Mathematics and Society
University of Isfahan
https://math-sci.ui.ac.ir

Mathematics and Society
ISSN (print): 2345-6493, ISSN (on-line): 2345-6507
Vol. 11 No. 1 (2026), pp. 109–131.
© 2026 The Author(s)
https://www.ui.ac.ir

EVALUATION OF THE MEASUREMENT ERROR OF IRANIAN LABOR FORCE INDICATORS DUE TO CHANGING DATA GATHERING MODE BASED ON THE MARKOV LATENT CLASS MODEL

LIDA KALHORI NADRABADI* AND PARIA TORABI KAHLAN

Abstract. One of the most important non-sampling errors is measurement error, which causes the measured value of the target variable to be different from its true value. The markov latent class model allows for the estimation of measurement error for categorical variables in panel surveys that have at least three repetition periods, with the advantage that it does not require re-interviews. In the Iranian labor force survey, which is conducted with rotational sampling panel with a 2-2-2 pattern, this model can be applied. In this article, the challenges of changing the survey mode in times of crisis and the advantages of telephone interviews compared to face-to-face interviews are reviewed. Also, in order to evaluate the results of the labor force survey in the context of the COVID-19 pandemic, when the survey mode was suddenly changed from face-to-face to telephone interviews, the markov latent class model was used to evaluate the measurement error using data from four survey periods from spring 2018 to summer 2021. The results indicate that the measurement error increased in the first period of changing the interview mode, but decreased in subsequent periods. Overall, it can be stated that telephone interviews can be a suitable method for collecting labor force survey data.

Keywords: Markov latent class model, Labor force survey, Measurement error

Article Type: Applied Paper.

Communicated by Majid Asadi.

*Corresponding author.

Received: 09-10-2024, Accepted: 20-05-2025, Published Online: 27-10-2025.

Cite this article: L. Kalhori Nadrabadi and P. Torabi Kahlan, Evaluation of the measurement error of Iranian labor force indicators due to changing data gathering mode based on the Markov latent class model, *Mathematics and Society*, **11** no. 1 (2026) 109–131. https://dx.doi.org/10.22108/msci.2025.142988.1700 .

1. Introduction

Iran's official statistics system, like other countries, is based on collecting data through filling out questionnaires by asking from sample households. Registered sources can be used is some cases as well. One of the most important challenges in this system is statistical error, especially measurement error, which has several sources. One of these source is respondent due to various reasons such as forgetting the real response, misunderstanding the question, not having enough information, or for intentional reasons.Therefore, accurate identification of error sources is required to control the survey error. The Labor Force Survey is one of the main surveys of the Statistical Center of Iran, which has been conducted quarterly and with a rotational sampling method since 2005.

The COVID-19 pandemic presented unprecedented challenges to national labor force surveys (LFS), as documented by the International Labour Organization (ILO, 2020). Given the variety of country contexts, appropriate solutions must ultimately be assessed at the country level and all challenges faced cannot be fully anticipated or prescribed. In addition, there may not be an ideal solution or perfect fit for the many challenges faced any given time. However, various nonetheless options can be considered to reduce the impact on generated statistics. Although there are many different sets of circumstances across countries, the main challenges facing most countries can be broadly categorized into two scenarios.

Scenario 1: Transitioning from fully face-to-face interviewing to telephone based interviewing. This scenario has been evaluated or implemented by many countries, some of whom have panel designs. Others have independent samples for each survey instance. These countries rely either on Computer Assisted Personal Interviewing (CAPI) or Paper and Pencil Interviewing (PAPI).

Scenario 2: Multi-mode data collection is being used (e.g. CAPI plus Computer Assisted Telephone Interviewing (CATI)). In these cases face-to-face CAPI is generally used for the first interview of a panel design, plus a limited number of households for subsequent interviews. The challenge is to move these interviews to either telephone or Computer Assisted Web Interviewing (CAWI). Either scenario 1 or 2 (or others) could involve temporary disruptions (short or long) in data collection, which in reality creates a number of variations within each scenario. Under each scenario there will be proposals to guide countries on how to assess the validity/robustness of different approaches, as this needs to be an important consideration regardless of scenario or approach to change. The range of challenges and responses reported by countries attempting to maintain the operation of their LFS is very wide. The severity of these challenges also varies heavily. Countries fully utilize CATI/CAWI generally facing minimal disruption to data collection (although potentially still some). By contrast, some countries losed their capacity to collect data through existing methods entirely. The challenge facing these countries cannot be underestimated and it is feasible, if not likely, that data collection may need to be suspended, either briefly or for a longer period of time, and indeed this has already happened in many

https://dx.doi.org/10.22108/msci.2025.142988.1700

countries. Scenario 1 reported by a large number of countries, whereby restrictions on movement have caused all face-to-face interviews to be halted, and face-to-face interviewing was the only method in use (CAPI or PAPI). Under this scenario there can be different periodicities (quarterly continuous collection, quarterly periodic, annual continuous, annual periodic, etc.). Some of the countries in this situation utilized independent samples for each survey instance, while others used a panel design. The challenges were mostly the same but those with a panel design were able to potentially benefit from it in different ways. It recommended to consider ways to begin the move to telephone interviewing as it is less disrupted in case of restrictions of movement. The main challenges faced under this scenario are the following: a) retrieving contact details for the final sampling units (FSUs – generally households); b) enabling capacity to carry out interviews by phone; and c) assessing the suitability of PAPI/CAPI questionnaire for telephone interviews. Scenario 2 covers many countries in Europe and selected countries in other parts of the world. In almost all cases there is a panel design. The most typical scenario is a face-to-face interview for the first interview, with CATI (and possibly CAWI) for subsequent interviews. Some face-to-face interviewing will typically continue to take place for the subsequent interviews, but limited to those for whom CATI or CAWI are not possible, for example households with no phone, or respondents who state a preference for a face-to-face interview. As with those countries under scenario 1, the immediate response for the majority of these countries is to move any remaining face-to-face interviewing to one of the other modes, mostly CATI. Essentially the range of challenges and options is broadly be similar to those outlined for scenario 1, but some specific issues still warrant attention.

In Iran, as in many countries, the outbreak of COVID-19 caused the data collection method to change suddenly from face-to-face interviews to telephone interviews, which can affect data quality and typical error rates. In order to estimate the measurement error, gold-standard method, reliability analysis, latent class analysis, and markov latent class analysis are mainly used. Gold standard method measures the bias between a measurement with error and a measurement without error (the gold standard). The disadvantage of the gold standard method is that this value rarely exists, and in cases where it does, there may be latent (unobserved) error in the gold standard estimate, which can lead to invalid results. Reliability analysis decomposes the variance to determine the portion of the variance in the estimate that is due to measurement error. Latent class analysis uses the maximum likelihood method for estimates to determine the bias between the survey and maximum likelihood estimates.

Markov latent class models can be a good alternative for costly and time consuming methods such as re-interview, so as a result of which improvement in error estimation methodology and data quality are obtained In this article, markov latent class model ia applied as an alternative to the re-interview method, to estimate of the measurement error of the Iranian labor force survey which was conducted differently in aspect of data gathering mode under the conditions of the COVID-19 pandemic.

2. Main Results

With the sudden onset of the Covid-19 pandemic, the mode of collecting data for the Iranian Labor Force Survey suddenly changed from face-to-face interviews to telephone interviews due to health restrictions. This change could have affected the accuracy and quality of the data and increased the possibility of measurement errors. The markov latent class model is a statistical tool that allows for the estimation and analysis of measurement error for longitudinal categorical variables with at least three time periods. Employment status in laboure force survey is a categorical varible which is contain 3 level as Employed(e), Unemployed(U) and Not in the laboue force(N). The survey is conducted with a rotating panel design and the response variable is categorical, thus providing the conditions for using the hidden Markov class model. To apply the markov latent class model, data of four survey rounds from spring 2018 to summer 2020 are considered. Independence of classification error, time homogeneity of classification error for different time periods, and homogeneous error probabilities or different time periods are assumed as well as the Markov property $P(X_t|X_{t-1}) = P(X_{t+1}|X_t)$.

Based on the 2-2-2 rotation pattern of the survey, sample units are included in the survey for two consecutive seasons, excluded from the survey for two seasons, and then included again in the survey for two consecutive seasons. Therefore, in order to satisfy the assumption of three time periods of data, which is necessary to implement the Markov model an additional assumption is made in the model. We suppose that no changes in the activity status occurs between the second and third periods. Therefore, those observations without changes in activity status from the second to third quarters are considered in the study. Accordingly, the average difference between the expected probability of stability and the observed probability of no change in the activity status between two consecutive quarters was considered as a criterion for assessing the amount of measurement error due to the change in the survey mode. Given that using weighted data does not have a significant effect on relative estimates and complicates calculations, unweighted data has been used.

In this study, data from four labor force survey periods (from spring 2018 to summer 2021) were analyzed with this model to examine the effect of changing the data collection mode on measurement error. Table 1 shows that in fact 96.2 percent of those employed in the spring of 2018 remained employed in the following quarter. Similarly, 70.2 percent of those unemployed in the spring of 2018 remained unemployed in the following quarter.

TABLE 1. Matrix of transition probabilities without classification error from spring to summer 2018

		Summer 2018		
		E	U	N
Spring 2018	E	0.962	0.007	0.032
	U	0.204	0.702	0.095
	N	0.023	0.005	0.973

https://dx.doi.org/10.22108/msci.2025.142988.1700

TABLE 2. Measurement error matrix of labor force status from spring to summer 2018

		Summer 2018		
		E	U	N
Spring 2018	E	0.984	0.006	0.066
	U	0.079	0.876	0.045
	N	0.009	0.003	0.988

According to Table 2, a person who is actually employed in the spring of 2018 has a 0.6 percent probability of being falsely reported as unemployed. There is also a 1.2 percent probability of being falsely reported as not labor force. Similarly, a person who is actually unemployed has a 7.9 percent probability of being falsely reported as employed and a 4.5 percent probability of being falsely reported as not labor force.

TABLE 3. Matrix of transition probabilities without classification error from spring to summer 2019

		Summer 2019		
		E	U	N
Spring 2019	E	0.950	0.009	0.041
	U	0.191	0.671	0.138
	N	0.016	0.004	0.979

Table 3 shows that in fact 95 percent of those employed in the spring of 2019 remained employed in the following quarter. Similarly, 67.1 percent of those unemployed in the spring of 2019 remained unemployed in the following quarter.

TABLE 4. Measurement error matrix of labor force status from spring to summer 2019

		Summer 2019		
		E	U	N
Spring 2019	E	0.983	0.008	0.008
	U	0.060	0.857	0.083
	N	0.007	0.002	0.991

According to Table 4, a person who is actually employed in the spring of 2019 has a 0.8 percent probability of being falsely reported as unemployed. There is also a 0.8 percent probability of being falsely reported as not labor force. Similarly, a person who is actually unemployed has a 0.6 percent probability of being falsely reported as employed and an 8.3 percent probability of being falsely reported as not labor force.

TABLE 5. Matrix of transition probabilities without classification error from spring to summer 2020

		Summer 2020		
		E	U	N
Spring 2020	E	0.958	0.006	0.036
	U	0.243	0.653	0.104
	N	0.021	0.006	0.973

https://dx.doi.org/10.22108/msci.2025.142988.1700

Table 5 shows that in fact 95.8 percent of those employed in the spring of 2020 remained employed in the following quarter. Similarly, 65.3 percent of those unemployed in the spring of 2020 remained unemployed in the following quarter.

