Information Theory

Problem Set 10 - Advanced Information Measures

Luís Felipe Ramos Ferreira

lframos_ferreira@outlook.com

1. (a) Given a set of secret values, we can define the set of all posible probability distributions over this set of secret values. A state of knowledge of a agent about the secret (or state of the world) can be defined as one of such probability distributions that the set of posible secrets can have.

   (b) A information measure is a mapping from a state of knowledge to a real number, i.e, it's a function $f : \mathcal{X}\mathbb{D} \rightarrow \mathbb{R}$, where $\mathcal{X}\mathbb{D}$ is a set of probability distributions about a set of secrets $\mathcal{X}$, f is the function itself and $\mathbb{R}$ is a real number that represents how much knowledge the agent had about the secret before knowing it's value.

   (c) The mathematical definiton of information measure is a function that maps every posible input of possible states of knowledge to a real number that represents the value of that knowledge. The operational significance of information measure it's the meaning of the real number returned by the function mentioned, i.e., what that value means and how can we interpretate it.

   (d) i. Shannons's entropy is formally defined as the function

   $$H(\pi) = -\sum_{x \in \mathbb{X}} \pi_x log_2(\pi_x)$$

   , where π is a probability distribution over the set of secrets $\mathbb{X}$ Shannon's entropy operational significance is the expected number of questions an agent would ask to discover the secret when using an optimal binary search.

   ii. Bayes vulnerability is formally defined as the function

   $$V(\pi) = \max_{x \in \mathbb{X}} \pi_x$$

   . Bayes vulnerability operational significance is the probability that an agent can guess the secret with only one try.

   iii. Guessing entropy is formally defined as the function

   $$G(\pi) = \sum_{k} \pi_k k$$

, where k defines a non increasing ordering of the probability distribution π_i. Gyessing entropy operational significance is the expected number of guesses an agent would need to try in order to discover the secret in a optimal linear search.

(e) We can define $\mathbb{X}$ the set of secrets and $mathbbW$ the set of posible actions the adversary can take in order to discover the secrets we want to hide. The function

$$g : \mathbb{W} \times \mathbb{X} \to \mathbb{R}$$

can be seen as a gain function that represents the gain of the adversary when he takes action w and the secret value is x. In this context, the g-vulnerability of a distribution π is a information measure formally defined as

$$V_g[\pi] = \max_{w\in\mathbb{W}\sum_{x\in\mathbb{X}}\pi_x g(w,x)}$$

It represents the expected gain of a rational adversary taking a best action.

(f) When the adversary observes the output of a channel, it's state of knowledge about the world changes, i.e., it's prior probability distribution that represented his knowledge about the secret becomes a set of conditional probabilites given by the observation of the output of the channel.

(g) With a prior probability distribution pi as the prior state fo the knowledge about the secret, g as a gain function and C a channel matrix from $\mathcal{X}$ to $\mathcal{Y}$, the posterior g-vulnerability can be formally defined as a function that maps a posterior state of knowledge to a real number and is defined by:

$$V_g[\pi\rangle C] = \sum_{y\in\mathcal{Y}\text{ and }p(y)\neq 0} p(y)V_g(p(X \mid y)$$

It is interpreted as the expected optimal gain of a rational adversary given all possible outputs of the channel.

(h) Leakage, in this conext, represents the amount of information the adversary gains about the secret by the output of the chanenl. Given a prior state of knowledge probaility distribution π, a gain function g and a channel C, the leakage of information of a channel is formally defined by:

