



ÉCOLE
D'INGÉNIEURS
PARIS-LA DÉFENSE

Complément de formation Statistiques

Variables Aléatoires Discrètes

Mohamad Ghassany

École supérieure d'ingénieurs Léonard-de-Vinci

1. Introduction aux probabilités
2. Notion de variable aléatoire réelle
3. Variables aléatoires discrètes
4. Moments d'une variable aléatoire discrète
5. Couple de variables aléatoires discrètes
6. Loi Uniforme Discrète $\mathcal{U}(n)$
7. Loi de Bernoulli $\mathcal{B}(p)$
8. Loi Binomiale $\mathcal{B}(n, p)$
9. Loi de Poisson $\mathcal{P}(\lambda)$

Introduction aux probabilités

Hasard

Exemple fondamental: Considérons le jeu du lancé d'un dé.

- ▶ **Expérience aléatoire** ε : "lancer un dé équilibré" \leftarrow **Action**.
- ▶ **Univers:** l'ensemble de tous les résultats possibles de cette expérience aléatoire

$$\Omega = \{1, 2, 3, 4, 5, 6\}$$

- ▶ **Événements:** Dans cette expérience aléatoire, on peut s'intéresser à des événements plus complexes qu'un simple résultat élémentaire.
- ▶ L'**ensemble de parties de Ω** , appelé $\mathcal{P}(\Omega)$, est l'ensemble des sous-ensembles de Ω .
- ▶ Une **famille \mathcal{A} de parties** (i.e. de sous ensembles) de Ω . Ces parties sont appelées des événements. On dit que l'événement A s'est réalisé si et seulement si le résultat ω de Ω qui s'est produit appartient à A .
- ▶ **Tribu:** On appelle **tribu** sur Ω , toute famille \mathcal{A} de parties de Ω vérifiant:
 1. $\Omega \in \mathcal{A}$.
 2. si $A \in \mathcal{A}$, alors $\bar{A} \in \mathcal{A}$.
 3. si $(A_n)_{n \in \mathbb{N}}$ est une suite d'éléments de \mathcal{A} , alors $\bigcup_{n \in \mathbb{N}} A_n \in \mathcal{A}$.
- ▶ (Ω, \mathcal{A}) est un **espace probabilisable**.

- ▶ Soit (Ω, \mathcal{A}) un espace probabilisable:
 - L'ensemble \mathcal{A} est appelé **tribu des événements**. Les éléments de \mathcal{A} s'appellent les **événements**.
 - L'événement Ω est appelé **événement certain**. L'événement \emptyset est appelé **événement impossible**.
- ▶ **Opérations sur les événements**. Soient A et B deux événements:
 - \bar{A} est l'événement **contraire** de A (on note aussi A^c). $\bar{A} = \Omega \setminus A$.
 \bar{A} se réalise si et seulement si A ne se réalise pas.
 - $A \cap B$ est l'événement «**A et B**».
 $A \cap B$ se réalise lorsque les deux événements se réalisent.
 - $A \cup B$ est l'événement «**A ou B**».
 $A \cup B$ se réalise lorsque au moins un des deux événements se réalise.
- ▶ **Incompatibilité**: A et B sont incompatibles si leur réalisation simultanée est impossible: $A \cap B = \emptyset$.
- ▶ **Implication**: A implique B signifie que si A se réalise, alors B se réalise aussi: $A \subset B$.

- Soit (Ω, \mathcal{A}) un espace probabilisable. On appelle **probabilité** sur (Ω, \mathcal{A}) , toute application

$$P : \mathcal{A} \rightarrow \mathbb{R}$$

vérifiant:

1. $\forall A \in \mathcal{A}, P(A) \geq 0$.
2. $P(\Omega) = 1$.
3. $\forall (A_n)_{n \in \mathbb{N}^*} \in \mathcal{A}^{\mathbb{N}^*}$, une suite d'éléments de \mathcal{A} deux à deux incompatibles, on a:

$$P\left(\bigcup_{n \in \mathbb{N}^*} A_n\right) = \sum_{n=1}^{+\infty} P(A_n)$$

- Le triplet (Ω, \mathcal{A}, P) est appelé **espace probabilisé**.

1. $P(\emptyset) = 0$.
2. $P(A_1 \cup A_2) = P(A_1) + P(A_2) - P(A_1 \cap A_2)$.
3. Si A_1 et A_2 sont incompatibles, $A_1 \cap A_2 = \emptyset$, $P(A_1 \cup A_2) = P(A_1) + P(A_2)$.
4. $P(A_1 \cup A_2 \cup A_3) = P(A_1) + P(A_2) + P(A_3) - P(A_1 \cap A_2) - P(A_1 \cap A_3) - P(A_2 \cap A_3) + P(A_1 \cap A_2 \cap A_3)$.
5. $P(\bar{A}) = 1 - P(A)$.
6. $P(B \setminus A) = P(B) - P(B \cap A)$.
7. $A \subset B \Rightarrow P(A) \leq P(B)$.

Probabilité uniforme sur Ω fini

- Soit Ω un univers fini. On dit que P est la **probabilité uniforme** sur l'espace probabilisable $(\Omega, P(\Omega))$ si:

$$\forall \omega, \omega' \in \Omega, \quad P(\{\omega\}) = P(\{\omega'\})$$

On dit aussi qu'il y a **équiprobabilité** des événements élémentaires.

- Soit $(\Omega, \mathcal{P}(\Omega), P)$ un espace probabilisé fini. Si P est la probabilité uniforme, alors

$$\forall A \in \mathcal{A}, \quad P(A) = \frac{\text{Card}(A)}{\text{Card}(\Omega)}$$

- Soit (Ω, \mathcal{A}, P) un espace probabilisé et $B \in \mathcal{A}$ tel que $P(B) > 0$. L'application P_B définie sur \mathcal{A} par:

$$P_B(A) = P(A|B) = \frac{P(A \cap B)}{P(B)}, \quad \forall A \in \mathcal{A}$$

est une probabilité sur (Ω, \mathcal{A}) ; elle est appelée la **probabilité conditionnelle** sachant B . C'est la probabilité pour que l'événement A se produise sachant que l'événement B s'est produit.

- **Remarque:** $(A|B)$ n'est pas un événement! On utilise la notation $P(A|B)$ par simplicité, mais c'est $P_B(A)$ qui est correcte.
- **Formule des probabilités composées:**

$$P(A \cap B) = P(A|B)P(B) = P(B|A)P(A)$$

- **Formule des probabilités totales:**

- $\forall A \in \mathcal{A}, \quad P(A) = P(A \cap B) + P(A \cap \bar{B})$
- On appelle **système complet d'événements (SCE)**, toute partition dénombrable de Ω formée d'éléments de \mathcal{A} ; c-à-d tout ensemble dénombrable d'événements, deux à deux incompatibles et dont l'union dénombrable est l'événement certain.
- Soit $(B_n)_{n \geq 0}$ un SCE de Ω . On a:

$$\forall A \in \mathcal{A}, \quad P(A) = \sum_{n \geq 0} P(A \cap B_n)$$

- **Indépendance:** Les événements A et B sont indépendants ssi $P(A \cap B) = P(A)P(B)$.

