

# Dr. Spock and the Case of the Vanishing Women Jurors

Hans Zeisel†

When the one hundred jurors whose names the clerk had pulled from the court's jury box appeared before the judge who was to try Dr. Benjamin Spock and his associates for conspiring to violate the Military Service Act of 1967,<sup>1</sup> there were only nine women among them.<sup>2</sup> Since the jurors were allegedly drawn more or less at random<sup>3</sup> from the list of residents over 21, the nine women posed a puzzle: how could such a selection reduce the proportion of women from over 50 per cent in the eligible population<sup>4</sup> to 9 per cent in the jury venire? On this ground the defense challenged the array of jurors, and the inquiry into this perplexing sex distribution was undertaken in aid of that challenge.<sup>5</sup>

By what transpired in the course of that inquiry, there hangs a tale of three-fold significance. First, there was the essential problem of solving the mystery of the vanishing women; the first five sections of this article trace the clues which lead to the solution. The sixth section considers the relevance of the solution to a more general subject that is attracting increasing interest in the courts—proof of an event through statistical inference. In the final section, the beneficial effects of the recent changes in the law governing the selection of federal juries<sup>6</sup> are brought into focus against the background provided by this study of the procedure under the old statute.

---

† Professor of Law and Sociology, The University of Chicago Law School.

<sup>1</sup> 50 U.S.C. §§ 451-73 (Supp. IV, 1969).

<sup>2</sup> *United States v. Coffin*, Crim. No. 68-1-F (D. Mass., Aug. 14, 1968), *rev'd sub nom. United States v. Spock*, 416 F.2d 165 (1st Cir. 1969).

<sup>3</sup> See text following note 8 *infra*.

<sup>4</sup> Brief for Appellant Spock at 50, 416 F.2d 165 (1st Cir. 1969). "The 1960 census shows that there are roughly 1,721,000 women over the age of 21 and 1,524,000 men over the age of 21 in the district. U.S. Department of Commerce, *Census of Population: 1960*, Volume I, Part 23, Massachusetts, pp. 31-32." *Id.* at n.38.

<sup>5</sup> The data of this study were compiled and analyzed at the request of defense counsel for Dr. Spock. They are published here with his permission, now that the case is no longer *sub judice*.

<sup>6</sup> 28 U.S.C. §§ 1861-71 (Supp. IV, 1969), *amending* 28 U.S.C. 1861-71 (1964).

## I. JURY SELECTION IN THE DISTRICT COURT

At the time of the Spock trial, jury venires in the United States District Court for the District of Massachusetts were selected in two stages: (1) A relatively large contingent of names of potential jurors were selected by the clerk of the court from the police lists<sup>7</sup> of the court's district. These potential jurors received questionnaires from the clerk of the court who, after eliminating those disqualified by statute from jury service, placed batches of 300 names into a central jury box; (2) from this box, the clerk, prior to the scheduled jury trial, drew a venire of jurors. The names so drawn were ordered to appear in court on the day of the trial.

To support its challenge to the array, the defense called the clerk of the court, who under examination described phase (1) of the selection process as follows:<sup>8</sup>

Answer: . . . I put my finger on the place and on a name on the page and then I make a mark next to it with a pen.

Question: Do you do that by not looking at the page?

Answer: I have to look at it enough to know where it is in relation to my finger.

Question: Yes.

Answer: I do not intend to look carefully at the name . . . .

Question: I assume that at some point you have to look at the name in order to send out the questionnaire?

Answer: Correct.

\* \* \*

Question: Do you have any explanation for that [the disparity between the number of questionnaires sent to the men and women] except the possibility that you might have seen the name and recognized it as a woman's name and figured it is a little more efficient not to send out too many questionnaires to women?

Answer: That is the only possible explanation other than pure chance. . . .

---

<sup>7</sup> A record of all residents 21 years and older, compiled and kept up-to-date by the police; hence, the name.

<sup>8</sup> Record at 466-7, 472-3, *United States v. Coffin*, Crim. No. 68-1-F (D. Mass., Aug. 14, 1968).

The court, after hearing the evidence, overruled the challenge to the array and began the trial proper.

## II. THE JURY VENIRES

Lists of the 46 jury venires selected in the above manner and summoned for trial before the seven judges of the district court were made available for this study.<sup>9</sup> The trials had taken place between April 4, 1966, and October 22, 1968. Synopses of the venires by judge are given in Table 1, which includes the following items:

- (i) month and year in which they were drawn from the jury box;
- (ii) size, that is, the total number of potential jurors who were to report to the court;
- (iii) the percentage of women in the venire.

The venires are arranged in time sequence for each of the seven judges for whose trials they had been selected. The venires range in size from 35 to 200 jurors, with an average size of  $(2975 \div 46 =) 65$ . The total number of jurors in these 46 venires was 2975.

The number of available venire lists varies from judge to judge. For Judge D we had only two venire lists; for Judge F and the trial judge, we were given lists of nine venires. The last line for each judge indicates the total number of jurors in all his listed venires, and the very last figure denotes the average percentage of women found in all his venires.<sup>10</sup>

Table 2 summarizes the percentage of women jurors found in the venires of the seven judges. The average percentage of women jurors in the venires of the six colleagues of the trial judge ranges between 34.1 and 26.8 per cent, with an average distance between them of  $(7.3 \div 5 =) 1.8$  percentage points. The average proportion of women in the venires for these six judges was 29.0 per cent. The venires of the seventh judge, the trial judge, show a surprisingly different picture: in his venires, the average proportion of women jurors was 14.6 per cent, exactly half that of his six colleagues. The distance between the trial judge and the judge nearest to him was about seven times as great (12.2%) as the average distance between any other two proximate judges (1.8%).

