Aqila Habib (47162)

Project Title

(Poisson distribution Probability)

Problem statement:

A big hall houses 50 shops. The daily visit from each shop is estimated to be on average 120. The average per shop is 150 on Saturday. We choose a shop and a day at random. Compute the chance of having between 0 to 200 visitors.
How is the distribution if the day happens to be Saturday?

$$\Pr(0 <= x <= 200) = \sum \frac{\lambda^x e^{-\lambda}}{x!} = \frac{120^{200} * e^{-120}}{120!}$$

$$\Pr(0 <= x <= 200) = \sum \frac{\lambda^x e^{-\lambda}}{x!} = \frac{150^{200} * e^{-150}}{150!}$$

X= c(0 : 200)

$\lambda = 120$ (on average per shop on all days)

$\lambda = 150$ (on average per shop on Saturday)

Methods:

How do you solve the problem, and which tools do you use?

Since it is a Poisson Distribution Probability, we can solve it using the

- Poisson Distribution Formula
- Using ggplot2 in the R language program for a better analysis of the problem.

Since the values of Lambda and the visitors are high and unable to solve using a calculator, we can use Rstudio for a better way to solve of the problem.

Solution:

- 50 Shops in a big hall house
- Each shop's daily visit on average is 120.
- On Saturday 150 people visit each shop
- Random day, random shop

  Lambda = 120 and x = 200 <<<<<<<< $\Pr(0 <= x <= 200) = \sum \frac{\lambda^x e^{-\lambda}}{x!}$

Lambda= 150 and x = 200 <<<<<<<<< Pr (0<= x <= 200) =$\sum \frac{\lambda^x e^{\wedge -\lambda}}{x!}$

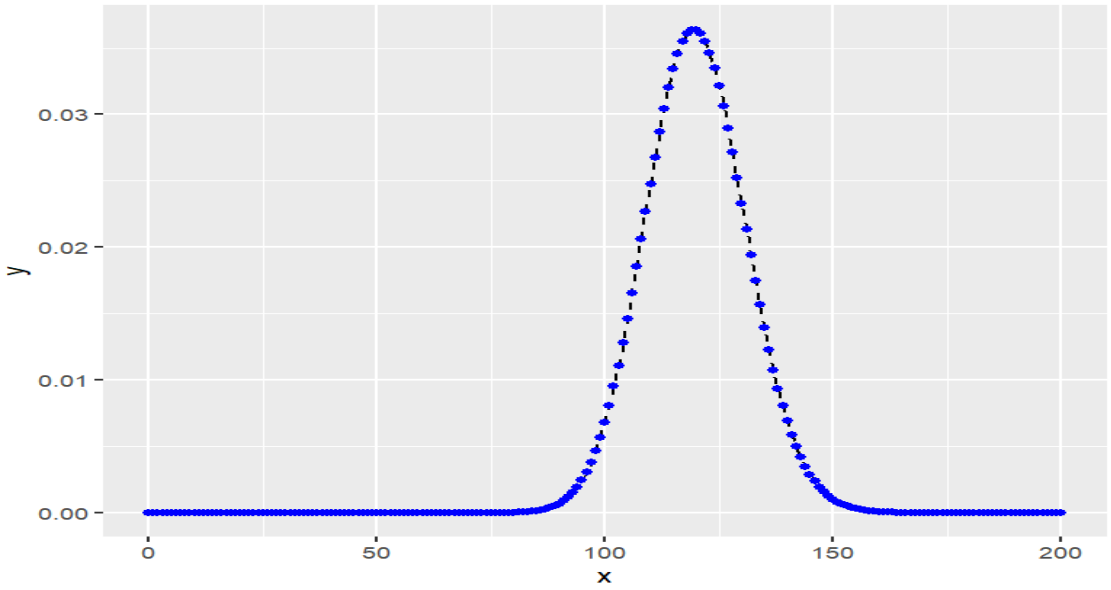
R Codes:

PMF analysis codes

Lambda = 120

Victor x= c(0: 200)

