CSE512 Fall 2018 - Machine Learning - Homework 1

Name : Astitv Nagpal

Solar ID : 112008011

NetID email address : anagpal@cs.stonybrook.edu

Names of people whom you discussed the homework with :
Sukriti Mittal

→ Solving for question 1.

(*) Expectation $E(X)$ || Variance $Var(X)$ || Co-Variance $Cov(X, X_1)$

The given information is:

→ X_1, X_2 → independent, discrete random variables, uniformly distributed.

→ N → number over which X_1, X_2 are uniformly distributed.

→ $f(x) \Rightarrow \boxed{X = \max(X_1, X_2) - X_1}$

⇒ Trying to work around with our $f(x)$ & looking for patterns & flow.

* Example / sample 1: Lets consider $N = 5$.

Hence, we have $X_1 = [1, 2, 3, 4, 5]$

$X_2 = [1, 2, 3, 4, 5]$

→ $f(x) = \max(X_1, X_2) - X_1$ → from this we can conclude a point which says value of "X" will always be greater than equal to "0" & less than or equal to "N - 1"

→ We will now observe how $f(x)$ will react against values of x_1 & x_2.

→ We have :

$$f(x) = x = \begin{cases} 0 & ; \; x_1 \geq x_2 \\ x_2 - x_1 & ; \; x_1 < x_2 \end{cases}$$

∴ In our example we have $N = 5$, hence we have.

Note : x can reach a maximum of '$N-1$' as value

$(x_1, x_2) \longrightarrow$

$$x = 0 \begin{cases} (1,1) \\ (2,1) \quad (2,2) \\ (3,1) \quad (3,2) \quad (3,3) \\ (4,1) \quad (4,2) \quad (4,3) \quad (4,4) \\ (5,1) \quad (5,2) \quad (5,3) \quad (5,4) \quad (5,5) \end{cases}$$

Total = 15 values.

$$x = 1 \begin{cases} (1,2) \\ (2,3) \\ (3,4) \\ (4,5) \end{cases}$$

Total = 4 values.

$x = 2 \begin{cases} (1, 3) \\ (2, 4) \\ (3, 5) \end{cases}$

Total = 3 values.

$x = 3 \begin{cases} (1, 4) \\ (2, 5) \end{cases}$

Total = 2 values.

$x = 4 \begin{cases} (1, 5) \end{cases}$

Total = 1 value.

→ Hence, our table become. for N = 5.

$x = 0 \rightarrow$ 15 times
$x = 1 \rightarrow$ 4 times
$x = 2 \rightarrow$ 3 times
$x = 3 \rightarrow$ 2 times
$x = 4 \rightarrow$ 1 times

Total frequency = 25

→ From the above data we can compute & generalize the propability of $x \rightarrow Pr(x)$.

$x = 0 \rightarrow P_R = \cancel{\text{}}\ 15/25$; $x = 2 \rightarrow P_R = 3/25$; $x = 4 \rightarrow P_R = 1/25$

$x = 1 \rightarrow P_R = 4/25$; $x = 3 \rightarrow P_R = 2/25$

→ For any number between [1, N−1] we have a pattern.

$$Pr(X = x) = \frac{N-x}{N^2}$$

→ $Pr(X = 0) = \frac{N(N+1)}{2N^2}$

Note:
Here we are not worried about the $Pr(X=0)$ because expectation will always cancel out to zero.

⇒ Hence, the probability of any number ranging from

$[1, N-1]$ ⇒ $\boxed{Pr(X=x) = \frac{N-x}{N^2}}$ [My Generalized Formula.]

① Finding the expectation of X → E(X).

→ Mathematically we know :

$$E(X) = \sum_{x=1}^{N} x \cdot Pr(X=x)$$ —— (i) formula for expectation.

