PDF and Statistical Parameters of Continuous random variable

Continuous random variable

Examples of continuous RV are

- Arrival time of phone call
- Arrival time of customer at Petrol pump
- Measuring of room temperature at certain time
- Initial phase of sine wave generator

Discrete random variable has finite sample space. Thus, each event can be described by a probability number. The sample space of continuous random variable is infinite. Thus, it is not possible to associate probability with each sample point as $1/\infty = 0$. Therefore, continuous random variable can better be explained by probability distribution function. An example of PDF of continuous random variable is

There are three main parameters describing continuous RV

- Mean
- Variance
- Distribution

Gaussian distribution function

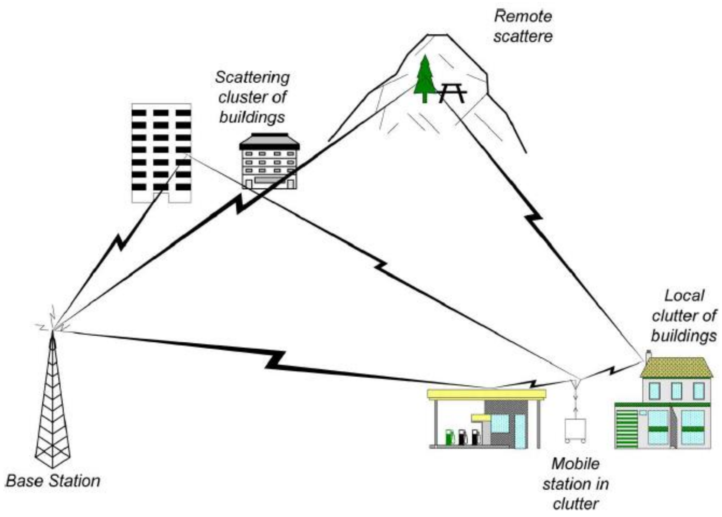

Figure 1: Received Signal modeled as Guassian Distribution

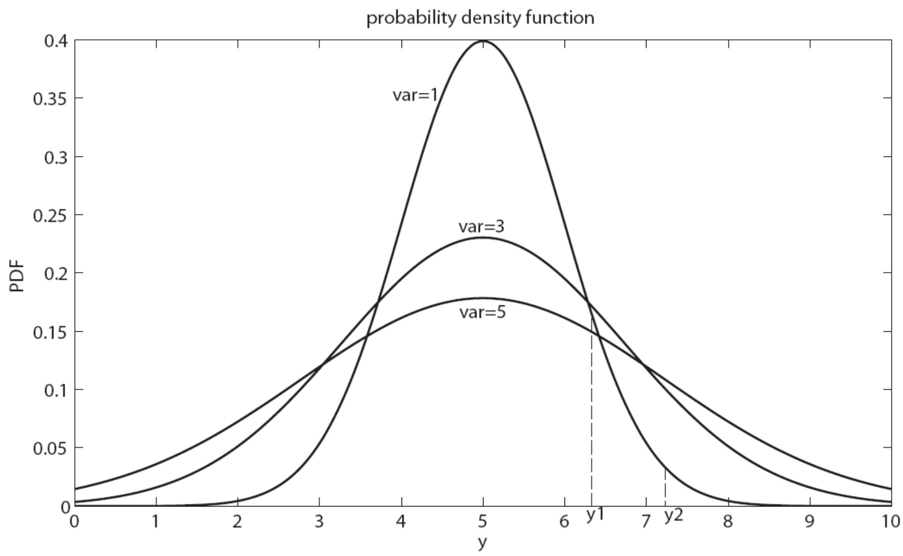

Figure 1.1 Gaussian PDF for mean 'm=5'

$$f_Y(y) = \frac{1}{\sqrt{2\pi\sigma_Y^2}} . \exp\left[- \frac{(\mathrm{y} - \mathrm{m_y})^2}{2\sigma_Y^2} \right] \tag{1}$$

where m and σ_Y^2 are mean and variance of random variable Y respectively defined as:

$$\mathrm{m_y} = E[Y] = \int_{-\infty}^{\infty} y \times f_Y(y) dy \tag{2}$$

$$\sigma_Y^2 = E[(y - m_y)^2] = \int_{-\infty}^{\infty} (y - m_y)^2 \times f_Y(y) dy$$

$$E[(y-m_y)^2] = E[y^2 + m_y^2 - 2ym_y] = E[y^2] + E[m_y^2] - 2E[m_y^2] = E[y^2] - E[m_y^2] \quad (3)$$

For continuous random variable, we are interested in "what is the probability that a random variable Y is lying between certain range, say, y1 to y2. This probability is defined as

$$P(y_1 \le Y \ge y_2) = \int_{y_1}^{y_2} f_X(y)dy$$

We note the following property of probability distribution function.

$$\int_{-\infty}^{\infty} f_X(y)dy = 1$$

Cumulative distribution function (CDF) or simply distribution function is given as

$$F_X(x) = P(X \le x) \quad (4)$$

Central-Limit theorem

This says that if we have n independent and identically distributed random-variable $X_1, X_2, ...X_n$ where n is very large and we define Y such that

$$Y = X_1 + X_2 + X_3 + ... + X_n$$

then, Y will be Gaussian distributed. Note that we have not put any condition on the distribution of $X_1, X_2 ......$. We have only said that whatever distribution they have, all should have same distribution. For example, if one has a square distribution, then, other also should have square distribution. Hence, we can call $X_1, X_2 ......$ as independent and identically distributed (iid). It is found that if distribution is square, then, distribution of summation converges fast towards Gaussian.

