EPITECH
PROBABILITES ET STATISTIQUES
Cours203
Loi discrètes

Dominique Neveu

Année 2009-2010

Table des matières

Résumé du cours

Les lois discrètes les plus utilisées que nous présentons ici sont : la loi de Bernouilli ou loi du succès/échec, la loi binomiale ou somme d'une série d'épreuves de Bernouilli, la loi de Poisson qui convient particulièrement au phénomène de comptage d'évènements rares, la loi hypergéométrique ou loi du tirage sans remise.

Chapitre 1

Lois discrètes

1.1 Avertissement au lecteur

Dans ce cours, on utilise les nombres factorielles, notés n !, ainsi que les combinaisons, notés C_n^k. Voici deux petits paragraphes de rappels au sujet de ces nombres.

1.1.1 Factorielle

Définition 1 *Une permutation d'un ensemble fini E d'éléments distincts est une séquence ordonnée de tous les éléments de E, chaque élément apparaissant exactement une fois.*

Exemple 1 *Il existe 6 permutations de l'ensemble de trois lettres* $E = \{a,b,c\}$ *:*

$$abc, acb, bac, bca, cab, cba$$

Le nombre de permutations d'un ensemble de n éléments (ou n-uplet) est :

$$p(n) = n.(n-1).(n-2)...3.2.1 = n!$$

puisqu'il y a n façons de choisir le premier élément, n-1 de choisir le deuxième, et ainsi de suite. Traditionnellement on note ce nombre n ! ce qui se lit "factorielle n". Par convention, on a : $0\,! = 1$. On peut citer quelques valeurs :

$1\,! = 1$	$2\,! = 2$	$3\,! = 6$	$4\,! = 24$
$5\,! = 120$	$6\,! = 720$	$7\,! = 5\;040$	$8\,! = 40\;320$
$9\,! = 362\;880$	$10\,! = 3\;628\;800$	...	

1.1.2 Combinaison

Dans la plupart des jeux de cartes, l'ordre dans lequel sont tirées les cartes n'a aucune importance. On s'intéresse alors aux sous-ensembles à k éléments d'un ensemble de n éléments. On appelle un tel sous-ensemble une combinaison sans répétition de n éléments pris k à k.

Définition 2 *Une combinaison de k éléments d'un ensemble fini E est un sous-ensemble de E possédant k éléments.*

L'ordre des éléments dans la combinaison n'intervient pas.

Exemple 2 *Les combinaisons de 2 lettres parmi les lettres de l'ensemble E = {a,b,c} sont les trois suivantes :*

$$ab, ac, bc$$

Le nombre de combinaisons est égal à :

$$C_n^k = \frac{n!}{k!(n-k)!}$$

On C_n^k le nombre de combinaisons de k éléments d'un ensemble de n éléments. Une autre notation couramment utilisée est

$$C_n^k = \begin{pmatrix} n \\ k \end{pmatrix}$$

On peut citer quelques formules :

. $C_n^0 = 1$
. $C_n^1 = n$
. $C_n^2 = \frac{n.(n-1)}{2}$
. ...

1.2 Loi uniforme discrète

1.2.1 Présentation

La loi uniforme discrète s'applique à la plupart des jeux de hasard comme le jeu de dé, le lancer de pièce, la roulette, le loto ...

1.2.2 Loi de probabilité

La loi uniforme discrète est la loi d'une variable aléatoire X qui prend les valeurs x_i, i=1, 2, ..., n avec des probabilités uniformément distribuées :

$$P(X = x_i) = \frac{1}{n}$$

L'espérance et la variance de la loi uniforme sont données par :

$$\begin{aligned} \mathrm{E(X)} &= \frac{1}{n}\sum_{i=1}^{n} x_i = \overline{x} \\ \mathrm{Var(X)} &= \frac{1}{n}\sum_{i=1}^{n}(x_i - \overline{x})^2 = \sigma(x)^2 \end{aligned}$$

On retrouve les grandeurs statistiques de la moyenne et de la variance des valeurs (x_i).