∴ Substituting the values in formula (i) we get,

$$E(X) = \sum_{x=0}^{N-1} x\frac{(N-x)}{N^2}$$ ⟶ now opening the formula & solving for X.

$$= \sum_{x=0}^{N-1} \frac{Nx}{N^2} - \sum_{x=0}^{N-1} \frac{x^2}{N^2}$$ ⟶ here 'N' is a constant hence, we can take it out.

$$= \frac{1}{N}\sum_{x=0}^{N-1} x - \frac{1}{N^2}\sum_{x=0}^{N-1} x^2$$

We know,

$$\sum_{x=1}^{N}(x) = \frac{N(N+1)}{2} \quad \& \quad \sum_{x=1}^{N}(x^2) = \frac{N(N+1)(2N+1)}{6}$$

→ Substituting the above formulas in our equation, we get :

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

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

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

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

⇒ Hence, the generic formula for Expectation of x is :

$$\boxed{E(x) = \frac{N^2-1}{6N} = \frac{(N+1)(N-1)}{6N}}$$ ⇒ Answer to part (1)

⇒ We can also verify by putting some value of N in it & verifying the value with code. For example if $\boxed{N=100}$

$$E(X) = \frac{N^2-1}{6N} = \frac{(N+1)(N-1)}{6N} = \frac{(101)(99)}{600} = \boxed{16.665}$$

⟹ This value can be verified through my code.

2 Finding the Variance of X → Var(X)

→ Mathematically we know :

$$Var(X) = E(X^2) - (E(X))^2 \quad \text{——— (i)}$$

→ And from earlier example we already know that.

$$E(X) = \sum_{\text{all } X} X \cdot P(X) \quad ; \quad P(X) = \frac{N-X}{N^2}$$

$$\boxed{E(X) = \frac{N^2-1}{6N}} \rightarrow \text{derived in last part.}$$

$$E(X^2) = \sum_{\text{all } X} X^2 \cdot P(X) \quad ;$$

→ Hence using the above formulas & substituting in equation (i) we get.

$$Var(X) = E(X^2) - (E(X))^2$$

$$= \sum_{X=0}^{N-1} X^2 \cdot \left(\frac{N-X}{N^2}\right) - \left[\frac{N^2-1}{6N}\right]^2$$

$$= \frac{1}{N^2} \sum_{X=0}^{N-1} NX^2 - \frac{1}{N^2} \sum_{X=0}^{N-1} X^3 - \left[\frac{N^2-1}{6N}\right]^2$$

$$= \frac{1}{N} \sum_{X=0}^{N-1} X^2 - \frac{1}{N^2} \sum_{X=0}^{N-1} X^3 - \left[\frac{(N+1)(N-1)}{6N}\right]^2$$

We know,

$$\sum_{\text{all } X} X^2 = \frac{N(N+1)(2N+1)}{6} \quad \& \quad \sum_{\text{all } X} X^3 = \left[\frac{N(N+1)}{2}\right]^2$$

→ Substituting the above formulas in our equation we get,

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

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

$$= \frac{2N^2+3N+1}{6} - \left[\frac{N^2+1+2N}{4}\right] - \left[\frac{(N+1)(N-1)}{6N}\right]^2$$

$$= \frac{4N^2+\cancel{6N}+2-3N^2-3-\cancel{6N}}{12} - \left[\frac{(N+1)(N-1)}{6N}\right]^2$$

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

$$= \frac{N^2-1}{12} - \left[\frac{N^4+1-2N^2}{36N^2}\right]$$

$$= \frac{3N^4-3N^2-N^4-1+2N^2}{36N^2}$$

$$= \frac{2N^4-N^2-1}{36N^2} \Rightarrow \boxed{Var(X) = \frac{2N^4-N^2-1}{36N^2}}$$

→ This can be verified by substituting a value of N. Lets say $N = 100$. $Var(x) = \frac{2\times(100)^4-(100)^2-1}{36\times(100)^2} = \underline{555.5277}$

This value can be verified from the code.

3 Finding the Co-Variance of $X, X_1 \rightarrow Cov(X, X_1)$

→ Mathematically we know if given two numbers A, B we have :

$$Cov(A, B) = E((A)(B)) - E(A)E(B)$$

→ In our case we have :

$A \rightarrow max(X_1, X_2) - X_1$

$B \rightarrow X_1$

$$\therefore Cov(X, X_1) = E((max(X_1, X_2) - X_1)X_1) - E(X)E(X_1)$$

→ We know ;

$$f(X) = X = \begin{cases} 0 ; & X_1 \geq X_2 \\ X_2 - X_1 ; & X_1 < X_2 \end{cases}$$

⇒ We can skip for this case as it is equal to 0.

$\therefore max(X_1, X_2) - X_1$ can be re-written as :

$\rightarrow X_2 - X_1 \quad ; \forall X_1 < X_2$. Hence, our equation is :

$$\boxed{Cov(X, X_1) = E((X_2 - X_1)X_1) - E(X)E(X_1) \quad \text{—— (i)}}$$

$$\boxed{Cov = E(X_1 X_2) - E(X_1^2) - E(X)E(X_1)}$$

★ → In our case we have constraints over $X_1 < X_2$.

