CSE357 Assignment 3 Solutions

September 29, 2020

1 Question 1

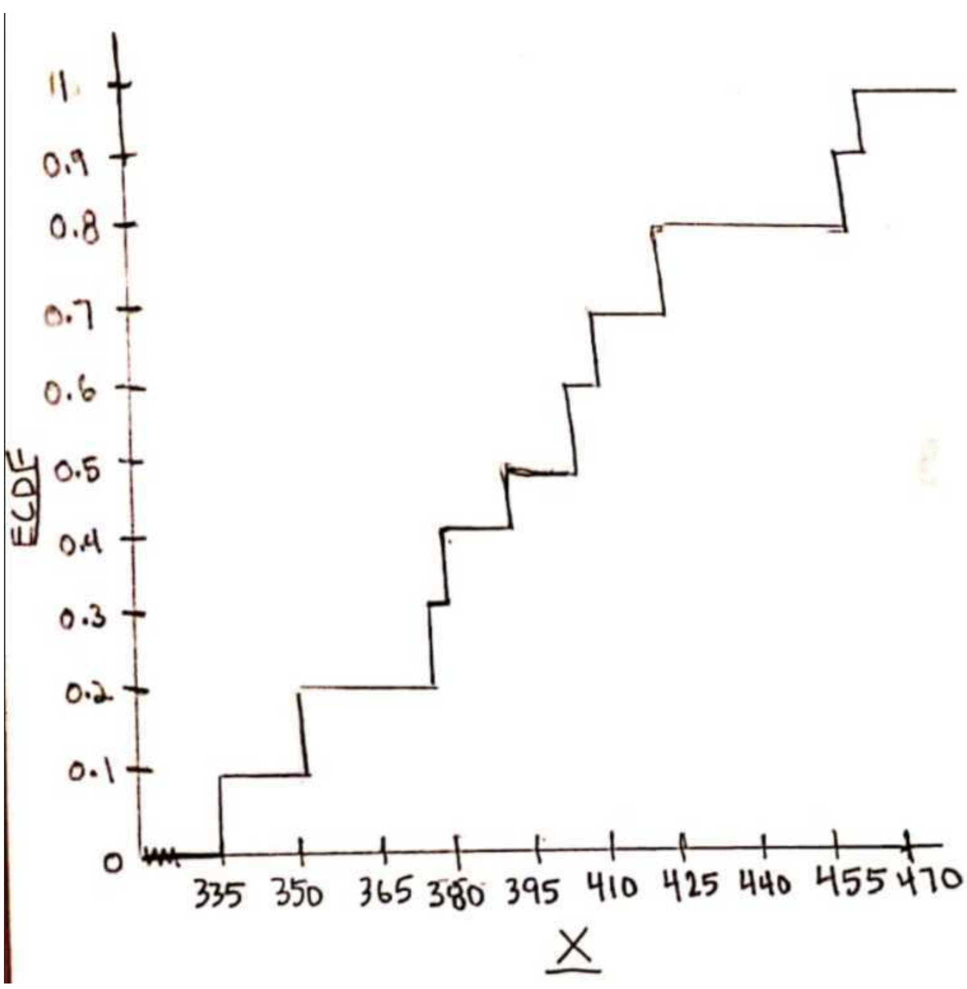
$$MSE = E[(\hat{\theta} - \theta)^2] = E[\hat{\theta}^2 - 2\hat{\theta}\theta + \theta^2]$$
$$MSE = E[\hat{\theta}^2] - 2\theta E[\hat{\theta}] + \theta^2$$

