23MAT204

MATHEMATICS FOR INTELLIGENT SYSTEMS-3

Dr. Sarada Jayan
9448455760
6364333760

2 coins:-

$\{\underset{x=0}{HH}, \underset{x=1}{HT}, \underset{x=1}{TH}, \underset{x=2}{TT}\}$

$X \rightarrow$ R.V that takes no. of tails in each outcome.

X	0	1	2
$f(x)$	$1/4$	$1/2$	$1/4$

Random Variables

Defn:- A Numerical value assigned for every possible outcome of an experiment.

1 2 3 4 5

Discrete R.V

$X = 1, 2,$

$X = 0.5, 3, -1, 4$

$X = a$

Continuous R.V.

$0 \leq X \leq 1$

$a \leq X \leq b$

a b

Discrete Random variables

$S = \{1, 2, 3, 4, 5, 6\}$ $f(x) = P(X = x) = \frac{1}{6}, x = 1, 2, .. 6$
$X = \{1, 2, 3, 4, 5, 6\}$

1. The outcome while throwing a die
2. Number of heads while tossing 'n' number of coins
3. The number of customers arriving in a book store
4. The position of the participants in a running race.

3 coins are tossed, X = # heads.

$$\left.\begin{array}{c|cccc} X & \overset{TTT}{0} & \overset{HTT,\,THT,\,TTH}{1} & \overset{HHT,\,HTH,\,THH}{2} & \overset{HHH}{3} \\ \hline f(x) & 1/8 & 3/8 & 3/8 & 1/8 \end{array}\right\} \text{pmf.}$$

p.m.f, $f(x) = P(X = x)$

<u>Probability mass fn:-</u> probability distrn. of a discrete random variable is

<u>characteristics of pmf:-</u>
1. $0 \le f(x) \le 1$
2. $\sum_{x_i} f(x_i) = 1$

AMRITA VISHWA VIDYAPEETHAM UNIVERSITY Established u/s 3 of UGC Act 1956

Continuous Random variables

1. Heights of randomly selected Person from a population say with age greater than 16.
2. Weights of randomly selected Person from a population say with Age greater than 16.
3. The time one has to weight in a queue(railway ticket queue) over a period of time
4. The service time in a public service system for different customers
5. The life time distribution of a particular brand of electric bulbs

Probability distrn. of conts. r.v.s is called Probability Density function (pdf)

$P(a \le x \le b) = \int_a^b f(x)\,dx$.

Characteristics of pdf:-

① $f(x) \geq 0$

② $\int_x f(x)\,dx = 1$

For conts. R.V. X

$P(X=a) = 0$

$P(a \le X \le b)$
$= P(a < X < b)$
$= P(a \le X < b)$
$= P(a < X \le b)$

Random Variables

Check whether the following can be pmf

Discrete

① $f(x) = \frac{x-2}{2}$, for $x = 1, 2, 3, 4$

x	1	2	3	4
f(x)	$-1/2$	0	$1/2$	1

$\Rightarrow$ f(x) is not a pmf as $f(1) < 0$

② $g(x) = \frac{x^2}{30}$, $x = 0, 1, 2, 3, 4$

x	0	1	2	3	4
g(x)	0	$\frac{1}{30}$	$\frac{4}{30}$	$\frac{9}{30}$	$\frac{16}{30}$

$\Rightarrow$ Here $0 \le g(x) \le 1 \ \forall x$

$\sum g(x) = 1$

$\therefore$ g(x) is a pmf.

③ $f(x) = \frac{x^2}{25}$, $x = 0, 1, 2, 3, 4$

x	0	1	2	3	4
f(x)	0	$\frac{1}{25}$	$\frac{4}{25}$	$\frac{9}{25}$	$\frac{16}{25}$

$\sum f(x) \ne 1$

$\Rightarrow$ f(x) is not a pmf

④ $f(1) = 0.15$; $f(-1) = 0.28$; $f(0) = 0.29$; $f(3) = 0.28$

R.V can be negative

$0 \le f(x) \le 1$ ✓

& $\sum_x f(x) = 1$

Check whether the following can be pdf.

① $f(x) = 2e^{-2x}$, $x \ge 0$

$f(x) \ge 0 \ \forall x$

$\int_x f(x)\,dx = 1$

$f(x) \ge 0$

$$\int_0^\infty f(x)\,dx = \int_0^\infty 2e^{-2x}\,dx = 2\frac{e^{-2x}}{-2}\Big)_0^\infty = -1(0-1) = 1$$

$\therefore$ f(x) is a pdf

② $f(x) = \frac{1}{4}$, $-2 \le x \le 0$

$f(x) \ge 0 \ \forall x$

$$\int_{-2}^0 \frac{1}{4}\,dx = \frac{1}{4}x\Big)_{-2}^0 = \frac{2}{4} = \frac{1}{2} \ne 1$$

$\Rightarrow$ f(x) is not a pdf.

③ $f(x) = \frac{1}{4}$, $-4 \le x \le 0$

$f(x) \ge 0 \ \forall x$

$$\int_{-4}^0 f(x)\,dx = \int_{-4}^0 \frac{1}{4}\,dx = \frac{1}{4}x\Big)_{-4}^0 = 1$$

$\Rightarrow$ f(x) is a pdf

④ $f(x) = \frac{x}{4}$, $-1 \le x \le 0$

f(x) is -ve for some values of x

$\Rightarrow$ f(x) is not a pdf.

⑤ $f(x) = 3e^{-2x}$, $x \ge 0$

$f(x) \ge 0 \ \forall x$.

$$\int_0^\infty 3e^{-2x}\,dx = \frac{3e^{-2x}}{-2}\Big)_0^\infty = \frac{-3}{2}(0-1) = 1.5 \ne 1$$

$\Rightarrow$ f(x) is not a pdf.

1. A random variable X has the following probability distribution. (pmf)

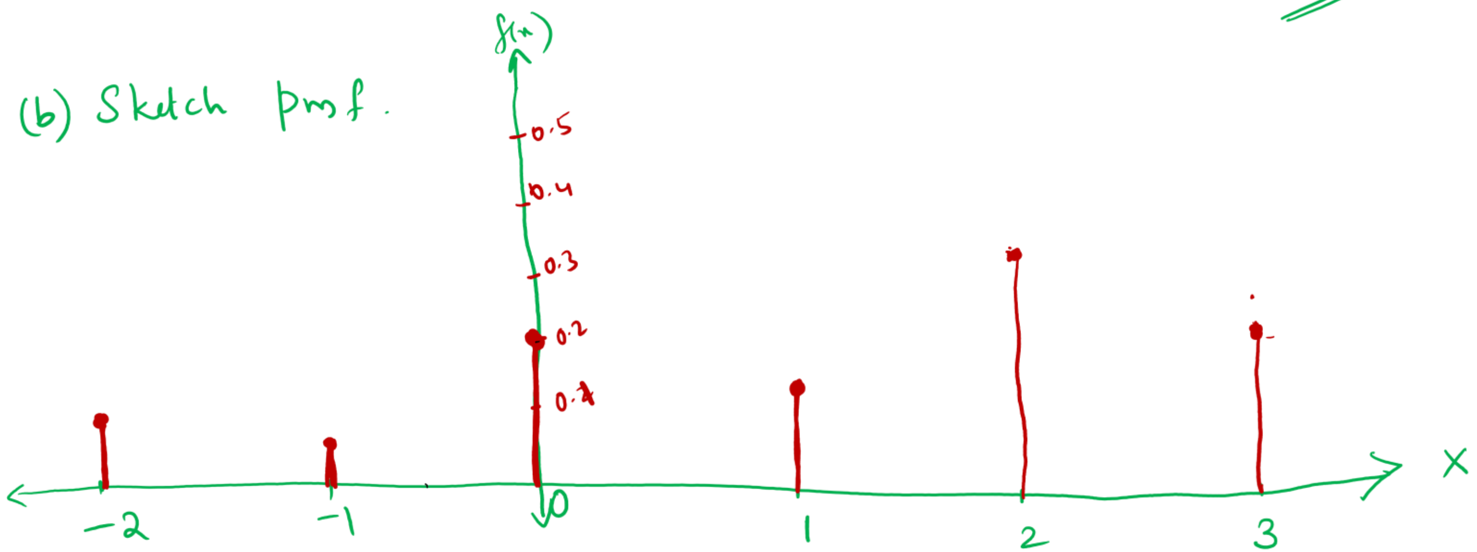
X:	-2	-1	0	1	2	3
f(X):	0.1	K	0.2	2K	0.3	3K

(a) Find 'K'.

$$\Sigma f(x) = 1$$

$$\frac{1}{10} + k + \frac{2}{10} + 2k + \frac{3}{10} + 3k = 1 \implies k = \frac{1}{15} = 0.067$$

(b) Sketch pmf.



(c) $P(-1 \le x \le 2) = P(x=-1) + P(x=0) + P(x=1) + P(x=2) = 0.701$

(d) $P(x \geq 2) = P(x=2) + P(x=3) = 0.3 + \frac{3}{15} = 0.501$

(e) $P(X \le -1) = P(x=-2) + P(x=-1) = 0.1 + 1/15 = 0.167$

(f) $P(0 < x < 3) = P(x=1) + P(x=2) = \frac{2}{15} + 0.3 = 0.401$

Prob. density function

2. . X is a continuous RV with PDF given by $f(x)=\begin{cases} kx & in \quad 0 \le x \le 2 \\ 2k & in \quad 2 \le x \le 4 \\ 6k-kx & in \quad 4 \le x \le 6 \\ 0 & elsewhere \end{cases}$ (a) Find the value of k.

$$\int_0^6 f(x)\,dx = 1 \Rightarrow \int_0^2 kx\,dx + \int_2^4 2k\,dx + \int_4^6 (6k-kx)\,dx = 1$$

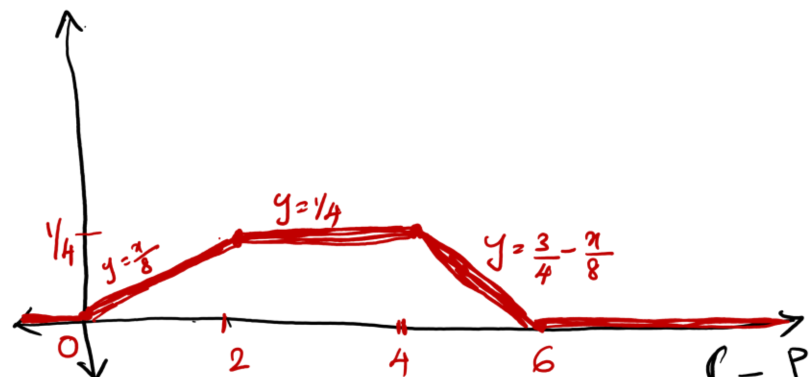
$$\Rightarrow k\frac{x^2}{2}\Big)_0^2 + 2kx\Big)_2^4 + 6kx\Big)_4^6 - \frac{k}{2}x^2\Big)_4^6 = 1 \Rightarrow 2k+4k+12k-10k=1$$

$$\Rightarrow 8k=1 \Rightarrow \boxed{k=1/8}$$

k.

(b) Sketch pdf.

$$f(x)=\begin{cases} x/8, & in \quad 0 \le x \le 2 \\ 1/4, & in \quad 2 \le x \le 4 \\ \frac{3}{4}-\frac{x}{8} & in \quad 4 \le x \le 6 \\ 0 & elsewhere. \end{cases}$$



(c) $P(1 \le x \le 3) = \int_1^3 f(x)\,dx$

$$= \int_1^2 \frac{x}{8}\,dx + \int_2^3 \frac{1}{4}\,dx = \frac{x^2}{16}\Big)_1^2 + \frac{x}{4}\Big)_2^3 = \frac{3}{16} + \frac{1}{4} = \frac{7}{16}$$

$= P(1 < x \le 3)$
$= P(1 < x < 3)$
$= P(1 \le x < 3)$

(d) $P(3 < x < 4) = \int_3^4 \frac{1}{4}\,dx = 1/4$

(e) $P(3 \le x < 5) = \int_3^4 \frac{1}{4}\,dx + \int_4^5 \left(\frac{3}{4} - \frac{x}{8}\right) dx = 7/16$

Random Variables

3. If the probability distribution of X is given as

X:	1	2	3	4
f(X):	0.4	0.3	0.2	0.1

(a) Sketch the pmf. (b) Find $P\left(\underbrace{\frac{1}{2} < X < \frac{7}{2}}_{A} / \underbrace{X > 1}_{B}\right) = P(A|B)$

$= \frac{P(A \cap B)}{P(B)} = \frac{P(1 < X < 7/2)}{P(X > 1)}$

$= \frac{0.3 + 0.2}{0.6} = \frac{5}{6}$

Random Variables

4. An assembly consists of three mechanical components. Suppose that the probablties that the first, second and third components meet specifications are 0.95, 0.98 and 0.99. Assume that the components are independent. Determine the probability mass function of the number of components in the assembly that meet specifications.

