

ECOLE NATIONALE POLYTECHNIQUE
DE CONSTANTINE

COURS DE PROBABILITES
(SEMESTRE 2)

ZAHER MOHDEB

E-mail: z.mohdeb@gmail.com, zaher.mohdeb@umc.edu.dz

Chapitre 1 Notion de variables aléatoires réelles

1- Introduction et définition

En théorie des probabilités, **une variable aléatoire réelle** est une application définie sur l'ensemble des résultats possibles d'une expérience aléatoire noté Ω à valeurs dans \mathbb{R} . Les jeux de hasard tels que le lancer d'un dé, d'un tirage à pile ou face, d'une roulette, . . . qui amènent à concevoir les variables aléatoires, en associant à une éventualité un gain. Cette association éventualité-gain a donné lieu par la suite à la conception d'une fonction de portée plus générale. La fonction décrivant les valeurs possibles d'une variable aléatoire et leur probabilité est connue sous le nom de la **distribution de probabilité**.

Notion de variables aléatoires réelles

Les variables aléatoires peuvent être **discrètes**, qui est, en prenant tous les éléments d'une liste finie ou dénombrable de valeurs spécifiée, doté d'une fonction de masse de probabilité, caractéristique d'une distribution de probabilités ; ou **continues**, en prenant une valeur numérique dans un intervalle ou d'une famille d'intervalles, par l'intermédiaire d'une fonction de densité de probabilité qui est une caractéristique de la distribution de probabilités ; ou un mélange des deux types.

Ainsi la notion de variable aléatoire permet de créer une relation entre un espace probabilisé et un espace mesurable qui est presque toujours l'espace \mathbb{R} ou \mathbb{R}^n . Par cette relation, l'espace qui n'était que mesurable devient aussi probabilisé. Dans ce cas, la variable aléatoire présente le grand intérêt pratique d'associer des nombres à une expérience soumise au hasard dont les issues sont abstraites. Pour un expérimentateur, l'espace probabilisé est le phénomène étudié et l'espace mesurable sera l'ensemble de tous les résultats chiffrés des expériences portant sur le phénomène en question.

Définition

Soit X l'application définie de Ω dans \mathbb{R} , on note X^{-1} l'application de $\mathcal{P}(\mathbb{R})$ dans $\mathcal{P}(\Omega)$ qui à $B \subset \mathbb{R}$ associe son *image réciproque* par X : $X^{-1}(B) = \{\omega \in \Omega / X(\omega) \in B\}$.

Remarque : Il s'agit d'une *application réciproque ensembliste* mais pas de l'application réciproque d'une application bijective.

Propriétés

Soient X une application de Ω dans \mathbb{R} , A, B deux sous-ensembles de \mathbb{R} et $A_i \subset \mathbb{R}, \forall i \in I \subset \mathbb{N}$, alors

$$1) X^{-1}(\emptyset) = \emptyset$$

$$2) X^{-1}(\mathbb{R}) = \Omega$$

$$3) X^{-1}(C_{\mathbb{R}}B) = C_{\Omega}X^{-1}(B)$$

$$4) X^{-1}\left(\bigcap_{i \in I} A_i\right) = \bigcap_{i \in I} X^{-1}(A_i)$$

$$5) X^{-1}\left(\bigcup_{i \in I} A_i\right) = \bigcup_{i \in I} X^{-1}(A_i)$$

$$6) X^{-1}(A \Delta B) = X^{-1}(A) \Delta X^{-1}(B).$$

Tribu engendrée, tribu borélienne sur \mathbb{R}

Proposition

(**et définition**) i) L'intersection d'une collection non vide quelconque de tribus de parties de Ω est elle-même une tribu.

ii) Pour toute classe \mathcal{C} de parties de Ω , l'intersection de toutes les tribus contenant \mathcal{C} est donc une tribu : elle est appelée la plus petite tribu contenant \mathcal{C} , ou *tribu engendrée par \mathcal{C}* et notée $\sigma(\mathcal{C})$:

$$\sigma(\mathcal{C}) := \bigcap_{\mathcal{A} \text{ tribu, } \mathcal{C} \subseteq \mathcal{A}} \mathcal{A}.$$

Définition

On appelle *tribu borélienne sur \mathbb{R}* (ou *tribu des boréliens de \mathbb{R}*) la tribu engendrée par les intervalles ouverts de \mathbb{R} et on note $\mathcal{B}(\mathbb{R})$ ou \mathcal{B} .

Proposition

La tribu $\mathcal{B}_{\mathbb{R}}$ des boréliens de \mathbb{R} est engendrée par

- i) la famille des intervalles ouverts bornés,
- ii) la famille des intervalles de la forme $] - \infty, a[$ avec $a \in \mathbb{R}$,
- iii) la famille des intervalles de la forme $] - \infty, a]$ avec $a \in \mathbb{R}$.

Définition

Soit (Ω, \mathcal{A}, P) un espace probabilisé. On appelle **variable aléatoire réelle** définie sur (Ω, \mathcal{A}) , l'application X de Ω dans \mathbb{R}

$(X : (\Omega, \mathcal{A}) \rightarrow (\mathbb{R}, \mathcal{B}_{\mathbb{R}}))$, telle que l'image réciproque par X de tout intervalle $] - \infty, x]$ est un événement de \mathcal{A} , c'est-à-dire $X^{-1}(] - \infty, x]) \in \mathcal{A}, \forall x \in \mathbb{R}$.

Remarque : Dire que X est une **variable aléatoire réelle** sur (Ω, \mathcal{A}) revient à dire $\forall B \in \mathcal{B}_{\mathbb{R}}, X^{-1}(B) \in \mathcal{A}$, puisque la famille des intervalles de la forme $] - \infty, x]$, avec $x \in \mathbb{R}$, engendrent la tribu des boréliens $\mathcal{B}_{\mathbb{R}}$.

Notion de variables aléatoires réelles

Exemple : Soient (Ω, \mathcal{A}, P) un espace probabilisé et A un événement de \mathcal{A} . On appelle **fonction indicatrice de l'événement A** la fonction notée \mathbb{I}_A (ou I_A) définie de (Ω, \mathcal{A}) dans $(\mathbb{R}, \mathcal{B}_{\mathbb{R}})$ par

$$\mathbb{I}_A(\omega) = \begin{cases} 1 & \text{si } \omega \in A, \\ 0 & \text{si } \omega \notin A. \end{cases}$$

La fonction indicatrice \mathbb{I}_A est clairement une variable aléatoire car pour tout borélien B dans $\mathcal{B}_{\mathbb{R}}$, on a

$$(\mathbb{I}_A)^{-1}(B) = \{\omega \in \Omega : \mathbb{I}_A(\omega) \in B\} = \begin{cases} A & \text{si } 0 \notin B \text{ et } 1 \in B, \\ C_{\Omega}A & \text{si } 0 \in B \text{ et } 1 \notin B, \\ \Omega & \text{si } 0 \in B \text{ et } 1 \in B, \\ \emptyset & \text{si } 0 \notin B \text{ et } 1 \notin B \end{cases} \quad (1)$$

et \emptyset , A , $C_{\Omega}A$ et Ω sont des événements de \mathcal{A} .

Ou bien, on peut faire le raisonnement suivant. Soit x un nombre réel quelconque, on a :

$$\begin{aligned}(\mathbb{I}_A)^{-1}(] - \infty, x]) &= \{\omega \in \Omega : \mathbb{I}_A(\omega) \leq x\} \\ &= \begin{cases} \emptyset & \text{si } x < 0, \\ C_{\Omega}A & \text{si } 0 \leq x < 1, \\ \Omega & \text{si } x \geq 1 \end{cases}\end{aligned}$$

et \emptyset , $C_{\Omega}A$ et Ω sont des événements de \mathcal{A} .