Eg: $\cdot X_1 = 1 \rightarrow$ possible pairs $(X_1, X_2) \rightarrow (1, 2) (1.3) (1.4) \ldots (1.N)$

$\rightarrow [(N-1)$ terms$]$

$X_1 = 2 \rightarrow$ possible term $\rightarrow (2.3) (2.4) (2.5) \ldots (2.N)$

$\rightarrow [N-2$ terms$]$

$\vdots$

$X_1 = N-1 \rightarrow$ possible term $\rightarrow (N-1 . N)$

$\therefore$ this can be generalized as all the numbers have <u>equal likelihood</u> of all the events.

→ Hence for $x_1 = 1$, $x_1 = 2$, ... $x_1 = N-1$ can be written as :

$$E(x_1 x_2) = \frac{1}{N^2}\left[2+3+4+\cdots+N\right] + \frac{2}{N^2}\left[3+4+\cdots+N\right] + \cdots\cdots + \frac{N-1}{N^2}[N]$$

$\left\{ \begin{array}{l} \text{as } x_1 \text{ can be} \\ \text{a maximum of } N-1 \\ \text{as a value.} \end{array} \right.$

⇒ <u>We</u> know sum of a series with 'n' numbers & first element 'a' & last element 'l' is equal to :

$$\text{Sum} = \frac{n}{2}\left[a + l\right]$$ → Summation in a Arithmatic Progression.

→ Using the above formula to calculate $E(x_1 x_2)$, we get:

$$E(x_1 x_2) = \frac{1}{N^2}\left[\frac{N-1}{2}[N+2]\right] + \frac{2}{N^2}\left[\frac{N-2}{2}[N+3]\right] + \cdots\cdots + \frac{N-1}{N^2}\left[N \times \frac{2}{2}\right]$$

⇒ 2 was multiplied & divided so that we can take $\frac{1}{2N^2}$ common.

$$= \frac{1}{2N^2}\Big[1(N-1)(N+2) + 2(N-2)(N+3) + \cdots\cdots + (N-1)(N-(N-1))(N+(N-1)+1)\Big]$$

↳ written in a generalized way.

→ The equation can be written as :

$$\rightarrow \frac{1}{2N^2}\Big[x(N-x)(N+x+1)\Big]_{x=1}^{N-1}$$

→ Generalized form. Here we need to take summation of it.

$$\Rightarrow \sum_{x=1}^{N-1} \frac{1}{2N^2}\Big(x(N-x)(N+x+1)\Big)$$

$$\Rightarrow \frac{1}{2N^2}\sum_{x=1}^{N-1}\Big((Nx-x^2)(N+x+1)\Big)$$

$$= \frac{1}{2N^2}\sum_{x=1}^{N-1}\Big(N^2x + \cancel{Nx^2} + Nx - \cancel{x^2N} - x^3 - x^2\Big)$$

$$= \frac{1}{2N^2}\left[\sum_{x=1}^{N-1} N^2x + \sum_{x=1}^{N-1} Nx - \sum_{x=1}^{N-1} x^3 - \sum_{x=1}^{N-1} x^2\right]$$

$$= \frac{1}{2N^2}\left[N^2\sum_{x=1}^{N-1} x + N\sum_{x=1}^{N-1} x - \sum_{x=1}^{N-1} x^3 - \sum_{x=1}^{N-1} x^2\right] \text{ — (a)}$$

⇒ We know :

$$\sum_{x=1}^{N-1} x = \frac{N(N-1)}{2} \quad \Big| \quad \sum_{x=1}^{N-1} x^2 = \frac{N(N-1)(2N-1)}{6}$$

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

→ Substituting the formulas in equation (a) we get:

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

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

$$= \frac{N-1}{4N^2}\left[\frac{6N^2 + 6N^2 - 3N^3 + 3N^2 - 4N^2 + 2N}{6}\right]$$

$$= \frac{N-1}{4N^{\cancel{2}}}\left[\frac{\cancel{N}(6N^2 + 6N - 3N^2 + 3N - 4N + 2)}{6}\right]$$

$$= \frac{N-1}{24N}\left[3N^2 + 5N + 2\right] = E(x_1 x_2) \text{ ——— } (I)$$

* Now, Calculating $E(x,2)$: we have if:

$x_1 = 1 \rightarrow (1.2)\ (1.3) \ldots (1.N) \rightarrow N-1$ term

$x_1 = 2 \rightarrow (2.3)\ (2.4) \ldots (2.N) \rightarrow N-2$ terms.

