Downloaded from rspa.royalsocietypublishing.org on October 19, 2014

PROCEEDINGS OF THE ROYAL SOCIETY A | MATHEMATICAL, PHYSICAL & ENGINEERING SCIENCES

Exponentially Decreasing Distributions for the Logarithm of Particle Size

O. Barndorff-Nielsen

Proc. R. Soc. Lond. A 1977 **353**, doi: 10.1098/rspa.1977.0041, published 25 March 1977

References

Article cited in:
http://rspa.royalsocietypublishing.org/content/353/1674/401#related-urls

Email alerting service

Receive free email alerts when new articles cite this article - sign up in the box at the top right-hand corner of the article or click **here**

To subscribe to *Proc. R. Soc. Lond. A* go to:
http://rspa.royalsocietypublishing.org/subscriptions

Downloaded from rspa.royalsocietypublishing.org on October 19, 2014

Proc. R. Soc. Lond. A. **353**, 401–419 (1977)
Printed in Great Britain

Exponentially decreasing distributions for the logarithm of particle size

By O. Barndorff-Nielsen
Aarhus University, Denmark

(*Communicated by D. G. Kendall, F.R.S. – Received* 2 *July* 1976)

The family of continuous type distributions such that the logarithm of the probability (density) function is a hyperbola (or, in several dimensions, a hyperboloid) is introduced and investigated. It is, among other things, shown that a distribution of this kind is a mixture of normal distributions. As to applications, the paper focuses on the mass-size distribution of aeolian sand deposits, with particular reference to the findings of R. A. Bagnold. The distribution family seems, however, to be of some potential usefulness in other concrete contexts too.

1. Introduction

The cue for the work reported on here was the remarkable regularities in the distributions according to size of the particles of wind blown sands, discovered by R. A. Bagnold, F.R.S. The dynamical processes leading to those regularities are still far from being fully understood and modelled. The present paper is, however, not directly addressed to the intriguing and complex question of these processes, but to the much more modest task of providing a fairly simple family of probability distributions for graduating what Bagnold has termed 'regular sands'. Letting, throughout, $s\,(>0)$ denote size – in the case of the sand samples s stands for the diameter of a particle – and v the natural logarithm of s, the proposed family is that for which the logarithm of the probability (density) function for v is a hyperbola.

As will be further indicated in §§5 and 7, this distribution family appears to be applicable in various contexts other than that of aeolian sand deposits, in particular where the empirical distributions considered are size distributions, e.g. of mined diamonds or of personal incomes.

Often, when size phenomena are studied, the sizes of the single items are not observed directly, but the data consists of grouped 'moment distributions', i.e. letting $f(s)$ denote a theoretical probability function for the size s, what is observed is an empirical counterpart of some grouping of the distribution having probability function proportional to $s^l f(s)$ where l is a known non-negative number, typically 1, 2 or 3.

Thus, samples of particles of sand, as well as of many other kinds of material, are usually analysed with respect to particle sizes by dividing each sample into subsamples by means of a series of sieves of diminishing size (i.e. mesh width) and

The Royal Society is collaborating with JSTOR to digitize, preserve, and extend access to
Proceedings of the Royal Society of London. Series A, Mathematical and Physical Sciences.
www.jstor.org

Downloaded from rspa.royalsocietypublishing.org on October 19, 2014

then weighing, rather than counting the number of particles in, each subsample. For each sample a grouping of its, so-called, mass-size distribution is thereby observed. Viewing the particles as spheres of a homogeneous matter, the theoretical mass-size distribution is then the third-moment distribution ($l = 3$) derived from the theoretical size distribution.

Among the possibilities and suggestions which the availability of a simple mathematical–statistical description of the (mass-) size distribution of regular sands yields is that of microscopic or small sample studies of the texture of sands, as opposed to the more macroscopic studies based on the above-mentioned sieving-weighing procedure. (An investigation along this line is presently going on at the universities of Aarhus and Copenhagen.)

In dealing with (moment-) size distributions it should be kept in mind that the sizes of the single items sampled can, in many cases, not be considered as completely independent random variates.

The general problem of fitting a parametric distribution to an observed size or mass–size distribution is briefly discussed in §2. Section 3 contains a summary of the part of Bagnold's investigations of wind-blown sand which is relevant here, and the family of 'hyperbolic' distributions is introduced in §4. The log-normal distribution, which is a standard distribution in size statistics, is a limiting form of the 'log-hyperbolic' family and, moreover, this family can be generated as a mixture of log-normal distributions, cf. §5. An application to a series of sand samples from the Danish west coast (Kuhlman 1957) is presented in §6, and §7 consists of a number of complementary remarks, including a multivariate extension of the 'hyperbolic' family.

2. Maximum likeness estimation of the parameters of (mass-) size distributions

In graduating observed, grouped size or mass–size distributions by a member of a given parametric family of distributions, the estimates of the parameters have commonly been obtained by inspection of graphs of the empirical distribution, by fitting moments, and the like.

Theoretically, it seems in several respects preferable, in general, to select as an estimate of the parameter vector, θ say, that value $\hat{\theta}$ which maximizes the expression

$$\Sigma r_i \ln p_i(\theta), \tag{2.1}$$

where $p_i(\theta)$ denotes the theoretical probability mass in the ith group interval while r_i is the empirical counterpart of $p_i(\theta)$, i.e. for a size distribution, $r_i = n_i/n$ where n_i is the observed number of elements in the ith interval and $n = \Sigma n_i$, and for a mass–size distribution, $r_i = w_i/w$ with w_i being the weight for the ith interval and $w = \Sigma w_i$.

This estimation procedure I shall term 'maximum likeness estimation', and $\hat{\theta}$ will be called 'the maximum likeness estimate', with the following motivation.

Downloaded from rspa.royalsocietypublishing.org on October 19, 2014

In the case of a size distribution, if the n observations can be considered as independent then the procedure coincides with that of maximum likelihood. However, as stressed in the Introduction, independence is sometimes not a tenable assumption in size distribution investigations.

Maximizing (2.1) is the same as minimizing

$$I(r, p(\theta)) = \Sigma r_i \ln \frac{r_i}{p_i(\theta)}.$$

Generally, for two multinomial distributions π and p with cell probabilities π_i and p_i $(i = 1, 2, \ldots, k)$ the quantity

$$I(\pi, p) = \Sigma \pi_i \ln (\pi_i / p_i)$$

is known as 'the information of π with respect to p' (see Savage (1954), p. 50) or as 'the discrimination information between π and p' (cf. Kullback 1959). It may be thought of as a directed measure of the dissimilarity or distance between π and p. In particular, $I(\pi, p)$ is non-negative and it equals 0 if and only if p is identical to π. (The functional $I(.,.)$ is not symmetrical in its two arguments, and minimization of $I(p(\theta), r)$ would yield an estimate of θ different from $\hat{\theta}$, but corresponding to the 'minimum discrimination information estimates' proposed for analysis of contingency tables by Kullback *et al.*, see for instance Ku & Kullback (1974).)

Determination of the maximum likeness estimate $\hat{\theta}$ will nearly always have to be done numerically, by means of a calculator and using an iterative algorithm, e.g. that of Davidon–Fletcher–Powell or Newton–Raphson (see, for instance, Chambers (1973)).

3. Bagnold's investigations

'From the results of many analyses of sand samples it seems clear that if sand grading is a random phenomenon, from which conclusion it is difficult to escape, then some special probability function must be looked for'. (Bagnold 1954).

The mass–size distributions of aeolian sand deposits laid down under steady wind conditions exhibit certain, highly striking regularities, revealed and extensively studied by R. A. Bagnold. A brief summary of his investigations, as discussed in Bagnold (1954, chapters 9 and 10), will be given here.

