SI 140A: Probability and Statistics for EECS
Spring 2025, Instructor: Xavier Lagorce, TA: Junye Wang, Chuyan Zhou

Homework 3

Due: 23:59 (CST), Mar 21, 2025
Full Mark: 50 points, 7 problems including 1 bonus.

Read all the instructions below carefully before you start working on the assignment, and before you make a submission.

- You are required to write down all the major steps towards making your conclusions; otherwise you may obtain limited points of the problem.
- Write your homework in English; otherwise you will get no points of this homework.
- Any form of plagiarism will lead to 0 point of this homework.
- If there are any descriptions unclear, we recommend you to ask on Piazza.

Problem 1

(Bayesian Story *5'***)**

Please reinterpret the following story from the Bayesian perspective.

The Boy Who Cried Wolf

There is a naughty boy in a village. He likes telling lies. One day he wants to make fun of the farmers. So he shouts out , "Wolf! Wolf! Wolf is coming!" The kind farmers are working on the land. They hear the yelling and hurry up to help the boy. But when they get there, the boy bursts into laughter and says: "There is not any wolf at all. I'm just joking. The farmers are very angry and go back to their land. After a while the boy shouts again, "Wolf! Wolf! Wolf is coming!" And those warm-hearted farmers rush up to the mountain , finding that they are cheated once again. The boy laughs and laughs. The farmers say to the boy angrily, "You are telling lies. We will not believe you any more." And they get back to their work,leaving the laughing boy behind.
Later a wolf really comes. The boy is very scared. He yells out to the farmers down along the mountain to seek help:"Wolf! Wolf! Wolf is coming!" the boy shouts and shouts. "Help! Help!" But no one comes. And the wolf eats the naughty boy.

狼来了：从前有个放羊娃，每天都把羊群带到山上去吃草, 山里有狼出没．第一天，放羊娃觉得无聊，想要作弄山下耕作的村民。他朝着山下大喊"狼来了!狼来了"，村民们信以为真，冲上山来准备帮助他，发现被欺骗了，大家很生气。第二天，放羊娃故技重施，村民们虽然有点迟疑，但还是冲上山来准备打狼，结果又一次发现被欺骗了，大家非常生气. 第三天，狼真的来了，此时放羊娃慌了，哭着向山下大喊"狼来了!狼来了!"，请求村民的帮助．但这一次村民们认为他又在撒谎，无人相信他。最后他所有的羊都被狼吃掉了。

Problem 2

(Fair Die, Recurrent Relations *5'***)**

A fair die is rolled repeatedly, and a running total is kept (which is, at each time, the total of all the rolls up until that time). Let p_n be the probability that the running total is ever exactly n (assume the die will always be rolled enough times so that the running total will eventually exceed n, but it may or may not ever equal n).

(a) Write down a recursive equation for p_n (relating p_n to earlier terms p_k in a simple way). Your equation should be true for all positive integers n, so give a definition of p_0 and p_k for $k < 0$ so that the recursive equation is true for small values of n.

(b) Find p_7.

(c) Give an intuitive explanation for the fact that $p_n \sim \frac{1}{3.5} = \frac{2}{7}$ as $n \to \infty$.

Problem 3

(PMF *5'+5'***)**

Let X be the number of purchases that Fred will make on the online site for a certain company (in some specified time period). Suppose that the PMF of X is $P(X = k) = \frac{e^{-\lambda}\lambda^k}{k!}$ for $k = 0, 1, 2, \ldots$. This distribution is called the *Poisson distribution* with parameter λ, and it will be studied extensively in later chapters.

(a) Find $P(X \geq 1)$ and $P(X \geq 2)$ without summing infinite series.

