Some Angular-Linear Distributions and Related Regression Models

RICHARD A. JOHNSON and THOMAS E. WEHRLY*

Parametric models are proposed for the joint distribution of bivariate random variables when one variable is directional and one is scalar. These distributions are developed on the basis of the maximum entropy principle and by the specification of the marginal distributions. The properties of these distributions and the statistical analysis of regression models based on these distributions are explored. One model is extended to several variables in a form that justifies the use of least squares for estimation of parameters, conditional on the observed angles.

KEY WORDS: Directional data; Angular-linear distribution; Regression; Trigonometric regression; Entropy.

1. INTRODUCTION

Researchers are sometimes confronted with bivariate data, one component of which is an angle and the other a real number. For instance, the wind direction and the level of a pollutant may occur in environmental studies, such as Ohta, Marita, and Mizoguchi (1976) and De Wiest and Della Fiorentina (1975). There is an extensive literature on directional data, including the monograph by Mardia (1972), but there have been only limited attempts to handle such angular-linear data. We attempt to fill this void by presenting some new models for angular-linear distributions based on the principle of maximum entropy in Section 2, and on the property of having specified marginal distributions and on wrapping one of a pair of random variables in Section 3. We have tried to develop models that are useful for statistical inference and, conveniently, the entropy arguments lead to nice exponential families. Some of these are angular-linear generalizations of the von Mises distribution, which is the most useful distribution for statistical inference of angular data. We denote by $\mathrm{VM}(\mu, \kappa)$ the von Mises distribution with density

$$f(\theta) = [2\pi I_0(\kappa)]^{-1} \exp\{\kappa \cos(\theta - \mu)\} ,$$

where $0 \le \theta < 2\pi$, $\kappa \ge 0$, $0 \le \mu < 2\pi$, and $I_0(\kappa)$ is the modified Bessel function of the first kind and order zero.

Our primary interest is in regression models, which we emphasize in Sections 4 and 5. Some potentially important new models for multiple regression are studied in the latter section, and a numerical example is presented. A somewhat different approach designed for obtaining measures of dependence for angular-linear distributions appears in Johnson and Wehrly (1977).

2. MAXIMUM ENTROPY DISTRIBUTIONS

In this section, we use information theory concepts to find new angular-linear distributions which maximize the entropy subject to constraints on certain moments. Four such families of distributions with interesting properties are discussed. The first family is presented in the following theorem.

Theorem 1: The density function of (Θ, X) given by

$$f(\theta, x) = (\lambda^2 - \kappa^2)^{\frac{1}{2}}(2\pi)^{-1} \cdot \exp\{-\lambda x + \kappa x \cos(\theta - \mu)\} , \quad (2.1)$$

where $0 \le \theta < 2\pi$, $x > 0$, $0 < \kappa < \lambda$, and $0 \le \mu < 2\pi$, is the maximum entropy distribution subject to $E(X)$, $E(X \cos \Theta)$, and $E(X \sin \Theta)$ taking specified values which are consistent with expectation with respect to the distribution (2.1).

Proof: First we show that (2.1) is a density. Obviously, $f(\theta, x) \ge 0$. Also,

$$\int_0^{2\pi} \int_0^\infty f(\theta, x) dx d\theta$$
$$= \frac{(\lambda^2 - \kappa^2)^{\frac{1}{2}}}{2\pi} \int_0^{2\pi} \int_0^\infty \exp\{-(\lambda - \kappa \cos(\theta - \mu))x\} dx d\theta$$
$$= \frac{(\lambda^2 - \kappa^2)^{\frac{1}{2}}}{2\pi} \int_0^{2\pi} [\lambda - \kappa \cos(\theta - \mu)]^{-1} d\theta$$
$$= \frac{1}{2\pi} \int_0^{2\pi} \frac{1 - \rho^2}{1 + \rho^2 - 2\rho \cos(\theta - \mu)} d\theta = 1 ,$$

