Journal of Progressive Research in Mathematics(JPRM)
ISSN: 2395-0218

SCITECH
RESEARCH ORGANISATION

Volume 13, Issue 1
Published online: February 15, 2018|

Journal of Progressive Research in Mathematics
www.scitecresearch.com/journals

Inclusion and Exclusion probability

William W.S. Chen
Department of Statistics
The George Washington University
Washington D.C.20013
E-mail: williamwschen@gmail.com

Abstract.

We use mathematical induction method to prove the Poincare Formula. To demonstrate the usefulness of this formula, we provide five examples. This formula is related to a broad class of counting problems in which several interacting properties either all must hold, or none must hold. When there are only two or three events that need to be counted, we usually use a Venn diagram. In section 4, we present a general mathematical formula to count any finite number of inclusion and exclusion events. This leads to an easy way to apply the Poincare Formula to define the probability

Mathematical Subject Classification: 60C05

Key Words and Phases: Begins with a vowel; Consecutive appearance; Ends with a vowel; Inclusion and exclusion probability; Kolomogorov axioms. Poincare Formula; Probability space; Randomly selected 5-letter word.

I Introduction

The inclusion and exclusion principle is a counting technique which generalizes the familiar method of obtaining the number of elements in the union of two finite sets.

The name comes from the idea that the principle is based on over by generous inclusion, followed by compensating exclusion. Generalizing the results to the union of n sets:

we include the cardinalities of the sets, then exclude the cardinalities of the pair-wise intersections, followed by inclusion of the cardinalities of the triple-wise intersections,

and exclusion of the cardinalities of the quadruple-wise intersections. Continuing, until the cardinality of the n-tuple-wise intersection is included or excluded. The formulas for the principle of inclusion-exclusion remain valid when the cardinalities of sets are replaced by finite probabilities. In this article, we adopt the fundamental inclusion and exclusion rule but write in the probabilities Poincare Formula. In this way, a broad class of the complicated counting probabilities can be solved. For example, if we ask:" what is the chance that for n people randomly reaching into a dark closet to retrieve their hats, no person will pick their own hat?" Another example: "What is the chance that, if we randomly select a positive integer ,i, between 1 through n, are there in which never selected immediately followed number i+1, for i=1,2,….n-1. Questions like these naturally involve counting the subsets of outcomes in which various combinations of the properties hold. Usually, when events are small, we use Venn Diagrams to depict these different combinations. When the events turn out to be large, we give out a more general mathematical formula counting rule. This formula is given in section 4. In the concluding remarks we provide a more general theorem due to Ch. Jordan that will turn the Poincare Formula to be a special case.

2. DEFINE THE PROBABILITY SPACES

In this section, we define the fundamental properties of the probability. Later, we use these properties to find the complicated events probabilities.

DEFINITION (Kolmogorov axioms)

The pair (Ω, a) is said to be a probabilizable space if there is a non-negative real valued function defined for each $A \in a$ such that

(1) $0 \leq P(A) \leq 1, \quad for all A \in a;$

(2) $\mathrm{P}(\Omega) = 1;$

(3) if $\{\mathrm{A_n}\}$ is a sequence of pairwise disjoint events in a, then

$$\mathrm{P}(\bigcup_{n \geq 1} \mathrm{A_n}) = \sum_{n \geq 1} \mathrm{P}(\mathrm{A_n})$$

The triple (Ω, a, P) is called a probability space.

next, we state some of the basic properties of the probability measure P, but do not prove it in this article. All of the proofs can be found in standard textbooks.

lemma Let $\mathrm{A,B,A_1,A_2,....} \in a$ *.then*

1) If $\mathrm{B} \subset \mathrm{A}$ then $\mathrm{P(B)} \leq \mathrm{P(A)}$.

2) $\mathrm{P}(\overline{\mathrm{A}}) = 1 - P(A)$.

3) $\mathrm{P}(\Phi) = 0$.

4) $\mathrm{P}(\mathrm{A} \bigcap \overline{\mathrm{B}}) = \mathrm{P(A)} - \mathrm{P}(\mathrm{A} \bigcap \mathrm{B})$

5) $\mathrm{P}(\mathrm{A} \bigcup \mathrm{B}) = \mathrm{P(A)} + \mathrm{P(B)} - \mathrm{P}(\mathrm{A} \bigcap \mathrm{B})$

6) $\mathrm{P}(\bigcup_{n=1}^{\infty} \mathrm{A_n}) \leq \sum_{n=1}^{\infty} \mathrm{P}(\mathrm{A_n})$

7) $\mathrm{P}(\mathrm{A} \bigcap \mathrm{B}) \geq 1 - \mathrm{P}(\overline{\mathrm{A}}) - P(\overline{B})$