$\vdots$

$x_1 = N-1 \rightarrow (N-1)(N) \longrightarrow N-(N-1)$ terms.

→ Here every number as equal likelihood of appearing & hence it can be generalized as:

→ ~~$\&$~~ We know $E(X_1^2) = \sum_{\text{all } n} x_1^2 \cdot P(x)$

→ $\frac{1^2(N-1)}{N^2} + \frac{2^2(N-2)}{N^2} + \cdots + \frac{(N-1)^2(N-(N-1))}{N^2}$

→ This can be re-written in a generalized way as:

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

$$= \frac{1}{N^2}\left[N\sum_{x=1}^{N-1} x^2 - \sum_{x=1}^{N-1} x^3\right] \quad \text{—— (b)}$$

→ We know:

$$\sum_{x=1}^{N-1} x^2 = \frac{N(N-1)(2n-1)}{6} \;\& \quad \sum_{x=1}^{N-1} x^3 = \left[\frac{N(N-1)}{2}\right]^2$$

→ Substituting the formulas in equation (b) we get:

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

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

$$= \frac{(N-1)(2N-1)}{6} - \frac{(N-1)^2}{4}$$

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

$$= \frac{8N^2 - \cancel{12}N + 4 - 6N^2 - 6 + 1\cancel{2}N}{24}$$

$$= \frac{2N^2 - 2}{24} = \boxed{\frac{N^2-1}{12} = E(x_1^2)} \text{ --- } II$$

$\Rightarrow$ We know $Cov(x, x_1) = E(x_1 x_2) - E(x_1^2) - E(x)\,E(x_1)$

Solving for $E(x_1 x_2) - E(x_1^2)$ we get.

$\rightarrow$ Substitute values from eq. I & II we get:

$$E(x_1 x_2) = \frac{(N-1)(3N^2 + 5N + 2)}{24N} \quad \Big| \quad E(x_1^2) = \frac{N^2-1}{12}$$

$\therefore$ Substituting them we get:

$$= (N-1)\left[\frac{3N^2 + 5N + 2}{24N} - \left(\frac{(N+1)(2N)}{24N}\right)\right]$$

$$= \frac{N-1}{24N}\left[3N^2 + 5N + 2 - 2N^2 - 2N\right]$$

$$= \frac{N-1}{24N}\left[N^2 + 3N + 2\right] \text{ --- } III \Rightarrow E(x_1 x_2) - E(x_1^2)$$

$\rightarrow$ Now we already know $E(x)$ as we calculated it in part 1 of question.

$$\boxed{E(x) = \frac{N^2-1}{6N}}$$

→ Now we need to calculate $E(x_1)$. In this case there is no restriction on x_1 & x_1 can range from $1 \rightarrow N$ with each number equally probable.

$$\therefore\ P(x_1) = \frac{1}{N}$$

$$\therefore\ E(x_1) = \sum_{x=1}^{N} x \cdot P(x_1) = \sum_{x=1}^{N} x \cdot \frac{1}{N}$$

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

$$\therefore\ \boxed{E(x_1) = \frac{N+1}{2}} \text{ ——— IV}$$

Now Substituting values III & IV in equation (i) we get:

$$Cov(x, x_1) = \frac{(N-1)(N^2+3N+2)}{24N} - \left(\frac{N^2-1}{6N}\right)\left(\frac{N+1}{2}\right)$$

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

$$= \frac{(N+1)(N-1)}{24N}\left[N+2 - 2(N+1)\right]$$

$$= \frac{(N+1)(N-1)}{24N}\left[N+2-2N-2\right]$$

$$\frac{(N+1)(N-1)}{24\cancel{N}}\left(-\cancel{N}\right)$$

$$\therefore \operatorname{Cov}(X, X_1) = \frac{(N+1)(N-1)(-1)}{24}$$

$$\boxed{\operatorname{Cov}(X, X_1) = \frac{1-N^2}{24}}$$

This can be verified by substituting N with a value.

