

MAPSI – Cours 1 : Rappels de probabilités et statistiques

Nicolas Thome & Pierre Henri Wuillemin

LIP6 / ISIR – Sorbonne Université, France

- 1 Fonctionnement de l'UE MAPSI
- 2 Variables aléatoires & probabilités : vocabulaire, définitions
- 3 Description d'une population, d'un échantillon
- 4 Variables multiples, loi jointe, conditionnelle
- 5 Indépendance probabiliste
- 6 Conclusion

- **MAPSI** : Méthodes et algorithmes de probabilité et statistique en informatique
 - **Code UE** : MU4IN601
- **Calendrier** :
<https://cal.ufr-info-p6.jussieu.fr/master/>
⇒ Cocher M1-IMA
- **Ressources sur moodle** :
<https://moodle-sciences-24.sorbonne-universite.fr>
- Répartition dans les groupes quasi-figée
- Mail & nouvelles fraîches pour les informations de dernière minute
- **Mattermost** (=slack gratuit) pour échanger sur les problèmes scientifiques
 - <http://tiny.cc/M1DAC24> pour s'inscrire
 - Channel MAPSI

Organisation :

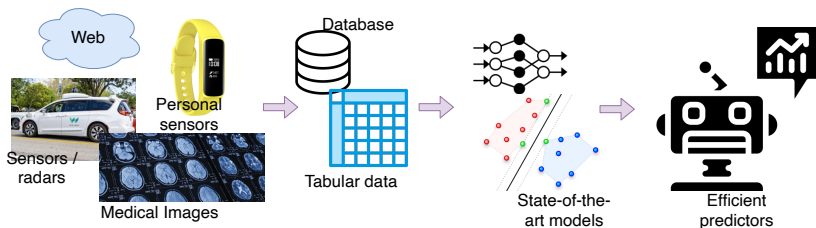
- Cours : théorie & concepts, exemples
- TD : applications & calculs sur feuille
- TME : mise en oeuvre des méthodes sur des exemples concrets

Notation :

- Examen final : 50%
 - Questions sur des formulations analytiques + calcul
 - Questions algo/code
- Partiel : 35%
- Notes de participation (contrôle continu, CC) : 15%
 - Attention : l'essentiel de la note est constitué du travail effectué **durant la séance**
 - **Soumission obligatoire** du code de TME en fin de séance...
 - ... Et **commentaires bienvenus** pour faciliter la correction
- session 2 : $\max(\text{rattrapage}, 15\% \text{ CC} + 85\% \text{ rattrapage})$

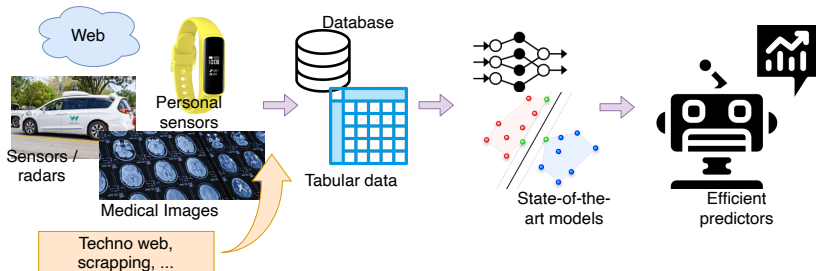
Pourquoi faire MAPSI ?

- Parce que c'est obligatoire
- Parce que c'est un bon rappel de statistiques pour...
 - Comprendre la littérature scientifique en générale
 - Comprendre comment fonctionne l'analyse de données
- Parce que c'est la porte d'entrée vers les sciences des données !



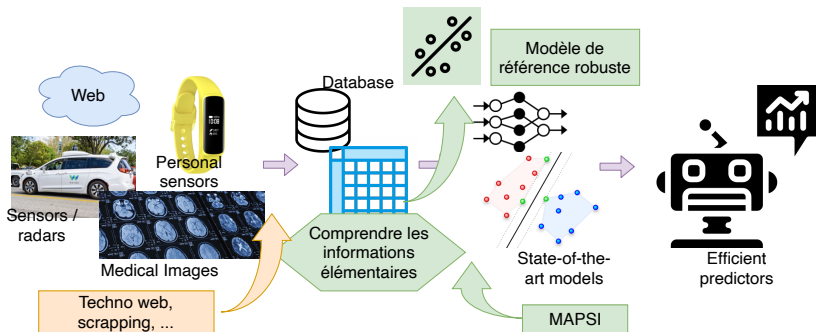
Pourquoi faire MAPSI ?

