This definition corresponds to the negative of the entropy of a probability distribution as defined by Shannon [1].

The fact that the information is non-negative is no liability, since we shall primarily be interested in the differences of information when we compare two distributions actually seldom interested in the absolute information of a distribution only in differences.

add:

[1] §2 pp. 5. Addendum that a good discussion of information theory is to be found in Shannon [1] and Woodward [1]. Note, however, that in the theory of communication one defines information a state \( \omega_i \) which has a priori probability \( P_i \) to be \( \log P_i \). We prefer, however, to regard information as referring to the information of the distribution itself. Our information is then simply the expectation of the information contained in the state. —

[add to pg. 10, 76]

[1] §2 pp. 5. Addendum that a good discussion of information theory is to be found in Shannon [1] and Woodward [1]. Note, however, that in the theory of communication one defines information a state \( \omega_i \) which has a priori probability \( P_i \) to be \( \log P_i \). We prefer, however, to regard information as referring to the information of the distribution itself. Our information is then simply the expectation of the information contained in the state.

\[
S = \sum \omega_i [\log \omega_i]
\]

where, in [\( \omega_0 \) up]

\[
S = \left( \sum_{m=1}^{n} \omega_m [\log \omega_m] \right) = \sum \omega_m \left( \log \omega_m [\log \omega_m] \right)
\]

... completing the proof of Theorem (p. 9) i.e. The underlying theorem

\[
\beta(U\Lambda U^{-1}) = U\beta(A)U^{-1}\quad \text{see von Neumann pg. 388}
\]