MATH50011 Statistical Modelling 1
Midterm Examination

1. A random sample of n individuals is selected from a Bernoulli(π) distribution, where π is the proportion of individuals with an attribute A. The i-th individual spins a spinner that points to A with *known* probability θ and not to A with probability $1-\theta$, where $0 < \theta < 1$ and $\theta \neq 0.5$. The i-th individual responds "yes" if the spinner points to the group they belong to (either A or *not* A). Otherwise, the i-th individual responds "no."

   For $i = 1, 2, \dots, n$, let the random variable X_i take the value 1 if the i-th individual responds "yes," and let X_i take the value 0 if the i-th individual responds "no."

   (a) Show that $P(X_i = 1) = \pi(2\theta - 1) + (1 - \theta)$. (2 marks)

   (b) Derive an explicit expression for $\hat{\pi}$, the maximum likelihood estimator (MLE) of π. Clearly write the likelihood function that you use. (7 marks)

   (c) Develop explicit expressions for *bias*$(\hat{\pi})$ and *Var*$(\hat{\pi})$, the bias and variance of the MLE $\hat{\pi}$. (4 marks)

   (d) Does there exist an unbiased estimator of π with lower mean square error than the MLE $\hat{\pi}$? Justify your answer. (2 marks)

2. In a medical study, a random sample of n pairs of twins is enrolled. For each pair, one twin receives treatment X and the other receives treatment Y. Let the i-th pair of twins be denoted by (X_i, Y_i) where X_i denotes the outcome of the twin receiving treatment X and Y_i denotes the outcome of the twin receiving treatment Y. Let $E(X_i) = \mu_X$ and $E(Y_i) = \mu_Y$. Suppose it is known that $Var(X_i) = Var(Y_i) = \sigma^2$ and $Cov(X_i, Y_i) = \tau > 0$.

   Consider testing the hypothesis $H_0 : \mu_X = \mu_Y$ against $H_1 : \mu_X \neq \mu_Y$ based on

   $$T = \sqrt{n}\frac{\bar{X} - \bar{Y}}{\sqrt{S_X^2 + S_Y^2}}$$

   where notations such as $\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$ and $S_X^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2$ denote the usual sample mean and variance of the corresponding random variable.

   The test based on T rejects if $|T| > z_{0.975}$, where $P(N(0,1) \leq z_{0.975}) = 0.975$. Determine whether the probability of making a *type I error* is less than, greater than or equal to $\alpha = 0.05$ as $n \to \infty$.

   Justify your answer, and state carefully any results that you use from the lectures. (5 marks)

(Total 20 marks)