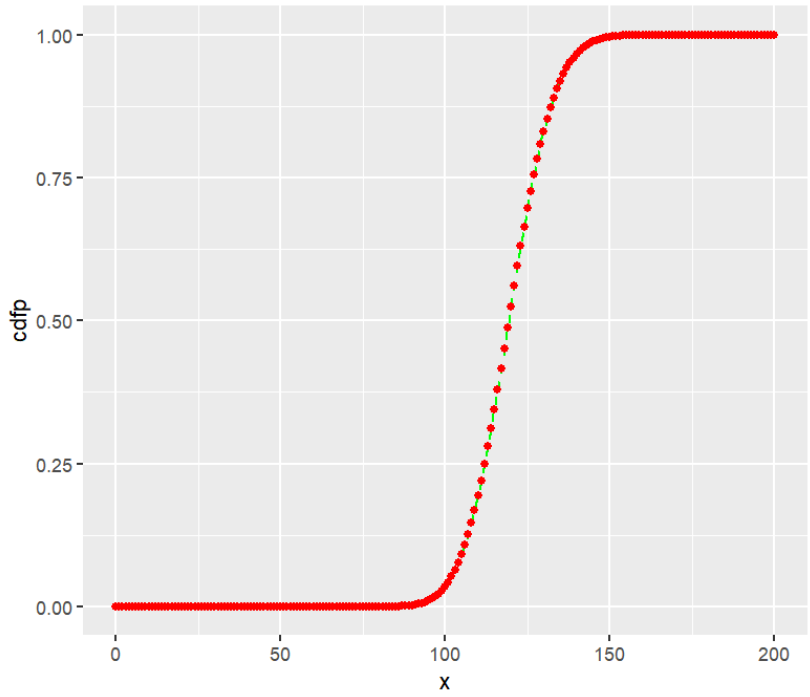
```
> library(ggplot2)
> help("dpois")
> dpois(200, 120)
[1] 6.668417e-12  <<< The probability is nearly 0.
> x= seq(0, 200, by=1)
> x
  [1]   0   1   2   3   4   5   6   7   8   9  10  11  12  13  14  15  16  17
18
 [20]  19  20  21  22  23  24  25  26  27  28  29  30  31  32  33  34  35  36
37
 [39]  38  39  40  41  42  43  44  45  46  47  48  49  50  51  52  53  54  55
56
 [58]  57  58  59  60  61  62  63  64  65  66  67  68  69  70  71  72  73  74
75
 [77]  76  77  78  79  80  81  82  83  84  85  86  87  88  89  90  91  92  93
94
 [96]  95  96  97  98  99 100 101 102 103 104 105 106 107 108 109 110 111 112
113
[115] 114 115 116 117 118 119 120 121 122 123 124 125 126 127 128 129 130 131
132
[134] 133 134 135 136 137 138 139 140 141 142 143 144 145 146 147 148 149 150
151
[153] 152 153 154 155 156 157 158 159 160 161 162 163 164 165 166 167 168 169
170
[172] 171 172 173 174 175 176 177 178 179 180 181 182 183 184 185 186 187 188
189
[191] 190 191 192 193 194 195 196 197 198 199 200
> y= dpois (x, 120)
> dt= as.data.frame(cbind(x, y))
> g1= ggplot(data= dt, aes(x=x, y=y))+
+    geom_line(linetype= 'dashed', color= 'black')+
+    geom_point(color= 'blue')
> g1
```



It is estimated that each point on the graph is the probability of visitors visiting daily per shop. The average probability is around 120 at a maximum point of nearly 0 over the y-axis and the x-axis.

Cumulative Distribution Probability (CDF)

```
> #Cumulative prob distribution (cdfp)
> #prob(x<= 200)
> ppois( 200, 120)
[1] 1  <<< The chance of getting 200 visitor per shop on daily is estimated n
early 1 as the cumulative distribution of the probability. Or in simple the e
stimation of 200 visitor per daily shop is 1 as the CDFP.
> cdfp= ppois(x, 120, lower.tail = TRUE)
> dt1= as.data.frame(cbind(x, cdfp))
> dt= as.data.frame(cbind(x, y))
> g1= ggplot(data = dt, aes(x=x, y= cdfp))+
+   geom_line(linetype= 'dashed', color = 'green')+
+   geom_point(color= 'red')
> g1
```

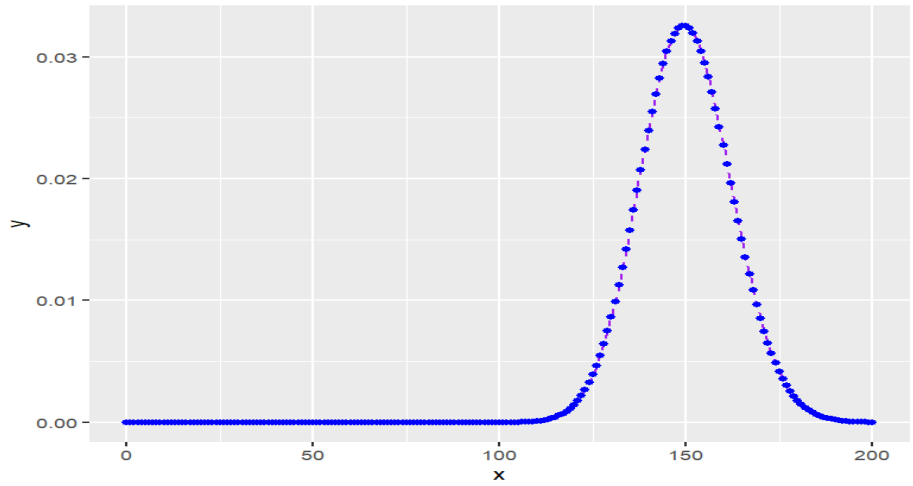


PMF analysis codes

Lambda = 150 (On Saturday)

