CSE512 Fall 2018 Machine Learning - Homework 1

Your Name: Caitao Zhan

Solar ID: 111634527

NetID email address: caitao.zhan@stonybrook.edu

Names of people whom you discussed the homework with: Wenzhe Cui

1 Question 1

1. The expectation: E(X) = $\frac{N^2-1}{6N}$

   The problem is discrete. There are N^2 cases with equal probability, since X1 and X2 are independent and both are uniformly distributed.

   When $X2 \leq X1, X = 0$, so only need to consider cases all cases where $X2 > X1$.
   $P(X = 1) = \frac{N-1}{N^2}$
   $P(X = 2) = \frac{N-2}{N^2}$
   $P(X = i) = \frac{N-i}{N^2}$
   $E(X) = \sum_{i=1}^{N} iP(i) = \sum_{i=1}^{N} i \cdot \frac{N-i}{N^2} = \frac{N^2-1}{6N}$

2. The variance: Var(X) = $\frac{2N^4-N^2-1}{36N^2}$

   $Var(X) = E(X^2) - E(X)^2$
   We already derived $E(X)$, so we only need to derive $E(X^2)$. We just need to slightly alter the process of computing $E(X)$
   $P(X = 1^2) = \frac{N-1}{N^2}$
   $P(X = 2^2) = \frac{N-2}{N^2}$
   $P(X = i^2) = \frac{N-i}{N^2}$
   $E(X^2) = \sum_{i=1}^{N} i^2 P(i) = \sum_{i=1}^{N} i^2 \cdot \frac{N-i}{N^2} = \frac{N^2-1}{12}$
   $Var(X) = E(X^2) - E(X)^2 = \frac{2N^4-N^2-1}{36N^2}$

3. The covariance: Cov(X, X1) = $\frac{-N^2+1}{24}$

   $Cov(X, X1) = E[(X - E(X) \cdot (X1 - E(X1))] = E(X \cdot X1) - E(X)E(X1)$
   We already derived $E(X)$, and $E(X1) = \frac{N+1}{2}$. So we need to derive $E(X \cdot X1)$

$$
\begin{aligned}
E(X \cdot X1) &= \frac{1}{N^2} \sum_{X1=1}^{N} \sum_{X2=1}^{N} [max(X2 - X1, 0) \cdot X1] \\
&= \frac{1}{N^2} \sum_{X1=1}^{N} \sum_{X2=X1+1}^{N} [(X2 - X1) \cdot X1] + \frac{1}{N^2} \sum_{X1=1}^{N} \sum_{X2=1}^{X1} \cdot 0 \\
&= \frac{1}{N^2} \sum_{X1=1}^{N} [\frac{(X1 + 1 + N)(N - X1)}{2} - X1(N - X1)]X1 \\
&= \frac{1}{N^2} \sum_{X1=1}^{N} [X1^3 - (2N + 1)X1^2 + (N^2 + N)X1] \\
&= \frac{N^3 + 2N^2 - N - 2}{24N}
\end{aligned}
$$

   $Cov(X, X1) = Cov(X, X1) - E(X)E(X1) = \frac{N^3+2N^2-N-2}{24N} - \frac{N^2-1}{6N} \cdot \frac{N+1}{2} = \frac{-N^2+1}{24N}$

