Assignment – 5

1. In a study, 44 randomly chosen students were asked to guess, to the nearest metre, the width of the lecture hall in which they were sitting. The data collected are given below:
   8, 9, 10, 10, 10, 10, 10, 10, 11, 11, 11, 11, 12, 12, 13, 13, 13, 14, 14, 14, 15, 15, 15, 15, 15, 15, 15, 15, 16, 16, 16, 17, 17, 17, 17, 18, 18, 20, 22, 25, 27, 35, 38, 40

(a) Can the above random sample be assumed to be drawn from a normal distribution?
(b) Let μ be the population mean. Test the null hypothesis H_0: $\mu = 13.1$ against the alternative H_1: $\mu \neq 13.1$
(c) Obtain a 95% confidence interval of μ using the CLT.
(d) Obtain a 95% bootstrap pivotal confidence interval for μ.

2. The data given below are from Charles Darwin's study of cross- and self-fertilization. Pairs of seedlings of the same age, one produced by cross-fertilization and the other by self-fertilization, were grown together so that the members of each pair were reared under nearly identical conditions. The aim was to demonstrate the greater vigour of the cross-fertilized plants. The data are the final heights of each plant after a fixed period of time. Suppose the vigour of a plant is measured by the height of the plant.

Pair	Cross-fertilized	Self-fertilized	Pair	Cross-fertilized	Self-fertilized
1	23.5	17.4	9	18.3	16.5
2	12.0	20.4	10	21.6	18.0
3	21.0	20.0	11	23.3	16.3
4	22.0	20.0	12	21.0	18.0
5	19.1	18.4	13	22.1	12.8
6	21.5	18.6	14	23.0	15.5
7	22.1	18.6	15	12.0	18.0
8	20.4	15.3			

(i) Let $D_i = C_i - S_i$ where C_i and S_i are the heights of the i-th pair of Cross-fertilized and Self-fertilized plants. Can the Di's be considered as a random sample from the normal distribution?
(ii) Test the null hypothesis $H_0 : \mu_C = \mu_S$ against $H_1 : \mu_C > \mu_S$ at 5% level of significance, where μ_C is the average height of the cross-fertilized plants and μ_S is the average height of the self-fertilized plants.
(iii) Obtain a 95% confidence interval for $\mu_C - \mu_S$.

3. The data given below come from a study investigating a new method of measuring body composition, and give the body fat percentage (% fat) and age for 14 normal adult females aged between 23 and 61 years.

Age	%fat
23	27.9
39	31.4
41	25.9
49	25.2
50	31.1
53	34.7
53	42.0
54	29.1
56	32.5
57	30.3
58	33.0
58	33.8
60	41.1
61	34.5

(a) Fit a simple linear regression model with "%fat" as the response variable and "Age" as the predictor variable. Report the obtained results.
(b) Comment on the goodness of the regression model fit.
(c) Write a small paragraph of 50 words communicating your findings to the layman (for e.g. as a statement to a local newspaper/ magazine).

4. These data come from Canadian records of pure-bred dairy cattle. They give butterfat (the natural fat contained in milk and dairy products) percentages for random samples of 10 mature cows from each of the two breeds - Ayrshire and Guernsey.

Ayrshire	Guernsey
3.74	4.54
4.01	5.18

3.77	5.75
3.78	5.04
4.10	4.64
4.06	4.79
4.27	4.72
3.94	3.88
4.11	5.28
4.25	4.66

(a) Can the random sample of butterfat percentages of Ayrshire breed cows be assumed to follow a normal distribution?
(b) Can the random sample of butterfat percentages of Guernsey breed cows be assumed to follow a normal distribution?
(c) Let μ_A be the average butterfat percentage for Ayrshire breed cows and μ_G be the same for the Guernsey breed cows. Test the null hypothesis $H_0 : \mu_A \leq \mu_G$ against the alternative $H_1 : \mu_A > \mu_G$ at 5% level of significance.
(d) Write a small paragraph of 50 words communicating your findings to the layman (for e.g. as a statement to a local newspaper/ magazine).

5. The data gives the yields of winter wheat (in Kg. per unit area) for the years 1970 and 1973 at twelve different sites in England. Has there been an increase in yield of Winter wheat in the year 1973 compared to that in the year 1970?

Location	Yield (1970)	Yield(1973)
Cambridge	46.81	32.61
Cockle Park	46.49	41.02
Rarpers Adams	44.03	50.23
Head1ey Hall	52.24	34.56
Morley	36.55	43.17
Myerscough	34.88	50.08
Rosemaund	56.14	38.99
Seale-Hayne	45.67	50.32
Sparsholt	42.97	47.49
Sutton Bonington	54.44	46.94
Terrington	54.95	39.13
Wye	48.94	59.72

(a) Formulate the null and alternative hypothesis appropriately.
(b) Carry out a test of the hypothesis at 5% level of significance and report the findings.
(c) Write a small paragraph of 50 words communicating your findings to the layman (for e.g. as a statement to a local newspaper/ magazine).