Proposition

Soit X une variable aléatoire réelle définie sur un espace probabilisé (Ω, \mathcal{A}, P) à valeurs dans un espace probabilisable $(\mathbb{R}, \mathcal{B}_{\mathbb{R}})$. Alors la famille des parties $\{X^{-1}(B) : B \in \mathcal{B}_{\mathbb{R}}\}$ est une tribu (ou σ -algèbre) de Ω , appelée *tribu engendrée par X* .

On note $\sigma(X)$ cette tribu de parties de Ω et on remarquera que $\sigma(X)$ est une sous-tribu de la tribu \mathcal{A} de Ω .

Exemple : Reprenons l'exemple de la fonction indicatrice \mathbb{I}_A définie de (Ω, \mathcal{A}) dans $(\mathbb{R}, \mathcal{B}_{\mathbb{R}})$.

D'après la relation (1), la tribu engendrée par la fonction indicatrice \mathbb{I}_A est donnée par $\sigma(\mathbb{I}_A) = \{\emptyset, \Omega, A, C_{\Omega}A\}$.

2- Variables aléatoires discrètes

2.1- Préliminaires- Définitions : - Considérons l'expérience qui consiste à jeter trois fois une pièce, la probabilité de tomber sur "pile" étant p à chaque fois. Supposons qu'à chaque jet donnant "pile", on gagne un Dinar et qu'à chaque jet donnant "face" on en perde un.

Intéressons nous au gain total, soit X cette quantité. L'espace des épreuves Ω correspondant est formé des éléments suivant :

PPP, PPF, PFP, FPP, PFF, FPF, FFP, FFF.

La quantité X ne peut prendre que les valeurs $+3, +1, -1, -3$; mais nous ne pouvons dire avec certitude laquelle; cela dépend de résultat de notre expérience.

Posons $q = 1 - p$ et résumons dans le tableau suivant la liste des valeurs de X correspondant à chacun des événements avec leurs probabilités respectives.

Notion de variables aléatoires réelles

ω	<i>PPP</i>	<i>PPF</i>	<i>PFP</i>	<i>FPP</i>	<i>PFF</i>	<i>FPF</i>	<i>FFP</i>	<i>FFF</i>
$X(\omega)$	3	1	1	1	-1	-1	-1	-3
$P(\{\omega\})$	p^3	p^2q	p^2q	p^2q	pq^2	pq^2	pq^2	q^3

X est donc une fonction à valeurs réelles définie sur l'espace de probabilité correspondant à l'expérience.

Considérons l'événement $\{\omega \in \Omega / X(\omega) = -1\}$, il est formé de trois éléments *PPF*, *FPF*, *FFP* et donc la probabilité de cet événement est égale à $3pq^2$ et nous écrivons en abrégé

$$P(X = -1) = 3pq^2.$$

A chaque valeur possible x de X est attaché le nombre $P(X = x)$.
 X est un exemple de variable aléatoire (en abrégé v.a.) discrète.

Définition

Une *variable aléatoire (v.a.) discrète* sur un espace probabilisé (Ω, \mathcal{A}, P) est une application définie sur Ω prenant un nombre fini ou dénombrable de valeurs réelles $\{x_1, x_2, \dots\}$ telle que $\forall x_i \in X(\Omega)$, $X^{-1}(\{x_i\}) \in \mathcal{A}$.

Définition

Soit X une variable aléatoire réelle discrète sur (Ω, \mathcal{A}, P) et $X(\Omega) = \{x_i, i \in I\}$. On appelle *tribu engendrée par X* , notée $\sigma(X)$ (ou \mathcal{A}_X), la sous-tribu de \mathcal{A} engendrée par la famille des $\{X^{-1}(\{x_i\}), i \in I\}$.

Remarque : La tribu engendrée par X apparaît comme l'ensemble des événements qui peuvent être décrits au moyen de X .

2.2- Loi de probabilité d'une variable aléatoire réelle

Définition

Soit X une variable aléatoire définie sur l'espace probabilisé (Ω, \mathcal{A}, P) à valeurs dans l'espace probabilisable $(\mathbb{R}, \mathcal{B}_{\mathbb{R}})$, (où $\mathcal{B}_{\mathbb{R}}$ est la tribu des boréliens de \mathbb{R}), on appelle *loi de probabilité (ou distribution)* de la variable aléatoire X , la *probabilité* P_X sur $(\mathbb{R}, \mathcal{B}_{\mathbb{R}})$ définie par

$$\begin{aligned}\forall B \in \mathcal{B}_{\mathbb{R}}, \quad P_X(B) &= P[X^{-1}(B)] \\ &= P(\{\omega \in \Omega / X(\omega) \in B\}) \stackrel{\text{notation}}{=} P(X \in B).\end{aligned}$$

Notation : $P_X(\{x\}) = P(X = x)$, $P_X(B) = P(X \in B)$.

Remarque : Si X est une variable aléatoire réelle discrète définie sur (Ω, \mathcal{A}, P) telle que $X(\Omega) = \{x_i, i \in I \subseteq \mathbb{N}\}$, la *loi de probabilité de X* est déterminée par la donnée de tous les $(P(X = x_i), x_i \in X(\Omega))$.

Définition

La fonction f définie sur \mathbb{R} par $f(x) = P(X = x)$ est appelée *fonction densité discrète* de X .

Propriétés : La densité discrète f d'une variable aléatoire X vérifie

a) $f(x) \geq 0, \forall x \in \mathbb{R}$,

b) $\{x \in \mathbb{R}/f(x) \neq 0\}$ est un sous-ensemble fini ou dénombrable $\{x_1, x_2, \dots\}$ de \mathbb{R} .

c) $\sum_i f(x_i) = 1$.

Définition

Deux variables aléatoires réelles discrètes X et Y définies sur le même espace probabilité (Ω, \mathcal{A}, P) sont dites *presque sûrement égales* si

$$P(\{\omega \in \Omega/X(\omega) \neq Y(\omega)\}) = 0, \quad \text{on note } X = Y \quad \text{p.s.}$$

2.3- Loi usuelles discrètes

- Loi de Bernoulli

Définition

Une expérience qui a *deux résultats possibles* (représentés en général par *échec et succès*) est appelée *expérience de Bernoulli*.

Remarque : Si on note échec par "e" et succès par "s", l'espace probabilisé correspondant à cette expérience est (Ω, \mathcal{A}, P) , où $\Omega = \{e, s\}$, $\mathcal{A} = \mathcal{P}(\Omega)$ et la probabilité P est déterminée par $P(\{s\}) = p$ et $P(\{e\}) = q = 1 - p$.

Définition

L'application X de Ω dans $\{0, 1\}$ telle que $X(e) = 0$, $X(s) = 1$, $P(X = 1) = p$ et $P(X = 0) = 1 - p$ est appelée *variable aléatoire de Bernoulli de paramètre p* notée $\mathcal{B}(p)$.

- Loi binomiale

Considérons n répétitions indépendantes d'une expérience de Bernoulli (pile-face, succès-échec, boule rouge - boule blanche) de probabilités respectives p et $q = 1 - p$.

$$X_i = \begin{cases} 1 & \text{si succès à la } i^{\text{eme}} \text{ épreuve avec une probabilité } p \\ 0 & \text{si échec à la } i^{\text{eme}} \text{ épreuve avec une probabilité } q = 1 - p, \end{cases}$$

$i = 1, 2, \dots, n$.

Soit $S_n = \sum_{i=1}^n X_i$: le nombre de succès au cours des n répétitions. Il est clair que S_n peut prendre les valeurs $0, 1, 2, \dots, n$. Essayons de déterminer la probabilité $P(S_n = k)$, où $k \in \{0, 1, \dots, n\}$.

On a $S_n = k$, si les k variables X_i prenant la valeur 1, les $(n - k)$ autres prenant la valeur 0. L'événement $\{S_n = k\}$ peut donc être considéré comme la réunion de C_n^k événements disjoints tels que k et k seulement des X_i prennent la valeur 1. Ces événements sont équiprobables et il suffit de calculer la probabilité de l'un d'eux.

