List of exercises

Exercise 1 *Global warming.* 1
Exercise 2 *Gifted Children Part III.* 2
Exercise 3 *Paired or not.* 2
Exercise 4 *Math scores of 13 year old.* 2
Exercise 5 *Math scores of 13 year old. Part 2.* 2
Exercise 6 *Virus prevalence by gender.* 3
Exercise 7 *Gifted children Part IV.* 3
Exercise 8 *Maturity Variances.* 3

* **Ex. 1 —** **Global warming.** Is there strong evidence of global warming? Let's consider a small scale example, comparing how temperatures have changed in the US from 1968 to 2008. The daily high temperature reading on January 1 was collected in 1968 and 2008 for 51 randomly selected locations in the continental US. Then the difference between the two readings (**temperature in 2008 - temperature in 1968**) was calculated for each of the 51 different locations. The average of these 51 values was 1.1 degrees with a standard deviation of 4.9 degrees. We are interested in determining whether these data provide strong evidence of temperature warming in the continental US.

1. Is there a relationship between the observations collected in 1968 and 2008? Or are the observations in the two groups independent? Explain.
2. Write hypotheses for this research in symbols and in words.
3. Check the conditions required to complete this test.
4. Calculate the test statistic and find the p-value.
5. What do you conclude? Interpret your conclusion in context.
6. What type of error might we have made? Explain in context what the error means.
7. Calculate a 95% confidence interval for the average difference between the temperature measurements between 1968 and 2008.
8. Interpret this interval in context.

** **Ex. 2 — Gifted Children Part III.** Researchers investigating characteristics of gifted children collected data from schools in a large city about children who were identified as gifted children soon after they reached the age of four and their parents. The sample was taken at random. Import the dataset **gifted.csv** and address the following questions:

1. Are the IQs of mothers and the IQs of fathers in this data set related? Explain.
2. Conduct a hypothesis test to evaluate if the scores are equal on average. Define the difference as 'mother' - 'father'. Make sure to clearly state your hypotheses, check the relevant conditions, and state your conclusion in the context of the data.

** **Ex. 3 — Paired or not.** In each of the following scenarios, determine if the data are paired.

1. We would like to know if Intel's stock and Southwest Airlines' stock have similar rates of return. To find out, we take a random sample for Intel's stock and for Southwest's stock in the same period.
2. We randomly sample 50 items from Target stores and note the price for each. Then we visit Walmart and collect the price for each of those same 50 items.
3. A school board would like to determine whether there is a difference in average SAT scores for students at one high school versus another high school in the district. To check, they take a simple random sample of 100 students from each high school.

** **Ex. 4 — Math scores of 13 year old.** The National Assessment of Educational Progress tested a simple random sample of 1,000 thirteen year old students in both 2004 and 2008 (two separate simple random samples). The average and standard deviation in 2004 were 257 and 39, respectively. In 2008, the average and standard deviation were 260 and 38, respectively.

1. Calculate a 90% confidence interval for the change in average scores from 2004 to 2008 (use $\mu_{2008} - \mu_{2004}$), and interpret this interval in the context of the application. (Reminder: check conditions.)
2. Do these data provide strong evidence that the average math score for 13 year old students has changed from 2004 to 2008? Use a 10% significance level. Assume that the given standard deviations are the population standard deviations
3. It is possible that your conclusion is incorrect? What type of error is possible for this conclusion? Explain.
4. Are your conclusions on the test consistent with the 90% confidence interval? Explain.

** **Ex. 5** — **Math scores of 13 year old. Part 2.** Consider the exercise about 'Math Scores of 13 year olds children'

1. Perform the test again assuming that the given standard deviations are the sample standard deviations (so the population standard deviations are unknown). Assume variances are equal. Please use the t distribution even if the sample size is greater than 30. Use a 10% significance level.
2. Perform the test again assuming that the given standard deviations are the sample standard deviations but they are unequal. Please use the t distribution even if the sample size is greater than 30. Use a 10% significance level.

** **Ex. 6** — **Virus prevalence by gender.** In a random sample, 42 of 921 women and 51 of 957 men were infected by a new virus.

1. Is there a significant difference between the proportion of women and the proportion of men infected by the new virus?

** **Ex. 7** — **Gifted children Part IV.**Researchers investigating characteristics of gifted children collected data from schools in a large city about children who were identified as gifted children soon after they reached the age of four. The sample was taken at random. In this study, along with variables on the children, the researchers also collected data on the mother's and father's IQ of the randomly sampled gifted children. Import the dataset **gifted.csv** and address the following questions:

1. Perform a hypothesis test to evaluate if these data provide convincing evidence that the proportion of mothers whose IQ is greater than 120 is different from that of fathers whose IQ is greater than 120. Use a significance level of 0.01.
2. Calculate a 99% confidence interval for the difference in proportions of mothers whose IQ is greater than 120 and fathers whose IQ is greater than 120.
3. Do your results from the hypothesis test and the confidence interval agree? Explain.

** **Ex. 8** — **Maturity Variances.** The research staff of Investors Z, a financial trading firm,was interested in determining if there is a difference in the variance of the maturities of AAA-rated industrial bonds compared to CCC-rated industrial bonds. A random sample of 17 AAA-rated bonds resulted in a sample variance $s_x^2 = 123.35$, and a random sample of 11 CCC-rated bonds resulted in a sample variance $s_y^2 = 8.02$. Use $\alpha = 0.02$