i) = R - \hat{\alpha}_L - \hat{\beta}_L \bar{x} \quad ,$$

where $D \equiv V_0 V_2 - V_1^2 > 0$ with probability 1 by the Cauchy-Schwarz inequality. Observe that $\hat{\beta}_L$ involves the usual covariance term in its numerator, but its denominator is quadratic in $\underset{\sim}{y}$. Further

$$-\left.\frac{\partial^2 \ell}{\partial \underset{\sim}{\psi} \partial \underset{\sim}{\psi}'}\right|_{\hat{\underset{\sim}{\psi}}_L} = n\hat{\lambda}_L \begin{bmatrix} V_0 & V_1 & 0 \\ V_1 & V_2 & 0 \\ 0 & 0 & \frac{1}{2}\hat{\lambda}_L^{-3} \end{bmatrix} .$$

Since this is positive definite, $\hat{\underset{\sim}{\psi}}_L$ locates the unique maximum of ℓ. We will call $\hat{\underset{\sim}{\psi}}_L$ the maximum likelihood root estimator (MLRE).

In order to obtain the maximum likelihood estimator (MLE), one needs to examine whether or not the root lies in the parameter space Ω. The last equation in (4) shows that $\hat{\lambda}_L > 0$. Also from the reduced versions of the first two, it is clear that at most one of $\hat{\alpha}_L$ and $\hat{\beta}_L$ can be negative. Thus, a violation of the constraints can occur in no more than one component of $\hat{\underset{\sim}{\psi}}_L$. It is easy to construct examples where one of $\hat{\alpha}_L$ and $\hat{\beta}_L$ can indeed be negative. In such cases, a search for the MLE requires a maximization of ℓ on the boundaries $\alpha = 0$ and $\beta = 0$. Substituting $\alpha = 0$ in (3), we find that ℓ is maximized at $\hat{\beta}_a = \bar{x}V_2^{-1}$, $\hat{\lambda}_a^{-1} = R - \bar{x}^2 V_2^{-1}$, and its maximum value is $\ell_a = c - n/2 + (n/2) \log \hat{\lambda}_a$ where c is a function of $\underset{\sim}{y}$. Similarly, with $\beta = 0$, the maximum value of ℓ is $\ell_b = c - n/2 + (n/2) \log \hat{\lambda}_b$ which is attained at $\hat{\alpha}_b = V_0^{-1}$ and $\hat{\lambda}_b^{-1} = R - V_0^{-1}$. In either case, the maximizing solutions are positive and $\hat{\lambda}_{\bullet}^{-1}$ has the form $R - \hat{\alpha}_{\bullet} - \hat{\beta}_{\bullet}\bar{x}$. Finally, a comparison of these maximized likelihoods shows that $\ell_a \geq \ell_b$ if and only if $\bar{x}^2 V_0 \geq V_2$. Collecting these results together, a formal characterization of the MLE $\hat{\underset{\sim}{\psi}}$ can be stated as follows:

$$(7) \qquad \begin{aligned} (\hat{\alpha},\hat{\beta}) &= (\hat{\alpha}_L,\hat{\beta}_L) && \text{if } V_1 < \min(\bar{x}V_0, \bar{x}^{-1}V_2) \\ &= (0, \bar{x}V_2^{-1}) && \text{if } \bar{x}^{-1}V_2 \leq V_1 < \bar{x}V_0 \\ &= (V_0^{-1}, 0) && \text{if } \bar{x}V_0 \leq V_1 < \bar{x}^{-1}V_2 \\ \hat{\lambda}^{-1} &= R - \hat{\alpha} - \hat{\beta}\bar{x} \ , \end{aligned}$$

where $\hat{\alpha}_L$ and $\hat{\beta}_L$ are given in (6). Note that if the root $\hat{\alpha}_L < 0$, the MLE of α gets pulled to zero, as one would anticipate. However, at the same time, $\hat{\beta}$ changes its functional form. The situation is similar when $\hat{\beta}_L < 0$.

