National Chiao Tung University
Department of Computer Science

Fall 2020: 1175 Probability – Midterm

2020/11/04, 10:10AM-12PM

- This exam contains 6 problems. Total of points is 110.
- Write your solutions in the booklet, and only the solutions in the booklet will be graded.
- You are allowed to bring a two-sided, A4 formula sheet.
- When giving a formula for a CDF/PMF/PDF, make sure to specify the range over which the formula holds.
- For Problem 1, please mark each subproblem with 'O' or 'X'.
- Problem 1 (true/false questions) will receive no partial credit. Partial credit for Problems 2-6 will be awarded.
- Good luck! May the force be with you.

1. (28 points) Please mark whether each of the following arguments is true (O) or false (X). If you mark a subproblem as false (X), please briefly explain the reason or provide a counterexample. If you mark a subproblem as true (O), then no justification is required:

   (a) (4 points) Let X be a standard normal random variable. Then, for any $a \neq 0, b \in \mathbb{R}$, $Y = aX + b$ is a normal random variable with $E[Y] = b$ and $\mathrm{Var}[Y] = a^2$.

   (b) (4 points) Let Ω be the universal set and B, C be two sets that satisfy $B \subseteq \Omega$ and $C \subseteq \Omega$. Let $\{F_k\}_{k=1}^{\infty}$ denote the Fibonacci sequence, i.e. $F_1 = F_2 = 1$ and $F_{k+1} = F_k + F_{k-1}$, for $k \geq 2$. Define a countably infinite sequence of sets $A_1, A_2, A_3, \cdots$ as

$$A_n = \begin{cases} B, & \text{if } n \text{ is in the Fibonacci sequence } \{F_k\}, \\ C, & \text{otherwise.} \end{cases}$$

   Then, $\bigcup_{m=1}^{\infty} \bigcap_{n=m}^{\infty} A_n = C$.

   (c) (4 points) Let $X_1, \cdots, X_n$ be n independent Geometric random variables with parameters $p_1, \cdots, p_n$, respectively (n is a positive integer, and $p_i > 0$ for each $i \in \{1, \cdots, n\}$). Define $X = \min(X_1, \cdots, X_n)$. Then, X is also a Geometric random variable.

   (d) (4 points) Let X be a random variable and suppose $P(X \in [0, 8/(4+n)]) = (1 + e^{-n})/6$, for all $n \in \mathbb{N}$. Then, we know $P(X = 0) = 0$.

   (e) (4 points) Let X be a continuous random variable with CDF $F_X(t)$. Then, the PDF of X is unique and can be derived by taking the derivative of $F_X(t)$.

   (f) (4 points) Let Y be a discrete random variable with PMF $p_Y(k)$ defined as

$$p_Y(k) = \begin{cases} 1/(2k(k+1)), & \text{if } k = 1, 2, 3, \cdots, \\ 1/(2k(k-1)), & \text{if } k = -1, -2, -3, \cdots, \\ 0, & \text{otherwise.} \end{cases}$$

   Then, we have $E[Y] = 0$.

   (g) (4 points) For any $n \in \mathbb{N}$, we have $C_0^n + C_1^n + \cdots + C_n^n = 2^n$.

2. (10 points) Suppose we toss the same fair coin twice and define two events $E_1 = \{\text{1st toss is a head}\}$ and $E_2 = \{\text{2nd toss is a tail}\}$. We also define another event $D = \{\text{2 tosses have different results}\}$. Are the three events E_1, E_2, D pairwise independent? Moreover, are the three events E_1, E_2, D independent? Please clearly explain your answer.

3. (10 points) Let U be a continuous uniform random variable between 0 and 1. Consider a function $G(z) = \ln(z/(1-z))$ and accordingly define another random variable $V = G(U)$. Please find out the CDF of V and clearly explain your answer. (Hint: Inverse transform sampling)

4. (20 points) Suppose X is a standard normal random variable, i.e. $X \sim \mathcal{N}(0, 1)$.

   (a) (12 points) Find the expected value and variance of e^X. (Hint: use LOTUS, i.e. $E[g(X)] = \int g(x) f_X(x) dx$)

(b) (8 points) Define another random variable Y as

$$Y = \begin{cases} 2X+3, & \text{if } X > 1, \\ X-5, & \text{otherwise.} \end{cases}$$

Please write down the CDF of Y. Is Y a normal random variable? (Note: You may use $\Phi(\cdot)$ to denote the CDF of a standard normal)

5. (26 points) In a wireless communication system, a binary message ('+' or '−') is transmitted as a wireless signal, whose value is either $+1$ ('+') or -1 ('−'). At each transmission trial, the transmitter sends either a '+' with probability p, or a '−' with probability $1-p$ ($p \in (0,1)$).

   (a) (8 points) Let X be the wireless signal of the first transmitted message. Suppose the wireless channel corrupts the transmission with an additive continuous uniform noise $Y \sim \text{Unif}(-2,2)$, independent from X. The received signal Z can thereby be expressed as $Z = X + Y$. The receiver concludes that the signal '+' (or '−') was transmitted if $Z \geq 0$ (or $Z < 0$, respectively). Given that $X = +1$, what is the probability of error (i.e. $P(Z < 0|X = +1)$)? Similarly, given that $X = -1$, what is the probability of error (i.e. $P(Z \geq 0|X = -1)$)?

   (b) (8 points) Based on (a), given that the receiver gets a '+', what is the probability that '−' is sent by the transmitter? (Hint: Bayes' rule)

   (c) (10 points) Suppose the number of transmissions within a given observation window T has a Poisson PMF with average rate λ ($\lambda > 0$). Define a random variable Q to be the number of +'s *transmitted* in that observation window. Please write down the PMF of Q. Is Q also a Poisson random variable? Please clearly explain your answer. (Hint: Define V = total number of transmissions in the given interval. Leverage the total probability theorem $P(Q = k) = \sum_{n=0}^{\infty} P(Q = k|V = k+n) \cdot P(V = k+n)$)

6. (16 points) It is well-known that there is a secret path connecting NCTU and NTHU, and there is a long-lasting debate over the name of this path (either called "交清小徑" or "清交小徑"). To better understand the status quo, we conduct the following experiments:

   (a) (10 points) Suppose we randomly sample 20 students walking in NCTU and X of them vote for "交清小徑". Suppose each student in NCTU shares her/his opinion independently and votes for "交清小徑" with probability 0.8. What is the most likely outcome of X? (Hint: Define $p_k = P(X = k)$, for $k = 0, 1, \cdots, 20$ and find the k that maximizes p_k).

   (b) (6 points) Next, we move to NTHU and conduct a similar experiment. We assume that each student in NTHU shares her/his opinion independently and votes for "交清小徑" with probability $\alpha \in [0,1]$. However, α is unknown to us, and we would like to infer the value of α from the experiment. We achieve this by considering the approach as follows:

   - We randomly sample 20 students walking in NTHU and observe that only 2 of them vote for "交清小徑". For simplicity, we use E to denote this event.
   - Define the likelihood function $L(\alpha)$ to be the probability of event E under an α.

   Please write down $L(\alpha)$ and find the α that maximizes $L(\alpha)$ (Note: Such estimate of α is called the Maximum Likelihood Estimate).