1980

Mathematical Models and Legal Realities: Some Comments on the Poisson Model of Jury Behavior

David H. Kaye
Penn State Law

Recommended Citation

David H. Kaye, *Mathematical Models and Legal Realities: Some Comments on the Poisson Model of Jury Behavior*, 13 *Conn. L. Rev.* 1 (1980).

CONNECTICUT
LAW REVIEW

VOLUME 13 FALL 1980 NUMBER 1

SPECIAL SECTION
MODELS AND LEGAL POLICY

MATHEMATICAL MODELS AND LEGAL REALITIES: REFLECTIONS ON THE POISSON MODEL OF JURY BEHAVIOR

*by David Kaye**

Mathematical models are great fun. With them, we can trace the development of the large-scale features of the universe,[1] the struggles of predator and prey,[2] the evolution of the sun and other stars,[3] and the path of the world economy into the year 2000.[4] Indeed, the range

---

* Professor of Law, Arizona State University, S.B. 1968, M.I.T.; A.M. 1969, Harvard University; J.D. 1972, Yale Law School. This work was supported by the Arizona State University Faculty Grant-In-Aid Program and the Cleary Research Fellowship Fund. The author wishes to acknowledge the assistance of Dennis Young, Dennis Karjala, and Alan Gelfand in the development of this article.

1. *E.g.*, Barrow & Silk, *The Structure of the Early Universe*, SCIENTIFIC AM., April 1980, at 118; Norman & Silk, *On the Formation and Evolution of Clumps of Galaxies in an Expanding Universe*, 224 ASTROPHYSICAL J. 293 (1978).
2. *E.g.*, Tanner, *The Stability and the Intrinsic Growth Rates of Prey and Predator Populations*, 56 ECOLOGY 855 (1975). *Cf.* Hirsch & Nasell, *The Transmission and Control of Schistosome Infections*, in MATHEMATICAL ANALYSIS OF DECISION PROBLEMS IN ECOLOGY 271 (A. Charnes & W. Lynn eds. 1975) (modeling parasite endemicity).
3. *E.g.*, D. CLAYTON, PRINCIPLES OF STELLAR EVOLUTION AND NUCLEOSYNTHESIS (1968).
4. Leontief, *The World Economy of the Year 2000*, SCIENTIFIC AM., Sept. 1980, at 206.

of phenomena that might be modeled seems bounded only by imagination and ingenuity. However, if little is known about the reality that the model is intended to simulate, even the most mathematically elegant model may not be very useful in practice. As Michael Finkelstein, perhaps the preeminent advocate of quantitative models in law, has cautioned:

> A mathematical model cannot reflect all the elements of reality and need not do so to produce usable estimates. However, modeling involves drastic simplifications, and care is needed to avoid conclusions that are wide of the mark. In particular, there is a tendency—which must be scrutinized in each case—to sweep away complexity to permit mathematical accessibility. This drive for quantification sometimes tempts the mathematician to ignore or reject complicating factors that are nonetheless essential to the legal picture. The lawyer may be an unwitting accomplice in this process if he limits his role to a passive understanding of his expert's work, and fails to pursue a critical evaluation from a carefully focused legal point of view. The records of judicial and administrative proceedings are already strewn with the disasters that this uncritical acceptance sometimes allows.[5]

I wish to underscore and illuminate this warning by considering in more detail than Finkelstein has done[6] a mathematical model that seems to be in vogue among mathematicians,[7] economists,[8] and political scientists[9] interested in jury decisionmaking. This model began

5. M. FINKELSTEIN, QUANTITATIVE METHODS IN LAW 11-12 (1978) (footnote omitted).

6. *Id.* at 11 n.27.

7. Gelfand & Solomon, *Considerations in Building Jury Behavior Models: An Argument in Favor of the Twelve Member Jury*, 17 JURIMETRICS J. 292 (1977); Gelfand & Solomon, *An Argument in Favor of 12-Member Juries*, in MODELING THE CRIMINAL JUSTICE SYSTEM 205 (S. Nagel ed. 1977); Gelfand & Solomon, *Rejoinder*, 72 J. AM. STATISTICAL A. 536 (1977); Gelfand & Solomon, *Analyzing the Decision Making Process of the American Jury*, 70 J. AM. STATISTICAL A. 305 (1975); Gelfand & Solomon, *Modeling Jury Verdicts in the American Legal System*, 69 J. AM. STATISTICAL A. 32 (1974); Gelfand & Solomon, *A Study of Poisson's Model for Jury Verdicts in Criminal and Civil Cases*, 68 J. AM. STATISTICAL A. 271 (1973).

