TA Lecture 01 - Probability and Counting

March 13-14

School of Information Science and Technology,
ShanghaiTech University

上海科技大学
ShanghaiTech University

Outline

Main Contents Recap

HW Problems

More Exercices

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

e.g. $n=2$

Team Captain (Story Proof)

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

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

$$\binom{n}{1} \qquad \binom{n}{k}$$

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

离散

$\begin{cases} \text{order} \\ \text{replacement} \end{cases}$

Theorem

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

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

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

eg. $x_1 \geq 5 \xrightarrow{\text{convert}} a_1 = x_1 - 5$

integer, $x > 5 \longrightarrow ?$

Naive Definition of Probability

- Assumption 1: finite sample space
- Assumption 2: all outcomes occur equally likely 出现是等概率的 ✓

否则个数 X

Definition

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

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

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:

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

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

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

不是 disjoint? 容斥

独立 & disjoint 关系，

Bonferroni's Inequality

Theorem

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

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

Venn 图

Boole's Inequality

Theorem

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

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

取等 $\Rightarrow$ disjoint

Inclusion-Exclusion Formula

☆☆ 容斥 probably 最难 in chapter 1

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

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

交集是更易刻画的概率

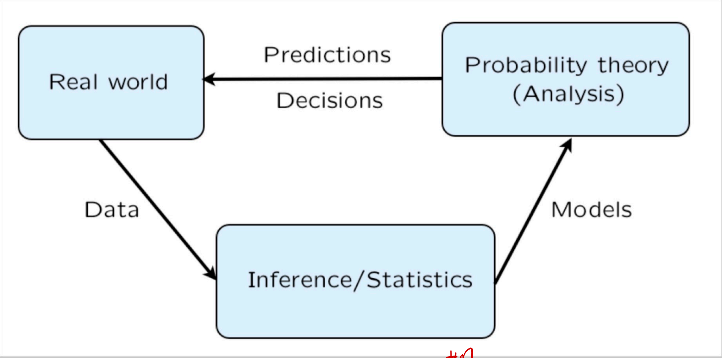
计算量↑ 思维↓

分类讨论

Probability and Statistics

A framework for analyzing phenomena with uncertain outcomes:

- Rules for consistent reasoning
- Used for predictions and decisions



贯穿始终　　概率与统计

Outline

Main Contents Recap

HW Problems

More Exercices

Problem 1

Second Stirling 数

Define $\left\{ \begin{matrix} n \\ k \end{matrix} \right\}$ 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:

自己一组

$$\binom{n}{0}$$

$$\binom{k}{1}$$

(a)

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

Hint: I'm either in a group by myself or I'm not. 没有Hint也要想到

Problem 1 Solution

Problem 1 Countined

(b)

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

剩k个组

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

不和我一组

Problem 1 Solution

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

Sample space: # no repeated word

$= \sum_{k=1}^{26} \#$ · · · of length k

$$\frac{\# \, 126 \cdots}{\sum \# \, 1i \cdots} = \frac{26!}{\sum \binom{26}{k} \cdot k!} = \frac{}{\sum_{i=1}^{26} (n-k)!} \approx \frac{1}{e}$$

$1 \sim 26$

$0 \sim 25$ ✓ $\quad 0 \sim 26$ × $\quad 1 \sim 26$ ×

$0! = 1$

Problem 3

Given $n \geq 2$ numbers $(a_1, a_2, \dots, a_n)$ with no repetitions, a bootstrap sample is a sequence $(x_1, x_2, \dots, x_n)$ 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 $(1,1)$.

Problem 3 Continued

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

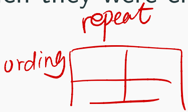
n^n

just definition

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)?



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

$$\Rightarrow k = n$$

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

Problem 3 Solution

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 都出现 possible, and an unordered bootstrap sample $\mathbf{b}_2$ that is as unlikely 只 1种 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 ?

Problem 3 Solution

完全不一样 $\frac{n!}{n^n}$

一一 $\frac{1}{n^n}$ 这一种情况出现的概率

得到其中一个 ✓ 共n个

甚至没有问原因 → intuitively

Problem 4

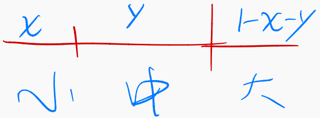
You get a stick and break it randomly into three pieces. What is the probability that you can make a triangle using such three pieces?

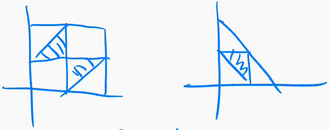
几何概形

线性规划

Problem 4 Solution

$l=1$

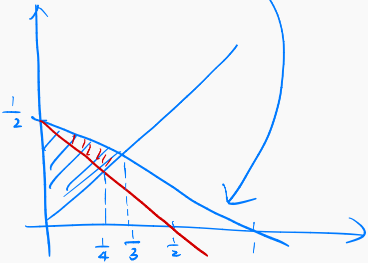




不同定义

$x<y<1-x-y$

$y<\frac{1}{2}-\frac{1}{2}x$



$x+y>1-x-y$

$y>-x+\frac{1}{2}$

$1-\frac{\frac{1}{2}\times\frac{1}{2}\times\frac{1}{4}}{\frac{1}{2}\times\frac{1}{2}\times\frac{1}{3}}=1-\frac{3}{4}=\frac{1}{4}$

