Applied Probability for Computer Science [CM0546]
Master's Degree in Computer Science, a.y. 2022/23

Lecturer: Isadora Antoniano-Villalobos

Mock Exam **With Solutions**

Last Name: ______________________ First Name: ______________________

ID (Matr.): ______________________ Signature: ______________________

READ ALL INSTRUCTIONS CAREFULLY AND SIGN YOUR AGREEMENT.

Exams with missing or incomprehensible name, ID and/or signature will not be marked

Only answers on the question sheets will be graded. Use the spaces provided, continuing on the last page if required (clearly indicate it). All sheets (including all scrap paper, which however will not be graded) must be turned in.

An outline of the procedure and calculations is required for each question, including the use of R functions, with the exception of multiple choice questions. Use at least a 4 decimal precision for all calculations. Other than correctness of response and justification, **order and clarity will also be taken into account when marking.**

This is an **open book** exam: you may use R, the textbook, your notes and other teaching material, but obtaining targeted help from dedicated websites, apps, each other, or any third parties is estrictly forbiden and will be severely penalized.

This exam must be solved in **90 minutes**. It is composed of **6 questions**, adding up to **30 points**(depending on your team assignment and quiz results, you could loose some points and still achieve full marks)

Question:	1	2	3	4	5	6	Total
Points:	5	5	5	5	5	5	30
Score:							

Question 1 (5 points)
Consider two discrete random variables, X and Y, such that:

x	$P_X(x)$
0	0.2
1	0.5
2	0.3

	$y=1$	$y=2$
$P_{Y\|X}(y\|0)$	0.5	0.5
$P_{Y\|X}(y\|1)$	0.1	0.9
$P_{Y\|X}(y\|2)$	0.4	0.6

(a) Find the joint probability mass function of X and Y.

Solution: For each $i = 0, 1, 2$ and $j = 1, 2$,

$$\mathbb{P}[X = i, Y = j] = \mathbb{P}[Y = j|X = i]\mathbb{P}[X = i].$$

Therefore, the required joint pmf is given in the following table

$X \setminus Y$	1	2
0	0.1	0.1
1	0.05	0.45
2	0.12	0.18

(b) Find the marginal pmf of Y.

Solution: The marginal distribution for Y is found by adding the joint over all possible values of $X = x$, therefore,

y	1	2
$P_Y(y)$	0.27	0.73

(c) Find $\mathbb{P}[X < Y]$

Solution:

$$\mathbb{P}[X < Y] = \mathbb{P}[X = 0, Y = 1] + \mathbb{P}[X = 0, Y = 2] + \mathbb{P}[X = 1, Y = 2] = 0.65.$$

(d) Are the two variables independent? Justify your answer.

Solution: The distribution of $Y|X = x$ changes with x, thus the variables are not independent.

Question 2 (5 points)
The customers of an internet service provider connect to the network according to a Poisson process with a rate of 8 customers per minute. As part of a promotion, the provider offers special discounts to each connecting customer with probability 0.2. Compute:

(a) The probability that no offer is made during 2 consecutive minutes

Solution: Let $N = \{N(t) : t \geq 0\} \sim$PP$(\lambda = 8)$ be the arrivals process and let $\tilde{N}(t)$ be the number of discounts offered by time t. Then the process $\tilde{N} = \{\tilde{N}(t) : t \geq 0\}$ is obtained by thinning N with probability $p = 0.2$. Therefore, $\tilde{N} \sim$PP$(\lambda p = 1.6)$ and the time $\tilde{X}$ between two consecutive discount offers follows an exponential distribution with parameter 1.6. So the probability that no offer is made during 2 consecutive minutes is given by

$$\mathbb{P}[\tilde{X} > 2] = 1 - \texttt{pexp}(2, 1.6) = 0.0408$$

(b) The expected time (in minutes) between consecutive offers

Solution: The expected time (in minutes) between consecutive offers is

$$\mathbb{E}[\tilde{X}] = 1/1.6 = 0.6250$$

(c) The probability that the first customer to connect after the activation of the promotion is offered a discount

Solution: The probability that the first customer to connect after the activation of the promotion is offered a discount is simply $p = 0.2$, since that is the probability with which any connected client is offered a discount.