8. Klevorick & Rothschild, *A Model of the Jury Decision Process*, 8 J. LEGAL STUD. 141, 150 (1979).

9. Grofman, *Jury Decision-Making Models*, in MODELING THE CRIMINAL JUSTICE SYSTEM 191, 196 (S. Nagel ed. 1977); Grofman, *Not Necessarily Twelve and Not Necessarily Unanimous*, in PSYCHOLOGY AND THE LAW (G. Bermont & N. Vidmur eds. 1976).

with the renowned mathematician Simeon Poisson. With it, Poisson calculated that in his day 47% of the criminally accused in France were not guilty and that the individual jurors erred in one-third of their verdicts.[10] In more recent years, Poisson's model has been ably refined and extended, principally by Gelfand and Solomon, who have employed it with data from the 1950's on jury verdicts in the United States.[11] Among other things, Gelfand and Solomon purport to show (with varying degrees of confidence) that the probable (apparent) guilt of an accused brought before a jury in this country is between .66 and .76, that there is "essentially no difference" in the probability of conviction by a six- as opposed to a twelve-person jury, and that the probabilities of both false convictions and false acquittals are substantially greater with the smaller jury. Conclusions such as these are of obvious interest in connection with the issue of the constitutionality of departures from the traditional twelve-member jury, and Gelfand and Solomon conclude that their "fairly sophisticated probabilistic models"[12] undermine cases like *Williams v. Florida*.[13]

As I have intimated, I believe that at least some of these deductions and predictions are not warranted by the current jury models, especially as used with existing empirical data on jury decisions. I focus on the Poisson model as refined by Gelfand and Solomon because I believe it to be the most powerful and defensible model thus far constructed.[14] Part I reviews the Poisson model in order to enucleate its many assumptions. Although a number of mathematical equations appear in this section, the analysis should be accessible to readers whose mathematical expertise is limited to high school algebra. Part II is directed to more mathematically inclined readers. It explains

---

10. S. POISSON, RECHERCHES, SUR LA PROBABILITÉ DES JUGEMENTS EN MATIÈRE CRIMINELLE ET EN MATIÈRE CIVILE 373 (1837).

11. *See* note 7 *supra*.

12. *Considerations in Building Jury Behavior Models*, *supra* note 7, at 311.

13. 399 U.S. 78 (1970). A careful and very helpful discussion of the relationship between social science research and the constitutional issue is provided in Lempert, *Uncovering "Nondiscernible" Differences: Empirical Research and the Jury-Size Cases*, 73 MICH. L. REV. 643 (1975). The Court's most recent decisions on the constitutionality of smaller juries are Ballew v. Georgia, 435 U.S. 223 (1978), and Burch v. Louisiana, 441 U.S. 130 (1979). These rulings, and the role of statistical reasoning in resolving the jury size question are discussed in Kaye, *And Then There Were Twelve: The Supreme Court, Statistical Reasoning, and the Size of the Jury*, 68 CALIF. L. REV. 1004 (1980).

14. Most contemporary models are discussed in Penrod & Hastie, *Models of Jury Decision Making: A Critical Review*, 86 PSYCH. BULL. 462 (1979). For a more detailed criticism of the models cited in *Ballew v. Georgia* (and a less elaborate analysis of the Poisson model than is presented here), see Kaye, *supra* note 13.

why these assumptions may introduce serious errors into the probabilities calculated according to the model. To emphasize the difficulties involved, it also exhibits a model that is more plausible but quite intractable. Part III indicates one reason that even a relatively error-free model would have limited usefulness in resolving the constitutional issue posed by juries composed of fewer than twelve persons.

I. THE POISSON MODEL AND ITS ASSUMPTIONS

To appreciate the Poisson model and its limitations, it is helpful to see how it is derived. An understanding of the essential features of the model requires more patience than mathematical acumen. Although the derivation of the equations is not mathematically complex, there are many symbols to keep track of.

A. *Notation*

To start with, the mathematics can be illustrated most easily by considering the simplest possible cases—a one-member "jury" and a two-member panel. The probability that a one-member "jury" will convict in a randomly selected case can be called P_C. This probability of conviction can be written, as we will soon see, in terms of two parameters: the probability μ that the decision will be correct and the probability θ that the defendant is legally guilty.[15] To express P_C in terms of μ and θ, some further notation is helpful. Let G represent the event that the defendant is guilty, and let $\bar{G}$ designate the complementary event that he is not guilty. The probabilities associated with these events are P_G and $P_{\bar{G}}$, respectively. Similarly, we can denote the probabilities of conviction and non-conviction as P_C and $P_{\bar{C}}$.

