



UNIVERSITY OF
ILLINOIS PRESS

Base Rates in Bayesian Inference: Signal Detection Analysis of the Cab Problem

Author(s): Michael H. Birnbaum

Source: *The American Journal of Psychology*, Vol. 96, No. 1 (Spring, 1983), pp. 85-94

Published by: University of Illinois Press

Stable URL: <http://www.jstor.org/stable/1422211>

Accessed: 03-05-2017 00:07 UTC

JSTOR is a not-for-profit service that helps scholars, researchers, and students discover, use, and build upon a wide range of content in a trusted digital archive. We use information technology and tools to increase productivity and facilitate new forms of scholarship. For more information about JSTOR, please contact support@jstor.org.

Your use of the JSTOR archive indicates your acceptance of the Terms & Conditions of Use, available at
<http://about.jstor.org/terms>



University of Illinois Press is collaborating with JSTOR to digitize, preserve and extend access to *The American Journal of Psychology*

Base rates in Bayesian inference: Signal detection analysis of the cab problem

MICHAEL H. BIRNBAUM

University of Illinois at Urbana-Champaign

Several investigators concluded that humans neglect base rate information when asked to solve Bayesian problems intuitively. This conclusion is based on a comparison between normative (calculated) and subjective (responses by naive judges) solutions to problems such as the cab problem. The present article shows that the previous normative analysis was incomplete. In particular, problems of this type require both a signal detection theory and a judgment theory for their proper Bayesian analysis. In Bayes' theorem, posterior odds equals prior odds times the likelihood ratio. Previous solutions have assumed that the likelihood ratio is independent of the base rate, whereas signal detection theory (backed up by data) implies that this ratio depends on base rate. Before the responses of humans are compared with a normative analysis, it seems desirable to be sure that the normative analysis is accurate.

Recent papers contend that humans do not make statistical inferences by means of Bayes' theorem (Hammerton, 1973; Kahneman & Tversky, 1973). In particular, it has been argued that judges neglect base rate information. This conclusion, which is called the "base rate fallacy," is based on the finding that when judges are asked to solve a statistical problem intuitively, the modal response differs from the Bayesian solution. A demonstration, which has already become something of a classic, is the cab problem, variations of which have been investigated by Kahneman and Tversky (1973), Bar-Hillel (1980), Lyon and Slovic (1976), Fischhoff, Slovic, and Lichtenstein (1979),

and others. Tversky and Kahneman (1980) state the cab problem as follows:

A cab was involved in a hit-and-run accident at night. Two cab companies, the Green and the Blue, operate in the city. You are given the following data:

- (i) 85 % of the cabs in the city are Green and 15 % are Blue.
- (ii) A witness identified the cab as a Blue cab. The court tested his ability to identify cabs under the appropriate visibility conditions. When presented with a sample of cabs (half of which were Blue and half of which were Green) the witness made correct identifications in 80 % of the cases and erred in 20 % of the cases.

Question: What is the probability that the cab involved in the accident was Blue rather than Green?

The modal response by untrained judges is usually observed to be about .8, whereas the so-called “normative” solution is supposed to be .41. The purposes of this article are to question the previous solution and to call attention to the fact that the proper normative solution to the cab problem is a bit more complicated than previously supposed in papers on the “base rate fallacy.” In particular, it will be argued that the solution called “normative” in previous papers on this topic makes (implicitly) a very unrealistic assumption. The basic purpose of this paper will be to discuss the implications of signal detection theory for the normative solution.

Bayes’ theorem

Bayes’ theorem can be written in odds form as follows:

$$\Omega_1 = \Omega_0 \frac{P(\text{“B”}|B)}{P(\text{“B”}|G)} \tag{1}$$

where Ω_1 = posterior odds [$P(B|\text{“B”})/(1 - P(B|\text{“B”}))$], Ω_0 is the prior odds of a Blue cab [from the base rate, $P(B)/P(G)$]; $P(\text{“B”}|B)$ is the probability that the witness reports “Blue” given the cab is actually Blue (hit rate); and $P(\text{“B”}|G)$ is the probability the witness reports “Blue” given the cab is actually Green (false alarm rate). The answer to the cab problem, $P(B|\text{“B”})$, is given by the expression $\Omega_1/(1 + \Omega_1)$.