- Parce que c'est obligatoire
- Parce que c'est un bon rappel de statistiques pour...
 - Comprendre la littérature scientifique en générale
 - Comprendre comment fonctionne l'analyse de données
- Parce que c'est la porte d'entrée vers les sciences des données !



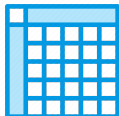
Pourquoi faire MAPSI ?

- Parce que c'est obligatoire
- Parce que c'est un bon rappel de statistiques pour...
 - Comprendre la littérature scientifique en générale
 - Comprendre comment fonctionne l'analyse de données
- Parce que c'est la porte d'entrée vers les sciences des données !



Faire le lien entre une situation réelle et un modèle Statistique

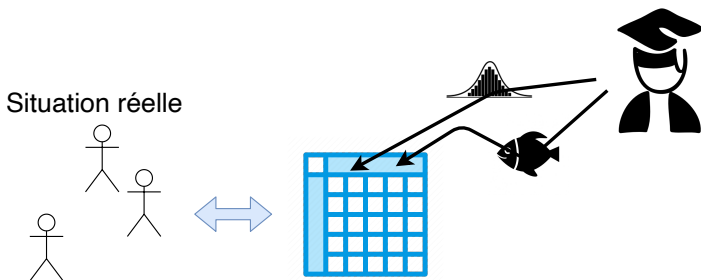
① avec un modèle paramétrique



MAPSI, un peu plus concrètement

Faire le lien entre une situation réelle et un modèle Statistique

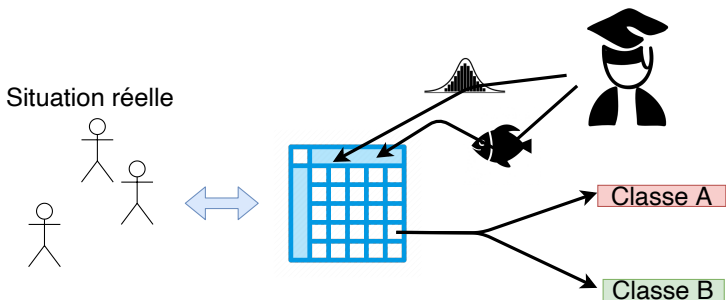
- 1 avec un modèle paramétrique



MAPSI, un peu plus concrètement

Faire le lien entre une situation réelle et un modèle Statistique

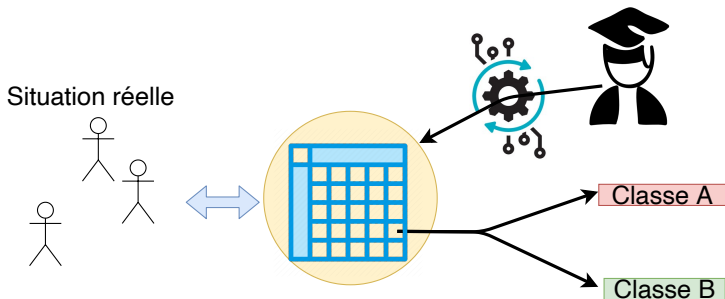
- 1 avec un modèle paramétrique



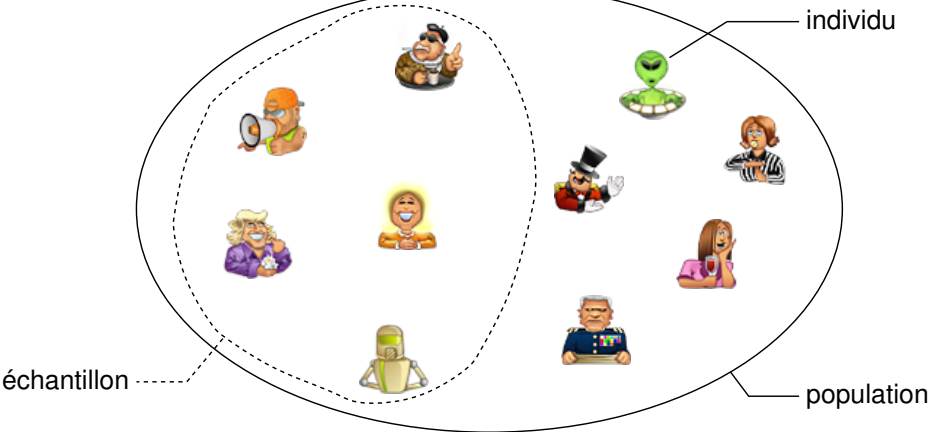
MAPSI, un peu plus concrètement

Faire le lien entre une situation réelle et un modèle Statistique

- 1 avec un modèle paramétrique
- 2 avec un modèle agnostique

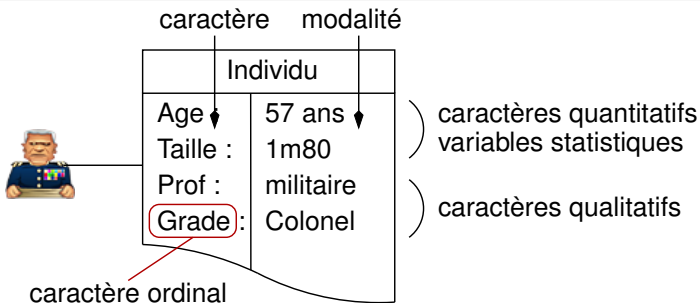


- 1 Fonctionnement de l'UE MAPSI
- 2 Variables aléatoires & probabilités : vocabulaire, définitions**
- 3 Description d'une population, d'un échantillon
