SMAI - Assignment 1

M KRISHNA PRANEET
2020113010

1) (i) PMF with finite range (example)

let 'X' be a random variable denoting the heights of students in a class in cm.

It can take only integer values in the range [160, 170).

It is a uniform distribution.

i.e. $P(v_i) = \frac{1}{10}$ $v_i \in \{160, 161, 162, 163, 164, 165, 166, 167, 168, 169\}$

→ It follows the conditions of a PMF because,

① $\forall v_i \in \{160, 161, \ldots, 169\}$, $P(v_i) \geq 0$ $(= \frac{1}{10})$

② $\sum_{i=1}^{10} P(v_i) = \frac{1}{10} + \frac{1}{10} + \frac{1}{10} + \frac{1}{10} + \frac{1}{10} + \frac{1}{10} + \frac{1}{10} + \frac{1}{10} + \frac{1}{10} + \frac{1}{10} = 1$

∴ it satisfies both conditions of PMF and has a finite range.

(ii) Consider a random variable X which denotes the number of coin tosses required till a 'Head' is obtained.

Let's assume 'p' is the prob. of 'H', (1-p) is the prob. of 'T'.

X can take any ~~non-~~ positive integer value. Thus the range is infinite.

But since ~~the~~ ~~integer~~ $\mathbb{N}$ (natural no.) set is countable, $\Rightarrow$ the range of X is countably infinite.

Range of X = $\{1, 2, 3, \cdots\}$ = $\mathbb{N}$

PMF can be defined as,

$$P(x) = \begin{cases} p(1-p)^{x-1} & x = 1, 2, 3, \cdots \\ 0 & \text{o.w} \end{cases}$$

$(\because P(1) = p \; , \; P(2) = (1-p)p \; , \cdots)$

To show that it satisfies conditions of PMF.

① since x is defined to be a natural number,

and we know that $0 \le p \le 1$

$(\because p$ is probability$)$,

$\Rightarrow 0 \le 1-p \le 1$

$\Rightarrow \quad 0 \le (1-p)^{x-1} \le 1$

$\Rightarrow \quad 0 \le p \cdot (1-p)^{x-1} \le 1$

$\therefore \quad P(x) \ge 0 \quad \forall \; x \in \{1, 2, 3, \ldots\}$

② to find $\sum P(x)$

$= \sum_{x=1}^{\infty} p(1-p)^{x-1}$

$= p + p(1-p) + p(1-p)^2 + \cdots$

(It's a G.P)

$= \dfrac{p}{1-(1-p)}$ (G.P formula)

$= 1$

$\Rightarrow \sum_{\forall x} P(x) = 1$

$\therefore$ it satisfies both conditions of PMF and also has infinetly countable range.

To prove: $\sigma^2 = \frac{(b-a)^2}{12}$ for Uniform Density PDF.

from eq. (f), we know that

$$\sigma^2 = E(x^2) - [E(x)]^2$$

(i) $E(x) = \int_a^b x \cdot P(x) \cdot dx$ $\quad \left[\because U(a,b) = \begin{cases} \frac{1}{b-a} & a \le x \le b \\ 0 & o.w. \end{cases} \right]$

$$= \int_a^b x \cdot \left(\frac{1}{b-a}\right) \cdot dx$$

$$= \left(\frac{1}{b-a}\right) \int_a^b x\,dx$$

$$= \frac{1}{b-a} \left[x^2/2 \right]_a^b$$

$$E(x) = \frac{b+a}{2} \quad \text{—— (1)}$$

(ii) $E(x^2) = \int_a^b x^2 \cdot P(x) \cdot dx$

$$= \int_a^b x^2 \cdot \left(\frac{1}{b-a}\right) \cdot dx$$

$$= \frac{1}{b-a} \int_a^b x^2\,dx$$

$$= \frac{1}{b-a} \left[x^3/3 \right]_a^b$$

$$= \frac{1}{b-a} \times \frac{b^3 - a^3}{3}$$

$$= \frac{1}{\cancel{b-a}} \times \frac{\cancel{(b-a)}(b^2+ba+a^2)}{3} \quad (\text{if } a \neq b)$$

$$E(x^2) = \frac{b^2+ba+a^2}{3} \quad \text{—— } ②$$

from ① & ②.

