10-601A Machine Learning (Spring 2014)
Homework 5 — Information Theory
Out: 31st Jan 2014
Due: 5th Feb 2014, 11:59pm

TA: Sujay Kumar Jauhar
`sjauhar@cs.cmu.edu`

Instructions: This homework on Information Theory contains 22 questions for a total of 100 points. You will use Blackboard (as in Homework 1 on Course Policies) to answer these structured questions. The homework is available at the following link: `https://blackboard.andrew.cmu.edu/webapps/blackboard/content/launchAssessment.jsp?course_id=_3871435_1&content_id=_614801_1&mode=cpview`. Note that this time, you are only allowed to **SUBMIT** your work **ONCE**, although you can save your work on Blackboard at any point and come back to it later – up until the deadline.

Whenever asked for a numeric solution to a question:

1. Make sure your answer is accurate to a precision of at least 4 significant digits.
2. Do **NOT** give your output in scientific notation, since your answers will be checked for correctness against decimal point numbers.

In the popular smartphone game Fruity Ninjas, random fruits appear on your screen and you are required to slice them. There is a steady stream of these random fruits. The only fruits that appear are Watermelons and Oranges. The game randomly generates a watermelon with a probability 0.6 and an orange with a probability 0.4.

Q.1 [3 pts] **How surprised are you (in bits) to observe a watermelon?**

Q.2 [3 pts] **How surprised are you (in bits) to observe an orange?**

Q.3 [4 pts] **What is the expected information content of a fruit showing up on your screen?**

You roll two fair six-sided dice. Let S be the sum of the two faces of the dice after a roll.

Q.4 [4 pts] **How surprised are you in bits to observe S = 3?**

Q.5 [4 pts] **How surprised are you in bits to observe S = 12?**

Q.6 [4 pts] How surprised are you in bits to observe S = 6?

Q.7 [4 pts] How surprised are you in bits to observe S = 8?

Q.8 [5 pts] What is the entropy of S?

An expedition of mountaineers heads up Mount Everest to study the effects of altitude on human physiology. A mountaineer volunteers to make a series of recordings of her Blood Oxygen Level(O) and Pulse Rate(P) numbers once she reaches Death Zone($>$ 26,000ft). Back in her lab, she analyses the data and finds that her readings fluctuate somewhat. She decides to treat this uncertainity in readings as a probability distribution. She groups the O and P readings into CRITICAL, ABNORMAL and NORMAL categories and creates the following probability table.

	Pr(P=CRITICAL)	Pr(P=ABNORMAL)	Pr(P=NORMAL)
Pr(O=CRITICAL)	0.05	0.025	0.0
Pr(O=ABNORMAL)	0.025	0.4	0.2
Pr(O=NORMAL)	0.0	0.2	0.1

Q.9 [3 pts] What is the marginal probability for Pr(P=ABNORMAL)?

Q.10 [3 pts] What is the marginal probability for Pr(O=CRITICAL)?

Q.11 [4 pts] What is the entropy H(P)?

Q.12 [4 pts] What is the entropy H(O)?

Q.13 [5 pts] What is the conditional entropy H(P|O)?

Q.14 [5 pts] What is the conditional entropy H(O|P)?

Q.15 [5 pts] What is the mutual infromation I(P;O)?

Q.16 [3 pts] The entropy of a random variable can be measured in _____.

Q.17 [5 pts] Select all that apply. The following are true about the entropy of a random variable X with non-zero probability for all $x \in X$:

(a) $H(X) = -\sum_{x \in X} p(x)log(p(x))$

(b) $H(X) \geq 0$

(c) $H(X) = \sum_{x \in X} p(x)log(p(x))$

(d) $H(X) = E_p[log(\frac{1}{p(X)})]$

(e) $H(X) > 0$

(f) $H(X) = log(E_p[\frac{1}{p(X)}])$

Q.18 [3 pts] The entropy of a random variable is maximized when its probability distribution is _____.

Q.19 [6 pts] Among the following, select all the identities that are guaranteed to always hold with respect to the entropy of several random variables:

(a) $H(X,Y) = H(Y,X)$

(b) $H(X|Y) = H(Y|X)$

(c) $H(X) - H(X|Y) = H(Y) - H(Y|X)$

(d) $H(X,Y) \leq H(X) + H(Y)$

Q.20 [5 pts] Select all that apply. The following are true about Mutual Information:

(a) $I(X;Y) = H(X) - H(X|Y)$

(b) $I(X;Y) > 0$

(c) $I(X;Y) = H(X) + H(Y) + H(X,Y)$

Q.21 [8 pts] Let X and Y be discrete random variables which are identically distributed (which implies that H(X) = H(Y)), but not necessarily independent. Then define $r = 1 - \frac{H(Y|X)}{H(X)}$. Select all that apply to r:

(a) $r = -\sum_{v \in X} p(v)^2 log(p(v)^2)$

(b) $0 \leq r \leq 1$

(c) $r = 0$ when $X \perp Y$

(d) $r = \frac{I(X;Y)}{H(X)}$

(e) $r = 1$ when $X \perp Y$

where $X \perp Y$ means that X and Y are independent.

Q.22 [10 pts] Let X, Y and Z be jointly distributed random variables. Which of the following inequalities are true?

[NOTE: this questions involves concepts that will be taught in the lecture on Tuesday, 4th Feb 2014. You might want to wait until then before answering it.]

(a) $H(X,Y,Z) - H(X,Y) \leq H(X,Z) - H(X)$

(b) $H((X,Y)|Z) \geq H(X|Z)$

(c) $H(X,Y|Z) = H(X|Z) + H(Y|X,Z)$

(d) $I((X,Y);Z) \geq I(X;Z)$

(e) $I(X;Y|Z) \geq 0$