Let components meet specifications be A, B, C

0.95 0.98 0.99
A B C

Let X = no: of components in the assembly that meet specifications.

X	0	1	2	3
$f(x)$	0.00001	0.00167	0.07663	0.92169

$f(0) = P(X=0) = P(\text{each component fail}) = P(A^c \cap B^c \cap C^c) = P(A^c)\,P(B^c)\,P(C^c)$

$= (0.05)(0.02)(0.01) = 0.00001$

$f(1) = P(X=1) = P(AB'C', A'BC', A'B'C) = (0.95)(0.02)(0.01) + (0.05)(0.98)(0.01) + (0.05)(0.02)(0.99)$

$= 0.00167$

$f(2) = P(X=2) = P(ABC', A'BC, AB'C) = (0.95)(0.98)(0.01) + (0.05)(0.98)(0.99)$

$+ (0.95)(0.02)(0.99) = 0.07663$

$f(3) = P(X=3) = P(ABC) = (0.95)(0.98)(0.99)$

$= 0.92169$

Random Variables

5. Let a continuous random variable X denote the current measured in a thin copper wire in milliamperes. Assume that the range of X is [0,20mA] and assume that the probability density function of X is f(x)=0.05 for 0≤x ≤20. → This is a uniform distn.

(a) Sketch the pdf.

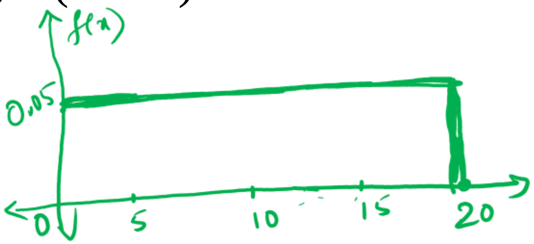
(b) What is the probability that the current measured is less than 10 amperes.

(c) What is the probability that the current measured is greater than 5 amperes.

(d) What is the probability that the current measured is between10 and 15 amperes.

(e) P(X=12)

(a)



(b) $P(X<10) = \int_0^{10} f(x)\,dx = \int_0^{10} 0.05\,dx = 0.5$

$P(X\le 10)$

$P(0\le X\le 10)$

In Discrete, $f(x)=P(X=x)$

in conts; $f(x) \neq P(X=x)$

$P(a<X<b) = \int f(x)\,dx$

(c) $P(X>5) = \int_5^{20} 0.05\,dx = 0.75$

(d) $P(10\le X\le 15) = \int_{10}^{15} 0.05\,dx = 0.25$

(e) $P(X=12) = 0$

$\boxed{\text{*** For any continuous random variable, } X,\ P(X=a)=0}$

Probability Mass function
and
Probability density Function

Warning! Continuous random variables can lead to confusion. First, note that if X is continuous then $\mathbb{P}(X = x) = 0$ for every x. Don't try to think of $f(x)$ as $\mathbb{P}(X = x)$. This only holds for discrete random variables. We get probabilities from a PDF by integrating. A PDF can be bigger than 1 (unlike a mass function). For example, if $f(x) = 5$ for $x \in [0, 1/5]$ and 0 otherwise, then $f(x) \geq 0$ and $\int f(x)dx = 1$ so this is a well-defined PDF even though $f(x) = 5$ in some places. In fact, a PDF can be unbounded. For example, if $f(x) = (2/3)x^{-1/3}$ for $0 < x < 1$ and $f(x) = 0$ otherwise, then $\int f(x)dx = 1$ even though f is not bounded.

Cumulative Distribution function for discrete and continuous random variables

{HH, HT, TH, TT}

pdf

X	0	1	2
$f(x)$	1/4	1/2	1/4

$$\boxed{F(x) = P(X \le x)} \quad -\infty \le x \le \infty$$

Conts fn.

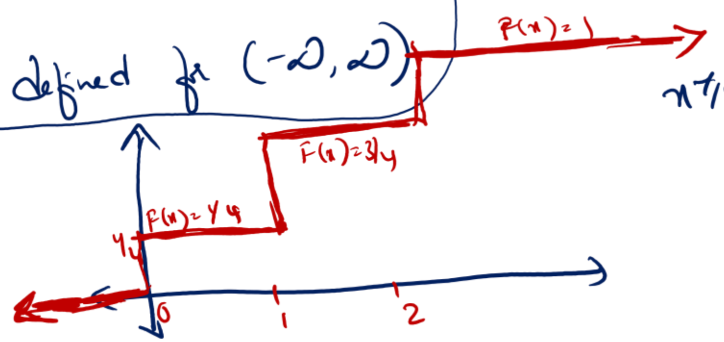
$$F(x) = \begin{cases} 0, & x < 0 \\ 1/4 & 0 \le x < 1 \\ 3/4 & 1 \le x < 2 \\ 1 & x \ge 2 \end{cases}$$

$0 \le x < 1$
$$\left.\begin{aligned} F(0) &= P(X \le 0) = P(X=0) \\ F(0.2) &= P(X \le 0.2) = P(X=0) \\ F(0.99) &= P(X \le 0.99) = P(X=0) \end{aligned}\right\} = 1/4$$

$1 \le x < 2$
$$\left.\begin{aligned} F(1) &= P(X \le 1) = P(X=0) + P(X=1) \\ F(1.1) &= P(X \le 1.1) = \text{"} \quad \text{"} \\ F(1.5) &= P(X \le 1.5) = \text{"} \quad \text{"} \\ F(1.99) &= P(X \le 1.99) = \text{"} \quad \text{"} \end{aligned}\right\} = \frac{1}{4} + \frac{1}{2} = \frac{3}{4}$$

Cumulative Distribution function of a discrete random variable X is defined as $F(x) = P(X \le x)$

where x is defined for $(-\infty, \infty)$



$x \ge 2$
$$\begin{aligned} F(2) &= P(X \le 2) \\ &= P(X=0) + P(X=1) + P(X=2) \\ &= \frac{1}{4} + \frac{1}{2} + \frac{1}{4} = 1 \end{aligned}$$
$$F(3.5) = P(X \le 3.5) = P(X=0,1,2)$$
$$F(100) = P(X \le 100) = P(X=0,1,2)$$

AMRITA VISHWA VIDYAPEETHAM UNIVERSITY Established u/s 3 of UGC Act 1956

Cumulative Distribution function for discrete and continuous random variables

(2) Find CDF of X if pmf of X is $f(x) = \begin{cases} 1/3, & x = -5 \\ \frac{1}{6}, & x = 0 \\ \frac{1}{4}, & x = 5 \\ \frac{1}{4}, & x = 10 \end{cases}$

$$F(x) = \begin{cases} 0, & x < -5 \\ 1/3, & -5 \le x < 0 \\ 1/2, & 0 \le x < 5 \\ 3/4, & 5 \le x < 10 \\ 1, & x \geq 10 \end{cases}$$

Sketch the pdf & cdf of X

3. If the random variable X takes values 1,2,3 and 4, such that 2P(X=1)=3P(X=2)=P(X=3)=5P(X=4), find the PMF and cumulative distribution function / cumulative mass function(cmf) for X. Also sketch the pmf and cmf of X.

pmf:-

x	1	2	3	4
$f(x)$	$\frac{30}{122}$	$\frac{30}{183}$	$\frac{30}{61}$	$\frac{30}{305}$

Let $P(X=3)=k$ $\quad P(X=2)=k/3$

$P(X=1)=\frac{k}{2}$ $\quad P(X=4)=k/5$

$\Sigma f(x)=1 \Rightarrow k+\frac{k}{2}+\frac{k}{3}+\frac{k}{5}=1$

$\Rightarrow k=\frac{30}{61}$

cmf:-

Cumulative Distribution function for continuous random variables – Cumulative Density Function (CDF)

2.11 Definition. *A random variable X is* **continuous** *if there exists a function f_X such that $f_X(x) \geq 0$ for all x, $\int_{-\infty}^{\infty} f_X(x)dx = 1$ and for every $a \leq b$,*

$$\mathbb{P}(a < X < b) = \int_a^b f_X(x)dx. \quad (2.2)$$

The function f_X is called the **probability density function** (PDF). *We have that*

$F(x) = P(X \leq x)$

$$F_X(x) = \int_{-\infty}^{x} f_X(t)dt$$

and $f_X(x) = F'_X(x)$ at all points x at which F_X is differentiable.

$If\ f_X(x)\ is\ PDF\ of\ continuous\ RV\ X, then$

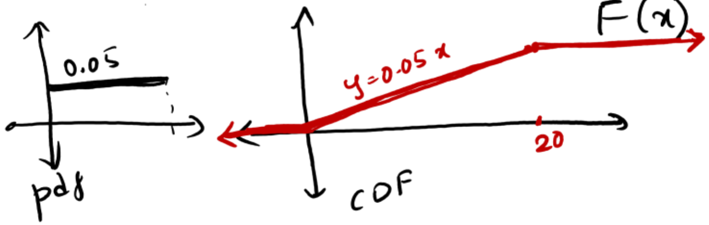
$$F_X(x) = P(X \leq x) = \int_{X=-\infty}^{X=x} f_X(x)dx$$

$is\ called\ Cumulative\ Density\ function\ or\ CDF\ of\ X.$

Cumulative Distribution function for continuous random variables – Cumulative Density Function (CDF)

Find the cumulative density function for the random variable X that had probability density function as given below.

1. f(x) = 0.05, 0 ≤ x ≤ 20



$$F(x) = \begin{cases} 0, & x < 0 \\ 0.05x, & 0 \le x < 20 \\ 1, & x \geq 20 \end{cases}$$

$0 \le a < 20$

$F(10) = P(X \le 10)$

$F(a) = P(X \le a)$

$= \int_{-\infty}^{a} f(x)\,dx$

$= \int_{0}^{a} 0.05\,dx$

$F(a) = 0.05a$

$F(x) = 0.05x$

$F(20) = P(X \le 20)$

$= \int_0^{20} 0.05\,dx = 1$

2. $$f(x) = \begin{cases} \frac{x}{2}, & 0 \le x \le 1 \\ \frac{1}{2}, & 1 \le x \le 2 \\ \frac{3}{2} - \frac{x}{2}, & 2 \le x \le 3 \\ 0, & \text{elsewhere} \end{cases}$$

If $0 \le x < 1$, $F(x) = \int_{-\infty}^{x} f(a)\,da$

$= \int_0^x \frac{a}{2}\,da = \left.\frac{a^2}{4}\right)_0^x = \frac{x^2}{4}$

$1 \le x < 2$, $F(x) = \int_{-\infty}^{x} f(a)\,da$

$= \int_0^1 f(x)\,dx + \int_1^x f(x)\,dx = \int_0^1 \frac{x}{2}\,dx + \int_1^x \frac{1}{2}\,dx = \frac{1}{4} + \frac{x}{2}$

$2 \le x < 3$, $F(x) = \int_0^1 f(a)\,da + \int_1^2 f(a)\,da + \int_2^x f(a)\,da$

$= \int_0^1 \frac{a}{2}\,da + \int_1^2 \frac{1}{2}\,da + \int_2^x \left(\frac{3}{2} - \frac{a}{2}\right)da$

$= \left.\frac{a^2}{4}\right)_0^1 + \left.\frac{a}{2}\right)_1^2 + \left.\frac{3a}{2} - \frac{a^2}{4}\right)_2^x = \frac{1}{4} + \frac{1}{2} + \frac{3}{2}(x-2) - \frac{1}{4}(x^2-4)$

$$F(x) = \begin{cases} 0, & x < 0 \\ \frac{x^2}{4}, & 0 \le x < 1 \\ \frac{x}{2} - \frac{1}{4}, & 1 \le x < 2 \\ -\frac{x^2}{4} + \frac{3x}{2} - \frac{5}{4}, & 2 \le x < 3 \\ 1, & x \geq 3 \end{cases}$$

Properties of CDF for continuous random variables (CDF - cumulative density Function)

Only for continuous random variables.