Première formule de Bayes

Soit (Ω, \mathcal{A}, P) une espace probabilisé. Pour tous événements A et B tels que $P(A) \neq 0$ et $P(B) \neq 0$, on a :

$$P(B|A) = \frac{P(A|B)P(B)}{P(A)}$$

Deuxième formule de Bayes

Soit (Ω, \mathcal{A}, P) une espace probabilisé et $(B_n)_{n \geq 0}$ un SCE de Ω t.q. pour tout $n \geq 0$ $P(B_n) \neq 0$. On a pour tout $A \in \mathcal{A}$ t.q. $P(A) \neq 0$

$$P(B_i|A) = \frac{P(A|B_i)P(B_i)}{\sum_{n \geq 0} P(A|B_n)P(B_n)} \quad \forall i \geq 0$$

Notion de variable aléatoire réelle

Définition

Soient ε une expérience aléatoire et (Ω, \mathcal{A}, P) un espace probabilisé lié à cette expérience. Dans de nombreuses situations, on associe à chaque résultat $\omega \in \Omega$ un nombre réel noté $X(\omega)$; on construit ainsi une application $X : \Omega \rightarrow \mathbb{R}$. Historiquement, ε était un jeu et X représentait le gain du joueur.

Exemple: Jeu de dé

Un joueur lance un dé équilibré à 6 faces numérotées de 1 à 6, et on observe le numéro obtenu.

- ▶ Si le joueur obtient 1, 3 ou 5, il gagne 1 euro.
- ▶ S'il obtient 2 ou 4, il gagne 5 euros.
- ▶ S'il obtient 6, il perd 10 euros.

Analyse

- ▶ ε : “lancer d'un dé équilibré”.
- ▶ $\Omega = \{1, 2, 3, 4, 5, 6\}$.
- ▶ $\mathcal{A} = \mathcal{P}(\Omega)$.
- ▶ P l'équiprobabilité sur (Ω, \mathcal{A}) .

Soit X l'application de Ω dans \mathbb{R} qui à tout $\omega \in \Omega$ associe le gain correspondant.

On a donc

- ▶ $X(1) = X(3) = X(5) = 1$
- ▶ $X(2) = X(4) = 5$
- ▶ $X(6) = -10$

On dit que X est une **variable aléatoire** sur Ω .

On peut s'intéresser à la probabilité de gagner 1 euro:

$$\Rightarrow X(\omega) = 1.$$

- ▶ ce qui se réalise si et seulement si $\omega \in \{1, 3, 5\}$.
- ▶ La probabilité cherchée est donc $P(\{1, 3, 5\}) = 1/2$.
- ▶ On écrira aussi $P(X = 1) = 1/2$.

On pourra donc considérer l'événement:

$$\{X = 1\} = \{\omega \in \Omega / X(\omega) = 1\} = \{\omega \in \Omega / X(\omega) \in \{1\}\} = X^{-1}(\{1\}) = \{1, 3, 5\}.$$

On aura du même:

- ▶ $P(X = 5) = 1/3$.
- ▶ $P(X = -10) = 1/6$.

On peut présenter les probabilités précédentes dans un tableau:

x_i	-10	1	5
$p_i = P(X = x_i)$	1/6	1/2	1/3

Cela revient à considérer un nouvel ensemble d'événements élémentaires:

$$\Omega_X = X(\Omega) = \{-10, 1, 5\}$$

et à munir cet ensemble de la probabilité P_X définie par le tableau des $P(X = x_i)$ ci dessus. Cette nouvelle probabilité s'appelle **loi de la variable aléatoire** X .

Remarquer que

$$P\left(\bigcup_{x_i \in \Omega_X} \{X = x_i\}\right) = \sum_{x_i \in \Omega_X} P(X = x_i) = 1$$

Variables aléatoires discrètes

Définition

On dit qu'une variable aléatoire réelle (v.a.r.) X est **discrète** (v.a.r.d.) si l'ensemble des valeurs que prend X est fini ou infini dénombrable.

Si on suppose $X(\Omega)$ l'ensemble des valeurs de X qui admet un plus petit élément x_1 . Alors la v.a.r.d. X est entièrement **définie** par:

- ▶ L'ensemble $X(\Omega)$ des valeurs prises par X , rangées par ordre croissant: $X(\Omega) = \{x_1, x_2, \dots, x_i, \dots\}$ avec $x_1 \leq x_2 \leq \dots \leq x_i \leq \dots$
- ▶ La **loi de probabilité** définie sur $X(\Omega)$ par

$$p_i = p(x_i) = P(X = x_i) \quad \forall i = 1, 2, \dots$$

Remarques:

- ▶ $B \subset \mathbb{R}$, $P(X \in B) = \sum_{i/x_i \in B} p(x_i)$.
- ▶ $P(a < X \leq b) = \sum_{i/a < x_i \leq b} p(x_i)$.
- ▶ $p(x_i) \geq 0$ et $\sum_{i=1}^{\infty} p(x_i) = 1$.
- ▶ Si X ne prend qu'un petit nombre de valeurs, cette loi est généralement présentée dans un tableau.

Définition

On appelle fonction de répartition de la v.a. X , qu'on note $F(a)$ de la v.a.r.d. X , ou $F_X(a)$, la fonction définie pour tout réel a , $-\infty < a < \infty$, par

$$F(a) = P(X \leq a) = \sum_{i/x_i \leq a} P(X = x_i)$$

Cette valeur représente la probabilité de toutes les réalisations inférieures ou égales au réel a .

Propriétés

1. C'est une fonction en escalier (constante par morceaux).
2. $F(a) \leq 1$ car c'est une probabilité.
3. $F(a)$ est continue à droite.
4. $\lim_{a \rightarrow -\infty} F(a) = 0$ et $\lim_{a \rightarrow \infty} F(a) = 1$

La fonction de répartition caractérise la loi de X , autrement dit: $F_X = F_Y$ si et seulement si les variables aléatoires X et Y ont la même loi de probabilité.

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

$$P(a < X \leq b) = F(b) - F(a) \quad \text{pour tout } a < b$$

On peut mieux s'en rendre compte en écrivant $\{X \leq b\}$ comme union des deux événements incompatibles $\{X \leq a\}$ et $\{a < X \leq b\}$, soit

$$\{X \leq b\} = \{X \leq a\} \cup \{a < X \leq b\}$$

et ainsi

$$P(X \leq b) = P(X \leq a) + P(a < X \leq b)$$

ce qui établit l'égalité ci dessus.

Remarque

On peut déduire de F les probabilités individuelles par:

$$p_i = P(X = x_i) = F(x_i) - F(x_{i-1}) \quad \text{pour } 1 \leq i \leq n$$

Exemple

On joue trois fois à pile ou face \Rightarrow

- ▶ $\Omega = \{P, F\}^3$.
- ▶ $\text{card}(\Omega) = |\Omega| = 2^3 = 8$.

