1. To compare the content calories of hamburgers and chicken nuggets, following data are collected. Please use Box-Whisker Plot to exhibit the two sets of data and compare. （10%）

	calories
hamburgers	19 31 34 35 39 39 43
chicken nuggets	7 9 15 16 16 18 22 25 27 33 39

2. A production planner wants to see if the operating rates for the 2 factories is the same.　At present, only limited operation data are available. For factory 1, the rates (% of capacity) are 71, 82, 77, 92, 88.　For factory 2, the rates are 85, 82, 94 & 97.　Do the factories have the same rates at the 0.10 significance level? Answer the question based on the following sequence. (15%)

   a. What is the hypothesis you need to test? An average? A Median? Or a Mode? (3%)

   b. What are the assumptions you have to make? (5%)

   c. If the lower and upper critical values are 12 and 28, what is the conclusion you reach?　(7%)

3. As production manager, you want to see if 3 filling machines have different mean filling times. You assign 15 workers with varied experience into 5 groups of 3 based on similarity of their experience, and assigned each group of 3 workers with similar experience to the machines. At the .05 significance level, is there a difference in mean filling times? (15%)

Machine1	Machine2	Machine3
25.40	23.40	20.00
26.31	21.80	22.20
24.10	23.50	19.75
23.74	22.75	20.60
25.10	21.60	20.40

   a. What kind of method do you suggest to analyze the question? (5%)

   b. If we get part of computer output as the following: (5%)

Source of Variation	degrees of freedom	sum of squares	mean squares	F statistics
Among group	?	47.164	?	?
Error	?	8.403	?	

   Please complete the above computer output by filling the numbers in the ? places.

   c. The critical range was computed 1.8506 by Tukey procedure, what is the conclusion would you make? (5%)

4. John is a building contractor with a record of a total of 24 single family homes constructed over a 6-year period. Provide John with a 3-year moving average graph (5%).

Year	1998	1999	2000	2001	2002	2003
units	3	5	3	2	7	5

5. To describe distribution of the characteristics of the data collected, we may use the so-called empirical rule (i.e. approaching a normal distribution) and Bienayme-Chebyshev Rule .　If the average and the standard deviation of the scores of statistical course are 76 and 7.　Determine if the following two data 98 and 58 are outliers ?　(Note: It is defined that a datum is regarded as an outliers if the probability of occurrence is less than 1%)　(10%)

注意：背面有試題

6. Concerning a binomial distribution, it is defined that n is sample size, p is the probability of success, and N is the population size. (10%)

    a. Please use diagrams to illustrate a binomial distribution under the following conditions: if n is finite and p < 0.1, p=0.5, and p>0.9, respectively. (8%)

    b. Under what circumstances, the mean, median, and mode of a binomial distribution will be the same? (2%)

7. While conducting a survey on a trial of a new cosmetic product for a company, you choose a confidence coefficient of 0.9 and an allowable error of 10%. On the conservative side, what sample size is needed? （5%）

8. Following table is the simple price indices based on the year 1980. (5%)

Year	1980	1985	1990	1995	2000
Indices	0.692	0.684	0.719	0.835	0.896

Please change the base year of the simple price indices from 1980 to 1995.

9. Concerning multiple linear regression analysis, some computer output is provided as the following table. Please answer the following questions: (25%)

    a. Explain the difference between R square and Adjusted R square in details? (5%)

    b. What is the conclusion you make based on the ANOVA table? You have to state if the model and regression coefficients significant, the appropriateness of sign and intercept. (5%)

    c. To check if the assumptions for multiple linear regression analysis are met, state as much as you know to perform. (5%)

    d. To detect outliers, we may use influence analysis. Please state the criteria you use. (5%)

    e. To check collinearity among independent variables, suggest an indicator (factor) to detect and explain how to use. (5%)

Regression Analysis

Regression Statistics	
Multiple R	0.986157697
R Square	0.972507004
Adjusted R Square	0.965008914
Standard Error	24.29377937
Observations	15

ANOVA

	df	SS	MS	F	Significance F
Regression	3	229643.1645	76547.72149	129.7006349	7.26403E-09
Residual	11	6492.064875	590.1877159		
Total	14	236135.2293			

	Coefficients	Standard Error	t Stat	P-value	Lower 95%	Upper 95%
Intercept	624.5864209	42.43515952	14.71860664	1.39085E-08	531.1872173	717.9856245
Temp	-5.362603095	0.317128467	-16.90987611	3.20817E-09	-6.060598498	-4.66460769
Insulation	-44.58678859	14.9546884	-2.981458884	0.012486902	-77.50185248	-11.6717247
Insulation Squares	1.866704651	1.123755228	1.661131004	0.124891755	-0.606665181	4.340074482