2007

Judgement for Judges: What Traditional Statistics Don't Tell You About Causal Claims

Steven N. Goodman

JUDGMENT FOR JUDGES: WHAT TRADITIONAL STATISTICS DON'T TELL YOU ABOUT CAUSAL CLAIMS

Steven N. Goodman, M.D.[*]

INTRODUCTION

The determination of the likelihood that a given agent or exposure caused an injury to another person is a critical foundation of the tort system. When this determination is based on evidence derived from statistical analyses of scientific studies, it is critical that a judge or jury reviewing these analyses or the testimony of experts understand where statistics leave off and judgment begins. Too often, the determination of causality in the tort setting is left to a formulaic misapplication of what are regarded as scientific criteria for proof. This essay will review how scientific and legal judgment must augment statistical measures in addressing questions of both general and specific causation. This article will explore the traditional method of statistical inference, based on hypothesis testing and P-values, then an alternative based on Bayes Theorem, with the Bayes Factor as a measure of evidence.

I. TRADITIONAL STATISTICAL SIGNIFICANCE APPROACH VS. BAYES THEOREM

First, this article will explore in some depth the meaning of a finding of "statistical significance," which is the cornerstone

---

[*] Johns Hopkins University, Associate Professor of Oncology, Pediatrics, Epidemiology and Biostatistics; sgoodman@jhmi.edu.

of virtually all claims of scientific proof. It will show that standard statistical indices do not answer the question most judges are interested in, for example, how likely an observed relationship reflects a true one. It will set aside for the moment the question of whether that relationship, once established, is causal. If no relationship is found, then the causal question is moot. So it is reasonable first to examine the question of whether an observed statistical relationship is due to chance, or an alternative, non-chance explanation.

A simple example will serve as the departure point for this exploration. Let us imagine that a plaintiff claims that proximity to a local power plant is responsible for an increased leukemia risk in a town, and the court is interested in the truth of that claim. An epidemiologic study addressing this question is designed and conducted.[1] Traditional statistical methods approach the process of inference as follows:

1. **State a null hypothesis (Ho)**: There is no effect of the proximity to plant on leukemia risk.

2. **Calculate the rarity of the observed geographic pattern of leukemia cases under that null hypothesis**: The measure of "rarity" is measured by an index called a "P-value."[2] The P-values for the pattern is reported as equal to 0.03.

3. **If observed data are "rare enough" under the null hypothesis then "statistical significance" is declared**. Statistical significance is typically defined as a P-value less than 5 percent. As the reported P-value of 3 percent here is less than this, one "rejects the null hypothesis" and declares that the association between leukemia risk and power plant proximity has been scientifically demonstrated.

On the surface, the above procedure seems appealingly

---

[1] One aspect of this process which will not be discussed here is how the study design or statistical analysis eliminates variables other than distance from the plant as determinants of outcome, factors known as "confounders," and whose control is critical to causal inference.

[2] *See infra* Part II.A. for further discussion.

logical and "scientific" (i.e., objective). However, the question might arise, where is there room for judgment? The short answer is that, in a formal sense, there is none. Judgment makes its appearance only when we recognize that the procedure above does not, in fact, address the question posed earlier: "What is the probability that chance is the explanation for the observed pattern of illness around the plant?" The procedure described above is a decision rule, not directly addressing the question we need answered.[3] We will see this as we closely examine the logic of standard statistical procedures, starting with the meaning of the central statistical index, the P-value.

A. The P-value

Introduced as an inferential tool by R.A. Fisher in the 1920s, the P-value is the central evidential index that undergirds the calculations of traditional, "frequentist" statistics.[4] Its definition is as follows:

> Under the hypothesis of no effect, the P-value is the probability of observing a result equal to or more extreme than the observed data (relative to the null hypothesis).

