Cours Commun Scientifique

de

Probabilités & Statistiques

Résumé de cours

Suivi des fiches d'exercices

Responsable du cours : Laurent Tournier

Ce document, ainsi que d'autres documents liés au cours, peut être trouvé au format PDF à l'adresse suivante :

http://www.math.univ-paris13.fr/~tournier/enseignement/

1 Espaces de probabilités

1.1 Définitions

Définition

> *Un* ***espace de probabilité*** (Ω,P) *est constitué de*
> — Ω*, un ensemble*
> — P*, une probabilité sur* Ω*.*

Ω représente l'**ensemble des résultats possibles d'une expérience aléatoire**.
Un élément $\omega \in \Omega$ est appelé une **réalisation**, c'est un résultat possible de cette expérience.
Un sous-ensemble $A \subset \Omega$ est appelé un **événement**. C'est un ensemble de réalisations (par exemple, celles qui vérifient une certaine condition). L'ensemble des événements est donc l'ensemble $\mathcal{P}(\Omega)$ des sous-ensembles (ou « parties ») de Ω.
La **probabilité** P donne, pour chaque événement A, la « proportion de chance » $P(A)$ pour qu'il se produise. Elle doit vérifier deux propriétés :

Définition

> *Une* ***probabilité*** *sur* Ω *est une application* $P : \mathcal{P}(\Omega) \to [0,1]$ *telle que*
>
> 1. $P(\Omega) = 1$
> 2. *pour toute suite finie ou infinie* $(A_n)_n$ *d'événements disjoints deux à deux,*
>
> $$P\left(\bigcup_n A_n\right) = \sum_n P(A_n).$$
>
> *(l'union et la somme portent sur toute la suite)*
>
> *Si un événement A vérifie $P(A) = 0$, on dit que A est* ***négligeable*** *; et si $P(A) = 1$, on dit que A est* ***presque sûr****, ou que A a lieu* ***presque sûrement*** *(abrégé « p.s. »).*

Propriétés

> a) $P(\emptyset) = 0$
> b) *Pour tout événement A,* $P(A^c) = 1 - P(A)$
> c) *Si $A \subset B$, alors* $P(A) \leq P(B)$
> d) *Pour tous événements A et B,* $P(A \cup B) = P(A) + P(B) - P(A \cap B)$
> e) *Pour toute suite finie ou infinie $(A_n)_n$ d'événements,* $P\left(\bigcup_n A_n\right) \leq \sum_n P(A_n)$.

NB. Les opérations usuelles sur des événements A et B peuvent s'interpréter intuitivement :

Notation	Sens mathématique	Interprétation en probabilités
$A^c (= \Omega \setminus A)$	complémentaire de A	contraire de A, « non A »
$A \cup B$	réunion de A et B	« A ou B »
$A \cap B$	intersection de A et B	« A et B »
$A \cap B = \emptyset$	A et B sont disjoints	« A et B sont incompatibles »
$A \subset B$	A est inclus dans B	« A implique B ».

> Pour simplifier, on considère dans ce cours que tout ensemble de réalisations est un événement. En réalité, dans beaucoup d'exemples (comme $\Omega = \mathbb{R}$ avec une probabilité à densité), on ne pourrait pas définir $P(A)$ pour tous les ensembles, mais seulement pour des ensembles « mesurables » ; cependant, ceci ne pose pas de problème pratique car tous les ensembles rencontrés seront mesurables.

1.2 Cas symétrique : équiprobabilité

On suppose que l'expérience n'a que n résultats possibles : $\Omega = \{\omega_1,\omega_2,\ldots,\omega_n\}$. Si chacun de ces résultats a la même probabilité alors, comme

$$1 = P(\Omega) = P(\{\omega_1\}) + \cdots + P(\{\omega_n\}),$$

on doit avoir $P(\{\omega_1\}) = \cdots = P(\{\omega_n\}) = \frac{1}{n}$. Ceci définit la probabilité uniforme :

Définition

> *La* ***probabilité uniforme*** *sur Ω (ou distribution équiprobable) est la probabilité P définie par : pour tout $A = \{\omega_{i_1},\omega_{i_2},\ldots,\omega_{i_k}\} \subset \Omega$,*
>
> $$P(A) = \frac{k}{n} = \frac{\mathrm{Card}\, A}{\mathrm{Card}\, \Omega}.$$

On résume parfois ceci par : pour la probabilité uniforme,

$$P(\text{événement}) = \frac{\text{nombre de cas favorables}}{\text{nombre de cas possibles}}.$$

NB. L'expression courante « au hasard » fait souvent référence à un tirage selon la probabilité uniforme. Il en va ainsi lorsque les différents résultats jouent des rôles symétriques.

Rappels de dénombrement :

Proposition

> *Soit E un ensemble fini.*
>
> — *Une* ***permutation*** *de E est une façon d'ordonner les éléments de E.*
> *Le nombre de permutations d'un ensemble à n éléments est*
>
> $$n! = n(n-1)\cdots 2\cdot 1.$$
>
> — *Un* ***arrangement*** *de k éléments de E est une façon de choisir et d'****ordonner*** *k éléments de E : c'est une suite de k éléments de E distincts 2 à 2.*
> *Le nombre d'arrangements de k éléments parmi n éléments (où $0 \leq k \leq n$) est*
>
> $$A_n^k = \underbrace{n(n-1)\cdots(n-k+1)}_{k \text{ termes}} = \frac{n!}{(n-k)!}.$$
>
> — *Une* ***combinaison*** *de k éléments de E est une façon de choisir k éléments de E,* ***sans spécifier d'ordre*** *: c'est un sous-ensemble de E à k éléments.*
> *Le nombre de combinaisons de k éléments parmi n éléments (où $0 \leq k \leq n$) est*
>
> $$\binom{n}{k} = C_n^k = \frac{n(n-1)\cdots(n-k+1)}{k!} = \frac{n!}{k!(n-k)!}.$$

On peut dire aussi qu'un arrangement correspond à un tirage de k éléments *successivement* en mémorisant l'ordre de tirage, tandis qu'une combinaison correspond à un tirage de k éléments *simultanément*.
Un arrangement de n éléments parmi n est une permutation, donc $A_n^n = n!$.

1.3 Probabilités conditionnelles

Définition

Soit B un événement tel que $P(B) > 0$. L'application $P(\cdot|B) : \mathcal{P}(\Omega) \longrightarrow [0,1]$ définie par

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

est une probabilité sur Ω.
$P(A|B)$ est appelée la ***probabilité conditionnelle de** A **sachant** B.*

Définition

Une ***partition*** *de Ω est une suite finie ou infinie $(A_n)_n$ d'événements disjoints, dont la réunion est Ω :*

$$\text{pour tous } i \neq j,\ A_i \cap A_j = \emptyset, \qquad \text{et} \qquad \Omega = \bigcup_n A_n.$$

Par exemple, pour tout événement B, le couple (B,B^c) est une partition de Ω.

Théorème (*Théorème des probabilités totales*)

Soit $(A_n)_n$ une partition de Ω. Pour tout événement A,

$$P(A) = \sum_n P(A|A_n)P(A_n).$$

En particulier, pour tous événements A et B,

$$P(A) = P(A|B)P(B) + P(A|B^c)P(B^c).$$

Théorème (*Formule de Bayes*)

Soit $(A_n)_n$ une partition de Ω. Pour tout événement A, et tout événement A_i de la partition,

$$P(A_i|A) = \frac{P(A|A_i)P(A_i)}{\sum_n P(A|A_n)P(A_n)}.$$

En particulier, pour tous événements A et B,

$$P(B^c|A) = \frac{P(A|B^c)P(B^c)}{P(A|B)P(B) + P(A|B^c)P(B^c)}\ .$$

1.4 Événements indépendants

Définition

Deux événements A et B sont ***indépendants*** *si*

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

On a alors $P(A|B) = P(A)$ et $P(B|A) = P(B)$, si $P(A) \neq 0$ et $P(B) \neq 0$.

Ainsi, A et B sont indépendants si l'information que A est réalisé n'a pas d'influence sur la probabilité que B est réalisé aussi : cela correspond à la notion intuitive d'indépendance.

Proposition (*Indépendance du complémentaire*)

Si A et B sont indépendants, alors A^c et B^c le sont aussi, de même que A et B^c.

On étend la définition à davantage d'événements :

Définition

Une famille $(A_i)_{i\in I}$ d'événements est ***indépendante*** *si pour toute sous-famille finie $A_{i_1}, \ldots, A_{i_k}$ on a*

$$P(A_{i_1} \cap A_{i_2} \cdots \cap A_{i_k}) = P(A_{i_1})P(A_{i_2}) \cdots P(A_{i_k}).$$

En particulier, des événements A, B et C sont indépendants si

$$P(A \cap B) = P(A)P(B), \qquad P(B \cap C) = P(B)P(C), \qquad P(A \cap C) = P(A)P(C)$$
$$\textbf{et} \qquad P(A \cap B \cap C) = P(A)P(B)P(C).$$

Proposition (*Indépendance du complémentaire*)

Si une famille $(A_i)_{i\in I}$ d'événements sont indépendants, et que, pour tout $i \in I$, on définit un événement B_i par $B_i = A_i$ ou $B_i = A_i^c$, alors la famille $(B_i)_{i\in I}$ est indépendante.

Ainsi, si A,B,C sont indépendants, alors A^c,B,C aussi, de même que A^c,B,C^c, etc.

Corollaire (*Loi binomiale*)

Pour toute suite de n événements ***indépendants*** *$A_1, \ldots, A_n$ ayant tous* ***la même probabilité*** *$P(A_i) = p$ on a, pour $0 \leq k \leq n$,*

$$P\big(\text{exactement } k \text{ événements parmi } A_1, \ldots, A_n \text{ se réalisent}\big) = \binom{n}{k} p^k (1-p)^{n-k}.$$

2 Variables aléatoires

Soit (Ω,P) un espace de probabilité.

Définition

*Une **variable aléatoire** est une application $X:\Omega\to\mathbb{R}$.*

Définition

*Soit X une variable aléatoire. La **loi** de X est la probabilité P_X sur $\mathbb{R}$ définie par :*

$$\text{pour tout } B\subset\mathbb{R},\qquad \begin{aligned}P_X(B)&=P(\{\omega\in\Omega\,|\,X(\omega)\in B\})\\&=P(X\in B).\end{aligned}$$

*P_X peut aussi être vue comme une probabilité sur l'ensemble $X(\Omega)$ des valeurs prises par X, aussi appelé **support** de P_X. On abrège $X\sim P_X$ et on dit que X **suit** la loi P_X.*

La seconde égalité est une nouvelle notation : on note $\{X\in B\}$ l'événement formé des réalisations ω pour lesquelles $X(\omega)\in B$, et on abrège $P(\{X\in B\})=P(X\in B)$.

Définition

*Si A est un événement, on introduit la variable aléatoire **fonction indicatrice de** A, notée $\mathbf{1}_A$, qui indique si l'événement A est réalisé :*

$$\text{pour tout } \omega\in\Omega,\qquad \mathbf{1}_A(\omega)=\begin{cases}1 & \text{si } \omega\in A\\ 0 & \text{si } \omega\notin A.\end{cases}$$

2.1 Lois discrètes

Une variable aléatoire X est dite **discrète** si l'ensemble $X(\Omega)$ des valeurs qu'elle prend est *dénombrable* (c'est-à-dire que l'on peut trouver une suite qui énumère tous les éléments de $X(\Omega)$: c'est le cas notamment si $X(\Omega)$ est un ensemble fini, $\mathbb{N}$, $\mathbb{Z}$ ou $\mathbb{Q}$, **mais pas** l'intervalle $[0,1]$ ni $\mathbb{R}$). On dit aussi que la *loi* de X est discrète.
Si X est discrète, alors, pour tout $B\subset\mathbb{R}$, on peut calculer

$$P_X(B)=P(X\in B)=P\Big(\bigcup_{x\in B\cap X(\omega)}\{X=x\}\Big)=\sum_{x\in B\cap X(\Omega)}P(X=x)=\sum_{x\in B}P(X=x).$$

Pour caractériser une loi discrète, il suffit donc de se donner les **probabilités élémentaires**

$$p_X(x)=P(X=x)\qquad\text{pour tout } x\in X(\Omega).$$

Définition-Proposition

*Soit $E\subset\mathbb{R}$. Une famille $(p(x))_{x\in E}$ est une **famille de probabilités élémentaires** si*

1. *pour tout $x\in E$, $p(x)\geq 0$*
2. $\sum_{x\in E}p(x)=1$.

Dans ce cas, on peut définir une variable aléatoire X (sur un espace de probabilité (Ω,P)), à valeurs dans E, de probabilités élémentaires $(p(x))_{x\in E}$, c'est-à-dire telle que,

$$\text{pour tout } x\in E,\qquad P(X=x)=p(x).$$

Exemple : loi uniforme sur un ensemble fini E. Pour un ensemble fini E, on dit qu'une variable aléatoire X suit la loi uniforme sur E si sa loi est la probabilité uniforme sur E :

$$P(X=x)=\frac{1}{\operatorname{Card}E}\qquad\text{pour tout } x\in E.$$

Exemple : loi de Bernoulli de paramètre p. Pour tout événement A de probabilité p, la fonction indicatrice $\mathbf{1}_A$ est une variable aléatoire à valeurs dans $\{0,1\}$, donc elle est discrète et sa loi est donnée par les probabilités

$$P(\mathbf{1}_A=1)=P(A)=p\quad\text{et}\quad P(\mathbf{1}_A=0)=1-p.$$

Cette loi est appelée **loi de Bernoulli** de paramètre p (où $p\in[0,1]$), notée $\mathcal{B}(p)$.
Si $p=1/2$, c'est la loi uniforme sur $\{0,1\}$.

Exemple : loi binomiale de paramètres n et p. Si $A_1,\ldots,A_n$ sont n événements indépendants et ayant la même probabilité p (c'est-à-dire que $P(A_1)=\cdots=P(A_n)=p$), alors la variable aléatoire

$$S_n=\mathbf{1}_{A_1}+\cdots+\mathbf{1}_{A_n}=\text{« nombre d'événements parmi } A_1,\ldots,A_n \text{ qui se réalisent »}$$

est à valeurs dans $\{0,1,\ldots,n\}$, donc est discrète, et sa loi est donnée par les probabilités

$$P(S_n=k)=\binom{n}{k}p^k(1-p)^{n-k}\qquad\text{pour } k=0,1,\ldots,n.$$

Cette loi est appelée **loi binomiale** de paramètres n et p (où $n\in\mathbb{N}$ et $p\in[0,1]$), notée $\mathcal{B}(n,p)$.
C'est donc la loi du nombre de piles en n tirages d'une pièce qui a probabilité p de donner pile.

Exemple : loi géométrique de paramètre p. Si $A_1,A_2,\ldots$ sont des événements indépendants et ayant la même probabilité $p>0$, alors la variable aléatoire

$$N=\min\{n\in\mathbb{N}^*\,|\,\mathbf{1}_{A_n}=1\}=\text{« premier indice } n \text{ tel que } A_n \text{ se réalise »}$$

est à valeurs dans $\mathbb{N}^*$ (on a $P(N=\infty)=P(A_1^c\cap A_2^c\cap\cdots)=(1-p)(1-p)\cdots=0$) et sa loi est donnée par

$$P(N=n)=(1-p)^{n-1}p,\qquad\text{pour } n=1,2,\ldots.$$

Cette loi est appelée **loi géométrique** de paramètre p (où $p\in]0,1]$), notée $\mathcal{G}(p)$. C'est donc la loi du nombre de lancers d'une pièce pour obtenir pile, si sa probabilité de donner pile est p.

