

Relative information relative to a given information measure.

§5 Information for general distributions:

Although we have given a definition of correlation applicable for all probability distributions, we have not yet extended the definition of information past ~~finite~~ finite distributions. In order to make this extension we first generalize the definition that we gave for discrete distributions, to a definition of ~~information relative to an information measure~~ ^{relative}. If we assign a measure to the set of values of a random variable X , which is simply the assignment of a positive number a_i to each value x_i in the finite case, we define the information of a probability distribution $P(x_i)$ relative to this information measure to be

$$(5.1) \quad I_X = \sum_i P(x_i) \ln \frac{P(x_i)}{a_i} = \text{Exp} \left[\ln \frac{P(x_i)}{a_i} \right]$$

If we have a joint distribution of random variables X, Y, \dots, Z , with ~~information~~ measures on their values $\{a_{ij}\}, \{b_{ij}\}, \dots, \{c_{ij}\}$ respectively, then we define the total information relative to these measures:

$$(5.2) \quad I_{X_1 \dots X_n} = \sum_{i,j \dots k} P(x_i, y_j, \dots, z_k) \ln \frac{P(x_i, y_j, \dots, z_k)}{a_i b_j \dots c_k} = \text{Exp} \left[\ln \frac{P(x_i, y_j, \dots, z_k)}{a_i b_j \dots c_k} \right]$$

so that the ~~basic~~^{information} measure on the cartesian product set is always taken to be the product measure of the individual ~~set~~^{information} measures.

We should like now, after our previous position slightly and consider information ~~as~~ always being defined relative to some ~~basic~~ information measure, so that our previous definition of

of information is to be regarded as the information relative to the measure for which all the c_i 's by $i = 1, \dots, C_K$ are taken to be unities, which we shall henceforth call the uniform measure.

Let us now compute the correlation $\{X, Y, \dots, Z\}'$ ^{by (3.4)} using the relative information:

$$(5.3) \quad \{X, Y, \dots, Z\}' = I'_{XY\dots Z} - I'_X - I'_Y - \dots - I'_Z$$

$$= \exp \left[\ln \frac{P(X_1, Y_1, \dots, Z_K)}{c_1 c_2 \dots c_K} \right] - \exp \left[\ln \frac{P(X_1)}{c_1} \right] - \exp \left[\ln \frac{P(Y_1)}{c_2} \right] - \dots - \exp \left[\ln \frac{P(Z_K)}{c_K} \right]$$

$$= \exp \left[\ln \frac{P(X_1, Y_1, \dots, Z_K)}{P_1 P_2 \dots P_K} \right] = \{X, Y, \dots, Z\}$$

So that the correlation for discrete distributions, as defined by (3.4) is independent of the choice of information measure, and the correlation remains an absolute, not relative quantity. If we now consider refinements of our distributions, as before, and realize that such a refinement is also a refinement of the ~~info.~~ ^{information} measure, then we can prove a relation analogous to theorem 2:

Theorem 4: The information of a distribution relative to a ~~info.~~ given measure never decreases under refinement. (proof in appendix) $\textcircled{Q} P' \in P \Rightarrow I'^P \geq I^P$

Therefore, just as for correlation, we can define the information of an arbitrary probability measure M_P on the cartesian product of arbitrary sets X, Y, \dots, Z_n , relative to the ~~info.~~ measures U_X, U_Y, \dots, U_Z

by considering, as before, finite partitions \mathcal{P} into the sets $\{X_i\}, \{Y_j\}, \dots \{Z_k\}$, for which we take as definition of information:

$$(5.4) \quad I_{XY\dots Z}^{\mathcal{P}} = \sum_{ij\dots k} M_p(X_i, Y_j, \dots, Z_k) \ln \frac{M_p(X_i, Y_j, \dots, Z_k)}{M_X(X_i) M_Y(Y_j) \dots M_Z(Z_k)}$$

$I_{XY\dots Z}^{\mathcal{P}}$ is then, as ~~seen~~^{was}, a monotone function upon the directed set of partitions (by thm 4), and as before, we take the directed set limit^(which is always defined) for our well definition:

$$(5.5) \quad I_{XY\dots Z} = \lim_{\mathcal{P}} I_{XY\dots Z}^{\mathcal{P}} = \sup_{\mathcal{P}} I_{XY\dots Z}^{\mathcal{P}}$$

which is then the information relative to the ~~basic~~ information measures M_X, M_Y, \dots, M_Z .

Now, for functions f, g on a directed set the existence of $\lim f$ and $\lim g$ is a sufficient condition for the existence of $\lim(f+g)$, which is then $\lim f + \lim g$, provided that this is not indeterminate ($\stackrel{f \text{ b/w}}{s-\infty}$). Therefore

$$\begin{aligned} \text{Theorem} \quad \{X, \dots, Y\} &= \lim \{X, \dots, Y\}^{\mathcal{P}} = \lim [I_{X\dots Y}^{\mathcal{P}} - I_X^{\mathcal{P}} - \dots - I_Y^{\mathcal{P}}] \\ &= I_{X\dots Y} - I_X - \dots - I_Y \end{aligned}$$

where the information is taken relative to any ~~basic~~^{Information} measure for which the expression is not indeterminate. It is sufficient for the validity of the above that the basic measures M_X, \dots, M_Y be such that^{expression} the marginal informations I_X, \dots, I_Y shall be ^{none of} positively infinite.



wouldn't
be
(still)
be
appropriate

no!

The latter statement holds, ^{since} because of the general relation $I_{X \dots Y} \geq I_X + \dots + I_Y$, the determinateness of I_X, I_Y, I_Z is guaranteed so long as all of the I_X, I_Y are $< +\infty$.

Henceforth, unless otherwise noted, we shall understand that information is to be computed with respect to the uniform measure for discrete distributions and Lebesgue measure for continuous distributions over real variables. In case of a mixed distribution with a continuous density $P(X, Y, \dots Z)$ plus discrete "lumps" $P'(X_i, Y_j, \dots Z_k)$ we shall understand the ^{information} base measure to be the uniform one over the discrete range, and Lebesgue measure over the continuous range. These conventions then lead us to the expressions:

$$(5.6) \quad I_{XYZ} = \begin{cases} \sum_{ijk} P(X_i, Y_j, Z_k) \ln P(X_i, Y_j, Z_k) & (\text{discrete case}) \\ \int P(X, Y, Z) \ln P(X, Y, Z) dX dY dZ & (\text{continuous case}) \\ \sum_{ijk} P'(X_i, Y_j, Z_k) \ln P'(X_i, Y_j, Z_k) \\ + \int P(X, Y, Z) \ln P(X, Y, Z) dX dY dZ & (\text{mixed case}) \end{cases}$$

(unless otherwise noted)

The mixed case occurs often in quantum mechanics, for quantities ^{which} have both a discrete and continuous spectrum.

- 17 for introduction see Feller ✓
- 18 reference to article on detailed
Balancing Rev and Phy (1958-55
article on time coordination.
Wendell?)
- 19
- 20
- 21
- 22
- 23
- 24
- 25
- 26 26 Export Information about
Conservation of Information in Stockless Processes
- 27 Conservation of Information in class mech
(contradiction to stock processes)

Invert 1

In order to make this extension we first generalize the definition that we gave for discrete distributions to a definition of relative information, relative to a given underlying measure, on the values of the random variable, which we shall call an information measure for a random variable.

28

underlying

29

30