V. A FAMILY OF DISTRIBUTIONS DESCRIBING SEISMIC DATA

A. Generalized Gaussian Distribution

The reflection seismogram can be modeled as the output of a linear filter, given the earth's reflectivity series as input. The linear filter generally represents the source waveform and other undesirable averaging effects. Deconvolution is an attempt to estimate the inputs given the outputs. To solve the problem some property of the filter or inputs must be assumed.

Here it is assumed reflection amplitudes and averaging effects are well described by members of a family of symmetrical probability distributions defined by

$$f(x,\beta,\alpha) = \frac{\alpha}{2\beta\Gamma\left(\frac{1}{\alpha}\right)} e^{-\left(\left|\frac{x}{\beta}\right|\right)^{\alpha}} \tag{5-1}$$

where

$-\infty < x < \infty$ is a random reflection amplitude;

$\Gamma(\cdot)$ is the gamma function;

$\beta > 0$ is the scale parameter; and

$\alpha > 0$ is the shape parameter.

The family will be called the *generalized Gaussian* after Miller and Thomas (14) who used it in detection theory as a model for non-Gaussian noise. The statisticians initially called it *Subbotin's distribution* (15). It was rarely used until recently when Box and Tiao (16) examined its

properties as a prior distribution for Bayesian analysis. Regression models based on the family have been investigated by Goodman and Kotz (17).

The generalized Gaussian covers a wide range of distributions. For $\alpha = 1$, we have the double exponential distribution, also called Laplace's distribution,

$$f(x,\beta,1) = \frac{1}{2\beta} e^{-\frac{|x|}{\beta}} , \quad -\infty < x < \infty . \tag{5-2}$$

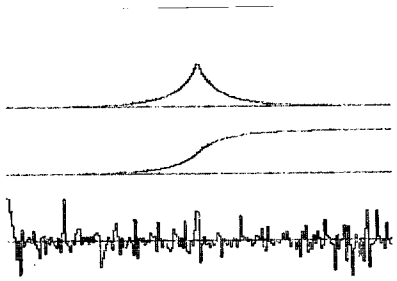
For $\alpha = 2$ we have the Gaussian distribution,

$$f(x,\beta,2) = \frac{1}{\beta\sqrt{\pi}} e^{-\left(\frac{|x|}{\beta}\right)^2} , \quad -\infty < x < \infty \tag{5-3}$$

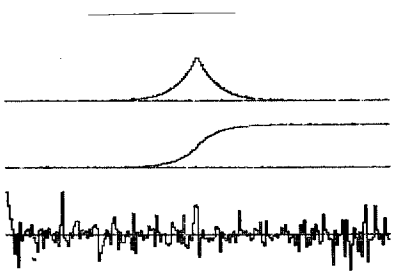
where $\beta^2 = 2\sigma^2$. As α tends toward infinity, we have the uniform distribution,

$$f(x,\beta,\infty) = \frac{1}{2\beta} , \quad -\beta \leq x \leq +\beta \tag{5-4}$$

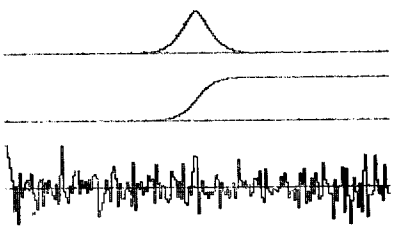
and as α tends towards zero the chances of finding an event in a finite sample goes to zero. Such a distribution can be loosely described as the certain event. Figure 5-1 displays members of the family as a function of the shape parameter.



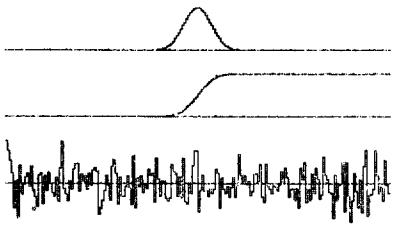
$\alpha = .8$



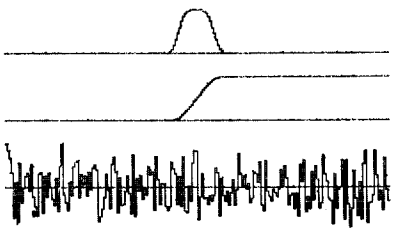
$\alpha = 1.$
(double exponential)



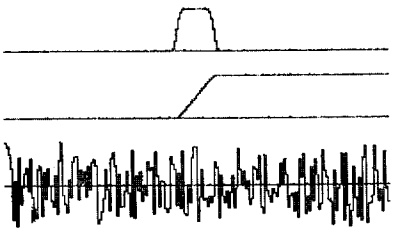
$\alpha = 1.5$



$\alpha = 2.$
(Gaussian)



$\alpha = 4.$



$\alpha = 8.$

FIGURE 5-1. Members of the generalized Gaussian family of distributions. For each shape parameter the probability density function is plotted above the cumulative distribution and a representative random sample. The shape parameter determines the rate of exponential decay for the pdf.