Exemple : loi de Poisson de paramètre λ. Soit $\lambda>0$. Une variable aléatoire X suit la loi de Poisson de paramètre λ, notée $\mathcal{P}(\lambda)$, si X est à valeurs dans $\mathbb{N}$ et

$$\text{pour tout } k\in\mathbb{N},\qquad P(X=k)=e^{-\lambda}\frac{\lambda^k}{k!}\ .$$

C'est la loi limite de la loi binomiale $\mathcal{B}(n,p)$, avec n grand et $np\simeq\lambda$:

Proposition

Si, pour tout n, S_n suit la loi $\mathcal{B}(n,p_n)$, et $np_n\underset{n\to\infty}{\longrightarrow}\lambda$, alors

$$\text{pour tout } k\in\mathbb{N},\qquad P(S_n=k)\underset{n\to\infty}{\longrightarrow}e^{-\lambda}\frac{\lambda^k}{k!}.$$

Dans la pratique, on peut souvent approcher la loi binomiale par une loi de Poisson lorsque les conditions $n\geq 50$ et $p\leq 0{,}1$ sont vérifiées *(l'erreur dans les calculs de probabilités est alors inférieure à 5 % ; elle devient inférieure à 1 % si par exemple $n\geq 50$ et $p\leq 0{,}01$).*

2.2 Lois continues (ou à densité)

Une variable aléatoire X est dite **continue** ou **à densité** s'il existe une fonction $f_X : \mathbb{R} \to \mathbb{R}$ telle que, pour tout $B \subset \mathbb{R}$,

$$P_X(B) = P(X \in B) = \int_B f_X(x)dx.$$

La fonction f_X est appelée la **densité** de X.

Définition-Proposition

Une fonction $f : \mathbb{R} \to \mathbb{R}$ est une ***fonction de densité (de probabilité)*** *si*

1. *pour tout $x \in \mathbb{R}$, $f(x) \geq 0$*
2. $\displaystyle\int_{\mathbb{R}} f(x)dx = 1.$

Dans ce cas, on peut définir une variable aléatoire X (sur un espace de probabilité (Ω,P)) de densité $f_X = f$.

Propriétés

Si la variable aléatoire X a pour densité f_X, alors

a) pour tout $x \in \mathbb{R}$, on a $P(X = x) = 0$

b) $P(X \in \mathrm{Supp}(f_X)) = 1$, où le support de la fonction f_X est défini par

$$\mathrm{Supp}(f_X) = \{x \in \mathbb{R} \,|\, f_X(x) > 0\}.$$

Par le point b), on pourra considérer que les valeurs de X sont dans l'ensemble $\mathrm{Supp}(f_X)$.

Exemple : loi uniforme sur un segment $[a,b]$. Soit $a < b$. La loi uniforme sur $[a,b]$, notée $\mathcal{U}([a,b])$, est la loi de densité

$$f(x) = \frac{1}{b-a}\mathbf{1}_{[a,b]}(x) = \begin{cases} \frac{1}{b-a} & \text{si } a \leq x \leq b \\ 0 & \text{si } x \notin [a,b]. \end{cases}$$

Comme $f(x) = 0$ si $x \notin [a,b]$, une variable aléatoire X de loi $\mathcal{U}([a,b])$ est à valeurs dans $[a,b]$.

Exemple : loi exponentielle de paramètre λ. Soit $\lambda > 0$. La loi exponentielle de paramètre λ, notée $\mathcal{E}(\lambda)$, est la loi de densité

$$f(x) = \lambda e^{-\lambda x}\mathbf{1}_{\mathbb{R}_+}(x).$$

Comme $f(x) = 0$ si $x < 0$, une variable aléatoire X de loi $\mathcal{E}(\lambda)$ est à valeurs dans $\mathbb{R}_+$.
La loi exponentielle est une loi « sans mémoire ». En effet, si $X \sim \mathcal{E}(\lambda)$, pour tous $s,t \geq 0$,

$$P(X \geq s+t \,|\, X \geq s) = \frac{P(\{X \geq s+t\} \cap \{X \geq s\})}{P(X \geq s)} = \frac{e^{-\lambda(s+t)}}{e^{-\lambda s}} = e^{-\lambda t} = P(X \geq t),$$

en utilisant le fait que l'événement $\{X \geq s+t\}$ est inclus dans l'événement $\{X \geq s\}$.
De ce fait, les lois exponentielles sont utilisées pour modéliser des durées de vie de machines sans vieillissement.

2.3 Fonction de répartition

Définition

Soit X une variable aléatoire. La **fonction de répartition de** *X est la fonction $F_X : \mathbb{R} \to \mathbb{R}$ définie par*

$$\text{pour tout } x \in \mathbb{R}, \qquad F_X(x) = P(X \leq x).$$

Proposition

a) Soit X une variable aléatoire. Sa fonction de répartition F_X est une fonction croissante,

$$\lim_{x \to -\infty} F_X(x) = 0 \qquad \text{et} \qquad \lim_{x \to +\infty} F_X(x) = 1.$$

b) Si X et Y sont deux variables aléatoires telles que $F_X(t) = F_Y(t)$ pour tout $t \in \mathbb{R}$, alors X et Y ont même loi.

c) Si X est une variable aléatoire discrète, alors F_X est une fonction constante par morceaux, dont les sauts se situent aux points de $X(\Omega)$ et, pour $x \in X(\Omega)$, la hauteur du saut au point x est égale à $P(X = x)$.

d) Si X est une variable aléatoire de densité f_X, alors on a

$$\text{pour tout } x \in \mathbb{R}, \qquad F_X(x) = \int_{-\infty}^{x} f_X(t)dt,$$

donc en particulier F_X est continue, et on a la dérivée $(F_X)'(x) = f_X(x)$ (sauf aux points x où f_X n'est pas continue).
Inversement, si F_X est continue, et dérivable (sauf éventuellement en un nombre fini de points), alors X a pour densité $f_X(x) = (F_X)'(x)$.

2.4 Loi de $Y = \varphi(X)$

Soit X une variable aléatoire à valeurs réelles, et $\varphi : X(\Omega) \to \mathbb{R}$ une fonction. On peut définir la variable aléatoire $Y = \varphi(X)$.
• Si X est discrète, alors Y aussi, et sa loi se calcule directement : Y est à valeurs dans $\varphi(X(\Omega))$ et, pour tout $y \in \varphi(X(\Omega)))$,

$$P(Y = y) = P(\varphi(X) = y) = \sum_{x \in \mathbb{R} \text{ tel que } \varphi(x) = y} P(X = x).$$

• Si X a une densité f_X, on peut souvent montrer que Y a une densité et la calculer :
- *Via la fonction de répartition.* Si φ est strictement croissante sur un intervalle I tel que $X \in I$ presque sûrement, alors φ est une bijection de I sur $J = \varphi(I)$ et on a, pour $y \in J$,

$$F_Y(y) = P(Y \leq y) = P(\varphi(X) \leq y) = P(X \leq \varphi^{-1}(y)) = F_X(\varphi^{-1}(y)).$$

On pourra alors vérifier la continuité de F_Y et dériver (si possible) pour obtenir la densité de Y.
- *Par changement de variable.* Sous les mêmes conditions : pour tous $a,b \in J$ avec $a < b$ on a, en posant $y = \varphi(x)$, c'est-à-dire $x = \varphi^{-1}(y)$, (si φ^{-1} est de classe $\mathcal{C}^1$)

$$P(a < Y < b) = P(\varphi^{-1}(a) < X < \varphi^{-1}(b)) = \int_{\varphi^{-1}(a)}^{\varphi^{-1}(b)} f_X(x)dx = \int_a^b f_X(\varphi^{-1}(y))(\varphi^{-1})'(y)dy$$

donc Y a pour densité $f_Y(y) = f_X(\varphi^{-1}(y))(\varphi^{-1})'(y)$. On retrouverait le même résultat.

2.5 Espérance d'une variable aléatoire

Définition

*L'**espérance** d'une variable aléatoire X, notée $E[X]$, est la moyenne de ses valeurs, pondérées par leurs probabilités.*
Si X est discrète,

$$E[X] = \sum_{x \in X(\Omega)} xP(X = x).$$

Si X est continue, de densité f_X,

$$E[X] = \int_{\mathbb{R}} x f_X(x) dx.$$

***Attention.** L'espérance n'est pas toujours définie. Il faut pour cela que la série ou l'intégrale ci-dessus converge absolument.*

Propriétés

a) *Si X est constante, égale à $c \in \mathbb{R}$ (pour tout $\omega \in \Omega$, $X(\omega) = c$), alors $E[X] = E[c] = c$.*
b) *Pour tout événement $A \subset \Omega$, $E[\mathbf{1}_A] = P(A)$.*
c) *L'espérance est linéaire : pour toutes variables aléatoires X et Y, et tout réel a,*

$$E[aX] = aE[X] \qquad \text{et} \qquad E[X+Y] = E[X] + E[Y].$$

d) *L'espérance est croissante : si $X \leq Y$, alors $E[X] \leq E[Y]$.*

Proposition

Soit X une variable aléatoire, et $\varphi : \mathbb{R} \to \mathbb{R}$ une fonction.

— *Si X est discrète, alors*

$$E[\varphi(X)] = \sum_{x \in X(\Omega)} \varphi(x)P(X = x).$$

— *Si X est continue, alors*

$$E[\varphi(X)] = \int_{\mathbb{R}} \varphi(x) f_X(x) dx.$$

(À condition que la série et l'intégrale soient bien définies)

2.6 Variance d'une variable aléatoire

Définition

*Soit X une variable aléatoire. La **variance** de X est l'espérance des carrés des écarts de X à sa moyenne :*

$$\mathrm{Var}(X) = E\left[\left(X - E[X]\right)^2\right] \geq 0.$$

*L'**écart type** de X est $\sigma(X) = \sqrt{\mathrm{Var}(X)}$.*
***Attention.** La variance n'est pas toujours définie. Il faut que l'espérance $E[X]$ soit définie **et** que l'espérance ci-dessus converge. Tout ceci revient à demander à ce que $E[X^2]$ converge.*

NB. L'écart type $\sigma(X)$ est *homogène* à X : si par exemple X est une distance, alors $\sigma(X)$ est une distance aussi. Ceci justifie l'intérêt de l'écart type par rapport à la variance.

Propriétés

Pour toutes variables aléatoires X et Y ayant une variance, et toute constante a,

a) $\mathrm{Var}(X) = E[X^2] - E[X]^2$

b) $\mathrm{Var}(aX) = a^2 \mathrm{Var}(X)$

c) $\mathrm{Var}(X+a) = \mathrm{Var}(X)$

d) $\mathrm{Var}(X+Y) = \mathrm{Var}(X) + 2\,\mathrm{Cov}(X,Y) + \mathrm{Var}(Y)$, *où la* ***covariance*** *est définie par*

$$\mathrm{Cov}(X,Y) = E\Big[\big(X - E[X]\big)\big(Y - E[Y]\big)\Big] = E[XY] - E[X]E[Y].$$

NB. Si X a une variance, alors $Y = \dfrac{X - E[X]}{\sigma(X)}$ est **centrée** ($E[Y] = 0$) et **réduite** ($\mathrm{Var}(Y) = 1$).

Pour $r > 0$, on définit (s'ils existent) le **moment d'ordre** r et le **moment centré d'ordre** r :

$$m_r(X) = E[X^r] \qquad \text{et} \qquad \mu_r(X) = E\Big[\big(X - E[X]\big)^r\Big].$$

2.7 Inégalités

Proposition (*Inégalité de Markov*)

Soit X une variable aléatoire. Pour tout $a > 0$,

$$P(|X| \geq a) \leq \frac{E\big[|X|\big]}{a}.$$

Plus généralement, pour tout $a > 0$ et $r > 0$,

$$P(|X| \geq a) \leq \frac{E\big[|X|^r\big]}{a^r}.$$

Démonstration : Dans le cas où X a pour densité f_X,

$$\begin{aligned} E[|X|^r] = \int_{\mathbb{R}} |x|^r f_X(x)dx &= \int_{-\infty}^{-a} |x|^r f_X(x)dx + \int_{-a}^{a} |x|^r f_X(x)dx + \int_{a}^{+\infty} |x|^r f_X(x)dx \\ &\geq \int_{-\infty}^{-a} a^r f_X(x)dx \quad + \quad 0 \quad + \quad \int_{a}^{+\infty} a^r f_X(x)dx \\ &= a^r P(X \leq -a) + a^r P(X \geq a) = a^r P(|X| \geq a), \end{aligned}$$

d'où l'inégalité annoncée (la première correspond à $r = 1$). ■

Proposition (*Inégalité de Tchebychev*)

Soit X une variable aléatoire ayant une variance. Pour tout $a > 0$,

$$P\Big(\big|X - E[X]\big| \geq a\Big) \leq \frac{\mathrm{Var}(X)}{a^2}.$$

Autre écriture : pour tout $A > 0$,

$$P\Big(E[X] - A\sigma(X) \leq X \leq E[X] + A\sigma(X)\Big) \geq 1 - \frac{1}{A^2}.$$

Démonstration : Appliquer l'inégalité de Markov pour $r = 2$ à $X - E[X]$. ■

2.8 Indépendance de variables aléatoires

Définition

Des variables aléatoires $X_1,\ldots,X_n$ sont **indépendantes** *si, pour tous $B_1,\ldots,B_n \subset \mathbb{R}$,*

$$P(X_1 \in B_1,\ldots,X_n \in B_n) = P(X_1 \in B_1)\cdots P(X_n \in B_n).$$

Ici, les virgules se lisent « et » : $P(X_1 \in B_1,\ldots,X_n \in B_n) = P(\{X_1 \in B_1\}\cap\cdots\cap\{X_n \in B_n\})$.

Par exemple, deux variables aléatoires X et Y sont indépendantes si les événements qui ne dépendent que de X sont indépendants des événements qui ne dépendent que de Y.

Ainsi, X et Y sont indépendantes si connaître la valeur de X ne renseigne pas sur Y. Les résultats de deux tirages à pile-ou-face ou de deux lancers de dés sont en général indépendants. De même, deux tirages de boules numérotées dans une même urne *avec remise* sont indépendants, tandis que deux tirages *sans remise* ne sont pas indépendants car dans ce cas le second tirage ne peut pas être égal au premier.

Proposition

a) *Si $X_1,\ldots,X_n$ sont indépendantes, alors les variables aléatoires $f_1(X_1),\ldots,f_n(X_n)$ sont indépendantes, quelles que soient les fonctions $f_1,\ldots,f_n$.*

b) ***« Indépendance par paquets ».*** *Si $X_1,\ldots,X_n$ sont indépendantes alors, par exemple, les variables aléatoires $f(X_1,X_2)$, $g(X_4)$, $h(X_3,X_5,X_6)$,... sont indépendantes : les fonctions de « paquets disjoints » de variables sont indépendantes.*

c) *Si des événements $A_1,\ldots,A_n$ sont indépendants, alors leurs fonctions indicatrices $\mathbf{1}_{A_1},\ldots,\mathbf{1}_{A_n}$ sont des variables aléatoires indépendantes ; et réciproquement.*

Proposition

Si $X_1,\ldots,X_n$ sont des variables aléatoires indépendantes, alors

a) *si leurs espérances sont bien définies,*

$$E[X_1\cdots X_n] = E[X_1]\cdots E[X_n]$$

b) *si leurs variances sont bien définies, on a $\mathrm{Cov}(X_i,X_j) = 0$ pour tous $i \neq j$, d'où*

$$\mathrm{Var}(X_1+\cdots+X_n) = \mathrm{Var}(X_1)+\cdots+\mathrm{Var}(X_n).$$

NB. Les réciproques sont fausses !

