A Bayesian Eigenvalue Game

Kent E. Morrison

Department of Mathematics
California Polytechnic State University
San Luis Obispo, CA 93407
kmorriso@calpoly.edu

June 22, 1999

Imagine a gambling device that randomly generates a 10 by 10 matrix whose square is the identity with entries in the field $\mathbf{F}_3$. There are 1.55967×10^{24} such matrices, a small fraction of the more than 10^{47} matrices of that size but still a sizeable number to draw from. Each of these square roots of the identity is diagonalizable with eigenvalues ± 1. The device computes the eigenvalues and drops 10 tokens into a closed box. For each eigenvalue equal to 1 it puts into the box a red token, and for each eigenvalue equal to -1 it puts into the box a blue token. The tokens are in the box but no one knows what they are. At this point the "croupier" takes the box in his hands and shakes it several times. He then begins to remove the tokens, one at a time, without revealing those that remain in the box. The first token removed is red. The second is also red, as is the third, the fourth, and the fifth. Now the croupier asks you, the player: "Would you care to place a bet at even odds that there is another red token in the box?"

Should you place such a bet?

We shall return to this question after analyzing a similar but simpler game. In this case the gambling device flips a coin ten times internally and places into the closed box a red token for each heads and a blue token for each tails. Again suppose that five tokens are removed and that they are

1991 Mathematics Subject Classification. Primary 60C05; Secondary 15A18, 15A21
Key words and phrases. Eigenvalues, matrices over finite fields, Bayes's Theorem

all red. Would you care to bet that there is still another red token in the box? At even odds? You bet! In fact, the probability that there is at least one more red token is 31/32. The computation is an exercise with Bayes's Theorem of elementary probability. Here is how it goes.

Let H_k, $0 \leq k \leq 10$ be the event that k heads occur in the 10 coin tosses, and hence that k red tokens are in the closed box. Let R_5 be the event that five tokens are removed from the box and that they are all red. Then the conditional probability $\mathrm{P}(H_5|R_5)$ is the probability that no red token remains in the box when you are asked whether you want to bet.

$$\begin{aligned} \mathrm{P}(H_5|R_5) &= \frac{\mathrm{P}(H_5 \cap R_5)}{\mathrm{P}(R_5)} \\ &= \frac{\mathrm{P}(H_5 \cap R_5)}{\sum_{k=0}^{10} \mathrm{P}(H_k \cap R_5)} \end{aligned}$$

since the H_k are disjoint events partitioning the sample space. Then

$$\mathrm{P}(H_5|R_5) = \frac{\mathrm{P}(R_5|H_5)\mathrm{P}(H_5)}{\sum_{k=0}^{10} \mathrm{P}(R_5|H_k)\mathrm{P}(H_k)} \tag{1}$$

This last equation is the statement of Bayes's Theorem. At this point we invite the reader to pause for a few minutes to evaluate the right side of the equation in order to see that $\mathrm{P}(H_5|R_5) = 1/32$.

Back to the eigenvalue game: let E_k, $0 \leq k \leq 10$ be the event that the random matrix has eigenvalue 1 with multiplicity k. We need to compute the right side of

$$\mathrm{P}(E_5|R_5) = \frac{\mathrm{P}(R_5|E_5)\mathrm{P}(E_5)}{\sum_{k=0}^{10} \mathrm{P}(R_5|E_k)\mathrm{P}(E_k)} \tag{2}$$

The conditional probabilities are the same as those for the coins: $\mathrm{P}(R_5|E_k) = \mathrm{P}(R_5|H_k)$. For $k \geq 5$ these probabilities are equal to $\frac{k(k-1)(k-2)(k-3)(k-4)}{10 \cdot 9 \cdot 8 \cdot 7 \cdot 6}$, and for $k \leq 4$ they are 0.

Claim Let A be an $n \times n$ matrix satisfying $A^2 = I$. Then A is similar to a diagonal matrix having diagonal entries 1 or -1.