Lets say $N = 100$

$$\therefore \operatorname{Cov}(X, X_1) = \frac{1-(100)^2}{24}$$

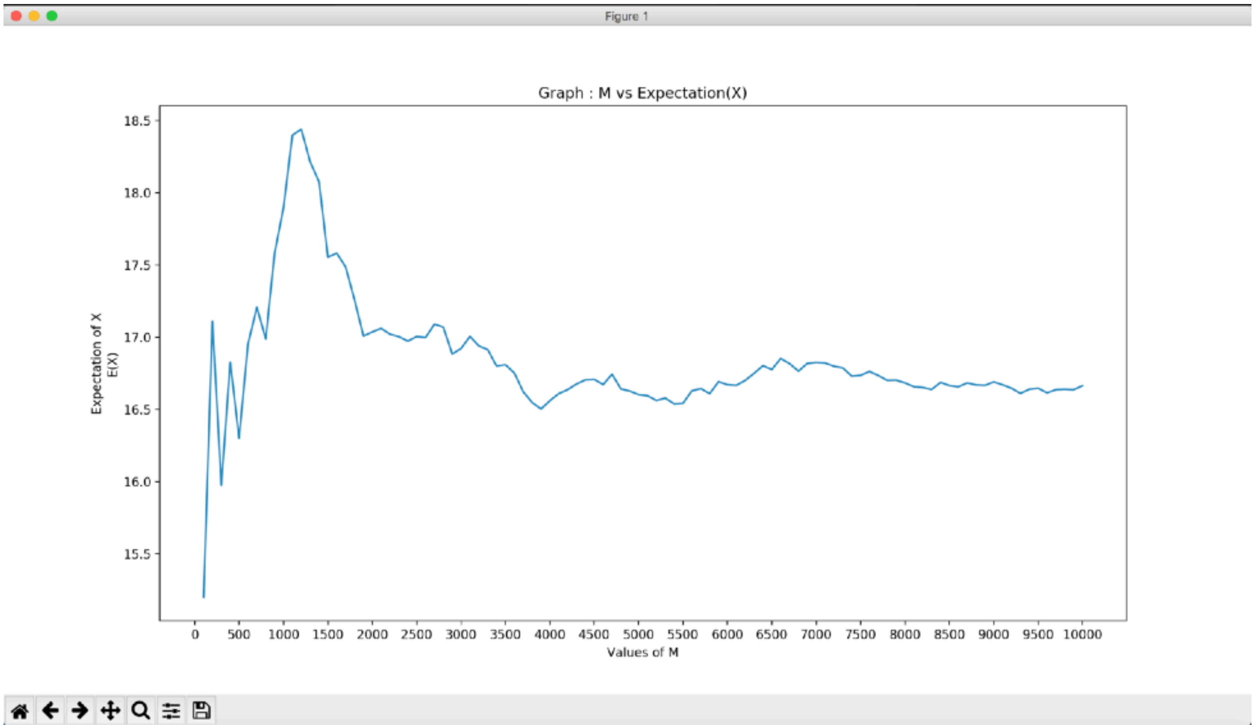
$$\operatorname{Cov}(X, X_1) = \underline{-416.625}$$

$\Rightarrow$ FINAL ANSWERS :

