# EVALUATION OF THE HUMAN DEVELOPMENT INDEX (HDI) USING THE METHODS OF SAMPLE THEORY

### Marija Antonijević[*]
*Institute of Economic Sciences Belgrade, Republic of Serbia*

### Đina Ivanović
*Institute of Economic Sciences Belgrade, Republic of Serbia*

***Abstract:*** *This paper aims to determine which of the two sample plans, i.e., a simple random sample without replacement, or a stratified sample, gives a more accurate estimate of the feature's mean. The feature that was the subject of this research is the human development index in 2018. The analysis included 189 countries globally, classified into specific categories according to the United Nations development classification. The research results showed that a more accurate estimate of the mean of the human development index was obtained by applying a stratified sampling since the mean of the human development index is close to the population mean. Also, the variance of the sample mean is lower than the value obtained by applying a simple random sampling without replacement. Therefore, it was justified to approach stratification, which indicates that in the case of conducting research, the use of a stratified sampling should be considered since it provides a more precise estimate of the mean.*

***Keywords:*** *simple random sampling without replacement, stratified sampling, human development index, sample theory*

***JEL classification:*** *C80, C83*

---

[*] marija.antonijevic@ien.bg.ac.rs

# OCENA INDEKSA LJUDSKOG RAZVOJA (HDI INDEX) PRIMENOM METODA TEORIJE UZORAKA

***Sažetak:*** *Cilj ovog rada je da se utvrdi koji od dva plana uzorkovanja daje tačniju procenu srednje vrednosti nekog obeležja - jednostavan slučajni uzorak bez zamene ili stratifikovani uzorak. Predmet ovog istraživanja je indeks humanog razvoja u 2018. godini. Analizom je obuhvaćeno 189 zemalja sveta, razvrstanih u posebne kategorije prema klasifikaciji razvoja Ujedinjenih nacija. Rezultati istraživanja pokazali su da se tačnija procena srednje vrednosti indeksa humanog razvoja dobija primenom stratifikovanog uzorka, budući da je srednja vrednost indeksa humanog razvoja bliska srednjoj vrednosti populacije. Takođe, varijansa srednje vrednosti uzorka je niža od vrednosti dobijene primenom jednostavnog slučajnog uzorkovanja bez zamene. Stoga je bilo opravdano da se pristupi stratifikaciji, što ukazuje da u slučaju sprovođenja istraživanja treba da se razmotri korišćenje stratifikovanog uzorkovanja, budući da ono daje precizniju procenu srednje vrednosti.*

***Ključne reči****: jednostavno nasumično uzorkovanje bez zamene, stratifikovano uzorkovanje, indeks ljudskog razvoja, teorija uzorka*

## 1. INTRODUCTION

Examining all the population elements is often costly, time-consuming, and technically challenging to perform in scientific research. For these reasons, the study is conducted on a sample, i.e., on one part of the population. Choosing an adequate sampling method becomes crucial for obtaining reliable data that should present the characteristics of the population as accurately as possible. According to the sample theory, some of the following sample plans can be applied to the observed population: simple random sampling (without replacement and with replacement), stratified sampling, sample with unequal probabilities, systematic, multi-stage, cluster sampling, etc. (Petrović, 2013; Alvi, 2016). Some studies have shown that different sample designs provide different precision of estimates (Alvarez, del Nero Velasco, Barbin, Lima & Do Couto, 2005; Elsayir, 2014; Okororie & Otuonye, 2015; Wibowo, 2015; Koprivica, 2017; Wang, Zhang, Xue, Xu, Ren & Chen, 2020). When designing and conducting research, the selection of an adequate sample plan is essential. The accuracy of the estimate of the mean of the human development index has not been examined by comparing different sample plans. In this regard, this paper aims to determine which of the two sample plans (simple random sample without replacement and stratified sample) is better in

estimating the mean of the human development index. Therefore, the research should determine the following:

1. Which sampling plan provides the mean closest to the value of the population mean;

2. Which sampling plan provides the lower value of the variance of the human development index.

We used data of the human development index for 2018, based on the Human Development Report covering 189 member states of the United Nations. Most empirical research dealing with similar topics has shown that stratification provides a more accurate estimate of mean compared to the estimate obtained by applying a simple random sample without replacement (Elsayir, 2014; Okororie & Otuonye, 2015; Wibowo, 2015; Koprivica, 2017; Wang et al., 2020). Therefore, the results of this research should contribute by expanding the literature related to examining the accuracy of estimates using the different sampling plans. Also, the importance of this research is reflected in providing guidelines for applying an adequate sampling plan, especially for research dealing with economic indicators.