1.2.3 Exemple

Exemple 3 *Prenons l'exemple du jeu de dé, X étant la variable aléatoire égale au point obtenu. On a :*

$$P(X = i) = \frac{1}{6}, \text{ pour } i{=}1,\dots 6$$
$$E(X) = \frac{1}{6}\sum_{i=1}^{6} i = 3{,}5$$
$$Var(X) = \frac{1}{6}\sum_{i=1}^{6}(i-3,5)^2 \simeq 2,917$$

1.3 Loi de Bernouilli

1.3.1 Présentation

Une expérience aléatoire n'ayant que deux résultats possibles "succès" ou "échec" est dite de Bernouilli. On peut citer comme exemples le sexe d'un bébé à la naissance, la qualité bonne ou défectueuse d'un objet extrait d'une production...

1.3.2 Loi de probabilité

Soit A l'évènement "succès" de l'expérience aléatoire. Sa fonction indicatrice notée 1_A est la variable aléatoire de l'expérience :

$$1_A(x) = \begin{cases} 1 & \text{si } x \in A, \\ 0 & \text{si } x \in A^C \end{cases}$$

La loi de probabilité correspondante est la loi de Bernouilli :

$$\begin{array}{lll} \mathrm{P}(1_A = 1) & = \mathrm{P(A)} & = \mathrm{p} \\ \mathrm{P}(1_A = 0) & = \mathrm{P}(A^C) & = 1 \text{ - p} \end{array}$$

L'espérance et la variance de la loi de Bernouilli sont données par :

$$\begin{array}{l} \mathrm{E(X) = p} \\ \mathrm{Var(X) = p(1 - p)} \end{array}$$

1.3.3 Exemple

Exemple 4 *Soit une urne qui contient une boule blanche B et trois boules vertes V, le tirage d'une boule en considérant qu'obtenir une boule blanche est un succès est une épreuve de Bernouilli. Pour cette expérience, on a :*

$$p = \frac{1}{4} = 0{,}25$$
$$E(X) = 0{,}25$$
$$Var(X) = 0{,}1875$$

1.4 Loi binomiale

1.4.1 Présentation

Lorsque l'on répète n fois une épreuve de Bernouilli dans des conditions identiques, le résultat d'une épreuve étant indépendant des résultats précédents, on dit que l'on a une suite d'épreuves de Bernouilli.

L'exemple classique de la loi binomiale est celui de l"urne de Bernouilli". Cette urne contient des boules blanches et des boules noires. On tire successivement une boule "avec remise" dans l'urne. Ainsi les proportions de boules ne changent pas.

1.4.2 Loi de probabilité

Soit A l'évènement succès de l'épreuve de Bernouilli de probabilité p. Soit X la variable aléatoire définie par le nombre d'apparitions de l'évènement A au cours des n épreuves (0 ≤ X ≤ n). On dit alors que X suit une loi binomiale de paramètres n et p. On note X ∼ B(n,p).
Afin de chercher la loi de probabilité, c'est à dire l'expression de P(X=k) de la loi binomiale, remarquons que toutes les configurations telles que k épreuves présentent le résultat A, et n-k épreuves présentent le résultat A^C, sont équiprobables et qu'il y a C_n^k configurations de cette sorte. Typiquement, un évènement formé de k "succès" et (n-k) "échecs" (dans n'importe quel ordre), a comme probabilité une valeur proportionnelle à $p^k.(1-p)^{n-k}$. On obtient donc la loi de probabilité suivante :

$$P(X=k) = C_n^k p^k (1-p)^{n-k}$$

qui est la probabilité d'obtenir k succès parmi n épreuves.

L'expression est facilement calculable quand n est petit, par exemple, n plus petit que 10 ou 12. Quand n est plus grand, le calcul est plus élaboré et demande plus d'efforts. Pour des valeurs modérées de n, inférieures à 25 ou 30, des tables de probabilités binomiales sont disponibles. Quand n est grand, supérieur à 25 ou 30, on peut utiliser des approximations (voir plus loin) : la distribution de Poisson.