- 4 Variables multiples, loi jointe, conditionnelle
- 5 Indépendance probabiliste
- 6 Conclusion



Définitions

- **population (statistique)** : ensemble des objets (ou personnes) sur lesquels porte l'étude
- **individu** : chaque élément de la population



Définitions

- **Caractères** : critères d'étude de la population
- **Modalités** : les valeurs que peuvent prendre les caractères
- **Caractère quantitatif ou Variable statistique** : ensemble de modalités = des nombres + échelle mathématique
- **Caractère qualitatif ou Variable catégorielle** : caractère non quantitatif
- **Caractère ordinal** : les modalités sont ordonnées



Individu	
Age :	57 ans
Taille :	1m80
Prof :	militaire
Grade :	Colonel

— variable discrète

— variable continue

Définitions sur les variables statistiques

- **Variable discrète** : définie sur un espace discret (par exemple des entiers)
- **Variable continue** : définie sur un continuum (toutes les valeurs numériques d'un intervalle)

Quelques définitions de statistiques

- X : caractère défini sur une population de N individus
- $\{x_1, \dots, x_I\}$ modalités de X
- $N_i = \text{effectif}$ de x_i
= nombre d'individus pour lesquels X a pris la valeur x_i
- **fréquence ou effectif relatif** : $f_i = \frac{N_i}{N}$
- **distribution** de X : ensemble des couples $\{(x_1, f_1), (x_2, f_2), \dots\}$
Représentation usuelle sous forme de tableau

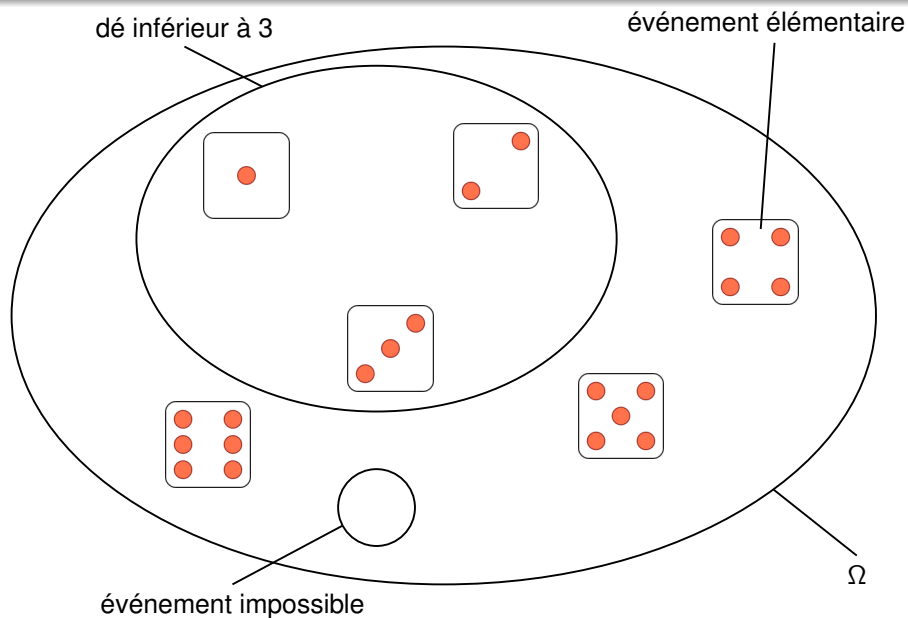
Statistiques = description d'un échantillon

