REFERENCES

Fulks, W. (1978), *Advanced Calculus* (3rd ed.), New York: Wiley.

MCCUE, K. F. (2001), "THE STATISTICAL FOUNDATIONS OF THE EI METHOD," *THE AMERICAN STATISTICIAN*, 55, 106–110: COMMENT BY LEWIS AND REPLY

I read with great interest "The Statistical Foundations of the EI Method" in the May 2001 issue of *The American Statistician*. I write to point out a substantial error in McCue's analysis. McCue claims that King "miscalculates the distribution of $\beta_i^b|T_i$" with the result that "both King's precinct-level and district-level estimates are also incorrect" (p. 108). This is a serious assertion that cuts right to the heart of the usefulness of King's method and the software that implements it. However, McCue's claim is incorrect.

Consider a set of voting precincts, $i = 1, \ldots, p$, and the following definitions. Let T_i be the fraction of citizens in precinct i that vote, X_i be the fraction of citizens that are African American, and β_i^b and β_i^w be the fraction of African Americans and non-African Americans that vote, respectively. The following identity must then hold

$$T_i = \beta_i^b X_i + \beta_i^w (1 - X_i).$$

This identity is the core of King's method. Given a set of observations on T and X, the assumption that the joint distribution of (β_i^b, β_i^w) is truncated bivariate normal (TBVN), and the assumption that the β's are independent of X, King is able to estimate the parameters of the assumed TBVN distribution. The distribution of the β's is truncated such that (β_i^b, β_i^w) lies on the unit square reflecting the fact that the fraction of citizens (of any type) that vote must be in the $[0, 1]$ interval. Given estimates of the TBVN parameters, King then generates a posteriori estimates of each (β_i^b, β_i^w) conditional on T_i. King asserts that conditional on T_i (and, implicitly, on X_i), the distribution of β_i^b is univariate truncated normal (TN). McCue's proof asserts that the law of iterated expectations would fail to hold "in general" if $\beta_i^b|T_i$ was TN though he offers no demonstration of such a failure (p. 108).

To prove that McCue's claim is false, begin with the notation and definitions employed by King (1997) and McCue (2001). In particular, let

$$\boldsymbol{\beta} = \begin{bmatrix} \beta_i^b \\ \beta_i^w \end{bmatrix}$$

be distributed truncated bivariate normal (TBVN) on the unit square with location vector

$$\mathcal{B} = \begin{bmatrix} \mathcal{B}^b \\ \mathcal{B}^w \end{bmatrix}$$

and dispersion matrix

$$\boldsymbol{\Sigma} = \begin{bmatrix} \sigma_b & \sigma_{bw} \\ \sigma_{bw} & \sigma_{bw} \end{bmatrix}.$$

By the definition of the TBVN,

$$f(\boldsymbol{\beta}) = C_i^{-1}(2\pi)^{-1}|\boldsymbol{\Sigma}|^{-1/2} \times \exp\left[-\frac{1}{2}(\boldsymbol{\beta}_i - \mathcal{B})'\boldsymbol{\Sigma}^{-1}(\boldsymbol{\beta}_i - \mathcal{B})\right],$$

for $(\beta_i^b, \beta_w^b) \in [0, 1] \times [0, 1]$ and 0 otherwise where

$$C_i = (2\pi)^{-1}|\boldsymbol{\Sigma}|^{-1/2}\int_0^1\int_0^1 \exp\left[-\frac{1}{2}(\boldsymbol{\beta}_i - \mathcal{B})'\boldsymbol{\Sigma}^{-1}(\boldsymbol{\beta}_i - \mathcal{B})\right] d\beta_i^w d\beta_i^b.$$