Of course, the probability of a jury conviction does not necessarily equal the probability of a defendant's guilt. Some of the guilty escape punishment, and some of the innocent are convicted. To account for such unpleasant possibilities, we let $P_{C/G}$ stand for the probability that a juror votes to convict given that defendant is actually guilty, and we take $P_{C/\bar{G}}$ to be the probability of a vote for convic-

15. On the meaning of legal and factual guilt, see Aranella, *Reforming the Federal Grand Jury and the State Preliminary Hearing to Prevent Conviction Without Adjudication*, 78 MICH. L. REV. 463, 465-66 n.6 (1980). Gelfand and Solomon prefer to define θ as the probability that a defendant is very likely to be guilty given the evidence adduced, *i.e.*, that a defendant is "convictable" rather than "truly guilty." Nevertheless, interpreting θ as the proportion of defendants who are actually guilty is equally compatible with the equations used in the model. That there is no practical method to verify the value of θ deduced from the model does not preclude this interpretation of θ. *See* Kaye, *supra* note 13, at 410-11 n.43.

tion given that defendant is not guilty. In an ideal world, one in which all the guilty were convicted and none of the innocent were convicted, $P_{C/G}$ would equal one, and $P_{C/\bar{G}}$ would equal zero.

Another way to express the likelihood of false convictions is to define P_{CG} as the probability, before trial, that a randomly selected defendant is guilty and that the jury convicts. Likewise, $P_{C\bar{G}}$ is the probability of a conviction and an innocent defendant. Whereas the previous formulation pertained to the conditional probabilities (of conviction given guilt and of conviction given innocence), the probabilities P_{CG} and $P_{C\bar{G}}$ are unconditional. They pertain to events occurring jointly. To illustrate this distinction, consider the events "a cloudy day" and "a rainy day." Suppose we are asked to estimate the conditional probability that tomorrow will be rainy given that tomorrow is cloudy. Knowing that it will be cloudy, we might pick a fairly high value for this probability. On the other hand, our estimate of the probability that it will be both cloudy and rainy may be lower, since we are not assured that tomorrow will be cloudy.

The joint probabilities and the conditional probabilities are closely intertwined, however. By the definition of conditional probability,

$$P_{CG} = P_{C/G}P_G \tag{1a}$$

and

$$P_{C\bar{G}} = P_{C/\bar{G}}P_{\bar{G}} \tag{1b}$$

Put otherwise, (1a) states that the probability that defendant is not only guilty, but also convicted, is simply the probability that he is convicted given that he is guilty times the probability that he is guilty. Equation (1b) states in an analogous way the probability that defendant is not guilty but convicted. So much for notation.

B. *First Ballot Probabilities*

Let us consider the one-person "jury." The probability P_C that the "juror" will convict can be written as the sum of two terms: (1) the probability that he will convict and that the defendant is guilty, and (2) the probability that he will convict and that the defendant is not guilty. In symbols,

$$P_C = P_{CG} + P_{C\bar{G}} \tag{2}$$

The first term on the right is an "error free" term, while the second represents the likelihood of an erroneous conviction. Substituting (1a) and (1b) gives

$$P_C = P_{C/G}P_G + P_{C/\bar{G}}P_{\bar{G}} \tag{3}$$

Now C/G is a correct decision, and $C/\bar{G}$ is an incorrect one. Since, by definition, μ is the probability of a correct decision, we can substitute $P_{C/G} = \mu$ and $P_{C/\bar{G}} = 1-\mu$ into (3) to obtain

$$P_C = P_G\mu + P_{\bar{G}}(1-\mu) \tag{4}$$

But P_G is the same as θ, the probability before trial of guilt, so

$$P_C = \theta\mu + (1-\theta)(1-\mu) \tag{5}$$

Equation (5) expresses P_C in terms of θ and μ, as promised, but is otherwise uninteresting. The two-person jury case is more revealing. Let $\gamma_{2,i}$ represent the probability that exactly i jurors in the two-member panel will vote for acquittal on the first ballot. The probability that both jurors will vote to convict is thus $\gamma_{2,0}$. By an extension of the reasoning that led to (5), it follows that

$$\gamma_{2,0} = \theta\mu^2 + (1-\theta)(1-\mu)^2 \tag{6}$$

The "error-free" term now contains a factor μ^2 instead of μ, reflecting the probability that two jurors, rather than one, vote correctly to convict. Likewise, the $(1-\mu)^2$ in the "error" portion of (6) reflects the possibility that both votes for conviction are incorrectly cast.[16]

