CS215 Assignment 2

Aquib Nawaz(190050023),Rajesh Dasari(190050030),
Paavan Kumar(190050051)

14 September 2020

Question 1

Given , $X_1, X_2, X_3, ...X_n$ are n identically distributed random variables with cdf $F_X(x)$ and pdf $f_X(x) = F'_X(x)$.

And $Y_1 = max(X_1, X_2, X_3, ...X_n)$
We know that the cdf of a variable is $F_X(x) = P(X \leq x)$
Therefore, cdf of Y_1 will be ,$F_{Y_1}(y) = P(Y_1 \leq y)$

$F_{Y_1}(y) = P(max(X_1, X_2, X_3, ...X_n) \leq y)$
If $max(X_1, X_2, X_3, ...X_n)$ is less than y ,then each of $X_1, X_2, X_3, ...X_n$ must be less than y

$F_{Y_1}(y) = P(X_1 \leq y \& X_2 \leq y \& ....\& X_n \leq y)$
Since these are independent random variables

$F_{Y_1}(y) = P(X_1 \leq y)P(X_2 \leq y) \dots P(X_n \leq y)$
Each of these is cdf of X_i for i ϵ $1 \dots n$

$F_{Y_1}(y) = F_X(x)F_X(x) \dots F_X(x)$
Therefore,

$F_{Y_1}(y) = F_X(x)^n$
For finding the pdf ,differentiating is sufficient

$f_{Y_1}(y) = nF_X(x)^{n-1}f_X(x)$

And $Y_1 = min(X_1, X_2, X_3, ...X_n)$
We know that the cdf of a variable is $F_X(x) = P(X \leq x)$
Therefore, cdf of Y_1 will be ,$F_{Y_2}(y) = P(Y_2 \leq y)$

$F_{Y_2}(y) = P(min(X_1, X_2, X_3, ...X_n) \leq y)$

$F_{Y_2}(y) = 1 - P(min(X_1, X_2, X_3, ...X_n) \geq y)$
If $min(X_1, X_2, X_3, ...X_n)$ is greater than y ,then each of $X_1, X_2, X_3, ...X_n$ must be greater than y

$F_{Y_2}(y) = 1 - P(X_1 \geq y \& X_2 \geq y \& ....\& X_n \geq y)$
Since these are independent random variables

$F_{Y_2}(y) = 1 - P(X_1 \geq y)P(X_2 \geq y) \dots P(X_n \geq y)$

$F_{X_i}(x) = P(X_i \leq x) = 1 - P(X_i \geq x)$
Hence ,

$F_{Y_2}(y) = 1 - (1 - F_X(x))(1 - F_X(x)) \dots (1 - F_X(x))$
Therefore,

$F_{Y_2}(y) = 1 - (1 - F_X(x))^n$
For finding the pdf ,differentiating is sufficient

$f_{Y_2}(y) = -n(1 - F_X(x))^{n-1}(-f_X(x))$

$f_{Y_2}(y) = n(1 - F_X(x))^{n-1} f_X(x)$

Question 2

Given k mixing probabilities p_i's where $\sum_{i=1}^{k} p_i = 1$ and $\forall i,\ 0 \le p_i \le 1$
Since $X \sim \sum_{i=1}^{k} p_i \mathcal{N}(\mu_i, \sigma_i^2)$

We have $\phi_X(t) = \sum_{i=1}^{k} p_i \phi_{X_i}(t)$ (using property of mgf)

$E(X) = \frac{d\phi_X(t)}{dt}|_{t=0} = \sum_{i=1}^{k} p_i \frac{d\phi_{X_i}(t)}{dt}|_{t=0}$

But we know that $\frac{d\phi_{X_i}(t)}{dt}|_{t=0} = \mu_i$

So, $E(X) = \sum_{i=1}^{k} p_i \mu_i$

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

$E(X^2) = \frac{d^2\phi_X(t)}{dt^2}|_{t=0} = \sum_{i=1}^{k} p_i \frac{d^2\phi_{X_i}(t)}{dt^2}|_{t=0} = \sum_{i=1}^{k} p_i(\mu_i^2 + \sigma_i^2)$ as $E(X_i^2) = \mu_i^2 + \sigma_i^2$

$Var(X) = \sum_{i=1}^{k} p_i(\mu_i^2 + \sigma_i^2) - (\sum_{i=1}^{k} p_i \mu_i)^2$

MGF of X_i is $\exp(\mu_i t + \frac{\sigma_i^2 t^2}{2})$

MGF of X is $\phi_X(t) = \sum_{i=1}^{k} p_i \exp(\mu_i t + \frac{\sigma_i^2 t^2}{2})$

Given $Z = \sum_{i=1}^{k} p_i X_i$

Consider $Y_i = p_i X_i$ let G_i and g_i be its CDF and PDF respectively. then,

Also $F_i(x)$ and $f_i(x)$ is CDF and PDF of X_i respectively.

