



**IJIRCCCE**

e-ISSN: 2320-9801 | p-ISSN: 2320-9798



# INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

Volume 9, Issue 5, May 2021

**ISSN** INTERNATIONAL  
STANDARD  
SERIAL  
NUMBER  
INDIA

**Impact Factor: 7.488**

 9940 572 462

 6381 907 438

 [ijirccce@gmail.com](mailto:ijirccce@gmail.com)

 [www.ijirccce.com](http://www.ijirccce.com)

# Statistical Analysis of Unemployment in India

Shruti Shrivastava<sup>1</sup>, Siddharth Nanda<sup>2</sup>

U.G Student, School of Engineering, Ajeenkya DY Patil University, Pune, Maharashtra, India<sup>1</sup>

Faculty, School of Engineering, Ajeenkya DY Patil University, Pune, Maharashtra, India<sup>2</sup>

**ABSTRACT:** This paper focuses on a statistical analysis of the unemployment rates in India, which have been severely impacted during Covid-19 crisis. Since the unemployment rate greatly impacts our GDP its imperative to analyse their trends. We divide these unemployment rates into two subsets of urban and rural population to determine which region is more severely hit during the pandemic. We explore the use of Stratified Sampling in this paper by using the dataset and Microsoft Excel as our tool. It is shown that how Covid has spiked the rates and how investment and employment generation is to be implemented.

**KEYWORDS:** Stratified Sampling, Trends, Analysis, Unemployment, Rates, Covid-19.

## I. INTRODUCTION

Unemployment is a hinderance to the economic growth of any nation and is a grave issue to focus on. It is the cause for poverty, poor health, malnutrition and many other causes that hinder the development of our country. Covid-19 has aggressively triggered the unemployment rates in India and has resulted in an economic pitfall. In this paper we analyse urban and rural regions to see which region has been worst hit by the pandemic. For our study, I have taken a dataset with 721 entries that have state-wise collected during the time period of 2019-2020 from both urban and rural areas. The dataset has attributes namely, State, Region, Date, Frequency, Estimated Unemployment Rate (%), Estimated Employed and Estimated Labour Participation Rate (%).

We draw a sample from this dataset to estimate the population parameters and use Stratified Sampling method to analyse the Unemployment rates in urban and rural areas separately, so that we can get equal representation of the data. We have 360 rural regions and 361 urban regions in the given dataset. We estimate the population's mean and variance and then we will calculate the standard error, critical value, and margin of error to compute the confidence interval. With the sample mean and confidence interval, we will get an idea of India's unemployment rate during the pandemic. If the mean is close to 4-5% it is considered to be an ideal unemployment rate by many economists.

## II. LITERATURE SURVEY

- 1) Poverty and unemployment in India: An analysis of recent evidence.

This paper inspects the relationship amidst poverty and unemployment using the data from Indian National Sample Survey dating October 1972-September 1973, for the states Gujarat and Maharashtra. The author was an economist in World Bank during the publication, has involved the concepts of statistics accompanied by his study of economics. In this paper it is observed that to the contrary of popular opinion that poor are too poor to remain unemployed, it is found that the association between unemployment and poverty (as measured by the per capita expenditure of households), is undeniable and significant. However, poverty in these states is more widespread than unemployment.

- 2) Toward and explanation of interregional variations in poverty and unemployment in rural India.

This study inspects the interregional variations in poverty and unemployment rates in 1972-73 in India, across 56 agroclimatic regions. It focused on the analysis of the interconnections and determinants that factored towards the changing rates. It was found that the rate of unemployment was in proportion to the wage dependent households in the rural area. The interregional variations in average domestic consumption expenditures were explained by the average value of household assets. These factors were influenced by the population density and the inequality in distribution of

assets which in turn contributed significantly to the explanation of interregional differences in case of poverty and unemployment. The statistical regression model is employed to obtain and describe these results.

- 3) Trends in Rural Unemployment in India: An Analysis with Reference to Conceptual and Measurement Problems.

This paper is an integrative study intended to understand and resolve the problem of rural unemployment in India, published in the year 1976. It is divided into three segments for better understanding of the problem to the core, first is a review of the concepts and methods proposed for measurement of rural unemployment followed by a brief comment on dataset used and third an interpretation of the outcomes from the given data in the light of reforming and institutional changes taking place within the rural sector in the Indian economy.

