Permutations & Combinatorics

How many 1 letter words with letter A?

A = 1

How many 2 letter words with letters A & B?

AB BA = 2

How many 3 letter words with letters A, B & C?

ABC ACB

BAC BCA = 6

CAB CBA

4 letters: 24 words

1. 4 choices for 1^{st} letter, after this choice
2. 3 choices for 2^{nd} letter, after this choice
3. 2 choices for 3^{rd} letter
4. 1 choice for 4^{th} letter

For a N-letter word with N unique letters, there are N! possible permutations

$$N! = N \cdot (N-1) \cdot (N-2) \cdot \ldots$$

We can write this as "N choose N"

$$_nP_n = \binom{n}{n}$$

Suppose we want to know how many 2-letter words we can make with 4 letters

AB AC AD

BA BC BD = 12

CA CB CD

DA DB DC

This is a somewhat different set of permutations and we say that $4 \cdot 3 = \frac{4!}{2!} = {}_4P_2$

In general: $${}_NP_k = \frac{N!}{(N-k)!}$$

All of these examples assume ordering is important that is: ABCD $\neq$ BACD $\neq$ ABDC

Suppose all we care about is how many combinations instead of permutations.
That is, if I reach into a box and pull out letters and don't care what order they are in:

$${}_NC_k = \frac{\overbrace{{}_NP_k}^{\text{ways of picking the letters}}}{\underbrace{{}_kP_k}_{\text{ways of arranging the letters we've picked}}} = \frac{N!}{(N-k)!\ k!}$$

How many ways can we divide up 5 molecules between 2 flasks. (that is 5 objects into groups of 2 & 3)

$${}_5C_3 = \frac{5!}{2!\ 3!} = \frac{5\cdot4\cdot3\cdot2\cdot1}{(3\cdot2\cdot1)(2\cdot1)}$$

$$= 10$$

abc : de	abd : ce	abe : cd
acd : be	ace : bd	ade : bc
bcd : ae	bce : ad	bde : ac
cde : ab		

Your book concentrates on a 2-box partitioning scheme where $N_1 + N_2 = N$

$$\therefore \quad {}_N C_{N_1} = \frac{N!}{(N-N_1)!\,N_1!} = \frac{N!}{N_2!\,N_1!}$$

This isn't quite the whole story.

Consider a cheater coin that lands on heads 60% of the time

$$P_H = 0.6$$
$$P_T = 1 - P_H = 0.4$$

		Cheater	Fair Coin
For 2 flips:	$P_{HH} = (P_H)^2 =$	0.36	0.25
	$P_{HT} = P_H(1-P_H) =$	0.24	0.25
	$P_{TH} = (1-P_H)P_H =$	0.24	0.25
	$P_{TT} = (1-P_H)^2 =$	0.16	0.25

If we want to know the probability of n_H heads in N coin tosses (in any order)

$$P(n_H, N) = \underbrace{P_H^{\,n_H}\,(1-P_H)^{N-n_H}}_{\text{probability of a specific outcome}} \cdot \underbrace{{}_N C_{n_H}}_{\text{\# of combinations of the outcomes that match the } n_H \text{ criteria.}}$$

$$P(n, N) = P^n (1-P)^{N-n} \frac{N!}{n!\,(N-n)!}$$

binomial distribution.

For more general cases, we use the <u>multinomial</u> distribution:

$$p(n_1, n_2, n_3 \dots n_t, N) = p_1^{n_1}\, p_2^{n_2} \dots p_t^{n_t} \left(\frac{N!}{n_1!\, n_2! \dots n_t!} \right)$$

probability of arranging

Lottery example:

Powerball : 5 white balls chosen out of numbered pool (1-59)

1 red ball chosen out of pool (1-35)

Total combinations:

$$_{59}C_5 \cdot {}_{35}C_1$$

$$= \frac{59!}{5!\,54!} \cdot \frac{35!}{1!\,34!} = 5,006,386 \times 35$$

$$= 175,223,510$$

why no $p_1^{n_1} \dots$? because these are all <u>unique</u> outcomes <u>no</u> winning combos with 2 balls numbered 5

Expectation values:

Consider that each powerball ticket costs \$2

What are your expected winnings for buying a ticket when the jackpot is \$540,000,000?

$$EV = \frac{\$540M}{175,223,510} - \$2 = \$1.08$$

Why isn't that correct?
Nearly 1.5 billion tickets were played; and if multiple people land on a combination the prize is split.

$$\frac{1.5 \times 10^9}{175223510} = 8.56 \text{ players / combination}$$

$$EV = \$540M \times \frac{1}{175{,}223{,}570} \cdot \frac{1 \text{ combination}}{8.56 \text{ players}} - \$2$$

$$EV = -\$1.64$$

(still not quite right.
Lump sum cash is 62% of jackpot
Indiana tax is 3.4% of payout)

$$EV = \$540M\,(0.62)(0.966)\frac{1}{175{,}223{,}510}\cdot\frac{1}{8.56} - \$2$$

$$EV = -\$1.78$$ ← still a bad bet

70% of players use the computer
30% pick meaningful numbers to them (usually dates)

How can we increase EV into positive territory?