Par le a) de la proposition précédente on déduit, si $X_1,\ldots,X_n$ sont indépendantes,

$$E[f_1(X_1)\cdots f_n(X_n)] = E[f_1(X_1)]\cdots E[f_n(X_n)].$$

On rappelle que l'espérance est toujours linéaire : même si $X_1,\ldots,X_n$ ne sont pas indépendantes,

$$E[X_1+\cdots+X_n] = E[X_1]+\cdots+E[X_n].$$

2.9 Théorème (ou « Loi ») des grands nombres

Le résultat suivant est fondamental :

Théorème (***Loi des grands nombres***)

Soit $(X_n)_{n\geq 1}$ une suite de variables aléatoires ***indépendantes et de même loi****, d'espérance m et de variance σ^2. On définit la variable aléatoire $\overline{X}_n$, par*

$$\overline{X}_n = \frac{X_1 + \cdots + X_n}{n}.$$

$\overline{X}_n$ est appelée ***moyenne empirique****. On a :*

$$\text{pour tout } \varepsilon > 0, \qquad P\Big(\big|\overline{X}_n - m\big| < \varepsilon\Big) \underset{n\to\infty}{\longrightarrow} 1.$$

Démonstration : Par linéarité de l'espérance, $E[\overline{X}_n] = \frac{1}{n}(E[X_1] + \cdots + E[X_n]) = m$, et les variables sont indépendantes donc

$$\mathrm{Var}\left(\overline{X}_n\right) = \frac{1}{n^2}\sum_{i=1}^{n}\mathrm{Var}(X_i) = \frac{\sigma^2}{n}.$$

L'inégalité de Tchebychev pour $\overline{X}_n$ s'écrit donc, pour tout $\varepsilon > 0$,

$$P\Big(\big|\overline{X}_n - m\big| \geq \varepsilon\Big) \leq \frac{\sigma^2}{n\varepsilon^2},$$

ce qui donne, en passant au complémentaire,

$$P\Big(\big|\overline{X}_n - m\big| < \varepsilon\Big) \geq 1 - \frac{\sigma^2}{n\varepsilon^2}.$$

Comme le terme de droite converge vers 1 quand $n \to \infty$, et que le terme de gauche est toujours ≤ 1, on obtient l'énoncé. ■

NB. Si $(A_n)_{n\geq 1}$ est une suite d'événements *indépendants* et qui ont *même probabilité* p (par exemple, dans une suite de tirages à pile-ou-face, $A_n = \{\text{le } n\text{-ième tirage est pile}\}$, et $p = \frac{1}{2}$), alors en posant $X_i = \mathbf{1}_{A_i}$, on a

$$\overline{X}_n = \frac{\mathbf{1}_{A_1} + \cdots + \mathbf{1}_{A_n}}{n} = \frac{\text{nombre d'événements réalisés parmi } A_1, \ldots, A_n}{n}$$

donc $\overline{X}_n$ est la **fréquence de réalisation** des événements $A_1, \ldots, A_n$, et la loi des grands nombres montre que, si n est grand, cette fréquence a de grandes chances d'être proche de $E[X_1] = p$, qui est la probabilité commune des événements A_n. On retrouve la définition intuitive d'une probabilité.

Ainsi, la fréquence d'apparition de pile dans une suite de tirages à pile-ou-face indépendants converge vers $\frac{1}{2}$ (ou vers la probabilité d'obtenir pile, si la pièce est biaisée).

2.10 Loi normale de moyenne m et de variance σ^2

Les lois normales joueront un rôle important dans la partie Statistiques (Section 5.4).

Définition

*La **loi normale centrée** ($m = 0$) **réduite** ($\sigma = 1$), notée $\mathcal{N}(0,1)$, est la loi de densité*

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

*Si $m \in \mathbb{R}$ et $\sigma \in]0, +\infty[$, la **loi normale de moyenne** m **et de variance** σ^2, notée $\mathcal{N}(m,\sigma^2)$, est la loi de la variable aléatoire $X = m + \sigma Z$, où Z suit la loi $\mathcal{N}(0,1)$.*
*Lorsque X suit une loi normale, on dit que X est une variable aléatoire **gaussienne**.*

La courbe représentative de f est la « courbe en cloche »(voir fin du poly). Si $Z \sim \mathcal{N}(0,1)$, sa fonction de répartition est

$$\Phi(x) = P(Z \leq x) = \int_{-\infty}^{x} e^{-t^2/2} \frac{dt}{\sqrt{2\pi}}.$$

Cette intégrale $\Phi(x)$ ne pouvant pas s'exprimer à l'aide des fonctions usuelles, on utilise une table (imprimée, comme à la fin du polycopié, ou dans un logiciel de calcul numérique), ou une approximation : par exemple,

$$P(Z > x) = 1 - \Phi(x) \underset{x \to \infty}{\sim} \frac{e^{-\frac{x^2}{2}}}{x\sqrt{2\pi}}$$

(avec une erreur relative inférieure à 0,2 si $x > 1{,}9$ (et inférieure à 0,1 si $x > 2{,}9$)).
Les valeurs d'une variable de loi $\mathcal{N}(0,1)$ sont fortement concentrées près de 0 : si Z suit la loi $\mathcal{N}(0,1)$, on obtient, avec la table (ou une approximation),

$$P(-2 < Z < 2) \simeq 0{,}954 \qquad \text{et} \qquad P(-3 < Z < 3) \simeq 0{,}997.$$

De la loi normale centrée réduite se déduisent les autres lois normales. Avec la section 2.4, on obtient que la densité de la loi $\mathcal{N}(m,\sigma^2)$ est

$$f_{m,\sigma}(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{(x-m)^2}{2\sigma^2}}.$$

Pour les calculs, on se ramène à la loi $\mathcal{N}(0,1)$: par exemple, si X suit la loi $\mathcal{N}(m,\sigma^2)$,

$$P(m - 2\sigma < X < m + 2\sigma) = P\Big(-2 < \frac{X - m}{\sigma} < 2\Big) \simeq 0{,}954.$$

Proposition

Toute combinaison linéaire de variables aléatoires gaussiennes indépendantes est une variable aléatoire gaussienne.
Plus précisément, si $X_1, \ldots, X_n$ sont indépendantes et $X_i \sim \mathcal{N}(m_i,\sigma_i^2)$ alors, pour tous $a_1, \ldots, a_n \in \mathbb{R}$, la variable aléatoire

$$X = a_1 X_1 + \cdots + a_n X_n$$

suit la loi $\mathcal{N}(M,\Sigma^2)$, où

$$M = E[X] = \sum_{i=1}^{n} a_i m_i \qquad \text{et} \qquad \Sigma^2 = \mathrm{Var}(X) = \sum_{i=1}^{n} a_i^2 \sigma_i^2.$$

3 Couples de variables aléatoires

3.1 Loi du couple, lois marginales

Définition

> *Soit X,Y deux variables aléatoires. La* **loi du couple** *(X,Y) est la probabilité $P_{(X,Y)}$ sur $\mathbb{R}^2$ qui vérifie :*
>
> $$\text{pour tous } A,B \subset \mathbb{R}, \qquad P_{(X,Y)}(A \times B) = P(X \in A,\, Y \in B).$$
>
> *Les lois de X et de Y se déduisent de $P_{(X,Y)}$: ainsi, pour $A \subset \mathbb{R}$, $P_X(A) = P_{(X,Y)}(A \times \mathbb{R})$. P_X et P_Y sont appelées les* **lois marginales** *de $P_{(X,Y)}$.*

Si X et Y sont indépendantes, la loi du couple est fournie par les lois de X et de Y :

$$P_{(X,Y)}(A \times B) = P_X(A)P_Y(B).$$

La loi du couple contient davantage d'information que P_X et P_Y : elle indique aussi la façon dont les variables dépendent l'une de l'autre (connaître X peut renseigner sur Y).

Cas de deux variables discrètes. Si X et Y sont discrètes alors la loi de (X,Y) est donnée par les probabilités élémentaires :

$$p_{(X,Y)}(x,y) = P(X = x,\, Y = y) \quad \text{pour tous } x \in X(\Omega),\, y \in Y(\Omega).$$

Elles vérifient $p_{(X,Y)}(x,y) \in [0,1]$ pour tous x,y, et

$$\sum_{x \in X(\Omega)} \sum_{y \in Y(\Omega)} p_{(X,Y)}(x,y) = 1.$$

Inversement, les lois marginales se déduisent des $(p_{(X,Y)}(x,y))$:

$$\text{pour tout } x \in X(\Omega), \qquad p_X(x) = P(X = x) = \sum_{y \in Y(\Omega)} P(X = x, Y = y) = \sum_{y \in Y(\Omega)} p_{(X,Y)}(x,y),$$

$$\text{pour tout } y \in Y(\Omega), \qquad p_Y(y) = P(Y = y) = \sum_{x \in X(\Omega)} P(X = x, Y = y) = \sum_{x \in X(\Omega)} p_{(X,Y)}(x,y).$$

NB. X et Y sont indépendantes si, et seulement si $p_{(X,Y)}(x,y) = p_X(x)p_Y(y)$ pour tous x,y.

Cas où $P_{(X,Y)}$ a une densité. On dit que le couple (X,Y) a une densité s'il y a une fonction $f_{(X,Y)} : \mathbb{R}^2 \to \mathbb{R}$ telle que

$$\text{pour tout } D \subset \mathbb{R}^2, \qquad P_{(X,Y)}(D) = \iint_D f_{(X,Y)}(x,y)dx\,dy.$$

$f_{(X,Y)}$ est appelée la **densité** du couple (X,Y). Alors $f_{(X,Y)}(x,y) \geq 0$ pour tous $x,y \in \mathbb{R}$, et

$$\int_{\mathbb{R}} \int_{\mathbb{R}} f_{(X,Y)}(x,y)dx\,dy = 1.$$

Inversement, une telle fonction définit la loi d'un couple (X,Y).
Comme pour les simples variables aléatoires, si (X,Y) a pour densité $f_{(X,Y)}$, alors, presque sûrement, $(X,Y) \in \mathrm{Supp}(f_{(X,Y)})$ où le support de la fonction $f_{(X,Y)}$ est défini par

$$\mathrm{Supp}(f_{(X,Y)}) = \{(x,y) \in \mathbb{R}^2 \,|\, f_{(X,Y)}(x,y) > 0\}.$$

On déduit les lois marginales de la loi du couple et, *dans le cas indépendant*, on déduit la loi du couple des lois marginales :

Proposition

a) Si (X,Y) a pour densité $f_{(X,Y)}$, alors X et Y ont des densités f_X et f_Y données par

$$f_X(x) = \int_{\mathbb{R}} f_{(X,Y)}(x,y)dy \qquad \text{et} \qquad f_Y(y) = \int_{\mathbb{R}} f_{(X,Y)}(x,y)dx.$$

b) Si X et Y ont des densités f_X et f_Y et sont indépendantes, alors (X,Y) a pour densité

$$f_{(X,Y)}(x,y) = f_X(x)f_Y(y).$$

Réciproquement, si $f_{(X,Y)}(x,y) = f(x)g(y)$ pour deux fonctions f et g, alors X et Y sont indépendantes, et les densités de X et Y sont proportionnelles à f et g.

3.2 Calculs d'espérances

Connaître la loi du couple (X,Y) permet de calculer l'espérance de fonctions de X et Y :

Proposition

Soit $\varphi : \mathbb{R}^2 \to \mathbb{R}$ une fonction.

— *Si X et Y sont discrètes, alors*

$$E[\varphi(X,Y)] = \sum_{x\in X(\Omega)} \sum_{y\in Y(\Omega)} \varphi(x,y)P(X=x,\, Y=y).$$

— *Si (X,Y) a pour densité $f_{(X,Y)}$, alors*

$$E[\varphi(X,Y)] = \int_{\mathbb{R}}\int_{\mathbb{R}} \varphi(x,y)f_{(X,Y)}(x,y)dx\,dy.$$

(À condition que les séries et les intégrales soient bien définies)

3.3 Somme de variables aléatoires indépendantes à densité.

Proposition

Si X et Y sont des variables aléatoires indépendantes, de densités f_X et f_Y, alors la variable aléatoire $Z = X+Y$ a une densité, donnée par le produit de convolution de f_X et f_Y :

$$f_{X+Y}(z) = \big(f_X * f_Y\big)(z) = \int_{\mathbb{R}} f_X(x)f_Y(z-x)dx.$$

Démonstration : Comme X,Y sont indépendantes, (X,Y) a pour densité $f_{(X,Y)}(x,y) = f_X(x)f_Y(y)$ donc, pour tout $A \subset \mathbb{R}$, en faisant le changement de variable $y \mapsto z = x+y$, c'est-à-dire $y = z-x$,

$$P(X+Y\in A) = E[\mathbf{1}_A(X+Y)] = \int_{\mathbb{R}}\left(\int_{\mathbb{R}} \mathbf{1}_A(x+y)f_X(x)f_Y(y)dy\right)dx = \int_{\mathbb{R}}\left(\int_{\mathbb{R}} \mathbf{1}_A(z)f_X(x)f_Y(z-x)dz\right)dx$$

$$= \int_{\mathbb{R}} \mathbf{1}_A(z)\Big(\int_{\mathbb{R}} f_X(x)f_Y(z-x)dx\Big)dz = \int_{\mathbb{R}} \mathbf{1}_A(z)(f_X * f_Y)(z)dz.$$

■

NB. Le même principe permet de calculer la densité de $Z = \varphi(X,Y)$ pour diverses fonctions φ.

4 Statistiques descriptives

Les statistiques descriptives visent à décrire un ensemble, en général important, de données, c'est-à-dire à en résumer certaines particularités, sous la forme de représentations graphiques ou de grandeurs numériques. L'interprétation des résultats est ensuite propre à chaque champ d'application. NB. Cette partie n'utilise pas de probabilités (à la différence de la prochaine).

4.1 Vocabulaire, représentations simples

En statistiques, les données dont on dispose associent à chaque **individu** d'une certain ensemble, appelé la **population**, une ou plusieurs **variables** qui quantifient ou qualifient certains caractères des individus (remarque : les individus peuvent être des personnes, mais aussi des objets, des dates, des lieux...). Ces données sont aussi appelées une **série statistique**.
On dispose de données $(x_1,y_1,\ldots),(x_2,y_2,\ldots),\ldots,(x_n,y_n,\ldots)$ où n est l'**effectif total** (taille de la population), et $(x_i,y_i,\ldots)$ sont les **observations** des variables associées au i-ième individu.

Une variable à valeurs numériques est dite **quantitative**. Dans le cas contraire, elle est dite **qualitative**, c'est-à-dire que ses valeurs sont des catégories et non pas des nombres : par exemple, des couleurs, ou des pays. Dans ce cas, les valeurs possibles sont appelées les **modalités** de la variable. Une variable quantitative est **discrète** si ses valeurs possibles sont restreintes à un ensemble fini, et elle est **continue** sinon. Une variable quantitative discrète peut d'ailleurs se voir comme une variable qualitative.
Pour une variable qualitative (ou quantitative discrète),

l'**effectif** d'une modalité (ou d'une valeur) est le nombre de fois où elle est présente dans la population. On représente graphiquement les effectifs par un **diagramme en bâtons**.

la **fréquence** d'une modalité (ou d'une valeur) est le quotient de l'effectif de cette valeur par l'effectif total. On représente graphiquement les fréquences par un **diagramme circulaire** (ou par un diagramme en bâtons, éventuellement empilés).

Pour représenter graphiquement des observations de variables quantitatives continues, on choisit en général un nombre fini d'intervalles et on classifie les individus selon l'intervalle qui contient leur valeur : ainsi, on approche la variable continue par une variable qualitative. La représentation des effectifs constitue alors un **histogramme**. Attention, le choix des intervalles est essentiel et ne suit pas une règle générale.

4.2 Répartition des observations

Supposons que l'on dispose d'observations $x_1,\ldots,x_n$ d'une variable quantitative. Pour tout $x\in\mathbb{R}$, la **fréquence cumulée jusqu'à** x est la proportion d'observations inférieures à x :

$$F(x)=\frac{1}{n}\,\mathrm{Card}\{i\in\{1,\ldots,n\}\,|\,x_i\le x\}.$$

Pour $\alpha\in[0,1]$, un **quantile d'ordre** α est un réel q_α tel qu'une proportion $\ge\alpha$ des données est dans $]-\infty,q_\alpha]$, et une proportion $\ge 1-\alpha$ des données est dans $[q_\alpha,+\infty[$.
En pratique, pour $0<\alpha<1$, si on ordonne les données $x_{(1)}\le\cdots\le x_{(n)}$,

— si $\alpha n\notin\mathbb{N}$, et $k-1<\alpha n<k$ avec $k\in\mathbb{N}$, alors $x_{(k)}$ est l'unique quantile d'ordre α ;
— si $\alpha n=k\in\mathbb{N}^*$, les quantiles d'ordre α sont les éléments de $[x_{(k)},x_{(k+1)}]$.

