Variables aléatoires discrètes

Module MAP - Telecom nancy

Apprentissage

1 Modélisation

2 Caractéristiques d’une v.a.r.

3 Quelques lois usuelles discrètes

4 Vecteur aléatoire discret

Variable aléatoire discrète

Definition

Soit un espace de probabilité $(\Omega, \mathcal{A}, \mathbb{P})$ et une v.a.r. X définie sur cet espace. X est dite discrète si $X(\Omega) = \{x_i\}_{i \in I}$, l'ensemble des valeurs prises par X, est fini ou dénombrable ($I = [\![1, n]\!]$ ou $I = \mathbb{N}$, ...).

Si on note $A_i = X^{-1}(\{x_i\}) = \{\omega \in \Omega; X(\omega) = x_i\}$ l'image réciproque de $\{x_i\}$ par X, la famille $(A_i)_{i \in I}$ est un système complet d'événements de Ω.

En effet, X étant une application, on a :

- $A_i \cap A_j = \emptyset, \forall i, j \in I, i \neq j$ (un élément de Ω n'a qu'une seule image)
- $\cup_{i \in I} A_i = \Omega$ (tout élément de Ω a une image)

On note $\mathbb{P}(A_i) = \mathbb{P}(X = x_i) = p_i$.

Definition

Soit un espace de probabilité $(\Omega, \mathcal{A}, \mathbb{P})$ et une v.a.r. X discrète définie sur cet espace. On appelle loi de probabilité de X la distribution des probabilités $(p_i)_{i \in I}$ associées aux valeurs $\{x_i, I \in I\}$.

Remarque : La distribution des $p_i = \mathbb{P}(X = x_i)$, $i \geq 1$ vérifie les

conditions $(\star)$ $\begin{cases} 0 \leq p_i \leq 1 \text{(par définition)} \\ \displaystyle\sum_{i \in I} p_i = 1 \text{(La famille } \{X = x\}_{x \in X(\Omega)} \text{ est un s.c.e.,} \end{cases}$

en particulier $\Omega = \cup_{x \in X(\Omega)} \{X = x\}$, alors

$1 = \mathbb{P}(\Omega) = \mathbb{P}(\cup_{x \in X(\Omega)} \{X = x\}) = \displaystyle\sum_{x \in X(\Omega)} \mathbb{P}(X = x)$.

Une suite $(x_i, p_i)_{i \in I}$ peut être considérée comme la loi de probabilité d'une v.a. si les deux conditions $(\star)$ sont vérifiées.

- Dans le cas d’une v.a.r. discrète finie (et raisonnable), la loi de probabilité peut se représenter sous la forme d’un tableau.
- *Exemple : X nombre de Piles obtenus en jetant trois pièces non biaisées.*

Theorem

Soit un espace de probabilité $(\Omega, \mathcal{A}, \mathbb{P})$ et une v.a.r. X discrète définie sur cet espace. La fonction de répartition de X est déterminée par : $F_X(x) = \sum_{x_i \leq x, i \in I} p_i, \ \forall x \in \mathbb{R}$.

C'est une fonction constante par morceaux.

Dans l'exemple :

Opérations sur les v.a.r. discrètes

Theorem

Soit un espace de probabilité $(\Omega, \mathcal{A}, \mathbb{P})$ *et des v.a.r.* X, Y *discrètes définies sur cet espace,* $\lambda \in \mathbb{R}$. *Alors* $X + Y$, λX *et* XY *dont des v.a.r. discrètes, définies par :*

$$\begin{array}{rcl} X + Y : \Omega & \to & \mathbb{R} \\ \omega & \mapsto & X(\omega) + Y(\omega) \end{array}$$

$$\begin{array}{rcl} \lambda X : \Omega & \to & \mathbb{R} \\ \omega & \mapsto & \lambda X(\omega) \end{array}$$

$$\begin{array}{rcl} XY : \Omega & \to & \mathbb{R} \\ \omega & \mapsto & X(\omega)Y(\omega) \end{array}$$

Il est clair que ce sont bien des fonctions mesurables de Ω dans $\mathbb{R}$.
Si on note $X(\Omega) = \{x_i\}_{i\in I}$ et $Y(\Omega) = \{y_j\}_{j\in J}$ avec $I, J \subset \mathbb{N}$, alors $(X + Y)(\Omega) = \{x_i + y_j\}_{(i,j)\in I\times J}$, $(\lambda X)(\Omega) = \{\lambda x_i\}_{i\in I}$ et $(XY)(\Omega) = \{x_i y_j\}_{(i,j)\in I\times J}$ avec $I \subset \mathbb{N}$, $I \times J \subset \mathbb{N}^2$ donc les v.a.r. $X + Y, \lambda X, XY$ sont discrètes.

