Clutter modeling and analysis

Introduction

In the last session we saw how probabilistic models of targets and clutter could be used to calculate radar performance; obviously such calculations are only as good as the models on which they are based. So, by improving the clutter model, we should get a more realistic, and useful, performance calculation capability. This is the subject of today's session. Although we will pay particular attention to sea clutter the models we will discuss can be adapted to land clutter and other non-Gaussian noise processes. Some of the material covered is fairly close to the 'cutting edge' and should provide you with a point of entry into the research literature.

When a maritime scenario is interrogated by a radar system the returned signal invariably contains contributions scattered by the sea surface, as well as those from any target that may be present. This sea clutter may partially or wholly obscure the target signature and reduce its detectability quite dramatically. Much of this loss in detection can be recovered if we are able to identify the characteristic features of the clutter and exploit these in our signal processing. To do this we need a realistic and tractable model of the clutter. Attempts to characterise the microwave back scatter from the sea surface in terms of the structure and dynamics of that surface and an appropriate solution of Maxwell's equations describing the electromagnetic field would not be particularly useful in this context. The fundamental problems involved are so complicated that the likelihood of a realistic solution being achieved is very small; such a solution would also be far too detailed and complex to be of use in practical data processing. Suitably parameterised statistical models, that specify the probability of the clutter making a given contribution to the back scatter, are much more useful in these applications. Ideally such a model should be realistic, tractable and incorporate an underlying phenomenology that makes direct contact with our physical understanding of radar operation and clutter and target returns.

Gaussian clutter models (revisited)

Under many circumstances the complex Gaussian or speckle process provides just such a model. For this reason this has formed the basis of the conventional signal processing and clutter modelling described in standard texts. A radar return is represented by a complex signal, with in phase and quadrature components E_I and E_Q. When a low resolution radar system illuminates the ocean, there will be effectively random contributions to this return from the many independent scattering structures in the footprint. These will add up to give the resultant signal that corresponds to the radar return.

We start by looking at a simple discrete scatterer model for scattering that gives rise to the standard Gaussian or thermal noise. The scattered field is taken to be composed of contributions from discrete scatterers:

$$\mathbf{E} = \sum_{n=1}^{N} \mathbf{a}_n \tag{1}$$

(written like this it looks like a random walk; we should not be surprised by normal statistics)

What are the statistics of **E**? In general, for an arbitrary N it would be impossible to give a useful answer to this question without detailed knowledge of the **a** . When N gets large the central limit theorem tells us that the I Q components of the electric field will be normally distributed; we also know that these have zero means and the same variance and are uncorrelated. So what is the central limit theorem? Basically it states that the sum of a <u>large</u> number of <u>independent</u> random variables is normally distributed, as long as the distributions of the constituent random variables are well behaved. (As you may recall, a simple example of this last requirement is provided by the Cauchy distribution; the sum of any number of Cauchy random variables is itself Cauchy distributed. However the second moment of the Cauchy distribution is infinite while the existence of the '2+a bit' moment of a distribution is the condition it must satisfy if the Central Limit Theorem is to hold.)

How is this proved? Roughly as follows: we form the characteristic function of the distribution of the scattered electric field

$$\begin{aligned}\langle \exp(i\mathbf{U}.\mathbf{E}\rangle &= \left\langle \exp(i\sum_{n=1}^{N}\mathbf{U}.\mathbf{a}_n)\right\rangle \\ &= \left\langle \prod_{n=1}^{N}\exp(i\mathbf{U}.\mathbf{a}_n)\right\rangle \\ &= \prod_{n=1}^{N}\langle \exp(i\mathbf{U}.\mathbf{a}_n)\rangle \text{ (independence)} \\ &= \langle \exp(i\mathbf{U}.\mathbf{a})\rangle^N\end{aligned} \tag{2}$$

where we have assumed the **a** to have identical statistics (this just makes the algebra tidier). To ensure that the power in the scattered field remains finite we scale a with the square root of N, $a \to a/\sqrt{N}$ and expand up the characteristic function of the distribution of **a**

$$\langle \exp(i\mathbf{U}.\mathbf{a})\rangle = 1 + i\frac{\langle \mathbf{U}.\mathbf{a}\rangle}{\sqrt{N}} - \frac{\langle (\mathbf{U}.\mathbf{a})^2\rangle}{2N} \tag{3}$$

We expect the **a** to be distributed isotropically (i.e. with random phase) so the first term goes out; we let N get big and find that

$$\begin{aligned}\langle \exp(i\mathbf{U}.\mathbf{E}\rangle &= \langle \exp(i\mathbf{U}.\mathbf{a})\rangle^N \\ &= \left(1 - \frac{\langle (\mathbf{U}.\mathbf{a})^2\rangle}{2N}\right)^N\end{aligned} \tag{4}$$

which tends to $\exp\left(-U^2\langle a^2\rangle/4\right)$ as $N \to \infty$

This is the familiar Gaussian characteristic function of the normal distribution. Normal distribution results of this type occur all over the place in physics and engineering; Maxwell's distribution of velocities, the equipartition theorem, thermal noise anywhere where you get a lot of independent contributions to an extensive property of a system.

In this way we are led to the Gaussian speckle model

$$P(E_\mathrm{I}, E_\mathrm{Q}) = \frac{1}{\pi\langle I\rangle}\exp\left(-\left(E_\mathrm{I}^2 + E_\mathrm{Q}^2\right)/\langle I\rangle\right); \tag{5}$$

the corresponding pdfs of the envelope and intensity of the signal are

$$P(E) = \frac{2E}{\langle I \rangle} \exp\left(-E^2 / \langle I \rangle\right); \quad E = \sqrt{E_I^2 + E_Q^2}$$
$$P(I) = \frac{1}{\langle I \rangle} \exp\left(-I / \langle I \rangle\right); \quad I = E^2 \tag{6}$$

The probability that the envelope exceeds some threshold E_T, and so is a simple probability of false alarm, is given by

$$\mathrm{Prob}(E > E_T) = \int_{E_T}^{\infty} dE\, P(E) = \exp\left(-E_T^2 / \langle I \rangle\right). \tag{7}$$

$\langle I \rangle$ is the mean intensity of the clutter return, and is the parameter that specifies this model. This cannot be calculated directly and so must ultimately be derived from experimental data, from which empirical formulae relating $\langle I \rangle$ to prevailing conditions can be constructed. Let us assume that we have N independent measurements of intensity $\{I_k\}$; we could quite sensibly assign the value of their arithmetic mean to $\langle I \rangle$. In fact this is the 'best' we can do in the context of the Gaussian model. To see that this is so, and to introduce a technique that will be useful in the less intuitively accessible case of non-Gaussian clutter, we identify our best estimate of $\langle I \rangle$ as that which is most probable, given our model for the clutter pdf and the data $\{I_k\}$. On the basis of the Gaussian model the likelihood that the data take the values they do, for a given value of $\langle I \rangle$, is given by

$$P\left(\{I_k\} | \langle I \rangle\right) = \frac{1}{\langle I \rangle^N} \exp\left(-\frac{1}{\langle I \rangle} \sum_{k=1}^{N} I_k\right); \tag{8}$$

Bayes theorem tells us that this is proportional to the probability the data $\{I_k\}$ are described by a given value of $\langle I \rangle$. Thus to find the most probable value of $\langle I \rangle$, given, $\{I_k\}$ we must find that value which maximises (4). Thus we have

$$\frac{\partial \log\left(P\left(\langle I \rangle | \{I_k\}\right)\right)}{\partial \langle I \rangle} = -\frac{N}{\langle I \rangle} + \frac{1}{\langle I \rangle^2} \sum_{k=1}^{N} I_k = 0$$
$$\langle I \rangle_{\mathrm{ML}} = \frac{1}{N} \sum_{k=1}^{N} I_k \tag{9}$$

The subscript 'ML' denotes the 'maximum likelihood' estimate of $\langle I \rangle$. In this Gaussian case, where optimum estimation procedures are usually also the most straightforward, this approach may seem excessive; its full power and usefulness is only evident when it is applied to non-Gaussian noise models.

