Journal of Multivariate Analysis **68**, 283–298 (1999)
Article ID jmva.1998.1796, available online at http://www.idealibrary.com on IDEAL®

Characterization Theorems when Variables Are Measured with Error

John P. Holcomb, Jr.

Youngstown State University
E-mail: jholcomb@math.ysu.edu

Received November 13, 1997; revised July 23, 1998

Linear regression models are studied when variables of interest are observed in the presence of measurement error. Techniques involving Fourier transforms that lead to simple differential equations with unique solutions are used in the context

One characterization involves the Fisher score of the observed vector. A second characterization involves the Hessian matrix of the observed density. © 1999 Academic Press

AMS 1991 subject classifications: 62J05, 62H05.

Key words and phrases: Measurement error model; conditional expectation; conditional variance.

1. INTRODUCTION

The standard linear regression model with one independent random variable is defined by

$$y_i = \alpha + \beta x_i + \varepsilon_i, \qquad i = 1, 2, ..., n \tag{1.1}$$

where ε_i are independent, and $N(0, \sigma_\varepsilon^2)$. In practice, the measured x_i contain errors. The traditional additive model has w_i consisting of two independent components: x_i, the variable of interest, and u_i, the measurement error of x_i, i.e., $W = X + U$.

Measurement error in linear regression has a long history that includes Adcock (1877, 1878), Lindley (1947), Kendall (1951, 1952), and Rao (1947, 1975). Cochran (1968), Fuller (1987), and Carroll, Ruppert, and Stefanski (1995) provide a review of measurement error models.

The work of Lindley and Rao deals with the characterization of the error-free variable X and the measurement, error U. Their techniques involve the use of Fourier transforms that lead to unique solutions of

0047-259X/99 $30.00

differential equations involving the characteristic functions of X and U. Kendall also used similar techniques that are summarized in Kendall, Stuart, and Ord (1979). Kagan (1983) utilizes similar techniques to prove necessary and sufficient conditions for the characterization of both X and U. This paper does not follow the measurement-error paradigm, but a Bayesian one. In the Kagan paper, X represents a location parameter and U represents some type of independent error. The equation for the characteristic functions obtained by Kagan are the same as equations obtained by Laha and Lukacs (1960) and Morris (1982).

Further work in understanding the effect of measurement error as it relates to conditional expectation and variance leads to an extension of an observation made by Stefanski and Carroll (1990). They state that if f_W denotes the density of W and U is normally distributed, then

$$E[U \mid (W=w)] = -\sigma_U^2 \frac{f'_W(w)}{f_W(w)}. \tag{1.2}$$

A substantial amount of literature exists in constructing estimates of the Fisher score, i.e., $f'_W(x)/f_W(w)$. This includes Stefanski and Carroll (1985, 1991), Whittemore (1989), Rosner, Willett, and Spiegelman (1989), Carroll and Stefanski (1990), and Gleser (1990).

This paper shows that if Eq. (1.2) holds for some σ, then the measurement error is characterized as normal. It can be shown that the error can be characterized as normal if Eq. (1.2) holds for all $\sigma > 0$ and $w \in \Re$. Stefanski, in personal correspondence with the author, showed this by first writing $W = X + \sigma U$, where U has mean 0 and variance 1. Utilizing the assumption that X and U are independent, the basic identity can be written in terms of the densities of X and U. The resulting expression depends on w and σ. After some manipulation, the limit of w to infinity and σ to infinity can be taken in such a way that the ratio w/σ approaches t. Then, by dominated convergence type arguments, the differential equation

$$g'(t) = -tg(t)$$

is obtained where g is the density function of U. The standard normal density function is the unique solution to this differential equation.

The relaxation of the condition that Eq. (1.2) must hold for all σ allows investigators to use the results of previous work in estimating $f_W^{(1)}/f_W$ to characterize the distribution of the measurement error. Thus when Eq. (1.2) is seen to hold, the measurement error distribution is characterized as normal. The result also provides insight into the nature of the conditional expectation in the multivariate setting. In fact, Section 2 proves a generalized version of Eq. (1.2), i.e., if $\mathbf{X}$ is a random vector with $\mathbf{U}$ the associated vector of independent measurement errors, then under general conditions

$$E[\mathbf{U} \mid (\mathbf{W}=\mathbf{w})] = -\boldsymbol{\Sigma}\frac{\nabla f_{\mathbf{W}}(\mathbf{w})}{f_{\mathbf{W}}(\mathbf{w})} \qquad \text{iff} \quad \mathbf{U} \sim MVN(\mathbf{0}, \boldsymbol{\Sigma}), \tag{1.3}$$

where $\nabla f_{\mathbf{W}}(\mathbf{w})$ denotes the gradient vector of $f_{\mathbf{W}}$ evaluated at $\mathbf{w}$. This proof utilizes the techniques of Lindley, Rao, and Kagan, which apply Fourier transforms to obtain a differential equation involving the characteristic function of $\mathbf{U}$. It is then shown that the characteristic function of a multivariate normal distribution is the unique solution of the differential equation.

Similar techniques used to prove the previous theorem are adapted to show that the second moment matrix also characterizes the measurement error as multivariate normal. Spiegelman (1986) examined conditional variance when variables are measured with error. Under general conditions, Section 3 proves the following:

$$E[\mathbf{U}\mathbf{U}' \mid (\mathbf{W}=\mathbf{w})] = \boldsymbol{\Sigma}\frac{\mathbf{H}_{f_{\mathbf{W}}}(\mathbf{w})}{f_{\mathbf{W}}(\mathbf{w})}\boldsymbol{\Sigma} + \boldsymbol{\Sigma}, \qquad \text{iff} \quad \mathbf{U} \sim MVN(\mathbf{0}, \boldsymbol{\Sigma}),$$

where $\mathbf{H}_{f_{\mathbf{W}}}(\mathbf{w})$ denotes the Hessian matrix of $f_{\mathbf{W}}$ evaluated at $\mathbf{w}$.