Transformation d'une v.a.r. discrète

Theorem

Soit un espace de probabilité $(\Omega, \mathcal{A}, \mathbb{P})$ *et une v.a.r.* X *discrète définie sur cet espace, soit* $g : X(\Omega) \to \mathbb{R}$ *une application.*
L'application $g(X)$ *définie par*

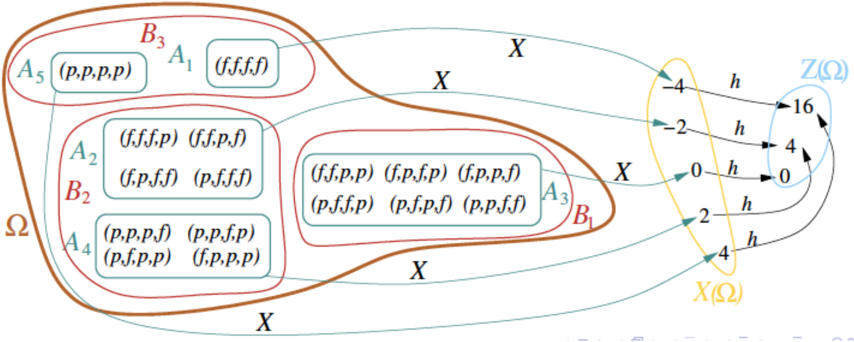
$$\begin{array}{rcl} g(X) : \Omega & \to & \mathbb{R} \\ \omega & \mapsto & g(X(\omega)) \end{array}$$

est une v.a.r. discrète dont la loi est donnée par : $g(X)(\Omega) = \{g(x_i)\}_{i\in I}$ *et* $\forall y \in g(X)(\Omega)$, $\mathbb{P}(g(X) = y) = \sum_{x_i \in X(\Omega), g(x_i)=y} \mathbb{P}(X = x_i)$.

$g(X)(\Omega) = g(X(\Omega)) = \{g(x_i)\}_{i\in I}$ ainsi $g(X)$ est une v.a.r. discrète.
Soit $y \in g(X)(\Omega)$. Alors $\{g(X) = y\} = \{\omega \in \Omega; g(X(\omega)) = y\} = \{\omega \in \Omega; X(\omega) \in \{x_i; i \in I, g(x_i) = y\}\} = \cup_{i\in I; g(x_i)=y}\{X = x_i\}$. On est face à une union dénombrable d'événements deux à deux incompatibles ; donc :
$\mathbb{P}(g(X) = y) = \sum_{i\in I; g(x_i)=y} \mathbb{P}(X = x_i)$.

Exemple 1 : $h : x \to x^2$

- e.a. : 4 lancers d'une pièce équilibrée $\Omega = \{P, F\}^4$, $P(\omega) = \frac{1}{16}, \forall\omega$.
- $X = Nb\ Pile - Nb\ Face$ $X(\Omega) = \{x_1 := -4, x_2 := -2, x_3 := 0, x_4 := 2, x_5 := 4\}$;
- $Z = X^2$, $Z(\Omega) = \{z_1 = 0, z_2 = 4, z_3 = 16\}$.



Exemple 2 : $g : x \to e^x$

$Y(\Omega) = g(X)(\Omega) = \{g(x); x \in X(\Omega)\} = \{e^x; x \in X(\Omega)\}$.

Soit $y \in Y(\Omega)$. *On a :* $\mathbb{P}(Y = y) = \mathbb{P}(e^X = y) = \mathbb{P}(X = ln(y))$. *Ceci est bien défini car* $y \in Y(\Omega) \subset \mathbb{R}_+^*$.

Donc : $\forall y \in Y(\Omega)$, $\mathbb{P}(Y = y) = \mathbb{P}(e^X = y) = \mathbb{P}(X = ln(y))$.

1 Modélisation

2 Caractéristiques d’une v.a.r.

3 Quelques lois usuelles discrètes

4 Vecteur aléatoire discret

Espérance (mathématique) d'une v.a.r. discrète

Soit X une variable aléatoire discrète $X(\Omega) = \{x_i ; i \geq 1 \text{ ou } i \in I\}$ définie sur un espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$.
Alors on dit que la v.a. X est intégrable si la série est absolument convergente : $\sum_{i \in I} | x_i | \mathbb{P}(X = x_i) < +\infty$.

Definition

Soit X une variable aléatoire discrète définie sur un espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$. Si X est intégrable, l'espérance (mathématique) de X est définie par :

$$\mathbb{E}(X) = \sum_{i \in I} x_i \, \mathbb{P}(X = x_i)$$

Elle représente la valeur moyenne de la variable X.

- *Retour à notre exemple :* $\mathbb{E}(X) = \frac{0\times1+1\times3+2\times3+3\times1}{8} = \frac{12}{8} = 1.5$
- Si $X(\Omega)$ est fini, X est intégrable. Donc une v.a.r. discrète finie admet toujours une espérance.
- Dans le cas infini, l'hypothèse d'absolue convergence est fondamentale.

Definition

Une v.a.r. X est dite centrée si $\mathbb{E}(X) = 0$.

$\hookrightarrow$ $X - \mathbb{E}(X)$ est une v.a.r. centrée.

Propriétés de l'espérance

Soit X, Y deux v.a.r. discrètes définies sur un espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$, admettant toutes deux une espérance. Soit $\alpha, \beta \in \mathbb{R}$.

- $X + Y$ et αX admettent une espérance et
$$\left.\begin{array}{l}\mathbb{E}(\alpha X) = \alpha \mathbb{E}(X) \\ \mathbb{E}(X+Y) = \mathbb{E}(X) + \mathbb{E}(Y)\end{array}\right\} \text{linéarité}$$

  Pour αX : notons $X(\Omega) = \{x_i\}_{i \in I}$. Si $\alpha = 0$, la v.a.r. αX est identiquement nulle. Donc $\mathbb{E}(\alpha X) = \mathbb{E}(0) = 0$. Si $\alpha \neq 0$, l'ensemble image de αX est $(\alpha X)(\Omega) = \{\alpha x_i\}_{i \in I}$. On peut admettre que X intégrable conduit à αX intégrable, donc αX admet une espérance et $\mathbb{E}(\alpha X) = \sum_{i \in I} \alpha x_i \, \mathbb{P}(X = x_i) = \alpha \, \sum_{i \in I} x_i \, \mathbb{P}(X = x_i) = \alpha \mathbb{E}(X)$.

