Prof. Dr. Rudolf Mathar, Dr. Michael Reyer

Tutorial 4
- Proposed Solution -

Friday, May 10, 2019

Solution of Problem 1

Theorem 4.3 shall be proven.

a) X is a discrete random variable with $p_i = P(X = x_i)$, $i = 1, \dots, m$. It holds

$$H(X) = -\sum_i p_i \log(p_i) \geq 0,$$

as $p_i \geq 0$ and $-\log(p_i) \geq 0$ for $0 < p_i \leq 1$ and $0 \cdot \log 0 = 0$ per definition. Equality holds, if all addends are zero, i.e.,

$$p_i \log(p_i) = 0 \Leftrightarrow p_i \in \{0, 1\} \quad i = 1, \dots, m,$$

as $p_i > 0$ and $-\log(p_i) > 0$, thus, $-p_i \log(p_i) > 0$ for $0 < p_i < 1$.

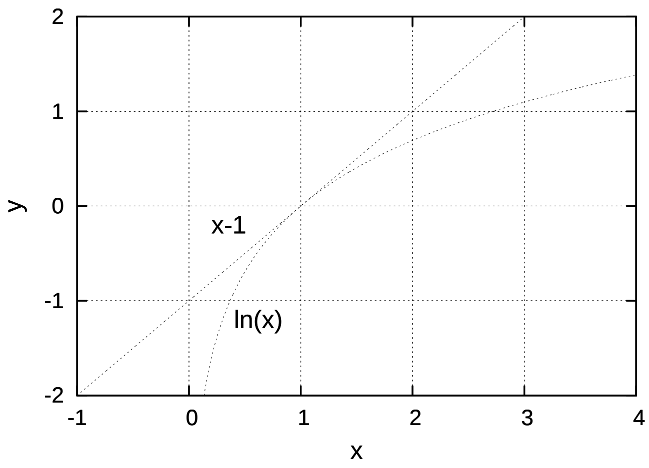
b) It holds

$$\begin{aligned}
H(X) - \log(m) &= -\sum_i p_i \log(p_i) - \underbrace{\sum_i p_i}_{=1} \log(m) \\
&= \sum_{i:p_i>0} p_i \log\left(\frac{1}{p_i\, m}\right) \\
&= (\log e) \sum_{i:p_i>0} p_i \ln\left(\frac{1}{p_i\, m}\right) \\
&\overset{\ln(x) \leq x-1}{\leq} (\log e) \sum_{i:p_i>0} p_i \left(\frac{1}{p_i\, m} - 1\right) \\
&= (\log e) \sum_{i:p_i>0} \left(\frac{1}{m} - p_i\right) = 0
\end{aligned}$$

As $\ln(x) = x - 1$ only holds for $x = 1$ it follows that equality holds iff $p_i = 1/m$, $i = 1, \dots, m$. In particular, as $p_i = \frac{1}{m}$, it follows $p_i > 0$, $i = 1, \dots, m$.

c) Define for $i = 1, \dots, m$ and $j = 1, \dots, d$

$$p_{i|j} = P(X = x_i \mid Y = y_j).$$



Show $H(X \mid Y) - H(X) \leq 0$ which is equivalent to the claim.

$$\begin{aligned}
H(X \mid Y) - H(X) &= -\sum_{i,j} p_{i,j} \log(p_{i|j}) + \sum_i p_i \log(p_i) \\
&= -\sum_{i,j} p_{i,j} \log\left(\frac{p_{i,j}}{p_j}\right) + \sum_i \underbrace{\sum_j p_{i,j}}_{=p_i} \log(p_i) \\
&= (\log e) \sum_{i,j:p_{i,j}>0} p_{i,j} \ln\left(\frac{p_i\, p_j}{p_{i,j}}\right) \\
&\overset{\ln(x)\leq x-1}{\leq} (\log e) \sum_{i,j:p_{i,j}>0} p_{i,j} \left(\frac{p_i\, p_j}{p_{i,j}} - 1\right) \\
&= (\log e) \sum_{i,j:p_{i,j}>0} (p_i\, p_j - p_{i,j}) = 0
\end{aligned}$$

Note that from $p_{i,j} > 0$ it follows $p_i, p_j > 0$. Equality hold for $p_i\, p_j = p_{i,j}$ which is equivalent to X and Y being stochastically independent.

This means that the mutual information $I(X,Y) = H(X) - H(X \mid Y)$ is nonnegative.

d) It holds

$$\begin{aligned}
H(X,Y) &= -\sum_{i,j} p_{i,j} \log(p_{i,j}) \\
&= -\sum_{i,j} p_{i,j} \left[\log(p_{i,j}) - \log(p_i) + \log(p_i)\right] \\
&= -\sum_{i,j} p_{i,j} \log \underbrace{\left(\frac{p_{i,j}}{p_i}\right)}_{p_{j|i}} - \sum_i \underbrace{\sum_j p_{i,j}}_{=p_i} \log(p_i) \\
&= H(Y \mid X) + H(X).
\end{aligned}$$

e) It holds

$$H(X,Y) \overset{(d)}{=} H(X) + H(Y \mid X) \overset{(c)}{\leq} H(X) + H(Y)$$

with equality as in (c) iff X and Y are stochastically independent.

