※ 考生請注意：本試題不可使用計算機。　請於答案卷(卡)作答，於本試題紙上作答者，不予計分。

I.　　True or False (30 points, 3 pts each)

1. The probabilities must be the same for the following two events: (1) you are gonna bring an umbrella tomorrow (2) you are gonna bring an umbrella tomorrow if it rains tomorrow.

2. It is better using nonparametric tests when collected data demonstrates irregular behavior or the currently known distributions cannot satisfactorily capture the characteristics of data.

3. Once the sample size increases, the sample average definitely approaches the population mean even if observations are drawn from the one side of the population.

4. For a non-negative random variable with continuous density function, the expected value can be obtained using distribution function.

5. An estimator is claimed to be relatively more efficient than the other estimator if the skewness of this estimator is lower than of the other estimator.

6. If we sum up several random variables, the variance of the sum will increase as the number of random variables in the sum increases.

7. If one subject is added to Group 2 and the average of Group 2 increases correspondingly, then we can be sure that the value of this subject is higher than the original average of Group 2 without this subject.

8. Thirty consecutive stock price observations of a specific stock were observed and had only two runs for daily ups and downs in stock prices. Then, we can be sure that the observed stock prices are highly correlated using the runs test.

9. If a population is positively skewed, it is possible that an investigator could come up with a sample mean of 45.6 and a sample median of 48.2.

10. The mean and the median will be the same under the symmetric distribution even when the associated distribution has higher kurtosis than the normal distribution.

II.　Choose the BEST answer (45 points, 3 pts each)

Use the following information to answer Questions 1 to 5.

A statistician collected three data points from an experiment which are normally distributed as follows:

25, 8, 40.

1.　What is the arithmetic mean of these data points?

(A) 24.33
(B) 24.23
(C) 24.12
(D) 24.67

2.　What is the geometric mean of these data points?

(A) 35
(B) 30
(C) 25
(D) 20

3.　What is the variance of the population?

(A) 7.45
(B) 8.66
(C) 9.68
(D) 10.12

4.　What is the coefficient of variation of these data points?

(A) 0.24

(B) 0.18

(C) 0.12

(D) 0.06

5.　What is the possible value at 97.75% level of the associated distribution?

(A) 28.10
(B) 30.10
(C) 32.10
(D) 34.10

Use the following information to answer Questions 6 to 10.

An investigator tries to understand the relationship between advertisement expenditures and sales during the good and bad economic conditions. The variance of advertisement expenditures is 25 (hereafter, unit is million) in the bad economic condition while the variance of advertisement expenditures is 16 in the good economic condition. According to available information, the variance of sales in the good economic condition is 9 and the correlation coefficient between the advertisement expenditures and sales is 0.25 in the good economic condition. The regression line between the advertisement expenditures and sales is stable whether in the good or bad economic condition.

6. If we run a regression between advertisement expenditures and sales using sales as the dependent and advertisement expenditures as independent variables, respectively, what is the estimated regression coefficient?

   (A) 0.1875
   (B) 0.1934
   (C) 0.2037
   (D) 0.2175

7. What is the variance of sales in the bad economic condition?

   (A) 6.12
   (B) 7.53
   (C) 8.24
   (D) 9.32

8. What is the correlation between advertisement expenditures and sales in the bad economic condition?

   (A) 0.31
   (B) 0.43
   (C) 0.52
   (D) 0.68

9. What is R-squared for the regression line between advertisement expenditures and sales in the good economic condition?

   (A) 0.0412
   (B) 0.0534
   (C) 0.0625
   (D) 0.0723

10. If the expected advertisement expenditures are 102 and expected sales are 125, respectively, what is expected sales if there are no advertisement expenditures?

   (A) 104.77
   (B) 105.88
   (C) 106.99
   (D) 107.22

Use the following information to answer Questions 11 to 15.

The following is a contingency table which shows the smartphone purchase behavior of people with or without computer software knowledge.

	Computer Software Knowledge		
	With	*Without*	**Total**
iPhone	30	51	81
Android	50	44	94
Total	80	95	175

The basic prediction is that people with computer software knowledge tend to buy Android-based smartphones while people without computer software knowledge tend to buy iPhones. We would like to know whether the computer software knowledge affects the smartphone purchase behavior using two correlation measures: Lambda and Gamma Y.

11. What is the Lambda value for this contingency table using the category of smartphones as the dependent variable?

   (A) 3.62%
   (B) 4.94%
   (C) 5.32%
   (D) 6.12%

12. What is symmetric Lambda value for this contingency table?

   (A) 8.07%
   (B) 9.12%
   (C) 10.23%
   (D) 11.12%

13. What is Gamma Y value of this contingency table?

   (A) 29.56%
   (B) 30.22%
   (C) 31.78%
   (D) 32.45%

14. What is the expected frequency for people with computer software knowledge who purchased iPhones?

   (A) 42.97
   (B) 37.03
   (C) 43.97
   (D) 51.03

15. What is the value of the chi-square test statistic?

   (A) 7.23
   (B) 6.27
   (C) 5.12
   (D) 4.58

III.　Partial Credit Questions and Fill in the Blanks (25points, 5 pts each)

Notes:

(1) Write down your answers along with associated blanks.

(2) Label blanks in alphabetical order.

Questions 1 to 3 belong to the same group of questions.

Stock price ratios $\{s_t\}$ ($s_t = \frac{S_t}{S_0}$, S_t and S_0 are stock prices at time t and 0, respectively) are believed to follow the lognormal distribution. (Note: Answers can be expressed as the nth power of Euler's number, e. For example, the square of e can be expressed as e^2 or $\exp(2)$)

1. An economist proposes the following equation which describes the relationship among the GDP (gross domestic productivity, GDP_t) level, the stock index value (S_t) and other explanatory variables (X_{1t} and X_{2t}).

$$GDP_t = \beta_0 S_t^{\beta_1} X_{1t}^{\beta_2} X_{2t}^{\beta_3} \varepsilon_t$$

(1) If this economist tries to make this nonlinear relationship become linear, what is the transformed linear relationship?＿＿＿(a)＿＿＿

2. If the current stock index value is 144 and $\ln s_1$, natural log of the stock price ratio at time1, has mean of 10% and variance of 20%,

(1) What is the expected value of the stock price ratio s_1? ＿＿(b)＿＿

(2) What is the variance of the stock price ratio s_1? ＿＿(c)＿＿

(3) What is the expected stock price at time 1? ＿＿(d)＿＿

3. What is the expected GDP level at time 1 if $\beta_0 = 1$, $\beta_1 = 0.5$ and expected logged values of other explanatory variables are zero? ＿(e)＿