SOME PROBABILITY PROBLEMS Concerning the Game of BINGO

By **GERALD G. BROWN**
and **HERBERT C. RUTEMILLER**

California State University, Fullerton
Fullerton, California

WE HAVE found the well-known game of bingo to be an excellent device for illustrating to students various laws of probability and a source for hypergeometric probability problems to supplement the familiar playing card examples. In particular, solution of the question "What is the expected length of a game of bingo as a function of the number of players?" provides students an opportunity to develop the expectation of a random variable theoretically and then to verify the result by computer simulation of the game.

A 5×5 bingo card with the center free space is shown in figure 1. The first column of each player's card is filled with five random digits from the set 1–15, the second column with five from the set 16–30, the third with four digits from 31–45, the fourth with five digits from 46–60, and the fifth with five digits from 61–75.

Numbers are drawn at random, without replacement, from the set 1–75 until a player has had all the numbers called in a row, column, or major diagonal. We shall refer to this as completing a string. There are 12 strings for each player, as indicated in figure 1. They have been labeled $H_1, \cdots, H_4, V_1, \cdots, V_4$, representing the eight wins for which five numbers must be called, plus F_1, F_2, F_3, and F_4, which make use of the free space and require four numbers.

Some sample problems are:

1. How many unique bingo cards are there? How many unique cards are lost because of the free space? Without the

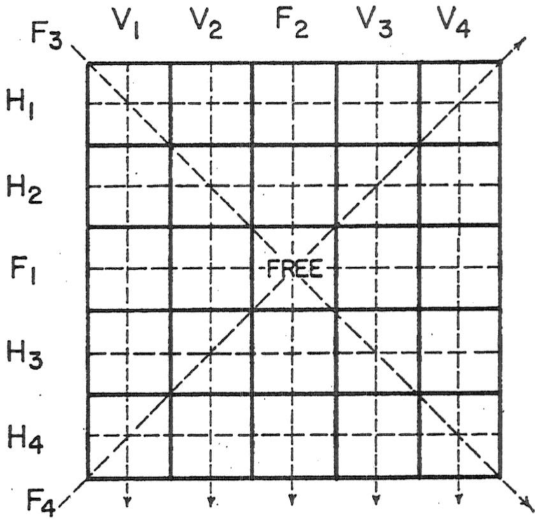

Fig. 1. A bingo card illustrating the notation used for the twelve possible winning strings

free space, each column may be selected in ${}_{15}P_5$ ways, so there are $({}_{15}P_5)^5$ cards. With the constraint, the number becomes ${}_{15}P_4({}_{15}P_5)^4$. The first number is approximately 6.08×10^{26}, and the second 5.52×10^{26}. Hence about 56,000,000,000,000,-000,000,000,000 cards are lost by deleting one square.

2. What is the probability that a particular player wins on the fourth draw? This can occur only by obtaining F_1, F_2, F_3, or F_4. Since these four events intersect only at the free space, they are mutually exclusive and their union may be obtained by adding the individual probabilities. Each of the four has the probability $\frac{1}{\binom{75}{4}}$, so the probability is $\frac{4}{\binom{75}{4}} = .0000033$.

3. If n players participate, what is the probability that the game ends on the fourth draw? How large must n be to make this probability appreciable, say 0.5?

The game continues if no one wins. If we make the simplifying assumption that the ratio of the number of players, n, to the number of possible cards, 5.52×10^{26}, is small, then the probability that no one wins on the fourth draw is approximately $(1 - .0000033)^n$. Hence, the probability that the game ends is about $1 - (1 - .0000033)^n$.

To answer the second question, we must solve, for n, $1 - .9999967^n = 0.5$:

$$n = 210{,}000, \text{ approximately.}$$

4. If a player wins the game on the kth draw, what is the probability that his winning configuration uses the free space?

This question requires use of the hypergeometric analog of the negative binomial distribution, since the game must end with completion of a row, column, or diagonal on draw k. We ignore, for simplicity, the slim probability that the game will end with a simultaneous completion of two or more strings.

For $F_1, \cdots, F_4$, we require that $(k - 1)$ draws contain three of the four winning numbers and that the kth draw completes the string. This probability is given by

$$\frac{\binom{4}{3}\binom{71}{k-4}}{\binom{75}{k-1}} \cdot \frac{1}{75 - k + 1}.$$

For the $H_1, \cdots, H_4, V_1, \cdots, V_4$, where five members are required, the probability is

$$\frac{\binom{5}{4}\binom{70}{k-5}}{\binom{75}{k-1}} \cdot \frac{1}{75 - k + 1}.$$

Finally, there are four of the first type and eight of the second. Therefore, the total probibility of a win on the kth draw is

$$\frac{1}{\binom{75}{k-1}}\left[4\cdot\binom{4}{3}\binom{71}{k-4} + 8\cdot\binom{5}{4}\binom{70}{k-5}\right]\frac{1}{75 - k + 1}.$$

Now, if we know that a win has occurred, the conditional probability that it involves the free space is

$$\frac{4\cdot\binom{4}{3}\binom{71}{k-4}}{4\cdot\binom{4}{3}\binom{71}{k-4} + 8\cdot\binom{5}{4}\binom{70}{k-5}},$$

which simplifies to

$$\frac{1136}{976 + 40k}.$$

Now, k may vary from 4 to 70. When k is 4, the probability is unity; and, as k increases, the probability that the free space is involved decreases. The interesting feature of this equation is that, for large k, it is more probable than not that the free space is not involved in a win. For example, at $k = 60$, the probability of using the free space is only about 1/3.

