Statistical Methods in AI
Assignment 1

Megha Bose
2019111021

---

1. Give an example each of probability mass functions with finite and infinite ranges. Show that the conditions on PMF are satisfied by your example.

Ans: Probability mass function with finite range $\{0, \frac{1}{n}\}$:

$$P_X(x) = \begin{cases} \frac{1}{n} & \text{if } x = 0, 1, ... n-1 \\ 0 & \text{otherwise} \end{cases} \tag{1}$$

where $n \in \mathbb{N}$ and $n > 1$

Checking the conditions taking $P(v_i)$ as the probability that random variable $x \in Z$ takes value v_i,

$$\forall i, P(v_i) \geq 0 \text{ as } n > 0 \implies \frac{1}{n} > 0 \tag{2}$$

$$\sum_{i=1}^{n} P(v_i) = \sum_{i=1}^{n} p_i = n * \frac{1}{n} + 0 = 1 \tag{3}$$

Probability mass function with infinite range $[0, 1)$:

$$P_X(x) = \begin{cases} p(1-p)^{x-1} & \text{if } x \in \mathbb{N} \\ 0 & \text{otherwise} \end{cases} \tag{4}$$

where $p \in (0, 1)$

Checking the conditions taking $P(v_i)$ as the probability that random variable x takes value v_i,

$$\forall i, P(v_i) \geq 0 \text{ as } 0 < p < 1 \implies (1-p)^{x-1} > 0 \tag{5}$$

$$\sum_{i=1}^{n} P(v_i) = \sum_{i=1}^{n} p_i \tag{6}$$

$$= \sum_{x=1}^{\infty} p(1-p)^{x-1} = p \sum_{y=0}^{\infty} (1-p)^y = p \frac{1}{1-(1-p)} + 0 = p * \frac{1}{p} = 1 \tag{7}$$

2. Show with complete steps that the variance of uniform density is given by equation 10. (Hint: use the expression for variance in equation 5.)

Ans: For uniform distribution,

Uniform Distribution :

$$U(a,b) = \begin{cases} \frac{1}{(b-a)} & \text{if } a \le x \le b \\ 0 & \text{otherwise} \end{cases}$$

To show, $\sigma^2 = (b-a)^2/12$.

We know,

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

$$= \int_a^b x^2 P(x)\, dx - \left(\int_a^b x P(x)\, dx \right)^2$$

$$= \frac{1}{b-a} \frac{x^3}{3} \Big|_a^b - \frac{1}{(b-a)^2} \left(\frac{x^2}{2} \Big|_a^b \right)^2$$

$$= \frac{1}{b-a} \cdot \frac{b^3 - a^3}{3} - \frac{1}{(b-a)^2} \frac{(b^2-a^2)^2}{4}$$

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

$$= \frac{4b^2 + 4ab + 4a^2 - 3b^2 - 3a^2 - 6ab}{12}$$

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

(Proved)

3. Show examples of two density functions (draw the function plots) that have the same mean and variance, but clearly different distributions. Plot both functions in the same graph with different colours.

Ans: Two different distributions with same mean and variance can be normal distribution with mean 0 and variance 1 ($X \sim N(0,1)$), and uniform distribution defined as:

