

# MATH-232: Midterm 2022 - Solutions A

April 12th, 2022

**Exercise 1.** Let  $(X, Y)$  be a joint random variable whose probability mass function is given by the following table:

		Y	
	X	-1	1
0		0.2	0.4
1		0.3	0.1

1. What are the marginal laws of  $X$  and  $Y$ ?
2. Are  $X$  and  $Y$  independent? (As always, justify your answer.)
3. Compute  $\mathbb{E}[X]$ ,  $\mathbb{E}[Y]$ ,  $\mathbb{E}[X + Y]$  and  $\mathbb{E}[XY]$ .

*Solution:* (4 points)

1. (1pt)

$$f_Y(y) = 0.5, \quad y = -1, +1; \quad (1)$$

$$f_X(x) = \begin{cases} 0.6 & \text{if } x = 0, \\ 0.4 & \text{if } x = 1 \end{cases} \quad (2)$$

2. (0.5pt) No. For instance,

$$f_{X,Y}(0, -1) = 0.2 \quad (3)$$

$$f_X(0) \cdot f_Y(-1) = 0.6 \cdot 0.5 = 0.3. \quad (4)$$

Since,  $f_{X,Y}(0, -1) \neq f_X(0) \cdot f_Y(-1)$ , it follows that  $X$  and  $Y$  are not independent.

3. (2.5pt)

$$\mathbb{E}[X] = 0.4 \cdot 1 = 0.4 \quad (5)$$

$$\mathbb{E}[Y] = 0.5 \cdot 1 + 0.5 \cdot -1 = 0 \quad (6)$$

$$\mathbb{E}[X + Y] = \mathbb{E}[X] + \mathbb{E}[Y] = 0.4 \quad (7)$$

$$\mathbb{E}[X \cdot Y] = -1 \cdot 0.3 + 1 \cdot 0.1 = -0.2. \quad (8)$$

□

**Exercise 2.** Let  $f : \mathbb{R} \rightarrow \mathbb{R}$  be the function defined by the formula

$$f(x) = \begin{cases} 4x \exp(-2x) & \text{if } x \geq 0, \\ 0 & \text{if } x < 0. \end{cases}$$

1. Justify that  $f$  is a probability density.

2. What is the expectation of a random variable  $X$  with density  $f$ ?

*Solution.* (2.5 points)

1. (1.5pt)  $f(x) \geq 0$  for all  $x$ . One can check by integration by part that  $\int_{-\infty}^{\infty} f(x)dx = 1$ .

2. (1pt)

$$\mathbb{E}[X] = \int_0^{\infty} 4x^2 \exp(-2x)dx \quad (9)$$

$$= -2x^2 \exp(-2x) \Big|_0^{\infty} + \int_0^{\infty} 4x \exp(-2x)dx \quad (10)$$

$$= 1. \quad (11)$$

□

**Exercise 3.** Let  $F : \mathbb{R} \rightarrow [0, 1]$  be the function defined by

$$F(x) = \begin{cases} 0 & \text{if } x < 0, \\ x^2 & \text{if } 0 \leq x \leq 1, \\ 1 & \text{if } x > 1. \end{cases}$$

1. Justify that  $F$  is a cumulative distribution function.
2. Justify that the law associated to  $F$  is continuous.
3. Compute the density  $f$  associated to this law.
4. Let  $X$  be a random variable with cumulative distribution function  $F$ .
  - (a) Compute  $\mathbb{P}(X > 1/2)$ .
  - (b) Compute  $\mathbb{E}[X]$ .
  - (c) Compute the variance of  $X$ .

*Solution.* (5 points)

1. (1pt) One can verify that  $F$  is such that

- $\lim_{x \rightarrow -\infty} F(x) = 0$ ;
- $\lim_{x \rightarrow \infty} F(x) = 1$ ;
- $F$  is non-decreasing;
- $F$  is continuous on the right.

Alternatively, one can check that  $f(x) = \frac{d}{dx} F(x)$  is a valid density function.

2. (0.5pt) Since  $F(x)$  is continuous, it follows that the law associated to  $F$  is continuous.
3. (1pt)

$$f(x) = \begin{cases} 0 & \text{if } x < 0, \\ 2x & \text{if } 0 \leq x \leq 1, \\ 0 & \text{if } x > 1. \end{cases} \quad (12)$$

4. (2.5pt) One can compute that:

- a) (0.5pt)  $\mathbb{P}(X > 1/2) = 3/4$
- b) (1pt)  $\mathbb{E}[X] = 2/3$
- c) (1pt)  $\text{Var}(X) = \mathbb{E}[X^2] - \mathbb{E}[X]^2 = 1/2 - 4/9 = 1/18$ .

