Analyse et Traitement de l'Information

TP5: Information Theory.

---

1 Quantifiers of information (50%)

In information theory, the amount of information present in a random variable (r.v.) is quantified with **entropy** which is a function, only of the probability mass function. Intuitively, entropy indicated how surprising the outcome of an experiment, or equivalently the observed value of an r.v. would be.

Mutual information, on the other hand, specifies the amount of information between two r.v.'s. In other words, it indicates how much the knowledge of one r.v. helps in predicting the value of another r.v.

Using the definitions in the lecture notes, try to solve the below problem. Pay careful attention to the differences between marginal, conditional and joint versions of entropy and mutual information.

U, V and W are three binary random variables. Their joint probability mass function is depicted as below:

U	V	W	$p(u, v, w)$
0	0	0	$\frac{1}{4}$
0	0	1	0
0	1	0	$\frac{1}{4}$
0	1	1	$\frac{1}{8}$
1	0	0	0
1	0	1	$\frac{1}{8}$
1	1	0	0
1	1	1	$\frac{1}{4}$

Try to calculate the information measures below:

a $H(U)$, $H(V)$ and $H(W)$

b $H(U|V)$, $H(V|U)$ and $H(W|U)$

c $I(U;V)$, $I(U;W)$, $I(U;V,W)$

d $H(U,V,W)$

For each of the above items you could directly use the definition. However, because they are related to each other in many ways, you could calculate some of them and derive the rest by using their relations. Chain rules for entropy and mutual information and drawing the Venn diagram could especially be helpful.

2 Communication through noisy channels (50%)

In this part we are investigating the problem of data transmission through noisy channels and try to experiment with a very simple channel coding method to combat the errors.

Suppose you have an alphabet of 32 words and you want to communicate a message based on them, i.e. your (random) message consists of the concatenation of these words. Your communication channel is binary, so what you send to the channel as the representative of each word of the alphabet should be a binary sequence as well, i.e. they should consist only of '0's and '1's.

Because you want your communication to be as fast as possible, you want to have your binary sequences for each word as short as possible.

As a natural way to choose these binary sequences, assign a 5-bit representative for each of the 32 words in the alphabet and keep them in a database. Therefore, you will have necessarily all possible combinations of '0's and '1' in your database.

As the medium of communication, You have a Binary Symmetric Channel with probability of bit error equal to 0.1 (BSC, $P_{be} = 0.1$). This means that every bit which is fed to this channel will be flipped in the output to its opposite due to noise, with probability 0.1. You could use MATLAB object **comm.BinarySymmetricChannel** or define it yourselves.

i. Suppose you have a string of 1000 words that you want to send to your destination through this channel. Generate the 5-bit binary representations of these words so that you will have a string of 5000 digits and see the output of the BSC. Convert the output binary string to index numbers. Calculate the fraction of erroneous output words to the total number of communicated words (1000) by comparing the output string to the input string.

ii. Redo the above experiment this time by generating binary representations of the words by repeating each bit 3 times, so that you will have 15 bit representations for each word. As an example, suppose the binary representation of the 27th item in the database was '11010' in the previous part. Now by repeating each bit 3 times, this 27th item will be **encoded** as '111111000111000'. Send a string of 1000 words to the channel (a binary string of length 15000).

In the output of the channel, divide the received binary sequence to 500, 3-bit long pieces and for each of these triples perform a majority vote between them. This means that if your received triple is '101', you should **decode** it as '1', because you have two '1's and one '0'.

Based on this encoding-decoding scheme, calculate the fraction of erroneous words to the total.

iii. Redo the previous part this time by repeating 5, 11 and 23 times and derive the error fraction for each. Your code should have the ability to be easily generalized to different values of repetitions.

iv. In Information Theory, to quantify how fast a communication is, we define rate of communication as:

$$R = \frac{log_2 \mathcal{M}}{n}, (0 \leq R \leq 1) \tag{1}$$

where $\mathcal{M}$ is the size of the alphabet and n is the length of the binary description of the words in the alphabet.

Calculate the rate of communication for each of the experiments above and compare it with the corresponding error fraction. Try to explain what happens when we increase the number of repetitions.

The above scenarios, follow essentially the general idea of **Channel Coding**, a very important branch of Information Theory with lots of achievements.

Assessment Please archive your report and codes in "Prénom Nom.zip" (replace "Prénom" and "Nom" with your real name), and upload to "`Upload TP5 - Information Theory`" on `https://moodle.unige.ch` before Monday, November 25 2019, 23:59 PM. Note, the assessment is mainly based on your report, which should include your answers to all questions and the explanations of your experimental results.

Supplements

1. Explain what is entropy, conditional entropy and the chain rule.

2. Explain what is entropy and mutual information with Venn diagrams.

3. Explain what is a communication channel in the Shannon's theory.

4. What are the differences between discrete and differential entropies?