

※ 考生請注意：本試題 可 不可 使用計算機

1. (10%) The following observations are given for two variables x and y .

| | | | | | | | | |
|-----|---|----|----|----|----|----|----|----|
| x | 3 | 13 | 4 | 7 | 12 | 20 | 19 | 10 |
| y | 6 | 9 | 19 | 21 | 23 | 31 | 11 | 8 |

- (a) (5%) Compute and interpret the sample covariance for the above data.
- (b) (5%) Compute and interpret the sample correlation coefficient.
2. (10%) A diagnostic test for a certain disease is not 100% reliable in that, if a person has the disease, the test will detect it with probability 0.95. And, if a person does not have the disease, the test will report that he or she does not have it with probability 0.98. Only 1% of the population has the disease in question. If a person is chosen at random from the population and the diagnostic test indicates that he or she has the disease, what is the conditional probability that he or she does, in fact, have the disease?
3. (30%) Answer the following questions:
- (a) (4%) What is pooled variance? Under what condition shall the pooled variance be used?
- (b) (4%) What is the main purpose of matched sample design (compared to independent sample design)?
- (c) (8%) The sampling error is the absolute value of the difference between the value of an unbiased point estimator and the value of the population parameter it estimates. But, because the value of the population parameter is unknown, how will the sampling error be measured?
- (d) (4%) Is the statement "The joint probability of two events is the probability of the union of these two events." correct? Please explain.
- (e) (5%) What is the purpose of the power curve?
- (f) (5%) What are the situations in which hypothesis testing procedures are commonly employed?

(背面仍有題目,請繼續作答)

系所組別： 資訊管理研究所甲組

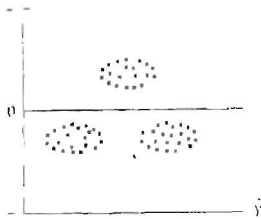
考試科目： 統計學

考試日期： 0220，節次： 3

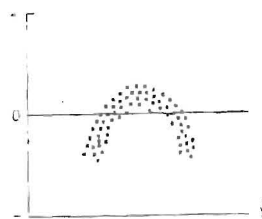
※ 考生請注意：本試題 可 不可 使用計算機

4. (12%) Please explain the following terms in detail and state the relationship for each pair:
- Standardized Residual vs. Studentized Deleted Residual
 - Outlier vs. Influential Observation
 - Autocorrelation vs. Multicollinearity
 - Comparisonwise Type I Error Rate vs. Experimentwise Type I Error Rate
5. (11%) Consider a completely randomized design involving three treatment A, B, and C. Instead of using ANOVA to compare the three means, multiple regression approach can be used to deal with this problem. Please write a multiple regression equation and define all variables to complete the analysis of hypothesis testing.
6. (8%) Please state the rationales for performing a (i) Parametric analysis, and (ii) Nonparametric analysis, i.e., under what situations, for what reasons, and with what objectives, will we use them, and also state the pros and cons for both of the analyses.
7. (10%) Please describe how to use χ^2 test to perform both tests of (i) goodness of fits, and (ii) independence. You may use illustrative examples to clarify your descriptions.
8. (9%) Regression analysis may generate the following residual plots. Please state the possible causes and remedy strategies for each case:

(i).



(ii).



(iii).

