PROBABILITES

Généralités

Les probabilités sont conçues pour être une mesure de la vraisemblance des événements.

Définitions générales

On considère Ω l'ensemble des événements élémentaires qui constituent le phénomène que l'on veut étudier
par exemple , pour un jeu de dés ordinaire , à 6 faces
Ω sera {1,2,3,4,5,6}

les parties A de Ω représentent les événements que l'on peut être amené à étudier
A = {2,4,6} dans l'exemple précédent représente l'événement : ''la valeur obtenue par le lancé du dé est un nombre pair ''

F sera la famille des parties A que l'on veut pouvoir considérer .

$\Omega \in F$
$A \in F$ et $B \in F$ implique :
(complémentaire de B dans A) $\in F$

$A_n \in F \ \forall n$ implique :
$(\cup A_n) \in F$

Une famille F qui possède ces propriétés est appelée : une tribu .

Une loi de probabilité est alors
une application de *F* dans [0 , 1]
A -> Prob(A) qui vérifie les 2 propriétés

Prob(Ω) = 1
et
$(A_n)_n$ une partition de A alors
Prob(A) = $\sum_n$ Prob(A_n)

Remarque: il s'ensuit que Prob(complémentaire de A) =
1 – Prob(A)

Exemple : loi uniforme sur un ensemble fini

Ici Ω est un ensemble fini
F est la tribu formée par
toutes les parties de Ω

Prob({x}) = 1 / card(Ω)
pour toute partie {x} de Ω réduite à un point .

Prob(A) = card(A) / card(Ω) qui se traduit souvent par la célèbre formule :

Prob(A) = nombre de cas favorables / nombre de cas possibles .

<u>Probabilités conditionnelles</u>

On notera Prob(A I B) la probabilité de A si l'on suppose que B a lieu.

Ce qui donne dans le cas de la loi uniforme sur un ensemble fini

Prob(A I B) =
card(A ∩ B) / card(B) =
card(A ∩ B) / card(Ω) multiplié par
card(Ω) / card(B) =
Prob(A I B) = Prob(A ∩ B) / Prob(B)

Prob(A I B) =
Prob(A ∩ B) / Prob(B)
Formule de BAYES

<u>Evénements indépendants</u>

Prob(AIB)=Prob(A)
ou bien
Prob(A∩B)=Prob(A) . Prob(B)

Cette dernière formule se généralise de manière évidente au cas de n événements indépendants .

<u>Formule des probabilités totales</u>

Des considérations ci-dessus ,il découle la formule :
$(B_n)_n$ étant une partition de Ω ,
avec Prob(B_n) > 0 on a

Prob(A) =
$$\sum_n \mathbf{Prob(A\ I\ B_n) . Prob(B_n)}$$

<u>Variable aléatoire</u>

<u>Définitions générales</u>

Une variable aléatoire X est un « objet mathématique » (en fait, on peut l'interpréter comme une fonction définie sur un ensemble probabilisé Ω à valeur dans un ensemble noté Ω_X) qui, lors d'une <u>réalisation</u>, peut prendre différentes valeurs possibles . Par exemple pour le jeu de dé ordinaire à 6 faces : on peut considérer la variable aléatoire X = valeur que l'on obtiendrait si on lançait le dé .

Alors l'événement :
''obtenir un nombre pair lors d'un lancé du dé ''
s'écrit : '' $X \in \{2,4,6\}$ ''

à une variable aléatoire X , on associe une probabilité sur Ω_X :
A -> Prob(X^{-1} (A)) =
Prob($X \in A$)
celle ci s'appelle la loi de X .

X peut être à valeur vectorielle (de dimension n) si l'on veut considérer des n-uplets de variables aléatoires comme une seule variable aléatoire .
X pourrait être non numérique (par exemple : une couleur ;une réponse ''oui '' ou ''non '' …) mais dans ce cas, on la numérise symboliquement .

Pour un couple (X , Y) de variables aléatoires, on appelle loi conjointe du couple, la loi de probabilité :

$$\Delta \;\; \text{->} \;\; \text{Prob}(\ (X,Y) \in \Delta\ )$$

(souvent la considération des parties Δ de la forme : $\Delta = A \quad B$ suffit).