Probabilités = description d'une population.

Notations ensemblistes :

- événements = sous-ensembles de Ω
- \emptyset = événement impossible
- $A \cup B$ = événement qui est réalisé si A ou B est réalisé
- $C \cap D$ = événement qui est réalisé si C et D sont réalisés
- $\overline{C \cup D}$ = complémentaire de $C \cup D$ dans Ω
= événement qui est réalisé ssi $C \cup D$ ne l'est pas
- $A \cap B = \emptyset$ = 2 événements qui ne peuvent se réaliser simultanément

Probabilités : approche événementielle



Les événements : approche évènementielle

événement = sous-ensemble de Ω

A

\emptyset = événement impossible

$$A \cap B = \emptyset$$

B

$A \cup B$

$\overline{C \cup D}$

Ω

C

D

$$C \cap D$$

Définition des probabilités (Kolmogorov)

- Ω = ensemble fini ou dénombrable
d'événements élémentaires $e_k, k \in K \subseteq \mathbb{N}$
- $\mathcal{A} = 2^{|\Omega|}$ = ensemble des événements
- Mesure de probabilité : $\forall A \in \mathcal{A}, 0 \leq P(A) \leq 1$
- $P(\Omega) = 1$
- $A = \bigcup_{k \in L} A_k$, avec L ensemble dénombrable et, $\forall j, k \in L, j \neq k, A_j \cap A_k = \emptyset$,
$$P(A) = \sum_{k \in L} P(A_k).$$

\Rightarrow Les probabilités des événements élémentaires déterminent entièrement P

conséquence : $P(A \cup B) = P(A) + P(B) - P(A \cap B)$

Quelques définitions de probabilités

- **Univers** Ω : un ensemble dénombrable (fini ou infini)
- **Evènement** : un sous-ensemble de l'univers Ω
- **Mesure de probabilité** :
une fonction qui associe à chaque évènement une valeur entre 0 et 1, la probabilité de Ω est 1, et la probabilité d'une union dénombrable d'évènements incompatibles (ensembles disjoints) est la somme de leurs probabilités.
- **Espace probabilisé** :
un couple (Ω, P) où P est une mesure de probabilité sur Ω .
- **Variable aléatoire**

Variable aléatoire

Lorsqu'on est face à une expérience aléatoire, on s'intéresse plus souvent à une *valeur* attribuée au résultat qu'au résultat lui-même.

Exemples

- Lorsque l'on joue à un jeu de hasard on s'intéresse plus au gain que l'on peut obtenir qu'au résultat du jeu.
 - Nombre de pannes dans un ensemble de systèmes plutôt que l'état exact des systèmes
- Solution : “traduire” l'univers en évènements “compréhensibles”.
⇒ variable aléatoire : application de l'univers Ω vers un autre ensemble.

