Project IND ENG 241 - Fall 25.

Individual project. **Due date: Monday December 15, 2025 at 11:59 PM.**
Online submission on Gradescope. Submit a PDF with Python code and simulation.
Advice: PDF generated from a Jupyter Notebook is encouraged.

AI policy: AI tools can be used to help you to respond but it has to be used as an assistant and not a creative tool. The answers should reflect your own thinking, may be assisted with AI, and be presented in a pedagogical way. Presentation and clarity are part of the grade.
This project contains **theoretical** questions, noted (**T**), and questions requiring numerical simulations to be performed in Python, noted (**S**).

1 Microcredit and Markov Chains

In this project, we study a microcredit model introduced by G. Tedeshi, in which obtaining a loan depends on the financial situation of the investment project. The first part of this project presents a very simple model. The second part extends this first case to include credit exclusion in the event of a loan refusal.

1.1 An example of microcredit with state changes

Let us begin with a very simple example where a financial agent can be in two states at any time t:

- a loan applicant, denoted by state A ("applicant"),
- a loan recipient, denoted by state B ("beneficiary").

We assume that if the financial agent receives a loan at time k, they repay it in the event of project success with probability $\beta \in [0, 1]$. Their loan is then automatically renewed in the next period $k + 1$, hence they remain in state B. If the agent defaults (with probability $1 - \beta$) and cannot repay the loan, they lose their automatic renewal and become a loan applicant again (state A) at $k + 1$.

As a loan applicant (that is, if the agent is in state A), they obtain a loan with probability $\alpha \in [0, 1]$.

We denote by $(E_k)_{k\in\mathbb{N}}$ a sequence of random variables taking values in $\{A, B\}$.

Question 1: (T) What does E_k represent?

Question 2: (T) For every integer $k \geq 0$, define

$$P_{1,1} = \mathbb{P}(E_{k+1} = A|E_k = A),\ P_{2,1} = \mathbb{P}(E_{k+1} = A|E_k = B),$$

$$P_{1,2} = \mathbb{P}(E_{k+1} = B|E_k = A),\ P_{2,2} = \mathbb{P}(E_{k+1} = B|E_k = B).$$

Compute $P_{i,j}, i, j \in \{1, 2\}$ as functions of α, β and write the matrix P. What does it represent?

Question 4: (S) Assign the value 0 when the agent is in state A and 1 when in state B. Assume that $\mathbb{P}(E_0 = A) = \frac{1}{2}$. Let $k = 5000$, $\alpha = \beta = \frac{1}{2}$. Plot one possible trajectory[1] of $(E_k)_{k\geq 0}$. Compute $\frac{1}{1000}\sum_{k=0}^{1000} E_k$ for this trajectory. Was this result predictable? (In particular, show that in this case, $(E_k)_{k\geq 0}$ is a sequence of identically distributed and independent random variables.)

Question 5: (T) Let $\pi_0 = (\mathbb{P}(E_0 = A), \mathbb{P}(E_0 = B))$ be the initial distribution vector. Define

$$\pi_k = (\mathbb{P}(E_k = A), \mathbb{P}(E_k = B)).$$

What does π_k represent?

Questions 6 to 11: take $\beta = \frac{2}{3}$ and $\alpha = \frac{1}{4}$.

Question 6: (S) Assume $\pi_0 = (\frac{1}{2}, \frac{1}{2})$. Plot 100 trajectories of $(E_k)_{1\leq k\leq 500}$ assigning 0 to state A and 1 to state B. How would you approximate π_{500} in this case? What do you obtain?

Question 7: (S) Using the previous method, plot π_{500} as a function of an arbitrary π_0 in $[0, 1]^2$ such that the sum of its components equals 1. What do you observe?

Question 8: (T) Show that

$$\pi_k = \pi_0 P^k.$$

Question 9: (T) Compute $\pi_* = (\pi_*^1, \pi_*^2)$ in $[0, 1]^2$ such that

$$\pi_* = \pi_* P.$$

Question 10: (T) Study the convergence of the components of π_k.