Now define the basic "accounting identity" to be

$$T_i = \beta_i^b X_i + \beta_i^w (1 - X_i)$$

for $0 \leq X_i < 1$ (for $X_i = 1$, $\beta_i^b = T_i$ and the posterior distribution of β_i^b given T_i is degenerate). The linear transformation matrix $\mathbf{H}$ that takes $\boldsymbol{\beta}$ to

$$\mathbf{Z}_i = \begin{bmatrix} \beta_i^b \\ T_i \end{bmatrix}$$

is

$$\mathbf{H}_i = \begin{bmatrix} 1 & 0 \\ X_i & 1 - X_i \end{bmatrix}.$$

Using standard change-of-variables methods,

$$f(\mathbf{Z}_i) = C_i^{-1}(2\pi)^{-1}|\mathbf{H}_i|^{-1}|\boldsymbol{\Sigma}|^{-1/2} \times \exp\left[-\frac{1}{2}\left(\mathbf{H}_i^{-1}\mathbf{Z}_i - \mathcal{B}\right)'\boldsymbol{\Sigma}^{-1}\left(\mathbf{H}_i^{-1}\mathbf{Z}_i - \mathcal{B}\right)\right]$$

for $X_i\beta_i^b < T_i < (1 - X_i) + X_i\beta_i^b$ and $0 < \beta_i^b < 1$ and 0 otherwise. Rearranging and applying properties of determinants,

$$f(\mathbf{Z}_i) = C_i^{-1}(2\pi)^{-1}|\mathbf{H}_i\boldsymbol{\Sigma}\mathbf{H}_i'|^{-1/2} \times \exp\left[-\frac{1}{2}(\mathbf{Z}_i - \mathbf{H}_i\mathcal{B})'(\mathbf{H}_i\boldsymbol{\Sigma}\mathbf{H}_i')^{-1}(\mathbf{Z}_i - \mathbf{H}_i\mathcal{B})\right] \quad (1)$$

for $X_i\beta_i^b < T_i < (1 - X_i) + X_i\beta_i^b$ and $0 < \beta_i^b < 1$ and 0 otherwise. Applying the same transformation to the definition of C_i we have

$$C_i^* = (2\pi)^{-1}|\mathbf{H}_i\boldsymbol{\Sigma}\mathbf{H}_i'|^{-1/2}\int_0^1\int_{X_i\beta^b}^{(1-X_i)+X_i\beta^b} \times \exp\left[-\frac{1}{2}(\mathbf{Z}_i - \mathbf{H}_i\mathcal{B})'(\mathbf{H}_i\boldsymbol{\Sigma}\mathbf{H}_i')^{-1}(\mathbf{Z}_i - \mathbf{H}_i\mathcal{B})\right] dT d\beta^b. \quad (2)$$

Equations (1) and (2) establish that the distribution of (β_i^b, T_i) is truncated bivariate normal with support on a parallelogram P with vertices $\{(0,0), (0, 1 - X), (1, 1), (1, X)\}$ and having location parameters $\mathbf{H}_i\mathcal{B}$ and dispersion parameters $\mathbf{H}_i\boldsymbol{\Sigma}\mathbf{H}_i'$. Let g be an untruncated bivariate normal density such that

$$f(\mathbf{Z}_i) = \begin{cases} C_i^{*-1}g(\beta_i^b, T_i) & \text{if } (\beta_i^b, T_i) \in P \\ 0, & \text{otherwise} \end{cases}$$

Suppressing the i subscripts, $\beta^b|T$ is univariate truncated normal if

$$\begin{aligned}
\frac{g(\beta^b|T)}{\int_l^u g(\beta^b|T)d\beta^b} &= \frac{f(\mathbf{Z})}{\int_l^u f(\mathbf{Z})d\beta^b} \\
&= \frac{C^{*-1}g(\beta^b|T)g(T)}{\int_l^u C^{*-1}g(\beta^b|T)g(T)b\beta^b} \\
&= \frac{C^{*-1}g(\beta^b|T)g(T)}{C^{*-1}g(T)\int_l^u g(\beta^b|T)d\beta^b} \\
&= \frac{g(\beta^b|T)}{\int_l^u g(\beta^b|T)d\beta^b},
\end{aligned}$$

where $u = \min[1, T/X]$ and $l = \max[0, (T + X - 1)/X]$ completing the proof.

Jeffrey B. LEWIS
University of California, Los Angeles

REFERENCES

King, G. (1997), *A Solution to the Ecological Inference Problem: Reconstructing Individual Behavior from Aggregate Data*, Princeton, NJ: Princeton University Press.