Now, we use the above lemma to derive the main theorem, namely, "Poincare Formula",

as follows:

Let $A_1, A_2, \qquad A_n$, be arbitrary events then

$$P(\bigcup_{i=1}^{n} A_i) = \sum_{i=1}^{n} P(A_i) - \sum_{1\le i\le j\le n} P(A_i \bigcap A_j) + \sum_{1\le i\le j\le k\le n} P(A_i \bigcap A_j \bigcap A_k) - .....$$
$$+(-1)^{n+1} P(A_1 \bigcap A_2 \bigcap .... \bigcap A_n)$$

Proof: We prove this formula by induction on n. For n=1, the formula is trivial. The case

N=2 has been established in lemma (5). Now, let us assume that the formula is true for n=m. We need to prove it is true for n=m+1. We have

$$P(\bigcup_{i=1}^{m+1} A_i) = P[(\bigcup_{i=1}^{m} A_i) \bigcup A_{m+1}) = P(\bigcup_{i=1}^{m} A_i) + P(A_{m+1}) - P(\bigcup_{i=1}^{m} A_i) \bigcap P(A_{m+1})$$

$$= \sum_{i=1}^{m} P(A_i) - \sum_{1\le i\le j\le m} P(A_i \bigcap A_j) + \sum_{1\le i\le j\le k\le m} P(A_i \bigcap A_j \bigcap A_k) - .....(-1)^{m+1} P(A_1 \bigcap A_2 \bigcap .... \bigcap A_m)$$
$$+ P(A_{m+1}) - P\{\bigcup_{i=1} (A_i \bigcap A_{m+1})\}$$

$$= \sum_{i=1}^{m+1} P(A_i) - \sum_{1\le i\le j\le m} P(A_i \bigcap A_j) + \sum_{1\le i\le j\le k\le m} P(A_i \bigcap A_j \bigcap A_k) - .....(-1)^{m+1} P(A_1 \bigcap A_2 \bigcap .... \bigcap A_m)$$
$$- P\{\bigcup_{i=1}^{m} (A_i \bigcap A_{m+1})\} \qquad (2.1)$$

$$P\{\bigcup_{i=1}^{m} (A_i \bigcap A_{m+1})\} = \sum_{i=1}^{m} P(A_i \bigcap A_{m+1}) - \sum_{1\le i\le j\le m} P(A_i \bigcap A_j \bigcap A_{m+1}) + \sum_{1\le i\le j\le k\le m} P(A_i \bigcap A_j \bigcap A_k \bigcap A_{m-1})$$
$$(-1)^{m+1} P(A_1 \bigcap A_2 \bigcap .... \bigcap A_m \bigcap A_{m+1}) \qquad (2.2)$$

substituting (2.2) into (2.1), we have

$$P(\bigcup_{i=1}^{m+1} A_i) = \sum_{i=1}^{m+1} P(A_i) - \{\sum_{1\le i\le j\le m} P(A_i \bigcap A_j) + \sum_{1 i=1} P(A_i \bigcap A_{m+1})\}$$

$$+\{\sum_{1\le i\le j\le k\le m} P(A_i \bigcap A_j \bigcap A_k) + \sum_{1\le i\le j\le m} P(A_i \bigcap A_j \bigcap A_{m+1})\}$$

$$-...... (-1)^{n+2} P(A_1 \bigcap A_2 \bigcap .... \bigcap A_m \bigcap A_{m+1}) \quad \ldots$$

$$= \sum_{i=1}^{m+1} P(A_i) - \sum_{1\le i\le j\le m+1} P(A_i \bigcap A_j) + \sum_{1\le i\le j\le k\le m+1} P(A_i \bigcap A_j \bigcap A_k) -.....$$

$$(-1)^{n+2} P(A_1 \bigcap A_2 \bigcap .... \bigcap A_m \bigcap A_{m+1})$$

This establishes the Poincare Formula.

3. SOME SAMPLE EXAMPLES

Example 3.1: Find the probability that for a randomly selected 5-letter word, letters may be repeated, either beginning or ending with a vowel ?

Let B be an event that begins with a vowel, and E event that ends with a vowel.

Let N total possible outcomes,

$$\textit{then}\text{ we may find that } N = 26^5,\ N(B) = 5*26^4,\ N(E) = 5*26^4,\text{ and } N(B\bigcap E) = 5^2*26^3$$

$$\textit{use}\text{ the lemma } 5:$$

$$P(B\bigcup E) = P(B) + p(E) - P(B\bigcap E) = \frac{1}{26^5}(5*26^4 + 5*26^4 - 5^2*26^3) = 0.3476$$

Example 3.2

Find the probability that randomly selected n-digit ternary (0,1,2) sequences are there with at least one 0, at least one 1, and at least one 2?

