1. What is the Central Limit Theorem ? (10%)

2. Given the following data set,

Data on Lot Size and Number of Man-Hours, Company A

Production Run	Lot Size	Man-Hours
i	Xi	Yi
1	30	73
2	20	50
3	60	128
4	80	170
5	40	87
6	50	108
7	60	135
8	30	69
9	70	148
10	60	132

(a) Sketch the scatter plot of (X,Y). (5%)

(b) What can you say about the relationship between X and Y from this scatter plot ? (5%)

(c) Suppose after looking at the scatter plot, you decide to run a simple linear regression and obtain the regression line as

$$Y = 10.0 + 2.0 \times X$$

What are the meanings of 10.0 and 2.0 ? (10%)

(d) Complete the following ANOVA Table for Company A. (10%)

ANOVA Table

s.o.v	S.S	d.f	MS	F
Regression	13600			
Error				
Total	13660			

(e) What are the degrees of freedoms for F in (d) ? We can use this F value to test a hypothesis. What is that hypothesis ? (5%)

(f) Suppose the standard error for $\hat{\beta_1}$ is 0.04697. Perform t-test for Hypothesis $H_0 : \beta_1 = 0$ v.s. $H_1 : \beta_1 \neq 0$ (5%)

(g) What is the relation between the t-value in (f) and the F-value in (d) ? (5%)

(h) Can we perform test for lack of fit on Company A data set ? (5%)

3. Customers A, B, and C are asked to evaluate the quality of 4 PCs, IBM, DELL, ACER, and ASUS, on a scale ranging from 0 to 10. Subject A's evaluative scores on 4 PCs are 1, 2, 3, 4, and subject B's scores on 4 PCs are 7, 8, 9, 10, respectively. Is the degree of consistency (i.e. correlation coefficient) between A's scores and C's scores the same as that between B's and C's? (Assume all scores are interval scales) (5%)
A) No
B) Yes
C) Have to depend on customer C's scores
D) Have to depend on standardized scores for all customers

4. Let X be a random variable with the normal distribution, N(1000, 0.001). That is, the mean of the corresponding normal distribution is 1000, and the variance is 0.001. Which of the following statements is correct? (5%)
A) X is positive
B) X is negative
C) X can not be 1000.001
D) X can be -1000

5. The sales volume (in thousand dollars) for three different types of in-store promotions is shown below. To test the effect of in-store promotion on sales, please fill the one-way ANOVA table. (10%)

Sales Volume for Three Types of In-Store Promotions
Promotion-Type 1: 3 4 5 6
Promotion-Type 2: 6 7 8 9
Promotion-Type 3: 4 4 8 8

ANOVA Table

Source	df	SS	MS	F
Treatment	(A)	(D)	(G)	(I)
Error	(B)	(E)	(H)	
Total	(C)	(F)		

Note: df-degree of freedom, SS-sum of square, MS-mean sum of square, F-F value

6.   Which of the following statements regarding one-way ANOVA is correct?   (5%)
A) F value can be negative
B) ANOVA is particularly designed to exam the linear relationship between X and Y
C) When population means are all the same, F value from sampling will be very small
D) Population variance can not be different

7.   In a one-way ANOVA testing 100 population means, if the null hypothesis that all of one hundred population means are equal is rejected, then which of the following statements is correct.   (5%)
A) None of the treatment mean is the same as the other one
B) At least 50 treatment means are not the same

C) At least one treatment mean is different from the other ones
D) The variation among treatment means is smaller than that within treatment means

8.   182 students who were majoring in either Math or English were asked a test question, and the researcher recorded whether they answered the question correctly. The test results are given below. To exam whether an association exists between response and major, please calculate chi-square. $\chi^2 = ?$ (5%)

	Correct	Incorrect
Math	11	7
English	73	91

9.   Which of the following statements about chi-square is correct?   (5%)
A) Chi-square is a random variable ranging from negative to positive
B) The chi-square will be significant when there is a significant match between the observed and expected values.
C) The two variables under consideration are sex and blood type. If the two variables are not associated, we would expect that the proportion of women in the sample with a given blood type would be roughly equal to the proportion of men in the sample with the same blood type
D) The two variables under consideration are sex and blood type. If the two variables are not associated, we would expect that the number of women with a given blood type would be roughly equal to the number of men with the same blood type?