Valley International Journals
Open Access Journal
New Thinking New Innovation

INTERNATIONAL JOURNAL OF MEDICAL SCIENCE AND CLINICAL INVENTIONS
Volume 2 issue 06 2015 page no. 892-901 ISSN: 2348-991X
Available Online At: http://valleyinternational.net/index.php/our-jou/ijmsci

Estimation Of Occupational Accident Rate Using Raking Ratio Method

Young Sun Kim [1], Kyung Yong Rhee [2]
[1,2]Safety and Health Policy Research Department, Occupational Safety and Health Research Institute, KOSHA, Ulsan 681-230, Korea

Correspondence : Kyung Yong Rhee
Occupational Safety and Health Research Institute, Occupational Safety and Health Agency, Ulsan, Korea
E-mail: appleyskim@gmail.com

Abstract:
Objectives:
While the occupational accident rate is calculated based on the Industrial Accidents Compensation Data in Korea, because there are establishments without registered insurance, it is necessary to revise the estimation method. This study adjusted the employee proportion distribution to estimate the occupational accident rate.
Methods:
The employee distribution within the Industrial Accidents Compensation Data by industrial sector, region and size was approximated to the population distribution using the raking ratio method, which was estimated through error rate against the Economically Active Population Survey data. The occupational accident rate was calculated through each stratum and adjusted employee proportion.
Results:
As a result of the raking ratio method, when employee proportions within the industrial sector were compared to the Economically Active Population Survey, the error and occupational accident rates were reduced. However, there was a problem with under-sampling of the construction industry establishment in the survey population where the raking ratio method was used. As a result of analysis by region and establishment size, the error rate was reduced while the occupational accident rate increased.
Conclusion:
As shown in this study, when data are incomplete, the reliability index can still be estimated using population data. In particular, when the occupational accident rate is calculated based on industrial accident compensation insurance data, the index could be diversified. Further study will examine cases of incomplete population data.

Keyword: Occupational Accident, Raking Ratio Method, Workers' Compensation Insurance

I. INTRODUCTION

Occupational accidents have increased as employees have become exposed to new environments and unknown risk factors associated with increased industrial development. In Korea, occupational accident statistics are based on application of the Industrial Accident Compensation Insurance Act. At its inception in early 1964, the Act was applied to establishments having more than 500 employees in the mining and manufacturing industries. Later in 1989, the Act was expanded to include all establishments and then expanded again in 2000 to cover all businesses having at least 1 full-time employee. Nonetheless, a review of the data published in the Census on Establishments and Economically Active Population Survey charting the distributional proportions of employees and establishments across industrial sector, establishment size and region shows that Korea's representative population data differ from data on industrial accidents, calling into question the reliability of occupational accident status statistics. For example, there are many cases of unregistered occupational accident insurance claim for short-term construction operations that were closed as soon as the objective was achieved. Accordingly, it is necessary to estimate an occupational accident index while reconciling data on industrial sector, establishment size and regional distribution from Industrial Accidents Compensation Data with information about the overall population distribution.

If occupational accident insurance is not calculable for all employees, then industrial accident insurance data can be regarded not as population data but as sample data. A sample, however, might deviate from the population distribution due to the extraction and response rates in the design process. In other words, these deviations may be due to the employees who were not captured in the industrial accident insurance dataset. Fortunately, this difference can be adjusted by weight using extraction probability to obtain a representative population sample. Without an adjustment, an unbiased parameter estimate is obtained and distorted statistics may result. Thus, a weight adjustment procedure is required to control bias (Kalton, 1983).

Weights are constructed on design or benchmarking calibration variables. In the past, weights were calculated easily in sample design processes; however, modern methods are more complex and require the consideration of many factors. Deville and Sarndal's (1992) calibration estimation theory is a well-known adjustment method of estimation using auxiliary variables to modify weights (Deville et al., 1992). This theory has also been applied to post-stratification procedures such as the raking ratio method. The raking ratio method and general regression estimator are representative estimation methods using auxiliary information. The advantages of using auxiliary information include reliability of estimation and processing cost reduction, strengths that have been verified through the Canadian Monthly Wholesale and Retail Trade Survey (Hidiroglou et al., 2006). In this study, the raking ratio method ("RRM") was used to make weight adjustments to account for differences in the industrial sector, establishment size and geographical region of employees between the target population and the Industrial Accident Compensation Data.