$G_i(y) = P|Y_i \le y| = P|p_i X_i \le y| = P|X_i \le \frac{y}{p_i}| = F_i(\frac{y}{p_i})$

$g_i(y) = F_i'(\frac{y}{p_i}) = \frac{1}{p_i} f_i(\frac{y}{p_i}) = \frac{exp(\frac{-(\frac{y}{p_i} - \mu_i)^2}{2\sigma_i^2})}{p_i\sqrt{2\pi}\sigma_i} = \frac{exp(\frac{-(y - p_i\mu_i)^2}{2p_i^2\sigma_i^2})}{\sqrt{2\pi}p_i\sigma_i}$

So Y_i is also gaussian random variable with mean $p_i\mu_i$ and variance $p_i^2\sigma_i^2$

Now $Z = \sum_{i=1}^{k} Y_i$

MGF of sum of random variables is product of respective mgf.

$\phi_Z(t) = \prod_{i=1}^{k} \phi_{Y_i}(t) = \prod_{i=1}^{k} \exp(p_i\mu_i t + \frac{p_i^2\sigma_i^2 t^2}{2}) = \exp(\sum_{i=1}^{k}(p_i\mu_i t + \frac{p_i^2\sigma_i^2 t^2}{2}))$

$\phi_Z(t)$ can also be written as $\exp(\mu_Z t + \frac{\sigma_Z^2 t^2}{2})$

where $\mu_Z = \sum_{i=1}^{k} p_i \mu_i$ and $\sigma_Z^2 = \sum_{i=1}^{k} p_i^2 \sigma_i^2$

so by uniqueness of mgf Z is a gaussian random variable with parameters(μ_Z, σ_Z^2)

Now $E(Z) = \mu_Z$
$Var(Z) = \sigma_Z^2$
PDF of Z is $f_Z(x) = \frac{exp(-\frac{(x-\mu_Z)^2}{2\sigma_Z^2})}{\sqrt{2\pi}\sigma_Z}$

Question 3

Consider $\tau > 0$,
By markov's inequality , we have for any non zero random variable X,
$P(X \geq a) \leq \frac{E(X)}{a}$
We have $X - \mu > \tau$,where $\tau > 0$ Now consider for any $b > 0$,
we have $X - \mu + b > \tau + b$
Since , both the quantities are positive, the solution set of $X - \mu > \tau$ must be a subset of
The solution set S of $(X - \mu + b)^2 > (\tau + b)^2$
Therefore ,we have $P(X - \mu > \tau) \leq P((X - \mu + b)^2 > (\tau + b)^2)$
Applying **Markov's inequality**

$P(X - \mu > \tau) \leq \frac{E((X-\mu+b)^2)}{(\tau+b)^2}$

$E((X - \mu + b)^2) = E((X - \mu)^2) + E(2(X - \mu)b) + E(b^2)$
The term in the middle is anyway zero since $E(X - \mu) = 0$

$E((X - \mu + b)^2) = \sigma^2 + b^2$

$P(X - \mu > \tau) \leq \frac{\sigma^2+b^2}{(\tau+b)^2}$
Let $f(b) = \frac{\sigma^2+b^2}{(\tau+b)^2}$
Since this equation is true for all values of b, so differentiating it w.r.t b,
$f'(b) = \frac{2(\tau b-\sigma^2)}{\tau+b)^3}$
Setting $f'(b) = 0$ yields $b = \frac{\sigma^2}{\tau}$
The equation$f(b)$ yields the value $\frac{\sigma^2}{\sigma^2+\tau^2}$ at $b = \frac{\sigma^2}{\tau}$

