TA Lecture 01 - Probability and Counting

Oct

School of Information Science and Technology,
ShanghaiTech University

上海科技大学
ShanghaiTech University

Outline

Main Contents Recap

HW Problems

Summary of Counting

Choose k objects out of n objects, the number of possible ways:

	Order Matters	Order Not Matter
with replacement	n^k	$\binom{n+k-1}{k}$
without replacement	$n(n-1)\cdots(n-k+1)$	$\binom{n}{k}$

Multinomial Theorem

Theorem

$$(x_1 + x_2 + \cdots + x_r)^n = \sum_{n_1, n_2, \cdots, n_r \geq 0} \frac{n!}{n_1! n_2! \cdots n_r!} x_1^{n_1} x_2^{n_2} \cdots x_r^{n_r}$$

where $n_1 + n_2 + \cdots + n_r = n$.

Team Captain (Story Proof)

For any positive integers n and k with $k \leq n$,

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

Vandermonde's Identity (Story Proof)

A famous relationship between binomial coefficients, called *Vandermonde's identity*, says that

$$\binom{m+n}{k} = \sum_{j=0}^{k} \binom{m}{j}\binom{n}{k-j}$$

Bose-Einstein Counting

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

Theorem

There are $\binom{n+k-1}{n-1}$ *distinct* nonnegative *integer-valued vectors* $(x_1, x_2, \ldots, x_n)$ *satisfying the equation*

$$x_1 + x_2 + \cdots + x_n = k, x_i \geq 0, i = 1, 2, \ldots, n.$$

Naive Definition of Probability

- Assumption 1: finite sample space
- Assumption 2: all outcomes occur equally likely

Definition

Let A be an event for an experiment with a finite sample space S. The naive probability of A is

$$P_{naive}(A) = \frac{|A|}{|S|} = \frac{\text{number of outcomes favorable to } A}{\text{total number of outcomes in } S}.$$

$\frac{A}{S}$

General Definition of Probability

Definition

A *probability space* consists of a *sample space* S and a *probability function* P which takes an event $A \subseteq S$ as input and returns $P(A)$, a real number between 0 and 1, as output. The function P must satisfy the following axioms:

P

1. $P(\emptyset) = 0$, $P(S) = 1$.
2. If $A_1, A_2, \ldots$ are disjoint events, then

$P(\phi) = 0$ $P(S) = 1$

$$P\left(\bigcup_{j=1}^{\infty} A_j\right) = \sum_{j=1}^{\infty} P(A_j)$$

A_1, b

$P(A_1) + \cdots P(A_n)$

(Saying that these events are disjoint means that they are mutually exclusive: $A_i \cap A_j = \emptyset$ for $i \neq j$.)

$i \neq j$

Bonferroni's Inequality

Theorem

For any n events $A_1, \ldots, A_n$, we have

$$P(A_1 \cap A_2 \cap \cdots \cap A_n) \geq P(A_1) + P(A_2) + \cdots + P(A_n) - (n-1).$$

Boole's Inequality

$A_1 \cup A_2 \cup \cdots \cup A_{\infty} \le P(A_1)$

$+ \cdots + P(A_{\infty})$

Theorem

For any events $A_1, A_2, \ldots,$ we have

$$P\left(\bigcup_{i=1}^{\infty} A_i\right) \le \sum_{i=1}^{\infty} P(A_i).$$

Inclusion-Exclusion Formula

For any events $A_1, \ldots, A_n$:

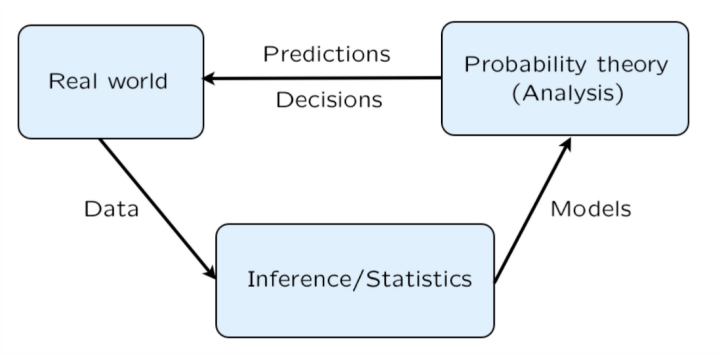
$$P(\bigcup_{i=1}^{n} A_i) = \sum_i P(A_i) - \sum_{i<j} P(A_i \cap A_j) + \sum_{i<j<k} P(A_i \cap A_j \cap A_k)$$
$$+ \ldots + (-1)^{n+1} P(A_1 \cap \cdots \cap A_n).$$

