Probabilités(2).Variables aléatoires discrètes

Un peu d'histoire : le postulat des parallèles. « Par un point on peut mener une et une seule parallèle à une droite donnée ». Tel est le cinquième postulat servant de fondement à la géométrie dans les éléments d'Euclide. Pendant 2300 ans les mathématiciens se sont posés la question suivante : ce postulat est il une conséquence des 4 autres (auquel cas on peut le supprimer, sans rien changer à la géométrie) ou bien en est il indépendant (auquel cas il existe des géométries dans lequel il n'est pas vrai, des géométries non euclidiennes). C'est Bolyai et Lobatchevski qui répondront au 19eme siècle en créant une géométrie (appelée hyperbolique) ne vérifiant pas ce postulat. Puis Riemann en construira une autre (la géométrie Riemannienne ou sphérique). On sait aujourd'hui que ces 3 « modèles »recouvrent toutes les géométries possibles. Dans ces géométries, certaines propriétés usuelles ne sont plus vérifiées, par exemple la somme des angles d'un triangle ne vaut plus 180 degrés, mais la plupart restent vraies (par exemple : par deux points passent une et une seule droite). L'existence de géométries disctinctes a des conséquences fondamentales en physique : laquelle de ces géométries s'applique à notre univers ? Pour la physique Newtonnienne c'est la géométrie d'Euclide. Pour la physique relativiste tout dépend de la courbure de l'univers, c'est à dire de sa masse (aujourd'hui inconnue)

Généralités sur les variables aléatoires et l'espérance.

1. Limite inférieure et supérieure d'une suite de variables aléatoires réelles.
   Soit (X_n) une suite de variables aléatoires réelles discrètes définies sur un espace probabilisé $(\Omega, (A), P)$ et minorées par une variable aléatoire Y.
   On pose pour tout n, $Y_n = \inf\{X_k, k \geq n\}$ puis $\underline{X} = \sup\{Y_k, k \geq 0\}$.
   Montrer que $\underline{X}$ est une variable aléatoire.

2. Limite de variables aléatoires :
   Soit (X_n) une suite de variables aléatoires réelles définies sur un espace probabilisé $(\Omega, (A), P)$. On suppose que la suite (X_n) converge simplement vers une application X à valeurs discrètes. Montrer que X est une variable aléatoire.

3. Temps d'arrêt
   Soit $(X_n)_n$ une suite de variables aléatoires discrètes, mutuellement indépendantes à valeurs dans $\mathbb{N}$. Si A est un sous ensemble de $\mathbb{N}$, on note $T = \min\{n,\ X_n \in A\}$ (le minimum étant infini si l'ensemble est vide).
   (a) Démontrer que T est une variable aléatoire.
   (b) Démontrer que X_T est une variable aléatoire.
   (c) On suppose que les X_n ont toutes la même loi. Déterminer la loi de X_T. Comment s'interprète ce résultat ?

4. Fonction de répartition d'une variable aléatoire discrète réelle.

   Soit X une variable aléatoire discrète à valeurs réelles. On pose
   $$F_X(x) = P(X \leq x)$$
   F_X s'appelle la fonction de répartition de X.

   (a) Montrer que F_X est une fonction croissante donner ses limites en $\pm\infty$.
   (b) Montrer que F_X est continue à droite en tout point.
   (c) Etablir que $P(X = x) = F_X(x) - \lim\limits_{y \to x^-} F_X(y)$. Quels sont les points de discontinuité de F_X ?
   (d) En déduire que F_X caractérise la loi de X (ie : deux variables aléatoires ont même loi si et seulement si elles ont même fonction de répartition.
   (e) Un exemple (difficile) : soit (a_n) une numérotation des rationnels. On pose $H(x) = \sum\limits_{a_n \leq x} \dfrac{1}{2^n}$. Exhiber une variable aléatoire dont H est la fonction de répartition et trouver ses points de continuité.

Généralités sur l'espérance

5. Un exercice théorique :
   Soient X et Y deux variables aléatoires réelles discrètes indépendantes.
   On suppose que $X+Y$ a une espérance. Montrer que X aussi.

   indication : écrire la définition et choisir y_0 tel que $p(Y=y_0)$ soit non nul.
   Le résultat est il toujours vrai si les variables ne sont pas indépendantes ?

6. Une caractérisation des variables aléatoires ayant une espérance.
   cet exercice utilise le lemme de Borel Cantelli
   Soit (X_n) une suite de variables aléatoires à valeurs réelles iid.
   On note $A = [\{n, |X_n| > n\}\text{est fini}]$
   Démontrer que X_1 possède une espérance si et seulement si $P(A) = 1$
   Indication : On commencera par le cas des variables aléatoires à valeurs entières et on utilisera la formule calculant l'espérance dans ce cas.

7. Une utilisation géométrique de l'espérance.

   Soient $v_1, \ldots, v_n$ n vecteurs de norme (euclidienne) égale à 1.
   Soit $p_1, \ldots, p_n$ des éléments de $[0,1]$ et $v = \sum p_i v_i$.
   Enfin pour $I \subset \{1, \ldots, n\}$ on note $w_I = \sum_I v_i$.

   On se propose de montrer que v est à distance au plus $\dfrac{\sqrt{n}}{2}$ de l'un des w_I.

   (a) On considèrera des variables indépendantes X_i suivant la loi $\mathcal{B}(p_i)$. Calculer l'espérance de la variable aléatoire
   $$||\sum_1^n (X_i - p_i)v_i||^2$$

   (b) Conclure.

