UNIVERSITÉ DE GENÈVE

Introduction à la Statistique

Indépendance et variables aléatoires

Stéphane Guerrier, Mucyo Karemera, Samuel Orso & Lionel Voirol

Data Analytics Lab

Règles de Probabilité

1. Probabilité conditionnelle de *A sachant B*

$$\mathbb{P}(A|B) = \frac{\mathbb{P}(A \textbf{ et } B)}{\mathbb{P}(B)} = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)}$$

   Probabilité conditionnelle de *B sachant A*

$$\mathbb{P}(B|A) = \frac{\mathbb{P}(A \textbf{ et } B)}{\mathbb{P}(A)} = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(A)}$$

2. Règle de multiplication

$$\mathbb{P}(A \textbf{ et } B) = \mathbb{P}(A \cap B) = \mathbb{P}(B) \cdot \mathbb{P}(A|B)$$

$$\mathbb{P}(A \textbf{ et } B) = \mathbb{P}(A \cap B) = \mathbb{P}(A) \cdot \mathbb{P}(B|A)$$

Règles de Probabilité

3. Règle générale d'addition

$$\mathbb{P}(A \textbf{ ou } B) = \mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B)$$

4. Probabilité totale

$$\mathbb{P}(A) = \mathbb{P}(A \cap B) + \mathbb{P}(A \cap B^c)$$

5. Règle du complément

$$\mathbb{P}(A) = 1 - \mathbb{P}(A^c)$$

$$\mathbb{P}(A|B) = 1 - \mathbb{P}(A^c|B)$$

Exemple

Paquet standard de cartes

Un paquet standard de cartes contient 52 cartes.

4 couleurs (Cœurs ♥, Piques ♠, Carreaux ♦, Trèfles ♣)

Chaque couleur contient : Des numéros de 1 à 10

Des figures (Roi , Dame , Valet )

Nous tirons une carte au hasard.
Quelle est la probabilité de tirer un Roi ou un Cœur ?

Exemple

Utilisez la règle générale de l'addition :

$$\mathbb{P}(\text{“}Roi\text{”}) = \frac{4}{52}$$

$$\mathbb{P}(\text{“}Coeurs\text{”}) = \frac{13}{52}$$

$$\mathbb{P}(\text{“}Roi\text{”} \cap \text{“}Coeurs\text{”}) = \mathbb{P}(\text{“}Roi\ de\ Coeurs\text{”}) = \frac{1}{52}$$

$$\begin{aligned}\mathbb{P}(\text{“}Roi\text{”} \cup \text{“}Coeurs\text{”}) &= \mathbb{P}(\text{“}Roi\text{”}) + \mathbb{P}(\text{“}Coeurs\text{”}) - \mathbb{P}(\text{“}Roi\text{”} \cap \text{“}Coeurs\text{”}) \\ &= \frac{4}{52} + \frac{13}{52} - \frac{1}{52} = \frac{16}{52} = \frac{4}{13}\end{aligned}$$

Exemple

Considérons un dé équilibré.
Quelle est la probabilité de lancer un *nombre pair* ou un 3 ?

$$\mathbb{P}(\textit{“Lancer Pair”}) = \mathbb{P}(\{2,4,6\}) = \frac{3}{6} = \frac{1}{2}$$

$$\mathbb{P}(\textit{“Lancer 3”}) = \mathbb{P}(\{3\}) = \frac{1}{6}$$

“Lancer un nombre pair” et *“Lancer un* 3*”* sont deux *événements mutuellement exclusifs*.

$$\begin{aligned}\mathbb{P}(\textit{“Pair ou 3”}) &= \mathbb{P}(\textit{“Lancer Pair”}) + \mathbb{P}(\textit{“Lancer 3”}) \\ &= \frac{1}{6} + \frac{3}{6} = \frac{4}{6} = \frac{2}{3}\end{aligned}$$

Indépendance

A et *B* sont des événements *indépendants* si

$$\mathbb{P}(B|A) = \mathbb{P}(B) \quad \text{ou} \quad \mathbb{P}(A|B) = \mathbb{P}(A).$$

