1 CSE512 Fall 2018 - Machine Learning - Homework 1

Name: Rohit Bhal
Solar ID: 112073893
NetID email address: rbhal@cs.stonybrook.edu
Names of people whom you discussed the homework with: Sayontan

2 Question 1 : Probability

2.1 1.1

Consider a positive integer N and let X1 and X2 be independent, discrete random variables uniformly distributed from 1 to N. Let

$$X = max(X1, X2) - X1$$

. Compute:
1. (20 points) The expectation: E(X).
2. (20 points) The variance: Var(X).
3. (20 points) The covariance: Cov(X, X1).

Answer: Let

$$\begin{aligned} g(X1, X2) &= X2 - X1 \quad , \quad X2 >= X1 \\ &= 0 \quad , \quad X1 > X2 \end{aligned} \tag{1}$$

Now,

$$Pr(X1 = xi) = 1/N; \tag{2}$$

$$Pr(X2 = xi) = 1/N; \tag{3}$$

Since, X1 & X2 are independent random variables uniformly distributed. Therefore,

$$\begin{aligned} Pr(X1 = xi, X2 = yi) &= Pr(X1 = xi) * Pr(X2 = yi) \\ &= (1/N) * (1/N) \\ &= \frac{1}{N^2} \end{aligned} \tag{4}$$

Also, the Pr(X1,X2) = 0 for X1 < X2.

Therefore, E(X) can be written as :

$$
\begin{aligned}
E(X) &= \sum_{X2=1}^{n} \sum_{X1=1}^{X2} Pr(X1, X2) * g(X1, X2) \\
&= \sum_{X2=1}^{n} \sum_{X1=1}^{X2} \frac{1}{N^2} * (X_2 - X_1) \\
&= \frac{1}{2 * N^2} \sum_{X2=1}^{n} X_2^2 - X_2 \\
&= \frac{1}{2 * N^2} * [\frac{(N * (N+1) * (2 * N + 1))}{6} - \frac{(N * (N+1)}{2}] \\
&= \frac{N * (N+1)}{2 * N^2} * [\frac{(2 * N + 1))}{6} - \frac{1}{2}] \\
&= \frac{N^2 - 1}{6 * N}
\end{aligned} \tag{5}
$$

Now,

$$
Var(X) = E(X^2) - E(X)^2 \tag{6}
$$

$$
\begin{aligned}
E(X^2) &= \sum_{X2=1}^{n} \sum_{X1=1}^{X2} Pr(X1, X2) * g(X1, X2)^2 \\
&= \sum_{X2=1}^{n} \sum_{X1=1}^{X2} \frac{1}{N^2} * (X_2 - X_1)^2 \\
&= \frac{1}{N^2} \sum_{X2=1}^{n} \frac{X_2 * (X_2 + 1)(2 * X_2 + 1)}{6} + X_2^3 - \frac{(2 * X_2 * X_2 * (X_2 + 1))}{2} \\
&= \frac{1}{6 * N^2} \sum_{X2=1}^{n} [2 * (X_2^3) - 3 * (X_2^2) + X_2] \\
&= \frac{1}{6 * N^2} [\frac{2 * N^2 * (N+1)^2}{4} - \frac{(3 * N * (N+1) * (2 * N + 1)}{6} - \frac{N * (N+1)}{2}] \\
&= \frac{N^2 - 1}{12}
\end{aligned} \tag{7}
$$

$$
\begin{aligned}
Var(X) &= E(X^2) - E(X)^2 \\
&= \frac{N^2-1}{12} - \frac{(N^2-1)^2}{(6*N)^2} \\
&= \frac{N^2-1}{12}[\frac{1}{1} - \frac{N^2-1}{3*N^2}] \\
&= \frac{N^2-1}{12}[\frac{3*N^2-N^2+1}{3*N^2}] \\
&= \frac{(2*N^2+1)*(N^2-1)}{36*N^2}
\end{aligned} \tag{8}
$$

