Bloom, D. M. (1973), "A Birthday Problem," *American Mathematical Monthly*, 80, 1141–1142.
Diaconis, P., and Mosteller, F. (1989), "Methods for Studying Coincidences," *Journal of the American Statistical Association*, 84, 853–861.
Feller, W. (1968), *An Introduction to Probability Theory and Its Applications* (3rd ed.), New York: John Wiley.
Gail, M. H., Weiss, G. H., Mantel, N., and O'Brien, S. J. (1979), "A Solution to the Generalized Birthday Problem With Application to Allozyme Screening for Cell Culture Contamination," *Journal of Applied Probability*, 16, 242–251.
Mathis, F. H. (1991), "A Generalized Birthday Problem," *SIAM Review*, 33, 265–270.
Mosteller, F. (1962), "Understanding the Birthday Problem," *The Mathematics Teacher*, 55, 322–325.
Munford, A. G. (1977), "A Note on the Uniformity Assumption in the Birthday Problem," *The American Statistician*, 31, 119.
Naus, J. I. (1968), "An Extension of the Birthday Problem," *The American Statistician*, 22, 27–29.
Rust, P. F. (1976), "The Effect of Leap Years and Seasonal Trends on the Birthday Problem," *The American Statistician*, 30, 197–198.
Schwarz, W. (1988), "Approximating the Birthday Problem," *The American Statistician*, 42, 195–196.
Shaughnessy, J. M. (1977), "Misconceptions of Probability: An Experiment With a Small-Group, Activity-Based Model Building Approach to Introductory Probability at the College Level," *Educational Studies in Mathematics*, 8, 295–316.
United States Department of Health and Human Services (1989), *Vital Statistics of the United States, 1987. Vol. 1: Natality*, Washington, DC: U.S. Government Printing Office, p. 46.

Bayesian Resolution of the "Exchange Paradox"

RONALD CHRISTENSEN and JESSICA UTTS*

In this article we present a paradox that can be used to illustrate Bayesian principles in the classroom. The paradox is also resolved using a frequentist argument and illustrates how the misapplication of a symmetry argument causes problems.

KEY WORDS: Newcomb's paradox; Noninformative prior; Symmetry arguments.

One of the arguments in favor of using Bayesian methods is that prior assumptions are made explicit rather than being incorporated implicitly into the solution of a problem. In this article we discuss an interesting paradox and show how explicit use of prior assumptions leads to its resolution. The paradox is also resolved using a frequentist argument and illustrates the pitfalls of using symmetry incorrectly. It shows that careful thinking is important in solving probability problems and elucidates the difference between the Bayesian and frequentist approaches.

1. THE PARADOX

This paradox apparently dates back to Kraitchik (1953). It is called the "Wallet Game" by Gardner (1982) and the "Exchange Paradox" by Zabell (1988). None of these authors provides a resolution to the paradox.

A swami puts m dollars in one envelope, and $2m$ dollars in another. He hands one envelope to you and one to your opponent, so that the probability is 1/2 that you get either envelope. You open your envelope and find x dollars. Let Y be the amount in your opponent's envelope. You reason that since the envelopes were handed out with equal probability, $Y = x/2$ or $Y = 2x$, each with probability $\frac{1}{2}$. Thus, your expected winnings from a trade are $(\frac{1}{2})\,(x/2 + 2x) = 5x/4$ which is obviously greater than the x dollars you currently possess. With a gleam in your eye, you offer to trade envelopes with your opponent. Since she has made the same calculation, she readily agrees.

The paradox of this problem is that the rule indicating that one should always trade is intuitively unreasonable while the method of arriving at the rule seems very reasonable. The actual structure of the problem is hidden from the casual observer by the simplicity and intuitiveness of the arguments.

2. A BAYESIAN RESOLUTION

The conclusion that trading envelopes is always optimal is based on the assumption that there is no information obtained by observing the contents of the envelope. From a Bayesian perspective, the key to a successful analysis is in recognizing the potential information to be gained from the observation. The paradox can be arrived at as the consequence of using an improper "noninformative" prior, and thus reaffirms the dangers of blindly using such priors. This example also illustrates an idea often associated with complete class theorems, i.e., reasonable decision rules are Bayes's rules and one ignores the implicit prior at one's peril.

The parameter in this problem is m, the amount of money placed in the first envelope. A Bayesian approach uses prior information to model beliefs about the value of m. Some prior information for the value of m certainly exists. For example, m is highly unlikely to be greater than the gross national product. (When this problem was presented to a class of graduate students with the instructor placing the money in the en-

*Ronald Christensen is Associate Professor of Statistics, Department of Mathematics and Statistics, University of New Mexico, Albuquerque, NM 87131. Jessica Utts is Associate Professor, Division of Statistics, University of California, Davis, CA 95616. The authors would like to thank two anonymous referees and an associate editor for useful comments. The second author would like to thank Sandy Zabell for describing the envelope problem and Edwardo Engel for a helpful discussion.

velope, the graduate students had a high prior probability that the value of m would be near zero.)

