STEWART, M. B. and WALLIS, J. F. (1981). *Introductory Econometrics*, 2nd ed. Blackwell, Oxford.
STONE, C. J. (1981). Admissible selection of an accurate and parsimonious normal linear regression model. *Ann. Statist.* **9** 475–485.
SWAMY, P. A. V. B. and VON ZUR MUEHLEN, P. (1987). Further thoughts on causality with econometric models. Special Study Paper. Federal Reserve Board, Washington.
THEIL, H. (1971). *Principles of Econometrics*. Wiley, New York.
WOLINS, L. (1978). Sex differentials in salaries: Faults in analysis of covariance. *Science* **200** 723.

Comment

Gail Blattenberger

I am pleased to see attention directed to the theoretical principles underlying the statistical reasoning process in legal proceedings. Arthur Dempster has raised important questions regarding statistical analysis in employment discrimination cases. Too often statistical advice has been given in legal cases purporting to estimate a discrimination effect without explicit statement or understanding of what is being estimated. Here I offer comment on statistical evidence of discrimination in the legal context. I focus on interpretational questions rather than specific model specification issues such as reverse regression.

Continued active controversies over the meaning of probability render a universally accepted standard definition impossible. Nonetheless, the understanding of probabilistic language has implications for the interpretation of the evidential content of the analysis. Explicit statement of probabilistic modeling assumptions becomes necessary for communication not only between statistical experts and lawyers, but even among statisticians. In this I agree with Dempster.

My own position on probability adopts basically the personal measure of uncertainty meaning that Dempster advocates in this context, but my position is perhaps more extreme in this direction than his. Following de Finetti, I view probabilities as representations of uncertain opinions about the value of unknown but observable quantities. In this context it becomes important to specify whose opinions the probabilistic structure represents and under what circumstances. Probability for me is not a physical property, and estimation of unknown and inherently unmeasurable constructs lacks substance. This has relevance for the specification and interpretation of the probabilistic model.

Within this probabilistic perspective a linear model of the form specified in Dempster's equations (1) and (2) might represent a linear belief structure of an analyst. This model is specified:

$$Y_i = G_i\alpha + X_i\beta + \theta_i. \tag{1}$$

I will play the role of the analyst. From my position equation (1) denotes the process by which I base my opinion about the measurable value of Y_i, employee i's salary, given the measurable values of G_i, the ith employee's gender, and given X_i, a vector of other measured characteristics of the ith employee. Although this equation has the same form as the standard model in Dempster's discussion, the interpretation is different.

Dempster expands his model to include information known by the employer but not the statistician, X^*, and a more comprehensive vector of characteristics, X^{**}, needed to determine the employee's "true worth," Y^{**}. Undoubtedly, the employer does use information available to him, but unknown to me in setting salaries; it is also true that the employer may provide nonmonetary fringe benefits which are unknown to me. I could incorporate recognition of this into my belief structure. I would question, however, the role of unmeasurable or unmeasured characteristics, X^{**}, the existence of the unmeasured and inherently unmeasurable, Y^{**}, and its expected value, Y^*. Dempster admits that the realism of these concepts is questionable, but he assumes that they exist. He proceeds to develop a model based on these concepts and examines its implications for assessing discrimination.

Economists, adhering to the human capital approach, have used the idea of an individual's marginal productivity to indicate the "true worth" of that employee. I have argued elsewhere, Blattenberger and Michelson (1984), that individual marginal productivity is not an intrinsic property possessed by an individual. It is inherently unmeasurable. Dempster does not use the term marginal productivity, but the same arguments are applicable to "true worth." I personally have had recent experience with this issue. In response to state budget cuts, I have participated in a committee

Gail Blattenberger is Associate Professor, Department of Economics, University of Utah, Salt Lake City, Utah 84112.

Institute of Mathematical Statistics is collaborating with JSTOR to digitize, preserve, and extend access to *Statistical Science*. www.jstor.org

setting work standards within my department. Classes taught and articles published were agreed to be important components of any index of work; these were simply generally acceptable, observable measures. Disputes were, of course, rampant on relative weights and specific definitions. No one among the academic economists on the committee believed that there existed a "true worth" of which we could define an unbiased measure. All information available to the employer was available to the department and more.

In assessing discrimination I reassert that the equation (1) is a potential representation of my (the analyst's) belief structure. It does not represent the reality of how salaries are set at a firm. In this I agree with Dempster; I do not accept the implicit causal structure in the standard model of the employer applying some mathematical formula and operating a chance mechanism to determine an employee's salary. What my belief structure implies is that observable characteristics provide me information on worker's salary levels. Additional information might lead me to change my belief structure. For example obtaining information concerning post-and-bid data on filling open positions would induce a different opinion structure. The accuracy of the predictions of alternative belief structures could be empirically evaluated.

Dempster assesses discrimination in a parametric context. He distinguishes theoretically among α, α^* and the components of α^*, α' and α''. He shows that α^* cannot be decomposed empirically into α' and α'' with his model structure. But even the value of α is dependent in Dempster's representation on effectively infinite populations; it is inherently unknown and unknowable. The function of α in the linear belief structure defined above is limited to locating my predictive distribution. It is a nuisance parameter which can be integrated out of this distribution. A "true" value of α does not exist even theoretically. In order for a unique value of α to exist applied uniformly over the relevant time period by personnel officers and plant managers, it would have to be established in explicit company policy. The job of a statistician would be merely to measure how effectively this policy was carried out.

