

所別：資訊管理學系碩士班 甲組 科目：統計學
乙組

注意：請依題號順序回答，跳答每題扣5分。

I. Multiple Choice Questions (30%, 3% for each question. There is only ONE correct answer for each question.)

- () 1. Whenever $p = 0.15$, the binomial distribution will
(A) always be symmetric.
(B) Be symmetric only if n is large.
(C) Be right-skewed.
(D) Be left-skewed.
- () 2. A financial analysis is presented with information on the past records of 60 start-up companies and told that in fact only 3 of them have managed to become highly successful. He selected 3 companies from this group as the candidates for success. To analyze his ability to spot the companies that will eventually become highly successful, he will use what type of probability distribution?
(A) binomial distribution.
(B) Poisson distribution.
(C) normal distribution.
(D) hypergeometric distribution.
- () 3. Concerning random sampling, choose the incorrect description from the followings:
(A) A sample unit in a population may have unequal probability to be selected.
(B) Sampling error may be reduced by choosing more sampling units.
(C) It is an objective sampling technique.
(D) The estimation of population parameters from statistics can be unbiased.
- () 4. Concerning median and mode, choose the incorrect description from the followings:
(A) The median of the four data, (1,2,3,4), is 2.5.
(B) The median of the five data, (2,1,2,1,1), is 1.
(C) The mode of the four data, (1,2,3,4), is 4.
(D) The mode of the four data, (1,2,1,2) are 1 and 2.
- () 5. To compare the failure rate (i.e. the proportion of the score below 60 points) of statistics course between classes A and B, following data are collected: the average scores are 75 and 72 points, the standard deviations are 7.5 and 4 points for the classes A and B respectively. Choose the correct description from the followings:
(A) The failure rate of the class A is higher.
(B) The failure rate of the class B is higher.
(C) The two classes have the same failure rate.
(D) There is insufficient data to justify.
- () 6. The defective rate for a product is 0.1 consistently. We choose 4 products randomly and independently for inspection. Choose the correct description on population parameters or the shape of the probability of the number of defects:
(A) The shape is skewed to the left.
(B) The shape is symmetrical.
(C) The mode is 0.
(D) The median is greater than the average.
- () 7. For the length of a product, the average and the standard deviation are specified to be 100 cm and 4 cm, respectively. We choose 4 products for inspection every 30 minutes. The tolerance limits are set 4 cm away from the mean of the average. According to the empirical rule for a normal distribution, what is the probability between the tolerance limits?
(A) 0.9974
(B) 0.9772
(C) 0.9544
(D) 0.6826
- () 8. Because there exists strong competition among 4 candidates for presidency, one of the four candidates asks a survey company to implement a survey of support with a marginal error of 2.5%. Based on the previous experience, the candidate possesses the rate of support between 20% and 30%. At the level of confidence of 0.95 ($Z_{0.975} = 1.96$, i.e. $\Pr(Z \leq 1.96) = 0.975$), what is the conservative sample size to be chosen?
(A) 984
(B) 1291
(C) 1537
(D) 2017

注意：背面有試題

所別：資訊管理學系碩士班 甲組 科目：統計學
乙組

- () 9. To test the difference between the means of two samples, choose the incorrect description from the followings:
 (A) Random sampling has to be used for both independent and related sampling.
 (B) If the two populations are normal distributions with two unknown variances and are sampled independently, the sampling distribution is normal.
 (C) In case one of the population distribution is distribution-free with a small sample size, e.g., less than 10, for a related sampling, a signed test may be used.
 (D) If the two populations are normal distributions with two unknown variances and are relation-sampled, the sampling distribution is a student distribution.
- () 10. To test population variances with normal distributions for one-sample or two-sample testing, choose the incorrect description from the followings:
 (A) For a one-sample test, a chi-square distribution can be used.
 (B) For a two-sample test, an F distribution can be used.
 (C) A chi-square distribution has the degree of freedom equal to the sample size minus one.
 (D) An F distribution has the degree of freedom equal to the summation of the two sample sizes minus one.

II. Short essay questions (70%)

1. A videocassette recorder (VCR) repair service wanted to study the effect of VCR brand and service center on the repair time measured in minutes. Three VCR brands (A,B,C) were specifically selected for analysis. Three service centers (1,2,3) were also specifically selected. Each service center was assigned to perform a particular repair on two VCRs of each brand. The results are presented in the following tables. (22%)

Service Center	VCR brand		
	A	B	C
1	52	48	59
2	57	39	67
3	51	61	58
	43	52	64
	37	44	65
	46	50	69

SOURCE	degrees of freedom	sum of squares	mean of squares	F	p
Center	2	27.4	13.7	(a)	0.6173
Brand	2	945.8	472.9	(b)	0.0008
Interaction	4	361.2	90.3	(c)	0.0607
Error	9	242.0			
Total	17	1576.4			

- (A) What are the values of (a), (b), and (c)? (1% each)
 At the 0.05 level of significance:
 (B) Is there an interaction between service center and VCR brand? What conclusion can you make regarding on the interaction effect (5%)
 (C) Plot a graph of mean service time for each service center for each VCR brand. Please mark the cell means in your plot. (10%)
 (D) What conclusions can you reach concerning mean service time? (4%)
2. What is the difference between "independent samples" and "related samples"? Please give an example for each. (5%)
3. As marketing manager, you want to see if 3 newly-designed computer keyboards have different mean degree of preference to users. You assign 15 similarly trained & experienced users, 5 per keyboard, to the computer keyboards. At the 0.05 significance level, is there a difference in mean degree of preference? Likert five-scale measure is used to survey the degree of preference, the result are shown in the following table. (11%)

Keyboard1	Keyboard2	Keyboard3
4.540	4.340	4.000
4.631	4.180	4.220
4.410	4.350	3.975
4.374	4.275	4.060
4.510	4.160	4.040

Assuming the three population distributions are normal, please answer the questions (A) and (B).

- (A) Given that the global average is 4.271, and the averages for each types of keyboards are 4.493, 4.261, and 4.059 respectively. The SSA(Sum of squares among groups) is 0.4716, SSW(Sum of squares within groups) is 0.1105. The critical value of F distribution is 3.89. Please construct a summary table and make your conclusion. (8%)
 (B) Given that the critical range is 0.1618, show the result of the follow-up study if there exists. (3%)
4. What are the differences between regression analysis and correlation analysis? How to determine whether regression analysis or correlation analysis is more proper for one research analysis? (10%)
5. What is multicollinearity? What impacts will multicollinearity cause on regression models? How to avoid such impacts? (12%)
6. List the steps for stepwise regressions. What are the major differences between stepwise regression methods and stagewise regression methods? (10%)