IJIRCCE

INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

ISSN INTERNATIONAL STANDARD SERIAL NUMBER INDIA

Impact Factor: 7.488

A Novel Study and Analysis on Covid-19 Cases in India

Ayushi Jain[1], Siddharth Nanda[2]

U.G Student, School of Engineering, Ajeenkya DY Patil University, Pune, Maharashtra, India [1]

Faculty, School of Engineering, Ajeenkya DY Patil University, Pune, Maharashtra, India [2]

ABSTRACT: In this paper, we've a dataset of national level covid-19 cases in India from 30th January to 2th august 2020. It has various parameters like date, total confirmed,total deceased, total recovered, daily confirmed, daily deceased, daily recovered cases. I'll be analysing this dataset using systematic sampling method to find out the recovery rate and death rate throughout the time period so that we can know the pattern and can take precautions accordingly.

KEYWORDS: covid-19,systematic sampling, recovery rate, death rate.

I. INTRODUCTION

Corona virus is a disease that spreads through touch or droplets of saliva etc. some of its symptoms are cough, cold, high fever, loss of taste etc. this disease originated from China and is now spreading rapidly throughout the world. To prevent the spread of this disease, the government has thaken various measures like a nation-wide lockdown. We can take precautions like going out only when necessary, wearing masks and maintaining social distancing at all times.

In this paper, I've used systematic sampling. Systematic sampling is a probability sampling method where we choose elements from a target population by selecting a random starting point and selects ample members after a fixed 'sampling interval. I used the following steps for systematic sampling:

Step one: we import our datasetStep two: figure out the ideal size of the sample Step three: by using iseven formula in excel and taking even values, we filtered out the true values and considered them from further analysis.Step four: we filtered out the columns not needed for our analysisStep five: we select each column and using data analysis in excel we find out the required values.

I've selected a dataset of 126 days out of which I've taken a sample of 95 days, in my dataset I've filtered out the needed columns that are total confirmed,total deceased,total recovered cases for further analysis. Using systematic sampling, ill be calculating various parameters of our dataset.

II. LITERATURE SURVEY

1. **The Outbreak of Novel Coronavirus in India: Psychological Impact** by Bibin John Thurackal,Eleanora NalluChith and PriscillaMascarenhas on March 27, 2020

Description: The paper focuses on the impact corona virus in India, the remedial measures to stop the spread of COVID19 and the possible psychological consequences. It also tells us about the state by state increase in the number of cases.

2. **Coronavirus: Hotspot on coronavirus disease 2019 in India by** Rameshwar S. Cheke, Sachin Shinde,Jaya Ambhore, Vaibhav Adhao, and Dnyaneshwar Chekeon2020 Apr 30

Description: in the following report, they have very well summarized the records and origin, microbiology and taxonomy, mode of transmissions, goal receptor, medical features, diagnosis, prevention, and remedies for COVID-19.

3. **Outbreak Trends of Coronavirus Disease–2019 in India**: A PredictionbySunita Tiwari,Sushil KumarandKalpnaGuleria on 22 April 2020

Description: This paper's goal is to put together the authorities and residents of India to take measures proactively to lessen the effect of coronavirus ailment 2019 (COVID-19). They've taken a dataset of confirmedinstancesthrough date, recovered instances, and demiseinstances of China are filtered from the dataset. Additionally, the to be hadfacts for India also are filtered for validation. It predicts that Indiamight also additionally have almostone millionshowedinstancesthrough the end of May 2020. This can bemanaged if the climatic situations and Government of India guidelinesend upfavorableto manipulate the virus.

4. **India under COVID-19 lockdown**, by The Lancet on 2020 Apr 23
Description: This paper evaluates India's circumstancethroughout lockdown and concludes that Theinstanttask is to keep infections at potentialstages and make certain the capacity to test, hint contacts, isolate patients, put into effect COVID care plans, and disseminate timely information. The primaryauthoritiesmust loosen itsmanage and deliver states greater autonomy over their investment and selection making.

5. **Prudent public health intervention strategies to control the coronavirus disease 2019 transmission in India: A mathematical model-based approach**, By Sandip Mandal,Tarun Bhatnagar, NimalanArinaminpathy, Anup Agarwal, Amartya Chowdhury, Manoj Murhekar, Raman R. Gangakhedkar, and Swarup Sarkaron 2020 Feb-Mar**Description**: This paper uses hypothetical epidemic curves, to inform us if itsfeasible to prevent, or delay, the local outbreaks of COVID-19 viaregulations on tour from overseas and if the virus has already installed in-country transmission, to what volumecould its effect be mitigated via quarantine of symptomatic patients?

