MIDTERM

An electronic calculator, the distributions tables as well as a sheet of personnal notes are allowed for the exam. Any kind of dictionary are permitted.

Exercise 1

1. *Poker (again!) : in a 52 cards poker game calculate the probabilities of obtaining :*

- **a.** *A pair (two cards of the same kind and three cards with different kinds).*
- **b.** *A full house (three cards of a kind and two cards of another kind).*
- **c.** *A flush (5 cards of the same suit). To be more precise you can exclude the Straight Flush which is 5 cards of the same suit forming a consecutive sequence.*
- **d.** *A Staight (5 cards forming a consecutive sequence). If you want to be precise you should exclude the Straight Flush.*

2. *Seating (again!) : In how many ways cand 8 people be seated in a row if*

- **a.** *there are no restrictions on the seating arrangement.*
- **b.** *persons A and B must sit next to each other.*
- **c.** *there are 4 men and 4 women and no 2 men or 2 women can sit next to each other.*
- **d.** *there are 5 men and they must sit next to each other.*

Exercise 2

The goal of this exercise is to prove that the geometric distribution is the only one discrete random variable which has the property of lake of memory. In the first part of the exercise you will prove that the geometric distribution has this property. Then in the second part you will prove that a distribution which has the property of lake of memory is geometric.

1. *Let* $X \sim \mathcal{G}(p)$.

- **a.** *Recall the value of the sum* $1 + u + u^2 + \cdots + u^n$.
- **b.** *Prove that* $P(X \geqslant k) = q^k$ *with* $q = 1 - p$.
- **c.** *Conclude that* X *has the property of lake of memory, i.e.*

$$P(X \geqslant n+m) | X \geqslant m) = P(X \geqslant n)$$

2. *Now we consider* Z *a discrete random variable such that* $Z(\Omega) = \mathbb{N}$, *for all* $m \in \mathbb{N}$, $P(Z \geqslant m) > 0$. *We suppose that* Z *satisfies the property of lake of memory for all* $(n, m) \in \mathbb{N} \times \mathbb{N}$ *:*

$$P(Z \geqslant m + n | Z \geqslant n) = P(Z \geqslant m)$$

Let p *denote the value* $p = P(Z = 0)$ *with* $p > 0$.

- **a.** *Show that* $P(Z \geqslant 1) = 1 - p$. *We denote by* q *the number* $1 - p$.
- **b.** *Show that* $P(Z \geqslant n + m) = P(Z \geqslant n)P(Z \geqslant m)$.
- **c.** *For all* $n \in \mathbb{N}$ *we denote by* u_n *the sequence* $u_n = P(Z \geqslant n)$.
  - *i. Prove that* (u_n) *is a geometric sequence, i.e.* $u_{n+1} = ru_n$. *What is the value of* r *?*
  - *ii. Determine* u_n *in terms of* q, n.
  - *iii. Express* $P(Z = n)$ *in terms of* u_n *and* u_{n+1}.

iv. Conclude $P(Z = n) = p^n q$. *What is the distribution Z ?*

Exercise 3

In this exercise we consider f *defined over* $\mathbb{R}$ *by :*

$$f(x) = \begin{cases} 3(1-x)^2 & \text{if } x \in [0,1] \\ 0 & \text{otherwise} \end{cases}$$

1. *Sketch the graph of the function* f.

2. *Show that* f *is a density function. We call* X *the random variable whose density is* f.

3. *Compute the cumulative distribution function* F_X *of* X *and find the value of* $P(-1 < X < \frac{1}{2})$. *Give on the graph of question* **1** *the geometric interpretation of* $P(-1 < X < \frac{1}{2})$

We now want to calculate $E(X)$ *and* $V(X)$. *For this purpose we introduce a new random variable* $Y = -\ln(1-X)$.

4. *Distribution of* Y *:*

a. *Show that the event* $(Y \leqslant y)$ *is equivalent to the event* $(X \leqslant 1 - e^{-y})$.

b. *Let us denote by* $G_Y(y)$ *the cumulative distribution function of* Y. *Deduce from the previous question that* $G_Y(y) = F_X(1 - e^{-y})$.

c. *Conclude that* $Y \sim \mathcal{E}(3)$.

5. *Expectation of* X *:*

a. *Show that* $\int_{-\infty}^{\infty} e^{-y} f_Y(y) dy = \frac{3}{4}$.

b. *Let* g *be a function. If* f_Y *is the density function of* Y, *the expectation of* $g(Y)$ *can be calculated from* $E(g(Y)) = \int_{-\infty}^{+\infty} g(y) f_Y(y) dy$. *Show that* $E(e^{-Y}) = \dfrac{3}{4}$.

c. *Find the function* g *such that* $X = g(Y)$ *and then calculate* $E(X)$.

