.

# Probabilités et Statistiques (2)

Le présent document fournit le support des notions abordées en *Probabilités* du cours PRBS2.

## 1 Rappels synthétiques des notions antérieures

Le symbole $\nabla$ mentionne les notions n'étant pas exigibles mais dont la connaissance peut permettre une meilleure compréhension.

### 1.1 Expériences Aléatoires

On admet que les étudiants comprennent la locution *Expérience Aléatoire*. On considérera comme expériences aléatoires :

- Les lancés de dés, peu important leur nombre de faces, ce que l'on y inscrit ou s'ils sont (ou non) pipés
- Le fait de tirer une (ou plusieurs) carte(s) d'un jeu mélangé
- Tirer une (ou plusieurs) boule(s) indiscernables au toucher depuis une urne
- Selectionner aléatoirement un individu d'une population
- Le résultat obtenu lors de l'utilisation d'une commande de type `RAND()`.

Aucune discussion ne sera développée dans ce cours sur le sens de *Aléatoire* et on admettra simplement que les outils mathématiques abordés en permettent une description fiable et commode.
On pourra employer le terme **pseudo-aléatoire** pour distinguer un phénomène étudié par la théorie des probabilités et pour lequel le caractère purement *alétoire* peut être remis en question.

### 1.2 Vocabulaire des modèles d'expériences aléatoires

Le présent vocabulaire est supposé être en grande partie connu :

☐ **Issue** : Toute possibilité observable résultant d'une expérience aléatoire.

☐ **Univers** : Un univers $\mathcal{U}$ est l'ensemble des issues d'une expérience aléatoire. Il dépend donc de l'expérience considérée. *On utilise aussi souvent la lettre $\Omega$ pour désigner un univers dans ce contexte*.

☐ **Evénement** : Partie $E$ de l'univers $\mathcal{U}$ recevable dans l'étude.

☐ **Evénement atomique (ou élémentaire)** : Evénement à une seule issue (un singleton).

☐ **Evénement impossible :** L'ensemble $\emptyset$ dans le contexte des expériences aléatoires.

☐ **Evénements incompatibles :** Deux événements $A$ et $B$ vérifiant $A \cap B = \emptyset$
On pourra dire *compatibles* dans le cas contraire.

☐ **Evénement contraire :** L'ensemble $A^c$ noté $\overline{A}$ dans le contexte des probabilités modélise l'événement contraire de $A$.

$\nabla$ **Tribu $\mathcal{A}$ des événéments :** Ensemble $\mathcal{A} \subset \mathcal{P}(\Omega)$ de tous les événements associés à l'expérience d'univers $\Omega$. Elle doit vérifier les propriétés axiomatiques suivantes :

- $\emptyset \in \mathcal{A}$ et $\Omega \in \mathcal{A}$
- Stabilité par unions dénombrables :

$$\forall (A_n)_{n\in\mathbb{N}} \in \mathcal{A}^{\mathbb{N}} \quad \bigcup_{n\in\mathbb{N}} A_n \in \mathcal{A}$$

c'est-à-dire que toute union d'une suite d'événements est un événement

- Stabilité par passage au complémentaire :

$$\forall A \in \mathcal{P}(\Omega) \quad A \in \mathcal{A} \Rightarrow \overline{A} \in \mathcal{A}$$

c'est-à-dire que tout complémentaire d'un événement est un événement.

$\nabla$ **Espace probabilisable :** Le couple $(\Omega\ ;\ \mathcal{A})$ formé par l'univers $\Omega$ et la tribu $\mathcal{A}$.

□ **Probabilité :** Application $\mathbb{P}$ définie sur l'ensemble $\mathcal{A}$ des événements d'un univers $\Omega$ (en fait, une tribu) à valeurs dans $[0;1]$ vérifiant :
- $\mathbb{P}(\Omega) = 1$
- $\forall A \in \mathcal{A} \quad \mathbb{P}(\overline{A}) = 1 - \mathbb{P}(A)$
- Pour toute suite $(A_n)_{n\in\mathbb{N}}$ d'événements pris dans $\mathcal{A}$ *deux à deux incompatibles* on a :

$$\mathbb{P}\left[\bigcup_{k\in\mathbb{N}} A_k\right] = \sum_{k\in\mathbb{N}} \mathbb{P}[A_k] = \lim_{n\to+\infty} \sum_{k=0}^{n} \mathbb{P}[A_k]$$

$\nabla$ **Espace probabilisé :** Le triplet $(\Omega\ ;\ \mathcal{A}\ ;\ \mathbb{P})$ où $\mathbb{P}$ est une probabilité sur l'espace probabilisable $(\Omega\ ;\ \mathcal{A})$

□ **Système Complet d'Evénements (cas général) :** Toute suite d'événements $(A_n)_{n\in\mathbb{N}}$ d'un même univers $\Omega$ vérifiant :
1. $\bigcup_{n\in\mathbb{N}} A_n = \Omega$
2. $\forall i \neq j \quad A_i \cap A_j = \emptyset$

*Il s'agit du vocabulaire employé en probabilités pour désigner une partition de l'ensemble* $\Omega$.
*Remarque :* On peut remplacer "$n \in \mathbb{N}$" par $n \in I$ (ou $i \in I$ en exploitant un indice $i$) avec $I$ en bijection avec un sous-ensemble de $\mathbb{N}$.

## 1.3 Vocabulaire associé au calcul de probabilités

Dans cette section, on se place dans un espace probabilisé $(\Omega\ ;\ \mathcal{A}\ ;\ \mathbb{P})$ et on notera $A$ un événement de cet espace de probabilité non nulle.

□ **Probabilités élémentaires :** Valeur de probabilité d'un événement élémentaire.

□ **Evénement négligeable :** Evénement de probabilité nulle. *Attention !* Ce n'est pas nécessairement $\emptyset$.

□ **Evénement presque sûr :** Evénement de probabilité 1 (un). *Attention !* Ce n'est pas nécessairement $\Omega$.

□ **Equiprobabilité (cas $\Omega$ fini) :** Lorsque $\Omega$ est fini de cardinal $n \in \mathbb{N}^*$, parle d'équiprobabilités lorsque $\mathbb{P}$ renvoie les mêmes valeurs pour chaque événement élémentaire :

$$\forall (\omega_1\ ;\ \omega_2) \in \Omega^2 \quad \mathbb{P}[\{\omega_1\}] = \mathbb{P}[\{\omega_2\}]$$

□ **Probabilité Conditionnelle (sachant $A$) :** Il s'agit de la probabilité notée $\mathbb{P}_A$ vérifiant :

$$\forall E \in \mathcal{A} \quad \mathbb{P}_A(E) = \frac{\mathbb{P}(A \cap E)}{\mathbb{P}(A)} \qquad (\mathbb{P}(A) \neq 0)$$

□ **Evénements indépendants :** On dit que $B$ est indépendant de $A$ lorsque $\mathbb{P}_A(B) = \mathbb{P}(B)$.
De façon équivalente, $A$ et $B$ sont indépendants lorsqu'ils satisfont : $\mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}(B)$. Ceci permet d'inclure les cas où $A$ ou $B$ pourraient être vides.

□ **Indépendance mutuelle d'événements :** Pour $(E_n)_{n\in I}$ avec $I \subset \mathbb{N}$, famille d'événements de $(\Omega\ ;\ \mathcal{A}\ ;\ \mathbb{P})$, on parlera d'indépendance mutuelle lorsque toute sous-famille finie $(E_{i_1}\ ;\ E_{i_2}\ ;\ \ldots\ ;\ E_{i_n})$ vérifie :

$$\mathbb{P}\left(E_{i_1} \cap E_{i_2} \cap \ldots \cap E_{i_n}\right) = \prod_{k=1}^{n} \mathbb{P}[E_{i_k}]$$

## Vocabulaire (et notations) des variables aléatoires

Dans cette section, on se place dans un espace probabilisé $(\Omega\ ;\ \mathcal{A}\ ;\ \mathbb{P})$.

☐ **Variable Aléatoire Réelle (VAR) :** Application $X : \Omega \longrightarrow \mathbb{R}$ vérifiant, pour tout intervalle $I$ de $\mathbb{R}$ : $X^{-1}(I) \in \mathcal{A}$.

☐ **VAR finie :** On dira que $X$ est finie lorsque $X(\Omega)$ est de cardinal fini.

☐ **VAR discrète :** On dira que $X$ est discrète lorsque $X(\Omega)$ lorsque $\inf\{|x_1 - x_2|\ ;\ (x_1; x_2) \in X(\Omega)^2\} > 0$.
Le terme de *discret* provient de la branche mathématique dite de le *topologie* et ne sera donc pas développée ici.