- Multiplicative g-leakage: the ratio of increase in the information the adversary has

$$\mathbb{L}_g^x(\pi, C) = \frac{V_g[\pi\rangle C]}{V_g(\pi)}$$

- Additive g-leakage: the absolute increase in the information the adversary has

$$\mathbb{L}_g^+(\pi, C) = V_g[\pi \rangle C] - V_g(\pi)$$

2. (a) We call $\mathbb{X}$ the set of posible secrets. Since each voter can only make a binary choice, candidate A or B, there are 2^k elements in $\mathbb{X}$. Since the prior state of knowledge is the uniform distribution π_u, we have that the Bayes vulnerability of the system is:

$$V(\pi_u) = \max_{x \in \mathbb{X}} \pi_u(x) = 2^{-k}$$

(b) Since $k = 3$, i.e, there are 3 voters, the set of posible secrets/votes is $\mathbb{X} = \{AAA, AAB, ABB, BBB, ABA, BAB, BAA, BBA\}$, where each triple $v_1v_2v_3$ represents for which candidate each voter voted. The output of the channel can be represented by the set $\mathbb{Y} = \{(0,3), (1,2), (2,1), (3,0)\}$, where each tuple (a, b) represents the number of votes candidates A and B got. The channel C can be represented by the following table:

C	(0, 3)	(1, 2)	(2, 1)	(3, 0)
AAA	0	0	0	1
AAB	0	0	1	0
ABA	0	0	1	0
ABB	0	1	0	0
BAA	0	0	1	0
BAB	0	1	0	0
BBA	0	1	0	0
BBB	1	0	0	0

Table 1: C

(c) Since π_u is a prior uniform distribution and $k = 3$, the prior Bayes vulnerability is:

$$V[\pi_u] = \max_x \pi_u(x) = \frac{1}{8}$$

Using the values for π_u and C, we can construct the joint distribution $\mathbb{J}$ below:

J(x, y)	(0, 3)	(1, 2)	(2, 1)	(3, 0)
AAA	0	0	0	$\frac{1}{8}$
AAB	0	0	$\frac{1}{8}$	0
ABA	0	0	$\frac{1}{8}$	0
ABB	0	$\frac{1}{8}$	0	0
BAA	0	0	$\frac{1}{8}$	0
BAB	0	$\frac{1}{8}$	0	0
BBA	0	$\frac{1}{8}$	0	0
BBB	$\frac{1}{8}$	0	0	0

Table 2: $J(x, y)$

With these values, we can then compute the set of posterior distributions of X given the value of the output Y, and also the values of the posterior probabilities on Y.

p(x — y)	(0, 3)	(1, 2)	(2, 1)	(3, 0)
AAA	0	0	0	1
AAB	0	0	$\frac{1}{3}$	0
ABA	0	0	$\frac{1}{3}$	0
ABB	0	$\frac{1}{3}$	0	0
BAA	0	0	$\frac{1}{3}$	0
BAB	0	$\frac{1}{3}$	0	0
BBA	0	$\frac{1}{3}$	0	0
BBB	1	0	0	0

Table 3: Posterior hyper distribution

y	p(y)
(0, 3)	$\frac{1}{8}$
(1, 2)	$\frac{3}{8}$
(2, 1)	$\frac{3}{8}$
(3, 0)	$\frac{1}{8}$

Table 4: Posterior probabilites on Y

With this values calculated, we can compute the posterior Bayes vulnerability:

$$V[\pi_u, C] = \sum_y p(y) max_x p(x \mid y) = \frac{1}{8} + \frac{3}{8}\frac{1}{3} + \frac{3}{8} + \frac{1}{8} = \frac{1}{2}$$

And the multiplicative Bayes leakage, which can be calculated with:

$$\frac{V[\pi_u, C]}{V[\pi_u]} = \frac{1}{2}\frac{8}{1} = 4$$

(d) d

(e) e

(f) f

References

[] David J. C. MacKay. *Information Theory, Inference and Learning Algorithms.* 7th edition, 2005.

[] Thomas M. Cover and Joy A. Thomas. *Elements of Information Theory 2nd Edition (Wiley Series in Telecommunications and Signal Processing).* Wiley-Interscience, July 2006. ISBN 0471241954.

[] Mário S Alvim, Konstantinos Chatzikokolakis, Annabelle McIver, Carroll Morgan, Catuscia Palamidessi, and Geoffrey Smith. *The Science of Quantitative Information Flow.* Springer, 2020.