Continuous Probability Distribution

Introduction

A continuous distribution describes the probabilities of the possible values of a continuous random variable. A continuous random variable is a random variable with a set of possible values (known as the range) that is infinite and uncountable.

Probabilities of continuous random variables (X) are defined as the area under the curve of its PDF. Thus, only ranges of values can have a nonzero probability. The probability that a continuous random variable equals some value is always zero.

Example: Normal Distribution, Exponential Distribution

Normal Distribution

The Normal Distribution is defined by the probability density function for a continuous random variable in a system. Let us say, f(x) is the probability density function and X is the random variable.
$f(x) \geq 0$ for all $x \in (-\infty, \infty)$ and $\int_{-\infty}^{\infty} f(x)dx = 1$

The probability density function of normal or Gaussian distribution is given by;

$$f(x, \mu, \sigma) = \frac{1}{\sigma\sqrt{2\pi}} e^{\frac{-(x-\mu)^2}{2\sigma^2}}$$

Where,
x is the variable
μ is the mean
σ is the standard deviation

Characteristics of Normal Distribution:

- It is symmetric, unimodal (i.e., one mode), and asymptotic
- The values of mean, median, and mode are all equal.
- A normal distribution is quite symmetrical about its center. That means the left side of the center of the peak is a mirror image of the right side. There is also only one peak (i.e., one mode) in a normal distribution.

Procedure:

- Generating the data set
- Determine the probabilities of the random variable using Normal distribution in R
- Visualize the probability distribution using R functions

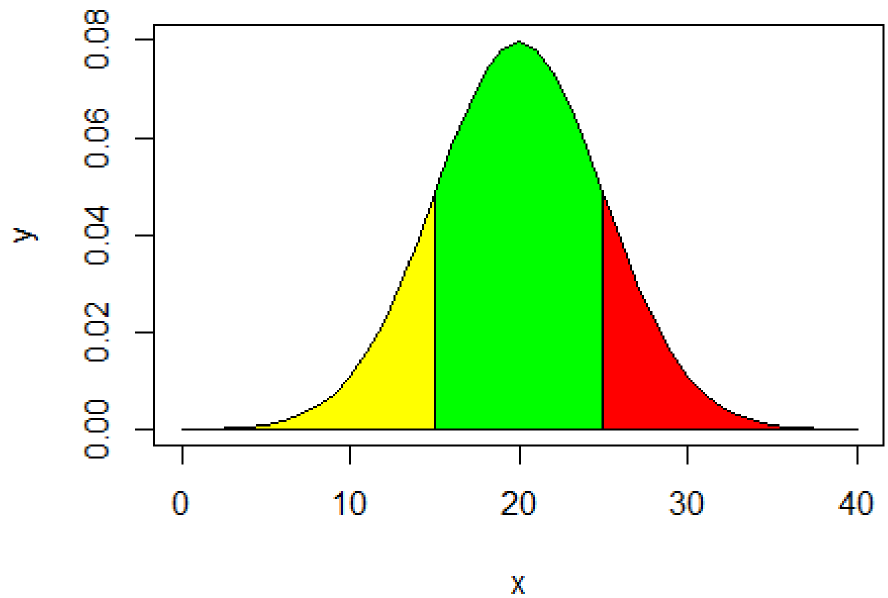
Problem:

A company finds that the time taken by one of its engineers to complete or repair job has a normal distribution with mean 20 minutes and S.D 5 minutes. State what proportion of jobs take:

i. Less than 15 minutes
ii. More than 25 minutes
iii. Between 15 and 25 minutes
iv. Plot the distribution
v. Table the distribution

Code and Results:

```
# Generating the data x
x=seq(0,40)
# find the density function of x
y=dnorm(x,mean=20,sd=5)
# plot the normal distribution curve
plot(x,y,type='l')
# Proportion of jobs take less than 15 minutes
p1=pnorm(15,mean=20,sd=5)
x2=seq(0,15)
y2=dnorm(x2,mean=20,sd=5)
polygon(c(0,x2,15),c(0,y2,0),col='yellow')
#Proportion of jobs take more than 25 minutes
p2=pnorm(40,mean=20,sd=5)-pnorm(25,mean=20,sd=5)
x1=seq(25,40)
y1=dnorm(x1,mean=20,sd=5)
polygon(c(25,x1,40),c(0,y1,0),col='red')
#Proportion of jobs take between 15 and 25 minutes
p3=pnorm(25,mean=20,sd=5)-pnorm(15,mean=20,sd=5)
x3=seq(15,25)
y3=dnorm(x3,mean=20,sd=5)
polygon(c(15,x3,25),c(0,y3,0),col='green')
```




```
# Probability distribution
data.frame(p1,p2,p3)

##          p1        p2        p3
## 1 0.1586553 0.1586236 0.6826895
```

Exponential Distribution

A continuous random variable X is said to follow exponential distribution if its probability density function is given by,

$$f(x) = \begin{cases} \alpha e^{-\alpha x}, & x > 0 \\ 0, & otherwise \end{cases}$$

Procedure:

- Generate the data set
- Determine the probabilities of the random variable using Exponential distribution in R
- Visualize the probability distribution using R functions

Problem:

The mileage which car owners get with a certain kind of radial tire is a RV having an exponential distribution with mean 4,000km. Find the probabilities that one of these tires will last (i) at least 2,000km and (ii) at most 3,000km.

Code and Results:

```r
#Exponential Distribution
# mean value
m=4000
# exponential distribution parameter
alpha=1/m
# probability of tires will last at most 3000 kms
pexp(3000,alpha,lower.tail=T)
```

```
## [1] 0.5276334
```

```r
#probability of tires will last at least 3000 kms
pexp(2000,alpha,lower.tail=F)
```

```
## [1] 0.6065307
```