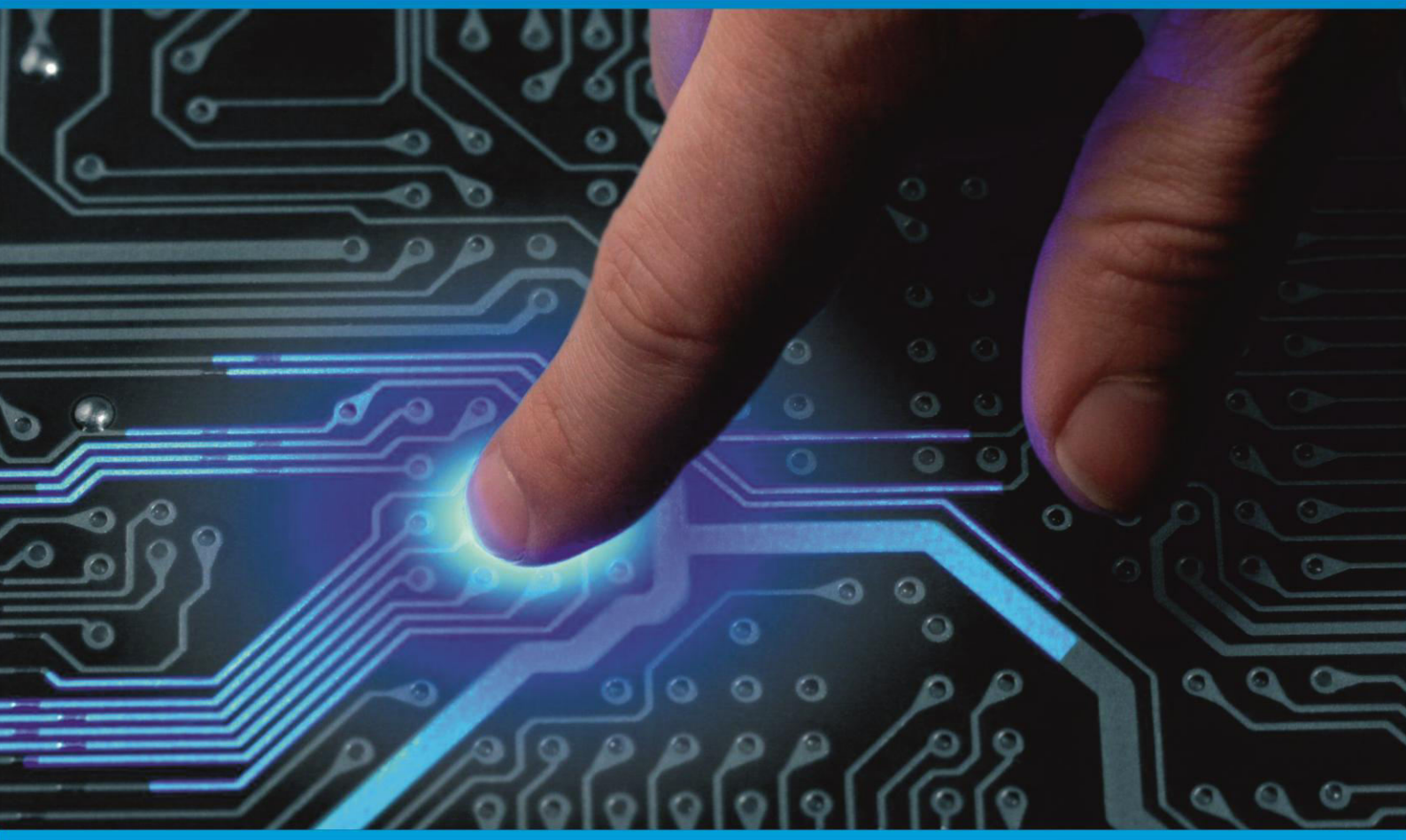
IJIRCCE

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



INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

Volume 9, Issue 5, May 2021

INTERNATIONAL STANDARD SERIAL NUMBER INDIA

Impact Factor: 7.488

International Journal of Innovative Research in Computer and Communication Engineering

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 7.488 |

|| Volume 9, Issue 5, May 2021 ||

| DOI: 10.15680/IJIRCCE.2021.0905052 |

Statistical Analysis of Sex-Ratio of India

Nikhil Kudupudi[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 : The sex ratio is an essential social indicator for any country. For our research, we would draw the sample from the 2011 census data of India. For sampling, the simple random sampling technique is used as its easy and reliable technique. Some Statistical calculations are used to analyze the sample and compute the estimated mean of the population. With the estimated population mean and confidence interval, we will get an idea of the sex ratio of India.

