Parameter Estimation with Least-Squares Method for the Inverse Gaussian distribution Model Using Simplex and Quasi-Newton Optimization Methods

Khizar Hayat Khan

Department of Mathematics, College of Science and Humanities, Prince Sattam Bin Abdulaziz University, Al-Kharj, Kingdom of Saudi Arabia

Abstract

We find Survival rate estimates; parameter estimates for the inverse Gaussian distribution model using least-squares estimation method. We found these estimates for the case when partial derivatives were available and for the case when partial derivatives were not available. The simplex optimization (Nelder and Mead, and Hooke and Jeeves) methods were used for the case when first partial derivatives were not available and the Quasi – Newton optimization (Davidon-Fletcher-Powel (DFP) and the Broyden-Fletcher-Goldfarb-Shanno (BFGS) methods were applied for the case when first partial derivatives were available. The medical data sets of 21 Leukemia cancer patients with time span of 35 weeks were used.

Keywords

Inverse Gaussian distribution model, Nelder and Mead, and Hooke and Jeeves, DFP and BFGS optimization methods, Parameter estimation, Least Square method, Kaplan-Meier estimates, Survival rate Estimates, Variance-Covariance matrix.

1. Introduction

The methods of Nelder and Mead [1] and Hooke and Jeeves [2] do not require first or higher partial derivatives and in these methods, the function values are compared to find the optimal value of the objective function. Whereas, the Quasi –Newton Methods (DFP and BFGS) require only first partial derivatives of the function of n variables [3], we find the survivor rate estimates, parameter estimates, optimal function values using simplex optimization methods and Quasi-Newton optimization methods. These optimization methods were applied using medical data sets of cancer patients [4].

The method of linear least squares requires that a straight line be fitted to a set of data points such that the sum of squares of the vertical deviations from the points to be minimized [5], [7].

Adrien Merie Legendre (1752-1833) is generally credited for creating the basic ideas of the method of least squares. Some people believe that the method was discovered at the same time by Karl F. Gauss (1777-1855), Pierre S. Laplace (1749-1827) and others. Furthermore, Markov's name is also included for further development of these ideas. In recent years, [6], [9] an effort have been made to find better methods of fitting curves or equations to data, but the least-squares method remained dominant, and is used as one of the important methods of estimating the parameters. The least-squares method consists of finding those parameters that minimize a particular objective function based on squared deviations.

It is to be noted that for the least-squares estimation method, we are interested to minimize some function of the residual, that is, we want to find the best possible agreement between the observed and the estimated values. To define the objective function F, we set up a vector of residuals

$$r_i = y_i^{obs} - y_i^{est}, i = 1,2,\cdots,m \ . \tag{1}$$

Then the objective function is a sum of squared residuals - the term 'least-squares' derives from the function:

$$F = \sum_{i=1}^{m} r_i^2 = \sum_{i=1}^{m}\left(y_i^{obs} - y_i^{est}\right)^2. \tag{2}$$

The objective function is the sum of the squares of the deviations between the observed values and the corresponding estimated values [6]. The maximum absolute discrepancy between observed and estimated values is minimized using optimization methods.

We used non-parametric Kaplan-Meier estimates $(KM(t_i))$, [10], [11] as the observed values (y_i^{obs}) of the objective function and the survivor rate estimates $(S(t_i))$ of inverse Gaussian [8], [20] model as the estimated value (y_i^{est}) of the objective function F [10]. We considered the objective function for the model of the form

$$F = \sum_{i=1}^{n} f_i\left(KM(t_i) - S(t_i)\right)^2 \tag{3}$$

where f_i is the number of failures at time t_i and m is the number of failure groups [20].

Of course it is not so easy to work with an account of the relatively complicated form of Inverse Gaussian Distribution models for its density and survivor functions.

We used the following procedure:

- Note that the Kaplan-Meier method [10] is independent of parameters, so for a particular value of time t_i, we find the value of the Kaplan-Meier estimate $(KM(t_i))$, of the survival function.

