Abstract

This thesis consider the normal distribution with its important appearance in many statistical fields of applications. Some mathematical and statistical properties of the distribution have been collected and illustrated with moments and higher moments. Six related theorems have been studied in the applications of this type of distribution.

The estimation manner and its properties have been illustrated throughout two methods (Moment and Maximum Likelihood methods) which are used to estimate the distribution parameters theoretically. Equality and properties of estimation have been studied throughout many well-known theorems.

Five methods to approximate the cumulative distribution function have been used namely: Trapezoidal, Simpson, Gaussian, Hit or Miss and Sample mean rules. The results of these rules have been compared in its behavior and error of approximation resulted from each method. The comparison shows clearly that the last method "Sample mean rule" is the best method among of five methods for approximating the solution for this type of functions. In addition to that the results of each method have been represented by curves line and numerical tables for helping in reading and comparing the results of each method with each other.

Finally four procedures for generating random varieties from normal distribution are discussed which are Box-Muller, Acceptance-Rejection, Central limit theorem and Tocher procedures and their efficiencies which are compared theoretically and practically by Monte Carlo simulation. The results of comparison shows that the Box-Muller procedure is the best one among three methods for this type of generation in sense of time consuming.

Acknowledgments

Praise is to Allah , the Cherisher and Sustainer of the Worlds who enabled me to achieve this research work .

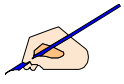
I am deeply indebted to my Supervisor ***Assist.Prof. Dr. Akram M. Al-Abood*** *for his time and patience in helping me to achieve this important goal in my life.*

I would like to thank the ***staff members of the Mathematics and Computer Applications Department of Al- Nahrain University*** *for giving me the chance to complete my graduate study.*

Thank are also express to several people who have helped me to edit this thesis: ***Dr. Akram M. Al-Abood*** *(Head of the Mathematics and Computer Applications Department),* ***Dr. Fadhil S. Fadhil.*** *and* ***Mr. Mushtaq Kareem.***

A special world of thanks goes to my family for their care, sacrifice, respect and love during my study.

Finally, to all my friends ……..……I present my thanks.



Israa Abdul Ameer Resen

June, 2007

Appendix

A

Approximation Programs

A1: Approximation by Trapezoidal Rule

```
program trapezod;
uses crt;
var
  j,i,a,n:integer;
  er,ex,b,x,h,sum,g,z,s:real;
  g1,f1,f2:text;
  function f(x:real):real;
  begin
    f:=exp((-1/2)*sqr(x));
  end;
  begin
 clrscr;
  assign(f1,'h:\new\ex.dat');
  reset(f1);
  assign(f2,'h:\new\x.dat');
  reset(f2);
  a:=0;
  b:=0;
    writeln('   x       g(x)        Erorr');
    writeln('------------------------------');
    j:=0;
   while not eof(f2) do
    begin
        readln(f2,b);
        h:=(b-a)/10;
        sum:=0;
        for i:=1 to 9 do
          begin
```

```
        x:=a+i*h;
        sum:=sum+2*f(x);
      end;
    g:=(h/2)*(f(a)+sum+f(b));
    z:=(1/sqrt(2*pi))*g;
    s:=(1/2)+z;
    readln(f1,ex);
    er:=abs(ex-s);
    writeln(b:6:3,'  ',s:15:8,'    ',er:15:8);
    {b:=b+0.05;}
    j:=j+1;
    if j=33 then begin j:=0; readln; end;
  end;
  readln;
end.
```

A2: Approximation by Simpson Rule

```
program simpson;
uses crt;
var
  j,i,a,n:integer; b,x,h,sum1,sum2,g,z,s,ex,er:real;
  g1,f1:text;
  function f(x:real):real;
  begin
    f:=exp((-1/2)*sqr(x));
  end;
  begin
  clrscr;
    a:=0;
    assign(f1,'h:\new\ex.dat');
    reset(f1);
    b:=0;
    writeln('   x       g(x)        Erorr');
    writeln('------------------------------');
     j:=0;
    while not eof(f1) do
    begin
    h:=(b-a)/10;
    sum1:=0; sum2:=0;
    for i:=1 to 9 do
```

```
   begin
     x:=a+i*h;
     if i mod 2 =0  then
     sum1:=sum1+2*f(x)
     else sum2:=sum2+4*f(X)
     end;
     g:=(h/3)*(f(a)+sum1+sum2+f(b));
     z:=(1/sqrt(2*pi))*g;
     s:=(1/2)+z;
     readln(f1,ex);
     er:=abs(ex-s);
     writeln(b:6:2,'  ',s:6:3,'    ',er:15:8);
      b:=b+0.05;
      j:=j+1;
      if j=5 then begin readln; j:=0; end;
      end;
      readln;
end.
```

A3: Approximation by Gaussian Quadrature Rule

```
program gauss;
{uses crt;}
type
c=array[1..100]of real;
var
j,i,a,n:integer;
b,sum,v,g,z,ex,er:real;
x,w:c;
f1,f2:text;
function f(x:real):real;
begin
  f:=(1/(sqrt(2*pi)))*exp(-sqr(x)/2);
end;
  begin
  x[1]:=0.577350269189626;
  x[2]:=-x[1];
  w[1]:=1.000000000000000;
  w[2]:=w[1];
  assign(f1,'h:\new\ex.dat');
  reset(f1);
```

```
  assign(f2,'h:\new\x.dat');
  reset(f2);
  a:=0;
  b:=0;
  writeln('   x       g(x)        Erorr');
  writeln('------------------------------');
  n:=2;
  j:=0;
  while not eof(f2) do
  begin
  sum:=0;
  readln(f2,b);
  for i:=1 to n do
     begin
     v:=((b-a)/2)*x[i]+(b+a)/2;
     sum:=sum+f(v)*w[i];
     g:=((b-a)/2)*sum;
     z:=0.5+g;
    end;
  readln(f1,ex);
  er:=abs(ex-z);
  writeln(b:6:2,'  ',z:15:8,'    ',er:15:8);
  {b:=b+0.05;}
  j:=j+1;
  if j=30 then begin j:=0; readln; end;
  end;
end.
```

A4: Approximation by Hit or Miss Rule

```
program HitMiss;
var
        N,i,j,NH:integer;
        U:array [1..3500] of real;
        U1:array [1..3500] of real;
        xa:array [1..3500] of real;
        g1,z,a,b,x,s,ex,er:real;
        f1,f2:text;
function g(x:real):real;
begin
    g:=exp((-1/2)*sqr(x));
```

```
end;
begin
N:=40;
a:=0;

randomize;
  assign(f1,'c:\tp\bin\ex.dat');
  reset(f1);
  assign(f2,'c:\tp\bin\x.dat');
  reset(f2);
  writeln('   x       g(x)        Erorr');
  writeln('-------------------------------');
while not eof(f2) do
  begin
      readln(f2,b);
      readln(f1,ex);
      {writeln(pi:10:5);}
      NH:=0;
       for i:=1 to N do
          begin
            U[i]:=random;
            xa[i]:=a+U[i]*(b-a);
          end;
       for i:=1 to N do
            U1[i]:=random;
      for i:=1 to N do
             if g(xa[i])>(g(b)*U1[i]) then
                 NH:=NH+1;
      g1:=g(b)*((b-a)*NH)/N;
      z:=(1/sqrt(2*pi))*g1;
      s:=(1/2)+z;
  er:=abs(ex-s);
  writeln(b:6:2,'  ',s:10:8,'    ',er:10:8);
  j:=j+1;
  if j=30 then begin j:=0; readln; end;
  end;
end.
```

A๐: Approximation by Sample Mean Rule

```
program Samplemean;
```

```
var
        N,i,j,NH:integer;
        U:array [1..3500] of real;
        xa:array [1..3500] of real;
        z,g1,a,b,x,s,sum,ex,er:real;
        f1,f2:text;
function g(x:real):real;
begin
    g:=exp((-1/2)*sqr(x));
end;
begin
N:=300;
a:=0;
randomize;
  assign(f1,'c:\tp\bin\ex.dat');
  reset(f1);
  assign(f2,'c:\tp\bin\x.dat');
  reset(f2);
  writeln('   x       g(x)        Erorr');
  writeln('-------------------------------');
while not eof(f2) do
  begin
      readln(f2,b);
      readln(f1,ex);
       sum:=0;
       for i:=1 to N do
           begin
             U[i]:=random;
             xa[i]:=a+U[i]*(b-a);
             sum:=sum+g(xa[i]);
           end;
          g1:=(b-a)*(1/N) * sum;
          z:=(1/sqrt(2*pi))*g1;
          s:=(1/2)+z;
      er:=abs(ex-s);
      writeln(b:6:2,'  ',s:10:8,'    ',er:10:8);
      j:=j+1;
      if j=30 then begin j:=0; readln; end;
  end;
end.
```

Appendix

B

Estimation Programs

B1: N-1 Procedure

```
program acc_rej;
uses crt;
type
 a=array[1..2000]of real;
var
  grd,grm,ercd,i,nx,ny,r,n:integer;
  x:a; u,u1,u2,sumx,sumy,sumr,xb,c,v,sg,y,y2,avr:real;
begin
clrscr;
  randomize;
  n:=0;
  while n<1000 do
  begin
    n:=n+100;
  sumx:=0; sumy:=0; sumr:=0;
   for i:=1 to n do
  begin
    r:=0;
    repeat
    r:=r+1;
    u1:=random;
    u2:=random;
    y:=-ln(u2);
    y2:=exp((-1/2)*sqr(y-1))
    until not(u1>y2);
    u:=random;
    if u1<=1/2 then x[i]:=y
    else x[i]:=-y;
```

```
    sumx:=sumx+x[i];
    sumy:=sumy+sqr(x[i]);
    sumr:=sumr+r;
    end;
    xb:=sumx/n;
    sg:=(sumy-n*sqr(xb))/(n-1);
    avr:=sumr/n;
    c:=1/avr;
    writeln(n:3,'Xb=',xb:7:3,'    Segma=  ',sg:7:3,avr:7:3,c:7:3);
    end;
    readln;
end.
```

B2: N-2 Procedure

```
program BoxMuller;
uses crt,dos;
type
a=array[1..1000]of real;
var
  h2,m2,s2,ms2:word;
  i,n,k:integer;
  x,y,v:a; u1,u2,z,sum,sumx,xb,sg,x1,y1,sumt,avr,c:real;
  begin
  clrscr;
  randomize;
  n:=0;
  sumt:=0;
  while n<100 do
  begin
  settime(0,0,0,0);
  n:=n+10;
  for i:=1 to n do
 begin
    u1:=random;
    u2:=random;
    x[i]:=sqrt(-2*ln(u1))*cos(2*pi*u2);
    y[i]:=sqrt(-2*ln(u1))*sin(2*pi*u2);
    writeln(x[i],y[i]);
    end;
   gettime(h2,m2,s2,ms2);
```

```
  sumt:=sumt+s2*100+ms2;
  writeln(n:3,s2:3,ms2:3);
  readln;
  end;
  avr:=sumt/10;
  c:=1/avr;
  writeln(avr:6:3,c:6:3);
  for i:=1 to 200 do
   begin
   if i<=100 then
   v[i]:=x[i]
   else v[i]:=y[i-100];
   z:=(1/sqrt(2*pi))*exp((-1/2)*sqr(v[i]));
   end;
  n:=0;
  while n<200 do
  begin
  n:=n+20;
  sum:=0; sumx:=0;
  for i:=1 to n do
  begin
  sum:=sum+v[i];
  sumx:=sumx+sqr(v[i]);
  end;
  xb:=sum/n;
  sg:=(sumx-n*sqr(xb))/(n-1);
  writeln;
  writeln('Xb=',xb:7:3,'    Segma=  ',sg:7:3);
  end;
  readln;
  readln;
 end.
```

B3: N-3 Procedure

```
program Central;
uses crt,dos;
type
 a=array[1..1000]of real;
var
h2,s2,m2,ms2:word;
```

```
i,j,nx,ny,n:integer;
  x:a; u,u1,u2,sumx,sumy,sumu,xb,v,vb,ub,y,y2,avr:real;
 begin
  clrscr;
  randomize;
  n:=0;
  while n<200 do
  begin
  settime(0,0,0,0);
  n:=n+20;
  sumx:=0; sumy:=0;
   for i:=1 to n do
  begin
    sumu:=0;
    for j:=1 to 10 do
    begin
      u1:=random;
      sumu:=sumu+u1;
    end;
    ub:=sumu/10;
    x[i]:=sqrt(12*10)*(ub-0.5);
    writeln(x[i]:9:3);
    sumx:=sumx+x[i];
    sumy:=sumy+sqr(x[i]);
    end;
    xb:=sumx/n;
    vb:=(sumy-n*sqr(xb))/(n-1);
    gettime(h2,m2,s2,ms2);
    writeln(n:3,s2:3,ms2:3,'xb',xb:9:4,'vb',vb:9:4);
    readln;
    end;
    readln;
end.
```

B4: N-4 Procedure

```
program tocher;
uses crt,dos;
type
 a=array[1..1000]of real;
var
```

```
 h2,m2,s2,ms2:word;
 i,nx,ny,n:integer;
 x:a; u1,u2,sumx,sumy,xb,v1,v2,vb,y,k:real;
begin
 clrscr;
 randomize;
 n:=0;
 while n<200 do
 begin
 n:=n+20;
 settime(0,0,0,0);
 sumx:=0; sumy:=0;
  for i:=1 to n do
 begin
    u1:=random;
    u2:=random;
    y:=0.5*sqrt(pi/2)*ln((1+u1)/(1-u2));
    if u2<=0.5 then x[i]:=y
    else x[i]:=-y;
  writeln(x[i]:9:3);
  sumx:=sumx+x[i];
  sumy:=sumy+sqr(x[i]);
  end;
  xb:=sumx/n;
  vb:=(sumy-n*sqr(xb))/(n-1);
  gettime(h2,m2,s2,ms2);
  writeln(n:3,s2:3,ms2:3,'xb',xb:9:4,'vb',vb:9:4);
  readln;
  end;
  readln;
 end.
```

CHAPTER

1

The Normal Distribution

1.1 Introduction

Basic mathematical and statistical properties of normal distribution are discussed in this chapter which is involve six sections as follows. Section 1.2 basic properties of normal distribution are given, while in section 1.3 we illustrated the moments of normal distribution. Some related theorems are given in section 1.4, while in section 1.5 two methods of parameters estimation (Moment and Maximum Likelihood) are theoretically discussed. Finally, section 1.6 equality of estimators are studied to reach to the best estimator for the unknown parameters.

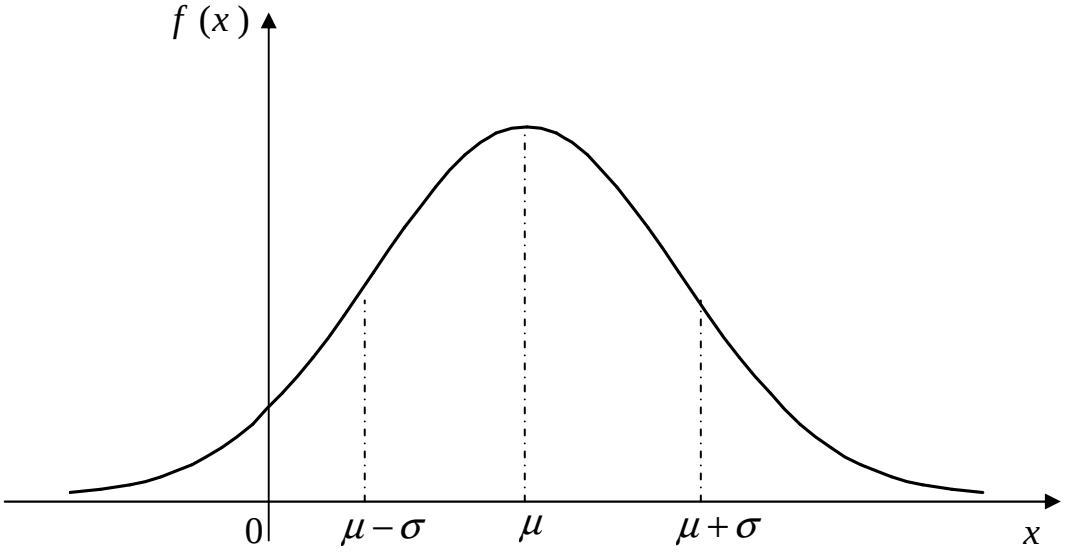
1.2 Basic Properties of Normal Distribution

1.2.1 Definition [26]

A continuous r.v. X is said to have a normal distn., denoted by $X \sim N(\mu,\sigma^2)$, if X has p.d.f.

$$f(x)=\frac{1}{\sqrt{2\pi}\sigma}e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2},\ -\infty<x<\infty \qquad (1.1)$$

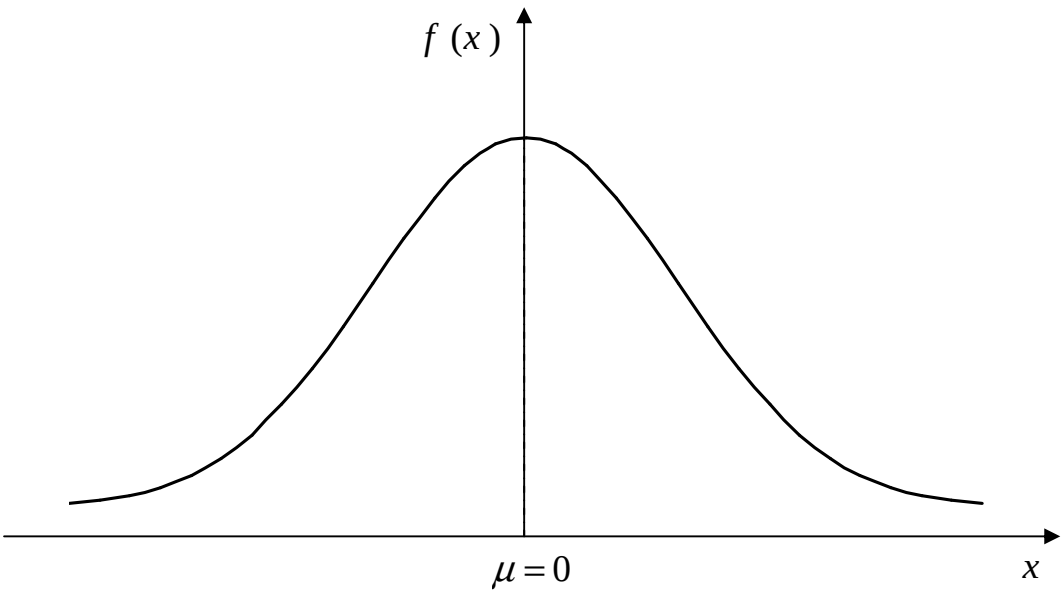
Where $-\infty < x < \infty$, $\sigma^2 > 0$ are known as location and scale parameters respectively with $e = 2.7183$ and $\pi = 3.1416$. The graphical representation of eq.(1.1) is



Figure(1.1): The Normal Curve

When $\mu = 0$ and $\sigma^2 = 1$, then $X \sim N(0,1)$ is said to have a standard normal distn. as shown in Figure(1.2) and the p.d.f. of eq.(1.1) reduce to

$$f(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}x^2}, \quad -\infty < x < \infty \tag{1.2}$$



Figure(1.2): The Standard Normal Curve

To verify that the function $f(x)$ of eq.(1.1) is valid p.d.f., we first note that $f(x)>0$ for all $x\in(-\infty,\infty)$.

We also need to show that integral of $f(x)$ over $-\infty<x<\infty$ is unity. Viz

Consider the integral

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

Make a simple transformation by setting $y=\dfrac{x-\mu}{\sigma}$ with $dx=\sigma dy$, we have

$$I=\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{-\frac{1}{2}y^2}dy$$

Now,

$$I^2=\left\{\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{-\frac{1}{2}x^2}dx\right\}\left\{\frac{1}{\sqrt{2\pi}}\int_{-\infty}^{\infty} e^{-\frac{1}{2}y^2}dy\right\}$$

$$=\frac{1}{2\pi}\int_{-\infty}^{\infty}\int_{-\infty}^{\infty} e^{-\frac{1}{2}\left(x^2+y^2\right)}dxdy$$

Changing to polar coordinates by setting

$x=r\cos\theta$, $y=r\sin\theta$, where $0<r<\infty$, $0<\theta<2\pi$

The Jacobian of transformation

$$J=\frac{\partial(x,y)}{\partial(r,\theta)}=\begin{vmatrix}\dfrac{\partial x}{\partial r} & \dfrac{\partial x}{\partial\theta}\\ \dfrac{\partial y}{\partial r} & \dfrac{\partial y}{\partial\theta}\end{vmatrix}=\begin{vmatrix}\cos\theta & -r\sin\theta\\ \sin\theta & r\cos\theta\end{vmatrix}=r\cos^2\theta+r\sin^2\theta=r$$

and that implies $dxdy=rdrd\theta$, therefore

$$I^2=\frac{1}{2\pi}\int_{\theta=0}^{2\pi}\int_{r=0}^{\infty} re^{-\frac{1}{2}r^2}drd\theta=-\frac{1}{2\pi}\int_{\theta=0}^{2\pi} e^{-\frac{1}{2}r^2}\Big|_0^{\infty}d\theta=\frac{1}{2\pi}\int_{\theta=0}^{2\pi}d\theta=1$$

since $I^2 > 0$, it follow that $I = 1$.