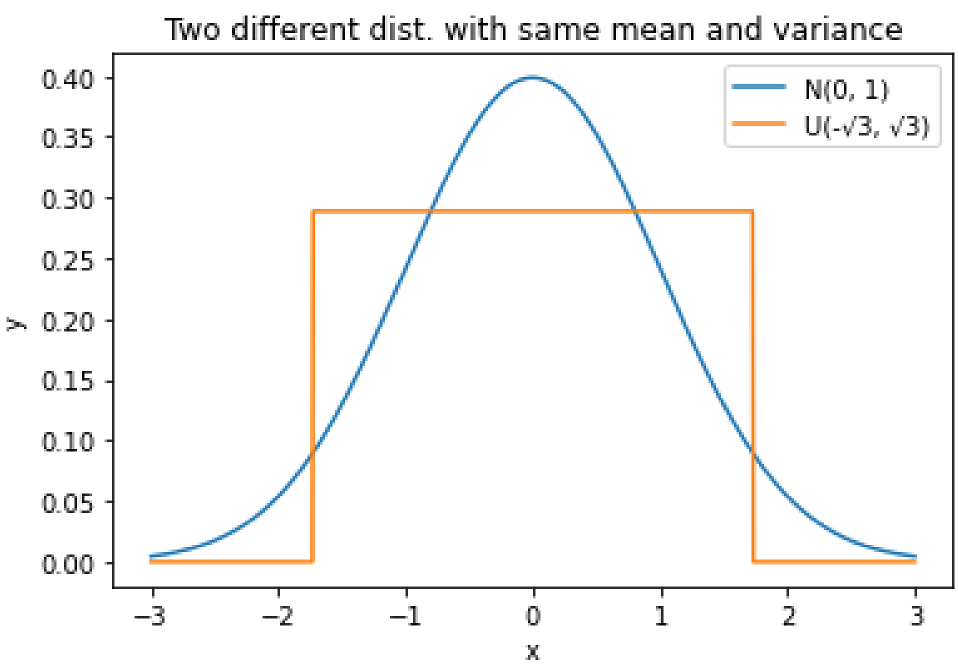
$$U(a,b) = U(-\sqrt{3}, \sqrt{3}) = \begin{cases} \frac{1}{2\sqrt{3}} & \text{if } -\sqrt{3} \le x \le \sqrt{3} \\ 0 & \text{otherwise} \end{cases} \tag{8}$$

The mean of this uniform distribution is

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

and variance is

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



4. Show that the alternate expression for variance given in equation 5 holds for discrete random variables as well.

Ans: For discrete random variables,

To show:–

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

for discrete random variables.

For discrete RV,

$$\mu = E[x] = \sum_{x \in X} x P(x) = \sum_{i=1}^{n} v_i \cdot P(v_i)$$

$$\sigma^2 = E[(x-\mu)^2] = \sum_{i=1}^{n} (v_i - \mu)^2 P(v_i)$$

$$\therefore \sigma^2 = \sum_{i=1}^{n} (v_i^2 - 2\mu v_i + \mu^2) P(v_i)$$

$$= \sum_{i=1}^{n} v_i^2 P(v_i) - 2\mu \sum_{i=1}^{n} v_i P(v_i) + \mu^2 \sum_{i=1}^{n} P(v_i)$$

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

$$\left[\text{since, } \sum_{i=1}^{n} v_i^2 P(v_i) = E[x^2], \ \sum_{i=1}^{n} P(v_i) = 1 \text{ and } \sum_{i=1}^{n} v_i P(v_i) = \mu\right]$$

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

(Proved) $= \sigma^2$

5. Prove that the mean and variance of a normal density, $N(\mu, \sigma^2)$ are indeed its parameters, μ and σ^2.

Ans: For normal distribution $N(\mu, \sigma^2)$,

__/__/____

We know,

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

$\therefore E[x]$

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

Substituting $t = x-\mu$,

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

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

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

Taking the first integral,

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

$$= \int_{-\infty}^{0} \frac{t}{\sigma\sqrt{2\pi}} e^{\frac{-t^2}{2\sigma^2}} dt$$

$$+ \int_{0}^{\infty} \frac{t}{\sigma\sqrt{2\pi}} e^{\frac{-t^2}{2\sigma^2}} dt$$

$$= -\int_{0}^{-\infty} \frac{t}{\sigma\sqrt{2\pi}} e^{\frac{-t^2}{2\sigma^2}} dt + \int_{0}^{\infty} \frac{t}{\sigma\sqrt{2\pi}} e^{\frac{-t^2}{2\sigma^2}} dt$$

