SMAI Assignment-1

2019115007

Q1)

- finite case:

  consider a discrete case where a coin is tossed. Let X be the random variable representing Heads or Tails (H or T) on the coin then

$$range(X) = [H, T]$$

$$p(X = H) = \frac{1}{2}$$
$$p(X = T) = \frac{1}{2}$$

  - conditions Check

$$\forall x \epsilon X, \quad p(X) > 0$$

$$\Sigma p(x) = \frac{1}{2} + \frac{1}{2} = 1$$

    therefore both the conditions are satisfied

- Infinite case:

  Let X be number of tosses until we get a head then

$$range(X) = [1, \infty]$$

$$p(X = n) = \frac{1}{2^n}$$

  - condition check

$$\forall x \epsilon X, \quad p(X) > 0 \quad [since\, 2^n > 0]$$

$$\Sigma p(x) = \frac{1}{2} + \frac{1}{2^2} + \frac{1}{2^3} + \ldots\ldots \frac{1}{2^\infty} = 1$$

    therefore both the conditions are satisfied

Q2)

we know that

$$U(a,b) = \left\{ \frac{1}{b-a} \quad if \quad a <= x <= b \right\}$$

$$0 \qquad otherwise$$

$$var(U) = E[U^2] - E[U]^2$$

$$\Rightarrow \int_{-\infty}^{\infty} x^2 p(x)dx - (\int_{-\infty}^{\infty} xp(x)dx)^2$$

$$\Rightarrow \int_a^b x^2 \frac{1}{b-a} dx - (\int_a^b x \frac{1}{b-a} dx)^2 \ \ (as\ for\ x < a \ \ and \ \ x > b, \ \ p(x) = 0)$$

solving this simple integral we get

$$var(U) = \frac{(b-a)^2}{12}$$

Q3)

Consider a **Normal density** graph with

$$\mu = 0 \ , \ \sigma = 1$$

$$\Rightarrow N(0, 1)$$

Also consider a **Uniform density** graph with parameters

$$a = -\sqrt{3} \ , \ b = \sqrt{3}$$

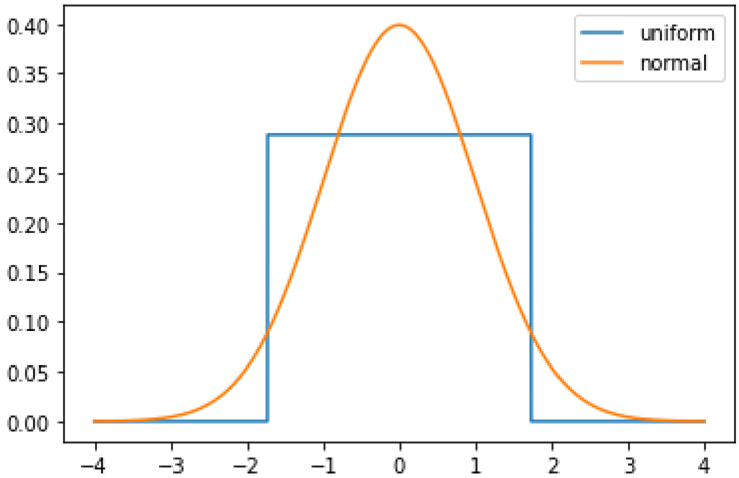
$$\Rightarrow U(-\sqrt{3}, \sqrt{3})$$

then the corresponding mean and variance of this uniform density will be

$$\sigma^2 = \frac{(b-a)^2}{12} = \frac{(2\sqrt{3})^2}{12} = 1$$

$$\mu = \frac{b+a}{2} = \frac{-\sqrt{3}+\sqrt{3}}{2} = 0$$

Clearly both of them has the same mean and variance, but one is Normal Density and the other is Uniform which are different. The graph has been shown below.