Soit X la variable aléatoire "nombre de pile obtenus" $\Rightarrow X(\Omega) = \{0, 1, 2, 3\}$.

- ▶ Calculons par exemple $P(X = 1)$.
 - ▶ $X^{-1}(1) = \{(P, F, F), (F, P, F), (F, F, P)\}$.
- $\Rightarrow P(X = 1) = \frac{3}{8}$

En procédant de la même façon, on obtient la loi de probabilité de X :

k	0	1	2	3
$P(X = k)$	1/8	3/8	3/8	1/8

La fonction de répartition de X est donc donnée par:

$$F(x) = \begin{cases} 0 & \text{si } x < 0 \\ 1/8 & \text{si } 0 \leq x < 1 \\ 1/2 & \text{si } 1 \leq x < 2 \\ 7/8 & \text{si } 2 \leq x < 3 \\ 1 & \text{si } x \geq 3 \end{cases}$$

On peut présenter la fonction de répartition dans le tableau de la loi de probabilité de X :

k	0	1	2	3
$P(X = k)$	1/8	3/8	3/8	1/8
$F_X(x)$	1/8	1/2	7/8	1

Le graphe de cette dernière est représentée dans la figure suivante:

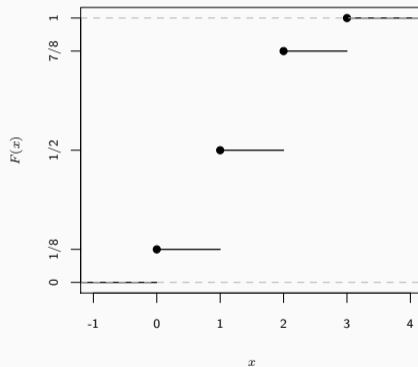


Figure 1: Fonction de répartition

Une autre représentation de la fonction de répartition:

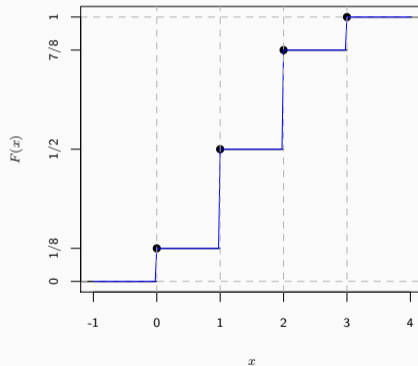


Figure 2: Fonction de répartition

Définition

Soit A un événement quelconque. On appelle *variable aléatoire indicatrice* de cet événement A , la variable aléatoire notée $X = \mathbb{1}_A$ et définie par:

$$X(\omega) = \begin{cases} 1 & \text{si } \omega \in A \\ 0 & \text{si } \omega \in \bar{A} \end{cases}$$

Ainsi:

- ▶ $P(X = 1) = P(A) = p$
- ▶ $P(X = 0) = P(\bar{A}) = 1 - p$

La fonction de répartition de X est donc donnée par:

$$F(x) = \begin{cases} 0 & \text{si } x < 0 \\ 1 - p & \text{si } 0 \leq x < 1 \\ 1 & \text{si } x \geq 1 \end{cases}$$

Exemple

- ▶ Soit U une urne contenant 2 boules blanches et 3 boules noires.
- ▶ On tire une boule au hasard.
- ▶ Soit A : “obtenir une boule blanche”.
- ▶ Soit X la variable indicatrice de A .

Déterminez la loi de probabilité de X ainsi que sa Fonction de répartition.

La loi de probabilité de X est

k	0	1
$P(X = k)$	$\frac{3}{5}$	$\frac{2}{5}$

et sa fonction de répartition est:

$$F(x) = \begin{cases} 0 & \text{si } x < 0 \\ 3/5 & \text{si } 0 \leq x < 1 \\ 1 & \text{si } x \geq 1 \end{cases}$$

Moments d'une variable aléatoire discrète

Définition

Pour une variable aléatoire discrète X de loi de probabilité $p(\cdot)$, on définit l'espérance de X , notée $E(X)$, par l'expression

$$E(X) = \sum_{i \in \mathbb{N}} x_i p(x_i)$$

En termes concrets, l'espérance de X est la moyenne pondérée des valeurs que X peut prendre, les poids étant les probabilités que ces valeurs soient prises.

Exemples

1. Dans l'exemple où on joue 3 fois à pile ou face. L'espérance de X = "nombre de pile obtenus" est égal à:

$$E(X) = 0 \times \frac{1}{8} + 1 \times \frac{3}{8} + 2 \times \frac{3}{8} + 3 \times \frac{1}{8} = 1.5$$

2. Pour la variable aléatoire indicatrice de A :

$$E(X) = 0 \times P(X = 0) + 1 \times P(X = 1) = P(A) = p$$

Ceci signifie que l'espérance de la variable indicatrice pour l'événement A est égale à la probabilité que A se produise.

Théorème

Si X est une variable aléatoire discrète pouvant prendre ses valeurs parmi les valeurs x_i , $i \geq 1$, avec des probabilités respectives $p(x_i)$, alors pour toute fonction réelle g on a

$$E(g(X)) = \sum_i g(x_i)p(x_i)$$

Exemple

Soit X une variable aléatoire qui prend une des trois valeurs $\{-1, 0, 1\}$ avec les probabilités respectives

$$P(X = -1) = 0.2 \quad P(X = 0) = 0.5 \quad P(X = 1) = 0.3$$

Calculer $E(X^2)$.

Solution

Première approche: Soit $Y = X^2$. La distribution de Y est donnée par

$$P(Y = 1) = P(X = -1) + P(X = 1) = 0.5$$

$$P(Y = 0) = P(X = 0) = 0.5$$

Donc

$$E(X^2) = E(Y) = 1(0.5) + 0(0.5) = 0.5$$

Deuxième approche: En utilisant le théorème

$$\begin{aligned} E(X^2) &= (-1)^2(0.2) + 0^2(0.5) + 1^2(0.3) \\ &= 1(0.2 + 0.3) + 0(0.5) = 0.5 \end{aligned}$$

Remarque

$$0.5 = E(X^2) \neq (E(X))^2 = 0.01$$

Propriétés

1. $E(X + a) = E(X) + a$, $a \in \mathbb{R}$

résultat qui se déduit de:

$$\sum_i p_i(x_i + a) = \sum_i p_i x_i + \sum_i a p_i = \sum_i p_i x_i + a \sum_i p_i = \sum_i p_i x_i + a$$

2. $E(aX) = aE(X)$, $a \in \mathbb{R}$

il suffit d'écrire:

$$\sum_i p_i a x_i = a \sum_i p_i x_i$$

3. $E(X + Y) = E(X) + E(Y)$, X et Y étant deux variables aléatoire.

On peut résumer ces trois propriétés en disant que l'espérance mathématique est linéaire:

$$E(\lambda X + \mu Y) = \lambda E(X) + \mu E(Y), \quad \forall \lambda \in \mathbb{R}, \forall \mu \in \mathbb{R}.$$

Définition

Il s'agit d'un indicateur mesurant la dispersion des valeurs x_i que peut prendre la v.a. X et son espérance $E(X)$. On appelle **variance** de X , que l'on note $V(X)$, la quantité, lorsqu'elle existe,

$$V(X) = E[(X - E(X))^2]$$

C'est l'espérance mathématique du carré de la v.a. centrée $X - E(X)$.