## 2. LITERATURE REVIEW

The importance of research design is reflected in the fact that at the end of the nineteenth century, there was a discussion called the "representative method", which was related to determining the validity of different sampling methods. One of the leading proponents of using the sample was the Norwegian statistician Anders Nicolai Kiær, who empirically found that a stratified sampling can provide good estimates of the mean and total (Smith, 1976). Also, it is considered that stratification contributes to efficiency by taking into account the cost and accuracy of population parameter estimates (Parsons, 2005).

The potential to increase the accuracy of population parameter estimates using a stratified sampling (if the strata are homogeneous) has been pointed out by numerous authors (Pope, 1956; Jensen, 1991; Petrović, 2013; McRoberts, Tomppo & Czaplewski, 2015, Yadav & Tailor, 2020 ), as well as the US Environmental Protection Agency (EPA, 2002). Furthermore, if applied adequately, there are claims that stratification almost always provides lower variance for the estimated mean than in the case of a simple random sampling without replacement (Cochran, 1977). The characteristics of the stratified sample and the simple random sample without replacement are presented below.

The advantages of a stratified sampling over a simple random sampling without replacement are as follows (Freese, 1962):

1. It enables obtaining the mean value and variance of each stratum;

2. It often gives more precise estimates of population parameters if the sample is the same size, whereby it is necessary that within each stratum, the values of the features be approximate and that the values of the features of elements from different strata differ significantly.

The disadvantages of stratified sampling are that it is not useful to apply when homogeneous groups cannot be formed within the population (Berndt, 2020) and if the criterion characteristic used for classification is not chosen correctly (Alvi, 2016). In addition, there is a possibility that the strata may not be proportional in size (Berndt, 2020).

The advantages of a simple random sampling without replacement are:

1. It is a simple process (Berndt, 2020);

2. It enables avoidance of bias when selecting units from the sample (Alvi, 2016);

3. It allows obtaining a representative sample (Alvi, 2016), while the disadvantages of simple random sampling without replacement are the following:

   a. The process can be costly and time-consuming, especially in the case when the respondents are geographically widespread and difficult to access (Alvi, 2016);

   b. It requires much effort, especially in the case of a large population (Alvi, 2016);

   c. Sometimes, it is impossible to obtain or prepare a complete list of elements (Berndt, 2020).

A study conducted in Serbia revealed that stratification provides a more accurate estimate of mean in the case of the in forest inventory if the sample size is the same (Koprivica, 2017). Also, the study that assessed deforestation in Brazil revealed that a more precise estimate of the population total was obtained by applying a stratified sampling (Broich, Stehman, Hansen, Potapov & Shimabukuro 2009). On the other hand, the results of a study conducted in the Santa Cecília district of Brazil indicate that it is more appropriate to apply a simple random sampling without replacement for urban tree inventory estimation. The reason is that it provides better estimates of the total number of

trees and trees per kilometer of the sidewalk (Alvarez et al., 2005). The research results related to six geopolitical zones in Nigeria showed that stratified sampling is more efficient than simple random sampling and cluster sampling when estimating the mean using the criterion of minimum variance (Okororie & Otuonye, 2015). A lower value of the variance of sample mean was also obtained in a study that analyzed the production of three types of agricultural products in Sudan (Elsayir, 2014). A recent study comparing the estimates of the mean of the abundance indices of three macrozoobenthic species and species diversity index obtained using a simple random sampling, a stratified sampling, and a cluster sampling showed that the highest accuracy of the estimates of mean was provided by a stratified sampling (Wang et al., 2020). The accuracy of the estimate achieved by applying the stratified sampling was also confirmed by a study that analyzed the number of minimarkets in Indonesia (Wibowo, 2015). UK National Audit Office found that on a sample of 50 units at a confidence interval of 95%, a more accurate estimate of the mean was obtained using stratified sampling than simple random sampling without replacement (National audit office, n.d.). A similar conclusion was found in a study on a population that included 344 scientific journals in economics on the SCI list classified according to the impact factor in 2015 (Lojanica, 2017).