- $\mathbb{E}(\alpha) = \alpha$; $\alpha X + \beta$ admet une espérance et $\mathbb{E}(\alpha X + \beta) = \alpha \mathbb{E}(X) + \beta$
- Si $X \geq 0$, alors $\mathbb{E}(X) \geq 0$ positivité
- Si $X \geq Y$, alors $\mathbb{E}(X) \geq \mathbb{E}(Y)$ croissance

Théorème du transfert

Soit X une variable aléatoire discrète $X(\Omega) = \{x_i; i \geq 1 \text{ ou } i \in I\}$ définie sur un espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$. Soit $g : X(\Omega) \to \mathbb{R}$ une application. La v.a. $g(X)$ est intégrable si la série est absolument convergente : $\sum_{i \geq 1} | g(x_i) | \mathbb{P}(X = x_i) < +\infty$.

Sous cette hypothèse, l'espérance de $g(X)$ est définie par :

$$\mathbb{E}(g(X)) = \sum_{i \in I} g(x_i) \, \mathbb{P}(X = x_i)$$

Cette formule est importante car elle nous évite de devoir établir la loi de Z !

Exercice : soit X telle que $X(\Omega) = [\![-n, n]\!]$
et $\mathbb{P}(X = i) = 1/(2n + 1), \forall i \in [\![-n, n]\!]$.

Calculer $\mathbb{E}(X)$ et $\mathbb{E}(X^2)$ (cette dernière en utilisant la formule de transfert et aussi en établissant au préalable la loi de X^2).

Aide : $\displaystyle\sum_{k=1}^{n} k^2 = \frac{n(n+1)(2n+1)}{6}$.

Variance d'une v.a. discrète

Definition

Soit X une variable aléatoire définie sur un espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$, et de carré intégrable. On appelle variance de la v. a. X :
$\mathrm{Var}(X) = V(X) = s_X^2 = \sigma_X^2 = \mathbb{E}[(X - \mathbb{E}(X))^2]$.
Si X une variable aléatoire discrète $X(\Omega) = \{x_i; i \in I\}$, la variance s'écrit : $\mathrm{Var}(X) = \sum_{i \in I} (x_i - \mathbb{E}(X))^2 p_i$.

- estimation de la dispersion de la variable autour de la moyenne ("moyenne des carrés des écarts à la moyenne")
- Autre formulation :

Théorème de Koenig-Huygens

$\mathrm{Var}(X) = \mathbb{E}(X^2) - (\mathbb{E}(X))^2$("moyenne des carrés - carré de la moyenne")

$\mathrm{Var}(X) = \sum_{i \in I} x_i^2 p_i - (\mathbb{E}(X))^2$ (cas discret)

Par linéarité de l'espérance :

$$\begin{aligned}\mathrm{Var}(X) &= \mathbb{E}((X-\mathbb{E}(X))^2) = \mathbb{E}(X^2 - 2\mathbb{E}(X)X + \mathbb{E}(X)^2) \\ &= \mathbb{E}(X^2) - 2\mathbb{E}(X)\mathbb{E}(X) + \mathbb{E}(X)^2 \\ \mathrm{Var}(X) &= \mathbb{E}(X^2) - \mathbb{E}(X)^2 = (\textstyle\sum_{i\in I} x_i^2 \mathbb{P}(X = x_i)) - \mathbb{E}(X)^2\end{aligned}$$

Cette formulation est plus rapide pour les calculs "à la main".

Definition

Soit X une variable aléatoire définie sur un espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$, et de carré intégrable. On appelle écart-type de la variable X : $\sigma_X = s_X = \sqrt{s_X^2} = \sqrt{\mathrm{Var}(X)}$.

- Plus lisible que la variance car dans la même unité que la variable (si X est une mesure en *cm*, Var(X) est en cm^2).
- *Retour à notre exemple :* $s_X^2 = 0.75$; $s_X = 0.866$

Propriétés de la variance

On se place dans un espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$ sur lequel sont définis des v.a. discrètes X et Y admettant toutes deux une variance. Soit $\alpha \in \mathbb{R}$.

- $\text{Var}(X) \geq 0$
- $\text{Var}(\alpha X) = \alpha^2 \text{Var}(X)$
- $\text{Var}(\alpha) = 0$
- $\text{Var}(X + Y) = \text{Var}(X) + \text{Var}(Y) + 2\,\text{Cov}(X, Y)$ où :

Definition

Soit X, Y des v.a. définies sur un même espace de probabilité $(\Omega, \mathcal{A}, \mathbb{P})$ et admettant chacune une variance. On appelle covariance de X et Y :

$$\begin{aligned}\text{Cov}(X, Y) = s^2_{X,Y} \quad &= \text{E[(X - E(X))(Y - E(Y))]} \\ &= \sum_{i,j \geq 1} (x_i - \mathbb{E}(X))\,(y_j - \mathbb{E}(Y))\,\mathbb{P}(X = x_i, Y = y_j)\end{aligned}$$

Propriétés de la covariance

- C'est une mesure de la dépendance (linéaire) entre X et Y.
- $\mathrm{Cov}(X, Y) = \mathbb{E}(X \times Y) - \mathbb{E}(X) \times \mathbb{E}(Y)$

$$
\begin{aligned}
\mathrm{Cov}(X, Y) &= \mathbb{E}[(X - \mathbb{E}(X))(Y - \mathbb{E}(Y))] \\
&= \mathbb{E}[XY - \mathbb{E}(Y)X - \mathbb{E}(Y)X + \mathbb{E}(X)\mathbb{E}(Y)] \\
&= \mathbb{E}(XY) - \mathbb{E}(Y)\mathbb{E}(X) - \mathbb{E}(Y)\mathbb{E}(X) + \mathbb{E}(X)\mathbb{E}(Y) \\
&= \mathbb{E}(XY) - \mathbb{E}(X)\mathbb{E}(Y)
\end{aligned}
$$

