General comments: Problems 4 and 5 give some practice with basic mathematical techniques we’ll need. If you have routine in this kind of work, they should be fairly straightforward. If not, feel free to seek additional help from textbooks, the net etc.

The parameters in the Bonus Problem are chosen so that there should be no issues with numerical accuracy or computation time, if you think a little about how to arrange the computations sensibly.

1. Explain briefly your background and what made you interested in this course. What is your major subject (computer science, math, ...), and at what stage of your studies you are? How much mathematics have you studied (name some courses you have taken)? How about programming—what language(s) you use, and how would you describe your programming skills? If you have studied something else relevant, like statistics, signal processing, etc., you can also mention that. Have you studied any information theory before?

2. Find out what compression algorithms are used in the gadgets you use—DVD player, digital TV, CD, iPod, cell-phone, laptop (WinZip), etc. You can simply list the names of the algorithms, you don’t have to study how these algorithms work.

3. In a programming language of your choice, find as short a program as you can for printing out the sequence 141592653589793...375678 (the first 10000 digits of \( \pi \) after the decimal point). Search the web to find an algorithm for calculating \( \pi \). Your solution to the problem should contain your source code and a brief description of the algorithm you use (the formula and where it’s from).

4. For \( 0 \leq p \leq 1 \), we define the binary entropy \( H(p) \) as

\[
H(p) = -p \log_2 p - (1 - p) \log_2 (1 - p).
\]

This is the entropy of a random variable that gets value 0 with probability \( p \) and value 1 with probability \( 1 - p \), and will appear frequently during the rest of the course.

Prove that \( H(p) \) is strictly concave and achieves its maximum at \( p = 1/2 \). (Hint: Remember that a sufficient condition for concavity is that the second derivative is negative.)

Continues on the next page!
5. We write $\mu = \mathbb{E}[X]$ for the expectation and $\sigma^2 = \text{Var}[X]$ for the variance of a random variable $X$. (Recall that $\sigma^2 = \mathbb{E}[(X - \mu)^2].$)

(a) Prove Markov’s Inequality: for any non-negative random variable $X$ and any $\delta > 0$, we have
$$P(X \geq \delta) \leq \frac{\mu}{\delta}.$$ 

(b) Apply Markov’s Inequality to the random variable $(X - \mu)^2$ to obtain Chebyshev’s Inequality: for any $\delta > 0$ we have
$$P(|X - \mu| \geq \delta) \leq \frac{\sigma^2}{\delta^2}.$$ 

(c) Consider now $X$ that follows binomial distribution with parameters $n$ and $p$. That is, $X$ is the number of heads in $n$ independent tosses of a biased coin with probability $p$ of getting heads. We know from a basic course in probability that now $\mu = np$ and $\sigma^2 = np(1 - p)$. Use Chebyshev’s Inequality to give an upper bound for the probability that the value of $X$ differs from its expectation by at least 5%; that is, $|X - \mu|/\mu \geq 0.05$. Give a general formula and calculate numerical values for $p = 0.5$ and $p = 0.8$, with $n = 20$, $n = 100$ and $n = 1000$. 

**Warning:** This gives very loose upper bounds. In practice you should just calculate the values exactly (if you need specific numerical values) or use Chernoff bounds or similar (if you need an asymptotically fairly accurate closed-form formula for theoretical purposes); see the Bonus Problem. 

**Bonus Problem** Consider a binomially distributed random variable $X$ with parameters $n$ and $p$, as in Problem 5(c). 

(a) We know that for $k \in \{0, \ldots, n\}$ we have
$$P(X = k) = \binom{n}{k} p^k (1 - p)^{n-k}.$$ 

Use this to calculate the probability for $X$ differing from its expectation by at least 5%, for the same combinations of $n$ and $p$ as in 5(c). That is, you should calculate the actual values for the probabilities you upper bounded in 5(c) using Chebyshev’s Inequality. 

(b) Compare the probabilities you calculated in part (a) with upper bounds you get from the Chernoff bound
$$P(|X - \mu| \geq \delta \mu) \leq 2 \exp(-\mu \delta^2 / 3).$$ 

**Remark:** In the literature you can find a large number of inequalities called “Chernoff bound.” The one given here is one of the simplest, but perhaps not the tightest.