Although the inverse Gaussian is not a location-scale parameter family, certain equivariance properties hold for the MLE's under scale changes in x or y, and location change in x. To describe these, we refer to the MLE's as functions of $\underset{\sim}{x}$ and $\underset{\sim}{y}$ by writing, for instance, $\hat{\alpha}(\underset{\sim}{x},\underset{\sim}{y})$ in place of $\hat{\alpha}$, and consider first the effects of scalar multiplications. For constants $d_1 > 0$ and $d_2 > 0$, the following relations hold

$$\begin{aligned}
\hat{\alpha}(d_1\underset{\sim}{x},\ d_2\underset{\sim}{y}) &= d_2^{-1}\ \hat{\alpha}(\underset{\sim}{x},\underset{\sim}{y}) \\
\hat{\beta}(d_1\underset{\sim}{x},\ d_2\underset{\sim}{y}) &= d_1^{-1}\ d_2^{-1}\ \hat{\beta}(\underset{\sim}{x},\underset{\sim}{y}) \\
\hat{\lambda}(d_1\underset{\sim}{x},\ d_2\underset{\sim}{y}) &= d_2\ \hat{\lambda}(\underset{\sim}{x},\underset{\sim}{y}) \quad .
\end{aligned}$$

Also, the effect of a translation of $\underset{\sim}{x}$ on the MLRE's is readily apparent. If $\underset{\sim}{x}$ is changed to $\underset{\sim}{x} + x_0\underset{\sim}{1}$ where $x_0 > -\min(x_1,\ldots,x_n)$, then $\hat{\alpha}_L$ changes to $\hat{\alpha}_L - x_0\hat{\beta}_L$ while $\hat{\beta}_L$ remains unchanged. However, $\hat{\alpha}$ and $\hat{\beta}$ could change their functional forms. For instance, if with the original $\underset{\sim}{x}$, $\hat{\alpha} > 0$, $\hat{\beta} > 0$ and if $x_0 > \hat{\alpha}\hat{\beta}^{-1}$, then we have $\hat{\beta} = (\bar{x}V_0 - V_1)D^{-1}$, whereas with $\underset{\sim}{x}$ translated to $\underset{\sim}{x} + x_0\underset{\sim}{1}$, the new $\hat{\alpha}$ is 0 and the new $\hat{\beta}$ is $(\bar{x} + x_0)(V_2 + 2x_0V_1 + x_0^2V_0)^{-1}$. The MLE $\hat{\lambda}$ changes accordingly. These properties help relate the MLE's under different choices of scales for the x and y variables.

4. Asymptotic Theory

Strong consistency and asymptotic normality of the maximum likelihood estimators are established in this section under some mild conditions on the limiting behavior of the design points. Letting $F_n(x) = \#\{x_i \le x,\ i=1,\ldots,n\}/n$, we henceforth assume that, as $n \to \infty$, F_n converges weakly to a proper distribution function F on $(0,\infty)$. For brevity, we will only treat the case when the true parameter point $\underset{\sim}{\psi}$ is in the interior of Ω; that is, neither α nor β is 0. Unless specified otherwise, all limits are taken as $n \to \infty$. We introduce the notation

$$\tau_j(n) = \int_0^\infty x^j (\alpha + \beta x)^{-1} \, dF_n \ , \quad \tau_j = \int_0^\infty x^j \ (\alpha + \beta x)^{-1} \, dF$$

(8)

$$c_j(n) = \int_0^\infty x^j \, dF_n \ , \quad c_j = \int_0^\infty x^j \, dF \ , \quad j = 0,1,2 \quad .$$

LEMMA. Assume that x^2 is uniformly integrable in F_n. Then R converges almost surely (a.s.) to $\alpha + \beta c_1 + \lambda^{-1}$, and V_j to τ_j, $j = 0,1,2$.

