MATHEMATIQUES TD N°3 : VARIABLES DISCRETES - Corrigé.

R&T Saint-Malo - 2nde année - 2011/2012

Loi d'une variable discrète

On notera :
$X(\Omega)$ l'ensemble des valeurs possibles de X
$\mathbb{E}[X]$ l'espérance de X
$\sigma(X)$ l'écart type de X
$F(x)$ la fonction de répartition de la variable aléatoire X

I.

1°. $X(\Omega) = \{-2; -1; 0; 3\}$
$\mathbb{E}[X] = -2 \times \frac{1}{6} - \frac{1}{6} + 0 \times \frac{1}{3} + 3 \times \frac{1}{3} = \frac{1}{2}$
$\mathbb{E}[X^2] = (-2)^2 \times \frac{1}{6} - \frac{1}{6} + 0^2 \times \frac{1}{3} + 3^2 \times \frac{1}{3} = \frac{23}{6}$
$Var(X) = \mathbb{E}[X^2] - \mathbb{E}[X]^2 = \frac{43}{12}$ et $\sigma(X) = \frac{1}{2}\sqrt{\frac{43}{3}}$

2°. $F(x) = \mathbb{P}[X \leq x]$

La fonction F est une fonction constante par intervalle donnée par :

$$\begin{cases} F(x) = 0 & \text{si} \quad x < -2 \\ F(x) = 1/6 & \text{si} \quad -2 \leq x < -1 \\ F(x) = 1/3 & \text{si} \quad -1 \leq x < 0 \\ F(x) = 2/3 & \text{si} \quad 0 \leq x < 3 \\ F(x) = 1 & \text{si} \quad x \geq 3 \end{cases}$$

II.

1°. $X(\Omega) = \{-1; 0; 1\}$ $\mathbb{E}[X] = 1/4$, $\mathbb{E}[X^2] = 3/4$,
$Var(X) = \mathbb{E}[X^2] - \mathbb{E}[X]^2 = 11/16$ et $\sigma(X) = \sqrt{11}/4$

2°. $F(x) = \mathbb{P}[X \leq x]$
La fonction F est une fonction constante par intervalle donnée par :

$$\begin{cases} F(x) = 0 & \text{si} \quad x < -1 \\ F(x) = 1/4 & \text{si} \quad -1 \leq x < 0 \\ F(x) = 1/2 & \text{si} \quad 0 \leq x < 1 \\ F(x) = 1 & \text{si} \quad x \geq 1 \end{cases}$$

3°. $X(\Omega) = \{-1; 0; 1\}$
$\mathbb{P}[M = -1] = \mathbb{P}[X = -1] \times \mathbb{P}[Y = -1] = 1/16$
$\mathbb{P}[M = 1] = \mathbb{P}[X = 1] + \mathbb{P}[Y = 0] - \mathbb{P}[X = 1]\mathbb{P}[Y = 1] = 12/16$ $\mathbb{P}[M = 0] = 1 - 13/16 = 3/16$

III.

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

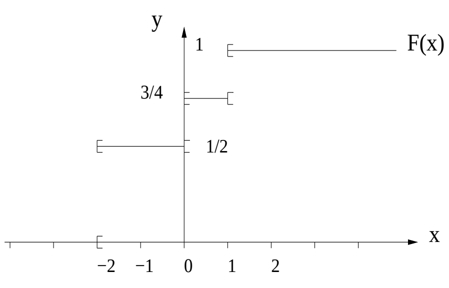

FIGURE 1 –

1°. Déterminer la loi de X.

$\mathbb{P}[X = k] \neq 0 \iff k$ point de discontinuité de F et $\mathbb{P}[X = k] = F(k^+) - F(k^-)$
Ainsi, $\mathbb{P}[X = -2] = \frac{1}{2}$, $\mathbb{P}[X = 0] = \frac{1}{4}$, $\mathbb{P}[X = 1] = \frac{1}{4}$

2°. Calculer $\mathbb{E}[X]$, $var(X)$ et $\sigma(X)$
$\mathbb{E}[X] = -\frac{3}{4}$, $\mathbb{E}[X^2] = \frac{9}{4} \Rightarrow Var(X) = \frac{27}{16}$ et
$\sigma(X) = \frac{3\sqrt{3}}{4}$

IV.

$\mathbb{E}(X) = 5/4$, $\mathrm{var}(X) = 27/16$, $\sigma(X) = 3\sqrt{3}/4$ et $\mathbb{P}[X \leq 5/2] = 3/4$

V.

$\mathbb{P}[X = k] = \frac{1}{n} \ \forall \ k = 1, .., n$

1°. Déterminer $\mathbb{E}[X]$ et $Var(X)$

$\mathbb{E}[X] = \frac{1}{n}(1 + 2 + ... + n) = \frac{n+1}{2}$
$\mathbb{E}[X^2] = \frac{1}{n}(1^2 + 2^2 + ... + n^2) = \frac{1}{n}\frac{n(n+1)(2n+1)}{6}$
$= \frac{(n+1)(2n+1)}{6}$
$\Rightarrow Var(X) = \mathbb{E}[X^2] - \mathbb{E}[X]^2 = \frac{n+1}{2}\left(\frac{2n+1}{3} - \frac{n+1}{2}\right) = \frac{n^2-1}{12}$

2°. Pour le dé tétraédrique, on applique les résultats ci-dessus avec $n = 4$. On a alors, $\mathbb{E}[X] = 5/2$ et $Var(X) = 5/4$

VI.