2. THE FIRST MOMENT VECTOR

Let f_W denote the density of W and f'_W denote its derivative. Assume U is independent of X and U has a second moment. Stefanski and Carroll (1990) state

$$E[X \mid (W=w)] = w + \sigma_U^2 \frac{f'_W(w)}{f_W(w)} + O(\sigma_U^4).$$

They also state that if $U \sim N(0, \sigma_U^2)$, then $O(\sigma_U^4) = 0$.

Under reasonable assumptions it will be shown that

$$E[X \mid (W=w)] = w + \sigma_U^2 \frac{f'_W(w)}{f_W(w)}$$

is a necessary and sufficient condition for $U \sim N(0, \sigma_U^2)$. In fact, the result will be generalized to n-dimensional random vectors $\mathbf{X}$ and $\mathbf{U}$. The following notation is used throughout this section.

Let $\mathbf{X}$ represent an n-dimensional absolutely continuous random vector with density function $f_{\mathbf{X}}$. Let $\mathbf{U}$ represent independent, absolutely continuous measurement error for $\mathbf{X}$ with density function g. The density of $\mathbf{W}$ is again denoted as f_W. Since $\mathbf{X}$ and $\mathbf{U}$ are independent, $\mathbf{W} = \mathbf{X} + \mathbf{U}$ has density

$$f_{\mathbf{W}}(\mathbf{w}) = \int_{-\infty}^{\infty} g(\mathbf{w} - \mathbf{x})\, f_{\mathbf{X}}(\mathbf{x})\, d\mathbf{x}.$$

The gradient vector of a density function will be denoted with ∇ and $\mathbf{H}$ will denote its Hessian matrix. Thus

$$[\mathbf{H}_{f_{\mathbf{W}}}(\mathbf{w})]_{j,k} = \frac{\partial^2 f_{\mathbf{W}}(\mathbf{w})}{\partial w_k\, \partial w_j}.$$

Lastly, let $\mathbf{w}_{[-k]}$ represent the vector $\mathbf{w}$ omitting the kth component and let $\mathbf{w}_{[-j,-k]}$ represent the vector $\mathbf{w}$ omitting the j and k components.

The following lemmas are necessary to prove the theorem where it is assumed that $\mathbf{X}$ and $\mathbf{U}$ are jointly absolutely continuous. The first two proofs are left to the reader.

LEMMA 1. *Let* $\mathbf{X}$ *be an n-dimensional random vector with measurement error* $\mathbf{U}$. *Assume* $\mathbf{X}$ *and* $\mathbf{U}$ *are independent. Let* f^* *be a matrix-valued function such that* $E[f^*(U)]$ *exists. Then for* $\mathbf{w}$ *in the support of* $f_{\mathbf{W}}$,

$$E[f^*(\mathbf{U}) \mid (\mathbf{W} = \mathbf{w})] = \frac{\int_{\Re^n} f^*(\mathbf{w} - \mathbf{x})\, g(\mathbf{w} - \mathbf{x})\, f_{\mathbf{X}}(\mathbf{x})\, d\mathbf{x}}{f_{\mathbf{W}}(\mathbf{w})}.$$

LEMMA 2. *Assume* $\mathbf{\Sigma}$ *is positive definite. Let* $\mathbf{X}$ *be a random variable with measurement error* $\mathbf{U} \sim MVN(\mathbf{0}, \mathbf{\Sigma})$. *Assume* $\mathbf{X}$ *and* $\mathbf{U}$ *are independent. Then*

$\nabla f_{\mathbf{W}}(\mathbf{w})$ *exists and* $\nabla f_{\mathbf{W}}(\mathbf{w}) = \int_{\Re^n} \nabla g(\mathbf{w} - \mathbf{x})\, f_{\mathbf{X}}(\mathbf{x})\, d\mathbf{x}$,
$\mathbf{H}_{f_{\mathbf{W}}}(\mathbf{w})$ *exists and* $\mathbf{H}_{f_{\mathbf{W}}}(\mathbf{w}) = \int_{\Re^n} \mathbf{H}_g(\mathbf{w} - \mathbf{x})\, f_{\mathbf{X}}(\mathbf{x})\, d\mathbf{x}$,

$$\int_{\Re^n} \left| \frac{\partial f_{\mathbf{W}}(\mathbf{w})}{\partial w_j} \right| d\mathbf{w} < \infty$$

for almost all $\mathbf{w}_{[-j]} \in \Re^{n-1}$, *and*

$$\int_{\Re^n} \left| \frac{\partial^2 f_{\mathbf{W}}(\mathbf{w})}{\partial w_k w_j} \right| d\mathbf{w} < \infty$$

for almost all $\mathbf{w}_{[-j,-k]} \in \Re^{n-2}$.

The proof follows from the fact that $f_{\mathbf{W}}$ is defined by the convolution involving a normal density. The following lemmas are necessary to show that integration by parts can be done in the context of convolution. The proofs can be found in the appendix.

LEMMA 3. *If* $h(\mathbf{w}) \in L^1(\Re^n)$ *and* $\lim_{w_k \to \pm\infty} h(\mathbf{w})$ *exists, taking the limit in the kth component, then* $\lim_{w_k \to \pm\infty} h(\mathbf{w}) = 0$ *except on a set of* $(n-1)$ *dimensional measure* 0 *in* $\mathbf{w}_{[-k]}$ *space.*

LEMMA 4. *If*

$$\frac{\partial^2 f_{\mathbf{W}}(\mathbf{w})}{\partial w_k \, \partial w_j} \in L^1(\Re^n) \qquad \text{for} \quad j, k = 1, 2, ..., n,$$

for almost all $\mathbf{w}_{[-j, -k]} \in \Re^{n-2}$, *and*

$$\frac{\partial f_{\mathbf{W}}(\mathbf{w})}{\partial w_j} \in L^1(\Re^n) \qquad \text{for} \quad j = 1, 2, ..., n,$$

for almost all $\mathbf{w}_{[-j]} \in \Re^{n-1}$, *then*

$$\int_{\Re^n} \nabla f_{\mathbf{W}}(\mathbf{w}) \exp(i\mathbf{t}'\mathbf{w}) \, d\mathbf{w} = -i\mathbf{t} \int_{\Re^n} f_{\mathbf{W}}(\mathbf{w}) \exp(i\mathbf{t}'\mathbf{w}) \, d\mathbf{w} \tag{2.4}$$

and

$$\int_{\Re^n} \mathbf{H}_{f_{\mathbf{W}}}(\mathbf{w}) \exp(i\mathbf{t}'\mathbf{w}) \, d\mathbf{w} = -\mathbf{t}\mathbf{t}' \int_{\Re^n} f_{\mathbf{W}}(\mathbf{w}) \exp(i\mathbf{t}'\mathbf{w}) \, d\mathbf{w}. \tag{2.5}$$

With these facts, we have the following theorem containing necessary and sufficient conditions for $\mathbf{U}$ to be normal.