KEYWORDS: Sex-Ratio, Sampling, Simple Random Sampling, Population

I. INTRODUCTION

The ratio of females to males or vice-versa is called the sex ratio. The sex ratio for any country should 1:1 or 1000 females per 1000 males. Gender imbalance can cause many social and political issues in that country. For our study, I have taken data on the sex ratio of India which is collected from the 2011 Census of India. This data contains the sex ratio of every district in India. There are seven columns, Sr.no, District, State, Population, Sex Ratio, and 640 rows in this dataset.

In this paper, a sample has been drawn from the population to estimate the population parameters. The simple Random Sampling technique is used to draw a sample from the population, as simple Random Sampling is an effortless technique to draw a sample from the population. At the same time, it provides an accurate representation of the population. In simple random sampling, every element/row has an equal chance of being selected in the sample. After estimating the population's mean and variance, 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 sex ratio. If the mean is close to 1000 (1000 females per 1000 males), we would say India's sex ratio is perfect.

II. LITERATURE SURVEY

[1] Female Sex-Ratio in India
In this paper author, 'Vijay Kumar Sarabu' discusses the declining sex ratio of India. For this research, he compared the sex ratio of India of last one decade (1901-2011) and found out that there was a steady drop of sex ratio from 1901 to 1941 and from 2001 there is a slight upward trend. He compared the sex ratio of India with the neighbouring countries and saw that Afghanistan ranked last, followed up by India being second last. This paper also analyzed the 2001 sex-ratio of Indian states data with 2011 data and Indian rural sex ratio with urban cities. The author described some factors responsible for India's low female sex ratio like married couples preferring boy over a girl, discrimination against girls in every aspect of life like education, marriage, and employment. This paper also discussed the implication of this, like the marriage squeeze, re-emergence of bride price, and increased crime against women. Some suggestions given by the author to stabilize the Indian sex ratio are change in the attitude of men towards women, stopping gender discrimination by imposing strict rules, and by spreading awareness.

[2] Sex Ratio of Maharashtra state: a geographical analysis
This paper's primary focus is to analyze the sex ratio of Maharashtra State with the 2011 census report. To understand the decline of the sex ratio of Maharashtra state author has studied at the district level to identify the core problem. According to the author, the Vidarbha region has a high sex ratio with an average of 977. In contrast, the Konkan region of Maharashtra has a low sex ratio with an average of 940. The average Sex ratio of North Maharashtra, Khandesh, Western Maharashtra, and Marathwada region are 962, 964, 967, 959, respectively.

[3] A study of Correlation between Literacy Rate and Sex Ratio in Haryana: Trend and Emerging Issues
This paper's main focus was to study sex ratio trends in India and Haryana and inspect the correlation between sex ratio and literacy rate in the Haryana district. The author used Spearman's Rank Correlation method to calculate the

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 7.488 |

|| Volume 9, Issue 5, May 2021 ||

| DOI: 10.15680/IJIRCCE.2021.0905052 |

correlation between sex ratio and literacy rate among Haryana districts. The author got a negative correlation, i.e., r = -0.54, which tells us if the literacy rate increases, the sex ratio decreases and vice-versa.

[4] The Issue of gender ratio in India: An analysis
This paper aims to find a relationship between the per-capita income and sex ratio of Indian states and union territory. The author divided the states and UTs of India into two groups. States and UTs with higher per-capita income were in the first group and others in the second group when compared to the average per-capita of India. The author found a negative correlation between per-capita income and sex ratio. States with a high per-capita income had a lousy sex ratio when compared with low per-capita states.

[5] Determining sample size
Author Glenn D. Israel has discussed three criteria: level of precision, Confidence level, and Degree of variability, which is required to determine the appropriate sample size. This paper uses various approaches to determine the sample size, like Using a census for a small Population, Using a sample size of a similar study, Using Published Tables, and Using formulas to calculate the sample size.

