Why Your Friends Are More Popular Than You (Statistically Speaking)

By Joseph Kasongo and Jordan Lin

Agenda

1. Introduction to Basic Probability Concepts
   a. Counting Principle (Slide 1), Probability Axioms (Slide 2)
2. Permutation and Combination In Probability
   a. Permutation (Slide 3) and Combinations (Slide 4)
3. Random Variables and Expected Value
   a. Random Variable (Slide 5) and Expected Value (Slide 6)
4. The Friendship Paradox: Concept and Setup
   a. Paradox Introduction (Slide 7) and Setup (Slide 8)
5. Expected Friends: Random Person vs. Random Friend
   a. Random Person X (Slide 9) vs Random Friend Y (Slide 10)
6. Simplified inequality and Key Takeaways
   a. Inequality (Slide 11) and Takeaways (Slide 12)

Basic principle of counting

- If experiment 1 has m outcomes and experiment 2 has n outcomes, then there are $m \times n$ total possible outcomes for both.

- An outcome is (i,j) if experiment 1 gives its *i*th result and experiment 2 gives its *j*th, forming m rows of n outcomes, confirming there are $m \times n$ total possibilities.

Proof of the Basic Principle: The basic principle may be proven by enumerating all the possible outcomes of the two experiments; that is,

$$\begin{array}{l} (1,1),\ (1,2),\ \ldots,\ (1,n) \\ (2,1),\ (2,2),\ \ldots,\ (2,n) \\ \vdots \\ (m,1),\ (m,2),\ \ldots,\ (m,n) \end{array}$$

Example question: In a small community 10 women with 3 children, a women and one of her children are chosen as Parent and Child of the Year. How many different choices are possible?

Axioms of probability

Probability is based on three fundamental Axioms:

1.Non-negativity: For any event A, $P(A) \geq 0$.

2. Normalization: The probability of the entire sample space is 1, so $P(\Omega) = 1$.

3. Additivity: If A and B are mutually exclusive events, then $P(A \cup B) = P(A) + P(B)$.

Permutations

- Number of ways to order distinct objects
- There are 6 different ordered arrangements (permutations) of the letters a, b, and c because there are 3 choices for the first letter, 2 for the second, and 1 for the third, giving 3×2×1=6 possible permutations.

> Suppose now that we have n objects. Reasoning similar to that we have just used for the 3 letters then shows that there are
>
> $$n(n-1)(n-2)\cdots 3 \cdot 2 \cdot 1 = n!$$
>
> different permutations of the n objects.

- $n!$ (read as “n factorial”) is defined to equal $1 \cdot 2 \cdots n$ when n is a positive integer.

Example question: How many different batting orders are possible for a baseball team consisting of 9 players?

Solution: There are 9! = 362,880 possible batting orders.

Combinations

We are often interested in how many different groups of r objects can be formed

from a total of n objects, regardless of order. This is known as a combination.

General rule:

$$\frac{n(n-1)(n-2)\cdots(n-r+1)}{r!}$$

Notation:

$$\binom{n}{r} = \frac{n!}{r!(n-r)!}$$

Example question: A committee of 3 is to be formed from a group of 20 people.How many different

Committees are possible?

Random Variable

A random variable assigns a value to an outcome based on random processes.

- Rolling a 6 sided die, X = number on the die
- Probabilities: P (X = i) = $\frac{1}{6}$ for
  - I - 1,2… 6

Then:

- $P(X = 1) = \frac{1}{6}$
- $P(X = 2) = \frac{1}{6}$
- ...
- $P(X = 6) = \frac{1}{6}$

Expected Value

The expected value is the average possible value a random variable can take.

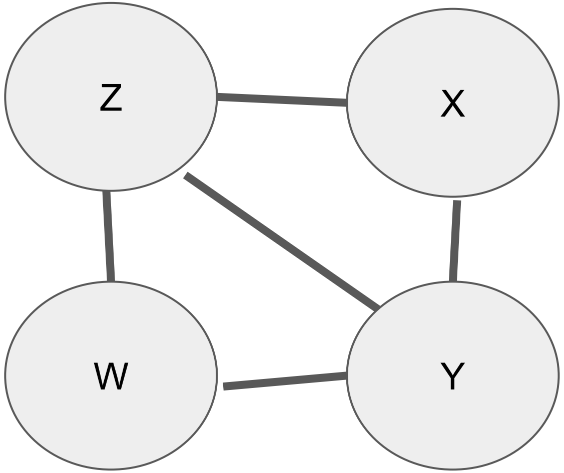
Using the same die:

$$E[X] = \sum_{i=1}^{6} i \cdot P(X = i) = \frac{1+2+3+4+5+6}{6} = \frac{21}{6} = 3.5$$

While we can’t roll a 3.5, the expected value of a die roll is **3.5**

The Friendship Paradox

1. On average, your friends have more friends that you do
2. Below is a sketch of 4 nodes (friend) and edges connection them (“friendships”)



Set up

- n people labeled $1, ..., n$
- $f(i)$ = number of friends of person i
- Total friendships: $f = \sum_{i=1}^{n} f(i)$
- Example: $f(1) = 3,\ f(2) = 2,\ f(3) = 1,\ f(4) = 2,\ f = 8$

Expected Friend of a Random Person X

- Let X be a person chosen at random from a population:

$$P(X = i) = \frac{1}{n} \quad \text{for each } i = 1, 2, ..., n$$

  - The number of friends of person i is $f(i)$
  - Expected number of friends of a random person:

$$E[f(X)] = \sum_{i=1}^{n} f(i) \cdot \frac{1}{n} = \frac{1}{n} \sum_{i=1}^{n} f(i)$$

  - This is just the **average number of friends** across all individuals
- This is just the average number of friend across all individuals

Expected Friend of a Random Friend Y

- Let Y be a randomly chosen friend from the network. THe probability of picking person i as a friend is proportional to their number of friends, *f*(i)

$$f = \sum_{i=1}^{n} f(i)$$

  - So:

$$P(Y = i) = \frac{f(i)}{f}$$

  - Expected number of friends of a randomly chosen friend:

$$E[f(Y)] = \sum_{i=1}^{n} f(i) \cdot \frac{f(i)}{f} = \frac{1}{f} \sum_{i=1}^{n} f(i)^2$$

- This value is generally higher than the Expected value of f(x)

Simplified Inequality

$$E[f^2(X)] = \sum_{i=1}^{n} \frac{f(i)^2}{n} \geq (\frac{\sum f(i)}{n})^2 = E[f(X)]^2$$

- Therefore

$$E[f(Y)] = \frac{E[f^2(X)]}{E[f(X)]} \geq E[f(X)]$$

Overall Takeaways

- Samping via friendship will overweight popular individuals because they tend to have more friends
- This results in… $E[f(Y)] \geq E[f(X)]$
- Applies to social media, epidemiology, survey etc
- It’s purely statistical not a personal popularity