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

$\mathbb{E}[Y] = \mathbb{E}[2X + 3] = 2\mathbb{E}[X] + 3 = 7$
$Var(Y) = Var(2X + 3) = Var(2X) = 4Var(X) = 8$

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.

Puisqu'on ne connaît pas le nombre de jetons dans la poignée, il faut la considérer comme une partie de l'ensemble des jetons. Ainsi, $Card\ \Omega = 2^n$

$[X_k = k]$="la poignée contient le jeton n°k"
$\Rightarrow \ Card[X_k = k] = 2^{n-1}$ donc $\mathbb{P}[X_k = k] = \frac{2^{n-1}}{2^n} = \frac{1}{2}$

2°. Calculer $E[X_k]$ et $E[S]$

$\Rightarrow \mathbb{E}[X_k] = 0 \times \mathbb{P}[X_k = 0] + k \times \mathbb{P}[X_k = k] = \frac{k}{2}$
$\mathbb{E}[S] = \mathbb{E}[X_1 + ... + X_n] = \sum_{k=1}^{n} \mathbb{E}[X_k] = \frac{1}{2}\sum k = \frac{n(n+1)}{4}$

VIII.

Soit $p \in [0, 1]$ et $q = 1 - p$.
Soit S une va pouvant prendre les valeurs 1 et -1 avec les probas p et q.

1°. Déterminer $\mathbb{E}[X]$ et $var(X)$

$\mathbb{E}[X] = \mathbb{P}[X = 1] - \mathbb{P}[X = -1] = p - q$

$\mathbb{E}[X^2] = \mathbb{P}[X=1] + \mathbb{P}[X=-1] = p+q = 1$
$Var(X) = 1-(p-q)^2 = 4pq$ et $\sigma(X) = 2\sqrt{pq}$

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 ? Retrouver $E[S]$ et $var(S)$ par le calcul.

$\mathbb{E}[S] = \mathbb{E}[X+Y] = \mathbb{E}[X] + \mathbb{E}[Y] = 2(p-q)$
$Var(S) = Var(X) + Var(Y) = 8pq$ par indépendance de X et Y

S peut prendre comme valeurs $-2, 0, 2$ et sa loi est :
$\mathbb{P}[S=-2] = \mathbb{P}([X=-1] \cup [Y=-1]) =$
$\mathbb{P}[X=-1] \times \mathbb{P}[Y=-1] = q^2$ par indépendance de X et Y $\mathbb{P}[S=2] = \mathbb{P}([X=1] \cup [Y=1])$
$= \mathbb{P}[X=1] \times \mathbb{P}[Y=1] = p^2$
$\mathbb{P}[S=0] = 1-p^2-q^2 = 2pq$

IX. Loi géométrique.

1°. Quelles valeurs peut prendre N ?

Pile peut apparaître à n'importe quel numéro de tirage et peut éventuellement ne jamais apparaître, ainsi,
$N(\Omega) = \mathbb{N}^* \cup \infty$
$N = 1, 2, 3, ..., +\infty$

2°.
$[X_1 = 0]$=« La personne fait face au 1er lancer. »
$[X_2 = 0]$=« La personne fait face au 2nd lancer. »
$[X_{k-1} = 0]$=« La personne fait face au $(k-1)$ième lancer. »
$[X_k = 1]$=« La personne fait face au $k$ième lancer. »
$[N = k]$=« Le premier pile est apparu au $k$ième lancer. »

$[N=k] = [X_1=0] \cap [X_2=0] \cap ... \cap [X_{k-1}=0] \cap [X_k=1]$

3°. $P[N=k] =$
$\mathbb{P}([X_1=0] \cap [X_2=0] \cap ... \cap [X_{k-1}=0] \cap [X_k=1])$
$= \mathbb{P}[X_1=0] \times \mathbb{P}[X_2=0] \times ... \times \mathbb{P}[X_{k-1}=0] \times \mathbb{P}[X_k=1]$
par indépendance des variables aléatoires X_k
$\mathbb{P}[N=k] = q \times q \times ... \times q \times p = q^{k-1}p$
ceci caractérise la loi géométrique.

4°. $P[N \le n] = \displaystyle\sum_{k=1}^{n} q^{k-1}p = p \times \frac{1-q^n}{1-q} = 1-q^n$
qui est la probabilité que le premier pile apparaisse avant l'instant n. On en déduit alors que la probabilité que pile n'apparaisse jamais est égale à $\mathbb{P}[N=\infty] = \lim\limits_{n\to+\infty} q^n = 0$ à condition que $q \neq 1$

5°. Se reporter au cours pour la démonstration qui utilise les séries entières : $\mathbb{E}[N] = \dfrac{1}{p}$ et $Var(N) = \dfrac{q}{p^2}$

6°. Applications.

