II. Variational Distance and Error Probabilities

When testing two simple hypotheses (measures) P and Q the minimal possible sum $\inf\{\alpha+\beta\}$ of both error probabilities satisfies a simple relation

$$\inf\{\alpha+\beta\} = 1 - \frac{1}{2}\|P-Q\| = 1 - \frac{1}{2}\int_{\mathcal{X}} |dP - dQ|. \quad (7)$$

Relation (7) (and its natural generalization through the convex hull of measures for composite hypotheses) was proved first by C. Kraft [4]. Much later (but independently !) it was obtained also in [3], where (see also [1]) some examples of application of a generalized version of relation (7) in testing of "very composite" hypotheses are presented.

A good collection of various estimates for $\|P-Q\|$ can be found in [5, Ch. 4] (where the author has learned about the reference [4] for relation (7)).

Due to relation (7) we can reformulate Corollary 1 in a pure geometrical form that supplements the collection in [5].

Corollary 2: The following bounds for $\|P-Q\|$ are valid:

$$2\left(1 - \exp\left\{\mu(s^*) - \frac{1}{2}\sqrt{\mu''(s^*)}\right\}\right) \le \|P-Q\| \le 2(1-\exp\{\mu(s^*)\}) \quad (8)$$

Acknowledgment

The author wishes to thank the anonymous referee for many good comments that helped to improve the presentation of the material.

References

[1] L. A. Bassalygo and M. V. Burnashev, "Authentication, identification and pairwise separated measures," *Probl. Inform. Transm.*, vol. 32, no. 1, pp. 41–47, 1996.
[2] W. Bounjakowsky, "Mémoires de l'Académie des Sciences de St-Pétersbourg, 7 série," vol. 1, no. 9, 1859.
[3] M. V. Burnashev, "Minimax detection of inaccurately known signal in the background of white Gaussian noise," *Theory Prob. and Its Appl.*, vol. 24, no. 1, pp. 106–118, 1979.
[4] C. H. Kraft, "Some conditions for consistency and uniform consistency of statistical procedures," *Univ. California Publ. in Statistics*, no. 1, pp. 125–142, 1955.
[5] L. L. Cam, *Asymptotic Methods in Statistical Decision Theory*. New York: Springer-Verlag, 1986.
[6] V. V. Petrov, *Sums of Independent Random Variables.* Berlin: Springer-Verlag, 1975.
[7] C. E. Shannon, R. G. Gallager, and E. R. Berlekamp, "Lower bounds to error probability for coding on discrete memoryless channels," *Inform. Contr.*, pt. 1, vol. 10, pp. 65–103, 1967.
[8] A. J. Viterbi and J. K. Omura, *Principles of Digital Communication and Coding*. New York: McGraw-Hill, 1979.

The Real-Complex Normal Distribution

A. van den Bos, *Senior Member, IEEE*

***Abstract*—An expression is derived for the distribution of a mixture of real and complex normal variates.**

***Index Terms*— Complex distributions, complex stochastic variables, normal distribution.**

I. Introduction

In [1] an expression is derived for the multivariate complex normal distribution. It generalizes complex normal distributions proposed earlier and specialized to a limited class of covariance matrices. However, the distribution in [1] cannot be used if one or more of the variates are real. An example is the asymptotic distribution of the estimates of the parameters of the complex-valued exit wave of a periodic crystal specimen from noise disturbed intensity observations in transmission electron microscopy [2]. The parameters in this problem are the Fourier coefficients of the wave and both spatial periods; their estimates are used to reconstruct the complex-valued wave. One of the Fourier coefficients is real and so are the periods; all further Fourier coefficients are complex. In [3] it is shown how these real and complex parameters are estimated *directly* as real and complex quantities. As compared with separate estimation of the real and imaginary parts, this considerably simplifies the expressions involved, in particular those for the first-order and second-order partial derivatives used in the numerical maximization of the likelihood function concerned. As a result, the pertinent code is simplified correspondingly [4]. The asymptotic distribution of the resulting real-complex maximum-likelihood estimates is the real-complex normal distribution derived in this correspondence. The covariance matrix of this distribution is particularly important. It is the asymptotic covariance matrix for maximum-likelihood estimates and the Cramér–Rao lower bound on the variance of the real-complex estimates in general. From this covariance matrix, the variance of the reconstructed complex-valued exit wave then follows using the pertinent propagation formulas. The resulting expressions show the dependence of the variance on the free microscope parameters used for experimental design.

