Tutorial Worksheet 3 - Testing of Hypothesis

Objective Questions

1. Consider a hypothesis H_0 where $\phi_0 = 5$ against H_1 where $\phi_1 > 5$. The test is?

   (a) Right tailed (b) Left tailed (c) Two tailed (d) Cross tailed

2. Type 1 error occurs when?

   (a) We reject H_0 if it is True (b) We reject H_0 if it is False (c) We accept H_0 if it is True (d) We accept H_0 if it is False

3. The probability of Type 1 error is referred as?

   (a) 1- α (b) β (c) α (d) Variance

4. What is the assumption made for performing the hypothesis test with T-distribution?

   (a) the distribution is non-symmetric (b) the distribution has more than one modal class (c) the distribution has a constant variance (d) the distribution follows a normal distribution

5. If a hypothesis is rejected at 0.6 Level of Significance then

   (a) it will be rejected at any level (b) it must be rejected at 0.5 level (c) it may be rejected at 0.5 level (d) it cannot be rejected at 0.5 level

6. In a two-tailed test when a Null Hypothesis is rejected for a True Alternative Hypothesis then it has

   (a) Type 1 error (b) Type 2 error (c) No error (d) Many errors

7. In a hypothesis test, what does the p value signify?

   (a) smallest level of significance for rejection of Null Hypothesis.
   (b) largest level of significance for rejection of Null Hypothesis.
   (c) smallest level of significance for acceptance of Null Hypothesis.
   (d) smallest level of significance for acceptance of Null Hypothesis.

8. The choice of one-tailed test and two-tailed test depends upon

   (a) Null hypothesis (b) Alternative hypothesis (c) Composite hypothesis (d) None of these

9. Arrange the following steps in the process of hypothesis testing in proper sequence:

   (A) Select the level of significance
   (B) Setup null and alternative hypothesis
   (C) Establish the decision rule
   (D) Performance computations
   (E) Select test statistics
   (F) Draw conclusion

   Choose the correct answer from the options given below

   (a) A, B, C, D, E, F (b) A, B, E, D, C, F (c) B, A, C, D, E, F (d) B, A, E, C, D, F

10. Match the items of List - II with the items of List - I and select the code of correct matching Some matchings from Column A and Column B are given below. Select the correct matching?

    (A) (a) - (i), (b) - (ii), (c) - (iii), (d) - (iv) (B) (a) - (iv), (b) - (iii), (b) - (ii), (d) - (i) (C) (a) - (i), (b) - (ii), (c) - (iv), (d) - (iii) (D) (a) - (ii), (b) - (i), (c) - (iii), (d) - (iv)

List - I		List - II	
(a)	Chi-square Test	(i)	Testing the significance of the differences of the mean values among more than two sample groups.
(b)	ANOVA (F-test)	(ii)	Testing the goodness of fit of a distribution
(c)	Z-test	(iii)	Testing the significance of the difference of the mean values between two large sized samples
(d)	t - test	(iv)	Testing the significance of the difference of the mean values between two small sized samples when population standard deviation is not available

Subjective Questions

Problem 1.

A drug company is testing a drug intended to increase heart rate. A sample of 100 yielded a mean increase of 1.4 beats per minute, with a standard deviation known to be 3.6. Since the company wants to avoid marketing an ineffective drug, it proposes a 0.001 significance level. Should it market the drug?

Problem 2.

The accompanying data on the breakdown voltage of electrically stressed circuits was read from a normal probability plot that appeared in the article "Damage of Flexible Printed Wiring Boards Associated with Lightning-Induced Voltage Surges" (IEEE Transactions on Components, Hybrids, and Manuf. Tech., 1985: 214-220). The straightness of the plot gave strong support to the assumption that breakdown voltage is approximately normally distributed.

$$1470, 1510, 1690, 1740, 1900, 2000, 2030, 2100, 2190, 2200, 2290, 2380, 2390, 2480, 2500, 2580, 2700$$

(a) $P(\bar{X} \geq 2200)$

(b) $P(\bar{X} \leq 2000)$

(c) $P(1900 \leq \bar{X} \leq 2100)$

(d) Confidence Interval of the population mean at 95%

Problem 3.

A company manufacturing automobile tyres finds that tyre-life is normally distributed with a mean of 40,000 km and a standard deviation of 3000 km. It is believed that a change in the production process will result in a better product and the company has developed a new tyre. A sample of 100 new tyres has been selected. The company has found that the mean life of these new tyres is 40,900 km. Can it be concluded that the new tyre is significantly better than the old one, using the significant level of 0.01?

Problem 4.

The Environmental Protection Department requires a mill to aerate its effluent so that the mean dissolved oxygen (DO) level is above 6.0 mg/L . To maintain compliance, the Department collects air samples at 12 randomly selected dates. The data collected is given as

$$5.85, 6.28, 6.50, 6.21, 5.94, 6.12, 6.65, 6.14, 6.34, 6.19, 6.29, 6.40$$

The Department requires strong evidence that the mean DO is high. It is required to test it with significance level, $\alpha = 1\%$. Answer the following questions. Assume that the population data follows normal distribution.