Let M be the subjectively random amount of money placed in the first envelope and let $g(m)$ be the prior density for M. Let X be the random amount of money in your envelope. The sampling distribution of X is $\Pr(X = m|M = m) = \Pr(X = 2m|M = m) = \frac{1}{2}$. Since X equals either M or $2M$, on observing $X = x$, M can take on only two values, x and $x/2$. Applying Bayes's Theorem, we find that

$$\Pr(M = x|X = x)$$
$$= \frac{\Pr(X = x|M = x)g(x)}{\Pr(X = x|M = x)g(x) + \Pr(X = x|M = x/2)g(x/2)}$$
$$= \frac{g(x)}{g(x) + g(x/2)}$$

and

$$\Pr(M = x/2|X = x) = \frac{g(x/2)}{g(x) + g(x/2)}.$$

The Bayesian can now compute his expected winnings from the two actions. If he keeps the envelope he has, he wins x dollars. If he trades envelopes, he wins $x/2$ if he currently has the envelope with $2M$ dollars, i.e., if $M = x/2$ and he wins $2x$ if he currently has the envelope with M dollars, i.e., $M = x$. His expected winnings from a trade are

$$E(W|\text{Trade}) = E(Y|X = x)$$
$$= \frac{g(x/2)}{g(x) + g(x/2)} \cdot \frac{x}{2} + \frac{g(x)}{g(x) + g(x/2)} \cdot 2x.$$

It is easily seen that when $g(x/2) = 2g(x)$, $E(W|\text{Trade}) = x$. Therefore, if $g(x/2) > 2g(x)$ it is optimal to keep the envelope and if $g(x/2) < 2g(x)$ it is optimal to trade envelopes. For example, if your prior distribution on M is exponential λ, so that $g(m) = e^{-\lambda m}$, then it is easily seen that it is optimal to keep your envelope if $x > 2\log(2)/\lambda$.

The intuitive value of the expected winnings when trading envelopes was shown to be $5x/4$. This value can be obtained by assuming that $g(x)/[g(x) + g(x/2)] = \frac{1}{2}$ for all x. In particular, this implies that $g(x) = g(x/2)$ for all x, i.e., $g(x)$ is a constant function. In other words, the intuitive expected winnings assumes an improper "noninformative" uniform density on $[0, \infty)$. It is of interest to note that the improper noninformative prior for this problem gives a truly noninformative (maximum entropy) posterior distribution.

Finally, before demonstrating some frequentist calculations relating to this problem, we would like to suggest that the real paradox here is not that the noninformative rule is to always trade, but rather that people want noninformative rules to give informative conclusions. If you really have no idea whether M is x or $x/2$, it is better to trade. When $M = x$, you win twice as much by trading as you lose when $M = x/2$. The problem is that if you get an X that you think is big, you (quite correctly) do not want to trade but a noninformative rule tells you to trade. The problem here is not the noninformative rule, it is trying to make a decision ignoring prior information.

3. SOME FREQUENTIST CALCULATIONS

The correct thinking from a frequentist perspective is easily seen by writing the sample space for (X, Y) and noting that the two sample points are equally likely:

$$S = \{(x,y)\colon (m,2m), (2m,m)\}.$$

The paradox is elicited by assuming that the conditional distribution of Y given $X = x$ is

$$\Pr(Y = y|X = x) = \frac{1}{2} \quad y = \frac{x}{2}, 2x$$
$$= 0 \quad \text{otherwise.}$$

In order to be correct, this would have to be correct for all x. In fact, there is no value of x for which it is true. The correct conditional distribution should take the value of X into account:

$$\text{For } X = m \ \Pr(Y = y|X = x) = 1 \quad y = 2x$$
$$= 0 \quad \text{otherwise.}$$

$$\text{For } X = 2m \ \Pr(Y = y|X = x) = 1 \quad y = \frac{x}{2}$$
$$= 0 \quad \text{otherwise.}$$

We can then write the conditional expectation as

$$E(Y|X = x) = 2x\, I_{[X=m]} + \frac{x}{2} I_{[X=2m]}. \qquad (1)$$

This is distinct from the intuitive result that $E(Y|X = x) = 5x/4$. Since Equation (1) involves unknown quantities (the values of the indicator functions), the frequentist has no basis for evaluating the worth of a trade.

Taking the expectation over X gives the expected winnings for the strategy of always trading envelopes:

$$E(Y) = E[E(Y|X{=}x] = 2m\left(\frac{1}{2}\right) + \frac{2m}{2}\left(\frac{1}{2}\right)$$
$$= m + \frac{m}{2} = \frac{3m}{2}.$$

This is obviously the same as the expected winnings for the strategy of always keeping your original envelope.

