本份統計考題總分為 100 分。

一、單選題(共 30 題，每題 3 分)：※ 注意：請於試卷上「選擇題作答區」依序作答。

1. Find the sampling distribution of the sample mean $\overline{X}$, $P(\overline{X}=0)$, if samples of size 2 are drawn from the following population:

x	-2	0	2
$p(x)$	0.2	0.6	0.2

   A. 0.44   B. 0.55   C. 0.6   D. 0.66   E. None of the above

2. The following data give the number of pets owned for a population of 4 families. Samples of size 2 will be drawn at random from the population without replacement. Calculate the standard deviation of the sampling distribution of the sample means $\overline{X}$.

Family	A	B	C	D
Number of Pets Owned	2	1	4	3

   A. 2.5   B. 1.118   C. 0.6455   D. 2.236   E. None of the above

3. In performing a hypothesis test where the null hypothesis is that the population mean is 4.8 against the alternative hypothesis that the population mean is not equal to 4.8, a random sample of 25 items is selected. The sample mean is 4.1 and the sample standard deviation is 1.4. It can be assumed that the population is normally distributed. What is the computed "t" value for this problem?
   A. -12.5   B. 12.50   C. -2.5   D. -0.7   E. None of the above

4. A null hypothesis is $P \geq 0.65$. To test this hypothesis, a sample of 400 is taken and alpha is set at 0.05. If the true proportion is $P = 0.60$, what is the probability of a type II error?
   A. 0.17   B. 0.45   C. 0.95   D. 0.67   E. None of the above

5. Suppose a researcher wants to conduct a study to estimate the population mean. He plans to use a 95% level of confidence to estimate the mean, and the population standard deviation is approximately 34. The researcher wants the error to be no more than 4. The sample size should be at least _____.
   A. 196   B. 278   C. 1110   D. 1421   E. None of the above

6. To apply the Friedman test to determine whether the locations of two or more populations are the same, the samples must be:
   A. from multinomial experiment   B. from normal populations
   C. independent   D. larger than 20
   E. None of the above

7. A researcher is interested in testing the difference in two population means. Previous studies indicate no difference in population standard deviation. He randomly selected 49 samples from each population and reported the following: $\overline{x}_1 = 7$ days, $\overline{x}_2 = 8$ days, $s_1 = 1$, and $s_2 = 2$. Assuming a two-tail test for and $\alpha = .01$, the calculated z value is _____.
   A. -16.33   B. -3.13   C. -9.80   D. -4.04   E. None of the above

8. A researcher is interested in estimating the difference in two population proportions. A sample of 400 from each population results in sample proportions of .61 and .64. A 90% confidence interval for the difference in the population proportions is _____.
   A. -0.10 to 0.04   B. -0.09 to 0.03
   C. -0.11 to 0.05   D. -0.07 to 0.01   E. None of the above

9. A variable contains five categories. It is expected that data are uniformly distributed across these five categories. To test this, a sample of observed data is gathered on this variable resulting in frequencies of 27, 30, 29, 21, 24. Using $\alpha = .01$, the calculated value of chi-square is _____.
   A. 2.09   B. 9.82   C. 1.62   D. 17.81   E. None of the above

10. Use the following set of observed frequencies test the independence of the two variables. Variable one has values of "A" and "B"; variable two has values of "C", "D", and "E". Using $\alpha = 0.05$, the calculated test value is _____.

	C	D	E
A	12	10	8
B	20	24	26

    A. 0   B. 0.69   C. 1.54   D. 21.28   E. None of the above

接背面

11. Charles Clayton monitors the daily performance of his investment portfolio by recording a "+" or a "-" to indicate whether the portfolio's value increased or decreased from the previous day. His record for the last eighteen business days is "- + + - - - + - - + + + - + + + + -". The number of runs in this sample is _____.

