MATH 180A: INTRO TO PROBABILITY (FOR DATA SCIENCE)

www.math.ucsd.edu/~tkemp/180A

Today: § 1.2 – 1.3

Next: § 1.4 – 2.1

HW.0 due TONIGHT!

HW.1 due FRIDAY, 10/04

Lab.1 due MONDAY, 10/07

Sections / Labs:

B01, B02, B03, B04 : ~~CENTR 207~~ → ERC 117

B05, B06 : ~~HSS 1128A~~ → CSB 115

Mathematical thinking

Probability Space $(\Omega, \mathcal{F}, \mathbb{P})$

sample space → Ω, events → $\mathcal{F}$, probability measure → $\mathbb{P}$

Determined by the experiment being modeled

Key property: if $\underbrace{A_1, A_2, A_3, \ldots}_{\in \mathcal{F}}$ are disjoint $(A_i \cap A_j = \phi)$ $i \neq j$

then $\mathbb{P}(A_1 \cup A_2 \cup A_3 \cup \cdots) = \sum_j \mathbb{P}(A_j)$

Important special case: if $\# \Omega < \infty$ then

$$\Omega = \{\omega_1, \omega_2, \omega_3, \ldots, \omega_n\}$$

the singleton sets $\{\omega_1\}, \{\omega_2\}, \ldots, \{\omega_n\}$ are disjoint

$\therefore \mathbb{P}(A) = \{a_1\} \cup \{a_2\} \cup \cdots \cup \{a_r\}$ $\qquad \mathbb{P}(A) = \sum_{j=1}^{r} \mathbb{P}(\{a_j\})$

$A = \{a_1, a_2, \ldots, a_r\}$

Uniform Probability Measure (Sampling)

1.2

If Ω is finite, the **uniform probability measure** is defined by:

For each $\omega \in \Omega$, $\mathbb{P}(\{\omega\}) = \dfrac{1}{\#\Omega}$

$\Rightarrow$ For any event A, $\mathbb{P}(A) = \dfrac{\#A}{\#\Omega}$

This means calculating probabilities in such models is tantamount to <u>Counting</u>.

E.g. A fair die is cast 2 times. What is the probability that the sum is 4?

$\Omega = \{ (i,j) : 1 \le i, j \le 6 \}$ $\quad \#\Omega = 36$

$A = \{ (1,3), (2,2), (3,1) \}$ $\quad \#A = 3$

(sum is 4)

$\mathbb{P}(A) = \dfrac{3}{36} = \dfrac{1}{12} \simeq 8.3\%$

E.g. A fair coin is tossed 3 times.

$A = \{\text{at least two tails}\}$

$B = \{\text{exactly two tails}\}$

$\Omega = \{HHH, HHT, \cdots, TTT\}$ $\quad \#\Omega = 8$

$= \{(i, j, k) : \quad i, j, k \in \{T, H\}\}$

$A = \{TTH, THT, HTT, TTT\}$ $\quad \mathbb{P}(A) = \frac{\#A}{8} = \frac{4}{8} = \frac{1}{2}$

$B = \{TTH, THT, HTT\}$ $\quad \mathbb{P}(B) = \frac{3}{8}.$

THINK PAIR SHARE

There are 10 people on a committee.
How many different ways are there to select a subcommittee of 4 people?

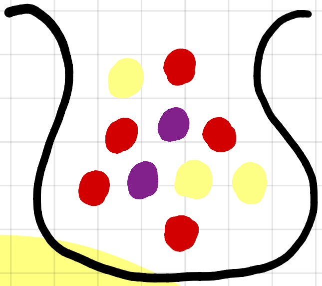
(a) $10 \cdot 10 \cdot 10 \cdot 10 = 10^4 = 10{,}000$

(b) $10 \cdot 9 \cdot 8 \cdot 7 = 5{,}040$

(c) $\binom{10}{4} = \dfrac{10 \cdot 9 \cdot 8 \cdot 7}{4!} = 210$

(d) $\dfrac{10!}{4!} = 151{,}200$

Combinatorics



A collection of labeled balls $\{1, 2, \dots, n\}$ are in an urn. k are taken out one by one

* with replacement, or
* without replacement $(k \leq n)$

$b_i \neq b_j$ if $i \neq j$

They are lined up

* in the order they came out, or
* disregarding order.

$$\Omega = \{(b_1, \dots, b_k) : 1 \leq b_j \leq n\} = \{1, \dots, n\}^k$$

$$\Omega = \{\{b_1, \dots, b_k\} : \}$$

How many ways?

* with replacement : n^k
* without : $\begin{cases} \text{\& ordered: } n \cdot (n-1) \cdot (n-2) \cdots (n-k+1) \\ \text{not: } \dfrac{n(n-1)(n-2)\cdots(n-k+1)}{k!} = \dfrac{n!}{k!\,(n-k)!} = \dbinom{n}{k} \end{cases}$

when $k=n$
$\Downarrow$ $n!$

Sampling with Replacement

Toss a fair coin n times; record a statistic observing # H vs. # T.

E.g. $n = 10$, $\mathbb{P}\{\text{odd rolls are all H}\}$.

$$\Omega = \{(c_1, c_2, c_3, \ldots, c_{10}) : \text{“}\forall_j\text{”}\ c_j \in \{H, T\}\}$$

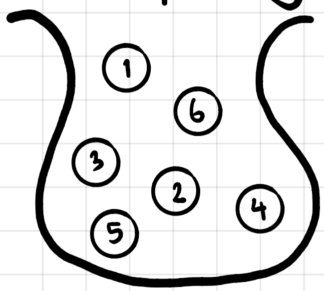
$$\#\Omega = 2^{10}$$

$$A = \{c_1 = c_3 = c_5 = c_7 = c_9 = H\}$$

$$= \{(H, *, H, *, H, * H, *, H, *)\} \qquad \#A = 2^5$$

$$\mathbb{P}(A) = \frac{2^5}{2^{10}} = \frac{1}{2^5} = \frac{1}{32}.$$

Sampling without Replacement (order matters)



There are 6 labeled balls in an urn. 3 are removed in sequence (without replacement), and lined up in order.

What is the probability that the first two are $(3,6)$?

$$\Omega = \{ (b_1, b_2, b_3) : \quad 1 \le b_j \le 6, \quad b_1 \neq b_2, \ b_2 \neq b_3, \ b_1 \neq b_3 \}$$

$$\# \Omega = 6 \cdot 5 \cdot 4 = 120$$

$$A = \{ (3, 6, \underbrace{*}_{\in \{1,2,4,5\}}) \} \qquad \# A = 4$$

$$\mathbb{P}(A) = \frac{4}{120} = \frac{1}{30}.$$