View metadata, citation and similar papers at core.ac.uk

brought to you by CORE

provided by ethesis@nitr

STUDY ON DELAY DISTRIBUTION OF TRAINS

A thesis submitted in partial fulfilment

Of the requirements for the degree of

Master of Technology

In

Civil engineering

(With Specialization in Transportation Engineering)

By

KANNEMADUGU REDDIKRISHNA

(213CE3091)

Under the Guidance of

Dr. Ujjal Chattaraj

Department of Civil Engineering

National Institute of Technology, Rourkela

Odisha-769008

MAY 2015

DEDICATED

TO

THE BEST TEACHERS OF THE WORLD

Department of Civil Engineering

National Institute of Technology

Rourkela-769008, Odisha, India

CERTIFICATE

This is to certify that the thesis entitled **"Study on delay distributions of train"** submitted by ***Mr.Kannemadugu ReddiKrishna (Roll No:213CE3091)*** in partial fulfillment of the requirements for the award of the degree of Master of Technology in Transportation engineering during the session 2013-2015 in the department of Civil engineering, National Institute of Technology Rourkela is an authentic work carried out by him under my supervision and guidance.

To the best of my knowledge, the matter embodied in the project report has not been submitted to any other University/Institute for the award of any Degree.

Date:

Dr. Ujjal Chattaraj

Asst. Professor

Department of Civil Engineering

National Institute of Technology, Rourkela

ACKNOWLEDGEMENTS

First and foremost I take this opportunity to express my deepest sense of gratitude to my guide ***Prof. U. Chattaraj for*** his able guidance during my project work. I would also like to extend my gratitude to ***Prof. M.Panda****, and* ***Prof. P.Bhuyan*** *Department of Civil Engineering,* for their kind permission and co-operation in data collection. This project would not have been possible without their direct and indirect help.

I am very thankful to Indian Railways by updating railway timetable that is helpful for my data collection.

I would like to thank Director ***Prof. S.Sarangi ,*** **NIT Rourkela** for giving me complete freedom to use their resources and work in such a challenging environment. The project would not have completed in stipulated time without his grace and support.

Last but not the least I would like to thank ***Prof. S.K.Sahu,*** *Head civil engineering department,* all of my friends and teachers who have been very cooperative and helped to me in completing my project in time.

DATE: -

KANNEMADUGU REDDIKRISHNA

ABSTRACT

This Thesis presents the reduction of strength to the Indian railway system by identifying passenger train delay. Now days, not all trains come on time at every station; it may delay anywhere in the given route due to operational problems or due to another train. The current timetable in India of Indian Railways demands dense traffic at various nodes on a network in sections and complex national and international dependencies.

Delay studies are very useful for finding out the efficiency of work, system and other operations. Identifying delay distribution of train is helpful to measure reliability of trains and railway system. Adding more trains to the calendar could cause danger to its stability in delay timings. This analysis is very useful to locate and eliminate some serious disturbances at various locations. A full range analysis of this data can lead to better identification of various delays and analysis their context. Contrast the distribution of arrival and departure timings of the train in different stations by using delay studies.

CONTENTS

ITEM PAGE NO.

List of Figures…… vi

List of Tables……. vii

CHAPTER 1: INTRODUCTION…….

1.1 General…… 1

1.2 Brief history about railway system……. 2

1.3 Classification of delay…..…… 3

1.4 Groups of Railway Statistics……... 3

CHAPTER 2: LITERATURE REVIEW…….

2.1 Literature review……. 4

2.2 Motivation…… 6

2.3 Problem statement…… 7

CHAPTER 3: DATA COLLECTION AND ANALYSIS……..

3.1Route selection…… 8

3.2 Data collection…… 9

3.3 Methods of Data collection …….……. 10

3.4 Data extraction…….……. 11

3.5 Studies of various statistical distributions……. 12

3.6 Calculation of various distribution from the collected data... 18

3.7 Discussions…… 40

CHAPTER 4: CONCLUSIONS……..…….. 42

REFERENCES

LIST OF FIGURES

List of figures	Abbreviation	page no.
Fig 1.1	Selected Route map	9
Fig 3.1	Normal Probability Distribution Curve	13
Fig 3.2	Lognormal distribution curve	15
Fig 3.3	Normal distribution curve	15
Fig 3.5	Weibull distribution curve	17
Fig 3.6	Various distribution graphs of Azad Hind Express	30-32
Fig 3.7	Various distribution graphs of Gitanjali Express	33-35
Fig 3.8	Various distribution graphs of Howrah Mail	36-38
Fig 3.9	Various distribution graphs of Jnaneswari Express	38-39

LIST OF TABLES

List of Tables	Abbreviation	page no.
Table 3.1	List of Train Data Collected Dates	10
Table 3.2	Azad Hind express Arrival distributions P-values	19
Table 3.3	Azad Hind express Departure distributions P-values	20
Table 3.4	Gitanjali express Arrival distributions P-values	21
Table 3.5	Gitanjali express Departure distributions P-values	22
Table 3.6	Howrah Mail Arrival distributions P-values	23
Table 3.7	Howrah Mail Departure distributions P-values	24
Table 3.8	Jnaneswari express Arrival distributions P-values	25
Table 3.9	Jnaneswari express Departure distributions P-values	25
Table 3.10	Karmabhoomi express Arrival distributions P-values	26
Table 3.11	Karmabhoomi express Departure distributions P-values	27
Table 3.12	Samarsata express Arrival distributions P-values	28
Table 3.13	Samarsata express Departure distributions P-values	29
Table 3.14	Skewness coefficient of different Trains at Rourkela station	40

Chapter 1

INTRODUCTION

1.1 General

In traffic control system travel time management is an important problem to improve railway system. Determining the amount of time delay to travel from any two points along the route requires a travel time study and collected various causes for delay. Data obtained from travel time and delay studies gives good indication of the LOS on the section. The travel time and delay data uses to improve operation efficiency and safety, and also locates where design and operational problem ceases to traffic. Since from historical days IR transport is suffering some of the major problems, one of the major problems is delays and their effects around various types of networks. These will sometimes cause accidents and various kinds of disruptions. Some of the operations of a signaling system, the IR use colour signal lights, but in some areas of operation, the discs-based signaling and older semaphores are still in use that causes delays. This transportation will cause difficulty to the passengers and affect their daily duties. Small disturbances in one rail section have an impact at another rail; this impact should be more in single track segments compared to multiple track segments. Therefore it is important to study on passenger train delays problem, which helps to minimize delay timings.

Some of the theory is copied from Google, Wikipedia and papers like Methods of data collection and limitations of statics are from railway statistics.

1.2 The railway system: a brief history

Indian Railways (IR) play a vital role in not only meeting the transport needs of the country, but also in cover mutually dispersed areas and strengthening social integration.IR have been prime movers in the community and produce the detail about largest railway systems in the world under a single management. Railways are the best suitable for bulk commodities transport for a long distance. In the national emergency, railways have been in the forefront in rushing relief material to disaster-stricken regions.

The IR system is manageable through zones and operating divisions, six production units, wheels and axles and other ancillary components to meet railways requirements.

Transport is an important part of India's economy, the major part from transport considers is covered by railways. Rail transport is one of the major transports used for travelling long distances within a country. In India almost all rail operations are handled by a state-owned organization, Indian Railways. Indian rail transport is 4th largest rail network in the world carrying billion of passengers, and this number of passengers increases annually. Railways were introduced to India in 1853 from Mumbai to Thane and they had been growing to large extent. In past days steam locomotives have been used and later replaced with diesel and electric locomotives. The maximum number of travelers choosing rails as their main transport mode. Rail transportation plays significant role in the development of sustainable transportation opportunities.

Train delays are more complex to study as they are increases in a selected route. Delays are classified according to their functioning.

1.3 Classification of delays

a) Track related delays: - Those delays which caused by train having to slow down due to track problems and delays caused by complete stoppage for a particular time.

b) Train dependent delays: - The delays which caused by a train that breaks down or is forced to slow down in line sections.

c) Terminal/schedule stop delay: - The delays which occur at scheduled stops and at origin station. This may involve delays associated with loading/unloading, train connections, fueling and crew changes.

1.4 Groups of Railway Statistics

1. Economic and financial statistics: - Include approximate figures of earnings, various funds, expenditure etc. The figures of gross receipts, gain/ loss etc. can be seen with the help of financial statistics.
2. Operating statistics

   a) Traffic: - It studies on wagon load, wagon mobility, wagon uses, punctuality etc.

   b) Power: - It studies on engine uses, coal and fuel consumption, engine failure statistics.
3. Commercial statistics: - In this statistics we studied on passenger and goods carried, passenger kms, average load of traffic etc.
4. Rolling stock and workshop Repair statistics
5. Administrative Statistics: - These statistics relating the staff matters, physical statistics like no. of stations, curvature of p-way, gradients etc.

Chapter 2

As explained briefly in the previous chapters, understanding train movement in a particular segment were more station points and traffic density is high; it is difficult to study delay variation of trains due to many factors. Possibly passenger trains run slowly those cause delay of express and other trains. In India train delay is more compared to other countries where population increases rapidly that causes one of the main traffic problem. Some trains cause accidents due to delays at some other locations in a segment.

In this chapter a review is presented on past work related to delay distributions, distributions type suited for various types of delay i.e. arrival and departure delays at every station in the selected route.

2.1 Literature Review

Schwanhausser (1974) is probably the first study resolving that the tail of arrival delays observes the negative exponential distribution for arrival delays of passenger trains at the stations. Later studies are confirmed this result by using distribution models, like Weibull distribution, the gamma distribution and the lognormal distributions. These models better capture the distribution of late arrival and departure times

Goverde et al. (2001) use a software tool to estimate precise arrival and departure times of trains at a station and extracts the occupancy and clearance times of each train. They proved that

the arrival and dwell times fitted normality statistical distribution, and departure follows negative exponential distribution. They determined that most of the delay times follow standard and negative exponential distributions respectively will allow an improved design of timetable by using historical data.

Goverde (2001) performs accurate operation from variation in actual departure times, arrival times and fluctuations due to other reasons. For these services data a software tool, TNV-Prepare, has been developed which helps to recovers successive track occupancies and clearances of trains from the record. Using regression analysis interrelationships between arrivals and departures of a particular train are identified and prediction models from independent event times are derived.

Ullius (2005) uses Open Timetable analytical tool to operational data that automatically captured by the SBB network. The results can be used by railways to measures relating operations and for optimization of the timetable. Open Timetable is one of series of programming to planning, scheduling on existing network to help railroad planners. It displays statistical output data by comparing actual and planned timetables.

Yuan (2006) compares the goodness-of-fit among several distribution models selected for train event times and process times by fine-tuning the distribution parameters for data recorded at The Hague railway station. He finds that the Weibull distribution can be considered as the best distribution model for arrival delays, departure delays and the free dwell times of trains.

Güttler (2006) fits a normal-lognormal mixed distribution to assess running times of trains between two stations using the data obtained for the German railway.

Murali et al. (2009) presents a simulation-based technique to generate delay estimates by taking a function of traffic and topology in a network is taken into account. They estimated delay

data for a network by using Los Angeles and simulation model developed by Lu et al. (2004) which has been representative of the real world values and compared them. They present a delay estimation methodology that defines an exponential relation between train delays and their mix.

Dingler et al. (2010) use simulation software tool known as Rail Traffic Controller (RTC) to determine best capacity of the project and analyze train delay. Reduction in delay is often used by the railroads to calculate the benefit of a project or operational change. RTC is also used to calculate the impact of various mechanisms that affect train delays. The primary output from the simulation software is train delay; delays were categorized by conflict and source.

Ghosh et al. (2013) perform statistical analysis of the run-time delays of express trains in IR, which are caused not being able efficiently to handle the rapidly increasing amounts of traffic. They also develop a Web-based application for visualizing traffic-flow and identify several problems in IR in the IGP zone.

2.2 MOTIVATION

It is indisputable from the research look over in this chapter that till now many numbers of experiments had done on train delays by using different software's. Train delay parameters change spatially and temporally, notably, in response to daily changes of speed. Also, there is no experimental study in India to find out train delay distributions. These motivated the empirical studies conducted in this thesis.

2.3 PROBLEM STATEMENT

The problem of this thesis can be stated as “study on train delay distribution”. Trains running in India observed on various days in railway stations on selected route that have significant density. Here we are studying about last type of delay as given in above chapters and also calculation and best fitting distribution curve.

Chapter 3

The IR network consists of 115,000 track-km, of which IR administration owns and manages approximately 95 percent. The railway network in India spans the entire country, but due to very uneven population distribution, the railway network is denser in the metropolitan areas. The major part of the Indian railway network has single tracks, and only around 30 percent of the system has two or more parallel tracks. The proportion of passenger train is about 70 percent of passengers; however the extent of traffic blend varies a lot between different routes. IR network disturbances due to vulnerable massive railway infrastructure.

3.1 ROUTE SELECTION

In this section, we identify the major sections in the India, there are around 8000 railway stations, from those I selected the route that connected between two main stations i.e. Mumbai Central to Howrah station (Kolkata). The distance between two selected stations is around 1968 km that connects 26 railway stations in between them. I picked this route because the way that passes through Rourkela that connecting central station and having regular, weekly and special trains. The route Howrah to Mumbai via Nagpur line is a part of the golden quadrilateral, carryout half the freight and nearly half the passenger traffic. These cities are metropolitan cities that having more population and large number of population travel in this route.