Généralisation immédiate au cas de n variables aléatoires.

<u>Variables aléatoires indépendantes</u>

X et Y sont dites indépendantes

ssi

‘’ $X \in A$ ‘’ et ‘’ $Y \in B$ ‘’ sont des événements indépendants pour tout A et B

ce qui se traduit par :

Prob($X\in A$ et $Y\in B$)=
Prob($X\in A$) . Prob($Y\in B$)

Généralisation immédiate au cas de n variables aléatoires indépendantes.

Loi marginale

Si Ω_X désigne l'ensemble des valeurs possibles pour X alors :

si on connaît la loi conjointe du couple (X ,Y) on a :

$$\text{Prob}(Y \in B) = \text{Prob}(X \in \Omega_X \text{ et } Y \in B)$$

La loi de probabilité

B -> Prob ($Y \in B$)

ainsi récupérée s'appelle :

loi marginale de Y .

Remarque:

Dans le cas de v.a indépendantes les lois marginales donnent la loi conjointe

Parmi toutes les lois de probabilité que l'on peut théoriquement définir , 2 catégories sont utilisées dans la pratique (Statistiques , Traitement du Signal , Théorie de l'information , Recherche Opérationnelle , Complexité d'algorithmes en moyenne , Recouvrement aléatoire de graphes , etc…) ce sont : les probabilités discrètes et les probabilités diffuses avec densité continue.

Loi de probabilité discrète

Construction générale

Ω est dénombrable, F est la tribu formée par toutes les parties de Ω .

on considère une suite de nombres réels $p(w_n)$ positifs t.q $\sum_n p(w_n) = 1$

alors A -> $\text{Prob}(A) = \sum p(w_n)$ la sommation $\sum$ portant sur les indices n t .q $w_n \in A$

ceci donne la forme de toutes les lois de probabilités appelées : discrètes .

Variable aléatoire discrète

On appelle variable aléatoire discrète une variable aléatoire dont la loi est discrète .

Espérance mathématique d'une variable aléatoire discrète

Définition

X étant une variable aléatoire discrète , l'ensemble de ses valeurs possibles Ω_X est dénombrable, donc on peut le représenter par une suite x_n .

On appelle Espérance mathématique d'une variable aléatoire discrète X , et on la note E(X) la valeur :

$$\mathbf{E(X) = \sum_{n} x_n \ Prob(X = x_n)}$$

avec la condition :
la série de terme général x_n Prob(X = x_n) est absolument convergente.

Pour signaler cette condition , on emploiera simplement la terminologie : ''E(X) existe''

Pourquoi demande t-on la convergence ABSOLUE ???

Formule de transfert

f étant une fonction numérique définie sur Ω_X = l'ensemble des valeurs possibles pour une variable aléatoire discrète X ; l'ensemble des valeurs possibles pour f(X) est dénombrable , on peut donc le représenter par une suite x'_n

on a par définition :

$E(f(X))=\sum_n x'_n \, Prob(f(X)=x'_n)$

formule de transfert **:**

$\mathbf{E(f(X)) =}$

$\sum_n \mathbf{f(x_n)\ Prob(X = x_n)}$

Linéarité de E

Pour α , β 2 scalaires et X , Y 2 variables aléatoires :

$\mathbf{E(\alpha X + \beta Y) =}$

$\mathbf{\alpha E(X) + \beta E(Y)}$

<u>Variable aléatoire constante</u>

On assimile une constante ct et une variable aléatoire X dont la loi est :

Prob(X= ct) = 1 .

Ainsi E(ct) = E(X) = ct

Donc E(Y – E(Y)) = 0 pour toute variable aléatoire Y .

On dit que Y –E(Y) est centrée .

<u>Formule à propos de E(XY)</u>

Si X et Y sont indépendantes alors on a :

E(XY) = E(X) E(Y)

<u>Variance, covariance</u>

Dans ce qui suit les notations sont adaptées au cas de variables aléatoires à valeurs réelles , l'adaptation au cas de variables aléatoires à valeurs complexes est aisée, il suffit de remplacer la forme bilinéaire symétrique

$C_{X,Y}$ = E(XY) – E(X) E(Y)

par la forme sesquilinéaire hermitienne correspondante .

