

※ 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。

1. A study was conducted to assess the potential association between a certain chemical in drinking water and death from esophageal cancer. Data were obtained from two geographic areas; the water in area A contained large amounts of the chemical under investigation, while the water in area B contained only trace amounts of the chemical. Data from the study, categorized by age, are presented below:

Age	Area A			Area B			Standard Population		
	Persons	Deaths	Rate	Persons	Deaths	Rate	Persons	Deaths	Rate
40-49	2000	4	0.002	4000	2	0.001	3000,000	3,000	0.001
50-59	3000	12	0.004	4000	6	0.015	6000,000	12,000	0.002
60-69	5000	50	0.01	2000	5	0.025	11,000,000	33,000	0.003
Totals	10,000	66		10,000	13		20,000,000	48,000	

- (1) What is the expected number of deaths from esophageal cancer among subjects aged 60-69 in area A? (5%)
- (2) What is the SMR (standardized mortality ratio) for area B? (5%)

2. Clinical trials:

- (1) What is the purpose of randomization in a clinical trial? [Choose one best answer.] (5%)
- To make the diseased and non-diseased as similar as possible with respect to all variables except the exposure of interest
 - To reduce the number of subjects who are lost to follow-up
 - To isolate the effect of the exposure of interest
 - To encourage compliance with the assigned treatment regimen
- (2) A clinical trial was conducted to compare a new chemotherapy protocol to the standard therapy for leukemia. At the end of the follow-up period, the investigators performed two separate analyses. For the first analysis, they followed the “intention to treat” rule. For the second analysis, they used the “per-protocol analysis.” The results of the two analyses differed substantially. The most likely explanation for the discrepancy is: [Choose one best answer] (5%)
- The randomization was unsuccessful
 - The new therapy was not effective
 - There was a significant degree of recall bias
 - There was a problem with patient compliance

3. A case-control study was conducted to assess whether coffee consumption (high vs. low) is associated with peptic ulcer disease. The results, stratified on smoking status, are summarized below:

	Smokers			Non-smokers	
	High	Low		High	Low
Ulcer	40	60	Ulcer	150	50
No Ulcer	50	150	No Ulcer	60	40

(1) Calculate the crude odds ratio, stratum-specific odds ratios, and the odds ratio adjusted for smoking. Show your calculations. Based upon these calculations, is smoking status a confounder in these data? Justify your answer. (7%)

(2) 請說明一個 confounder 通常需要滿足哪些條件？(7%)

4. What is the purpose of the control group in a case-control study? [Choose one best answer.] (3%)

- A. To provide an estimate of the background risk or rate of disease.
- B. To provide an estimate of the exposure frequency among the population that produced the cases.
- C. To provide an estimate of the magnitude of the placebo effect.
- D. To provide an estimate of the expected number of cases among the unexposed.

5. Which of the following would **NOT** be a useful technique for minimizing the likelihood of selection bias in an epidemiologic study? [Choose one best answer] (3%)

- A. Strive for high response and follow-up rates (cohort study)
- B. Conduct population-based rather than hospital-based case-control studies
- C. Increase the sample size
- D. Study incident rather than prevalent cases of disease

6. An investigator was interested in studying the relationship between a new drug and gastrointestinal bleeding. He assembled a cohort without past history of gastrointestinal bleeding; half of them used this new drug in the recent one month and the others did not. He followed them over 6 months for evidence of gastrointestinal bleeding.

Suppose that the following table represents the TARGET population (i.e. no misclassification) for the study of the new drug and gastrointestinal bleeding:

	Use of New Drug	
	Yes	No
GI bleed	400	300
no GI bleed	200	300

Assume that when the study was carried out, GI bleeds were classified with sensitivity = 0.9 and specificity = 0.75 for subjects who were using the drug. For subjects not using the drug, GI bleeds were classified with sensitivity = 0.6 and specificity = 0.75.

(1) Use this information to fill in the table below: (5%) 請將下表畫於答案卷中作答

	Use of New Drug	
	Yes	No
GI bleed		
no GI bleed		

(2) Calculate the appropriate measure of association for the OBSERVED (misclassified) data. Indicate whether there is bias, and if so, in what direction. (5%)

7.

(1) 請下列樣本的數據計算中位數 (median), 平均數 (mean), 四分位間距 (inter-quartile range), 標準差 (standard deviation) 與變異係數 (coefficient of variation)。 (請列出公式與計算過程, 15%)

5, 7, 8, 9, 9, 12, 12, 14, 14

(2) 請計算此平均值的 95% 信賴區間 (5%)。

($Z_{0.95} = 1.64$; $Z_{0.975} = 1.96$; $t_{0.95, 8} = 1.86$; $t_{0.975, 8} = 2.31$; $t_{0.95, 9} = 1.83$; $t_{0.975, 9} = 2.26$)

(3) 假設母群體平均值為 9，標準差為 2，請以統計方法檢定上述樣本平均是否顯著高於母群體平均值？ (10%)

8.

- (1) 某地區前一次總統大選有 50% 的投民主黨候選人，有 45% 投共和黨候選人，有 5% 投其他政黨參選人。已知前一次投民主黨者這次投民主黨候選人的機會為 90%，前一次投共和黨者這次投民主黨候選人的機會為 10%，而前一次投其他政黨參選人這次投民主黨候選人的機會為 50%。某人這次投票投給民主黨候選人，請問他平時是共和黨支持者的機會為何？(5%)
- (2) 若選前隨機抽樣 100 位選民，其中 60 位表示這次會投民主黨，有 35 位表示這次會投共和黨，另外 5 位會投其他政黨參選人。請問民主黨這次得票率 95% 的信賴區間為何？($Z_{0.95} = 1.64$; $Z_{0.975} = 1.96$)(5%)

9. 某學校四個班級學生的生物統計期末考成績之平均與標準差如下：

	A 班	B 班	C 班	D 班
人數	50	50	50	50
平均成績	80	84	86	90
標準差	5	5	6	6

- (1) 請問這四個班級之生統成績是否有差異？($\alpha = 0.05$, $F_{0.95;3,196} = 2.66$, $F_{0.95;4,196} = 2.42$, $F_{0.95;3,197} = 2.66$)(6%)
- (2) 後續欲進行兩兩多重比較，若 type I error 要維持在 0.05，以 Bonferroni 法調整後多重檢定之顯著性水準應為多少？(4%)