RRM is an adjustment method commonly used in survey sampling data in which population distributions are widely used for auxiliary information (Tillé, 1998). RRM can also be used for expansion to multivariate (Oh, 1978) or non-linear parameter estimation for Taylor linearization in terms of partial derivatives (Demnati, 2004). RRM is used extensively in large data surveys. The 1991 population census in Canada, for example, used evaluations of weight area level (Bankier et al., 1992), and the Health Interview Survey carried out in California used RRM as the final stage against 12 adjusted weights for benchmarking sampling weight (Cervantes, 2008). Additionally, the New York City Social Indicators Survey developed an estimation method to effectively improve sample and non-sample biases based on RRM (Hao Lu et al., 2003). In the US census, RRM has been used for error estimation and adjustment arising from Dual System Estimation (differences between census and post-census survey) (Haines et al., 1998). RRM has also been effectively used to estimate drug use in the entire US and smaller regions. Recently, new methods have been introduced based on RRM. Generalized Raking Procedures in Survey Sampling, for example, complimented limitations to the existing RRM (Deville et al., 1993), and studies on cases of non-convergence when zero is included in the sample cell have also been conducted (Kalton et al., 1998).

In this study, the proportions of employees by industrial sector, geographical region and establishment size were adjusted via RRM to minimize the proportional differences in survey data between the Industrial Accident Compensation Data and The Census on Establishments. By adjusting for omitted establishments in the Industrial Accident Compensation Data, this study attempted to minimize sample bias and to increase accuracy of the occupational accident index.

II. MATERIALS AND METHOD

1 Material

1.1 Industrial Accidents Compensation Data

Occupational accidents have increased as employees have become more exposed to new environments and unknown risk factors in conjunction with increased industrial development. Occupational accidents include job-related death, injury or disease due to structures, facilities, raw materials, gas, steam or dust related to employees' job functions or by other works. In Korea, an industrial victim is defined as an employee on the job fatality or an instance of more than 4 days of medical attention due to occupational-related injuries or diseases. The Industrial Accidents Compensation Data is a database of establishments that have reported industrial accident insurance claims and industrial accident victims. In 2011, the number of such establishments stood at 1,738,196, and the number of employees was 14,362,372. Furthermore, the number of reported industrial victims was 93,292, yielding an occupational accident rate of 0.65%. The purpose of this study was to revise and evaluate employee Industrial Accidents Compensation Data by weighting employee distribution and then estimating the occupational accident rate.

1.2 The Census on Establishments

This study also incorporated population data from The Census on Establishments in analyzing RRM. The Census on Establishments is a basic database used for national and local government policy planning and business management planning for private enterprises, as well as academic studies and population data for various statistical surveys of establishments to identify their related industrial sectors, employment status and regional characteristics. However, the database excludes private agricultural, forestry and fishery establishments (establishments run by corporations or non-corporations are subject to survey), national defense, household service, international and foreign organizations and temporary sales having no fixed facility or place of sales.

1.3 Economically Active Population Survey

To evaluate RRM results, data from the Economically Active Population Survey were used. Economically Active Population Survey data describe economic characteristics such as economic activity, employment and unemployment to provide basic information for macroeconomic analysis and human resource development policy such as labor supply, employment structure, available working hours and human resource utilization levels. It also provides basic data for government employment policy planning and evaluation. Despite its limited number of samples, the population survey is advantageous in many ways. First, it is seasonally adjusted through monthly survey by interlocked samples. Second, the survey is unaffected by closure of an establishment because the survey is made of the household. Third, the survey data can approximate population distributions as random sampling even though it is not a complete enumeration-based survey (Census). Accordingly, even though the Economically Active Population Survey cannot be used for sample survey or sampling frame, it is used in this study as the criteria for evaluation under the assumption that it approximates the population distribution.

2. Raking Ratio Method

RRM is the representative post hoc adjustment method used after the survey sampling. While RRM is known as marginal distribution in categorical data, when no joint distribution is known, it is a post-stratification method for estimating joint distribution by using marginal total and known marginal distributions. It adjusts to correspond with marginal strata information instead of all joined strata on the joint of variables if there are more than 2 stratification variables.

The estimation process of RRM uses the population marginal total of auxiliary variables that have strong correlations with the variable of interest. Available auxiliary variables could be data estimated in the previous stage in cases of complete enumeration survey data or survey by stage.