where $\rho = \kappa[\lambda + (\lambda^2 - \kappa^2)^{\frac{1}{2}}]^{-1}$. The last equality holds since $f(\theta) = (2\pi)^{-1}(1 - \rho^2)(1 + \rho^2 - 2\rho \cos(\theta - \mu))^{-1}$ is the density of the wrapped Cauchy distribution, which is defined by letting $\Theta = Z \pmod{2\pi}$, where Z has the Cauchy distribution with density $f(z) = \alpha[\pi(\alpha^2 + z^2)]^{-1}$ and $\alpha = -\ln \rho$.

Next, we define a_1, a_2, a_3 by

$$E(X) = a_1 , \quad E(X \cos \Theta) = a_2 , \quad E(X \sin \Theta) = a_3 , \quad (2.2)$$

where the expectations are taken with respect to the

* Richard A. Johnson is Professor, Department of Statistics, University of Wisconsin, Madison, WI 53706. Thomas E. Wehrly is Assistant Professor, Institute of Statistics, Texas A&M University, College Station, TX 77843. This research was supported by the National Aeronautics and Space Administration under Grant NSG-1196. The authors gratefully acknowledge helpful comments from G.K. Bhattacharyya, the Editor, and the referees.

© Journal of the American Statistical Association
September 1978, Volume 73, Number 363
Theory and Methods Section

distribution (2.1). Then by standard arguments (see, e.g., Mardia 1975, p. 352), the density (2.1) maximizes the entropy over all angular-linear distributions for which (2.2) holds.

Remark: The density function (2.1) has marginal density functions

$$f_1(\theta) = \frac{(\lambda^2 - \kappa^2)^{\frac{1}{2}}}{2\pi} \frac{1}{\lambda - \kappa \cos(\theta - \mu)} = \frac{1}{2\pi} \frac{1 - \rho^2}{1 + \rho^2 - 2\rho \cos(\theta - \mu)} \quad (2.3)$$

and

$$f_2(x) = (\lambda^2 - \kappa^2)^{\frac{1}{2}} I_0(\kappa x) e^{-\lambda x} , \quad (2.4)$$

where $\rho = \kappa[\lambda + (\lambda^2 - \kappa^2)^{\frac{1}{2}}]^{-1}$ and $I_0(\cdot)$ is the modified Bessel function of the first kind and order zero.

The conditional distribution functions obtained from (2.1) are

$$g_1(\theta | x) = [2\pi I_0(\kappa x)]^{-1} \exp\{\kappa x \cos(\theta - \mu)\} \quad (2.5)$$

and

$$g_2(x | \theta) = [\lambda - \kappa \cos(\theta - \mu)] \cdot \exp\{-[\lambda - \kappa \cos(\theta - \mu)]x\} . \quad (2.6)$$

The density (2.1) has several interesting properties. The marginal distribution (2.3) of Θ is the wrapped Cauchy distribution defined above, while the marginal distribution (2.4) of X does not follow a familiar distribution. The conditional density (2.5) of Θ given x is a von Mises density VM$(\mu, \kappa x)$. The conditional density of (2.6) of X given θ is an exponential density with mean $(\lambda - \kappa \cos(\theta - \mu))^{-1}$.

Another characterization using the second moment of the linear variable produces a different joint distribution. The proof is similar to that of Theorem 1 and is therefore omitted.

Theorem 2: Let (Θ, X) have the joint density

$$f(\theta, x) = c \cdot \exp\left\{-\frac{x^2}{2\sigma^2} + \frac{\lambda x}{\sigma^2} + \frac{\kappa x}{\sigma^2} \cos(\theta - \mu)\right\} , \quad (2.7)$$

where $c > 0$ is a constant of integration, $-\infty < x < \infty$, $0 \le \theta < 2\pi$, $-\infty < \lambda < \infty$, $\kappa > 0$, and $0 \le \mu < 2\pi$. Then $f(\theta, x)$ is the maximum entropy angular-linear distribution subject to $E(X)$, $E(X^2)$, $E(X \cos \Theta)$, and $E(X \sin \Theta)$ taking specified values consistent with expectation with respect to the distribution (2.7).

