Chapitre III. VARIABLES ALEATOIRES

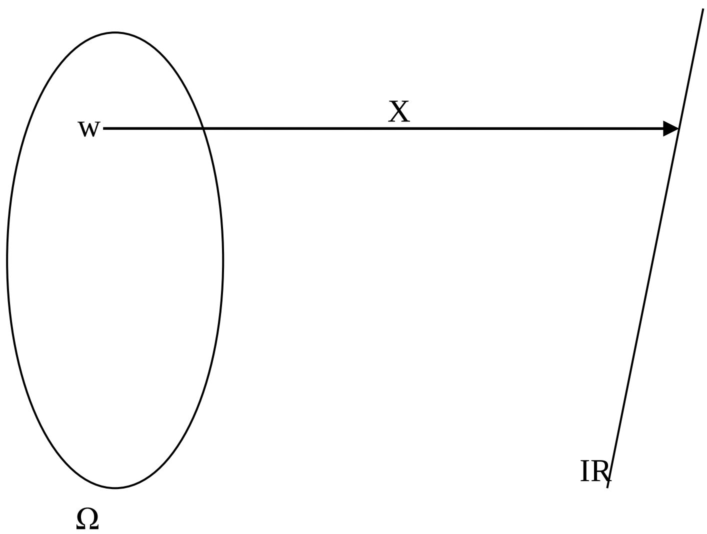
1. *Généralités sur les variables aléatoires*

Dans la plupart des phénomènes aléatoires, le résultat d'une épreuve peut se traduire par une « grandeur » mathématique, très souvent représentée par un nombre entier ou un nombre réel, par exemple :
La proportion de pièces défectueuses fabriquées par une machine ;
Le temps désintégration d'un atome radioactif ;
La notion mathématique qui représente efficacement ce genre de concrète est celle de variable aléatoire. On se limitera ici au cas des variables aléatoires réelles.

Définition

Etant donné un espace probabilisé d'espace fondamentale Ω et de mesure de probabilité P, on appelle variable aléatoire (réelle) sur cet espace toute application X de Ω dans R :

$X : w \in \Omega \mapsto X(w) \in R$



Exemple

Une urne contient 4 boules, 6 boules rouges on extrait de cette urne 4 boules, Soit Y un variables aléatoire qui désigne le nombre de boules blanches parmi ces quatre 4 boules extraites. Y va prendre les valeurs 0, 1, 2, 3,4.

Y	0	1	2	3	4

P(Y=k)	15 /210	80 /210	90 /210	24 /210	1 /210

$$P(Y = k) = \frac{C_4^k \times C_6^{4-k}}{C_{10}^4}$$

Notation

Soit un espace probabilisé d'espace fondamental Ω et de mesure de probabilité P, et X une variable aléatoire. Si B est une partie de R, et x un nombre de R, on définit :

$$P(X \in B) = P(X^{-1}(B))$$
$$P(X = x) = P(X^{-1}(\{x\}))$$

Définition

Etant donnée une variable aléatoire X dont an a précisé l'ensemble des valeurs prises, on appelle ***loi*** de cette variable aléatoire **une expression des probabilités** relatives aux **valeurs**.

Le mot ***distribution*** (d'une variable aléatoire) est très souvent employé comme synonyme de loi.

3. Variables aléatoires discrètes

3.1 Loi de probabilité discrète

Variable aléatoire discrète une variable qui ne prend qu'un nombre fini ou dénombrable (x_1, x_2, x_3, …x_i,….. ;) de valeurs.

La loi d'une variable aléatoire discrète peut être donnée :

- Soit par la liste de probabilités ;
- Soit par une formule générale permettant de calculer les probabilités ponctuelles. Elle doit satisfaire la condition que la probabilité totale soit égale à 1 :
  - $\sum_{i=1}^{n} x_i \times p_i = 1$

2.2 Diagramme en bâtons, fonction de répartition

Il est souvent souhaitable de pouvoir visualiser **une distribution de probabilités**.