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

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

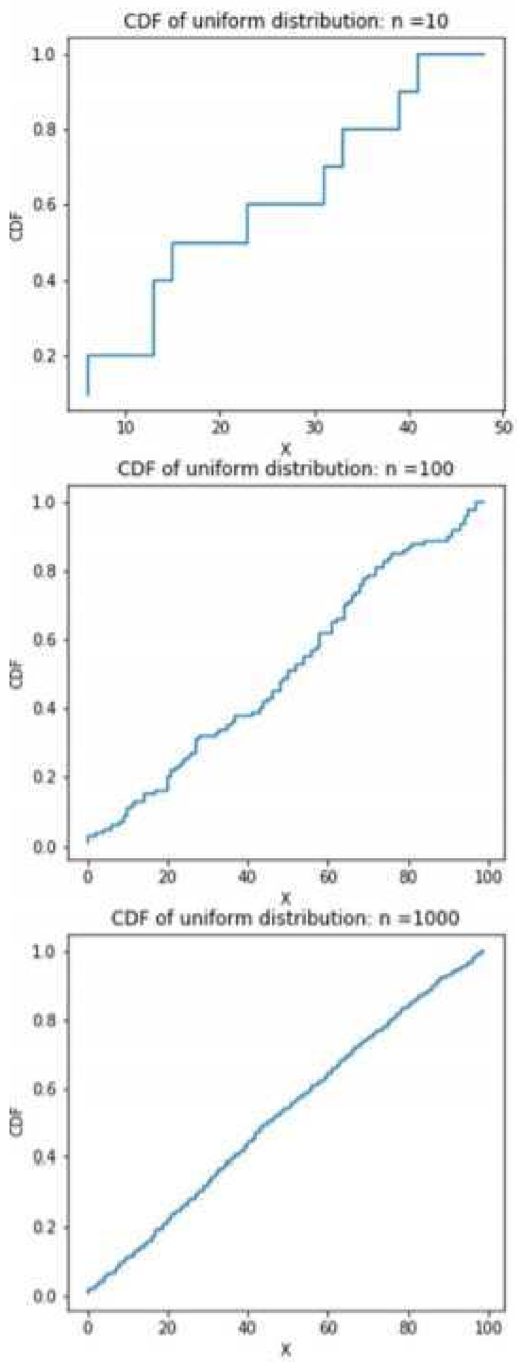
$$MSE = Var(\hat{\theta}) + bias^2[\hat{\theta}]$$
$$MSE = E[\hat{\theta^2}] - (E[\hat{\theta}])^2 + (E[\hat{\theta}])^2 - 2\theta E[\hat{\theta}] + \theta^2$$
$$MSE = E[\hat{\theta^2}] - 2\theta E[\hat{\theta}] + \theta^2$$
$$E[\hat{\theta}^2] - 2\theta E[\hat{\theta}] + \theta^2 = E[\hat{\theta^2}]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