Variance

On montre (en majorant séparément les $|x_n|$ dans la série

$\sum_n |x_n|$ Prob(X = x_n)

selon que $|x_n| > 1$ ou que $|x_n| \leq 1$) que :

$E(X^2)$ existe implique $E(X)$ existe

(la réciproque est fausse en général) .

Dans le cas où $E(X^2)$ existe ;
on appelle variance de X :

$$\mathbf{Var(X) = E(X^2) - (E(X))^2}$$

on a la formule ci dessous :

$$\mathbf{Var(X) = E(X^2) - (E(X))^2}$$
$$\mathbf{= E((X- E(X))^2)}$$

C'est cette formule qui permet de donner la signification heuristique de la variance

α étant un scalaire on a :

$Var(\alpha X) = \alpha^2 \ Var(X)$

pour cette raison on est amené à introduire la notion d'écart type de X

$$\boldsymbol{\sigma(X) = \sigma_X = (Var(X))^{1/2}} \ .$$

on a : $\sigma(\alpha X) = \alpha\, \sigma(X)$

pour tout scalaire α positif .

Covariance

Considérons un couple X,Y t.q
$E(X^2)$ et $E(Y^2)$ existent
alors E(XY) existe
Dans ce cas on appelle covariance du couple X,Y et on note $C_{X,Y}$ la forme bilinéaire symétrique , positive

$$\mathbf{C_{X,Y} = E(XY) - E(X)\,E(Y)}$$
$$\mathbf{= E(\ (X-E(X))\ (Y-E(Y))\ )}$$

Remarque $\mathrm{Var}(X) = C_{X,X}$

si X et Y sont indépendantes alors
$E(XY) = E(X)\,E(Y)$ d'où
$C_{X,Y} = 0$
(la réciproque de cette propriété est fausse en général)

Coefficient de corrélation

On appelle coefficient de corrélation

$$\rho_{X,Y} = C_{X,Y} \,/\, (\sigma_X \,.\, \sigma_Y)$$

L'intérêt de cette notion est, que ce coefficient de corrélation est invariant par changement d'unités.
Attention à ne pas confondre les notions de corrélation et de causalité.

L'inégalité de Cauchy-Schwartz appliquée à la forme bilinéaire symétrique positive $C_{X,Y}$ donne :

$$C_{X,Y}{}^2 \leq C_{X,X} \; C_{Y,Y}$$

$$\mathrm{I}\, C_{X,Y} \,\mathrm{I} \leq \sigma_X \; \sigma_Y$$

D'où $\mathbf{I}\, \rho_{X,Y} \,\mathbf{I} \leq 1$

<u>Variance d'une somme</u>

$Var(X + Y) =$
$Var(X) + Var(Y) + 2\, C_{X,Y}$

Corollaire : si X et Y sont indépendantes alors $Var(X + Y) = Var(X) + Var(Y)$.

(car $C_{X,Y} = 0$)

Généralisation immédiate au cas de N variables aléatoires indépendantes, ce qui donne le résultat important suivant :

$X_1 , X_2 , \ldots X_N$ étant N variables aléatoires indépendantes qui suivent la même loi que X (on dit que $X_1 , X_2 , \ldots X_N$ est un échantillon de X, de taille N)

$$\bar{X} = 1/N \sum_n X_n$$

$\bar{X}$ est appelé estimateur de moyenne

formules très importantes:

$$E(\bar{X}) = E(X)$$

$$Var(\bar{X}) = 1/N \; Var(X)$$

autre formule importante :
si on pose :

$$\mathbf{V^* = 1/(N-1) \sum_i (X_i - \bar{X})^2}$$

alors

$$\mathbf{E(V^*) = Var(X)}$$

Calculs :

