Pairwise Adjusted Mutual Information Supplementary Material

Anonymous Author(s)
Affiliation
Address
email

Proof of Proposition 1

The first equality follows from the fact that X and X_σ have the same distribution. Specifically, we have for any positive integer k,

$$\mathrm{P}(X_\sigma = k) = \mathrm{P}((X \circ \sigma)^{-1}(k)) = \mathrm{P}(\sigma^{-1}(X^{-1}(k))) = \mathrm{P}(X^{-1}(k)) = \mathrm{P}(X = k).$$

For the second, we observe that if σ is a random permutation of $\{1, \ldots, n\}$, chosen uniformly at random, so is σ^{-1} which implies:

$$\mathrm{E}(H(X, Y_\sigma)) = \mathrm{E}(H(X, Y_{\sigma^{-1}})) = \mathrm{E}(H(X_\sigma, Y)),$$

where we have used the first equality and the fact that $(X, Y_{\sigma^{-1}}) \circ \sigma = (X_\sigma, Y)$. The third equality is a direct consequence of the two first.

Proof of Proposition 2

If Y is constant, then $Y = Y_\sigma$ for all permutations σ and the result follows from (**??**). Now assume that Y is a permutation of $\{1, \ldots, n\}$. Then $H(X, Y) = H(Y) = \log(n)$ and $I(X, Y) = H(X)$, for any random variable X. It then follows from (**??**) (and the symmetry in X and Y) that

$$\Delta I(X, Y) = I(X, Y) - \mathrm{E}(I(X_\sigma, Y)) = 0.$$

Proof of Proposition 3

If $\Delta H(X) = 0$, then $d((X, X_\sigma)) = 0$ for all permutation σ. In particular, there exists some bijection f such that $X_\sigma = f(X)$. Now assume that for some integer i, the event $A = \{\omega : X(\omega) = i\}$ is such that $1 < |A| \ < n$. Then there exists some $j \neq i$ such that the event $B = \{\omega : X(\omega) = j\}$ is not empty. Choose $a \in A$ and $b \in B$ and define σ as the permutation of a and b. Then $X_\sigma(a) = X(b) = j$ while $X_\sigma(a') = X(a') = i$ for all $a' \in A \setminus \{a\}$. So $X_\sigma(a) \neq X_\sigma(a')$ while $X(a) = X(a')$ for all $a' \in A \setminus \{a\}$, which contradicts the existence of some mapping f that $X_\sigma = f(X)$. Thus for each integer i, the cardinal of the event $A = \{\omega : X(\omega) = i\}$ is $0, 1$ or n. This implies that X is constant or equal to some permutation of $\{1, \ldots, n\}$.

Proof of Theorem 1

Consider two items selected uniformly at random in $\{1, \ldots, n\}$. Let A_{i_1}, B_{j_1} be the clusters of the first item, A_{i_2}, B_{j_2} be the clusters of the second item. In particular, these items belong respectively to the sets $A_{i_1} \cap B_{j_1}$ and $A_{i_2} \cap B_{j_2}$. The probability of this event is:

$$\frac{n_{i_1 j_1} n_{i_2 j_2}}{n^2}.$$

Submitted to 35th Conference on Neural Information Processing Systems (NeurIPS 2021). Do not distribute.

Now assume that these items exchange their labels for the first clustering, so that the first item move to set A_{i_2} while the second item move to the set A_{i_1}. If $i_1 = i_2$ or $j_1 = j_2$, the new contingency matrix remains unchanged; now if $i_1 \neq i_2$ and $j_1 \neq j_2$, the new contingency matrix n'_{ij} remains unchanged except for the following entries:

$$n'_{ij} = \begin{cases} n_{ij} - 1 & \text{for } i,j = i_1, j_1 \text{ and } i_2, j_2, \\ n_{ij} + 1 & \text{for } i,j = i_1, j_2 \text{ and } i_2, j_1. \end{cases}$$