- The time T required to cover a distance d is a random variable with pdf for Inverse Gaussian,

$$f(t) = \frac{1}{\sqrt{2\pi\beta t^3}} d\exp\left(-\frac{(d-vt)^2}{2\beta t}\right), t > 0$$

and for ease of computation, we put $\kappa = \frac{d^2\rho}{\beta}$, $\rho = \frac{v}{d}$, then we have

$$f(t) = \sqrt{\frac{\kappa}{2\pi\rho t^3}} \exp\left[-\frac{\kappa}{2\rho t}\left\{(\rho t - 1)^2\right\}\right].$$

- We suppose that the survivor function of inverse Gaussian distribution model at time t_i is

$$S(t) = \Phi\left(-\sqrt{\frac{\kappa}{\rho t}}(\rho t - 1)\right) - e^{2\kappa}\Phi\left(-\sqrt{\frac{\kappa}{\rho t}}(\rho t + 1)\right),$$

where Φ is the standard normal distribution function and $S(t_i, \alpha_0, \beta_0)$ is the survivor function [9] at the starting point (α_0, β_0).

- From the numerical values of the Kaplan-Meier estimates $(KM(t_i))$, and the survivor function $S(t_i, \alpha_0, \beta_0)$ of the inverse Gaussian distribution model at time t_i, we can evaluate errors $|S(t_i, \alpha_0, \beta_0) - KM(t_i)|$.
- The function value with a suitable starting point (α_0, β_0) is given by

$$F(t_i, \alpha_0, \beta_0) = \max_i |S(t_i, \alpha_0, \beta_0) - KM(t_i)| \ .$$

- We find numerical value of the objective function, F at initial point (α_0, β_0) and this function value can be used in numerical optimization search (Simplex or Quasi-Newton) methods to find the optimal point (α^*, β^*), and then parameter estimates at this optimal point.

For practical applications of least square method, we considered some medical data sets [4]. The drug 6-mercaptopurine (6-MP) was compared to a placebo to maintain remission in acute leukemia patients. The following table 1 gives remission times for two groups of twenty-one patients each; one group was given the placebo and was given the other the drug 6-MP.

Table 1. Remission times for two groups of twenty-one patients

	Length of remission (in weeks) of leukemia patients
6-MP for 21 patients	6,6,6,6* ,7, 9*,10,10*, 11*, 13,16, 17*, 19*, 20*, 22,23,25*, 32*,32*,34*,35*
Placebo for 21 patients	1, 1, 2, 2, 3, 4, 4, 5, 5, 8, 8, 8, 8, 11, 11, 12, 12, 15, 17, 22, 23

* Censored observations

We considered only the data of twenty-one patients who were given 6-MP drug and there were 7 failures at times (weeks) 6,7,10, 13, 16, 22 and 23; and 12 of the 21 patients were censored. The data of twenty-one leukemia patients were used for assessing the appropriateness of Inverse Gaussian model to find the survivor rate estimates. The results of each Inverse Gaussian model using Least-Squares methods and applying simplex methods are presented by giving the values of the variance-covariance matrices, parameter estimates and other related information in tables 2, table 3 and table 4.

2. Inverse Gaussian Distribution model using Least-Squares Method and Applying Simplex Methods (Nelder and Mead and Hooke and Jeeves Search Methods)

For a practical application of the least-squares estimation method when assuming that, the partial derivatives of the objective function F were not available.

Nelder and Mead [12], [13], [14] and Hooke and Jeeves [2], [15] are simplex methods and are useful for optimizing the nonlinear programming problems. These numerical methods optimize the objective function without calculating the derivatives at all. These methods do not require first partial derivatives (gradients) so may converge very slow or even may diverge at all [15], [18]. The numerical results of Inverse Gaussian distribution model using Nelder and Mead and Hooke and Jeeves search methods have been presented in this paper. The results include function values, parameter estimates, survivor-rate estimates; Kaplan-Meier estimates [10], [11] and other information have been given in Table 2 and Table 3 and in Figure 1 and Figure 2.

Numerical and Graphical Results for Inverse Gaussian Probability Distribution Model using Least-Squares Methods and Applying Nelder and Mead (NM) and Hooke and Jeeves (HJ) Methods.

