CSE 544: Probability & Statistics for Data Science,
Spring 2021

Assignment - 3

Submitted by

Harmanpreet Singh Khurana, SBU ID: 113262379,
Mayank Jain, SBU ID: 113263864,
Rajadorai DS, SBU ID: 113259773,
Venkatesan Ravi, SBU ID: 113263484

1) MSE in terms of bias :

To prove : $bias^2(\hat{\theta}) + Var(\hat{\theta})$

We know, $MSE = E[(\theta - \hat{\theta})^2]$

$= E[\theta^2 + \hat{\theta}^2 - 2\theta.\hat{\theta}]$

$\overset{LOE}{=} E[\theta^2] + E[\hat{\theta}]^2 - 2E[\theta].E[\hat{\theta}]$

But θ is the true value, a constant,

(i) $\Rightarrow \therefore MSE = \theta^2 + E[\hat{\theta}]^2 - 2\theta E[\hat{\theta}]$ $\rightarrow$ Eqn (i)

We know, $Bias(\hat{\theta}) = E[\hat{\theta}] - \theta$

$\therefore Bias^2(\hat{\theta}) = E^2[\hat{\theta}] + \theta^2 - 2\theta E[\hat{\theta}]$ (ii)

(Squaring on both sides)

We know, $Var(\hat{\theta}) = E[\hat{\theta}^2] - E^2[\hat{\theta}]$ $\rightarrow$ (iii)

Add (ii), (iii) $\Rightarrow Var(\hat{\theta}) + Bias^2\theta = E[\hat{\theta}^2] - \cancel{E^2[\hat{\theta}]} + \cancel{E^2[\hat{\theta}]} + \theta^2 - 2\theta E[\hat{\theta}]$

$Var(\hat{\theta}) + Bias^2(\hat{\theta}) = E[\hat{\theta}^2] + \theta^2 - 2\theta E[\hat{\theta}]$

which is also = Equation (i)

$$\boxed{\therefore MSE = Var(\hat{\theta}) + Bias^2(\hat{\theta})}$$

2) First 10 samples from collision.csv

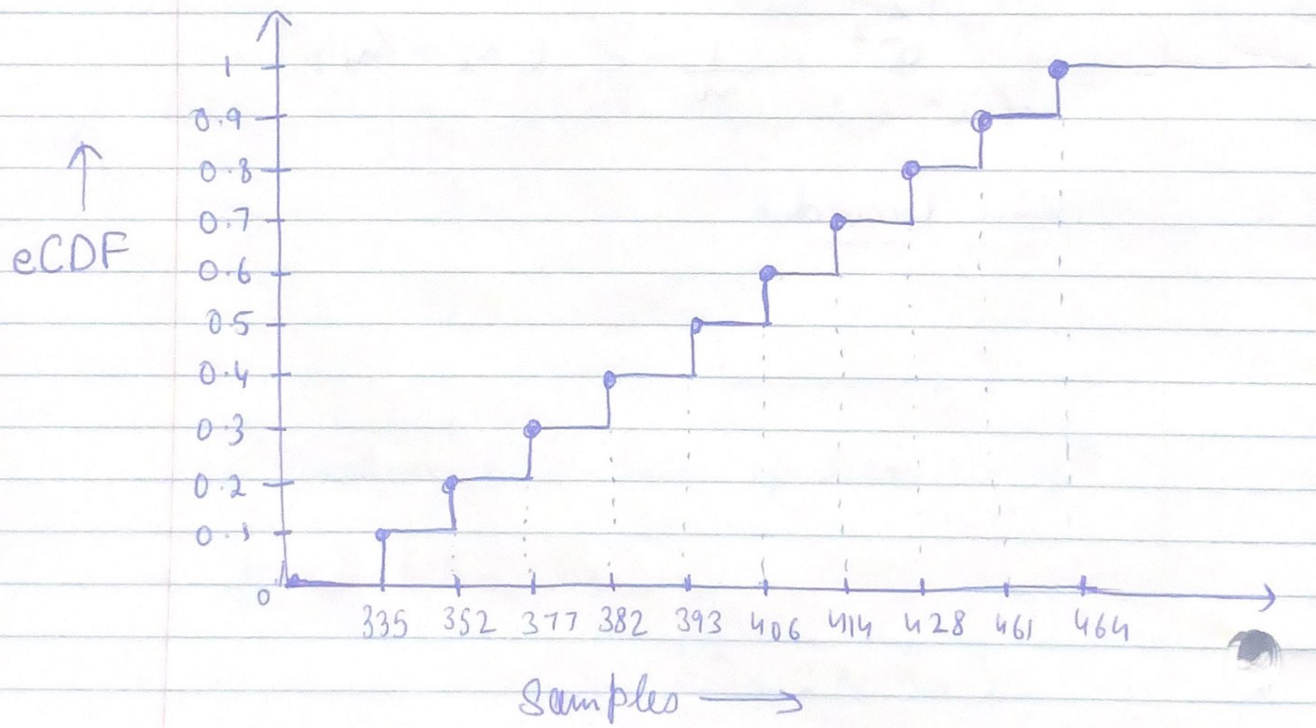
$= \{393, 377, 414, 382, 335, 461, 428, 406, 464, 352\}$

After sorting in ascending order:

$\rightarrow \{335, 352, 377, 382, 393, 406, 414, 428, 461, 464\}$

Here we total $n = 10$
with $\{X_1, X_2, ---- X_n\}$ samples
each occurring once

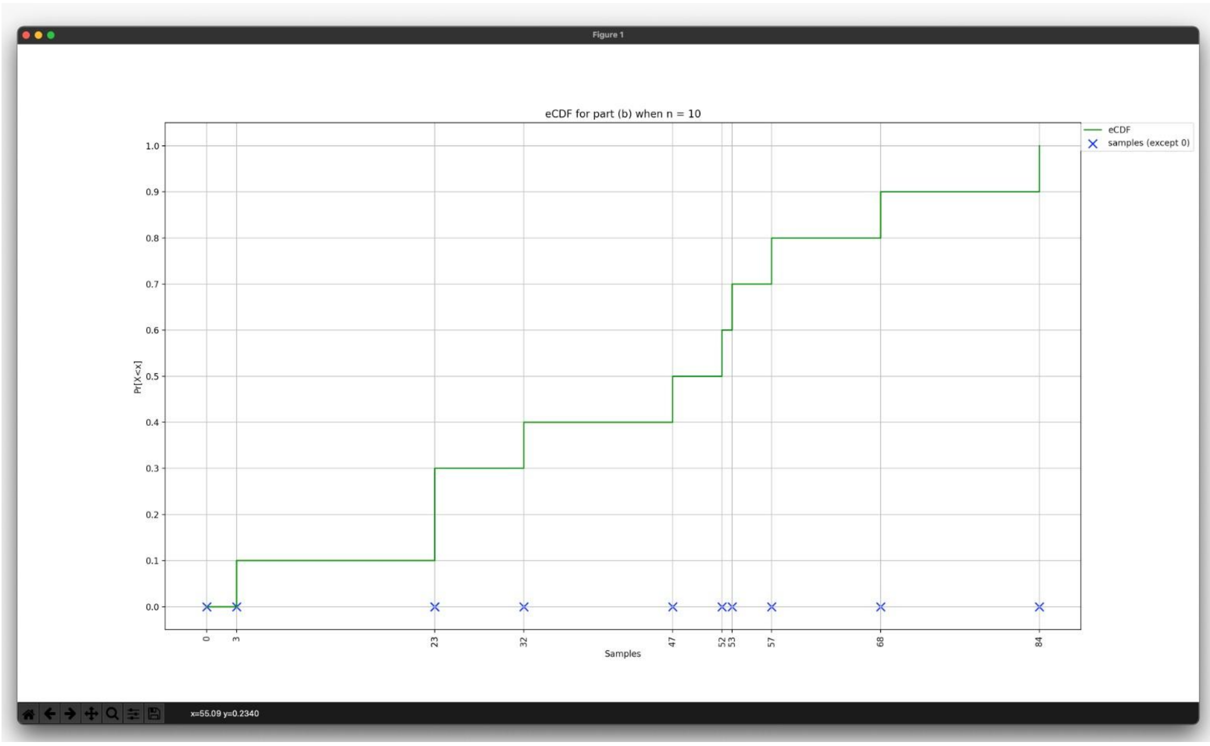
$\therefore P(X_i) = \frac{1}{n} = \frac{1}{10} = 0.1$