Of course the frequentist might also realize that different values of x call for different strategies and could design a rule to take that into account. For example: Choose a value c that you think is likely to be between m and $2m$. Suppose $X = x$ is observed.

If $x \le c$, trade envelopes.

If $x > c$, keep x. (2)

If you have indeed chosen c such that $m \le c < 2m$, you will trade when $X = m$ and not trade when $X = 2m$, so your expected winnings will be $2m$. If $c < m$, you will never trade, and if $c \ge 2m$ you will always trade. In both cases, as shown above, your expected winnings will be $3m/2$. Thus, the strategy in (2) is at least as good as always trading or never trading and can be better depending on how your choice of c relates to m.

4. CONCLUSION

We believe the Bayesian has an advantage in this game. As mentioned above, if the observed value of X is very large, the desire to trade should be small. At the other extreme, if there is a single dollar (or penny) in the envelope, the smart decision would be to trade. The Bayesian solution allows these considerations to be explicitly taken into account, but the frequentist solution does not.

Another well-known paradox, called Newcomb's Paradox, (e.g., Skyrms 1984, p. 66) is resolved in a similar manner in Christensen and Utts (1991).

[*Received January 1991. Revised January 1992.*]

REFERENCES

Christensen, R., and Utts, J. (1991), "Bayesian Resolution of Classical Paradoxes: Two Examples," Technical Report 220, University of California, Davis, Division of Statistics.

Gardner, M. (1982), *Gotcha: Paradoxes to Puzzle and Delight*, New York: W. H. Freeman.

Kraitchik, M. (1953), *Mathematical Recreations* (2nd ed.), New York: Dover.

Skyrms, B. (1984), *Pragmatics and Empiricism*, New Haven: Yale University Press.

Zabell, S. L. (1988), "Symmetry and Its Discontent," in *Causation, Chance, and Credence*, eds. B. Skyrms and W. L. Harper, Dordrecht: Kluwer Academic Publishers.

An Invariance Property of the Poisson Process

MARCEL F. NEUTS*

The points of a homogeneous Poisson process that fall in each of a string of consecutive intervals are uniformly and independently redistributed over these intervals. It is shown that the resulting point process is again a homogeneous Poisson process. The two processes are stochastically dependent and their superposition is not even stationary. The proofs use only elementary properties and yield useful examples for educational use.

KEY WORDS: Dependence; Stationarity.

This article deals with an invariance property of the homogeneous Poisson process found while investigating a construction for a more general class of random point processes. That construction, which is called "local poissonification," is motivated by the study of "burstiness" of point processes. It is discussed in greater detail in Neuts, Liu, and Narayana (1992). The educational appeal of the proof of our main result lies in its use of geometric insight combined with simple uses of all the classical properties of the Poisson process discussed in standard texts (Karlin and Taylor 1975; Parzen 1962). In addition, the article provides a number of elementary examples of point processes with "unusual" properties.

We consider a homogeneous Poisson process of rate λ on the entire real line. Without loss of generality, we may choose λ to be one. We further consider a monotone sequence of real constants a_k having $-\infty$ and ∞ as its only accumulation points. Again without loss of generality, the $\{a_k\}$ may be assumed to be distinct. Let N_k be the number of points of the Poisson process in the interval $(a_k, a_{k+1}]$, then since the Poisson process has stationary independent increments, the random variables $\{N_k\}$ are independent and have Poisson densities with mean $a_{k+1} - a_k$. Moreover, a classical property of the Poisson process states that, conditionally on the event $N_k = n$, with $n \geq 1$, the locations of the n Poisson events in the interval $(a_k, a_{k+1}]$, are distributed as the order statistics of n independent random variables with the uniform density on $(a_k, a_{k+1}]$.

We now perform the following construction (see Fig. 1). Let us refer to the points of the original Poisson process as the "blue" points. The values of the random variables N_k are noted and for any interval $(a_k, a_{k+1}]$ for which the corresponding N_k is positive, an *equal number* of "red" points are placed independently and at random in that interval.

Theorem 1. The point process of red points is a homogenous Poisson process of unit rate.

Proof. There are a number of ways of proving this statement, but it turns out that with appropriate geometric insight, a direct verification of the axioms defining the Poisson process is easiest. Let $N_r(\cdot)$ and $N_b(\cdot)$ stand for the counting processes of the red and the blue points. We need to show that the red points form a *stationary* point process with *stationary independent increments* and that

$$P\{N_r(t + h) - N_r(t) \geq 2\} = o(h), \text{ as } h \to 0.$$

Let us consider the sequence of squares in the (x_1, x_2) plane, obtained by completing the squares with vertices at the points (a_k, a_k) and whose diagonals of slope one

*Marcel F. Neuts is Professor, Department of Systems and Industrial Engineering, University of Arizona, Tucson, AZ 85721. This research was supported in part by Grants DDM-8915235 from the National Science Foundation and AFOSR-88-0076 from the Air Force Office of Scientific Research.