Q4)

we have to prove that

$$\sigma^2 = E[X^2] - E[X]^2$$

we know that the variance

$$var(X) = E[(X - \mu)^2]$$

$$\Rightarrow \sum_{i=1}^{i=n} x^2 p(X = x) + \mu^2 \sum_{i=1}^{i=n} p(X = x) - 2\mu \sum_{i=1}^{i=n} xp(X = x)$$

$$\Rightarrow E[X^2] + \mu^2 - 2\mu^2$$

$$\Rightarrow E[X^2] - \mu^2$$

$$\Rightarrow \sigma^2 = E[X^2] - E[X]^2$$

$$Hence \;\; Prooved$$

Q5)

We have to find the Mean and variance of Gaussian pdf,

$$p(x) = \frac{1}{\sigma\sqrt{2\pi}} exp(-\frac{(x-\mu)^2}{2\sigma^2})$$

Proof:

$$E(x) = \int_{-\infty}^{\infty} xp(x)dx$$

$$\Rightarrow \int_{-\infty}^{\infty} x\frac{1}{\sigma\sqrt{2\pi}} exp(-\frac{(x-\mu)^2}{2\sigma^2})dx$$

$$\Rightarrow \int_{-\infty}^{\infty} x\frac{1}{\sigma\sqrt{2\pi}} exp(-\frac{x^2}{2\sigma^2})dx + \int_{-\infty}^{\infty} \mu\frac{1}{\sigma\sqrt{2\pi}} exp(-\frac{x^2}{2\sigma^2})dx$$

$$\Rightarrow I1 + I2$$

now clear $I1 = 0$ as it is a odd function so now we have to solve for $I2$ or we can say that

$$E[x] = \int_{-\infty}^{\infty} \mu\frac{1}{\sigma\sqrt{2\pi}} exp(-\frac{x^2}{2\sigma^2})dx$$

$$\Rightarrow \frac{2}{\sqrt{\pi}}\mu\int_{-\infty}^{\infty} e^{-x^2}dx$$

$$\Rightarrow \frac{2}{\sqrt{\pi}}\mu\frac{\sqrt{\pi}}{2}$$

$$\Rightarrow E[x] = \mu$$

For Variance

$$Var(x) = \int_{-\infty}^{\infty} (x-\mu)^2 p(x)dx$$

$$\Rightarrow \int_{-\infty}^{\infty} (x-\mu)^2\frac{1}{\sigma\sqrt{2\pi}} exp(-\frac{(x-\mu)^2}{2\sigma^2})dx$$

now using the same methods as mentioned in the mean derivation this integral boils down to

$$\Rightarrow \int_{-\infty}^{\infty} x^2 \frac{1}{\sigma\sqrt{2\pi}} exp(-\frac{x^2}{2\sigma^2})dx$$

$$\Rightarrow \sigma\sqrt{2}\int_{-\infty}^{\infty} (\sigma\sqrt{2}x)^2 \frac{1}{\sigma\sqrt{2\pi}} exp(-\frac{(\sigma\sqrt{2}x)^2}{2\sigma^2})dx$$