Let F be the CDF *for a random variable X. Then:*

$p(x \leq X \leq y) = p(x \leq X < y)$

$= P(x < X < y)$

$$= \mathbb{P}(x < X \leq y) = F(y) - F(x);$$

$1 - P(X \leq x)$

$$\mathbb{P}(X > x) = 1 - F(x);$$

$p(X \geq x)$

$$\begin{aligned} F(b) - F(a) &= \mathbb{P}(a < X < b) = \mathbb{P}(a \leq X < b) \\ &= \mathbb{P}(a < X \leq b) = \mathbb{P}(a \leq X \leq b). \end{aligned}$$

Cumulative Density function

1. Given the cumulative density function for a random variable X, F(x) = $\begin{cases} 0, & x < 0 \\ 1 - e^{-x}, & x \geq 0 \end{cases}$,

find (a) the probability ~~mass~~ density function

$f(x) = F'(x) = -e^{-x}(-1) = e^{-x}, \quad x \geq 0 \quad \boxed{f(x) = e^{-x}, x \geq 0}$

$F(x) = P(X \leq x) = \int_{-\infty}^{x} f(x)dx$

$f(x) = F'(x)$

(b) P(1<x<3) from CDF

$P(1 < x < 3) = F(3) - F(1) = (1 - e^{-3}) - (1 - e^{-1}) = e^{-1} - e^{-3} = 0.319$

Check:- $P(1<x<3) = \int_1^3 f(x)dx = \int_1^3 e^{-x}dx = \left.\frac{e^{-x}}{-1}\right)_1^3 = -(e^{-3} - e^{-1}) = e^{-1} - e^{-3}$

(c) P(1 ≤x<5) from CDF

$P(1 \leq x < 5) = F(5) - F(1) = e^{-1} - e^{-5} = 0.361$

(d) P(2<x ≤ 5) from CDF

$P(2 < x \leq 5) = F(5) - F(2) = e^{-2} - e^{-5} = 0.1286$

(e) P(3≤x ≤ 5) from CDF

$P(3 \leq x \leq 5) = F(5) - F(3) = e^{-3} - e^{-5} = 0.043$

(f) P(x ≤ 5) from CDF

$P(x \leq 5) = F(5) = 1 - e^{-5} = 0.993$

(g) P(x ≥ 3) from CDF

$P(x \geq 3) = 1 - P(x < 3) = 1 - F(3) = 1 - (1 - e^{-3}) = e^{-3} = 0.049$

Cumulative Density function

2. Given the pdf of a random variable X, $f(x) = \begin{cases} \frac{x}{2}, & 0 \le x \le 1 \\ \frac{1}{2}, & 1 \le x \le 2 \\ \frac{3}{2} - \frac{x}{2}, & 2 \le x \le 3 \\ 0, & \text{elsewhere} \end{cases}$, find cumulative density function(CDF) for X and evaluate the following probabilities using CDF.

(a) $P(x \ge 2)$ (b) $P(x \le 1.5)$ (c) $P(0.5 < x \le 2.5)$ (d) $P(1 \le x \le 3)$ (e) $P(2 \le x \le 5)$

$$F(x) = \begin{cases} 0, & x < 0 \\ \frac{x^2}{4}, & 0 \le x < 1 \\ \frac{x}{2} - \frac{1}{4}, & 1 \le x < 2 \\ \frac{3}{2}x - \frac{x^2}{4} - \frac{5}{4}, & 2 \le x < 3 \\ 1, & x \ge 3 \end{cases}$$

(a) $P(x \ge 2) = 1 - F(2) = 1 - \left(\frac{3}{4}\right) = \underline{\underline{1/4}}$

(b) $P(x \le 1.5) = F(1.5) = \frac{1.5}{2} - \frac{1}{4} = 1/2 //$

(c) $P(0.5 < x < 2.5) = F(2.5) - F(0.5)$
$= \left(\frac{3}{2}(2.5) - \frac{2.5^2}{4} - \frac{5}{4}\right) - \frac{(0.5)^2}{4} = \underline{\underline{0.0625}}$

(d) $P(1 \le x \le 3) = F(3) - F(1) = 1 - \left(\frac{1}{2} - \frac{1}{4}\right) = 1/4 //$

(e) $P(2 \le x \le 5) = F(5) - F(2) = 1 - \left(\frac{3}{2}(2) - \frac{2^2}{4} - \frac{5}{4}\right) =$

Mean and Variance of discrete and continuous random variables

3.1 Definition. *The* **expected value,** *or* **mean,** *or* **first moment,** *of* X *is defined to be*

$$\mathbb{E}(X) = \int x\, dF(x) = \begin{cases} \sum_x x f(x) & \text{if } X \text{ is discrete} \\ \int x f(x) dx & \text{if } X \text{ is continuous} \end{cases} \qquad (3.1)$$

X = Marks	25	20	15	17	23
f(x)	1/5	1/5	1/5	1/5	1/5

Average = 20

$= 25 \times \frac{1}{5} + 20 \times \frac{1}{5} + 15 \times \frac{1}{5} + \frac{17 \times 1}{5} + 23 \times \frac{1}{5}$

$= \sum_x x f(x) dx \longrightarrow$ Discrete RV

Conti $E(X) = \text{Average} = \text{Mean} = \mu$

$= \int_x x f(x)\, dx$

x	$(x-\mu)^2$
100	$+50^2$
100	$+50^2$
0	$(-50)^2$
0	$(-50)^2$

Variance :-

$\text{Variance}(X) = \text{Var}(X) = \sigma_x^2$

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

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

$= E(X^2) + \mu^2 - 2\mu^2$

$= E(X^2) - [E(X)]^2$ //

Mean and Variance of discrete and continuous random variables

If X is a random variable with pmf/pdf $f_x(x)$ and if g(x) is a function of X. Then,

$E(x) = \int x f(x) dx$
$E(g(x)) = \int g(x) f(x) dx$

$$E\big(g(x)\big) = \int g(x) f_x(x) dx$$

3.14 Definition. *Let X be a random variable with mean μ. The* **variance** *of X — denoted by σ^2 or σ_X^2 or $\mathbb{V}(X)$ or $\mathbb{V}X$ — is defined by*

$$\sigma^2 = \mathbb{E}(X-\mu)^2 = \int (x-\mu)^2 dF(x) \tag{3.7}$$

assuming this expectation exists. The **standard deviation** *is* $\mathrm{sd}(X) = \sqrt{\mathbb{V}(X)}$ *and is also denoted by σ and σ_X.*

σ^2 = Variance
σ = std. deviation.

Mean and Variance of discrete random variables

1. Find the mean and variance of the number heads obtained while tossing three coins.

$$S = \{\underbrace{TTT}_{0}, \underbrace{HTT, THT, TTH}_{1}, \underbrace{THH, HTH, HHT}_{2}, \underbrace{HHH}_{3}\}$$

X	0	1	2	3
$f(x)$	$\frac{1}{8}$	$\frac{3}{8}$	$\frac{3}{8}$	$\frac{1}{8}$

proof Verification :- $\sum f(x) = 1$

Mean:- $E(X) = \mu = \sum_x x f(x) = 0 \times \frac{1}{8} + 1 \times \frac{3}{8} + 2 \times \frac{3}{8} + 3 \times \frac{1}{8} = \frac{3}{2} = 1.5$

Variance:- $Var(X) = \sigma_x^2 = E\left((X-\mu)^2\right) = (-1.5)^2 \times \frac{1}{8} + (-0.5)^2 \times \frac{3}{8} + (0.5)^2 \times \frac{3}{8} + (1.5)^2 \times \frac{1}{8}$

$$= \frac{3}{4} = 0.75$$

$(x-\mu)^2$	$(-1.5)^2$	$(-0.5)^2$	$(0.5)^2$	$(1.5)^2$
$f(x)$	$\frac{1}{8}$	$\frac{3}{8}$	$\frac{3}{8}$	$\frac{1}{8}$

OR

$$\sigma^2 = E(X^2) - \mu^2 = \left(0^2 \times \frac{1}{8} + 1^2 \times \frac{3}{8} + 2^2 \times \frac{3}{8} + 3^2 \times \frac{1}{8}\right) - \left(\frac{3}{2}\right)^2$$

$$= \left(\frac{3}{8} + \frac{12}{8} + \frac{9}{8}\right) - \frac{9}{4}$$

$$= \frac{3}{4} = 0.75$$

Std. deviation $= \sigma_x = \sqrt{Var(X)} = \sqrt{\frac{3}{4}} = \frac{\sqrt{3}}{2}$

Mean and Variance of discrete random variables

2. There is a chance that a bit transmitted through a digital transmission channel is received in error. Let X equals the number of bits in error in the next four bits transmitted. The possible values of X are {0,1,2,3,4} and their corresponding probabilities are P(X=0)=0.6561; P(X=1)=0.2916; P(X=2)=0.0486; P(X=3)=0.0036 and P(X=4)=0.0001. Find the mean, variance and standard deviation of X.

x	0	1	2	3	4
$f(x)$	0.6561	0.2916	0.0486	0.0036	0.0001

$\mu = E(X) = \text{Mean} = \sum x f(x) = 0.4$

$\text{Variance} = \sum_x x^2 f(x) - \mu^2$ or $\sum_x (x-\mu)^2 f(x)$

$= 0.36$

$\text{Std deviation} = \sqrt{\text{Var}(X)} = 0.6$

Mean and Variance of continuous random variables

1. Find the mean and variance of X if the probability density function of X is f(x)=0.125x, 0<x<4.

$$f(x) = 0.125x, \quad 0 < x < 4$$

$$\text{Mean} = \mu = E(X) = \int_x x\, f(x)\, dx = \int_0^4 x \cdot (0.125\, x)\, dx = (0.125) \left.\frac{x^3}{3}\right)_0^4 = \frac{8}{3} = 2.666$$

$$\text{Variance} = \sigma^2 = E(x^2) - \mu^2 = \int_x x^2 f(x)\, dx - \left(\frac{8}{3}\right)^2$$

$$= \int_{x=0}^{4} x^2 (0.125\, x)\, dx - \left(\frac{8}{3}\right)^2$$

$$= 0.125 \left.\frac{x^4}{4}\right)_0^4 - \frac{64}{9}$$

$$= (4^3)(0.125) - \left(\frac{64}{9}\right)$$

$$= 0.8888$$

2. Given the pdf of X is $f(x)=ax^2$, $0 \le x \le 1$.

HW

(a) Find 'a'.

(b) Find the value of 'k' if $P(X \le k) = P(X > k)$. $\int_0^k f(x)\,dx = \int_k^1 f(x)\,dx$

(b) Find the mean and standard deviation of X.

Important Distributions

Discrete

-Binomial Distribution

-Poisson Distribution

Continuous

-Continuous Uniform Distribution

-Exponential Distribution

-Gaussian/Normal Distribution

Binomial Distribution

Binomial Distribution

The Bernoulli Process

Strictly speaking, the Bernoulli process must possess the following properties:

1. The experiment consists of repeated trials.
2. Each trial results in an outcome that may be classified as a success or a failure.
3. The probability of success, denoted by p, remains constant from trial to trial.
4. The repeated trials are independent.

Binomial Distribution A Bernoulli trial can result in a success with probability p and a failure with probability $q = 1 - p$. Then the probability distribution of the binomial random variable X, the number of successes in n independent trials, is

$$b(x; n, p) = \binom{n}{x} p^x q^{n-x}, \quad x = 0, 1, 2, \ldots, n.$$

$n = 2$, $p = 1/2$
Coin tossed 2 times
Success = getting a head.

	TT	TH, HT	HH
x	0	1	2
$f(x)$	$\frac{1}{4}$	$\frac{1}{2}$	$\frac{1}{4}$

$\frac{1}{4} \to 2C_0 \left(\frac{1}{2}\right)^0 \left(1-\frac{1}{2}\right)^{2-0}$

$\frac{1}{2} \to 2C_1 \left(\frac{1}{2}\right)^1 \left(1-\frac{1}{2}\right)^{2-1}$

$\frac{1}{4} \to 2C_2 \left(\frac{1}{2}\right)^2 \left(\frac{1}{2}\right)^{2-2}$

Binomial Distribution

1. The chance that a bit transmitted through a digital transmission channel is received in error is 0.1. Also, assume that the transmission trials are independent. Let X the number of bits in error in the next four bits transmitted. Determine P(X=2).

Success = receiving error.

$p = 0.1, \quad n = 4, \quad x = 0, 1, 2, 3, 4.$

$P(X=2) = {}^nC_x (0.1)^x (1-0.1)^{4-x}$

$= {}^4C_2 (0.1)^2 (1-0.1)^{4-2}$ (using formula)

$= 0.0486$

Outcome	x	Outcome	x
OOOO	0	*EOOO*	1
OOOE	1	*EOOE*	2
OOEO	1	*EOEO*	2
OOEE	2	*EOEE*	3
OEOO	1	*EEOO*	2
OEOE	2	*EEOE*	3
OEEO	2	*EEEO*	3
OEEE	3	*EEEE*	4

From table:-

$P(x=2) = 6 \times (0.1)^2 (0.9)^2$

P(OOEE, OEOE, OEEO, EOOE, EOEO, EEOO)

$P(\text{no bits are in error}) = P(X=0)$

```
>> pdf('bino',0,4,0.1)
ans =
   0.6561
```

$P(\text{at-least 1 bit is in error}) = P(X=1,2,3,4) = 1 - P(X=0) = 1 - {}^4C_0 (0.1)^0 (0.9)^4$

$= 0.3439$

```
>> 1-pdf('bino',0,4,0.1)
ans =
   0.3439
```

Binomial Distribution

1. The chance that a bit transmitted through a digital transmission channel is received in error is 0.1. Also, assume that the transmission trials are independent. Let X the number of bits in error in the next four bits transmitted. Determine P(X=2).