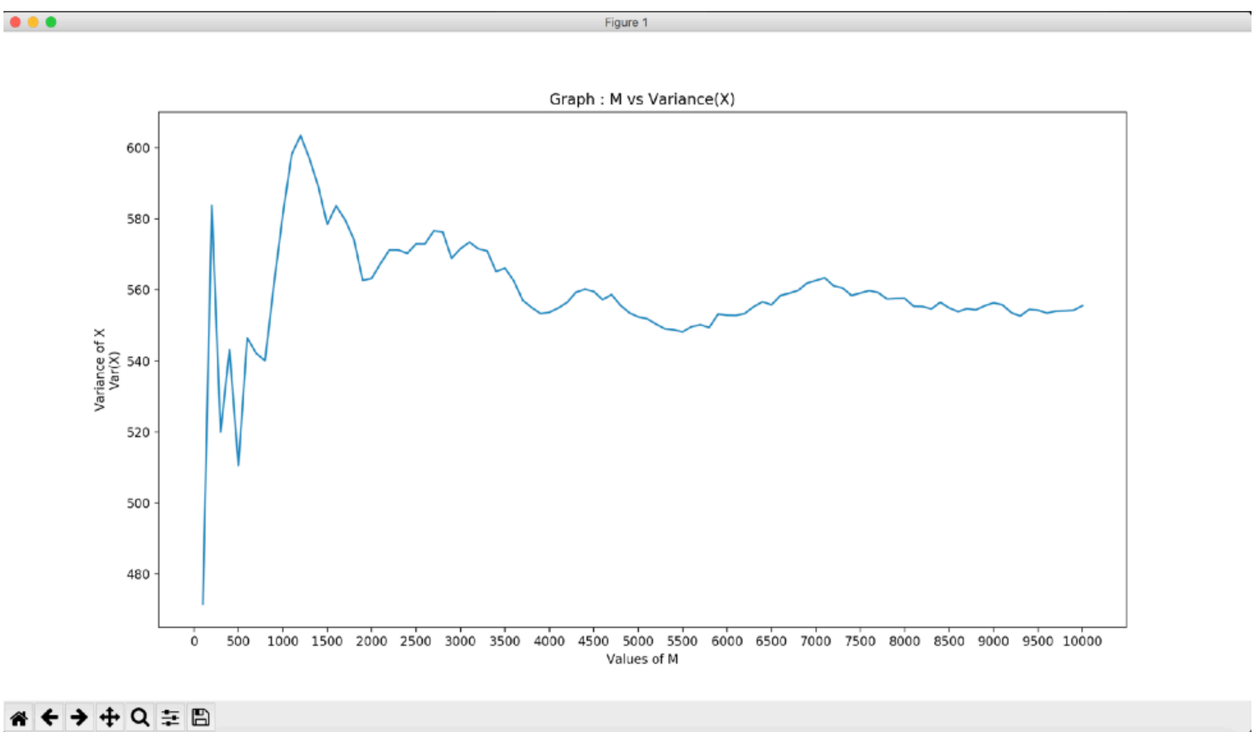
① $E(X) = \dfrac{N^2-1}{6N} = \dfrac{(N+1)(N-1)}{6N}$

② $\operatorname{Var}(X) = \dfrac{2N^4 - N^2 - 1}{36N^2} = \dfrac{(N^2-1)(2N^2+1)}{36N^2}$

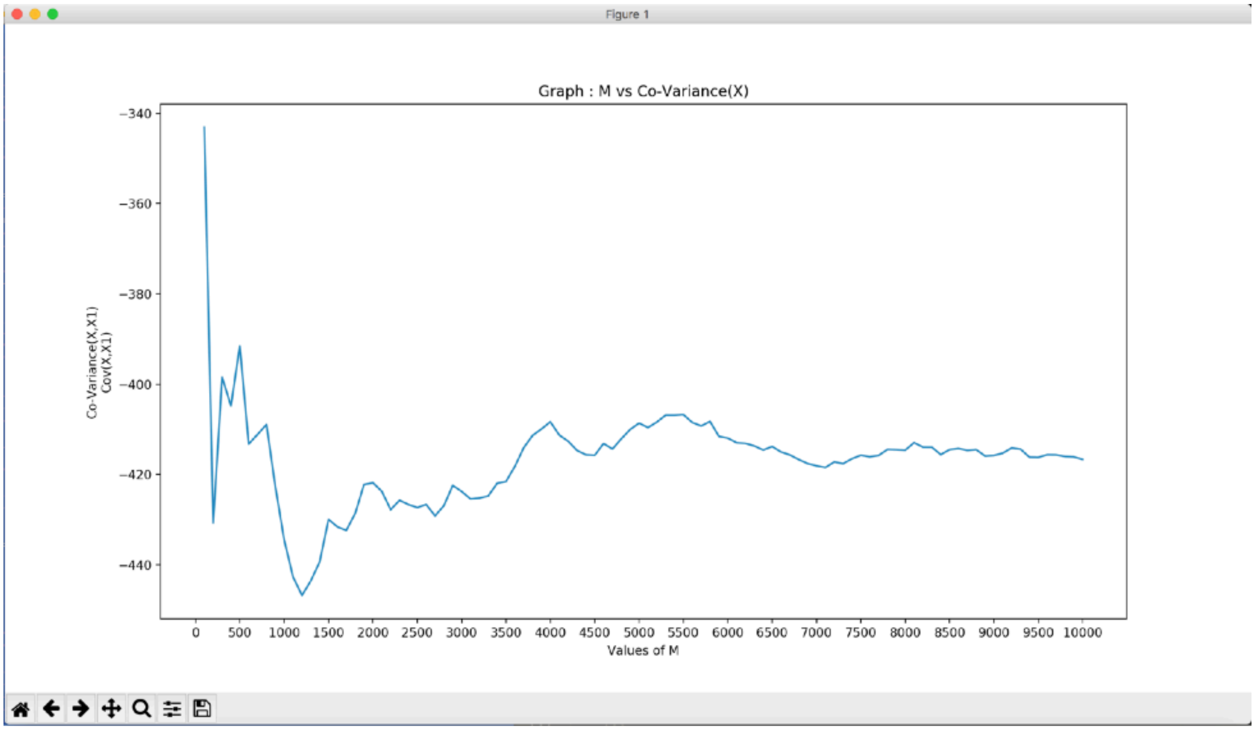
③ $\operatorname{Cov}(X, X_1) = \dfrac{1-N^2}{24} = \dfrac{(1+N)(1-N)}{24}$



