The estimators of normalized moments of the gamma distribution

1994 J. Phys. A: Math. Gen. 27 7883

(http://iopscience.iop.org/0305-4470/27/23/031)

J. Phys. A: Math. Gen. 27 (1994) 7883–7892. Printed in the UK

The estimators of normalized moments of the gamma distribution

R J A Tough

Defence Research Agency, Malvern, Worcs WR14 3PS, UK

Received 18 July 1994

Abstract. Exact results for the bias and variance of estimators of the normalized moments of gamma-distributed random variables, based on N independent samples, are derived. Some of the implications of these results for the analysis of K and gamma-distributed data are discussed; an illustrative example of such an analysis is included. These results extend and, in the large-N limit, reduce to the earlier perturbation calculations of Jakeman and Oliver.

1. Introduction

Since its introduction by Jakeman and Pusey [1] and, independently, by Ward [2] the K distribution has provided a robust and versatile model for the non-Gaussian statistical properties of radiation scattered from and propagated through a wide variety of random media [3]. The representation of a K-distributed random variable as a complex Gaussian noise process, decorrelating on a time scale τ_c, with a power x which is itself a gamma-distributed random variable and decorrelates over a much longer time τ_p, has proved to be particularly valuable, both as a source of physical insight and as an aid to calculation [4]. Many of the applications of the K distribution to empirical data take advantage of this, its so-called compound form, to reduce the statistical analysis of the signal to that of its power averaged over a period τ, $\tau_c \ll \tau \ll \tau_p$. As τ_p can be of the order of many seconds, the number of independent samples of x obtained experimentally may not be very large and due attention must be paid to the spread and bias expected in measurements that result from this restricted number of available samples. More specifically, it is customary to analyse putatively gamma-distributed data sets by calculating their estimated normalized moments (based on N samples) and comparing them with those derived from the probability density function (PDF)

$$P(x) = \frac{b^{\nu}}{\Gamma(\nu)} x^{\nu-1} \exp(-bx) \tag{1}$$

of the values x taken by the gamma-distributed variable x. As these estimated normalized moments are based on a finite number of samples, each displays a bias and a spread characterized by a variance, both of which depend on N. Previous workers [5, 6] have derived expressions for these biases and variances based on an expansion in powers of N^{-1}, typically neglecting all terms $O(N^{-2})$ and smaller. In this paper we derive exact, closed-form expressions for the mean, variance and expectation value of the mth power of the estimator of the nth normalized moment based on N independent samples $\{x_i\}$ of the gamma-distributed random variable x. These results form the basis of a discussion of the behaviour of the normalized moment plots, frequently used to display the results of the statistical analysis of sea clutter and other K-distributed signals [4], as a function of the number of samples. Finally we present an analysis of some typical microwave scattering data; the results we have derived allow us to accommodate their extremely non-Gaussian statistical behaviour within the familiar context of the gamma/K distribution model. This is done simply by taking account of the finite number of samples available in the measurements.

2. Evaluation of the bias and variance of the estimator of the nth normalized moment

Given the set $\{x_i\}$ of N independent samples of a random variable, an estimator of the nth moment of its distribution is given by

$$\hat{x}^n = \frac{1}{N}\sum_i x_i^n. \tag{2}$$

As is well known [7], in the case of the gamma distribution (1), this provides an unbiased estimate of

$$\begin{aligned}\langle x^n\rangle &= \frac{b^\nu}{\Gamma(\nu)}\int_0^\infty \mathrm{d}x\, x^{\nu+n-1}\exp(-bx)\\ &= \frac{\Gamma(\nu+n)}{\Gamma(\nu)b^n}\end{aligned} \tag{3}$$

i.e. $\langle \hat{x}^n\rangle = \langle x^n\rangle$. However, the corresponding estimator for the normalized moment is biased so that

$$\left\langle \frac{\hat{x}^n}{(\hat{x})^n}\right\rangle \neq \frac{\langle x^n\rangle}{\langle x\rangle^n} = \frac{\Gamma(\nu+n)}{\Gamma(\nu)\nu^n}.$$