The probability that the two-member jury will split is

$$\gamma_{2,1} = \theta\mu(1-\mu) + (1-\theta)\mu(1-\mu) \tag{7}$$

---

16. The full derivation of (6) requires additional notation. Let C_1C_2 denote the joint event that both juror number one and juror number two vote to convict on the first ballot. Then the probability associated with this event can be written

$$\gamma_{2,0} = P_G P_{C_1C_2/G} + P_{\bar{G}} P_{C_1C_2/\bar{G}}$$

Assuming that each juror votes independently,

$$\gamma_{2,0} = P_G P_{C_1/G} P_{C_2/G} + P_{\bar{G}} P_{C_1/\bar{G}} P_{C_2/\bar{G}}$$

If each juror is equally likely to vote correctly, so that $P_{C_1/G} = P_{C_2/G} = \mu$ and $P_{C_1/\bar{G}} = P_{C_2/\bar{G}} = 1-\mu$, we have

$$\gamma_{2,0} = P_G\mu^2 + P_{\bar{G}}(1-\mu)^2$$

Finally, identifying θ with P_G and $1-\theta$ with $P_{\bar{G}}$, we arrive at (6).

The Poisson model differs from a more simplistic "one parameter" model in which each juror has the same probability of voting to convict, and this probability is independent of whether defendant is guilty. The probability of a vote to convict in the Poisson model depends on whether defendant is guilty. If there is guilt, the conditional probability is μ; otherwise, it is $1-\mu$. One consequence of this difference is that although a one parameter model requires that $\gamma_{2,0} = \gamma_{1,0}^2$, the more flexible Poisson model does not.

Here the "error-free" and the "error" term alike involve the probability of one correct vote, μ, and the probability of one incorrect vote, $1-\mu$.

The remaining possibility is that both jurors vote to acquit on the first ballot. The probability of this event is given by

$$\gamma_{2,2} = \theta(1-\mu)^2 + (1-\theta)\mu^2 \tag{8}$$

The values of θ and μ can now be deduced—provided that we have measurements of the way two-member juries divide on their first ballot votes in a representative cross-section of cases. Thinking of the observed data on the ballot distributions of two-member juries as having been generated according to the probability model (6-8), we can ask what specific values of θ and μ would be the most likely to produce the observed sample of juror votes. The values deduced in this fashion are known as maximum likelihood estimates.[17]

The generalization of this procedure to deal with larger juries is mathematically straightforward. Without belaboring the details any further, it turns out that for a jury of size n, the probability $\gamma_{n,i}$ of i initial votes for acquittal can be written[18]

$$\gamma_{n,i} = \binom{n}{i} \left[\theta\mu^{n-i}(1-\mu)^i + (1-\theta)\mu^i(1-\mu)^{n-i}\right] \tag{9}$$

Using data on the first ballot voting patterns of the twelve-member juries studied by Kalven and Zeisel, Gelfand and Solomon find the maximum likelihood estimators of θ and μ to be .69 and .88, respectively.

C. *Final Verdict Probabilities*

To arrive at the probability P_C that a jury will convict on the final ballot, we must relate the initial votes, whose probabilities $\gamma_{n,i}$ we

---

17. Other methods of deducing the values of the parameters θ and μ are also used by Gelfand and Solomon. I mention maximum likelihood estimation because it is intuitively accessible and appealing. To take a simpler example, suppose we know that a coin has been flipped 10 times and has come up heads nine times. What is the probability π that this coin will come up heads on any given toss? The maximum likelihood estimator for π is simply the observed proportion of heads, $\hat{p}=.9$. The likelihood that a coin will show 9 out of 10 heads is greatest if that coin is weighted so that in each toss the chance of a head is .9.

Of course, π might not be .9. If our prior experience with coins led us to expect that the coin would be balanced, then we might choose a number closer to .5 for our estimate of π. The formal procedure for incorporating prior experience into parameter estimation is known as Bayesian estimation.

