

# Lec 1 basic combinatorics

## 1.1 permutations

### Def 1.1 (permutations)

一个 permutation 就是对一组 objects 的一个 rearrangement (这些 objects 中可以有 same 的也可以有 distinct 的).

对于  $n$  个 distinct objects, 一共存在

$$n! = n(n-1)(n-2)\cdots$$

个 permutations.



**Example 1.1** 求 "STATISTICS" 的 # distinct permutations.

**Sol.** 这里一共有 10 个 objects. 但问题是: 其中有 3 个  $S$ , 3 个  $T$ , 2 个  $I$  是相同的.

于是: 我们首先假设它们都是 distinct 的, 则存在  $10!$  个 permutations. 而, 每个 permutation 都包含了对 3 个  $S$  的一个子 permutation. 而对 3 个  $S$  的任意 permutation 都是相同的! 同样的道理 apply to 3 个  $T$  和 2 个  $I$ .

所以这个结果是真实结果的  $3!3!2!$  倍.

同样地, 由于因而, 正确结果是:

$$\frac{10!}{3!3!2!1!1!}$$

**Remark** 求一组 objects 的 distinct permutations 的数量时, 对于其中相同的 objects, 我们只需要除去它们的重复数量的 factorial (即: 它们自己内部有多少个 permutation 都算作一个 permutation).

公式:

$$\frac{n!}{n_1!n_2!\cdots n_k!}$$

其中  $n_1, n_2, \dots, n_k$  是每个 object 的重复数量.

这个式子又叫做 multinomial coefficient.

### Def 1.2 (multinomial coefficient)

令  $n \in \mathbb{N}, n_1, n_2, \dots, n_k \in \mathbb{N}, n_1 + n_2 + \dots + n_k = n$ , 我们定义 multinomial coefficient:

$$\binom{n}{n_1, n_2, \dots, n_k} = \frac{n!}{n_1!n_2!\cdots n_k!}$$



### Proposition 1.1

如果我们需要  $n_1$  个 object 1,  $n_2$  个 object 2,  $\dots$ ,  $n_k$  个 object  $k$ , 那么它们的 distinct permutations 的数量为:

$$\binom{n_1 + n_2 + \dots + n_k}{n_1, n_2, \dots, n_k}$$



## 1.2 combinations

### Def 1.3 (combinations)

一个 combination 就是从一个 set 中选取若干个 elements, 而忽略它们的顺序.



### Proposition 1.2

从  $n$  个 distinct objects 中选取  $k$  个的 combinations 的数量为:

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



**Proof** 我们可以将问题转化为: 从  $n$  个 distinct objects 中选取  $k$  个的 permutations 的数量, 然后再除去重复的 permutations. 而, 从  $n$  个 distinct objects 中选取  $k$  个的 permutations 的数量为:

$$n \times (n-1) \times \cdots \times (n-k+1) = \frac{n!}{(n-k)!}$$

而其中, 对于每个 valid combination, 都包含了它的所有 ordered permutations, 即重复了  $k!$  次. 因此, 最终的结果为:

$$\frac{n!}{k!(n-k)!}$$

### 1.2.1 binomial theorem

### Theorem 1.1 (Binomial Theorem)

令  $x, y \in \mathbb{R}, n \in \mathbb{N}$ , 则有:

$$(x+y)^n = \sum_{k=0}^n \binom{n}{k} x^{n-k} y^k$$



**Proof** 我们可以 prove this by combinatorial interpretation. 因为把  $x+y$  展开即  $n$  个  $(x+y)$  的乘积. 即: 对于每个被乘项, 我们都是在  $x$  和  $y$  之间选择一个.

因而:  $x^k y^{n-k}$  的系数就是从  $n$  个  $(x+y)$  中选取  $k$  个  $x$  的 combinations 的数量, 即  $\binom{n}{k}$ .