Une **médiane** est un quantile d'ordre 1/2 : elle sépare les données en deux parties égales.
Un **premier quartile** est un quantile d'ordre 1/4.
Un **troisième quartile** est un quantile d'ordre 3/4.
Le **minimum** est le plus grand quantile d'ordre 0, et le **maximum** est le plus petit d'ordre 1.

On peut représenter ces valeurs par une **boîte de dispersion** (ou « boîte à moustaches »).

4.3 Moyenne, variance

Si $x_1,\dots,x_n$ sont les observations d'une variable quantitative, la **moyenne** de cette série statistique est

$$\overline{x}=\frac{x_1+\cdots+x_n}{n}.$$

Pour une variable quantitative discrète, de valeurs $v_1,\dots,v_r$, ayant pour effectifs $n_1,\dots,n_r$ et fréquences $f_1,\dots,f_r$, on a aussi

$$\overline{x}=\frac{n_1v_1+\cdots+n_rv_r}{n}=f_1v_1+\cdots+f_rv_r.$$

Moyenne et médiane sont des valeurs qui « approchent » au mieux toutes les données :

Proposition

La moyenne est l'unique réel c qui minimise $\sum_{i=1}^n(x_i-c)^2$.
Les médianes sont les réels c qui minimisent $\sum_{i=1}^n|x_i-c|$.

Remarquons que la moyenne est sensible aux valeurs aberrantes, à la différence de la médiane.
La **variance** de la série statistique est

$$s_x^2=\frac{1}{n}\sum_{i=1}^n(x_i-\overline{x})^2=\frac{1}{n}\sum_{i=1}^n x_i^2-\overline{x}^2\quad(=\overline{x^2}-\overline{x}^2),$$

et son **écart type** est $\sigma_x=\sqrt{s_x^2}$. C'est une mesure de la dispersion des valeurs autour de la moyenne : si le quotient $\frac{\sigma_x}{|\overline{x}|}$ est petit (et $\overline{x}\neq 0$), les données sont concentrées autour de $\overline{x}$.

4.4 Observation d'un couple de variables, régression linéaire

On suppose que l'on observe deux variables pour chaque individu, si bien que l'on dispose de données $(x_1,y_1),\dots,(x_n,y_n)$.
La **covariance** des deux variables est

$$s_{x,y}=\sum_{i=1}^n(x_i-\overline{x})(y_i-\overline{y})=\sum_{i=1}^n x_iy_i-\overline{x}\,\overline{y}\quad(=\overline{xy}-\overline{x}\,\overline{y}),$$

(donc $s_x^2=s_{x,x}$) et leur **corrélation** est $\rho_{x,y}=\frac{s_{x,y}}{\sigma_x\sigma_y}$. Une corrélation positive (et proche de 1) indique que les variables ont tendance à être simultanément grandes, ou simultanément petites, tandis qu'une corrélation négative (et proche de -1) indique des variations en sens opposés.
Si les variables sont fortement corrélées (corrélation proche de ±1), on peut penser qu'une variable est la cause de l'autre, et on peut chercher à approcher y par une fonction de x. Le cas le plus simple est le cas affine ($y_i\simeq ax_i+b$), qui peut s'aborder par **régression linéaire** :

Proposition (*Droite des moindres carrés*)

La droite $y=ax+b$ qui minimise la somme des carrés des erreurs

$$E(a,b)=\sum_{i=1}^n\big(y_i-(ax_i+b)\big)^2$$

est donnée par

$$a=\frac{s_{x,y}}{s_x^2}=\frac{\overline{xy}-\overline{x}\,\overline{y}}{\overline{x^2}-\overline{x}^2}\quad\text{et}\quad b=\overline{y}-a\overline{x}.$$

NB. Si on cherche une approximation de la forme $y=Cx^\alpha$, alors $\ln(y)=\alpha\ln(x)+\ln(C)$, ce qui permet de déterminer C et α par régression linéaire entre $\ln(x)$ et $\ln(y)$.

5 Estimation

5.1 Principe, statistiques classiques

Définition

> *Soit X une variable aléatoire. Un* **échantillon de taille n de X** *est une famille $X_1,\dots,X_n$ de n variables aléatoires indépendantes et de même loi que X.*

On souhaite étudier la loi de X. *Par exemple, X est la taille en centimètres d'un individu choisi uniformément dans la population adulte française. Son espérance est donc la taille moyenne d'un Français adulte, que l'on peut vouloir estimer.*
On ne dispose pour cela que d'une **observation** d'un échantillon de taille n : une réalisation $(x_1,\dots,x_n) \in \mathbb{R}^n$ de n variables aléatoires indépendantes $(X_1,\dots,X_n)$ qui ont la même loi que X.
À défaut de pouvoir mesurer toute la population, ce qui serait long, coûteux et compliqué, on se contente de mesurer la taille de n personnes ***choisies au hasard*** *parmi les Français adultes.*
L'objectif de l'estimation statistique est de déduire certaines propriétés de la loi de X (son espérance, sa variance, ses paramètres...) à partir des valeurs observées $x_1,\dots,x_n$. On va notamment appliquer des méthodes de statistiques descriptives à ces données.
Pour étudier la distribution (la loi) des quantités qui dépendent des données, on les considère comme des variables aléatoires, donc des fonctions de $X_1,\dots,X_n$. Lors de l'utilisation pratique, on remplacera $X_1,\dots,X_n$ par l'observation $x_1,\dots,x_n$ dont on dispose.

Statistiques simples. Les quantités les plus classiques pour décrire un échantillon sont

— la **moyenne empirique** :

$$\overline{X}_n = \frac{X_1 + \cdots + X_n}{n} = \frac{1}{n}\sum_{i=1}^{n} X_i$$

— la **variance empirique** :

$$S_n^2 = \frac{1}{n}\sum_{i=1}^{n} \left(X_i - \overline{X}_n\right)^2$$

— la **variance empirique modifiée** :

$$\Sigma_n^2 = \frac{1}{n-1}\sum_{i=1}^{n} \left(X_i - \overline{X}_n\right)^2.$$

Proposition

> *Si X a pour espérance m et pour écart type σ, et $X_1,\dots,X_n$ est un échantillon de X, alors*
>
> $$E\left[\overline{X}_n\right] = m, \qquad \mathrm{Var}\left(\overline{X}_n\right) = \frac{\sigma^2}{n},$$
>
> $$E\left[S_n^2\right] = \frac{n-1}{n}\sigma^2, \qquad E\left[\Sigma_n^2\right] = \sigma^2.$$

5.2 Estimateurs

On suppose que X suit une loi P_θ qui dépend d'un paramètre $\theta \in \Theta$, où $\Theta \subset \mathbb{R}$ est l'ensemble des valeurs a priori possibles du paramètre. On ignore la valeur de θ, et on souhaite l'estimer.
NB. Dans ce qui suit, on pourrait noter E_θ au lieu de E pour indiquer que la loi de X et donc de $X_1,\ldots,X_n$ dépend de θ ; on ne le fait pas, afin d'alléger la notation, mais il faut le garder à l'esprit.

Définition

*Un **estimateur** de θ est une variable aléatoire $T_n = f(X_1,\ldots,X_n)$ qui dépend d'un échantillon $X_1,\ldots,X_n$ de X. On utilise souvent la notation $\widehat{\theta}$ pour un estimateur de θ.*
*Une **estimation** de θ est la valeur réelle $t_n = f(x_1,\ldots,x_n)$ prise par une réalisation particulière de l'échantillon.*

NB. La définition d'estimateur peut paraître curieuse à plusieurs titres. On note que n n'apparaît pas dans l'appellation « estimateur de θ » mais apparaît dans la définition. En fait, un estimateur peut être vu comme une *suite* $(T_n)_n$ de variables aléatoires où T_n dépend de $X_1,\ldots,X_n$, et on utilisera la variable T_n adaptée à la taille de l'échantillon dont on dispose. De plus, θ n'apparaît pas dans la définition : n'importe quelle fonction de $X_1,\ldots,X_n$ est donc un estimateur de θ. En revanche θ intervient pour mesurer la qualité de l'estimateur :

Définition

Soit T_n un estimateur de θ.
*Le **biais** de T_n est la différence $E[T_n] - \theta$.*
*On dit que T_n est **sans biais** si $E[T_n] = \theta$, quel que soit $\theta \in \Theta$.*
*On dit que T_n est **asymptotiquement sans biais** si $E[T_n] \underset{n\to\infty}{\longrightarrow} \theta$, quel que soit $\theta \in \Theta$.*
*On dit que T_n est **convergent** si, quel que soit $\theta \in \Theta$,*

$$\text{pour tout } \alpha > 0, \qquad P\big(|T_n - \theta| > \alpha\big) \underset{n\to\infty}{\longrightarrow} 0.$$

Proposition

Tout estimateur asymptotiquement sans biais dont la variance tend vers 0 est convergent.

Définition

*Le **risque quadratique** d'un estimateur T_n de θ est*

$$R_{T_n}(\theta) = E\big[(T_n - \theta)^2\big].$$

*On dit que l'estimateur S_n est **meilleur** que T_n si, quel que soit θ,*

$$R_{S_n}(\theta) \le R_{T_n}(\theta).$$

Par l'inégalité de Markov, un estimateur dont le risque quadratique tend vers 0 (quel que soit θ) est convergent.
NB. Si T_n est sans biais, alors $R_{T_n}(\theta) = \mathrm{Var}(T_n)$.

NB. On peut parfois vouloir estimer non pas un paramètre réel, mais une fonction de la loi de X : par exemple, les fréquences cumulées de l'échantillon (voir p. 16) estiment la fonction de répartition de X. Et un histogramme peut servir à estimer la densité de X, si X a une densité.

5.3 Construction d'estimateurs

Méthode des moments Le principe est d'utiliser la loi des grands nombres pour estimer les moments, et d'utiliser ensuite ces estimateurs des moments pour estimer θ.
Par la loi des grands nombres, on a :

Proposition

Soit X une variable aléatoire d'espérance m et de variance σ^2.

a) La moyenne empirique est un estimateur sans biais et convergent de m.

b) La variance empirique est un estimateur asymptotiquement sans biais et convergent de σ^2, et la variance empirique modifiée est un estimateur sans biais et convergent de σ^2.

c) Pour tout $r > 0$, le moment empirique d'ordre r,

$$\widehat{m}_r = \frac{1}{n}\sum_{i=1}^{n}(X_i)^r$$

est un estimateur sans biais et convergent de $m_r = E[X^r]$. (Si m_r est bien défini)

On en déduit la méthode des moments : exprimer (si possible) θ à l'aide des moments $(m_r)_{r>0}$, puis remplacer dans cette expression les moments par les moments empiriques. Ceci fournit un estimateur convergent de θ. L'expression peut aussi faire intervenir σ^2, remplacé par S_n^2 ou Σ_n^2. En pratique, on calcule $E[X]$, $E[X^2]$, etc., jusqu'à obtenir une expression faisant intervenir θ (souvent, $E[X]$ suffit), et on inverse pour obtenir θ en fonction de $E[X]$, $E[X^2]$,etc. Il ne reste plus qu'à remplacer $m_1 = E[X]$ par $\overline{X}_n$, $m_2 = E[X^2]$ par $\widehat{m}_2$, etc.

Méthode du maximum de vraisemblance Le principe est d'estimer θ par la valeur qui maximise la densité du vecteur $(X_1, \dots, X_n)$.

Définition

*La **vraisemblance** de l'échantillon $(X_1, \dots, X_n)$ est la fonction L définie par :*

— *si X est discrète, de probabilité élémentaire P_θ, pour tous $x_1, \dots, x_n$,*

$$L(x_1, \dots, x_n; \theta) = \prod_{i=1}^{n} P_\theta(x_i)$$

— *si X est continue, de densité f_θ, pour tous $x_1, \dots, x_n$,*

$$L(x_1, \dots, x_n; \theta) = \prod_{i=1}^{n} f_\theta(x_i).$$

*Un **estimateur du maximum de vraisemblance** (ou EMV) pour θ est un estimateur $h(X_1, \dots, X_n)$ tel que, pour toutes les valeurs $x_1, \dots, x_n \in X(\Omega)$,*

$$L\big(x_1, x_2, \dots, x_n; h(x_1, \dots, x_n)\big) = \max_{\theta\in\Theta} L(x_1, \dots, x_n; \theta).$$

NB. Pour tous $x_1, \dots, x_n$, $h(x_1, \dots, x_n)$ est la ou l'une des valeurs de θ où $L(x_1, \dots, x_n; \theta)$ est maximum. Ceci définit l'estimation. L'*estimateur* est la variable aléatoire $h(X_1, \dots, X_n)$.
Sous des hypothèses assez générales, on montre que ceci définit un bon estimateur convergent.
Pour le calcul, on est amené à maximiser $L(x_1, \dots, x_n; \theta)$ selon θ. Vu que le logarithme est strictement croissant, c'est équivalent à maximiser la **log-vraisemblance** $\ln(L(x_1, \dots, x_n; \theta))$, souvent plus pratique. Pour maximiser, on étudie les variations de la fonction $\theta \mapsto \ln(L(x_1, \dots, x_n; \theta))$.

5.4 Intervalles de confiance

Définition

*Un **intervalle de confiance de niveau** $1-\alpha$ **pour** θ est un intervalle I, qui dépend de $X_1,\ldots,X_n$, tel que*

$$P(\theta \in I) \geq 1-\alpha.$$

Intervalle de confiance pour la moyenne m. Soit X une variable aléatoire de moyenne m et de variance σ^2.

Soit $(X_1,X_2,\cdots,X_n)$ un échantillon de X. On cherche un intervalle de confiance à partir des estimateurs sans biais $\overline{X}_n$ et Σ_n^2.

Théorème (***Théorème central limite***)

Soit $(X_n)_n$ une suite de variables aléatoires indépendantes, de même loi de moyenne m et d'écart type σ. Soit Z_n la variable aléatoire définie par

$$Z_n = \frac{\sqrt{n}(\overline{X}_n - m)}{\sigma}.$$

Lorsque $n \to +\infty$, Z_n converge en loi vers une variable $Z \sim \mathcal{N}(0,1)$. On a donc

$$\text{pour tout intervalle } I \subset \mathbb{R}, \qquad P(Z_n \in I) \underset{n\to\infty}{\longrightarrow} P(Z \in I) = \frac{1}{\sqrt{2\pi}} \int_I e^{-\frac{x^2}{2}} dx.$$

Dans la pratique on applique ce résultat dès que n est suffisamment grand ($n \geq 30$).
Soit Z une variable aléatoire de loi $\mathcal{N}(0,1)$ et a et α définis par

$$P(-a \leq Z \leq a) = 1-\alpha$$

Par exemple, on sait que (voir la table)

$$P(|Z| \leq 1{,}96) = 95\% \quad \text{et} \quad P(|Z| \leq 2{,}576) = 99\%.$$

Si $n \geq 30$, le théorème central limite permet d'écrire

$$P(-a \leq \frac{\sqrt{n}(\overline{X}_n - m)}{\sigma} \leq a) \simeq 1-\alpha.$$

On obtient, vu que Σ_n^2 est un estimateur convergent de σ^2,

$$P(-a \leq \frac{\sqrt{n}(\overline{X}_n - m)}{\sqrt{\Sigma_n^2}} \leq a) \simeq 1-\alpha,$$

ce qui se réécrit

$$P\Big(\overline{X}_n - a\frac{\sqrt{\Sigma_n^2}}{\sqrt{n}} \leq m \leq \overline{X}_n + a\frac{\sqrt{\Sigma_n^2}}{\sqrt{n}}\Big) \simeq 1-\alpha.$$

Si $\overline{x}_n$ est la moyenne observée et σ_n^2 la variance modifiée observée, on en déduit que

$$[\pi_1,\pi_2] = \left[\overline{x}_n - a\frac{\sqrt{\sigma_n^2}}{\sqrt{n}},\, \overline{x}_n + a\frac{\sqrt{\sigma_n^2}}{\sqrt{n}}\right]$$

est un intervalle de confiance de m de niveau $1-\alpha$.

