NOAA Technical Memorandum ERL WPL-84

DEPARTMENT OF COMMERCE
UNITED STATES OF AMERICA

THE PROBABILITY DISTRIBUTION FOR THE RATIO OF VARIABLES HAVING A NORMAL DISTRIBUTION

R. M. Jones

Wave Propagation Laboratory
Boulder, Colorado
November 1981

noaa NATIONAL OCEANIC AND ATMOSPHERIC ADMINISTRATION / Environmental Research Laboratories

NOAA Technical Memorandum ERL WPL-84

THE PROBABILITY DISTRIBUTION FOR THE RATIO OF VARIABLES HAVING A NORMAL DISTRIBUTION

R. M. Jones

Wave Propagation Laboratory
Boulder, Colorado
November 1981

UNITED STATES
DEPARTMENT OF COMMERCE
Malcolm Baldrige,
Secretary

NATIONAL OCEANIC AND
ATMOSPHERIC ADMINISTRATION
John V. Byrne,
Administrator

Environmental Research
Laboratories
George H. Ludwig
Director

CONTENTS

THE PROBABILITY DISTRIBUTION FOR THE RATIO OF VARIABLES HAVING A NORMAL DISTRIBUTION

R.M. Jones

ABSTRACT

The probability density function is derived for the ratio of two independent random variables, each of which has a normal distribution. When the variances of the normal distributions are small, a simplifying approximation to the probability density function for the ratio of the two variables is valid; the lower tail of the approximate distribution is approximately normal. This result validates the bias-correction calculation for minimum-envelope processing of skywave radar sea-echo spectra presented in a previous report. The distribution function for the ratio approaches a normal distribution as the variances for the numerator and denominator distributions approach zero, in agreement with previous, related calculations.

1. INTRODUCTION

When we process skywave sea-echo spectra, we average many (usually 84) independent power spectra (Georges et al., 1981). By the central limit theorem (the probability distribution of the sample mean tends to become Gaussian as the number of statistically independent samples increases without limit, regardless of the probability distribution of the random variable or process being sampled — Davenport and Root, 1958, page 81), we would expect the distribution function for the power in each of the FFT bins to be approximately normal. As a further part of our processing (minimum-envelope processing, Georges et al., 1981), we normalize several average spectra so that the first-order peaks have the same power spectral density. Such a normalization will change the distribution function for the second-order part of the spectrum. The purpose of this report is to calculate that distribution function, and find out under what conditions it is nearly a normal distribution.

For that purpose, we assume that the distribution function for each spectral line is exactly normal. The normalization described above is equivalent to taking the ratio of a second-order FFT line in the spectrum to the dominant first-order line. We are thus considering the ratio of two stochastic variables, each of which has a normal distribution, and we want to find out the distribution function for the ratio.

2. SUMMARY

If we have two independent random variables, each of which is normally distributed according to the probability density functions

$$P_1'(A) = \exp(-(A - \bar{A}_1)^2/2\sigma_1^2)/\sigma_1\sqrt{2\pi}$$

and

$$P_2'(A) = \exp(-(A - \bar{A}_2)^2/2\sigma_2^2)/\sigma_2\sqrt{2\pi} \quad ,$$

then the probability density function for the ratio of those two variables (subscript 1 for the denominator and subscript 2 for the numerator) is

$$P_3'(A) = \frac{\sigma_1\,\sigma_2}{2\pi\,\sigma_3^2\,\bar{A}_1^2}\,\exp(-(A - \bar{A})^2/2\sigma_3^2)\,(e^{-\gamma^2} + \gamma\sqrt{\pi}(1 + \mathrm{erf}(\gamma)))$$

where

$$\bar{A} \equiv \bar{A}_2/\bar{A}_1 \quad ,$$

$$\sigma_3^2 \equiv (\sigma_2^2 + \sigma_1^2\,A^2)/\bar{A}_1^2 \quad ,$$

and

$$\gamma \equiv \frac{\sigma_2^2 + \sigma_1^2\,A\,\bar{A}}{\sigma_1\,\sigma_2\,\sigma_3\sqrt{2}} \quad .$$

When the normalized variances of the initial distributions are small, (as they will be for averages of 84 spectra) the distribution for the ratio is approximately

$$P_3'(A) \approx \frac{\sigma_2^{\,2} + \sigma_1^{\,2}\, A\,\bar{A}}{\sigma_2^{\,2} + \sigma_1^{\,2}\, A^2}\, \exp(-(A - \bar{A})^2/2\sigma_3^{\,2})/\sigma_3\sqrt{2\pi} \quad .$$

The distribution of the lower tail is approximately normal. As a consequence, the bias-correction calculation derived by Georges et al. (1981) for minimum-envelope processing is valid. That calculation assumed a normal distribution in the lower tail of the distribution for the normalized average spectra.