This can be written as follows (where Prob=probability):

> **Prob** (Future observed effect   Current observed effect, given that the True effect = 0)

A graphical counterpart to this equation can be found in Figure 1. The "Future" effect refers to an effect that might be observed if we exactly repeated the study under the identical conditions. When examining this definition carefully, it does not have a clear common sense interpretation. To underscore this, let us consider the guidance offered in a textbook titled *Intuitive*

---

[3] *See generally* Steven N. Goodman & Richard Royall, Commentary, *Evidence and Scientific Research*, 78 AM. J. PUB. HEALTH 1568 (1988); *See also* RICHARD ROYALL, STATISTICAL EVIDENCE: A LIKELIHOOD PARADIGM 1-33 (Chapman & Hall 1997).

[4] *See* RONALD A. FISHER, STATISTICAL METHODS FOR RESEARCH WORKERS 24-76 (13th ed. 1958).

96   *JOURNAL OF LAW AND POLICY*

Biostatistics. "Thinking about P-values seems quite counterintuitive at first, as you must use backwards, awkward logic. Unless you are a lawyer or a Talmudic scholar . . . you will probably find this sort of reasoning a bit uncomfortable."[5]

The "backwards" logic of the P-value is related to the fact that it is a *deductive* probability statement. It assigns a probability to the data under the assumption that we know the null hypothesis to be true. But the question at hand requires an *inductive* statement, a statement about the probability of a hypothesis based on the evidence, the reverse or "backwards" direction. The difference is equivalent to the contrast between the probability of the evidence, assuming a defendant were innocent (deductive), and the probability of innocence given the evidence (inductive). The route from the P-value to a probability of a given claim being right or wrong is circuitous, and not part of conventional statistical approaches.

To see this, let us contrast the P-value with the mathematical counterpart of the question posed earlier: given the observed effect, what is the chance that the true effect is zero (or conversely, non-zero)? This can be written as follows:

Probability of a claim: Prob (True effect = 0 given the Observed effect)

This probability cannot be calculated with standard statistical methods. However, there is an inductive inferential calculus, known as Bayes Theorem, about which much has been written in application to the law.[6] Figure 2 shows Bayes Theorem, first written using the legal analogy and then its statistical counterpart, where the hypothesis of "innocence" is instead the hypothesis of "no effect."

---

[5] HARVEY MOTULSKY, INTUITIVE BIOSTATISTICS 96 (Oxford Univ. Press 1995).

[6] *See* Stephen E. Fienberg & Joseph B. Kadane, *The Presentation of Bayesian Statistical Analyses in Legal Proceedings*, 32 STATISTICIAN 88 (1983); Donald A. Berry, *Inferences Using DNA Profiling in Forensic Identification*, 6 STAT. SCI. 175 (1991); Donald A. Berry & Seymour Geisser, *Inference in Cases of Disputed Paternity*, *in* STATISTICS AND THE LAW 353-82 (Morris H. DeGroot, Stephen E. Fienberg & Joseph B. Kadane eds., John Wiley & Sons 1986).

B. Bayes Theorem

Bayes Theorem is a fascinating mathematical and philosophic entity. The mathematics of the theorem are simple and incontrovertible, but its implications profound. If we are dealing with modeling games of chance, or medical diagnoses or other situations where all the relevant probabilities are well described, there is no question as to its relevance and correctness. Where its application becomes trickier, and more controversial, is in the realms of inference highlighted here: statistical and legal. The controversy stems from its requirement for a "prior probability" of a hypothesis. In the legal realm this hypothesis could be one of innocence, and in the scientific arena a hypothesis of no effect. Thus, how to assign and then interpret probabilities on these hypotheses is controversial. It will not be the application of Bayes Theorem that we will focus on here, but rather how it illustrates the flaws in logic—and room for judgment—in the standard approaches to statistical proof.

It is worth first contrasting to the Bayesian measure of evidence—the Bayes Factor—with the P-values. The Bayes factor is simply a comparison of how likely the evidence is under two competing hypotheses. It is different from the P-values in two critical ways. First, it is comparative. Evidence that is rare under the null hypothesis is not considered evidence against it unless that same evidence can be shown to be more common under the competing hypothesis. In contrast, the competing or "alternative" hypothesis has no role in calculating the P-values. Second, because the Bayes Factor is comparative, it can be negative or positive, i.e. support either hypothesis relative to the other. In contrast, the P-values is only negative, i.e., against the null hypothesis, making it impossible to quantify evidence that supports the hypothesis of no effect (or of innocence).[7] Both of these features are captured in colorful quotations from a noted epidemiologist from the 1940's:

---

[7] *See* Goodman and Royall, *supra* note 3, at 1569 (citing Jerome Cornfield, *Sequential Trials, Sequential Analysis, and the Likelihood Principle*, 20 AM. STATIST. 18 (1996)); ROYALL, *supra* note 3, at 68-71.

