Cover page for answers.pdf
CSE512 Fall 2018 - Machine Learning - Homework 1

Your Name: Sriram Reddy Kalluri

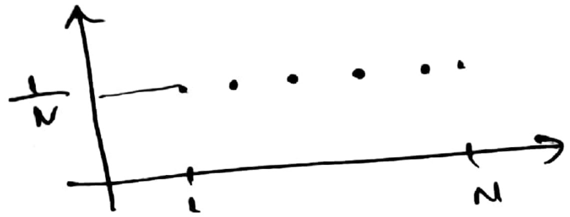
Solar ID: 111878857

NetID email address: skalluri@cs.stonybrook.edu & sriram.kalluri@stonybrook.edu

Names of people whom you discussed the homework with:

1.1) After discussing with professor, I have taken the discrete uniform random variables for integers in range [1,N]. So,

X_1 and X_2 has the following distribution.



Validating the summation property i.e $\sum_{-\infty}^{\infty} P_x(x) = 1$ for ay above R.V's. $\frac{1}{N} \sum_{i=1}^{N} 1 = N \times \frac{1}{N} = 1$.

Coming to the distribution of X i.e. $\max(X_1, X_2) - X_1$.

Both the RV's have same probability for all values [1,N].

So, generating the random variable X distribution for all values of X_1 and X_2 in range [1,N].

X_2 \ X_1	1	2	3	4	…	N
1	0	0	0	0	…	0
2	1	0	0	0	…	0
3	2	1	0	0	…	0
4	3	2	1	0	…	0
⋮					⋱	0
N	N−1	N−2	N−3	N−4		0

So, checking all the possible values of x and counting the number of times x occured in the distribution, I have generated the probability distribution for x

$$
x = \begin{cases}
\text{Value} & \text{probability} \\
0 & \frac{n^2+n}{2n^2} \\
1 & \frac{n-1}{n^2} \\
2 & \frac{n-2}{n^2} \\
\vdots & \\
k & \frac{n-k}{n^2} \\
\vdots & \\
N-1 & \frac{1}{n^2}
\end{cases}
$$

$\boxed{\text{here } n = N}$

checking whether it is valid pdf or not, again using the summation property of probabilities,

$$\sum_{-\infty}^{\infty} P_x(x) = \sum_{k=0}^{N-1} P_x(x=k) = \frac{n^2+n}{2n^2} + \sum_{k=1}^{n-1}\left(\frac{n-k}{n^2}\right)$$

$$= \frac{n^2+n}{2n^2} + \frac{n(n-1)}{n^2} + - \frac{(n-1)n}{2n^2}$$

$$= \frac{n^2+n + 2n^2-2n - n^2 + n}{2n^2}$$

$$= \frac{2n^2}{2n^2} = 1.$$

It's a valid PDF.

1.1). Calculating the expectation of x.

$$E[x] = \sum_{-\infty}^{\infty} x P_x(x),$$

$$= 0 \cdot \left(\frac{n^2+n}{2n^2}\right). + \sum_{k=1}^{N-1} \frac{k(N-k)}{N^2}$$

$$= \frac{1}{N^2} \cdot \sum_{k=1}^{N-1} (N \cdot k - k^2)$$

$$= \frac{1}{N^{\cancel{2}}}\left[\frac{N(N-1).\cancel{N}}{2} - \frac{(N-1)\cancel{N}(2N-1)}{6}\right]$$

$$= \frac{1}{N}\left[\frac{3N^2-3N-(2N^2-3N+1)}{6}\right]$$

$$= \left[\frac{N^2-1}{6N}\right]$$

1.2). Calculating the variance of x.

$$Var(x) = E[x^2] - (E[x])^2.$$

First calculating $E[x^2]$,

$$E[x^2] = \sum_{-\infty}^{\infty} x^2 P(x).$$

$$= 0\left(\frac{N^2+2N}{2N^2}\right). + \sum_{k=1}^{N-1} \frac{k^2(N-k)}{N^2}$$

Scanned by CamScanner

$$= \frac{1}{N^2} \sum_{k=1}^{N-1} k^2 (N-k).$$

$$= \frac{1}{N^2} \left[\frac{N(N-1)N(2N-1)}{6} - \frac{((N-1)N)^2}{4} \right]$$

$$= [N-1] \left[\frac{2N-1}{6} - \frac{N-1}{4} \right].$$

$$= [N-1] \left[\frac{4N-2-(3N-3)}{12} \right]$$

$$= [N-1] \left[\frac{4N-2-3N+3}{12} \right] = \frac{N^2-1}{12}.$$

$$E[x^2] = \frac{N^2-1}{12}.$$

$$Var(x) = E[x^2] - (E[x])^2 = \left[\frac{N^2-1}{12}\right] - \left[\frac{N^2-1}{6N}\right]^2.$$