Considering all the above, the following research hypotheses are defined:

**H1$_0$:** A simple random sampling without replacement provides a more accurate estimate of the mean than the stratified sampling.

**H1$_a$:** *A simple random sampling without replacement does not provide a more accurate estimate of the mean than the stratified sampling.*

**H2$_0$:** The stratified sampling provides a more accurate estimate of the mean than a simple random sampling without replacement.

**H2$_a$:** *The stratified sampling does not provide a more accurate estimate of the mean than a simple random sampling without replacement.*

# 3. EMPIRICAL RESEARCH

## 3.1. DATA ANALYSIS OF POPULATION

The feature examined in this study is the Human Development Index (HDI). It is a composite index that measures progress in three basic life dimensions: long

and healthy life (life expectancy index), knowledge (education index) and a decent standard of living (Gross National Income index).[†]

The United Nations Development Program (UNDP), starting in 1990, prepares the Human Development Report (HDR) each year and ranks UN member states according to the value of the human development index (Bhanojirao, 1991). The index has been calculated as the arithmetic mean of the three components since 1990. However, the calculation was changed, so in the 2010 report, the index was calculated as the geometric mean (Gaye, 2011). Table 1 shows the four tiers and cutoff points based on which countries are grouped. Classification of countries based on HDI values was introduced in the 2014 Report.

Table 1

*Classification of countries based on their level of development*

| Level of human development | HDI values |
|---|---|
| Very high human development | ≥ 0.800 |
| High human development | 0.700 – 0.799 |
| Medium human development | 0.550 – 0.699 |
| Low human development | < 0.550 |

*Note.* Human Nations Development Programme, Retrieved from http://hdr.undp.org.

We used data on the human development index for 2018 from the Human Development Report.[‡] The survey covered 189 UN member states. The countries not included in the analysis (since data are not available) are primarily from Europe (Vatican, Monaco, San Marino) and Oceania (Nauru, Tonga, Tuvalu), while Somalia is the only African country for which the data were not available. Based on the available data, the mean and the variance of the human development index were calculated. After that, the mean and variance of the human development index were estimated using a simple random sampling without replacement and stratified sampling. Finally, we compared these two sampling plans.

Based on the Human Development Report data, population parameters were calculated for 189 countries (Table 2).

---

[†] https://www.rs.undp.org/content/serbia/sr/home/presscenter/articles/2019/izve_taj-o-ljudskom-razvoju-za-2019--svet-se-suoava-sa-novim-nej.htm (Accessed on April 28, 2020).
[‡] http://hdr.undp.org/en/data (Accessed on April 20, 2020).

For the purpose of the analysis, we used generally accepted formulas according to Petrović (2013) to calculate the values of population parameters, i.e., total, mean and variance.

Table 2

*Population parameters*

| Description | Value |
|---|---|
| Number of observations $(N)$ | 189 |
| **Mean** $(\bar{Y})$ | **0.713** |
| Total $(Y)$ | 134.842 |
| Maximum | 0.954 |
| Minimum | 0.377 |
| Variance $(S^2)$ | 0.023 |
| Standard deviation $(S)$ | 0.151 |

*Note.* Author's calculation.

The calculated values of the parameters indicate that the average value of the human development index in 2018 was 0.713. Norway had the highest value of the human development index, while Niger had the lowest value.