考虑所有的 possible  $k$  值, 我们得到:

$$(x+y)^n = \sum_{k=0}^n \binom{n}{k} x^{n-k} y^k$$

这是 combinatorial 的 proof.

另外一种更轮椅的思路是 prove by induction. 这需要一个辅助的 proposition:

### Proposition 1.3

$$\binom{n}{k} = \binom{n-1}{k-1} + \binom{n-1}{k}$$



这个等式的 combinatorial interpretation 很 trivial: 对于其中的任意一个 object:

- 这个 object 被选中的情况, combinations 的数量:  $\binom{n-1}{k-1}$  (从其他里面选  $k-1$  个);
- 这个 object 不被选中的情况, combinations 的数量:  $\binom{n-1}{k}$  (从其他里面选  $k$  个).

## problems

**Example 1.2** 一个 52-card deck, 取 5 张随机牌, 我们获得:

- 4 张同 rank 的牌, 最后一张不同 rank 的牌
- a full house (3 张同 rank 的牌, 2 张同 rank 的牌)

的概率是多少?

**Sol.** 一共有  $\binom{52}{5}$  种取法. 取 4 张同 rank 的牌: 13 种取法. 取最后一张不同 rank 的牌:  $52-4 = 48$  种取法. 因而, 概率是:

$$\frac{13 \times 48}{\binom{52}{5}}$$

如果是取 3 张同 rank 的牌 + 两张 different 同 rank 的牌: 我们首先在 4 个花色里面选 3 个, 有  $\binom{4}{3}$  种取法. 因而选取 3 cards of the same rank 的数量为:  $13 \cdot \binom{4}{3}$ .

然后选取剩余的两张: 然后故技重施, 从剩下的 12 个 rank 里面选 1 个, 而选择它们的花色有  $\binom{4}{2}$  种取法. 因而, 概率是:

$$\frac{13 \cdot \binom{4}{3} \cdot 12 \cdot \binom{4}{2}}{\binom{52}{5}}$$

**Example 1.3** 我们有  $n$  把钥匙, 其中有一把是正确的. 尝试  $k$  次, 能够成功开门的概率是多少?

**Sol.** 一共有  $n!$  种钥匙的 permutations. 我们需要的情况: 正确的钥匙出现在前  $k$  个位置:

- 正确钥匙出现在第 1 个位置, 其他  $n-1$  随便排列:  $\cdot (n-1)!$  种
- ...
- 正确钥匙出现在第  $k$  个位置, 其他  $n-1$  随便排列:  $k \cdot (n-1)!$  种

因而正确的 permutations 的数量为:

$$k \cdot (n-1)!$$

因而, 概率是:

$$\frac{k \cdot (n-1)!}{n!} = \frac{k}{n}$$

**Remark** 正确钥匙出现在每个位置上的概率都是  $1/n$ . 因而它出现在前  $k$  个位置的概率就是  $k/n$ .

这是一类典型的问题: 不放回的抽取. 在只关心“是否在前  $k$  次成功”这一事件时, 这个过程等价于“正确钥匙在一个随机排列中的位置”. 这一结果只和比例有关, 与过程细节无关. 只要没有信息 bias, 没有偏好, 完全随机, 那么成功概率只取决于:

- 你允许的尝试次数  $k$
- 总可能性数  $n$

而如果是放回则是不同的情况. 通过补事件容易得:

$$\mathbb{P}(A) = 1 - \left(1 - \frac{1}{n}\right)^k$$

放回和不放回最大的区别是: 放回时, 每次尝试都是独立的, 而不放回时, 每次尝试都不是独立的. 最明显的例子就是, 如果不放回,  $k = n$  时概率为 1. 而如果放回, 不论  $k$  有多大, 概率都小于 1; 当  $k \ll n$  的时候, 这两个概率相近 (符合直觉, 因为  $n$  很大时放回和不放回几乎没区别.)

**Example 1.4** 一个篮子里有 10 个 red balls 和 5 个 blue balls. 我们随机从中取出 3 个 balls, exactly 其中 1 个是 blue ball 的概率是多少? 如果每次都放回呢?

**Sol.** 不放回:

$$\mathbb{P}(\text{exactly one blue ball}) = \binom{10}{2} \binom{5}{1} / \binom{15}{3} = \frac{45}{91}$$

放回: 5 ways to choose the blue ball, 10 ways to choose the red ball, 以及 3 positions to place the blue ball,

$$\mathbb{P}(\text{exactly one blue ball}) = 3 \cdot \frac{3 \cdot 10^2}{15^3}$$

## 1.3 combinations with repetition

### Def 1.4 (combinations with repetition)

一个 combination with repetition 就是从一个 set 中选取若干个 elements, 而忽略它们的顺序, 并且允许重复选取.



### Proposition 1.4

从  $n$  个 distinct objects 中选取  $k$  个的 combinations with repetition 的数量为:

$$\binom{n+k-1}{k}$$



**Remark** 关于这个问题我们最开始可能会犯一个错误: 认为 combinations with  $k$  repetitions 的数量为:

$$\binom{kn}{k}$$

但是想一下就知道这是错的. 因为我们相当于给 repeated 的同一个 objects 赋予了顺序, 从而计入了额外的数量.

**Proof** 这个问题比较巧妙. 我们上面错误的尝试已经表明: 用 "make copies" 的方法行不通. 我们需要变换一下思路. 原问题是 "要选哪几个元素, 每个元素要选几个". 而我们可以把这个问题理解为: 一共有  $k$  个位置,  $n$  个组, 我们给每个组分配多少个位置?

Formalize 这个想法即: 对于第  $i$  个 object, 我们给它分配  $x_i$  个位置. 所有满足条件的 combinations 可以 represent by:

$$\{y = (x_1, x_2, \dots, x_n) \in \mathbb{Z}_{\geq 0}^n : x_1 + x_2 + \dots + x_n = k\}$$

到这里我们想到一个经典的问题: stars and bars. 即: 把  $k$  个星星分成  $n$  个组, 每个组至少有 1 个星星. 这个问题等价于: 把  $k$  个星星和  $n-1$  个隔板排成一排, 然后选择  $n-1$  个隔板的位置.

问题是: 我们这里, 一个组可以有 0 个 stars; 但是这是小问题. 因为我们可以 set  $x'_i = x_i + 1$ , 问题等价转化为:

$$\{y = (x'_1, x'_2, \dots, x'_n) \in \mathbb{Z}_{\geq 1}^n : x'_1 + x'_2 + \dots + x'_n = k+n\}$$

这就强制每个组至少有一个 star, 于是可以使用 stars and bars 的方法来解决. 即: 用  $n-1$  个隔板隔开  $k+n$  个星星 (有  $k+n-1$  个空档). 因而, 满足条件的 combinations 的数量为:

$$\binom{k+n-1}{n-1} = \binom{k+n-1}{k}$$

**Example 1.5** 有 5 种口味的 ice creams. 一个人随机选择 20 个 scoops. 求: 每种口味至少被选中一次的概率.

**Sol.** 即从 5 种口味中选取 20 个 combinations with repetition. 于是 sample space 的大小:  $\binom{25-1}{20}$ .

而满足条件的 combinations: 即每种口味我们都预选一个. 然后再从 5 种口味中选取 15 个 combinations with repetition.

$$\mathbb{P}(\text{each flavor is selected at least once}) = \frac{\binom{19}{15}}{\binom{24}{20}} = \frac{\binom{24}{20}}{\binom{24}{20}}$$

## 1.4 inclusion-exclusion principle

### Proposition 1.5 (inclusion-exclusion principle)

如果  $\Omega$  是一个 finite measure space, 那么对于任意  $A_1, \dots, A_n \subseteq \Omega$  的, 有:

$$|\cup_{i=1}^n A_i| = \sum_{i=1}^n |A_i| - \sum_{i < j} |A_i \cap A_j| + \sum_{i < j < k} |A_i \cap A_j \cap A_k| - \dots + (-1)^{n+1} |A_1 \cap \dots \cap A_n|.$$