☐ Notations : Les événements associés aux manipulations de VAR disposent de leurs notations propres :
- Pour $I$, intervalle de $\mathbb{R}$, on préfère noter $[X \in I]$ l'événement $X^{-1}(I) = \{\omega \in \Omega\ ;\ X(\omega) \in I\}$
- On peut écrire $[X = a]$ au lieu de $[X \in \{a\}]$
- On peut écrire $[X \leq a]$ au lieu de $[X \in ]-\infty\ ;\ a]]$
- On peut écrire $[X < a]$ au lieu de $[X \in ]-\infty\ ;\ a[]$
- On peut écrire $[X \geq a]$ au lieu de $[X \in [a\ ;\ +\infty[]$
- On peut exploiter les notations $[a < X < b]$, $[a \leq X \leq b]$, $[a < X \leq b]$ ou encore $[a \leq X < b]$ de façon analogue.

☐ **Loi de VAR :** La loi d'une VAR notée $X$ est la donnée des valeurs $(\mathbb{P}[X \in I])_{I:\text{intervalle}}$.
*NB :* En pratique, on cherche d'autres moyens équivalents et plus efficaces de décrire une loi de probabilité.

☐ **Fonction de Répartition d'une VAR :** Fonction $F_X$ associée à la Variable Aléatoire $X$ définie par :

$$\forall t \in \mathbb{R} \quad F_X(t) = \mathbb{P}[X \leq t]$$

### 1.4 Indicateurs des VAR discrètes :

Dans cette section, on se place dans un espace probabilisé $(\Omega\ ;\ \mathcal{A}\ ;\ \mathbb{P})$ et on considère que $X$ est une VAR discrète définie sur cet espace, c'est-à-dire que $X(\Omega)$ est un ensemble $D \subset \mathbb{R}$ discret.
En pratique, un ensemble $D$ dénombrable suffit pour mettre en application ce que suit.

☐ **Espérance :** Valeur, sous couvert d'existence définie par la formule : $\mathbb{E}[X] = \sum\limits_{x \in X(\Omega)} x\mathbb{P}[X = x]$
*Remarque :* Existence assurée si $X$ est (presque-sûrement) finie.

☐ **Variance :** Valeur, sous couvert d'existence définie par la formule : $\mathbb{V}[X] = \mathbb{E}\left[(X - \mathbb{E}[X])^2\right]$
*Remarque :* Existence assurée si $X$ est (presque-sûrement) finie.

☐ **Ecart-type :** Sous couvert d'existence de $\mathbb{V}[X]$, valeur $\sigma(X) = \sqrt{\mathbb{V}[X]}$.

☐ **Moment (d'ordre $r$) :** Avec $r \in \mathbb{N}^*$, valeur, sous couvert d'existence définie par la formule : $m_r[X] = \mathbb{E}[X^r]$
*Remarque :* Existence assurée si $X$ est (presque-sûrement) finie.

☐ **VAR centrée :** On dit de $X$ qu'elle est centrée lorsque $\mathbb{E}[X] = 0$.

☐ **VAR réduite :** On dit de $X$ qu'elle est réduite lorsque $\mathbb{V}[X] = 1$.

☐ **VAR centrée-réduite :** On dit de $X$ qu'elle est centrée lorsque $\mathbb{E}[X] = 0$ et que $\mathbb{V}[X] = 1$.

☐ **VAR centrée-réduite associée à $X$ :** On définit $X^*$ associée à $X$ où $\sigma(X) \neq 0$ comme : $X^* = \dfrac{1}{\sigma_X} \cdot (X - \mathbb{E}[X])$

### 1.5 Rappels des lois discrètes usuelles

Le choix de présenter ou non les démonstrations relevant de l'étude de ces lois est laissée à libre appréciation. Le calcul des indicateurs peut constituer des exercices types.

## Loi Certaine

**Utilisation :** Intérêt théorique, rejet d'hypothèse d'aléatoire. Issue déterministe

**Notation :** On peut écrire $\mathbb{1}$ pour désigner une variable aléatoire certaine égale à 1.
On pourra donc écrire $a\mathbb{1}$ pour désigner une variable aléatoire certaine égale à $a \in \mathbb{R}$.

**Ensemble de valeurs** On a $\mathbb{1}(\Omega) = \{1\}$ et donc $a\mathbb{1}(\Omega) = \{a\}$

**Loi de probabilité** : $a$ étant un réel fixé, si $X$ suit la même loi certaine que $a\mathbb{1}$ alors :

| $x$ | $a$ | $\neq a$ |
|---|---|---|
| $\mathbb{P}[X = x]$ | 1 | 0 |

**Indicateurs :** Le couple espérance-variance est donné par :

- $\mathbb{E}[\mathbb{1}] = 1$ et (par linéarité) $\mathbb{E}[a\mathbb{1}] = a$
- $\mathbb{V}[\mathbb{1}] = 0$ et $\mathbb{V}[a\mathbb{1}] = 0$

## Loi Uniforme (discrète)

**Utilisation :** Situation d'équiprobabilité générant des nombres entiers de $a$ à $b$.

**Notation :** On écrit $\mathcal{U}[\![a\ ;\ b]\!]$ avec $a \leq b$ entiers.

**Ensemble de valeurs** Si $X$ suit une loi $\mathcal{U}[\![a\ ;\ b]\!]$ alors $X(\Omega) = [\![a\ ;\ b\ ]\!]$ de cardinal $N = b - a + 1$.

**Loi de probabilité** : Les paramètres $a$ et $b$ étant des entiers naturels avec $a < b$ on a :

$$\forall k \in [\![a\ ;\ b]\!] \quad \mathbb{P}[X = k] = \frac{1}{b - a + 1}$$

; sous forme de tableau :

| $x$ | $a$ | $a+1$ | $\dots$ | $b$ |
|---|---|---|---|---|
| $\mathbb{P}[X = x]$ | $\frac{1}{b-a+1}$ | $\frac{1}{b-a+1}$ | $\dots$ | $\frac{1}{b-a+1}$ |

*Remarque :* Si $a = b$, la loi est certaine.
**Indicateurs :** Le couple espérance-variance de $X \hookrightarrow \mathcal{U}[\![a\ ;\ b]\!]$ est donné par :

- $\mathbb{E}[X] = \dfrac{b + a}{2}$
- $\mathbb{V}[X] = \dfrac{N^2 - 1}{12}$

où $N = b - a + 1$.

**Attendu :** Savoir retrouver une loi uniforme $\mathcal{U}[\![a\ ;\ b]\!]$ avec $a < b$ entiers, à partir de la connaissance particulière de $\mathcal{U}[\![1\ ;\ n]\!]$, au moyen de la transformation :

$$X = U - 1 + a$$

avec $U \hookrightarrow \mathcal{U}[\![1\ ;\ n]\!]$ en posant $n = b - a + 1$. Les valeurs $a$ et 1 sont assimilables aux variables certaines $a\mathbb{1}$ et $a\mathbb{1}$ respectivement.

## Loi de Bernoulli

**Utilisation :** Etude d'une expérience générique avec discrimination type *succès / échec*.
On encode en binaire la réalisation du succès :

| événement | Succès | Echec |
|---|---|---|
| VAR | 1 | 0 |

**Notation :** On écrit $\mathcal{B}(p)$ avec $p$ probabilité du succès.

**Ensemble de valeurs** Si $X$ suit une loi $\mathcal{B}(p)$ alors $X(\Omega) = \{0; 1\}$.

**Loi de probabilité** : Le paramètre $p$ étant dans $[0; 1]$ on a :

| $x$ | $0$ | $\neq 1$ |
|---|---|---|
| $\mathbb{P}[X = x]$ | $1-p$ | $p$ |

**Indicateurs :** Le couple espérance-variance de $X \hookrightarrow \mathcal{B}(p)$ est donné par :
- $\mathbb{E}[X] = p$
- $\mathbb{V}[X] = p(1-p)$

## Loi Binomiale

**Utilisation :** Nombre de succès d'un schéma de Bernoulli à $n$ répétitions.
On peut aussi la voir comme somme de variables de Bernoulli indépendantes et de même paramètre $p$.

**Notation :** On écrit $\mathcal{B}(n;p)$ avec $p$ probabilité du succès et $n$ le nombre d'épreuves réaliées.

**Ensemble de valeurs** Si $X$ suit une loi $\mathcal{B}(n;p)$ alors $X(\Omega) = [\![0\ ;\ n\,]\!]$.

**Loi de probabilité** : Le paramètre $p$ étant dans $[0; 1]$ et $n$ étant un entier naturel on a :

$$\forall k \leq n \quad \mathbb{P}[X = k] = \binom{n}{k} p^k (1-p)^{n-k}$$

**Indicateurs :** Le couple espérance-variance de $X \hookrightarrow \mathcal{B}(n;p)$ est donné par :
- $\mathbb{E}[X] = np$
- $\mathbb{V}[X] = np(1-p)$

**Théorème de convergence Binomiale** :
Soit $X_n$ une VAR de loi $\mathcal{B}(n;p_n)$ pour $n \geq 1$. Si l'on a $\lim\limits_{n \to +\infty} np_n = \lambda > 0$ alors $(X_n)$ converge (en loi) vers une VAR notée $X$ et dont la loi peut être caractérisée par :

$$\forall n \in \mathbb{N} \quad \mathbb{P}[X = n] = e^{-\lambda} \cdot \frac{\lambda^n}{n!}$$

## Loi de Poisson

**Utilisation :** On nomme aussi *Loi des événements rares* la loi de Poisson. Elle permet de modéliser la survenue du nombre d'événements "rares" d'une série d'essais "en grands nombres".
Ceci s'explique par le théorème de convergence binomiale : toute situation relevant d'un schéma de Bernoulli avec $n$ arbitrairement grand et où la probabilité d'un *succès* générique est arbitrairement petite peut relever d'une loi de Poisson. On garde alors à l'esprit qu'il s'agit d'un modèle d'approximation plus commode à exploiter que le modèle binomial.