PROOF. Since $E(y_i) = (\alpha + \beta x_i)^{-1}$ and $E(y_i^{-1}) = \alpha + \beta x_i + \lambda^{-1}$ (cf. Tweedie, 1957), we have $E(V_j) = \tau_j(n)$ and $E(R) = \alpha + \beta c_1(n) + \lambda^{-1}$. The assumptions $\alpha > 0$ and $\beta > 0$ imply that $x^j(\alpha + \beta x)^{-1} \leq \min\{\alpha^{-1}x^j, \ \beta^{-1}x^{j-1}\}$, so $\lim \tau_j(n) = \tau_j$ and $\lim E(R) = \alpha + \beta c_1 + \lambda^{-1}$ by the uniform integrability of x. Noting that $\mathrm{Var}(y_i^{-1}) = \lambda^{-1}(\alpha + \beta x_i + 2\lambda^{-1})$, the stated a.s. convergence of R would follow by an application of the Kolmogorov strong law once we verify that $\sum_{i=1}^{\infty} x_i/i^2 < \infty$. Because $c_2(n) \to c_2 < \infty$, for a given $\varepsilon > 0$ there exists an n_0 such that $n^{-1} x_n^2 \leq \varepsilon$ for $n \geq n_0$. Consequently,

$$\sum_{i=1}^{\infty} x_i/i^2 \leq \sum_{i=1}^{n_0} x_i/i^2 + \varepsilon^{1/2} \sum_{i=n_0+1}^{\infty} i^{-3/2} < \infty \quad ,$$

which establishes the desired result. For an application of the strong law to V_2 we require that $\sum_{i=1}^{\infty} x_i^4 \, \mathrm{Var}(y_i) i^{-2} < \infty$. This follows from the facts that $\mathrm{Var}(y_i) = \lambda^{-1}(\alpha + \beta x_i)^{-3}$ and $x^4(\alpha + \beta x)^{-3} \leq \beta^{-3}x$. The treatment of V_0 and V_1 are similar, and the proof is concluded.

To show that both $\hat{\underset{\sim}{\psi}}_L$ and $\hat{\underset{\sim}{\psi}}$ are strongly consistent for $\underset{\sim}{\psi}$, we first refer to (6) and use the lemma to each component of $\hat{\underset{\sim}{\psi}}_L$. In particular, $\hat{\beta}_L \to g(\underset{\sim}{\tau})$, a.s. where

$$g(\underset{\sim}{\tau}) = (c_1\tau_0 - \tau_1)(\tau_0\tau_2 - \tau_1^2)^{-1} \quad .$$

The relations $\beta\tau_1 = 1 - \alpha\tau_0$ and $\beta\tau_2 = c_1 - \alpha\tau_1$ can be verified from the definition of τ_j, and a substitution of these yields $g(\tau) = \beta$, so $\hat{\beta}_L \to \beta$, a.s. Referring to (7) we observe that $\hat{\beta}_L \neq \hat{\beta}$ if and only if $\bar{x}V_0 - V_1 < 0$. However, using the lemma we have with probability one, $\bar{x}V_0 - V_1 \to c_1\tau_0 - \tau_1 > 0$. Thus, $\lim\sup|\hat{\beta}_L - \hat{\beta}| = 0$, with probability one, hence $\hat{\beta} \to \beta$, a.s. The proofs for $\hat{\alpha}$ and $\hat{\lambda}$ are analogous.

We now turn to the limiting distribution of the maximum likelihood estimators. In view of the asymptotic a.s. equivalence of $\hat{\underset{\sim}{\psi}}$ and $\hat{\underset{\sim}{\psi}}_L$, it suffices to consider the limiting distribution of $\hat{\underset{\sim}{\psi}}_L$. Referring to (6), the most direct approach would be to first establish the joint asymptotic normality of the V_j's and R and then use the δ-method. However, this process incurs some formidable expressions whose simplifications are quite tedious. Instead, we examine the first and second partial derivatives of the log-likelihood, and observe an interesting relation

$$n^{-1/2}\partial\ell/\partial\underset{\sim}{\psi} = \underset{\sim}{M}[n^{1/2}(\hat{\underset{\sim}{\psi}}_L - \underset{\sim}{\psi})] \tag{9}$$

where

$$\underset{\sim}{M} = \begin{bmatrix} \lambda\underset{\sim}{S} & \underset{\sim}{0} \\ \underset{\sim}{e}' & (\lambda\hat{\lambda}_L)^{-1}/2 \end{bmatrix} \quad , \quad \underset{\sim}{S} = \begin{bmatrix} V_0 & V_1 \\ V_1 & V_2 \end{bmatrix}$$

$$\underset{\sim}{e}' = (e_1, e_2)$$

$$e_1 = -1 + \tfrac{1}{2}\{V_0(\hat{\alpha}_L + \alpha) + V_1(\hat{\beta}_L + \beta)\}$$

$$e_2 = -\bar{x} + \tfrac{1}{2}\{V_1(\hat{\alpha}_L + \alpha) + V_2(\hat{\beta}_L + \beta)\} \quad .$$

The exact relation (9) is more convenient to work with than the usual Taylor expansion of $\partial\ell/\partial\underset{\sim}{\psi}$ which involves the second derivatives evaluated at some undetermined intermediate point between $\hat{\underset{\sim}{\psi}}_L$ and $\underset{\sim}{\psi}$.

THEOREM. Assume that x^3 is uniformly integrable in F_n. Then $n^{1/2}(\hat{\underset{\sim}{\psi}}_L - \underset{\sim}{\psi})$ is asymptotically trivariate normal $N_3(\underset{\sim}{0},\underset{\sim}{\Sigma})$ where

$$(10) \qquad \underset{\sim}{\Sigma}^{-1} = \begin{bmatrix} \lambda\underset{\sim}{\Delta} & \underset{\sim}{0} \\ \underset{\sim}{0}' & \lambda^{-2}/2 \end{bmatrix}, \qquad \underset{\sim}{\Delta} = \begin{bmatrix} \tau_0 & \tau_1 \\ \tau_1 & \tau_2 \end{bmatrix} .$$

PROOF. As a consequence of the lemma, $e_i \to 0$, a.s. for $i = 1,2$, so $\underset{\sim}{M} \to \underset{\sim}{\Sigma}^{-1}$, a.s. Therefore, it suffices to show that $n^{-1/2}\partial\ell/\partial\underset{\sim}{\psi}$ has the limiting distribution $N_3(\underset{\sim}{0},\underset{\sim}{\Sigma}^{-1})$. Referring to (4) we observe that $\partial\ell/\partial\underset{\sim}{\psi}$ is of the form $\sum_{i=1}^{n} \underset{\sim}{q}_i$ where $\underset{\sim}{q}_i$'s are independent but non-identically distributed random vectors with means $\underset{\sim}{0}$. Focusing on an arbitrary linear function $Z_n = n^{-1/2}\, \underset{\sim}{h}'\, \partial\ell/\partial\underset{\sim}{\psi}$ with $\underset{\sim}{h} \neq \underset{\sim}{0}$ we readily observe that $\lim \operatorname{Var}(Z_n) = \underset{\sim}{h}'\underset{\sim}{\Sigma}^{-1}\underset{\sim}{h} > 0$. Therefore, the Liapounov central limit theorem would apply once we establish that $\lim \zeta_n = 0$ where

$$\zeta_n = n^{-3/2} \sum_{i=1}^{n} E|\underset{\sim}{h}'\underset{\sim}{q}_i|^3 .$$

Denoting $\omega_i = y_i(\alpha + \beta x_i)$ and referring to (4), one can find constants $a_j \geq 0$ depending only on $\underset{\sim}{h}$, α, β and λ such that

$$|\underset{\sim}{h}'\underset{\sim}{q}_i|^3 \leq (1 + x_i)^3 \sum_{j=0}^{6} a_j \omega_i^{j-3} .$$

Consequently, $\zeta_n \leq n^{-1/2} \sum_{j=0}^{6} a_j \varepsilon_n$ where $\varepsilon_n = n^{-1} \sum_{i=1}^{n} (1 + x_i)^3\, E(\omega_i^{j-3})$.