For seismic observations several random independent reflection amplitudes are described by a joint distribution. Figure 5-2 displays two-dimensional distributions for several members of the generalized Gaussian. In n-dimensional space, contours of equal probability for the uniform are hypercubes; the Gaussian, hyperspheres; and the double exponential, regular octahedra. When the shape parameter drops below two, the contours sag inward towards the axes. As the parameter tends to zero, the probability mass concentrates along the coordinate axes.

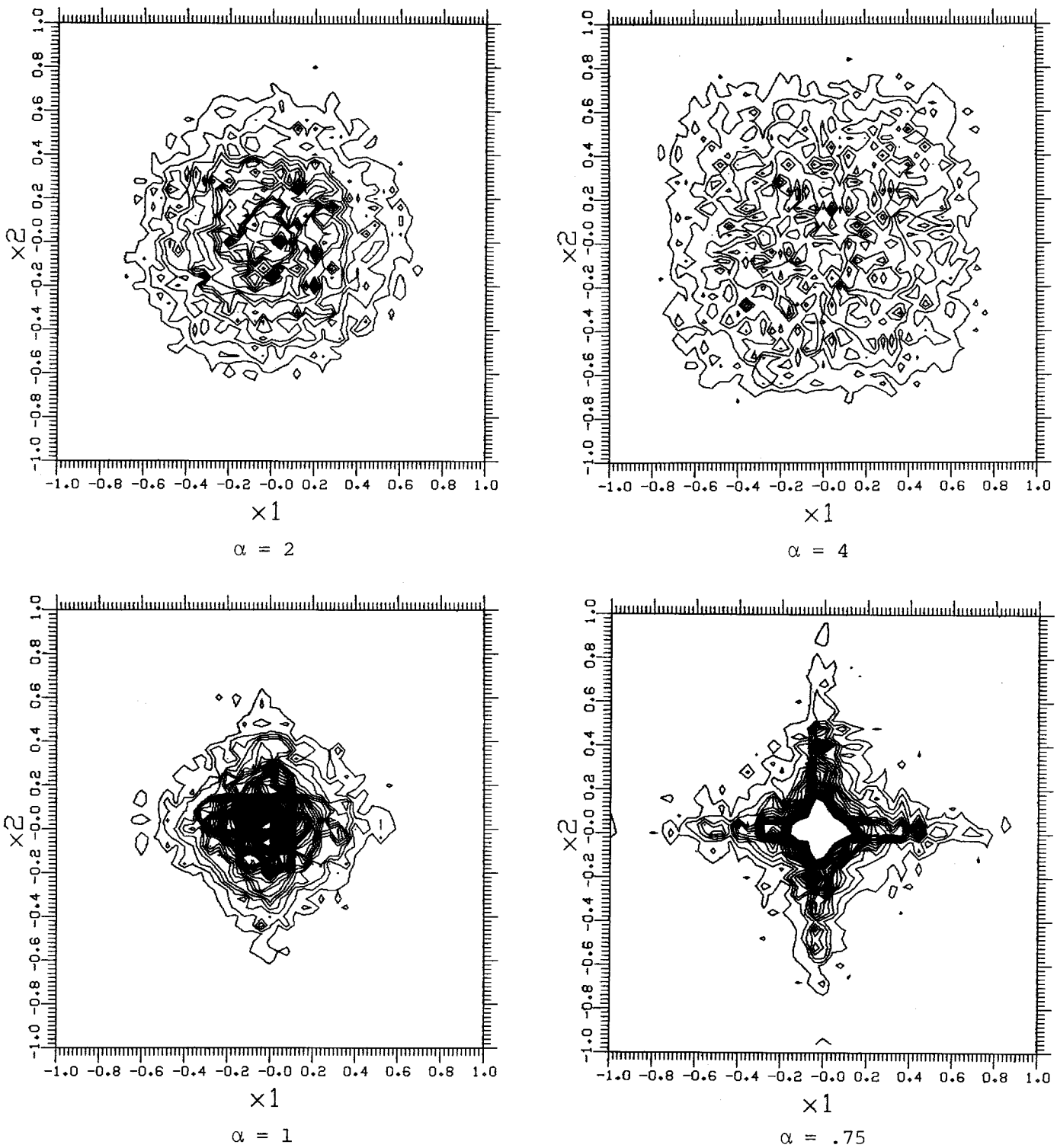


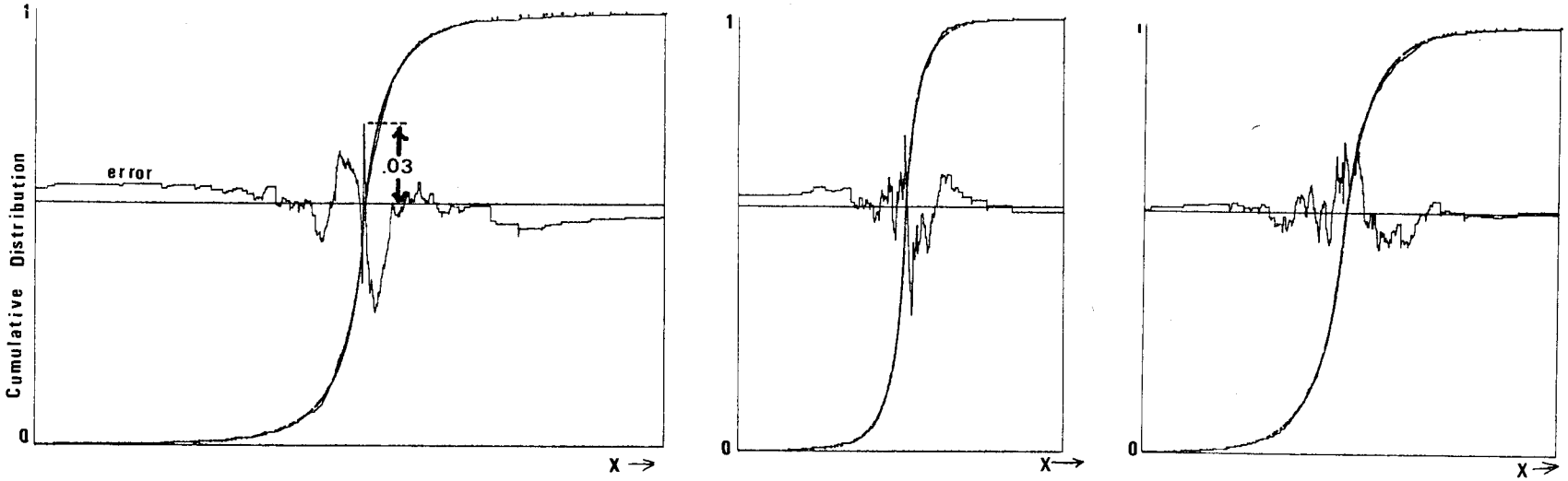
FIGURE 5-2. Joint density functions for members of the generalized Gaussian family. Five thousand pairs of random numbers for each member were placed into bins indexed by (x1, x2); the contour interval is four events. A 50 x 50 grid was used where the edges corresponded to the largest value in the sample.

A measure of the parsimony or spikiness for a seismogram is given by estimating its shape parameter. The smaller the estimate, the more parsimonious or spiky the seismogram. A brief explanation of this assumes a two-point