CS 215 - Assignment 2

Niraj Mahajan
180050069

Raaghav Raaj
180050082

August 28, 2019

Home...

$X_1, X_2, \ldots X_n$: independent events

1. $cdf = F_X(x)$ and pdf $f_X(x) = F_X'(x)$

Let $Y_1 = \max(X_1, X_2, \ldots X_n)$ and $Y_2 = \min(X_1, X_2, \ldots X_n)$

For, Y_1,

$$P(Y_1 \le y) = P(\max(X_1, X_2, \ldots X_n) \le y)$$
$$= P(X_1 \le y, X_2 \le y, \ldots X_n \le y)$$

Now, as the variables $X_1, X_2 \ldots X_n$ are independent,

$$P(Y_1 \le y) = \prod_{i=1}^{n} P(X_i \le y)$$

$$\Rightarrow P(Y_1 \le y) = \underbrace{F_X(y) \cdot F_X(y) \ldots F_X(y)}_{n \text{ times}}$$

$$F_{Y_1}(y) = [F_X(y)]^n$$

Hence,

$$f_{Y_1}(y) = \frac{d F_{Y_1}(y)}{dy} = n[F_X(y)]^{n-1} f_X(y)$$
$$= n[F_X(y)]^{n-1} F_X'(y)$$

For, Y_2,

$$P(Y_2 > y) = P(\min(X_1, X_2, \ldots, X_n) > y)$$
$$= P(X_1 > y, X_2 > y, \ldots X_n > y)$$

As the variables are independent,

$$1 - P(Y_2 \le y) = \prod_{i=1}^{n} P(X_i > y)$$
$$= (1 - F_X(y))^n$$

$$F_{Y_2}(y) = 1 - [1 - F_X(y)]^n$$

Hence,

$$f_{Y_2}(y) = n[1 - F_X(y)]^{n-1} f_X(y)$$
$$= n[1 - F_X(y)]^{n-1} \cdot F_X'(Y)$$

2. A random variable,

x belongs to a Gaussian mixture model (GMM)

if $X \sim \sum_{i=1}^{K} p_i \mathcal{N}(\mu_i, \sigma_i^2)$,

where $\underbrace{\sum_{i=1}^{K} p_i = 1}$.

$$\mathcal{N}(\mu_i, \sigma_i^2) = \frac{1}{\sqrt{2\pi\sigma_i^2}} \exp\left(-\frac{(x-\mu_i)^2}{2\sigma_i^2}\right)$$

Now, as per the questions,

$x \in GMM$

Hence,

$$\cancel{E(x) = \sum \int_{-\infty}^{\infty} x \cdot \mathcal{N}(\mu_i, \sigma_i^2)}$$

$$E(x) = \int_{-\infty}^{\infty} x \cdot \sum_{i=1}^{K} p_i \mathcal{N}(\mu_i, \sigma_i^2)\, dx$$

$$\Rightarrow E(x) = \sum_{i=1}^{K} \left[p_i \int_{-\infty}^{\infty} x \cdot \mathcal{N}(\mu_i, \sigma_i^2)\, dx \right]$$

$$\Rightarrow E(x) = \sum_{i=1}^{K} p_i \cdot \mu_i$$

Now, to determine the Var(x),

we will find the MGF(x).

$$f_x(x) = \sum_{i=1}^{K} p_i \mathcal{N}(\mu_i, \sigma_i^2)$$

$$\therefore \cancel{MGF(x)}\ MGF(x) = \phi(t) = E(e^{tx})$$

$$\Rightarrow \cancel{\phi(x)}\ \phi(t) = \int_{-\infty}^{\infty} e^{tx} \left(\sum_{i=1}^{K} p_i \mathcal{N}(\mu_i, \sigma_i^2) \right) \cdot dx.$$

$$= \sum_{i=1}^{K} \left[p_i \int_{-\infty}^{\infty} e^{tx} \mathcal{N}(\mu_i, \sigma_i^2) \cdot dx \right]$$

$$\cancel{= \sum_{i=1}^{K} \left[p_i \cdot \exp\left(\mu_i t + \frac{1}{2}\sigma_i t^2\right) \right]}.$$

We know that the moment generating function, $\phi_i(t)$ of $\mathcal{N}(\mu_i, \sigma_i^2)$,

$$\phi_i(t) = \exp\left(\mu_i t + \frac{1}{2}\sigma_i^2 t^2\right)$$

~~Thus, the moment~~

Hence, the moment generating function of the given Gaussian mixture model is given by,

$$\phi(t) = \sum_{i=1}^{K} p_i \exp\left(\mu_i t + \frac{1}{2}\sigma_i^2 t^2\right)$$

- We know that,

$$E(X) = \phi'(0)$$

and $E(X^2) = \phi''(0)$.

Hence, $\phi'(t) = \sum_{i=1}^{K} p_i (\mu_i + \sigma_i^2 t) \cdot \exp\left(\mu_i t + \frac{1}{2}\sigma_i^2 t^2\right)$

and, $\phi''(t) = \sum_{i=1}^{K} p_i \left\{\sigma_i^2 \exp\left(\mu_i t + \frac{1}{2}\sigma_i^2 t^2\right) + (\mu_i + \sigma_i^2 t)^2 \cdot \exp\left(\mu_i t + \frac{1}{2}\sigma_i^2 t^2\right)\right\}$

Hence,

$$E(X^2) = \phi''(0) = \sum_{i=1}^{K} p_i (\sigma_i^2 + \mu_i^2)$$

Thus,

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

$$\Rightarrow \boxed{Var(X) = \sum_{i=1}^{K} p_i (\sigma_i^2 + \mu_i^2) - \left[\sum_{i=1}^{K} p_i \mu_i\right]^2}$$