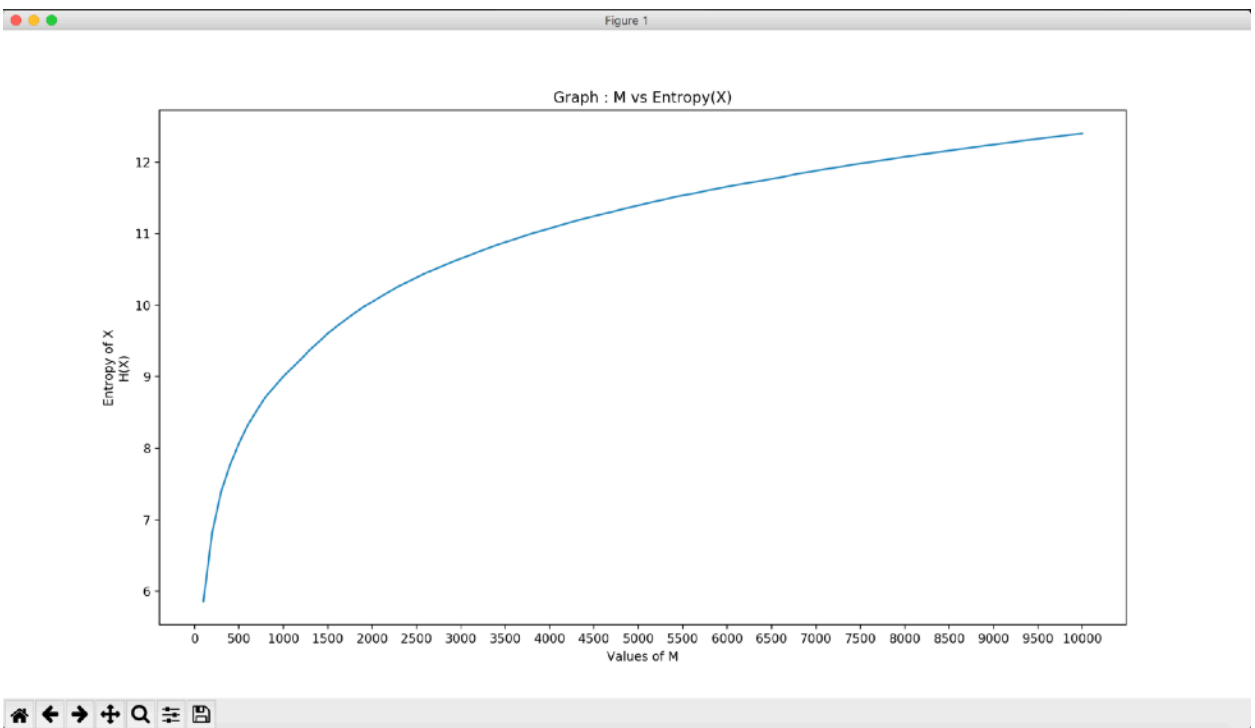
M vs Expectation of X -> E(X)



M vs Variance of X -> Var(X)



M vs Co-Variance of X -> Cov(X)



M vs Entropy of X -> H(X)

References:

1) https://machinelearningmastery.com/introduction-to-expected-value-variance-and-covariance

2) https://docs.scipy.org/doc/scipy/reference/generated/scipy.stats.entropy.html

3) https://en.wikipedia.org/wiki/Entropy_(information_theory)

4) https://www.youtube.com/user/jbstatistics