$$= \frac{(N^2-1)}{12} \left[1 - \frac{N^2-1}{3N^2} \right].$$

$$= \left(\frac{N^2-1}{12}\right) \left[\frac{3N^2 - N^2 + 1}{3N^2} \right].$$

$$= \left[\frac{N^2-1}{12}\right] \left[\frac{2N^2+1}{3N^2} \right].$$

1.3). Covariance of. x, x_1

$$= E\left[(x - E[x])(x_1 - E[x_1])\right].$$

$$= E[x x_1] - E[x]\,E[x_1].$$

To calculate the expectation of $E[x x_1]$; I'm generating the distribution of $x x_1$ by checking the possible values for $x x_1$ based on x_1 and x_2 as x is a function of x_1 and x_2. Possibilities of $x x_1$ based on x_1 and x_2 is given below

x_2 \ x_1	1	2	3	4	5	6	...	N
1	0	0	0	0	0	.		0
2	1	0	0	0	0			
3	2	2	0	0	0			
4	3	4	3	0	0			
5	4	6	6		0			
6	5	8	9					
⋮								
N	N−1	2(N−2)	3(N−3)	...			(N−1)	0

The expectation of XX_1 can be written as

$$= \frac{\text{sum of all possible values}}{N^2}.$$

Here you can observe a pattern for different vertical columns.

$E[XX_1]$ = Summing all the columns seperately and adding them.

$$= \frac{1}{N^2}\left[1\cdot\frac{(N-1)N}{2} + 2\cdot\frac{(N-2)(N-1)}{2} + 3\frac{(N-3)(N-2)}{2} + \dots\dots + (N-2)\frac{3\cdot 2}{2} + (N-1)\cdot\frac{2\times 1}{2}\right]$$

$$= \frac{1}{2N^2}\sum_{i=1}^{N-1} i(N-i)(N-i+1).$$

$$= \frac{1}{2N^2}\sum_{i=1}^{N-1} i\left((N-i)^2 + N-i\right). = \frac{1}{2N^2}\sum_{i=1}^{N-1}\left[i\,(N^2+i^2-2iN-i+N)\right]$$

$$= \frac{1}{2N^2}\sum_{i=1}^{N-1} i\left[(N^2+N) - (2N+1)i + i^2\right]$$

$$= \frac{1}{2N^2}\left[(N^2+N)\frac{(N-1)N}{2}\right. - (2N+1)\frac{(N-1)N(2N-1)}{6} \left. + \frac{((N-1)N)^2}{4}\right].$$

$$= \frac{(N-1)N}{2N^2}\left[\frac{N^2+N}{2} - \left[\frac{4N^2-1}{6}\right] + \frac{N^2-N}{4}\right]$$

$$= \frac{(N-1)N}{2N^2}\left(\frac{6N^2+6N-[8N^2-2]+3N^2-3N}{12}\right)$$

$$= \frac{(N-1)N}{2N^2}\left[\frac{N^2+3N+2}{12}\right]$$

$$E[x x_1] = \frac{(N-1)N(N+1)(N+2)}{24N^2}.$$

Covariance of x, x_1

$$Cov(x,x_1) = E[x x_1] - E[x]E[x_1]$$

$$= \frac{(N-1)\cancel{N}(N+1)(N+2)}{24N^{\cancel{2}}} - \frac{(N^2-1)}{6N}\left(\frac{N+1}{2}\right)$$