TABLE 6. Measurement error matrix of labor force status from spring to summer 2020

		Summer 2020		
		E	U	N
Spring 2020	E	0.984	0.006	0.01
	U	0.061	0.891	0.047
	N	0.005	0.002	0.993

According to Table 6, a person who is actually employed in the spring of 2020 has a 0.6 percent probability of being falsely reported as unemployed. There is also a 1 percent probability of being falsely reported as not labor force. Similarly, a person who is actually unemployed has a 6.1 percent probability of being falsely reported as employed and a 4.7 percent probability of being falsely reported as not labor force. The total measurement error estimate for the period from spring 2018 to summer 2018, which is before the start of the COVID-19 pandemic, is 0.153, for the period from spring 2019 to summer 2019, which coincides with the outbreak of the COVID-19 pandemic and the sudden change in the survey mode, it is 0.168, and for the period from spring 2020 to summer 2019, it is 0.131. Since the spring of 2020, with the decrease in the prevalence of Covid-19 and nationwide vaccination due to the easing of restrictions, the survey was conducted in the previous method and face-to-face, therefore, the examination of the amount of measurement error due to the change in the survey mode from face-to-face interviews to telephone interviews is not relevant. Therefore, it seems that with the change in the survey mode, the measurement error initially increased, but with the continued implementation of the new mode, the measurement error decreased even compared to when the survey was only conducted face-to-face. Overall, the markov latent class model was introduced as an efficient and low-cost alternative to traditional error assessment methods (such as re-interview) and showed that telephone interviews can be a suitable option for collecting labor force data in crisis situations.

3. Conclusions

Sudden changes in the mode of data collection without the necessary opportunity to test and evaluate them can have significant effects on the quality of the survey results. One of the most important non-sampling errors is measurement error, which represents the difference between the response received from the respondent and the actual value of the under study variable. Obviously, due to the nature of this type of error and the lack of access to true values of responses, its calculation is difficult. In some special cases where administrative data are accessible or conditions of repeating part of the survey with very high accuracy is feasible in practice, the measurement error can be estimated. As an alternative the markov latent class model allows for the estimation of measurement error for

https://dx.doi.org/10.22108/msci.2025.142988.1700

categorical variables in panel surveys that have at least three repetition periods without the need for re-interview. Therefore in the Iranian labor force survey, which is conducted with rotational sampling with a 2-2-2 pattern, the markov latent class model is applicable to estimate the survey measurement error. In this article, in order to evaluate the results of the labor force survey during the Covid-19 pandemic, when the mode of conducting the survey was changed suddenly from face-to-face to telephone interviews or a combination of the two, the markov latent class model was used to evaluate the measurement error using data from four survey periods from spring 2018 to summer 2021. The results indicated that the measurement error increased in the first period of the change in the interview mode but decreased in the subsequent periods. In addition, it was observed that the share of employed and unemployed people joining the not in the labor force population was significant at the beginning of the COVID-19 pandemic, but it is interesting that as the epidemic has prolonged, more employed people have preferred to stay employed rather than become unemployed or not in the labor force and take the risk of looking for a better job. The share of unemployed people in the inactive population has also decreased, meaning that economic hardship and the state of the labor force are returning to pre-COVID-19. Based on the results of this study, it can be stated that telephone interviews, in addition to reducing costs, provide more accurate monitoring tools and can be a suitable alternative to face-to-face interviews for collecting labor force survey data. Using the telephone as a data collection tool allows interviewers to get access to employed individuals in order to fill out the questionnaire, thus reducing proxy response rates.

It should be noted that telephone surveys are methodologically different from surveys in which only interviews are conducted by telephone, and it is necessary to provide the necessary infrastructure for designing and implementing telephone surveys. For telephone surveys to be successful, attention must be paid to telephone coverage of the target population and appropriate sample design to prevent bias. The results show that, by observing the technical and methodological requirements, telephone surveys can maintain the quality of labor force data and be a reliable solution in critical situations. Also, the use of the markov latent class model is recommended for assessing and controlling measurement error.

Lida Kalhori Nadrabadi
Department of Technical Designs and Statistical Methods, Statistical Research and Training Center, Tehran, Iran
Email: Kalhori@srtc.ac.ir

Paria Torabi Kahlan
Department of Department of Technical Designs and Statistical Methods, Statistical Research and Training Center, Tehran, Iran
Email: torabiparya@yahoo.com

ریاضی و جامعه
شاپا (چاپی): ۲۳۴۵-۶۴۹۳، شاپا(الکترونیکی): ۲۳۴۵-۶۵۰۷
جلد ۱۱، شماره ۱ (۱۴۰۵)، صص. ۱۰۹-۱۳۱
© ۱۴۰۵ دانشگاه اصفهان

https://www.ui.ac.ir
https://math-sci.ui.ac.ir

ارزیابی خطای اندازه‌گیری شاخص‌های نیروی کار ایران با تغییر شیوه‌ی آمارگیری بر اساس مدل رده‌ی نهان مارکوف

لیدا کلهری ندرآبادی* و پریا ترابی کهلان

چکیده. یکی از مهم‌ترین خطاهای غیرنمونه‌گیری، خطای اندازه‌گیری است که سبب می‌شود مقدار اندازه‌گیری‌شده‌ی متغیر مورد مطالعه برای یک واحد نمونه متفاوت با مقدار واقعی آن متغیر باشد. مدل رده‌ی نهان مارکوف امکان برآورد خطای اندازه‌گیری برای متغیرهای رسته‌ای در آمارگیری‌های پانلی که حداقل سه دوره تکرار دارند را بدون نیاز به مصاحبه‌ی مجدد فراهم می‌کند. در آمارگیری نیروی کار ایران که با نمونه‌گیری چرخشی با الگوی ۲-۲-۲ اجرا می‌شود، امکان استفاده از این مدل فراهم است. در این مقاله، به مرور چالش‌های تغییر شیوه‌ی آمارگیری در زمان بحران و مزایای مصاحبه‌ی تلفنی در مقایسه با مصاحبه‌ی رودررو پرداخته می‌شود. همچنین به‌منظور ارزیابی نتایج آمارگیری نیروی کار در شرایط همه‌گیری کووید ۱۹ که شیوه‌ی اجرای آمارگیری به یک‌باره از مصاحبه‌ی رودررو به مصاحبه‌ی تلفنی تغییر یافت، مدل رده‌ی نهان مارکوف برای ارزیابی خطای اندازه‌گیری با استفاده از داده‌های چهار دوره‌ی آمارگیری از بهار ۱۳۹۷ تا تابستان ۱۴۰۰ به کار گرفته شده است. نتایج حاکی از آن است که خطای اندازه‌گیری در دوره‌ی اول تغییر شیوه‌ی مصاحبه با افزایش همراه بوده ولی در دوره‌های بعدی کاهش یافته است و در مجموع می‌توان اظهار داشت که مصاحبه‌ی تلفنی می‌تواند شیوه‌ی مناسبی برای گردآوری داده‌های آمارگیری نیروی کار باشد.

۱. مقدمه

نظام آمار رسمی ایران مثل سایر کشورهای جهان بر اساس جمع‌آوری اطلاعات از طریق پرسش از افراد در محل اقامت (خانوار) یا در محل کار (کارگاه) و یا از طریق منابع ثبتی است. خطای آماری در ذات جمع‌آوری اطلاعات آماری است و از بین بردن آن یا کاستن آن مستلزم شناسایی دقیق منابع خطا و کنترل عوامل تأثیرگذار است. مرکز آمار ایران همواره برای شناسایی و کنترل خطاها و بهبود کیفیت آمارها در تلاش بوده است. خطای اندازه‌گیری یکی از مهم‌ترین خطاهای غیرنمونه‌گیری است که به دلایل مختلفی مانند فراموشی، درست نفهمیدن سؤال یا در دست نداشتن اطلاعات یا به دلایل عمدی دیگر رخ می‌دهد. محاسبه‌ی خطای اندازه‌گیری به دلیل در دست نداشتن مقادیر واقعی متغیرهای پاسخ، همواره یک چالش اساسی برای روش‌شناسان آمارگیری بوده است و ارائه‌ی راهکارهایی برای کمی‌کردن و کاهش این خطاها در آمارگیری‌ها از اهمیت بسیاری برخوردار است. از جمله طرح‌های آمارگیری مستمر مرکز آمار ایران، طرح آمارگیری فصلی نیروی کار خانوارهای شهری و روستایی است. آمارگیری نیروی کار ایران از سال ۱۳۸۴ با استفاده از نمونه‌گیری چرخشی و با هدف دستیابی به شاخص‌های فصلی و سالانه‌ی نیروی کار و تغییرات

عبارات و کلمات کلیدی: مدل رده‌ی نهان مارکوف، آمارگیری نیروی کار، خطای اندازه‌گیری.
نوع مقاله: کاربردی
دبیرتخصصی رابط: مجید اسدی
* نویسنده مسئول
تاریخ دریافت: ۱۴۰۳/۰۷/۱۸ تاریخ پذیرش: ۱۴۰۴/۰۲/۳۰ تاریخ انتشار آنلاین: ۱۴۰۴/۰۴/۰۸
ارجاع به مقاله: ل. کلهری ندرآبادی و پ. ترابی کهلان، ارزیابی خطای اندازه‌گیری شاخص‌های نیروی کار ایران با تغییر شیوه‌ی آمارگیری بر اساس مدل رده‌ی نهان مارکوف، ریاضی و جامعه، ۱۱ شماره ۱ (۱۴۰۵) ۱۰۹-۱۳۱. https://dx.doi.org/10.22108/msci.2025.142988.1700 .

آن در سطح کل کشور و استان‌ها، به‌صورت فصلی اجرا شده است. نرخ‌های اشتغال و بیکاری و همچنین برآوردهای تعداد بیکاران وشاغلان که مرکز آمار ایران به صورت فصلی وسالانه ارائه می‌دهد، از این آمارگیری به‌دست می‌آیند. روش نمونه‌گیری در آمارگیری نیروی کار، نمونه‌گیری خوشه‌ای دو مرحله‌ای با طبقه‌بندی است که واحد نمونه‌گیری مرحله‌ی اول، خوشه و واحد نمونه‌گیری مرحله دوم، گروهی شامل سه خانوار معمولاً مجاور است. الگوی چرخش ۲-۲-۲ یعنی از هر خانوار نمونه چهار بار آمارگیری به عمل می‌آید، به این ترتیب که خانوار، دو فصل متوالی در نمونه است سپس به‌طور موقت برای دو فصل متوالی از نمونه خارج می‌شود، بعد مجدد برای دو فصل متوالی به نمونه باز می‌گردد و پس از آن برای همیشه از نمونه خارج می‌شود. به این ترتیب میزان تداخل نمونه بین دو فصل متوالی و دو فصل یکسان از دو سال متوالی، ۵۰٪ و بین دو سال متوالی، ۵۵٪ است. شیوه‌ی معمول تکمیل پرسشنامه از طریق مراجعه‌ی حضوری مأمور آمارگیر و مصاحبه با خانوار است. اما شیوع یکباره بیماری کووید ۱۹ و لزوم حفظ قرنطینه، آمارگیری‌های رودررو با مراجعه‌ی حضوری را به شدت تحت تأثیر قرار داد و انجام این آمارگیری‌ها در برخی از کشورها به‌صورت کامل تعطیل شد و تنها آمارگیری‌هایی که لغو آن‌ها امکان‌پذیر نبود، به شیوه‌های تلفنی یا اینترنتی اجرا شدند. در مورد اجرای آمارگیری‌های منظم که متغیرهای اجتماعی اقتصادی کوتاه‌مدت مانند شاخص‌های نیروی کار را اندازه‌گیری می‌کند و این متغیرها برای اتخاذ تصمیم‌گیری، بسیار مهم است، به تعویق انداختن عملیات آمارگیری یک گزینه مناسب نیست. مرکز آمار ایران نیز از این قاعده مستثنی نبوده است و برای رفع این نیاز، مرکز آمار ایران به‌عنوان مرجع رسمی تهیه و انتشار آمارهای رسمی کشور شیوه‌ی گردآوری داده‌های آمارگیری نیروی کار را از بهار ۱۳۹۹ از مصاحبه‌ی رودررو به مصاحبه‌ی تلفنی تغییر داد. با وجود تهیه‌ی دستورالعمل‌ها و در نظر گرفتن تمهیدات لازم برای اجرای این آمارگیری، تغییر شیوه‌ی گردآوری داده‌ها به‌صورت ناگهانی و بدون آزمایش‌های قبلی صورت گرفت که می‌تواند بر نتایج آمارگیری اثر گذار باشد. لذا ارزیابی مؤلفه‌های خطای آمارگیری برای ارزیابی تغییر شیوه‌ی گردآوری اطلاعات بر نتایج آمارگیری ضروری است. خطای اندازه‌گیری به‌عنوان یکی از مهم‌ترین خطاهای غیرنمونه‌گیری، نمایانگر تفاوت بین مقدار کمیت مورد نظر و مقدار گزارش‌شده از سوی پاسخ‌گو است. محاسبه‌ی مقدار دقیق خطای اندازه‌گیری تنها در صورتی امکان‌پذیر است که مقدار واقعی کمیت مورد نظر معلوم و در دسترس باشد. برای کمی کردن خطای اندازه‌گیری در متغیرهای رسته‌ای، آماردانان معمولاً از روش‌های زیر استفاده می‌کنند:

- روش استاندارد طلایی [1]: این روش، اریبی بین اندازه‌گیری با خطا و اندازه‌گیری بدون خطا (استاندارد طلایی) را می‌سنجد. نقطه ضعف روش استاندارد طلایی این است که این مقدار به ندرت وجود دارد و در مواردی نیز که وجود دارد، ممکن است خطای پنهان (مشاهده نشده) در برآورد استاندارد طلا وجود داشته باشد که می‌تواند منجر به نتایج نامعتبر شود.
- تحلیل قابلیت اطمینان: واریانس را تجزیه می‌کند تا بخشی از واریانس برآورد را که به دلیل خطای اندازه‌گیری است، مشخص کند.
- تحلیل رده‌ی نهان (LCA[2]): از روش حداکثر درستنمایی برای برآوردها استفاده می‌کند تا بتواند اریبی بین برآوردهای آمارگیری و حداکثر درستنمایی را تعیین کند.
- تحلیل رده‌ی نهان مارکوف (MLCA[3]) : از رویکردی مشابه با LCA استفاده می کند، اما برای آمارگیری‌های پانلی یا طولی طراحی شده است [۱].