THEOREM 1. *Assume* $\mathbf{\Sigma}$ *positive definite.* $\mathbf{U} \sim MVN(\mathbf{0}, \mathbf{\Sigma})$ *iff*

$$E[\mathbf{U} \mid (\mathbf{W} = \mathbf{w})] = -\mathbf{\Sigma} \frac{\nabla f_{\mathbf{W}}(\mathbf{w})}{f_{\mathbf{W}}(\mathbf{w})}$$

for w *in the support of* $f_{\mathbf{W}}$,

$$\frac{\partial f_{\mathbf{W}}(\mathbf{w})}{\partial w_k} \in L^1(\Re^n) \qquad \text{for} \quad k = 1, 2, ..., n,$$

for almost all $\mathbf{w}_{[-k]} \in \Re^{n-1}$, *and* $\mathbf{U}$ *has a finite first-moment vector.*

Proof. Assume $\mathbf{U}$ is $MVN(\mathbf{0}, \boldsymbol{\Sigma})$. Then $\mathbf{U}$ has a first-moment vector, indeed $E(\mathbf{U}) = 0$. Letting $f^*(\mathbf{U}) = \mathbf{U}$ in Lemma 1 gives

$$E[\mathbf{U} \mid (\mathbf{W} = \mathbf{w})] = \frac{\int_{\Re^n} (\mathbf{w} - \mathbf{x})\, g(\mathbf{w} - \mathbf{x})\, f_{\mathbf{X}}(\mathbf{x})\, d\mathbf{x}}{f_{\mathbf{W}}(\mathbf{w})}.$$

Since g is the density of a multivariate normal distribution,

$$\nabla g(\mathbf{w} - \mathbf{x}) = -\boldsymbol{\Sigma}^{-1}(\mathbf{w} - \mathbf{x})\, g(\mathbf{w} - \mathbf{x}).$$

This implies

$$E[(\mathbf{U}) \mid (\mathbf{W} = \mathbf{w})] = -\frac{\int_{\Re^n} \boldsymbol{\Sigma}\, \nabla g(\mathbf{w} - \mathbf{x})\, f_{\mathbf{X}}(\mathbf{x})\, d\mathbf{x}}{f_{\mathbf{W}}(\mathbf{w})}.$$

By Lemma 2,

$$\nabla f_{\mathbf{W}}(\mathbf{w}) = \int_{\Re^n} \nabla g(\mathbf{w} - \mathbf{x})\, f_{\mathbf{X}}(\mathbf{x})\, d\mathbf{x}.$$

Therefore,

$$E[\mathbf{U} \mid (\mathbf{W} = \mathbf{w})] = -\boldsymbol{\Sigma} \frac{\nabla f_{\mathbf{W}}(\mathbf{w})}{f_{\mathbf{W}}(\mathbf{w})}.$$

In Lemma 2 it was shown that

$$\int_{\Re^n} \left| \frac{\partial f_{\mathbf{W}}(\mathbf{w})}{\partial w_k} \right| d\mathbf{w} < \infty$$

for almost all $\mathbf{w}_{[-k]} \in \Re^{n-1}$, and the necessity follows.

Now assume

$$E[\mathbf{U} \mid (\mathbf{W} = \mathbf{w})] = -\boldsymbol{\Sigma} \frac{\nabla f_{\mathbf{W}}(\mathbf{w})}{f_{\mathbf{W}}(\mathbf{w})}, \qquad \int_{\Re} \left| \frac{\partial f_{\mathbf{W}}(\mathbf{w})}{\partial w_k} \right| dw_k < \infty \tag{2.6}$$

for almost all $\mathbf{w}_{[-k]} \in \Re^{n-1}$, and $\mathbf{U}$ has a finite first-moment vector.

Multiplying each side of Eq. (2.6) by $\exp(i\mathbf{t}'\mathbf{w})$ generates the Fourier transform, and taking the expectation yields

$$E[E[\mathbf{U} \mid (\mathbf{W} = \mathbf{w})] \exp(i\mathbf{t}'\mathbf{w})] = E\left[-\boldsymbol{\Sigma} \frac{\nabla f_{\mathbf{W}}(\mathbf{w})}{f_{\mathbf{W}}(\mathbf{w})} \exp(i\mathbf{t}'\mathbf{w}) \right]. \tag{2.7}$$

Since $\mathbf{X}$ and $\mathbf{U}$ are independent and $\mathbf{U}$ has a finite first moment,

$$E[E[\mathbf{U} \mid (\mathbf{W}=\mathbf{w})] \exp(i\mathbf{t}'\mathbf{w})] = E[\mathbf{U} \exp(i\mathbf{t}'\mathbf{w})]$$

$$= -i \nabla\phi_{\mathbf{U}}(\mathbf{t}) \, \phi_{\mathbf{X}}(\mathbf{t}).$$

The independence of $\mathbf{X}$ and $\mathbf{U}$ also gives $\phi_{\mathbf{W}}(\mathbf{t}) = \phi_{\mathbf{U}}(\mathbf{t}) \, \phi_{\mathbf{W}}(\mathbf{t})$. Since characteristic functions are equal to 1 at 0 and continuous, there exists $\delta_1 > 0$ such that $\phi_{\mathbf{U}}(\mathbf{t}) \neq 0$ for $\|\mathbf{t}\| < \delta_1$ ($\|\cdot\|$ denotes the Euclidean norm). Thus for $\|\mathbf{t}\| < \delta_1$,

$$E[E[\mathbf{U} \mid (\mathbf{W}=\mathbf{w})] \exp(i\mathbf{t}'\mathbf{w})] = -\frac{\nabla\phi_{\mathbf{U}}(\mathbf{t})}{\phi_{\mathbf{U}}(\mathbf{t})} \phi_{\mathbf{W}}(\mathbf{t}). \tag{2.8}$$