Question 11: (T) Show that there exists a positive constant C and a constant $\eta \in [0, 1)$ such that for any initial distribution π_0,

$$|\mathbb{P}(E_k = \varepsilon_i) - \pi_*^i| \leq C\eta^k,\ i \in \{1, 2\},\ \varepsilon_1 = A,\ \varepsilon_2 = B.$$

What can you deduce about the convergence of the variable E_k?

Question 12: (T) Suppose now that a financial agent starts at time 0 with a state E_0 following a Bernoulli distribution with parameter $p_0 \in [0, 1]$ (where value 1 corresponds to state B and 0 to state A). The probability of success β of their project is known, but the agent does not know their probability α of obtaining the loan as an applicant in state E_1. After 500 attempts, the agent obtained the loan 60% of the time. Noting that E_1 follows a Bernoulli distribution whose parameter must be expressed in terms of the given data, provide a 95% confidence interval for α as a function of β and p_0.

[1] A trajectory associated with a sequence $(X_n)_{n\geq 0}$ of random variables is the mapping $n \mapsto X_n(\omega)$ for a fixed $\omega \in \Omega$. Here, this means randomly drawing E_0 and then the subsequent states according to the transition rules.

1.2 Extension to a model with fixed exclusion period

Now suppose that if the bank refuses a borrower's loan application, the borrower enters an exclusion period of length N, meaning that they cannot obtain a loan for N years. Thus:

- The probability that an agent benefiting from a loan at time k sees it renewed from k to $k+1$ depends only on project success and equals α,
- If the agent has a loan at time k and defaults with probability $1-\alpha$, they enter an exclusion state for the next N years,
- If the agent is in the last year of exclusion, the probability of obtaining a loan the following year is β, otherwise they remain excluded for one more year before reapplying.

We denote by E_k the state in which the financial agent is at year k, which can be B (loan beneficiary), A^1 (loan applicant), or A^j, $j \in \{2, N\}$ (excluded for j remaining years).

Note that here, E_k depends only on the agent's state at $k-1$, not on the entire history. Hence, $(E_k)_{k\geq 1}$ is a Markov chain on the state space $\{B, A^1, \ldots, A^N\}$.

We say that a real matrix Q of size $(N+1)\times(N+1)$ is a transition matrix on a finite space $E = \{x_1, \ldots, x_{N+1}\}$ if

- $Q(x_i, x_j) \geq 0$ for all $(i,j) \in \{1, \ldots, N+1\}^2$,
- $\sum_{j\in E} Q(x_i, x_j) = 1$.

It follows that $Q(x_i, x_j) \in [0,1]$. We denote $Q_{i,j} := Q(x_i, x_j)$.

We say that Q is the transition matrix associated with $(E_k)_{k\geq 0}$ if

$$Q(x_i, x_j) = \mathbb{P}(E_{k+1} = x_j | E_k = x_i).$$

Question 13: (T) Compute

$$\begin{aligned}
P_{1,1} &= \mathbb{P}(E_{k+1} = B | E_k = B), \\
P_{1,N+1} &= \mathbb{P}(E_{k+1} = A^N | E_k = B), \\
P_{1,j} &= \mathbb{P}(E_{k+1} = A^j | E_k = B),\ j \in \{2, \ldots, N-1\}, \\
P_{2,1} &= \mathbb{P}(E_{k+1} = B | E_k = A^1), \quad P_{2,2} = \mathbb{P}(E_{k+1} = A^1 | E_k = A^1), \\
P_{i+1,i} &= \mathbb{P}(E_{k+1} = A^{i-1} | E_k = A^i), \quad i \in \{2, \ldots, N\}, \\
P_{2,j+1} &= \mathbb{P}(E_{k+1} = A^j | E_k = A^1),\ j \in \{2, \ldots, N\}, \\
P_{i,j} &= \mathbb{P}(E_{k+1} = A^{j-1} | E_k = A^{i-1}),\ i \in \{3, \ldots, N+1\},\ i \neq j+1.
\end{aligned}$$

Then write the matrix P in $\mathcal{M}_{N+1\times N+1}([0,1])$.