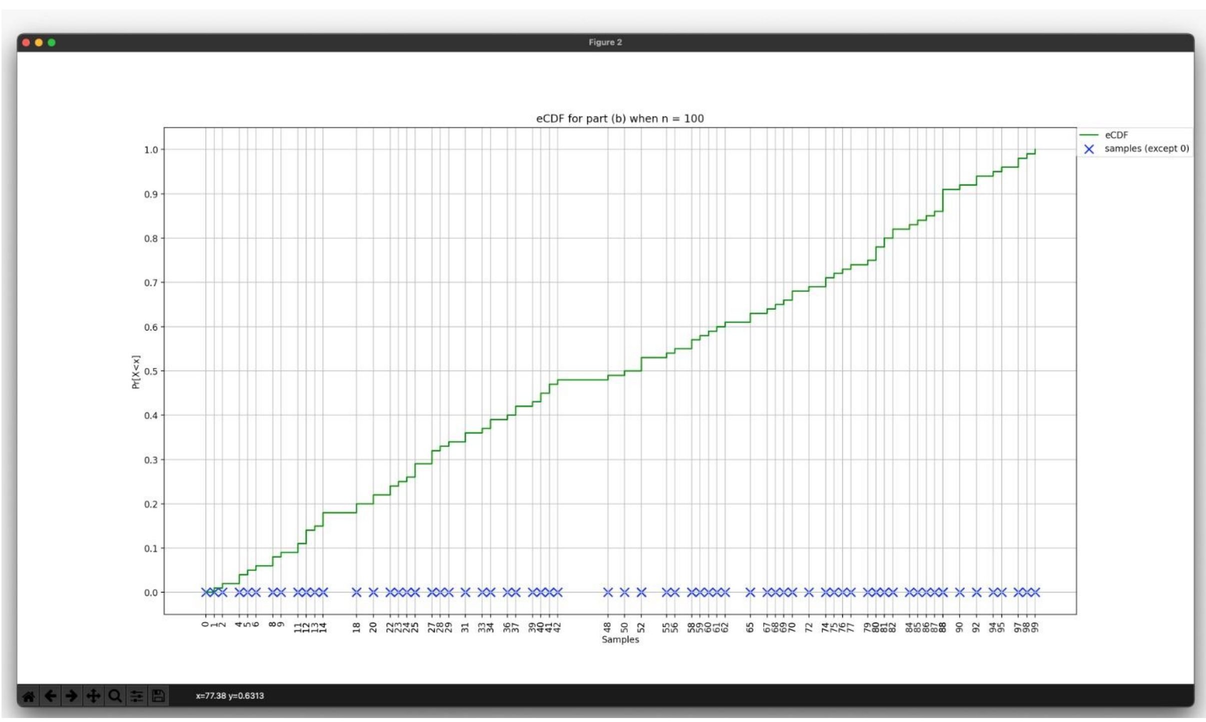


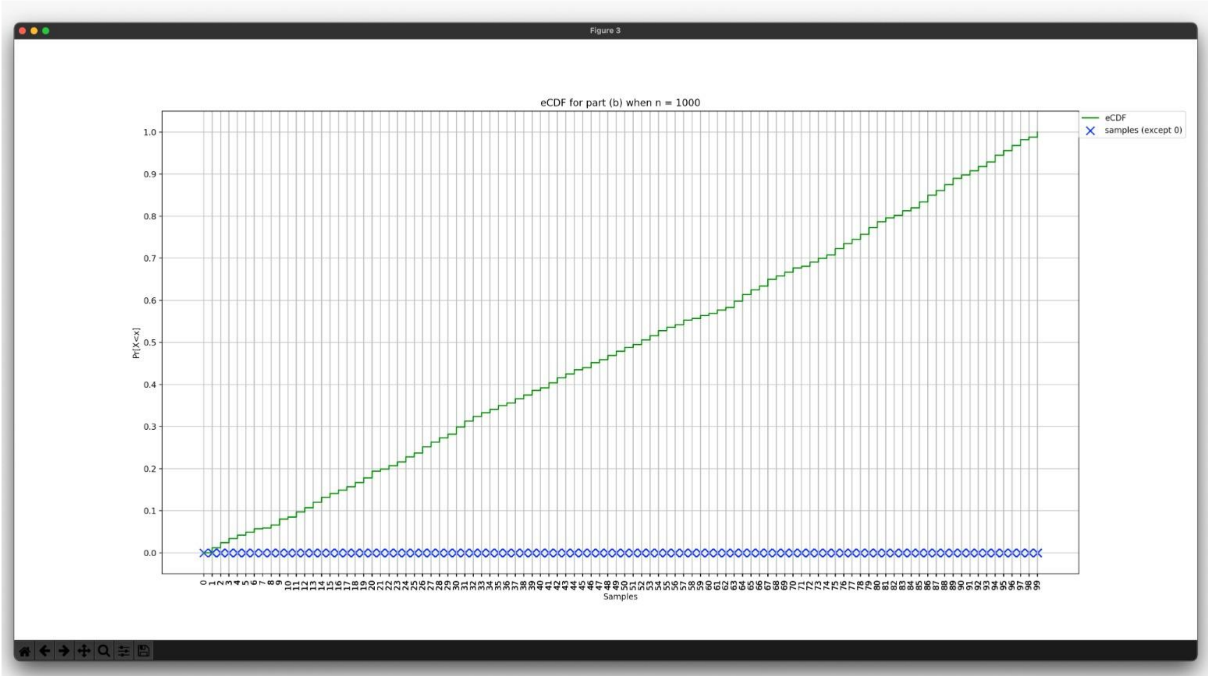
3)

a) Code has been implemented

b)



