Capture-Recapture Sampling Techniques: Artificial and Real Population Data Analysis

Neelam Bukhari 1, Sadia Bukhari 2, Muhammad Ahmed Shehzad 3

1,3 Department of Statistics, Bahauddin Zakariya, University, Multan
2 Higher Education Commission of Pakistan

* Corresponding author: ahmad.shehzad@bzu.edu.pk

INTRODUCTION

Capture-Recapture sampling methods are being used to estimate the population size of an unknown population. These methods are widely used for determining the population size of animals and birds. Literature shows that these methods were purely used by ecologist in the past. But now-a-days, these methods are also used to determine the chronic diseases, for example, cancer patients, HIV Aid patients, etc. Not only this, many studies had conducted to determine the road accidents ratios on different points. This study covered two examples; one was the total number of female drivers at Bahauddin Zakariya University, Multan and the second example was about the total number of male smokers at Bahauddin Zakariya University, Multan. On these two examples, the basic capture-recapture sampling methods were used which involved Lincoln Peterson Index, Chapman Estimate and Schnable Estimate. Then an artificial data with known population was created and same methods were used for analysis. In the end, the results of artificial data and two examples were compared with each other through analysis on excel sheet.

Keywords: Capture-Recapture sampling, Artificial P(population), Real Population

Capture-Recapture sampling methods are being used to estimate the population size of an unknown population. These methods are widely used for determining the population size of animals and birds. Literature shows that these methods were purely used by ecologist in the past. But now-a-days, these methods are also used to determine the chronic diseases, for example, cancer patients, HIV Aid patients, etc. Not only this, many studies had conducted to determine the road accidents ratios on different points. This study covered two examples; one was the total number of female drivers at Bahauddin Zakariya University, Multan and the second example was about the total number of male smokers at Bahauddin Zakariya University, Multan. On these two examples, the basic capture-recapture sampling methods were used which involved Lincoln Peterson Index, Chapman Estimate and Schnable Estimate. Then an artificial data with known population was created and same methods were used for analysis. In the end, the results of artificial data and two examples were compared with each other through analysis on excel sheet.

Keywords: Capture-Recapture sampling, Artificial Population, Real Population

INTRODUCTION

Capture-Recapture sampling methods are being used to estimate the population size of an unknown population. These methods are widely used for determining the population size of animals and birds. Literature shows that these methods were purely used by ecologist in the past. But now-a-days, these methods are also used to determine the chronic diseases, for example, cancer patients, HIV Aid patients, etc. Not only this, many studies had conducted to determine the road accidents ratios on different points. This study covered two examples; one was the total number of female drivers at Bahauddin Zakariya University, Multan and the second example was about the total number of male smokers at Bahauddin Zakariya University, Multan. On these two examples, the basic capture-recapture sampling methods were used which involved Lincoln Peterson Index, Chapman Estimate and Schnable Estimate. Then an artificial data with known population was created and same methods were used for analysis. In the end, the results of artificial data and two examples were compared with each other through analysis on excel sheet.

Keywords: Capture-Recapture sampling, Artificial Population, Real Population

Capture-Recapture sampling methods are being used to estimate the population size of an unknown population. These methods are widely used for determining the population size of animals and birds. Literature shows that these methods were purely used by ecologist in the past. But now-a-days, these methods are also used to determine the chronic diseases, for example, cancer patients, HIV Aid patients, etc. Not only this, many studies had conducted to determine the road accidents ratios on different points. This study covered two examples; one was the total number of female drivers at Bahauddin Zakariya University, Multan and the second example was about the total number of male smokers at Bahauddin Zakariya University, Multan. On these two examples, the basic capture-recapture sampling methods were used which involved Lincoln Peterson Index, Chapman Estimate and Schnable Estimate. Then an artificial data with known population was created and same methods were used for analysis. In the end, the results of artificial data and two examples were compared with each other through analysis on excel sheet.

Keywords: Capture-Recapture sampling, Artificial Population, Real Population

Capture-Recapture sampling methods are being used to estimate the population size of an unknown population. These methods are widely used for determining the population size of animals and birds. Literature shows that these methods were purely used by ecologist in the past. But now-a-days, these methods are also used to determine the chronic diseases, for example, cancer patients, HIV Aid patients, etc. Not only this, many studies had conducted to determine the road accidents ratios on different points. This study covered two examples; one was the total number of female drivers at Bahauddin Zakariya University, Multan and the second example was about the total number of male smokers at Bahauddin Zakariya University, Multan. On these two examples, the basic capture-recapture sampling methods were used which involved Lincoln Peterson Index, Chapman Estimate and Schnable Estimate. Then an artificial data with known population was created and same methods were used for analysis. In the end, the results of artificial data and two examples were compared with each other through analysis on excel sheet.

Keywords: Capture-Recapture sampling, Artificial Population, Real Population

Capture-Recapture sampling methods are being used to estimate the population size of an unknown population. These methods are widely used for determining the population size of animals and birds. Literature shows that these methods were purely used by ecologist in the past. But now-a-days, these methods are also used to determine the chronic diseases, for example, cancer patients, HIV Aid patients, etc. Not only this, many studies had conducted to determine the road accidents ratios on different points. This study covered two examples; one was the total number of female drivers at Bahauddin Zakariya University, Multan and the second example was about the total number of male smokers at Bahauddin Zakariya University, Multan. On these two examples, the basic capture-recapture sampling methods were used which involved Lincoln Peterson Index, Chapman Estimate and Schnable Estimate. Then an artificial data with known population was created and same methods were used for analysis. In the end, the results of artificial data and two examples were compared with each other through analysis on excel sheet.