18. The expression $\binom{n}{i}$ is the binomial coefficient that gives the number of distinct ways n objects can be selected i at a time.

know from (9), to the final votes. Gelfand and Solomon use data on both mock and actual twelve-member juries to calculate the probability $P_{C/12,i}$ that a jury of size twelve that starts out with i votes for acquittal ultimately will convict. The probability of conviction is then given by

$$P_C = P_{C/12,0}\gamma_{12,0} + P_{C/12,1}\gamma_{12,1} + \cdots + P_{C/12,12}\gamma_{12,12} \tag{10}$$

To make predictions about other sized juries, Gelfand and Solomon use the analogous relation

$$P_C = P_{C/n,0}\gamma_{n,0} + \cdots + P_{C/n,n}\gamma_{n,n} \tag{11}$$

where n is, of course, the jury size in question. They assume that the transition probabilities $P_{C/12,i}$ found from data on unanimous twelve-member juries hold, *mutatis mutandis*, for nonunanimous juries and for smaller juries. For example, they assume that the probability that a six-member jury will convict if it starts out with one vote for acquittal is the same as the probability that a twelve-member jury will convict if it begins with two votes for acquittal.[19] In other words, Gelfand and Solomon calculate the transition probabilities on the basis of a "proportionality hypothesis" that states

$$P_{C/n,ni\div12} = P_{C/12,i} \tag{12}$$

Having estimated the probability of conviction for various sized juries, Gelfand and Solomon find the conditional probabilities of false convictions and false acquittals by various applications of the definition of conditional probability and its close cousin, Bayes's rule.

D. *Scrutinizing the Assumptions*

The mathematical manipulations outlined above are mathematically impeccable. Since Gelfand and Solomon (not to mention Poisson) are deservedly respected mathematicians, this is hardly surprising. If the model has any weakness, it must lie in the plausibility and appropriateness of the assumptions built into it. Many of these assumptions are carefully stated by Gelfand and Solomon, but evaluating their significance is another matter. In this section, I shall argue that the assumptions are so drastically simplified that the numerical predictions of the model must be approached with the greatest caution.

To begin with, let us consider the meaning of θ in more detail.

---

19. *See Considerations in Building Jury Behavior Models*, *supra* note 7, at 310.

This parameter, as we have seen, is the probability before trial that a defendant is guilty.[20] Although θ undoubtedly varies from case to case, it is assumed to be the same for all defendants. In principle θ could be allowed to take on different values for various types of defendants or offenses, but current data is too limited to permit this.[21]

In addition, θ is assumed to be independent of n. At first blush, this supposition seems plausible, but it ignores a conceivable feedback effect of the jury size n on θ. Some writers have speculated that smaller juries are more prone to convict an innocent defendant but less likely to fail to convict a guilty person than larger juries.[22] Certainly, some plaintiffs' attorneys believe that it is easier to convince six jurors than to prevail before the traditional panel of twelve.[23] Moreover, much of the commentary on the Supreme Court's jury size decisions argues that the decisions of smaller juries are likely to be more erratic than those of larger bodies.[24] If many prosecutors share the view that smaller juries are more prone to err, especially by way of conviction, then these prosecutors might be willing to proceed with less compelling evidence, and a lower value of θ might come to characterize a regime of smaller juries.

The second parameter μ is the probability that a juror will vote correctly on the first ballot. Although μ plainly varies from juror to juror and from case to case, it, too, is taken to be constant.[25] Moreover, there is a possible feedback effect between μ and θ. If θ is approximately one as opposed to, say, one-half, jurors may be more inclined to reason that "the defendant is guilty, or else he wouldn't be

20. *See* note 15 *supra*.

21. *Considerations in Building Jury Behavior Models*, *supra* note 7, at 300; *Modeling Jury Verdicts in the American Legal System*, *supra* note 7; *A Study of Poisson's Model*, *supra* note 7.

22. Friedman, *Trial by Jury: Criteria for Convictions, Jury Size and Type I and Type II Errors*, 26 AM. STATISTICIAN 21 (1972); Lempert, *supra* note 3; Nagel & Neef, *Deductive Modeling to Determine an Optimum Jury Size and Fraction Required to Convict*, 1975 WASH. U.L.Q. 933.

23. ROSCOE POUND—AMERICAN TRIAL LAWYERS FOUNDATION, ANNUAL CHIEF JUSTICE EARL WARREN CONFERENCE ON ADVOCACY IN THE UNITED STATES: THE AMERICAN JURY SYSTEM, FINAL REPORT 20 (1977).

24. *E.g.*, Kaye, *supra* note 13; Lempert, *supra* note 13; Zeisel, "*. . . And Then There Were None": The Diminution of the Federal Jury*, 38 U. CHI. L. REV. 710 (1971).