□

**Exercise 4.** At the end of a video game, the player must kill a monster sampled randomly as follows:

- 1 time out of 10, it is a *dragon*,
- 3 times out of 10, it is a *troll*,
- the rest of the time, it is a *giant*.

When the monster dies, the player gets a chance to get a *ruby*:

- the dragon always gives a ruby,
- the troll gives a ruby 1 time out of 2,
- the giant gives a ruby 1 time out of 5.

We assume that the game is very easy; thus the players always succeed at killing the monster. Two different games are independent.

Alice plays the game. We denote  $D$  (respectively  $T$ ,  $G$ ) the event “the monster is a dragon” (respectively a troll, a giant). We denote  $R$  the event “Alice wins a ruby”, and  $p = \mathbb{P}(R)$ .

1. Compute  $p$ .
2. Alice won a ruby! What is the probability that the monster was a troll?
3. Bob decides to play 6 games. We denote  $S$  the number of rubies that he gets.
  - (a) What is the law of  $S$ ?
  - (b) What is the expectation of  $S$ ?
  - (c) What is the probability that Bob wins exactly 3 rubies?
  - (d) What is the probability that Bob wins at least 1 ruby?
4. Charlie decides to play until he wins one ruby. We denote  $X$  the number of games Charlie plays.
  - (a) What is the law of  $X$ ?
  - (b) What is the expectation of  $X$ ?
  - (c) What is the probability that Charlie does exactly 3 games?

*Solution.* (8 points)

1. (1.5pt) Using the law of total probability we write,

$$\begin{aligned}
 p = \mathbb{P}(R) &= \mathbb{P}(R|D)\mathbb{P}(D) + \mathbb{P}(R|T)\mathbb{P}(T) + \mathbb{P}(R|G)\mathbb{P}(G) \\
 &= 1 \cdot \frac{1}{10} + \frac{1}{2} \cdot \frac{3}{10} + \frac{1}{5} \cdot \frac{6}{10} \\
 &= \frac{1}{10} \cdot \frac{10 + 15 + 12}{10} \\
 &= \frac{37}{100}.
 \end{aligned}$$

2. (1pt) Using Bayes rule,

$$\mathbb{P}(T|R) = \frac{\mathbb{P}(R|T)\mathbb{P}(T)}{p} = \frac{1}{2} \cdot \frac{3}{10} \cdot \frac{100}{37} = \frac{15}{37}.$$

3. (3pt)

- a) (1pt) Since  $S$  is the sum of 6 independent events, with probability of success  $p$ , it follows that  $S \sim \text{Binomial}(p, 6)$

- b) (0.5pt)  $\mathbb{E}[S] = 6p$
- c) (1pt)  $\mathbb{P}(S = 3) = \binom{6}{3}p^3(1-p)^3$
- d) (0.5pt)  $\mathbb{P}(S \geq 1) = 1 - \mathbb{P}(S = 0) = 1 - (1-p)^6$

4. (2.5pt)

- a) (1pt) Since  $X$  denotes the number of trials to get the first success of independent events of probability  $p$ , it follows that  $X \sim \text{Geom}(p)$
- b) (0.5pt)  $\mathbb{E}[X] = 1/p = 100/37$
- c) (1pt)  $\mathbb{P}(X = 3) = (1-p)^2p$ .

□

**Exercise 5.** *This exercise is significantly harder than the previous ones, but not worth many points. Try it once you have finished the rest of the midterm.*

We remind that:

- an exponential random variable with rate  $\lambda$  has density  $f(x) = \lambda e^{-\lambda x} I(x > 0)$ ,
- a gamma random variable with shape parameter  $\alpha$  and rate  $\lambda$  has density

$$f(x) = \frac{\lambda^\alpha}{\Gamma(\alpha)} x^{\alpha-1} e^{-\lambda x} I(x > 0),$$

- if  $n$  is a non-negative integer,  $\Gamma(n+1) = n!$ , and
- a Poisson random variable with rate  $\lambda$  has probability mass function  $f(n) = e^{-\lambda} \lambda^n / n!$ ,  $n = 0, 1, \dots$

Let  $X_1, X_2, X_3, \dots$  be an infinite sequence of independent random variables of exponential law with rate parameter 1. We denote  $T_0 = 0$  and for all  $n \geq 1$ ,  $T_n = X_1 + \dots + X_n$ . For all  $t \geq 0$ , we denote  $N_t = \max\{n \geq 0 \mid T_n \leq t\}$ .

1. (a) Compute the cumulative distribution function of  $T_2$ . Deduce that  $T_2$  is a gamma random variable with shape parameter 2 and rate 1. (Like everywhere else, a proof is required.)  
 (b) Let  $n \geq 1$ . Using the same method, compute the law of  $T_n$ .
2. Let  $t > 0$ . Compute the law of  $N_t$ .
3. Let  $n \geq 1$ .  
 (a) Compute the joint law of  $(T_1, \dots, T_n)$ .  
 (b) Compute the conditional law of  $(T_1, \dots, T_n)$  given that  $N_t = n$ .