> [T]he argument does not seem to accord with what would be the mode of reasoning in ordinary rational discourse . . . . Suppose I said, "Albinos are very rare in human populations . . . . Therefore, if you have taken a random sample of 100 . . . and found in it an albino, the population is not human." . . . I believe the rational retort would be, "If the population is not human, what *is* it?" . . . With the *corpus delicti* in front of you, you do not say, "Here is evidence against the hypothesis that no one is dead." You say, "Evidently someone has been murdered."[8]

In comparing standard methods to Bayesian approaches to inference, it first must be noted that they are asking different questions and have different aims. It is in understanding these different aims that we will see why judgment seems to play a role in one but not the other. The aim of the Bayesian inference is to calculate the probability that a knowledge claim or hypothesis is true or false. It does this, however, at a price—a price that requires specification of the probability of that same relationship in the absence of the current empirical evidence. In many instances, that prior probability will be unavoidably subjective, and differ among experts, leading to this inferential approach being tarred as "subjective" or "nonscientific."

The central question that arises is, what would an inferential system look like that avoided the seeming subjectivity of Bayes Theorem? This was exactly the conundrum faced by R.A. Fisher in the 1920s and Jerzy Neyman and Egon Pearson in the 1930s, which leads them to develop a new "frequentist" statistical approach to this question, the model of which was presented earlier.[9]

"Frequentist" is a term referring to a definition of probability that requires a well defined, mathematically

---

[8] Joseph Berkson, *Tests of Significance Considered as Evidence,* 37 J. AM. STAT. ASS'N 325, 326 (1942) (emphasis in original).

[9] *See supra* Part II.A. and *infra* note 11 for a further discussion and explanation.

justifiable, empirically observable foundation for the calculation of the probability. The definition, most powerfully articulated by Richard von Mises in the late 1920s, defines probability as the relative frequency of objects with a particular trait amidst a large collective of otherwise indistinguishable objects, such as the fraction of all 60-year olds over 6 feet tall.[10] In theory, any scientist observing the same collective would measure the same relative frequency. This definition represented an effort to put probability on scientific par with other fundamental measures of nature, like mass, velocity and length.

Thus, a system of inference based on frequentist probability would have to be based on probabilities that were indisputable, i.e., governed by clearly defined probability distributions, such that any scientist armed with the same distribution would calculate the same number. That is a characteristic of deductive probabilities: if you accept the premises, you accept the subsequent calculations. The notion of a probability of a *hypothesis* is anathema; such a number is not deemed a probability at all, since there can be no "long run" or "collective" of hypotheses. So the frequentist abandons, at the outset, any notion that they will be assessing the credibility of a truth claim. Instead, the frequentist is concerned with long-run probabilities, i.e., probabilities of possible *outcomes,* defined against a theoretical infinite number of repetitions of an experiment.

The question then becomes, if the frequentist declines to calculate the probability of truth, what is the goal of the system of traditional statistics? The answer is, to control the number of errors over the long run, but not calculate the chance of error in any particular case. This was articulated clearly and forcefully by Neyman and Pearson in their classic paper:

> [N]o test based upon the theory of probability can by itself provide any valuable evidence of the truth or falsehood of that hypothesis.
>
> But we may look at the purpose of tests from another

---

[10] RICHARD VON MISES, PROBABILITY, STATISTICS AND TRUTH 11-29 (Dover Publ. ed., George Allen & Unwin Ltd. 2d ed. 1957).