Examining the right-hand side of Eq. (2.7) gives

$$E\left(-\boldsymbol{\Sigma} \frac{\nabla f_{\mathbf{W}}(\mathbf{w})}{f_{\mathbf{W}}(\mathbf{w})} \exp(i\mathbf{t}'\mathbf{w}) \right) = -\int_{\mathfrak{R}^n} \boldsymbol{\Sigma} \nabla f_{\mathbf{W}}(\mathbf{w}) \exp(i\mathbf{t}'\mathbf{w}) \, d\mathbf{w}.$$

Lemma 4 showed that integration by parts yields

$$-\int_{\mathfrak{R}^n} \boldsymbol{\Sigma} \nabla f_{\mathbf{W}}(\mathbf{w}) \exp(i\mathbf{t}'\mathbf{w}) \, d\mathbf{w} = i\boldsymbol{\Sigma}\mathbf{t} \int_{\mathfrak{R}^n} f_{\mathbf{W}}(\mathbf{w}) \exp(i\mathbf{t}'\mathbf{w}) \, d\mathbf{w}$$

$$= i\boldsymbol{\Sigma}\mathbf{t}\phi_{\mathbf{W}}(\mathbf{t}). \tag{2.9}$$

Equating the equivalent expressions of both sides of Eq. (2.7) given in (2.8) and (2.9) yields

$$-\frac{\nabla\phi_{\mathbf{U}}(\mathbf{t})}{\phi_{\mathbf{U}}(\mathbf{t})} \phi_{\mathbf{W}}(\mathbf{t}) = \boldsymbol{\Sigma}\mathbf{t}\phi_{\mathbf{W}}(\mathbf{t}).$$

Take $\delta_2 < \delta_1$ so that for $\|\mathbf{t}\| < \delta_2$

$$\nabla \log(\phi_{\mathbf{U}}(\mathbf{t})) = -\boldsymbol{\Sigma}\mathbf{t}.$$

This implies that

$$\log(\phi_{\mathbf{U}}(\mathbf{t})) = -\frac{\mathbf{t}'\boldsymbol{\Sigma}\mathbf{t}}{2} + k_1, \tag{2.10}$$

where k_1 is a constant of integration. Since $\log(\phi_U(0))=0$, $k_1=0$. Thus Eq. (2.10) gives

$$\phi_{\mathbf{U}}(\mathbf{t})=\exp\left(-\frac{\mathbf{t}'\boldsymbol{\Sigma}\mathbf{t}}{2}\right). \tag{2.11}$$

Thus for $\mathbf{t}$ such that $\|\mathbf{t}\|<\delta_2$, the characteristic function of $\mathbf{U}$ is the characteristic function of a multivariate normal distribution with mean $\mathbf{0}$ and variance covariance matrix $\boldsymbol{\Sigma}$. This implies that $\mathbf{U}\sim MVN(\mathbf{0},\boldsymbol{\Sigma})$. Q.E.D

Proving the normality of the error involved a simple differential equation of the characteristic function of the measurement error. Since this equation has a unique solution, the normality of the measurement error is proved.

This result can now be utilized to estimate f'_W/f_W. In the one-dimensional case, density smoothing methods for W are used to estimate f'_W/f_W (Sepanski, Carroll, and Knickerbocker, 1994; Stefanski and Carroll, 1990, 1991). Estimation of $E[\mathbf{U}\mid(\mathbf{W}=\mathbf{w})]$ in the multivariate setting has received little attention in the literature.

3. THE SECOND MOMENT MATRIX

From a mathematical standpoint, it is interesting that the conditional expectation of the second moment matrix of $\mathbf{U}$ characterizes normality as well. The proof of the following theorem requires a different approach than the one utilized for Theorem 1. In this proof, a second-order differential equation involving the characteristic function of the measurement error is obtained. This differential equation has a unique solution because it is linear, $\phi_U(\mathbf{0})=1$ and $\nabla\phi_U(\mathbf{0})=\mathbf{0}$. With these facts and the condition that

$$\mathbf{H}_\phi(\mathbf{t})=\phi(\mathbf{t})(\boldsymbol{\Sigma}\mathbf{t}\mathbf{t}'\boldsymbol{\Sigma}-\boldsymbol{\Sigma}),$$

where $\mathbf{H}_\phi(\mathbf{t})$ is the Hessian matrix of $\phi(\mathbf{t})$ and $\boldsymbol{\Sigma}$ is a matrix of constants, using standard differential equations techniques (for a reference see Ford, 1955) it is easily shown that $\phi(\mathbf{t})$ is unique.

THEOREM 2. *Assume $\boldsymbol{\Sigma}$ is positive definite.* $\mathbf{U}\sim MVN(\mathbf{0},\boldsymbol{\Sigma})$, *iff*

$$E[\mathbf{U}\mathbf{U}'\mid(\mathbf{W}=\mathbf{w})]=\boldsymbol{\Sigma}\,\frac{\mathbf{H}_{f_\mathbf{W}}(\mathbf{w})}{f_\mathbf{W}(\mathbf{w})}\,\boldsymbol{\Sigma}+\boldsymbol{\Sigma} \tag{3.12}$$

for w in the support of $f_\mathbf{W}$,

$$\frac{\partial f_\mathbf{W}(\mathbf{w})}{\partial w_j}\in L^1(\Re^n)\qquad \forall j\in 1,2,\ldots,n,$$

and almost all $\mathbf{w}_{[-j]} \in \Re^{n-1}$,

$$\frac{\partial^2 f_{\mathbf{W}}(\mathbf{w})}{\partial w_j \, \partial w_k} \in L^1(\Re^n) \qquad \forall j, k \in 1, 2, ..., n,$$

and almost all $\mathbf{w}_{[-j,\,-k]} \in \Re^{n-2}$, *and* $\mathbf{U}$ *has a finite second moment matrix.*