6.1. $p = 1-q$ est la proba de transmission avec succès. L'instant X de passage avec succès suit une loi géométrique de paramètre $p = 0,9$ (par indépendance).
$\mathbb{P}[X=4] = q^3p = 0,0009$ et $\mathbb{P}[X<3] = p+pq = 0,99$
Le délai moyen d'attente est la valeur moyenne de X soit $1/p = 1,1$: les messages passent donc en moyenne dès le premier essai. Mais ce système est inefficace car il bloque toute communication dès que deux messages entrent en collision. Pour éviter cela, on utilise une technique présentée dans l'exercice Aloha et Ethernet.

6.2. Soit X le nombre de tirages nécessaires avant de tirer une boule rouge. X suit une loi géométrique de paramètre $p = 0,3$ (proportion de boules rouges). On cherche $\mathbb{P}[X \le 4] = p + pq + .. + pq^3 = 1-(0,7)^4 = 0,75$

X. Loi hypergéométrique.

1°. Le nombre de boules tirées est n, le nombre de boules blanches est Np ; le nombre maximum de boules blanches que l'on peut tirer est donc $M = \inf(n, Np)$
X peut donc prendre n'importe quelle valeur entière entre 0 et M (inclus tous deux).

2°. Les tirages étant simultanés, l'ordre n'importe pas. Soit k le nombre de boules blanches tirées ; on peut les choisir de C_{Np}^k façons ; les $n-k$ boules noires peuvent alors être choisies de C_{Nq}^{n-k} façons et le nombre total de poignées de n boules parmi N est C_N^n. En faisant l'hyprothèse d'équiprobabilité sur toutes les configurations, en déduit :
$P[X=k] = \dfrac{C_{Np}^k C_{Nq}^{n-k}}{C_N^n}$ si $k \in \{0, 1, 2, ..., M\}$

3°. Si $N \to +\infty$, $\mathbb{E}[X] = np$ d'après la formule donnée. et $Var(X) \to npq$ qui sont les caractéristiques d'une loi binomiale (Attention, ceci ne prouve pas que la limite est celle d'une loi binomiale ! !). En fait, on peut considérer que quand N devient très grand, on a peu de chances de tirer deux fois la même personne et les tirages peuvent être considérés comme sans remise.

4°. X suit une loi hypergéométrique $\mathcal{H}(1000; 10; 0,02)$ mais on a tout intérêt à utiliser l'approximation par la loi binomiale. En ce cas, $\mathbb{P}[X=0] \approx 0,98^{10} = 0,81$,
$\mathbb{E}[X] = 10 \times 0,02 = 0,2$ et $Var(X) \approx 0,194$
$\mathbb{P}[X=1] \approx 10 \times 0,02 \times 0,98^9 = 0,17$,
$\mathbb{P}[X=2] \approx 45 \times 0,02^2 \times 0,98^8 = 0,015$ et donc
$\mathbb{P}[X \le 2] \approx 0,999$

XI.

1°. $\mathbb{P}[X \ge k] = \displaystyle\sum_{i=k}^{+\infty} q^{i-1}p = q^{k-1}$ après simplifications.
Ainsi,
$\mathbb{P}[X \ge k] = q_1^{k-1}$ et $\mathbb{P}[Y \ge k] = q_2^{k-1}$

2°. Soit $Z = \inf(X, Y)$; $[Z \ge k] = [X \ge k] \cap [Y \ge k]$
En effet, si $Z \ge k$, alors X et Y le sont aussi puisque Z est le plus petit des deux, et si les deux sont plus grands que k alors le plus petit des deux l'est également !
Puisque l'on a supposé X et Y indépendants, on aura donc $\mathbb{P}[Z \ge k] = \mathbb{P}[X \ge k] \times \mathbb{P}[Y \ge k] = (q_1q_2)^{k-1}$

3°. $[Z \ge k] = [Z=k] \cup [Z \ge k+1]$ et ces deux évènements sont disjoints. Ainsi,
$\mathbb{P}[Z \ge k] = \mathbb{P}[Z=k] + \mathbb{P}[Z \ge k+1]$
$\Rightarrow \mathbb{P}[Z=k] = \mathbb{P}[Z \ge k] - \mathbb{P}[Z \ge k+1]$
$= (q_1q_2)^{k-1}(1-q_1q_2)$
Z suit donc une loi géométrique de paramètre $1-q_1q_2$

XII.

Soit une variable aléatoire X suivant une loi de Poisson de paramètre $\lambda = 3$. Calculer $P([X \le 4]/[X \ge 2])$

$$\mathbb{P}[X \le 4/X \ge 2] = \frac{\mathbb{P}([X \le 4] \cap [X \ge 2])}{\mathbb{P}[X \le 2]} = \frac{0,616}{0,8} \approx 0,77$$

loi binomiale et loi de Poisson

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.