  Dans le cas discret :

$$\mathrm{Cov}(X, Y) = \sum_{i,j \geq 1} x_i\, y_j\, \mathbb{P}(X = x_i; Y = y_j) - \mathbb{E}(X) \times \mathbb{E}(Y)$$

- $\mathrm{Cov}(X, X) = \mathrm{var}(X) \geq 0$ (positivité)
- $\mathrm{Cov}(X, Y) = \mathrm{Cov}(Y, X)$ (symétrie)
- $\mathrm{Cov}(X, \alpha) = 0$

- Rappel dans le cas discret : On dit que X et Y sont indépendantes si

$$\forall x \in X(\omega), y \in Y(\Omega),\ \mathbb{P}(X = x, Y = y) = \mathbb{P}(X = x).\mathbb{P}(Y = y)$$

- Si X, Y indépendants $\Rightarrow \mathrm{Cov}(X, Y) = 0$ et donc $\mathrm{Var}(X + Y) = \mathrm{Var}(X) + \mathrm{Var}(Y)$

  $\mathrm{Cov}(X, Y) = \mathbb{E}(X.Y) - \mathbb{E}(X).\mathbb{E}(Y) = 0$ par l'indépendance.

  Attention, la réciproque est fausse !

  Contre-exemple : soit la v.a.r. X définie sur $X(\Omega) = \{-1, 0, 1\}$ par $\mathbb{P}(X = -1) = \mathbb{P}(X = 0) = \mathbb{P}(X = 1) = \dfrac{1}{3}$

  et la v.a. Y définie sur $Y(\Omega) = \{0, 1\}$ par $Y = \begin{cases} 0 \text{ si } X \neq 0 \\ 1 \text{ si } X = 0 \end{cases}$

  Clairement X et Y ne sont pas indépendantes et pourtant $\mathrm{Cov}(X, Y) = 0$, en effet :

  $\mathbb{E}(X) = \dfrac{1}{3}(-1 + 0 + 1) = 0$, $\mathbb{E}(Y) = 0 \times \dfrac{2}{3} + 1 \times \dfrac{1}{3} = \dfrac{1}{3}$

  et $\mathbb{E}(XY) = 0$ car $X.Y \equiv 0$, donc $\mathrm{Cov}(X, Y) = 0$.

Definition

Une v.a.r. X est dite réduite si $s_X = 1$.

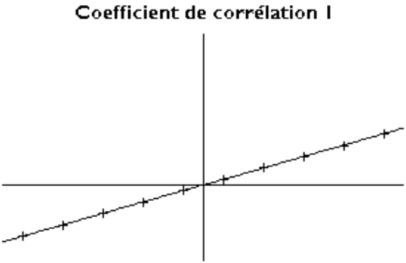
$\hookrightarrow \quad \dfrac{X - \mathbb{E}(X)}{s_X}$ est une v.a.r. centrée et réduite.

Definition

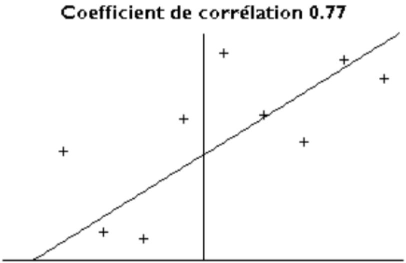
On se place dans un espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$ sur lequel sont définis des v.a. discrètes X, Y admettant toutes une variance. [a] On appelle coefficient de corrélation linéaire : $\rho(X, Y) = \dfrac{\mathrm{Cov}(X, Y)}{\sigma_X \sigma_Y}$.

a. ou soit (X, Y) un couple aléatoire de loi conjointe $\mathbb{P}$

- $-1 \leq \rho(X, Y) \leq 1$
- $X \perp\!\!\!\perp Y \Rightarrow \rho(X, Y) = 0$
- Si $\rho(X, Y) \simeq 1$, alors dépendance linéaire positive entre X et Y, i.e. $Y = aX + b, a > 0$ (idem avec -1, $a < 0$).



X = a*Y (corrélation linéaire)



Nuage de point

Definition

Soit une variable aléatoire X définie un espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$, soit $k \in \mathbb{N}^*$. On dit que X admet un **moment d'ordre k** si X^k est intégrable. Le moment d'ordre k est défini par $\mathbb{E}(X^k)$, on le notera m_k.

L'espérance est donc le moment d'ordre 1.

Definition

Soit une variable aléatoire X définie un espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$, soit $k \in \mathbb{N}^*$. Le **moment centré d'ordre k** est défini par $\mathbb{E}((X - \mathbb{E}(X))^k)$, on le notera μ_k.

La variance est donc le moment centré d'ordre 2.

1 Modélisation

2 Caractéristiques d'une v.a.r.

3 Quelques lois usuelles discrètes

4 Vecteur aléatoire discret

Quelques lois usuelles discrètes

Loi de Bernoulli

Soit E une épreuve comportant deux résultats possibles S et E (nommés "succès" et "échec") : épreuve ou schéma de Bernoulli [a]. On définit la variable aléatoire indicatrice de l'événement "'succès" : $X(\Omega) = \{0, 1\}$

$$X \mapsto \begin{cases} 1 \text{ si "succès"} & \mathbb{P}(X = 1) = p \\ 0 \text{ si "échec"} & \mathbb{P}(X = 0) = 1 - p \end{cases}$$

Notation : $\mathcal{L}(X) = \mathcal{B}e(p)$.