In Section II, the main result, the general expression for the real-complex normal distribution, is derived. Special cases are also presented.

II. Real-Complex Normal Distribution

Assume that the vector of normally distributed variates $\boldsymbol{w} \in \mathbb{R}^{(K+2L)\times 1}$ is described by

$$\boldsymbol{w} = (r_1 \ \cdots \ r_K \ \ x_1 \ \ y_1 \ \cdots \ x_L \ \ y_L)^T \quad (1)$$

where the superscript T denotes transposition. Then the probability density function of $\boldsymbol{w}$ is

$$\frac{1}{(2\pi)^{(K+2L)/2}(\det \boldsymbol{W})^{1/2}} \exp\left(-\frac{1}{2}\boldsymbol{w}^T\boldsymbol{W}^{-1}\boldsymbol{w}\right) \quad (2)$$

Manuscript received February 2, 1997; revised January 15, 1998.
The author is with the Department of Applied Physics, Delft University of Technology, P.O. Box 5046, 2600 GA Delft, The Netherlands (e-mail: a.vandenbos@tn.tudelft.nl).
Publisher Item Identifier S 0018-9448(98)03635-9.

where $\boldsymbol{W} \in \mathbb{R}^{(K+2L)\times(K+2L)}$ is the covariance matrix of $\boldsymbol{w}$. For simplicity it will be assumed that the expectation $E[\boldsymbol{w}]$ of $\boldsymbol{w}$ is equal to the null vector. Next define the vector of real and complex variates $\boldsymbol{v} \in \mathbb{C}^{(K+2L)\times 1}$ by

$$\boldsymbol{v} = (r_1 \cdots r_K \;\; z_1 \;\; z_1^* \cdots z_L \;\; z_L^*)^T \tag{3}$$

where $z_\ell = x_\ell + jy_\ell$ and z_ℓ^* is the conjugate of z_ℓ with $j^2 = -1$. Then

$$\begin{pmatrix} z_\ell \\ z_\ell^* \end{pmatrix} = J \begin{pmatrix} x_\ell \\ y_\ell \end{pmatrix} \tag{4}$$

where the matrix $\boldsymbol{J} \in \mathbb{C}^{2\times 2}$ is defined by

$$\boldsymbol{J} = \begin{pmatrix} 1 & j \\ 1 & -j \end{pmatrix}. \tag{5}$$

Therefore,

$$\boldsymbol{v} = \boldsymbol{B}\boldsymbol{w} \tag{6}$$

where $\boldsymbol{B} \in \mathbb{C}^{(K+2L)\times(K+2L)}$ is defined as the block diagonal matrix

$$\boldsymbol{B} = \mathrm{diag}\,(\boldsymbol{I} \; \boldsymbol{A}) \tag{7}$$

where $\boldsymbol{I} \in \mathbb{R}^{K\times K}$ is the identity matrix of order K while $\boldsymbol{A} \in \mathbb{C}^{2L\times 2L}$ is defined as the block diagonal matrix

$$\boldsymbol{A} = \mathrm{diag}\,(\boldsymbol{J} \cdots \boldsymbol{J}). \tag{8}$$

