SI140 Discussion 02

Li Zeng, Tao Huang, Xinyi Liiu

ShanghaiTech University, China
{zengli,huangtao1,liuxy10}@shanghaitech.edu.cn

1 Naive Probability

1.1 Set

A set is a collection of non-repeating objects. Given two sets A, B, key concepts include

- Union of A and B: $A \cup B$
- Intersection of A and B: $A \cap B$
- Complement of A: A^c
- Cardinality of A: $\mid A \mid$
- De Morgan's laws:

$$\begin{aligned}(A \cup B)^c = A^c \cap B^c \\ (A \cap B)^c = A^c \cup B^c\end{aligned} \tag{1}$$

1.2 Naive Definition of Probability

- *The sample space* S of an experiment is the set of all possible outcomes of the experiment.
- *An event* A is a subset of the sample space S, and we say that A occurred if the actual outcome is in A.
- Let A be an event for an experiment with a finite sample space S. The *naive probability of A* is

$$P_{naive}(A) = \frac{\mid A \mid}{\mid S \mid} = \frac{number\, of\, outcomes\, favorable\, to\, A}{total\ number\, of\, outcomes\, in\, S} \tag{2}$$

Note: be careful when using this formula, because you have to make sure that all your outcomes have equal probability! To use eqn.2 efficiently, you need to know what are counting rules, permutation and combination, which will be explained in the following section.

2 Counting and Combinatorics

2.1 Basic Counting Rules

- **Sum Rule:** If the outcome of an experiment can either be one of m outcomes or one of n outcomes, where none of the outcomes in the set of m outcomes is the same as the any of the outcomes in the set of n outcomes, then there are $m + n$ possible outcomes of the experiment.
- **Product/Multiplication Rule:** If an experiment has two parts, where the first part can result in one of m outcomes and the second part can result in one of n outcomes regardless of the outcome of the first part, then the total number of outcomes for the experiment is mn.

2.2 Inclusion-Exclusion Principle on Sets

Basic idea of IE Principle: If the outcome of an experiment can either be drawn from set A or set B, and sets A and B may potentially overlap (i.e., it is not guaranteed that $A \cap B = \Phi$), then the number of outcomes of the experiment is $|A \cup B| = |A| + |B| - |A \cap B|$.

Exercise 1. An 8-bit string (one byte) is sent over a network. The valid set of strings recognized by the receiver must either start with 01 or end with 10. How many such strings are there?

Solution 1. The potential bit strings that match the receiver's criteria can either be the 64 strings that start with 01 (since that last 6 bits are left unspecified, allowing for $2^6 = 64$ possibilities) or the 64 strings that end with 10 (since the first 6 bits are unspecified). Of course, these two sets overlap,since strings that start with 01 and end with 10 are in both sets. There are $2^4 = 16$ such strings (since the middle 4 bits can be arbitrary). Casting this description into corresponding set notation, we have: $|A| = 64$, $|B| = 64$, and $|A \cap B| = 16$, so by the Inclusion-Exclusion Principle, $64 + 64 - 16 = 112$ strings that match the specified receiver's criteria.

2.3 Permutations

- **Permutation Rule:** A permutation is an ordered arrangement of n distinct objects. Those n objects can be permuted in $n \cdot (n-1) \cdot (n-2) \cdot \cdots \cdot 2 \cdot 1 = n!$ ways.
- **Permutation of Indistinct Objects:** Generally when there are n objects and n_1 are the same (indistinguishable), n_2 are the same, ..., and n_r are the same, then there are $\frac{n!}{n_1! n_2! \cdots n_r!}$ distinct permutations of the objects.

2.4 Combinations:

1. **Combinations:** A combination is an un-ordered selection of r objects from a set of n objects. If all objects are distinct, then the number of ways of making the selection is: $\frac{n!}{r!(n-r)!} = \binom{n}{r}$, read as n choose r.
2. **Some useful identities** for combinations are as follows:
   (a) The Team Captain:
   $$n\binom{n-1}{r-1} = r\binom{n}{r}$$
   Exercise 2. calculate $\sum_{r=1}^{n} r^2 \binom{n}{r}$.