---

a. Mathématicien suisse J.Bernoulli 1654-1705, "Ars Conjectandis"

Caractéristiques

$\mathbb{E}(X) = p$
$s_X^2 = p\,(1 - p)$

La v.a. étant finie, elle admet une espérance et une variance.
$\mathbb{E}(X) = 1\times p + 0\times(1-p) = p$; $\mathrm{Var}(X) = [1\times p + 0\times(1-p)] - p^2 = p - p^2 = p(1-p)$.

Exemple 1 : jet d'une pièce de monnaie

$$X = \begin{cases} 1 \text{ si on obtient Pile} \\ 0 \text{ sinon} \end{cases} \qquad X \sim \mathcal{B}e\left(\frac{1}{2}\right)$$

Exemple 2 : tirage d'une boule dans une urne contenant r boules rouges et v boules vertes

$$X = \begin{cases} 1 \text{ si on obtient une boule rouge} \\ 0 \text{ sinon} \end{cases} \qquad X \sim \mathcal{B}e\left(\frac{r}{r+v}\right)$$

Loi binomiale

Une variable aléatoire Y est dite binomiale de paramètres $n \in \mathbb{N}^*$ et $p \in [0,1]$ notée $\mathcal{B}(n,p)$ si $Y(\Omega) = \{0, 1, 2, ..., n\}$ et

$$\mathbb{P}(Y = k) = C_n^k\, p^k\, (1-p)^{n-k}, \text{ pour } k = 0, 1, 2, ..., n$$

- La loi binomiale modélise le nombre de succès dans une expérience aléatoire que l'on peut décomposer en n épreuves de Bernoulli $\mathcal{B}e(p)$ indépendantes. Ainsi, pour le calcul de probabilité : k succès (de probabilité p), $n-k$ échecs (de probabilité $1-p$), et C_n^k possibilités pour placer les k succès parmi les n épreuves.
- On peut écrire cette v.a. comme $Y = X_1 + ... + X_n$ avec $X_1, ..., X_n$ n variables de même loi $\mathcal{B}e(p)$ indépendantes.

- *Exemple : la variable X de notre exemple précédent (nombre de Piles obtenus lors du jet de 3 pièces) suit une loi Binomiale $\mathcal{B}(3, \frac{1}{2})$.*
- Distribution symétrique si $p = 0,5$; dissymétrique sinon.

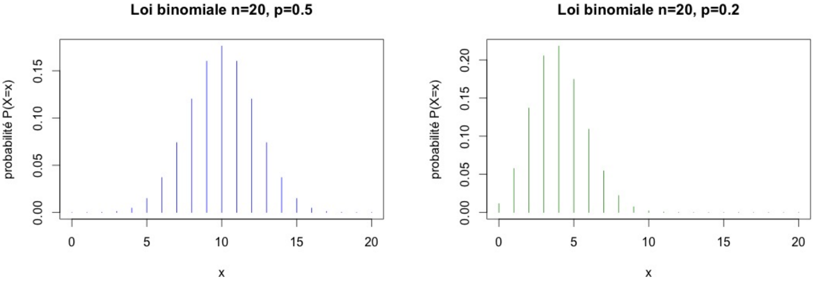


Figure – loi binomiale avec $n = 20$; et $p = 0.5$; 0.2

Caractéristiques

$\mathbb{E}(Y) = n \times p$
$\mathrm{Var}(Y) = n \times p \times (1 - p)$

- Si $X_1, ..., X_N$ sont des variables mutuellement indépendantes de lois respectives $\mathcal{B}(n_i, p)$ (pour $i = 1, ..., N$), alors $Y = X_1 + ... + X_N$ est une variable de loi $\mathcal{B}(n_1 + ... + n_N, p)$
- $\mathcal{L}(n - Y) = \mathcal{B}(n, 1 - p)$ en choisissant de compter les échecs.

On se place toujours dans le même contexte de répétition d'une e.a. de type Bernoulli de même paramètre. Mais on ne limite plus le nombre de répétition ; on s'arrête au premier Succès.

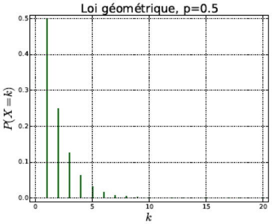
Loi géométrique

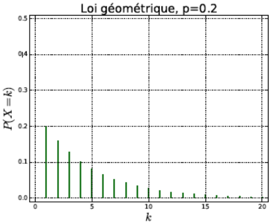
Une variable aléatoire X suit une loi géométrique de paramètre $p \in [0, 1]$ si $X(\Omega) = \mathbb{N}^*$ et

$$\forall k \in \mathbb{N}^*, \mathbb{P}(X = k) = (1 - p)^{k-1} p$$

Notation : $X \sim \mathcal{G}(p)$

- *Exemple : on lance une pièce de monnaie jusqu'à ce qu'elle tombe sur Pile. La variable du nombre de lancés effectués* $X \sim \mathcal{G}(\frac{1}{2})$





- C'est une loi sans mémoire :

$$\mathbb{P}(X = k + t / X > k) = \mathbb{P}(X = t), \ \forall k \in \mathbb{N}, t \in \mathbb{N}^*$$

Caractéristiques

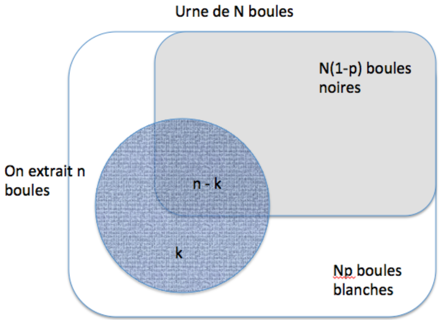
$$\mathbb{E}(Y) = \frac{1}{p}$$

$$\mathrm{Var}(Y) = \frac{1-p}{p^2}$$

Que se passe-t-il lorsque la probabilité varie entre chaque répétition de l'expérience aléatoire (on n'a plus indépendance) ?