Previous solution

The cab problem is interpreted to imply that when the two colors are presented with equal frequency, the hit rate is .80 and the false alarm rate is .20. Lyon and Slovic (1976) pointed out that early versions of the problem did not clearly identify the hit rates and false alarm rates, giving the reader instead only the percentage of correct identifications. In newer versions of the problem, the judge is informed that the witness made an equal number of errors on each type of cab.¹

The hit and false alarm rates are represented in Figure 1 by the open circle labeled a . The problem also states that $P(B)$, the probability of a blue cab, is .15, so the prior odds are $.15/.85 = .176$.

If it is assumed that the ratio of hit rate to false alarm rate is independent of the proportion of each cab color (dashed straight line in Figure 1), Equation 1 implies that $\Omega_1 = (.176)(.80/.20) = .706$ or $P(B|“B”) = \Omega_1/(1 + \Omega_1) = .414$, which is the so-called “normative” answer utilized in this work as the number to be compared with the judge’s response. Judges often respond “.8” instead of .41, which led experimenters to propose that judges neglect base rate information (Bar-Hillel, 1980; Tversky & Kahneman, 1974, 1980, 1982). However, the previous solution (.41) implicitly assumes that $P(“B”|B)/P(“B”|G)$ is independent of the base rate, which is inconsistent with theories and experiments in the area of signal detection. It will be shown that the answer .8 is compatible with Equation 1 and signal detection theory.

Signal detection theory

Although it may be reasonable to assume for some sources of information that the ratio of hit rate to false alarm rate is independent of signal probability, this assumption would not be realistic for the human witness² (Schum, 1981). A good deal of research has shown that a better approximation to the behavior of witnesses would be the curve shown in Figure 1 (Green & Swets, 1966). The curve in Figure 1 is based on the theory that each color of cab produces a normal distribution on a discriminial continuum, as shown in the inset of Figure 1. The witness responds “Blue” when the value on the Green-Blue continuum exceeds the response criterion, indicated by one of the vertical lines in the inset. For simplicity, the two variances are assumed to be equal and can be set to 1.0.

This signal detection theory can be written:

$$\begin{aligned} P(“B”|B) &= N(b - t) \\ P(“B”|G) &= N(g - t) \end{aligned} \quad (2)$$

where N represents the cumulative standard normal density function, g is the mean value of Green cabs on the discriminial continuum, b is the corresponding value for Blue cabs, and t is the criterion for responding “Blue” or “Green.” It follows that $d' = b - g = N^{-1}[P(“B”|B)] - N^{-1}[P(“B”|G)]$ where N^{-1} is the inverse cumulative standard normal and d' is discriminability. For hit and false alarm rates of .80 and .20, $d' = 1.68$, which was the value used to derive the curve in Figure 1.³

In order to solve the cab problem by means of Equation 1, it is necessary to generalize the hit/false alarm ratio from the court’s signal detection experiment, where $P(G) = P(B)$, to the city conditions,