The variance of the estimator

$$\left\langle \frac{(\hat{x}^n)^2}{(\hat{x})^{2n}}\right\rangle - \left\langle \frac{\hat{x}^n}{(\hat{x})}\right\rangle^2 \tag{4}$$

provides a quantitative measure of the expected spread in its observed value. We will now show how the bias and variance of the estimator of the nth moment can be expressed in a relatively simple closed form and, for completeness, evaluate the expectation value of its mth power:

$$I_{n,m} = \left\langle \frac{(\hat{x}^n)^m}{(\hat{x})^{mn}}\right\rangle. \tag{5}$$

In the conventional analysis of this problem [5, 6] the reciprocal of the estimator $\hat{x}$ is expanded about the value $\langle x\rangle^{-1}$ in powers of N^{-1}. We can avoid many of the difficulties inherent in this approach by introducing an integral representation of inverse powers of $\hat{x}$:

$$\frac{N^n}{(\Sigma_i x_i)^n} = \frac{N^n}{(n-1)!}\int_0^\infty \mathrm{d}s\, s^{n-1} \exp\left(-s\sum_i x_i\right). \tag{6}$$

In this way we are able to obtain results in closed form (and not as a series in N^{-1}) and reduce the combinatorial problems involved to a level where the evaluation of $I_{n,m}$ is tractable. As the N samples $\{x_i\}$ are independent we may write their joint PDF as follows:

$$P(\{x_i\}) = \frac{b^{\nu N}}{[\Gamma(\nu)]^N}\prod_i (x_i^{\nu-1}\exp(-bx_i)). \tag{7}$$

Thus, if we wish to evaluate the expectation value of the estimator of the nth normalized moment we have

$$\begin{aligned} I_{n,1} &= \frac{N^{(n-1)}}{(n-1)!}\int \mathrm{d}x_1 \ldots \int \mathrm{d}x_N\, P(\{x_i\})\sum_j x_j^n \int_0^\infty \mathrm{d}s\, s^{n-1}\prod_i \exp(-sx_i) \\ &= \frac{N^n b^{\nu N}}{[\Gamma(\nu)]^N (n-1)!}\int_0^\infty \mathrm{d}s\, s^{n-1}\int_0^\infty \mathrm{d}x\, x^{\nu+n-1}\exp(-(b+s)x) \\ &\quad \times\left[\int_0^\infty \mathrm{d}x\, x^{\nu-1}\exp(-(b+s)x)\right]^{N-1} \\ &= \frac{N^n b^{\nu N}}{(n-1)!}\,\frac{\Gamma(\nu+n)}{\Gamma(\nu)}\int_0^\infty \mathrm{d}s\,\frac{s^{n-1}}{(s+b)^{\nu N+n}}. \end{aligned}$$

The final integral over s can be evaluated in terms of gamma functions using the standard result

$$\int_0^\infty \mathrm{d}z\,\frac{z^{p-1}}{(1+z)^{p+q}} = \frac{\Gamma(p)\Gamma(q)}{\Gamma(p+q)}$$

to give

$$\begin{aligned} I_{n,1} &= \frac{\nu(\nu+1)\ldots(\nu+n-1)}{\nu^n\left(1+\frac{1}{N\nu}\right)\ldots\left(1+\frac{n-1}{N\nu}\right)} \\ &= \frac{\langle x^n\rangle}{\langle x\rangle^n}\prod_{j=1}^{n-1}\left(1+\frac{j}{N\nu}\right)^{-1}. \end{aligned} \tag{9}$$