1.2.2 Properties of Normal Curve

The graph of Figure (1.1) represents the curve of the normal p.d.f. $f(x)$ given by eq.(1.1), where curve has the following properties:

1. The curve is symmetric about a vertical axis through the mean μ.
2. The curve has the line $y = 0$ (x-axis) as a horizontal asymptote.
3. The curve increasing for $-\infty < x < \mu$ and decreasing for $\mu < x < \infty$.
4. The curve has maximum point at $x = \mu$.
5. The curve have points of inflection at $x = \mu \pm \sigma$.
6. The curve concave upward for $-\infty < x < \mu - \sigma$, $\mu + \sigma < x < \infty$ and concave downward for $\mu - \sigma < x < \mu + \sigma$.
7. The total area under the curve and above the horizontal axis is equal to one.

1.2.3 Relation Between $N(\mu,\sigma^2)$ and $N(0,1)$

The relation between the normal and the standard normal distributions is given by the following theorem

Theorem [14]

The r.v. $X \sim N(\mu,\sigma^2)$ iff the r.v. $Y = (\frac{x-\mu}{\sigma}) \sim N(0,1)$.

Proof:

$\Rightarrow$ Let $X \sim N(\mu,\sigma^2)$, then X has p.d.f.

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

the function $y = \left(\frac{x-\mu}{\sigma}\right)$ define one-to-one transformation that maps

the space $A = \{x : -\infty < x < \infty\}$ onto $B = \{y : -\infty < y < \infty\}$ with inverses $x = \sigma y + \mu$ and the Jacobian is $J = \frac{dx}{dy} = \sigma$

Then, the p.d.f. of r.v. y, say $g(y)$, is

$$g(y) = f(y)|J| = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2}y^2} \sigma = \frac{1}{\sqrt{2\pi}} e^{-\frac{1}{2}y^2}, \quad -\infty < y < \infty$$

which is the p.d.f. of the theorem is obvious.

$\Leftarrow$ The converse of the theorem is obvious.

1.2.4 The Cumulative Distribution Function

The c.d.f. of r.v. $X \sim N(\mu, \sigma^2)$ defined as

$$\Phi(x) = pr(X \le x) = \int_{-\infty}^{x} f(t)dt$$

$$= \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{1}{2}\left(\frac{t-\mu}{\sigma}\right)^2} dt \qquad (1.3)$$

The integral side of eq.(1.3) can not be evaluated analytically because the derivative of $e^{-\frac{1}{2}\left(\frac{t-\mu}{\sigma}\right)^2}$ is not available under the integral sign therefore, table(1.1) below is prepared in most of statistical books which evaluate $pr(X \le x) = pr\left(\frac{X-\mu}{\sigma} \le \frac{x-\mu}{\sigma}\right) = pr\left(Y \le \frac{x-\mu}{\sigma}\right)$

where $y = \frac{x-\mu}{\sigma} \sim N(0,1)$. In this case the c.d.f of Y is

$$\Phi(y) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{y} e^{-\frac{1}{2}t^2} dt \qquad (1.4)$$

More discussion is given in chapter two about several suggested procedures for approximating the integral side of eq.(1.4).

Table (1.1): The Normal Distribution of parameter $\mu = 0$, $\sigma^2 = 1$

$$\Phi(y) = pr(Y \le y) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{y} e^{-\frac{1}{2}t^2} dt$$

$$\Phi(-y) = 1 - \Phi(y)$$

y	Φ(y)	y	Φ(y)	y	Φ(y)
0.00	0.500	1.10	0.864	2.05	0.980
0.05	0.520	1.15	0.875	2.10	0.982
0.10	0.540	1.20	0.885	2.15	0.984
0.15	0.560	1.25	0.894	2.20	0.986
0.20	0.579	1.282	0.900	2.25	0.988
0.25	0.599	1.30	0.903	2.30	0.989
0.30	0.618	1.35	0.911	2.326	0.990
0.35	0.637	1.40	0.919	2.35	0.991
0.40	0.655	1.45	0.926	2.40	0.992
0.45	0.674	1.50	0.933	2.45	0.993
0.50	0.691	1.55	0.939	2.50	0.994
0.55	0.709	1.60	0.945	2.55	0.995
0.60	0.726	1.645	0.950	2.576	0.995
0.65	0.742	1.65	0.951	2.60	0.995
0.70	0.758	1.70	0.955	2.65	0.996
0.75	0.773	1.75	0.960	2.70	0.997
0.80	0.788	1.80	0.964	2.75	0.997
0.85	0.802	1.85	0.968	2.80	0.997
0.90	0.816	1.90	0.971	2.85	0.998
0.95	0.829	1.95	0.974	2.90	0.998
1.00	0.841	1.960	0.975	2.95	0.998
1.05	0.853	2.00	0.977	3.00	0.999

1.3 The Moments of $N(\mu,\sigma^2)$ *[26]*

The m.g.f. of r.v. $X \sim N(\mu,\sigma^2)$ is defined by

$$M_X(t) = E(e^{tX}) = \int_{-\infty}^{\infty} e^{tx} f(x)dx = \int_{-\infty}^{\infty} \frac{1}{\sqrt{2\pi}\sigma} e^{tx} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} dx$$

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

Consider the exponent

$$(x-\mu)^2 - 2\sigma^2 tx = x^2 - 2\mu x + \mu^2 - 2\sigma^2 tx$$

$$= x^2 - 2(\mu+\sigma^2 t)x + (\mu+\sigma^2 t)^2 + \mu^2 - (\mu+\sigma^2 t)^2$$

$$= \left[x - (\mu+\sigma^2 t)\right]^2 + \mu^2 - \mu^2 - 2\mu\sigma^2 t - \sigma^4 t^2$$

$$= \left[x - \left(\mu+\sigma^2 t\right)\right]^2 - 2\sigma^2\left(\mu t + \frac{1}{2}\sigma^2 t^2\right)$$

Therefore,

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

Setting $y = \dfrac{x - \left(\mu+\sigma^2 t\right)}{\sigma}$ that implies

$x = \sigma y + \left(\mu+\sigma^2 t\right)$ and $dx = \sigma dy$

So,

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

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

The integral side of the above equation is unity

Thus,

$$M_X(t) = e^{\mu t + \frac{1}{2}\sigma^2 t^2} \tag{1.5}$$

According to the theorem of section 1.2.3, the r.v. $Y = \frac{X - \mu}{\sigma} \sim N(0,1)$ has

m.g.f.

$$M_Y(t) = e^{\frac{1}{2}t^2} \tag{1.6}$$

Maclaurian series expansion of $e^{\frac{1}{2}t^2}$ leads to

$$M_Y(t) = e^{\frac{1}{2}t^2} = 1 + \frac{t^2}{2} + \frac{1}{2!}\left(\frac{t^2}{2}\right)^2 + \frac{1}{3!}\left(\frac{t^2}{2}\right)^3 + \ldots + \frac{1}{r!}\left(\frac{t^2}{2}\right)^r + \ldots$$

$$= 1 + 1 \cdot \frac{t^2}{2!} + 1 \cdot 3\frac{t^4}{4!} + 1 \cdot 3 \cdot 5\frac{t^6}{6!} + \ldots + 1 \cdot 3 \cdot 5 \ldots (2r-1)\frac{t^{2r}}{(2r)!} + \ldots$$

So $E(Y^{2r})$ is the coefficient of $\frac{t^{2r}}{(2r)!}$. That is

$$E(Y^{2r}) = 1 \cdot 3 \cdot 5 \ldots (2r-1)$$

$$= 1 \cdot 3 \cdot 5 \ldots (2r-1)\frac{2 \cdot 4 \cdot 6 \ldots 2r}{2 \cdot 4 \cdot 6 \ldots 2r} = \frac{1 \cdot 2 \cdot 3 \ldots 2r}{2^r (1 \cdot 2 \ldots r)}$$

$$E(Y^{2r}) = \frac{(2r)!}{2^r r!}, \; r = 1,2,3,\ldots \tag{1.7}$$

and

$$E\left(Y^{2r-1}\right) = 0, \; r = 1,2,3,\ldots \tag{1.8}$$

1.3.1 Central Moments of $N(\mu,\sigma^2)$

Setting $Y=\frac{X-\mu}{\sigma}$, in eq.(1.7) and eq.(1.8), we have

$E\left[\left(\frac{X-\mu}{\sigma}\right)^{2r}\right]=\frac{(2r)!}{2^r r!}$ leads to

$$E\left[(X-\mu)^{2r}\right]=\frac{(2r)!}{2^r r!}\sigma^{2r},\ r=1,2,3,\ldots \tag{1.9}$$

and

$$E\left[(X-\mu)^{2r-1}\right]=0,\ r=1,2,3,\ldots \tag{1.10}$$

(I) Mean

Use of eq.(1.10) with $r=1$, we have

$E(X-\mu)=0$ and that implies $E(X)=\mu$, where μ is called the mean of r.v. X (or distn.). It is a measure of central tendency.

(II) Variance

Use of eq.(1.9) with $r=1$, we have

$E\left[(X-\mu)^2\right]=\sigma^2$, where σ^2 is called the variance of r.v. X (or distn.). It is a measure of dispersion.

(III) Coefficient of Skewness

$\gamma_1=\frac{E\left[(X-\mu)^3\right]}{\left(\sigma^2\right)^{3/2}}$ is called the coefficient of skewness. It is a measure of departure from symmetry of frequency curve.

Use of eq. (1.10) with $r=2$, we have $\gamma_1=0$.

(IV) Coefficient of Kurtosis

$\gamma_2 = \dfrac{\mathrm{E}\left[(\mathrm{X}-\mu)^4\right]}{\left(\sigma^2\right)^2} - 3$ is called the coefficient of kurtosis. It is a measure of degree of flatting of frequency curve.

Use of eq. (1.9) with $r = 2$, we have

$$\gamma_2 = \frac{3\sigma^4}{\sigma^4} - 3 = 0.$$

1.3.2 Other Central Moments

(I) Mode

A mode of a distn. is defined to be the value of x which maximize the p.d.f. $f(x)$. For continuous distributions, the mode x is the solution of $\dfrac{df(x)}{dx} = 0$ and $\dfrac{d^2 f(x)}{dx^2} < 0$. A mode is a measure of location.

For $\mathrm{N}(\mu,\sigma^2)$ distn. with p.d.f. of eq.(1.1), we have

$$f'(x) = -\frac{1}{\sqrt{2\pi}}\sigma^{-3}(x-\mu)e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$$

$$f'(x) = 0 \Rightarrow x - \mu = 0 \Rightarrow x = \mu$$

$$f''(x) = -\frac{1}{\sqrt{2\pi}}\sigma^{-3}\left[-\sigma^{-2}(x-\mu)^2 + 1\right]e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$$

$$f''(\mu) = -\frac{1}{\sqrt{2\pi}}\sigma^{-3} < 0$$

Thus, the distn. mode is μ.

(II) Median

A median of a distn. is defined to be the value of x such that the c.d.f. $F(x) = \dfrac{1}{2}$.

The median is a measure of location.

For $N(\mu,\sigma^2)$ with c.d.f. of eq.(1.3), we have

$$\frac{1}{2} = pr(X \le x) = pr\left(\frac{X-\mu}{\sigma} \le \frac{x-\mu}{\sigma}\right) = pr\left(Y \le \frac{x-\mu}{\sigma}\right)$$

where $Y \sim N(0,1)$ and that implies from table(1.1) $\frac{x-\mu}{\sigma} = 0 \Rightarrow x = \mu$

thus, the distn. median is μ.

1.4 Some Related Theorems [14]

Theorem (1.4.1)

If the r.v. $X \sim N(0,1)$ then the r.v. $Y = X^2 \sim \chi^2(1)$, where the p.d.f. of y is $g(y) = \frac{1}{\Gamma(\frac{1}{2})2^{\frac{1}{2}}} x^{\frac{1}{2}-1} e^{-\frac{x}{2}}$, $0 < x < \infty$

$= 0, e.w.$.

Theorem (1.4.2)

If $X_1, X_2, \ldots, X_n$ are indep. r.v.'s with $X_i \sim N(\mu_i, \sigma_i^2)$, $i = 1,2,\ldots,n$,

then the r.v. $Y = \sum_{i=1}^{n} k_i X_i \sim N\left[\sum_{i=1}^{n} k_i \mu_i, \sum_{i=1}^{n} k_i{}^2 \sigma_i{}^2\right]$.

Theorem (1.4.3)

If $X_1, X_2, \ldots, X_n$ are indep. r.v.'s with $X_i \sim \chi^2(r_i)$, $i = 1,2,\ldots,n$,

then the r.v. $Y = \sum_{i=1}^{n} X_i \sim \chi^2\left(\sum_{i=1}^{n} r_i\right)$.

Theorem (1.4.4)

If $X_1, X_2, \ldots, X_n$ is a r.s. of size n from $N(\mu,\sigma^2)$, then the r.v.

$Y = \sum_{i=1}^{n} \left(\frac{X_i - \mu}{\sigma} \right)^2 \sim \chi^2(n)$ where the p.d.f. of y is

$$g(y) = \frac{1}{\Gamma(\frac{n}{2}) 2^{\frac{n}{2}}} x^{\frac{n}{2}-1} e^{-\frac{x}{2}}, \; 0 < x < \infty, \; n = 1,2,3,\ldots$$

$$= 0, e.w.$$

Theorem (1.4.5)

If the r.v.'s X and Y are stochastically independent with $X \sim N(0,1)$ and $Y \sim \chi^2(r)$. Then, the r.v. $T = \frac{X}{\sqrt{\frac{Y}{r}}} \sim t(r)$.

1.4.1 Independence of $\bar{X}$ and $\frac{(n-1)S^2}{\sigma^2}$ [26]

For normal case, there are many techniques can be found throughout the literature providing the independency of the statistics $\bar{X}$ and $\frac{(n-1)S^2}{\sigma^2}$ [26].

To the best of our knowledge the following approach seem to be knew.

Let $X_1, X_2, \ldots, X_n$ be a r.s. of size n from $N(\mu, \sigma^2)$ and let $\bar{X} = \frac{1}{n} \sum_{i=1}^{n} X_i$ and $S^2 = \frac{1}{n-1} \sum_{i=1}^{n} \left(X_i - \bar{X} \right)^2$ be the sample mean and sample variance respectively.

The aims are to show that $\bar{X}$ and $\frac{(n-1)S^2}{\sigma^2}$ are stochastically independent.

The joint p.d.f. of the sets $\{X_i\}$ is

$$f(\underset{\sim}{x}) = \prod_{i=1}^{n} f(x_i) = (2\pi)^{-\frac{n}{2}} \sigma^{-n} \exp\left[-\frac{1}{2} \sum_{i=1}^{n} \left(\frac{x_i - \mu}{\sigma} \right)^2 \right], -\infty < x_i < \infty$$

$$i = 1,2,\ldots,n \tag{1.11}$$

If we consider the transformation $W_i = \dfrac{X_i - \mu}{\sigma}$, $i = 1,2,\ldots,n$, then according to the theorem of section 1.2.3, the r.v. W_i □ N(0,1) and the joint p.d.f. of r.v.'s $W_1, W_2, \ldots, W_n$ is

$$g(w_1, w_2, \ldots, w_n) = (2\pi)^{-\frac{n}{2}} \exp\left[-\frac{1}{2} \sum_{i=1}^{n} w_i \right], \ -\infty < w_i < \infty, \ i = 1,2,\ldots,n \tag{1.12}$$

Now,

Let $\overline{W} = \left(\dfrac{\overline{X} - \mu}{\sigma} \right)$ and $\sum_{i=1}^{n} \left(W_i - \overline{W} \right)^2 = \sum_{i=1}^{n} \left(\dfrac{X_i - \overline{X}}{\sigma} \right)^2 = \dfrac{(n-1)S^2}{\sigma^2}$

Consider the transformation

$$\left.\begin{aligned}
y_1 &= \frac{1}{\sqrt{n}} W_1 + \frac{1}{\sqrt{n}} W_2 + \ldots + \frac{1}{\sqrt{n}} W_n \\
y_2 &= \frac{1}{\sqrt{2 \cdot 1}} W_1 - \frac{1}{\sqrt{2 \cdot 1}} W_2 \\
y_3 &= \frac{1}{\sqrt{3 \cdot 2}} W_1 + \frac{1}{\sqrt{3 \cdot 2}} W_2 - \frac{2}{\sqrt{3 \cdot 2}} W_3 \\
y_4 &= \frac{1}{\sqrt{4 \cdot 3}} W_1 + \frac{1}{\sqrt{4 \cdot 3}} W_2 + \frac{1}{\sqrt{4 \cdot 3}} W_3 - \frac{3}{\sqrt{4 \cdot 3}} W_4 \\
&\vdots \\
&\vdots \\
y_n &= \frac{1}{\sqrt{n(n-1)}} W_1 + \frac{1}{\sqrt{n(n-1)}} W_2 + \ldots + \frac{1}{\sqrt{n(n-1)}} W_{n-1} - \frac{(n-1)}{\sqrt{n(n-1)}} W_n
\end{aligned}\right\} \tag{1.13}$$

The system of eqs.(1.13) can be written in a matrix form as

$\underset{\sim}{Y} = \underset{\sim}{A} \underset{\sim}{W}$, where

$\underset{\sim}{Y}^{\mathrm{T}} = (y_1, y_2, \ldots, y_n)$, $\underset{\sim}{\overline{W}} = (W_1, W_2, \ldots, W_n)$ and $\underset{\sim}{\mathrm{A}} = \left(a_{ij}\right)_{n\times n}$ with

$$a_{1i} = \frac{1}{\sqrt{n}},\ a_{ii} = \frac{-(i-1)}{\sqrt{i(i-1)}},\ i = 2,3,\ldots,n$$

$$a_{ij} = \frac{1}{\sqrt{i(i-1)}},\ i = 2,3,\ldots,n,\ j = 1,2,\ldots,(i-1)$$

The Jacobian of this transformation is J, where

$$\frac{1}{J} = \frac{\partial(y_1, y_2, \ldots, y_n)}{\partial(w_1, w_2, \ldots, w_n)}$$

$$= \begin{vmatrix} \frac{1}{\sqrt{n}} & \frac{1}{\sqrt{n}} & \frac{1}{\sqrt{n}} & \cdots & \cdots & \cdots & \frac{1}{\sqrt{n}} \\ \frac{1}{\sqrt{2\cdot 1}} & -\frac{1}{\sqrt{2\cdot 1}} & 0 & 0 & \cdots & \cdots & 0 \\ \frac{1}{\sqrt{3\cdot 2}} & \frac{1}{\sqrt{3\cdot 2}} & -\frac{2}{\sqrt{3\cdot 2}} & 0 & \cdots & \cdots & 0 \\ \frac{1}{\sqrt{4\cdot 3}} & \frac{1}{\sqrt{4\cdot 3}} & \frac{1}{\sqrt{4\cdot 3}} & -\frac{3}{\sqrt{4\cdot 3}} & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & & \vdots \\ \vdots & \vdots & \vdots & \vdots & & \ddots & \vdots \\ \frac{1}{\sqrt{n(n-1)}} & \frac{1}{\sqrt{n(n-1)}} & \frac{1}{\sqrt{n(n-1)}} & \frac{1}{\sqrt{n(n-1)}} & \cdots & \cdots & \frac{-(n-1)}{\sqrt{n(n-1)}} \end{vmatrix}$$

$$= \frac{1}{\sqrt{n}}\frac{1}{\sqrt{2\cdot 1}}\frac{1}{\sqrt{3\cdot 2}}\frac{1}{\sqrt{4\cdot 3}}\cdots\frac{1}{\sqrt{n(n-1)}} \cdot$$

$$\begin{vmatrix} 1 & 1 & 1 & 1 & 1 & \cdots & \cdots & \cdots & 1 \\ 1 & -1 & 0 & 0 & 0 & \cdots & \cdots & \cdots & 0 \\ 1 & 1 & -2 & 0 & 0 & \cdots & \cdots & \cdots & 0 \\ 1 & 1 & 1 & -3 & 0 & \cdots & \cdots & \cdots & 0 \\ 1 & 1 & 1 & 1 & -4 & 0 & \cdots & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & & & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & & \ddots & & \vdots \\ 1 & 1 & 1 & 1 & 1 & \cdots & \cdots & -(n-2) & 0 \\ 1 & 1 & 1 & 1 & 1 & 1 & \cdots & 1 & -(n-1) \end{vmatrix}$$

Multiply the 1^{st} row by (-1) and add to the i^{th} row $(i = 2,3,\ldots,n)$, we have

$$= k \begin{vmatrix} 1 & 1 & 1 & 1 & 1 & 1 & \ldots & \ldots & 1 & 1 \\ 0 & -2 & -1 & -1 & -1 & -1 & \ldots & \ldots & -1 & -1 \\ 0 & 0 & -3 & -1 & -1 & -1 & \ldots & \ldots & -1 & -1 \\ 0 & 0 & 0 & -4 & -1 & -1 & \ldots & \ldots & -1 & -1 \\ 0 & 0 & 0 & 0 & -5 & -1 & \ldots & \ldots & -1 & -1 \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots & & & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & & \ddots & & \vdots & \vdots \\ \vdots & \vdots & \vdots & \vdots & \vdots & & & \ddots & \vdots & \vdots \\ 0 & 0 & 0 & 0 & 0 & 0 & \ldots & \ldots & -(n-1) & -1 \\ 0 & 0 & 0 & 0 & 0 & 0 & \ldots & \ldots & 0 & -n \end{vmatrix}$$

where $k = \frac{1}{\sqrt{n}} \frac{1}{\sqrt{2 \cdot 1}} \frac{1}{\sqrt{3 \cdot 2}} \frac{1}{\sqrt{4 \cdot 3}} \cdots \frac{1}{\sqrt{(n-2)(n-1)}} \frac{1}{\sqrt{n(n-1)}}$

The above is the determent of the upper triangular matrix whose value obtained by multiplying the elements of the main diagonal.

$$\frac{1}{J} = k\,(1)(-2)(-3)(-4)\ldots\ldots[-(n-1)](-n)$$

$$= \frac{(-1)^{n-1} 1 \cdot 2 \cdot 3 \cdot 4 \cdot 5 \ldots\ldots (n-1)n}{\sqrt{1 \cdot 2 \cdot 3 \ldots (n-1)n}\sqrt{1 \cdot 2 \cdot 3 \ldots (n-1)n}}$$

$$= \frac{(-1)^{n-1} n!}{\sqrt{n!}\sqrt{n!}} = 1$$

$\frac{1}{J} = 1$ implies $J = 1$ and

$$\sum_{i=1}^{n} W_i^{\,2} = \sum_{i=1}^{n} Y_i^{\,2}$$

This show that the set $\{Y_i\}_{i=1}^{n}$ represent a r.s. of size n from $N(0,1)$.