This is a startling result, given the fact that the figures for the trial

<sup>9</sup> It was several months after the end of the trial that I obtained, through the help of Leonard Boudin, Esq., counsel on the appeal for Dr. Spock, lists of 46 venires. These, as the letter of transmittal from the clerk of the court stated, were all such lists still on file with the court.

<sup>10</sup> This average percentage is not the average of the individual percentages, but what is called the weighted average, treating the jurors in all venires as if they had been in one large venire.

TABLE I  
THE VENIRES OF THE SEVEN DISTRICT JUDGES

Judge A (5 Venires)			Judge B (6 Venires)		
Date	Size of Venire	Per Cent Women Jurors	Date	Size of Venire	Per Cent Women Jurors
5/67	45	40	5/66	55	36
9/67	50	30	10/66	60	32
2/68	45	16	1/67	66	32
4/68	40	35	3/67	55	27
9/68	46	50	4/68	56	29
Total	226	34.1%	10/68	62	45
			Total	354	33.6%

Judge C (8 Venires)			Judge D (2 Venires)		
Date	Size of Venire	Per Cent Women Jurors	Date	Size of Venire	Per Cent Women Jurors
10/66	50	34	4/66	50	24
1/67	50	30	5/66	61	30
2/67	50	32	Total	111	27.0%
4/67	35	29			
9/67	50	24			
11/67	50	28			
9/68	55	20			
10/68	65	35			
Total	405	29.1%			

Judge E (7 Venires)			Judge F (9 Venires)		
Date	Size of Venire	Per Cent Women Jurors	Date	Size of Venire	Per Cent Women Jurors
9/66	70	33	5/66	70	22
11/66	55	36	9/66	80	21
1/67	71	28	11/66	70	31
3/67	51	20	3/67	70	27
9/67	45	18	5/67	70	17
11/67	200	22	10/67	70	29
9/68	60	40	6/68	80	26
Total	552	27.0%	6/68	95	29
			9/68	125	34
			Total	730	26.8%

TABLE 1 (Continued)

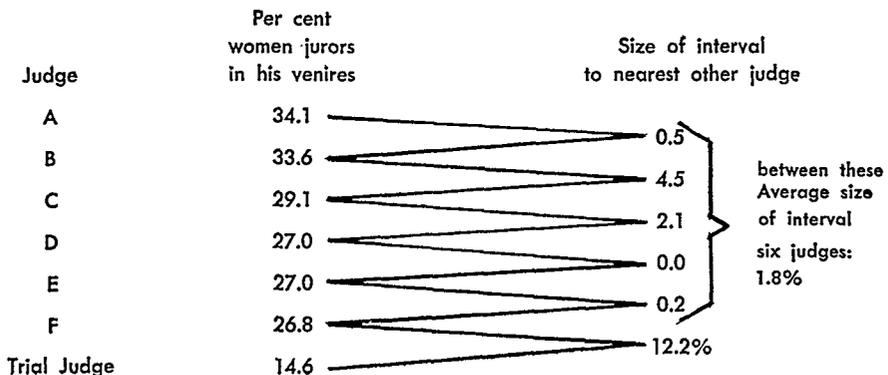
The Trial Judge (9 Venires)		
Date	Size of Venire	Per Cent Women Jurors
5/66	50	16
9/66	50	18
1/67	50	14
5/67	53	6
10/67	50	18
3/68	129	15
3/68	70	15
5/68	100	9
10/68	45	24
Total	597	14.6%

judge are based on 9 different venires, nearly 600 jurors. The figures for each of the other six judges are also based on large numbers: on the average,  $(37 \div 6 =)$  6 different venires or  $(2378 \div 6 =)$  about 400 jurors.

III. CHANCE OR . . . ?

There was thus a serious difference between the trial judge and his colleagues in the percentage of women found in their respective jury venires. Is this difference of a kind that allows us to assume that it had arisen by chance, by the luck of the draw, or is it so large that we must conclude that the venires for the trial judge's court were drawn by a method different from that used by his colleagues on the bench?

TABLE 2  
PROPORTION OF WOMEN JURORS IN THE VENIRES OF EACH JUDGE IN ORDER OF MAGNITUDE



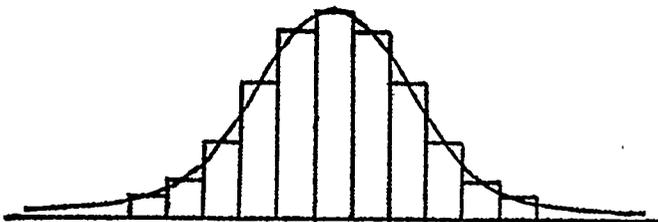
In order to help answer this question, it will be useful to show the crucial percentages of women jurors in each of the 46 venires in graph form. Graph 1 consists of two parts: the upper part lists the 37 venires for the six colleagues of the trial judge; the lower part lists the 9 venires of the trial judge himself. Each rectangle represents one venire. The letter in the rectangle designates the judge for whom the venire was drawn. All rectangles have the same width; their height, however, varies according to the number of jurors in the particular venire. The scale at the left hand margin allows a measure of the size of these venires, individually and cumulatively if more than one is piled on the other. Thus, the first venire in the upper part of Graph 1 represents a venire of Judge F's consisting of 70 jurors of which the percentage of women was either 16 or 17 (16/17). Below it is a venire of Judge A's consisting of 45 jurors, of which 16 or 17 per cent were women, and so forth.<sup>11</sup> If two or more venires are piled on top of each other, it means that all these venires had the same proportion—plus or minus 1 per cent—of women jurors.