Bit in Error = E
Bit with error = O

Outcome	x	Outcome	x
OOOO	0	*EOOO*	1
OOOE	1	*EOOE*	2
OOEO	1	*EOEO*	2
OOEE	2	*EOEE*	3
OEOO	1	*EEOO*	2
OEOE	2	*EEOE*	3
OEEO	2	*EEEO*	3
OEEE	3	*EEEE*	4

$n = 4$, $p = P(\text{bit received is error}) = 0.1$

$$P(X=2) = B(x=2;\, n=4,\, p=0.1)$$
$$= 4C_2 (0.1)^2 (1-0.1)^{4-2}$$
$$= \frac{4\times 3}{1\times 2}(0.1)^2(0.9)^2$$
$$= 6(0.01)(0.81)$$
$$= 0.0486$$

```
>> pdf('bino',2,4,0.1) % f(2)
ans =
    0.0486
```

$P(\text{no bits are received in error}) = P(X=0) = 4C_0 (0.1)^0 (0.9)^4 = 0.6561$

```
>> pdf('bino',0,4,0.1)
ans =
    0.6561
```

$P(\text{at most 2 are in error}) = P(X=0,1,2) = 4C_0(0.1)^0(0.9)^4 + 4C_1(0.1)^1(0.9)^3 + 4C_2(0.1)^2(0.9)^2$

$= P(X\le 2)$ $= 0.9963$

```
>> cdf('bino',2,4,0.1) % F(2)
ans =
    0.9963
```

$P(\text{at least 1 in error}) = P(X=1,2,3,4) = 1 - P(X=0)$

$= 0.3439$

```
>> 1-pdf('bino',0,4,0.1)
ans =
    0.3439
```

Binomial Distribution

2. The probability that a patient recovers from a rare blood disease is 0.4. If 15 people are known to have contracted this disease, what is the probability that (a) at least 10 survive, (b) from 3 to 8 survive, and (c) exactly 5 survive?

$p = 0.4$, $n = 15$ success = patient recovers from the disease.

(a) $P(X \geq 10) = 1 - P(X < 10) = 1 - P(X \leq 9) = 1 -$ cdf('bino', 9, 15, 0.4)
$= 0.0338$

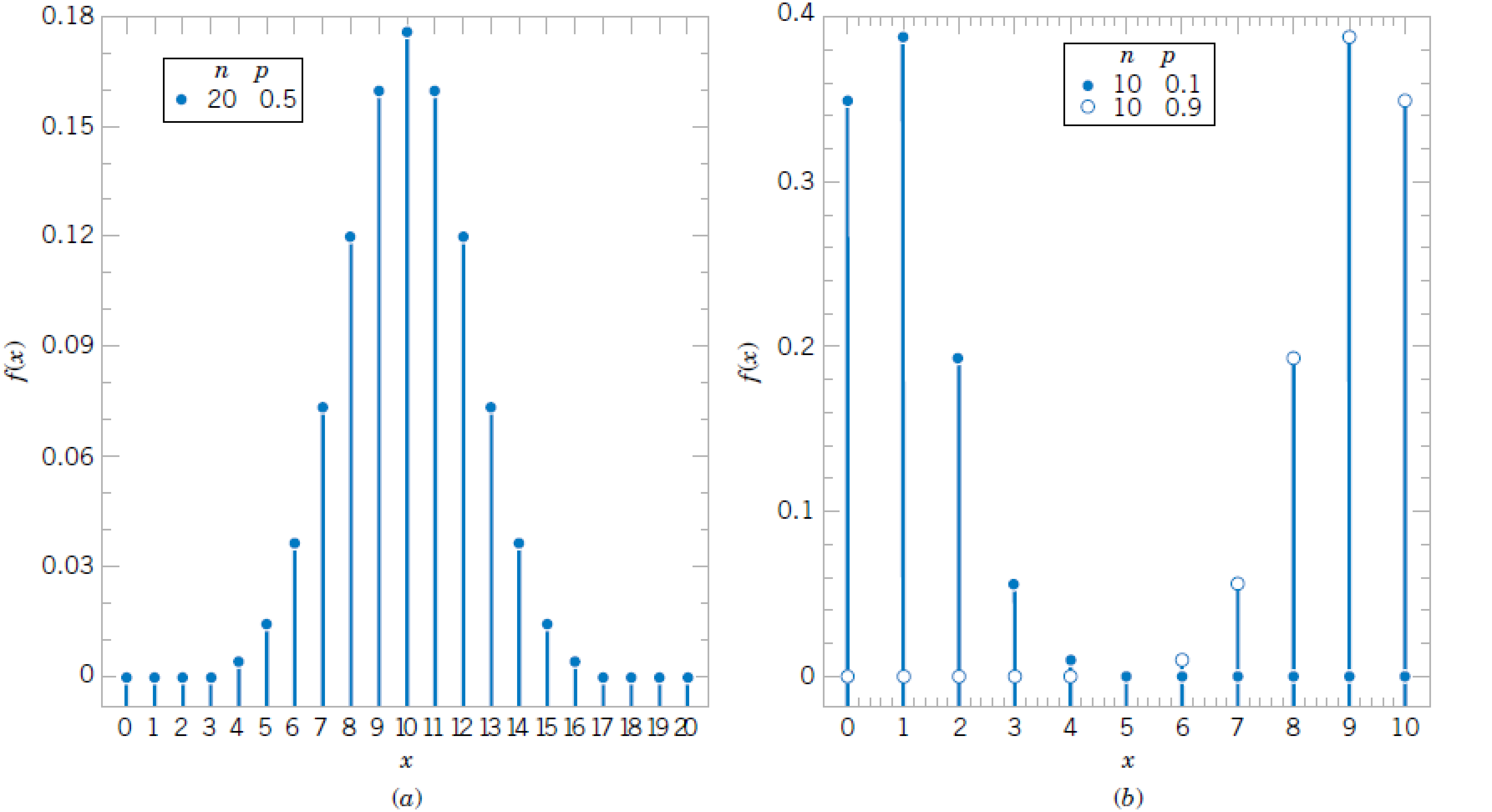
↳ $P(X = 10, 11, 12, 13, 14, 15)$

↳ >> pdf('bino', 10, 15, 0.4) + pdf('bino', 11, 15, 0.4) + - - - + pdf('bino', 15, 15, 0.4)

(b) $P(3 \leq X \leq 8) = P(X = 3, 4, 5, 6, 7, 8) = 0.8778$

(c) $P(X = 5) = 0.1859$

Binomial distributions for selected values of n and p



MGF, mean and Variance of Binomial Distribution

$$\sum_{x=0}^{n} {}^nC_x\, p^x q^{n-x} = (p+q)^n$$

$$MGF = M_X(t) = E(e^{tx}) = \sum_{x=0}^{n} {}^nC_x\, p^x (1-p)^{n-x} e^{tx}$$

$$= \sum_{x=0}^{n} {}^nC_x\, (pe^t)^x (1-p)^{n-x} =$$

$$\boxed{MGF = (pe^t + 1-p)^n}$$

$$Mean = E(X) = \frac{d}{dt}\left(M_X(t)\right)\Big|_{t=0} = \frac{d}{dt}\left(pe^t+1-p\right)^n\Big|_{t=0} = n(pe^t+1-p)^{n-1}\cdot pe^t\Big|_{t=0} = n(p+1-p)^{n-1}p$$

$$\boxed{Mean = np}$$

$$E(X^2) = \frac{d^2}{dt^2}M_X(t)\Big|_{t=0} = \frac{d}{dt}\left(npe^t\left(pe^t+1-p\right)^{n-1}\right)\Big|_{t=0} = \Big[npe^t(n-1)(pe^t+1-p)^{n-2}\cdot pe^t + npe^t\left(pe^t+1-p\right)^{n-1}\Big]_{t=0}$$

$$= np(n-1)\cdot p + np = np^2(n-1) + np = n^2p^2 - np^2 + np$$

$$Variance = E(X^2) - [E(X)]^2 = n^2p^2 - np^2 + np - n^2p^2 = -np^2 + np = np(1-p)$$

$$\boxed{\therefore\ Variance = np(1-p)}$$

If X is a binomial random variable with parameters p and n,

$$\mu = E(X) = np \qquad \text{and} \qquad \sigma^2 = V(X) = np(1-p)$$

Binomial Distribution – Mean and Variance

1. Find the mean and standard deviation of the number of times a 6 is obtained while throwing a die 9 times?

X = no: of times 6 is obtained.

$n = 9$; $p = 1/6$.

$\text{mean} = np = \frac{9}{6} = 1.5$

$\text{Variance} = np(1-p) = \frac{9}{6}\left(1-\frac{1}{6}\right) = \frac{5}{4} = 1.25$; std. dev. $= \sqrt{\frac{5}{4}} = \frac{\sqrt{5}}{2}$

2. Find the mean and variance of the number of heads obtained while tossing a coin 100 times.

X = no: of heads, $n = 100$; $p = 1/2$

$\text{Mean} = np = 50$

$\text{Variance} = np(1-p) = 25$

3. Two dice are thrown 120 times. Find the average number of times in which the number on the first dice exceeds the number on the second dice.

$n = 120$, $p = P(1^{st}$ die shows no: $> 2^{nd}$ die$) = P(A)$, where $A = \{(2,1), (3,1), (3,2), (4,1), (4,2), (4,3), (5,1), (5,2), (5,3), (5,4), (6,1), (6,2), (6,3), (6,4), (6,5)\}$

$= \frac{15}{36}$.

$\text{Mean} = np = 120 \times \frac{15}{36} = 50$

$\text{Variance} = np(1-p) = 50 \times \left(1 - \frac{15}{36}\right) = \frac{175}{6} = 29.1667$

Binomial Distribution

4. Given a binomial random variable has mean 4 and variance 3. Find P(X=1), P(X≤5), P(X≥10), P(3<X<9).

Given:- $\left.\begin{array}{l} np = 4 \\ np(1-p) = 3 \end{array}\right\}$ $1-p = \frac{3}{4}$; $\boxed{p = \frac{1}{4}}$ and $\boxed{n = 16}$

$\therefore$ Binomial Distribution is $B(x; 16, \frac{1}{4}) = f(x)$

$P(x=1) = \text{pdf}(\text{'bino'}, 1, 16, 0.25) = 0.0535$

$P(x \le 5) = F(5) = \text{cdf}(\text{'bino'}, 5, 16, 0.25) = 0.8103$

$P(x \geq 10) = 1 - P(x < 10)$

$= 1 - P(x \le 9)$

$= 1 - F(9) = 1 - \text{cdf}(\text{'bino'}, 9, 16, 0.25) = 0.0016$

$P(3 < x < 9) = P(x = 4, 5, 6, 7, 8)$

$= F(8) - F(3)$ OR $f(4) + f(5) + f(6) + f(7) + f(8)$

$= 0.5875$

$F(8) = P(x \le 8) = P(x = 0, 1, \ldots 8)$

$F(3) = P(x \le 3) = P(x = 0, \ldots 3)$

$F(8) - F(3) = P(x = 4, 6, \ldots 8)$

Binomial Distribution - Exercise

1. An electronic product contains 40 integrated circuits. The probability that any integrated circuit is defective is 0.01, and the integrated circuits are independent. The product operates only if there are no defective integrated circuits. What is the probability that the product operates?
2. The phone lines to an airline reservation system are occupied 40% of the time. Assume that the events that the lines are occupied on successive calls are independent. Assume that 10 calls are placed to the airline.
   a. What is the probability that for exactly three calls the lines are occupied?
   b. What is the probability that for at least one call the lines are not occupied?
   c. What is the expected number of calls in which the lines are all occupied?
3. Human error is given as a reason for 75% of all accidents in a plant. Use the formula for the binomial distribution to find the probability that human error will be given as the reason for 2 of the next 4 accidents.
4. An agricultural cooperative claims that 90% of the watermelons shipped out are ripe and ready to eat. Find the probabilities that among 18 watermelons shipped out,
   a. all 18 are ripe and ready to eat.
   b. Atleast 16 are ripe and ready to eat.
   c. Atmost 14 are ripe and ready to eat
5. Among the 12 solar collectors on display at a trade show, 9 are flat-plate collectors and the others are concentrating collectors. If a person visiting the show randomly selects 4 of the solar collectors to check out, what is the probability that 3 of them will be flat-plate collectors?