seismogram. When the underlying distribution is uniform, the realizations (0,1) and (1,1) are equally likely. As the shape parameter of the underlying distribution decreases, the realization (1,1) becomes less likely while realizations near the coordinate axes increase in likelihood. The deconvolution method proposed tries to estimate a linear operator that will yield an output with the smallest possible shape parameter.

B. Testing the Distribution of Seismic Data

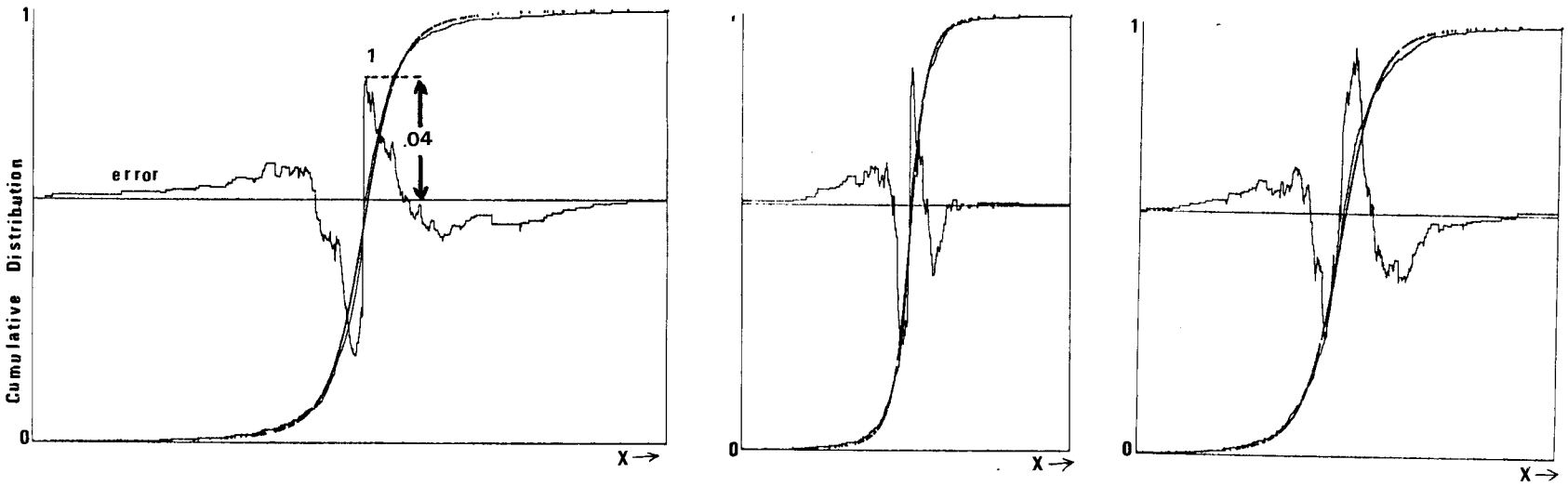
According to Algazi and Lerner (18) the family for $.1 < \alpha < .6$ is representative of atmospheric processes. Using the Kolmogorov-Smirnov (19) test, Godfrey, a coworker, has shown the family to be fairly representative of both reflectivity series and reflection seismograms. The reflectivity series of Figure 4-1 were tested by Godfrey and his results are displayed in Figure 5-3. They show fair agreement between the estimated members of the generalized Gaussian and the actual distributions. Typically, reflectivity series have a shape parameter in the range .4 to 1.5. Gained seismograms, being more Gaussian, have shape parameters in the range .9 to 2.5.

