1. (a) Consider an independent identically distributed sequence $X_1, X_2, \cdots, X_{n+1}$ taking values 0 or 1 with probability distribution

$$P\{X_i = 1\} = 1 - P\{X_i = 0\} = p.$$

Uniformly choose M *fragments* $F_1, F_2, \cdots, F_M$ of length 2 starting in the interval [1, n], that is, $F_i = (X_{j_i}, X_{j_i+1})$ for some $1 \le j_i \le n$. Let $\mathbf{W} = (1, 1)$.

- Let $N_{\mathbf{W}}$ be the number of times the word $\mathbf{W}$ occurs among the M fragments. Calculate $E(N_{\mathbf{W}})$.
- Calculate the probability $P(F_1 = \mathbf{W}, F_2 = \mathbf{W})$.
- Calculate $\mathrm{Var}(N_{\mathbf{W}})$.

(Note: Due to time constraints, you can ignore the boundary effect.)

2. Let T and C be independent Geometric random variables with success probability of r and s, respectively. That is

$$P[T = j] = r(1-r)^{j-1}; j = 1, 2, \cdots,$$
$$P[C = j] = s(1-s)^{j-1}; j = 1, 2, \cdots,$$

Let $X = (\min(T, C), I(T \le C))$. Denote $X_1 = \min(T, C), X_2 = I(T \le C)$, where $I(\cdot)$ is the indicator function.

(a) What is the joint distribution of X?

(b) Calculate $EX = (EX_1, EX_2)$ and the covariance matrix of $X = (X_1, X_2)$.

(c) Let $T_1, T_2, \cdots, T_n$ be a random sample from T, and $C_1, C_2, \cdots, C_n$ be a random sample from C. Define

$$S_1 = \sum_{i=1}^{n} \min(T_i, C_i)$$

$$S_2 = \sum_{i=1}^{n} I(T_i \le C_i).$$

What is the maximum likelihood estimate $(\hat{r}, \hat{s})$ of (r, s), in terms of S_1 and S_2?