Project Report 2
Disturbing Distributions

Varendra Maurya
200010055@iitdh.ac.in

Department of Computer Science, IIT Dharwad

September 24, 2021

1 Problem Statement

We are given with dataset of 50000 samples of a random varilable Z. We also know that $Z = X + 10Y$, where X is a uniform random varilable between -3 and 3, such that

$$Y = \sum_{i=1}^{k} W_i \tag{1}$$

where k can take values $2, 3$ and 4, W_i's are independent and identically distributed (i.i.d.) and belong to one of the following:
• Exponential distribution characterized by its mean $\frac{1}{\lambda}$.
• Rayleigh distribution characterized by σ.
• Half-normal distribution characterized by σ.

2 Working Meachanism

From now onwards, consider $mean\_Z$ as mean of Z.

We know,

$$Z = X + 10Y$$
$$E[Z] = E[X] + 10E[Y]$$

Since, our dataset is quite large we can consider $mean\_Z$ as $E[X]$.
i.e.

$$mean\_Z = E[Z] = E[X] + 10E[Y] \tag{2}$$

where $E[Y]$ is,

$$E[Y] = \sum_{i=1}^{k} E[W_i] \tag{3}$$

Here, each W_i has same distribution as they are identically distributed, therefore each of W_i has same mean and variance. Therefore $E[Y]$ can be written as:

$$E[Y] = kE[W] \tag{4}$$

Now, we also know that X is uniform random varilable between -3 and 3, therefore,

$$E[X] = (-3 + 3)/2 = 0 \tag{5}$$
$$var(X) = (-3 - 3)^2/12 = 3 \tag{6}$$

Using equation 2, 4 and 5 we can say that,

$$mean\_Z = 10kE[W] \tag{7}$$

Now, for variance,

$$var(Z) = var(X + 10Y)$$

$$var(Z) = var(X) + 2covariance(X, 10Y) + 100var(Y)$$

Since X and Y are independent of each other so $2covariance(X, 10Y) = 0$ and from equation 6 var(X) = 3.

Hence,

$$var(Z) = 3 + 100var(Y) \tag{8}$$

For $var(Y)$, since W_i is (i.i.d.) we can conclude that,

$$var(Y) = 100kvar(W) \tag{9}$$

Finally,

$$var(Z) = 3 + 100kvar(W) \tag{10}$$

Now using equation 7, 10 and considering all three cases one by one:

• For exponential destribution

$$Expectation = \frac{1}{\lambda}$$

$$Variance = \frac{1}{\lambda^2}$$

$$k = \frac{(mean\_Z)^2}{var(Z) - 3} \tag{11}$$

$$\lambda = \frac{10(mean\_Z)}{var(Z) - 3} \tag{12}$$

• For Rayleigh distribution

$$Expectation = \sigma\sqrt{\pi/2}$$

$$Variance = \frac{4 - \pi}{2}\sigma^2$$

$$k = \frac{(4 - \pi)(mean\_Z)^2}{4(var(Z) - 3)} \tag{13}$$

$$\sigma = \frac{\sqrt{\pi}(var(Z) - 3)}{5\sqrt{2}(mean\_Z)(4 - \pi)} \tag{14}$$

• For half-normal distribution:

$$Expectation = \sigma\sqrt{2/\pi}$$

$$Variance = \sigma^2(1 - \frac{2}{\pi})$$

$$\sigma = \frac{\sqrt{2\pi}(var(Z) - 3)}{10(mean\_Z)(4 - \pi)} \tag{15}$$

$$k = \frac{(mean\_Z)^2(\pi - 2)}{2(var(Z) - 3)} \tag{16}$$

Using all these equation and implementing in python we will get our final result with least error as:

- $k = 3$
- Distribution: **Exponential destribution**
- $\lambda = 1$