Remarque

On peut établir une autre formule pour le calcul de $V(X)$:

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

Or:

$$\begin{aligned} V(X) &= E[X^2 - 2XE(X) + E^2(X)] \\ &= E(X^2) - E[2XE(X)] + E[E^2(X)] \\ &= E(X^2) - 2E^2(X) + E^2(X) \\ &= E(X^2) - E^2(X) \end{aligned}$$

Exemple

On cherche $V(X)$ où X est le nombre obtenu lors du jet d'un dé équilibré.

On a vu précédemment que $E(X) = \frac{7}{2}$.

De plus,

$$\begin{aligned} E(X^2) &= \sum_i x_i^2 p(x_i) \\ &= 1^2 \left(\frac{1}{6}\right) + 2^2 \left(\frac{1}{6}\right) + 3^2 \left(\frac{1}{6}\right) + 4^2 \left(\frac{1}{6}\right) + 5^2 \left(\frac{1}{6}\right) + 6^2 \left(\frac{1}{6}\right) \\ &= \left(\frac{1}{6}\right)(91) = \frac{91}{6} \end{aligned}$$

Et donc

$$\begin{aligned} V(X) &= E(X^2) - E^2(X) \\ &= \frac{91}{6} - \left(\frac{7}{2}\right)^2 = \frac{35}{12} \end{aligned}$$

Propriétés

1. $V(X) \geq 0$

2. $\forall a \in \mathbb{R}, \quad V(X + a) = V(X)$
en effet:

$$\begin{aligned}V(X + a) &= E \left[[X + a - E(X + a)]^2 \right] \\&= E \left[[X + a - E(X) - a]^2 \right] \\&= E \left[[X - E(X)]^2 \right] = V(X)\end{aligned}$$

3. $\forall a \in \mathbb{R}, \quad V(aX) = a^2 V(X)$
en effet:

$$\begin{aligned}V(aX) &= E \left[[aX - E(aX)]^2 \right] \\&= E \left[[aX - aE(X)]^2 \right] \\&= E \left[a^2 [X - E(X)]^2 \right] \\&= a^2 \left[E [X - E(X)]^2 \right] = a^2 V(X)\end{aligned}$$

Définition

La racine carrée de $V(X)$ est appelée l'**écart-type** de X , qui se note σ_X ¹. On a

$$\sigma_X = \sqrt{V(X)}$$

σ_X s'exprime dans les mêmes unités de mesure que la variable aléatoire X .

- ▶ L'écart type sert à mesurer la dispersion d'un ensemble de données.
- ▶ Plus il est faible, plus les valeurs sont regroupées autour de la moyenne.
- ▶ **Exemple:** La répartition des notes d'une classe. Plus l'écart type est faible, plus la classe est homogène.
- ▶ L'espérance et l'écart-type sont reliés par l'**inégalité de Bienaymé-Tchebychev**.

¹ou $\sigma(X)$, ou tout simplement σ .

Théorème

Soit X une variable aléatoire d'espérance μ et de variance σ^2 . Pour tout $\varepsilon > 0$, on a l'inégalité suivante:

$$P(|X - E(X)| \geq \varepsilon) \leq \frac{\sigma^2}{\varepsilon^2}$$

Remarque

On peut l'écrire autrement. Soit $k = \varepsilon/\sigma$.

$$P(|X - E(X)| \geq k\sigma) \leq \frac{1}{k^2}$$

Importance

Cette inégalité relie la probabilité pour X de s'écarter de sa moyenne $E(X)$, à sa variance qui est justement un indicateur de dispersion autour de la moyenne de la loi. Elle montre quantitativement que "plus l'écart type est faible, plus la probabilité de s'écarter de la moyenne est faible".

Inégalité de Markov

Soit X une variable aléatoire à valeur non négatives. Pour tout réel $a > 0$

$$P(X > a) \leq \frac{E(X)}{a}$$

Définition

On appelle *moment non centré* d'ordre $r \in \mathbb{N}^*$ de X la quantité, lorsqu'elle existe:

$$m_r(X) = \sum_{i \in \mathbb{N}} x_i^r p(x_i) = E(X^r).$$

Définition

Le *moment centré* d'ordre $r \in \mathbb{N}^*$ est la quantité, lorsqu'elle existe:

$$\mu_r(X) = \sum_{i \in \mathbb{N}} p_i [x_i - E(X)]^r = E[X - E(X)]^r.$$

Remarque

Les premiers moments sont:

- ▶ $m_1(X) = E(X)$, $\mu_1(X) = 0$.
- ▶ $\mu_2(X) = V(X) = m_2(X) - m_1^2(X)$.

Couple de variables aléatoires discrètes

Nous avons traité jusqu'ici des variables isolées. Or, il est souvent nécessaire de considérer des événements relatifs à deux variables simultanément, ou même à plus de deux variables.

Définition

Soit X et Y deux variables aléatoires réelles discrètes, définies sur un espace probabilisé (Ω, \mathcal{A}, P) et que $X(\Omega) = \{x_1, x_2, \dots, x_l\}$ et $Y(\Omega) = \{y_1, y_2, \dots, y_k\}$, l et $k \in \mathbb{N}$.

La **loi du couple** (X, Y) est entièrement définie par les probabilités:

$$p_{ij} = P(X = x_i; Y = y_j) = P(\{X = x_i\} \cap \{Y = y_j\})$$

On a

$$p_{ij} \geq 0 \quad \text{et} \quad \sum_{i=1}^l \sum_{j=1}^k p_{ij} = 1$$

Le couple (X, Y) s'appelle *variable aléatoire à deux dimensions* et peut prendre $l \times k$ valeurs.

Les probabilités p_{ij} peuvent être présentées dans un tableau à deux dimensions qu'on appelle table de probabilité conjointe:

Table 1: Table de probabilité conjointe

$X \setminus Y$	y_1	y_2	\dots	y_j	\dots	y_k
x_1	p_{11}	p_{12}		p_{1j}		p_{1k}
x_2	p_{21}	p_{22}		p_{2j}		p_{2k}
\vdots						
x_i	p_{i1}	p_{i2}		p_{ij}		p_{ik}
\vdots						
x_l	p_{l1}	p_{l2}		p_{lj}		p_{lk}

A la première ligne figure l'ensemble des valeurs de Y et à la première colonne figure l'ensemble des valeurs de X . La probabilité $p_{ij} = P(X = x_i; Y = y_j)$ est à l'intersection de la i^e et de la j^e colonne.

Exemple

On tire au hasard 3 boules d'une urne contenant 3 boules rouges, 4 blanches et 5 noires. X et Y désignent respectivement le nombre de boules rouges et celui de boules blanches tirées. Déterminer la loi de probabilité conjointe du couple (X, Y) .

