

Sample models in monitoring survey UniDOS.

Mykola Sydorov¹ and Oleksiy Sereda²

¹Taras Shevchenko National University of Kyiv, e-mail: myksyd@knu.ua

²Taras Shevchenko National University of Kyiv, e-mail: as_sereda@knu.ua

Abstract

The paper presents models of sampling in the 9th wave of the monitoring survey UniDOS 2013. The problems in sample construction comprise inaccurate and incomplete information about the population, hard access to some groups of students and small size of strata.

Keywords: Multilevel sampling, strata

1 Introduction

Since 2009 the Faculty of Sociology of Taras Shevchenko National University of Kyiv conducts a monitoring survey of students. The survey examines issues of motivation for obtaining higher education, plans for the future after graduation, expectations of students regarding the labour market and further employment, the attitudes to the educational process at the university faculties, etc. Every year we have the problem with constructing an optimal sample model. In this paper, we present sample models of poll in the spring of 2013.

2 General Population

The general population consists of full-time students of 17 faculties and institutes of the Taras Shevchenko National University of Kyiv, which is 17484 students. Part-time students was not included in the population, because, education is not their main activity and the time of stay at the university is considerably limited compared with the students of full-time education, as well as their limited stay and communication in the community. At the beginning of the survey we did not have all relevant data on the composition of groups and courses in the faculties of KNU, as based upon estimates for September 2013. The structure of the population is given in Appendix 1. The biggest problem was the quality of information for the first year students. Also the first year can be considered a special category for which the questionnaire is different from other years of study questionnaire because of first year students can't be asked about the quality of education and other aspects of the university because they studied just for several weeks.

Thus, the sample was calculated only for 2+ year students, of which 13658, excluding

those absent because of field practice. Based on the purpose and objectives of the research, the research team was interested in the opinion of the representatives of all faculties and institutes, therefore, at this stage, a continuous selection is applied.

3 Sampling

To achieve the goals and objectives, as well as to implement the methodological plan of the study, we used multilevel sample selection scheme.

1st level - continuous selection of faculties, volume - in proportion to the number of students at the faculty: faculty - stratum.

Level 2 - stratification at the year of study (bachelor 2+, masters – all) - in proportion to the number of students available for each year: every year of study - stratum

Level 3 - nested method of selection, "nest" corresponds to the selected groups (group) of a year of study at the faculty.

Level 4 - random selection of respondents in each selected group (using two-colours cards).

The choice of such a model is due to the inability to obtain lists of students from all faculties and institutes, which would allow the use of the random selection model. At the same time, students are distributed at auditorium time by groups, in which the number and composition of students at each faculty are approximately the same. Thus, we have only 2 variables to describe the population: the approximate number of students in the faculties and the approximate number of students in the groups. We do not even have the distribution of students by gender.

The next step was to calculate the sample size. Based on the formula for simple random selection, for the general population 13658 we should have a sample of 374 respondents for a sampling error of 5% with a confidence level of 0.95. But, if we make conclusions about each faculty separately, then the sample size should be calculated for every faculty separately and we will receive a slightly larger number (Appendix 2). The total number of all respondents in all faculties in this case is 4105. This significantly exceeds the client's ability, which is about 1200 questionnaires.

Next, we propose two approaches to the formation of a sample population.

3.1. Sample approach #1

Since during the construction of a sample of 1,200 respondents in some departments there were not enough respondents in sample, for the adequate representation within the faculty, we decided to secure a minimum number of 50 respondents for each faculty / institute.

The sample was divided into 2 parts: proportional and additional. The proportional number was 1085 respondents and the additional - another 112 for those faculties,

where the number of respondents did not reach 50 (Appendix 3.). However, after the calculation of the size proportional to the courses selection at the faculties, it turned out that the size of additional selection should be 117 (Appendix 4). This is due to the rounding of the calculated numbers of students in the groups.

Thus, we obtained a sample that allows us to conduct a representative survey of students at the KNU. The sample size was 1202 respondents. Sample weights are given in Appendix 5.

3.2. Sample approach #2

Since the minimum number of respondents for each faculty must be at least 50 people, from the maximum sample size we can select $50 * 17 = 850$ respondents. The residual of $1200 - 850 = 350$ is distributed among the faculties, in proportion to the difference in between the number of students with the smallest faculty. The smallest number of students study at the Faculty of Sociology – 138 students available for surveying, therefore, to construct the proportions, we will subtract the number of students of each faculty from the number of faculty of sociology. Afterwards we calculate proportions. The general population, the proportion of the faculties and the estimated sample sizes by faculty are given in Appendix 6.

