Date: 16 April, 2022

Practical 14: To fit a Poisson distribution to the given data of mutated DNA segment.

Question: The following mutated DNA segments were observed in 325 individuals:

Mutated DNA segments	0	1	2	3	4
Number of individuals	211	90	19	5	0

Fit a Poisson distribution to the data.

Methods and Formula:

- **Poisson Distribution:** The Poisson distribution is a discrete probability distribution that expresses the probability of a given number of events occurring in a fixed interval of time or space if these events occur with a known constant mean rate and independently of the time since the last event.

$$X \sim P(\lambda)$$

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

Mean = λ

The recurrence relation for Poisson probabilities is

$$p(x+1) = \frac{\lambda}{x+1}\, p(x), \quad p(0) = e^{-\lambda}$$

Expected Frequencies = N*p(x)

- **Bar graph:** A bar graph is a chart that plots data using rectangular bars or columns (called bins) that represent the total amount of observations in the data for that category.

Analysis:

To fit a Poisson distribution to the given data of mutated DNA segment. The tables 14.1 have been prepared in MS Excel.

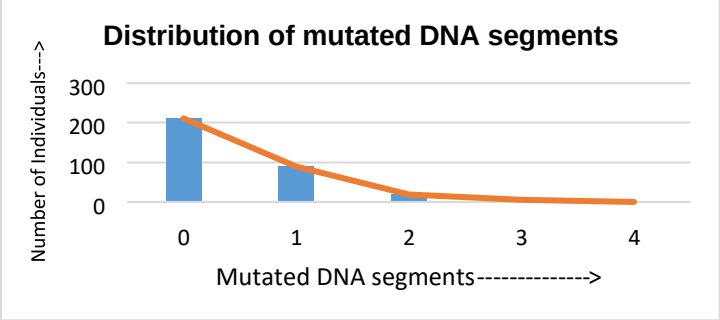
Table 14.1: Tabular form of Data for calculating Poisson distribution of the given data

Mutated DNA segments	Number of individuals	FX	Lamda /(x+1)	p(x)	Exp. Frequencies	Exp. Frequencies round off

0	211	0	0.44	0.64403642	209.311837	209
1	90	90	0.22	0.28337603	92.0972082	92
2	19	38	0.14666667	0.06234273	20.2613858	21
3	5	15	0.11	0.0091436	2.97166992	3
4	0	0	0.088	0.0010058	0.32688369	0
Total	**325**	**143**		**0.99990457**		**325**

From Table 14.1, this represents the calculation of recurrence relation for probabilities of Poisson distribution of the given data of mutated DNA segment.

Fig 14.1: Graphical representation of distribution of mutated DNA segment



From Fig14.1, it shows the graphical representation of the distribution of the mutated DNA segment. It shows the data is positively skewed.

CONCLUSION:

A mutation occurs when a DNA gene is damages or changed in such a way as to alter genetic message carried by that gene. The individuals having mutate DNA segments faces problems like colour- blindness, Tay-Sachs disease and many other. All of these disorders are caused by mutation of single gene.

From the above analysis we concluded that there are less than 50% chances of a person having mutation. As in the given data we can see that out of 325 individual 209 individual have 0 mutated DNA segments and if we talk about more than 4 mutated DNA segments then there is no individual present over here. It means on an average, there are either 0 or 1 mutated DNA segments found in individual.

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 bar graph, select the data from the table then click on INSERT=>COLUMNS=> select the bar graph.

And then add graph title and suitable headings to the y-axis and x-axis.

Now, the bar-graph is created.

4. Calculate probabilities under Poisson distribution

In excel, you can use any of the following three methods

Method 1: using pmf

Method 2: using recurrence relation for probabilities

$$p(x+1) = \frac{\lambda}{x+1} p(x), \quad p(0) = e^{-\lambda}$$

Method 3: using Excel function for binomial Probabilities (Poisson.dist())

Obtain the expected frequencies = N*p(x).

NAME-TANYA

ID-ABBSM21198

ROLL NO-2111793

SUBMITTED TO:-

DR.GARGI TYAGI