Notion de variables aléatoires réelles

On a aussi, compte tenu de l'indépendance des X_i (car pour tout $i = 1, \dots, n$, X_i est lié à la i^{ieme} répétition et que les n répétitions sont indépendantes)

$$\begin{aligned} & P\{(X_1 = 1) \cap (X_2 = 1) \cap \dots \cap (X_k = 1) \cap (X_{k+1} = 0) \cap \dots \cap (X_n = 0)\} \\ &= P(X_1 = 1)P(X_2 = 1) \dots P(X_k = 1)P(X_{k+1} = 0) \dots P(X_n = 0) \\ &= p^k q^{n-k}, \end{aligned}$$

d'où

$$P(S_n = k) = \begin{cases} \binom{n}{k} p^k q^{n-k} & \text{si } k = 0, 1, \dots, n \\ 0 & \text{sinon.} \end{cases} \quad (2)$$

Cette distribution est l'une des plus importantes de la théorie des probabilités, appelée **distribution binomiale** car son k^{ieme} terme $b(k, n, p)$ est le k^{ieme} terme du développement de $(p + q)^n$. On note $S_n \sim \mathcal{B}(n, p)$.

Remarques

i) La loi binomiale décrit des phénomènes ne pouvant prendre que deux états s'excluant mutuellement, par exemple : succès ou échec dans un jeu, bonne pièce ou pièce défectueuse dans une fabrication, lot acceptable ou lot refusé, défaillance ou fonctionnement d'un matériel, etc. . .

ii) Elle est utilisée dans le domaine technique pour déterminer la probabilité de défaillance à la sollicitation de matériels, en contrôle qualité, mais elle ne peut s'appliquer rigoureusement que si les expériences sont non exhaustives, c'est la loi du tirage avec remise.

iii) Les événements considérés doivent être indépendants et la probabilité de réalisation d'un événement doit être constante.

- Loi hypergéométrique

Considérons une population de N objets parmi lesquels M sont d'un type et $N - M$ d'un second type.

Supposons que nous tirions un échantillon aléatoire (tirage sans remise) de taille $n \leq N$ de la population.

Soit X le nombre d'objets du 1^{er} type dans l'échantillon de taille n . Il est clair que le nombre k d'objets du premier type et donc $(n - k)$ objets du second type dans l'échantillon de taille n doivent vérifier les conditions suivantes :

$$\left\{ \begin{array}{l} 0 \leq k \leq n \\ 0 \leq k \leq M \\ 0 \leq n - k \leq N - M \end{array} \right. \iff \left\{ \begin{array}{l} 0 \leq k \leq n \\ 0 \leq k \leq M \\ n - N + M \leq k \leq n. \end{array} \right.$$

Notion de variables aléatoires réelles

Donc X est une variable aléatoire dont les valeurs possibles varient entre $\max(0, n - N + M)$ et $\min(n, M)$, avec

$$P(X = k) = \begin{cases} \frac{C_M^k C_{N-M}^{n-k}}{C_N^n} & \text{si } k = \max(0, n - N + M), \dots, \min(n, M) \\ 0 & \text{sinon,} \end{cases} \quad (3)$$

où n, M, N sont des entiers tels que $n \leq N$ et $M \leq N$. On dira que X suit une loi hypergéométrique de paramètres $N, n, \frac{M}{N}$ et on note $X \sim \mathcal{H}(N, n, \frac{M}{N})$.

Remarque : Notons que si $n \leq M$ et $n \leq N - M$, alors la formule donnée par (3) s'écrit

$$P(X = k) = \begin{cases} \frac{C_M^k C_{N-M}^{n-k}}{C_N^n} & \text{si } k = 0, 1, \dots, n \\ 0 & \text{sinon.} \end{cases} \quad (4)$$

- Loi de Poisson

Définition

Soit λ un nombre réel positif. On dit que la variable aléatoire X suit une loi de Poisson de paramètre λ , si X est une variable aléatoire discrète prenant la valeur entière n avec la probabilité

$$P(X = k) = \begin{cases} \frac{e^{-\lambda} \lambda^k}{k!} & \text{si } k = 0, 1, 2, \dots \\ 0 & \text{sinon.} \end{cases} \quad (5)$$

On note $X \sim \mathcal{P}(\lambda)$.

Remarques : 1) - Notons que la suite $p_n = P(X = n)$ définit bien une loi de probabilité. En effet, on a : i) $p_n \geq 0, \forall n \in \mathbb{N}$,

$$\text{ii) } \sum_{n=0}^{\infty} p_n = \sum_{n=0}^{\infty} \frac{e^{-\lambda} \lambda^n}{n!} = e^{-\lambda} \sum_{n=0}^{\infty} \frac{\lambda^n}{n!} = e^{-\lambda} e^{\lambda} = 1.$$

2) - Sous certaines conditions, une distribution binomiale peut être approchée par une distribution de Poisson ($n > 50$ et $p < 0,10$).

En effet, si $X \sim \mathcal{B}(n, p)$, on a

$$\begin{aligned}P(X = k) &= C_n^k p^k (1 - p)^{n-k} \\&= \frac{1}{k!} n(n-1) \dots (n-k+1) p^k (1-p)^{n-k} \\&= \frac{n^k}{k!} \left(1 - \frac{1}{n}\right) \left(1 - \frac{2}{n}\right) \dots \left(1 - \frac{k-1}{n}\right) p^k (1-p)^{n-k}.\end{aligned}$$

Posons $\lambda = np$ et supposons que λ reste **constant** lorsque n augmente (ce qui sous-entend que p décroît).

Notion de variables aléatoires réelles

On a

$$P(X = k) = \frac{1}{k!} \left(1 - \frac{1}{n}\right) \left(1 - \frac{2}{n}\right) \dots \left(1 - \frac{k-1}{n}\right) \lambda^k \left(1 - \frac{\lambda}{n}\right)^n \frac{1}{\left(1 - \frac{\lambda}{n}\right)^k}$$

Lorsque $n \rightarrow \infty$, $\left(1 - \frac{\lambda}{n}\right)^n \rightarrow e^{-\lambda}$, d'où

$$P(X = k) = C_n^k p^k (1 - p)^{n-k} \rightarrow \frac{e^{-\lambda} \lambda^k}{k!}, \quad \text{quand } n \rightarrow \infty.$$

La distribution binomiale $\mathcal{B}(n, p)$ peut être approximée par une distribution de Poisson de paramètre $\lambda = np$, dès que $n > 50$ et $p < 0,10$.

- Loi géométrique

Supposons qu'une expérience admette deux résultats possibles appelés succès et échec notés respectivement "s" et "e" tels que $P(s) = p$ et $P(e) = 1 - p = q$.

Répetons cette expérience (répétitions indépendantes) et définissons la variable aléatoire X représentant "le nombre de répétitions de l'expérience nécessaires à l'obtention du premier succès".

Pour calculer $P(X = k)$, on note A_k l'événement : "obtenir un succès à la $k^{ième}$ répétition de l'expérience". A cause de l'indépendance des répétitions de l'expérience, il est clair que

$$\begin{aligned}P(X = k) &= P(\bar{A}_1 \cap \dots \cap \bar{A}_{k-1} \cap A_k) \\ &= P(\bar{A}_1) \dots P(\bar{A}_{k-1})P(A_k) \\ &= (1 - p)^{k-1}p.\end{aligned}$$

Définition

On dit qu'une variable aléatoire X suit une *loi géométrique de paramètre p* et on note $X \sim \mathcal{G}(p)$, si

$$P(X = k) = (1 - p)^{k-1}p, \quad \forall k \in \mathbb{N}^* = \{1, 2, \dots\}. \quad (6)$$

Remarque : Notons que la suite $p_k = P(X = k)$, $k \in \mathbb{N}^*$ définit bien une loi de probabilité.