La fonction de répartition F d'une variable aléatoire X est définie pour tout réel x $-\infty \le x \le +\infty$, $F(x)$ est la probabilité que X prend une valeur plus petite ou égale à x :

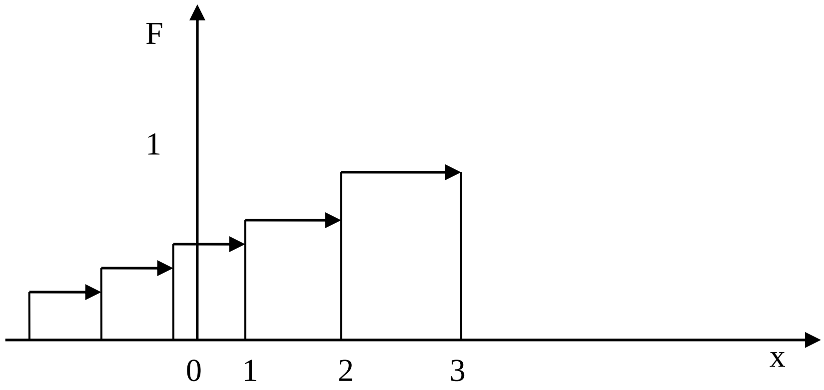
$$F(x) = P(X \le x) = \sum_{x_i \le x} P_j$$

Cette représentation graphique, concentrée sur les valeurs prises, est trop « ponctuelle », et il est très important de pouvoir associer à la variable aléatoire X une fonction qui soit définie pour tout x réel. Cette fonction est la fonction de répartition de X, et elle est définie par :

$$F(x) = P(X \leq x) = \sum_{x_i \leq x} P_j$$

L'un des intérêts de cette fonction est que l'on peut la définir pour n'importe quelle variable aléatoire, discrète, continue ou mixte.

Dans le cas d'une variable aléatoire discrète, c'est une fonction « en escalier »



Propriétés

1. F est une fonction croissante
2. F est constante sur tout l'intervalle $[x_k, x_{k+1}[$
3. F est continue à droite en tout point x de R
4. $\lim_{x \to -\infty}(F(x)) = 0$ et $\lim_{x \to +\infty}(F(x)) = 1$
5. $F(x) = P(X \leq x) = \sum_{x_i \leq x} P_j$
6. $P(a \leq X \leq b) = F(b) - F(a)$

2.3 Espérance mathématique

Définition

Soit X une variable aléatoire discrète dont l'ensemble des valeurs prises est xi, i=1,2, 3,…….n, et dont la loi de probabilité est $P(X = x_i) = p_i$. L'espérance mathématique de X est le nombre réel

$$E(X) = \sum_{i=1}^{n} x_i P(X = x_i) = \sum_{i=1}^{n} x_i p_i$$

Moment d'ordre k de X

$$m_k = E\left[\left(X^k\right)\right] = \sum_{i=1}^{n} p_i \left(x_i\right)^k$$

Moment centré d'ordre k de X

$$\mu_k = E\left[\left(X - m_1\right)^2\right] = \sum_{i=1}^{n} p_i \left(x_i - m_1\right)^k$$

2.3 __Variance, écart-type__

__Variance__

Soit X une variable aléatoire discrète dont l'ensemble des valeurs prises est xi, i=1,2,……n, et dont la loi de probabilité est $P(X = x_i) = p_i$. La variance de X est le nombre réel positif.

$$Var(X) = E\left(\left(X - E(X)^2\right)\right) = \sum_{i=1}^{n} \left(x_i - E(X)\right)^2 p_i$$

C'est le moment centré d'ordre 2

__Ecart type__

$$\sigma^2 = Var(X) = E\left(X^2\right) - \left[E(X)\right]^2$$

__Variable aléatoire absolument continue__

__Loi de probabilité__
__Définition__

Soit un espace probabilisé (Ω, ε, P), et F la fonction de répartition de X définie (Ω, ε, P). X est dite absolument continue si, de plus, il existe une fonction $f(x)$ définie sur R et possédant les propriétés :

1) $f(x) > 0$ pour tout $x \in R$
2) $f(x)$ est continue sur R sauf, peut être, en un nombre fini de points
3) $\int_{-\infty}^{+\infty} f(x)dx$ est converge et $\int_{-\infty}^{+\infty} f(x)dx = 1$
4) $F(x) = \int_{-\infty}^{x} f(t)dt$

f est dite densité de la variable aléatoire X ;

Pour tous les réel a et b tels que a<b, on a : $P(a < X \leq b) = \int_{a}^{b} f(t)dt = F(b) - F(a)$

$$P(X = a) = 0$$

F est continue sur R et si f est continue en x en R, F est dérivable et $F^{'}(x)=f(x)$, la densité $f(x)$ de x définit la loi de probabilité de x.

Caractéristiques de la variable aléatoire

1) Espérance mathématique de X

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

2) Moment d'ordre k de X

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

Si $E(X)=m_1$ existe, on appelle moment centré d'ordre k l'intégrale :

$$\mu_k = E\left[(X-m_1)^2\right]=\int_{-\infty}^{+\infty}(x-m_1)^k f(x)dx$$

3) Variance

$$\sigma^2 = V(x)=E\left[(X-m_1)^2\right]=\int_{-\infty}^{+\infty}(x-m_1)^2 f(x)dx = m_2 - m_1$$

Exemple 1

Une urne contient 5 billes rouges et 6 billes noires. On tire, sans remise, 3 billes de cette urne. Etablir la suite de répartition de la variable aléatoire X correspondant au nombre de boules noires extraites.

1. Déterminer la fonction de répartition de X.
2. Calculer l'espérance mathématique $E(X)$ et l'écart type $\sigma(X)$
3. Calculer $P(X=2)$ et $P(1\leq X\leq 3)$

Solution

En notant par *Ni* (i=1, 2, 3) l'événement :
Ni : « la boule extraite au ieme tirage est noire » et par conséquent
$x_1=0$, $x_2=1$, $x_3=2$, $x_4=3$ la différentes valeurs que peut prendre la variable aléatoire X, on a :

$$P(X=0)=P\left(\overline{N}_1\overline{N}_2\overline{N}_3\right)=\frac{5}{11}\cdot\frac{4}{10}\cdot\frac{3}{9}=0.06$$
$$P(X=1)=P\left(N_1\overline{N}_2\overline{N}_3+\overline{N}_1N_2\overline{N}_3+\overline{N}_1\overline{N}_2N_3\right)=0.36$$
$$P(X=2)=P\left(N_1N_2\overline{N}_3+N_1\overline{N}_2N_3+\overline{N}_1N_2N_3\right)=0.46$$
$$P(X=3)=P\left(N_1N_2N_3\right)=0.12$$

x_i	0	1	2	3
p_i	0.06	0.36	0.46	0.12

$$F(x)=\begin{cases} 0 \quad si \quad x<0 \\ 0.06\, si\ 0\le x<1 \\ 0.42\, si\ 1\le x<2 \\ 0.88\, si\ 2\le x<3 \\ 1 \quad si \quad 3\le x \end{cases}$$

$$E(X)=\sum_{i=1}^{4} p_i \cdot x_i = 1.64$$

$$V(X)=E(X^2)-E(X)^2=0.59 \text{ et } \sigma(X)=0.77$$

La probabilité d'extraire 2 boules noires parmi les 3 est : $P(X=2)=0.46$
La probabilité de tirer au moins une boule noire est :
$P(1\le X\le 3)=P(X=1)+P(X=2)+P(X=3)=0.94$

Exemple 2

La durée de fonctionnement d'un ordinateur avant sa première panne est une variable aléatoire continue de densité donnée par :

$$f(x)=\begin{cases} \lambda e^{-x/100} & x\ge 0 \\ 0 & x<0 \end{cases}$$

a) quelle est la probabilité que cette durée de fonctionnement soit comprise entre 50 et 150 heures ?
b) quelle est la probabilité que l'ordinateur fonctionne moins de 100 heures ?

Solution

a) comme $1=\int_{-\infty}^{+\infty} f(x)dx=\lambda\int_{0}^{+\infty}\lambda e^{-x/100}dx$

On obtient $1=\lambda(100)\lambda e^{-x/100}\Big|_0^{\infty}=100\lambda$ ou $\lambda=\dfrac{1}{100}$

Ainsi la probabilité que la durée de fonctionnement de l'ordinateur soit comprise entre 50 et 150 heures est donnée par :

$$P(50<X<150)=\int_{50}^{150}\frac{1}{100}e^{-x/100}dx=-e^{-x/100}\Big|_{50}^{150}=e^{-1/2}-e^{-3/2}\approx 0.383$$

b) de la même manière

$$P(X<100)=\int_{0}^{100}\frac{1}{100}e^{-x/100}dx=-e^{-x/100}\Big|_{50}^{100}=1-e^{-1}\approx 0.632$$

En d'autres termes, l'ordinateur tombera en panne avant sa 100-iéme heure de service 63.3 fois sur 100 en moyenne.

Exemple

Trouver E[X] lorsque la densité de X est :

$$f(x) = \begin{cases} 2x & si\, 0 \le x \le 1 \\ 0 & \sin on \end{cases}$$

Solution

$$E(x) = \int xf(x)dx = \int_0^1 2x^2 dx = \frac{2}{3}$$

Exemple

La loi de probabilité d'une variable aléatoire X est donnée par $p(i) = \frac{c\lambda^i}{i!}$, i=0,1,3 ,……. ;, ou λ est un réel positif. Trouver

a) $P(X = 0)$;

b) $P(X > 2)$.

Solution

Puisque $\sum_{i=0}^{+\infty} p(i) = 1$, nous avons que $c\sum_{i=0}^{+\infty} \frac{c\lambda^i}{i!} = 1$ ce qui implique, puisque $e^x = \sum_{i=0}^{+\infty} \frac{x^i}{i!}$, que $ce^{\lambda} = 1$ ou $c = e^{-\lambda}$

Donc

a) $P(X = 0) = e^{-\lambda} \frac{\lambda^0}{0!} = e^{-\lambda}$

b) $P(X > 0) = 1 - P(X \le 2) = 1 - [P(X = 0) + P(X = 1) + P(X = 3)] = 1 - e^{-\lambda} - \lambda e^{-\lambda} - \lambda^2 \frac{e^{-\lambda}}{2!}$

6- Lois de probabilités

Nous allons nous intéresser à des familles de fonctions réelles (f) très spéciales.

$$\begin{aligned} f : E \subset R &\to F \subset R \\ x &\to f(x) \quad \text{vérifiant} \end{aligned}$$

i. $\forall x \in E,\ f(x) \ge 0$

ii. $\int_a^b f(x).dx = 1$

E étant un intervalle ouvert : $[a, b]$, semi - ouvert [a, b[, ou fermé $[a, b]$ ou un nombre

Une loi de probabilité est un **MODELE** représentant "au mieux", une distribution de fréquences d'une variable statistique ou aléatoire.

Une classification des lois

a. Les lois "discrètes"

b. Les lois "continues"

Comme l'étude de ces lois n'est pas simple, n'étudierons-nous que quelques unes d'entre elles.

• Pour chacune d'elles il faut connaître :

a. L'espérance mathématique E(X) (ou moyenne μ)et la variance Var(X) (notée σ)
b. La forme de la fonction de répartition
c. Les domaines et conditions d'utilisation

Dans le cas discret il faut pouvoir calculer les nombres :

Prob(X=k)

Dans les cas continus les nombres :

$$\text{Prob}(a < X < b) = \int_a^b f(x)dx = [F(x)]_a^b = F(b) - F(a)$$

La fonction de densité sera notée f(x), la fonction de répartition F(x)

Exemples de lois "discrètes"

1. Loi "Uniforme discrète"
2. Loi de "Bernoulli"
3. Loi "Binomiale"
4. Loi de "Poisson"
5. Loi "Hypergéométrique"
6. Loi "Binomiale négative"
7. Loi "Géométrique"
8. Loi "Multinomiale"

I- Loi UNIFORME

X est une variable aléatoire qui prend les valeurs x_1, x_2, …, x_n avec les probabilités pi=1/n

II- Loi de BERNOULLI

Dans un tel schéma, les épreuves satisfont aux conditions suivantes :

- l'issue de chaque épreuve est l'une ou l'autre de deux éventualités complémentaires
- la probabilité de succès (p) est la même pour chaque épreuve
- l'issue d'une épreuve est indépendante du résultat de chacune des épreuves précédentes

III - Loi BINOMIALE

Le nombre de succès (k) au cours d'une série de (n)épreuves répondant à un schéma de Bernoulli est une variable aléatoire discrète, appelée loi Binomiale.

Conditions :

Échantillon de taille n , 2 issues à chaque tirage de probabilité p et 1-p, Indépendance des tirages.

$$\text{Prob}(X = k) = C_n^k . p^k .(1 - p)^{n-k}$$

Remarques :

$$\boxed{La\ moyenne\ \text{est}\ m = np\ et\ l'\acute{e}cart - type\ \sigma = \sqrt{\text{np}}}$$

Une relation pratique :

$$\frac{\Pr ob(X=k+1)}{\Pr ob(X=k)}=\frac{n-k}{k+1}\cdot\frac{p}{1-p}$$

<u>Exercice</u>

La probabilité pour qu'un tireur atteigne une cible est 1/3.

• Sachant qu'il tire 5 fois, quelle est la probabilité pour qu'il atteigne la cible au moins deux fois ?

• Combien de fois doit-il tirer pour que la probabilité d'atteindre au moins une fois la cible soit plus grande que 0,9 ?

<u>Réponses</u> : - P(X≥2)= 131/243 - n=6

IV. Loi de POISSON

La Loi de densité ou de répartition est donnée par la fonction :

$$P(X=k)=\frac{m^k}{k!}e^{-m}$$

La loi de POISSON est une bonne approximation de la loi binomiale lorsque p est « très petit » (de l'ordre de 1/100 par exemple)ou p « voisin de 1 » et n « grand » .((n≥20 et p<1/30)

P(X=k) donne la fréquence des valeurs égale à k pour une loi normale réduite.

m = np

n nombre d'épreuves

La moyenne est m, la variance est m.

• Si n « grand », et p non voisin de 0 ou 1 alors

$$\Pr ob(X=k)=\frac{1}{\sigma\sqrt{2\pi}}e^{-\frac{(k-m)^2}{2\sigma^2}}dx \text{ avec m = np et } \sigma=\sqrt{np(p-1)}$$

Exemples de lois "continues"

a. La loi "Normale"
b. La loi du "Chi-deux"
c. La loi de "Student"
d. La loi de "Fisher"
e. La loi de "Laplace"
f. La loi "Beta"
g. La Loi de "Cauchy"
h. La loi "Exponentielle"
i. La loi "Gamma"
j. La loi "Lognormale"

I- Loi de LAPLACE-GAUSS

Loi normale de densité ou de répartition est donnée par la fonction :

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

f(z) donne l'effectif d'une variable quantitative continue qui suit la loi normale réduite N(0,1)

F(z) donne la fréquence des valeurs inférieures ou égales à z pour une loi normale réduite.

$$F(z) = \frac{1}{\sqrt{2\pi}} \int_{-\infty}^{z} e^{-\frac{1}{2}x^2} dx$$

$$x \to z = \frac{x-m}{\sigma} \quad \text{et} \quad z \to \sigma x + m$$

Une série $(x_i)_{0 \le i \le N}$:$(x_1,\ x_2,\ \dots\ ,x_N)$ se transforme en une série $(z_i)_{0 \le i \le N}$: $(z_1,\ z_2,\ \dots\ ,z_N$
$)$
Ce qui permettra d'utiliser les tables de la loi normale réduite

Théorème :

Si

La série $(x_i)_{0 \le i \le N}$ a pour moyenne **m** et pour écart-type σ alors on démontre que

La série $(z_i)_{0 \le i \le N}$ a pour moyenne **0** et pour écart-type **1**

Théorème :

La loi de LAPLACE-GAUSS $N(np, \sqrt{np(1-p)})$ est une approximation satisfaisante de la loi BINOMIALE $B(n,p)$ lorsque les conditions suivantes sont satisfaites : n≥30 et 0,2<p<0,8