1. A person tosses a fair coin until she obtains 2 heads in a row. She then tosses a fair die the same number of times as she tossed the coin. What is the expectation and variance for the number of 1s in the sequence of die rolls? How would your answer change if the coin was unfair, and the probability of obtaining heads was p?

2. Let X and Y be independent random variables with common distribution function F and density function f.
   (a) Show that $V = \max\{X, Y\}$ has distribution function $F_V(v) = F(v)^2$ and density function $f_V(v) = 2f(v)F(v)$.
   (b) Find the density function of $U = \min\{X, Y\}$.
   (c) Let X and Y be independent random variables each having the uniform distribution on $[0, 1]$. Find the expectation $E(U)$ and $\mathrm{cov}(U, V)$.
   (d) Let X and Y be independent exponential random variables with mean 1. Show that U is an exponential random variable with mean $\frac{1}{2}$. Find $E(V)$ and $\mathrm{var}(V)$.

3. The following questions concern different types of convergence.
   (a) Prove that weak limits are unique (weak convergence is sometimes also known as convergence in distribution).
   (b) Prove that almost sure convergence implies convergence in probability.
   (c) Does weak convergence imply convergence in probability? Either prove or provide a counterexample.

4. On the state space $S = \{1, 2, 3\}$, consider a Markov chain W_n with transition matrix $Q = (Q_{ij})$ and with initial probability distribution q. Define $V_n = (W_n, W_{n+1})$.
   (a) Prove that V_n is a Markov chain with state space $S \times S$. Compute its transition matrix G.
   (b) What is the initial probability distribution p of V_n?
   (c) Assume that q is a probability distribution invariant by Q. Prove that p is then invariant by G.

5. Let X and Y be random variables taking values in the set $A = \{0, 1\}$ and having the following joint probability distribution:

$$\mu(x, y) = \mathrm{Prob}(X = x \text{ and } Y = y) = \frac{1}{c} e^{x+y-xy}$$

   for all (x, y) in the set $S = A \times A$.
   (a) Compute the value of the constant c, the means $E(X)$ and $E(Y)$, and the covariance $\mathrm{Cov}(X, Y)$.
   (b) Compute the (marginal) probability distribution of X and Y, as well as the conditional distributions of Y given $X = x$ and of X given $Y = y$.
   (c) Compute the probability distribution of $Z = \frac{XY}{1+X+Y}$.

6. The notations and hypotheses are the same as in Question 5.
   (a) On the state space $S = A \times A$, consider two arbitrary Markov chain transition matrices G and H. Prove that the product $M = GH$ is also a transition matrix.
   (b) Define a 4×4 matrix $Q((x, y), (u, v))$, for all (x, y) and (u, v) in S, by the formula

$$Q((x, y), (u, v)) = 1_{\{u=x\}} \times \frac{e^{v(1-x)}}{1 + e^{1-x}}.$$

   By symmetry we define another matrix R by the formula

$$R((x, y), (u, v)) = 1_{\{v=y\}} \times \frac{e^{u(1-y)}}{1 + e^{1-y}}.$$

   Write Q explicitly to verify that Q is the transition matrix of a Markov chain on the state space S. By symmetry this shows also that R is a transition matrix.

(c) Prove that the Markov chain with transition matrix Q admits the probability μ as a stationary probability distribution. Determine all the other stationary probability distributions of this Markov chain. By symmetry, R has similar properties.

(d) Prove that the Markov chain with transition matrix $K = QR$ is ergodic and has a unique stationary probability distribution equal to μ.

7. (a) Let Y be a random variable having a Poisson distribution with parameter $\lambda > 0$. For $t \geq 0$, compute explicitly the Laplace transform $g(t) = E(e^{tY})$.

   (b) Explain why the following relation is valid for any $r \geq 0$ and any $t \geq 0$:
$$P(Y > r) = P(e^{-tr}e^{tY} > 1) \leq g(t)e^{-tr}.$$

   (c) For any $r > 0$, compute $G(r) = \min_{t \geq 0} g(t)e^{-tr}$ and use (b) to prove that
$$P(Y > r) \leq G(r).$$

   (d) Let $Y_1, \ldots, Y_n$ be independent random variables having the same distribution as Y. Let $Z_n = \frac{Y_1 + \cdots Y_n}{n}$. Fix $a > \lambda$. Use (c) to derive an upper bound for $(1/n) \log P(Z_n > a)$ and compute the limit of this upper bound as $n \to \infty$.