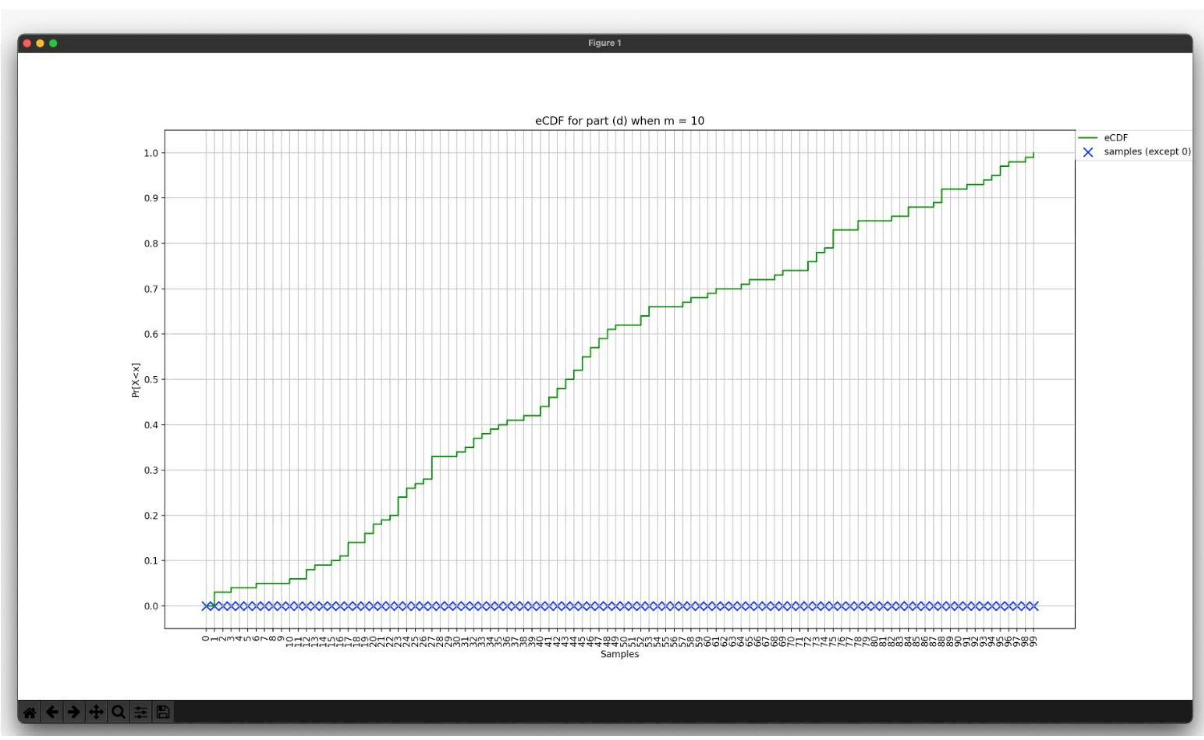




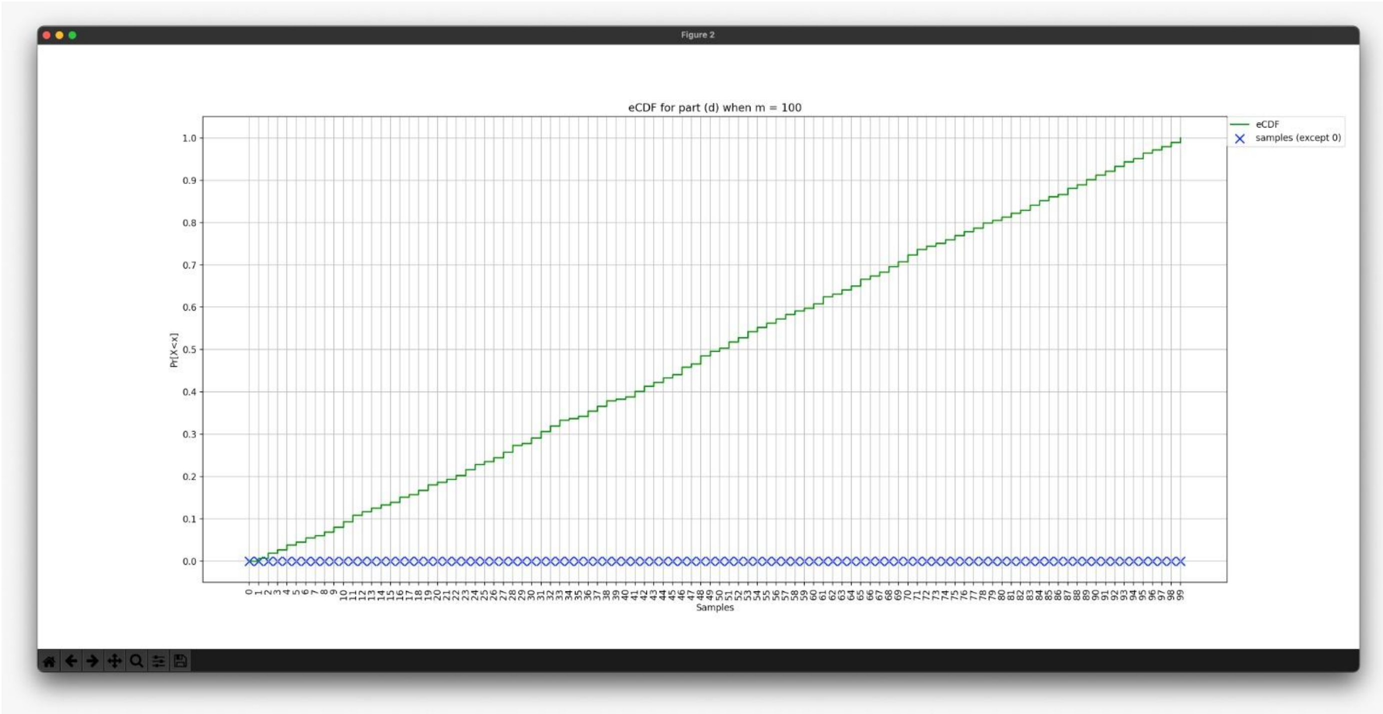
We observe that as the number of samples increases, the cdf starts looking like a straight line

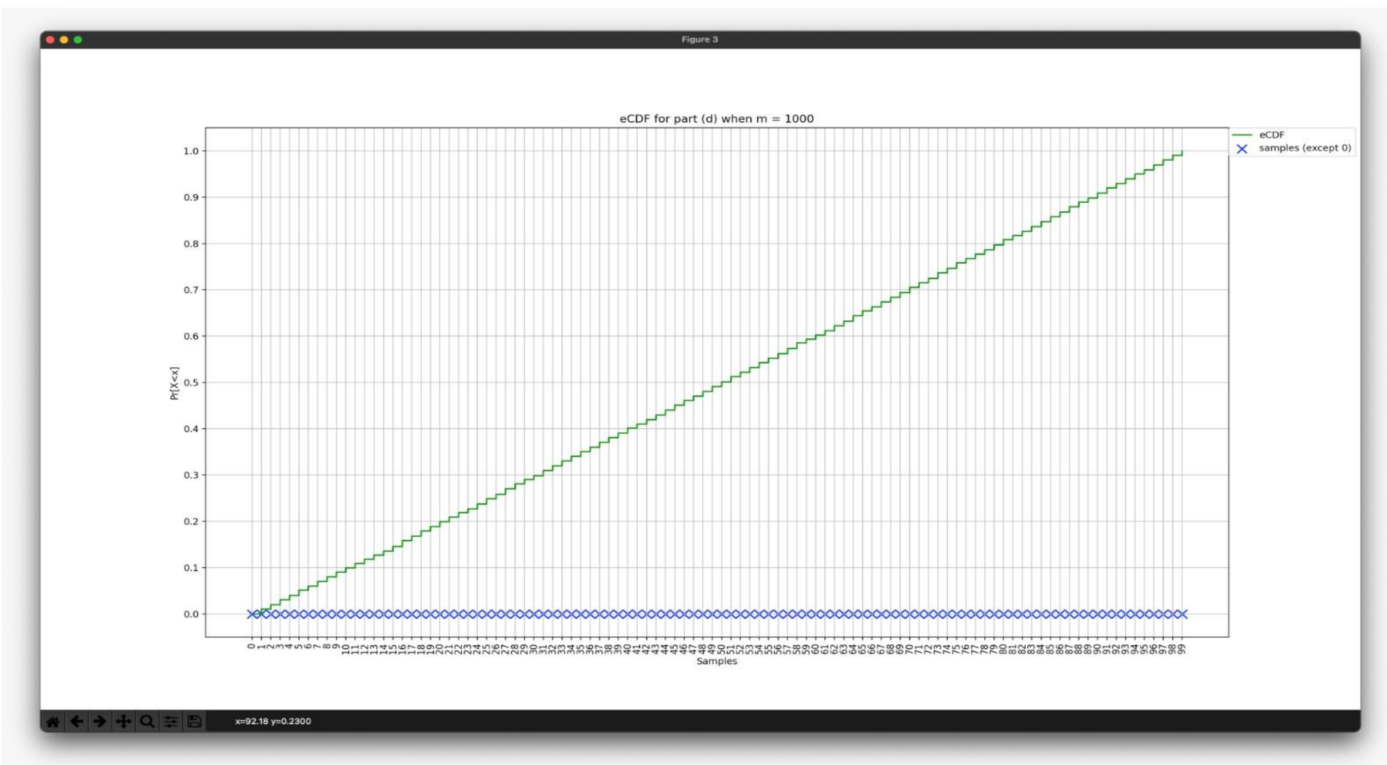
c) Code has been implemented

d)



‘’





We observe that as the number of samples increases, the cdf starts looking like a straight line e)