- 4) Trade liberalization and unemployment: Theory and evidence from India

The aim of this study is to provide a statistical analysis on the speculated view that trade liberalization increases unemployment. In this research, state and industry-level trade and unemployment data from India was deployed, but no effects of unemployment increasing effect of trade reforms was found. Moreover, the state level analysis revealed that unemployment in urban region declines with trade liberalization with flexible labour markets and larger employment shares in majority of exporter industries. Industry level analysis further indicated that workers experiencing reductions in trade protection were less likely to become unemployed. The results can also be explained within a theoretical framework and search-generated unemployment and some institutional features of the Indian economy.

- 5) Modelling suicide and unemployment: a longitudinal analysis covering 63 countries, 2000–11

This study tests the relationship between unemployment and suicide rates association after the 2008 economic crisis. Since unemployment has a direct effect on an individual's assets and thereby, his health, it is unsurprising that studies have shown an association between suicide and unemployment. As a statistical model that describes this relationship by considering specific time trends among age-gender-country subgroups conducted over the world regions is not yet available, but this study aims to enhance the knowledge of specific effect of unemployment in relation with suicide by analysing the global public data that has been classified into different world regions. The public data for population, suicide and economy was analysed from the WHO mortality database and International Monetary Fund's world economic outlook database from the year 2000 to 2011 of over 63 countries based on sample size and completeness of the respective data and extracted the information about four age groups and sex. It was found that the overall model for all world regions, adjusted for unemployment rate showed a increase in annual relative risk of suicide by 20-30% during the study period. Overall, 41,148 suicides were associated with unemployment in 2007 and 46,131 in 2009, indicating 4983 excess suicides since the economic crisis in 2008. This study also helps us to strategize and employ unemployment prevention strategies not only in difficult times but also in times of stable economy.

- 6) Econometric Analysis to Examine the Relationship between Unemployment and Macroeconomics Aggregates. Evidence from Kosovo

The aim of this study is to analyse the correlation and affect that macroeconomic aggregates in unemployment in Kosovo. The primary attributes of this study were Gross Domestic Product, Inflation, Exports and Foreign Direct Investment, and to prove their interrelationship an econometrical model was constructed with multiple regressions analysis based on the Ordinary Least Square (OLS) method, utilizing secondary data from World Bank Indicators for the time period of 2001 till 2018. The conclusion obtained is that in Kosovo, the link between unemployment and economic growth is significant and the results show that 1% growth of Gross Domestic Product will lower the unemployment rate by 1.7%, for the given time period. Also, the results of the regression model show that the relationship between Export and unemployment rate is significant while the relation between Foreign Direct Investment and unemployment rate results negative as growth for 1% of Foreign Direct Investment will increase the Unemployment rate to 1.25. Meanwhile, the results between unemployment and inflation showed an insignificant relationship between them.

7) A longitudinal analysis of ethnic unemployment differentials in the UK

This study focuses on the relatively higher rate of unemployment of ethnic minorities, like in many developed countries, the white British majority has comparatively lower unemployment rate, which may be due to the higher possibility of minorities entering unemployment by losing a job, or to a lower probability of exiting unemployment by finding a job. This paper analyses what job or personal characteristics cause unemployment, what generates employment, and to what extent ethnic unemployment depends on these characteristics. Segregation in occupations is triggered by low wages and less stable jobs and does not seem to contribute to the higher unemployment rate of ethnic minorities. For both men and women, the results show no significant ethnic differences in the probability to transition from employment into unemployment. But Indian UK born women seem more likely to transition than British majority women, while for other groups the small differences are in favour of ethnic minorities. The main factor of ethnic unemployment is found to be the longer duration of unemployment for ethnic minorities, however the reason for the same remains unexplained after the inclusion of individual and domestic characteristics.

8) Youth employment and unemployment in India

This paper focuses on a grave issue that is a challenge for future economic growth in India, that is, youth unemployment. Although there is increase in school and college enrolment rates, the proportion of youth in labour force has been declining for quite some time, these high proportions in the labour force indicate that the problem of youth unemployment and underemployment would remain a serious policy issue for the future of India. This paper studies these unemployment issues through the year 1983 to 2008. It analyses the trends in labour force and workforce participation rates, unemployment, joblessness, working poor and growth. The paper also offers policy recommendations for increasing productive employment and reduction in unemployment for the youth. The poor employability of the workforce would hamper the advantages due to demographic dividend if measures are not taken to improve the educational attainment and skill development of the youth.