Intervalle de confiance pour une proportion p. Ici X est une variable aléatoire de Bernoulli de paramètre p, d'où $E[X] = p$ et $\mathrm{Var}(X) = p(1-p)$.
On a, pour n grand,

$$\frac{\sqrt{n}(\overline{X}_n - p)}{\sqrt{p(1-p)}} \sim \mathcal{N}(0,1)$$

et l'intervalle de confiance pour p de niveau $1-\alpha$ précédent est donc

$$I = \left[\overline{x}_n - a\frac{\sqrt{\overline{x}_n(1-\overline{x}_n)}}{\sqrt{n}}, \overline{x}_n + a\frac{\sqrt{\overline{x}_n(1-\overline{x}_n)}}{\sqrt{n}}\right].$$

Remarques.

1. Attention : pour que ces approximations soient justifiées, les valeurs de $n\pi_1$, $n\pi_2$, $n(1-\pi_1)$ et $n(1-\pi_2)$ doivent être toutes les quatre supérieures ou égales à 5.
2. Les intervalles de confiance donnés ci-dessus permettent aussi de déterminer la taille n de l'échantillon nécessaire pour avoir une précision donnée pour l'estimation d'une proportion.
3. Vu que, pour tout $p \in [0,1]$, $p(1-p) \leq 1/4$, on a $I \subset J$ où J est donc aussi de niveau α et plus simple mais un peu plus large :

$$J = \left[\overline{x}_n - \frac{a}{2\sqrt{n}}, \overline{x}_n + \frac{a}{2\sqrt{n}}\right].$$

5.5 Tests d'hypothèse

On souhaite, à partir de l'échantillon observé $x_1,\ldots,x_n$, savoir si l'on peut raisonnablement conclure qu'une certaine hypothèse sur la loi de X est fausse (en vue de prendre une décision). L'hypothèse est appelée **hypothèse nulle**, et notée $\mathcal{H}_0$. C'est une hypothèse que l'on veut avoir « peu de chance » de rejeter si elle est vraie (erreur de première espèce).

Définition

*Un **test de $\mathcal{H}_0$ au seuil de risque** α est une condition portant sur $X_1,\ldots,X_n$ permettant de décider si $\mathcal{H}_0$ est rejetée ou non, et telle que,*

$$\text{si } \mathcal{H}_0 \text{ est vraie,} \qquad P(\text{rejet de } \mathcal{H}_0) \leq \alpha.$$

Le risque α doit être petit, en général on demande $\alpha = 5\%$.

On peut aussi spécifier un condition alternative $\mathcal{H}_1$ (si elle n'est pas précisée, c'est que $\mathcal{H}_1$ est le contraire de $\mathcal{H}_0$), et on parle de **test de $\mathcal{H}_0$ contre $\mathcal{H}_1$**. Le risque α ne dépend que de $\mathcal{H}_0$.

Définition

*Rejeter $\mathcal{H}_0$ alors que $\mathcal{H}_0$ est vraie est l'**erreur de première espèce**.*
*Ne pas rejeter $\mathcal{H}_0$ alors que $\mathcal{H}_1$ est vraie est l'**erreur de deuxième espèce**.*
*On dit que le test de $\mathcal{H}_0$ contre $\mathcal{H}_1$ a une **puissance** $1-\beta$ si*

$$\text{si } \mathcal{H}_1 \text{ est vraie,} \qquad P(\text{rejet de } \mathcal{H}_0) \geq 1-\beta.$$

NB. Le « test » qui consiste à ne *jamais* rejeter $\mathcal{H}_0$ a un seuil de risque 0. Par contre, sa puissance est nulle. Il ne fait jamais d'erreur de première espèce mais toujours une erreur de deuxième espèce. Un bon test a un risque faible et une puissance proche de 1.

Test de comparaison à une moyenne théorique

Soit $\overline{x}_n$ la moyenne observée sur un échantillon de taille n. On veut savoir si cette moyenne est conforme à la moyenne théorique (c'est-à-dire l'espérance) annoncée, μ.

$$\text{Hypothèse } \mathcal{H}_0 : \text{la moyenne théorique est } \mu.$$

Sous l'hypothèse $\mathcal{H}_0$, $E[X] = \mu$, si bien que, si n est suffisamment grand, approximativement,

$$Z_n = \frac{\sqrt{n}(\overline{X}_n - \mu)}{\sqrt{\Sigma_n^2}} \sim \mathcal{N}(0{,}1),$$

d'où par exemple

$$P\Big(- 1{,}96 \leq \frac{\sqrt{n}(\overline{X}_n - \mu)}{\sqrt{\Sigma_n^2}} \leq 1{,}96 \Big) \simeq 95\,\%.$$

Soit σ_n^2 la variance empirique (modifiée ou non) observée sur l'échantillon. Alors on considère le test suivant :

$$\text{Test } \textit{bilatéral} : \quad \text{si } \left| \frac{\sqrt{n}(\overline{x}_n - \mu)}{\sigma_n} \right| > 1{,}96, \text{ alors on rejette l'hypothèse } \mathcal{H}_0.$$

C'est un test dont le seuil de risque est 5 %. (On remplace 1,96 par a pour un seuil de risque α)
Si l'hypothèse est rejetée, on peut donc conclure que l'espérance n'est pas μ, et on se trompe dans seulement 5 % des cas.
NB. Si le test est faux, on ne peut pas conclure que l'espérance *est* μ, mais simplement que l'échantillon *ne permet pas d'exclure* qu'elle vaut μ. On utilise ce test pour détecter (avec grande probabilité) les cas où la moyenne observée *n'est pas conforme* à μ. Par exemple pour vérifier si un fabricant fournit bien des pièces qui ont une précision donnée, ou des médicaments qui ont une certaine efficacité, etc.

Néanmoins, ce test a heureusement tendance à conduire au rejet de $\mathcal{H}_0$ lorsque $\mathcal{H}_0$ est fausse : en effet, si $E[X] = m$ avec $m \neq \mu$, alors $\frac{\overline{X}_n - \mu}{\sqrt{\Sigma_n^2}} \simeq \frac{m-\mu}{\sigma} \neq 0$ donc $\left| \frac{\sqrt{n}(\overline{X}_n - m)}{\sqrt{\Sigma_n^2}} \right| \underset{n\to\infty}{\longrightarrow} +\infty$ et par conséquent le test rejette $\mathcal{H}_0$ avec une probabilité qui tend vers 1.
Ceci revient à dire que le test précédent est un test de l'hypothèse $\mathcal{H}_0$ « l'espérance est μ » contre $\mathcal{H}_1$ « l'espérance est $\neq \mu$ » de risque 5 % et dont la puissance tend vers 1 quand $n \to \infty$.

Souvent, le seul cas considéré à risque est lorsque $E[X] > \mu$, c'est-à-dire que l'hypothèse alternative est $\mathcal{H}_1$ « l'espérance est $> \mu$ » . On utilise alors un test **unilatéral** pour augmenter la puissance : par la table,

$$P(Z_n \leq 1{,}65) \simeq 95\,\%,$$

d'où

$$\text{Test } \textit{unilatéral} : \quad \text{si } \frac{\sqrt{n}(\overline{x}_n - \mu)}{\sigma_n} > 1{,}65, \text{ alors on rejette l'hypothèse } \mathcal{H}_0.$$

C'est toujours un test de $\mathcal{H}_0$ de risque 5 %, mais par rapport au test précédent il a davantage tendance à conduire au rejet de $\mathcal{H}_0$ si $\mathcal{H}_1$ est vrai ; il est donc plus puissant.

On définit symétriquement un test de $\mathcal{H}_0$ « $E[X] = \mu$ » contre $\mathcal{H}_1$ « $E[X] < \mu$ ».

Test de comparaison à une fréquence théorique

Soit $\overline{x}_n$ la fréquence observée d'un caractère C sur un échantillon de taille n. On cherche à savoir si cette proportion paraît conforme à une fréquence théorique (probabilité) annoncée, π.

Hypothèse $\mathcal{H}_0$: la fréquence théorique est π.

Sous l'hypothèse $\mathcal{H}_0$, X suit la loi de Bernoulli de paramètre π. Si n est suffisamment grand,

$$Z_n = \frac{\sqrt{n}(\overline{X}_n - \pi)}{\sqrt{\pi(1-\pi)}} \sim \mathcal{N}(0,1),$$

d'où par exemple

$$P\Big(-1{,}96 \leq \frac{\sqrt{n}(\overline{X}_n - \pi)}{\sqrt{\pi(1-\pi)}} \leq 1{,}96 \Big) \leq 95\,\%.$$

Comme $\overline{x}_n$ est l'estimation associée à $\overline{X}_n$, ceci conduit au test suivant, qui a un seuil de risque de 5 % :

$$\text{Test } \textit{bilatéral} : \quad \text{si } \left| \frac{\sqrt{n}(\overline{x}_n - \pi)}{\sqrt{\pi(1-\pi)}} \right| > 1{,}96 \text{ alors on rejette l'hypothèse } \mathcal{H}_0.$$

On peut ici aussi considérer le test de « la fréquence théorique est π » contre « la fréquence théorique est $> \pi$ », auquel cas on utilise le test suivant, de risque 5 % et plus puissant :

$$\text{Test } \textit{unilatéral} : \quad \text{si } \frac{\sqrt{n}(\overline{x}_n - \pi)}{\sqrt{\pi(1-\pi)}} > 1{,}65 \text{ alors on rejette l'hypothèse } \mathcal{H}_0.$$

et de même symétriquement si $\mathcal{H}_1$ est « la fréquence théorique est $< \pi$ ».

Fiche 1 – Espaces de probabilités, probabilités conditionnelles, loi binomiale

---

Exercice 1. Combien de nombres peut-on former avec les chiffres 0, 1, 2, 3, 4, 5, 6 et 7, en tenant compte de l'ensemble des contraintes suivantes :
• chaque nombre est composé de chiffres différents ;
• chaque nombre commence par un 7 ;
• chaque nombre est divisible par 5,
1. si les nombres sont de 8 chiffres ?
2. si les nombres sont de 6 chiffres ?

Exercice 2. On lance un dé à 6 faces. Vérifier que les événements $A = \{\text{le résultat est pair}\}$ et $B = \{\text{le résultat est 1 ou 2}\}$ sont indépendants.

Exercice 3. On lance un dé jusqu'à obtenir un résultat différent de 6.
1. Quelle est la probabilité que ce résultat soit égal à 1 ?
2. Soit $n \in \mathbb{N}^*$. Quelle est la probabilité qu'il ait fallu effectuer n lancers ?

Exercice 4. Pour une soirée, on propose à chacun des 3 invités de contribuer au dîner en apportant au choix une entrée ou un dessert et, s'ils le souhaitent, une boisson. Les choix des invités sont supposés indépendants.
On suppose de plus que chaque invité a une chance sur deux d'apporter une entrée (plutôt qu'un dessert), et a une chance sur cinq d'apporter à boire ; enfin, ces deux événements sont supposés indépendants.
1. Quelle est la probabilité pour que le premier invité
— apporte une entrée et à boire ?
— apporte une entrée ou à boire ? (« ou » inclusif)
— apporte une entrée ou à boire, mais pas les deux ? (« ou » exclusif)
2. Calculer la probabilité pour que quelqu'un apporte à boire.
3. Quelle est la probabilité pour qu'il y ait, au total :
— aucune entrée ?
— exactement deux entrées ?
— au moins deux entrées ?
— au plus deux entrées ?
— exactement une entrée ?
— au moins une entrée et un dessert ?
4. Justifier que les événements « il y a au moins une entrée et un dessert » et « il y a à boire » sont indépendants.
5. En déduire la probabilité de pouvoir faire un repas complet avec ce que les invités auront apporté.

Exercice 5. Lors d'une brève conversation vous apprenez d'un homme qu'il a deux enfants dont au moins une fille. Quelle est la probabilité pour qu'il ait deux filles ?

Exercice 6. Une population est composée de 40 % d'hommes et de 60 % de femmes. Dans cette population, 50 % des femmes et 30 % des hommes lisent plus de 10 romans par an. Quelle est la probabilité pour qu'un lecteur de plus de 10 romans par an, choisi au hasard, soit un homme ?

Exercice 7. Deux usines fabriquent les mêmes pièces. La première en produit 70 % de bonnes et la deuxième 90 %. Les deux usines fabriquent la même quantité de pièces.
1. Quel est le pourcentage de bonnes pièces sur l'ensemble du marché, alimenté par les deux usines ?
2. On achète une pièce, elle est bonne ; quelle est la probabilité pour qu'elle provienne de la 2ᵉ usine ?
3. Mêmes questions lorsque la première usine produit 2,5 fois plus que la deuxième.

Exercice 8. On considère un avion de 50 places. La probabilité pour qu'un voyageur ayant réservé ne se présente pas à l'embarquement est de 2 % . Un jour la compagnie a enregistré 52 réservations. Quelle est la probabilité pour qu'elle se trouve dans une situation embarrassante ?

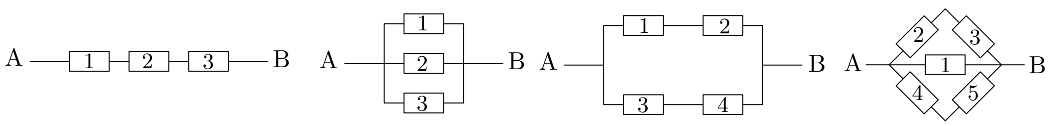
Exercice 9. Un nouveau test de dépistage d'une maladie rare, touchant environ une personne sur 100 000, vient d'être mis au point. Pour tester sa validité, on a effectué un test statistique : sur 534 sujets sains, le test a été positif 1 fois et, sur 17 sujets malades, il a été positif 16 fois.
Une personne effectue ce test ; le résultat est positif. Donner une estimation de la probabilité pour que cette personne soit atteinte par la maladie. Au vu de ce résultats, peut-on commercialiser le test ?

Exercice 10. La probabilité pour que l'injection d'un vaccin à un individu choisi au hasard provoque une réaction allergique est de 0,1 %. Quelle est la probabilité pour que, sur 900 individus vaccinés, on observe l'allergie dans :
1. exactement trois cas ?
2. au plus trois cas ?
3. au moins trois cas ?

Exercices complémentaires

Exercice 11. Dans une université, sur 500 étudiants en première année de licence, 100 font le trajet à pieds, 350 viennent en transports en commun, et 50 arrivent en voiture. On a constaté que ceux qui viennent à pied sont toujours à l'heure, ceux qui prennent les transports en commun sont en retard 15% des jours, et ceux qui viennent en voiture sont en retard un jour sur 10. On suppose que le fait pour un étudiant d'arriver en retard est indépendant d'un jour à l'autre.
1. Quelle est la probabilité pour qu'un étudiant, pris au hasard, arrive en retard ?
2. Un étudiant vient d'arriver en retard. Quelle est la probabilité qu'il soit venu en voiture ?
3. Sur les 10 premiers jours, un étudiant est arrivé en retard 3 fois.
3.a) Selon son moyen de transport, quelle est la probabilité que ceci arrive ?
3.b) Quelle est la probabilité qu'il soit venu en voiture ?
4. En 10 jours, un étudiant a toujours été à l'heure. Quelle est la probabilité qu'il vienne à pieds ?