If the normalized variances of the numerator and denominator distributions are very small, then the distribution for the ratio is approximately normal throughout its range, with a normalized variance that is the sum of the normalized variances of the numerator and denominator. The conditions for this approximation are not usually met in practice for spectra that are averages of only 84 individual spectra. The distribution functions for the ratio of second-order to first-order average power spectral density deviate significantly from a normal distribution near the peak and the upper tail of the distribution. This has negligible consequences for our minimum-envelope processing (Georges et al., 1981) because that processing is sensitive mostly to the lower tail of the distribution.

3. BRIEF SUMMARY OF METHOD AND FORMULAS

Barrick (1980) (see also Papoulis, 1965, page 196) gives the formula for calculating the probability density function $P_3'(A)$ for the ratio of two stochastic variables when the numerator has the probability density function $P_2'(A)$, and the denominator has the probability density function $P_1'(A)$. The result is

$$P_3'(A) = \frac{1}{A^2}\int_0^\infty P_2'(X)\, P_1'(X/A)\, X dX \quad . \tag{1}$$

The method for deriving the above formula is straightforward. One first converts the probability density functions (pdf's) P_1' and P_2' to the appropriate pdf's in log A. Now we are interested in the distribution function for the sum of $\log A_2$ and $-\log A_1$. This is simply the convolution of the individual pdf's (Bracewell, 1965, page 301). This yields the pdf for $\log A_3$. We then convert from a pdf in log A_3 to a pdf in A_3 to get (1). Jones (1981) demonstrates this method to calculate the distribution function for the product of two stochastic variables.

We are interested in the case where the distribution function for the first-order part of the average spectrum is normal,

$$P_1'(A) = \frac{1}{\sigma_1 \sqrt{2\pi}} \exp(-(A - \bar{A}_1)^2/2\, \sigma_1^2) \,, \tag{2}$$

as is the distribution function for the second-order part of the average spectrum.

$$P_2'(A) = \frac{1}{\sigma_2 \sqrt{2\pi}} \exp(-(A - \bar{A}_2)^2/2\, \sigma_2^2) \,. \tag{3}$$

Substituting (2) and (3) into (1) gives

$$P_3'(A) = \frac{1}{2\pi\, \sigma_1\, \sigma_2\, A^2} \int_0^\infty \exp(- \frac{(X - \bar{A}_2)^2}{2\, \sigma_2^2} - \frac{(X/A - \bar{A}_1)^2}{2\, \sigma_1^2})\, XdX \,. \tag{4}$$

The Appendix evaluates the above integral to give

$$P_3'(A) = \frac{\sigma_1\, \sigma_2}{2\, \pi\, \sigma_3^2\, \bar{A}_1^2} \, (e^{-\gamma^2} + \gamma\sqrt{\pi}(1 + \mathrm{erf}(\gamma)))\, \exp(-(A - \bar{A})^2/2\, \sigma_3^2) \,, \tag{5}$$

where

$$\bar{A} \equiv \bar{A}_2/\bar{A}_1 \,, \tag{6}$$

$$\sigma_3^2 \equiv \frac{\sigma_2^2 + \sigma_1^2\, A^2}{\bar{A}_1^2} \,, \tag{7}$$

and

$$\gamma = \frac{\sigma_2^2 + \sigma_1^2\, A\, \bar{A}}{\sigma_1\, \sigma_2\, \sigma_3 \sqrt{2}} \,. \tag{8}$$

Equation (5) gives the probability density function for the ratio of two stochastic variables when each of those variables has the normal distributions given in (2) and (3).

4. APPROXIMATION FOR SMALL VARIANCE

When the normalized variances $\sigma_1^2/\bar{A}_1^2$ and $\sigma_2^2/\bar{A}_2^2$ are small, γ in (8) becomes large, so that the first term in (5) is negligible, and the error function in (5) is nearly unity. Under these approximations, (5) becomes

$$P_3'(A) \approx \frac{\sigma_2^2 + \sigma_1^2 A \bar{A}}{\sigma_2^2 + \sigma_1^2 A^2} \exp(-(A - \bar{A})^2/2\sigma_3^2)/\sigma_3 \sqrt{2\pi} . \tag{9}$$

The above approximation is nearly always valid for spectra that are the average of 84 or more individual spectra.