To explain the estimation process of raking ratio U as it pertains to the population, U is split into a cross-tabulation with $D_1 \times D_2$, while $U_{d_1 d_2}$ represents the population column with (d_1, d_2) column size in $N_{d_1 d_2}$ units, D_1 and D_2 indicate the numbers of rows and columns in each contingency table, and d_1 and d_2 represent column and row, respectively. In this study, there were three stratification variables that were cross tabulated when a variable was fixed. The sample is s, and the sample column can be expressed as $s_{d_1 d_2} = s \cap U_{d_1 d_2}$. In addition, w_k represents the survey weight of unit k. The variable of interest is y, and the unit k value is y_k. $\sum_{k \in s} w_k y_k$ is regarded as the estimation value of the population total $\sum_{k \in U} y_k$. Generally, the weight adjustment is made by revising the corresponding weight of the interest variable to calculate an estimation of the population total.

$N_{d_1 d_2}$ is the number of population units in the column defined by column d_1 and row d_2, which is an unknown variable. However, the marginal total, which is the total of the population row and column, is a known variable as defined in Equation (1).

$$N_{d_1+} = \sum_{d_2=1}^{D_2} N_{d_1 d_2}, d_1 = 1,2,\ldots,D_1, N_{d_2+} = \sum_{d_1=1}^{D_1} N_{d_1 d_2}, d_2 = 1,2,\ldots,D_2 \quad \cdots\cdots (1)$$

The unknown population column $N_{d_1 d_2}$ is then estimated from the sample using the process described below. First, it begins with the survey weighted value $w_0 = N/n$, where $N = \sum_{d_1}^{D_1} \sum_{d_2}^{D_2} N_{d_1 d_2}$ is the population size, and n is the number of the sample. The weighted values are then adjusted as illustrated below using the known population marginal total from the weighted value w_0.

$$w^{(1,1)} = \left[\frac{N_{d_1}}{\sum_{d_2=1}^{D_2} \widehat{N_{d_1 d_2}}^{(0)}} \right] \times w_0, \quad w^{(1,2)} = \left[\frac{N_{d_2}}{\sum_{d_1=1}^{D_1} \widehat{N_{d_1 d_2}}^{(1,1)}} \right] \times w^{(1,1)} \quad \cdots\cdots (2)$$

In Equation (2), $\widehat{N_{d_1 d_2}}^{(0)}$ is an estimated value of population column $N_{d_1 d_2}$ using the weighted value w_0. Then, the revised weighted value $w^{(1,1)}$ against the column is used to calculate $\widehat{N_{d_1 d_2}}^{(1,1)}$, and $\widehat{N_{d_1 d_2}}^{(1,1)}$ is also an estimated value of $N_{d_1 d_2}$. The adjusted weighted value $w^{(1,2)}$ is obtained after completion of the first repetition, and the process is repeated until the specific convergence is satisfied.

Ultimately, the adjusted weighted value becomes $w^{(f)}$, and this value is then applied to the total population estimation formula $\sum_{k \in s} w_k y_k$, completing the estimation process where y_k is the post-survey observation value.

Deming and Stephan (1940) first introduced RRM (Deming, 1940), which was also used for non-response adjustment by Oh and Scheuren (1987).

3. Analysis Design and Validation of the Raking Ratio Method

This study used the Industrial Accidents Compensation Data as sample data while The Census on Establishments data as well as Economically Active Population Survey data were used as population data. In stage 1, the marginal distribution was extracted on D_1 (type), D_2 (size), and D_3 (area) from the census on establishments data, and the cell distribution at cross sample by D_1, D_2, D_3 was extracted using RRM. Weight was calculated using the sample cell distribution extracted from D_1, D_2, D_3 and the population cell distribution. The analysis unit is establishment, and weight is calculated according to D_1, D_2, D_3 of the individual establishment. In doing so, employee distribution by D_1 (type), D_2 (size), and D_3 (area) is calculated, and the validation process is then carried out by comparing the analysis of the economically active population survey data as another population. The comparison of proportional differences within each category against variable D {type, region, size} becomes the criteria for evaluating the analytical results. A_i is the proportion of category i within the Economically Active Population Survey data, and O_i is the proportion of category i data adjusted using RRM. Accordingly, the error rate is explained by the ratio against differences within the Economically Active Population Survey Data over the design output.

