Energy, Mines and Resources Canada
Canada Centre for Remote Sensing

Énergie, Mines et Ressources Canada
Centre Canadien de Télédétection

A Method for the Estimation of the Parameters of a Gamma Distribution

I. K. Crain
Applications Division

Technical Note 73-5
October 1973

A Method for the Estimation of the Parameters of a Gamma Distribution

Technical Note 73-5
(Formerly Ref 73-12)
October 1973

I. K. Crain
Applications Division

Canada Centre for Remote Sensing
Department of Energy, Mines and Resources, Ottawa

ABSTRACT

The gamma distribution is a useful stochastic model for linear geographic features such as geological lineaments on air-photos or remotely sensed images.

A method for estimating the parameters of this distribution is given utilizing the well known and desirable Method of Maximum Likelihood. The simultaneous solution of the two non-linear likelihood equations is aided by the fact that one is separable so that the system can be reduced to a single non-linear equation. The solution requires the use of a series approximation for the logarithmic derivative of the gamma function, and a computer iterative procedure to determine the parameters.

In studies of the statistical nature of linear features on the Earth's surface, it has been shown that the gamma distribution is an accurate model for stream link-lengths (James and Krumbein, 1969; Shreve, 1969). Crain (1973) has shown that this same distribution can be applied in the analysis of geotectonic lineaments and should be useful in the study of the length of linear features on imagery of planetary surfaces. The distribution has also been applied in seismology (Shlien and Toksoz, 1970) and in urban geography (Dacey, 1968).

The desirability of maximum likelihood estimates of parameters of distributions is well known (Fisher, 1924), but unfortunately, in this case the method leads to non-linear equations, one of which involves a function which must be approximated by a series expansion. For this reason, the use of computer computation is almost unavoidable. This note gives an outline of the method as a practical guide to those who wish to make sample estimates of these parameters.

Following the notation of Freund (1962), the gamma density function is defined as

$$g(x;\alpha,\beta) = \frac{1}{\Gamma(\alpha)\beta^{\alpha}} x^{\alpha-1} e^{-x/\beta} \qquad (1)$$

where x is the random variable and α and β are the two parameters

The moments of the density about the origin are given by

$$\mu'_k = \frac{\Gamma(\alpha+k)\beta^k}{\Gamma(\alpha)} \qquad (2)$$

so that $\mu'_1 = \alpha\beta$ and $\mu'_2 = \beta^2(\alpha+1)\alpha$

The mean and variance are thus $\mu = \alpha\beta$ and $\sigma^2 = \alpha\beta^2$. The mode of the density is at $x = (\alpha-1)\beta$. Standardization is customarily done by converting to a mean of 1, (as did Shreve, 1969) by dividing by μ.

By replacing the moments μ'_k Eq. (2) by the sample moments m'_k the "method of moments" estimators for α and β may be obtained.

$$\hat{\alpha} = \frac{(m'_1)^2}{m'_2 - (m'_1)^2} \qquad \hat{\beta} = \frac{m'_2 - (m'_1)^2}{m'_1} \qquad (3)$$

These estimators are known to be inefficient (Kendall and Stuart, 1961, vol. 2, p. 67).

The maximum likelihood estimators seek to maximize

$$L(\alpha,\beta) = \prod_{i=1}^{N} g(x_i;\alpha,\beta) \qquad (4)$$

or, more conveniently,

$$L'(\alpha,\beta) = \sum_{i=1}^{N} \ln g(x_i;\alpha,\beta) \qquad (5)$$

$$= -N \ln\Gamma(\alpha) - N\alpha\ln\beta + (\alpha-1)\sum_{i=1}^{N} \ln x_i - \frac{1}{\beta}\sum_{i=1}^{N} x_i$$

Differentiating Eq. (5) with respect to α and β and equating the results to zero gives the likelihood equations

$$-N \ln\beta - N \frac{d}{d\alpha} \ln\Gamma(\alpha) + \sum_{i=1}^{N} \ln x_i = 0 \quad \text{a}$$

$$\frac{-N\alpha}{\beta} + \frac{\sum_{i=1}^{N} x_i}{\beta^2} = 0 \quad \text{b} \qquad (6)$$