Poisson Distribution

Poisson Distribution

In binomial distribution, suppose that the number of bits transmitted increases and the probability of success decreases exactly enough that 'np' remains equal to a constant, $\lambda = np$.

i.e., $E(X)$ remains constant. Then, it can be proved that:

$$\lim_{n\to\infty} P(X = x) = \frac{e^{-\lambda}\lambda^x}{x!}, \qquad x = 0, 1, 2, \ldots$$

Thus the binomial distribution tends to a new distribution called the Poisson distribution that has only one parameter, λ (=np) and the probability mass function as:

$$f(x) = \frac{e^{-\lambda}\lambda^x}{x!} \quad x = 0, 1, 2, \ldots$$

AMRITA VISHWA VIDYAPEETHAM UNIVERSITY Established u/s 3 of UGC Act 1956

MGF, Mean and Variance of Poisson Distribution

$$M_X(t) = E(e^{tx}) = \sum_{x=0}^{\infty} e^{tx} \frac{e^{-\lambda}\lambda^x}{x!} = \sum_{x=0}^{\infty} \frac{(\lambda e^t)^x e^{-\lambda}}{x!}$$

$$\left(\because \sum_{x=0}^{\infty} \frac{a^x}{x!} = e^a = 1 + \frac{a}{1!} + \frac{a^2}{2!} + \frac{a^3}{3!} + \dots \right)$$

$$= e^{-\lambda} \sum_{x=0}^{\infty} \frac{(\lambda e^t)^x}{x!}$$

$$= e^{-\lambda} e^{\lambda e^t}$$

$$\text{MGF} = e^{\lambda(e^t-1)}$$

$$\text{Mean} = \frac{d}{dt}\left(M_X(t)\right)\Big|_{t=0} = \frac{d}{dt}\left(e^{\lambda(e^t-1)}\right)\Big|_{t=0} = e^{\lambda(e^t-1)} \cdot \lambda e^t \Big|_{t=0} = \lambda$$

$$E(X^2) = \frac{d^2}{dt^2} M_X(t)\Big|_{t=0} = \frac{d^2}{dt^2}\left(e^{\lambda(e^t-1)}\right)\Big|_{t=0} = \frac{d}{dt}\left(\lambda e^{t+\lambda e^t-\lambda}\right)\Big|_{t=0}$$

$$= \lambda e^{t+\lambda e^t-\lambda}\left(1+\lambda e^t\right)\Big|_{t=0}$$

$$= \lambda e^{\lambda-\lambda}(1+\lambda) = \lambda + \lambda^2$$

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

$$= \lambda + \lambda^2 - \lambda^2 = \lambda$$

Mean $= \lambda$

Variance $= \lambda$

Poisson Distribution

1. If X is a Poisson random variable with mean 5, find (P(X=0), P(X=5), P(X≤2), P(X>3).

$\lambda = 5$

$$P(X=0) = \frac{e^{-5}5^0}{0!} = e^{-5} = 0.0067$$

```
>> pdf('pois',0,5)
ans =
   0.0067
```

$$P(X=5) = \frac{e^{-5}5^5}{5!} = 0.1755$$

```
>> pdf('pois',5,5)
ans =
   0.1755
```

$$P(X\le 2) = F(2) = 0.1247$$

$$\text{OR} = P(X=0,1,2) = e^{-5}\left(\frac{5^0}{0!}+\frac{5^1}{1!}+\frac{5^2}{2!}\right)$$
$$= e^{-5}\left(1+5+\frac{25}{2}\right) = 0.1247$$

```
>> cdf('pois',2,5)
ans =
   0.1247
```

$$P(X>3) = 1-P(X\le 3) = 1-F(3) = 0.735$$

$$\text{OR} = 1-P(X=0,1,2,3) = 0.735$$

```
>> 1-cdf('pois',3,5)
ans =
   0.7350
```

Poisson Distribution

2. Suppose that the number of customers that enter a bank in an hour is a Poisson random variable such that, P(X=0) = 0.05. Determine:
   (a) Mean and variance of X
   (b) P(X≥3)

$$P(x=0)=0.05$$

$$\frac{e^{-\lambda}\lambda^0}{0!}=0.05 \Rightarrow e^{-\lambda}=0.05 \Rightarrow -\lambda=\ln(0.05)$$

$$\Rightarrow \lambda=-\ln(0.05)$$

$$\lambda=2.995$$

(a) Mean $=\lambda=2.995$

Variance $=\lambda=2.995$

(b) $P(x\geq 3)=1-P(x<3)$

$=1-P(x\leq 2)$

$=1-F(2)$

$=1-\text{cdf}(\text{'pois'}, 2, 2.995)=0.5757$

OR

$=1-P(x=0,1,2)=1-(f(0)+f(1)+f(2))=1-0.4243=0.5757$

Poisson Distribution

3. The number of telephone calls that arrive at a phone exchange is modelled as a Poisson random variable. Assume that on the average there are 10 calls per hour.
   (a) What is the probability that there are exactly 5 calls in one hour.
   (b) What is the probability that there are exactly 20 calls in two continuous hours.

(a) Mean $= \lambda = 10/\text{hr}.$

$$P(X=5) = \frac{e^{-10}\,10^5}{5!} = 0.0378$$

OR $=$ pdf('pois', 5, 10) $= 0.0378$

(b) $P(20 \text{ calls in 2 hrs}) = P(X=20, \lambda=20)$ (In 2 hrs, $\lambda=20$)

$$= \frac{e^{-20}(20)^{20}}{20!} \approx 0.0888$$

$=$ pdf('pois', 20, 20)

$$\frac{e^{-2\lambda}(2\lambda)^{x}}{x!}$$

Poisson Distribution - Exercise

1. The number of messages sent to a computer bulletin board is a Poisson random variable with mean of 5 messages per hour. What is the probability that at least two messages are received in 1 hour?

2. If X is a Poisson variate such that 2P(X=0) + P(X=2) = 2P(X=1), find E(X).

3. If X is a Poisson variate such that $E(X^2) = 6$, find E(X).

4. If X is a Poisson variate such that P(X=0) = 0.5, find variance of X.

5. Patients arrive randomly and independently at a doctor's consulting room from 5 P.M at an average rate of one in 5 min. The waiting room can hold 12 persons. What is the probability that the room will be full, when the doctor arrives at 6 P.M.?

Conti. R.V $\quad a \leq X \leq b$.

$f(x)$ is a pdf if

$$F(b)-F(a) = \left\{\begin{array}{l} P(a \leq X \leq b) \\ P(a < X < b) \\ P(a \leq X < b) \\ P(a < X \leq b) \end{array}\right\} = \int_a^b f(x)\,dx.$$

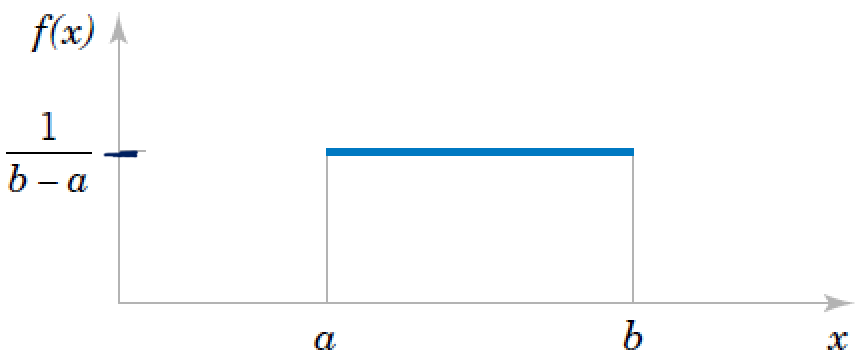
Continuous Uniform Distribution

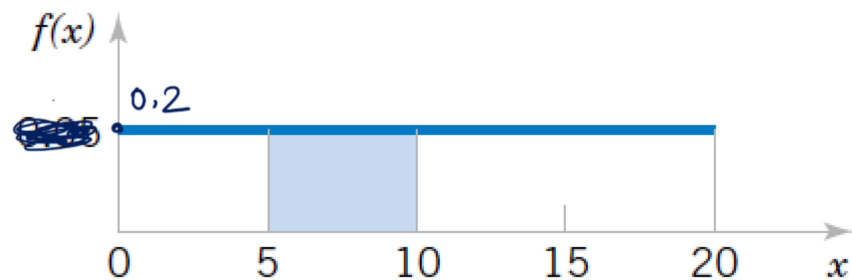
Continuous Uniform Distribution

A continuous random variable X with probability density function

$$f(x) = 1/(b - a), \qquad a \le x \le b$$

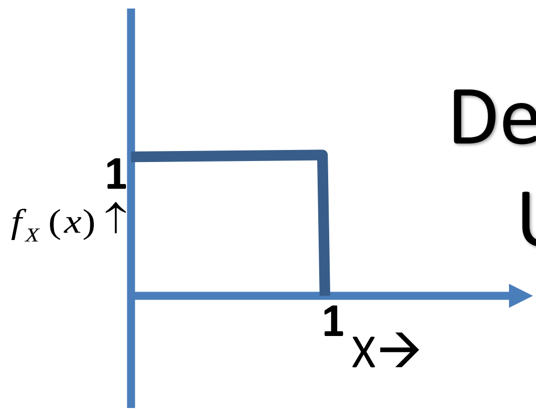
is a **continuous uniform random variable.**





$\frac{1}{5} = 0.2$

Continuous Uniform Distribution



Density Function Uniform(0,1)

rand(1) → no: for uniformly distributed b/w 0 & 1

$$U(0,1) = \frac{1}{1-0}, \quad 0 \le x \le 1$$
$$= 1, \quad 0 \le x \le 1$$

$$U(5,10) = \frac{1}{5}, \quad 5 \le x \le 10$$
$$= 0.2, \quad 5 \le x \le 10$$

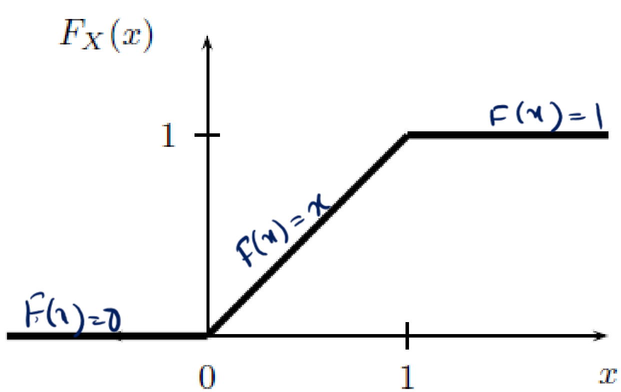


FIGURE . CDF for Uniform (0,1).

Cumulative Density Function Uniform(0,1)

$$F(x) = \begin{cases} 0, & x < 0 \\ x, & 0 \le x \le 1 \\ 1, & x \ge 1 \end{cases}$$

For U(0,1)

$$\int_0^x f(u)\,du$$
$$\int_0^x 1\,du = x$$

Mean, Variance and MGF of Continuous Uniform Distribution

$$f(x) = 1/(b-a), \qquad a \le x \le b$$

$$\mu_U = \frac{a+b}{2}$$

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

$$M_X(t) = \frac{e^{bt}-e^{at}}{t(b-a)}$$

$$E(X) = \int_a^b \frac{x}{b-a}\,dx = \frac{0.5x^2}{b-a}\bigg|_a^b = \frac{(a+b)}{2}$$

$\int x f(x)\,dx$

$$V(X) = \int_a^b \frac{\left(x-\left(\frac{a+b}{2}\right)\right)^2}{b-a}\,dx = \frac{\left(x-\frac{a+b}{2}\right)^3}{3(b-a)}\Bigg|_a^b = \frac{(b-a)^2}{12}$$

$E(X-\mu)^2$
or
$E(X^2)-\mu^2$

$$M_X(t) = E(e^{tx}) = \int_a^b e^{tx}\left(\frac{1}{b-a}\right)dx = \frac{e^{bt}-e^{at}}{t(b-a)}$$

$$= \frac{1}{b-a}\left[\frac{e^{tx}}{t}\right]_a^b = \frac{1}{t(b-a)}\left[e^{bt}-e^{at}\right]$$

Continuous Uniform Distribution

1. Let the continuous random variable *X* denote the current measured in a thin copper wire in milliamperes. Assume that the range of *X* is [0, 20 mA], and assume that the probability density function of *X* is f(x) = 0.05, 0 ≤ x ≤ 20. What is the probability that a measurement of current is between 5 and 10 milliamperes?

$f(x) = \frac{1}{20-0}, \; 0 \le x \le 20$ $\boxed{f(x) = 0.05, \; 0 \le x \le 20}$

$P(5 \le x \le 10) = \int_5^{10} 0.05\, dx = 0.05\, x \Big)_5^{10}$

$= 0.25$

```
>> cdf('unif',10,0,20)-cdf('unif',5,0,20)
ans =
   0.2500
```

$x \quad a \quad b$

$F(10) - F(5)$

2. Suppose the time it takes a data collection operator to fill out an electronic form for a database is uniformly between 1.5 and 2.2 minutes.What is the mean and variance of the time it takes an operator to fill out the form?

$f(x) = \frac{1}{2.2-1.5} = \frac{1}{0.7} = \frac{10}{7}, \; 1.5 \le x \le 2.2$

$\text{Mean} = \frac{a+b}{2} = \frac{1.5+2.2}{2} = 1.85$

$\text{Variance} = \frac{(b-a)^2}{12} = \frac{(2.2-1.5)^2}{12} = 0.04083$

Continuous Uniform Distribution

3. The thickness of a flange on an aircraft component is uniformly distributed between 0.95 and 1.05 millimeters.

   (a) Determine the cumulative distribution function of flange thickness.

   (b) Determine the proportion of flanges that exceeds 1.02 millimeters.

   (c) What thickness is exceeded by 90% of the flanges?