Figure 5-4 is a set of results obtained by testing the t-family (20) of distributions. This family ranges from the heavy-tailed Cauchy distribution to the Gaussian distribution. Comparison of Figures 5-3 and 5-4 shows a much closer fit achieved by the generalized Gaussian. This is believed due to its peakedness at the origin for $\alpha < 2$. Members of the



Well A	Well B	Well C
$\hat{\alpha}$ = .583	$\hat{\alpha}$ = .644	$\hat{\alpha}$ = .791
Rms error = .0155	Rms error = .0123	Rms error = .0099
K^+ = 1.104	K^+ = .891	K^+ = .816
K^- = 1.369	K^- = .899	K^- = .381

FIGURE 5-3. Results obtained by Godfrey using the K-S test to determine the fit of the generalized Gaussian to the reflectivity series of Figure 5-1. Displayed are the actual cumulative distributions, those estimated using the generalized Gaussian, and the rescaled error. The K-S statistics K^+ and K^- indicate a reasonable fit.



Well A	Well B	Well C
Rms error = .0271	Rms error = .0260	Rms error = .0254
K^+ = 1.373	K^+ = 1.157	K^+ = 1.538
K^- = 1.786	K^- = 1.268	K^- = 1.186

FIGURE 5-4. Similar results to those of Figure 5-3. The t-family fits the actual distribution poorly in the center relative to the generalized Gaussian. They both agree closely on the tails. The K-S statistics, which are based on the maximum positive and negative errors, indicate the fit is not as good as that achieved with the generalized Gaussian.

t-family have a bell shape at the origin which is not typical of seismic data.

These results are based on a small number of tests. Other distributions may be found that fit particular datasets closer. The generalized Gaussian has computational and mathematical properties which make it attractive. A survey article by Hettmansperger and Keeran (21) gives an excellent discussion of other families of distributions and tailweight.

C. Estimators

In the last section the fit of the generalized Gaussian to actual reflectivity series was examined. Here we show how the parameters of the distribution are estimated from an arbitrary sample. First, a technique for estimating the shape parameter is developed.

The λ-th absolute moment for the generalized Gaussian is derived from

$$E(|x|^{\lambda}) = \int_{-\infty}^{\infty} |x|^{\lambda} f(x,\alpha,\beta)\, dx. \tag{5-5}$$

where $E(\cdot)$ denotes the expectation operator.

Performing the integration yields

$$E(|x|^{\lambda}) = \frac{\Gamma\left(\frac{\lambda+1}{\alpha}\right)}{\Gamma\left(\frac{1}{\alpha}\right)} \beta^{\lambda} . \tag{5-6}$$

The variance, σ^2, is the second absolute moment, (λ=2)

$$\sigma^2 = \frac{\Gamma\left(\frac{3}{\alpha}\right)}{\Gamma\left(\frac{1}{\alpha}\right)} \beta^2 . \tag{5-7}$$

Kurtosis is traditionally defined as the ratio of the fourth sample moment and second sample moment squared.

$$K(X) = \frac{\frac{1}{n}\sum_{i=1}^{n} x_i^4}{\left(\frac{1}{n}\sum_{i=1}^{n} x_i^2\right)^2} \quad . \tag{5-8}$$

It has long been used as a statistic to describe the tailweight of a random sample and for testing departure from normality.

"Generalized kurtosis," $K(\rho,X)$, is also a measure of tailweight. It is defined as

$$K(\rho,X) = \frac{\frac{1}{n}\sum_{i=1}^{n} |x_i|^{2\rho}}{\left(\frac{1}{n}\sum_{i=1}^{n} |x_i|^{\rho}\right)^2} \quad . \tag{5-9}$$

This statistic is useful here because it provides an efficient scale invariant method of estimating a shape parameter best describing a given random sample. Using Equation (5-6) gives

$$\frac{E(|x|^{2\rho})}{[E(|x|^{\rho})]^2} = \frac{\Gamma\left(\frac{2\rho+1}{\alpha}\right)\Gamma\left(\frac{1}{\alpha}\right)}{\Gamma\left(\frac{\rho+1}{\alpha}\right)^2} \quad . \tag{5-10}$$

Combining Equations (5-9) and (5-10) gives

$$\frac{\frac{1}{n}\sum_{i=1}^{n}|x_i|^{2\rho}}{\left(\frac{1}{n}\sum_{i=1}^{n}|x_i|^{\rho}\right)^2} \cong \frac{\Gamma\left(\frac{2\rho+1}{\alpha}\right)\Gamma\left(\frac{1}{\alpha}\right)}{\Gamma\left(\frac{\rho+1}{\alpha}\right)^2} \tag{5-11}$$

where equality is obtained as the sample size goes to infinity.