Exercice 12. Chaque dessin ci-dessous schématise un système mécanique ou informatique, où un rectangle numéroté correspond à une partie pouvant tomber en panne avec probabilité $p = 1\,\%$. Le système fonctionne tant que l'on peut relier A à B par un chemin sans partie en panne. Les parties sont indépendantes. Calculer la probabilité de fonctionnement dans chaque cas.



Exercice 13. Maryse joue à un jeu télévisé. Elle a, face à elle, trois portes (A, B et C) identiques. Derrière l'une d'elles se trouvent 5000 euros et derrière les deux autres rien du tout.
Maryse choisit une des portes (la porte A par exemple). L'animateur, qui connaît la porte gagnante, ouvre une autre porte (disons la C) et lui montre qu'il n'y a rien derrière. Il demande alors à Maryse si elle maintient son choix ou si elle préfère la porte B. Quel choix donne à Maryse la plus grande probabilité de gagner ? Justifiez votre réponse.

Fiche 2 – Variables aléatoires

Exercice 1. Soit $n \in \mathbb{N}^*$. On considère une variable aléatoire X, à valeurs dans $\{1,2,\ldots,n\}$. On suppose que la loi de X est donnée par :

$$\text{pour } k = 1,\ldots,n, \qquad P(X = k) = Ck,$$

où C est une constante.
1. Justifier que l'on peut calculer C. *On rappelle que, pour tout* $m \in \mathbb{N}$, $1 + 2 + \cdots + m = \frac{m(m+1)}{2}$.
2. Pour $k = 1,\ldots,n$, calculer $P(X \leq k)$.
3. Une urne contient $n+1$ boules, numérotées de 0 à n. On tire simultanément deux boules. Quelle est la loi du plus grand des deux numéros tirés ?

Exercice 2. Déterminer si les fonctions suivantes (définies sur $\mathbb{R}$) sont des densités de probabilité, et préciser leur support dans ce cas : *(on rappelle que* $\mathbf{1}_I(x) = 1$ *si* $x \in I$ *et* $\mathbf{1}_I(x) = 0$ *sinon)*

$$f_1(x) = \frac{1}{2}\sin(x)\mathbf{1}_{[0,\pi]}(x), \qquad f_2(x) = \frac{1}{2}\sin(x)\mathbf{1}_{[0,3\pi]}(x), \qquad f_3(x) = -\ln(x)\mathbf{1}_{]0,1[}(x),$$

$$f_4(x) = \frac{1}{1+x^2}, \qquad f_5(x) = \begin{cases} \frac{1}{2}e^x & \text{si } x \leq 0 \\ 0 & \text{si } 0 < x < 1 \text{ ou } x > 2 \\ x - 1 & \text{si } 1 \leq x \leq 2. \end{cases}$$

Exercice 3. Soit $\lambda > 0$. Soit X une variable aléatoire de loi exponentielle de paramètre λ. On note Y sa partie entière (on peut noter $Y = \lfloor X \rfloor$).
1. Quel est l'ensemble des valeurs possibles de Y ? Que faut-il calculer pour se donner la loi de Y ?
2. Déterminer la loi de Y.

Exercice 4. Soit $X_1,X_2,\ldots,X_5$ des variables aléatoires de densités $f_1,f_2,\ldots,f_5$ données dans l'exercice 2 (lorsque ce sont bien des densités).
1. Calculer leurs fonctions de répartition. Les représenter graphiquement.
2. Calculer leurs espérances.
3. Calculer leurs variances.

Exercice 5. Soit T une variable aléatoire, de densité de probabilité f_T donnée par :

$$f_T(t) = \lambda(1 - t^2)\mathbf{1}_{[-1,1]}(t).$$

1. Calculer λ et représenter graphiquement $f_T(t)$.
2. Déterminer la fonction de répartition $F_T(t)$ et tracer sa courbe représentative.
3. Calculer la probabilité de l'événement $\{|T| \geq \frac{1}{2}\}$ et la représenter sur les deux graphiques précédents.
4. Calculer l'espérance et la variance de T. Quelle est sa médiane ? (c'est un réel t tel que $P(T \leq t) = \frac{1}{2}$)

Exercice 6 – ***Recherche d'erreur.*** On sait que parmi les n composants électroniques dont on dispose, 2 sont défaillants, mais on ignore lesquels. On teste les composants un par un en les choisissant uniformément au hasard et en mettant de côté les composants déjà testés.
On note X le nombre de composants qu'il faudra tester jusqu'à trouver le premier composant défaillant (on a donc $X = 1$ si le premier composant testé est l'un des deux défaillants).
1. Que vaut $P(X = 1)$? $P(X = 2)$? $P(X = 3)$?
2. Calculer $P(X = k)$ pour tout $k \in \{1,\ldots,n\}$. On vérifiera que $P(X = k) = \frac{2(n-k)}{n(n-1)}$.

3. Calculer $E[X]$. *On pourra utiliser :* $\sum_{k=1}^{n} k = \frac{n(n+1)}{2}$ *et* $\sum_{k=1}^{n} k^2 = \frac{n(n+1)(2n+1)}{6}$.
4. On s'intéresse à Y, le nombre de composants à tester jusqu'à trouver le second composant défaillant. Imaginons qu'on ait testé les n composants ; on remarque que si on les avait testés en ordre inverse, alors on aurait trouvé le deuxième composant défaillant au bout de $n+1-X$ tests. En déduire que $P(Y=k) = P(X=n+1-k)$ pour $k = 1,\dots,n$, et calculer $E[Y]$.

Exercice 7 – ***Calculs de lois via la fonction de répartition.***
1. Soit X une variable aléatoire de loi exponentielle de paramètre λ. Soit $a > 0$. On pose $Y = aX$.
1.a) Rappeler le calcul la fonction de répartition de X.
1.b) Calculer la fonction de répartition de Y. En déduire la loi de Y.
2. Soit U une variable aléatoire de loi uniforme sur $[0,1]$. Soit $\lambda > 0$. On définit $Y = -\frac{1}{\lambda} \ln U$.
2.a) Que vaut $P(U \leq 0)$? En déduire que Y est bien définie.
2.b) Rappeler le calcul de la fonction de répartition de U.
2.c) Calculer la fonction de répartition de Y. En déduire la loi de Y.
3. Soit X une variable aléatoire de densité

$$f_X : x \mapsto f_X(x) = \frac{1}{\pi(1+x^2)}.$$

On définit $Y = \frac{1}{X}$.
3.a) Que vaut $P(X = 0)$? En déduire que Y est bien définie.
3.b) Calculer la fonction de répartition de X.
3.c) En déduire la fonction de répartition de $Y = \frac{1}{X}$. Quelle est sa loi ?

Exercice 8. Chaque page du site web de l'institut Galilée comporte un grand nombre de caractères, et un gros effort a été fait pour éviter coquilles et autres erreurs. On constate cependant qu'on y trouve en moyenne 1,5 fautes par page. Déterminez la probabilité pour que la page que vous êtes en train de consulter présente au moins trois fautes.

Exercice 9. Le conseil d'administration d'une banque décide d'organiser sa gestion de manière à ce qu'il y ait 999 chances sur 1 000 de pouvoir faire face aux demandes de retrait de ses clients.
La banque a 20 000 clients, qui ont tous 5 000 euros sur leur compte. La probabilité pour qu'un client retire tout son argent un jour donné est 0,005 %. Dans ces conditions, combien la banque doit-elle conserver de liquidités journalières pour suivre le principe de gestion qui a été posé ? (On pourra utiliser une loi de Poisson.)

Exercice 10. Un avion peut transporter 100 passagers et leurs bagages. Il pèse sans les passagers mais avec l'équipage et le carburant 120 tonnes. Les consignes de sécurité interdisent le décollage si le poids de l'appareil dépasse 129,42 tonnes.
Les 100 places ont été occupées. Le poids d'un voyageur suit une loi d'espérance 70 kg et d'écart type 10 kg. Le poids de ses bagages suit une loi d'espérance 20 kg et d'écart type 10 kg. Toutes ces variables sont supposées indépendantes.
1. Calculer l'espérance du poids de l'avion au décollage. Est-elle conforme aux normes de sécurité ?
2. Calculer l'écart type du poids total de l'appareil.
3. En utilisant l'inégalité de Tchebychev, trouver un majorant de la probabilité pour que le poids réel de l'appareil au décollage dépasse 129,42 tonnes.

Exercice 11. Soit X une variable aléatoire de loi $\mathcal{E}(\lambda)$, où $\lambda > 0$ est un réel donné, et ε une variable aléatoire indépendante de X, de loi donnée par

$$P(\varepsilon = +1) = \frac{1}{2} = P(\varepsilon = -1).$$

On définit $Y = \varepsilon X$.
1. Avec quelle probabilité Y est-elle positive ?
2. On souhaite calculer la fonction de répartition F_Y de Y.

2.a) Montrer que, pour $y < 0$, $F_Y(y) = P(\varepsilon = -1)P(X > y)$ et calculer cette valeur.
2.b) Montrer que, pour $y \geq 0$, $F_Y(y) = P(\varepsilon = -1) + P(\varepsilon = +1)P(X < y)$ et calculer cette valeur.
3. En déduire que Y a une densité, et la calculer.

Exercice 12. Philippe joue à un jeu de fléchettes, avec une cible circulaire de rayon r. Chaque lancer qui touche la cible rapporte un nombre de points égal à 3 si la distance au centre est inférieure à $\frac{1}{3}r$, 2 si cette distance est comprise entre $\frac{1}{3}r$ et $\frac{2}{3}r$, et 1 sinon.
Philippe lance sa fléchette sans regarder la cible. On considère qu'il a 3 chances sur 4 de toucher la cible et, s'il la touche, son lancer suit la loi uniforme sur la cible : autrement dit, sachant que la cible est atteinte, la probabilité que la fléchette atteigne une partie A de la cible est proportionnelle à son aire donc vaut $\frac{\text{aire}(A)}{\pi r^2}$.
1. Quelle est la loi du nombre de points obtenus ? Calculer son espérance.
2. Philippe effectue un second lancer, indépendant du premier. Quelle est la probabilité
— qu'il obtienne le même score que la première fois ?
— qu'il obtienne un total de points au moins égal à 5 ?
— qu'il améliore (strictement) son score ?

Exercice 13. Une urne contient des boules, dont une proportion p $(0 < p < 1)$ de boules blanches, les autres boules étant rouges. On tire les boules une à une avec remise, et on note $X_1, X_2, \ldots$ les couleurs obtenues : $X_n = 1$ si le n-ième tirage est blanc, et $X_n = 0$ sinon.
1. Soit S_n la variable aléatoire égale au nombre de boules blanches obtenues au cours des n premiers tirages. Quelle est la loi de S_n ? Rappeler son espérance.
2. Soit N_1 la variable aléatoire égale au rang d'apparition de la première boule blanche. Quelle est la loi de N_1 ? Retrouver pour tout entier $n \geq 1$ la probabilité $P(N_1 = n)$ et calculer l'espérance de N_1.
3. Soit $k > 1$. On note N_k la variable aléatoire égale au rang d'apparition de la k-ième boule blanche.
3.a) Quelles sont les valeurs prises par N_k ?
3.b) Après avoir remarqué que, pour tout $n \geq k$,

$$\{N_k = n\} = \{S_{n-1} = k - 1\} \cap \{X_n = 1\},$$

donner la loi de N_k.
3.c) Calculer l'espérance de N_k *(on pourra dériver k fois la série entière* $\displaystyle\sum_{n=0}^{\infty} x^n = \frac{1}{1-x}$*)*. Commenter.

Exercice 14. Lors d'un salon de l'artisanat, un artisan a un stand où il ne vend que deux produits. Le produit A au prix unitaire de 8 euros, le produit B au prix unitaire de 12 euros.
On suppose dans la suite de l'exercice,
— que les quantités achetées par les différents clients sont indépendantes,
— que le nombre de produits achetés par une personne s'arrêtant au stand est :
 — pour le produit A, une variable aléatoire de Poisson de paramètre 2 notée X,
 — pour le produit B, une variable aléatoire de Poisson de paramètre 1 notée Y,
— que les variables aléatoires X et Y sont indépendantes.
1. Calculer les probabilités, pour une personne s'arrêtant au stand :
— qu'elle n'achète rien
— qu'elle achète uniquement des produits A
— qu'elle achète uniquement des produits B
— qu'elle achète des produits A et des produits B
— qu'elle achète des produits A ou des produits B (ou inclusif).
2. On note $Z = 8X + 12Y$ la variable aléatoire égale au montant en euros de l'achat d'une personne s'arrêtant au stand. Calculer $E[Z]$ et $\text{Var}(Z)$.
3. On a remarqué que 350 personnes en moyenne s'arrêtaient au stand chaque jour. Soit W la variable aléatoire égale au chiffre d'affaire sur une journée de l'artisan.
3.a) Déterminer la moyenne et la variance de W.
3.b) En utilisant l'inégalité de Tchebychev, trouver un minorant de $P(9000 < W < 10600)$.

3.c) On suppose que la loi de W peut être approchée par une loi normale de moyenne $m = 9800$ et d'écart type $\sigma = 309$.
• Quelle est la loi suivie par la variable aléatoire $W^* = \frac{W-9800}{309}$?
• Avec la table de la loi gaussienne centrée réduite, trouver, si $Z \sim \mathcal{N}(0{,}1)$, $P(Z \leq 0{,}16)$ et $P(Z \leq 0{,}49)$.
• En déduire la probabilité pour que le chiffre d'affaire W d'une journée soit compris entre 9750 et 9951 euros.

Exercices complémentaires

Exercice 15. Dans un jeu, on lance deux dés à six faces, et on obtient :

- 1 point si au moins un dé est impair
- 2 point la somme des dés est supérieure ou égale à 10
- 3 point si les deux résultats sont égaux

et ces points sont cumulables. Calculer la loi du nombre de points, son espérance et sa variance.

Exercice 16. Soit $\lambda,\mu > 0$. Soit X,Y deux variables aléatoires indépendantes, de loi $\mathcal{P}(\lambda)$ et $\mathcal{P}(\mu)$. On définit

$$Z = X + Y.$$

1. Pour tout $n \in \mathbb{N}$, calculer $P(Z = n)$. Déterminer la loi de Z.
2. Pour tous $0 \leq k \leq n$, calculer $P(X = k \,|\, Z = n)$.

Exercice 17 – *Ampoules.* À cause des imperfections dans les matières premières et dans le processus de fabrication, les ampoules à filament produites par une usine présentent des caractéristiques variables d'une ampoule à l'autre.
La durée de vie T d'une ampoule suit une loi exponentielle, et a une durée de vie moyenne de 5 mois.
1. Quelle est la probabilité que l'ampoule dure moins de 3 mois ?
2. On a acheté un lot de 5 ampoules du type précédent. Leurs durées de vie sont supposées indépendantes. Si on les utilise toutes à partir du même moment, quelle est la probabilité qu'il faille en remplacer exactement 2 avant 3 mois ?
3. La luminosité L et la durée de vie T d'une ampoule sont approximativement reliées par la relation

$$T = L^{-4}$$

(où la luminosité est exprimée dans une unité appropriée).
3.a) Quelles sont les valeurs possibles de L ?
3.b) Calculer la fonction de répartition F_L de L.
3.c) Justifier que L admet une densité et la calculer.

Exercice 18. Une urne contient N boules, dont S boules vertes. On tire simultanément n boules de l'urne, et on note X le nombre de boules vertes parmi celles-ci.
1. Montrer que X suit la loi hypergéométrique :

$$\text{pour } 0 \leq k \leq n, \quad P(X = k) = \frac{C_S^k \, C_{N-S}^{n-k}}{C_N^n}.$$

2. Vérifier que c'est bien une probabilité. Pour cela, on pourra développer les deux membres de l'identité :

$$(1 + x)^N = (1 + x)^S \, (1 + x)^{N-S}$$

et comparer les coefficients.
3. Soit $p \in [0{,}1]$. Quelle est la limite de $P(X = k)$ lorsque $N \to \infty$ et $S \to \infty$ avec $\frac{S}{N} \to p$, et n fixé ? Expliquer le résultat intuitivement.
4. $\star$ Calculer la moyenne $E[X]$ et la variance $\mathrm{Var}(X)$. On montrera que :

$$E[X] = np \quad \text{et} \quad \mathrm{Var}(X) = np(1-p)\,\frac{N-n}{N-1}$$

où p désigne la proportion de boules vertes : $S = pN$.