$P(X - \mu > \tau) \leq \frac{\sigma^2}{\sigma^2+\tau^2}$

Consider $\tau < 0$
We have $X - \mu > \tau$,where $\tau < 0$ Now consider for any $b > 0$,
we have $-X + \mu + b < -\tau + b$
Since , both the quantities are positive, the solution set of $-X + \mu > -\tau$ must be a subset of
The solution set S of $(-X + \mu + b)^2 < (-\tau + b)^2$
Therefore ,we have $P(-X + \mu < -\tau) \leq P((-X + \mu + b)^2 > (-\tau + b)^2)$
Applying **Markov's inequality**

$P(-X + \mu > -\tau) \leq \frac{E((-X+\mu+b)^2)}{(-\tau+b)^2}$

$E((-X + \mu + b)^2) = E((-X + \mu)^2) + E(2(-X + \mu)b) + E(b^2)$
The term in the middle is anyway zero since $E(X - \mu) = 0$

$E((-X + \mu + b)^2) = \sigma^2 + b^2$

$P(-X+\mu > -\tau) \leq \frac{\sigma^2+b^2}{(-\tau+b)^2}$
Let $f(b) = \frac{\sigma^2+b^2}{(-\tau+b)^2}$
Since this equation is true for all values of b, so differentiating it w.r.t b,
$f'(b) = \frac{2(-\tau b-\sigma^2)}{-\tau+b)^3}$
Setting $f'(b) = 0$ yields $b = \frac{\sigma^2}{-\tau}$
The equation$f(b)$ yields the value $\frac{\sigma^2}{\sigma^2+\tau^2}$ at $b = \frac{\sigma^2}{-\tau}$

$P(-X+\mu > -\tau) \leq \frac{\sigma^2}{\sigma^2+\tau^2}$
$P(X-\mu > \tau) = 1 - P(\tau < X-\mu) = 1 - P(-X+\mu > -\tau)$
Therefore,

$P(X-\mu > \tau) \geq 1 - \frac{\sigma^2}{\sigma^2+\tau^2}$

Question 4

We know that $e^{tX} \geq e^{tx}, t > 0 \iff X \geq x$, since e^x is an increasing function
Applying **Markov's inequality**

$P(X \geq x) = P(e^{tX} \geq e^{tx}) \leq \frac{E(e^{tX})}{e^{tx}}$, for $t > 0$

We also know that $\phi_X(t) = E(e^{tX})$ for $t > 0$
Therefore

$P(X \geq x) \leq e^{-tx}\phi_X(t)$
We know that $e^{tX} \geq e^{tx}, t < 0 \iff X \leq x$, since e^x is an increasing function
Applying **Markov's inequality**

$P(X \leq x) = P(e^{tX} \geq e^{tx}) \leq \frac{E(e^{tX})}{e^{tx}}$, for $t < 0$

We also know that $\phi_X(t) = E(e^{tX})$ for $t < 0$
Therefore

$P(X \leq x) \leq e^{-tx}\phi_X(t)$
Applying the above result for the random variable $X = \sum_1^n X_i$,
We get

$P(X \leq (1+\delta)\mu) \leq e^{-(1+\delta)\mu}\phi_X(t)$
So the problem essentially boils down to calculating an upper bound for MGF of X

$\phi_X(t) = E(e^{tX}) = E(e^{t(X_1+X_2+\ldots+X_n)})$

$\phi_X(t) = \prod_1^n E(e^{tX_i})$
Since, the expectation of a bernoulli random variable is $1 + p(e^t - 1)$

$\phi_X(t) = \prod_1^n E(e^{tX_i}) = \prod_1^n (1 + p_i(e^t - 1))$
Using the equation $1 + x < e^x$

$\phi_X(t) = \prod_1^n (1 + p_i(e^t - 1)) \leq \prod_1^n e^{p_i(e^t-1)}$

$\phi_X(t) \leq e^{\sum_1^n p_i(e^t-1)} = e^{\mu(e^t-1)}$

$P(X \leq (1+\delta)\mu) \leq e^{-(1+\delta)\mu}\phi_X(t)$

$P(X \leq (1+\delta)\mu) \leq \frac{e^{\mu(e^t-1)}}{e^{(1+\delta)\mu}}$

Question 5

a)

The code for this part is attached with the name **q5a.m**

The code on execution produces 10 images with names as N.png where n belongs to [5,10,20,100,200,500,1000,5000,10000]