Table 2: Comparison of Survival Rate estimates for Inverse-Gaussian Distribution Model

Failure Time (Weeks)	Number of Failures	Nelder and Mead Method		Hooke and Jeeves Method	
		Inverse-Gaussian	Kaplan Meier	Inverse-Gaussian	Kaplan Meier
0	0	1	1	1	1
6	3	0.877412837850	0.857142857143	0.888879444169	0.857142857143
7	1	0.841704584033	0.806722689076	0.853157673967	0.806722689076
10	1	0.744553376981	0.752941176471	0.753023268335	0.752941176471
13	1	0.664633654815	0.690196078431	0.668558805251	0.690196078431
16	1	0.599291159714	0.627450980392	0.598642626794	0.627450980392
22	1	0.499796188983	0.537815126050	0.491380175350	0.537815126050
23	1	0.486228723453	0.448179271709	0.476720366932	0.448179271709

Table-3: Parameter Estimates and Optimal Function Value for Inverse-Gaussian Distribution Model

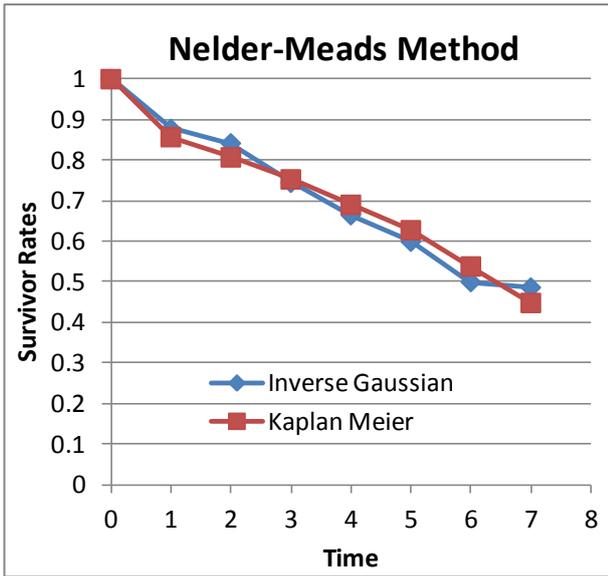|	Nelder and Mead Method	Hooke and Jeeves Method
Parameters Estimates	0.3131781774389110 0.018238423150123389	0.419982779999999999 0.2199999999999997
Optimal Functional value	0.038049451744791629	0.046434984891636665

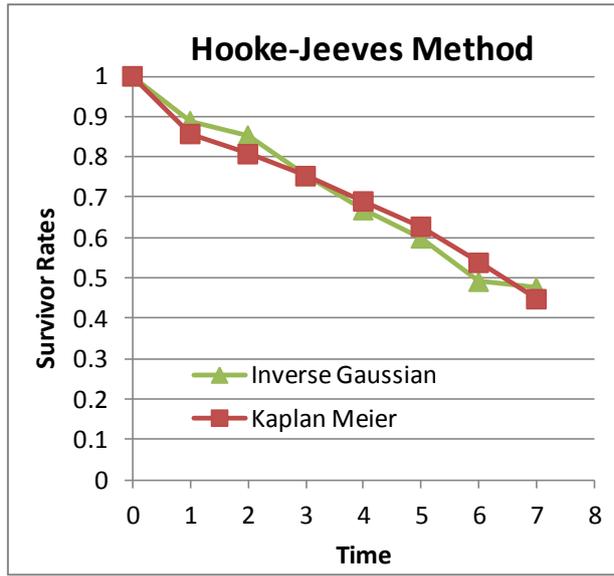**Figure 1. Survivor rate estimates using Nelder-Meads method**

Figure 2. Survivor rate estimates using Hooke-Jeeves method

3. Inverse Gaussian distribution model using Least-Squares Methods and Applying Quasi-Newton Optimization Methods (DFP and BFGS Methods)

We know that the survivor function, the probability of no failure before time t, for the two-parameter inverse Gaussian distribution model is

$$S(t) = \Phi\left(-\sqrt{\frac{\kappa}{\rho t}}(\rho t - 1)\right) - e^{2\kappa}\Phi\left(-\sqrt{\frac{\kappa}{\rho t}}(\rho t + 1)\right) \qquad (4)$$

where Φ is the usual standard normal distribution function (Chhikara et al (1974)).