8. Let $T_1 < T_2 < \cdots < T_k < \cdots$ be the sequence of random successive car passage times at a toll booth. We assume that the waiting times $S_n = T_n - T_{n-1}$ are independent and have the same Gaussian distribution with mean m and standard deviation s.

   (a) What is the probability distribution of T_k?

   (b) For each positive t, call $N(t)$ the random number of cars which went through the toll booth between times 0 and t. For each integer k, express the event $N(t) < k$ in terms of T_k and t in order to compute $\text{Prob}(N(t) < k)$.

   (c) Use the result from (b) to compute an integral formula for the probability $\text{Prob}(N(t) = k)$.

   (d) Explain briefly how one could attempt a similar approach to compute $\text{Prob}\{(N(t) < k) \cap (N(t+u) < r)\}$. The full computation is not requested here.

9. State both cases of the Borel-Cantelli Lemma. Then prove the following: Let $X_1, X_2, X_3, \ldots, X_n, \ldots$ be independent identically distributed random variables with $E[|X_i|] < \infty$ and $\text{Var}(|X_i|) < \infty$. Show that
$$P(|X_n| > n^{1/2+\epsilon} \text{ i.o.}) = 0$$
for any $\epsilon > 0$.

10. (About limiting theorems.)

    (a) State and prove the Weak Law of Large Numbers, under the assumption that all random variables X_n have finite variance.

    (b) State and prove the Central Limit Theorem.

11. The following question concerns some basic properties of probability:

    (a) State the Law of Total Probability.

    (b) Suppose that you pick a number from the unit interval with uniform probability. If the number is greater than $r \in [0, 1]$, toss a fair six-sided die. If it is smaller than $r \in [0, 1]$, then toss a six-sided die where rolling an even number is twice as likely as rolling an odd number. What is the probability of getting an even number at the end of the experiment? Explain all the steps.

    (c) Suppose that the k events $\{A_1, A_2, \ldots, A_k\}$ form a partition of a sample space Ω. Assume that $P(B) > 0$. Prove that if $P(A_1|B) < P(A_1)$, then $P(A_i|B) > P(A_i)$ for some $i = 2, \ldots, k$.

12. Let Z_n be a sequence of independent random variables having the same cumulative distribution function $F(z)$. Assume that $F(z)$ is strictly increasing and continuous, so that it has an inverse function $G(y)$ verifying the identity $F(G(y)) = y$.

    Define new random variables U_n with values in the interval $[0, 1]$ by $U_n = 1 - F(Z_n)$.

    (a) Prove that U_n is uniformly distributed on the interval $[0, 1]$.

    (b) Use the result in (a) to compute the density function of $Y_n = \log U_n$; compute also the expectation and the variance of Y_n.

    (c) Show that as n tends to infinity, the empirical means $M_n = (Y_1 + Y_2 + \cdots + Y_n)/n$ converge almost surely to a deterministic limit w, and compute w.

    (d) State the central limit theorem and explain why it applies to $n^{1/2}(M_n - w)$ as n tends to infinity.

13. This problem will the preceding result 12(d), but you DO NOT need to solve Problem 12 in order to solve this problem.

    (a) Let Z_n be a sequence of independent random variables having the same exponential density with mean 1. Compute the cumulative distribution function $F(z)$ of Z_n.

    (b) We now observe two kinds of random samples of n observations: either "error free" samples of the form $(Z_1, Z_2, \ldots, Z_n)$ or "polluted" samples, where one single observation Z_k has been erroneously replaced by $S_k = 4 + Z_k$. The proportion of polluted samples among all observed samples is 1%, and polluted samples occur at random among the observed samples. When we observe a new sample $(Z_1, Z_2, \ldots, Z_n)$, we do not know if this sample is polluted or not; so we compute the variables $Y_j = \log(1 - F(Z_j))$ and the sum $M_n = (Y_1 + Y_2 + \cdots + Y_n)/n$. We then reject the current sample whenever $M_n > K$, where K is a fixed threshold. Use the stated result 12(d) above to compute the probability that a true sample get rejected.

    (c) Use the stated result 12(d) above to compute the probability that a polluted sample is not rejected.

14. Consider a Markov chain Z_n on the finite state space $[1, 2]$, with transition matrix H given by

    $$H_{1,1} = 1 - H_{1,2} = a; \qquad H_{2,2} = 1 - H_{2,1} = b$$

    where $0 < a < 1$ and $0 < b < 1$ are unknown parameters.

    Assume that one has observed the following finite sequence of data $Z_0, Z_1, \ldots, Z_{27}$:

    $$2, 1, 1, 2, 2, 2, 1, 1, 2, 1, 2, 1, 1, 1, 2, 2, 2, 2, 1, 2, 1, 2, 1, 1, 2, 1, 1, 1$$

    (a) Compute the log-likelihood of these observations.

    (b) Compute the associated maximum likelihood estimators $\hat{a}$ and $\hat{b}$ of the parameters a and b.

    (c) Compute the stationary probability distribution $\gamma = [\gamma(1), \gamma(2)]$ of the Markov chain Z_n in terms of the unknown parameters a and b.

    (d) Compute an estimator of the probability $\gamma(1)$.

15. A gambler systematically bets one dollar on a single number at roulette. At each time n, the probability of a win in 1/36, and if the gambler wins, the casino gives him a payback of $W_n = 36$ dollars (including his one dollar bet); if the gambler loses, his bet is lost and his payback is $W_n = -1$ dollar. The gambler's total gain at time N is denoted $X_N = W_1 + \cdots + W_N$ and is hence the sum of all paybacks received after N successive bets. Note that the "gain" X_N can be negative as well as positive.

    (a) Compute the mean and variance of X_N. This does not require computing the probability distribution of X_N.

    (b) Compute the probability distribution of X_N.

    (c) State the central limit theorem, and explain why it applies to X_N for N large. Apply the central limit theorem to X_N.

(d) Use (c) to evaluate, for N large, the probability g_N that after N bets the gambler's total gain X_N is inferior to $-\sqrt{N}$, and compute the limit g_∞ of g_N as $N \to \infty$.

(e) Another gambler systematically bets one dollar on the color red, so that the probability of a win is $1/2$, and each win yields a payback $W_n = 2$ dollars, while each loss yields $W_n = -1$ dollar. As in (d), compute g_∞ for this case.

(f) For a gambler who cannot afford to lose more than $\sqrt{N}$ dollars after N bets, which of the two preceding gambling plans is the safest?

16. Suppose that you are given a number of independent samples from a Poisson distribution with unknown parameter λ.

(a) What is the estimate of the mean and the unbiased estimate of the variance?

(b) How would you compute the confidence interval for the mean?

(c) What is a good, unbiased estimator for the parameter λ? Prove that your estimator is unbiased.

(d) Does the variance of the estimator you gave in the previous part meet the Cramér-Rao lower bound?

17. (About conditional expectations.)

(a) Let X, Y, Z be independent and uniform in $[0, 1]$. Calculate

$$E[(X + 2Y + Z)^2 | X].$$

(b) Pick the point (X, Y) uniformly in the triangle

$$\{(x, y) \mid 0 \leq x \leq 1,\ 0 \leq y \leq x\}.$$

i. Calculate $E[X|Y]$.

ii. Calculate $E[Y|X]$.

iii. Calculate $E[(X - Y)^2|X]$.

(c) Assume that the two random variables X and Y are such that $E[X|Y] = Y$ and $E[Y|X] = X$. Show that $P(X = Y) = 1$.

18. Consider a three-dimensional cube composed of equally sized smaller cubes (chambers). There are doors on all the internal faces of the constituent chambers, however no openings on the faces on the exterior of the cube. The transition probability of going from chamber A_i to chamber A_j is nonzero only if they share a face, in which case it equals $1/N_i$, with N_i the number of openings on chamber A_i. Consider the Markov chain defined by these transition probabilities.

(a) Find the stationary probability distribution if there are 8 chambers. This corresponds to a cube with two chambers per side, or dimensions $2 \times 2 \times 2$.

(b) Find the stationary probability distribution if there are 27 chambers, that is, three chambers per side. (Hint: Use the symmetry of the situation to considerably reduce the problem.)

(c) Set up, but do not solve the following problem: There is an equal chance of starting in any of the 9 chambers forming one face of the cube in (b). Find equation(s) that give the probability that you will reach one of the 9 chambers on the opposite side of the cube before returning to one of the chambers on the starting side.

19. On the state space $S = \{1, 2, 3\}$, consider a Markov chain X_n having the 3×3 transition matrix Q. Denote the three rows of Q by Q_1, Q_2, Q_3, and assume that $Q_1 = [a, b, c]$, $Q_2 = [c, a, b]$, and $Q_3 = [b, c, a]$, where a, b, c are positive and satisfy $a+b+c = 1$. Call $m = [m(1), m(2), m(3)]$ the probability distribution of X_0.

(a) Let $p_n(i) = P(X_n = i)$ for $i = 1, 2, 3$. Give a matrix formula expressing the vector p_n in terms of m and Q. State a theorem explaining why as n tends to infinity, the $p_n(i)$ must have a limit $q(i)$ for $i = 1, 2, 3$.

(b) Write the system of linear equations determining the probability q, and prove that $q(1) = q(2) = q(3) = 1/3$.

(c) For the next two parts, we set $a = 1/6$, $b = 1/3$, and $c = 1/2$. Compute the exponential speed at which the $p_n(i)$ converge to $q(i)$ as n tends to infinity.

(d) Let T be the first time n superior or equal to zero such that $X_n = 1$. For $i = 1, 2, 3$, define $s(i) = E(T|X_0 = i)$ and compute $s(1)$ directly. To study $s(2)$ and $s(3)$, use an intermediary conditioning by the possible values of X_1 to obtain two linear equations verified by $s(2)$ and $s(3)$, and then solve this linear system.

20. Consider a Markov process U_n on the state space $S = (1, 2, 3)$, with transition matrix Q equal to

$$\begin{pmatrix} 1/2 & 0 & 1/2 \\ 0 & 1/2 & 1/2 \\ 1/2 & 1/2 & 0 \end{pmatrix}.$$

(a) Let T be the first (random) time such that $U_T = 3$. Define the conditional expectations $f(j) = E(T|U_0 = j)$ and note that $f(3) = 0$. Use the conditioning by the position reached at time 1 to derive a system of two linear equations verified by $f(1)$ and $f(2)$. Compute $f(1)$ and $f(2)$.

(b) Admit without calculations the following numerical results:
The eigenvalues of Q are $\mu_1 = -1/2$, $\mu_2 = 1/2$, and $\mu_3 = 1$, with associated eigenvectors equal to the columns of the matrix V given by

$$\begin{pmatrix} 1 & 1 & 1 \\ 1 & -1 & 1 \\ -2 & 0 & 1 \end{pmatrix}.$$

The inverse of V is the matrix $V^{-1} = \frac{1}{3}U$, where the matrix U is given by

$$\begin{pmatrix} 1/2 & 1/2 & -1 \\ 3/2 & -3/2 & 0 \\ 1 & 1 & 1 \end{pmatrix}.$$

(c) Let D be the 3×3 diagonal matrix with diagonal terms μ_1, μ_2, μ_3. Explain briefly why Q verifies the product formula $Q = VDV^{-1}$.

(d) Assume that the initial state U_0 of the Markov chain U_n is equal to 2. Use the result of part (a) to compute explicitly the probabilities $q(j, n) = P(U_n = j)$ for $j = 1, 2, 3$.

(e) Compute the limits $r(j) = \lim_{n\to\infty} q(j, n)$. What is the interpretation of the probabilities $r(j)$?

(f) For $j = 1, 2, 3$, call $G(j, N)$ the random number of instants $n \le N$ such that $U_n = j$. Verify that

$$G(j, n) = \sum_{n=0}^{N} 1_{U_n = j}$$

where 1_A is the indicator function of the event A.
Deduce from this formula that

$$E(G(j, n)) = \sum_{n=0}^{N} q(j, n).$$

21. Let W_n be a sequence of independent random variables having the same standard Gaussian distribution. Define a sequence of random variables U_n by $U_0 = W_0$ and $U_n = aU_{n-1} + W_n$ for all $n \ge 1$, where a is a fixed real number.

(a) For all n, compute the mean $b(n)$ and standard deviation $\sigma(n)$ of U_n, as well as the probability distribution and the characteristic function of U_n.

(b) Prove that if $|a| < 1$, the probability distribution of U_n has a limit as $n \to +\infty$, and compute this limit.

(c) For $a > 1$, define $V_n = \frac{U_n - b(n)}{\sigma(n)}$. Compute the characteristic function of V_n and the limit of the probability distribution of V_n as $n \to \infty$.

(d) For $a > 1$ and large n, determine $k(n)$ such that $\text{Prob}(|U_n| > k(n))$ is approximately equal to $3/1000$.

22. Let X and Y be random variables verifying $0 < X < 1$ and $0 < Y < 1$, with joint density function $f(x, y) = \frac{1}{c}(x^2 + y^2 + xy)$.

    (a) Compute the constant c, the means $E(X) = E(Y)$, and the covariance $\text{Cov}(X, Y)$.

    (b) Compute the density of X and the conditional density of Y given $X = x$.

    (c) Compute the density function of $V = X + Y$.

23. (Coin flipping problems.)

    (a) There are two coins. The first coin is fair. The second coin is such that $P(H) = 0.6 = 1 - P(T)$. You are given one of the two coins, with equal probabilities between the two coins. You flip the coin four times, and three of the four outcomes are H. What is the probability that your coin is the fair one?

    (b) A man has five coins in his pocket. Two are double-headed, one is double-tailed, and two are normal. The coins cannot be distinguished unless one looks at them.

        i. The man shuts his eyes, chooses a coin at random, and tosses it. What is the probability that the lower face of the coin is heads?

        ii. He open his eyes and sees that the upper face of the coin is a head. What is the probability that the lower face is a head?

        iii. He shuts his eyes again, picks up the same coin, and tosses it again. What is the probability that the lower face is a head?

24. Suppose that you are given a number of independent samples from a normal distribution with mean μ and variance σ^2.

    (a) Show that the sample mean is an unbiased estimate of the mean.

    (b) How would you compute the confidence interval for the mean? Give an interpretation of the confidence interval.

    (c) Give an unbiased estimator of the variance. Show that it is unbiased.

    (d) Does the variance of the estimator you gave in the previous part meet the Cramér-Rao lower bound?

25. A gambler systematically bets one dollar on a single number at roulette. The probability of a win is then $1/37$, and each time the gambler wins, the casino gives him a payback of 36 dollars (including his one dollar bet). When the gambler loses, his bet is lost. Call X_N the sum of all paybacks received by the gambler after N successive bets.

    (a) Give the classical formula for the probability $p(N, k) = P(X_N = 36k)$, where $0 \leq k \leq N$. Compute the mean and variance of X_N.

    (b) State the central limit theorem, and recall how it applies to X_N for N large.

    (c) Use (b) to evaluate, for N large, the probability g_N that after N bets, the gambler's loss L_N is larger than $\sqrt{N}$, and compute the limit g_∞ of g_N as $N \to \infty$.

    (d) Another gambler systematically bets one dollar on the color red, so that the probability of a win is $18/37$, and each win yields a payback of two dollars. As in (c), compute g_∞ for this case. For a gambler who cannot afford to lose more than $\sqrt{N}$ dollars after N bets, which of the two preceding gambling plans is the safest?

26. Let Z_n be a Markov chain on the state space $S = (1, 2)$ with transition matrix Q given by $Q(1, 1) = 1/3$, $Q(1, 2) = 2/3$, $Q(2, 1) = Q(2, 2) = 1/2$.
    (a) Assume that the initial distribution is given by $P(Z_0 = 1) = 1/4$ and $P(Z_0 = 2) = 3/4$. Explain why the probabilities $a(n) = P(Z_n = 1)$ and $b(n) = P(Z_n = 2)$ have limits as $n \to +\infty$, and compute these limits by solving a linear system.
    (b) For any finite n, compute explicitly the probabilities $a(n)$ and $b(n)$.
    (c) Assume now that $P(Z_0 = 2)$ is equal to 1. Let T be the first (random) time $T \geq 1$ such that $Z_T = 2$. For any $k \geq 1$, describe the event $T = k$ in terms of $Z_1, \ldots, Z_k$, and use this description to compute $P(T = k)$.
    (d) Use the result of (c) to compute $E(T)$.

27. (Convergence of random variables.)
    (a) Let X_n and X_0 be random variables on a probability space $(\Omega, \mathcal{F}, P)$, for $n \geq 1$. Give the definition of the following convergence:
        i. $X_n \to X_0$ almost surely;
        ii. $X_n \to X_0$ in probability;
        iii. $X_n \to X_0$ in distribution;
        iv. $X_n \to X_0$ in L_2.
    (b) Give an example showing that almost sure convergence does not imply convergence in L_2.

28. First explain what it means for an event to occur infinitely often (i.o.), and state the Borel-Cantelli Lemma. Then prove the following:

    Let $\mu = E[X_i]$, and $\sigma^2 = \text{Var}[X_i]$ for a sequence of i.i.d. random variables $\{X_i\}$. Assume that $|X_i - \mu| \leq M$ for all i, $\overline{X}_n = \frac{1}{n}\sum_{i=1}^{n} X_i$, and that Bernstein's inequality holds:

    $$P(|\overline{X}_n - \mu| \geq t) \leq 2\exp\left(-\frac{nt^2}{2(\sigma^2 + tM/3)}\right).$$

    Prove that there is a decreasing sequence t_n such that $P(|\overline{X}_n - \mu| \geq t_n \text{ i.o.}) = 0$.