$P(A_1 \cup A_2 \cup A_n)$

Probability and Statistics

A framework for analyzing phenomena with uncertain outcomes:

- Rules for consistent reasoning
- Used for predictions and decisions



Outline

Problem 1

Define $\left\{ \begin{matrix} n \\ \underbrace{k} \end{matrix} \right\}$ as the number of ways to partition $\{1, 2, \ldots, n\}$ into k non-empty subsets, or the number of ways to have n students split up into k groups such that each group has at least one student. For example, $\left\{ \begin{matrix} 4 \\ 2 \end{matrix} \right\} = 7$ because we have the following possibilities:

- $\{1\}, \{2, 3, 4\}$
- $\{2\}, \{1, 3, 4\}$
- $\{3\}, \{1, 2, 4\}$
- $\{4\}, \{1, 2, 3\}$
- $\{1, 2\}, \{3, 4\}$
- $\{1, 3\}, \{2, 4\}$
- $\{1, 4\}, \{2, 3\}$

Problem 1 Continued

Prove the following identities:

(a)

$n \to n+1$

$$\left\{ \begin{matrix} n+1 \\ k \end{matrix} \right\} = \underline{\left\{ \begin{matrix} n \\ k-1 \end{matrix} \right\} + k \left\{ \begin{matrix} n \\ k \end{matrix} \right\}}.$$

Hint: I'm either in a group by myself or I'm not.

Problem 1 Solution

① 新项为新子集

n项只能固定在 k-1子集中 $1 \times \left\{ {n \atop k-1} \right\} = \left\{ {n \atop k-1} \right\}$

② 新项为 $\left\{ {n \atop k} \right\}$ 子集中

k种方法中的任一种 $k \left\{ {n \atop k} \right\} +$

$$\left\{ {n+1 \atop k} \right\} = \left\{ {n \atop k-1} \right\} + k \left\{ {n \atop k} \right\}$$

Problem 1 Countined

(b)

$$\sum_{j=k}^{n} \binom{n}{j} \left\{ \begin{matrix} j \\ k \end{matrix} \right\} = \left\{ \begin{matrix} n+1 \\ k+1 \end{matrix} \right\}.$$

Hint: First decide how many people are not going to be in my group.

Problem 1 Solution

2) n+1 个物品分到 k+1 组中

一个物品被分到了 (k+1)th 组中

剩 n 个物品

j个物品不在第 (k+1)th 组中 $\Rightarrow \binom{n}{j}$

j个物品分成 k 组中 $\Rightarrow \left\{ {j \atop k} \right\}$

每个群组中，存在物品，$j \geq k$

(k+1)th 组中，存在物品，$j \leq n$

j 的数量

$$\sum_{j=k}^{n} \binom{n}{j} \left\{ {j \atop k} \right\} = \left\{ {n+1 \atop k+1} \right\}$$

Problem 2

A *norepeatword* is a sequence of at least one (and possibly all) of the usual 26 letters $\mathrm{a}, \mathrm{b}, \mathrm{c}, \ldots, \mathrm{z}$, with repetitions not allowed. For example, “course” is a norepeatword, but “statistics” is not. Order matters, e.g., “course” is not the same as “source”. A norepeatword is chosen randomly, with all norepeatwords equally likely. Show that the probability that it uses all 26 letters is very close to $1/e$.

Problem 2 Solution

A: norepeatword with 26 letters

$$A = 26! \qquad 26 \; \underline{25} \; \cdots \; 1$$

S: 所有 norepeatword

$$S = \sum_{n=1}^{26} \binom{26}{n} A_n^n$$

$$P = \frac{26!}{\sum_{n=1}^{26} \binom{26}{n} A_n^n} = \frac{1}{\sum_{n=0}^{25} \frac{1}{n!}} \qquad P \approx \frac{1}{e}$$

由 $e^x = \sum_{n=0}^{\infty} \frac{x^n}{n!}$

$x = 1$

$$e = \sum_{n=0}^{\infty} \frac{1}{n!} \approx \sum_{n=0}^{25} \frac{1}{n!}$$

Problem 3

Given $n \geq 2$ numbers $(a_1, a_2, \ldots, a_n)$ with no repetitions, a bootstrap sample is a sequence $(x_1, x_2, \ldots, x_n)$ formed from the a_j's by sampling with replacement with equal probabilities. Bootstrap samples arise in a widely used statistical method known as the bootstrap. For example, if $n = 2$ and $(a_1, a_2) = (3, 1)$, then the possible bootstrap samples are $(3, 3), (3, 1), (1, 3)$, and $(1, 1)$.