Proof. Assume $\mathbf{U} \sim MVN(\mathbf{0}, \boldsymbol{\Sigma})$. Then $\mathbf{U}$ has a second moment matrix. In fact, $E(\mathbf{U}\mathbf{U}') = \boldsymbol{\Sigma}$. Lemma 1 ($f^*(\mathbf{U}) = \mathbf{U}\mathbf{U}'$), implies

$$E[\mathbf{U}\mathbf{U}' \mid (\mathbf{W} = \mathbf{w})] = \frac{\int_{\Re^n} [(\mathbf{w}-\mathbf{x})(\mathbf{w}-\mathbf{x})'] \, g(\mathbf{w}-\mathbf{x}) \, f_{\mathbf{X}}(\mathbf{x}) \, d\mathbf{x}}{f_{\mathbf{W}}(\mathbf{w})}. \tag{3.13}$$

Since g is the density of a multivariate normal distribution,

$$\nabla g(\mathbf{w}-\mathbf{x}) = -g(\mathbf{w}-\mathbf{x})[\boldsymbol{\Sigma}^{-1}(\mathbf{w}-\mathbf{x})].$$

The Hessian of g is $\mathbf{H}_g = \nabla(\nabla g)'$, which implies

$$\mathbf{H}_g(\mathbf{w}-\mathbf{x}) = -g(\mathbf{w}-\mathbf{x}) \, \boldsymbol{\Sigma}^{-1} + g(\mathbf{w}-\mathbf{x})[\boldsymbol{\Sigma}^{-1}(\mathbf{w}-\mathbf{x})(\mathbf{w}-\mathbf{x})' \, \boldsymbol{\Sigma}^{-1}].$$

Thus

$$\boldsymbol{\Sigma}\mathbf{H}_g(\mathbf{w}-\mathbf{x}) \, \boldsymbol{\Sigma} = -g(\mathbf{w}-\mathbf{x}) \, \boldsymbol{\Sigma} + g(\mathbf{w}-\mathbf{x})[(\mathbf{w}-\mathbf{x})(\mathbf{w}-\mathbf{x})'].$$

Utilizing this result in Eq. (3.13) gives

$$E[\mathbf{U}\mathbf{U}' \mid (\mathbf{W} = \mathbf{w})]$$

$$= \frac{1}{f_{\mathbf{W}}(\mathbf{w})} \int_{\Re_n} [\boldsymbol{\Sigma}\mathbf{H}_g(\mathbf{w}-\mathbf{x}) \, \boldsymbol{\Sigma} + g(\mathbf{w}-\mathbf{x}) \, \boldsymbol{\Sigma}] \, f_{\mathbf{X}}(\mathbf{x}) \, d\mathbf{x}. \tag{3.14}$$

$$= \frac{1}{f_{\mathbf{W}}(\mathbf{w})} \boldsymbol{\Sigma} \left[\int_{\Re^n} \mathbf{H}_g(\mathbf{w}-\mathbf{x}) \, f_{\mathbf{X}}(\mathbf{x}) \, d\mathbf{x} \right] \boldsymbol{\Sigma} + \boldsymbol{\Sigma}. \tag{3.15}$$

By Lemma 2,

$$\mathbf{H}_{f_{\mathbf{W}}}(\mathbf{w}) = \int_{\Re^n} \mathbf{H}_g(\mathbf{w}-\mathbf{x}) \, f_{\mathbf{X}}(\mathbf{x}) \, d\mathbf{x}.$$

So Eq. (3.15) may be rewritten as

$$E[\mathbf{U}\mathbf{U}' \mid (\mathbf{W} = \mathbf{w})] = \frac{\boldsymbol{\Sigma}\mathbf{H}_{f_{\mathbf{W}}}(\mathbf{w}) \, \boldsymbol{\Sigma}}{f_{\mathbf{W}}(\mathbf{w})} + \boldsymbol{\Sigma},$$

which is Eq. (3.12) in the statement of the theorem. Lemma 2 showed that

$$\frac{\partial^2 f_{\mathbf{W}}(\mathbf{w})}{\partial w_j\,\partial w_k} \in L^1(\Re^n) \qquad \forall r, s \in 1, 2, ..., n.$$

Conversely, assume $\mathbf{U}$ is such that for $\mathbf{w}$ on the support of $f_{\mathbf{W}}$,

$$E[\mathbf{U}\mathbf{U}' \mid (\mathbf{W}=\mathbf{w})] = \boldsymbol{\Sigma}\,\frac{\mathbf{H}_{f_{\mathbf{W}}}(\mathbf{w})}{f_{\mathbf{W}}(\mathbf{w})}\,\boldsymbol{\Sigma} + \boldsymbol{\Sigma}, \tag{3.16}$$

and assume the regularity conditions listed in the statement of the theorem. Taking the Fourier transforms as in Theorem 1 of both sides of Eq. (3.16) yields

$$E[E[\mathbf{U}\mathbf{U}' \mid (\mathbf{W}=\mathbf{w})]\exp(i\mathbf{t}'\mathbf{w})] = E\left[\left(\boldsymbol{\Sigma}\,\frac{\mathbf{H}_{f_{\mathbf{W}}}(\mathbf{w})}{f_{\mathbf{W}}(\mathbf{w})}\,\boldsymbol{\Sigma} + \boldsymbol{\Sigma}\right)\exp(i\mathbf{t}'\mathbf{w})\right]. \tag{3.17}$$

Since $\mathbf{X}$ and $\mathbf{U}$ are independent and $\mathbf{U}$ has a finite second moment matrix,

$$\begin{aligned} E[E[\mathbf{U}\mathbf{U}' \mid (\mathbf{W}=\mathbf{w})]\exp(i\mathbf{t}'\mathbf{w})] &= E[\mathbf{U}\mathbf{U}'\exp(i\mathbf{t}'\mathbf{w})] \\ &= -\mathbf{H}_{\phi_{\mathbf{U}}}(\mathbf{t})\,\phi_{\mathbf{X}}(\mathbf{t}). \end{aligned}$$

