Année universitaire 2001-2002

UNIVERSITÉ D'ORLÉANS

Olivier GARET

MA4.04a : Probabilité et Statistiques

Table des matières

Chapitre 1

Statistique descriptive

Faire des statistiques, c'est recueillir des données et les interpréter.

1.1 Définitions

- L'ensemble sur lequel porte l'étude statistique est appelé population.
- Chaque élément de la population s'appelle un individu.
- L'aspect étudié s'appelle la variable ou le caractère.
- Une variable peut prendre plusieurs valeurs ("les x_i").
- Les classes sont des ensembles (souvent des intervalles) regroupant plusieurs valeurs prises par la variable.
- Si une variable prend des valeurs numériques (je veux dire par là que l'on peux comparer ces valeurs), on dit su'elle est quantitative. Sinon on dit qu'elle est qualitative.
- Le nombre d'individus pour lesquels la variable prend une valeur donnée s'appelle l'effectif de cette valeur (noté n_i)
- L'effectif total est la somme des effectifsde toutes les valeurs de la variable, c'est à dire le nombre d'individus de la population :
$$N = n_1 + n_2 + \dots n_p.$$
- La fréquence d'une valeur du caractère est le quotient de l'effectif relatif à cette valeur par l'effectif total : $f_i = \frac{n_i}{N}$
- Les fréquences cumulées croissantes sont :
$$f_1, f_1 + f_2, \dots, f_1 + f_2 + \cdots + f_p = 1.$$
- Les fréquences cumulées décroissantes sont :
$$f_1 + f_2 + \cdots + f_p = 1, f_2 + f_3 + \cdots + f_p, f_3 + f_4 + \cdots + f_p, \dots, f_p.$$

1.2 Moyenne

On considère un caractère x prenant les valeurs $x_1, x_2, \dots x_p$ avec les effectifs $n_1, N_2, \dots, n_p$. On appelle moyenne de cette série statistique x le nombre $\overline{x}$ défini par

$$\overline{x} = \frac{n_1 x_1 + n_2 x_2 + \cdots + n_p x_p}{n_1 + n_2 + \cdots + n_p} = \frac{1}{N} \sum_i n_i x_i$$

Remarque : On a aussi

$$\overline{x} = \sum_i n_i f_i.$$

Quelques propriétés Pour des caractères quantitatifs x et y étudiés sur la même population, on peut s'intéresser au caractère $z = x + y$. On peut démontrer la formule suivante : $\overline{x+y} = \overline{x} + \overline{y}$

De même, si λ est une constante réelle (ou une variable quantitative constante dans la population observée), on a $\overline{\lambda} = \lambda$.

Qui plus est, on a la propriété : $\overline{\lambda x} = \lambda \overline{x}$.

Exercice Pour a et b réels, que vaut $\overline{ax+b}$?

1.3 Médiane

On appelle médiane d'un caractère quantitatif $x = (x_1, \dots, x_p)$ tout nombre μ tel que

- le nombre d'individus pour lesquels le caractère x est strictement inférieur à μ est au plus $N/2$.
- le nombre d'individus pour lesquels le caractère x est strictement inférieur à μ est au plus $N/2$. Ainsi, si a est l'entier tel que

  $$n_1 + n_2 + \cdots + n_a \leq N/2 < n_1 + n_2 + \cdots + n_{a+1}$$

  et b l'entier tel que

  $$n_b + n_{b+1} + \cdots + n_p \leq N/2 < n_{b-1} + n_b + \cdots + n_p,$$

  alors les médianes de x sont les réels

  $$\mu \in [x_a, x_b].$$

Lorsque $x_a \neq x_b$ et que l'on veut, par simplicité, parler de "la" médiane de x, on peut prendre $\frac{x_a + x_b}{2}$ comme médiane, mais ce choix est arbitraire.

1.4 Mode

On appelle mode d'une série statistique une valeur du caractère dont l'effectif est le plus grand.

Exemple : On demande à des patients en ophtalmologie quelle couleur de teinture des lunettes leur est le plus agréable. On obtient les résultats suivants :

x_i	rouge	vert	bleu	jaune
n_i	41	6	41	12

Cette série statistique possède deux modes : le rouge et le bleu.

Exercice A-t-on étudié ici un caractère quantitatif ou qualitatif ?

1.5 Paramètres de dispersion

1.5.1 Variance, écart-type

Si x est un caractère quantitatif, on appelle variance de x la moyenne du carré de la distance de x à sa moyenne. Autrement dit

$$\text{Var } x = \overline{(x-\overline{x})^2} = \sum_i f_i(x_i - \overline{x})^2.$$

On a la formule de König :

$$\text{Var } x = \overline{x^2} - \overline{x}^2 = \sum_i f_i(x_i)^2 - \big(\sum_i f_i x_i\big)^2$$

On définit l'écart type $\sigma(x)$ comme étant la racine carrée de la variance : $\sigma(x) = \sqrt{\text{Var } x}$.

1.5.2 Étendue

L'étendue d'un caractère quantitatif est la différence entre la plus grande valeur prise et la plus petite valeur prise.
Exercice Pourquoi ce paramètre est-il peu significatif ?

1.5.3 Quartiles

Coupons en deux une population sur laquelle un caractère quantitatif est étudié suivant la position de leur valeur du caractère par rapport à la médiane. (ex : après une interrogation écrite : la première moitié et la deuxième moitié de la classe)

- La première moitié de la première moitié est appelée premier quartile.
- La deuxième moitié de la première moitié est appelée deuxième quartile.
- La première moitié de la deuxième moitié est appelée troisième quartile.
- La deuxième moitié de la deuxième moitié est appelée dernier quartile.

Notons Q_2 une médiane de l'effectif total, Q_1 une médiane de la première moitié, Q_3 une médiane de la deuxième moitié.
On appelle écart interquartile le nombre $Q_3 - Q_1$.

1.6 Exercices

1. On a classé 5000 groupes de 8 individus d'après le nombre d'individus du groupe ayant une caractéristique génétique donnée. On a obtenu le tableau suivant :

Nombre d'individus du groupe x_i	0	1	2	3	4	5	6	7
Nombre de groupes n_i	500	1331	1561	1037	435	114	19	3

   (a) Quelle est la population étudiée ? Quelle est la variable ? Quelle est sa nature ? Quels sont les valeurs possibles de cette variable sur la population ?

   (b) Déterminer la moyenne m et l'écart-type σ de cette série statistique.

2. Une étude du rendement sur cent parcelles de terre (de $25m2$) d'un meme champ de mais a donné les résultats suivants, exprimés en kgs :

Rendements	16.2	16.4	16.6	16.8	17	17.2	17.4	17.6	18	18.2	18.4
Effectif	1	5	7	14	23	25	12	8	2	2	1

   Calculer la moyenne, la variance et l'écart-type de cette série statistique.
   Déterminer la médiane, le mode et les quantiles.

Chapitre 2

Rappels de dénombrement

2.1 Rappels de vocabulaire ensembliste

Un ensemble Ω est constitué de points, tous distincts. On dit qu'un ensemble A est inclus dans Ω, et l'on écrit $A \subset \Omega$ lorsque tous les éléments de A appartiennent à Ω.

On rappelle que l'ensemble vide (noté $\varnothing$) ne contient aucun élément et est inclus dans tous les ensembles

Pratiquement, si l'on veut montrer le résultat $A \subset \Omega$, la preuve ressemblera donc à "Soit $x \in A$...(raisonnement)...donc $x \in \Omega$. Comme on a choisi x quelconque dans A, on peut conclure que $A \subset \Omega$."

Si A est inclus dans Ω, on dit que A est un sous-ensemble, ou encore une partie de Ω.

Si A et B sont des parties de Ω, l'ensemble $A \cup B$ est constitué des éléments de Ω qui sont dans A ou dans B, éventuellement dans les deux. Plus généralement, si I est un ensemble quelconque et $(A_i)_{i \in I}$ une famille de parties de Ω indexé par I, $\underset{i \in I}{\cup} A_i$ est constitué des points de Ω qui sont dans au moins un des A_i.

Pratiquement, si l'on veut montrer le résultat $x \in \underset{i \in I}{\cup} A_i$, , la preuve ressemblera donc à "...(raisonnement)...Il existe donc $i_0 \in I$ tel que $x \in A_{i_0}$. Donc $x \in \underset{i \in I}{\cup} A_i$."

Si A et B sont des parties de Ω, l'ensemble $A \cap B$ est constitué des éléments de Ω qui sont dans A et dans B. Plus généralement, si I est un ensemble quelconque et $(A_i)_{i \in I}$ une famille de parties de Ω indexé par I, $\underset{i \in I}{\cap} A_i$ est constitué des points de Ω qui sont dans tous les A_i.

Pratiquement, si l'on veut montrer le résultat $x \in \bigcap_{i\in I} A_i$, , la preuve ressemblera donc à "Soit $i \in I$...(raisonnement)...Donc $x \in A_i$. Comme i est quelconque, on a donc $x \in \bigcap_{i\in I} A_i$."

2.2 Applications et cardinaux : définitions et notations

Pour A, D deux ensembles non vides quelconques, on note A^D ou $\mathcal{F}(D, A)$ l'ensemble des fonctions de D (ensemble de départ) vers A (ensemble d'arrivée). Soit f une application de A dans D. On dit que f est

- *injective* si $\forall x, y \in D \quad x \neq y \Longrightarrow f(x) \neq f(y)$
- *surjective* si $\forall z \in A \quad \exists x \in D \quad f(x) = z$.
- *bijective* si elle elle à fois injective et surjective.

Une application injective (resp. surjective, bijective) est une injection (resp. surjection, bijection).

Une bijection d'un ensemble Ω dans lui-même est appelée *permutation* de Ω.

Un ensemble est Ω dit *fini* si

- ou bien c'est l'ensemble vide. ($\varnothing$)
- ou bien il existe un entier n tel qu'il existe une bijection entre Ω et $\{1, \dots, n\}$.

Cet entier n est unique : on l'appelle le *cardinal* de l'ensemble Ω. On le note $|\Omega|$. De manière intuitive, c'est le nombre d'éléments de Ω.

Le cardinal de l'ensemble vide est zéro.

Pour Ω fini de cardinal n, et $p \in \{0, \dots, n\}$, on note $\mathcal{B}_p(\Omega)$ l'ensemble des parties de Ω de cardinal p. Par exemple $\mathcal{B}_2(\{a, b, c\}) = \{\{a, b\}, \{b, c\}, \{a, c\}\}$. De même, on note $\mathcal{P}(\Omega)$ l'ensemble des parties de Ω, quelque cardinal qu'elles aient. Par exemple $\mathcal{P}(\{a, b, c\}) = \{\varnothing, \{a\}, \{b\}, \{c\}, \{a, b\}, \{b, c\}, \{a, c\}, \{a, b, c\}\}$.

Soit A et D deux ensembles finis. On admettra les résultats suivants :

- Il existe (au moins) une bijection de D dans A si et seulement si $|A| = |D|$.
- Il existe (au moins) une injection de D dans A si et seulement si $|A| \geq |D|$.
- Il existe (au moins) une surjection de D dans A si et seulement si $|A| \leq |D|$.

Le premier des trois résultats énoncés est évidemment le plus utilisé lorsque l'on veut des dénombrements exacts, alors que les deux autres sont plutôt utilisés dans les cas trop complexes où l'on peut juste espérer des encadrements.

2.3 Principes de base du dénombrement

2.3.1 Principe de bijection

Dans la pratique, lorsque l'on veut compter les éléments d'un ensemble, on montre que cet ensemble est en bijection avec un ensemble dont on connaît (par coeur !) le nombre d'éléments. La section suivante énoncera un certain nombre de résultats qu'il faut connaître.

2.3.2 Principe d'indépendance

C'est juste la formule

$$|A \times B| = |A|.|B| \quad .$$

Considérée isolément, elle peut paraître sans intérêt mais elle est souvent utilisée en association avec le principe de bijection.

2.3.3 Principe de partition

On dit que les ensembles $(A_i)_{i\in I}$ forment une partition de A si l'on a $A = \underset{i\in I}{\cup} A_i$ et $i \neq j \Longrightarrow A_i \cap A_j = \varnothing$. On a alors

$$|A| = \sum_{i\in I} |A_i|.$$

Exemple fondamental : Si B est un ensemble quelconque inclus dans Ω, les ensembles B et B^c forment une partition de Ω.

Le résultat élémentaire suivant peut souvent être utile.

Théorème 1. *Soit Ω un ensemble quelconque, I un ensemble d'index fini ou dénombrable et $(\Omega_i)_{i\in I}$ une partition de Ω. Alors, si l'on pose $A_i = A \cap \Omega_i$, les ensembles $(A_i)_{i\in I}$ forment une partition de A.*

Démonstration. Comme $\Omega = \underset{i\in I}{\cup} \Omega_i$, on a $A = A\cap\Omega = \underset{i\in I}{\cup} A\cap\Omega_i = \underset{i\in I}{\cup} A_i$. D'autre part, pour $i \neq j$, on a $A_i \cap A_j \subset \Omega_i \cap \Omega_j = \varnothing$, d'où $A_i \cap A_j = \varnothing$. $\square$

2.3.4 Lemme des bergers

Le lemme suivant peut également être utile

Lemme 1 (Lemme des bergers). *Soit ϕ une application* surjective *de A dans D. On suppose qu'il existe un entier $a \geq 1$ tel que*

$$\forall y \in A \quad |\{x \in D; \phi(x) = y\}| = a$$

(autrement dit si tout élément de A a exactement a antécédents), on a

$$|A| = \frac{|D|}{a}.$$

Démonstration. On applique le principe de partition avec $I = A$: si l'on pose, pour $y \in A$: $D_y = \{x \in D; \phi(x) = y\}$, les D_y forment clairement une partition de D, d'où

$$|D| = \sum_{y \in A} |D_y| = \sum_{y \in A} a = |A|a.$$

□

Le nom du lemme est du à la procédure prétendûment employée par les bergers chaldéens pour compter le nombre de leurs moutons : il s'agit de compter le nombre de pattes et de diviser par 4. Dans cet exemple A est l'ensemble des moutons, D l'ensemble des pattes de mouton, et ϕ l'application qui a une patte associe le mouton auquel elle appartient.

2.4 Quelques résultats incontournables

2.4.1 Choix indépendant (avec remise) de p objets dans un ensemble de cardinal $|A| = n$.

Il existe exactement n^p manières de choisir p fois un objet dans un ensemble de n objets, autrement dit il existe n^p résultats possibles lorsqu'on tire au hasard avec remise p objets dans un lot de n objets.

Exemple : Un professeur note chaque étudiant d'une classe de 30 étudiants par une note entière de 0 à 20. Le nombre de résultats possibles est le nombre de manières de choisir de manière indépendante 30 élémentsde l'ensemble $A = \{0, \ldots, 20\}$ des notes possibles. Comme $|A| = 21$, il y a donc 21^{30} résultats possibles.

2.4.2 Nombre de permutations de Ω

Le nombre de permutations d'un ensemble est le nombre de manière d'ordonner ses éléments.

On pose $|\Omega| = n$. Le nombre de permutations de Ω est

$$n(n-1)\dots 1.$$

On note $n! = n(n-1)\dots 1$.

Remarque : $n!$ se lit "*factorielle n*" .

Exemple : Un professeur doit faire passer dans la journée 5 étudiants à l'oral de contrôle. Il a 5 !=120 manières de choisir l'ordre dans lequel il va les interroger.

2.4.3 Nombre de tirages sans remise

On appelle tirage sans remise de p éléments d'un ensemble A tout tirage successif de p éléments de A, chaque élément ne pouvant être tiré plus d'une fois.

Bien évidemment, pour qu'un tel tirage puisse exister, il faut avoir $p \leq |A| = n$.

Alors, le nombre de tirages sans remise est

$$n(n-1)\dots(n-p+1).$$

Remarques :

- Comme on l'a vu dans la preuve, ce nombre peut s'écrire aussi $\frac{n!}{(n-p)!}$.
- Lorsque $n = p$, on trouve $n!$. En fait, choisir successivement n éléments dans un ensemble de cardinal n revient à les ordonner.

Exemple : 3500 personnes se présentent au concours de l'Agrégation de Biologie. 300 places sont mises aux concours. Combien y a-t-il de palmarès possibles, en supposant qu'il n'y ait pas d'ex-æquos ?

Réponse : $3500 \times 3499 \times \cdots \times 3202 \times 3201$.

2.4.4 Nombre de parties de Ω possédant p éléments

On pose $|\Omega| = n$. Par définition, on note $\binom{n}{p}$ le nombre de parties à p éléments d'un ensemble de n éléments. Il s'agit donc de calculer $|\mathcal{B}_p(\Omega)|$. C'est exactement le nombre de manières de choisir p objets dans un ensemble de n objets, l'ordre n'ayant aucune importance. On peut montrer que

$$\binom{n}{p} = \frac{n(n-1)\dots(n-p+1)}{p(p-1)\dots 1} = \frac{n!}{p!(n-p)!}.$$

Exemple : 3500 personnes se présentent au concours de l'Agrégation de Biologie. 300 places sont mises aux concours. Combien y a-t-il de listes alphabétiques des reçus possibles ? Réponse : $\binom{3500}{300}$. Ici Ω est l'ensemble des candidats et $p = 300$ le nombre de reçus.

2.4.5 Nombre total de parties de Ω

Le nombre total de parties de Ω est $|\mathcal{P}(\Omega)| = 2^{|\Omega|}$.

Démonstration. Il suffit de remarquer que l'application

$$\begin{array}{rcl} \mathcal{P}(\Omega) & \to & \{0;1\}^{\Omega} \\ A & \mapsto & \mathbb{1}_A \end{array}$$

est une bijection. On rappelle que pour $A \subset \Omega$, l'application $\mathbb{1}_A$ (appelée indicatrice de A) est définie sur Ω par

$$\mathbb{1}_A(x) = \begin{cases} 1 \text{ si } x \in A \\ 0 \text{ si } x \notin A \end{cases}$$

$\square$

Exemple : 200 étudiants se présentent à un examen. Combien y a-t-il de listes alphabétique des reçus possibles ? Réponse : 2^{200}. Ici Ω est l'ensemble des candidats. La grande différence avec l' exemple précédent est qu'ici, le nombre de reçus n'est pas fixé à l'avance.

2.5 Exercices

1. Au tapis vert, il faut choisir 4 cartes (as,roi, dame, valet, dix, neuf, huit ou sept) par couleur (pique, coeur, carreau ou trèfle). On appelle grille l'un des résultats possibles.
   exemple de grille : as de pique, dame de cœur, 7 de carreau, 7 de trèfle.
   Dénombrer les grilles contenant :
   (a) la dame de cœur
   (b) une dame et une seulement
   (c) deux dames et deux seulement
   (d) aucune dame
   (e) au moins une dame
   (f) l'as de pique et une dame seulement
2. 5 prévenus sont amenés à choisir un avocat dans une liste de 10 avocats commis d'office.
   (a) Combien y a-t-il de choix possibles ?
   (b) Combien y a-t-il de choix tels que les 5 prévenus choisissent le même avocats ?
   (c) Combien y a-t-il de choix tels que 2 avocats soient appelés ?
3. De combien de manières peut-on disposer n personnes autour d'une table ronde ? Précision : seule la position relative à la maîtresse de maison importe, une rotation ne change donc rien.
   Si $n = 2k$, avec k hommes et k femmes, combien de dispositions permettent une alternance des sexes ?
4. Dans une assemblée de n personnes, on note la date d'anniversaire de chacun.
   (a) Donner le cardinal de toutes les répartitions possibles.
   (b) Calculer le cardinal des ensembles suivants :
      i. tous les participants ont la meme date d'anniversaire.
      ii. Pierre et Paul ont la meme date d'anniversaire.
      iii. deux personnes ont la meme date d'anniversaire.
      iv. toutes les personnes ont des dates d'anniversaires différentes.
5. On considère une population de n individus.
   (a) Combien de groupe de p individus peut-on former à partir de cette population.

(b) On fixe un individu de cette population. Combien y-a-t-il de groupes distincts de p individus contenant l'individu fixé ? Combien y-a-t-il de groupes distincts de p individus ne contenant pas l'individu fixé ?

(c) En déduire la formule :

$$C_{n-1}^{p-1} + C_{n-1}^{p} = C_n^p.$$

6. On effectue 20 fois de suite une experience qui n'a que deux issues possibles : réussite ou échec.
   (a) Combien de suites différentes possibles d'observations des réussites et des échecs sur les 20 essais a-t-on ?
   (b) Combien de suites différentes possibles d'observations comportant 5 réussites (et 15 échecs) a-t-on ?
7. On possède 10 drapeaux. Un signal maritime est fait en hissant le long d'un mat ces 10 drapeaux.
   (a) Si tous les drapeaux sont différents, combien de signaux peut-on faire ?
   (b) Si 4 de ces drapeaux sont bleus et identiques, et les 6 autres rouges et identiques, combiens de signaux peut-on faire ?
8. Une urne contient 49 boules numérotées de 1 à 49, dont 6 boules rouges (numéros gagnants) et une jaune (numéro complémentaire) et les autres blanches. Un échantillon de 6 boules est obtenu par tirage au hasard sans remise, l'ordre n'intervenant pas.
   (a) Combien d'échantillons peut-on obtenir ?
   (b) Combien d'échantillons y-a-t-il contenant 3, 4, 5 ou 6 boules rouges ?
   (c) Combien d'échantillons y-a-t-il contenant 3 boules rouges et la jaune ? 4 boules rouges et la jaune ? 5 boules rouges et la jaune ?

Chapitre 3

Formalisme probabiliste, notion d'indépendance

3.1 Introduction

Il y a dans le monde de nombreuses expériences à venir dont on ne peut prévoir l'issue avec certitude.

1. La pièce que je lance tombera-t-elle sur pile ou sur face ?
2. Jacques Chirac sera-t-il réélu Président de la République ?
3. Combien de temps me faudra-t-il pour obtenir ma licence ?

Le présent cours ne prétend pas répondre à ces questions. Si même ce cours vous permettait de tenter une prédiction sur l'un de ces événements et que celle-ci se révélait inexacte, ce ne saurait cependant remettre en question la pertinence de ce cours. Si vous êtes un téléspectateur assidu des soirées électorales, je gage que vous avez déjà entendu des arguments similaires dans la bouche des spécialistes des instituts de sondage.

Pourquoi un résultat isolé ne saurait remettre en question la théorie ? Parce que la théorie est basée sur le postulat suivant : si une même expérience est répétée un grand nombre de fois, la fréquence d'obtention d'un même résultat tend à se stabiliser vers un nombre appelé probabilité théorique du résultat. Une seule expérience ne peut donc prétendre approcher l'infini. Il faut souligner que ce postulat n'est pas un postulat mathématique – comme les postulats d'Euclide ou l'axiome du choix – mais un postulat physique, c'est à dire sur la nature de l'univers. Évidemment, ce postulat a comme origine la pratique : on s'est aperçu que si on lançait 1000 fois une pièce et que l'on divisait le nombre de faces obtenues par 1000, le résultat obtenu était proche de 1/2 – quoique rarement égal à 1/2.

L'objectif d'une théorie des probabilités est de bâtir des objets mathématiques dont au pourra calculer la probabilité – mot dont le sens mathématique reste encore à définir. Pour que la construction alors faite ait un sens, il est nécessaire dans ce cadre théorique, l'on puisse démontrer que si une même expérience est répétée une infinité fois de manière indépendantes , la fréquence d'obtention d'un même résultat converge (en un sens à préciser) vers la probabilité théorique alors définie. Cette vérification est nécessaire si l'on veut pouvoir relier les nombreuses pages de calcul que nous allons faire à la réalité. Cet objectif majeur est l'une des grandes nouveautés de ce cours par rapport à ce que vous avez pu voir dans les années antérieures.

3.2 Des événements aux espaces probabilisés

La théorie des probabilités décrit les événements comme des sous-ensembles d'un ensemble Ω représentant tous les résultats possibles *a priori* – même s'il peut s'avérer ensuite que certains n'arrivent jamais. Remarquons bien qu'il n'est pas possible de modéliser un phénomène aléatoire quelconque si l'on ne connaît pas les résultats possibles *a priori*.

Soit donc Ω un ensemble. Pour tout $A \subset \Omega$, on note A^c le complémentaire de A dans Ω :

$$A^c = \{x \in \Omega; x \notin A\}.$$

On dit qu'une partie $\mathcal{F} \subset \mathcal{P}(\Omega)$ est une famille d'observables – on dit aussi une tribu – si elle vérifie les propriétés suivantes :

1. $\varnothing \in \mathcal{F}$.
2. $\forall A \in \mathcal{F} \quad A^c \in \mathcal{F}$.
3. Pour toute suite $(A_i)_{i\in\mathbb{N}}$ d'éléments de $\mathcal{F}$, $\bigcup_{i=1}^{+\infty} A_i \in \mathcal{F}$.

Une fois que Ω et $\mathcal{F}$ sont fixés, on appelle événement tout élément de $\mathcal{F}$. Voyons maintenant la définition d'une probabilité sur $(\Omega, \mathcal{F})$.

On appelle probabilité sur $(\Omega, \mathcal{F})$ toute application

$$P : \mathcal{F} \to [0, 1]$$

vérifiant les propriétés suivantes :

1. $P(\Omega) = 1$.
2. Pour toute suite $(A_i)_{i\in\mathbb{N}}$ d'éléments de $\mathcal{F}$ deux à deux disjoints, $P(\bigcup_{i=1}^{+\infty} A_i) = \sum_{i=1}^{+\infty} P(A_i)$.

Alors, le triplet $(\Omega, \mathcal{F}, P)$ est appelé espace probabilisé.

Les propositions suivantes sont alors des conséquences relativement faciles de ces définitions :

- $P(\varnothing) = 0$
- $\forall A \in \mathcal{F} \quad P(A^c) = 1 - P(A)$
- $\forall A, B \in \mathcal{F} \quad A \cap B = \varnothing \Longrightarrow P(A \cup B) = P(A) + P(B)$
- $\forall A, B \in \mathcal{F} \quad P(A \cup B) = P(A) + P(B) - P(A \cap B)$
- $\forall A, B \in \mathcal{F} \quad P(A \cup B) \leq P(A) + P(B)$
- $\forall A, B \in \mathcal{F} \quad A \subset B \Longrightarrow P(A) \leq P(B)$
- $\forall A, B \in \mathcal{F} \quad P(A \cap B) \leq \min(P(A), P(B))$
- $\forall A, B \in \mathcal{F} \quad P(A \cup B) \geq \max(P(A), P(B))$
- Pour toute suite $(A_i)_{1 \leq i \leq n}$ d'éléments de $\mathcal{F}$ deux à deux disjoints, $P(\bigcup_{i=1}^{n} A_i) = \sum_{i=1}^{n} P(A_i)$.
- Pour toute suite $(A_i)_{1 \leq i \leq n}$ d'éléments de $\mathcal{F}$, $P(\bigcup_{i=1}^{n} A_i) \leq \sum_{i=1}^{n} P(A_i)$.

3.3 Partitions et probabilités

Le théorème très simple qui suit est très fréquemment utilisé. Il traduit le fait que pour calculer une probabilité, il faut parfois diviser les cas.

Théorème 2. *Soit $(\Omega, \mathcal{F}, P)$ un espace probabilisé. Soit I un ensemble d'index fini ou dénombrable et $(\Omega_i)_{i \in I}$ une partition de Ω. Alors on a*

$$\forall A \in \mathcal{F} \quad P(A) = \sum_{i \in I} P(A \cap \Omega_i).$$

Démonstration. D'après le théorème 1, la famille $(A \cap \Omega_i)_{i \in I}$ est une partition de A. A est donc réunion disjointe des $(A \cap \Omega_i)_{i \in I}$, donc $P(A) = \sum_{i \in I} P(A \cap \Omega_i)$. $\square$

3.4 Événements indépendants

3.4.1 Indépendance de deux événements

On dit que deux événements observables A et B sont indépendants si on a

$$P(A \cap B) = P(A)P(B).$$

Théorème 3. *Soient A et B deux événements observables.*
A est indépendant de B si et seulement si A est indépendant de B^c.

Démonstration. Calculons de deux manières différentes $P(A)$: on a d'une part

$$P(A) \quad = \quad P(A \cap B) + P(A \cap B^c)$$

D'autre part $1 = P(B) + P(B^c)$ entraîne $P(A) = P(A)P(B) + P(A)P(B^c)$. On a donc

$$P(A \cap B) + P(A \cap B^c) = P(A)P(B) + P(A)P(B^c).$$

Cette égalité entraîne immédiatement que $P(A \cap B) = P(A)P(B)$ est équivalent à $P(A \cap B^c) = P(A)P(B^c)$. □

Remarque : Les événement de probabilité 0 ou de probabilité 1 sont indépendants de n'importe quels événements – même d'eux-mêmes.

Démonstration. Soit B un événement de probabilité nulle : soit A un événement observable quelconque : on a $P(A \cap B) \leq P(B) = 0$ donc $P(A \cap B) = 0 = P(A)P(B)$. B est donc indépendant de n'importe quel événement observable.

Si B est de probabilité 1, B^c est de probabilité 0 et est donc indépendant de n'importe quel événement, d'après la première partie de notre raisonnement. . Le théorème 3 permet alors de conclure. □

3.4.2 Indépendance d'une famille d'événements

Soit $(A_i)_{i \in G}$ une partie d'éléments de $\mathcal{F}$ indexés par un ensemble G. On dit que les événements constituant la famille $(A_i)_{i \in G}$ sont globalement indépendants si l'on a pour tout ensemble fini $I \subset G$:

$$P(\bigcap_{i \in I} A_i) = \prod_{i \in I} P(A_i).$$

Remarque : si un ou plusieurs des événements est remplacé par son complémentaire, cela ne change rien à l'indépendance de la famille. La preuve est laissée au lecteur ; il faut utiliser le théorème 3 et effectuer une récurrence sur $|I|$.

3.5 Exercices

1. *Opérations sur les évènements*
   Trois boules sont tirées successivement d'une urne contenant des boules blanches et des boules rouges. On définit les événements :

$$\begin{aligned} A &= \{\text{la première boule est blanche}\}, \\ B &= \{\text{la deuxième boule est blanche}\}, \\ C &= \{\text{la troisième boule est blanche}\}. \end{aligned}$$

   Exprimer à l'aide des événements A, B et C les événements suivants :

$$\begin{aligned} D &= \{\text{toutes les boules tirées sont blanches}\}, \\ E &= \{\text{les deux premières sont blanches}\}, \\ F &= \{\text{au moins une boule est blanche}\}, \\ G &= \{\text{seule la troisième est blanche}\}, \\ H &= \{\text{une seule boule est blanche}\}. \end{aligned}$$

2. *Dur métier.*
   Une princesse est enfermée dans un château. Un prince est envoyé pour la délivrer. Dans la cour du château se trouvent plusieurs portes. L'une mène chez la princesse, une autre dans l'antre d'un dragon. Toutes les autres cachent des sorcières. Le prince choisit une porte au hasard. S'il trouve la princesse il la délivre[1], s'il trouve le dragon il se fait dévorer. S'il tombe sur une sorcière, elle le reconduit dans la cour après lui avoir fait boire un philtre qui lui fait oublier la porte qu'il a choisie. Le prince recommence ses tentatives jusqu'à ce qu'il délivre la princesse ou soit dévoré.
   On définit pour $k \in N^*$ les événements :

$$\begin{aligned} D_k &= \{\text{Le prince est dévoré à la } k\text{-ième tentative}\}, \\ F_k &= \{\text{Le prince délivre la princesse à la } k\text{-ième tentative}\}, \\ S_k &= \{\text{Le prince trouve une sorcière à la } k\text{-ième tentative}\}. \end{aligned}$$

   (a) En utilisant les D_k, F_k et S_k, exprimer les événements :

$$\begin{aligned} E &= \{\text{Le prince échoue, il est dévoré}\}, \\ R &= \{\text{Le prince réussit à délivrer la princesse}\}, \\ I &= \{\text{Le prince recommence indéfiniment ses tentatives}\}. \end{aligned}$$

[1] Et ils seront heureux et contribueront fortement à la démographie galopante...

(b) Donner une représentation de l'espace Ω des événements élémentaires de cette *expérience.*

(c) Que valent $\bigcap_{i=1}^{k-1} S_i \cap D_k$ et $E \cup R \cup I$?

(d) Exprimer de deux manières différentes l'événement :

$$M_k = \{\text{Le prince effectue au moins } k \text{ tentatives}\}.$$

3. *Paradoxe du chevalier de Méré*
   Est-il plus probable d'avoir un six en lançant un dé 4 fois de suite ou un double six en lançant deux dés 24 fois de suite ?
4. On lance 4 dés. Quelle est la probabilité d'obtenir :
   (a) un carré (aaaa) ;
   (b) un brelan (aaab) ;
   (c) une double paire (aabb) ;
   (d) une paire (aabc) ;
   (e) une disposition simple (abcd).
5. Le jeu de Keno consiste à cocher de 4 à 10 numéros par grille de 70 numéros. A chaque tirage, il y a 20 numéros gagnants.
   (a) Si on a coché 4 numéros, quelle est la probabilité qu'ils soient tous gagnants ? [57/10787]
   (b) Si on a coché 4 numéros, quelle est la probabilité que 3 soient gagnants ? [approximativement 1/16]
   (c) Si on a coché 10 numéros, quelle est la probabilité que 7 soientgagnants ? [approximativement 1/261]
6. On effectue de façon indépendante trois fois une même expérience et on s'intéresse au nombre de succès.
   (a) Donner l'espace Ω des résultats de ce test et son cardinal.
   (b) On suppose que la probabilité de réussite p est la même à chaqueessai
      i. Que signifie mathématiquement ¡¡ independance des trois essais ¿¿ ?
      ii. Donner la probabilité de tous les événements de Ω.
      iii. Pour $0 \leq i \leq 3$, soit A_i l'événement ¡¡ i reussite au bout des trois essais ¿¿. Déterminer les $P(A_i)$, et donner leur valeur pour $p = 0,3$.

(c) On appelle p_1, p_2 et p_3 les probabilités de reussite au premier, deuxième et troisième essai. Recalculer les $P(A_i)$, et donner leur valeur pour $p_1 = 0,75$, $p_2 = 0,5$ et $p_3 = 0,25$.

7. Au sein d'une population de 1000 individus, on en a dénombré 120 atteints d'une maladie M, les 880 autres étant indemnes de cette maladie.
   (a) Quelle est la probabilité pour qu'un individu tiré au hasard au sein de cette population soit atteint de la maladie M ?
   (b) On observe un petit échantillon de 5 sujets tirés au hasard au sein de cette population. Quelle est la probabilité que les 5 sujets soient indemnes de la madadie M ?
8. Dans une maternité, on constate que sur l'ensemble des accouchements, 20% présentent des complications et 10% ont lieu avant le terme normal (40 semaines).
   (a) Si le terme est indépendant de l'existence de complications, quelle est la probabilité pour qu'une femme ait un accouchement normal à terme ?
   (b) En fait, il y a 40% de complications quand l'accouchement à lieu avant terme. Dans ces conditions, quelle est la probabilité :
      i. d'un accouchement avant terme et avec complications ?
      ii. d'un accouchement normal à terme.

Chapitre 4

Variables aléatoires

4.1 Généralités

4.1.1 Définition

Soit $(\Omega, \mathcal{F}, P)$ un espace probabilisé. On dit qu'une application X définie sur Ω est à valeurs dans $\mathbb{R}$ est une variable aléatoire si elle vérifie la propriété suivante :

$$\forall t \in \mathbb{R} \quad \{\omega \in \Omega; X(\omega) \leq t\} \in \mathcal{F}.$$

Dès lors, le nombre

$$P(\{\omega \in \Omega; X(\omega) \leq t\})$$

est bien défini. Pour $x \in \mathbb{R}$, on pose

$$F_X(t) = P(\{\omega \in \Omega; X(\omega) \leq t\}).$$

La fonction $t \mapsto F_X(t)$ est appelée fonction de répartition. Nous allons y revenir bientôt.

Comme les écritures du type $P(\{\omega \in \Omega; X(\omega) \leq t\})$ sont assez lourdes, on choisit de les alléger en $P(X \leq t)$. Ainsi, dans la pratique courante et à notre niveau, l'existence d'Ω peut parfois être oubliée.

4.1.2 Loi d'une variable aléatoire

Théorème 4. *Soit X une variable aléatoire définie sur un espace probabilisé $(\Omega, \mathcal{F}, P)$ L'application*

$$\begin{aligned} P_X : \mathcal{B}(\mathbb{R}) &\to [0,1] \\ A &\mapsto P(X \in A) \end{aligned}$$

définit une probabilité sur $(\mathbb{R}, \mathcal{B}(\mathbb{R}))$.

On dit que la probabilité P_X est la loi de la variable aléatoire X sous la probabilité P, ou, plus briévement, que P_X est la loi de X.

4.1.3 Fonction de répartition

On remarque que la fonction de répartition de X est liée à la loi de X par la formule
$$F_X(t) = P_X(]-\infty, t]).$$
Ainsi, lorsque l'on connaît la loi d'une variable aléatoire, on peut retrouver la fonction de répartition. Réciproquement, on admettra que la fonction de répartition détermine de manière unique la loi de la variable aléatoire.

Nous allons maintenant énoncer les propriétés caractéristiques d'une fonction de répartition.

Théorème 5. *La fonction de répartition F_X d'une variable aléatoire vérifie les propriétés suivantes*
- F_X *est à valeurs dans* $[0, 1]$
- F_X *est croissante sur* $\mathbb{R}$.
- $\lim_{t \to -\infty} F_X(t) = 0$.
- $\lim_{t \to +\infty} F_X(t) = 1$.
- *En tout point,* F_X *est continue à droite.*
- *En tout point,* F_X *admet une limite à gauche.*

4.1.4 Fonction d'une variable aléatoire

Il est très simple de fabriquer, à partir d'une variable aléatoire, de nouvelles variables aléatoires. On admettra le théorème suivant :

Théorème 6. *Soit X une variable aléatoire définie sur un espace probabilisé $(\Omega, \mathcal{F}, P)$ et f une fonction continue par morceaux définie sur $X(\Omega)$. Alors, la fonction Y définie par*
$$\forall \omega \in \Omega \quad Y = f(X(\omega))$$
est une variable aléatoire sur $(\Omega, \mathcal{F}, P)$.

4.1.5 Variables aléatoires indépendantes

On dit que des variables aléatoires $X_1, \dots, X_n$ sont indépendantes si pour tout n-uplet $(A_1, \dots, A_n) \in \mathcal{B}(\mathbb{R})^n$, les événements

$$\{X_1 \in A_1\}, \dots, \{X_n \in A_n\}$$

sont indépendants.

Théorème 7. *Des variables aléatoires $X_1, \dots, X_n$ sont indépendantes si et seulement si pour tout n-uplet $(A_1, \dots, A_n) \in \mathcal{B}(\mathbb{R})^n$, on a*

$$P(\bigcap_{i=1}^{n} A_i) = \prod_{i=1}^{n} P(A_i).$$

Démonstration. La preuve est laissée en exercice au lecteur. Une des implications est évidente, l'autre mérite un peu de réflexion. □

4.1.6 Fonctions de variables aléatoires indépendantes

Théorème 8. *Soient $X_1, \dots, X_n$ des variables aléatoires indépendantes et $f_1, \dots, f_n$ des fonctions continues par morceaux respectivement définies sur $X_1(\Omega), \dots, X_n(\Omega)$. Alors les variables aléatoires $f_1(X_1), \dots, f_n(X_n)$ sont indépendantes*

La preuve est laissée en exercice.

Exemple : : si X, Y, Z sont des variables aléatoires indépendantes, alors $\sin X$, $\cos Y$ et Z^3 sont des variables aléatoires indépendantes.

Théorème 9. *Soient $X_1, \dots, X_{n+p}$ des variables aléatoires indépendantes et f, g des fonctions continues par morceaux respectivement définies sur $X_1(\Omega) \times \dots \times X_n(\Omega)$ et $X_{n+1}(\Omega) \times \dots \times X_{n+p}(\Omega)$. Alors les variables aléatoires $f(X_1, \dots, X_n)$ et $g(X_{n+1}, X_{n+p})$ sont indépendantes*

Exemple : : si X, Y, Z, T sont des variables aléatoires indépendantes, alors $\sin(XY)$ et $T^3 + \cos Z$ sont des variables aléatoires indépendantes.

4.2 Variables aléatoires discrètes

On dit que X est une variable aléatoire discrète s'il existe un ensemble D fini ou dénombrable inclus dans $\mathbb{R}$ tel que $X(\Omega) = D$.

Alors, si l'on pose

$$\forall i \in D \quad p_i = P(X = i)$$

La famille $(p_i)_{i\in D}$ est une famille de réels positifs vérifiant

$$\sum_{i\in D} p_i = 1.$$

La connaissance de D est des p_i permet de reconstituer la loi de X. En effet, on a le théorème suivant :

Théorème 10. *Soit X est une variable aléatoire discrète, et D un ensemble D fini ou dénombrable inclus dans $\mathbb{R}$ tel que $X(\Omega) = D$. Alors, pour tout $A \in \mathcal{B}(\mathbb{R})$, on a*

$$P_X(A) = \sum_{i\in D\cap A} p_i,$$

où l'on a posé $p_i = P(X = i)$.

Démonstration. Si l'on pose pour $i \in D$

$$\Omega_i = \{\omega \in \Omega; X(\omega) = i\},$$

il est facile de voir que les Ω_i forment une partition de Ω. On a donc

$$\{\omega \in \Omega; X(\omega) \in A\} = \underset{i\in D}{\cup} \Omega_i \cap \{\omega \in \Omega; X(\omega) \in A\})$$

Mais si $i \notin A$, on a $\Omega_i \cap \{\omega \in \Omega; X(\omega) \in A\}) = \varnothing$. Inversement, si $i \in A$, on a $\Omega_i \cap \{\omega \in \Omega; X(\omega) \in A\}) = \Omega_i$ d'où

$$\{\omega \in \Omega; X(\omega) \in A\} = \underset{i\in D;i\in A}{\cup} \Omega_i,$$

puis

$$P_X(A) = P(\{\omega \in \Omega; X(\omega) \in A\}) = \sum_{i\in D;i\in A} P(\Omega_i) = \sum_{i\in D\cap A} p_i.$$

□

En particulier, on a donc

$$F_X(t) = \sum_{i\in D;i\leq t} p_i.$$

On a la réciproque suivante :

Théorème 11. *Soit D un ensemble fini ou dénombrable, $(p_i)_{i\in D}$ une famille de réels positifs vérifiant*

$$\sum_{i\in D} p_i = 1.$$

Alors, on peut construire un espace probabilisé $(\Omega, \mathcal{F}, P)$ est une variable aléatoire X sur cet espace telle que

$$\forall i \in D \quad p_i = P(X = i)$$

Démonstration. On peut prendre $\Omega = D$, $\mathcal{F} = \mathcal{P}(\Omega)$ et $X(\omega) = \omega$ avec

$$P(A) = \sum_{i\in D\cap A} p_i.$$

(La fin de la preuve est laissée au lecteur). □

4.2.1 Fonction d'une variable aléatoire discrète

Théorème 12. *Soit X une variable aléatoire discrète définie sur un espace probabilisé $(\Omega, \mathcal{F}, P)$ et f une fonction quelconque définie sur $X(\Omega)$. Alors, la fonction Y définie par*

$$\forall \omega \in \Omega \quad Y = f(X(\omega))$$

est une variable aléatoire discrète sur $(\Omega, \mathcal{F}, P)$.

De manière plus concise, on écrira $Y = f(X)$.

Exemple : Soit X une variable aléatoire vérifiant $X(\Omega) = \{-1; 0; 1\}$, avec $P(X = -1) = P(X = 0) = P(X = 1) = \frac{1}{3}$.

On pose $Y = X^2$.

On a $Y(\Omega) = \{0; 1\}$, avec

$$\{Y = 0\} = \{X = 0\}$$

et

$$\{Y = 1\} = \{X = 1\} \cup \{X = -1\},$$

d'où

$$P(Y = 0) = P(X = 0) = \frac{1}{3}$$

et

$$P(Y = 1) = P(X = 1) + P(X = -1) = \frac{1}{3} + \frac{1}{3} = \frac{2}{3}.$$

4.2.2 Variables et lois discrètes classiques

4.2.3 Indicatrice d'un évévenement

On rappelle que pour $A \subset \Omega$, l'application $\mathbb{1}_A$ (appelée indicatrice de A) est définie sur Ω par

$$\mathbb{1}_A(x) = \begin{cases} 1 \text{ si } x \in A \\ 0 \text{ si } x \notin A \end{cases}$$

1_A est une variable aléatoire à valeurs dans $\{0; 1\}$.

4.2.4 Loi uniforme sur un ensemble

Soit E un ensemble fini. On dit qu'une variable aléatoire X suit la loi uniforme sur E si l'on a

$$\forall x \in E \quad P(X = x) = \frac{1}{|E|}.$$

Exemple : La variable aléatoire X représentant le résultat du lancer d'un dé non truqué suit la loi uniforme sur l'ensemble $\{1, 2, 3, 4, 5, 6\}$.

4.2.5 Loi binomiale

On dit qu'une variable aléatoire X suit la loi binomiale de paramètres n et p si l'on a

$$\forall i \in \{0, \dots, n\} \quad P(X = k) = \binom{n}{k} p^k (1-p)^{n-k}.$$

Théorème 13. *Soit $A_1, \dots, A_n$ n événements indépendants de même probabilité p. On pose*

$$X = \sum_{k=1}^{n} \mathbb{1}_{A_k}.$$

Alors X suit la loi binomiale de paramètres n et p .

Démonstration. L'événement $\{X = k\}$ se partitionne ainsi :

$$\{X = k\} \quad = \quad \bigcup_{B \in \mathcal{B}_k(\{1,\dots,n\})} \Big(\bigcap_{i \in B} A_i \Big) \cap \Big(\bigcap_{i \in \{1,\dots,n\} \setminus B} A_i^c \Big)$$

Or, pour tout $B \in \mathcal{B}_k(\{1, \dots, n\})$, on a

$$\begin{aligned} P\Big(\Big(\bigcap_{i \in B} A_i \Big) \cap \Big(\bigcap_{i \in \{1,\dots,n\} \setminus B} A_i^c \Big)\Big) &= \Big(\prod_{i \in B} P(A_i) \Big) . \Big(\prod_{i \in \{1,\dots,n\} \setminus B} P(A_i^c) \Big) \\ &= p^k (1-p)^{n-k} \end{aligned}$$

(On utilise ici l'indépendance des (A_i).) On en déduit

$$P(X=k) = \sum_{B \in \mathcal{B}_k(\{1,\dots,n\})} p^k(1-p)^{n-k} = |\mathcal{B}_k(\{1,\dots,n\})|p^k(1-p)^{n-k} = \binom{n}{k} p^k(1-p)^{n-k}.$$

$\square$

Remarque : X est le nombre de A_k qui sont réalisés.

Exemple : On lance n fois une pièce de monnaie équilibrée. Le nombre X de "pile" obtenus suit une loi binomiale de paramètres n et $\frac{1}{2}$.

4.2.6 Loi géométrique

On dit qu'une variable aléatoire X suit la loi géométrique de paramètre p si l'on a

$$\forall k \in \mathbb{N}^* \quad P(X=k) = p(1-p)^{n-1}.$$

Théorème 14. *Soit $(A_i)_{i \in \mathbb{N}^*}$ une suite d'événements indépendants de même probabilité p. On pose*

$$X(\omega) = \inf\{k \in \mathbb{N}^* \quad \omega \in A_k\}.$$

Alors X suit la loi géométrique de paramètre p . De plus

$$\forall k \in \mathbb{N} \qquad F_X(k) = 1-(1-p)^k.$$

Démonstration.

$$\{X > k\} = \bigcap_{i \in \{1,\dots,k\}} A_i^c.$$

Comme les A_i sont dans $\mathcal{F}$, les A_i^c le sont aussi, et donc comme on peut l'écrire comme intersection finie d'éléments de $\mathcal{F}$,$\{X > k\}$ est dans $\mathcal{F}$, et donc, par passage au complémentaire

$$\{X \le k\} \in \mathcal{F}.$$

En utilisant l'indépendance des (A_i), on obtient

$$P(X > k) = (1-p)^k.$$

Comme

$$\{X = +\infty\} = \bigcap_{k \ge 1} \{X > k\},$$

on obtient, par continuité séquentielle décroissante

$$P(X=+\infty) = \lim_{k \to +\infty} P(\{X > k\}) = \lim_{k \to +\infty} (1-p)^k = 0.$$

Ainsi X est bien une variable aléatoire, et l'on a

$$\forall k \in \mathbb{N}^* \quad P(X = k) = P(X > k-1) - P(X > k) = (1-p)^{k-1} - (1-p)^k = (1-p)^{k-1}p$$

De plus

$$\forall k \in \mathbb{N} \quad F_X(k) = P(X \leq k) = 1 - P(X > k) = 1 - (1-p)^k.$$

□

Exemple : On lance une pièce de monnaie équilibrée jusqu'à obtention de "pile". Le nombre de lancers effectués suit une loi géométrique de paramètre $\frac{1}{2}$.

4.2.7 Loi de Poisson

On dit qu'une variable aléatoire X suit la loi géométrique de paramètre λ si l'on a

$$\forall k \in \mathbb{N} \quad P(X = k) = e^{-\lambda}\frac{\lambda^k}{k!}.$$

La construction de telles variables est bien possible car $e^{-\lambda}\frac{\lambda^k}{k!} \geq 0$ et

$$\sum_{k=0}^{+\infty} e^{-\lambda}\frac{\lambda^k}{k!} = e^{-\lambda}e^{\lambda} = 1.$$

La loi de Poisson est aussi appelée loi des événements rares.

Pa exemple, si on sait que dans une région, il y a en moyenne 10 cas d'une certaine maladie rare pendant une période donnée, alors le nombre de cas observés suit une loi de Poisson de paramètre 10.

4.3 Convolution de variables aléatoires discrètes

Le but de cette section est de déterminer la loi de la somme de deux variables aléatoires discrètes indépendantes.

Théorème 15. *Soient X et Y deux variables aléatoires indépendantes sur $(\Omega, \mathcal{F}, P)$ à valeurs dans $\mathbb{Z}$ On pose $Z = X + Y$. Alors, Z est une variable aléatoire discrète vérifiant*

$$\forall n \in \mathbb{Z} \quad P(Z = n) = \sum_{i \in \mathbb{Z}} P(X = i)P(Y = n - i).$$

Si X et Y sont en fait à valeurs dans $\mathbb{N}$, la formule ci-dessous se simplifie en

$$\forall n \in \mathbb{N} \quad P(Z=n) = \sum_{i=0}^{n} P(X=i)P(Y=n-i).$$

Démonstration. Soit $n \in \mathbb{Z}$.

$$\begin{aligned} P(Z=n) &= \sum_{i\in\mathbb{Z}} P(\{Z=n\}\cap\{X=i\}) &&= \sum_{i\in\mathbb{Z}} P(\{X+Y=n\}\cap\{X=i\}) \\ &= \sum_{i\in\mathbb{Z}} P(\{Y+i=n\}\cap\{X=i\}) &&= \sum_{i\in\mathbb{Z}} P(\{Y=n-i\}\cap\{X=i\}) \\ &= \sum_{i\in\mathbb{Z}} P(X=i)P(Y=n-i) \end{aligned}$$

Si X est à valeurs dans $\mathbb{N}$, $P(X=i)=0$ pour $i<0$. De même, si Y est à valeurs dans $\mathbb{N}$, $P(Y=n-i)=0$ pour $i>n$. On en déduit les simplifications annoncées. □

4.4 Exercices

1. On sait que dans une région, il y a en moyenne 10 cas d'une certaine maladie pendant une période donnée. On suppose que la probabilité qu'une personne ait cette maladie est très faible .
   (a) Proposer, en la justifiant une modélisation pour la loi du nombre de maladies observées.
   (b) En accord avec cette modélisation , calculer la probabilité qu'il y ait strictement moins de 3 malades une année donnée.
2. Un liquide contient 10^5 bactéries par litre réparties au hasard. On en prélève 1 mm^3.
   (a) Quelle est la probabilité que ce prélèvement ne contienne aucune bactérie ?
   (b) Quelle est la probabilité que ce prélèvement ne contienne au moins 3 bactéries ?
3. Soient $A, B \subset \Omega$. On note $A\Delta B = \{x \in A; x \notin B\} \cup \{x \in B; x \notin A\}$. Montrer $\mathbb{1}_{A\Delta B} = (\mathbb{1}_A - \mathbb{1}_B)^2$.
4. On lance deux fois un dé à 6 faces non truqué. On note X le résultat du premier lancer et Y le résultat du second lancer.
   (a) Pour $k \in \{0, \dots, n\}$, calculer $P(X \leq k)$.
   (b) On pose $Z = \max(X, Y)$. Pour $k \in \{0, \dots, n\}$, calculer $P(Z \leq k)$.
   (c) $k \in \{1, \dots, n\}$, calculer $P(Z = k)$.
5. Soit X, Y deux variables aléatoires indépendantes suivant une loi de Poisson de paramètre λ. Montrer que $\{X = 0\} \cap \{Y = 0\} \subset \{X = Y\}$. En utilisant le fait que $\forall x \geq 0, e^{-x} \geq 1 - x$, en déduire
$$P(X \neq Y) \leq 2\lambda.$$
6. Soient $(A_1, \dots, A_{2n})$ $2n$ événements globalement indépendants de même probabilité $p \in]0, 1[$. On pose $X = \sum\limits_{k=1}^{n} \mathbb{1}_{A_k}$, $Y = \sum\limits_{k=n+1}^{2n} \mathbb{1}_{A_k}$ et $Z = X + Y$.
   Calculer de deux manières différentes $P(Z = n)$. En déduire l'identité
$$\sum_{k=0}^{n} \binom{n}{k}^2 = \binom{2n}{n}.$$
7. Alain et Bernard jettent chacun n fois une pièce de monnaie équilibrée. On note X le nombre de "pile" obtenus par Alain, et Y le nombre de "pile" obtenus par Bernard.

(a) Loi de X ? Loi de Y ?

(b) Calculer $P(X = Y)$.
Indication : On rappelle l'identité

$$\sum_{k=0}^{n} \binom{n}{k}^2 = \binom{2n}{n}.$$

8. Soit X, Y deux variables aléatoires indépendantes suivant des lois de Poisson de paramètres respectifs α et β. Montrer que $X + Y$ suit une loi de Poisson de paramètre $\alpha + \beta$.

9. Sur une chaîne de fabrication de plaquettes de comprimés d'un médicament X, il y a une proportion D de plaquettes défectueuses. Un contrôle de qualité de fabrication conduit à prélever un échantillon de n plaquettes. On observe R plaquettes défectueuses

   (a) Exprimer la probabilité que $R = 0$ en fonction de D et n.

   (b) Si $D = 10\%$, quelles valeurs faut il donner à n pour que $P(R = 0) < 0.05$?

   (c) Si $D = 5\%$, combien faut-il prélever de plaquettes pour que l'on ait au moins 90% de chances d'en avoir au moins une défectueuse ?

   (d) Si $n = 50$, pour quelles valeurs de D a-t-on $P(R = 0) < 0.01$? Si $n = 10$, quelles sont les valeurs de D qui permettent d'observer au moins une plaquette défectueuse dans au moins 95% ? des échantillons contrôlés ?

   (e) On a controlé 100 plaquettes dont $R = 2$ ont été trouvées défectueuses. Des 3 conjectures $D = 1\%$, $D = 3\%$ et $D = 5\%$, quelle est celle qui vous paraît la plus vraisemblable ?

Chapitre 5

Probabilités conditionnelles

Dans la vie scientifique courante, il est souvent nécessaire de savoir si des événements sont liés. Quelques exemples :

- Être contrôlé à plus de 110 km/h sur la route Pérenchies-Vervicq et avoir réellement commis un excès de vitesse.
- Avoir la moyenne au partiel et réussir l'examen.
- Se déclarer, lors d'un sondage, hostile à l'immigration et voter pour un parti fasciste.

Il faut bien remarquer que les liens entre ces événements ne sont pas symétriques. Par exemple, il est fort probable que quelqu'un contrôlé à 110km/h ait effectivement oublié son pied sur l'accélérateur, alors que – à l'heure où ses lignes sont écrites ! – il est peu probable que quelqu'un dépassant la limitation de vitesse sur la route Pérenchies-Vervicq se fasse prendre, vu le faible nombre de contrôles.

5.1 Définition

Soit $(\Omega, \mathcal{F}, P)$ un espace de probabilité et B un événement observable de probabilité non nulle. On appelle probabilité conditionnelle sachant B l'application

$$\begin{aligned} P(.|B) : \mathcal{F} &\to \mathbb{R} \\ A &\mapsto P(A|B) = \frac{P(A \cap B)}{P(B)} \end{aligned}$$

$P(A|B)$ se lit "Probabilité de A sachant B".

On a évidemment

$$P(A \cap B) = P(B)P(A|B). \tag{5.1}$$

Remarques :

- L'application "probabilité conditionnelle" est une probabilité. Elle vérifie donc toutes les propriétés énoncées au chapitre 2.
- Si $P(B) \neq 0$, alors A est indépendant de B si et seulement si $P(A|B) = P(A)$.

5.2 Conditionnements en chaîne

Si A, B sont deux événements observables avec $A \subset B$ et $P(B) \neq 0$, la formule (5.1) devient

$$P(A) = P(B)P(A|B). \tag{5.2}$$

On a la généralisation suivante :

Théorème 16. *Soient $n \geq 2$ et $E_1, \dots, E_n$ des événements observables vérifiant*

$$E_n \subset E_{n-1} \subset \cdots \subset E_1$$

et $P(E_{n-1}) > 0$. Alors on a

$$P(E_n) = P(E_n|E_{n-1})P(E_{n-1}|E_{n-2}) \dots P(E_2|E_1)P(E_1)$$

Démonstration. La formule se montre par récurrence sur n. Pour $n = 2$, c'est une conséquence immédiate de (5.2). Pour $n > 2$, on applique d'abord la formule pour $n = 2$ aux événements E_n et E_{n-1} :

$$P(E_n) = P(E_n|E_{n-1})P(E_{n-1}),$$

puis on applique la propriété de récurrence au rang $n - 1$. □

Exemple : (d'après André Franquin) Chez les papous, il y a les papous à poux et les papous pas à poux. La probabilité pour qu'un papou ait des poux vaut 0.1. De plus, chez les papous, il y a les papous papas et les papous pas papas. La probabilité pour qu'un papou à poux soit papa vaut 0.6. Or, chez les poux, il y a les poux papas et les poux pas papas : la probabilité pour qu'un papou à poux possède au moins un pou papa est de 0.8.

Question : on tire au hasard un papou. Quelle est la probabilité pour que ce soit un papa papou à poux papa ? Réponse : $0.8 \times 0.6 \times 0.1 = 0.048$.

Ce théorème est parfois énoncé sous la forme plus compliquée – mais équivalente – suivante.

Théorème 17. *Soient $n \geq 2$ et $A_1, \dots, A_n$ des événements observables avec $P(A_1 \cap A_2 \cap \dots \cap A_{n-1}) > 0$. Alors*

$$P(A_1 \cap A_2 \cap \dots \cap A_n) = (\prod_{k=1}^{n-1} P(A_1 \cap A_2 \cap \dots \cap A_{k+1} | A_1 \cap A_2 \cap \dots \cap A_k))P(A_1)$$

Démonstration. Il suffit de poser, pour $1 \leq k \leq n$, $E_k = \bigcap\limits_{k=1}^{n} A_k$ et d'appliquer le théorème précédent. □

5.3 Conditionnement par tous les cas possibles

Ceci est une reformulation à l'aide des probabilités conditionnelles du théorème 2.

Théorème 18. *Soit $(\Omega, \mathcal{F}, P)$ un espace probabilisé. Soit I un ensemble d'index fini ou dénombrable et $(\Omega_i)_{i \in I}$ une partition de Ω. Alors on a*

$$\forall A \in \mathcal{F} \quad P(A) = \sum_{i \in J} P(A|\Omega_i)P(\Omega_i),$$

où $J = \{i \in I; P(\Omega_i) > 0\}$.

Démonstration. D'après le théorème 2, on a

$$\begin{aligned} P(A) &= \sum_{i \in I} P(A \cap \Omega_i) \\ &= \sum_{i \in J} P(A \cap \Omega_i) + \sum_{i \in I \setminus J} P(A \cap \Omega_i) \\ &= \sum_{i \in J} P(A \cap \Omega_i) \\ &= \sum_{i \in J} P(A|\Omega_i)P(\Omega_i) \end{aligned}$$

□

5.4 Formule de Bayes

Théorème 19. *Soit $(\Omega, \mathcal{F}, P)$ un espace probabilisé. Soit I un ensemble d'index fini ou dénombrable et $(\Omega_i)_{i \in I}$ une partition de Ω telle que pour tout $i \in I$, $P(\Omega_i)$ soit non nul. Soit A un élément de probabilité non nulle.*

Alors on a, pour tout $j \in I$, la formule

$$P(\Omega_j|A) = \frac{P(A|\Omega_j)P(\Omega_j)}{\sum_{i \in I} P(A|\Omega_i)P(\Omega_i)}.$$

Démonstration.

$$\begin{aligned} P(\Omega_j|A) &= \frac{P(A \cap \Omega_j)}{P(A)} \\ &= \frac{P(A|\Omega_j)P(\Omega_j)}{P(A)} \end{aligned}$$

et on applique le théorème précédent.

□

Exemple :

- 60% des étudiants qui vont en T.D. obtiennent l'examen.
- 10% des étudiants qui ne vont pas en T.D. obtiennent l'examen.
- 70% des étudiants vont en T.D.

Quelle proportion des lauréats a séché les cours ? On note A l'événement "être assidu en cours". On a $P(A) = 0.7$, et donc $P(A^c) = 0.3$. On note L l'événement "obtenir l'examen" : on a $P(L|A^c) = 0.1$ et $P(L|A) = 0.6$. On a alors

$$P(A^c|L) = \frac{P(L|A^c)P(A^c)}{P(L|A^c)P(A^c) + P(L|A)P(A)} = \frac{0.1 \times 0.3}{0.1 \times 0.3 + 0.6 \times 0.7} = \frac{3}{45} = \frac{1}{15}.$$

5.5 Exercices

1. Alain et Bernard jouent à pile ou face. Chacun des deux possède une pièce de monnaie. Ils lancent successivement et simultanément leurs pièces de monnaie jusqu'à ce que Alain obtienne pile ou/et que Bernard obtienne face. Si les deux réussissent en même temps, la partie est nulle, sinon c'est le premier qui a réussi qui gagne. On cherche a calculer la probabilité pour que la partie soit nulle.
   (a) Calculer cette probabilité lorsque les deux pièces sont identiques et honnêtes (pile et face sont équiprobables).
   (b) Calculer cette probabilité lorsque les deux pièces sont identiques et malhonnêtes (pour un lancer de chaque pièce, la probabilité de pile est deux). Montrer quelle est majorée par $\frac{1}{3}$.
2. Soient P, C tels que $P(C) \neq 0$. Montrer $P(B|B \cup C) \geq P(B)$, puis $P(B|B \cup C) \geq P(B|C)$.
3. Un avion a disparu et la région où il s'est écrasé est divisée pour sa recherche en trois zones de même probabilité. Pour $i = 1, 2, 3$, notons $1 - \alpha_i$ la probabilité que l'avion soit retrouvé par une recherche dans la zone i s'il est effectivement dans cette zone. Les constantes α_i représentent les probabilités de manquer l'avion et sont généralement attribuables à l'environnement de la zone (relief, végétation,...). On notera A_i l'événement *l'avion est dans la zone i*, et R_i l'événement *l'avion est retrouvé dans la zone i* ($i = 1, 2, 3$).
   (a) Pour $i = 1, 2, 3$, déterminer les probabilités que l'avion soit dans la zone i sachant que la recherche dans la zone 1 a été infructueuse.
   (b) Etudier brièvement les variations de ces trois probabilités conditionnelles considérées comme fonctions de α_1 et commenter les résultats obtenus.
4. Des étudiants se préparent à un examen. Le sujet sera fabriqué par l'un de leurs trois professeurs : Xxxx, Yyyy ou Zzzz.
   Or, les étudiants redoutent qu'un certain chapitre soit posé à l'examen et ils évaluent à 10% la probabilité pour que le chapitre en question sorte si c'est Xxxx qui fait le sujet, 40% si c'est Yyyy qui le fait, et 80% si c'est Zzzz (Inutile de dire à quel prof vont les sympathies des étudiants...).
   Yyyy leur a dit : Il y a une chance sur deux pour que ce soit moi qui fasse le sujet, et si je ne le fais pas il y a trois chances sur cinq pour que ce soit Xxxx.
   Le jour J arrive, et le chapitre fatidique est posé à l'examen ! Sachant

cela, calculer les probabilités pour que l'examen ait été posé par Xxxx, Yyyy ou Zzzz.

5. Dans un élevage de moutons, on décèle 15% d'animaux malades. La probabilité qu'un mouton qui n'est pas malade ait une réaction négative à un test donné est 0.90. Par contre, s'il est malade, la réaction sera positive avec une probabilité 0.80. Quelle est la probabilité qu'un mouton choisi au hasard et ayant une réaction positive au test soit malade ?
6. **La propriété d'absence de mémoire**
   (a) Montrer que si X est une v. a. de loi géométrique, elle vérifie la propriété d'absence de mémoire suivante :
   $$\forall k \in \mathbb{N},\ \forall n \in \mathbb{N}, \qquad P(X > n + k \mid X > n) = P(X > k). \quad (5.3)$$
   Interpréter ce résultat en considérant une suite d'épreuves répétées.
   (b) Trouver toutes les lois qui vérifient la propriété (5.3).
   Indication : On notera $G(n) = P(X > n)$ et on montrera que (5.3) se traduit par une relation simple entre $G(n + k)$, $G(n)$ et $G(k)$.
7. **Mélange de lois**
   On supose que le nombre N d'oeufs pondus par un insecte suit une loi de Poisson de paramètre α :
   $$P(N = k) = \frac{\exp(-\alpha)\alpha^k}{k\,!} \qquad k \in \mathbb{N}$$
   On suppose également que la probabilité de développement d'un oeuf est p et que les oeufs sont mutuellement indépendants. On note S le nombre (aléatoire) de survivants. Montrer que S suit une loi de Poisson de paramètre $p\alpha$.
8. **Rhumes**
   Le nombre de fois où un individu attrape un rhume en une année donnée est une variable de Poisson de paramètre $\lambda = 5$. Un nouveau médicament vient d'être mis sur le marché. Il réduit le paramètre de Poisson à $\lambda = 3$ pour 75% de la population. Pour le reste de la population, le médicament est sans effet notable sur les rhumes. Si un individu essaie le médicament pendant un an et attrape 2 rhumes au cours de cette période, quelle est la probabilité que le médicament lui ait été bénéfique ?
9. On considère l'ensemble Ω des incidents possibles sur un réseau électrique. On sait que :

   10% des incidents provoquent la coupure du courant. . .

25% des incidents sont dûs à la foudre et, parmi ceux-là, 30% provoquent la coupure du courant.

Le courant étant coupé, quelle est la probabilité que ce soit dû à la foudre ?

10. Une boîte contient 3 boules rouges et 7 boules blanches. On en tire une au hasard et on la remplace dans la boîte par une boule de l'autre couleur. Puis, on tire une deuxième boule.
Quelle est la probabilité qu'elle soit rouge ?

11. Une compagnie d'assurance répartit ses clients en trois classes R_1, R_2 et R_3 : les bons risques, les risques moyens et les mauvais risques. Les effectifs des ces trois classes représentent 20% de la population totale pour la classe R_1, 50% pour R_2 et 30% pour R_3. Les statistiques indiquent que les probabilités d'avoir un accident au cours de l'année pour une personne de l'une de ces trois classes sont respectivement de 0.05, 0.15 et 0.30.
    (a) Quelle est la probabilité qu'une personne choisie au hasard dans la population ait un accident dans l'année ?
    (b) Si M. Martin n'a pas eu d'accident cette année, quelle est la probabilité qu'il soit un bon risque ?

12. Dans une fabrique, on a prélevé 6 lots égaux de pièces, parmi lesquels :

    Trois lots comportent 5% de pièces défectueuses.

    Deux lots en comportent 10%

    Un lot en comporte 13%

    Le contenu des 6 lots étant mélangé, on tire une pièce au hasard et on constate qu'elle est défectueuse. Quelle est la probabilité qu'elle provienne d'un lot à 5% ? d'un lot à 10% ? du lot à 13% ?

13. *Alcootest*
On s'intéresse à la fiabilité d'un alcootest pour automobilistes. Grâce à des études statistiques sur un grand nombre d'automobilistes, on sait que $0,5\%$ d'entre eux dépassent la dose d'alcool autorisée.
Aucun test n'est fiable à 100%. Avec celui que l'on considère, la probabilité que le test soit positif quand la dose d'alcool autorisée est dépassée, et la probabilité que le test soit négatif quand elle ne l'est pas, valent toutes deux $\rho = 0,95$.
Quelle est la probabilité qu'un automobiliste ayant un test positif ait réellement dépassé la dose d'alcool autorisée ?
Quelle devrait être la valeur de ρ pour que cette probabilité soit de 95% ?

14. Un garage peut réparer huit voitures par jour. Chaque jour, huit conducteurs indépendants ont rendez-vous, chacun ayant une probabilité de 90% de venir effectivement.
    (a) Quelle est la loi du nombre X de voitures présentes au garage un jour donné ? En déduire la probabilité que le garage ne soit pas plein (moins de huit voitures à réparer).
    (b) Un véhicule qui entre a 70% de chances d'être réparé (tous ne sont pas réparables). On note Y le nombre de voitures réparées un jour donné. Calculer $P(Y = k|X = n)$ (distinguer le cas $0 \leq k \leq n$ du cas $k > n$). En déduire que la loi de Y est une binômiale dont on précisera les paramètres.
    (c) Pour i de 1 à 8, on définit la variable Z_i par :
    $Z_i = 0$ si la i^{e} voiture ne vient pas ou n'est pas réparable.
    $Z_i = 1$ si la i^{e} voiture se présente et est réparée.
    En exprimant Y en fonction des Z_i, retrouver la loi de Y.

Chapitre 6

Variables aléatoires à densité

On dit que X est une variable aléatoire à densité s'il existe une fonction f positive et continue par morceaux sur $\mathbb{R}$ telle que

$$\forall t \in \mathbb{R} \quad F_X(t) = \int_{-\infty}^{t} f(x)\ dx.$$

On dit que f est la densité de (la loi de) la variable aléatoire

6.1 Propriétés

Soit X une variable aléatoire de densité f. f a les propriétés suivantes :

- $\forall a, b \in \mathbb{R} \quad a \leq b \Longrightarrow \quad P(a \leq X \leq b) = P(a \leq X < b) = P(a < X \leq b) = P(a < X < b) = \int_a^b f(x)\ dx.$
- $\forall a \in \mathbb{R} \quad P(a \leq X) = P(a < X) = \int_a^{+\infty} f(x)\ dx.$
- $\forall a \in \mathbb{R} \quad P(a \geq X) = P(a > X) = \int_{-\infty}^{a} f(x)\ dx.$
- $\int_{-\infty}^{+\infty} f(x)\ dx = 1.$

6.2 Lois usuelles

6.2.1 Loi uniforme sur un intervalle

On dit qu'une variable aléatoire X suit la loi uniforme sur l'intervalle $[a, b]$ si elle admet la densité

$$x \mapsto \frac{1}{b-a} 1\!\!1_{[a,b]}(x)$$

6.2.2 Loi gaussienne de paramètres m et σ^2

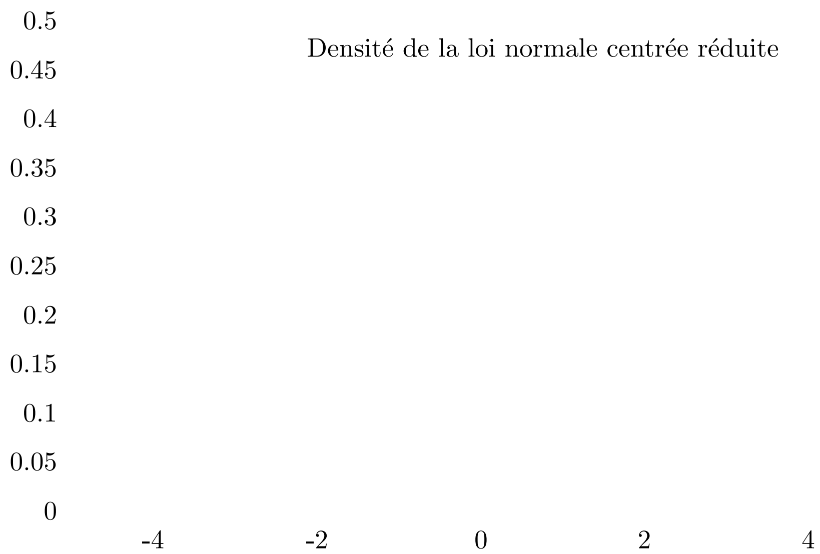
Soit $m \in \mathbb{R}$ et $\sigma^2 > 0$. On dit qu'une variable aléatoire X suit la loi gaussienne $\mathcal{N}(m, \sigma^2)$ de paramètres m et σ^2 si elle admet la densité

$$x \mapsto \frac{1}{\sqrt{2\pi}\sigma} \exp(-\frac{(x-m)^2}{2\sigma^2}).$$

On emploie également parfois le mot "normale" à la place de "gaussienne" : ces deux mots signifient exactement la même chose.

On dit qu'une variable gaussienne est centrée lorsque $m = 0$.
On dit qu'une variable gaussienne est centrée lorsque $\sigma^2 = 1$.

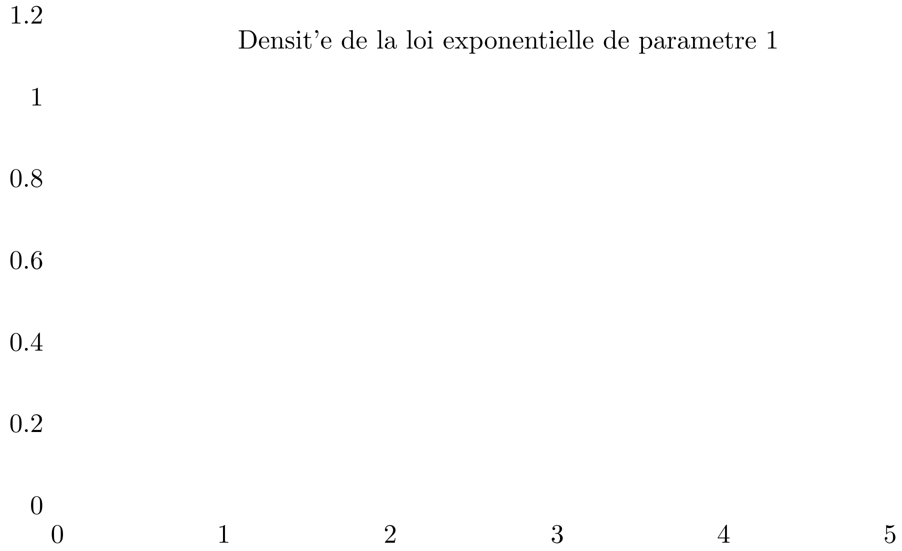
Remarques importantes : si $X \sim \mathcal{N}(m, \sigma^2)$, alors $aX+b \sim \mathcal{N}(m+b, a^2\sigma^2)$. En particulier, si $X \sim \mathcal{N}(m, \sigma^2)$, alors $\frac{X-m}{\sigma} \sim \mathcal{N}(0,1)$; et si $X \sim \mathcal{N}(0,1)$, alors $\sigma X + m \sim \mathcal{N}(m, \sigma^2)$.



6.2.3 Loi exponentielle de paramètres a

Soit $a > 0$. On dit qu'une variable aléatoire X suit la loi exponentielle de paramètres a si elle admet la densité

$$x \mapsto a \exp(-ax) \mathbb{1}_{\mathbb{R}^+}(x).$$


Densit'e de la loi exponentielle de parametre 1
1.2
1
0.8
0.6
0.4
0.2
0
0
1
2
3
4
5

6.3 Exercices

1. On choisit de manière uniforme sur $[0,1]$ un réel Y. Quelle est la probabilité pour que le polynome
$$p(x) = x^2 + x + Y$$
ait des racines complexes ? des racines distinctes ?
2. Dans le segment $[AB]$ de longueur 1, on choisit au hasard un point M. Quelle est la probabilité pour que l'on ait $AM.MB \geq \frac{2}{9}$?
3. Soit X une variable aléatoire suivant la loi uniforme sur $[a,b]$ et f une application affine. Que peut-on dire de la loi de $Y = f(X)$? Montrer que c'est faux dans le cas général.
4. Déterminer A tel que $f(x) = A\cos^2 x \mathbb{1}_{[-\pi/2,\pi/2]}(x)$ soit une densité de probabilité.
Si X est une loi de densité f, calculer $P(\frac{\pi}{4} \leq X \leq \frac{\pi}{3})$.
On rappelle que
$$\cos^2 x = \frac{1+\cos 2x}{2}.$$
5. Soient $X_1, \ldots, X_n, \ldots$ des variables aléatoires indépendantes suivant la loi uniforme sur $[0,1]$. On pose $M_n = \max(X_1, \ldots, X_n)$. Déterminer la fonction de répartition de M_n. Montrer que M_n admet une densité que l'on déterminera.
6. Soit X une variable aléatoire suivant la loi uniforme sur $[0,1]$. Déterminer la loi de $Z = -\ln(1-X)$.
7. Soit X une variable aléatoire de densité f. Montrer que la variable aléatoire $|X|$ admet comme densité
$$x \mapsto \mathbb{1}_{R_+}(x)(f(x) + f(-x)).$$
8. Soit X une variable aléatoire positive de densité f. Montrer que la variable aléatoire X^2 admet comme densité
$$x \mapsto \mathbb{1}_{R_+}(x)2x(f(x^2)).$$
Indication : utiliser un changement de variable.
9. Soit X une variable aléatoire normale centrée réduite. Montrer que la variable aléatoire X^2 est à densité et la déterminer.

10. La longueur des tiges de chrysanthèmes en fleurs coupées intervient dans le classement par catégorie. Pour simplifier, on supposera ici que c'est le seul critère de qualité employé. Un chrysanthème sera classé en catégorie extra si la longueur de sa tige est supérieure ou égale à 80 cm. Au 1er décembre, on évalue la production d'une certaine serre à 6000 chrysanthèmes pour le mois. A cette époque, les chrysanthèmes classés en catégorie extra sont payés au producteur 26 F les dix, tandis que les autres sont seulement payés 16F.

    La qualité de production a été étudiée sur de grands échantillons, et on en a conclut que la longueur des tiges coupées est une variable aléatoire qui suit une loi normale de moyenne 92cm et d'écart type 8cm.

    (a) Quelle est la probabilité pour qu'une fleur soit classée en catégorie extra ?

    (b) Quelle est l'espérance mathématique du nombre de fleurs qui seront classées en catégorie extra sur les 6000 fleurs de la production de décembre de cette année ?

    (c) En déduire l'espérance mathématique de la recette pour le total de la production de ce mois.

Chapitre 7

Moments, lois des grands nombres

On peut associer à la plupart des variables alétoires X définies sur un espace probabilisé $(\Omega, \mathcal{F}, P)$ un nombre représentant la “valeur moyenne” de X sous la probabilité P. De telles variables aléatoires sont dites intégrables et on note $\mathbb{E}X$ cette valeur moyenne. On note $\mathcal{L}^{\infty}(\Omega, \mathcal{F}, P)$ l'ensemble des variables aléatoires intégrables.

$\mathbb{E}X$ est appelée espérance de X.

7.1 Quelques propriétés

- L^1 est un espace vectoriel.
- $\forall X, Y \in L^1 \quad \mathbb{E}(X+Y) = \mathbb{E}X + \mathbb{E}Y$.
- $\forall X \in L^1, \forall a \in \mathbb{R} \quad \mathbb{E}aX = a\mathbb{E}X$.
- $\forall A \in \mathcal{F} \quad \mathbb{E}(\mathbb{1}_A) = P(A)$.
- La variable aléatoire X est intégrable si et seulement si $|X|$ est intégrable.
- $\forall X \in L^1 \quad P(X \geq 0) = 1 \Longrightarrow \mathbb{E}X \geq 0$
- $\forall X \in L^1 \quad P(X \leq a) = 1 \Longrightarrow \mathbb{E}X \leq a$.
- $\forall X \in L^1 \quad P(X \geq b) = 1 \Longrightarrow \mathbb{E}X \geq b$.
- $\forall X \in L^1 \quad P(X = a) = 1 \Longrightarrow \mathbb{E}X = a$.
- $\forall X \in L^1 \quad P(|X| \leq a) = 1 \Longrightarrow \mathbb{E}|X| \leq a$.
- Soient X, Y deux variables aléatoires vérifiant $0 \leq X \leq Y$. Si Y est intégrable, alors X est intégrable.
- $\forall X, Y \in L^1 \quad P(X \leq Y) = 1 \Longrightarrow \mathbb{E}X \leq \mathbb{E}Y$.

7.2 Calcul de l'espérance d'une variable aléatoire discrète

Théorème 20. *Soit X est une variable aléatoire discrète, et D un ensemble fini ou dénombrable inclus dans $\mathbb{R}$ tel que $X(\Omega) = D$. Alors, X est intégrable si et seulement si*

$$\sum_{i \in D} |i| p_i < +\infty,$$

où l'on a posé $p_i = P(X = i)$. Si cette somme est finie, on a alors

$$\mathbb{E}X = \sum_{i \in D} i p_i.$$

Il admet la généralisation suivante, appelée *théorème de transfert*.

Théorème 21. *Soit X est une variable aléatoire discrète, et D un ensemble fini ou dénombrable inclus dans $\mathbb{R}$ tel que $X(\Omega) = D$. Soit g une fonction quelconque de D dans $\mathbb{R}$. Alors, la variable aléatoire $Y = g(X)$ est intégrable si et seulement si*

$$\sum_{i \in D} |g(i)| p_i < +\infty,$$

où l'on a posé $p_i = P(X = i)$. Si cette somme est finie, on a alors

$$\mathbb{E}Y = \mathbb{E}g(X) = \sum_{i \in D} g(i) p_i.$$

7.3 Calcul de l'espérance d'une variable aléatoire à densité

Théorème 22. *Soit X est une variable aléatoire admettant la fonction f comme densité. Alors, X est intégrable si et seulement si*

$$\int_{\mathbb{R}} |x| f(x)\, dx < +\infty.$$

Si cette intégrale est convergente, on a alors

$$\mathbb{E}X = \int_{\mathbb{R}} x f(x)\, dx.$$

Voici maintenant le théorème de transfert pour les variables à densité

Théorème 23. *Soit X est une variable aléatoire admettant la fonction f comme densité, et g une fonction continue par morceaux définie sur $X(\Omega)$. Alors, la variable aléatoire $Y = g(X)$ est intégrable si et seulement si*

$$\int_{\mathbb{R}} |g(x)|f(x)\ dx < +\infty.$$

Si cette intégrale est convergente, on a alors

$$\mathbb{E}Y = \mathbb{E}g(X) = \int_{\mathbb{R}} g(x)f(x)\ dx.$$

Remarque importante : si la densité de X est paire, alors $\mathbb{E}X = 0$.

Démonstration. On a

$$\begin{aligned} \mathbb{E}X &= \int_{-\infty}^{0} xf(x)\ dx + \int_{0}^{+\infty} xf(x)\ dx \\ &= \int_{0}^{+\infty} -xf(-x)\ dx + \int_{0}^{+\infty} xf(x)\ dx \\ &= \int_{0}^{+\infty} x(f(x) - f(-x))\ dx \\ &= 0 \end{aligned}$$

car $f(x) = f(-x)$. □

7.4 Moments d'ordre 2

On dit qu'une variable aléatoire X admet un moment d'ordre 2 si elle est de carré intégrable, c'est à dire si $X^2 \in L^1$.

On note $L^2(\Omega, \mathcal{F}, P)$ (ou encore L^2) l'ensemble des variables aléatoires de carré intégrable.

Lemme 2. *Soient $X, Y \in L^2$. Alors la variable aléatoire XY est intégrable.*

Démonstration. Pour tous les réels a, b, on a $|ab| \leq \frac{1}{2}(a^2 + b^2)$.
(En effet, $a^2+b^2+2ab = (a+b)^2 \geq 0$, d'où $(a^2+b^2)/2 \geq -ab$ et $a^2+b^2-2ab = (a-b)^2 \geq 0$, d'où $(a^2+b^2)/2 \geq ab$.)
On a donc

$$0 \leq |XY| \leq \frac{1}{2}(X^2 + Y^2),$$

Comme $X^2 + Y^2$ est intégrable, on en déduit que $|XY|$ est intégrable, ce qui est ce que l'on voulait montrer □

Corollaire 1. *$L^2(\Omega, \mathcal{F}, P)$ est un espace vectoriel.*

Démonstration. La stabilité par multiplication ne pose pas de problème. Pour la stabilité par addition, il faut remarquer que

$$(X+Y)^2 = X^2 + Y^2 + 2XY,$$

puis utiliser le lemme précédent est le fait que L^1 est un espace vectoriel. □

7.4.1 Covariance et variance

Soient X et Y deux variables aléatoires. On appelle covariance du couple (X, Y) le nombre

$$\text{Covar } (X, Y) = \mathbb{E}(X - \mathbb{E}X)(Y - \mathbb{E}Y)$$

On appelle variance de X le nombre

$$\text{Var } X = \text{Covar } (X, X) = \mathbb{E}(X - \mathbb{E}X)^2.$$

On appelle écart-type de X le nombre

$$\sigma(X) = (\text{Var } X)^{1/2}.$$

On a les propriétés suivantes
- Covar $(X, Y) =$ Covar (Y, X).
- $\forall a, b \in \mathbb{R}$ Covar $(X - a, Y - b) =$ Covar (X, Y).
- Var $X + Y =$ Var $X +$ Var $Y + 2$Covar (X, Y).
- Covar $(X, Y) = \mathbb{E}XY - (\mathbb{E}X)(\mathbb{E}Y)$.
- Var $X = \mathbb{E}X^2 - (\mathbb{E}X)^2$.
- $|\mathbb{E}XY|^2 \le \mathbb{E}X^2\mathbb{E}Y^2$ (inégalité de Cauchy-Schwarz)
- $|$Covar $(X, Y)| \le \sigma(X)\sigma(Y)$.

Démonstration.

Évident.

Si l'on pose $\overline{X} = X - \mathbb{E}X$, on a Covar $(X, Y) = \mathbb{E}\overline{XY}$. Or $\overline{X - a} = X - a - \mathbb{E}(X - a) = X - a - \mathbb{E}X + \mathbb{E}a = X - a - \mathbb{E}X + a = \overline{X}$. De même pour Y. La formule s'ensuit.

Var $(X + Y) =$ Covar $(X + Y, X + Y) = \mathbb{E}\overline{X+Y}^2$. Il est facile de voir que $\overline{X+Y} = \overline{X} + \overline{Y}$. On a donc Var $X + Y = \mathbb{E}(\overline{X} + \overline{Y})^2 = \mathbb{E}\overline{X}^2 + \overline{Y}^2 + 2\overline{XY} =$ Covar $(X, X) +$ Covar $(Y, Y) + 2$Covar $(X, Y) =$ Var $X +$ Var $Y + 2$Covar (X, Y).

$(X-\mathbb{E}X)(Y-\mathbb{E}Y) = XY+\mathbb{E}X\mathbb{E}Y-(\mathbb{E}X)Y-(\mathbb{E}Y)X$. D'où $\mathbb{E}(X-\mathbb{E}X)(Y-\mathbb{E}Y) = \mathbb{E}XY+\mathbb{E}(EX\mathbb{E}Y)-\mathbb{E}X\mathbb{E}Y-\mathbb{E}Y\mathbb{E}X = \mathbb{E}XY+\mathbb{E}X\mathbb{E}Y-2\mathbb{E}X\mathbb{E}Y = \mathbb{E}XY-(\mathbb{E}X)(\mathbb{E}Y)$.

Il suffit d'appliquer la formule précédente avec $X = Y$.

Considérons, pour $t \in \mathbb{R}$:

$$f(t) = \mathbb{E}(tX+Y)^2.$$

Comme $(tX+Y)^2 \geq 0$, on a $f(t) \geq 0$ pour tout t. Or

$$f(t) = t^2\mathbb{E}X^2 + 2t\mathbb{E}XY + \mathbb{E}Y^2$$

est un polynôme du second degré. Comme il ne change pas de signe, son discriminant réduit est négatif :

$$\Delta' = b'^2 - ac = (\mathbb{E}XY)^2 - \mathbb{E}X^2\mathbb{E}Y^2 \leq 0,$$

d'où le résultat

Il suffit d'appliquer l'inégalité de Cauchy-Schwarz à $\overline{X}$ et $\overline{Y}$, puis de prendre la racine carrée. □

Lorsque $\sigma(X)$ et $\sigma(Y)$ sont non nuls, on définit le coefficient de corrélation de X et Y par

$$\text{Corr }(X,Y) = \frac{\text{Covar }(X,Y)}{\sigma(X)\sigma(Y)}.$$

D'après ce qui précède, Covar $(X,Y) \in [-1;1]$. Lorsque Covar $(X,Y) = 0$ (ce qui implique Corr $(X,Y) = 0$ si $\sigma(X)$ et $\sigma(Y)$ sont non nuls), on dit que X ne sont pas corrélées.

Lorsque Covar $(X,Y) \geq 0$ (ce qui implique Corr $(X,Y) \geq 0$ si $\sigma(X)$ et $\sigma(Y)$ sont non nuls), on dit que X et Y sont positivement corrélées.

Lorsque Covar $(X,Y) \leq 0$ (ce qui implique Corr $(X,Y) \leq 0$ si $\sigma(X)$ et $\sigma(Y)$ sont non nuls), on dit que X et Y sont négativement corrélées.

7.4.2 Un exemple de calcul

On suppose que X prend les valeurs 1 et 2, Y les valeurs 0 et 1. On suppose que l'on a

$$\begin{aligned} P(\{X=1\}\cap\{Y=0\}) &= \tfrac{1}{3} \\ P(\{X=1\}\cap\{Y=1\}) &= \tfrac{1}{6} \\ P(\{X=2\}\cap\{Y=1\}) &= \tfrac{1}{4} \\ P(\{X=2\}\cap\{Y=0\}) &= \tfrac{1}{4} \end{aligned}$$

La matrice associée à la loi du couple est

$$\begin{pmatrix} 1/3 & 1/4 \\ 1/6 & 1/4 \end{pmatrix}$$

En faisant la somme ligne par ligne, on obtient la loi de Y :

$$\begin{pmatrix} P(Y=0) \\ P(Y=1) \end{pmatrix} = \begin{pmatrix} 1/3+1/4 \\ 1/6+1/4 \end{pmatrix}$$

En faisant la somme colonne par colonne, on obtient la loi de X :

$$\begin{pmatrix} P(X=0) & P(X=1) \end{pmatrix} = \begin{pmatrix} 1/3+1/6 & 1/4+1/4 \end{pmatrix} = \begin{pmatrix} 1/2 & 1/2 \end{pmatrix}$$

On a alors

$$\begin{aligned}
\mathbb{E}Y &= 0 \times \frac{7}{12} + 1 \times \frac{5}{12} = \frac{5}{12} \\
\mathbb{E}X &= 1 \times \frac{1}{2} + 2 \times \frac{1}{2} = \frac{3}{2} \\
\mathbb{E}XY &= (0 \times 1) \times \frac{1}{3} + (0 \times 2) \times \frac{1}{4} + (1 \times 1) \times \frac{1}{6} + (1 \times 2) \times \frac{1}{4} = \frac{2}{3}
\end{aligned}$$

D'où Covar $(X,Y) = \mathbb{E}XY - \mathbb{E}X\mathbb{E}Y = \frac{2}{3} - \frac{5}{12}\frac{3}{2} = \frac{1}{24}$

$$\begin{aligned}
\mathbb{E}Y^2 &= 0^2 \times \frac{7}{12} + 1^2 \times \frac{5}{12} = \frac{5}{12} \\
\mathbb{E}X^2 &= 1^2 \times \frac{1}{2} + 2^2 \times \frac{1}{2} = \frac{5}{2} \\
\text{Var } Y &= \mathbb{E}Y^2 - (\mathbb{E}Y)^2 = \frac{35}{(12)^2} \\
\text{Var } X &= \mathbb{E}X^2 - (\mathbb{E}X)^2 = \frac{35}{(12)^2} = 4
\end{aligned}$$

$$\text{Corr }(X,Y) = \frac{\text{Covar }(X,Y)}{\sqrt{\text{Var } X \text{Var } Y}} = \frac{1}{24} \times \frac{1}{2} \times \frac{12}{\sqrt{35}} = \frac{1}{4\sqrt{35}}$$

7.4.3 Espérance et indépendance

On admettra le théorème suivant très important :

Théorème 24. *Soient X, Y deux variables aléatoires intégrables indépendantes. Alors, leur produit XY est une variable aléatoire intégrable et l'on a*

$$\mathbb{E}XY = \mathbb{E}X\mathbb{E}Y.$$

Corollaire 2. *Soient X, Y deux variables aléatoires intégrables indépendantes. Alors X et Y ne sont pas corrélées.*

Démonstration. On a Covar $(X, Y) = \mathbb{E}XY - \mathbb{E}X\mathbb{E}Y = 0$. □

Remarque importante : des variables aléatoires peuvent être non corrélées sans être indépendantes.
Exemple : soient deux variables aléatoires vérifiant

$$P(\{X=1\}\cap\{Y=1\}) = P(\{X=1\}\cap\{Y=-1\}) = P(\{X=-1\}\cap\{Y=0\}) = 1/3.$$

La matrice associée à la loi du couple est

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

Corollaire 3. *Soient X, Y deux variables aléatoires indépendantes de carré intégrable. Alors on a*

$$Var\ (X+Y) = Var\ X + Var\ Y.$$

Démonstration. On a toujours Var $X+Y =$ Var $X+$Var $Y+2$Covar (X, Y).Comme X et Y sont indépendantes, elles ne sont pas corrélées, d'où le résultat. □

7.5 Calcul des premiers moments des lois discrètes usuelles

7.5.1 Indicatrice d'un événement

On rappelle que pour $A \subset \Omega$, l'application $\mathbb{1}_A$ (appelée indicatrice de A) est définie sur Ω par

$$\mathbb{1}_A(x) = \begin{cases} 1 \text{ si } x \in A \\ 0 \text{ si } x \notin A \end{cases}$$

1_A est une variable aléatoire à valeurs dans $\{0; 1\}$. Il est important de remarquer que, comme $\forall x \in \{0; 1\} \quad x^2 = x$, on a $\mathbb{1}_A^2 = \mathbb{1}_A$. Maintenant, on a

- $\mathbb{E}\mathbb{1}_A = P(A)$.
- Var $\mathbb{1}_A = \mathbb{E}\mathbb{1}_A^2 - (\mathbb{E}\mathbb{1}_A)^2 = \mathbb{E}\mathbb{1}_A - (\mathbb{E}\mathbb{1}_A)^2 = P(A) - P(A)^2 = P(A)(1 - P(A))$.