For a low-resolution radar system the speckle process provides a model for the single point statistics of the clutter. Many signal processing techniques exploit the correlation properties of the clutter and target returns. We will now show how the autocorrelation function of a signal is directly related to its power spectrum. Let us consider a complex signal

$$Z(t) = E_I(t) + iE_Q(t); \quad -T/2 < t < T/2$$
$$= 0; \text{ otherwise} \tag{10}$$

The Fourier transform of this signal is given by

$$\hat{Z}(\omega) = \int_{-\infty}^{\infty} dt \exp(i\omega t) Z(t), \tag{11}$$

from which we form the power spectrum as

$$S(\omega) = \frac{\left|\hat{Z}(\omega)\right|^2}{2\pi T} \tag{12}$$

whose interpretation as the power in a given frequency component is evident. The power spectrum can be written in terms of the I and Q components of the signal as follows

$$\begin{aligned} S(\omega) &= \frac{1}{2\pi T} \int_{-T/2}^{T/2} dt_1 \int_{-T/2}^{T/2} dt_2 \exp(i\omega(t_1 - t_2))(E_I(t_1) + iE_Q(t_1))(E_I(t_2) - iE_Q(t_2)) \\ &= \frac{1}{2\pi} \int_{-\infty}^{\infty} dt \exp(i\omega t)\left[\langle E_I(0)E_I(t)\rangle + \langle E_Q(0)E_Q(t)\rangle + i\left(\langle E_I(0)E_Q(t)\rangle - \langle E_Q(0)E_I(t)\rangle\right)\right] \end{aligned} \tag{13}$$

In the second step we have assumed that Z is a stationary process and that the correlation functions of its I and Q components decay rapidly on the timescale of measurement T. The statistical equivalence and independence of the I and Q components of the signal result in

$$\begin{aligned} \langle E_I(0)E_I(t)\rangle &= \langle E_Q(0)E_Q(t)\rangle = \frac{1}{2}\int_{-\infty}^{\infty} d\omega S(\omega)\cos(\omega t) \\ \langle E_Q(0)E_I(t)\rangle &= -\langle E_I(0)E_Q(t)\rangle = \frac{1}{2}\int_{-\infty}^{\infty} d\omega S(\omega)\sin(\omega t) \end{aligned} \tag{14}$$

while the integral of the power spectrum over all frequencies yields the total mean power or intensity in the signal

$$\int_{-\infty}^{\infty} d\omega\, S(\omega) = \left\langle E_Q{}^2\right\rangle + \left\langle E_I{}^2\right\rangle = \left\langle E^2\right\rangle = \langle I\rangle . \tag{15}$$

These relationships between the power spectrum and correlation functions of a signal, known as Weiner-Khintchine theorems, are not dependent on Z being a Gaussian process. However, the power spectrum is of particular relevance in the Gaussian special case as it determines all the higher order correlation functions of the signal, through the Gaussian factorisation property.

Non-Gaussian clutter

The complex Gaussian process is a reasonable model for sea clutter in a low resolution maritime radar system. In a high resolution radar system however, there may be only a few independent scattering structures within a range cell. In this case the arguments put forward in support of the Gaussian model (many contributions from independent scatterers, so the Central Limit theorem applies) can break down as the radar effectively resolves some of the larger scale structure of the sea surface. The clutter observed in this case is subject to much greater fluctuations (sea spikes) than are seen in Gaussian clutter. Thus the improvement in signal to noise ratio and potential detection achieved by increasing the resolution of the system can be more than offset by the clutter becoming effectively target-like and regularly exceeding thresholds based on a Gaussian model of the clutter. Realistic non-Gaussian clutter models are needed if the performance lost to the sea spikes is to be regained. To test a non-Gaussian clutter model against available data we focus on quantities, the normalised moments of intensity data, that highlight the 'spiky' non-Gaussian behaviour that is most evident in the 'tails' of the pdf. In the Gaussian limit we have

$$\frac{\langle I^n \rangle}{\langle I \rangle^n} = n!\,; \qquad (16)$$

values of normalised intensity moments greater than these are characteristic of spiky data like high resolution sea clutter.

The structure of the sea surface is very complex and is characterised by many length scales, ranging from 1 cm or less, typical of foams and ripples, to many metres, typical of swell structures. Similarly the time-scales characteristic of sea surface motions range from less than a second to many seconds, if not minutes. Consequently there will be many effectively independent small-scale structures within a high-resolution range cell that give rise to a speckle-like clutter. This decorrelates over a short period of time in which the small scale structures move through a distance of the order of half a radar wavelength . However these small-scale structures will be modulated by the more slowly changing large scale structures; this results in a changing 'local power' of the Rayleigh-like returns from the small scale structure. This decomposition of the non-Gaussian clutter into a rapidly decorrelating, locally Rayleigh process, whose local power is modulated by a much more slowly varying (but nonetheless random) process, forms the basis the K distribution and other compound models for clutter.

The compound representation of NG clutter

The physical arguments of the previous section can be put into a simple mathematical framework as follows. The pdf of the envelope of the locally Rayleigh process, with a local power x can be written as

$$P(E|x) = \frac{2E}{x}\exp\left(-E^2/x\right). \qquad (17)$$

This local power is itself a random variable, whose values have a pdf $P_c(x)$. Thus the full pdf of the envelope of the signal is obtained by integrating the locally Rayleigh pdf (17) with power x over $P_c(x)$

$$\begin{aligned} P(E) &= \int_0^\infty dx\, P(E|x) P_c(x) \\ &= 2E\int_0^\infty \frac{dx}{x}\exp\left(-E^2/x\right) P_c(x) \end{aligned}. \qquad (18)$$

This is the mathematical expression of the compound form of the pdf of non-Gaussian clutter. We have already identified normalised intensity moments as a useful diagnostic for non-Gaussian behaviour. Within this compound model we find that

$$\left\langle I^n \right\rangle = \int_0^\infty dx P_c(x) \left\{ 2\int_0^\infty \frac{dE}{x} E^{2n+1} \exp\left(-E^2/x\right) \right\}$$
$$= n! \int_0^\infty dx P_c(x) x^n \qquad (19)$$
$$= n! \left\langle x^n \right\rangle_c$$

so that

$$\frac{\left\langle I^n \right\rangle}{\left\langle I \right\rangle^n} = n! \frac{\left\langle x^n \right\rangle_c}{\left\langle x \right\rangle_c^n} . \qquad (20)$$

Similarly the probability that the signal envelope exceeds a threshold E_{T} is given by

$$\mathrm{Prob}\left(E > E_{\mathrm{T}}\right) = \int_0^\infty dx \exp\left(-E_{\mathrm{T}}^2/x\right) P_c(x) . \qquad (21)$$

These results show how straightforward the generalisation of standard Gaussian-derived results is within the compound representation of non-Gaussian clutter. One simply takes the required Gaussian based result for a given local speckle power, then integrates this over the appropriate distribution of local power. This remark applies equally well to modelling target plus clutter returns and evaluating probabilities of detection. However, before any real progress can be made, we must identify the local power pdf $P_c(x)$.