$$\sigma^2 = \frac{b^2+ba+a^2}{3} - \left(\frac{b+a}{2}\right)^2$$

$$= \frac{b^2+ba+a^2}{3} - \frac{b^2+2ab+a^2}{4}$$

$$= \frac{b^2-2ab+a^2}{12}$$

$$\boxed{\sigma^2 = \frac{(b-a)^2}{12}} \quad \text{QED}$$

---

④ By definition,

$$\sigma^2 = E\left[(x-\mu)^2\right]$$

(we know that in DRV, $E(x) = \sum_{\forall x_i} x_i \cdot P(x_i)$.)

$$\sigma^2 = E\left[x^2 - 2x\mu + \mu^2\right]$$

$$= \sum_{\forall x_i} (x_i^2 - 2x_i\mu + \mu^2)\, P(x_i) \quad (\because \mu \text{ is a constant})$$

$$= \sum_{\forall x_i} x_i^2 \cdot P(x_i) - 2\mu \sum_{\forall x_i} x_i \cdot P(x_i) + \mu^2 \sum_{\forall x_i} P(x_i)$$

$$= E[x^2] - 2\mu(E[x]) + \mu^2$$

$\left(\because \sum_{\forall x_i} P(x_i) = 1 \text{ by definition} \right)$

also, replace $\mu = E[x]$

$$\sigma^2 = E[x^2] - 2(E[x])^2 + (E[x])^2$$

$$\boxed{\sigma^2 = E[x^2] - (E[x])^2}$$

---

(5) $N(\mu, \sigma^2) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$

(i) find Expectation:

$$E = \int_{-\infty}^{\infty} x \cdot f(x) \cdot dx.$$

$$= \int_{-\infty}^{\infty} \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} \cdot x \cdot dx$$

$$= \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} x \cdot e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} \cdot dx$$

Subs. $t = \frac{x-\mu}{\sqrt{2}\sigma}$

$dt = \frac{1}{\sqrt{2}\sigma} dx$

$$= \frac{1}{\sqrt{\pi}} \int_{-\infty}^{\infty} (\sqrt{2}\sigma t + \mu) \cdot e^{-t^2} \cdot dt$$

$$= \frac{1}{\sqrt{\pi}}\left[\sqrt{2}\sigma\int_{-\infty}^{\infty} t\cdot e^{-t^2} dt + \mu\int_{-\infty}^{\infty} e^{-t^2} dt\right]$$

~~put u=~~

$$= \frac{1}{\sqrt{\pi}}\left[-\frac{\sqrt{2}}{2}\sigma\int_{-\infty}^{\infty} -2t\, e^{-t^2} dt + \mu\int_{-\infty}^{\infty} e^{-t^2} dt\right]$$

$$= \frac{1}{\sqrt{\pi}}\left[\underbrace{-\frac{\sigma}{\sqrt{2}}\left[e^{-t^2}\right]_{-\infty}^{\infty}}_{0} + \mu\int_{-\infty}^{\infty} e^{-t^2} dt\right]$$

$$= \frac{1}{\sqrt{\pi}}\left[\mu\underbrace{\int_{-\infty}^{\infty} e^{-t^2} dt}_{\text{Gaussian integral } (=\sqrt{\pi})}\right]$$

$$= \frac{1}{\cancel{\sqrt{\pi}}} \times \mu \times \cancel{\sqrt{\pi}}$$

$$\underline{E = \mu}$$

$\therefore$ the expectation of the Gaussian function is μ.

(ii) Variance.

$$\text{Variance } V = E[x^2] - (E[x])^2$$

$$V = \int_{-\infty}^{\infty} x^2 \cdot f(x)\, dx - \mu^2$$

already proved that μ is $E(x)$.

$$= \int_{-\infty}^{\infty} x^2 \left(\frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} \right) dx - \mu^2$$

$$= \frac{1}{\sigma\sqrt{2\pi}} \int_{-\infty}^{\infty} x^2 e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} dx - \mu^2$$

