Response to McCue’s Reply to Lewis’s Comment on “The Statistical Foundations of the EI Method”

Jeffrey B. Lewis
Department of Political Science
University of California, Los Angeles
August 2002

In his reply to my comment, McCue repeats the same error he made in his original paper. To quickly restate the central issue, let \( \hat{\beta}^b \) and \( \hat{\beta}^w \) be distributed truncated normal on the unit square and let \( T = \beta^b X + \beta^w (1 - X) \) for a some constant \( X \in (0, 1) \). The question is: what is the distribution of \( \hat{\beta}^b \) given \( T \)? King (1997) claims that \( \hat{\beta}^b | T \) is univariate truncated normal on the interval \( (\max(0, (T + X - 1)/X), \min(1, T/X)) \). McCue’s paper disputes this claim. My comment shows that King’s assertion is correct. McCue’s reply to my comment rebuts my proof and reasserts McCue’s original objection to King’s claim about \( \hat{\beta}^b | T \).

My proof presented in a comment in the TAS directly derived the joint distribution of \( \beta^b \) and \( T \) and then derived the conditional distribution of \( \beta^b | T \) from that joint distribution. In what follows, I show why McCue’s objection to my proof is groundless. However, before working through the algebra, consider a few Monte Carlo simulations that demonstrate my contention. The point of departure for these Monte Carlo experiments is the joint distribution of \( T \) and \( \beta^b \). Monte Carlo techniques for drawing conditional distributions directly from joint distributions are well-known and widely employed in the simulation literature. McCue does not contest my derivation of the joint distribution of \( \beta^b \) and \( T \) and, therefore, he should not object to sampling \( \beta^b | T \) directly from this joint distribution. Figure 1 shows histograms.

---


2The simulations presented here use rejection sampling in which the joint density is used as the kernel of the conditional density. The true truncated normal conditional distribution (which McCue contests) is a convenient covering distribution from which to draw proposal values for the rejection sampling. See
Two Examples of The conditional distribution of $\beta^b|T$

Figure 1: Each panel shows a histogram of 100,000 simulated values of $\beta^b|T$ with a corresponding truncated normal density superimposed. In the left-hand panel the parameters of the underlying EI model are those used by McCue in his numerical example. Location parameters of the distribution of $\beta^b$ and $\beta^w$ are (0.33, 0.66), the dispersion matrix is equal to the identity matrix, $X = 0.45$, and $T = 0.8$. In the right-hand panel, the same means were used but the dispersion matrix was set equal to $\begin{bmatrix} 0.010 & 0.002 \\ 0.002 & 0.010 \end{bmatrix}$, $X = 0.8$, and $T = 0.1$. Note that in both panels the truncated normal describes the distribution of the simulated data.
of the results of two experiments in which values of $\beta^b|T$ were sampled in this way. The solid lines superimposed on the histograms are the truncated normal densities which McCue claims do not describe the distribution of these data. Like all simulations these two only show that a truncated normal distribution describes the variation in $\beta^b|T$ for two sets of parameter values. Of course, it is hypothetically possible that other values would yield conditional distributions that were not truncated normal, supporting McCue’s claim. I am quite confident that no such values exist, but I challenge the reader to prove me wrong and will gladly provide a simple computer program that can be used to simulate these conditional distributions for any parameter values.\textsuperscript{3}

I now turn to McCue’s claim that $\beta^b|T$ is not truncated normal. Using standard change of variables techniques, it is straight-forward to show that $\beta^b$ and $T$ are truncated bivariate normal on a parallelogram $P = \{(\beta^b,T) \text{ s.t. } 0 < \beta^b < 1 \& \beta^b X < T < 1 - (1 - \beta^b) X\}$. In his reply to my comment, McCue accepts this formulation of the joint distribution. The following derivation is presented in McCue’s reply:

\begin{align}
  f(Z) & \propto |\Psi|^{-1/2} \exp \left[-\frac{1}{2}(Z - \theta)' \Psi^{-1}(Z - \theta)\right] I_p \\
  & \propto \left\{ \frac{1}{\gamma_1} \exp \left[-\frac{1}{2} \left(\frac{\beta^b - \mu_1(T)}{\gamma_1} \right)^2 \right] \right\} \times \left\{ \frac{1}{\gamma_2} \exp \left[-\frac{1}{2} \left(\frac{T - \mu_2}{\gamma_2} \right)^2 \right] \right\} I_p \\
  & \propto g_1(\beta^b|T) g_2(T) I_p \\
  & \propto g_1(\beta^b|T) g_2(T) I_p
\end{align}

where $Z = (\beta^b, T)$, $\theta$ and $\Psi$ are the location and dispersion parameters of the joint (truncated bivariate normal) distribution of $\beta^b$ and $T$. Further, $\mu_1(T) = \theta_1 + \frac{\psi_{12}}{\psi_{22}} (T - \theta_2)$, $\gamma_1^2 = \psi_{11} - \frac{\psi_{12}^2}{\psi_{22}}$, $\mu_2 = \theta_2$, $\gamma_2^2 = \psi_{22}$, and $I_p$ is an indicator function that defines the region of support for the distribution. The first line of the derivation is the joint distribution found using the change of variables. The second line is a common factoring of the function in the first line. In the third line, the first bracketed term of the second line is defined as $g_1(\beta^b|T)$ and the second bracketed term of the second line is defined as $g_2(T)$. McCue accepts the first

\textsuperscript{3}The program written in the R computer language is available at www.bol.ucla.edu/~jblewis. The program allows the user to pick the parameters of the underlying model and the value of $T$ to condition on. It returns a histogram with the truncated normal density superimposed and a table comparing quantiles of the simulated draws and the corresponding truncated normal distributions.

3
three lines. What he objects to is the fourth line in which \( g_1(\beta^b, T) \) is written as \( g_1(\beta^b|T) \) and \( g_1 \) and \( g_2 \) are taken to be normal density functions. McCue claims that because of the truncation, \( g_1 \) and \( g_2 \) cannot be “read off” as densities.

At first glance, this claim is bizarre. Consider line 2 of the above derivation, which McCue accepts. The factorization in line 2 isolates \( \bar{b} \) in the first bracketed term. Noting that conditional distributions are proportional to joint distributions and leaving out the second bracketed term of line 2 which is not a function of \( \bar{b} \),

\[
f(\beta^b|T) \propto \left\{ \frac{1}{\gamma_1} \exp \left[ -\frac{1}{2} \left( \frac{\beta^b - \mu_1(T)}{\gamma_1} \right)^2 \right] \right\} I_p.
\]

The right-hand side of this equation is a normal density function multiplied by an indicator function equal to one over an interval of values of \( \beta^b \). Therefore, \( f(\beta^b|T) \) is univariate truncated normal.¹ This conclusion follows directly from simple algebraic manipulations and does not turn on properties of untruncated bivariate normal distributions as McCue claims. The factorization in line 2 exists quite apart from whether the objects of the function \( f \) are random variables, much less random variables following a particular law. Thus, my proof and King’s original claim are correct.

The question then is how does McCue purport to show that \( \beta^b|T \) is not truncated normal? Rather than directly undermining the logic of the previous paragraph, McCue instead attacks what he believes is the another implication of lines 1 through 4. Focusing on the function \( g_2 \), McCue asserts that line 4 implies that the marginal distribution of \( T \) is univariate truncated normal. He then goes about showing that the marginal distribution of \( T \) is not truncated normal and concludes that the factorization in 4 is (somehow) invalid. This tact is the same one that McCue takes in his original article, though the method of proof in his reply to my comment differs from the method used in his article.²

¹Intersecting the set of \( (\beta^b, T) \) pairs that have a given value of \( T \) and the set that fall in \( P \), yields the region of support for the distribution of \( \beta^b \) given in my comment.