The gamma distribution of local power and the K distribution

As we have noted the local speckle process decorrelates in a time characteristic of the motion of small scale structure on the sea surface through a distance of half a radar wavelength; typically this is of the order of several milliseconds. Immediate decorrelation can be effected by the use of a frequency agile waveform; frequency agility does not, however, decorrelate the more slowly varying background modulation due to large scale structure. Thus we are able to obtain many independent samples of the local power in a time in which it does not change appreciably. These give an estimate of the local power *x*. By analysing a sufficiently large quantity of data a large set of independent measurements of *x* can be built up. These can be used to identify a good model for $P_c(x)$. It has been found that the gamma distribution provides the best fit to most of the available data i.e.

$$P_c(x) = \frac{b^\nu}{\Gamma(\nu)} x^{\nu-1} \exp(-bx) . \qquad (22)$$

Other choices for $P_c(x)$ could be made, and would provide other potentially useful clutter models that would retain many of the attractive features of the K distribution. However, it is unlikely that they would result in distributions that are as well characterised in terms of tabulated functions, nor that they would have the property of infinite divisibility possessed by the K distribution.

When (22) is substituted into (18) we find that the pdf of the clutter envelope is given by

$$
\begin{aligned}
P(E) &= \frac{2Eb^{\nu}}{\Gamma(\nu)}\int_0^{\infty} dx x^{\nu-2}\exp\left(-bx - E^2/x\right) \\
&= \frac{4b^{(\nu+1)/2}E^{\nu}}{\Gamma(\nu)} K_{\nu-1}\left(2E\sqrt{b}\right)
\end{aligned} \tag{23}
$$

We see that this integral can be evaluated in terms of the modified Bessel or K function that gives its name to the model. Fortunately, no knowledge of the properties of these Bessel functions is required if we are to evaluate quantities of interest, such as probabilities of false alarm or intensity moments. In each case we merely take the Gaussian result and integrate it over the gamma distribution of x . In this way we find that

$$
\begin{aligned}
\mathrm{Prob}\left(E > E_{\mathrm{T}}\right) &= \frac{b^{\nu}}{\Gamma(\nu)}\int_0^{\infty} dx \exp\left(-E_{\mathrm{T}}^2/x\right) x^{\nu-1}\exp(-bx) \\
&= \frac{2b^{\nu/2}}{\Gamma(\nu)} E_{\mathrm{T}}^{\nu} K_{\nu}\left(2E_{\mathrm{T}}\sqrt{b}\right)
\end{aligned} \tag{24}
$$

and

$$
\frac{\left\langle I^n \right\rangle}{\left\langle I \right\rangle^n} = n!\frac{\Gamma(n+\nu)}{\Gamma(\nu)\nu^n} = n!\prod_{k=1}^{n-1}\left(1 + k/\nu\right). \tag{25}
$$

Fluctuating populations of scatterers and NG statistics

In our earlier ‘derivation’ of the central limit theorem we considered a large fixed number of scatterers (i.e. steps in our random walk). Let us now consider the case where the number of scatterers/ steps in the walk fluctuates. The simplest such model for this fluctuation is the Poisson or shot-noise distribution for which the probability of there being N scatterers present is given by

$$
P(N) = \exp(-\overline{N})\frac{(\overline{N})^N}{N!} \tag{26}
$$

The characteristic function of the distribution of the electric field distribution is found by averaging the fixed scatterer number (N) result over this distribution of N, having scaled a with the square root of the *mean* number of scatterers in the population:

$$\begin{aligned}\left\langle \exp(i\mathbf{U}.\mathbf{E}\right\rangle &= \sum_{N=0}^{\infty} P(N)\left\langle \exp(i\mathbf{U}.\mathbf{a})\right\rangle^{N} \\ &= \sum_{N=0}^{\infty} P(N)\left(1-\frac{\left\langle (\mathbf{U}.\mathbf{a})^2\right\rangle}{2\overline{N}}\right)^{N} \\ &= \exp(-\overline{N})\sum_{N=0}^{\infty}\frac{\overline{N}^{N}}{N!}\left(1-\frac{\left\langle (\mathbf{U}.\mathbf{a})^2\right\rangle}{2\overline{N}}\right)^{N}\end{aligned} \tag{27}$$

which tends to $\exp\left(-U^2\left\langle a^2\right\rangle/4\right)$ as $\overline{N}\to\infty$

Thus merely having a Poisson fluctuating population of scatterers does not affect the statistics of the scattered field. Poisson statistics characterise the number of events occurring randomly (i.e. without 'bunching') in a given interval; introducing some bunching into these events (i.e. the occurrence of a scatterer in the illuminated volume) might cause the CLT to break down and give rise to non-Gaussian statistics. Consider a population subject to the un-correlated processes of birth, death and immigration with the rates of birth and death being proportional to the population number and rate of immigration being constant, but only from regions of lower to higher population. The equation of balance for the population number is then given by

$$\frac{dP(N,t)}{dt} = \mu(N+1)P(N+1,t) - [(\lambda+\mu)N+\nu]P(N,t) + [\lambda(N-1)+\nu]P(N-1,t) \tag{28}$$

λ, μ and ν characterise the rates of birth death and migration respectively. (Economists, ecologists and others use models of this type for populations of living creatures and ascribe various interpretations to the constants in the rate equation. An identical equation occurs in the description of photon statistics.) The equilibrium population distribution can be determined by setting this time derivative to zero, and introducing a 'characteristic function' appropriate to a discrete distribution:

$$\begin{aligned}&\mu(N+1)P(N+1) - [(\lambda+\mu)N+\nu]P(N) + [\lambda(N-1)+\nu]P(N-1) = 0 \\ &C(s) = \sum_{N=0}^{\infty} P(N)s^N\end{aligned} \tag{29}$$

C(s) can be shown to satisfy a simple differential equation

$$\frac{dC}{ds}(\mu-(\lambda+\mu)s+\lambda s^2) = (1-s)\nu C \tag{30}$$

whose solution, that satisfies the normalisation condition $C(1)=1$, is

$$C(s) = \frac{(1-\lambda/\mu)^{\alpha}}{(1-s\lambda/\mu)^{\alpha}} \text{ where } \alpha = \nu/\lambda\,. \tag{31}$$

Direct expansion of this result gives us

$$P(N) = (1-\lambda/\mu)^{\alpha}\left(\frac{\lambda}{\mu}\right)^{N}\frac{\Gamma(N+\alpha)}{N!\,\Gamma(\alpha)}; \qquad (32)$$

it follows that the mean and variance of the population number are given by

$$\langle N\rangle = \overline{N} = \frac{\nu}{\mu-\lambda} \text{ and } \frac{\langle N^2\rangle}{\langle N\rangle^2} - 1 = \frac{1}{\alpha} + \frac{1}{\overline{N}}. \qquad (33)$$

(You should check all this out as an exercise; it's quite straightforward.)

This equilibrium population distribution is called the negative binomial distribution. The non-vanishing of the normalised variance of population in the limit of a large mean population is characteristic of bunching, or correlation, between its members.

Now let's work out the characteristic function of the distribution of electric field scattered from a negative binomially distributed population of scatterers, scaling *a* and letting the mean number of scatterers get large:

$$\begin{aligned}\langle \exp(i\mathbf{U}.\mathbf{E}\rangle &= \sum_{N=0}^{\infty} P(N)\langle \exp(i\mathbf{U}.\mathbf{a})\rangle^{N} \\ &= \sum_{N=0}^{\infty} P(N)\left(1-\frac{\langle(\mathbf{U}.\mathbf{a})^2\rangle}{2\overline{N}}\right)^{N} \\ &= \frac{1}{\left(1+\overline{N}/\alpha\right)^{\alpha}}\sum_{N=0}^{\infty}\frac{\Gamma(N+\alpha)}{\Gamma(\alpha)N!}\left(\frac{\overline{N}/\alpha}{1+\overline{N}/\alpha}\right)^{N}\left(1-\frac{\langle(\mathbf{U}.\mathbf{a})^2\rangle}{2\overline{N}}\right)^{N}\end{aligned} \qquad (34)$$

which tends to $\dfrac{1}{\left[1+U^2\langle a^2\rangle/4\alpha\right]^{\alpha}}$ as $\overline{N}\to\infty$

Thus we see that the characteristic function does not take the Gaussian form in this limit. Fourier inversion of this result gives us the familiar K distribution. Show for yourself that

$$P(E) = \frac{E^{\alpha}\beta^{\alpha+1}K_{\alpha-1}(\beta E)}{2^{\alpha-1}\Gamma(\alpha)};\ \beta = \sqrt{\frac{4\alpha}{\langle a^2\rangle}} \qquad (35)$$

We note that this distribution is infinitely divisible; this reflects the infinite divisibility of the gamma distribution revealed in the compound form of the K distribution.