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

(Taking $-t$ as t)

$$= -\underbrace{\int_{0}^{\infty} \frac{t}{\sigma\sqrt{2\pi}} e^{\frac{-t^2}{2\sigma^2}} dt} + \underbrace{\int_{0}^{\infty} \frac{t}{\sigma\sqrt{2\pi}} e^{\frac{-t^2}{2\sigma^2}} dt}$$

$$= 0$$

__/__/____

$\therefore$

$$E[x] = 0 + \int_{-\infty}^{\infty} \mu \frac{1}{\sigma\sqrt{2\pi}} e^{\frac{-t^2}{2\sigma^2}} dt$$

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

$$= \frac{2\mu}{\sigma\sqrt{2\pi}} \int_{0}^{\infty} e^{-t^2/2\sigma^2} dt$$

(as integrand is even function)

$$= \frac{2\mu\sqrt{2\sigma^2}}{\sigma\sqrt{2\pi}} \int_{0}^{\infty} e^{-z^2} dz$$

$\left(\text{Taking } z = \frac{t}{\sqrt{2\sigma^2}}\right)$

$$= \frac{2\mu}{\sqrt{\pi}} \int_{0}^{\infty} e^{-z^2} dz$$

Now, we need to calculate

$$\int_{0}^{\infty} e^{-z^2} dz$$

__/__/__

Let

$$I = \int_{-\infty}^{\infty} e^{-z^2} dz = \int_{-\infty}^{\infty} e^{-y^2} dy$$

$$\therefore I^2 = \int_{-\infty}^{\infty} \int_{-\infty}^{\infty} e^{-(x^2+y^2)} dx dy$$

Changing to polar coordinates,

$$I^2 = \int_0^{2\pi} \int_0^{\infty} e^{-r^2} r dr d\theta$$

Taking $p = r^2$,

$$I^2 = 2\pi \int_0^{\infty} e^{-p} dp/2$$

$$= 2\pi \cdot \frac{1}{2} = \pi$$

$$\therefore I = \sqrt{\pi}$$

$$\therefore \int_0^{\infty} e^{-z^2} dz = \frac{1}{2} I = \sqrt{\pi}/2$$

__/__/____

$$\therefore E[x] = \frac{2\mu}{\sqrt{\pi}} \cdot \frac{\sqrt{\pi}}{2}$$

$$= \mu$$

$$\therefore \boxed{E[x] = \mu} \quad \text{(Proved)}$$

Now, for variance,

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

Taking $x - \mu = \sigma t$,

$$= \int_{-\infty}^{\infty} t^2 \frac{1}{\sigma\sqrt{2\pi}} e^{-t^2/2\sigma^2} dt$$

Taking $t = \sigma\sqrt{2}\, z$,

$$= \sigma\sqrt{2} \int_{-\infty}^{\infty} (\sigma\sqrt{2}\, z)^2 \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(\sigma\sqrt{2} z)^2}{2\sigma^2}} dz$$

$$= \frac{4\sigma^2}{\sqrt{\pi}} \int_0^\infty z^2 e^{-z^2} dz$$

Taking $z^2 = y$, $\Rightarrow 2z\,dz = dy$

$$= \frac{4\sigma^2}{\sqrt{\pi}} \int_0^\infty z^2 e^{-z^2} dz$$

$$= \frac{4\sigma^2}{\sqrt{\pi}} \int_0^\infty y\, e^{-y} \frac{dy}{2\sqrt{y}}$$

$$= \frac{4\sigma^2}{\sqrt{\pi}} \underbrace{\int_0^\infty y^{3/2-1} e^{-y} dy}_{\text{gamma function}}$$

$$= \frac{4\sigma^2}{\sqrt{\pi}} \cdot \Gamma\left(\frac{3}{2}\right)$$

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

$\therefore \boxed{Var(X) = \sigma^2}$ (Proved)

6. Using the inverse of CDFs, map a set of 10, 000 random numbers from $U[0, 1]$ to follow the following pdfs:

   1. Normal density with $\mu = 0$, $\sigma = 3.0$.
   2. Rayleigh density with $\sigma = 1.0$.
   3. Exponential density with $\lambda = 1.5$.

   Once the numbers are generated, plot the normalized histograms (the values in the bins should add up to 1) of the new random numbers with appropriate bin sizes in each case; along with their pdfs. What do you infer from the plots?

   Note: see rand() function in C for $U[0, INTMAX]$.

Ans: Let X be a random variable whose distribution can be described by the cumulative distribution function F_X. We want to generate values of X which are distributed according to this distribution.

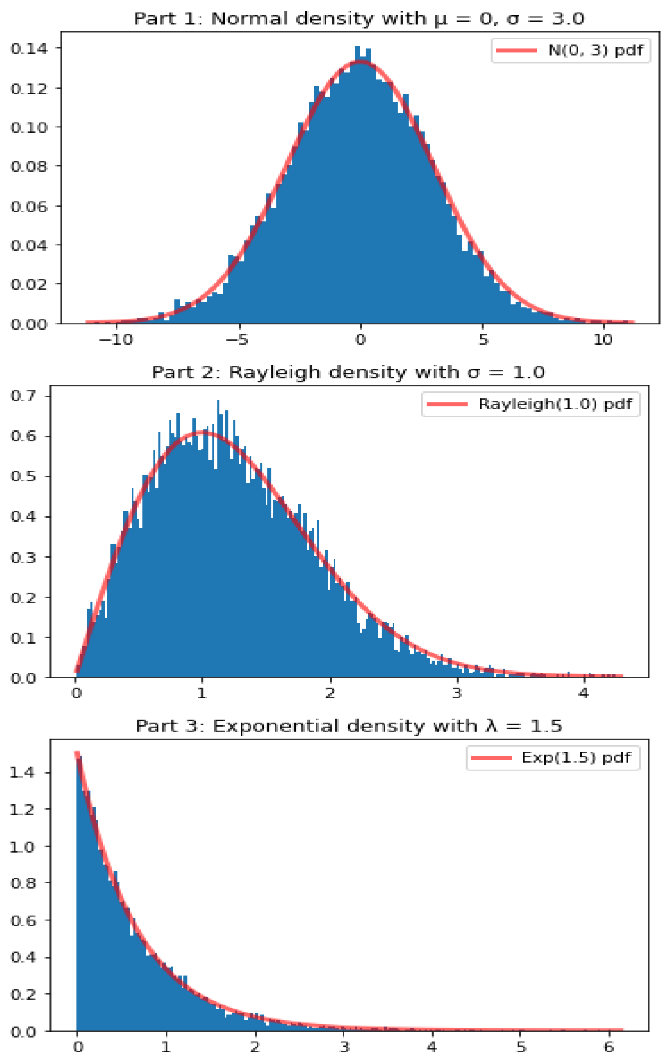
We use the inverse transform sampling method that works as follows:

We generate a random number u from the standard uniform distribution in the interval $[0, 1]$ like $U[0, 1]$.

Then we find the inverse of the desired CDF, $F_X^{-1}(x)$. We compute $X = F_X^{-1}(u)$ and this computed random variable X has distribution $F_X(x)$.

Here we do it using 10000 random numbers from $U[0, 1]$.