The next step was to calculate the sample size for each faculty and for each year of study. Due to rounding, the proportion has slightly changed and is shown in appendix 7. It also shows weight ratios.

To calculate this sample option, we used R (R Core Team, 2018) package and the surveyplanning package (Breidaks, Liberts, & Jukams, 2017).

It should be noted that during the field stage the first approach to the sample construction was used.

3.3. Final stage of sampling

The next step was random selection of groups, which was carried out for each year of the study of each faculty separately. We did not have information about the principle of grouping students into groups: in some departments, the division was based on an alphabet; in some - in separate groups there were students with higher grades, in other - students living in dormitories in one group, and local residents of Kyiv - up to other, etc. Therefore, the only approach to groups selection was random.

In order to select respondents in the group, the method of labelled cards was used as the only suitable method of randomization. The pre-interviewer receives a set of white cards and a set of white cards marked with a red square. The number of white cards was equal to the number of students presented in the group minus the number of respondents who should be interviewed in this group. The number of white cards marked with a red square is equal to the number of students to be interviewed in this group. The pile of cards with cards of both types (white and coloured) was well mixed and handed out to

the students.

At the stage of elaboration of the sampling design, several ways of respondent selection at the last stage (in groups) were proposed. Among the main ones were the following methods of selection:

- random selection of respondents by student lists. But the lists of students enrolled in the university are confidential information, so there are significant problems with the access;
- step-by-step selection of respondents directly in the classrooms. For this selection procedure, information is needed on the number of students in the group and the number of students to be interviewed. Based on this information, the interviewer calculates the step and makes selection. This selection procedure requires an interviewer to make mathematical calculations and takes a long time. This increases the possibility that the interviewer will make a mistake in the calculations, or in the process of calculating the step. There is a high probability that the interviewer will not be able to calculate the step at all and distribute the questionnaire to all interested persons.

The selection procedure with the use of cards was chosen because: the use of cards does not require serious mental calculations, as well as the availability of lists of students studying in this group.

- Working with multi-coloured cards and the selection procedure reminds a certain lottery game. Increased interest in the method of selection increases the motivation of the interviewer to follow all the requirements of this procedure.
- Red cards received by respondents are an important element in controlling the work of interviewers.

4. Conclusions

Both of the above procedures for sampling the population allow us to conduct a representative survey of students at the faculties of the KNU. The peculiarities of the design of the sample are related to the tasks set for the research team and the strict constraints on the customer's resources. In 2013, the 1st approach was implemented, although the 2nd one we consider simpler and more convenient. In addition, it has a smaller range of weights. The second method we plan to apply in the study, which will take place in September 2018.

References

- R Core Team (2018). *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria. URL <https://www.R-project.org/>
- Breidaks J, Liberts M, Jukams J (2017). *_surveyplanning: Survey planning tools_*. R package version 2.9, <URL:<https://csblatvia.github.io/surveyplanning/>>

Appendices

Appendix 1.

Table # 1. Structure of the general population

Faculty / Institute	Number of students	B1 (Bach. degree stud. 1 year)	B2	B3	B4	S1 (specialist degree)	M1 (MA stud. 1 year)	M2
Educational and Scientific Centre "Institute of Biology"	933	196	172	129	151	21	133	131
Geography Faculty	1051	207	178	143	195	53	140	135
Geological Faculty	401	75	90	46	73	0	56	61
Faculty of Economics	1671	364	322	241	278	0	244	222
Historical Faculty	754	161	140	97	138	64	78	76
Faculty of Cybernetics	913	210	175	139	148	55	107	79
Mechanical and Mathematical Faculty	673	126	124	110	117	30	88	78
Faculty of Psychology	669	161	130	108	115	19	74	62
Radiophysics Faculty	750	142	157	163	126	18	78	66
Faculty of Sociology	347	84	72	44	63	18	33	33
Faculty of Physics	713	143	116	103	106	19	109	117
Philosophy Faculty	770	181	174	98	157	25	73	62
Chemical Faculty	481	91	88	73	83	20	66	60
Faculty of Law	2153	415	413	273	409	87	282	274
Institute of Journalism	1150	235	261	170	226	15	120	123
Institute of International Relations	1762	175	336	294	348	110	254	245
Institute of Philology	2293	612	510	476	25	43	305	322
Total	17484	3578	3458	2707	2758	597	2240	2146

In the table # 1, years in which the number of students indicated for 0 is those for which

there was no set ("Specialist" of the Geological and Economic Faculties). There were also years in which the students were in practice for the period of the survey, so they were difficult to access (they are highlighted in grey). For further calculations, their number was also counted as 0, because the ability to involve these students in participating in the survey is absent.