*Solution.* (5 points)

1. (2pt)

- a) (1pt) The CDF of  $T_2$  can be computed as follows:

$$F_{T_2}(t) = \mathbb{P}(X_1 + X_2 \leq t) = \int_0^t dx_1 \int_0^{t-x_1} dx_2 f_{(X_1, X_2)}(x_1, x_2) \quad (13)$$

$$\stackrel{(\text{indep.})}{=} \int_0^t dx_1 \int_0^{t-x_1} dx_2 f_{X_1}(x_1) f_{X_2}(x_2) \quad (14)$$

$$= \int_0^t dx_1 e^{-x_1} \int_0^{t-x_1} dx_2 e^{-x_2} \quad (15)$$

$$= \int_0^t dx_1 e^{-x_1} (1 - e^{-(t-x_1)}) \quad (16)$$

$$= \int_0^t dx_1 (e^{-x_1} - e^{-t}) \quad (17)$$

$$= 1 - e^{-t}(t+1). \quad (18)$$

Deriving the above with respect to  $t$ , we obtain

$$\frac{d}{dt}F_{T_2}(t) = -e^{-t} + e^{-t}(t+1) = te^{-t} \quad (19)$$

that is the pdf of a Gamma(2,1) random variable.

- b) (1pt) We show by induction that  $T_n \sim \text{Gamma}(n, 1)$ . The base case is given by point a). The inductive step is as follows: assume that  $T_n \sim \text{Gamma}(n, 1)$ , then

$$F_{T_{n+1}}(t) = \mathbb{P}(T_n + X_{n+1} \leq t) = \int_0^t dy \int_0^{t-y} dx f_{(T_n, X_{n+1})}(y, x) \quad (20)$$

$$= \int_0^t dy \int_0^{t-y} dx f_{T_n}(y) f_{X_{n+1}}(x) \quad (21)$$

$$= \int_0^t dy \frac{1}{(n-1)!} y^{n-1} e^{-y} \int_0^{t-y} dx e^{-x} \quad (22)$$

$$= \int_0^t dy \frac{1}{(n-1)!} y^{n-1} e^{-y} (1 - e^{-(t-y)}) \quad (23)$$

$$= \int_0^t dy \frac{1}{(n-1)!} y^{n-1} e^{-y} - e^{-t} \int_0^t dy \frac{1}{(n-1)!} y^{n-1}. \quad (24)$$

Deriving the above with respect to  $t$ , we obtain

$$\frac{d}{dt}F_{T_{n+1}}(t) = \frac{1}{(n-1)!} t^{n-1} e^{-t} - e^{-t} \frac{1}{(n-1)!} t^{n-1} + e^{-t} \int_0^t dy \frac{1}{(n-1)!} y^{n-1} \quad (25)$$

$$= \frac{1}{n!} t^n e^{-t}, \quad (26)$$

that is the pdf of a Gamma distribution with shape parameter  $n+1$  and scale parameter 1.

2. (2pt) For all  $n \in \mathbb{N}$ , we have  $\{N_t = n\} = \{T_n \leq t, T_{n+1} > t\}$ , since the sequence  $(T_n)_{n \geq 0}$  is increasing. Note that  $T_n$  and  $T_{n+1}$  are not independent, since they both depend on  $X_1, \dots, X_n$ . However,  $T_n$  and  $X_{n+1}$  are independent. Thus, for  $n \geq 1$ , we get

$$\mathbb{P}(N_t = n) = \mathbb{P}(T_n \leq t, T_n + X_{n+1} > t) \quad (27)$$

$$= \mathbb{P}(T_n \leq t, X_{n+1} > t - T_n) \quad (28)$$

$$= \int_0^t dy \int_{t-y}^{\infty} dx f_{(T_n, X_{n+1})}(y, x) \quad (29)$$

$$= \int_0^t dy f_{T_n}(y) \int_{t-y}^{\infty} dx f_{X_{n+1}}(x) \quad (30)$$

$$= \int_0^t dy \frac{1}{(n-1)!} y^{n-1} e^{-y} \int_{t-y}^{\infty} dx e^{-x} \quad (31)$$

$$= \int_0^t dy \frac{1}{(n-1)!} y^{n-1} e^{-y} e^{-(t-y)} \quad (32)$$

$$= e^{-t} \int_0^t \frac{1}{(n-1)!} y^{n-1} = \frac{1}{n!} t^n e^{-t}, \quad (33)$$

thus  $N_t \sim \text{Poisson}(t)$ .