L'espérance mathématique et la variance d'une variable binomiale sont données par :

E(X) = np
Var(X) = np(1 - p)

1.4.3 Exemple

Exemple 5 *Un jury est composé de 12 personnes choisies au hasard et d'une façon indépendante à partir de la liste électorale d'une commune. Il il y a quatre fois plus d'hommes que de femmes dans la liste. Quelle est la probabilité que le jury soit composé d'autant de femmes que d'hommes ?*

On choisit successivement 12 personnes dans une liste. A chaque choix, il y a une probabilité $p=\frac{1}{5}$ *de choisir une femme. La constitution du jury suit donc une loi binomiale B(n,p) de paramètres n=12 et* $p=\frac{1}{5}$ *Soit X la variable aléatoire représentant le nombre k de femmes chosies, on a la loi de probabilité suivante :*

$$P(X=k)=C_{12}^{k}\left(\frac{1}{5}\right)^{k}\left(\frac{4}{5}\right)^{12-k}$$

En particulier, le jury est formé d'autant de femmes que d'hommes s'il y a exactement 6 femmes. La probabilité de cet évènement est donc :

$$P(X=6)=\frac{12!}{6!6!}\left(\frac{1}{5}\right)^{6}\left(\frac{4}{5}\right)^{6}\simeq 0,0155$$

1.5 Loi de Poisson

1.5.1 Présentation

La loi de Poisson est un modèle probabiliste qui convient particulièrement au phénomène de comptage d'évènements rares situés dans le temps ou dans l'espace.

Exemple 6 (Dans le temps) *S'agissant du temps, on peut citer : le nombre de particules émises par une substance radioactive, le nombre d'erreurs téléphoniques enregistrées par une centrale téléphonique, le nombre d'accidents intervenus sur une autoroute par jour, ou encore le nombre d'arrivées à un guichet.*

Exemple 7 (Dans l'espace) *En ce qui concerne l'espace, on peut étudier le nombre de bactéries contenues dans une préparation microscopique, le nombre d'éléphants dans une jungle,...*

En général, nous pouvons étudier toute distribution de "points" lorsque ces points se positionnent au hasard soit dans le temps, soit dans l'espace.

1.5.2 Loi de probabilité

On suppose qu'un seul évènement arrive à la fois, que le nombre d'évènements se produisant pendant une période ne dépend que de la durée de cette période, et que les évènements sont indépendants. Alors une variable aléatoire X suit une loi de Poisson de paramètre λ si :

$$P(X=k)=\frac{e^{-\lambda}.\lambda^k}{k!} \qquad k=0,1,2,...$$

où λ représente le nombre moyen d'évènements par unité de temps (ou d'espace). On note X $\sim$ P(λ).

L'espérance et la variance de la loi de Poisson sont données par :

$$\mathrm{E(X)} = \lambda$$
$$\mathrm{Var(X)} = \lambda$$

Ainsi s'agissant de la loi de Poisson, espérance mathématique et variance sont égales.

1.5.3 Exemple

Exemple 8 *Si le nombre moyen d'arrivées de clients à un guichet par minute est égal à 1,9, calculons la probabilité d'observer 5 arrivées dans une minute donnée, supposant que les arrivées sont indépendantes les unes des autres.*
Dans notre problème, la valeur de λ est égale à 1,9, et la valeur de k est égale à 5. Nous aurons donc :

$$P(X=5)=\frac{e^{-1,9}.1,9^5}{5!}\simeq 0,0309$$

La probabilité de voir arriver 5 clients au guichet dans une minute donnée est donc de 3,09%.

1.5.4 Approximation de la loi binomiale par la loi de Poisson

Quand la probabilité binomiale p est très faible (en pratique p<0.1 : l'évènement est rare) et n est grand (n>50), et que np $\leq$ 10 alors la loi binomiale B(n,p) peut être bien approchée par la loi de Poisson de paramètre λ = np.

