Homework 1

Jinyuan Wu

September 22, 2022

Exercise 9 in chapter 1 (**) Let us show now one simple way to produce the realizations of B_m knowing the realizations of one example of $B_{1/2}$. From the set of the realizations of $B_{1/2}$, which we can view as a real number $r \in [0, 1[$ by forming the binary digit number $0.\beta^{(1)}\beta^{(2)}\beta^{(3)}\beta^{(4)}\beta^{(5)}\ldots$ (example: $0.011010110001111\ldots$), we can obtain the real numbers $r^{(1)}, r^{(2)}, r^{(3)}$ etc... from the formula:

$$r^{(n)} = (2^n r) \bmod 1$$

The realisations $\beta_m^{(n)}$ will be obtained from $r^{(n)}$ by the expression

$$\beta_m^{(n)} = 0 \text{ if } r^{(n)} \geqslant m$$
$$\beta_m^{(n)} = 1 \text{ if } r^{(n)} < m$$

Prove that this last expression yields the expected properties for B_m.

Solution Since

$$r^{(n)} = \beta^{(1)}\beta^{(2)}\cdots\beta^{(n)}.\beta^{(n+1)}\beta^{(n+2)}\cdots,$$

we know

$$r^{(n)} = 0.\beta^{(n+1)}\beta^{(n+2)}\cdots. \tag{1}$$

Each digit of $r^{(n)}$ is 0 or 1, and thus the possible range of $r^{(n)}$ is $[0, 1]$.[1] Suppose

$$x = 0.x^{(1)}x^{(2)}\cdots \in [0, 1],$$

we have

$$
\begin{aligned}
P(r^{(n)} < x) &= P(\beta^{(n+1)} < x^{(1)}) + P(\beta^{(n+1)} = x^{(1)})P(\beta^{(n+2)} < x^{(2)}) + \cdots \\
&= \frac{1}{2}\delta_{x^{(1)},1} + \frac{1}{2} \times \frac{1}{2}\delta_{x^{(2)},1} + \cdots \\
&= 0.x^{(1)}x^{(2)}\cdots = x,
\end{aligned}
$$

so $r^{(n)}$ has a uniform probabilistic distribution on $[0, 1]$. So the probability of $r^{(n)} < m$ i.e. $\beta_m^{(n)} = 1$ is exactly m, and therefore $\beta_m^{(n)}$ is a realization of B_m, regardless of what n is.

Exercise 14 in chapter 1 (**) Explain the link between the binomial distribution and the expansion of $(a + b)^N$.

Solution The binomial distribution can be derived by an intermediate step used to derive the expansion of $(a + b)^N$.

The binomial coefficient $\binom{N}{n}$ gives the number of ways to pick n points in N different points. Without invoking the commutative property of multiplication, there are 2^N terms in the expansion of $(a + b)^N$, each of which is like

$$aabbabba\cdots.$$

Now by the definition of the binomial coefficient, there are $\binom{N}{n}$ terms that have n a's and $(N-n)$ b's.

From this conclusion we can derive the expansion of $(a + b)^N$: there are $\binom{N}{n}$ terms in the total 2^N terms which has n a's and $(N - n)$ b's, and we have

$$(a + b)^N = \sum_{n=0}^{N} \binom{N}{n} a^n b^{N-n}. \tag{2}$$

---

[1] It's actually possible to have $r^{(n)} = 1$, because the binary $0.11111\cdots$ is actually 1, in the same way $0.9999\cdots = 1$ in the decimal case. But the probability to have such a $r^{(n)}$ is $1/2 \times 1/2 \times \cdots = 0$. That is, the event that $r^{(n)} = 1$ is possible but is a null set.