$$E_R = \sum_{i=1}^{k} \sqrt{\left(\frac{A_i - O_i}{A_i} \right)^2} \quad \cdots\cdots (3)$$

4. Estimation of Occupational Accident Rate

The occupational accident rate is calculated as the ratio between the total employees N and total industrial victims M. In terms of industrial sector, region and size, when the formula for the occupational accident rate was examined, the employee proportion n_i/N, which each category possesses, indicated the weighted average of the individual category of occupational accident rate p_i having weight w_i.

$$P_{total} = \frac{m_1 + \cdots + m_k}{n_1 + \cdots + n_k} = \sum_{i=1}^{k} \frac{n_i}{N} \times \frac{m_i}{n_i} = \sum_{i=1}^{k} w_i p_i \quad \cdots\cdots (4)$$

where m_i is the number of victims in the i category, and n_i is the number of employees in the i category.

This study calculated adjusted w_i^*, and the adjusted total occupational accident rate P^* could be estimated using w_i^* instead of using w_i, which was used in the calculation of the total occupational accident rate.

III. RESULT

1. Adjusted employee distribution and occupational accident rate in the industrial sector using the raking ratio method(Table 1)

As for the proportion by industrial sector in the Industrial Accidents Compensation Data, construction (24.20%) was the highest, followed by manufacturing (23.33%) and wholesale/retail (9.38%). When compared with the proportions by industrial sector found in the Economically Active Population Survey Data, the largest difference could be confirmed as the error rate. The error rate for construction was the highest with 1.99, followed by public administration, national defense & social welfare administration (0.69) and education service (0.65). The error rate for the total industrial sector was 6.76. When RRM was conducted using The Census on Establishments data, the proportion by type of the Industrial Accidents Compensation Data was changed. For construction, it was changed to 6.04%, and the error rate was revised to 0.25 when it was compared with the Economically Active Population Survey Data. The proportion with public administration, national defense & social welfare was 4.26%, with an error rate reduced to 0.23. The error rate for total types was reduced to 5.08. For the total occupational accident rate, calculated using the employee proportion and the occupational accident rate of the industrial sector, the occupational accident rate for the industrial accident compensation insurance data was 0.65%, but occupational accident rates using adjusted proportion by type through RRM was reduced to 0.64%. This is because of the reduction in the proportion of employees in the construction where the occupational accident rate was rather high, which results in total occupational accident rate reduction.

2. Adjusted employee distribution and accident rate by region using the raking ratio method(Table 2)

As for the regional proportions within the Industrial Accidents Compensation Data, Seoul (33.05%) was the highest, followed by Gyeonggi (20.00%) and Gyeongnam (6.34%). When compared with regional proportions of the Economically Active Population Survey data, the greatest different area was Seoul (0.50), followed by Daegu (0.44) and Jeonnam (0.29). The error rate of the total area was 3.43, which was lower than the error rate for the industrial sector. When RRM was conducted using the Census on Establishments data, the employee proportion from Seoul within the industrial accidents compensation data was changed to 23.83%. When compared to the data from the Economically Active Population Survey data, the error rate was reduced to 0.08. The employee proportion of Daegu was 4.13%, and the error rate was reduced to 0.16, while the error rate for the total area was reduced to 1.72. Regarding the total occupational accident rate calculated using each regional proportion and accident rate, the accident rate in the Industrial Accident Compensation Data was 0.65%, and the occupational accident rate calculated using the adjusted regional proportion with RRM increased to 0.70%. This is because of the decrease in the proportions of employees in Seoul, where the occupational accident rate is generally lower than other regions, resulting in the increased total occupational accident rate.