Université Paris 13, Institut Galilée — CCS – Probabilités & Statistiques

Année universitaire 2018–2019

Fiche 3 – Couples de variables aléatoires

Exercice 1. Un atelier fonctionne avec deux équipes d'ouvriers, une du matin et l'autre du soir. Chaque jour on enregistre le nombre d'ouvriers absents. On note X (respectivement Y) le nombre d'absences dans l'équipe de jour (respectivement de nuit). La loi jointe de (X,Y) est donnée par

X \ Y	0	1	2	3
0	c	$2c$	0	$3c$
1	0	$3c$	$4c$	0
2	0	$4c$	$2c$	c

1. Déterminer la constante c.
2. Donner les lois marginales de X et Y ainsi que leurs espérances.
3. Une absence coûte 30 € à l'usine. Quelle est la perte journalière moyenne due aux absences ?

Exercice 2. On effectue deux tirages avec remise dans une urne contenant des boules numérotées de 1 à 3. On note X le plus petit numéro, et Y le plus grand.
1. Calculer la loi jointe de (X,Y), ainsi que les lois de X et de Y.
2. Calculer la covariance de X et Y.

Exercice 3. Soit D le domaine du plan

$$D = \{(x,y) \mid x > 0,\, y > 0,\, y^2 > x\}.$$

1. Dessiner D.
2. Soit (X,Y) un couple de variables aléatoires à valeurs réelles de densité de probabilité

$$f(x,y) = \begin{cases} \frac{1}{2\sqrt{x}} e^{-y} & \text{si } (x,y) \in D \\ 0 & \text{ailleurs.} \end{cases}$$

2.a) Déterminer les lois de X et Y, et leurs espérances. *Aide : pour* $n \in \mathbb{N}$, $\int_0^\infty y^n e^{-y} dy = n!$
2.b) Les variables aléatoires X et Y sont-elles indépendantes ?
2.c) Calculer $\mathrm{Cov}(X,Y)$.

Exercice 4. Pour rentrer chez lui, un étudiant a le choix entre la ligne 361, qui passe toutes les 4 minutes, et la ligne 356 qui passe toutes les 9 minutes. On suppose que les bus passent à intervalle régulier, et que les lignes sont indépendantes entre elles. On cherche le temps moyen d'attente avant un bus quelconque, quand on arrive à l'arrêt sans connaître les horaires.
On note T_4 le temps d'attente du 361, et T_9 le temps d'attente du 356. Le temps d'attente d'un bus quelconque est donc

$$T = \min(T_4,T_9).$$

1. Quelles lois suivent T_4 et T_9 ? Quel est l'ensemble des valeurs possibles de T ?
2. Calculer, pour tout $t \in \mathbb{R}$, $P(T > t)$ à l'aide des fonctions de répartition de T_4 et T_9.
3. En déduire la fonction de répartition de T, puis sa loi.
4. Répondre à la question initiale.

Exercice 5. Soit X,Y deux variables aléatoires indépendantes, de densités f_X et f_Y. On pose

$$Z = X + Y.$$

1. Rappeler la densité f_Z de la variable Z.
2. Soit $\lambda > 0$. Calculer f_Z lorsque X et Y suivent la loi $\mathcal{E}(\lambda)$.
3. Soit $\lambda > 0$. Calculer f_Z lorsque X suit la loi $\mathcal{E}(\lambda)$ et Y suit la loi $\mathcal{U}([0,1])$.

Exercice 6. Soit $\lambda > 0$. Soit X,Y deux variables aléatoires indépendantes, de loi $\mathcal{E}(\lambda)$. On pose

$$Z = \frac{X}{Y}.$$

1. Justifier que Z est bien définie. Quel est l'ensemble des valeurs possibles de Z ?
2. Montrer que, pour tout $t > 0$,

$$P(Z \leq t) = \int_0^\infty \left(\int_0^{ty} \lambda e^{-\lambda x} dx \right) \lambda e^{-\lambda y} dy,$$

et calculer cette intégrale double.
3. En déduire que Z admet une densité et la calculer.
4. Procéder de même pour calculer la densité de $U = \frac{X}{X+Y}$.

Exercices complémentaires

Exercice 7. Soit (X,Y) un couple de variables aléatoires de densité

$$f(x,y) = \frac{1}{2x} \quad \text{si } (x,y) \in D, \quad f(x,y) = 0 \quad \text{sinon}$$

avec

$$D = \left\{ (x,y) \;\middle|\; 0 < y \leq x \quad \text{et} \quad 0 < y \leq \frac{1}{x} \right\}.$$

1. Donner une représentation graphique de D.
2. Déterminer les densités des variables aléatoires X et Y. Les variables aléatoires X et Y sont-elles indépendantes ?

Exercice 8. Deux personnes se donnent rendez-vous entre 12h et 13h. On suppose que l'on peut modéliser les instants d'arrivée de ces deux personnes par des variables aléatoires X et Y, indépendantes et de loi uniforme sur $[0,1]$. On souhaite étudier la variable aléatoire $U = |X - Y|$, qui est égale au temps d'attente de la première personne arrivée.
1. Quelle est la densité du couple (X,Y) ? Pour $D \subset \mathbb{R}^2$, que vaut donc $P((X,Y) \in D)$?
2. Montrer que la fonction de répartition de U est donnée par

$$F_U(u) = \begin{cases} 0 & \text{si } u < 0 \\ D(u) & \text{si } u \in [0,1] \\ 1 & \text{si } u > 1, \end{cases}$$

où, pour $0 \leq u \leq 1$, $D(u)$ est l'aire de l'ensemble $\{(x,y) \in [0,1]^2 \,|\, |x - y| \leq u\}$, que l'on calculera.
3. En déduire la densité de U et son espérance $E[U]$.
4. Ces deux personnes conviennent que la première arrivée s'en ira après un temps d'attente égal à $2E[U]$. Quelle est la probabilité pour que le rendez-vous ait lieu ?

Fiche 4 – Statistiques descriptives, Estimateurs

Exercice 1. Chaque jour, on relève le nombre de clients entrés dans une boutique entre 9h et 9h30. Sur une période de 100 jours, on obtient les données suivantes :

4, 6, 6, 0, 6, 6, 3, 6, 2, 5, 5, 7, 9, 5, 4, 8, 5, 7, 3, 1, 6, 7, 3, 7, 5, 9, 6, 4, 8, 4, 9, 5, 5, 3, 3, 6, 3, 3, 3, 7, 7, 11, 3, 4, 7, 4, 3, 1, 8, 3, 4, 8, 3, 7, 3, 9, 6, 7, 5, 11, 4, 2, 7, 7, 5, 5, 3, 6, 8, 6, 4, 6, 6, 4, 6, 5, 4, 3, 5, 5, 2, 4, 2, 4, 5, 5, 7, 4, 9, 9, 3, 3, 5, 4, 5, 4, 1, 6, 4, 4.

1. Calculer les effectifs de chaque valeur, et leurs fréquences.
2. Calculer le nombre moyen de clients, et sa variance.
3. Calculer minimum, maximum, médiane, premier et troisième quartiles de cette série statistique.
4. Représenter les données par un diagramme en bâtons et un diagramme circulaire.

Exercice 2. On a relevé la température chaque jour du mois de mars à la même heure : en °C,

4,1 ; 3,6 ; 5,5 ; 9,8 ; 9,1 ; 3,9 ; 6,1 ; 5,3 ; 10,4 ; 4,6 ; 6,5 ; 2,3 ; 9,0 ; 5,5 ; 8,1 ; 7,0 ; 8,5 ; 5,9 ; 4,4 ; 9,1 ; 5,0 ; 2,3 ; 5,9 ; 6,4 ; 4,6 ; 4,3 ; 6,7 ; 6,4 ; 4,3 ; 5,7 ; 4,5.

1. Calculer la température moyenne, et sa variance.
2. Calculer minimum, maximum, médiane, premier et troisième quartiles de cette série statistique.
3. Représenter les températures par un histogramme, en les regroupant par classes de largeur 1°C.

Exercice 3. On cherche à retrouver la 3e loi de Kepler, qui relie le demi-grand axe R de l'orbite d'une planète (distance maximale au soleil) à sa période de révolution T (durée d'une rotation autour du soleil), en sachant que cette relation est de la forme $T = CR^\alpha$ où $C,\alpha \in \mathbb{R}$. On dispose des données approximatives suivantes (1 UA = 149 600 000 km) :

Planète	Demi-grand axe (en UA)	Période (en années)
Mercure	0,39	0,24
Vénus	0,72	0,61
Terre	1	1
Mars	1,5	1,9
Jupiter	5,2	12
Saturne	9,5	29
Uranus	19	84
Neptune	30	165

On note $r_1,\dots,r_8$ les demi-grand axes et $t_1,\dots,t_8$ les périodes ci-dessus.
1. Comment déduire de l'équation $T = CR^\alpha$ une relation de la forme $Y = aX + b$?
2. Calculer dans un tableau les logarithmes $x_i = \ln r_i$ et $y_i = \ln t_i$, et représenter graphiquement les points (x_i,y_i).
3. Effectuer une régression linéaire entre x_i et y_i et conjecturer alors une relation entre la période T et le demi-grand axe R d'une planète du système solaire. Calculer la corrélation $\rho_{x,y}$ et commenter.
4. La période de Pluton est de 248 ans. Que peut-on prédire de son demi-grand axe ?

Exercice 4. Soit T_n un estimateur d'un paramètre θ. Montrer que si T_n est asymptotiquement sans biais et que sa variance tend vers 0, alors T_n est convergent. *Utiliser l'inégalité de Markov.*

Exercice 5. Soit X une variable aléatoire suivant une loi de Poisson de paramètre $\lambda > 0$. On considère un échantillon $(X_1,X_2,\ldots,X_n)$ de X.
1. Rappeler $E[X]$ et $\mathrm{Var}(X)$. En déduire deux estimateurs de λ par la méthode des moments.
2. On cherche un estimateur de λ par la méthode du maximum de vraisemblance.
2.a) Écrire la vraisemblance $L(k_1,k_2,\ldots,k_n;\lambda)$ de l'échantillon $(X_1,X_2,\ldots,X_n)$.
2.b) En déduire que l'estimateur T_n de λ par maximum de vraisemblance est égal à $\overline{X}_n$.

Exercice 6. Soit X une variable aléatoire admettant pour densité

$$f(x) = \frac{1}{\theta} x^{\frac{1}{\theta}-1} \mathbf{1}_{]0,1]}(x),$$

où $\theta > 0$ est un paramètre. Soit $(X_1,\ldots,X_n)$ un échantillon de X.
1. Calculer $E[X]$ et $\mathrm{Var}(X)$. En déduire un estimateur de θ par la méthode des moments.
2. On pose $Y = \ln X$. Quel est l'ensemble des valeurs possibles de Y ? Déterminer, sa fonction de répartition et sa densité, et calculer $E[Y]$.
3. On cherche un estimateur par la méthode du maximum de vraisemblance.
3.a) Écrire la vraisemblance $L(x_1,x_2,\ldots,x_n;\theta)$ de l'échantillon pour θ.
3.b) En déduire un estimateur $\widehat{\theta}$ de θ par la méthode du maximum de vraisemblance, et montrer que cet estimateur est sans biais.

Exercice 7. On dispose d'un échantillon $(X_1,\ldots,X_n)$ d'une variable aléatoire X admettant pour densité

$$f(x) = \begin{cases} \frac{3x^2}{\theta^3} & \text{si } 0 < x < \theta \\ 0 & \text{ailleurs.} \end{cases}$$

1. Calculer $E[X]$, en déduire un estimateur T_1 de θ par la méthode des moments. Quel est son biais ?
2. On définit $W = \max(X_1,\ldots,X_n)$.
2.a) Calculer la vraisemblance $L(x_1,\ldots,x_n;\theta)$, et montrer que W est l'estimateur de θ donné par la méthode du maximum de vraisemblance.
2.b) Calculer la fonction de répartition de W, puis sa densité et son espérance. En déduire un estimateur T_2 sans biais de θ sous la forme $T_2 = cW$, où c est une constante.
3. Comparer T_2 à T_1, au sens du risque quadratique.

Exercice 8. Soit X une variable aléatoire de loi uniforme sur $[0,\theta]$, où $\theta > 0$ est un paramètre. Soit $(X_1,X_2,\ldots,X_n)$ un échantillon de X. Vérifier que :
1. la méthode des moments fournit l'estimateur $T_1 = 2\overline{X}_n$ de θ, et que T_1 est un estimateur sans biais et convergent ;
2. la méthode du maximum de vraisemblance fournit l'estimateur $T_2 = \max(X_1,X_2,\ldots,X_n)$ de θ, et que T_2 est un estimateur asymptotiquement sans biais et convergent.

Exercice 9. Soit X une variable aléatoire dont la loi dépend de deux paramètres p_1 et p_2 par :

$$P(X = 0) = 1 - p_1 - p_2, \quad P(X = 1) = p_1 \quad \text{et} \quad P(X = 2) = p_2.$$

1. Trouver les conditions à vérifier par p_1 et p_2 pour que le support de la loi de X soit égal à $\{0,1,2\}$.
2. Calculer $E[X]$, $E[X^2]$, $\mathrm{Var}(X)$.
3. Soit $(X_1,X_2,\ldots,X_n)$ un échantillon de X. Déterminer des estimateurs L_1 et L_2 de p_1 et p_2 par la méthode des moments. Montrer qu'ils sont sans biais.
4. Pour $j = 0, 1, 2$, on désigne par N_j le nombre de X_k égaux à j. Écrire la vraisemblance $L(x_1,\ldots,x_n;p_1,p_2)$ de l'échantillon en fonction de p_1, p_2, n_0, n_1, n_2 (où n_j est le nombre de x_k égaux à j). Déterminer les estimateurs Z_1 et Z_2 de p_1 et p_2 par la méthode du maximum de vraisemblance.
5. Montrer que $L_1 = Z_1$ et $L_2 = Z_2$.
6. Un échantillon de taille $n = 100$ de X a donné les observations suivantes :

$$n_0 = 20, \quad n_1 = 50, \quad n_2 = 30 \ .$$

À quelles estimations de p_1 et p_2 conduisent les estimateurs L_1 et L_2 ?

Fiche 5 – Intervalles de confiance, tests d'hypothèse

Exercice 1. Soit Y une variable aléatoire de Bernoulli de paramètre p et $\overline{Y}_n$ la variable aléatoire

$$\overline{Y}_n = \frac{1}{n}(Y_1 + Y_2 + \cdots + Y_n),$$

où $(Y_1, \ldots, Y_n)$ est un échantillon de Y, c'est-à-dire que $Y_1, \ldots, Y_n$ sont des v.a. indépendantes, de même loi que Y.
1. À l'aide du théorème central limite, définir à partir de $\overline{Y}_n$ une variable aléatoire Z_n qui suit approximativement la loi $\mathcal{N}(0,1)$ si n est suffisamment grand.
2. À l'aide de la table de la loi gaussienne,
2.a) trouver un réel A tel que $P(Z_n > A) \simeq 0{,}05$;
2.b) trouver un réel α tel que $P(-\alpha \leq Z_n \leq \alpha) \simeq 0{,}95$.
3. On suppose $n = 100$.
3.a) Si $p = 0{,}1$, trouver un réel a tel que $P(\overline{Y}_{100} > a) \simeq 0{,}05$.
3.b) Trouver un réel ε tel que $P\Big(\overline{Y}_{100} - \varepsilon\sqrt{\overline{Y}_{100}(1-\overline{Y}_{100})} \leq p \leq \overline{Y}_{100} + \varepsilon\sqrt{\overline{Y}_{100}(1-\overline{Y}_{100})}\Big) \simeq 0{,}95$.
3.c) Justifier que $P\big(\overline{Y}_{100} - \frac{\varepsilon}{2} \leq p \leq \overline{Y}_{100} + \frac{\varepsilon}{2}\big) \geq 0{,}95$.
4. Application. Un fabricant de vis prélève 100 vis dans sa production, et trouve que 10 sont défectueuses.
4.a) Estimer la proportion p de vis défectueuses dans toute la production.
4.b) Donner un intervalle de confiance de niveau 95 % pour p.
5. Application. Un acheteur veut tester les livraisons de vis qu'il reçoit et savoir si elles sont conformes à ce que dit le fabricant. Le fabricant assure que moins de 10 % des vis sont défecteuses. L'acheteur prélève 100 vis de sa livraison et relève la proportion π de vis défectueuses.
Pour quelles valeurs de π l'acheteur peut-il légitimement refuser la livraison au seuil de risque de 5 % ?