On peut considérer l'expérience suivante :
Une urne contient N boules : m blanches, et $N - m$ noires. On tire n boules sans remise, donc les tirages sont non indépendants. La variable aléatoire X du nombre de boules blanches tirées suit une loi hypergéométrique $\mathcal{H}(N, n, p)$ avec $p = \frac{m}{N}$

Construisons-la...



boules noires : $N - m = N(1-p)$; boules blanches : $m = Np$

On suppose bien sûr $n \leq N$. On a $X \leq \min\{n, Np\}$. D'autre part si $N(1-p) < n$, on aura nécessairement $X \geq n - N(1-p)$, soit $X \geq \max\{0, n - N(1-p)\}$. Conclusion :
$X(\Omega) = [\![\max\{0, n - N(1-p)\}, \min\{n, Np\}]\!]$

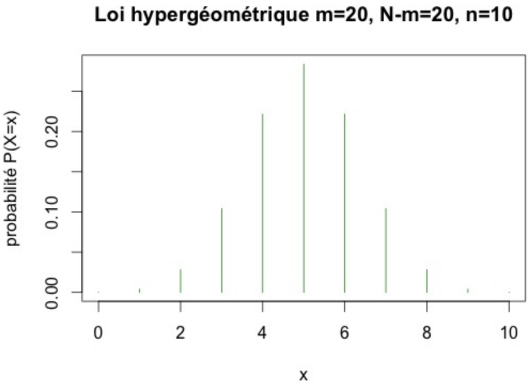
Dans les cas usuels $n \leq \min\{K, N(1-p)\}$, alors $X(\Omega) = [\![0, n]\!]$.

Loi hypergéométrique

Une variable aléatoire X suit une loi hypergéométrique de paramètres N, n, p, avec $N, n \in \mathbb{N}^*$, $p \in [0, 1]$ tels que $Np \in \mathbb{N}$ et $N \geq n$, si $X(\Omega) = \{0, 1, ..., n\}$ et

$$\forall k \in 0, 1, ..., n,\ \mathbb{P}(X = k) = \frac{C_{Np}^{k}\ C_{N(1-p)}^{n-k}}{C_{N}^{n}}$$

Condition : $\max\{0; n - (N - Np)\} \leq k \leq \min\{n; Np\}$
Notation : $X \sim \mathcal{H}(N, n, p)$



$\mathcal{H}(40, 10, 1/2)$

Caractéristiques

$\mathbb{E}(Y) = np$

$\mathrm{Var}(Y) = np(1-p)\dfrac{N-n}{N-1}$

- Lorsque $N \to \infty$, la loi hypergéométrique $\mathcal{H}(N, n, p)$ converge vers la loi binomiale $\mathcal{B}(n, p)$. Ceci est admis dès que N est grand devant n : $N \geq 10n$.

Loi de Poisson

Une v.a. discrète X suit une loi de Poisson de paramètre $\lambda > 0$ lorsque $X(\Omega) = \mathbb{N}$ et

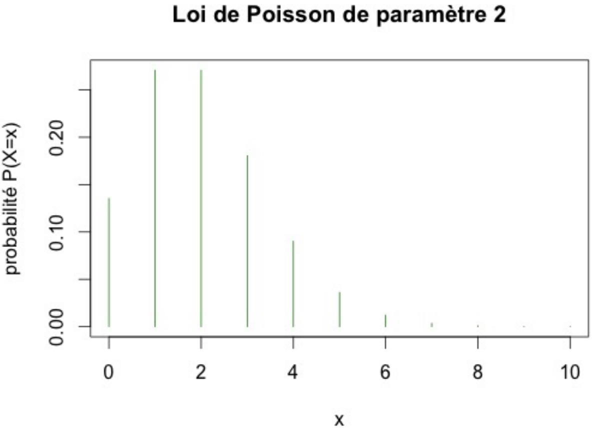
$$\forall k \in \mathbb{N}, \ \mathbb{P}(X = k) = e^{-\lambda} \frac{\lambda^k}{k!}$$

Notation : $\mathcal{L}(X) = \mathcal{P}(\lambda)$.