Remark: The marginal distributions are not of familiar form, but the conditional distributions have densities

$$g_1(\theta | x) = (2\pi I_0(\kappa x/\sigma^2))^{-1} \exp\{(\kappa x/\sigma^2) \cos(\theta - \mu)\} \quad (2.8)$$

and

$$g_2(x | \theta) = (2\pi\sigma^2)^{-\frac{1}{2}} \cdot \exp\{-(1/2\sigma^2)[x - (\lambda + \kappa \cos(\theta - \mu))]^2\} . \quad (2.9)$$

The density (2.8) is a von Mises density VM$(\mu, \kappa x/\sigma^2)$. The density (2.9) is normal with mean $\lambda + \kappa(\cos(\theta - \mu))$ and variance σ^2.

A major limitation of the two previous densities is that if X and Θ are independent, then Θ is forced to be uniformly distributed on the circle. The following density has von Mises and exponential marginal distributions when X and Θ are independent, but otherwise the marginal and conditional distributions are not as tractable as those for the density (2.1).

Theorem 3: Let (Θ, X) have the joint density

$$f(\theta, x) = c \exp\{-\lambda x + \kappa x \cos(\theta - \mu_1) + \nu \cos(\theta - \mu_2)\} , \quad (2.10)$$

where $0 \le \theta < 2\pi$, $0 < x < \infty$, $c = (\lambda^2 - \kappa^2)^{\frac{1}{2}}(2\pi)^{-1} \cdot \{I_0(\nu) + 2\sum_{p=1}^{\infty} \rho^p I_p(\nu) \cos[p(\mu_1 - \mu_2)]\}^{-1}$, $\lambda > \kappa > 0$, $0 \le \mu_1, \mu_2 < 2\pi$, $\rho = \kappa[\lambda + (\lambda^2 - \kappa^2)^{\frac{1}{2}}]^{-1}$, and $I_p(\cdot)$ is the modified Bessel function of the first kind and order p.

The joint density (2.10) is a maximum entropy density subject to $E(X)$, $E(\cos \Theta)$, $E(\sin \Theta)$, $E(X \cos \Theta)$, and $E(X \sin \Theta)$ taking specified values consistent with expectation with respect to the distribution (2.10).

Remark: The distribution (2.10) has conditional density functions

$$g_1(\theta | x) = [2\pi I_0(\kappa^*)]^{-1} \exp\{\kappa^* \cos(\theta - \mu^*)\} \quad (2.11)$$

and

$$g_2(x | \theta) = [\lambda - \kappa \cos(\theta - \mu_1)] \cdot \exp\{-[\lambda - \kappa \cos(\theta - \mu_1)]x\} , \quad (2.12)$$

where κ^* and μ^* are the solutions to

$$\kappa^* \cos \mu^* = \kappa x \cos \mu_1 + \nu \cos \mu_2 ,$$
$$\kappa^* \sin \mu^* = \kappa x \sin \mu_1 + \nu \sin \mu_2 .$$

The conditional densities (2.11) and (2.12) are von Mises and exponential densities, respectively. For this bivariate distribution, the independence of Θ and X is equivalent to the parameter κ being equal to zero. If $\kappa = 0$, then Θ and X are independent von Mises and exponential random variables, respectively. By applying a similar argument, we could also obtain a maximum entropy distribution which has von Mises and normal conditional distributions.

Mardia and Sutton (1976) construct an angular-linear distribution by conditioning a trivariate normal distribution. This distribution has the property that the distributions of X and Θ are normal and von Mises, respectively, when X and Θ are independent. The dependence is somewhat different from the corresponding model suggested above using entropy arguments.

