Aris Spanos [Virginia Tech, Department of Economics]
The primary objective of this paper is to argue against the widely held thesis that: traditional statistical techniques cannot be applied to disciplines like economics because the available data are too heterogeneous, resulting from complex interactions of millions of agents in an on-going and highly complicated process; the economy.
The main argument is that such a thesis is misinformed for several reasons, the most crucial being:

[a] the theory of stochastic processes has greatly extended the intended scope of statistical modeling to include data that exhibit both heterogeneity and dependence, rendering the above potential heterogeneity problem testable, and

[b] the thesis conflates two different dimensions of empirical modeling:

(i) the substantive model $M_\varphi(z)$, aiming to shed light on [describe, explain, predict] on a phenomenon of interest, and

(ii) the statistical model $M_\theta(z)$ aiming to capture the chance regularities
exhibited by data $z_0$. The appropriateness of $\mathcal{M}_\theta(z)$ pertains only to whether the data $z_0$ constitute a ‘truly typical realization’ of the stochastic process \( \{Z_t, \ t=1, 2, \ldots, n, \ldots\} \) whose probabilistic structure is (implicitly) defined by the substantive model $\mathcal{M}_\varphi(z)$.

**Error statistics** [a refinement/extension of the Fisher-Neyman-Pearson approach] views empirical modeling as a *piecemeal process* that distinguishes between the **statistical** $\mathcal{M}_\theta(z)$ and **substantive model** $\mathcal{M}_\varphi(z)$, clearly delineating the following two questions:

- **(a) statistical adequacy:** does $\mathcal{M}_\theta(z)$ account for the chance regularities in $z_0$?
- **(b) substantive adequacy:** does the model $\mathcal{M}_\varphi(z)$ shed adequate light on (describe, explain, predict) the phenomenon of interest?

Substantive inadequacy can arise from impractical *ceteris paribus* clauses, confounding factors, false causal claims, etc.

- It’s one thing to claim that a model is *not* an exact picture of reality in a substantive sense (*realisticness*), and totally another to claim that this statistical model $\mathcal{M}_\theta(z)$ could *not* have generated data $z_0$ because the assumed probabilistic structure of the process $\{Z_t, \ t \in \mathbb{N}\}$ underlying this data is **invalid**.
Example 1. Simple measurement model:

\[ Z_t = \mu + \varepsilon_t, \; \varepsilon_t \sim \text{NIID}(0, \sigma^2), \; t = 1, 2, \ldots, n. \]

The implicit statistical model is the simple Normal:

\[ \mathcal{M}_\theta(z): Z_t \sim \text{NIID}(\mu, \sigma^2), \; \theta = (\mu, \sigma^2) \in \mathbb{R} \times \mathbb{R}_+, \; z_t \in \mathbb{R}, \; t \in \mathbb{N} \]

What goes wrong when \( \mathcal{M}_\theta(z) \) is statistically misspecified?

misspecified \( \mathcal{M}_\theta(z) \Rightarrow \text{false likelihood } L(\theta; z_0) \Rightarrow \text{Unreliable Inferences} \)

Bayesian: erroneous posterior \( \pi(\theta|z_0) = \pi(\theta) L(\theta; z_0) \).

Frequentist: [a] erroneous fit/prediction measures, [b] false error probabilities.
Why is the substantive vs. statistical distinction important?

(1) A substantive model $M_\phi(z)$ may always come up short in fully capturing or explaining a phenomenon of interest, but a statistical model $M_\theta(z)$ may be entirely adequate to reliably test and assess the appropriateness of $M_\phi(z)$, and/or provide reliable answers to any substantive adequacy questions of interest. Indeed, without statistical adequacy the answers to the substantive questions of interest will be unreliable.

(2) Learning from data depends crucially on establishing a sound link between the process generating data $z_0$ and the assumed $M_\theta(z)$, by securing statistical adequacy; the latter ensures the actual are close enough to the nominal (assumed) error probabilities.

(3) Statistical adequacy depends only on $M_\theta(z)$ and $z_0$ and can be established independently by different practitioners using thorough Mis-Specification testing.

(4) Good fit/prediction is neither necessary nor sufficient for statistical adequacy, but it is relevant for substantive adequacy in the sense that it provides a measure of the structural model’s comprehensiveness (explanatory capacity) vis-a-vis the phenomenon of interest.
What about the notion of a model being adequate for a purpose? That notion makes good sense only in relation to substantive adequacy. Unreliastic models can lead to good insights, but misspecified model would not; the reliability of any inferential claim based on a misspecified model will be questionable. The types of errors one needs to probe for and guard against are very different in the two cases. Substantive adequacy calls for additional probing of (potential) errors in bridging the gap between theory and data.

Contrary to the above claim, it is argued that the combination of the complex nature of economic phenomena and the observational nature of economic data call for more not less reliance on statistical methods. In the social sciences, where the actual DGM is highly complex and the available data are (often) observational, a statistically adequate $M_{θ}(z)$ could play a more crucial role in guiding the search for substantively adequate explanations (theories) by delineating ‘what there is to explain’.

It is argued that learning from observational data is possible when modeling and inference are viewed in the context of the error statistical framework that emphasizes the importance of probing, eliminating or ‘controlling’ potential errors.