Answers:

$$E(X) = \sum_{i=1}^{K} \mu_i p_i$$

$$Var(X) = \sum_{i=1}^{K} p_i (\sigma_i^2 + \mu_i^2) - \left[\sum_{i=1}^{K} p_i \mu_i\right]^2$$

$$\phi(t) = \sum_{i=1}^{K} p_i \exp\left(\mu_i t + \frac{1}{2}\sigma_i^2 t^2\right)$$

Given, $X_i \sim N(\mu_i, \sigma_i^2)$

and, $Z = \Sigma_{i=1}^{K} p_i X_i$

Hence,

$$E(Z) = E(\Sigma_{i=1}^{K} p_i X_i)$$
$$= \Sigma_{i=1}^{K} E(p_i X_i)$$
$$= \Sigma_{i=1}^{K} p_i E(X_i)$$
$$= \Sigma_{i=1}^{K} p_i \mu_i$$

~~and~~, Now, for Var(z) and PDF(z), we calculate ~~Var(z)~~ the MGF first -

$$\Rightarrow \phi(t) = E(e^{tz})$$
$$\cancel{= \int_{-\infty}^{\infty} e^{tz}}$$

$$\Rightarrow \phi(t) = E(e^{tp_1X_1} \cdot e^{tp_2X_2} \cdots e^{tp_KX_K})$$
$$\cancel{= \prod_{i=1}^{K} p_i \exp(\mu_i t + \frac{1}{2}\sigma_i^2 t^2)}$$

$$\Rightarrow \cancel{\phi(t) = \left(\prod_{i=1}^{K} p_i\right) \exp((\sum_{i=1} \mu_i) t + \frac{1}{2}(\sum_{i=1} \sigma_i^2) t^2)}$$

$$\phi(t) = \prod_{i=1}^{K} \exp(\mu_i p_i t + \frac{1}{2}\sigma_i^2 p_i^2 t^2)$$

Thus, MGF of Z:

~~Hence~~, $\phi(t) = \exp\left[(\Sigma \mu_i p_i)t + \frac{1}{2}(\Sigma \sigma_i^2 p_i^2)t^2\right]$

~~Thus~~ Now, comparing with MGF of general Gaussian expression,

$$\sigma_z^2 = \sum_{i=1}^{K} \sigma_i^2 p_i^2$$

and, $\mu_z = \sum_{i=1}^{K} \mu_i p_i$

Therefore,

the variance of z:

$$Var(z) = \Sigma_{i=1}^{K} p_i^2 \sigma_i^2$$

and, using the uniqueness property,

PDF of Z:

$$f_Z(z) = \frac{1}{\sqrt{2\pi\sigma_z^2}} \exp\left\{-\frac{(z-\mu_z)^2}{2\sigma_z^2}\right\}$$

$$\text{where, } \mu_z = \sum_{i=1}^{K} \mu_i p_i$$

$$\sigma_z^2 = \Sigma_{i=1}^{K} p_i^2 \sigma_i^2$$

Answers:

$$\mu_z = E(z) = \Sigma_{i=1}^{K} p_i \mu_i$$

$$\sigma_z^2 = Var(z) = \Sigma_{i=1}^{K} \sigma_i^2 p_i^2$$

$$PDF(z) = \frac{1}{\sqrt{2\pi\sigma_z^2}} \exp\left\{-\frac{(z-\mu_z)^2}{2\sigma_z^2}\right\}$$

$$\phi(t) = \exp\left[(\Sigma_{i=1}^{K} p_i \mu_i)t + \frac{1}{2}(\Sigma_{i=1}^{K} p_i^2 \sigma_i^2)t^2\right]$$

3. Markov's inequality states that,

if X is a random variable that takes ~~on~~ only non negative values, then for any value $a>0$,

$$P(X \geq a) \leq \frac{E(X)}{a}$$

To prove: With mean μ, and variance σ^2:

$$P(X-\mu \geq \tau) \leq \frac{\sigma^2}{\sigma^2+\tau^2}, \quad \tau > 0$$

and,

$$P(X-\mu \geq \tau) \geq 1-\frac{\sigma^2}{\sigma^2+\tau^2}, \quad \tau < 0$$

Proof: ~~We have that~~

$$\cancel{P[X \geq a] \leq \frac{E(X)}{a}}$$

~~For~~ Case: $\tau > 0$

Clearly, for

$$P[(X-\mu)^2 \geq \tau^2] = P[|X-\mu| \geq \tau], \quad \text{—(1)}$$

$$\forall \tau > 0$$

Let Y be a random variable such that,

$$Y = X-\mu.$$

Now,

$$P[X-\mu \geq \tau] = P[Y \geq \tau]$$

$$\Rightarrow P[X-\mu \geq \tau] = P[Y+\alpha \geq \tau+\alpha], \text{ for some } \alpha > 0 \quad \text{—(2)}$$

Now, clearly, by using (1),

$$P[Y+\alpha \geq \tau+\alpha] \leq P[(Y+\alpha)^2 \geq (\tau+\alpha)^2] \quad \text{—(3)}$$

Now, as per Markov's Inequality,

$$P[X \geq a] \leq \frac{E(X)}{a}$$

$$\Rightarrow P[(Y+\alpha)^2 \leq (\tau+\alpha)^2] \leq \frac{E[(Y+\alpha)^2]}{(\tau+\alpha)^2} \quad \text{—(4)}$$

Scanned with CamScanner

using ②, ③, and ④,

$$P(X-\mu \geq \tau) \leq \frac{E((Y+\alpha)^2)}{(\alpha+\tau)^2} \quad — ⑤$$

~~As i~~ We have that,

$$E[(Y+\alpha)^2] = E(Y^2 + 2\alpha Y + \alpha^2)$$

$$\Rightarrow E[(Y+\alpha)^2] = E(Y^2) + 2\alpha E(Y) + \alpha^2$$

Since, $Y = X - \mu$,

$$E(Y^2) = \sigma^2$$

and, $E(Y) = 0$.

$$\Rightarrow E[(Y+\alpha)^2] = \sigma^2 + \alpha^2 \quad — ⑥$$

using ⑤ and ⑥

$$P(X-\mu \geq \tau) \leq \frac{\sigma^2+\alpha^2}{(\tau+\alpha)^2} \quad — ⑦$$

Since, the above inequality is true $\forall \alpha \geq 0$;

we ~~cho~~ select an α, such that

RHS is minimum.

