So, H. J. et al. (Eds.) (2020). Proceedings of the 28th International Conference on Computers in Education. Australia: Asia-Pacific Society for Computers in Education

An educational system with functions of guidance and adaptive advice to support problem solving based on basic concepts of statistics

Kanako SUZUKI [a*], Tatsuhiro KONISHI [b]
[a]*Graduate School of Integrated Science and Technology, Shizuoka University, Japan*
[b]*Faculty of Informatics, Shizuoka University, Japan*
*suzuki.kanako.16@shizuoka.ac.jp

Abstract: Recently, statistics is a necessary subject not only for experts but for common people as well. In the field of education on statistics, it is required to present that improve learners' motivation and promote deep understanding. We thought that problems using statistical concepts on the real world can satisfy such a condition. In this study, we have developed a educational system equipped with a guidance function that asks learners such questions and organizes data, and an adaptive advice function that gives advice according to the type of the wrong answer. Learners tackle the problem using our system and system-independent statistical tools. In addition, we conducted an evaluation experiment to evaluate the usefulness of the system. The results showed that the guidance function and the advisory function of this system are effective for improving the exercise achievement rate for the learner to proceed with learning, and that the learner feels that the acquired knowledge is useful.

Keywords: Statistics, Guidance, Adaptive advice, Learning support system

1. Introduction

In modern society, statistics is a subject required by many other than the experts. However, in the field of education, mainly rote learning is employed. There have been various discussions on ways for learners to improve their statistical skills. It is argued that it is important to learn not only statistical calculation process but also total problem solving process using statistical skills (Wild & Pfannkuch, 1999). Wild & Pfannkuchh propose the PPDAC cycle as the total process of statistical problem solving. In the PPDAC cycle, the problem solving process consists of the following 5 steps; Problem, Plan, Data, Analysis, and Conclusion. Smith also points out the importance of total problem solving process (Smith, 1998). He recommends to incorporate active learning strategies in order to develop students' statistical inference ability. The active learning strategy is to let learners carry out activities such as study design, data collection, and result analysis. From the above, we design exercises that learners solve more than just the statistical calculation process. Here, since exercises of the study design or the data collection must be relatively long-term ones, we design our system to handle exercises of the result analysis process that can be finished within one lesson.

It is difficult to say that the teachers give encouraging lessons to the learners to develop motivation and a deep understanding of statistics. In this study, we define exercises which can encourage learners' motivation and deepen understanding ore ones of finding properties of concepts in real world using statistics and data of the real world (e.g., Let's find the most important property of the team in order to increase winning rate by analyzing the stats data of the 12 Japan professional baseball

teams in 2015!). We developed a learning support system which acts some of the teacher's roles in such exercises. We suppose that there are learners who don't know how to solve exercises and find an erroneous conclusion. These learners need individual guidance and adaptive advice. However, teachers generally do not have enough resources to pay attention to individual needs because there are too many learners per teacher. There is also a limitation on school hours. In this study, we define a teacher's role as conducting guidance and giving the adaptive advice to complete the exercise. We developed a learning support system which has the following two functions to help learners:
(1) Guidance of learner's behavior toward completion of exercises.
(2) Evaluation and adaptive advice of answers obtained from learners.
We evaluated our system experimentally and found that it works effectively.

There are many existing learning support systems for statistics. For example, the Web-based statistical learning support system (Baharun & Porter, 2012) aims to support mathematical problem solving rather than real world problems that our system supports. A learning support system by an expert system (Barbara & William, 1996) deals with real-world problems. However, the system does not have the functions for resolving learners' impasse (the function of guiding the problem-solving process or the function of giving advice according to the error of the learner) that our system equips.

2. Methods

2.1 Discussion on Exercises Employed in the Learning Support System

Exercises in this learning support system are the exercises to find properties of a dataset using concepts of statistics. Learners complete the exercises using a statistical analysis tool (We adapt BellCurve for Excel in this study. It is statistical analysis software that adds statistical analysis functions to the Excel menu, such as calculating correlation coefficient, drawing box plot or scatter plot and so on).