Typically, a plot of the histogram of a distribution of the kind in question, with both scales logarithmic, will strongly suggest a theoretical curve which first increases almost linearly and then, after a smooth transition, decreases almost linearly. An example of this is provided by table 3.1 and figure 3.1, the data stemming from Bagnold (1954).

For such regular diagrams Bagnold introduced four characterizing quantities: the slopes f and $-g$ of the two linear asymptotes, which he termed, respectively ,the 'small-grade coefficient' and the 'coarse-grade coefficient', the abscissa m of the intersection point between the asymptotes, $\exp m$ being called the 'peak diameter'

Downloaded from rspa.royalsocietypublishing.org on October 19, 2014

of the sample, and finally, as a measure of spread of the distribution, the 'width' w defined as the ratio of the height to the length of the base line of any of the triangles formed by the two asymptotes and a horizontal line. One sees that

$$w^{-1} = f^{-1} + g^{-1},$$

and thus w is, except for a factor of 2, equal to the harmonic mean of f and g.

TABLE 1. THE MASS-SIZE DISTRIBUTION OF A SAND SAMPLE, DETERMINED BY SIEVING

(With calculations for the double-logarithmic histogram plot shown in figure 1. Data from Bagnold 1954.)

sieve diameter mm (s_i)	logarithmic interval midpoint $\frac{1}{2}(\ln s_i + \ln s_{i+1})$	mass (%) $100\,r_i$	$100 \times$ height of histogram, $100\,h_i$	$\ln h_i$
0.054		0.008		
0.073	−2.77	0.024	0.080	−7.14
0.099	−2.46	0.082	0.269	−5.92
0.144	−2.13	0.404	1.078	−4.52
0.191	−1.80	1.105	3.91	−3.24
0.261	−1.50	3.86	12.36	−2.09
0.318	−1.24	6.08	30.8	−1.18
0.414	−1.01	20.3	76.9	−0.26
0.592	−0.70	51.8	144.8	0.37
0.755	−0.40	14.12	58.1	−0.54
0.915	−0.18	1.855	9.65	−2.34
1.17	0.03	0.338	1.375	−4.29
1.58	0.31	0.043	0.143	−6.55
		0.005		

Observations from deserts and beaches and from wind-tunnel experiments (Bagnold 1954; Kuhlman 1957) indicate that the magnitudes of f and g are to a considerable extent dependent on the wind conditions at the time and place of deposition of the sand. Bagnold, however, noted a tendency for g to be near to 9.

Moreover, the two quantities f and g are differentially related to the two prevailing manners, 'saltation' and 'creep', in which the sand particles are transported by the wind. The finer sand grains are moved by 'saltation' during which process they are lifted or bounce into the air where they follow a trajectory of a characteristic shape and of a limited length. On hitting the sand surface such a grain may either rebound or strike another grain into saltation, or if it hits a coarser grain may start the latter rolling or 'creeping' along the surface.

Through his wind-tunnel experiments Bagnold discovered a further marked regularity which appears essential for a deeper understanding of the dynamics of aeolian sand transport. A sand placed at the entrance to the tunnel will be sorted and transported, by a steady wind, down through the tunnel in such a way that

Downloaded from rspa.royalsocietypublishing.org on October 19, 2014

at an arbitrary time after the start of the experiment there exists a log–linear relation between the peak diameter and the amount of sand, as functions of the distance from the entrance. For details, see Bagnold (1954, chapter 10).

Some further discussion of the dynamics of the wind transport of sand grains is contained in Bagnold (1956).

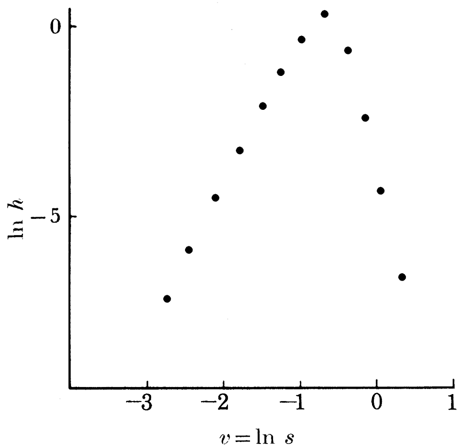

FIGURE 1. Double-logarithmic histogram plot of the mass-size distribution of a sand sample, given in table 1. (Only the top midpoints of the histogram slabs are shown.)

4. A family of distributions with both tails exponential

The simplest mathematical curve having a form like that indicated by figure 1 is the hyperbola. Except for a change of origin and scale, the general expression for the ordinate at v of such a hyperbola is

$$-\alpha\sqrt{(1+v^2)}+\beta v, \tag{4.1}$$

the asymptotes for v tending to $-\infty$ and $+\infty$ being, respectively, ϕv and $-\gamma v$, where

$$\phi = \alpha+\beta,$$
$$\gamma = \alpha-\beta.$$

The relevant domain of variation for (ϕ, γ) is $(0, \infty)^2$.

The probability function of v, corresponding to (4.1), is

$$a(\alpha,\beta)\exp\left(-\alpha\sqrt{(1+v^2)}+\beta v\right), \tag{4.2}$$

where $a(\alpha,\beta)$ denotes a norming constant. By changing to the parametrization given by (ϕ,γ), the probability function takes the form

$$a(\phi,\gamma)\exp\left(-\tfrac{1}{2}\{\phi(\sqrt{(1+v^2)}-v)+\gamma(\sqrt{(1+v^2)}+v)\}\right) \tag{4.3}$$

or

$$a(\phi,\gamma)\exp\left(-\tfrac{1}{2}\{\phi(\sqrt{(1+v^2)}+v)^{-1}+\gamma(\sqrt{(1+v^2)}+v)\}\right). \tag{4.4}$$

Downloaded from rspa.royalsocietypublishing.org on October 19, 2014

An expression for the norming constant in terms of a Bessel function is obtainable via the monotone transformation

$$x(v) = \surd(1+v^2)+v,$$

as follows

$$\int_{-\infty}^{+\infty} \exp\left(-\tfrac{1}{2}\{\phi x(v)^{-1}+\gamma x(v)\}\right) \mathrm{d}v = \tfrac{1}{2}\int_0^{\infty} (1+x^{-2}) \exp\left(-\tfrac{1}{2}\{\phi x^{-1}+\gamma x\right\} \mathrm{d}x$$
$$= \kappa K_1(\kappa)/\omega,$$

where

$$\kappa = \surd(\phi\gamma), \tag{4.5}$$

$$\omega = 1/(\phi^{-1}+\gamma^{-1}) \tag{4.6}$$

and $K_1(\cdot)$ is the modified Bessel function of the third kind and with index $\nu = 1$. Finally, introducing a location parameter μ and a scale parameter δ (>0) and writing ϕ for ϕ/δ and γ for γ/δ one obtains the probability function

$$f(v;\phi,\gamma,\mu,\delta) = \frac{\omega}{\delta\kappa K_1(\delta\kappa)} \exp\left(-\tfrac{1}{2}(\phi+\gamma)\surd(\delta^2+(v-\mu)^2)+\tfrac{1}{2}(\phi-\gamma)(v-\mu)\right). \tag{4.7}$$

Here ϕ and γ give the slopes of the two linear asymptotes for the graph of $\ln f(\cdot\,;\phi,\gamma,\mu,\delta)$, and κ and ω are determined by (4.5) and (4.6). The abscissa of the intersection point of the two asymptotes is μ, and the ordinate ν, say, equals the logarithm of the norming constant, i.e.

$$\nu = \ln \frac{\omega}{\delta\kappa K_1(\delta\kappa)}. \tag{4.8}$$

Thus $\exp(\mu)$ and ω are the theoretical counterparts to Bagnold's peak diameter $\exp(m)$ and width w, respectively. Note, however, that μ is not the mode of (4.7), except when $\phi = \gamma$. Denoting the mode by υ, one has

$$\upsilon = \mu+\delta(\phi-\gamma)/(2\kappa)$$

and

$$f(\upsilon;\phi,\gamma,\mu,\delta) = \frac{\omega}{\delta\kappa K_1(\delta\kappa)} \exp(-\delta\kappa). \tag{4.9}$$

The distribution with probability function (4.7) may be referred to as the 'hyperbolic' distribution. When v follows this distribution, the distribution of $s = \exp(v)$ is 'log-hyperbolic'.

