CPSC 418/MATH 318 Practice Problems

Probability, Entropy and Perfect Secrecy, part II

Author: Janet Leahy (with modifications by Renate Scheidler)

1. A fair six-sided die is rolled, and the outcome is captured by a random variable X. You win \$2 if the result is a "1", you win \$1 if the result is a "6", but otherwise you lose \$1.

   (a) Find a random variable Y that properly describes the amount of money won or lost with each die roll. State the function $f : \mathcal{X} \to \mathcal{Y}$ such that $Y = f(X)$.

   (b) Calculate the expected value for the amount of money won or lost with each die roll. This is given by the formula

   $$\mathbb{E}[Y] = \sum_{y \in \mathcal{Y}} y \Pr(Y = y) \ .$$

2. Alice is constructing random variables in her lab. First, she flips two coins whose outcomes are represented by random variables X and Y. The coin for X is biased, with $p(0) = \frac{5}{6}$, and the coin for Y is fair, with $p(0) = \frac{1}{2}$. Then she defines a third random variable $Z = X \oplus Y$, where $\oplus$ denotes exclusive or.[1],

   (a) Are X and Y independent?

   (b) What is $\Pr(Z = 0 \mid X = 0)$?

   (c) Compute the probability distribution of Z from the probability distributions on X and Y.

   (d) What is the entropy of Z?

   (e) What is $\Pr(X = 0 \mid Z = 0)$?

   (f) Are X and Z independent?

3. Alice (Janet *cough*) is trying to design a encryption system with plaintext space $\mathcal{M} = \{a, b, c\}$ and ciphertext space $\mathcal{C} = \{1, 2, 3\}$.

   (a) This encryption table describes her first attempt, using key space $\mathcal{K} = \{k_1, k_2\}$. What is (very) wrong with this one?

	a	b	c
k_1	1	2	3
k_2	2	3	2

   (b) Alice's second attempt, using same key space, fixes the problem of part (a) via the followig ciphertext table:

[1] Note: This is a small example of the idea behind zero-knowledge proofs, which have applications in cryptography, complexity theory, entanglement theory in quantum information and more. The key idea is that a second party, say Eve, can learn something about the outcomes X and Y by accessing Z (in this case, she learns their parity), without gaining any other information about the individual outcomes themselves.

	a	b	c
k_1	1	2	3
k_2	2	3	1

Keys are chosen uniformly at random, so $P(K = k_1) = P(K = k_2) = \frac{1}{2}$. Does there exist a distribution over the message space such that the system provides perfect secrecy? If so, give an example of such a distribution, and if not, explain why.

(c) Finally, fixing the problem of part (b), Alice settles on the encryption system with key space $\mathcal{K} = \{k_1, k_2, k_3, k_4\}$, where ciphertexts given by the following table:

	a	b	c
k_1	1	2	3
k_2	2	3	1
k_3	1	3	2
k_4	3	2	1

The distributions over the plaintext space and the key space are both uniform, i.e. every plaintext and every key occurs with equal probability.

i. What is the probability distribution over ciphertextspace $\mathcal{C} = \{1, 2, 3\}$?

ii. Suppose Eve sees the ciphertext $C = 3$. What is her best guess for the message?

iii. What is the probability that this guess is correct?

iv. Does Alice's system provide perfect secrecy?

v. What is Eve's best guess for the key, given that she has seen the ciphertext $C = 3$? What is the probability that this guess is correct?

4. (a) Which has a higher entropy?

- Outcome of tossing a fair coin
- Outcome of tossing a biased coin

(b) Which has a higher entropy?

- Outcome of tossing a fair coin
- Outcome of tossing a biased coin
- Outcome of rolling a fair 6-sided die
- Outcome of rolling a biased 6-sided die

5. Consider a password system where passwords must be 5 characters long, and can consist of lowercase letters and/or digits (i.e. characters are drawn from the set $\{a, b, ...z\} \cup \{1, 2, ...0\}$).

(a) What is the entropy of this password system, assuming all passwords are equally likely?

(b) Now, suppose that to reduce the success of dictionary attacks, the system has ruled out a set of 500 most common passwords, including things like "test1" and "pwd00". Now, users cannot choose any of these as their password. What is the entropy of the new password system? How does it compare to the previous entropy?