$$\Rightarrow \sigma^2 \frac{4}{\sqrt{\pi}} \int_{0}^{\infty} x^2 e^{-x^2} dx$$

$$\Rightarrow \sigma^2 \frac{4}{\sqrt{\pi}} \frac{1}{2} \frac{\sqrt{\pi}}{2}$$

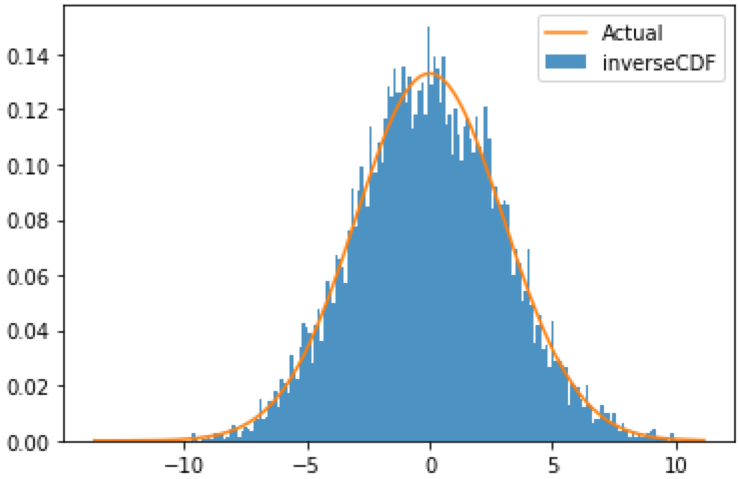
$$\Rightarrow Var(x) = \sigma^2$$

$$Hence\ \ Prooved$$

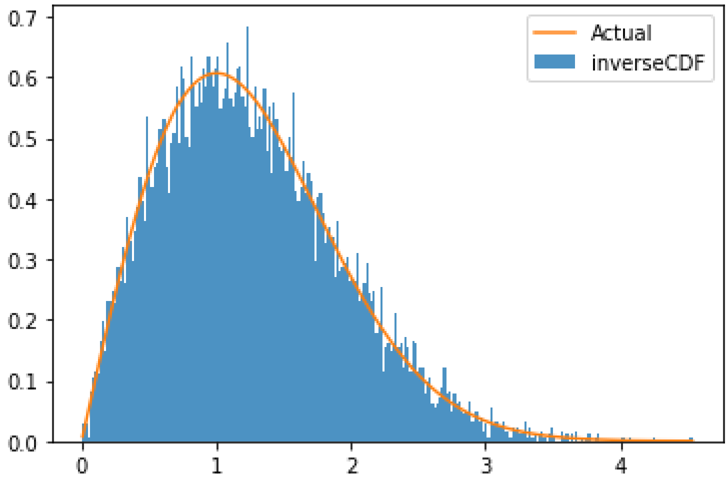
NOTE: I directly make use of gamma function results for finding complex integral values in the above derivation!

Q6)

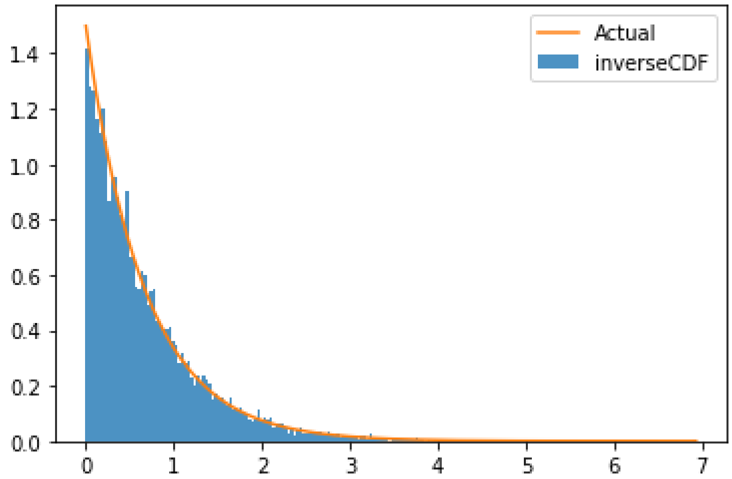
a) Normal density with μ = 0, σ = 3.0.



b) Rayleigh density with σ = 1.0.