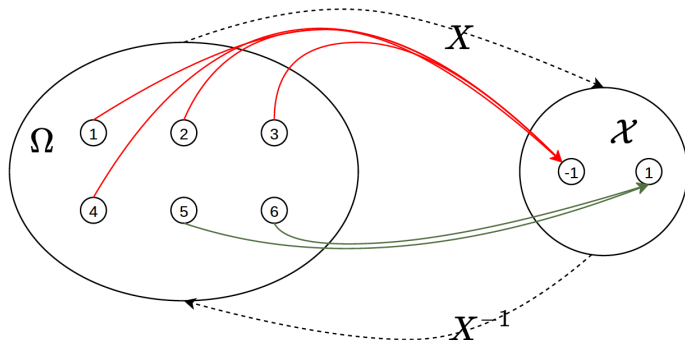
Exemple du lancer de dé

On lance un dé après avoir misé 1 EUR. Si le résultat est un 5 ou un 6 on double la mise, sinon perd la mise. Dans ce cas :

- $\Omega = \{1, 2, 3, 4, 5, 6\}$
- $\text{Card } \Omega = 6$, et $\forall e \in \Omega, P(e) = \frac{1}{6}$
- Soit X la v.a. qui associe à tout résultat du dé un *gain* :
 $X(1) = X(2) = X(3) = X(4) = -1$ et $X(5) = X(6) = (2 - 1) = 1$
 X est à valeur dans l'ensemble noté $\mathcal{X} = \{-1, 1\} \subset \mathbb{R}$,
 $X : \Omega \rightarrow \mathcal{X}$
- Question :
Comment calculer la probabilité de gagner 1 EUR ?
- Réponse : Définir une probabilité sur \mathcal{X} , notée \mathbb{P} , en retournant dans l'espace probabilisé $(\Omega, \mathcal{P}(\Omega), P)$
i.e. utiliser $P(\text{résultat du dé} = 5 \text{ ou } 6)$ pour estimer $\mathbb{P}(\{1\})$.

Univers, Evènements et Variable Aléatoire (3/5)

- $\Omega = \{1, 2, 3, 4, 5, 6\}$, $\text{Card } \Omega = 6$, et $\forall e \in \Omega, P(e) = \frac{1}{6}$
- Soit X la v.a. qui associe à tout résultat du dé un *gain* :
 $X(1) = X(2) = X(3) = X(4) = -1$ et $X(5) = X(6) = (2 - 1) = 1$
 X est à valeur dans l'ensemble noté $\mathcal{X} = \{-1, 1\} \subset \mathbb{R}$, $X : \Omega \rightarrow \mathcal{X}$
- Question : **Comment calculer la probabilité de gagner 1 EUR ?**
- Réponse : Définir une probabilité \mathbb{P} sur \mathcal{X} , en retournant dans l'espace probabilisé $(\Omega, \mathcal{P}(\Omega), P)$, i.e. utiliser $P(\text{résultat du dé} = 5 \text{ ou } 6)$ pour estimer $\mathbb{P}(\{1\})$.



Définition Variable aléatoire à valeurs discrètes

Soit Ω un ensemble dénombrable, et P une mesure de probabilité sur Ω .

Soit Ω' , un ensemble discret. Une variable aléatoire est une fonction X de Ω muni de la mesure P vers Ω' .

Exemples

- Lancer d'un dé :

Soit $\Omega = \{1, \dots, 6\}$ muni de la probabilité uniforme P .

$$X : i \mapsto \begin{cases} 1 & \text{si } i \text{ est pair} \\ 0 & \text{sinon} \end{cases}$$

est une variable aléatoire de (Ω, P) vers $\Omega' = \{0, 1\}$.

- Lancer de deux dés :

Soit $\Omega = \{1, \dots, 6\}^2$ muni de la probabilité uniforme P .

$$X : (i, j) \mapsto i + j$$

est une variable aléatoire de (Ω, P) vers $\Omega' = \{2, \dots, 12\}$

Définitions : Loi de probabilité

Soit (Ω, P) un espace probabilisé où Ω est dénombrable.

Soit Ω' un ensemble discret, et X une v.a. de (Ω, P) vers Ω' .

- P_X définit une mesure de probabilité sur Ω' :

$$\forall E' \subset \Omega', \quad P_X(E') = P(X^{-1}(E'))$$

avec $X^{-1}(E') = \{\omega \in \Omega | X(\omega) \in E'\}$

- L'ensemble des valeurs $P_X(\{\omega'\})$ pour $\omega' \in \Omega'$ s'appelle la *loi de probabilité* de X .