   Solution 2.

   (b) Vandermonde's Identity:
   $$\binom{m+n}{k} = \sum_{j=0}^{k} \binom{m}{j}\binom{n}{k-j}$$
   Exercise 3. Show a story-proof of the equation above.

   Solution 3.

   (c) Choosing a Committee:
   $$\binom{n}{k}\binom{k}{m} = \binom{n}{m}\binom{n-m}{k-m}, \quad \forall n \geq k \geq m \geq 0$$
   Exercise 4. Show a story-proof of the equation above.

   Solution 4. We will show that both sides of the equation count the number of ways to choose a committee of k students from a student body of n students, where, in addition, a subcomittee of m of the k students form the executive committee. The counting process for LEFT term: First, we choose k students from the student body of n students, to form the committee. There are $\binom{n}{k}$ ways to do this. Then we choose m students from among those k to form the subcommittee. There are $\binom{k}{m}$ ways to do this. The counting process for RIGHT term: First, we choose m students from the student body of n students, to form the executive committee. There are $\binom{n}{m}$ ways to do this. Then we choose km of the remaining portion of the student body (which consists of nm students), to form the non-executive part of the committee. There are $\binom{n-m}{k-m}$ways to do this.

(d) The One that is Special:

$$\binom{n}{r} = \binom{n-1}{r-1} + \binom{n-1}{r}, \quad \forall 0 \le r \le n$$

Exercise 5. Show a story-proof of the equation above.

Solution 5. This identity can be proved via a combinatorial argument. When we select a group of size r from n distinct objects, then any particular object (say, object 1) will either be part of that group or not part of that group. We can then define sets A and B, where A is the number of ways of selecting a group that contains object 1, and B is the number of ways of selecting a group that does not contain object 1. For Set A, if we decide to include object 1, then we must select $r-1$ of remaining $n-1$ objects (since the membership of object 1 in our selection is already decided), or $\binom{n-1}{r-1}$. For Set B, if we decide to exclude object 1, then we only have $n-1$ possible objects to select from to create a group of size r, or $\binom{n-1}{r}$. These sets are mutually exclusive, and therefore by the Sum Rule of Counting the total number of possibilities are as above.

3. **(Important) Bose-Einstein:** How many ways are there to choose k times from a set of n objects with replacement, if order does not matter (we only care about how many times each object was chosen, not the order in which they were chosen)?
The only thing we care is the times each object was chosen, or a length-n and sum-k non-negative integer-valued vector. We finally are looking for how many these vectors are.
Two important Results for this:
(a) There are $\binom{r-1}{n-1}$ distinct positive integer-valued vectors $(x_1, x_2, \cdots, x_n)$ satisfying

$$x_1 + x_2 + \cdots + x_n = r, \quad x_i > 0, i = 1, 2, \cdots, n$$

(b) There are $\binom{n+k-1}{n-1}$ distinct non-negative integer-valued vectors $(x_1, x_2, \cdots, x_n)$ satisfying

$$x_1 + x_2 + \cdots + x_n = k, \quad x_i \ge 0, i = 1, 2, \cdots, n$$

Exercise 6. Say you are a startup incubator and you have \$10 million to invest in 4 companies (in 1 million increments). How many ways can you allocate this money? What if you know you want to invest at least \$3 million in Company 1? What if you don't have to invest all \$10 M? (The economy is tight, say, and you might want to save your money.)

