1. Consider the following McNemar's table from "Agent Orange" study. This is a pair-matched case-control study, where the cases are babies born with genetic anomalies and controls are babies born without such anomalies. The matching variables are hospital, time period of birth, mother's age, and health insurance status. The exposure factor is status of father (Vietnam veteran = 1 or nonveteran = 0). (10%)

		Case	
		Exposed	Not exposed
Control	Exposed	2	121
	Not exposed	125	8254

a. For the above data, carry out McNemar's test for the significance of exposure and compute the estimate odds ratio. What are your conclusions?

The following printout results from using conditional maximum likelihood estimation of an appropriate logistic model for analyzing the Agent Orange data:

					95% CI for OR	
Variable	β	s_β	P-value	OR	L	U
Exposure	0.032	0.128	0.901	1.033	0.804	1.326

b. How does the odds ratio obtained from the printout given above compared with the odds ratio computed using McNemar's formula?

2. Please answer the following questions using three risk factors (behavior type, systolic blood pressure, and total cholesterol) related to survival of 35 individuals with high cholesterol. Multi-variable model: analysis of the impact of behavior type (AB; A = 1 = aggressive individuals with perceived stressful lives and B = 0 = more relaxed and less aggressive individuals), systolic blood pressure (SYSBP) in mmHg, and total cholesterol (CHOL) in mg/dL on survival time (in days) using a proportional hazards model. (10%)

Analysis of Maximum Likelihood Estimates					
Variable	DF	Parameter Estimate	Standard Error	Chi-Square	Pr>ChiSq
AB	1	0.91405	0.84991	1.1566	0.2822
SYSBP	1	-0.01018	0.02986	0.1162	0.7332
CHOL	1	0.00732	0.00491	2.2198	0.1362

a. Calculate the approximate 95% confidence interval for the regression coefficient (**behavior type A versus B**).
b. Estimate the hazard ratio (**1 mg/dL increase in total cholesterol**).
c. Calculate the approximate 95% confidence interval for the hazard ratio (**1 mg/dL increase in total cholesterol**).

3. Why is a high response rate important in a cross-sectional survey? How may it be improved? How may the consequences of non-response be assessed? (10%)

4. You are planning a case-control study of breast cancer in relation to post-menopausal use of estrogen. Discuss the selection of cases and controls. (10%)

5. What is meant by confounding in epidemiology and what are the principal methods used to try to control its effects? (10%)

6. 在某項研究中，發現孩童血中鉛濃度(X)與智商表現(Y)有線性關係，得知結果如下：(10%)
X_{mean}=10*ug*/dL, Y_{mean}=100, $\sum(X_i-X_{mean})^2$=400, $\sum(Y_i-Y_{mean})^2$=100, $\sum(X_i-X_{mean})(Y_i-Y_{mean})$=-300,
a. 試計算回歸係數 b，及相關係數 r。
b. r^2 為判定係數(coefficient of determination)，試述其意義。
c. 計算一個血中鉛 15 *ug*/dL 的孩童的期望智商值。

見背面

7. 某一大腸癌篩檢結果與臨床確診結果比較如下表所示，請計算敏感度(sensitivity)、特定度(specificity)、陽性預測值(positive predictive value)、陰性預測值(negative predictive value)。並請說明在篩檢工具的選擇上，這些參數的意義與判定基準的原則為何？(10%)

		大腸癌診斷		
		確診	未確診	合計
篩檢結果	陽性	150	850	1000
	陰性	50	6150	6200
		200	7000	7200

8. From the viewpoint of epidemiology, how do you interpret the results of logistic multiple regression analysis for the occurrence of myocardial infarction (MI) as presented in the following table. Also please explain the relationship among coefficients in regression model, odds ratio and 95% confidence interval (CI) of odds ratio. (10%)

Variables	Coefficients in Regression Model	Odds Ratio	95% CI of Odds Ratio
Constant	0.12		
Gender (Male = 0, Female = 1)	0.13	1.35	[0.93, 1.95]
Smoking (No = 0, Yes = 1)	0.37	2.34	[1.32, 4.17]
Work History (Office = 0, Smelter = 1)	0.67	4.68	[2.24, 9.77]

9. 簡答題：

a. 請簡述標準常態分佈的特性，並概要說明第一類錯誤(α-error)、第二類錯誤（β-error）及檢力(power)。(8%)
b. Healthy worker effect (4%)
c. Misclassification (4%)
d. Proportion/Rate/Ratio (4%)

試題隨卷繳回