2 Question 2

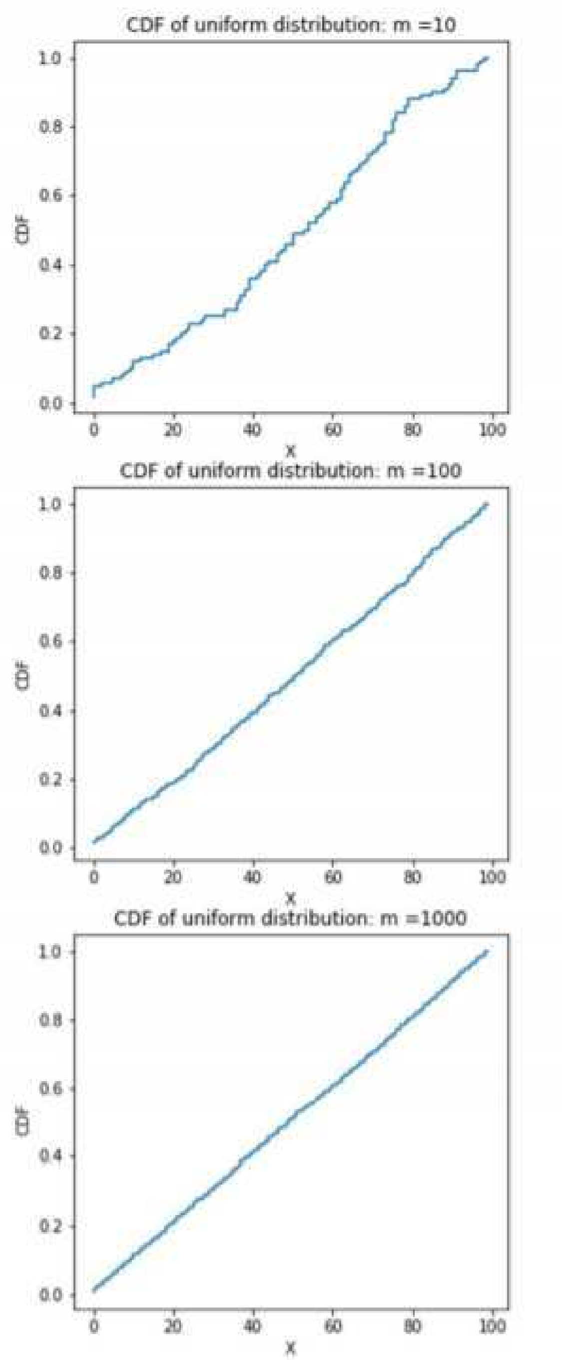


3 Question 3

3b

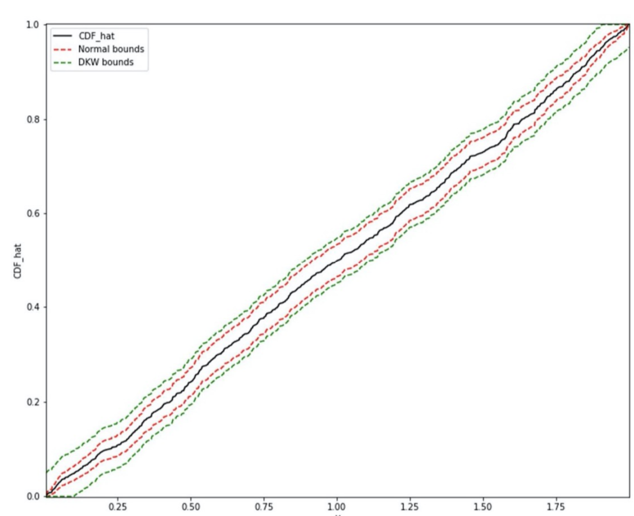


3d



OBSERVATIONS
3b) As value of n (sample size) increases the CDF estimate becomes smoother and the estimated CDF approaches the true CDF.
3d) As value of m (# of rows) increases the CDF estimate becomes smoother and approaches the true CDF even with small sample size because m list of n samples is equivalent to a sample of size n*m.

3e,3f)



From the figure we can see that Normal bound is tighter than DKW bound.

4 Question 4

a) Let $\hat{\sigma}$ be plugin estimator of σ & $\bar{x}_n$ be plugin estimator for mean μ.

We know, $E[X] = \sum_i x_i \, p(x_i)$. Using plugin estimator for $p(x_i)$

we get $P(x_i) = 1/n$ where n = sample size.

$\therefore E[X] = \frac{1}{n}\sum_i x_i = \bar{x}_n$; $E[X^2] = \sum_i x_i^2 \, p(x_i) = \sum_i x_i^2 \hat{p}(x_i)$

$= \frac{1}{n}\sum x_i^2$

$\neq \cancel{\frac{1}{n}}$

$\therefore \hat{\sigma}^2 = E[X^2] - (E[X])^2 = \frac{1}{n}\sum_i x_i^2 - (\frac{1}{n}\sum x_i)^2 = \frac{1}{n}\sum_i x_i^2 - \bar{x}_n^2$ — (1)

$$R.H.S. = \frac{1}{n}\sum_i (x_i - \bar{x}_n)^2 = \frac{1}{n}\sum_i (x_i^2 - 2x_i\bar{x}_n + \bar{x}_n^2)$$

$$= \frac{1}{n}\sum x_i^2 - \frac{2\bar{x}_n}{n}\sum x_i + \frac{\bar{x}_n^2}{n}\sum_i 1$$

$$= \frac{1}{n}\sum_i x_i^2 - 2\bar{x}_n . \bar{x}_n + \bar{x}_n^2 . \frac{\cancel{n}}{\cancel{n}}$$

$$= \frac{1}{n}\sum_i x_i^2 - \bar{x}_n^2 \quad \text{— (2)}$$

From (1) & (2)

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

(b)