The previous arguments for obtaining maximum entropy distributions can be extended to obtain distributions for several angular and linear variables, which will be useful in providing a model for trigonometric regression. Let $\mathbf{\Theta}' = (\Theta_1, \ldots, \Theta_p)$, $\mathbf{X}' = (X_1, \ldots, X_q)$, and

$$\mathbf{H}(\mathbf{\Theta}) = \begin{pmatrix} \cos\Theta_1 & \dots & \cos n\Theta_1 & \sin\Theta_1 & \dots & \sin n\Theta_1 \\ \vdots & & \vdots & \vdots & & \vdots \\ \cos\Theta_p & \dots & \cos n\Theta_p & \sin\Theta_p & \dots & \sin n\Theta_p \end{pmatrix}.$$

Theorem 4: Let $\mathbf{\Theta}$ and $\mathbf{X}$ have the joint density function

$$f(\boldsymbol{\theta}, \mathbf{x}) = c\cdot\exp\{-\tfrac{1}{2}\mathbf{x}'\Sigma^{-1}\mathbf{x} + \boldsymbol{\lambda}'\Sigma^{-1}\mathbf{x} + \mathbf{a}(\boldsymbol{\theta})'\Sigma^{-1}\mathbf{x}\}\ , \quad (2.13)$$

where c is a constant of integration, $\mathbf{a}(\boldsymbol{\theta})' = (a_1(\boldsymbol{\theta}), \dots, a_q(\boldsymbol{\theta}))$,

$$a_i(\boldsymbol{\theta}) = \sum_{j=1}^{p}\sum_{k=1}^{n} a_{ijk}\cos[k(\theta_j - \mu_{ijk})]$$
$$= \sum_{j=1}^{p}\sum_{k=1}^{n}[\alpha_{ijk}\cos(k\theta_j) + \beta_{ijk}\sin(k\theta_j)]\ ,$$
$$i = 1, \dots, q\ , \quad (2.14)$$

$\mathbf{x} \in R^q$, $\boldsymbol{\theta} \in [0, 2\pi)^p$, and Σ^{-1} is positive definite. Then $f(\boldsymbol{\theta}, \mathbf{x})$ maximizes the entropy of multivariate angular-linear distributions subject to $E[\mathbf{X}\ \mathbf{X}']$, $E(\mathbf{X})$, and $E[\mathbf{X} \otimes \mathbf{H}(\mathbf{\Theta})]$, where $\otimes$ is the Kronecker product, taking specified values consistent with expectation with respect to the distribution (2.13).

Remark: The conditional distribution of $\mathbf{X}$ given $\mathbf{\Theta} = \boldsymbol{\theta}$ is q-dimensional multivariate normal with mean $\boldsymbol{\lambda} + \mathbf{a}(\boldsymbol{\theta})$ and covariance matrix Σ.

The distribution (2.13) is a member of an exponential family with sufficient statistic $\{X_i,\ X_iX_j,\ X_i\cos m\Theta_k,\ X_i\sin m\Theta_k;\ 1 \le i \le j \le q,\ 1 \le k \le p,\ 1 \le m \le n\}$, and parameters which take all values in a $[q + \frac{1}{2}q(q+1) + 2pqn]$-dimensional rectangle. Hence we can apply the usual theory for complete exponential families (cf. Lehmann 1959) to find optimal tests for a particular (natural) parameter in the joint distribution.

3. OTHER ANGULAR-LINEAR DISTRIBUTIONS

We now present a method of obtaining angular-linear distributions with specified marginal distributions.

Theorem 5: Let $f_1(\theta)$ be a density on the circle and $f_2(x)$ be a density on the line with cumulative distribution functions (cdf's) $F_1(\theta)$ and $F_2(x)$, respectively, where F_1 is with respect to a fixed, arbitrary origin. Let $g(\cdot)$ be a density on the circle. Then

$$f(\theta, x) = 2\pi g[2\pi(F_1(\theta) - F_2(x))]f_1(\theta)f_2(x)\ , \quad (3.1)$$

where $0 \le \theta < 2\pi$, $-\infty < x < \infty$, is a density for an angular-linear distribution having the specified marginal densities $f_1(\theta)$ and $f_2(x)$.