   (d) Determine the mean and variance of flange thickness.

$$\boxed{f(x) = 10 \quad , 0.95 \le x \le 1.05}$$

(a) $F(x) = \begin{cases} 0, & x < 0 \\ 10x - 9.5, & 0.95 \le x \le 1.05 \\ 1, & x \geq 1.05 \end{cases}$

In $0.95 \le x \le 1.05$

$F = \int_{0.95}^{x} 10\, dx$

$= 10(x - 0.95)$

(b) $P(X > 1.02) = \int_{1.02}^{1.05} f(x)\,dx = \underline{\underline{0.3}}$ OR $P(x > 1.02) = 1 - \text{cdf}(\text{'Unif'}, 1.02, 0.95, 1.05)$

$= \underline{\underline{0.3}}$

(c) $P(X > a) = 0.9$ (Find 'a')

$\int_{a}^{1.05} f(x)\,dx = 0.9 \Rightarrow 10(1.05 - a) = 0.9 \Rightarrow 10a = 10.5 - 0.9$

$\Rightarrow \boxed{a = 0.96}$

(d) Mean $= \dfrac{a+b}{2}$; Variance $= \dfrac{(b-a)^2}{12}$

$= \underline{\underline{1}}$; $= \dfrac{0.01}{12} = \underline{\underline{0.00083}}$

Continuous Uniform Distribution - Exercise

1. On a certain city transport route buses ply uniformly every 30 minutes between 6 am and 9 pm. If a person reaches a bus stop on this route at a random time during this period, what is the probability that he will have to wait for at least 10 minutes?

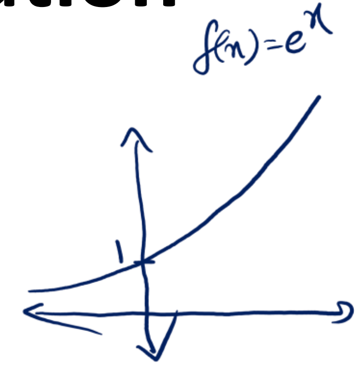
2. Suppose X has a uniform distribution over the interval [1.5, 5.5]. Determine (i) the mean and variance (ii) $P(X < 2.5)$ (iii) $P(X > 2 / X < 2.5)$ and (iv) $P(X > 1.5 / X < 2.5)$

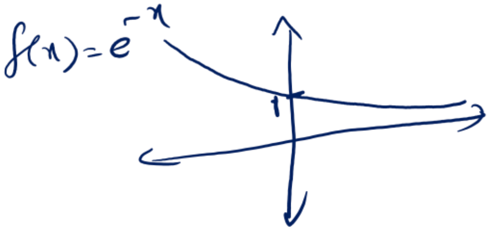
3. Suppose X has a continuous uniform distribution over the interval [-1,1]. Determine (i) the mean and standard deviation (ii) the value of x such that $P(-x < X < x) = 0.90$

4. If X is a random variable uniformly distributed in the interval (a, 9) and $P(3 < X < 5) = 2/7$. Find ‘a’ and compute $P(|X - 5| < 2)$.

5. If X is a random variable uniformly distributed in the interval (-a, a), determine ‘a’ if $P(-1 < X < -2) = 0.75$.

Exponential Distribution





Exponential Distribution

Let N (number of arrivals/ counts) be a random variable having Poisson distribution with mean λx and X be the random variable that gives the time between each arrival/count. Then,

$$P(X > x) = P(N = 0) = \frac{e^{-\lambda x}(\lambda x)^0}{0!} = e^{-\lambda x}$$

$1 - F(x) = e^{-\lambda x}$

$F(x) = 1 - e^{-\lambda x}$ → exp. dist

i.e., the CDF of X is:

$$F(x) = P(X \le x) = 1 - e^{-\lambda x}, \quad x \ge 0$$

Differentiating this we get the pdf of X as:

$$f(x) = \lambda e^{-\lambda x}, x \ge 0$$

which is called the **Exponential Distribution**.

pdf of Exp. distr.

$f(x) = \lambda e^{-\lambda x}, x \geq 0$

$f(x) = \frac{1}{\beta} e^{-x/\beta}, x \geq 0$

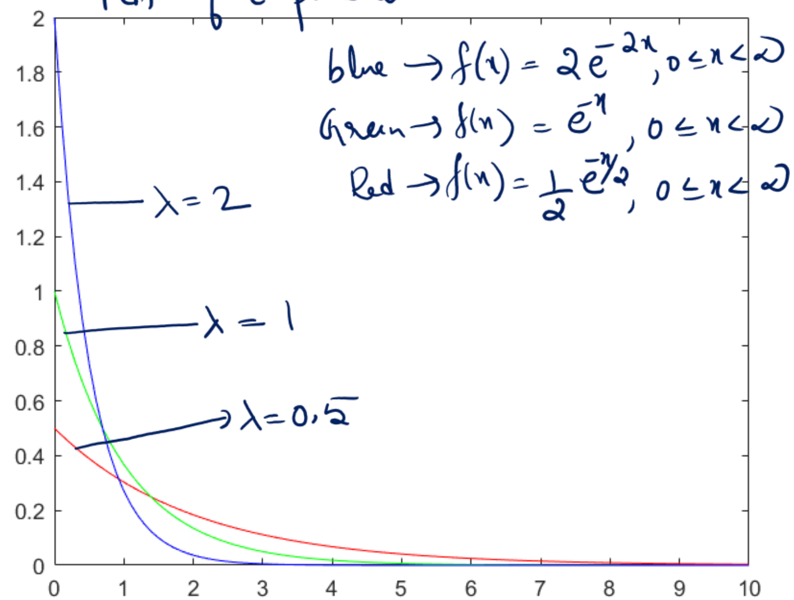
$\int_0^\infty \lambda e^{-\lambda x} dx = \lambda \left.\frac{e^{-\lambda x}}{-\lambda}\right)_0^\infty = -1 (0 - 1) = 1$

The random variable X that equals the distance between successive counts of a Poisson process with mean $\lambda > 0$ is an **exponential random variable** with parameter λ. The probability density function of X is

$$f(x) = \lambda e^{-\lambda x} \quad \text{for} \quad 0 \le x < \infty \tag{4-14}$$

Pdf and cdf of Exponential Distribution

Pdf of Exponential distr.



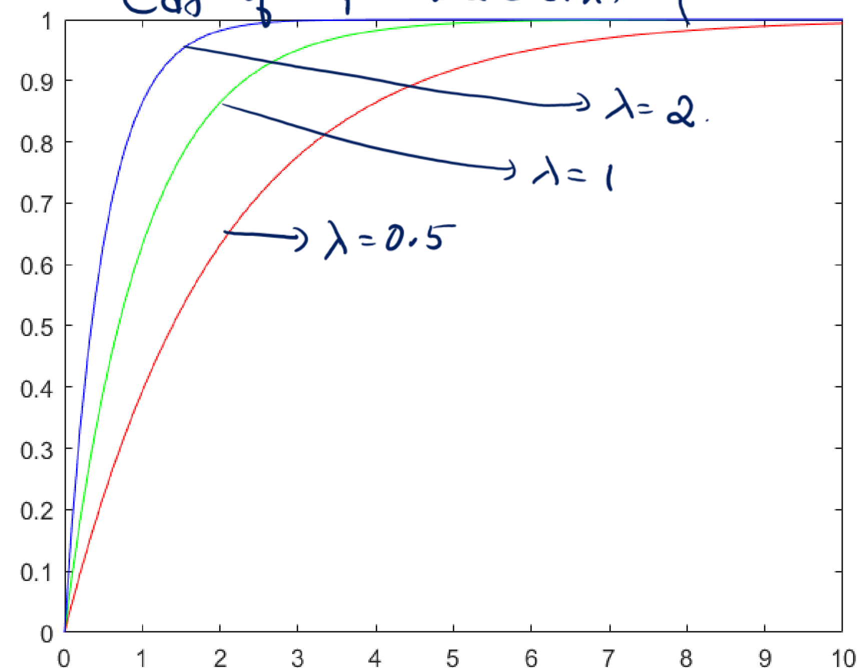
$$f(x) = \lambda e^{-\lambda x}, \quad 0 \le x < \infty.$$

$$F(x) = \begin{cases} 0, & x < 0 \\ \int_0^x \lambda e^{-\lambda u}\, du, & x < 0 \end{cases}$$

$$\int_0^x \lambda e^{-\lambda u}\, du = \left.\frac{\lambda e^{-\lambda u}}{-\lambda}\right)_0^x = (-1)\left(e^{-\lambda x} - 1\right) = 1 - e^{-\lambda x}$$

$$F(x) = \begin{cases} 0, & x < 0 \\ 1 - e^{-\lambda x}, & x \ge 0 \end{cases}$$

Cdf of exponential distn.



MGF, Mean and Variance of Exponential Distribution

<u>MGF</u> :- $M_x(t) = E(e^{tx}) = \int_x e^{tx} f(x)\,dx = \int_0^\infty e^{tx} \lambda e^{-\lambda x}\,dx = \int_0^\infty \lambda\, e^{-x(\lambda-t)}\,dx$

$= \left.\dfrac{\lambda e^{-x(\lambda-t)}}{-(\lambda-t)}\right)_0^\infty = \dfrac{-\lambda}{\lambda-t}(0-1) = \dfrac{\lambda}{\lambda-t}$

$f(x) = \lambda e^{-\lambda x}$
$M_x(t) = \dfrac{\lambda}{\lambda-t}$
Mean $= 1/\lambda$
Variance $= 1/\lambda^2$

$f(x) = \dfrac{1}{\beta} e^{-x/\beta}$
mean $= \beta$.
variance $= \beta^2$

Mean :- $E(X) = \left.\dfrac{d}{dt} M_x(t)\right)_{x=0} = \left.\dfrac{d}{dt}\left(\dfrac{\lambda}{\lambda-t}\right)\right|_{t=0} = \left.\dfrac{-\lambda(-1)}{(\lambda-t)^2}\right)_{t=0} = \dfrac{\lambda}{\lambda^2} = \dfrac{1}{\lambda}$

$E(X^2) = \dfrac{d^2}{dt^2} M_x(t) = \left.\dfrac{d}{dt}\left(\dfrac{\lambda}{(\lambda-t)^2}\right)\right)_{t=0} = \left.\dfrac{\lambda(-2)(-1)}{(\lambda-t)^3}\right)_{t=0} = \dfrac{2\lambda}{\lambda^3} = \dfrac{2}{\lambda^2}$

<u>Variance</u> :- $Var(X) = E(X^2) - [E(X)]^2$

$= \dfrac{2}{\lambda^2} - \dfrac{1}{\lambda^2}$

$= \dfrac{1}{\lambda^2}$

For Exp. Distr. :- $f(x) = \lambda e^{-\lambda x}, x \geq 0$
Mean $= 1/\lambda$
Variance $= 1/\lambda^2$
Std. deviation $= 1/\lambda$

$$f(x) = \lambda e^{-\lambda x}, x \geq 0; \quad M_X(t) = \frac{\lambda}{\lambda - t}; \quad \mu = E(X) = \frac{1}{\lambda} \quad \text{and} \quad \sigma^2 = V(X) = \frac{1}{\lambda^2}$$

Exponential Distribution

1. Suppose X has an exponential distribution with mean 0.5. Determine the following:

   (a) $P(X \leq 0)$ (b) $P(X \geq 2)$

   (c) $P(X \leq 1)$ (d) $P(1 < X < 2)$

   (e) Find the value of x such that $P(X < x) = 0.05$.

Mean of exp. distrn. $= 0.5 \Rightarrow \frac{1}{\lambda} = 0.5 \Rightarrow \boxed{\lambda = 2} \Rightarrow \boxed{f(x) = 2e^{-2x},\ x \geq 0}$

(a) $P(X \leq 0) = F(0) = 0$

(b) $P(X \geq 2) = \int_2^\infty 2e^{-2x}\,dx = \left.\frac{2e^{-2x}}{-2}\right)_2^\infty$

$= (-1)(0 - e^{-4})$

$= e^{-4} = 0.0183$

OR $P(X \geq 2) = 1 - F(2)$

$= 1 - \text{cdf}(\text{'exp'}, \underset{x}{2}, \underset{\text{mean}=\frac{1}{\lambda}}{0.5})$

$= 0.0183$

(c) $P(X \leq 1) = F(1) = \text{cdf}(\text{'exp'}, 1, 0.5) = 0.8647$

OR $= \int_0^1 2e^{-2x}\,dx = e^{-2}$

(d) $P(1 < X < 2) = F(2) - F(1) = \text{cdf}(\text{'exp'}, 2, 0.5) - \text{cdf}(\text{'exp'}, 1, 0.5) = 0.117$

OR $= \int_1^2 2e^{-2x}\,dx = e^{-4} - e^{-2}$

(e) $P(X < x) = 0.05 \Rightarrow \int_0^x f(u)\,du = 0.05 \Rightarrow \int_0^x 2e^{-2u}\,du = 0.05$

$\Rightarrow \left.\frac{2e^{-2u}}{-2}\right)_0^x = 0.05$

$(-1)(e^{-2x} - 1) = 0.05$

$1 - e^{-2x} = 0.05$

$e^{-2x} = 0.95$

$-2x = \ln(0.95)$

$\Rightarrow x = -\frac{1}{2}\ln(0.95)$

$x = 0.0256$

OR $x = \text{icdf}(\text{'exp'}, 0.05, 0.5) = 0.0256$

Exponential Distribution

2. The time to failure (in hours) for a laser in a cytometry machine is modeled by an exponential distribution with λ=0.00004. (a) What is the probability that the laser will last at least 20,000 hours? (b) What is the probability that the laser will last at most 30,000 hours? (c) What is the probability that the laser will last between 20,000 and 30,000 hours?