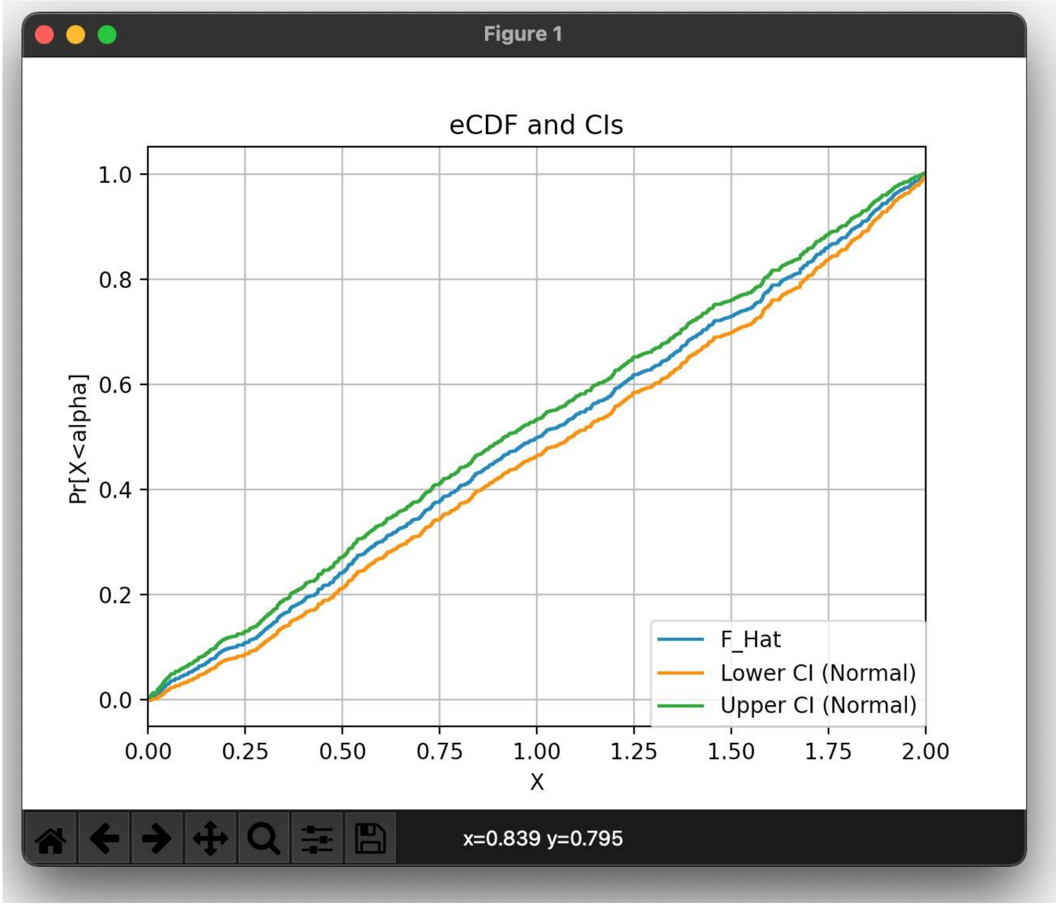


Figure 1
eCDF and CIs
Pr[X<alpha]
X
F_Hat
Lower CI (Normal)
Upper CI (Normal)
x=0.839 y=0.795

f)

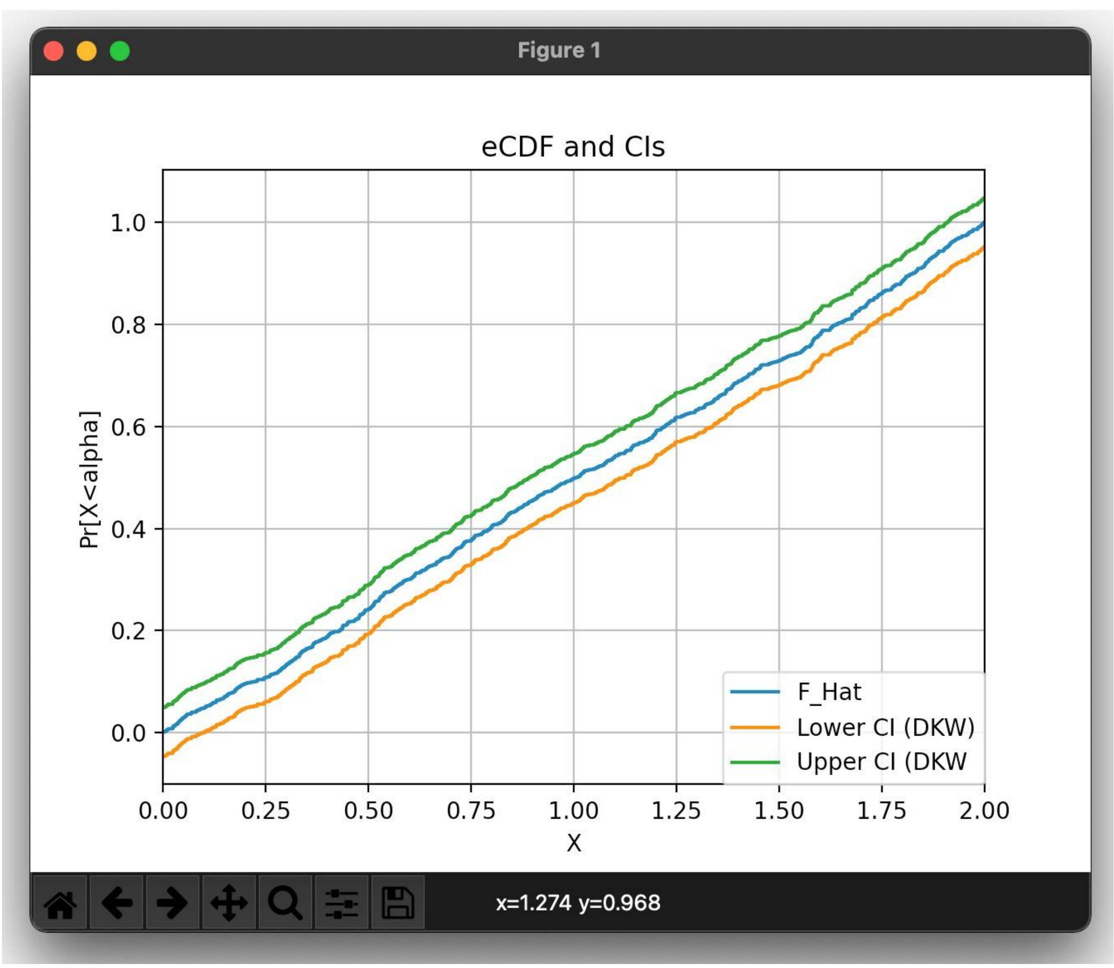
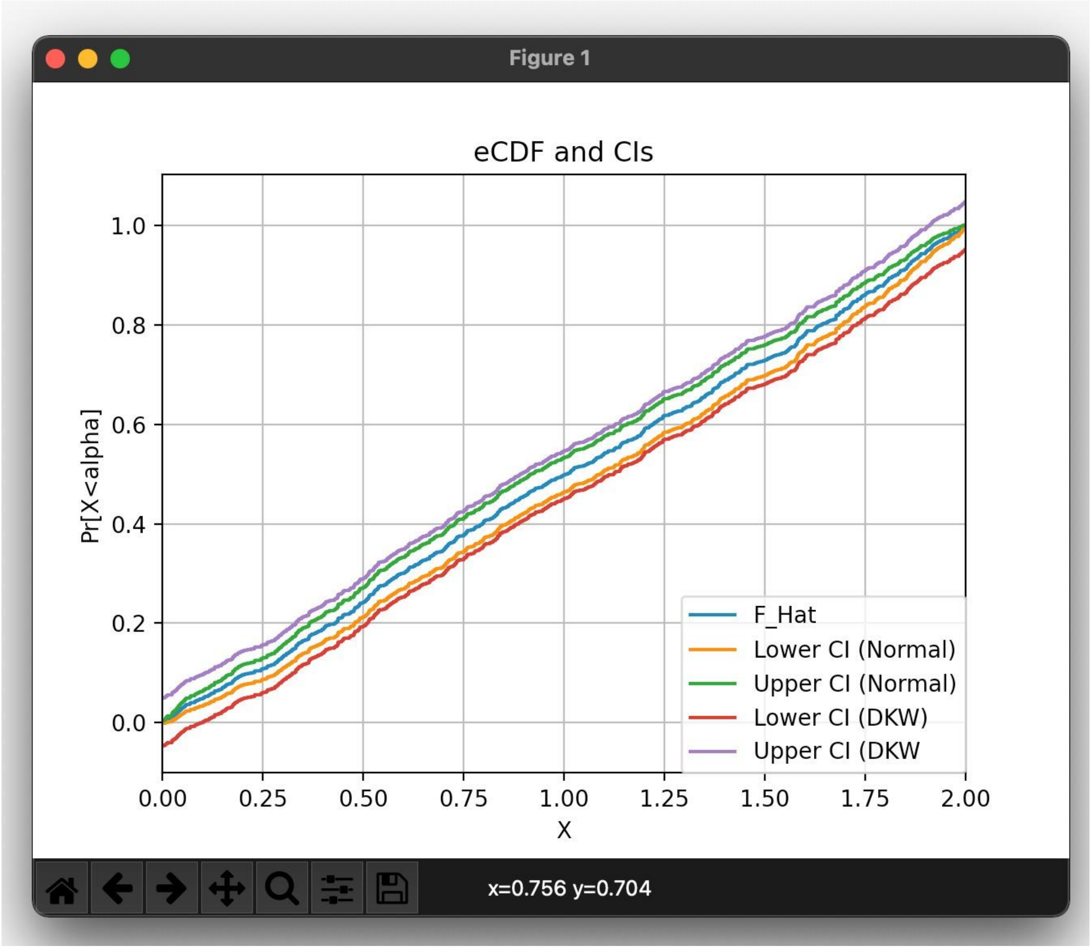