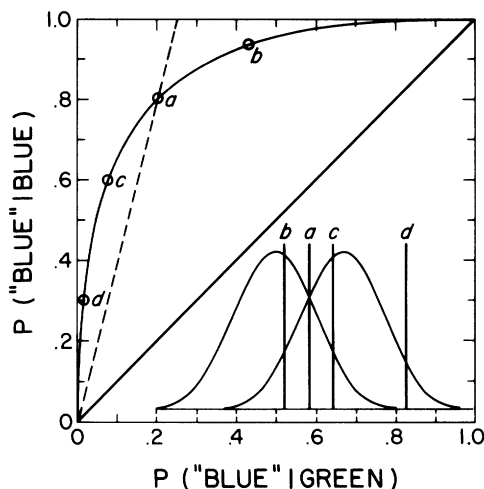


Figure 1. Signal detection analysis of the witness in the cab problem. Hit rate $[P(\text{"B"}|\text{B})]$ plotted against false alarm rate $[P(\text{"B"}|\text{G})]$. Point *a* shows values specified in the problem. Dashed line shows loci of points for which ratio of hit rate to false alarm rate is constant. The curve was derived from signal detection theory assuming each color of cab produces a normal distribution on a sensory continuum (inset). Point *b* is based on range-frequency theory. Point *c* is based on probability matching. Point *d* is close to a point that maximizes percentage of correct responses. The Bayesian solution to the cab problem changes depending on the theory of the witness. Assuming signal detection theory, virtually any answer from $P(\text{B})$ to 1.0 would be acceptable. The slope of the chord from the origin through the point is the likelihood ratio $[P(\text{"B"}|\text{B})/P(\text{"B"}|\text{G})]$. If point *d* is assumed, then $P(\text{B}|\text{"B"})$ would be .82, rather than .41.

where $P(\text{B})$ is only .15. To accomplish this generalization, one needs both signal detection theory (e.g., Equation 2) for the witness and a theory of judgment to predict how the witness will adjust his/her criterion (t in Equation 2) in response to changes in the signal probability (the proportion of Blue cabs). In other words, one needs to know the curve in Figure 1 and also how the point along the curve depends on $P(\text{B})$. Given only the curve, virtually any value from $P(\text{B})$ to 1.0 would be an acceptable solution to the cab problem. To pin down the solution requires a judgment theory, i.e., a theory of t .

Unfortunately, competing theories of judgment have received supporting evidence, though under different conditions. Several theories predict that when the signal frequency is increased, the response criterion will move to increase the probability of a correct response (Friedman, Carterette, Nakatani, & Ahumada, 1968; Green & Swets,

1966). Range-frequency theory (Parducci & Sandusky, 1970) predicts that when signal frequency is increased, the criterion will move the *opposite direction*, tending to equalize the two response frequencies.

Range-frequency solution

Range-frequency theory states that judges have a tendency to use response categories with equal frequency. Parducci and Sandusky (1970) have shown that when the witness is not given information about signal probabilities and is not given feedback on each trial, the response probabilities are largely independent of stimulus probabilities. If the witness reports half the cabs are “Blue” and half “Green” and if d' is constant (consistent with the results of Parducci & Sandusky, 1970), then the hit rate and false alarm rate will be .933 and .429, respectively, shown by point b in Figure 1. If such a witness reports “Blue,” then by Equation 1, $\Omega_1 = (.176)(.933/.429) = .383$, or $P(B|“B”) = .28$. Thus, assuming range-frequency theory, the “correct” answer of previous investigators (.41) is too large. The calculations are summarized in Table 1.

Probability matching solution

However, the witness in the street may be aware that only 15% of the cabs are Blue. The witness may have learned this from news reports or from feedback: An observer may judge a cab’s color from the distance and then see it up close, see cabs during the day, and/or hear consensus judgments of others concerning color. Should the witness operate according to range-frequency theory, the probability of a correct identification would be only .63. The witness could do better to remain at point a , in which case the witness would be correct 80% of the time. However, the witness could do still better by *never* saying “Blue,” in which case the witness would be right 85% of the time, since 85% of the cabs are Green. Friedman et al. (1968) found that when the observer is given feedback, the response criterion shifts in the opposite direction from that predicted by range-frequency theory.

With feedback therefore, it seems plausible that the witness might match response probabilities to the stimulus probabilities. When $P(“B”|B) = .599$ and $P(“B”|G) = .076$, $P(“B”)$ is approximately equal to $P(B)$. This strategy is shown by point c in Figure 1, and row c in Table 1. Assuming constant d' and assuming probability matching, Equation 1 implies that the normative posterior probability [$P(B|“B”)$] is .58, a normative solution actually greater than .5.