**Notation :** On écrit $\mathcal{P}(\lambda)$ avec $\lambda > 0$ paramètre

**Ensemble de valeurs** Si $X$ suit une loi $\mathcal{P}(\lambda)$ alors $X(\Omega) = \mathbb{N}$.

**Loi de probabilité** : Le paramètre $\lambda$ étant dans $\mathbb{R}_+^*$ on a :

$$\forall n \in \mathbb{N} \quad \mathbb{P}[X = n] = e^{-\lambda} \cdot \frac{\lambda^n}{n!}$$

<u>**Indicateurs :**</u> Le couple espérance-variance de $X \hookrightarrow \mathcal{P}(\lambda)$ est donné par :

- $\mathbb{E}[X] = \lambda$
- $\mathbb{V}[X] = \lambda$

# 2 L'analyse fonctionnelle pour les probabilités

## 2.1 Fonctions usuelles complémentaires

**Définition (fonctions indicatrices)** :
Soit $A \subset \mathbb{R}$ une partie donnée. On nomme *fonction indicatrice de* $A$ notée $\mathbb{1}_A$ ou encore $_A$ la fonction à valeurs booléennes définie comme :

$$\begin{array}{rccl} \mathbb{1}_A : \mathbb{R} & \longrightarrow & \mathbb{B} \\ x & \mapsto & \begin{cases} 1 & \text{si } x \in A \\ 0 & \text{sinon} \end{cases} \end{array}$$

On utilisera fréquement les fonctions indicatrices dans les cas où $A$ est un intervalle de $\mathbb{R}$ ou encore un ensemble de référence tel $\mathbb{N}$ ou $\mathbb{Z}$. On assimilera $\mathbb{B}$ à $\{0; 1\} \subset \mathbb{R}$.

<u>Convention :</u> On désignera par $f \cdot \mathbb{1}_A$ la fonction dont l'expression $f(x) \cdot \mathbb{1}_A(x)$ canonique associée désigne l'écriture :

$$\begin{array}{rccl} f \cdot \mathbb{1}_A : \mathbb{R} & \longrightarrow & \mathbb{R} \\ x & \mapsto & \begin{cases} f(x) & \text{si } x \in A \\ 0 & \text{sinon} \end{cases} \end{array}$$

Ainsi, si $f$ est une fonction de $\mathbb{R}$ dans $\mathbb{R}$ avec $\mathcal{D}_f$ son ensemble de définition, on considérera prioritaire l'affection à 0 de la valeur $f(x)\mathbb{1}_A(x)$ si $x \notin A$ et $x \notin \mathcal{D}_f$.

**Exemple :** la fonction $h$ définie comme $h(x) = \dfrac{1}{x} \cdot \mathbb{1}_{\mathbb{R}^*}(x)$ est définie sur tout $\mathbb{R}$ et $h(0) = 0$. Si $x \neq 0$ alors $h(x) = x^{-1}$.

**Propriété.** *Si $f$ est une fonction définie sur $D \subset \mathbb{R}$ alors la fonction $f\mathbb{1}_D$ est une application de $\mathbb{R}$ dans $\mathbb{R}$ qui coïncide avec $f$ sur $D$.*

**Point Code** Ces conventions et résultats permettent l'étude ou l'écriture de programmes de ce type :

```
def y = appli_h(x)
        if [x in D] :           # écriture d'une condition équivalente
                y = f(x)                # fonction f supposée déjà existante
        else:
                y=0
        return(y)
```

On voit que le test de la condition $x \in D$ est priopritaire devant l'évaluation (éventuellement interdite)

**Exemple :** la fonction $h$ définie comme $h(x) = \dfrac{1}{x}\mathbb{1}_{\mathbb{R}^*}(x)$ représente le code :

```
def y = h(x)
        if x<>0 :
                y = 1/x
        else :
```

```
            y=0
        return(y)
```

**Exercice-Exemple.** *On considère la fonction $h_a(x) = e^{-ax}\mathbb{1}_{[0;a]}(x)$ pour $a > 0$ donné et $x \in \mathbb{R}$.*
*Procéder à l'étude et la représentation graphique de $h_a$. Calculer $\displaystyle\int_{-\infty}^{+\infty} h_a(x)\,dx$.*

**Définition (partie entière)** :
La fonction partie entière est la fonction qui à $x \in \mathbb{R}$ associe l'unique entier relatif $\lfloor x \rfloor$ vérifiant :

$$\lfloor x \rfloor \leq x < \lfloor x \rfloor + 1$$

*Remarque :* on peut aussi désigner cette fonction par l'appelation `floor` utilisée comme commande dans de nombreux langages de programmation. Cette fonction est commode pour tester le prédicat *is integer* :

**Point Code** On va créer un programme qui calcule $u_n = (-1.02)^n$ si $n \in \mathbb{N}$ mais qui renvoie un code d'erreur si l'entrée $n$ n'est pas un entier :

```
n= float(input("saisir valeur réelle"))
if (n==floor(n)) and (n>=0) :
        u=(-1.02**n)
else :
        print("error code FFFF")
```

**Propriété.** *La fonction `floor` est continue et constante sur chaque intervalle $I_k = [k; k+1[$ où elle vaut $k \in \mathbb{Z}$. En chaque valeur $a$ de $\mathbb{Z}$, elle admet une discontinuité par la gauche.*
*Etant en escalier, elle reste intégrable sur tout segment $I \subset \mathbb{R}$.*

**Exercice-Exemple.** *Calculer $\displaystyle\int_0^x \lfloor t \rfloor\,\mathtt{d}t$ en fonction de $x \in \mathbb{R}_+$ puis $\displaystyle\int_0^x \lfloor 2t-5 \rfloor\,\mathtt{d}t$.*

**Propriété. décomposition de floor en indicatrices**
*On a, pour tout $x \geq 0$ réel :*

$$\lfloor x \rfloor = \sum_{k=0}^{+\infty} \mathbb{1}_{[k;+\infty[}(x) = \sum_{k=0}^{+\infty} k \cdot \mathbb{1}_{[k;k+1[}(x)$$

L'écriture d'une formule analogue pour $x < 0$ peut être faite en guise d'exercice.

**Exercice-Exemple.** *Calculer $\displaystyle\int_0^x \lfloor e^t \rfloor\,\mathtt{d}t$ et en déduire $\displaystyle\int_0^{+\infty} \lfloor e^{-t} \rfloor\,\mathtt{d}t$*

## 2.2 Fonctions de densité

**Définition (densité de probabilités)** : On appelle *densité de probabilité* toute application $f$ de $\mathbb{R}$ dans $\mathbb{R}_+$, intégrable sur $\mathbb{R}$ vérifiant :

$$\int_{-\infty}^{+\infty} f(t)\,\mathtt{d}t = 1$$

Remarques :
- Dire que $f$ est une application définie sur $\mathbb{R}$ impose $\mathcal{D}_f = \mathbb{R}$. Ainsi, $f(x)$ admet une valeur en chaque $x \in \mathbb{R}$
- Ecrire que $f : \mathbb{R} \longrightarrow \mathbb{R}_+$ revient à demander $f \geq 0$ (positivité)
- On ne développera pas la locution $f$ *intégrable* mais on gardera à l'esprit que l'on peut intégrer $f$ et que tout intégrale obtenue converge bien.

**Exercice-Exemple.** *Les fonctions* $\mathbb{1}_{[0;1]}$ *et* $x \mapsto \dfrac{\frac{1}{\pi}}{1+x^2}$ *sont des densités de probabilités.*

**Propriété.** *Si $f$ est une densité de probabilité, alors il existe une variable aléatoire réelle $X$ dont la fonction de répartition vérifie :*

$$\forall x \in \mathbb{R} \quad F_X(x) = \mathbb{P}[X \leq x] = \int_{-\infty}^{x} f(t)\ \mathtt{d}t$$

*et, en particulier, on aura* $\mathbb{P}[a \leq X \leq b] = \displaystyle\int_a^b f(t)\ \mathtt{d}t$

**Vocabulaire :** On dit dans un tel cas que $X$ est une variable aléatoire à densité.

**Propriété.** *Soit $X$ une variable aléatoire réelle.*
*On a l'équivalence entre* $X$ est à densité *et* la fonction de répartition $F_X$ de $X$ est continue sur tout $\mathbb{R}$ et de classe $\mathcal{C}^1$ sur $\mathbb{R}$ éventuellement privé d'un ensemble *négligeable.*

Sans rentrer dans la description formelle et exhaustive de la théorie de la mesure permettant d'introduire avec rigueur la notion d'ensemble négligeable, on pourra retenir que tout ensemble fini ou dénombrable se trouve être dans ce cas pour les intégrales usuelles.