Solution 6. i. This is just like putting 10 balls into 4 urns. Using the Divider Method we get: $\binom{10+4-1}{10} = 286$. This problem is analogous to solving the integer equation $x_1 + x_2 + x_3 + x_4 = 10$, where x_i represents the investment in company i such that $x_i \ge 0$ for all $i = 1, 2, 3, 4$.

ii. There is one way to give \$3 million to Company 1. The number of ways of investing the remaining money is the same as putting 7 balls into 4 urns: $\binom{7+4-1}{7} = 120$. This problem is analogous to solving the integer equation $x_1 + x_2 + x_3 + x_4 = 10$, where $x1 \ge 3$ and $x2, x3, x4 \ge 0$. To translate this problem into the integer solution equation that we can solve via the divider method, we need to adjust the bounds on x_1 such that the problem becomes $x_1 + x_2 + x_3 + x_4 = 7$.

iii. Imagine that you have an extra company: yourself. Now you are investing \$10 million in 5 companies. Thus, the answer is the same as putting 10 balls into 5 urns: $\binom{10+5-1}{10} = 1001$. This problem is analogous to solving the integer equation $x_1 + x_2 + x_3 + x_4 + x_5 = 10$, such that $x_i \ge 0$ for all $i = 1, 2, 3, 4, 5$.

Exercise 7. How many distinct integer-valued vectors $(x_1, x_2, \cdots, x_n)$ satisfying $x_1 + x_2 + \cdots + x_n = r, \quad x_i \ge -1, i = 1, 2, \cdots, n$?

Solution 7.

3 General Definition of Probability

3.1 Axioms and Corollaries of Probability

$$\begin{aligned}
&\text{Axiom } 1: 0 \leq P(E) \leq 1 \\
&\text{Axiom } 2: P(S) = 1 \\
&\text{Axiom 3: If E and F are mutually exclusive } (E \cap F = \emptyset), \text{ then } P(E) + P(F) = P(E \cup F) \\
&\text{Corollary } 1: P(E^c) = 1 - P(E) \quad (= P(S) - P(E)) \\
&\text{Corollary } 2: \text{ If } E \subseteq F, \text{ then } P(E) \leq P(F) \\
&\text{Corollary } 3: P(E \cup F) = P(E) + P(F) - P(EF)
\end{aligned}$$

3.2 Two perspectives of Probability

- *Frequentist*: P(E)=$\lim_{n\to\infty} \frac{n(E)}{n}$
- *Bayesian*: Probability represents a degree of belief about the event in question(Next topic).

3.3 Equally Likely Outcomes in Sample Space

We like those sample spaces with equally likely outcomes; they make it simple to compute probabilities using naive definition of probability.

1. Coin flip: $S = \{$ Heads, Tails $\}$.
2. Flipping two coins: $S = \{(H,H),(H,T),(T,H),(T,T)\}$

Exercise 8. There are 4 oranges and 3 apples in a bag . You draw out 3. What is the probability that you draw 1 orange and 2 apples? *Hint: Consider this problem with 2 ways – whether using equally likely outcome sample space or not*

Solution 8. We have two different perspectives to solve this problem.

- *Method 1 - Create a ordered group to have an equally likely outcome sample space:*If we treat the oranges and apples as indistinct, we do not have a space with equally likely outcomes. We therefore treat all objects as distinct. Suppose we treat each outcome in the sample space as an ordered list of three distinct items. The size of the sample space, S, is simply the total number of ways to order 3 of 7 distinct items: $|S| = 7\cdot6\cdot5 = 210$. We can then decompose the event, E, into three mutually exclusive events, where we pick the orange first, second, or third, respectively: $|E| = 4\cdot3\cdot2+3\cdot4\cdot2+3\cdot2\cdot4 = 72$ The probability of our event is therefore $P(E) = 72/210 = 12/35$
  Note that experiments with indistinguishable objects often have sample spaces that are not equally likely. However, if we make the objects distinguishable, then we can create a sample space that with equally likely outcomes, because we do not need to consider overcounting/non-unique outcomes. In probability, since we are just looking to take a ratio of number of outcomes in a sample space with equally likely outcomes of our own design, we do not need to explicitly recreate the indistinguishable case.
- *Method 2 - Treat each outcome in the sample space as an unordered group:* The size of the sample space, S, is the total number of ways to choose any 3 of 7 distinct items: $|S| = \binom{7}{3}$. The event space is then the way to pick 1 distinct orange (out of 4) and 2 distinct apples (out of 3), which we combine with the product rule: $|E| = \binom{4}{1}\binom{3}{2}$. The probability of our event is therefore
  $$P(E) = \frac{\binom{4}{1}\binom{3}{2}}{\binom{7}{3}} = 12/35$$