By identifying $x_1 = B$ and $x_i = A^{i-1}$ for $2 \leq i \leq N+1$, show that P is the transition matrix associated with $(E_k)_{k\geq 1}$.

Question 14: (T) Solve $\pi_\star P = \pi_\star$ for the unknown row vector $\pi_\star$ of size $N+1$ with values in $[0,1]^{N+1}$ such that the sum of its components equals 1.

Question 15: (T) Let $\pi_k := (\mathbb{P}(E_k = B), \mathbb{P}(E_k = A^1), \ldots, \mathbb{P}(E_k = A^N))$. Give the relation between π_k and π_{k+1}, then deduce π_k as a function of π_0 and P.

Question 16: (S) Let $N = 5$, $\alpha = \frac{1}{3}$ and $\beta = \frac{3}{4}$. Compute numerically each component of $\pi_0 P^n$ where $n = 1000$ and $\pi_0 := (\frac{1}{N+1}, \ldots, \frac{1}{N+1})$ of size $N+1$. What do you observe, using the two previous questions as guidance? Test this result with other vectors π_0 of size $N+1$. What do you notice?

Reference: Diener, F., Diener, M., Khodr, O., & Protter, P. (2009, December). Mathematical models for microlending. In Proceedings of the 16th Mathematical Conference of the Bangladesh Mathematical Society (Vol. 1).

2 Ruin Theory

Ruin theory is considered a branch of insurance risk management. It studies the probability that an unfavorable event occurs for an insured party under an insurance contract. Such events generate risks the insurer must evaluate in order to pay indemnities to the insured.

We typically model a claim as a jump that (negatively) impacts the insurer's wealth. In a first step, we analyze the modeling of claims as jump processes. Then, we extend this to modeling the total cost associated with these claims. Finally, we see how to calibrate the insurer's initial reserve so as to minimize its probability of ruin, *i.e.*, the probability that its wealth becomes negative.

2.1 Modeling random claims

The modeling of the arrival of random events such as incidents linked to a risky project, earthquakes, or even customer arrivals in a store must satisfy certain criteria to be as realistic as possible. The Poisson processes we study in this section are one way to meet these modeling requirements.

We model the occurrence of random events by a sequence $(\tau_n)_n$ of i.i.d. random variables with exponential distribution of parameter $\lambda > 0$. Recall that the exponential distribution with parameter λ has density f_λ defined by

$$\gamma_\lambda(x) = \lambda e^{-\lambda x} \mathbf{1}_{x \geq 0}, \quad x \in \mathbb{R}.$$

Let $T_1 = \tau_1$ be the time of the first event and $T_n := \sum_{i=1}^n \tau_i$ for $n > 0$ the time at which the n-th event occurs.

Question 1: (T) By using the convolution product, show that T_n has density

$$\gamma_{\lambda,n}(x) = \frac{\lambda^n}{(n-1)!} x^{n-1} e^{-\lambda x} \mathbf{1}_{x \geq 0}, \quad x \in \mathbb{R}.$$

Question 2: (T) Define a Poisson process N by

$$N_0 = 0, \qquad N_t = \sum_{i=1}^{+\infty} \mathbf{1}_{T_i \leq t}.$$

Noting that $N_t = \sup\{n \in \mathbb{N} : \; T_n \leq t\}$, give the interpretation of N_t.

Question 3: (T) Give a necessary and sufficient condition such that

$$\mathbb{P}\big(N \text{ only has jumps of size } 1\big) > 0.$$

Deduce that

$$\mathbb{P}\big(N \text{ only has jumps of size } 1\big) = 1.$$

Hint: note that

$$\{N \text{ only has jumps of size } 1\} = \{\forall i \geq 1, \; T_i < T_{i+1}\} = \bigcap_{i \geq 1} \{\tau_i > 0\}.$$

Question 4: (T) Show that N_t has a Poisson distribution with parameter λt.
Hint: use $\{N_t = n\} = \{T_n \leq t < T_{n+1}\}$.

