Bingo Problem

Brian Tomasik, 17,18, 20, 22 June 2003

Question:

Given a bingo board with five spaces across, five spaces down, and no free spaces. A bingo requires five spaces in a row horizontally, vertically, or diagonally. Each of the 25 squares has one unique letter from A to Y, as is shown in Figure 1. If a different letter from A to Y is called out each time without repetition, what is the chance of getting a bingo after 5, 6, 7, etc. letters are called? How many times does the probability improve each time the number of letters called increases? Is there a general formula for the probability depending on the number of letters called?

Possible Solution:

If only five letters are called:

I began by thinking about the probability of a bingo in just one row of the board. The chance that the first letter called will be in any particular square is 1/25, and since it doesn't matter which square of the row is chosen first, the probability that any of the squares will be chosen is 5/25. Once the first square is chosen, there remain only 24 squares that can be picked, and if the first square selected was one of the five in that particular row, there are only four possible squares left in that row, making the probability that the next letter is also on a square in that row 4/24. If the second square chosen is also in the row, the probability that the third letter will fall in the row is 3/23, for the fourth letter, it's 2/22, and for the fifth letter, it's 1/21. Thus, the probability for a bingo in one row after five letters are called is (5/25)(4/24)(3/23)(2/22)(1/21) = 1/53130 $1.882175795 \cdot 10^{-5}$.

There are twelve ways to get a bingo, with 5 rows across, 5 rows down, and 2 diagonal rows. Therefore, twelve times the above number should be approximately the probability for a bingo any of the twelve ways, although it isn't exact because, for example, if there were 100,000 ways to get a bingo and the above number were multiplied by 100,000, the result would exceed one. 12(1/53130) = 2/8855 $2.258610954 \cdot 10^{-4}$.

I next thought about the problem in terms of the probability of getting a bingo depending on the square chosen. Since there are only five letters called, each square chosen must be in a single row for a success. Once any square, except the nine on diagonal lines, is selected, which has a probability of 16/25, there remain only eight squares that could be called next that would be in the same row as the first (see Figure 2). Once one of these is selected, there are only three other squares in the row to be filled, and subsequently, there will be only two and then only one. Thus, the probability for five in a row beginning with one of the non-diagonal squares and with only five letters called is (16/25)(8/24)(3/23)(2/22)(1/21).

If the square chosen is on the lines of one of the diagonals, which has a 9/25 chance of occurring, there are twelve other squares that, if chosen, would be in the same row (see Figure 3). However, once one of these twelve spaces is selected, the rest of the squares must be in one certain row, so the number of possible squares after two letters have been chosen is again three, then two, then one, making the probability of five in a row with only five letters called and beginning with a diagonal square (9/25)(12/24)(3/23)(2/22)(1/21). Adding this to the other probability should be the chance of getting five in a row with five letters called, regardless of the location of the first square: (16/25)(8/24)(3/23)(2/22)(1/21) + (9/25)(12/24)(3/23)(2/22)(1/21) = 59/265650 2.220967438 · 10^{-4}, which is quite close, although slightly lower (as it probably should be) than the probability obtained by the first method.

Because they only differ by the numerators of the first and second fractions in each term, I attempted to simplify the two fraction products used above into one term by finding the average of the eight and twelve in the numerators of the second fraction. Sixteen times the numerator is eight and nine times it is twelve, out of a total of 25 times: (16 · 8 + 9 · 12)/25 = 236/25 = 9.44. Since it is guaranteed that one of the squares on the entire board will be chosen, the first fraction in the product will always be 25/25. Substituting 9.44 for both eight and twelve into one term: (25/25)(9.44/24)(3/23)(2/22)(1/21) 2.220967438 · 10^{-4}, which is exactly the same answer as above.

