The Annals of Mathematical Statistics
1972, Vol. 43, No. 6, 2000-2003

ON THE CORRELATION COEFFICIENT OF A BIVARIATE, EQUAL VARIANCE, COMPLEX GAUSSIAN SAMPLE[1]

BY TOBY BERGER

Cornell University

Let u_n denote the sample correlation coefficient for n observations from a bivariate, equal variance, complex Gaussian distribution. In this note we derive the exact distribution of u_n by extending a method of Mehta and Gurland to the complex case. The asymptotic behavior of $E|u_n|^k$ as $n \to \infty$ is determined via the method of steepest descent. Applicability of the results to the analysis of certain estimators of spectral parameters of stationary time series is discussed.

1. Density of u_n. Let $\xi_1, \xi_2, \cdots, \xi_n$ be an independent sample from a zero mean, bivariate, equal variance, complex Gaussian distribution with correlation matrix

$$\Sigma_\xi = E\xi\bar{\xi}' = \begin{pmatrix} \sigma^2 & \sigma^2\rho \\ \sigma^2\bar{\rho} & \sigma^2 \end{pmatrix}. \tag{1}$$

The Hermitian sample correlation matrix $A = (A_{ij}) = n^{-1} \sum_{k=1}^n \xi_k \bar{\xi}_k'$, $1 \leqq i, j \leqq 2$, then has the bivariate complex Wishart density

$$p(A) = |A|^{n-2}[\pi\Gamma(n)\Gamma(n-1)|\Sigma_\xi|^n]^{-1} \exp[-\operatorname{tr}(\Sigma_\xi^{-1}A)]. \tag{2}$$

(See Goodman (1963) for a detailed discussion of the complex Gaussian and complex Wishart distributions.) The function $p(A)$, defined over the domain where A is Hermitian positive semi-definite, is a compact way of writing the joint density of the four real random variables A_{11}, A_{22}, A_{12R}, and A_{12I}. The usual estimator u_n of the complex correlation coefficient ρ is a function of the elements of A, namely

$$u_n = 2A_{12}[A_{11} + A_{22}]^{-1}.$$

Throughout the remainder of this section we suppress the subscript n on u_n. The joint probability density of the real random variables u_R and u_I defined by the relation $u = u_R + iu_I$ may be found by extending the method of Mehta and Gurland (1969) to the complex case as follows. First, two auxiliary variables $v = A_{11} + A_{22}$ and $w = A_{22}$ are introduced. A simple calculation reveals that the magnitude of the Jacobian of the transformation $(A_{11}, A_{22}, A_{12R}, A_{12I}) \to (u_R, u_I, v, w)$ equals $(v/2)^2$. It follows from (1), (2), and some algebra that $p(u_R, u_I, v, w) = K_1(v/2)^2[(v-w)w - (|u|v/2)^2]^{n-2} \exp[-v\sigma^{11} + v\operatorname{Re}(\sigma^{12}\bar{u})]$, where $K_1^{-1} = \Gamma(n)\Gamma(n-1)\pi\sigma^{4n}(1-|\rho|^2)^n$, σ^{jk} is the (j, k)th element of Σ_ξ^{-1} and the

---

Received October 4, 1971; revised April 12, 1972.

[1] *AMS* 1970 *subject classifications*. Primary, 62.10; Secondary, 62.15.

Key words and phrases. Complex correlation coefficient, bivariate complex Gaussian distribution, complex Wishart distribution.

density vanishes outside the region $0 \leqq w \leqq v < \infty$, $|u| \leqq 2[(w/v) - (w/v)^2]^{\frac{1}{2}}$. Although $p(u_R, u_I)$ may be obtained by integrating $p(u_R, u_I, v, w)$ over this region with respect to v and w, it proves more convenient first to introduce the variable t defined by

$$w = (v/2)[1 + t(1 - |u|^2)^{\frac{1}{2}}] .$$

Then $p(u_R, u_I, v, t) = p(u_R, u_I, v, w)|\partial w/\partial t|$ is supported on the cylinder $0 \leqq |u| \leqq 1$, $0 \leqq v < \infty$, $-1 \leqq t \leqq 1$, wherein it assumes the form

$$p(u_R, u_I, v, t) = K_1(1 - t^2)^{n-2}(1 - |u|^2)^{(2n-3)/2}(v/2)^{2n-1}\exp[-v\sigma^{11} + v\,\mathrm{Re}(\sigma^{12}\bar{u})] .$$

Integrating over v and t yields

$$(3) \qquad p(u_R, u_I) = K_2(1 - |u|^2)^{(2n-3)/2}[\sigma^{11} - \mathrm{Re}(\sigma^{12}\bar{u})]^{-2n}, \quad |u| \leqq 1 ,$$

where $K_2 = \Gamma(2n)[\pi^{\frac{1}{2}}\Gamma(n)\Gamma(n - \frac{1}{2})\sigma^{4n}(1 - |\rho|^2)^n 2^{2n-1}]^{-1}$.