Solving this equation for α is probably impossible but in a computer the RHS can be evaluated for many values of α and saved in a table. The LHS is computed from the random sample and a shape parameter determined by interpolation on the table. Figure 5-5 is a plot of K(2,X) versus α.

K(ρ,X) is a family of estimators. Geary (22) has shown that K(2,X) is most efficient when no information on the distribution of the random sample is available. For seismic data we have information on the expected distribution and from Monte Carlo testing we have determined that K(1,X) is a more efficient estimator than K(2,X). Box (23) also concluded that K(1,X) would be a better estimator for a finite sample having greater kurtosis than the Gaussian.

An estimator for the scale is now derived using maximum likelihood. The estimate is valid assuming the reflectivity series behaves as a random sample from the distribution characterized by α. Define L, the log likelihood for the generalized Gaussian, as

$$L = \log \prod_{i=1}^{n} f(x_i,\beta,\alpha) = \log \alpha^n - \log\left[2\beta\Gamma\left(\frac{1}{\alpha}\right)\right]^n - \frac{\sum_{i=1}^{n}|x_i|^{\alpha}}{\beta^{\alpha}} . \tag{5-12}$$

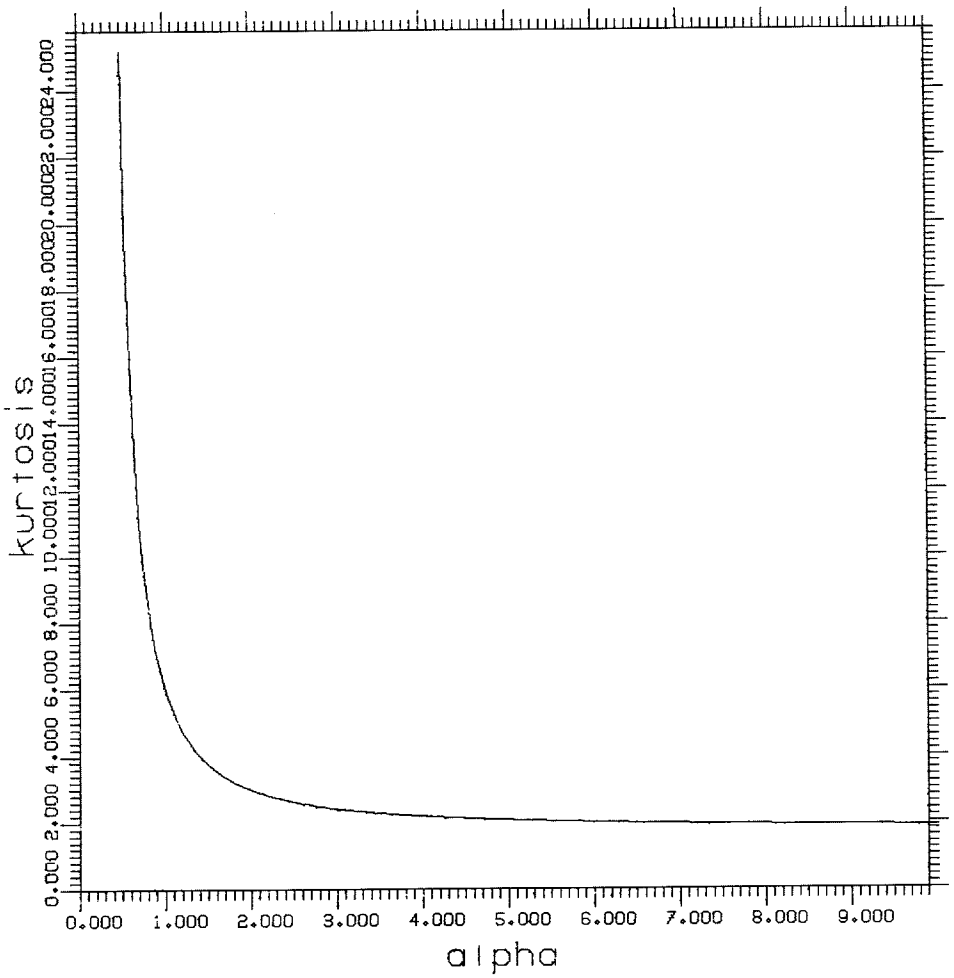


FIGURE 5-5. The generalized Gaussian family is ordered by kurtosis. K(2,X) is plotted versus α, the shape parameter.

The maximum likelihood estimator for the scale parameter, β, is the solution of

$$\frac{\partial L}{\partial \beta} = 0 \ . \tag{5-13}$$

Solving the equation gives the estimator

$$\beta = \left(\frac{\alpha}{n} \sum_{i=1}^{n} |x_i|^{\alpha} \right)^{\frac{1}{\alpha}} \ . \tag{5-14}$$

This yields the mean deviation in the case of the double exponential ($\alpha = 1$). For the normal distribution ($\alpha = 2$), β^2 is estimated by twice the square of the standard deviation and for the uniform ($\alpha \to \infty$), β is estimated by one-half of the sample range.