INDIA
RAILWAY MAP
Srinagar
Jammu Tawi
Amritsar
Jalandhar City
Ludhiana
Chandigarh
Ambala Cantt.
Shimla
Dehradun
Kurukshetra
Panipat
Moradabad
NewDelhi
Kasganj
Bareilly
Shahjahanpur
Farrukhabad
Lucknow
Kanpur
Agra
Mathura
Bharatpur
Jaipur
Ajmer
Marwar
Jodhpur
Bikaner
Jaisalmer
Phalodi
Merta Rd.
Sawar Madhopur
Jhansi
Allahabad
Varanasi
Patna
Gangtok
Siliguri
New Jalpaiguri
Dekargaon
Amguri
Haibargaon
Lumding
Dispur
Naharlagun
Itanagar
North Lakhimpur
Dangari
Silchar
Agartala
Kota
Chittaurgarh
Udaipur
Palanpur
Bhildi
Ruthiyai
Guna
Bina
Satna
Chunar
Mughalsarai
Gaya
Madhupur
Bhagalpur
Katihar
Sitarampur
Barddhaman
Calcutta
Ranchi
Bokaro
Tatanagar
Rourkela
Kharagpur
Jharsuguda
Rupsa
Jakhapura
Bhubaneshwar
Puri
Titlagarh
Abhanpur
Raipur
Durg
Nagpur
Bilaspur
Raigarh
Katni
Jabalpur
Bhopal
Ujjain
Indore
Khandwa
Itarsi
Nainpur
Chhindwara
Tumsar Rd
Gondia
Balaghat
Sambalpuri
Badampahar
Ahmedabad
Vadodara
Rajkot
Dhola
Jetalsar
Porbandar
Prachi Rd.
Bharuch
Surat
Udhna
Jalgaon
Akola
Achalpur
Amla
Arvi
Khamgaon
Aurangabad
Yavatmal
Tadali
Rajhara
Dalli
Dhamtari
Manikgarh
Singapur Rd.
Gunupur
Koraput
Naupada
Vizianagaram
Kothavalasa
Samalkot
Vijayawada
Gudivada
Kotipalli
Guntur
Nadikudi
Hyderabad
Dornakal
Kazipet
Nizamabad
Mudkhed
Parbhani
Kalyan
Mumbai
Pune
Kurduwadi
Baramati
Miraj
Wadi
Belgaum
Hubli
Hospet
Guntakal
Nandyal
Cuddapah
Krishnapatnam
Gudur
Renigunta
Chennai (Madras)
Bangalore
Hosur
Jolarpettai
Villupuram
Puducherry
Cuddalore
Vriddhachalam
Tiruchchirappalli
Nagore
Agasthiampalli
Dindigul
Madurai
Virudhunagar
Rameswaram
Tuticorin
Tirunelveli
Tiruchendur
Kanyakumari
Nagercoil
Thiruvananthapuram
Kayankulam
Tenkasi
Ernakulam
Thrissur
Pollachi
Karur
Coimbatore
Salem
Ooty
Kozhikode
Shoranur
Mysore
Hassan
Mangalore
Chikjajur
Dharmavaram
Birur
Chikballapur
Udupi
Bhatkal
Vasco-da-gama
Miraj
Ratnagiri
Rajapur Rd.
Shri Chhatrapati Sahu Maharaj
Panaji
Logda
Roha
Veer
Khed
Delvadi
Porbandar
Jetalsar
Patliana
Rajkot
Kanalus
New Bhuj
Gandhinagar
Ahmedabad
Ratlam
Nagda
Maksi
Mavli
Udaipur
Khedbrahma
Bali Sadri
Phulera
Alwar
Rohtak
Rewari
Sardarshahar
Lalgarh
Suratpura
Sikar
Churu
Sri Ganga Nagar City
Bathinda
Jakhal
Suratgarh
Anupgarh
Hisar
Dhuri
Fazilka
Patiala
Jalandhar City
Amritsar
Dera Baba Nanak
Lahara
Hoshiarpur
Udhampur
Katra
Banihal
Qazigund
Baramulla
Munabao
Barmer
Luni
Sheopur Kalan
Khajuraho
Manikpur
Allahabad
Sultanpur
Lucknow
Faizabad
Nepalganj
Jarwa
Nautanwa
Narkatiaganj
Seamarhi
Bihar
Jayanagar
Siwan
Chhapra
Gorakhpur
Kathgodam
Lalkua
Kotdwar
Raiwala
Dehradun
Shimla
Chandigarh
Ambala Cantt.
Dhuri
Patiala
Kurukshetra
Panipat
Rohtak
Moradabad
Bareilly
Shahjahanpur
Farrukhabad

each departing trains. However, the delay given in the website is not clear that the delay of the train is in a minute is between a delay of 0 and 1 minute or between 1 and 2 minutes.

TRAIN NAME	DATES	
	FROM	TO
AZAD HIND EXPRESS	14-01-2015	13-02-2015
GITANJALI EXPRESSS	14-01-2015	14-02-2015
HOWRAH MAIL	14-01-2015	15-02-2015
JNANESWARI DELUX	16-01-2015	16-02-2015
KARMABHOOMI EXPRESS	14-01-2015	18-02-2015
LTT SHALIMAR EXPRESS	13-01-2015	18-02-2015
SAMARSATA EXPRESS	15-01-2015	18-02-2015

TABLE 3.1 List of Train Data Collected Dates

3.3 METHODS OF DATA COLLECTION

Following are the important methods of data collecting for required information.

1) Station to station Method
2) Density method
3) Car method
4) Exemption method

Station to station Method: - The method consists of compiling the quantity of traffic, which has passed between each pair of station during the period.

Density Method: - In this method we record against name of station of the traffic (a) Received at the station (b) Forward from the station (c) Passing the station.

Car Method: - This method we worked out with the help of combined train engine ticket report, guard's prepared by loco pilots and guard of the train.

Exemption Method: - In this method we worked out total units arising from the standard running of train variations (plus or minus) from the standard are reported every month by

officials. This method can be successfully employed for items such as seats kms, passenger vehicle kms etc.

The following limitations of statistics should not altogether be ignored: -

(a) Statistical methods are more dangerous tools in expert hands.

(b) The manipulated figures, arrangement and design of effective data when presented in support of argument.

(c) Statistical laws are not like the exact laws of physical science which are said to hold true in every case that subjected to them.

(d) Planning on the basis of inadequate and in accurate statistics is worse than no planning at all.

3.4 DATA EXTRACTION

After collection of all data required for projects in the selected route. The data collected over 30 days of all trains that are travel from origin to destination and so on. The data collected having scheduled time of arrival and departure of all stations, actual arrival and departure timing that is update by the IR. From these data we calculate the delay timings of trains in various stations; run-time delay on the station is the difference between actual time and scheduled time (d_a-d_s). The delay will be negative, zero or positive values, if the train is arrived/departure early than the scheduled time. Plot the graph between delay timing and days for collected data and find out the best fitting curve to the data. The average run-time delays of train for 30days are computed by using type of best fitting curve.

3.5 STUDIES OF VARIOUS STATISTICAL DISTRIBUTIONS

Probability theory provides models of probability distributions to be tested by statistical methods and applied on engineering applications, long list field applications etc. We have to handle obtained data from field numerically or in terms of graphs, these are influenced by chance by factors whose effect we cannot predict exactly, therefore we have to implement probability theory. Graphical representation like stem and leaf plot, Histogram, Box plot, Median, Interquartile range Outlier.

The average size of the data values can be measured by using

$$\overline{\mathrm{x}} = 1/\mathrm{n}\sum_{\mathrm{j}=1}^{\mathrm{n}} x_j = 1/\mathrm{n}\,(\mathrm{x}_1 + x_2 + \cdots . \mathrm{x}_n)$$

Similarly standard deviation's' and variance are obtained as follows

$$\mathrm{S}^2 = \frac{1}{\mathrm{n}-1}\sum_{\mathrm{j}=1}^{\mathrm{n}}\left(\mathrm{x}_j - \overline{\mathrm{x}}\right)^2 = \frac{1}{\mathrm{n}-1}\left[(\mathrm{x}_1 - \overline{\mathrm{x}})^2 + \cdots + (x_n - \overline{\mathrm{x}})^2\right]$$

The frequency distribution of data shows relative frequency that is similar to probability distribution, showing the probabilities of the events in an experiment.

Let quantity that we observe in an experiment will be denoted as "X" which called as random variable. We have discrete random variable and distribution when we count the data; we have continuous random variable and distribution when we measure the data.

Discrete function $\mathrm{F(x)} = \mathrm{P}(\mathrm{x} \leq x)$

Note: - The "X" will assume any value that not exceed x to generate probability in the trail.

From the above formula we obtain the fundamental formula for probability corresponding to an interval $a < x \leq b$.

$$p(a < X \leq b) = F(b) - F(a)$$

The value for discrete random variable distribution function is

$$F(x) = \sum_{X_{j \leq X}} f(x_j) = \sum_{Xj \leq X} p_j$$

The value for continuous random variable distribution function is

$$p(a < X \leq b) = F(b) - F(a) = \int_a^b f(v)\, dv$$

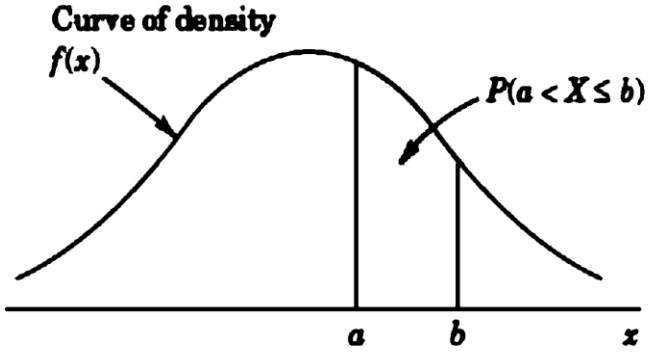


Fig 3.1 Normal Probability Distribution Curve

Mean and variance of a distribution

$\mu = \sum_j x_j f(x_j) \quad and\ \sigma^2 = \sum_j (x_j - \mu)^2 f(x_j)$ For discrete distribution

$\mu = \int_{-\infty}^{\infty} x\, f(x) dx \ \ and \ \ \sigma^2 = \int_{-\infty}^{\infty} (x_j - \mu)^2 f(x) dx$ for continuous distribution

Bernoulli distribution

The distribution of "X" with probability (X=x) of X=x, that is of obtaining a precisely x times in n trails. Hence "X" has the probability function of

$$f(x) = \left\langle {n \atop x} \right\rangle p^x q^{n-x} \qquad (x=0, 1, 2\ldots n)$$

The distribution function $f(x) \neq 0$ is called the binomial distribution or Bernoulli distribution.

Mean and variance of the Binomial distribution

$$\mu = np \qquad and \quad \sigma^2 = npq$$

Lognormal distribution:

When random variable of Y=In(X) has a normal distribution if X having nonnegative random variable is said to be lognormal distribution. The probability distribution function of random variable when In(X) is normally distributed with parameters μ and σ is

$$f(X;\mu,\sigma) = \begin{cases} \frac{1}{\sqrt{2\pi\sigma X}} e^{-\frac{[\ln(x)-\mu]^2}{2\sigma^2}} & X \leq 0 \\ 0 \end{cases}$$

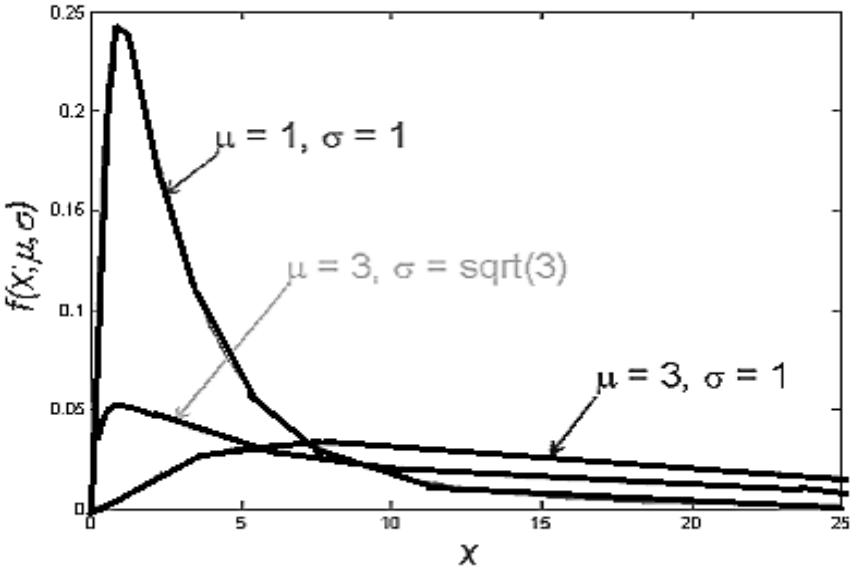


Fig 3.2 Lognormal distribution curve

Normal distribution

In this we discuss the normal distribution that turns from discrete to continuous distributions. The normal distribution or Gauss distribution is defined as the distribution with density.