In this study, as mentioned in Section 1, our system supports exercises to find properties of real world by using statistics in order to increase learners' motivation and to deepen understanding. We define "deep understanding" as the understanding of how to use a statistical concept in the real world. We think that it is not deep understanding to know a statistical concept as a simple numerical formula. It tends to be hard for learners to understand how to use the statistical concept in real life because learners recognize only the mathematical aspect of the statistical concept as performed in in-class exercises. Therefore, our learning support system consists of situations from daily experience that learners can solve using the statistical concept. Our educational support system gives learners guidance after asking to execute such examples. It focuses on the method of statistical concepts using which learners solve the exercises. Our learning support system sets exercises that learners can use in real life. Learners solve these exercises with a statistical concept called the correlation coefficient.

2.2 Methods Guiding Learners for Solving Exercises

It is our idea, learners have abilities to solve exercises by themselves through organizing data and deciding which statistical concept to use. However, it is difficult to say that all learners can solve exercises by themselves. Therefore, our learning support system guides learners who cannot solve exercises by themselves. It shows what to do next in a given exercise to achieve a predefined goal. First, we defined the common learning processes independent of learning content.
(Step 1) Our learning support system gives a rough guide to learners for solving exercises.
(Step 2) Learners visualize the given data and information using the statistical analysis tool.
(Step 3) Learners solve sub-exercises to organize data from our learning support system.
The sub-exercise means smaller exercise that is necessary for solving the exercise.
(Step 4) Learners answer given exercises based on organized information in Step 3.

At first, our learning support system gives a rough guide to learners for solving exercises (Step 1). An example of an exercise for learning about correlation coefficient is shown below.

Let us compare the properties of the team with a high winning rate from the team results of the 12 professional baseball teams of Japan in 2015 and compare it with that of 1950!
Learners should organize given data and visualize them using BellCurve for Excel first. Our system gives learners instructions on how to use BellCurve for Excel because in order to get the output correlation matrixes and heat maps of correlation matrixes (Step 2).

In Step 3, our learning support system guides the learners by setting sub-exercises (Figure 1). The sub-exercises let leaners focus on a parameter or correlation between parameters that is useful for solving the exercise. Then it gives learners adaptive advice according to mistakes that they make. Our system repeats setting sub-exercise in order to guide learners to achieve the goal of the exercise (Figure 1). Finally, it set a kind of another sub-exercise that has the same goal as the final goal of the exercise. We call such sub-exercise "final sub-exercise". Figure 1 shows example of dialogue setting sub exercise (in Step 3). The messages in Figure 1 are translated from Japanese into English for this paper.