Now,

$$\sum_{i=2}^{n} \mathrm{Y}_i^{\,2} = \sum_{i=1}^{n} \mathrm{Y}_i^{\,2} - \mathrm{Y}_1^{\,2} = \sum_{i=1}^{n} W_i^{\,2} - \left(\sqrt{n}\overline{W}\right)^2$$

$$\sum_{i=1}^{n} W_i^{\,2} - n\overline{W}^{\,2} = \sum_{i=1}^{n} (W_i - \overline{W})^2$$

since $\{Y_i\}_{i=1}^{n}$ are indep. r.v.'s from $\mathrm{N}(0,1)$, then according to theorems (1.4.3) and (1.4.4)

$$\sum_{i=2}^{n} \mathrm{Y}_i^{\,2} = \sum_{i=1}^{n} (W_i - \overline{W})^2 = \frac{(n-1)S^2}{\sigma^2} \sim \chi^2(n-1) \qquad (1.14)$$

Also, we have

Y_1 is distributed independently of $\sum_{i=2}^{n} \mathrm{Y}_i^{\,2} = \frac{(n-1)S^2}{\sigma^2}$

where $\mathrm{Y}_1 = \sqrt{n}\overline{W} = \sqrt{n}\left(\frac{\overline{X} - \mu}{\sigma}\right)$

It follows that $\frac{(n-1)S^2}{\sigma^2}$ is indep. of $\sqrt{n}\,\frac{\overline{\mathrm{X}} - \mu}{\sigma}$.

1.4.2 Distribution of $\overline{\mathrm{X}}$ and $\frac{(n-1)S^2}{\sigma^2}$

Let $\mathrm{X}_1, \mathrm{X}_2, \ldots, \mathrm{X}_n$ be a r.s. of size $n \geq 2$ from $\mathrm{N}(\mu, \sigma^2)$, we shall consider first the distn. of the sample mean $\overline{\mathrm{X}} = \frac{1}{n}\sum_{i=1}^{n} \mathrm{X}_i$.

According to theorem (1.4.2) with $k_i = \frac{1}{n}$, $\mu_i = \mu$, $\sigma_i^{\,2} = \sigma^2$, $\forall i = 1,2,\ldots,k$

Then $\overline{\mathrm{X}} \sim \mathrm{N}\left(\mu, \frac{\sigma^2}{n}\right)$ (1.15)

Also,

$$S^2 = \frac{1}{n-1}\sum_{i=1}^{n} (\mathrm{X}_i - \overline{\mathrm{X}})^2 = \frac{1}{n-1}\left[\sum_{i=1}^{n} \mathrm{X}_i^{\,2} - n\overline{\mathrm{X}}^2\right]$$

Consider

$$\sum_{i=1}^{n}(X_i-\mu)^2=\sum_{i=1}^{n}\left[(X_i-\bar{X})+(\bar{X}-\mu)\right]^2$$

$$=\sum_{i=1}^{n}(X_i-\bar{X})^2+2(\bar{X}-\mu)\sum_{i=1}^{n}(X_i-\bar{X})+n(\bar{X}-\mu)^2$$

but $\sum_{i=1}^{n}(X_i-\bar{X})=\sum_{i=1}^{n}X_i-n\bar{X}=n\bar{X}-n\bar{X}=0$

So,

$$\sum_{i=1}^{n}(X_i-\mu)^2=\sum_{i=1}^{n}(X_i-\bar{X})^2+n(\bar{X}-\mu)^2$$

$$=(n-1)S^2+n(\bar{X}-\mu)^2 \qquad (1.16)$$

by dividing both sides of eq.(1.16) by σ^2, we have

$$\sum_{i=1}^{n}\frac{(X_i-\mu)^2}{\sigma^2}=\frac{(n-1)S^2}{\sigma^2}+\frac{n(\bar{X}-\mu)^2}{\sigma^2}$$

$$\frac{(n-1)S^2}{\sigma^2}=\sum_{i=1}^{n}\frac{(X_i-\mu)^2}{\sigma^2}-\frac{n(\bar{X}-\mu)^2}{\sigma^2} \qquad (1.17)$$

since $X_i \sim N(\mu,\sigma^2)$, then according to the theorem of section 1.2.3, the r.v. $\frac{X_i-\mu}{\sigma} \sim N(0,1)$.

Also, according to theorems(1.4.1) and (1.4.4), the r.v. $\left(\frac{X_i-\mu}{\sigma}\right)^2 \sim \chi^2(1)$

and the r.v. $Y=\sum_{i=1}^{n}\left(\frac{X_i-\mu}{\sigma}\right)^2 \sim \chi^2(n)$

Also, since the r.v. $\bar{X} \sim N\left(\mu,\frac{\sigma^2}{n}\right)$, then the r.v. $\frac{\bar{X}-\mu}{\frac{\sigma}{\sqrt{n}}} \sim N(0,1)$

and the r.v. $\left(\frac{\sqrt{n}(\overline{X}-\mu)}{\sigma}\right)^2 = n\left(\frac{\overline{X}-\mu}{\sigma}\right)^2 \sim \chi^2(1)$.

and the r.v.'s $\overline{X}$ and $\frac{(n-1)S^2}{\sigma^2}$ are stochastically independent. Then, the r.v.'s

$\sum_{i=1}^{n}\left(\frac{X_i-\mu}{\sigma}\right)^2$ and $n\left(\frac{\overline{X}-\mu}{\sigma}\right)^2$ are stochastically independent.

and since from eq.(1.17), the r.v. $\frac{(n-1)S^2}{\sigma^2} = \sum_{i=1}^{n}\frac{(X_i-\mu)^2}{\sigma^2} - \frac{n(\overline{X}-\mu)^2}{\sigma^2}$

then, the r.v. $\frac{(n-1)S^2}{\sigma^2} \sim \chi^2(n-1)$.

1.5 Estimation

The problem of estimation can be defined as follows:

Let $X_1, X_2, \ldots, X_n$ be a r.s. of size n from a distn. Whose p.d.f. is $f(x, \underset{\sim}{\theta})$, where $\underset{\sim}{\theta} = (\theta_1, \theta_2, \ldots, \theta_k)$ is a vector of unknown parameters, we assume that the values $x_1, x_2, \ldots, x_n$ of r.v.'s $X_1, X_2, \ldots, X_n$ are observed. It is desired to estimate $\underset{\sim}{\theta}$ on the basis of the observed values $x_1, x_2, \ldots, x_n$.
This estimation can be made in two ways:

1.5.1 Interval estimation

Is to find two statistics say $U_1 = u_1(X_1, X_2, \ldots, X_n)$ and $U_2 = u_2(X_1, X_2, \ldots, X_n)$, $(U_1 \le U_2)$ such that the unknown parameter say θ lie in between that $U_1 \le \theta \le U_2$ with certain prob. Say $1-\alpha$ (α is small).

For normal case, we have two unknown parameters μ and σ^2. We assume $X_1, X_2, \ldots, X_n$ be a r.s. of size n from $N(\mu, \sigma^2)$ is available and a confidence interval for the distn. parameters are required with prob. $1-\alpha$.

(I) Confidence Interval for the Mean μ

There are two cases:

Case (1): when σ^2 is known

According to section 1.4.2, we have

$$\overline{X} \sim N\left(\mu, \frac{\sigma^2}{n}\right) \Rightarrow Z = \frac{\overline{X}-\mu}{\frac{\sigma}{\sqrt{n}}} = \frac{\sqrt{n}(\overline{X}-\mu)}{\sigma} \sim N(0,1)$$

So, we can find from N(0,1) table two no.'s say $\pm Z_{1-\frac{\alpha}{2}}$ such that

$$pr\left(-Z_{1-\frac{\alpha}{2}} < Z < Z_{1-\frac{\alpha}{2}}\right) = 1-\alpha$$

Now, consider the event

$$-Z_{1-\frac{\alpha}{2}} < Z < Z_{1-\frac{\alpha}{2}} \equiv -Z_{1-\frac{\alpha}{2}} < \frac{\sqrt{n}(\overline{X}-\mu)}{\sigma} < Z_{1-\frac{\alpha}{2}}$$

$$\overline{X} - \frac{\sigma}{\sqrt{n}} Z_{1-\frac{\alpha}{2}} < \mu < \overline{X} + \frac{\sigma}{\sqrt{n}} Z_{1-\frac{\alpha}{2}}$$

therefore, the $100(1-\alpha)\%$ C.I for μ is:

$$\left(\overline{X} - \frac{\sigma}{\sqrt{n}} Z_{1-\frac{\alpha}{2}},\ \overline{X} + \frac{\sigma}{\sqrt{n}} Z_{1-\frac{\alpha}{2}}\right).$$

Case (2): when σ^2 is unknown

According to section 1.4.2, we have

1) $\overline{X} \sim N\left(\mu, \frac{\sigma^2}{n}\right) \Rightarrow \frac{\sqrt{n}(\overline{X}-\mu)}{\sigma} \sim N(0,1)$

2) $\frac{(n-1)S^2}{\sigma^2} \sim \chi^2(n-1)$

3) $\overline{X}$ and S^2 are stochastically indep.

Then, according to theorem (1.4.5)

the r.v. $\mathrm{T}=\dfrac{\sqrt{n}\,(\overline{\mathrm{X}}-\mu)/\sigma}{\sqrt{\dfrac{(n-1)S^2}{\sigma^2}/(n-1)}}=\dfrac{\sqrt{n}\,(\overline{\mathrm{X}}-\mu)}{S}\sim t(n-1)$

So, from t-distn. table with $(n-1)$ dof, we can find two no.'s say $\pm t_{1-\frac{\alpha}{2}}$

such that $pr\left(-t_{1-\frac{\alpha}{2}}<\mathrm{T}<t_{1-\frac{\alpha}{2}}\right)=1-\alpha$

Now, consider the event

$$-t_{1-\frac{\alpha}{2}}<\mathrm{T}<t_{1-\frac{\alpha}{2}}\equiv -t_{1-\frac{\alpha}{2}}<\frac{\sqrt{n}\,(\overline{\mathrm{X}}-\mu)}{S}<t_{1-\frac{\alpha}{2}}$$

$$\overline{\mathrm{X}}-\frac{S}{\sqrt{n}}t_{1-\frac{\alpha}{2}}<\mu<\overline{\mathrm{X}}+\frac{S}{\sqrt{n}}t_{1-\frac{\alpha}{2}}$$

therefore, the $100(1-\alpha)\%$ C.I for μ is:

$$\left(\overline{\mathrm{X}}-\frac{S}{\sqrt{n}}t_{1-\frac{\alpha}{2}},\ \overline{\mathrm{X}}+\frac{S}{\sqrt{n}}t_{1-\frac{\alpha}{2}}\right).$$

(II) Confidence Interval for the Variance σ^2

There are two cases:

Case (1): when μ is known

According to theorem of section 1.2.3 and theorems (1.4.1) and (1.4.4), we have

$\mathrm{X}_i\sim \mathrm{N}(\mu,\sigma^2),\ \forall i=1,2,\ldots,n$

$\dfrac{\mathrm{X}_i-\mu}{\sigma}\sim \mathrm{N}(0,1)$ and $\left(\dfrac{\mathrm{X}_i-\mu}{\sigma}\right)^2\sim \chi^2(1)$

$$Y = \sum_{i=1}^{n} \left(\frac{X_i - \mu}{\sigma} \right)^2 \sim \chi^2(n)$$

so, we can find from χ^2-distn. table two no.'s say $\chi^2_{\frac{\alpha}{2}}$ and $\chi^2_{1-\frac{\alpha}{2}}$ such that

$$pr\left(\chi^2_{\frac{\alpha}{2}} < Y < \chi^2_{1-\frac{\alpha}{2}} \right) = 1-\alpha$$

Now, consider the event

$$\chi^2_{\frac{\alpha}{2}} < Y < \chi^2_{1-\frac{\alpha}{2}} \equiv \chi^2_{\frac{\alpha}{2}} < \frac{\sum_{i=1}^{n} (X_i - \mu)^2}{\sigma^2} < \chi^2_{1-\frac{\alpha}{2}}$$

$$\frac{\sum_{i=1}^{n} (X_i - \mu)^2}{\chi^2_{1-\frac{\alpha}{2}}} < \sigma^2 < \frac{\sum_{i=1}^{n} (X_i - \mu)^2}{\chi^2_{\frac{\alpha}{2}}}$$

therefore, the $100(1-\alpha)\%$ C.I for σ^2 is

$$\left(\frac{\sum_{i=1}^{n} (X_i - \mu)^2}{\chi^2_{1-\frac{\alpha}{2}}}, \frac{\sum_{i=1}^{n} (X_i - \mu)^2}{\chi^2_{\frac{\alpha}{2}}} \right).$$

Case (2): when μ is unknown

According to section 1.4.2, we have

$Y = \frac{(n-1)S^2}{\sigma^2} \sim \chi^2(n-1)$ when $S^2 = \frac{1}{(n-1)} \left[\sum_{i=1}^{n} X_i^2 - n\overline{X}^2 \right]$ is the sample variance

So, we can find two no.'s from χ^2-distn. table with $(n-1)$ dof, say $\chi^2_{\frac{\alpha}{2}}$ and $\chi^2_{1-\frac{\alpha}{2}}$ such that

$$pr\left(\chi^2_{\frac{\alpha}{2}} < \mathrm{Y} < \chi^2_{1-\frac{\alpha}{2}}\right) = 1-\alpha$$

Now, consider the event

$$\chi^2_{\frac{\alpha}{2}} < \mathrm{Y} < \chi^2_{1-\frac{\alpha}{2}} \equiv \chi^2_{\frac{\alpha}{2}} < \frac{(n-1)S^2}{\sigma^2} < \chi^2_{1-\frac{\alpha}{2}}$$

$$\frac{(n-1)S^2}{\chi^2_{1-\frac{\alpha}{2}}} < \sigma^2 < \frac{(n-1)S^2}{\chi^2_{\frac{\alpha}{2}}}$$

therefore, the $100(1-\alpha)\%$ C.I for σ^2 is

$$\left(\frac{(n-1)S^2}{\chi^2_{1-\frac{\alpha}{2}}}, \frac{(n-1)S^2}{\chi^2_{\frac{\alpha}{2}}}\right).$$

1.5.2 Point Estimation [21]

Point estimation is concerned with inference about the unknown parameters of a distribution from a sample. It provides a single value for each unknown parameter. Point estimation admits two problems:

First, developing methods of obtaining a statistics whose values can be used to estimate the unknown parameters of the distribution, such statistics are called point estimators.

Second, selecting criteria and technique to obtain a best estimator among possible estimators.

1.5.2.1 Definition (Estimator) [21]

Any statistic whose values are used to estimate the unknown parameter θ or some function of θ, say $\tau(\theta)$ is called point estimator.

1.5.3 Methods of Finding Estimators

Many techniques have been proposed in the literatures for finding estimators of the distn. parameters such as Moments, Maximum likelihood, Minimum chi-square, Minimum distance, Least square and Bayesian method.

These methods provide a single value for each unknown parameter of the distribution. For normal case, we shall discuss two methods: The method of Moments and the Maximum likelihood method.

1.5.3.1 Moments Method

Let $X_1, X_2, \ldots, X_n$ be a r.s. of size n from a distribution whose p.d.f. $f(x, \underset{\sim}{\theta})$, where $\underset{\sim}{\theta} = (\theta_1, \theta_2, \ldots, \theta_k)$ is a vector of unknown parameters.

Let $\mu_r' = E(X^r)$ be the r^{th} moment about origin of the distribution and let $\mu_r = \frac{1}{n}\sum_{i=1}^{n} X_i^r$ be the r^{th} moment about origin of the sample. The method of moments can be described follows:

Since, we have k unknown parameters, equate

μ_r' to μ_r at $\theta = \hat{\theta}$. That is

$\mu_r' = \mu_r$ at $\theta = \hat{\theta}$, $r = 1, 2, 3, \ldots, k$.

For these k equations, we find a unique solution for $\hat{\theta}_1, \hat{\theta}_2, ..., \hat{\theta}_k$ and we say that $\hat{\theta}_r$, $(r = 1, 2, \ldots, k)$ is an estimate of θ_r obtained by method of moments and the corresponding statistic $\hat{\theta}_r$ is the method of moments estimator of θ_r.

Now, to estimate μ and σ^2 for normal case by method of moments we let $X_1, X_2, \ldots, X_n$ be a r.s. of size n from $N(\mu, \sigma^2)$ is taken.

Since $N(\mu, \sigma^2)$ distribution involve two unknown parameters

We set $\mu'_r = \mu_r$ at $\mu = \hat{\mu}$, $\sigma^2 = \hat{\sigma}^2$, $r = 1, 2$

$r = 1$ implies

$$\mu'_1 = E(X) = \mu, \ \mu_1 = \frac{1}{n}\sum_{i=1}^{n} X_i = \overline{X}$$

$r = 2$ implies

$$\mu'_2 = E(X^2) = \mu^2 + \sigma^2, \ \mu_2 = \frac{1}{n}\sum_{i=1}^{n} X_i^2 = \frac{(n-1)}{n} S^2 + \overline{X}^2 \text{ where:}$$

$$S^2 = \frac{1}{n-1}\left[\sum_{i=1}^{n} X_i{}^2 - n\overline{X}^2\right]$$

$r = 1$ implies $\mu'_1 = \mu_1$ at $\mu = \hat{\mu}$, $\sigma^2 = \hat{\sigma}^2$, we obtain

$$\hat{\mu} = \overline{X} \tag{1.18}$$

$r = 2$ implies $\mu'_2 = \mu_2$ at $\mu = \hat{\mu}$, $\sigma^2 = \hat{\sigma}^2$, we obtain

$$\hat{\mu}^2 + \hat{\sigma}^2 = \frac{n-1}{n} S^2 + \overline{X}^2 \tag{1.19}$$

solving eqs. (1.18) and (1.19) we get

$$\hat{\sigma}^2 = \frac{n-1}{n} S^2 \text{ and } \hat{\mu} = \overline{X} \tag{1.20}$$

are respectively the estimators of μ and σ^2 obtained by method of moments.

1.5.3.2 Maximum Likelihood Method [14]

Definition (likelihood function)

The likelihood of a r.s. $X_1, X_2, \ldots, X_n$ of size n from a distribution having p.d.f. $f(X, \underset{\sim}{\theta})$ where $\underset{\sim}{\theta} = (\theta_1, \theta_2, \ldots, \theta_k)$ is a vector of unknown

parameters is defined to be the joint p.d.f. of the n r.v.'s $X_1, X_2, \ldots, X_n$ which is considered as a function of $\underset{\sim}{\theta}$ and denoted by $L(\underset{\sim}{\theta}, \underset{\sim}{X})$, that is

$$L = L(\underset{\sim}{\theta}, \underset{\sim}{X}) = f(\underset{\sim}{X}, \underset{\sim}{\theta}) = \prod_{i=1}^{n} f(X_i, \underset{\sim}{\theta}).$$

Now,

Let $L(\underset{\sim}{\theta}, \underset{\sim}{X})$ be the likelihood function of a r.s. $X_1, X_2, \ldots, X_n$ of size n from a distribution whose p.d.f. $f(X, \underset{\sim}{\theta})$, $\underset{\sim}{\theta} = (\theta_1, \theta_2, \ldots, \theta_k)$ is a vector of unknown parameters.

