

Qualifying Exam: 2026 Spring

考试课程: Probability & Statistics 姓名: _____ 学号: _____

- There are 11 problems in this exam (4 pages). You need to choose 8 of them to solve. If you select more than 8, only the first 8 that you have worked on will be graded. Note that 4 of the problems are worth 15 points each and the rest 10 points each.
- You must follow all the rules of exam taking. Misconducts will be subject to proper disciplinary actions by the Center.
- You must provide all necessary details for full credits. A final answer with no or little explanation/derivation, even if correct, receives a minimal credit.
- \mathbb{R} denotes the set of real numbers and $\mathbb{N} = \{1, 2, 3, \dots\}$ denotes the set of positive integers. $\xrightarrow{(d)}$ and $\stackrel{(d)}{=}$ mean “converges in distribution” and “equal in distribution”, respectively.
- The pmf of a Poisson(λ) random variable X is given by $P(X = x | \lambda) = \frac{e^{-\lambda} \lambda^x}{x!}; x = 0, 1, \dots; \lambda > 0$.

1. (10 points) Let $X = (X_1, \dots, X_n)^T$ be an \mathbb{R}^n -valued Gaussian random vector under a probability measure \mathbb{P} , with mean vector $\mu \in \mathbb{R}^n$ and covariance matrix $V \in \mathbb{R}^{n \times n}$. Let $\alpha \in \mathbb{R}^n$ and define a new probability measure \mathbb{Q} by

$$\frac{d\mathbb{Q}}{d\mathbb{P}} = \frac{\exp(\alpha^T X)}{\mathbb{E}_{\mathbb{P}}[\exp(\alpha^T X)]}.$$

Show that X is still Gaussian under \mathbb{Q} and identify the mean and covariance matrix of X under \mathbb{Q} .

2. (15 points) Let $\{X_n\}_{n \in \mathbb{N}}$ be a sequence of independent random variables with distribution:

$$\begin{cases} \mathbb{P}[X_n = n^2] = \mathbb{P}[X_n = -n^2] = \frac{1}{6n^2}, \\ \mathbb{P}[X_n = n] = \mathbb{P}[X_n = -n] = \frac{1}{6}, \\ \mathbb{P}[X_n = 0] = \frac{2}{3} - \frac{1}{3n^2}. \end{cases}$$

Let $S_n = \sum_{j=1}^n X_j$. Find a sequence of positive real numbers $\{b_n\}_{n \in \mathbb{N}}$ such that the law of S_n/b_n converges to standard normal distribution $\mathcal{N}(0, 1)$.

3. (10 points) We have 100 noodles in a bowl. Each noodle has two free ends, so initially there are 200 free ends in total. We are blindfolded and repeatedly perform the following operation: choose uniformly at random two free ends among all available free ends and connect them together. We continue until there are no free ends left. At the end of this procedure, the noodles form a random collection of disjoint loops. Compute the expectation of the number of loops.
4. (10 points) Let $\{B_t\}_{t \geq 0}$ be a one-dimensional standard Brownian motion started from the origin. Find and prove the limit (in distribution) of

$$\left[\int_0^{T^2} \exp(B_t) dt \right]^{1/T} \text{ as } T \rightarrow \infty.$$

Write your answer as a function of the standard normal distribution.

5. (10 points) A Galton-Watson branching process is a stochastic process $\{Z_n\}_{n \in \mathbb{N} \cup \{0\}}$ which evolves according to the recurrence formula

$$Z_0 := 1, Z_{n+1} := \sum_{j=1}^{Z_n} X(j, n),$$

where $\{X(j, n) : j \in \mathbb{N}, n \in \mathbb{N} \cup \{0\}\}$ is a set of i.i.d. $\mathbb{N} \cup \{0\}$ -valued random variables with mean $m := \mathbb{E}[X(1, 1)] < \infty$. Let $p_k := \mathbb{P}[X(1, 1) = k]$, $k \in \mathbb{N} \cup \{0\}$, and we assume $p_0 > 0$, $p_0 + p_1 < 1$. Let $f(s) := \mathbb{E}[s^{X(1,1)}]$ for $s \in [0, 1]$ (with the convention $0^0 = 1$). Let

$$\mathcal{F}_n := \sigma(Z_0, \dots, Z_n), n \in \mathbb{N} \cup \{0\}.$$

- (a) Compute

$$\mathbb{E}[Z_{n+1} | \mathcal{F}_n], \mathbb{E}[Z_n], \mathbb{E}[s^{Z_{n+1}} | \mathcal{F}_n], \mathbb{E}[s^{Z_n}], n \in \mathbb{N} \cup \{0\}, s \in [0, 1].$$

- (b) Determine the number of solutions of $f(s) = s$ on $[0, 1]$.

- (c) Let $\tau := \inf\{n \in \mathbb{N} : Z_n = 0\}$ (with $\inf \emptyset := \infty$). Call the event $\{\tau < \infty\}$ *extinction*. Compute $\mathbb{P}[\tau < \infty]$. Write your answer as a function of q , where q is the smallest solution of $f(s) = s$ on $[0, 1]$.

6. (15 points) In the above definition of Galton-Watson branching process, Z_n can be thought of as the number of descendants in the n -th generation, and $X(j, n)$ can be thought of as the number of children of the j -th of these descendants. By tracking genealogical relationships, we obtain a tree T rooted at the single individual in generation 0 with a vertex for each individual in the progeny and an edge for each parent-child relationship. T is called a *Galton-Watson tree*.

