Your first exam is on October 18th. It will cover Chapters 1-4 of the book – everything up to and including the Noisy Coding Theorem, excluding sections 1.6 and 1.7, which we didn't cover. Note that this means I decided to leave Chapter 5 (CRCs) off of this test, but that generosity comes at a price – it means all of Chapter 5 (instead of just the easy stuff) will be on the next exam.

The early chapters are tough for some people, myself included, because doing all of the basic ideas about probability would easily fill up an entire course (like Math 4653 or 5651), and we don't have the time or desire to do that in this course. We just need a few ideas and terms so we can talk intelligently about how many words a code has, what the probability of transmission error is, and so on. That means we necessarily have a whole bunch of definitions and theorems compressed together. Random variables might not appear until you're in the second month of a normal probability course, but they showed up during the second *week* of this course. Not having studied everything in as much depth, some of our "results" relied on intuition and elegant hand-waving.

On the test, I'll stay away from these intuitive things which can be hard to explain; for example, you won't have to estimate the different informational rates in a novel or a technical article. You'll have some computational problems – like homework – involving counting, probability, or other formulas. You'll also have to use some of the big results (like the Kraft-McMillan inequalities or Noisy Coding Theorem) to draw certain conclusions. A list of possible topics would include:

Chapter 1: Basics of functions, sets, counting, and probability, including the ideas in 1.4 and 1.5 (but not variance).

Chapter 2: A basic understanding of the issues in 2.1 is important for the course, but hard to test on an exam. Definitions of information, entropy, and the properties of entropy starting on p38. Computing entropy given a list of probabilities (whether from a probability space or as the distribution of values of a random variable).

Chapter 3: The definitions in 3.1 are important, but try not to be confused by the abuses of language. (Random variables now taking on values in W instead of $\mathbb{R}$, etc.) The Kraft-McMillan inequalities are important. You should know them, but also be able to use them. (You don't need to know the proofs, even though we did them in class.) Same goes for the Noiseless Coding Theorem. Huffman codes are also important. You'll probably have to construct a Huffman Code and/or know its properties. **Note**: a Huffman Code can have words of different lengths, or they might all be the same length. I should have done more examples in class to stress this, but ask me on Monday in class and I can clear this up.

Chapter 4: I've already written two pages about what I think the important issues are in Chapter 4. It was handed out in class and is posted on the web site; I won't repeat myself here, but please read through it. Formulas for rate, capacity (especially for a Binary Symmetric Channel), Hamming weight and distance are all likely to show up on the exam. The Noisy Coding Theorem is the reason this field exists, so you should know and understand its statement, even if we skipped the proof. Also realize that a few basic error detection or correction schemes (like parity checks or repetition codes) are likely to show up; otherwise I don't have anything for you to compute the rate of! The probability calculations in section 4.2 (showing how effective a parity check is) are also a good source of problems.