8. Inégalité de Jensen.

   Soit X une variable aléatoire réelle discrète et f une fonction convexe.
   On rappelle que si $p_1, \ldots, p_n$ sont des réels positifs de somme 1 alors pour tout $(x_1, \ldots, x_n)$ on a
   $$f(\sum_1^n x_i p_i) \leq \sum_1^n p_i f(x_i)$$

   (a) Démontrer que si les $(p_n)_{n\in\mathbb{N}}$ sont des réels positifs de somme 1, on a pour toute famille (x_n) sous réserve de sommabilité l'inégalité
   $$f(\sum_{i\in\mathbb{N}} x_i p_i) \leq \sum_{i\in\mathbb{N}} p_i f(x_i)$$

   (b) On suppose que X et $f(X)$ ont une espérance.
   Démontrer que
   $$f(\mathbb{E}(X)) \leq \mathbb{E}(f(X))$$

9. Définition de l'entropie.
   cet exercice utilise l'inégalité de Jensen démontrée à l'exercice précédent. '

   (a) Soit $(x_1, \ldots, x_n) \in (\mathbb{R}^{+*})^n$ tel que $\sum_{i=1}^n x_i = 1$. Montrer que $-\sum_{i=1}^n x_i \ln x_i \leq \ln n$, et caractériser l'égalité.

   (b) Soit $(x_n)_n \in\,]0,1[^{\mathbb{N}}$ une suite telle que la série de terme général nx_n converge. Montrer que la série de terme général $x_n \ln(x_n)$ converge.

(c) Soit X une variable aléatoire entière possédant une espérance. On pose $p_n = P(X = n)$. Montrer que la somme $S(X) = \sum -p_n \ln(p_n)$ est bien définie (Par convention, on pose $x \ln x = 0$ lorsque $x = 0$). $S(X)$ s'appelle l'entropie de la variable X.

(d) Démontrer que $S(X)$ est positive. Caractériser les cas où elle est nulle.

Limites d'espérances : les 3 exercices qui suivent sont assez difficiles : ce sont des adaptations au cas des variables aléatoires des théorème d'interversion

10. Le théorème de convergence dominée pour les variables aléatoires entières.

Soit (X_n) une suite de variables aléatoires à valeurs entières, et X une variable aléatoire. On suppose que pour tout a, $\mathbb{P}(X_n = a)$ tend vers $\mathbb{P}(X = a)$. On suppose en outre la suite $(|X_n|)$ majorée par une variable aléatoire Y ayant une espérance. On se propose de montrer que X a une espérance et que $E(X_n)$ tend vers $E(X)$

(a) Montrer que pour tout entier a, $\mathbb{P}(X_n > a)$ tend vers $\mathbb{P}(X > a)$ (regarder l'événement complémentaire).

(b) Etablir pour tout N l'inégalité

$$|\mathbb{E}(X_n) - \mathbb{E}(X)| \leq \sum_{k=0}^{N} |\mathbb{P}(X_n > k) - \mathbb{P}(X > k)| + 2 \sum_{k=N}^{\infty} \mathbb{P}(Y > k)$$

(c) Conclure.

11. Le théorème de Beppo Levi pour les variables aléatoire à valeurs entières.

On démontre dans cet exercice le théorème suivant :

Soit $(Y_n)_n$ une suite de variables aléatoires à valeurs <u>entières</u>. On pose $S_n = \sum_{k=0}^{n} Y_k$ et $S = \sum_{n=0}^{\infty} Y_n$ avec la convention que $S = +\infty$ si la série diverge.

Alors on a équivalence entre :

(i) *S est d'espérance finie*

(ii) *La série $\sum \mathbb{E}(Y_n)$ converge*

de plus en cas de convergence on a :

$$\mathbb{E}(S) = \sum_{0}^{\infty} \mathbb{E}(Y_n)$$

(a) Démontrer que (i) implique (ii)

Dans les deux questions suivantes, on suppose que (ii) est réalisée.

(b) Justifier que pour tout entier k on a $\mathbb{P}(S > k) = \lim_{n \to \infty} \mathbb{P}(S_n > k)$ et que pour tout entier K et tout entier n on a

$$\sum_{k=0}^{K} \mathbb{P}(S_n > k) \leq \sum_{p=0}^{\infty} \mathbb{E}(Y_p)$$

(c) Conclure.

12. La variable aléatoire de comptage :

cet exercice utilise l'exercice précédent

Soit (A_n) une famille d'événements. On note S la variable aléatoire égale au nombre d'événements parmi les A_n qui sont réalisés.

(a) Vérifier que S est bien une variable aléatoire.

(b) Montrer que $\mathbb{E}(S) = \sum_{n} \mathbb{P}(A_i)$

13. Un problème de prise de sang.

On considère un groupe de N personnes souffrant ayant la probabilité p (petite) de souffrir d'une maladie. On dispose d'un test fiable par prise de sang pour déceler cette maladie. Mais le test est couteux. On souhaite trouver une stratégie pour minimiser le coût total. On note X la variable aléatoire égale au nombre de tests effectués.

La première stratégie consiste à tester tout le monde. La variable X est constante égale à N.