    A. four   B. five   C. seven   D. eight   E. None of the above

12. Most "Before and After" types of experiments should be analyzed using _____.
    A. Chi-square goodness of fit test   B. Kruskal-Wallis test
    C. Mann-Whitney U test   D. Sign test   E. None of the above

13. A person wishes to compare the means of three populations. The data is ordinal. Which of the following should be used?
    A. One-way analysis of variance   B. Chi-square test
    C. Wilcoxon   D. Mann-Whitney   E. None of the above

14. A Mann-Whitney U test was performed. U was calculated to be 38.78 based on sample sizes of 24 and 20. Calculate the z value for this.
    A. -4.74   B. 0.1333   C. 240   D. 42.43   E. None of the above

15. The Wilcoxon test was used on 18 pairs of data. The total of the ranks (T) were computed to be 111 (for + ranks) and 60 (for - ranks). Calculate the z value that would be used with this.
    A. -1.11   B. -0.05   C. -0.07   D. 0.033   E. None of the above

16. Which one of the following statements is true?
    A. In a histogram, the highest bar represents the class with the highest cumulative frequency.
    B. If data are grouped into intervals and the number of items in each group is listed, this could be called a stem and leaf plot.
    C. One rule that must always be followed in constructing frequency distributions is that the width of each class is equal to the range.
    D. A histogram is best to show the percentage of a total budget that is spent on each category of items.
    E. None of the above

17. Consider the following 3 numbers: 10, 20, 30. In calculating measures of central tendency and variability for this set of data, which of the following is true?
    A. You can calculate the mean only if you know if this is a sample or a population.
    B. You can calculate the median only if you know if this is a sample or a population.
    C. You can calculate the mode only if you know if this is a sample or a population.
    D. You can calculate the variance only if you know if this is a sample or a population.
    E. None of the above.

18. Liz Chapa manages a portfolio of 250 common stocks. Her staff compiled the following performance statistics for two new stocks and the coefficient of correlation between these two stocks is 0.35. If Liz decides to invest equally to both stocks, what is the standard deviation of rate of return?

	Rate of Return	
Stock	Mean	Standard Deviation
Salas Products, Inc.	15%	5%

    A. In the range of (0.40, 0.45]   B. In the range of (0.45, 0.50]
    C. In the range of (0.50, 0.55]   D. In the range of (0.55, 0.60]   E. None of the above

19. In a set of 15 aluminum castings, two castings are defective (D), and the remaining thirteen are good (G). A quality control inspector randomly selects three of the fifteen castings without replacement, and classifies each as defective (D) or good (G). How many elementary events are in the sample space for this experiment?
    A. 3,375   B. 2,730   C. 210   D. 15   E. None of the above

20. Abel Alonzo, Director of Human Resources, is exploring employee absenteeism at the Plano Power Plant. Ten percent of all plant employees work in the finishing department; 20% of all plant employees are absent excessively; and 7% of all plant employees work in the finishing department and are absent excessively. A plant employee is selected randomly; F is the event "works in the finishing department;" and A is the event "is absent excessively." What is the value of P(F|A)?
    A. 0.35   B. 0.70   C. 0.13   D. 0.37   E. None of the above

接次頁

21. Max Sandlin is exploring the characteristics of stock market investors. He found that sixty percent of all investors have a net worth exceeding $1,000,000; 20% of all investors use an online brokerage; and 10% of all investors a have net worth exceeding $1,000,000 and use an online brokerage. An investor is selected randomly, and E is the event "networth exceeds $1,000,000," and O is the event "uses an online brokerage." Which of the following is true?
   A. E and O are dependent.
   B. E and O are collectively exhaustive.
   C. E and O are independent.
   D. E and O are mutually exclusive.
   E. None of the above

22. Dorothy Little purchased a mailing list of 2,000 names and addresses for her mail order business, but after scanning the list she doubts the authenticity of the list. She randomly selects five names from the list for validation. If 40% of the names on the list are not authentic, and x is the number of non-authentic names in her sample, What is the value of $P(x<2)$?
   A. 0.2333   B. 0.4370   C. 0.9785   D. 0.9853   E. None of the above

23. Meagan Dubean manages a portfolio of 200 common stocks. Her staff classified the portfolio stocks by "industry sector" and "investment objective." If a stock is selected randomly from Meagan's portfolio, What is the value of $P(\text{Airlines}|\text{Income})$?