Soit X la variable égale au nombre de boules rouges tirées. Les tirages se font avec remise et X suit donc une loi binomiale $\mathcal{B}(n,p)$ avec $n=7$ et $p=25/40=5/8$; c'est le nombre de succès au cours de n expériences de Bernoulli indépendantes. Ainsi
$\mathbb{P}[X=4]=C_7^4(5/8)^4(3/8)^3 \approx 0,28$

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. Calculer la loi de X puis la proba de l'évènement $A=$« 2 ou 3 ordinateurs sont défectueux. »

X suit une loi binomiale de paramètre $n=5$ et $p=0,9$
$\mathbb{P}[X=3]=C_5^3\times 0,9^3\times 0,1^2 \approx 0,073$
$\mathbb{P}[X=2]=C_5^2\times 0,9^2\times 0,1^3 \approx 0,0081$ donc
$\mathbb{P}(A)=0,0811$

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 ?

Le nombre de cas suit une loi binomiale (car on peut supposer la population assez grande) avec $n=500$ et $p=1/100$. Comme $np=5$, on peut encore approcher cette loi par une loi de Poisson de paramètre $\lambda=5$. Soit X la variable égale au nombre de cas. On cherche :
$\mathbb{P}[X\leq 4]\approx e^{-5}\left(1+5+\frac{5^2}{2!}+..+\frac{5^4}{4!}\right)\approx 0,44$
$\mathbb{P}[X\geq 5]\approx 0,559$

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 ?

Soit X le nombre de pannes annuelles.
$\mathbb{P}[X\geq 1]=1-\mathbb{P}[X=0]=1-e^{-2}\approx 0.865$

5°. X suit une loi binomiale de paramètre $n=7$ et $p=0.9$. Ainsi :
$\mathbb{P}[X\geq 4]=0.023+0.124+0.37+0.48=0.99$

XIV.

1°. $\mathbb{P}[X=k+1]=e^{-\lambda}\frac{\lambda^{k+1}}{(k+1)!}$ et $\mathbb{P}[X=k]=e^{-\lambda}\frac{\lambda^k}{k!}$
donc $\frac{\mathbb{P}[X=k+1]}{\mathbb{P}[X=k]}=\frac{\lambda}{k+1}$
On passe donc d'une proba à l'autre en multipliant par $\lambda/(k+1)$

2°. $\mathbb{P}[X=0]=e^{-\lambda}$; une fois calculée cette probabilité, il suffit d'appliquer la formule de récurrence ci-dessus.

3°. Notons $u_k=\mathbb{P}[X=k]$. On cherche le sens de variation de la suite de terme général u_k.
$u_{k+1}-u_k=e^{-\lambda}\frac{\lambda^k}{k!}(\frac{\lambda}{k+1}-1)$ ainsi, la suite sera décroissante dès que $\lambda<k+1$ ie. $k>\lambda-1$

4°. $\mathbb{P}[X\leq 1]=0,95$
$\mathbb{P}[X\leq 1]=\mathbb{P}[X=0]+\mathbb{P}[X=1]=e^{-\lambda}(\lambda+1)=0,95$
$\Rightarrow \ln(e^{-\lambda})+\ln(\lambda+1)=\ln 0,95$
$\Rightarrow \lambda$ est solution de l'équation
$\ln(1+x)-x=\ln 0,95\simeq -0,05$

Soit $\phi(x)=\ln(1+x)-x$. On voit facilement que cette fonction est décroissante sur $\mathbb{R}^+$, continue donc bijective et que l'équation $\phi(x)=\ln 0,95$ a donc une solution unique. Un calcul approché de $\phi(0,3)$ et $\phi(0,4)$ donne $\lambda\approx 0,4$

XV.

1°. X suit une loi binomiale de paramètres $n=300$ et $p=0,01$. Ainsi, $\mathbb{E}[X]=npq\approx 3$ et $\sigma=\sqrt{npq}\approx 1,72$

2°. $\lambda=np$ dans la loi des évènements rares, ie $\lambda=0,01n$. Si
$A=[X\geq 3]\Rightarrow \mathbb{P}(A)=1-\mathbb{P}[X\leq 2]=1-e^{-\lambda}(1+\lambda+\frac{\lambda^2}{2})$

Soit $f(x)=1-e^{-x}(1+x+\frac{x^2}{2})$

f est continue et dérivable sur $\mathbb{R}^+$, $f'(x)=-\frac{x^2}{2}e^{-x}$ et l'on en déduit donc que f est décroissante.
$f(6,1)\approx 0,06, f(6,2)\approx 0,05$ et $f(6,3)\approx 0,049$
$\mathbb{P}(A)\approx 0,95 \iff 1-f(\lambda)=0,95 \iff f(\lambda)=0,05$
$\iff \lambda\approx 6,2 \iff 0,01n=6,2 \iff n\geq 621$

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 ?
Par indépendance des erreurs, il s'agit d'une loi binomiale de paramètres n et p

2°. Déterminer la probabilité qu'un fichier ne contienne aucune erreur.
$\mathbb{P}[X=0]=q^{10000}\approx 4.5\times 10^{-5}$

3°. Déterminer le nombre moyen d'erreurs par fichier.
$\mathbb{E}[X]=np=10$

4°. On améliore la qualité de la transmission afin que $p=10^{-5}$. Par quelle loi peut-on approcher X ? Pourquoi ? Calculer alors $P[X\leq 1]$
$np=1\Rightarrow$ on peut approcher la loi de X par une loi de Poisson de paramètre $\lambda=1$ (en utilisant la loi des évènements rares). On a alors $\mathbb{P}[X\leq 1]=2e^{-1}$

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 $k=223$. A 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 et que les bits sont transmis de façon indépendante. 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 ces paramètres ? Quel est le nombre moyen d'erreurs par image transmise ?