We obtain the similarity between clusterings A and B:

$$\begin{aligned} s_{\mathrm{p}}(A,B) = \sum_{i_1 \neq i_2, j_1 \neq j_2} \frac{n_{i_1 j_1} n_{i_2 j_2}}{n^2} \Bigg(& \frac{n_{i_1 j_1}}{n} \log \frac{n_{i_1 j_1}}{n} - \frac{n_{i_1 j_1} - 1}{n} \log \frac{n_{i_1 j_1} - 1}{n} \\ & + \frac{n_{i_2 j_2}}{n} \log \frac{n_{i_2 j_2}}{n} - \frac{n_{i_2 j_2} - 1}{n} \log \frac{n_{i_2 j_2} - 1}{n} \\ & + \frac{n_{i_1 j_2}}{n} \log \frac{n_{i_1 j_2}}{n} - \frac{n_{i_1 j_2} + 1}{n} \log \frac{n_{i_1 j_2} + 1}{n} \\ & + \frac{n_{i_2 j_1}}{n} \log \frac{n_{i_2 j_1}}{n} - \frac{n_{i_2 j_1} + 1}{n} \log \frac{n_{i_2 j_1} + 1}{n} \Bigg), \end{aligned}$$

where by convention, $x \log x = 0$ for any $x \leq 0$. Finally,

$$\begin{aligned} s_{\mathrm{p}}(A,B) = {} & 2 \sum_{i,j} \frac{n_{ij}(n - a_i - b_j + n_{ij})}{n^2} \left(\frac{n_{ij}}{n} \log \frac{n_{ij}}{n} - \frac{n_{ij} - 1}{n} \log \frac{n_{ij} - 1}{n} \right) \\ & + 2 \sum_{i,j} \frac{(a_i - n_{ij})(b_j - n_{ij})}{n^2} \left(\frac{n_{ij}}{n} \log \frac{n_{ij}}{n} - \frac{n_{ij} + 1}{n} \log \frac{n_{ij} + 1}{n} \right). \end{aligned}$$

Proof of Corollary 1

The proof follows on observing that the second sum in the expression of $s_{\mathrm{p}}(A,B)$ in Theorem 1 can be written:

$$\begin{aligned} S & \stackrel{d}{=} \sum_{i,j} \frac{(a_i - n_{ij})(b_j - n_{ij})}{n^2} \left(\frac{n_{ij}}{n} \log \frac{n_{ij}}{n} - \frac{n_{ij} + 1}{n} \log \frac{n_{ij} + 1}{n} \right) \\ & = \sum_{i,j : n_{ij} > 0} \frac{(a_i - n_{ij})(b_j - n_{ij})}{n^2} \left(\frac{n_{ij}}{n} \log \frac{n_{ij}}{n} - \frac{n_{ij} + 1}{n} \log \frac{n_{ij} + 1}{n} \right) \\ & \quad - \sum_{i,j : n_{ij} = 0} \frac{(a_i - n_{ij})(b_j - n_{ij})}{n^2} \frac{1}{n} \log \frac{1}{n}. \end{aligned}$$

We obtain:

$$\begin{aligned} S = {} & \sum_{i,j : n_{ij} > 0} \frac{(a_i - n_{ij})(b_j - n_{ij})}{n^2} \left(\frac{n_{ij}}{n} \log \frac{n_{ij}}{n} - \frac{n_{ij} + 1}{n} \log \frac{n_{ij} + 1}{n} + \frac{1}{n} \log \frac{1}{n} \right) \\ & - \left(n^2 - \sum_i a_i^2 - \sum_j b_j^2 + \sum_{i,j} n_{ij}^2 \right) \frac{1}{n} \log \frac{1}{n}. \end{aligned}$$

Proof of Corollary 2

The proof follows from Theorem 1 applied to the diagonal contingency matrix $n_{ij} = a_i \delta_{ij} = b_j \delta_{ij}$, where δ_{ij} denotes the Kronecker symbol.