Solution of Problem 2

Show for any function $f: X(\Omega) \times Y(\Omega) \to \mathbb{R}$, that $H(X, Y, f(X, Y)) = H(X, Y)$.

By definition, we have:

$$H(X,Y,Z=f(X,Y)) \stackrel{\text{Def.}}{=} -\sum_{x,y,z} P(X=x,Y=y,Z=z)\log\left(P(X=x,Y=y,Z=z)\right)$$

With

$$P(X=x,Y=y,Z=z) = \begin{cases} P(X=x,Y=y) & \text{, if } z=f(x,y) \\ 0 & \text{, if } z \neq f(x,y) \end{cases},$$

it follows that

$$H(X,Y,Z=f(X,Y)) = -\sum_{x,y} P(X=x,Y=y)\log(P(X=x,Y=y)) = H(X,Y)\,.$$

Note: It holds $0 \cdot \log 0 = 0$.

Solution of Problem 3

Prove Theorem 4.13 '⇒' (sufficient solution):

Recall that each element of these sets has a positive probability:

$$\mathcal{M}_+ := \{M \in \mathcal{M} \mid P(\hat{M}=M) > 0\},$$
$$\mathcal{C}_+ := \{C \in \mathcal{C} \mid P(\hat{C}=C) > 0\}.$$

Lemma 4.12 provides conditions of perfect secrecy on $\mathcal{M}_+$, $\mathcal{K}_+$, $\mathcal{C}_+$.
With Lemma 4.12 a), we obtain:

$$|\mathcal{M}_+| \leq |\mathcal{C}_+| \stackrel{(I)}{\leq} |\mathcal{C}| \stackrel{(II)}{=} |\mathcal{M}| \stackrel{(III)}{=} |\mathcal{M}_+|.$$

(I): With $P(\hat{C}=C) > 0 \Rightarrow \mathcal{C}_+ \subseteq \mathcal{C}$.
(II): Given by assumption $|\mathcal{M}| = |\mathcal{K}| = |\mathcal{C}|$.
(III): Given by assumption $P(\hat{M}=M) > 0,\ \forall M \in \mathcal{M}$.

By the 'sandwich theorem', i.e., the upper and lower bounds are both equal to $|\mathcal{M}_+|$:

$$\Rightarrow |\mathcal{C}_+| = |\mathcal{C}| \Rightarrow \mathcal{C}_+ = \mathcal{C},$$
$$\Rightarrow P(\hat{C}=C) > 0,\ \forall C \in \mathcal{C}.$$

Let $M \in \mathcal{M}$, $C \in \mathcal{C}$:

$$\begin{aligned} 0 < P(\hat{C}=C) &\stackrel{(IV)}{=} P(\hat{C}=C \mid \hat{M}=M) = P(e(\hat{M},\hat{K})=C \mid \hat{M}=M) \\ &\stackrel{(V)}{=} P(e(M,\hat{K})=C) = \textstyle\sum_{K\in\mathcal{K}:\, e(M,K)=C} P(\hat{K}=K) \neq 0 \qquad (1) \\ \Rightarrow\ & \forall M \in \mathcal{M},\ C \in \mathcal{C}\ \exists K \in \mathcal{K} : e(M,K) = C. \end{aligned}$$

(IV): With perfect secrecy as given by Corollary 4.11.
(V): Given by the assumption that $\hat{M}, \hat{K}$ are stochastically independent.

However, (1) is not shown to be unique yet!

(i) Fix $M \in \mathcal{M}$:

$$|\mathcal{C}_+| = |\mathcal{C}| = |\{e(M,K) \mid K \in \mathcal{K}_+ = \mathcal{K}\}| \leq |\mathcal{K}| \overset{(II)}{=} |\mathcal{C}|$$
$$\Rightarrow K \text{ is unique with } K = K(M,C) \text{ by the 'sandwich theorem'.}$$

(II) Given by assumption $|\mathcal{M}| = |\mathcal{K}| = |\mathcal{C}|$.
Let $M \in \mathcal{M}$, $C \in \mathcal{C}$:

$$\Rightarrow P(\hat{C} = C) \overset{(1)}{=} P(\hat{K} = K(M,C)),$$

because of perfect secrecy this expression is independent of M.

(ii) Fix $C_0 \in \mathcal{C}$:

$$\Rightarrow \{K(M,C_0) \mid M \in \mathcal{M}\} = \mathcal{K},$$

because of injectivity of $e(\cdot, K)$, i.e., $e(M,K) = C_0$, and by the assumption $|\mathcal{M}| = |\mathcal{C}|$.

$$\Rightarrow P(\hat{C} = C) = P(\hat{K} = K) \; \forall C \in \mathcal{C}, K \in \mathcal{K}$$
$$\Rightarrow P(\hat{K} = K) = \tfrac{1}{|\mathcal{K}|} \; \forall K \in \mathcal{K}. \quad \square$$