# Topics in Physical Chemistry and Biophysics

## 1 Review of probability

18.04.23 lec 1

**Definition 1.1**

*Probability.* If $N$ is the total number of outcomes, and $n_A$ fall in category $A$, then

$$p_A = \frac{n_A}{N} = \frac{\text{outcomes cat. } A}{\text{all outcomes}}$$

Rules of composite events:

1. Mutually exclusive: outcomes $(A_1, A_{2,\dots})$ are *mutually exclusive* if one outcome precludes another outcomes. (Event $A_1$ prevents even $A_2$ from happening simultaneously.)
2. Collectively exhaustive: if all known outcomes are also all possible outcomes. $\sum p_i = 1$.
3. Independence: outcomes do not depend on each other.
4. Multiplicity: total number of ways in which outcomes occur.

Rules of calculation:

1. Let there be 3 outcomes $A, B, C$ with probability $p_A, p_B, p_C$. What is the probability that either one occurs ($A$ or $B$ or $C$)?

   $$p(A \cup B \cup C) = p_A + p_B + p_C$$

   That's the addition rule.
2. Probability that all outcomes occur? (Assuming independence)

   $$p(A \cap B \cap C) = p_A p_B p_C$$

3. Probability that an event $A$ is not happening? $p = 1 - p_A$

**Example.** We roll a die twice. What is the probability of rolling a 1 first **or** a 4 second?

Split the problem to parts. Note that the events are not mutually exclusive. Condition applies if:

- 1 first and not a 4 second: $\frac{1}{6} \cdot \frac{5}{6}$
- not a 1 first and a 4 second: $\frac{5}{6} \cdot \frac{1}{6}$
- 1 first and 4 second: $\frac{1}{6} \cdot \frac{1}{6}$

Now sum up all of the options to get result.

**Definition 1.2**

*Correlated events.* $p(B|A)$ is the probability that $B$ occurs given $A$ has occurred.

*Joint probability.* $p(AB)$ that both $A$ and $B$ occur.

**Definition 1.3**

*General multiplication rule.*

$$p(AB) = p(B|A)\, p(A)$$

$P(A)$ is called the a priori probability and $p(B|A)$ is called the a posterior probability

**Theorem 1.4**

*Bayes theorem.*

$$p(B|A)\,p(A) = p(A|B)\,p(B)$$

**Example.** 1% of population has breast cancer. We use mammography to detect cancer.

Event $A$: breast cancer. $p(A) = 0.01$. $p(\bar{A}) = 1 - p(A) = 0.99$.

Event $B$: diagnosis. $p(B|A) = 0.8$. $p(B|\bar{A}) = 0.096$. (i.e. false positive0

What is the chance that a doctor has diagnosed someone with cancer? i.e. $p(A|B)$

$$p(A|B) = \frac{p(B|A)\,p(A)}{p(B)}$$

$p(B)$ is the diagnosis of breast cancer irrespective whether it's there or not there.

$$p(B) = p(BA) + p(B\bar{A}) = p(B|A)\,p(A) + p(B|\bar{A})\,p(\bar{A}) = 0.8 \cdot 0.01 + 0.096 \cdot 0.99 = 0.103$$

$$p(A|B) = \frac{0.8 \cdot 0.01}{0.103} = 0.078 = 7.8\%$$

The reason that $p(A|B)$ is so small is that the rate of false positive is really low and the rate of having breast cancer is really low.

**Combinatorics.** Concerned with composition of events, and not with their order.

**Example.** How many combinations there are of $N$ amino acids?

$$W = N! = N\,(N-1)\,(N-2)\cdots$$

**Example.** Distinguish or not Distinguish: What are the possible number of ways to arrange $N$ amino acids? Divide all permutations (assuming objects are distinguishable) by the number of permutations of objects that are indistinguishable.

$$W = \frac{N!}{N_A}$$

In general, for $N$ objects consisting of $t$ categories in which the objects are indistinguishable:

$$W = \frac{N!}{(n_1!)\,(n_2!)\cdots(n_t!)}$$

So, if $t = 2$, (e.g. possible number of ways to arrange three acids A,A,H)

$$W = \frac{N!}{n_1! \cdot n_2!} = \frac{N!}{n_1!\,(N-n_1)!} = \binom{N}{n}$$

**Definition 1.5**

*Distribution functions.* Describe collections of probabilities. Relevant for continuous variables.

$$\sum_i p_i \longrightarrow \int_a^b p(x)\,\mathrm{d}x$$

Popular distributions:

1. *Binomial Distribution.* Relevant when there are only two outcomes.

**Example.** What is the probability that a series of $N$ trials has $n_H$ heads and $n_T$ tails in any order? $p_H, p_T$ are mutually exclusive, so the probability of one sequence is

$$p_H^{n_H} \cdot p_T^{n_T} = p_H^{n_H} (1 - p_H)^{N - n_H}; \quad N = n_H + n_T$$

and the number of ways to arrange the coins is

$$W = \frac{N!}{n_H! (N - n_H)!}$$

therefore, the possibility for the outcome (getting $n_H$ and $n_T$) in any order is

$$\boxed{p(n_H, N) = \binom{N}{n_H} p_H^{n_H} (1 - p_H)^{N - N_H}}$$

that's the binomial distribution.

**Example.** Given the molecule $C_{27}H_{44}O$ such that 1.1% is $^{13}C$ and the rest are $^{12}C$, the fraction of molecules without a single $^{13}C$ is given by the binomial distribution.

2. *Multinomial distribution.* Basically the extension of the binomial distribution.

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

**Definition 1.6**

*Moments of distributions.* Averages and Variances of distribution functions.

Given $p(i)$ s.t. $\sum_i p_{(i)} = 1$, the **Average** is defined as

$$\langle i \rangle = \sum_i i\, p(i) \longrightarrow \langle x \rangle = \int x\, p(x)\, \mathrm{d}x$$

*Given* $f(x)$,

$$\langle f(x) \rangle = \int f(x)\, p(x)\, \mathrm{d}x$$

*Given* $a \in \mathbb{R}$

$$\langle a f(x) \rangle = \int a f(x)\, p(x)\, \mathrm{d}x = a \langle f(x) \rangle$$

Given 2 functions $f(x), g(x)$,

$$\langle f(x) + g(x) \rangle = \langle f(x) \rangle + \langle g(x) \rangle$$

$$\langle f(x) \cdot g(x) \rangle \neq \langle f(x) \rangle \langle g(x) \rangle$$

The 2nd and 3nd **Moments** of the distributions $p(x)$ are

$$\langle x^2 \rangle = \int x^2\, p(x)\, \mathrm{d}x$$

$$\langle x^3 \rangle = \int x^3\, p(x)\, \mathrm{d}x$$

The **Variance** of the distribution, $\sigma^2$ is defined as

$$\sigma^2 = \langle x^2 \rangle - \langle x \rangle^2 = \langle (x - \langle x \rangle)^2 \rangle$$