5.3.3 MaxEnt and the binomial distribution

Although we have already derived the binomial distribution from elementary considerations in Section 5.1, it's instructive to see how it emerges from the use of the MaxEnt principle. Suppose that we are given (only) the expected number of success in M trials, $\langle N \rangle = \mu$. What should we assign for the probability of a specific number of favourable outcomes, $\text{prob}(N|M,\mu)$?

According to the MaxEnt principle, we need to maximize the entropy S of eqn (5.28) subject to the testable information

$$\langle N \rangle = \sum_{N=0}^{M} N \, \text{prob}(N|M,\mu) = \mu \tag{5.41}$$

and normalization. Following an earlier calculation, this optimization yields the pdf of eqn (5.32); specifically, we obtain

$$\text{prob}(N|M,\mu) \propto m(N)\, \mathrm{e}^{-\lambda N}, \tag{5.42}$$

where λ is a Lagrange multiplier and $m(N)$ is the measure. The former is, of course, determined by the constraint on the mean, but we must first assign $m(N)$.

From our discussion of the Lebesgue measure in Section 5.2.2, $m(N)$ is proportional to the pdf which reflects gross ignorance about the details of the situation. Given only that there are M trials, the principle of indifference tells us to assign equal probability to each of the 2^M possible outcomes. The number of different ways of obtaining N successes in M trials, or ${}^{M}\mathrm{C}_N$ in eqn (5.8), is therefore an appropriate measure for this problem:

$$m(N) = \frac{M!}{N!\,(M-N)!}\,. \tag{5.43}$$

All we need to do now is to substitute this into eqn (5.42), and to impose the constraints of normalization and eqn (5.41). The related algebra is simplified by noting that

$$\sum_{N=0}^{M} m(N)\, \mathrm{e}^{-\lambda N} = \left(\mathrm{e}^{-\lambda} + 1\right)^{M}, \tag{5.44}$$

which follows from eqn (5.9) on putting $a = \mathrm{e}^{-\lambda}$ and $b = 1$. Its reciprocal yields the constant of proportionality in eqn (5.42), and its implicit differentiation with respect to λ, giving

$$\sum_{N=0}^{M} N\, m(N)\, \mathrm{e}^{-\lambda N} = M\left(\mathrm{e}^{-\lambda} + 1\right)^{M-1} \mathrm{e}^{-\lambda},$$

allows eqn (5.41) to be reduced to $M\left(1+\mathrm{e}^{\lambda}\right)^{-1} = \mu$. A little algebraic rearrangement then shows eqn (5.42) to be a binomial pdf:

$$\text{prob}(N|M,\mu) = \frac{M!}{N!\,(M-N)!} \left(\frac{\mu}{M}\right)^{N} \left(1 - \frac{\mu}{M}\right)^{M-N}. \tag{5.45}$$