### 3.2. A SAMPLE RANDOM SAMPLING WITHOUT REPLACEMENT

The most straightforward sample design is simple random sampling. Depending on selecting elements from the population, a simple random sampling without replacement and a simple random sampling with replacement are distinguished. When a simple random sample is chosen without replacement, each element of the population has the same probability of being selected into the sample, whereby there is no possibility for a unit that has already been chosen to re-enter the selection process (Petrović, 2013). The sample of 40 units was selected using Excel 2010 and the RAND, LARGE and VLOOKUP functions. The following formulas, according to Petrović (2013), were used to calculate estimates of the mean and variance:

Sampling fraction: $f = \dfrac{n}{N}$ (1)

Total: $y = \displaystyle\sum_{i=1}^{n} y_i$ (2)

Mean: $\bar{y} = \dfrac{1}{n} \sum_{i=1}^{n} y_i$ (3)

Variance: $s^2 = \dfrac{1}{(n-1)} \sum_{i=1}^{n} \left(y_i - \bar{y}\right)^2$ (4)

Variance of the sample mean:

$$V\left(\bar{y}\right) = \dfrac{S^2}{n}\left(1 - f\right)$$ (5)

Estimate of variance of the sample mean:

$$\hat{V}\left(\bar{y}\right) = \dfrac{s^2}{n}\left(1 - f\right)$$ (6)

The sampling fraction is 0.21 since $n = 40$, while N=189.

Estimates of population parameters obtained by a simple random sampling without replacement are presented in Table 3.

Table 3

*Statistics of the simple random sampling without replacement*

| Description | Value |
|---|---:|
| Sample size $\left(n\right)$ | 40 |
| **Mean** $\left(\bar{y}\right)$ | **0.700575** |
| Total $\left(y\right)$ | 28.023 |
| Maximum | 0.930 |
| Minimum | 0.434 |
| $n$/N | 0.211640212 |
| $s^2$ | 0.020488758 |
| $v\left(\bar{y}\right)$ | 0.000448209 |
| $\hat{V}\left(\bar{y}\right)$ | 0.000403813 |
| $\sqrt{\hat{V}\left(\bar{y}\right)}$ | 0.020095095 |

*Note.* Author's calculation.

The average value of the human development index obtained based on a simple random sampling without replacement is 0.700575. In 2018, Denmark had the highest and Eritrea the lowest value of the human development index. In order to check the accuracy of the estimate of the mean, it is necessary to determine the confidence interval, which is claimed with some certainty to contain values for the population. In this paper, a 95% confidence interval is applied, which means that α = 0.05. Since the size of a simple random sample without replacement is 40, a 95% confidence interval based on the normal distribution can be used according to the Central Limit Theorem. Table 4 presents the procedure for calculating the confidence interval.

Table 4

*The procedure for calculating the confidence interval*

| Description | Value |
|---|---|
| $\overline{y}$ | 0.700575 |
| $z$ | 1.96 |
| $\sqrt{\hat{V}\left(\overline{y}\right)}$ | 0.020095 095 |
| $\overline{y} - z\sqrt{\hat{V}\left(\overline{y}\right)}$ | **0.661** |
| $\overline{y} + z\sqrt{\hat{V}\left(\overline{y}\right)}$ | **0.740** |

*Note.* Author's calculation.

The calculated interval (0.661; 0.740) is claimed to include the population mean value. The obtained values will be compared with the values for the stratified sampling to assess the accuracy of the estimates of the mean of the human development index.

### 3.3. STRATIFIED SAMPLE

Stratification aims to divide the population into several parts, i.e., strata. Strata should be different from each other, while the elements of each stratum should be homogeneous. In this paper, the strata are formed based on the UNDP classification. Thus, according to the mentioned classification, the following four strata were developed (Table 5).

Table 5

*Population division into strata according to the value of human development index*

| Stratum | Stratum description | Index criteria/ value index | Number of countries/size of stratum ($N_h$) |
|---|---|---|---|
| **1** | Very high human development | HDI $\geq$ 0.8000 | 62 |
| **2** | High human development | $0.700 \leq$ HDI $\leq 0.799$ | 54 |
| **3** | Medium human development | $0.550 \leq$ HDI $\leq 0.699$ | 37 |
| **4** | Low human development | HDI $< 0.550$ | 36 |
| **Total** | | | **189** |

*Note.* Author's calculation.

The results analysis showed that as much as a third of the population had a very high human development index. In contrast, about 29% of countries had a high level. The data for each stratum are presented in Table 6.