Proof: Make the change of variable $U = 2\pi F_1(\Theta)$ and integrate with respect to U to find the marginal density of X. Similarly, let $V = 2\pi F_2(X)$ to find the marginal density of Θ.

Example 1: Circular Uniform and Normal Marginal Distributions. Set $f_1(\theta) = (2\pi)^{-1}$, $0 \le \theta < 2\pi$, $f_2(x) = \phi(x)$, where $\phi(x)$ is the standard normal density, and $g(\xi) = (2\pi I_0(\kappa))^{-1}\exp\{\kappa\cos(\xi - \mu)\}$. Then by Theorem 5, the density

$$f(\theta, x) = (2\pi I_0(\kappa))^{-1}\cdot\exp\{\kappa\cos(\theta - 2\pi\Phi(x) - \mu)\}\phi(x)\ , \quad (3.2)$$

where $\Phi(x)$ is the standard normal cdf, has the circular uniform distribution and the normal distribution as marginal distributions.

Example 2: von Mises and Normal Marginal Distributions. Set $f_1(\theta) = (2\pi I_0(\kappa_1))^{-1}\exp\{\kappa_1\cos(\theta - \mu_1)\}$, $0 \le \theta < 2\pi$, and let $f_2(x)$ and $g(\xi)$ be as defined in Example 1. Then by Theorem 5, the density

$$f(\theta, x) = (I_0(\kappa))^{-1}\cdot\exp\{\kappa\cos[2\pi(F_1(\theta) - \Phi(x)) - \mu]\}f_1(\theta)\phi(x)\ , \quad (3.3)$$

where $F_1(\theta) = \int_0^\theta f_1(\xi)d\xi$, has the von Mises distribution and the normal distribution as marginal distributions.

The following theorem shows that the families of distributions given in the above examples can also be characterized as maximum entropy distributions. The proof follows immediately from a standard argument (see, e.g., Mardia 1975, p. 352).

Theorem 6:

(a) The distribution with density (3.2) maximizes the entropy of angular-linear distributions subject to

$$E(X) = 0\ ,\quad E(X^2) = 1\ ,$$
$$E\{\cos[\Theta - 2\pi\Phi(X)]\} = A(\kappa)\cos\mu\ ,$$
$$E\{\sin[\Theta - 2\pi\Phi(X)]\} = A(\kappa)\sin\mu\ ,$$

where $A(\kappa) = I_1(\kappa)/I_0(\kappa)$.

(b) The distribution with density (3.3) maximizes the entropy of angular-linear distributions subject to

$$E(X) = 0\ ,\quad E(X^2) = 1\ ,$$
$$E(\cos\Theta) = A(\kappa_1)\cos\mu_1\ ,\quad E(\sin\Theta) = A(\kappa_1)\sin\mu_1,$$
$$E\{\cos[2\pi(F_1(\Theta) - \Phi(X))]\} = A(\kappa)\cos\mu\ ,$$
$$E\{\sin[2\pi(F_1(\Theta) - \Phi(X))]\} = A(\kappa)\sin\mu\ .$$

We conclude this section by noting that another method of forming angular-linear distributions from bivariate linear distributions is to wrap one of the linear variables around the circle; that is, if (Y_1, Y_2) has a bivariate distribution, define $\Theta = Y_1 \pmod{2\pi}$, $X = Y_2$. Then (Θ, X) has an angular-linear distribution. Statistical inference for the singly wrapped normal distribution is discussed in Johnson and Wehrly (1977).