$\left(\text{put } t = \dfrac{x-\mu}{\sqrt{2}\,\sigma}\right)$ $\Rightarrow \sqrt{2}\,\sigma\, dt = dx.$

$$= \frac{\cancel{\sqrt{2}}\,\cancel{\sigma}}{\cancel{\sigma}\,\cancel{\sqrt{2}}\sqrt{\pi}} \int_{-\infty}^{\infty} (\sqrt{2}\,\sigma t + \mu)^2 e^{-t^2} dt - \mu^2$$

$$= \frac{1}{\sqrt{\pi}} \left(2\sigma^2 \int_{-\infty}^{\infty} t^2 e^{-t^2} dt + 2\sqrt{2}\,\sigma\mu \left(\int_{-\infty}^{\infty} t e^{-t^2} dt \right) + \mu^2 \underbrace{\int_{-\infty}^{\infty} e^{-t^2} dt}_{\text{Gaussian integral } = \sqrt{\pi}} \right) - \mu^2$$

same integral in (i) $= 0$

$$= \frac{1}{\sqrt{\pi}} \left(2\sigma^2 \int_{-\infty}^{\infty} t^2 e^{-t^2} dt + 2\sqrt{2}\,\sigma\mu(0) + \mu^2\sqrt{\pi} \right) - \mu^2$$

$$= \frac{1}{\sqrt{\pi}}\left(2\sigma^2\int_{-\infty}^{\infty} t^2 e^{-t^2}\,dt\right) + \cancel{\mu^2} - \cancel{\mu^2}$$

$$= \frac{2\sigma^2}{\sqrt{\pi}} \int_{-\infty}^{\infty} t^2 e^{-t^2}\,dt$$

using integration by parts,

$$= \frac{2\sigma^2}{\sqrt{\pi}}\left(\underbrace{\left(\frac{-t \cdot e^{-t^2}}{2}\right)_{-\infty}^{\infty}}_{=0} + \frac{1}{2}\underbrace{\int_{-\infty}^{\infty} e^{-t^2}\,dt}_{=\sqrt{\pi}}\right)$$

$$= \frac{\cancel{2}\sigma^2}{\cancel{\sqrt{\pi}}} \times \frac{1}{\cancel{2}} \times \cancel{\sqrt{\pi}}$$

$$V = \sigma^2$$

$\therefore$ Variance $= \sigma^2$

$\therefore$ In $N(\mu, \sigma^2)$, the μ & σ^2 indeed represent mean and variance of the distribution.

SMAI Assignment 1 Coding Part

Q3

(3) The two plots I've chosen are
Normal Distribution & Uniform distribution.

→ Uniform Distribution:

$a = -1, \quad b = 11$ $\quad \left(\because \sigma^2 = \frac{(b-a)^2}{12}, \ \mu = \frac{a+b}{2} \right)$

$\Rightarrow \mu = 6, \quad \sigma^2 = 12$

This has been plotten in orange line.

→ Normal distribution

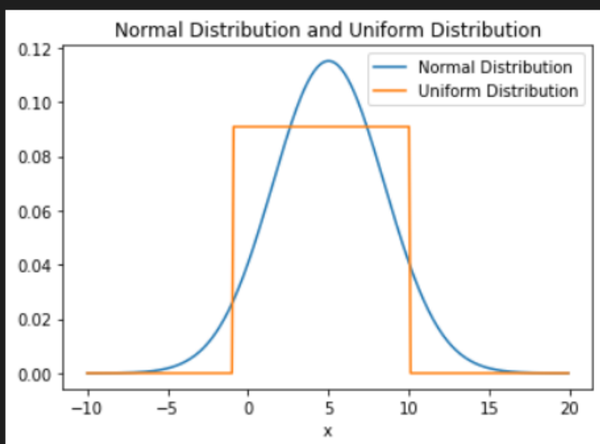
As we know, the parameters for the normal distribution are μ & σ^2.
I've given the same

$\mu = 6, \quad \sigma^2 = 12$ as parameters.

We can clearly see that the two distributions are different.