Let $y = \dfrac{\sigma^2+\alpha^2}{(\tau+\alpha)^2}$

For minimum,

$$\therefore \frac{dy}{d\alpha} = \frac{2\alpha}{(\tau+\alpha)^2} - \frac{2(\sigma^2+\alpha^2)}{(\alpha+\tau)^3} = 0$$

$$\Rightarrow \alpha(\alpha+\tau) - (\sigma^2+\alpha^2) = 0.$$

$$\therefore \boxed{\alpha = \frac{\sigma^2}{\tau}}$$

Scanned with CamScanner

Substituting , $\alpha = \frac{\sigma^2}{\tau}$ in ⑦ ,

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

· Now,

Case : $\tau < 0$

We proceed as in the previous , with $\beta = -\tau > 0$, and $\alpha \geq 0$.

$$\Rightarrow P(X-\mu < \tau) = P(Y < -\beta)$$
$$= P(-Y > \beta) \leq \frac{\sigma^2}{\beta^2+\sigma^2}$$

Hence, $P(X-\mu < \tau) \leq \frac{\sigma^2}{\tau^2+\sigma^2}$ $[\because \beta = -\tau]$

· Taking compliment,

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

Hence, proved.

4. Let X be the random variable, such that the PDF of X is $f_X(x)$.

The moment generating function is given by,

$$\phi_X(t) = \int_{-\infty}^{\infty} e^{tx} \cdot f_X(x) \cdot dx$$

Part 1:

~~For t>0;~~

a) We have that,

$$P[X \geq x] = \int_{x}^{\infty} f_X(z) \cdot dz$$

$$\Rightarrow P[X \geq x] = e^{-tx} \cdot e^{tx} \int_{x}^{\infty} f_X(z) \cdot dz$$

$$= e^{-tx} \int_{x}^{\infty} e^{tx} f_X(z) \cdot dz \quad \text{— (1)}$$

Now, as $t>0$,

$$\int_{x}^{\infty} e^{tx} f_X(z) \cdot dz \leq \int_{-\infty}^{\infty} e^{tx} f_X(z) \cdot dz$$

$$\leq \int_{-\infty}^{\infty} e^{tz} f_X(z) \cdot dz \quad \text{— (2)}$$

(as $t>0$)

$\Rightarrow$ using (1) and (2)

$$P[X \geq x] \leq e^{-tx} \int_{-\infty}^{\infty} e^{tz} f_X(z) \cdot dz$$

$$\Rightarrow \boxed{P[X \geq x] \leq e^{-tx} \cdot \phi_X(t)}$$

b) Here,

$$P[X \leq x] = \int_{-\infty}^{x} f_X(z) dz$$

$$\Rightarrow P[X \leq x] = e^{-tx} \cdot e^{tx} \int_{-\infty}^{x} f_X(z) \cdot dz$$

$$= e^{-tx} \int_{-\infty}^{x} e^{tx} f_X(z) \cdot dz \quad \text{— (3)}$$

As, $t<0$,

$$\int_{-\infty}^{x} e^{tx} f_X(z) \cdot dz \leq \int_{-\infty}^{x} e^{tz} f_X(z) dz$$

$$\leq \int_{-\infty}^{\infty} e^{tz} f_X(z) \cdot dz = \phi_X(t). \quad \text{— (4)}$$

(as $t<0$)

Hence, using ③ and ④

$$\boxed{P[X \le x] \le e^{-tx}\phi_X(t)}, \text{ for } t<0$$

Part 2:

$$X = \sum_{i=1}^{n} X_i, \quad E(X_i) = p_i$$

$$\mu = \Sigma_{i=1}^{n} p_i = E(x)$$

To prove:

$$P[x > (1+\delta)\mu] \le \frac{e^{\mu(e^t-1)}}{e^{(1+\delta)t\mu}} \quad \forall t \ge 0, \ \forall \delta > 0$$

Proof: As proved in the previous question,

for $t \ge 0$

$$P[x > (1+\delta)\mu] \le e^{-t\mu(\delta+1)} \cdot \phi_x(t) \quad \text{—(i)}$$

~~Now,~~ $\cancel{\phi_X(t)} = \cancel{B}$

As $X = \Sigma_{i=1}^{K} X_i$, and are Bernoulli variables, that are indepent.

Hence,

$$\phi_X(t) = \prod_{i=1}^{K} \phi_{X_i}(t) \quad \text{—(ii)}$$

For a Bernoulli variable, the moment generating function,

$$\phi_{X_i}(t) = 1 - p_i + p_i e^t$$

$$= 1 + p_i(e^t - 1)$$

Now, $1 + x \le e^x$

$$\Rightarrow \phi_{X_i}(t) \le e^{p_i(e^t-1)} \quad \text{—(iii)}$$

Hence, using (ii) and (iii)

$$\phi_X(t) \le \exp\left(\sum_{i=1}^{K} p_i(e^t - 1)\right) \quad \text{—(iv)}$$

$$\phi_x(t) \le \exp(\mu(e^t - 1)) \quad \text{— (iv)}$$

Hence, by using (i) and (iv),

$$P[X > (1+\delta)\mu] \le \frac{e^{\mu(e^t-1)}}{e^{(1+\delta)t\mu}}$$

Now, to tighten the bounds, we minimise the RHS.