(Kendall and Stuart, 1961, vol.2, p.66)

Substituting the simple equation 6b into 6a gives the single non-linear equation in α.

$$\ln\alpha - \frac{d}{d\alpha} \ln\Gamma(\alpha) - \ln \bar{x} + \overline{\ln x} = 0 \qquad (7)$$

where $\bar{x} = \sum_{i=1}^{N} x_i/N$

and $\overline{\ln x} = \sum_{i=1}^{N} \ln x_i/N$

One of the difficulties in solving Eq. (7) is the calculation of the second term, the logarithmic derivative of the gamma function. A converging series expansion is given by Morse and Feshback, (1953, p.422) as follows:

$$f(\alpha) = \frac{d}{d\alpha} \ln\Gamma(\alpha)$$

$$= -\gamma - \frac{1}{\alpha} + \sum_{i=1}^{\infty} \left(\frac{1}{i} - \frac{1}{i+\alpha}\right) \qquad (8)$$

where $\gamma = 0.5772157$ (the Euler-Mascheroni constant)

A recursion relation allows the range of α in Eq. (8) to be kept between 0 and 1, for most rapid covergence of the series.

$$f(\alpha) = f(\alpha+1) - \frac{1}{\alpha} \qquad (9)$$

Since Eq. (8) represents a monotonically decreasing alternating series, the absolute accuracy of the approximation is less than the last term retained, (Wylie, 1953, p.417) and so is at least as good as 1/N where N is the number of terms calculated. In practice, the computer calculation of Eq. (8) is extremely rapid so that an N of the order of 3000 creates no hardship.

Having a method to calculate $f(\alpha)$, the next problem is, obviously, to solve for α the non-linear equation (7). There are many techniques available for the solution of non-linear equations (see for example Ralston, 1965, pp.318-347). The initial guess required by most iterative methods is provided by the method of moments estimators of Eq. (3)

The simplest solution to Eq. (7) is by a "trial and error" iteration. The maximum likelihood estimators seldom differ from the method of moments estimators by more than 30%. This region is then searched incrementally until the left-hand side of Eq. (7) changes sign, thus confining the solution to the region of one increment. This region is then subdivided, and the result further refined. This procedure may be repeated until any desired accuracy is obtained. With an α in the vicinity of 2 and using 10 subdivisions of each region, only four repetitions are required to obtain three figure accuracy. If considerably more accuracy is desired, this method becomes inefficient, and more sophisticated iterative techniques should be employed.

REFERENCES

1. CRAIN, I.K., (1973). A Statistical Approach to the Analysis of Geotectonic Elements, Ph.D. Thesis, Australian National University, Canberra.

2. DACEY, M.F., (1968). A Family of Density Functions for Losch's Measurements on Town Distribution, in: Spatial Analysis, ed. B.J.L. Berry and D.F. Marble, Englewood Cliffs, N.J., Prentice-Hall, 168-171.

3. FISHER, R.A., (1924). The Conditions Under Which X^2 Measures the Discrepancy Between Observation and Hypothesis, J. Roy.Soc., vol. 87, p. 442.

4. FREUND, J.E., (1962). Mathematical Statistics, Englewood Cliffs, N.J., Prentice-Hall Inc.

5. JAMES, W.R., and KRUMBEIN, W.C., (1969). Frequency Distributions of Stream Link-lengths: Jour. Geology, vol. 77. pp. 544-565.

6. KENDALL, M.G., and STUART, A., (1961). The Advanced Theory of Statistics, vol. 2, London, Charles Griffin and Co.

7. MORSE, P.M., and FESHBACK, H., (1953). Methods of Theoretical Physics: New York, McGraw-Hill.

8. RALSTON, A., (1965). A First Course in Numerical Analysis, New York, McGraw-Hill.

9. SHLIEN, S., and TOKSOZ, M.N., (1970) (1970). A Clustering Model for Earthquake Occurences, Bull. Seis. Soc. Am., vol. 60, pp. 1765-1787.

10. SHREVE, R.L., (1969). Stream Lengths and Basin Areas in Topologically Random Channel Networks, Jour. Geology, vol. 77, pp. 397-414.

11. WYLIE, C.R., (1953). Calculus, New York, McGraw-Hill.