Using the positive and negative moments of the inverse Gaussian distribution (cf. Tweedie, 1957), it can be seen that ε_n is a fixed (not depending on n) linear combination of the terms

$$\nu_{jk}(n) = \int_0^\infty x^k(\alpha+\beta x)^{-j}\, dF_n(x)$$

with $0 \leq j, k \leq 3$. From the uniform integrability assumption, each of these $\nu_{jk}(n)$ has a finite limit. Hence $\lim \zeta_n = 0$ and the proof is concluded.

The limiting normal distribution of the MLE's, along with the a.s. convergence of the sample information matrix, can be used to construct large sample confidence intervals for the parameters. Specifically, the approximate variances of $\hat{\alpha}, \hat{\beta}$ and $\hat{\lambda}$ are $(n\hat{\lambda}D)^{-1}V_2$, $(n\hat{\lambda}D)^{-1}V_0$ and $2n^{-1}\hat{\lambda}^2$, respectively. Also, the reciprocal mean failure time $[\theta(x^*)]^{-1} = \alpha+\beta x^*$, at a specified stress level x^*, is estimated by $\hat{\alpha}+\hat{\beta}x^*$, whose approximate variance is given by $(nD\hat{\lambda})^{-1}$ $(V_2 - 2x^* V_1 + x^{*2}V_0)$.

5. A Least Squares Approach for Replicated Designs

Although closed form expressions were obtained in Section 3, an analytical treatment of the exact mean and variance of the MLE's does not appear to be feasible. In this section, we consider experiments with replicated observations and construct some unbiased estimators by employing a combination of the maximum likelihood and least squares principles. A similar procedure has been used by Singpurwalla (1973) when the underlying life distribution is exponential.

Consider k stress settings $x_1,\ldots,x_k$ and n_i independent failure times $(y_{i1},\ldots,y_{in_i})$ observed at x_i, $i = 1,\ldots,k$. The random variables y_{ij}, $j = 1,\ldots, n_i$; $i = 1,\ldots,k$ are all independent with y_{ij} distributed as $IG(\theta_i,\lambda)$, and $\theta_i^{-1} = \alpha+\beta x_i$. Let

$$\bar{y}_i = n_i^{-1} \sum_{j=1}^{n_i} y_{ij} \quad , \quad N = \sum_{i=1}^{k} n_i \quad ,$$

$$(11) \qquad Q = \sum_{i=1}^{k} \sum_{j=1}^{n_i} (y_{ij}^{-1} - \bar{y}_i^{-1}) \quad .$$

Our method consists of two steps. Disregarding the regression structure, we first consider the θ_i's as free parameters in which case $\bar{y}_i$, $i=1,\ldots,k$ and Q constitute a set of complete sufficient statistics. As shown by Tweedie (1957), these statistics are all independent, $\bar{y}_i$ is distributed as $IG(\theta_i, n_i\lambda)$ and λQ is distributed as χ^2 with (N-k) degrees of freedom. Moreover $E(\bar{y}_i^{-1}) = \theta_i^{-1} + (n_i\lambda)^{-1}$ and $Var(\bar{y}_i^{-1}) = (\theta_i n_i \lambda)^{-1} + 2(n_i\lambda)^{-2}$. Defining

$$\tilde{\lambda} = (N-k)\, Q^{-1} \quad ,$$

$$t_i = \bar{y}_i^{-1} - (n_i\tilde{\lambda})^{-1} \quad ,$$

it then follows that t_i and $\tilde{\lambda}^{-1}$ are the uniformly minimum variance unbiased estimators of θ_i^{-1} and λ^{-1}, respectively. Focusing now on the t_i's which are, in essence, the bias-corrected reciprocal means, we have the linear model $E(t_i) = \alpha + \beta x_i$ with the covariance structure

$$Var(t_i) = (\alpha + \beta x_i)(n_i\lambda)^{-1} + 2(n_i\lambda)^{-2}\,[1 + (N-k)^{-1}] \quad ,$$

$$Cov(t_i, t_{i'}) = 2[n_i n_{i'}(N-k)]^{-1}\,\lambda^{-2} \quad , \quad i \neq i' \quad .$$

