General model based on complexity

Classical Information Theory

$\Phi(x,y)$; $x \in X, y \in Y$ a *complexity function*,

$x = P$, $y = \varkappa$

$x \curvearrowright \hat{x}$ a map of X into Y (often the identity).

$P \curvearrowright \varkappa = \hat{P}$

Φ: average code length: $\langle \varkappa, P \rangle \geq \langle \hat{P}, P \rangle$

Assume [diag], the *diagonal condition*:

$$\forall x \; : \; \min_{y \in Y} \Phi(x,y) = \Phi(x,\hat{x})$$

and minimum is only attained on diagonal ($y = \hat{x}$).

Shannon-Boltzmann-Gibbs entropy:

Define *Φ-entropy* as *minimal complexity*:

$$H(x) = \min_{y \in Y} \Phi(x,y)$$

$H(P) = -\sum p_i \ln p_i$

and *Φ-divergence or redundancy* as the difference.
Obviously then, the *linking identity* holds:

$$\Phi(x,y) = H(x) + D(x,y) \, .$$

Kullback-Leibler divergence

$D(P,Q) = \sum p_i \ln \frac{p_i}{q_i}$

Typical extra conditions, indicated briefly here, are: