Méthodes Statistiques

M1 MIAGE
2019/20

Notes de cours

Shuyan LIU

Shuyan.Liu@univ-paris1.fr
http://samm.univ-paris1.fr/-Shuyan-LIU

1er octobre 2019

Table des matières

Chapitre 1

Introduction

Faire de la statistique suppose que l'on étudie un ensemble d'objets équivalents sur lesquels on observe des caractéristiques appelées "variables". Ainsi en contrôle de fabrication on prélèvera un ensemble de pièces dans une production homogène et on mesurera leur poids, leur diamètres, etc.

La notion fondamentale en statistique est celle de groupe ou d'ensemble d'objets équivalents que l'on appelle **population**. Ce terme hérité des premières applications de la statistique qui concernaient la démographie est employé pour désigner toute collection d'objets à étudier ayant des propriétés communes. Ces objets sont appelés des **individus** ou unités statistiques.

Généralement la population à étudier est trop vaste pour pouvoir être observée exhaustivement : c'est évidemment le cas lorsque la population est infinie, par exemple l'ensemble de toutes les pièces métalliques que pourrait sortir une machine dans des conditions de fabrication déterminées, mais c'est aussi le cas lorsque les observations sont coûteuses (contrôle destructif entre autres). Lorsque l'on n'observe qu'une partie de la population on parle de **sondage**, la partie étudiée s'appelant l'**échantillon**.

Chaque individu d'une population est décrit par un ensemble de caractéristiques appelées **variables** ou caractères. Ces variables peuvent être classées selon leur nature : variables **quantitatives** ou numériques, par exemple taille, poids, volume, s'expriment par des nombres réels sur lesquels les opérations arithmétiques courantes (somme, moyenne, ...) ont un sens. Certaines peuvent être **discrètes** (nombre fini ou dénombrable de valeurs) : nombre de défauts d'une pièce, de véhicules passant en une heure à un péage, etc. ou **continues** si toutes les valeurs d'un intervalle de $\mathbb{R}$ sont acceptables. Variables **qualitatives** s'exprimant par l'appartenance à une **catégorie** ou **modalité** d'un ensemble fini, par exemple type de traitement thermique subi par un alliage, catégorie socio-professionnelle d'un actif (ouvrier, cadre, employé ...).

Après le recueil des données, la démarche statistique consiste à traiter et interpréter les informations recueillies. Elle comporte deux grands aspects :

l'aspect descriptif ou exploratoire et l'aspect inférentiel ou décisionnel.

La statistique exploratoire Son but est de synthétiser, résumer, structurer l'information contenue dans les données. Elle utilise pour cela des représentations des données sous forme de tableaux, de graphiques, d'indicateurs numériques.

Connue sous le nom de **statistique descriptive** cette phase s'est enrichie ces dernières années de nombreuses techniques de visualisation de données multidimensionnelles (n individus, p variables) c'est l'**analyse des données** qui est une des branches les plus vivantes de la discipline statistique.

Le rôle de la statistique exploratoire est de mettre en évidence des propriétés de l'échantillon et de suggérer des hypothèses. Les modèles probabilistes ne jouent ici qu'un rôle très restreint voire même nul.

La statistique inférentielle Son but est d'étendre les propriétés constatées sur l'échantillon à la population toute entière et de valider ou d'infirmer des hypothèses a priori ou formulées après une phase exploratoire. Le calcul des probabilités joue souvent un rôle fondamental.

Ce document est un extrait du livre [2] et du polycopié [1].

Chapitre 2

Le modèle probabiliste

2.1 Expérience aléatoire et événements

Une **expérience** est qualifiée d'**aléatoire** si on ne peut prévoir par avance son résultat et si, répétée dans des conditions identiques, elle peut donner lieu à des résultats différents. On représente le résultat de cette expérience comme un élément ω de l'ensemble Ω de tous les résultats possibles : Ω **est appelé l'ensemble fondamental** ou encore l'univers des possibles. Par exemple à l'expérience aléatoire qui consiste à lancer deux dés, on peut associer l'ensemble $\Omega = \{(1,1),(1,2),(1,3),\ldots,(6,6)\}$ à 36 éléments. L'obtention de deux fois 6 peut s'écrire $\omega = (6,6)$.

Un **événement** est une assertion ou proposition logique relative au résultat de l'expérience. C'est un sous-ensemble de Ω, c.a.d. un sous-ensemble de tous les résultats possibles. Par exemple, la somme des points supérieure ou égale à 10 est l'ensemble de résultats suivants :

$$\{(4,6),(5,5),(5,6),(6,4),(6,5),(6,6)\}.$$

Si A est un ensemble, la notation $\mathcal{P}(A)$ désigne l'**ensemble des parties** de A, c.a.d. l'ensemble dont les éléments sont tous les sous-ensembles possibles de A (ce qui inclut notamment $\emptyset$ et A). Un événement est donc un élément de $\mathcal{P}(\Omega)$. On dira qu'un événement est réalisé ou non suivant que la proposition est vraie ou fausse une fois l'expérience accomplie. Désormais nous identifierons un événement à la partie de Ω pour laquelle cet événement est réalisé. On appelle **événement élémentaire** une partie de Ω réduite à un seul élément. Un événement élémentaire est donc de la forme $\{\omega\}$ pour tout $\omega \in \Omega$.

2.2 Vocabulaire probabiliste

Bien que l'univers d'une expérience et les événements associés soient décrits grâce aux concepts de la théorie des ensembles, la théorie des probabilités utilise un vocabulaire différent détaillé ci-dessous (ce vocabulaire est en partie partagé avec la logique mathématique).

Vocabulaire probabiliste	Vocabulaire ensembliste	Notation
événement de l'univers Ω	sous-ensemble de Ω	$A \subset \Omega$
événement **impossible**	ensemble vide	$\emptyset$
événement **certain**	ensemble plein Ω	Ω
événement **contraire** de A	complémentaire de A dans Ω	$\overline{A}$
A et B sont **incompatibles**	A et B sont disjoints	$A \cap B = \emptyset$
A **implique** B	A est inclus dans B	$A \subset B$
A **et** B	intersection de A et B	$A \cap B$
A **ou** B	union de A et B	$A \cup B$

Donnons encore une définition utile :

Système complet d'événements. $A_1, A_2, \ldots, A_n$ forment un système complet d'événements si les parties $A_1, A_2, \ldots, A_n$ de Ω constituent une partition de Ω :

$$\begin{cases} \forall i \neq j \qquad A_i \cap A_j = \emptyset, \\ \cup A_i = \Omega. \end{cases}$$

Rappels de propriétés des opérations ensemblistes

commutativité $A \cap B = B \cap A$,
$A \cup B = B \cup A$.

associativité $(A \cap B) \cap C = A \cap (B \cap C)$,
$(A \cup B) \cup C = A \cup (B \cup C)$.

élément neutre $\emptyset \cup A = A$,
$\Omega \cap A = A$.

involution $\overline{\overline{A}} = A$.

distributivité $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$,
$A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$.

dualité $\overline{A \cup B} = \overline{A} \cap \overline{B}$,
$\overline{A \cap B} = \overline{A} \cup \overline{B}$.

2.3 Probabilité

A chaque événement on associe un nombre positif compris entre 0 et 1, sa probabilité. La définition d'une probabilité est la suivante.

On appelle **probabilité** sur Ω (ou mesure/loi de probabilité) une fonction $\mathbb{P}$ de $\mathcal{P}(\Omega)$ dans $[0,1]$ telle que
- $\mathbb{P}(\Omega) = 1$;
- pour tout ensemble dénombrable d'événements incompatibles $A_1, A_2, \ldots, A_n$, on a $\mathbb{P}(\cup A_i) = \sum \mathbb{P}(A_i)$.

Soit une expérience aléatoire, son univers Ω et une probabilité $\mathbb{P}$ sur Ω. Alors $\mathbb{P}$ vérifie les propriétés suivantes :
1. $\mathbb{P}(\emptyset) = 0$.
2. $\mathbb{P}(\overline{A}) = 1 - \mathbb{P}(A)$.
3. $\mathbb{P}(A) \leq \mathbb{P}(B)$ si $A \subset B$.
4. $\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B)$.
5. $\mathbb{P}(\cup A_i) \leq \sum \mathbb{P}(A_i)$.
6. Si $A_i \downarrow \emptyset$, alors $\lim \mathbb{P}(A_i) = 0$ (continuité monotone séquentielle).
7. Théorème des probabilités totales :
   Soit B_i un système complet d'événements alors $\forall A : \mathbb{P}(A) = \sum \mathbb{P}(A \cap B_i)$.

Exercice 2.3.1. *Une certaine expérience aléatoire est modélisée par l'ensemble fondamental Ω et la probabilité $\mathbb{P}$. On considère deux événements A et B tels que $\mathbb{P}(A) = 0.8$, $\mathbb{P}(B) = 0.4$ et $\mathbb{P}(A \cap B) = 0.2$. Calculer les probabilités de $A \cup B$, $\overline{A}$, $\overline{B}$, $\overline{A} \cap B$, $A \cup \overline{B}$, $\overline{A} \cup \overline{B}$, $\overline{A} \cap \overline{B}$, $\overline{A} \cup \Omega$.*

La définition d'une mesure de probabilité peut sembler complexe car on doit a priori définir une fonction sur l'ensemble de tous les sous-ensembles de l'univers considéré. Dans le cas où l'univers est un ensemble fini, la situation est beaucoup plus simple. Se donner une probabilité sur un univers fini à n éléments revient à se donner n nombres réels de $[0,1]$ correspondant aux probabilités de n événements élémentaires de l'univers. La somme des n réels doit être 1.

La situation se simplifie encore s'il est naturel de faire une hypothèse de symétrie sur l'expérience aléatoire. Considérons en effet le lancer d'un dé non truqué. Par symétrie matérielle de l'objet, on s'attend à tomber avec autant de chance sur chacun des faces. En termes de fréquences, on s'attend, lors de lancers répétés, à obtenir environ autant de fois chacune des six faces. Ceci se traduit naturellement en supposant que les probabilités de chacun des événements élémentaires sont égales.

Soit une expérience aléatoire et son univers **fini** Ω. La probabilité **uniforme** sur Ω est celle qui associe à chaque événement élémentaire la même probabilité. Elle est définie par

$$\forall \omega \in \Omega, \quad \mathbb{P}(\{\omega\}) = \frac{1}{|\Omega|}.$$

On parle alors d'**équiprobabilité** pour l'expérience concernée.

Les propriétés classiques des probabilités simplifient grandement le calcul de la probabilité d'un événement dans le cas d'équiprobabilité.

Soit une expérience aléatoire et son univers **fini** Ω, muni de la probabilité $\mathbb{P}$ uniforme. Pour tout événement $A \subset \Omega$, on a

$$\mathbb{P}(A) = \frac{|A|}{|\Omega|}.$$

Le calcul de la probabilité d'un événement dans le cas d'équiprobabilité se ramène donc au calcul du cardinal d'un ensemble, c.a.d. à du dénombrement. Un cas particulier important d'expérience aléatoire illustre parfaitement les liens entre dénombrement et probabilité uniforme. Considérons l'expérience suivante : on place n jetons numéroté de 1 à n dans une urne, puis on choisit au hasard p jetons dans cette urne. On note $J = \{1, \ldots, n\}$ l'ensemble des jetons. En fonction du mode de tirage, l'expérience est modélisée par des univers différents.

Tirages successifs avec remise : il s'agit du mode de tirage qui conduit au modèle le plus simple. On choisit les jetons un par un, en remettant le jeton choisi dans l'urne après chaque tirage. On obtient donc un p-uplet dont les éléments sont choisis dans J, sans contrainte particulière. On peut notamment tomber plusieurs fois sur le même jeton. Formellement, l'univers de l'expérience est donc $\Omega = J^p$ et on a donc $|J^p| = |J|^p = n^p$. L'expérience est caractérisée par la prise en compte de l'ordre (on obtient une liste de p jetons) et par la remise qui autorise à avoir plusieurs fois le même jeton.

Tirages successifs sans remise : dans cette situation, on choisit les jetons un par un, en ne remettant pas les jetons après tirage. On obtient ainsi un p-uplet (une liste de p jetons) dont les éléments sont tous distincts.

Formellement, l'univers de l'expérience est donc

$$\Omega = \{(j_1, \ldots, j_p) \in J^p \mid \forall (k,l), k \neq l \Rightarrow j_k \neq j_l\}.$$

L'ensemble Ω est en fait l'ensemble des arrangements de p éléments choisis parmi n et donc $|\Omega| = A_n^p$. Les deux aspects importants de l'expérience qui conduisent à ce résultat sont la prise en compte de l'ordre des tirage (et donc l'obtention d'un p-uplet) et de l'absence de remise dans l'urne (et donc le fait que les éléments du p-uplet sont distincts).

Tirage simultané : le dernier mode classique de tirage consiste à prendre en une seule fois un paquet de p jetons. Comme dans le mode précédent, chaque tirage ne peut contenir qu'une seule fois un jeton. Cependant, on ne peut pas déduire du tirage un ordre sur les p jetons choisis. L'univers de l'expérience est donc le suivant

$$\Omega = \{K \subset J \mid |K| = p\} = \{\{j_1, \ldots, j_p\} \subset J \mid \forall (k,l), k \neq l \Rightarrow j_k \neq j_l\}.$$

L'ensemble Ω est donc l'ensemble des combinaisons de p éléments choisis parmi n et donc $|\Omega| = C_n^p$. Ce mode de tirage est caractérisé par son aspect simultané qui conduit à la fois à l'absence d'ordre et de répétition.

Rappels sur le dénombrement

1. Constructions d'objets complexes à partir des éléments d'un ensemble :
   (a) Une liste est une collection *ordonnée* d'éléments d'un ensemble (attention, ordonnée veut dire qu'on a choisi un ordre, pas qu'on range les éléments du plus petit au plus grand) ;
   (b) Une p-liste d'un ensemble A est une liste de p éléments de A. Si $p = 2$, on parle de couple ;
   (c) Un arrangement de p éléments de A est une p-liste d'éléments *distincts* de A. Si $p = n$ on parle de permutation ou d'ordre des éléments de A ;
   (d) Un n-uplet sur les ensembles $A_1, A_2, \ldots, A_n$ est une liste $(a_1, a_2, \ldots, a_n)$ telle que pour tout $i \in \{1, 2, \ldots, n\}, a_i \in A_i$.
   L'ensemble des n-uplets formés à partir de $A_1, A_2, \ldots, A_n$ est le produit cartésien $A_1 \times A_2 \times \ldots \times A_n$;
   (e) Les combinaisons de p éléments d'un ensemble A sont les parties à p éléments de A, ou autrement dit, les sous-ensembles de A cardinal p.
2. Factorielle : $n! = n \times (n-1) \times (n-2) \times \ldots \times 3 \times 2 \times 1 = \prod_{i=1}^{n} i$. On pose $0! = 1$.
3. Nombre de p-listes d'un ensemble à n éléments : n^p.

4. Nombre d'arrangements de p éléments (ou p-listes d'éléments distincts) d'un ensemble à n éléments : $A_n^p = \frac{n!}{(n-p)!}$.
5. Nombre de permutations (d'ordres) d'un ensemble à n éléments : $n!$.
6. Nombre de combinaisons de p éléments : $C_n^p = \frac{n!}{p!(n-p)!} = \frac{A_n^p}{p!}$.

Exercice 2.3.2. *Dans une course de chevaux, il y a* 15 *participants au départ. Combien y a-t-il d'ordres d'arrivée possibles ? Combien y a-t-il d'ordres d'arrivée possibles avec le cheval a finissant premier, le cheval b deuxième et le cheval c troisième ? Quelle est la proportion d'ordres d'arrivée possibles avec comme trio de tête a, b, c (dans cet ordre) par rapport au total d'ordres arrivées possibles ? Mêmes questions pour n chevaux en fixant l'ordre des k premiers.*

Exercice 2.3.3. *Reprenons l'exemple de la course de chevaux avec* 15 *participants au départ. Combien y a-t-il de tiercés (trios) gagnants dans l'ordre possibles ? Combien y a-t-il de quintés (les* 5 *premiers) gagnants dans l'ordre possible ? Mêmes questions pour n participants.*

Exercice 2.3.4. *Le programme d'une épreuve d'examen comporte* 30 *sujets. Chaque candidat tire trois sujets au hasard. Quelle est la probabilité, n'ayant étudié que le tiers des sujets,*

1. *d'avoir préparé les trois sujets ?*
2. *d'avoir préparé exactement deux des sujets ?*
3. *de n'en avoir préparé aucun ?*
4. *d'en avoir préparé au moins un ?*
   (Remarque : $C_{30}^3 = C_{10}^3 + C_{10}^2 C_{20}^1 + C_{10}^1 C_{20}^2 + C_{20}^3$*)*
5. *d'en avoir préparé au moins deux ?*
   (Indication : $C_{10}^2 C_{20}^1 + C_{10}^3$ *?* $C_{10}^2 C_{30}^1$ *?* $C_{10}^2 C_{28}^1$ *?)*

Exercice 2.3.5. *On étudie une urne contenant 6 pièces bleues, 7 pièces vertes et 5 pièces jaunes. Attention, dans cet exercice les questions sont indépendantes : on suppose qu'avant chaque tirage (au début d'une question), l'urne contient les pièces indiquées et on ne tient pas compte des questions précédentes. Les résultats peuvent être donnés sous forme fractionnaire avec des multiplications sans avoir simplifiées. Mais les notations* A_n^k, C_n^k, $\binom{k}{n}$ *et* $n!$ **ne sont pas autorisées**.

1. *On tire 2 pièces dans l'urne simultanément. Quelle est la probabilité d'obtenir au moins une pièce jaune ?*
2. *On tire 2 pièces dans l'urne successivement et avec remise. Quelle est la probabilité d'obtenir exactement une pièce bleue ?*

2.4 Lois de probabilités conditionelles, indépendance

Supposons que l'on s'intéresse à la réalisation d'un événement A, tout en sachant qu'un événement B est réalisé. Si A et B sont incompatibles la question est tranchée : A ne se réalisera pas, mais si $A \cap B \neq \emptyset$, il est possible que A se réalise ; cependant, l'univers des possibles n'est plus Ω tout entier, mais est restreint à B ; en fait, seule nous intéresse la réalisation de A à l'intérieur de B, c.a.d. $A \cap B$ par rapport à B. Ceci justifie la définition suivante :

> Soit B un événement de probabilité non nulle. On appelle **probabilité conditionnelle** de A sachant B le rapport noté $\mathbb{P}(A|B)$:
>
> $$\mathbb{P}(A|B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)}.$$

Dans certaines situations, avoir une information sur une partie d'une expérience aléatoire sous la forme d'un événement réalisé n'entraîne pas de révision de la probabilité d'un autre événement, comme le montre l'exemple simple suivant.

On lance deux dés à six faces non truqués, un dé rouge et un dé noir. Soit l'événement A "les deux dés donnent des résultats identiques" et l'événement B "le dé rouge donne 1". On étudie $\mathbb{P}(A)$ et $\mathbb{P}(A|B)$. L'univers est

$$\Omega = \{1, 2, 3, 4, 5, 6\}^2.$$

L'événement A est la diagonale de Ω, soit

$$A = \{(1, 1), (2, 2), (3, 3), (4, 4), (5, 5), (6, 6)\}.$$

L'événement B s'écrit simplement

$$B = \{1\} \times \{1, 2, 3, 4, 5, 6\}.$$

On a donc $\mathbb{P}(A|B) = \mathbb{P}(A) = \frac{1}{6}$ ce qui montre que savoir que B est réalisé n'apporte pas de connaissance sur A. Dans une telle situation, on parle d'indépendance entre les deux événements, selon la définition suivante.

> A est **indépendant** de B si $\mathbb{P}(A|B) = \mathbb{P}(A)$.

Une autre définition est la suivante.

A et B sont **indépendants** si et seulement si

$$\mathbb{P}(A \cap B) = \mathbb{P}(A) \times \mathbb{P}(B).$$

On note alors $A \perp\!\!\!\perp B$.

Présentons maintenant les formules de Bayes qui ont pour but d'exprimer $\mathbb{P}(A|B)$ en fonction de $\mathbb{P}(B|A)$.

Première formule de Bayes :

$$\mathbb{P}(B|A) = \frac{\mathbb{P}(A|B)\mathbb{P}(B)}{\mathbb{P}(A)}.$$

Soit B_i un système complet d'événement. On peut écrire : $\mathbb{P}(A \cap B_i) = \mathbb{P}(A|B_i)\mathbb{P}(B_i)$. On obtient donc la **règle des probabilités totales** :

$$\mathbb{P}(A) = \sum_i \mathbb{P}(A|B_i)\mathbb{P}(B_i).$$

On en déduit alors la **deuxième formule de Bayes** :

$$\mathbb{P}(B_i|A) = \frac{\mathbb{P}(A|B_i)\mathbb{P}(B_i)}{\sum\limits_k \mathbb{P}(A|B_k)\mathbb{P}(B_k)}.$$

Exercice 2.4.1. *Reprenons l'exercice 2.3.1. Calculer les probabilités suivantes* $\mathbb{P}(A \mid B)$, $\mathbb{P}(B \mid A)$, $\mathbb{P}(B \mid \overline{A})$, $\mathbb{P}(\overline{B} \mid A)$, $\mathbb{P}(A \mid \Omega)$.

Exercice 2.4.2. *Un questionnaire à choix multiples propose quatre réponses pour chaque question. Soit* $p = 0.3$ *la probabilité qu'une étudiante connaisse la réponse à une question donnée. Si l'étudiante ignore la réponse, elle choisit au hasard l'une des réponses proposées. Quel est pour le correcteur la probabilité qu'une étudiante connaisse vraiment la bonne réponse lorsqu'elle l'a donnée ?*

Exercice 2.4.3. *Une compagnie d'assurances répartit ses clients en trois classes de risque :* R_1 *(risque faible),* R_2 *(risque moyen) et* R_3 *(risque élevé). Les effectifs de chaque classe représentent* 20%, 50% *et* 30% *de la population totale. Les probabilités d'avoir un accident dans l'année pour chacune des classes sont* 0.05, 0.15 *et* 0.30.

1. *Quelle est la probabilité pour qu'une personne choisie au hasard ait un accident dans l'année ?*
2. *Si M. Martin n'a pas eu d'accident cette année, quelle est la probabilité qu'il soit dans la classe de risque faible* R_1 *?*

Chapitre 3

Variables aléatoires

3.1 Loi de probabilité et moments d'une variable aléatoire réelle

Le concept de v.a. formalise la notion de grandeur variant selon le résultat d'une expérience aléatoire.

Considérons le lancer de deux dés parfaitement équilibrés : cette expérience se traduit par l'ensemble Ω de tous les couples de chiffres de 1 à 6 :

$$\Omega = \{(1,1),(1,2),(1,3),\ldots,(6,6)\}$$

muni de la loi de probabilité $\mathbb{P}$ telle que $\mathbb{P}(\omega) = \frac{1}{36}$, $\forall \omega \in \Omega$.

Intéressons nous à la somme des points marqués par les deux dés. On définit ainsi une application S de Ω dans l'ensemble $E = \{2,3,\ldots,12\}$.

Pour obtenir la probabilité d'une valeur quelconque de S il suffit de dénombrer les ω qui réalisent cette valeur. Ainsi :

$$\mathbb{P}(S=5) = \mathbb{P}(\{(1,4),(2,3),(3,2),(4,1)\}) = \frac{4}{36}$$

et généralement $\mathbb{P}(S=s) = \mathbb{P}(\{S^{-1}(s)\})$.

On voit que pour définir la loi de probabilité sur S on transporte la loi de probabilité de Ω sur E par l'application S.

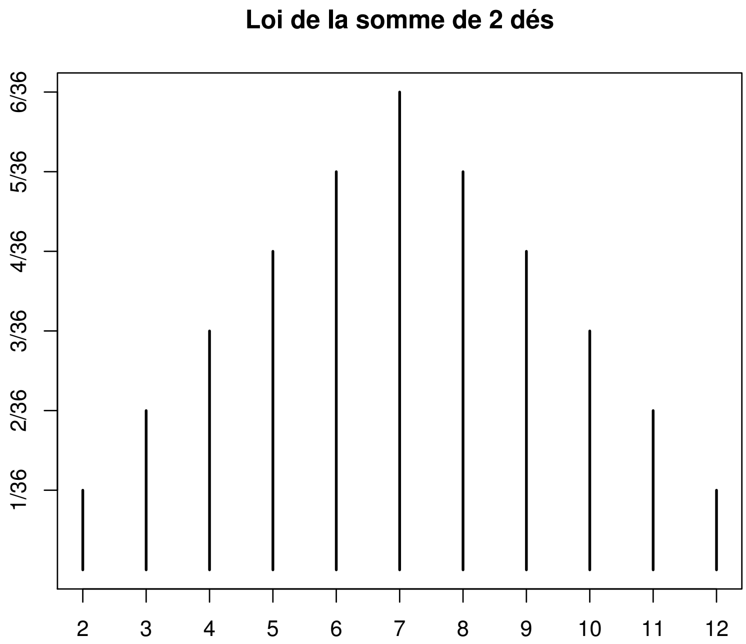
> Soit une expérience aléatoire d'univers Ω et de probabilité associée $\mathbb{P}$. Soit un ensemble quelconque E. Une **variable aléatoire** sur $(\Omega,\mathbb{P})$, X, est une fonction de Ω dans E. X est dite **à valeurs dans** E.

Soit X une v.a. sur $(\Omega,\mathbb{P})$ et à valeurs dans E. On définit une probabilité sur E, **la loi de** X, notée $\mathbb{P}_X$, par

$$\forall A \subset E, \quad \mathbb{P}_X(A) = \mathbb{P}(\{\omega \in \Omega \,|\, X(\omega) \in A\}) = \mathbb{P}(X^{-1}(A)).$$

On note aussi
$$\mathbb{P}_X(A) = \mathbb{P}(X \in A),$$
et pour tout $x \in E$,
$$\mathbb{P}_X(\{x\}) = \mathbb{P}(X = x).$$

Pour une variable discrète, c.a.d. une variable ne pouvant prendre qu'un nombre fini (ou dénombrable) de valeurs $x_1, x_2, \ldots, x_n$ la loi $\mathbb{P}_X$ est constituée de masses ponctuelles. $\mathbb{P}_X$ peut alors être représentée par un diagramme en bâtons. Ainsi pour l'exemple du lancer de deux dés on a la figure ci dessous.



Exercice 3.1.1. *Dans cet exercice on va découvrir* `R` *en générant des entiers aléatoires qui représentent la somme de deux dés avec les commandes ci-dessous.*

```
n = 100
d1 = sample(1:6, n, replace = T)
d2 = sample(1:6, n, replace = T)
s = d1+d2; tab.s = table(s); plot(tab.s)
```

1. *Exécuter ces commandes. Quelles sont les valeurs de* `d1`*,* `d2`*,* `s` *et* `tab.s` *? Qu'obtenez vous en tapant la commande* `sum(tab.s)` *?*

2. *Que font les fonctions* `sample()` *et* `table()` *?*
3. *Lancer ces commandes plusieurs fois. Les graphiques sont ils identiques ? Pourquoi ?*
4. *Augmenter* `n` *jusqu'à* `1000`*, puis à* `10000`*. Pour chaque valeur de* `n`*, lancer plusieurs fois ces commandes et observer les changements des graphiques. Commenter ces changements.*
5. *Créer un vecteur* `v` *de longueur* 11 *qui représente la loi de la somme de deux dés en utilisant* `v = c(1:6,5:1)/36`*. Qu'observez vous en tapant* `rbind(tab.s/n, v)` *? Commenter le résultat.*
6. *Pouvez vous prédire les valeurs de* `tab.s[1]` *et* `tab.s[6]` *?*

Exercice 3.1.2. *Soit l'ensemble Ω des entiers compris (au sens large) entre* 1 *et* 11*. On munit Ω de la mesure de probabilité $\mathbb{P}$ uniforme. Soit X la fonction de Ω dans {pair, impair} la fonction qui à un entier de Ω associe sa parité. Déterminer la loi de X.*

Exercice 3.1.3. *Soit une v.a. X à valeurs dans $\{a, b, c, d\}$. On suppose que $\mathbb{P}(X = a) = \frac{1}{3}$ et que $\mathbb{P}(X \in \{b, c\}) = \frac{1}{3}$.*

1. *Calculer $\mathbb{P}(X = d)$.*
2. *On suppose maintenant que $\mathbb{P}(X \in \{c, d\}) = \frac{4}{9}$. Calculer $\mathbb{P}(X = b)$ et $\mathbb{P}(X = c)$.*

Exercice 3.1.4. *On suppose que le nombre de livres achetés par un client dans une librairie est une v.a. notée X. On connaît :*

$$\mathbb{P}(0 \leq X \leq 1) = 0.6,\ \mathbb{P}(1 \leq X \leq 2) = 0.5,\ \mathbb{P}(0 \leq X < 3) = 0.8,$$

$$\mathbb{P}(X = 3) = \mathbb{P}(X \geq 4) \ \ et\ \ \mathbb{P}(X = 4) = \mathbb{P}(X \geq 5).$$

Calculer $\mathbb{P}(X = i)$ pour $i \in \{0, 1, 2, 3, 4\}$.

La fonction de répartition d'une v.a. X est l'application F de $\mathbb{R}$ dans $[0, 1]$ définie par :

$$F(x) = \mathbb{P}(X \leq x).$$

La fonction F vérifie les quatre propriétés fondamentales suivantes :

1. $\lim_{x\to-\infty} F(x) = 0$;
2. $\lim_{x\to+\infty} F(x) = 1$;
3. F est croissante : $s \leq t \Rightarrow F(s) \leq F(t)$;
4. F est continue à droite en tout point : $\forall x \in \mathbb{R},\ \lim\limits_{h\to0^+} F(x+h) = F(x)$.

L'importance pratique de la f.d.r. vient de ce qu'elle permet de calculer la probabilité de tout intervalle de $\mathbb{R}$:

$$\mathbb{P}(a < X \leq b) = F(b) - F(a).$$

La notion de **variable continue** se confond avec celle de variable admettant une densité de probabilité.

Une loi de probabilité $\mathbb{P}_X$ admet une **densité** f pour tout intervalle I de $\mathbb{R}$ on a : $$\mathbb{P}_X(I) = \int_I f(x)dx.$$

F est alors dérivable et admet f pour dérivée. On a donc :

$$\mathbb{P}(a < X \leq b) = \int_a^b f(x)dx = F(b) - F(a).$$

Une densité f est donc une fonction positive d'intégrale égale à 1 :

$$\int_{\mathbb{R}} f(x)dx = 1.$$

On remarque que pour une variable à densité :

$$\mathbb{P}(X = x) = 0, \ \ \forall x.$$

Soient X et Y deux v.a. réelles définies sur le même espace probabilisé. **X et Y sont indépendantes** si on a

$$\mathbb{P}(X \in A \text{ et } Y \in B) = \mathbb{P}(X \in A)\mathbb{P}(Y \in B), \ \ \forall A, B \subset \mathbb{R}.$$

Les v.a. discrètes X et Y sont indépendantes si et seulement si $$\forall x \in X(\Omega),\ \forall y \in Y(\Omega), \quad \mathbb{P}(X = x, Y = y) = \mathbb{P}(X = x)\mathbb{P}(Y = y).$$

Exercice 3.1.5. *On lance deux dés à six faces non truqués, un dé rouge et un dé noir. On définit deux v.a. X et Y où X représente la différence entre les points obtenus de deux dés et Y représente le résultat du dé rouge. Dans la section 2.4 nous avons étudié l'événement A "les deux dés donnent des résultats identiques" et l'événement B "le dé rouge donne 1".*

1. *Réécrire l'indépendance des événements A et B en utilisant les v.a. X et Y.*
2. *Calculer la loi de X.*
3. *X et Y sont elles indépendantes ?*

Une loi de probabilité peut être caractérisée par certaines valeurs typiques associées aux notions de valeur centrale, de dispersion et de forme de la distribution.

Pour une variable discrète on définit l'**espérance** $\mathbb{E}(X)$ par la formule :

$$\mathbb{E}(X) = \sum_i x_i \mathbb{P}(X = x_i)$$

(si cette expression a un sens). $\mathbb{E}(X)$ est la moyenne arithmétique des différentes valeurs de X pondérées par leurs probabilités.

Pour une variable continue admettant une densité f, l'espérance $\mathbb{E}(X)$ est définie par

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

si l'intégrale converge.

Les propriétés élémentaires de l'espérance mathématique sont celles des intégrales et se déduisent de la linéarité. Si a est une constante :

1. $\mathbb{E}(a) = a$;
2. $\mathbb{E}(aX) = a\mathbb{E}(X)$;
3. $\mathbb{E}(X + a) = \mathbb{E}(X) + a$;
4. $\mathbb{E}(X_1 + X_2) = \mathbb{E}(X_1) + \mathbb{E}(X_2)$.

Une autre propriété importante de l'espérance est la suivante

$$\boxed{X \text{ et } Y \text{ sont indépendantes} \Rightarrow \mathbb{E}(XY) = \mathbb{E}(X)\mathbb{E}(Y).}$$

Attention : La réciproque est fausse et $\mathbb{E}(XY) = \mathbb{E}(X)\mathbb{E}(Y)$ n'entraîne pas en général l'indépendance de X et Y.

On appelle **variance** de X notée $\mathbb{V}(X)$ ou σ^2 la quantité définie par

$$\mathbb{V}(X) = \mathbb{E}((X - m)^2) = \int_{\mathbb{R}} (x - m)^2 f(x) dx$$

où $m = \mathbb{E}(X)$. σ s'appelle l'**écart-type** de X. La variance est donc le moment centré d'ordre 2 de la distribution et est une mesure de la dispersion de X autour de m. Une relation usuelle est

$$\mathbb{V}(X) = \mathbb{E}(X^2) - m^2.$$

Les propriétés de la variance sont les suivantes : si a est une constante,

1. $\mathbb{V}(X+a)=\mathbb{V}(X)$;
2. $\mathbb{V}(aX)=a^2\mathbb{V}(X)$;
3. $\mathbb{V}(X)=0 \Leftrightarrow X=a$ (presque sûrement).

On appelle **covariance** de X et Y la quantité :

$$\mathrm{cov}(X,Y)=\mathbb{E}(XY)-\mathbb{E}(X)\mathbb{E}(Y)=\mathbb{E}((X-\mathbb{E}(X))(Y-\mathbb{E}(Y)))$$

donc :

$$\mathbb{V}(X+Y)=\mathbb{V}(X)+\mathbb{V}(Y)+2\,\mathrm{cov}(X,Y).$$

En particulier

$$\boxed{X \text{ et } Y \text{ sont indépendantes} \Rightarrow \mathbb{V}(X+Y)=\mathbb{V}(X)+\mathbb{V}(Y)}$$

mais la réciproque est ici encore inexacte en général.

Exercice 3.1.6. *Soit X une v.a. discrète réelle de f.d.r. F_X donnée par*

$x \in$	$]-\infty,-5[$	$[-5,5[$	$[5,9[$	$[9,11[$	$[11,\infty[$
$F_X(x)$	0	1/9	1/3	2/3	1

Calculer $\mathbb{P}(X=-5)$, $\mathbb{P}(X\leq \frac{13}{2})$, $\mathbb{P}(X>\frac{3}{2})$ *et* $\mathbb{E}(X)$.

Exercice 3.1.7. *Soit F la fonction définie par*

$$F(t)=\begin{cases} a & si\, t<0 \\ b & si\, 0\leq t<1 \\ 2b+c & si\, 1\leq t<2 \\ d & si\, t\geq 2 \end{cases}$$

1. *A quelle conditions sur a, b, c et d, F est-elle une f.d.r. ?*
2. *On suppose que F est la f.d.r. d'une v.a. X. Calculer les probabilités suivantes :* $\mathbb{P}(X\leq 3)$, $\mathbb{P}(X>1)$ *et* $\mathbb{P}(0.5<X\leq 1)$.
3. *En considérant* $\mathbb{P}(X>1)=0.3$ *et* $\mathbb{P}(0.5<X\leq 1)=0.6$, *calculer a, b, c et d.*
4. *Déterminer la loi de X, son espérance et sa variance.*

Exercice 3.1.8. *Soit*

$$f(x)=\begin{cases} \frac{1}{3} & si\, x\in[0,3] \\ 0 & sinon \end{cases}$$

Vérifier que f est bien la densité d'une v.a., calculer la f.d.r., l'espérance et la variance.

Exercice 3.1.9. *Soit X une v.a. de densité*

$$f(x) = \begin{cases} \frac{2x}{\theta^2} & si\ x \in [0,\theta] \\ 0 & sinon \end{cases}$$

1. *Calculer $\mathbb{E}(X)$, $\mathbb{E}(X^2)$ puis $\mathbb{V}(X)$ en fonction de θ.*
2. *Calculer $\mathbb{P}(X < \frac{\theta}{4})$*

3.2 Lois de probabilité d'usage courant

Nom	Valeurs prises	Loi	Espérance	Variance
Uniforme $X \sim \mathcal{U}_n$	$\{1,\ldots,n\}$	$\mathbb{P}(X=k) = \frac{1}{n}$	$\frac{n+1}{2}$	$\frac{n^2-1}{12}$
Bernoulli $X \sim \mathcal{B}(p)$	$\{0,1\}$	$\mathbb{P}(X=0) = 1-p$ $\mathbb{P}(X=1) = p$	p	$p(1-p)$
Binomiale $X \sim \mathcal{B}(n,p)$	$\{0,\ldots,n\}$	$\mathbb{P}(X=k) = C_n^k p^k (1-p)^{n-k}$	np	$np(1-p)$
Géométrique $X \sim \mathcal{G}(p)$	$\mathbb{N}$	$\mathbb{P}(X=k) = (1-p)^k p$ $F(k) = 1-(1-p)^{k+1}$	$\frac{1-p}{p}$	$\frac{1-p}{p^2}$
Poisson $X \sim \mathcal{P}(\lambda)$	$\mathbb{N}$	$\mathbb{P}(X=k) = \frac{\lambda^k}{k!}e^{-\lambda}$	λ	λ

Loi de Bernoulli $\mathcal{B}(p)$

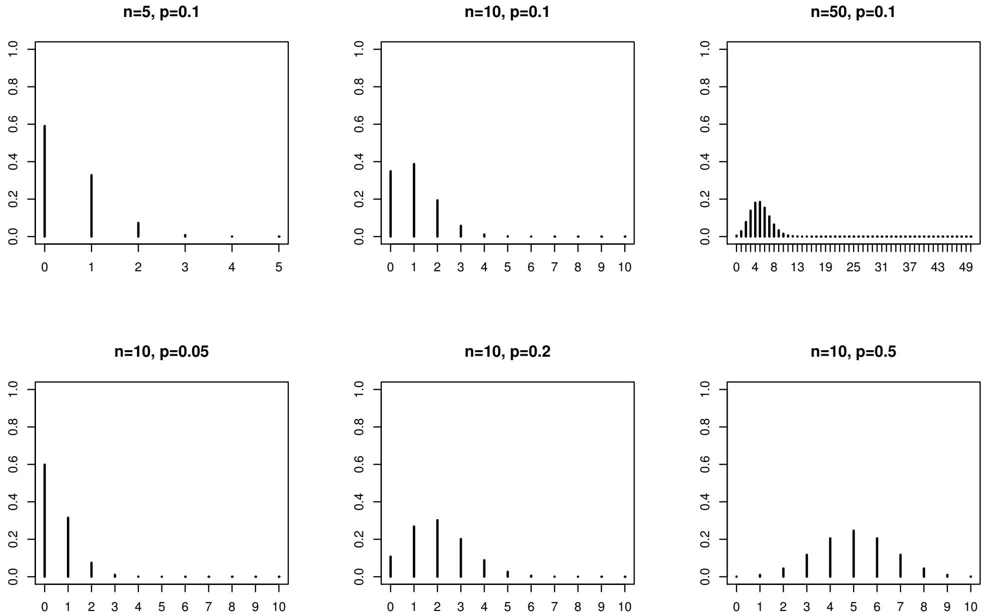
C'est la loi d'une variable X ne pouvant prendre que les deux valeurs 1 ou 0 avec les probabilités p et $1-p$.

Loi binomiale $\mathcal{B}(n,p)$

Supposons que l'on répète n fois dans des conditions identiques une expérience aléatoire, dont l'issue se traduit par l'apparition ou non-apparition d'un événement A de probabilité p, le résultat d'une expérience étant indépendant des résultats précédents ; soit X le nombre d'apparitions de l'événement A parmi ces n expérience ($0 \leq X \leq n$). On dit alors que X suit une **loi binomiale** de paramètres n et p notée $\mathcal{B}(n,p)$. La deuxième définition de la loi binomiale est la suivante.

X suit une **loi binomiale** $\mathcal{B}(n,p)$ si X est une somme de n variables de Bernoulli indépendantes et de même paramètre p.

La figure ci dessous représente quelques diagrammes en bâtons correspondant à diverses valeurs de n et p. On notera que la distribution est symétrique si $p = 1/2$ et le devient approximativement sinon, dès que n est assez élevé. Pour n grand on verra plus loin que la loi binomiale peut être approximée soit par une loi de Poisson (si p est petit) soit par une loi de Gauss.



Loi de Poisson $\mathcal{P}(\lambda)$

On obtient la loi de Poisson comme approximation de la loi binomiale dans le schéma suivant :

Lorsque n est "grand" et np "petit", on peut remplacer la loi binomiale $\mathcal{B}(n, p)$ par la loi de Poisson $\mathcal{P}(\lambda)$ où $\lambda = np$. Cette règle pratique est justifiée par le théorème de convergence de la loi binomiale vers la loi de Poisson : si $(p_n)_{n\geq 1}$ est une suite de réels dans $[0, 1]$ vérifiant

$$np_n \to \lambda \in]0, +\infty[\text{ quand } n \to +\infty$$

alors

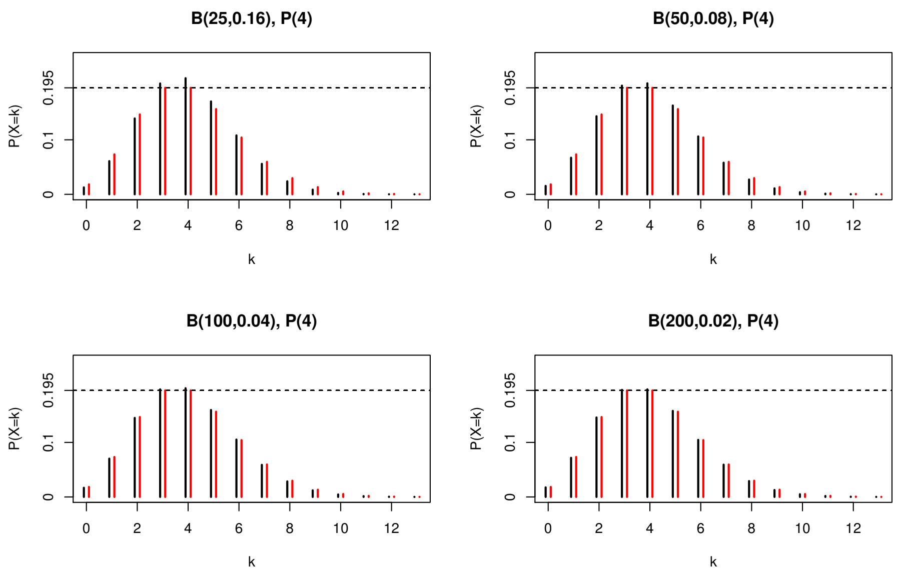
$$\forall k \in \mathbb{N}, \;\; C_n^k p^k (1-p)^{n-k} \to e^{-\lambda}\frac{\lambda^k}{k!}, \;\text{ quand } n \to +\infty.$$

Nous allons, en fait, établir ce résultat sous la forme mathématique suivante.

Théorème
Soit X_n une suite de variables binomiales $\mathcal{B}(n,p)$ telles que $n \to \infty$ et $p \to 0$ de manière à ce que le produit np tende vers une limite finie λ. Alors la suite de v.a. X_n converge en loi vers une variable de Poisson $\mathcal{P}(\lambda)$.

Pour des variables discrètes la convergence en loi vers une variable discrète s'exprime par $\mathbb{P}(X_n = x) \longrightarrow \mathbb{P}(X = x)$. Les notions de convergence seront étudiées en détail dans la section suivante.

Les diagrammes en bâtons ci-dessous représentent la loi binomiale $\mathcal{B}(n,p)$ et la loi de Poisson approximante $\mathcal{P}(\lambda)$ avec $\lambda = np = 4$. Les segments verticaux (les bâtons) du diagramme représentant la loi d'une variable discrète X ont une hauteur égale à $\mathbb{P}(X = k)$ avec une extrémité inférieure au point d'abscisse k de l'axe horizontal. Pour la lisibilité, on a légèrement décalé vers la droite les bâtons de la loi de Poisson (en rouge) et vers la gauche ceux de la loi binomiale (en noir). La dernière graduation verticale donne la valeur de la plus grande probabilité de Poisson. Remarquons que $\mathbb{P}(X = 3) = \mathbb{P}(X = 4)$ pour $X \sim \mathcal{P}(4)$. On constate que pour $n = 200$ (figure en bas à droite), la différence entre les deux diagrammes n'est pratiquement plus discernable visuellement.



Exercice 3.2.1. *On lance dix fois une pièce de monnaie. Quelle est la probabilité d'obtenir "pile" au moins neuf fois ?*

Exercice 3.2.2. *Un magasin reçoit* 3 *réclamation en moyenne par jour ouvré. Supposant poissonnienne la loi de survenance de ces réclamations, calculer la probabilité pour que le premier lundi de septembre soient enregistrées :*

1. *0 réclamation*
2. *2 réclamations*
3. *strictement plus de 4 réclamations*

Exercice 3.2.3. *A part la fonction* `sample()`*, le logiciel* `R` *possède d'autres générateurs des entiers aléatoires pour simuler les v.a. suivant les lois de probabilité usuelles par exemple* `rbinom()`*,* `rgeom()` *et* `rpois()`*.*

1. *On a tapé les commandes ci-dessous. Pouvez vous prédire les valeurs de* `mean(x.b)`*,* `mean(x.g)` *et* `mean(x.p)`*, puis* `var(x.b)`*,* `var(x.g)` *et* `var(x.p)` *?*

```
n = 100; size = 10; prob = 0.6; lambda = 3
x.b = rbinom(n, size, prob); x.g = rgeom(n, prob);
x.p = rpois(n, lambda)
```

2. *Les codes suivants renvoient une matrice de dimension 2×3 dont la première ligne est les moyennes de* `x.b`*,* `x.g` *et* `x.p` *et la deuxième ligne est leur espérance. Qu'observez vous ? Modifier ces codes pour renvoyer une matrice dont la première ligne est les variances de* `x.b`*,* `x.g` *et* `x.p` *et la deuxième ligne est leur variance théorique.*

```
m.b = mean(x.b); m.g = mean(x.g); m.p = mean(x.p)
e.b = size*prob; e.g = (1-prob)/prob; e.p = lambda
v.m = c(m.b, m.g, m.p); v.e = c(e.b, e.g, e.p)
rbind(v.m, v.e)
```

3. *Modifier dans les codes précédents* `rbind` *pour* `cbind` *et observer le changement de résultat.*

Exercice 3.2.4. *Nous utilisons* `dbinom()` *et* `pbinom()` *dans cet exercice pour tracer les courbes de la fonction de densité et la f.d.r. de loi binomiale.*

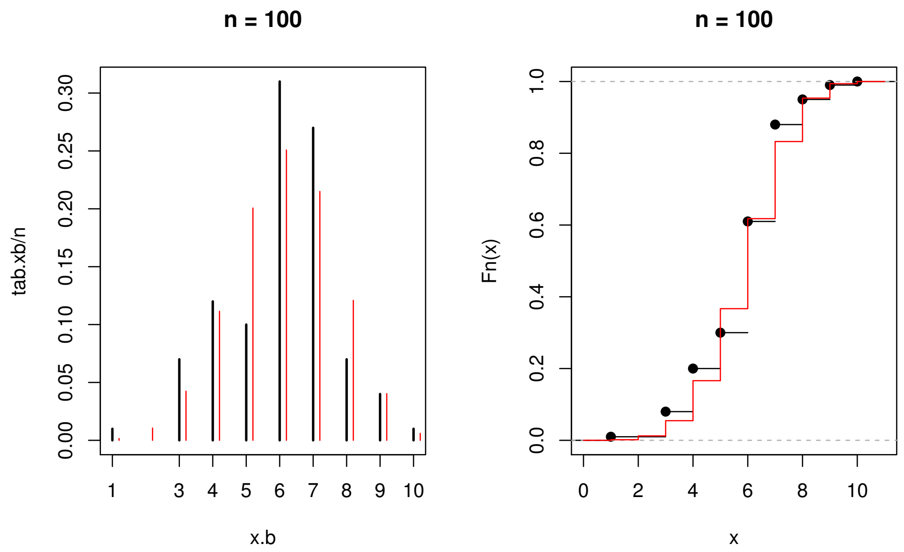
1. *Créer un vecteur* `x.b` *qui contient $n = 100$ entiers aléatoires de loi $\mathcal{B}(10, 0.6)$.*
2. *Calculer les occurrences des entiers apparus avec* `table()`*. Présenter le résultat obtenu par un diagramme en bâtons.*
3. *Nommer les occurrences obtenues dans la question précédente* `tab.xb`*. Lancer les codes suivants. Commenter le graphique obtenu.*

```
plot(tab.xb/n)
p = dbinom(0:10, 10, 0.6); p.x = 0:10+0.2
for(i in 1:(10+1))
lines(rep(p.x[i], 2), c(0, p[i]), col = "red")
```

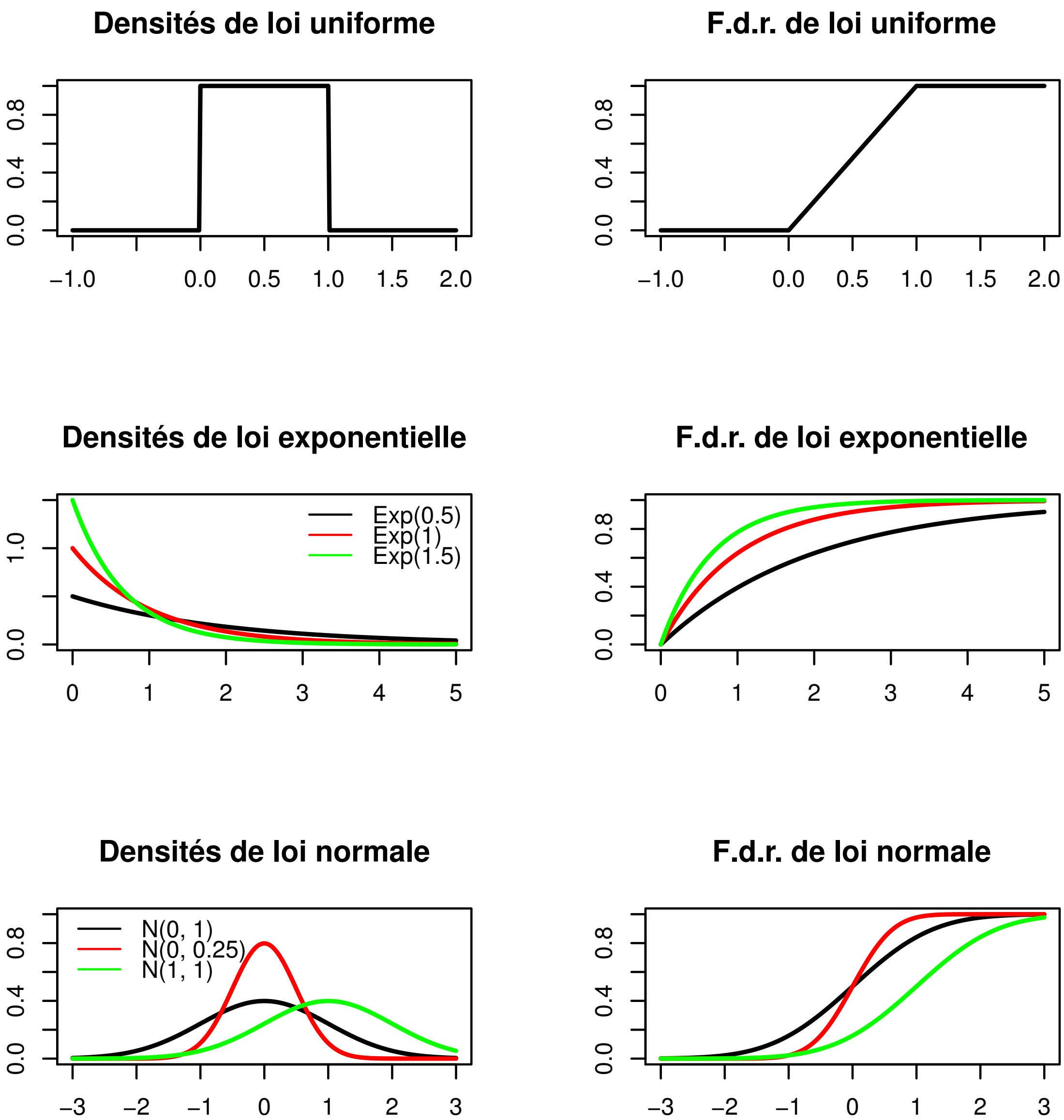
4. *Remplacer dans les codes ci-dessus* `lines` *par* `plot`*, que se passe t il ? Quelle est alors la différence entre ces deux fonctions ?*
5. *Lancer les codes suivants. Commenter le graphique obtenu.*

```
plot.ecdf(x.b); qf = seq(0, 11, 0.1)
fdr = pbinom(qf, 10, 0.6)
lines(qf, fdr, col = "red", type = "s")
```

6. *Augmenter* n *jusqu'à* 1000*, puis à* 10000*. Relancer les codes des questions précédentes et observer les changement.*
7. *Utiliser* `par(mfrow = c(1,2))` *pour mettre deux graphiques obtenus dans les questions 3 et 5 côte à côte. La figure ci-dessous est un exemple pour* $n = 100$*.*



Nom	Valeurs prises	Densité et Fonction de répartition	Espérance	Variance
Uniforme $X \sim \mathcal{U}([a,b])$	$[a,b]$	$f(x) = \begin{cases} \frac{1}{b-a}, & \text{si } x \in [a,b] \\ 0, & \text{sinon} \end{cases}$ $F(x) = \begin{cases} 0 & \text{si } x < a \\ \frac{x-a}{b-a}, & \text{si } x \in [a,b] \\ 1, & \text{si } x \geq b \end{cases}$	$\frac{a+b}{2}$	$\frac{(b-a)^2}{12}$
Exponentielle $X \sim \mathcal{E}(\lambda)$	$\mathbb{R}^+$	$f(x) = \begin{cases} \lambda \exp(-\lambda x), & \text{si } x \geq 0 \\ 0, & \text{sinon} \end{cases}$ $F(x) = \begin{cases} 1 - \exp(-\lambda x), & \text{si } x \geq 0 \\ 0, & \text{sinon} \end{cases}$	$\frac{1}{\lambda}$	$\frac{1}{\lambda^2}$
Normale $X \sim \mathcal{N}(m, \sigma^2)$	$\mathbb{R}$	$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp(-\frac{(x-m)^2}{2\sigma^2})$	m	σ^2



Exercice 3.2.5. *Le temps d'attente (en minutes) pour accéder à des données suit une loi uniforme* $\mathcal{U}[1, 6]$.

