1. (5 points) Prove the equality on page 72 of our text that for direct chaining hashing,
\[ s^2(A_n) = \frac{n(m-1)}{m^2} \]

2. (3 points) Given two sorted lists,
\[ a_1, \ldots, a_{n_a}, b_1, \ldots, b_{n_b} \]
show that any algorithm which merges them into one sorted list of length \( n_a + n_b \) on the basis of pairwise comparisons needs, in the worst case,
\[ C_{n_a,n_b}^{\text{MM}} \geq \hat{\text{E}} \log_2 \left( \frac{n_a + n_b}{n_a} \right) \]
pairwise comparisons. (Note: this is the top result of Gonnet 4.3.3, pg. 187).

Hint: You may proceed by
- noting that any algorithm whose basic computation is a pairwise comparison may be modeled as a binary tree,
- establishing a lower bound on the number of outputs of the algorithm (leaves of the binary decision tree),
- finding the minimum height of a binary tree which has at least a certain number of leaves.
1. \( s^2(A') = \mathbb{E}[(A')^2] - n^2/m^2 = \sum_{i \geq 0} i^2 \mathbb{C}(n,i) \frac{1}{m} [(m-1)/m]^{n-i} \frac{n^2}{m^2} \)

To remove the \( i^2 \), we note that
\( \mathbb{C}(n,i) = \frac{n}{i} \mathbb{C}(n-1,i-1) = \frac{n*(n-1)/(i*(i-1))}{i} \mathbb{C}(n-2,i-2) \), so
\( i^2 \mathbb{C}(n,i) = (n^2-n) \mathbb{C}(n-2,i-2) + i \mathbb{C}(n,i) \)

\( \mathbb{E}[(A')^2] = \sum_{i \geq 0} (n^2-n) \mathbb{C}(n-2,i-2) \frac{1}{m} [(m-1)/m]^{n-i} + \sum_{i \geq 0} i \mathbb{C}(n,i) \frac{1}{m} [(m-1)/m]^{n-i} - n^2/m^2 \)

The second term is seen to be identical to \( \mathbb{E}[A'] \), which we showed in class to be \( n/m \). So we now have
\( s^2(A') = n/m - n^2/m^2 + \sum_{i \geq 0} (n^2-n) \mathbb{C}(n-2,i-2) \frac{1}{m} [(m-1)/m]^{n-i} \)

To get the final term into a form to use the binomial theorem, we replace
\( i \) by \( i+2 \), yielding
\( \sum_{i \geq 0} (n^2-n) \mathbb{C}(n-2,i) \frac{1}{m} [(m-1)/m]^{n-i-2} \)

Pulling out the factor \( (n^2-n) \), and noting that \( \mathbb{C}(n-2,i)=0 \) for \( i<0 \), we note that this term is
\( (n^2-n) \sum_{i \geq 0} \mathbb{C}(n-2,i) \frac{1}{m} [(m-1)/m]^{n-i} \)

Pulling out the factor \( (1/m)^2 \), we can apply the binomial theorem to
\( [(n^2-n)/m^2] \sum_{i \geq 0} \mathbb{C}(n-2,i) \frac{1}{m} [(m-1)/m]^{n-i-2} \) yielding
\( [(n^2-n)/m^2] [(1/m) + (m-1)/m]^{n-2} = (n^2-n)/m^2 \), and plugging back in to the original equation,
\( s^2(A') = n/m - n^2/m^2 + (n^2-n)/m^2 = (mn-n^2+n^2-n)/m^2 = n*(m-1)/m^2 \)

2. There output of any merge algorithm is a list \( x_1,...,x_m \), where
\( m = n_a + n_b \), and the list of a's is a sublist of \( x_1,...,x_m \) of length \( n_a \).

There are \( \mathbb{C}(n_a + n_b \ n_a) \) ways to choose this sublist, so the algorithm must be able to distinguish between \( \mathbb{C}(n_a + n_b \ n_a) \) outputs. The corresponding decision tree must have at least \( \mathbb{C}(n_a + n_b \ n_a) \) leaves. Any binary tree of \( \geq \mathbb{C}(n_a + n_b \ n_a) \) leaves must have have height \( \geq \log_2 \mathbb{C}(n_a + n_b \ n_a) \).