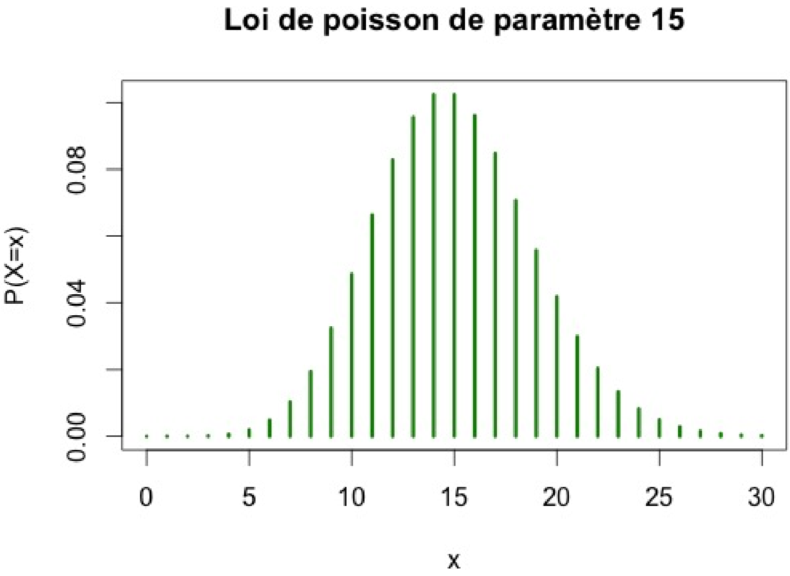
- introduite par le mathématicien français Simeon Denis Poisson en 1838 (*Recherche sur la probabilités des jugements en matière criminelle et en matière civile*).
- "Loi des événements rares", lorsque l'on veut compter des événements indépendants, se produisant sur un laps de temps donné tels que la probabilité d'apparition ne dépend que du laps de temps

Exemples : une autoroute sur laquelle se produit en moyenne 1.8 accidents par semaine (le nombre d'accidents par semaine suit une loi $\mathcal{P}(1.8)$), nombre de particules émises par un matériau radioactif dans un intervalle de temps défini...

- Distribution dissymétrique étalée à droite



- Distribution tend à devenir symétrique lorsque λ augmente.



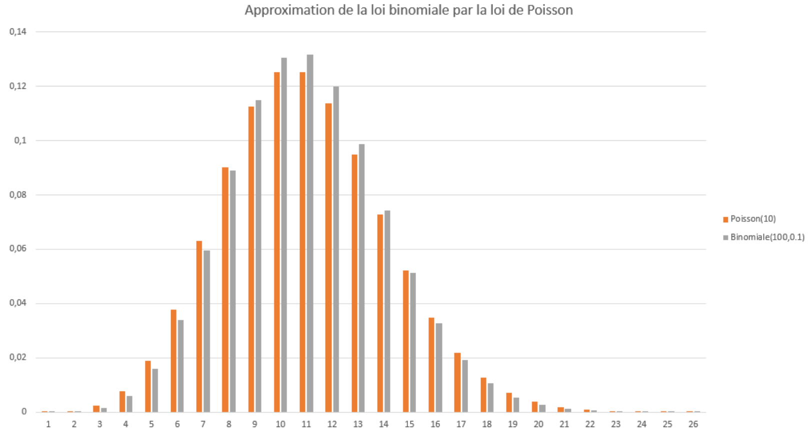
Caractéristiques

$\mathbb{E}(X) = \lambda$
$s_X^2 = \lambda$

- On remarque que l'espérance et la variance sont identiques. C'est une propriété importante de la loi de Poisson qui peut permettre de l'identifier lorsque l'on est dans le contexte d'identification d'une loi.

- Si X, Y sont deux variables indépendantes de lois respectives $\mathcal{P}(\lambda)$ et $\mathcal{P}(\mu)$, alors $Z = X + Y$ est une variable de loi $\mathcal{P}(\lambda + \mu)$ (Propriété de stabilité)
- Si n est suffisament grand, la loi de Poisson $\mathcal{P}(np)$ est une bonne approximation de la loi binomiale $\mathcal{B}(n, p)$ $(n \geq 20,\ p \leq 0,05$ usuellement admis)[1].

1. vous trouverez presque autant de conditions que d'auteurs !



Graphique de Valentin P-W

Fonctions génératrices

Nous allons voir un outil assez pratique pour les **v.a. discrètes dont les valeurs sont des entiers positifs** ($X(\Omega) \subset \mathbb{N}$). On les appelle v.a. entières.

Soit X une telle v.a. entière et $s \in [0,1]$ fixé. Alors s^X est une v.a. discrète à valeur dans $\{s^k, k \in X(\Omega)\}$ qui est intégrable.

En effet : $E(|s^X|) = E(s^X) = \sum_{k \in X(\Omega)} \underbrace{s^k}_{\in[0,1]} P(X=k) \leq \sum_{k \in X(\Omega)} P(X=k) = 1$

Definition (Fonction génératrice)

La fonction $s \in [0,1[\mapsto g_X(s) := E(s^X)$ s'appelle fonction génératrice de la v.a. entière X. Elle s'écrit :

$$g_X(s) = \sum_{k \in X(\Omega)} s^k P(X=k) = \sum_{k \in X(\Omega)} p_k s^k$$

1 Modélisation

2 Caractéristiques d'une v.a.r.

3 Quelques lois usuelles discrètes

4 Vecteur aléatoire discret

Vecteur aléatoire discret

Definition

- Soient n variables aléatoires $X_i : \Omega \to E_i$ définies sur $(\Omega, \mathcal{A}, \mathbb{P})$ et discrètes. On appelle vecteur aléatoire l'application mesurable : $\begin{array}{rcl} X : \Omega & \to & X(\Omega) \subset E_1 \times \ldots \times E_n \\ \omega & \mapsto & X(\omega) = (X_1(\omega), \ldots, X_n(\omega)) \end{array}$
- Loi de probabilité (con)jointe de X : $\{\mathbb{P}(X = (x_1, \ldots x_n); x_i \in E_i \, \forall i\} = \{\mathbb{P}(X_1 = x_1, \ldots, X_n = x_n); x_i \in E_i \, \forall i\}$
- Pour $j \in [\![1; n]\!]$, la loi marginale de X_j est donnée par :

$$\mathbb{P}(X_j = x_j) = \sum_{(x_1, \ldots, x_{j-1}, x_{j+1} \ldots x_n) \in E_1 \times \ldots E_{j-1} \times E_{j+1} \times \ldots \times E_n} \mathbb{P}(X = (x_1, \ldots, x_n))$$