Similarly, if we consider the probabilistic distribution of

$$X_{m,N} = \sum_{k=1}^{N} B_{m,k}, \tag{3}$$

we will find the probability of the event that $X_{m,N} = x$ is the sum of the probability of all outputs of $\{B_{m,k}\}$ in which there are x 1 outputs and $N - x$ 0 outputs, and for each possible output, the probability is

$$p(1)^x p(0)^{N-x} = m^x (1-m)^{N-x},$$

and we have

$$p_{m,N}(x) = \binom{N}{x} m^x (1-m)^{N-x}. \tag{4}$$

So the relation between the binomial distribution and the $(a + b)^N$ expansion is they both involve the notion of "picking x points from N points". Indeed, by considering the normalization condition of (4), which is

$$1 = \sum_x p_{m,N}(x) = \sum_{x=0}^{N} \binom{N}{x} m^x (1-m)^{N-x}, \tag{5}$$

we rediscover the expansion of $(a + b)^N$, where we set $a = m$ and $b = 1 - m$.

Exercise 3 in chapter 2 (**) (a) Show that the above expression (2.15) for $w(x, t)$ with $t > 0$ satisfies this equation. (b) By using a double Fourier transform in x and t show that the Green's function of the Smoluchowsky equation (2.26) is indeed the above expression (2.15) for $w(x, t)$ with $t \geq 0$.

Solution

(a) From (2.15) we have

$$\frac{\partial}{\partial t} w(x, t) = -\frac{1}{2} \sqrt{\frac{1}{4\pi Dt^3}} e^{-\frac{(x-v_d t)^2}{4Dt}} - \sqrt{\frac{1}{4\pi Dt}} e^{-\frac{(x-v_d t)^2}{4Dt}} \frac{1}{4Dt^2}(2v_d(v_d t - x)t - (x - v_d t)^2)$$

$$= -\sqrt{\frac{1}{4\pi Dt}} e^{-\frac{(x-v_d t)^2}{4Dt}} \left(\frac{1}{2t} + \frac{(v_d t - x)(v_d t + x)}{4Dt^2} \right),$$

$$\frac{\partial}{\partial x} w(x, t) = -\sqrt{\frac{1}{4\pi Dt}} e^{-\frac{(x-v_d t)^2}{4Dt}} \frac{x - v_d t}{2Dt},$$

and

$$\frac{\partial^2}{\partial x^2} w(x, t) = -\sqrt{\frac{1}{4\pi Dt}} e^{-\frac{(x-v_d t)^2}{4Dt}} \left(\frac{1}{2Dt} - \left(\frac{x - v_d t}{2Dt} \right)^2 \right),$$

The RHS of the Smoluchowski equation is

$$D\frac{\partial^2 w}{\partial x^2} - v_d \frac{\partial w}{\partial x} = -\sqrt{\frac{1}{4\pi Dt}} e^{-\frac{(x-v_d t)^2}{4Dt}} \left(\frac{1}{2t} - \frac{(x - v_d t)^2}{4Dt^2} - v_d \frac{x - v_d t}{2Dt} \right)$$

$$= -\sqrt{\frac{1}{4\pi Dt}} e^{-\frac{(x-v_d t)^2}{4Dt}} \left(\frac{1}{2t} - \frac{x^2 - v_d^2 t^2}{4Dt^2} \right),$$

so we have

$$\frac{\partial}{\partial x} w(x, t) = D\frac{\partial^2 w}{\partial x^2} - v_d \frac{\partial w}{\partial x}.$$

(b) The initial condition is

$$\lim_{t \to 0} w = \delta(x),$$

which can be imposed to (2.26) by adding an "impact":

$$\frac{\partial w}{\partial t} = D\frac{\partial^2 w}{\partial x^2} - v_d \frac{\partial w}{\partial x} + \delta(x)\delta(t). \tag{6}$$

2

Now by Fourier transformation we have

$$w(x,t) = \int \frac{\mathrm{d}k\,\mathrm{d}\omega}{(2\pi)^2} \mathrm{e}^{-\mathrm{i}(\omega t - kx)} \tilde{w}(k,\omega),$$

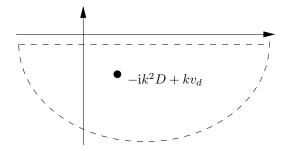$$-\mathrm{i}\omega\tilde{w} = D(\mathrm{i}k)^2\tilde{w} - \mathrm{i}kv_d\tilde{w} + 1.$$