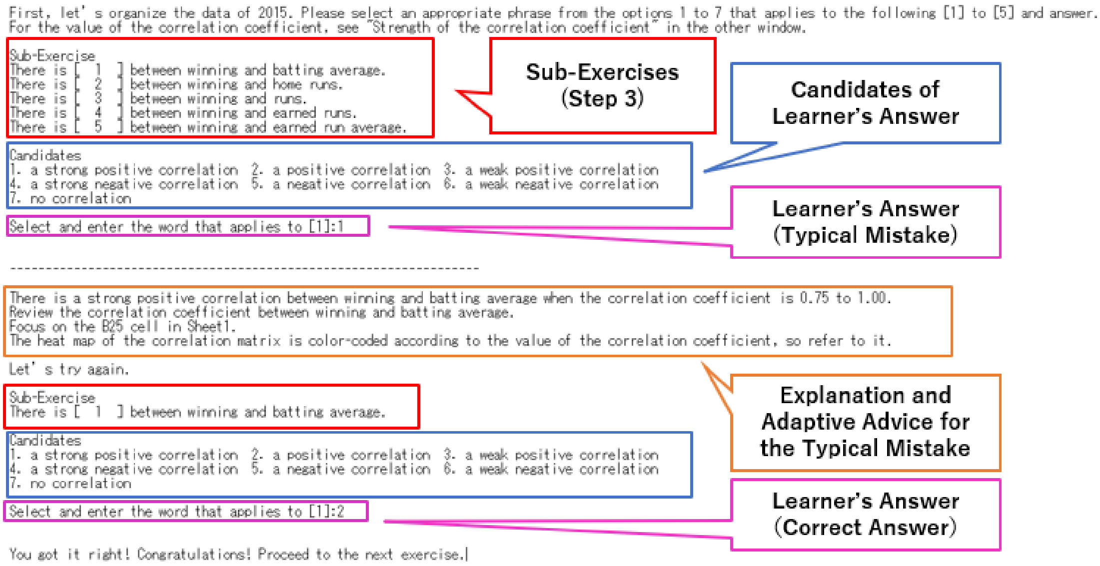

Figure 1. Example of dialogue

2.3 Generation Method of Adaptive Advice According to Mistakes

Our learning support system checks the learners' answers and gives learners adaptive advice according to their mistakes. Our system gives advices not only to suggest their answer is wrong simply, but also to show how they should correct their wrong answer. It prevents them from finding answer by trial-and-error. Variations of the correlation coefficient advice implemented in our system are shown below.

- **<u>If</u>** learners focus on the wrong cell, **<u>then</u>** point out that it's wrong and that they should focus on another cell.
- **<u>If</u>** learners answer suggests that they misunderstand how to interpret positive/negative correlation coefficient, **<u>then</u>** let learners confirm meaning of positive/negative correlation coefficient by reading the textbook of statistics again.
- **<u>If</u>** learners answer suggests that they misunderstand how to interpret of strength correlation coefficient, **<u>then</u>** let learners confirm meaning of strength of correlation coefficient by reading the textbook of statistics again.

And variations of the box plot advice implemented in our system are shown below.

- **<u>If</u>** learners focus on the wrong box plot, **<u>then</u>** point out that it's wrong and that they should focus on another box plot.
- **<u>If</u>** learners answered a number in wrong cell instead of another cell that has correct answer, **<u>then</u>** point out that you are focusing on wrong cell and read the textbook of statistics again if you miss some concepts of statistics.

First, we explain how to generate a method of adaptive advice according to mistake. Our learning support system has candidate for answers to all sub-exercises Learners select one of the candidates to answer the sub-exercises. Our system has functions to diagnose learners' answers. The learners are allowed to go to the next sub-exercises when his/her answer is correct. In case that the answer is wrong, our educational system gives adaptive advice according to his/her mistake. We divide all choices into correct answers and several types of typical mistakes. We attach adaptive advice to each typical mistake. The advice is designed to help learners fix their wrong answers. The above process is as shown in Figure 2. The messages in Figure 2 are also translated from Japanese into English for this paper. We attach adaptive advice to each typical mistake.