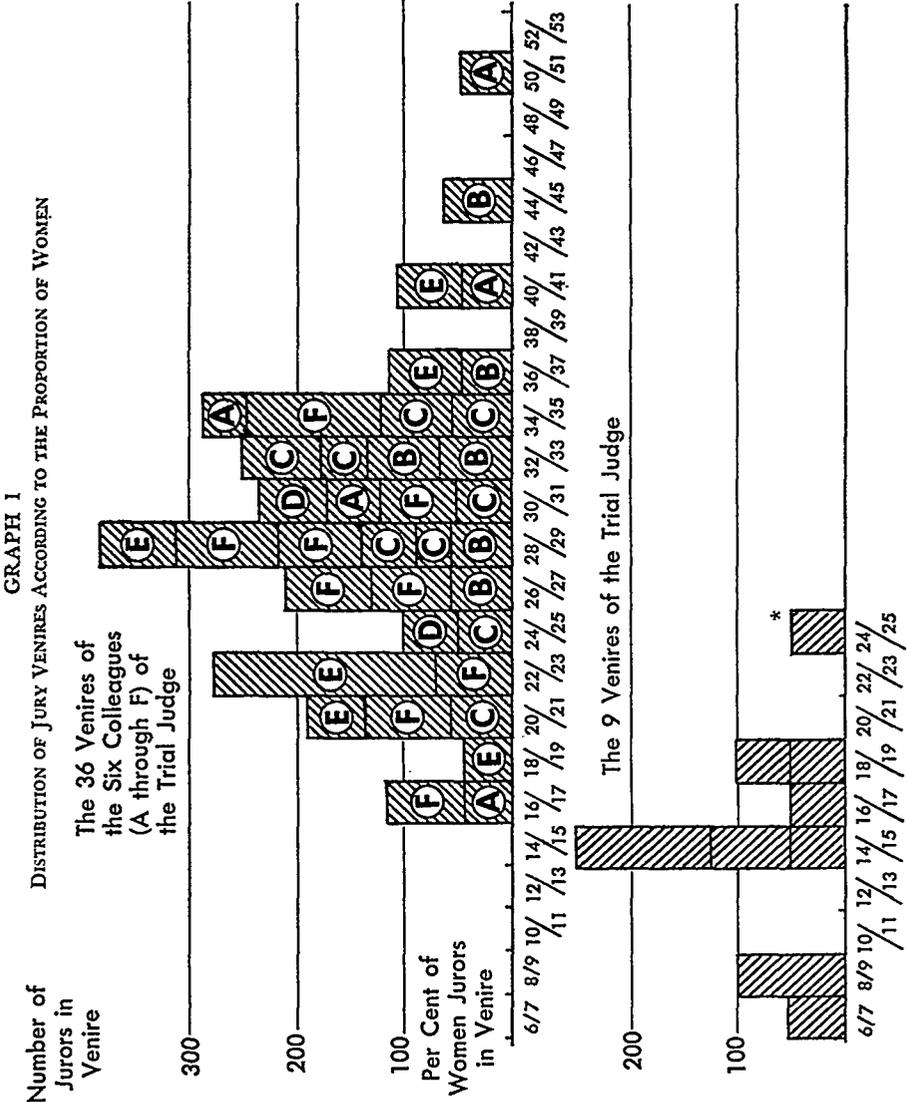
The pattern that emerges in this upper part of Graph 1 is more or less bell-shaped, which is typical for repeated lottery-type drawings from a drum that contains a certain percentage of women (as against men), or black balls (as against white balls), or of any other collection of units of which a certain proportion is "A" while the remainder is "non-A."<sup>12</sup> The peaks of such a distribution are expected to fall around the average (percentage of women) in the drum; that is, most of the draws (venires) are expected to be close to this "true" average of 29.0 per cent women, and so they are. Some draws will result in percentages that deviate somewhat more from the true average, but the greater the distance from the average, the rarer will be the occurrence of such a deviant draw. In our case, only two out of the 37 venires came as low as 16/17 per cent, and only one as high as 50/51 per cent.

The distribution of the 37 venires of the six colleagues of the trial

<sup>11</sup> The precise figures, as can be seen in Table 2, were 17% for the venire of Judge F and 16% for Judge A.

<sup>12</sup> The bell-shaped distribution is depicted and discussed in every standard statistical text. Here, for instance, is an illustration from chapter 7, *The Normal Distribution*, in H. BLALOCK, *SOCIAL STATISTICS* 7 (1960).

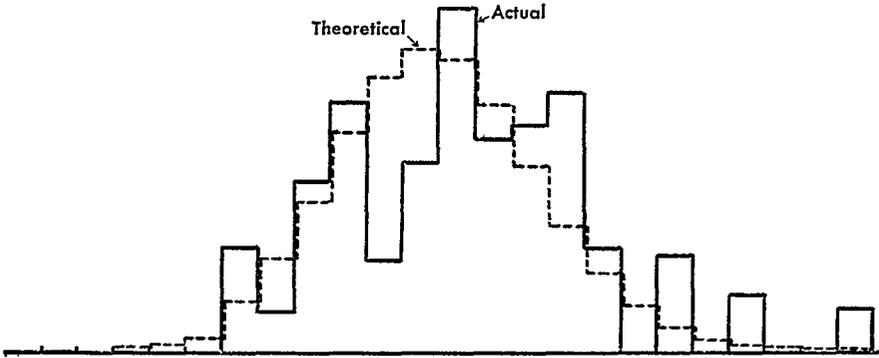




judge is fairly close to the bell-shaped distribution expected by statistical theory. Graph 2 compares the actual distribution with the expected one.<sup>13</sup>

<sup>13</sup> The "expected distribution," designed to approximate the actual distribution, was computed on the following simplified assumptions derived from our data: from a drum that contains a large number of names, 29% of which are female, 37 draws are made; 12 of these draws had 100 jurors per venire, the remaining 25 had 50 jurors per venire. The resulting probabilities were multiplied by 2378, the total number of jurors in all venires.