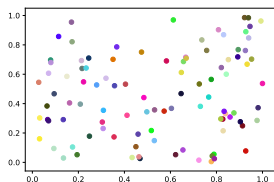
Notations

- L'événement $X \in]-\infty, a]$ sera noté par $X \leq a$
- L'événement $X \in]a, b]$ sera noté par $a < X \leq b$
- L'événement $X \in \{a\}$ sera noté par $X = a$
- On a donc $P_X(B) = P(X^{-1}(B)) = P(X \in B)$

- 1 Fonctionnement de l'UE MAPSI
- 2 Variables aléatoires & probabilités : vocabulaire, définitions
- 3 Description d'une population, d'un échantillon**
- 4 Variables multiples, loi jointe, conditionnelle
- 5 Indépendance probabiliste
- 6 Conclusion

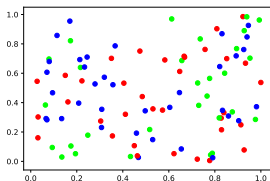
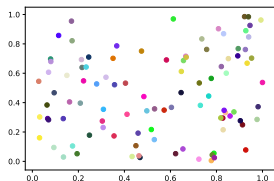
DESCRIPTION D'UNE POPULATION

- A partir d'un échantillon
- En simplifiant les données continues
- Simplifiant les différentes dimensions, ...



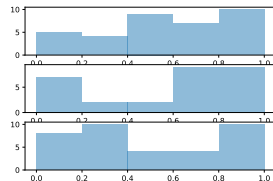
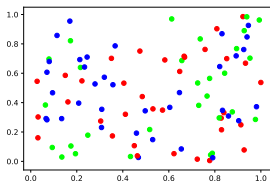
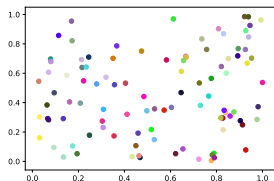
DESCRIPTION D'UNE POPULATION

- A partir d'un échantillon
- En simplifiant les données continues
- Simplifiant les différentes dimensions, ...



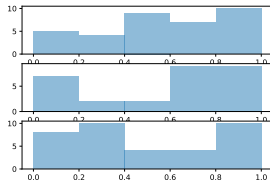
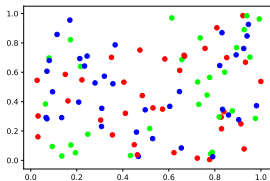
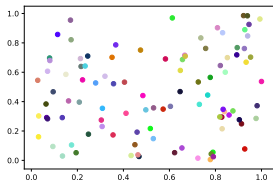
DESCRIPTION D'UNE POPULATION

- A partir d'un échantillon
- En simplifiant les données continues
- Simplifiant les différentes dimensions, ...



DESCRIPTION D'UNE POPULATION

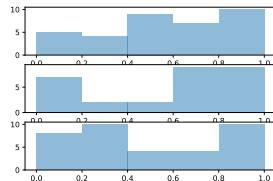
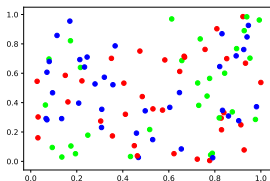
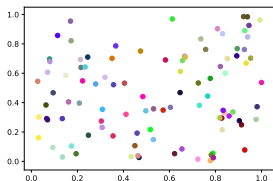
- A partir d'un échantillon
- En simplifiant les données continues
- Simplifiant les différentes dimensions, ...



Moyennes : 0.56 0.59 0.47

DESCRIPTION D'UNE POPULATION

- A partir d'un échantillon
- En simplifiant les données continues
- Simplifiant les différentes dimensions, ...

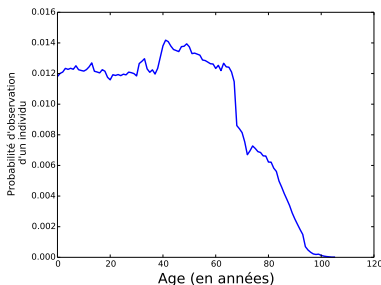


Moyennes : 0.56 0.59 0.47
Ou même une moyenne générale : 0.54

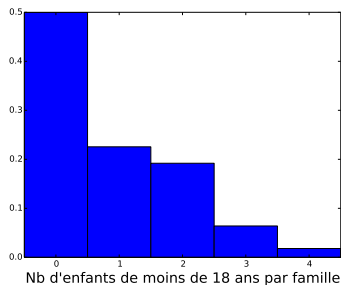
Décrire parfaitement une population = connaître sa loi de probabilité

Exemples selon la nature des variables :

en continu :



en discret :



Décrire parfaitement une population = connaître sa loi de probabilité

Problème général :

Comment déduire la loi sur la population si on ne connaît qu'un échantillon ?