From (6) it follows that the covariance matrix $\boldsymbol{V} \in \mathbb{C}^{(K+2L)\times(K+2L)}$ of $\boldsymbol{v}$, defined as $E[\boldsymbol{v}\;\boldsymbol{v}^H]$, is equal to

$$\boldsymbol{V} = \boldsymbol{B}E[\boldsymbol{w}\;\boldsymbol{w}^T]\boldsymbol{B}^H = \boldsymbol{B}\boldsymbol{W}\boldsymbol{B}^H \tag{9}$$

and hence

$$\boldsymbol{W} = \boldsymbol{B}^{-1}\boldsymbol{V}\boldsymbol{B}^{-H}. \tag{10}$$

In these expressions, the superscript H denotes complex conjugate transposition. Since $\boldsymbol{J}^{-1} = \frac{1}{2}\boldsymbol{J}^H$, it follows that $\boldsymbol{B}^{-1} = \mathrm{diag}\,(\boldsymbol{I}\;\frac{1}{2}\boldsymbol{A}^H)$ and $\boldsymbol{B}^{-H} = \mathrm{diag}\,(\boldsymbol{I}\;\frac{1}{2}\boldsymbol{A})$. Then, by (10)

$$\det \boldsymbol{W} = (\tfrac{1}{2}\,j)^L \det \boldsymbol{V}(-\tfrac{1}{2}\,j)^L = \det \boldsymbol{V}/4^L \tag{11}$$

since $\det \boldsymbol{J} = -2j$ and $\det \boldsymbol{J}^H = 2j$. Furthermore, since $\boldsymbol{w} = \boldsymbol{B}^{-1}\boldsymbol{v}$ and $\boldsymbol{w}^T = \boldsymbol{w}^H = \boldsymbol{v}^H\boldsymbol{B}^{-H}$

$$\begin{aligned}\boldsymbol{w}^T\boldsymbol{W}^{-1}\boldsymbol{w} &= \boldsymbol{v}^H\boldsymbol{B}^{-H}\boldsymbol{W}^{-1}\boldsymbol{B}^{-1}\boldsymbol{v} \\ &= \boldsymbol{v}^H(\boldsymbol{B}\boldsymbol{W}\boldsymbol{B}^H)^{-1}\boldsymbol{v}.\end{aligned} \tag{12}$$

Therefore, by (9)

$$\boldsymbol{w}^T\boldsymbol{W}^{-1}\boldsymbol{w} = \boldsymbol{v}^H\boldsymbol{V}^{-1}\boldsymbol{v}. \tag{13}$$

Substituting (11) and (13) in (2) yields

$$\frac{1}{2^{K/2}\pi^{(K/2+L)}(\det \boldsymbol{V})^{1/2}} \exp\left(-\frac{1}{2}\,\boldsymbol{v}^H\boldsymbol{V}^{-1}\boldsymbol{v}\right). \tag{14}$$

Next, rearrange the elements of $\boldsymbol{v}$ as follows:

$$\boldsymbol{u} = \boldsymbol{P}\boldsymbol{v} \tag{15}$$

where $\boldsymbol{P} \in \mathbb{R}^{(K+2L)\times(K+2L)}$ is the permutation matrix such that

$$\boldsymbol{u} = (r_1 \cdots r_K \;\; z_1 \cdots z_L \;\; z_1^* \cdots z_L^*)^T. \tag{16}$$

Next, let $\boldsymbol{U} \in \mathbb{C}^{(K+2L)\times(K+2L)}$ be the covariance matrix of $\boldsymbol{u}$. Then by (15)

$$\boldsymbol{U} = E[\boldsymbol{u}\boldsymbol{u}^H] = \boldsymbol{P}\boldsymbol{V}\boldsymbol{P}^T. \tag{17}$$

Because permutation matrices are orthogonal and since the absolute value of their determinant is equal to one [5, p. 360 and p. 25], it follows from (15) and (17) that

$$\det \boldsymbol{V} = \det \boldsymbol{U} \tag{18}$$

and

$$\boldsymbol{v}^H\boldsymbol{V}^{-1}\boldsymbol{v} = \boldsymbol{u}^H\boldsymbol{U}^{-1}\boldsymbol{u}. \tag{19}$$

Then substituting (18) and (19) in (14) yields

$$\frac{1}{2^{K/2}\pi^{(K+2L)/2}(\det \boldsymbol{U})^{1/2}} \exp\left(-\frac{1}{2}\,\boldsymbol{u}^H\boldsymbol{U}^{-1}\boldsymbol{u}\right). \tag{20}$$

This is the expression for the normal probability density function of the K real variates $r_1, \cdots, r_K$ and the $2L$ complex variates $z_1, \cdots, z_L, z_1^*, \cdots, z_L^*$. It is the main result of this correspondence.

