Problems

1. **(10 points)** For each of the following questions, exactly one answer is correct. Each correct answer gives 1 point, and each incorrect answer results in a 1/2 point reduction. The minimal possible total score for the full problem is 0.

   a) Let $(S_n)_{n=0,\dots,N}$ be a random walk. Which of the following is NOT a stopping time?

   1. $\inf\{n : S_n \geq 5\} \wedge N$
   2. $\inf\{n : S_{n+1} \geq 5\} \wedge N$
   3. $\inf\{n : S_{n-1} \geq 5\} \wedge N$

   b) Let $(S_n)_{n=0,..,N}$ be a random walk. Which of the following statements is TRUE?

   1. $\mathbb{E}(S_n^2)$ is increasing in n.
   2. For any stopping time T, $\mathbb{E}(S_T^2) = \mathbb{E}(S_0^2)$.
   3. $\text{Var}(S_n^2) = n^2$.

   c) Let μ and μ_n, $n \in \mathbb{N}$, be distributions on $\mathbb{R}$. Let F, F_n be their respective distribution functions. Which of the following statements is equivalent to $\mu_n \to \mu$ weakly?

   1. $F_n(x) \to F(x)$ for every $x \in \mathbb{R}$ at which F is continuous.
   2. $\int_{\mathbb{R}} f(x) d\mu_n(x) \to \int_{\mathbb{R}} f(x) d\mu(x)$ for every function $f : \mathbb{R} \to \mathbb{R}$.
   3. $F_n(x) \to F(x)$ for every $x \in \mathbb{R}$ at which F is strictly positive.

   d) Which of the following statements is FALSE?

   1. Sum of independent Bernoulli random variables is still a Bernoulli random variable.
   2. Sum of independent Poisson random variables is still a Poisson random variable.
   3. Sum of independent Normal random variables is still a Normal random variable.

   e) Let X_1 and X_2 be independent random variables with characteristic function ϕ_1 and ϕ_2, respectively. What is the characteristic function ϕ of $X_1 - X_2$?

   1. $\phi(u) = \phi_1(u) - \phi_2(u)$
   2. $\phi(u) = \phi_1(u)/\phi_2(u)$
   3. $\phi(u) = \phi_1(u)\phi_2(-u)$

   f) If X_n converges in probability to X, then which of the following statements follows?

   1. X_n converges almost surely to X.
   2. $\lim_{n\to\infty} \mathbb{E}(X_n) = \mathbb{E}(X)$.
   3. X_n converges weakly to X.

g) The characteristic function ϕ of a Binomial(n, p) random variable is given by:

1. $\phi(u) = (e^{ui}n + (1-n))^p$
2. $\phi(u) = (e^{ui}p + (1-p))^n$
3. $\phi(u) = (e^{ni}p + (1-p))^u$

h) Let $(X_n)_{n\in\mathbb{N}}$ be an i.i.d. sequence of $N(0, \sigma^2)$ random variables. Define $\overline{\sigma}_n = \frac{1}{n}\sum_{i=1}^n X_n^2$ and $\widetilde{\sigma}_n = \frac{1}{n-1}\sum_{i=1}^n X_n^2$. Which property is TRUE about these estimators?

1. $\widetilde{\sigma}_n$ is unbiased.
2. $\overline{\sigma}_n$ is consistent.
3. $\mathrm{Var}(\overline{\sigma}_n) \geq \mathrm{Var}(\widetilde{\sigma}_n)$.

i) Suppose a statistical test resulted in a p-value of 0.025. Which of the following statements is TRUE?

1. The null hypothesis could not be rejected at significance level 0.01.
2. The probability that the null hypothesis is true is 0.025.
3. The null hypothesis could not be rejected at significance level 0.05

j) Let A and B be two events with positive probability. Which of the following equations is Bayes' rule?

1. $\mathbb{P}(B \mid A) = \mathbb{P}(A \mid B)\mathbb{P}(B)/\mathbb{P}(A)$
2. $\mathbb{P}(B \mid A) = \mathbb{P}(A \mid B)\mathbb{P}(A)/\mathbb{P}(B)$
3. $\mathbb{P}(B \mid A) = \mathbb{P}(A \cap B)\mathbb{P}(B)/\mathbb{P}(A)$

2. (15 points) Let $X_1, X_2, \ldots$ be an i.i.d. sequence of Poisson(λ) distributed random variables for some fixed $\lambda > 0$.

a) How does $\frac{1}{n}\sum_{i=1}^n X_i$ behave as $n \to \infty$?

b) How does $\frac{1}{\sqrt{n}}\sum_{i=1}^n (X_i - \lambda)$ behave as $n \to \infty$?