(b) Suppose that the company only knows about people who have made at least one purchase on their site (a user sets up an account to make a purchase, but someone who has never made a purchase there doesn't appear in the customer database). If the company computes the number of purchases for everyone in their database, then these data are draws from the conditional distribution of the number of purchases, given that at least one purchase is made. Find the conditional PMF of X given $X \geq 1$. (This conditional distribution is called a *truncated Poisson distribution.*)

Problem 4

(CDF *5'+5'***)**

Let F_1 and F_2 be CDFs, $0 < p < 1$, and $F(x) = pF_1(x) + (1-p)F_2(x)$ for all x.

(a) Show directly that F has the 3 properties of a valid CDF (see Slide 41 of Lecture 3). The distribution defined by F is called a *mixture* of the distributions defined by F_1 and F_2.

(b) Consider creating an r.v. in the following way. Flip a coin with probability p of Heads. If the coin lands Heads, generate an r.v. according to F_1; if the coin lands Tails, generate an r.v. according to F_2. Show that the r.v. obtained in this way has CDF F.

Problem 5

(Function of Random Variables *3'+3'+4'***)**

For x and y binary digits (0 or 1), let $x \oplus y$ be 0 if $x = y$ and 1 if $x \neq y$ (this operation is called exclusive or (often abbreviated to XOR), or addition mod 2).

(a) Let $X \sim Bern(p)$ and $Y \sim Bern(1/2)$, independently. What is the distribution of $X \oplus Y$?

(b) With notation as in sub-problem (a), is $X \oplus Y$ independent of X? Is $X \oplus Y$ independent of Y? Be sure to consider both the case $p = 1/2$ and the case $p \neq 1/2$.

(c) Let $X_1, ..., X_n$ be i.i.d. (i.e., independent and identically distributed) $Bern(1/2)$ random variables. For each nonempty subset J of $\{1, 2, ..., n\}$, let

$$Y_J = \bigoplus_{j \in J} X_J.$$

Show that $Y_J \sim Bern(1/2)$ and that these $2^n - 1$ random variables are pairwise independent, but not independent.

Problem 6

(Extended Monty Hall *1'+4'+5'***)**

In Monty Hall problem, now suppose the car is not placed randomly with equal probability behind the three doors. Instead, the car is behind door one with probability p_1, behind door two with probability p_2, and behind door three with probability p_3. Here $p_1+p_2+p_3 = 1$ and $p_1 \geq p_2 \geq p_3 > 0$. Recap that you are to choose one of the three doors, after which Monty will open a door he knows to conceal a goat. Monty always chooses randomly with equal probability among his options in those cases where your initial choice is correct.

(a) Recap what is covered in the lecture. What strategy would you follow if $p_1 = p_2 = p_3$?

(b) What strategy should you follow if the assumption in a) does not exist? You should answer in a "Choose which, then switch or stay" form.

Consider the following generalized version of the Monty Hall problem. There are $n \geq 3$ doors, behind one of which there is a car (which you want), and behind the rest of which there are goats (which you don't want). Initially, all possibilities are equally likely for where the car is. You choose a door. Monty then opens m goat doors, and offers you the option of switching to any of the remaining $n - m - 1$ doors.

(c) Assume that Monty knows which door has the car, will always open m goat doors and offer the option of switching, and that Monty chooses with equal probabilities from all his choices of which goat doors to open. Should you switch? What is your probability of success if you switch to one of the remaining $n - m - 1$ doors?

(Bonus: Tool for problems with recursive structures *5'***)**

Problem 7

By LOTP for problems with recursive structure, we generate many difference equations. To solve the difference equation in the form of

$$f_{i+1} = b \cdot f_i + a \cdot f_{i-1}, \quad i \geq 1.$$

where a and b are constants, we turn to the so-called characteristic equation:

$$x^2 = bx + a.$$

If such an equation has two distinct roots r_1 and r_2, then the general form of f_i is

$$f_i = c \cdot r_1^i + d \cdot r_2^i.$$

If there is only one distinct root r, then the general form of f_i is

$$f_i = c \cdot r^i + d \cdot i \cdot r^i.$$

Show the mathematical principle behind the method of characteristic equation.