Cette propriété peut aussi suggérer l'emploi de la dénomination *VAR continues* au lieu de *à densité*.

**Méthode :** Pour établir qu'une VAR est à densité à partir de sa fonction de répartition $F_X$, on peut exprimer sa dérivée $F'_X(x)$, valable sur $\mathbb{R}$ sauf éventuellement sur $D$, négligeable, puis établir que la fonction obtenue est une densité de probabilités.

Alternativement, on peut utiliser la caractérisation des fonctions de répartitions, puis la propriété qui précède.

**Propriété. (caractérisation des fonctions de répartition)**
*$F$ est une fonction de répartition si, et seulement si, $F$ vérifie simultanément :*
*a)* *$F$ est croissante sur $\mathbb{R}$ (au sans large)*
*b)* $\lim\limits_{x\to-\infty} F(x) = 0$ *et* $\lim\limits_{x\to+\infty} F(x) = 1$
*c)* *$F$ est continue à droite en chaque $a \in \mathbb{R}$.*

**Exercice-Exemple.** *1. On définit la famille de fonctions $(F_n)$ par :*

$$F_n(x) = \frac{n + \lfloor x \rfloor}{2n} \mathbb{1}_{[-n;n]}(x) + \mathbb{1}_{]n;+\infty[}(x) \quad (n \in \mathbb{N}^*)$$

*Etablir que chaque $F_n$ est la fonction de répartition d'une variable aléatoire $X_n$ mais que cette dernière n'est pas à densité.*

2. *Etablir que la fonction* $H(x) = \dfrac{\arctan(x)}{\pi} + \dfrac{1}{2}$ *est la fonction de répartition d'une variable aléatoire $Y$ à densité.*
3. *Pour $n \in \mathbb{N}^*$, la fonction $G_n(x) = x^n \mathbb{1}_{[0;1]}(x) + \mathbb{1}_{]1;+\infty]}(x)$ définit une fonction de répartition de VAR à densité pour $n \in \mathbb{N}^*$.*

**Propriété. (densité induite)** *Toute fonction $f$ positive définie, intégrable sur un intervalle $I$ de $\mathbb{R}$ de longueur $l = \lambda(I) > 0$ induit une densité $h$ en posant :*

$$h(x) = \frac{1}{c} f(x) \mathbb{1}_I(x)$$

*où la valeur $c$ désigne* $\int_I f(x)\ \mathtt{d}x$

**Exercice-Exemple.** *Déterminer la densité induite par la fonction $f$ définie sur $]0;1]$ par* $f(x) = \dfrac{3}{2\sqrt{x}}$

**Propriété.** *Soir $X$ une variable aléatoire à densité $f$. Alors $\mathbb{P}[X = a] = 0$ pour tout $a \in \mathbb{R}$.*
*Ainsi :*

$$\forall(a;b) \in \mathbb{R}^2 \;\; a \leq b \implies \mathbb{P}[a < X < b] = \mathbb{P}[a < X \leq b] = \mathbb{P}[a \leq X < b] = \mathbb{P}[a \leq X \leq b]$$

On observera que ceci permettra de calculer une probabilité de type $\mathbb{P}[a \leq X \leq b]$ comme différence $F_X(b) - F_X(a)$ d'images de la fonction de répartition, d'où l'utilité de sa connaissance.

De plus, ceci rend la valeur ponctuelle $f(x_0)$ concrètement inutile pour $x_0$ fixé en quantité dénombrable. Autrement dit, si $f_1$ et $f_2$ sont deux densités égales partout sauf sur un ensemble négligeable (dénombrable suffira), alors elles définissent une même loi de probabilités et en conséquence, une variable aléatoire $X$ admet plusieurs densités.

Retenir qu'en revanche, la fonction de répartition $F_X$ caractérise quant à elle la loi de $X$ : chaque VAR admet une unique fonction de répartition.

> **Définition (support)** : Soit $X$ une VAR à densité $f$. On appelle support de $X$ le plus petit ensemble fermé contenant l'ensemble :
> $$supp(f) = \sigma_f = \{x \in \mathbb{R} \mid f(x) > 0\}$$

**Exemple :** Les fonctions $\mathbb{1}_{[0;1]}$, $\mathbb{1}_{]0;1]}$, $\mathbb{1}_{[0;1[}$ et $\mathbb{1}_{]0;1[}$ sont toutes des densités associées à une même loi. Notons $X$ une VAR admettant ces densités. Le support de $X$ est $[0;1]$.

Dans la suite, on peut donc définir des VAR sur des intervalles notés $(a;b)$ pour éviter de préciser le statut des bornes (fermés ou ouverts) mais cette notation est ambigüe (ce n'est pas le couple $(a;b)$). On préférera écrire les bornes explicitement et garder à l'esprit qu'elles peuvent être fermées ou ouvertes indifféremment sans changer l'étude.

# 3 Lois et générations de VAR à densité

## 3.1 Indicateurs usuels

Dans cette section, on se place dans un espace probabilisé $(\Omega \; ; \; \mathcal{A} \; ; \; \mathbb{P})$ et on considère que $X$ est une VAR <u>à densité</u> définie sur cet espace.
On remarquera que, en dehors de l'espérance (encore que . . . ), les définitions sont formellement identiques à celles des variables aléatoires discrètes. Il s'agit de garder à l'esprit que, la formule de l'espérance $\mathbb{E}[X]$ étant distincte, les concepts qui en découlent se fondent sur cette dernière.
On pourra désigner par $f_X$ une densité de $X$ et par $F_X$ la fonction de répartition de $X$.

- ☐ **<u>Espérance :</u>** Valeur, sous réserve de *convergence absolue*, définie par la formule : $\mathbb{E}[X] = \displaystyle\int_{-\infty}^{+\infty} x f_X(x) \, \mathrm{d}x$
  *Remarque :* Existence assurée si $X$ est de support un segment $I$.

- ☐ **<u>Variance :</u>** Valeur définie, sous couvert d'existence, comme : $\mathbb{V}[X] = \mathbb{E}\left[(X - \mathbb{E}[X])^2\right]$
  *Remarque :* Définition invariante selon les cas discrets et continus.

- ☐ **<u>Ecart-type :</u>** Sous-couvert d'existence de $\mathbb{V}[X]$, valeur $\sigma(X) = \sqrt{\mathbb{V}[X]}$.

- ☐ **<u>Moment (d'ordre $r$) :</u>** Avec $r \in \mathbb{N}^*$, valeur, sous couvert d'existence définie par la formule : $m_r[X] = \mathbb{E}[X^r]$
  *Remarque :* Existence assurée si $X$ est (presque-sûrement) finie.

- ☐ **<u>VAR centrée :</u>** On dit de $X$ qu'elle est centrée lorsque $\mathbb{E}[X] = 0$.

- ☐ **<u>VAR centrée-réduite :</u>** On dit de $X$ qu'elle est centrée lorsque $\mathbb{E}[X] = 0$ et que $\mathbb{V}[X] = 1$.

  *Remarque :* L'emploi de *VAR réduite* pour le seul cas $\mathbb{V}[X] = 1$ est un abus, non reconnu universellement.

☐ **VAR centrée-réduite associée :** On définit $X^*$ associée à $X$ comme : $X^* = \frac{1}{\sigma_X} \cdot (X - \mathbb{E}[X])$ dès lors que $\sigma_X \neq 0$ ce qui revient à demander $X$ non presque-sûrement certaine

**Exercice-Exemple.** *On donne la fonction $f$ définie sur $[0;1]$ par $f(x) = x(x-1)$. Déterminer la densité induite $h$. Si $X$ est une VAR à densité $h$, quelles sont les valeurs d'espérance et de variance de $X$ associés ?*

## Fondements des générateurs pseudo-aléatoires

## 3.2 Loi uniforme (continue) sur $[0;1[$

On admet que la commande `RAND()` permet de générer une réalisation d'une VAR notée $U$ dont une densité est $f_U = \mathbb{1}_{[0;1[}$. En pratique, elle pourra être appelée via l'importation d'une bibliothèque.
La construction d'un tel générateur ne fait pas l'objet de ce cours.

**Définition** : On dit que $U$ suit la loi uniforme sur $[0;1[$ lorsqu'elle admet pour densité la fonction $\mathbb{1}_{[0;1[}$.