5. THE LOWER TAIL OF THE DISTRIBUTION IS NORMAL

For the normal distributions we are considering, the normalized variance is about 2.5/84 (Georges et al., 1981). That is,

$$\frac{\sigma_1^2}{\bar{A}_1^2} \approx \frac{2.5}{84} \tag{10}$$

and

$$\frac{\sigma_2^2}{\bar{A}_2^2} \approx \frac{2.5}{84} . \tag{11}$$

Under these conditions, (7) can be rewritten to give

$$\frac{\sigma_3^2}{\bar{A}^2} \approx \frac{2.5}{84} (1 + \frac{A^2}{\bar{A}^2}) , \tag{12}$$

and the fraction in (9) can be rewritten to give

$$\frac{\sigma_2^{\,2} + \sigma_1^{\,2}\, A\,\bar{A}}{\sigma_2^{\,2} + \sigma_1^{\,2}\, A^2} \approx \bar{A}\,\frac{\bar{A} + A}{\bar{A}^2 + A^2} \quad . \tag{13}$$

For $A \ll \bar{A}$ (the lower tail of the distribution), (12) becomes

$$\frac{\sigma_3^{\,2}}{\bar{A}^2} \approx \frac{2.5}{84} \approx \frac{\sigma_2^{\,2}}{\bar{A}_2^2} = \text{constant}, \tag{14}$$

and (13) becomes

$$\frac{\sigma_2^{\,2} + \sigma_1^{\,2}\, A\,\bar{A}}{\sigma_2^{\,2} + \sigma_1^{\,2}\, A^2} \approx 1 \quad . \tag{15}$$

Thus, for the lower tail of the distribution, (9) is approximately

$$P_3'(A) \approx \exp(-(A - \bar{A})^2/2\sigma_3^{\,2})/\sigma_3\sqrt{2\pi} \quad , \tag{16}$$

which is a normal distribution because the variance given by (14) is a constant.

Not only is the lower tail of the distribution normal; it also has the same normalized variance as the numerator. Thus, the lower tail of the distribution for the ratio has the same distribution as the numerator. Georges et al. (1981) make use of this property in their analysis of minimum-envelope processing of average spectra.

6. THE DISTRIBUTION REDUCES TO A NORMAL DISTRIBUTION FOR VERY SMALL VARIANCE

From (A-4) (which is equivalent to (7)), we have

$$\frac{\sigma_3^{\,2}}{\bar{A}^2} = \frac{\sigma_2^{\,2}}{\bar{A}_2^2} + \frac{\sigma_1^{\,2}}{\bar{A}_1^2}\,\frac{A^2}{\bar{A}^2} \quad . \tag{17}$$

All other parameters being equal, as the normalized variances $\sigma_1^{\ 2}/\bar{A}_1^2$ and $\sigma_2^{\ 2}/\bar{A}_2^2$ approach zero, the normalized variance $\sigma_3^{\ 2}/\bar{A}^2$ will also approach zero according to (17). When this happens, the exponential in (9) will dominate so that (9) will be small except when

$$A \approx \bar{A} \ . \tag{18}$$

We may thus use the approximation (18) everywhere except in the exponential of (9) without losing much accuracy. In particular, substituting (18) into (17) gives

$$\frac{\sigma_3^{\ 2}}{\bar{A}^2} \approx \frac{\sigma_2^{\ 2}}{\bar{A}_2^2} + \frac{\sigma_1^{\ 2}}{\bar{A}_1^2} \quad , \tag{19}$$

showing that the effective normalized variance for the distribution of the ratio is the sum of the normalized variances of the numerator and denominator and is effectively constant. Substituting (18) into all but the exponential part of (9) gives

$$P_3'(A) \approx \exp(-(A - \bar{A})^2/2\ \sigma_3^{\ 2})/\sigma_3\ \sqrt{2\pi} \quad , \tag{20}$$

showing that the distribution is normal when the variance is small enough.

7. RELATION TO THE F DISTRIBUTION

Barrick (1980) considered the distribution function for the ratio of stochastic variables when the distribution functions for both the numerator and denominator are chi-squared (Abramowitz and Stegun, 1964, page 940). As pointed out by Barrick, the distribution function for such a ratio is an F distribution (Abramowitz and Stegun, 1964, page 946). Chi-square distributions are relevant for distribution of power from sea-echo backscatter (Barrick and Snider, 1977).

For skywave backscatter, the distribution function is more complicated and has a larger normalized variance because of the variability of the ionospheric reflection coefficient (Georges et al., 1981). The distribution function is no longer chi-square with 2 degrees of freedom (a Rayleigh distribution), but is closer to a Hankel distribution (Jones, 1981).

Still, we would expect the distribution function considered in the present report and the F distribution considered by Barrick to approach each other in the limit of small variance because of the central limit theorem. This is, in fact what happens. Barrick shows that for small variance the F distribution approaches a normal distribution with the same normalized variance as that given by (19).