**Remark** 这里的  $\sum_{i < j}$  等 index 意思就是考虑所有可能的 combinations, 不考虑顺序, 和下标的顺序没有关系. 比如一共有三个集合  $A_1, A_2, A_3$ , 那么  $\sum_{i < j} |A_i \cap A_j|$  就是  $A_1 \cap A_2, A_1 \cap A_3, A_2 \cap A_3$  的并集. 这个定理是 countable additivity 的直接推广.

### problems

**Example 1.6** (Divisibility) 令  $n \in \mathbb{N}$ , 我们随机取一个  $x \in \{1, 2, \dots, n\}$ , 求  $x$  is divisible by 2 or 3 or 5 的概率.

**Sol.** 令  $A_2, A_3, A_5$  为  $x$  是 2, 3, 5 的倍数的 events. 即:

$$A_i = \{k \in \{1, 2, \dots, n\} \mid k \text{ is divisible by } i\}$$

于是我们要计算的是:

$$\begin{aligned} \mathbb{P}(A_2 \cup A_3 \cup A_5) &= \mathbb{P}(A_2) + \mathbb{P}(A_3) + \mathbb{P}(A_5) - \mathbb{P}(A_2 \cap A_3) - \mathbb{P}(A_2 \cap A_5) - \mathbb{P}(A_3 \cap A_5) + \mathbb{P}(A_2 \cap A_3 \cap A_5) \\ &= \frac{\lfloor \frac{n}{2} \rfloor + \lfloor \frac{n}{3} \rfloor + \lfloor \frac{n}{5} \rfloor - \lfloor \frac{n}{6} \rfloor - \lfloor \frac{n}{10} \rfloor - \lfloor \frac{n}{15} \rfloor + \lfloor \frac{n}{30} \rfloor}{n} \end{aligned}$$

**Example 1.7** (matching problem) 假设有  $n$  个人参加一个 event, 每个人都上交了一顶帽子; 现在再把帽子随机地发给每个人, 求没有人拿回自己的帽子的概率.

**Sol.** 令  $A_i$  为第  $i$  个人拿回自己的帽子的事件. 则我们要求的概率是:  $\mathbb{P}(\bigcap_{i=1}^n A_i^c)$ .

$$\begin{aligned} \mathbb{P}\left(\bigcap_{i=1}^n A_i^c\right) &= \mathbb{P}\left(\left(\bigcup_{i=1}^n A_i\right)^c\right) \\ &= 1 - \mathbb{P}\left(\bigcup_{i=1}^n A_i\right) \\ &= 1 - \frac{|\bigcup_{i=1}^n A_i|}{n!} \end{aligned}$$

由于:

$$\begin{aligned}
 |\bigcup_{i=1}^n A_i| &= \sum_{i=1}^n |A_i| - \sum_{i < j} |A_i \cap A_j| + \sum_{i < j < k} |A_i \cap A_j \cap A_k| - \dots + (-1)^{n+1} |A_1 \cap \dots \cap A_n| \\
 &= \binom{n}{1}(n-1)! - \binom{n}{2}(n-2)! + \binom{n}{3}(n-3)! - \dots + (-1)^{n+1} \\
 &= n! \sum_{k=1}^n (-1)^{k-1} \frac{1}{k!}
 \end{aligned}$$

我们可以得到:

$$\mathbb{P}\left(\bigcap_{i=1}^n A_i^c\right) = \sum_{k=2}^n \frac{(-1)^k}{k!}$$

**Remark** 当  $k \rightarrow \infty$  的时候, 这个结果趋向于  $e^{-1} \approx 0.367879$ . (By Taylor's expansion of  $e^x$ .)