Here are the images attached for N = 100,1000 and 10000, using nsamp = 2000

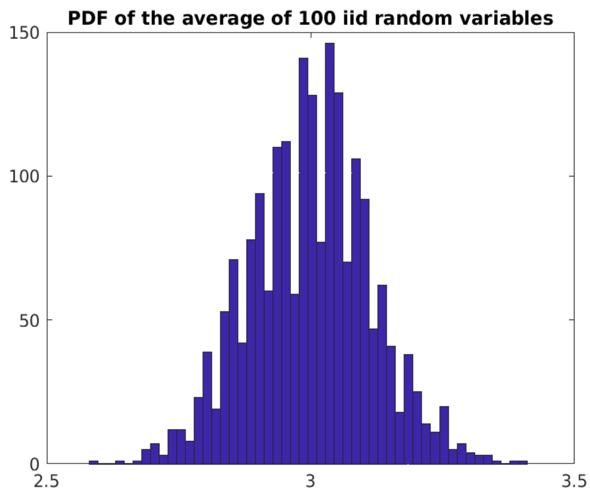

Figure 1: Plot for N = 100

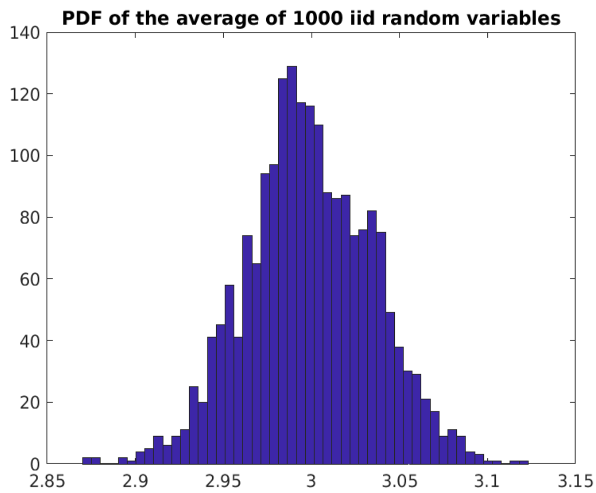

Figure 2: Plot for N = 1000

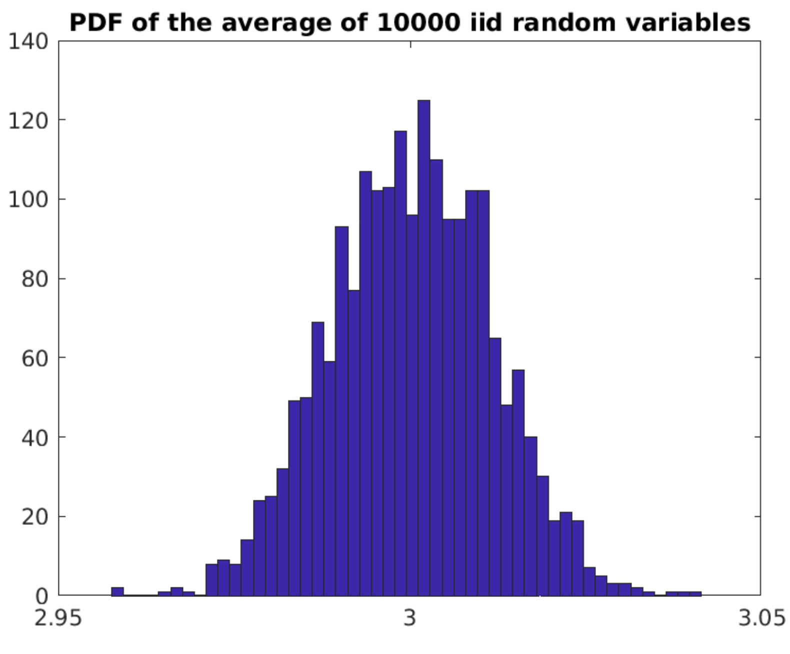

Figure 3: Plot for N = 10000

b),c)

The code for this part is attached with the name **q5bc.m**

The code on execution produces 11 images with the first 10 images for part **b)** and the last image for part **c)** Here are the images attached for N = 100,1000 and 10000.