The bias of estimator $\hat{\sigma}^2$ is

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

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

$$= E\left[\frac{1}{n}\sum_i (x_i - \mu)^2 - \frac{2}{n}(\bar{x}_n - \mu)\sum_i (x_i - \mu) + (\bar{x}_n - \mu)^2\right]$$

$$= E\left[\frac{1}{n}\sum_i (x_i - \mu)^2 - \frac{2}{n}(\bar{x}_n - \mu) . n . (\bar{x}_n - \mu) + (x_n - \mu)^2\right]$$

$$= E\left[\frac{1}{n}\sum_i (x_i - \mu)^2 - (\bar{x}_n - \mu)^2\right]$$

$$= E\left[\frac{1}{n}\sum_i (x_i - \mu)^2\right] - E[(\bar{x}_n - \mu)^2]$$

$$= \sigma^2 - E[(\bar{x}_n - \mu)^2]$$

$$= \sigma^2 - Var(\bar{x}_n) = \sigma^2 - Var(\frac{1}{n}\sum_i x_i)$$

$$= \sigma^2 - \frac{1}{n^2}\sum_i Var(x_i) = \sigma^2 - \frac{1}{n} . \sigma^2$$

Thus bias is $E[\hat{\sigma}^2] - \sigma^2 = -\frac{1}{n}\sigma^2$

(c) Let $\hat{\sigma}^2$, $\hat{\mu}$ be plugin estimator for σ^2 & μ.

From part A we know $\hat{\sigma}^2 = \frac{1}{n}\sum_i (x_i - \bar{x}_n)^2$ — ①

& $\hat{\mu} = \bar{x}_n$

$$E[(x-\mu)^4] = \sum_i (x_i-\mu)^4 \cdot P(x_i) = \sum_i (x_i-\mu)^4 \, \hat{P}(x_i)$$

$$= \frac{1}{n}\sum_i (x_i - \hat{\mu})^4 \quad \{\text{Plugin estimator}\} \qquad — ②$$

From ①

$$\hat{\sigma}^4 = \frac{1}{n^2}\left(\sum_i (x_i - \bar{x}_n)^2\right)^2 \qquad — ③$$

From ② & ③ & by definition of Kurt[x], we have:

$$\text{Kurt}[x] = \frac{\frac{1}{n}\sum_i (x_i - \bar{x}_n)^4}{\frac{1}{n^2}\left(\sum_i (x_i - \bar{x}_n)^2\right)^2}$$

$$= \frac{n \sum_i (x_i - \bar{x}_n)^4}{\left(\sum_i (x_i - \bar{x}_n)^2\right)^2}$$

(d)

$$\rho = \frac{E[XY] - E[X]E[Y]}{\sigma_x \sigma_y}$$

$$\rho = \frac{E[(X - E[X])(Y - E[Y])]}{\sigma_x \sigma_y}$$

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

5 Question 5

(a)

$$\hat{F}(\alpha) = \frac{\sum_{i=1}^{n} I(X_i < \alpha)}{n}$$

$$E[\hat{F}(\alpha)] = E[\frac{\sum_{i=1}^{n} I(X_i < \alpha)}{n}]$$

$$E[\hat{F}(\alpha)] = \frac{\sum_{i=1}^{n} E[I(X_i < \alpha)]}{n} \qquad \text{By L.O.E}$$

$$E[\hat{F}(\alpha)] = \frac{n * E[I(X_i < \alpha)]}{n} \qquad X_i s \text{ are iid}$$

$$E[\hat{F}(\alpha)] = E[I(X_i < \alpha)] = Pr(X_i < \alpha)$$

$$E[\hat{F}(\alpha)] = F(\alpha)$$

(b)

$$Bias(\hat{F}(\alpha)) = E[\hat{F}(\alpha)] - F(\alpha) = F(\alpha) - F(\alpha) = 0$$

(c)

$$SE(\hat{F}(\alpha)) = \sqrt{Var(\hat{F}(\alpha))}$$

$$Var(\hat{F}(\alpha)) = Var(\frac{\sum_{i=1}^{n} I(X_i < \alpha)}{n})$$

$$Var(\hat{F}(\alpha)) = \frac{1}{n^2} Var(\sum_{i=1}^{n} I(X_i < \alpha))$$

$$Var(\hat{F}(\alpha)) = \frac{n}{n^2} Var(I(X_i < \alpha)) \qquad X_i s \text{ are iid}$$

$$Var(\hat{F}(\alpha)) = \frac{1}{n} Var(I(X_i < \alpha))$$

(d) $As\ n \to \infty\ Bias\left(\hat{F}(\alpha)\right) = 0\ and\ Se\left(\hat{F}(\alpha)\right) \to 0, \therefore \hat{F}$ is consistent estimator of F.