If this result is expanded to lowest order in N^{-1} we find that

$$\begin{aligned} I_{n,1} &= \frac{\langle x^n\rangle}{\langle x\rangle^n}\left(1-\frac{1}{N\nu}\sum_{j=1}^{n-1} j\right) \\ &= \frac{\langle x^n\rangle}{\langle x\rangle^n}\left(1-\frac{n(n-1)}{2N\nu}\right) + \mathrm{O}(N^{-2}). \end{aligned} \tag{10}$$

This agrees with the result obtained from the series expansion analyses of Jakeman [5] and Oliver [6]. We now evaluate $I_{n,2}$ by noting that

$$(\hat{x}^n)^2 = \frac{1}{N^2} \sum_{i,j} x_i^n x_j^n$$

$$\frac{1}{(\hat{x})^{2n}} = \frac{N^{2n}}{(2n-1)!} \int_0^\infty \mathrm{d}s\, s^{2n-1} \exp\left(-s \sum_i x_i\right)$$

so that we obtain, much as before,

$$I_{n,2} = \frac{b^{\nu N} N^{2n-2}}{(2n-1)!} \left\{ N \frac{\Gamma(\nu+2n)}{\Gamma(\nu)} + N(N-1) \left[\frac{\Gamma(\nu+n)}{\Gamma(\nu)} \right]^2 \right\} \int_0^\infty \mathrm{d}s \frac{s^{2n-1}}{(b+s)^{\nu N + 2n}}$$

$$= \left\{ \left(1 - \frac{1}{N}\right) \frac{\langle x^n \rangle^2}{\langle x \rangle^{2n}} + \frac{1}{N} \frac{\langle x^{2n} \rangle}{\langle x \rangle^{2n}} \right\} \prod_{j=1}^{2n-1} \left(1 + \frac{j}{N\nu}\right)^{-1}. \tag{11}$$

Thus the variance of the estimator of the nth normalized moment is given by

$$I_{n,2} - (I_{n,1})^2. \tag{12}$$

This can be expanded to lowest order in N^{-1} as

$$\frac{1}{N}\left(\frac{\langle x^{2n}\rangle - \langle x^n \rangle^2}{\langle x \rangle^{2n}}\right) + \frac{\langle x^n \rangle^2}{\langle x \rangle^{2n} N\nu} \left[2 \sum_{j=1}^{n-1} j - \sum_{k=1}^{2n-1} k \right] = \frac{\langle x^n \rangle^2}{\langle x \rangle^{2n} N} \left\{ \frac{\langle x^{2n} \rangle}{\langle x^n \rangle^2} - 1 - \frac{n^2}{\nu} \right\}; \tag{13}$$

this agrees with the series expansion result given in [6].

The methods we have just described can be applied to the evaluation of $I_{n,m}$, the only difficulty encountered coming from the combinatorix of the term arising from $(\hat{x}^n)^m$.

Thus we have

$$I_{n,m} = \frac{N^{m(n-1)}}{(mn-1)!} \int \mathrm{d}x_1 \ldots \int \mathrm{d}x_N\, P(\{x_i\}) \left(\sum_k x_k^n\right)^m \int_0^\infty \mathrm{d}s\, s^{mn-1} \exp\left(-s \sum_i x_i\right)$$

$$= \frac{N^{m(n-1)} b^{N\nu}}{(mn-1)! [\Gamma(\nu)]^N} \int_0^\infty \mathrm{d}\, s^{mn-1} \sum_{r=1}^{N} \frac{N!}{(N-r)!}$$

$$\times \left[\int_0^\infty \mathrm{d}x\, \mathrm{x}^{\nu-1} \exp(-(b+s)x) \right]^{N-r} \sum_{\{a_j\}} (m; \{a_j\}) \prod_j$$

$$\times \left[\int_0^\infty \mathrm{d}x\, x^{\nu+nj-1} \exp(-(b+s)x) \right]^{a_j}$$

where $\{a_j\}$ is any set of integers satisfying the conditions $\Sigma_j a_j = r$ and $\Sigma_j j a_j = m$ and $(m; \{a_j\})$ is the number of ways of partitioning m different objects into a_k subsets, each containing k objects, for $k = 1, 2, \ldots, m$. Noting that $0! = 1$, $(m; \{a_j\})$ may be written as

