# COMPSCI 650: Applied Information Theory
# Exam I: Mar. 10, 2016

Instructor: Arya Mazumdar

OPEN BOOK, OPEN NOTES. CALCULATORS OK. USE OF PHONES AND COMPUTERS NOT ALLOWED. PLEASE BE RIGOROUS AND PRECISE AND SHOW YOUR WORK. THIS EXAM CONSISTS OF FOUR PROBLEMS THAT CARRY A TOTAL OF 100 POINTS, AS MARKED (PERFECT SCORE = 100). DURATION: 75 MINUTES. GOOD LUCK!

**NAME (LAST, FIRST):**
**Student Id Number:**

| | Points Available | Points Achieved |
|---|---|---|
| Problem 1 : | 30 | |
| Problem 2 : | 15 | |
| Problem 3 : | 40 | |
| Problem 4 : | 15 | |
| Totals : | 100 | |

**Problem 1:** (30 points Data Compression) Let the random variable $X$ have five possible outcomes $\{1, 2, 3, 4, 5\}$. Consider two distributions $p(x)$ and $q(x)$ on this random variable.

| Symbol | p(x) | q(x) |
|---|---|---|
| 1 | 0.5 | 0.5 |
| 2 | 0.25 | 0.125 |
| 3 | 0.125 | 0.125 |
| 4 | 0.0625 | 0.125 |
| 5 | 0.0625 | 0.125 |

1) Compute $H(p), H(q), D(p\|q)$ and $D(q\|p)$.
2) Find the Huffman code for two distributions. What are the two respective average lengths?
3) Now assume that we use the code designed for $q$ when the distribution is $p$. What is the average length of the codewords? By how much does it exceed the entropy of $p$?

**Problem 2:** (5+10 points) Consider the following game played using a dice: a single dice is rolled and we gain a dollar if the outcome is 2, 3, 4 or 5, and lose a dollar if its 1 or 6.

1) What is our expected gain assuming all outcomes in $\{1, 2, 3, 4, 5, 6\}$ are equally likely.
2) Find the maximum entropy distribution over the universe $U = \{1, 2, 3, 4, 5, 6\}$ such that the expected gain is at least $T$ (say $T$ is greater than the expected gain for the uniform distribution).

**Problem 3:** (40 points) Consider parametric family of Gaussian distributions: $X \sim \mathcal{N}(0, \theta)$, where $\theta$ is the variance of the distribution. That is

$$f(x; \theta) = \frac{1}{\sqrt{2\pi\theta}} e^{-\frac{x^2}{2\theta}}.$$

1) Find out the score function $V(X)$.
2) Find out the Fisher Information $J(\theta)$. (Hint: $\int x^4 f(x; \theta) dx = 3\theta^2$).
3) What is the Cramér-Rao bound on the mean-square error of the best estimator?
4) Suppose we see iid samples from the distribution $X_1, X_2, \ldots, X_n$. Give an example of an unbiased estimator.

**Problem 4:** (15 points Very biased coin ) Suppose, we have two coins, probability of heads for the first is $\epsilon$ and probability of heads for the second is $1-\epsilon$. Suppose we randomly and uniformly picked up a coin and toss $m$ times. We aim to identify whether this is coin 1 or 2.

1) Use Le Cam's identity and Pinsker's inequality to find a lower bound on the probability of error.
2) What is the minimum value of $m$ so that the probability of error is at most $1/100$?