Let $\underset{\sim}{\hat{\theta}} = u(\underset{\sim}{X})$

$= (u_1(\underset{\sim}{X}), u_2(\underset{\sim}{X}), \ldots, u_k(\underset{\sim}{X}))$

be a vector function of the observations $\underset{\sim}{X} = (X_1, X_2, \ldots, X_n)$

If $\underset{\sim}{\hat{\theta}}$ have the value of $\underset{\sim}{\theta}$ which maximizes $L(\underset{\sim}{\hat{\theta}}, \underset{\sim}{X})$ then $\underset{\sim}{\hat{\theta}}$ is the m.l.e of $\underset{\sim}{\theta}$ and the corresponding statistic $\underset{\sim}{\hat{\Theta}}$ is the M.L.E of $\underset{\sim}{\theta}$. We note that:

(I) Many likelihood function satisfy the condition that the m.l.e is a solution of likelihood equations

$$\frac{\partial L(\underset{\sim}{\theta}, \underset{\sim}{X})}{\partial \theta_r} = 0, \text{ at } \underset{\sim}{\theta} = \underset{\sim}{\hat{\theta}},\ r = 1, 2, 3, \ldots, k\ .$$

(II) Since $L(\underset{\sim}{\theta}, \underset{\sim}{X})$ and $\ln L(\underset{\sim}{\theta}, \underset{\sim}{X})$ have their maximum at the same value of $\underset{\sim}{\theta}$ so sometimes it is easier to find the maximum of the logarithm of the likelihood.

In such case, the m.l.e $\underset{\sim}{\hat{\theta}}$ of $\underset{\sim}{\theta}$ which maximizes $L(\underset{\sim}{\theta}, \underset{\sim}{X})$ may be given the solution of the likelihood equations

$$\frac{\partial \ln L(\underset{\sim}{\theta}, \underset{\sim}{X})}{\partial \theta_r} = 0, \text{ at } \underset{\sim}{\theta} = \underset{\sim}{\hat{\theta}},\ r = 1, 2, 3, \ldots, k\ .$$

For normal case

Let $X_1, X_2, \ldots, X_n$ be a r.s. of size n from $N(\mu, \sigma^2)$ where the distribution p.d.f. is given by eq.(1.1), the likelihood function is

$$L(\mu, \sigma^2, \underset{\sim}{x}) = f(\underset{\sim}{x}, \mu, \sigma^2)$$

$$= \prod_{i=1}^{n} f(x_i, \mu, \sigma^2) = \prod_{i=1}^{n} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2\sigma^2}(x_i - \mu)^2}$$

$$= (2\pi)^{-\frac{n}{2}} (\sigma^2)^{-\frac{n}{2}} e^{-\frac{1}{2}\sum_{i=1}^{n}(x_i - \mu)^2}$$

$$\ln L = -\frac{n}{2}\ln(2\pi) - \frac{n}{2}\ln\sigma^2 - \frac{1}{2\sigma^2}\sum_{i=1}^{n}(x_i - \mu)^2$$

$$\frac{\partial \ln L}{\partial \mu} = \frac{1}{\sigma^2}\sum_{i=1}^{n}(x_i - \mu) \tag{1.21}$$

$$\frac{\partial \ln L}{\partial \sigma^2} = -\frac{n}{2}\frac{1}{\sigma^2} + \frac{1}{2(\sigma^2)^2}\sum_{i=1}^{n}(x_i - \mu)^2 \tag{1.22}$$

we set $\frac{\partial \ln L}{\partial \mu} = 0$ and $\frac{\partial \ln L}{\partial \sigma^2} = 0$ at $\mu = \hat{\mu}$, $\sigma^2 = \hat{\sigma}^2$

we have

$$\sum_{i=1}^{n}(x_i - \hat{\mu}) = 0 \Rightarrow \hat{\mu} = \frac{1}{n}\sum_{i=1}^{n} x_i \tag{1.23}$$

and $$-\frac{n}{2\hat{\sigma}^2} + \frac{1}{2(\hat{\sigma}^2)^2}\sum_{i=1}^{n}(x_i - \hat{\mu})^2 = 0 \tag{1.24}$$

This implies that:

$$\hat{\mu} = \overline{X} \tag{1.25}$$

from eq.(1.24) and eq.(1.25), we have

$$\hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n}\left(x_i - \overline{x}\right)^2$$

implies $\hat{\sigma}^2 = \dfrac{(n-1)S^2}{n}$ (1.26)

1.6 Equality of Estimations [21]

In this section, we shall introduce some definitions and theorems concern the equality of estimators which reach to the best estimators for the unknown parameters.

1.6.1 Definition [21]

Let the statistic $\hat{\theta} = u(X_1, X_2, \ldots, X_n)$ be an estimator of the unknown parameter θ, then $\hat{\theta}$ is said to be an

(I) Unbiased estimator if and only if $E(\hat{\theta}) = \theta$, otherwise $\hat{\theta}$ is called biased estimator for θ. The term $E(\hat{\theta}) - \theta$ is called the bias term.

(II) Consistent estimator if $\underset{n \to \infty}{Lim}\ pr\left(\left|\hat{\theta} - \theta\right| < \varepsilon\right) = 0$.

(III) Asymptotically unbiased if $\underset{n \to \infty}{Lim}\ E(\hat{\theta}) = \theta$.

(IV) Minimum variance unbiased estimator (MVUE) if

1. $\hat{\theta}$ is an unbiased estimator for θ.

2. The variance of $\hat{\theta}$ is less than or equal to the variance of every other unbiased estimators of θ.

1.6.2 Definition (Sufficient Statistics) [2, 13]

Let $X_1, X_2, \ldots, X_n$ be a r.s. of size n from a distn. whose p.d.f. $f(\underset{\sim}{x}, \underset{\sim}{\theta})$, where $\underset{\sim}{\theta} = (\theta_1, \theta_2, \ldots, \theta_m)$ is a vector of unknown parameters and $Y_i = u_i(X_1, X_2, \ldots, X_n)$, $i = 1, 2, \ldots, m$ be K statistics whose joint p.d.f. $g(\underset{\sim}{y}, \underset{\sim}{\theta})$. Then the K statistics are called jointly sufficient statistics for $\underset{\sim}{\theta}$ iff

$$\frac{f(\underset{\sim}{x},\underset{\sim}{\theta})}{g(\underset{\sim}{y},\underset{\sim}{\theta})} = \mathrm{H}(\underset{\sim}{x})$$

where $\mathrm{H}(x)$ does not depend on $\underset{\sim}{\theta}$ for all fixed values of $y_i = u_i(x_i)$, $i = 1,2,\ldots,m$.

1.6.3 Theorem (Neymann Factorization Theorem) [2]

Let $\mathrm{X}_1, \mathrm{X}_2, \ldots, \mathrm{X}_n$ be a r.s. of size n from a distn. whose p.d.f. $f(x,\underset{\sim}{\theta})$, where $\underset{\sim}{\theta} = (\theta_1, \theta_2, \ldots, \theta_m)$ is a vector of unknown parameters. A set of statistics $Y_i = u_i(\underset{\sim}{X})$, $i = 1,2,\ldots,m$ is jointly sufficient statistics for $\underset{\sim}{\theta}$ iff, we can find two nonnegative functions k_1 and k_2 such that

$$f(\underset{\sim}{x};\underset{\sim}{\theta}) = f(x_1, x_2, \ldots, x_n; \theta_1, \theta_2, \ldots, \theta_m)$$
$$= k_1[u_1(\underset{\sim}{x}), u_2(\underset{\sim}{x}), \ldots, u_m(\underset{\sim}{x}); \theta_1, \theta_2, \ldots, \theta_m] \cdot k_2(\underset{\sim}{x})$$

where $k_2(\underset{\sim}{x})$ is free of $\underset{\sim}{\theta}$ for every values of $y_1, y_2, \ldots, y_k$ of $\mathrm{Y}_1, \mathrm{Y}_2, \ldots, \mathrm{Y}_m$.

For normal $\mathrm{N}(\mu,\sigma^2)$ case, we have two unknown parameters μ and σ^2, where we assume a r.s. $\mathrm{X}_1, \mathrm{X}_2, \ldots, \mathrm{X}_n$ is a available, then the joint p.d.f. can be written as

$$f(\underset{\sim}{x};\mu,\sigma^2) = \prod_{i=1}^{n} f\left(x_i,\mu,\sigma^2\right) = \prod_{i=1}^{n} (2\pi)^{-\frac{1}{2}} (\sigma^2)^{-\frac{1}{2}} e^{-\frac{1}{2\sigma^2}(x_i-\mu)^2}$$

$$= (\sigma^2)^{-\frac{n}{2}} e^{-\frac{1}{2\sigma^2}\left[\sum_{i=1}^{n} x_i^2 - 2\mu\sum_{i=1}^{n} x_i + n\mu^2\right]} \cdot (2\pi)^{-\frac{n}{2}}$$

$$= k_1\left[\sum_{i=1}^{n} x_i, \sum_{i=1}^{n} x_i^2, \mu, \sigma^2\right] \cdot k_2(\underset{\sim}{x})$$

where $k_2(\underset{\sim}{x}) = (2\pi)^{-\frac{n}{2}}$.

Thus according to factorization theorem (1.6.3), the statistics $Y_1 = \sum_{i=1}^{n} X_i$ and $Y_2 = \sum_{i=1}^{n} X_i^2$ are jointly sufficient statistics for μ and σ^2.

Remark

If $\{Y_i = u_i(\underset{\sim}{X})\}$, $i = 1,2,\ldots,m$ is a set of jointly sufficient statistics for $\underset{\sim}{\theta} = (\theta_1, \theta_2, \ldots, \theta_m)$, then any set of one-to-one functions or transformations of $Y_1, Y_2, \ldots, Y_m$ are also jointly sufficient statistics for $\underset{\sim}{\theta}$.

For normal $N(\mu, \sigma^2)$ case, we have $\sum_{i=1}^{n} X_i$ and $\sum_{i=1}^{n} X_i^2$ are jointly sufficient statistics for μ and σ^2.

But $\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$ and $S^2 = \frac{1}{n-1}\left[\sum_{i=1}^{n} X_i{}^2 - n\overline{X}^2\right]$ are one-to-one functions of $\sum_{i=1}^{n} X_i$ and $\sum_{i=1}^{n} X_i^2$. Then $\overline{X}$ and S^2 are jointly sufficient statistics for μ and σ^2.

1.6.4 Definition (Completeness) [2, 13]

Let X be a r.v. of either type (continuous or discrete) defined on S.S A and having p.d.f. as a member of the family $\{f(x;\theta), \theta \in \Omega\}$ of p.d.f.'s, and let $u(X)$ be a continuous function (not a function of θ). If $E[u(X)] = 0$, $\forall \theta \in \Omega$ implies $u(x) = 0$, $\forall x \in A$, then the family $\{f(x;\theta), \theta \in \Omega\}$ is called a complete family of p.d.f.'s.

Remark

If $Y = u(\underset{\sim}{X})$ is a sufficient statistic for θ whose p.d.f. belong to the complete family of p.d.f.'s, then Y is called a complete sufficient statistic for θ.

1.6.5 Theorem (Lehmann-Scheffé Theorem-1st Theorem) [2]

Let $X_1, X_2, \ldots, X_n$ be a r.s. of size n from a distn. whose p.d.f. $f(\underset{\sim}{x};\underset{\sim}{\theta}), \theta \in \Omega$. Let $Y = u(\underset{\sim}{X})$ be a sufficient statistic for θ whose p.d.f. belong to the complete family $\{g(y;\theta), \theta \in \Omega\}$.

If $\Phi(Y)$ is a function of Y which is an unbiased estimator for θ, then $\Phi(Y)$ is a unique MVUE for θ.

1.6.6 Definition (The Exponential Family of p.d.f.'s)[2, 13]

Several parameter cases

Consider the family $\{f(x;\underset{\sim}{\theta}), \underset{\sim}{\theta} \in \Omega^m\}$ of p.d.f.'s which can be expressed as

$$f(x;\theta) = \exp\left[\sum_{j=1}^{m} p_j(\theta) k_j(x) + q(\underset{\sim}{\theta}) + s(x)\right], \; a < x < b$$

=0, e.w.

Such p.d.f. is said to be a member of exponential class of p.d.f.'s and satisfying the following conditions:

(i) Neither a nor b depends on $\underset{\sim}{\theta} = (\theta_1, \theta_2, \ldots, \theta_m)$.

(ii) $p_j(\underset{\sim}{\theta})$ is nontrivial, functionally independent, continuous functions of θ_j, $j = 1, 2, \ldots, m$.

(iii) $k_j'(x) \neq 0$ and $s(x)$ is continuous function of x for $a < x < b$.

Now, if a r.s. $X_1, X_2, \ldots, X_n$ is taken from a distn. whose p.d.f. $f(x;\underset{\sim}{\theta})$.

Then the joint p.d.f. of the sample set $\{X_i\}$ is

$$f(\underset{\sim}{x},\underset{\sim}{\theta}) = \prod_{i=1}^{n} f(x_i,\underset{\sim}{\theta}) = \prod_{i=1}^{n} \exp\left[\sum_{j=1}^{m} p_j(\underset{\sim}{\theta}) k_j(x_i) + q(\underset{\sim}{\theta}) + s(x_i)\right]$$

$$= Exp\left[\sum_{j=1}^{m} p_j(\underset{\sim}{\theta}) \sum_{i=1}^{n} k_i(x_i) + nq(\underset{\sim}{\theta}) + \sum_{i=1}^{n} s(x_i)\right]$$

$$= Exp\left[\sum_{j=1}^{m} p_j(\underset{\sim}{\theta})\sum_{i=1}^{n} k_i(x_i) + nq(\underset{\sim}{\theta})\right]\cdot Exp\left[\sum_{i=1}^{n} s(x_i)\right]$$

Then according to the Factorization theorem (1.6.3),

The statistics $Y_1 = \sum_{i=1}^{n} k_1(x_i)$, $Y_2 = \sum_{i=1}^{n} k_2(x_i)$,..., $Y_m = \sum_{i=1}^{n} k_m(x_i)$

are jointly sufficient statistics for the m parameters $\theta_1, \theta_2, \ldots, \theta_m$.

Note

If can be shown easily [2] that the joint p.d.f. of the sufficient statistics $Y_1, Y_2, \ldots, Y_m$ take the form

$$R(y_1, y_2, \ldots, y_m)\exp\left[\sum_{i=1}^{m} p_j(\underset{\sim}{\theta})y_i + nq(\underset{\sim}{\theta})\right] \qquad (1.27)$$

This p.d.f. of eq.(1.27) expressed as a member of the exponential family.

1.6.6 Theorem (Lehmann-Scheffé-2^{nd} Theorem) [2]

Let $X_1, X_2, \ldots, X_n$ be a r.s. of size n from a distn. whose p.d.f. $f(x;\underset{\sim}{\theta})$, $\underset{\sim}{\theta} = (\theta_1, \theta_2, \ldots, \theta_m)$ belong to the exponential family and let $Y_1, Y_2, \ldots, Y_m$ be jointly sufficient statistics for $\theta_1, \theta_2, \ldots, \theta_m$, then the family of p.d.f.'s $\{g(y;\underset{\sim}{\theta}), \underset{\sim}{\theta} \in \Omega^m\}$ is complete and the statistics $Y_1, Y_2, \ldots, Y_m$ are jointly complete sufficient statistics for $\theta_1, \theta_2, \ldots, \theta_m$.

*For $N(\mu, \sigma^2)$ with p.d.f.

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

which can be written as a member of the exponential family as

$$f(x;\mu,\sigma^2) = \exp\left[-\frac{1}{2}\ln(2\pi) - \frac{1}{2}\ln(\sigma^2) - \frac{1}{2\sigma^2}(x^2 - 2\mu x + \mu^2)\right]$$

$$= \exp\left[\frac{\mu}{\sigma^2}x - \frac{1}{2\sigma^2}x^2 - \left(\frac{\mu^2}{2\sigma^2} + \frac{1}{2}\ln\sigma^2\right) - \frac{1}{2}\ln(2\pi)\right]$$

Where $p_1(\mu,\sigma^2) = \frac{\mu}{\sigma^2}$, $p_2(\mu,\sigma^2) = \frac{-1}{2\sigma^2}$, $k_1(x) = x$, $k_2(x) = x^2$,

$q(\mu,\sigma^2) = -\left(\frac{\mu^2}{2\sigma^2} + \frac{1}{2}\ln\sigma^2\right)$, $s(x) = -\frac{1}{2}\ln(2\pi)$

Now, if a sample set $\{X_i\}$ is available, then the statistic $Y_1 = \sum_{i=1}^{n} k_1(x_i) = \sum_{i=1}^{n} X_i$ and $Y_2 = \sum_{i=1}^{n} k_2(x_i) = \sum_{i=1}^{n} X_i^2$ are jointly sufficient statistics for μ and σ^2.

Since $\overline{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$ and $S^2 = \frac{1}{n-1}\left[\sum_{i=1}^{n} X_i^2 - n\overline{X}^2\right]$ is a one-to-one functions of Y_1 and Y_2, then $\overline{X}$ and S^2 are jointly sufficient statistics for μ and σ^2.

As shown in sections 1.4.1 and 1.4.2 that

1. $\overline{X} \sim N\left(\mu, \frac{\sigma^2}{n}\right)$ and $\frac{(n-1)S^2}{\sigma^2} \sim \chi^2(n-1)$.
2. $\overline{X}$ and S^2 are stochastically independent.

Then the joint p.d.f. of $\overline{X}$ and S^2 can be expressed as a member of the exponential family and that implies that $\overline{X}$ and S^2 are minimal jointly sufficient statistics for μ and σ^2.

Further more, we have

$E(\overline{X}) = \mu$ and $E(S^2) = \sigma^2$, then from completeness we see that $\overline{X}$ and S^2 are M.V.U.E's for μ and σ^2 respectively.

CHAPTER

3

Generating Random Varieties from Normal Distribution

3.1 Introduction

The first step in studying a certain problem under consideration is building a mathematical model; the next step is driving a solution from this model. The solution may be obtained analytically or numerically. The analytic solution is usually obtained directly from it's mathematical representation in the form of the formula, while a numerical solution is generally an approximate solution obtained as a result of substitution of numerical values for the variables and parameters of the model. Many numerical methods are iterative, that is, each successive step in the solution uses the results from the previous step, such as Newton-Raphson method for approximating the roots of a nonlinear equation. Two special types of numerical methods are simulation and Monte Carlo designed for a solution of deterministic and stochastic problems.

Simulation "in the wide scene" is defined as numerical technique for conducting experiments on a digital computer, which involve certain types of mathematical and logical models that describe the system behavior over

extended periods of time, for example, simulating football game, supersonic jet flight, a telephone communication system, a wind tunnel, a large scale military battle (to evaluate defensive or offensive weapon system), or a maintenance operations (to evaluate the optimal size of repair crews). Simulation is often viewed as a " Method of Last Resort" to be used when everything else has failed, software building and technical developments have made simulation one of the most widely used and accepted tools for designer in system analysis and operational research.

Simulation "in a narrow sence" (also called stochastic simulation) is defined as experimenting with the model over time, it includes sampling stochastic varieties from probability distribution. Because sampling from a particular distribution involve the use of random numbers, stochastic simulation sometimes called Monte Carlo Simulation.

Historically, the Monte Carlo method was considered as a technique using random or pseudorandom numbers for solution of a model. These random numbers are essentially independent random variables uniformly distributed over unit interval [0, 1].

Actually there are arithmetic codes available at computer center (0 to 9) occurs with approximately equal probability "imagine flips of a fair ten-side die". Such codes are called random number generators.

In the beginning of the 20^{th}-century the Monte Carlo was used to examine the Boltzmann equation.

In 1908 the famous statistician Gosset (student) uses the Monte Carlo method for estimating the correlation coefficient in his t-distribution, [9].

One of the earliest problems connected with Monte Carlo method is the famous Buffon's needle problem, who found the probability of a needle of length L thrown randomly onto a floor composed of parallel planks of width

$D > L$ is $P = \frac{2L}{\pi D}$ which can be estimated as the ratio of the number of throws hitting the crack to the total number of throws.

A. N. Kolmogororv (1931) applies Monte Carlo method and showed the relationship between Markov stochastic processes and certain Integro-differential equations, [8].

In 1948 S.Ulam used Monte Carlo method for estimation of the eigenvalues of Schrodinger equation, [7].

The terms "Monte Carlo" was introduced by Von Neumann and Ulam during World War II, as a code word for secret work at Los Alamos, it was suggested by the gambling casinos at the city of Monte Carlo in Monaco. The Monte Carlo method was then applied to problems related to the atomic bomb [3] where the work involve direct simulation of behavior concerned with random neutron diffusion in fissionable material. Shortly thereafter Monte Carlo method were used to evaluate complex multidimensional integrals, solution of certain differential and integral equations stochastic problems, deterministic problems if they have the same formal expression as some stochastic process, evaluating parameters of queues and networks, sampling random varieties from probability distributions, and analyzing complex problems. A useful reference related to Monte Carlo simulation is given by Rubinstein [27] and Norman [22].