Subsequent discussion is facilitated by transforming (3) to polar coordinates. Letting $u = |u|e^{i\theta_u}$ and $\rho = |\rho|e^{i\theta_\rho}$, and using the explicit expressions for σ^{11} and σ^{12} from (1), we obtain

$$(4) \qquad p(|u|, \theta_u) = C|u|(1 - |u|^2)^{(2n-3)/2}[1 - |u|\,|\rho|\cos(\theta_u - \theta_\rho)]^{-2n}$$
$$(0 \leqq |u| \leqq 1, -\pi \leqq \theta_u \leqq \pi) ,$$

where

$$(5) \qquad C = (1 - |\rho|^2)^n\Gamma(2n)[\pi^{\frac{1}{2}}\Gamma(n)\Gamma(n - \tfrac{1}{2})2^{2n-1}]^{-1} .$$

Note that $|u|$ and θ_u are *not* statistically independent. Extensions of certain of the above results to the unequal variances case appear in Goodman (1957).

2. Asymptotic behavior of $E|u_n|^k$. The asymptotic behavior of $E|u_n|^k$ for large n may be obtained as follows. From (4) we have

$$(6) \qquad E|u_n|^k = C\int_{-\pi}^{\pi} d\theta \int_0^1 dx \frac{x^{k+1}}{(1 - x^2)^{\frac{3}{2}}}\left[\frac{1 - x^2}{(1 - x|\rho|\cos\theta)^2}\right]^n .$$

For $|\theta| < \pi/2$ the integrand has a high peak centered at $x = |\rho|\cos\theta$ when n is large, while for $\pi/2 < |\theta| < \pi$ the integrand contains the nth power of a factor that is smaller than 1 for all $0 \leqq x \leqq 1$. Accordingly, the error associated with limiting the range of integration to $|\theta| < \pi/2$ is negligible when n is large. For each such θ the inner integral in (6) is of the form

$$I = \int_0^1 g(x)\exp[-nf(x)]\,dx ,$$

where $g(x) = x^{k+1}(1 - x^2)^{-\frac{3}{2}}$, $f(x) = \log[(1 - bx)^2/(1 - x^2)]$, and $b = |\rho|\cos\theta$. Since $f(x)$ has a minimum at $x = b$ and $f''(b) = 2(1 - b^2)^{-2}$, the method of steepest descent yields $I \sim (\pi/n)^{\frac{1}{2}}b^{k+1}/(1 - b^2)^{n+\frac{1}{2}}$. Therefore

$$(7) \qquad E|u_n|^k \sim \left(\frac{\pi}{n}\right)^{\frac{1}{2}} C\int_{-\pi/2}^{\pi/2}\frac{(|\rho|\cos\theta)^{k+1}}{(1 - |\rho|^2\cos^2\theta)^{n+\frac{1}{2}}}\,d\theta .$$

Since $|\rho| < 1$ in nondegenerate cases, the integrand in (7) peaks sharply at $\theta = 0$

for large n. Steepest descent therefore is applicable again. In this regard we note that, to second order in θ,

$$(1-|\rho|^2\cos^2\theta)^{-n} = \exp[-n\log(1-|\rho|^2\cos^2\theta)] \sim (1-|\rho|^2)^{-n}\exp\left(-\frac{n|\rho|^2\theta^2}{1-|\rho|^2}\right).$$

It follows that

$$E|u_n|^k \sim \left(\frac{\pi}{n}\right)^{\frac{1}{2}} C\frac{|\rho|^{k+1}}{(1-|\rho|^2)^{n+\frac{1}{2}}}\int_{-\infty}^{\infty}\exp\left(-\frac{n|\rho|^2\theta^2}{1-|\rho|^2}\right)d\theta = \frac{\pi C|\rho|^k}{n(1-|\rho|^2)^n}.$$

Substituting for C from (5) yields the desired result,

$$E|u_n|^k \sim \frac{\pi^{\frac{1}{2}}\Gamma(2n)|\rho|^k}{n\Gamma(n)\Gamma(n-\frac{1}{2})2^{2n-1}}. \tag{8}$$

Asymptotic expansion of the gamma functions in (8) verifies the intuitively obvious fact that $E|u_n|^k \to |\rho|^k$ as $n \to \infty$, whereupon the L_r-convergence theorem (Loève (1963)) implies that the $|u_n|$ converge in rth mean to the constant $|\rho|$ for all $r > 0$.