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

x_axis = np.arange(-10, 20, 0.1)

mean = 5
std = np.sqrt(12)

plt.plot(x_axis, norm.pdf(x_axis, mean, std), label='Normal Distribution')
plt.xlabel('x')
# plt.ylabel('p(x)')
plt.title('Normal Distribution and Uniform Distribution')
plt.plot(x_axis, uniform.pdf(x_axis, -1, 11), label='Uniform Distribution')
plt.legend()
plt.show()
```

[71] 0.3s



Q6

(6) ~~y = c(t)~~ y follows a PDF, $U[0,1]$,

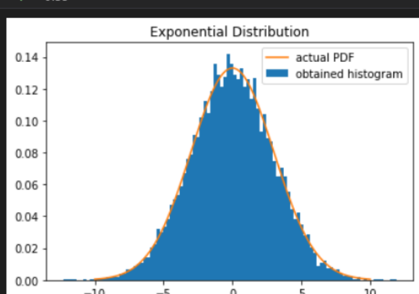
$x = C^{-1}(y)$ has a PDF whose CDF is $C()$.

The plots show that the normalised histograms obtained of x (got by C^{-1} on $U[0,1]$) indicate the PDF of x. This is confirmed by the coinciding of the theoretical PDF of the distributions (orange line) with the histograms (blue bars).

Normal Density

```python
''' normal distribution '''
from scipy import special

y = np.zeros(10000)
y = np.random.uniform(0, 1, 10000)

mu_val = 0.0
sigma_val = 3.0

x = (mu_val + sigma_val * np.sqrt(2) * special.erfinv(2*y - 1))

plt.hist(x, bins=100, density=True, label='obtained histogram')

axis = np.arange(-10, 10, 0.001)

plt.plot(axis, (1/(sigma_val*np.sqrt(2*np.pi))) * ( np.exp( (-1/2) * ((axis-mu_val)/sigma_val)**2 ) ) ,  label='actual PDF')
plt.legend()
plt.title('Exponential Distribution')

plt.show()
```

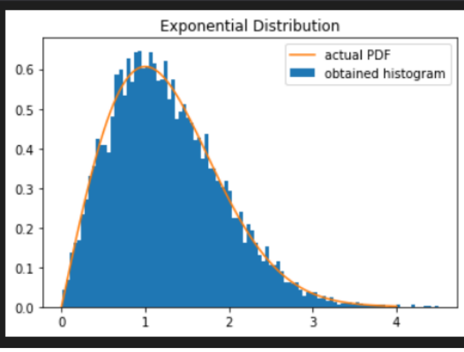
✓ 0.3s



Rayleigh Density

```
''' rayleigh distribution '''

y = np.zeros(10000)
y = np.random.uniform(0, 1, 10000)


sigma_val = 1
x = (np.sqrt(-2*(np.log(1 - y))) * (sigma_val))

plt.hist(x, bins=100, density=True, label='obtained histogram')

axis = np.arange(0, 4, 0.001)

plt.plot(axis, (axis/(sigma_val*sigma_val)*(np.exp((-axis*axis)/(2*sigma_val*sigma_val)))),  label='actual PDF')
plt.legend()
plt.title('Exponential Distribution')

plt.show()
```

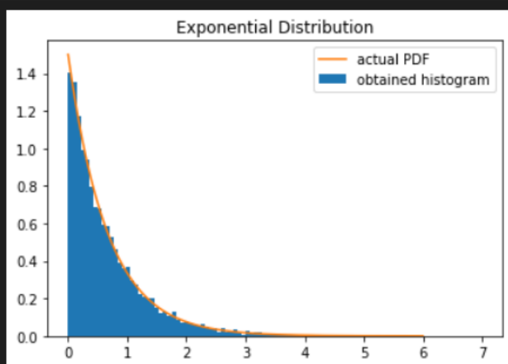
0.4s



Exponential Density

```python
''' exponential distribution '''
y = np.zeros(10000)
y = np.random.uniform(0, 1, 10000)

lambda_val = 1.5
x = np.log(1 - y) / (-lambda_val)

plt.hist(x, bins=100, density=True, label='obtained histogram')

axis = np.arange(0, 6, 0.001)

plt.plot(axis, lambda_val * np.exp(-lambda_val * axis),  label='actual PDF')
plt.legend()
plt.title('Exponential Distribution')

plt.show()
```

✓ 0.4s



Q7

(7) The shape of the resulting histogram seems to be that of a normal distribution with μ around 250.

```python
def gen500():
    return np.sum(np.random.uniform(0, 1, 500))


arr = np.zeros(50000)

for i in range(50000):
    arr[i] = gen500()
# print(arr)
plt.hist(arr, bins=500, range=(0, 500), density=True)
plt.xlabel('generated random number')
plt.xlim(200, 300)
plt.show()
```

✓ 1.4s