Remarques : Si A et B sont des événements *indépendants*, alors (i) A^c et B, (ii) A et B^c, (iii) B^c et A, et (iv) A^c et B^c sont aussi des événements indépendants, c'est-à-dire,

$$\mathbb{P}(A^c|B) = \mathbb{P}(A^c), \quad \mathbb{P}(A|B^c) = \mathbb{P}(A),$$
$$\mathbb{P}(B^c|A) = \mathbb{P}(B^c), \quad \mathbb{P}(A^c|B^c) = \mathbb{P}(A^c).$$

- Probabilité jointe de deux événements indépendants :

$$\begin{aligned}\mathbb{P}(A \text{ et } B) &= \mathbb{P}(A) \cdot \mathbb{P}(B|A) = \mathbb{P}(B) \cdot \mathbb{P}(A|B) \quad \textit{(Règle de multiplication)}\\ &= \mathbb{P}(A) \cdot \mathbb{P}(B) \quad \textit{(Indépendance)}.\end{aligned}$$

Indépendance : “Roi” et “Coeur”

$$\mathbb{P}(\textit{“Roi”}) = \frac{4}{52}$$

$$\mathbb{P}(\textit{“Coeurs”}) = \frac{13}{52}$$

$$\mathbb{P}(\textit{“Roi”} \cap \textit{“Coeurs”}) = \mathbb{P}(\textit{“Roi de Coeurs”}) = \frac{1}{52}$$

$$\mathbb{P}(\textit{“Roi”}) \cdot \mathbb{P}(\textit{“Coeurs”}) = \frac{4}{52} \cdot \frac{13}{52} = \frac{52}{52^2} = \frac{1}{52}$$

Ils sont indépendants !

Indépendance : “Roi” et “Coeur”

Une autre façon de vérifier l'indépendance :

$$\mathbb{P}(\textit{“Roi”}) = \frac{4}{52} = \frac{1}{13}$$

$$\mathbb{P}(\textit{“Roi”} \mid \textit{“Coeurs”}) = \frac{\mathbb{P}(\textit{“Roi”} \cap \textit{“Coeurs”})}{\mathbb{P}(\textit{“Coeurs”})} = \frac{\mathbb{P}(\textit{“Roi de Coeurs”})}{\mathbb{P}(\textit{“Coeurs”})}$$

$$= \frac{\frac{1}{52}}{\frac{13}{52}} = \frac{1}{13}$$

$$\mathbb{P}(\textit{“Roi”} \mid \textit{“Coeurs”}) = \mathbb{P}(\textit{“Roi”})$$

Ils sont indépendants !

Indépendance : "Figure" et "Roi"

$$\mathbb{P}(\textit{“Roi”}) = \frac{4}{52}$$

$$\mathbb{P}(\textit{“Figure”}) = \mathbb{P}(\{\textit{“Roi coeur”}, \textit{“Dame coeur”}, \ldots, \textit{“Valet Pique”}\}) = \frac{12}{52}$$

$$\mathbb{P}(\textit{“Roi”} \cap \textit{“Figure”}) = \mathbb{P}(\textit{“Roi”}) = \frac{4}{52}$$

$$\mathbb{P}(\textit{“Roi”}) \cdot \mathbb{P}(\textit{“Figure”}) = \frac{4}{52} \cdot \frac{12}{52} = \frac{48}{52^2}$$

Ils ne sont PAS indépendants !

Indépendance : “Figure” et “Roi”

Une autre façon de vérifier l’indépendance :

$$\mathbb{P}(\textit{“Roi”}\,) = \frac{4}{52} = \frac{1}{13}$$

$$\begin{aligned}
\mathbb{P}(\textit{“Roi”}\,|\textit{“Figure”}\,) &= \frac{\mathbb{P}(\textit{“Roi”} \cap \textit{“Figure ”})}{\mathbb{P}(\textit{“Figure”}\,)} \\
&= \frac{\mathbb{P}(\textit{“Roi”}\,)}{\mathbb{P}(\textit{“Figure”}\,)} \\
&= \frac{\frac{4}{52}}{\frac{12}{52}} = \frac{4}{12} = \frac{1}{3}
\end{aligned}$$

$$\mathbb{P}(\textit{“Roi”}\,|\textit{“Figure”}\,) \neq \mathbb{P}(\textit{“Roi”}\,)$$

Ils ne sont PAS indépendants !