La fonction de répartition $F_U$ associée est alors définie par :

$$\forall x \in \mathbb{R} \quad F_U(x) = x \cdot \mathbb{1}_{[0;1[} + \mathbb{1}_{[1;+\infty[}(x)$$

**Point Code.** *On peut générer des réalisations de $U$ de la façon suivante, parmi d'autres :*

```
import numpy.random as rd
U=rd.random()
```

Nous considérerons que, pour des raisons de clarté et de simplicité, nous écrirons simplement :

```
U=rand()
```

**Méthode** On peut utiliser ce type de VAR pour simuler une condition (de sortie de boucle par exemple) aléatoire.

```
U=rand()
p=float(input("probabilité de réussite"))       #supposée connue
n=1
while U > p:
        n=n+1
        U=rand()
print(n)
```

Un tel code compte le nombre de tentatives effectuées au cours d'une simulation pour obtenir la première réalisation d'un certain événement $A$ dont la probabilité $p = \mathbb{P}[A]$ est connue.

## 3.3 Lois uniformes (continues) sur $[a;b]$

**Utilisation :** Sortie de l'usage d'une commande `RAND()`

**Notation :** On écrit $\mathcal{U}([a\ ;\ b])$ avec $a \leq b$ réels.

**Support :** Si $X$ suit une loi $\mathcal{U}([a\ ;\ b])$ alors son support est $[a;b]$.

**Densité** : La fonction $\frac{1}{b-a}\mathbb{1}_{[a;b]}$.
**Fonction de répartion** : La fonction définie par $F(x) = \frac{x-a}{b-a}\mathbb{1}_{[a;b]}(x) + \mathbb{1}_{]b;+\infty[}(x)$.
*Remarque :* Si $a = b$, la loi est certaine et n'est plus continue.
**Indicateurs :** Le couple espérance-variance de $\mathcal{U}([a\ ;\ b])$ est donné par :

- $\mathbb{E}[X] = \dfrac{a+b}{2}$
- $\mathbb{V}[X] = \dfrac{(b-a)^2}{12}$

**Point Code.** *On propose une instruction simple de génération d'une variable aléatoire $X$ suivant une loi $\mathcal{U}([a\ ;\ b])$ à partir de $U$*

```
a=float(input("borne inférieure"))      #supposée connue
b=float(input("borne inférieure"))      #supposée connue
U=rand()
X=(b-a) U +a
```

**Exercice-Exemple.** *Démontrer que, si $U$ suit la loi $\mathcal{U}([0\ ;\ 1])$ alors $X = (b-a)U + a$ suit la loi $\mathcal{U}([a\ ;\ b])$.*
*On utilisera pour cela les fonctions de répartition qui caractérisent les lois.*

**Méthode** Pour simuler des tirages discrets, on peut utiliser la fonction `floor`

**Point Code.** *Simulation de loi uniforme discrète :*

```
a=int(input("valeur inférieure"))
b=int(input("valeur supérieure"))
U=rand()
X=(b+1-a)U+a
D=floor(X)
```

**Exercice-Exemple.** *Démontrer que, si $U$ suit la loi $\mathcal{U}([0\ ;\ 1[)$ alors $D = \lfloor (b+1-a)U + a \rfloor$ suit la loi $\mathcal{U}([\![a\ ;\ b]\!])$.*
*Décrire la programmation d'une* fonction $\texttt{ALEA}(\texttt{a};\texttt{b})$ *qui renvoie un entier (pseudo)-aléatoire compris entre $a$ et $b$ lors de son appel.*

*Remarque :* La simulation d'un "jet de dé à $N$ faces" se fait donc en choisissant $a = 1$ et $b = N$.

**Point Code.** *Déterminer le rôle du code suivant (on pourra l'insérer dans un contexte* ludique*)*

```
a=1
b=6
U=rand()
X=(b+1-a)U+a
D=floor(X)
n=1
while D<6 :
        n=n+1
        U=rand()
        X=(b+1-a)U+a
        D=floor(X)
print("n=",n)
```

**Propriété.** *Soit $T$ le nombre de tentatives effectuées dans un schéma de Bernoulli jusqu'à obtention du premier succès.*
*Si $p \in ]0;1[$ est la probabilité d'un succès générique dans ce schéma, alors on a :*

$$\forall n \in \mathbb{N}^* \quad \mathbb{P}[T = n] = p(1-p)^{n-1}$$

*et $T(\Omega) = \mathbb{N}^*$, ce qui fait de $T$ une VAR discrète.*

On nomme *Loi Géométrique* de paramètre $p$ la loi ainsi caractérisée pour une telle variable aléatoire $T$.
Ces paramètres d'espérance et de variances sont alors : $\mathbb{E}[T] = \dfrac{1}{p}$ et $\mathbb{V}[T] = \dfrac{1-p}{p^2}$

**Exercice-Exemple.** *Déterminer le temps d'attente moyen de sortie de boucle dans le programme qui précède (en nombre de tours de boucle).*
*Comment quantifier l'écart (attendu) à la moyenne dans l'exécution (à venir) de ce programme ?*

**Point Code. simulation de la loi binomiale**
*Nous savons maintenant programmer une simulation d'une épreuve de Bernoulli. Nous pouvons donc générer une loi Binomiale et cette fois, nous adopterons le point de vue fonctionnel :*

```
def binomial(n,p):
        if 0<= p and p<=1 and n==floor(n) and n>0:
                k=0
                for t in range(n):
                        U=rand()
                        if U <= p:
                                k=k+1
                return(k)
        else :
                print("error input parameters")
                return(-1)                      #code d'erreur :-)
```

**Exercice-Exemple.** *Justifier que le programme suivant permet également la génération d'une VAR $X$ de loi $\mathcal{B}(n;p)$ :*

```
def binomial(n,p):
        if 0<= p and p<=1 and n==floor(n) and n>0:
                X=0
                for k in range(n):
                        U=rand()
                        X=X+floor(1+p-U)
                return(X)
        else :
                print("error input parameters")
                return(-1)
```

## 3.4 Lois à densité usuelles

On présente de façon synthétique chaque loi à densité et on considère exigibles les résultats énoncés.

## Loi Exponentielle

**Utilisation :** Durée de *survie* d'un élément radioactif, d'une machine, sous hypothèse d'absence de vieillissement.
**Notation :** On écrit $\mathcal{E}(\lambda)$ avec $\lambda > 0$ paramètre inverse de la durée de vie moyenne attendue.
**Support :** Si $X$ suit une loi $\mathcal{E}(\lambda)$ alors sont support est $\mathbb{R}_+$.

**Densité** : La fonction $f_\lambda : x \mapsto \lambda e^{-\lambda x} \mathbb{1}_{\mathbb{R}_+}(x)$.
**Fonction de répartion** : La fonction définie par $F(x) = \left(1 - e^{-\lambda x}\right) \mathbb{1}_{\mathbb{R}_+}(x)$.
**Indicateurs :** Le couple espérance-variance de $X$ suivant le loi $\mathcal{E}(\lambda)$ est donné par :
- $\mathbb{E}[X] = \lambda$

- $\mathbb{V}[X] = \lambda^2$

**Propriété.** *Les variables aléatoires $X$ à densité qui possèdent la propriété de* durée de vie sans vieillissement *(DVSV) s'énonçant :*

$$\forall h \geq 0 \forall x \geq 0 \quad \mathbb{P}[X \geq h] = \mathbb{P}_{[x \geq x]}[X \geq x+h]$$

*sont exactement les VAR qui suivent une loi $\mathcal{E}(\lambda)$ (avec $\lambda$ décrivant $\mathbb{R}_+^*$).*

**Point Code.** *On propose un programme de simulation de réalisation d'une VAR suivant une loi exponentielle de paramètre $\lambda = 1$ :*

```
U=rand()
X=-log(1-U)
```

**Exercice-Exemple.** *1. Démontrer que si $U$ suit une loi $\mathcal{U}([0;1[)$ alors $X = -\ln(1-U)$ suit une loi $\mathcal{E}(1)$.*

*2. On pose à présent $Y = \dfrac{1}{\lambda}X$ avec $\lambda > 0$. Etablir que :*

$$Y \hookrightarrow \mathcal{E}(\lambda) \iff X \hookrightarrow \mathcal{E}(1)$$

## Loi Gamma

**Utilisation :** Durée de *survie* d'un organisme vivant, contrat d'assurance, sévérité du risque. Cas de régression.
**Notation :** On écrit $\Gamma(a;p)$ avec $a > 0$ dit paramètre de forme et $p > 0$ dit paramètre d'échelle.
**Support :** Si $X$ suit une loi $\Gamma(a;p)$ alors son support est $\mathbb{R}_+$.