(d) The probability that 6 customers were offered the discount in the first minute, given that 12 customers were offered the discount in the first 2 minutes.

Solution: For any $0 < s < t$, we can write $\tilde{N}(t) = [\tilde{N}(t) - \tilde{N}(s)] + \tilde{N}(s)$, where $[\tilde{N}(t) - \tilde{N}(s)]$ ~Poisson$(1.6\,(t-s))$ is independent of $\tilde{N}(s)$ ~Poisson$(1.6\,s)$. So by the properties of the Poisson distribution,

$$[\tilde{N}(t) - \tilde{N}(s) | \tilde{N}(t) = k] \sim \text{Binom}(k, 1.6\,(t-s)/1.6\,t).$$

Letting $t = 2$, $s = 1$ and $k = 12$, the probability that 12 customers were offered the discount in the first 2 minutes, given that 6 customers were offered the discount in the first minute is

$$\mathbb{P}[\tilde{N}(t) - \tilde{N}(s)] = 6 | \tilde{N}(t) = 12] = \texttt{dbinom}(6, 12, 1/2) = 0.2256$$

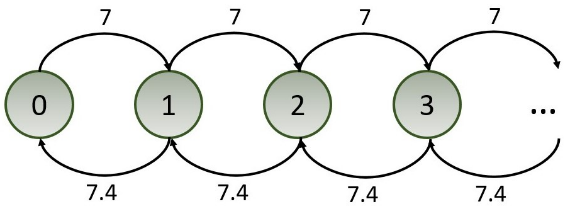
.

Question 3 (5 points)

Small aircrafts arrive at an airport according to a Poisson process with rate $a = 7$ per hour. The time that each aircraft stays at the airport is exponentially distributed with rate $d = 7.4$. Let $N(t)$ denote the number of aircrafts at the airport at time t.

(a) Find G, the generator of the process $N = \{N(t), t \geq 0\}$ and sketch a graph that represents it.

Solution: The state space for N is $\{0, 1, 2, \ldots\}$ and the process can be represented by the following graph:



Therefore, the generator for N is given by

$$\mathbf{G} = \begin{bmatrix} -7 & 7 & 0 & 0 & 0 & \dots \\ 7.4 & -14.4 & 7 & 0 & 0 & \dots \\ 0 & 7.4 & -14.4 & 7 & 0 & \dots \\ \vdots & \vdots & \vdots & \vdots & \vdots & \ddots \end{bmatrix}$$

(b) In the long run, what is the probabilty that the airport has only one small aircraft at any given moment?

HINT: if $0 < |r| < 1$, then $\sum_{n=0}^{\infty} r^n = \frac{1}{1-r}$

Solution: We must find the stationary distribution of the process (if it exists!).

We recognize N as a birth-death process with constant birth rate $\lambda_i = 7$ for $i \geq 0$ and constant death rate $\mu_i = 7.4$ for $i \geq 1$ and $\mu_0 = 0$

We notice that

$$1 + \sum_{i=1}^{\infty} \frac{\lambda_0 \cdots \lambda_{i-1}}{\mu_1 \cdots \mu_i} = \sum_{i=0}^{\infty} \left(\frac{7}{7.4}\right)^i \sum_{i=0}^{\infty} 0.9459^{\text{``}i} = \frac{1}{1 - 0.9459} < \infty$$

so, as seen in class, the stationary distribution exists and the required probability is $\pi_1 = \lambda_0 \pi_0 / \mu_1 = 7(1 - 0.9459)/7.4 = 0.0512$

(c) Do you consider this an adequate model for a small airport? Justify.

Solution: Not really. The model considers an infinite state space, which may be reasonable for a large airport, but not for a small one, for which we could assume a limited capacity to receive aircrafts, no matter how small they are.