9) Post-pandemic Penury of the Financially Marginalized in India: Coping with the New Normal

COVID-19 has hardly hit the poor in India, as per World Bank 2020 estimates, there are around 50.7 million people in India in extreme poverty. Although the government of India provided several remedial measures, the effective implementation of the same is a major challenge due to the population density. The different segments of poor have suffered from the impact of the pandemic and subsequent lockdown, in different proportions, this paper studies these effects of unemployment. At the lowest end of the spectrum, the migrant labours are worst affected, they are required to travel for their homes due to loss of jobs and require immediate relief. At the other end, are the skilled labours, who have not lost their jobs but are facing reduction in wages. In order to make the systematic assessment of the requirements, the study is broken down and grouped the people in three categories, high impact, medium impact and low impact.

10) Livelihoods in COVID times: Gendered perils and new pathways in India

This paper presents the argument that to understand the gendered impact of livelihood loss due to COVID-19, we need not only examine the direct effects on women's earnings but also the indirect effects on intra-household dynamics, such as food insecurity, depletion of savings and assets, social isolation, and mobility loss. These are faced by not only women who have lost jobs but also by unpaid workers which triggers unemployment rates directly. Drawing on telephone surveys and other emerging evidence on the pandemic in India, this piece examines the direct and indirect effects on women of livelihood losses by both genders, especially in poor households, as well as the lessons offered by women-led group approaches for charting new developmental pathways.



### III. PROPOSED ANALYSIS APPROACH

We have taken the dataset from Kaggle which has the following Urban: Rural distribution-

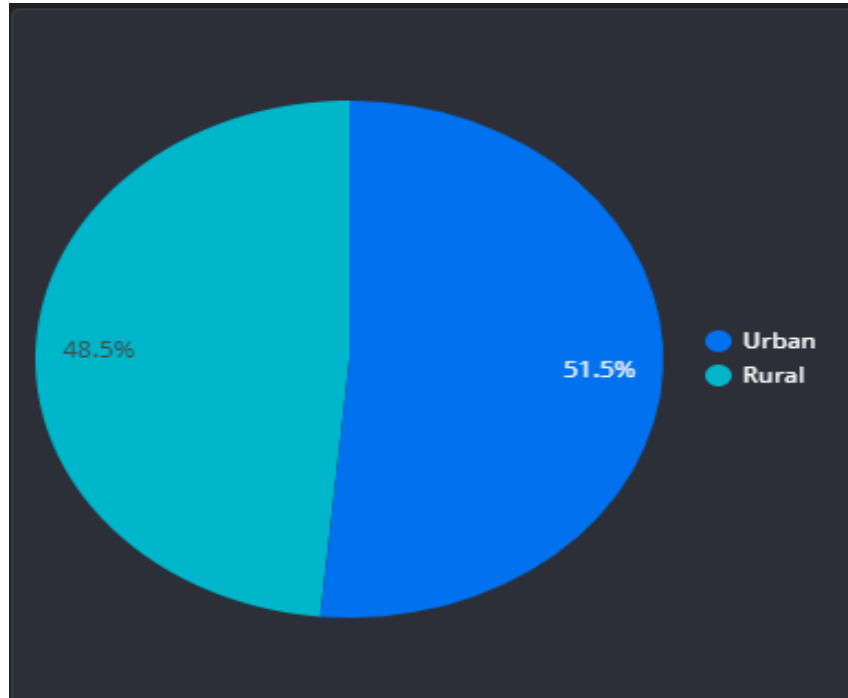


Fig 1: Data Distribution

We are using the Simplified Formula for Proportions to find the suitable sample size for our study. In this, the assumptions are: Proportion (p) = 0.5 and a confidence of 95%.

