

1. Assume $f(x, y)$ exists and is continuous.
2. For a fixed Δ , for all integer i and j , define the intervals

$$A_x^i = [i\Delta, (i+1)\Delta) \quad \text{and} \quad A_y^j = [j\Delta, (j+1)\Delta)$$

and the rectangle

$$A_{xy}^{i,j} = A_x^i \times A_y^j.$$

3. Define discrete r.v.'s

$$\begin{cases} \hat{X}_\Delta = i & \text{if } X \in A_x^i \\ \hat{Y}_\Delta = j & \text{if } Y \in A_y^j \end{cases}$$

4. \hat{X}_Δ and \hat{Y}_Δ are quantizations of X and Y , respectively.

5. For all i and j , let $(x_i, y_j) \in A_x^i \times A_y^j$.

6. Then

$$\begin{aligned} I(\hat{X}_\Delta; \hat{Y}_\Delta) &= \sum_i \sum_j \Pr\{(\hat{X}_\Delta, \hat{Y}_\Delta) = (i, j)\} \log \frac{\Pr\{(\hat{X}_\Delta, \hat{Y}_\Delta) = (i, j)\}}{\Pr\{\hat{X}_\Delta = i\} \Pr\{\hat{Y}_\Delta = j\}} \\ &\approx \sum_i \sum_j f(x_i, y_j) \Delta^2 \log \frac{f(x_i, y_j) \Delta^2}{(f(x_i) \Delta)(f(y_j) \Delta)} \\ &= \sum_i \sum_j f(x_i, y_j) \Delta^2 \log \frac{f(x_i, y_j)}{f(x_i) f(y_j)} \\ &\approx \int \int f(x, y) \log \frac{f(x, y)}{f(x) f(y)} dx dy \\ &= I(X; Y) \end{aligned}$$

7. Therefore, $I(X; Y)$ can be interpreted as the limit of $I(\hat{X}_\Delta; \hat{Y}_\Delta)$ as $\Delta \rightarrow 0$.

8. This interpretation continues to be valid for general distribution for X and Y .