This theorem has very wide application. In communication, the noise at the input of the receiver is modeled as additive-white Gaussian noise (AWGN). It is because of the fact that the noise consists of infinite number of small but independent noise sources. Hence, it is the result of cumulative effect of the entire noise source and the resultant noise distribution tends towards Gaussian.

Let X_i, $i = 1,2,...n$ being summed are statistically independent and identically distributed, each having a finite mean m_X and finite variance c σ_X^2 . Let us define U_i such that

$$U_i = \frac{X_i - m_X}{\sigma_X}, \qquad i = 1,2,....n$$

Hence, random variable U_i has the mean value as 0 and variance as 1.

$$Y = \frac{1}{\sqrt{n}}\left[U_1 + U_2 + ... + U_n\right] = \frac{1}{\sqrt{n}}.\left[\sum_{i=1}^{n} U_i\right]$$

So, Y will be Gaussian with zero mean & unit variance.

Example:

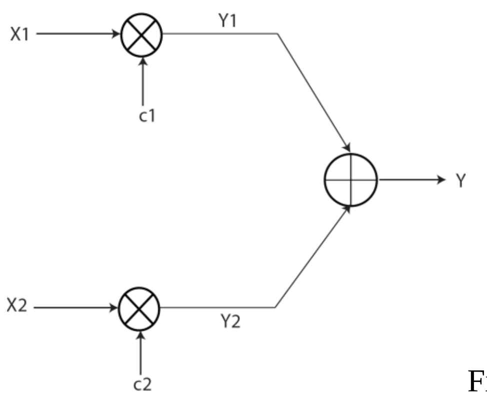

Figure 2

In this problem, X_1 and X_2 are independent and having zero mean and σ^2. Parameters C_1 and C_2 are constant. Determine the mean and variance of Y.

Solution:

$$Y = Y_1 + Y_2 = C_1X_1 + C_2X_2$$

$$\Rightarrow \quad E[Y] = E[C_1X_1 + C_2X_2]$$

$$= C_1E[X_1] + C_2E[X_2]$$

$$= 0.$$

$$\sigma_Y^2 = E[(Y - m_Y)^2]$$

$$= E[(C_1X_1 + C_2X_2 - 0)^2]$$

$$= E[C_1^2X_1^2 + C_2^2X_2^2 + 2C_1C_2X_1X_2]$$

$$= C_1^2E[X_1^2] + C_2^2E[X_2^2] + 2C_1C_2E[X_1X_2] = C_1^2\sigma_{X1}^2 + C_2^2\sigma_{X2}^2 + 2C_1C_2E[X_1]E[X_2]$$

$$= C_1^2\sigma_{X1}^2 + C_2^2\sigma_{X2}^2 = \sigma_{Y_1}^2 + \sigma_{Y_2}^2 \qquad since\ X_1\ and\ X_2\ are\ independent.$$

If C_1 and C_2 are assumed to be one, then, $\sigma_Y^2 = \sigma_{X_1}^2 + \sigma_{X_2}^2$.

Example: Repeat the same above problem if X_1 and X_2 are having mean m_{X_1}, m_{X_2} and varaiance as $\sigma_{X_1}^2, \sigma_{X_2}^2$.

Solution:

$$E[Y] = C_1m_{X_1} + C_2m_{X_2}$$

$$\sigma_Y^2 = E\left[\left(C_1X_1 + C_2X_2 - C_1m_{X_1} - C_2m_{X_2}\right)^2\right]$$

$$= E\left[\left\{C_1(X_1 - m_{X_1}) + C_2(X_2 - m_{X_2})\right\}^2\right]$$

$$= \mathrm{C}_1^2\mathrm{E}\left[\left(X_1 - m_{X_1}\right)^2\right] + \mathrm{C}_2^2\mathrm{E}\left[\left(X_2 - m_{X_2}\right)^2\right] + 2C_1C_2\mathrm{E}\left[\left(X_1 - m_{X_1}\right)\left(X_2 - m_{X_2}\right)\right]$$

$$= C_1^2\sigma_{X_1}^2 + C_2^2\sigma_{X_2}^2 + 2C_1C_2\mathrm{I}$$

Let $$\mathrm{I} = \mathrm{E}\left[\left(X_1 - m_{X_1}\right)\left(X_2 - m_{X_2}\right)\right]$$

$$\mathrm{I} = \mathrm{E}\left[X_1X_2 - m_{X_2}X_1 - m_{X_1}X_2 + m_{X_1}m_{X_2}\right]$$

$$= \mathrm{E}[X_1X_2] - m_{X_2}\mathrm{E}[X_1] - m_{X_1}\mathrm{E}[X_2] + m_{X_1}m_{X_2}$$

$$= m_{X_1}m_{X_2} - m_{X_1}m_{X_2} - m_{X_1}m_{X_2} + m_{X_1}m_{X_2}$$

$= 0$

$\Rightarrow \sigma_Y^2 = C_1^2 \sigma_{X_1}^2 + C_2^2 \sigma_{X_2}^2$

if $C_1 = C_2 = 1$

$$\Rightarrow \sigma_Y^2 = \sigma_{X_1}^2 + \sigma_{X_2}^2$$

Hence in general; if there are n random variable such that $X = X_1 + X_2 \cdots + X_n = \sum_{\mathrm{i=1}}^{\mathrm{n}} \mathrm{X_i}$ and each random variable has mean as m_{X_i} and variance as $\sigma_{X_i}^2$, then mean of X is $m_X = \sum_{\mathrm{i=1}}^{\mathrm{n}} m_{X_i}$ and variance $\sigma_X^2 = \sigma_{X_1}^2 + \sigma_{X_2}^2 + \cdots + \sigma_{X_n}^2 = \sum_{\mathrm{i=1}}^{\mathrm{n}} \sigma_{X_i}^2$. It may be noted that this is true only under the condition that all the random variables are statistically independent.