8. REFERENCES

Abramowitz, M. and I.A. Stegun (eds.), Handbook of Mathematical Functions, National Bureau of Standards Applied Mathematical Series 55, U.S. Government Printing Office, Washington, D.C., Second printing, November 1964.

Barrick, D.E., Accuracy of parameter extraction from sample-averaged sea-echo Doppler spectra, IEEE Trans. Antennas Propagation AP-28, pp. 1-11, 1980.

Barrick, D.E. and J.B. Snider, The statistics of HF sea-echo Doppler spectra, IEEE Trans. Antennas Propagation AP-25, pp. 19-24, 1977.

Bracewell, R., The Fourier Transform and its Applications, McGraw-Hill, New York, 1965.

Davenport, W.B., Jr. and W.L. Root, An Introduction to the Theory of Random Signals and Noise, McGraw-Hill, New York, 1958.

Georges, T.M., J.W. Maresca, C.T. Carlson, J.P. Riley, R.M. Jones, and D.E. Westover, Recovering ocean wave height from HF radar sea echoes distorted by imperfect ionospheric reflection, NOAA Tech. Memo. ERL WPL-73, 1981.

Jones, R.M., The Hankel probability distribution--the probability distribution for the product of two random variables, each of which has a Rayleigh distribution, NOAA Tech. Memo. ERL WPL-79, 1981.

Papoulis, Athanasios, Probability, Random Variables, and Stochastic Processes, McGraw Hill, New York, 1965.

APPENDIX

DERIVATION OF THE PROBABILITY DISTRIBUTION

We can rewrite (4) (leaving out considerable algebra) as

$$P'_3(A) = \frac{\sigma_1 \sigma_2 \alpha}{2 \pi \sigma_3^2 \bar{A}_1^2} \exp(-(A - \bar{A})^2/2\sigma_3^2) \int_0^\infty \exp(-\alpha(X-\beta)^2/2) \, XdX \qquad \text{(A-1)}$$

where

$$\bar{A} \equiv \bar{A}_2/\bar{A}_1 \ , \qquad \text{(A-2)}$$

$$\sigma_3^2 = \frac{\sigma_2^2 + \sigma_1^2 A^2}{\bar{A}_1^2} \qquad \text{(A-3)}$$

(which can also be expressed as $\frac{\sigma_3^2}{\bar{A}^2} = \frac{\sigma_2^2}{\bar{A}_2^2} + \frac{\sigma_1^2}{\bar{A}_1^2} \frac{A^2}{\bar{A}^2}$) , (A-4)

$$\alpha \equiv \frac{\sigma_3^2 \bar{A}_1^2}{\sigma_1^2 \sigma_2^2 A^2} \ , \qquad \text{(A-5)}$$

and

$$\beta \equiv A \bar{A}_1 \frac{\sigma_2^2 + \sigma_1^2 A \bar{A}}{\sigma_2^2 + \sigma_1^2 A^2} = A \frac{\sigma_2^2 + \sigma_1^2 A \bar{A}}{\bar{A}_1 \sigma_3^2} \ . \qquad \text{(A-6)}$$

The integral in (A-1) can be rewritten with an obvious change in variable as

$$\int_{-\beta}^\infty (y + \beta) \exp(-\alpha y^2/2) \, dy \ , \qquad \text{(A-7)}$$

which in turn can be broken up into two integrals as

$$\int_{-\beta}^{\infty} y \exp(-\alpha\, y^2/2)dy + \beta \int_{-\beta}^{\infty} \exp(-\alpha\, y^2/2)dy \; . \tag{A-8}$$

The integral on the left is exact, and the integral on the right can be evaluated in terms of the error function to give

$$\frac{1}{\alpha}\,(\exp(-\gamma^2) + \gamma\sqrt{\pi}(1 + \text{erf}(\gamma))) \; , \tag{A-9}$$

where

$$\gamma \equiv \beta\,\sqrt{\alpha/2} = \frac{\sigma_2^{\,2} + \sigma_1^{\,2}\, A\, \bar{A}}{\sigma_1\, \sigma_2\, \sigma_3\, \sqrt{2}} \; . \tag{A-10}$$

Substituting (A-10) for the integral in (A-1) gives

$$P_3'(A) = \frac{\sigma_1\, \sigma_2\, \exp(-(A - \bar{A})^2/2\sigma_3^{\,2})}{2\pi\, (\sigma_2^{\,2} + \sigma_1^{\,2}\, A^2)}\,(e^{-\gamma^2} + \gamma\sqrt{\pi}(1 + \text{erf}(\gamma))) \; . \tag{A-11}$$