La seconde stratégie est la suivante : on fixe un entier k divisant N . On groupe les N personnes en paquets de k personnes dont on mélange le sang. On teste chacun des paquets, puis, le cas échéant, on teste chaque personne dans les paquets malsains.

a) Montrer que avec la seconde stratégie $E(X) = \dfrac{N}{k}((1-p)^k + (k+1)(1-(1-p)^k))$

b) Montrer que pour p très petit, cette espérance est minimale pour k proche de $\sqrt{\dfrac{1}{p}}$ et qu'alors elle vaut environ $2\sqrt{p}N$

c) Comparer avec la première stratégie.

Lois classiques.

14. (mines) Exemple de Loi d'un temps d'arrêt.

Une personne effectue une série de sauts , la probabilité de réussir le $i$ème saut étant $\dfrac{1}{i}$. On s'arrête au premier échec. On note X le nombre de saut effectués

Montrer que X est presque surement finie, déterminer sa loi, son espérance et sa variance.

15. Loi du premier retour d'un résultat.

On considère une famille de $n+1$ variables aléatoires à valeurs dans un ensemble fini de cardinal n et suivant la loi uniforme. On note $T = \min\{j \geq 2, X_j \in \{X_1, \ldots, X_{j-1}\}\}$

(a) Calculer $P(T > j)$ puis déterminer la loi de T.

(b) Montrer que $E(T) = \dfrac{n!}{n^n} \displaystyle\sum_{j=0}^{n} \dfrac{n^j}{j!}$

(c) En déduire

$$E(T) = \sqrt{n} \int_0^{\infty} (1 + \frac{t}{\sqrt{n}})^n e^{-\sqrt{n}t} dt$$

(d) *difficile* Donner un équivalent de $E(T)$ quand$n$ tend vers l'infini.On pourra noter que la limite de l'intégrande est égale à $e^{\frac{-t^2}{2}}$ la domination est difficile)

16. Loi Dzéta.

a)Existe t'il une variable aléatoire Y à valeurs dans $\mathbb{N}^*$ telle que $P(Y = k)$ soit égale à $\dfrac{a}{k}$, avec a constante ?

b) Soit maintenant $\alpha > 1$. Montrer qu'il existe une variable aléatoire Y telle que $P(Y = k) = \dfrac{a}{k^\alpha}$. Combien vaut la constante a ?.

A quelle condition Y possède t-elle une espérance ?

c) On suppose que la variable Y suit la loi précédente. Montrer que les événements "p divise Y" où p décrit l'ensemble des nombres premiers, sont indépendants.

d) En déduire que l'on a $\zeta(\alpha) = \displaystyle\prod_{p \in \mathbb{P}} (1 - \frac{1}{p^\alpha})^{-1}$ (développement Eulérien de la fonctions zéta)

e) Démontrer que la probabilité que Y ne soit pas divisible par un carré est égale à $\dfrac{1}{\zeta(2\alpha)}$

17. Loi de poisson versus loi géométrique.

    Dans une salle d'attente de médecin, on suppose que les personnes qui arrivent sont aléatoirement des femmes et des hommes. On suppose aussi que le nombre de personne qui arrive durant une période d'une heure suit une loi de Poisson.

    a) Je parie que le nombre de personnes qui arrivent avant la première femme est pair. Prenez vous le pari ?
    b) Je parie que le nombre de personnes arrivant pendant la première heure est impair. Prenez vous le pari ?

18. Approximation d'une loi binomiale par une loi de Poisson(1).

    On choisit 500 personnes au hasard.

    a) Quelle est la probabilité que 3 exactement soient nées le premier mars ?
    b) Donner une valeur approchée à l'aide de l'approximation de la loi binomiale par la loi de Poisson.

19. Approximation par une loi de Poisson (2)

    L'ensemble des applications de $\{1, \ldots, m\}$ dans $\{1, \ldots, n\}$ est muni de la loi uniforme.

    (a) Déterminer la loi du nombre X d'antécédents de 1
    (b) Déterminer l'espérance de X
    (c) On fixe λ , on prend $m = \lfloor n\lambda \rfloor$. Quelle est la limite quand $n$tend vers l'infini de $P(X = k)$?

20. Minimum de lois indépendantes : cas géométriques.

    Soit $(Z_n)_n$ une famille de variables aléatoires de même loi à valeurs entières mutuellement indépendantes.

    On pose $X_n = \min(Z_1, \ldots, Z_n)$.

    (a) Vérifier que X_n est une variable aléatoire.
    (b) Déterminer $P(X_n \geq k)$ puis la loi de X_n.
    (c) Soient $(Z_1, \ldots, Z_p)$ des variables aléatoires mutuellement indépendantes suivant une loi $\mathcal{G}(p)$. Démontrer que leur minimum suit aussi une loi géométrique.
    Application : Supposons que la durée de vie moyenne d'une ampoule électrique est de n années. Sachant que cette durée de vie suit une loi géométrique, au bout de combien de temps devrez vous en moyenne changer une ampoule dans une pièce éclairée par dix luminaires ?

    les questions qui suivent sont plus difficiles :
    (d) On revient au cas général : Trouver la limite de $P(X_n = k)$ pour tout entier k.
    (e) On suppose que Z_1 possède une espérance, trouver la limite de $E(X_n)$.

    (f) Traiter de même le cas du maximum $Y_n = \max(Z_1, \ldots, Z_n)$.
    (g) Dans le cas d'une suite (Z_n)de variables aléatoires vérifiant la loi uniforme sur $[1, b]$, donner un développement asymptotique de $E(Y_n)$ quand n tend vers l'infini.

21. Conditionnement de lois géométriques.

    Soient X et Y deux variables aléatoires indépendantes suivant une loi $\mathcal{G}(p)$. Déterminer la loi de X conditionnellement à l'événement $X + Y = n$.

22. *classique.* Poissonisation.

On considère une suite tirages pile ou face indépendants, l a probabilité de tirer pile étant égale à p . On note T_i la variable aléatoire qui prend la valeur 1 si le $i$ème tirage est pile et zéro sinon. Pour $n \in \mathbb{N}*$, On pose X_n et Y_n les variables aléatoires égales respectivement au nombre de pile et au nombre de face tirés pendant les n premiers tirages. (autrement dit : $X_n = \sum_1^n T_i$ et $Y_n = n - X_n$)

(a) Dans cette question on fixe un entier n. Quelle est la loi de X_n ? de Y_n ?
Ces deux variables aléatoires n'ont aucune chance d'être indépendantes. Pourquoi ?

On souhaite y remédier, pour cela on va faire varier le nombre n de tirage. On introduit donc une variable aléatoire N à valeurs dans $\mathbb{N}^*$. On suppose cette variable aléatoire indépendante des T_i. On note alors X_N la variable aléatoire qui prend la valeur X_n lorsque $N = n$. On définit de même Y_N. En particulier on a la formule $X_N + Y_N = N$

(b) Dans cette question on suppose que N suit une loi de Poisson de paramètre λ.
(i) Montrer que $\mathbb{P}(X_N = k) = \sum_{n=k}^{\infty} \mathbb{P}(N = n)\mathbb{P}(X_n = k)$. En déduire $\mathbb{P}(X_N = k)$.
(ii) Démontrer que X_N et Y_N sont indépendantes et suivent une loi de Poisson

(c) Réciproque. Dans cette question on suppose que X_N et Y_N sont indépendantes.

(i) Soit $g(u, v) = \mathbb{E}[u^{X_N} v^{Y_N}]$. Montrer que $g(u, v) = g(u, 1)g(1, v)$
(ii) Montrer que $g(u, v) = G(up + vq)$ où G désigne la fonction caractéristique de N
(iii) Etablir que la fonction $h : x \mapsto G(x + 1)$ vérifie sur un intervalle non trivial la relation

$$h(x + y) = h(x)h(y)$$

(iv) Un déduire l'existence de λ tel que $G(u) = e^{\lambda(u-1)}$ puis conclure que N suit une loi de Poisson.

23. Le collectionneur de coupons

Un collectionneur souhaite acquérir tous les coupons d'une série de n, mais lors de chaque achat (les achats sont indépendants) il ignore quel coupon il achète. On s'intéresse au temps nécessaire pour obtenir l'ensemble des coupons.

Soit $X_{i,n}$ la variable aléatoire égale au temps nécessaire pour obtenir le $i$ème coupon après avoir obtenu les $i - 1$ premiers et T_n la variable aléatoire égale au temps d'attente de la collection complète.

(a) Justifier que $X_{i,n} \sim \mathcal{G}(\frac{n - i + 1}{n})$
(b) Calculer l'espérance de T_n et en donner un équivalent quand n tend vers l'infini.

24. Loi hypergéométrique.
On tire au hasard 5 paires d'écouteurs dans un lot de 15 dont 3 sont défectueux.
Déterminer la loi et l'espérance du nombre d'écouteurs non défectueux tirés.

25. Loi hypergéométrique.
Soient $N \geq n$ des entiers et $p \in [0, 1]$ tel que $Np = N_1$ soit un entier.
On dit qu'une variable aléatoire X vérifie la loi hypergéométrique $\mathcal{H}(N, n, p)$ si $X(\Omega) \subset [0, n]$ et

$$P(X = k) = \frac{\binom{N_1}{k}\binom{N_2}{n-k}}{\binom{N}{n}}$$

(a) Interpréter la loi hypergéométrique à l'aide de tirage de boules dans une urne contenant des boules de deux couleurs.
(b) Montrer que son espérance vaut np et sa variance $npq\frac{N - n}{N - 1}$
(c) Quelle est la limite d'une loi hypergéométrique quand Ntend vers l'infini, p restant constant ?

26. Matrice aléatoire.

Soient X, Y, Z trois variables aléatoires indépendantes suivant une loi $\mathcal{P}(\lambda)$

On note $M = \begin{pmatrix} X & Y \\ 0 & Z \end{pmatrix}$

Déterminer la probabilité que M soit diagonalisable.

Fonction génératrice.

27. (CCP) Un sac contient 4 boules, une numérotée 0, 2 numérotées 1 et une numérotée 2.

    (a) Rappeler le lien entre les fonctions génératrices de X, Y et $X+Y$ lorsque X et Y sont indépendantes.

    (b) On tire n fois dans le sac et on calcule la somme S_n des numéros tirés. Calculer l'espérance et la variance de S_n.

28. On considère la variable X_n égale au nombre de cycles disjoints d'une permutation de l'ensemble $[1, n]$

    (a) Calculer la loi de X_3 et son espérance.

    (b) Etablir la formule $P(X_{n+1} = i) = P(X_n = i-1)\dfrac{1}{n+1} + P(X_n = i)\dfrac{n}{n+1}$ pour tout $i, n+1 > i > 1$.

    (c) En déduire que si G_n désigne la fonction génératrice de la loi de X_n on a $G_{n+1}(t) = \dfrac{n+t}{n+1} G_n(t)$.

    (d) En déduire l'expression factorisée du polynôme G_N

    (e) Calculer l'espérance de X_n

29. *classique.* Peut on piper deux dés de façon que la somme des points des deux dés suive la loi uniforme sur $\{2, ..., 12\}$?