Optimal observer

Given feedback, it is not inconceivable that the witness could adjust the response criterion to maximize the probability of a correct iden-

Table 1. Summary of calculations

Limen value	Name	$^aP(\text{"B"} \text{B})$	$^aP(\text{"B"} \text{G})$	bL	$^cP(\text{"B"})$	$^dP(\text{Correct})$	$^eP(\text{B} \text{"B"})$
a	Conditional independence	.800	.200	4.00	.29	.80	.41
b	Equal response frequency	.933	.429	2.17	.50	.63	.28
c	Probability matching	.599	.076	7.88	.15	.88	.58
d	Optimal witness	.302	.012	25.17	.06	.89	.82

$^aP(\text{"B"}|\text{B})$ and $P(\text{"B"}|\text{G})$ are given by Equations 2, with $b-g = 1.68$.
 $^bL = P(\text{"B"}|\text{B})/P(\text{"B"}|\text{G}) = \text{slope of line from origin to point on ROC curve in Figure 1.}$
 $^cP(\text{"B"}) = P(\text{"B"}|\text{B}) \cdot P(\text{B}) + P(\text{"B"}|\text{G})P(\text{G}).$
 $^dP(\text{Correct}) = P(\text{"B"}|\text{B}) \cdot P(\text{B}) + (1-P(\text{"B"}|\text{G}))P(\text{G}).$
 $^eP(\text{B}) = .15.$
 $^fP(\text{B}|\text{"B"}) = \Omega_1/(1 + \Omega_1)$ where $\Omega_1 = \Omega_0L$; $\Omega_0 = .176$.

tification, or to minimize some index representing the costs of different types of errors, such as the expected cost (Green & Swets, 1966). The witness could conceivably perform so that $P(\text{"B"}|B) = .302$ and $P(\text{"B"}|G) = .012$, corresponding to point d in Figure 1, giving 89% correct identification, close to the maximal value. In this case, the witness would be highly diagnostic, and if the witness says "Blue," then the probability the cab was actually Blue, by Equation 1, is .82! Thus, it is not necessarily unreasonable for a judge to give .80 as the solution to the cab problem.

Now, suppose the base rate were .50; i.e., $P(B) = P(G)$. Under these conditions, the witness would likely remain at point a , which simultaneously equalizes the two response frequencies, matches stimulus and response frequencies, and maximizes the percentage of correct responses. Thus, the Bayesian solution would be .8, virtually the same as for the condition in which the base rate was .15 and the witness was assumed to be an ideal observer. The point is that Bayes' theorem does not necessarily imply an effect of the base rate unless a specific theory of the witness is assumed.

DISCUSSION

Theories of signal detection and judgment are required to generalize from the court's test of the witness to the performance in the street. In many signal detection experiments, data show that L varies as base rate is manipulated. When information concerning base rate is presented or the witness is given feedback, L increases as base rate decreases. When the cab problem is analyzed assuming d' is constant, the "correct" answer would be .28 if one believes the witness will respond "Blue" half the time; it would be .58 if the witness matched the probability of responding "Blue" to the proportion of Blue cabs in the street; it would be .82 if the witness will respond "Blue" less often, in response to the less frequent presence of Blue cabs and/or the possible high cost of a false alarm in the courtroom. In sum, the modal response of .8 by untrained subjects may be closer to the best normative solution than the value of .41 used in previous research. Therefore, there may be no evidence for a base rate fallacy.