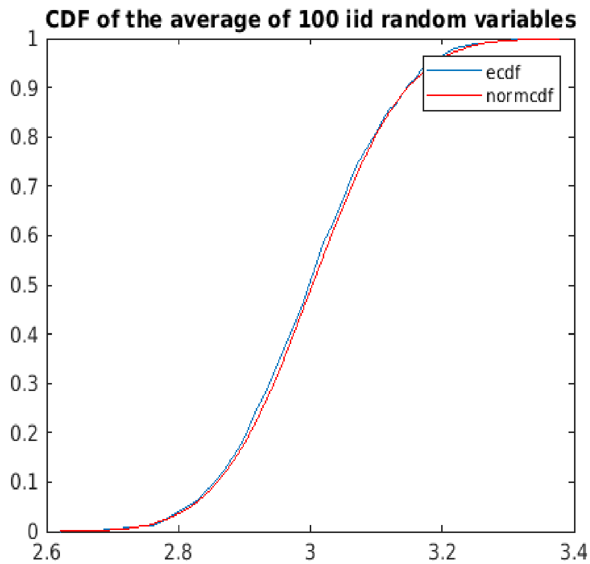

Figure 4: Plot for N = 100

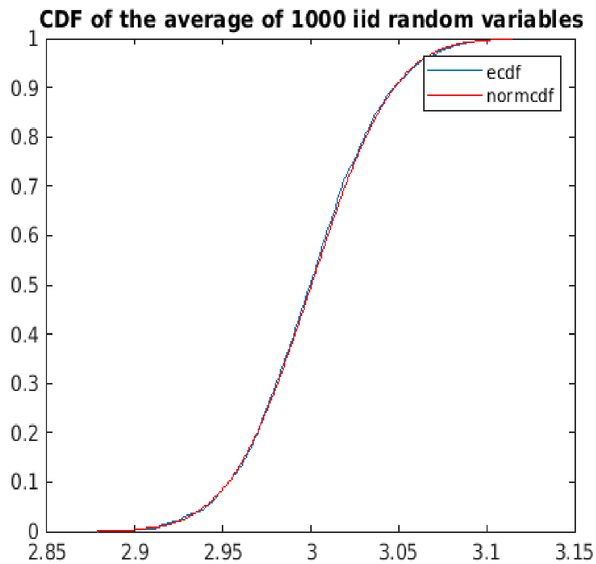

Figure 5: Plot for N = 1000

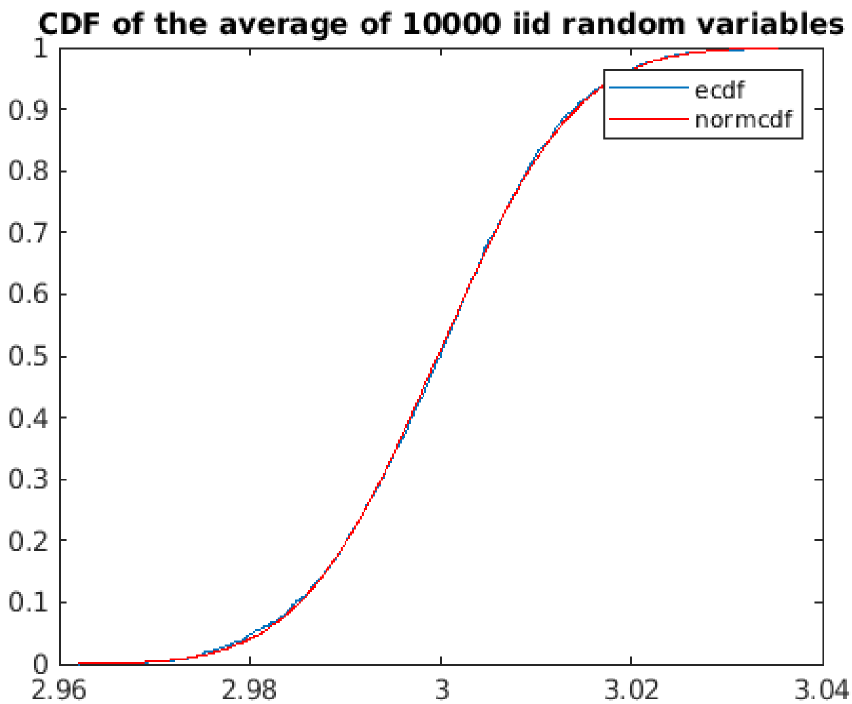

Figure 6: Plot for N = 10000

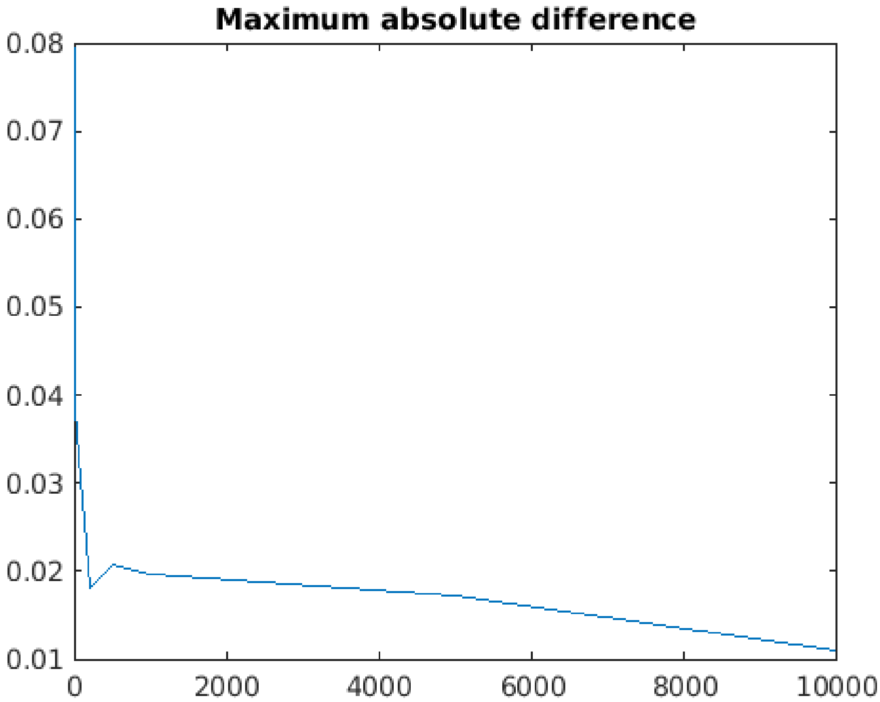

Figure 7: Plot for MAD versus N

Question 6

Correlation Coefficient

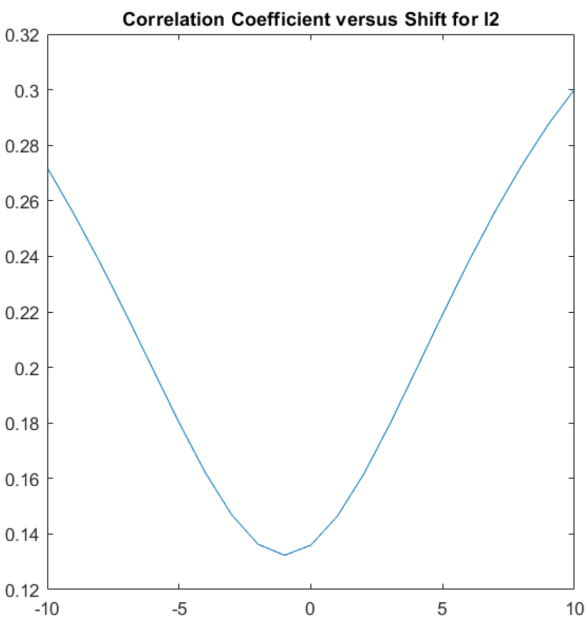

Figure 8: Plot of ρ versus t_i for I2

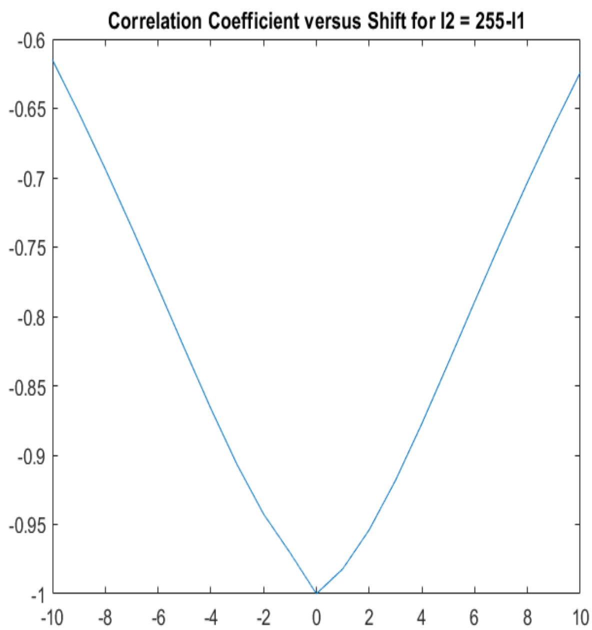

Figure 9: Plot of ρ versus t_i for I2 =255-I1

For **I2**
The correlation coefficient is minimum at $t_i = -1$ and it is decreasing as t_i goes from -10 to -1 since the distance b/w aligned pixels is