**Densité** : On rappelle que la fonction $\Gamma$ est définie comme $\Gamma(x) = \displaystyle\int_0^{+\infty} e^{-t}t^{x-1}\,\mathtt{d}t$.
On donne alors la densité $f_{a;p}$ de la loi Gamma :

$$f_{a;p}(x) = \frac{p^a x^{a-1} e^{-px}}{\Gamma(a)} \mathbb{1}_{\mathbb{R}_+^*}(x)$$

**Fonction de répartion** : La fonction définie par :

$$F_{a;p}(x) = \frac{p^a}{\Gamma(a)} \int_0^x e^{-pt}t^{x-1}\,\mathtt{d}t \cdot \mathbb{1}_{\mathbb{R}_+^*}(x)$$

.
**Indicateurs :** Le couple espérance-variance de $X$ suivant le loi $\mathcal{E}(\lambda)$ est donné par :

- $\mathbb{E}[X] = \dfrac{a}{p}$
- $\mathbb{V}[X] = \dfrac{a}{p^2}$

*Remarque :* La fonction $\Gamma$ est une fonction de référence bien connue et pour laquelle l'étude pourra être revue si besoin. En particulier $\Gamma(n+1) = n!$ pour les valeur de $n \in \mathbb{N}$.
Le lien avec la transformée de Laplace $\mathcal{L}(id^{\alpha-1})(1) = \Gamma(\alpha)$ n'est pas non plus à exclure pour relier les formules de calcul entre elles.

**Propriété.** *Soient $X_1 \ldots X_n$ des VAR mutuellement indépendantes qui suivent une même loi $\mathcal{E}(\lambda)$ avec $\lambda > 0$.*
*Alors $Z = X_1 + \cdots + X_n$ suit une loi $\Gamma(n;\lambda)$.*

Cette propriété sera établie à l'aide du produit de convolution dans le TDn°3.

**Point Code.** *On propose une simulation d'instances particulières de la loi Gamma $\Gamma(n;p)$ à partir de la construction de lois exponentielles. Voir TDn°3.*

```
n=int(input("paramètre entier Gamma"))
p=float(input("paramètre scaling Gamma"))
X=zeros(n)
for k in range(n):
        U=rand()
        X[k]=-log(1-U)/p
T=np.sum(X)
```

## Lois Normales

Beaucoup de comportements aléatoires sont assimilés à du *gaussien* ; peut-être même un peu trop justement.
Un des objectifs de ce cours vise tant à pouvoir assimiler un jeu de données aléatoires à une loi normale qu'à être en mesure de critiquer et remettre en question une telle hypothèse. On gardera à l'esprit que ce modèle est souvent un lissage prolongé sur $\mathbb{R}$ par commodité d'un véritable phénomène et n'en est donc pas la description exacte.
**Utilisation :** Dispersion de l'erreur des mesures physiques. Etude du comportement asymptotique d'une moyenne normalisée de tirages successifs. Répartition de critères biologiques quantifiés dans une population : taille, masse etc...
**Notation :** On écrit $\mathcal{N}(\mu; \sigma^2)$ avec $\mu \in \mathbb{R}$ la moyenne et $\sigma > 0$ l'écart-type.
**Support :** Si $X$ suit une loi $\mathcal{N}(\mu; \sigma^2)$ alors sont support est $\mathbb{R}$.

**Densité** : On rappelle que la fonction $\varphi$ dite de Gauss-Laplace est définie comme $\varphi(x) = \dfrac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$. Elle représente la densité de la loi $\mathcal{N}(0; 1)$.
On donne alors la densité $f_{\mu;\sigma}$ de la loi Normale plus généralement :

$$f_{\mu;\sigma}(x) = \frac{1}{\sigma}\varphi\left(\frac{x-\mu}{\sigma}\right) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left(-\frac{(x-\mu)^2}{2\sigma^2}\right)$$

**Fonction de répartion** : On note fréquemment $\Phi$ la fonction définie comme $\Phi(x) = \displaystyle\int_{-\infty}^{x} \varphi(t)\ \mathtt{d}t$, fonction de répartition de la loi $\mathcal{N}(0; 1)$
**Indicateurs :** Le couple espérance-variance de $X$ suivant le loi $\mathcal{N}(\mu; \sigma)$ est donné par :
- $\mathbb{E}[X] = \mu$
- $\mathbb{V}[X] = \sigma^2$

*RemarqueS importantes :*
- Les valeurs $\mu$ et $\sigma$ sont nommées *a posteriori* moyenne et écart-type mais il convient de démontrer que tel est le cas.
- On préfère alors définir $Z \hookrightarrow \mathcal{N}(0; 1)$ via sa densité $\varphi$ et établir que $\mathbb{E}[Z] = 0$ et $\mathbb{V}[Z] = 1$, donc que l'écart-type est bien 1.
- On définira enfin $X$ de loi $\mathcal{N}(\mu; \sigma^2)$ en passant par :

$$X \hookrightarrow \mathcal{N}(\mu; \sigma^2) \iff Z = \frac{X-\mu}{\sigma} \hookrightarrow \mathcal{N}(0; 1)$$

  L'étude de la loi en découle

**Propriété. intégrale de Gauss** *On a la convergence et la valeur alors induite de :*

$$\int_{-\infty}^{+\infty} e^{-x^2}\ \mathtt{d}x = \sqrt{\pi}$$

On pourra proposer une démonstration sous forme de problème en TD.
Il en découle que $\varphi$ est une densité, comme densité induite par $x \mapsto e^{-x^2/2}$, donc la convergence et la valeur d'intégrale sur $\mathbb{R}$ seront obtenues par changement de variables $u = \dfrac{x}{\sqrt{2}}$ dans l'intégrale de Gauss.

**Point Code.** *On propose un programme de simulation de réalisation d'une VAR suivant une loi normale centrée-réduite :*

```
#Algorithme de Box-Muller tronqué :
U=rand()
V=rand()
X= sqrt(-2*log(U)) * cos(2*pi* V)
```

**Propriété. Algorithme de Box-Muller**
*Soient $U$ et $V$ deux variables aléatoires indépendantes suivant toutes deux une loi uniforme sur $[0;1[$. Alors le couple $(X;Y)$ défini par :*

$$X = \sqrt{-2\ln U}\cos(2\pi V) \quad ; \quad Y = \sqrt{-2\ln U}\sin(2\pi V)$$

*est couple de VAR indépendantes qui suivent des lois normales centrées-réduites.*

Ce résultat se démontre alors d'intégrales doubles sur lesquelles on opère un changement de couple de variables : ceci fait appel au Jacobien.

# Lois induites

On peut construire de nouvelles lois en partant de lois déjà connues et en appliquant sur les VAR des transformations, comme déjà fait jusqu'à présent. Certaines lois sont alors définies comme telles :

## Lois Log-Normales

**Utilisation :** Linguistique : approximativement, la longueur d'une phrase. On trouve aussi comme champ d'application la masse d'un humain adulte, la distribution des richesses.
**Notation :** On écrit $\Lambda(\mu;\sigma^2)$ avec $\mu \in \mathbb{R}$ la moyenne et $\sigma > 0$ l'écart-type.
**Support :** Si $X$ suit une loi $\Lambda(\mu;\sigma^2)$ alors sont support est $\mathbb{R}_+^*$.

**transformation caractéristique** : On dit que $X$ suit une loi LogNormale $\Lambda(\mu;\sigma^2)$ lorsque :

$$Y = \ln X \quad \hookrightarrow \quad \mathcal{N}(\mu;\sigma^2)$$

**Indicateurs :** Le couple espérance-variance de $X$ suivant le loi $\Lambda(\mu;\sigma^2)$ est donné par :
- $\mathbb{E}[X] = \mu$
- $\mathbb{V}[X] = \sigma^2$

**Propriété. Densités et fonction de répartition LogNormales :** *Les trois assertions suivantes sont équivalentes :*
- *$X$ est une VAR suivant la loi LogNormale $\Lambda(\mu;\sigma^2)$*
- *$X$ est une VAR à densité dont une des densités est :*

$$f(x) = \frac{1}{x\sigma\sqrt{2\pi}}\exp -\frac{(\ln x\ -\mu)^2}{2\sigma^2}\Big) \cdot \mathbb{1}_{\mathbb{R}_+^*}(x)$$

- *$X$ admet pour fonction de répartition :*

$$F(x) = \int_{-\infty}^{x} \frac{1}{t\sigma\sqrt{2\pi}}\exp -\frac{(\ln t\ -\mu)^2}{2\sigma^2}\Big) \cdot \mathbb{1}_{\mathbb{R}_+^*}(t)\ \mathtt{d}t \quad = \int_{0}^{\ln x} \frac{1}{\sigma\sqrt{2\pi}}\exp -\frac{(t\ -\mu)^2}{2\sigma^2}\Big)\ \mathtt{d}t \cdot \mathbb{1}_{\mathbb{R}_+^*}(x)$$

On pourra proposer une démonstration sous forme d'exercice en TD.

**Point Code.** *On propose un programme de simulation de réalisation d'une VAR suivant une loi LogNormale centrée-réduite :*

```
#Algorithme de Box-Muller tronqué :
U=rand()
V=rand()
Y= sqrt(-2*log(U)) * cos(2*pi* V)
X = exp(Y)
```

**Exercice :** Déterminer, dans le programme qui précède, la variable renvoyant une réalisation de la loi LogNormale, en justifiant.