For the description of special cases of this probability density, the covariance matrix $\boldsymbol{U}$ is partitioned as follows:

$$\boldsymbol{U} = \begin{pmatrix} \boldsymbol{R} & \boldsymbol{Q} & \boldsymbol{Q}^* \\ \boldsymbol{Q}^H & \boldsymbol{Z} & \boldsymbol{S} \\ \boldsymbol{Q}^T & \boldsymbol{S}^H & \boldsymbol{Z}^* \end{pmatrix} \tag{21}$$

where $\boldsymbol{R} \in \mathbb{R}^{K\times K}$ is the covariance matrix $E[\boldsymbol{r}\;\boldsymbol{r}^T]$ of $\boldsymbol{r} = (r_1 \cdots r_K)^T$, $\boldsymbol{Q} \in \mathbb{C}^{K\times 2L}$ is the covariance matrix $E[\boldsymbol{r}\;\boldsymbol{z}^H]$ of $\boldsymbol{r}$ and $\boldsymbol{z} = (z_1 \cdots z_L)^T$, $\boldsymbol{Z} \in \mathbb{C}^{L\times L}$ is the covariance matrix of $\boldsymbol{z}$, and $\boldsymbol{S} \in \mathbb{C}^{L\times L}$ is the covariance matrix $E[\boldsymbol{z}\;\boldsymbol{z}^T]$ of $\boldsymbol{z}$ and $\boldsymbol{z}^*$. First, consider the special case that the elements of $\boldsymbol{r}$ are uncorrelated with those of $\boldsymbol{z}$. Then $\boldsymbol{Q}$ is equal to the null matrix and

$$\boldsymbol{U}^{-1} = \mathrm{diag}\,(\boldsymbol{R}^{-1}\;\boldsymbol{N}^{-1}) \tag{22}$$

with

$$\boldsymbol{N} = \begin{pmatrix} \boldsymbol{Z} & \boldsymbol{S} \\ \boldsymbol{S}^H & \boldsymbol{Z}^* \end{pmatrix} \tag{23}$$

and the probability density function becomes

$$\frac{1}{2^{K/2}\pi^{(K+2L)/2}(\det \boldsymbol{R})^{1/2}(\det \boldsymbol{N})^{1/2}} \cdot \exp\left\{-\tfrac{1}{2}\left(\boldsymbol{r}^T\boldsymbol{R}^{-1}\boldsymbol{r} + \boldsymbol{n}^H\boldsymbol{N}^{-1}\boldsymbol{n}\right)\right\} \tag{24}$$

where $\boldsymbol{n} = (\boldsymbol{z}^T\;\boldsymbol{z}^H)^T$ with covariance matrix $\boldsymbol{N} \in \mathbb{C}^{2L\times 2L}$. If, in addition, as is often assumed in the literature [1], $E[z_p\;z_q] = E[z_p^*\;z_q^*] = 0$, the matrix $\boldsymbol{S}$ is equal to the null matrix. Then it may be shown that [1]

$$\frac{1}{\pi^L(\det \boldsymbol{N})^{1/2}} \exp\left(-\frac{1}{2}\,\boldsymbol{n}^H\boldsymbol{N}^{-1}\boldsymbol{n}\right) = \frac{1}{\pi^L \det \boldsymbol{Z}} \exp\left(-\boldsymbol{z}^H\boldsymbol{Z}^{-1}\boldsymbol{z}\right) \tag{25}$$

and, therefore, the probability density becomes

$$\frac{1}{2^{K/2}\pi^{(K/2+L)}(\det \boldsymbol{R})^{1/2}\det \boldsymbol{Z}} \cdot \exp\left(-\tfrac{1}{2}\,\boldsymbol{r}^T\boldsymbol{R}^{-1}\boldsymbol{r} - \boldsymbol{z}^H\boldsymbol{Z}^{-1}\boldsymbol{z}\right). \tag{26}$$

