1. (10%) Approximately 1/125 of all births are fraternal twins and 1/300 of all births are identical twins. John (who is a male) had a twin brother. What is the probability that John was an identical twins? You may approximate the probability of a boy or girl birth as 1/2. [Hint: Use conditional probability].

2. (10%) Suppose we break a stick of unit length according to the distribution $f(x) = 6x(1-x)$, $0 < x < 1$. Here X is the break point. Find the expected value of the longer piece.

3. (24%) Suppose $X_1, X_2, \ldots$ is a sequence of independent, identically distributed random variables. Suppose N is an integer valued random variable, independent of $X_1, X_2, \ldots$. Define $S_N = \sum_{i=1}^{N} X_i$. Hence, there is a random number of terms in the sum (S_N is called a *random sum* of independent random variables).

   (a) (5%) Show $E(S_N) = E(N)E(X)$ and also find $Var(S_N)$.

   (b) (6%) Roll a die until a 6 appears. Let let N = *number of <u>trials</u>* until the six appears. Show that $P(N = k) = \left(\frac{1}{6}\right)\left(\frac{5}{6}\right)^{k-1}$, $k = 1, 2, \ldots$, and show the moment generating function $M_N(t) = e^t/(6 - 5e^t)$.

   Suppose now we conduct an experiment with a die and a coin in two steps:

   step 1. roll a die until a 6 appears and let N = *number of <u>trials</u>* until the six appears;

   step 2. given that $N = n$, toss a fair coin n times and record the number of heads in n trials.

   Let $X_i = 1$ if heads and 0 otherwise, then $S_N = \sum_{i=1}^{N} X_i$ is the number of heads recorded in the experiment.

   (c) (5%) Find the numerical values of $E(S_N)$ and $Var(S_N)$ for the experiment.

   (d) (8%) Show the moment generating function of S_N, $M_{S_N} = \dfrac{1+e^t}{7-5e^t}$.

4. (14%) Let X and Y be independent normal $N(0, \sigma^2)$ variables.

   (a) (8%) Show that $W = X^2 + Y^2$ and $V = X/\sqrt{W}$ are independent.

   (b) (6%) Prove $U = \sin^{-1} V$ has the uniform $(-\pi/2,\ \pi/2)$ distribution.

5. (14%) Let $X_1, X_2, \ldots$ be iid with pdf

$$f(x;\theta) = \frac{1}{\theta} e^{-x/\theta}, \quad x > 0, \quad \theta > 0$$

(a) (7%) Derive a uniformly most power (UMP) test for testing $H_0 : \theta \geq \theta_0$ against the alternative $H_1 : \theta < \theta_0$ at level of significance α.

(b) (7%) Use Central Limit Theorem to determine the minimum sample size n required to obtain power at least 0.95 against the alternative $\theta_1 = 500$ when $\theta_0 = 1000$ and $\alpha = 0.05$.

6. (18%) Consider the following genetic (linkage) data model

Type	AA	Aa	aA	aa	$Total$
Probability	$(1+\theta)/4$	$(1-\theta)/4$	$(1-\theta)/4$	$(1+\theta)/4$	1
Observed frequency	n_1	n_2	n_3	n_4	n

where θ $(-1 \leq \theta \leq 1)$ is the linkage parameter.

(a) (6%) Derive an equation which can be used to find the maximum likelihood estimator (MLE) $\hat{\theta}$ of θ. Find the MLE $\hat{\theta}$ and also compute its mean and variance.

(b) (6%) What is the (exact) distribution of $\hat{\theta}$?

(c) (6%) Is $\hat{\theta}$ the uniformly minimum variance unbiased estimate (UMVUE) of θ?

7. (10%) Let $X_1, X_2, \ldots$ be a random sample of size n from a $N(\mu, \sigma^2)$ population. Let $\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$, $S^2 = \frac{1}{n-1}\sum_{i=1}^{n}(X_i - \bar{X})^2$, and

$$T_{n-1} = \frac{\sqrt{n}(\bar{X} - \mu)}{S}.$$

Calculate the correlation $Corr(\bar{X}, T_{n-1})$, and numerically evaluate this expression for $n = 3$ and 4.
You may use the fact that if $Y \sim \chi^2_\nu$, then

$$E(Y^r) = 2^r \frac{\Gamma(\frac{\nu}{2} + r)}{\Gamma(\frac{\nu}{2})}, \quad \frac{\nu}{2} + r > 0$$

where Γ denotes the gamma function.