$$\text{RHS}: \quad \exp[\mu e^t - \mu - t\mu - \delta t\mu]$$

- Taking logarithm with base e,

$$\log_e(\text{RHS}) = \mu e^t - \mu - t\mu - \delta t\mu.$$

→ Differentiating the above expression with respect to t, and equating to zero.

$$\mu e^t = \cancel{\mu(1+t)}\ \mu(1+\delta).$$

$$\Rightarrow \boxed{t = \ln(1+\delta)}$$

Hence, to tighten the bounds, the value of

$$t = \ln(1+\delta).$$

Thus, the inequality converts to,

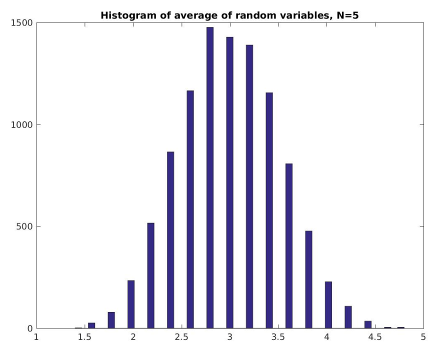
$$P[X > (1+\delta)\mu] \le \frac{e^{\mu\delta}}{(1+\delta)^{\mu\delta}} = \left(\frac{e}{1+\delta}\right)^{\mu\delta}$$

QED.

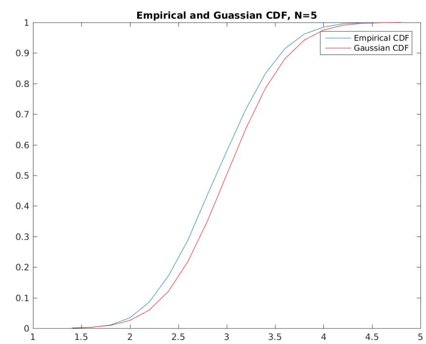
1 Question 5

Usage of MATLAB Code

- Load code in the following path 'matlab_code/Q5/q5.m'
- Run the code
- This should create two figures simultaneously (probably one on the other, you may have to separate them). They both will iterate over all N values and show the respective histograms and cdfs
- After this, two more figures will be created. Both will have MAD plotted against N. But one will have it plotted in the usual way, but the other will have it plotted against semilogx, which will have better visibility.
- All plots are inlcuded in this report, at the end, (after question 6).
- I have also saved these plots in the directory 'matlab_code/Q5/jpgs'



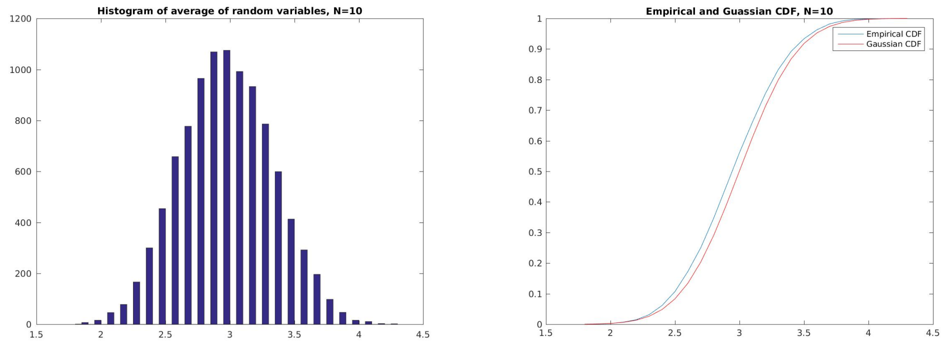
(a) Histogram.



(b) More Histogram.

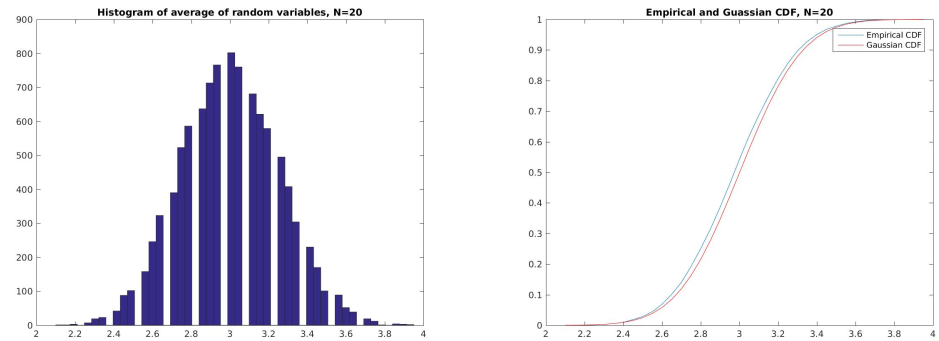
Figure 1: N=5

As the value of N increases, the histogram starts resembling a bell curve



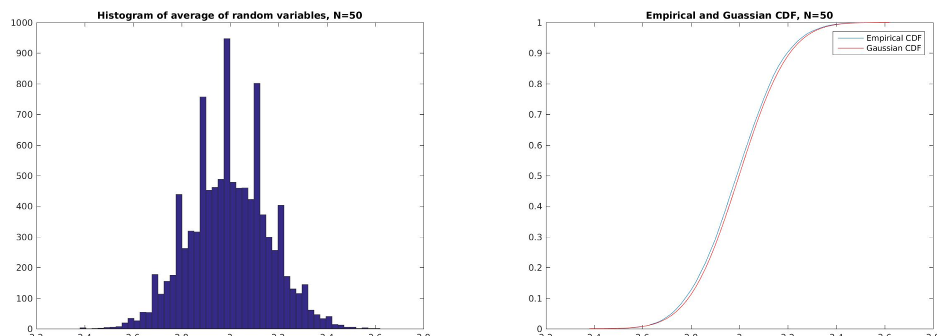
(a) Histogram. (b) More Histogram.