$$(m; \{a_j\}) = \frac{m!}{\prod_{k=1}^{m} (k!)^{a_k} a_k!}.$$

The integrals over x and s can now be expressed in terms of gamma functions as before to give as the final result:

$$I_{n,m}=\frac{N^{m(n-1)}\Gamma(N\nu)}{\Gamma(N\nu+mn)}\sum_{r=1}^{m}\frac{N!}{(N-r)!}\sum_{\{a_j\}}(m;\{a_j\})\prod_j\left[\frac{\Gamma(\nu+jn)}{\Gamma(\nu)}\right]^{a_j}$$
$$=\frac{1}{\langle x\rangle^{mn}}\prod_{k=1}^{mn-1}\left(1+\frac{k}{N\nu}\right)^{-1}\sum_{r=1}^{m}\frac{N!}{(N-r)!N^m}\sum_{\{a_j\}}(m;\{a_j\})\prod_j\langle x^{nj}\rangle^{a_j}.$$

A list of unrestricted partitions $\{a_j\}$ for $m\leqslant 10$ and the corresponding values of $(m;\{a_j\})$ is given in [8], and allows $I_{n,m}$ to be evaluated explicitly for any reasonable value of m. The following special case values of $I_{n,m}$ are of particular interest:

$$N=1 \qquad I_{n,m}=1$$

and

$$I_{n,m}\rightarrow N^{m(n-1)} \qquad \text{as } \nu\rightarrow 0. \tag{15}$$

The first of these provides us with a check on our analysis as, when $N=1$, $\hat{x}^n/(\hat{x})^n=1$ and $I_{n,m}$ should take the value unity. When $\nu\rightarrow 0$, the normalized moments tend to large but finite values which depend on the finite number of samples taken; this should be contrasted with the unbiased result $\langle x^n\rangle^m/\langle x\rangle^{nm}\rightarrow\infty$ as $\nu\rightarrow 0$.

3. Numerical results and discussion

As is discussed in [4], a common method of assessing the applicability of the gamma distribution as a model of sets of data is to plot the estimators of the third, fourth and, occasionally, higher-order normalized moments of the data sets as a function of their estimated normalized variances. For convenience, these results are best presented in a log–log format. The availability of only finite sets of data will induce a bias and spread (described by the results we have just derived) in these plots which will not coincide with the universal curves based on

$$\frac{\langle x^n\rangle}{\langle x\rangle^n}=\frac{\Gamma(n+\nu)}{\Gamma(\nu)\nu^n} \qquad \text{and} \qquad \frac{\langle x^2\rangle}{\langle x\rangle^2}=\frac{1}{\nu}+1. \tag{16}$$

These effects of finite sample size are evident in the plots of $I_{3,1}$ and $I_{4,1}$ against $I_{2,1}$ for $N=50$, 100, 500 and 1000 shown in figures 1 and 2. For a given value of N, $I_{n,m}$ cannot exceed the value given by (15), so that the curves terminate at end points (shown in figures 1 and 2 for the cases $N=50$, 100 and 500). This is in contrast to the $N\rightarrow\infty$ curves based on (16), which tend to infinity as $\nu\rightarrow 0$. These are shown as dashed lines.