In the symmetric case $\phi = \gamma\,(=\kappa)$ the distribution (4.7) tends to the Laplace distribution

$$\tfrac{1}{2}\kappa \exp(-\tfrac{1}{2}\kappa|v-\mu|)$$

for $\delta \to 0$, and to the Gauss distribution

$$1/\{\surd(2\pi)\,\sigma\}\exp(-\tfrac{1}{2}(v-\mu)^2/\sigma^2)$$

for $\delta \to \infty$ and $\delta/\kappa \to \sigma^2$.

With (4.7) as mass–size distribution the corresponding size distribution is of the same kind but having parameters $(\phi-3, \gamma+3, \mu, \delta)$.

Downloaded from rspa.royalsocietypublishing.org on October 19, 2014

Rough estimates of ϕ, γ and μ may be read off directly from a diagram such as figure 1. The same holds for the ordinate ν of the intersection point of the theoretical asymptotes, so that (cf. formula (4.8)) an estimate of δ is simple to obtain via a table of the function $xK_1(x)$. If, in addition to ν, one is prepared to estimate also the log–mode–size $\ln f(\upsilon;\phi,\gamma,\mu,\delta)$ from the diagram then it is even simpler to find an estimate of δ since, by (4.8) and (4.9),

$$\delta\kappa = \nu - \ln f(\upsilon;\phi,\gamma,\mu,\delta).$$

(With modern pocket calculators it is, of course, a trivial matter to plot (4.7) or its logarithm, given any set of parameter values.)

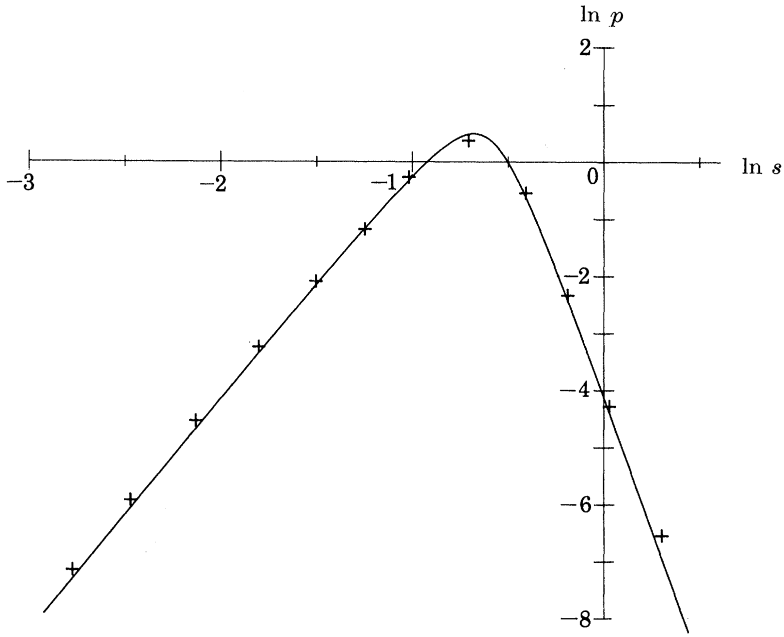

FIGURE 2. The mass-size distribution of a sand sample, together with the fitted 'log-hyperbolic' probability function, in double-logarithmic plotting. (Same data as in table 1 and figure 1.)

These rough estimates may be used as starting values for an iterative numerical algorithm leading to the maximum likeness estimates defined in §2.

Applying this procedure to the data of table 1 one obtains as maximum likeness estimates

$$\hat{\phi} = 4.11, \quad \hat{\gamma} = 9.54, \quad \hat{\mu} = -0.59, \quad \hat{\delta} = 0.20.$$

The log-probability function determined by these values is shown together with the data in figure 2. The letter p in the label of the ordinate axis in figure 2, and in figures 3 and 4 below, is short for probability function.

In considering a diagram, such as figure 2, which compares an observed, grouped

Downloaded from rspa.royalsocietypublishing.org on October 19, 2014

distribution with the log-probability function of a member of the family of distributions discussed here, it should be kept in mind that the fitted curve typically will lie below the observed points in both tails of the distribution and above the highest observed point(s), provided the model gives a good description. To be more specific, consider the ideal case where

$$r_i = p_i = \int_{v_0+(i-\frac{1}{2})l}^{v_0+(i+\frac{1}{2})l} f(\cdot\,;\phi,\gamma,\mu,\delta) \quad (i = 0, \pm 1, \pm 2, \ldots),$$

l being the constant interval length. The difference, d_i say, between $\ln h_i = (p_i/l)$ and the corresponding ordinate of the log-probability function satisfies

$$\exp\left[d_i = p_i/\{lf(v_0+il;\phi,\gamma,\mu,\delta)\}\right] \to \begin{cases} \sinh(\frac{1}{2}\phi l)/(\frac{1}{2}\phi l) & \text{for} \quad i \to -\infty, \\ \sinh(\frac{1}{2}\gamma l)/(\frac{1}{2}\gamma l) & \text{for} \quad i \to \infty. \end{cases}$$

Thus, in particular, the difference is asymptotically constant and positive. Actually, d_i will be positive as soon as the interval is sufficiently far from the mode for $f(\cdot\,;\phi,\gamma,\mu,\delta)$ to be a convex function of v, as may be seen by Jensen's inequality.

5. Relation to the log-normal distribution

The theoretical probability distribution most commonly used in connection with empirical particle size or mass-size distributions is the log-normal

$$\frac{1}{\sqrt{(2\pi)}\,\sigma}\exp\left(-\tfrac{1}{2}(v-\xi)^2/\sigma^2\right) \tag{5.1}$$

(v denoting the logarithm of the particle size). As has already been pointed out above, this distribution occurs as a limit case for the family (4.7). The log-normal distribution does not seldom give a satisfactory fit to the data (but it is quite inadequate for describing the mass-size distribution of well-sorted sands, as is apparent for instance from figure 2), and in some contexts it may, to a greater or lesser extent, be backed by simple probabilistic models of physical breakage processes, as shown by Kolmogoroff (1941), Halmos (1944), Epstein (1947) and Rényi (1960). Briefly described, the former three authors demonstrate that if a particle of initial size s_0 is repeatedly diminished by the breaking off of random proportions, so that after n steps in the breaking process the size is $s_n = s_0 \cdot x_1 \cdot \ldots \cdot x_n$, where $x_1, \ldots, x_n$ are random variates taking values between 0 and 1, then for large n the distribution of $\ln s_n$ will, under mild regularity conditions and by the central limit theorem, be approximately normal. Rényi's model is of a somewhat different type.

