Chapitre 4

Variables aléatoires

4.0.9 Définition d'une variable aléatoire

Aprés la réalisation d'une expérience, on s'intéresse le plus souvent à une fonction du résultat qu'au résultat lui-même.

Exemples

- Jet de 2 dés: on s'intéresse à la somme obtenue sur 2 dés plutôt qu'au résultat de chaque dé.

- Jet d'une pièce plusieurs fois de suite: on s'intéresse au nombre de fois que pile est apparu plutôt qu'à suite de piles et faces obtenus.

Ces grandeurs sont des fonctions réelles définies sur l'ensemble fondamental et sont appelés "variables aléatoires" (v.a.). On peut attribuer une probabilité aux différentes valeurs que la v.a. peut prendre.

4.0.10 Exemples de variables aléatoires

Exemple 1

On considère le jet de 3 pièces de monnaie. On s'intéresse au nombre de piles qu'on va obtenir, qu'on désigne par X.

X est une v.a., elle peut prendre 4 valeurs $X = 0, 1, 2, 3$.

L'ensemble fondamental est $S = \{(ppp), (ppf), \cdots, (fff)\}$.

Les différentes probabilités associées à chaque valeurs de X sont:

$$\begin{aligned}
P(\{X=0\}) &= P(\{fff\}) = 1/8 \\
P(\{X=1\}) &= P(\{(ffp),(fpf),(pff)\}) = 3/8 \\
P(\{X=2\}) &= P(\{(fpp),(pfp),(ppf)\}) = 3/8 \\
P(\{X=3\}) &= P(\{(ppp)\}) = 1/8
\end{aligned}$$

On a $P(\cup_{i=0}^{3}\{X=i\}) = \sum_{i=0}^{3} P(\{X=i\}) = 1$, car X prend nécessairement l'une des valeurs $0, 1, 2, 3$.

Exemple 2

On répéte le jet d'une pièce jusqu'à ce que pile apparaîsse, mais au plus n fois (c'est à dire qu'on arrête l'expérience au n-ième jet quel que soit le résultat obtenu).

Les jets sont indépendants. Pile apparaît avec une probabilité p.

On désigne par X le nombre de jets effectués jusqu'à l'arrêt de l'expérience.

X est une variable aléatoire, elle peut prendre les valeurs $1, 2, 3, \cdots, n$.

Les différentes probabilités associées à chaque valeurs de X sont:

$$\begin{aligned}
P(\{X=1\}) &= P(\{(p)\}) = p \\
P(\{X=2\}) &= P(\{(fp)\}) = (1-p)p \\
P(\{X=3\}) &= P(\{(ffp)\}) = (1-p)^2 p \\
&\vdots \\
P(\{X=n-1\}) &= P(\{(ff\cdots fp)\}) = (1-p)^{n-2} p \\
P(\{X=n\}) &= P(\{(ff\cdots ffp),(ff\cdots fff)\})
\end{aligned}$$

$$= \quad (1-p)^{n-1}p + (1-p)^n = (1-p)^{n-1}$$

On vérifie bien que pour n fixé:

$$P(\cup_{i=1}^{n}\{X=i\}) = \sum_{i=1}^{n} P(\{X=i\}) = 1$$

On remarque que les différents événements qu'on a considéré:

$$(p), (fp), (ffp), \cdots, (ff\cdots fff)$$

appartiennent à des ensembles fondamentaux différents contrairement à l'Exemple 1 précédent:

$$\begin{array}{l}(p) \in \{p, f\} = S_1 \\ (fp) \in \{ff, fp, pf, pp\} = S_2 \\ \vdots \\ \{(ff\cdots ffp), (ff\cdots fff)\} \subset S_n\end{array}$$

4.0.11 Fonctions de répartition

Définition

La fonction de répartition F d'une variable aléatoire X (discrète ou continue) est définie pour tout nombre réel $a(-\infty \leq a \leq +\infty)$ par:

$$F(a) = P(\{X \leq a\})$$

Propriété

- F est une fonction croissante: si $a < b$, alors $F(a) \leq F(b)$
- $\lim_{a\to+\infty} F(a) = 1$
- $\lim_{a\to-\infty} F(a) = 0$
- En plus, on suppose que F est continue à droite

Fonction de répartition et probabilité sur X

Tous les calculs de probabilité concernant X peuvent être traités en termes de fonction de répartition. Par exemple:

$$P(a < X \leq b) = F(b) - F(a), \quad \forall a < b$$

4.0.12 Types de variables aléatoires discrétes

Définition, lois de probabilité

Une v.a. ne pouvant prendre qu'une quantité dénombrable de valeurs est dite "discrète".

