STRATEGIES TO MAKE COUNTER-EXAMPLES WHILE COMPARING TWO GROUPS

Kai-Lin Yang and Wen-Han Chang
Department of Mathematics, National Taiwan Normal University, Taiwan
anniecs@hotmail.com

In this paper, we aim at analyzing how pre-service and in-service mathematics teachers make a judgment about comparing two groups which requires several important concepts. The research questions include: (1) Which concepts are applied to make a judgment about comparing two groups? (2) How are the counter-examples made? In our study, teachers were asked to make a judgment about comparing two groups individually and then explained their judgments in class. The result shows that most teachers prefer using concept of mean to represent the outline of the whole distribution. And the strategies of successfully making counter-examples include thinking with (1) proportion-based, (2) properties-oriented, (3) statistics-modified, and (4) unfixing a sub-group. Implications and further research will be suggested in relation to roles of making counter-examples for learning statistical thinking.

INTRODUCTION

Makar and Confrey (2002) think comparing two groups becomes a powerful tool to draw conclusions and towards a consideration of statistical reasoning. There are many examples of tasks that ask students to compare two or more groups of data in the literature. Watson and Moritz (1999) used the tasks which help students in grades 3 through to 9 make inferences about group different by comparing two data sets. Besides the statistics properly used to compare two groups, we set up another goal of using probabilistic models to compare two groups. In our study, we design research questions that include: (1) Which concepts are applied to make a judgment about comparing two groups? (2) How are the counter-examples made?

METHOD

There are two tasks which are provided for pre-service teachers and in-service teachers respectively in this study. Both of the two tasks include three activities. The first activity is comparing two groups, the second activity is making a counter-example to disprove some wrong statements, and the last activity is explaining this counter-example. However the distributions which should be compared and the wrong statement which should be disproved by counter-examples are not the same in the two tasks.

The distributions and the questions of the first task are shown in Figure 1. There are 13 students in each of the two groups. Regarding the distributions of the two classes, all of the three statistics (mean, mode, median) in class A are larger than those in class B. In addition, the probability that one student chosen randomly from class A gets more scores than one from class B is larger than the probability that one student chosen randomly from class B gets more scores than one from class A (e.g., $P(A_i>B_j) > P(B_j>A_i)$ for some i,j=1,2,…,22, where A_i and B_j represent the score of the students who were chosen randomly from class A and class B respectively).

After reading the distributions of the two classes, 22 in-service teachers were asked to make a judgment about comparing the two classes. The question is "*If we randomly choose one student from each of the two classes, which class has the larger probability to win?*" Makar and Confrey (2005) and Hammerman and Rubin (2004) suggest that teacher' understanding of the basic statistical analysis, such as comparing two groups, can be very confused (e.g., wanting to compare individual data points rather than group trends). Hence we provided teachers some common statistics as choices for judging. The choices include "*It depends on the chosen student's ability of solving problems.*", "*By the means of two classes*", "*By the total scores of two classes*", "*By the number of having score four of two classes*", "*By the medians of two classes*" ,"*Other opinions*". In-service teachers could choose more than one choice as long as they thought the reasons they chose were correct. Then the first author guided them to discuss whether the following statement is correct or not:"If the mean of one distribution is larger than the mean of the other, then its probability of winning is larger". We found out that few teachers could successfully construct

In C. Reading (Ed.), *Data and context in statistics education: Towards an evidence-based society. Proceedings of the Eighth International Conference on Teaching Statistics (ICOTS8, July, 2010), Ljubljana, Slovenia*. Voorburg, The Netherlands: International Statistical Institute. www.stat.auckland.ac.nz/~iase/publications.php

appropriate counter-examples within about ten minutes. For this reason, we decided to modify the task by providing more assistance. That was the reason why the second task was provided.

Question:

If we randomly choose one student from each of the two classes, which class has the larger probability to win?

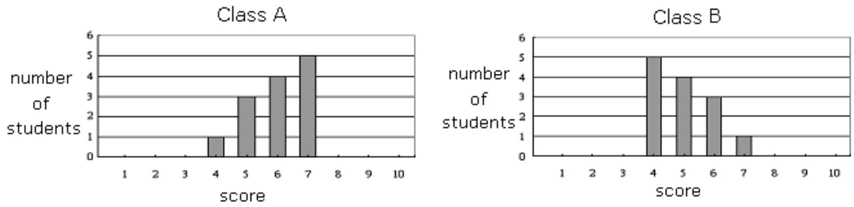

Figure 1. The first task

In the second task, we also provided two distributions which are shown in Figure 2. The question is the same as the first task. There are 20 students in each of two classes. Regarding the distributions of the two classes, all of the three statistics (mean, mode, median) in class B are larger than those in class A. However, the probability that one student chosen randomly from class B gets more scores than one from class A is smaller than the probability that one student chosen randomly from class A gets more scores than one from class B (e.g., $P(A_i>B_j) > P(B_j>A_i)$). 37 pre-service teachers were asked to compare two classes firstly. The choices are similar to the first task. For encouraging pre-service teachers to make counter-examples to refute that if all of three statistics (mean, mode and median) of one distribution is larger than those of the other, then its probability of winning is larger, the first author guided them to discuss the conditions which should be satisfied by counter-examples. Then they were asked to create a counter-example in which the number of each class was 15.

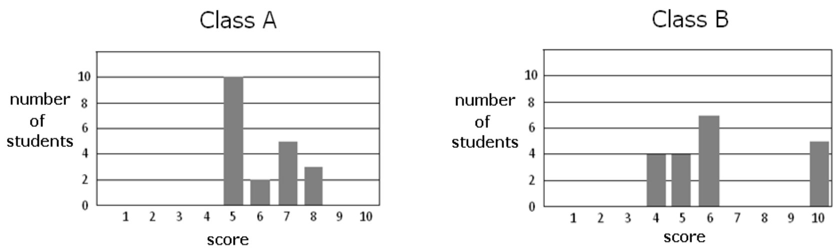

Figure 2. The second task

RESULTS

Task I

Table 1 shows that most teachers tended to regard the mean as a representative of the whole group. Only three teachers made their judgment by calculating the probability. Most teachers had wrong myths that if class A had larger mean, the probability that a student randomly chosen from class A than one randomly chosen from class B would be larger (e.g, mean(A)>mean(B) => $P(A_i>B_j) > P(B_j>A_i)$, for some i,j=1,2,…,22). We found that they were not aware that using the concept of the mean was not appropriate.

In order to disprove the above misconception, these teachers were asked to make a counter-example in which the mean of class A was larger than the mean of class B, but the probability of winning in class A was smaller than that in class B (i.e. mean(A)>mean(B) and $P(A_i>B_j) < P(B_j>A_i)$). The following Figure 3 is a counter-example made from one of them. The mean of class A is larger than the mean of class B (6 > 4.2), but the probability of winning in class A is smaller than that in class B (0.467 < 0.533). His thinking of making a counter-example has a main idea that the number of class B students who scored better than all class A students should exceed half of total number of class B students. After this teacher presented and explained his counter-example, there were still many teachers rejected to accept that using the mean was not appropriate. They thought those counter-examples were just only special cases, they couldn't represent all situations.

Table 1. 22 in-service teachers' responses of comparing two groups

How do you make this judgment? (Multiple selections accepted)	number	percent
1. It depends on the chosen student's ability of solving problems.	1	4.5%
2. By the means of two classes.	18	81.8%
3. By the total scores of two classes.	7	31.8%
4. By the number of having score four.	0	0%
5. By the medians of two classes.	4	18.1%
6. Other opinions. (correct opinion: by probability)	7 (3)	31.8% (13.6%)

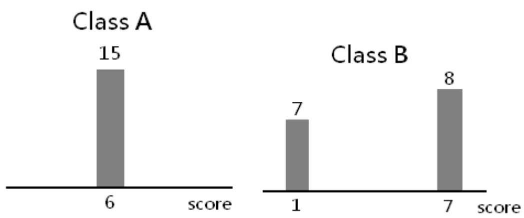

Figure 3. One counter-example in Task I

Task II

Table 2 shows that more than half of pre-service teachers regarded that the mean of one class could represent this class. Eleven of fifteen teachers who chose "other opinions" made their judgments by calculating the probability. After one of eleven students calculated the probability on the blackboard, all of the others who chose wrong reasons accepted the fact that judging by the probability was correct. They accepted it just because they couldn't find anything wrong with the probability, but they couldn't realize why judging by the means were wrong. After a period of discussing, they realized that they didn't recognize the meaning of the question correctly－the key point was the probability of winning not the scores of winning.

Table 2. 37 pre-service teachers' responses of comparing two groups

How do you make this judgment? (Multiple selections accepted)	number	percent
1. I have no idea	0	0%
2. It depends on the chosen student's ability of solving problem	1	2.7%
3. By the means of two classes	20	54.1%
4. By the modes of two classes	8	21.6%
5. By the medians of two classes.	11	29.7%
6. Other opinions. (correct opinion: by probability)	15 (11)	40.5% (29.7%)

In the second activity, after teachers knew that the statistics couldn't replace the probability, the first author asked them to create another counter-example in which the total number of students was different from the original one. This counter-example fills the following conditions: "mean(A)<mean(B), median(A)<median(B), mode(A)<mode(B), but $P(A_i>B_j) > P(B_j>A_i)$". The result shows that there were two different types of thinking when creating counter-examples. One is interactively changing both distributions and the other is controlling one distribution to adjust the other.

Type I: interactively changing both distributions

The original number of the original question was twenty (in one class), but we asked these teachers to construct a counter-example which had fifteen students in one class. These teachers tried to decrease the figures according to proportion. Even teachers in this type had the same thinking; there were different strategies that they made. Three different strategies in this type are try-and-error, proportion-based and properties-oriented.

Type II: controlling one distribution to adjust the other

Pre-service teachers in this type set up class A by dividing these students into two nearby categories and making sure that the mode and the median could be in the lower-scores category. The most difficult problem which teachers met was the balance between the probability and the mean. If the number of high-score students is larger, the mean will be larger and the probability of that class may be larger too. Therefore, students in the second type mainly tried to solve this problem. Two different strategies in this type are statistics-modified and unfixing a subgroup. Description of each strategy will be elaborated in oral presentation.

CONCLUSION

While in-service teachers were asked to make counter-examples, few of them could make an example to show the inconsistent relations between mean and probability of two distributions in about ten minutes. For encouraging pre-service teachers to make examples, the first author guided them to discuss the conditions which should be satisfied by examples. After about fifteen minutes, many of them could make an example. Their strategies can be identified into two types of interactively changing both distributions and controlling one distribution to adjust the other. The strategies of changing both distributions include try-and-error, proportion-based, and property-oriented. The strategies of controlling one distribution include statistics-modified and unfixing one sub-group.

Most of in-service teachers did not think this task was good to understand how and why some statistics or probability was applied in context. However, pre-service agreed that this was good for them to resolve their own misunderstanding while comparing two groups, and to make the statuses of premise and conclusion from implicit to explicit was a new experience in learning mathematics.

REFERENCES

Hammerman, J. K., & Rubin, A. (2004, November). Strategies for managing statistical complexity with new software tools. *Statistics Education Research Journal, 3*(2), 17-41. Online: www.stat.auckland.ac.nz/~iase/serj/SERJ3(2)_Hammerman_Rubin.pdf.

Makar, K., & Confrey, J. (2002). Comparing two distributions: Investigating secondary teachers' statistical thinking. In B. Phillips (Ed.), *Proceedings of the Sixth International Conference on Teaching of Statistics*, Cape Town. Voorburg, The Netherlands: International Statistical Institute.

Makar, K., & Confrey, J. (2005). using distributions as statistical evidence in well-structured and ill-structured problems. In K. Makar (Ed.), *Reasoning about distribution: A collection of current research studies. Proceedings of the Fourth International Research Forum on Statistical Reasoning, Thinking, and Literacy (SRTL-4)*, University of Auckland, New Zealand, 2-7 July. Brisbane: University of Queensland.

Watson, J. M., & Moritz, J. B. (1999). The beginning of statistical inference: Comparing two data sets. *Educational Studies in Mathematics*, *37*, 145-168.