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} exp\left[-\frac{1}{2}\left(\frac{x-y}{\sigma}\right)^2\right]$$

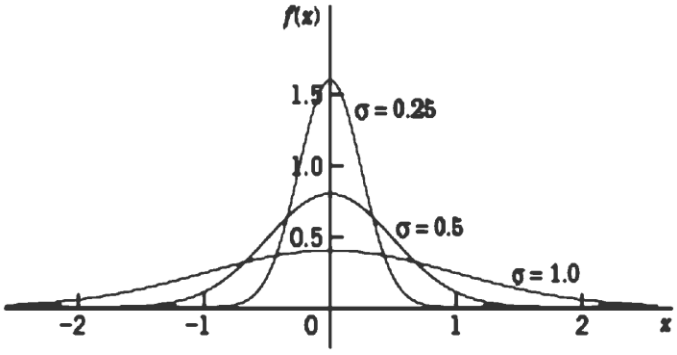


Fig 3.3 Normal distribution curve

Distribution parameters to be estimated by using two methods (i) Method of moments. (ii) Maximum likelihood Method.

Method of moments: - In this method the parameters to be estimated are expressed in terms of the moments of the distribution. From the basic formula K_{th} moment of a sample $x_1 \ldots x_n$ is

$$m_k = \frac{1}{n}\sum_{j=1}^{n} x_j^k$$

Maximum Likelihood Method: - In this method probability function depends on a single parameterθ, when we consider a discrete random variables.

Let $x_1 \ldots x_n$ are the n independent samples

Sample size $l = f(x_1)f(x_2) \ldots \ldots f(x_n)$

If l is the differentiable function of θ, a necessary condition for l to have maximum in an interval is$\frac{dl}{d\theta} = 0$. This is called maximum likelihood estimate

Poisson distribution

The discrete distribution with finitely many possible values and probability function is called the Poisson distribution.

$$f(x) = \frac{\mu^x}{x!}e^{-\mu} \qquad (x=0, 1 \ldots)$$

Mean and variance of Poisson distribution $\mu = \sigma^2 = np$ for θ.

Weibull distribution

Random variable X in a sample taken having a Weibull distribution with parameters α and β (α>0, β>0).

The probability distribution function of X is

$$f(x;\alpha,\beta) = \begin{cases} \frac{\alpha}{\beta^{\alpha}} x^{\alpha-1} e^{-(x/\beta)^{\alpha}} & x \geq 0 \\ 0 & x < 0 \end{cases}$$

We can explain a special cases of gamma and Weibull distribution by using exponential distribution with parameter $\lambda = \frac{1}{\beta}$ and $\propto = 1$ for the above equation.

Some of the gamma distributions are not Weibull distributions, therefore we conclude that one family is not a subset of other.

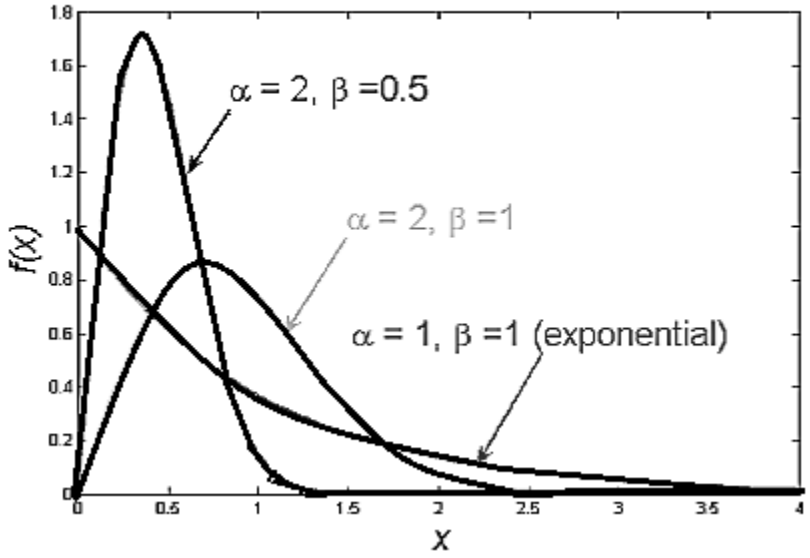


Fig 3.4 Weibull distribution curve

3.6 CALCULATION OF VARIOUS DISTRIBUTION FROM THE COLLECTED DATA

As we discussed in the previous chapter extraction of data and statistical studies. In this chapter we studied the application of statistics on data extraction. Before starting calculation first set the significance level and method of estimation (Moments method or Maximum Likelihood Method), calculate mean and variance of the data then calculate P-values of various distribution. The distribution having maximum P-value is the best fitting curve for the data having.

P-VALUES								
STATION NAME	Binomial	Exponential	Log-normal	Logistic	Normal	Poisson	Weibull	REMARKS/BEST FITTING CURVE
Daund Jn	-	-	-	< 0.0001	< 0.0001	-	**0.012**	Weibull
Ahmad Nagar	< 0.0001	0.033	**0.239**	< 0.0001	< 0.0001	< 0.0001	0.180	Log-normal
Belapur	< 0.0001	0.006	**0.248**	< 0.0001	< 0.0001	< 0.0001	0.160	Log-normal
Kopargaon	-	-	-	< 0.0001	< 0.0001	-	**0.031**	Weibull
Manmad Jn	-	0.013	0.070	< 0.0001	< 0.0001	-	**0.180**	Weibull
Jalgaon Jn	< 0.0001	0.015	**0.103**	< 0.0001	< 0.0001	< 0.0001	0.018	Log-normal
Bhusaval Jn	-	-	-	< 0.0001	< 0.0001	-	**0.010**	Weibull
Malkapur	< 0.0001	0.015	**0.176**	< 0.0001	< 0.0001	< 0.0001	0.029	Log-normal
Shegaon	< 0.0001	0.016	**0.204**	< 0.0001	< 0.0001	< 0.0001	0.075	Log-normal
Akola Jn	< 0.0001	0.014	**0.060**	< 0.0001	0.000	< 0.0001	0.024	Log-normal
Badnera Jn	-	-	-	< 0.0001	< 0.0001	-	**0.001**	Weibull
Wardha Jn	-	-	-	< 0.0001	< 0.0001	-	**0.406**	Weibull
Nagpur	-	-	-	< 0.0001	< 0.0001	-	**0.015**	Weibull
Bhandara Road	< 0.0001	< 0.0001	0.082	< 0.0001	< 0.0001	< 0.0001	**0.104**	Weibull
Tumsar Road	< 0.0001	< 0.0001	**0.055**	< 0.0001	< 0.0001	< 0.0001	0.048	Log-normal
Gondia Jn (G)	< 0.0001	0.000	**0.300**	< 0.0001	< 0.0001	< 0.0001	0.100	Log-normal
Raj Nandgaon	< 0.0001	0.003	-	< 0.0001	0.001	< 0.0001	**0.048**	Weibull
Durg	-	-	-	< 0.0001	0.000	-	**0.166**	Weibull
Raipur Jn	< 0.0001	< 0.0001	-	< 0.0001	0.000	< 0.0001	**0.180**	Weibull
Bilaspur Jn	-	-	-	< 0.0001	< 0.0001	-	**0.347**	Weibull
Champa	-	-	-	< 0.0001	0.000	-	**0.099**	Weibull
Raigarh	< 0.0001	0.010	0.141	< 0.0001	< 0.0001	< 0.0001	**0.211**	Weibull
Jharsuguda Jn	-	-	-	< 0.0001	< 0.0001	-	**0.076**	Weibull
Rourkela	< 0.0001	0.024	**0.051**	< 0.0001	0.000	< 0.0001	0.017	Log-normal
Chakradharpur	< 0.0001	0.048	**0.214**	0.055	0.000	< 0.0001	0.071	Log-normal
Tata-gar Jn	-	-	-	0.022	0.001	-	**0.197**	Weibull
Kharagpur Jn	< 0.0001	0.042	-	0.040	0.001	< 0.0001	**0.138**	Weibull

Table 3.2 Azad Hind express Arrival distributions P-values

P-VALUES								
STATION NAME	Binomial	Exponential	Log-normal	Logistic	Normal	Poisson	Weibull	REMARKS\BEST FITTING CURVE
Daund Jn	< 0.0001	< 0.0001	-	< 0.0001	< 0.0001	0.015	0.006	Poisson
Ahmad nagar	< 0.0001	0.045	0.289	< 0.0001	0.000	< 0.0001	0.201	Log-normal
Belapur	< 0.0001	0.008	0.247	< 0.0001	< 0.0001	< 0.0001	0.190	Log-normal
Kopargaon	-	< 0.0001	0.001	< 0.0001	< 0.0001	-	0.019	Weibull
Manmad Jn	< 0.0001	0.010	0.058	< 0.0001	< 0.0001	< 0.0001	0.018	Log-normal
Jalgaon Jn	< 0.0001	0.013	0.082	< 0.0001	< 0.0001	< 0.0001	0.010	Log-normal
Bhusaval Jn	< 0.0001	0.009	0.289	< 0.0001	< 0.0001	< 0.0001	0.229	Log-normal
Malkapur	< 0.0001	0.022	0.274	< 0.0001	0.000	< 0.0001	0.261	Log-normal
Shegaon	< 0.0001	0.023	0.255	< 0.0001	0.000	< 0.0001	0.101	Log-normal
Akola Jn	< 0.0001	0.022	0.115	< 0.0001	0.000	< 0.0001	0.050	Log-normal
Badnera Jn	< 0.0001	< 0.0001	-	< 0.0001	< 0.0001	< 0.0001	0.014	Weibull
Wardha Jn	< 0.0001	< 0.0001	-	< 0.0001	< 0.0001	0.164	0.000	Poisson
Bhandara Road	< 0.0001	< 0.0001	0.043	< 0.0001	< 0.0001	< 0.0001	0.022	Log-normal
Tumsar Road	< 0.0001	< 0.0001	0.059	< 0.0001	< 0.0001	< 0.0001	0.047	Log-normal
Gondia Jn (G)	< 0.0001	0.000	0.301	< 0.0001	< 0.0001	< 0.0001	0.100	Log-normal
Raj Nandgaon	< 0.0001	0.001	-	< 0.0001	0.001	< 0.0001	0.049	Weibull
Durg	< 0.0001	< 0.0001	-	< 0.0001	0.000	< 0.0001	0.104	Weibull
Raipur Jn	< 0.0001	< 0.0001	-	< 0.0001	0.000	< 0.0001	0.258	Weibull
Bilaspur Jn	< 0.0001	0.007	-	< 0.0001	0.000	< 0.0001	0.249	Weibull
Champa	< 0.0001	0.018	-	< 0.0001	0.000	< 0.0001	0.125	Weibull
Raigarh	< 0.0001	0.014	0.185	< 0.0001	< 0.0001	< 0.0001	0.035	Log-normal
Jharsuguda Jn	< 0.0001	0.008	0.151	< 0.0001	< 0.0001	< 0.0001	0.032	Log-normal
Rourkela	< 0.0001	0.025	0.058	< 0.0001	0.000	< 0.0001	0.020	Log-normal
Chakradharpur	< 0.0001	0.048	0.212	0.055	0.000	< 0.0001	0.069	Log-normal
Tata-gar Jn	< 0.0001	0.000	-	0.006	0.002	< 0.0001	0.104	Weibull
Kharagpur Jn	< 0.0001	0.023	-	0.034	0.001	< 0.0001	0.352	Weibull

Table 3.3 Azad Hind express Departure distributions P-values

P-VALUES									
STATION NAME	Beta	Binomial	Exponential	Gamma	GEV	Log-normal	Logistic	Weibull	REMARKS\BEST FITTING CURVE
Kalyan Jn	0.125	< 0.0001	< 0.0001	**0.763**	0.635	0.690	0.614	0.572	Gamma
Igatpuri	< 0.0001	-	-	-	< 0.0001	-	**0.000**	< 0.0001	Logistic
Jalgaon Jn	0.579	< 0.0001	0.000	**0.996**	0.993	0.983	0.925	0.992	Gamma
Bhusaval Jn	0.469	< 0.0001	0.091	0.706	**0.894**	-	0.648	0.638	GEV
Malkapur	0.573	< 0.0001	0.004	0.801	0.866	**0.953**	0.573	0.735	Log-normal
Shegaon	**0.913**	< 0.0001	0.021	0.822	0.677	0.656	0.637	0.776	Beta
Akola Jn	**0.560**	< 0.0001	0.001	0.506	0.331	0.283	0.312	0.371	Beta
Badnera Jn	0.008	-	-	-	0.054	-	**0.068**	0.016	Logistic
Wardha Jn	0.716	-	-	-	0.740	-	0.849	**0.989**	Weibull
Nagpur	0.363	-	-	-	**0.439**	-	0.285	0.258	GEV
Bhandara Road	**0.992**	< 0.0001	0.019	0.876	0.844	0.925	0.833	0.945	Beta
Gondia Jn (G)	**0.940**	-	-	-	0.539	-	0.895	0.661	Beta
Raj Nandgaon	0.006	-	-	-	0.152	-	**0.526**	0.050	Logistic
Durg	0.357	-	-	-	**0.911**	-	0.786	0.824	GEV
Raipur Jn	0.062	< 0.0001	0.898	0.228	0.000	-	0.200	**0.954**	Weibull
Bilaspur Jn	0.023	-	-	-	0.000	-	**0.491**	0.431	Logistic
Raigarh	0.002	< 0.0001	0.063	0.005	0.000	**0.371**	< 0.0001	0.116	Log-normal
Jharsuguda Jn	0.002	-	-	-	0.000	-	< 0.0001	**0.107**	Weibull
Rourkela	0.006	-	-	-	0.000	-	< 0.0001	**0.297**	Weibull
Chakradharpur	0.012	-	-	-	0.000	-	0.168	**0.579**	Weibull
Kharagpur Jn	0.002	-	-	-	0.000	-	< 0.0001	**0.079**	Weibull