Pour une telle v.a. X, on définit sa loi de probabilité $p(.)$ par la fonction:

$$p(a) = P(\{X = a\})$$

Si X peut prendre les valeurs $x_1, x_2, \cdots$ alors $p(x_i) \geq 0,\ i = 1, 2, \cdots$ et $p(x) = 0$ pour toutes les autres valeurs de x.

On aura $\sum_{i=1}^{\infty} P(\{X = i\}) = \sum_{i=1}^{\infty} p(x_i) = 1$.

Les v.a. discrètes sont réparties en plusieurs catégories selon le type de leur loi.

Variables de Bernouilli

On considére une expérience qui peut donner 2 résultats: A ou B.

p est la probabilité de A ($0 \leq p \leq 1$) et $(1-p)$ sera la probabilité de B.

On définit la variable aléatoire X, en lui donnant la valeur $X = 1$ si A est sorti et $X = 0$ si c'est B qui est sorti.

La loi de probabilité de X est alors:

$$\begin{aligned} p(0) &= P\{X = 0\} = 1 - p \\ p(1) &= P\{X = 1\} = p \end{aligned}$$

Une variable aléaoire est dite de Bernouilli s'il existe $p \in [0,1]$ tel que la loi de probabilité soit donner comme auparavant. On a clairement $p(0) + p(1) = 1$.

Variables binomiales

Comme avant, on considére une expérience qui peut donner 2 résultats: A ou B, p est la probabilité de A ($0 \leq p \leq 1$) et $(1-p)$ la probabilité de B.

On réalise maintenant n epreuves indépendantes.

On définit la variable aléatoire X qui compte le nombre de fois que A est sorti sur ces n épreuves.

X peut prendre les valeurs $0, 1, 2, \cdots, n$.

X est appelée v.a. binomiale de paramètres (n, p).

On remarque alors qu'une variable de Bernouilli est une v.a. binomiale de paramètres $(1, p)$.

La loi de probabilité d'une v.a. binomiale de paramètres (n, p) est donnée par:

$$p(i) = P(\{X = i\}) = C_i^n p^i (1-p)^{n-i}, \qquad i = 0, 1, \cdots, n$$

Il y a C_i^n manières d'avoir i résultats A sur les n épreuves.

On remarquera que:

$$\sum_{i=0}^{n} p(i) = \sum_{i=0}^{n} C_i^n p^i (1-p)^{n-i} = 1$$

en utilisant la formule du binôme de Newton.

Variables de Poisson

Une variable aléatoire X pouvant prendre les valeurs $0, 1, 2, \cdots$ est dite de Poisson avec paramètre $\lambda > 0$ si sa loi de probabilité est donnée par:

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

On a bien une loi de probabilité car:

$$\sum_{i=0}^{\infty} p(i) = \sum_{i=0}^{\infty} e^{-\lambda} \lambda^i / i! = 1$$

Variables géométriques

On considére une expérience qui peut donner 2 résultats: A ou B, p est la probabilité de A ($0 \leq p \leq 1$) et $(1-p)$ la probabilité de B.

On exécute une séries d'épreuves indépendantes jusqu'à obtenir le premier résultat A.

On désigne par X la variable aléatoire qui compte le nombre d'épreuves nécessaires jusqu'à ce résultat.

X peut prendre les valeurs $1, 2, 3, \cdots$.

La loi de probabilité est donnée par:

$$p(i) = P(\{X = i\}) = (1-p)^{i-1} \cdot p, \qquad i = 1, 2, \cdots, n$$

car pour $\{X = i\}$, on $(i-1)$ B et un seul A.