30. Un pion se déplace sur des cases numérotées par des entiers naturels. Initialement il se trouve sur la case 0 et à chaque étape il se déplace d'un nombre strictement positif de cases. On note Y_i la variable aléatoire donnant le nombre de cases parcourues lors de la $i$ème étape. On suppose que les Y_i sont indépendantes de même loi.

    On note $S_n = \sum_1^n Y_i$, $f_i = P(Y_1 = i)$ et $u(t) = \sum_1^\infty f_i t^i$.

    On note E_k l'événement : "le pion passe par la case k" et on pose $u_k = P(E_k)$

    (a) Calculer $P(E_k | Y_1 = j)$ pour $k \geq j$

    (b) En déduire pour tout $k > 0$ $u_k = \sum_1^k f_j u_{k-j}$

    (c) On pose $f(t) = \sum_0^\infty u_k t^k$. Montrer que $u(t) = \dfrac{1}{1 - f(t)}$

    (d) Calculer $f(t)$ dans le cas ou $Y_1 - 1$ suit une loi de Bernoulli. Calculer dans ce cas la limite de la suite (u_n)

    (e) *difficile.* On suppose que Y_1 prend un nombre fini de valeurs, et que les entiers k pour lesquels $P(Y_1 = k)$ est non nul sont premiers entre eux dans leur ensemble.

        Montrer que la suite u_n converge vers $\dfrac{1}{E(Y_1)}$

31. Loi binomiale négative (ou loi de Pascal).

    Considérons une suite $(X_n)_n$ de variables indépendantes vérifiant une loi de Bernoulli à valeurs dans $\{0, 1\}$ de paramètre p. Soit r un entier. On note T_r la variable aléatoire égale à l'instant n ou se produit le $r$ième succès. Autrement dit : $T_r = \inf\{n, X_1 + ... + X_n = r\}$. Cette variable aléatoire est à valeurs dans $[r, +\infty]$.

    (a) Etablir que $P([T_r = n])) = \begin{pmatrix} n-1 \\ r-1 \end{pmatrix} p^r q^{n-r}$.

        Cette loi est appelée loi binomiale négative ou loi de Pascal.

    (b) Montrer que cette loi est aussi celle de la somme de r variables aléatoires indépendantes suivant la loi géométriques de paramètre p.

(c) En déduire son espérance et la somme de sa série génératrice.

(d) Retrouver la somme de sa série génératrice en utilisant les propriétés des séries entières.

32. *classique.* Somme aléatoire. Identité de Wald
Soit N une variable aléatoire à valeurs entières , et $(X_n)_n$ une suite de variables aléatoires entières mutuellement indépendantes suivant la loi d'une variable X. On suppose également ces variables indépendantes de N. On note S la variable aléatoire $X_1 + ...X_N$.

    (a) Démontrer que $G_S = G_N \circ G_X$.

    (b) Démontrer $E(S) = E(X)E(N)$ (identité de Wald)

33. Soient $X_1, \ldots, X_n$ des variables indépendantes suivant la loi $\mathcal{B}(1/2)$.
On note N_n le nombres de blocs de 0 ou de 1 consécutifs dans la suite $X_1, \ldots, X_n$.

    (a) Exprimer $P(N_n = k, X_n = 0)$ en fonction de $P(N_{n-1} = k, X_{n-1} = 0)$ et $P(N_{n-1} = k-1, X_{n-1} = 1)$

    (b) En déduire que $E(t^{N_n}) = \dfrac{t(1+t)^{n-1}}{2^{n-1}}$

    (c) Donner la loi, l'espérance et la variance de N_n

Variance, Moments. Couples de variables , covariance.

34. Montrer que pour une variable aléatoire à valeurs dans $\mathbb{N}$ on a $V(X) = \displaystyle\sum_0^{\infty}(2k+1)P(X > k)$

35. Propriété extrémale liée à la variance.
Soient X, et Y deux variables aléatoires réelles telles que $V(X) > 0$ et $V(Y) > 0$.
Trouver $(a, b) \in \mathbb{R}^2$ rendant minimale la quantité $f(a, b) = E((Y - aX - b)^2)$ (pour cet exercice on pourra utiliser sans justification que le minimum est atteint quand les deux dérivées partielles sont nulles).

36. On considère deux événements A, B et les variables aléatoires X, Y indicatrices de ces événements. Démontrer qu'elles sont indépendantes si et seulement si elles ne sont pas corrélées.

37. Déterminant aléatoire.
Soit $(X_{i,j})_{1\leq i,j\leq 2}$ une famille de variables aléatoires indépendantes de même loi, centrées et ayant une variance.
a) Déterminer l'espérance de leur déterminant et la variance de leur déterminant.
b) (plus difficile) Reprendre la question avec n quelconque.

38. Soient X, Y des variables aléatoires indépendantes, centrées réduites.

    Démontrer que $E(\max(X, Y)) \leq \dfrac{1}{\sqrt{2}}$.

    indication : Majorer $E(|X - Y|)$ en utilisant l'inégalité de Cauchy Schwarz

39. Inégalité de Cantelli

    (a) Soit Y une variable aléatoire non presque nulle, ayant une variance, et d'espérance positive.
    Montrer que
    $$P(Y > 0) \geq \frac{E(Y)^2}{E(Y^2)}$$
    on pourra introduire la variable aléatoire $Y.\mathbf{1}_{Y>0}$ et appliquer l'inégalité de Schwarz.

    (b) Soit X une variable aléatoire ayant un moment d'ordre 2 Déduire de ce qui précède que pour tout $\epsilon > 0$ on a l'inégalité :
    $$P(X - E(X) < -\epsilon) \leq 1 - \frac{\epsilon^2}{V(X) + \epsilon^2}$$

loi conjointe. Couples aléatoires.

40. *classique.*

Soit N une variable aléatoire suivant la loi de Poisson de paramètre λ.

Un promeneur ramasse N champignons. Chaque champignon est avec probabilité p comestible. On note X le nombre de champignons comestibles.

Déterminer la loi du couple X, N et en déduire la loi de X.

Les variables $X, N-X$ sont elles indépendantes ?

41. On considère un entier naturel n et l'ensemble I_n défini par

$$I_n = \{(k,\ell)/\ k \in [\![0,n]\!] \quad \text{et} \quad \ell \in [\![0,n]\!] \quad \text{et} \quad k+\ell \leqslant n\}$$

Un espace probabilisé $(\Omega, \mathcal{T}, P)$ étant donné, ainsi que trois réels strictement positifs p, q et r vérifiant $p+q+r=1$, on considère un couple aléatoire (X_n, Y_n) défini sur $(\Omega, \mathcal{T}, P)$, à valeurs dans I_n, et tel que, pour tout couple $(k,\ell) \in I_n$: $P\left((X_n, Y_n) = (k,\ell)\right) = \dfrac{n!}{k!l!(n-k-l)!} p^k q^\ell r^{n-k-\ell}$

(a) Vérifier que : $\displaystyle\sum_{(k,\ell)\in I_n} \frac{n!}{k!l!(n-k-l)!} p^k q^\ell r^{n-k-\ell} = 1.$

(b) Montrer que les variables aléatoires X_n et Y_n suivent toutes deux une loi binomiale (en préciser les paramètres respectifs). Ces variables aléatoires sont elles indépendantes ?

(c) On se propose de calculer la covariance du couple (X_n, Y_n).

i. On suppose que $n \geqslant 2$. En utilisant la relation

$$kl \frac{n!}{k!l!(n-k-l)!} = n(n-1) \frac{(n-2)!}{(k-1)!(l-1)!(n-k-l)!}$$

Montrer que $E(X_n Y_n) = n(n-1)pq$.

ii. Cette relation est-elle encore vraie si $n = 0$? si $n = 1$?

iii. En déduire la valeur de la covariance $\text{Cov}(X_n, Y_n)$ du couple (X_n, Y_n) et leur coefficient de corrélation.

42. On se donne deux variables X et Y à valeurs entières. On suppose que X a une espérance finie, que Y est majorée par X et suit la loi uniforme sur $[0, X]$.

a) Calculer l'espérance de Y en fonction de celle de X

b) On suppose que Y suit une loi géométrique . Montrer que les variables Y et $X-Y$ sont indépendantes.

43. On considère un couple de variables aléatoires (X,Y) à valeurs entières dont la loi conjointe est donnée par $P(X=j, Y=k) = \dfrac{(j+k)}{e2^{j+k} j!k!}$.

a) Trouver la loi marginale de X

c) Calculer $\mathbb{E}(2^{X+Y})$

44. Un exemple de loi conditionnelle :

On lance indéfiniment une pièce amenant pile avec la probabilité p et face avec la probabilité $1-p$. Soit r un entier. On note X la variable aléatoire égale au nombre de lancers nécessaires pour obtenir r pile consécutifs. On note Y la variable aléatoire égale au rang du premier face. On admet que X possède une espérance finie.

(a) Déterminer la loi conditionnelle de X sachant $Y > r$, puis $\mathbb{E}(X|[Y > r])$

(b) Pour tout $i \in \{1, .., r\}$ exprimer $\mathbb{E}(X|[Y = i])$ à l'aide de $\mathbb{E}(X)$

(c) En déduire l'espérance de X

Inégalités de Markov et Bienaymé Tchebychev.

45. Surbooking

    Une compagnie aérienne a constaté que 8% des passagers ayant réservé leur vol ne se présentent pas à l'embarquement. Elle décide donc d'accepter 420 passagers pour un vol sur un avion ne comportant que 400 places. En utilisant l'inégalité de Bienaymé Tchebychev, majorer la probabilité qu'elle ne puisse pas embarquer tous les passagers se présentant le jour du vol.

46. Une variable aléatoire possède une moyenne égale à 10 et un écart type de 1.

    (a) Montrer que la variable X prend ses valeurs dans $[0, 20]$ avec une probabilité d'au moins $0,99$.

    (b) On répète de façon indépendante 25 fois l'expérience modélisée par la variable X. Avec quelle probabilité la moyenne des résultats se trouve t-elle dans $[8, 12]$?

47. On suppose que X est une variable aléatoire ayant un moment d'ordre r. Montrer que la suite $\mathbb{P}(|X| > n)$ est un $O(\frac{1}{n^r})$.

    On suppose que e^X possède une espérance, montrer que pour tout entier r, la suite $\mathbb{P}(|X| > n)$ est un $O(\frac{1}{n^r})$.

48. Points fixes d'une permutation.

    Soit n un entier. Sur l'ensemble σ_n des permutations, on note X_i la variable aléatoire égale à 1 si i est un point fixe et zéro sinon. On pose aussi $S_n = X_1 + ... + X_n$ ainsi S_n est la variable aléatoire qui donne le nombre de points fixes.

    (a) Pour chaque X_i calculer l'espérance et la variance.

    (b) Calculer $\mathbb{E}(X_k X_l)$. Calculer le coefficient de corrélation de X_k, X_l et interpréter le signe.

    (c) Calculer l'espérance et la variance de S_n.

    (d) On suppose $n > 11$ Montrer que la probabilité que S_n soit au moins égal à 11 est inférieure à 0.01

49. Nombre de zéros d'un polynôme aléatoire.

    (a) Soit, pour $\lambda \in \mathbb{R}^{+*}$, X_λ une variable aléatoire suivant une loi de Poisson de paramètre λ. Soit $\epsilon > 0$. Montrer que $P(|X_\lambda - \lambda| \geq \epsilon\lambda) \to 0$ lorsque $\lambda \to +\infty$.

    (b) *difficile.* Soient, pour $\lambda \in \mathbb{R}^{+*}$, $X_\lambda, Y_\lambda, Z_\lambda$ des variables aléatoires indépendantes suivant une loi de Poisson de paramètre λ. Déterminer la limite, lorsque $\lambda \to +\infty$, de la probabilité que le polynôme $A_\lambda X^2 + B_\lambda X + C_\lambda$ ait toutes ses racines réelles.

Etude asymptotique. Inégalités de Markov et Bienaymé Tchebytchev. Loi faible des grands nombres.

50. Soit X une variable aléatoire suivant une loi de poisson de paramètre λ.

    a) Montrer que $P(X > 2\lambda) \leq \frac{1}{\lambda}$. Cette majoration n'est en général pas très bonne. On cherche à l'améliorer.

    b) Soit $G_X(t) = e^{\lambda(t-1)}$ la fonction génératrice de X. Démontrer que pour tout a , et tout $t > 1$ $P(X \geq a) \leq \frac{G_X(t)}{t^a}$

    c) En déduire $P(X > 2\lambda) \leq (\frac{e}{4})^\lambda$

    d) Dernière méthode : Montrer que le reste R_n de la série exponentielle est majoré par $\frac{n+1}{n+1-\lambda}$ fois son premier terme.

    En déduire une nouvelle majoration de $P(X > 2\lambda)$

51. Soit a un réel positif.
    on pose

    $$x_n(a) = \sum_{k=0}^{[an]} \frac{n^k}{k!}$$

    Démontrer que pour $a < 1$ $x_n(a) = o(e^n)$.

    Que se passe t'il pour $a > 1$?

52. loi faible appliquée à la loi binomiale.

Un jury est composé de $2n+1$ personnes. Chacun prend sa décision de façon indépendante et possède la probabilité $p$de rendre une décision juste.

(a) Déterminer la probabilité p_n pour que le jury prenne à la majorité une décision juste.

(b) Déterminer la limite de p_n quand n tend vers l'infini.

53. Points isolés d'un graphe aléatoire

Soit G un ensemble de cardinal n. Un graphe aléatoire sur G est une famille de $\frac{n(n-1)}{(2)}$ variables aléatoire de Bernoulli de paramètre p_n mutuellement indépendantes. (deux points de G sont reliés par une arête avec probabilité p_n).

(a) Soit X_n le nombre de sommets isolés. Vérifier que X_n est la somme de n variables de Bernoulli. Quelle est l'espérance de X_n ?

(b) Démontrer que La variance de $V(X_n) = n(n-1)(1-p_n)^{2n-3} + E(X_n) - E(X_n)^2$.

(c) On suppose que $np_n - \ln n$ tend vers l'infini. Trouver la limite de $E(X_n)$. En déduire que $P(X_n \geq 1)$ tend vers 0.

(d) On suppose plutôt $np_n \to a$ avec$a > 0$. Donner un équivalent de $E(X_n)$, $V(X_n)$ puis montrer que $P(X_n \geq 1)$ tend vers 1.

(e) Commenter le résultat obtenu.

54. Convergence presque sure.
Soit (X_n) une suite de variables aléatoire définies sur un même espace probabilisé $(\Omega, \mathfrak{A}P,)$.

(a) Soit X une autre variable aléatoire.
Montrer que l'ensemble $\{\omega \in \Omega, X_n(\omega) \to X(\omega)\}$ est un événement.

On introduira l' événement $E_n(\epsilon) = \{|X_n - X| > \epsilon\}$, et on écrira l'ensemble cherché à l'aide d'inclusions réunions.

On dit que la suite X_n converge presque surement (p.s) vers X lorsque cet événement est presque sûr.

(b) On note $Y_n = \sup\limits_{p \geq n} |X_p - X|$. (cette variable aléatoire est à valeurs dans $\overline{\mathbb{R}}$)
Montrer que X_n CV presque surement vers X si et seulement si $P(Y_n > \epsilon)$ tend vers 0 pour tout $\epsilon > 0$.

(c) En déduire que si la série de terme général $P(|X_n - X| > \epsilon)$ converge pour tout ϵ alors la suite (X_n) converge p.s vers X.

55. La loi forte des grands nombres dans le cas L^2.

Cet exercice utilise les propriétés de la convergence presque sure vue dans l'exercice précédent.

Soit $(X_n)_n$ une suite de variables aléatoires centrées i.i.d possédant un moment d'ordre 2.

On pose $S_n = \sum\limits_{k=1}^{n} X_k$.

(a) Soit $\varepsilon > 0$. Montrer que la suite $\mathbb{P}(\frac{|S_n|}{n} > \varepsilon)$ est un $O(\frac{1}{n})$.