This chapter involves three sections. Section (3.2) illustrated random number generation. While section (3.3) discussed random varieties generation. Generation random varieties from normal distribution are studied theoretically and applied practically in section (3.4).

3.2 Random Number Generation

Many techniques for generating random numbers on digital computer by Monte Carlo method and simulation have been suggested tested and used in recent years. Some of these methods are based on random phenomena, others on deterministic recurrence procedures.

Initially manual methods were used to generate a sequence of numbers such as coin flipping, dice rolling, card shuffling, and roulette wheels, but these methods were to slow for general use, moreover the generated sequence not reproduced.

Shortly following with the computer aid it become possible to obtain random numbers. In 1951 Von Neumann [30] suggests the mid-square method using the arithmetic operations of a computer. His idea is to take the square of the preceding random number and extract the middle digits. For instance, suppose we wish to generate 4-digits numbers

1- Choose any 4-digit number, say 5232.
2- Square it, we have 27373824.
3- The next 4-digit number is the middle 4-digits of step2, that is, 3738.
4- Repeat the process.

Von Neumann's method proved slow and awkward for statistical analysis, furthermore the sequences tend to cyclicity and once a zero is encountered the sequence terminates. One method of generating random numbers on a digital computer is published in 1951 by RAND Corporation [24], the method consists of preparing a well known table of a million digits and. storing it in the memory of the computer. The advantage of this method is reproducibility and it’s disadvantage is its lack of speed and risk of exhausting the table.

It is noted in the literature that the random numbers generated by any

method is good one if the random numbers are uniformly distributed, statistically independent, reproducible, fast, and requires minimum capacity in the computer memory.

The congruential methods for generating pseudorandom numbers are designed specifically to satisfy as many of these requirements as possible.

These methods produce a non-random sequence of numbers according to some recursive formula based on calculating the residues modulo of some integer m of a linear transformation Although these processes are completely deterministic, Knuth in 1969 [16] show that the numbers generated by such sequence appear to be uniformly distributed and statistically independent.

The congruential methods [19] are based on a fundamental congruence relationship, which may be formulated as:

$$X_{i+1} = (aX_i + c) \pmod{m},\ i = 1, 2, \ldots, m \qquad (3.1)$$

Where a is a multiplier, c is the increment, and m is the modulus (a, c, m are non-negative integers), (mod m) means that eq.(3.1) can be written as:

$$X_{i+1} = aX_i + c - m\left[\frac{aX_i + c}{m}\right] \qquad (3.2)$$

Where [Z] is the largest positive integer in Z.

Given an initial starting value X_1 with fixed values of a, c, and m, then eq.(3.2) yields congruence relationship (modulo m) for any value i of the sequence $\{X_i\}$. The sequence $\{X_i\}$ will repeat itself in at most m steps and will be therefore periodic. For example,

Let $a = c = X_1 = 3$, and $m = 5$, then the sequence obtained from the recursive formula

$X_{i+1} = (3X_i + 3) \pmod 5$ is $X_i = 3, 2, 4, 0, 3, \ldots$

The random numbers on the unit interval [0, 1] can be obtained by:

$$U_i = \frac{X_i}{m} \tag{3.3}$$

It fellow's from eq.(3.3) that $X_i \le m$, $\forall i$, this inequality mean that the period of the generator can not exceed m, that is, the sequence $\{X_i\}$ contains at most m distinct numbers. So m must be chosen as large as possible to ensure a sufficiently large sequence of distinct numbers in the cycle.

It is noted in the literature [10, 18, 20] that good statistical results can be achieved from a computer by choosing $a = 2^7+1$, $c=1$ and $m = 2^{35}$.

3.3 Random Varieties Generation

Two well-known methods for generating random varieties form continuous distribution, namely the inverse transform method and acceptance-rejection method.

3.3.1 The Inverse Transform Method

Recall the properties of the c.d.f

$\Pr(X \le x)=F(x)$ of r.v. X

(i) $0 \le F(x) \le 1$.

(ii) $F(-\infty)=0$, $F(\infty)=1$.

(iii) $F(x)$ is non-decreasing function of x.

(iv) $F(x)$ is continuous function to the right at each x.

The inverse transform method is based on the following theorem:

Theorem (3.3.1.1) [18]

The r.v $U = F(X) \sim U(0,1)$ if and only if the r.v. $X = F^{-1}(U)$ has c.d.f $pr(X \le x) = F(X)$.

Proof

$\Rightarrow$ Consider the r.v. $U = F(X) \sim U(0,1)$ then the c.d.f. of U is

$$G(u) = pr(U \le u) = \begin{cases} 0, & u \le 0 \\ u, & 0 < u < 1 \\ 1, & u \ge 1 \end{cases}$$

$$pr(X \le x) = pr(F^{-1}(U) \le x) = pr(U \le F(x)) = F(x)$$

$\Leftarrow$ Consider the r.v. X has c.d.f $F(X) = pr(X \le x)$

$$G(u) = pr(U \le u) = pr(F(X) \le u)$$

$$= pr\left[X \le F^{-1}(u)\right]$$

$$= F\left[F^{-1}(u)\right] = u$$

The IT algorithm describe the necessary steps for generating r.v. by Inverse Transform Method

1) Generate U from U(0, 1).
2) Set $X = F^{-1}(U)$.
3) Deliver X as a r.v. generated from the p.d.f f(x).

We note that, this method is valid when the c.d.f. F(x) exists in a form for which the corresponding inverse transform can be solved analytically.

3.3.2 The Acceptance-Rejection Method [30]

This method consists of sampling a r.v. from an appropriate distn. and subjecting it to a test to determine whether or not it will be acceptance for use.

To carry out this method, the p.d.f. f(x) of the generated r.v. X represented as $f(x) = ch(x)g(x)$, where $c \ge 1$, $h(x)$ is also p.d.f. and $0 < g(x) \le 1$. Then we generate two $r.v.'s$ U and Y from U (0, 1) and h(y) respectively and test to see whether or not the inequality $U \le g(Y)$ holds:

1. If the inequality hold, then accept Y=X as a r.v. generated from f(x).
2. if the inequality violated, then reject the pair (U,Y) and try again.

The theory behind this method is based on the following theorem**.**

Theorem (3.3.2.1) [27]

Let the p.d.f of r.v. X represented as $f(x)=ch(x)g(x)$, where $c\geq 1$, $h(x)$ as also p.d.f., and $0<g(x)\leq 1$. Let U and Y be distributed U (0, 1) and h(y) respectively, then $pr[Y=x \mid U\leq g(Y)]=f(x)$.

Proof

$$pr[Y=x \mid U\leq g(Y)]=\frac{pr[Y=x, U\leq g(Y)]}{pr[U\leq g(Y)]}$$

$$=\frac{pr[Y=x, U\leq g(Y)]}{\int_x pr[Y=x, U\leq g(Y)]dx}$$

Using Bayes theorem [2], we have

$$pr[Y=x \mid U\leq g(Y)]=\frac{pr[U\leq g(Y)\mid Y=x]\,pr(Y=x)}{\int_x pr[U\leq g(Y)\mid Y=x]\,pr(Y=x)dx}$$

Since $pr[U\leq g(Y)\mid Y=x]=pr[U\leq g(x)]=g(x)$ and $pr(Y=x)=h(x)$. Then:

$$pr[Y=x \mid U\leq g(y)]=\frac{g(x)h(x)}{\int_{x\neq 0} g(x)h(x)dx}=\frac{g(x)h(x)}{\int_x \frac{f(x)}{c}dx},\ c\neq 0$$

$$=\frac{g(x)h(x)}{\frac{1}{c}}=cg(x)h(x)=f(x).$$

The efficiency of Acceptance-Rejection is determined by the inequality $U\leq g(Y)$, where the efficiency$=pr[U\leq g(Y)]=\frac{1}{c}$.

Since the trails are independent, the probability of success in each trail is $P=\frac{1}{c}$. If N is a random variable represent the number of trials before a successful pair (U, Y), then N has geometric distribution with p.d.f.

$$pr[N = n] = p(1-p)^{n-1},\ n = 1,2,\ldots$$

$$= 0, \text{ e.w.}$$

and the expected number of trails is

$$\mathrm{E}(N) = \frac{1}{p} = c$$

The AR-Algorithm describes the necessary steps of generating random varieties by Acceptance-Rejection method.

AR-Algorithm

1) Generate U from $U(0,1)$.
2) Generate Y from $h(y)$.
3) If $U \le g(Y)$, deliver (we accept) Y=X as a random varieties generated from $f(x)$.
4) Go to step (1).
5) Stop.

Remark

For acceptance-Rejection method to be of practical interest, the following conditions must be satisfied:

1. It should be easy to generate a r.v. X from $h(x)$.
2. The efficiency (probability) of the procedure $\frac{1}{c}$ should be large, that is c should close to one.

Illustration of the Acceptance-Rejection method, we choose $c \ge 1$ such that $f(x) \le ch(x) = \phi(x)$.

The problem then is to find the function $\phi(x)$ and the function $h(x) = \frac{1}{c}\phi(x)$ from which the random variable can be easily generated.

3.4 Generating Random Varieties from Normal Distribution

3.4.1 Procedure N-1

This approach is due to Box and Muller (1958) [26]. Viz

If U_1, U_2 is a r.s. of size 2 from $U(0,1)$, then the r.v.'s $X_1 = (-2\ln U_1)^{1/2}\cos(2\pi U_2)$, $X_2 = (-2\ln U_1)^{1/2}\sin(2\pi U_2)$ represent a r.s. of size 2 from $N(0,1)$.

Proof

The joint distribution of U_1 and U_2 are

$$g(u_1,u_2) = 1,\ 0 < u_i < 1,\ i = 1,2$$

$$= 0, \text{ e.w.}$$

The functions $x_1 = (-2\ln u_1)^{1/2}\cos(2\pi u_2)$, $x_2 = (-2\ln u_1)^{1/2}\sin(2\pi u_2)$ that maps

$$A = \{(u_1,u_2): 0 < u_i < 1, i = 1,2\} \underset{onto}{\overset{1-1}{\rightarrow}} B = \{(x_1,x_2): -\infty < x_i < \infty, i = 1,2\}$$

with inverses $x_1^2 + x_2^2 = (-2\ln u_1)\cos^2(2\pi u_2) + (-2\ln u_1)\sin^2(2\pi u_2)$

$$= -2\ln u_1\left(\cos^2(2\pi u_2) + \sin^2(2\pi u_2)\right)$$

$$= -2\ln u_1$$

$$\ln u_1 = -\frac{1}{2}\left(x_1^2 + x_2^2\right) \text{ implies } u_1 = e^{-\frac{1}{2}\left(x_1^2 + x_2^2\right)}$$

$$\frac{x_2}{x_1} = \tan(2\pi u_2) \Rightarrow 2\pi u_2 = \tan^{-1}\left(\frac{x_2}{x_1}\right) \text{ implies } u_2 = \frac{1}{2\pi}\tan^{-1}\left(\frac{x_2}{x_1}\right)$$

The Jacobian of this transformation is

$$J=\frac{\partial(u_1,u_2)}{\partial(x_1,x_2)}=1\begin{vmatrix}\frac{\partial u_1}{\partial x_1} & \frac{\partial u_1}{\partial x_2}\\ \frac{\partial u_2}{\partial x_1} & \frac{\partial u_2}{\partial x_2}\end{vmatrix}=\begin{vmatrix}-x_1 e^{-\frac{1}{2}\left(x_1^2+x_2^2\right)} & -x_2 e^{-\frac{1}{2}\left(x_1^2+x_2^2\right)}\\ \frac{1}{2\pi}\frac{\frac{-x_2}{(x_1)^2}}{1+\left(\frac{x_2}{x_1}\right)^2} & \frac{1}{2\pi}\frac{\frac{1}{x_1}}{1+\left(\frac{x_2}{x_1}\right)^2}\end{vmatrix}$$

$$=\frac{-e^{-\frac{1}{2}\left(x_1^2+x_2^2\right)}}{2\pi\left[1+\left(\frac{x_2}{x_1}\right)^2\right]}-\frac{\left(\frac{x_2}{x_1}\right)^2 e^{-\frac{1}{2}\left(x_1^2+x_2^2\right)}}{2\pi\left[1+\left(\frac{x_2}{x_1}\right)^2\right]}=\frac{-e^{-\frac{1}{2}\left(x_1^2+x_2^2\right)}}{2\pi\left[1+\left(\frac{x_2}{x_1}\right)^2\right]}.\left[1+\left(\frac{x_2}{x_1}\right)^2\right]$$

$$=\frac{-e^{-\frac{1}{2}\left(x_1^2+x_2^2\right)}}{2\pi}$$

Then, the joint distn. of x_1 and x_2 is

$$f\left(x_1,x_2\right)=g\left[e^{-\frac{1}{2}\left(x_1^2+x_2^2\right)},\frac{1}{2\pi}\tan^{-1}\left(\frac{x_2}{x_1}\right)\right]|J|$$

$$=(1)\left|\frac{-e^{-\frac{1}{2}\left(x_1^2+x_2^2\right)}}{2\pi}\right|=\frac{1}{2\pi}e^{-\frac{1}{2}\left(x_1^2+x_2^2\right)},\ -\infty<x<\infty$$

$X=\left(X_1,X_2\right)$ distributed as a random vector of size 2 from $N(0,1)$. That is, $X_i \sim N(0,1)$, $i=1,2$.

Algorithm N-1

1) Generate U_1 and U_2 from $U(0,1)$.
2) Set $X_1=\left(-2\ln U_1\right)^{1/2}\cos\left(2\pi U_2\right)$, $X_2=\left(-2\ln U_1\right)^{1/2}\sin\left(2\pi U_2\right)$.
3) Deliver $\underset{\sim}{X}=(X_1,X_2)$ as a random vector of size 2 generated from $N(0,1)$.

A computer program as written in Appendix B1 for generating a sample of size n from normal distribution written in Pascal language using Microcomputer Pentium IV, CPU 1.7 GHz and the run size 2 is taken. The results of mean and variance of standard normal distribution with efficiency and time are tabulated in table (3.1).

Table (3.1): Values of Time, $\hat{\mu}$ and $\hat{\sigma}^2$ with Difference n-Samples Using N-1 Procedure

n	$\hat{\mu}$	$\hat{\sigma}^2$	Time	Average Time
4	0.066	0.997	00	10.9
6	0.035	0.957	05	
8	0.072	0.856	05	
10	0.058	0.951	05	
12	0.103	0.939	10	
14	0.023	0.997	10	
16	0.043	0.947	16	
18	0.053	0.973	21	
20	0.038	1.014	21	
22	0.062	0.981	21	

3.4.2 Procedure N-2

This procedure is based on the Acceptance-Rejection method. Let the r.v. X be distributed

$$f(x) = \sqrt{\frac{2}{\pi}} e^{-\frac{1}{2}x^2}, \quad 0 < x < \infty \qquad (3.4)$$

$$= 0, \text{ e.w.}$$

Since the standard normal distribution is symmetrical about zero, we can assign a random sign to the r.v. generated from eq.(3.4) and obtain an r.v.

from $N(0,1)$.

To generate an r.v. from (3.4) write $f(x)$ as

$f(x) = ch(x)g(x)$

Set $(x-1)^2 \geq 0 \Rightarrow x^2 - 2x + 1 \geq 0 \Rightarrow x^2 \geq 2x - 1$

$\frac{1}{2}x^2 \geq x - \frac{1}{2} \Rightarrow -\frac{1}{2}x^2 \leq \frac{1}{2} - x$ Implies $e^{-\frac{1}{2}x^2} \leq e^{\frac{1}{2}-x}$

$$f(x) = \sqrt{\frac{2}{\pi}} e^{-\frac{1}{2}x^2} \leq \sqrt{\frac{2}{\pi}} e^{\frac{1}{2}-x} = \phi(x)$$

$$= \sqrt{\frac{2e}{\pi}} e^{-x} = \phi(x) \tag{3.5}$$

$$ch(x) = \phi(x) \Rightarrow c = \sqrt{\frac{2e}{\pi}} \int_0^\infty e^{-x} dx = \sqrt{\frac{2e}{\pi}} \tag{3.6}$$

$$h(x) = \frac{\phi(x)}{c} = e^{-x}, \quad 0 < x < \infty \tag{3.7}$$

$$= 0, \text{ e.w.}$$

$$H(x) = \begin{cases} 0, & x \leq 0 \\ 1 - e^{-x}, & 0 < x < \infty \\ 1, & x = \infty \end{cases}$$

Set $u = H(x) \Rightarrow x = -\ln u$

$$g(x) = \frac{f(x)}{\phi(x)} = \frac{e^{-\frac{1}{2}x^2}}{e^{\frac{1}{2}-x}} = e^{-\frac{1}{2}\left(x^2 - 2x + 1\right)} = e^{-\frac{1}{2}(x-1)^2} \tag{3.8}$$

The efficiency of the method is equal to $\frac{1}{c} = \sqrt{\frac{\pi}{2e}} \approx 0.76$.

The acceptance condition $U \leq g(Y)$ is $U \leq \exp\left[-(Y-1)^2/2\right]$ (3.9)

Algorithm N-2

1) Generate U_1 and U_2 from $U(0,1)$.
2) Set $Y = -\ln U_2$.
3) If $U > g(Y) = \exp\left[-(Y-1)^2/2\right]$, go to step (1).
4) Generate U from $U(0,1)$.
5) If $U \le \frac{1}{2}$, Deliver $Y = X$ as a r.v. generated from $N^+(0,1)$.
6) Deliver $Y = -X$ as a r.v. generated from $N^-(0,1)$.

A computer program is made in Appendix B2 for computation the values of efficiency with run size of 2 is taken. The results of mean and variance of standard normal distribution with efficiency are displayed in Table (3.2) relative to the theoretical efficiency value of 0.760.

Table (3.2): Values of c, $\hat{\mu}$ and $\hat{\sigma}^2$ with Difference n-Samples Using N-2 Procedure

N	$\hat{\mu}$	$\hat{\sigma}^2$	Simulation Efficiency	Theoretical Efficiency	Error	Average Error
4	0.169	0.929	0.798	0.760	0.038	0.014
6	0.170	0.944	0.772		0.012	
8	0.174	0.983	0.769		0.009	
10	0.211	0.955	0.797		0.037	
12	0.202	1.004	0.776		0.016	
14	0.181	0.916	0.770		0.010	
16	0.178	0.980	0.758		0.002	
18	0.211	0.993	0.757		0.003	
20	0.204	0.960	0.756		0.004	

3.4.3 Procedure N-3

This procedure relies on the Central Limit Theorem, which says that if $X_1, X_2, \ldots, X_n$ be a r.s. of size n from any distribution (continuous or discrete) having mean μ and variance σ^2 with existence of moment generating function $M(t)$, then the r.v.

$$X = \frac{\sqrt{n}\left(\bar{X} - \mu\right)}{\sigma} \tag{3.10}$$

Converges asymptotically with n to $N(0,1)$. Consider the particular case when all X_i, $i = 1,2,\ldots,n$ are from $U(0,1)$. We find that

$$\mu = \mathrm{E}(u) = \frac{1}{2}.$$

$$\sigma^2 = \mathrm{var}(u) = \frac{1}{12}.$$

$$X = \frac{\sqrt{n}\left(\bar{U} - \frac{1}{2}\right)}{\sqrt{\frac{1}{12}}} = \sqrt{12n}\left(\bar{U} - \frac{1}{2}\right) \tag{3.11}$$

A good approximation can already be obtained for $n = 12$.

Algorithm N-3

1) Generate $U_1, U_2, \ldots, U_n$ from $U(0,1)$.
2) Set $\bar{U} = \frac{1}{n}\sum_{i=1}^{n} U_i$.
3) Set $X = \sqrt{12n}\left(\bar{U} - \frac{1}{2}\right)$.
4) Deliver X as a random vector from $N(0,1)$.

A computer program as written in Appendix B3 for generating a sample of size n from normal distribution written in Pascal language using Microcomputer Pentium IV, CPU 1.7 GHz and the run size 2 is taken. The

results of mean and variance of standard normal distribution with efficiency and time are tabulated in table (3.3).

Table (3.3): Values of Time, $\hat{\mu}$ and $\hat{\sigma}^2$ with Difference n-Samples Using N-3 Procedure

n	$\hat{\mu}$	$\hat{\sigma}^2$	Time	Average Time
2	0.040	0.978	00	24.1
4	0.012	0.987	05	
6	0.007	0.655	10	
8	0.174	1.064	21	
10	0.110	0.905	21	
12	0.158	1.125	27	
14	0.024	0.997	32	
16	0.079	0.950	38	
18	0.102	0.894	38	
20	0.087	1.010	49	

3.4.4 Procedure N-4

This procedure due to Tocher (1963). The technique for this procedure from generating from $N(0,1)$ is made by approximating the p.d.f.

$$\sqrt{\frac{2}{\pi}} e^{-\frac{x^2}{2}} \approx \frac{2ke^{-kx}}{\left(1+e^{-kx}\right)^2} \tag{3.12}$$

Where $0 < x < \infty$ and $k = \sqrt{\frac{8}{\pi}}$.

