Entropy is the (logarithm of the) number of states of a system for which some *macroscopic* measure is left invariant. The logarithm is there just so we can add entropies, instead of multiplying the number of possible states.

For example take 100 coins. If we describe the *macrostate* by the number of heads, states with high entropy are those around 50, for which there are a huge number of possible *microstates* leading to the same measure. States around 0 and 100 are low entropy, since there are only few possible such arrangements.

Note that the definition of entropy depends on which measure(s) we are using to describe a macrostate.

Quick, some airplane notes. Maximum entropy distribution for the interval $[0, 1]$. Suppose the solution is $p(x) = 1$. The entropy is then given as $H_p = \int_0^1 1 \log 1 = 0$. We take another distribution $q(x) = 1 + \epsilon q'(x)$ with $\int_0^1 q'(x) = 0$. Using the linear approximation $\log(x) \approx 1 + x$, the linear approximation of the entropy is $H_q \approx -\int_0^1 (1+\epsilon q'(x))\epsilon q'(x) = -\int_0^1 \epsilon^2 q'(x)^2$, which is always less than $0$, so $p(x)$ has maximal entropy.