Proof If λ is an eigenvalue of A, then $\lambda = \pm 1$. Now, let E_1 and E_{-1} be the corresponding eigenspaces. We know $E_1 \cap E_{-1} = \{0\}$ because eigenvectors of distinct eigenvalues are linearly

independent. Suppose $E_1 \oplus E_{-1} \neq V$. Then there exists $x \in V$ and not in $E_1 \oplus E_{-1}$. Let $y = Ax$. Since x is not an eigenvector of A, x and y are linearly independent. Furthermore, y is not in $E_1 \oplus E_{-1}$, for otherwise $x = Ay$ would be in $E_1 \oplus E_{-1}$. Thus, x and y are linearly independent and not in $E_1 \oplus E_{-1}$. Now, $A(x + y) = y + x$, and so $x + y \in E_1$, which is impossible. It must be the case that $E_1 \oplus E_{-1} = V$, and hence A is diagonalizable. □

Let D_k be the diagonal matrix with k repetitions of 1 and $n-k$ repetitions of -1. The matrices in the similarity class of D_k have the form PD_kP^{-1} where P is an element of the group of invertible matrices $GL_n(q)$. Hence the cardinality of the similarity class is the quotient of $|GL_n(q)|$ by the order of the subgroup of P such that $PD_kP^{-1} = D_k$. This subgroup consists of matrices in block form

$$\begin{bmatrix} A & 0 \\ 0 & B \end{bmatrix}$$

where $A \in \mathrm{GL}_k(q)$ and $B \in \mathrm{GL}_{n-k}(q)$. Thus, the number of matrices similar to D_k is

$$\frac{|GL_n(q)|}{|GL_k(q)||GL_{n-k}(q)|}$$

The entire sample space of $n \times n$ matrices whose square is the identity has cardinality

$$\sum_{k=0}^{n} \frac{|GL_n(q)|}{|GL_k(q)||GL_{n-k}(q)|}$$

As is well-known, $|GL_n(q)| = (q^n - 1)(q^n - q)(q^n - q^2) \cdots (q^n - q^{n-1})$. The reason is that an invertible matrix has any non-zero vector in the first column and the jth column has any vector not in the span of the $j - 1$ columns to the left, giving $q^n - q^{j-1}$ choices for that column.

With this we have the probabilities for the E_k

$$\mathrm{P}(E_k) = \frac{|GL_n(q)|}{|GL_k(q)||GL_{n-k}(q)|} \left(\sum_{m=0}^{n} \frac{|GL_n(q)|}{|GL_m(q)||GL_{n-m}(q)|} \right)^{-1}$$

Now we can answer the croupier's question: would you like to bet at even odds that there is still a red token remaining inside the box? With a bit of electronic help we calculate the conditional

probability $\mathrm{P}(E_5|R_5) = 0.5995\ldots$, which is the probability that all the remaining tokens are blue. We can understand why this might be so when we examine the probabilities $\mathrm{P}(E_k)$ and see that

$$\begin{aligned}
\mathrm{P}(E_0) = \mathrm{P}(E_{10}) &\approx 6.4 \times 10^{-25} \\
\mathrm{P}(E_1) = \mathrm{P}(E_9) &\approx 3.7 \times 10^{-16} \\
\mathrm{P}(E_2) = \mathrm{P}(E_8) &\approx 2.0 \times 10^{-9} \\
\mathrm{P}(E_3) = \mathrm{P}(E_7) &\approx 1.2 \times 10^{-4} \\
\mathrm{P}(E_4) = \mathrm{P}(E_6) &\approx .09 \\
\mathrm{P}(E_5) &\approx .82
\end{aligned}$$

Finally, we note that the binomial probabilities for the coin tossing game bear a formal similarity to those of the eigenvalue game. The number of ways to get k heads in n tosses is $\binom{n}{k}$, which is $\frac{|S_n|}{|S_k||S_{n-k}|}$, where S_n is the symmetric group on n letters, and we have seen that the cardinality of the set E_k is $\frac{|GL_n(q)|}{|GL_k(q)||GL_{n-k}(q)|}$. The sample space of matrices whose square is the identity is bijective with the space of ordered splittings of an n-dimensional vector space V. The matrix A corresponds to the pair (V_1, V_{-1}) where $V_{\pm 1}$ are the ± 1 eigenspaces and $V_1 \oplus V_{-1} = V$. Likewise, the coin tossing sample space corresponds to the ordered splittings of an n element set into two disjoint subsets with the first subset corresponding to the tosses in which heads occur. (In both cases we count the trivial splittings by allowing the zero subspace or the empty subset.) Of course, we have a simple expression for the cardinality of the sample space and the identity $\sum_{k=0}^{n} \binom{n}{k} = 2^n$, while we have no simple expression for the total number of splittings of a vector space of dimension n.

The number of heads in n coin tosses is a binomial random variable. The number of 1's in n rolls of an m sided die is a multinomial random variable. An analog of a multinomial random variable is the number of 1's among the eigenvalues of a random $n \times n$ matrix over a finite field

that satisfies the equation $A^m = I$. We should require that the field contains m distinct roots of unity in order that all these matrices are diagonalizable.