

# Final Exam – Part 2

Jasmeet Singh Saini - 0758054

2023-04-17

## Highschool Senior Citizens

The National Center of Education Statistics conducted a survey of high school seniors, collecting test data on reading, writing, and several other subjects. Here we examine a simple random sample of 200 students from this survey. Side-by-side box plots of reading and writing scores as well as a histogram of the differences in scores are shown below. Also provided below is a histogram of randomized averages of paired differences of scores (read - write), with the observed difference ( $\bar{x}_{\text{read-write}} = -0.545$ ) marked with a red vertical line. The randomization distribution was produced by doing the following 1000 times: for each student, the two scores were randomly assigned to either read or write, and the average was taken across all students in the sample.

a. Is there a clear difference in the average reading and writing scores?

Yes, there is a difference in the average reading and writing scores. It can be seen from the box plot that the median of writing score is greater than the median of reading score. Thus, there is a difference in mean scores of reading and writing.

b. Are the reading and writing scores of each student independent of each other?

No, the reading and writing scores of each student are not independent of each other. Both the reading as well as the writing are related with each other. Also, if the reading score decreases then there will be changes in the writing score.

c. Create hypotheses appropriate for the following research question: is there an evident difference in the average scores of students in the reading and writing exam?

### Hypothesis

The  $H_0$  is Null hypothesis and  $H_A$  is Alternative hypothesis:

$H_0$  : There is no significant difference in the average scores of students in the reading and writing exam

vs.

$H_A$  : There is a significant difference in the average scores of students in the reading and writing exam

**d. Is the average of the observed difference in scores ( $\bar{X}$  read-write=-0.545) consistent with the distribution of randomized average differences? Explain.**

Yes, the average of the observed difference in scores ( $\bar{X}$  read - write = -0.545) is consistent with the distribution of randomized average differences. The reason for this is that the observed difference, that is,  $\bar{X}$  read - write = - 0.545 (red line) fall in the range of randomization distribution.

**e. Do these data provide convincing evidence of a difference between the average scores on the two exams? Estimate the p-value from the randomization test, and conclude the hypothesis test using words like “score on reading test” and “score on writing test.”**

In the histogram of “1000 means of randomized differences”, we have observe that the observed difference of -0.545 (i.e., red line) value is near the center and from this we can say that  $p$ -value will be more than 0.05 as the most of the data is near to center. Thus, we can estimate the p-value from the randomization test which will be more than the level of significance (0.05). Therefore, we fail to reject the null hypothesis and conclude that there is not enough evidence to suggest a difference between the average scores on the reading and writing exams.