The independence of $\mathbf{X}$ and $\mathbf{U}$ also gives $\phi_{\mathbf{W}}(\mathbf{t}) = \phi_{\mathbf{U}}(\mathbf{t})\phi_{\mathbf{X}}(\mathbf{t})$. Since characteristic functions are equal to 1 at 0 and continuous, there exists $\delta_1 > 0$ such that $\phi_{\mathbf{U}}(\mathbf{t}) \neq 0$ for $\|t\| < \delta_1$. Thus for $\|t\| < \delta_1$,

$$E[E[\mathbf{U}\mathbf{U}' \mid (\mathbf{W}=\mathbf{w})]\exp(i\mathbf{t}'\mathbf{w})] = -\frac{\mathbf{H}_{\phi_{\mathbf{U}}}(\mathbf{t})}{\phi_{\mathbf{U}}(\mathbf{t})}\,\phi_{\mathbf{W}}(\mathbf{t}). \tag{3.18}$$

Examining the right-hand side of Eq. (3.17) gives

$$\begin{aligned} &E\left[\left(\boldsymbol{\Sigma}\,\frac{\mathbf{H}_{f_{\mathbf{W}}}(\mathbf{w})}{f_{\mathbf{W}}(\mathbf{w})}\,\boldsymbol{\Sigma} + \boldsymbol{\Sigma}\right)\exp(i\mathbf{t}'\mathbf{w})\right] \\ &\quad = \boldsymbol{\Sigma}\left[\int_{\Re^n} \frac{\mathbf{H}_{f_{\mathbf{W}}}(\mathbf{w})}{f_{\mathbf{W}}(\mathbf{w})}\exp(i\mathbf{t}'\mathbf{w})\,f_{\mathbf{W}}(\mathbf{w})\,d\mathbf{w}\right]\boldsymbol{\Sigma} + \boldsymbol{\Sigma}\int_{\Re^n}\exp(i\mathbf{t}'\mathbf{w})\,f_{\mathbf{W}}(\mathbf{w})\,d\mathbf{w} \\ &\quad = \boldsymbol{\Sigma}\left[\int_{\Re^n}\mathbf{H}_{f_{\mathbf{W}}}(\mathbf{w})\exp(i\mathbf{t}'\mathbf{w})\,d\mathbf{w}\right]\boldsymbol{\Sigma} + \boldsymbol{\Sigma}\phi_{\mathbf{W}}(\mathbf{t}). \end{aligned}$$

Lemma 4 showed that performing integration by parts twice yields

$$\int_{\Re^n} \mathbf{H}_{f_\mathbf{W}}(\mathbf{w}) \exp(i\mathbf{t}'\mathbf{w})\, d\mathbf{w} = -\mathbf{tt}' \int_{\Re^n} f_\mathbf{W}(\mathbf{w}) \exp(i\mathbf{t}'\mathbf{w})\, d\mathbf{w}$$
$$= -\mathbf{tt}'\phi_\mathbf{W}(\mathbf{t}). \tag{3.19}$$

Equating the equivalent expressions of both sides of Eq. (3.17) given in (3.18) and (3.19) and dividing both sides by $\phi_W(\mathbf{t}) \neq 0$ yields

$$\frac{\mathbf{H}_{\phi_\mathbf{U}}(\mathbf{t})}{\phi_\mathbf{U}(\mathbf{t})} = \mathbf{\Sigma tt'\Sigma} - \mathbf{\Sigma}. \tag{3.20}$$

Note that

$$\phi_\mathbf{U}(t) = \exp\left(-\frac{\mathbf{t}'\mathbf{\Sigma t}}{2}\right) \tag{3.21}$$

is a solution to the differential equation given in Eq. (3.20). Since $\phi_\mathbf{U}(\mathbf{t})$ is a characteristic function,

$$\phi_\mathbf{U}(\mathbf{0}) = 1 \qquad \text{and} \qquad \nabla\phi_\mathbf{U}(\mathbf{0}) = \mathbf{0},$$

which implies that the expression for $\phi_\mathbf{U}(\mathbf{t})$ given in Eq. (3.21) is unique. Thus for $\mathbf{t}$ such that $\|\mathbf{t}\| < \delta$ the characteristic function of $\mathbf{U}$ is the characteristic function of a multivariate normal distribution with mean $\mathbf{0}$ and variance covariance matrix $\mathbf{\Sigma}$. Therefore, $\mathbf{U} \sim MVN(0, \mathbf{\Sigma})$. Q.E.D

4. DISCUSSION

Using techniques requiring the Fourier transform and knowledge of differential equations has defined the necessary and sufficient conditions for the measurement error to be characterized as normal. Initially, the author wished to show that

$$E[X \mid (W = w)] = w + \sigma^2 \frac{f'_W(w)}{f_W(w)}$$

characterized the distribution of the measurement error. This is indeed true, but as Section 2 demonstrated the result holds in the multivariate setting as well. Further research may lead to using this result as a diagnostic tool in determining when an assumption of the normality of the measurement error is reasonable in a multivariate setting.

The techniques used in proving the first theorem were also applied to investigating the second moment matrix of the measurement error. Thus knowledge of the covariance structure conditioned on the observed vector can also characterize the distribution of the measurement error.

APPENDIX

Proof of Lemma 3. By assumption,

$$\lim_{w_k \to -\infty} h(\mathbf{w})$$

exists. Let m be an integer and M a negative-valued constant. Define the set

$$A_m = \left\{ \mathbf{w}_{[-k]} \colon |h(\mathbf{w})| > \frac{1}{m} \text{ and } w_k < M \right\}.$$

Since $h(\mathbf{w}) \in L^1(\Re^n)$, there exists positive constant c so that invoking Tonnelli's Theorem to permute the order of integration yields

$$c = \int_{\Re^n} |h(\mathbf{w})| \, d\mathbf{w} \geqslant \int_{-\infty}^{M} \int_{A_m} |h(\mathbf{w})| \, d\mathbf{w}. \tag{5.22}$$

