14

Statistical Inference and Methodology in the Postwar Era

The fundamentals of statistical inference—frequentist, Bayesian, Fisherian —were set in place by the end of the first half of the twentieth century, as discussed in Part I of this book. The postwar era witnessed a massive expansion of statistical methodology, responding to the data-driven demands of modern scientific technology. We are now at the end of Part II, "Early Computer-Age Methods," having surveyed the march of new statistical algorithms and their inferential justification from the 1950s through the 1990s.

This was a time of opportunity for the discipline of statistics, when the speed of computation increased by a factor of a thousand, and then another thousand. As we said before, a land bridge had opened to a new continent, but not everyone was eager to cross. We saw a mixed picture: the computer played a minor or negligible role in the development of some influential topics such as empirical Bayes, but was fundamental to others such as the bootstrap.

Fifteen major topics were examined in Chapters 6 through 13. What follows is a short scorecard of their inferential affinities, Bayesian, frequentist, or Fisherian, as well as an assessment of the computer's role in their development. None of this is very precise, but the overall picture, illustrated in Figure 14.1, is evocative.

Empirical Bayes

Robbins' original development of formula (6.5) was frequentistic, but most statistical researchers were frequentists in the postwar era so that could be expected. The obvious Bayesian component of empirical Bayes arguments is balanced by their frequentist emphasis on (nearly) unbiased estimation of Bayesian estimators, as well as the restriction to using only current data for inference. Electronic computation played hardly any role in the theory's development (as indicated by blue coloring in the figure). Of course mod-

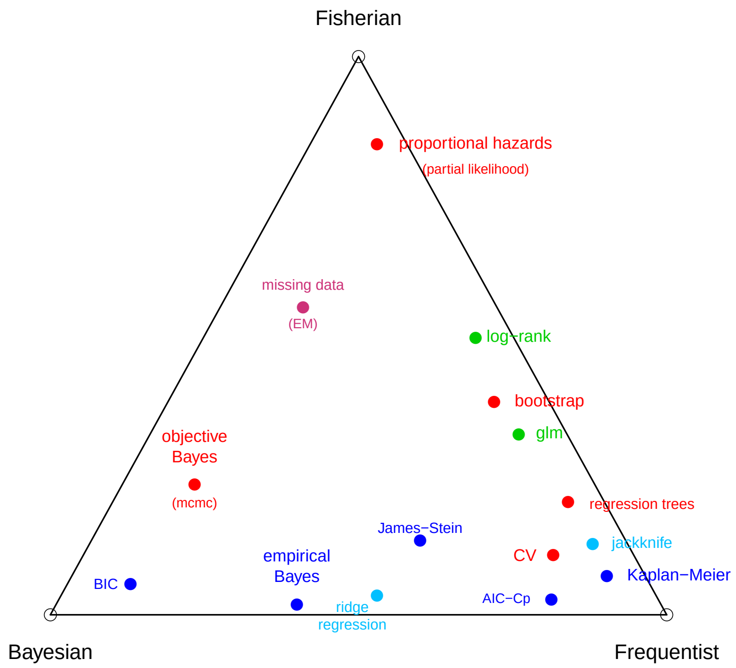

Figure 14.1 Bayesian, frequentist, and Fisherian influences, as described in the text, on 15 major topics, 1950s through 1990s. Colors indicate the importance of electronic computation in their development: red, crucial; violet, very important; green, important; light blue, less important; blue, negligible.

ern empirical Bayes applications are heavily computational, but that is the case for most methods now.

James–Stein and Ridge Regression

The frequentist roots of James–Stein estimation are more definitive, especially given the force of the James–Stein theorem (7.16). Nevertheless, the empirical Bayes interpretation (7.13) lends James–Stein some Bayesian credibility. Electronic computation played no role in its development. This was less true for ridge regression, colored light blue in the figure, where the matrix calculation (7.36) would have been daunting in the pre-electronic age. The Bayesian justification (7.37)–(7.39) of ridge regression carries more weight than for James–Stein, given the absence of a strong frequentist theorem.

Generalized Linear Models

GLM development began with a pronounced Fisherian emphasis on likelihood[1] modeling, but settled down to more or less standard frequentist regression theory. A key operational feature, low-dimensional sufficient statistics, limited its computational demands, but GLM theory could not have developed before the age of electronic computers (as indicated by green coloring).

Regression Trees

Model building by means of regression trees is a computationally intensive enterprise, indicated by its red color in Figure 14.1. Its justification has been mainly in terms of asymptotic frequentist properties.

Survival Analysis

The Kaplan–Meier estimate, log-rank test, and proportional hazards model move from the frequentist pole of the diagram toward the Fisherian pole as the conditioning arguments in Sections 9.2 through 9.4 become more elaborate. The role of computation in their development increases in the same order. Kaplan–Meier estimates can be done by hand (and were), while it is impossible to contemplate proportional hazards analysis without the computer. Partial likelihood, the enabling argument for the theory, is a quintessential Fisherian device.

Missing Data and the EM Algorithm

The imputation of missing data has a Bayesian flavor of *indirect evidence*, but the "fake data" principle (9.44)–(9.46) has Fisherian roots. Fast computation was important to the method's development, particularly so for the EM algorithm.

Jackknife and Bootstrap

The purpose of the jackknife was to calculate frequentist standard errors and biases. Electronic computation was of only minor importance in its

[1] More explicitly, *quasilikelihoods*, an extension to a wider class of exponential family models.

development. By contrast, the bootstrap is the archetype for computer-intensive statistical inference. It combines frequentism with Fisherian devices: plug-in estimation of accuracy estimates, as in (10.18)–(10.19), and correctness arguments for bootstrap confidence intervals, (11.79)–(11.83).

Cross-Validation

The renaissance of interest in cross-validation required fast computation, especially for assessing modern computer-intensive prediction algorithms. As pointed out in the text following Figure (12.3), cross-validation is a *strongly* frequentist procedure.

BIC, AIC, and C_p

These three algorithms were designed to *avoid* computation, BIC for Bayesian model selection, Section (13.3), AIC and C_p for unbiased estimation of frequentist prediction error, (12.76) and (12.50).

Objective Bayes and MCMC

In addition to their Bayesian provenance, objective Bayes methods have some connection with fiducial ideas and the bootstrap, as discussed in Section 11.5. (An argument can be made that they are at least as frequentist as they are Bayesian—see the notes below—though that has not been acted upon in coloring the figure.) Gibbs sampling and MCMC, the enabling algorithms, epitomize modern computer-intensive inference.

Notes

Figure 14.1 is an updated version of Figure 8 in Efron (1998), “R. A. Fisher in the 21st Century.” There the difficulty of properly placing *objective Bayes* is confessed, with Erich Lehmann arguing for a more frequentist (decision-theoretic) location: “In fact, the concept of uninformative prior is philosophically close to Wald’s least favorable distribution, and the two often coincide.”

Figure 14.1 shows a healthy mixture of philosophical and computational tactics at work, with all three edges (but not the center) of the triangle in play. All new points will be red (computer-intensive) as we move into the twenty-first century in Part III. Our triangle will have to struggle to accommodate some major developments based on machine learning, a philosophically atheistic approach to statistical inference.

Part III

Twenty-First-Century Topics