In view of this rôle of the log-normal distribution it seems of some interest that the distribution (4.7) is obtainable as a mixture of normal distributions. Specifically, if for given σ^2 a random variate v follows the normal distribution (5.1) with

$$\xi = \mu + \beta\sigma^2 \tag{5.2}$$

Downloaded from rspa.royalsocietypublishing.org on October 19, 2014

and if σ^2 is endowed with the distribution having probability function

$$\frac{1}{2}\frac{\kappa}{\delta K_1(\delta\kappa)}\exp\left(-\tfrac{1}{2}(\delta^2\sigma^{-2}+\kappa^2\sigma^2)\right) \tag{5.3}$$

then the resulting distribution for v is given by (4.7) (as in the foregoing, $\beta = \frac{1}{2}(\phi-\gamma)$ and $\kappa = \sqrt{(\phi\gamma)}$).

Table 2. Size distribution of diamonds from a large mining area in South West Africa

(Empirical data from Sichel (1973); 1 carat/stone = 205.6 mg.)

class intervals (carats/stone)	observed no. of stones	observed (normed) distribution	fitted 'log-hyperbolic' distribution
0–0.15	9	0.0088	0.0104
0.15–0.25	39	0.0382	0.0338
0.25–0.75	358	0.3503	0.3558
0.75–1.25	257.5	0.2520	0.2509
1.25–1.75	137	0.1341	0.1286
1.75–2.25	69.5	0.0680	0.0702
2.25–2.75	40.5	0.0396	0.0417
2.75–3.25	28	0.0274	0.0266
3.25–3.75	20.5	0.0201	0.0180
3.75–4.25	16.5	0.0161	0.0127
4.25–4.75	7.5	0.0073	0.0093
4.75–5.25	7	0.0068	0.0070
5.25–5.75	2.5	0.0024	0.0054
5.75–6.25	4.5	0.0044	0.0042
6.25–6.75	2	0.0020	0.0034
6.75–7.25	2	0.0020	0.0028
7.25–7.75	3	0.0029	0.0023
7.75–8.25	2	0.0020	0.0019
8.25–8.75	1	0.0010	0.0016
8.75–9.25	1.5	0.0015	0.0014
9.25–9.75	0.5	0.0005	0.0012
9.75–14.75	7	0.0068	0.0061
14.75–21.75	3	0.0029	0.0026
21.75–32.25	1	0.0010	0.0012
32.25–71.25	2	0.0020	0.0010
total	1022	1.0001	1.0001

The relation (5.2) between mean value and variance is such as occurs for a Brownian motion with drift, and herein may lie a clue to a derivation of the theoretical distribution of §4 by way of a stochastic process model.

Another concomitant remark is that the above mixture representation was hit upon by rather formal manipulation of the formula for the Laplace transform of the inverse Gaussian distribution. With u as the transform variable, this Laplace transform is

$$\exp\left(-\sqrt{(2\chi u+\chi\psi)}+\sqrt{(\chi\psi)}\right),$$

Downloaded from rspa.royalsocietypublishing.org on October 19, 2014

while the corresponding probability function has the form

$$\frac{\sqrt{\chi}}{\sqrt{(2\pi)}} \exp \sqrt{(\chi\psi)}\, x^{-\frac{3}{2}} \exp\left(-\tfrac{1}{2}(\chi x^{-1} + \psi x)\right) \quad (x > 0). \tag{5.4}$$

As is well known, this distribution has an interpretation as a first passage distribution for a Brownian motion with drift (cf. for instance, Cox & Miller (1965)).

(A few further remarks on the distributions (5.3) and (5.4) are given in §7 (v).)

On this background it was interesting to discover that Sichel (1973) has constructed another mixture of log-normal distributions, in fact specifying a linear relation

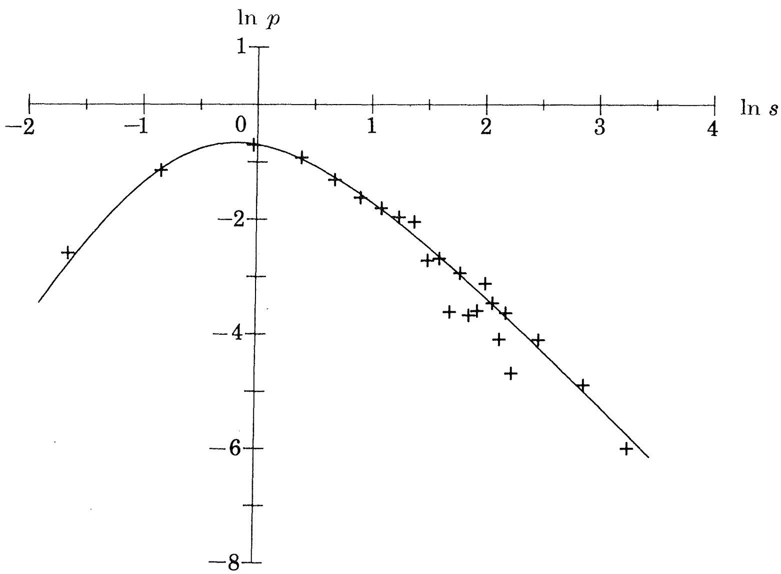

FIGURE 3. Double-logarithmic plot of observed distribution of diamond sizes and of the fitted 'log-hyperbolic' distribution. (The same distributions are described in table 2.)

between mean and variance as in (5.2) but taking for σ^2 the gamma distribution

$$\frac{\rho^{\tau}}{\Gamma(\tau)} (\sigma^2)^{\tau-1} \exp(-\rho\sigma^2).$$

This leads to another four-parameter distribution for v, having probability function

$$\frac{\sqrt{(2/\pi)}}{\Gamma(\tau)} \frac{\rho^{\tau}}{(2\rho+\beta^2)^{\tau-\frac{1}{2}}} |v-\mu|^{\tau-\frac{1}{2}} K_{\tau-\frac{1}{2}}\left(\sqrt{(2\rho+\beta^2)}\,|v-\mu|\right) \exp(\beta(v-\mu)), \tag{5.5}$$

where $K_{\tau-\frac{1}{2}}(\cdot)$ is the modified Bessel function of the third kind and index $\tau-\frac{1}{2}$. (The special case of (5.5) for which $\mu = \beta = 0$ was given in Teichroew (1957).)

The practical problems, with which Sichel (1973) dealt, concerned size distributions of diamonds. He remarked that 'diamond sizes in the marine deposit of South

Downloaded from rspa.royalsocietypublishing.org on October 19, 2014

West Africa are well represented by a two-parameter log-normal distribution provided the stones originate from a small compact mining block, on one and the same beach horizon'. For larger mining areas, however, marked deviations from log-normality are observed. The observed distribution given in table 2, and reproduced from Sichel (1973), demonstrates this. In order to handle data of this latter kind, Sichel introduced the mixture (5.5). He argued as follows for a linear specification as in (5.2): 'From the remarks with respect to the sorting action of the sea and the trapping by the bed rock formation on land, it is not unreasonable to assume that, *within* a particular beach deposit, parameters ξ and σ^2 are both functions of the distance d from the mouth of the Orange River. Furthermore, in view of sorting, ξ and σ^2 should be related inversely to d and should be related directly to each other. As a first assumption we may formalize the latter proposition as $\xi = a + b\sigma^2$, where a and b are two constants.'

The empirical distribution of diamond sizes, in table 2, has been fitted by (4.7) via maximum likeness estimation and the result is shown in that table and in figure 3. The estimated values of the parameters are

$$\hat{\phi} = 3.61, \quad \hat{\gamma} = 2.12, \quad \hat{\mu} = -0.54, \quad \hat{\delta} = 1.30.$$

Finally, it may be mentioned that for $\phi = \gamma$ and $\mu = 0$ the distribution (4.7) belongs, as follows from the above, to the class of scale mixtures of the symmetric normal distribution which was treated by Andrews & Mallows (1974), with a view to Monte Carlo evaluations.