3. (1pt)

- a) (0.5pt) First, consider the case where  $n = 2$ .

$$f_{(T_1, T_2)}(t_1, t_2) = f_{(X_1, X_1 + X_2)}(t_1, t_2) \quad (34)$$

$$= f_{X_1}(t_1) f_{X_1 + X_2 | X_1}(t_2 | t_1) \quad (35)$$

$$= f_{X_1}(t_1) f_{X_2 | X_1}(t_2 - t_1 | t_1). \quad (36)$$

Since  $X_1$  and  $X_2$  are independent we can write  $f_{X_2|X_1}(t_2 - t_1|t_1) = f_{X_2}(t_2 - t_1)$ , and plug in the density functions of  $X_1$  and  $X_2$

$$f_{(T_1, T_2)}(t_1, t_2) = f_{X_1}(t_1)f_{X_2}(t_2 - t_1) \quad (37)$$

$$= e^{-t_1}\mathbf{1}(t_1 > 0)e^{-(t_2-t_1)}\mathbf{1}(t_2 > t_1) \quad (38)$$

$$= e^{-t_2}\mathbf{1}(0 < t_1 < t_2). \quad (39)$$

We now prove by induction that for general  $n$ ,

$$f_{(T_1, \dots, T_n)}(t_1, \dots, t_n) = e^{-t_n}\mathbf{1}(0 < t_1 < t_2 < \dots < t_n) \quad (40)$$

The base case is given by the above. For the inductive step, assume that (40) holds for  $n$ . Then,

$$f_{(T_1, \dots, T_n, T_{n+1})}(t_1, \dots, t_n, t_{n+1}) = f_{(T_1, \dots, T_n, T_n + X_{n+1})}(t_1, \dots, t_n, t_{n+1}) \quad (41)$$

$$= f_{(T_1, \dots, T_n)}(t_1, \dots, t_n)f_{T_n + X_{n+1}|(T_1, \dots, T_n)}(t_{n+1}|t_1, \dots, t_n) \quad (42)$$

$$= f_{(T_1, \dots, T_n)}(t_1, \dots, t_n)f_{X_{n+1}|(T_1, \dots, T_n)}(t_{n+1} - t_n|t_1, \dots, t_n). \quad (43)$$

Observe one more time that  $X_{n+1}$  is independent from  $T_1, \dots, T_n$ , which gives

$$f_{(T_1, \dots, T_n, T_{n+1})}(t_1, \dots, t_n, t_{n+1}) = f_{(T_1, \dots, T_n)}(t_1, \dots, t_n)f_{X_{n+1}}(t_{n+1} - t_n) \quad (44)$$

$$= e^{-t_n}\mathbf{1}(0 < t_1 < t_2 < \dots < t_n)e^{-(t_{n+1}-t_n)}\mathbf{1}(t_{n+1} > t_n) \quad (45)$$

$$= e^{-t_{n+1}}\mathbf{1}(0 < t_1 < t_2 < \dots < t_{n+1}), \quad (46)$$

which concludes the proof.

- b) (0.5pt) Recall that  $\{N_t = n\} = \{T_n \leq t, T_{n+1} > t\}$ . We first compute the joint distribution of  $(T_1, \dots, T_n, N_t)$  :

$$f_{(T_1, \dots, T_n, N_t)}(t_1, \dots, t_n, n) = \int_0^\infty dt_{n+1} f_{(T_1, \dots, T_n, T_{n+1}, N_t)}(t_1, \dots, t_n, t_{n+1}, n) \quad (47)$$

$$= \int_0^\infty dt_{n+1} f_{(T_1, \dots, T_n, T_{n+1})}(t_1, \dots, t_n, t_{n+1})\mathbf{1}(t_n \leq t, t_{n+1} > t) \quad (48)$$

$$= \int_0^\infty dt_{n+1} e^{-t_{n+1}}\mathbf{1}(0 < t_1 < \dots < t_{n+1})\mathbf{1}(t_n \leq t, t_{n+1} > t) \quad (49)$$

$$= \int_t^\infty dt_{n+1} e^{-t_{n+1}}\mathbf{1}(0 < t_1 < \dots < t_n \leq t) \quad (50)$$

$$= e^{-t}\mathbf{1}(0 < t_1 < t_2 < \dots < t_n \leq t). \quad (51)$$

Applying the formula for the conditional probability we get

$$\begin{aligned} f_{(T_1, \dots, T_n)|N_t}(t_1, \dots, t_n|n) &= \frac{f_{(T_1, \dots, T_n, N_t)}(t_1, \dots, t_n, n)}{\mathbb{P}(N_t = n)} \\ &= \frac{e^{-t}\mathbf{1}(0 < t_1 < t_2 < \dots < t_n \leq t)}{e^{-t}t^n/n!} \\ &= n!t^{-n}\mathbf{1}(0 < t_1 < t_2 < \dots < t_n \leq t). \end{aligned}$$

□