Statistical uncertainty and information [tln37]

An experiment has n possible outcomes that occur with probabilities $P_1, P_2, \ldots, P_n$.

Properties that must be satisfied by any quantitative measure of *uncertainty*:

1. The uncertainty is a function of the probabilities of all possible outcomes: $\Sigma = \Sigma(P_1, P_2, \ldots, P_n)$.
2. The uncertainty is symmetric under all permutations of the P_i.
3. The maximum uncertainty occurs if all P_i are equal.
4. The uncertainty is zero if one of the outcomes has probability $P_i = 1$.
5. The combined outcome of two independent experiments has an uncertainty equal to the sum of the uncertainties of the outcomes of each experiment.

$$\Rightarrow \ \Sigma(P_1, P_2, \ldots, P_n) = -\sum_{i=1}^{n} P_i \ln P_i = -\langle \ln P \rangle.$$

Information comes in messages: $A_1, A_2, \ldots$. A message carries information only if it contains some news, i.e. something not completely expected.

$P(A)$: probability that message A is sent.
$I(A)$: information gain if message is indeed received.

The less likely the message, the greater the information gain if the message is received:

If $P(A) < P(B)$ then $I(A) > I(B)$, if $P(A) = 1$ then $I(A) = 0$.

If two independent messages are received, then the information gain is the sum of the information gains pertaining to each individual message:

$$P(A \cap B) = P(A)P(B) \ \Rightarrow \ I(A \cap B) = I(A) + I(B).$$

The information content of a message is equal to the change in (statistical) uncertainty at the receiver:

$$P_1, P_2, \ldots, P_n \xrightarrow{A} \bar{P}_1, \bar{P}_2, \ldots, \bar{P}_n \ \Rightarrow \ I(A) = \Sigma(P_1, P_2, \ldots, P_n) - \Sigma(\bar{P}_1, \bar{P}_2, \ldots, \bar{P}_n)$$

Information as used here refers only to the scarcity of events. Any aspects of usefulness and meaningfulness are disregarded.