25. Gelfand and Solomon have also devised a "three parameter" model in which μ is allowed to take on two values, μ_1 and μ_2, for the probability of a correct vote given that defendant is guilty and for the probability of a correct vote given that defendant is innocent. The criticisms advanced here apply, with minor emendations, to this three parameter model.

here." Where θ really is high, acting on this surmise may lead to fewer mistakes and thereby result in a larger value of μ. Finally, μ itself seems likely to be a function of n and to depend as well on whether the jury is required to achieve unanimity. Social-psychological studies have shown that within a broad range of sizes, larger groups perform better than smaller ones on tasks involving recall or insight.[26] Where a unanimous verdict is not essential, the contributions of some jurors may be ignored, making the jury act as if n were smaller. Of course, to the extent a jury proceeds to a first ballot prior to any discussion among the jurors, the influence of n on μ will be mitigated, but since smaller juries may be more likely to discuss matters informally before proceeding to a formal expression of opinion, even the magnitude of this mitigating factor may be affected by n.

As shown in section A, however, Gelfand and Solomon estimate θ and μ by searching for the single value of θ and the single value of μ that jointly generate the function $\gamma_{n,i}$ that best fits the data of Kalven and Zeisel on the first ballot votes of 225 unanimous, twelve-member juries in Chicago and Brooklyn. Consequently, the probabilities $\gamma_{n,i}$ given by (9) should be considered speculative, especially when the jury size n differs from twelve.

The transition probabilities $P_{C/n,i}$ that are used to calculate the likelihood that the jury as a whole will err rest on equally troublesome premises. These probabilities may be more stable across the spectrum of defendants and cases, but they too seem connected with the size and unanimity requirement of a jury. To account for size effects, Gelfand and Solomon rely on a "proportionality hypothesis" under which two out of six first ballots for acquittal produce the same likelihood of a final verdict of conviction as do $2i$ out of twelve.[27] Yet, this assumption is almost surely false. It contradicts generally accepted social-psychological findings about the probability of a group

---

26. Thomas & Fink, *Effects of Group Size*, 60 PSYCH. BULL. 371 (1963).

27. *See* text accompanying note 19 *supra*. In Kaye, *supra* note 13, at 1031 n.105, I characterized this view as a resurrection of the proportionality thesis advanced in Williams v. Florida, 399 U.S. 78, 100-01 n.49 (1970). This characterization was imprecise and unfortunate. The *Williams* Court thought that regardless of jury size, "jurors in the minority on the first ballot are likely to be influenced by the proportional size of the majority aligned against them." *Id.* This type of proportionality implies that the transition probabilities for two different size juries are proportional, as Gelfand and Solomon assume. The difference is that it also fixes the values for these transition probabilities uniquely and unrealistically. Gelfand and Solomon reject this facet of the *Williams* Court's claim of proportionality.

of i dissenters' conforming to a position held by a majority of n-i group members.[28] Since neither the calculated values of $\gamma_{n,i}$ nor those of $P_{C|n,i}$ can be accepted with much confidence, the conclusions about the error rates for various jury sizes and voting protocols must be viewed with some healthy skepticism.

II. Errors in the Estimation and Predictions of Means

The previous section challenged several of the assumptions built into the Poisson model as employed by Gelfand and Solomon. Nevertheless, since these researchers do not use equations (9) and (11) to make predictions for specific cases, it might be argued that these assumptions are not as serious as I have implied. For example, although I have stated that the model assumes that μ does not vary from one juror to another, it is open to Gelfand and Solomon to respond that the correct value for each juror can be conceived of as an independent observation from a common distribution with mean μ.[29] More generally, perhaps all the "constants" can be viewed as the means of random variables, and (9) and (11) can be understood as predicting mean probabilities of conviction, etc.[30]

Such a rejoinder would be false for at least two reasons. First, (9) is nonlinear in θ and μ. In general, for a nonlinear function g of m random variables X_j,

$$E[g(X_1,\ldots,X_m)] \neq g[E(X_1),\ldots,E(X_m)]$$

More specifically, let us write an equation for $\gamma_{n,i}$ that expresses the interdependence of θ, μ and n more clearly than (9) and (11). If we assume that the feedback effects between θ and n and between θ and μ eventually result in stable equilibrium values $\tilde{\theta}$ and $\tilde{\mu}$, and if we explicitly note the dependence of θ and μ on n, we obtain,

$$\gamma_{n,i} = \binom{n}{i}\left[\tilde{\theta}_n\tilde{\mu}_n^{n-i}(1-\tilde{\mu}_n)^i + (1-\tilde{\theta}_n)\tilde{\mu}_n^i(1-\tilde{\mu}_n)^{n-i}\right] \qquad (13)$$

To account for the variation of $\tilde{\theta}_n$ from case to case and defendant to defendant, we now regard $\tilde{\theta}_n$ as a random variable. Likewise, to recognize the variation of $\tilde{\mu}_n$ from case to case and from juror to juror,

---

28. Lempert, *supra* note 13, at 678-79.

29. *See A Study of Poisson's Models*, *supra* note 7; *Considerations in Building Jury Behavior Models*, *supra* note 7, at 300.