In figure 3 we show a plot of

$$\frac{I_{3,2}-I_{3,1}^3}{I_{3,1}^2}$$

as a function of $I_{2,1}$ for several values of N, again in a log–log format. The corresponding plots for the estimator of the fourth normalized moment are shown in figure 4. We note that, in some circumstances, the standard deviation of the distribution of the estimators exceed the value of its mean which, as the estimators are necessarily positive, is symptomatic of the distribution having a very long 'tail'. Tapster *et al* [9] discuss the

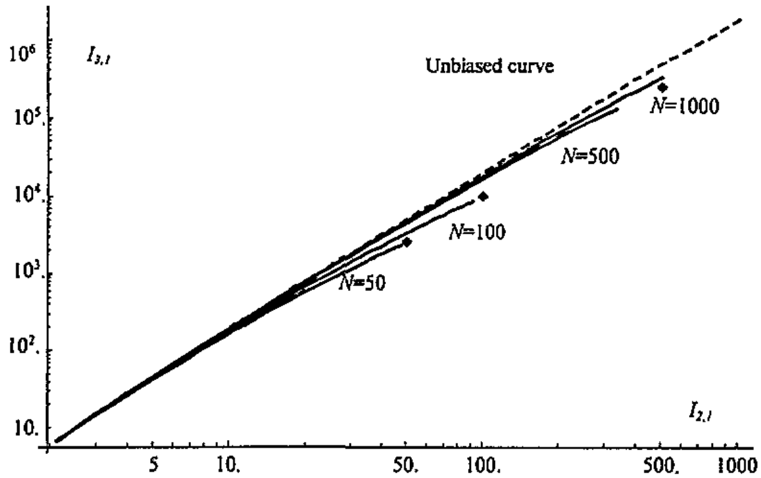

Figure 1. Plots of $I_{3,1}$ versus $I_{2,1}$ for $N=50$, 100, 500, 1000 and the unbiased ($N\to\infty$) limit. The end points of the $N=50$, $N=100$ and $N=500$ curves are also shown.

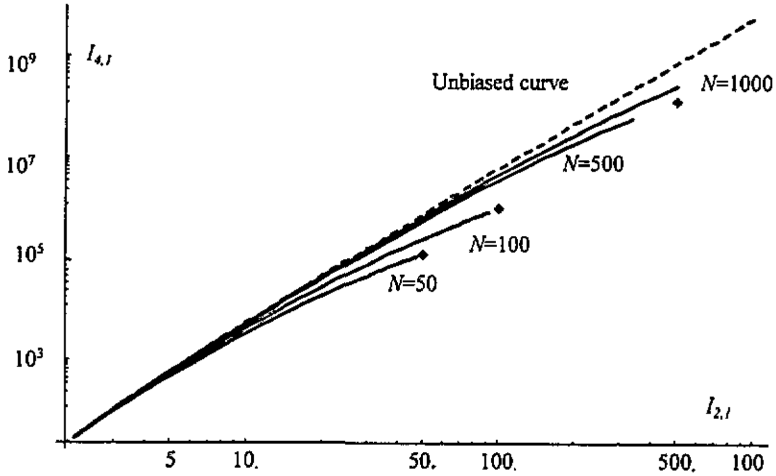

Figure 2. Plots of $I_{4,1}$ versus $I_{2,1}$ for $N=50$, 100, 500, 1000 and the unbiased ($N\to\infty$) limit. The end points of the $N=50$, $N=100$ and $N=500$ curves are also shown.

consequences of this feature for the experimental determination of moments from finite sets of data; the detailed results we present here allow an immediate extension of their analysis of un-normalized (and unbiased) moments to the estimation of the biased normalized moments. Finally we note that, in accordance with the behaviour of the estimators (see (15)) as $\nu\to 0$, the variance of the estimator of a normalized moment tends to zero in this limit, illustrating a dramatic effect of finite sample size.

In our analysis we have assumed that our N samples are independent. Effects of correlation among the samples are difficult to analyse in detail, though a parametrization in terms of an effective number N_{eff} of independent samples (where $N_{\text{eff}} < N$) provides a useful basis for qualitative discussion [5, 6]. The incorporation of such an N_{eff} into the results of the present analysis would be equally valid and entirely straightforward.