Solution

- ▶ ε : "tirer 3 boules d'une urne contenant 12 boules".
- ▶ $|\Omega| = C_{12}^3 = 220$.
- ▶ $X(\Omega) = \{0, 1, 2, 3\}$ et $Y(\Omega) = \{0, 1, 2, 3\}$.
- ▶ $p(X = 0, Y = 0) = p(0, 0) = C_5^3 / C_{12}^3 = \frac{10}{220}$.
- ▶ $p(0, 1) = C_4^1 C_5^2 / C_{12}^3 = \frac{40}{220}$.
- ▶ $p(1, 0) = C_3^1 C_5^2 / C_{12}^3 = \frac{30}{220}$.

Exemple

On tire au hasard 3 boules d'une urne contenant 3 boules rouges, 4 blanches et 5 noires. X et Y désignent respectivement le nombre de boules rouges et celui de boules blanches tirées. Déterminer la loi de probabilité conjointe du couple (X, Y) .

Solution

Table 2: Table de probabilité conjointe

$X \backslash Y$	0	1	2	3
0	$\frac{10}{220}$	$\frac{40}{220}$	$\frac{30}{220}$	$\frac{4}{220}$
1	$\frac{30}{220}$	$\frac{60}{220}$	$\frac{18}{220}$	0
2	$\frac{15}{220}$	$\frac{12}{220}$	0	0
3	$\frac{1}{220}$	0	0	0

Lorsqu'on connaît la loi conjointe des variables aléatoires X et Y , on peut aussi s'intéresser à la loi de probabilité de X seule et de Y seule. Ce sont les lois de probabilité marginales.

- Loi marginale de X :

$$p_{i.} = P(X = x_i) = P[\{X = x_i\} \cap \Omega] = \sum_{j=1}^k p_{ij} \quad \forall i = 1, 2, \dots, l$$

- Loi marginale de Y :

$$p_{.j} = P(Y = y_j) = P[\Omega \cap \{Y = y_j\}] = \sum_{i=1}^l p_{ij} \quad \forall j = 1, 2, \dots, k$$

On peut calculer les lois marginales directement depuis la table de la loi conjointe.

Table 3: Table de probabilité conjointe avec les lois marginales

$X \setminus Y$	y_1	y_2	\dots	y_j	\dots	y_k	Marginale de X
x_1	p_{11}	p_{12}		p_{1j}		p_{1k}	$p_{1.}$
x_2	p_{21}	p_{22}		p_{2j}		p_{2k}	$p_{2.}$
\vdots							
x_i	p_{i1}	p_{i2}		p_{ij}		p_{ik}	$p_{i.}$
\vdots							
x_l	p_{l1}	p_{l2}		p_{lj}		p_{lk}	$p_{l.}$
Marginale de Y	$p_{.1}$	$p_{.2}$		$p_{.j}$		$p_{.k}$	1

Exemple

On tire au hasard 3 boules d'une urne contenant 3 boules rouges, 4 blanches et 5 noires. X et Y désignent respectivement le nombre de boules rouges et celui de boules blanches tirées. Déterminer les lois marginales de X et Y .

Solution

Table 4: Table de probabilité conjointe

$X \setminus Y$	0	1	2	3	$p_{i.} = P(X = x_i)$
0	$\frac{10}{220}$	$\frac{40}{220}$	$\frac{30}{220}$	$\frac{4}{220}$	$\frac{84}{220}$
1	$\frac{30}{220}$	$\frac{60}{220}$	$\frac{18}{220}$	0	$\frac{108}{220}$
2	$\frac{15}{220}$	$\frac{12}{220}$	0	0	$\frac{27}{220}$
3	$\frac{1}{220}$	0	0	0	$\frac{1}{220}$
$p_{.j} = P(Y = y_j)$	$\frac{56}{220}$	$\frac{112}{220}$	$\frac{48}{220}$	$\frac{4}{220}$	1

Définition

Pour chaque valeur y_j de Y telle que $p_{.j} = P(Y = y_j) \neq 0$ on peut définir la loi conditionnelle de X sachant $Y = y_j$ par

$$p_{i/j} = P(X = x_i / Y = y_j) = \frac{P(X = x_i; Y = y_j)}{P(Y = y_j)} = \frac{p_{ij}}{p_{.j}} \quad \forall i = 1, 2, \dots, l$$

De même on définit la loi de Y sachant $X = x_i$ par

$$p_{j/i} = P(Y = y_j / X = x_i) = \frac{P(X = x_i; Y = y_j)}{P(X = x_i)} = \frac{p_{ij}}{p_{i.}} \quad \forall j = 1, 2, \dots, k$$

Définition

On dit que deux v.a.r.d sont indépendantes si et seulement si

$$P(X = x_i; Y = y_j) = P(X = x_i)P(Y = y_j) \quad \forall i = 1, 2, \dots, l \text{ et } j = 1, 2, \dots, k$$

On montre que

$$P(\{X \in A\} \cap \{Y \in B\}) = P(\{X \in A\})P(\{Y \in B\}) \quad \forall A \text{ et } B \in \mathcal{A}$$

Propriétés

Soit deux v.a.r.d. X et Y ,

1. $E(X + Y) = E(X) + E(Y)$
2. Si X et Y sont indépendantes alors $E(XY) = E(X)E(Y)$. Mais la réciproque n'est pas toujours vraie.

Définition

Soit X et Y deux v.a.r.d. On appelle **covariance** de X et de Y la valeur si elle existe de

$$\text{Cov}(X, Y) = E[(X - E(X))(Y - E(Y))] = \sum_i \sum_j (x_i - E(X))(y_j - E(Y))p_{ij}$$

qu'on peut calculer en utilisant la formule suivante

$$\text{Cov}(X, Y) = E(XY) - E(X)E(Y)$$

Propriétés

- ▶ $\text{Cov}(X, Y) = \text{Cov}(Y, X)$
- ▶ $\text{Cov}(aX_1 + bX_2, Y) = a\text{Cov}(X_1, Y) + b\text{Cov}(X_2, Y)$
- ▶ $V(X + Y) = V(X) + V(Y) + 2\text{Cov}(X, Y)$
- ▶ Si X et Y sont indépendantes alors
 - $\text{Cov}(X, Y) = 0$ (la réciproque n'est pas vraie)
 - $V(X + Y) = V(X) + V(Y)$ (la réciproque n'est pas vraie)

Définition

On appelle coefficient de corrélation linéaire de X et de Y la valeur définie par

$$\rho = \rho(X, Y) = \frac{\text{Cov}(X, Y)}{\sqrt{V(X)V(Y)}} = \frac{\text{Cov}(X, Y)}{\sigma_X \sigma_Y}$$

On peut montrer que

$$-1 \leq \rho(X, Y) \leq 1$$

Interprétation de ρ

- ▶ Le coefficient de corrélation est une mesure du degré de linéarité entre X et Y .
- ▶ Les valeurs de ρ proches de 1 ou -1 indiquent une linéarité quasiment rigoureuse entre X et Y .
- ▶ Les valeurs de ρ proche de 0 indiquent une absence de toute relation linéaire.
- ▶ Lorsque $\rho(X, Y)$ est positif, Y a tendance à augmenter si X en fait autant.
- ▶ Lorsque $\rho(X, Y) < 0$, Y a tendance à diminuer si X augmente.
- ▶ Si $\rho(X, Y) = 0$, on dit que ces deux statistiques sont non corrélées.