Exercice 2. La durée de fonctionnement d'une ampoule suit une loi normale $\mathcal{N}(m,\sigma^2)$. On examine 50 ampoules et on enregistre les résultats $x_1 = X_1(\omega), \ldots, x_{50} = X_{50}(\omega)$. Ils vérifient :

$$\overline{x}_{50} = \overline{X}_{50}(\omega) = 58 \text{ jours} \quad \text{et} \quad \sigma_{50}^2 = S_{50}^2(\omega) = \frac{1}{50}\sum_{i=1}^{50}\big(x_i - 58\big)^2 = 99 \text{ jours}^2.$$

1. Calculer dans ce cas les bornes d'un intervalle de confiance I de niveau 95 % de la moyenne m.
2. Énoncer un test, au seuil de risque 5 % de l'hypothèse $\mathcal{H}_0$: « $m = 65$ jours » contre $\mathcal{H}_1$: « $m \neq 65$ jours ». Dans le cas présent, quelle est la conclusion du test ?

Exercice 3. Un revendeur a observé que les chargeurs de téléphone contrefaits ont, selon l'exemplaire, un temps de charge (en heures) distribué selon la loi $\mathcal{N}(6,2)$, tandis que les chargeurs authentiques ont un temps de charge de loi $\mathcal{N}(4,1)$. Autrement dit, si on note X le temps de charge d'un chargeur contrefait sélectionné au hasard de façon uniforme, alors X suit la loi $\mathcal{N}(6,2)$; et de même le temps de charge Y d'un chargeur authentique choisi au hasard suit la loi $\mathcal{N}(4,1)$.
1. Calculer la probabilité pour qu'un chargeur authentique ait un temps de charge supérieur à 6 heures.
2. Pour décider si un chargeur est authentique, on mesure son temps de charge.
Soit $\mathcal{H}_0$ l'hypothèse : « le chargeur est authentique » et $\mathcal{H}_1$ l'hypothèse alternative : « le chargeur est une contrefaçon ».
On décide de rejeter l'hypothèse $\mathcal{H}_0$ si le temps de charge est supérieur à $A = 6$ heures.
2.a) Qu'est-ce que l'erreur de 1$^{\text{ère}}$ espèce ? Calculer sa probabilité α. (α est le *seuil de risque* du test)
2.b) Qu'est ce que l'erreur de 2$^{\text{e}}$ espèce ? Calculer sa probabilité β. ($1-\beta$ est la *puissance* du test)
2.c) On souhaite un risque de 1$^{\text{ère}}$ espèce égal à 5 %. Comment choisir A ? Que vaut alors β ?

Exercice 4.
1. Un grossiste reçoit des pelotes de laine provenant de deux usines différentes. Toutes les pelotes d'une même usine n'ont pas exactement le même poids, mais le poids (en grammes) d'une pelote prise au hasard dans un lot de même provenance suit une loi normale $\mathcal{N}(m,\sigma^2)$. Les usines A et B annoncent respectivement

$$m_A = 100\,\mathrm{g}, \quad \sigma_A = 8\,\mathrm{g} \quad ; \quad m_B = 96\,\mathrm{g}, \quad \sigma_B = 4\,\mathrm{g}.$$

1.a) Un premier client refuse toute pelote dont le poids est inférieur à 90 g. De quelle usine provient le lot donnant le plus faible pourcentage de rebut ?
1.b) Pour un autre client, le poids minimum exigé est inconnu ; mais on constate qu'il est tel que le choix entre les deux provenances est indifférent. Déterminer ce poids minimum exigé ainsi que le pourcentage de rebut (commun au deux lots).
2. Le grossiste de son côté souhaite vérifier si les lots reçus sont conformes aux qualifications annoncées (ni plus, ni moins). Pour cela il prélève un échantillon de 100 pelotes dans le lot provenant de l'usine A et détermine le poids moyen $\overline{x}_A$ d'une pelote de cet échantillon. On suppose qu'il connaît déjà la valeur de $\sigma_A(= 8\,\mathrm{g})$.
2.a) Soient $(X_1,X_2,\ldots,X_{100})$, 100 variables aléatoires indépendantes de même loi $\mathcal{N}(m_A,\sigma_A^2)$. On note $\overline{X}_{100}$ la moyenne de ces variables aléatoires :

$$\overline{X}_{100} = \frac{1}{100}\sum_{i=1}^{100} X_i.$$

Définir à partir de $\overline{X}_{100}$ une variable aléatoire Y dont la loi est approximativement $\mathcal{N}(0,1)$.
2.b) Quelles sont les valeurs de $\overline{x}_A$ pour lesquelles il pourra rejeter, au seuil de risque 5 %, l'hypothèse $\mathcal{H}_0$: « $m_A = 100\,\mathrm{g}$ » ?
2.c) On ne suppose plus que le grossiste connaît σ_A. Par contre, il a calculé, sur son échantillon, l'écart-type empirique

$$\sigma_{100} = \sqrt{\frac{1}{99}\sum_{i=1}^{100}\left(X_i(\omega) - \overline{x}_A\right)^2} \simeq 8{,}5\,\mathrm{g}.$$

Répondre à nouveau à la question précédente.
2.d) On suppose maintenant que le grossiste veut seulement s'assurer que les pelotes ne sont excessivement légères : il veut tester $\mathcal{H}_0$ contre $\mathcal{H}_1$: « $m_A < 100$ g ». Répondre de nouveau à la question.

Exercice 5. Un directeur de centre de loisirs s'intéresse à l'évolution de la proportion de personnes choisissant la planche à voile par rapport à l'an dernier.
L'an dernier, $\pi_0 = 12\,\%$ des personnes fréquentant le centre de loisirs avaient choisi la planche à voile. Pour cette année, il dispose seulement du pourcentage $\overline{x}_n = 16\,\%$ des personnes ayant choisi la planche à voile parmi les n premières venues en juillet. Il souhaite estimer avec un risque de 10 % si $\overline{x}_n$ est conforme aux observations de l'année passée, c'est-à-dire tester l'hypothèse $\mathcal{H}_0$: « $\pi_1 = \pi_0$ », où π_1 sera la proportion de cette année, contre l'hypothèse $\mathcal{H}_1$: « $\pi_1 \neq \pi_0$ ».
On définit n variables aléatoires indépendantes et de même loi $(X_i)_{1\leq i\leq n}$ par :

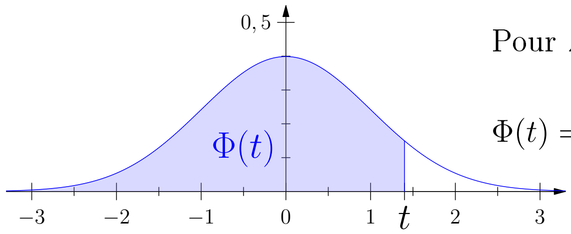
$$X_i = \begin{cases} 1 & \text{si la } i\text{-ième personne choisit la planche à voile} \\ 0 & \text{si la } i\text{-ième personne choisit un autre sport.} \end{cases}$$

1. On se place sous l'hypothèse que la proportion de personnes optant pour la planche à voile reste π_0.
1.a) Que valent $E[X_i]$, $\mathrm{Var}(X_i)$, $E[\overline{X}_n]$ et $\sigma(\overline{X}_n)$?
1.b) Déterminer en fonction de n un intervalle I_n centré en π_0 et un réel A_n tels que

$$P(\overline{X}_n \in I_n) = 90\,\% \qquad \text{et} \qquad P(\overline{X}_n \geq A_n) = 10\,\%.$$

2. Que vaut I_n lorsque $n = 30$? Le pourcentage $\overline{x}_n$ est-il donc significativement différent, au risque de 10 %, du pourcentage π_0 obtenu l'année dernière ? (c.-à-d. que l'on peut rejeter $\mathcal{H}_0$ et accepter $\mathcal{H}_1$) Pour quelles valeurs de n aurait-on pu le conclure ?
3. Si le nombre d'inscrits augmente par rapport à l'an passé, il faudra avoir recruté un nouveau moniteur. On considère donc plus précisément l'hypothèse alternative $\mathcal{H}_1$: « $\pi_1 > \pi_0$ ». Que vaut A_n lorsque $n = 30$? Le pourcentage $\overline{x}_n$ est-il donc significativement supérieur, au risque de 10 %, au pourcentage π_0 obtenu l'année dernière ? Pour quelles valeurs de n aurait-on pu le conclure ?

Table de la fonction de répartition de la loi normale centrée réduite $\mathcal{N}(0,1)$



Pour Z de loi $\mathcal{N}(0,1)$,

$$\Phi(t) = F_Z(t) = P(Z \leq t) = \int_{-\infty}^{t} e^{-\frac{x^2}{2}} \frac{dx}{\sqrt{2\pi}}$$

t	0,00	0,01	0,02	0,03	0,04	0,05	0,06	0,07	0,08	0,09
0,0	0,5000	0,5040	0,5080	0,5120	0,5160	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	0,5753
0,2	0,5793	0,5832	0,5871	0,5910	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	0,6480	0,6517
0,4	0,6554	0,6591	0,6628	0,6664	0,6700	0,6736	0,6772	0,6808	0,6844	0,6879
0,5	0,6915	0,6950	0,6985	0,7019	0,7054	0,7088	0,7123	0,7157	0,7190	0,7224
0,6	0,7257	0,7291	0,7324	0,7357	0,7389	0,7422	0,7454	0,7486	0,7517	0,7549
0,7	0,7580	0,7611	0,7642	0,7673	0,7704	0,7734	0,7764	0,7794	0,7823	0,7852
0,8	0,7881	0,7910	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	0,8340	0,8365	0,8389
1,0	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	0,8770	0,8790	0,8810	0,8830
1,2	0,8849	0,8869	0,8888	0,8907	0,8925	0,8944	0,8962	0,8980	0,8997	0,9015
1,3	0,9032	0,9049	0,9066	0,9082	0,9099	0,9115	0,9131	0,9147	0,9162	0,9177
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	0,9370	0,9382	0,9394	0,9406	0,9418	0,9429	0,9441
1,6	0,9452	0,9463	0,9474	0,9484	0,9495	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	0,9750	0,9756	0,9761	0,9767
2,0	0,9772	0,9778	0,9783	0,9788	0,9793	0,9798	0,9803	0,9808	0,9812	0,9817
2,1	0,9821	0,9826	0,9830	0,9834	0,9838	0,9842	0,9846	0,9850	0,9854	0,9857
2,2	0,9861	0,9864	0,9868	0,9871	0,9875	0,9878	0,9881	0,9884	0,9887	0,9890
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	0,9920	0,9922	0,9925	0,9927	0,9929	0,9931	0,9932	0,9934	0,9936
2,5	0,9938	0,9940	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	0,9960	0,9961	0,9962	0,9963	0,9964
2,7	0,9965	0,9966	0,9967	0,9968	0,9969	0,9970	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	0,9980	0,9981
2,9	0,9981	0,9982	0,9982	0,9983	0,9984	0,9984	0,9985	0,9985	0,9986	0,9986
3,0	0,9987	0,9987	0,9987	0,9988	0,9988	0,9989	0,9989	0,9989	0,9990	0,9990
3,1	0,9990	0,9991	0,9991	0,9991	0,9992	0,9992	0,9992	0,9992	0,9993	0,9993
3,2	0,9993	0,9993	0,9994	0,9994	0,9994	0,9994	0,9994	0,9995	0,9995	0,9995
3,3	0,9995	0,9995	0,9995	0,9996	0,9996	0,9996	0,9996	0,9996	0,9996	0,9997
3,4	0,9997	0,9997	0,9997	0,9997	0,9997	0,9997	0,9997	0,9997	0,9997	0,9998
3,5	0,9998	0,9998	0,9998	0,9998	0,9998	0,9998	0,9998	0,9998	0,9998	0,9998
3,6	0,9998	0,9998	0,9999	0,9999	0,9999	0,9999	0,9999	0,9999	0,9999	0,9999
3,7	0,9999	0,9999	0,9999	0,9999	0,9999	0,9999	0,9999	0,9999	0,9999	0,9999
3,8	0,9999	0,9999	0,9999	0,9999	0,9999	0,9999	0,9999	0,9999	0,9999	0,9999
3,9	1,0000	1,0000	1,0000	1,0000	1,0000	1,0000	1,0000	1,0000	1,0000	1,0000

Exemples : $\Phi(0{,}25) \simeq 0{,}5987$, $\quad \Phi(-0{,}32) = 1 - \Phi(0{,}32) \simeq 1 - 0{,}6255 = 0{,}3745$

Memento des lois usuelles

Lois discrètes

Nom	Paramètres	Support	Définition : $P(A) = \sum_{a \in A} p(a)$	Espérance	Variance
Loi de Dirac δ_a	$a \in \mathbb{R}$	$\{a\}$	$p(a) = 1$	a	0
Loi de Bernoulli $\mathcal{B}(p)$	$p \in [0,1]$	$\{0,1\}$	$p(0) = 1-p$, $p(1) = p$	p	$p(1-p)$
Loi binomiale $\mathcal{B}(n,p)$	$n \in \mathbb{N}$, $p \in [0,1]$	$\{0,\dots,n\}$	$p(k) = \binom{n}{k} p^k (1-p)^{n-k}$	np	$np(1-p)$
Loi géométrique $\mathcal{G}(p)$	$p \in]0,1]$	$\mathbb{N}^*$	$p(k) = (1-p)^{k-1} p$	$\dfrac{1}{p}$	$\dfrac{1-p}{p^2}$
Loi de Poisson $\mathcal{P}(\lambda)$	$\lambda \in]0,+\infty[$	$\mathbb{N}$	$p(k) = e^{-\lambda} \dfrac{\lambda^k}{k!}$	λ	λ

Lois continues

Nom	Paramètres	Support	Définition : $P(A) = \int_A f(x)dx$	Espérance	Variance
Loi uniforme $\mathcal{U}([a,b])$	$a < b$	$[a,b]$	$f(x) = \dfrac{1}{b-a} \mathbf{1}_{[a,b]}(x)$	$\dfrac{a+b}{2}$	$\dfrac{(b-a)^2}{12}$
Loi exponentielle $\mathcal{E}(\lambda)$	$\lambda \in]0,\infty[$	$]0,+\infty[$	$f(x) = \lambda e^{-\lambda x} \mathbf{1}_{]0,+\infty[}(x)$	$\dfrac{1}{\lambda}$	$\dfrac{1}{\lambda^2}$
Loi normale/gaussienne $\mathcal{N}(m,\sigma^2)$	$m \in \mathbb{R}$, $\sigma^2 \in]0,+\infty[$	$\mathbb{R}$	$f(x) = \dfrac{1}{\sqrt{2\pi\sigma^2}} \exp\left(-\dfrac{(x-m)^2}{2\sigma^2}\right)$	m	σ^2