An estimator for the shape parameter α, assuming the scale is known, was derived from

$$\frac{\partial L}{\partial \alpha} = 0 \ . \tag{5-15}$$

Its effectiveness for estimating α was evaluated and surprisingly the method given by Equation (5-11) was more accurate. Equation (5-11) is also computationally more efficient because the solution to Equation (5-15) involves derivatives of the gamma function and is fairly complex.

D. Relation to Entropy

Perhaps the most interesting property of the generalized Gaussian family is that for a fixed α-th absolute moment the distribution $f(x,\beta,\alpha)$ has maximum entropy. This is shown by determining a function $f(x)$ such

that the entropy

$$-\int_{-\infty}^{\infty} f(x) \log f(x)\, dx \tag{5-16}$$

is maximized subject to the constraints that the function has unit area

$$\int_{-\infty}^{\infty} f(x)\, dx = 1 \tag{5-17}$$

and the α-th absolute moment, μ_α, is fixed.

$$\int_{-\infty}^{\infty} |x|^\alpha f(x)\, dx = \mu_\alpha \quad . \tag{5-18}$$

This problem is solved using the calculus of variations. Define the integral,

$$I = \int_{-\infty}^{\infty} [-f(x) \log f(x) + \lambda_1 f(x) + \lambda_2 |x|^\alpha f(x)]\, dx \tag{5-19}$$

where λ_1 and λ_2 are Lagrange multipliers. Its variation is

$$\delta I = -\frac{\partial f(x)}{\partial f} \log f(x) + \lambda_1 \frac{\partial f(x)}{\partial f} + \lambda_2 |x|^\alpha \frac{\partial f(x)}{\partial f} \tag{5-20}$$

or

$$\delta I = -[1 + \log f(x)] + \lambda_1 + \lambda_2 |x|^\alpha \quad .$$

Entropy is maximized when $\delta I = 0$ so

$$-[1 + \log f(x)] + \lambda_1 + \lambda_2 |x|^\alpha = 0 \tag{5-21}$$

is solved for f, yielding

$$f(x) = e^{(\lambda_1 - 1)} \; e^{\lambda_2 |x|^{\alpha}} \quad . \tag{5-22}$$

Substituting this result in Equation (5-17) we have

$$\int_{-\infty}^{\infty} e^{(\lambda_1 - 1)} \; e^{\lambda_2 |x|^{\alpha}} \, dx = 1 \quad . \tag{5-23}$$

Solving the equation gives

$$e^{\lambda_1 - 1} = \frac{\alpha(-\lambda_2)^{\frac{1}{\alpha}}}{2\Gamma(\frac{1}{\alpha})} \tag{5-24}$$

which when substituted into Equation (5-22) gives the function in terms of λ_2:

$$f(x) = \frac{\alpha(-\lambda_2)^{\frac{1}{\alpha}}}{2\Gamma(\frac{1}{\alpha})} \; e^{\lambda_2 |x|^{\alpha}} \quad . \tag{5-25}$$

The second constraint, Equation (5-18), is used to determine λ_2:

$$\mu_{\alpha} = \int_{-\infty}^{\infty} |x|^{\alpha} \left[\frac{\alpha(-\lambda_2)^{\frac{1}{\alpha}}}{2\Gamma(\frac{1}{\alpha})} \; e^{\lambda_2 |x|^{\alpha}} \right] dx \quad . \tag{5-26}$$

Solving this equation gives

$$\lambda_2 = - \frac{1}{\alpha \mu_{\alpha}} \tag{5-27}$$

and the function can be written

$$f(x) = \frac{\alpha}{2\beta\Gamma(\frac{1}{\alpha})} \; e^{-\left(\frac{|x|}{\beta}\right)^{\alpha}} \tag{5-28}$$

where

$$\beta = (\alpha\mu_\alpha)^{\frac{1}{\alpha}}$$

which is the generalized Gaussian.

This result implies that entropy is maximized by the double exponential for fixed L_1 norm, by the Gaussian when L_2 is fixed, and by the uniform when L_∞ (the maximum value) is fixed.

For a given variance the family of distributions can be parameterized by Shannon's entropy. Let entropy be defined by

$$H(\beta,\alpha) = -\int_{-\infty}^{\infty} f(x,\beta,\alpha) \log f(x,\beta,\alpha)\, dx \qquad (5\text{-}29)$$

where f is the distribution given by Equation (5-1). Performing the integration yields

$$H(\beta,\alpha) = \log\left(\frac{2\beta\Gamma(\frac{1}{\alpha})}{\alpha}\right) + \frac{1}{\alpha} \quad . \qquad (5\text{-}30)$$

Substituting Equation (5-7) into this result gives entropy as a function of variance and the shape parameter,

$$H(\alpha,\sigma^2) = \log\left(\frac{2}{\alpha}\sqrt{\frac{\sigma^2\Gamma(\frac{1}{\alpha})^3}{\Gamma(\frac{3}{\alpha})^2}}\right) + \frac{1}{\alpha} \qquad (5\text{-}31)$$

This function is shown in Figure 9 for different members of the family with constant variance. The Gaussian has maximum entropy as expected. Entropy decreases on both sides. As $\alpha \to 0$ it rapidly goes to $-\infty$ which is the entropy of the certain event. On the other side it falls off slowly to another minimum, the uniform distribution. A similar function was derived for entropy as a function of the first absolute moment, and also shown in Figure 5-6.

