Maximal Entropy principle - Problem 2.6 from Kardar

Amir Bar

January 28, 2013

Based on *Kardar, Statistical physics of particles*, problem 2.6; and the course's notes section 8.3.

1 Introduction

In this tutorial we shall use the maximal entropy principle in order to calculate specific distributions given specific constraints. The maximal entropy principle states that given a set of constraints, the distribution $\rho(x,p)$ which represents only the provided information is the distribution which maximizes the entropy functional

$$S\left[\rho(x,p)\right] = -I\left[\rho(x,p)\right] = -\int \rho(x,p)\log\left[\rho(x,p)\right]dxdp$$

This principle is effective not only as an alternative derivation of Boltzmann distribution, which is the maximal entropy distribution given average energy, but also in many other fields as diverse as neuroscience, economics etc.

2 Problem 2.6 from Kardar

consider the velocity of a gas particle in one dimension ($-\infty < v < \infty$).

(a) Find the unbiased probability density $\rho_1(v)$, subject only to the constraint that the average speed is c, that is, $\langle |v| \rangle = c$.

(b) Now find the probability density $\rho_2(v)$, given only the constraint of average kinetic energy, $\left\langle \frac{1}{2}mv^2 \right\rangle = \frac{1}{2}mc^2$

(c) Which of the above statements provides more information on the velocity? Quantify the difference in information in terms of $I_2 - I_1 = \left(\rho_2(v) - \rho_1(v)\right)/\log 2$

3 Solution

I will solve this problem following the method of the course's notes, which are slightly different from Kardar's method. I will then solve (a) again according to Kardar and find the same expression.

3.1 Solution 1

(a) According to Eq.(286) in the notes

$$\rho_1(v) = Z^{-1}\exp\left(-\lambda\left|v\right|\right)$$

with

$$\begin{aligned} Z &= \int_{-\infty}^{\infty}\exp\left(-\lambda\left|v\right|\right)dv \\ &= \int_{-\infty}^{0}\exp\left(\lambda v\right)dv + \int_{0}^{\infty}\exp\left(-\lambda v\right)dv \\ &= \left[\frac{1}{\lambda}\exp\left(\lambda v\right)\right]_{-\infty}^{0} - \left[\frac{1}{\lambda}\exp\left(-\lambda v\right)\right]_{0}^{\infty} \\ &= \frac{2}{\lambda} \end{aligned}$$

Now λ is given by

$$c = \langle |v| \rangle = -\frac{\partial \log Z}{\partial \lambda} = \frac{\partial \log\left(\lambda/2\right)}{\partial \lambda} = \frac{1}{\lambda}$$

So $Z = 2c$

$$\rho_1(v) = \frac{1}{2c}\exp\left(-\frac{|v|}{c}\right)$$

(b) Following the same logic

$$\rho_2(v) = Z^{-1}\exp\left(-\lambda\frac{mv^2}{2}\right)$$

with

$$Z = \int_{-\infty}^{\infty}\exp\left(-\lambda\frac{mv^2}{2}\right)dv = \sqrt{\frac{2\pi}{\lambda m}}$$

using the Gaussian integral formula.λ is given by

$$\frac{mc^2}{2} = \left\langle\frac{mv^2}{2}\right\rangle = -\frac{\partial \log Z}{\partial \lambda} = \frac{\frac{1}{2}\partial \log\left(\lambda m/2\pi\right)}{\partial \lambda} = \frac{1}{2\lambda}$$

So

$$\begin{aligned} \lambda &= \frac{1}{mc^2} \\ Z &= \sqrt{2\pi c^2} \end{aligned}$$

$$\rho_2(v) = \frac{1}{\sqrt{2\pi c^2}} \exp\left(-\frac{v^2}{2c^2}\right)$$

(c) Let's calculate I_1 and I_2:

$$\begin{aligned}
I_1 &= \int_{-\infty}^{\infty} \rho_1(v) \log\left[\rho_1(v)\right] dv \\
&= \frac{2}{2c} \int_0^{\infty} \exp\left(-\frac{v}{c}\right) \left(-\log(2c) - \frac{v}{c}\right) - dv \\
&= -\frac{\log(2c)}{c} \left[-c \exp\left(-\frac{v}{c}\right)\right]_0^{\infty} + \frac{1}{c^2} \frac{\partial}{\partial\left(c^{-1}\right)} \left[-c \exp\left(-\frac{v}{c}\right)\right]_0^{\infty} \\
&= -\log(2c) - \frac{c^2}{c^2} \\
&= -\left(\log(c) + \log(2) + 1\right)
\end{aligned}$$