Entropy

Suppose we have 2 flasks & 4 molecules

If we put all of the molecules in the left flask, there's only one way to do this:

12 34	

← one combination $(W = 1)$

If we are interested in the "macro state" which has 2 molecules in each flask:

12	34
13	24
14	23
23	14
24	13
34	12

6 combinations $= {}_4C_2 = \dfrac{4!}{2!\,2!}$

or "microstates"

Last time we defined the entropy of a macrostate as:

$$S = k_B \ln W$$

number of microstates contributing to that macrostate

Now, suppose we want to know ΔS for the expansion of these molecules into both flasks:

$$\Delta S = S_{2\&2} - S_{4\&0}$$

$$= k_B \ln 6 - k_B \ln 1 = k_B \ln 6$$

$$\Delta S = 1.79\ k_B$$

Extensivity

What was the probability of the microstate with all molecules on left?

total # of possible combinations:

$$\frac{4!}{4!\,0!} + \frac{4!}{5!\,1!} + \frac{4!}{2!\,2!} + \frac{4!}{1!\,3!} + \frac{4!}{0!\,4!}$$

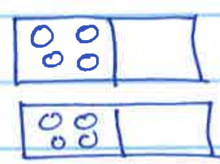
all on right 1 on left 2 on left 3 on left 4 on left

$$1 + 4 + 6 + 4 + 1$$

$$N_{combinations} = 16$$

$$P_{all\ on\ left} = \frac{1}{16} \qquad P_{2\&2} = \frac{6}{16} = \frac{3}{8}$$

Now, suppose we have an exact replica of the 2 flask problem



System A

system B

We want the entropy to be extensive, so

$$S_{A+B} = S_A + S_B$$

However, probabilities are multiplicative

$$P_{\substack{all\ molecules \\ on\ left}} = P_{\substack{molecules \\ in\ A\ on \\ left}} \times P_{\substack{molecules \\ in\ B\ on \\ left}}$$

$$= \frac{1}{16} \cdot \frac{1}{16} = \frac{1}{256}$$

$$S_{A+B} = S_A + S_B$$

$$P_i^{A+B} = P_i^A \; P_i^B$$

Boltzmann noticed that the function relating the probabilities to the entropy had to have a form:

$$S_A = -k \sum_i P_i^A \ln P_i^A$$

in order to be extensive $\ln P^A P^B = \ln P^A + \ln P^B$

So this is then the more general entropy expression:

$$S = -k_B \sum_i P_i \ln P_i$$

probability of microstate i

$\sum$ sum over microstates i contributing to macrostate

when all of the microstates are equally likely, we get:

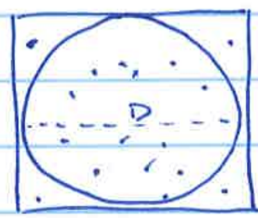
$$S = +k_B \ln W$$

Monte Carlo

Suppose we were stuck on a desert island and forgot the value of π.

Here's a way we might estimate it:

1) Draw a square in the sand:



2) Inscribe a circle of diameter equal to the length of the square

3) Wait for a light rainfall

How this

4) Count the number of drops landing in the circle & in the square.

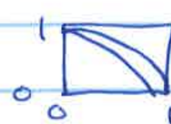
$$A_{circle} = \pi \left(\frac{D}{2}\right)^2$$

$$A_{square} = D^2$$

$$\frac{N_{\text{hits in circle}}}{N_{\text{hits in square}}} = \frac{\pi D^2/4}{D^2} = \frac{\pi}{4}$$

$$\therefore \quad \pi = 4 * \frac{N_{circle}}{N_{square}}$$

(The modern version uses:



← why?

1) Pick 2 random numbers on interval [0, 1]
   call them x & y
2) if $x^2 + y^2 < 1$, $N_{circle} = N_{circle} + 1$
3) $N_{trials} = N_{trials} + 1$
4) If $N_{trials} < 1000$, go to 1)

$$\pi = 4 \frac{N_{circle}}{N_{trials}}$$

These are both versions of a Monte Carlo experiment (one which uses an inherently random process to estimate an integral).