مدل‌های رده‌ی نهان مارکوف را می‌توان به‌عنوان مدلی مناسب و جایگزین مصاحبه‌ی مجدد در برآورد خطای رده‌بندی در داده‌های نیروی کار، به کار گرفت [۲] و [۱۱]. مشکل اصلی مصاحبه‌ی مجدد هزینه اجرای قابل توجه و فراهم کردن شرایط ایده‌آل در اجرای آمارگیری از جمله آمارگیر خبره و نظارت دقیق است که در عمل اجرای آن را با مشکلاتی همراه کرده است. در مطالعاتی نیز از رویکرد رده‌ی نهان در یک آمارگیری کوچک‌مقیاس برای محاسبه‌ی خطای رده‌بندی [۱۴] و ارزیابی خطای اندازه‌گیری آمارگیری سلامت روان ایرانیان [۴] استفاده شده است. تحقیقات تجربی در مورد خطاهای اندازه‌گیری در آمارگیری‌ها

[1]gold-standard method [2]latent class analysis [3]Markov latent class analysis

https://dx.doi.org/10.22108/msci.2025.142988.1700

گسترده نیستند. به‌طور مثال، [۱۲] از داده‌های یک آمارگیری کوچک‌مقیاس که در یکی از دانشکده‌های دانشگاهی در تهران انجام شد، برای برآورد خطای کل آمارگیری استفاده کردند. آن‌ها پیشنهاد دادند که واحدهای تولیدکننده‌ی آمار با بازنگری در طرح‌های آمارگیری امکان محاسبه‌ی خطای کل آمارگیری را به‌وجود آورند [۵]. نشان دادند که خطای اندازه‌گیری، اریبی زیادی در داده‌های آمارگیری ایجاد می‌کنند به‌طوری‌که در این مطالعه اریبی میانگین بیش از سه برابر بیشتر از بی‌پاسخ‌دهی آمارگیری بوده است.

در این مقاله، ضمن تشریح مدل‌های تحلیل رده‌ی نهان مارکوف به‌عنوان جایگزینی برای روش مصاحبه‌ی مجدد، برآورد خطای اندازه‌گیری آمارگیری نیروی کار ایران در شرایط شیوع بیماری کووید ۱۹ که منجر به تغییر سریع شیوه‌ی گردآوری داده‌ها شد، انجام شده است. در ادامه، چالش‌های اصلی کشورها در اجرای طرح آمارگیری نیروی کار در زمان همه‌گیری کووید ۱۹ مرور می‌شود. بخش سوم مقاله، شامل روش‌شناسی برآورد خطای اندازه‌گیری در آمارهای نیروی کار در زمان همه‌گیری کووید ۱۹ و رویکرد رده‌ی نهان مارکوف است و بخش چهارم، کاربرد این مدل در آمارگیری نیروی کار با استفاده از داده‌های چهار دوره‌ی آمارگیری از بهار ۱۳۹۷ تا تابستان ۱۴۰۰ را دربر دارد. بخش پنجم، به بحث و نتیجه‌گیری اختصاص یافته است.

۲. چالش‌های اصلی کشورها در اجرای طرح آمارگیری نیروی کار در زمان همه‌گیری کووید ۱۹

در این بخش، چالش‌های کلیدی و برخی از راهکارهایی را که کشورها برای حفظ عملیات طرح آمارگیری نیروی کار خود و کاهش تأثیر کووید ۱۹ بر آمارهای تولید شده در نظر گرفتند، ارائه می‌شود [۳]. شایان ذکر است که با توجه به تنوع چالش‌ها در زمان همه‌گیری کووید ۱۹، راه حل‌های مناسب در نهایت بایستی در سطح هر کشور ارزیابی می‌شد و همه چالش‌های پیش‌رو، قابل پیش‌بینی، یا راهکار آن قابل تجویز برای تمام کشورها نبود. اما به‌طور کلی، چالش‌های اصلی پیش‌روی اکثر کشورها در زمان همه‌گیری کووید ۱۹ را می‌توان در دو سناریو بررسی کرد. در سناریوی اول حرکت از مصاحبه‌ی رودررو کامل به سمت مصاحبه‌ی تلفنی مطرح می‌شد، که مسائل متفاوتی برای کشورهای دارای طرح‌های پانلی یا دارای نمونه‌های مستقل برای هر دوره‌ی آمارگیری داشت. در سناریوی دوم، استفاده از روش چندمُدی [۴] برای گردآوری داده‌ها مطرح بود. موضوع اصلی در این‌حالت، حرکت از این نوع مصاحبه‌ها به سمت مصاحبه‌ی تلفنی یا مصاحبه‌ی وبی رایانه‌یار (CAWI[۵]) بود.

در هر یک از سناریوها، پیشنهادهایی توسط سازمان بین‌المللی کار برای راهنمایی کشورها در مورد نحوه‌ی ارزیابی اعتبار یا استواری رویکردهای مختلف ارائه شد، زیرا این موضوع، صرف‌نظر از سناریوی به‌کار گرفته شده، به‌دلیل به‌کارگیری رویکردهای تغییر یافته برای گردآوری داده‌ها، موضوع مهمی بود. باید این نکته را مد نظر قرار داد که در حالت ایده‌آل، هر تغییری از طریق یک فرایند اساسی توسعه و آزمون، ایجاد می‌شود. بنابراین، حتی در شرایط ویژه‌ای که در اثر همه‌گیری بیماری کووید ۱۹ ایجاد شده بود و تغییرات اساسی یکباره تحمیل شده بودند، برخی آزمون‌ها و ارزیابی‌ها قبل از اجرای تغییرات، حتی در مقیاس کوچک، باید در برنامه‌ها گنجانده می‌شدند تا مشکلات حین اجرا کاهش یابد. کشورهایی که پیش از همه‌گیری کووید ۱۹ از مصاحبه‌ی وبی رایانه‌یار CAWI یا مصاحبه‌ی تلفنی رایانه‌یار (CATI[۶]) کامل استفاده می‌کردند، به‌طور معمول با حداقل اختلال در گردآوری داده‌ها مواجه شدند. در مقابل، برخی از کشورها توانایی خود را برای گردآوری داده‌ها از طریق روش‌های موجود به‌طور کامل از دست دادند. چالش پیش روی این کشورها را نمی‌توان دست کم گرفت و ممکن بود گردآوری داده‌ها برای مدت زمان کوتاه یا برای مدت زمان طولانی‌تر به تعلیق درآید که در واقع این امر در زمان همه‌گیری در بسیاری از کشورها اتفاق افتاد.

سناریوی ۱ توسط تعداد زیادی از کشورها گزارش و به موجب آن تمام مصاحبه‌های رودررو (مصاحبه‌ی شخصی یارانه‌یار (CAPI[۷]) یا مصاحبه‌ی قلم و کاغذی (PAPI[۸]) در زمان همه‌گیری کووید ۱۹ متوقف شدند. تحت این سناریو، دوره‌های مختلف

[4]multi-mode [5]computer-assisted web interviewing (CAWI) [6]computer-assisted telephone interview (CATI)) [7]computer-assisted personal interview (CAPI) [8]pen and paper interviewing (PAPI)

گردآوری داده‌ها (گردآوری فصلی پیوسته [9]، دوره‌ای فصلی [10]، پیوسته‌ی سالانه [11]، دوره‌ای سالانه [12] و غیره) وجود داشت. در این وضعیت، برخی از کشورها از نمونه‌های مستقل برای هر آمارگیری و برخی دیگر از طرح پانلی استفاده کردند. چالش‌ها اغلب یکسان بودند، اما کشورهایی که طرح پانلی داشتند، به‌طور بالقوه توانستند به طُرق مختلف از آن بهره ببرند. گرچه ممکن بود انجام مصاحبه‌ی رودررو ضروری به‌نظر رسد، اما با این وجود سازمان بین‌المللی کار توصیه کرد راه‌هایی برای شروع انتقال به مصاحبه‌ی تلفنی در نظر گرفته شود، زیرا در زمان محدودیت‌ها، مانند تجربه‌ی همه‌گیری کووید ۱۹، کمتر دچار اختلال می‌شوند. چالش‌های اصلی که کشورها تحت این سناریو با آن مواجه شدند، عبارتند از:

الف) بازیابی اطلاعات تماس برای واحدهای نمونه‌ی نهایی،

ب) توانمندسازی افراد و افزایش ظرفیت برای انجام مصاحبه‌های تلفنی،

ج) ارزیابی مناسب بودن پرسش‌نامه‌ی PAPI/CAPI برای مصاحبه‌های تلفنی.

سناریوی ۲ در بسیاری از کشورهای اروپایی و کشورهای منتخب در سایر نقاط جهان اتفاق افتاد. معمول‌ترین سناریو، مصاحبه‌ی رودررو برای اولین مصاحبه و CATI (و به‌طور احتمالی) CAWI برای مصاحبه‌های بعدی بود. برای خانوارهایی که به دلایلی (مانند نداشتن تلفن یا ترجیح به مصاحبه‌ی رودررو) اجرای CATI یا CAWI برای آن‌ها امکان‌پذیر نبود، مصاحبه‌ی رودررو در سایر نوبت‌های آمارگیری نیز تکرار شد. همانند کشورهای تحت سناریوی ۱، پاسخ فوری برای اکثریت این کشورها این بود که هرگونه مصاحبه‌ی رودرروی باقی‌مانده را به یکی از شیوه‌های دیگر، به‌طور عمده CATI، منتقل کنند. در این وضعیت طیف چالش‌ها و گزینه‌ها به‌طور کلی مشابه مواردی است که برای سناریوی ۱ ذکر شد.

اما پس از گردآوری داده‌ها نیز کشورها با چالش‌هایی مواجه بودند. به دلایل بسیاری، ممکن بود نتوان به نمونه‌ی نمایانگر با سطح‌های پاسخ دلخواه، دست یافت. این امر به طریق مختلف می‌توانست منجر به اریبی در پوشانش جامعه شود. پوشانش خوب از طریق ادامه‌ی انجام مصاحبه‌ها با شیوه‌ای متفاوت در طول قرنطینه، یا تکمیل عملیات گردآوری داده‌ها در مرحله‌ی بعد، یا به‌طور بالقوه ترکیبی از این دو روش قابل حفظ بود. با توجه به این‌که بسیاری از کشورها در حال تلاش برای معرفی یا افزایش مصاحبه‌های تلفنی در زمان شیوع بیماری کووید ۱۹ بودند، برخی از موارد خاص ایجاد اریبی نیز وجود داشت، مانند برخی از گروه‌هایی که تلفن ثابت یا همراه نداشتند یا نمی‌خواستند شماره‌ی آن‌ها در فهرست تلفن‌ها منتشر شود و گروه‌هایی که گرچه اطلاعات تماس برای برخی از آن‌ها در دسترس بود، اما تماس با آن‌ها دشوار بود. برقراری تماس و دسترسی دشوار به این گروه‌ها به دلیل حضور نداشتن آن‌ها در محل سکونتشان در اکثر اوقات در طول روز و هفته بود.