We find

$$\tilde{w} = \frac{1}{-\mathrm{i}\omega + k^2 D + \mathrm{i}kv_d},$$

and thus

$$w(x,t) = \int \frac{\mathrm{d}k\,\mathrm{d}\omega}{(2\pi)^2} \mathrm{e}^{-\mathrm{i}(\omega t - kx)} \frac{1}{-\mathrm{i}\omega + k^2 D + \mathrm{i}kv_d}.$$

We first complete the integral over ω, with the following contour:

$$\int \mathrm{d}\omega\, \mathrm{e}^{-\mathrm{i}(\omega t - kx)} \frac{1}{\omega + \mathrm{i}Dk^2 - kv_d} = -2\pi\mathrm{i}\mathrm{e}^{-\mathrm{i}(-\mathrm{i}k^2 Dt + kv_d t - kx)}.$$

Thus

$$\begin{aligned}
w(x,t) &= \frac{\mathrm{i}}{(2\pi)^2} \int \mathrm{d}k\,(-2\pi\mathrm{i})\mathrm{e}^{-\mathrm{i}(-\mathrm{i}k^2 Dt + kv_d t - kx)} \\
&= \frac{1}{2\pi} \int \mathrm{d}k\, \mathrm{e}^{-k^2 Dt - \mathrm{i}k(v_d t - x)} \\
&= \frac{1}{2\pi} \cdot \sqrt{\frac{2\pi}{2Dt}} \mathrm{e}^{\frac{1}{2}\frac{1}{2Dt}(-\mathrm{i}(v_d t - x))^2} \\
&= \sqrt{\frac{1}{4\pi Dt}} \mathrm{e}^{-\frac{(x - v_d t)^2}{4Dt}}.
\end{aligned}$$

This is exactly (2.15).

Exercise 5 in chapter 2 (**) Explain in detail how, by measuring for the first time the position diffusion constant of a small Brownian sphere immerged in water, the physicist Jean Perrin, using the Einstein relation, was able to measure Avogadro's Number N_A, thereby confirming the existence of atoms (Jean Perrin received the Nobel prize for this work in 1926, see his Nobel lecture on the Nobel website). Use Stokes' law stating that a sphere of radius R moving at a velocity V feels in a fluid with viscosity η a frictional force

$$F = 6\pi R\eta V$$

Remember that Avogadro's Number N_A is involved in the ideal gas constant, defined by the relation

$$\frac{\text{pressure}\ \ \text{volume}}{\text{temperature}} = nR_{\mathrm{ig}}$$

where n is the number of moles of the volume of gas considered. In the kinetic theory of gases, R_{ig} is given by

$$R_{\mathrm{ig}} = N_A k_B$$

Solution The Stokes' law
$$F = 6\pi R \eta v \tag{7}$$
connects two physical quantities arising from the same dissipation process in the fluid: the viscosity μ and the response coefficient
$$\mu = \frac{v}{F}. \tag{8}$$
The relation between the two is imposed by the Navier-Stokes equation. Since we also have
$$\mu = \frac{D}{k_B T}, \tag{9}$$
we have
$$\frac{1}{6\pi R \eta} = \frac{D}{k_B T}. \tag{10}$$
This equation can be used to measure k_B: each quantities involved in the equation can be measured separately. The viscosity η can be measured by standard fluid dynamic methods. The radius R can be measured by letting the particles fall in the fluid and recording its terminal velocity, and then we have
$$R = \frac{mg}{6\pi \eta v_{\text{terminal}}}. \tag{11}$$
The diffusion coefficient D can be measured by looking at the trajectory of a Brownian particle. The temperature is measured by a thermometer. Now we find k_B, and by the ideal gas equation
$$pV = nR_{\text{ig}}T \tag{12}$$
we can measure R_{ig}, so finally, by
$$R_{\text{ig}} = N_A k_B, \tag{13}$$
the Avogadro constant is found.

Exercise 4 in lecture 3 Treat the case of the Shrapnell process in dimension 2.