Appendix 2.

Table # 2. Calculation of the sample size for each faculty separately

Faculty / Institute	N of general population	n of sample
Educational and Scientific Centre "Institute of Biology"	737	253
Geography Faculty	844	264
Geological Faculty	326	176
Faculty of Economics	1307	297
Historical Faculty	593	233
Faculty of Cybernetics	703	248
Mechanical and Mathematical Faculty	547	226
Faculty of Psychology	508	219
Radiophysics Faculty	608	235
Faculty of Sociology	138	102
Faculty of Physics	570	229
Philosophy Faculty	589	233
Chemical Faculty	390	194
Faculty of Law	1738	315
Institute of Journalism	792	259
Institute of International Relations	1587	309
Institute of Philology	1681	313
Total	13658	4105

Appendix 3.

Table # 3. Calculation primary and additional selection in sample

Faculty / Institute	Total number for 2+	B2 (Bach. degree stud. 1 year)	B3	B4	S1 (specialist degree)	M1 (MA stud. 1 year)	M2	Number of additional selection
Educational and Scientific Centre "Institute of Biology"	59	14	10	12	2	11	10	0
Geography Faculty	67	14	11	15	4	11	11	0

Faculty / Institute	Total number for 2+	B2 (Bach. degree stud. 1 year)	B3	B4	S1 (specialist degree)	M1 (MA stud. 1 year)	M2	Number of additional selection
Geological Faculty	26	7	4	6	0	4	5	24
Faculty of Economics	104	26	19	22	0	19	18	0
Historical Faculty	47	11	8	11	5	6	6	3
Faculty of Cybernetics	56	14	11	12	4	9	6	0
Mechanical and Mathematical Faculty	43	10	9	9	2	7	6	7
Faculty of Psychology	40	10	9	9	2	6	5	10
Radiophysics Faculty	48	12	13	10	1	6	5	2
Faculty of Sociology	11	6	0	0	0	3	3	39
Faculty of Physics	45	9	8	8	2	9	9	5
Philosophy Faculty	47	14	8	12	2	6	5	3
Chemical Faculty	31	7	6	7	2	5	5	19
Faculty of Law	138	33	22	32	7	22	22	0
Institute of Journalism	63	21	14	18	1	10	0	0
Institute of International Relations	126	27	23	28	9	20	19	0
Institute of Philology	134	41	38	2	3	24	26	0
Total	1085							112

Appendix 4.

Table # 4. Final calculation of the sample in approach 1.

Faculty / Institute	Total number for 2+	B2 (Bach. degree stud. 1 year)	B3	B4	S1 (specialist degree)	M1 (MA stud. 1 year)	M2
Educational and Scientific Centre "Institute of Biology"	59	14	10	12	2	11	10

Faculty / Institute	Total number for 2+	B2 (Bach. degree stud. 1 year)	B3	B4	S1 (specialist degree)	M1 (MA stud. 1 year)	M2
Geography Faculty	66	14	11	15	4	11	11
Geological Faculty	50	14	7	11	0	9	9
Faculty of Economics	104	26	19	22	0	19	18
Historical Faculty	50	12	8	12	5	7	6
Faculty of Cybernetics	56	14	11	12	4	9	6
Mechanical and Mathematical Faculty	50	11	10	11	3	8	7
Faculty of Psychology	50	13	11	11	2	7	6
Radiophysics Faculty	52	13	13	10	5	6	5
Faculty of Sociology	50	26	0	0	0	12	12
Faculty of Physics	50	10	9	9	2	10	10
Philosophy Faculty	49	15	8	13	2	6	5
Chemical Faculty	50	11	9	11	3	8	8
Faculty of Law	138	33	22	32	7	22	22
Institute of Journalism	68	21	14	18	5	10	0
Institute of International Relations	126	27	23	28	9	20	19
Institute of Philology	134	41	38	2	3	24	26
Total	1202	315	223	229	56	199	180

Appendix 5.

Table # 5. Weights coefficients for the sample in approach 1.