On a : $\forall x_i \in E_i$, $\mathbb{P}(X_1 = x_1, \ldots, X_n = x_n) \in [0; 1]$ et
$$\sum_{\{(x_1, \ldots, x_n) \in E_1 \times \ldots \times E_n\}} \mathbb{P}(X_1 = x_1, \ldots, X_n = x_n) = 1.$$

On se place en dimension 2 : couple de v.a.r., pour simplifier les écritures suivantes. *Exemple : jet de 3 pièces non biaisées*
$Z = (X, Y); X = 1\!\!\mathrm{I}_A; Y = 1\!\!\mathrm{I}_B$ *avec les événements :*
A = "On obtient au plus une fois Pile"
B= "On obtient au moins une fois Pile et au moins une fois Face"

Loi conjointe de $Z = (X, Y)$ *:*

Loi de la somme de deux variables aléatoires X et Y discrètes

Soit un couple de v.a. (X, Y) défini sur un espace probabilisé $(\Omega, \mathcal{A}, \mathbb{P})$.

- $\mathbb{P}(X + Y = z) = \displaystyle\sum_{x \in X(\Omega), z-x \in Y(\Omega)} \mathbb{P}(X = x, Y = z - x)$
- Si X et Y sont indépendantes, alors $\mathbb{P}(X + Y = z) = \displaystyle\sum_{x \in X(\Omega), z-x \in Y(\Omega)} \mathbb{P}(X = x)\mathbb{P}(Y = z - x)$ C'est le produit de convolution des lois de X et Y : $\mathbb{P}_{X+Y} = \mathbb{P}_X * \mathbb{P}_Y$.

Suite de l'exemple : loi de $X + Y$

ω	$(X, Y)(\omega)$	$X(\omega) + Y(\omega)$
(P,P,P)	(0,0)	0
(P,P,F)	(0,1)	1
(P,F,P)	(0,1)	1
(P,F,F)	(1,1)	2
(F,P,P)	(0,1)	1
(F,P,F)	(1,1)	2
(F,F,P)	(1,1)	2
(F,F,F)	(1,0)	1

Ceci peut nous permettre aussi d'établir des propriétés de stabilité de certaines lois.

Alors la loi de $X + Y$ est :

On l'obtient aussi avec le produit de convolution :

Loi de probabilité conditionnelle

Definition

Soit un couple de v.a.r. (X, Y) défini sur un espace de probabilité $(\Omega, \mathcal{A}, \mathbb{P})$ de loi jointe $(\mathbb{P}(X = x_i, Y = y_j) = p_{ij})_{i \in I, j \in J}$. Notons $(x_i, p_i)_{i \in I}$ et $(y_j, p_j)_{j \in J}$ les lois marginales. La loi Y conditionnelle à $\{X = x_i\}$ est l'application de $X(\Omega)$ dans $\mathbb{R}$, définie par :

$$b_i(j) = \mathbb{P}(Y = y_j / X = x_i) = \frac{\mathbb{P}(X = x_i, Y = y_j)}{\mathbb{P}(X = x_i)} = \frac{p_{ij}}{p_i}$$

Elle permet une définition simple de l'espérance conditionnelle : si Y admet une espérance finie, alors on définit l'espérance conditionnelle par rapport à $\{X = x_i\}$ par :

$$\mathbb{E}(Y/X = x_i) = \sum_{j \in J} b_i(j) y_j$$

Remarque : $\mathbb{E}(\mathbb{E}(Y/X)) = \mathbb{E}(Y)$.

Retour à l'exemple :
Loi de X sachant $Y = 0$:
Loi de X sachant $Y = 1$:

- On peut retrouver les lois marginales via les lois conditionnelles, par exemple pour X :

$$\forall i \in I, \mathbb{P}(X = x_i) = \sum_{j \in J} \mathbb{P}(Y = y_j)\mathbb{P}(X = x_i / Y = y_j)$$

Dans l'exemple, on retrouve facilement la loi marginale de X :

Extension de la loi binomiale : loi multinomiale

Supposons qu'une e.a. admet m résultats possibles notés $R_1, ..., R_m$, avec probabilités respectives connues $p_1, ..., p_m$ ($p_1 + ... + p_m = 1$). On réalise n épreuves indépendantes. Quelle est la probabilité d'obtenir n_1 résultats R_1, ... n_m résultats R_m ?

Loi multinomiale

Une variable aléatoire Y est dite multinomiale de paramètres $n \in \mathbb{N}^*$ et $p_1, ..., p_m \in [0, 1]$ satisfaisant $p_1 + ... + p_m = 1$, si $Y = (X_1, ...X_m)$, chaque X_i étant le nombre d'occurence du résultat R_i, $Y(\Omega) = [\![0, n]\!]^m$ et $\forall n_1, ..., n_m \in \mathbb{N}^*; n_1 + ... + n_m = n$:

$$\mathbb{P}(X_1 = n_1, ..., X_m = n_m) = \frac{n!}{n_1!...n_m!} p_1^{n_1} ... p_m^{n_m}$$

Remarquons que la loi marginale de X_i est une binomiale $\mathcal{B}(n, p_i)$.