Suppose now you observe a noisy version of the X_i's. Specifically, let $Y_i = X_i + Z_i$, where $Z_i \sim N(0, \sigma^2)$ for some $\sigma \geq 0$. Assume all random variables are mutually independent.

c) How does $\frac{1}{\sqrt{n}}\sum_{i=1}^n (Y_i - \lambda)$ behave as $n \to \infty$?

d) Suppose you observe a sample $Y_1, \ldots, Y_n$, and would like to test for the presence of noise. Specifically, you consider the hypotheses

$$\begin{cases} \text{Null:} & \sigma^2 = 0; \\ \text{Alternative:} & \sigma^2 > 0. \end{cases}$$

Design a test at significance level $\alpha = 0$, with power 1 at any $\sigma^2 > 0$.

3. (15 points) Consider the probability density function

$$f_X(x) = \begin{cases} (\alpha + \alpha^2) x^{\alpha - 1}(1 - x) & x \in (0, 1) \\ 0 & \text{otherwise} \end{cases}$$

where $\alpha > 0$ is a parameter. The corresponding distribution is called the Beta$(\alpha, 2)$ distribution.

a) Show that f_X is indeed a probability density function.

b) Show that $\mathbb{E}(-\log X) = \dfrac{1 + 2\alpha}{\alpha + \alpha^2}$, where $X \sim f_X$.

c) Let $\widehat{\alpha}_n$ denote the maximum likelihood estimator of α based on an i.i.d. sample $X_1, \ldots, X_n$ from f_X. Show that $\widehat{\alpha}_n$ exists, is unique, and is the solution of the equation

$$\frac{1 + 2\widehat{\alpha}_n}{\widehat{\alpha}_n + \widehat{\alpha}_n^2} = -\frac{1}{n} \sum_{i=1}^{n} \log X_i.$$

d) Show that the maximum likelihood estimator is consistent.

4. (15 points) Let $X_1, X_2, \ldots$ be i.i.d. Exponential(λ) for some fixed $\lambda > 0$, and consider the maxima $M_n = \max(X_1, \ldots, X_n)$ for each n.

a) Show that $\mathbb{P}(M_n \le x) = (1 - e^{-\lambda x})^n$ for $x \ge 0$.

b) Show that the distribution of $\lambda M_n - \log n$ converges weakly to the *Gumbel distribution*, whose distribution function is $F(x) = e^{-e^{-x}}$, $x \in \mathbb{R}$.

c) Let $(a_n)_{n\in\mathbb{N}}$ be a positive increasing sequence with $\lim_{n\to\infty} \frac{\log n}{a_n} = c$ for some $c \in [0, \infty]$. Show that $\lim_{n\to\infty} a_n^{-1} M_n = \lambda^{-1} c$ in probability.

d) Show that $\lim_{n\to\infty} M_n = \infty$ almost surely.

5. **(15 points)** Define $S_0 = 0$ and $S_n = \sum_{i=1}^n X_i$ for $n \geq 1$, where $X_1, X_2, \ldots$ are i.i.d. random variables with $\mathbb{P}(X_i = +1) = \mathbb{P}(X_i = -1) = \frac{1}{2}$. Let $(\Omega, \mathcal{A}, \mathbb{P})$ denote the probability space on which these objects are defined.

a) Fix any $N \in \mathbb{N}$ and consider $(S_n)_{n=0,\ldots,N}$, which is a random walk of length N. For each $n = 0, \ldots, N$, let $\mathcal{A}_n$ denote the set of all events of the form $\{\omega\colon (S_0(\omega), \ldots, S_n(\omega)) \in D\}$ with $D \subseteq \mathbb{R}^n$ measurable. State the definition of a stopping time.

Recall the following two facts, which you may use later on:

(i) For any stopping time T, one has $\mathbb{E}(S_T) = 0$. Here it is crucial that $T \leq N$ almost surely for some deterministic number N.

(ii) For any $a \in \mathbb{Z}$, letting $T_a = \inf\{n > 0 : S_n = a\}$ one has $\lim_{N\to\infty} \mathbb{P}(T_a > N) = 0$.

Fix integers $a > 0 > b$ and let $T_{a,b} = \inf\{n > 0 : S_n = a \text{ or } S_n = b\}$ denote the first time S_n hits either a or b. If this never happens, set $T_{a,b} = \infty$.

b) Show that $\mathbb{P}(T_{a,b} < \infty) = 1$.

c) Show that $\mathbb{E}(S_{T_{a,b}}) = 0$.

Hint: Consider the stopping time $T = \min(T_{a,b}, N)$, apply the above facts about stopping times, and take limits.

d) What is the probability that S_n reaches a before it reaches b?