4. THE REGRESSION OF AN ANGULAR VARIATE ON A LINEAR VARIATE

Before illustrating the versatility of our models, we point out some serious drawbacks of one regression model introduced by Gould (1969). He introduced a regression analysis procedure for angular variates which corresponds to a multiple regression analysis for normally distributed variates, but for simplicity we discuss the simple regression model and its shortcomings. For this model, $\Theta_1, \ldots, \Theta_n$ are independently distributed with $\text{VM}(\mu_0 + \beta x_i, \kappa)$ distributions, $i = 1, \ldots, n$, respectively, where $x_1, \ldots, x_n$ are known concomitant variables, and μ_0, β, and κ are unknown parameters. Gould finds the maximum likelihood estimators of these parameters from the log-likelihood,

$$-n \ln 2\pi + \kappa \sum_{i=1}^{n} \cos(\theta_i - \mu_0 - \beta x_i) - n \ln I_0(\kappa) ,$$

using a straightforward iterative procedure, and develops an approximate test for $\beta = 0$. For this model, maximum likelihood estimation coincides with least squares estimation.

The most serious drawback to Gould's approach is that the likelihood function has infinitely many equally large peaks. If the x_i's are equally spaced with δ the space between adjacent x_i's, then the likelihood function is periodic with period $2\pi/\delta$. For several different samples, we set the parameter μ_0 equal to zero and plotted the quantity $\sum_{i=1}^{n} \cos(\theta_i - \beta x_i)$ as a function of β using several choices of the concomitant variables x_i. The plots revealed global maxima occurring every $2\pi/\delta$ with several local maxima within each period, so that an iteration procedure for obtaining a maximum likelihood estimator could lead to any of the peaks.

As an alternative to the Gould approach, we introduce two examples of new models for regression. The first has the concomitant variables X_i, $i = 1, \ldots, n$, conditionally centering the Θ_i's. In the second, the X_i's serve as scale parameters in the conditional distributions of Θ_i given X_i.

Example 3: Conditional Centering. Let (Θ_i, X_i), $i = 1, \ldots, n$, have the distribution (3.1) specialized to the case where Θ is uniform,

$$g(\xi) = [2\pi I_0(\kappa)]^{-1} \exp\{\kappa \cos(\xi - \mu)\}$$

is the von Mises density, and $f_2(x) = f(x)$ is a known, completely specified density with cdf F. The conditional distribution of $\mathbf{\Theta}$ given $\mathbf{X} = \mathbf{x}$ is

$$f_1(\boldsymbol{\theta} | \mathbf{x}) = (2\pi I_0(\kappa))^{-n} \exp\{\kappa \sum_{i=1}^{n} \cos(\theta_i - 2\pi F(x_i) - \mu)\} .$$

For this distribution the conditional maximum likelihood estimates of μ and κ are $\hat{\mu} = \bar{x}_0$ and $\hat{\kappa} = A^{-1}(\bar{R})$, where

$$\bar{R} \cos \bar{x}_0 = (1/n) \sum_{i=1}^{n} \cos(\theta_i - 2\pi F(x_i))$$

and

$$\bar{R} \sin \bar{x}_0 = (1/n) \sum_{i=1}^{n} \sin(\theta_i - 2\pi F(x_i)) ,$$

$$A(t) = I_1(t)/I_0(t) ,$$

and $I_p(t)$ is the modified Bessel function of the first kind and order p. As in Gould's model, the X_i's serve to center the Θ_i's conditionally. However, this model is not a multiplicative model in the parameters, so the centering values for the Θ's do not progress many times around the circle as X increases. Hence the associated problems in parameter estimation are eliminated.