## La loi Bêta

L'étude théorique de cette loi pourra être détaillée en TD ou admise.
On définit pour tout couple $(r, s)$ de réels strictement positifs le nombre $B(r; s)$ par :

$$B(r; s) = \int_0^1 x^{r-1}(1-x)^{s-1}dx$$

On peut vérifier que, pour tout $(r; s) \in (\mathbb{R}_+^*)^2$ l'application $x \mapsto x^{r-1}(1-x)^{s-1}$ est bien de classe $\mathcal{C}^\infty$ sur $]0; 1[$.
On peut le vérifier en traitant plusieurs cas, selon que $r$ ou $s$ sont inférieurs ou supérieurs à 1.

**Propriété.** *L'intégrale qui définit $B(r; s)$ est convergente pour tout $(r; s) \in (\mathbb{R}_+^*)^2$ à valeur strictement positive.*

En conséquence, la fonction $f$ définie sur $\mathbb{R}$ par :

$$f(x) = \frac{1}{B(r; s)} x^{r-1}(1-x)^{s-1} \mathbb{1}_{]0;1[}(x)$$

est une densité de probabilités.

On peut relier les fonctions Bêta et Gamma à l'aide d'une formule de calcul utile (la démonstration est complètement facultative) :

**Propriété.** *Pour $p > 0$ et $q > 0$ on a :*

$$B(p; q)\Gamma(p+q) = \Gamma(p)\Gamma(q)$$

L'analogie de cette formule est à faire avec $\binom{p+q}{p}(p+q)! = p!q!$, plus aisé à voir avec la notation multinomiale : $\binom{p+q}{p} = \binom{p+q}{q} = \binom{p+q}{p;q}$.

*Cas Particulier :* On peut reconnaître une loi déjà étudiée dans le cas où $r = s = 1$.
**Définition :** Une variable aléatoire $X$ suit une loi Bêta de paramètres $r > 0$ et $s > 0$ lorsqu'elle admet la fonction $f$ décrite plus haut comme densité.

On peut décrire la fonction de répartition $F$ associée à la densité $f$ (pour $(r; s) \in (\mathbb{R}_+^*)^2$ fixé) à l'aide d'une intégrale.
**Notation :** On écrit $\mathtt{B}(r; s)$ avec $r > 0$ et $s > 0$.
**Support :** Si $X$ suit une loi $\mathtt{B}(r; s)$ alors son support et $[0; 1]$.

**Indicateurs :** Le couple espérance-variance de $X$ suivant le loi $\mathtt{B}(r; s)$ est donné par :
- $\mathbb{E}[X] = \dfrac{r}{r+s}$
- $\mathbb{V}[X] = \dfrac{rs}{(r+s)^2(r+s+1)}$

## Une étude particulière

On présente des résultats utiles dans le cas où $r$ et $s$ sont entiers.

**Propriété.**

$$\forall (p; q) \in (\mathbb{N}^*)^2 \qquad \frac{1}{B(p; q)} = p\binom{p+q-1}{p} = q\binom{p+q-1}{q}$$

Ceci peut se démontrer en établissant que $\forall (p; q) \in \mathbb{N}^* \quad B(p; q+1) = \dfrac{q}{p} B(p+1; q)$ puis en calculant la valeur de $B(p; 1)$ pour $p \in \mathbb{N}^*$. La propriété s'en déduit.
*Utilisations :*

- Si $(X_1 \dots X_n)$ sont des variables aléatoires indépendantes suivant chacune une loi $\mathcal{U}[0;1[$, alors $M = \max(X_i)$ suit une loi $B(n;1)$.
- Si $(X_1 \dots X_n)$ sont des variables aléatoires indépendantes suivant chacune une loi $\mathcal{U}[0;1[$, alors $W = \min(X_i)$ suit une loi $B(1;n)$.
- De façon générale, la loi Beta décrit la proportion de succès asymptotique du modèle de tirages successifs avec remise renforcée :
  *Dans une urne ayant initialement $b$ boules blanches et $r$ boules rouges, on procède à des tirages successifs avec remise de la boule sélectionnée tout en ajoutant une boule de même couleur que celle piochée.*
  *On s'intéresse à l'évolution des proportions de boules blanches et rouges de l'urne dans une telle expérience.*

# 4 Produit de convolution

**Définition (Produit de convolution)** : On se donne deux fonctions $f$ et $g$, intégrables sur $\mathbb{R}$. Le produit de convolution $f * g$ de $f$ par $g$ est défini comme fonction par :

$$f * g : \begin{array}{ccc} \mathbb{R} & \longrightarrow & \mathbb{R} \\ x & \mapsto & \displaystyle\int_{-\infty}^{+\infty} f(t)g(x-t)\mathtt{d}t \end{array}$$

**Exemple :** On donne pour exemple calculatoire le produit de convolution de $f(x) = e^{-x}\mathbb{1}_{\mathbb{R}_+}(x)$ avec $\mathbb{1}_{[-1;1]}$ :

$$\begin{aligned}
(f*g)(x) &= \int_{-\infty}^{+\infty} e^{-t}\mathbb{1}_{\mathbb{R}_+}(t)\mathbb{1}_{[-1;1]}(x-t)\mathtt{d}t \\
&= \mathbb{1}_{\mathbb{R}_+}(x)\left(\int_0^{+\infty} e^{-t}\mathbb{1}_{[-1;1]}(x-t)\mathtt{d}t\right) \\
&= \mathbb{1}_{\mathbb{R}_+}(x)\left(\int_{\max(x-1;0)}^{x+1} e^{-t}\right) \\
&= \mathbb{1}_{\mathbb{R}_+}(x)\left(e^{-\max(x-1;0)} - e^{-x-1}\right)
\end{aligned}$$

ce qui peut se réécrire :

$$(f*g)(x) = \begin{cases} 0 & \text{si } x < 0 \\ 1 - e^{-x-1} & \text{si } x \in [0;1] \\ e^{-x}\left(e - e^{-1}\right) & \text{si } x \geq 1 \end{cases}$$

**Propriété.** *Le produit de convolution de deux fonctions $f$ et $g$ nulles sur $\mathbb{R}_-$ est une fonction nulle sur $\mathbb{R}_-$. On a alors :*

$$\forall x \geq 0 \quad (f*g)(x) = \int_0^x f(t)g(x-t)\mathtt{d}t$$

Cette propriété permet d'avoir une formule simplifée de calcul de $(f * g)$ pour deux densités de probabilités à support inclus dans $\mathbb{R}_+$.

**Exercice-Exemple.** *Déterminer le produit de convolution des fonctions de densités de deux variables aléatoires indépendantes $U$ et $V$ suivant une même loi $\mathcal{U}[0;1[$.*

**Propriété.** *Le produit de convolution est une loi associative, commutative et admettant un élément neutre $\delta$ qui n'est pas une fonction de $\mathbb{R}$ dans $\mathbb{R}$ (mais une distribution au sens de Schwartz)*

On ne démontrera pas ce théorème qui manipule des objets hors programme (ne serait-ce que pour dire "interne" si on en avait envie)

**Théorème.** *Soient deux variables aléatoires $X$ et $Y$ indépendantes suivant des lois à densités respectives $f_X$ et $f_Y$.*
*Alors la fonction $f_X * f_Y$ est une densité de $X + Y$.*

**Exemple :** La loi de $S = U + V$ où $U$ et $V$ suivent des lois $\mathcal{U}[0; 1[$ est à densité, de support $[0; 2]$ et une densité est donnée par :

$$\forall x \in [0; 2] \quad f_S(x) = (\mathbb{1}_{[0;1]} * \mathbb{1}_{[0;1]})(x) = \int_{\max(x-1;0)}^{\min(1;x)} \mathrm{d}t$$

Il vient $f_S(x) = x\mathbb{1}_{[0;1[}(x) + (2 - x)\mathbb{1}_{[1;2]}(x)$ de façon plus tangible. Ce type de fonction est parfois appelé *triangle* ou *chapeau de clown*.

# 5 Introduction aux estimateurs

## 5.1 Présentation de la problématique

1. **Contexte :** On considère un phénomène aléatoire et on s'intéresse à une variable aléatoire $X$ réelle qui lui est liée.
   La loi de probabilité associée à $X$ appartient à une famille de lois dépendant d'un paramètre $\theta$ décrivant un ensemble $\Theta \subset \mathbb{R}$ (ou $\mathbb{R}^2$). *Exemple :* $X$ suit une loi exponentielle de paramètre $\lambda \in [0.2\ ;\ 1]$, associée à l'étude de machines dont la durée de vie espérée est une valeur fixée entre 1 et 5 ans. On cherche estimer cette espérance de vie.