Now the objective function for the inverse Gaussian distribution model using least-squares estimation method is

$$F = \sum_{i=1}^{n} f_i\big(S(t_i) - KM(t_i)\big)^2 \ , \qquad (5)$$

where f_i is the number of failures and $KM(t_i)$ is the Kaplan-Meier estimate at the failure time t_i.

We used the DFP and the BFGS optimization methods to find the parameter estimates for the objective function. Evaluating the first partial derivatives of the objective function F w.r.t. κ and ρ we obtain

$$\frac{\partial F}{\partial \kappa} = 2\sum_{i=1}^{m} f_i\left(S(t_i) - KM(t_i)\right)\frac{\partial S(t_i)}{\partial \kappa} \qquad (6)$$

and

$$\frac{\partial F}{\partial \rho} = 2\sum_{i=1}^{m} f_i\left(S(t_i) - KM(t_i)\right)\frac{\partial S(t_i)}{\partial \rho} \ , \qquad (7)$$

where

$$\frac{\partial S(t)}{\partial \kappa} = -\left[\left(\frac{v_3\exp(-v_1^2/2) - v_4\exp(2\kappa - v_2^2/2)}{2\sqrt{2\pi}\rho t}\right) + \exp(2\kappa)\Phi(-v_2)\right]$$

and

$$\frac{\partial S(t)}{\partial \rho} = -\left[\left(\frac{v_2 \exp(-v_1^2/2) - v_1 \exp(2\kappa - v_2^2/2)}{2\sqrt{2\pi}\rho}\right)\right].$$

here

$$V_1 = \sqrt{\frac{\kappa}{\rho t}}(\rho t - 1) \text{ and } V_2 = \sqrt{\frac{\kappa}{\rho t}}(\rho t + 1)$$

$$\frac{\partial V_1}{\partial \kappa} = \frac{V_3}{2\rho t} \quad , \quad \frac{\partial V_1}{\partial \rho} = \frac{V_2}{2\rho}$$

$$\frac{\partial V_2}{\partial \kappa} = \frac{V_4}{2\rho t} \quad , \quad \frac{\partial V_2}{\partial \rho} = \frac{V_1}{2\rho} ,$$

where

$$V_3 = \sqrt{\frac{\rho t}{\kappa}}(\rho t - 1) \text{ and } V_4 = \sqrt{\frac{\rho t}{\kappa}}(\rho t + 1).$$

Using the objective function eq. (5), and the first partial derivatives eq. (6) and eq. (7) in the DFP and the BFGS optimization method, we can find the estimated value of the parameters for which the least-squares function gives the minimum value for inverse Gaussian distribution model and the results are presented in the table 4.

Table 4. Numerical Results for the inverse Gaussian Probability Distribution Model using Least-Squares Method and Applying Quasi Newton Methods

Quasi Methods	Parameters Estimates	Optimal Functional value	Gradient at Optimal (κ^*, ρ^*)	The Variance-Covariance at Optimal (κ^*, ρ^*)
DFP Model	0.18341065 0.01220035	0.565086365E-02	-0.63533E-05 -0.46768E-04	6.240512 0.3346758 0.3346758 0.018077
BFGS Model	0.18346625 0.01220656	0.565086341E-02	-0.27184E-06 0.59512E-05	6.0990794 0.3257961 0.3257961 0.0175183

5. Conclusion

The Survival rate estimates for the 21 Leukemia patients for the period of 35 weeks under observations with drug 6-MP [5], [19] were compared using parametric Inverse Gaussian distribution model and non-parametric Kaplan Meier Model [12]. We found that the results like parameter estimates, optimal function value, variance covariance matrix, gradient vector at the optimal point (α^*, β^*) were approximately same for both the cases when the derivatives of an objective

function were available (using Quasi-Newton method (DFP and BFGS methods)) and when first partial derivatives of the objective function were not available (using the Hooke and Jeeves, and Nelder and Mead method). We also noted that the parameter estimates for the Nelder and Mead method are very close to the Quasi-Newton methods (DFP or BFGS). Using unconstrained optimization method (DFP-method and BFGS), we presented parameter estimates, inverse – Hessian matrices, survivor-rate estimates and the maximum likelihood function values for the inverse Gaussian distribution model using survival time cancer data sets. With the help of DFP-method and BFGS – method variance–covariance matrix can be found automatically which is very difficult if we apply some other method (e.g. Newton-Raphson method) for the inverse Gaussian distribution model.

