MSI – Semestre 5 – LST Informatique
Chapitre I – Probabilités discrètes – Ce Qu'il Faut Retenir
Ensemble des définitions, propositions et théorèmes à connaître

Généralités

Définition espace de probabilité discret : Une espace de probabilité discret est un couple (Ω, Pr) , où Ω est un ensemble non vide, au plus dénombrable, et Pr une application de Ω dans $[0,1]$ telle que $\sum_{w \in \Omega} Pr(w) = 1$ (somme des probabilités des éléments de $\Omega = 1$).

Définition variable aléatoire : Soit (Ω, Pr) un espace de probabilité discret et soit Ω' un ensemble non vide au plus dénombrable. Une variable aléatoire (v.a) X à valeurs dans Ω' est une application de Ω dans Ω' . Nous prenons souvent pour Ω' un sous ensemble de N ou R . On pourra munir Ω' d'une loi de probabilité Pr_X en posant $\forall w' \in \Omega' : Pr_X(w') = Pr(X^{-1}(\{w'\}))$.

Propositions élémentaires sur la variable aléatoire :

- Pr_X est une loi de probabilité sur $\Omega' \Leftrightarrow \sum_{w' \in \Omega'} Pr_X(w') = 1$.
 - Nous avons de plus, pour tout $A' \subset \Omega' : Pr_X(A') = Pr(X^{-1}(A'))$.
-

Espérance et variance

Définition Espérance mathématique : Une variable aléatoire admet un certain nombre de valeurs typiques. Nous considérons dans la suite les v.a à valeurs dans R . En arithmétique usuelle, la valeur moyenne de n nombre est définie comme leur somme divisée par n . En calcul de probabilité, l'espérance d'une variable aléatoire est définie comme la somme des valeurs prises pondérées par les probabilités respectives c'est-à-dire $E_X = \sum_{x \in X(\Omega)} x \cdot Pr(X=x)$ lorsque cette somme converge absolument.

$X(\Omega)$ est l'ensemble des valeurs prises par la variable aléatoire X

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

Linéarité de l'espérance : Soient X_1 et X_2 deux v.a définies sur le même espace de probabilité discret (Ω, Pr) et admettant toutes deux une espérance E . Soit $\alpha \in R$ alors :

- $E(\alpha X_1) = \alpha E(X_1)$;
 - $E(X_1 + X_2) = E(X_1) + E(X_2)$.
-

Distribution conjointe : On peut considérer plusieurs v.a sur le même espace probabilisé discret (Ω) . Soient les v.a X et Y définies sur (Ω, Pr) . La loi (ou distribution conjointe) est la donnée de : $Pr_{XY} = (X=x, Y=y) = Pr(\{w \in \Omega / X(w)=x, Y(w)=y\}) \forall x, y$.

X et Y sont dites indépendantes, si $\forall x, \forall y$ on a $Pr_{XY}(W=x, Y=y) = Pr_X(x) \cdot Pr_Y(y)$.

Si X et Y sont deux valeurs indépendantes admettant une espérance, alors la valeur aléatoire produite XY admet une espérance et $E(XY) = E(X) \cdot E(Y)$.

Variance : Un paramètre important qui vient tout de suite après l'espérance est la Variance. Elle mesure la dispersion d'une valeur aléatoire. Si X est une v.a définie sur (Ω, Pr) , sa variance est définie par $V_X = E((X - E_X)^2)$ lorsque celle-ci existe.

Nous avons $V_X = E(X^2) - E(X)^2$.

Si X et Y sont 2 v.a indépendantes admettant chacune une variance, alors la v.a de $X+Y$ admet une variance qui est la somme des deux variances.

La racine carrée de la variance est appelée écart type et est notée par σ : $\sigma_X = \sqrt{V_X}$

Fonctions génératrices

Définition : Lorsque les V.a prennent des valeurs dans N , l'utilisation des fonctions génératrices pour les suites des probabilités et les techniques associées deviennent les outils favoris dans l'étude des distributions discrète. La fonction génératrice d'une v.a X à valeurs dans N est définie par : $G_X(z) = \sum_{k \geq 0} Pr(X=k).z^k$

Cette série entière en z contient toutes les informations sur la v.a. X . Elle est par ailleurs l'espérance de z^X , puisque : $G_X(z) = \sum_{w \geq 0} Pr(w).z^{X(w)}$.