If six letters are called:

As always, the chance that any square will be chosen is 25/25. I began by investigating the possibilities if the second square chosen is not in line with the first, which has a probability of 14.56/24, or one minus the probability that it will be in the same row: 1-(9.44/24) = 14.56/24. If this situation occurs, either the remaining four squares must fall into a row with the first one chosen, the probability of which is (25/25)(14.56/24)(9.44/23)(3/22)(2/21)(1/20), or the remaining four squares will be part of a row created by the second, rather than the first, square chosen, which has a probability of (25/25)(14.56/24)(9.44/23)(3/22)(2/21)(1/20); the chance of either of these two possibilities, therefore, is the same. Figures 4a and 4b illustrate these two possibilities (the numbers represent the order in which the squares are chosen, so the square containing "1" is the first square selected).

Those situations, however, only show the result if the "extra" (sixth) square comes first or second. Once two squares chosen lie in the same row, there is no possibilitiy that a third square outside of that row could successfully form its own row since only six letters are called. Therefore, there is only way the third square called could be the extra square: (25/25)(9.44/24)(20/23)(3/22)(2/21)(1/20). (The fraction for the extra square was 20/23 because there were only three possible squares out of 23 total squares that would be within the row at that point.) Following the same procedure, the probability if the extra square is fourth is (25/25)(9.44/24)(3/23)(20/22)(2/21)(1/20), if the extra square is fifth, (25/25)(9.44/24)(3/23)(2/22)(20/21)(1/20), and if the extra square is sixth, (25/25)(9.44/24)(3/23)(2/22)(1/21)(20/20). The sum of all of these terms (counting the first and second ones separately, even though they are the same result, because they are obtained two different ways), is approximately 0.0012117598, which, assuming the two probabilities to be correct, is about 5.456000001 times better than the probability when only five letters are called.

It is important to note that each of the above products of six fractions is the probability for the situation in which the extra square is in each of the six positions that are possible around five fixed fractions, as Figure 5 illustrates. When there are seven letters called, there will be six possible positions of the first extra square, which, when it chosen, creates another "fixed point"; then, since there are now six fixed points, the second extra square can go in one of seven possible places in between, making the number of possible positions $6 \cdot 7 = 42$, because every one of the six positions for the first extra square is accompanied by seven possible positions for the second. When there are eight letters called, the third extra square will have eight possible positions in between the other seven fixed points, which means there are $6 \cdot 7 \cdot 8$ ways the three extra squares can combine. A general formula for the number of possible positions of the extra fractions, n, based on the number of letters called, w, when there are five squares in a row is $n = {}_{w}P_{w-5}$, because this is equivalent to w(w-1)(w-2)…(6). An even more general formula, in which s represents the number of squares in a row, would be $n = {}_{w}P_{w-s}$ because w-s is the number of extra squares that are present when a certain number of letters are called.

General form:

I first thought about the fact that the denominators of all of the fractions are the same (as long as the number of fractions multiplied is, as well) and that the denominator decreases by one each time a new fraction is multiplied. Therefore, the denominator for the entire formula should be a permutation: ${}_{24}P_{w-1}$. For example, when there are six letters (w = 6 and w-1 = 5), the denominator is always ${}_{24}P_5 = 24 \cdot 23 \cdot 22 \cdot 21 \cdot 20$. The one is subtracted from w because it represents the denominator of the first fraction, 25/25, which will be ignored throughout the formula, as it always equals one; for the same reason, I started with 24 instead of 25 before the P.

I next observed that the numerators are always $9.44 \cdot 3 \cdot 2 \cdot 1 \cdot$ extra(s) outside of the row, excluding the 25, of course. $9.44 \cdot 3 \cdot 2 \cdot 1 = 56.64$. For the two possibilities that arise when the second square chosen is not in a row with the first, the extra number to multiply is 14.56, and if there are multiple extras, 13.56, 12.56, etc. Although a permutation of a decimal isn't possible, I used it in the formula to indicate that it is necessary to multiply the number preceding the P by that number minus one, then that number minus one, and so on as many times as the number following the P shows is needed, just as a normal permutation is done: ${}_{14.56}P_{w-s}$. $(56.64)({}_{14.56}P_{w-s})$ is the product of the numerators for one of the terms, and since both of the combinations possible when the second square is not in line with the first are the same, the sum of them should be able to be expressed as $2(56.64)({}_{14.56}P_{w-s}) = 113.28({}_{14.56}P_{w-s})$.