The points of this paper may also apply to other problems employed in the literature on the "base rate fallacy" such as the light bulb problem or the engineer vs. lawyer problem (see Kahneman & Tversky, 1973). In the light bulb problem, the judge knows the probability of a defective bulb and the hit and false alarm rates of a mechanical light bulb tester. Suppose that the light bulb tester measures the current in each bulb for a given voltage. Suppose among bulbs that light, the distri-

bution of current readings is normal for both good bulbs and defective bulbs, but the means of the distributions differ. It seems reasonable that any profit-oriented factory would adjust the criterion for deciding “defective” as a function of the costs of false alarms (discarding a good bulb) or misses (replacing a bad one) and the probability of a defective bulb.

In the engineer vs. lawyer problem, the judge is told that there is a group of 30 engineers and 70 lawyers that were tested by psychologists, who then wrote thumbnail sketches. The judges’ task is to guess whether a sketch describes a lawyer or engineer. The psychologists gave information such as “he shows no interest in political and social issues . . . spends most of his free time on . . . mathematical puzzles.” Kahneman and Tversky (1973) found that the judged probability that the sketch described an engineer varied as a function of the proportion of engineers, but the effect of base rate was much less than predicted by Bayes’ theorem. However, their calculations assumed that the hit to false alarm ratio for the sketches would be independent of the base rate. The problem is that the probability of a given person *being described* as “interested in mathematical puzzles” would presumably depend on the proportion of engineers and lawyers. For example, if the entire population were engineers, a witness might describe Joe as “weak in mathematics.” However, if Joe were the only engineer in a population of lawyers, the same witness would be unlikely to select this description of Joe.

A recent paper by Schum (1981) gives a very thorough analysis of the evidential impact of testimony in Bayesian inference. Schum makes an important distinction between direct testimony concerning the hypothesis and testimony concerning a datum relevant to the hypothesis. Similarly the distinction between the datum and the report of the datum should be maintained. The report may convey more information than the datum itself (Schum, 1981). For example, it may be that interest in mathematical puzzles is not highly diagnostic of being an engineer or lawyer, but the *report* of interest in mathematical puzzles may be highly diagnostic of a rare engineer in a population consisting mostly of lawyers.

The theoretical analysis of this paper shows that many previous investigations of the “base rate fallacy” were based on a normative solution that rested on a shaky and perhaps unrealistic assumption concerning the independence of the hit/false alarm ratio from base rate. It should also be noted that the empirical evidence for neglect of base rate is also inconsistent. When the base rate is manipulated in a within-subjects design, it appears to be utilized by subjects (Birnbbaum & Mellers, Note 1; Fischhoff, Slovic, & Lichtenstein, 1979). Birnbbaum & Mellers (Note 1) found that the weight of the base rate even exceeds that of highly diagnostic sources. Subjects in within-

subjects designs attend to base rate, to source bias, and to source diagnosticity (see also Birnbaum & Stegner, 1979). Tversky and Kahneman (1982) also discuss between-subject results that were interpreted as evidence of subjects attending to base rate information.

CONCLUSION

The moral of this paper is not that subjects are necessarily accurate intuitive statisticians who understand signal detection theory, assume the witness is an ideal observer, and correctly solve the cab problem. Subjects in between-subject studies may indeed neglect the base rate and respond “.8” because this value seems to characterize the believability of the source, as argued by Tversky and Kahneman (1980, 1982), Bar-Hillel (1980), and others. However, this note does call attention to the fact that the normative solution to the cab problem requires the assumption of a theory of the witness, whether by the subject or the experimenter.

Notes

Thanks are due to Barbara Mellers, David Noreen, Amos Tversky, and Robert Sorkin for comments on an earlier draft.

Requests for offprints should be addressed to Michael H. Birnbaum, Department of Psychology, University of Illinois, 603 East Daniel St., Champaign, IL 61820. Received for publication September 11, 1981; revision received April 28, 1982.

