X MATHEMATICAL SUPPLEMENT

X.1 *Relative information entropies*

The Shannon-entropy S (or, equivalently, the Gibbs-entropy as it is called in the statistical physics community rather than in the information theory community)

$$S = -\sum_i p_i \ln p_i \tag{X.571}$$

can be extended to measuring the relative entropy between two discrete distributions p_i and q_i, yielding the Kullback-Leibler divergence ΔS,

$$\Delta S = -\sum_i p_i \ln\left(\frac{p_i}{q_i}\right) = \left\langle \ln\left(\frac{p_i}{q_i}\right)\right\rangle \tag{X.572}$$

which really plays its strength when generalised to continuous distributions $p(x)\mathrm{d}x$ and $q(x)\mathrm{d}x$,

$$\Delta S = -\int \mathrm{d}x\, p(x) \ln\left(\frac{p(x)}{q(x)}\right) = -\left\langle \ln\left(\frac{p(x)}{q(x)}\right)\right\rangle \tag{X.573}$$

The relative entropy comes with a large advantage as it is invariant under transformations of the random variable in the continuous case (the problem does not arise in the discrete case, anyways). The transformation law is commonly written as $p(x)\mathrm{d}x = p(z)\mathrm{d}z$ and results from integration by substitution:

$$\int \mathrm{d}x\, p(x) = \int \mathrm{d}z\, p(x(z))\frac{\mathrm{d}x}{\mathrm{d}z} \tag{X.574}$$

with a transformation Jacobian $J = \mathrm{d}x/\mathrm{d}z$. In contrast to the straightforward entropy $S = -\int \mathrm{d}x\, p(x) \ln p(x)$ which transforms to $-\int \mathrm{d}z\, p(z)[\ln(p(z)) + \ln J]$ and picks up an additional term depending on the transformation, this additional term cancels in the ratio $p(z)/q(z) = p(x)/q(x)$ of the relative entropy. That effectively means, that the continuum limit of the Shannon-entropy can not be defined in an invariant way:

$$S = -\int \mathrm{d}x\, p(x) \ln p(x) = -\langle \ln p(x)\rangle\,. \tag{X.575}$$

Related entropy measures, that are likewise (i) positive definite and bounded by 0, (ii) additive for independent random processes and (iii) growing with the number of possible outcomes are Rényi-entropies S_α

$$S_\alpha = -\frac{1}{\alpha-1}\ln\int \mathrm{d}x\, p(x)p(x)^{\alpha-1} = -\frac{1}{\alpha-1}\ln\left\langle p(x)^{\alpha-1}\right\rangle \tag{X.576}$$

for any constant $0 < \alpha \neq 1$. There are corresponding definitions of relative entropies ΔS_α

$$\Delta S_\alpha = -\frac{1}{\alpha-1}\ln\int \mathrm{d}x\, p(x)\left(\frac{p(x)}{q(x)}\right)^{\alpha-1} = -\frac{1}{\alpha-1}\ln\left\langle\left(\frac{p(x)}{q(x)}\right)^{\alpha-1}\right\rangle \tag{X.577}$$

One often runs into problems with Rényi-entropies when dealing with conditional and joint probabilities, which miraculously works with Kullback-Leibler divergences: Joint probabilities $p(x,z)$ can be generated in in a two-step random process as

$$p(x,z) = p(x|z)p(z) = p(z|x)p(x) \tag{X.578}$$

with conditional probabilities, which are obviously connected through Bayes' law. The conditional entropy $S(z|x)$ of $p(x,z)$ relative to $p(x)$ is given by

$$\begin{aligned} S(z|x) = -\int dx \int dz\, p(x,z) \ln\left(\frac{p(x,z)}{p(x)}\right) = \\ -\int dx \int dz\, p(x,z) \ln p(x,z) + \int dx \int dz\, p(x,z) \ln p(x) = \\ -\int dx \int dz\, p(x,z) \ln p(x,z) + \int dx\ \ln p(x) \int dz\, p(x,z) = \\ -\int dx \int dz\, p(x,z) \ln p(x,z) + \int dx\, p(x) \ln p(x) = S(x,z) - S(x) \end{aligned} \tag{X.579}$$

because of the marginalisation $\int dz\, p(x,z) = p(x)$ in the second term, such that we can write down the entropy-version of Bayes' law, making use of the symmetry of $S(x,z)$:

$$S(z|x) + S(x) = S(z,x) = S(x,z) = S(x|z) + S(z) \tag{X.580}$$

which is impossible to formulate in terms of Rényi-entropies due to the logarithm acting on an integral.

