It is possible to break a choice of one event from the set $\{E_1, E_2, \ldots, E_n\}$ into two steps: at first, to choose a group where the chosen event is and then to choose the event from the already chosen group. By Axiom A3, this gives us the following equality

$$H(1/n, 1/n, \ldots, 1/n) = H(p_1, p_2, \ldots, p_n) + \sum_{i=1}^{k} p_i H(1/n_i, 1/n_i, \ldots, 1/n_i) \tag{3.5.10}$$

It is proved that

$$H(1/n, 1/n, \ldots, 1/n) = \log_2 n$$

and thus,

$$H(1/n_i, 1/n_i, \ldots, 1/n_i) = \log_2 n_i$$

for all $i = 1, 2, 3, \ldots, n$.

This and the equality (3.5.10), give us

$$\log_2 n = H(p_1, p_2, \ldots, p_n) + \sum_{i=1}^{k} p_i \log_2 n_i \tag{3.5.11}$$

Consequently, we have

$$\begin{aligned} H(p_1, p_2, \ldots, p_n) &= \log_2 n - \sum_{i=1}^{k} p_i \log_2 n_i \\ &= -\sum_{i=1}^{k} p_i log_2 n_i/n = -\sum_{i=1}^{k} p_i \log_2 p_i \end{aligned}$$

This proves formula (3.2.4) for all rational numbers. As by Axiom A1, $H(p_i, p_2, \ldots, p_n)$ is a continuous function and the set of all rational numbers is dense in the set of all real numbers, we have

$$H(p_i, p_2, \ldots, p_n) = -\sum_{i=1}^{k} p_i \log_2 p_i$$

for all real numbers.

Theorem is proved.

Of we take Axiom A4 away, we have a weaker result.

Theorem 3.5.2 (Shannon, 1948). The only function that satisfies Axioms A1-A3 and A5 is the information entropy $H(p_i, p_2, \ldots, p_n)$ determined by the formula

$$H(p_i, p_2, \ldots, p_n) = K \sum_{i=1}^{k} p_i \log_2 p_i \tag{3.5.12}$$

where K is an arbitrary constant.