A somewhat more complicated question, which requires access to a computer for solution, is the following:

5. If n players participate, what is the expected length of a game?

Consider, first, the state of a specific individual. Let us determine the probability that this player has *not* completed any winning strings on draw k. We shall obtain this probability by counting the winning configurations and taking the complement. Each of the eight 5-number strings can occur in $\binom{5}{5}\binom{70}{k-5}$ ways, while each of the four 4-number strings can occur in $\binom{4}{4}\binom{71}{k-4}$ ways. We require the probability of the union of all the twelve strings at draw k. One could obtain the probability of one or more strings being

present approximately by simply computing

$$P_1 = \frac{8\binom{5}{5}\binom{70}{k-5} + 4\binom{4}{4}\binom{71}{k-4}}{\binom{75}{k}}. \quad (1)$$

However, the twelve strings are not mutually exclusive, so that many of the outcomes have been counted doubly, trebly, and so forth. Equation (1) can be corrected for the presence of two and three strings simultaneously by computing 2-string intersections:

$$P_2 = \frac{30\binom{8}{8}\binom{67}{k-8} + 24\binom{9}{9}\binom{66}{k-9}}{\binom{75}{k}} + \frac{12\binom{10}{10}\binom{65}{k-10}}{\binom{75}{k}} \quad (2)$$

The first term in the numerator represents intersections of type H_iF_2, H_iF_3, H_iF_4, V_iF_1, V_iF_3, V_iF_4, F_iF_j; the second term represents intersections of type H_iF_1, V_iF_2, H_iV_j; the third term represents intersections of type H_iH_j and V_iV_j.

In a similar manner, the 3-string intersections may be enumerated, giving

$$P_3 = \frac{1}{\binom{75}{k}} \left[16\binom{11}{11}\binom{64}{k-11} + 132\binom{12}{12}\binom{63}{k-12} + 48\binom{13}{13}\binom{62}{k-13}\right] \quad (3)$$

There are similar correction terms for 4, 5, and so forth strings present simultaneously, but the probabilities are so small that these terms may be ignored. We have then P (no win at trial k) $= 1 - (P_1 - P_2 + P_3)$. The probability distribution for each value of k may be obtained as first differences of these "no win" probabilities.

Finally, if n players participate with randomly drawn cards and if n is small relative to 5.5×10^{26}, then the probability that no one has won is approximately the product of the individual probabilities. The "no win" probabilities generated for a single player may be raised to the nth power to generate the distribution for n players.

The similarity of the combinatorial terms in equations (1), (2), and (3) makes their computation on a computer extremely simple. Evaluation of

$$\binom{75 - j + 1}{k - j + 1}$$

for $j = 1, \cdots, 14$ and $k = 4, \cdots, 50$ will generate all of the combinatorial terms required. Actually, k could be as large as 70, but the probability of no win after draw 50 is essentially zero.

A typical result is shown in figure 2,

Draw Number	Probability No Wins	Probability of Win on This Draw
4	.9998	.0002
5	.9991	.0007
6	.9974	.0017
7	.9938	.0036
8	.9873	.0065
9	.9766	.0107
10	.9603	.0163
11	.9370	.0233
12	.9050	.0320
13	.8628	.0422
14	.8096	.0532
15	.7449	.0647
16	.6696	.0753
17	.5854	.0842
18	.4953	.0901
19	.4034	.0919
20	.3144	.0890
21	.2330	.0814
22	.1631	.0699
23	.1070	.0561
24	.0652	.0418
25	.0367	.0285
26	.0188	.0179
27	.0087	.0101
28	.0036	.0051
29	.0013	.0023
30	.0004	.0009
31	.0001	.0003
32	.0000	.0001

Fig. 2. Expected length of game 18.28 draws (50 players)

for 50 players. The distribution is quite symmetrical, with mean, median, and mode all in the vicinity of 18 draws.

The theoretical results may be verified readily by using a computer to play the game. For n players, set up n 5×5 matrices, filling the five columns of each with random rectangular numbers drawn from 1–15, 16–30, 31–45, 46–60, 61–75. Then, random numbers are drawn without replacement from the set 1–75, and each number is checked for presence on each card. If present, the twelve strings are examined for a win. Figure 3 shows some results of 200 games played with 1, 5, 10, 15, and 20 players, along with the theoretical expectations.

The bingo problems described herein are suggested for students who are near completion of a first course in discrete probability. The computation of expected game length and the simulation of the game require a first course in a problem-oriented computer language such as FORTRAN or ALGOL. We find that students are particularly intrigued by the concept of simulation presented in this manner, involving analysis of a familiar game for which the empirical answers can be checked.

Number of Players	Mathematical Expectation	Mean Length of 200 Simulated Games
1	40.26	40.51
5	29.35	29.94
10	25.50	25.28
15	23.43	23.97
20	22.08	21.31

Fig. 3.—Comparison of analytical and empirical measures of expected game length