Figure 1
eCDF and CIs
Pr[X<alpha]
X
F_Hat
Lower CI (DKW)
Upper CI (DKW
x=1.274 y=0.968



Normal Based CI is tighter

4) (a) To prove : $\hat{\sigma}^2 = \frac{1}{n} \sum_{i=1}^{n} (X_i - \bar{X}_n)^2$

Given → $\bar{X}_n = \frac{1}{n} \sum_{i=1}^{n} X_i$

We know that,

variance i.e. $\sigma^2 = E[X^2] - (E[X])^2$

∴ Estimator $\hat{\sigma}^2 = \sum_{i=1}^{n} X_i^2 \hat{P}(X_i) - \left(\sum_{i=1}^{n} X_i \hat{P}(X_i) \right)^2$ —— (1.)

Also, $\hat{P}(X_i) = \frac{1}{n}$ (ePMF)

∴ Substituting this in (1)

$$\Rightarrow \hat{\sigma}^2 = \sum_{i=1}^{n} X_i^2 \left(\frac{1}{n}\right) - \left(\sum_{i=1}^{n} X_i \left(\frac{1}{n}\right) \right)^2$$

$$= \frac{1}{n} \sum_{i=1}^{n} X_i^2 - \frac{1}{n^2} \left(\sum_{i=1}^{n} X_i \right)^2$$

LHS

$$= \frac{1}{n} \sum_{i=1}^{n} X_i^2 - \frac{1}{n^2} (n \bar{X}_n)^2 \qquad \left(\because \bar{X}_n = \frac{1}{n} \sum_{i=1}^{n} X_i \right)$$

$$= \frac{1}{n} \sum_{i=1}^{n} X_i^2 - (\bar{X}_n)^2$$ —— (2.)

Expanding RHS from question we get :

$$\Rightarrow \frac{1}{n} \sum_{i=1}^{n} (X_i - \bar{X}_n)^2$$

$$= \frac{1}{n} \sum_{i=1}^{n} (X_i^2 + \bar{X}_n^2 - 2 X_i \bar{X}_n)$$

$\Rightarrow$ Separate out the terms

$$\Rightarrow \frac{\sum_{i=1}^{n} x_i^2}{n} + \frac{\sum_{i=1}^{n} \bar{x}_n^2}{n} - \frac{2\bar{x}_n}{n} \sum_{i=1}^{n} x_i$$

$$\Rightarrow \frac{\sum_{i=1}^{n} x_i^2}{n} + \frac{n.(\bar{x}_n)^2}{n} - \frac{2\bar{x}_n}{n} \sum_{i=1}^{n} x_i$$

$$\Rightarrow \frac{1}{n} \sum_{i=1}^{n} x_i^2 + (\bar{x}_n)^2 - \frac{2\bar{x}_n}{n} (n\bar{x}_n) \qquad \left(\because \bar{x}_n = \frac{\sum_{i=1}^{n} x_i}{n}\right)$$

$$\Rightarrow \frac{1}{n} \sum_{i=1}^{n} x_i^2 - (\bar{x}_n)^2$$

③

We can see from ② and ③ that LHS = RHS

$$\therefore \hat{\sigma}^2 = \frac{1}{n} \sum_{i=1}^{n} (x_i - \bar{x}_n)^2$$

Hence Proved.

4.) (b) To prove $bias(\hat{\sigma}^2) = -\frac{\sigma^2}{n}$

We know $bias(\hat{\sigma}^2) = E[\hat{\sigma}^2] - \sigma^2$ —— ①

($\because bias(\hat{\theta}) = E[\hat{\theta}] - \theta$)

$\therefore$ Now we have to find

$\underline{\underline{E[\hat{\sigma}^2]}}$

From 4(a) we know $\hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X}_n)^2$

$$\therefore E[\hat{\sigma}^2] = E\left[\frac{1}{n}\sum_{i=1}^{n}(X_i - \bar{X}_n)^2\right]$$

$$= E\left[\frac{1}{n}\sum_{i=1}^{n}(X_i^2 + \bar{X}_n^2 - 2X_i\bar{X}_n)\right]$$

$$= E\left[\frac{1}{n}\left(\sum_{i=1}^{n}X_i^2 + \sum_{i=1}^{n}\bar{X}_n^2 - 2\bar{X}_n\sum_{i=1}^{n}X_i\right)\right]$$

$$= E\left[\frac{1}{n}\left(\sum_{i=1}^{n}X_i^2 + n\cdot\bar{X}_n^2 - 2\bar{X}_n(n\bar{X}_n)\right)\right] \quad \left(\because \frac{\sum X_i}{n} = \bar{X}_n\right)$$

$$= E\left[\frac{1}{n}\left(\sum_{i=1}^{n}X_i^2 + n\cdot\bar{X}_n^2 - 2n\bar{X}_n^2\right)\right]$$

$$= E\left[\frac{1}{n}\left(\sum_{i=1}^{n}X_i^2 - n\bar{X}_n^2\right)\right]$$

$$\stackrel{LOE}{=} \frac{1}{n}\left[E\left[\sum_{i=1}^{n}X_i^2\right] - nE\left[\bar{X}_n^2\right]\right]$$

$$\stackrel{LOE}{=} \frac{1}{n}\left[\sum_{i=1}^{n}E[X_i^2] - n\cdot E[\bar{X}_n^2]\right]$$

$$= \frac{1}{n}\left[n\,E[X_1^2] - nE[\bar{X}_n^{\,2}]\right]$$ (∵ X_i are iid so all $E[X_i]$ are equal)

$$= E[X_1^2] - \frac{n}{n}E[\bar{X}_n^2] \quad \text{(2)}$$

We know variance i.e. $\sigma^2 = E[X^2] - (E[X])^2$

$$\Rightarrow E[X^2] = \sigma^2 + (E[X])^2$$

So we can write (2) as:

$$\Rightarrow \sigma^2 + (E[X_1])^2 - \frac{n}{n}\left(Var(\bar{X}_n) + (E[\bar{X}_n])^2\right)$$

$$\Rightarrow \sigma^2 + (E[X_1])^2 - 1\left(Var\left(\frac{\sum_{i=1}^{n} X_i}{n}\right) + \left(E\left[\frac{\sum_{i=1}^{n} X_i}{n}\right]\right)^2\right)$$

$\Rightarrow$ (substituting $\bar{X}_n = \frac{\sum_{i=1}^{n} X_i}{n}$)

$$\Rightarrow \sigma^2 + (E[X_1])^2 - Var\left(\frac{\sum_{i=1}^{n} X_i}{n}\right) - \left(E\left(\frac{\sum_{i=1}^{n} X_i}{n}\right)\right)^2$$

using LOV and LOE $\Rightarrow \sigma^2 + (E[X_1])^2 - \frac{1}{n^2}\left(\sum_{i=1}^{n} Var(X_i)\right) - \frac{1}{n^2}\left(\sum_{i=1}^{n} E[X_i]\right)^2$

(Taking out $1/n$ from Var() becomes $1/n^2$)

$\rightarrow$ using iid we know all $Var(X_i)$ and $E(X_i)$ are equal

$$\Rightarrow \sigma^2 + (E[X_1])^2 - \frac{1}{n^2}(n \cdot Var(X_1)) - \frac{1}{n^2}(n E[X_1])^2$$

(Substituting $\sum_{i=1}^{n} Var(X_i)$ as $n \cdot Var(X_1)$ and same for $(\sum_{i=1}^{n} E[X_i])^2$)

$$\Rightarrow \sigma^2 + (E[X_1])^2 - \frac{1}{n^2}(n\sigma^2) - \frac{1}{n^2}(n^2 E[X_1]^2) \quad (\because Var(X) = \sigma^2)$$

$$\Rightarrow \sigma^2 - \frac{\sigma^2}{n} + \cancel{(E[X_1])^2} - \cancel{(E[X_1])^2}$$