1. Other aspects of the cab problem as stated above are also unclear. For example, the statement that 85% of the cabs in the city are Blue does not necessarily imply that 85% of the cabs involved in hit and run accidents at night are Blue. The decision-maker is presumably supposed to assume that both cab companies operate at night, have drivers with comparable skill, etc., so that the base rate can be interpreted as relevant to the problem. Tversky and Kahneman (1980) discuss effects of variation in the phrasing of the base rate information and contend that if the base rate information seems “causal” then it will be utilized. However, their experiments have not yet unconfounded the so-called “causal” version of the problem from the noncausal but clearly relevant version. Tversky and Kahneman (1982) used a modified form of the cab problem in which the court test is described as being done “under the same circumstances” as existed on the night of the accident, but the base rate and payoff conditions of the test are not explicitly specified.

2. Even mechanical devices such as a light bulb tester may be amenable to the same analysis. Suppose the threshold for the light bulb tester is adjustable, and it is assumed that normal and defective bulbs produce normal distributions of resistance values with different means, the same analysis would apply as in Figure 1. Thus, the ratio of hit rate to false alarm rate will vary for some mechanical devices as well as for human witnesses.

3. Other signal detection theories assuming unequal variance or using other functions, such as the logistic instead of the normal, would lead to similar conclusions for the present analysis. Empirical curves in signal detection studies with varying base rates typically have a shape similar to the curve in Figure 1 and do not resemble the dashed line in Figure 1 (e.g., Green & Swets, 1966, p. 88, p. 95). However, even very different theories, such as high threshold theory or a linear theory with slope = 1 imply that the ratio of $P("B"|B)/P("B"|G)$ will change as a function of $P(B)$.

Reference note

1. Birnbaum, M. H., & Mellers, B. A. *Bayesian inference: Combining base rates with opinions of sources who vary in credibility*. Paper presented at the Midwestern Psychological Association, 1978.

References

- Bar-Hillel, M. The base rate fallacy in probability judgments. *Acta Psychologica*, 1980, 44, 211-233.
- Birnbaum, M. H., & Mellers, B. A. Bayesian inference: Combining base rates with opinions of sources who vary in credibility.
- Birnbaum, M. H., & Stegner, S. E. Source credibility: Bias, expertise, and the judge's point of view. *Journal of Personality and Social Psychology*, 1979, 37, 48-74.
- Friedman, M. P., Carterette, E. C., Nakatani, L., & Ahumada, A. Feedback and frequency variables in signal detection. *Perception & Psychophysics*, 1968, 3, 5-10.
- Fischhoff, B., Slovic, P., & Lichtenstein, S. Subjective sensitivity analysis. *Organizational Behavior and Human Performance*, 1979, 23, 339-359.
- Green, D. M., & Swets, J. A. *Signal detection theory and psychophysics*. New York: Wiley, 1966.
- Hammerton, M. A. A case of radical probability estimation. *Journal of Experimental Psychology*, 1973, 101, 252-254.
- Kahneman, D., & Tversky, A. On the psychology of prediction. *Psychological Review*, 1973, 80, 237-351.
- Lyon, D., & Slovic, P. Dominance of accuracy information and neglect of base rates in probability estimation. *Acta Psychologica*, 1976, 40, 287-289.
- Parducci, A., & Sandusky, A. Limits on the applicability of signal detection theory. *Perception & Psychophysics*, 1970, 7, 63-64.
- Schum, D. Sorting out the effects of witness sensitivity and response-criterion placement upon the influential value of testimonial evidence. *Organizational Behavior and Human Performance*, 1981, 27, 153-196.
- Tversky, A., & Kahneman, D. Judgment under uncertainty: Heuristics and biases. *Science*, 1974, 185, 1124-1131.
- Tversky, A., & Kahneman, D. Causal schemata in judgments under uncertainty. In M. Fishbein (Ed.), *Progress in Social Psychology*. Hillsdale, N.J.: Erlbaum, 1980.
- Tversky, A., & Kahneman, D. Evidential impact of base rates. In D. Kahneman, P. Slovic, & A. Tversky (Eds.), *Judgment under uncertainty: Heuristics and biases*. Cambridge University Press, 1982.