Solution Now the damage is
$$X = \frac{\Omega}{r}, \tag{14}$$
and the probability per unit surface is
$$n = \frac{1}{\pi R^2}.$$
The condition $X < x$ is equivalent to
$$r > \frac{\Omega}{x}. \tag{15}$$
We have
$$p(r > \Omega/x) = \frac{R^2 - (\Omega/x)^2}{R^2}, \tag{16}$$
so the probability density is
$$w(x) = \frac{\mathrm{d}p(r > \Omega/x)}{\mathrm{d}x} = \frac{2\Omega^2}{R^2 x^3}. \tag{17}$$
There is a minimum of X: it's Ω/R, because explosion doesn't happen outside the circle. Now the first momentum is
$$\langle X \rangle = \int_{\Omega/R}^{\infty} \mathrm{d}x \, x \frac{2\Omega^2}{R^2 x^3} = \frac{2\Omega}{R} < \infty, \tag{18}$$
and the second momentum is
$$\langle X^2 \rangle = \int_{\Omega/R}^{\infty} \mathrm{d}x \, x^2 \frac{2\Omega^2}{R^2 x^3} = \infty. \tag{19}$$

So still the high order momenta of the variable diverges, and thus the central limit theorem fails.

Problem 2

Solution

(a) For a single bit we have

$$H(\mathrm{B}_{1/2}) = -\frac{1}{2}\log_2\frac{1}{2} - \frac{1}{2}\log_2\frac{1}{2} = 1. \tag{20}$$

For N independent bits, we have

$$H(\otimes_{j=1}^N \mathrm{B}_{1/2,j}) = -\sum_i \left(\frac{1}{2}\right)^N \log_2\left(\frac{1}{2}\right)^N = -2^N \times \frac{1}{2^N} \times (-N) = N. \tag{21}$$

(b) We have

$$S = -\frac{\partial F}{\partial T}, \quad F = -k_\mathrm{B} T \ln Z.$$

Now with the definition of the partition function

$$Z = \sum_i \mathrm{e}^{-E_i/k_\mathrm{B}T}, \tag{22}$$

we have

$$\begin{aligned}
\frac{\partial}{\partial T} T \ln Z &= \ln Z + \frac{T}{Z}\frac{\partial Z}{\partial T} \\
&= \ln Z + \frac{T}{Z}\sum_i \frac{E_i}{k_\mathrm{B}T^2}\mathrm{e}^{-E_i/k_\mathrm{B}T} \\
&= \ln Z + \frac{1}{k_\mathrm{B}T}\sum_i p_i E_i.
\end{aligned}$$

Thus

$$S = -k_\mathrm{B}\ln Z - \frac{1}{T}\sum_i p_i E_i = -k_\mathrm{B}\ln Z - \frac{\langle E \rangle}{T}. \tag{23}$$

On the other hand, we have

$$\begin{aligned}
\sum_i p_i \ln p_i &= \sum_i p_i \underbrace{\left(-\frac{1}{k_\mathrm{B}T}\sum_i E_i - \ln Z\right)}_{\ln p_i} \\
&= -\frac{1}{k_\mathrm{B}T}\langle E \rangle - \ln Z,
\end{aligned}$$

so we get

$$S = -k_\mathrm{B}\sum_i p_i \ln p_i = k_\mathrm{B}\ln 2 \times H. \tag{24}$$

(c) In the high temperature limit $E_i/k_\mathrm{B}T \to 0$ for every E_i, so energy is no longer important in determining the probabilistic distribution and each configuration has the same probability. The energy of N indistinguishable random bits, in this case, is therefore

$$E = \sum_{j=1}^N \mathrm{B}_{1/2,j}. \tag{25}$$

The probability of $E = \epsilon$ is

$$p(E = \epsilon) = \binom{N}{\epsilon} \times \frac{1}{2^N}, \tag{26}$$

which reaches its peak when $\epsilon = N/2$, so

$$\mathcal{E} = N/2. \tag{27}$$

There are $\binom{N}{\epsilon}$ microstates in the macrostate $(N, \mathcal{C})$, so

$$\lim_{N \to \infty} \frac{1}{N} S(N, \mathcal{E}) = k_{\mathrm{B}} \frac{1}{N} \ln \binom{N}{N/2}$$
$$= \frac{k_{\mathrm{B}}}{N} (\ln N! - 2 \ln (N/2)!)$$
$$\approx \frac{k_{\mathrm{B}}}{N} (N \ln N - 2(N/2) \ln N/2)$$
$$= k_{\mathrm{B}} \ln 2.$$