$$n = \frac{N}{1 + N(e)^2}$$

Where n - sample size  
 N – Population size  
 e – precision level

$$n = \frac{N}{1 + N(e)^2} = \frac{721}{1 + 721(.05)^2} = 251$$

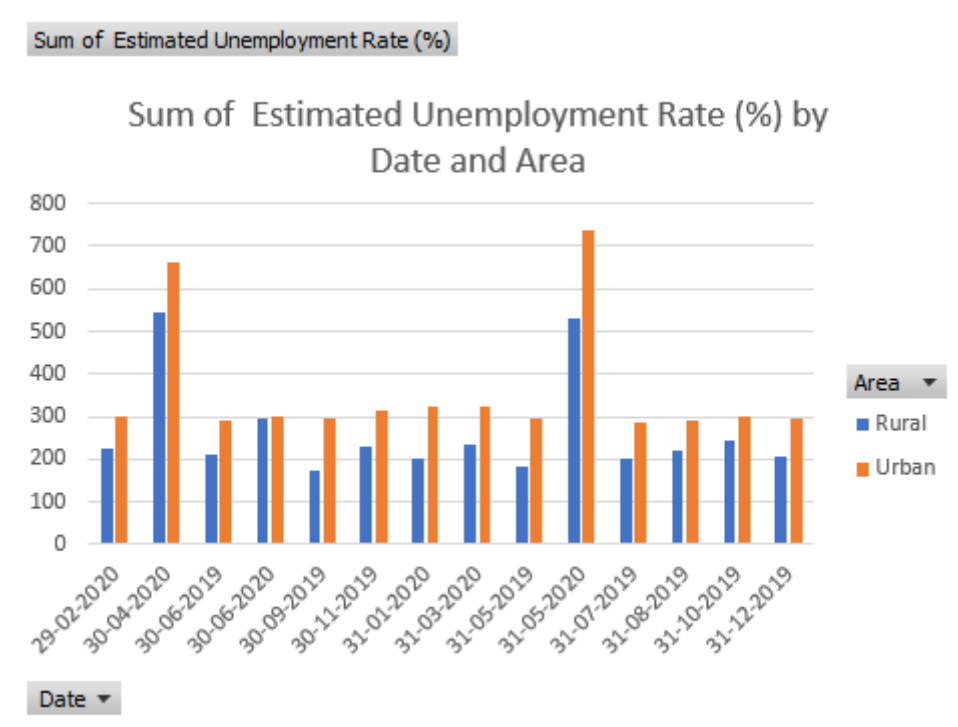
Hence our ideal sample size comes out to be 251.

Since our rural and urban regions are almost equally proportionate the division of sample is almost going to be the same, that is, 125 for rural and 126 for urban.

Here is a date-wise tabular as well as pictorial representation of our sample data:



Date	Sum of Estimated Unemployment Rate (%)	
	Rural	Urban
29-02-2020	227.56	300.57
30-04-2020	543.65	662.07
30-06-2019	213.23	289.15
30-06-2020	295.63	299.55
30-09-2019	175.92	294.77
30-11-2019	227.67	315.09
31-01-2020	203.91	323.48
31-03-2020	232.4	324.03
31-05-2019	183.77	295.44
31-05-2020	530.27	738.37
31-07-2019	201.29	286.54
31-08-2019	221.08	289.73
31-10-2019	244.38	300.17
31-12-2019	205.84	297.52



A representation of data helps us visualize the problem but clearly this doesn't help us arrive at a conclusion hence we continue our analysis by estimating the population parameter.

**Estimating Population Parameter**

In Statistics, we use sample statistics to estimate population parameters. The sample mean will be used to estimate the population parameter. We calculate the mean of the strata and then the sample mean.



Urban strata mean: 12.857

Rural strata mean: 9.105

$$\bar{x} = \frac{\sum_{i=1}^n x_i}{n}$$

$$\bar{x} = 9.6117$$

### Estimating population variance

Sample variance is used to estimate population variance.

$$S^2 = \frac{\sum_{i=1}^n [(x)_i - \bar{x}]^2}{n - 1}$$

Urban strata variance: 108.13

Rural strata variance: 54.24

$$S^2 = 54.64$$

Sample variance is 54.64

### Calculating Standard Deviation

$$\text{Standard Deviation}(SD) = \sqrt{\text{Variance}} (S^2)$$

$$SD = 9.8595$$

Therefore, the standard deviation is 9.8595

### Computing Standard Error

Standard error represents the variation in the sample means of a sampling distribution.

$$SE = \frac{SD}{\sqrt{n}}$$

$$SE = 0.7229$$

### Confidence Level

Confidence interval is the percentage of all possible samples expected to include the true population parameter. A confidence level of 95% means the confidence interval would include the true population parameter, which is what we have assumed.



### Finding the Critical Value

The critical value is a factor used to compute the margin of error.

$$\alpha = 1 - \left( \text{confidence} \frac{\text{level}}{100} \right)$$

$$\alpha = 1 - \left( \frac{95}{100} \right)$$

$$\alpha = 0.05$$

Critical Probability( $p^*$ )

$$p^* = 1 - \alpha/2$$

$$p^* = 1 - 0.05/2$$

$$p^* = 0.975$$

Here we use Z-test for our sample size is large. For small samples, t-test is used.

$$\text{Critical value} = 1.96$$

### Margin of Error

The margin of error gives the amount of random sampling error in the result.