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

pts = np.random.rand(10000)

norm_transform = stats.norm.ppf(pts, loc=0, scale=3.0)
norm_actual = stats.norm(loc=0, scale=3.0)
n1 = 100

plt.title("Part 1: Normal density with μ = 0, σ = 3.0")
plt.hist(norm_transform, density=True, bins = n1)
x = np.linspace(norm_actual.ppf(0.0001), norm_actual.ppf(0.9999), 100)
plt.plot(x, norm_actual.pdf(x), 'r-', lw=3, alpha=0.6, label='N(0, 3) pdf')
plt.legend()
plt.show()


ray_result = stats.rayleigh.ppf(pts, loc=0, scale=1.0)
ray_actual = stats.rayleigh(loc=0, scale=1.0)
n2 = 200

plt.title("Part 2: Rayleigh density with σ = 1.0")
plt.hist(ray_result, density=True, bins = n2)
x = np.linspace(ray_actual.ppf(0.0001), ray_actual.ppf(0.9999), 100)
plt.plot(x, ray_actual.pdf(x), 'r-', lw=3, alpha=0.6, label='Rayleigh(1.0) pdf')
plt.legend()
plt.show()


lmda = 1.5
std = 1/lmda
expo_result = stats.expon.ppf(pts, loc=0, scale=std)
expo_actual = stats.expon(loc=0, scale=std)
n3 = 200

plt.title("Part 3: Exponential density with λ = 1.5")
plt.hist(expo_result, density=True, bins = n3)
x = np.linspace(expo_actual.ppf(0.0001), expo_actual.ppf(0.9999), 100)
plt.plot(x, expo_actual.pdf(x), 'r-', lw=3, alpha=0.6, label='Exp(1.5) pdf')
plt.legend()
plt.show()
```



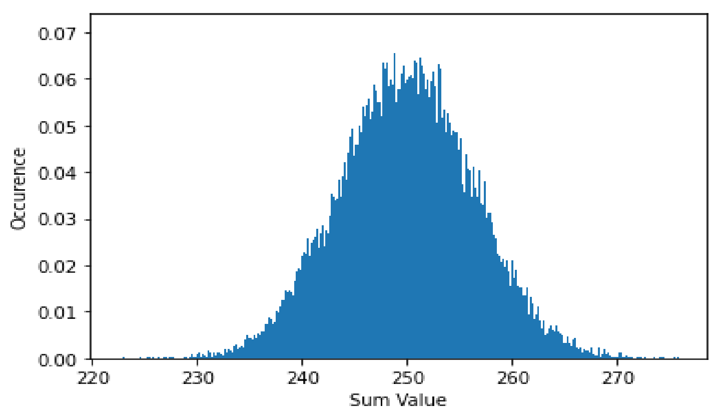
We infer from these plots that the random variable X computed from the values generated by inverse cdf of the given distributions with with random values from uniform distribution as input, follows the pdf of that distribution. This is because, given a random value p, the cdf(x) gives us the probability that p is less than or equal to x, where x follows the distribution. Thus, p is given by the inverse cdf value of x. Hence, when we map uniformly random p values using inverse pdf function, it gives us x values which follows the given distribution.

A suitable bin size while plotting helps us visualise the distribution better. If the bin size is too small, then the blocks are too wide and do not capture the distribution well while if the bins are too many in number, they make the graph too jagged.

7. Write a function to generate a random number as follows: Every time the function is called, it generates 500 new random numbers from $U[0,1]$ and outputs their sum. Generate $50,000$ random numbers by repeatedly calling the above function, and plot their normalized histogram (with bin-size = 1). What do you find about the shape of the resulting histogram?

Ans: The function and the required graph go as follows:

```
import matplotlib.pyplot as plt
import numpy as np
import math

def sum500Uni():
  nums = np.random.uniform(0, 1, size=500)
  return sum(nums)

uni_sums = []
total = 0
for _ in range(0, 50000):
  val = sum500Uni()
  uni_sums.append(val)
  total += val

plt.xlabel("Sum Value")
plt.ylabel("Occurence")
plt.hist(uni_sums, density=True, bins = 500)
plt.show()
```



The graph looks like normal distribution with mean at $\frac{500}{2} = 250$. It is actually Irwin-Wall distribution or uniform sum distribution. The normal approximation to this distribution is done with a normal distribution which has the same mean and standard deviation.