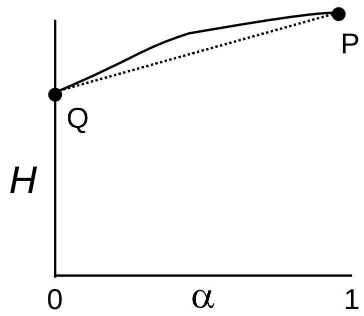
Multivariate Methods

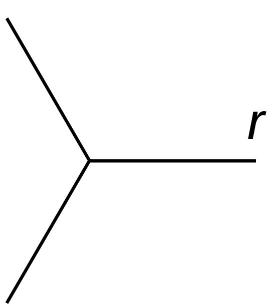
Homework #3 (2020-2021), Questions

Q1. Here we show that the entropy of a mixture is no less than the mixture of the entropies. Given two distributions P and Q, with entropies $H(P) = -\sum_i p_i \log p_i$ and $H(Q) = -\sum_i q_i \log q_i$, a mixture distribution $M_\alpha = \alpha P + (1-\alpha)Q$ is defined by the probabilities $m_{\alpha,i} = \alpha p_i + (1-\alpha) q_i$, for $0 \le \alpha \le 1$.
Show $H(M_\alpha) \ge \alpha H(P) + (1-\alpha) H(Q)$. Note that, since $H(M_0) = H(Q)$ and $H(M_1) = H(P)$, it suffices to show that $\frac{d^2}{d\alpha^2} H(M_\alpha) \le 0$, as this means that $H(M_\alpha)$ (solid line) is concave downward, and therefore above the line (dashed) of mixtures of entropies.



Q2: Here we show that the entropy of a joint distribution is maximized when the variables are independently distributed. Let P be a discrete probability distribution on a set of M values $\{x_i\}$, i.e., P_i is the probability that a random draw chooses the value x_i. Similarly, let Q be a discrete probability distribution on a set of N values $\{y_j\}$, i.e., Q_j is the probability that a random draw chooses the value y_j. Let R be a discrete distribution on a set of $M \times N$ values $\{(x_i, y_j)\}$, i.e., $R_{i,j}$ is the probability that a random draw chooses the pair of values (x_i, y_j). Find the joint distribution R that maximizes entropy, subject to the constraints that its marginals are compatible with P and Q, i.e., that $P_i = \sum_j R_{i,j}$ and that $Q_j = \sum_i R_{i,j}$. Lagrange multipliers will work nicely.

Q3: ICA: toy examples with cubic and quartic surrogates for entropy.



A. Consider the above distribution for bivariate data (centered at the origin), and its projection onto a line whose orientation with respect to the horizontal is given by θ. Determine the angular dependence of the second moment M_2, the third moment M_3, and the fourth moment M_4. Which of these is sensitive to the structure in the data?

B. Same as A, but for this distribution.