Table 3.4 Gitanjali express Arrival distributions P-values

P-VALUES									
STATION NAME	Beta	Negative binomial	Exponential	Gamma	GEV	Log-normal	Logistic	Weibull	REMARKS\BEST FITTING CURVE
Dadar	< 0.0001	< 0.0001	< 0.0001	< 0.0001	**0.000**		< 0.0001	< 0.0001	GEV
Kalyan Jn	0.022	0.152	< 0.0001	0.144	0.222	0.200	**0.333**	0.078	Logistic
Igatpuri	0.000	< 0.0001	< 0.0001	< 0.0001	0.000	< 0.0001	0.000	**0.000**	Weibull
Nasik Road	0.531	**0.992**	< 0.0001	0.886	0.718	0.877	0.799	0.927	Negative binomial
Jalgaon Jn	0.344	0.257	0.000	0.985	**0.994**	0.978	0.934	0.942	GEV
Bhusaval Jn	0.485	0.093	0.081	0.465	0.694	-	**0.844**	0.820	Logistic
Malkapur	0.639	0.464	0.002	0.838	0.810	**0.950**	0.629	0.802	Log-normal
Shegaon	**0.886**	0.411	0.007	0.781	0.639	0.718	0.618	0.739	Beta
Akola Jn	0.613	0.383	0.002	**0.660**	0.472	0.409	0.389	0.491	Gamma
Badnera Jn	0.015	< 0.0001	**0.064**	0.064	0.043	-	0.058	0.035	Exponential
Wardha Jn	**0.955**	0.757	0.324	0.304	0.058	-	0.443	0.952	Beta
Bhandara Road	0.966	0.881	0.023	0.936	0.861	0.970	0.875	**0.978**	Weibull
Raj Nandgaon	0.004	0.000	0.011	0.075	0.132	-	**0.214**	0.069	Logistic
Durg	0.428	0.459	0.317	0.629	0.614	-	0.732	**0.938**	Weibull
Bilaspur Jn	0.011	< 0.0001	0.319	0.008	< 0.0001	-	< 0.0001	**0.628**	Weibull
Raigarh	0.002	< 0.0001	0.063	0.005	0.000	**0.369**	< 0.0001	0.115	Log-normal
Jharsuguda Jn	0.002	< 0.0001	< 0.0001	< 0.0001	< 0.0001	-	< 0.0001	**0.032**	Weibull
Rourkela	0.007	0.001	0.090	0.053	0.000	-	< 0.0001	**0.346**	Weibull
Chakradharpur	0.020	0.020	0.045	0.057	< 0.0001	-	0.042	**0.671**	Weibull
Kharagpur Jn	0.014	**0.074**	< 0.0001	< 0.0001	< 0.0001	-	< 0.0001	0.007	Negative binomial

Table 3.5 Gitanjali express Departure distributions P-values

P-VAUE								
STATION NAME	Beta	Exponential	GEV	Log-normal	Logistic	Poisson	Weibull	REMARKS\BEST FITTING CURVE
Dadar	< 0.0001	< 0.0001	< 0.0001	< 0.0001	**< 0.0001**	< 0.0001	< 0.0001	Logistic
Thane	< 0.0001	< 0.0001	0.000	0.001	**0.001**	< 0.0001	< 0.0001	Logistic
Kalyan Jn	0.839	< 0.0001	0.750	0.840	0.615	**0.840**	0.696	Poisson
Nasik Road	0.187	0.000	0.307	**0.755**	0.467	< 0.0001	0.647	Log-normal
Manmad Jn	0.060	0.005	< 0.0001	**0.958**	0.897	< 0.0001	0.935	Log-normal
Chalisgaon Jn	0.131	0.010	< 0.0001	**0.983**	0.837	< 0.0001	0.947	Log-normal
Jalgaon Jn	0.084	0.005	0.583	0.434	**0.872**	< 0.0001	0.556	Logistic
Bhusaval Jn	0.150	0.032	0.842	0.736	**0.887**	< 0.0001	0.726	Logistic
Malkapur	0.237	0.006	0.646	0.442	**0.935**	< 0.0001	0.653	Logistic
Nandura	0.204	0.004	0.769	0.618	**0.892**	< 0.0001	0.764	Logistic
Jalamb Jn	0.264	0.001	0.687	0.598	**0.953**	< 0.0001	0.671	Logistic
Shegaon	0.309	0.001	0.937	0.888	**0.999**	< 0.0001	0.961	Logistic
Akola Jn	0.283	0.001	0.608	0.511	**0.886**	< 0.0001	0.683	Logistic
Badnera Jn	0.011	-	0.009	-	**0.435**	-	0.144	Logistic
Dhamangaon	0.253	-	0.992	-	**0.999**	-	0.884	Logistic
Pulgaon Jn	0.430	0.274	< 0.0001	0.553	0.681	< 0.0001	**0.955**	Weibull
Wardha Jn	0.289	-	0.733	-	**0.873**	-	0.712	Logistic
Bhandara Road	0.616	0.002	0.394	0.799	0.604	< 0.0001	**0.812**	Weibull
Tumsar Road	0.755	0.015	0.797	0.890	0.772	< 0.0001	**0.897**	Weibull
Gondia Jn	0.579	-	0.527	-	0.700	-	**0.985**	Weibull
Dongargarh	0.854	0.098	0.764	0.466	0.573	< 0.0001	**0.943**	Weibull
Durg	0.047	-	< 0.0001	-	**0.541**	-	0.201	Logistic
Raipur Jn	0.017	-	0.000	-	< 0.0001	-	**0.262**	Weibull
Bilaspur Jn	0.005	-	0.000	-	< 0.0001	-	**0.391**	Weibull
Sakti	0.001	0.193	< 0.0001	**0.929**	< 0.0001	< 0.0001	0.154	Log-normal
Kharsia	0.001	0.131	< 0.0001	**0.756**	< 0.0001	< 0.0001	0.470	Log-normal
Raigarh	0.001	0.113	< 0.0001	**0.439**	< 0.0001	< 0.0001	0.177	Log-normal
Brajraj-gar	0.056	0.005	< 0.0001	**0.477**	0.427	< 0.0001	0.315	Log-normal
Chakradharpur	0.085	-	0.000	-	0.330	-	**0.640**	Weibull
Tatanagar Jn	0.006	-	0.001	-	0.019	-	**0.037**	Weibull
Kharagpur Jn	0.069	-	0.001	-	0.121	-	**0.408**	Weibull

Table 3.6 Howrah Mail Arrival distributions P-values

P-VAUE									
STATION NAME	Beta	Binomial	Negative binomial	Exponential	GEV	Log-normal	Logistic	Weibull	REMARKS\BEST FITTING CURVE
Dadar	0.000	**0.003**	< 0.0001	< 0.0001	0.000	-	0.000	-	Binomial
Thane	0.000	< 0.0001	< 0.0001	< 0.0001	< 0.0001	0.000	0.001	**0.001**	Weibull
Kasara	0.368	< 0.0001	**0.985**	< 0.0001	0.180	0.459	0.158	0.155	Negative binomial
Igatpuri	0.014	< 0.0001	< 0.0001	< 0.0001	< 0.0001	0.004	0.008	**0.017**	Weibull
Devlali	0.207	< 0.0001	< 0.0001	< 0.0001	0.570	0.555	**0.605**	0.460	Logistic
Nasik Road	0.274	< 0.0001	0.024	0.001	0.375	0.493	**0.654**	0.471	Logistic
Manmad Jn	0.145	< 0.0001	< 0.0001	0.005	< 0.0001	**0.966**	0.842	0.658	Log-normal
Chalisgaon Jn	0.135	< 0.0001	< 0.0001	0.010	< 0.0001	**0.980**	0.823	0.799	Log-normal
Jalgaon Jn	0.095	< 0.0001	0.017	0.003	0.516	0.373	**0.794**	0.591	Logistic
Bhusaval Jn	0.337	< 0.0001	0.240	0.004	0.901	0.839	**0.998**	0.894	Logistic
Malkapur	0.240	< 0.0001	0.000	0.005	0.634	0.434	**0.774**	0.693	Logistic
Nandura	0.216	< 0.0001	0.007	0.004	0.720	0.579	**0.934**	0.780	Logistic
Jalamb Jn	0.268	< 0.0001	0.026	0.001	0.656	0.569	**0.947**	0.731	Logistic
Shegaon	0.300	< 0.0001	0.000	0.001	0.958	0.916	**0.998**	0.925	Logistic
Akola Jn	0.305	< 0.0001	0.010	0.001	0.770	0.647	**0.954**	0.678	Logistic
Murtajapur	0.222	< 0.0001	0.000	0.000	**0.859**	0.692	0.659	0.474	GEV
Pulgaon Jn	0.436	< 0.0001	0.080	0.534	< 0.0001	0.725	0.700	**0.969**	Weibull
Wardha Jn	0.108	< 0.0001	**0.169**	0.000	< 0.0001	-	0.016	-	Negative binomial
Raj Nandgaon	0.113	< 0.0001	0.243	0.072	0.951	-	**0.973**	-	Logistic
Durg	0.022	< 0.0001	**0.082**	< 0.0001	< 0.0001	-	< 0.0001	-	Negative binomial
Bhatapara	0.019	< 0.0001	0.000	0.452	0.000	**0.720**	0.212	0.432	Log-normal
Bilaspur Jn	0.016	< 0.0001	< 0.0001	0.273	< 0.0001	-	< 0.0001	**0.502**	Weibull
Champa	0.001	0.016	< 0.0001	0.196	< 0.0001	-	< 0.0001	**0.391**	Weibull
Sakti	0.001	< 0.0001	< 0.0001	0.197	< 0.0001	**0.932**	< 0.0001	0.440	Log-normal
Kharsia	0.001	< 0.0001	< 0.0001	0.140	< 0.0001	**0.787**	< 0.0001	0.240	Log-normal
Raigarh	0.002	< 0.0001	< 0.0001	0.144	< 0.0001	**0.562**	< 0.0001	0.144	Log-normal
Brajraj-gar	0.053	< 0.0001	< 0.0001	0.004	< 0.0001	**0.344**	0.298	0.131	Log-normal
Jharsuguda Jn	0.001	-	< 0.0001	0.007	< 0.0001	-	< 0.0001	-	Poisson
Rourkela	0.142	< 0.0001	0.163	0.662	0.000	-	0.219	**0.666**	Weibull
Kharagpur Jn	0.106	< 0.0001	0.096	0.001	0.000	-	0.014	**0.162**	Weibull

Table 3.7 Howrah Mail Departure distributions P-values

P-VALUES								
STATION NAME	Beta	Exponential	Gamma	GEV	Logistic	Normal	Weibull	REMARKS\BEST FITTING CURVE
Kalyan Jn	**0.840**	< 0.0001	0.574	0.468	0.592	0.436	0.765	Beta
Bhusaval Jn	0.276	0.009	0.944	0.858	**0.981**	0.785	0.616	Logistic
Akola Jn	0.544	0.014	0.682	0.962	**0.986**	0.913	0.974	Logistic
Badnera Jn	0.256	-	-	0.482	**0.553**	0.523	0.345	Logistic
Raipur Jn	0.765	-	-	0.310	0.733	0.627	**0.848**	Weibull
Bilaspur Jn	0.959	0.962	0.952	0.286	0.800	0.609	**0.992**	Weibull
Rourkela	**0.975**	0.711	0.884	0.864	0.830	0.796	0.930	Beta
Chakradharpur	**0.993**	0.545	0.708	0.747	0.750	0.779	0.850	Beta
Tatanagar Jn	**0.937**	-	-	0.663	0.656	0.645	0.811	Beta
Kharagpur Jn	0.536	0.459	**0.862**	0.403	0.433	0.256	0.604	Gamma

Table 3.8 Jnaneswari express Arrival distributions P-values

P-VALUES							
STATION NAME	Beta	Exponential	Log-normal	Logistic	Normal	Weibull	REMARKS\BEST FITTING CURVE
Kalyan Jn	**0.959**	0.000	0.775	0.871	0.838	0.776	Beta
Bhusaval Jn	0.050	0.004	0.577	**0.727**	0.320	0.250	Logistic
Akola Jn	0.681	0.024	0.281	**0.986**	0.984	0.953	Logistic
Badnera Jn	0.379	0.021	-	0.343	0.169	**0.499**	Weibull
Durg	0.898	0.265	-	0.665	0.518	**0.931**	Weibull
Bilaspur Jn	0.564	0.293	**0.639**	0.324	0.228	0.545	Log-normal
Rourkela	0.904	0.571	0.872	0.970	0.974	**0.980**	Weibull
Chakradharpur	**0.993**	0.540	-	0.765	0.787	0.901	Beta
Tatanagar Jn	0.130	0.003	-	**0.226**	0.186	0.210	Logistic
Kharagpur Jn	**0.838**	0.581	-	0.597	0.387	0.820	Beta