Let us define A_i as the number of n-digit ternary sequences with no i's . where i=0,1,2.

Let a be the set of all n-digit ternary sequences. Then the probability we seek with at least one of each digit will be $P(\overline{A_0}\bigcap\overline{A_1}\bigcap\overline{A_2})$. The number of n-digit ternary sequences is $N = 3^n$. The number of n-digit ternary sequences with no 0's is simply the number of n-digit sequences of 1's and 2's. Thus $N(A_0) = 2^n$. Similary $N(A_1) = N(A_2) = 2^n$. The only n-digit sequence with no 0's or 1's is sequence of all 2's . Thus $N(A_0\bigcap A_1) = 1$, *also* $N(A_1\bigcap A_2) = N(A_0\bigcap A_2) = 1$, Finally, there is no ternary sequence with no 0's or 1's or 2's. Thus we can compute the joint probability as follows:

$$P(A_0\bigcup A_1\bigcup A_2) = \sum_{i=1}^{3} P(A_i) - \sum_{1\le i<j\le 3} P(A_i\bigcap A_j) + P(A_1\bigcap A_2\bigcap A_3)$$

$$= \frac{1}{3^n}(3*2^n - 3 + 0)$$

$$P(\overline{A_0}\bigcap\overline{A_1}\bigcap\overline{A_2}). = 1 - P(A_0\bigcup A_1\bigcup A_2) = 1 - \frac{1}{3^n}(3*2^n - 3 + 0)$$

$$= \frac{1}{3^n}(3^n - 3*2^n + 3)$$

Example 3.3.

Find the probability that a randomly selected positive integer between 1 and 280 is relatively prime to 280.

$$280 = 2^3 * 5 * 7$$

$$\begin{aligned}
&\textit{Let } \mathrm{A}_1 \textit{ is } \text{a selected } \textit{number of multiple } \text{of } 2 \le 280 \\
&\textit{Let } \mathrm{A}_2 \textit{ is } \text{a selected } \textit{number of multiple } \text{of } 5 \le 280 \\
&\textit{Let } \mathrm{A}_3 \textit{ is } \text{a selected } \textit{number of multiple } \text{of } 7 \le 280
\end{aligned}$$

$$\begin{aligned}
&\mathrm{N}(\mathrm{A}_1) = \frac{280}{2} = 140,\ \mathrm{N}(\mathrm{A}_2) = \frac{280}{5} = 56,\ \mathrm{N}(\mathrm{A}_3) = \frac{280}{7} = 40, \\
&\mathrm{N}(\mathrm{A}_1 \bigcap A_2) = \frac{280}{2*5} = 28,\ \mathrm{N}(\mathrm{A}_1 \bigcap A_3) = \frac{280}{2*7} = 20,\ \mathrm{N}(\mathrm{A}_2 \bigcap A_3) = \frac{280}{5*7} = 8 \\
&\mathrm{N}(\mathrm{A}_1 \bigcap A_2 \bigcap A_3) = \frac{280}{2*5*7} = 4
\end{aligned}$$

$$\begin{aligned}
&\mathrm{P}(\mathrm{A}_1 \bigcup \mathrm{A}_2 \bigcup \mathrm{A}_3) = \sum_{i=1}^{3} P(A_i) - \sum_{1 \le i < j \le 3} P(A_i \bigcap A_j) + P(A_1 \bigcap A_2 \bigcap A_3) \\
&= \frac{1}{280}(140 + 56 + 40 - 28 - 20 - 8 + 4) = \frac{184}{280}
\end{aligned}$$

$$P(\overline{A_1} \bigcap \overline{A_2} \bigcap \overline{A_3}) = 1 - \mathrm{P}(\mathrm{A}_1 \bigcup \mathrm{A}_2 \bigcup \mathrm{A}_3) = 1 - \frac{184}{280} = \frac{96}{280}$$

4. Define the Notation

Often we meet the case where n is large, (say, n is greater than 5). It would be tedious to write the Poincare Formula. In this section we introduce the summation notation to replace the unions or intersections.