6. Application to a sand sample series

An investigation of the mass-size distribution of particles of dune and beach sands at a locality – Skallingen – on the Danish west coast is reported by Kuhlman (1957). Samples were taken in July–August of 1954 and 1955, along a line of 200 m, orthogonal to the shore and running from the sea, across the flat beach and to a fairly large dune (rising to about 7 m above the flat beach level). The surface samples taken in 1955 and treated in tables 2 and 3 of Kuhlman (1957) were selected for the present application (excluding, however, the 4 samples from the top and the lee side of the large dune). This set consists of 16 samples, and the numbers 41, 43, 17, etc., for these, employed below, correspond to Kuhlman's notation. The samples are listed in order corresponding to their position on the above-mentioned transect, no. 41 being nearest to the sea, etc.

The distribution (4.7) was fitted to each of the 16 samples by maximum likeness estimation, using the computational procedure indicated earlier, and the results are shown in figure 4 and table 3. In every case the fit is excellent.

The strikingly small variation in $\hat{\mu}$ between samples corresponds to a variation in grain size diameter between 0.169 and 0.192 mm.

An impression of what the differences between the $\hat{\phi}$-values and between the $\hat{\gamma}$-values mean physically may be gained from the information (H. Kuhlman,

Downloaded from rspa.royalsocietypublishing.org on October 19, 2014

private communication) that samples 10 and 9 were taken on the same day and about one meter apart, within a fairly flat area that looked homogeneous, whereas the sites of samples 5 and 4, which were also taken near each other on the same day but on the slope of the large dune, looked very different to the naked eye, that of sample 4 appearing as a finer sand, as indeed it turned out to be, cf. table 3.

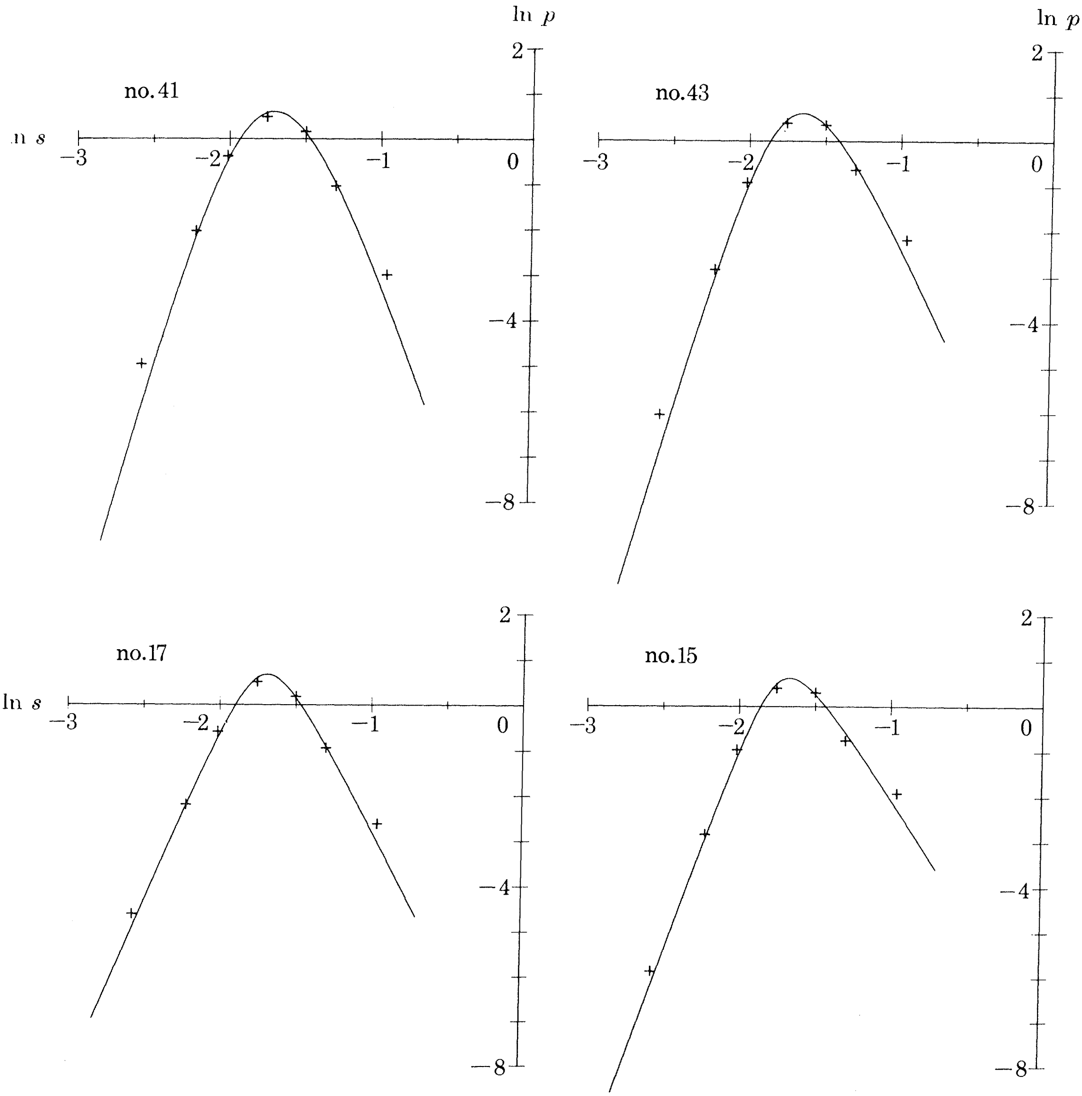

FIGURE 4*a* (For description see page 415)

Downloaded from rspa.royalsocietypublishing.org on October 19, 2014

7. Additional remarks

(i) Pareto's Law, celebrated in economics, states that distributions of persons according to income, of firms according to size, etc., have a strong tendency to decrease, in the upper tail, like a power of the variate studied (i.e. personal income,