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

$$\therefore E[\sigma'^2] = \sigma^2 - \frac{\sigma^2}{n}$$

Substituting in (1)

$$bias(\sigma'^2) = \sigma^2 - \frac{\sigma^2}{n} - \sigma^2$$

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

Hence Proved

4.) (c) $\text{Kurt}[X] = E[(X-\mu)^4] / \sigma^4$

To find $\hat{\text{Kurt}}[X]$?

$$E[(X-\mu)^4] = \sum_{i=1}^{n} (X_i - \mu)^4 P(X_i) \qquad \left(\because E[z^4] = \sum z^4 (P(z))\right)$$

$$\therefore \hat{\text{Kurt}}[X] = \frac{\sum_{i=1}^{n} (X_i - \hat{\mu})^4 \hat{P}(X_i)}{\hat{\sigma}^4}$$

From 4(a) $\hat{\sigma}^2 = \frac{1}{n} \sum_{i=1}^{n} (X_i - \bar{X}_n)^2$

$$\therefore (\hat{\sigma})^4 = \left(\frac{1}{n} \sum_{i=1}^{n} (X_i - \bar{X}_n)^2\right)^2$$

& $\hat{P}(X_i) = 1/n$

We also know $\hat{\mu} = \bar{X}_n$, so replacing this in $\hat{\text{Kurt}}[X]$ expression

$$\Rightarrow \hat{\text{Kurt}}[X] = \frac{\sum_{i=1}^{n} (X_i - \bar{X}_n)^4 \frac{1}{n}}{\left(\frac{1}{n} \sum_{i=1}^{n} (X_i - \bar{X}_n)^2\right)^2}$$

$$= n \cdot \frac{\sum_{i=1}^{n} (X_i - \bar{X}_n)^4}{\left(\sum_{i=1}^{n} (X_i - \bar{X}_n)^2\right)^2}$$

4.) (d.) $\rho = \frac{E[XY]}{\sigma_x \sigma_y} - \frac{E[X]\,E[Y]}{\sigma_x \sigma_y}$

$D = \{ (x_1, y_1)\ (x_2, y_2)\ \cdots\ (x_n, y_n) \}$

Here ePMF we can define as

$$\hat{P}_{XY}(x_i y_j) \begin{cases} 1/n & \text{if } i = j \\ 0 & \text{otherwise} \end{cases}$$

$$\therefore \hat{P}(x_i y_i) = \frac{1}{n}$$

$$\rho = \sum_{i=1}^{n} \frac{x_i y_i \, P(x_i y_i)}{\sigma_x \sigma_y} - \sum_{i=1}^{n} \frac{x_i P(x_i)}{\sigma_x} \cdot \sum_{i=1}^{n} \frac{y_i P(y_i)}{\sigma_y}$$

$$\left(\because E[X] = \sum_{i=1}^{n} x_i P(x_i) \right)$$

$$\Rightarrow \hat{\rho} = \sum_{j=1}^{n} \frac{x_i y_i \, \hat{P}(x_i y_i)}{\hat{\sigma}_x \hat{\sigma}_y} - \sum_{i=1}^{n} \frac{x_i \hat{P}(x_i)}{\hat{\sigma}_x} \cdot \sum_{i=1}^{n} \frac{y_i \hat{P}(y_i)}{\hat{\sigma}_y}$$

Also

$$\hat{P_x}(\alpha) = \hat{P}_x(x_i) \begin{cases} 1/n & \alpha = x_i \\ 0 & \text{otherwise} \end{cases}$$

$$\hat{P_y}(\alpha) = \hat{P}_y(y_i) \begin{cases} 1/n & \alpha = y_i \\ 0 & \text{otherwise} \end{cases}$$

$$\Rightarrow \hat{\rho} = \sum_{i=1}^{n} \frac{X_i Y_i}{\hat{\sigma}_X \hat{\sigma}_Y} \left(\frac{1}{n}\right) - \sum_{i=1}^{n} \frac{X_i (1/n)}{\hat{\sigma}_X} \cdot \sum_{i=1}^{n} \frac{Y_i (1/n)}{\hat{\sigma}_Y}$$

$$= \frac{1}{n} \sum_{i=1}^{n} \frac{X_i Y_i}{\hat{\sigma}_X \hat{\sigma}_Y} - \frac{1}{n^2} \frac{\sum_{i=1}^{n} X_i}{\hat{\sigma}_X} \frac{\sum_{i=1}^{n} Y_i}{\hat{\sigma}_Y}$$

From 4(a) $\hat{\sigma}_X = \sqrt{\frac{1}{n} \sum_{i=1}^{n} (X_i - \bar{X}_n)^2}$

and $\hat{\sigma}_Y = \sqrt{\frac{1}{n} \sum_{i=1}^{n} (Y_i - \bar{Y}_n)^2}$

$$\therefore \hat{\rho} = \frac{1}{n} \sum_{i=1}^{n} \frac{X_i Y_i}{\hat{\sigma}_X \hat{\sigma}_Y} - \frac{1}{n^2} \frac{\sum_{i=1}^{n} X_i \sum_{i=1}^{n} Y_i}{\sqrt{\frac{1}{n} \sum_{i=1}^{n} (X_i - \bar{X}_n)^2} \sqrt{\frac{1}{n} \sum_{i=1}^{n} (Y_i - \bar{Y}_n)^2}}$$

where $\bar{X}_n = \frac{\sum_{i=1}^{n} X_i}{n}$

and $\bar{Y}_n = \frac{\sum_{i=1}^{n} Y_i}{n}$

$$\Rightarrow \hat{\rho} = \frac{1}{n \sqrt{\frac{1}{n} \sum_{i=1}^{n} (X_i - \bar{X}_n)^2} \sqrt{\frac{1}{n} \sum_{i=1}^{n} (Y_i - \bar{Y}_n)^2}} \cdot \left(\sum_{i=1}^{n} X_i Y_i - \frac{1}{n} \sum_{i=1}^{n} X_i \sum_{i=1}^{n} Y_i \right)$$

5) Consistency of eCDF :

(a) $D = \{X_1 .. X_n\} \overset{iid}{\sim}$ CDF ⇒ F

eCDF ⇒ $\hat{F}$

We know, $\hat{F} = \dfrac{\sum_{i=1}^{n} I(X_i \le \alpha)}{n}$

↳ expectation on both sides,

$E(\hat{F}) = E\left[\dfrac{\sum_{i=1}^{n} I(X_i \le \alpha)}{n}\right]$

$= \dfrac{1}{n} E\left[\sum_{i=1}^{n} I(X_i \le \alpha)\right]$ (∵ constant taken outside)

$\underset{=}{LOE} \; \dfrac{1}{n} \cdot \sum_{i=1}^{n} E[I(X_i \le \alpha)]$

$X_1 .. X_n$ $\underset{=}{iids}$ $\dfrac{1}{n} \times n \cdot E[I(X_1 \le \alpha)]$

$= E[I(X_1 \le \alpha)] = P_r(X \le \alpha) = F_X(\alpha)$

(∵ Assignment 1 → Q.5(a))

(b) $Bias(\hat{F}) = E(\hat{F}) - F = F_X(\alpha) - F_X(\alpha) = 0$

(from a)

(c) $se(\hat{F}) = \sqrt{Var(\hat{F}_X(i))} = \sqrt{var\left(\dfrac{\sum_{i=1}^{n} I(X_i \le \alpha)}{n}\right)}$

$\dfrac{1}{n}$ comes out of the 'Var' expression as $\dfrac{1}{n^2}$,

$$= \sqrt{\frac{1}{n^2} \text{Var}\left[\sum_{i=1}^{n} I(X_i \le \alpha)\right]}$$

$$\overset{\text{Linearity of variance}}{=} \sqrt{\frac{1}{n^2} \cdot \sum_{i=1}^{n} \text{Var}\left[I(X_i \le \alpha)\right]}$$

We know, $X_1 \ldots X_n$ are iids and $\text{var}[I[X \le \alpha]) = P.(1-P)$
(From Assignment 1, Qn 5(b))

$$\therefore \Rightarrow \sqrt{\frac{1}{n^2} \cdot n \; Pr(X_i \le \alpha) \cdot (1 - Pr(X_i \le \alpha))}$$

But, $Pr(X_i \le \alpha) = F_X(\alpha)$

$$\therefore se(\hat{F}) = \sqrt{\frac{F_X(\alpha) \cdot (1 - F_X(\alpha))}{n}}$$

As $n \to \infty$, $se(\hat{F}) = 0$

(d) We know as $n \to \infty$, $se(\hat{F}) = 0$ (from (c))
and $\text{bias}(\hat{F}) = 0$ (from (b))

$\therefore \hat{F}$ is a consistent estimator of F.