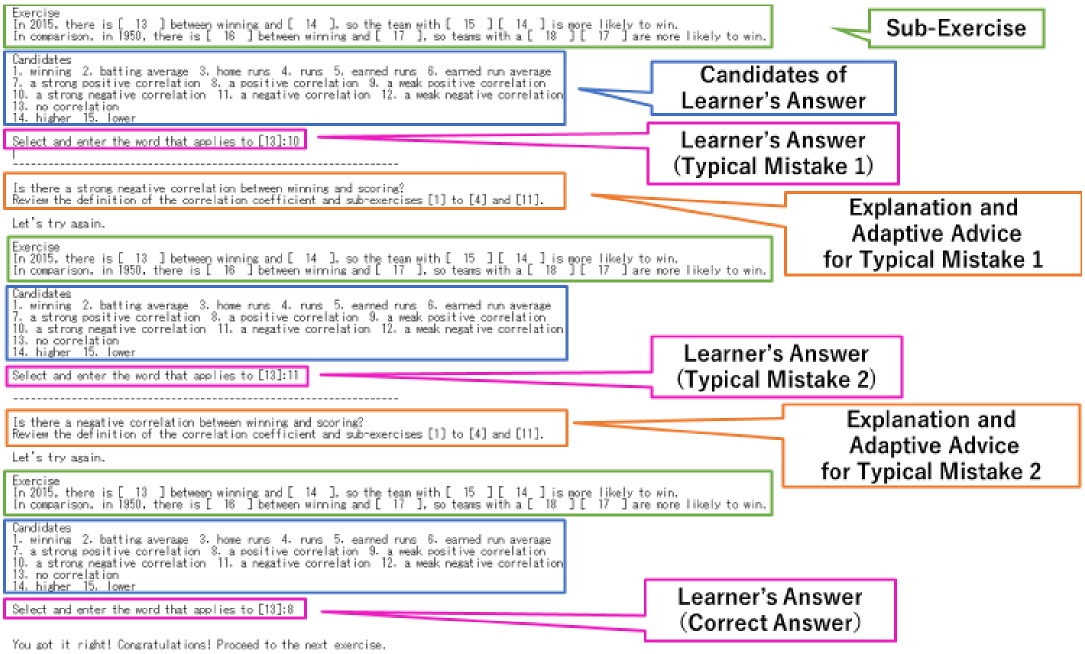

Figure 2. Example of advice according to mistake

3. Design of our System

Our system is based on the assumption that the learner's personal computer has a statistical analysis tool independent of the system. We independently developed our system and statistical analysis tool because by incorporating a statistical analysis tool in the system, the system itself has the advantage of being able to monitor the behavior of the learner, but we thought that it would be very costly to change the statistical analysis tool. For the learners to proceed smoothly with the exercise, we created slides explaining how to use and explaining the statistical terms used in the exercise. Learners refer to these slides in addition to our system and to proceed with learning.

We mainly devised the message display when implementing our system. The points that we have devised specifically are as follows.

- Admixed words of praise according to the behavior of the learners
- Directives and commentary output in a colloquial style

4. Experimental Evaluation of the Experiment

4.1 *Hypothesis*

In order to evaluate the effectiveness and usefulness of the system, we conducted experimental evaluation on the following two hypotheses:

Hypothesis 1: Guidance and advisory functions are useful for learners to proceed with learning.

Hypothesis 1-1: The guidance function is effective.

Hypothesis 1-2: The advisory function is effective.
Hypothesis 2: Subjective evaluation by the subjects suggests our system is useful
Hypothesis 2-1: The guidance function of this system is useful for learners. The subjects feel the guidance function is useful.
Hypothesis 2-2: The subjects feel the advisory function is useful. Hypothesis
Hypothesis 2-3: The subjects feel our system totally useful.

4.2 Outline of the Experiment

First, we describe the design of the experiment. The subjects are ten persons (six master's students, one university student, two university graduates and one vocational school graduate). To verify hypotheses 1-1 and 1-2, we compare the incidence of impasses with and without guidance and advisory functions. Specifically, the subject is made to solve the problem under the following three conditions, and the accuracy rate of the final sub-exercises is compared (final sub-exercise is defined in 3.2).
(Condition 1) Solve exercises using a system without guidance and advisory functions.
(Condition 2) Solve exercises using a system equipped with only guidance function.
(Condition 3) Solve exercises using a system equipped with both guidance and advisory functions. In order to verify hypotheses 2-1, 2-2 and 2-3 a questionnaire was given to the subjects after the experiment (Condition 3). The questions in the questionnaire are as fellows.
(Q1) Was the guidance function useful?
(Q2) Was the advisory function useful?
(Q3) Was the system equipped with both guidance function and advisory function useful?

The subjects select one from 5 (useful), 4 (relatively useful), 3 (fair), 2 (relatively useless) and 1 (useless). There are six exercises to be solved by the subject, which consist of three exercises solved using the correlation coefficient and three exercises solved using the box plot. The difficulty level of each exercise is different, with number 1 being the easiest and number 3 being the most difficult. We let subjects solve each exercise under the above conditions (1)-(3) and compare the correct answer rate of the final sub exercises. The time limit for the subjects to solve an exercise is seven minutes. Subjects are allowed to give up solving exercises.

