Lecture #23
2009

Probability and statistics

We usually talk about "probability" and "random events" when we lack information about the event and its outcome. A <u>random</u> event does not depend on initial conditions (or we don't know the initial conditions).

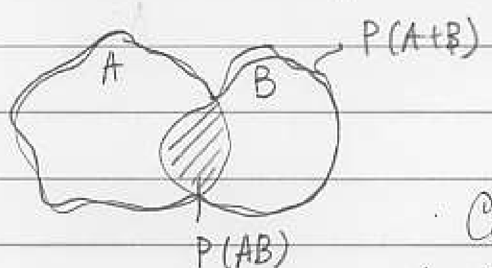
Sample space ⟶ the combination of all possible <u>mutually</u> <u>exclusive</u> events.

To assign a probability to a particular outcome, we have to repeat the measurement many times

$$P_A = \frac{\text{number of outcomes } A}{\text{total number of events.}}$$

If A, B, C are <u>all</u> the possible mutually exclusive outcomes, then $P_A + P_B + P_C = 1$; each $0 \le P_A, P_B, P_C \le 1$

Non-mutually exclusive events



$P(A+B) = P(A \text{ or } B) = P(A) + P(B) - P(AB)$

$P(AB) = P(A \text{ and } B)$

Conditional probability (what is the probability of B happening is A happens)

$$P(A \cdot B) = P_A \cdot P_A(B) \Rightarrow P_A(B) = \frac{P(A \cdot B)}{P(A)}$$

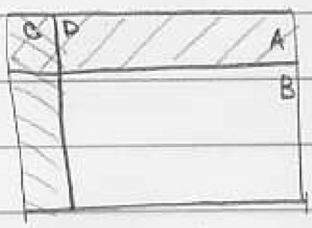
If two events are <u>mutually exclusive</u>

$P(A \cdot B) = 0 \quad P_A(B) = 0$

If two event are <u>independent</u> $P(A \cdot B) = P(A) P(B)$

Example: In the university X 95% of entering class graduates, and 5% drops out. 97% of graduates entered with high SAT, while 80% of drop-outs entered with low SAT. What is the probability of a person with <u>low</u> SAT to graduate?

Sample space



Mutually exclusive outcomes:

[A - low SAT | C - graduate]
[B - high SAT | D - drop out]

$P_A + P_B = 1$ $\qquad$ $P_C = P_D = 1$

We know: $P_C = 95\%$, $P_D = 5\%$

$P_C(B) = 97\%$, $P_D(A) = 80\%$

[thus $P_C(A) = 3\%$, $P_D(B) = 20\%$]

Let's first figure out what's the probability of an entering student to have high or low SAT

$$P_A = P_C \cdot P_C(A) + P_D \cdot P_D(A) = 0.95 \cdot 0.03 + 0.8 \cdot 0.05 = 0.0685 \quad (6.85\%)$$

$$P_B = P_C \cdot P_C(B) + P_D \cdot P_D(B) = 0.95 \cdot 0.97 + 0.05 \cdot 0.2 = 0.9315 \quad (93.15\%)$$

The probability that the person enters with high / low SAT <u>and</u> graduates

$$P(A \cdot C) = P_C \cdot P_C(A) = 0.95 \cdot 0.03 = 0.0285$$

$$P(B \cdot C) = P_C \cdot P_C(B) = 0.95 \cdot 0.97 = 0.9215$$

Thus $$P_A(C) = \frac{P(A \cdot C)}{P(A)} = \frac{0.0285}{0.0685} \approx 42\%$$

$$P_B(C) = \frac{P(B \cdot C)}{P(B)} = \frac{0.9215}{0.9315} \approx 99\%$$

Permutations and Combinations

It is a common problem in statistics is how to distribute k particles in n boxes. Or, in thermodinamics, how to distribute k particles among n energy levels or degrees of freedom. In this case the statistics of the particles is crucial

1. Maxwell-Boltzman statistics — distinguishable particles, unlimited possible occupation for each state.

   For each particle — n possible positions

   For k particles — n^k possible combinations

2. Fermi-Dirac statistics — indistinguishable particles, one particle per site

$$\overbrace{\square \quad \square \; . \; \square \; .\;.\;.\;. \quad \square}^{n \text{ sites}}$$

n possible positions for particle #1
$(n-1)$ — " — for — " — #2
. . . . . .
$(n-k+2)$ — " — for — " — #k-1
$(n-k+1)$ — " — for — " — #k

\# of permutations $P(n,k) = n(n-1)\ldots(n-k+2)(n-k+1) = \frac{n!}{(n-k)!}$

If sites are also indistinguishabel

\# of combinations $C(n,k) = \frac{P(n,k)}{k!} = \frac{n!}{(n-k)!\,k!}$

It is not surprizing that $C(n,k) = \binom{n}{k}$ – binomial coefficient,

$(a+b)^n = \underbrace{(a+b)(a+b)\ldots(a+b)}_{n \text{ sites}} = \ldots a^k b^{n-k}$. [# of combinations to pick k "a"s and $n-k$ "b"s]

Random variables and propability functions

For many processes we can find that the number of combinations to put k particles to n boxes is

$$C_{BE} = C(n+k-1, k) = \frac{(n+k-1)!}{(n-1)!\,k!}$$

We can (sort of) prove it using induction

1 step: k particle, 2 boxes

$\square$	$\square$	
k	0	$(k+1)$ possible combinations
$k-1$	1	
$k-2$	2	$C(k+1,k) = k+1$
…		
2	$k-2$	
1	$k-1$	
0	k	

2 step k particles, 3 boxes

$\square$	$\square\ \square$	
k	0	1 combination
$k-1$	1	$1+1$ combination
$k-2$	2	$2+1$ — " —
…		
2	$k-2$	$k-1$ — " —
1	$k-1$	k — " —
0	k	$k+1$ — " —

Total number of combinations is

$$\underbrace{1+2+\ldots+(k+1)}_{k+1 \text{ terms}} = \frac{(k+1)(k+2)}{2} = \frac{(k+2)!}{k!\cdot 2!} = C(k+2, k)$$

For k particles in the n boxes

$\square$	$\overbrace{\square \cdots \square}^{n-1}$		
k	0	$n'=n-1$, $k=0$	$C(n-1,0) = 1$
$k-1$	1	$n'=n-1$, $k'=1$	$C(n-1,1) = (n-1)$ combinations
$k-2$	2	$n'=n-1$, $k'=2$	$C(n,2)$
$\cdots$	$\cdots$		
1	$k-1$	$n'=n-1$, $k'=k-1$	$C(n+k-3, k-1)$
0	k	$n'=n-1$, $k'=k$	$C(n+k-2, k)$

Total number of combinations

$$\sum_{i=0}^{k} C(n+i-2, i) = \sum_{i=0}^{k} \frac{(n+i-2)!}{(n-2)!\, i!} =$$

Using $$\sum_{i=0}^{k} \frac{(i+a)!}{i!\, a!} = \frac{(k+a+1)!}{k!\,(a+1)!}$$

$$= \{a = n-2\} = \frac{(k+n-1)!}{k!\,(n-1)!} = C(n+k-1, k)$$