Assume the sets $A_1, A_2, \ldots\ldots A_n$ are n different events.

$$\begin{aligned}
&\text{Let } \mathrm{S}_\mathrm{k} : \textit{total} \text{ sum of the sizes of the k-intersection of the sets of } A_1, A_2, \ldots\ldots A_n \\
&S_1 = N(A_1) + N(A_2) + \ldots . N(A_n) \\
&S_2 = \sum_{1 \le i \le j \le N} N(A_i \bigcap A_j) \\
&\ldots\ldots\ldots \\
&S_n = N(A_1 \bigcap A_2 \ldots . \bigcap A_n)
\end{aligned}$$

Number of elements in none of the sets

$$= N(\overline{A_1}\bigcap\overline{A_2}......\bigcap\overline{A_n})$$
$$= N - S_1 + S_2 - S_3.......+(-1)^n S_n$$

Suppose the number in each k-intersection for any k is the same n_k , then

$$N(\overline{A_1}\bigcap\overline{A_2}......\bigcap\overline{A_n})$$
$$= S_0 - \binom{n}{1}n_1 + \binom{n}{2}n_2 - \binom{n}{3}n_3 + ...+(-1)^n\binom{n}{n}n_n$$

Example 4.1.

What are the chances if we wish an arrangement of “a,a,a,b,b,b,c,c,c” without three consecutive letters being the same.

Let A_1 be three letter a's consecutive appearance,
Let A_2 be three letter b's consecutive appearance.
Let A_3 be three letter c's consecutive appearance.

$$\text{total possible number of outcome N(n)} = \frac{9!}{3!3!3!} = 1680$$
$$S_1 = \sum_{i=1}^{3} N(A_i) = \binom{3}{1}\frac{7!}{3!3!} = 420$$
$$S_2 = \sum_{1\le i\le j\le 3} N(A_i\bigcap A_j) = \binom{3}{2}\frac{5!}{3!} = 60$$
$$S_3 = 3! = 6$$

$$P(A_1\bigcup A_2\bigcup A_3) = \frac{1}{1680}(420-60+6) = 0.2179$$
we seek the probability
$$\mathrm{P}(\overline{A_1}\bigcap\overline{A_2}\bigcap\overline{A_3}) = 1 - P(A_1\bigcup A_2\bigcup A_3)$$
$$= 1 - 0.2179 = 0.7821$$

Example 4.2

What is the chance that if a random letter is selected from secret codes that can be made by assigning each letter of the alphabet a unique different letter?

Let A_i by assigning each letter of the alphabet a same letter.

$$P(A_1 \bigcup A_2 .... \bigcup A_{26})$$
$$= \frac{1}{26!}[\binom{26}{1}25! - \binom{26}{2}24! + \binom{26}{3}23! - ....(-1)^n \binom{26}{26}0!]$$

$$P(\overline{A_1} \bigcap \overline{A_2} .... \bigcap \overline{A_{26}}) = 1 - P(A_1 \bigcup A_2 .... \bigcup A_{26})$$
$$= \frac{1}{26!}\{26! - [\binom{26}{1}25! - \binom{26}{2}24! + \binom{26}{3}23! - ....(-1)^n \binom{26}{26}0!]\}$$

$$= \frac{1}{26!}\{\sum_{k=0}^{26}(-1)^k \binom{26}{k}(26-k)!\}$$
$$= \frac{1}{26!}\{\sum_{k=0}^{26}(-1)^k \frac{26!}{k!(26-k)!}(26-k)!\}$$
$$= \sum_{k=0}^{26}(-1)^k \frac{1}{k!} \approx e^{-1}$$

5. Concluding Remarks:

The Poincare formula is a special case of a more general theorem due to Ch. Jordan.

JORDAN THEOREM.

Let P(n, r) denote the probability of the occurrence of exactly r among the n given events

$A_1, A_2, ...A_n$ then

$$P(n;r) = \sum (-1)^k \binom{r+k}{k} S_{r+k}^{(n)}$$
$$where\ S_0^{(n)} = 1$$
$$S_j^{(n)} = \sum_{1<i_1<i_2....<i_j<n} P(A_{i_1} \bigcap A_{i_2} \bigcap ....A_{i_j})$$

J=1, 2, 3,… and the summation is to be extended over all combinations of the members

1,2,….n. at any time, repetitions are not allowed.

References

[1] Feller W. (1957) An Introduction to Probability Theory and Its Applications. Volume I, second edition, John Wiley & Sons, Inc.

[2] Feller W. (1965) An Introduction to Probability Theory and Its Applications. Volume II, John Wiley & Sons,

[3] Hoel P.G., Port S.C. and Stone C.J.(1971) Introduction To Probability Theory. Houghton Mifflin Company.

[4] Hoel P.G., Port S.C. and Stone C.J.(1971) Introduction To Statistical Theory. Houghton Mifflin Company.

[5] Hoel P.G., Port S.C. and Stone C.J.(1972) Introduction To Stochastic Processes. Houghton Mifflin Company.

[6] Renyi A. (1970) Probability Theory. North-Holland, Publishing Company-Amsterdam.

[7] Tucker A. (1980) Applied Combinatorics. John Wiley & Sons.