Événements mutuellement exclusifs vs. événements indépendants

Événements indépendants	*Événements mutuellement exclusifs*
$\mathbb{P}(A \text{ et } B) = \mathbb{P}(A) \cdot \mathbb{P}(B)$	$\mathbb{P}(A \text{ et } B) = 0$
$\mathbb{P}(A \mid B) = \mathbb{P}(A)$	$\mathbb{P}(A \mid B) = 0$
$\mathbb{P}(A \text{ ou } B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A) \cdot \mathbb{P}(B)$	$\mathbb{P}(A \text{ ou } B) = \mathbb{P}(A) + \mathbb{P}(B)$

Règles de base

- "Non" $\longrightarrow$ 1 – Probabilité
- "Ou" & Événements mutuellement exclusifs $\longrightarrow$ Additionner
- "Et" & Indépendance $\longrightarrow$ Multiplier

Événements successifs

- $\mathbb{P}(\textit{“Face puis Pile”}) = \left(\frac{1}{2}\right)\left(\frac{1}{2}\right) = \frac{1}{4}$

- $\mathbb{P}(\textit{“Lancer deux 1”}) = \left(\frac{1}{6}\right)\left(\frac{1}{6}\right) = \frac{1}{36}$

- $\mathbb{P}(\textit{“Lancer 1 puis un nombre pair”}) = \left(\frac{1}{6}\right)\left(\frac{1}{2}\right) = \frac{1}{12}$

Événements successifs

- A = *"Lancer Face 6 fois"*

$$\begin{aligned}\mathbb{P}(A) &= \frac{1}{2} \times \frac{1}{2} \times \frac{1}{2} \times \frac{1}{2} \times \frac{1}{2} \times \frac{1}{2} \\ &= \left(\frac{1}{2}\right)^6 = \frac{1}{64}\end{aligned}$$

- B = *"Obtenir FPFPPF"*

$$\mathbb{P}(B) = \frac{1}{64}$$

Échantillonnage sans remplacement

- Un bocal contient 6 billes bleues et 6 billes rouges.
- Vous en tirez 3.
- $\mathbb{P}($*“1ère bille est bleue”*$) = \frac{6}{12} = 0.5$
- $\mathbb{P}($*“2ème bille est bleue”* $\mid$ *“1ère bille est bleue”*$) = \frac{5}{11} = 0.45$
- $\mathbb{P}($*“3ème bille est bleue”* $\mid$ *“les 2 premières étaient bleues”*$) = \frac{4}{10} = 0.4$

$\mathbb{P}($*“Les 2 premières sont bleues”*$) = \mathbb{P}($*“1ère bille est bleue”*$)\times$

$\mathbb{P}($*“2ème bille est bleue”* $\mid$ *“ 1ère bille est bleue”* $)$

$= (0.5)(0.45) = 0.225$

Échantillonnage sans remplacement

Tirer deux cartes d'un jeu standard :

A = *"Première carte est un Roi"*
B = *"Deuxième carte est un Roi"*

$$\mathbb{P}(B|A) = \frac{3}{51}$$

$$\begin{aligned}\mathbb{P}(\textit{“Tirer deux Rois”}) &= \mathbb{P}(A \text{ et } B)\\ &= \mathbb{P}(A)\cdot\mathbb{P}(B|A)\\ &= \frac{4}{52}\cdot\frac{3}{51} = 0.0045\end{aligned}$$

Variables Aléatoires

Une *variable aléatoire* est une variable dont la valeur dépend de l'issue d'un *phénomène aléatoire*.

Chaque issue correspond à un nombre ou à un état/condition :

Valeur (X)	1	2	3
Probabilité ($\mathbb{P}$)	0.2	0.6	0.2

Nommons la variable aléatoire X :

$$\mathbb{P}(X = 1) = 0.2, \quad \mathbb{P}(X = 2) = 0.6, \quad \mathbb{P}(X = 3) = 0.2.$$

