NTHU MATH 2810 Final Examination - A Dec 17, 2024

Note. There are 6 problems in total. **For problems 2 to 6, to ensure consideration for partial/full scores, write down necessary intermediate steps.** Correct answers with inadequate or no intermediate steps may result in zero credit.

Some useful formula.

- The probability mass function (pmf) of a binomial(n, p) distribution is

$$p(x) = \binom{n}{x} p^x (1-p)^{n-x}, \text{ for } x = 0, 1, \ldots, n,$$

  and zero, otherwise. Its mean and variance are np and $np(1-p)$, respectively. When $n = 1$, it is called Bernoulli(p) distribution.

- The pmf of a hyper-geometric(n, N, R) distribution is

$$p(x) = \frac{\binom{R}{x}\binom{N-R}{n-x}}{\binom{N}{n}}, \text{ for } x = 0, 1, \ldots, n.$$

  Its mean and variance are nR/N and $[nR(N-R)(N-n)]/[N^2(N-1)]$, respectively.

- The pmf of a negative binomial(r, p) distribution is

$$p(x) = \binom{x-1}{r-1} p^r (1-p)^{x-r}, \text{ for } x = r, r+1, r+2, \ldots$$

  Its mean and variance are r/p and $r(1-p)/p^2$, respectively.

- The probability density function (pdf) of a uniform(α, β) distribution is

$$f(x) = \frac{1}{\beta - \alpha}, \text{ for } \alpha < x < \beta,$$

  and zero, otherwise. The cumulative distribution function (cdf) of a uniform(α, β) distribution is

$$F(x) = \frac{x - \alpha}{\beta - \alpha}, \text{ for } \alpha < x < \beta,$$

  and $F(x) = 0$ when $x \leq \alpha$ and $F(x) = 1$ when $x \geq \beta$. Its mean and variance are $\frac{\alpha+\beta}{2}$ and $\frac{(\beta-\alpha)^2}{12}$, respectively.

- The pmf of a Poisson(λ) distribution is

$$p(x) = \frac{\lambda^x e^{-\lambda}}{x!}, \text{ for } x = 0, 1, 2, \ldots,$$

  and zero, otherwise. Its mean and variance are λ and λ, respectively.

- The pdf of an exponential(λ) distribution is

$$f(x) = \lambda e^{-\lambda x}, \text{ for } 0 < x < \infty,$$

and zero, otherwise. The cdf of an exponential(λ) distribution is

$$F(x) = 1 - e^{-\lambda x}, \text{ for } 0 < x < \infty,$$

and $F(x) = 0$ when $x \leq 0$. Its mean and variance are $\frac{1}{\lambda}$ and $\frac{1}{\lambda^2}$, respectively.

- The pdf of a gamma(α, λ) distribution is

$$f(x) = \frac{\lambda^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\lambda x}, \text{ for } x \geq 0,$$

and zero, otherwise. Its mean and variance are α/λ and α/λ^2, respectively. Notice that $\Gamma(\alpha) = (\alpha - 1)!$ if α is a positive integer.

- The pdf of a normal(μ, σ^2) distribution is

$$f(x) = \frac{1}{\sqrt{2\pi}\sigma} e^{-\frac{(x-\mu)^2}{2\sigma^2}}, \text{ for } -\infty < x < \infty.$$

Its mean and variance are μ and σ^2, respectively.

- The joint pmf of a multinomial(n, m, $p_1, \ldots, p_m$), where $p_1 + \cdots + p_m = 1$, is

$$p(x_1, \ldots, x_m) = \binom{n}{x_1, \cdots, x_m} p_1^{x_1} \times \cdots \times p_m^{x_m},$$

for $0 \leq x_i \leq n$, $i = 1, \ldots, m$, and $x_1 + \cdots + x_m = n$.

---

1. (*13 points*) For each of the random observation(s) $\boldsymbol{X}$ below,

   - determine the type of the (joint) distribution (i.e., normal(μ, σ^2), exponential(λ), gamma(α, λ), beta(α, β), uniform, Weibull(α, β, ν), Cauchy(μ, σ), Poisson(λ), hyper-geometric(n, N, R), binomial(n, p), Bernoulli(p), multinomial($n, m, p_1, \ldots, p_m$), negative binomial(r, p), geometric(p), ..., etc.) which best models $\boldsymbol{X}$, and
   - identify the values of the parameters for the chosen distribution **when the question supplies enough information**.

(a) (*5 points*) A survey selects 50 male students and 60 female students, each from different families, to participate in a study about which two of ten given attributes are most desirable in their fathers. Each student is asked to choose two attributes out of the ten. The following table shows the counts for how the attribute "being a college graduate" is rated:

	Chosen	Not Chosen
Male	X_1	X_2
Female	Y_1	Y_2

where $X_1 + X_2 = 50$ and $Y_1 + Y_2 = 60$. Set $\boldsymbol{X} = (X_1, X_2, Y_1, Y_2)$.

(b) (*4 points*) George, Mary, and Hilary live in a certain house in which phone calls are received as a Poisson process with parameter λ per hour, where λ is unknown. Someday, in order to receive an important phone call, they came up with a schedule: George waited by the phone between 9AM-10AM, Mary waited by the pone between 10AM-12PM, and Hilary waited by the phone between 12PM-5PM. Let X_1, X_2, X_3 be the numbers of phone calls that George, Mary, and Hilary received respectively. Suppose that the total number of phone calls they received between 9AM-5PM was 12. Given this information, set $\boldsymbol{X} = (X_1, X_2, X_3)$.