Keywords: Capture-Recapture sampling, Artificial Population, Real Population
Results

The results of first real population example (female drivers at Bahauddin Zakariya University, Multan) showed that highest value of population (N=30) and lowest value of population was (N=22) when the procedure of Peterson Estimate was repeated for the six times. The mean was calculated as (mean=21.74) and highest variance was observed as (var=13). Chapman estimate results for first real population example showed that the highest value of estimated population was (N=29) and the lowest value was (N=21). The mean was observed as (mean=21.41) and highest value of variance was recorded as (var=14). Again, the procedure of estimating the population size by Chapman Estimate was run for the six times. The results of Schnable estimate indicated that estimated population for example 1 was (N=25) with a variance (var=5). Real population Example 2 (male smokers at Usman Hall of Bahauddin Zakariya University, Multan) was analyzed by using the Lincoln-Index, Chapman and Schnable Estimate. The analysis was run on excel sheet. For Peterson-Index, the highest value of population estimate was (N=40) and the lowest was (N=23). The average estimated population was (mean=26) with a highest variance (var=15). To estimate the population size and average value, the procedure was repeated for the six times. The results of Chapman Estimate indicated that the highest value of estimated population was (N=39) and lowest estimated population was (N=23) and mean was calculated as (mean=25) with a highest variance (var=16). The estimation of population through Schnable estimate was found (N=35) with a variance (var=3).

The real population data analysis was done to estimate the population size. Then researcher generated the artificial data with a known population size (N=100). The Peterson, Chapman and Schnable procedure were applied on this artificial data and it was analyzed that how much deviation in results exist for the already known population. The highest value of population estimate was found (N=106) and the lowest value was (N=97) for Peterson Index. The average value was found (mean=100) with a highest variance (var=19). The estimated population for Chapman was recorded as (N=101), the highest value and (N=92) as lowest value. The mean was (mean=97) with a highest variance (var=5). The Schnable Estimate results indicated the estimated population as (N=106) with a variance (var=2).

Table 1 represents the comparison of both real population examples with artificial data. The analysis of artificial data indicated that the estimation of population size was very close to the actual size of the population. Hence, in example 1, it is concluded that total number of female drivers at Bahauddin Zakariya University are 25 to 30. Similarly, in example 2, it is concluded that total number of male smokers at Usman Hall of Bahauddin Zakariya University, Multan are 35 to 40.

Discussion

The findings of the current study depicts that the result of all methods are more or less similar. Later, the artificial data analysis with known population also confirms the accuracy of results. It is evident from literature that these basic methods are widely used in wild life and by ecologist. However, this study is unique in nature as none of the study has been with artificial data analysis with known population by using the Peterson Index, Chapman estimate and Schnable estimate. This study also has limitations as it is conducted in closed population and in Bahauddin Zakariya University while if the study would conducted at city level then result would definitely be different. Further, this study is done by using the simplest methods of capture-recapture techniques. There are lots of other methods for such techniques. They can also be used on these real population examples and later comparison can be made with artificial generated data.

References:

Chapman, D. G. (1951). Some Properties of the Hypergeometric Distribution with Applications to Zoological Sample Censuses, University of California Pabul. Stat. 17(1):13-1160.

Jibasen, D. (2011). Capture-recapture Type Models for Estimating the Size of an Elusive Population. A PhD Thesis submitted to the University of Ilofin, Ilofin, Nigeria.

Pollock KH (2010). Modeling capture, Recapture and Removal Statistics for Estimation of Demographic Parameters for Fish and Wildlife Populations: past, present and future. J. Amer Stat Assoc.

Mineau, P. & Whiteside, M. (2013). Pesticide Acute Toxicity Is a Better Correlate of US Grassland Bird Declines than Agricultural Intensification. Plos One, 8.

Smallwood, K.S. (2013) Comparing bird and bat fatality-rate estimates among North American wind-energy projects. Wildlife Society Bulletin, 37, 19-33

Table 1: Comparison Table

| Artificial Data | Example 1 | Example 2 |
|-----------------|-----------|-----------|
| Real Population | Peterson  | Chapman   | Schnable  |
| Mean Estimate   | 101.02    | 97.89     | 106.96    |
| Variance        | 21.74     | 21.41     | 25.85     |
| Artificial Data | Peterson  | Chapman   | Schnable  |
| Mean Estimate   | 25.45     | 21.74     | 26.12     |
| Variance        | 25.45     | 25.85     | 35.85     |

Table 1 represents the comparison of both real population examples with artificial data. The analysis of artificial data indicated that the estimation of population size was very close to the actual size of the population. Hence, in example 1, it is concluded that total number of female drivers at Bahauddin Zakariya University are 25 to 30. Similarly, in example 2, it is concluded that total number of male smokers at Usman Hall of Bahauddin Zakariya University, Multan are 35 to 40.