Proof

By inverse transform method

$$F(x)=\int_0^x f\ (t)dt=2\int_0^x \frac{ke^{-kt}}{\left(1+e^{-kt}\right)^2}dt$$

$$=\frac{2}{1+e^{-kt}}\Bigg|_0^x$$

$$F(x)=\begin{cases}0, & x\le 0\\ \dfrac{2}{1+e^{-kx}}-1, & 0<x<\infty\\ 1, & x=\infty\end{cases}$$

set $u=F(x)\Rightarrow \dfrac{2}{1+e^{-kx}}-1=u$

$$\frac{u+1}{2}=\frac{1}{1+e^{-kx}}\Rightarrow 1+e^{-kx}=\frac{2}{u+1}$$

$$e^{-kx}=\frac{2}{u+1}-1=\frac{1-u}{1+u}$$

$$-kx=\ln\left(\frac{1-u}{1+u}\right)\Rightarrow x=-\frac{1}{k}\ln\left(\frac{1-u}{1+u}\right)$$

implies $x=\dfrac{1}{k}\ln\left(\dfrac{1+u}{1-u}\right)$.

Algorithm N-4

1) Generate U_1 and U_2 from $U(0,1)$.
2) Set $X=\sqrt{\dfrac{\pi}{8}}\ln\left(\dfrac{1+U_1}{1-U_2}\right)$.
3) If $U_2\le\dfrac{1}{2}$, Deliver $X=Y$ as a r.v. generated from $N^{+}(0,1)$.
4) Deliver $X=-Y$ as a r.v. generated from $N^{-}(0,1)$.

A computer program as written in Appendix B4 for generating a sample of size n from normal distribution written in Pascal language using Microcomputer Pentium IV, CPU 1.7 GHz and the run size 2 is taken. The

results of mean and variance of standard normal distribution with efficiency and time are tabulated in table (3.4).

Table (3.4): Values of Time, $\hat{\mu}$ and $\hat{\sigma}^2$ with Difference n-Samples Using N-4 Procedure

n	$\hat{\mu}$	$\hat{\sigma}^2$	Time	Average Time
2	0.285	1.077	00	27.5
4	0.370	1.033	00	
6	0.304	0.742	16	
8	0.448	0.714	21	
10	0.365	1.008	27	
12	0.343	0.951	32	
14	0.324	0.905	38	
16	0.348	1.091	43	
18	0.530	1.115	49	
20	0.340	0.862	49	

CHAPTER

2

An Approximation to the Cumulative Distribution Function of Normal Distribution

2.1 Introduction

The importance of good numerical integration schemes is evident. There are many deterministic quadrature formulas can be found throughout the literature for computation of ordinary integrals with well behaved integrands. It is often more convenient to compute such integrals by Monte Carlo method, which, although less accurate than conventional quadrature formulas, but is much simpler to use.

This chapter involves three numerical procedures and two reduction techniques for approximating of the N(0,1) c.d.f. of eq.(1.4). The three numerical procedure namely Trapezoidal, Simpson and Gauss-Quadrature Rules are discussed in sections 2.2, 2.3 and 2.4 respectively. While the two reduction techniques namely Hit or Miss Monte Carlo and Sample Mean Monte Carlo Rules are discussed in sections 2.5 and 2.6 respectively.

For simplicity and computing purposes, we write eq.(1.4) as

$$\Phi(t) = \frac{1}{2} + \frac{1}{\sqrt{2\pi}} \int_0^x e^{-\frac{1}{2}t^2} dt \qquad (2.1)$$

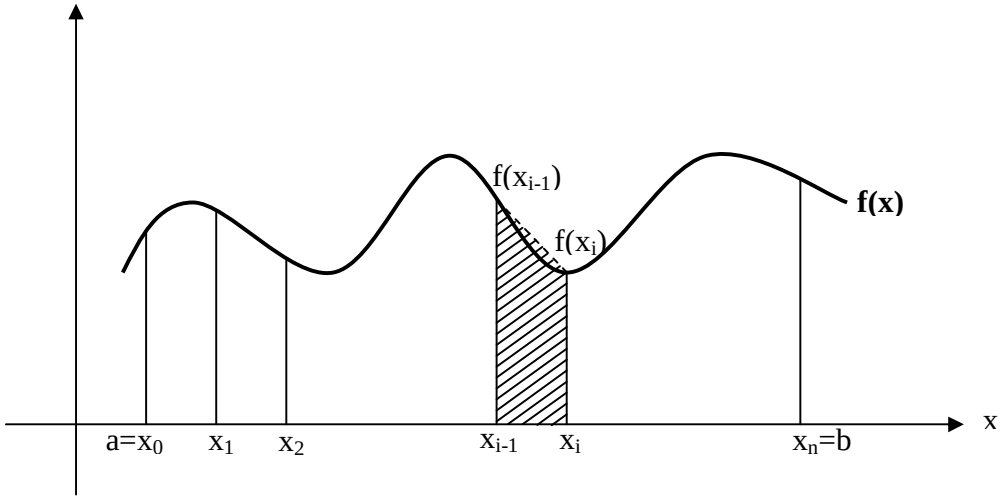
Finally section 2.7, the results of the five methods are tabulated and compared with the literature table(1.1) given in chapter one in order to obtain the more efficient and accurate procedure.

2.2 Approximation by Trapezoidal Rule [12]

Trapezoidal method is used for approximating the area under a curve by series of trapezoids. It has been shown theoretically that using an infinite number of trapezoids give prefect accuracy, but rounding of error will give us problems.
The trapezoidal rule procedure can be illustrated as follows:
Suppose we wish to approximate the integral

$I = \int_a^b f(x)dx$, by using trapezoidal rule (2.2)

We divide the interval from a to b into n equal parts as shown in figure(2.1), where the boundaries of the trapezoids are $x_0, x_1, \ldots, x_n$.



Figure(2.1): Integration by the Trapezoidal Rule.

Let $h = \frac{b-a}{n}$ be the width of the ith trapezoid that lies between x_{i-1} and x_i whose heights at the left and right side are respectively $f(x_{i-1})$ and $f(x_i)$.

The area of the ith trapezoid is:

$$A_i = \frac{h}{2}[f(x_{i-1}) + f(x_i)]$$

The total area of all n trapezoids is the trapezoidal approximation to the integral I. that is:

$$I \approx \sum_{i=1}^{n} A_i \ .$$

$$\cong \frac{h}{2}[f(x_0) + f(x_1)] + \frac{h}{2}[f(x_1) + f(x_2)] + \frac{h}{2}[f(x_2) + f(x_3)] +$$

$$\ldots + \frac{h}{2}[f(x_{n-1}) + f(x_n)]$$

$$\cong \frac{h}{2}\left[f(x_0) + f(x_n) + 2\sum_{i=1}^{n-1} f(x_i)\right] \qquad (2.3)$$

Composite Trapezoid Rule Algorithm

Step1: Input a, b (Interval of integration)

n (Number of subintervals)

Step2: Define f(x) (integrand)

Step3: Set $h = \frac{(b-a)}{n}$

Step4: sum=0

Step5: For i=1 to n-1

Step6: $x = a + i * h$

Step6: sum=sum+2f(x)

Step7: $g = \frac{h}{2}(f(a) + sum + f(b))$

Step8: Output g

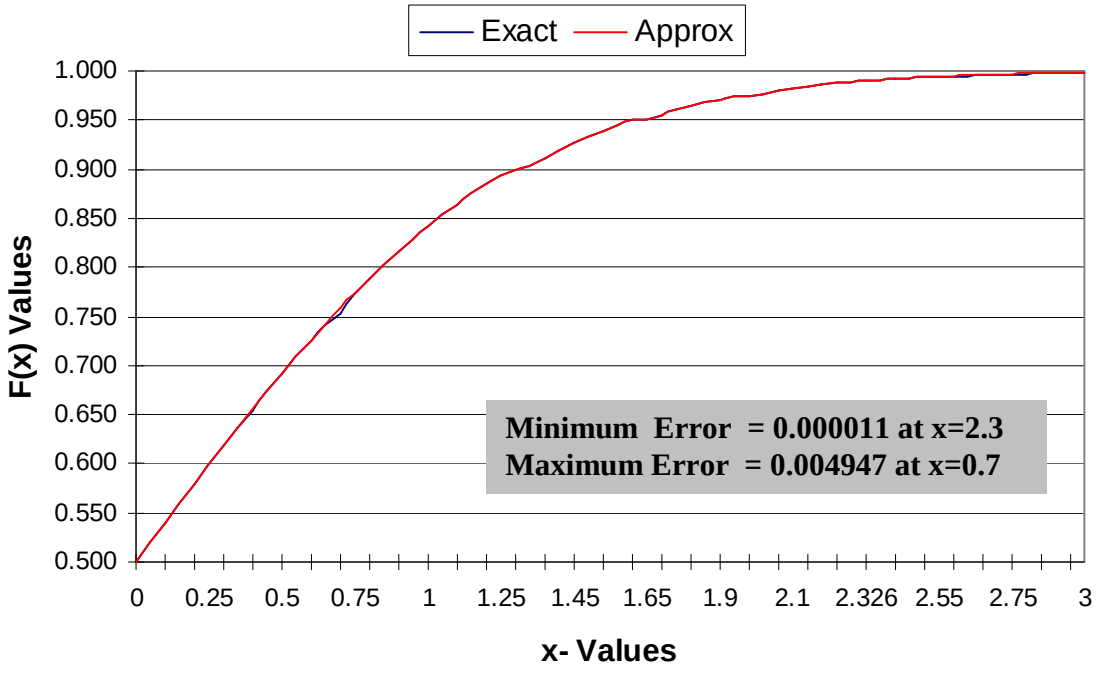
Step9: Stop

Appendix (A1) involves a computer program written in Pascal language using composite trapezoid rule for approximating the integral side of eq.(2.1). The x-values of the upper limit of the integral is taken from the normal distribution table (1.1). Table(2.1) gives together the approximate and the errors value of F(x) while figure(2.2) shows the difference between the exact and approximate results of the N(0,1) c.d.f. by Trapezoidal method. Also this figure shows the minimum and maximum errors between the exact and approximate result.

Table(2.1): Approximation by Trapezoidal Method

$\Pr(X \le x)$					
x	**Approx.**	**Error**	**x**	**Approx.**	**Error**
0.00	0.500000	0.000000	1.60	0.944822	0.000178
0.05	0.519939	0.000061	1.645	0.949633	0.000367
0.10	0.539828	0.000172	1.65	0.950146	0.000854
0.15	0.559617	0.000383	1.70	0.955049	0.000049
0.20	0.579257	0.000257	1.75	0.959559	0.000441
0.25	0.598701	0.000299	1.80	0.963686	0.000314
0.30	0.617903	0.000097	1.85	0.967463	0.000537
0.35	0.636817	0.000183	1.90	0.970909	0.000091
0.40	0.655402	0.000402	1.95	0.974044	0.000044
0.45	0.673617	0.000383	1.960	0.974636	0.000364
0.50	0.691426	0.000426	2.00	0.976890	0.000110

0.55	0.708793	0.000207	2.05	0.979468	0.000532
0.60	0.725687	0.000313	2.10	0.981796	0.000204
0.65	0.742080	0.000080	2.15	0.983895	0.000105
0.70	0.757947	0.004947	2.20	0.985782	0.000218
0.75	0.773267	0.000267	2.25	0.987475	0.000525
0.80	0.788021	0.000021	2.30	0.988989	0.000011
0.85	0.802195	0.000195	2.326	0.989712	0.000288
0.90	0.815778	0.000222	2.35	0.990341	0.000659
0.95	0.828762	0.000238	2.40	0.991545	0.000455
1.00	0.841143	0.000143	2.45	0.992615	0.000385
1.05	0.852919	0.000081	2.50	0.993563	0.000437
1.10	0.864092	0.000092	2.55	0.994401	0.000599
1.15	0.874667	0.000333	2.576	0.994797	0.000203
1.20	0.884651	0.000349	2.60	0.995141	0.000141
1.25	0.894053	0.000053	2.65	0.995792	0.000208
1.282	0.899771	0.000229	2.70	0.996363	0.000637
1.30	0.902886	0.000114	2.75	0.996864	0.000136
1.35	0.911163	0.000163	2.80	0.997301	0.000301
1.40	0.918901	0.000099	2.85	0.997682	0.000318
1.45	0.926116	0.000116	2.90	0.998014	0.000014
1.50	0.932828	0.000172	2.95	0.998302	0.000302
1.55	0.939057	0.000057	3.00	0.998551	0.000449



Figure(2.2): Shows the difference between the exact and approximate solutions to the c.d.f. $\Phi(x)$ by Trapezoidal Rule.

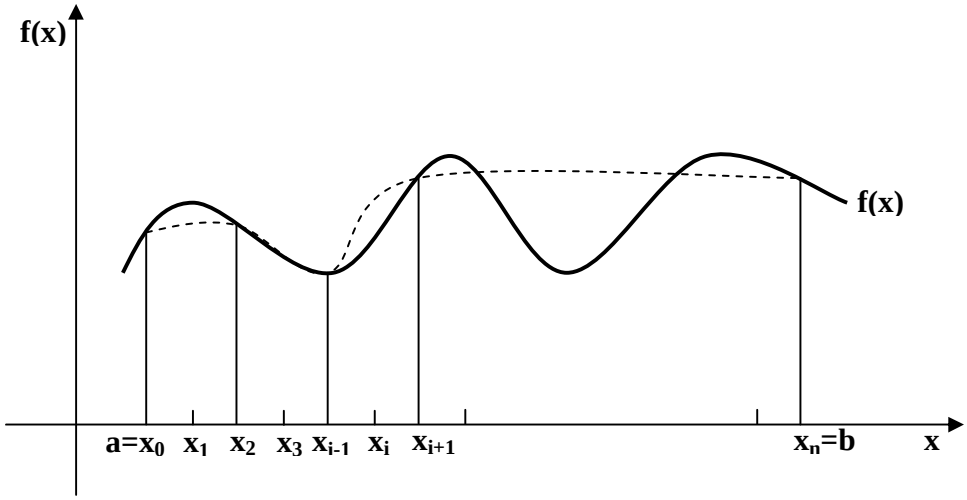
2.3 Approximation by Simpson Rule [17]

In Simpson method, we try to approximate $\int_a^b f(x)dx$ by a series of parabolic segments hoping that parabola will more closely much to a given curve of f(x) than it would be straight line in the trapezoidal method.

Simpson rule (or Simpson $\frac{1}{3}$ rule) is given by the equation $A_1 = \frac{h}{3}(f_0 + 4f_1 + f_2) + O(h^5)$ where A_1 denotes the area under the graph of f(x) from the point x_0 to the point x_2 and $h = \frac{x_n - x_0}{n}$, (n=1, 2, …). This equation calculates the integral over two segments of integration. Repeated

application of Simpson $\frac{1}{3}$ rule over segment pairs of segments, and summation of all the formulas over the total interval, gives the multiple segments Simpson $\frac{1}{3}$ Rule:

$$A = \sum_{i=1}^{n} A_i = \frac{h}{3}\left[f_0 + 4\sum_{i=1}^{n-1} f_i + 2\sum_{i=2}^{n-2} f_i + f_n \right] + O(h^5) \qquad (2.4)$$



Figure(2.3): Integration by the Simpson Rule.

Since, Simpson $\frac{1}{3}$ Rule fits pairs of segments, the total intervals must be divided into an even number of segments. The first summation terms in eq.(2.4) sums up to odd-subscripted terms and the second summation odds up to even-subscripted terms, the order of the error of the multiple-segment Simpson $\frac{1}{3}$ rule was reduced by one order of magnitude of $o(h^4)$ for the same reason as in the trapezoidal rule. For more details see [15, 25].

Simpson Rule Algorithm

Step1: Input a, b (Interval of integration)

n (Number of subintervals)

Step2: Define f(x) (integrand)

Step3: Set $h = \frac{(b-a)}{n}$

Step4: sum1=0, sum2=0

Step5: For i=1 to n-1

Step6: $x = a + i * h$

Step7: If i is even then sum1=sum1+2f(x)

Else sum2=sum2+4f(x)

Step8: $g \cong \frac{h}{3}[f(a) + sum1 + sum2 + f(b)]$

Step9: Output g

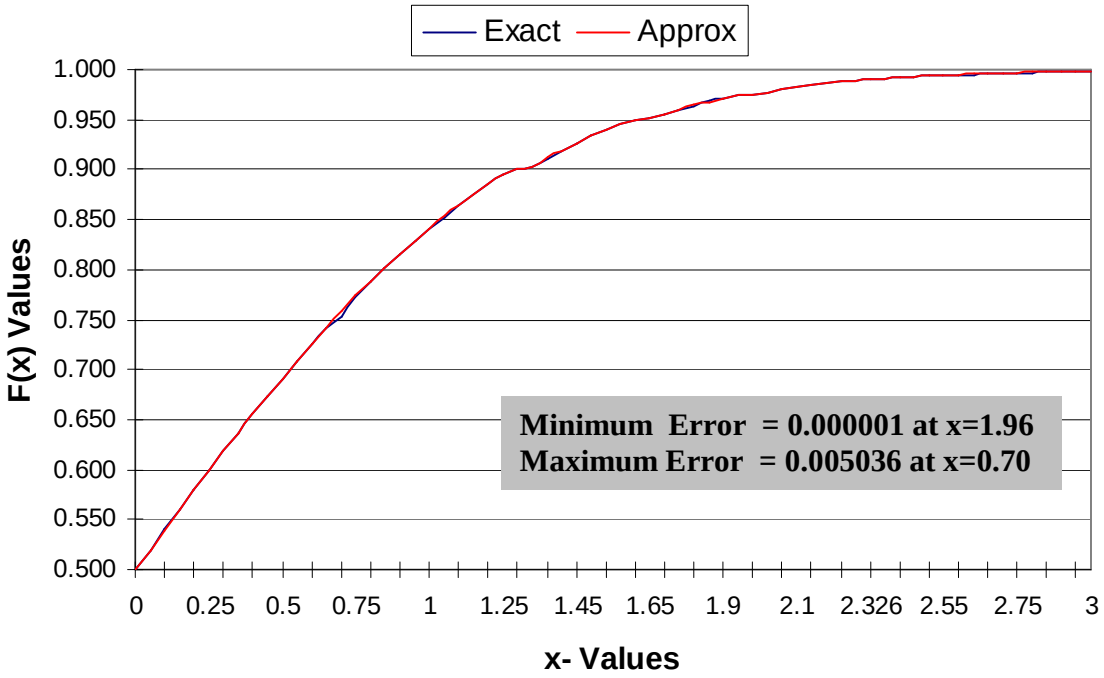
Step10: Stop

Appendix (A2) involves a computer program written in Pascal language using Simpson rule for approximating the integral side of eq.(2.1). The x-values of the upper limit of the integral is taken from the normal distribution table (1.1). Table (2.2) gives together the approximate and the errors value of F(x). while figure(2.3) shows the difference between the exact and approximate results of the N(0,1) c.d.f. by Simpson method. Also this figure shows the minimum and maximum errors between the exact and approximate result.

Table(2.2): Approximation by Simpson Method

Pr (X ≤ x)					
x	**Approx.**	**Error**	**x**	**Approx.**	**Error**
0.00	0.500000	0.000000	1.60	0.945201	0.000201
0.05	0.519939	0.000061	1.645	0.950015	0.000015
0.10	0.539828	0.000172	1.65	0.950529	0.000471
0.15	0.559618	0.000382	1.70	0.955435	0.000435

0.20	0.579260	0.000260	1.75	0.959941	0.000059
0.25	0.598706	0.000294	1.80	0.964069	0.000069
0.30	0.617911	0.000089	1.85	0.967843	0.000157
0.35	0.636831	0.000169	1.90	0.971283	0.000283
0.40	0.655422	0.000422	1.95	0.974411	0.000411
0.45	0.673645	0.000355	1.960	0.975001	0.000001
0.50	0.691462	0.000462	2.00	0.977249	0.000249
0.55	0.708840	0.000160	2.05	0.979817	0.000183
0.60	0.725747	0.000253	2.10	0.982134	0.000134
0.65	0.742154	0.000154	2.15	0.984221	0.000221
0.70	0.758036	0.005036	2.20	0.986095	0.000095
0.75	0.773373	0.000373	2.25	0.987773	0.000227
0.80	0.788145	0.000145	2.30	0.989274	0.000274
0.85	0.802338	0.000338	2.326	0.989988	0.000012
0.90	0.815940	0.000060	2.35	0.990611	0.000389
0.95	0.828944	0.000056	2.40	0.991800	0.000200
1.00	0.841345	0.000345	2.45	0.992854	0.000146
1.05	0.853141	0.000141	2.50	0.993787	0.000213
1.10	0.864334	0.000334	2.55	0.994611	0.000389
1.15	0.874928	0.000072	2.576	0.994999	0.000001
1.20	0.884931	0.000069	2.60	0.995335	0.000335
1.25	0.894351	0.000351	2.65	0.995972	0.000028
1.282	0.900079	0.000079	2.70	0.996529	0.000471
1.30	0.903200	0.000200	2.75	0.997017	0.000017
1.35	0.911492	0.000492	2.80	0.997441	0.000441
1.40	0.919244	0.000244	2.85	0.997810	0.000190
1.45	0.926471	0.000471	2.90	0.998131	0.000131
1.50	0.933193	0.000193	2.95	0.998408	0.000408
1.55	0.939430	0.000430	3.00	0.998647	0.000353



Figure(2.3): Shows the difference between the exact and approximate solutions to the c.d.f. $\Phi(x)$ by Simpson Rule.