En effet, on a :

i) $p_k \geq 0, \forall k \in \mathbb{N}^*$,

ii)
$$\sum_{k=1}^{+\infty} p_k = \sum_{k=1}^{+\infty} (1 - p)^{k-1}p = p \sum_{k=1}^{+\infty} (1 - p)^{k-1} = p \frac{1}{1 - (1 - p)} = 1.$$

3- Variables aléatoires continues, fonction de répartition

3.1- Définition, fonction de répartition

Dans la section précédente, nous avons considéré des variables aléatoires discrètes, mais on peut trouver de nombreuses situations dans lesquelles les variables aléatoires naturelles à considérer sont continues, en particulier les mesures de quantités physiques comme par exemple : les coordonnées spatiales, le poids, le temps, la température, la durée de vie d'une batterie de voiture, l'heure d'arrivée des voitures à un péage donné d'autoroute, sont mieux décrites par des variables aléatoires continues, c'est-à-dire des variables aléatoires dont les valeurs possibles sont des intervalles (ou une réunion d'intervalles) de \mathbb{R} .

Définition

Soit (Ω, \mathcal{A}, P) un espace probabilisé. Une variable aléatoire X continue est une application de (Ω, \mathcal{A}, P) dans \mathbb{R} pour laquelle la fonction F_X qui à x appartenant à \mathbb{R} associe $P(\{\omega \in \Omega / X(\omega) \leq x\}) \stackrel{\text{notation}}{=} P(X \leq x)$ est bien définie et continue.

La fonction F_X est appelée *fonction de répartition* (en abrégé f.r.) de la variable aléatoire X ,

$$F_X(x) = P(X \leq x), \quad \forall x \in \mathbb{R}.$$

Propriétés

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

En effet, on a $F_X(+\infty) = \lim_{x \rightarrow +\infty} P(\{\omega \in \Omega / X(\omega) \leq x\}) = P(\Omega) = 1$

$$\text{et } F_X(-\infty) = \lim_{x \rightarrow -\infty} P(\{\omega \in \Omega / X(\omega) \leq x\}) = P(\emptyset) = 0.$$

ii) $\forall x_1, x_2 \in \mathbb{R}$ tels que $x_1 < x_2$, on a

$$P(x_1 < X \leq x_2) = F_X(x_2) - F_X(x_1).$$

En effet, on a

$$\{\omega \in \Omega / X(\omega) \leq x_2\} = \{\omega \in \Omega / X(\omega) \leq x_1\} \cup \{\omega \in \Omega / x_1 < X(\omega) \leq x_2\}$$

$$\implies P(X \leq x_2) = P(X \leq x_1) + P(x_1 < X \leq x_2)$$

$$\implies P(x_1 < X \leq x_2) = F_X(x_2) - F_X(x_1).$$

iii) $P(x_1 < X < x_2) = F_X(x_2^-) - F_X(x_1)$, où $F_X(x_2^-) = \lim_{x \nearrow x_2} F_X(x)$.

En effet, on a $\{\omega \in \Omega / X(\omega) < x_2\} = \{\omega \in \Omega / X(\omega) \leq x_1\} \cup \{\omega \in \Omega / x_1 < X(\omega) < x_2\}$,

comme $P(X < x_2) = \lim_{x \nearrow x_2} P(X \leq x) = \lim_{x \nearrow x_2} F_X(x) = F_X(x_2^-)$,

d'où

$$F_X(x_2^-) = F_X(x_1) + P(x_1 < X < x_2).$$

iv) F_X est monotone non décroissante.

En effet, soit $x_1, x_2 \in \mathbb{R}$ tels que $x_1 < x_2$, on a

$$0 \leq P(x_1 < X \leq x_2) = F_X(x_2) - F_X(x_1) \implies F_X(x_1) \leq F_X(x_2).$$

v) F_X est continue à droite.

En effet, soit $(x_n)_{n \in \mathbb{N}}$ une suite de nombres réels tels que $x_n \xrightarrow{\geq} x$, quand $n \rightarrow +\infty$, il s'agit de montrer que $\lim_{n \rightarrow +\infty} F_X(x_n) = F_X(x)$.

On a $x_n \searrow x$ et $x_0 > x_1 > x_2 > \dots > x_n > \dots$

Considérons les intervalles $]x_i, x_{i-1}]$, $i = 1, 2, \dots$. Ces intervalles sont deux à deux disjoints et $]x, x_0] = \bigcup_{i \geq 1}]x_i, x_{i-1}]$, donc

$$\begin{aligned} F_X(x_0) - F_X(x) &= P(X \in]x, x_0]) = P\left(X \in \bigcup_{i \geq 1}]x_i, x_{i-1}]\right) \\ &= \sum_{i \geq 1} P(X \in]x_i, x_{i-1}]) = \sum_{i \geq 1} P(x_i < X \leq x_{i-1}) \\ &= \sum_{i \geq 1} [F_X(x_{i-1}) - F_X(x_i)] \\ &= \lim_{n \rightarrow +\infty} \sum_{i=1}^n [F_X(x_{i-1}) - F_X(x_i)] \end{aligned}$$

d'où

$$\begin{aligned}F_X(x_0) - F_X(x) &= \lim_{n \rightarrow +\infty} \left[F_X(x_0) - F_X(x_1) + F_X(x_1) - F_X(x_2) \right. \\ &\quad \left. + \dots + F_X(x_{n-1}) - F_X(x_n) \right] \\ &= \lim_{n \rightarrow +\infty} \left[F_X(x_0) - F_X(x_n) \right] \\ &= F_X(x_0) - \lim_{n \rightarrow +\infty} F_X(x_n),\end{aligned}$$

ainsi

$$F_X(x) = \lim_{n \rightarrow +\infty} F_X(x_n).$$

Notion de variables aléatoires réelles

vi) $P(X = a) = 0 \iff F_X$ est continue en a .

En effet,

a) Supposons que $P(X = a) = 0$, montrons que F_X est continue en a .

On sait que F_X est continue à droite en tout point de \mathbb{R} , il s'agit donc de montrer que F_X est continue à gauche en a .

On a

$$\begin{aligned} \lim_{x \nearrow a} F_X(x) &= \lim_{x \nearrow a} P(X \leq x) = P(X \leq a^-) \\ &= P(X \leq a^-) + P(X = a) = P(X \leq a). \end{aligned}$$

b) Supposons maintenant que F_X est continue en a , montrons que $P(X = a) = 0$.

On a, d'après la continuité de F_X en a ,

$$\lim_{x \nearrow a} F_X(x) = F_X(a) = P(X \leq a)$$

et

$$\begin{aligned}P(X \leq a) &= P(X = a) + P(X \leq a^-) \\F_X(a) &= P(X = a) + \lim_{x \nearrow a} F_X(x) \\&= P(X = a) + F_X(a),\end{aligned}$$

d'où $P(X = a) = 0$.

Remarque

On en déduit que si X est une v.a. réelle continue, alors $P(X = x) = 0$, $\forall x \in \mathbb{R}$.

Types de fonction de répartition

(i) Fonction de répartition d'une variable aléatoire discrète

Soit X une variable aléatoire discrète définie de (Ω, \mathcal{A}, P) dans $(\mathbb{R}, \mathcal{B}_{\mathbb{R}})$, avec $X(\Omega) = \{x_1, x_2, \dots, x_n, \dots\}$. La fonction de répartition de la variable aléatoire X s'écrit

$$F_X(x) = P(X \leq x) = P\left(\bigcup_{x_i \leq x} \{X = x_i\}\right) = \sum_{x_i \leq x} P(X = x_i).$$

F_X est donc une **fonction en escalier** constante sur tout intervalle ne contenant pas de points de $X(\Omega)$ et admettant un saut d'amplitude égal à $P(X = x_i)$ en chaque point x_i , $i = 1, 2, \dots, n, \dots$. Le **saut en un point x_i** est égal à

$$F_X(x_i) - F_X(x_i^-) = P(x_i^- < X \leq x_i) = P(X = x_i).$$

Notion de variables aléatoires réelles

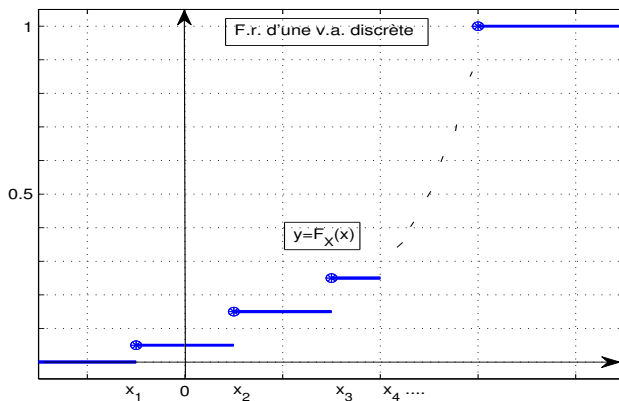


FIGURE: Fonction de répartition d'une variable aléatoire réelle discrète.