30. *See Considerations in Building Jury Behavior Models*, *supra* note 7, at 300 ("If each juror is conceived as being chosen randomly from the population of possible jurors then the average or mean value of this characteristic is appropriate to use in describing mean or average performance for a jury scheme.").

we can view $\tilde{\mu}_n$ as another random variable. Denoting the joint probability density of these variables by $f(\tilde{\theta}_n, \tilde{\mu}_n)$, the expected value of $\gamma_{n,i}$ is given by

$$E(\gamma_{n,i}) = \binom{n}{i}\int_0^1\int_0^1 \left[\tilde{\theta}_n\tilde{\mu}_n^{n-i}(1-\tilde{\mu}_n)^i + (1-\tilde{\theta}_n)\tilde{\mu}_n^i(1-\tilde{\mu}_n)^{n-i}\right]f(\tilde{\theta}_n, \tilde{\mu}_n)d\tilde{\theta}_n d\tilde{\mu}_n \tag{14}$$

$E(\gamma_{n,i})$ given by equation (14) does not generally equal $\gamma_{n,i}$ found from (13). Depending on the nature of f, the magnitude of the discrepancy can be substantial.[31]

The second difficulty with viewing Gelfand and Solomon's development of the Poisson model as a correct procedure for calculating expectation values occurs when the values of θ_n and μ_n found from data for juries of size n are used to make predictions for juries composed of some other number n' of jurors. These predictions are, of course, of great interest to the legal community, but as the notation of equation (13) emphasizes, it is not clear that $\tilde{\theta}_n = \tilde{\theta}_{n'}$ or that $\tilde{\mu}_n = \tilde{\mu}_{n'}$. In addition, as discussed in section C of part I, the transition probabilities needed to arrive at P_C and the like also seem to depend on n.

The magnitude of the errors resulting from these matters could best be appreciated by testing the predictions of the Poisson model against empirical conviction rates and ballot distributions for the six-member juries used in some states. In the absence of such data, we are left with an unvalidated model of jury decisionmaking. By the same token, its predictions cannot be disproved at present, but the number of doubtful assumptions built into it gives us little reason to be confident in the predictions thus far developed.[32]

---

31. Gelfand and Solomon argue that this difference will be negligible if jurors are selected at random from a sufficiently homogeneous population (and, by implication, if defendants and criminal cases are also homogeneous). *Considerations in Building Jury Behavior Models*, *supra* note 7, at 300. They offer no reason beyond mathematical convenience for making this assumption of homogeneity. Of course, the observation that Gelfand and Solomon may not be able to estimate accurately the mean probabilities associated with various size juries and voting protocols does not establish that the probabilities deduced from the Poisson model are meaningless. For example, $\gamma_{n,i}$ does represent the probability that i out of n identical, "average" jurors voting independently will vote to acquit on the first ballot. Although not the "average" probability for juries of size n, $\gamma_{n,i}$ is a probability for a perfectly homogeneous, "average" jury.

32. *But see* note 33 *infra*.

III. From Mathematics to Policy

In the end, it may be that the present form of the model is the best we can hope for. And, if some assumptions—even heroic ones—must be made, is it not better to make them and to generate some results than to leave the resolution of the legal issues to which these results pertain solely to judicial hunch? The short answer is that, as long as the model's predictions are suitably qualified and explained, they may appropriately serve a function in judicial decisionmaking.[33]

But lawyers distrust predictions derived from mathematical models, even when the results are plainly relevant to legal issues, partly because they are like other human beings—they fear the unfamiliar. Furthermore, they worry that judges, not being schooled in such modes of thought, may place more reliance on superficially impressive statistical demonstrations than can be justified by more careful analysis. As Justice Blackmun's opinion in *Ballew v. Georgia*[34] unfortunately demonstrates, this latter concern is not entirely without foundation.[35] Indeed, even an accurate model of size effects on jury verdicts and error rates can be seriously misleading. Suppose, for instance, that such a model "proved" that smaller juries do not differ substantially from larger ones, or even that smaller juries can be expected to minimize a weighted sum or erroneous verdicts, as some researchers have claimed.[36] If these results are restricted to the "average jury" or the "average defendant" or the "average case"—in short, if they are stated in terms of expected values which collapse

