THE PRINCIPLE OF MINIMAL FREE ENERGY. Assume that the probability that a system is in a state \(i\) is \(p_i\), and that the energy of the state \(i\) is \(E_i\). By a fundamental principle, nature tries to minimize the free energy \(F = -\sum p_i \log(p_i)\) when the energies \(E_i\) are fixed. The free energy is the difference of the entropy \(S(p) = -\sum p_i \log(p_i)\) and the energy \(E(p) = \sum p_i E_i\). The probability distribution \(p_i\) satisfying \(\sum p_i = 1\) minimizing the free energy is called the Gibbs distribution.

**SOLUTION:** \(\nabla f = (-1, -\log(p_1), \ldots, -1, -\log(p_n))\), \(\nabla g = (1, \ldots, 1)\). The Lagrange equations are \(-1 - \log(p_1) = \lambda p_1 + \ldots + p_n = 1\), from which we get \(p_i = e^{-\lambda}/(1 - e^{-\lambda})\). The last equation \(\sum p_i = 1\) fixes \(\lambda = -\log(1/e) = 1\). The distribution, where each event has the same probability is the distribution of maximal entropy.

REMARK. Maximal entropy means least information content. A dice which is fixed (asymmetric weight distribution for example) allows a cheating gambler to gain profit. Cheating through asymmetric weight distributions can be avoided by making the dice transparent.

THE PRINCIPLE OF MAXIMAL ENTROPY. Consider a dice showing \(i\) with probability \(p_i\), where \(i = 1, \ldots, 6\). The entropy of the probability distribution is defined as \(S(p) = -\sum p_i \log(p_i)\). Find the distribution \(p_i\) which maximizes entropy under the constraint \(\sum p_i = 1\).

**SOLUTION:** \(\nabla f = (-1 - \log(p_1), \ldots, -1 - \log(p_n))\), \(\nabla g = (1, \ldots, 1)\). The Lagrange equations are \(-1 - \log(p_i) = \lambda p_i + \ldots + p_n = 1\), from which we get \(p_i = e^{-\lambda}/(1 - e^{-\lambda})\). The last equation \(\sum p_i = 1\) fixes \(\lambda = -\log(1/e) = 1\). The distribution, where each event has the same probability is the distribution of maximal entropy.

REMARK. Maximal entropy means least information content. A dice which is fixed (asymmetric weight distribution for example) allows a cheating gambler to gain profit. Cheating through asymmetric weight distributions can be avoided by making the dice transparent.

**EXAMPLE.** To find the shortest distance from the origin to the curve \(x^2 + 3y^2 = 1\), we minimize \(f(x,y) = x^2 + y^2\) with constraint \(g(x,y) = x^2 + 3y^2 = 1\). The Lagrange equations are \(2x = 2\lambda x\), \(6y = 2\lambda y\), \(x^2 + 3y^2 = 1\). From the constraint equation, we obtain \(y = \sqrt{1/3}\) So, we have the solutions \((0, \pm \sqrt{1/3})\) and \((1, 0), (-1, 0)\). To see which is the minimum, just evaluate \(f\) on each of the points. We see that \((0, \pm \sqrt{1/3})\) are the minimum.

**HIGHER DIMENSIONS.** The above constrained extremal problem works also in more dimensions. For example, if \(f(x,y,z)\) is a function of three variables and \(g(x,y,z) = c\) is a surface, we solve the system of 4 equations

\[
\nabla f = (2x, 2y, 2z) \nabla g = (6x^2, 6y, 6y^2) \]

\(x^2 + 3y^2 = 1\) to find the 4 unknowns \((x, y, z, \lambda)\). In \(n\) dimensions, we have \(n+1\) equations and \(n+1\) unknowns \((x_0, \ldots, x_n, \lambda)\).

**THE PRINCIPLE OF MAXIMAL ENTROPY.** Consider a dice showing \(x, y, z\) with probability \(p_{xyz}\), where \(x, y, z = 1, \ldots, 6\). The entropy of the probability distribution is defined as \(S(p) = -\sum p_{xyz} \log(p_{xyz})\). Find the distribution \(p_{xyz}\) which maximizes entropy under the constraint \(\sum p_{xyz} = 1\).

**SOLUTION:** \(\nabla f = (3(2x+y), 3x, 3y, 3z) \nabla g = (6x^2, 6y, 6y^2) \)

\(x^2 + 3y^2 = 1\) so that \(\sum p_{xy} = 1\) giving \(\sum p_{xy} = 1\). The Gibbs distribution is \(p_{xy} = \exp(-E)\) where \(C = \exp(-\lambda)\). The additional equation \(\sum p_{xy} = 1\) gives \(\sum p_{xy} = 1\) so that \(C = 1/\sum p_{xy}\). The Gibbs distribution is \(p_{xy} = \exp(-E)\) \(\sum p_{xy} = 1\).

**EXAMPLE.** The Gaussian distribution (normal distribution) is a probability distribution that is maximized by the intersection \(x = 0\) and \(y = 0\). It is also an example of a distribution that is symmetric under the constraint \(\sum p_{xy} = 1\). The Gaussian distribution is \(p_{xy} = \exp(-E)\) \(\sum p_{xy} = 1\).

**SOURCE:** (a) The Book of the Dead, p. 2. (b) The California Heat Engine, p. 3.

**REMARK.** Other factors can influence the shape also. For example, the can has to withstand pressure forces up to 100 psi.