Thus

$$\lim_{N \to \infty} \frac{1}{N} S(N, \mathcal{E}) = k_{\mathrm{B}} \ln 2. \tag{28}$$

(d) An output of E – we denote it as ϵ – may be decomposed into

$$\epsilon = \sum_i n_i \xi_i, \tag{29}$$

so its probability is

$$p(\epsilon) = \Omega p_1^{n_1} p_2^{n_2} \cdots p_M^{n_M}, \tag{30}$$

where

$$\sum_i n_i = N, \tag{31}$$

and

$$\Omega = \binom{N}{n_1} \binom{N - n_1}{N - n_1 - n_2} \cdots \binom{N - n_1 - \cdots - n_{M-1}}{n_M}$$
$$= \frac{N!}{(N - n_1)! n_1!} \frac{(N - n_1)!}{(N - n_1 - n_2)! n_2!} \cdots \frac{(N - n_1 - \cdots - n_{M-1})!}{n_M! 0!} \tag{32}$$
$$= \frac{N!}{n_1! n_2! \cdots n_M!}.$$

Since N is large, we view $\{n_i\}$ as continuous variables. Then, to maximize $p(\epsilon)$, we can maximize

$$\ln p(\epsilon) = \ln \Omega + \sum_i n_i \ln p_i \tag{33}$$

under the constraint of (31). The expression of $\ln \Omega$ is given by the Stirling approximation:

$$\ln \Omega = \ln N! - \sum_i \ln n_i!$$
$$\approx N \ln N - N - \sum_i (n_i \ln n_i - n_i) \tag{34}$$
$$= N \ln N - \sum_i n_i \ln n_i.$$

By the method of Lagrange multiplier, the minimum is given by the equations

$$0 = \frac{\partial}{\partial n_j} \left(\ln \Omega + \sum_i n_i \ln p_i - \lambda \left(\sum_i n_i - N \right) \right)$$
$$= -\ln n_j - 1 + \ln p_j - \lambda.$$

This means there is a constant C such that for all i, we have

$$n_i = C p_i.$$

By (31), we know

$$n_i = N p_i. \tag{35}$$

So

$$\ln \Omega(N, \mathcal{E}) = N \ln N - \sum_i N p_i \ln(N p_i)$$

$$= N \ln N - N \sum_i p_i \ln p_i - \sum_i N p_i \ln N = -N \sum_i p_i \ln p_i. \tag{36}$$

So in the $N \to \infty$ limit, under $\log_2$, (36) means

$$\lim_{N \to \infty} \frac{1}{N} \log_2 \Omega(N, \mathcal{E}) = - \sum_i p_i \log p_i =: H(X). \tag{37}$$

The Boltzmann entropy of $(N, \mathcal{E})$ is

$$S(N, \mathcal{E}) = k_{\mathrm{B}} \ln \Omega(N, \mathcal{E}) = -k_{\mathrm{B}} N \sum_i p_i \ln p_i, \tag{38}$$

and thus

$$\lim_{N \to \infty} \frac{S(N, \mathcal{E})}{N} = -k_{\mathrm{B}} \sum_i p_i \ln p_i = k_{\mathrm{B}} \ln 2 \times H. \tag{39}$$

Thus if we put a large amount of identical copies of the same system together and calculate its Boltzmann entropy when the total energy is in its most possible value, then the average Boltzmann entropy only differs with the Shannon entropy with a prefactor $k_{\mathrm{B}} \ln 2$.

(e) To summarize:

– In the canonical ensemble, the thermodynamic entropy only differs with the Shannon entropy of the system configuration with a prefactor $k_{\mathrm{B}} \ln 2$ – see (24).

– In the microcanonical ensemble, when a large amount of systems are put together and their total energy is in the most possible value, then the average Boltzmann entropy only differs with the Shannon entropy of the system configuration with a prefactor $k_{\mathrm{B}} \ln 2$ – see (39).