[6] Determining Sample Size For Research Activities
This paper used a formula published in the article "Small Sample Techniques" to determine the sample size of different population sizes. The author created a table where he computed sample size for different population sizes and plotted the graph of the same. As the population increases, the author found that the sample size increased at a slow rate and remained somewhat constant from 380 sample size.

[7] Sampling: Why and how of it?
This paper explains why sampling is essential and its two types, probability sample and Non-Probability Sample. Further, the author explained the subtypes of probability sampling with examples and their advantages and disadvantages. The author concluded that probability sampling methods are the best approach for the generalizability of the results to the target population and ensure the sample is representative.

[8] Pros and cons of different sampling techniques
This paper clarifies the concept of sampling. The author explains subtypes and different categories present in those subtypes of sampling. This paper states the pros and cons of each type of sampling and tells which sampling technique to use when.

[9]Standard Deviation and Standard error of the mean
This paper explains the difference between Standard Deviation and Standard Error of mean with their formulas and plotted graphs to explain the same. This paper concluded that the Standard Error of Mean could describe the characteristics of a sample. When analyzing statistical results, the Standard Error of Mean coupled with sample size is more beneficial because it allows for a visual comparison of the estimated population through graphs and tables.

[10] Using the confidence interval confidently
This paper gives detailed information about Confidence interval, how to calculate them, their practical use in clinical trials and some misconceptions regarding confidence interval. The author illustrates how the confidence interval's width increases with the increase in confidence level and how the confidence interval's width is inversely related to sample size.

III. PROPOSED ANALYSIS APPROACH

We are using the Simplified Formula for Proportions [5] to find the appropriate sample size for our study. For this, we have assumed Proportion (p) = 0.5 because 0.5 indicates maximum variability in a population and a confidence of 95%.

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

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

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 7.488 |

|| Volume 9, Issue 5, May 2021 ||

| DOI: 10.15680/IJIRCCE.2021.0905052 |

$$n = \frac{N}{1 + N(e)^2} = \frac{640}{1 + 640(.05)^2} = 246$$

So, our sample size is 246.