The original idea was called Buffon's theorem (1777)

← Equally spaced parallel lines drawn on floor

Drop needles of length ℓ on the floor

The probability a needle crosses a line

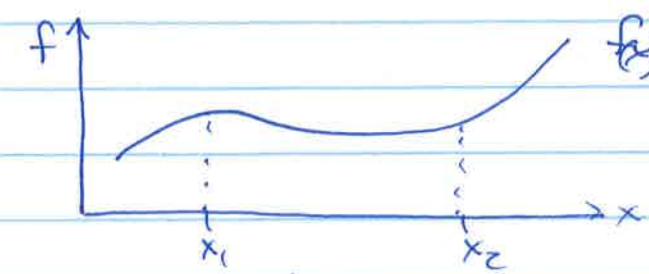
$$P_{cross} = \frac{2\ell}{\pi d} \quad \leftarrow \text{Buffon's theorem}$$

In 1901 Mario Lazzarini did the experiment (he said) with a 2.5 cm needle on a 3cm grid. He got 1808 crossings out of 3408 trials.

$$\pi = 3.1415929$$

We're now suspicious that this was a case of mathematical fraud, however. Lazzarini knew of $\frac{355}{113}$, a rational estimate of π from Tsu Ch'ung-Chi known since 500 AD

Monte Carlo estimates of integrals



$$I = \int_{f_2}^{f_1} f(x)\,dx$$

One way to do this is to introduce a probabilistic measure $p(x)$ ← some random way of sampling x with probability p on this interval.

$$I = \int_{x_1}^{x_2} f(x)\,\frac{p(x)}{p(x)}\,dx$$

$$= \int_{x_1}^{x_2} \frac{f(x)}{p(x)}\,p(x)\,dx$$

$$= \left\langle \frac{f(x)}{p(x)} \right\rangle$$ ← an <u>average</u> of the random samples.

Suppose $p(x)$ is the uniform distribution on $[x_1, x_2]$

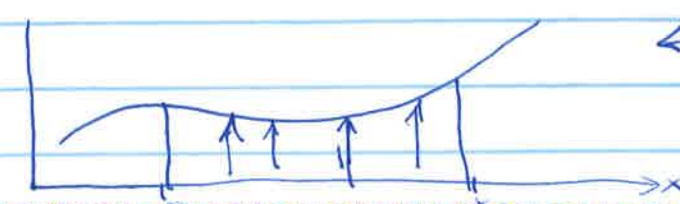
$$p(x) = \frac{1}{x_2 - x_1} \qquad x_1 \le x \le x_2$$

A monte carlo approach would involve choosing <u>random</u> points on this interval:

$$I = \frac{1}{N_{trials}} \sum_{i=1}^{N_{trial}} \frac{f(x_i)}{p(x_i)}$$

$$= \frac{1}{N_{trials}} \sum_{i=1}^{N_{trial}} \frac{f(x_i)}{\frac{1}{x_2 - x_1}} =$$

$$I = \frac{x_2 - x_1}{N_{trials}} \sum_{i=1}^{N_{trial}} f(x_i)$$

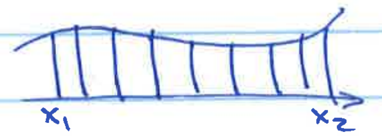


← Shooting random points between x_1 & x_2 and accumulating the sum of the function there.

This works remarkably well in <u>higher</u> dimensions

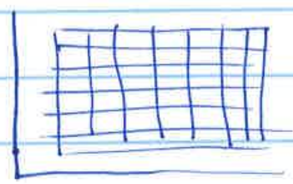
why?

To compute an integral on a '<u>grid</u>'



$$I = \sum_{i=1}^{N_x} f(x_i) \underbrace{\frac{x_2 - x_1}{N_x}}_{\Delta x}$$

In 2D:



$$I = \sum_{i=1}^{N_x} \sum_{j=1}^{N_y} f(x_i, y_j) \Delta x \Delta y$$

Now we need $N_x \times N_y$ <u>samples</u>

If $N_x \approx 20$, and $N_y \approx 20$

$N_x N_y \approx 400$ $\qquad$ $N_x N_y N_z \approx 8000$

Monte Carlo integration <u>converges</u> very quickly

$O\left(\frac{1}{\sqrt{N}}\right)$

so the number of samples required is actually quite <u>small</u> compared to a <u>grid</u> integration.

There are even a number of tricks out there to make sure MC samples the <u>most important</u> parts of an integral. (i.e. where it has the most variation.)