اریبی ناشی از بی‌پاسخی موضوع جدیدی نیست، اما در شرایط همه‌گیری کووید ۱۹ با اعمال قرنطینه‌های گسترده و افزایش احتمالی بی‌پاسخی، این مشکل اهمیت بیشتری را به خود اختصاص داد. یکی از راه حل‌های اولیه‌ی مورد استفاده‌ی کشورها برای اجتناب از اریبی‌های ناشی از بی‌پاسخی، استفاده از شیوه‌های آمیخته‌مُد برای گردآوری داده‌ها بود. مدیریت بی‌پاسخی پس از گردآوری داده‌ها از طریق ارزیابی اریبی احتمالی شامل مقایسه‌ی ویژگی‌های نمونه‌ی واقعی به‌دست آمده با چارچوب نمونه‌گیری یا نمونه‌های گذشته بود. اگر ارزیابی‌ها نشان‌دهنده‌ی اریبی‌های احتمالی بودند، اداره‌های آمار ملی (NSOs[13]) حداقل دو گزینه‌ی تلاش برای انجام مصاحبه‌های مضاعف با گروه‌هایی که کمتر پوشش داده شده‌اند و تلاش برای کاهش اریبی با تعدیل بی‌پاسخی و وزن‌دهی را پیش‌رو داشتند. چالش بعدی، اجرای یک رویکرد گزارش‌دهی کیفیت جامع یا عناصری از آن برای NSOها بود. بسیاری از بُعدها و شاخص‌ها می‌توانند برای توصیف کیفیت یک آماره مورد استفاده قرار گیرند، اما مواردی که بیشترین ارتباط با اختلالات ناشی از همه‌گیری کووید ۱۹ را داشتند، درستی داده‌ها [14] (خطاهای نمونه‌گیری، خطاهای پوشانش، خطاهای اندازه‌گیری، خطاهای پردازش و خطاهای بی‌پاسخی) و قابلیت مقایسه‌ی آن‌ها در طول زمان بود. سایر شاخص‌ها مانند به‌موقع بودن [15] و

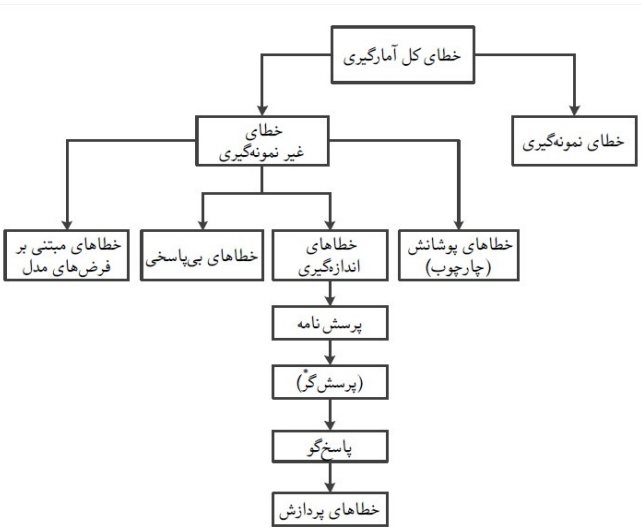
[9]quarterly continuous collection [10]quarterly periodic [11]annual continuous [12]annual periodic [13]national statistical offices
[14]accuracy of data [15]punctuality

https://dx.doi.org/10.22108/msci.2025.142988.1700

هم‌خوانی [۱۶]، گرچه کمتر تحت تأثیر اختلالات یا تغییرات در گردآوری داده‌های آمارگیری نیروی کار قرار داشتند، اما بدون شک مهم بودند. ارزیابی کیفیت آمارهای تولیدشده هم برای در نظر گرفتن تعدیل‌های ممکن در روش‌شناسی مانند جانهی [۱۷]، تعدیل بی‌پاسخی [۱۸]، وزن‌دهی و برآورد و هم برای ارزیابی مناسب کیفیت داده‌ها قبل و در حین انتشار آن‌ها مهم است.

۳. خطاهای آمارگیری

سازمان‌های آمار ملی و آمارشناسان همواره در تلاش هستند که با توجه به امکانات و محدودیت‌های مالی، آمارهایی تولید کنند که سطح مطلوبی از کیفیت را دارا باشند. کیفیت در آمارگیری‌های نمونه‌ای، مفهومی گسترده و چندبُعدی است و در بیش‌تر آمارگیری‌های نمونه‌ای امکان دستیابی به آمارهایی با بهترین سطح کیفیت در تمام جنبه‌های آن به‌طور تقریبی غیرممکن است. چنان‌چه جنبه‌های مختلف کیفیت شناخته نشود، امکان دارد تولیدکنندگان به اشتباه تنها بر یک بُعد کیفیت تأکید کنند و با این‌کار کیفیت را در سایر بُعدها کاهش دهند. در بیش‌تر مطالعه‌های آماری، اندازه‌ی جامعه به حدی بزرگ است که امکان گردآوری اطلاعات از تمام اعضای آن وجود ندارد؛ در این‌حالت، می‌توان یک نمونه‌ی کوچک ولی نمایانگر را برای برآورد پارامترهای جامعه‌ای به‌کار برد. در این مواقع، خصیصه‌های مورد نظر، برای واحدهای نمونه‌ای اندازه‌گیری شده و سپس آماره‌های حاصل به کل جامعه تعمیم داده می‌شوند. خطای نمونه‌گیری، از انتخاب آگاهانه‌ی قسمتی از جامعه به‌جای کل جامعه، ناشی می‌شود. سایر خطاها، به غیر از خطای نمونه‌گیری، خطای غیرِنمونه‌گیری نامیده می‌شوند. این خطاها، ناشی از عامل‌هایی هستند که چه به‌صورت عمدی و چه به‌صورت غیرعمدی در هنگام طراحی و در حین اجرای فعالیت‌های یک آمارگیری، رخ داده و موجب کاهش کیفیت آماره‌های آمارگیری می‌شوند. در بسیاری از موارد، بروز خطاهای غیرِنمونه‌گیری در آمارگیری‌ها اجتناب‌ناپذیر است ولی می‌توان با شناختن علت‌های بروز آن‌ها، این منبع‌ها را کنترل کرد. شکل ۱ مؤلفه‌های خطای کل آمارگیری، خطای نمونه‌گیری و خطاهای غیرِنمونه‌گیری را نشان می‌دهد [۱۳].



شکل ۱. مؤلفه‌های خطای کل آمارگیری، خطای نمونه‌گیری و خطاهای غیرِنمونه‌گیری [۱۳].

Figure 1: Components of total survey error, sampling error, and non-sampling errors [10].

[16]coherence [17]imputation [18]non-response adjustment

https://dx.doi.org/10.22108/msci.2025.142988.1700

۱.۳. منابع خطای اندازه‌گیری و برآورد آن در آمارگیری نیروی کار در زمان همه‌گیری کووید ۱۹. خطای اندازه‌گیری یا خطای مشاهده‌ای، هر خطایی است که سبب شود مقدار اندازه‌گیری‌شده‌ی متغیر مورد مطالعه برای یک واحد نمونه، متفاوت با مقدار واقعی آن متغیر باشد. شش مؤلفه وجود دارند که هر یک از آن‌ها می‌تواند سبب بروز خطای اندازه‌گیری شود. این شش مؤلفه عبارتند از پرسش‌گر، پاسخ‌گو، روش گردآوری داده‌ها، پرسش‌نامه (ابزار سنجش)، نظام بازیابی اطلاعات و محیط مصاحبه.

پرسشنامه یا مشخصات پرسشنامه یعنی نحوه‌ای که سؤالات طرح شده‌اند و شیوه‌ای که سؤال‌ها پرسش می‌شوند در نحوه پاسخ اثرگذار است. خطای اندازه‌گیری بر حسب روش جمع‌آوری اطلاعات میزان و اثر مختلفی بر نتایج دارد. سه نوع اصلی روش جمع‌آوری اطلاعات آماری شامل روش حضوری، روش تلفنی، روش خوداظهاری است. روش‌های دیگر در موارد خاص استفاده می‌شوند. آمارگیری حضوری روش مرسوم آمارگیری‌های مرکز آمار ایران است. در این نوع آمارگیری، آمارگیر سؤال را بر اساس پرسشنامه طرح شده از پاسخ‌گو می‌پرسد و پاسخ را در پرسشنامه ثبت می‌نماید. پرسشنامه معمولاً به‌صورت کاغذی یا به‌صورت تبلتی است. با همگانی شدن بیش از بیش شبکه‌ی تلفنی در ایران و سایر کشورها قابلیت جمع‌آوری اطلاعات آماری از طریق تلفن بیشتر می‌شود. آمارگیری تلفنی به‌خصوص در طرح‌های پانلی یا در طرح‌هایی با نمونه‌ی چرخشی می‌تواند مؤثر باشد. در این نوع طرح‌ها، مصاحبه‌ی اول به‌صورت رودررو انجام می‌شود و چنان‌چه خانوار دارای تلفن باشد و رضایت او جلب شود، مصاحبه‌های دوره‌های بعد را می‌توان از طریق تلفن انجام داد. با حذف شدن مخارج رفت و آمد، آمارگیری تلفنی نسبت به آمارگیری حضوری کم‌هزینه‌تر است. در آمارگیری تلفنی نیز نظارت بیشتری بر عملکرد آمارگیر می‌توان انجام داد و در نتیجه خطاهای اندازه‌گیری را بیشتر کنترل کرد. در مقابل، احتمال عدم همکاری و بی‌پاسخی در آمارگیری‌های تلفنی بیش از آمارگیرهای حضوری است، زیرا در آمارگیری تلفنی به دلایل مختلفی که در ادامه خواهد آمد پاسخ‌گو آسان‌تر می‌تواند با آمارگیر قطع ارتباط نماید. آمارگیری تلفنی نیز زمان کوتاه‌تری را در اختیار آمارگیر برای مصاحبه می‌گذارد و در نتیجه این روش جمع‌آوری اطلاعات آماری برای آمارگیری‌های طولانی مناسب نیست.

آمارگیر یکی از حساس‌ترین نقش‌ها را در جمع‌آوری اطلاعات آماری در آمارگیری‌های از طریق مصاحبه ایفا می‌کند. انتخاب صحیح آمارگیر برای اجرای مصاحبه و جمع‌آوری آمار، اولین قدم در کنترل اثر آمارگیر بر خطای اندازه‌گیری است. پاسخ‌گو، نقطه نهایی سیر جمع‌آوری اطلاعات آماری است. او کسی است که اطلاعات مورد نیاز را ارائه می‌نماید. شناسایی مراحلی که در ذهن پاسخ‌گو، بین شنیدن سؤال و دادن پاسخ صورت می‌گیرد، برای کنترل خطای پاسخ‌گویی اساسی است. در پاسخ‌گویی، ذهن پاسخ‌گو اساساً چهار مرحله، درک سؤال، استخراج اطلاعات از حافظه، قضاوت و تصمیم در نحوه پاسخ‌گویی و ارائه پاسخ نهایی را طی می‌کند و در هر کدام از این مراحل چهارگانه پاسخ‌گو می‌تواند خطایی انجام دهد که منجر به خطای اندازه‌گیری شود [۱۵].

