Measuring information beyond communication theory. Why some generalized information measures may be useful, others not

J. ACZEL

April 20, 2011

Common model

- Messages
- Frequencies
- Coding
- Entropy

$$H_n(p_l, \cdots, p_n) = -\sum^n p_k log p_k$$

Properties

- Bounded, nonnegativity
- Subadditivity(Additivity): $H(PQ) \leq H(P) + H(Q)$
- Conditional Entropy: $H(PQ) = H(P) + H(Q|P)$
- Mutual Information: $I(P, Q) = H(Q) - H(Q|P)$

Source Entropy

- Source entropy:
  $H^{\infty} = lim_{r \to \infty} H(P^r)/r = lim_{r \to \infty} H(P/P^{r-1})$
- Bounded, nonnegative.
- Expansibility
- Recursivity(Branching property)
  $H_{n+1}(p_1 * q_1, p_2 * q_2, p_2, \cdots, p_n) =$
  $H_n(p_1, \cdots, p_n) + p_1 * H_2(q_1, q_2)$

Other Measures

- Subadditivity, additivity, expansibility:
  $a * log\sharp(p_k \neq 0) + b\sum^n p_k * logp_k \ (a \geq 0 \geq b)$
- Replace the subadditivity by its generalization:
  $H(PQ|R) \leq H(P|R) + H(Q|R)$ then $a = 0$

Forecasting theory. Divergence

- $\sum^n p_k * f(q_k) \leq \sum^n p_k * f(p_k)$
- $f(q) = a * logq + b$
- $\sum^n p_k * f(p_k) = a * \sum^n p_k * logp_k + b$
- Directed divergence $\sum^n p_k * log(p_k/q_k)$

Sum-Form

- $H_n(p_1, \cdots, p_n) = \sum^n \phi(p_k)$
- Branching Property: $H_n(p_1, p_1, p_2, \cdots, p_n) = H_{n-1}(p_1 + p_2, \cdots, p_n) + J_n(p_1, p_2)$

Expected Information

- ${}^{\Psi}H_n(p_1,\cdots,p_n) = \Psi^{-1}(\sum^n p_k * \Psi(-logp_k))$ where $\Psi]0,\infty[\rightarrow R$ is continuous and strictly increasing.
- Renyi Entropy: ${}_aH_n(p_1,\cdots,p_n) = \frac{1}{1-a} * log \sum^n {p_k}^a$

Mixed theory of information

- Events+Messages
- $H_n\{\frac{E_1,\cdots E_n}{p_1,\cdots p_n}\}$ $(E_j \bigcap E_k = \emptyset$ if $j \neq k, p_k \geq 0, \sum^n p_k = 1)$
- $a * \sum^n p_k * logp_k + \sum^n p_k * g(E_k) - g(\bigcup^n E_k)$