Acknowledgement

References

[1]   J.A. Nelder, R. Mead, "A Simplex method for function minimization", Computer Journal 7, (1965), 308-313.

[2]   Robert Hooke and T.A. Jeeves, Westinghouse Research Laboratories, Pittsburgh, Pennsylvania, "Direct Research Solution of Numerical and statistical Problem research", International Journal of Statistics, Vol.16, pp.670-683,1960.

[3]   Fletcher,   Roger   (1987), Practical   methods   of   optimization (2nd   ed.),   New   York: John   Wiley   & Sons, ISBN 978-0-471-91547-8

[4]   Freireich et al. (1963) "The Effect of 6-mercaptopurine on the duration of steroid induced remission in acute leukemia", Blood 21, pp. 699-716.

[5]   Plackett, R.L. (1972). "The discovery of the method of least squares. Biometrika", 59,239–251.

[6]   J. T. Betts, "Solving the Nonlinear Least Square Problems: Application of a General Method", Journal of Optimization Theory and Applications, vol. 8, no. 4, 1976.

[7]   J.E. Dennis Jr., R.B. Schnabel. "Nonlinear least squares Numerical Methods for Unconstrained Optimization and Nonlinear Equations", SIAM, Philadelphia, PA (1996), pp. 218–238.

[8]   J. L. Folks and R. S. Chhikara", The Inverse Gaussian Distribution and Its Statistical Application--A Review" Journal of the Royal Statistical Society. Series B (Methodological), Vol. 40, No. 3, pp. 263-289 (1978).

[9]   Ledvij, M. "Curve Fitting Made Easy." Industrial Physicist 9, 24-27, May 2003.

[10] Kaplan E.L., Meier P. (1958). Nonparametric estimation from incomplete observations, J. Amer. Statist. Assoc., 53, pp. 457-481.

[11] Cox D.R.,Oaks D. (1984) Analysis of Survival Data. London Chapman and Hall.

[12] C.J. Price, I.D. Coope, D. Byatt, "A convergent variant of the Nelder-Mead algorithm", Journal of Optimization Theory and Applications, 113, (2002), 5–19.

[13] L. Han, M. Newmann, "Effect of dimensionality on the Nelder-Mead simplex method", Optimization Methods and Software, 21, 1 (2006), 1-16.

[14] Nam P., Bogdan M.Wilamow "improved Nelder Mead's Simplex method and Applications", Journal Of Computing, Volume 3, Issue 3, ISSN 2151-9617, March 2011.

[15] G.S Kirat , A.N Surde "Review of Hooke and Jeeves Direct Search Solution Method Analysis Applicable To Mechanical Design Engineering" Novateur Publications International Journal of Innovations in Engineering Research and Technology [IJIERT], ISSN: 2394-3696, Volume 1, Issue 2, Dec-2014.

[16] Lisawadi, S.," Parameter Estimation for the Two-Sided BS and IG Distributions", VDM Verlag, Saarbrücken, 2009.

[17] Zhenlin Yang, "Maximum Likelihood Predictive Densities for The Inverse Gaussian Distribution with Applications to Reliability and Lifetime Predictions", Microelectronics Reliability, 39, 1413-1421, 1999.

[18] Seshardi, V. The Inverse Gaussian Distribution. Oxford Science Publications. ,1993.

[19] P. Basak, N. Balakrishnan, Estimation for the three-parameter inverse Gaussian distribution under progressive Type-II censoring, Journal of the Statistics Computation and Simulation 82 (7), 1055-1072, 2012.

[20] M. C. K.Tweedie, Statistical properties of inverse Gaussian distribution, Annals of Mathematical Statistics 28 (2), 362-377,1957.