در واقع در آمارگیری مصاحبه‌ای دو عامل اصلی پاسخ‌گو و مصاحبه‌گر هر دو بر کیفیت داده‌ها اثرگذار هستند. مسائلی مانند این‌که آیا مصاحبه از طریق تلفن (تلفن ثابت یا همراه) یا رودررو انجام می‌شود، آیا مصاحبه در منزل پاسخ‌گو انجام می‌شود یا در یک فضای بی‌طرف، آیا مصاحبه‌گر و پاسخ‌گو به گروه‌های اجتماعی مختلف تعلق دارند، آیا مصاحبه در حضور دیگران انجام می‌شود، زمینه‌ی اجتماعی و تأثیر آن‌ها بر پاسخ‌گو در نتایج آمارگیری تأثیرگذار هستند. دریافت مشارکت و ایجاد انگیزه در پاسخ‌گو، درک پرسش‌ها و دستورالعمل‌ها، میزان دانش مصاحبه‌گر در مواجهه با مواقعی که پاسخ‌گو نیاز به درک بیشتر پرسش دارد، از جمله وظایف مصاحبه‌گرها در آمارگیری‌ها است. در مصاحبه‌ی تلفنی برخی محدودیت‌ها وجود دارد. به‌عنوان مثال بر خلاف مصاحبه‌ی رودررو، ارتباط چشمی بین مصاحبه‌گر و مصاحبه‌شونده در مصاحبه‌ی تلفنی وجود ندارد. از این‌رو، نحوه‌ی ایجاد ارتباط با نشانه‌های ارتباطی صریح و ضمنی با پاسخ‌گویان تلفنی، می‌تواند تأثیرات قابل‌توجهی بر موفقیت ارتباط و در نتیجه کیفیت داده‌های آمارگیری داشته باشد. مصاحبه‌گری که بتواند درک پاسخ‌گو را تقویت کند، داده‌های آمارگیری دقیق‌تری را گرداوری می‌کند.

از سوی دیگر ویژگی‌های مصاحبه‌گرهای تلفنی از طریق صدایشان اثر قابل توجهی بر نرخ پاسخ دارند. این‌که صدایشان جذاب، قابل اعتماد، و حرفه‌ای باشد، همه این موارد و بسیاری نکات ظریف دیگر باعث می‌شود که مصاحبه‌گران و مصاحبه‌شوندگان با

https://dx.doi.org/10.22108/msci.2025.142988.1700

کلمات و ویژگی‌های شنیداری صدا، ارتباط برقرار کنند. در مقابل، مدت مکث طولانی، تن صدای پایین، قدرت صدای کم، توضیح دادن نابجا، همگی در پایین آوردن نرخ پاسخ تأثیر دارند. نقش مصاحبه‌گر در بی‌پاسخی مصاحبه‌شونده در مصاحبه‌ی تلفنی از مصاحبه‌ی رودررو بیشتر است و مصاحبه‌شوندگان خیلی سریع‌تر تصمیم می‌گیرند که بی‌پاسخ باشند یا به پرسش جواب بدهند.

روش‌های مختلفی برای برآورد خطاهای اندازه‌گیری وجود دارد، از جمله طرح آزمایش تصادفی [۱۹]، مقایسه اطلاعات ثبتی [۲۰] و مصاحبه‌ی مجدد [۲۱] و رده‌ی نهان مارکوف (MLC) [۱۵]. در روش آزمایش تصادفی، حالت‌های مختلف متغیر که به خطای اندازه‌گیری موردنظر مربوط می‌شوند، با زیرنمونه‌های مختلف در شرایط کنترل‌شده، آزمایش و نتایج، با هم مقایسه می‌شوند. در روش مقایسه با اطلاعات ثبتی، داده‌های حاصل از پاسخ در طرح آمارگیری با اطلاعات موجود در منبع ثبتی که فرض می‌شود از دقت بیشتری برخوردار است، مقایسه می‌شوند. در روش مصاحبه‌ی مجدد، پرسش یا پرسش‌های مورد نظر در مصاحبه‌ی اصلی مجدداً در شرایط مشابه از همان واحد آماری پرسش می‌شود و پاسخ در مصاحبه‌ی مجدد با پاسخ در مصاحبه‌ی اصلی مقایسه می‌شود. برای کنترل شرایط و صرفه‌جویی مالی به‌طور معمول مصاحبه‌ی مجدد با زیرنمونه‌ی محدودی به‌طور مثال ۵٪ نمونه‌ی اصلی و در فاصله‌ی زمانی کوتاهی بعد از مصاحبه‌ی اصلی صورت می‌گیرد، به‌طوری‌که شرایط مصاحبه‌ی مجدد با شرایط مصاحبه‌ی اصلی حتی‌الامکان یکسان باشد. روش آزمایش تصادفی و مصاحبه‌ی مجدد نیازمند اجرای مجدد آمارگیری هر چند در سطح محدود است، و لذا نیازمند صرف هزینه و زمان می‌باشد. علاوه بر آن لازم است اجرا، کاملاً در شرایط کنترل‌شده انجام شود تا قابلیت مقایسه و ارزیابی نتایج وجود داشته باشد. همچنین استفاده از روش مقایسه‌ی اطلاعات ثبتی تنها در صورتی قابل استفاده است که منابع اطلاعات ثبتی درخصوص متغیر مورد بررسی موجود، معتبر و در دسترس باشند. در مقابل این روش‌ها، روش رده‌ی نهان مارکوف قرار می‌گیرد که در آمارگیری‌های طولی با حداقل سه دوره تکرار، قابل استفاده است. مزیت این روش آن است که با استفاده از اطلاعات حاصل از واحدهای نمونه‌ای مشترک در دوره‌های مختلف آمارگیری امکان ارزیابی خطای اندازه‌گیری را فراهم می‌سازد و نیازی به تکرار آمارگیری در شرایط کنترل‌شده یا دسترسی به اطلاعات ثبتی ندارد. از آنجایی‌که تمرکز این پژوهش، برآورد خطای اندازه‌گیری با استفاده از روش رده‌ی نهان مارکوف است در ادامه، این مدل به اختصار معرفی می‌شود. با توجه به این‌که در واقع مدل رده‌ی نهان مارکوف، در زمره‌ی مدل‌های مارکوف آشکار (MM) است که در آن متغیرهای پاسخ با خطا ثبت شده‌اند، برای درک مدل نهان ابتدا مفهوم مدل مارکوف آشکار، به اختصار بیان می‌شود. فرض کنید که برآورد X بدون خطای اندازه‌گیری باشد. چنان‌چه خواهیم دید مدل‌های (MM) مؤلفه‌های ساختاری مدل‌های (MLC) است. در حقیقت، در ادامه با اضافه کردن یک مؤلفه‌ی اندازه‌گیری، چارچوب مدل‌بندی (MM) به مدل‌های (MLC) تعمیم داده خواهد شد.

۲.۳. مدل مارکوف آشکار. فرض کنید که X متغیر رسته‌ای مورد بررسی باشد که در هر دوره بدون خطا ثبت شده است. برای هر واحد دلخواه در جامعه، فرض کنید x_t مقدار مشاهده شده X_t، $t = ۱, ۲, \ldots, T$، باشد. فرض کنید K تعداد وضعیت‌هایی باشد که هر واحد در زمان t می‌تواند اختیار کند. همچنین، فرض کنید که تعداد وضعیت‌ها در همه زمان‌ها یکسان باشد. تغییرات از وضعیت $x_{t-۱}$ به x_t را انتقال و $P(X_t = x_t|X_{t-۱} = x_{t-۱})$ را احتمال انتقال گویند و با $\pi^{X_t|X_{t-۱}}_{x_t|x_{t-۱}}$ نشان می‌دهند. این احتمال نشان می‌دهد که واحدی که زمان t در وضعیت x_t است، زمان $t-۱$ در وضعیت $x_{t-۱}$ بوده است. به‌عنوان مثال فرض کنید که X_t وضع فعالیت یک فرد در زمان t را مشخص می‌کند که $X_t = ۳, ۲, ۱$ به ترتیب به وضعیت اشتغال (EMP[۲۲])، بیکاری (UNE[۲۳]) و غیرفعال بودن (NLF[۲۴]) اشاره می‌کند. بنابراین، یک فرد در هر زمان می‌تواند تنها یکی از سه وضعیت را تصرف کند و در هر دوره $K = ۳$ است. فرض کنید سه دوره‌ی پیاپی در آمارگیری پانلی وجود دارد $(T = ۳)$ و $n_{x_۱x_۲x_۳}$ مشخص کننده تعداد افرادی که در دوره‌ی اول در وضعیت $x_۱$، در دوره‌ی دوم در وضعیت $x_۲$ و در دوره‌ی سوم در وضعیت $x_۳$ هستند، باشد. نمایش داده‌ها در یک جدول سه‌طرفه مانند جدول ۱ خواهد بود.

[19] randomized experiment [20] record checking [21] reinterview survey [22] employed [23] unemployed [24] not in labour force

جدول ۱. رده‌بندی متقاطع وضعیت اشتغال در جدول سه‌طرفه اشتغال (E)، بیکاری (U) و غیرفعال بودن (N)

Table 1: Cross-classification of employment status in a three-way table of Employment(E), Unemployment(U), and Not in Labor Force(N)

		دوره سوم									
		N			U			E			
		N	U	E	N	U	E	N	U	E	دوره دوم
دوره اول	E	$n_{۱۳۲}$	$n_{۱۲۳}$	$n_{۱۱۳}$	$n_{۱۳۲}$	$n_{۱۲۲}$	$n_{۱۱۲}$	$n_{۱۳۱}$	$n_{۱۲۱}$	$n_{۱۱۱}$	E
	U	$n_{۲۳۳}$	$n_{۲۲۳}$	$n_{۲۱۳}$	$n_{۲۳۲}$	$n_{۲۲۲}$	$n_{۲۱۲}$	$n_{۲۳۱}$	$n_{۲۲۱}$	$n_{۲۱۱}$	U
	N	$n_{۳۳۳}$	$n_{۳۲۳}$	$n_{۳۱۳}$	$n_{۳۳۲}$	$n_{۳۲۲}$	$n_{۳۱۲}$	$n_{۳۳۱}$	$n_{۳۲۱}$	$n_{۳۱۱}$	N

فرض کنید $\pi^{X_۱X_۲X_۳}_{x_۱x_۲x_۳}$ احتمال اینکه یک مشاهده متعلق به خانه $(x_۱, x_۲, x_۳)$ باشد، اشاره می‌کند. عام‌ترین شکل مدل به‌صورت

$$\pi^{X_۱X_۲X_۳}_{x_۱x_۲x_۳} = \pi^{X_۱}_{x_۱}\pi^{X_۲|X_۱}_{x_۲|x_۱}\pi^{X_۳|X_۱X_۲}_{x_۳|x_۱x_۲}$$

است که با لحاظ کردن ویژگی مارکوف به‌صورت زیر خلاصه می‌شود و به آن مدل مارکوف آشکار استاندارد گویند.

$$\pi^{X_۱X_۲X_۳}_{x_۱x_۲x_۳} = \pi^{X_۱}_{x_۱}\pi^{X_۲|X_۱}_{x_۲|x_۱}\pi^{X_۳|X_۲}_{x_۳|x_۲}$$

برآورد پارامترها، با در نظر گرفتن یک مدل لگ‌خطی و روش ماکسیمم درستنمایی به‌دست می‌آیند. در صورتی‌که احتمال، انتقال زمان همگن یا ایستا باشد، یعنی احتمال‌های انتقال برای دوره‌های پیاپی یکسان هستند، تعداد پارامترهای مدل کاهش یافته و برآورد آن‌ها با پیچیدگی کمتری مواجه خواهد بود.

۳.۳. مدل رده‌ی نهان مارکوف. مدل‌های رده‌ی نهان مارکوف ابتدا توسط [۱۰] پیشنهاد شدند، با این‌وجود [۶]، [۸] و [۷] سهم به‌سزایی در تعمیم آن در مسائل کاربردی داشته‌اند. در مدل‌های رده‌ی نهان مارکوف استاندارد فرض می‌شود که برخی ویژگی‌های مورد بررسی، X، برای $T \geq ۳$ دوره‌ی پانلی اندازه‌گیری شده است و برای هر متغیر X حداقل یک متغیر نشان‌گر[۲۵] مانند A، وجود دارد. اگر متغیر X در شرایط ایده‌آل قابل اندازه‌گیری باشد، آن‌گاه میزان خطا در اندازه‌گیری X با استفاده از متغیرهای نشان‌گر به‌صورت زیر قابل تعریف است:

$$A = X + e_A$$

$$B = X + e_B$$

$$C = X + e_C$$

که e_A، e_B و e_C مؤلفه‌های خطای تصادفی و مستقل از یکدیگر هستند. در ادامه، توضیحات برای یک مطالعه با سه دوره‌ی زمانی آمده است. تعمیم این مدل‌ها به چهار یا بیشتر از چهار دوره به‌صورت مشابه انجام می‌شود. در مدل رده‌ی نهان مارکوف استاندارد فرض‌های زیر در نظر گرفته می‌شود:

(۱) **ویژگی مارکوف**: این ویژگی، فرض می‌کند که

$$\pi^{X_۳|X_۱X_۲}_{x_۳|x_۱x_۲} = \pi^{X_۳|X_۲}_{x_۳|x_۲}$$

[25] indicator

https://dx.doi.org/10.22108/msci.2025.142988.1700

(۲) **استقلال خطای رده‌بندی:** این فرض می‌تواند به‌عنوان تعمیمی از استقلال موضعی در داده‌های طولی در نظر گرفته شود و عبارت است از

$$\pi^{A_۱A_۲A_۳|X_۱X_۲X_۳}_{a_۱a_۲a_۳|x_۱x_۲x_۳} = \pi^{A_۱|X_۱}_{a_۱|x_۱}\pi^{A_۲|X_۲}_{a_۲|x_۲}\pi^{A_۳|X_۳}_{a_۳|x_۳}$$

این فرض به این معنی است که خطاهای رده‌بندی برای سه متغیر نشانگر در هر دوره مستقل از یکدیگر هستند.

