Problem Sheet-3

1. A sample of 10 fish were caught at lake A and their PCB concentrations were measured using a certain technique. The resulting data in parts per million were

   Lake A:$11.5, 10.8, 11.6, 9.4, 12.4, 11.4, 12.2, 11, 10.6, 10.8$

   In addition, a sample of 8 fish were caught at lake B and their levels of PCB were measured by a different technique than that used at lake A. The resultant data were

   Lake B:$11.8, 12.6, 12.2, 12.5, 11.7, 12.1, 10.4, 12.6$

   If it is known that the measuring technique used at lake A has a variance of .09 whereas the one used at lake B has a variance of .16, could you reject (at the 5 percent level of significance) a claim that the two lakes are equally contaminated?

2. A method for measuring the pH level of a solution yields a measurement value that is normally distributed with a mean equal to the actual pH of the solution and with a standard deviation equal to .05. An environmental pollution scientist claims that two different solutions come from the same source. If this were so, then the pH level of the solutions would be equal. To test the plausibility of this claim, 10 independent measurements were made of the pH level for both solutions, with the following data resulting.
   Measurements of solution A: $6.24, 6.31, 6.28, 6.30, 6.256.26, 6.24, 6.29, 6.22, 6.28$
   Measurements of solution B: $6.27, 6.25, 6.33, 6.27, 6.246.31, 6.28, 6.29, 6.34, 6.27$

   (a) Do the data disprove the scientist's claim? Use the 5 percent level of significance.

   (b) What is the p$-$value?

3. The following are the values of independent samples from two different populations.
   Sample 1: $122, 114, 130, 165, 144, 133, 139, 142, 150$
   Sample 2: $108, 125, 122, 140, 132, 120, 137, 128, 138$
   Let μ_1 and μ_2 be the respective means of the two populations. Find the p$-$value of the test of the null hypothesis $H_0 : \mu_1 \leq \mu_2$ versus the alternative $H_1 : \mu_1 > \mu_2$ when the population standard deviations are $\sigma_1 = 10$ and

   (a) $\sigma_2 = 5$;

(*b*) $\sigma_2 = 10$.

4. The viscosity of two different brands of car oil is measured and the following data resulted:
   Brand 1:$10.62, 10.58, 10.33, 10.72, 10.44, 10.74$
   Brand 2:$10.50, 10.52, 10.58, 10.62, 10.55, 10.51, 10.53$

   Test the hypothesis that the mean viscosity of the two brands is equal, assuming that the populations have normal distributions with equal variances.

5. Twenty-five men between the ages of 25 and 30, who were participating in a wellknown heart study carried out at AIIMS Jodhpur, were randomly selected, where assuming that the populations have normal distributions with equal variances. Of these, 11 were smokers and 14 were not. The following data refer to readings of their systolic blood pressure.
   Smoker: $124, 134, 136, 125, 133, 127, 135, 131, 133, 125, 118$
   Nonsmoker: $130, 122, 128, 129, 118, 122116, 127, 135, 120, 118122, 120, 115, 123$
   Use these data to test the hypothesis that the mean blood pressures of smokers and nonsmokers are the same.

6. In a 1943 experiment (Whitlock and Bliss, "A Bioassay Technique for Antihelminthics," Journal of Parasitology, 29, pp. 48–58) 10 albino rats were used to study the effectiveness of carbon tetrachloride as a treatment for worms, where assuming that the populations have normal distributions with equal variances. Each rat received an injection of worm larvae. After 8 days, the rats were randomly divided into two groups of 5 each; each rat in the first group received a dose of .032 cc of carbon tetrachloride, whereas the dosage for each rat in the second group was .063 cc. Two days later the rats were killed, and the number of adult worms in each rat was determined. The numbers detected in the group receiving the .032 dosage were $421, 462, 400, 378, 413$, whereas they were $207, 17, 412, 74, 116$ for those receiving the .063 dosage. Do the data prove that the larger dosage is more effective than the smaller?

7. In a certain experimental laboratory, a method A for producing gasoline from crude oil is being investigated. Before completing experimentation, a new method B is proposed. All other things being equal, it was decided to abandon A in favor of B only if the average yield of the latter was clearly greater. The yield of both processes is assumed to be normally distributed. However, there has been insufficient time to ascertain their

true standard deviations, although there appears to be no reason why they cannot be assumed equal. Cost considerations impose size limits on the size of samples that can be obtained. If a 5 percent significance level is all that is allowed, what would be your recommendation based on the following random samples? The numbers represent percent yield of crude oil.

$$A: \ 23.2, 26.6, 24.4, 23.5, 22.6, 25.7, 25.5$$

$$B: \ 25.7, 27.7, 26.2, 27.9, 25.0, 21.4, 26.1$$

Assuming that the populations have normal distributions with equal variances.