With large n_i's, the covariances are negligible compared to the variances. The leading term in $Var(t_i)$ is proportional to n_i^{-1} but its dependence on $(\alpha + \beta x_i)$ is the major deterrant to an application of the usual weighted least squares. In order to get unbiased estimators we consider the particular weighted least squares which minimizes $\sum_{i=1}^{k} n_i(t_i - \alpha - \beta x_i)^2$. Letting $m_j = N^{-1} \sum_{i=1}^{k} n_i x_i^j$, $j \geq 1$, the resulting unbiased estimators are

$$\tilde{\alpha} = N^{-1} \sum_{i=1}^{k} n_i \bar{y}_i^{-1} - \tilde{\beta} m_1 + \tilde{\lambda}^{-1} k N^{-1} \quad ,$$

$$\tilde{\beta} = (m_2 - m_1^2)^{-1} N^{-1} \sum_{i=1}^{k} (x_i - m_1)(n_i \bar{y}_i^{-1} - \tilde{\lambda}^{-1}) \quad .$$

It can be seen that the behaviors of $\tilde{\psi} = (\tilde{\alpha}, \tilde{\beta}, \tilde{\lambda})'$ under different scalings of x and/or y are identical to those of $\hat{\psi}_L$. For the balanced design $n_1 = \cdots = n_k$, $\tilde{\beta}$ is independent of $\tilde{\lambda}$. The exact χ^2 distribution of $(N-k)\lambda\tilde{\lambda}^{-1}$ and the simple linear forms of $\tilde{\alpha}$ and $\tilde{\beta}$ also enable us to derive the exact variance and covariance expressions. Dropping terms of order N^{-2}, we obtain

$$\mathrm{Var}(\tilde{\lambda}^{-1}) = 2(N-k)^{-1} \lambda^{-2}, \ \mathrm{Cov}(\tilde{\alpha}, \tilde{\lambda}^{-1}) = \mathrm{Cov}(\tilde{\beta}, \tilde{\lambda}^{-1}) = 0 \ ,$$

$$\begin{aligned} \mathrm{Var}(\tilde{\alpha}) &= (N\lambda)^{-1} s_2^{-2} [\alpha m_2 s_2 + \beta m_1 s_3] \ , \\ \mathrm{Var}(\tilde{\beta}) &= (N\lambda)^{-1} s_2^{-2} [\alpha s_2 + \beta m_1^{-1} (s_2^2 + s_3)] \ , \\ \mathrm{Cov}(\tilde{\alpha}, \tilde{\beta}) &= -(N\lambda)^{-1} s_2^{-2} [\alpha m_1 s_2 + \beta s_3] \ , \end{aligned} \tag{12}$$

where $s_2 = m_2 - m_1^2$ and $s_3 = m_1 m_3 - m_2^2$.

The asymptotic (as $N \to \infty$, $n_i N^{-1} \to r_i$, and k fixed) properties of $\tilde{\psi}$ follow directly from the asymptotics for Q and $\bar{y}_i$, $i = 1, \ldots, k$. In particular, $\tilde{\psi}$ is strongly consistent, and $N^{1/2}(\tilde{\psi} - \psi)$ is asymptotically distributed as $N_3(0, \Gamma)$ where the entries of Γ are equal to the limits of N times the corresponding expressions given in (12).

Referring to (12) and the covariance matrix Σ given in (10), we obtain the asymptotic efficiencies (AE's) of the least squares estimators. Specifically,

$$AE(\tilde{\lambda}) = 1,$$

$$AE(\tilde{\alpha}) = \tau_2(\tau_2 - c_1\tau_1)^{-1}(c_2 - c_1^2)^2\{c_2(c_2 - c_1^2) + \nu(c_1c_3 - c_2^2)\}^{-1} ,$$

$$AE(\tilde{\beta}) = \tau_0(\tau_2 - c_1\tau_1)^{-1}(c_2 - c_1^2)^2\{(c_2 - c_1^2) + \nu(c_1^2 - 2c_2 + c_1^{-1}c_3)\}^{-1} ,$$

where $\nu = \beta\alpha^{-1}c_1$ and $\tau_j = \alpha^{-1}\sum_{i=1}^{k} r_i x_i^{\,j}(1 + \nu c_1^{-1}x_i)^{-1}$, $c_j = \sum_{i=1}^{k} r_i x_i^{\,j}$, $j = 0,1,2,3$.