X.2 *Volume and area of the unit sphere*

In phase space integrals one needs to carry out an integration over momentum space. With the assumption of isotropic momenta this leads naturally to the question what the volume and surface area of a unit sphere would be. While everyone is familiar with 4π as the solid angle subtended by the unit sphere's surface in $3d$, this result can be generalised to higher dimensions.

Surface area $S_n(r)$ and volume $V_n(r)$ are naturally related by:

$$V_n(r) = \int_0^r dr\, S_n(r) \quad \rightarrow \quad S_n(r) = \frac{dV_n(r)}{dr} \tag{X.581}$$

If the volume scales $V = C(n)r^n$ with a function $C(n)$ yet to be determined, one needs the relation

$$S_n(r) = nC(n)r^{n-1}, \tag{X.582}$$

from which we can make the identification that $nC(n)$ is the integrated solid angle $\int d\Omega_{n-1}$, because $\int dr^{n-1} = r^n/n$ from the radial integration.

The volume element needed for this integration, expressed in Cartesian and spherical coordinates reads:

$$\mathrm{d}x_1\mathrm{d}x_2\ldots\mathrm{d}x_n = r^{n-1}\mathrm{d}r\mathrm{d}\Omega_{n-1} \tag{X.583}$$

with a yet unknown solid angle element $\mathrm{d}\Omega_{n-1}$ in n dimensions.

There is a neat trick to proceed for evaluating volume integrals without having the strict boundary condition that $x_1^2 + x_2^2 + \ldots + x_n^2 = r^2$, which is to integrate over an n-dimensional Gaussian:

$$\int \mathrm{d}x_1 \int \mathrm{d}x_2 \ldots \int \mathrm{d}x_n \, \exp\left(-(x_1^2 + x_2^2 + \ldots x_n^2)\right) = \int_0^\infty \mathrm{d}r\, r^{n-1} \exp(-r^2) \int \mathrm{d}\Omega_{n-1}. \tag{X.584}$$

The $\mathrm{d}x_n$-integrals factorise and give a contribution of $\sqrt{\pi}$ each, while the $\mathrm{d}r$-integral can be solved by substitution $r^2 \rightarrow s$, with $2\mathrm{d}r = \mathrm{d}s$:

$$\int_0^\infty \mathrm{d}r\, r^{n-1} \exp(-r^2) = \frac{1}{2}\int_0^\infty \mathrm{d}s\, s^{\frac{n-1}{2}} \exp(-s) = \frac{\Gamma(n/2)}{2}, \tag{X.585}$$

because the integral is exactly the representation of the Γ-function. Then,

$$\pi^{n/2} = \frac{\Gamma(n/2)}{2}\, n\mathrm{C}(n) = \Gamma(n/2)\frac{n}{2}\,\mathrm{C}(n) = \Gamma(n/2+1)\mathrm{C}(n) \tag{X.586}$$

using $n\Gamma(n) = \Gamma(n+1)$ as the generalisation of the factorial. Collecting results yields

$$\mathrm{C}(n) = \frac{\pi^{n/2}}{\Gamma(n/2+1)} \tag{X.587}$$

for the solid angle element, implying that volume and surface area of unit spheres in n dimensions, shown in Fig. 40, would be

$$\mathrm{V}_n = \mathrm{C}(n)r^n = \frac{\pi^{n/2}}{\Gamma(n/2+1)}\, r^n \quad \text{and} \quad \mathrm{S}_n = n\mathrm{C}(n)r^{n-1} = \frac{2\pi^{n/2}}{\Gamma(n/2)}\, r^{n-1}. \tag{X.588}$$

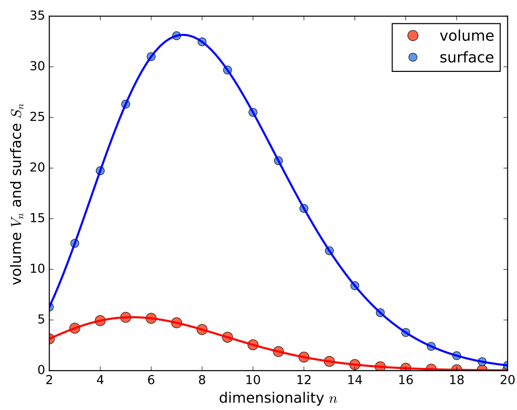

Figure 40: Volume V_n and surface S_n of a unit sphere, as a function of dimension n.