Table 3.9 Jnaneswari express Departure distributions P-values

P-VALUES									
STATION NAME	Binomial	Gamma	GEV	Log-normal	Logistic	Normal	Poisson	Weibull	REMARKS\BEST FITTING CURVE
Kalyan Jn	0.000	0.732	0.740	0.695	0.700	0.652	0.404	**0.905**	Weibull
Manmad Jn	-	-	0.476	-	0.448	0.256	-	**0.630**	Weibull
Bhusaval Jn	< 0.0001	0.917	0.030	0.858	0.974	0.865	< 0.0001	**0.987**	Weibull
Akola Jn	< 0.0001	0.761	0.962	0.623	**0.989**	0.928	< 0.0001	0.892	Logistic
Gondia Jn (G)	< 0.0001	0.754	**0.943**	0.789	0.914	0.924	0.018	0.933	GEV
Raipur Jn	-	-	0.651	-	0.999	**1.000**	-	0.999	Normal
Bilaspur Jn	-	-	< 0.0001	-	0.971	0.926	-	**0.998**	Weibull
Raigarh	-	-	0.584	-	0.926	0.847	-	**0.990**	Weibull
Rourkela	< 0.0001	0.730	0.702	0.701	0.587	0.538	< 0.0001	**0.865**	Weibull
Tatanagar Jn	-	-	**0.980**	-	0.977	0.966	-	0.951	GEV
Kharagpur Jn	-	-	0.030	-	**0.960**	0.825	-	0.941	Logistic
Howrah Jn	-	-	0.851	-	**0.854**	0.818	-	0.772	Logistic
Barddhaman Jn	< 0.0001	0.916	0.612	0.872	0.940	0.887	< 0.0001	**0.948**	Weibull
Bolpur	< 0.0001	0.991	0.775	**0.993**	0.879	0.846	< 0.0001	0.941	Log-normal
Rampur Hut	-	-	0.818	-	0.963	0.929	-	**0.998**	Weibull
Pakur	0.219	0.414	0.325	0.382	0.810	0.588	**0.930**	0.501	Poisson
Malda Town	-	-	0.437	-	**0.997**	0.967	-	0.917	Logistic
Siliguri	< 0.0001	0.309	0.447	-	0.807	0.551	< 0.0001	**0.945**	Weibull
Hasimara	< 0.0001	**0.980**	0.030	0.962	0.949	0.970	< 0.0001	0.952	Gamma
Alipur Daur Jn	-	-	0.525	-	**0.997**	0.992	-	0.993	Logistic
New Bongaigaon	< 0.0001	0.949	0.896	0.889	**0.999**	0.999	0.036	0.992	Logistic

Table 3.10 Karmabhoomi express Arrival distributions P-values

P-VALUES								
STATION NAME	Beta	Exponential	Gamma	Log-normal	Logistic	Normal	Weibull	REMARKS\BEST FITTING CURVE
Manmad Jn	0.611	0.692	0.454	0.628	0.618	0.516	**0.743**	Weibull
Bhusaval Jn	0.642	0.437	0.785	0.713	**0.915**	0.720	0.887	Logistic
Akola Jn	0.827	0.321	0.770	0.632	**0.990**	0.934	0.901	Logistic
Gondia Jn	0.877	0.546	0.808	0.820	0.942	**0.961**	0.905	Normal
Bilaspur Jn	**0.928**	0.812	0.891	-	0.839	0.807	0.901	Beta
Raigarh	0.882	0.684	0.962	-	0.880	0.790	**0.994**	Weibull
Jharsuguda Jn	0.499	0.226	0.447	0.610	0.488	0.453	**0.804**	Weibull
Rourkela	0.766	**0.825**	0.777	0.769	0.615	0.564	0.762	Exponential
Howrah Jn	0.323	0.183	0.523	0.457	0.545	0.289	**0.853**	Weibull
Barddhaman Jn	0.956	0.906	0.954	**0.991**	0.963	0.910	0.942	Log-normal
Bolpur	0.942	0.899	0.978	**0.992**	0.860	0.831	0.873	Log-normal
Pakur	0.463	0.192	0.299	0.296	**0.725**	0.451	0.350	Logistic
Hasimara	0.896	0.196	**0.970**	0.942	0.943	0.961	0.790	Gamma
New Bongaigaon	0.988	0.643	0.949	0.889	**0.999**	0.999	0.992	Logistic

Table 3.11 Karmabhoomi express Departure distributions P-values

P-VALUES								
STATION NAME	Beta	Exponential	Gamma	Log-normal	Logistic	Normal	Weibull	REMARKS\BEST FITTING CURVE
Kalyan Jn	**0.997**	0.060	0.991	0.995	0.979	0.993	0.974	Beta
Igatpuri	0.752	0.355	0.372	0.426	0.630	0.584	**0.762**	Weibull
Nasik Road	0.869	0.024	0.926	**0.935**	0.887	0.921	0.877	Log-normal
Bhusaval Jn	0.980	0.501	0.908	0.790	**0.995**	0.991	0.994	Logistic
Akola Jn	**0.980**	0.142	0.808	0.863	0.909	0.879	0.864	Beta
Badnera Jn	0.870	0.372	0.758	0.804	0.874	0.898	**0.906**	Weibull
Nagpur	**0.865**	-	-	-	0.763	0.785	0.764	Beta
Gondia Jn	0.471	0.124	0.356	0.246	**0.860**	0.566	0.470	Logistic
Raj Nandgaon	0.687	0.609	0.968	0.932	**0.996**	0.873	0.955	Logistic
Durg	0.793	-	-	-	0.831	0.697	**0.919**	Weibull
Raipur Jn	0.773	0.573	**0.911**	0.453	0.697	0.630	0.648	Gamma
Bilaspur Jn	0.990	-	-	-	0.993	**0.993**	0.967	Normal
Rourkela	0.894	0.447	0.765	-	0.964	**0.985**	0.854	Normal
Chakradharpur	**0.506**	-	-	-	0.494	0.380	0.347	Beta
Bankura	**0.865**	0.581	0.693	-	0.708	0.702	0.775	Beta
Midnapore	0.532	-	-	-	0.535	0.410	**0.883**	Weibull
Kharagpur Jn	0.831	-	-	-	0.941	0.917	**0.979**	Weibull

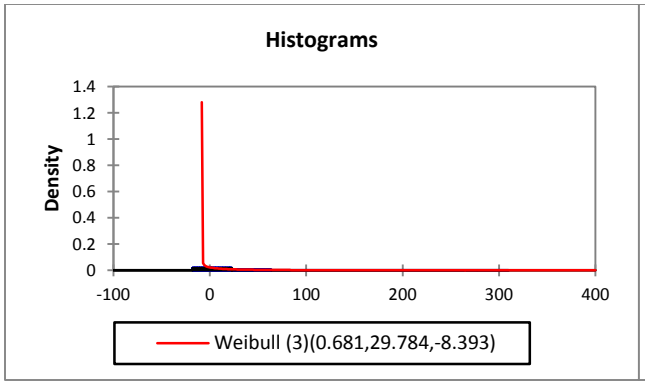
Table 3.12 Samarsata express Arrival distributions P-values

P-VALUES									
STATION NAME	Beta	Negative binomial	Exponential	Gamma	Log-normal	Logistic	Poisson	Weibull	REMARKS\BEST FITTING CURVE
Kalyan Jn	0.631	0.066	0.075	0.762	0.688	**0.996**	0.320	0.724	Logistic
Igatpuri	**0.768**	0.073	0.100	0.662	0.602	0.670	0.245	0.545	Beta
Nasik Road	0.869	0.041	0.022	0.926	**0.934**	0.887	0.256	0.877	Log-normal
Bhusaval Jn	0.947	< 0.0001	0.352	0.972	0.913	1.000	< 0.0001	**1.000**	Weibull
Akola Jn	**0.983**	< 0.0001	0.144	0.836	0.882	0.929	< 0.0001	0.898	Beta
Badnera Jn	0.488	0.394	0.203	0.529	0.685	0.584	**0.728**	0.427	Poisson
Gondia Jn	0.623	< 0.0001	0.132	0.395	0.265	**0.882**	< 0.0001	0.521	Logistic
Raj Nandgaon	0.588	< 0.0001	0.597	0.950	0.913	**0.971**	< 0.0001	0.921	Logistic
Durg	0.891	< 0.0001	0.985	0.985	-	0.962	< 0.0001	**0.996**	Weibull
Raipur Jn	0.767	< 0.0001	0.931	**0.931**	-	0.721	< 0.0001	0.723	Gamma
Bilaspur Jn	0.950	< 0.0001	0.730	0.962	-	**0.986**	< 0.0001	0.985	Logistic
Jharsuguda Jn	0.431	< 0.0001	0.580	0.656	**0.757**	0.422	< 0.0001	0.750	Log-normal
Chakradharpur	0.505	< 0.0001	0.222	0.215	-	**0.572**	< 0.0001	0.419	Logistic
Bankura	**0.865**	< 0.0001	0.581	0.693	-	0.708	< 0.0001	0.775	Beta

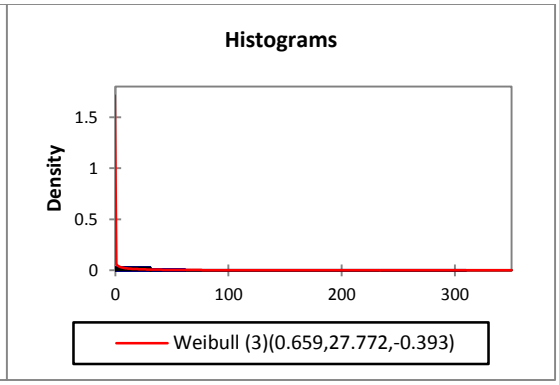
Table 3.13 Samarsata express Departure distributions P-values

GRAPHS

Bilaspur Jn (BSP)