(ii) Fonction de répartition d'une variable aléatoire continue

Définition

Soit F la fonction de répartition d'une variable aléatoire réelle X , s'il existe une fonction f positive, intégrable au sens de Riemann et telle que

$$\forall x \in \mathbb{R}, \quad F(x) = \int_{-\infty}^x f(t) dt,$$

f est appelée *densité de probabilité* de la loi de X .

Propriétés

- $F'(x) = f(x)$, $\left(\frac{dF(x)}{dx} = f(x) \right)$,
- $\int_{-\infty}^{+\infty} f(t) dt = F(+\infty) - F(-\infty) = 1$,

Notion de variables aléatoires réelles

c) $P(a \leq X \leq b) = P(a < X < b) = \int_a^b f(t) dt,$

d) $P(x < X \leq x + dx) = f(x) dx,$

e) Soit g une fonction positive, telle que $\int_{-\infty}^{+\infty} g(t) dt = 1,$ alors la

fonction définie par $G(x) = \int_{-\infty}^x g(t) dt$ est la fonction de répartition d'une loi dont g est la densité de probabilité.

Exemple : Soit X une variable aléatoire continue dont la loi admet pour densité de probabilité

$$f(x) = \begin{cases} c(4x - 2x^2) & \text{si } 0 < x < 2 \\ 0 & \text{sinon.} \end{cases}$$

- 1) Déterminer la valeur de la constante réelle c .
- 2) Calculer $P(X > 1)$.

Solution

1) f étant une densité de probabilité de la loi de la variable aléatoire X , on doit avoir

i) $f(x) \geq 0, \forall x \in \mathbb{R}$, ce qui entraîne que $c \geq 0$,

ii) $\int_{-\infty}^{+\infty} f(t) dt = 1$, ce qui entraîne que

$$1 = c \int_0^2 (4x - 2x^2) dx = c \left[2x^2 - \frac{2x^3}{3} \right]_0^2 = \frac{8}{3} c \implies c = \frac{3}{8}.$$

$$2) P(X > 1) = \int_1^{+\infty} f(x) dx = \frac{3}{8} \int_1^2 (4x - 2x^2) dx = \frac{1}{2}.$$

3.2- Lois usuelles continues

- Loi uniforme

Définition

- Une loi de probabilité P de support $[a, b]$ (avec a et b deux nombres réels distincts) définie sur $(\mathbb{R}, \mathcal{B}_{\mathbb{R}})$ est dite **uniforme** sur cet intervalle, si elle admet une densité de probabilité de la forme

$$f(x) = \begin{cases} \frac{1}{b-a} & \text{si } x \in [a, b] \\ 0 & \text{sinon.} \end{cases} \quad (7)$$

On note $\mathcal{U}([a, b])$ la loi uniforme sur $[a, b]$.

- On dira qu'une v.a. réelle X suit une **loi uniforme sur l'intervalle $[a, b]$** , si sa loi de probabilité admet une densité de probabilité de la forme (7) et on note $X \sim \mathcal{U}([a, b])$.

- Loi normale

Définition

- Une loi de probabilité P définie sur $(\mathbb{R}, \mathcal{B}_{\mathbb{R}})$ est appelée *loi normale de paramètres m et σ* , si elle admet une densité de probabilité de la forme

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-m}{\sigma}\right)^2}, \quad \forall x \in \mathbb{R}. \quad (8)$$

On note $\mathcal{N}(m, \sigma^2)$ la loi normale de paramètres m et σ .

- On dira qu'une v.a. réelle X suit une loi normale $\mathcal{N}(m, \sigma^2)$, si sa loi de probabilité admet une densité de probabilité de la forme (8) et on note $X \sim \mathcal{N}(m, \sigma^2)$.

- Si $m = 0$ et $\sigma^2 = 1$, la loi est dite *loi normale centrée réduite* notée $X \sim \mathcal{N}(0, 1)$.

Remarques : *i)* La loi normale est également appelée loi de Laplace-Gauss où simplement loi gaussienne.

ii) Si $X \sim \mathcal{N}(m, \sigma^2)$, alors $Y = \frac{X - m}{\sigma} \sim \mathcal{N}(0, 1)$.

En effet, soit F_Y la fonction de répartition de la v.a. Y . Pour tout $y \in \mathbb{R}$, on a

$$F_Y(y) = P(Y \leq y) = P\left(\frac{X - m}{\sigma} \leq y\right) = P(X \leq \sigma y + m) = F_X(\sigma y + m),$$

où F_X est la fonction de répartition de la v.a. X . La densité de probabilité f_Y de la loi de Y est donnée par

$$f_Y(y) = \frac{d}{dy} F_Y(y) = \frac{d}{dy} F_X(\sigma y + m) = \sigma F'_X(\sigma y + m) = \sigma f_X(\sigma y + m),$$

où f_X est la densité de probabilité de la loi normale $\mathcal{N}(m, \sigma^2)$, d'où

$$f_Y(y) = \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}}, \quad \forall y \in \mathbb{R}.$$

- Loi de Cauchy

Définition

- Une loi de probabilité P définie sur $(\mathbb{R}, \mathcal{B}_{\mathbb{R}})$ est appelée *loi de Cauchy de paramètres réels a et b* (avec $a > 0$), si elle admet une densité de probabilité de la forme

$$f(x) = \frac{1}{\pi} \left[\frac{a}{a^2 + (x - b)^2} \right], \quad \forall x \in \mathbb{R}. \quad (9)$$

Remarque : La courbe de la fonction de densité d'une loi de Cauchy rappelle celle d'une loi normale, à savoir une forme de cloche, mais avec un étalement plus large. En effet, on peut le constater sur les courbes de la représentation (2) suivante pour une loi de Cauchy avec $a = 1$ et $b = 0$ et une loi normale centrée réduite $\mathcal{N}(0, 1)$.

Notion de variables aléatoires réelles

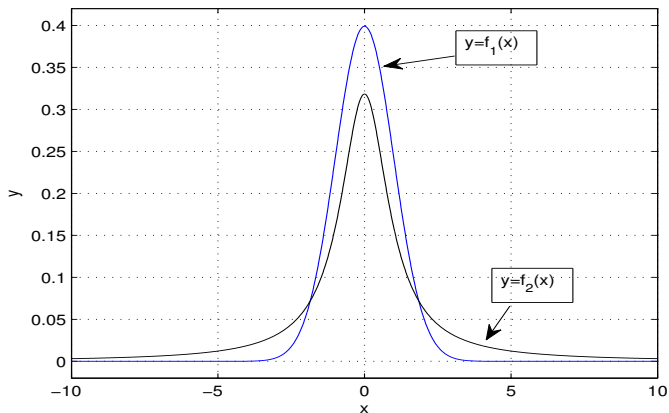


FIGURE: Courbe de la densité d'une loi normale $\mathcal{N}(0, 1)$: $f_1(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}}$, ainsi que celle de la densité d'une loi de Cauchy : $f_2(x) = \frac{1}{\pi(1+x^2)}$.