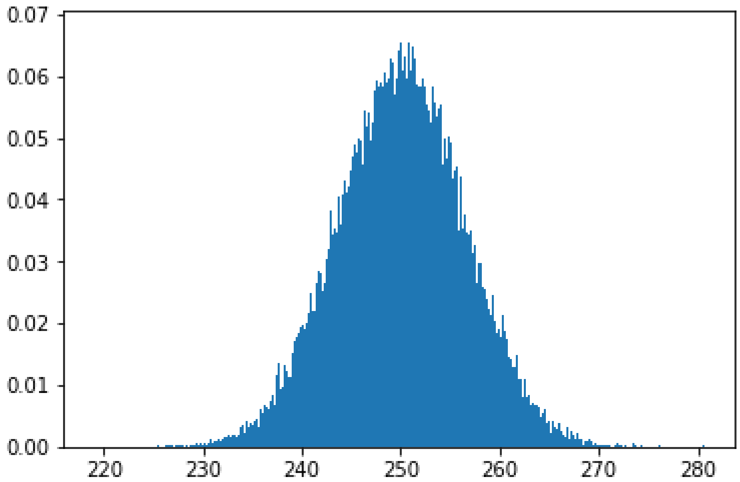
c) Exponential density with λ = 1.5



If we start with a set of Random Variables with in a Uniform density $U(0, 1)$. Now if we need to map these numbers to a particular distribution of random variables with CDF C . We can approximately achieve this by using the inverse function of desired CDF, C^{-1} . These graphs are solid pieces of evidence for proving the approximation methods discussed in the tutorial.

Q7)

The Required histogram is shown below. The shape of this histogram is very much similar to a Normal Density PDF graph.



Code Snippets

Q3)

```
from matplotlib import pyplot as plt
import random
import numpy as np

def sumOfFiveHundred():
    ans=0
    times=500
    while(times > 0):
        ans = ans + random.uniform(0, 1)
        times = times-1
    return ans

times = 50000
templist =[]
while(times > 0):
    templist.append(sumOfFiveHundred())
    times = times-1

X = np.array(templist)

plt.hist(X,density=True, bins=500, alpha=1)
plt.show()
```

Q6)

a) Normal Distribution

```
import matplotlib.pyplot as plt
from scipy.stats import uniform
from scipy.stats import norm
import numpy as np


def normal(n,mean,stdeviation,bins):
    U=uniform.rvs(size=n)
    X=norm.ppf(U,loc=0,scale=stdeviation)
    X=np.sort(X)
    plt.hist(X,density=True, bins=bins, alpha=0.8, label="inverseCDF")
    plt.plot(X, norm.pdf(X,mean,stdeviation), alpha=1, label="Actual")

    plt.legend()
    plt.show()

normal(n=10000,mean=0,stdeviation=3.0, bins=200)
```

b) Rayleigh Distribution

```
import matplotlib.pyplot as plt
from scipy.stats import uniform
from scipy.stats import rayleigh
import numpy as np


def rayleigh(n=1,mean=1,stdeviation=1, bins = 50):
    U=uniform.rvs(size=n)
    X=rayleigh.ppf(U,loc=mean,scale=stdeviation)
    X=np.sort(X)
    plt.hist(X,density=True, bins=bins, alpha=0.8, label="inverseCDF")
    plt.plot(X, rayleigh.pdf(X,mean,stdeviation), label="Actual")
    plt.legend()
    plt.show()

rayleigh(n=10000,mean=0,stdeviation=1.0, bins=250)
```

c) Exponential Distribution

```
import matplotlib.pyplot as plt
from scipy.stats import expon
from scipy.stats import uniform
import numpy as np


def exp(n,labda,bins):
    labda=(1/labda)
    U=uniform.rvs(size=n)
    X=expon.ppf(U,loc=0,scale=labda)
    X=np.sort(X)
    plt.hist([X],bins,density=True,label="inverseCDF",alpha=0.8)
    plt.plot(X, expon.pdf(X,0,labda), label="Actual")
    plt.legend()
    plt.show()

exp(n=10000,labda=1.5, bins=200)
```

Q7)

```
from matplotlib import pyplot as plt
import random
import numpy as np

def sumOfFiveHundred():
    ans=0
    times=500
    while(times > 0):
        ans = ans + random.uniform(0, 1)
        times = times-1
    return ans

times = 50000
templist =[]
while(times > 0):
    templist.append(sumOfFiveHundred())
    times = times-1

X = np.array(templist)

plt.hist(X,density=True, bins=500)
plt.show()
```