GRAPH 2  
ACTUAL AND THEORETICAL DISTRIBUTION OF THE 36 VENIRES



The similarity of the actual and the expected distribution supports two conclusions:<sup>14</sup>

(i) The proportion of women in the drum was on the average about 29 per cent, and this average was fairly constant from one filling of the drum to the next.

(ii) Allowing for such minor variations of the percentage of women in each filling of the drum, the six judges drew their venires as true random samples, that is, as a true lottery.

In contrast, the venires of the trial judge, as one glance at the lower part of Graph 1 reveals, fall into a pattern that is different from that of the other six judges. With one exception, his venires also fall fairly symmetrically around the average; but the average itself (14.6 per cent) is markedly lower.<sup>15</sup>

Before posing and answering the hard question that confronts us here, it will be useful to outline the scientific laws on which the answer will be based. We might start with what may look like a reasonable objection to the whole enterprise: since all of these venires are by definition the outcome of a lottery, that is, of pure chance, why may not *any* result be explained on the ground that it was a result of chance? The answer is that while it is true that a lottery may produce any conceivable distribution, some results are less probable than others, and the degree of these probabilities can be calculated with precision.

<sup>14</sup> It would not do trying to explain the minor deviations of the actual distribution from the hypothetical without further information on the details of how the central box was filled each time, and how the venires were drawn. The evidence elicited in the trial is not sufficiently specific. See Brief for Appellant, *supra* note 4, at 50-53.

<sup>15</sup> The exception, the one venire of the trial judge that seems to fall into the pattern of the other six judges (24/25%), is, in a way, the most interesting one. It happens to be the trial judge's most recent venire on our list, dated October 5, 1968.

This is the paradox of the so-called laws of chance: the individual event (whether the next name drawn from the jury drum will be a man or a woman, or whether the next spin of the roulette will bring a black or a red number) is by definition unpredictable; however, the long-run behavior of the roulette wheel, as the long-run ratio of the women jurors drawn from the drum, can be predicted with considerable accuracy. In fact, few profit margins can be predicted with such reliability as those provided by a true roulette wheel or other game of chance.

Thus, the more exact formulation of our problem is: recognizing that it is *possible* that the selection method used by the trial judge has been the same as that of his six colleagues, what is the probability that a difference as large as that between 15 and 29 per cent could arise by chance, if their methods of selecting the venires had been indeed the same? The computation<sup>16</sup> yields the following result: Only once in 1,000,000,000,000,000 will it happen that the distribution obtained by the trial judge will occur by the "luck of the draw."<sup>17</sup> The conclusion, therefore, is virtually inescapable that the clerk must have drawn the venires for the trial judge from the central jury box in a fashion that somehow *systematically* reduced the proportion of women jurors.

#### IV. THE LINK IN A FATAL CHAIN

The systematic deviation of the jury venires must be seen in context, that is, as one link in the chain that formed the jury selection process in the district court, as described in Graph 3.

The chain begins when the clerk of the court draws names from the police list to fill the court's central jury box. At this point he has reduced the proportion of potential women jurors from over 50 per cent on the police list to an average of about 29 per cent of all the names in the central jury box.

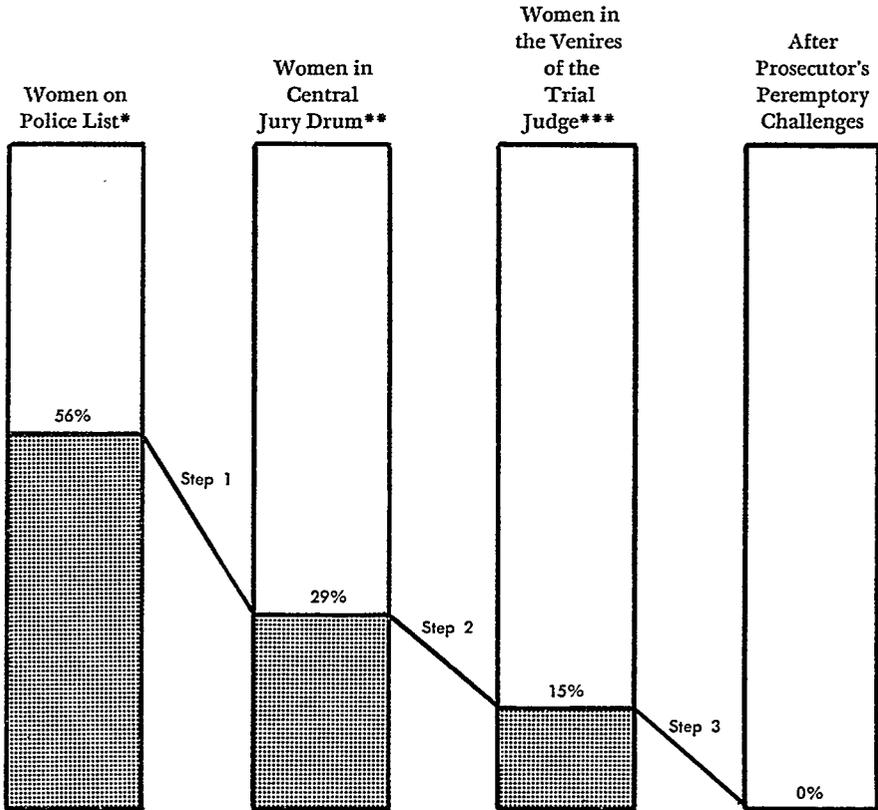
At the time of trial, the defense did not suspect the existence of a second link, namely the special reduction of women jurors in the venires of the trial judge. In these venires, the proportion of women was further reduced from 29 per cent to an average of 15 per cent of all names, and in the venire for the Spock trial to 9 per cent. This reduction in the trial judge's venire set the scene for the last act. With, on the average, only 15 per cent—or one out of seven—women jurors on the venire, the prosecution had no difficulty in preventing any women

---

<sup>16</sup> By the chi-square method, a standard method devised for such tests of approximation.