Modeling the power spectrum of K distributed clutter

We have seen how the compound model, and in particular the K distribution, is able to describe the single point statistics of a non-Gaussian process. This compound K representation can be extended to the modelling of power spectra. The first requirement of the model is that the total power in the spectrum is gamma distributed. Thus we expect

$$\int_{-\infty}^{\infty} d\omega S(\omega) = x \qquad (36)$$

to be gamma distributed i.e. that the pdf of x takes the form (22). As an illustration we take the simple Gaussian spectrum with unit power

$$\hat{S}(\omega,\sigma,\omega_0) = \frac{1}{\sqrt{2\pi}\sigma}\exp\left(-(\omega-\omega_0)^2/2\sigma^2\right) \tag{37}$$

as our fundamental building block. We then form the clutter power spectrum as

$$S(\omega) = x\hat{S}(\omega,\sigma,\omega_0);$$

This will now satisfy (36) automatically. By letting the parameters σ and ω_0 depend on x, either deterministically or stochastically, we can model the statistics of the returns in given Doppler bins (i.e. the value taken by the power spectrum for given values of the frequency) in a variety of ways. Thus we can write the n[th] moment of the power at a given frequency as

$$\left\langle S(\omega)^n \right\rangle = \frac{1}{(2\pi)^{n/2}}\int dx \int d\omega_0 \int d\sigma P_c(x) P\left(\omega_0,\sigma|x\right) x^n \frac{\exp\left(-n(\omega-\omega_0)^2/2\sigma^2\right)}{\sigma^n} \tag{38}$$

The conditional probability $P\left(\omega_0,\sigma|x\right)$ accommodates a wide variety of models. Rather than consider the general case further we will consider several special cases, discussing the extent to which they make contact with experimental data and models already discussed in the literature. Much of the published analysis of Doppler spectra of clutter presents the data in terms of an effective ν parameter that is in general frequency dependent and is defined by

$$\frac{1}{\nu_{\text{eff}}} = \frac{\left\langle S(\omega)^2 \right\rangle}{\left\langle S(\omega) \right\rangle^2} - 1 \tag{39}$$

ν_{eff} is commonly observed to decrease (i.e. the frequency component in the power spectrum becomes more 'spiky' or non-Gaussian) as the Doppler frequency increases. It is this feature that we particularly wish to reproduce in our model.

The simplest model merely represents the power spectrum as the product of a gamma variate and $\hat{S}$, with σ and ω_0 taking fixed values. While this model is straightforward to analyse and simulate it necessarily implies that

$$\frac{\left\langle S(\omega)^2 \right\rangle}{\left\langle S(\omega) \right\rangle^2} = \frac{\left\langle x^n \right\rangle_c}{\left\langle x \right\rangle_c^n} \tag{40}$$

and so cannot reproduce the frequency dependence of the effective shape parameter.

Introduction of a dependence of the spectral width on the local power x of the clutter is physically reasonable (a breaking wave feature way produces a larger cross section and a greater spread of velocities/Doppler frequencies) The simplest way to incorporate this into our model is through a deterministic dependence i.e. $\sigma = \sigma(x)$ or $P\left(\sigma|x\right) = \delta\left(\sigma - \sigma(x)\right)$ so that

$$\left\langle S(\omega)^n \right\rangle = \frac{1}{(2\pi)^{n/2}}\int dx P_c(x) x^n \frac{\exp\left(-n\omega^2/2\sigma(x)^2\right)}{\sigma(x)^n} \tag{41}$$

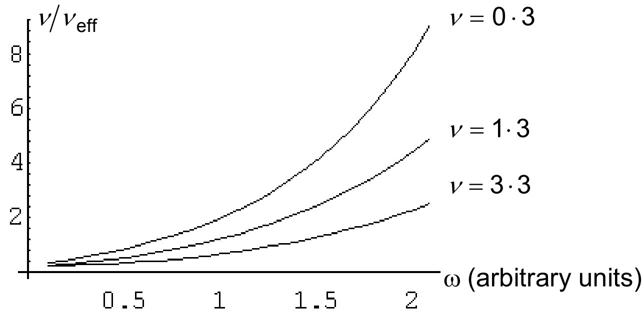
We now make the choice $\sigma(x) = \sqrt{2x}$, primarily to facilitate the analysis; nonetheless the monotonic growth of σ with *x* captured by this model is sensible. Thus we have

$$\begin{aligned}\left\langle S(\omega)^n \right\rangle &= \frac{1}{\pi^{n/2}} \frac{b^\nu}{\Gamma(\nu)} \int dx \exp(-bx) x^{\nu+n/2-1} \exp\left(-n\omega^2/x\right) \\ &= \frac{2}{\pi^{n/2}} \frac{b^{\nu/2-n/4}}{\Gamma(\nu)} \omega^{\nu+n/2} n^{(\nu/2+n/4)} K_{\nu+n/2}(2\omega\sqrt{bn})\end{aligned} \tag{42}$$

so that the effective shape parameter can be calculated from

$$\frac{\nu}{\nu_{\mathrm{eff}}} = \nu\left(\frac{\Gamma(\omega)}{\omega^\nu} \frac{2^{\frac{\nu-1}{2}}}{b^{\nu/2}} \frac{K_{\nu+1}\left(2\omega\sqrt{2b}\right)}{K_{\nu+1/2}\left(2\omega\sqrt{b}\right)^2} - 1\right); \tag{43}$$

as usual *K* is the modified Bessel function. The following plot of the effective shape parameter as a function of frequency shows that this simple model is able to reproduce qualitatively the behaviour seen in coherent clutter data.



The application of the compound representation to the modelling of coherent clutter is much less thoroughly developed than is its application to single point statistics and is the subject of an on-going research effort. We shall see how it is exploited in simulation studies in the next session

Other NG models for clutter: Class A, BAM and Weibull

At this point we might mention some other models of sea clutter, that exploit the compound representation, but introduce rather different physical mechanisms for the non-Gaussian statistics. The more widely used of these models is the Middleton Class A model. The Class A model gives the pdf of the intensity of the combined noise processes, normalised to have a unit mean, as

$$P(z) = \exp[-A] \sum_{m=0}^{\infty} \frac{A^m}{m!} \frac{\exp[-z/\langle I_m \rangle]}{\langle I_m \rangle} \tag{44}$$

where

$$\langle I_m \rangle = \frac{m/A + \Gamma}{1+\Gamma},$$

with Γ representing the ratio of the powers in the Gaussian and non-Gaussian processes. This pdf consists of a set of exponential distributions with powers characterised by the integers *m;* their contributions have been weighted by a Poisson distribution of these integers (i.e. numbers of scatterers).