Table 6

*Basic data of strata*

| Description / Stratum | 1 | 2 | 3 | 4 |
|---|---|---|---|---|
| Stratum size $\left(N_h\right)$ | 62 | 54 | 37 | 36 |
| Mean value $\left(\overline{Y}_h\right)$ | 0.875 | 0.746 | 0.622 | 0.480 |
| Stratum total $\left(Y_h\right)$ | 54.268 | 40.263 | 23.018 | 17.293 |
| Maximum value | 0.954 | 0.799 | 0.698 | 0.549 |
| Minimum value | 0.801 | 0.700 | 0.557 | 0.377 |
| Variance in stratum $h$ $\left(S_h^2\right)$ | 0.002 | 0.001 | 0.002 | 0.002 |
| Standard deviation of $h$ stratum $\left(S_h\right)$ | 0.047 | 0.031 | 0.042 | 0.048 |
| Stratum weight $\left(W_h\right)$ | 0.328 | 0.286 | 0.196 | 0.190 |

*Note.* Author's calculation.

The division into strata is performed correctly if the variance between the strata is greater than the variance within the stratum. According to the data presented in Table 7, it can be concluded that this condition was fulfilled, which was the aim of stratification.

Table 7

*Population variance, variance within the stratum and variance between the strata*

| Description / Stratum | 1 | 2 | 3 | 4 |
|---|---|---|---|---|
| Variance $\left(S_h^2\right)$ | 0.002 | 0.001 | 0.002 | 0.002 |
| $(N_h-1)S_h^2$ | 0.137 | 0.050 | 0.063 | 0.081 |
| $\sum(N_h-1)S_h^2$ | 0.3310065 | | | |
| $1/(N-1)$ | 0.0053191 | | | |
| **Variance within the stratum** $\left(S_u^2\right)$ | **0.00176** | | | |
| $\left(\overline{Y}_h-\overline{Y}\right)$ | 0.162 | 0.032 | -0.091 | -0.233 |
| $\left(\overline{Y}_h-\overline{Y}\right)^2$ | 0.026 | 0.001 | 0.008 | 0.054 |
| $\sum N_h\left(\overline{Y}_h-\overline{Y}\right)$ | 3.943809 | | | |
| **Variance between the strata** $\left(S_i^2\right)$ | **0.02098** | | | |
| **Variance of stratified population** $\left(S^2\right)$ | **0.02274** | | | |

*Note.* Author's calculation.

The presented results show that the stratification was correctly performed so that the units within the stratum are similar and the strata are different. Also, the sum of the variance within the stratum and the variance between the strata is equal to the total variance of the stratified population. After correct stratification, the determination of the sample size for each stratum was performed (Table 8).

Table 8

*Sample size for each stratum*

| Description/Stratum | 1 | 2 | 3 | 4 |
|---|---|---|---|---|
| **Sample size** $\left(n_h\right)$ | 15 | 8 | 8 | 9 |

*Note.* Author's calculation.

The following formula was used to calculate the sample size:

$$\left( \frac{n}{\sum\limits_{h=1}^{4} W_h S_h} W_h S_h \right) \tag{7}$$

where $n = 40$.

Thus, the division into strata and the sample size of each stratum can be graphically presented as follows (Figure 1), where $N_h$ ( $x = 1, ..., 4$) represents the size of the stratum, while $_{NH}$ ($h = 1, ..., 4$) represents the sample size of the stratum.
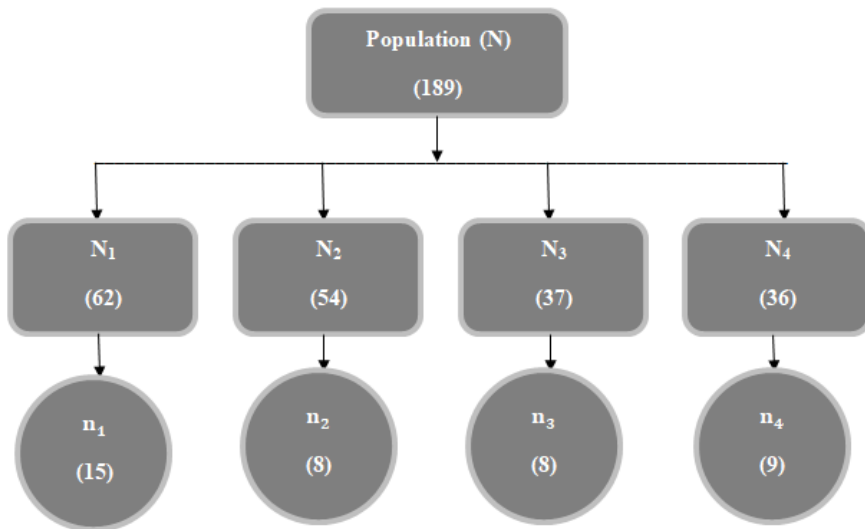


*Figure 1.* The division into strata and sample size by strata

*Note.* Made by authors.