On a bien:

$$\sum_{i=1}^{\infty} p(i) = p \cdot \sum_{i=1}^{\infty} (1-p)^{i-1} = \frac{p}{1-(1-p)} = 1$$

X est alors appelée v.a. géométrique de paramètres p.

Variables binomiales négatives

On considére une expérience qui peut donner 2 résultats: A ou B, p est la probabilité de A $(0 \leq p \leq 1)$ et $(1-p)$ la probabilité de B.

On exécute une séries d'épreuves indépendantes jusqu'à obtenir un total de r résultats A.

On désigne par X la variable aléatoire qui compte le nombre d'épreuves nécessaires jusqu'à ce résultat.

X peut prendre les valeurs $r, r+1, r+2, \cdots$.

La loi de probabilité est donnée par:

$$p(i) = P(\{X = i\}) = C_{r-1}^{n-1} (1-p)^{i-r} \cdot p^r, \qquad i = r, r-1, r-2, \cdots, n$$

car pour $\{X = i\}$, le dernier résultat est A et parmi les $(n-1)$ précédents résultats, il y a $(r-1)$ A et $(n-r)$ B d'où le C_{r-1}^{n-1}.

On peut montrer:

$$\sum_{i=r}^{\infty} p(i) = p \cdot \sum_{i=r}^{\infty} C_{r-1}^{n-1} (1-p)^{i-r} = 1$$

Une telle variable X est dite binomiale négative de paramètres (r, p). Une variable géométrique est binomiale négative de paramètres $(1, p)$.

Variables hypergéométriques

On tire sans remise un échantillon de n boules d'une boite en contenant N. Parmi ces N boules, Np sont blanches et $N - Np$ sont noires.

On désigne par X le nombre de boules blanches tirées. On aura:

$$p(i) = P(\{X = i\}) = \frac{C_i^{Np} \cdot C_{n-i}^{N-Np}}{C_n^N}$$

avec $i = max(0, n - N + Np), \cdots, min(n, Np)$.

Une telle v.a. est appelée hypergéométrique.

<u>Exercice</u>

Une boite conteint 20 boules numérotées de 1 à 20.

On tire sans remplacement 3 boules.

Quelle est la probabilité qu'au moins une des 3 boules tirées porte un numéro supérieur à 17?

Réponse:

Si X représente le plus grand numéro tiré.

X prend les valeurs $3, 4, \cdots, 20$.

Les C_3^{20} tirages sont tous équiprobables.

L'événement $\{X = i\}$ correspond au tirage de la boule i et de 2 boules quelconques, portant un numéro entre 1 et $(i-1)$. Il y a clairement $C_1^1 \cdot C_2^{i-1}$ nombres de tirages.

Donc:

$$P(\{X = i\}) = \frac{C_2^{i-1}}{C_3^{20}}, \qquad i = 3, 4, \cdots, 20$$

On aura donc:

$$P(\{X \geq 17\}) = P(\{X = 17\}) + \cdots + P(\{X = 20\}) = 51\%$$

<u>Exercice</u>

D'une boite contenant 3 boules blanches, 3 rouges et 5 noires, on tire 3 boules au hasard.

Supposons que l'on gagne 1 point pour chaque boule blanche tirée et qu'on perd 1 point pour chaque boule rouge.

On désigne par X le nombre de points marqués pour un tirage donné.

X est une v.a. qui prend les valeurs $0, \pm1, \pm2, \pm3$ avec les probabilités respectives:

$$\begin{array}{rcl} P(\{X=0\}) & = & \dfrac{C_3^5 + C_1^3 C_1^3 C_1^5}{C_3^{11}} = \dfrac{55}{165} \\ P(\{X=1\}) & = & P(\{X=-1\}) = \dfrac{C_1^3 C_2^5 + C_2^3 C_1^3}{C_3^{11}} = \dfrac{39}{165} \\ P(\{X=2\}) & = & P(\{X=-2\}) = \dfrac{C_2^3 C_2^5}{C_3^{11}} = \dfrac{15}{165} \\ P(\{X=3\}) & = & P(\{X=-3\}) = \dfrac{C_3^3}{C_3^{11}} = \dfrac{1}{165} \end{array}$$

On s'assure que:

$$\sum_{i=-3}^{3} p(i) = 1$$

La probabilité de gagner à ce jeu est:

$$P(\{X \geq 1\}) = \sum_{i=1}^{3} p(i) = \frac{55}{165} = \frac{1}{3} = 33,3\%$$

Exercice

La probabilité qu'un tireur atteigne une cible est de $5/6 = 83,3\%$ (dépend de l'habilité du tireur). Il tire 9 coups sur la cible.