Visualization of Population Sizeof Different States

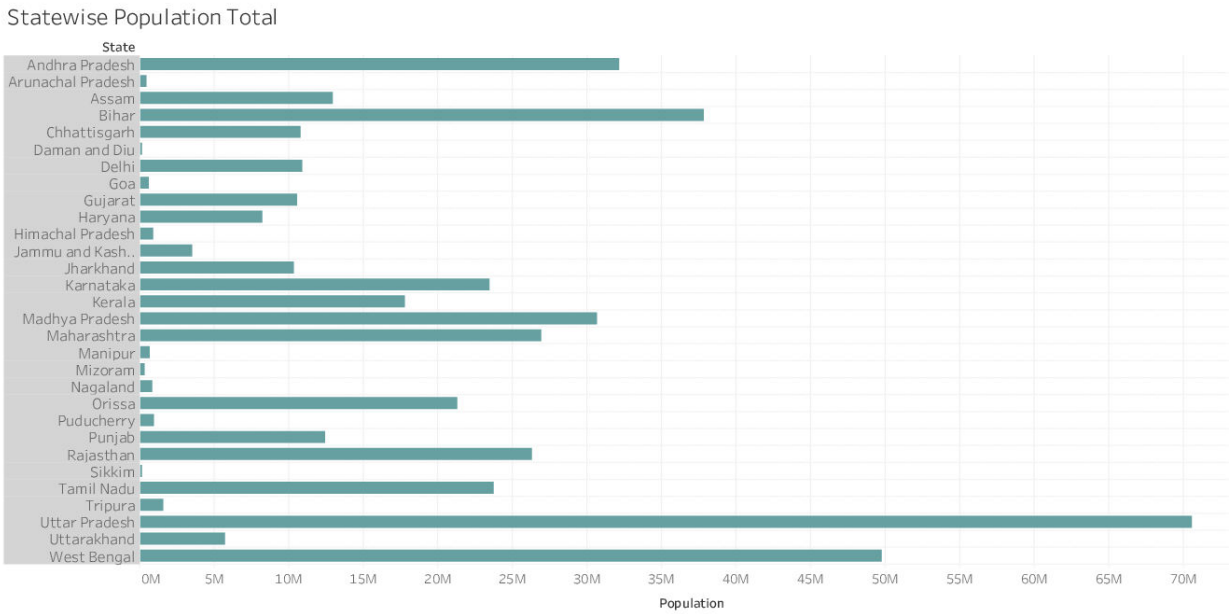


Figure 1. Bar Graph of Statewise Population of India. Visualization: Tableau

Visualization of AverageSexRatio of Different States

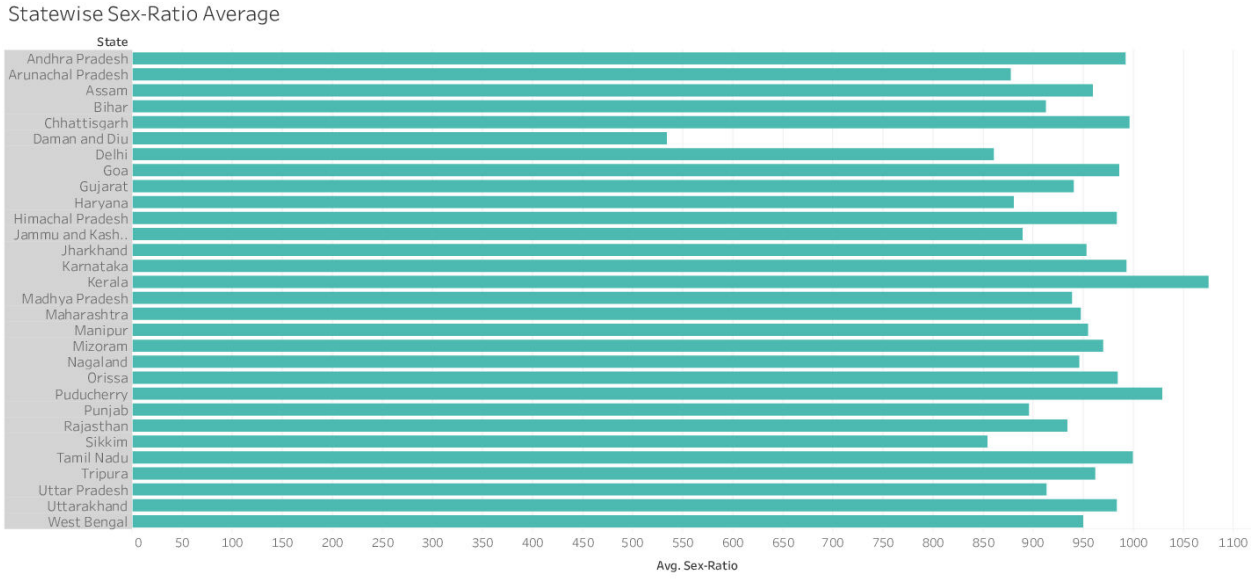


Figure 2. StatewiseAverage Sex-Ratio of India. Visualization: Tableau

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 7.488 |

|| Volume 9, Issue 5, May 2021 ||

| DOI: 10.15680/IJIRCCE.2021.0905052 |

Estimating Population Parameter

In Statistics, we use sample statistics to estimate population parameters. The sample mean will be used to estimate the population parameter

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

$$\bar{x} = 946.38$$

Estimating population variance

Sample variance is used to estimate population variance.

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

$$S^2 = 3970.16$$

Sample variance is 3970.16

Calculating Standard Deviation

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

$$SD = \sqrt{3970.16}$$

$$SD = 63.00$$

Therefore, the standard deviation is 62.84

Computing Standard Error

Standard error represents the variation in the sample means of a sampling distribution [9]. The standard error is the expected value of the standard deviation of means of several samples.

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

$$SE = \frac{63}{\sqrt{246}}$$

$$SE = 4.01$$

Confidence Level

Confidence interval is the percentage of all possible samples expected to include the true population parameter. For our study, we would select a Confidence level of 95%. A confidence level of 95% means the confidence interval would include the true population parameter.

Findingthe Critical Value

The critical value is a factor used to compute the margin of error. To compute the critical value, first, we have to calculate α

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

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

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 7.488 |

|| Volume 9, Issue 5, May 2021 ||

| DOI: 10.15680/IJIRCCE.2021.0905052 |

$$\alpha = 0.05$$

Critical Probability(p*)

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

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

$$p* = 0.975$$

For finding the critical value, there are two tests, T-test and Z-test. We use T-test when Sample is smaller than 15 and the population variance is unknown. Z-test is used when population size is big or greater than 15, and population variance is known.

