UNIFORM DISTRIBUTIONS ON THE INTEGERS: A CONNECTION TO THE BERNOUILLI RANDOM WALK

JOSEPH B. KADANE AND JIASHUN JIN,
IN HONOR OF ARNOLD ZELLNER

Abstract. Associate to each subset of the integers its almost sure limiting relative frequency under the Bernouilli random walk, if it has one. The resulting probability space is purely finitely additive, and uniform in the sense of residue classes and shift-invariance. However, it is not uniform in the sense of limiting relative frequency.

Keywords: almost sure limiting relative frequency, finitely additive probability, residue class, relative frequency, uniform distribution on the integers

1. Finitely Additive Probability

Zellner (1971) was particularly interested in priors that are not proper (see Example 2.2 on p. 20, for instance) following Jeffreys (1961). There are two approaches in the literature to understand these mathematically. The first, explored by Renyi (1970) and Hartigan (1983), work directly with integrating measures that are σ-finite, but possibly integrate to infinity. However, these lack probabalistic intuition, and have been criticized as falling prey to the marginalization paradox (Dawid et al., 1973).

The second approach is to give up countable additivity and allow probabilities that are only finitely additive. In this approach, the usual axiom,

$$\text{If } A_1, A_2, \ldots, A_n \text{ are disjoint, then } P(\cup_{i=1}^{\infty} A_i) = \sum_{i=1}^{\infty} P(A_i) \tag{1}$$

is replaced by the weaker assumption

$$\text{If } A_1, A_2, \ldots, A_n \text{ are disjoint, then } P(\cup_{i=1}^{n} A_i) = \sum_{i=1}^{n} P(A_i). \tag{2}$$

DeFinetti (1970, 1974) championed this approach, and showed that both a betting argument and a method through Brier scoring led to (2), but not (1). Dubins and Savage (1965) found finite additivity mathematically convenient in their study of the gambling game "red and black."

It turns out that even on a simple infinite set like the integers, there is more than one way to express the idea of a uniform distribution. In particular, Schirokauer and Kadane (2007), following Kadane and O'Hagan (1995) study three senses of uniformity on the line, namely

(1) residue classes: Each residue class $\mod k$ has probability $1/k$.
(2) shift-invariance: If A has probability p, then $A+1$ and $A-1$ (equal, respectively to $\{x \mid x-1 \epsilon A\}$ and $\{x \mid x+1 \epsilon A\}$) also have probability p.
(3) relative frequency: If $\lim_{n\to\infty} \frac{\{x|x\epsilon A, -n\leq A\leq n\}}{2n+1}$ exists, and equals p, then $P(A)=p$.

The results of Shirokauer and Kadane show that these are in order of strictness (thus relative frequency is the most restrictive, and residue classes the least restrictive) and, by example, that the inclusions are strict. The examples offered to show strict inclusion are number-theoretic in character, and hence unfamiliar to statisticians, probabalists and econometricians. The purpose of this note is to give an example that satisfies (1) and (2) but not (3) above, and is much more familiar.

2. Almost sure limiting relative frequency

Consider an arbitrary discrete stochastic process w on an arbitrary space Ω. Suppose $w_i \epsilon \Omega$ indexes a countably infinite sequence of w's. Then a set A is said to have almost sure limiting relative frequency (aslrf) p just in case $\exists\, N_A$ such that $P\{N_A\}=0$, and for all $\omega \notin N_A$,

$$\lim_{n\to\infty} \sum_{i=1}^{n} I_A(\omega_i)/n \tag{3}$$

exists and equals p.

Proposition 1. *Suppose A and B are disjoint sets with aslrf a and b respectively. Then $A\cup B$ has aslrf $a+b$.*

Proof: Since A and B have aslrf's, $\exists\, N_A$ and N_B such that $P\{N_A\}=P\{N_B\}=0$, satisfying (3). If we let $N_{A\cup B}=N_A\cup N_B$, then $P\{N_{A\cup B}\}=0$ and $\forall\, \omega \notin N_{A\cup B}$,

$$\begin{aligned}
&\lim_{n\to\infty} \sum_{i=1}^{n} I_{A\cup B}(\omega_i)/n \\
&= \lim_{n\to\infty} \sum_{i=1}^{n} (I_A(\omega_i)+I_B(\omega_i))/n && \text{[A and B are disjoint]} \\
&= \lim_{n\to\infty} \sum_{i=1}^{n} I_A(\omega_i)/n + \lim_{n\to\infty} \sum_{i=1}^{n} I_B(\omega_i)/n && \text{[all summands are non-negative]} \\
&= a+b && \text{[by definition]}
\end{aligned}$$

□

By induction, the same result applies to a finite number of disjoint sets. This proposition is not special to random walks, but rather applies to every stochastic process.