We admit that $(N_t)_{t \geq 0}$ is a homogeneous process with independent and stationary increments:

1. (stationarity) If $0 \leq s \leq t$, then $N_t - N_s$ has the same distribution as N_{t-s}. Thus $N_t - N_s$ is Poisson with parameter $\lambda(t-s)$.

2. (independence) If $0 < t_1 < t_2 < \cdots < t_n$, then the random variables $N_{t_1}, \; N_{t_2} - N_{t_1}, \; \ldots, \; N_{t_n} - N_{t_{n-1}}$ are independent.

Question 5: (T) Let $n \in \mathbb{N}$, $t > 0$, and $U_1, \ldots, U_n$ be i.i.d. uniform on $[0, t]$. Let $U_{(1)}, \ldots, U_{(n)}$ be the order statistics of $(U_1, \ldots, U_n)$, i.e., $(U_{(1)}, \ldots, U_{(n)})$ is the vector $(U_1, \ldots, U_n)$ rearranged in increasing order. In particular,

$$U_{(1)} = \min\{U_1, \ldots, U_n\}, \qquad U_{(n)} = \max\{U_1, \ldots, U_n\}.$$

Show that the random vector $(U_{(1)}, \ldots, U_{(n)})$ has a density, and compute it.

Question 6: (T) Show that the conditional law of $(T_1, \ldots, T_n)$ given $N_t = n$ is the same as that of $(U_{(1)}, \ldots, U_{(n)})$.

Question 7: (S) Using the previous question, simulate a Poisson process with $\lambda = 5$ and $t = 1$.

2.2 Cost of indemnities

Consider a Poisson process with parameter λ as defined above. Let $(J_n)_n$ be a sequence of i.i.d. random variables independent of all the random variables N_t, $t \geq 0$.

Consider the following process, called a *compound Poisson process*:

$$\forall t \geq 0, \quad C_t = \sum_{i=1}^{N_t} J_i.$$

We admit that (C_t) is also a process with independent and stationary increments.

Question 8: (T) What do J_i and C_t represent?

Question 9: (T) Show that C is a (simple) Poisson process if and only if

$$J_1 \sim \mathrm{Ber}(p), \quad p \in (0,1),\ t \geq 0,$$

where $\mathrm{Ber}(p)$ denotes the Bernoulli distribution with parameter p.
Hint: equivalently, $\mathbb{P}[C \text{ only has jumps of size 0 or 1}] = 1$.

Question 10: (T) Suppose J_1 has finite variance and set

$$\mu = \mathbb{E}[J_1] \quad \text{and} \quad \sigma^2 = \mathrm{Var}(J_1).$$

For $t \geq 0$, compute the expectation and variance of C_t in terms of μ, σ^2, λ, and t.

Question 11: (S) Inspired by Question 7, simulate a trajectory of C_t with $t = 1$, $\lambda = 5$, assuming the J_i follow a centered normal distribution with variance 5.

2.3 Probability of ruin

The processes introduced above are particularly relevant for an insurance company that covers a certain class of risks. Assume claims occur according to a Poisson process N_t with parameter $\lambda > 0$ (with initial date $t = 0$) and that J_i represents the amount of the indemnity for the i-th claim that the company must pay. Assume (J_i) are positive i.i.d. random variables independent of (N_t), so that the total amount of indemnities paid by the company between times 0 and t is the compound Poisson process:

$$C_t = \sum_{i=1}^{N_t} J_i.$$

Let R_t denote the company's wealth (or reserve) at time t. We assume

$$R_t = u + ct - C_t = u + ct - \sum_{i=1}^{N_t} J_i.$$

Here $u > 0$ is the initial wealth and the term ct corresponds to a steady inflow of funds (thus $c > 0$).
The natural question for the insurance company is how to choose u so that the *probability of ruin* remains below a given threshold.
Define:

- $\varphi(u;t) = \mathbb{P}(\exists s \in [0,t] : R_s < 0)$, the probability of ruin before time t given initial wealth u;
- $\varphi(u) = \mathbb{P}(\exists s > 0 : R_s < 0)$, the ultimate probability of ruin with initial wealth u.