For fixed c_j's, these AE's are monotonically decreasing functions of ν. Their bounds can therefore be established by considering the limits $\beta \to 0$ and $\alpha \to 0$. As $\beta \to 0$, $\tau_j \to \alpha^{-1}c_j$ and the limiting AE's are each equal to 1. By continuity, the AE's are high when α is much larger than βc_1, i.e., when the major contribution to the reciprocal mean lifetime, at the center of the design, is due to α. As $\alpha \to 0$, $\tau_j \to \beta^{-1}c_{j-1}$ and the lower bounds are given by

$$AE(\tilde{\alpha}) \geq \frac{(c_2 - c_1^2)^2}{(c_1c_3 - c_2^2)(c_1c_{-1} - 1)} , \quad AE(\tilde{\beta}) \geq \frac{c_{-1}(c_2 - c_1^2)^2}{(c_1^3 - 2c_1c_2 + c_3)(c_1c_{-1} - 1)} .$$

Here, $c_{-1} = \sum_{i=1}^{k} r_i x_i^{-1}$ is well-defined since we have assumed all $x_i > 0$. When some x_i values are close to 0, the quantity c_{-1} gets large and it forces the lower bounds to become small. For example, the three point design with $(x_1, x_2, x_3) = (0.1, 1.0, 2.8)$ and $(r_1, r_2, r_3) = (.4, .4, .2)$ gives $AE(\tilde{\alpha}) \geq .311$ and $AE(\tilde{\beta}) \geq .679$. This drawback of the least squares estimators can be overcome by suitably translating the x-scale. For instance, in the previous example, the translation to $\underset{\sim}{x}_* = \underset{\sim}{x} + \underset{\sim}{1}$ gives the lower bounds .866 and .888 for $AE(\tilde{\alpha})$ and $AE(\tilde{\beta})$, respectively.

Straightforward computations show that, independent of the design, the AE of the estimated reciprocal mean lifetime at the center of the design is equal to 1, i.e., $AE(\tilde{\alpha} + \tilde{\beta}c_1) = 1$. Also, for any two point design we have $AE(\tilde{\alpha}) = AE(\tilde{\beta})$

= 1. This last result is explained by noting that when $k=2$, $\tilde{\alpha}=\hat{\alpha}_L+N^{-1}\tilde{\lambda}^{-1}b_1$ and $\tilde{\beta}=\hat{\beta}_L+N^{-1}\tilde{\lambda}^{-1}b_2$ where b_1 and b_2 are constants. Evidently, the least squares estimators are then asymptotically equivalent to the MLE's.

6. Example

Nelson (1971) reports data on the failure times of an insulation material in a motorette test performed at four elevated temperature settings ranging from 190^0C to 260^0C. The original goal of the experiment was to determine if the mean time to failure at the design temperature of 180^0C exceeded a specified minimum requirement. However, the 260^0C data were taken on a batch of insulation different from the batch used at the other temperature settings, and it became important to investigate whether or not the data from the two batches were consistent. Nelson establishes the inconsistency of the 260^0C data by employing a combination of graphical and analytic techniques based upon the assumptions that the log-failure times are normally distributed with a constant variance and the mean depending on temperature through the Arrhenius relationship.

For illustrative purposes, we fit an IG regression model to these data excluding the 260^0C setting. Although the physical properties of the insulation material are unknown to us, it seems reasonable to hypothesize that the wear of the insulation increases until a critical amount has disintegrated or ceased to perform adequately. Thus the assumption of an underlying IG distribution is plausible. The assumption of a constant λ is tenable since the materials used in the first three levels are known to have a common source. In the absence of a mechanistic model relating the mean life (θ) and the temperature (T), we base our choice on the following observations: (i) regressing $\bar{y}_i^{-1}$ on T_i^3 gives a value of $R^2=99.9\%$, (ii) consistent with the IG assumption, the sample variances raised to the power $-\frac{1}{3}$ are approximately linear in T_i^3, and (iii) the transformation to $T_i^3-(180)^3$ is convenient in that it produces positive MLE's. Without a translation of T_i^3, the restriction $\alpha \geq 0$ imposed in the definition of Ω would not be meaningful.