- Loi exponentielle

Définition

- Une loi de probabilité P définie sur $(\mathbb{R}, \mathcal{B}_{\mathbb{R}})$ est appelée *loi exponentielle de paramètre réel* $\lambda > 0$, si elle admet une densité de probabilité de la forme

$$f(x) = \begin{cases} \lambda e^{-\lambda x} & \text{si } x > 0 \\ 0 & \text{sinon.} \end{cases} \quad (10)$$

Remarques :

- La loi de probabilité exponentielle peut être associée aux processus de Poisson. En effet, un tel processus génère des événements dont les temps d'occurrence sont indépendants et distribués suivant une loi exponentielle.
- La loi de probabilité exponentielle de paramètre λ est également utilisée en fiabilité. Le paramètre λ représente le taux de défaillance d'un appareil et son inverse $1/\lambda$ désigne le temps moyen de bon fonctionnement.

- Loi Gamma

Définition

- Soient a et b deux nombres réels strictement positifs. Une loi de probabilité P définie sur $(\mathbb{R}, \mathcal{B}_{\mathbb{R}})$ est appelée *loi Gamma de paramètres a et b* , si elle admet une densité de probabilité de la forme

$$f(x) = \begin{cases} \frac{b^a}{\Gamma(a)} x^{a-1} e^{-bx} & \text{si } x > 0 \\ 0 & \text{sinon,} \end{cases} \quad (11)$$

où $\Gamma(\cdot)$ est la fonction Gamma d'Euler définie, pour $a > 0$, par $\Gamma(a) = \int_0^{+\infty} e^{-x} x^{a-1} dx$.

Cette loi de probabilité est identifiée par la notation $\Gamma(a, b)$.

Remarques :

i) Dans le cas particulier $a = 1$, on retrouve la loi exponentielle de paramètre $\lambda = b$.

ii) La loi $\Gamma(a, b)$ peut être utilisée pour modéliser un certain nombre de phénomènes aléatoires. Par exemple :

- Dans la théorie des files d'attente, la loi gamma représente la loi de probabilité d'occurrence de a événements (a étant un entier), dans un processus poissonnien. Si le temps T , entre les défaillances successives d'un système, suit une loi exponentielle, le temps cumulé d'apparitions de b défaillances suit une loi $\Gamma(a, b)$,
- En fiabilité, la loi $\Gamma(a, b)$ peut être utilisée pour modéliser les temps de défaillance d'un matériel.

4- Moments d'une variable aléatoire

4.1- Espérance mathématique d'une variable aléatoire discrète

Soit X une variable aléatoire réelle définie sur un espace probabilisé (Ω, \mathcal{A}, P) , telle que $X(\Omega) = \{x_1, x_2, \dots, x_n, \dots\}$ et $p_n = P(X = x_n)$, $n = 1, 2, \dots$ les probabilités correspondantes. Si on fait un grand nombre N d'observations, on aura, d'après la signification de la probabilité, environ Np_1 observations $X = x_1$, Np_2 observations $X = x_2$, etc

En faisant la moyenne arithmétique des valeurs de X fournies par les N observations, on obtiendra une valeur proche de

$$\frac{1}{N} \left[(Np_1)x_1 + (Np_2)x_2 + \dots + (Np_n)x_n + \dots \right] = \sum_{k \geq 1} p_k x_k, \quad (12)$$

avec $\sum_{i \geq 1} Np_i = N$.

Notion de variables aléatoires réelles

Nous sommes amenés ainsi à considérer la série $\sum_{k \geq 1} x_k p_k$. Si on définit $E(X)$ comme étant la moyenne arithmétique des valeurs x_k pondérées par les p_k , il faut supposer que la série du second membre de (12) soit absolument convergente.

Si X peut prendre une infinité de valeurs, $E(X)$ peut ne pas exister ; par exemple pour la variable aléatoire X telle que

$$p_n = P(X = n) = \begin{cases} \frac{1}{n(n+1)} & \text{si } n = 1, 2, \dots \\ 0 & \text{sinon.} \end{cases}$$

On vérifie que la suite (p_n) forme bien une probabilité, en effet

$$\begin{aligned} \sum_{n \geq 1} p_n &= \sum_{n \geq 1} \frac{1}{n(n+1)} = \sum_{n \geq 1} \left(\frac{1}{n} - \frac{1}{n+1} \right) \\ &= \lim_{n \rightarrow +\infty} \sum_{k=1}^n \left(\frac{1}{k} - \frac{1}{k+1} \right) = \lim_{n \rightarrow +\infty} \left(1 - \frac{1}{n+1} \right) = 1. \end{aligned}$$

Mais l'espérance correspondant n'est pas finie, car

$$\sum_{k \geq 1} kp_k = \sum_{k \geq 1} \frac{1}{k+1} = +\infty.$$

Ceci nous conduit à la définition suivante.

Définition

Soit X une variable aléatoire réelle discrète, prenant les valeurs x_i , $i = 1, 2, \dots$, si $\sum_{k \geq 1} |x_k|P(X = x_k) < +\infty$, on dit que X a une espérance finie. L'espérance de X est alors

$$E(X) = \sum_{k \geq 1} x_k P(X = x_k).$$

Sinon, on dit que X n'a pas d'espérance.

On appelle aussi espérance l'application, $E : X \longrightarrow E(X)$.

Remarques : - L'espérance mathématique d'une variable aléatoire Z est souvent utilisée pour représenter le centre de la distribution de Z .

- L'espérance est exactement analogue au concept physique de centre de gravité d'une distribution de masse. Imaginons l'ensemble \mathbb{R} comme une corde avec une masse $P(Z = z_1)$ en z_1 , $P(Z = z_2)$ en z_2, \dots . La masse unité est distribuée le long de \mathbb{R} . Le point équilibre est le centre de gravité situé en $E(Z)$.

Exemples

a) Espérance mathématique de la distribution binomiale

Si $X \sim \mathcal{B}(n, p)$, alors $E(X) = np$. En effet, on a

$$E(X) = \sum_{k=0}^n kP(X = k) = \sum_{k=0}^n kC_n^k p^k (1-p)^{n-k} = \sum_{k=1}^n kC_n^k p^k (1-p)^{n-k}.$$

Or, $kC_n^k = nC_{n-1}^{k-1}$, donc

$$\begin{aligned} E(X) &= \sum_{k=1}^n nC_{n-1}^{k-1} p^k (1-p)^{n-k} = np \sum_{k=1}^n C_{n-1}^{k-1} p^{k-1} (1-p)^{n-k} \\ &= np \sum_{j=0}^{n-1} C_{n-1}^j p^j (1-p)^{n-1-j} = np [p + (1-p)]^{n-1} = np, \end{aligned}$$

où l'on a posé $j = k - 1$.

b) Espérance mathématique de la distribution hypergéométrique

Soit X une variable aléatoire de distribution hypergéométrique de paramètres $N, n, \frac{M}{N}$, alors $E(X) = \frac{Mn}{N}$.

En effet, supposons que $n \leq M$ et $n \leq N - M$, on a

$$P(X = k) = \frac{C_M^k C_{N-M}^{n-k}}{C_N^n}, \quad k = 0, 1, \dots, n \text{ et}$$

$$E(X) = \sum_{k=0}^n k P(X = k) = \sum_{k=0}^n k \frac{C_M^k C_{N-M}^{n-k}}{C_N^n} = \sum_{k=1}^n k \frac{C_M^k C_{N-M}^{n-k}}{C_N^n}$$

Or $k C_M^k = M C_{M-1}^{k-1}$ et $n C_N^n = N C_{N-1}^{n-1}$, donc

$$E(X) = \frac{Mn}{N} \sum_{k=1}^n \frac{C_{M-1}^{k-1} C_{N-M}^{n-k}}{C_{N-1}^{n-1}} = \frac{Mn}{N} \sum_{j=0}^{n-1} \frac{C_{M-1}^j C_{N-M}^{n-1-j}}{C_{N-1}^{n-1}} = \frac{Mn}{N},$$

où l'on a posé $j = k - 1$.

c) Espérance mathématique de la distribution de Poisson $\mathcal{P}(\lambda)$

Soit X une variable aléatoire de loi de Poisson de paramètre λ , alors $E(X) = \lambda$.

En effet, on a

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

où l'on a posé $j = k - 1$.