3.4 Properties

- *Boole's Inequality*: $P\left(\bigcup_{i=1}^{\infty} A_i\right) \leq \sum_{i=1}^{\infty} P\left(A_i\right)$
- *Bonferroni's Inequality*: $P\left(A_1 \cap A_2 \cap \cdots \cap A_n\right) \geq P\left(A_1\right)+P\left(A_2\right)+\cdots+P\left(A_n\right)-(n-1)$
- *General Inclusion-Exclusion (IE) Formula in Probability:* For any events $A_1, A_2, \cdots, A_n$:

$$P(\bigcup_{i=1}^{n} A_i)=\sum_{i=1} P(A_i)-\sum_{i<j} P(A_i \cap A_j)+\sum_{i<j<k} P(A_i \cap A_j \cap A_k)+\cdots+(-1)^{n+1} P(A_1 \cap \cdots A_n)$$

Or more compactly,

$$P\left(\bigcup_{i=1}^{n} A_i\right)=\sum_{r=1}^{n}(-1)^{r+1} \sum_{i_1<\cdots<i_r} P\left(A_{i_1} A_{i_2} \ldots A_{i_r}\right)$$

Exercise 9. (Optional) Derive Bonferroni's Inequality above.

Solution 9. See note 1-11.

Exercise 10. There are N passengers in a plane with N assigned seats (N is a positive integer), but after boarding, the passengers take the seats randomly. Assuming all seating arrangements are equally likely.

a. Let us define A_i : the first i passengers sit in the assigned seats. What is $P(A_i)$?
b. What is the probability that no passenger is in their assigned seat? Compute the probability when $N \to \infty$?
Hint: Use the inclusion-exclusion principle
Another Hint: Use conclusion from a).

Solution 10. a. The probability of that the first i passengers sit in the assigned seats is :

$$P(A_i)=\frac{(N-i)!}{N!}$$

b. We know that the probability of that no passengers is in their assigned seats is

$$P(B)=1-P\left(\bigcup_{n=1}^{N} A_i\right)$$

Using the inclusion-exclusion principle,

$$P\left(\bigcup_{n=1}^{N} A_i\right)=\sum_{i} P\left(A_i\right)-\sum_{i<j} P\left(A_i \cap A_j\right)+\sum_{i<j<k} P\left(A_i \cap A_j \cap A_k\right)-\ldots+(-1)^{N-1} P\left(A_1 \cap A_2 \ldots \cap A_N\right)$$

Let's define $p_i=P(A_i)$ since we notice that all seats are the same, therefore,

$$P\left(\bigcup_{n=1}^{N} A_i\right)=\sum_{i=1}^{N}(-1)^{i-1}\binom{N}{i} p_i=\sum_{i=1}^{N}(-1)^{i-1} \frac{N!}{i!(N-i)!} \frac{(N-i)!}{N!}=\sum_{i=1}^{N}(-1)^{i-1} \frac{1}{i!}$$

So

$$P(B)=1-\sum_{i=1}^{N}(-1)^{i-1} \frac{1}{i!}=\sum_{i=0}^{N}(-1)^{i} \frac{1}{i!}$$

Let $N \to \infty$, it becomes

$$P(B)=\sum_{i=0}^{\infty}(-1)^{i} \frac{1}{i!}=1/e$$