Loi Uniforme Discrète $\mathcal{U}(n)$

Définition

Une distribution de probabilité suit une loi uniforme lorsque toutes les valeurs prises par la variable aléatoire sont équiprobables. Si n est le nombre de valeurs différentes prises par la variable aléatoire alors on a :

$$P(X = x_i) = \frac{1}{n} \quad \forall i \in \{1, \dots, n\}$$

On dit $X \sim \mathcal{U}(n)$.

Exemple

La distribution des chiffres obtenus au lancer de dé (si ce dernier est non pipé) suit une loi uniforme dont la loi de probabilité est la suivante :

x_i	1	2	3	4	5	6
$P(X = x_i)$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$	$\frac{1}{6}$

Cas particulier

Dans le cas particulier d'une loi uniforme discrète où chaque valeur de la variable aléatoire X correspond à son rang, i.e. $x_i = i \forall i \in \{1, \dots, n\}$, on a:

$$E(X) = \frac{n+1}{2} \quad \text{et} \quad V(X) = \frac{n^2-1}{12}$$

Démonstration

La démonstration de ces résultats est établie en utilisant les égalités:

$$\sum_{i=1}^n i = \frac{n(n+1)}{2} \quad \text{et} \quad \sum_{i=1}^n i^2 = \frac{n(n+1)(2n+1)}{6}.$$

Vous avez la démonstration de ces égalités dans l'Annexe.

Exemple

L'exemple du lancer du dé: on peut calculer directement les moments de X :

$$E(X) = \frac{6+1}{2} = 3.5 \quad \text{et} \quad V(X) = \frac{6^2-1}{12} = \frac{35}{12} \simeq 2.92.$$

Loi de Bernoulli $\mathcal{B}(p)$

Variable Indicatrice

Soit A un événement quelconque; on appelle v.a. indicatrice de l'événement A , la v.a. définie par $X = \mathbb{1}_A$, c'est à dire:

$$X(\omega) = \mathbb{1}_A(\omega) = \begin{cases} 0 & \text{si } \omega \in \bar{A} \\ 1 & \text{si } \omega \in A \end{cases}$$

Ainsi $X(\Omega) = \{0, 1\}$ avec:

$$P(X = 1) = P\{\omega \in \Omega / X(\omega) = 1\} = P(A) = p$$

$$P(X = 0) = P\{\omega \in \Omega / X(\omega) = 0\} = P(\bar{A}) = 1 - P(A) = q$$

$$\text{avec } p + q = 1$$

Définition

On dit que X suit une loi de Bernoulli de paramètre $p = P(A)$, ce qu'on écrit symboliquement $X \sim \mathcal{B}(p)$. Une distribution de Bernoulli est associée à la notion "épreuve de Bernoulli", qui est une épreuve aléatoire à deux issues: succès ($X = 1$) et échec ($X = 0$).

Fonction de répartition de loi de Bernoulli

$$F(x) = \begin{cases} 0 & \text{si } x < 0 \\ 1 - p & \text{si } 0 \leq x < 1 \\ 1 & \text{si } x \geq 1. \end{cases}$$

Espérance de loi de Bernoulli

$$E(X) = 1 \times P(A) + 0 \times P(\bar{A}) = P(A) = p$$

Variance de loi de Bernoulli

$$V(X) = E(X^2) - E^2(X) = p - p^2 = p(1 - p) = pq$$

car

$$E(X^2) = 1^2 \times P(A) + 0^2 \times P(\bar{A}) = P(A) = p$$

Loi Binomiale $\mathcal{B}(n, p)$

- ▶ Décrite pour la première fois par *Isaac Newton* en 1676 et démontrée pour la première fois par le mathématicien suisse *Jacob Bernoulli* en 1713.
- ▶ La loi binomiale est l'une des distributions de probabilité les plus fréquemment rencontrées en statistique appliquée.
- ▶ On exécute n épreuves **indépendantes** de **Bernoulli**.
- ▶ Chaque épreuve a p pour probabilité de succès et $1 - p$ pour probabilité d'échec.

A A \bar{A} A \bar{A} ... \bar{A} A A

S S E S E ... E S S

- ▶ $X =$ le **nombre de succès** sur l'ensemble des n épreuves.
- ▶ X dépend de deux paramètres n et p .

S S E S E ... E S S

- ▶ X = le nombre de succès sur l'ensemble des n épreuves.
- ▶ $X(\Omega) = \{0, 1, \dots, n\}$

$$P(X = k) = \binom{n}{k} p^k (1-p)^{n-k} \quad 0 \leq k \leq n$$

- ▶ $\binom{n}{k}$ est le nombre d'échantillons de taille n comportant exactement k succès, de probabilité p^k , indépendamment de l'ordre, et donc $n - k$ échecs, de probabilité $(1-p)^{n-k}$.
- ▶ On écrit $X \sim \mathcal{B}(n, p)$.

Remarque

Une variable de Bernoulli n'est donc qu'une variable binomiale de paramètres $(1, p)$.

$$X \sim \mathcal{B}(p) \iff X \sim \mathcal{B}(1, p)$$

Triangle de Pascal

$$\binom{n}{k} = \binom{n-1}{k-1} + \binom{n-1}{k} \quad \forall n \geq 1 \text{ et } 1 \leq k \leq n-1$$

$n = 0$										1										
$n = 1$										1	1									
$n = 2$										1	2	1								
$n = 3$										1	3	3	1							
$n = 4$										1	4	6	4	1						
$n = 5$										1	5	10	10	5	1					
$n = 6$										1	6	15	20	15	6	1				
										0	1	2	3	4	5	6				

Binôme de Newton

$$(x + y)^n = \sum_{k=0}^n \binom{n}{k} x^{n-k} y^k$$

Cette formule permet de vérifier que la loi **Binomiale** est une loi de probabilité:

$$\sum_{k=0}^n P(X = k) = \sum_{k=0}^n \binom{n}{k} p^k (1-p)^{n-k} = [p + (1-p)]^n = 1$$

Exemple

On jette **cinq pièces équilibrées**. Les résultats sont supposés indépendants. Donner la loi de probabilité de la variable X qui compte **le nombre de piles obtenus**.