X suit clairement une loi binomiale de paramètres $n = 2.10^7$ et $p = 5.10^{-2}$ (nombre de réalisations de n succès indépendants avec proba de succès égale à p). On a donc $\mathbb{E}[X] = np = 10^6$, ce qui est beaucoup pour une transmission...

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 $\mathbb{P}[X \leq 20]$

$p = 10^{-6} \Rightarrow \mathbb{E}[X] = 20$; comme $np \geq 5$, on ne peut pas approcher par une loi de Poisson.

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 $\mathbb{P}[X \leq 2]$

$p = 10^{-6} \Rightarrow \mathbb{E}[X] = 2 \leq 5$, on peut donc approcher par une loi de Poisson de paramètre $\lambda = 2$. On a alors $\mathbb{P}[X \geq 1] = 1 - \mathbb{P}[X = 0] = 1 - e^{-2}$ et $\mathbb{P}[X \leq 2] = 5/e^2$

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

$\mathbb{P}[X = 0] = 3/4$, $\mathbb{P}[X = 1] = 1/4$, $\mathbb{P}[Y = 0] = 1/2$, $\mathbb{P}[X = 1] = 1/2$

2°. Ces variables sont-elles indépendantes ?

$\mathbb{P}[X = 1, Y = 1] = 0 \neq \mathbb{P}[X = 1] \times \mathbb{P}[Y = 1]$
X et Y ne sont donc pas indépendantes.

3°. Calculer $cov(X, Y)$

$\mathbb{E}[XY] = 0 \Rightarrow \ Cov(X, Y) = -1/8$

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°. $\mathbb{P}[X = 1] = 1/3$, $\mathbb{P}[X = -1] = 2/3$, $\mathbb{P}[Y = 0] = 2/3$, $\mathbb{P}[Y = 1] = 1/3$, $\mathbb{E}[X] = -1/3$, $\mathbb{E}[Y] = 1/3$, $\mathbb{E}[X^2] = 1$, $\mathbb{E}[Y^2] = 1/3$, $\Rightarrow Var(X) = 8/9$ et $\Rightarrow Var(Y) = 2/9$
$Cov(X, Y) = \mathbb{E}[XY] - \mathbb{E}[X]\mathbb{E}[Y] = -2/9 \neq 0$
X et Y ne sont donc pas indépendantes.

2°. On pose $S = X + Y$. Déterminer la loi de S, son espérance et sa variance.

$\mathbb{P}[S = 1] = 1/3$, $\mathbb{P}[S = -1] = 1/3$, $\mathbb{P}[S = 0] = 1/3$, $\mathbb{P}[S = 2] = 0$, $\mathbb{E}[S] = 0$ et $Var(S) = 2/3$

XX.

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

$U \backslash V$	-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°.2°.

$\mathbb{P}[U = 2] = \mathbb{P}[U = -2] = 1/4$, $\mathbb{P}[U = 0] = 1/2$ et idem pour V. $\mathbb{E}[U] = \mathbb{E}[V] = 0$, $\mathbb{E}[U^2] = \mathbb{E}[V^2] = 2$ et donc $Var(U) = Var(V) = 2$

$\mathbb{E}[UV] = 0 \Rightarrow Cov(U, V) = 0$ mais U et V ne sont pas indépendants car
$\mathbb{P}[U = 2]\mathbb{P}[V = 2] = 1/16 \neq \mathbb{P}[U = 2; V = 2]$

3°. Construisons X et Y à partir d'un tableau :

U	V	X	Y	$\mathbb{P}$
-2	0	-1	-1	1/4
2	0	1	1	1/4
0	-2	-1	1	1/4
0	2	1	-1	1/4

D'après le tableau, $\mathbb{P}[X = 1] = \mathbb{P}[X = -1] = 1/2$,
$\mathbb{P}[Y = 1] = \mathbb{P}[Y = -1] = 1/2$
avec $X = (U + V)/2$ et $Y = (U - V)/2$
$\mathbb{E}[XY] = 1/4 + 1/4 - 1/4 - 1/4 = 0 \Rightarrow Cov(X, Y) = 0$
X et Y sont indépendants car
$\mathbb{P}[X = \epsilon; Y = \epsilon'] = \mathbb{P}[X = \epsilon]\mathbb{P}[X = \epsilon']$, $\forall \epsilon, \epsilon' = -1, 1$

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.

$\mathbb{P}[X=-1]=1/3$, $\mathbb{P}[X=1]=2/3$,
$\mathbb{P}[Y=-1]=\mathbb{P}[Y=1]=\mathbb{P}[Y=0]=1/3$
$\mathbb{E}[X]=1/3$, $\mathbb{E}[X^2]=1$, $\mathbb{E}[Y]=0$, $\mathbb{E}[Y^2]=2/3$,
$Var(X)=8/9$, $Var(Y)=2/3$

2°. Déterminer la covariance $Cov(X,Y)$. X et Y sont-elles indépendantes ?