6. **Novel Coronavirus (COVID-19) in India: Current Scenario**, by Varsha Kachroo on 3 march 2020
Description: This evaluation in short covers the introduction, viable mode of transmission, definitions, a few basic advices, diagnosis, remedy and control protocol being observed proper now in India; but subjected to alternate in due path of time as is the variety of instances and mortality.

7. **Epidemic and Challenges of Coronavirus Disease-2019 (COVID-19): India Response** by Prem Chaurasiya, Pragati Pandey, Upendra Rajak, KrishnakantDhakar, Manoj Verma, Tikendranath Verma on 7 Apr 2020
Description: This examinationindicates the up to datequantity of showedinstancestogether with death rates globally (as of 26 March 2020). It additionallyprovides the present daysituation and status in India. Additionally, it discusses the demanding situations and containment measures followedwith the aid of using India to limition the unfold of COVID-19 virus, and additionallyprovides the treatmentmovement to be followedin opposition toeffect of coronavirus on intellectual health.

8. **Coronavirus disease (COVID-19) pandemic: Furnishing experiences from India**, by Iqbal, Naved Dar, Kaiser Ahmadin Aug 2020
Description:this paper tends to the issues and consequences of the quick spread of the Covid illness (COVID-19) in India. Since the time COVID-19 was announced a pandemic, prompting a cross country lockdown, a lion's share of individuals in India have confronted a heap of difficulties, with emotional well-being emergency being the most unmistakable one.

9. **An overview of mobile applications (apps) to support the coronavirus disease 2019 response in India** by, Abhinav Bassi**,** SumaiyaArfin**,** Oommen Johnand Vivekanand Jhaon 20-Jun-2020
Description: this paper methodicallly surveys COVID-19 related portable applications and feature holes to illuminate the improvement regarding future mHealth activities to reason that 346 potential COVID-19 applications, of which 50 met the incorporation measures. Scattering of untargeted COVID-19-related data on safeguard procedures and observing the developments of isolated people was the capacity of 27 (54%) and 19 (32%) applications, individually. Eight (16%) applications had a contact following and area of interest recognizable proof capacity.

10. **A Review of Coronavirus Disease-2019 (COVID-19)**, By Tanu Singhal on 13 March 2020. **Description:** this paper investigations the monetary, clinical and general wellbeing framework of Covid in china and what its meaning for india and what safety measures we can take to restrict its spread.

III. PROPOSED ANALYSIS AND APPROACH

Question: In India, a survey of corona patients was conducted. A sample size of 96 students was taken out of 127 students for analysing the recovery date, and death rate. Using sample data, estimate the mean of the population. Find the margin of error and the confidence interval. Assume 95% confidence interval.
Using data analytics tool in excel, we found the following values:

Total Confirmed		Total Recovered		Total Deceased	
Mean	303460.9	Mean	183350.7	Mean	7744.779
Standard Er	49985.47	Standard Er	32551.24	Standard Er	1156.749
Median	42778	Median	11763	Median	1463
Mode	3	Mode	3	Mode	0
Standard De	487198.1	Standard De	317270.2	Standard De	11274.59
Sample Vari	2.37E+11	Sample Vari	1.01E+11	Sample Vari	1.27E+08
Kurtosis	2.700684	Kurtosis	3.311202	Kurtosis	0.997316
Skewness	1.871033	Skewness	2.01121	Skewness	1.466388
Range	1963252	Range	1327134	Range	40754
Minimum	1	Minimum	0	Minimum	0
Maximum	1963253	Maximum	1327134	Maximum	40754
Sum	28828782	Sum	17418320	Sum	735754
Count	95	Count	95	Count	95
Confidence	99247.33	Confidence	64631.25	Confidence	2296.751

We'll be using these values in our further calculations.