4.3 Result on Hypothesis 1

Table 1 shows the average of the correct answer rate of each exercise. Some final sub-exercises ask multiple items. “equipped” means the condition that the system equipped the function written in the topline of Table 1.

Table 1. *Problem difficulty level / correct answer rate by system*

Themes	Exercise Level	Correct answer rate of the final sub exercise					
		Guidance function	Advisory function	Guidance function	Advisory function	Guidance function	Advisory function
				equipped		equipped	equipped
Correlation Coefficient	Level1	85.00%		98.33%		98.33%	
	Level2	30.00%		80.00%		100%	
	Level3	95.00%		100%		100%	
Box Plot	Level1	83.33%		98.33%		100%	
	Level2	86.66%		100%		100%	
	Level3	100%		100%		100%	

In all of the exercises, the correct answer rate of the “condition with guidance/without advice” is higher than (or equal to) that of the “condition without guidance/without advice”. This result suggests that the guidance function was effective in resolving the impasses.

In all of the exercises, the correct answer rate of the "condition with guidance/without advice" is higher than (or equal to) that of the "condition without guidance/without advice". This result suggests that the advisory function was effective in resolving impasses. As a result, Hypothesis 1 was supported.

4.4 Result on Hypothesis 2

Table 2 shows the distribution of responses and average values for the three questions in the questionnaire. The results of (Q1) and (Q2) indicate that all subjects have positive evaluation of both the guidance function and the advisory function. We interviewed subjects who answered "fair" to (Q3). They had high grade of problem-solving ability, so they correctly answered the final exercises without any advice. Therefore, they felt the system is not so effective. As a reason for subjects' evaluation, they wrote the following messages (We translate the original messages in Japanese into English). "It was easier to carry out the exercises because I could find what I should do next and how my answer was wrong.", "I think (Condition 3) is upwardly compatible with (Condition 2), but I felt the detailed messages from (Condition 3) a little bothersome when I could solve exercises without advice."

Table 2. *The distribution of responses and average values for the three questions in the questionnaire*

Selection / Question Number	useless (=1)	relatively useless (=2)	fair (=3)	relatively useful (=4)	useful (=5)	Average Point
(Q1)	0	0	0	1	9	4.9
(Q2)	0	0	0	1	9	4.9
(Q3)	0	0	2	2	6	4.4

This result suggests that most of the subjects feel the guidance and advisory functions of our system useful. Furthermore, it is suggested that the subjects feel that the system having both functions is useful for learning. As a result, Hypothesis 2 was supported.

5. Conclusion

In this study, we set out the exercise of finding out the property from the data by using the statistical concepts, we have developed a learning support system to act as a teacher. We developed two functions. One is a guidance function that promotes data organization. The other is an advisory function that gives adaptive advice according to learners' typical mistakes. In addition, we evaluated the usability and educational effectiveness of our system experimentally. The results show that guidance and advisory functions increase the exercises achievement rate. Furthermore, the results of questionnaires on the guidance function and the advisory function show usefulness of our system. In the future, we will tackle the issues pointed out by the subjects. For example, introducing GUI to our system.

References

Baharun, N. & Porter, A.L. (2012). A web-based learning support to import student' learning of statistics. *ISM International Statistical Conference,* 359-367

Grabowski, B.L. & Harkness, W.L. (1996). Enhancing Statistics Education with Expert Systems: More than an Advisory System. *Journal of Statistics Education[Online], 4*(3).

Smith, G. (1998). Learning statistics by doing statistics. *Journal of Statistics Education[Online]*, *6*(3).

Wild, C.J. & Pfannkuch, M. (1999). Statistical Thinking in Empirical Enquiry. *International Statistical Review, 67*(3), 223-265