6. *Calculate* $V(X)$ *(you can first start to explain which computations you will need to do, and then try to do it).*

Correction

Exercise 1

1. *The number of possible hands is* $\binom{52}{5}$. *In order to calculate the probabililty of obtaining a given type of hand we count the number of corresponding hands :*

a. *A pair : choose a kind* $\binom{13}{1}$, *then the suits* $\binom{4}{2}$ *for the pair. Choose the corresponding three other kinds* $\binom{12}{3}$ *and for each kind the suit* $\binom{4}{1}$. *The number of hands corresponding to a pair is* $\binom{13}{1} \times \binom{4}{2} \times \binom{12}{13} \times 4^3$. *The probability of having a pair is* $\dfrac{\binom{13}{1} \times \binom{4}{2} \times \binom{12}{3} \times 4^3}{\binom{52}{5}} = 0.422$.

b. *A full house : select one kind* $\binom{13}{1}$ *and three cards out of the* 4 *of the this kind* $\binom{4}{3}$, *then select the other kind* $\binom{12}{1}$ *and two cards of this kind* $\binom{4}{2}$. *One obtains* $\binom{13}{1} \times \binom{4}{3}\binom{12}{1}\binom{4}{2}$ *possible hands. Thus the probability of obtaining a full house is* $\dfrac{\binom{13}{1} \times \binom{4}{3}\binom{12}{1}\binom{4}{2}}{\binom{52}{5}}$.

c. *A flush : select a suit* $\binom{4}{1}$, *then we choose* 5 *cards out of the* 13 *of the same suits,* $\binom{13}{5}$. *This makes* $\binom{4}{1} \times \binom{13}{5}$ *possibilities. But the hands corresponding to straight flush are counted in that calculation and there should be removed. There are* 40 *of them (see test 1). Thus the probabililty of having a flush (which is not a straight flush) is* $\dfrac{\binom{4}{1} \times \binom{13}{5} - 40}{\binom{52}{5}} = 0.198$.

d. *A straight : select the lowest card of your straight. There are* 10 *possibilities (if you consider the first possible sequence* $1-2-3-4-5$ *and the last one* $10-J-Q-K-1$). *Then for each card there are* 4 *possibilities (suit). It makes* 10×4^5. *But again the straight flush should be removed. Thus the probability is* $\dfrac{10 \times 4^5 - 40}{\binom{52}{5}} = 0.394$.

2. **a.** *Without restriction the number of arrangement is* $8!$ *(*8 *possibilities for the choice of first person,*7 *for the second,...).*

b. *Persons* A *and* B *must seat together. One consider persons* A *and* B *as a unique person and we consider there are* 7 *people to be seated. Then one obtains* $2 \times 7!$ *possibilities (we multiply by* 2 *because we have* AB *or* BA).

c. *If the first person on the left is a man there are* 4 *possibilities, then there are* 4 *possibilities for the second person who has to be a woman, then* 3 *possibilities for the third person who is a man,.... It makes* $4!4!$ *possibilities. If the first personn on the left is a woman one obtains the same number. Thus the number of arrangement is* $2 \times 4!4!$.

d. *There are four possibilities to place the first of the five men (first, second or third position). Once the first man is placed there are* $5!$ *ways to arrange the men by permutation among them and* $3!$ *ways to arrange the women. Thus the number of possibilities is* $4 \times 5!3!$

Exercise 2

1. $X \sim \mathcal{G}(p)$.

a. $1+u+u^2+\cdots+u^n = \dfrac{1-u^{n+1}}{1-u}$.

b. $P(X \geqslant k) = 1 - P(X \leqslant k-1) = 1 - (P(X=0) + P(X=1) + \cdots + P(X = k-1)) = 1 - (p + (1-p)p + \cdots + (1-p)^{k-1}p) = 1 - p(1 + (1-p) + \ldots (1-p)^{k-1}) = 1 - p(1 + q + \cdots + q^{k-1}) = 1 - p\dfrac{1-q^k}{1-q} = 1 - p\dfrac{1-q^k}{p} = 1 - 1 + q^k = q^k$.

c. *We use the definition of conditional probabilities :*

$$P(X \geqslant n+m | X \geqslant m) = \frac{P((X \geqslant n+m) \cap (X \geqslant n)}{P(X \geqslant m)} = \frac{P(X \geqslant n+m)}{P(X \geqslant m)} = \frac{q^{m+n}}{q^m} = q^n$$