Question 4 (5 points)

Consider a continuous random variable X with the following cumulative distribution function:

$$F(x) = \begin{cases} 0 & \text{if } x \leq 0, \\ cx & \text{if } 0 < x \leq 1, \\ x + 1 - c(x+4) & \text{if } 1 < x \leq 2, \\ 1 & \text{otherwise.} \end{cases}$$

(a) Find the value of c that makes F a continuous cumulative distribution function (cdf).

Solution: In order to make F continuous at $x = 2$ we need

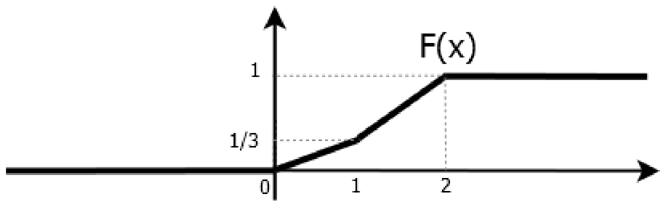
$$1 = F(2) = 2 + 1 - c(2+4) = 3 - 6c \Leftrightarrow c = \frac{1}{3}.$$

If $c = 1/3$ and we can verify that F is continuous also at $x = 1$ because

$$cx = \frac{1}{3} \qquad \text{and} \qquad x + 1 - c(x+4) = 1 + 1 - \frac{1}{3}(1+4) = \frac{1}{3}.$$

Notice that the resulting F is a non-decreasing function, as required:

$$F(x) = \begin{cases} 0 & \text{if } x \leq 0, \\ x/3 & \text{if } 0 < x \leq 1, \\ 2x/3 - 1/3 & \text{if } 1 < x \leq 2, \\ 1 & \text{otherwise} \end{cases}$$



(b) X is used as a quality index of a production line for hardware components. The number of defective components produced in a day is considered to be Poisson with parameter $\lambda_1 = 1$ if $0 < X \leq 1$ and $\lambda_2 = 2$ if $1 < X \leq 2$. If the number of defective components produced yesterday was 2 what is the probability that the quality index value was greater than 1?

Solution: We know that $\mathbb{P}[0 < X \leq 1] = \mathbb{P}[X \leq 1] = F(1) = 1/3$ and $\mathbb{P}[1 < X \leq 2] = \mathbb{P}[X > 1] = 1 - F(1) = 2/3$,since the support of X is $(0, 2]$.
Let N be the number of defective components produced in a day, then

$$\mathbb{P}[N = 2|0 < X \leq 1] = \mathbb{P}[N = 2|X \leq 1] = \texttt{dpois}(2, 1) = 0.1839,$$

$$\mathbb{P}[N = 2|1 < X \leq 2] = \mathbb{P}[N = 2|X > 1] = \texttt{dpois}(2, 2) = 0.2707,$$

Therefore, we can find the required probability by Bayes rule

$$\begin{aligned}\mathbb{P}[X > 1|N = 2] &= \frac{\mathbb{P}[N = 2|X > 1]\mathbb{P}[X > 1]}{\mathbb{P}[N = 2|X \leq 1]\mathbb{P}[X \leq 1] + \mathbb{P}[N = 2|X > 1]\mathbb{P}[X > 1]} \\ &= \frac{2(0.2707)/3}{0.1839/3 + 2(0.2707)/3} = 0.7464\end{aligned}$$

Question 5 (5 points)
Write the formal statement of the Central Limit Theorem, and provide an interpretation in your own words in terms of one or more Markov processes.

Solution: For the formal statement of the CLT refer to one of the course textbooks. Regarding the interpretation, the correct answer is not unique, but it should somehow make a reference to either the sum or the mean of n i.i.d. random variables, both of which are discrete-time Markov processes.

Question 6 (5 points)
This question consists of 5 single choice queries. Read carefuly and make your single selected response is indicated clearly. Justifications are not required and will not be marked if provided (so don't waste your time).