Margin of error = 7.938

$$ME = \text{Critical Value} \times \text{Standard Error}$$

$$ME = 1.96 \times 0.7229$$

$$ME = 1.4168$$

The margin of error of our sample is 1.4168

### Confidence Interval

The lower limit of the Confidence interval

$$Cl_{min} = \bar{x} - ME$$

$$Cl_{min} = 9.6117 - 1.4168$$

$$Cl_{min} = 8.1949$$

Upper Limit of the Confidence interval

$$Cl_{max} = \bar{x} + ME$$

$$Cl_{max} = 9.6117 + 1.4168$$

$$Cl_{max} = 11.0285$$





Confidence interval – 8.19949 to 11.0285

#### IV. FUTURE SCOPE

According to the study based on the dataset obtained we can draw a plan based on the current requirements in the perspective of the given trends. In our study we used Stratified Sampling method instead of Random Sampling method in order to analyse the data on the basis of its true distribution into urban and rural region. Stratified Sampling method was used because the data was divided into subgroups and this method provides more precise representation of our data. In this paper we found that pandemic has adversely affected the unemployment rate of our country and we need to take necessary steps to prevent the rising of these rates.

#### V. CONCLUSION

The aim of this paper was to get an idea about the unemployment rates of India and which among the urban and rural regions is more affected by the pandemic. For this we conducted a statistical analysis on the states of India and drew a sample of 251 random regions; 125 from rural and 126 from urban region; and used Stratified sampling for precise representation. The estimated mean was found out to be 9.6117 with a confidence interval of 8.19 - 11.02. The ideal unemployment rate is considered to be 4-5%, hence we can say that India has been severely affected by the pandemic in terms of unemployment.

#### REFERENCES

- 1) Visaria, P., 1981. Poverty and unemployment in India: An analysis of recent evidence. *World Development*, 9(3), pp.277-300.
- 2) Sundaram, K. and Tendulkar, S.D., 1988. Toward and explanation of interregional variations in poverty and unemployment in rural India. *Rural poverty in South Asia.*, pp.316-362.
- 3) Raj, K.N., 1976. Trends in rural unemployment in India: an analysis with reference to conceptual and measurement problems. *Economic and Political Weekly*, pp.1281-1292.
- 4) Hasan, R., Mitra, D., Ranjan, P. and Ahsan, R.N., 2012. Trade liberalization and unemployment: Theory and evidence from India. *Journal of Development Economics*, 97(2), pp.269-280.
- 5) Nordt, C., Warnke, I., Seifritz, E. and Kawohl, W., 2015. Modelling suicide and unemployment: a longitudinal analysis covering 63 countries, 2000–11. *The Lancet Psychiatry*, 2(3), pp.239-245.
- 6) Ziberi, B. and Avdiu, M., 2020. Econometric analysis to examine the relationship between unemployment and macroeconomics aggregates. Evidence from Kosovo. *Academic Journal of Economic Studies*, 6(2), pp.33-41.
- 7) Longhi, S., 2020. A longitudinal analysis of ethnic unemployment differentials in the UK. *Journal of Ethnic and Migration Studies*, 46(5), pp.879-892.
- 8) Venkatanarayana, M. and Mahendra Dev, S., 2012. Youth employment and unemployment in India.
- 9) Singh, A.S., Venkataramani, B. and Ambarkhane, D., 2021. Post-pandemic penury of the financially marginalized in India: Coping with the new normal. *The Future of Service Post-COVID-19 Pandemic, Volume*, p.201.
- 10) Agarwal, B., 2021. Livelihoods in COVID times: Gendered perils and new pathways in India. *World Development*, 139, p.105312.

Other links:

[economictimes.indiatimes.com/definition/stratified-sampling/](http://economictimes.indiatimes.com/definition/stratified-sampling/) as accessed on March 29<sup>th</sup>  
<https://www.scribbr.com/methodology/stratified-sampling/> as accessed on March 29<sup>th</sup>



INNO  SPACE  
SJIF Scientific Journal Impact Factor

Impact Factor:  
7.488

**ISSN** INTERNATIONAL  
STANDARD  
SERIAL  
NUMBER  
INDIA



# INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

 9940 572 462  6381 907 438  [ijircce@gmail.com](mailto:ijircce@gmail.com)



[www.ijircce.com](http://www.ijircce.com)

Scan to save the contact details