$\mathbb{E}[XY]=1/6+0-1/6-1/6+0+1/6=0$
$\Rightarrow Cov(X,Y)=\mathbb{E}[X]\mathbb{E}[Y]=0$
$\mathbb{P}[Y=0]\mathbb{P}[X=-1]=1/3\times 1/3=1/9\neq 0$
$\mathbb{P}[Y=0;X=-1]=0$ donc X et Y ne sont pas indépendants.

3°. On pose $U=XY$ et $V=Y/X$.
Déterminer la loi du couple (U,V). U et V sont-elles indépendantes ?

X	Y	U	V
-1	-1	1	1
-1	0	0	0
-1	1	-1	-1
1	-1	-1	-1
1	0	0	0
1	1	1	1

$U\backslash V$	-1	0	1	TOT
-1	1/3	0	0	1/3
0	0	1/3	0	1/3
1	0	0	1/3	1/3
TOT	1/3	1/3	1/3	1

U et V ne sont pas indépendantes.

Applications.

XXII. Protocole Aloha et Ethernet.

1°. Chaque message apparaît selon une loi de Bernoulli de paramètre α. Le nombre de messages à t suit donc une loi binomiale $\mathcal{B}(N,\alpha)$. Puisque les messages arrivant sont immédiatement transmis, il y a collision si $X>1$ avec probabilité :

$\mathbb{P}[X>1]=1-\mathbb{P}[X=0]-\mathbb{P}[X=1]$
$=1-(1-\alpha)^N-N\alpha(1-\alpha)^{N-1}$

2°. L'instant où le message essaie à nouveau de passer une première fois suit une loi géométrique (sinon, à chaque instant $t+1$, $t+2$, ... la proba reste la même : α). On a :

$p(t+1)=p$, $p(t+2)=qp$, ...
Pour une variable géométrique, l'espérance est $1/p$

3°. Un nouveau message est émis avec succès s'il arrive **seul** et si les messages en attente n'émettent pas. La probabilité cherchée est donc :

$(N-k)\alpha(1-\alpha)^{N-k-1}(1-p)^k$
En effet, $(1-p)^k$ est la proba qu'aucun message en attente n'émette. $(1-\alpha)^{n-k-1}$ indique qu'il reste $N-k$ stations dont $N-k-1$ ne doivent pas émettre.

Le plus simple est de regarder lorsqu'il n'y a pas de collision. Notons $\bar{C}$ cet évènement. Il peut se décrire par "aucun message en attente n'émet et aucun nouveau n'arrive" OU "aucun message en attente n'émet et un seul arrive" OU "un seul des messages en attente n'émet et aucun nouveau n'arrive". Ainsi,

$\mathbb{P}(\bar{C})=(1-p)^k(1-\alpha)^{N-k}+...$
$...+(1-p)^k(N-k)\alpha(1-\alpha)^{N-k-1}+...$
$...+kp(1-p)^{k-1}(1-\alpha)^{N-k}$

4°. On cherche $\mathbb{P}[X>1]=1-e^{-\lambda}-\lambda e^{-\lambda}=0,0046$
En ce cas, il y a collision puisque tout message arrivant est immédiatement transmis.
$[N(t+1)=k-1/N(t)=k]$ se décrit par "un message a été transmis avec succès"
On notera $\alpha_k=\mathbb{P}[X=k]=e^{-\lambda}\lambda^k/k!$ la probabilité que k messages arrivent à un instant donné.
$\mathbb{P}[N(t+1)=k-1/N(t)=k]=$
$\alpha_0 kp(1-p)^{k-1}+\alpha_1(1-p)^k$
de même, $\mathbb{P}[N(t+1)=k/N(t)=k]=$
$(1-p)^k\alpha_1+\alpha_0(1-kp(1-p)^{k-1})$
et cet évènement se lit "le nombre de messages ne change pas si aucun n'émet et un message arrive ou si aucun message n'arrive et qu'aucun message en attente n'est transmis avec succès".
$\mathbb{P}[N(t+1)=k+1/N(t)=k]=\alpha_i$

5°. On re-émet à $t+2$ avec proba $p/2$ si l'on a déjà tenté en $t+1$ et avec proba p sinon.

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

$$I(A_k)=-\log_2\mathbb{P}(A_k)=-\log_2\mathbb{P}[X=x_k]$$

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 :

$$H(X)=\mathbb{E}[I(X)]=-\sum_{k=1}^n p_k\log_2 p_k=-\sum_{k=1}^n\mathbb{P}[X=x_k]\log_2\mathbb{P}[X=x_k]$$

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

1°. Montrer que $H(X)\geq 0$. A quelle condition a-t-on $H(X)=0$?

$H(X)=-\sum_k p_k\log_2 p_k$; $0\leq p_k\leq 1\Rightarrow\log_2 p_k\leq 0$
$\Rightarrow-\log_2 p_k\geq 0\Rightarrow H(X)\geq 0$
$H(X)=0\iff p_k\log_2 p_k=0\ \forall k,\iff p_k=0$ ou $1\ \forall k$
$p_k=1\iff X=$cte.