$$E(\bar{X}^2) = E\left(1/N^2 \sum_i X_i \sum_j X_j\right)$$

$$= 1/N^2 \, E\left(\sum_i X_i^2 + \sum_{i,j;i\neq j} X_i X_j\right)$$

$$= 1/N^2 \left(N \, E(X^2) + (N^2 - N) \, E(X)^2\right)$$

$$= 1/N \, E(X^2) + (N-1)/N \, E(X)^2$$

$$\mathbf{V^* = 1/(N-1) \sum_i (X_i - \bar{X})^2}$$

$$\mathbf{V^* = N/(N-1)} \; \overline{(X - \bar{X})^2}$$

$$\mathbf{= N/(N-1) (} \, \overline{X^2} - \bar{X}^2 \, \mathbf{)}$$

$E(V^*) = N/(N-1) \, E(\overline{X^2} - \bar{X}^2) =$

or $E(\bar{X}^2) = 1/N \, E(X^2) + (N-1)/N \, E(X)^2$

$E(V^*) = N/(N-1)$ multiplié par

$(E(X^2) - 1/N \, E(X^2) - (N-1)/N \, E(X)^2)$

$= E(X^2) - E(X)^2 = Var(X)$

<u>Inégalité de Tchebycheff</u>

Théorème : $b > 0$ et Y une variable aléatoire (t.q $E(Y^2)$ existe) ; on note σ son écart type alors
$\text{Prob}(| Y - E(Y) | > b\,\sigma) \leq 1/b^2$

cette inégalité a surtout une utilisation théorique

<u>Théorème des grands nombres :</u>

Soit X une variable aléatoire t.q E(X) existe,
$X_1, X_2, \dots X_N$ désigne un échantillon de X de taille N
Alors
$\forall \varepsilon > 0 \;\; \text{Prob}(| \bar{X} - E(X) | > \varepsilon)$
converge vers 0 lorsque N tend vers $\propto$.

on dit que $\bar{X}$ converge en probabilité vers E(X) lorsque N tend vers $\propto$, on écrira ceci sous la forme symbolique :

pour N suffisamment grand
$\bar{X}$ eq.prob E(X)

$V^* = \mathbf{N/(N-1)}\ (\overline{X^2} - \bar{X}^2)$ donne aussi :

pour N suffisamment grand
V* eq.prob Var(X)

Y étant une variable aléatoire,
d'espérance = $E(Y)$
souvent notée aussi m,
d'écart type σ

$$\mathbf{X = (Y - m)/\sigma}$$

est dite
variable aléatoire réduite associée à Y,
elle est telle que
$E(X)=0$ et $\sigma(X)=1$

loi Binomiale

Au cours d'une certaine expérience , un événement A , a la probabilité d de se réaliser .
On considère la variable aléatoire X = le nombre d'occurrences de cet événement A que l'on obtiendrait si l'on effectuait N telles expériences de manière indépendante .
Les variables aléatoires caractéristiques de cet événement A au cours de ces expériences sont appelées : variables aléatoires de Bernoulli

La loi de X est :
$$\mathrm{Prob}(X=k) = C^k_N \, d^k \, (1-d)^{N-k}$$

$$E(X) = N\, d$$
$$\mathrm{Var}(X) = Nd\,(1-d)$$

Loi de Poisson de pramètre a>0

C'est la loi entière :
Prob(X = k) = exp(-a) . (a^k / k !)
pour tout entier positif k .

E(X) = a et Var(X) = a

Théorème : posons $a = N d$
a étant fixé .
lorsque N tend vers l'infini
$C^k_N \ d^k \ (1 - d)^{N-k}$
converge vers $\exp(-a) . (a^k / k !)$

car $C^k_N \ d^k \ (1 - d)^{N-k} =$
$N/N. (N-1)/N ..(N-k+1)/N. (1-a/N)^{-k}$.
$(1-a/N)^N . (a^k /k!)$

Dans la pratique dès que N est assez grand ,
a restant < 10 (environ) on est amené à approcher la loi Binomiale de paramètre d et d'ordre N , par la loi de Poisson de paramètre $a = N d$
<u>Théorème</u> : Si X et Y sont indépendantes et suivent des lois de Poissons de paramètres respectifs a et b Alors
$S = X + Y$ suit la loi de Poisson de paramètre $a + b$.

Généralisation immédiate au cas de N variables aléatoires .

exemple : Soit $X_1 , X_2 , \ldots X_N$ un échantillon de X qui suit la loi de poisson de paramètre a, alors $N \bar{X}$ suit la loi de Poisson de paramètre $N a$.