3. Applicability to estimation of spectral parameters. The above results find direct application to the analysis of certain radar estimates of spectral parameters of distributed-velocity media such as storm clouds and clear air turbulence. Specifically, the in-phase and quadrature signals returned from the portion of such a medium that is located at a fixed range from the transmitter may be modeled as the real and imaginary parts, respectively, of a zero mean complex Gaussian random process $\{Z_t\}$. It follows that range gating of the returns from a pair of narrow radar pulses spaced T seconds apart produces a bivariate complex Gaussian random variable ξ. If there is no clutter from ambiguous range cells, then $\xi = (Z_t, Z_{t+T})$. Moreover, if $\{Z_t\}$ is wide-sense stationary, which usually is the case, then Z_t and Z_{t+T} have equal variances. Their correlation matrix Σ_ξ then is of the form (1) with $\sigma^2 = \int dF(\gamma)$ and $\sigma^2\rho = \int e^{i2\pi\gamma T}\,dF(\gamma)$, where F is the spectral distribution function of $\{Z_t\}$. An independent sample $\xi_j = (Z_{t_j}, Z_{t_j+T})$, $1 \leqq j \leqq n$, of such bivariate, equal variance, complex Gaussian random variables may be obtained either by frequency-stepping the radar carrier frequency or by inserting sufficiently long delays between successive pulse pairs. In many applications it is of interest to estimate the centroid γ_0 and spread σ_γ^2 of F, which are defined by $\int(\gamma-\gamma_0)\,dF(\gamma) = 0$ and $\sigma^2\sigma_\gamma^2 = \int(\gamma-\gamma_0)^2\,dF(\gamma)$. Reasonable estimators of these quantities proposed by Rummler (1968a) are expressible as functions of the complex statistic $u = |u|e^{i\theta_u}$, specifically $\hat{\gamma}_0 = (2\pi T)^{-1}\theta_u$ and $\hat{\sigma}_\gamma^2 = (2\pi^2T^2)^{-1}(1-|u|)$. Some properties of these estimators and of certain extensions of them to cases in which $\{Z_t\}$ is corrupted by additive, independent, "white" receiver noise have been explored by Rummler (1968b), Hofstetter (1970), and Berger (1971).

REFERENCES

[1] Berger, T. (1971). Exact distributions and asymptotic moments of pulse-pair estimators. Unpublished tech. memo. TB-110, Raytheon Company, Wayland, Mass.

[2] Goodman, N. R. (1957). On the joint estimation of spectrum, cospectrum, and quadrature spectrum of a two-dimensional stationary Gaussian process. Ph. D. thesis, Princeton Univ.

[3] Goodman, N. R. (1963). Statistical analysis based on a certain multivariate complex Gaussian distribution (an introduction). *Ann. Math. Statist.* **34** 152–177.

[4] Hofstetter, E. M. (1970). Simple estimates of wake velocity parameters. Unpublished technical note 1970–11, Lincoln Laboratory, Lexington, Mass.

[5] Mehta, J. S. and Gurland J. (1969). Some properties and applications of a statistic arising in testing correlation. *Ann. Math. Statist.* **40** 1736–1745.

[6] Loève, M. (1963). *Probability Theory*, (3rd ed.). Van Nostrand, Princeton. 163.

[7] Rummler, W. D. (1968a). Introduction of a new estimator for velocity spectral parameters. Unpublished tech. memo. MM68-4121-5, Bell Telephone Laboratories, Whippany, N.J.

[8] Rummler, W. D. (1968b). Accuracy of spectral width estimators using pulse pair waveforms. Unpublished tech. memo. MM68-4121-14, Bell Telephone Laboratories, Whippany, N.J.

Phillips Hall
Cornell University
Ithaca, New York 14850