$l = n$

(a) How many possible bootstrap samples are there for $\underline{(a_1, \ldots, a_n)}$?

$l = n$

n^n

$\underbrace{n \times n \cdots n}_{n}$

n^n

Problem 3 Solution

Problem 3 Continued

(b) How many possible bootstrap samples are there for $(a_1, \ldots, a_n)$, if order does not matter (in the sense that it only matters how many times each a_j was chosen, not the order in which they were chosen)?

Problem 3 Solution

b) $\cancel{a_j}$ a_j t_j

a_1 t_1

$t_1 + \cdots t_n = n$

a_n t_n

$$\binom{n+n-1}{n-1} = \binom{2n-1}{n-1}$$

Problem 3 Continued

(c) One random bootstrap sample is chosen (by sampling from $a_1, \ldots, a_n$ with replacement, as described above). Show that not all unordered bootstrap samples (in the sense of (b)) are equally likely. Find an unordered bootstrap sample $\mathbf{b}_1$ that is as likely as possible, and an unordered bootstrap sample $\mathbf{b}_2$ that is as unlikely as possible. Let p_1 be the probability of getting $\mathbf{b}_1$ and p_2 be the probability of getting $\mathbf{b}_2$ (so p_i is the probability of getting the specific unordered bootstrap sample $\mathbf{b}_i$). What is p_1/p_2 ? What is the ratio of the probability of getting an unordered bootstrap sample whose probability is p_1 to the probability of getting an unordered sample whose probability is p_2 ?

c) b_1：具有不同元素样本 $\quad p_1 = \dfrac{n!}{n^n}$

b_2：具有相同的 $\quad p_2 = \dfrac{1}{n^n}$

$$\frac{p_1}{p_2} = n!$$

b_1：n个数字的 bootstrap样本

$$p_1 = \frac{n}{n^n}$$

$$\frac{p_1}{p_2} = (n-1)!$$

Problem 4

You get a stick and break it randomly into three pieces. What is the probability that you can make a triangle using such three pieces?

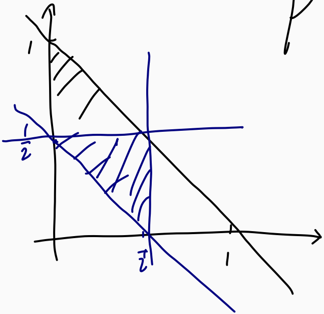
Problem 4 Solution

$$l = 1$$

$$x,\ y,\ 1-(x+y)$$

$$\begin{cases} x+y > 1-(x+y) \\ x+1-(x+y) > y \\ y+1-(x+y) > x \end{cases} \Longrightarrow \begin{cases} x+y > \frac{1}{2} \\ x < \frac{1}{2} \\ y < \frac{1}{2} \end{cases}$$

$$\begin{cases} 0 < x < 1 \\ 0 < y < 1 \\ x+y > 0 \\ x+y < 1 \end{cases}$$



$$P = \frac{\frac{1}{2} \times \frac{1}{2} \times \frac{1}{2}}{1 \times 1 \times \frac{1}{2}} = \frac{1}{4}$$

Problem 5

In the birthday problem, we assumed that all 365 days of the year are equally likely (and excluded February 29). In reality, some days are slightly more likely as birthdays than others. For example, scientists have long struggled to understand why more babies are born 9 months after a holiday. Let $p = (p_1, p_2, ..., p_{365})$ be the vector of birthday probabilities, with p_j the probability of being born on the jth day of the year (February 29 is still excluded, with no offense intended to Leap Dayers). The kth elementary symmetric polynomial in the variables $x_1, ..., x_n$ is defined by

$$e_k(x_1, \ldots, x_n) = \sum_{1 \le j_1 < j_2 < \cdots < j_k \le n} x_{j_1} \ldots x_{j_k}.$$

Problem 5 Continued

This just says to add up all of the $\binom{n}{k}$ terms we can get by choosing and multiplying k of the variables. For example, $e_1(x_1, x_2, x_3) = x_1 + x_2 + x_3$, $e_2(x_1, x_2, x_3) = x_1x_2 + x_1x_3 + x_2x_3$, and $e_3(x_1, x_2, x_3) = x_1x_2x_3$. Now let $k \geq 2$ be the number of people.

(a) Find a simple expression for the probability that there is at least one birthday match, in terms of $\mathbf{p}$ and an elementary symmetric polynomial.

p -

Problem 5 Solution

a) 无人同一天生日

$e_k(\vec{p})$：k个不同日子的乘积和，每次有对k个日子进行映射

p(无同一天生日) $= k!\,e_k(\vec{p})$

p(至少有两人同一天生日) $= 1 - k!\,e_k(\vec{p})$

Problem 5 Continued

(b) Explain intuitively why it makes sense that P(at least one birthday match) is minimized when $p_j = \frac{1}{365}$ for all j, by considering simple and extreme cases.