(c) (*4 points*) A trained flea sits on the real line at $t = 3$, and its master begins flipping a biased coin. The probability of landing heads is $1/4$, and the probability of landing tails is $3/4$. Each time the coin shows a head, the flea hops one unit to the right, each time a tail shows it hops one unit to the left. Let T be the *discrete* random variable representing "the flea's position on the real line after n flips." Let $\boldsymbol{X}$ be a *continuous* random variable such that, as n becomes large, the distribution of $\boldsymbol{X}$ nearly match the distribution of T, i.e., $\boldsymbol{X}$ and T have almost identical cumulative distribution functions (cdfs). [**Hint**. If $U \sim \text{Bernoulli}(p)$ and $V = 2U - 1$, then $V = 1$ with probability p, and $V = -1$ with probability $1 - p$.]

2. (*24 points*) Compute the following probabilities. A correct answer without intermediate steps will receive no credit.

   (a) (*6 points*) Let X and Y be independent random variables each uniformly distributed on $(0, 1)$. Find $P\left(\min(X, Y) = X \,\middle|\, Y \geq \frac{3}{4}\right)$.

   (b) (*8 points*) The time that it takes to service a car is an exponential random variable with rate $\lambda = 1$. If A.J. brings his car in a time 0 and M.J. brings her car in a time

$t > 0$, what is the probability that M.J.'s car is ready before A.J.'s car? (assume that service times are independent and service begins upon arrival of the car.)

(c) (*10 points*) Choose two points independently and uniformly on the circumference of the unit circle. Find the probability that the *squared distance* between the two points is larger than 3. [**Hint.** (i) express the two points as $(\cos(\Theta_1), \sin(\Theta_1))$ and $(\cos(\Theta_2), \sin(\Theta_2))$, where Θ_1 and Θ_2 are random variables; (ii) $\cos(\Theta_1 - \Theta_2) = \cos(\Theta_1)\cos(\Theta_2) + \sin(\Theta_1)\sin(\Theta_2)$.]

3. (*17 points*) Let $X_1, \ldots, X_n$ be independent uniform$(0,1)$ random variables. Define $X_{(1)} = \min(X_1, \ldots, X_n)$ and $X_{(n)} = \max(X_1, \ldots, X_n)$. Let

$$R = X_{(n)} - X_{(1)} \quad \text{and} \quad M = \frac{X_{(n)} + X_{(1)}}{2}.$$

(a) (*3 points*) Provide an explanation or demonstrate that the joint pdf of $X_{(1)}$ and $X_{(n)}$ is

$$f_{X_{(1)},X_{(n)}}(s,t) = \begin{cases} n(n-1)(t-s)^{n-2}, & \text{if } 0 < s < t < 1, \\ 0, & \text{otherwise.} \end{cases}$$

(b) (*6 points*) Find the joint cumulative distribution function (joint cdf) $F_{X_{(1)},X_{(n)}}(u,v)$ of $X_{(1)}$ and $X_{(n)}$.

(c) (*5 points*) Compute the joint pdf of R and M.

(d) (*3 points*) Derive and express the covariance of R and M as a function of $Var(X_{(1)})$ and $Var(X_{(n)})$.

4. (*12 points*) Suppose that X and Y are independent binomial random variables with identical parameters n and p. Let $U = X$ and $V = X + Y$. Show analytically that the conditional distribution of U given $V = m$, is the hyper-geometric distribution, and identify the parameters of this hyper-geometric distribution using n, m, p.

5. (*17 points*) Suppose that W, the amount of moisture in the air on a given day, is a gamma random variable with parameters (α, λ), where α is a positive integer. Suppose also that given $W = w$, the number of accidents during that day — call it N — has a Poisson distribution with mean w.

(a) (*3 points*) Use multiplication law to find the joint distribution of W and N.

(b) (*5 points*) Use the law of total probability to show that $N + \alpha$ is a negative binomial random variable with parameters $\left(\alpha, \frac{\lambda}{\lambda+1}\right)$.

(c) (*5 points*) Use Bayes theorem to show that the conditional distribution of W given that $N = n$ is the gamma distribution with parameters $(\alpha + n, \lambda + 1)$.

(d) (*4 points*) Use the law of total expectation to find the mean of N.

6. (*17 points*) Suppose that k, where $k \geq 2$, married couples are randomly seated at a round table with $2k$ different seats. Let the random variable N be the number of wives sitting next to their husbands. Let I_i, $i = 1, \ldots, k$, be the indicator functions with $I_i = 1$ if the couple i sits together, and 0 if they do not.

(a) (*5 points*) Show that

$$P(I_i = 1) = \frac{2}{2k-1},$$

and

$$P(I_i = 1, I_j = 1) = \frac{2}{(k-1)(2k-1)},$$

for $i \neq j$.

(b) (*5 points*) What is the expectation of N?

(c) (*7 points*) What is the variance of N?

[**Hint**. (i) Represent N as a function of the I_i's. (ii) Note that the random variables I_i's are **not** independent as shown in (a).]