Figure 2: N=10



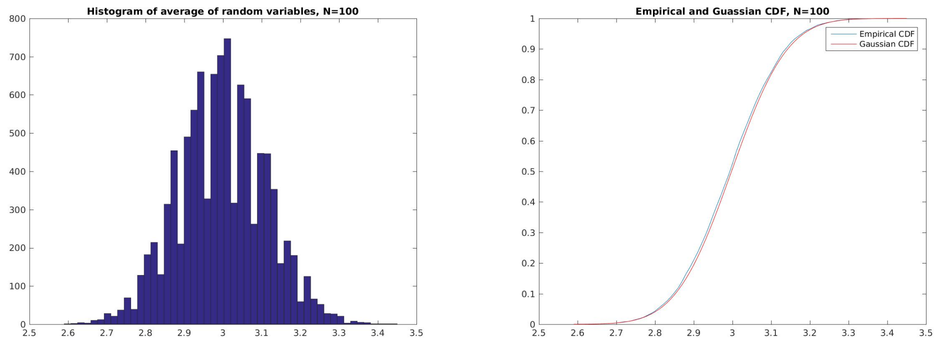
(a) Histogram. (b) More Histogram.

Figure 3: N=20



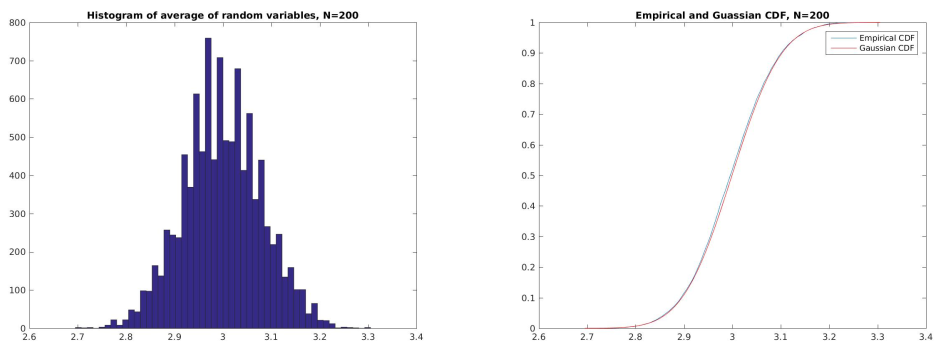
(a) Histogram. (b) More Histogram.

Figure 4: N=50



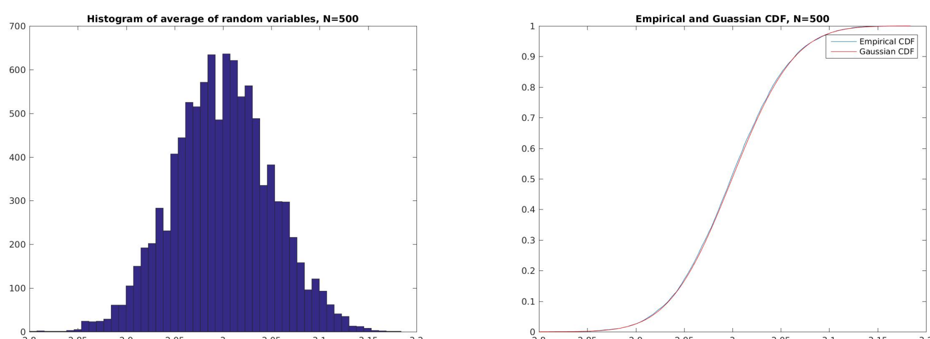
(a) Histogram. (b) More Histogram.

Figure 5: N=100



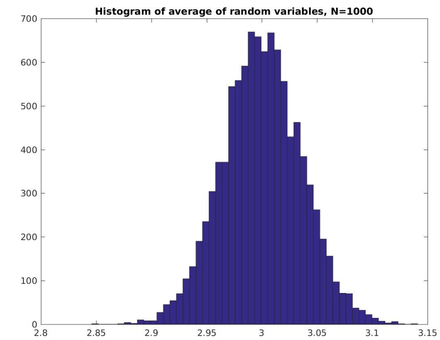
(a) Histogram. (b) More Histogram.

Figure 6: N=200

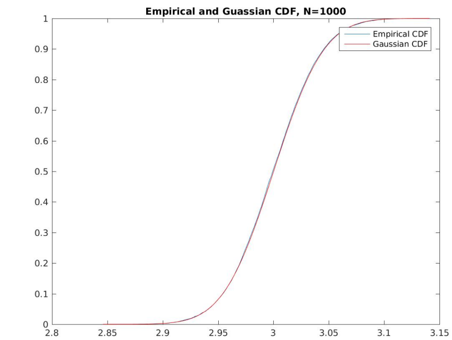


(a) Histogram. (b) More Histogram.

Figure 7: N=500

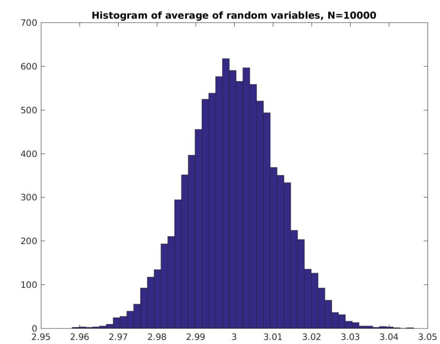


(a) Histogram.

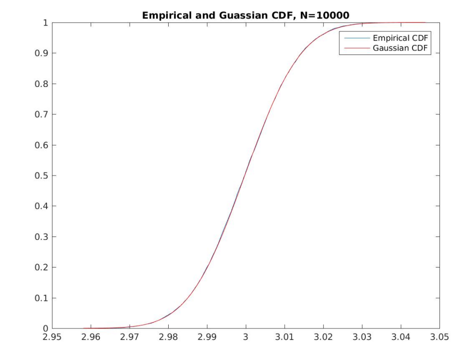


(b) CDF.