3. Adjusted employee distribution and accident rate by establishment size using the raking ratio method
(Table 3)

As for employee proportions within the Industrial Accident Compensation Data by establishment size, the establishments with 1～4 employees were 14.30%, and those with 5～9 employees were 11.23%, while those with 10～99, 100～299, and over 300 employees were 39.62%, 14.00%, and 20.85%, respectively. When employee proportion by establishment size was compared to that within the Economically Active Population Survey data, the greatest difference was shown in the establishments with over 300 employees and an error rate of 0.83, and those with 100～299 employees and an error rate of 0.41. The total error rate was 1.90, which was lower than the error rates for industrial sector and geographical region. When RRM was conducted using The Census on Establishments data, the proportion by size of the establishments within the Industrial Accidents Compensation Data was changed, and the error rates in all categories were reduced. The employee proportion of the establishments with 1～4 employees was changed to 17.77%, and the error rate was reduced to 0.07 compared with the Economically Active Population Survey data. The employee proportion of the establishments with 5～9 employees was 13.49%, and the error rate was 0.21. The employee proportion of those with 10～99 employees was 42.54%, and the error rate was 0.00. The employee proportion of those with 100～299 was 12.71%, and the error rate was 0.28. The employee proportion of those with over 300 was 13.50%, and the error rate was 0.18. For the total region, the error rate was reduced to 0.75. As for the total accident rate calculated by the employee proportion of the size of the establishment and accident rate, the occupational accident rate in Industrial Accident Compensation data was 0.65%, while the occupational accident data by revised proportion by size through RRM increased to 0.73%.

Table 3. Comparison of the employee proportion and occupational accident rate by establishment size

IV. DISCUSSION

The results of this study showed that the distribution of employees within the Industrial Accidents Compensation data was different from that of the Economically Active Population Survey, similar to the target population. The employee proportion in the Industrial Accidents Compensation Data was higher in the construction industry as well as in the Seoul region and in establishments with over 300 employees compared with that of the Economically Active Population Survey data. When it was adjusted using RRM, the error rate was reduced compared with Industrial Accidents Compensation Data to improve sample distribution. The adjusted accident rate also increased or decreased, focusing on the official accident rate of 0.65% in accordance with the changed categorical accident rate. When the occupational accident rate was calculated in the event of low error rates based on overall size and region of the establishments, the total accident rate was higher than the current official occupational accident rate. The occupational accident rate was low when using RRM to adjust employee proportion by industrial sector. Because the adjusted employee proportion is dramatically changed in the construction industry, it is caused by the tendency of the under-sampling of the construction industry at the survey population data (establishment basic survey) this study used. As for the final estimated occupational accident rate, the error rate for the industrial sector was relatively higher than for region and size. The adjusted occupational accident rate by adjusting proportion of the industrial sector was relatively low. The adjusted occupational accident rate incurred by adjusting the employee proportion of the region and size was higher than the 2011 official accident rate of 0.65%. On this basis, it is necessary to consider the possibility of higher accident rates that more realistically reflect on the status of industrial accidents than the official accident rate of 0.65%.

Based on such results, first, when the occupational accident rate is calculated based on the current industrial accidents compensation data, we could figure out a sampling weakness besides the current official occupational accident rate extraction method and apply it to the RRM to obtain an adjusted occupational accident rate for the study of such differences. Such differences enable one to examine which stratum the establishment without registered insurance is mainly distributed to. This also requires the study of industrial accident status for individuals without registered insurance. Second, besides the calculation of the occupational accident rate based on the Industrial Accidents Compensation Data, the system of reporting in the event of an industrial accident is necessary to figure out victim size. And a new occupational accident index needs to be used by applying RRM to the population survey data and Industrial Accidents Compensation Data of the establishment survey data or the Economically Active Population Survey data. While the occupational accident rate in Korea is calculated based on the medical period, other OECD member countries calculate the accident rate based on the off-working days. Countries that use Industrial Accident Compensation Data as a data source like Korea

adopt the Labor Force Survey, a type of sample survey taken to calculate the accident rate.
If the sample distribution did not correspond to the population distribution, the study attempted to reflect population characteristics by using weights. However, the populations we used also had a limit as the survey population was not a target population. For example, even though The Census on Establishments data are the representative population data for Korea, the closed establishment at the time of the survey is excluded from the survey. Additionally, just as in the construction industry, while many establishments are created and closed repeatedly, there are many establishments that are closed at the survey even if they existed at the reference point of the survey. This study suggests that RRM can be used as a complementary method for such incomplete samples and incomplete populations. It is necessary to improve population surveys using the Bayesian method in future studies.

Acknowledgments

This study was supported by the intramural research fund of the Occupational Safety and Health Research Institute (OSHRI).

V. References

[1] Bankier, M. D., Rathwell, S. and Majkowski, M. 1992 Two step generalized least squares estimation in the 1991 Canadian Census. Proceedings of the Workshop on Uses of Auxiliary Information in Surveys. 5-7.

[2] Cervantes, I. F. and Brick, J. M. 2008 Empirical Evaluation of Raking Ratio Adjustments for Nonresponse. Denver, CO.