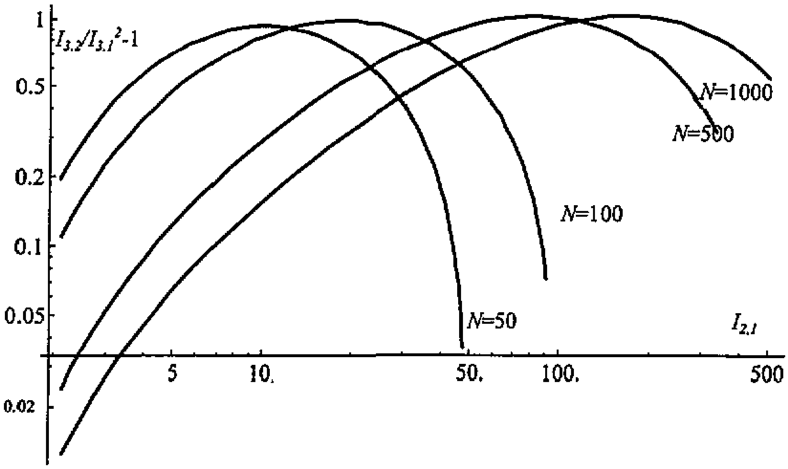

Figure 3. Plots of $I_{3,2}/(I_{3,1})^2-1$ versus $I_{2,1}$ for $N=50$, 100, 500 and 1000.

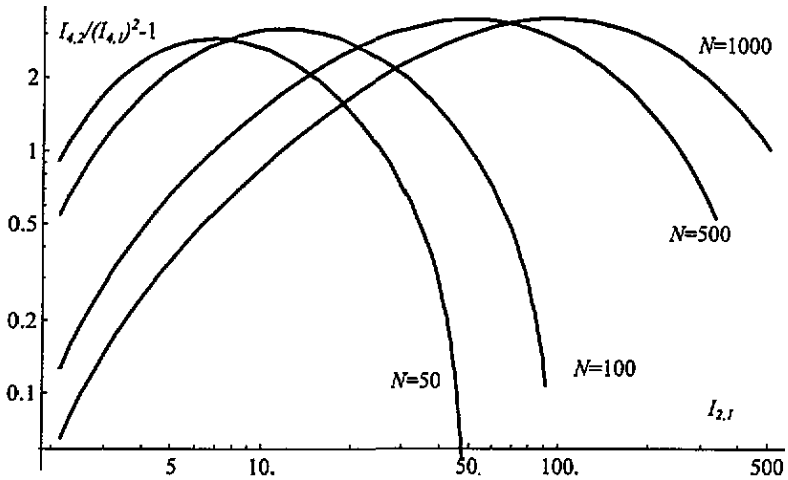

Figure 4. Plots of $I_{4,2}/(I_{4,1})^2-1$ versus $I_{2,1}$ for $N=50$, 100, 500 and 1000.

The results that we have presented thus far are essentially formal; the widespread use of the gamma and K distributions to model the statistical variation of scattering cross section and the corresponding fluctuations in the scattered field provides both a motivation for their derivation and a source of data with which they may be validated. To conclude our discussion we will present analyses of a set of data that illustrate these effects of finite sample number.

Jakeman [9] has presented an attractive phenomenology that may underpin the widespread applicability of these models. He considers the coherent illumination of a fluctuating population of scatterers which tend to bunch together so that the scattered light no longer has the Gaussian statistics characteristic of speckle. This bunching of the scatterers is manifest in the gamma distribution of their effective scattering cross section. 'Spiky' data, which sporadically attain very large values, are modelled by distributions with small values of ν; in Jakeman's model the spikes correspond to the presence of bunches of scatterers appearing occasionally in an otherwise unoccupied scattering volume. We see from the series expansion result (10) that finite sample effects are most pronounced when the product $N\nu$ is small and so should be evident in the analysis of strongly non-Gaussian spiky data. Such data are obtained when the illuminated area containing fluctuating scatterers is very small. In optical experiments this is achieved by focusing down a laser beam onto a suitable random medium; a corresponding situation arises in microwave scattering as a result of the range-gating of returns in a high-resolution radar system. The practicalities of these measurements need not concern us here; we should note, however, that it is in the areas of radar clutter and optical scattering and propagation that the gamma/K distribution model has proved to be most useful [3].