To focus on the non-Gaussian component of the mode, we set $\Gamma = 0$. The moments of *z* are then given by

$$\begin{aligned}\langle z^n \rangle &= n!\,e^{-A}\sum_{m=0}^{\infty}\frac{A^m}{m!}\left(\frac{m}{A}\right)^n \\ &= n!\frac{e^{-A}}{A^n}\left(A\frac{d}{dA}\right)^n e^A \\ &= n!\,f_n(A)\end{aligned} \tag{45}$$

The polynomial in inverse powers of *A*, $f_n(A)$, satisfies the simple recurrence relation

$$f_{n+1}(A) = \left(1+\frac{n}{A}\right)f_n(A) + \frac{\partial}{\partial A}f_n(A)\ ; \tag{46}$$

this can be 'pushed forward', with the help of *'Mathematica'* if necessary, to give the following results

n	$A^{n-1}f_n(A)$
1	1
2	$A+1$
3	A^2+3A+1
4	A^3+6A^2+7A+1
5	$A^4+10A^3+25A^2+15A+1$
6	$A^5+15A^4+65A^3+90A^2+31A+1$

Higher order moments can be generated as required. Compare these results with those obtained from the gamma distribution:

n	$\nu^n\,\Gamma(\nu+n)/(\Gamma(\nu)$
1	1
2	$1+\nu$
3	$\nu^2+3\nu+2$
4	$\nu^3+6\nu^2+11\nu+6$
5	$\nu^4+10\nu^3+35\nu^2+50\nu+24$
6	$\nu^5+15\nu^4+85\nu^3+225\nu^2+274\nu+120$

The two sets of moments have the same Gaussian limiting forms as A (Class A) or ν (K distribution) tends to infinity. The first order corrections to the Gaussian limits (expanded in inverse powers of A or ν) are identical if A and ν are identified one with the other. A special case of this is the corresponding identity of the second normalised moments of the two distributions; this identity is not found in the higher order moments. As A (or ν) tends to zero the model pdfs describe increasingly non-Gaussian statistics. For the Class A model we see that:

$$\frac{\langle z^n\rangle}{\langle z\rangle^n} \sim n!\frac{1}{A^{n-1}} \tag{47}$$

while for the K distribution

$$\frac{\langle z^n\rangle}{\langle z\rangle^n} \sim \frac{n!(n-1)!}{\nu^{n-1}}. \tag{48}$$

The Class A model pdf

$$\begin{aligned} P(z) &= Ae^{-A}\sum_{m=0}^{\infty}\frac{A^m}{m!}\frac{\exp\left[-\frac{Az}{m}\right]}{m} \\ &= 2e^{-A}\delta(z) + Ae^{-A}\sum_{m=1}^{\infty}\frac{A^m}{m!}\frac{\exp\left[-\frac{Az}{m}\right]}{m} \end{aligned} \tag{49}$$

can be interpreted as the intensity distribution derived from a population of scatterers whose number has a Poisson distribution. We have already seen that the K distribution describes the intensity distribution obtained by coherently illuminating a negative binomially distributed population of scatterers, if the limit of a large mean number of scatterers is taken appropriately. The Poisson distribution is a limit of the negative binomial distribution. Thus a negative binomial population of scattterers can provide us with a more general model, from which the Class A and K distributions emerge as special limiting cases.

The probability of the number of scatterers, in a negative binomially distributed population, being N is given by

$$P(N|\overline{N},\alpha) = \frac{\Gamma(N+\alpha)}{N!\,\Gamma(\alpha)}\frac{\left(\overline{N}/\alpha\right)^N}{\left(1+\overline{N}/\alpha\right)^{N+\alpha}}. \tag{50}$$

Let the parameter α tend to infinity:

$$\begin{gathered} \frac{\Gamma(N+\alpha)}{\Gamma(\alpha)} \to \alpha^N, \quad \frac{1}{\left(1+\overline{N}/\alpha\right)^{N+\alpha}} \to \exp(-\overline{N}) \\ \text{so that} \\ P(N|\overline{N},\alpha) \to \frac{\exp(-\overline{N})\,\overline{N}^N}{N!} \end{gathered} \tag{51}$$

i.e. we have the Poisson distribution. Letting $\overline{N}$ tend to infinity, while keeping α constant , gives us the gamma distibution:

$$N = \overline{N}x;\text{define } P(x) = \overline{N}P(\overline{N}x \mid \overline{N},\alpha)\text{ then } P(x) \to \frac{\alpha^{\alpha}}{\Gamma(\alpha)}\exp(-\alpha x)x^{\alpha-1}\text{ as } \overline{N} \to \infty \qquad (52)$$

Thus the non-Gaussian characters that are manifest in the Class A and K distribution models arise from different causes: the finite number of 'unbunched' scatterers in Class A and the bunched (correlated) but very numerous scatterers in the K model. We have demonstrated the connection between the two models and their essentially complementary nature.

The so-called breaking area model (BAM) identifies the source of the local power in the compound model with breaking wave events; these are assumed to occur when the vertical acceleration of the sea surface, which is itself modeled as a Gaussian process, exceeds a threshold. It can be shown that, in the large threshold limit, the number n of excursions above that threshold in the area A is Poisson distributed

$$P(n) = \frac{\langle N \rangle^{n}}{n!}\exp(-\langle N \rangle)$$

The area s over which the excursion above the threshold u takes place has a negative exponential distribution:

$$P(s) = \gamma u^{2}\exp(-\gamma u^{2}s) \qquad (53)$$

Thus if there are n excursions above the threshold in the area A, the pdf of their total area is given by

$$P(s,n) = \frac{(\gamma u^{2})^{n}s^{n-1}}{(n-1)!}\exp(-\gamma u^{2}s)\,. \qquad (54)$$

It should be stressed that these results only hold in the limit where the threshold is large.

These simple results can be brought together to give the marginal pdf of the breaking area within A; the breaking area per unit area of ocean surface is identified with the local power *x.* The formula manipulation which leads to the BAM pdf (of x) is similar to that which produces the Middleton Class A model; in each case a basic pdf, associated in some way with n events, is averaged over a Poisson distribution of n. In the case of the BAM, this sum can be identified with a known special function, which is quite fortunate. Thus we have:

$$\begin{aligned} P(s) &= \sum_{n} P(s,n)P(n) \\ &= \exp(-\langle N \rangle)\sum_{n} P(s,n)\frac{\langle N \rangle^{n}}{n!} \end{aligned} \qquad (55)$$

For n=0, $P(s,n)$ is tentatively identified with the delta function $2\delta(s)$, just as in the case of the Class A model. Thus we write

$$P(s) = 2\exp\left(-\langle N\rangle\right)\delta(s) + \exp\left(-\langle N\rangle - \gamma u^2 s\right)\sum_{n=1}^{\infty}\frac{\langle N\rangle^n}{n!}\frac{\left(\gamma u^2\right)^n s^{n-1}}{(n-1)!}$$

$$= 2\exp\left(-\langle N\rangle\right)\delta(s) + \langle N\rangle\gamma u^2\exp\left(-\langle N\rangle - \gamma u^2 s\right)\sum_{n=0}^{\infty}\frac{\langle N\rangle^n}{n!}\frac{\left(\gamma u^2\right)^n s^n}{(n+1)!} \qquad (56)$$

The infinite series occurring here can be related to the expansion of the modified Bessel function of the first kind and first order

$$I_1(z) = \frac{z}{2}\sum_{n=0}^{\infty}\frac{\left(z^2/2\right)^n}{n!(n+1)!}$$

Thus we find that

$$P(s) = 2\exp\left(-\langle N\rangle\right)\delta(s) + \exp\left(-\langle N\rangle - \gamma u^2 s\right)\sqrt{\frac{\langle N\rangle\gamma u^2}{s}}I_1\left(2\sqrt{\langle N\rangle\gamma u^2 s}\right) \qquad (57)$$

If we set

$$x = \frac{s}{A}$$

then we can show quite easily that

$$\langle x\rangle = \frac{\langle N\rangle}{A\gamma u^2} = 2\frac{\upsilon}{A\gamma u^2}\,. \qquad (58)$$

If we now substitute these results into the pdf (57) we find that

$$P(x) = 2\upsilon\frac{1}{\sqrt{x\langle x\rangle}}\exp\left(-2\upsilon\left(1 + x/\langle x\rangle\right)\right)I_1\left(4\upsilon\sqrt{\frac{x}{\langle x\rangle}}\right) + 2\exp(-2\upsilon)\delta(x) \qquad (59)$$

This model of the variation of local power can now be fed into a compound model of the clutter statistics. (This breaking area approach has the advantage that the parameters occurring in the model can be related directly to the properties of the Gaussian field modeling the sea surface acceleration, which can in turn be related to environmental conditions.) Perhaps you would like to make a comparison of the normalised moments of the local power implicit in this model with those we derived for the gamma and Class A models.

