SOLVING SOME THEORETICAL DISTRIBUTION PROBLEMS BY USING C++ PROGRAM

J.Sathyamalar[#1], M.Vadivukkarasi[*2]
[#1] *Assistant Professor, K.S.G College of Arts & science,*
Coimbatore, Tamil Nadu, India.
[1] sathyajmalar@gmail.com
[*2] *Assistant Professor, K.S.G College of Arts & science,*
Coimbatore, Tamil Nadu, India.
[2] vadivukkarasi153@gmail.com

Abstract - **In this paper we study the notions of Binomial, Poisson and Normal Distribution in statistical methods and also we introduce the C++ program in Binomial, Poisson and Normal Distribution problems.**

I. INTRODUCTION

In Statistics the binomial distribution also known as "Bernoulli Distribution" is associated with the name of a Swiss mathematician James Bernoulli also known as Jacques (1654-1705).Binomial distribution is a probability distribution expressing the probability of one set of dichotomous alternatives. i.e., success or failure. Thus distribution has been used to describe a wide variety of processes in business and the social science as well as other areas.

A C++ program is a collection of commands. This collection of commands is usually called C++ source code, source code or just code**.** Commands are either "functions" or "keywords". Keywords are a basic building block of the language, while functions are, in fact. Every program in C++ has one function, always named main, that is always called when your program first executes. [1]

The #include is a "pre-processor" directive that tells the compiler to put code from the header called iostream into our program before actually creating the executable. By including header files, you gain access to many different functions. By including this line at the top of a file. The "curly braces" ({and}) signal the beginning and end of functions and other code blocks [1]. In C++, however, the cout object is used to display text ("pronounced out"). It uses the <<symbols, known as "insertion operators", to indicate what to output. Cout<<results in a function call with the ensuing text as an argument to the function. [1]

A. Binomial Distribution

The Binomial Distribution describes the distribution of binary data from a finite sample. Thus it gives the probability of getting r events out of n trials. [2]

1) General Form Of The Binomial Distribution [4]:

$P(r) = {}^nC_r\ q^{n-r}\ p^r$
P = Probability of success in a single trial
q=1-p
n=Number of trails
r=Number of success in a trials

2) Binomial Expansion [4]:

$(q + p)^n = q^n + {}^nC_1\ q^{n-1}\ p + {}^nC_2\ q^{n-2}\ p^2 + \ldots + {}^nC_r\ q^{n-r}\ p^r + \ldots p^n.$

3) Constants of Binomial distribution:

Mean = np
Standard deviation = $\sqrt{npq}$
First moment or $\mu_1 = 0$
Second moment or $\mu_2 = npq$
Third moment or $\mu_3 = npq(q-p)$
Fourth moment or
$\mu_4 = 3n^2\, p^2 q^2 + npq(1-6pq)$
$\beta_1 = \frac{(q-p)^2}{npq} = \frac{\mu_3^2}{\mu_2^2}$
$\beta_2 = 3 + \frac{1-6pq}{npq} = \frac{\mu_4}{\mu_2^2}$

4) Example

If the probability of defective bolts is 0.1 find
(a) The mean and standard deviation for the distribution of defective bolts in a total of 500, and
(b) The moment coefficient of skewness and kurtosis of the distribution.

Solution:

(a) p = 0.1 n = 500
Mean = np = 500X.1 = 50.
Thus we can expect 50 bolts to be defective.
n = 500,p = 0.1 and q = 0.9
$\sigma = \sqrt{npq}$
$\sigma = \sqrt{500X0.1X0.9}$ = 6.71

(b) Moment coefficient of skewness, i.e.,

$$\boldsymbol{\gamma}_1 = \sqrt{\boldsymbol{\beta}} = \sqrt{\frac{(q-p)}{npq}} = \frac{q-p}{\sqrt{npq}} = \frac{(0.9-0.1)}{6.71} = \frac{0.8}{6.71} = \mathbf{0.}$$

Since γ_1 is more than zero the distribution is positively skewed. However, skewness is very moderate.

Moment coefficient of kurtosis γ_1-β_2-3

$\beta_2 = 3+\frac{1-6pq}{npq} = 3+\frac{1-6(0.1)(0.9)}{44.9}$

$= 3+\frac{0.46}{44.9} = 3.0$

γ_2=3.01-3=+0.01

Since γ_2 is positive, the distribution is platykuritic,

B. POISSON DISTRIBUTION

The Poisson distribution describes the distribution of binary data from a infinite sample. Thus it gives the probability of getting r events in a population [2].