Solution

- ▶ $X =$ nombre de piles (*succès*).
- ▶ $n = 5$.
- ▶ $p = 1/2$.
- ▶ $X \sim \mathcal{B}(5, \frac{1}{2})$.
- ▶ $X(\Omega) = \{0, 1, \dots, 5\}$
- ▶ $P(X = 0) = \binom{5}{0} \left(\frac{1}{2}\right)^0 \left(1 - \frac{1}{2}\right)^{5-0} = \frac{1}{32}$
- ▶ $P(X = 1) = \binom{5}{1} \left(\frac{1}{2}\right)^1 \left(1 - \frac{1}{2}\right)^4 = \frac{5}{32}$
- ▶ $P(X = 2) = \binom{5}{2} \left(\frac{1}{2}\right)^2 \left(1 - \frac{1}{2}\right)^3 = \frac{10}{32}$
- ▶ $P(X = 3) = \binom{5}{3} \left(\frac{1}{2}\right)^3 \left(1 - \frac{1}{2}\right)^2 = \frac{10}{32}$
- ▶ $P(X = 4) = \binom{5}{4} \left(\frac{1}{2}\right)^4 \left(1 - \frac{1}{2}\right)^1 = \frac{5}{32}$
- ▶ $P(X = 5) = \binom{5}{5} \left(\frac{1}{2}\right)^5 \left(1 - \frac{1}{2}\right)^0 = \frac{1}{32}$

Si $X \sim \mathcal{B}(n, p)$ alors $E(X) = np$ et $V(X) = np(1 - p)$

Démonstration

Première approche: On associe à chaque épreuve i , $1 \leq i \leq n$, une v.a. de Bernoulli.

$$\mathbb{1}_A = X_i = \begin{cases} 1 & \text{si } A \text{ est réalisé} \\ 0 & \text{si } \bar{A} \text{ est réalisé} \end{cases}$$

On peut écrire alors: $X = \sum_{i=1}^n X_i = X_1 + X_2 + \dots + X_n$

Donc

$$E(X) = E\left(\sum_{i=1}^n X_i\right) = \sum_{i=1}^n E(X_i) = np$$

et

$$V(X) = V\left(\sum_{i=1}^n X_i\right) = \sum_{i=1}^n V(X_i) = np(1 - p)$$

car les v.a. X_i sont indépendantes.

Deuxième approche: Calcul direct.

- ▶ $E(X) = \sum_{k=0}^n k \binom{n}{k} p^k (1-p)^{n-k} = \dots = np$
- ▶ $V(X) = E(X^2) - E^2(X)$
- ▶ Pour obtenir $E(X^2)$ par un procédé de calcul identique, on passe par l'intermédiaire du moment factoriel $E[X(X-1)]$.
- ▶ $V(X) = E(X^2) - E^2(X) = E[X(X-1)] + E(X) - E^2(X)$
- ▶ $E[X(X-1)] = \sum_{k=0}^n k(k-1) \frac{n!}{k!(n-k)!} p^k (1-p)^{n-k} = \dots = n(n-1)p^2$
- ▶ $V(X) = n(n-1)p^2 + np - (np)^2 = np(1-p)$

Exemple

Le nombre de résultats pile apparus au cours de n jets d'une pièce de monnaie suit une loi binomiale $\mathcal{B}(n, 1/2)$:

$$P(X = k) = \binom{n}{k} \left(\frac{1}{2}\right)^k \left(\frac{1}{2}\right)^{n-k} = \frac{\binom{n}{k}}{2^n}, \quad 0 \leq k \leq n$$

avec $E(X) = n/2$ et $V(X) = n/4$.

Exemple

Le nombre N de boules rouges apparues au cours de n tirages avec remise dans une urne contenant deux rouges, trois vertes et une noire suit une loi binomiale $\mathcal{B}(n, 1/3)$:

$$P(N = k) = \binom{n}{k} \left(\frac{1}{3}\right)^k \left(\frac{2}{3}\right)^{n-k} = \binom{n}{k} \frac{2^{n-k}}{3^n}, \quad 0 \leq k \leq n$$

avec $E(X) = n/3$ et $V(X) = 2n/9$.

Remarque

Si $X_1 \sim \mathcal{B}(n_1, p)$ et $X_2 \sim \mathcal{B}(n_2, p)$, les v.a. X_1 et X_2 étant **indépendantes**, alors $X_1 + X_2 \sim \mathcal{B}(n_1 + n_2, p)$. Ceci résulte de la définition d'une loi binomiale puisqu'on totalise ici le résultat de $n_1 + n_2$ épreuves indépendantes.

Loi de Poisson $\mathcal{P}(\lambda)$

Définition

Une v.a. X suit une loi de Poisson de paramètre $\lambda > 0$ si c'est une variable à valeurs entières, $X(\Omega) = \mathbb{N}$, donc avec une infinité de valeurs possibles, de probabilité:

$$P(X = k) = e^{-\lambda} \frac{\lambda^k}{k!}, \quad k \in \mathbb{N}$$

Cette loi ne dépend qu'un seul paramètre réel positif λ , avec l'écriture symbolique $X \sim \mathcal{P}(\lambda)$.

Remarque

$$e^x = \sum_{i=0}^{+\infty} \frac{x^i}{i!}$$

Donc

$$\sum_{k=0}^{\infty} P(X = k) = \sum_{k=0}^{\infty} e^{-\lambda} \frac{\lambda^k}{k!} = e^{-\lambda} \sum_{k=0}^{\infty} \frac{\lambda^k}{k!} = e^{-\lambda} e^{\lambda} = 1$$

Si $X \sim \mathcal{P}(\lambda)$ alors $E(X) = \lambda$ et $V(X) = \lambda$

Espérance de loi de Poisson

$$\begin{aligned} E(X) &= \sum_{k=0}^{\infty} kP(X = k) \\ &= \dots \\ &= \lambda. \end{aligned}$$

Variance de loi de Poisson

- ▶ On calcule d'abord $E(X^2) = \sum_{k=0}^{\infty} k^2 P(X = k) = \dots = \lambda(\lambda + 1)$.

Ensuite

$$V(X) = \lambda(\lambda + 1) - \lambda^2 = \lambda$$

Exemple

- ▶ X = nombre de micro-ordinateurs vendus chaque jour dans un magasin.
- ▶ On suppose $X \sim \mathcal{P}(5)$.
- ▶ La probabilité associée à la vente de 5 micro-ordinateurs est

$$P(X = 5) = e^{-5} \frac{5^5}{5!} = e^{-5} \simeq 0.1755$$

- ▶ La probabilité de vendre au moins 2 micro-ordinateurs est

$$P(X \geq 2) = 1 - \left(e^{-5} \frac{5^0}{0!} + e^{-5} \frac{5^1}{1!} \right) \simeq 0.9596$$

- ▶ Le nombre moyen de micro-ordinateurs vendus chaque jour dans le magasin est égal à 5 puisque $E(X) = \lambda = 5$.

Propriétés

Si X et Y sont deux variables **indépendantes** suivant des lois de Poisson, $X \sim \mathcal{P}(\lambda)$ et $Y \sim \mathcal{P}(\mu)$, alors leur somme suit aussi une loi de Poisson: $X + Y \sim \mathcal{P}(\lambda + \mu)$.

Si $n \rightarrow \infty$ et $p \rightarrow 0$ alors $X : \mathcal{B}(n, p) \sim \mathcal{P}(\lambda)$.

Remarque

Une bonne approximation est obtenue si $n \geq 50$ et $np \leq 5$.

Dans ce contexte, la loi de Poisson est souvent utilisée pour modéliser le nombre de succès lorsqu'on répète un très grand nombre de fois une expérience ayant une chance très faible de réussir par une loi de Poisson.