X	1	2	3
$\mathbb{P}$	0.2	0.6	0.2

Lancer une pièce équilibrée 3 fois

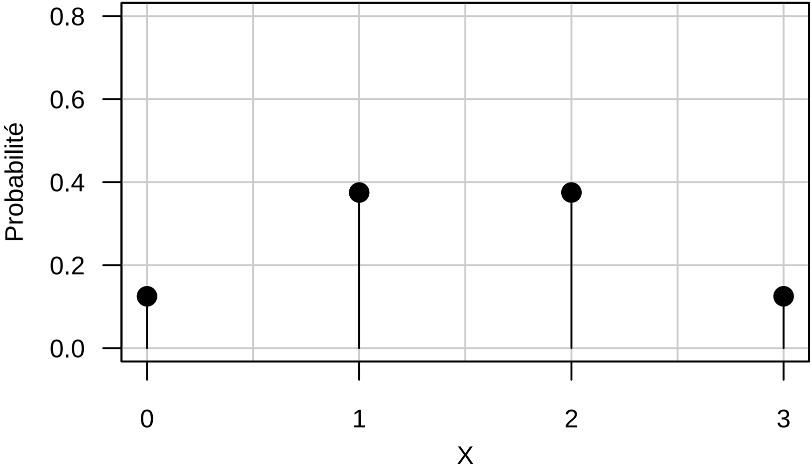
Résultats possibles :

FFF	FFP	FPF	FPP
PFF	PFP	PPF	PPP

X = Nombre de faces (“F”)

Fonction de masse de probabilité (FMP)

k	$X = 0$	$X = 1$	$X = 2$	$X = 3$
$\mathbb{P}(X = k)$	1/8	3/8	3/8	1/8



Plus généralement

X prend des valeurs $\{x_1, \ldots, x_k\}$

FMP

k	$X = x_1$	$\ldots$	$X = x_k$
$\mathbb{P}(X = k)$	p_1	$\ldots$	p_k

$p_1 = P(X = x_1)$,

$\ldots$

$p_k = P(X = x_k)$

Remarques

1. Toutes les probabilités somment à 1.

$$p_1 + \ldots + p_k = 1$$

2. Pour chaque k,

$$0 \leq p_k \leq 1$$

3. Tous les autres nombres ont une probabilité de 0.

Lancer deux dés équilibrés

Résultats :

$(1,1)$	$(1,2)$	$(1,3)$	$(1,4)$	$(1,5)$	$(1,6)$
$(2,1)$	$(2,2)$	$(2,3)$	$(2,4)$	$(2,5)$	$(2,6)$
$(3,1)$	$(3,2)$	$(3,3)$	$(3,4)$	$(3,5)$	$(3,6)$
$(4,1)$	$(4,2)$	$(4,3)$	$(4,4)$	$(4,5)$	$(4,6)$
$(5,1)$	$(5,2)$	$(5,3)$	$(5,4)$	$(5,5)$	$(5,6)$
$(6,1)$	$(6,2)$	$(6,3)$	$(6,4)$	$(6,5)$	$(6,6)$

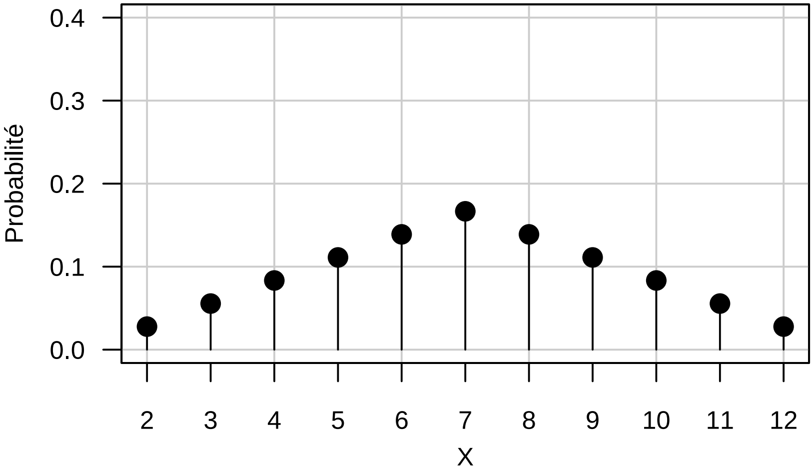
X = somme des deux dés

FMP

k	2	3	4	5	6	7	8	9	10	11	12
$\mathbb{P}(X=k)$	$\frac{1}{36}$	$\frac{1}{18}$	$\frac{1}{12}$	$\frac{1}{9}$	$\frac{5}{36}$	$\frac{1}{6}$	$\frac{5}{36}$	$\frac{1}{9}$	$\frac{1}{12}$	$\frac{1}{18}$	$\frac{1}{36}$