2°. Montrer que si X suit une loi uniforme sur $\{x_1,...,x_n\}$, alors $H(X)=\log_2 n$

$H(X)=-\sum_k\frac{1}{n}\log_2\frac{1}{n}=n\times\frac{1}{n}\log_2 n$

3°. Soient $(p_1,...,p_n)$ et $(q_1,...,q_n)$ deux mesures de probabilités. Montrer que $\sum_{k=1}^n p_k\log_2\frac{q_k}{p_k}\leq 0$

$$\sum_{i=1}^n p_i\log_2\frac{q_i}{p_i}=\frac{1}{\ln 2}\sum p_i\ln(q_i/p_i)$$
$$\leq\frac{1}{\ln 2}\sum p_i(q_i/p_i-1)=\frac{1}{\ln 2}(\sum q_i-\sum p_i)=0$$

4°. En déduire que $H(X) \leq \log_2 n \ \forall X$. En déduire la valeur maximale de l'entropie. Interprétation ?

Si $q_i = 1/n$, on a $\sum p_i \log_2(1/(np_i)) \leq 0$
$\Rightarrow -\sum p_i \log_2 n - \sum p_i \log_2 p_i \leq 0$
$-\log_2, +H(X) \leq 0 \Rightarrow H(X) \leq \log_2 n$
$H(X)$ est max si X suit une loi uniforme. En ce cas, l'incertitude sur X est la plus grande possible.

5°. Si X suit une loi de Bernouilli, calculer $H(X)$. Etudier la fonction $p \to H(X)$.
$H(X) = -p \log_2 p - q \log_2 q$ donc
$H(p) = -p \log_2 p - (1-p) \log_2 (1-p)$
Pour l'étude de cette fonction, on se reportera au TD de première année sur les fonctions (H est la fonction d'entropie binaire).

6°. Calculer $H(X)$ lorsque X suit une loi de Poisson de paramètre λ.

$$H(X) = \sum_{i=0}^{+\infty} \left[e^{-\lambda_i} \frac{\lambda^i}{i!} (\lambda_i + \log(i!) - i \log \lambda) \right]$$

XXIV. Entropie et information mutuelle.

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 dont 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.

XXVI. Fonction génératrice.

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

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

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 $]-1,1[$.

G est la somme d'une série entière de rayon de convergence supérieur ou égal à 1. D'où la réponse.

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.

• Si X suit une loi de Bernoulli $b(p)$: $G(x) = px + q$
• Si X suit une loi binomiale $\mathcal{B}(n,p)$:

$$G(x) = \sum_{k=0}^{n} C_n^k (px)^k q^{n-k} = (px+q)^n$$

• Si X suit une loi géométrique $\mathcal{G}(p)$:

$$G(x) = \sum_{k \geq 1} pq^{k-1} x^k = \frac{px}{1-qx}$$

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

$$\mathbb{E}[X] = \sum_{k \geq 1} k\mathbb{P}[X=k] \text{ et } G'(x) = \sum_{k=1} k\mathbb{P}[X=k] x^{k-1}$$

Il suffit de prendre la limite lorsque x tend vers 1 par valeur inférieure.

4°. La fonction génératrice d'une loi de Poisson de paramètre λ est $G(x) = e^{-\lambda} \sum\limits_{k \geq 1} \frac{\lambda^k}{k!} x^k = e^{\lambda(x-1)}$
En admettant la propriété, la fonction génératrice de la somme $X + Y$ est donc $e^{(\lambda+\mu)(x-1)}$ qui est la fonction génératrice d'une loi de Poisson de paramètre $\lambda + \mu$.

De la même façon, la fonction génératrice de la somme de deux lois binomiales est $G(x) = (px+q)^{n+m}$ qui est la fonction d'une loi binomiale de paramètres $(n+m, p)$

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 $\frac{\lambda}{n} \sim p$ 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.

A vous de le faire !

XXVII. Fonction caractéristique.

Cet exercice est très voisin du précédent et les techniques de démonstrations sont les mêmes.

XXVIII.

Cet exercice est assez difficile. Vous pouvez le laisser de côté dans un premier temps.

1°. Déterminer la loi du couple (X, Y)

Par indépendance de X et Y, $\mathbb{P}[X = k, Y = n] = p^2 q^{n+k}$

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

$[Z \leq k] = [X \leq k] \cap [Y \leq k] \Rightarrow \mathbb{P}[Z \leq k] = \mathbb{P}[X \leq k]^2$par indépendance et identique distribution des deux variables aléatoires.

$\mathbb{P}[Z \leq k] = \left(p \sum_{i=0}^{k} q^i\right)^2 = (1 - q^{k+1})^2$ pour $k \geq 1$

On en déduit :

$\mathbb{P}[Z = k] = \mathbb{P}[Z \leq k] - \mathbb{P}[Z < k] = (1 - q^{k+1})^2 - (1 - q^k)^2$
$= pq^k(2 - (1+q)q^k)$

$\mathbb{E}[Z] = 2p \sum_{k \geq 0} kq^k - p(1+q) \sum_{k \geq 0} kq^{2k}$

Soit $\alpha = \sum kq^k$ et $\beta = \sum kq^{2k}$

En utilisant les techniques classiques des séries entières, on a $\alpha = \dfrac{q}{p^2}$ et $\beta = \left(\dfrac{q}{p}\right)^2 \dfrac{1}{(1+q)^2}$ Ainsi,

$\mathbb{E}[Z] = \alpha - \beta = \dfrac{q}{p}\dfrac{q+2}{q+1}$

On laisse tomber le calcul de $Var(Z)$

XXIX.