Question 12: **(T)** Show that $\varphi(u) = \lim_{t\to+\infty} \varphi(u,t)$.

Question 13: **(T)** Assume additionally that the *safety loading condition* holds:

$$c - \lambda\mu > 0.$$

By applying the strong law of large numbers, show that

$$\lim_{t\to+\infty} R_t = +\infty \quad \text{a.s.}$$

Interpret this result.

Question 14: **(S)** Conversely, when $c-\lambda\mu < 0$, the process R_t converges to $-\infty$ almost surely. Now take J_i standard normal with mean $\mu = 1$, $\lambda = 5$, and $c = 7$. Simulate a trajectory of R_t and give the time at which ruin occurs.

Reference: Søren Asmussen, *Ruin Probabilities*.

3 Volatility calibration with data

1. **Question 1** **(T)** We model the price of a financial asset as a stochastic process S with price dynamic given as follows:

   $$S_k = S_{k-1} + \sigma\xi_k,\ t_0 = 0 < gt_1 < \cdots < t_N = T$$

   where $(\xi_k)_{k\geq 0}$ is a sequence of independent standard Gaussian distributions and $\sigma > 0$ the unknown volatility of the asset. The goal of this section is to estimate σ.

   We set

   $$\hat{\sigma}_N^2 := \frac{1}{N}\sum_{k=1}^{N}(S_{t_k} - S_{t_{k-1}})^2.$$

   Compute $\mathbb{E}[\hat{\sigma}_N^2]$ and $\mathrm{Var}(\hat{\sigma}_N^2)$. What do you deduce?

2. **Question 2** **(S)**We assume that each t_k corresponds to a trading day. Go to Yahoo Finance for example and choose the historical prices of one asset during T =three months (to be converted in trading day).

   You can export the table from 'History' Time period 3M. Calibrate the volatility σ using the estimator provided in the previous section.

4 Simulation with rejection method

1. Example extracted from Problem 10.2 in the textbook[2] We want to simulate a random variable X with density

   $$f(x) = \begin{cases} 30(x^2 - 2x^3 + x^4),\ 0 < x < 1, \\ 0 \text{ otherwise.} \end{cases}$$

[2]the correction is available in the textbook.

We take $g(x) = 1$ for any $x \in [0, 1]$, that is considering a uniform distribution Y on $[0, 1]$. We have seen that

$$c = \max \frac{f(x)}{g(x)} = \frac{15}{8}.$$

> **Question 1 (S-T)** Create an algorithm to simulate X with $M = 10,000$ simulations. Compute the empirical mean and compare with the theoretical value.
>
> **Question 2 (S)** Compute the number of times that your algorithm has accepted the condition $U \leq \frac{f(Y)}{cg(Y)}$ and observe that the frequency to accept this condition is indeed $\frac{1}{c}$.

2. Normal distribution, see Example 2c in the textbook. To simulate a normal distribution with the rejection method, we first notice that

$$Z = |N|, \text{ where } N \sim \mathcal{N}(0, 1),$$

has the density

$$f_Z(x) = \frac{2}{\sqrt{2\pi}} e^{-x^2/2}, \ 0 < x < \infty.$$

Let Y be an exponential distribution with parameter 1, that you have simulated in a previous homework. We note that

$$\frac{f_Z(x)}{g(x)} \leq c := \sqrt{\frac{2e}{\pi}}, \text{ for any } x > 0.$$

Hence

$$\frac{f_Z(x)}{cg(x)} = e^{-\frac{(x-1)^2}{2}}.$$

Question 3 (S) Generate $M = 10,000$ simulation of Z from Y with the rejection method. Compute the empirical mean and sample variance.

Question 4 (S) then, generate M simulations of

$$N = \begin{cases} Z \text{ if } V \leq \frac{1}{2} \\ -Z \text{ if } V \geq \frac{1}{2}, \end{cases}$$

where V is an independent uniform random variable on $[0, 1]$. Draw the histogram of the $M = 10,000$ simulations of N, compute the empirical mean and sample variance.