2.4 Approximation by Gaussian Quadrature Method [25]

To introduce the ideas involved in Gaussian Quadrature, we consider the more general integral $\int_a^b w(x)f(x)dx$, where $w(x) > 0$ is a weight function.

We are interested only in the case $w(x) = 1$ but different choices do play very important roles in numerical integration and a discussion of these can be found in [17]. The orthogonal polynomials corresponding to this weight function are known as the Legendre polynomials. Quadrature using these polynomials is called Gauss-Legendre Quadrature or, simply, Gaussian Quadrature which have the general formula

$$\int_{-1}^{1} f(x)dx \cong \sum_{i=0}^{n} w_i f(x) \qquad (2.5)$$

The coefficients $w_i\ (i = 0,1,2,\ldots,n)$ could be calculated, but this is not necessary because they, and the points x_i, have already been tabulated for a large value of n see [12, 15]. Some of the roots of the Legendre polynomials and the corresponding weights are used. We need in this method transforming the interval [a, b] in to [-1, 1], by using the simple linear transformation.

$T = \left[\frac{1}{(b-a)}\right](2x - a - b)$ which provided $b > a$, the Legendre polynomials reduce to approximate

$$\int_{-1}^{1} \frac{(b-a)}{2} f\left(\frac{(b-a)t + (a+b)}{2}\right)dt$$

Where f is any function that can be evaluated at the required region of points.

Gaussian Quadrature Rule Algorithm

Step1: Input a, b (Interval of integration)

n≤6 (Number of subintervals)

Step2: Define f(x) (integrand)

Initialize Array x(n, i), w(n, i) for the Gauss nodes and weights in table(2.3) as shown below, x(n, i) is the ith nonnegative node for the Gauss n-point formula, and w(n, i) is the corresponding weight.

Step3: Set h:= (b-a)/2

m:= (a+b)/2

x:= h x(n, 1)

Step4: If n is odd then g: =h w(n, 1)f(x)

Else g: =h w(n, 1)(f(-x+m)+f(x+m))

Step5: For $i := 2$ *to* $\left[\frac{n+1}{2}\right]$ set x:= h x(n, i)

Step6: g:= g+h w(n, i) (f (-x+m) + f (x+m))

Step7: Output g

Step8: Stop

Table(2.3): Shows Numerical Values of the Gauss Nodes and Weights

N	$\pm x_i$	w_i
2	0.577350269189626	1.000000000000000
3	0.000000000000000	0.888888888888888
	0.774596669241483	0.555555555555556
4	0.339981043584856	0.652145154862546
	0.861136311594053	0.347854845137454
5	0.000000000000000	0.568888888888889
	0.538469310105683	0.478628670599366
	0.906179845938664	0.236926885056189
6	0.238619186083197	0.467913934572691
	0.661209386466265	0.360761573048139
	0.932469514203152	0.171324492379170
8	0.183434642495650	0.362683783378362
	0.525532409916329	0.313706645877887
	0.796666477413627	0.222381034453374
	0.960289856497536	0.101228536290376
10	0.148874338981631	0.295524224714753
	0.433395394129247	0.269266719309996
	0.679409568299024	0.219086362515982
	0.865063366688985	0.149451349150581
	0.973906528517172	0.066671344308688

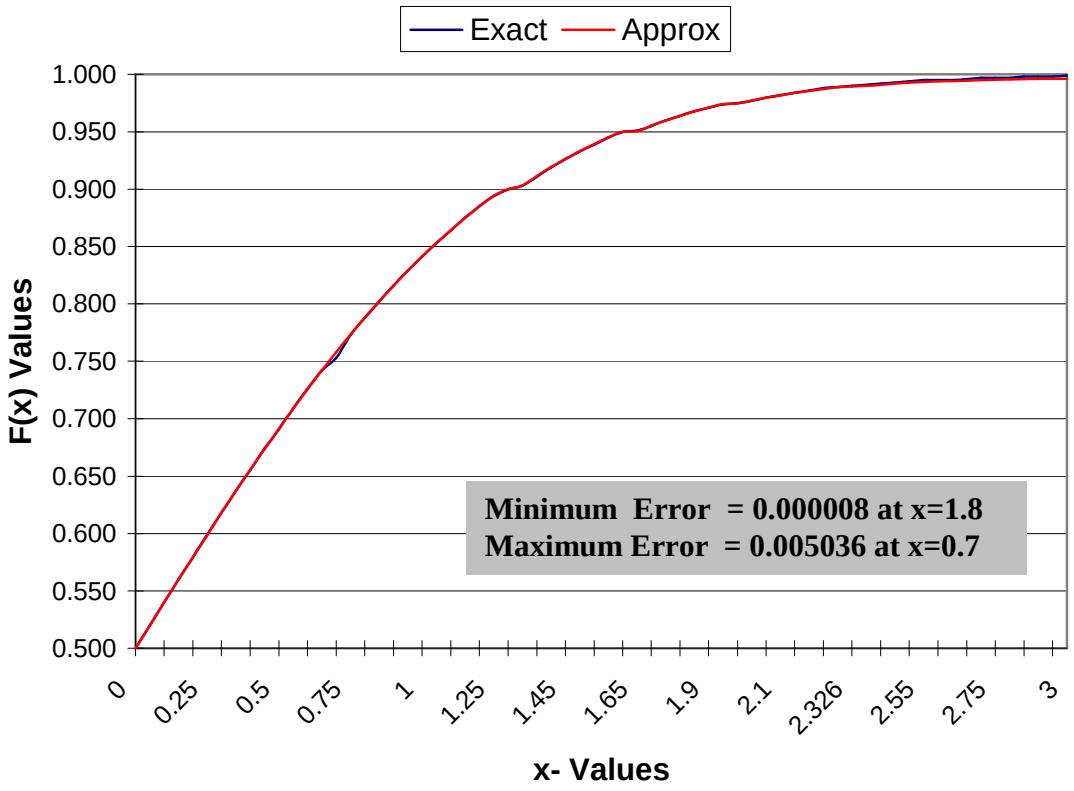
12	0.125233408511469	0.249147045813403
	0.367831498998180	0.233492536538355
	0.587317954286617	0.203167426723066
	0.769902674194305	0.160078328543346
	0.904117256370475	0.106939325995318
	0.981560634246719	0.047175336386512

Appendix (A3) involves a computer program written in Pascal language using Gaussian Quadrature Rule for approximating the integral side of eq.(2.1). The x-values of the upper limit of the integral is taken from the normal distribution table (1.1). Table (2.4) gives together the approximate and the errors value of F(x). while figure(2.4) shows the difference between the exact and approximate results of the N(0,1) c.d.f. by Gaussian Quadrature Method. Also this figure shows the minimum and maximum errors between the exact and approximate result.

Table (2.4): Approximation by Gaussian Quadrature Method

Pr (X ≤ x)					
x	**Approx.**	**Error**	**x**	**Approx.**	**Error**
0.00	0.500000	0.000000	1.60	0.945176	0.000176
0.05	0.519939	0.000061	1.645	0.949982	0.000018
0.10	0.539828	0.000172	1.65	0.950495	0.000505
0.15	0.559618	0.000382	1.70	0.955389	0.000389
0.20	0.579260	0.000260	1.75	0.959881	0.000119
0.25	0.598706	0.000294	1.80	0.963992	0.000008
0.30	0.617911	0.000089	1.85	0.967744	0.000256

0.35	0.636831	0.000169	1.90	0.971158	0.000158
0.40	0.655422	0.000422	1.95	0.974256	0.000256
0.45	0.673645	0.000355	1.960	0.974840	0.000160
0.50	0.691462	0.000462	2.00	0.977059	0.000059
0.55	0.708840	0.000160	2.05	0.979586	0.000414
0.60	0.725747	0.000253	2.10	0.981856	0.000144
0.65	0.742154	0.000154	2.15	0.983889	0.000111
0.70	0.758036	0.005036	2.20	0.985702	0.000298
0.75	0.773373	0.000373	2.25	0.987313	0.000687
0.80	0.788145	0.000145	2.30	0.988737	0.000263
0.85	0.802338	0.000338	2.326	0.989410	0.000590
0.90	0.815940	0.000060	2.35	0.989991	0.001009
0.95	0.828945	0.000055	2.40	0.991088	0.000912
1.00	0.841346	0.000346	2.45	0.992043	0.000957
1.05	0.853142	0.000142	2.50	0.992867	0.001133
1.10	0.864335	0.000335	2.55	0.993574	0.001426
1.15	0.874929	0.000071	2.576	0.993900	0.001100
1.20	0.884931	0.000069	2.60	0.994175	0.000825
1.25	0.894350	0.000350	2.65	0.994680	0.001320
1.282	0.900078	0.000078	2.70	0.995100	0.001900
1.30	0.903199	0.000199	2.75	0.995442	0.001558
1.35	0.911490	0.000490	2.80	0.995717	0.001283
1.40	0.919239	0.000239	2.85	0.995932	0.002068
1.45	0.926463	0.000463	2.90	0.996095	0.001905
1.50	0.933181	0.000181	2.95	0.996212	0.001788
1.55	0.939412	0.000412	3.00	0.996291	0.002709



Figure(2.4): Shows the difference between the exact and approximate solutions to the c.d.f. $\Phi(x)$ by Gaussian Rule.

2.5 Approximation by Hit or Miss Monte Carlo Method [27]

In this section, we consider a simple technique for computing the one-dimensional integral:

$$I = \int_a^b f(x)dx \tag{2.6}$$

By Monte Carlo method. Viz

For simplicity we assume that the integrand f(x) is bounded $0 \le f(x) \le c$, $a \le x \le b$. Let Ω denote the rectangle as shown in Figure(2.5)

$\Omega = \{(x, y): a \le x \le b, 0 \le y \le c\}$.

Let (X, Y) be a random vector uniformly distributed over the rectangle Ω with joint p.d.f.

$$g(x,y)=\begin{cases}\dfrac{1}{c(b-a)}, & (x,y)\in\Omega\\ 0, & otherwise\end{cases} \tag{2.7}$$

Let P be the probability that the random vector (X, Y) falls within the area under the curve f(x), and let S= {(x, y): y ≤ f(x)}. The area under the curve f(x) is:

$$\text{Area under f(x) = area S}=\int_a^b f(x)dx\ ,$$

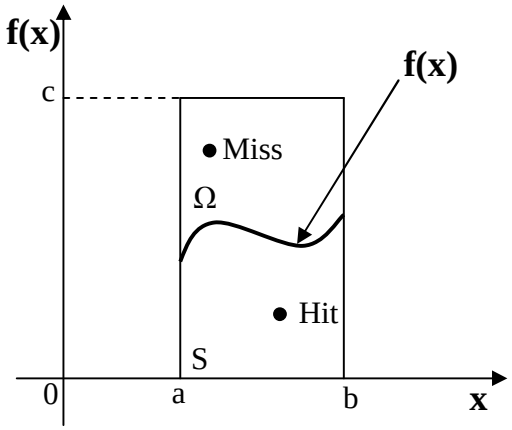


Figure (2.5): Graphical Representation of the Hit or Miss Monte Carlo method.

We obtain

$$P=\frac{area\ S}{area\ \Omega}=\frac{\int_a^b f(x)dx}{c(b-a)}=\frac{I}{c(b-a)} \tag{2.8}$$

Assume that N independent random vectors $(X_1,Y_1),(X_2,Y_2),\ldots,(X_n,Y_n)$ are generated. Then the probability P can be estimated by:

$$\hat{P} = \frac{No.of\ \ Hits}{Total\ \ no.of\ \ trials} = \frac{N_H}{N} \tag{2.9}$$

Where N_H is the number of occasions on which $f(x_i) \geq y_i$, i=1, 2, …, N, that is, the number of "hits" and N-N_H is the number of "misses"; we score a miss if

$f(x_i) < y_i$, i=1, 2, …, N.

It follows if $P \approx \hat{P}$ then from eqs.(2.8) and (2.9) that the integral I can be estimated by:

$$\frac{I}{c(b-a)} \approx \frac{N_H}{N} \text{ implies } I \approx \theta_1 = \frac{c(b-a)N_H}{N} \tag{2.10}$$

In other words we estimate the integral I by sampling N from the distn. of eq.(2.7), count the number N_H of hits and apply eq.(2.10).

Hit or Miss Monte Carlo Method Algorithm

Step1: Input a, b and c

Step2: Generate a sequence $\{U_j\}_{j=1}^{2N}$ of 2N random numbers.

Step3: Arrange the random numbers into N pairs

$(U_1, U_1'), (U_2, U_2'), \ldots, (U_N, U_N')$ in any fashion such that each random U_1 is used exactly once.

Step4: Set $X_i = a + U_i(b-a)$ and compute $f(x_i)$, i=1, 2, …, N.

Step5: Count the number of cases N_H for which $f(x_i) > cU_i'$.

Step6: Estimate the integral I by

$$\theta_1 = c(b-a)\frac{N_H}{N}.$$

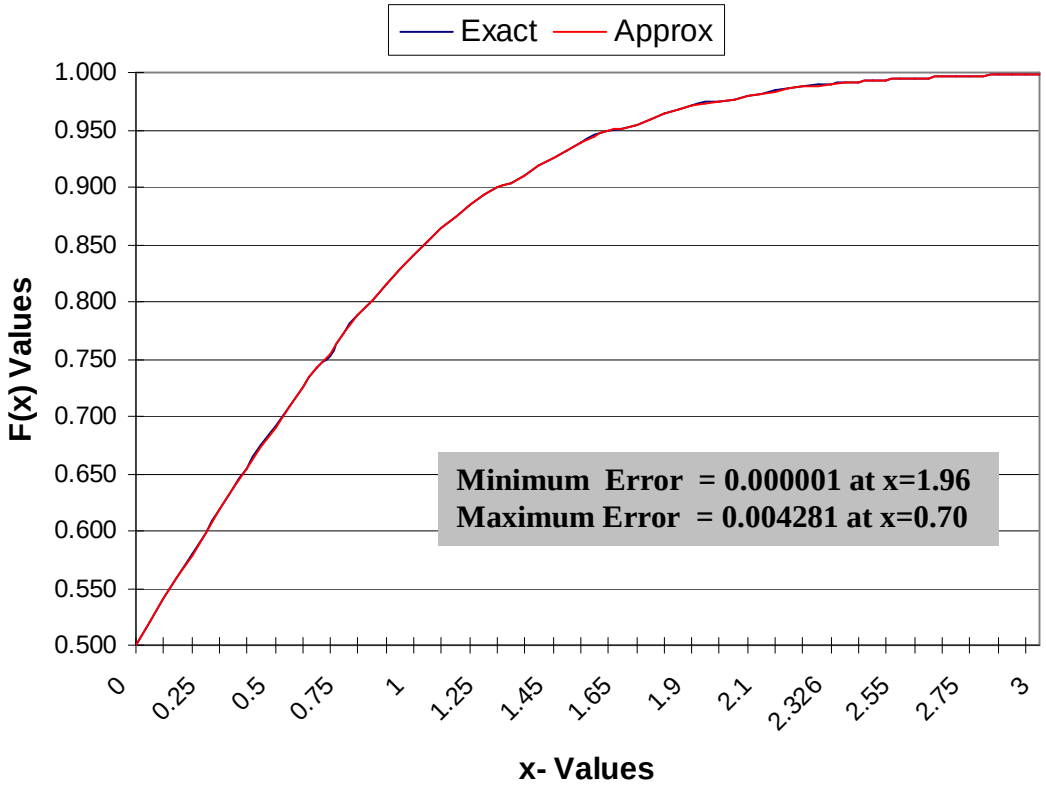
Appendix (A4) involves a computer program written in Pascal language using Hit or Miss Monte Carlo Method for approximating the integral side of eq.

(2.1). The x-values of the upper limit of the integral is taken from the normal distribution table (1.1). Table (2.5) gives together the approximate and the errors value of F(x). The run size is 1000 is made. While figure(2.6) shows the difference between the exact and approximate results of the N(0,1) c.d.f. by Hit or Miss Monte Carlo Method. Also this figure shows the minimum and maximum errors between the exact and approximate result.

Table (2.5): Approximation by Hit or Miss Monte Carlo Method

Pr (X ≤ x)					
x	**Approx.**	**Error**	**x**	**Approx.**	**Error**
0.00	0.500000	0.000000	1.60	0.944849	0.000151
0.05	0.519945	0.000055	1.645	0.949989	0.000011
0.10	0.539845	0.000155	1.65	0.950647	0.000353
0.15	0.559656	0.000344	1.70	0.954674	0.000326
0.20	0.578766	0.000234	1.75	0.959956	0.000044
0.25	0.598735	0.000265	1.80	0.963948	0.000052
0.30	0.617920	0.000080	1.85	0.967882	0.000118
0.35	0.636848	0.000152	1.90	0.970802	0.000198
0.40	0.654620	0.000380	1.95	0.973712	0.000288
0.45	0.673681	0.000320	1.960	0.974999	0.000001
0.50	0.690607	0.000393	2.00	0.976826	0.000174
0.55	0.708864	0.000136	2.05	0.979872	0.000128
0.60	0.725785	0.000215	2.10	0.981906	0.000094
0.65	0.741869	0.000131	2.15	0.983845	0.000155
0.70	0.753719	0.004281	2.20	0.985934	0.000067
0.75	0.772683	0.000317	2.25	0.987841	0.000159
0.80	0.787877	0.000123	2.30	0.988808	0.000192
0.85	0.801713	0.000287	2.326	0.989992	0.000008
0.90	0.815949	0.000051	2.35	0.990747	0.000253
0.95	0.828952	0.000048	2.40	0.991870	0.000130

1.00	0.840724	0.000276	2.45	0.992905	0.000095
1.05	0.852887	0.000113	2.50	0.993862	0.000138
1.10	0.863733	0.000267	2.55	0.994747	0.000253
1.15	0.874942	0.000058	2.576	0.994999	0.000001
1.20	0.884945	0.000055	2.60	0.994782	0.000218
1.25	0.893719	0.000281	2.65	0.995982	0.000018
1.282	0.899937	0.000063	2.70	0.996694	0.000306
1.30	0.902840	0.000160	2.75	0.996989	0.000011
1.35	0.910606	0.000394	2.80	0.996713	0.000287
1.40	0.918805	0.000195	2.85	0.997877	0.000124
1.45	0.925647	0.000353	2.90	0.997915	0.000085
1.50	0.932855	0.000145	2.95	0.997735	0.000265
1.55	0.938678	0.000323	3.00	0.998771	0.000229



Figure(2.6): Shows the difference between the exact and approximate solutions to the c.d.f. $\Phi(x)$ by Hit or Miss Monte Carlo Method

2.6 Approximation by the Sample-Mean Monte Carlo Method [27]

Another way of computing the integral

$$I = \int_a^b g(x)dx$$

is to represent it as an expected value of some random variable. Indeed, let us rewrite the integral as

$$I = \int_a^b \frac{g(x)}{f_x(x)} f_x(x)dx \tag{2.11}$$

Assuming that $f_x(x)$ is any p.d.f. such that $f_x(x) > 0$ when $g(x) \neq 0$.

Then

$$I = \mathrm{E}\left[\frac{g(X)}{f_x(X)}\right] \tag{2.12}$$

Where the random variable X is distributed according to $f_x(x)$.

Let us assume for simplicity

$$f_x(x) = \begin{cases} \dfrac{1}{(b-a)}, & a < x < b \\ 0, & otherwise \end{cases} \tag{2.13}$$

Then

$$\mathrm{E}[g(X)] = \frac{I}{b-a} \tag{2.14}$$

and

$$I = (b-a)\mathrm{E}[g(X)] \tag{2.15}$$

An unbiased estimator of I is its sample mean

$$\theta_2 = (b-a)\frac{1}{N}\sum_{i=1}^{N} g(X_i) \tag{2.16}$$

Sample-Mean Monte Carlo Method Algorithm

Step1: Input a, b and c

Step2: Generate a sequence $\{U_i\}_{i=1}^{N}$ of N random numbers.

Step3: Set $X_i = a + U_i(b-a)$, i=1, 2, …, N

Step4: compute $g(X_i)$, i=1, 2, …, N.