Let $\mathcal{A}$ be the collection of sets A that have a limiting relative frequency, and assign the limit as the probability measure of set A. Such an assignment satisfies the first two requirements of a probability measure: $P(\Omega)-1$ and $P(A)\geq$ for all $A\epsilon\mathcal{A}$. By Proposition 1 and its inductive extension, it satisfies finite addivity as well.

3. Bernouilli Random Walks

Consider now the special stochastic process on the integers defined as follows: Let $X_i, i = 1, \ldots$ be identically distributed independent random variables, with

$$P\{X_i = 1\} = p \text{ and } P\{X_1 = -1\} = q = 1 - p. \tag{4}$$

Then $S_n = \sum_{i=1}^{n} x_i$ is a Bernouilli random walk.

It is well-known that under a Bernouilli random walk each integer has an aslrf of zero (see e.g. Durrett (1996). However, the disjoint countable union of all integers has aslrf one. Hence this probability space is not countably additive.

4. Residue Classes

A residue class $r \mod k$ consists of all integers s such that $s = r + kn$, where $0 \leq r < k$ and n is an integer, or put differently, it is the set of all integers spaced k apart that includes r. For each k, there are k of these sets, and they are disjoint. The next proposition shows that under a Bernouilli random walk, each residue class $\mod k$ has aslrf $1/k$.

Proposition 2. *The almost sure limiting relative frequencies of the residue class mod k of the general (symmetric or asymmetric) Bernoulli random walk on the line exist and are $1/k$.*

Proof: For $k = 1$ there is nothing to prove. For $k = 2$, the random walk is alternatively even and odd, and hence the residue classes mod 2 have limiting relative frequencies of $1/2$. For $k > 2$, we separate the cases of k even and odd.

First, suppose that k is odd, and consider the Markov Chain where the states are the residue classes mod k. This Markov chain has transition probability matrix

$$P = \begin{pmatrix} 0 & q & 0 & . & . & . & . & . & . & p \\ p & 0 & q & 0 & . & 0 & . & . & . & 0 \\ 0 & p & 0 & q & . & . & . & . & . & . \\ & & & & & & & & & \\ & & & & & & & p & 0 & q \\ q & 0 & & & & & & 0 & p & 0 \end{pmatrix}.$$

This chain is irreducible, and, because k is odd, aperiodic. Therefore, it has a unique left eigen-vector associated with the eigenvalue 1. Because the column sums are identically 1, this eigenvector has the same probability in each co-ordinate. Since there are k of them, each has probability $1/k$. Since the stationary distribution is this eigenvector, the residue classes mod k almost surely have limiting relative frequency $1/k$.

Now we consider the case of k even. Then the chain described above is periodic with period 2. Consider then the Markov Chain on the even numbered residue classes $0, 2, \ldots, (k/2) - 1$ generated by two steps of the random walk. This chain has transition probability matrix

$$P' = \begin{pmatrix} 2pq & q^2 & 0 & . & . & . & . & 0 & p^2 \\ p^2 & 2pq & q^2 & 0 & & & & 0 & 0 \\ 0 & p^2 & 2pq & q^2 & & & & & \\ q^2 & & & & & & & p^2 & 2pq \end{pmatrix}.$$

This chain is irreducible and aperiodic. Furthermore, its column sums are again 1. Thus by the argument above, its stationary distribution is again uniform, here on $k/2$ states, so each has limiting relative frequency, in the two-step process, of $2/k$. In the one-step process the odd numbered steps cannot result in an even number, so the limiting relative frequency of each of the even numbered residue classes mod k is $(1/2)(2/k) = 1/k$.

The analysis for the odd numbered cylinder sets $1, 3, \ldots, (k/2) - 1$ is exactly the same, with the same two-step matrix P'. Hence each of them has almost sure limiting relative frequency $1/k$. $\square$

5. Shift invariance

Suppose a set A has aslrf a. The question to be addressed is whether $S^{\mathbf{y}}(A) = \{\mathbf{x} + \mathbf{y} \mid \mathbf{x} \in A\}$ also has aslrf a.