Lancer deux dés équilibrés

k	2	3	4	5	6	7	8	9	10	11	12
$\mathbb{P}(X=k)$	$\frac{1}{36}$	$\frac{1}{18}$	$\frac{1}{12}$	$\frac{1}{9}$	$\frac{5}{36}$	$\frac{1}{6}$	$\frac{5}{36}$	$\frac{1}{9}$	$\frac{1}{12}$	$\frac{1}{18}$	$\frac{1}{36}$



Événement $A = \{X = k\}$

- Les événements $\{X = k\}$ et $\{X = j\}$ sont **mutuellement exclusifs**.
- Règle simple du “OU” :

$$\mathbb{P}(X = k \text{ ou } X = j) = \mathbb{P}(X = k) + \mathbb{P}(X = j)$$

Entretien d'embauche

On choisi 5 personnes au hasard dans une salle de 100 personnes. Parmi ces 100 personnes, 50 personnes étudient à Genève et 50 personnes étudient à Lausanne.

X = nombre de personnes sélectionnées qui étudient à Genève

On obtient la distribtuon de probabilité suivante (nous verrons comment obtenir ces chiffres par la suite) :

	$X = 0$	$X = 1$	$X = 2$	$X = 3$	$X = 4$	$X = 5$
$\mathbb{P}(X = x)$	0.03125	0.15625	0.3125	0.3125	0.15625	0.03125

Quelle est la probabilité que un ou deux personnes sélectionnées étudient à Genève ?

Entretien d'embauche

	$X=0$	$X=1$	$X=2$	$X=3$	$X=4$	$X=5$
$\mathbb{P}(X=x)$	0.03125	0.15625	0.3125	0.3125	0.15625	0.03125

$$\begin{aligned}
& \mathbb{P}(\text{“\textit{une ou deux personnes sélectionnées étudient à Genève}”}) \\
= \; & \mathbb{P}(X=1 \text{ ou } X=2) \\
= \; & \mathbb{P}(X=1)+\mathbb{P}(X=2) \\
= \; & 0.15625+0.3125 \\
= \; & 0.46875
\end{aligned}$$

Entretien d'embauche

	$X=0$	$X=1$	$X=2$	$X=3$	$X=4$	$X=5$
$\mathbb{P}(X=x)$	0.03125	0.15625	0.3125	0.3125	0.15625	0.03125

Quelle est la probabilité qu'*au moins 2* personnes sélectionnées étudient à Genève ?

$$\mathbb{P}(\text{“}\textit{au moins 2 personnes sélectionnées étudient à Genève}\text{”})$$
$$= \mathbb{P}(\text{“}2 \textit{ ou } 3 \textit{ ou } 4 \textit{ ou } 5 \textit{ personnes sélectionnées étudient à Genève}\text{”})$$

$$\begin{aligned}
\mathbb{P}(X \geq 2) &= \mathbb{P}(X=2 \text{ ou } X=3 \text{ ou } X=4 \text{ ou } X=5) \\
&= \mathbb{P}(X=2) + \mathbb{P}(X=3) + \mathbb{P}(X=4) + \mathbb{P}(X=5) \\
&= 0.3125 + 0.3125 + 0.15625 + 0.03125 \\
&= 0.8125
\end{aligned}$$

On peut aussi obtenir plus simplement :

$$\begin{aligned}
\mathbb{P}(X \geq 2) &= 1 - \mathbb{P}(X < 2) \\
&= 1 - \mathbb{P}(X=0) - \mathbb{P}(X=1) = 1 - 0.03125 - 0.15625 = 0.8125
\end{aligned}$$

Notation Sigma

Une abréviation pour écrire de longues sommes

$$\sum_{k=1}^{5} k = 1 + 2 + 3 + 4 + 5$$

$$\begin{aligned}\sum_{k=1}^{5} k^2 &= 1^2 + 2^2 + 3^2 + 4^2 + 5^2 \\ &= 1 + 4 + 9 + 16 + 25\end{aligned}$$

$$\begin{aligned}\sum_{0 \leq k \leq 3} (k-1)^2 &= (0-1)^2 + (1-1)^2 + (2-1)^2 + (3-1)^2 \\ &= 1 + 0 + 1 + 4\end{aligned}$$