1) General form of Poisson distribution [4]:

P(r) = $\frac{e^{-m}m^r}{r!}$ Where r = 0, 1, 2, 3,4 ,…

e = 2.7183(the base of natural logarithms)

m=the mean of the Poisson distribution, i.e., n p or the average number f occurrences of an event.

2) Constants of Poisson distribution:

$\mu = 0. \mu_2 = m, \mu_3 = m + 3m^2$

$\beta_1 = \frac{\mu_3^2}{\mu_2^3} = \frac{m^2}{m^3} = \frac{1}{m}$

$\beta_2 = \frac{\mu_4}{\mu_2^2} = \frac{m+3m^2}{m^2} = 3+\frac{1}{m}$

Mean = λ

Variance = λ (x = λ)

3) Poisson distribution as an approximation of the Binomial Distribution:

The Poisson distribution can be a reasonable approximation of the binomial under certain conditions like:

* Number of trails, i.e., n is indefinitely large. i.e., n→ ∞.

* p i.e., the probability of success for each trail is indefinitely small i.e., p→ 0.

* np = m(say) is finite.

The rule most often used by statisticians is that the Poisson a good approximation of the binomial when n is equal to or greater than 20 and p is equal to or less than .05. If above conditions hold good, we can substitute the mean of the binomial distribution (np) in place of the mean of the Poisson distribution (m), so that the formula becomes:

$$P(r) = \frac{e^{-np}(np)^r}{r!}$$

Proof:

In case of binomial distribution the probability of r success is given by

$p^{(r)} = {}^nC_r\, q^{n-r}\, p^r = \frac{n\,(n-1)\ldots\ldots(n-r+1)}{r!}p^r q^{n-r}$

put p=$\frac{m}{n}$ $\therefore q = 1-p = 1-\frac{m}{n}$.

We now get P(r)= $\frac{n(n-1\ldots\ldots n-r+1)}{r!}\left(\frac{m}{n}\right)\times\left(1-\frac{m}{n}\right)^{n-r}$

$= \frac{1\left(1-\frac{1}{n}\right)\left(1-\frac{2}{n}\right)\ldots\ldots\left(1-\frac{r-1}{n}\right)m^r}{r!}\left\{\frac{\left(1-\frac{m}{n}\right)}{\left(1-\frac{m}{n}\right)}\right\}$

For fixed r, as n→ ∞

$\left(1-\frac{1}{n}\right)\ldots\ldots\left(1-\frac{r-1}{n}\right)\left(1-\frac{m}{n}\right)^r$ all tend 1 and $\left(1-\frac{m}{n}\right)^n$ to $e^{-m}m^r$

Hence in the limiting case $p^{(r)} = \frac{e^{-m}m^r}{r!}$ Needless to point out that use of Poisson as an approximation to binomial probability distribution if the conditions given above are satisfied can simplify calculation work and save time with more or less similar result as one would expect from binomial distribution.

4) Example:

Suppose on an average 1 house in 1,000 in a certain district has a file during a year. If there are 2,000 houses in that district, what is the probability that exactly 5 house will have during the year.

Solution:

Applying the Poisson distribution

$\bar{X}$ = np n = 2000. p=$\frac{1}{1000}$

n = 2000× $\frac{1}{1000}$=2

P(r) = $e^{-m}\frac{m^r}{r!}$,r = 0,1,2…..

m = 2,r = 5 and e = 2.718

P(5) = $\frac{2.7183^{-2}}{5!}\times 2^5$

$= \frac{\boldsymbol{Rec.}[\boldsymbol{antilog}(2\times\boldsymbol{log}2.7183)]\times 32}{5\times4\times3\times2\times1}$

$= \frac{\boldsymbol{Rec.}[\boldsymbol{antilog}(2\times.4343)]\times 32}{120}$

$= \frac{\boldsymbol{Rec.}[\boldsymbol{antilog}(.8686)]\times 32}{120}$

$= \frac{\boldsymbol{Rec.}[7.389]\times 32}{120} = \frac{1352\times 32}{120} = \mathbf{.036}$

C. NORMAL DISTRIBUTION

The Normal distribution describes continues data which have a symmetric distribution, with a characteristic ‘bell’ shape.