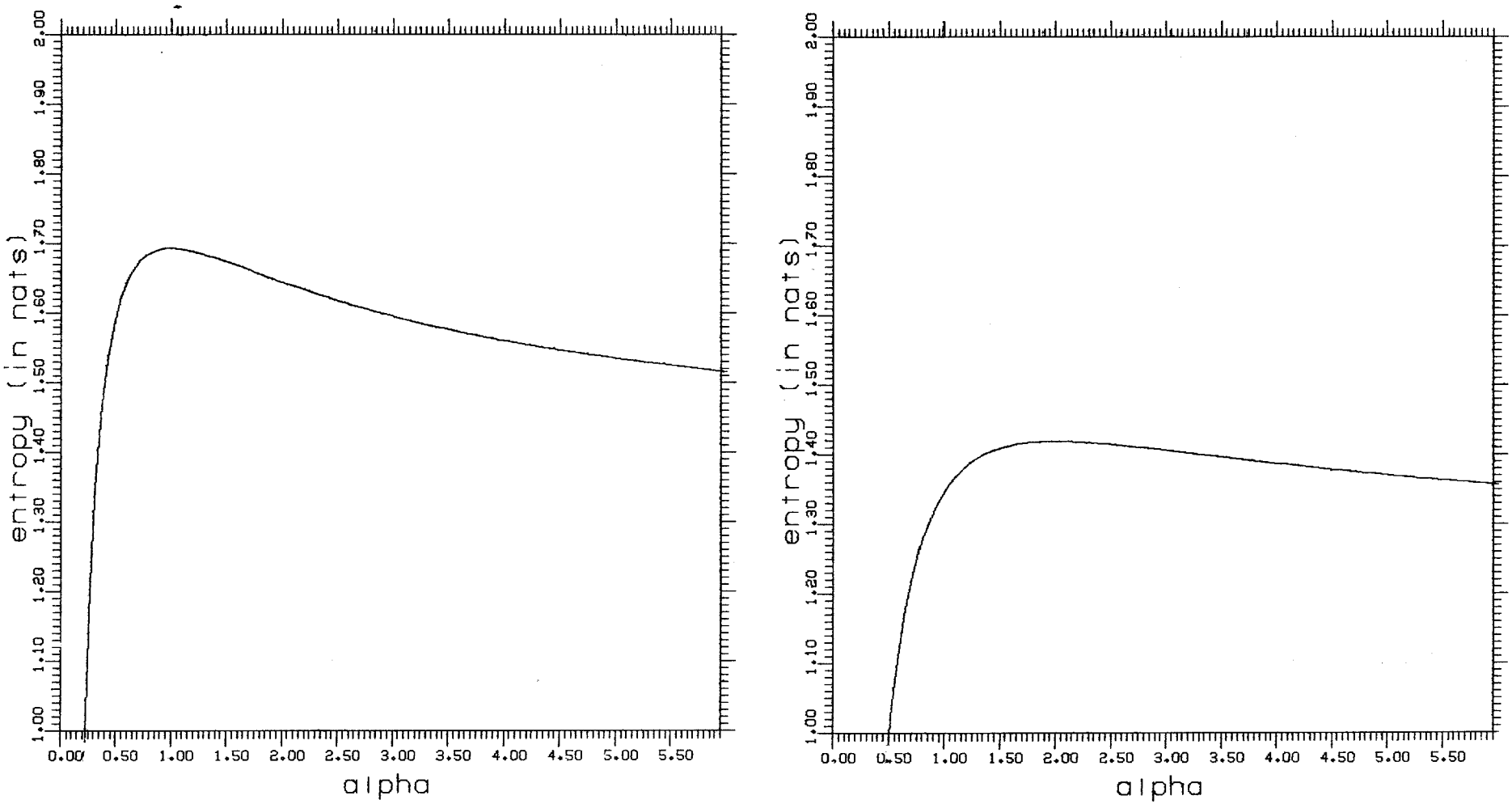


FIGURE 5-6. Entropy of members of the generalized Gaussian family. On the left the first absolute moment was held fixed and the double exponential distribution ($\alpha = 1$) has maximum entropy. On the right the second absolute moment (variance) was held fixed, and the Gaussian has maximum entropy.

E. Most Powerful Invariant Test

In statistics the most powerful test of the null hypothesis against another is one which maximizes discrimination and minimizes *Type II error*. Type II error is the probability of accepting the null hypothesis as true when actually the alternative is true. In information theory a statistic that maximizes detectability of a signal while minimizing the probability of a false alarm is most powerful. These concepts are illustrated by Figure 5-7.

