Date: 22 April, 2022

Practical 16: To fit a normal distribution and find the theoretical frequencies for the classes of the given data.

Question: The following table gives baseball throws for a distance by 303 first year high school girls:

Distance in feet	**Number of girls**
15-25	1
25-35	2
35-45	7
45-55	25
55-65	33
65-75	53
75-85	64
85-95	44
95-105	31
105-115	27
115-125	11
125-135	4

135-145	1

a) Fit a normal distribution and find the theoretical frequencies for the classes of the above frequency distribution.

b) Find the expected number of girls throwing baseballs at a distance exceeding 105 feet on the basis that the data fit a normal distribution.

Methods and Formula:

- **Normal Distribution:** The Normal distribution, also known as the Gaussian distribution, is a probability distribution that is symmetric about the mean, showing that data near the mean are more frequent in occurrence than data far from the mean. In graph form, normal distribution will appear as a bell curve.

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

Here, $\sigma = Standard\ deviation$

$\mu = mean$

- **Histogram:** A histogram is a graphical representation that organizes a group of data points into user-specified ranges. Similar in appearance to a bar graph, the histogram condenses a data series into an easily interpreted visual by taking many data points and grouping them into logical ranges or bins.

Analysis:

To fit a normal distribution and find the theoretical frequencies for the classes of the given data. The tables 16.1 have been prepared in MS Excel.

Table 16.1: Tabular form of Data for calculating Normal Distribution of the given data

Distance in feet	Number of girls	M	FM	|M-mean|^2	F(M-mean)^2	Limits	P(X-x)	Z	P(Z<z)	P_class interval	Exp frequency	Exp. Frequency round off
[15-25)	1	20	20	3675.6407	3675.6407	15	0.000866	-3.1326	0.000866	0.003096	0.938003597	1
[25-35)	2	30	60	2563.0995	5126.199	25	0.003962	-2.6553	0.003962	0.010743	3.255032016	3
[35-45)	7	40	280	1650.5582	11553.908	35	0.014705	-2.178	0.014705	0.029803	9.030232005	9
[45-55)	25	50	1250	938.01697	23450.424	45	0.044507	-1.7006	0.044507	0.066105	20.02974333	20
[55-65)	33	60	1980	425.47572	14040.699	55	0.110612	-1.2233	0.110612	0.11724	35.52359475	36
[65-75)	53	70	3710	112.93446	5985.5265	65	0.227852	-0.7459	0.227852	0.166266	50.37873119	50
[75-85)	64	80	5120	0.3932076	25.165289	75	0.394118	-0.2686	0.394118	0.188555	57.13221522	57
[85-95)	44	90	3960	87.851954	3865.486	85	0.582673	0.20874	0.582673	0.170994	51.81121407	52
[95-105)	31	100	3100	375.3107	11634.632	95	0.753667	0.68608	0.753667	0.124002	37.57259439	38

[105-115)	27	110	2970	862.76945	23294.775	105	0.877669	1.16341	0.877669	0.071906	21.78761766	22
[115-125)	11	120	1320	1550.2282	17052.51	115	0.949576	1.64075	0.949576	0.033341	10.10221972	10
[125-135)	4	130	520	2437.6869	9750.7477	125	0.982916	2.11809	0.982916	0.01236	3.74508076	4
[135-145)	1	140	140	3525.1457	3525.1457	135	0.995276	2.59543	0.995276	0.003663	1.109951118	1
Total	**303**		**24430**	**18205.112**	**132980.86**	145	0.99894	3.07277	0.99894			**303**

Where,

Mean	80.62706271
Variance	438.8807198
S.D	20.94948018

From Table 16.1, this represents the calculation of Normal Distribution of the given data of 303 first year high school girls throws baseball at a certain distance (in feet).

X: Distance pf baseball throw (in feet)

$$X \sim N(80.63, 438.88)$$

$$f(x) = \frac{1}{20.95\sqrt{2\pi}} e^{-\frac{1}{2*20.95}(x-80.63)}$$

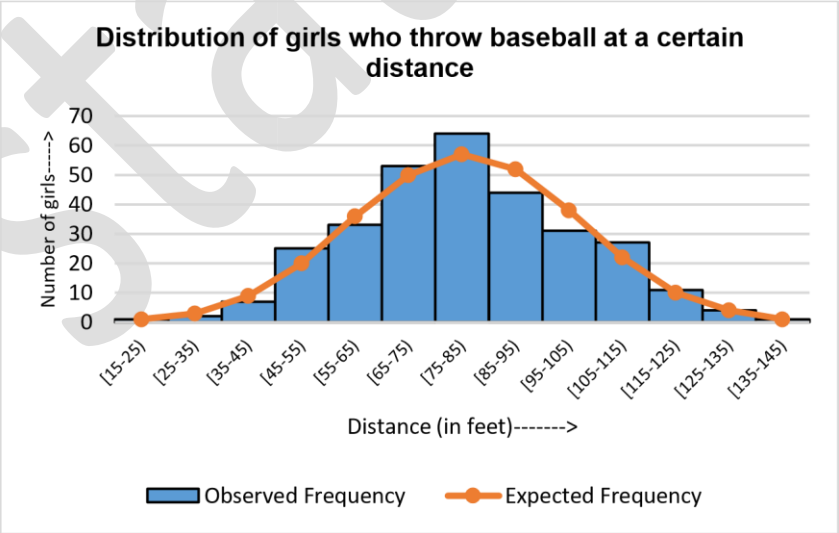
Expected frequency for say interval 15-25 = N* probability that X lies between 15 to 25

$$=N \times P[15 < X < 25]$$

$$=N \times [P[X < 25] - P[X < 15]]$$

$$=N \times [P\left[Z < \frac{25-80.63}{20.95}\right] - P\left[Z < \frac{15-80.63}{20.95}\right]]$$

Fig 16.1: Graphical representation of distribution of the given data



From Fig16.1, it shows the graphical representation of the distribution of no. of girls who throw baseball at a certain distance (in feet).

CONCLUSION:

From the above analysis the average distance is covered by the girls is 80.63 feet. Maximum no. of girls throw baseball between 75-85 feet.

The expected number of girls throwing baseballs at a distance exceeding 105 feet on the basis that the data fit a normal distribution is 37 girls.

EXCEL OPTION:

1. Enter data in column.
2.Then create table according to your need. (using all border option present in the HOME tab).
3. To create histogram, select the data from the table then click on INSERT=>COLUMNS=> select the bar graph and set the bar gap at 0.
And then add graph title and suitable headings to the y-axis and x-axis.
Now, the histogram is created.
4. using Excel function for Normal Distribution (NORM.S.DIST())
5. using Excel function for Normal Distribution (NORM.DIST())

NAME-TANYA
ID-ABBSM21198
ROLL NO-2111793
SUBMITTED BY:-GARGI TYAGI