(۳) **زمان‌همگنی خطای رده‌بندی:** فرض بر آن است که خطاهای رده‌بندی برای همه دوره‌ها یکسان است؛ یعنی

$$\pi^{A_t|X_t}_{a_t|x_t} = \pi^{A|X}_{a|x},\ a = a_t,\ x = x_t,\ t = ۱, ۲, ۳.$$

(۴) **احتمال‌های خطای همگن:** فرض کنید که همه آزمودنی‌ها در یک کلاس پنهان یکسان دارای احتمال برابر برای رده‌بندی هستند. یعنی، $\pi^{A_t|X_t}_{a_t|x_t}$ برای همه واحدها در کلاس $X_t = x_t$ برابر هستند.

فرض کنید $A_۱A_۲A_۳$ به جدول $K \times k \times k$ برای مشاهده‌های n آزمودنی در دوره‌های ۱، ۲ و ۳ اشاره می‌کند. احتمال توأم برای این‌که یک مشاهده در خانه‌ی $(a_۱a_۲a_۳)$ از جدول قرار گیرد، عبارت است از:

$$\pi^{A_۱A_۲A_۳}_{a_۱a_۲a_۳} = \sum_{x_۱}\sum_{x_۲}\sum_{x_۳}\pi^{X_۱X_۲X_۳}_{x_۱x_۲x_۳}\pi^{A_۱A_۲A_۳|X_۱X_۲X_۳}_{a_۱a_۲a_۳|x_۱x_۲x_۳}$$

با استفاده از ویژگی مارکوف مؤلفه‌ی ساختاری مدل می‌تواند به‌صورت زیر در نظر گرفته شود:

$$\pi^{X_۱X_۲X_۳}_{x_۱x_۲x_۳} = \pi^{X_۱}_{x_۱}\pi^{X_۲|X_۱}_{x_۲|x_۱}\pi^{X_۳|X_۲}_{x_۳|x_۲}$$

با استفاده از فرض‌های استقلال خطای طبقه‌بندی و احتمال‌های خطای همگن، مؤلفه‌ی اندازه‌گیری‌شده می‌تواند به‌صورت زیر در نظر گرفته شود:

$$\pi^{A_۱A_۲A_۳|X_۱X_۲X_۳}_{a_۱a_۲a_۳|x_۱x_۲x_۳} = \pi^{A_۱|X_۱}_{a_۱|x_۱}\pi^{A_۲|X_۲}_{a_۲|x_۲}\pi^{A_۳|X_۳}_{a_۳|x_۳}$$

علاوه بر این با استفاده از فرض ۳ به‌صورت زیر ساده‌تر می‌شود:

$$\pi^{A_۱A_۲A_۳|X_۱X_۲X_۳}_{a_۱a_۲a_۳|x_۱x_۲x_۳} = \pi^{A|X}_{a_۱|x_۱}\pi^{A|X}_{a_۲|x_۲}\pi^{A|X}_{a_۳|x_۳}$$

توجه کنید که بالانویس $A|X$ در هر احتمال به همگنی احتمال‌های طبقه‌بندی در طول زمان اشاره می‌کند. در این وضعیت احتمال‌های هر خانه برای مدل رده‌ی نهان مارکوف استاندارد به‌صورت زیر حاصل می‌شود:

$$\pi^{A_۱A_۲A_۳}_{a_۱a_۲a_۳} = \sum_{x_۱}\sum_{x_۲}\sum_{x_۳}(\pi^{X_۱}_{x_۱}\pi^{X_۲|X_۱}_{x_۲|x_۱}\pi^{X_۳|X_۲}_{x_۳|x_۲})(\pi^{A|X}_{a_۱|x_۱}\pi^{A|X}_{a_۲|x_۲}\pi^{A|X}_{a_۳|x_۳})$$

که در آن $a_۳, a_۲, a_۱, x_۳, x_۲, x_۱$ مقادیرشان را از ۲ ،۱، ... ،K اخذ می‌کنند. برآورد پارامترهای مدل با استفاده از الگوریتم بیشینه‌سازی امید ریاضی (EM) انجام می‌شود [۱].

۴. کاربرد مدل رده‌ی نهان مارکوف در آمارگیری نیروی کار در زمان شیوع کووید ۱۹

با توجه به محدودیت‌هایی که از فصل بهار سال ۱۳۹۹ در سراسر کشور برای پیشگیری از شیوع بیشتر بیماری کووید ۱۹ ایجاد شد، و با توجه به روش‌های اجرای آمارگیری‌های نمونه‌ای در کشور که اغلب با مصاحبه‌ی رودرروی پرسش‌گران با واحدهای آماری صورت می‌گرفت، اجرای آمارگیری‌ها در زمان شیوع بیماری با محدودیت مواجه شد. آمارگیری نیروی کار با مصاحبه‌ی رودرروی پرسش‌گر با خانوارهای نمونه و گردآوری و ثبت اطلاعات با استفاده از تبلت هنگام انجام مصاحبه اجرا می‌شود. با توجه به اهمیت گردآوری اطلاعات در خصوص بازار کار، به‌ویژه در زمان شیوع بیماری کووید ۱۹ که اثرات اجتماعی- اقتصادی فراوانی ایجاد

کرد، گردآوری اطلاعات به شیوه‌ی مصاحبه‌ی تلفنی به‌عنوان یک راهکار توسط مرکز آمار ایران در نظر گرفته شد. بر این اساس تکمیل پرسش‌نامه‌های آمارگیری نیروی کار از فصل بهار ۱۳۹۹، به شیوه‌ی رودررو، تلفنی یا ترکیبی از این دو روش انجام گرفت. با فرض امکان برقراری ارتباط با خانوارهای نمونه‌ی مورد نظر به‌صورت تلفنی، برخی دیگر از مشکلات و چالش‌های احتمالی در اجرای آمارگیری با توجه به تغییر سریع شیوه‌ی آمارگیری ایجاد می‌شود که می‌تواند بر نتایج اثرگذار باشد. لذا تغییر الگوی خطاهای اندازه‌گیری به دلیل تغییر در نحوه‌ی پرسش سؤال‌ها، روش ورود اطلاعات، نحوه‌ی بازبینی، شیوه‌ی کنترل اطلاعات و ... ایجاد می‌شود. در ادامه برآورد خطای اندازه‌گیری با استفاده از مدل‌های رده‌ی نهان مارکوف با به‌کارگیری برنامه ℓEM [۹] که یک برنامه برای تحلیل داده‌های رسته‌ای اسمی، ترتیبی و یا سطوح بازه‌ای است و لحاظ الگوی چرخش نمونه‌ها در آمارگیری نیروی کار ایران محاسبه می‌شود.

همان‌طور که در بخش قبل اشاره شد، در مدل‌های رده‌ی نهان مارکوف سه دوره‌ی متوالی برای مشاهدات در نظر گرفته شده و برآورد پارامترها بر این اساس به‌دست آمده است، در صورتی‌که الگوی چرخش نمونه‌ها در آمارگیری نیروی کار ایران ۲-۲-۲ است. در الگوی چرخش به کار گرفته‌شده توسط ایران بعد از تکرار خانوارها در دو فصل متوالی، خانوارها برای دو فصل از نمونه خارج می‌شوند و دوباره برای دو فصل وارد نمونه می‌شوند به‌عبارتی گام‌های بین فصل‌های مشترک یکسان نیستند. بر اساس مطالعه‌ی [۱۵] برای به‌کارگیری مدل رده‌ی نهان مارکوف در داده‌های نیروی کار ایران استفاده از داده‌های چهار دوره‌ی X_1، X_2، X_3 و X_4، با فرض برابری احتمالات $P(X_2|X_1) = P(X_4|X_3)$ در نظر گرفته می‌شود.

در این سناریو [۱۵]، از فرض احتمال‌های انتقال برابر استفاده نموده و با راهکاری مشکل عدم برابری گام‌ها بین دوره‌ها را برطرف کرده‌اند. طبق الگوی چرخش ۲-۲-۲ برای هر مشاهده چهار تکرار وجود دارد ولی همان‌طور که گفته شد فاصله زمانی بین تکرارها یکسان نیست (دو تکرار اول دو فصل پشت سرهم، تکرار دوم و سوم شش ماه فاصله زمانی و تکرار سوم و چهارم دوباره دوفصل پشت سر هم نمونه‌های مشترک قرار می‌گیرند). یک فرض اضافه در مدل در نظر گرفته شده است، به‌این صورت که فرض می‌شود بین دوره‌های دوم و سوم تغییر وضعیتی رخ نداده است. بنا بر این، در داده‌های اصلی مشاهداتی در نظر گرفته می‌شوند که وضع فعالیت آن‌ها در فصل دوم و سوم یکسان هستند. این فرض اولا سبب می‌شود مشکل گام‌های بین دوره‌ها حل شود و مهم‌تر این که بسیاری از تغییر وضعیت‌هایی که در بین تکرارهای دوم و سوم رخ می‌دهد، به علت طولانی بودن زمان بین دو دوره واقعی است و نباید به‌عنوان خطا لحاظ شود و فرض مذکور این مورد را نیز لحاظ می‌کند. بنا بر این در سناریوی مذکور با لحاظ برابری احتمال‌های انتقال بین دوره‌ها، چهار دوره در سه دوره تعریف می‌شود. با توجه به این‌که استفاده از داده‌های موزون تأثیر معنی‌داری بر برآوردهای نسبی ندارد و باعث پیچیدگی محاسبات می‌شود، از داده‌های بدون وزن استفاده شده است.

بر اساس تعاریف بین‌المللی جمعیت در سن کار به سه رده‌ی شاغل، بیکار و غیرفعال تقسیم می‌شود. در اجرای آمارگیری، در عمل بنا به عوامل متعددی از قبیل طراحی پرسش‌نامه، اثر مصاحبه‌گر، اثر پاسخ‌گو و سوء تعبیر پرسش در آمارگیری، خطای اندازه‌گیری به‌وجود می‌آید و فردی که در واقع شاغل است، بیکار طبقه‌بندی می‌شود، یا فردی که در واقع بیکار است، غیرفعال تشخیص داده می‌شود. تغییر شیوه‌ی آمارگیری در زمان شیوع کووید ۱۹ می‌تواند منبع جدیدی برای بروز خطا باشد. بنابراین در ادامه داده‌های مربوط به ۳ دوره کامل

۱. بهار ۱۳۹۷- تابستان ۱۳۹۷ـبهار ۱۳۹۸- تابستان ۱۳۹۸
۲. بهار ۱۳۹۸- تابستان ۱۳۹۸- بهار ۱۳۹۹ـتابستان ۱۳۹۹
۳. بهار ۱۳۹۹- تابستان ۱۳۹۹- بهار ۱۴۰۰ـتابستان ۱۴۰۰

مورد بررسی قرار گرفته است. دوره‌ی آمارگیری بهار ۱۳۹۷- تابستان ۱۳۹۷ـبهار ۱۳۹۸- تابستان ۱۳۹۸، به‌صورت کامل مربوط به دوره‌ی قبل از شیوع بیماری کووید ۱۹ است که آمارگیری به شیوه‌ی گذشته و بدون تغییر اجرا می‌شده است. در دوره‌ی آمارگیری بهار ۱۳۹۸- تابستان ۱۳۹۸- بهار ۱۳۹۹ـتابستان ۱۳۹۹، دو فصل اول آمارگیری به شیوه‌ی گذشته انجام می‌گرفته است و در

https://dx.doi.org/10.22108/msci.2025.142988.1700

دو فصل دوم، آمارگیری در شرایط جدید با در نظر گرفتن مصاحبه‌های تلفنی انجام شده است. در دوره‌ی آمارگیری بهار ۱۳۹۹- تابستان ۱۳۹۹- بهار ۱۴۰۰-تابستان ۱۴۰۰، آمارگیری در همه فصل‌ها با در نظر گرفتن شیوه‌نامه‌های دوران همه‌گیری کووید ۱۹ انجام شده است. مقایسه خطای اندازه‌گیری در این سه دوره‌ی زمانی می‌تواند معیاری برای ارزیابی اثر تغییر شیوه برخطای اندازه‌گیری فراهم کند.