4.2- Cas où la loi de la variable aléatoire X admet une densité de probabilité

Définition

Soit X une variable aléatoire dont la loi admet pour densité de probabilité f , si $\int_{-\infty}^{+\infty} |x|f(x) dx < +\infty$, l'espérance de X est

$$E(X) = \int_{-\infty}^{+\infty} xf(x) dx .$$

Exemples

a) Distribution exponentielle

Soit X une variable aléatoire de loi exponentielle de paramètre $\lambda > 0$, sa densité de probabilité s'écrit

$$f(x) = \begin{cases} \lambda e^{-\lambda x} & \text{si } x > 0 \\ 0 & \text{sinon.} \end{cases}$$

Il est facile de vérifier que f est bien une densité de probabilité ((i) $f(x) \geq 0$, $\forall x \in \mathbb{R}$ et (ii) $\int_{-\infty}^{+\infty} f(x) dx = 1$).

Par ailleurs, on a

$$E(X) = \int_{-\infty}^{+\infty} xf(x) dx = \int_0^{+\infty} x\lambda e^{-\lambda x} dx = \frac{1}{\lambda} \quad (\text{intégration par parties}).$$

b) Distribution uniforme sur $[a, b]$

Soit X une variable aléatoire de loi uniforme sur $[a, b]$, sa densité de probabilité a pour expression

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

On a

$$E(X) = \int_{-\infty}^{+\infty} xf(x) dx = \int_0^b x \frac{1}{b-a} dx = \frac{1}{b-a} \left(\frac{x^2}{2} \Big|_a^b \right) = \frac{a+b}{2}.$$

c) Distribution de Cauchy

Soit X une variable aléatoire de loi de Cauchy dont la loi a pour densité de probabilité donnée par

$$f(x) = \frac{1}{\pi(1+x^2)}, \quad \forall x \in \mathbb{R}.$$

La variable aléatoire X n'a pas d'espérance car

$$\begin{aligned} \int_{-\infty}^{+\infty} \frac{|x|}{\pi(1+x^2)} dx &= 2 \int_0^{+\infty} \frac{x}{\pi(1+x^2)} dx \\ &= \frac{2}{\pi} \left(\frac{1}{2} \text{Log}(1+x^2) \Big|_0^{+\infty} \right) = +\infty. \end{aligned}$$

4.3- Espérance d'une fonction d'une variable aléatoire

Soient X une variable aléatoire réelle définie sur l'espace probabilisé (Ω, \mathcal{A}, P) et $Z = g(X)$ une variable aléatoire fonction de la variable aléatoire X . Pour calculer $E(Z)$ on peut d'abord déterminer sa loi à partir de celle de X et ensuite utiliser la définition de l'espérance mathématique. Toutefois, il est possible de montrer que l'on peut directement calculer $E(Z)$ sur la loi de X .

Définition

Soient (E, \mathcal{E}) et (F, \mathcal{F}) deux espaces mesurables, on dira que la fonction φ définie de (E, \mathcal{E}) et (F, \mathcal{F}) est **mesurable** si pour tout $B \in \mathcal{F}$, $\varphi^{-1}(B)$ appartient à \mathcal{E} .

Théorème

Soit X une variable aléatoire réelle discrète prenant les valeurs x_1, x_2, \dots et φ une application mesurable de $(\mathbb{R}, \mathcal{B}_{\mathbb{R}})$ dans $(\mathbb{R}, \mathcal{B}_{\mathbb{R}})$, alors la variable aléatoire $Z = \varphi(X)$ a une espérance finie si et seulement si $\sum_{i \geq 1} |\varphi(x_i)| P(X = x_i) < +\infty$ et alors

$$E(Z) = E(\varphi(X)) = \sum_{i \geq 1} \varphi(x_i) P(X = x_i).$$

Démonstration : Soient X une variable aléatoire réelle discrète définie sur (Ω, \mathcal{A}, P) et φ une application mesurable de $(\mathbb{R}, \mathcal{B}_{\mathbb{R}})$ dans $(\mathbb{R}, \mathcal{B}_{\mathbb{R}})$. On note z_1, z_2, \dots les valeurs possibles de la variable aléatoire $Z = \varphi(X)$, c'est-à-dire $Z(\Omega) = \{z_1, z_2, \dots\}$.

Notion de variables aléatoires réelles

Pour tout $z_j \in Z(\Omega)$, il existe au moins $x_i \in \mathbb{R}$ tel que $\varphi(x_i) = z_j$, mais il peut y avoir plusieurs. Soient $A_j = \{x_i \in \mathbb{R} / \varphi(x_i) = z_j\}$ donc

$$\{\omega \in \Omega / Z(\omega) = z_j\} = \{\omega \in \Omega / X(\omega) \in A_j\}$$

et

$$P(Z = z_j) = P(X \in A_j) = P\left(\bigcup_{x \in A_j} \{X = x\}\right) = \sum_{x \in A_j} P(X = x).$$

Pour qu'on puisse parler de l'espérance de Z , il faut que $\sum_{j \geq 1} |z_j| P(Z = z_j)$ soit finie. Or

$$\begin{aligned} \sum_{j \geq 1} |z_j| P(Z = z_j) &= \sum_{j \geq 1} |z_j| \sum_{\{i \geq 1 / x_i \in A_j\}} P(X = x_i) \\ &= \sum_{j \geq 1} \sum_{\{i \geq 1 / x_i \in A_j\}} |z_j| P(X = x_i) \end{aligned}$$

et pour $x_i \in A_j$, on a $\varphi(x_i) = z_j$, donc

$$\begin{aligned}\sum_{j \geq 1} |z_j| P(Z = z_j) &= \sum_{j \geq 1} \sum_{\{i \geq 1 / x_i \in A_j\}} |\varphi(x_i)| P(X = x_i) \\ &= \sum_{i \geq 1} |\varphi(x_i)| P(X = x_i).\end{aligned}$$

En faisant un raisonnement analogue sans valeur absolue, on montre que

$$E(Z) = \sum_{j \geq 1} z_j P(Z = z_j) = \sum_{i \geq 1} \varphi(x_i) P(X = x_i).$$

Théorème

Si X est une variable aléatoire continue telle que sa loi admet pour densité de probabilité f et φ une application continue de \mathbb{R} dans \mathbb{R} , alors si

$$\int_{-\infty}^{\infty} |\varphi(x)| f(x) dx < \infty, \text{ on a : } E(\varphi(X)) = \int_{-\infty}^{\infty} \varphi(x) f(x) dx.$$

Notion de variables aléatoires réelles

Exemple : Considérons la variable aléatoire X de loi uniforme sur l'intervalle $[0, 1]$. Sa fonction de répartition est

$$F_X(x) = \begin{cases} 0 & \text{si } x < 0 \\ x & \text{si } 0 \leq x \leq 1 \\ 1 & \text{si } x > 1 \end{cases}$$

et sa densité

$$f_X(x) = \begin{cases} 1 & \text{si } 0 \leq x \leq 1 \\ 0 & \text{ailleurs.} \end{cases}$$

Considérons la variable aléatoire $Z = X^2$. Calculons d'abord $E(Z)$ en établissant la loi de Z . Soit F_Z la fonction de répartition de Z , on a

$$F_Z(z) = P(X^2 \leq z) = \begin{cases} P(-\sqrt{z} \leq X \leq \sqrt{z}) = F_X(\sqrt{z}) & \text{si } z \geq 0 \\ 0 & \text{sinon,} \end{cases}$$

puisque $F_X(x) = 0, \forall x < 0$.