After determining the sample size for each stratum, the units were selected using Excel 2010 and the RAND, LARGE, and VLOOKUP functions. Sample data for each stratum are presented in Table 9.

Table 9

*Sample data for the strata*

| Description / Stratum | 1 | 2 | 3 | 4 |
|---|---|---|---|---|
| The sample size of the stratum $(n_h)$ | 15 | 8 | 8 | 9 |
| Mean value of the sample of the stratum $(\bar{y}_h)$ | 0.880 | 0.759 | 0.612 | 0.473 |
| Total of the sample of the stratum $(y_h)$ | 13.198 | 6.075 | 4.896 | 4.256 |
| Maximum value | 0.946 | 0.799 | 0.693 | 0.528 |
| Minimum value | 0.808 | 0.708 | 0.563 | 0.377 |
| Sample variance of the stratum $(s_h{}^2)$ | 0.002 | 0.011 | 0.002 | 0.003 |

*Note.* Author's calculation.

The statistics of the stratified sample are presented below (Table 10).

Table 10

*Statistics of the stratified sample*

| Description | Value |
|---|---|
| Sample size $(n)$ | 40 |
| $\bar{y}_{st}$ | **0.71545** |
| Maximum | 0.946 |
| Minimum | 0.377 |
| $V(\bar{y}_{st})$ | 0.000033990 |
| $\hat{V}(\bar{y}_{st})$ | 0.000038709 |
| $\sqrt{\hat{V}(\bar{y}_{st})}$ | 0.006221683 |

*Note.* Author's calculation.

According to the data presented in Table 10, an estimate of the mean of the human development index has the value of 0.71545. In 2018, Switzerland had the maximum value of the human development index, while Niger had the minimum value of the index. A confidence interval was formed to determine the accuracy of the obtained estimate of the mean of the human development index. The average value of the human development index of the population must be within this interval. A 95% confidence interval ($\alpha = 0.05$) was used for

the analysis. The approximate 95% confidence interval for the population mean is (0.703; 0.728). The procedure for calculating the confidence interval is presented in Table 11.

Table 11

*The procedure for calculating the confidence interval*

| Description | Value |
|---|---|
| $\overline{y}_{st}$ | 0.71545 |
| $Z$ | 1.96 |
| $\sqrt{\hat{V}\left(\overline{y}_{st}\right)}$ | 0.006221683 |
| $\overline{y}_{st} - z\sqrt{\hat{V}\left(\overline{y}_{st}\right)}$ | **0.703** |
| $\overline{y}_{st} + z\sqrt{\hat{V}\left(\overline{y}_{st}\right)}$ | **0.728** |

*Note.* Author's calculation.

As can be seen, the estimate of the mean value is close to the value of the population mean of the human development index. To compare the values and determine which sampling plan is better, we compared the estimates of the mean of a simple random sampling without replacement and stratified sampling.

### 3.4. COMPARISON OF ESTIMATES OF MEAN: SAMPLE RANDOM SAMPLING WITHOUT REPLACEMENT VS. STRATIFIED SAMPLING

According to the data presented in Table 12, it can be concluded that the estimate of the mean obtained by stratification is more accurate since it is closer to the average value of the human development index of the population.

Table 12

*Comparison of a simple random sampling without replacement and a stratified sampling*

| Description | Population | Simple random sampling without replacement | Stratified sampling |
|---|---|---|---|
| Mean | 0.713 | 0.700575 | 0.71545 |
| Variance of sample mean | | 0.000448209 | 0.000033990 |
| Confidence interval | | (0.661;0.740) | (0.703;0.728) |

*Note.* Author's calculation.