REPLY

Lewis' method of proof (and hence proof) is invalid. Let Ψ be a 2 by 2 positive definite matrix, θ a 2 by 1 matrix, I_P be an indicator function which is one on a set P in R^2 and zero otherwise, and define $\gamma_1^2 = \Psi_{11} - \Psi_{21}\Psi_{22}^{-1}\Psi_{12}$, $\gamma_2^2 = \Psi_{22}$, $\mu_1(T) = \theta_1 + \Psi_{12}\Psi_{22}^{-1}(T - \theta_2)$, and $\mu_2 = \theta_2$, where Ψ_{ij} is the ijth element of Ψ. Then, dropping all precinct subscripts i and using the definitions in Lewis' letter, Lewis' reasoning can be stated as follows:

$$\begin{aligned}
f(\mathbf{Z}) &\propto |\Psi|^{-1/2}\exp\left[-\frac{1}{2}(\mathbf{Z} - \theta)'\Psi^{-1}(\mathbf{Z} - \theta)\right] I_P \quad (1) \\
&\propto \left\{\frac{1}{\gamma_1}\exp\left[-\frac{1}{2}\left(\frac{\beta^b - \mu_1(T)}{\gamma_1}\right)^2\right]\right\} \\
&\quad \times \left\{\frac{1}{\gamma_2}\exp\left[-\frac{1}{2}\left(\frac{T - \mu_2}{\gamma_2}\right)^2\right]\right\} I_P \quad (2) \\
&\propto g_1(\beta^b, T)g_2(T)I_P \quad (3) \\
&\propto g(\beta^b|T)g(T)I_P. \quad (4)
\end{aligned}$$

Equation (1) is Lewis' equation (1), with $\theta = \mathbf{H}\mathcal{B}$ and $\Psi = \mathbf{H\Sigma H'}$, and P being the parallelogram defined by Lewis. Equation (2) is a factorization of (1) which is shown in such elementary texts as Hogg and Craig (1970, sec. 3.5). Equation (3) reexpresses factorization as functions, where g_1 is the function defined by the expression within the first set of braces and g_2 is the expression within the second set. Equation (4) is the crux of Lewis' argument. He would like the g to represent probability densities, and in particular, $g(\beta^b|T)$ to be the univariate normal conditional on the T.

Put this way the error is clear; one cannot go from (3) to (4). For factorization to be a valid method of obtaining conditional and marginal densities one needs to know one of these densities as well as the bivariate density, and Lewis knows neither the conditional nor the marginal density. Rather, he infers both simultaneously, apparently by analogy with the usual conditional and marginal densities of the untruncated bivariate normal, and while there of course exists a factorization of the truncated bivariate normal into the densities of $\beta^b|T$ and T, it is not necessarily represented by the g in (4).

To show Lewis' factorization cannot represent the conditional and marginal densities, assume $\beta^b|T$ is univariate truncated normal as claimed by Lewis. Then defining $J_{[0,1]}$ as one if $T \in [0,1]$, zero otherwise, since $f(T) = f(\beta^b, T)/f(\beta^b|T)$ and both of these right-hand densities have now been defined by Lewis, one has

$$f(T) \;\propto\; \left\{ \frac{1}{\gamma_2} \exp\left[-\frac{1}{2}\left(\frac{T-\mu_2}{\gamma_2}\right)^2 \right] \right\} J_{[0,1]},$$

which is a truncated univariate normal distribution. The impossibility is that since $X\beta^b$ and $(1-X)\beta^w$ are both truncated univariate normal, $T = X\beta^b + (1-X)\beta^w$ cannot be truncated univariate normal. To see this, note the Laplace transform (denoted by $\mathcal{L}$) of a normal with parameters (μ, σ) which is truncated on the interval $[a, b]$ is of the form $[\Phi(d) - \Phi(c)]^{-1} \exp[\mu t + \sigma^2 t^2/2][\Phi(d - \sigma t) - \Phi(c - \sigma t)]$, where $d = (b-\mu)/\sigma$ and $c = (a - \mu)/\sigma$ (this can be checked by differenting with respect to t and observing that the first moment evaluated at zero gives the same value for the mean as formula (79) in Johnson and Kotz (1970)). Then the Laplace transforms of $X\beta^b$ and $(1-X)\beta^w$ are of this form, and, under the assumption of independence of (β^b, β^w), one has