Step5: compute the sample mean which estimates the integral I by

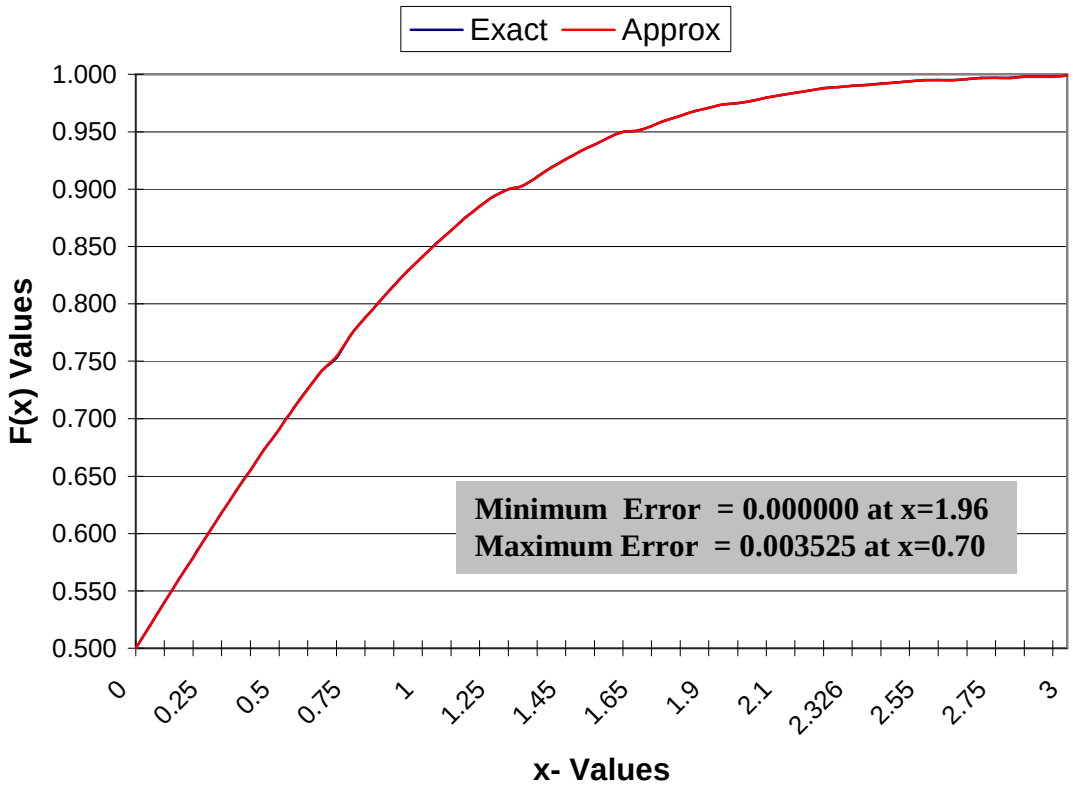
$$\theta_2 = (b-a)\frac{1}{N}\sum_{i=1}^{N} g(X_i)$$

Appendix (A5) involves a computer program written in Pascal language using Sample Mean Monte Carlo Method for approximating the integral side of eq.(2.1). The x-values of the upper limit of the integral is taken from the normal distribution table (1.1). Table(2.6) gives together the approximate and the errors value of F(x). while figure(2.7) shows the difference between the exact and approximate results of the N(0,1) c.d.f. by Sample Mean Monte Carlo Method. Also this figure shows the minimum and maximum errors between the exact and approximate result.

Table(2.6): Approximation by the Sample-Mean Monte Carlo Method

Pr (X ≤ x)					
x	**Approx.**	**Error**	**x**	**Approx.**	**Error**
0.00	0.500000	0.000000	1.60	0.944900	0.000101
0.05	0.519954	0.000046	1.645	0.949993	0.000008
0.10	0.539871	0.000129	1.65	0.950765	0.000236
0.15	0.559714	0.000287	1.70	0.954783	0.000218
0.20	0.578805	0.000195	1.75	0.959971	0.000030

0.25	0.598780	0.000221	1.80	0.963966	0.000035
0.30	0.617933	0.000067	1.85	0.967922	0.000079
0.35	0.636873	0.000127	1.90	0.970873	0.000127
0.40	0.654684	0.000317	1.95	0.973815	0.000185
0.45	0.673734	0.000266	1.960	0.975000	0.000000
0.50	0.690677	0.000323	2.00	0.976888	0.000112
0.55	0.708888	0.000112	2.05	0.979918	0.000082
0.60	0.725823	0.000177	2.10	0.981940	0.000060
0.65	0.741892	0.000108	2.15	0.983901	0.000099
0.70	0.754475	0.003525	2.20	0.985957	0.000043
0.75	0.772739	0.000261	2.25	0.987898	0.000102
0.80	0.787899	0.000102	2.30	0.988877	0.000123
0.85	0.801763	0.000237	2.326	0.989996	0.000004
0.90	0.815958	0.000042	2.35	0.990864	0.000136
0.95	0.828961	0.000039	2.40	0.991930	0.000070
1.00	0.840793	0.000207	2.45	0.992949	0.000051
1.05	0.852915	0.000085	2.50	0.993925	0.000075
1.10	0.863800	0.000200	2.55	0.994864	0.000136
1.15	0.874957	0.000043	2.576	0.995000	0.000000
1.20	0.884959	0.000041	2.60	0.994883	0.000117
1.25	0.893789	0.000211	2.65	0.995990	0.000010
1.282	0.899953	0.000047	2.70	0.996835	0.000165
1.30	0.902880	0.000120	2.75	0.996994	0.000006
1.35	0.910705	0.000295	2.80	0.996846	0.000154
1.40	0.918854	0.000146	2.85	0.997934	0.000067
1.45	0.925765	0.000236	2.90	0.997954	0.000046
1.50	0.932904	0.000097	2.95	0.997857	0.000143
1.55	0.938785	0.000215	3.00	0.998876	0.000124



Figure(2.7): Shows the difference between the exact and approximate solutions to the c.d.f. $\Phi(x)$ by Sample Mean Monte Carlo Method

2.7 Error of Approximation

In order to compare the five methods, table(2.8) shows the error of approximation of each method at a specific x-value. When x-values selected according to minimum and maximum errors resulted from each method.

Table(2.8): Error of approximation comparison for five methods

x-Value	Trapezoidal Rule	Simpson Rule	Gaussian-Quadrature Rule	Hit or Miss Rule	Sample Mean Rule
0.70	0.004947	0.005063	0.005063	0.004281	0.003525
1.80	0.000314	0.000069	0.000008	0.000052	0.000035
1.96	0.000364	0.000001	0.000160	0.000001	0.000000
2.30	0.000011	0.000274	0.000263	0.000192	0.000123

Conclusions and Future Work

Conclusion

1. Simpson rule is more accurate than trapezoidal rule because the function in Simpson rule is nearly quadratic on the close interval [a, b].
2. Gaussian Quadrature is more efficient than the trapezoidal and Simpson rule because if formula of degree n then the error will be of order (2n-1).
3. Hit or Miss Monte Carlo method is more efficient than Gaussian formula and the accuracy of this method increase as the sample size increases.
4. The Sample mean Monte Carlo method gives results superior than all methods of approximation for the c.d.f. of normal distribution which produce minimum errors and the accuracy increase as the sample size increases.
5. The best of our procedures for generating sample varieties from normal distribution is Box and Muller procedure which has less time consuming in comparison with the other method of generation.

Future Work

1. The methods of approximation to the c.d.f. for normal distribution can be used for other non-normal distributions.
2. It can be generate r.v.'^{s} from normal distribution by other new procedures which can be compare their efficiency with our used procedures.

Contents

Examining Committee's Certification

We certify that we read this thesis entitled ***"GENERATING NORMAL VARIETIES BY MONTE CARLO METHODS FOR ESTIMATING THE CUMULATIVE DISTRIBUTION FUNCTION AND PARAMETERS"*** and as examining committee examined the student, ***Israa Abdul Ameer Resen*** in its contents and in what it connected with, and that is in our opinion it meet the standard of thesis for the degree of Master of Science in Mathematics.

(Chairman)

Signature:

Name: Dr. Nadir G. Mansour

Date: / /2007

(Member)

Signature:

Name: Dr. Zainab A. Salman

Date: / /2007

(Member)

Signature:

Name: Dr. Radhi A. Zaboon

Date: / /2007

(Member and Supervisor)

Signature:

Name: Dr. Akram M. Al-Abood

Date: / /2007

Approved for the Dean of the College of Education

Signature:

Name: Assist. Prof. ***Dr. LAITH ABDUL AZIZ AL-ANI***

Dean of College of Science of Al-Nahrain University

Date: / / 2007

Introduction

The most important continuous probability distribution in the entire field of statistics is the "Normal Distribution", whose graph is symmetric bell shaped curve, extending indefinitely in both directions which describes so many sets of data that occur in nature, industry, research, and provided basis upon much of the theory of statistical inference has been developed.

The normal distribution was discovered by De Moiver in (1733) who derived the mathematical equation of the normal curve which is considered the best tool of statistical inference theory.

The normal distribution is often referred to as Gaussian distribution in honor of the German scientist Gauss (1773-1855), who also derived its equation from a study of errors in repeated measurements of the same quantity.

Laplace (1749-1827) studies in Astronomy gave results as Gauss obtained. Although the normal distribution may often give a very reasonable fit to an empirical distribution of some varieties, for example, height of people, component dimension, score in an aptitude test, etc, its primary importance occurs in connection with sampling theory. Thus when samples are drawn from populations, the distribution of measures commonly computed from samples, such as sample mean or proportions, often approach to normal distribution quite rapidly as the sample size increases (central limit theorem).

Sheppard (1903) [28] published accurate, famous and elaborate tables related to the basic integral of the standard normal c.d.f. $\Phi(x)$.

Adams A.G. (1969) [1] used the inverse Gussian distribution algorithm to evaluate the standard normal distribution.

Pettis (1974) [23] and Strecok (1968) [29] used a stable algorithm for computing the inverse error function in the "Tail-End" reign for standard normal integral.

B.D. Bunday, etal (1997)[4] developed new algorithms for the rapid, efficient and accurate evaluation of the standard normal integral and its tail.

Dmitry Danilov (2005) [6] generalized the problem of estimating the first K coefficients in a regression equation with K+1 variables to the case where the unknown variance is estimated by least squares and found that main properties of the Laplace estimator only change marginally. Therefore, he recommends the neutral Laplace estimator to be used in practice.

Jade Freeman and Reza Modarres (2006) [11] studied the moments of the power normal family and obtained expressions for its mean and variance. The quantile functions and a quantile measure of skewness are discussed to show that the power normal family is ordered with respect to the transformation parameter.

C.S. Withers and P.N. McGavin (2006) [5] gave a new expression for Mills' ratio and five expressions for repeated integrals of the univariate normal density, or equivalently for the Hermite functions and they also gave the derivatives of Mills' ratio and its inverse.

Yeh lam, etal (2006) [31] studied a sequential variable sampling plan. Supposed that the quality of an item in a batch in measured by a normally distributed random variable with a known variance, but the mean is unknown with a normal prior distribution. Then by using Bayesian approach and considering a Markoov decision process, the optimality equations for the minimum total expected cost are formulated. They showed that an optimal decision rule will have a control limit structure and they presented the statistical procedure for conducting the sequential sampling plan.

In this thesis mainly three chapters involved. The normal distribution is discussed in chapter one throughout six sections which include basic properties and moments of normal distribution, some related theorems, two methods of parameters estimation and equality of estimators. While, chapter two gives a full discussion for the approximation to the cumulative distribution function of the standard normal distribution by three numerical procedures namely: Trapezoidal, Simpson and Gaussian-Quadrature rules and two reduction techniques namely; Hit or Miss Monte Carlo and Sample Mean Monte Carlo rules. The approximation results of the five methods are tabulated and compared with the normal distribution table in order to obtain the more efficient and accurate procedure.

Chapter three presents three sections which include random number generation and four procedures for random varieties generation from normal distribution. The efficiency of these procedures is discussed theoretically and assessed destitution practically.

Notation and Abbreviation

Distn.	: Distribution
r.v.	: Random variable
p.d.f.	: Probability density function
c.d.f.	: Cumulative Distribution Function
m.g.f.	: Moment Generating Function.
Iff	: If and only if
r.s.	: Random sample
indep.	: Independent
M.L.E	: Maximum Likelihood Estimator
IT	: Inverse transform
$X \sim N(\mu, \sigma^2)$	: The r.v. X has normal distribution with mean μ and variance σ^2
$X \sim \chi^2(n)$	: The r.v. X has chi – Square distribution with parameter n
c	: Efficiency
Prob.	: probability
C.I	: confidence interval
dof	: degree of freedom
S.S	: sample space
M.V.U.E	: Minimum variance unbiased estimator

References

[1] Adams A.G., Algorithm 39, Computer Journal, 12, pp.197-198, (1969).

[2] Alexander M. Mood, Franklin A. Graybill and Duane C. Boes, Introduction to the Theory of Statistics, McGraw-Hill International Book Company, (1974).

[3] Averill M. Law and W. David Kelton, Simulation Modeling and Analysis, 3rd ed., McGraw-Hill series in industrial engineering and management science, (2000).

[4] B.D. Bunday, S. M. Husnain Bokhari and K. H. khan, A new algorithm for the normal distribution function, Sociedad de Estadística e Investigación Operativa, 6, pp.369-377, (1997).

[5] C.S. Withers and P.N. McGavin, Expressions for the normal distribution and repeated normal integrals, Statistics and Probability Letters, 76, pp.479-487, (2006).

[6] Dmitry Danilov, Estimation of the mean of a univariate normal distribution when the variance is not Known, Econometrics Journal, 8, pp.277-291, (2005).

[7] http://www.cybele.bu.edu/courses/gg645 Spring02/special topics

[8] http://www.groups.dcs.st-and.ac.uk/history/Mathematicians/ Kolmogrov.html

[9] http://www.groups.dcs.st-and.ac.uk/history/Mathematicians/Gosset.html

[10] Hull, T. E. and A. R. Dobell, Mixed Congruential Random Number Generators for Binary Machines, J. Assoc. Comp. Mach., 11, PP. 31-40, (1964).

[11] Jade Freeman and Reza Modarres, Inverse Box-Cox: The power-

normal distribution, Statistics and Probability Letters, 76, pp.764-772, (2006).

[12] James L. B. and Peter R. T., Numerical Methods and Analysis, McGraw-Hill Inc.,(1992).

[13] Jun Shao, Mathematical Statistics, Spring-Verlag New York, Inc., (1999).

[14] K. Krishnamoorthy, Handbook of Statistical Distribution with Applications, Taylor and Franics Group, LLC, (2006).

[15] Kendal Atkinson, Elementary Numerical Analysis, by John Wiley and Sons, Inc., (1985).

[16] Kunth, D. E., The Art of Computer Programming: Semi numerical Algorithm, 2, Addison-Wesley, Reading, Massa chuelts, (1969).

[17] L. V. Atkinson and P. J. Harley, An Introduction to Numerical Methods with Pascal, Addoson-Wesley, Publishers Limited, (1983).

[18] Law, A. M. and Keliton, W. D., Simulation Modeling and Analysis, 3rd ed., McGraw. Company, (2000).

[19] Lehmer, D. H., Mathematical Methods in Large-Scale Computing Units, Ann. Comp. Lab. Harvard Univ., 26, PP. 141-146, (1951).

[20] Maclaren, M. D. and G. Marsuglia, Uniform Random Number Generators, J. Assoc., Comp., Mach., 12, PP. 83-89, (1965).

[21] Norman L. Johnson, Adrienne W. Kemp and Samuel Kotz, Univariate Discrete Distributions, John Wiley and Son, Inc. All rights reserved, (2005).

[22] Norman S. M., Probability Modeling and Computer Simulation, Boston Publishing Company, Massachusetts, University of California., Inc. New York, (1988).

[23] Pettis H. E., A stable algorithm for computing the inverse error function in the "Tail-End" region, Mathematics of Computation, 28,

pp.585-587, (1974).

[24] Rand Corporation, A Million Random Digits with 1000,000 Normal Deviates, Free Press, Clence, Illinois, (1955).

[25] Richard L. Burden and J. Douglas Faires, Numerical Analysis, 3rd ed., Cover art "Reaction" by Masaaki Noda; asilk Screen print used with the permission of the artist, (1984).

[26] Robert V. Hogg and Elliot A. Tanis, Probability and Statistical Inference, Macmillan Publishing Co., (1983).

[27] Rubinstein R. Y., Simulation and Monte Carlo Methods, Wieley series in Probability and Mathematical Statistics, John Wieley and Sons, (1981).

[28] Sheppard, New Tables of the Probability Integral, Biomertrika, 2, pp.174-190, (1903).

[29] Strecok A. V., On the calculation of the inverse of the error function, Mathematics of Computation, 22, pp.144-158, (1968).

[30] Von Neuman J., Various Technique Used in Connection with Random Digits, U.S Nat. Bur. Stand. Appl. Math. Ser., 12, PP. 36-38, (1951).

[31] Yeh lam, Kim-Hung Li, Wai-Cheung Ip and Heung Wong, Sequential variable sampling plan for normal distribution, European Journal of operational Research, 172, pp.127-145, (2006).

Supervisor Certification

I certify that this thesis was prepared under my supervision at the department of mathematics and computer applications, ***College of Science, Al-Nahrain University*** as a partial fulfillment of the requirements for the degree of Master of Science in Mathematics.

Signature:

Name: Assist. Prof. Dr. Akram M. Al-Abood

Date: /6/2007

In view of the available recommendations; I forward this thesis for debate by examination committee.

Signature:

Name: Assist. Prof. Dr. Akram M. Al-Abood

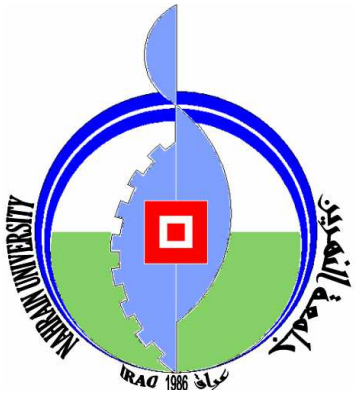
Head of the Mathematics and Computer Applications Department

Date: /6/2007

(وقُلْ رَبِّ أَدْخِلْنِي مُدْخَلَ صِدْقٍ
وَأَخْرِجْنِي مُخرَجَ صِدْقٍ وأجْعَلْ لي مِنْ
لدُنْكَ سُلْطاناً نَصيراً)

(سورة الاسراء/ ٧٩-٨٠)

Ministry of Higher Education
and Scientific Research
Al-Nahrain University
College of Science



Generating Normal Varieties by Monte Carlo Methods for Estimating the Cumulative Distribution Function and Parameters

A Thesis

Submitted to the Department of Mathematics, College of Science, Al-Nahrain University, as a Partial Fulfillment of the Requirements for the Degree of Master of Science in Mathematics

By

Israa Abdul Ameer Resen

(B.Sc. 2004)

Supervised by

Assist. Prof. Dr. Akram M. Al-Abood

June *Jamad Al-Awaal*

2007 *1428*

وزارة التعليم العالي والبحث العلمي
جامعـة النـهرين
كلية العلوم

توليد المتغيرات الطبيعية باستخدام طرق مونت كارلو لتخمين دالة التوزيع التجميعية ومعلماتها

رسالة
مقدمه إلى كلية العلوم في جامعة النهرين
وهي جزء من متطلبات نيل درجة ماجستير
علوم في الرياضيات

من قبل
إسراء عبد الأمير رسن
(بكالوريوس علوم ٢٠٠٤)

بإشراف
أ. د. أكرم محمد العبود

حزيران
٢٠٠٧

جمادي الأول
١٤٢٨

المستخلص

في هذه الأطروحة تطرقنا إلى التوزيع الطبيعي "Normal distribution" لأهميته في مجالات الإحصاء وتطبيقاته من حيث استعراض وتوحيد لخواص التوزيع الرياضية والإحصائية والعزوم والعزوم العليا. وتم دراسة ست نظريات مهمة في تطبيقات هذا التوزيع.

ثم التطرق إلى اسلوب التخمين وخواصه ومناقشة طريقتين لتخمين معلمات التوزيع نظريا وهما طريقة العزوم "Moment method" وطريقة الترجيح الأعظم "Maximum likelihood method" حيث نوقشت جودة المخمنات وخواصها من خلال مبرهنات التخمين المعروفة.

تم استخدام خمس طرائق لتقريب داله التوزيع التجميعية وهي طريقه "Trapezoidal, Simpson, Gaussian, Hit or Miss and Sample mean Rules" وتم مقارنه نتائج هذه الطرق من حيث التصرف والخطأ المتولد من كل طريقه والتوصل إلى أن آخر طريقه تم استخدامها هي الأفضل لتقريب الحل لمثل هذه الدالة. إضافة إلى ذلك تم تمثيل نتائج كل طريقة بمنحنيات وجداول عددية لسهولة مقارنة النتائج.

وأخيرا تطرقنا إلى أربع أساليب لتوليد متغيرات عشوائية وهي "Box-Muller, Acceptance-Rejection, Central Limit theorem and Tocher" حيث قورنت كفاءة هذه الطرق نظرياً وعملياً باستخدام محاكاة مونت كارلو وأظهرت النتائج بان "Box-Muller" هي أكفأ طريقه لتوليد هذه المتغيرات من حيث الوقت المستغرق في توليد العينة.