As our example of the application of our results we consider the statistics of the radar scattering cross sections, measured in individual range cells along the length of a vessel, subject to the aleatoric influence of the moving sea surface. In a high-resolution radar system the spatial extent of the vessel will occupy many range cells, each of which will contain relatively few scattering centres whose positions and orientations fluctuate. The cross sections associated with different range cells are found to have profoundly non-Gaussian statistical properties, for which the gamma/K distribution may provide a model. To test the applicability of this model, estimators of the third and fourth normalized moments of the measured cross section can be plotted as a function of that of its normalized variance. For each range cell a limited number (in our case $N \approx 200$) of sets of samples of the received back scatter are collected and N averaged effective cross sections are measured. These then give us estimators of the required normalized moments. In this way each range cell contributes a single point to each of the plots shown in figure 5. We see that the broken curves derived from the unbiased (infinite N) normalized moments do not describe the data in the strongly non-Gaussian regime. The unbroken curves derived from the biased results (9) and (11) are also shown. These describe the deviation of the behaviour of the estimators of the third and fourth normalized moments as a function of normalized variance from the unbiased curves very well. It is particularly noteworthy that, for relatively small values of the normalized variance, the estimators lie on both sides of the biased curve while, for moderate values of the normalized variance, the estimators tend to lie below the biased curve. This can be understood in terms of the variances of these estimators shown in figures 3 and 4. When the corresponding standard deviation is smaller than the mean, a scatter on both sides of the biased curve is expected. As is discussed by Tapster *et al* [10], the third and fourth normalized moments are necessarily positive and yet, for intermediate values of the normalized variance, have distributions whose means are significantly less than their standard deviations. This is indicative of these distributions having very long 'tails'. Consequently a given measurement of one of these estimators tends to lie below even the biased curve. In the extreme non-Gaussian limit we see from figures 3 and 4 (and indeed from (15)) that the variances of the distributions of these estimators decrease, so that the observed convergence of the experimental data to the biased curves as ν tends to zero is to be expected. It is perhaps worthy of note that the series expansion results of Jakeman and Oliver [5, 6] are quite inadequate to describe these results and predict negative values for the necessarily positive normalized moments in the extreme non-Gaussian regime. All in all, our results give an excellent description of the non-Gaussian fluctuations in effective cross section, merely by allowing for the effects of finite sample number, which, as $N\nu$ is of the order of unity or less, can be very significant.

