系所組別： 統計學系

考試科目： 數理統計

※ 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. **(40 points)** Given x_i, Y_i, $i = 1,2,\ldots,n$, independently follows $N(\beta_0+\beta_1 x_i, \sigma^2)$, β_0 and β_1 unknown, but σ^2 known. Also, not all the x_i's are equal, $\sum_{i=1}^{n} x_i = 0$, $n\geq 2$.

   (1) (8 points) Find the joint sufficient statistics for (β_0, β_1).

   (2) (8 points) Find the maximum likelihood estimators (MLEs) of (β_0, β_1), say$(\hat{\beta}_0, \hat{\beta}_1)$.

   (3) (8 points) What is the joint distribution of $(\hat{\beta}_0, \hat{\beta}_1)$?

   (4) If, for given x_i, Y_i, $i=1,2,\ldots,n$, independently follows Bernoulli $(\mu_i, \mu_i(1-\mu_i))$, where

$$\mu_i = \frac{e^{\beta_0+\beta_1 x_i}}{1+e^{\beta_0+\beta_1 x_i}},$$

   (a) (8 points) Show that $(\sum_{i=1}^{n} Y_i, \sum_{i=1}^{n} x_i Y_i)$ is sufficient for (β_0, β_1).

   (b) (8 points) Find the likelihood equations. Show or argue that the MLEs of (β_0, β_1) are functions of the sufficient statistics $(\sum_{i=1}^{n} Y_i, \sum_{i=1}^{n} x_i Y_i)$.

2. **(20 points)** Let X and Y be two independent random variables coming from Poisson distributions with parameters $E(X) = \lambda_1$, $E(Y) = \lambda_2$, respectively.

   (1) (10 points) Find the conditional distribution of $X|X+Y=m$.

   (2) (10 points) Suppose we have two random samples $X_1, X_2, \ldots, X_n$ and $Y_1, Y_2, \ldots, Y_n$ coming from Poisson(λ_1) and Poisson(λ_2), respectively. We are interested in testing H_0: $\lambda_1/\lambda_2 \leq 1$ vs H_1: $\lambda_1/\lambda_2 > 1$. Consider $X_1|X_1+Y_1=m_1$, $X_2|X_2+Y_2=m_2$, ..., $X_n|X_n+Y_n=m_n$. Based on $X_i|X_i+Y_i=m_i$, $i=1,2,\ldots,n$, construct an (conditional) uniformly most power test with level α.

3. **(24 points)**

   (1) (6 points) Let X be a continuous random variable with cumulative distribution function F_x. Show that $U = F_x(X)$ follows an uniform distribution, $U(0,1)$.

   (2) (6 points) Show that $-2\ln U$ follows a chi-square distribution with 2 degrees of freedom.

   (3) (6 points) Let $T(X)$ be a test statistic (with some probability density function) in testing H_0: $\theta = \theta_0$ vs H_1: $\theta > \theta_0$. The critical region for the test is determined by

$$P(T(X)) \geq c) = \alpha,$$

（背面仍有題目，請繼續作答）

※ 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。

where α is the level of significance, c is some constant. Show that

$$p\text{-}value = P(T(X) \geq t(x))$$

follows an uniform distribution, where $t(x)$ is the observed value of $T(X)$.

(4) (6 points) Let $P_1, P_2, \ldots, P_n$ be the p-values obtained by n medical centers in testing H_0: $\theta = \theta_0$ vs H_1: $\theta > \theta_0$. It is known that all the medical centers follow the same protocol in collecting the data. How to merge the above n p-values into one quantity so that the test H_0: $\theta = \theta_0$ vs H_1: $\theta > \theta_0$ can be performed more effectively?

4. **(16 points)** 內政部警政署資料顯示，近幾年國內第三級毒品的施用人數急遽增加，且有年輕化的傾向。為了防止毒品流入校園，採自願性的觀點推廣校園尿液檢測是個可行的措施。假設全國國、高中生的吸毒人口比例為 p，有吸毒且被正確檢測為陽性反應的機率（敏感度，sensitivity）為 s，沒有吸毒且被正確檢測為陰性反應的機率（明確度，specificity）為 q。

(1) (8 points) 求一次檢測下呈現陽性反應，但事實上該生未吸毒的機率，即偽陽率（false positive rate）。在 $p = 0.05, s = 0.9, q = 0.9$ 下，一次檢測的偽陽率為何？

(2) (8 points) 在(a)中所得的偽陽率可能很高，降低偽陽率的一個方式為重複檢驗。求檢驗 2 次皆為陽性，事實上並未吸毒的偽陽率。求在 $p = 0.05, s = 0.9, q = 0.9$ 下的偽陽率。