Since $|h(\mathbf{w})| > 1/m$ on A_m, Eq. (5.22) implies

$$c \geqslant \int_{-\infty}^{M} \frac{1}{m} \, v(A_m) \, dw_k, \tag{5.23}$$

where v denotes $(n-1)$-dimensional Lebesgue measure. As the integrand is not a function of w_k, $v(A_m) = 0$ in order that the integral in Eq. (5.23) be finite. This implies that

$$v\left(\bigcup_m A_m \right) = 0.$$

Define

$$B_l = \left\{ \mathbf{w}_{[-k]} \colon \lim_{m \to -\infty} |h(\mathbf{w})| < \frac{1}{l} \right\}.$$

Note that

$$B_l \subset \bigcup_m A_m \quad \forall l,$$

and so

$$v(B_l) = 0 \quad \forall l,$$

implying $v(\bigcup_l B_l) = 0$. Therefore, $\lim_{w_k \to -\infty} h(\mathbf{w}) = 0$ except on a set of $(n-1)$-dimensional Lebesgue measure 0 in $\mathbf{w}_{[-k]}$ space.

Using a similar argument, $\lim_{w_k \to \infty} h(\mathbf{w}) = 0$ except on a set of measure 0. Q.E.D

Proof of Lemma 4. To prove Eq. (2.4), it suffices to show that for $1 \leqslant j \leqslant n$,

$$\int_{\Re^n} \frac{\partial f_{\mathbf{W}}(\mathbf{w})}{\partial w_j} \exp(i\mathbf{t}'\mathbf{w})\, d\mathbf{w} = -it_j \int_{\Re^n} f_{\mathbf{W}}(\mathbf{w}) \exp(i\mathbf{t}'\mathbf{w})\, d\mathbf{w}.$$

Since

$$\frac{\partial f_{\mathbf{W}}(\mathbf{w})}{\partial w_j} \in L^1(\Re^n)$$

by assumption, Fubini's Theorem can be invoked to permute the order of integration. This yields

$$\int_{\Re^n} \frac{\partial f_{\mathbf{W}}(\mathbf{w})}{\partial w_j} \exp(i\mathbf{t}'\mathbf{w})\, d\mathbf{w} = \int_{\Re^{n-1}} \int_{\Re} \frac{\partial f_{\mathbf{W}}(\mathbf{w})}{\partial w_j} \exp(i\mathbf{t}'\mathbf{w})\, dw_j\, d\mathbf{w}_{[-j]}.$$

Performing integration by parts on the inner integral yields

$$\int_a^b \frac{\partial f_{\mathbf{W}}(\mathbf{w})}{\partial w_j} \exp(i\mathbf{t}'\mathbf{w})\, dw_j = f_{\mathbf{W}}(\mathbf{w}) \exp(i\mathbf{t}'\mathbf{w})|_a^b - it_j \int_a^b f_{\mathbf{W}}(\mathbf{w}) \exp(i\mathbf{t}'\mathbf{w})\, dw_j. \tag{5.24}$$

Now hold b fixed and take $\lim_{a \to -\infty}$ of both sides of Eq. (5.24). Since both

$$\frac{\partial f_{\mathbf{W}}(\mathbf{w})}{\partial w_j} \qquad \text{and} \qquad f_{\mathbf{W}}(\mathbf{w})$$

are elements of $L^1(\Re^n)$, the

$$\lim_{a \to -\infty} f_{\mathbf{W}}(\mathbf{w}) \exp(i\mathbf{t}'\mathbf{w})|_{w_j = a}^b \tag{5.25}$$

exists and is less than infinity.

Thus Lemma 3 can be invoked to conclude

$$\lim_{w_j \to -\infty} f(\mathbf{w}) = 0.$$

Similarly, hold a fixed and take the $\lim_{b \to \infty}$ of both sides of Eq. (5.24). An identical argument using Lemma 3 can be invoked to show

$$\lim_{w_j \to \infty} f(\mathbf{w}) = 0.$$

Thus

$$f(\mathbf{w}) \exp(i\mathbf{t}'\mathbf{w})|_{-\infty}^{\infty} = 0$$

and

$$\int_{\Re^n} \frac{\partial f_{\mathbf{W}}(\mathbf{w})}{\partial w_j} \exp(i\mathbf{t}'\mathbf{w})\, d\mathbf{w} = 0 - it_k \int_{\Re^n} f_{\mathbf{W}}(\mathbf{w}) \exp(i\mathbf{t}'\mathbf{w})\, d\mathbf{w}. \tag{5.26}$$

To prove Eq. (2.5), it suffices to show for j and k less than or equal n,

$$\int_{\Re^n} \frac{\partial^2 f_{\mathbf{W}}(\mathbf{w})}{\partial w_k\, \partial w_j} \exp(i\mathbf{t}'\mathbf{w})\, d\mathbf{w} = t_k t_j \int_{\Re^n} f_{\mathbf{W}}(\mathbf{w}) \exp(i\mathbf{t}'\mathbf{w})\, d\mathbf{w}.$$

Since

$$\frac{\partial^2 f_{\mathbf{W}}(\mathbf{w})}{\partial w_k\, \partial w_j} \in L^1(\Re^n),$$

the order of integration may again he permuted by Fubini's Theorem, and integration by parts of

$$\int_{\Re} \frac{\partial^2 f_{\mathbf{W}}(\mathbf{w})}{\partial w_k\, \partial w_j} \exp(i\mathbf{t}'\mathbf{w})\, dw_k$$

implies that

$$\int_a^b \frac{\partial^2 f_{\mathbf{W}}(\mathbf{w})}{\partial w_k\, \partial w_j} \exp(i\mathbf{t}'\mathbf{w})\, dw_k$$

$$= \frac{\partial f_{\mathbf{W}}(\mathbf{w})}{\partial w_j} \exp(i\mathbf{t}'\mathbf{w})|_a^b - it_k \int_a^b \frac{\partial f_{\mathbf{W}}(\mathbf{w})}{\partial w_j} \exp(i\mathbf{t}'\mathbf{w})\, dw_k. \tag{5.27}$$