Let x be an integer, and let $P^{\mathbf{x}}$ be the law of the random walk starting at $\mathbf{x}$. Consider the function

$$f_A(\mathbf{x}) = P^{\mathbf{x}}(\overline{\lim}_{n\to\infty} \sum_{k=0}^{n} I_A(X_k)/n = \underline{\lim}_{n\to\infty} \sum_{k=0}^{n} I_A(X_k)/n = a).$$

When A has aslrf a, $f_A(\mathbf{0}) = 1$. The function $f_A(\mathbf{x})$ satisfies the following equation:

$$f_A(\mathbf{x}) = \sum_{y} f_A(\mathbf{x} + \mathbf{y}) p^t(\mathbf{y}) \text{ for each finite } t \tag{5}$$

where $p^t(\mathbf{y})$ is the probability that the chain moves from $\mathbf{x}$ to $\mathbf{y}$ in t steps. This equation holds because the first t steps of the random walk are irrelevant to the aslrf. But this is the functional equation of the Choquet-Deny Theorem (see, for example, Rao and Shanbhag (1991)), and warrants the conclusion that $f_A(\mathbf{x}) = f_A(\mathbf{x}+\mathbf{y})$ for all $\mathbf{y}$ visited by the random walk starting from $\mathbf{x}$ with positive probability. Recalling $f_A(\mathbf{0}) = 1$, we have $f_A(\mathbf{x}) = 1$ for all $\mathbf{x}$ visited by the random walk starting at $\mathbf{0}$.

In summary, we have shown:

Proposition 3. *For each $\mathbf{x}$ visited with positive probability by starting at $x_0 = \mathbf{0}$, if A has aslrf a, then so does $S^{\mathbf{x}}(A)$.*

6. Relative frequency

The set B of positive integers has limiting relative frequency $1/2$. When $p > 1/2$ in (4), the random walk has positive drift, and B has aslrf equal to 1. Similarly, if $p < 1/2$, the drift is negative, and B has aslrf equal to 0. Hence, this random walk violates the relative frequency sense of uniformity.

7. Conclusion and Open Questions

Bernouilli random walks offer a familiar example of a process that satisfies the residue class and shift-invariance senses of uniformity of a finitely-additive probability, but not the relative-frequency sense. We had hoped that more general random walks on the line, or in the space of d-tuples of integers might provide an example that satisfied the residue-class criterion, but not shift-invariance However, this turned out not to be the case. Consequently, the discovery of a simple such example is an open question. More generally, improving our understanding the mathematics of even the simplest cases of uniform probability distributions on unbounded spaces is just in its infancy.

Acknowledgement

The authors thank Ramon van Handel for pointing us to the importance of the Choquet-Deny theorem in connection with shift invariance.

References

Dawid, P., Stone, M., and Zidek, J. (1973). "Marginalization Paradoxes in Bayesian and Structural Inference." *J. Roy. Stat. Soc.*, B, 35, 189–233.

DeFinetti, B. (1970). *Theory of Probability*. J. Wiley & Sons.

— (1974). *Theory of Probability*. J. Wiley & Sons.

Dubins, L. and Savage, L. (1965). *How to Gamble if You Must: Inequalities for Stochastic Processes*. McGraw-Hill.

Durrett, R. (1996). *Probability: Theory and Examples*. J. Wiley & Sons.

Hartigan, J. (1983). *Bayes Theory*. Springer-Verlag.

Jeffreys, H. (1961). *Theory of Probability*. Oxford University Press.

Kadane, J. B. and O'Hagan, A. (1995). "Using Finitely Additive Probability: Uniform Distributions on the Natural Numbers." *Journal of the American Statistical Association*, 626–631.

Rao, C. R. and Shanbhag, D. N. (1991). "An Elementary Proof for an Extended Version of the Choquet-Deny Theorem." *Journal of Multivariate Analysis*, 141–148.

Renyi, A. (1970). *Foundations of Probability*. Holden-Day.

Schirokauer, O. and Kadane, J. B. (2007). "Uniform Distributions on the Natural Numbers." *Journal of Theoretical Probability*, 20, 429–441.

Zellner, A. (1971). *An Introduction to Bayesian Inference in Econometrics*. J. Wiley & Sons.