Réponse dans les cours suivants...

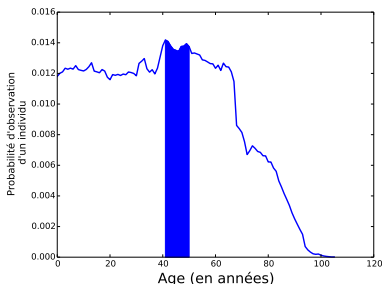
Cas général :

- Une distribution somme à 1
- Une probabilité est toujours ≥ 0

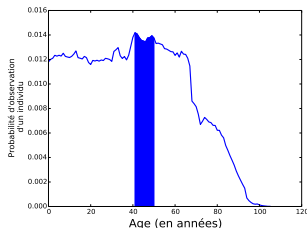
Cas Continu :

- Chaque événement élémentaire a une proba = 0
(eg : proba d'avoir 40 ans)
- Mais proba d'être dans un intervalle ≥ 0
(eg : proba d'avoir entre 40 et 50 ans)

$P(A)$ = surface délimitée par la
fonction de densité dans la zone
où les événements sont inclus
dans A



Probabilités : les détails dans le cas continu



$$P(X \in I) = \int_I p(x) dx$$

avec P = proba et p = fonction de densité

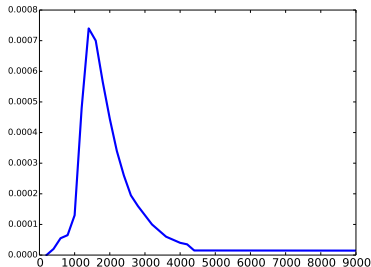
\Rightarrow connaître p = connaître P

intervalles $] -\infty, x[\Rightarrow$ fonction de répartition :

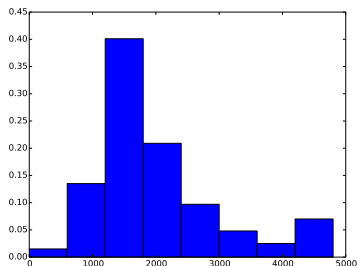
$$F(x) = P(X < x) = \int_{-\infty}^x p(y) dy$$

Continu \rightarrow Discret

Possibilité de discrétisation = regroupement par tranches
Modélisation approximative, mais manipulation plus facile



Distribution des salaires en France



Distribution des salaires en France
discrétisée

Caractéristiques

- **Espérance mathématique** ou **moyenne** : $E(X)$

X discrète : $E(X) = \sum x_k p_k$

X continue : $E(X) = \int x p(x) dx$



l'espérance mathématique n'existe pas toujours

- **Mode** : Mo de P (pas toujours unique) :

X discrète : $p(Mo) = \max_k p(x_k)$

X continue : $p(Mo) = \max_x p(x)$

Propriétés de l'espérance

- $E(aX + b) = aE(X) + b$
- $\forall X, Y, E(X + Y) = E(X) + E(Y)$

- **variance** : $V(X)$ ou σ^2 :

X discrète : $\sigma^2 = \sum [x_k - E(X)]^2 p_k$

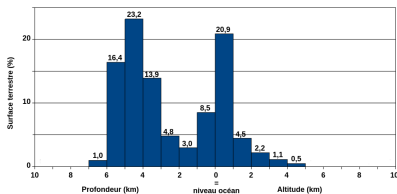
X continue : $\sigma^2 = \int [x - E(X)]^2 p(x) dx$

- moyenne des carrés des écarts entre les valeurs prises par X et son espérance $E(X)$
- **écart-type** : σ = racine carrée de la variance
- variance et donc écart-type n'existent pas toujours
- **Prop** : $Y = aX + b$, où a et b sont des nombres réels
 $V(Y) = a^2 V(X)$
- **Prop** : $V(X) = E[X^2] - E[X]^2$

Résumer les informations prépondérantes

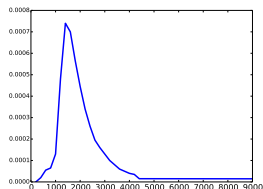
Idées :

- 1 Caractériser rapidement une distribution
 - 2 Donner en quelques chiffres une idée approchée de l'ensemble de la distribution de probabilité.
- Espérance, variance + moments statistiques



Niveau du sol sur la planète

- Quantiles (médianes, ...)