---

33. One method of presentation that can prove helpful when controversial assumptions about the values of parameters or variables are necessary is to display the results for a broad range of possible values. For instance, I have criticized the Poisson model partly on the ground that distinct values of θ and μ apply to different defendants. If, however, the model's output does not change substantially as θ and μ are varied, the force of this criticism is blunted. Thus, Gelfand remarks:

> Our greatest frustration is the lack of real data. There is considerable external data—conviction rates, etc., over various types of cases and jury schemes, but the only internal data [on preliminary balloting] is that of Kalven and Zeisel, and it is 20 years old. One fortunate aspect is the fact that the direction of our inferences is fairly insensitive . . . . [O]ver a broad range of social decision schemes [transition probabilities] and of θ, μ_1 [and] μ_2 values [used in the "three parameter" model], it is still the case that the six member jury will make more errors of each type than its twelve member counterpart.

Letter from Alan Gelfand to David Kaye, Nov. 17, 1980.

34. 435 U.S. 223 (1968).

35. *See* Kaye, *supra* note 13.

36. Nagel & Neef, *supra* note 22.

several dimensions of distinct legal interest into scalar quantities—they can seduce policymakers into ignoring crucial matters. For instance, finding that the expected number of false convictions and false acquittals are the same for twelve-member and six-member juries might lead a judge to think that a six-person jury is as good as the traditional twelve-member one. But if it is not also shown that the six-member jury makes no more errors when confronted with different categories of cases, the judge's first thought should be dismissed. If six-member juries are more likely to err when confronted with, say, a group of politically unpopular defendants in a "political trial," or a large corporate defendant in a personal injury action, they should not be upheld as constitutional even if their mean error rate (averaged across all types of defendants) is identical to that for twelve-member juries.

It is hard to avoid such mistakes in translating research work into policy conclusions. In fact, Gelfand and Solomon unwittingly may invite this type of mistake in reporting that twelve- and six-person juries can hardly be distinguished on the basis of their expected conviction rate P_C,[37] and that this finding "reinforces the judicial decision that there is essentially no difference in the prospects for conviction of an accused before a six-man vs. a twelve-man jury."[38] Even if the mathematical result is trustworthy, such a conclusion tempts a policymaker to forget that this is a mean figure and that important differences might be revealed by a more fine-grained analysis. Moreover, the differences Gelfand and Solomon describe in their earlier work as "negligible" include, for several values of θ and μ, as many as 42 convictions per 1000 jury trials. While such a number is small in relation to P_C itself, convicting, in the long run, an additional forty-two persons out of every thousand is not "negligible" for legal purposes.[39]

In presenting these thoughts, I do not wish to be understood as

---

37. *E.g.*, *Considerations in Building Jury Behavior Models*, *supra* note 7, at 308 (rates of conviction for six- and twelve-member juries are "approximately the same"). Gelfand and Solomon do find the different sized juries distinguishable on the basis of differences in $P_{C|G}$, $P_{C|\bar{G}}$ and similar quantities. *E.g.*, *Rejoinder*, *supra* note 7.

38. *Modeling Jury Verdicts*, *supra* note 7, at 36.

39. Similar observations have been made with regard to the confusion created by the different statistical and legal meanings of the word "significant." Lempert, *supra* note 13. In their most recent paper Gelfand and Solomon recognize that "if we compare a characteristic such as the rate of conviction over two schemes, [t]hough the percentage differences may be small the absolute number may not be so." *Considerations in Building Jury Behavior Models*, *supra* note 7, at 296.

arguing that legal policy should never be tainted by statistical modeling. To the contrary, I believe that the jury size question is an appropriate one for statistical analysis, that the model explicated by Gelfand and Solomon is intriguing, and that Gelfand and Solomon are unusually sensitive to some of the problems of model-building in this area and to the need for further research. My thesis, then, consists of two points. First, I believe that there are still major problems to be solved—perhaps in the collection of data and perhaps in further refinements of the Poisson model—before its predictions can be relied on with confidence in setting legal policy. Second, I hope that my observations will help create a better understanding of the obstacles to drawing policy conclusions from research efforts, and make for a more fruitful and constructive dialogue between the disciplines of law and statistics.