$$\begin{aligned}\mathcal{L}[T] &= \mathcal{L}[X\beta^b + (1-X)\beta^w] = \mathcal{L}[X\beta^b]\mathcal{L}[(1-X)\beta^w] \\ &= \exp\left[(\mu_b + \mu_w)t + (\sigma_b^2 + \sigma_w^2)t^2/2\right] \\ &\quad \frac{[\Phi(d_b - \sigma_b t) - \Phi(c_b - \sigma_b t)][\Phi(d_w - \sigma_w t) - \Phi(c_w - \sigma_w t)]}{[\Phi(d_b) - \Phi(c_b)][\Phi(d_w) - \Phi(c_w)]},\end{aligned}$$

where the subscript b refers to quantities related to $X\beta^b$ and w refers to quantities related to $(1-X)\beta^w$. The Laplace transform of T is not of the same form as a univariate truncated normal and thus by the uniqueness of the Laplace transform T is not truncated univariate normal as required by Lewis' claims and methods of proof.

Three other points. First, Lewis states that I offered no demonstration of the failure of law of iterated expectations, that is, of the failure of $E[E[\beta^b|T]] = \mathcal{B}^b$, where $\mathcal{B}^b$ is the mean of the truncated β^b. In fact, I derived an analytic formula (p. 108) for the actual value of $E[E[\beta^b|T]]$ under the assumption that $\beta^b|T$ is truncated univariate normal, as claimed by King and Lewis. It seems obvious that this analytical expression cannot be identically $\mathcal{B}^b$ for all parameter values of the underlying bivariate normal and all possible values of the X, and one point where the above equality does not hold is sufficient to show that the distribution $\beta^b|T$ cannot be univariate truncated normal. This point may be obtained by straightforward numerical techniques, and such a point exists when $X = .45$ and the underlying bivariate normal distribution has means (.33, .66) for the two variates and a covariance matrix equal to the identity. Then $\mathcal{B}^b = .3523$ and $E[E[\beta^b|T]] = .4674$ so T is not truncated univariate normal.

Second, there is one error in the article which has come to my attention (thanks to David James of Indiana University), and that is that the linear estimator of the $E[\beta^b|T]$can sometimes fall outside of the interval $[0, 1]$. I had stated that the extreme points of (T, X) could be mapped into $[0,1] \times [0,1]$, which is true, but since the mapping is nonlinear in X this is not enough to ensure the mapping stays within the bounds of $[0,1] \times [0,1]$. The simplest solution is to truncate the linear estimator when this happens. James has written a paper he is currently submitting for publication comparing the King estimator and the (corrected) linear estimator and finds that for the datasets used by King, the agreement between the two estimators is quite good.

Finally, Lewis states that my claim of King's miscalculation of $\beta^b|T$ is a "substantial error . . . that cuts right to the heart of the usefulness of King's method." Let us be clear—even had King correctly calculated $\beta^b|T$ it would have no bearing on the theoretical basis of EI (and hence its usefulness). The substantial error with EI (which is not disputed by Lewis) is the failure of King to recognize the relationship between his EI method and prediction. One example of how this failure leads to bad statistical practice is the technique of using the predicted precinct values from EI as dependent variables in additional regressions as in, say, Burden and Kimball (1998) (this technique was suggested by King—see p. 279). These estimates of $E[\beta_i^b|T]$, which are called $\hat{\beta}_i^b$ (the i indicates a precinct indexing), are essentially functions of the residual and making them dependent variables in a regression on some set of exogenous variables is just regressing random noise. As I showed in the original article, the linear estimator of $E[\beta^b|T]$ can be written in a form $\hat{\beta}_i^b = \mathcal{B}^b + \alpha_i[T_i - \mu_i]$, $\mu_i = E[T_i]$ and $\alpha_i = \text{cov}[\beta_i^b, T_i]/\text{var}[T_i]$. Then regressing the $\hat{\beta}_i^b$'s on an exogenous set of variables (call them $Z_i = (Z_{i1}, \ldots, Z_{ig})$), and letting τ be a g by 1 vector, Z be the p by g stacked matrix of the Z_i, $\hat{\beta}^b$ be the p by 1 stacked matrix of the $\hat{\beta}_i^b$'s, and v be a p by 1 stacked matrix of error terms, one has that the OLS estimate of $\hat{\beta}^b = Z\tau + v$ is