Now,

$$
Covariance(X, X1) = E(X*X1) - E(X)*E(X1) \tag{9}
$$

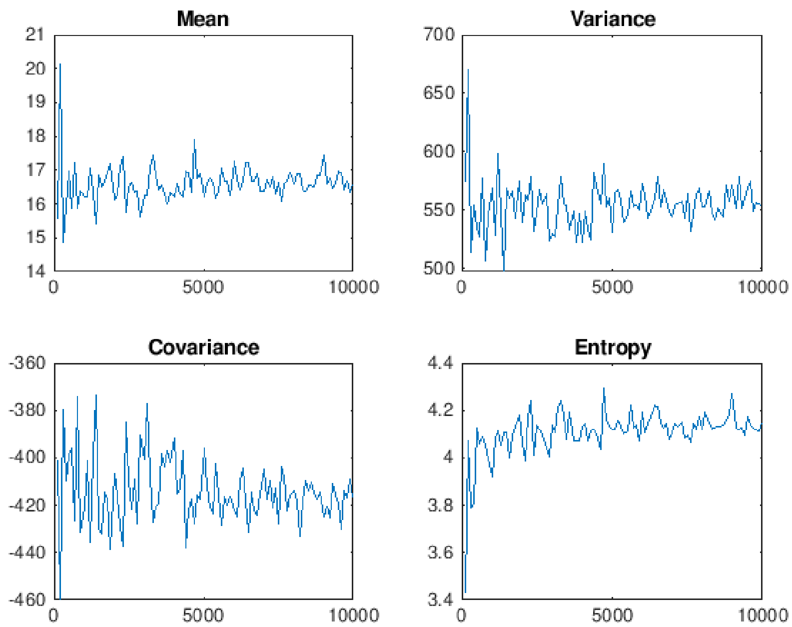
$$
\begin{aligned}
E(X*X1) &= \frac{1}{N^2}\sum_{X2=1}^{n}\sum_{X1=1}^{X2}(X_2*X_1 - X_1^2) \\
&= \frac{1}{N^2}\sum_{X2=1}^{n}\frac{X_2*X_2*(X_2+1)}{2} - \frac{X_2*(X_2+1)(2*X_2+1)}{6} \\
&= \frac{1}{6*N^2}\sum_{X2=1}^{n}(X_2^3 - X_2) \\
&= \frac{1}{6*N^2}[\frac{N^2*(N+1)^2}{4} - \frac{N*(N+1)}{2}] \\
&= \frac{N+1}{12*N}[\frac{N^2+N-2}{2}]
\end{aligned} \tag{10}
$$

$$
\begin{aligned}
Covariance(X, X1) &= E(X*X1) - E(X)*E(X1) \\
&= \frac{(N+1)*(N^2+N-2)}{24*N} - \frac{(N^2-1)*(N+1)}{12*N} \\
&= \frac{N^2-1}{12*N}[\frac{((N+2)}{2} - \frac{(N-1)}{1}] \\
&= \frac{(1-N^2)}{24}
\end{aligned} \tag{11}
$$

2.2 Question 2: Programming40 Points

Consider the random variables X1, X2, X defined as in Question 1. Write a Matlab (or equivalent Python) function with the below signature: [E, V, C, H] = question2(N, M) The inputs to the above function are positive integers N and M. M is the number of sample pairs of (X1, X2). The outputs of the function are four scalars E, V, C, H for E(X), Var(X), Cov(X, X1), and the entropy H(X). Plot E, V, C, H as the function of M = 100, 200, 300, $\cdots$, 9900, 10000 for N = 100. Produce four separate plots for E, V, C, H. You are allowed to use built-in functions of Matlab such as mean, var, cov. Other programming languages might have similar built-in functions, and you can use them if you want.

2.2.1 Plot



2.2.2 Code

```
function [MN,V,CV,H] = question2(N,M)

MN = [];
V = [];
CV = [];
H = [];

for m = 100 : 100: M
     R = unidrnd(N,m,2);
     X1 = R(:,1);
     X2 = R(:,2);
     X = max(X1,X2)-X1;
     MN = [MN, mean(X)];
     V = [V, var(X)];
     covariance = cov(X,X1);
     CV = [CV, covariance(1,2)];
     bs = transpose(repelem(N,size(X,1)));
     r = size(X,1);
     prob = [];
     found = [];
     for i = 1:r
          if(sum(found == X(i,1)) == 0)
               prob(i,1) = (sum(X == X(i,1))/r);
               found = [found, X(i,1)];
          else
               ;
          end
     end
     prob = prob(prob>0);
     H = [H, -sum( prob .* log2(prob) )];
end
%x = linspace(1,10,50);

figure
subplot(2,2,1)        % add first plot in 2 x 2 grid
a = linspace(100,M,size(MN,2));
b = MN;
plot(a,b)
title('Mean')

subplot(2,2,2)        % add first plot in 2 x 2 grid
a = linspace(100,M,size(V,2));
b = V;
plot(a,b)
```

```
title('Variance')

subplot(2,2,3)          % add first plot in 2 x 2 grid
a = linspace(100,M,size(CV,2));
b = CV;
plot(a,b)
title('Covariance')

subplot(2,2,4)          % add first plot in 2 x 2 grid
a = linspace(100,M,size(H,2));
b = H;
plot(a,b)
title('Entropy')


end
```