For $K = L = 1$, $\boldsymbol{r}$ and $\boldsymbol{z}$ become scalars r and z with probability density function

$$\frac{1}{\sqrt{2}\pi^{3/2}\sigma_r\sigma_z^2} \exp\left(-\frac{1}{2}\,r^2/\sigma_r^2 - zz^*/\sigma_z^2\right). \tag{27}$$

REFERENCES

[1] A. van den Bos, "The multivariate complex normal distribution—A generalization," *IEEE Trans. Inform Theory*, vol. 41, pp. 537–539, 1995.
[2] ——, "Complex electron wave reconstruction using parameter estimation," *IEEE Trans. Instrum. Meas.*, vol. 46, pp. 826–830, 1997.
[3] ——, "Volterra transfer models for electron wave reconstruction," in *Proc. 1997 IEEE Workshop on Nonlinear Signal and Image Processing NSIP'97* (Mackinac Island, MI, Sept. 1997), vol. 4, on CD-ROM.
[4] ——, "Complex gradient and Hessian," *Proc. Inst. Elec. Eng., Vision, Image and Signal Processing*, vol. 141, pp. 380–382, 1994.
[5] R. A. Horn and C. R. Johnson, *Matrix Analysis.* Cambridge, U.K.: Cambridge Univ. Press, 1992.

Zero-Crossing Rates of Mixtures and Products of Gaussian Processes

John T. Barnett and Benjamin Kedem

***Abstract*—Formulas for the expected zero-crossing rate of non-Gaussian mixtures and products of Gaussian processes are obtained. The approach we take is to first derive the expected zero-crossing rate in discrete time and then obtain the rate in continuous time by an appropriate limiting argument. The processes considered, which are non-Gaussian but derived from Gaussian processes, serve to illustrate the variability of the zero-crossing rate in terms of the normalized autocorrelation function $\rho(t)$ of the process. For Gaussian processes, Rice's formula gives the expected zero-crossing rate in continuous time as $\frac{1}{\pi}\sqrt{-\rho''(0)}$. We show processes exist with expected zero-crossing rates given by $\frac{\kappa}{\pi}\sqrt{-\rho''(0)}$ with either $\kappa \gg 1$ or $\kappa \ll 1$. Consequently, such processes can have an arbitrarily large or small zero-crossing rate as compared to a Gaussian process with the same autocorrelation function.**

***Index Terms*—Autocorrelation, cosine formula, expected zero-crossing rate, non-Gaussian processes, Rice's formula.**

I. INTRODUCTION

Consider a zero-mean, stricly stationary Gaussian process $\{Z(t)\}$, $-\infty < t < \infty$, with autocovariance $R(t)$ and autocorrelation function $\rho(t)$. We assume throughout that the variance of the underlying Gaussian process $\{Z(t)\}$ is one so that $R(0) = \rho(0) = 1$. If $\{Z(t)\}$ is mean-square-differentiable, that is, if $\rho''(0)$ exists and is finite, then the expected number of zero crossings per unit time is given by Rice's formula ([17], [19])

$$E[D_c] = \frac{1}{\pi}\sqrt{-\rho''(0)} \tag{1}$$

where D_c (c for continuous) is the number of zero crossings of $\{Z(t)\}$ for t in the unit interval $[0, 1]$, and $\rho''(0)$ is the second derivative of the autocorrelation function of $\{Z(t)\}$ at 0. In the sequel we shall continue to use D_c to denote the zero-crossing rate in continuous time regardless of the process.