Propriétés : Dans la suite nous retirons X de $G_X(z)$ lorsqu'il n'y a pas de danger de confusion. On peut retenir deux propriétés élémentaires :

- $G(1)=1$.
- On peut obtenir l'espérance et la variance d'une variable aléatoire à partir de sa fonction génératrice par :
 - $EX=G'(1)$.
 - $VX=G''(1)+G'(1) - [G'(1)]^2$.

Soient X et Y deux v.a indépendantes. La fonction génératrice de la v.a $X+Y$ est le produit des fonctions génératrices de X et de Y : $G_{X+Y}(z)=G_X(z). G_Y(z)$.

Quelques distributions discrètes

Loi de Bernoulli : La v.a prend deux valeurs : 1 avec la probabilité p ou 0 avec la probabilité q . On suppose que $p, q \in [0, 1]$ et $p+q=1$. Nous avons :

- $EX=p$
- $VX=pq$
- $\sigma X=\sqrt{pq}$

Cette loi intervient souvent de façon implicite lorsqu'on veut traiter une probabilité comme une espérance. En effet c'est la loi de v.a qui est la fonction indicatrice d'un évènement A de probabilité p : $Pr(A) = EI_A$.

Loi Binomiale : On effectue n épreuves identiques et indépendantes. La probabilité de succès dans chacune étant égale à p et celle d'échec $q=1-p$. Posons X = (le nombre total de succès obtenus dans les épreuves). X est

une v.a qui peut prendre la valeur k dans l'intervalle $[0, n]$ avec la probabilité :

$$p_k = \Pr(X = k) = \binom{n}{k} p^k q^{n-k}.$$

On dit alors que X suit une loi binomiale de paramètres n et p .

Il est clair qu'on peut considérer la v.a binomiale X comme la somme de n v.a de Bernouilli identiques et indépendantes de même paramètre p . Nous avons donc

- $EX = np$,
- $VX = npq$,
- $\sigma X = \sqrt{npq}$.

Loi géométrique : Soit $p > 0$ la probabilité de succès dans une épreuve. Posons $q = 1 - p$. Nous répétons la même épreuve indépendamment jusqu'à l'obtention d'un premier succès. Soit X la v.a désignant le nombre d'épreuves effectués. C'est une v.a qui peut prendre la valeur $k \in \mathbb{N}^*$ avec la probabilité $p_k = \Pr(X = k) = q^{k-1} p$.

On dit alors que X suit une loi géométrique de paramètre p . On calcule son espérance et sa variance soit par une méthode directe, soit en dérivant sa fonction géométrique :

- $EX = G'(z) = 1/p$
- $VX = G''(z) = q/p^2$

Loi de Poisson : Cette distribution intervient dans l'étude du nombre d'évènements intervenant dans un intervalle de temps (files d'attente). Elle propose aussi des modèles statistiques pour des évènements dits « rares » lorsqu'un grand nombre de cas a été étudié. La variable aléatoire X suit une loi de Poisson de

paramètre λ , si X peut prendre la valeur $k \in \mathbb{N}$ avec la probabilité : $p_k = \Pr(X = k) = \frac{\lambda^k}{k!} e^{-\lambda}$

L'espérance et la variance d'une distribution de Poisson se calculent aisément et on a :

- $EX = \lambda$
- $VX = \lambda$

Propriétés de la loi de Poisson : La somme de deux v.a de Poisson indépendantes de paramètres λ et μ est une variable de Poisson de paramètre $\lambda + \mu$.

On peut faire l'approximation de la loi binomiale par la loi de Poisson : Soit X une v.a binomiale de paramètres,

p, n . On suppose $np = \lambda > 0$. Alors $\Pr(X_n = k) = \frac{\lambda^k}{k!} e^{-\lambda}$. $K = [0, 1, \dots, n]$.