(a) Let X_1 and X_2 be two independent random variables with common mean 0 and variances σ_1^2 and σ_2^2, respectively and define $Y_1 = X_1 + X_2$, $Y_2 = X_1 - X_2$. Which of the following statements is surely true?

i) $\mathbb{E}[Y_1] > \mathbb{E}[Y_2]$ and $\text{Var}[Y_1] > \text{Var}[Y_2]$

ii) $Y_1 > Y_2$

iii) $\mathbb{E}[Y_1] = \mathbb{E}[Y_2] = 0$ and $\text{Var}[Y_1] > \text{Var}[Y_2]$

iv) $\text{Cov}[Y_1, Y_2] = \sigma_1^2 - \sigma_2^2$

v) $\text{Cov}[Y_1, Y_2] \leq 0$

Solution: *iv)*

(b) Let X and Y be two random variables taking values in $(0, \infty)$ and let $W = X + Y$. If $\mathbb{E}[W] = 80$, which of the following statements is certainly true?

i) $\max\{\mathbb{E}[X], \mathbb{E}[Y]\} < 80$ and $\min\{\mathbb{E}[X], \mathbb{E}[Y]\} > 0$

ii) $\max\{X, Y\} < 80$ and $\min\{X, Y\} > 0$

iii) None of the other options is correct

iv) $\max\{\mathbb{E}[X], \mathbb{E}[Y]\} > 80$

v) $\mathbb{E}[X] = \mathbb{E}[Y] = 40$

Solution: *i)*

(c) Which of the following could be the (conditional) transition probability matrix for a HCTMC?

i) All of them

ii) None of them

iii)

$$\tilde{P} = \begin{pmatrix} 0 & 1/3 & 0 & 0 \\ 0 & 0 & 1/2 & 0 \\ 0 & 2/3 & 0 & 0 \\ 0 & 0 & 1/2 & 0 \end{pmatrix}$$

iv)

$$\tilde{P} = \begin{pmatrix} -1 & 1 & 0 & 0 \\ 1/3 & -1 & 2/3 & 0 \\ 0 & 1/2 & -1 & 1/2 \\ 1 & 0 & 0 & -1 \end{pmatrix}$$

v)

$$\tilde{P} = \begin{pmatrix} 0 & 1/2 & 1/2 & 0 \\ 1/3 & 0 & 2/3 & 0 \\ 0 & 1/2 & 0 & 1/2 \end{pmatrix}$$

Solution: *ii)*

(d) If A and B are two disjoint events, which of the following statements must be true?

i) A and B are independent

ii) $\mathbb{P}[A \cap B] = 0$

iii) $\mathbb{P}[A] = \mathbb{P}[B]$

iv) $\mathbb{P}[A \cup B] = 1/2$

v) $\text{Cov}[A, B] = 0$

Solution: *ii)*

(e) Consider a CTMC $X = \{X(t) : t \geq 0\}$ with state-space $S = \{1, 2\}$ and assume that the infinitesimal generator for the process is provided in the form of a matrix G in R. Which of the following blocks of code can be used to simulate and plot a path of the process for the first 5 transitions, if the initial state is 1?

i)
```
t<-0
x<-1
for(j in 1:5)t<-c(t,rexp(1,G[x,x]))
plot(cumsum(t),c(2,1,2,1,2,1),type="s")
```

ii)
```
t<-0
x<-1
for(j in 1:5)t<-c(t,rexp(1,-G[x,x]))
plot(t,c(1,2,1,2,1,2),type="s")
```

iii)
```
t<-0
x<-1
for(j in 1:5)t<-c(t,t[j]+rexp(1,-G[x,x]))
plot(t,c(1,2,1,2,1,2),type="s")
```

iv)
```
t<-0
x<-2
for(j in 1:5)t<-c(t,t[j]+rexp(1,-G[x,x]))
plot(t,c(1,2,1,2,1,2),type="s")
```

v) None of the above

Solution: *iii)*