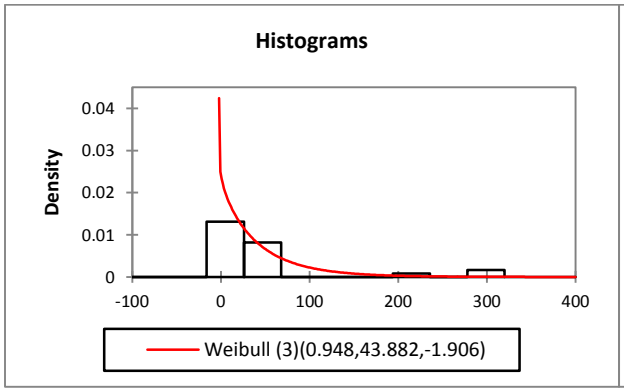


ARRIVAL

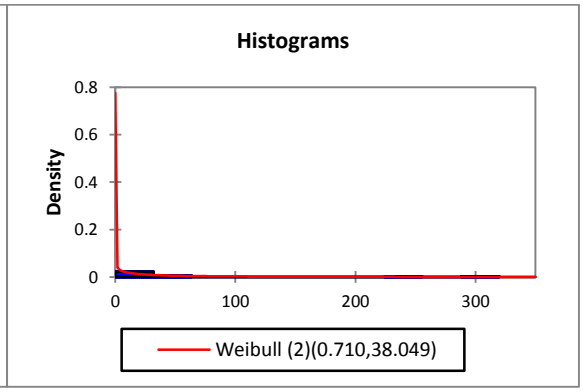


DEPARTURE

Champa

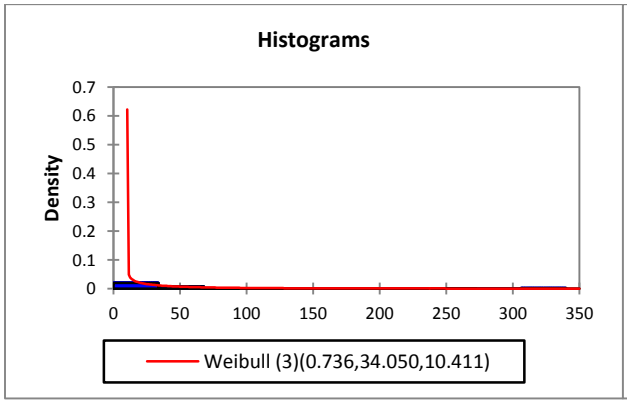


ARRIVAL

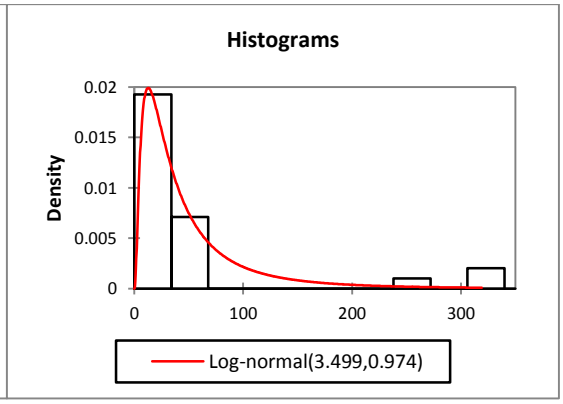


DEPARTURE

Raigarh

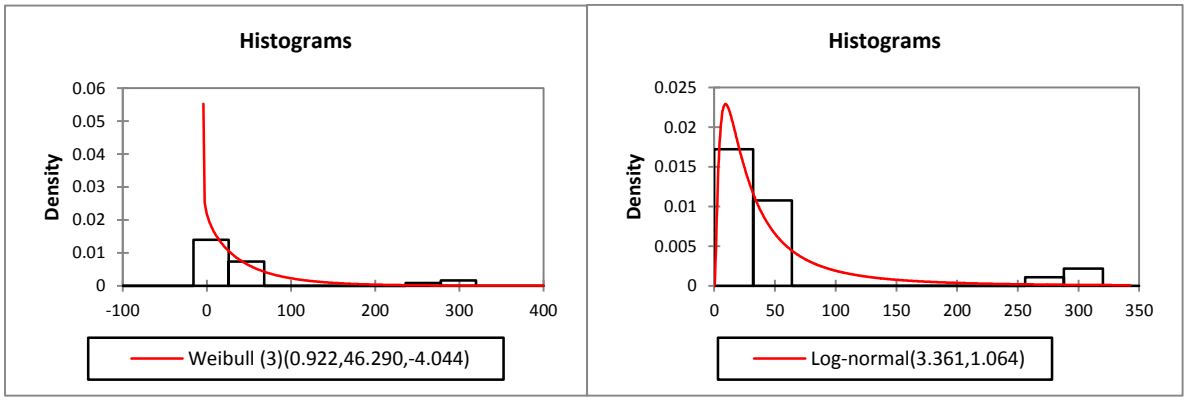


ARRIVAL



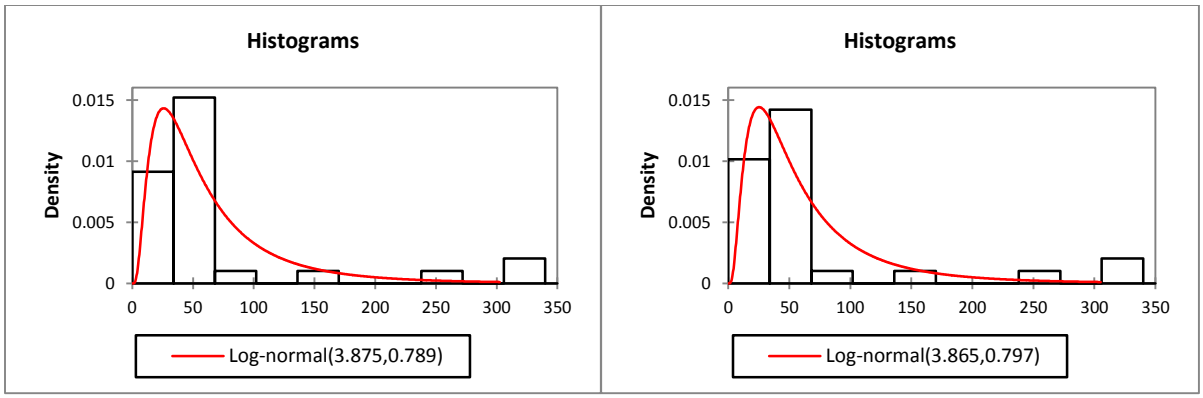
DEPARTURE

Jharsuguda



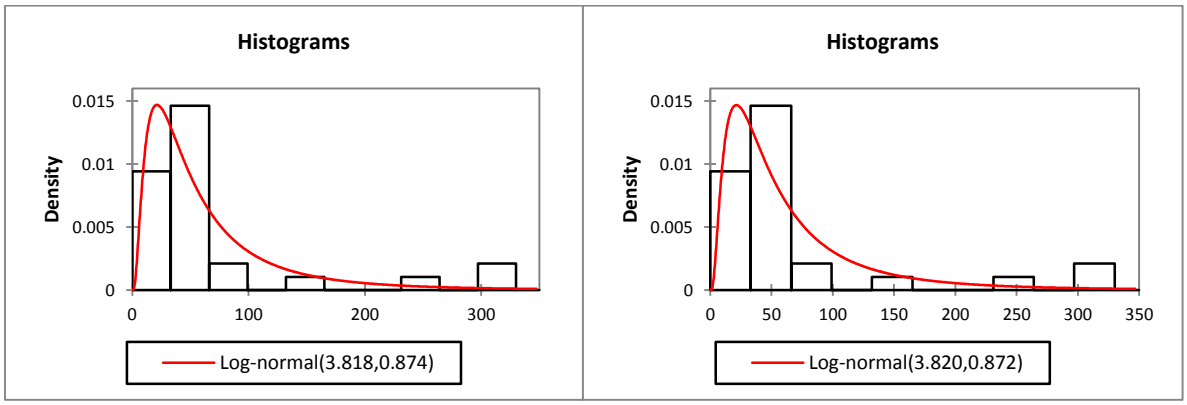
ARRIVAL DEPARTURE

Rourkela



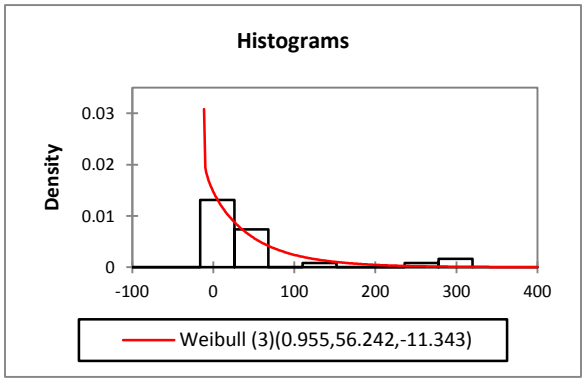
ARRIVAL DEPARTURE

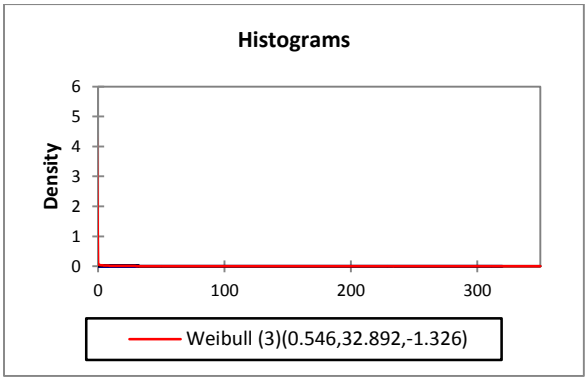
Chakradharpur



ARRIVAL DEPARTURE

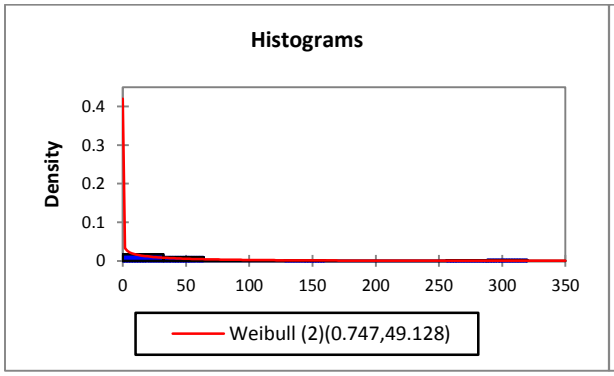
Tatanagar

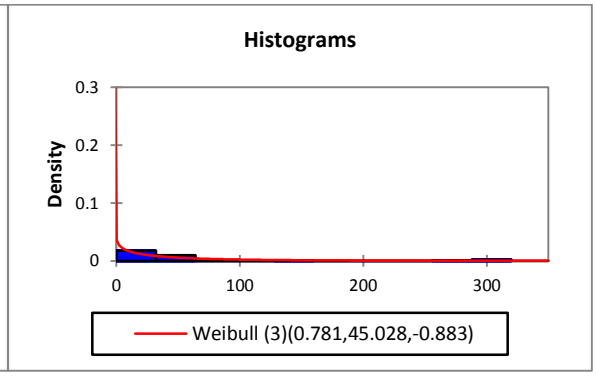




ARRIVAL DEPARTURE

Kharagpur





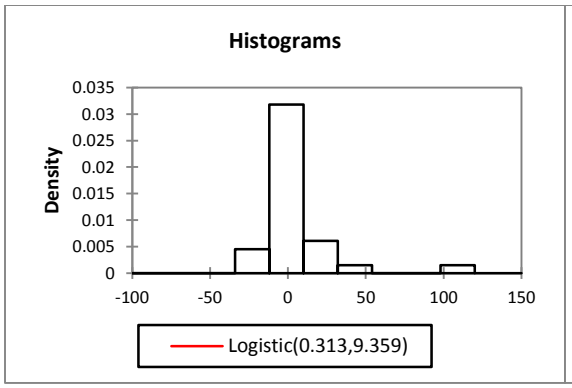
ARRIVAL DEPARTURE

Fig 3.5 Various distribution graphs of Azad Hind Express

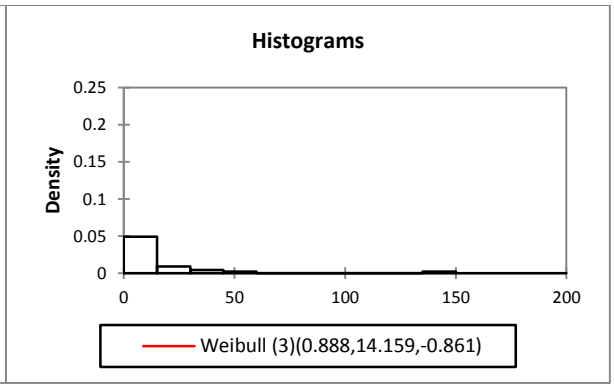
The arrival delay distribution of this train that fits best the data for the goodness of fit test is the Weibull distribution followed by log-normal distribution.

The departure delay distribution of this train that fits best the data for the goodness of fit test is the LOG-NORMAL distribution