$$\hat{\tau} = [Z'Z]^{-1} Z'\hat{\beta}^b. \tag{5}$$

To determine the values of $\hat{\tau}$, express $\hat{\beta}^b$ as $1\mathcal{B}^b + \Lambda[T - \mu_T]$, where 1 is a p by 1 matrix of 1's, Λ is a p by p matrix with α_i in the ith diagonal position, zero for other entries, T is the stacked p by 1 matrix of the T_i's, and μ_T is the stacked p by 1 matrix of the μ_i's. Then (5) may be reexpressed as

$$\hat{\tau} = [Z'Z]^{-1} Z'1\mathcal{B}^b + [Z'Z]^{-1} Z'\Lambda[T - \mu_T].$$

To determine the asymptotic behavior of $\hat{\tau}$, use the fact that $p^{-1}p = 1$ and take plims of both sides, so that

$$\begin{aligned}\text{plim}\,\hat{\tau} &= \text{plim}\left\{ \left[\frac{Z'Z}{p}\right]^{-1} \frac{Z'1}{p}\mathcal{B}^b \right\} \\ &\quad + \text{plim}\left\{ \left[\frac{Z'Z}{p}\right]^{-1} \frac{Z'\Lambda[T-\mu_T]}{p} \right\} \\ &= \Sigma_{Z'Z}^{-1}\mu_Z\mathcal{B}^b + \Sigma_{Z'Z}^{-1}\Sigma,\end{aligned}$$

where $\Sigma_{Z'Z}^{-1} = \text{plim}[(Z'Z)/p]^{-1}$, $\mu_Z = \text{plim}(Z'1)/p$, and $\text{plim}(Z'\Lambda[T - \mu_T])/p = \Sigma$ (all these plims are assumed to exist). As the plim of $\hat{\tau}$ depends upon distributions of the exogenous variables Z and the residuals $T_i - \mu_i$, it is clear that no substantive interpretation of the $\hat{\tau}$ is possible. This type of procedure would not be used if the nature of the $\beta^b|T$ were understood.

Historically, there have been many models which treat the coefficients as random and thus lend themselves to a straightforward application of the theory of prediction (see Hawkes 1969; Brown and Payne 1986; Lupia and McCue 1990; and Tam 1995 for a few such applications). This theory was not applied because no one thought the analysis of these case-level predicated values was more important or much different than residual analysis (in fact, Morrison (1976), who derived $\beta^b|T$ for the multivariate normal case, referred to the analysis of predicted values as an application of the field of residual analysis). What is necessary (in my opinion) for EI to live up to the claims of its adherents is to produce an analysis making an actual distinction between the residual values and predicted values in the EI formulation and show how the use of these predicted values is actually an improvement over results from the voluminous research literature on residual analysis. This type of analysis would start the process of creating the conditions for a reconciliation of EI to the statistical literature.

Kenneth F. McCue
California Institute of Technology

REFERENCES

Brown, P., and Payne, C. (1986), "Aggregate Data, Ecological Regression, and Voting Transitions," *Journal of the American Statistical Association*, 81, 452–460.

Burden, B. C., and Kimball, D. C. (1998), "A New Approach to the Study of Ticket Splitting," *American Political Science Review*, 92, 533–544.

Hawkes, A. G. (1969), "An Approach to the Analysis of Electoral Swing," *Journal of the Royal Statistical Society*, Ser. A, 132, 68–79.

Hogg, R. V., and Craig, A. T. (1970), *Introduction to Mathematical Statistics*, London: The MacMillan Company.

King, G. (1997), *A Solution to the Ecological Inference Problem*, Princeton, NJ: Princeton University Press.

Johnson, N. L., and Kotz, S. (1970), *Continuous Univariate Distributions—1*, Boston: Houghton Mifflin.

Lupia, A., and McCue, K. (1990), "Why the 1980s Measures of Racially Polarized Voting are Inadequate for the 1990s," *Law & Policy*, 12, 355–388.

McCue, K. (2001), "The Statistical Foundations of the EI Method," *The American Statistician*, 55, 106–110.

Morison, (1976), *Multivariate Statistical Methods*, New York: McGraw Hill.

Tam, W. (1995), "Asians: A Monolithic Voting Block?", *Political Behavior*, 17, 223–249.