Figure 8: N=1000

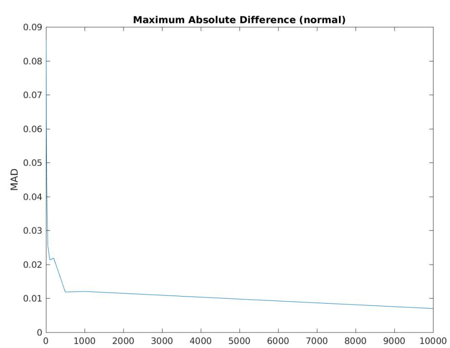


(a) Histogram.

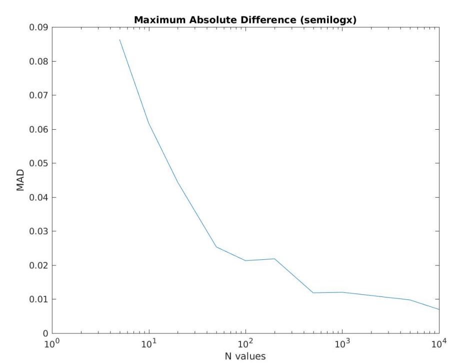


(b) CDF.

Figure 9: N=10000



(a) plotted against N



(b) plotted against semilogx(N)

Figure 10: MAD

2 Question 6

Case I : Comparing Figure 1 and Figure 2

The figures given 'T1.jpg' and 'T2.jpg' were loaded in matlab. The second image was shifted by a range from -10 to 10, and the correlation coefficient, and QMI was calculated and plotted for every shift.

Correlation Coefficient In the first case, we can observe that the plot of correlation coefficient achieves minimum at -1. By the defination, we can say that both the images are aligned 'the best' when the second image is shifted by -1. We can also conclude that the second image is apart by a pixel, as compared to the first image.

QMI The QMI is the squared summation of the difference between the joiont pdf and the product of the marginal pdfs. We can say that higher the values of the qmi, more will be the dependence between two images. In our plot, the best value of qmi is obtained when the shift is by -1, which supports our conclusion that the images are a pixel apart. We can also observe that the values of qmi are very low in the first case. This is because there is no 'explicit' relation between the two images.

Case II : Comparing Figure 1 and negative of Figure 1

The figures given 'T1.jpg' was loaded in matlab. The negative of the image was stored as another image and was shifted by a range from -10 to 10. The correlation coefficient, and QMI was calculated and plotted for every shift.

Correlation Coefficient In this first case, we can observe that the plot of correlation coefficient achieves a perfect -1 at a shift of zero. This is because both the images are physically the same, with just negative intensities. So, at zero shift, they will have the most common part and the least correlation coefficient coefficient. As the image is shifter further, the images get more and more misaligned which lead in an increase in the correlation coefficient.

QMI In this case, since both the images are just the negative of one another, the peak of the qmi is achieved when the shift is zero. Also, we can see that this qmi plot falls off faster than the one observed in the first case. This is because when we shift by, say tx, then the first tx columns are completely reduced to 0.

Usage of MATLAB Code

- Load code in the following path 'matlab_code/Q6/q6.m'
- Run the code.
- This will create 4 different plots for - Correlation coefficient, QMI, and both of these again when we compare with the first image with it's negative.

- These plots have been included in the report, as well as can be found in the directory 'matlab_code/Q6/jpgs'