Applications de la loi de Poisson

- ▶ Le nombre d'individus dépassant l'âge de 100 ans dans une communauté.
- ▶ Le nombre de faux numéros téléphoniques composés en un jour.
- ▶ Le nombre de clients pénétrant dans un bureau de poste donné en l'espace d'un jour.
- ▶ Le nombre de particules α émises par un matériau radioactif pendant un certain laps de temps.

La v.a. dans ces exemples est répartie de manière approximativement poissonnienne car: on approxime par là une variable binomiale.

Loi Géométrique ou de Pascal $\mathcal{G}(p)$

- ▶ ε : “On répète l'épreuve de Bernoulli jusqu'à avoir le premier succès”.
- ▶ Exemple:

$$\begin{array}{ccccccccccc} \bar{A} & \bar{A} & \bar{A} & \bar{A} & \bar{A} & \dots & \bar{A} & \bar{A} & A \\ E & E & E & E & E & \dots & E & E & S \end{array}$$

- ▶ Chaque épreuve a p pour probabilité de succès et $1 - p$ pour probabilité d'échec.
- ▶ X = “le nombre d'épreuves effectuées”.

$$\underbrace{E \ E \ E \ E \ E \ \dots \ E \ E}_{k-1} \ S$$

- ▶ $X(\Omega) = \mathbb{N}^* = \{1, 2, 3, \dots\}$. On dit $X \sim \mathcal{G}(p)$.
- ▶ $\forall k \in \mathbb{N}^* \quad P(X = k) = (1 - p)^{k-1} p$
- ▶ Attention: Parfois X = “nombre d'épreuves effectuées **avant** obtenir le premier succès”. Dans ce cas $X(\Omega) = \mathbb{N}$. On dit $X \sim \mathcal{G}(p)$ sur \mathbb{N} .
- ▶ Cette loi peut servir à modéliser des temps de vie, ou des temps d'attente, lorsque le temps est mesuré de manière discrète (nombre de jours par exemple).
- ▶ Série entière : $\sum_{k=0}^{\infty} x^k = 1/(1 - x)$ pour $|x| < 1$
- ▶ $\sum_{k=1}^{\infty} P(X = k) = \sum_{k=1}^{\infty} (1 - p)^{k-1} p = p \sum_{j=0}^{\infty} (1 - p)^j \sum_{k=1}^{\infty} (1 - p)^{k-1} p = p \sum_{j=0}^{\infty} (1 - p)^j = p \frac{1}{1 - (1 - p)} = 1$

Espérance de loi Géométrique

- ▶ $E(X) = \sum_{k=1}^{\infty} kP(X = k) = \sum_{k=1}^{\infty} kp(1-p)^{k-1} = p \sum_{k=1}^{\infty} k(1-p)^{k-1}$
- ▶ Série entière: $\sum_{k=0}^{\infty} x^k = 1/(1-x)$ pour $|x| < 1$
- ▶ Dérivée première de la série entière: $\sum_{k=1}^{\infty} kx^{k-1} = 1/(1-x)^2$
- ▶ Donc $E(X) = \frac{p}{[1-(1-p)]^2} = \frac{1}{p}$

En d'autres termes, si des épreuves indépendantes ayant une probabilité p d'obtenir un succès sont réalisés jusqu'à ce que le premier succès se produise, le nombre espéré d'essais nécessaires est égal à $1/p$. Par exemple, le nombre espéré de jets d'un dé équilibré qu'il faut pour obtenir la valeur 1 est 6.

Variance de loi Géométrique

- ▶ $V(X) = E(X^2) - E^2(X) = E[X(X - 1)] + E(X) - E^2(X)$. Or,

$$\begin{aligned} E[X(X - 1)] &= \sum_{k=2}^{\infty} k(k - 1)p(1 - p)^{k-1} \\ &= p(1 - p) \sum_{k=2}^{\infty} k(k - 1)(1 - p)^{k-2} \end{aligned}$$

- ▶ Dérivée première de la série entière: $\sum_{k=1}^{\infty} kx^{k-1} = 1/(1 - x)^2$
- ▶ Dérivée seconde de la série entière: $\sum_{k=2}^{\infty} k(k - 1)x^{k-2} = 2/(1 - x)^3$
- ▶ Donc $E[X(X - 1)] = \frac{2p(1-p)}{[1-(1-p)]^3} = \frac{2(1-p)}{p^2}$
- ▶ Et alors $V(X) = E[X(X - 1)] + E(X) - E^2(X) = \frac{1-p}{p^2}$.

Loi Binomiale Négative $\mathcal{BN}(r, p)$

- ▶ ε : “On répète l'épreuve de Bernoulli jusqu'à obtenir un total de r succès”.
- ▶ Exemple avec $r = 3$:

\bar{A} A \bar{A} \bar{A} \bar{A} A \bar{A} \bar{A} A
 E S E E E S E E S

- ▶ Mais on peut obtenir r succès d'autres façons:

S E E E E E S E S
 E E E E S E S E S

- ▶ Chaque épreuve a p pour probabilité de succès et $1 - p$ pour probabilité d'échec.
- ▶ Désignons X = “le nombre d'épreuves nécessaires pour attendre ce résultat”.

$r-1$ succès et $k-r$ échecs
 $\underbrace{E \ S \ E \ E \ E \ S \ E \ E \ S}_{X=k}$

- ▶ $X(\Omega) = \{r, r + 1, r + 2, \dots\}$. On dit $X \sim \mathcal{BN}(r, p)$.
- ▶ $\forall k \in X(\Omega)$,

$$P(X = k) = \binom{k-1}{r-1} p^r (1-p)^{k-r}$$

- ▶ ε : “On répète l'épreuve de Bernoulli jusqu'à obtenir un total de r succès”.
- ▶ Soit,

$E \dots E S E \dots E S \dots E \dots E S$

- ▶ Soit, Y_1 le nombre d'épreuves nécessaires jusqu'au premier succès, Y_2 le nombre d'épreuves supplémentaires nécessaires pour obtenir un deuxième succès, Y_3 celui menant au 3ème et ainsi de suite.
- ▶ Càd,

$\underbrace{E \dots E S}_{Y_1} \underbrace{E \dots E S}_{Y_2} \underbrace{\dots}_{\dots} \underbrace{E \dots E S}_{Y_r}$

- ▶ Les tirages étant indépendantes et ayant toujours la même probabilité de succès, chacune des variables Y_1, Y_2, \dots, Y_r est géométrique $\mathcal{G}(p)$.
- ▶ X = “le nombre d'épreuves nécessaires à l'obtention de r succès” = $Y_1 + Y_2 + \dots + Y_r$.
- ▶ Donc,

$$E(X) = E(Y_1) + E(Y_2) + \dots + E(Y_r) = \sum_{i=1}^r \frac{1}{p} = \frac{r}{p}$$

et

$$V(X) = \sum_{i=1}^r V(Y_i) = \frac{r(1-p)}{p^2}$$

car les Y_i sont indépendantes.