going to decrease and then starts increasing from 0 to 10 because the distance between aligned pixels is increasing resulting in positive corelation-coefficient as the two images are closely related to each other

For **I2 = 255 -I1**
The magnitude of correlation coefficient is maximum at $t_i = 0$ and it is increasing as t_i goes from -10 to -1 since the distance b/w aligned pixels is going to decrease and then starts decreasing from 0 to 10 because the distance between aligned pixels is increasing resulting in negative corelation-coefficient as the two images are opposites of each other, the negative corelation coefficient means that the images are negatively corelated and hence -ρ looks same like that plotted for $I2$

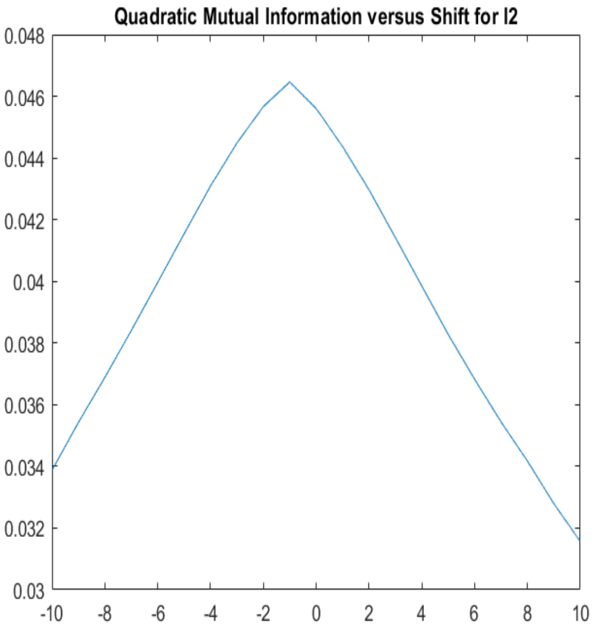

Figure 10: Plot of QMI versus t_i for I2

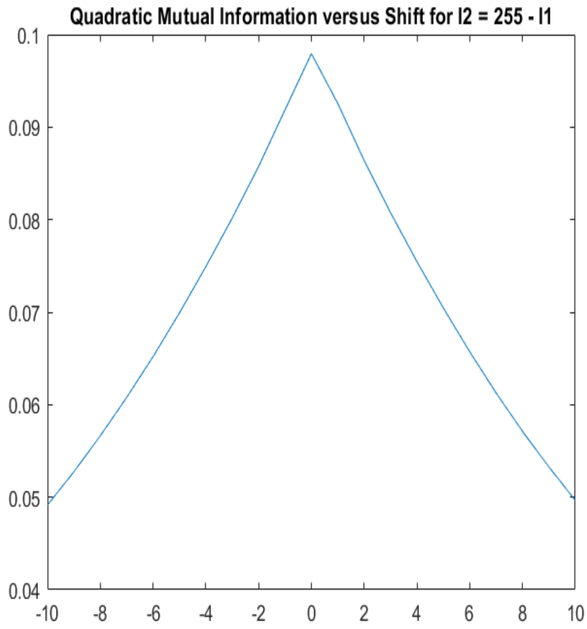

Figure 11: Plot of QMI versus t_i for I2 =255-I1

For **I2**
The Quadratic mutual information that is stored in QMI will be largest at $t_i = -1$, which also agrees with the fact that the correlation coefficient is minimum at $t_i = -1$ and the values of QMI decrease on either side of $t_i = -1$.This may be the charectersitic of the given images I1,I2. And also the QMI gradually falls to zero as the corelation coefficient increases