2 Question 2

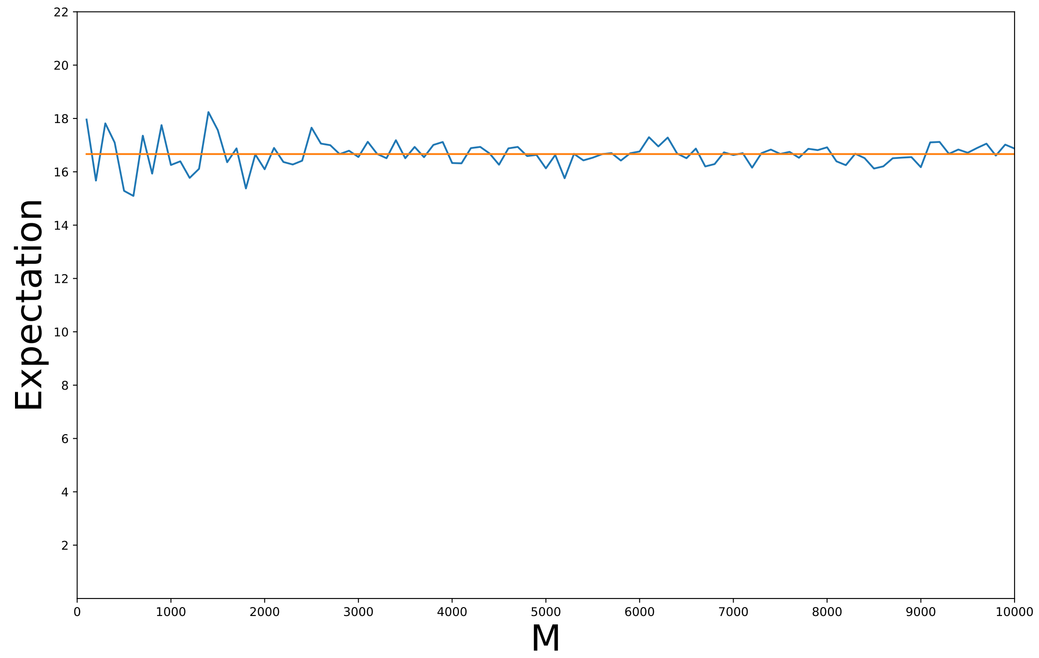

Figure 1: Expectation as a function of M. Theoretic E(X) = 16.665

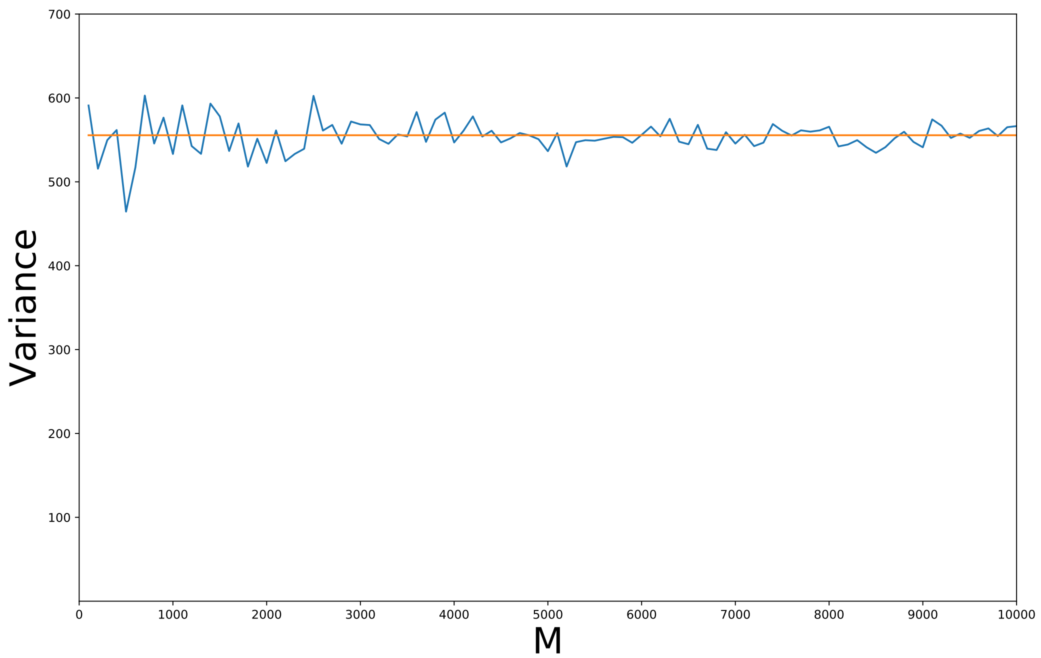

Figure 2: Variance as a function of M. Theoretic Var(X) = 555.527775