1. *Déterminer la probabilité d'attendre au moins 4 minutes.*
2. *Déterminer le temps d'attente moyen.*

Exercice 3.2.6. *En fiabilité (ou en assurance sur la vie), on considère la fonction de survie* $R(t)$ *définie comme la probabilité qu'un individu vive au* *delà d'une date t : $R(t) = \mathbb{P}(X > t)$. Cette fonction définit une loi de probabilité sur $\mathbb{R}^+$ et $\mathbb{P}(t_1 < X \leq t_2) = R(t_1) - R(t_2)$. La probabilité conditionnelle de défaillance (ou de décès) entre t_1 et t_2 sachant que l'individu a déjà fonctionné (ou vécu) jusqu'à t_1 est*

$$\mathbb{P}(t_1 < X \leq t_2 \,|\, X > t_1) = \frac{R(t_1) - R(t_2)}{R(t_1)}.$$

Supposons que la loi de survie est exponentielle de paramètre λ, c.a.d. $X \sim \mathcal{E}(\lambda)$. Montrer qu'il n'y a pas de vieillissement : la probabilité de fonctionner pendant $t_2 - t_1$ à partir de t_1 est la même qu'au démarrage, i.e.

$$\mathbb{P}(t_1 < X \leq t_2 \,|\, X > t_1) = \mathbb{P}(X \leq t_2 - t_1).$$

3.3 Convergences des suites de variables aléatoires

Une suite (X_n) de v.a. étant une suite de fonctions de Ω dans $\mathbb{R}$ il existe diverses façons de définir la convergence de (X_n) dont certaines jouent un grand rôle en calcul des probabilités.

La suite (X_n) **converge en probabilité** vers la constante a, si $\forall \varepsilon$ et η (arbitrairement petits) il existe n_0 tel que $n > n_0$ entraîne :

$$\mathbb{P}(|X_n - a| > \varepsilon) < \eta.$$

On note alors $X_n \xrightarrow{P} a$.

X et Y sont **égales presque sûrement** si $\mathbb{P}(\{\omega \,|\, X(\omega) \neq Y(\omega)\}) = 0$.

La suite (X_n) **converge presque sûrement** vers X si :

$$\mathbb{P}(\{\omega \,|\, \lim_{n \to \infty} X_n(\omega) \neq X(\omega)\}) = 0$$

et on note $X_n \xrightarrow{ps} X$.

Si $\mathbb{E}((X_n - X)^p)$ existe on a $(X_n) \longrightarrow X$ **en moyenne d'ordre** p si $\mathbb{E}(|X_n - X|^p) \longrightarrow 0$.

La plus utilisée est **la convergence en moyenne quadratique** si $p = 2$.

La suite (X_n) **converge en loi** vers la variable X de f.d.r. si en tout point de continuité de F la suite (F_n) des f.d.r. des X_n converge vers F. On note $X_n \xrightarrow{L} X$.

La hiérarchie des convergences est la suivante.

$$\begin{tikzcd}
\boxed{\text{Moyenne d'ordre } p} \arrow[dr] & & \\
& \boxed{\text{Probabilité}} \arrow[r] & \boxed{\text{Loi}} \\
\boxed{\text{Presque sûre}} \arrow[ur] & &
\end{tikzcd}$$

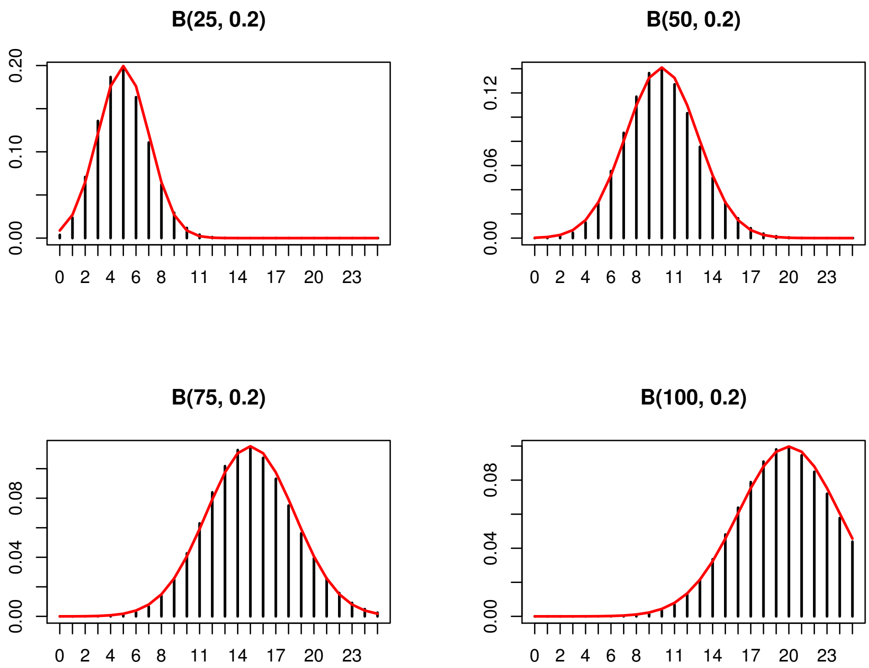
Nous présentons quelques applications de convergences des v.a..

Convergence en loi de $\mathcal{B}(n, p)$ vers $\mathcal{N}(m, \sigma^2)$: une suite de variables discrètes peut cependant converger en loi vers une variable continue.

Théorème de De Moivre-Laplace
X_n étant une suite de variables binomiales $\mathcal{B}(n, p)$ alors

$$\frac{X_n - np}{\sqrt{np(1-p)}} \xrightarrow{L} \mathcal{N}(0, 1).$$

La figure ci dessous illustre l'approximation de la loi binomiale $\mathcal{B}(n, p)$ par la loi normale de même espérance np et de même écart-type $\sqrt{np(1-p)}$.



La convergence de la loi binomiale vers la loi normale se traduit par le fait que les extrémités des bâtons du diagramme de la binomiale $\mathcal{B}(n, p)$ sont voisines de la courbe de densité de la loi $\mathcal{N}(np, \sqrt{np(1-p)})$

Somme de v.a. et théorème central-limite : l'étude de somme de variables indépendantes et de même loi joue un rôle capital en statistique. Le théorème suivant connu sous le nom de **théorème central-limite** établit la convergence vers la loi normale sous des hypothèses peu contraignantes.

Théorème
Soit (X_n) une suite de v.a. indépendantes de même loi d'espérance μ et d'écart-type σ. Alors

$$\frac{1}{\sqrt{n}}\left(\frac{X_1 + \cdots + X_n - n\mu}{\sigma}\right) \xrightarrow{L} \mathcal{N}(0,1).$$

Chapitre 4

Statistique descriptive

à compléter ...

Chapitre 5

Introduction aux statistiques inférentielles

à compléter ...

Bibliographie

[1] Fabrice Rossi. *Probabilités et statistique, Cours et TD de L2 de sciences économiques.*

[2] Gilbert Saporta. *Probabilités analyse des données et statistique.* 1990.