<sup>17</sup> There are possible refinements of this computation in view of the fact that the distribution of the six judges is not perfectly "homogeneous" as the technical term goes. But preliminary calculations reveal that we are discussing changes of the magnitude of only one of the eighteen zeros in the above result.

GRAPH 3  
THE THREE STEPS THAT ELIMINATE WOMEN JURORS



\* This is the percentage found in a study of the 1965 Boston police list which I had made earlier in a different context. It referred to the persons on the list who are eligible for jury service. The respective percentage for the police lists of the entire court district is not likely to be very different. See notes 4 & 7 *supra*.

\*\* This is the best available estimate from the venires drawn for the courts of the six colleagues of the trial judge. Since the central jury drum is periodically refilled, one must expect this percentage to differ somewhat from drum to drum. See note 14 *supra*.

\*\*\* This is the average for the venires of the trial judge's court.

from sitting on a jury if it so desired; the peremptory challenges saw to that.<sup>18</sup> The reduction of the number of women jurors in the venires of the trial judge thus assumed conclusive dimensions, for it was the key step in the elimination of *all* women jurors at the easy option of the prosecution or, for that matter, of the defense.

## V. LEGAL RELEVANCE

Defense counsel for Dr. Spock submitted the substance of the above evidence in the form of an affidavit appended to a motion requesting

<sup>18</sup> In the instant case, only one woman of the nine on the venire reached the voir dire stage and was promptly eliminated by a peremptory challenge of the prosecution.

a new trial. The motion emphasized the importance of the systematic exclusion of the women jurors in this particular trial by citing the results of a national Gallup Poll, taken in April 1968, just prior to the opening of the trial:<sup>19</sup>

“Question: ‘People are called ‘hawks’ if they want to step up our military effort in Vietnam. They are called ‘doves’ if they want to reduce our military effort in Vietnam. How would you describe yourself—as a ‘hawk’ or a ‘dove?’”

	Hawk	Dove	No Opinion	Total
	%	%	%	%
Men	50	33	17	100
Women	32	49	19	100

There were considerably more doves among the women than among the men.

On January 2, 1969, the defense received notification that the motion had been denied. On January 7, 1969, the Spock case was argued and on July 11, Dr. Spock was acquitted, not on the technical grounds of how the jury was selected, but on the grand design of the first amendment.<sup>20</sup>

## VI. STATISTICAL PROOF

Proof by statistical inference is not unknown to our courts. As a rule it comes up in the simple form of sampling operations from which inferences are urged about the populations from which the sample was drawn: sample surveys from which to estimate market shares in anti-trust litigation;<sup>21</sup> quality samples from produce,<sup>22</sup> or from mineral fields<sup>23</sup> from which to estimate the quality of the whole. This type of inference is now fairly well established in the law, although at times it runs into the barrier of the hearsay rule if the sample properties have been obtained by personal interviews.<sup>24</sup>

There is also another, more complicated, type of statistical inference

<sup>19</sup> Gallup Opinion Index 20 (May, 1968).

<sup>20</sup> *United States v. Spock*, 416 F.2d 165 (1st Cir. 1969).

<sup>21</sup> “If antitrust trials are to be kept manageable, samples must be used . . .” *United States v. United Shoe Machinery Corp.*, 110 F. Supp. 295, 305 (D. Mass. 1953) (Wyzanski, J.). See also ABA ANTITRUST SECTION, *STREAMLINING THE BIG CASE* 21 (1958).

<sup>22</sup> *Hardison Seed Co. v. Jones*, 149 F.2d 252, 257 (6th Cir. 1945).

<sup>23</sup> *United States v. Aluminum Co. of America*, 35 F. Supp. 820 (S.D.N.Y. 1940).

<sup>24</sup> See Zeisel, *Statistics as Legal Evidence*, 15 *INTERNATIONAL ENCYCLOPEDIA OF THE SOCIAL SCIENCES* 246-50 (1968); Zeisel, *The Uniqueness of Survey Evidence*, 45 *CORNELL L.Q.* 322-46 (1960).

that occasionally reaches the courts. It may be based on sampling operations, but it need not be. Its ultimate aim is not the measurement of a population, but proof of an individual event.