جدول ۲، ماتریس احتمالات واقعی $\pi_{(X_2|X_1)}$ بر اساس مدل‌سازی رده‌ی نهان مارکوف با فرض تساوی احتمال جابه‌جایی وضعیت شغلی بین فصل‌های بهار ۱۳۹۷و تابستان ۱۳۹۷و بین فصل‌های بهار ۱۳۹۸ و تابستان ۱۳۹۸ به‌دست آمده است. اگر این فرض درست باشد، جدول ۲ نشان می‌دهد که در حقیقت ۹۶٫۲٪ شاغلین بهار ۱۳۹۷در فصل بعدی شاغل باقی مانده‌اند. همین‌طور ۷۰٫۲٪ بیکاران بهار ۱۳۹۷در فصل بعدی بیکار باقی مانده‌اند.

جدول ۲. ماتریس احتمال‌های انتقال بدون خطای طبقه‌بندی در بهار به تابستان ۱۳۹۷

Table 2: Matrix of transition probabilities without classification error from spring to summer 2018

تابستان ۱۳۹۷				
N	U	E		
۰٫۰۳۲	۰٫۰۰۷	۰٫۹۶۲	E	
۰٫۰۹۵	۰٫۷۰۲	۰٫۲۰۴	U	بهار ۱۳۹۷
۰٫۹۷۳	۰٫۰۰۵	۰٫۰۲۳	N	

جدول ۳، ماتریس خطای اندازه‌گیری از بهار سال ۱۳۹۷به تابستان ۱۳۹۷را نشان می‌دهد. به‌طور کلی عناصر خارج قطر فرعی این ماتریس خطاهای اندازه‌گیری را برآورد می‌نماید.

جدول ۳. ماتریس خطای اندازه‌گیری وضعیت نیروی کار در بهار به تابستان ۱۳۹۷

Table 3: Measurement error matrix of labor force status from spring to summer 2018

تابستان ۱۳۹۷				
N	U	E		
۰٫۰۱۱	۰٫۰۰۶	۰٫۹۸۴	E	
۰٫۰۴۵	۰٫۸۷۶	۰٫۰۷۹	U	بهار ۱۳۹۷
۰٫۹۸۸	۰٫۰۰۳	۰٫۰۰۹	N	

بر اساس این جدول فردی که در بهار ۱۳۹۷ در واقع شاغل است ۰٫۶٪ احتمال دارد، به خطا بیکار گزارش شود. همین‌طور ۱٫۱٪ احتمال دارد به خطا غیر فعال گزارش شود. به‌همین ترتیب فردی که در حقیقت بیکار باشد ۷٫۹٪ احتمال دارد به غلط شاغل به حساب بیاید و ۴٫۵٪ احتمال دارد به غلط غیرفعال گزارش شود. بر اساس جداول ۴ و ۵، ماتریس احتمالات واقعی $\pi_{(X_2|X_1)}$ بر اساس مدل‌سازی رده‌ی نهان مارکوف با فرض تساوی احتمال جابه‌جایی وضعیت شغلی بین فصل‌های بهار ۱۳۹۸و تابستان ۱۳۹۸و بین فصل‌های بهار ۱۳۹۹ و تابستان ۱۳۹۹ به‌دست آمده است. بنابراین جدول ۴ نشان می‌دهد که در حقیقت ۹۵٪ شاغلین بهار ۱۳۹۸در فصل بعدی شاغل باقی مانده‌اند. همین‌طور ۶۷٫۱٪ بیکاران بهار ۱۳۹۸در فصل بعدی بیکار باقی مانده‌اند.

جدول ۵، ماتریس خطای اندازه‌گیری از بهار سال ۱۳۹۸به تابستان ۱۳۹۸را نشان می‌دهد.

بر اساس این جدول فردی که در بهار ۱۳۹۸ در واقع شاغل است ۰٫۸٪ احتمال دارد، به خطا بیکار گزارش شود. همین‌طور ۰٫۸٪ احتمال دارد به خطا غیر فعال گزارش شود. به‌همین ترتیب فردی که در حقیقت بیکار باشد ۰٫۶٪ احتمال دارد به غلط شاغل به حساب بیاید و ۸٫۳٪ احتمال دارد به غلط غیرفعال گزارش شود.

جدول ۴. ماتریس احتمال‌های انتقال بدون خطای طبقه‌بندی در بهار به تابستان ۱۳۹۸

Table 4: Matrix of transition probabilities without classification error from spring to summer 2019

تابستان ۱۳۹۸				
N	U	E		
۰٫۰۴۱	۰٫۰۰۹	۰٫۹۵۰	E	
۰٫۱۳۸	۰٫۶۷۱	۰٫۱۹۱	U	بهار ۱۳۹۸
۰٫۹۷۹	۰٫۰۰۴	۰٫۰۱۶	N	

جدول ۵. ماتریس خطای اندازه‌گیری وضعیت نیروی کار در بهار به تابستان ۱۳۹۸

Table 5: Measurement error matrix of labor force status from spring to summer 2019

تابستان ۱۳۹۸				
N	U	E		
۰٫۰۰۸	۰٫۰۰۸	۰٫۹۸۳	E	
۰٫۰۸۳	۰٫۸۵۷	۰٫۰۶۰	U	بهار ۱۳۹۸
۰٫۹۹۱	۰٫۰۰۲	۰٫۰۰۷	N	

جداول ۶ و ۷، از ماتریس احتمالات واقعی با فرض تساوی احتمال جابجایی وضعیت شغلی بین فصل‌های بهار ۱۳۹۹و تابستان ۱۳۹۹ و بین فصل‌های بهار ۱۴۰۰ و تابستان ۱۴۰۰ به‌دست آمده است. بنابراین جدول ۶ نشان می‌دهد که در حقیقت ۹۵٫۸٪ شاغلین بهار ۱۳۹۹در فصل بعدی شاغل باقی مانده‌اند. همین‌طور ۶۵٫۳٪ بیکاران بهار ۱۳۹۹در فصل بعدی بیکار باقی‌مانده‌اند.

جدول ۶. ماتریس احتمال‌های انتقال بدون خطای طبقه‌بندی در بهار به تابستان ۱۳۹۹

Table 6: Matrix of transition probabilities without classification error from spring to summer 2020

تابستان ۱۳۹۹				
N	U	E		
۰٫۰۳۶	۰٫۰۰۶	۰٫۹۵۸	E	
۰٫۱۰۴	۰٫۶۵۳	۰٫۲۴۳	U	بهار ۱۳۹۹
۰٫۹۷۳	۰٫۰۰۶	۰٫۰۲۱	N	

جدول ۷، ماتریس خطای اندازه‌گیری از بهار سال ۱۳۹۹به تابستان ۱۳۹۹را نشان می‌دهد.

جدول ۷. ماتریس خطای اندازه‌گیری وضعیت نیروی کار در بهار به تابستان ۱۳۹۹

Table 7: Measurement error matrix of labor force status from spring to summer 2020

تابستان ۱۳۹۹				
N	U	E		
۰٫۰۱	۰٫۰۰۶	۰٫۹۸۴	E	
۰٫۰۴۷	۰٫۸۹۱	۰٫۰۶۱	U	بهار ۱۳۹۹
۰٫۹۹۳	۰٫۰۰۲	۰٫۰۰۵	N	

https://dx.doi.org/10.22108/msci.2025.142988.1700

بر اساس این جدول فردی که در بهار ۱۳۹۹ در واقع شاغل است ۰/۶٪ احتمال دارد، به خطا بیکار گزارش شود. همین‌طور ۱٪ احتمال دارد به خطا غیرفعال گزارش شود. به‌همین ترتیب فردی که در حقیقت بیکار باشد ۶/۱٪ احتمال دارد به غلط شاغل به حساب بیاید و ۴/۷٪ احتمال دارد به غلط غیرفعال گزارش شود.

جدول‌های ۳، ۵ و ۷ نشان‌دهنده‌ی ماتریس‌های خطای اندازه‌گیری جابه‌جایی وضعیت نیروی کار برای دوره‌های فصل بهار ۱۳۹۷به تابستان ۱۳۹۷، بهار ۱۳۹۸به تابستان ۱۳۹۸ و بهار ۱۳۹۹به تابستان ۱۳۹۹هستند که همان‌طور که پیش‌تر ذکر شد برای برآورد آن‌ها از اطلاعات چهار دوره‌ی آمارگیری که با در نظر گرفتن فرض برابری احتمال انتقال بین‌دوره‌های آمارگیری و حفظ واحدهای بدون تغییر وضعیت از فصل دوم به فصل سوم، استفاده شده است. عناصر خارج از قطر فرعی این ماتریس در واقع برآورد خطای اندازه‌گیری وضعیت نیروی کار بین دو دوره‌ی مورد مطالعه را نشان می‌دهد. می‌توان مجموع عناصر غیرقطری هر یک از این ماتریس‌ها را به‌عنوان برآورد کل خطای اندازه‌گیری در نظر گرفت. بر این اساس، برآورد کل خطای اندازه‌گیری برای دوره‌ی بهار ۱۳۹۷به تابستان ۱۳۹۷ که پیش از شروع همه‌گیری کووید ۱۹ است، برابر با ۰/۱۵۳٪، برای دوره‌ی بهار ۱۳۹۸به تابستان ۱۳۹۸ که همزمان با شیوع همه‌گیری کووید ۱۹ و تغییر ناگهانی شیوه‌ی آمارگیری است برابر با ۰/۱۶۸٪، و برای دوره‌ی بهار ۱۳۹۹به تابستان ۱۳۹۹ برابر با ۰/۱۳۱٪ است. از بهار ۱۴۰۰ با کاهش شیوع بیماری کووید ۱۹ و واکسیناسیون سراسری به دلیل کاهش محدودیت‌ها آمارگیری به‌شیوه‌ی گذشته و به‌صورت رودررو انجام شد، بنا بر این بررسی میزان خطای اندازه‌گیری به دلیل تغییر شیوه‌ی آمارگیری از مصاحبه‌ی رودررو به تلفنی موضوعیت ندارد.

بنابراین به نظر می‌رسد با تغییر شیوه‌ی آمارگیری، خطای اندازه‌گیری در ابتدا با افزایش مواجه شده است ولی با تداوم اجرای شیوه‌ی جدید خطای اندازه‌گیری حتی نسبت به زمانی که آمارگیری فقط به‌صورت رودررو انجام می‌شده، کاهش یافته است. البته نکته‌ای که در تفسیر خطای اندازه‌گیری باید مورد توجه قرار گیرد این است که، در این روش منبع بروز خطای اندازه‌گیری لزوماً تغییر شیوه‌ی آمارگیری نیست و تغییرات می‌تواند ناشی از عوامل دیگر نیز باشد. به‌عنوان مثال، یکی از دلایل کاهش خطای اندازه‌گیری در سال ۱۳۹۹در مقایسه با سال‌های ۱۳۹۷ و ۱۳۹۸می‌تواند شیوع همه‌گیری کووید ۱۹ و کمتر شدن جابجایی افراد یا به‌عبارتی ثبات وضعیت نیروی کارشان به دلیل ماندگاری بیماری کووید ۱۹ باشد. در جدول ۶ ملاحظه می‌شود احتمال اینکه فردی که در بهار ۱۳۹۹ بیکار بوده، در تابستان ۱۳۹۹ نیز بیکار باشد، نسبت به دوره‌های قبلی تحت مطالعه بیشتر است.