[3] Deming, W. E. and Stephan, F. F. 1940 On a Least Squares Adjustment of a Sampled Frequency Table When the Expected Marginal Totals are Known. The Annals of Mathematical Statistics. 11(4), 427-444.

[4] Demnati, A. and Rao, J. 2004 Linearization variance estimators for survey data. Survey Methodology. 30, 17-26.

[5] Deville J. C., Särndal, C. E. and Sautory, O. 1992 Calibration estimators in survey sampling. Journal of the American Statistical Association. 87(418), 376-382.

[6] Deville, J. C., Särndal, C. E., and Sautory, O. 1993 Generalized Raking Procedures in Survey Sampling. Journal of the American Statistical Association. 88(423), 1013.

[7] Haines, D. E. and Hill, J. M. 1998 A Method for Evaluating Alternative Raking Control Variables, Census 2000 Sampling and Estimation II. Vacca, E.A, Washington, DC 20233. 647.

[8] Hidiroglou, M. A. and Patak, Z. 2006 Raking ratio estimation: an application to the Canadian retail trade survey. JOURNAL OF OFFICIAL STATISTICS-STOCKHOLM. 22, 71.

[9] Kalton, G. 1983 Introduction to survey sampling. Sage.

[10] Kalton, G., FloresCervantes, I., Zheng, H., Little, R. J., Wu, C., Luan, Y., Lu, H., Gelman, A., de Leeuw, E. D. and Hox, J. 1998 Weighting methods. New Methods for Survey Research. 79.

[11] Lu, H. and Gelman, A. 2003 A Method for Estimating Design-based Sampling Variances for Surveys with Weighting, Poststratification, and Raking. Journal of Official Statistics. 19, 133.

[12] Oh, H. and Scheuren, F. 1987 Modified raking ratio estimation. Survey Methodology. 13, 209-219.

[13] Oh, H. L. and Scheuren, F. 1978 Multivariate raking ratio estimation in the 1973 exact match study, Proceedings of the Section on Survey Research Methods. American Statistical Association. 716-722.

[14] Tillé, Y. 1998 Estimation in surveys using conditional inclusion probabilities: Simple random sampling. International Statistical Review. 66(3), 303-322.

Table1. Comparison of the employee proportion and occupational accident rate by industrial sector

Business type	Industrial accidents compensation data(A)		The Census on Establishments (B)	Economically active population survey (C)	Raking Ratio Method (D)		Occupational accident rate		
	Proportion	Error			Proportion	Error	Unit accident rate(E)	A×E	D×E
Construction	24.20%	1.99	6.16%	8.09%	6.04%	0.25	0.68%	0.16%	0.04%
Public administration, national defense & social welfare administration	1.73%	0.69	3.52%	5.52%	4.26%	0.23	0.81%	0.01%	0.03%
Mining	0.10%	0.25	0.09%	0.08%	0.04%	0.50	7.75%	0.01%	0.00%
Education service	2.79%	0.65	7.91%	8.05%	9.69%	0.20	0.40%	0.01%	0.04%
Financing and insurance	4.40%	0.07	3.97%	4.75%	2.71%	0.43	0.08%	0.00%	0.00%
Agriculture, forestry and fishery	0.87%	0.19	0.18%	1.08%	0.21%	0.81	1.79%	0.02%	0.00%
Wholesale/retail	9.38%	0.23	14.81%	12.25%	10.67%	0.13	0.47%	0.04%	0.05%
Health and social welfare service	6.21%	0.13	6.28%	7.14%	8.50%	0.19	0.27%	0.02%	0.02%
Real estate and rental	1.94%	0.11	2.49%	1.75%	2.39%	0.37	0.57%	0.01%	0.01%
Project facility management and facility support service	4.85%	0.20	4.48%	6.07%	3.80%	0.37	0.59%	0.03%	0.02%
Accommodation and restaurant	3.53%	0.42	10.17%	6.08%	6.35%	0.04	1.39%	0.05%	0.09%
Art, sports and leisure service	0.93%	0.31	1.85%	1.34%	1.44%	0.07	0.55%	0.01%	0.01%
Transportation	4.86%	0.18	5.50%	4.12%	4.34%	0.05	0.59%	0.03%	0.03%
Electricity, gas, steam and water	0.48%	0.20	0.37%	0.40%	0.50%	0.25	0.13%	0.00%	0.00%
Professional, science and technical service	4.75%	0.05	4.48%	4.98%	5.66%	0.14	0.18%	0.01%	0.01%
Manufacturing	23.33%	0.17	19.83%	19.93%	25.56%	0.28	0.94%	0.22%	0.24%
Publication, video, broadcasting and information service	2.98%	0.22	2.66%	3.80%	3.42%	0.10	0.12%	0.00%	0.00%
Wastewater, waste treatment, recycling and environmental restoration	0.43%	0.23	0.39%	0.35%	0.55%	0.57	1.68%	0.01%	0.01%
Association, organization, repair and other personal service	2.23%	0.47	4.86%	4.23%	3.88%	0.08	0.53%	0.01%	0.02%
Total error rate	-	6.76	-	-	-	5.08	-	-	-
Total occupational accident rate	-	-	-	-	-	-	-	0.65%	0.64%