1) General form of Normal distribution

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

X= values of the continuous random variable

μ = Mean of the normal random variable

e=Mathematical constant approximated by 2.7183, π=Mathematical constant approximated by 3.1416. ($\sqrt{2\pi}$=2.5066)

2) Graph of Normal Distribution [5]

(a) The normal distribution can have different shapes depending on different values of but there is one normal distribution for any given pair of values for

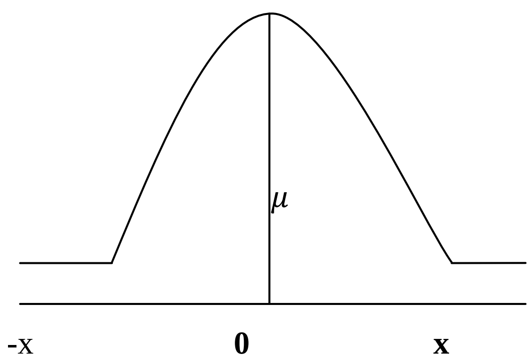
(b) Normal distribution is a limiting case of binomial distribution when

(c) n and

(d) Neither p nor q is very small.

(e) The mean of a normally distributed population lies at the centre of its normal curve.

(f) The two tails of the normal probability distribution extend infinitely and never touch the horizontal axis (which implies a positive probability for finding values of the random variable within any rang from minus infinity to plus infinite.



3) Example

A grinding machine is so that its production of shafts has an average diameter of 10.10 cm and a standard of 0.20 cm. the product specifications call for shaft diameters between 10.05 cm, and 10.20 cm. what proportion of output meets the specifications presuming normal distribution?

Solution:

Assuming that the distribution of diameter of shafts is normal.

We have $\mu = 10.10\ cm, \sigma = 0.20\ cm.$

For x = 10.05, $z = \frac{10.05-10.10}{0.20} = -0.25$

For x = 10.20, $z = \frac{10,20-10.10}{0.20} = 0.5$

Area of the normal curve between the ordinate -0.25 to 0.5 area between the ordinates -0.25 to 0+ area between the ordinates 0 to 0.50. 0987+0.1915=0.2902.thus, 29.02% of the output meets the specification.

II. SOLVING THE STATISTICAL PROBLEMS BY USING C++ PROGRAM

A. Binomial Distribution

1) Example

The probability of any ship of a company being destroyed on a certain voyage is 0.02. The company owns 6 ships for the voyage. What Is the probability of losing one ship?

Program

Coding

```
#include<iostream.h>
#include<conio.h>
#include<math.h>
long fact(int y);
void main()
{
int i,n,x;
float p,q;
long n1,n2;
double fx;
clrscr();
cout<<"Enter the n value";
cin>>n;
cout<<"Enter the x value";
cin>>x;
cout<<"Enter the p value";
cin>>p;
q=1-p;
float ncr=fact(n)/(fact(x)*fact(n-x));
cout<<"NCR value is = "<<ncr<<endl;;
fx=(ncr)*pow(p,x)*pow(q,(n-x));
cout<<"The Final value is "<<fx;
getch();
}
long fact(int y)
{
int i;
long f=1;
for(i=1;i<=y;i++)
{
f=f*i;
}
return(f);
}
```

Output

```
Enter the n value  : 6
Enter the x value  : 2
Enter the p value  : 0.02
Enter the q value  : 0.98
n factorial is 720
x factorial is 2
(n-x) factorial is 24
p^ x  is 0.0004
q ^(n-x) is 0.922368
THE BINOMIAL RESULT IS 0.005534.
```

2) Example

A coin is tossed six times. What is probability of obtaining 4 heads?

Program

Coding

```
#include<iostream.h>
#include<conio.h>
#include<math.h>
long fact(int y);
```

```
void main()
{
int i,n,x;
float p,q;
long n1,n2;
double fx;
clrscr();
cout<<"Enter the n value";
cin>>n;
cout<<"Enter the x value";
cin>>x;
cout<<"Enter the p value";
cin>>p;
q=1-p;
float ncr=fact(n)/(fact(x)*fact(n-x));
cout<<"NCR value is = "<<ncr<<endl;;
fx=(ncr)*pow(p,x)*pow(q,(n-x));
cout<<"The Final value is "<<fx;
getch();
}
long fact(int y)
{
int i;
long f=1;
for(i=1;i<=y;i++)
{
f=f*i;
}
return(f);
}
```

Output

Enter the n value: 6
Enter the x value: 4
Enter the p value: 0.5
Enter the q value: 0.5
n factorial is 720
x factorial is 24
(n-x) factorial is 2
p^ x is 0.0625
q ^(n-x) is 0.25
THE BINOMIAL RESULT IS 0.234375.