Faculty / Institute	For the Faculty / Institute	B2	B3	B4	S1	M1	M2
Educational and Scientific Centre "Institute of Biology"	1.0993	1.0812	1.1353	1.1074	0.9241	1.0641	1.0812
Geography Faculty	1.1254	1.1189	1.1441	1.1441	1.1661	1.1201	1.1189
Geological Faculty	0.5738	0.5658	0.5783	0.584	1	0.5476	0.5658
Faculty of Economics	1.106	1.0899	1.1163	1.1121	1	1.1302	1.0899
Historical Faculty	1.0438	1.0267	1.0671	1.0121	1.1265	0.9806	1.0267
Faculty of Cybernetics	1.1048	1.1001	1.1121	1.0854	1.2101	1.0463	1.1001
Mechanical and Mathematical Faculty	0.9628	0.9921	0.9681	0.9361	0.8801	0.9681	0.9921

Faculty / Institute	For the Faculty / Institute	B2	B3	B4	S1	M1	M2
Faculty of Psychology	0.8942	0.8801	0.8641	0.9201	0.8361	0.9304	0.8801
Radiophysics Faculty	1.029	1.0629	1.1035	1.1089	0.3168	1.1441	1.0629
Faculty of Sociology	0.2429	0.2437	1	1	1	0.242	0.2437
Faculty of Physics	1.0033	1.0209	1.0072	1.0365	0.8361	0.9593	1.0209
Philosophy Faculty	1.0579	1.0209	1.0781	1.0629	1.1001	1.0708	1.0209
Chemical Faculty	0.6865	0.7041	0.7138	0.6641	0.5867	0.7261	0.7041
Faculty of Law	1.1084	1.1014	1.0921	1.1248	1.0938	1.1281	1.1014
Institute of Journalism	1.025	1.0938	1.0687	1.105	0.264	1.0561	1.0938
Institute of International Relations	1.1085	1.0952	1.125	1.0938	1.0756	1.1177	1.0952
Institute of Philology	1.104	1.0947	1.1024	1.1001	1.2614	1.1184	1.0947
MIN	0.2429	0.2437	0.5783	0.584	0.264	0.242	0.2437
MAX	1.1254	1.1189	1.1441	1.1441	1.2101	1.1441	1.1189

Appendix 6.

Table # 6. Calculation of the sample size for each faculty separately for approach 2

Faculty / Institute	N of general population	General – number of stud. sociology faculty	n of sample
Educational and Scientific Centre "Institute of Biology"	737	599	69
Geography Faculty	844	706	72
Geological Faculty	326	188	56
Faculty of Economics	1307	1169	86
Historical Faculty	593	455	64
Faculty of Cybernetics	703	565	67
Mechanical and Mathematical Faculty	547	409	63
Faculty of Psychology	508	370	61
Radiophysics Faculty	608	470	65
Faculty of Sociology	138	0	50
Faculty of Physics	570	432	63
Philosophy Faculty	589	451	64
Chemical Faculty	390	252	58
Faculty of Law	1738	1600	100
Institute of Journalism	792	654	70
Institute of International Relations	1587	1449	95
Institute of Philology	1681	1543	98
Total	13658	11312	1201

Appendix 7.

Table # 7. Sample in approach 2 by faculty and years.

Faculty / Institute	For the Faculty / Institute	B2	B3	B4	S1	M1	M2
Educational and Scientific Centre "Institute of Biology"	68	16	12	14	2	12	12
Geography Faculty	73	15	12	17	5	12	12
Geological Faculty	56	15	8	13	0	10	10
Faculty of Economics	86	21	16	18	0	16	15
Historical Faculty	63	15	10	15	7	8	8
Faculty of Cybernetics	67	17	13	14	5	10	8
Mechanical and Mathematical Faculty	62	14	13	13	3	10	9
Faculty of Psychology	61	16	13	14	2	9	7
Radiophysics Faculty	64	17	17	13	2	8	7
Faculty of Sociology	50	26	0	0	0	12	12
Faculty of Physics	63	13	11	12	2	12	13
Philosophy Faculty	65	19	11	17	3	8	7
Chemical Faculty	58	13	11	12	3	10	9
Faculty of Law	101	24	16	24	5	16	16
Institute of Journalism	70	23	15	20	1	11	0
Institute of International Relations	96	20	18	21	7	15	15
Institute of Philology	99	30	28	1	3	18	19
Total	1202	314	224	238	50	197	179

Min and max weights of coefficients for the sample in approach 2.

	Weights for Faculties	Weights for Years of study
min	0.8108	0.2216
max	1.4696	1.6618