For **I2 = 255 -I1**
The Quadratic Mutual information stored in QMI will be largest when the two are least co-related since one of the image is inverse of the other image and this is due to the fact that each element is squared in calculation of QMI.The QMI hence decreases on either side of $t_i = 0$ since at $t_i = 0$,the images are negatives of each other,and one can observe that QMI falls to zero as the magnitude of corelation coefficient decreases or rather the corelation coefficient increases (since,it is negative)

From these graphs one can conclude that QMI is invesely dependent on corelation-coefficient

The code for this part is attached with the name **q6.m**

Question 7

The moment generating function of Multinomial random variable is given by,

$\phi_X(\vec{t}\,) = E(e^{\vec{t}\vec{X}}) = (\sum_1^n p_i e^{t_i})^n$ where p_i, n are parameters of multinomial random variable
Now , we have

$\frac{\partial \phi_X(\vec{t}\,)}{\partial t_i} = E(\frac{\partial e^{\vec{t}\vec{X}}}{\partial t_i}) = E(X_i e^{\vec{t}\vec{X}})$

And for i, j

$\frac{\partial}{\partial t_j}(\frac{\partial \phi_X(\vec{t}\,)}{\partial t_i}) = E(\frac{\partial}{\partial t_j}(X_i e^{\vec{t}\vec{X}})) = E(X_i X_j e^{\vec{t}\vec{X}})$

But also,

$\frac{\partial \phi_X(\vec{t}\,)}{\partial t_i} = \frac{\partial}{\partial t_i}((\sum_1^n p_i e^{t_i})^n) = n(\sum_1^n p_i e^{t_i})^{n-1}(\sum_1^n \frac{\partial}{\partial t_i}(p_i e^{t_i})) = np_i e^{t_i}(\sum_1^n p_i e^{t_i})^{n-1}$

For $i = j$

$\frac{\partial}{\partial t_i}(\frac{\partial \phi_X(\vec{t}\,)}{\partial t_i}) = \frac{\partial}{\partial t_i}(np_i e^{t_i}(\sum_1^n p_i e^{t_i})^{n-1}) = n(n-1)p_i e^{t_i}(\sum_1^n p_i e^{t_i})^{n-2}(\sum_1^n \frac{\partial}{\partial t_j}(p_i e^{t_i})) + np_i e^{t_i}(\sum_1^n p_i e^{t_i})^{n-1}$

$(\frac{\partial^2 \phi_X(\vec{t}\,)}{\partial t_i^2}) = n(n-1)p_i^2 e^{2t_i}(\sum_1^n p_i e^{t_i})^{n-2} + np_i e^{t_i}(\sum_1^n p_i e^{t_i})^{n-1}$

And for $i \neq j$

$\frac{\partial}{\partial t_j}(\frac{\partial \phi_X(\vec{t}\,)}{\partial t_i}) = \frac{\partial}{\partial t_j}(np_i e^{t_i}(\sum_1^n p_i e^{t_i})^{n-1}) = n(n-1)p_i e^{t_i}(\sum_1^n p_i e^{t_i})^{n-2}(\sum_1^n \frac{\partial}{\partial t_j}(p_i e^{t_i})) = n(n-1)p_i p_j e^{t_i} e^{t_j}(\sum_1^n p_i e^{t_i})^{n-2}$

We know that Variance$(X_i) = E(X_i^2) - E(X_i)^2$

$E(X_i^2) = n(n-1)p_i^2 + np_i$ letting $\vec{t} = \vec{0}$

$a_{ii} = n(n-1)p_i^2 + np_i - n^2 p_i^2$

$a_{ii} = np_i(1 - p_i)$

We know that Covariance$(X_i, X_j) = E(X_i X_j) - E(X_i)E(X_j)$

$E(X_i X_j) = n(n-1)p_i p_j$ for $i \neq j$ put $\vec{t} = \vec{0}$

$E(X_i)E(X_j) = (np_i)(np_j) = n^2 p_i p_j$

$a_{ij} = a_{ji} = n(n-1)p_i p_j - n^2 p_i p_j = -np_i p_j$ for $i \neq j$

$a_{ij} = -np_i p_j$ for $i \neq j$