Formule fondamentale de la probabilité

- X = variable aléatoire
- A = un ensemble des valeurs possibles de X (un événement)

$$P(\text{“}X \text{ prend une valeur dans } A\text{”}) = \sum_{\text{pour tout } k \text{ qui est dans } A} \mathbb{P}(X = k).$$

Formule fondamentale de la probabilité

- X = variable aléatoire
- A = un ensemble des valeurs possibles de X (un événement)

$$P(\text{“}X \text{ prend une valeur dans } A\text{”}) = \sum_{\text{pour tout } k \text{ qui est dans } A} \mathbb{P}(X = k).$$

k sont les résultats

probabilités de la FMP

Exemple de l'entretien d'embauche

Événement	Probabilité	Notation $\sum$	Calcul
Au moins 3	$\mathbb{P}(X \geq 3)$	$\sum_{k \geq 3} \mathbb{P}(X = k)$	$\mathbb{P}(X = 3) + \mathbb{P}(X = 4)$ $+\mathbb{P}(X = 5)$
Au plus 3	$\mathbb{P}(X \leq 3)$	$\sum_{k \leq 3} \mathbb{P}(X = k)$	$\mathbb{P}(X = 0) + \mathbb{P}(X = 1)$ $+\mathbb{P}(X = 2) + \mathbb{P}(X = 3)$
Moins de 3	$\mathbb{P}(X < 3)$	$\sum_{k < 3} \mathbb{P}(X = k)$	$\mathbb{P}(X = 0) + \mathbb{P}(X = 1)$ $+\mathbb{P}(X = 2)$
Plus de 3	$\mathbb{P}(X > 3)$	$\sum_{k > 3} \mathbb{P}(X = k)$	$\mathbb{P}(X = 4) + \mathbb{P}(X = 5)$

Exemple

Suposons que nous avons la distribution suivante :

	$X = 0$	$X = 1$	$X = 2$	$X = 3$	$X = 4$
$\mathbb{P}(X = x)$	0.0625	0.25	0.375	0.25	0.0625

Nous voulons calculer la probabilité suivante :

$$\begin{aligned} & \mathbb{P}(X \leq 3) \\ = \quad & \mathbb{P}(X = 0) + \mathbb{P}(X = 1) + \mathbb{P}(X = 2) + \mathbb{P}(X = 3) \\ = \quad & 0.0625 + 0.25 + 0.375 + 0.25 = \mathbf{0.9375}. \end{aligned}$$

Exemple

	$X = 0$	$X = 1$	$X = 2$	$X = 3$	$X = 4$
$\mathbb{P}(X = x)$	0.0625	0.25	0.375	0.25	0.0625

Une autre façon de calculer cette probabilité est d'utiliser la règle du complément :

$$\begin{aligned} \mathbb{P}(X \leq 3) &= 1 - \mathbb{P}(X > 3) \\ &= 1 - \mathbb{P}(X = 4) \\ &= 1 - 0.0625 = \mathbf{0.9375}. \end{aligned}$$

Espérance

$$\mathbb{E}(X) = \sum_{\text{tout } k} k \cdot \mathbb{P}(X = k)$$

Interprétation

1. Une moyenne pondérée par les probabilités.
2. La moyenne à long terme de X.
3. La valeur juste d'un pari.
4. Le point d'équilibre pour un histogramme ou un diagramme en barres de probabilité.

Comment calculer $\mathbb{E}(X)$

k	2	5
$\mathbb{P}(X = k)$	$1/3$	$2/3$

$$
\begin{aligned}
\mathbb{E}(X) &= 2 \cdot \mathbb{P}(X = 2) + 5 \cdot \mathbb{P}(X = 5) \\
&= 2\,(1/3) + 5\,(2/3) \\
&= 2/3 + 10/3 = 12/3 = 4
\end{aligned}
$$

$\mathbb{E}(X)$

k	2	5
$\mathbb{P}(X = k)$	1/3	2/3