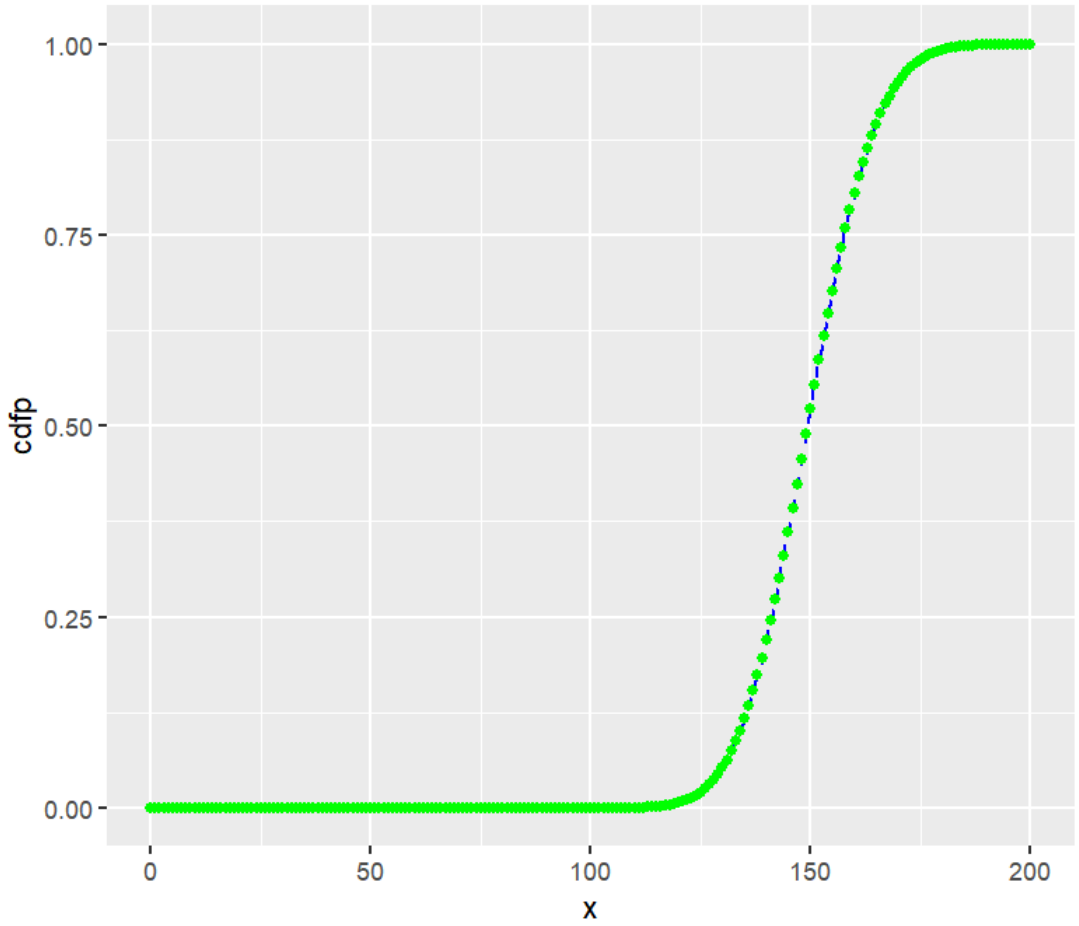
X= c(0 : 200)

```
> dpois(200, 150)
[1] 1.503803e-05 <<< The probability is nearly 0.
> y= dpois (x, 150)
> y
> dt1= as.data.frame(cbind(x, y))
> g1= ggplot(data= dt1, aes(x=x, y=y))+
+    geom_line(linetype= 'dashed', color= 'purple')+
+    geom_point(color= 'blue')
> g1
```



Cumulative Distribution Probability (CDF)

```
> #Cumulative prob distribution (cdfp)
> #prob(x<= 200)
> ppois( 200, 150)
[1] 0.9999579 <<<< the chance of getting 200 visitor per shop on Saturday is
estimated nearly 1 as the cumulative distribution of the probability.
> cdfp= ppois(x, 150, lower.tail = TRUE)
> dt2= as.data.frame(cbind(x, cdfp))
> dt= as.data.frame(cbind(x, y))
> g1= ggplot(data = dt, aes(x=x, y= cdfp))+
+   geom_line(linetype= 'dashed', color = 'blue')+
+   geom_point(color= 'green')
> g1
```


1.00
0.75
0.50
0.25
0.00
cdfp
0
50
100
150
200
x