(f) The terms in the RHS are

$$H(p_1 + p_2, \ldots, p_M) = -(p_1 + p_2) \log_2(p_1 + p_2) - \sum_{i \geq 3} p_i \log_2 p_i,$$

and

$$(p_1 + p_2) H\left(\frac{p_1}{p_1 + p_2}, \frac{p_2}{p_1 + p_2}\right)$$

$$= -(p_1 + p_2) \left(\frac{p_1}{p_1 + p_2} \log_2 \frac{p_1}{p_1 + p_2} + \frac{p_2}{p_1 + p_2} \log \frac{p_2}{p_1 + p_2}\right)$$

$$= -p_1 \log_2 p_1 + p_1 \log(p_1 + p_2) - p_2 \log p_2 + p_2 \log(p_1 + p_2),$$

so the RHS is

$$-p_1 \log_2 p_1 - p_2 \log_2 p_2 - \sum_{i \geq 3} p_i \log_2 p_i,$$

which is just the LHS. So

$$H(p_1, p_2, \ldots, p_M) = H(p_1 + p_2, \ldots, p_M) + (p_1 + p_2) H\left(\frac{p_1}{p_1 + p_2}, \frac{p_2}{p_1 + p_2}\right). \tag{40}$$

For example we have

$$H(1/2, 1/4, 1/4) = H(1/2, 1/2) + \frac{1}{2} H(1/2, 1/2) = 1 + \frac{1}{2} = 1.5. \tag{41}$$

(40) means the Shannon entropy is additive to some extent: if we first decide to ignore the difference between ξ_1 and ξ_1, then the resulting entropy is $H(p_1 + p_2, \ldots, p_M)$. To recover the original entropy, we just need to calculate the "inner" entropy of the ξ_1-or-ξ_2 possibility – that is, to calculate $H(p_1/(p_1 + p_2), p_2/(p_1 + p_2))$ – and then multiply a $(p_1 + p_2)$ weight to it, and after putting the two parts of entropies together, we get the original entropy corresponding to the full amount of information.

(g) We need to take the $\alpha \to 1$ limit of

$$H_\alpha(X) = \frac{1}{1-\alpha} \log_2 \left(\sum_i p_i^\alpha \right). \tag{42}$$

When $\alpha = 1$, both the numerator ($\log_2 1 = 0$) and the denominator are zero, so we can use the L'Hospital's rule:

$$\lim_{\alpha \to 1} H_\alpha = \lim_{\alpha \to 1} \frac{\frac{\sum_i \ln p_i p_i^\alpha}{\ln 2 \sum_i p_i^\alpha}}{-1} = -\sum_i p_i \log p_i = H(X).$$

(h) The eigenvalues of ρ are $(1 \pm |\boldsymbol{a}|)/2$. Thus

$$\begin{aligned} H_1(\rho) &= -\rho_1 \log_2 \rho_1 - \rho_2 \log_2 \rho_2 \\ &= 1 - \frac{1+|\boldsymbol{a}|}{2} \log_2(1+|\boldsymbol{a}|) - \frac{1-|\boldsymbol{a}|}{2} \log_2(1-|\boldsymbol{a}|), \end{aligned} \tag{43}$$

and

$$\begin{aligned} H_2(\rho) &= -\log_2 \left(\frac{(1+|\boldsymbol{a}|)^2}{4} + \frac{(1-|\boldsymbol{a}|)^2}{4} \right) \\ &= -\log_2 \frac{1+|\boldsymbol{a}|^2}{2}. \end{aligned} \tag{44}$$

The maximum 1 is reached when $|\boldsymbol{a}| = 0$, and the minimum 0 is reached when $|\boldsymbol{a}| = 1$. When $|\boldsymbol{a}| = 0$, ρ is essentially a classical $50\,\%$-$50\,\%$ probabilistic distribution, so the entropy is the same as the entropy of a random bit, which is 1. This is the most "noisy" case and indeed we get a maximal entropy here. When $|\boldsymbol{a}| = 1$, ρ is a pure state, and there is nothing uncertain about it, so its entropy is 0, which is the minimum.