7.5.2 Loi binomiale

On a vu que la loi binomiale de paramètres n et p était la loi de

$$X = \sum_{k=1}^{n} \mathbb{1}_{A_k},$$

où $A_1, \dots, A_n$ sont n événements indépendants de même probabilité p. On a donc

$$\mathbb{E}X = \sum_{k=1}^{n} \mathbb{E}\mathbb{1}_{A_k} = \sum_{k=1}^{n} P(A_k) = np,$$

et comme les variables aléatoires sont indépendantes

$$\text{Var } X = \sum_{k=1}^{n} \text{Var } \mathbb{1}_{A_k} = \sum_{k=1}^{n} P(A_k)(1 - P(A_k) = np(1-p).$$

7.5.3 Loi géométrique

Soit X une variable aléatoire suivant une loi géométrique de paramètre $p \in]0,1]$. On a

$$\begin{aligned}
\mathbb{E}X &= \sum_{k=0}^{+\infty} kP(X=k) \\
&= \sum_{k=1}^{+\infty} kP(X=k) \\
&= \sum_{k=1}^{+\infty} kp(1-p)^{k-1} \\
&= p\sum_{k=1}^{+\infty} k(1-p)^{k-1} \\
&= p\frac{1}{(1-(1-p))^2} \\
&= \frac{1}{p}.
\end{aligned}$$

$$\begin{aligned}
\mathbb{E}X(X-1) &= \sum_{k=0}^{+\infty} k(k-1)P(X=k) \\
&= \sum_{k=2}^{+\infty} k(k-1)P(X=k) \\
&= \sum_{k=2}^{+\infty} k(k-1)p(1-p)^{k-1} \\
&= p(1-p)\sum_{k=2}^{+\infty} k(k-1)(1-p)^{k-2} \\
&= p(1-p)\frac{2}{(1-(1-p))^3} \\
&= \frac{2(1-p)}{2p^2}.
\end{aligned}$$

On a alors $\mathbb{E}X^2 = \mathbb{E}X(X-1) + \mathbb{E}X = \frac{1}{p} + \frac{2(1-p)}{p^2}$ et
Var $X = \mathbb{E}X^2 - (\mathbb{E}X)^2 = \frac{1}{p} + \frac{2(1-p)}{p^2} - \frac{1}{p^2} = \frac{1-p}{p^2}$.

7.5.4 Loi de Poisson

$$\begin{aligned}
\mathbb{E}X &= \sum_{k=0}^{+\infty} kP(X=k) \\
&= \sum_{k=1}^{+\infty} kP(X=k) \\
&= \sum_{k=1}^{+\infty} ke^{-\lambda}\frac{\lambda^k}{k!} \\
&= e^{-\lambda}\lambda\sum_{k=1}^{+\infty} \frac{\lambda^{k-1}}{(k-1)!} \\
&= e^{-\lambda}\lambda e^{\lambda} \\
&= \lambda
\end{aligned}$$

$$\begin{aligned}
\mathbb{E}X(X-1) &= \sum_{k=0}^{+\infty} k(k-1)P(X=k) \\
&= \sum_{k=2}^{+\infty} k(k-1)P(X=k) \\
&= \sum_{k=2}^{+\infty} k(k-1)e^{-\lambda}\frac{\lambda^k}{k!} \\
&= e^{-\lambda}\lambda^2\sum_{k=2}^{+\infty} \frac{\lambda^{k-2}}{(k-2)!} \\
&= e^{-\lambda}\lambda^2 e^{\lambda} \\
&= \lambda^2
\end{aligned}$$

On a alors $\mathbb{E}X^2 = \mathbb{E}X(X-1) + \mathbb{E}X = \lambda^2 + \lambda$ et
Var $X = \mathbb{E}X^2 - (\mathbb{E}X)^2 = \lambda^2 + \lambda - \lambda^2 = \lambda$.

7.6 Calcul des premiers moments des lois à densité usuelles

7.6.1 Loi uniforme sur un segment

Soit X une variable aléatoire suivant la loi uniforme sur $[-1,1]$. La densité de X est donc

$$x \mapsto \frac{1}{2}\mathbb{1}_{[-1,1]}(x).$$

On a donc

$$\mathbb{E}X = \int_{-1}^{1} \frac{1}{2}x\ dx = 0$$

et

$$\mathbb{E}X^2 = \int_{-1}^{1} \frac{1}{2}x^2\ dx = \frac{1}{3}.$$

Comme X est centrée, on a Var $X = \mathbb{E}X^2$.

Passons au cas général : on pose $Y = \frac{a+b}{2} + \frac{b-a}{2}X$. X suit la loi uniforme sur $[-1,1]$ si et seulement Y suit la loi uniforme sur $[a,b]$. On a alors

- $\mathbb{E}Y = \mathbb{E}\frac{a+b}{2} + \frac{b-a}{2}\mathbb{E}X = \frac{a+b}{2}$.
- Var $Y = (\frac{b-a}{2})^2$Var $X = \frac{(b-a)^2}{12}$

7.6.2 Loi gaussienne

Soit X une variable aléatoire suivant la loi $\mathcal{N}(0,1)$. On rappelle que la densité de X est

$$f(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}.$$

Il est facile de vérifier que

$$\frac{\mathrm{d}}{\mathrm{dx}}(xf(x)) = (1-x^2)f(x).$$

On a donc

$$\forall a,b \in \mathbb{R} \quad bf(b) - af(a) = \int_a^b f(x)\ dx - \int_a^b x^2 f(x)\ dx$$

Comme $\lim_{a\to-\infty} af(a) = \lim_{b\to+\infty} bf(b) = 0$ et que $\int_{-\infty}^{+\infty} f(x)\ dx = 1$, on en déduit que

$$\int_{-\infty}^{+\infty} x^2 f(x)\ dx = 1.$$

Autrement dit, X admet un moment d'ordre 2 : $\mathbb{E}X^2 = 1$.

D'autre part, l'existence d'un moment d'ordre 2 implique celle d'un moment d'ordre 1. Comme la densité de X est paire, on en déduit que $\mathbb{E}X = 0$. On a donc Var $X = \mathbb{E}X^2 = 1$.

Passons au cas général. Si l'on a $Y = m + \sigma X$, on sait que Y suit la loi $\mathcal{N}(m, \sigma^2)$. On a alors $\mathbb{E}Y = m + \sigma\mathbb{E}X = m$ et Var $Y = \sigma^2$Var $X = \sigma^2$.

7.6.3 Lois exponentielles

Lemme 3. *Soit X une variable aléatoire suivant une loi exponentielle de paramètre 1 et f fonction croissante de classe C^1 de $\mathbb{R}_+$ dans $\mathbb{R}_+$ telle que $f'(X)$ soit intégrable et telle que $\lim_{t\to+\infty} e^{-t}f(t) = 0$. Alors $f(X)$ est intégrable et on a la formule l'intégration par parties suivante :*

$$\mathbb{E}f(X) = f(0) + \mathbb{E}f'(X)$$

Démonstration. Il suffit d'écrire

$$\begin{aligned}
\int_O^A e^{-t} f(t)\ dt &= \big[-e^{-t}f(t)\big]_0^A + \int_O^A e^{-t} f'(t)\ dt \\
&= -e^{-A}f(A) + f(0) + \int_O^A e^{-t} f'(t)\ dt
\end{aligned}$$

En faisant tendre A vers $+\infty$, on obtient

$$\int_O^{+\infty} e^{-t} f(t)\ dt = f(0) + \int_O^{+\infty} e^{-t} f'(t)\ dt,$$

soit

$$\mathbb{E}f(X) = f(0) + \mathbb{E}f'(X)$$

(Remarquer l'importance du fait que $f'(X)$ et $f(X)$ sont des variables aléatoires positives). □

On applique d'abord le lemme à $f(x) = x$: on obtient $\mathbb{E}X = 0 + \mathbb{E}1 = 1$. Ensuite on l'applique à $f(x) = x^2$: on obtient $\mathbb{E}X^2 = 0 + \mathbb{E}2X = 2\mathbb{E}X = 2$. On a donc Var $X = \mathbb{E}X^2 - (\mathbb{E}X)^2 = 2 - 1^2 = 1$.

Passons au cas général. Soit $a > 0$. La variable aléatoire $Y = \frac{1}{a}X$ suit la loi exponentielle de paramètre a. On a donc $\mathbb{E}Y = \frac{1}{a}\mathbb{E}X = \frac{1}{a}$ et Var $Y = \frac{1}{a^2}\text{Var } X = \frac{1}{a^2}$.

7.7 Approximation des lois

7.7.1 Approximation d'une loi binômiale par une loi gaussienne

Théorème 25. *Soit, pour $n \geq 1$ une variable aléatoire S_n suivant une loi binômiale de paramètres n et p. On note Z_n la variable aléatoire centrée réduite associée à S_n, c'est à dire*

$$Z_n = \frac{S_n - \mathbb{E}S_n}{\sqrt{n\,Var\ S_n}} = \frac{X_n - np}{\sqrt{np(1-p)}}.$$

Alors, pour tout intervalle $I \subset \mathbb{R}$, on a

$$\lim_{n\to\infty} P(Z_n \in I) = P(Z \in I),$$

où Z est une variable aléatoire suivant la loi gaussienne centrée réduite.

Application pratique Si n est "grand" et npq "pas trop petit", on peut remplacer la loi binomiale par une loi gaussienne de même espérance et de même variance. D'après une grand-mère statisticienne, n est grand à partir de 30 et npq n'est pas trot petit dès que $n \min(p, 1-p) \geq 5$.

7.7.2 Approximation d'une loi binômiale par une loi de Poisson

Théorème 26. *Soit, pour $n \geq 1$ une variable aléatoire X_n suivant une loi binômiale de paramètres n et p_n. On suppose que l'on a*

$$\lim_{n\to\infty} np_n = \lambda > 0$$

Alors, pour tout ensemble $A \subset \mathbb{N}$, on a

$$\lim_{n\to\infty} P(X_n \in A) = P(X \in A),$$

où X est une variable aléatoire suivant la loi de Poisson de paramètre λ.

Application pratique Si n est "grand" et np "pas trop grand", on peut remplacer la loi binomiale par une loi de Poisson. D'après une grand-mère statisticienne, n est grand à partir de 30 et np n'est pas trop grand jusqu'à 10. Ce théorème peut etre interprété de la manière suivante : la loi de Poisson est une bonne modélisation pour de nombre de fois où un événement rare survient.

7.8 Lois des grands nombres

7.8.1 * Inégalité de Markov

Théorème 27. *Soit X une variable aléatoire positive, intégrable. Alors, on a*

$$\forall a > 0 \quad P(X \geq a) \leq \frac{\mathbb{E}X}{a}.$$

Démonstration. On a les inégalités

$$\forall \omega \in \Omega a \mathbb{1}_{X \geq a}(\omega) \leq X(\omega)\mathbb{1}_{X \geq a}(\omega) \leq X(\omega).$$

Démontrons la première inégalité :

- si $X(\omega) \geq a$, on a $\mathbb{1}_{X \geq a}(\omega) = 1$, et donc l'inégalité à démontrer est exactement l'hypothèse $X(\omega) \geq a$.
- si $X(\omega) < a$, on a $\mathbb{1}_{X \geq a}(\omega) = 0$, et donc l'inégalité à démontrer se réduit à $0 \leq 0$.

Montrer la deuxième inégalité se résume à remarquer que $0 \leq \mathbb{1}_{X \geq a}(\omega)$ On prend alors l'espérance et on obtient

$$\mathbb{E} a \mathbb{1}_{X \geq a} \leq \mathbb{E}X.$$

Comme $\mathbb{E} a \mathbb{1}_{X \geq a} = a\mathbb{E}\mathbb{1}_{X \geq a} = aP(X \geq a)$, on obtient l'inégalité désirée. □

7.8.2 * Inégalité de Tchebytchef

Théorème 28. *Soit X une variable aléatoire admettant un moment d'ordre 2. Alors, on a*

$$\forall a > 0 \quad P(|X - \mathbb{E}X| \geq a) \leq \frac{\mathit{Var}\ X}{a^2}.$$

Démonstration.

$$P(|X - \mathbb{E}X| \geq a) = P(|X - \mathbb{E}X|^2 \geq a^2)$$

Il suffit alors d'appliquer l'inégalité de Markov à la variable aléatoire $Y = |X - \mathbb{E}X|^2$. Comme $\mathbb{E}Y = \text{Var } X$, l'inégalité s'ensuit. □

7.8.3 Loi faible des grands nombres

Théorème 29. *Soit $(X_n)_{n\geq 0}$ une suite de variables aléatoires de même loi, admettant un moment d'ordre 2 et deux à deux non corrélées. On pose $S_n = \sum_{k=1}^n X_k$ et $M_n = \frac{1}{n}S_n$. Alors*

1. $\mathbb{E}|M_n - \mathbb{E}X_0|^2 \to 0$. *On dit que M_n converge en moyenne quadratique vers $\mathbb{E}X_0$.*
2. $\forall \varepsilon > 0 \quad P(|M_n - \mathbb{E}X_0| \geq \varepsilon) \to 0$. *On dit que M_n converge en probabilité vers $\mathbb{E}X_0$.*

Démonstration.

1. $\mathbb{E}M_n = \frac{1}{n}\mathbb{E}S_n = \frac{1}{n}\sum_{k=1}^n EX_k = \frac{1}{n}n\mathbb{E}X_0 = \mathbb{E}X_0$. Par conséquent $|M_n - \mathbb{E}X_0|^2 = \text{Var } M_n = \frac{1}{n^2}\text{Var } S_n$. Comme les X_k sont 2 à 2 non corrélées, on a

$$\text{Var } S_n = \sum_{k=1}^n \text{Var } X_k = n\text{Var } X_1.$$

On a donc

$$\mathbb{E}|M_n - \mathbb{E}X_0|^2 = \frac{\text{Var } X_1}{n},$$

qui tend bien vers zéro.

2. D'après l'inégalité de Tchebitchev, on a

$$P(|M_n - \mathbb{E}X_0| \geq \varepsilon) \leq \frac{\text{Var } M_n}{\varepsilon^2},$$

qui tend vers zéro d'après ce qui précède.

□

7.8.4 Une loi forte des grands nombres

Théorème 30. *Soit* $(X_n)_{n\geq 0}$ *une suite de variables aléatoires indépendantes de même loi, admettant un moment d'ordre 4 (c'est à dire que* $\mathbb{E}X_1^4$ *est fini) . On pose* $S_n = \sum_{k=1}^n X_k$ *et* $M_n = \frac{1}{n}S_n$*. Alors il existe un événement observable* $\tilde{\Omega} \subset \Omega$ *avec* $P(\tilde{\Omega} \subset \Omega) = 1$ *et*

$$\forall \omega \in \tilde{\Omega} \subset \Omega \quad M_n(\omega) \to \mathbb{E}X_1.$$

7.8.5 Probabilités et fréquences asymptotiques

Théorème 31. *Soit* $(A_n)_{n\geq 0}$ *une suite d'événements observables indépendants de même probabilité* p*. Pour* ω *dans l'univers* Ω *On note* $N_n(\omega)$ *le nombre d'événements qui sont réalisés parmi* $A_1, \dots, A_n$*. (Ainsi, on a* $N_n = \sum_{k=1}^n \mathbb{1}_{A_k}$ *et* $f_n = \frac{1}{n}N_n$*. Alors il existe un événement observable* $\tilde{\Omega} \subset \Omega$ *avec* $P(\tilde{\Omega} \subset \Omega) = 1$ *et*

$$\forall \omega \in \tilde{\Omega} \subset \Omega \quad f_n(\omega).$$

Démonstration. Il suffit de poser $X_k = \mathbb{1}_{A_k}$ et d'appliquer le théorème précédent. X_k admet bien un moment d'ordre 4 car $0 \leq X_k \leq 1$ et l'on a $\mathbb{E}X_1 = P(A_1) = p$. □

7.9 Exercices

1. Soient A, B deux éléments observables. On note
$$A\Delta B = \{x \in A; x \notin B\} \cup \{x \in B; x \notin A\}.$$
Ce sont donc les éléments qui sont dans A ou dans B, mais pas dans les deux. Montrer $1\!\!1_{A\Delta B} = (1\!\!1_A - 1\!\!1_B)^2$. En déduire
$$P(A\Delta B) = P(A) + P(B) - 2P(A \cap B).$$

2. Soient X, Y deux variables aléatoires normales centrées réduites indépendantes. On pose
$$Z = \cos\theta X + \sin\theta Y.$$
Calculer $\text{Corr }(X, X), \text{Corr }(X, -X), \text{Corr }(X, Y), \text{Corr }(X, Z)$.