1. Population Parameter

Population parameter is a number thatdescribes something about a group or population.
Population mean= 489393.193

2. Population Variance
averageof the distances from each data point in a particular population to the meansquared, and it indicate show data points are spread out in the population.
Total confirmed : (303460.9-489393.193)sq./ 95-1= 367774655.108
Total recovered: (183350.7-489393.193)sq./95-1= 996404335.336
Total deceased : (7744.779-489393.193)sq./95-1= 2467927603.284

3. Standard Error
SE= SD/sq. root of N
SE tc= 487198.1/root 95= 49989.544
SE tr= 317270.2/root 95= 32553.888
SE td= 11274.59/root 95= 1156.842

4. Confidence Level
There are two tests, t-test and z-test. T-test is used when we have a small sample size (<50) and when population variance is unknown whereas z-test is used when we have large sample size (>50) and population variance is known.As the sample size is large, thus we will use z-test. Standard Normal Distribution Table is used to find critical z-score.
In this part of the analysis, usually researchers choose a confidence level and the most frequently chosen confidence level is 95%. Thus, we will use that only.

5. Critical Value
The critical value is a factor used to compute the margin of error. To find the critical value, we take these steps:

- Alpha(α):
  α = 1-(confidence level/100)
  α = 1-(95/100)
  = 0.05
- Critical Probability (p*) :

$p^* = 1 - (\alpha/2)$
$= 1 - (0.05/2)$
$= 0.975$

Using Standard Normal Distribution Table, we can see that the critical z-score value is 1.96.

6. Margin of Error

ME = (Critical Value * Standard Error)
ME totalconfirmed = 1.96 * 49989.544 = 97979.506
ME totalrecovery = 1.96 * 32553.888 = 63805.620
ME total deceased = 1.96 * 1156.842 = 2267.410

7. Confidence Interval

The minimum and the maximum values of the confidence interval are:
CI_{min} = x – Standard Error * Critical Value
CI_{max} = x + Standard Error * Critical Value

CI_{min}total confirmed = 205000
CI_{max}total confirmed = 401000
CI_{min}total recovery = 120000
CI_{max}total recovery = 247000
CI_{min}total deceased = 5470
CI_{max}total deceased = 10000

Summary:
Based on our sample data, we can summarize that theres a trend in the data, as time increases, the recovery and death rate also increase rapidly. When we look at the graphs below, we can see a substantial increase in the cases from jan to august. The reasons behind the rapid increase of cases are not following the guideliness given by the government.

comparison between total confirmed, recovered and deceased patients from jan to aug
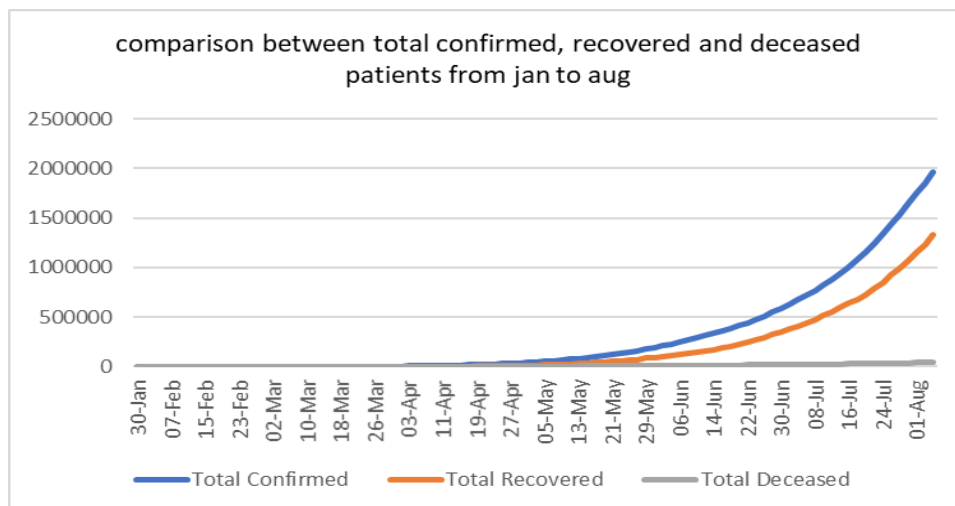
As we can see in this image, the number of cases have risen rapidly from 30th January to 1st august, the y-axis shows the number of cases and the x-axis shows the date.

