Problem Set 9

Statistical Methods In Engineering And Science

Due Date: 10:00 PM, June 2, 2023

Last Update: May 27, 2023

Prof. Alexander Giessing

Spring Quarter, 2023

Study Group: _____ _____ _____

Please upload your solution in a single pdf file on Canvas. Include all calculations, R-code, and figures (if applicable). All data sets are available on Canvas https://canvas.uw.edu/courses/1584511.

Question 1. The level of lead in the blood was determined for a sample of 152 male and a sample of 86 female hazardous-waste workers, resulting in a mean $\pm$ standard error of 5.5 ± 0.3 for men and 3.8 ± 0.2 for women. Let μ_d be the difference between true average blood lead levels for male and female workers.

(a) Is there substantial difference in the mean level of lead in the blood between these two groups? Conduct a hypothesis test at a significance level of 5%. Formulate the null and alternative hypothesis, propose a test statistic, describe the decision rule, and report your decision.

(b) What is the power of the test if the difference μ_d is $\Delta' = 0.5$? What does this mean?

(c) Construct a 95% confidence interval for the difference between true average blood levels for male and female workers. Using this interval alone, what do you conclude for the hypotheses in part (a)?

(d) What assumptions do you make in the above calculations?

Question 2. A survey of 81 high school teachers finds that 38 are very satisfied with their jobs.

(a) Is there statistical evidence that less than 50% of the high school teachers are very satisfied with their job? Conduct a hypothesis test at a significance level of 5%. Formulate the null and alternative hypothesis, propose a test statistic, describe the decision rule, and report your decision.

(b) Construct the 95% confidence interval that allows you to make the same decision about null and alternative hypothesis as in part (a).

(c) Show that the power function $f(p)$ of the test in part (a) is approximately

$$f(p) \approx \Phi\left(\frac{-0.8225}{\sqrt{p(1-p)}} + \frac{0.5-p}{\sqrt{p(1-p)/9}}\right).$$

(d) Calculate the power and the probability of a type II error of the test in part (a) at $p = 0.38$ and interpret these numbers in context.

(e) Additional 150 high school teachers are surveyed of which 71 report that they were very satisfied with their jobs. If the answers of all high school teachers were combined, how would this effect the probabilities of the type I and type II error of the test in part (a)?

Question 3. Measles is a highly contagious infectious disease which was declared eradicated from the US in 2000 due to the success of vaccination efforts. Consider the following data on vaccination rates for measles collected from students entering kindergarten in Washington.

Year	**Sample Size**	**Sample Vaccination Rate**
2000	8750	95.44%
2009	7875	93.60%

(a) Has the vaccination rate of students stayed the same or decreased since 2000? Conduct a hypothesis test at a significance level of 1% for $H_0 : p_{2000} = p_{2009}$ vs. $H_1 : p_{2000} > p_{2009}$, where p_{2000} and p_{2009} are the population vaccination rates for years 2000 and 2009, respectively. Propose the most appropriate test statistic and report your decision.

(b) To achieve herd immunity for measles, at least 93% of the population needs to be vaccinated. Was herd immunity achieved in the population of students entering kindergarten in 2009? Conduct a hypothesis test at a significance level of 1%. Formulate the null and alternative hypotheses, propose a test statistic, and report your decisions.

(c) How large should the sample size be so that with 99% confidence the population vaccination rate of 2009 is at least 93%? (Assume that adding additional samples does not change the sample vaccination rate.)

Question 4. Reconsider the data set `olympic1500m.txt` from Problem Set 8. We again want to investigate whether speed skaters have an advantage in the 1500m race if they start in the outer lane.

(a) Let p be the proportion of skaters who started in the outer lane and won the race. Formulate the appropriate null and alternative hypothesis about p. Test the null hypothesis at significance level 5% using a large sample test.

(b) On second thought, you realize that 23 measurements constitute a rather small sample. You therefore discard your approach from part (a). Develop a small sample test from first principle. Explain your reasoning. *(Hint: Slides 6f, Week 10 Lectures, Part 1.)*

Question 5. Please fill in the course evaluation. I would really appreciate your input! If at least 75% of you fill in the course evaluation, I will award each of you **20 extra points**. You should have received the link to the evaluation in an email. I've also posted the link on Canvas. Thank you for a great quarter :-)

– THE END –