Problem 4 Solution

随机切2刀 $\Longrightarrow$ 先切一刀，~~再~~切一刀

长为1

x random sample uniformly on $(0,1)$

y random — — — — — $(0,1-x)$

$x \sim \text{Unif}(0,1)$ $\quad y|x \sim \text{Unif}(0,1-x)$

$x+y > 1-x-y \Longrightarrow x+y > \frac{1}{2}$ $\quad y > \frac{1}{2} - x$

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

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

$$\int_0^{\frac{1}{2}} \int_{\frac{1}{2}-x}^{\frac{1}{2}} \frac{1}{1} \cdot \frac{1}{1-x}\, dy\,dx$$

$$= \int_0^{\frac{1}{2}} \frac{1}{1-x} \left(\frac{1}{2} - \left(\frac{1}{2} - x\right)\right) dx$$

$$= \int_0^{\frac{1}{2}} \frac{x}{1-x}\, dx$$

$$= \int_0^{\frac{1}{2}} \frac{1}{1-x} - 1\, dx$$

$$= -\ln(1-x) - x \Big|_0^{\frac{1}{2}} = \ln 2 + \frac{1}{2}$$

Problem 5

生日问题 ☆ thinking from the opposite.

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

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

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.

Problem 5 Solution

$K > 365$ $\quad P = 1$

$2 \leq k \leq 365$ $\quad$ at least one match $\Leftrightarrow 1 - P$ no match

$= 1 - \underline{k!}\, e_k(\vec{p})$

Problem 5 Continued

不好懂 ⇒ 取极限情况　简单例子描述

(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

eg. $k=2$ P(at least one birthday matches)

$$= 1 - 2e_2(P)$$

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

$$= \sum_{i=1}^{365} p_i^2$$

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

$$= \frac{1}{365}, \quad \forall i \; p_i = \frac{1}{365}$$

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

先证

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

Problem 5 Continued

show that

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

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

$$P(\cdots | p) = 1 - k!\, e_k(p)$$

$$= 1 - k!\,\big(p_1 p_2 e_{k-2}(p_3 \cdots p_n)$$

$$+ (p_1 + p_2) e_{k-1}(p_3 \cdots p_n)$$

$$+ e_k(p_3 \cdots p_n)\big)$$

$$r_1 = r_2 = \frac{p_1 + p_2}{2} \le \frac{(p_1+p_2)^2}{4}$$

$$\ge 1 - k!\,\Big(\frac{(p_1+p_2)^2}{4} e_{k-2}(p_3 \cdots p_n)$$

$$+ (p_1+p_2) e_{k-1}(p_3 \cdots p_n)$$

$$+ e_k(p_3 \cdots p_n)\Big)$$

$r_1 r_2$

$r_1 + r_2$

$$= P(\cdots | r)$$

Problem 6

小浣熊干脆面 集卡

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

$$P(\text{all collect})$$

$$= 1 - \underbrace{P\left(\bigcup_{i=1}^{108} Z_i^{\,n}\right)}_{} \Longrightarrow \sum_{i=1}^{108} (-1)^{i-1} \underbrace{P\left(\bigcap_{i=1}^{108} Z_i\right)}_{}$$

Z_i 第i张没被收集到

$$= 1 - \sum_{i=1}^{108} (-1)^{i-1} \binom{108}{i} \left(\frac{108-i}{108}\right)^n$$

$95\% \Rightarrow 823$

Outline

BH CH1 # 38 (Choose a Right Tool Given Keywords)

Standford

UI 包

BT ⇒ MIT

There are n balls in a jar, labeled with the numbers $1, 2, \cdots, n$. A total of $\boxed{k \text{ balls}}$ are drawn, one by one *with replacement*, to obtain a sequence of numbers.

(a) What is the probability that the sequence obtained is strictly increasing?

$$\frac{\binom{n}{k}}{n^k}$$

≠

1, 2, 3
1, 3, 2
3, 1, 2

order? X

replacement? X

BH CH1 # 38 Solution

order ✗

replacement ✓

(b) What is the probability that the sequence obtained is increasing?

$t_1 + \cdots + t_n = k$

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

BH CH1 # 38 Solution

BH CH1 # 52 (Inclusion-Exclusion)

Alice attends a small college in which each class meets only once a week. She is deciding between 30 non-overlapping classes. There are 6 classes to choose from for each day of the week, Monday through Friday. Trusting in the benevolence 仁慈 of randomness, Alice decides to register for 7 randomly selected classes out of the 30, with all choices equally likely. What is the probability that she will have classes every day, Monday through Friday? (This problem can be done either directly using the naive definition of probability, or using inclusion-exclusion.)

BH CH1 # 52 Solution

①

天	课
2	2
3	1
1	3
4	1

$$\frac{\binom{5}{2}\cdot\binom{6}{2}\cdot 6^3 + \binom{5}{1}\cdot\binom{6}{3}\cdot 6^4}{\binom{30}{7}}$$

$$\frac{144}{377} \approx 0.302$$

② 容斥，B_i：第i天无课

$P(A) = 1 - P(A^c)$

A：天天有课

$$P(A^c) = P\left(\bigcup_{i=1}^{5} B_i\right) = \sum_{i=1}^{5} (-1)^{i-1} P\left(\bigcap_{i=1}^{5} B_i\right)$$

$P(B_1) = \dfrac{\binom{24}{7}}{\binom{30}{7}}$

$P(B_1 \cap B_2) = \dfrac{\binom{18}{7}}{\binom{30}{7}}$ $\quad P(B_1 \cap \cdot B_3) = \dfrac{\binom{12}{7}}{\binom{30}{7}}$ ....