۵. بحث و نتیجه‌گیری

آمارگیری نیروی کار ایران یکی از مهم‌ترین آمارگیری‌های مرکز آمار ایران است که لغو اجرای آن به دلیل اطلاعات ارزشمندی که در خصوص وضعیت بازار کار ارائه می‌دهد، حتی در شرایط بحرانی، میسر نیست. بنابراین، لحاظ کردن تمهیدات لازم برای اجرای آن در شرایط بحرانی مانند همه‌گیری بیماری کووید ۱۹ اجتناب‌ناپذیر است. از سوی دیگر، باید این نکته را مد نظر قرار داد که تغییرات ناگهانی در شیوه‌ی گردآوری اطلاعات بدون اینکه فرصت لازم برای آزمایش و ارزیابی آن‌ها فراهم باشد، می‌تواند اثرات قابل توجهی بر کیفیت نتایج آمارگیری داشته باشد. یکی از مهم‌ترین خطاهای غیرنمونه‌گیری، خطای اندازه‌گیری است که معرف اختلاف بین پاسخ دریافت‌شده از پاسخ‌گو و مقدار واقعی متغیر مورد سؤال است. بدیهی است به دلیل ماهیت این نوع خطا و دسترسی نداشتن به پاسخ‌های واقعی، محاسبه آن با دشواری همراه است و معمولاً نیازمند دسترسی به اطلاعات ثبتی با شرایط خاص یا تکرار بخشی از آمارگیری با دقت بسیار بالا است که حتی اگر در عمل قابل پیاده‌سازی باشد، منجر به صرف زمان و هزینه خواهد شد. مدل رده‌ی نهان مارکوف امکان برآورد خطای اندازه‌گیری برای متغیرهای رسته‌ای در آمارگیری‌های پانلی که حداقل سه دوره تکرار دارند را بدون نیاز به مصاحبه‌ی مجدد فراهم می‌کند. در آمارگیری نیروی کار ایران که با نمونه‌گیری چرخشی با الگوی ۲-۲-۲ اجرا می‌شود، امکان استفاده از مدل رده‌ی نهان مارکوف فراهم است. در این مقاله، به‌منظور ارزیابی نتایج آمارگیری نیروی کار در شرایط همه‌گیری کووید ۱۹ که شیوه‌ی اجرای آمارگیری به یک‌باره از مصاحبه‌ی رودررو به مصاحبه‌ی تلفنی یا ترکیبی از این

دو تغییر یافت، مدل رده‌ی نهان مارکوف برای ارزیابی خطای اندازه‌گیری با استفاده از داده‌های چهار دوره‌ی آمارگیری از بهار ۱۳۹۷ تا تابستان ۱۴۰۰ به کار گرفته شد.

بر این اساس، اختلاف متوسط احتمال ثبات مورد انتظار و احتمال عدم تغییر وضعیت مشاهده‌شده در وضع فعالیت بین دو فصل متوالی به‌عنوان معیاری برای ارزیابی میزان خطای اندازه‌گیری با توجه به تغییر شیوه‌ی آمارگیری در نظر گرفته شد. نتایج حاکی از آن بود که خطای اندازه‌گیری در دوره‌ی اول تغییر شیوه‌ی مصاحبه با افزایش همراه بوده ولی در دوره‌های بعدی کاهش یافته است. علاوه بر آن مشاهده شد که سهم شاغلین و بیکارانی که به جمعیت غیر فعال پیوسته‌اند، در ابتدای همه‌گیری کووید ۱۹ قابل توجه است ولی جالب این است که با طولانی شدن اپیدمی افراد شاغل بیشتر ترجیح داده‌اند شاغل بمانند تا بیکار یا غیرفعال شوند و ریسک در جستجوی کار بهتر را داشته باشند. سهم بیکاران به غیرفعالان نیز کاسته شده است، به‌عبارتی دوباره مشکلات اقتصادی و وضعیت نیروی کار در حال بازگشت به زمان قبل از همه‌گیری کووید ۱۹ است.

بر اساس نتایج این مطالعه می‌توان اظهار داشت، مصاحبه‌ی تلفنی می‌تواند شیوه‌ی مناسبی برای گردآوری داده‌های آمارگیری نیروی کار باشد. همچنین از آن‌جایی‌که به‌طور معمول آمارگیری در ساعات کاری روز انجام می‌گیرد و احتمال می‌رود تمام افراد خانوار به‌خصوص افراد شاغل در منزل حضور نداشته باشند و امکان مراجعه‌های مکرر نیز اغلب وجود ندارد، استفاده از تلفن به‌عنوان ابزاری برای گردآوری داده‌ها امکان مصاحبه با افراد شاغل را فراهم می‌کند و به این‌ترتیب می‌تواند میزان پاسخ‌گویی جانشین را کاهش دهد. دلیل دیگری که آمارگیری تلفنی می‌تواند در این طرح بازدهی مثبتی داشته باشد، ثابت بودن بخشی از واحدهای نمونه در هر فصل از آمارگیری است. همچنین، بخش مهمی از هزینه‌ی اجرای آمارگیری نیروی کار صرف ایاب و ذهاب مربوط به تکمیل پرسش‌نامه‌ها و نظارت بر اجرای آمارگیری می‌شود. استفاده از تلفن علاوه بر اینکه می‌تواند منجر به کاهش هزینه‌های مربوط به ایاب و ذهاب و حمل و نقل به مناطق پرت و دور از دسترس شود، ابزارهای نظارتی دقیق‌تری در اختیار مدیران آمارگیری قرار می‌دهد.

از سوی دیگر، محبوبیت روزافزون تلفن‌های همراه، روش ارزیابی پوشانش تلفنی را تغییر می‌دهد. پیش از این فقط دو نوع خانوار وجود داشت: آن‌هایی که تلفن داشتند و آن‌هایی که تلفن نداشتند. ظهور تلفن‌های همراه گروه‌های جدیدی را ایجاد کرده است: کسانی که تلفن همراه، تلفن ثابت یا هر دو را دارند. طراحی مناسب نمونه‌ی آمارگیری تلفنی، نیازمند دانش ساختار پوشانش تلفنی است. با این‌حال، در بسیاری از کشورها تلفن‌های همراه در حال حاضر آنقدر محبوب هستند که باید در همه آمارگیری‌های تلفنی مورد توجه قرار گیرند. اگر آن‌ها از چارچوب نمونه‌گیری کنار گذاشته شوند، نتایج به‌دست آمده از آمارگیری‌ها اریب خواهند بود و میزان اریبی به نسبت خانوارهای جامعه‌ی هدف که فقط تلفن همراه دارند و به موضوع آمارگیری بستگی دارد. چارچوب‌های نمونه‌گیری شماره تلفن ثابت و تلفن همراه متفاوت است، شماره تلفن ثابت یک چارچوب خانواری و شماره تلفن همراه یک چارچوب در سطح فردی ایجاد می‌کند. باید توجه کرد که آمارگیری تلفنی به لحاظ روش‌شناسی با آمارگیری که در آن فقط مصاحبه‌ها به‌صورت تلفنی انجام می‌شود، تفاوت دارد و لازم است زیرساخت‌های لازم برای طراحی و اجرای آمارگیری تلفنی فراهم شود.

مراجع

[1] P. Biemer, *Latent class analysis of survey error*, Wiley Series in Survey Methodology, John Wiley & Sons, Inc., Hoboken, NJ, 2011.

[2] P. Biemer and M. Bushery, On the validity of markov latent class analysis for estimating classification error in labor force data, *Surv. Methodol.*, **26** no. 2 (2000) 139–152.

https://dx.doi.org/10.22108/msci.2025.142988.1700

[3] International Labour Organization (ILO), COVID-19: Guidance for labour statistics data collection, Guidance to data producers to maintain labour force survey data collection, *ILO Technical Note*, Geneva: ILO, 2020.

[4] K. Khalagi, M. A. Mansournia, A. Rahimi-Movaghar, K. Nourijelyani, M. Amin-Esmaeili, A. Hajebi, V. Sharif, R. Radgoodarzi, M. Hefazi and A. Motevalian, Assessing measurement error in surveys using latent class analysis: application to self-reported illicit drug use in data from the Iranian Mental Health Survey, *Epidemiology and Health*, **38** (2016) 1–11.

[5] B. D. Meyer and N. Mittag, An empirical total survey error decomposition using data combination, *J. Econometrics*, **224** no. 2 (2021) 286–305.

[6] C. A. Poulsen, *Latent structure analysis with choice modeling applications*, Arhus School of Business Administration and Economics, Arhus, Denmark, 1982.

[7] F. Van de Pol and R. Langeheine, Mixed markov latent class models, in C.C. Clogg ed., *Sociological Methodology*, **20** (1990) 213–247.

[8] F. Van de Pol and J. De Leeuw, A latent markov model to correct for measurement error, *Sociological Methods and Research*, **15** no. 1-2 (1986) 118–141.

[9] J. K. Vermunt, *LEM: a general program for the analysis of categorical data*, Department of Methodology and Statistics, Tilburg University, 1997.

[10] L. M. Wiggins, *Panel analysis*, latent probability models for attitude and behavior processing, Elsevier SPC, Amsterdam, 1973.

[۱۱] ز. فلاح محسن‌خانی و ل. کلهری ندرآبادی. استفاده از مدل‌های رده نهان مارکوف دربرآورد خطای رده‌بندی آمارهای جریان نیروی کار ایران، مجله علوم آماری، **۱۶** (۱۴۰۱) ۱۸۹–۲۰۷.

[۱۲] ح. نواب‌پور، ط. چگینی و ا. صفرنژاد بروجنی، کاربردی از برآورد خطای کل آمارگیری در یک آمارگیری کوچک‌مقیاس، مجله‌ی بررسی‌های آمار رسمی ایران، **۲۵** (۱۳۹۳) ۴۷–۶۸.

[۱۳] ح. نواب‌پور، م. صفاکیش و غ. ایزدی، آمار رسمی، پژوهشکده‌ی آمار، تهران، ۱۳۹۷.

[۱۴] ح. نواب‌پور، ا. صفرنژاد بروجنی و ط. چگینی، کاربردی از ارزیابی خطای اندازه‌گیری به‌روش تحلیل ردهٔ نهان، اندیشه آماری، **۲۲** (۱۳۹۶) ۸۵–۹۶.

[۱۵] ف. مهران، م. نورمحمدی، ز. رضایی قهرودی، ز. فلاح محسن‌خانی، ل. کلهری ندرآبادی، ا. یزدانی، م. ریحانی، ت. باغ فلکی، م. نقی‌خانی، ن. یوسفی و ر. علی‌اکبری صبا، شناسایی و کنترل خطاهای بی‌پاسخی و اندازه‌گیری طرح‌های آمارگیری، پژوهشکده‌ی آمار، تهران، ۱۳۹۶.

لیدا کلهری ندرآبادی
گروه طرح‌های فنی و روش‌های آماری، پژوهشکده‌ی آمار، تهران، ایران
Kalhori@srtc.ac.ir، lidakalhori@yahoo.com

لیدا کلهری ندرآبادی در سال ۱۳۸۰ وارد مقطع کارشناسی رشته آمار دانشگاه علامه‌طباطبایی و در سال ۱۳۸۴ وارد مقطع کارشناسی ارشد رشته آمار محض دانشگاه اصفهان شد. او مقطع دکتری خود را در سال ۱۳۹۶ در رشته‌ی آمار دانشگاه تربیت مدرس به اتمام رسانید.

پریا ترابی کهلان
گروه طرح‌های فنی وروش‌های آماری، پژوهشکده‌ی آمار، تهران، ایران
torabiparya@yahoo.com

پریا ترابی کهلان در سال ۱۳۸۴ وارد مقطع کارشناسی رشته‌ی آمار دانشگاه تبریز و در سال ۱۳۹۱ وارد مقطع کارشناسی ارشد آمار ریاضی دانشگاه علم و صنعت ایران شد. او مقطع دکتری خود را در سال ۱۴۰۰ در رشته‌ی آمار دانشگاه علامه طباطبایی به اتمام رسانید.

https://dx.doi.org/10.22108/msci.2025.142988.1700