$f(x) = 0.00004\, e^{-0.00004x}$, $\text{mean} = \frac{1}{\lambda} = \frac{1}{0.00004} = 25000$

(a) $P(X \geq 20{,}000) = 1 - \text{cdf}(\text{'exp'}, \underset{x}{20000}, \underset{\text{mean}}{25000}) = 0.4493$

OR $\int_{20000}^{\infty} 0.00004\, e^{-0.00004x}\, dx = e^{-0.8}$

```
>> 1-cdf('exp',20000,25000)
ans =
   0.4493
```

(b) $P(X \leq 30{,}000) = F(30000)$

OR $= \int_{0}^{30000} 0.00004 \cdot e^{-0.00004x}\, dx$

$= 1 - e^{-1.2}$

```
>> cdf('exp',30000,25000)
ans =
   0.6988
```

(c) $P(20000 \leq X \leq 30000)$

$= F(30000) - F(20000)$

OR $= \int_{20000}^{30000} f(x)\, dx = e^{-1.2} - e^{-0.8}$

```
>> cdf('exp',30000,25000)-cdf('exp',20000,25000)
ans =
   0.1481
```

Exponential Distribution

3. Let X denote the time between detections of a particle with a geiger counter and assume that X has an exponential distribution with mean 1.4 minutes. The probability that we detect a particle within 30 seconds of starting the counter is __________.

$\frac{1}{\lambda} = 1.4$ min., $\lambda = \frac{1}{1.4}$, $f(x) = \frac{1}{1.4} e^{-x/1.4}$, $x \geq 0$

$$P(X < 0.5 \text{ minute}) = F(0.5) = 1 - e^{-0.5/1.4} = 0.30$$

```
>> cdf('expo',0.5,1.4)
ans =
   0.3003
```

exp, x, $\mu = \frac{1}{\lambda}$

4. Now, suppose we turn on the Geiger counter and wait 3 minutes without detecting a particle. What is the probability that a particle is detected in the next 30 seconds?

$$P(\underbrace{X < 3.5}_{A} \mid \underbrace{X > 3}_{B}) = P(\underbrace{3 < X < 3.5}_{A \cap B})/P(\underbrace{X > 3}_{B})$$

$$P(3 < X < 3.5) = F(3.5) - F(3) = [1 - e^{-3.5/1.4}] - [1 - e^{-3/1.4}] = 0.0352$$

$$P(X > 3) = 1 - F(3) = e^{-3/1.4} = 0.117$$

$$P(X < 3.5 \mid X > 3) = 0.035/0.117 = 0.30$$

```
>> (cdf('expo',3.5,1.4)-cdf('expo',3,1.4))/(1-cdf('expo',3,1.4))
ans =
   0.3003
```

$P(x<0.5)$ is same as $P(x<3.5 \mid x>3)$ → This behaviour is typical in exp. r.vs. & is called the Lack of memory property.

Lack of memory property of Exponential Distribution

In the previous problem, after waiting for 3 minutes without a detection, the probability of a detection in the next 30 seconds is the same as the probability of a detection in the 30 seconds immediately after starting the counter. The fact that you have waited 3 minutes without a detection does not change the probability of a detection in the next 30 seconds.

$$\boxed{P(X<3.5/X>3) = P(X<0.5)}$$

$$\begin{aligned}
P(X > x + a \mid X > a) &= \frac{P(X > x + a, X > a)}{P(X > a)} \\
&= \frac{P(X > x + a)}{P(X > a)} \\
&= \frac{1 - F_X(x + a)}{1 - F_X(a)} \\
&= \frac{e^{-\lambda(x+a)}}{e^{-\lambda a}} \\
&= e^{-\lambda x} \\
&= P(X > x).
\end{aligned}$$

$F_X = 1 - e^{-\lambda x}$.

$\downarrow$

exp r.v

Exponential Distribution

If $F(a) = b$, $a = icdf('exp', b, u)$

5. Suppose that the log-ons to a computer network follow a Poisson process with an average of 3 counts per minute.

$\boxed{F(a) = 1 - e^{-\lambda a}}$

(a) What is the mean time between counts?

(b) What is the standard deviation of the time between counts?

(c) Determine *a* such that the probability that at least one count occurs before time *a* minutes is 0.95.

If arrivals (log-ons) follows Poisson process, time b/w counts follow exponential distribution.

$\lambda = 3/\text{minute}$

(a) Mean time b/w counts $= \frac{1}{\lambda} = \frac{1}{3}$ minutes $= 20$ seconds.

(b) Var.(time b/w counts) $= \frac{1}{\lambda^2} = \frac{1}{9} \Rightarrow$ std. dev. $= 1/3$ min $= 20$ seconds

(c) $P(x < a) = 0.95$

$F(a) = 0.95 \Rightarrow 1 - e^{-3a} = 0.95$

$e^{-3a} = 0.05$

$-3a = \ln(0.05)$

$a = -\frac{1}{3}\ln(0.05)$

$a = 0.9985$

$a = icdf('exp', 0.95, 1/3)$

$= \underline{\underline{0.9985}}$

6. The time between the arrival of electronic messages at your computer is exponentially distributed with a mean of two hours. $\lambda = 0.5/\text{hr}$.

for 2hrs no message $\Rightarrow$ waiting time > 2 hrs

(a) What is the probability that you do not receive a message during a two-hour period?

(b) If you have not had a message in the last four hours, what is the probability that you do not receive a message in the next two hours?

(c) What is the expected time between your fifth and sixth messages?

(a) $P(X > 2) = 1 - F(2) = 1 - cdf('exp', 2, 2) = 0.3679$

mean

$= 1 - (1 - e^{-2(0.5)}) = \underline{\underline{e^{-1}}} = 0.3679$

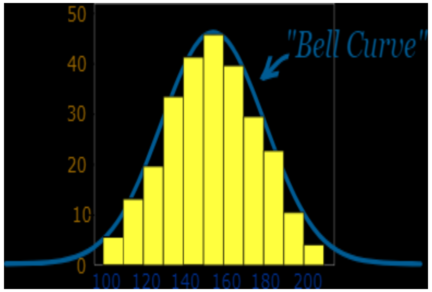
(b) $P(X > 6 / X > 4)$, Due to lack of memory property of exp. R.V, we can say that

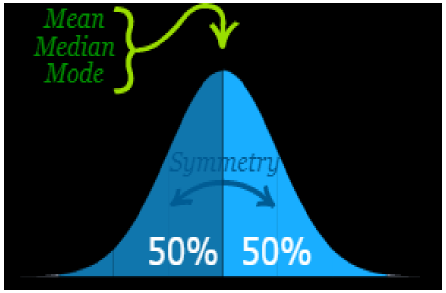
$P(X > 6 / X > 4) = P(X > 2) = \underline{\underline{0.3679}}$

(c) Expected time b/w 5th & 6th message = E(time b/w arrivals) $= E(X) = \frac{1}{\lambda} = \underline{\underline{2 \text{ hrs}}}$

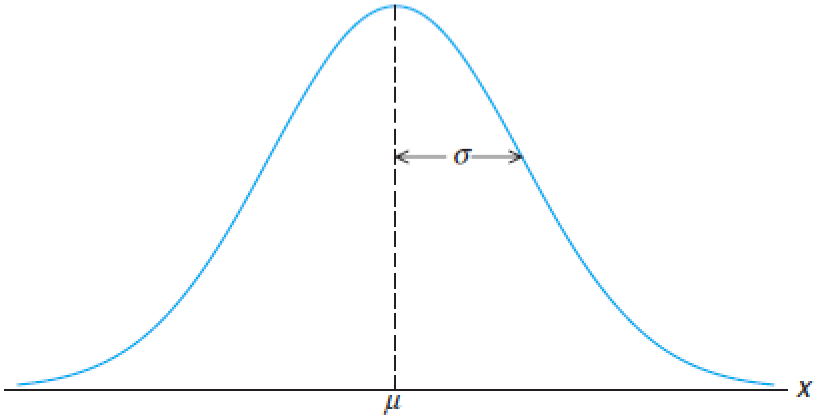
Exponential Distribution - Exercise

1. The length of a telephone conversation in a booth has been an exponential distribution and found on an average to be 5 minutes. Find the probability that a random call made from a booth (i) ends in less than 5 minutes. (ii) between 5 and 10 minutes.
2. In a certain town the duration of a shower is exponentially distributed with mean 5 minutes. What is the probability that a shower will last for 10 or more minutes.
3. Suppose X has an exponential distribution with mean equal to 10. Determine : (a) $P(X > 10)$ b) $P(X > 20)$ c) $P(X > 30)$ d) Find a such that $P(X < a) = 0.95$.
4. The time to failure in hrs of fans in a personal computer can be modeled by an exponential distribution with $\lambda = 0.0003$.
   (a)What proportion of fans will last at least 10,000 hrs?
   (b)What proportion of fans will last at most 7,000 hrs?
5. The amount of time that a surveillance camera will run without having to be reset is a RV having exponential distribution with mean 50 days. Find the probabilities that such a camera (a) will have to be reset in less than 20 days? (b) not have to be reset in at least 60 days?
6. The amount of time that a watch will run without having to be reset is a random variable having an exponential distribution with mean 120 days. Find the probability that such a watch will
   a. have to be set in less than 24 days and
   b. not have to be reset in at least 180 days.





Normal Distribution / Gaussian Distribution



Normal Distribution / Gaussian Distribution

Normal/Gaussian is One of the most widely used probability distribution

- **If you collect and plot histogram of heights/weights of student population in a college, it will be a bell shaped curve**
- **If you collect and plot histogram of mark of students in a subject, it will be approximately normally distributed.**
- **You generate 12 (it can be any large number) random numbers from Uniform distribution and add. Let it be y1.**
  Repeat the process and generate y2,y3,….., y1000.
  Plot histogram of y1,…,y1000. It will be normally distributed
  This is what is called central limit theorem in probability theory, that we shall study in the coming semesters

Normal Distribution

$\int e^{-x^2} dx =$ const do integration

A random variable X with probability density function

pdf of $N(\mu, \sigma)$,

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

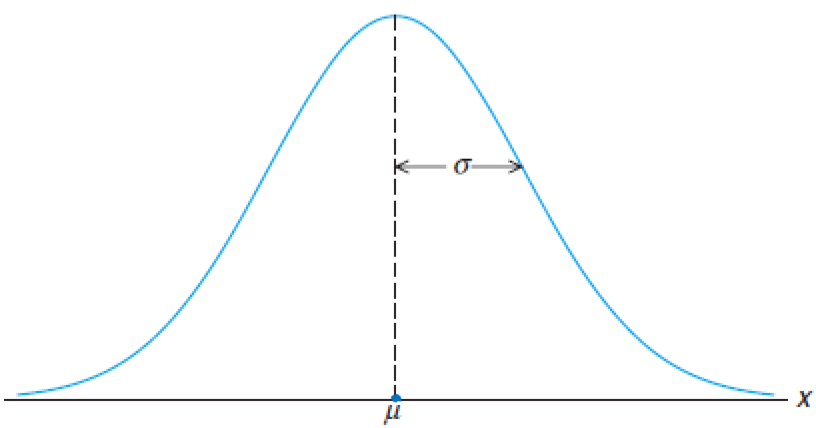
$f(x, \mu, \sigma^2)$

is a **normal random variable** with parameters μ, where $-\infty < \mu < \infty$, and $\sigma > 0$. Also,

$$E(X) = \mu \quad \text{and} \quad V(X) = \sigma^2$$

MGF of Normal Distribution

$$M_X(t) = e^{\mu t + \frac{\sigma^2 t^2}{2}}$$



Pdf of Normal Distribution with different means and variances

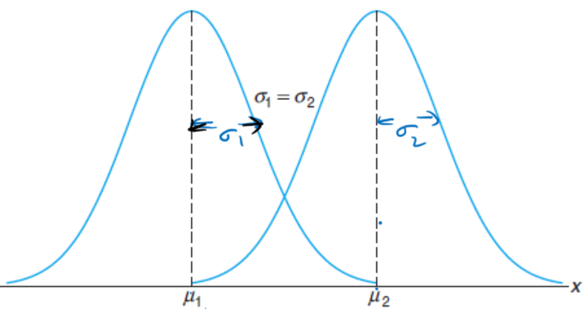


Figure 6.3: Normal curves with $\mu_1 < \mu_2$ and $\sigma_1 = \sigma_2$.