This type of inference is best exemplified by a case that reached the Swedish courts.<sup>25</sup> It concerned the innocuous charge of overtime parking at a place that had a time limit but no parking meter. To establish that the car had not moved in that time interval, the policeman had recorded the positions of the tire air-valves on one side of the car. The record was kept in the manner pilots indicate directions. The front-wheel valve was recorded pointing at 1 o'clock, the rear-wheel valve at 8 o'clock; at the second visit, after the allowed parking time had elapsed, the valves still pointed at 1 and 8 o'clock. The driver defended, claiming that he had been away in the meantime, had happened to return to the same parking place, and by chance the valves had come to rest in approximately the same position. The court accepted the defense, calculating the probability of such an accidental return to the same position, ( $\frac{1}{12} \times \frac{1}{12} =$ ) 1 in 144, as large enough to establish reasonable doubt. The court added that had all four valves been recorded and found in the same position, the probability of ( $\frac{1}{12} \times \frac{1}{12} \times \frac{1}{12} \times \frac{1}{12} =$ ) 1 in 20,736 would have been accepted as proof beyond reasonable doubt.<sup>26</sup>

The second case of this type is well known to law students:<sup>27</sup> a plaintiff established that she had been negligently run over by "a bus" and, without further evidence of identification, sued the company that had the only regular bus franchise, and thus owned the overwhelming majority of the buses running on the street where the accident occurred. Since she proved that "a bus" had run over her, the probability that it was the defendant's bus was proportionate to that overwhelming majority. The plaintiff claimed that the preponderance of the evidence justified a verdict in her favor. The court, however, agreeing with the reasoning in *Sargent v. Massachusetts Accident Company*<sup>28</sup> that it was "not enough that mathematically the chances somewhat favor a proposition to be proved,"<sup>29</sup> denied recovery, pointing out that a private or chartered bus could have caused the accident.

---

<sup>25</sup> Parkeringsfrågor, II. Tillförlitligheten av det S. K. klocksystemet för parkeringskontroll. *Svensk juristidning*, 47 (1962) 17-32. Cited in Zeisel, *Statistics as Legal Evidence*, 15 INTERNATIONAL ENCYCLOPEDIA OF THE SOCIAL SCIENCES 246 (1968).

<sup>26</sup> This method of computing the probability happened to be biased in favor of the defendant, since it is the formula for computing the combined probability of two independent events occurring simultaneously. The probability, for instance, of throwing two consecutive heads of a true coin is ( $\frac{1}{2} \times \frac{1}{2} =$ )  $\frac{1}{4}$ . But car wheels do not rotate independently, although because of curves, they do not run completely parallel either.

<sup>27</sup> *Smith v. Rapid Transit, Inc.*, 317 Mass. 469, 58 N.E.2d 754 (1945).

<sup>28</sup> 307 Mass. 246, 29 N.E.2d 825 (1940).

<sup>29</sup> *Id.* at 250, 29 N.E.2d at 827.

In the third case, one that did not go beyond the trial court,<sup>30</sup> the operator of a gambling wheel was accused of federal income tax evasion by underreporting his take from the wheel. Because of the underground character of these policy wheels, their mode of operation is not widely known in polite society. It will suffice here to report that in essence they operate in the manner of a roulette wheel in which the bank's winning ratio can be predicted with great precision.<sup>31</sup> In support of the charge, the prosecution offered statistical expert testimony to the effect that, in view of the rules that govern the operation of the wheel, a profit ratio as low as the one claimed by the defendant was extraordinarily improbable. The trial court, after listening to the expert testimony in the absence of a jury, refused to admit it, on the grounds that it was too hypothetical to be relevant to the issue in the case.

A more complicated argument of this kind recently came under the scrutiny of the Supreme Court of California.<sup>32</sup> In that case, one Collins and his common-law wife had been charged with robbery. Collins was a Negro, who at one time or other had worn a beard; his wife was a Caucasian with blond hair which she wore in a pony tail. Collins owned a yellow Lincoln with an off-white top. There was no confession nor a clear identification of the defendants. There was testimony to the effect that a blond woman who wore her hair in a pony tail had been seen running from the scene of the robbery and riding away in a yellow automobile driven by a bearded Negro. To bolster its case, the prosecution produced, in the trial, expert testimony by a professor of mathematics who computed the odds of finding in the Los Angeles area (where the crime had been committed) a second manifestation of that obviously rare combination of a partly yellow car, a blond woman with a pony tail, and a bearded Negro. The jury convicted the defendants, an appeal was denied, but the California Supreme Court reversed the conviction. The decision—unique in its even having a mathematical appendix—quite properly found three errors in the statistical argument.

The first of the court's objections was that the basic probabilities—

---

<sup>30</sup> *United States v. Scott* (unpublished, on file with the University of Chicago Law Review).

<sup>31</sup> If any of the numbers 1 through 36 shows, the bank pays the odds that correspond exactly to the probabilities of winning: 1:1 for even bets (black or red, etc.), 2:1 for bets covering one-third of the numbers, etc. On American wheels the bank wins if either 0 or 00 shows; these are 2 out of 38 positions, hence the banks wins, on the average, 5.26% of the placed bets. European wheels have only the 0-position, hence the bank wins only (1 in 37 =) 2.6% of the bets; this take is further reduced by the rule that if 0 shows, the bank takes only one-half of the even bets.

<sup>32</sup> *People v. Collins*, 68 Cal. 2d 319, 438 P.2d 33, 66 Cal. Rptr. 497 (1968). See also Mode, *Probability and Criminalistics*, 58 J. AM. STATIST. ASS'N 628-40 (1962). See also two earlier cases dealing with this problem area: *People v. Ripley*, 214 N.Y. 95, 108 N.E.2d 200 (1915), and *State v. Sneed*, 76 N.M. 349, 414 P.2d 858 (1966).

that is, the proportion of bearded Negroes and of partly yellow cars in the population—had not been empirically established. The prosecutor, in examining his expert witness, merely asked the witness “to assume” that the probabilities were the ones suggested by the prosecution. The court’s objection to such a procedure is, of course, well taken.<sup>33</sup>

The second of the court’s objections was that the expert witness, in computing the probability of the combined event, had merely multiplied the various individual probabilities, thus treating them as independent events.<sup>34</sup> The court correctly pointed out that, for the computation to be correct, one needed to know the interdependence of the various elements, for instance, whether—this is not the court’s example—a bearded Negro with a Lincoln would be more likely to be found associated with a blond Caucasian woman than a Negro without a beard or with another type car.