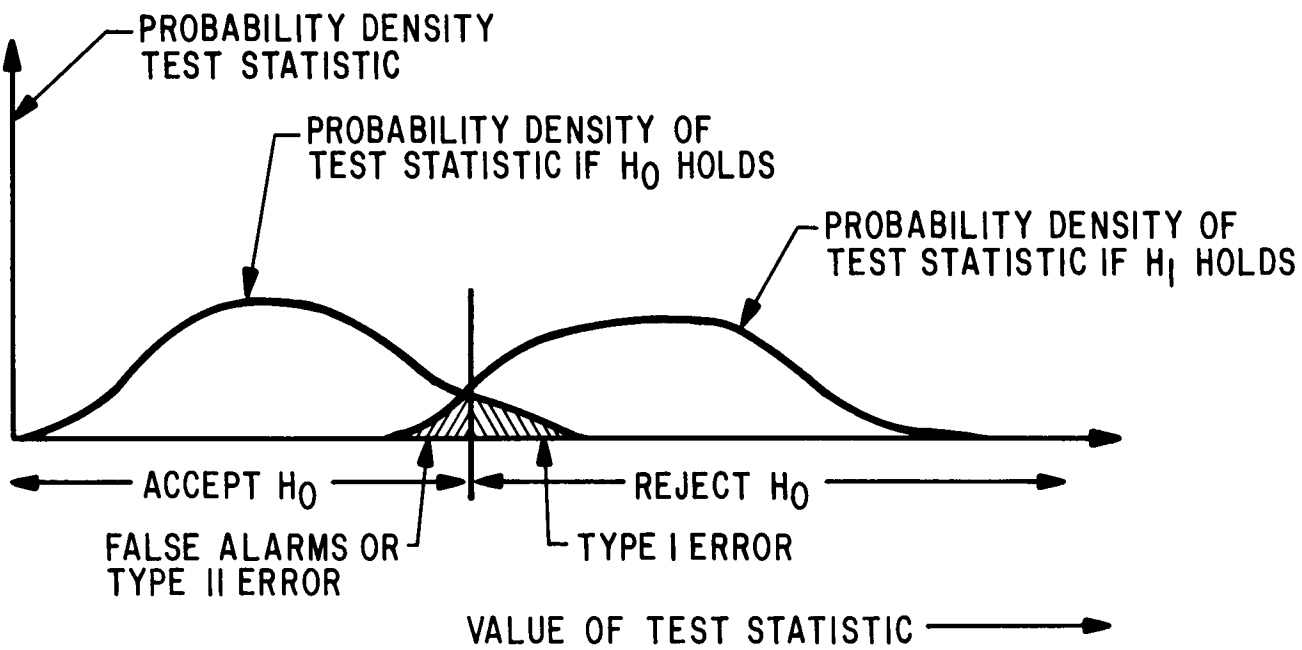


FIGURE 5-7. The most powerful test minimizes the false alarm rate. H_0 is the null hypothesis; H_1 the alternate.

Hogg (24) develops a scale-invariant, most powerful test for one member of the generalized Gaussian against another. He considers the family without a scale parameter

$$f(x,\alpha) = \frac{\alpha}{2\Gamma\left(\frac{1}{\alpha}\right)} e^{-|x|^{\alpha}} \tag{5-32}$$

and uses a procedure described by Hájek and Šidák (32) to derive the test. From the theory of rank statistics the most powerful test for a symmetrical distribution which is scale-invariant is found from the integral

$$I(X,\alpha) = \int_0^{\infty} \left[\prod_{i=1}^{n} f(\beta x_i,\alpha) \right] \beta^{n-1} \, d\beta \quad . \tag{5-33}$$

Substituting the density into the integral gives

$$I(X,\alpha) = \int_0^{\infty} \left[\frac{1}{2\Gamma\left(\frac{\alpha+1}{\alpha}\right)} \right]^{n} e^{-\left(\beta^{\alpha} \sum_{i=1}^{n} |x_i|^{\alpha} \right)} \beta^{n-1} \, d\beta \quad . \tag{5-34}$$

Using a change of variables

$$z = \beta^{\alpha} \sum_{i=1}^{n} |x_i|^{\alpha} \tag{5-35}$$

the integral is of the form

$$I(X,\alpha) = p(\alpha,n) \left(\frac{1}{n} \sum_{i=1}^{n} |x_i|^{\alpha} \right)^{-n/\alpha} \tag{5-36}$$

where

$$p(\alpha,n) = 2^{-n} \frac{\Gamma\left(\frac{n}{\alpha}\right)}{\left[\Gamma\left(\frac{1}{\alpha}\right)\right]^{n}} \alpha^{n-1} n^{-n/\alpha} \quad . \tag{5-37}$$

To determine if a random sample X is drawn from the distribution $f(X,\alpha_2)$ as opposed to $f(x,\alpha_1)$, the most powerful invariant test statistic is

$$V(X,\alpha_1,\alpha_2) = \frac{\left(\frac{1}{n}\sum_{i=1}^{n}|x_i|^{\alpha_1}\right)^{n/\alpha_1}}{\left(\frac{1}{n}\sum_{i=1}^{n}|x_i|^{\alpha_2}\right)^{n/\alpha_2}} \tag{5-38}$$

where the constants $p(\alpha_1,n)$ and $p(\alpha_2,n)$ are omitted because they are not dependent on the sample.

The larger V becomes, the more probable it is that the sample X is drawn from the distribution $f(x,\alpha_2)$ and not $f(x,\alpha_1)$.