> view-point. Without hoping to know whether each separate hypothesis is true or false, we may search for rules to govern our behaviour with regard to them, in following which we insure that, in the long run of experience, we shall not be too often wrong.[11]

It is the opening of the above passage that is critical to understand. The "theory of probability" referred to here is the frequentist definition. Thus, the standard statistical method of "hypothesis testing" is roughly akin to a judicial system where the goal is to minimize the collective number of incorrect verdicts, without regard to whether each individual is guilty or innocent. The problem encountered with this is analogous with the practice of "profiling" by police; it may indeed identify a class of individuals at a higher risk for some behavior, but to blindly apply a group characteristic to each member of that group is recognized as unjust. A practice that might work "on average" can be profoundly and recognizably wrong in particular cases.

The Bayesian definition of probability, in contrast, concerns itself precisely with what is eschewed above; the degree of belief that a specific hypothesis is true or false. Bayes Theorem, as defined earlier, tells us how this can be calculated. It tells us that the purpose of evidence, whether scientific or legal, is to change the probability that a given hypothesis is true. It tells us that if a hypothesis is more or less likely before seeing the evidence, it is correspondingly more or less likely afterwards; its prior plausibility affects its plausibility after considering the new evidence.

In both the legal and scientific settings, this requires a close look at the details of a particular case, instead of applying similar rules to all cases. In the legal realm, in a case based on circumstantial evidence, a key requirement affecting prior probability could be a motive for the crime. In the absence of such a motive, the prior probability might be so low that only

---

[11] Jerzy Neyman & Egon S. Pearson, *On the Problem of the Most Efficient Tests of Statistical Hypotheses*, 231 PHIL. TRANSACTIONS ROYAL SOC'Y 289, 291 (Series A. 1933).

extraordinary circumstantial evidence would be sufficient to convict.

In science the same degree of evidence can result in a different conclusion if the prior probability is different. The prior probability requires a close examination of many of the factors mentioned above, such as the biologic plausibility of the relationship, and the strength of prior empirical evidence. The reliability of that evidence is determined partly through the strength of the design and conduct of the experiments that produced it. The strength of an experiment typically is not describable with numbers. One must rely on expert judgment to help assess it. Various scientific groups have come up with crude aids to facilitate such assessment, such as the hierarchy of evidence used by the U.S. Preventive Services Task Force.[12]

In addition to the notion of prior probability, another component missing from standard approaches is a formal notion of statistical "evidence," the only language is that of procedural error rates.[13] In contrast, the Bayes Factor is a measure of evidence with appealing conceptual simplicity. It allows us to make the connection between traditional P-values and posterior probabilities. Table 2 shows a Bayesian-frequentist "Rosetta Stone," in that it demonstrates the maximum effect that a result with a given P-value could have on the prior probability of a hypothesis when viewed through a Bayesian lens.[14] Even though these results represent maximum effects, they are far lower than many judges, and even scientists, are aware of. For example, a P-value of 0.03 raises the probability of a 50:50 hypothesis to at most 91 percent, i.e., there is still almost a one-in-ten chance that it is wrong. If a hypothesis is implausible or initially unlikely, less than 25 percent probable, the table tells us that a P-value of 0.03 cannot raise the probability of such a hypothesis

---

[12] *See infra*, Table 1 for an illustration of the hierarchy of evidence.

[13] *See* Goodman, *supra* note 3, at 1569-70; Steven N. Goodman, *Towards Evidence-Based Medical Statistics, 1: The P-Value Fallacy*, 130 ANNALS OF INTERNAL MED. 995 (1999); Steven N. Goodman, *Towards Evidence-Based Medical Statistics 2: The Bayes Factor*, 130 ANNALS OF INTERNAL MED. 1005 (1999).

[14] *See infra*, Table 2.

to being more than 78 percent probable, i.e., a 22 percent chance of being wrong. Even if the P-value is 0.01, a hypothesis starting with 25 percent probability still has a 10 percent or greater chance of being false.

This table shows us both that (a) the prior probability is a critical factor in determining the probability that an observed relationship is true, given the evidence, and (b) that the evidential force of P-value is lower than their actual value suggests, and has little relationship to the probability of the truth of the null hypothesis.[15] Finally, this table only applies to "ideal" experiments, i.e., those with the strength of a randomized clinical trial. If these P-values are derived from observational studies, which are the typical designs used in toxic tort cases, the effect of the statistical evidence is weaker.