1°. $\mathbb{P}[X = k, Y = i] = q^k p^i$, $k, i \geq 1$

$\mathbb{P}[Y = i] = \sum_{k \geq 1} \mathbb{P}[X = k, Y = i] = p^i \sum_k q^k = qp^{i-1}$ et de la même façon, $\mathbb{P}[X = k] = q^{k-1}p$

X et Y suivent des lois géométriques de paramètres respectifs p et q. Pour tout k et tout i,
$\mathbb{P}[X = k, Y = i] = \mathbb{P}[X = k]\mathbb{P}[Y = i]$ et X, Y sont donc indépendants.

On en déduit alors $cov(X, Y) = 0$

2°. $S(x) = \dfrac{1}{1-x} = \sum_{k \geq 0} x^k$ avec $R = 1$

3°. $S'(x) = \dfrac{1}{(1-x)^2}$, $S''(x) = \dfrac{2}{(1-x)^3}$, ...,

$S^{(k)}(x) = \dfrac{k!}{(1-x)^{k+1}} = \sum_{n \geq k} n...(n-k+1)x^{n-k}$

$\phi(x) = \dfrac{1}{(1-x)^{k+1}} = \dfrac{1}{k!} S^{(k)}(x) = \sum_{n \geq k} C_n^k x^{n-k}$

avec un rayon de convergence égal à 1

4°. $\mathbb{P}[X = k, Y = n] = C_n^k p^{k+1} q^{2n-k}$, $0 \leq k \leq n$

$\mathbb{P}[X = k] = \sum_{n \geq k} C_n^k p^{k+1} q^{2n-k} = p^{k+1} q^k \phi(q^2)$

$= p^{k+1} q^k \dfrac{1}{(1-q^2)^{k+1}} = \alpha(1-\alpha)^k$ avc $\alpha = \frac{1}{1+q} \Rightarrow X$ suit une loi géométrique de paramètre α

$\mathbb{P}[Y = n] = \sum_{k=0}^{n} C_n^k p^{k+1} q^{2n-k} = pq^n$

$\Rightarrow Y$ suit une loi géométrique de paramètre p

$\mathbb{P}[X = k / Y = n] = \dfrac{\mathbb{P}[X = k, Y = n]}{\mathbb{P}[Y = n]} = C_n^k p^k q^{n-k}$

Qui est une loi binomiale de paramètres n et p

5°. NON car $\mathbb{P}[X = k / Y = n] \neq \mathbb{P}[X = k]$

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 $\sum_{n=k}^{+\infty} \dfrac{x^{n-k}}{(n-k)!}$

$e^x = \sum_{n \geq 0} \dfrac{x^n}{n!}$ avec $R = \infty$ et $\sum_{n \geq k} \dfrac{x^{n-k}}{(n-k)!} = e^x$

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

$\mathbb{P}[X = n] = \sum_{k=0}^{n} \mathbb{P}[X = n, Y = k] = e^{-\lambda} \dfrac{\lambda^n}{n!} \sum_{k=0}^{n} C_n^k p^k q^{n-k} = e^{-\lambda} \dfrac{\lambda^n}{n!}$

X suit une loi de Poisson de paramètre λ

$\mathbb{P}[Y = k] = \sum_{n \geq k} e^{-\lambda} \dfrac{\lambda^n}{n!} C_n^k p^k q^{n-k} = e^{-\lambda} \dfrac{(p\lambda)^k}{k!} \sum_{n \geq k} \dfrac{(\lambda q)^{n-k}}{(n-k)!}$

$= \dfrac{1}{k!} e^{-\lambda} (p\lambda)^k e^{\lambda q} = e^{-p\lambda} \dfrac{(p\lambda)^k}{k!}$

Y suit donc une loi de Poisson de paramètre λp

3°. $\mathbb{P}[Y = k / X = n] = \dfrac{\mathbb{P}[Y = k, X = n]}{\mathbb{P}[X = n]} = C_n^k p^k q^{n-k}$

4°. Les appels arrivent au standard suivant une loi de Poisson. Connaissant le nombre d'appels au standard, ceux-ci sont redirigés ou bien vers A ou vers ailleurs.
$A / X = n$ est donc binomiale.

XXXI.

1°. Soit (X, Y) un couple de variables aléatoires dont a 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 $[S = n] = \bigcup_{k=1}^{n-1} ([X = k] \cap [Y = n - k])$ et en déduire que $\mathbb{P}[S = n] = (n-1)p^2 q^{n-2}$ 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 nième tirage et 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 $g(s) = \sum_{n=1}^{+\infty} \mathbb{P}[T_1 = n] s^n$ 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 $h(s) = \sum_{k=1}^{+\infty} \mathbb{P}[T_2 = k] s^k$ de T_2
4°. Calculer $\mathbb{E}[T_2]$

5°. 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 $\mathbb{E}[U]$?