Admittedly, this position evades any determinant hard measure of discrimination and appears unsatisfactory in a legal context aspiring to determinacy. Lawyers in discrimination cases are accustomed to testimony on statistically significant measures of discrimination. Statisticians put up a facade of scientific rhetoric on the accuracy of their measures. Lawyers, however, would not find the softness of measures of uncertainty surprising in light of the many contradictory expert testimonies introduced in court. Lawyers are themselves masters of subtle distinctions in language; interpretational debates are not foreign to them. Furthermore, legal theory has also confronted disputes regarding faith in its foundational tenets. The confrontation set in motion by the field of critical legal studies brings into question the determinacy of the law and provides a framework for considering the role of statistical testimony in court.

Critical legal studies challenges the idealized model of the legal process presuming a distinct form of legal reasoning which when properly effected produces a unique correct decision. The law, in this view, is not immutable and absolute; indeterminacy in the interpretation of legal rules is prevalent. Legal reasoning itself is asserted to serve a legitimization function for the dominant social structure. This school has increased in prominence in recent years (see Kairys, 1982; Levinson, 1986).

Statisticians, viewing the law as absolute, are fond of using a judicial analogy in introducing the concept of type one and type two errors in introductory statistics courses. Type one error in this analogy corresponds to the probability of acquittal when guilty, type two to the probability of conviction when innocent. De Finetti (1980) has argued "probability does not exist." A critical legal theorist might paraphrase de Finetti and say "absolute guilt does not exist. We need to ask not only guilty of what, but also in what circumstances and evaluated by whom." Conviction is an observable; guilt is not. This understanding of the law complements the subjectivist position.

Returning to the specific application of wage discrimination, I consider how discrimination is understood within the framework of critical legal theory. Then I ask how to place Dempster's model in this context and how statistics can contribute to this understanding.

To a critical legal theorist the legal meaning of discrimination is a reflection of societal forces and has not been static. A fundamental distinction in this view of discrimination law has been between the perpetrator and the victim perspective. The perpetrator perspective which has dominated discrimination law presumes the aberrant behavior of a few individuals. Nonperpetrators are innocent. The victim perspective emphasizes effecting results instead of seeking out villains. Institutional structure can engender inequality as well as individual actions. School busing was designed toward inducing societal changes rather than punishing criminals, thus taking the victim perspective. *Griggs v. Duke Power Co.* also adopted this perspective. A facially neutral testing procedure with no demonstrated connection to job performance was found to be discriminatory. This case has fallen under disrepute as has school busing, but both illustrate legal departures from the predominant perpetrator

perspective. The court interpretation of discrimination law has vacillated in response to societal pressures.

Legal interpretations of discrimination imply some causal structure even if not individually motivated. The fundamental problem of causal inference discussed by Holland (1986a), who is cited by Dempster, makes statisticians generally timid about causal inference. Absence of any statements regarding implicit causal structure reflects this timidity. Some explicit statement seems imperative for communication. The causal structure in the above model is part of the opinion structure of the analyst represented by the probabilistic model; it is not a representation of reality. This rules out a perpetrator perspective as not statistically demonstrable. This is not an evasion, but an attempt at frankness. It does not support a victim perspective.

Usefully Dempster has explicitly articulated the causal structure underlying his model. Dempster's causal model of discrimination to some extent incorporates both of the discrimination perspectives acknowledged in legal theory. The parameter α', called prejudicial discrimination, clearly presumes an intentional action, although it does not differentiate between firm policy and individual action. The parameter α'', termed judgmental discrimination, could be associated with the victim perspective. In Dempster's context, however, this presumes a profit maximizing firm. This firm or its personnel department can ascertain an unbiased measure of productivity based on the information available to it even if the statistician can't. Discrimination in critical legal theory is a more encompassing construct. It is not simply a matter of individual action; it includes the effects of societal and institutional structure.

This discussion might be thought to imply that statistical analysis of discrimination is vacuous. I do not think so. Rather it needs reinterpretation of its meaning and a diminution of its claims to measurement.

Dempster has noted elsewhere (1985) that coherent formal probabilistic models are not innate. A statistician might assist the legal decision maker in the construction of a coherent formal belief structure and rooting out its inconsistencies. The statistician does not prescribe which belief structure is "true." None exists. The decision maker is required to honestly assess his/her own beliefs. Given a specific probabilistic representation of the decision maker's beliefs the statistician might assist the legal decision maker by computing how it would be coherently updated based on the observed data. Sensitivity analysis to competing priors could be performed.

Coherence does not imply accuracy and does not help in selecting among alternative, contending, coherent probabilistic representations. The statistician can, however, report the scores that the belief structures would achieve in predicting the observed data according to "proper scoring rules," a concept developed by Savage. These rules are relevant to both eliciting opinions and evaluating predictions. Inaccurate performance would not alter dogmatic beliefs; that is invariably true. Nonetheless, this record of performance could help a legal decision maker in selecting a personally plausible probabilistic representation of his/her beliefs.

This discussion does not concern practices that are currently employed in court. It does respond to some of the issues raised by Dempster. It suggests that relevant operational practices might be developed in this context.

ADDITIONAL REFERENCES

BLATTENBERGER, G. and MICHELSON, S. (1984). Reverse regression and employment discrimination. *J. Bus. Econ. Statist.* **2** 121–122.

DE FINETTI, B. (1980). Probability: Beware of falsifications. In *Studies in Subjective Probability* (H. E. Kyburg, Jr. and H. E. Smokler, eds.) 193–223. Wiley, New York.

DEMPSTER, A. (1985). Probability, evidence, and judgement. In *Bayesian Statistics 2* (J. M. Bernardo, M. H. DeGroot, D. V. Lindley and A. F. M. Smith, eds.) 119–132. North-Holland, Amsterdam.

KAIRYS, D. (1982). *The Politics of Law: A Progressive Critique.* Pantheon Books, New York.

LEVINSON, S. (1986). Professing law: Commitment of faith or detached analysis. *St. Louis Univ. Law J.* **31** 3–71.