II. JUDGMENT AND SPECIFIC CAUSATION

Specific causation is another domain in which mechanistic rules have seemingly eliminated the need for judgment. It is commonly known that for an exposed individual, a condition passes the "more likely than not" criteria for specific causation if the relative risk exceeds 2 (or $RR > 2$). The thinking behind this is based on a very simplistic model of causation. If the exposure doubles the baseline risk of contracting a condition, and an exposed person has the condition, then half of his or her risk is thought to be due to the baseline risk with an equal degree of additional risk due to the exposure. This additional risk is called the "attributable risk" in the exposed individual. Hence, at relative risks higher than 2, this attributable risk will be larger than the baseline risk, and the individual is regarded as "more likely than not" to have incurred the condition from his or her exposure.

Several epidemiologists have written eloquently on the flaw in this logic, which is easily demonstrated if one introduces a time dimension into the disease process.[16] Consider a process

---

[15] *See generally infra*, Table 2.

[16] *See* James M. Robins, *Should Compensation Schemes be Based on the*

that advances the appearance of disease by a decade. Everyone exposed who has developed the condition has lost 10 years of disease-free life, but there are no "excess cases" of the disease, since no one develops the disease who would not have otherwise. In such a situation, the relative risk is 1.0, but the exposure has had an adverse health impact in 100 percent of exposed individuals with the condition.

Figure 4 shows us that the probability that the exposure played an important role in the disease's occurrence cannot be discerned from the data without some knowledge about the mechanism by which the disease is produced.[17] The upper panel of the figure shows the situation described above, in that each individual has had the time of their illness advanced by 10 years.[18]

Thus, in an epidemiologic study, we would find exposed individuals with the disease between the ages of 40 to 80, whereas individuals not exposed to the disease are found to be between the ages of 50 to 90. We would not be able to tell from such data whether the exposure advanced the disease's appearance by a decade for each person, or whether the person who would have developed their disease at age 80 instead developed it at age 40, leaving everyone else unaffected. In both cases, there is a collective loss of four decades of disease-free life. But without knowledge of how the disease mechanism works, we cannot know how it is distributed among the individuals. So the fraction of individuals affected by the exposure—sometimes called the "probability of causation"—in this case varies from 20 percent to 100 percent.[19] Only if we had some measurable biomarker that told us the etiology of the

---

Probability of Causation or Expected Years of Life Lost?, 12 J. L. & POL'Y 537 (2004); Sander Greenland & James M. Robins, *Epidemiology, Justice and the Probability of Causation*, 40 JURIMETRICS 321 (2000); Sander Greenland, *Relation of Probability of Causation to Relative Risk and Doubling Dose: A Methodologic Error That Has Become a Social Problem*, 89 AM. J. PUB. HEALTH 1166 (1999).

[17] *See infra,* Figure 4.

[18] *Id.*

[19] *See* Robins, *supra* note 16, at 537-48.

disease in a particular person could we discern empirically which of these processes were operating.

This example also illustrates another problem with using the relative risk when time is a factor, as it almost always is. Let us assume that the individuals in question were exposed on the job, and that they were all 30 years old at the time. Let us imagine that an epidemiologic study is immediately mounted, and the individuals are followed forward in time, along with a similar-sized cohort of colleagues at work who were not exposed. After 20 years of follow-up, 2 cases have developed in the exposed group, and 1 in the unexposed, for a relative risk of 2. After 30 years, the corresponding numbers are 3 and 2, for a relative risk of 1.5. After 40 years, the relative risk is 4 over 3, after 50, 4 over 5, and finally, after 60 years, the relative risk is 5 over 5, or 1.0, since 5 subjects in both groups have developed the disease.

We could change the numbers here so they are as high as we wish at the beginning. The point is that the relative risk is not a constant number, and will vary according to how much time the subjects are observed. It will typically be highest with short follow-ups, and decrease over time. This is another reason why relative risk is a poor reflection of the likelihood that an illness observed in an exposed person is due to that exposure. In general, the relative risk serves as a lower bound on the fraction of cases induced by the exposure, but the upper bound is always 100 percent.

