Probability Theory Exercise 6

1. Let $X_1, X_2, \dots, X_{2n}$ be i.i.d. random variables with the distribution $P(X_i = 1) = P(X_i = -1) = 1/2$ for all $1 \le i \le 2n$. For every $1 \le i \le 2n$, define $S_i = X_1 + X_2 + \cdots + X_i$. What is the conditional probability $P(S_i \ge 0 \text{ for all } 1 \le i \le 2n \mid S_{2n} = 0)$?

2. Let $X_1, X_2, X_3, \dots$ be i.i.d. Bernoulli-p random variables, i.e., $P(X_i = 1) = p \in (0,1)$ and $P(X_i = 0) = 1 - p$ for all i. Define another sequence $\{L_n\}_{n=1}^{\infty}$ of integer-valued random variables as follows: If $X_n = 0$, then we define $L_n = 0$. If $X_n = 1$, then we define L_n as the unique positive integer such that $X_n = X_{n+1} = X_{n+2} = \cdots = X_{n+L_n-1} = 1, X_{n+L_n} = 0$. Define $L_{\max}^{(n)} := \max(L_1, \dots, L_n)$. Prove that there is a function of p, which we denote as $f(p)$, such that

$$\lim_{n\to\infty} P(L_{\max}^{(n)} \ge c\log(n)) = 1 \text{ for any constant } c < f(p),$$
$$\lim_{n\to\infty} P(L_{\max}^{(n)} \ge c\log(n)) = 0 \text{ for any constant } c > f(p).$$

Also solve the exact expression of $f(p)$. (We use natural logarithm in this problem.)

3. Consider a random walk on the integers. We start at $X_0 = 0$. In each step, we have $P(X_n = X_{n-1}+1) = P(X_n = X_{n-1} - 1) = 1/2$. Define the random variable U as the unique positive integer such that $X_U = 0$ and $X_i \ne 0$ for all $0 < i < U$. In other words, in step U, this random walk returns to the origin for the first time. Now let $m > 0$ be a positive integer. Define another random variable N_m as the number of times this random walk visits m before step U. More precisely, we have

$$N_m := |\{i : 0 < i < U, X_i = m\}|.$$

(For a set A, $|A|$ denotes the size of A.) Note that N_m is simply the number of times this random walk visits m before returning to 0.
(i) Find $P(N_m \ge 1)$.
(ii) For every positive integer n, find $P(N_m = n)$.

4. Passengers arrive at a train station as a Poisson process with rate λ.
(i) Suppose that there is only one train, and it departs at a deterministic time T. Let W be the sum of waiting time of all the passengers till time T, i.e., W does not include the waiting time of passengers arrived after time T. Compute $\mathbb{E}[W]$.
(ii) Now suppose that there are two trains. One departs at T, and the other departs at $S < T$. Compute $\mathbb{E}[W]$ in this case.

5. Let $G \sim \mathbb{G}(n,p)$ with $p = \frac{\log(n)}{n} + \frac{c}{n}$ for some constant c. (c can be negative.) Find

$$\lim_{n\to\infty} P(G \text{ is connected}).$$

6. Let $G \sim \mathbb{G}(n,p)$ with $p = n^{-\delta}$ for some constant $\delta > 0$. Prove that there exists a constant $\delta_0 > 0$ such that

$$\lim_{n\to\infty} P(G \text{ contains 4 vertices that are pairwise connected}) = 0 \text{ if } \delta > \delta_0,$$
$$\lim_{n\to\infty} P(G \text{ contains 4 vertices that are pairwise connected}) = 1 \text{ if } \delta < \delta_0.$$

Also find the value of δ_0. (Note that G contains a triangle means that G contains 3 vertices that are pairwise connected.)