The third of the court’s objections was the most interesting one. Prepared to overlook objections one and two, for argument’s sake, the court investigated the computation as it stood and then, applying the so-determined odds to the population of Greater Los Angeles, it concluded:

[E]ven accepting this conclusion as arithmetically accurate . . . . [t]he prosecution’s figures actually imply a likelihood of over 40 percent that the Collinses could be “duplicated” by at least *one other couple* . . . . Urging that the Collinses be convicted on the basis of evidence which . . . establishes no more than this seems as indefensible as arguing for the conviction of X on the ground that a witness saw either X or X’s twin commit the crime.<sup>35</sup>

The court’s argument is unimpeachable as it stands, but it raises several interesting questions for the future. First, to which population is the calculus to be applied? In this case, to the Los Angeles suburb where the crime occurred and the defendants lived, or the total Los Angeles metropolitan area to which the court applied it, or perhaps to the whole state of California? Clearly, this is a factual question that cannot be resolved in the abstract. It depends here on how far the defendants reasonably could have been expected to move.

<sup>33</sup> *People v. Collins*, 68 Cal. 2d at 325, 327-8, 438 P.2d at 36, 38-39, 66 Cal. Rptr. at 500, 502-3. The empirically established probabilities would probably have been smaller, and would hence have hurt the defendant more than the assumed ones. For instance, the prosecutor asked the expert witness to “assume” that the probability of a partly yellow automobile be 1/10. *Id.* at 325-6 n.10, 438 P.2d at 37 n.10, 66 Cal. Rptr. at 501 n.10.

<sup>34</sup> *Id.* at 328-9, 438 P.2d at 39, 66 Cal. Rptr. at 503.

<sup>35</sup> *Id.* at 331, 438 P.2d at 40-41, 66 Cal. Rptr. at 504-5.

But let us assume that all the difficulties can be resolved: empirical, not "assumed," probabilities; proper consideration for their interdependence; and a realistic determination of the population to which the combined probability is to be applied. Suppose that *then* the best estimate is that there is only *one* of these rare combinations. Does not the court's reasoning in the *Collins* case imply that, under these circumstances, the probability calculus would constitute valid evidence?

The next question is even more intriguing. In the *Collins* case, there was no other evidence that would help the jury to decide whether it was this pair or the other (computed) pair that committed the crime; the odds remained 50:50 for each. Suppose, however, there is some additional evidence, say one lone witness, by itself insufficient to eliminate "reasonable doubt," but sufficient to convince in conjunction with the statistical inference that one of the two suspects must be the culprit? Such a result is even more easily achieved in a civil trial, where mere "preponderance" suffices for a verdict. It would, thus, seem that the *Collins* case has opened the door for such a calculus of the probability of a single event.

Seen in the context of these cases, the statistical proof of how the women jurors for the Spock trial were made to disappear would have assumed a special significance had it come under judicial scrutiny. It might have offered the perfect opportunity for establishing the principle that statistical inference may constitute proof of an individual event.<sup>36</sup>

For one thing, the evidence in the present case comes exclusively from the court's own records, albeit summarized and evaluated with the help of statistical expertise. The summarization follows common-sense rules—counting the number of female names in each venire; by applying the chi-square test, the evaluation follows firmly established statistical methods. Secondly, the discrepancy between the percentage of women in the population (56 per cent) and the percentage in the jury venires of the trial judge (15 per cent) is so great as to obviate any argument as to how great a discrepancy must be before a court can be asked for redress. Thirdly, there can be no doubt as to whether

---

<sup>36</sup> The principle of accepting a fact as true if the probability of its being not true is very small is not alien to the law. Paternity blood tests, for example, are, in a very low proportion of cases, misleading because of intervening mutations. Ross, *The Value of Blood Tests as Evidence in Paternity Cases*, 71 HARV. L. REV. 466 (1958). Nevertheless it has been held that such blood tests, if properly made, "should be given such great weight by the court that the exclusion of the defendant as the father of the child follows irresistibly." See, e.g., *State ex. rel. Steiger v. Gray*, 3 Ohio Op. 2d 394, 399, 145 N.E.2d 162, 168 (Juv. Ct. 1957).

the discrepancy was accidental or the result of systematic discrimination: the odds were too formidable.

Finally, it would seem that the case would have bypassed the issue of intent. The prima facie proof of discrimination is so strong that it would have shifted the burden of proof to the court officials. One of them gave testimony on the first phase of the reduction of the proportion of women jurors, in which the percentage of women decreased from over 50 per cent in the original police list to 29 per cent in the central jury box. That testimony did nothing to dispel the evidence of bias. As to the second phase of the reduction, it might have been interesting to hear testimony from the clerk of the court as to how he managed to draw venires with on the average 15 per cent women from a jury box from which the other six judges drew samples with, on the average, 29 per cent women. Yet, while such testimony might have been interesting, it would have probably been irrelevant. This is to say that the fact of discrimination is so clearly established by the statistical evidence that the question of *motive* would have become moot. This was not true in the older cases where systematic intent to discriminate on the part of the jury commissioner or court clerk was essential before discrimination could be found in violation of the fourteenth amendment.<sup>37</sup> Since the older methods of compiling jury lists required the balanced use of many different sources for names—a difficult task to begin with—the law required proof of intent to discriminate before it would find unlawful discrimination.