Investment Objective	Industry Sector			Total
	Electronics	Airlines	Healthcare	
Growth	100	10	40	150
Income	20	20	10	50
Total	120	30	50	200

   A. 0.10   B. 0.25   C. 0.40   D. 0.67   E. None of the above

24. On Saturdays, cars arrive at Sami Schmitt's Scrub and Shine Car Wash at the rate of 6 cars per fifteen minute interval. What is the value of the probability that less than 10 minutes will elapse between car arrivals?
   A. 0.0183   B. 0.1535   C. 0.8465   D. 0.9817   E. None of the above

25. The results of SSR=700 and SST=1000 are obtained from a multiple regression analysis with $n = 35$ and four independent variables. What is the value of the adjusted r^2?
   A. 0.56   B. 0.70   C. 0.76   D. 0.80   E. None of the above

26. Yvonne Yang, VP of Finance at Discrete Components, Inc. (DCI), wants a regression model which predicts the average collection period on credit sales. Her data set includes two qualitative variables: sales discount rates (0%, 2%, 4%, and 6%), and total assets of credit customers (small, medium, and large). What is the number of dummy variables needed in Yvonne's regression model?
   A. 3   B. 4   C. 5   D. 6   E. None of the above

27. Which of the following methods may help to overcome the Autocorrelation problem in a regression forecasting model?
   A. Increasing the sample size
   B. Utilizing additional independent variables
   C. Increasing the level of significance for the F test
   D. Using the F test, instead of the t test
   E. None of the above.

28. Analysis of data for an autoregressive forecasting model produced the following tables. Which of the following statements is true?

	Coefficients	Standard Error	t Statistic	P-value
Intercept	11.27019	6.98597	1.61326	0.11368
y_{t-1}	0.789271	0.151681	5.203502	4.66E-06
y_{t-2}	-0.02648	0.150707	-0.17569	0.861323

	df	SS	MS	F	P-value
Regression	2	47962.61	23981.31	30.37912	5.96E-09
Residual	43	33944.24	789.4009		
Total	45	81906.85			

接背面

A. The first predictor, y_{t-1}, is not significant at the 5% level.

B. The second predictor, y_{t-2}, is significant at the 5% level.

C. All predictor variables are significant at the 5% level.

D. None of the predictor variables are significant at the 5% level.

E. None of the above.

29. Abby Kratz, a market specialist at the market research firm of Saez, Sikes, and Spitz, is analyzing household budget data collected by her firm. Abby's dependent variable is monthly household expenditures on groceries (in $'s), and her independent variables are annual household income (in $1,000's) and household neighborhood (0 = suburban, 1 = rural). Regression analysis of the data yielded the following table. For two households, one suburban and one rural, what will Abby's model predict ?

	Coefficients	Standard Error	t Statistic	p-value
Intercept	19.68247	10.01176	1.965934	0.077667
x_1 (income)	1.735272	0.174564	9.940612	1.68E-06
x_2 (neighborhood)	49.12456	7.655776	6.416667	7.67E-05

A. The rural household's monthly expenditures for groceries will be $49 more.

B. The suburban household's monthly expenditures for groceries will be $8 more.

C. The suburban household's monthly expenditures for groceries will be $49 more.

D. Equal monthly expenditures for groceries

E. None of the above.

30. The ratios of "actuals to moving averages" (seasonal indexes) for a time series are presented in the following table as percentages. What is the value of the seasonal index for Q_2?

	1999	2000	2001	2002	2003
Q_1		112.22	110.78	111.22	111.87
Q_2		100.65	108.68	103.78	101.95
Q_3	97.76	99.08	97.68	97.61	
Q_4	86.61	95.00	94.64	92.92	

A. 101.95   B. 103.765   C. 104.746   D. 102.865   E. None of the above

二、簡答題(共 2 題，每題 5 分)：請按照題號順序作答

1. Data from a randomized block design are shown in the following table.

	Treatment Levels			
	1	2	3	4
Block 1	5	2	3	8
Block 2	7	4	3	5
Block 3	4	3	1	9

(1) The Blocks Sum of Squares (SSA) is _____.

(2) The Treatment Sum of Squares (SSB) is _____.

(3) The Error Sum of Squares (SSE) is _____.

2. The Gross Domestic Product (GDP) is the total U.S. output of goods and services valued at market prices. The quarterly GDP values (in billions of dollars) for the period 2002 to 2004 are given in the following table.

Year	2002				2003				2004			
Quarter	1	2	3	4	1	2	3	4	1	2	3	4
GDP	9093	9161	9297	9522	9669	9858	9938	10028	10142	10203	10225	10480

(1) Please use 4-period centered moving average to find the seasonal indexes.

(2) If the simple linear regression model by using the de-seasonalized GDP data is GDP = 8992.016 + 124.524 t where $t = 1$ for the first quarter, 2002, use this model and the seasonal index found in (1) to forecast the 2005 quarterly GDP values.

試題隨卷繳回