Problem 5 Solution

b) $k=2$

P(至少两人同一天生日) $= 1-2!\,e_2(\vec{p}) = 1-2e_2(\vec{p})$

$= 1-2\sum_{1\le i<j\le n} p_i p_j$

$= \left(\sum_{i=1}^{365} p_i\right)^2 - 2\sum_{1\le i<j\le n} p_i p_j$

$= \sum_{i=1}^{365} p_i^2 \geq 365 \cdot \left(\frac{\sum_{i=1}^{365} p_i}{365}\right)^2$

$= \frac{1}{365}$

$p_i = \frac{1}{365}$ 成立 $\qquad P_2(p_1, p_n) = \frac{1}{365}$

Problem 5 Continued

(c) The famous arithmetic mean-geometric mean inequality says that for $x, y \geq 0$

$$\frac{x+y}{2} \geq \sqrt{xy}.$$

This inequality follows from adding $4xy$ to both sides of $x^2 - 2xy + y^2 = (x-y)^2 \geq 0$. Define $\mathbf{r} = (r_1, \ldots, r_{365})$ by $r_1 = r_2 = (p_1 + p_2)/2, r_j = p_j$ for $3 \leq j \leq 365$. Using the arithmetic mean-geometric mean bound and the fact, which you should verify, that

$$\begin{aligned} & e_k(x_1, \ldots, x_n) \\ = & x_1 x_2 e_{k-2}(x_3, \ldots, x_n) \\ & + (x_1 + x_2) e_{k-1}(x_3, \ldots, x_n) \\ & + e_k(x_3, \ldots, x_n) \end{aligned}$$

Problem 5 Continued

show that

$$\begin{aligned} &P(\text{at least one birthday match} \mid \mathbf{p}) \\ \geq &P(\text{at least one birthday match} \mid \mathbf{r}) \end{aligned}$$

with strict inequality if $\mathbf{p} \neq \mathbf{r}$, where the given $\mathbf{r}$ notation means that the birthday probabilities are given by $\mathbf{r}$. Using this, show that the value of $\mathbf{p}$ that minimizes the probability of at least one birthday match is given by $p_j = \frac{1}{365}$ for all j.

Problem 5 Solution

3) $x_1x_2e_{k-2}(x_3,\cdots x_n)$：从 $n-2$ 项中选择 x_1,x_2 和其 $k-2$ 项

$(x_1+x_2)e_{k-1}(x_1,\cdots,x_n)$：从 $n-1$ 项中选择 x_1 或 x_2 和其余 $\cancel{k-2}$ $k-1$ 项

$e_k(x_1,\cdots,x_n)$：从 n 项中选 k 项

p' 满足 $p($ $)$ 最小，且 $p'\neq p$，

$\vec{p}$ 使 $r_i'=r_j'=\frac{p_i'+p_j'}{2}$

$p(\quad|\vec{p'})\geq p(\quad\vec{r})$

p 不是最小的

$\vec{p}=(\frac{1}{365},\quad\cdots\quad,\frac{1}{365})$

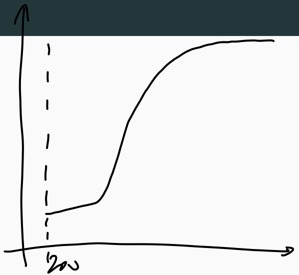
p(两人至少一天同天生日) 概率最小

Problem 6

If each box of a brand of crispy instant noodle contains a coupon, and there are 108 different types of coupons. Given $n \geq 200$, what is the probability that buying n boxes can collect all 108 types of coupons? You also need to plot a figure to show how such probability changes with the increasing value of n. When such probability is no less than 95%, what is the minimum number of n?

Problem 6 Solution

P



n个盒子，108种类型

108^n

n个盒子，划分108个非空集合，$\left\{ {n \atop 108} \right\}$

$$P = \frac{108!\left\{ {n \atop 108} \right\}}{108^n} = \sum_{k=0}^{108} (-1)^k \frac{108!}{k!\,(108-k)!} \cdot \left(\frac{108-k}{108}\right)^n$$

$$\left\{ {n \atop m} \right\} = \frac{1}{m!} \sum_{k=0}^{m} (-1)^k \binom{m}{k} (m-k)^n$$

$$= \sum_{k=0}^{m} (-1)^k \frac{(m-k)^n}{k!\,(m-k)!}$$

$P \geq 95\%$

$n = 823$