Example 4: Conditional von Mises Distribution. Both conditional von Mises distributions (2.5) and (2.8) have similar forms with the X_i's serving as scale parameters. The conditional distribution of Θ_i becomes more concentrated as X_i becomes larger. Limiting our consideration to (2.5), the conditional density of $\mathbf{\Theta}$ given $\mathbf{X} = \mathbf{x}$ is

$$f(\boldsymbol{\theta} | \mathbf{x}) = (2\pi)^{-n} (\prod_{i=1}^{n} I_0(\kappa x_i))^{-1} \exp\{\kappa \sum_{i=1}^{n} x_i \cos(\theta_i - \mu)\} .$$

The maximum likelihood estimates are $\hat{\mu} = \theta^*$ and $\hat{\kappa}$, where $\hat{\kappa}$, θ^*, and R^* satisfy $\sum_{i=1}^{n} x_i A(\kappa x_i) = R^*$, and

$$R^* \cos \theta^* = \sum_{i=1}^{n} x_i \cos \theta_i , \quad R^* \sin \theta^* = \sum_{i=1}^{n} x_i \sin \theta_i .$$

This solution is unique since $\sum_{i=1}^{n} x_i A(\kappa x_i)$ is a strictly increasing function of κ, taking all values on the interval $[0, \sum_{i=1}^{n} x_i]$.

5. REGRESSION OF LINEAR VARIATES ON OTHER LINEAR AND ANGULAR VARIATES

In this section we apply the distribution (2.13) as a population model for trigonometric regression. By looking at conditional distributions, we can use (2.13) as the population model for the regression of linear variates on other linear and angular variates.

Let $(\mathbf{\Theta}, \mathbf{X})$ have the joint density (2.13). We view $\mathbf{\Theta}$ as a vector of concomitant variables. The conditional distribution of $\mathbf{X}$ given $\mathbf{\Theta} = \boldsymbol{\theta}$ is q-variate normal with mean vector $\boldsymbol{\lambda} + \mathbf{a}(\boldsymbol{\theta})$ and covariance matrix Σ where $\boldsymbol{\lambda} \in R^q$, $\mathbf{a}(\boldsymbol{\theta})$ is defined by (2.14), and Σ is positive definite. This provides the usual model for multivariate trigonometric regression. Thus the full power of the multivariate multiple regression model can be brought to bear to estimate the λ's, α's, and β's. The residuals can be used to check the fit of the model. Results on optimal design for trigonometric regression (cf. Fedorov 1972, p. 94 and Laycock 1975) also apply to this conditional model.

An alternative to exact conditional inference would be maximum likelihood estimation using the joint distribution (2.13). The calculation of estimates would, in general, be iterative and the conditional least squares estimates should provide good starting values. However, exact distribution theory for the estimators would undoubtedly be difficult, and any inferences could be based on the limiting normal distribution for the maximum likelihood estimators.

The density (2.13) also provides a means for predicting $\mathbf{X}_1$ from $\mathbf{X}_2$ and $\mathbf{\Theta}$ by writing $\mathbf{X} = (\mathbf{X}_1' | \mathbf{X}_2')'$ and looking at $f(\mathbf{x}_1 | \mathbf{x}_2, \boldsymbol{\theta})$. Results for the conditional distribution of $\mathbf{X}_1$ given $\mathbf{x}_2$ and $\boldsymbol{\theta}$ follow from the usual results for the conditioning of one multivariate normal vector on another. If we partition Σ, $\boldsymbol{\lambda}$, and $\mathbf{a}(\boldsymbol{\theta})$ correspondingly, we find that the distribution of $\mathbf{X}_1 = (X_1, \ldots, X_r)'$

given $\mathbf{x}_2$ and $\boldsymbol{\theta}$ is the r-dimensional normal distribution with mean $\boldsymbol{\lambda}_1 + \Sigma_{12}\Sigma_{22}^{-1}[\mathbf{x}_2 - (\boldsymbol{\lambda}_2 + \mathbf{a}_2(\boldsymbol{\theta}))]$ and covariance matrix $\Sigma_{11} - \Sigma_{12}\Sigma_{22}^{-1}\Sigma_{21}$.