6)

(a) $\hat{\theta} = \frac{1}{n}\sum_{i=1}^{n} X_i$, where $X_i \overset{iid}{\sim}$ ~ Bernolli (θ)

Bias = $E[\hat{\theta}] - \theta$
(true value)

$E[\hat{\theta}] = E\left[\frac{1}{n} \cdot \sum_{i=1}^{n} X_i\right]$

taking constant outside and applying LOE,

$\Rightarrow \frac{1}{n}\sum_{i=1}^{n} E[X_i]$, $X_i \overset{iid}{\sim}$ $\Rightarrow$ all X are identical

Also, for Bernoulli $E[X_i] = p$.
Here, $p = \theta$

$E[\hat{\theta}] = \frac{1}{n} \times (n \times \theta) = \theta$

$\therefore$ Bias $= E[\hat{\theta}] - \theta = \theta - \theta = 0$.

$se(\hat{\theta}) = \sqrt{Var(\hat{\theta})}$

$var(\hat{\theta}) = var\left(\frac{1}{n}\sum_{i=1}^{n} X_i\right)$

$\frac{1}{n}$ comes out as $\frac{1}{n^2}$

$= \frac{1}{n^2} var\left(\sum_{i=1}^{n} X_i\right)$

Linearity of Variance $\Rightarrow \frac{1}{n^2}\sum_{i=1}^{n} var(X_i)$

X_i iid and $\sim$ Bernoulli (θ),

$\therefore \text{Var}(X_i) = p(1-p) \Rightarrow$ Here, $\theta.(1-\theta)$

$$\Rightarrow \frac{1}{n^2} \cdot n\left(\theta \cdot (1-\theta)\right) = \frac{\theta.(1-\theta)}{n} = \text{var}(\hat{\theta})$$

$$se = \sqrt{\text{Var}(\hat{\theta})} = \sqrt{\frac{\theta.(1-\theta)}{n}}$$

From Assignment 3, Question 1:

$$MSE = \text{Bias}^2(\hat{\theta}) + \text{Var}(\hat{\theta})$$
$$= 0 + \frac{\theta.(1-\theta)}{n} = \frac{\theta(1-\theta)}{n}$$

(b) Normal based CIs are applicable here if:

(i) $\rightarrow$ $\hat{\theta} \sim$ Normal distribution

(ii) $\rightarrow E[\hat{\theta}] = \theta$ ↳ where it is centered around true value.

For (ii), we know from (a) that $\text{Bias}(\hat{\theta}) = 0$. So, (ii) is satisfied.

For (i),

$$\hat{\Theta} = \frac{1}{n} \sum_{i=1}^{n} X_i$$

$X_1, X_2, X_3 \ldots X_n$ are iid RV with mean 'μ' and variance 'σ^2'.

★ By CLT, $\bar{X}_n$, (ie. $\sum \frac{X_i}{n}$) $\xrightarrow{n \to \infty}$ Nor $\left(\mu, \frac{\sigma^2}{n}\right)$

Hence, $\bar{X_n}$ here is $\hat{\Theta}$ which $\sim$ Normally distributed by CLT.

So, both (i) & (ii) are satisfied

∴ Normal based CIs are applicable here.

If $\hat{\Theta} \sim \text{Normal}(\Theta, se^2)$ then a $(1-\alpha)$ CI for $\hat{\Theta}$ is $(\hat{\Theta} - Z_{\alpha/2}\, se,\ \hat{\Theta} + Z_{\alpha/2}\, se)$

From part (a), $se(\hat{\Theta}) = \sqrt{\frac{\Theta(1-\Theta)}{n}}$

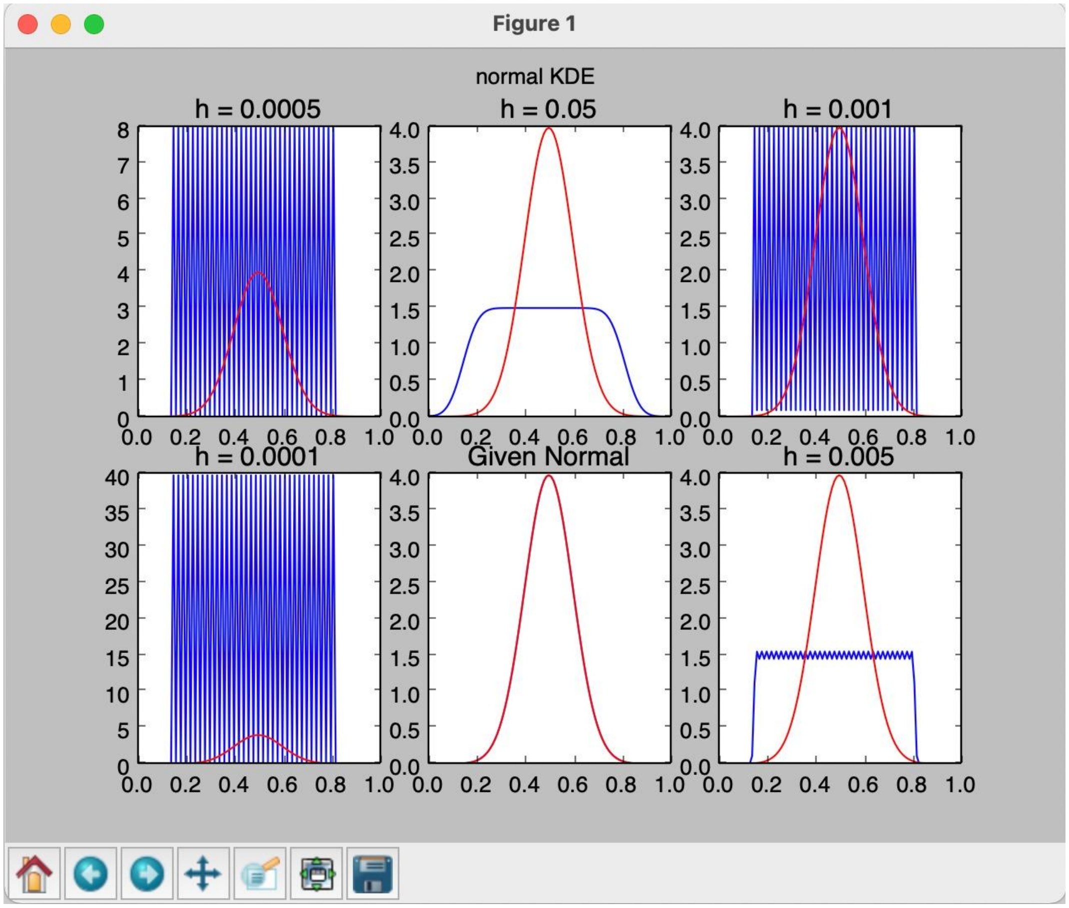
∴ CI for $\hat{\Theta} = \hat{\Theta} \pm Z_{\alpha/2} \sqrt{\frac{\Theta(1-\Theta)}{n}}$

7)

a) Code has been implemented

b)