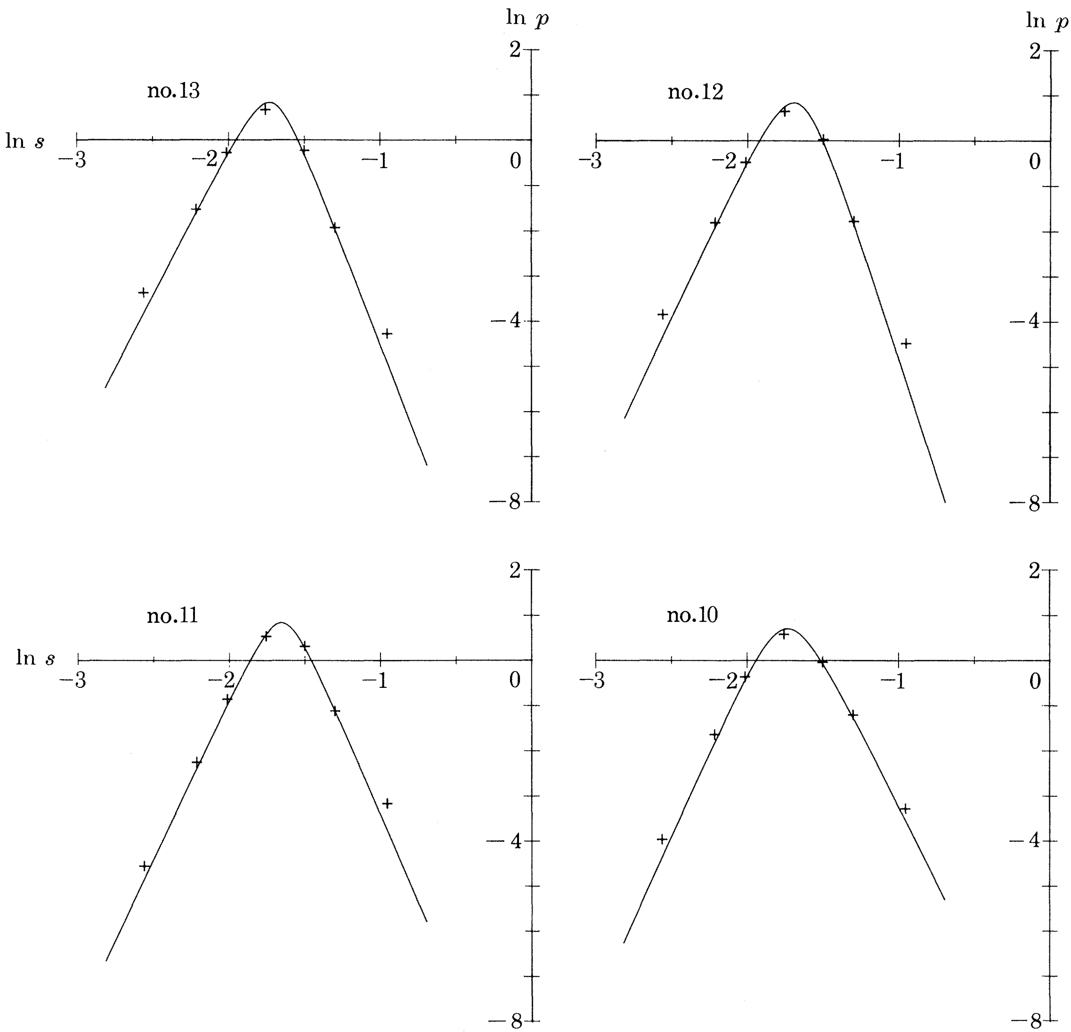

Figure 4*b* (For description see page 415)

firm size, etc.) (cf., for instance, Champernowne (1973); Steindl (1965).) In other words, the upper tail of the distribution of the log-variate falls off nearly exponentially. In fact, a similar tendency – perhaps somewhat less common – holds for the lower tail, and recent personal income distributions, for a number of nations, have

Downloaded from rspa.royalsocietypublishing.org on October 19, 2014

a shape more or less like that of (4.7) (see appendix to chapter 3 of Champernowne (1973)).

Reference is also made to Mandelbrot (1963), who discusses relations between Pareto's Law and the theory of, so-called, stable probability distributions, and who,

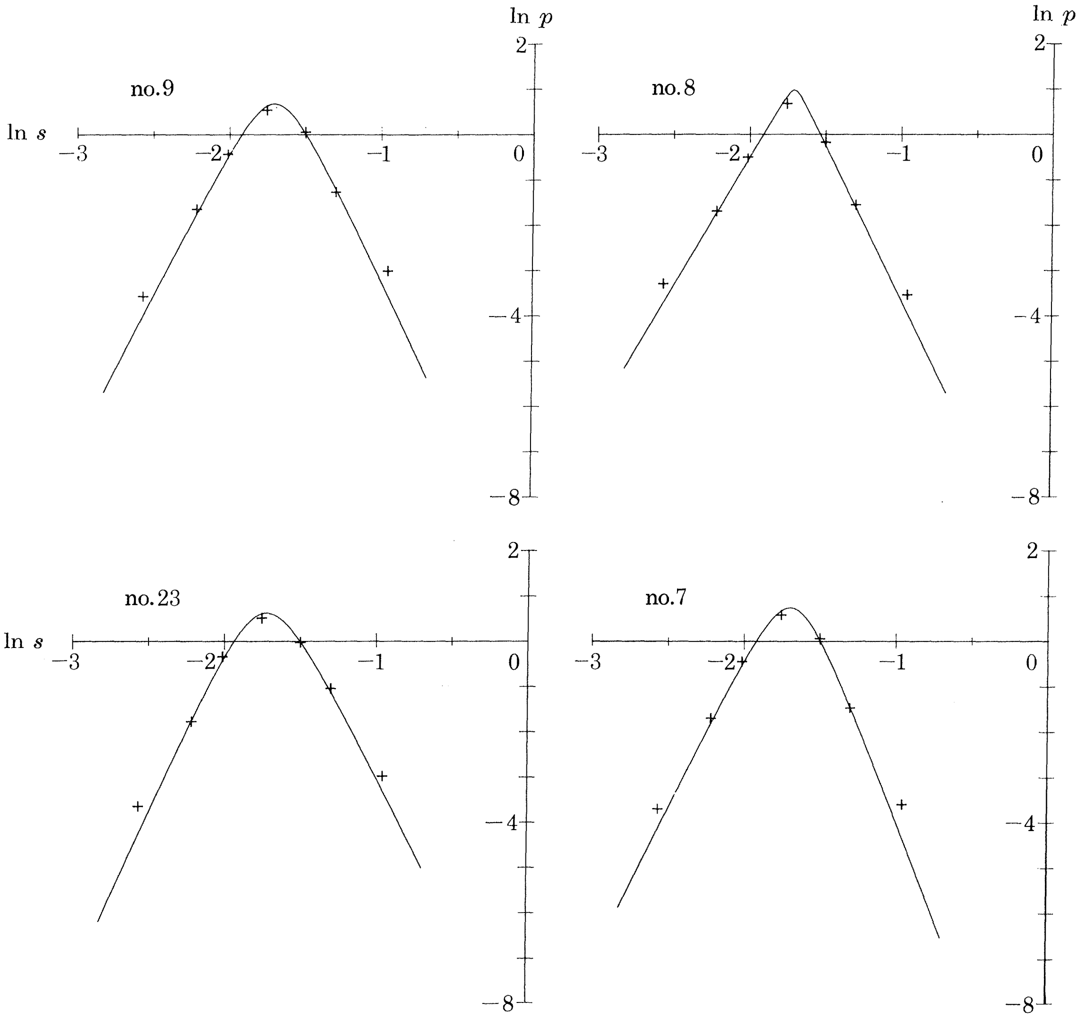

FIGURE 4c (For description see facing page)

besides treating economic questions, also indicate various Paretian phenomena in geology, geography, meteorology and physics.

(ii) In an investigation of the activity of single nerve cells in the visual system of cats, Herz, Creutzfeldt & Fuster (1964) recorded the time intervals between firings of a single neurone. For neurones from the visual cortex, the empirical