It is also possible to model the single point statistics of non-Gaussian processes without invoking the physically appealing and flexible compound representation. Any function can be pressed into service as a pdf of envelope or intensity, as long as it does not take negative values and is integrable. Thus in the case of the Weibull model the non-Gaussian process y is assigned the pdf

$$\begin{aligned} P(y) &= \gamma\frac{y^{\gamma-1}}{a^{\gamma}}\exp\left(-(y/a)^{\gamma}\right); \quad y \ge 0 \\ &= 0 \quad \text{otherwise} \end{aligned} \qquad (60)$$

This model includes the exponential and Rayleigh models as special cases (γ=1,2); the normalised moments of the distribution are useful in assessing its applicability to data and are given by

$$\frac{\left\langle y^n \right\rangle_W}{\left\langle y \right\rangle_W^n} = \frac{\Gamma\left(1+n/\gamma\right)}{\Gamma\left(1+1/\gamma\right)^n}. \tag{61}$$

It should be stressed, however, that the Weibull distribution does not incorporate the physically reasonable picture of a rapidly decorrelating speckle process, modulated over a much longer time-scale by a second stochastic process representing the effects of the large-scale structure of the sea surface. Consequently it cannot be extended straightforwardly to include contributions to the radar signal from steady or fluctuating target returns nor be used in the analysis of pulse to pulse integration.

Problems arising in the analysis of clutter measurements

So far we have reviewed the statistical modelling of clutter in general terms. We now turn to a couple of problems that relate directly to the analysis of data. These are the effects of added thermal noise and the correction of measured moments to compensate for its presence and the maximum likelihood estimation of parameters in non-Gaussian statistical models.

Contamination by thermal noise

Let us first consider the effects of thermal noise on measurements of moments of the clutter distribution. Our complex signal **S** now consists of the clutter contribution **Z** and a thermal noise component **n** (for convenience we represent the I and Q components in vector notation).

$$\mathbf{S} = \mathbf{Z} + \mathbf{n}$$

The characteristic function (or Fourier transform of the pdf) is very useful in the analysis of moments. If in the absence of thermal noise we form

$$C_{\mathrm{C}}(\mathbf{k}) = \left\langle \exp(i\mathbf{k}.\mathbf{Z}) \right\rangle_{\mathrm{C}} \tag{62}$$

As the phase of the signal **Z** is uniformly distributed the characteristic function can be written as

$$C_{\mathrm{C}}(\mathbf{k}) = \left\langle J_0(kZ) \right\rangle_{\mathrm{C}} = \sum_{r=0}^{\infty} \frac{\left(-k^2/4\right)^r}{r!r!} \left\langle Z^{2r} \right\rangle_{\mathrm{C}} = \sum_{r=0}^{\infty} \frac{\left(-k^2/4\right)^r}{r!r!} \left\langle I^r \right\rangle_{\mathrm{C}} \tag{63}$$

This can be thought of as a generating function for the intensity moments of the clutter. (The subscript C denotes an average over the clutter pdf)If we form the characteristic function when noise is present we obtain

$$C_{\mathrm{C+N}}(\mathbf{k}) = \left\langle \exp\left(i\mathbf{k}.\mathbf{S}\right) \right\rangle_{\mathrm{C+N}} = \left\langle \exp\left(i\mathbf{k}.(\mathbf{Z}+\mathbf{n})\right) \right\rangle_{\mathrm{C+N}} \tag{64}$$

As the clutter and noise processes are independent we can rewrite this as

$$\left\langle \exp\left(i\mathbf{k}.\mathbf{S}\right) \right\rangle_{\mathrm{C+N}} = \left\langle \exp\left(i\mathbf{k}.\mathbf{Z}\right) \right\rangle_{\mathrm{C}} \left\langle \exp\left(i\mathbf{k}.\mathbf{n}\right) \right\rangle_{\mathrm{N}}. \tag{65}$$

Here the subscripts N and C+N denote averaged over the pdfs for the noise and noise plus clutter signals. If we once again assume that all phases are uniformly distributed we find that

$$\sum_{r=0}^{\infty}\frac{\left(-k^2/4\right)^r}{r!r!}\left\langle S^{2r}\right\rangle_{\mathrm{C+N}} = \left\langle \exp\left(i\mathbf{k}.\mathbf{n}\right)\right\rangle_{\mathrm{N}}\sum_{p=0}^{\infty}\frac{\left(-k^2/4\right)^p}{p!p!}\left\langle I^r\right\rangle_{\mathrm{C}} \tag{66}$$

For thermal noise we have

$$P(\mathbf{n}) = \frac{1}{2\pi\left\langle n^2\right\rangle}\exp\left(-n^2/\left\langle n^2\right\rangle\right) \tag{67}$$

and

$$\left\langle \exp\left(i\mathbf{k}.\mathbf{n}\right)\right\rangle_{\mathrm{N}} = \exp\left(-k^2\left\langle n^2\right\rangle/4\right) \tag{68}$$

Therefore we can write

$$\sum_{q=0}^{\infty}\frac{\left(k^2/4\right)^q}{q!}\left\langle n^2\right\rangle^q\sum_{r=0}^{\infty}\frac{\left(-k^2/4\right)^r}{r!r!}\left\langle S^{2r}\right\rangle_{\mathrm{C+N}} = \sum_{p=0}^{\infty}\frac{\left(-k^2/4\right)^p}{p!p!}\left\langle I^p\right\rangle_{\mathrm{C}} \tag{69}$$

Simply by equating coefficients of powers of $k^2/4$ we can now express the moments of the clutter intensity in terms of the measured intensity moments of clutter plus noise and the (assumed known) corrupting noise power. Thus we have

$$\left\langle I^s\right\rangle_{\mathrm{C}} = \sum_{q=0}^{s}(-1)^q\frac{\left\langle n^2\right\rangle^q\left\langle S^{2(s-q)}\right\rangle_{\mathrm{C+N}}(s!)^2}{q!\left((s-q)!\right)^2}\,. \tag{70}$$

This formula applies to any clutter distribution, subject only to the assumption of uniformly distributed phase. In the case of K distributed clutter we can write the pdf of the intensity of the clutter plus noise signal in the attractive compound form (due to Watts)

$$P(I) = \frac{b^\nu}{\Gamma(\nu)}\int_0^\infty dx\, x^{\nu-1}\frac{\exp\left(-I/\left(x+\left\langle n^2\right\rangle\right)\right)}{\left(x+\left\langle n^2\right\rangle\right)} \tag{71}$$

Maximum likelihood estimation procedures for NG clutter

In the previous session we saw how the maximum likelihood technique could be applied to estimate the mean intensity parameter in the speckle model. This method can also be applied to the analysis of non-Gaussian data. We will demonstrate how this is done for the gamma and Weibull distributions. Consider first a set of independent data $\left\{x_j\right\}$ which we assume to be gamma distributed and to have the likelihood

$$P\left(\left\{x_j\right\}\middle|b,\nu\right) = \frac{b^{N\nu}}{\Gamma(\nu)^N}\left(\prod_{j=1}^{N}x_j^{\nu-1}\right)\exp\left(-b\sum_{j=1}^{N}x_j\right). \tag{72}$$