For every combination possible after a single row has been established, the numerator of the first extra term is twenty, the second is 19, and so on. Therefore, the product of the extra numerators in each term is ${}_{20}P_{w-s}$ and the entire numerator for one term is $(9.44 \cdot 3 \cdot 2 \cdot 1)({}_{20}P_{w-s}) = 56.64({}_{20}P_{w-s})$. Since there are four different combinations with the same numerator (when the first extra square is third, fourth, fifth, or sixth), the sum of all four of there combinations is $4(56.64)({}_{20}P_{w-s}) = 226.56({}_{20}P_{w-s})$.

I later realized, however, that the generalization made two paragraphs above is flawed because it assumes that the extra squares always come before a single row is established by two squares on the same line. This isn't true because while the first extra square would have to come before two squares in the same row are chosen (or else it would be one of the later extra square possibilities), the second extra square could come later, even last, and the probability of getting an extra square later is higher because a single row for the non-extra squares has been chosen. I attempted to rectify this by creating a weighted average, as I did with 9.44/24 instead of 8/24 and 12/24, so that the squares could go in any position, not just before or after the point when two squares in the same row are chosen. Two out of six times the probability of not choosing a square in the row is 14.56/(number of squres left) and the other four out of six times, it's 20/(number of sqaures left): $(14.56 \cdot 2 + 20 \cdot 4)/6 = 18.186666\ldots = 1364/75$. Since this probability decreases by one each time a square outside the row is chosen, the probability of each successive extra square is the previous probability minus one. This will again be expressed with a permutation even though a permutation of a decimal isn't possible: ${}_{1364/75}P_{w-s}$. Multiplying this by the other numerators in a term equals $9.44 \cdot 3 \cdot 2 \cdot 1 \cdot {}_{1364/75}P_{w-s} = 56.64({}_{1364/75}P_{w-s})$.

This is only the numerator for one possible term, and must therefore be multiplied by the n, the number of possible arrangements of the extra squares to determine the entire numerator: $n(56.64)({}_{1364/75}P_{w-s}) = ({}_{w}P_{w-s})(56.64)({}_{1364/75}P_{w-s})$. Dividing this by the denominator derived above yields: $({}_{w}P_{w-s})(56.64)({}_{1364/75}P_{w-s})/({}_{24}P_{w-1})$.

Testing the general formula:

The formula should not and will not work when $w < s$, as this results in a permutation with a negative number following the P. This exact formula would probably not be correct if $s \neq 5$ since the constants, including 56.64 and 1364/75, as well as the 24 in ${}_{24}P_{w-1}$, are based on a bingo board with 25 squares.

If $w = 5$ and $s = 5$, the probability is $({}_{5}P_{5-5})(56.64)({}_{1364/75}P_{5-5})/({}_{24}P_{5-1}) = ({}_{5}P_{0})(56.64)({}_{1364/75}P_{0})/({}_{24}P_{4}) = (1)(56.64)(1)/(255024)\ \ 2.220967438 \cdot 10^{-4}$. This is exactly the same number I found before.

If $w = 6$ and $s = 5$, the probability is $({}_{6}P_{6-5})(56.64)({}_{1364/75}P_{6-5})/({}_{24}P_{6-1}) = ({}_{6}P_{1})(56.64)({}_{1364/75}P_{1})/({}_{24}P_{5}) = (6)(56.64)(1364/75)/(5100480)\ \ 0.0012117598$. Again, this is the same answer as I obtained before.

--what happens with the 18..... constant when all of the 8/(number..) squares are used up or when it gets to 0.186666????

--why diff with 9.44 and 12 and 8 in 6 letter method???
--what about, once I have w=10, all of the extras could form a row by themselves….????
--is there any other way to write ${}_{1364/75}P_{w-s}$ in the formula?????

Formula so far:

$$56.64({}_{w}P_{w-s})({}_{1364/75}P_{w-s})/({}_{24}P_{w-1})$$

w = number of letters called
s = number of squares needed for a bingo; s = 5