Exemple 9 *Les lampes fabriquées par une usine, comme toute production, sont parfois défectueuses. Le taux de lampes défectueuses est de 3% pour l'usine en question. Quelle est la probabilité que dans un lot de 100 lampes, 8 soient défectueuses ?*
Soit X le nombre de lampes défectueuses dans un lot de 100 lampes : X est

une variable aléatoire qui suit une loi binomiale de paramètres n = 100 et p = 0,03, on écrit X ∼ B(100, 0,03).
Pour répondre à la question posée, nous calculons la probabilité :

$$P(X = 8) = C_{100}^{8}(0,03)^8(0,97)^{92} = \frac{100!}{8!92!}(0,03)^8(0,97)^{92} \simeq 0,0074$$

Comme n est grand (n > 50) et p est petit (p < 0.1), on peut approcher la loi binomiale par une loi de Poisson de paramètre

$$\lambda = 100.0,03 = 3$$

La probabilité recherchée est dans ce cas :

$$P(X = 8) = e^{-\lambda}\frac{\lambda^8}{8!} = e^{-3}\frac{3^8}{8!} \simeq 0,0081$$

Ce résultat 0,0081 est proche de la valeur exacte 0,0074 obtenue sur la base de la loi binomiale.

1.6 Loi hypergéométrique

1.6.1 Présentation

Soit une population de N individus parmi lesquels une proportion p possède un certain caractère. Il y a donc Np individus possédant ce caractère. On prélève un échantillon de n individus parmi cette population. Le tirage s'effectue d'un seul coup.

1.6.2 Loi de probabilité

Soit X la variable aléatoire égale au nombre d'individus prélevés possédant le caractère. Alors X suit la loi hypergéométrique, de loi de probabilité :

$$P(X = k) = \frac{C_{Np}^{k} C_{N(1-p)}^{n-k}}{C_{N}^{n}}$$

On note X ∼ H(N,n,p). La signification des différentes combinaisons entrant dans la formule est :

C_N^n	nombre d'échantillons possibles
C_{Np}^k	nombre de groupes de k d'individus possédant le caractère
$C_{N(1-p)}^{n-k}$	nombre de groupes de (n-k) individus ne possédant pas le caractère

Le nombre $\frac{n}{N}$ est appelé taux de sondage.

L'espérance et la variance de la loi hypergéométrique sont données par :

$$\begin{aligned} \mathrm{E(X)} &= \mathrm{np} \\ \mathrm{Var(X)} &= \frac{N-n}{N-1} np(1-p) \end{aligned}$$

L'espérance ne dépend pas de N et est la même que dans le cas du tirage avec remise (loi binomiale).

1.6.3 Tendance vers la loi binomiale

Lorsque N devient grand, la loi hypergéométrique H(N,n,p) peut être approchée par une loi binomiale B(n,p). En pratique, on effectue cette approximation dès que $\frac{n}{N} < 10\%$, c'est à dire dès que la population est dix fois plus grande que l'échantillon, ce qui arrive fréquemment dans le cas de sondages.

Exemple 10 *Un échantillon de 2 000 individus conviendra aussi bien pour faire un sondage dans une ville de 200 000 habitants que dans une ville de 2 millions d'habitants.*
Par exemple, choisissons p = 0,5 et k = 1000.
Prenons N = 200 000 et n = 2 000. On obtient :

$$P(X = 1000) = \frac{C_{10^5}^{10^3} C_{10^5}^{10^3}}{C_{2.10^5}^{2.10^3}} \simeq 0,1127$$

Prenons N = 2 millions et n = 2 000. On obtient :

$$P(X = 1000) = \frac{C_{10^6}^{10^3} C_{10^6}^{10^3}}{C_{2.10^6}^{2.10^3}} \simeq 0,1122$$

Par la loi binomiale, on a :

$$P(X = 1000) = C_{2000}^{1000}(0,5)^{1000}(0,5)^{1000} \simeq 0,1121$$