CONCLUSION

Simple rules governing either statistical or causal inference are invariably misleading or outright wrong. Scientific experts or legal arguments that invoke such rules are making implicit assumptions, which may not be defensible. Judges must be aware that the probability of the truth of causal claims is not calculable from the data alone, and any claim to the contrary is made out of ignorance or an intent to deceive. While judges cannot be expected to become methodologic experts themselves, they can play a critical role in eliciting the foundations for

judgments that are implicit in any causal claim. These foundations include:

For general causation:

1. The prior plausibility of the hypothesis being considered derived from prior studies, known biology and from mechanistic reasoning.

2. The strength of the design and conduct of the experiments in the evidence base.

3. The internal coherence of the evidence base with a proposed or known biologic mechanism.

For specific causation:

1. How well established is the biologic mechanism for effect.

2. Whether there is any biologic marker that allows inference about disease etiology in a specific case.

3. If the RR > 2 criterion is being used, whether the disease is an all-or-none phenomenon within the time period of observation (e.g., symptoms of food poisoning shortly after a group event), or if it emerges over an extended time with the timing and fact of occurrence both being relevant. If the latter scenario is true, the RR > 2 criterion is invalid.

The principles above apply with particular force to the toxic torts arena where one quite frequently encounters weak designs and poorly understood biologic processes, which translate into low prior probabilities and weak evidence. While rigorous thinking, formal analysis and systematic approaches to synthesis are hallmarks of the scientific approach, both scientific and legal judgment play prominent roles in ascertaining whether a claim of injury due to toxic exposure is likely to be true, and the nature of judgments being applied by the experts must be understood by the presiding judge.

106   *JOURNAL OF LAW AND POLICY*

Table 1: Hierarchy of Clinical Research Designs as Per the US Preventive Services Task Force

Critical research designs at the top of the table generally produce evidence of higher reliability or "strength" than those below.[20]

EVIDENCE GRADE	DESIGN
I	Evidence from at least one properly randomized controlled trial
II-1	Evidence from at least one well-designed, non-randomized controlled trial
II-2	Evidence from well-designed cohort or case-controlled studies, preferably from more than one center or research group
II-3	Evidence from multiple time series with or without the intervention
III	Opinions of respected authorities based on clinical experience, descriptive studies and case reports or reports of expert committees

---

[20] Adapted from Russell P. Harris et al., *Current Methods of the US Preventive Services Task Force: A Review of the Process,* 20 AM. J. PREVENTIVE MED. 21 (2001).

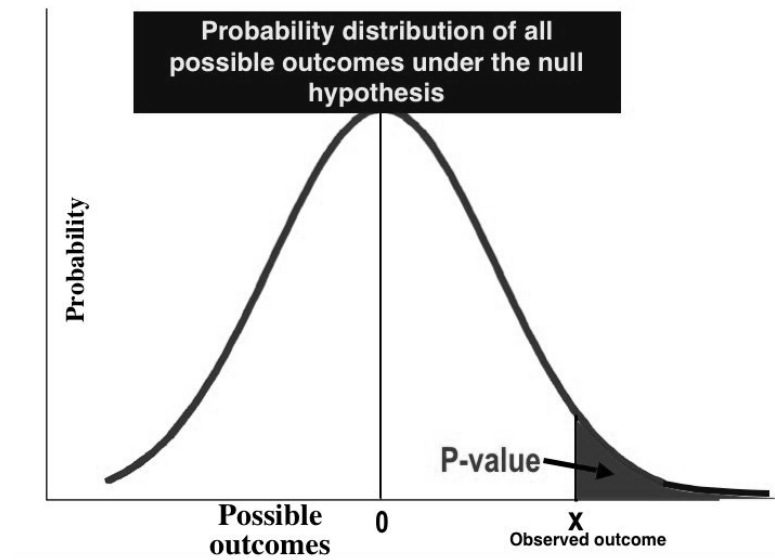## *JUDGMENT FOR JUDGES*   107

Table 2: A Frequentist-Bayesian Translation Table

Column 3 gives starting (or "prior") probabilities of non-null hypotheses, and Column 4 shows the maximum degree to which a given P-value can move a non-null hypothesis from a given starting probability to a final probability.