Arguing just as in the Gaussian case we identify the most probable, and in some sense, best, values of b, ν as those that satisfy the maximum likelihood conditions

$$\frac{\partial \log\left(P\left(b,\nu\middle|\{x_j\}\right)\right)}{\partial b} = \frac{N\nu}{b} - \sum_{j=1}^{N} x_j = 0$$
$$\frac{\partial \log\left(P\left(b,\nu\middle|\{x_j\}\right)\right)}{\partial \nu} = \sum_{j=1}^{N} \log(x_j) - N\psi(\nu) = 0 \qquad (73)$$

Here $\psi(\nu)$ is the logarithmic derivative of the gamma function,

$$\psi(\nu) = \frac{d\log(\Gamma(\nu))}{dz} = \frac{1}{\Gamma(\nu)}\frac{d\Gamma(\nu)}{d\nu} \qquad (74)$$

whose properties have been studied in detail. These maximum likelihood conditions provide us with an equation satisfied by ν

$$\log\nu + \frac{1}{N}\sum_{j=1}^{N}\log x_j - \log\left(\frac{1}{N}\sum_{j=1}^{N} x_j\right) - \psi(\nu) = 0\,; \qquad (75)$$

once this has been solved numerically to give our estimate ν_{ML} of ν this can be used to give the estimate of b as

$$b_{\mathrm{ML}} = \frac{N\nu_{\mathrm{ML}}}{\sum_{j=1}^{N} x_j}\,. \qquad (76)$$

In the case where the data $\{x_j\}$ are assumed to have a Weibull distribution their likelihood is given by

$$P\left(\{x_j\}\middle|a,\gamma\right) = \frac{\gamma^N}{a^{\gamma N}}\prod_{j=1}^{N} x_j^{\gamma-1}\exp\left(-\sum_{j=1}^{N} x_j^{\gamma} \Big/ a^{\gamma}\right) \qquad (77)$$

from which we derive the maximum likelihood conditions

$$\frac{\partial \log\left(P\left(a,\gamma\middle|\{x_j\}\right)\right)}{\partial a} = -\frac{N\gamma}{a} + \frac{\gamma}{a^{\gamma+1}}\sum_{j=1}^{N} x_j^{\gamma} = 0$$
$$\frac{\partial \log\left(P\left(a,\gamma\middle|\{x_j\}\right)\right)}{\partial \gamma} = \frac{N}{\gamma} - N\log a + \sum_{j=1}^{N}\log x_j + \frac{1}{a^{\gamma}}\sum_{j=1}^{N} x_j^{\gamma}\left(\log a - \log x_j\right) = 0 \qquad (78)$$

These can be simplified to

$$\frac{1}{\gamma}+\frac{1}{N}\sum_{j=1}^{N}\log x_j-\frac{\sum_{j=1}^{N}x_j^{\gamma}\log x_j}{\sum_{j=1}^{N}x_j^{\gamma}}=0$$
$$a=\left(\frac{1}{N}\sum_{j=1}^{N}x_j^{\gamma}\right)^{1/\gamma} \tag{79}$$

the first of which can be solve numerically to give γ_{ML}, which can then be substituted into the second to give a_{ML}. Attempts to set up a maximum likelihood scheme for estimating the parameters of a K distribution quickly come to grief, foundering on analytically intractable items such as $\partial K_\nu(x)/\partial\nu$. As we have already mentioned the analysis of K distributed data frequently reduces to that of the underlying gamma distribution of the integrated data, to which the ML estimation procedure embodied in (82) and (83) can be applied. Oliver has discussed these and related approximate, parameter estimation schemes in great detail. These invariably require the solution of a transcendental equation involving the digamma function. Blacknell and Tough happened upon a cunning scheme that avoids this difficulty. To this end they note that the non-integer moment of the K intensity distribution is given by

$$\begin{aligned}\left\langle z^{\mu}\right\rangle &=\frac{b^{\nu}}{\Gamma(\nu)}\int_0^{\infty}\exp(-bx)x^{\nu-2}\left(\int_0^{\infty}z^{\mu}\exp(-z/x)dz\right)dx\\ &=\frac{\Gamma(\mu+1)b^{\nu}}{\Gamma(\nu)}\int_0^{\infty}\exp(-bx)x^{\nu+\mu-1}\\ &=\frac{\Gamma(\mu+1)\Gamma(\mu+\nu)}{\Gamma(\nu)b^{\mu}}\end{aligned} \tag{80}$$

Differentiation with respect to μ yields

$$\left\langle z^{\mu}\log z\right\rangle=\frac{\partial\left\langle z^{\mu}\right\rangle}{\partial\mu}=\left\langle z^{\mu}\right\rangle\left(\psi(\mu+1)+\psi(\mu+\nu)-\log b\right) \tag{81}$$

Here we have again identified the digamma function and Euler's constant.

They then form the estimator

$$\frac{\left\langle z\log z\right\rangle}{\left\langle z\right\rangle}-\log\left\langle z\right\rangle=\psi(2)-\psi(1)+\psi(\nu+1)-\psi(\nu) \tag{82}$$

At first sight, this looks horrid; all those digamma functions must be bad news. However, we recall the fundamental property of the gamma function

$$\Gamma(z+1)=z\Gamma(z) \tag{83}$$

take its logarithm

$$\log(\Gamma(z+1))=\log z+\log(\Gamma(z)) \tag{84}$$

and differentiate to give us

$$\psi(z+1) = \psi(z) + \frac{1}{z} \tag{85}$$

Using this we find that

$$\frac{\langle z \log z \rangle}{\langle z \rangle} - \langle \log z \rangle = 1 + \frac{1}{\nu} \tag{86}$$

This fellow is easy to solve for the shape parameter, once we have plugged in estimators for the average values. A detailed analysis presented in I.E.E. Proc. Radar, Sonar and Navigation, **148**, 309-312, 2001, shows that this estimation technique performs at least as well as Oliver's methods

The effects of non-Gaussian clutter on 'peak pick' detection

As a final application of our compound modeling of the non-Gaussian character of sea clutter we consider what might be termed a poor man's CFAR, in which we pick out the largest of a set of returns and identifying that with a target return. Obviously, should there be no target present, the procedure gives a false alarm; if there is a target present we can calculate the probability that the largest signal is the target signal, as a function of signal to noise ratio. Initially we consider *N* signals, of which one contains target plus Gaussian noise (modeled by the Rice distribution); the clutter returns are modeled as identical, independent speckles, with exponentially distributed intensities. Thus the probability that all intensity values are less than *x* is given by the product of the cumulative probabilities

$$\begin{aligned}
C_{\max}(x) &= C_1(x) C_0(x)^{N-1} \\
C_1(x) &= \frac{1}{2\sigma^2} \int_0^x \exp\left(-\frac{x' + A^2}{2\sigma^2}\right) I_0\left(\frac{A\sqrt{x'}}{\sigma^2}\right) dx' \\
C_0(x) &= \frac{1}{2\sigma^2} \int_0^x \exp\left(-\frac{x'}{2\sigma^2}\right) dx' = \left(1 - \exp\left(-x/2\sigma^2\right)\right)
\end{aligned} \tag{87}$$

This in turn is the cumulative probability of the global maximum value; we obtain the associated pdf by differentiation

$$\begin{aligned}
P_{\max}(x) &= \frac{dC_{\max}(x)}{dx} = P_1(x) C_0(x)^{N-1} + (N-1) C_1(x) C_0(x)^{N-2} P_0(x) \\
P_1(x) &= \frac{1}{2\sigma^2} \exp\left(-\frac{x + A^2}{2\sigma^2}\right) I_0\left(\frac{A\sqrt{x}}{\sigma^2}\right); \quad P_0(x) = \frac{1}{2\sigma^2} \exp\left(-\frac{x}{2\sigma^2}\right)
\end{aligned} \tag{88}$$