BILASPUR

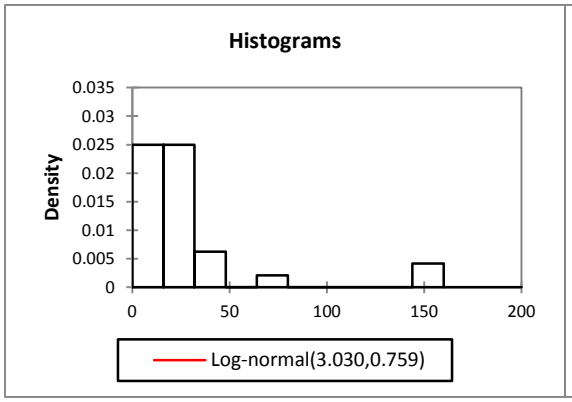


ARRIVAL

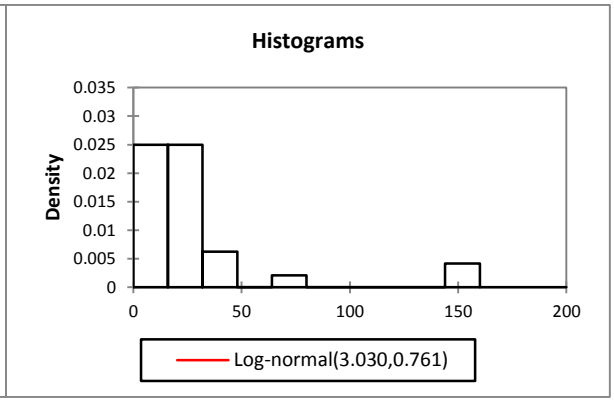


DEPARTURE

RAIGARH

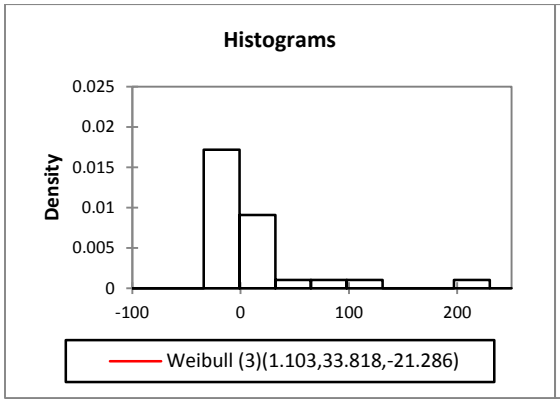


ARRIVAL

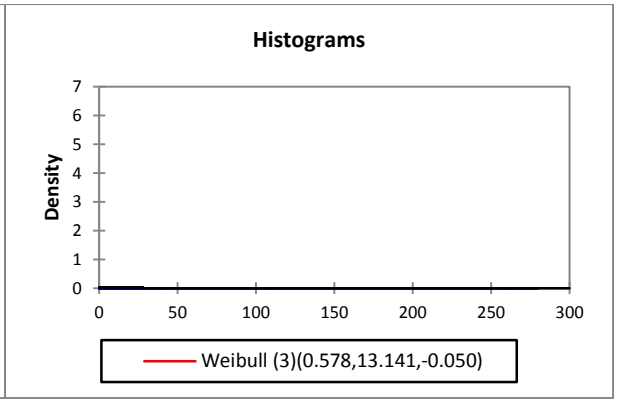


DEPARTURE

JHARSUGUDA

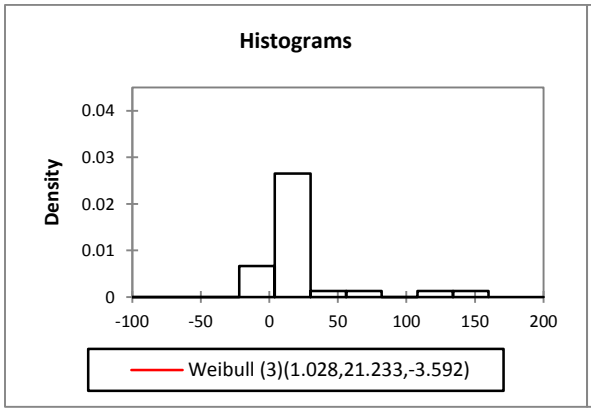


ARRIVAL

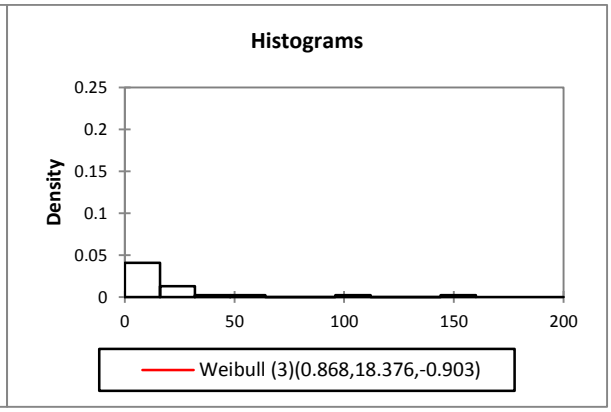


DEPARTURE

ROURKELA

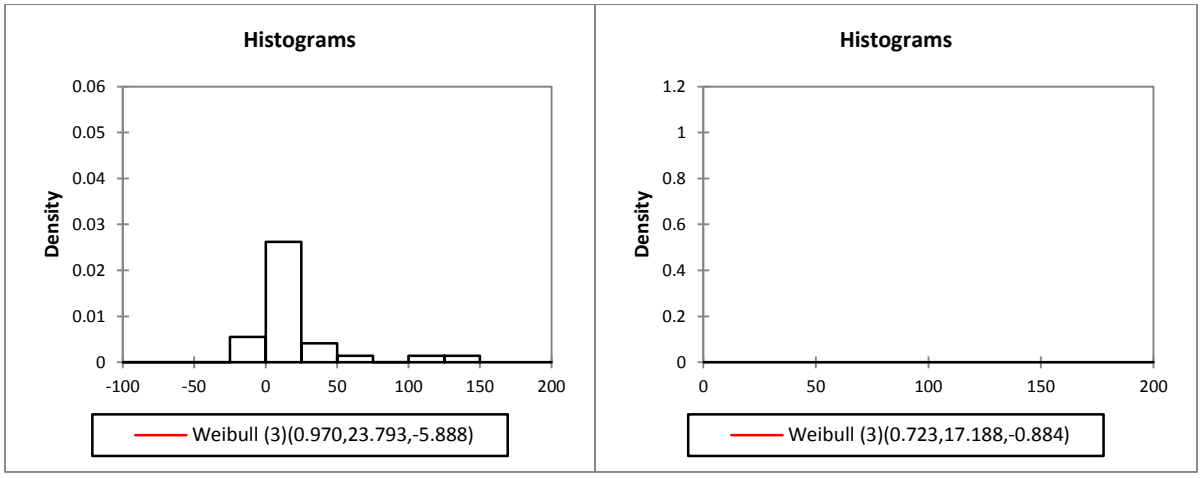


ARRIVAL



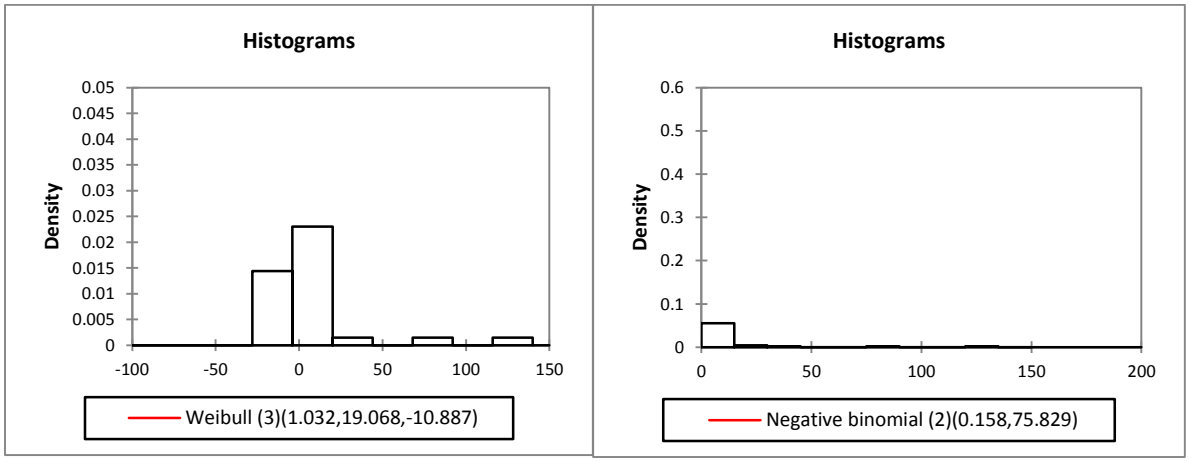
DEPARTURE

CHAKRADARPUR



ARRIVAL **DEPARTURE**

KHARAGPUR



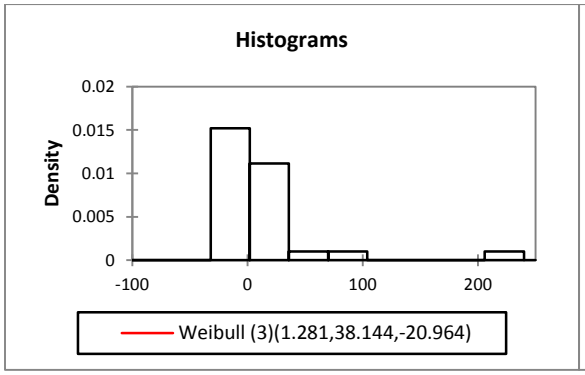
ARRIVAL **DEPARTURE**

Fig 3.6 Various distribution graphs of Gitanjali Express

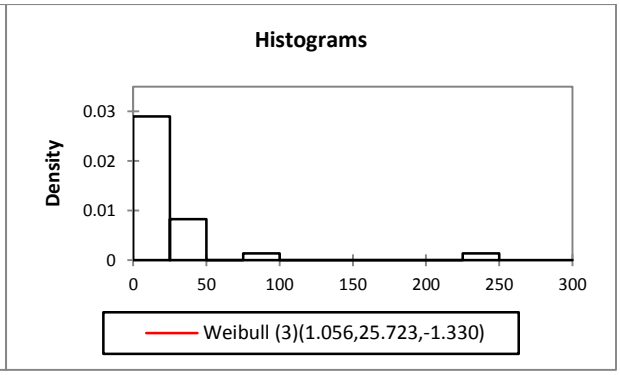
The arrival delay distribution of this train that fits best the data for the goodness of fit test is the Weibull distribution followed by Beta and logistic distribution.

The departure delay distribution of this train that fits best the data for the goodness of fit test is the Weibull distribution followed by log-normal and logistic distribution.