(b) En déduire que la suite $\frac{|S_{n^2}|}{n^2}$ converge presque sûrement vers 0.

(c) On pose $T_n = \frac{1}{n} \sum\limits_{[\sqrt{n}]^2+1}^{n} X_k$. Démontrer que $\mathbb{P}(|T_n| > \varepsilon)$ est un $O(\frac{1}{n^{3/2}})$.

(d) En déduire que la suite $(\frac{S_n}{n})_n$ converge presque sûrement vers 0.
(loi forte des grands nombres).

56. (grand classique) Démonstration probabiliste du théorème de Weierstrass

Soit f une fonction continue sur l'intervalle $[0,1]$

On note $B_n(f)$ le polynôme $\sum_0^n f(\frac{k}{n})\begin{pmatrix} n \\ k \end{pmatrix} x^k(1-x)^{n-k}$

et on considère une suite (X_n) de variables de Bernoulli indépendantes, et de paramètre x.

Enfin on pose $S_n = X_1 + ... + X_n$

a) Montrer que $B_n(f)(x) = \mathbb{E}(f(\frac{S_n}{n}))$ et $B_n(f)(x) - f(x) = \mathbb{E}(f(\frac{S_n}{n}) - f(x))$.

b) Pour $\delta > 0$ on pose $a(\delta) = \sup\{|f(x) - f(y)|, |x-y| \leq \delta\}$. Justifier que $a(\delta)$ tend vers zéro quand δ tend vers zéro.

c) Soit $A_{n,\delta}$ l'événement $|\frac{S_n}{n} - x| \geq \delta$. Démontrer que $\mathbb{E}(f(\frac{S_n}{n}) - f(x)) \leq a(\delta) + 2P(A_{n,\delta})||f||_\infty$

d) A l'aide de l'inégalité de Tchebychev, démontrer que$P(A_{n,\delta}) \leq \frac{x(1-x)}{n\delta^2} \leq \frac{1}{4n\delta^2}$

e) En déduire que la suite de polynômes $(B_n(f))$ converge uniformément vers f sur $[0,1]$

57. le théorème central limite dans le cas des variables vérifiant une loi de Poisson

Soit λ un réel positif. On considère une suite (X_n) de variables aléatoires indépendantes suivant toute une loi de Poisson de paramètre λ. On pose $S_n = \frac{X_1 + ... + X_n}{n}$

a) Rappeler la loi de $X1 + .. + X_n$. En déduire l' espérance et l'écart type σ_n de S_n. On pose ensuite $Y_n = \frac{S_n - \lambda}{\sigma_n}$ (cette transformation a pour effet de rendre la variable aléatoire centrée et réduite, ainsi Y_n peut être vue comme la moyenne "normalisée" des variables aléatoires $X_1, ...X_n$)

On se propose alors de démontrer le théorème suivant : pour tous réels x, y , lorsque n tend vers l'infini on a :

$$P(x \leq Y_n \leq y) \to \frac{1}{\sqrt{2\pi}} \int_x^y e^{\frac{-t^2}{2}} dt$$

Ce résultat très important est appelé le théorème central limite. Il est en fait vrai pour toute suite de variables aléatoires de même loi, du moment que ces variables aléatoires ont une variance. Il est clair (du moment qu'on choisit une bonne tribu sur $\mathbb{R}$) que la loi "limite" définie par l'égalité $P([x,y]) = \frac{1}{\sqrt{2\pi}} \int_x^y e^{\frac{-t^2}{2}} dt$ est bien une loi de probabilité (on dit qu'il y a convergence en Loi : cf autre exercice). Elle est appelée Loi gaussienne et le théorème central limite explique son importance : elle modélise le comportement moyen de toutes les suites de variables aléatoires indépendantes de même loi ayant une variance.

a) Soit X une VA suivant une loi de Poisson de paramètre μ. Démontrer que pour tout entier n on a $P(X \leq n) = \frac{1}{n!} \int_\mu^\infty e^{-t} t^n dt$

b) Dans la suite de l'exercice, un réel x est fixé. On note $p_n = P(Y_n \leq x)$. En déduire que pour tout réel y, $p_n = P(nS_n \leq x\sqrt{\lambda n} + \lambda n) = \frac{1}{N!} \int_{\lambda n}^\infty e^{-t} t^N dt$, où N est la partie entière de $x\sqrt{\lambda n} + \lambda n$.

c) A l'aide d'un changement de variable judicieux, Vérifier la formule

$$p_n = \frac{e^{-N} N^{N+\frac{1}{2}}}{N!} \int_{\frac{n\lambda - N}{\sqrt{N}}}^\infty e^{-\sqrt{N}u} (1 + \frac{u}{\sqrt{N}})^N du$$

d) Montrer que la suite $v_n = \frac{n\lambda - N}{\sqrt{N}}$ converge vers $-x$

e) Soit f_N la suite de fonction définie sur $\mathbb{R}$ par $f_N(u) = e^{-\sqrt{N}u}(1 + \frac{u}{\sqrt{N}})^N$ si $u \geq -x$ et égale à zéro sinon. Démontrer la convergence simple de f_N vers une fonction continue par morceaux sur $\mathbb{R}$ qu'on précisera.

f) (*) Soit a un réel positif et strictement plus grand que x. Etablir la majoration $f_n(t) \leq (1 + t/a)e^{-t/a}$ sur l'intervalle $[-a, +\infty[$ pour n assez grand

g) Conclure.