B Poisson distribution

1) Example

It is known from past experience that in a certain plant there are on the average 4 industrial accidents per month. Find the probability that in a given year there will be 2 accidents.

Program

Coding

```
#include<iostream.h>
#include<conio.h>
#include<math.h>
int fact(int m);
void main()
{
 int n,ft;
 float pd,lm;
 clrscr();
 cout<<"\n Enter the n value: ";
 cin>>n;
 cout<<"\n Enter the Lamda value: ";
 cin>>lm;
 pd=(exp(-lm)*pow(lm,n))/fact(n);
 cout<<"\nAns:"<<pd;
 getch();
}
int fact(int n)
{

  if(n==1)
  return(n);
  else
  n=n*fact(n-1);
  return(n);
}
```

Output

Enter the n value : 2
Enter the lamda value :4
Ans : 0.14656.

2) Example

A book contains 100 misprints distributed randomly throughout its 100 pages. What is the probability that a page observed at random contains one misprint.

Program

Coding

```
#include<iostream.h>
#include<conio.h>
#include<math.h>
int fact(int m);
void main()
{
 int n,ft;
 float pd,lm;
 clrscr();
 cout<<"\n Enter the n value: ";
 cin>>n;
 cout<<"\n Enter the Lamda value: ";
 cin>>lm;
 pd=(exp(-lm)*pow(lm,n))/fact(n);
 cout<<"\nAns:"<<pd;
 getch();
}
int fact(int n)
{
  if(n==1)
  return(n);
  else
  n=n*fact(n-1);
  return(n);
}
```

Output

Enter the n value: 1
Enter the lamda value: 1
Ans: 0.367879.

C. Normal Distribution

1) Example

The life time of a certain kind of battery has a mean of 300 hours and a standard deviation 35 hours. Assuming that the distribution of life times, which are measured to the nearest hour is normal find the percentage of batteries which have life time more than 370 hours.

Program
Coding

```
#include<iostream.h>
#include<conio.h>
#include<math.h>
void main()
{
clrscr ();
float z,x,m,sigma;
cout<<"\nEnter the X value: ";
cin>>x;
cout<<"\nEnter the Mu value: ";
cin>>m;
cout<<"\nEnter the Sigma Value: ";
cin>>sigma;
z=(x-m)/sigma;
cout<<"\nNormal distribution Z = "<<z;
getch();
}
```

Output

Enter the x value: 370
Enter the Mu value: 300
Enter the Sigma value: 35
Normal Distribution Z: 2

2) Example

If x=250, μ =200 and σ =35. Find the Normal variate.

Program
Coding

```
#include<iostream.h>
#include<conio.h>
#include<math.h>
void main()
{
clrscr();
float z,x,m,sigma;
cout<<"\nEnter the X value: ";
cin>>x;
cout<<"\nEnter the Mu value: ";
cin>>m;
cout<<"\nEnter the Sigma Value: ";
cin>>sigma;
z=(x-m)/sigma;
cout<<"\nNormal distribution Z = "<<z;
getch();
}
```

Output

Enter the x value: 250
Enter the Mu value: 200
Enter the Sigma value: 35
Normal Distribution Z: 1.4285714

III REFERENCES

[1] E. Balagurusamy, "Programming in C++", Tata McGraw Hill Education pvt.LTD, New Delhi, Fifth edition
[2] P. Bali, "Statistics", Laxmi Publications (p) ltd, New Edition, New Delhi.
[3] Dr. S. P. Gupta, "Statistical methods", Sultan Chand & Sons, New Delhi.
[4] S. C. Gupta, V.K. Kapoor," Fundamentals of Mathematical Statistics", Sultan Chand & Sons, Eleventh Thoroughly Revised edition, New Delhi.
[5] PA.Navnitham," Business Mathematics and Statistics", Jai Publishers-2011, Trichy.
[6] R. S. N. Pillai, V. Bagavathi," Statistics", S. Chand & Company LTD, New Delhi.