Homework Assignment 1

Information Theory, TBSI

October 9, 2017

1. **Entropy of NBA Playoffs.** In NBA Playoffs, two competing teams will play best of seven games, i.e., the team that has won 4 games wins, and ends the series of games. Denote the two teams by A and B, and the results of a game series by X. For instance, the possible values of X are $AAAA$, $BABABAA$, $BBAAAA$, etc. Let Y be the total number of games before the final result is determined. Assume that the result for each game is i.i.d, and the two teams are equivalently good. Calculate $H(X)$, $H(Y)$, $H(Y|X)$, $H(X|Y)$.

2. **Monty's hall.** Suppose you are a guest In a TV game show. There are three closed doors, and behind one and only one of them is a car while behind the other two are sheep. You will have a chance to win the car by opening the correct door. The host will let you pick one door, and them open one of the rest two doors without a car behind. Now you have a chance to change your mind. Shall you change your option or not? How much information has the host provide about the location of the car? (Suppose that the host knows the location of the car).

3. **Does conditioning reduce mutual information?** For random variables (r.v.'s) X, Y and Z, is inequality $I(X;Y|Z) < I(X;Y)$ always hold? If yes, proof it. if not, give a counter-example.
   Does conditioning reduce entropy? For random variables (r.v.'s) X, Y, is inequality $H(X|Y) \leq H(X)$ always hold? If yes, proof it. if not, give a counter-example. How about $H(X|Y=y) \leq H(X)$?

4. **Inequalities.** Let X, Y and Z are RVs. Proof the following inequalities and give the conditions for equality.
   (a) $H(X,Y|Z) \geq H(X|Z)$.
   (b) $I(X,Y;Z) \geq I(X;Z)$.
   (c) $H(X,Y,Z) - H(X,Y) \leq H(X,Z) - H(X)$.
   (d) $I(X;Z|Y) \geq I(Z;Y|X) - I(Z;Y) + I(X;Z)$.

5. **Calculation of K-L Divergence.** Let $p_1(x,y) = \frac{1}{2\pi\sigma_x\sigma_y}\exp[-\frac{1}{2}(\frac{x^2}{\sigma_x^2} + \frac{y^2}{\sigma_y^2})]$, and $p_2(x,y) = \frac{1}{2\pi\sigma_x\sigma_y\sqrt{1-\rho^2}}\exp[-\frac{1}{2(1-\rho^2)}(\frac{x^2}{\sigma_x^2} - 2\rho\frac{xy}{\sigma_x\sigma_y} + \frac{y^2}{\sigma_y^2})]$, compute $D(p_1||p_2)$, and $D(p_2||p_1)$. Is $D(p_1||p_2) = D(p_2||p_1)$? If not, please give an example of two distributions p, q such that $D(p||q) = D(q||p)$ (Other than the trivial case $p = q$).

6. **Mutual information and K-L Divergence.** $P_{X,Y}(x,y)$ is a joint distribution of discrete random variables X and Y. Assume $x_0 \in \mathcal{X}$ is a value of X, proof that

$$I(X;Y) = \sum_{x\in\mathcal{X}} P_X(x) D(P_{Y|X=x}||P_{Y|X=x_0}) - D(P_Y||P_{Y|X=x_0})$$

7. **Markov's inequality and Chebyshev's inequality.** If X is a nonnegative random variable and $a > 0$

   (a) Proof $\mathbb{P}(X \geq a) \leq \frac{\mathbb{E}(X)}{a}$

   (b) For $\forall d \geq 0$ Proof $\mathbb{P}(p(X) \leq d) log\frac{1}{d} \leq H(X)$

   (c) Proof $\mathbb{P}(|X - \mathbb{E}(X)| \geq a) \leq \frac{Var(X)}{a^2}$

8. **Entropy of Sums.** Let X and Y be two random variables. $\mathcal{X} = \{x_1, x_2, \cdots, x_r\}$ and $\mathcal{Y} = \{y_1, y_2, \cdots, y_s\}$. $Z = X + Y$.

   (a) Proof $H(Z|X) = H(Y|X)$, and if X and Y are independent, $H(Y) \leq H(Z)$ and $H(X) \leq H(Z)$.

   (b) Demonstrate by an example that, if X and Y are not independent, $H(X) > H(Z)$ and $H(Y) > H(Z)$.

   (c) Under what conditions $H(Z) = H(X) + H(Y)$?

9. **Measurement for distance.** A function $\rho(x,y)$ is a distance measurement, if the following four conditions are satisfied.

   - $\rho(x,y) \geq 0$,
   - $\rho(x,y) = \rho(y,x)$,
   - $\rho(x,y) = 0$, if and only if (iff.) $x = y$, and
   - $\rho(x,y) + \rho(y,z) \geq \rho(x,z)$.

   Define $\rho(X,Y) = H(X|Y) + H(Y|X)$, and $X = Y$, if there exists a one-to-one mapping between X and Y. Proof that $\rho(X,Y)$ is a distance measurement.

10. **Gaussian Random Vectors.** Suppose (X,Y) is a two-dimensional Gaussian Random Vector.

$$\begin{pmatrix} X \\ Y \end{pmatrix} \sim N_2\left(0, \begin{bmatrix} \sigma^2 & \rho\sigma^2 \\ \rho\sigma^2 & \sigma^2 \end{bmatrix}\right)$$

   Calculate $I(X;Y)$ when $\rho = -1, 0, 1$.

11. **Fano's inequality.** $\mathbb{P}(X = i) = p_i, i = 1, 2, ..., m$ and $p_1 \geq p_2 \geq ... \geq p_m$. The minimal probability of error predictor of X is $\hat{X} = 1$, with resulting probablity of error $P_e = 1 - p_1$. Maximize $H(\mathbf{p})$ subject to the constraint $1 - p_1 = P_e$ to find a bound on P_e in terms of H.