Indépendance

Définition : Deux évènements A et B sont dit indépendants si $\Pr(A \cap B) = \Pr(A) \cdot \Pr(B)$

Une famille $A_i, i \in 1 \dots n$ d'évènements est dite indépendante dans son ensemble si pour tout sous ensemble $J \subset \{1, \dots, n\}$ on a

$$\Pr\left(\bigcap_{i \in J} A_i\right) = \prod_{i \in J} \Pr(A_i)$$

Probabilité conditionnelle

Définition : Soit $(\Omega, \Pr)$ un espace de probabilité discret et soit A un évènement de probabilité non nulle. On définit par $(P(\Omega))$, l'application $\Pr(. / A)$ à valeurs dans $[0, 1]$ par : $\Pr(B / A) = \frac{\Pr(A \cap B)}{\Pr(A)}, \forall B \in P(\Omega)$.

On appelle $\Pr(B/A)$ probabilité (conditionnelle) de B sachant A .

Propriétés :

- Le couple $(\Omega, \Pr(. / A))$ est un espace de probabilité discret.
- $\Pr(\Omega / A) = \Pr(A) / \Pr(A) = 1$.
- Soit $A_1, \dots, A_n$, un partition de Ω . Si chacun de ces ensembles est de probabilité non nulle, alors :

$$\Pr(B) = \sum_{i=1}^n \Pr(A_i) \cdot \Pr(B / A_i), \forall B \in P(\Omega).$$

Probabilité des causes (théorème de Bayes) : Soit $A_1, A_2, \dots, A_n$ une partition de Ω telle que $\Pr(A_i) \neq 0, \forall i = 1, \dots, n$. Soit B un évènement de probabilité non nulle. Nous avons

$$\Pr(A_i / B) = \frac{\Pr(A_i) \Pr(B / A_i)}{\sum_{i=1}^n \Pr(A_i) \Pr(B / A_i)}, i = 1, 2, \dots, n$$

En particulier, pour deux évènements A et B de probabilité non nulle, nous avons :

$$\Pr(A / B) = \frac{\Pr(A) \Pr(B / A)}{\Pr(B)} \text{ et on appelle :}$$

- $\Pr(A)$: probabilité a priori de A ,
- $\Pr(A/B)$: probabilité a posteriori de A .

Loi conjointe

Définition : Soient X et Y deux v.a réelles sur le même espace Ω . Le couple (X, Y) peut être considéré comme un vecteur aléatoire sur Ω . Sa loi est appelée loi conjointe de (X, Y) . La connaissance de cette dernière permet de retrouver la loi de X et celle de Y individuellement. En effet, pour tout ensemble $A \subset X(\Omega)$ et tout ensemble $B \subset Y(\Omega)$:

- $\Pr_1(X \in A) = \Pr(X \in A, Y \in R)$ et
- $\Pr_2(Y \in A) = \Pr(X \in R, Y \in B)$.

Remarques :

- Les distributions $\Pr_1$ et $\Pr_2$ sont appelées lois marginales du vecteur (X, Y) .
- La connaissance des lois marginales ne permet pas de reconstruire la loi conjointe.

Espérance conditionnelle

Définition : Soit X une v.a et soit A un évènement de probabilité non nulle. L'espérance de X sachant A est définie par : $E(X / A) = E_{\Pr(\cdot / A)} X$, c'est-à-dire : $E(X / A) = \frac{1}{\Pr(A)} \sum_{\omega \in A} X(\omega) \Pr(\{\omega\})$.

Loi des grands nombres

Proposition (inégalité de Bienaymé-Tchebychev) : Soit X une variable aléatoire d'espérance μ et de variance V . Alors, pour tout $e > 0$ $\Pr(|X - \mu| \geq e) \leq \frac{V}{e^2}$.

Théorème (loi faible des grands nombres) : Soit X une variable aléatoire admettant l'espérance μ et la variance V et soit $X_1, X_2, \dots, X_n$ une suite de variable aléatoires indépendantes chacune suivant la même loi que X .

Désignons par S_n la valeur moyenne de la suite : $S_n = \frac{1}{n} \sum_{i=1}^n X_i$.

Alors S_n converge en probabilité vers son espérance μ , c'est-à-dire pour tout $e > 0$ on a $\lim_{n \rightarrow \infty} \Pr(|S_n - \mu| \geq e) = 0$.