		Increase in Probability of Ha	
P-value	**Strength of Evidence**	**From (%)**	**To at Most (%)**
0.10	Weak	25	56
		50	79
		83	95
0.05	Moderate to Weak	25	69
		50	87
		74	95
0.03	Moderate	25	78
		50	91
		67	95
0.01	Moderate to Strong	25	90
		50	96
		40	95
0.001	Strong to Very Strong	25	98.5
		50	99.5
		9	95

Figure 1: Graphical Representation of the P-value

The curve shows the probability of each possible outcome under the null hypothesis of no effect. The P-value is the probability of the observed outcome plus all more extreme outcomes upon exact repetitions of an experiment.

JUDGMENT FOR JUDGES   109

Figure 2: Bayes Theorem

The vertical line in the Bayes Factor equation should be read as "given that" or "if."

a.) "Legal" Version of Bayes Theorem

$$\begin{pmatrix} \text{Odds of innocence} \\ \text{after seeing evidence} \end{pmatrix} = \begin{pmatrix} \text{Odds of innocence} \\ \text{before seeing evidence} \end{pmatrix} \times \underbrace{\begin{pmatrix} \dfrac{\text{Prob(Evidence}|\text{ Innocent)}}{\text{Prob(Evidence}|\text{ Guilty)}} \end{pmatrix}}_{\text{Bayes Factor}}$$

b.) Statistical Version of Bayes Theorem

$$\begin{pmatrix} \text{Odds of No Effect} \\ \text{after seeing data} \end{pmatrix} = \begin{pmatrix} \text{Odds of No Effect} \\ \text{before seeing data} \end{pmatrix} \times \underbrace{\begin{pmatrix} \dfrac{\text{Prob(Data}|\text{ No effect)}}{\text{Prob(Data}|\text{ Some effect)}} \end{pmatrix}}_{\text{Bayes Factor}}$$

110   *JOURNAL OF LAW AND POLICY*

Figure 3: Graphical Representation of the Calculation of the Attributable Risk (AR$_e$) in Exposed Persons

If the risk is more than doubled in the exposed group then AR$_e$ > 50% of the total risk.
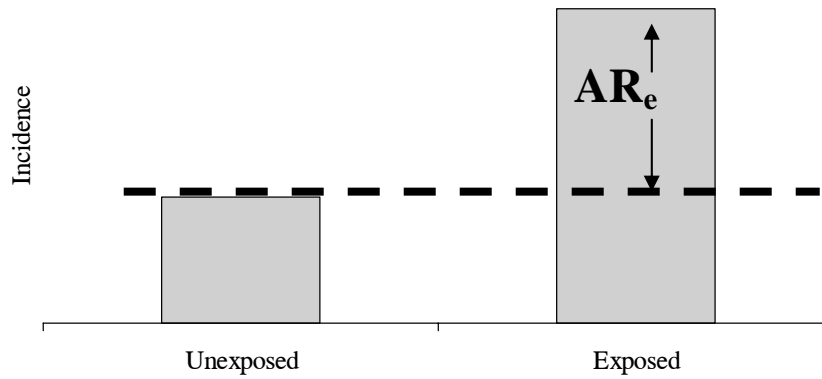
JUDGMENT FOR JUDGES   111

*Figure 4: Illustration of Different Patterns of Causation with the
Identical Patterns of Epidemiologic Data*

In Scenario 1, 100 percent of exposed subjects with disease lose
a decade of life. In Scenario 2, 20% of exposed subjects with
disease lose 50 years of life.[21]

	AGE OF DISEASE OCCURRENCE					
SCENARIO 1	40	50	60	70	80	90
Exposed subjects	1	1	1	1	1	
Identical subjects, not exposed		1	1	1	1	1
SCENARIO 2	40	50	60	70	80	90
Exposed subjects	1	1	1	1	1	
Identical subjects, not exposed		1	1	1	1	1

---

[21] Adapted from Robins, *supra* note 16, at 542.