Guided by these considerations, we take the distribution of failure times as $IG(\theta,\lambda)$ with $\theta^{-1} = \alpha + \beta x$ where $x = 10^{-8}[T^3 - 180^3]$. By using the results of Sections 3, 4, and 5, we calculate the maximum likelihood and least squares estimates of α, β and λ^{-1} and also their approximate standard errors. The results are given in Table 1 where the standard errors are shown in parentheses.

TABLE 1. The estimates and approximate standard errors

	Parameters		
	α	β	λ^{-1}
Maximum likelihood	.0371 (.0129)	7.3260 (.3557)	.0102 (.0026)
Least squares	.0320 (.0141)	7.4316 (.3747)	.0097 (.0026)

The MLE's of the mean lives (thousands of hours) at 190^0C, 220^0C, and 240^0C are, respectively, 8.902, 2.565 and 1.606, with the associated standard errors .094, .029 and .033. The corresponding least squares estimates are 8.863, 2.565 and 1.598, with the respective standard errors .101, .030 and .034. Both sets of estimates are comparable to the observed sample means 8.782, 2.638 and 1.581.

The MLE's obtained exclusively from the 260^0C data, specifically $\bar{y}_4$ and $\hat{\lambda}_4^{-1}$, are now compared with the estimates computed using only the first three levels. From Table 1, the estimates of the mean time to failure at the fourth level are $\hat{\theta}_4 = 1.112$ and $\tilde{\theta}_4 = 1.105$ with the respective standard errors .037 and .038. These estimates agree closely with the sample mean $\bar{y}_4 = 1.116$. However, $\hat{\lambda}_4^{-1} = .1310$ differs substantially from both $\hat{\lambda}^{-1}$ and $\tilde{\lambda}^{-1}$. We perform an exact test of the hypothesis of a common λ by using $\hat{\lambda}_4$ and $\tilde{\lambda}$. Under the null hypothesis of a common λ, $10\lambda\hat{\lambda}_4^{-1}$ has an exact $\chi^2(9)$ distribution and it is statistically independent of $27\lambda\tilde{\lambda}^{-1}$ which is distributed as $\chi^2(27)$. Thus, $10\hat{\lambda}_4^{-1}(9\tilde{\lambda}^{-1})^{-1}$ is distributed as $F(9,27)$. The observed value of this statistic is 15.01 which corresponds to a p-value <.001. In agreement with Nelson, we conclude that the two batches of insulating materials are significantly different.

ACKNOWLEDGEMENT

This work was supported by the Office of Naval Research Grant N00014-78-C-0722.

REFERENCES

Birnbaum, Z.W. and Saunders, S.C. (1969). A new family of life distributions. Journal of Applied Probability 6, 319-327.

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

Davis, A.S. (1977). Linear statistical inference as related to the inverse Gaussian distribution. Unpublished PhD thesis, Oklahoma State University.

Folks, J.L. and Chhikara, R.S. (1978). The inverse Gaussian distribution and its statistical application - a review. Journal of the Royal Statistical Society B 40, 263-275.

Mann, N.R., Schafer, R.E. and Singpurwalla, N.D. (1974). Methods for Statistical Analysis of Reliability and Life Data. Wiley, New York.

Nelson, W.B. (1971). Analysis of accelerated life test data. I.E.E.E. Transactions on Electrical Insulation EI-6, 165-181.

Singpurwalla, N.D. (1973). Inference from accelerated life tests using Arrhenius type re-parametrizations. Technometrics 15, 289-299.

Tweedie, M.C.K. (1957). Statistical properties of inverse Gaussian distributions. Annals of Mathematical Statistics 28, 362-377.