(This procedure is very similar to that adopted in the analysis of order statistics, in which all the contributing signals have the same statistics.) This pdf consists of two terms: the first, the probability that the target return takes the maximum value, while all the clutter returns take lesser values; the second, the probability that one of the clutter signals takes the maximum value, while the target and remaining clutter signals take lesser values. Thus we can write the probability that, given a maximum value *x* , it comes from the target, as

$$P(\text{target} \mid x) = \frac{P_1(x)C_0(x)^{N-1}}{P_1(x)C_0(x)^{N-1} + (N-1)C_1(x)C_0(x)^{N-2}P_0(x)} = \frac{P_1(x)C_0(x)^{N-1}}{P_{\max}(x)} \tag{89}$$

(This is a simple application of Bayes' theorem) To find the probability that the peak pick procedure identifies the target correctly we must average this over the distribution of global maximum values; the whole thing simplifies down most agreeably

$$P(\text{target}) = \int_0^\infty P(\text{target} \mid x')P_{\max}(x')dx' = \int_0^\infty P_1(x')C_0(x')^{N-1}dx' \tag{90}$$

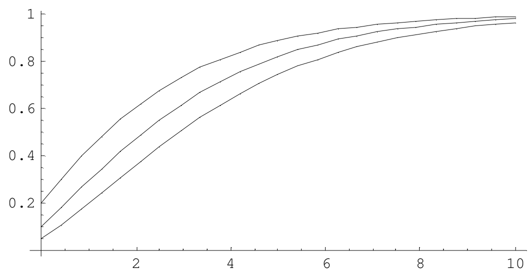
It now remains to evaluate this fellow for the simple target and clutter models we are using here. Significant analytic progress is possible, and is suitably gratifying. Thus we have

$$P(\text{target}) = \frac{1}{2\sigma^2}\int_0^\infty \exp\left(-\frac{x+A^2}{2\sigma^2}\right) I_0\left(\frac{A\sqrt{x}}{\sigma^2}\right)\left(1-\exp\left(-x/2\sigma^2\right)\right)^{N-1} \tag{91}$$

This can be expressed in the elementary form

$$P(\text{target}) = \sum_{p=0}^{N-1} \frac{(N-1)!}{p!(N-1-p)!}(-1)^p \frac{1}{(p+1)}\exp\left(-\frac{pA^2}{2(p+1)\sigma^2}\right) \tag{92}$$

which is reasonably compact and can be readily evaluated numerically. Some credence is lent to our analysis by the special cases where the SNR is zero and very large. In the latter case only the first term in the sum survives and the probability reduces to unity. In the former case the probability of correctly identifying the 'target' reduces to $1/N$, which again makes good sense.

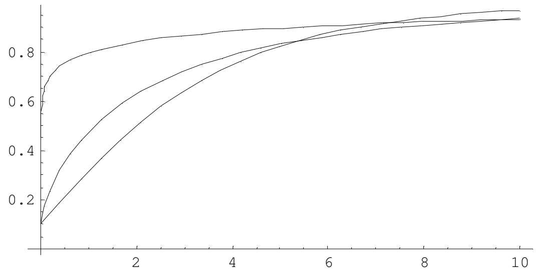


Here we plot the probability of correct identification as a function of SNR for N =5,10,20

If the clutter has a gamma distribution of powers then we can integrate this fellow up, giving an expression in terms of K Bessel functions

$$P(\text{target}) = \frac{2}{\Gamma(\nu)}\sum_{p=0}^{N-1} \frac{(N-1)!}{(p+1)!(N-1-p)!}(-1)^p \left(\frac{A^2bp}{p+1}\right)^{\nu/2} K_\nu\left(2\sqrt{\frac{A^2bp}{p+1}}\right) \tag{93}$$

which is again quite easy to evaluate. If we plot out the probability of correct identification for 10 looks, as a function of signal to clutter power ratio, we see that the tactic performs better at low SCR for smaller ν.



Plots of probability of correct identification, as a function of SCR, for $\nu = 10, 1, 1/10$. We see things working better for small SCR, and worse for large SCR, as the shape parameter ν gets smaller (and the clutter gets more spiky.). This makes sense, and ties in with the qualitative picture emerging from the analysis of CFAR performance in spiky clutter.

Exercises

As we get ever closer to the cutting edge, the exercises begin to make significant contact with the research literature. The following guide you into various extensions of the K distribution model; where appropriate, references are provided where you can find further details and, should you be pressed for time, the answers.

1 Trawl through the exercises and examples for the previous sessions, and bring together all the bits and pieces that have had you work out properties of the K distribution. Thus you should be able to work out the moments, characteristic function and probability of false alarm. Introduce Swerling 0 and 2 models for the target return and obtain expressions for the associated pdfs of the returned target plus clutter signal. Consider the effects of incoherent averaging, in which the intensities of the returns are added up. (Assume that the local power does not vary during this averaging process.)

K.D. Ward, C.J. Baker and S. Watts, '*Maritime surveillance radar, Parts 1 &2.*' IEE Proc. F, **137**, 51, 1990

2. Check the derivations of the results given for the negative binomial distribution and investigate how they reduce to those for the Poisson distribution in the appropriate limit. Starting with the discrete negative binomial rate equation, consider the limit in which the mean population becomes large. Introduce an effectively continuous variable *x* through $N = \overline{N}x$ (c.f. (52)) and derive a diffusion like equation for the pdf of *x*. Compare this with the Fokker Planck equation describing the Ornstein Uhlenbeck process introduced in session 6. Can you guess what the corresponding Langevin equation might be?

R.J.A. Tough, '*A Fokker Planck description of K distributed noise*', J. Phys. **A 20**, 551-567, 1987

3. Consider a 'compound Weibull' model in which the local power is Weibull, rather than gamma, distributed. What are the moments of the intensity associated with this distribution? Using Laplace's method, investigate the form taken by the intensity pdf for large values of its argument; to what extent is this reminiscent of the Weibull distribution itself?

4. In our discussion of the estimation of parameters characterising Gamma and K distributed quantities we encountered the 'mean of the log' on several occasions. Some insight into the utility of this quantity can be obtained from the following. Consider a complex signal represented by

$$E = a_1 \exp(i\phi_1) + a_2 \exp(i\phi_2)$$

(the *a*s and ϕs are real.) The first term represents a signal, the second a corrupting noise contribution. Assuming that the phase difference between the signal and noise terms is uniformly distributed, show that

$$\left\langle \log|E|^2 \right\rangle = 2\log|a|_>$$
$$|a|_> = \max(|a_1|,|a_2|)$$

Thus, if the amplitude of the noise is always less than that of the signal, the mean of the log of the intensity determines the amplitude of the signal directly. If the pdf of the amplitude of the noise is $p(a_2)$, show that

$$\left\langle \log|E|^2 \right\rangle = 2\log|a_1| \int_0^{a_1} p(a_2)da_2 + 2\int_{a_1}^{\infty} \log(a_2)p(a_2)da_2 .$$

For complex Gaussian noise we have

$$P(a_2) = \frac{2a_2}{x}\exp\left(-\frac{a_2^2}{x}\right);$$

in this case show that

$$\left\langle \log|E|^2 \right\rangle = 2\log a_1 + E_1\left(\frac{a_1^2}{x}\right);$$

$$E_1(z) = \int_z^{\infty} \frac{\exp(-t)dt}{t}$$

Compare this with $\log\left\langle |E|^2 \right\rangle$. Can you obtain these last few results directly from the Rice distribution? What might happen when x is gamma distributed? What conclusions can you draw about the utility of 'mean of the log' based methods when determining the shape parameter of clutter that might be corrupted by thermal noise?

K.D. Ward and R.J.A. Tough, '*Signal estimation through the mean of the log*', Electronics Letters, **24**, 85-87, 1988

5. Derive the results (92) and (93) in our discussion of the 'peak pick' procedure.