(6)

(a) $\text{Bias}(\hat{\theta}) = E[\hat{\theta}] - \theta$

$\Rightarrow \text{Bias}(\hat{\theta}) = E\left[\frac{1}{n}\sum_i X_i\right] - \theta$

$\text{Bias}(\hat{\theta}) = \frac{1}{n}\sum_i E[X_i] - \theta$ [By LOE]

$\text{Bias}(\hat{\theta}) = \frac{1}{n} n.\theta - \theta$

$= 0$

$\text{Var}(\hat{\theta}) = \text{Var}\left(\frac{1}{n}\sum_i X_i\right)$

$\text{Var}(\hat{\theta}) = \frac{1}{n^2}\sum_i \text{Var}(X_i)$ [$\because X_i$s are iid]

$\Rightarrow \text{Var}(\hat{\theta}) = \frac{1}{n^2} n\theta.(1-\theta)$

$= \frac{\theta(1-\theta)}{n}$

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

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

$MSE(\hat{\theta}) = \frac{\theta(1-\theta)}{n}$

(b) Assuming $\hat{\theta}$ is not near 1 or 0, & the sample size n is large enough, the CI can be estimated by a normal distribution and thus the CI would be

$$\hat{\theta} \pm Z_{1-\alpha/2}\sqrt{\frac{\hat{\theta}(1-\hat{\theta})}{n}}, \quad \text{where } \hat{\theta} = \frac{\Sigma X_i}{n}$$

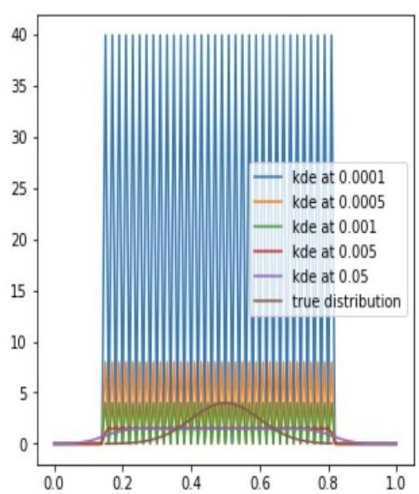
Now, normal based CIs are applicable since

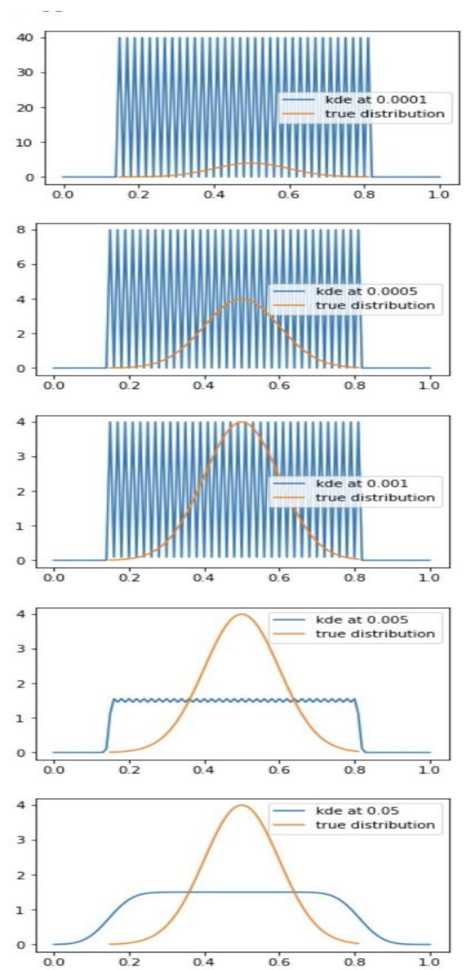
$\hat{\theta} = \frac{\Sigma X_i}{n}$ is normally distributed per CLT.