2. **Objectif :** Déterminer la valeur du paramètre $\theta$ ou en donner une information partielle.
   Le problème de l'estimation consiste donc à estimer $\theta$ ou une fonction $g(\theta)$. Cette dernière pourra représenter, en général, une valeur caractéristique de la loi inconnue -comme son espérance, sa variance, son étendue ...
   *Exemple :* Dans le problème précédant, on cherche à estimer cette espérance de vie, valant $g(\lambda) = \frac{1}{\lambda}$

3. **Moyens :** On se donnera un échantillon de données $x_1, \ldots, x_n$ obtenues en observant $n$ fois le phénomène. On supposera que cet échantillon est la réalisation d'une famille de variables aléatoires $X_1, \ldots, X_n$ définies sur un même espace probabilisable $(\Omega\ ;\ \mathcal{A})$ que l'on pourra munir de n'importe quelle probabilité $\mathbb{P}_\theta$ où $\theta \in \Theta$.
   Ces variables aléatoires seront de même loi et mutuellement indépendantes au sein de $\mathbb{P}_\theta$, et ce, pour n'importe quel $\theta \in \Theta$, ce que l'on pourra résumer en disant :

   *Les variables aléatoires $X_1,\ \ldots,\ X_n$ sont indépendantes et identiquement distribuées.*

   *Exemple :* Dans le problème précédant, on a observé sur trois machines du même type des durées de fonctionnement respectives de $x_1 = 2$, $x_2 = 2,5$ et $x_3 = 2,25$ années (format décimal).
   On vient de donner une réalisation de $X_1$, $X_2$ et $X_3$ indépendantes et identiquement distribuées de loi exponentielle de paramètre $\lambda \in [0.2\ ;\ 1]$

## 5.2 Estimateur ponctuel

**Vocabulaire :** Si $X$ est une variable aléatoire donnée, dont la loi $\mathcal{L}$ dépend d'un -voire deux- paramètre(s), on nomme $n$-échantillon de $X$, où $n \geq 2$ est un entier naturel, une famille $(X_1\ ;\ \ldots\ ;\ X_n)$ de variables aléatoires indépendantes et identiquement distribuées selon la loi suivie par $X$.

On pourra éventuellement dire que $(X_1\ ;\ \ldots\ ;\ X_n)$ est un $n$-échantillon de la loi $\mathcal{L}$, sans se référer à une variable $X$.

**Définition (Estimateur)** : On appellera *estimateur* de $g(\theta)$ toute variable aléatoire réelle $T_n$ de la forme $T_n = \varphi(X_1; \ldots; X_n)$ où $\varphi$ est une fonction de $\mathbb{R}^n$ dans $\mathbb{R}$ (éventuellement dépendante de $n$) et indépendante de $\theta$ dont la réalisation après expérience est envisagée comme estimation de $g(\theta)$.

*Remarques :*

1. Si $T_n$ est un estimateur, il peut admettre une espérance, une variance , des moments d'ordre $r$ en tant que variable aléatoire. En revanche, ces valeurs, si elles existent, dépendent de la probabilité $\mathbb{P}_\theta$.
2. Une *estimation* de $g(\theta)$ est donc une réalisation $(x_1 \ldots x_n)$ de $(X_1 \ldots X_n)$.

**Notation :** Si $T_n$ est un estimateur, on écrira, sous couvert d'existence, $\mathbb{E}_\theta[T_n]$ son espérance et $\mathbb{V}_\theta[T_n]$ sa variance.

**Exercice-Exemple.** Moyenne Empirique :
*Si* $(X_1\ ;\ \dots\ ;\ X_n)$ *est un* $n$*-échantillon de* $X$*, on appelle* moyenne empirique *notée* $\overline{X}_n$ *l'estimateur défini par :*

$$\overline{X}_n = \frac{1}{n}\sum_{k=1}^{n} X_k$$

Paramètres de la variance empirique.

*On considère* $X$ *une variable aléatoire suivant une loi* $\mathcal{L}$ *admettant une espérance* $\mu$ *et une variance* $\sigma^2$*. Pour* $n \geq 2$*, on se donne un* $n$*-échantillon* $(X_1\ ;\ \dots\ ;\ X_n)$ *de* $X$*.*

1. *On cherche à estimer* $\theta = \sigma^2$*.*
   *Etablir que la variance empirique* ${S_n}^2$ *admet une espérance et vérifier que* $\mathbb{E}_\theta[S_n^2] = \dfrac{n-1}{n}\sigma^2$*.*
2. *Expliquer pourquoi les logiciels de calcul numérique proposent une "variance corrigée" s'écrivant* ${\sigma^2}_{n-1} = \frac{n}{n-1}{S_n}^2$

**Exercice-Exemple.** Variance Empirique :
*Si* $(X_1\ ;\ \dots\ ;\ X_n)$ *est un* $n$*-échantillon de* $X$*, on appelle* variance empirique *notée* $\overline{X}_n$ *l'estimateur défini par :*

$${S_n}^2 = \frac{1}{n}\sum_{k=1}^{n}(X_k - \overline{X}_n)^2$$

Paramètres de la moyenne empirique
*On considère* $X$ *une variable aléatoire suivant une loi* $\mathcal{L}$ *admettant une espérance* $\mu$ *et une variance* $\sigma^2$*. Pour* $n \geq 2$*, on se donne un* $n$*-échantillon* $(X_1\ ;\ \dots\ ;\ X_n)$ *de* $X$*.*

1. *On cherche à estimer* $\theta = \mu$*. Etablir que la moyenne empirique* $\overline{X}_n$ *admet une espérance et vérifier que* $\mathbb{E}_\theta[\overline{X}_n] = \mu$*.*
2. *On cherche à estimer* $\theta = \sigma^2$*. Etablir que la moyenne empirique* $\overline{X}_n$ *admet un moment quadratique et vérifier que* $\mathbb{V}_\theta[\overline{X}_n] = \frac{\sigma^2}{n}$*.*

> **Définition (Biais)** :
> Soit $T_n$ un estimateur de $X$. Si pour tout $\theta \in \Theta$, l'estimateur $T_n$ admet une espérance $\mathbb{E}_\theta[T_n]$ alors on nommera *biais* le réel défini par :
> $$b_\theta(T_n) = \mathbb{E}_\theta[T_n] - g(\theta)$$

**Vocabulaire :** Si le biais d'un estimateur $T_n$ est nul ($b_\theta(T_n) = 0$) pour tout $\theta \in \Theta$, on dira que cet estimateur est *sans biais*

**Exercice-Exemple.** Moyenne Empirique :
*La moyenne empirique* $\overline{X}_n$ *d'une variable* $X$ *est un estimateur sans biais de l'espérance* $\mu$ *de la loi de* $X$*.*

**Exercice-Exemple.** Variance Empirique :

1. *La variance empirique* ${S^2}_n$ *d'une variable* $X$ *est un estimateur qui n'est pas sans biais de la variance* $\sigma^2$ *de la loi de* $X$*.*
   *On calculera le biais de la variance empirique* ${S_n}^2$*, estimateur d'une variable* $X$ *de loi* $\mathcal{L}$ *dont la variance est le paramètre* $\sigma^2$*.*
2. *L'estimateur* variance corrigée *défini comme* $\sigma_{n-1} = \dfrac{n}{n-1}{S_n}^2$ *est un estimateur sans biais de cette même variance.*

**Vocabulaire :** On pourra dire *biaisé* au lieu de "n'est pas sans biais".

> **Définition** :
> Une suite d'estimateurs $(T_n)_{n\in\mathbb{N}}$ est une suite de variables aléatoires telle que chaque $T_n$ est un estimateur associé à un $n$-échantillon d'une même variable aléatoire $X$, de la forme $T_n = \varphi(X_1; \dots; X_n)$.

**Propriété.** *Les estimateurs* moyenne empirique *et* variance empirique *d'une variable aléatoire* $X$ *induisent des suites d'estimateurs* $(\overline{X}_n)_{n\geq 1}$ *et* $({S_n}^2)_{n\geq 1}$ *respectivement, quitte à pose* $\overline{X}_1 = X_1$ *et* ${S_1}^2 = 0$

> **Définition** : Une suite d'estimateurs $(T_n)_{n\in\mathbb{N}}$ de $g(\theta)$ est dite convergente si, pour tout $\theta$ de $\Theta$, la suite de variables aléatoire $(T_n)_{n\in\mathbb{N}}$ converge en probabilités vers $g(\theta)$.

On pourra (par abus de langage) dire tout simplement que *l'estimateur est convergent*.

Ainsi, on pourra retenir que cela signifie formellement :

$$\forall\theta\in\Theta\ \forall\varepsilon>0\ \lim_{n\to+\infty}\mathbb{P}_\theta[|T_n-g(\theta)|\geq\varepsilon]=0$$

**Exercice-Exemple.** *Démontrer que l'estimateur $\overline{X}_n$ de l'espérance $\mu$ d'une variable aléatoire $X$ est convergent.*
Indication : *Utiliser la loi (faible) des grands nombres.*