Note that each component X_i, $i = 1, \ldots, r$, of $\mathbf{X}_1$ has a variance not depending on the conditioning variables and a mean of the form

$$\nu_0 + \sum_{i=r+1}^{q} \nu_i x_i + \sum_{i=r+1}^{q} \sum_{j=1}^{p} \sum_{k=1}^{n} [\gamma_{ijk} \cos (k\theta_j) + \delta_{ijk} \sin (k\theta_j)] \ . \quad (5.1)$$

Consequently, this model leads to a natural method of predicting a vector $\mathbf{X}$ from a vector $\boldsymbol{\theta}$ of directions. Here we fit the best Fourier series of nth degree in the individual $\boldsymbol{\theta}$'s to the $\mathbf{X}$'s. As shown, some X_i's may also be incorporated into the set of predictor variables.

Example 5: De Wiest and Della Fiorentina (1975) proposed a new air quality index. This index, along with the temperature, wind direction, and wind speed, was measured on several days. Here we wish to predict the air quality index as a function of the other variables. The data are given in the table.

Air Quality Index, Temperature, and Wind Direction

X_1 *A%-measure of pollution*	X_2 *Temperature (°C)*	θ *Wind direction (degrees)*
0.70	9.5	90
0.75	6.5	158
0.96	5.5	135
0.32	3.5	45
0.32	7.2	45
0.79	5.9	135
0.61	7.9	135
0.47	8.5	45
1.06	7.7	90
0.42	7.2	45
0.26	2.0	270
0.37	5.0	225
0.14	4.0	270
0.23	6.2	270
0.74	10.0	0
0.47	10.5	225

We apply the conditional normal distribution as the conditional distribution of X_1 given x_2 and θ. A least squares regression is used to estimate the parameters for the mean given in (5.1). The resulting regression equation is

$$\hat{x}_1 = 0.306 + 0.028x_2 - 0.179 \cos \theta + 0.216 \sin \theta \ .$$

Standard variable selection techniques may then be used to remove (conditionally) nonsignificant terms in the regression. Initially wind speed was also included in the regression, but it was removed because its contribution was (conditionally) highly nonsignificant.

6. CONCLUSIONS

The main emphasis of this article has been the development of new parametric models for angular-linear distributions and the use of the conditional distributions for regression. However, the joint distributions presented in Sections 2 and 3 also lend themselves to the development of parametric tests for independence of angular variates and linear variates. In particular, the usual theory of likelihood ratio tests can be applied to any of these distributions to find tests for independence.

[*Received September 1976. Revised January 1978.*]

REFERENCES

De Wiest, F., and Della Fiorentina, H. (1975), "Suggestions for a Realistic Definition of an Air Quality Index Relative to Hydrocarbonaceous Matter Associated with Airborne Particles," *Atmospheric Environment*, 951–954.

Fedorov, V.V. (1972), *Theory of Optimal Experiments*, New York: Academic Press.

Gould, A. Lawrence (1969), "A Regression Technique for Angular Variates," *Biometrics*, 25, 683–700.

Johnson, Richard A., and Wehrly, Thomas (1977), "Measures and Models for Angular Correlation and Angular-Linear Correlation," *Journal of the Royal Statistical Society*, Ser. B, 39, 222–229.

Laycock, P.J. (1975), "Optimal Design: Regression Models for Directions," *Biometrika*, 62, 305–311.

Lehmann, E.L. (1959), *Testing Statistical Hypotheses*, New York: John Wiley & Sons.

Mardia, K.V. (1972), *Statistics of Directional Data*, New York: Academic Press.

——— (1975), "Statistics of Directional Data," *Journal of the Royal Statistical Society*, Ser. B, 37, 349–393 (with discussion).

———, and Sutton, T.W. (1976), "A Model for Cylindrical Variables with Applications," Research Report No. 9, Department of Statistics, The University of Leeds, England.

Ohta, T., Marita, M., and Mizoguchi, I. (1976), "Local Distribution of Chlorinated Hydrocarbons in the Ambient Air in Tokyo," *Atmospheric Environment*, 10, 557–560.