²In his original article, McCue argues that the law of iterated expectations would fail if \( T \) were truncated normal. In his reply to my comment, he shows that the Laplace transform of the marginal density of \( T \) differs from the Laplace transform of a truncated normal density. McCue also presents a numerical example of his law of iterated expectations approach in his reply to my comment. Though I was unable to replicate his findings, I was able to show that the law of iterations holds when the correct marginal distribution of \( T \)
is: the marginal distribution of $T$ is not univariate truncated normal, if $\beta^b|T$ is univariate truncated normal then $T$ is univariate truncated normal, therefore $\beta^b|T$ is not truncated normal. The weak link in the logical chain is the second. McCue is quite correct that the marginal distribution of $T$ cannot be “read off” of the factorization in line 4, however my proof in no way implies that it can.

McCue’s conclusion that my derivation of $f(\beta^b|T)$ implies that $f(T)$ is truncated normal is arrived at in the following way. Letting $c$ and $c_1$ be appropriate constants of proportionality, if $f(\beta^b, T) = cg(\beta^b, T)I_p$ where $g$ is a bivariate normal density and $f(\beta^b|T) = c_1g_1(\beta^b|T)$, then the well-known relation

$$f(T) = \frac{f(T, \beta^b)}{f(\beta^b|T)}$$

implies that

$$f(T) = \frac{cg(T, \beta^b)}{c_1g_1(\beta^b|T)}I_p$$

and, therefore,

$$f(T) = \frac{c}{c_1} \left( g(T, \beta^b) \right) g_1(\beta^b|T)I_p = \frac{c}{c_1} g_2(T)I_p.$$

This is certainly true, but does not imply $f(T) \propto g(T)I_p$, and thus that $T$ is truncated normal, as McCue claims. The reason is that the “constant” of proportionality $c_1$ is a function of $T$. Therefore, $f(T)$ is not proportional to $g(T)I_p$ and, thus, $f(T)$ is not truncated normal even though $f(\beta^b|T)$ is univariate truncated normal and $f(\beta^b, T)$ is bivariate truncated normal.

To find the true distribution of $T$, I marginalize the joint distribution (as written in line 4) with respect to $\beta^b$.

$$f(T) \propto \left\{ \int_{l(T)}^{u(T)} g_1(\beta^b|T) d\beta^b \right\} g_2(T)I_p$$

where $u(T) = \min(1, T/X)$ and $l(T) = \max(0, (T + X - 1)/X)$. Because, the bracketed integral is a function of $T$, $f(T)$ is not proportional to $g_2(T)$ (a normal density) and, thus, is not univariate truncated normal. A complete rendering of the marginal distribution of $T$ derived in this way, is given in Appendix D of King (1997).

(implied by my proof) is employed.
Marginal Distribution of $T$ Given the Parameters in McCue’s Example

Figure 2: Shows the histogram of 100,000 Monte Carlo draws from the distribution of $T$ given that $\beta^b$ and $\beta^w$ are drawn from a truncated bivariate normal distribution with underlying untruncated means equal to 0.33 and 0.66 and an untruncated variance-covariance matrix equal to the identity matrix. The solid line is the marginal density function of $T$ derived in the text. The dashed line is the density that McCue claimed was implied by my derivation.
The distribution of $T$ (as defined here and in my reply) given the parameter values suggested in an example by McCue is shown in Figure 2 along with 100,000 draws from the $T$ formed by drawing $\beta^b$ and $\beta^w$ from their assumed distribution and then letting $T = \beta^b X + \beta^w (1 - X)$. Note that the distribution is not truncated normal as McCue asserts that my proof implies. McCue’s claim that King’s original derivation of the distribution of $\beta^b|T$ and my subsequent derivation are flawed, both rest on the false assumption that the truncated normality of $\beta^b|T$ implies the truncated normality of $T$. Once it is made clear that the truncated normality of $T$ is not implied by my derivation, McCue’s objection no longer stands.

References