The results can be graphically presented as follows (Figure 2), using the following symbols:

■ - $\overline{y}$

▼ - $\overline{Y}$

● - $\overline{y}_{st}$

—— the confidence interval for the mean - simple random sampling without replacement

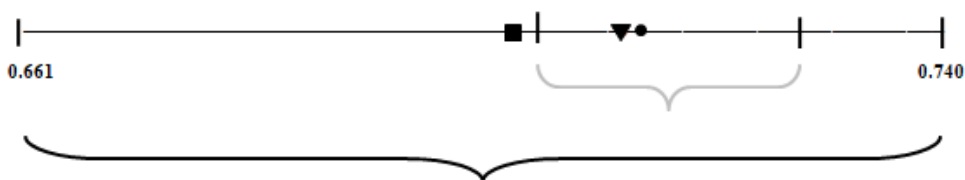—— the confidence interval for the mean-stratified sampling



*Figure 2.* Comparison of the two sample plans

*Note.* Author's calculation.

Stratification provides a confidence interval for the mean value narrower than the interval obtained based on a simple random sampling without replacement. In addition, the value of the variance of the mean of the human development index in the case of stratified sampling is less than that obtained from a simple random sampling without replacement. Considering the research results, we can conclude that the zero hypothesis is rejected within the first

group of hypotheses, and the alternative hypothesis $H1_a$ is accepted. This means that stratified sampling provides a more accurate estimate of the mean than simple random sampling without replacement. Within the second group of defined hypotheses, the zero hypothesis $H2_0$ is accepted. Therefore, the alternative hypothesis $H2_a$ is rejected, which implies that the stratified sampling provides a more accurate estimate of the mean of the human development index than a simple random sampling without replacement. Considering that no research has been conducted on the human development index, we compared our results with studies on similar topics. The results of this study are in line with the results obtained in several studies that compared estimates of the mean of simple random sampling without replacement and stratified sampling (Elsayir, 2014; Okororie & Otuonye, 2015; Wibowo, 2015; Koprivica, 2017; Wang et al., 2020), while in contrast to the results of the study by Alvarez et al. (2005).

Further research should focus on examining and comparing estimates of the mean of other sampling plans.

## 4. CONCLUSION

To conduct quality research, the selection of an adequate sampling plan is essential. Some studies have shown that different sampling plans provide estimates of different precision (Alvarez et al., 2005; Elsayir, 2014; Okororie & Otuonye, 2015; Wibowo, 2015; Koprivica, 2017; Wang et al., 2020). The authors of this paper achieved a defined objective with regard to the fact that they determined which of the two sampling plans (simple random sample without repetition and stratified sample) is better in terms of accuracy of estimates of the mean of human development index. Based on the results of this research, it can be determined which sampling plan provides the closest value of the sample mean of the human development index to the population mean. Also, it was determined which sampling plan provides the lower value of the variance of the mean of the human development index. The authors used data on the human development index for 2018 from the Human Development Report. The survey covered 189 member states of the United Nations. The research results show that a more precise estimate of the mean of the human development index is obtained by stratification. Opposed to a simple random sampling without replacement, the value of the mean of the human development index obtained by stratification is closer to the population mean. Also, the stratification determined a narrower confidence interval and a significantly lower variance of the sample mean.
All of the above indicates the advantage of the stratification process. However,

research related to comparing estimates of the mean of the human development index does not exist. Therefore, the results of this research can be compared with the results of research with a similar topic. In this regard, the results of this study are consistent with the results of several studies that compared the estimate of the mean obtained using a simple random sampling without replacement and stratified sampling (Elsayir, 2014; Okororie & Otuonye, 2015; Wibowo, 2015; Koprivica, 2017; Wang et al., 2020), while in contrast to the results of the study Alvarez et al. (2005).

Based on all of the above, when considering the choice of the sampling plan in economic analyses, using the stratified sampling should be taken into account, considering that it provides a more precise estimate of the mean than simple random sampling without replacement.

## REFERENCES

Alvarez, I. A., del Nero Velasco, G., Barbin, H. S., Lima, A. M. L. P., & Do Couto, H. Z. (2005). Comparison of two sampling methods for estimating urban tree density. *Journal of Arboriculture*, *31*(5), 209-2014.