$$= \frac{(N-1)(N+1)}{12N}\left[\frac{N+2}{2} - N+1\right].$$

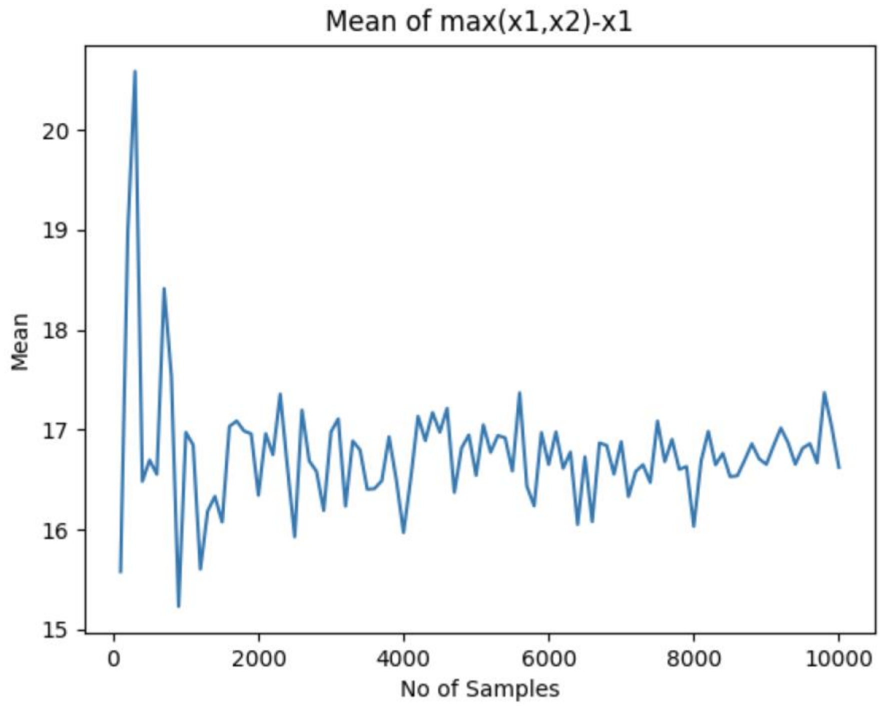
$$= \frac{(N-1)(N+1)}{12N}\left[\frac{N+2-2N-2}{2}\right].$$

