MDL exercises, fourth handout

(due March 20th)

1. Let $H(p) = -p \log p - (1-p)\log(1-p)$ denote the binary entropy of a Bernoulli[p] distribution when the probability of observing a zero is p. (The logarithm base is two.) Use Stirling's approximation $\ln(n!) = (n + \frac{1}{2})\ln n - n + \frac{1}{2}\ln 2\pi + O(1/n)$ to show that $\log \binom{n}{\gamma n} = nH(\gamma) - \frac{1}{2}\log n + O(1)$.

2. Suppose we observe an individual sequence X^n of 0s and 1s of length n, with n_1 ones. Let θ represent that the data are i.i.d. Bernoulli with probability of 1 given by θ. Suppose first (a) that $n_1/n = 1/3$. For what value of θ is $P_\theta(x^n)$ maximized? Now suppose (b) that $\theta = 1/3$. For what value of n_1 is $P_\theta(x^n)$ maximized? Finally, suppose (c) again that $\theta = 1/3$. We sample n outcomes from P_θ, and now we regard N_1 as a random variable. For what values of m is $P_\theta(N_1 = m)$ approximately maximized? (it is very hard to calculate this precisely, so it is sufficient to give two sets B_1 and B_2 (whose definition may depend on n) such that for large n, for some $m \in B_1$, the probability is much larger than for all $m \in B_2$).

3. Markov Chains.

   a) Compute the maximum likelihood estimator $\hat{\theta} = (p_{0\to 1}, p_{1\to 1})$ for a binary first order Markov chain.

   b) Draw X_1, X_2, X_3 from an order 1 Markov chain. Are X_1 and X_3 dependent? What if you know the value of X_2?