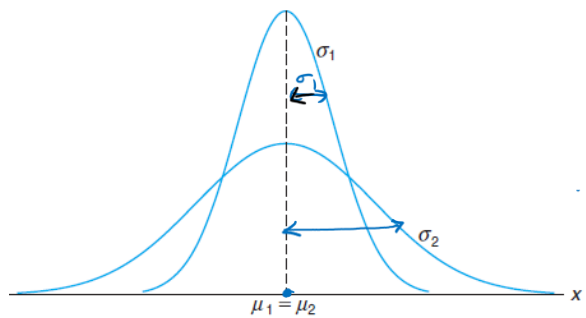


Figure 6.4: Normal curves with $\mu_1 = \mu_2$ and $\sigma_1 < \sigma_2$.

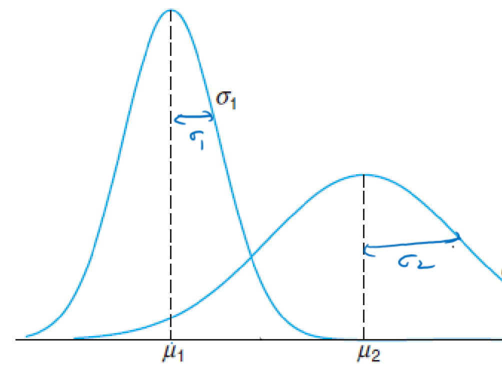
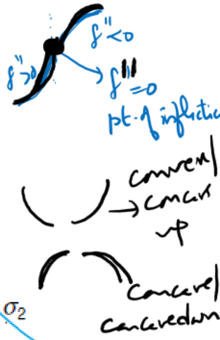


Figure 6.5: Normal curves with $\mu_1 < \mu_2$ and $\sigma_1 < \sigma_2$.



$\int_{-\infty}^{\infty} f(x)\,dx = 1$

1. The mode, which is the point on the horizontal axis where the curve is a maximum, occurs at $x = \mu$.
2. The curve is symmetric about a vertical axis through the mean μ.
3. The curve has its points of inflection at $x = \mu \pm \sigma$; it is concave downward if $\mu - \sigma < X < \mu + \sigma$ and is concave upward otherwise.
4. The normal curve approaches the horizontal axis asymptotically as we proceed in either direction away from the mean.
5. The total area under the curve and above the horizontal axis is equal to 1.

$\int_{-\infty}^{\infty} f(x)\,dx = 1$

Normal Distribution

1. Given that X has a normal distribution with $\mu = 300$ and $\sigma = 50$, find the probability that X assumes a value (a) greater than 362 (b) lesser than 290 (c) between 280 and 340

$\mu = 300, \quad \sigma = 50.$

(a) $P(X > 362) = 1 - F(362)$

$= 1 - \text{cdf}(\text{'norm'}, \underset{x}{362}, \underset{\mu}{300}, \underset{\sigma}{50})$

$= 0.1075$

```
>>1-cdf('norm',362,300,50)
ans =
   0.1075
```

(b) $P(X < 290) = F(290)$

$= \text{cdf}(\text{'norm'}, 290, 300, 50)$

$= 0.4207.$

```
>> cdf('norm',290,300,50)
ans =
   0.4207
```

(c) $P(280 < X < 340) = F(340) - F(280)$

$= \text{cdf}(\text{'norm'}, 340, 300, 50) - \text{cdf}(\text{'norm'}, 280, 300, 50)$

$= 0.4436$

```
>> cdf('norm',340,300,50)-cdf('norm',280,300,50)
ans =
   0.4436
```

Standard Normal Distribution

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

If $\mu = 0, \sigma = 1$

$\boxed{f(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}, \ -\infty < x < \infty}$ → St.d. N.D

A normal random variable with

$$\mu = 0 \quad \text{and} \quad \sigma^2 = 1$$

is called a **standard normal random variable** and is denoted as Z.

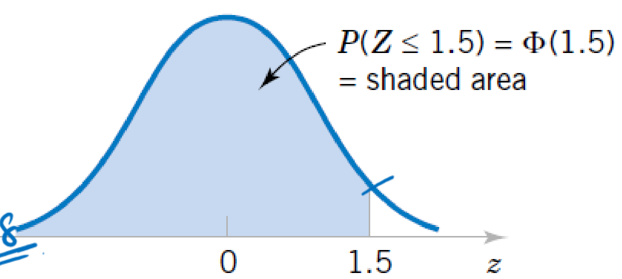
MGF of standard Normal Distribution

$$M_X(t) = e^{\frac{t^2}{2}}$$

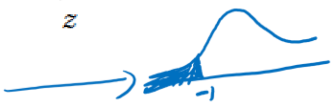
2. Given that Z has a standard normal distribution find the probability that Z assumes a value (a) greater than 1.5 (b) lesser than -1 (c) between -2 and 2.

$\mu_Z = 0$

$\sigma_Z^2 = 1 \Rightarrow \sigma_Z = 1$



(a) $P(Z > 1.5) = 1 - F(1.5)$
$= 1 - \text{cdf}(\text{'norm'}, 1.5, 0, 1) = 0.0668$



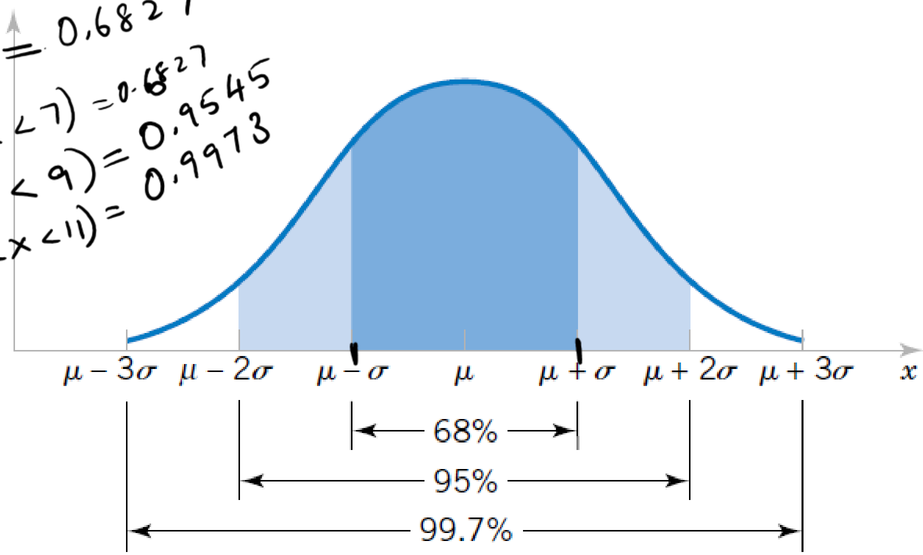
(b) $P(Z < -1) = F(-1) = \text{cdf}(\text{'norm'}, -1, 0, 1) = 0.1587$

(c) $P(-2 < x < 2) = F(2) - F(-2) = \text{cdf}(\text{'norm'}, 2, 0, 1) - \text{cdf}(\text{'norm'}, -2, 0, 1)$
$= 0.9545$

Normal Distribution

Eg:-1:- $\mu = 300$, $\sigma = 50$.
$P(\mu-\sigma < x < \mu+\sigma) = P(250 < x < 350) = 0.6827$

Eg:2:- $\mu = 5$, $\sigma = 2$
$P(\mu-\sigma < x < \mu+\sigma) = P(3 < x < 7) = 0.6827$
$P(\mu-2\sigma < x < \mu+2\sigma) = P(1 < x < 9) = 0.9545$
$P(\mu-3\sigma < x < \mu+3\sigma) = P(-1 < x < 11) = 0.9973$



$$P(\mu-\sigma < X < \mu+\sigma) = 0.6827$$

```
>> cdf('norm',350,300,50)-cdf('norm',250,300,50) % For μ=300, σ=50
ans =
   0.6827

>> cdf('norm',7,5,2)-cdf('norm',3,5,2) % For μ=5, σ=2
ans =
   0.6827

>> cdf('norm',-9,-15,6)-cdf('norm',-21,-15,6) % For μ=-15, σ=6
ans =
   0.6827
```

In the same way, verify
$P(\mu-2\sigma < X < \mu+2\sigma)$ **=0.95**
and
$P(\mu-3\sigma < X < \mu+3\sigma)$ **=0.997**
for various normal distributions.

Normal Distribution

1. Assume Z has a standard normal distribution. Find the value of 'a' if P(-a < Z < a) = 0.95.

$P(-a < Z < a) = 0.95$

$P(\mu - 2\sigma < Z < \mu + 2\sigma) = 0.95$

$-a = \mu - 2\sigma \; ; \quad a = \mu + 2\sigma$

$a = 0 + 2 \Rightarrow \boxed{a = 2}$

$P(-a < Z < a) = 0.95$

$2P(0 < Z < a) = 0.95 \Rightarrow P(0 < Z < a) = 0.475$

$\Rightarrow F(a) - F(0) = 0.475$

$\Rightarrow F(a) - 0.5 = 0.475 \Rightarrow F(a) = 0.975$

$a = \text{icdf}(\text{'norm'}, 0.975, 0, 1)$

$a = 2$

2. Assume X is a normal RV with mean 15 and variance 4. Find the value of 'b' if P(b-4<X<b) = 0.68.

$\mu = 15, \; \sigma = 2,$

$$\left.\begin{aligned} P(b-4 < X < b) &= 0.68 \\ P(\mu - \sigma < X < \mu + \sigma) &= 0.68 \end{aligned}\right\} \; b = 17$$

3. Assume X is a normal RV with mean 9 and variance 1. Find the value of 'c' if P(c-6<X<c) = 0.997.

$P(c-6 < X < c) = 0.997$

$P(\mu - 3\sigma < X < \mu + 3\sigma) = 0.997$

$\Rightarrow$

$\boxed{c = 12}$

Normal Distribution and Standard Normal Distribution

If X is a normal random variable with $E(X) = \mu$ and $V(X) = \sigma^2$, the random variable

$$Z = \frac{X - \mu}{\sigma}$$

is a normal random variable with $E(Z) = 0$ and $V(Z) = 1$. That is, Z is a standard normal random variable.

$E(Z) = E\left(\frac{X-\mu}{\sigma}\right) = \frac{1}{\sigma}\left[E(X) - E(\mu)\right] = \frac{1}{\sigma}\left[\mu - \mu\right] = 0$

$Var(Z) = Var\left(\frac{X-\mu}{\sigma}\right) = \frac{1}{\sigma^2}\left[Var(X) + Var(\mu)\right] = \frac{1}{\sigma^2}\left[\sigma^2 + 0\right] = 1$

- X is a normal RV with mean μ=5 and σ=2. So $Z = \frac{X-5}{2}$ is a standard normal RV.

  Verify this using MATLAB by finding P(X<6), P(X>3), P(4<X<7)

$P(X<6) = cdf('norm', 6, 5, 2) = 0.6915$

$P(X>3) = 1 - cdf('norm', 3, 5, 2) = 0.8413$

$P(4<X<7) = cdf('norm', 7, 5, 2) - cdf('norm', 4, 5, 2) = 0.5328$

$Z = \frac{X-\mu}{\sigma} = \frac{X-5}{2}$

$P(X<6) = P\left(\frac{X-\mu}{\sigma} < \frac{6-5}{2}\right) = P(Z < 0.5) = cdf('norm', 0.5, 0, 1) = 0.6915$

$P(X>3) = P(Z > -1) = 1 - F(-1) = 1 - cdf('norm', -1, 0, 1) = 0.8413$

$P(4<X<7) = P(-0.5 < Z < 1) = cdf('norm', 1, 0, 1) - cdf('norm', -0.5, 0, 1) = 0.5328$

Normal Distribution- Exercise

1. The mean and standard deviation of a certain group of 1000 high school grades that are normally distributed are 78% and 11% respectively. Find how many grades were above 90%?

   $\mu = 78, \quad \sigma = 11$

   $P(x > 90) = 1 - P(x \le 90) = 1 - F(90) = 1 - cdf('norm', 90, 78, 11) = 0.1377$

   $P(x>90) = 0.1377 \implies 13.77\%$ of grades are above $90\% \implies \frac{13.77}{100} \times 1000 \sim 138$ grades are above 90%

2. The fill volume of an automated filling machine used for filling cans of carbonated beverage is normally distributed with mean of 12.4 fluid ounces and a standard deviation of 0.1fluid ounce.
   a. What is the probability a fill volume is less than 12 fluid ounces?
   b. If all cans less than 12.1 or greater than 12.6 ounces are scrapped, what proportion of cans is scrapped?
3. The time it takes a cell to divide (called mitosis) is normally distributed with an average time of one hour and a standard deviation of 5 minutes.
   a. What is the probability that a cell divides in less than 45 minutes?
   b. What is the probability that it takes a cell more than 65 minutes to divide?
4. The time to microwave a bag of popcorn using the automatic setting can be treated as a RV having a normal distribution with standard deviation 10 seconds. If the probability is 0.8212 that the bag will take less than 282.5 seconds to pop, find the probabilities that
   (i) that it will take longer than 258.3 seconds to pop and
   (ii)that it will take on a value greater than 39.2 seconds.