Manuscript received December 22, 1995; revised November 1, 1997. This work was supported by the Independent Research Program Office of SPAWAR Systems Center, San Diego, CA, and under Grants AFOSR-89-0049, ONR N00014-92-C-0019, and NSF CDR-88-03012.

J. T. Barnett is with SPAWAR Systems Center, Electro-Optics Branch Code D743, San Diego, CA 92152 USA.

B. Kedem is with the Mathematics Department and Institute for Systems Research, University of Maryland, College Park, MD 20742 USA.

Publisher Item Identifier S 0018-9448(98)03469-5.

The analogous formula for a discrete-time, zero-mean, unit variance, stationary Gaussian sequence $\{Z(k)\}$, $k = 0, \pm 1, \pm 2 \cdots$ is given by ([14], [19], [9])

$$\rho_1 = \cos\frac{\pi E[D_1]}{N-1} \tag{2}$$

where D_1 is the number of sign changes or zero crossings in $\{Z(1), \cdots, Z(N)\}$, $\rho_k = E[Z(k+j)Z(j)]$ is the correlation sequence of $\{Z(k)\}$, and $E[D_1]/(N-1)$ is the expected zero-crossing rate in discrete time. We refer to (2) as the "cosine formula."

In this correspondence we present extensions of Rice's formula of the form $\frac{\kappa}{\pi}\sqrt{-\rho''(0)}$ where $\kappa \le 1$ or $\kappa \ge 1$, and $\rho(t)$ is the autocorrelation function of the process in question.

Our approach is to first derive the expected zero-crossing rate in discrete time (to obtain a cosine formula) and by an appropriate limiting argument arrive at the zero-crossing rate in continuous time. In particular, we derive analogs of the "cosine formula" and "Rice's formula" for a scaled-time mixture of a Gaussian process, for general mixtures of Gaussian processes, and for products of Gaussian processes.

Mixtures and products of Gaussian processes are used, in both engineering and physics, as models in such diverse areas as: rainfall, body weights, crushing processes, diffusive transport in random media, and multifractal processes (see [10], [7], and [16]). Hence, knowing the zero-crossing rates for such processes is of practical value.

To motivate our investigation, we first discuss a formal "orthant probability formula" for random processes satisfying mild stationarity requirements. Using a formal "cosine formula," a formal "orthant probability formula" is obtained from which we argue that, in general,

$$E[D_c] = \frac{\kappa}{\pi}\sqrt{-\rho''(0)} \tag{3}$$

for sufficiently smooth processes. Moreover, the fact that κ may be quite different than one in (3) serves as a warning that Rice's formula, (1), may not be indiscriminately applied in the non-Gaussian case (e.g. [3, p. 149], [8, p. 236], and [15, p. 1398]).

A. A Formal Orthant Probability Formula

Let $\{Z(t)\}$, $-\infty < t < \infty$, be a stochastic process consisting of continuous random variables with mean zero and satisfying the "stationarity" requirement

$$\Pr[Z(t) \ge 0] = \frac{1}{2}$$
$$\Pr[Z(t) \ge 0, Z(s) \ge 0] = g(|t-s|)$$

for some function $g(\cdot)$. For $t \in [0, 1]$ and for a positive integer $N > 2$ we define the discrete time process

$$Z_k \equiv Z((k-1)\Delta), \qquad k = 1, 2, \cdots, N$$

such that

$$(N-1)\Delta = 1. \tag{4}$$

The interval $(0, 1]$ is now partitioned into $N-1$ subintervals each of length Δ so that $\{Z_k\}$ is simply $\{Z(t)\}$ evaluated at the endpoints of the subintervals. Define the indicator

$$d_k = I_{[\text{Sign change in } Z_k, Z_{k-1}]} = I_{[Z_k \ge 0, Z_{k-1} < 0 \cup Z_{k-1} \ge 0, Z_k < 0]}.$$