1) Quelle est la probabilité qu'il atteigne cette cible au moins 7 fois?

2) S'il a atteind la cible 6 fois, quelle est la probabilité que les 3 echecs arrivent juste aprés 3 succés?

Réponse:

1) Désignons par X le nombre de fois que la cible a été atteinte en 9 coups.

X une variable binomiale de paramètres $(9, 5/6)$.

$$P(\{X=i\}) = C_i^9 \left(\frac{5}{6}\right)^i \left(\frac{1}{6}\right)^{9-i}$$

Probabilité d'atteindre la cible au moins 7 fois $= P(\{X \geq 7\}) = 82,2\%$

2) Il y a 4 cas tels que les 3 echecs arrivent juste aprés 3 succés. Il y a $C_6^9 = 84$ cas favorables pour atteindre 6 fois la cible.

D'où la probabilité cherchée $4/84 = 4,8\%$.

Exercice

Les vis fabriquées par une société peuvent avoir un défaut avec une probabilité 0,01. L'état d'une vis est indépendant des autres vis. La société accepte de rembourser les sachets de 10 vis qu'elle vend si plus d'une vis dans le sachet présente un défaut.

Quelle proportion de sachets vendus la société risque de devoir rembourser?

Réponse: On désigne par X le nombre de vis ayant un défaut dans sachet choisi au hasard. On peut voir que X est une variable aléatoire binomiale de paramètres $(n = 10, p = 0.01)$:

$$P(\{X=i\}) = C_j^{10} p^i (1-p)^{10-i}$$

La probabilité de rembourser ce sachet est:

$$P(\{X > 1\}) = 1 - P(\{X=0\}) - P(\{X=1\}) = 0.4\%$$

Propriétés de la loi binomiale et de Poisson

1- Croissance de la loi binomiale

Soit X une variable binomiale de paramètres (n,p) avec $0 < p < 1$. Lorsque k croit de 0 à n, $P(\{X=k\})$ grandit d'abord de manière monotone, puis décroit également de manière monotone, le maximum étant atteind lorsque k est égal à la partie entière de $p(n+1)$.

Preuve:

Considérons le rapport:

$$\frac{P(\{X=k\})}{P(\{X=k-1\})} = \frac{(n-k+1)p}{k(1-p)}$$

On étudie les valeurs de k pour lesquelles, ce rapport est plus grand ou plus petit que un. Par exemple $P(\{X=k\}) \geq P(\{X=k-1\})$ si $k \leq (n+1)p$.

2- Approximation Poissonienne de la loi binomiale

Supposons que X est une v.a. binomiale de paramètres (n,p) et posons $\lambda = np$. On a:

$$\begin{aligned} p(i) &= P(\{X=i\}) \\ &= \frac{n!}{(n-i)!i!}\, p^i(1-p)^{n-i} \\ &= \frac{n!}{(n-i)!i!}\, (\lambda/n)^i(1-\lambda/n)^{n-i} \\ &= \frac{n(n-1)\cdots(n-i+1)}{n^i}\, \frac{\lambda^i}{i!}\, \frac{(1-\lambda/n)^n}{(1-\lambda/n)^i} \end{aligned}$$

Si n est très grand et λ est fixé:

$$\begin{gathered} (1-\lambda/n)^n \sim e^{-\lambda} \qquad (1-\lambda/n)^i \sim 1 \\ \frac{n(n-1)\cdots(n-i+1)}{n^i} \sim 1 \end{gathered}$$

Donc: $p(i) = P(\{X=i\}) = e^{-\lambda}\frac{\lambda^i}{i!}$.

Si n est grand avec λ fixé, la loi binomiale tend vers une loi de Poisson.