Alvi, M. H. (2016). *A Manual for Selecting Sampling Techniques in Research.* Retrieved from https://mpra.ub.uni-muenchen.de

Berndt, A. E. (2020). *Sampling methods. Journal of Human Lactation*, *36*(2), 224-226.

Bhanojirao, V. V. (1991). Human development report 1990: review and assessment. *World Development*, *19*(10), 1451-1460.

Broich, M., Stehman, S. V., Hansen, M. C., Potapov, P., & Shimabukuro, Y. E. (2009). A comparison of sampling designs for estimating deforestation from Landsat imagery: A case study of the Brazilian Legal Amazon. *Remote Sensing of Environment*, *113*(11), 2448-2454.

Cochran, G. W. (1977). *Sampling Techniques*. New York: John Wiley & Sons.

Elsayir, H. A. (2014). Comparison of precision of systematic sampling with some other probability samplings. *American Journal of Theoretical and Applied Statistics*, *3*(4), 111-116.

Freese, F. (1962). *Elementary forest sampling (No. 232)*. US Department of Agriculture, Forest Service.

Gaye, A. (2011). *Contribution to Beyond Gross Domestic Product (GDP.)* Retrieved from https://ec.europa.eu

Jensen, A. L. (1991). Relative precision of stratified sampling, sampling with probability of selection proportional to size, and simple random sampling with ratio estimation. *Environmetrics*, *2*(4), 475-486.

Koprivica, M. (2017). Stratified sampling in forest inventory. *Šumarstvo*, *3*(4), 55-68.

Lojanica, N. (2017). Preciznost ocena prostog i stratifikovanog slučajnog uzorka na tržištu naučnih časopisa. *Škola biznisa*, 1/2017, 1-12.

McRoberts, R. E., Tomppo, E. O., & Czaplewski, R. L. (2015). *Sampling designs for national forest assessments. Knowledge Reference for National Forest Assessments*; FAO: Rome, Italy, 23-40.

National audit office, (n.d.). A Practical Guide to Sampling. Retrieved from https://www.nao.org.uk

Okororie C., & Otuonye E. (2015). Efficiency of some sampling techniques. *Journal of Scientific Research and Studies*, *2*(3), 63-69.

Parsons, V. (2005). *Stratified Sampling. Encyclopedia of Biostatistics*. John Wiley & Sons.

Petrović, Lj. (2013). *Teorija uzoraka i planiranje eksperimenata*. Beograd: Centar za izdavačku delatnost Ekonomskog fakulteta.

Pope, J. A. (1956). An outline of sampling techniques. An outline of sampling techniques. *Rapports et Proces-Verbaux des Reunions Consell International pour l'Exploration de la Mer*, *1*(140), 11-20.

Smith, T. M. F. (1976). The foundations of survey sampling: a review. *Journal of the Royal Statistical Society: Series A (General)*, *139*(2), 183-195.

UNDP. Data. (2019). Retrieved from https://www.rs.undp.org

UNDP. Human Development Data. (2018). Retrieved from http://hdr.undp.org

UNDP. Technical notes - Calculating the human development indices—graphical presentation. Retrieved from http://hdr.undp.org

United States Environmental Protection Agency. Guidance on Choosing a Sampling Design for Environmental Data Collection for Use in Developing a Quality Assurance Project Plan. Office of Environmental Information Washington, DC 20460. Retrieved from https://www.epa.gov

Wang, J., Zhang, C., Xue, Y., Xu, B., Ren, Y., & Chen, Y. (2020). Comparison of different sampling designs for macrozoobenthos survey in a tidal flat. *Regional Studies in Marine Science*, 35, 101-113.

Wibowo, A., Indahwati, I., & Sumertajaya, M (2015). Accuracy comparison of simple, systematic, and stratified random sampling for estimating population (minimarket case in Indonesia). In *Proceedings of International Conference On Research, Implementation And Education Of Mathematics And Sciences 2015* (17-19). Yogyakarta State University.

Yadav, R., & Tailor, R. (2020). Estimation of finite population mean using two auxiliary variables under stratified random sampling. *Statistics in Transition New Series*, *21*(1), 1-12.