National Chiao Tung University
Department of Computer Science

Fall 2020: 1175 Probability – Final Exam

2020/12/30, 10:10AM-12:10PM

- 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 MGF/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'.
- Partial credit could be awarded for the problems.
- Good luck! May the force be with you.

1. (32 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 Z_1 and Z_2 be two independent normal random variables. Then, Z_1 and Z_2 have a joint PDF that is bivariate normal with zero correlation coefficient.

   (b) (4 points) Two real numbers U and V are selected independently and uniformly at random from the interval $(0, 1)$. Then, we know $P(U^2 + V^2 \geq 1 \text{ and } U \leq V) = 1 - \frac{\pi}{4}$.

   (c) (4 points) Let X be a continuous uniform random variable between -1 and $+3$. Then, by Markov's inequality, for any $a > 0$, we have $P(X > a) \leq E[X]/a = 1/a$.

   (d) (4 points) Consider two sequences of random variables $X_1, X_2, \cdots$ and $Y_1, Y_2, \cdots$ defined on the same sample space. Suppose that X_n converges to $a \in \mathbb{R}$ and Y_n converges to $b \in \mathbb{R}$ almost surely, respectively. Then, $X_n^2 Y_n^3$ converges to $a^2 b^3$ almost surely.

   (e) (4 points) Let X_1 and X_2 be two Bernoulli random variables with $E[X_1^2] = \frac{1}{2}$ and $E[X_2^2] = \frac{1}{2}$. Then, X_1 and X_2 must be positively correlated.

   (f) (4 points) Let $\{X_n\}_{n \in \mathbb{N}}$ be a sequence of i.i.d. continuous uniform random variables between 0 and 1. For each $n \in \mathbb{N}$, define $Y_n = \min\{X_1, X_2, \cdots, X_n\}$. Then, Y_n converges to 0 in probability, but not almost surely.

   (g) (4 points) Let $D = \{Z_i\}_{i=1}^n$ be the observed data of i.i.d. normal random variables with an unknown mean μ and a known variance σ^2. Then, under the dataset D, the maximum likelihood estimate of μ is the empirical mean $\frac{1}{n}\sum_{i=1}^n Z_i$.

   (h) (4 points) Suppose we are given a special coin with an unknown head probability $\theta \in (0, 1)$. Consider the prior distribution over θ as $P(\theta = 0.3) = 0.4$ and $P(\theta = 0.7) = 0.6$. Suppose we toss the coin for 3 times and observe $\{\text{tail, head, tail}\}$. Then, the maximum a posteriori estimate of θ is $\theta = 0.7$. (Hint: posterior $\propto$ prior $\times$ likelihood)

2. (16 points) Let X be a continuous random variable with the PDF given by

$$f_X(x) = \begin{cases} \frac{1}{4}, & \text{if } x \in (-1, 3) \\ 0, & \text{else} \end{cases}$$

   (a) (6 points) Find $M_X(t)$, i.e. the moment generating function of X.

   (b) (10 points) Use $M_X(t)$ to find $E[X]$ and $\text{Var}[X]$. (Hint: When evaluating the first-order and second-order derivative of $M_X(t)$, you may need to leverage the L'Hôpital's rule)

3. (12 points) Let $X_1, X_2, \cdots, X_n$ be a sequence of i.i.d. standard normal random variables. Let $S_n = X_1^2 + X_2^2 \cdots + X_n^2$. Please leverage the central limit theorem to find an approximation for $P(S_n \leq n + \sqrt{2n})$. (Hint: What are the mean and variance of X_1^2? Recall that the MGF of a standard normal is $\exp(\frac{t^2}{2})$. You may express your answer using the symbol $\Phi(\cdot)$)

4. (22 points) Let X be a standard normal random variable and U be a Bernoulli random variable with mean 0.5. Suppose X and U are independent. Define $Z = (2U - 1)X$:

(a) (8 points) Show that Z is also a standard normal random variable. (Hint: Write down the CDF of Z by considering the two cases $U = 0$ and $U = 1$)

(b) (8 points) Please find out $\mathrm{Cov}(X, Z)$.

(c) (6 points) Are X and Z independent? Please justify your answer.

5. (12 points) Let X be an exponential random variable with mean $= 1$. For each $n \in \mathbb{N}$, define a random variable Y_n as

$$Y_n = \begin{cases} 1, & \text{if } X > n \\ 0, & \text{else} \end{cases}$$

Show that Y_n converges to 0 in probability. (Hint: We say that Y_n converges to 0 in probability if $\lim_{n\to\infty} P(\{\omega : |Y_n(\omega) - 0| \geq \varepsilon\}) = 0$ for every $\varepsilon > 0$)

6. (16 points) Suppose we have two binary classifiers A and B for predicting the labels of the images (e.g., whether there is a dog in the photo or not). Moreover, we are given a testing dataset of N images. For each image in the dataset, A and B would output the correct label independently with probability θ_a and θ_b, respectively.

(a) (8 points) Let $X_{A,N}$ denote the number of correct labels out of the N predicted labels provided by the classifier A. Then, by using the Chebyshev's inequality, for any $\delta > 0$, could you write down an upper bound for $P(\frac{X_{A,N}}{N} \leq \theta_a - \delta)$? (Please express your answer using N, θ_a, and δ)

(b) (8 points) Similarly, let $X_{B,N}$ denote the number of correct labels out of the N predicted labels provided by the classifier B. Suppose $\theta_a > \theta_b$ and define $\Delta = \theta_a - \theta_b$. By using Chebyshev's inequality, could you provide a non-trivial lower bound for the probability of the event that A has a better empirical accuracy than B (i.e., $\frac{X_{A,N}}{N} > \frac{X_{B,N}}{N}$)? (Hint: The event $\{\frac{X_{A,N}}{N} > \theta_a - \frac{\Delta}{2}\} \cap \{\frac{X_{B,N}}{N} \leq \theta_b + \frac{\Delta}{2}\}$ is a subset of the event $\{\frac{X_{A,N}}{N} > \frac{X_{B,N}}{N}\}$. Please express your answer using N, θ_a, θ_b, and Δ)