3. Soit A un réel. On considère la fonction f définie sur $\mathbb{R}$ par
$$f(x) = Ax(2 - x)1\!\!1_{[0,2]}(x).$$
   (a) Pour quelle(s) valeur(s) de A f est-elle la densité d'une loi de probabilité ?
   (b) On suppose maintenant que A a été choisi de telle manière que f soit la densité d'une loi de probabilité.
   Soit donc X une variable aléatoire admettant f comme densité.
   Calculer $\mathbb{E}X$, puis $\text{Var } X$.

4. On suppose que $Y = \ln X$ vérifie $X \sim \mathcal{N}(m, \sigma^2)$ (on dit alors que Y est log-normale). Calculer $\mathbb{E}X$ et $\text{Var } X$.

5. Calculer $\mathbb{E}\sin X$, où $P(X = \frac{\pi}{2}) = \frac{1}{6}, P(X = \frac{\pi}{3}) = \frac{1}{3}, P(X = \frac{\pi}{6}) = \frac{1}{2}$.

6. Soient X, Y deux variables aléatoires suivant chacune une loi uniforme sur $[a, b]$. Montrer que $\mathbb{E}|X - Y| \le \frac{b-a}{2}$. Indication : on pourra commencer par supposer que la loi est centrée (c'est à dire que $a + b = 0$) et faire une majoration simple). On s'y ramènera dans le cas général.
Que vaut $\mathbb{E}|X - Y|$ lorsque X et Y sont indépendantes ?

7. Soit X une variable aléatoire suivant la loi uniforme sur $\{0, \ldots, n\}$. Calculer $\mathbb{E}X$ et $\text{Var } X$.

Chapitre 8

Estimation

Dans le premier chapitre, on a vu comment calculer la moyenne et la variance d'un caractère d'une population. Néanmoins, dans la pratique, on ne dispose que très rarement des données sur la population totale : on a uniquement les données d'un échantillon tiré au hasard par sondage. Comment, à partir des données de cet échantillon, estimer certains paramètres – tels que la variance, la moyenne – de la population totale ? Tel est l'objectif de ce chapitre.
Définition On appelle estimateur non biaisé d'un paramètre p toute variable aléatoire Z dont l'espérance $\mathbb{E}Z$ vaut p.

8.1 Estimation de la moyenne et de la variance

Soient $X_1, X_2, \dots, X_n$ des variables aléatoires indépendantes suivant la même loi. On note m leur espérance commune et σ^2 leur variance commune.

Notre objectif est d'estimer m et σ^2 à partir de $X_1, X_2, \dots, X_n$.

Une idée naturelle est d'approcher m par la variable aléatoire

$$m_n = \frac{1}{n}\sum_{k=1}^{n} X_k = \frac{X_1 + X_2 + \cdots + X_n}{n}.$$

C'est une bonne idée !

En effet

$$\mathbb{E}m_n = \mathbb{E}\frac{1}{n}\sum_{k=1}^{n} X_k = \frac{1}{n}\mathbb{E}\sum_{k=1}^{n} X_k = \frac{1}{n}\sum_{k=1}^{n}\mathbb{E}X_k = \frac{1}{n}\sum_{k=1}^{n} m = \frac{1}{n}nm = m.$$

m_n est donc un estimateur non biaisé de m.

Maintenant, comment estimer σ^2 ? Une autre idée naturelle est d'approcher σ^2 par la variable aléatoire

$$\sigma_n^2 = \frac{1}{n}\sum_{k=1}^{n}(X_k - m_n)^2 = \frac{1}{n}\sum_{k=1}^{n}X_k^2 - \big(\frac{1}{n}\sum_{k=1}^{n}X_k\big)^2.$$

On a d'une part,

$$\mathbb{E}\frac{1}{n}\sum_{k=1}^{n}X_k^2 = \frac{1}{n}\sum_{k=1}^{n}\mathbb{E}X_k^2 = \frac{1}{n}\sum_{k=1}^{n}\text{Var } X_k+(\mathbb{E}X_k)^2 = \frac{1}{n}\sum_{k=1}^{n}\sigma^2+m^2 = \sigma^2+m^2.$$

D'autre part

$$\mathbb{E}\big(\frac{1}{n}\sum_{k=1}^{n}X_k\big)^2 = \frac{1}{n^2}\big(\sum_{k=1}^{n}X_k\big)^2.$$

On a

$$\big(\sum_{k=1}^{n}X_k\big)^2 = \sum_{k=1}^{n}X_k^2 + 2\sum_{1\le i<j}X_iX_k$$

On a pour tout k : $\mathbb{E}X_k^2 = \sigma^2 + m^2$ et, pour tous i, j distincts $\mathbb{E}X_iX_j = \mathbb{E}X_i\mathbb{E}X_j = m^2$ car X_i et X_j sont indépendantes. On en déduit

$$\begin{aligned}
\mathbb{E}\big(\sum_{k=1}^{n}X_k\big)^2 &= \sum_{k=1}^{n}\mathbb{E}X_k^2 + 2\sum_{1\le i<j}\mathbb{E}X_iX_k \\
&= \sum_{k=1}^{n}\sigma^2 + m^2 + 2\sum_{1\le i<j}m^2 \\
&= n(\sigma^2+m^2) + 2\frac{n(n-1)}{2}m^2 \\
&= n(\sigma^2+m^2) + n(n-1)m^2
\end{aligned}$$

Ainsi

$$\begin{aligned}
\mathbb{E}\sigma_n^2 &= \sigma^2 + m^2 - \frac{1}{n^2}\big(n(\sigma^2+m^2) + n(n-1)m^2\big) \\
&= \sigma^2 + m^2 - \frac{1}{n}\big((\sigma^2+m^2) + (n-1)m^2\big) \\
&= \frac{1}{n}\big(n(\sigma^2+m^2) - (\sigma^2+m^2) + (n-1)m^2\big) \\
&= \frac{n-1}{n}\sigma^2
\end{aligned}$$

Ce n'est pas tout à fait ce que l'on voulait, car on avait espéré trouver σ^2. Ce n'est pas grave : il suffit de considérer

$$\begin{aligned} s_n^2 = \frac{n}{n-1}\sigma_n^2 &= \frac{n}{n-1}\Big(\frac{1}{n}\sum_{k=1}^{n}(X_k - m_n)^2\Big) \\ &= \Big(\frac{1}{n-1}\sum_{k=1}^{n}X_k^2\Big) - \frac{n}{n-1}m_n^2. \end{aligned}$$

s_n^2 est un estimateur non biaisé de σ car $\mathbb{E}s_n^2 = \frac{n}{n-1}\mathbb{E}\sigma_n^2 = \sigma^2$.

Remarque importante : Si Z est un estimateur sans biais de p et f une fonction, en général, $f(Z)$ n'est PAS un estimateur sans biais de $f(p)$.

8.2 Application à l'estimation d'une proportion

On suppose qu'une proportion p de la population totale possède un caractère M (par exemple une maladie). Dans un échantillon de n personnes, on a trouvé k personnes possédant ce caractère.

Deux problèmes se posent :
- estimer p
- estimer $\sigma^2 = p(1-p)$

Il suffit d'appliquer le résultat précédent dans le cas où la variable aléatoire X_i vaut 1 si l'individu i possède le caractère M et 0 sinon. En d'autres termes, $X_i = \mathbb{1}_{M_i}$ où M_i est l'événement "le i-ème individu possède le caractère M".

On a alors pour $p = \mathbb{E}X_1$ l'estimateur sans biais

$$m_n = \frac{1}{n}\sum_{k=1}^{n}X_k = \frac{X_1 + X_2 + \cdots + X_n}{n} = \frac{k}{n}.$$

On pose alors $f = k/n$. Comme estimateur sans biais de $\sigma^2 = p(1-p)$, on a

$$\begin{aligned} s_n^2 &= (\frac{1}{n-1}\sum_{k=1}^{n} X_k^2) - \frac{n}{n-1}m_n^2 \\ &= (\frac{1}{n-1}\sum_{k=1}^{n} X_k) - \frac{n}{n-1}m_n^2 \\ &= (\frac{n}{n-1}m_n) - \frac{n}{n-1}m_n^2 \\ &= \frac{n}{n-1}f(1-f). \end{aligned}$$

8.3 Estimation d'une proportion par un intervalle de confiance

On reprend les hypothèses de la section précédente. On veut savoir si l'estimateur $m_n = f$ est souvent proche de la valeur réelle p.

Si l'on pose $S_n = X_1 + X_2 + \cdots + X_n$, on sait que S_n suit une loi binômiale de paramètres n et p, donc d'après le chapitre précédent, on sait que

si l'on pose

$$Z_n = \frac{S_n - \mathbb{E}S_n}{\sqrt{n\mathrm{Var}\ S_n}} = \frac{S_n - np}{\sqrt{np(1-p)}}.$$

Alors, pour tout intervalle $I \subset \mathbb{R}$, on a

$$\lim_{n\to\infty} P(Z_n \in I) = P(Z \in I),$$

où Z est une variable aléatoire suivant la loi gaussienne centrée réduite.

En particulier, pour tout $a > 0$, on a

$$\lim_{n\to\infty} P(|\frac{S_n - np}{\sqrt{np(1-p)}}| \le a) = P(|Z \le |a),$$

Comme $\frac{n}{n-1}f(1-f)$ un estimateur sans biais de $p(1-p)$, on subodore que

$$P(|\frac{k - np}{\sqrt{n\frac{n}{n-1}f(1-f)}}| \le a)$$

est correctement approché par $P(|Z| \le a)$. Autrement dit

$$P(|\frac{f - p}{\sqrt{\frac{1}{n-1}f(1-f)}}| \le a)$$

est correctement approché par $P(|Z| \leq a)$.

Ces approximations sont raisonnables lorsque n est grand à partir de 30 et que $n \min(f, 1-f) \geq 5$.

Application Sur 50 patients observés par une équipe de médecins, 14 sont atteints de la maladie A. On veut donner un intervalle pour la proportion de malades qui soit fiable à 90%.

Notons Φ la fonction de répartition d'une loi normale centrée réduite : on a $P(|Z| \leq a) = \Phi(a) - \Phi(-a) = \Phi(a) - (1 - \phi(a)) = 2\Phi(a) - 1$.
Pour avoir $P(|Z| \leq a) \geq 0.9$, il suffit d'avoir $\Phi(a) \geq 0.95$.
On voit sur la table que $a = 1.65$ convient.
On a $f = 14/50 = 0.28$. $50 > 30$ et $n \min(f, 1-f) = nf = 14 \geq 5$. Ainsi on peut considérer que

$$P(|\frac{0.28 - p}{\sqrt{\frac{1}{50-1} 0.28(1-0.28)}}| \leq 1.65) \geq 90\%.$$

Autrement dit

$$P(|0.28 - p| \leq 0.106) \geq 90\%.$$

Ou encore

$$P(p \in [0.17; 0.39]) \geq 90\%.$$

Fonction de répartition de la loi normale centrée réduite

La table ci-dessous donne la valeur, pour $x \geq 0$, de

$$\Phi(x) = P(Z \leq x)$$

où Z est une variable aléatoire normale centrée réduite.

Lorsque $x \leq 0$,on peut calculer $\Phi(x)$ grâce à la relation :
$\phi(-x) = 1 - \phi(x)$.

	0	0.01	0.02	0.03	0.04	0.05	0.06	0.07	0.08	0.09
0	0.5	0. 504	0.508	0.512	0.516	0.5199	0.5239	0.5279	0.5319	0.5359
0.1	0 .5398	0.5438	0.5478	0.5517	0.5557	0.5596	0.5636	0.5675	0.5714	0.5 753
0.2	0.5793	0.5832	0.5871	0.591	0.5948	0.5987	0.6026	0.6064	0.6103	0.6141
0.3	0.6179	0.6217	0.6255	0.6293	0.6331	0.6368	0 .6406	0.6443	0.648	0.6517
0.4	0.6554	0.6591	0.6628	0.6664	0.67	0.6736	0.6772	0.6808	0.6844	0.6879
0.5	0.6915	0.695	0.6985	0. 7019	0.7054	0.7088	0.7123	0.7157	0.719	0.7224
0.6	0.7257	0.729 1	0.7324	0.7357	0.7389	0.7422	0.7454	0.7486	0.7517	0.7549
0.7	0.758	0.7611	0.7642	0.7673	0.7704	0.7734	0.7764	0.7794	0.7823	0.7 852
0.8	0.7881	0.791	0.7939	0.7967	0.7995	0.8023	0.8051	0.8078	0.8106	0.8133
0.9	0.8159	0.8186	0.8212	0.8238	0.8264	0.8289	0 .8315	0.834	0.8365	0.8389
1	0.8413	0.8438	0.8461	0.8485	0.8508	0.8531	0.8554	0.8577	0.8599	0.8621
1.1	0.8643	0.8665	0.8686	0 .8708	0.8729	0.8749	0.877	0.879	0.881	0.883
1.2	0.8849	0.8869	0.8888	0.8907	0.8925	0.8944	0.8962	0.898	0.8997	0.9015
1.3	0.9 032	0.9049	0.9066	0.9082	0.9099	0.9115	0.9131	0.9147	0.9162	0.917 7
1.4	0.9192	0.9207	0.9222	0.9236	0.9251	0.9265	0.9279	0.9292	0.9306	0.9319
1.5	0.9332	0.9345	0.9357	0.937	0.9382	0.9394	0.9 406	0.9418	0.9429	0.9441
1.6	0.9452	0.9463	0.9474	0.9484	0.949 5	0.9505	0.9515	0.9525	0.9535	0.9545
1.7	0.9554	0.9564	0.9573	0.9582	0.9591	0.9599	0.9608	0.9616	0.9625	0.9633
1.8	0.9641	0. 9649	0.9656	0.9664	0.9671	0.9678	0.9686	0.9693	0.9699	0.9706
1 .9	0.9713	0.9719	0.9726	0.9732	0.9738	0.9744	0.975	0.9756	0.9761	0.9767
2	0.9772	0.9778	0.9783	0.9788	0.9793	0.9798	0.9803	0.98 08	0.9812	0.9817
2.1	0.9821	0.9826	0.983	0.9834	0.9838	0.9842	0.9846	0.985	0.9854	0.9857
2.2	0.9861	0.9864	0.9868	0.9871	0.9 875	0.9878	0.9881	0.9884	0.9887	0.989
2.3	0.9893	0.9896	0.9898	0.9901	0.9904	0.9906	0.9909	0.9911	0.9913	0.9916
2.4	0.9918	0 .992	0.9922	0.9925	0.9927	0.9929	0.9931	0.9932	0.9934	0.9936
2 .5	0.9938	0.994	0.9941	0.9943	0.9945	0.9946	0.9948	0.9949	0.9951	0.9952
2.6	0.9953	0.9955	0.9956	0.9957	0.9959	0.996	0.9961	0.9 962	0.9963	0.9964
2.7	0.9965	0.9966	0.9967	0.9968	0.9969	0.997	0.9971	0.9972	0.9973	0.9974
2.8	0.9974	0.9975	0.9976	0.9977	0 .9977	0.9978	0.9979	0.9979	0.998	0.9981
2.9	0.9981	0.9982	0.99 82	0.9983	0.9984	0.9984	0.9985	0.9985	0.9986	0.9986

Index