Notion de variables aléatoires réelles

Donc pour $\forall z \geq 0$, on a

$$F_Z(z) = \begin{cases} 0 & \text{si } z < 0 \\ \sqrt{z} & \text{si } 0 \leq \sqrt{z} \leq 1 & \iff 0 \leq z \leq 1 \\ 1 & \text{si } \sqrt{z} > 1 & \iff z > 1. \end{cases}$$

Ainsi la densité de probabilité de la loi de $Z = X^2$ est donnée par

$$f_Z(z) = \frac{dF_Z(z)}{dz} = \begin{cases} \frac{1}{2\sqrt{z}} & \text{si } 0 \leq z \leq 1 \\ 0 & \text{ailleurs.} \end{cases}$$

D'où,

$$E(Z) = \int_{-\infty}^{\infty} z f_Z(z) dz = \int_0^1 z \frac{1}{2\sqrt{z}} dz = \int_0^1 \frac{\sqrt{z}}{2} dz = \frac{1}{3}.$$

Calculons maintenant directement

$$E(Z) = E(X^2) = \int_{-\infty}^{\infty} x^2 f_X(x) dx = \int_0^1 x^2 dx = \frac{1}{3}.$$

4.3- Linéarité de l'opérateur $E(\cdot)$

Pour toute combinaison linéaire $\alpha g(X) + \beta h(X)$ de fonctions g et h de X , on a :

$$E\left(\alpha g(X) + \beta h(X)\right) = \alpha E\left(g(X)\right) + \beta E\left(h(X)\right).$$

Ceci découle immédiatement des théorèmes 1, 2 et de la linéarité de la sommation ou de l'intégration.

Voyons cela sur le cas particulier $\alpha X + \beta$, pour le cas d'une variable aléatoire réelle X discrète prenant les valeurs x_1, x_2, \dots :

$$\begin{aligned} E(\alpha X + \beta) &= \sum_{i \geq 1} (\alpha x_i + \beta) P(X = x_i) \\ &= \alpha \sum_{i \geq 1} x_i P(X = x_i) + \beta \sum_{i \geq 1} P(X = x_i) \\ &= \alpha E(X) + \beta. \end{aligned}$$

- Remarques :** - Notons que l'on peut voir β comme une variable aléatoire certaine, c'est-à-dire prenant cette seule valeur avec probabilité égale à 1 et en cohérence avec la définition de l'espérance mathématique, écrire par convention $E(\beta) = \beta$.
- On notera également que dans le cas général $E(g(X))$ n'est pas égal à $g(E(X))$, par exemple $E(X^2) \neq [E(X)]^2$.

4.5- Moments et moments centrés

Définition

Soit X une variable aléatoire réelle et r un entier positif ou nul. On dira que X a **un moment d'ordre r** si X^r a une espérance finie. Dans ce cas, on définit le moment d'ordre r par $E(X^r)$.

Remarques : - Si X est une variable aléatoire discrète prenant les valeurs x_1, x_2, \dots , d'après le Théorème 1,

$$E(X^r) = \sum_{i \geq 1} x_i^r P(X = x_i).$$

- Si X est une variable aléatoire continue dont la loi admet pour densité de probabilité f , d'après le Théorème 2,

$$E(X^r) = \int_{-\infty}^{\infty} x^r f(x) dx.$$

Définition

On appelle *moment centré d'ordre r* de la variable aléatoire X , où r est un entier positif ou nul, la valeur (si elle existe)

$$\mu_r = E\left((X - \mu)^r\right), \quad \text{où } \mu = E(X).$$

Remarques :

- *i*) Pour $r = 1$ on a $E(X - \mu) = E(X) - \mu = \mu - \mu = 0$ ce qui caractérise le centrage de X .
- *ii*) Si la variable aléatoire X a un moment d'ordre r , il a un moment d'ordre k pour $k \leq r$.

Notion de variables aléatoires réelles

En effet, on a

$$|X|^{r-1} \leq |X|^r + 1,$$

car

$$\begin{array}{ll} \text{si } |X| \geq 1, & |X|^{r-1} \leq |X|^r \implies |X|^{r-1} \leq |X|^r + 1, \\ \text{si } |X| \leq 1, & |X|^{r-1} \leq 1 \implies |X|^{r-1} \leq 1 + |X|^r. \end{array}$$

Ce qui entraîne que si X est une variable aléatoire discrète prenant les valeurs x_1, x_1, \dots , on a

$$\sum_{i \geq 1} |x_i|^{r-1} P(X = x) \leq \sum_{i \geq 1} |x_i|^r P(X = x) + \sum_{i \geq 1} P(X = x),$$

comme $\sum_{i \geq 1} |x_i|^r P(X = x) < \infty \implies \sum_{i \geq 1} |x_i|^{r-1} P(X = x) < \infty$.

4.5- Variance

Soit X une variable aléatoire réelle, son **moment centré d'ordre r** s'écrit

$$\mu_r = E\left(\left[X - E(X)\right]^r\right).$$

Pour $r = 2$, on a ce qu'on appelle la **variance de X** , qui est une caractéristique de mesure de dispersion de la distribution de X comme en statistique descriptive et qui mérite ainsi une attention particulière.

Définition

Soit X une variable aléatoire ayant un moment d'ordre deux fini. On définit **la variance de X** notée $\text{Var}(X)$ par

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

Notion de variables aléatoires réelles

En développant l'expression de la variance et en utilisant la linéarité de l'espérance, on a également une autre écriture de la variance utilisée souvent dans les calculs :

$$\begin{aligned} \text{Var}(X) &= E\left\{X^2 - 2XE(X) + [E(X)]^2\right\} \\ &= E(X^2) - 2E(X)E(X) + [E(X)]^2, \end{aligned}$$

d'où

$$\text{Var}(X) = E(X^2) - [E(X)]^2.$$

Définition

On définit l'écart-type de la variable aléatoire X noté σ_X par

$$\sigma_X = \sqrt{\text{Var}(X)}.$$

Remarque : La variance mesure l'étendue de la dispersion de la distribution de la variable aléatoire X par rapport à $\mu = E(X)$. C'est l'analogie du moment d'inertie en mécanique. Plus X dévie de sa valeur moyenne μ , plus $(X - \mu)^2$ augmente et donc plus la variance de X est grande.

Propriétés :

- *i*) $Var(a) = 0$, où a est une constante réelle.

En effet, on a $Var(a) = E\left(\left[a - E(a)\right]^2\right) = E(0) = 0$, car $E(a) = a$.

- *ii*) $Var(aX + b) = a^2 Var(X)$, où a et b sont des constantes réelles.

En effet, on a

$$\begin{aligned} Var(aX + b) &= E\left\{\left[(aX + b) - E(aX + b)\right]^2\right\} \\ &= E\left\{\left[aX - aE(X)\right]^2\right\} = a^2 E\left\{\left[X - E(X)\right]^2\right\}, \\ &= a^2 Var(X). \end{aligned}$$

Exemples :

- i) Soit X une variable aléatoire de loi de Poisson de paramètre λ , ($X \sim \mathcal{P}(\lambda)$). On a $E(X) = \lambda$ et

$$\text{Var}(X) = E(X^2) - [E(X)]^2 = E(X^2) - \lambda^2.$$

Par ailleurs

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

où l'on a remplacé k^2 par $k(k-1) + k$

$$E(X^2) = \sum_{k \geq 2} \frac{e^{-\lambda} \lambda^k}{(k-2)!} + \lambda = \lambda^2 e^{-\lambda} \sum_{k \geq 0} \frac{\lambda^j}{j!} + \lambda = \lambda^2 + \lambda,$$

où on a posé $j = k - 2$, d'où $\text{Var}(X) = \lambda$.

- ii) Soit X une variable aléatoire continue de **loi exponentielle de paramètre $\alpha > 0$** , sa densité de probabilité a pour expression

$$f(x) = \begin{cases} \alpha e^{-\alpha x} & \text{si } x > 0, \\ 0 & \text{sinon.} \end{cases}$$

On a

$$E(X^2) = \int_{-\infty}^{\infty} x^2 f(x) dx = \int_0^{\infty} x^2 \alpha e^{-\alpha x} dx$$

En intégrant deux fois par parties, on obtient $E(X^2) = \frac{2}{\alpha^2}$.

Par ailleurs, $E(X) = \frac{1}{\alpha}$, donc $\text{Var}(X) = \frac{2}{\alpha^2} - \frac{1}{\alpha^2} = \frac{1}{\alpha^2}$.