GEORGE CASELLA[1]

Cornell University

In his customary, penetrating way, Professor Brown has discovered and illuminated a fascinating admissibility paradox. This paradox brings together the elusive concept of ancillarity with the (still somewhat puzzling) Stein phenomenon and, through this synthesis, perhaps explains both a little better. The goal in this discussion is to understand and explain Brown's admissibility paradox in a simple intuitive way.

Using the notation of Section 3, we observe $Y_{n\times 1}$ and $V_{n\times r}$, where $EY = \alpha 1 + V\beta$, and we want to estimate α using an estimator that is a function of Y and V, say $d(Y, V)$. The loss function given by (3.1.2) is squared error loss

$$L(\alpha, d) = (\alpha - d)^2. \tag{1}$$

In a regression problem, we estimate α based on observing values $Y = y$ and $V = v$. Brown's paradox asserts that the admissibility of $\hat{\alpha}$, the least squares estimator, depends on whether V is treated as constant or as a realized value of an ancillary random variable.

An important distinction between the two problems lies in the risk functions: Although the loss function remains the same, the risk function changes depending on whether we consider the matrix V to be fixed or random. If V is fixed, then the risk of estimating α is conditional on the value $V = v$, that is,

$$R(\alpha, d|V = v) = E\Big[(\alpha - d(Y, v))^2 \Big| V = v\Big]. \tag{2}$$

Here the expectation is over the distribution of Y given $V = v$ which, of course, depends on α. If V is considered a random variable, then the risk of estimating α is unconditional on the value $V = v$, that is,

$$R(\alpha, d) = \int R(\alpha, d|V = v)\, f_V(v)\, dv, \tag{3}$$

where $f_V(\cdot)$ denotes the density of V.

Keeping the risk relationship (3) in mind, we can now reexamine the admissibility/inadmissibility results of Proposition 3.1.1 and Theorem 3.2.2 (or their predecessors, Proposition 2.1.1 and Theorem 2.1.2). The admissibility results relate to the risk function $R(\alpha, d|V = v)$ of (2), while the inadmissibility results relate to the risk function $R(\alpha, d)$ of (3). Furthermore, the relationship in (3) amplifies the paradoxical nature of Brown's results. Note that from (3) we immediately get the implication that if an estimator $d(Y, v)$ is dominated for every v by $d^*(Y, v)$ using $R(\alpha, d|V = v)$ of (2), it is inadmissible under $R(\alpha, d(Y, V))$ of (3). But this does not happen for $d(Y, v) = \hat{\alpha}$. Since the least squares estimator is admissible under $R(\alpha, \hat{\alpha}|V = v)$, this implies that it cannot be dominated in risk for every v by the same estimator.

[1]Research supported by NSF Grant No. DMS-89-0039.