Date wise- total recovered cases
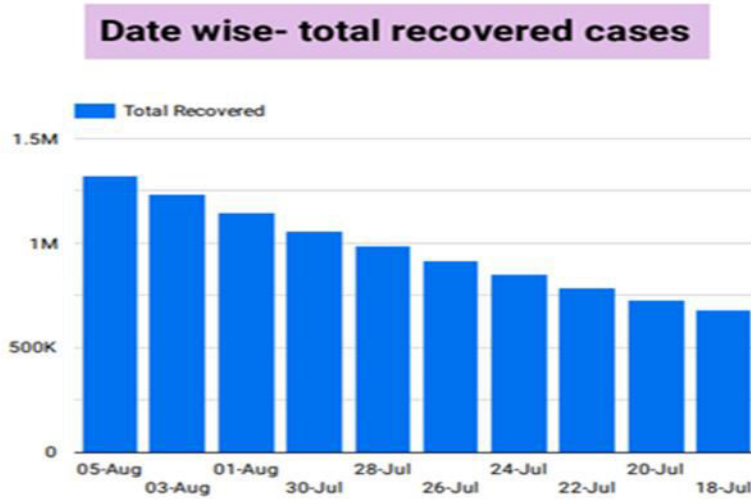
As we can see the no. of recovered cases have increased with time as the no. of cases have also increased.

Date wise- total confirmed cases
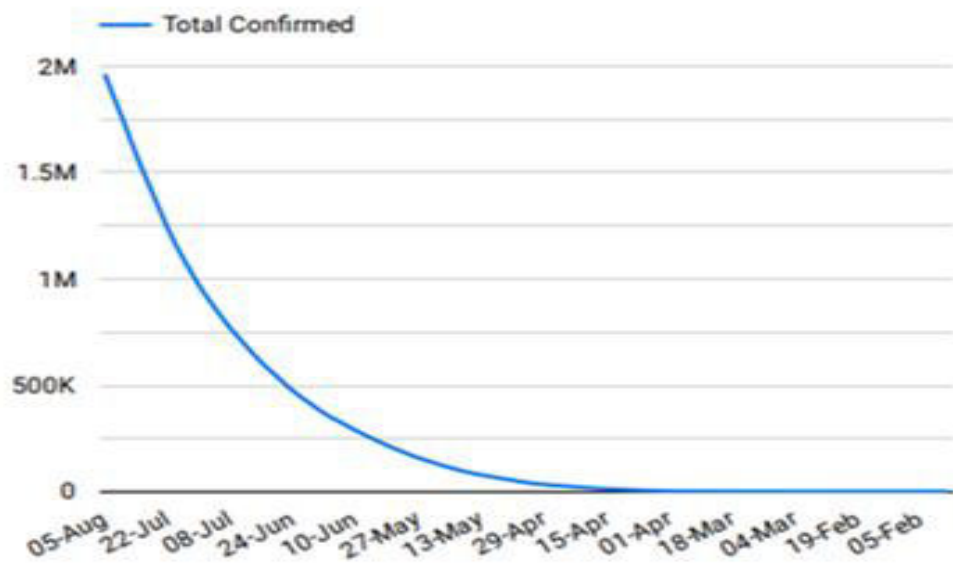
We have taken date on the x-axis and total confirmed cases on the y-axis. The no. of cases increase with time.

IV. FUTURE SCOPE

as we know, the spread of corona is increasing day by day. Prime minister Narendra Modi has taken precautions like nationwide lockdown, made wearing of masks mandatory etc. we cannot predict the exact outcome on how it will turn out but we have to follow the guidelines and be careful. It is said that it might take years for covid to be totally eradicated. It can be used in research related fields for further analysis.

V. CONCLUSION

In this paper, I've analysed a covid-19 dataset of 126 values taken from January 2020 to aug 2020 and by using systematic sampling I've filtered out the even values of the dataset and taken a sample of 95 values to calculate various parameters like mean, standard deviation, margin of error, confidence level, population variance, critical value etc. On analysing our dataset we can come to a conclusion that the recovery rate has increased rapidly with time and so did the number of deaths. We can take various precautions like wearing masks, going out only when necessary and maintaining social distancing to limit the spread of the disease.

REFERENCES

1. Guleria.k(**2020**)Outbreak Trends of Coronavirus Disease–2019 in India: A Prediction
2. Thurackal.B,Chith.E and Mascarenhas.P (2020) The Outbreak of Novel Coronavirus in India: Psychological Impact
3. The Lancet(2020)India under COVID-19 lockdown
4. Singhal.T (2020) A Review of Coronavirus Disease-2019 (COVID-19)
5. Mandal.S,Bhatnagar.T, Arinaminpathy.N,Agarwal.A, Chowdhury.A, Murhekar.M, Gangakhedkar.R, and Sarkar.S (2020) Prudent public health intervention strategies to control the coronavirus disease 2019 transmission in India: A mathematical model-based approach.

INTERNATIONAL JOURNAL
OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING