Loi d'une variable discrète.

I.

Soit X la variable aléatoire dont la loi est donnée par le tableau ci-dessous.

k	-2	-1	0	3
$\mathbb{P}[X = k]$	1/6	1/6	1/3	1/3

1°. Justifier le fait qu'il s'agisse d'une loi de probabilité, déterminer $X(\Omega)$, $\mathbb{E}[X]$, $var(X)$ et $\sigma(X)$.
2°. Tracer la courbe représentative de sa fonction de répartition F.

II.

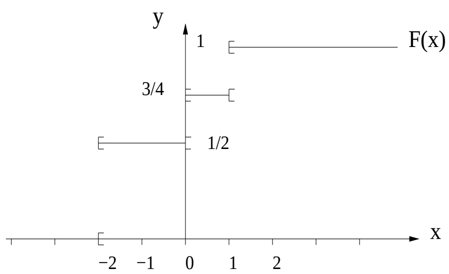
On considère la variable aléatoire X dont la loi est donnée par le tableau ci-dessous :

k	-1	0	1
$\mathbb{P}[X = k]$	1/4	1/4	1/2

1°. Déterminer l'espérance, la variance et l'écart type de X.
2°. Tracer la courbe représentative de sa fonction de répartition.
3°. On considère une variable aléatoire Y de même loi que X mais indépendante de celle-ci.
Déterminer la loi de $M = max(X, Y)$

III.

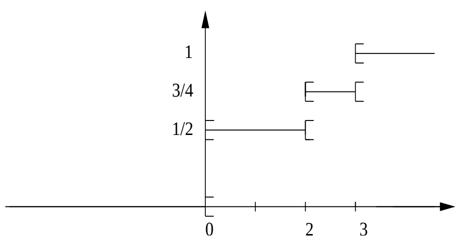
La figure ci-dessous représente la fonction de répartition d'une variable aléatoire X.



1°. Déterminer la loi de X.
2°. Calculer $\mathbb{E}[X]$, $var(X)$ et $\sigma(X)$.

IV.

On considère la fonction F donnée par la courbe représentative ci-dessous.
Soit X la variable aléatoire de fonction de répartition F.



1°. Déterminer la loi de X, calculer $\mathbb{E}[X]$, $var(X)$ et $\sigma(X)$.
2°. Calculer $\mathbb{P}[X \leq 5/2]$

V.

Une variable aléatoire X suit une loi équiprobable sur un ensemble à n éléments si la probabilité de chaque élément de l'ensemble est la même. C'est le modèle d'un jeu équitable.
Nous supposerons ici que $X(\Omega) = \{1, 2, ..., n\}$. On a alors $\mathbb{P}[X = k] = \dfrac{1}{n}\ \forall\ k = 1, .., n$

1°. Déterminer $\mathbb{E}[X]$ et $Var(X)$.
2°. On jette un dé tétraédrique dont les faces sont numérotées de 1 à 4.
On note X la variable aléatoire égale au numéro de la base. Calculer son espérance et son écart type.

VI.

X est une variable dont l'espérance et la variance valent 2 et $Y = 2X + 3$. Calculer $\mathbb{E}[Y]$ et $var(Y)$.

VII.

Une urne contient n jetons numérotés de 1 à n. On en tire une poignée et l'on note X_k la variable aléatoire égale à k si le jeton numéro k se trouve dans la poignée et 0 s'il ne s'y trouve pas. On note S la somme des numéros tirés dans la poignée.
1°. En considérant qu'une poignée est une partie de l'ensemble des jetons, déterminer la loi de X_k.
2°. Calculer $E[X_k]$ et $E[S]$.

VIII.

Soit $p \in [0, 1]$ et $q = 1 - p$.
Soit X une va pouvant prendre les valeurs 1 et -1 avec les probas p et q.
1°. Déterminer $\mathbb{E}[X]$ et $var(X)$.
2°. Soit Y une variable aléatoire de même loi que X et indépendante de celle ci.
Soit $S = X + Y$ la variable aléatoire égale à la somme de X et Y. Calculer $E[S]$ et $var(S)$.
Quelles sont les valeurs que peut prendre S ? Quelle est sa loi ? Retrouver $E[S]$ et $var(S)$ par le calcul.

IX. Loi géométrique.

On effectue une série de pile ou face indépendants dans lesquels la proba de faire pile est $p \in [0, 1]$. Cela revient à considérer une suite $X_1, X_2, ..., X_n, ...$ de variables aléatoires de Bernoulli indépendantes de paramètre p.

Soit N la variable aléatoire égale au rang du premier pile.
N est le nombre de lancers qu'il faut effectuer avant de voir apparaître pile pour la première fois.
On dit que N suit une loi géométrique de paramètre p et l'on se propose d'en étudier quelques propriétés.

1°. Quelles valeurs peut prendre N ?
2°. Soit $k \in \mathbb{N}^*$. Décrire par une phrase en français chacun des évènements suivants :
$[X_1 = 0]$, $[X_2 = 0]$, $[X_{k-1} = 0]$, $[X_k = 1]$ et $[N = k]$. En déduire une relation entre eux.
3°. Utiliser la question précédente pour calculer $\mathbb{P}[N = k]$ $\forall k \in \mathbb{N}^*$. Vérifier que la formule définit une loi de probabilité.
4°. Exprimer sous forme d'une somme de termes $\mathbb{P}[N \leq n]$ pour $n > 0$.
En déduire la probabilité qu'au cours du jeu pile n'apparaisse jamais.
5°. On sait que si $|x| < 1$, $\displaystyle\sum_{n\geq 0} x^n = \frac{1}{1-x}$
On admet de même que si $|x| < 1$, $\displaystyle\sum_{n\geq 1} nx^{n-1} = \frac{1}{(1-x)^2}$ et $\displaystyle\sum_{n\geq 2} n(n-1)x^{n-2} = \frac{2}{(1-x)^3}$
En déduire que si $p \neq 0$ alors $\mathbb{E}[N] = \dfrac{1}{p}$. Déterminer de même $\mathbb{E}[N(N-1)]$. En déduire $var(N)$.

6°. Applications.
6.1. Un terminal est relié à un réseau et souhaite émettre un message. La probabilité qu'il y ait une collision avec un message d'un autre terminal est de 0,1. En ce cas le terminal tente une réémission à l'instant suivant.
Quelle est la probabilité que le message passe avec succès à la 4^e tentative ? Avant la 3^e ?
Quel est le délai moyen d'attente d'un message dans ce terminal ? Pourquoi ce système pose-t-il des problèmes ?
6.2. Une urne contient 30% de boules rouges et 70% de boules blanches (on suppose que le nombre de boules est très grand). On tire des boules successivement, avec remise, et l'on s'arrête dès que l'on a tiré une boule rouge.
Quelle est la probabilité pour qu'il faille au plus quatre tirages ?

X. Loi hypergéométrique.

Considérons une urne de N boules contenant une proportion p de boules blanches et q de boules noires. On en extrait n simultanément et on appelle X la variable égale au nombre de boules blanches se trouvant parmi les n tirées.

On dit que X suit une loi hypergéométrique de paramètres $N \in \mathbb{N}$, $n \leq N$ et $p \in [0, 1]$ que l'on notera $\mathcal{H}(N, n, p)$.
Cette loi modélise les tirages effectués sans remise.

1°. Quelles sont les valeurs possibles de X (on notera M la plus grande de ces valeurs).
2° Montrer que $\mathbb{P}[X = k] = \dfrac{C_{Np}^k C_{Nq}^{n-k}}{C_N^n}$ si $k \in \{0, 1, 2, ..., M\}$
3°. On admet que si X suit une loi $\mathcal{H}(N, n, p)$, alors $\mathbb{E}[X] = np$ et $var(X) = npq\dfrac{N-n}{N-1}$.
Calculer $\mathbb{E}[X]$ et $var(X)$ lorsque $N \to +\infty$. À quoi cela correspond-t-il ?

4°. Dans un lot de 1000 composants, il y a une proportion 0,02 de composants défectueux.
On en prélève 10 et l'on note X le nombre de composants défectueux parmi les 10.
Calculer $\mathbb{P}[X = 0]$, $\mathbb{P}[X = 2]$, $\mathbb{P}[X \leq 2]$, $\mathbb{E}[X]$ et $\sigma(X)$.

XI.

Soient X et Y deux variables aléatoires indépendantes suivant une loi géométrique de paramètre respectif p_1 et p_2.
Soit Z la variable aléatoire définie par $Z = min(X, Y)$.

1°. Pour $k \geq 1$ calculer $\mathbb{P}[X \geq k]$ et $\mathbb{P}[Y \geq k]$.
2°. Exprimer l'évènement $[Z \geq k]$ en fonction des évènements $[X \geq k]$ et $[Y \geq k]$ et en déduire $\mathbb{P}[Z \geq k]$.
3°. Calculer alors $\mathbb{P}[Z = k]$ $\forall k \in \mathbb{N}^*$.

Loi binomiale et loi de Poisson.

XII.

Soit une variable aléatoire X suivant une loi de Poisson de paramètre $\lambda = 3$. Calculer $\mathbb{P}[X \leq 4/X \geq 2]$.
$\mathbb{P}(A/B)$ est la probabilité conditionnelle de A sachant B.

XIII.

1°. Une urne contient 40 boules dont 25 rouges. On en extrait 7 avec remise.
Calculer la proba pour qu'il y ait exactement 4 des 7 boules qui soient rouges.

2°. 10% des ordinateurs sortis d'une usine sont porteurs d'un composant défectueux. 5 ordinateurs sont choisis au hasard. Soit X le nombre d'ordinateurs défectueux de cet échantillon. Déterminer la loi de X puis la proba de l'évènement $A =$ « 2 ou 3 ordinateurs sont défectueux. »

3°. La proba d'une maladie infantile est de $1/100$. 500 naissances ont été enregistrées dans la ville. Quelle est la proba pour que l'on observe au moins 5 cas de cette maladie parmi les nouveaux-nés ?

4°. Le nombre de pannes annuelles d'une machine suit une loi de Poisson de paramètre $\lambda = 2$. Quelle est la probabilité pour que l'on observe au moins une panne dans l'année ?

5°. Un émetteur envoie des 0 et des 1 de façon indépendante sur un canal de communication. Avec probabilité $p = 0.9$, il n'y a pas d'erreur durant la transmission. Pour augmenter la fiabilité de la communication, on envoie sept 1 consécutifs au lieu d'un seul et on fait de même avec les 0. On considère qu'un 1 a été émis lorsqu'au moins quatre bits 1 sont reçus parmi les sept. Calculer la probabilité que la transmission d'un bit ait lieu sans erreur.

XIV.

Soit X une variable aléatoire suivant une loi de Poisson de paramètre λ.

1°. Calculer $\dfrac{\mathbb{P}[X = k+1]}{\mathbb{P}[X = k]}$ et en déduire un procédé de construction d'une table de Poisson.

2°. Montrer que l'application $\Lambda : \mathbb{N} \to [0,1]/k \to \mathbb{P}[X = k]$ est d'abord croissante puis décroît vers 0 quand $k \to +\infty$.
3°. On suppose que X vérifie $\mathbb{P}[X \leq 1] = 0,95$. Démontrer que λ est solution de l'équation $\ln(1 + x) - x = \ln 0,95$.
Étudier les variations de la fonction ϕ définie sur $[0, +\infty[$ par $\phi(x) = \ln(1 + x) - x$. En déduire λ à 0,1 près.

XV.

Afin de régler une liaison infrarouge, on dirige un rayon laser vers une cible à atteindre. D'une statistique préalable, on sait que la probabilité que ce rayon atteigne la cible est de 0,01. On fait l'expérience qui consiste à émettre n fois le rayon, les émissions étant deux à deux indépendantes et ayant même proba d'atteindre la cible.
Soit X la variable aléatoire égale au nombre de fois où la cible est atteinte.

1°. Déterminer la loi de X. Dans le cas où $n = 300$, calculer $\mathbb{E}[X]$ et $\sigma(X)$.

2°. Pour une expérience avec un nombre n d'essais, $n \geq 50$, on approche cette loi par une loi de Poisson.
2.1. Déterminer le paramètre λ de cette loi en fonction de n.
2.2. On estime que l'expérience est concluante lorsque le rayon atteint au moins trois fois la cible.
Donner les valeurs prises par X correspondant à l'évènement $A =$ « L'expérience est concluante. »
Exprimer la probabilité de cet évènement en fonction de λ.
2.3. Soit f la fonction définie pour x positif par $f(x) = (1 + x + \dfrac{x^2}{2})e^{-x}$
Étudier le sens de variation de f et calculer $f(6,1)$, $f(6,2)$ et $f(6,3)$
2.4. Quel est le nombre N d'essais à partir duquel la probabilité de l'évènement A est supérieure à 0,95 ?

XVI.

Une ligne de transmission transporte des fichiers, chaque fichier étant constitué de 10^5 bits.
La probabilité qu'un bit soit erroné est $p = 0,0001$ et l'on admet que les erreurs sont indépendantes les unes des autres. Soit X la variable aléatoire égale au nombre d'erreurs dans la transmission d'un fichier donné.

1°. Quelle est la loi de X ?

2°. Déterminer la probabilité qu'un fichier ne contienne aucune erreur.
3°. Déterminer le nombre moyen d'erreurs par fichier.
4°. On améliore la qualité de la transmission afin que $p = 10^{-5}$.
Par quelle loi peut-on approcher X ? Pourquoi ? Calculer alors $\mathbb{P}[X \leq 1]$.

XVII.

Le 14 janvier 2005, la sonde Cassini-Huygens a envoyé sur Terre les premières photos de la surface de Titan. Pour parvenir jusqu'à la terre, les signaux provenant de la sonde ont parcouru 1.5 milliards de kilomètres et ont voyagé durant plus d'une heure à une fréquence de 8.4 GHz. Ce trajet peut se modéliser par un canal très brouillé et les bits reçus sur la Terre peuvent donc être erronés. Pour un rapport signal sur bruit (RSB) de 0.7 dB (correspondant à un niveau de bruit très élevé) la probabilité qu'un bit arrive erroné est de $p = 5 \times 10^{-2}$.

Mais la sonde possède un code correcteur très puissant constitué de la concaténation d'un code convolutif avec un code de Reed-Solomon de longueur $n = 255$ et dimension $k = 223$. À l'aide de ce code, le taux d'erreur binaire tombe à une probabilité $p = 10^{-6}$ pour un rapport RSB de 0.7 dB.

Une image en noir et blanc de 500×635 pixels codée sur 64 niveaux de gris est transmise en direction de la Terre. Sa taille est donc de $n = 500 \times 635 \times 64 = 2 \times 10^7$ bits

1°. On suppose que la transmission ait lieu sans code correcteur. Soit X la variable aléatoire égale au nombre de bits erronés durant la transmission de cette image.
Quelle est la loi suivie par X ? Quels sont ses paramètres ? Quel est le nombre moyen d'erreurs par image transmise ?

2°. On suppose maintenant que la transmission utilise le code correcteur de Cassini. Pour un fichier de la même taille que l'image ci-dessus, quels sont alors les paramètres de la loi de X ? Quel est le nombre moyen d'erreurs par image transmise ?
Calculer $\mathbb{P}[X \leq 10]$, $\mathbb{P}[15 \leq X \leq 25]$ et $\mathbb{P}[X \leq 20]$

3°. On transmet une image plus petite de taille 2×10^6 bits. Quel est le nombre moyen d'erreurs ?
Calculer $\mathbb{P}[X \geq 1]$ et $\mathbb{P}[X \leq 2]$

Lois conjointes.

XVIII.

Soient X et Y deux variables aléatoires dont la loi conjointe est donnée par le tableau ci-dessous

$X \setminus Y$	0	1	TOT
0	$\frac{1}{4}$	$\frac{1}{2}$	$\frac{3}{4}$
1	$\frac{1}{4}$	0	$\frac{1}{4}$
TOT	$\frac{1}{2}$	$\frac{1}{2}$	1

1°. Déterminer les lois marginales en X et Y
2°. Ces variables sont-elles indépendantes ?
3°. Calculer $cov(X, Y)$

XIX.

On considère un couple (X,Y) de variables aléatoires dont la loi conjointe est donnée par le tableau ci-dessous

$X \setminus Y$	0	1	TOT
1	$\frac{1}{3}$	0	$\frac{1}{3}$
-1	$\frac{1}{3}$	$\frac{1}{3}$	$\frac{2}{3}$
TOT	$\frac{2}{3}$	$\frac{1}{3}$	1

1°. Déterminer les marginales en X et Y. Calculer $\mathbb{E}[X]$, $\mathbb{E}[Y]$ et $cov(X, Y)$. X et Y sont-elles indépendantes ?
2°. On pose $S = X + Y$. Déterminer la loi de S, son espérance et sa variance.

XX.

Soit (U, V) un couple de variables aléatoires dont la loi est donnée par le tableau ci-après :

$U \setminus V$	-2	0	2	TOT
-2	0	1/4	0	1/4
0	1/4	0	1/4	1/2
2	0	1/4	0	1/4
TOT	1/4	1/2	1/4	1

1°. Déterminer les lois marginales en U et V ainsi que leur espérance et leur variance.
2°. Déterminer la covariance $Cov(U, V)$. U et V sont-elles indépendantes ?
3°. On pose $X = \frac{1}{2}(U + V)$ et $Y = \frac{1}{2}(U - V)$
Déterminer les valeurs possibles du couple (X, Y) en fonction des valeurs de U et V.
En déduire la loi du couple (X, Y), puis les lois marginales de X et Y.
4°. Calculer $\mathbb{E}[XY]$ et en déduire $Cov(X, Y)$.
X et Y sont-ils indépendants ?

XXI.

Soit (X, Y) un couple de variables aléatoires dont la loi est donnée par le tableau ci-après :

$X\backslash Y$	-1	0	1	TOT
-1	1/6	0	1/6	1/3
1	1/6	1/3	1/6	2/3
TOT	1/3	1/3	1/3	1

1°. Déterminer les lois marginales en X et Y ainsi que leur espérance et leur variance.
2°. Déterminer la covariance $Cov(X, Y)$. X et Y sont-elles indépendantes ?
3°. On pose $U = XY$ et $V = Y/X$.
Déterminer la loi du couple (U, V). U et V sont-elles indépendantes ?

Applications aux télécommunications.

XXII. Protocoles Aloha et Ethernet.

On considère un réseau composé de N stations dans lequel le temps a été discrétisé en instants entiers $t = 1, 2, 3, ...$ Chaque station a une probabilité $\alpha \in [0, 1]$ de voir apparaître un message qu'elle essaie aussitôt d'émettre sur le réseau. Si plusieurs messages sont simultanément sur le réseau, il y a collision et ceux-ci sont détruits ; les stations devront les réémettre plus tard. Si un seul message est présent, il est transmis avec succès et disparait du réseau.

Lorsqu'un message est pris dans une collision et qu'il n'est donc pas transmis avec succès, la station dont il provient attend un temps aléatoire avant d'essayer de le réémettre.En cas de nouvel échec, elle recommence.

On supposera pour simplifier qu'une station ne peut recevoir de nouveaux messages tant qu'elle n'a pas émis celui qui est en attente. Le protocole décrit ci-dessus est Aloha (inventé vers 1968 à l'université d'Hawai).

1°. On suppose qu'à un instant t le canal est silencieux. Soit X le nombre de messages arrivant sur le canal à l'instant $t + 1$. Quelle est la loi de X ? Quelle est la probabilité qu'il y ait une collision à cet instant ?

2°. Un message est pris dans une collision. Quelle est la probabilité qu'il tente de réémettre une première fois à l'instant t+1 ? t+2 ? Avant t+4 ? Quel sera en moyenne le temps de présence de ce message sur le canal, en supposant qu'aucun autre message n'apparaisse ou ne tente d'émettre ?

3°. L'état du canal à l'instant t est le suivant : k messages sont en attente. Quelle est la probabilité qu'un nouveau message arrive et soit transmis avec succès ? Quelle est la probabilité qu'un des messages en attente soit transmis avec succès à t ? Quelle est la probabilité d'une nouvelle collision entre messages ?

4°. Lorsque n est grand et que p est petit, on peut considérer que X suit une loi de Poisson de paramètre $\lambda = np$, λ représentant la fréquence des arrivées.
4.1. Si $\lambda = 0,1$, quelle est la proba qu'à un instant donné plus d'un message arrive sur le canal ? Que se passe-t-il alors ?
4.2. Soit $N(t)$ la variable aléatoire représentant le nombre de messages présents à t (en attente ou arrivant) sur le canal. Calculer $\mathbb{P}[N(t + 1) = k - 1/N(t) = k]$, $\mathbb{P}[N(t + 1) = k/N(t) = k]$ et $\mathbb{P}[N(t + 1) = k + i/N(t) = k]$, $\forall\ i \in \{2, 3, ..., n - k\}$ en expliquant au préalable ce que signifient ces évènements.

5°. On décide d'un nouveau protocole : en cas de collision, on réessaye à chaque instant suivant et jusqu'au succès final avec une proba $p \in [0, 1]$ où p est divisé par 2 à chaque nouvelle collision. Ce protocole s'appelle Ethernet et est une version simplifiée de CSMA-CD.
5.1. On s'intéresse à un message M_1 arrivant sur le canal à t et entrant en collision avec un ou plusieurs autres messages. Quelle est la probabilité que M_1 tente une réémission à t+1 ? t+2 si l'on sait qu'une collision s'est produite à t+1 ?
5.2. Deux messages entrent en collision. Quelle est la proba qu'ils tentent une première réémission en même temps ?

XXIII. Entropie et information.

Nous reprenons la notion de quantité d'information vue dans la leçon précédente. Si X est une variable aléatoire prenant comme valeurs $\{x_1, ..., x_n\}$ et si A_k est l'évènement $A_k = [X = x_k]$ pour $k = 1, ..., n$, alors la quantité d'information de l'évènement A_k est

$$I(A_k) = -\log \mathbb{P}(A_k) = -\log \mathbb{P}[X = x_k]$$

Pour simplifier les notations, on posera $p_k = \mathbb{P}[X = x_k]$.

L'entropie $H(X)$ d'une variable aléatoire X est la valeur moyenne de la quantité d'information sur toutes les réalisations de X :

$$H(X) = \mathbb{E}[I(X)] = -\sum_{k=1}^{n} p_k \log p_k = -\sum_{k=1}^{n} \mathbb{P}[X = x_k] \log \mathbb{P}[X = x_k]$$

$H(X)$ mesure l'incertitude moyenne de la variable X.

1°. Montrer que $H(X) \geq 0$. À quelle condition a-t-on $H(X) = 0$?
2°. Montrer que si X suit une loi uniforme sur $\{x_1, ..., x_n\}$, alors $H(X) = \log_2 n$
3°. Soient $(p_1, ..., p_n)$ et $(q_1, ..., q_n)$ deux mesures de probabilités. Montrer que $\displaystyle\sum_{k=1}^{n} p_k \log_2 \frac{q_k}{p_k} \leq 0$
4°. En déduire que $H(X) \leq \log_2 n \ \forall X$. En déduire la valeur maximale de l'entropie. Interprétation ?
5°. Si X suit une loi de Bernouilli, calculer $H(X)$. Étudier la fonction $p \to H(X)$.
6°. Calculer $H(X)$ lorsque X suit une loi de Poisson de paramètre λ.

XXIV. Entropie et information mutuelle.

Considérons un couple de variables aléatoires (X, Y) dont l'ensemble des valeurs est $X(\Omega) = \{x_1, ..., x_n\}$ et $Y(\Omega) = \{y_1, ..., y_m\}$. On pose $\mathbb{P}[X = x_k; Y = y_i] = \mathbb{P}(x_k, y_i)$ et $\mathbb{P}[X = x_k/Y = y_i] = \mathbb{P}(x_k/y_i)$

On définit alors l'entropie conditionnelle $H(X/Y)$ et l'entropie conjointe $H(X, Y)$ par

$$H(X/Y) = -\sum_{i,k} \mathbb{P}(x_k/y_i) \log \mathbb{P}(x_k, y_i) \quad \text{et} \quad H(X;Y) = -\sum_{i,k} \mathbb{P}(x_k, y_i) \log \mathbb{P}(x_k, y_i)$$

$$I(X;Y) = \sum_{i,k} \mathbb{P}(x_k, y_i) \log \left(\frac{\mathbb{P}(x_k, y_i)}{\mathbb{P}(x_k)\mathbb{P}(y_i)} \right)$$

1°. Montrer que $H(X, Y) = H(Y, X)$ et $I(X;Y) = I(Y;X)$
2°. Montrer que $H(X, Y) = H(X) + H(Y/X) = H(Y) + H(X/Y)$. Montrer que $I(X;Y) = H(X) - H(X/Y) = H(Y) - H(Y/X)$
3°. Montrer que $f(x) = x \ln x$ est une fonction convexe.
4°. En déduire que $H(X/Y) \leq H(X) \leq H(X, Y)$, $\forall X, Y$.
5°. Montrer que $H(X, Y) \leq H(X) + H(Y)$ avec égalité si et seulement si X, Y sont indépendants.

XXV. Capacité d'un canal de communication.

Un canal de communication transmet des symboles (bits ou autres) depuis une source vers un destinataire. Le canal est sans mémoire si la transmission d'un bit ne dépend pas du précédent (la sortie ne dépend que de l'entrée). Le canal est sans bruit si la sortie correspond à l'entrée. En général, un canal est bruité et le bruit perturbe la communication : il peut effacer, atténuer ou changer la valeur des données.

Le but du codage de canal est de protéger les données contre le bruit. Ceci peut se faire à l'aide des codes correcteurs. En 1948, Shannon a démontré que tout canal possède une quantité C appelée sa capacité, telle qu'il est possible de transmettre sans erreurs à condition d'avoir un débit (en bits par seconde) inférieur à C. Le calcul et la connaissance de C sont donc très importants. On pose

$$C = \max_{\mathbb{P}(X)} I(X;Y)$$

C'est la valeur maximale de l'information mutuelle en sortie du canal et le maximum est calculé sur toutes les probabilités possibles de X en entrée.

1°. Si le canal transmet n valeurs et est sans perte, montrer que $C = \log_2 n$
2°. Le canal binaire symétrique BSC :
X et Y sont des variables aléatoires sur $\{0, 1\}$. La probabilité d'erreur est p (on l'appelle vraisemblance du canal) :

$\mathbb{P}[Y = 0/X = 1] = \mathbb{P}[Y = 1/X = 0] = p$

On supposera que $\mathbb{P}[X = 0] = \mathbb{P}[X = 1] = 1/2$

Déterminer la loi de (X, Y), calculer $H(Y)$ et $H(Y/X)$; en déduire C.
Etudier la fonction $C(p)$. Interprétation ?

3°. Le canal à effacement BEC :
On pose $\mathbb{P}[Y = 0/X = 1] = \mathbb{P}[Y = 0/X = -1] = \epsilon$ et $\mathbb{P}[X = 1] = \mathbb{P}[X = -1] = 1/2$
Calculer $H(Y)$ et $H(Y/X)$. En déduire C.

Approfondissements.

Les exercices qui suivent sont plus difficiles et proviennent des annales de l'IRMAR. Certains nécessitent l'utilisation de séries entières (se reporter au TD correspondant pour plus de précisions).

XXVI. Fonction génératrice d'une probabilité.

On appelle fonction génératrice d'une variable aléatoire X la fonction G définie par :

$$\begin{array}{rcl} G : [-1,1] & \longrightarrow & \mathbb{R} \\ x & \mapsto & G(x) = \sum_{k=0}^{+\infty} \mathbb{P}[X=k]\ x^k \end{array}$$

G est donc la somme d'une série entière de rayon de convergence ≥ 1.

1°. Justifier le fait que G est de classe C^∞ sur $]-1,1[$.

2°. Déterminer $G(x)$ lorsque X suit une loi de Bernoulli de paramètre p, puis une loi binomiale de paramètres n et p. Même question si X suit une loi géométrique de paramètre p.

3°. Démontrer que $\mathbb{E}[X] = \lim\limits_{x \to 1^-} G'(x)$ et en déduire les espérances des lois usuelles ci-dessus.

4°. On admet que si X et Y sont deux variables aléatoires indépendantes de fonction génératrice G et H, alors la fonction génératrice de $X+Y$ est la fonction $G(x)H(x)$. Déterminer alors la loi suivie par la somme de deux variables aléatoires indépendantes de loi de Poisson de paramètre respectif λ et μ, puis la loi suivie par deux variables indépendantes de loi binomiale de paramètres respectifs (n,p) et (m,p).

5°. Démonstration de la loi des évènements rares :

Soit Y une variable aléatoire de loi de Poisson de paramètre λ et de fonction génératrice H.
Soit X une variable aléatoire de loi binomiale de paramètres n et p et de fonction génératrice G_n.
On suppose que $\dfrac{\lambda}{n} \sim p$ quand $n \to +\infty$ (ie $\lim\limits_{n\to+\infty} np = \lambda$)
Démontrer que $\forall x \in [-1,1]$, $\lim\limits_{n\to+\infty} G_n(x) = H(x)$. Conclure.

XXVII. Fonction caractéristique.

On définit la fonction caractéristique d'une probabilité par $\phi_X(t) = \mathbb{E}[e^{itX}] = \displaystyle\sum_{k=0}^{+\infty} e^{itx_k}\mathbb{P}[X=x_k]$

C'est la série de Fourier dont les coefficients sont donnés par la loi de X

1°. Calculer $\phi_X(t)$ dans le cas où X suit une loi de Bernoulli de paramètre p, une loi binomiale de paramètres n et p, une loi de Poisson de paramètre λ et enfin une loi géométrique de paramètre p.

2°. Montrer que $|\phi_X(t)| \leq 1\ \forall t \in \mathbb{R}$

3°. Montrer que sous réserve d'existence $\mathbb{E}[X] = \phi'_X(0)$

4°. Montrer que si X et Y sont deux variables discrètes indépendantes, $Z = X+Y$ a pour fonction caractéristique $\phi_Z(t) = \phi_X(t)\phi_Y(t)$

XXVIII.

Soient X et Y deux vaiid données par $\mathbb{P}[X=n] = \mathbb{P}[Y=n] = pq^n$ avec $p+q=1, n \in \mathbb{N}$
On pose $Z = sup(X,Y)$ et on suppose X et Y indépendantes.

1°. Déterminer la loi du couple (X,Y)

2°. Déterminer la loi de Z, $\mathbb{E}[Z]$ et $var(Z)$ (on pourra développer $\dfrac{1}{1-x}$ en série entière autour de 0)

XXIX.

1°. On considère le couple de variables aléatoires (X,Y) dont la loi est donnée par $\mathbb{P}[X=k, Y=i] = q^k p^i$ pour $k, i \geq 1$. Déterminer les lois marginales en X et Y. Ces variables sont-elles indépendantes ? Calculer $Cov(X,Y)$

2°. Déterminer le développement en série entière de la fonction $\dfrac{1}{1-x}$ et préciser son rayon de convergence.

3°. Soit $\phi(x) = \dfrac{1}{(1-x)^{k+1}}$. Démontrer que $\phi(x) = \displaystyle\sum_{n\geq k} C_n^k x^{n-k}$ et préciser son rayon de convergence

4°. On considère un couple (X,Y) dont la loi est donnée par $\mathbb{P}[X=k, Y=n] = C_n^k p^{k+1} q^{2n-k}$ pour $\boxed{0 \leq k \leq n}$ avec $0 < p < 1$ et $q = 1-p$

Déterminer les lois marginales en X et Y.
Déterminer $\mathbb{P}[X=k/Y=n]$ qui est la loi conditionnelle de X sachant $[Y=n]$. Reconnaître cette loi.

5°. Les variables X et Y sont-elles indépendantes ?

XXX.

Soit X le nombre d'appels (indépendants) reçus à un standard téléphonique d'une entreprise durant un laps de temps fixé. Soit Y le nombre d'appels vers le service technique durant le même laps de temps.
Chaque appel arrivant au standard a une probabilité p de demander le service technique.
Soit $n \in \mathbb{N}$, $\lambda > 0$, $q = 1 - p$. On admet que la loi conjointe de (X, Y) est donnée par

$$\mathbb{P}[X = n, Y = k] = e^{-\lambda}\frac{\lambda^n}{n!}C_n^k p^k q^{n-k}$$

Cette formule étant valable pour $0 \leq k \leq n$

1°. Donner le développement en série entière de la fonction e^x en précisant le rayon de convergence.
En déduire $\displaystyle\sum_{n=k}^{+\infty} \frac{x^{n-k}}{(n-k)!}$

2°. Déterminer les marginales en X et Y ainsi que la probabilité $\mathbb{P}[Y = k/X = n]$
3°. Interpréter les résultats obtenus.

XXXI.

1°. Soit (X, Y) un couple de variables aléatoires dont la loi jointe est donnée par :
$\mathbb{P}[X = k, Y = i] = q^k p^i$ pour $i, k \in \mathbb{N}^*$
Déterminer les lois marginales en X et Y et en déduire que ces variables aléatoires sont indépendantes.
Calculer $Cov(X, Y)$ et en déduire $\mathbb{E}[XY]$

2°. On considère maintenant deux variables X et Y indépendantes, suivant chacune une loi géométrique de paramètre p. On pose $S = X + Y$

Démontrer que $[S = n] = \displaystyle\bigcup_{k=1}^{n-1} ([X = k] \cap [Y = n - k])$ et en déduire que $\mathbb{P}[S = n] = (n - 1)p^2q^{n-2}$ pour $n \geq 1$

3°. Démontrer que $\mathbb{P}[X = k/S = n] = \dfrac{1}{n-1}$ pour $k = 0, ..., n - 1$. Interpréter ce résultat.

4°. Cette question est plus difficile.
On considère n variables indépendantes $X_1, ..., X_n$ de même loi géométrique de paramètre p et on pose $S_n = \sum_{k=1}^{n} X_k$
Démontrer par récurrence sur n que

$$\mathbb{P}[S_n = k] = C_{k-1}^{n-1}p^n q^{k-n}$$

avec $k \geq n$
En se souvenant que X_k indique le nombre d'essais nécessaires avant le premier succès durant une succession de pile ou face indépendants, donner une interprétation de S_n
La loi de S_n s'appelle loi binomiale négative.

XXXII.

Un sac contient initialement une boule blanche et deux rouges. On répète indéfiniment l'opération qui consiste à tirer une boule, la remettre dans le sac si elle est blanche et l'éliminer si elle est rouge.
Soit X_n la variable aléatoire qui vaut 1 si l'on a tiré une boule blanche au n-ième tirage et 0 dans le cas contraire.
Soit T_1 l'instant où l'on tire la première boule rouge et T_2 celui où l'on tire la seconde boule rouge.

1°. Calculer $\mathbb{P}[T_1 = n]$, déterminer la fonction génératrice $g(s) = \displaystyle\sum_{n=1}^{+\infty} \mathbb{P}[T_1 = n]s^n$ de T_1 ainsi que $E[T_1]$ et $var(T_1)$
2°. Calculer $\mathbb{P}[T_1 = n, T_2 = k]$ pour $1 \leq n < k$
3°. En déduire $\mathbb{P}[T_2 = k]$ ainsi que la fonction génératrice $h(s) = \displaystyle\sum_{k=1}^{+\infty} \mathbb{P}[T_2 = k]s^k$ de T_2
4°. Calculer $\mathbb{E}[T_2]$ et en déduire $\mathbb{P}[X_n = 0]$ et déterminer la loi de X_n

XXXIII.

Soient $X_1, X_2, ..., X_n$ n vaiid suivant une loi de Bernouilli de paramètre p
Soit $S_n = X_1 + X_2 + ... + X_n$ pour $n \geq 1$
Soit $Y_1 = X_1$ et $Y_k = 1_{[X_{k-1}=0] \cap [X_k=1]}$ pour $k \geq 2$
Soit enfin T_n le nombre de suites ininterrompues de 1 dans la suite $(X_n)_{n\geq 1}$

1°. Rappeler la loi de S_n, son espérance et sa variance
2°. Déterminer la loi de Y_k, son espérance et sa variance
3°. Démontrer que Y_k et Y_{k+1} ne sont pas indépendantes mais que Y_k et Y_{k+l} le sont si $l \geq 2$
4°. Calculer $cov(Y_n, Y_{n+1})$
5°. Montrer que $T_n = Y_1 + ... + Y_n$ et calculer $\mathbb{E}[T_n]$

XXXIV.

