

# Department of Systems Innovation

2021 Online Written Examination:  
Problems Designed to Test Ability of Logical Thinking  
Field 2: Artificial intelligence, information science, and mathematical models

---

## Question 2-1

---

You have  $N$  mails in your mailbox. Each mail has its priority for replying, which is denoted by  $x \in (0, 1)$ . At each time step, you either choose the mail which has the highest priority with probability  $p$  or randomly choose a mail with probability  $1 - p$ . You reply to the mail you have chosen and delete it from your mailbox. After replying, you receive a new mail whose priority  $x$  is drawn from a uniform distribution on the interval  $(0, 1)$ .

- (1) First, let  $N = 2$  and  $p = 1$ . Second, let **A** and **B** denote two mails you initially have and let  $x_A$  and  $x_B$  be the priorities of these mails, respectively, in the ascending order (i.e.  $x_A < x_B$ ). Find the probability if the duration (the number of time steps) until you reply to the mail **A** is  $t$ .
  - (2) Next, let  $N \gg 1$  and  $p = 1/2$ . Consider the case that the above process (i.e. receiving and replying to mails) is repeated many times. Let  $f(x)$  be the distribution function of the priorities  $x$  of the mails in your mail box and let  $g(t)$  be the distribution function of the duration  $t$  (i.e. the number of time steps needed to reply to each mail). Then, describe the characteristic features of these two distribution functions  $f(x)$  and  $g(t)$ , and sketch their curves.
- 

## Question 2-2

---

There are  $m + n$  data points  $\{\mathbf{x}_1, \dots, \mathbf{x}_{m+n}\}$ , which consist of  $d$ -dimensional real values ( $d \geq 3$ ). Note that  $\{\mathbf{x}_1, \dots, \mathbf{x}_m\}$  belong to the class  $C_1$ , and  $\{\mathbf{x}_{m+1}, \dots, \mathbf{x}_{m+n}\}$  belong to the other class  $C_2$ .

Assuming  $m \gg 1$  and  $n \gg 1$ , answer the following questions regarding the method for determining the class of a new data point  $\mathbf{x}_0$ . If needed, you can draw figures.

- (1) Give two or more examples of the methods for determining the class of data points and describe the features of the methods.
- (2) Choose one method from your examples given in (1) and describe its algorithm using such as equations, pseudo-codes\*, or flow-charts.
- (3) Explain how to evaluate the accuracy of the method described in (2).

\*Pseudo-code is an informal description of the operating principle of a computer program or algorithm. Below is an example pseudo-code to give an order to study unless being sleepy.

```
if isSleepy == true:  
    Go to bed  
else:  
    Study
```