Normal KDE comparison

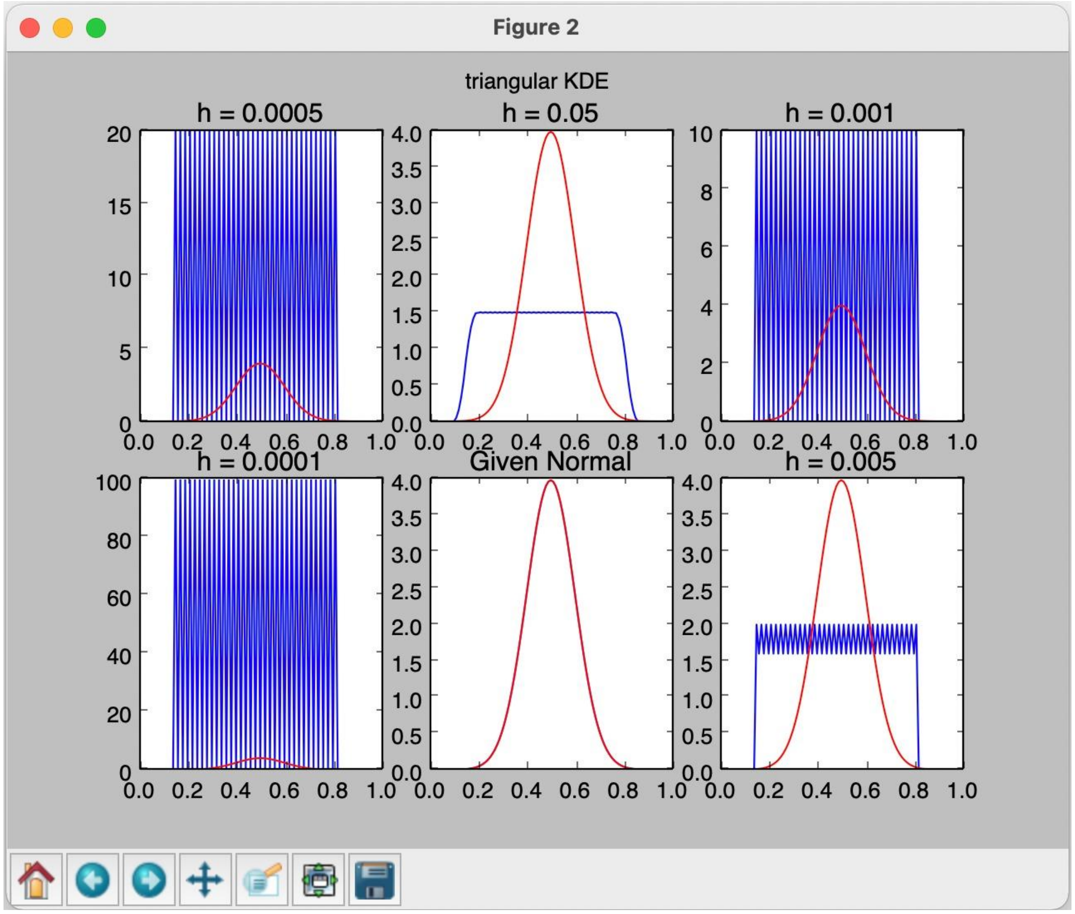


Difference in mean and standard deviation for Normal distribution

```
Type of KDE: normal

 h = 0.0001
Percentage Difference w.r.t True Mean: 1256.40434608
Percentage Difference w.r.t True Variance: 19506.4861551

 h = 0.0005
Percentage Difference w.r.t True Mean: 171.280877236
Percentage Difference w.r.t True Variance: 684.259422673

 h = 0.001
Percentage Difference w.r.t True Mean: 38.5654558441
Percentage Difference w.r.t True Variance: 91.8690520087

 h = 0.005
Percentage Difference w.r.t True Mean: 0.195733075916
Percentage Difference w.r.t True Variance: -72.5533203197

 h = 0.05
Percentage Difference w.r.t True Mean: -0.0025307461222
Percentage Difference w.r.t True Variance: -79.0731503347
```

c)

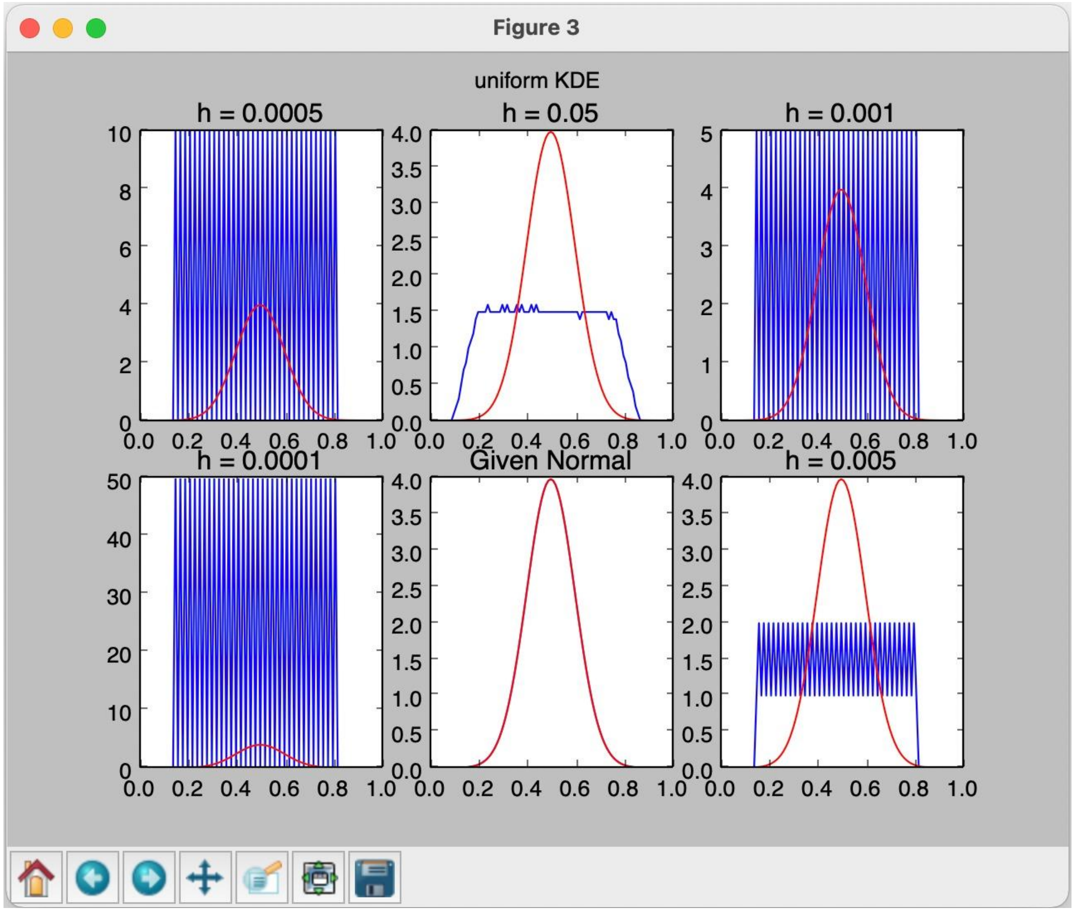
<u>Triangular KDE</u>



Difference in mean and standard deviation for Triangular distribution

```
Type of KDE: triangular

 h = 0.0001
Percentage Difference w.r.t True Mean: 3300.00148572
Percentage Difference w.r.t True Variance: 123091.185735

 h = 0.0005
Percentage Difference w.r.t True Mean: 580.000297143
Percentage Difference w.r.t True Variance: 4827.64742941

 h = 0.001
Percentage Difference w.r.t True Mean: 240.000148572
Percentage Difference w.r.t True Variance: 1131.91185735

 h = 0.005
Percentage Difference w.r.t True Mean: 20.8000527866
Percentage Difference w.r.t True Variance: -58.4900864304

 h = 0.05
Percentage Difference w.r.t True Mean: 4.36975475816e-05
Percentage Difference w.r.t True Variance: -74.9906979361
```

Uniform KDE



Difference in mean and standard deviation for Uniform distribution

```
Type of KDE: uniform

 h = 0.0001
Percentage Difference w.r.t True Mean: 1600.00074286
Percentage Difference w.r.t True Variance: 30697.7964338

 h = 0.0005
Percentage Difference w.r.t True Mean: 240.000148572
Percentage Difference w.r.t True Variance: 1131.91185735

 h = 0.001
Percentage Difference w.r.t True Mean: 70.0000742858
Percentage Difference w.r.t True Variance: 207.977964338

 h = 0.005
Percentage Difference w.r.t True Mean: 4.36975475816e-05
Percentage Difference w.r.t True Variance: -63.4103791622

 h = 0.05
Percentage Difference w.r.t True Mean: 0.400043872338
Percentage Difference w.r.t True Variance: -76.1004773381
```

We see that h = 0.05 is the best choice for all distributions. As h increases the value is accurate