8. A study was instituted to learn how the diets of women changed during the winter and the summer. A random group of 12 women were observed during the month of July and the percentage of each woman's calories that came from fat was determined. Similar observations were made on a different randomly selected group of size 12 during the month of January, where the populations have normal distributions with equal variances. The results were as follows:

$$\text{July: } 32.2, 27.4, 28.6, 32.4, 40.5, 26.2, 29.4, 25.8, 36.6, 30.3, 28.5, 32.0$$

$$\text{January: } 30.5, 28.4, 40.2, 37.6, 36.5, 38.8, 34.7, 29.5, 29.7, 37.2, 41.5, 37.0$$

Test the hypothesis that the mean fat percentage intake is the same for both months. Use the

   (*a*) 5 percent level of significance;

   (*b*) 1 percent level of significance.

9. The following data summary was obtained from a comparison of the lead content of human hair removed from adult individuals that had died between 1880 and 1920 with the lead content of present-day adults. The data is drawn from the populations have normal distributions with equal variances. Also the data are in units of micrograms, equal to one-millionth of a gram.

   (*a*) Do the following data establish, at the 5 percent level of significance, that the mean lead content of human hair is less today than it was in the years between 1880 and 1920? Clearly state what the null and alternative hypotheses are.

Sample Data		
	1880-1920	Today
Sample size	30	100
Sample mean	48.5	26.6
Sample standard deviation	14.5	12.3

(b) What is the p-value for the hypothesis test in part (a)?

10. To verify the hypothesis that blood lead levels tend to be higher for children whose parents work in a factory that uses lead in the manufacturing process, researchers examined lead levels in the blood of 33 children whose parents worked in a battery manufacturing factory (Morton, D., Saah, A., Silberg, S., Owens, W., Roberts, M., and Saah, M., "Lead Absorption in Children of Employees in a Lead-Related Industry," American Journal of Epidemiology, 115, 549–555, 1982). Each of these children was then matched by another child who was of similar age, lived in a similar neighborhood, had a similar exposure to traffic, but whose parent did not work with lead. The blood levels of the 33 cases (sample 1) as well as those of the 33 controls (sample 2) were then used to test the hypothesis that the average blood levels of these groups are the same. If the resulting sample means and sample standard deviations were

$$\bar{x_1} = .015, s_1 = .004, \bar{x_2} = .006, s_2 = .006$$

find the resulting p-value. Assume that variance of two populations are unequal.

11. A random sample of 10 observations from population A has a sample mean of 152.30 and a sample standard deviation of 1.83. A random sample of 8 observations from population B has a sample standard deviation of 1.94. If the p-value for the one-sided hypothesis test with an alternative hypothesis $H_A : \mu_A > \mu_B$ is less than 1%, what can you say about the sample mean of the observations from population B?

12. The breaking strengths of $n = 20$ bundles of wool fibers have a sample mean $\bar{x} = 436.5$ and a sample standard deviation $s_x = 11.90$. In addition, the breaking strengths of $m = 25$ bundles of synthetic fibers have a sample mean $\bar{y} = 452.8$ and a sample standard deviation $s_y = 4.61$. Answer the following questions without assuming that

the two population variances are equal. Does a one-sided hypothesis test with size $\alpha = 0.01$ accept or reject the null hypothesis that the synthetic fiber bundles have an average breaking strength no larger than the wool fiber bundles?

13. A random sample of $n = 16$ one-kilogram sugar packets of brand A have weights with a sample mean $\bar{x} = 1.053$ kg and a sample standard deviation $s_x = 0.058$ kg. In addition, a random sample of $m = 16$ one-kilogram sugar packets of brand B have weights with a sample mean $\bar{y} = 1.071$ kg and a sample standard deviation $s_y = 0.062$ kg. Is it safe to conclude that brand B sugar packets weigh slightly more on average than brand A sugar packets? Consider $\alpha = 0.05$.

14. The resilient moduli of $n = 10$ samples of a clay mixture of type A are measured and the sample mean is $\bar{x} = 19.50$. In addition, the resilient moduli of $m = 12$ samples of a clay mixture of type B are measured and the sample mean is $\bar{y} = 18.64$. Suppose that the experimenter wishes to use values $\sigma_A = \sigma_B = 1.0$ for the standard deviations of the resilient modulus of the two types of clay. What is the exact two-sided p-value for the null hypothesis that the two types of clay have equal average values of resilient modulus?

15. An experiment was conducted to investigate how the corrosion properties of chilled cast iron depend upon the chromium content of the alloy. A collection of $n = 12$ samples of chilled cast iron with 0.1% chromium content provided corrosion rates with a sample mean of $\bar{x} = 2.462$ and a sample standard deviation of $s_x = 0.315$, while a collection of $m = 13$ samples of chilled cast iron with 0.2% chromium content provided corrosion rates with a sample mean of $\bar{y} = 2.2964$ and a sample standard deviation of $s_y = 0.297$. Conduct a hypothesis test to investigate whether there is any evidence that the chromium content has an effect on the corrosion rate of chilled cast iron.

16. The mean lifetime of a sample of 9 light bulbs was observed to be 1309 hours with a standard deviation of 420 hours. A second sample of 16 bulbs chosen from a different batch showed a mean lifetime of 1205 hours with a standard deviation of 390 hours. Let us test to see whether there is a significant difference between the means of the two batches, assuming that the population variances are the same. What happen if the population variances are not equal?