Since both

$$\frac{\partial^2 f_{\mathbf{w}}(\mathbf{w})}{\partial w_k\, \partial w_j} \in L^1(\mathfrak{R}^n) \qquad \text{and} \qquad \frac{\partial f_{\mathbf{w}}(\mathbf{w})}{\partial w_j} \in L^1(\mathfrak{R}^n), \tag{5.28}$$

$$\lim_{a \to -\infty} f_{\mathbf{w}}(\mathbf{w}) \exp(i\mathbf{t}'\mathbf{w})\big|_{w_k=a}^{b}$$

exists and is less than infinity. Thus Lemma 3 can be invoked to conclude

$$\lim_{w_k \to -\infty} = \frac{\partial f(\mathbf{w})}{\partial w_j} = 0.$$

Similarly, hold a fixed and take the $\lim_{b \to \infty}$ of both sides of Eq. (5.27). An identical argument using Lemma 3 can be invoked to show

$$\lim_{w_k \to \infty} \frac{\partial f(\mathbf{w})}{\partial w_j} = 0.$$

Thus

$$\int_{-\infty}^{\infty} \frac{\partial^2 f_{\mathbf{w}}(\mathbf{w})}{\partial w_k\, \partial w_j} \exp(i\mathbf{t}'\mathbf{w})\, d\mathbf{w} = -it_k \int_{-\infty}^{\infty} \frac{\partial f_{\mathbf{w}}(\mathbf{w})}{\partial w_j} \exp(i\mathbf{t}'\mathbf{w})\, d\mathbf{w}.$$

Using vector notation to summarize,

$$\int_{\mathfrak{R}^n} \mathbf{H}_{f_{\mathbf{w}}}(\mathbf{w}) \exp(i\mathbf{t}'\mathbf{w})\, d\mathbf{w} = -i\mathbf{t} \int_{\mathfrak{R}^n} \nabla f_{\mathbf{w}}(\mathbf{w}) \exp(i\mathbf{t}'\mathbf{w})\, d\mathbf{w}.$$

Now it remains only to invoke the result of Eq. (5.26) to obtain

$$\int_{\mathfrak{R}^n} \mathbf{H}_{f_{\mathbf{w}}}(\mathbf{w}) \exp(i\mathbf{t}'\mathbf{w})\, d\mathbf{w} = -\mathbf{t}\mathbf{t}' \int_{\mathfrak{R}^n} f_{\mathbf{w}}(\mathbf{w}) \exp(i\mathbf{t}'\mathbf{w})\, d\mathbf{w}. \qquad \text{Q.E.D.}$$

REFERENCES

1. R. J. Adcock, Note on the method of least squares, *Analyst* **4** (1877), 183–184.
2. R. J. Adcock, A problem in least squares, *Analyst* **5** (1878), 53–54.
3. T. W. Anderson, “An Introduction to Multivariate Statistical Analysis,” Wiley, New York, 1958.
4. R. J. Carroll, D. Ruppert, and L. A. Stefanski, “Measurement Error in Nonlinear Models,” Chapman and Hall, New York, 1995.
5. W. G. Cochran, Errors of measurement in statistics, *Technometrics* **10** (1968), 637–666.
6. L. R. Ford, “Differential Equations,” McGraw–Hill, New York, 1955.
7. W. A. Fuller, “Measurement Error Models,” Wiley, New York, 1987.

8. L. J. Gleser, Improvements of the naive approach to estimation in nonlinear errors-in-variables regression models, *in* "Statistical Analysis of Measurement Error Models and Application" (P. J. Brown and W. A. Fuller, Eds.), American Mathematics Society, Providence, RI, 1990.
9. A. M. Kagan, Yu. V. Linnik, and C. R. Rao, "Characterization Problems in Mathematical Statistics," Wiley, New York, 1973.
10. M. G. Kendall, Regression, structure, and functional relationship, I, *Biometrika* **38** (1951), 11–25.
11. M. G. Kendall, Regression, structure,and functional relationship, II, *Biometrika* **39** (1952), 96–108.
12. M. G. Kendall, A. Stuart, and J. K. Ord, "The Advanced Theory of Statistics," 5th ed., Oxford Univ. Press, New York, 1987.
13. R. G. Laha and E. Lukacs, On a problem connected with quadratic regression, *Biometrika* **47** (1960), 335–343.
14. D. V. Lindley, Regression lines and the linear functional relationship, *J. Roy. Statist. Soc. Suppl.* **9** (1947), 218–244.
15. E. Lukacs and R. G. Laha, "Applications of Characteristic Functions," Hafner, New York, 1964.
16. C. Morris, Natural exponential families with quadratic variance functions, *Ann. of Statist.* **10** (1982), 65–80.
17. C. R. Rao, Note on a problem of Ragnar Frisch, *Econometrika* **15** (1947), 245–249 correction, *Ibid.* **17**, 212.
18. C. R. Rao, Some problems in the characterization of the multivariate normal distribution, *in* "Statistical Distributions in Scientific Work" Vol. 3, "Characterizations and Applications" (G. P. Patil, S. Kotz, and J. K. Ord, Eds.), Reidel, Dordrecht and Boston, 1975.
19. B. Rosner, W. C. Willett, and D. Spiegelman, Correction of logistic regression relative risk estimates and confidence intervals for systematic within-person measurement error, *Statist. Medicine* **8** (1989), 1051–1070.
20. J. H. Sepanski, R. J. Carroll, and R. Knickerbocker, A semiparametric correction for attenuation, *J. Amer. Statist. Assoc.* **89** (1994), 1366–1373.
21. C. H. Spiegelman, Two pitfalls of using standard regression diagnostics when both X and Y have measurement error, *Amer. Statist.* **40** (1986), 245–248.
22. L. A. Stefanski and R. J. Carroll, Covariate measurement error in logistic regression, *Ann. Statist.* **13** (1985), 1335–1351.
23. L. A. Stefanski and R. J. Carroll, Score tests in generalized linear measurement error models, *J. Roy. Statist. Soc. Ser. B* **52** (1990), 345–359.
24. L. A. Stefanski and R. J. Carroll, Deconvolution based score tests in measurement error models, *Ann. Statist.* **19** (1991), 249–259.
25. A. S. Whittemore, Errors in variables regression using Stein estimates, *Amer. Statist.* **43** (1989), 226–228.