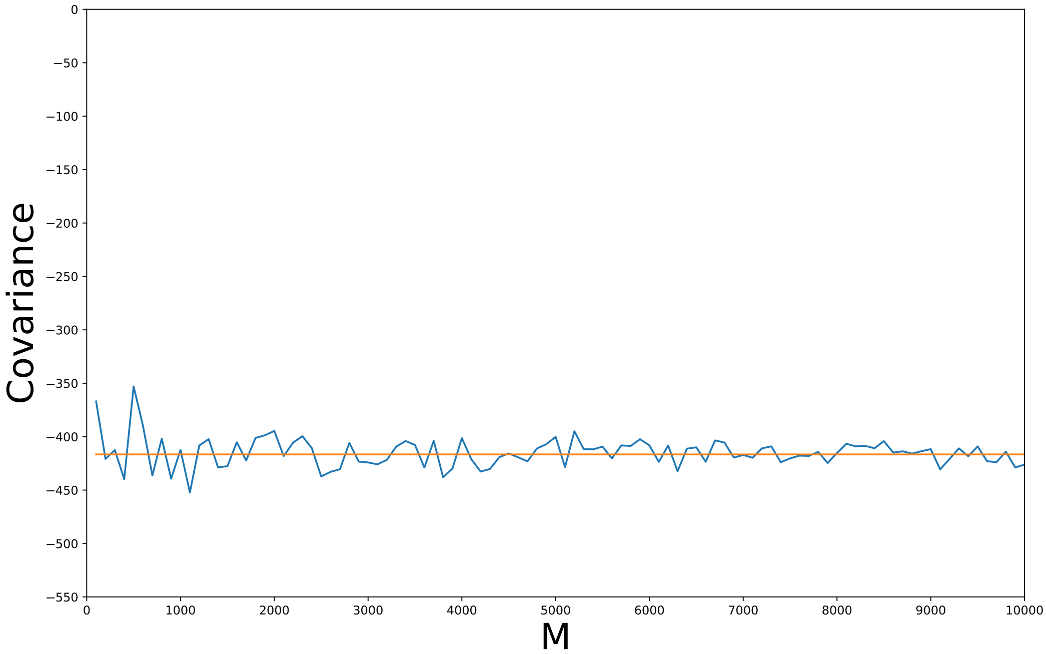

Figure 3: Variance as a function of M. Theoretic Cov(X) = -416.625

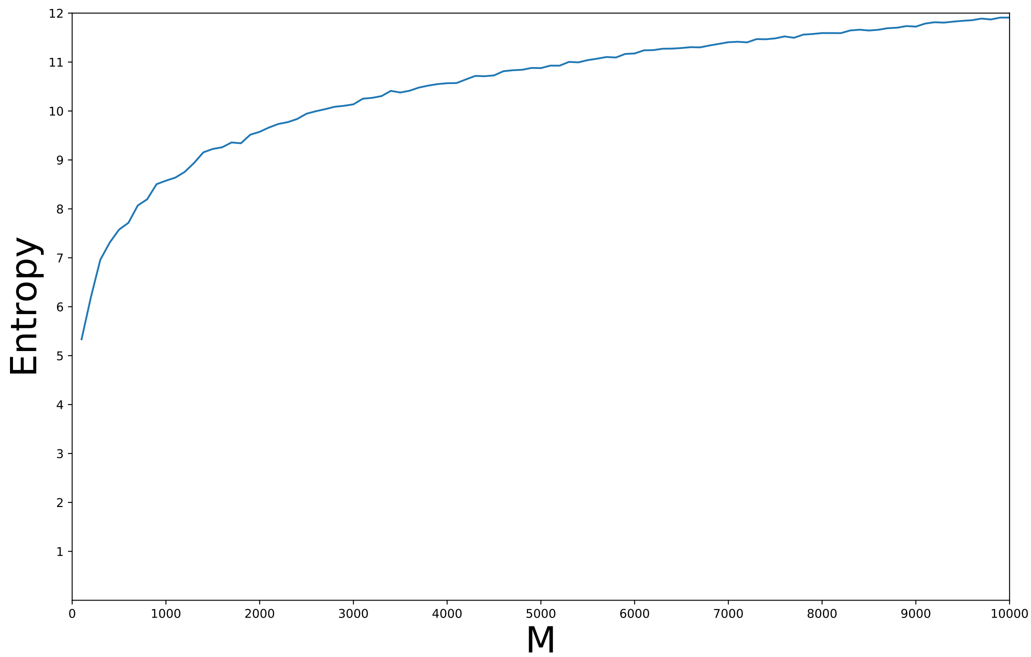

Figure 4: Entropy as a function of M.

Note: The python plotting script is in the jupyter notebook.