Dennis Kuzminer
CSCI-UA 310-001 PS6

1.

a. For any given pair, the probability of $X_i = X_j$ is 1/m. Considering all combinations, we take the union of the sets. We can see that from rewriting the union bound, the total number of $X_i = X_j$ is n(n-1)/2 (n-i combos for X_i, summing yeilds n(n-1)/2). Combining the two, we can bound the overall probability of a collision by $p_{n,m} \leq \frac{n(n-1)}{2m}$.

b. 1 - p = q, where q is Pr[$X_i \neq X_j$]. We can enumerate all options and possibilities of X_i and X_j being distinct $\frac{m-0}{m} * \frac{m-1}{m} * \ldots * \frac{m-n}{m}$ (Mutual independence)→ $\prod_{i=1}^{n} \frac{m-(i-1)}{m} \rightarrow \prod_{i=1}^{n} \frac{m}{m} - \frac{(i-1)}{m} \rightarrow \prod_{i=1}^{n} 1 - \frac{(i-1)}{m} \rightarrow \prod_{i=1}^{n} (1 - \frac{(i-1)}{m}) = 1 - p.$

c. $1 + x \leq e^x \rightarrow 1 - e^x \leq -x \rightarrow$ We know that $p_{n,m} \leq \frac{n(n-1)}{2m} \rightarrow -p_{n,m} \geq -\frac{n(n-1)}{2m} \rightarrow$ Stubstitute the first inequality with the second, $x = -p_{n,m}$. This results in $1 - e^{-\frac{n(n-1)}{2m}} \leq p_{n,m}$.

d. Prove that $-\frac{n(n-1)}{2m} \leq ln(.5),$ as $p_{n,m} \geq 1 - e^{ln(.5)} \rightarrow p_{n,m} \geq 1 - .5 \rightarrow p_{n,m} \geq .5$
Given that $n \geq \sqrt{2ln(2)m} + 1 \rightarrow n(n-1) \geq (n-1)^2 \geq 2ln(2)m \rightarrow$ $\frac{n(n-1)}{2m} \geq \frac{(n-1)^2}{2m} \geq ln(2) \rightarrow -\frac{n(n-1)}{2m} \leq -ln(2) \rightarrow -\frac{n(n-1)}{2m} \leq ln(1/2).$ Because this exponent is bounded, we can substitute and see that $p_{n,m} \geq .5$.