Downloaded from rspa.royalsocietypublishing.org on October 19, 2014

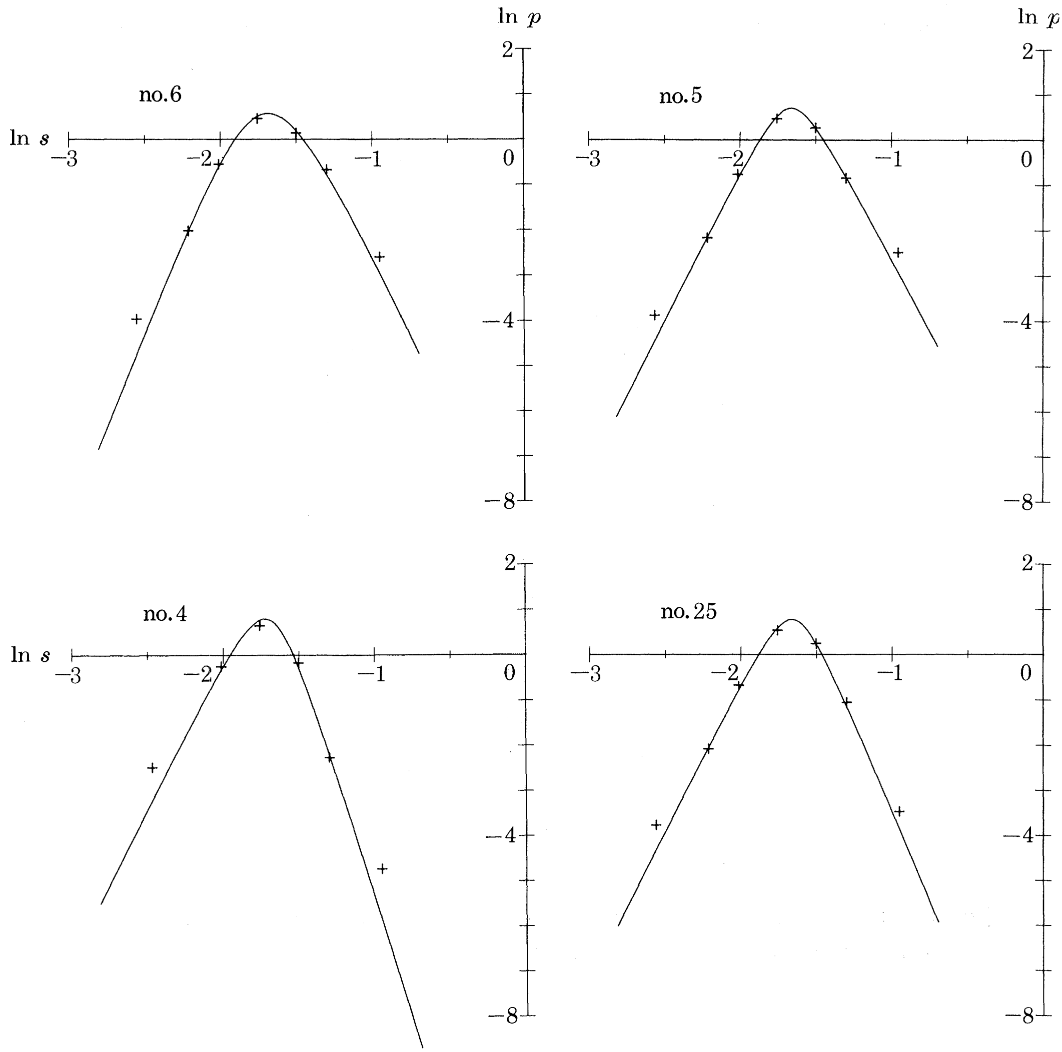

FIGURE 4. The mass-size distributions (histograms) of 16 sand samples from the Danish west coast, together with the fitted 'log-hyperbolic' distributions, in double-logarithmic plotting. (Numbering of samples as in table 3, and Kuhlman (1957).)

distributions of the time intervals typically had a shape close to that of the distribution (4.7) introduced here.

According to Herz *et al.* more than one frequency mechanism is active in determining the observed characteristic shape.

For a discussion relating measurements of a similar kind to the inverse Gaussian distribution (5.4), see Gerstein & Mandelbrot (1964).

(iii) For fixed μ and δ, and (ϕ, γ) varying in $(0, \infty)^2$ the class of distributions (4.7) is an exponential family of order 2 and regular (in the sense of Barndorff-Nielsen (1973)).

Downloaded from rspa.royalsocietypublishing.org on October 19, 2014

In particular, therefore, setting $\mu = 0$ and $\delta = 1$ the mean values (and higher order moments) of v and $\surd(1+v^2)$ are obtainable from the cumulant transform

$$\zeta(\phi, \gamma) = \ln(\omega^{-1}\kappa K_1(\kappa))$$

by taking derivatives with respect to $\frac{1}{2}(\phi-\gamma)$ and $-\frac{1}{2}(\phi+\gamma)$, respectively. One finds

$$Ev = \tfrac{1}{2}(\gamma^{-1}-\phi^{-1})\,(1-\kappa K_1'(\kappa)/K_1(\kappa)),$$

and $$E\surd(1+v^2) = \tfrac{1}{2}(\phi^{-1}+\gamma^{-1})\,(1-\kappa K_1'(\kappa)/K_1(\kappa))-2/(\phi+\gamma).$$

Table 3. Maximum likeness estimates for 16 sand samples from the Danish west coast

sample no.	$-\hat{\mu}$	$\hat{\delta}$	$\hat{\phi}$	$\hat{\gamma}$
41	1.78	0.46	13.2	9.8
43	1.74	0.34	12.2	7.4
17	1.72	0.20	8.2	6.5
15	1.73	0.19	9.7	5.3
13	1.71	0.13	6.6	8.8
12	1.66	0.17	7.2	10.5
11	1.65	0.13	7.2	7.9
10	1.74	0.18	7.6	6.9
9	1.70	0.18	6.8	7.1
8	1.71	0.03	5.7	6.8
23	1.74	0.23	7.7	6.8
7	1.68	0.18	7.0	8.7
6	1.71	0.31	8.6	7.3
5	1.66	0.15	6.7	6.3
4	1.68	0.18	6.8	11.1
25	1.65	0.15	6.7	8.1

(iv) Let (u, v) be a two dimensional, continuous type random variate following the isotropic distribution whose probability function is

$$\frac{\alpha^2}{2\pi}\exp(-\alpha\surd(u^2+v^2)).$$

Obviously, the conditional distribution of v given $u = \delta$ is of the type (4.7), with $\mu = \beta = 0$.

(v) The distributions (5.3) and (5.4) both belong to the family of 'generalized inverse Gaussian distributions' having parameter (λ, χ, ψ) and probability function

$$\frac{(\psi/\chi)^{\frac{1}{2}\lambda}}{2K_\lambda(\surd(\chi\psi))}x^{\lambda-1}\exp(-\tfrac{1}{2}(\chi x^{-1}+\psi x)) \quad (x>0) \tag{7.1}$$

(where $K_\lambda(\cdot)$ is the modified Bessel function of the third kind, with index λ). Furthermore, setting $\lambda = -2$ and $\psi = 0$, and truncating the distributions to an interval $(0, a)$ one obtains a distributional law proposed by Troesch & Grassmann (1963) as a model for the distribution of the diameters of droplets generated by

Downloaded from rspa.royalsocietypublishing.org on October 19, 2014

atomizers of various types. The authors derived the law by a reasoning analogous to the derivation of Boltzmann–Maxwell laws in statistical mechanics.

The family (7.1) has recently been considered by Sichel (1973, 1974, 1975), see also Wise (1975). Sichel used the family for constructing a flexible mixture of the Poisson distributions.

Mixing the normal distributions (5.1) by specifying (5.2) together with a generalized inverse Gaussian distribution (7.1) for σ^2 one obtains as probability function for v

$$(\surd(2\pi)\,(\chi/\psi)^{\frac{1}{2}\lambda}\,(\beta^2+\psi)^{\frac{1}{2}(\lambda-\frac{1}{2})}\,K_\lambda(\surd(\chi\psi))^{-1}\,(\chi+(v-\mu)^2)^{\frac{1}{2}(\lambda-\frac{1}{2})}$$
$$\times K_{\lambda-\frac{1}{2}}(\surd\{(\beta^2+\psi)\,(\chi+(v-\mu)^2)\})\exp(\beta(v-\mu)).$$

This includes (4.7) as a special case ($\lambda = 1$), cf. §5. Another case of interest, relating to subordination of stochastic processes (Feller 1966), is given by $\mu = \beta = \psi = 0$, $\lambda = -\frac{1}{2}$ which yields the scaled Cauchy distribution.

