

Week 4 Statistics Half of the Tutorial Questions

Part A

Alice recruited nine pairs of identical twins for a study of two cholesterol-reducing drugs, A and B, with the aim of showing that drug A gave higher reductions in cholesterol than drug B. One of the twins in each pair was given drug A and the other was given drug B, where the choice was made at random. The amount by which cholesterol was reduced in each subject (mg/dL) is given in the following table:

Pair	1	2	3	4	5	6	7	8	9
Drug A	74	55	61	47	53	74	52	40	50
Drug B	63	58	49	41	50	69	59	31	44
Difference	11	-3	12	6	3	5	-7	9	6

The sample mean difference in cholesterol reduction between drug A and drug B was 4.67 mg/dL with a sample standard deviation of 6.265 mg/dL.

- State the null and alternative hypotheses of interest in terms of μ , the mean difference in cholesterol reduction between drug A and drug B in twins in this population.
- Calculate the t statistic to test this null hypothesis.
- What is the corresponding p -value? What do you conclude?
- Based on this data, calculate a 95% confidence interval for the mean difference in cholesterol reduction between drug A and drug B.
- Suppose we wanted to carry out a new study that could estimate the mean difference in cholesterol reduction with a margin of error of 2 mg/dl. What sample size should the new study use if we keep the 95% confidence level? (Assume that population standard deviation, σ , is 6.265)

Part B

A simple alternative to the t test in Part A is to count the number of twins where there was a positive difference between drug A and drug B.

- a) State the null and alternative hypotheses of interest in terms of p , the probability that, in a random pair of identical twins, the one with drug A will have a greater reduction in cholesterol than the one with drug B.
- b) Assuming H_0 is true, what is the distribution of X , the number of positive differences in 9 trials?
- c) Based on the observed number of positive differences, what is the p -value for this test? What do you conclude?

Part C

The health officials carry out a survey for women aged 20-60 years and recorded the time in hours that pain lasted at the injection site after the second covid jab. Based on an earlier survey with a similar population of women, the pain duration is assumed to follow a normal distribution with a mean of 18.3 hours and a standard deviation of 6.9 hours.

- a) What is the probability that average pain duration time of a randomly selected 5 women in the survey had pains less than 12 hours?

b) What is the shortest time of experiencing pain that would be in the upper 10% of pain durations?

c) The health officials carried out a similar survey of 20 men aged between 20 and 60 years and recorded the sample mean of 15.5 hours with a standard deviation of 5.6 hours. They believe that mean pain duration for men is less than that of women. Suppose μ is the mean pain duration for men aged between 20 and 60 years who had the second job. Carry out a hypothesis test to test the health officials' belief.