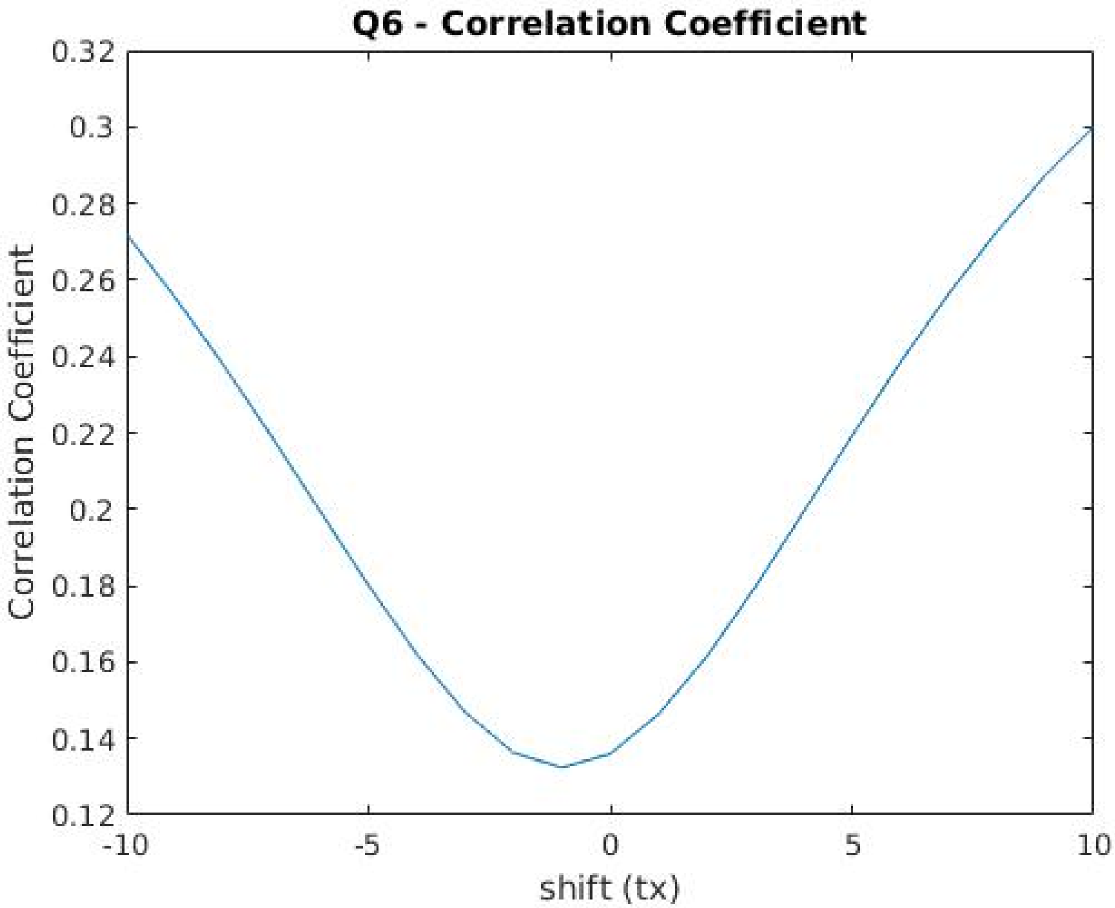


Figure 11

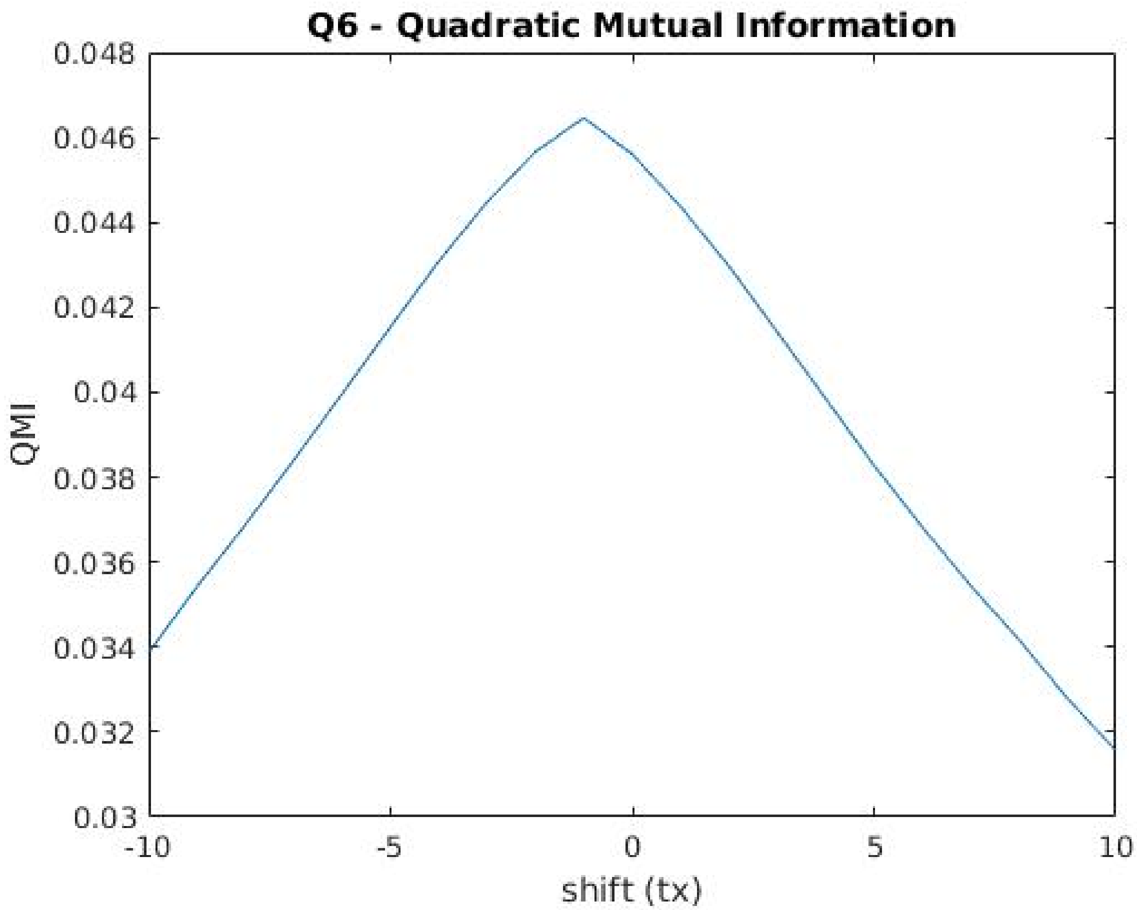

Q6 - Quadratic Mutual Information
0.048
0.046
0.044
0.042
0.04
0.038
0.036
0.034
0.032
0.03
QMI
-10
-5
0
5
10
shift (tx)