As the sample size is 246, which is significantly big, we will use Z-test. Using the Standard Normal Distribution table to find the critical value of the z-score, we got

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

Margin of Error

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

Margin of error = 7.938

$$ME = Critical\ Value \times Standard\ Error$$
$$ME = 1.96 \times 4.01$$
$$ME = 7.859$$

The margin of error of our sample is 7.85

Confidence Interval

An approximate range of values that is likely to include an unknown population parameter is given by confidence interval, which is calculated from sample data

The lower limit of the Confidence interval

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

$$Cl_{min} = 946.38 - 7.85$$

$$Cl_{min} = 938.53$$

Upper Limit of the Confidence interval

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

$$Cl_{max} = 946.38 + 7.85$$

$$Cl_{max} = 954.23$$

Confidence interval – 938.53 to 954.23

| e-ISSN: 2320-9801, p-ISSN: 2320-9798| www.ijircce.com | |Impact Factor: 7.488 |

|| Volume 9, Issue 5, May 2021 ||

| DOI: 10.15680/IJIRCCE.2021.0905052 |

IV. FUTURE SCOPE AND DISCUSSION

For this study, the data is taken from the 2011 Census of India. In 2021 Indian government will conduct the census of India. After the completion of the census 2021 data, we can compare our results with it and check if the sex ratio of India has got worse or much better. In this paper, we have used Simple Random Sampling as there were no clear subgroups in data to work with Stratified Random Sampling. Based on what criteria is used to segment the population into various subsets, Stratified Random Sampling may provide a more precise representation. If 2021 census data gives sex ratio in rural and urban parts of the district, we can utilize the Stratified Random Sampling technique to represent the population better.

V. CONCLUSION

We aimed to get an idea of the sex ratio of India, for that statistical analysis was necessary to estimate the population average. We draw a sample of 246 districts of India with their sex ratio from the population using a simple random sampling technique. The estimated population mean is 946.38 with a confidence interval of938.53- 954.23. So, we are 95% confident that the true population mean would lie between938.53 to 954.23. The ideal sex ratio for any country should be 1:1, or we can say 1000 females per 1000 males. The estimated sex ratio of India is less when compared to the ideal sex ratio. A low sex ratio can lead to major demographic problems in India.

REFERENCES

1.Sarabu, Vijay. (2012). FEMALE SEX -RATIO IN INDIA – A REVIEW. 8. 29-44.
2. Mane, Santosh. (2020). Sex Ratio of Maharashtra State: A Geographical Analysis. 19. 200- 206
3. Malik, Ruchi. (2018). A study of Correlation between Literacy Rate and Sex Ratio in Haryana: Trend and Emerging Issues. International Journal of Research. 05. 103-110.
4. Kumar, Sumesh, and Komal Sharma. "The Issue of Gender Ratio in India: An analysis." Journal of Management Research and Analysis 4.4 (2017): 128-132.
5. Israel, Glenn D. "Determining sample size." (1992): 2017.
6. Krejcie, Robert V., and Daryle W. Morgan. "Determining sample size for research activities." Educational and psychological measurement 30.3 (1970): 607-610.
7. Acharya, Anita S., et al. "Sampling: Why and how of it." Indian Journal of Medical Specialties 4.2 (2013): 330-333.
8.Sharma, Gaganpreet. "Pros and cons of different sampling techniques." International journal of applied research 3.7 (2017): 749-752.
9.Lee, Dong Kyu, Junyong In, and Sangseok Lee. "Standard deviation and standard error of the mean." Korean Journal of anesthesiology 68.3 (2015): 220.
10.Hazra, Avijit. "Using the confidence interval confidently." *Journal of thoracic disease* vol. 9,10 (2017): 4125-4130. doi:10.21037/jtd.2017.09.14

INNO SPACE
SJIF Scientific Journal Impact Factor

Impact Factor:
7.488

ISSN INTERNATIONAL STANDARD SERIAL NUMBER INDIA

निस्केयर
NISCAIR

INTERNATIONAL JOURNAL OF INNOVATIVE RESEARCH

IN COMPUTER & COMMUNICATION ENGINEERING

9940 572 462 6381 907 438 ijircce@gmail.com

www.ijircce.com

Scan to save the contact details