BILASPUR JN

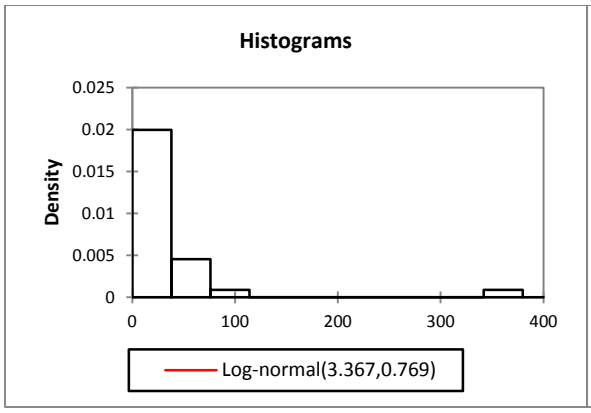


ARRIVAL

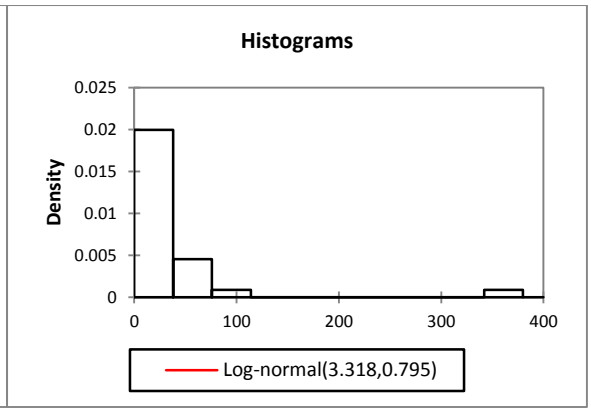


DEPARTURE

RAIGARH

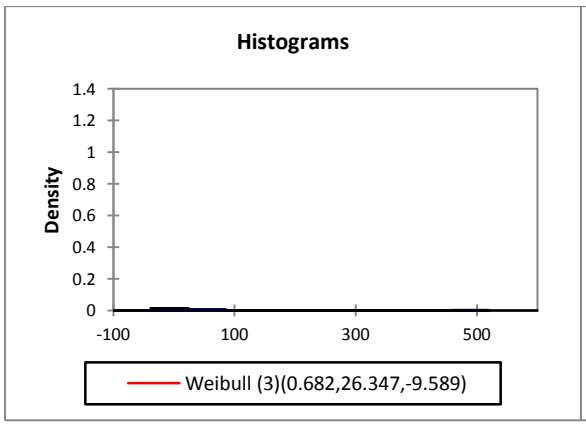


ARRIVAL

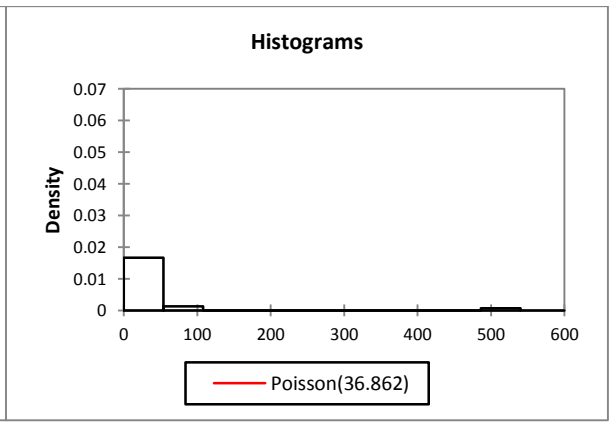


DEPARTURE

JHARSUGUDA

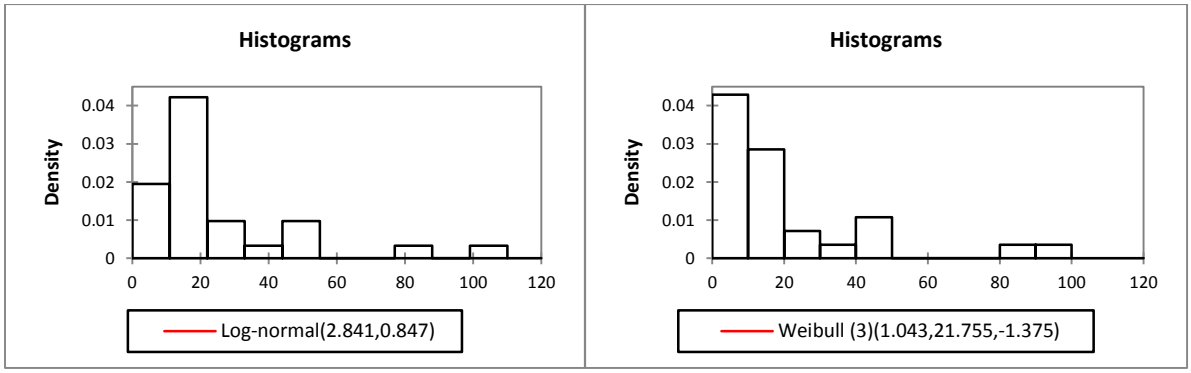


ARRIVAL



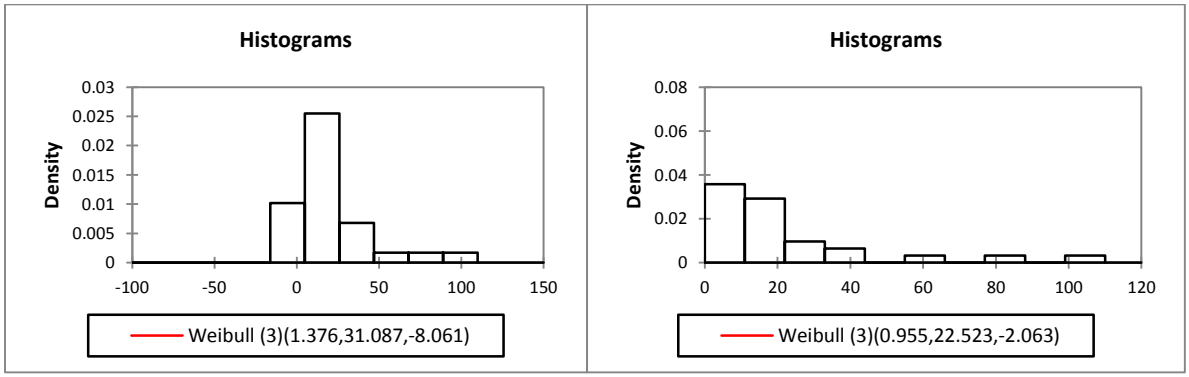
DEPARTURE

ROURKELA



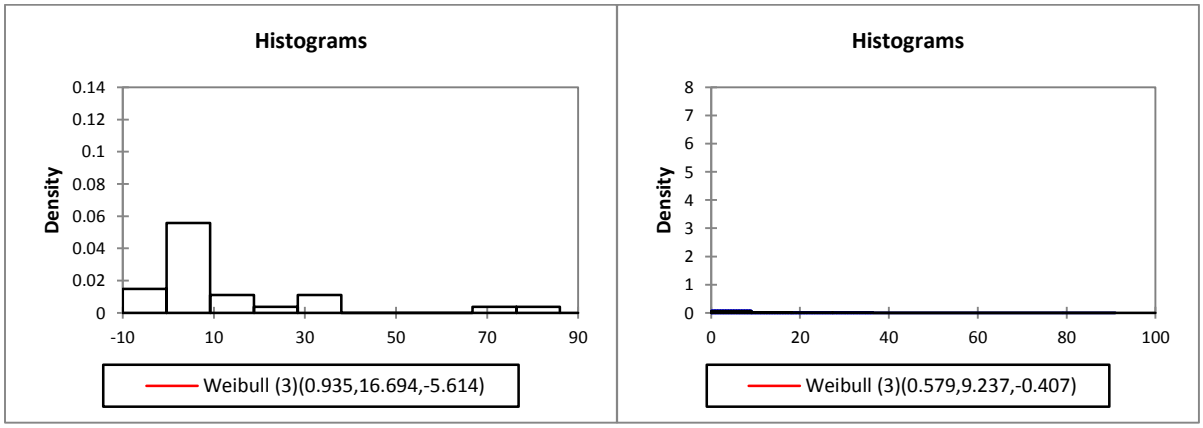
ARRIVAL **DEPARTURE**

CHAKRADHARPUR



ARRIVAL **DEPARTURE**

KHARAGPUR



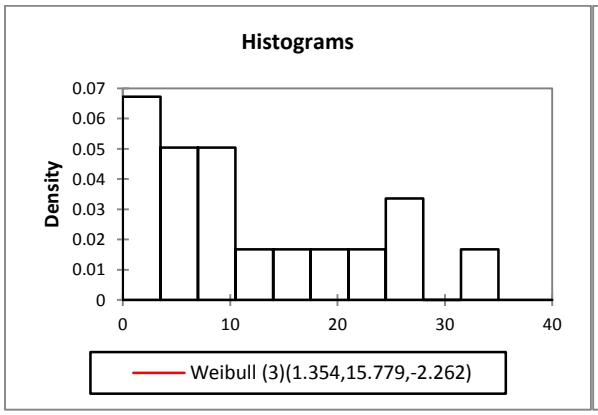
ARRIVAL **DEPARTURE**

Fig 3.7 Various distribution graphs of Howrah Mail

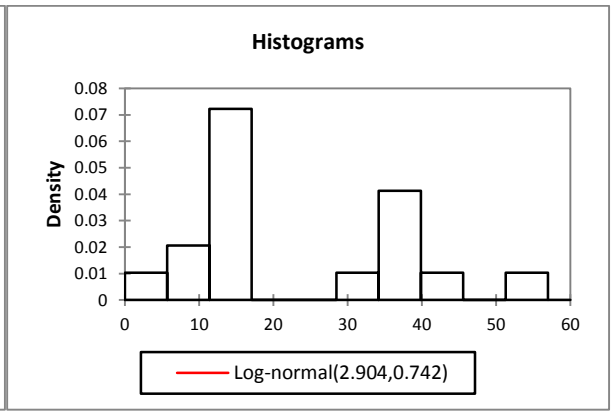
The arrival delay distribution of this train that fits best the data for the goodness of fit test is the logistic distribution followed by Weibull and log-normal distribution.

The departure delay distribution of this train that fits best the data for the goodness of fit test is the logistic distribution followed by log-normal distribution.

BILASPUR JN

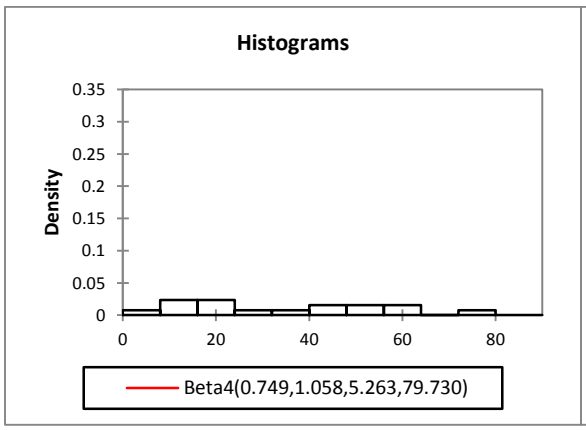


ARRIVAL

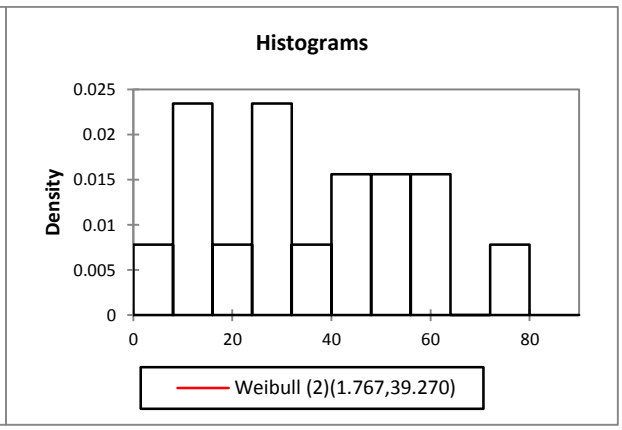


DEPARTURE

ROURKELA

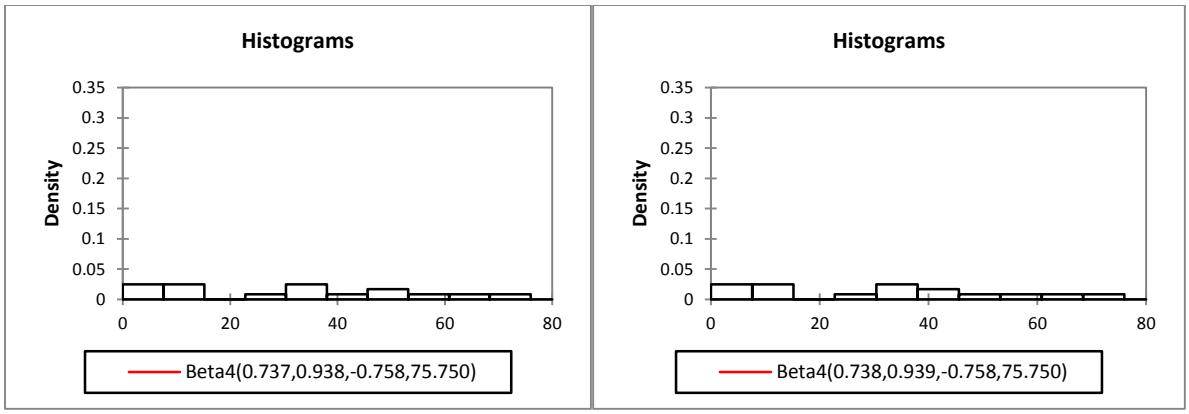


ARRIVAL



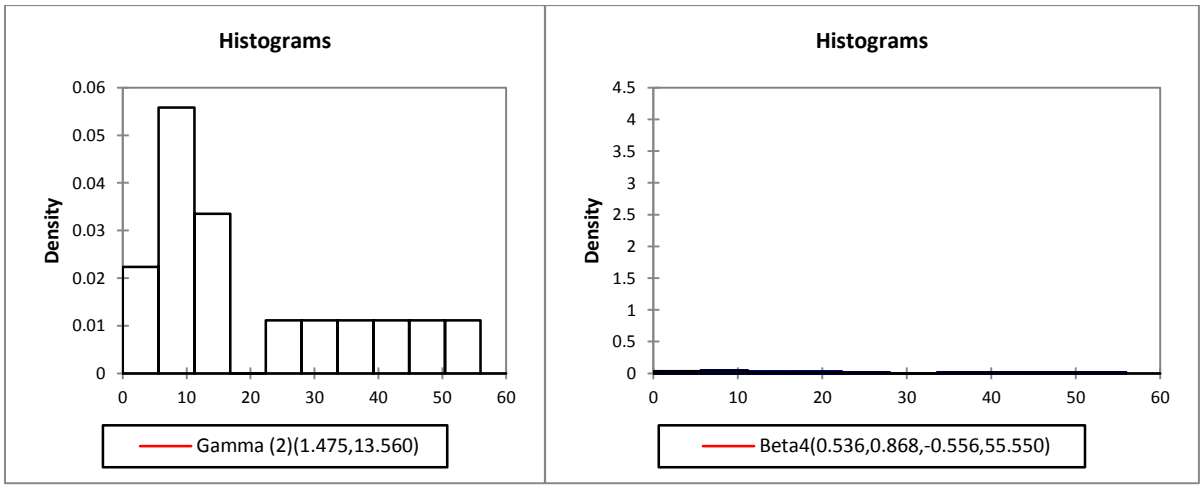
DEPARTURE

CHAKRADHARPUR



ARRIVAL **DEPARTURE**

KHARAGPUR JN



ARRIVAL **DEPARTURE**

Fig 3.8 Various distribution graphs of Jnaneswari Express

The arrival delay distribution of this train that fits best the data for the goodness of fit test is the Beta distribution followed by logistic distribution.

The departure delay distribution of this train that fits best the data for the goodness of fit test is the logistic distribution followed by Weibull distribution.

3.7 DISCUSSIONS

The arrival delay probability functions of various train lines are follows bell-shaped and skewed to the right (having a longer on right tail), see in above figures. Even though some of trains arrive early and other trains arrive more or less late, most arrival times falls around the mean values. Some of the trains depart early from preceding station without delay but early arrivals in some other stations are constrained due to low speed or unscheduled stops.

Some of the figures shown above having small skewness which indicates small deviation of delay distribution is symmetric and hence be approximated by a normal distribution.

The arrival and departure delays of trains having a skewness coefficient around as shown below.

Train Name	Skewness coefficient	
	Arrival	Departure
Azad Hind Express	2.188	2.199
Gitanjali Express	2.733	2.733
Howrah Mail	1.960	1.884
Jnaneswari Express	0.372	0.329
Karmabhoomi Express	0.479	0.469
Samarsata Express	-0.562	-0.467

Table 3.14 Skewness coefficient of different trains at Rourkela station

By which skewed distribution may be more appropriate like a Weibull, Erlang, or lognormal distribution. Also the hypothesis is tested that the arrival delays have a normal distribution with

unknown parameters. There is a possibility of slightly delayed trains lose their scheduled route setting to conflicting train and hence have to stop at the open track, by which delay is further increased.

By considering distribution of arrival lateness and early arrivals analysis of delay propagation is prime interest and help to detect the distribution shape from statistics only. The histogram and empirical density of late arrivals suggests a good fit by a negative-exponential distribution. This hypothesis is formally tested using the KS goodness-of-fit test of single exponentially for the estimated parameter (1/mean), with the mean. Hence, the late arrival delays of these train lines constitute most trains instead of just the right tail.

Train departure delays are expected to follow an exponential distribution, i.e. a train departs after the scheduled departure time with random delay. The departure delay histograms also suggest a good exponential fit of single exponentially of hypothesis test with a significance level of 5% for estimated parameter of the trains. Finally it can conclude that an exponential distribution is in a general a good stochastic model for train departure delays, as long as the departure is not restricted.

CHAPTER 4

CONCLUSIONS

The railway system is a significant change in India over midst. Rail Road's which are laid by government and others are more importantly, by customers, to provide better service and to be more efficient. For any success, the main key in this world is very carefully planning, scheduling, implementing planned and new existing works without any disruption to other works. Even if we planned carefully, there would be some time deviations.

Some of the studies on delay timings by using statistical distribution analysis give clarification about time delay distribution pattern and helps to rectify old timetable. The train delays should be measured and analyzed in minutes. The arrival and departure timing delay of trains at the stop positions are calculated. The statistical analysis of arrival times at various stations on the selected route, all most all points shows that about 30% to nearly 60% of trains delayed by more than 10 minutes.

The proven fit of most of the arrival delay times to Weibull and logistic distribution, departure delays to a log-normal distribution. This allows an improvement in the design of timetable and more reliable forecast of train delays during operations. In this research mainly focus on suitable models of train delay propagation in stations.

REFERENCESS

1. Goverde et al. (2000) TNV-prepare: Analysis of Dutch Railway operations based on Train detection data.
2. Goverde et al. (2001) Statistical Analysis of Train Traffic: The Eindhoven Case, TRAIL Studies in Transportation Science, No. S2001/1, Delft University Press, Delft.
3. Goverde et.al (2001) Delay Distributions in Railway Stations.
4. Huisman and Boucherie, Running times on railway sections with heterogeneous, train traffic. Transportation Research Part B 35 (2001), 271-292.
5. Nash and Ullius, Optimizing Railway Timetables with Open Timetable Paper for comprail 2004, ETH Zurich, 2004.
6. Ullius (2005) Delay Analysis of Rail 2000 1st Phase Using Open Timetable SBB AG I-TM-PS, Bern 2005.
7. Goverde (2005) Punctuality of Railway Operations and Timetable Stability Analysis, Ph.D. thesis, Delft University of Technology.
8. Mark Harris (2006) Analysis and modelling of train delay data.
9. Yuan (2006) capturing stochastic variations of train event times and process times with goodness-of-fit tests.
10. Yuan (2006) Stochastic Modelling of Train Delays and Delay Propagation in Stations. Ph.D Thesis, Delft University of Technology, the Netherlands.
11. Advance Engineering Mathematics by Erwin Kreyszig and Common probability distributions by D.Joyce, Clark University (2006).
12. Yuan and Hansen (2007) Optimizing capacity utilization of stations by estimating knock-on train delays. Transportation Research Part B 41, 202-217.

13. Murali et al. (2009) A Delay Estimation Technique for Single and Double-track Railroads. 3715 McClintock Ave., GER 240, Los Angeles, CA 90089-0193, USA
14. Gorman (2009), Statistical Estimation of Railroad Congestion Delay, Transportation Research Part E, Vol. 45, No. 3, pp. 446-456.
15. Dingler et al. (2010) Determining the Causes of Train Delay 205 N. Mathews Ave., Urbana IL, 61801 USA.
16. Ghosh et al. (2013) Run-time Delays in Indian Railways: is Traffic the Cause.