Figure 12

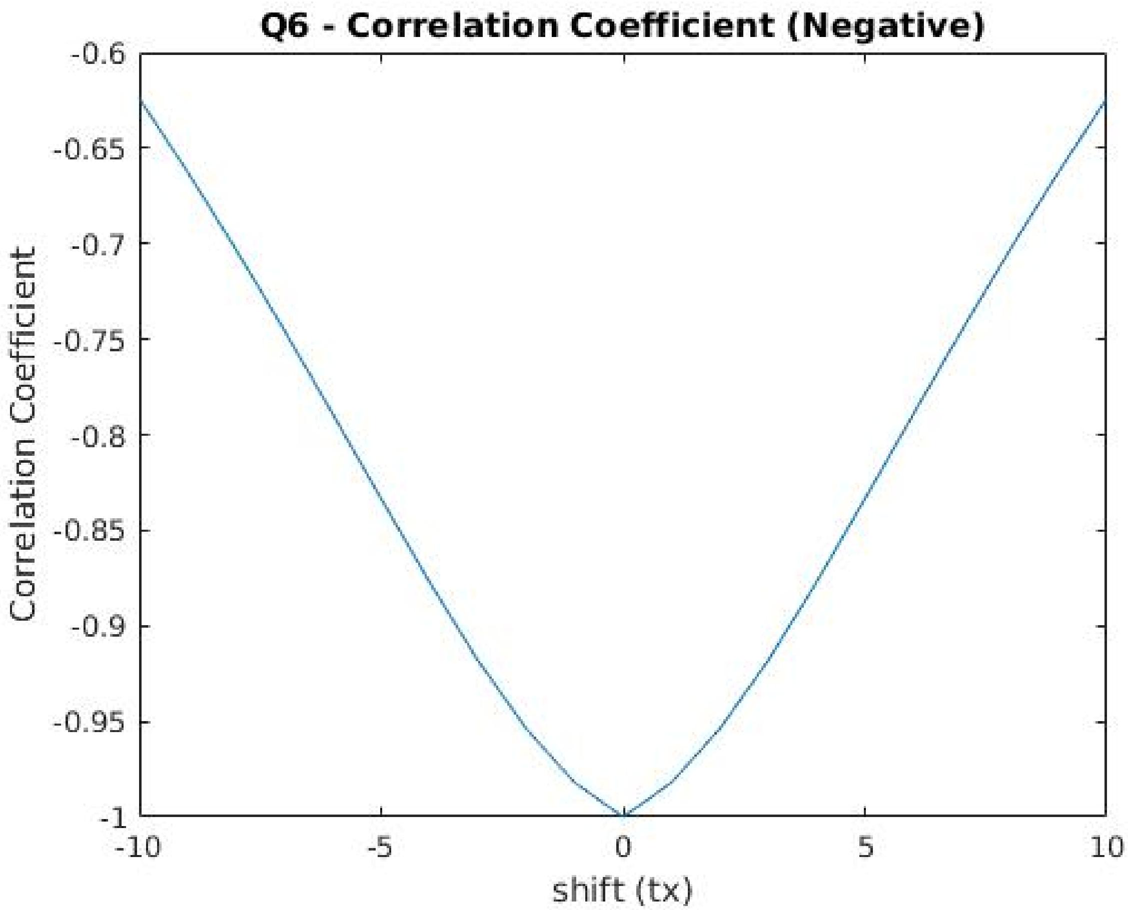
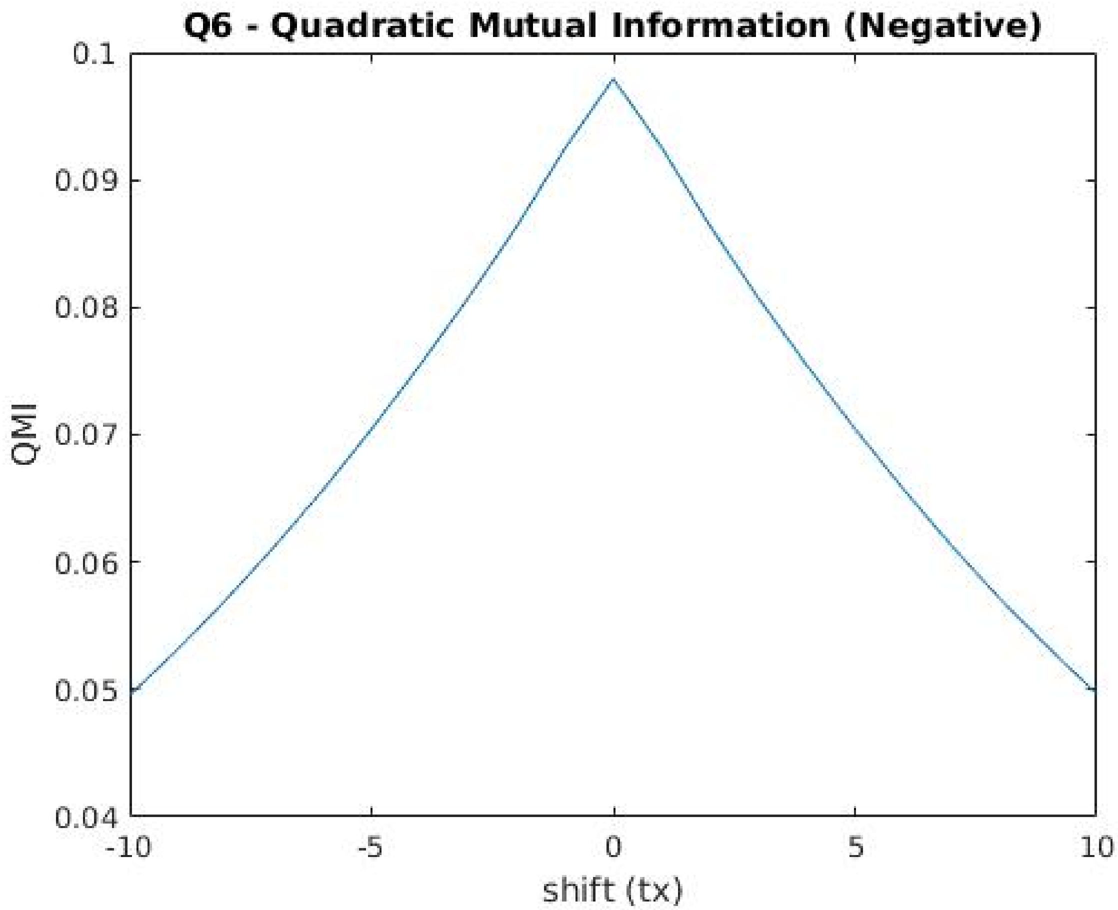

Q6 - Correlation Coefficient (Negative)
Correlation Coefficient
-0.6
-0.65
-0.7
-0.75
-0.8
-0.85
-0.9
-0.95
-1
-10
-5
0
5
10
shift (tx)

Figure 13



Q6 - Quadratic Mutual Information (Negative)
0.1
0.09
0.08
0.07
0.06
0.05
0.04
QMI
-10
-5
0
5
10
shift (tx)

Figure 14