Soit $(X_n)_{n\geq 1}$ une suite de va de Bernoulli de paramètre p indépendantes définies sur un même espace
On pose $T_1 = \{n \in \mathbb{N}/\ X_n = 1\}$ et $T_2 = \{n \in \mathbb{N}/\ n > T_1, X_n = 1\}$

1°. Déterminer $\mathbb{P}[T_2 - T_1 = k/T_1 = n]$
2°. Montrer que T_2 et $T_2 - T_1$ sont indépendantes et de même loi.
3°. En déduire la loi de T_2, son espérance et sa variance

XXXV.

Les objets manufacturés par une usine sont numérotés de 1 à n dans leur ordre de production. Ils sont fabriqués dans les mêmes conditions de manière indépendante avec pour chacun d'eux une probabilité p d'être défectueux.
Indépendamment de la fabrication, on procède à un contrôle des objets de la manière suivante : on fait un tirage aléatoire simulant le lancement d'un dé à six faces équilibré. On contrôle l'objet si on obtient un 1.
On désigne par S_n la variable aléatoire égale au nombre d'objets défectueux et par T_n le nombre d'objets contrôlés qui sont défectueux parmi les n premiers objets.

1°. Quelles sont les lois de S_1 et T_1 ?
2°. Déterminer pour $n \geq 2$ les lois de S_n et T_n
3°. Calculer l'espérance de chacune de ces deux variables aléatoires
4°. On considère la variable aléatoire N égale au numéro du premier objet défectueux et contrôlé. On convient que $N = +\infty$ s'il n'y en a aucun (la proba de cet évènement est supposée nulle).

Calculer $\mathbb{P}[N = 1]$ et $\mathbb{P}[N = 2]$, puis $\mathbb{P}[N = n]$ pour $n \geq 3$. Calculer l'espérance de N

5°. Soit U la variable aléatoire égale à 0 si $N = 1$ ou ∞ et à S_{n-1} si $N = n$
Ainsi, U est égale au nombre d'objets défectueux non décelés par le contrôle avant que n'apparaisse le premier objet défectueux et contrôlé.

Calculer $\mathbb{P}[U = 0, N = n]$ pour $n = 1$ puis pour $n \geq 2$. En déduire $\mathbb{P}[U = 0]$
Calculer $\mathbb{P}[U = 1, N = n]$ (distinguer le nième objet des précédents). En déduire $\mathbb{P}[U = 1]$
Calculer $\mathbb{P}[U = k, N = n]$ puis en déduire $\mathbb{P}[U = k]$ $\forall k \in \mathbb{N}$

6°. Que vaut $\mathbb{E}[U]$?

XXXVI.

Soient X, Y deux variables aléatoires indépendantes de même loi géométrique de paramètre p.

1°. En utilisant la définition de la somme de deux variables aléatoires, déterminer la loi de $S = X + Y$.
2°. Montrer que la loi conditionnelle de X sachant S est uniforme. C'est-à-dire que

$$\mathbb{P}[X = k/S = n] = \frac{1}{n-1} \quad \forall k \in \{1, ..., n-1\}$$

3°. Déterminer $\mathbb{P}[X > t]$ et $\mathbb{P}[X \geq t]$
4°. Déterminer $\mathbb{P}[X > t + s/X > s]$ et en déduire que X est sans mémoire.
5°. Soit $U = \min(X, Y)$. Déterminer la loi de U.

XXXVII.

On considère deux variables aléatoires indépendantes X et Y de même loi de Poisson de paramètre respectif λ et μ.

1°. Démontrer que $S = X + Y$ suit une loi de Poisson de paramètre $\lambda + \mu$.
2°. Déterminer la loi de X sachant $S = k$.
3°. Retrouver les résultats précédents en utilisant les fonctions génératrices.

XXXVIII.

Soient X, Y deux variables aléatoires indépendantes de loi géométrique de paramètre $p \in]0, 1[$.

1°. Déterminer les fonctions génératrices $\phi_U(x)$, $\phi_V(x)$ et $\phi_S(x)$ de $U = X - 1$, $V = Y - 1$ et $S = U + V$.
2°. Déterminer le rayon de convergence de ces trois séries entières.
3°. Démontrer que

$$\phi_S(x) = \frac{p^2}{q}\frac{d}{dx}\left(\frac{1}{1-qx}\right)$$

4°. Au cours d'une expérience, un évènement A se réalise avec probabilité $p \in]0, 1[$. On répète cette expérience de façon indépendante et l'on définit la variable Y_i comme suit : $Y_i = 1$ si A se produit et 0 sinon.
Soit N_1 le nombre d'expériences avant que A se réalise une première fois. Déterminer la loi de N_1.

5°. Soit N_2 le nombre d'expériences avant que A se réalise une seconde fois. Montrer que N_1 et $N_2 - N_1$ sont indépendantes et de même loi. En déduire la loi de N_2.