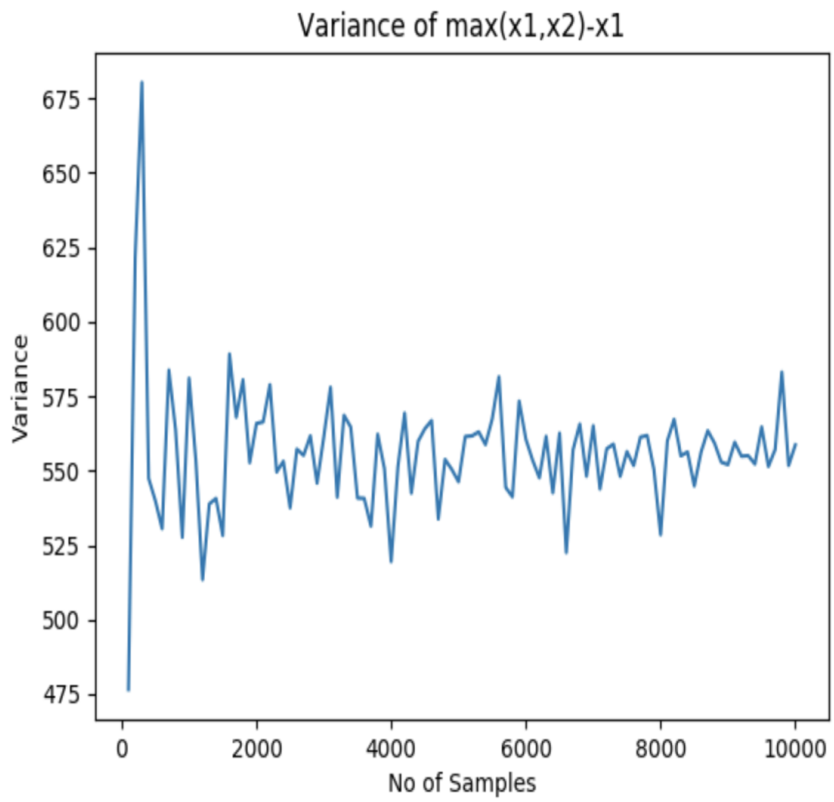
$$= -\frac{(N-1)(N+1)}{24}.$$

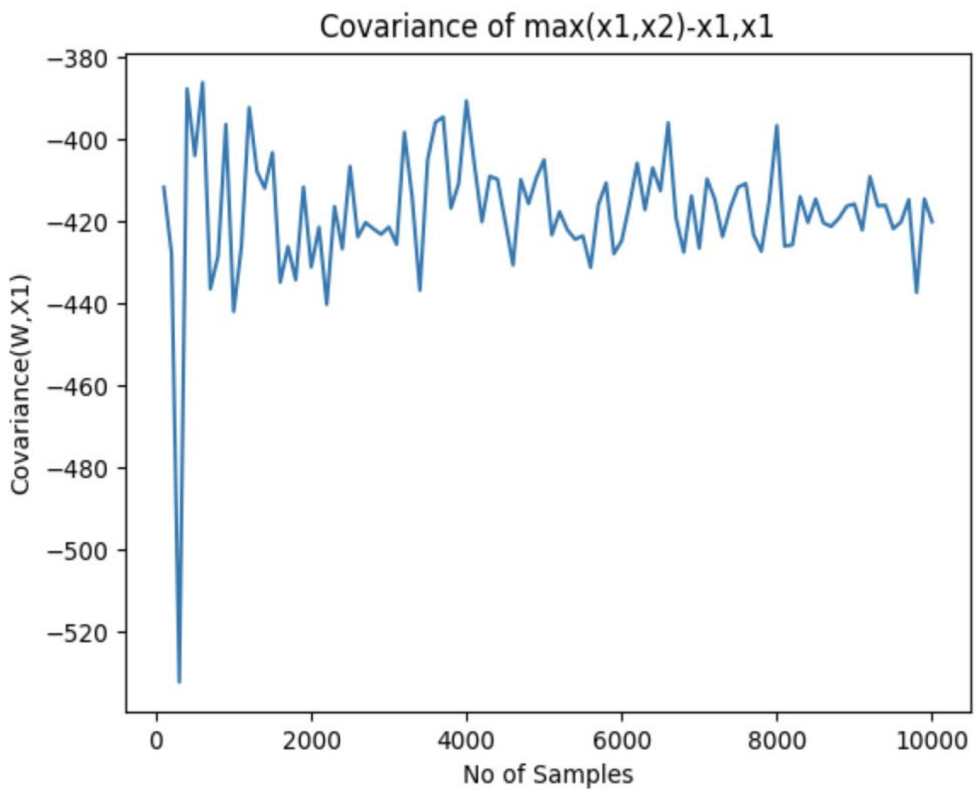
2 Question 2 – Programming

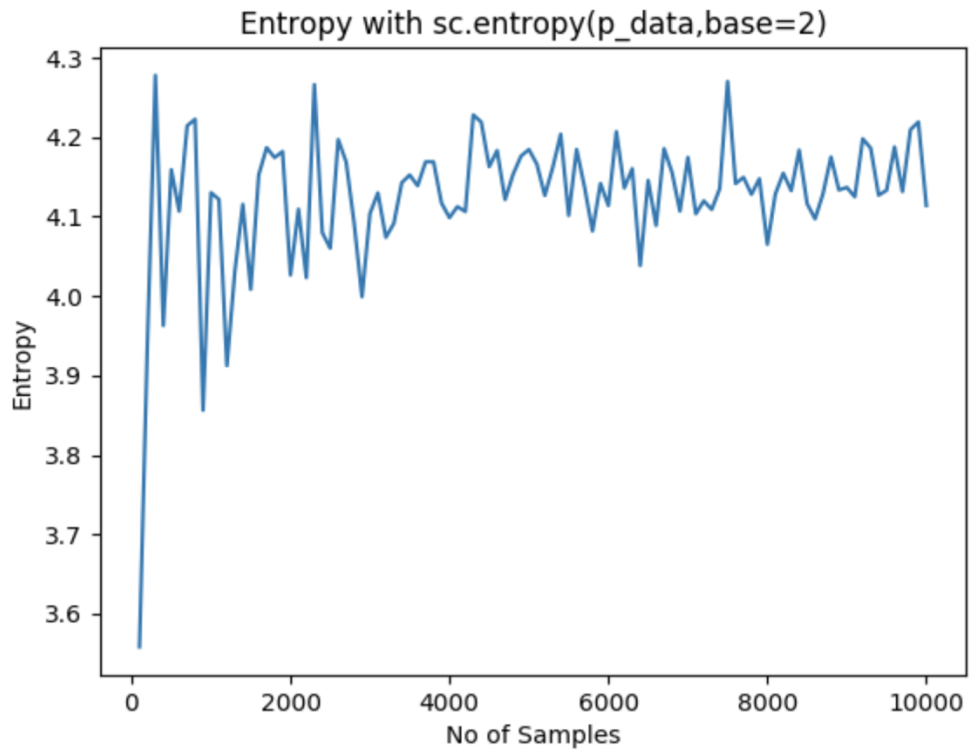
After having discussion with professor, I had used only integers for discrete random uniform distribution. I have implemented the program in python using numpy.

Please find the results below:









Please find the values below:

Mean	16.761206618894974
Variance	555.9771692565092

Covariance	-417.74842329795024
Entrophy	4.125863891073738