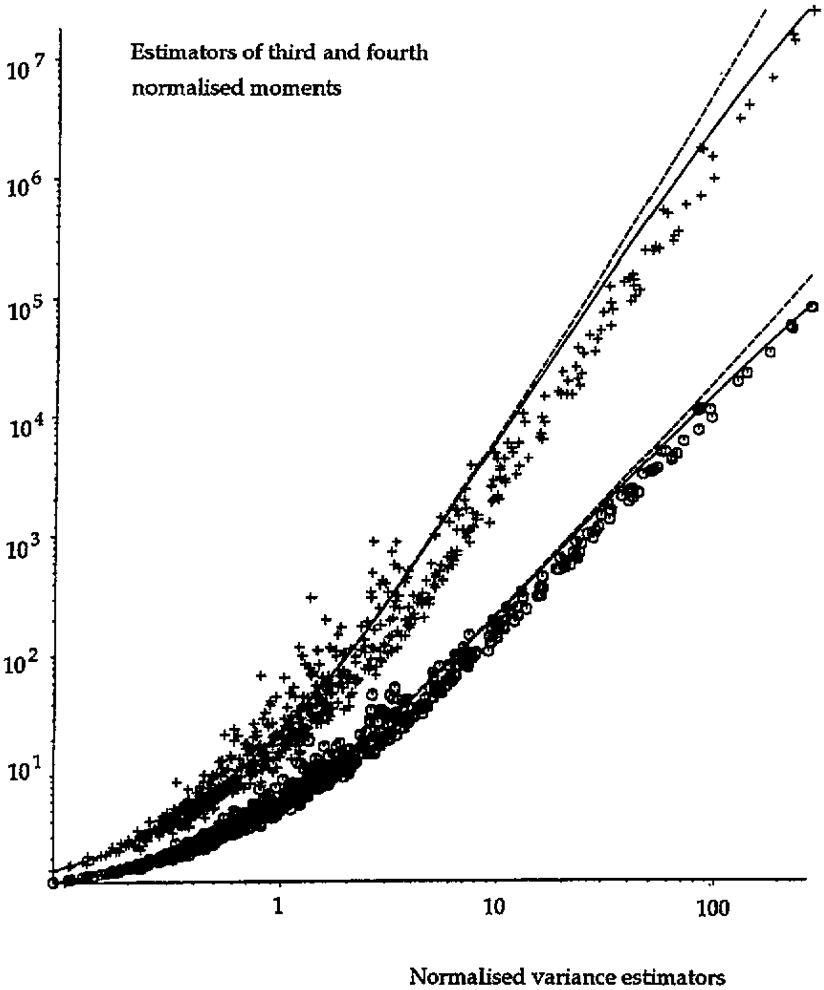

Figure 5. The third and fourth normalized moments of averaged radar cross section, presented as a function of the normalized variance. The circles and crosses denote measured third and fourth moments respectively; the unbiased curves are shown as broken lines while the biased curves, corrected for the effect of finite sample number, are shown as full lines.

Thus we have been able to accommodate these seemingly exceptional data within the familiar phenomenological framework of the K-distribution model.

To conclude, we re-emphasize that the results presented here have a much wider range of applicability than those given in [5, 6], which depend on $N\nu \gg 1$ to be valid. Consequently they will be of particular practical significance (in conjunction with the analysis of Tapster *et al* [10]) when the gamma distribution is used, with $\nu \approx 0$, to model very spiky data. We have presented an example that both demonstrates this and extends the application of the gamma/K distribution model. The gamma distribution is also used (occasionally as an approximation to less tractable distributions [11]) in other areas [12, 13], where our results may be of value in data analyses similar to those we have discussed here.

Acknowledgments

The author is grateful to Mr K J Hermiston for his provision of the radar data.

References

[1] Jakeman E and Pusey P N 1976 *IEEE Trans.* **AP-24** 806
[2] Ward K D 1981 *Electron. Lett.* **17** 561
[3] Jakeman E and Tough R J A 1988 *Adv. Phys.* **37** 471
[4] Ward K D, Baker C J and Watts S 1990 *IEE Proc. F* **137** 471
[5] Jakeman E 1974 *Photon Correlation and Light Beating Spectroscopy* ed H Z Cummins and E R Pike (London: Plenum) p 45
[6] Oliver C J 1986 *Inverse Problems* **2** 481
[7] Cramer H 1946 *Mathematical Methods of Statistics* (Princeton: Princeton University Press) ch 27
[8] Abramowitz M and Stegun I A 1965 *Handbook of Mathematical Functions* (New York: Dover) ch 24
[9] Jakeman E 1980 *J. Phys. A: Math. Gen.* **13** 31
[10] Tapster P R, Weeks A R, Pusey P N and Jakeman E 1989 *J. Opt. Soc. Am.* A **6** 782
[11] Meyer D P and Mayer H A 1973 *Radar Target Detection* (New York: Academic) sec 4.4
[12] Nakagami M 1960 *Statistical Methods in Radio Wave Propagation* ed W C Hoffman (London: Pergamon) p 3
[13] Hahn G J and Shapiro S S 1967 *Statistical Methods in Engineering* (New York: Wiley) sec 3.2