and

$$\begin{aligned}
I_2 &= \frac{1}{\sqrt{2\pi c^2}} \int_{-\infty}^{\infty} \exp\left(-\frac{v^2}{2c^2}\right) \left(-\frac{1}{2} \log\left(2\pi c^2\right) - \frac{v^2}{2c^2}\right) dv \\
&= -\frac{\log\left(2\pi c^2\right)}{2\sqrt{2\pi c^2}} \sqrt{2\pi c^2} + \frac{1}{\sqrt{2\pi} c^3} \frac{\partial}{\partial\left(c^{-2}\right)} \left(\sqrt{2\pi c^2}\right) \\
&= -\frac{1}{2} \log\left(2\pi c^2\right) - \frac{c^3}{2c^3} \\
&= -\left(\log(c) + \frac{1}{2}\log(2\pi) + \frac{1}{2}\right)
\end{aligned}$$

So

$$\frac{I_2 - I_1}{\log(2)} = -\frac{\log(\pi) - \log(2) - 1}{2\log(2)} \approx 0.3956$$

which implies that the constraint on the speed is less informative than the constraint on the energy. In section 3.3 we shall discuss the reason for this.

3.2 Solving (a) according to Kardar

Kardar's approach is a bit more general and he treats the normalization constraint $\int_{-\infty}^{\infty} \rho(v) dv = 1$ as an additional constraint rather than through Z. Hence for section (a) one must maximize $S\left[\rho(v)\right]$ with two constraints: $\langle |v| \rangle = c$ and $\int_{-\infty}^{\infty} \rho(v) dv = 1$, i.e. maximize

$$\tilde{S}\left[\rho(v)\right] = -\int_{-\infty}^{\infty} \rho(v) \log\left[\rho(v)\right] dv - \alpha \left(1 - \int_{-\infty}^{\infty} \rho(v) dv\right) - \beta \left(c - \int_{-\infty}^{\infty} |v| \rho(v) dv\right)$$

Hence

$$\begin{aligned}
0 &= \frac{\delta \tilde{S}\left[\rho\right]}{\delta \rho} = -\log\rho - 1 - \alpha - \beta |v| \Rightarrow \\
\rho_1(v) &= \exp\left(-1 - \alpha - \beta |v|\right)
\end{aligned}$$

Using the normalization constraint

$$\begin{aligned}
1 &= 2e^{-1-\alpha}\int_0^\infty \exp\left(-\beta v\right)dv \\
&= \frac{2}{\beta}e^{-1-\alpha} \Rightarrow \\
\rho_1\left(v\right) &= \frac{\beta}{2}\exp\left(-\beta v\right)
\end{aligned}$$

and

$$\begin{aligned}
c &= 2\frac{\beta}{2}\int_0^\infty v\exp\left(-\beta v\right)dv \\
&= -\beta\frac{\partial}{\beta}\left[\frac{1}{\beta}\right] = \frac{1}{\beta} \Rightarrow \\
\rho_1\left(v\right) &= \frac{1}{2c}\exp\left(-\frac{v}{c}\right)
\end{aligned}$$

as we found above.

3.3 Discussing the result

The result of section (c) might seem counter-intuitive at first site - how come constraining the energy and the average speed is not equivalent? It is the speed which sets the energy! But the point is that these constraints have different content from the distribution perspective, although the typical values of velocities in both distributions are the same ($|v| \approx c$).

From the mathematical perspective, it is clear that $\rho_1(v)$ has broader tails than $\rho_2(v)$ as the first is exponential and the second Gaussian, so it is more likely that improbable events will occur when constraining the speed and not the energy.

A more intuitive argument is to consider higher moments of the velocity ($\left\langle v^2\right\rangle$ is the second and $\langle|v|\rangle$ is sort of a first moment). Constraining $\left\langle v^{2n}\right\rangle$ it should be clear now that the tails of the distribution will go like $\rho_n \sim \exp\left(-\alpha v^{2n}\right)$ for some α. In the limit $n \to \infty$, $\rho_n \to \delta\left(v - c\right)$ which has 0 entropy and hence the highest I_∞ possible. This is because as n increases, we are less tolerant to large deviations in the speed.