(a) What should be the null and alternate hypothesis?

(b) Which test statistics: z, t or χ^2 test is applicable in this case?

(c) According to your test statistics, calculate the critical value.

(d) From the sample data, calculate the test value.

(e) Decide if the null hypothesis is to be accepted or rejected.

Problem 5.

It is claimed that an automobile is driven on average more than 20,000 kilometers per year. To test the claim, a random sample of 25 automobile owners is asked to keep a record of the kilometers they travel. The random sample showed an average of 23500 kilometers and a standard deviation of 3900 kilometers. It is planned to test the above with parametric-based hypothesis testing. Assume a 5% confidence level.

(a) What should be the null and alternate hypothesis?

(b) Which test statistics: z, t or χ^2 test is applicable in this case?

(c) According to your test statistics, calculate the critical value.

(d) From the sample data, calculate the test value.

(e) Decide if the null hypothesis is to be accepted or rejected.

Problem 6.

A manufacturer of electric batteries claims that the average capacity of a certain type of battery that the company produces is at least 140 ampere-hours with a standard deviation of 2.66 ampere-hours. An independent sample of 20 batteries gave a mean of 138.47 ampere-hour. Test at 5% significance level $H_0 = 140$ against $H_1 < 140$. Can the manufacturer's claim be sustained on the basis of this sample?

Problem 7.

A manufacturer of car batteries claims that the life of the company's batteries is approximately normally distributed with a standard deviation equal to 0.9 per year. If a random sample of 10 of these batteries has a standard deviation of 1.2 years, do you think that $\sigma > 0.9$ year? Use a 0.05 level of significance.

Problem 8.

The alkalinity, in milligrams per litre, of water in the upper reaches of rivers in a particular region is known to be normally distributed with a standard deviation of 10mg/l. Alkalinity readings in the lower reaches of rivers in the same region are also known to be normally distributed, but with a standard deviation of 25mg/l. Ten alkalinity readings are made in the upper reaches of a river in the region and fifteen in the lower reaches of the same river with the following results.

Upper reaches	91	75	91	88	94	63	86	77	71	69	-	-	-	-	-
Lower reaches	86	95	135	121	68	64	113	108	79	62	143	108	121	85	97

Investigate, at the 1% level of significance, the claim that the true mean alkalinity of water in the lower reaches of this river is greater than in the upper reaches.

Problem 9.

A chemical engineer claims that the population mean yield of a certain batch process is 500 grams per milliliter of raw material. To check this claim he samples 25 batches each month. If the computed t-value falls between $-t_{0.05}$ and $t_{0.05}$, he is satisfied with this claim. What conclusion should he draw from a sample that has a mean $\bar{x} = 518$ grams per milliliter and a sample standard deviation $s = 40$ grams? Assume the distribution of yields to be approximately normal.

Problem 10.

An experiment was performed to compare the abrasive wear of two different laminated materials. Twelve pieces of material 1 were tested by exposing each piece to a machine measuring wear. Ten pieces of material 2 were similarly tested. In each case, the depth of wear was observed. The samples of material 1 gave an average (coded) wear of 85 units with a sample standard deviation of 4, while the samples of material 2 gave an average of 81 with a sample standard deviation of 5. Can we conclude at the 0.05 level of significance that the abrasive wear of material 1 exceeds that of material 2 by more than 2 units? Assume the populations to be approximately normal with equal variances.

Problem 11.

Pumpkins were grown under two experimental conditions. Two random samples of 11 and 9 pumpkins show the sample standard deviations of their weights as 0.8 and 0.5 respectively. Assuming that the weight distributions are normal, test the hypothesis that the true variances are equal, against the alternative that they are not, at the 10% level.

Problem 12.

In testing for the difference in the abrasive wear of the two materials in Problem 10. we assumed that the two unknown population variances are equal. Were we justified in making this assumption? Use a 0.10 level of significance.

Problem 13.

In a sale of 300 units of an item the following preferences for colors are observed for customers:

Color	Brown	Grey	Red	Blue	White	Total
Customers	88	65	52	40	55	300

Test the hypothesis that all colors are equally popular.

Problem 14.

A survey was conducted among 500 students who are studying either in "government funded collages" (GVT) or "privately funded colleges" (PVT). The objective of the survey is to see the choice of "classroom-based learning" (C) over

"Internet-based learning" (I). It is proposed to apply the χ^2 test to verify if there exists any association between "colleges" and "learning".

Categories	C	I	Total
GVT	75	125	200
PVT	60	240	300

(a) Decide the null and alternate hypotheses in this case. Justify your answer.

(b) Calculate the χ^2-value from the sample data.

(c) Test the hypothesis with 5% confidence level.

Problem 15.

Let X be the number of defects in printed circuit boards. A random sample of $n = 60$ printed circuit boards is taken and the number of defects is recorded. The results were as follows: Does the assumption of a Poisson distribution seem

Number of Defects	Observed Frequency
0	32
1	15
2	9
3	4

appropriate as a model for these data?