A property of rooted trees is said to be *inherited* if this property holds for all finite trees, and whenever it holds for a tree, it also holds for all subtrees rooted at the children of the root.

- (a) Prove that for a Galton-Watson tree T , conditioned on non-extinction, an inherited property A has a probability of either 0 or 1. (Hint: One may use $s \leq f(s)$ implies $s \in [0, q] \cup \{1\}$.)
- (b) Let T be the Galton-Watson tree for an offspring distribution with mean $m > 1$. Perform percolation on T with density p (i.e., each edge of T is open with probability p and closed with probability $1 - p$, independently of all other edges). Let \mathcal{C}_0 be the maximal subtree of the root formed by open edges in T . Define the critical probability

$$p_c(T) := \sup\{p \in [0, 1] : \mathbb{P}_p[|\mathcal{C}_0| = \infty | T] = 0\},$$

where $|\mathcal{C}_0|$ is the number of vertices in \mathcal{C}_0 . Conditioned on non-extinction of T , compute $p_c(T)$.

7. (10 points) Let X_1, X_2, \dots, X_n be i.i.d. observations with pdf

$$f(x; \theta) = \frac{1}{2\theta\sqrt{x}} e^{-\sqrt{x}/\theta} I_{\{x>0\}},$$

where $\theta > 0$ is the unknown parameter.

- (a) Show that the family $f(x_1, \dots, x_n; \theta)$ ($\theta > 0$) has a monotone likelihood ratio in a certain statistic T .
- (b) Find a size α ($0 < \alpha < 1$) uniformly most powerful test for $H_0 : \theta \geq 1$ versus $H_1 : \theta < 1/2$. Make your rejection region expressed as explicitly as possible.
8. (15 points) Let X_1, X_2, \dots, X_n be i.i.d. Bernoulli random variables with unknown success probability $0 < \theta < 1$. Given $0 < p < 1$, let $\hat{\xi}_p$ be the sample quantile based on the n observations, i.e., $\hat{\xi}_p = \inf\{x : F_n(x) \geq p\}$, where F_n denotes the empirical CDF based on the data.
- (a) Does $\hat{\xi}_p$ converge in probability to the population quantile ξ_p for each $0 < p < 1$ as $n \rightarrow \infty$? Prove your answer.
- (b) Is it possible that for a non-degenerate discrete distribution F we have $\hat{\xi}_p$ converges in probability to ξ_p for all $0 < p < 1$? Prove your answer.
9. (10 points) Suppose a single observation $Y | \mu \sim N(\mu, 1)$, and the prior on μ is $\pi(\mu) = 0.5 \cdot I(\mu = -1) + 0.5 \cdot I(\mu = 1)$ where $I(\cdot)$ is the indicator function. That is, under this prior, μ is supported on $\{-1, 1\}$ and takes either value with a probability 0.5. For estimating μ , let the action space be $\{-1, 1\}$.
- (a) Write down the posterior $\pi(\mu | Y = y)$ and derive the Bayes rule δ under the absolute error loss $L(\mu, a) = |\mu - a|$.
- (b) Is δ above a minimax rule under the absolute error loss? Please justify your answer.

10. (10 points) Let X_1, X_2, \dots, X_n be i.i.d. random variables following $\text{Poisson}(\sqrt{\lambda})$ distribution where $\lambda > 0$.

(a) Find the BUE (UMVUE) $\tilde{\lambda}$ for λ .

(b) Let $\hat{\lambda}$ denote the MLE for λ based on X_1, X_2, \dots, X_n , show that $\sqrt{n}(\tilde{\lambda} - \hat{\lambda}) \rightarrow 0$ in probability as $n \rightarrow \infty$.

11. (15 points) Let $Y = \sigma(\rho|U| + \sqrt{1 - \rho^2}V)$ where U, V are independent $N(0, 1)$ random variables and $\rho \in (-1, 1), \sigma \in (0, \infty)$ are unknown parameters.

(a) Show that the pdf of Y is given by

$$f(y|\rho, \sigma) = \frac{a}{\sqrt{2\pi\sigma^2}} \exp\left(-\frac{y^2}{2\sigma^2}\right) \Phi\left(\frac{\rho y}{b\sigma\sqrt{1-\rho^2}}\right), \quad -\infty < y < \infty$$

for some positive constants a, b . Write down the values of a and b . In the above expression, $\Phi(z)$ denotes the CDF for a standard normal random variable.

(b) Consider n observations Y_1, \dots, Y_n , modeled as $Y_i \stackrel{\text{iid}}{\sim} f(y_i|\rho, \sigma), \rho \in (-1, 1), \sigma \in (0, \infty)$. Does a maximum likelihood estimate of $(\rho, \sigma) \in (-1, 1) \times (0, \infty)$ always exist? Please justify your answer.

(c) For the same model, suppose we are interested in testing $H_0 : \rho = 0$ versus $H_1 : \rho \neq 0$. What is the value of c such that the test that rejects H_0 if

$$\frac{\sqrt{n}|\bar{Y}|}{s_Y} > c$$

has size $\alpha = 0.05$? Identify c as a specific quantile of a named distribution. In the above expression, \bar{Y} is the sample mean and s_Y is the sample standard deviation.