Table2. Comparison of the employee proportion and occupational accident rate by the region

Region	Industrial accidents compensation data(A)		The Census on Establishments (B)	Economically active population survey(C)	Raking Ratio Method(D)		Occupational accident rate		
	Proportion	Error			Proportion	Error	Unit accident rate(E)	A×E	D×E
Gangwon	2.10%	0.18	2.78%	2.57%	2.57%	0.00	1.15%	0.02%	0.03%
Gyeonggi	20.00%	0.21	21.67%	25.34%	22.19%	0.12	0.79%	0.16%	0.18%
Gyeongnam	6.34%	0.04	6.65%	6.09%	7.00%	0.15	0.88%	0.06%	0.06%
Gyeongbuk	5.24%	0.12	5.23%	4.66%	5.41%	0.16	0.79%	0.04%	0.04%
Gwangju	1.85%	0.36	2.76%	2.89%	2.80%	0.03	0.79%	0.01%	0.02%
Daegu	2.73%	0.44	4.47%	4.90%	4.13%	0.16	0.80%	0.02%	0.03%
Daejeon	2.52%	0.19	2.82%	3.13%	2.77%	0.12	0.60%	0.02%	0.02%
Busan	5.11%	0.23	6.81%	6.64%	6.63%	0.00	0.80%	0.04%	0.05%
Seoul	33.05%	0.50	24.86%	22.09%	23.83%	0.08	0.34%	0.11%	0.08%
Ulsan	2.28%	0.09	2.50%	2.51%	2.76%	0.10	0.83%	0.02%	0.02%
Incheon	4.54%	0.24	4.69%	5.95%	4.83%	0.19	0.79%	0.04%	0.04%
Jeonnam	3.49%	0.29	3.24%	2.71%	3.14%	0.16	0.76%	0.03%	0.02%
Jeonbuk	2.86%	0.03	3.21%	2.94%	3.21%	0.09	0.93%	0.03%	0.03%
Jeju	0.80%	0.24	1.13%	1.05%	1.09%	0.04	0.76%	0.01%	0.01%
Chungnam	4.40%	0.20	4.14%	3.66%	4.41%	0.20	0.69%	0.03%	0.03%
Chungbuk	2.71%	0.06	3.05%	2.89%	3.23%	0.12	0.83%	0.02%	0.03%
Total error rate	-	3.43	-	-	-	1.72	-	-	-
Total occupational accident rate	-	-	-	-	-	-	-	0.65%	0.70%

Table3. Comparison of the employee proportion and occupational accident rate by the size of establishments

Size of establishment	Industrial accidents compensation data(A)		The Census on Establishments (B)	Economically active population survey(C)	Raking Ratio Method (D)		Occupational accident rate		
	Proportion	Error			Proportion	Error	Unit accident rate(E)	A×E	D×E
1-4 employees	14.30%	0.25	28.81%	19.13%	17.77%	0.07	1.55%	0.22%	0.28%
5-9 employees	11.23%	0.35	11.43%	17.18%	13.49%	0.21	1.00%	0.11%	0.14%
10-99 employees	39.62%	0.06	34.47%	42.33%	42.54%	0.00	0.61%	0.24%	0.26%
100-299 employees	14.00%	0.41	10.92%	9.95%	12.71%	0.28	0.26%	0.04%	0.03%
Over 300 employees	20.85%	0.83	14.38%	11.40%	13.50%	0.18	0.17%	0.04%	0.02%
Total error rate	-	1.90	-	-	-	0.75	-	-	-
Accident rate	-	-	-	-	-	-	-	0.65%	0.73%