Q.7.

Normal distribution:
h best for 0.05

```
Sample Mean 13.429740132325447
Sample Variance 355.411195337063

Sample Mean 2.6859481058726695
Sample Variance 14.216447386913261

Sample Mean 1.371934606875201
Sample Variance 3.478053567736609

Sample Mean 0.9920365276520545
Sample Variance 0.49753215114847577

Sample Mean 0.9900735203708438
Sample Variance 0.37934572239674874

Out[15]: <matplotlib.legend.Legend at 0x126755f10>
```



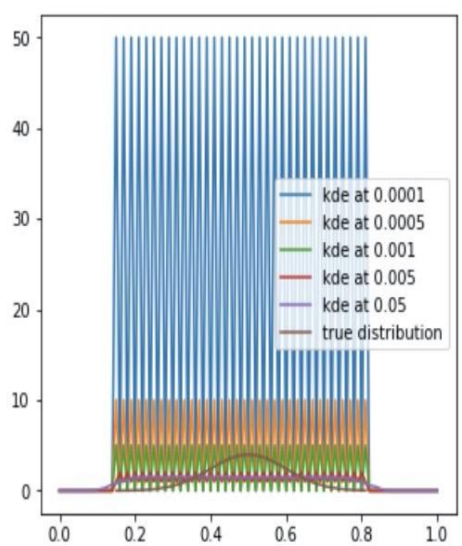


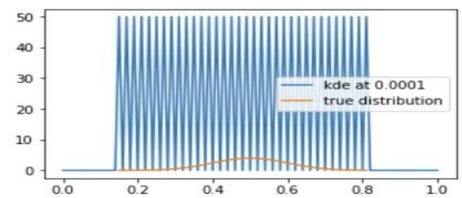
Uniform distribution:

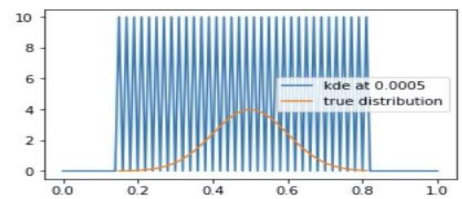
h best for 0.05

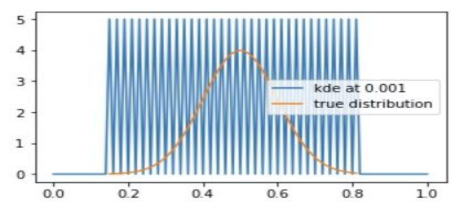
```
Sample Mean 16.831683168316832
Sample Variance 366.98441175719324

Sample Mean 3.366336633663366
Sample Variance 14.67937593566246

Sample Mean 1.683168316831683
Sample Variance 3.574919989949886

Sample Mean 0.9900990099009901
Sample Variance 0.4975359051235125

Sample Mean 0.9940594059405943
Sample Variance 0.3793616096805239
```

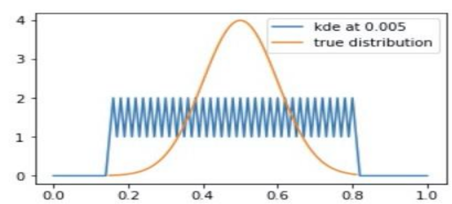
```
Out[16]: <matplotlib.legend.Legend at 0x1262aae10>
```

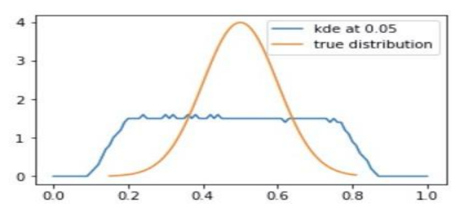












Triangular distribution:
h best for 0.05

```
Sample Mean 33.66336633662982
Sample Variance 764.810824712574

Sample Mean 6.732673267326579
Sample Variance 30.592431919257415

Sample Mean 3.366336633663327
Sample Variance 7.4556930121931995

Sample Mean 1.1960396039603993
Sample Variance 0.5391494062917452

Sample Mean 0.9900990099009902
Sample Variance 0.37934572304646486
```

Out[17]: <matplotlib.legend.Legend at 0x126a42250>