2. *We now assume Z has the hypothesis of the exercise (property of lake of memory and $Z(\Omega) = \mathbb{N}$ and $p = P(Z=0)$).*

a. $P(X \geqslant 1) = 1 - P(X=0) = 1-p$

b. $P(Z \geqslant n+m) = P(Z \geqslant n+m | X \geqslant m) P(X \geqslant m) = P(X \geqslant n) P(X \geqslant m)$ *because of the property of lake of memory.*

c. *Let* $u_n = P(Z \geqslant n)$

i. $u_{n+1} = P(X \geqslant n+1) = P(X \geqslant 1)P(X \geqslant n) = (1-p)u_n$.

ii. u_n *is a geometric sequence with* $r = q$ *and* $u_1 = q$ *then* $u_n = q^n$.

iii. $P(Z = n) = P(Z \geqslant n) - P(Z \geqslant n+1) = u_n - u_{n+1}$

iv. Thus $P(Z=n) = q^n - q^{n+1} = q^n(1-q) = q^n p$. *It is a geometric distribution of parameter* p.

Exercise 3

1. f *is nonnegative and we have* $\int_{-\infty}^{\infty} f(x)dx = \int_0^1 3(1-x)^2 dx = [-(1-t)^3]_0^1 = 1$. *Thus* f *is a density function.*

2. *The cumulative distribution function is defined by* $F_X(x) = \int_{-\infty}^{x} f(t)dt$. *If* $x \leqslant 0$ *we have* $F_X(x) = \int_{-\infty}^{x} 0dt = 0$. *If* $x \geqslant 1$ *we have* $F_X(x) = \int_{-\infty}^{x} f(t)dt = \int_0^1 3(1-t)^2 dt = 1$. *If* $0 \leqslant x \leqslant 1$ *one obtains* $F_X(x) = \int_{-\infty}^{x} f(t)dt = \int_0^x 3(1-t)^2 dt = [-(1-t)^3]_0^x = 1 - (1-x)^3$. *Thus the cumulative distribution function is*

$$F_X(x) = \begin{cases} 0 & \text{if } x \leqslant 0 \\ 1-(1-x)^3 & \text{if } x \in [0,1] \\ 1 & \text{if } x \geqslant 1 \end{cases}$$

$P(-1 < X < \frac{1}{2}) = F_X(\frac{1}{2}) - F_X(-1) = \frac{7}{8} - 0$.

3. *Let* $Y = -\ln(1-X)$

a. $(Y \leqslant y) \Leftrightarrow (-\ln(1-X) \leqslant y) \Leftrightarrow (\ln(1-X) \geqslant -y) \Leftrightarrow (1-X \geqslant e^{-y}) \Leftrightarrow (X \leqslant 1-e^{-y})$.

b. $G_y(y) = P(Y \leqslant y) = P(X \leqslant 1-e^{-y}) = F_X(1-e^{-y})$.

c. $1-e^{-y} \in [0,1]$ *for* $y \geqslant 0$ *thus for* $y \geqslant 0$ *we have* $G_Y(y) = 1-(1-(1-e^{-y}))^3 = 1-e^{-3y}$. *If* $y < 0$ *then* $1-e^{-y} < 0$ *and* $F_X(1-e^{-y}) = 0$. *Thus for* $y \leqslant 0$ *we have* $G_Y(y) = 0$. *We conclude that* G_Y *is the cumultative distribution of an exponential distribution of parameter* $\lambda = 3$.

4. $\int_{-\infty}^{\infty} e^{-}yf_Y(y)dy = \int_0^{\infty} e^{-y}3^{-2y}dy = 3\int_0^{\infty} e^{-3y}dy = 3[-\frac{e^{-4y}}{4}]_0^{\infty} = \frac{3}{4}$.

5. $E(e^{-Y}) = \int_{-\infty}^{\infty} e^{-y} f_Y(y)dy = \dfrac{3}{4}$.

6. $X = 1-e^{-Y}$ *with* $g(y) = 1-e^{-y}$. *Thus* $E(X) = E(1-e^{-Y})) = E(1)-E(e^{-Y})$ *(by linearity of the expectation). Thus* $E(X) = 1 - \frac{3}{4} = \frac{1}{4}$.

7. *To obtain the variance we need to calculate first* $E(X^2) = E((1 - e^{-Y})^2) = E(1 - 2e^{-Y} + e^{-2Y}) = E(1) - 2E(e^{-Y}) + E(e^{-2Y}) = 1 - 2 \times \frac{1}{4} + E(e^{-2Y})$. *But* $E(e^{-2Y}) = \int_{-\infty}^{\infty} e^{-2y} f_Y(y)dy = \frac{3}{5}$. *Thus* $E(X^2) = \dfrac{11}{10}$ *and* $E(X)^2 = \dfrac{1}{16}$. *Finally* $V(X) = \dfrac{11}{10} - \dfrac{1}{16}$.