It may also be noted that if x and y are independent random variates following gamma distributions with parameters $(\xi, 2\rho^{-1})$ and $(\eta, 2\rho)$, respectively, so that their joint probability function is

$$\frac{\rho^{\xi-\eta}}{\Gamma(\xi)\,\Gamma(\eta)\,2^{\xi+\eta}}x^{\xi-1}y^{\eta-1}\exp\left(-\tfrac{1}{2}(x\rho+y\rho^{-1})\right)$$

then the distribution of $t = \surd(y/x)$ given $u = \surd(xy)$ is of the form (7.1) with $\lambda = \eta-\xi$, $\chi = u\rho$ and $\psi = u/\rho$. For $\eta = \xi = 1$ this conditional distribution occurs in an example produced by Fisher (1958, ch. 6), to illustrate the concept of ancillarity and the fiducial argument, the statistic u being ancillary.

(vi) A possible multivariate extension of (4.7) is the n-dimensional distribution with probability function of the form

$$a(\alpha,\beta,\delta,\varDelta)\exp\{-\alpha\surd(\delta^2+(v-\mu)\,\varDelta^{-1}(v-\mu)')+\beta(v-\mu)'\}, \tag{7.2}$$

where v, μ and β are n-dimensional (row) vectors and $\varDelta$ is a positive definite $n\times n$ matrix, while the symbol $'$ indicates transposition. Moreover, with $\kappa^2 = \alpha^2-\beta\varDelta\beta'$ the norming constant is

$$a(\alpha,\beta,\delta,\varDelta) = \frac{1}{(2\pi)^{\frac{1}{2}(n-1)}\surd(|\varDelta|)}\cdot\frac{\kappa^{\frac{1}{2}(n+1)}}{2\alpha\,\delta^{\frac{1}{2}(n+1)}K_{\frac{1}{2}(n+1)}(\delta\kappa)},$$

$|\varDelta|$ denoting the determinant of $\varDelta$. An explicit expression for $K_{\frac{1}{2}(n+1)}(\delta\kappa)$ exists for n even (cf., for instance, Erdélyi *et al.* (1953)).

The conditional distribution given any subset of the coordinates of v is again of the type (7.2).

Like in the one dimensional case, the distribution may be generated as a mixture of normal distributions. Suppose that v given a scalar random variate u follows the n-dimensional normal distribution with mean (vector) $\mu+u\beta\varDelta$ and variance (matrix) $u\varDelta$. If u has the generalized inverse Gaussian distribution (cf. (7.1)) with parameters

Downloaded from rspa.royalsocietypublishing.org on October 19, 2014

$(\lambda, \delta^2, \kappa^2)$ then the marginal probability function for v is

$$\frac{1}{(2\pi)^{\frac{1}{2}n}\sqrt{|\Delta|}} \cdot \frac{\kappa^\lambda}{\delta^\lambda K_\lambda(\delta\kappa)} \cdot \frac{K_{\frac{1}{2}n-\lambda}\{\alpha\sqrt{(\delta^2+(v-\mu)\,\Delta^{-1}(v-\mu)')}\}}{\{\alpha\sqrt{(\delta^2+(v-\mu)\,\Delta^{-1}(v-\mu)')}\}^{\frac{1}{2}n-\lambda}} \exp(\beta(v-\mu)'). \quad (7.3)$$

For $\lambda = \frac{1}{2}(n+1)$ this reduces to (7.1).

From this mixture representation and from (7.3) it follows, on account of the (normal) form of the marginal distributions of v given u, that the distribution (7.2) does not have marginals of its own type. One sees, moreover, that an n-dimensional distribution whose m-dimensional marginals are of the type (7.2) is obtainable by setting $\lambda = \frac{1}{2}(m+1)$ in (7.3).

I am indebted to Hans Kuhlman for putting the original data from his Skallingen investigation at my disposal, and to Hanne Østergaard Kristensen and Jens Peter Andersen for assistance with the numerical and graphical work.

References

Andrews, D. F. & Mallows, C. L. 1974 Scale mixtures of normal distributions. *J. R. Statist. Soc.* **36**, 99–102.

Bagnold, R. A. 1954 *The physics of blown sand and desert dunes*. London: Methuen.

Bagnold, R. A. 1956 The flow of cohesionless grains in fluids. *Phil. Trans. R. Soc. Lond.* A **249**, 235–297.

Barndorff-Nielsen, O. 1973 Exponential families and conditioning. Sc.D. thesis, University of Copenhagen.

Chambers, J. M. 1973 Fitting nonlinear models: numerical techniques. *Biometrika* **60**, 1–13.

Champernowne, D. G. 1973 *The distribution of income between persons*. Cambridge University Press.

Cox, D. R. & Miller, H. D. 1965 *The theory of stochastic processes*. London: Methuen.

Epstein, B. 1947 The mathematical description of certain breakage mechanisms leading to the logarithmico-normal distribution. *J. Franklin Inst.* **224**, 471–477.

Erdélyi, A. *et al.* 1953 *Higher transcendental functions*, vol. II. New York: McGraw-Hill.

Feller, W. 1966 *An introduction to probability theory and its applications*, vol. II. London: Wiley.

Fisher, R. A. 1956 *Statistical methods and scientific inference*. Edinburgh: Oliver & Boyd.

Gerstein, G. L. & Mandelbrot, B. 1964 Random walk models for the spike activity of a single neuron. *Biophysical J.* **4**, 41–68.

Halmos, P. R. 1944 Random Alms. *Ann. Math. Statist.* **15**, 182–189.

Herz, A., Creutzfeldt, O. & Fuster, J. 1964 Statistische Eigenschaften der Neuronaktivität im ascendierenden visuellen System. *Kybernetik* **2**, 61–71.

Kolmogoroff, A. N. 1941 Über das logarithmisch normale Verteilungsgesetz der Dimensionen der Teilchen bei Zerstückelung. *C. R. Acad. Sci. U.R.S.S.* **31**, 99–101.

Ku, H. H. & Kullback, S. 1974 Log-linear models in contingency table analysis. *Am. Statistician* **28**, 115–122.

Kuhlman, H. 1957 Kornstørrelser i klit- og strandsand. (With an English summary). *Geogr. Tidsskr.* **56**, 20–56.

Kullback, S. 1959 *Information theory and statistics*. New York: Wiley.

Mandelbrot, B. 1963 New methods in statistical economics. *Bull. Inst. Int. Stat.* **40** (2), 699–721.

Rényi, A. 1960 Az apritás matematikai elméletéröl. *Épitöanyag*, pp. 1–8.

Downloaded from rspa.royalsocietypublishing.org on October 19, 2014

Savage, L. J. 1954 *The foundations of statistics*. New York: Wiley.
Sichel, H. S. 1973 Statistical valuation of diamondiferous deposits. *Jl. S. Afr. Inst. Min. Metall.* **73**, 235–243.
Sichel, H. S. 1974 On a distribution representing sentence-length in written prose. *J. R. Statist. Soc.* A **137**, 25–34.
Sichel, H. S. 1975 On a distribution law for word frequencies. *J. Am. statist. Ass.* **70**, 542–547.
Steindl, J. 1965 *Random processes and the growth of firms*. London: Griffin.
Teichroew, D. 1957 The mixture of normal distributions with different variances. *Ann. Math. Statist.* **28**, 510–512.
Troesch, H. A. & Grassmann, P. 1953 Zum Verteilungsgesetz der Tropfengrössen bei der Zerstäubung. *Z. angew. Math. Phys.* **4**, 81–85.
Wise, M. E. 1975 Skew distributions in biomedicine including some with negative powers of time. *Statistical distributions in scientific work*, vol. II (eds G. P. Patil *et al.*), pp. 241–262.