## VII. PUBLIC LAW 90-274

Recently, however, a more modern notion of how juror lists ought to be compiled has begun to gain ground. From a single list that contains all or most of the persons eligible for jury service, such as the voter lists, city directories, or police lists, a random sample of names was to be selected. This new procedure was in every way superior to the old one. Statisticians have developed techniques of random selection that can be applied without special skill. Most of the time taking every *n*th name will suffice. In 1968 the development culminated in the adoption of Public Law 90-274,<sup>38</sup> which prescribed new rules for selecting juries in the federal courts. The new law orders ran-

---

<sup>37</sup> See *State v. James*, 96 N.J.L. 132, 114 A. 533 (1921).

<sup>38</sup> 28 U.S.C. §§ 1861-71 (Supp. IV, 1969), amending 28 U.S.C. §§ 1861-71 (1964). See Professor Harry Kalven Jr.'s and the author's testimony supporting the bill in *Hearings on S. 383, S. 384, S. 385, S. 386, S. 387, S. 989, S. 1319 Before the Subcomm. on Improvements in Judicial Machinery of the Senate Comm. on the Judiciary*, 90th Cong., 1st Sess. 117-73 (1967).

dom selection of jurors, primarily from voter lists, to guarantee a true cross-section of the population eligible for jury service.

Besides simplifying the selection procedure, the adoption of strict random selection has another advantage. If the rules of random selection are violated, the resulting list reveals the flaw. In addition, because of the simplicity of its application, any infraction of the rule will itself—irrespective of whether it was intentional, or even accidental—violate the law. If the rules of random selection are violated, the motive should become irrelevant, and from this there should result a great simplification of the law of jury discrimination.<sup>39</sup>

Now that the new selection rules have been in force for some time, it is a pleasure to report that women have been restored to their rightful place on our federal juries. The percentage of women in the venires selected by the United States District Court of the District of Massachusetts under the new jury statute has markedly increased. In five venires for the colleagues of the judge who tried the Spock case, the proportion of women was 38 per cent, 41 per cent, 42 per cent, 44 per cent, 45 per cent—42 per cent on the average. The Spock trial judge himself, topping them all, had a venire in which the proportion of women was 49 per cent.

If ever there was a federal statute that should serve as a model for the states it is Public Law 90-274. It removes many of the traditional ambiguities<sup>40</sup> in the jury selection process by devising simple me-

---

<sup>39</sup> See *Ballard v. United States*, 329 U.S. 187, 193 (1946); *Thiel v. Southern Pacific Co.*, 328 U.S. 217 (1946); *McNabb v. United States*, 318 U.S. 332 (1943); *Kentucky v. Powers*, 139 F. 452 (C.C.E.D. Ky. 1905); *State v. James*, 96 N.J.L. 132, 114 A. 553 (1921); *Jaurez v. State*, 102 Tex. Crim. 297, 277 S.W. 1091 (Ct. Crim. App. 1925).

The "systematic or intentional exclusion" requirement necessary to show discrimination is also discussed in recent cases. See *Hansen v. United States*, 393 F.2d 763 (8th Cir. 1968); *United States v. Ramseur*, 378 F.2d 902 (6th Cir. 1967); *United States v. Birrell*, 276 F. Supp. 798 (S.D.N.Y. 1967); *Shelton v. Jones*, 272 F. Supp. 139 (W.D. Va. 1967). However, 1968 federal legislation imposes an objective procedure—random selection of jurors—which eliminates the need to show an intent to discriminate. See text at note 40 *infra*.

An interesting corollary to the issue of discrimination in jury selection is the question of who has standing to object to the make-up of the jury. Early cases dealing with racial discrimination held that only a Negro could claim that Negroes were discriminated against in the selection of grand or petit juries. See *State v. Jones*, 44 Del. 372, 57 A.2d 109 (1947); *Griffin v. State*, 183 Ga. 775, 190 S.E. 2 (1937); *State v. Carlino*, 98 N.J.L. 48, 118 A. 784 (1922), *aff'd*, 99 N.J.L. 292, 122 A. 830 (1923); *State v. James*, 96 N.J.L. 132, 114 A. 553 (1921); *Commonwealth v. Garletts*, 81 Pa. 271 (Super. Ct. 1923); *McKinney v. State*, 3 Wyo. 719, 30 P. 293 (1892). Today objection to discrimination in juror selection may be asserted by any litigant even though not a member of the excluded class, *Rabinowitz v. United States*, 266 F.2d 34 (5th Cir. 1966), and even though the litigant has not shown that the jury imbalance resulting from discrimination has prejudiced him. See *United States v. Dennis*, 183 F.2d 201, 216 (2d Cir. 1950), *aff'd*, 341 U.S. 494 (1951); *United States v. Cohen*, 274 F. Supp. 724, 737 (D.C. Md. 1967).

<sup>40</sup> Not all ambiguities, because there remains, one may hope not for long, the problem

chanically executed procedures. More and more states will have their voter lists on computer tape, and, as a consequence, all the selection—and all the paper work—will be done by machines. The new statute will have still another important effect: jury service, traditionally limited to a small section of the population, recalled for service every two years, will now be spread among many more people. This should make it possible to increase considerably the time between tours of jury duty and thereby make jury service a tolerable burden, even for those who in the past have been excused from it for reasons of hardship.<sup>41</sup> Thus, by expanding jury service for the first time to the total eligible population and thereby greatly reducing the burden to the individual jurors, Public Law 90-274 will also correct the system's major shortcoming: the hitherto reduced participation of the lower income classes.

---

of regions where the voter list does not yet represent the population. In these regions, the statute demands that the voter lists be supplemented by other sources.

<sup>41</sup> The increased jury fee, now \$20.00 per day in the federal courts, should also work toward this end.