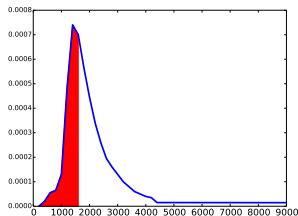
Médiane = 1650, Moyenne = 2000

Médiane d'une variable statistique continue

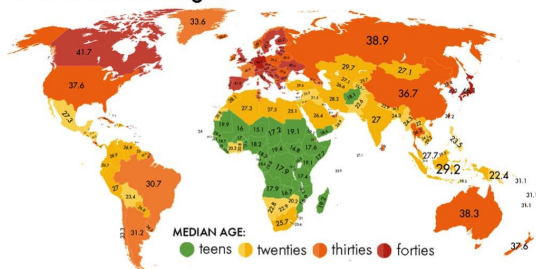
Médiane d'une variable statistique continue

- X : variable statistique continue
- **Médiane** = le nombre δ tel que les aires situées de part et d'autre de ce nombre dans l'histogramme représentant X sont égales

Médiane M : $P(X \leq M) \geq \frac{1}{2}$ et $P(X \geq M) \geq \frac{1}{2}$



World Median Ages



YOUNGEST: 1. Niger (15.1) 2. Uganda (15.5) 3. Mali (16) 4. Malawi (16.3) 5. Zambia (16.7)
OLDEST: 1. Germany & Japan (46.1) 2. Italy (44.5) 3. Austria (44.3) 4. Virgin Islands (44.2)

Quantile d'une variable discrète

- X : variable statistique discrète, modalités $\{x_1, \dots, x_I\}$
- population de N individus ($N_i = \text{effectif de } x_i$)
- **quantile d'ordre** $\alpha = \delta$ tel que :

$$\sum_{i \in \{j: x_j < \delta\}} N_i \leq \alpha N \quad \text{et} \quad \sum_{i \in \{j: x_j > \delta\}} N_i \leq (1 - \alpha)N$$

Quantile d'une variable continue

- X : variable statistique continue
- **quantile d'ordre** $\alpha =$ le nombre δ tel que les aires situées de part et d'autre de ce nombre dans l'histogramme représentant X sont égales respectivement à $\alpha \times \text{aire totale}$ et $(1 - \alpha) \times \text{aire totale}$

- 1 Fonctionnement de l'UE MAPSI
- 2 Variables aléatoires & probabilités : vocabulaire, définitions
- 3 Description d'une population, d'un échantillon
- 4 Variables multiples, loi jointe, conditionnelle**
- 5 Indépendance probabiliste
- 6 Conclusion

- Chaque individu de la population est décrit sur plusieurs caractères

Exemples :

- Carte à jouer : Couleur (Trefle, Carreau, Coeur, Pique), Valeur (7, ..., Roi)
- Sportifs : Age (< 20 , ..., > 50), Sport pratiqué (Natation...)

Loi jointe $P(A, B)$: décrire toutes les intersections possibles

Exemple :

Sport \ Age	< 20	$[20, 30[$	$[30, 40[$	$[40, 50[$	≥ 50
Natation	0.02	0.05	0.09	0.08	0.08
Jogging	0.10	0.15	0.10	0.07	0.05
Tennis	0.02	0.03	0.06	0.07	0.03

NB : l'ensemble de l'univers Ω somme toujours à 1

Définition

La marginalisation consiste à projeter une loi jointe sur l'une des variables aléatoires.

Par exemple, extraire $P(A)$ à partir de $P(A, B)$.

$$P(A) = \sum_i P(A, B = b_i)$$

Sport \ Age	< 20	[20, 30[[30, 40[,	[40, 50[≥ 50	Marginale
Natation	0.02	0.05	0.09	0.08	0.08	0.32
Jogging	0.10	0.15	0.10	0.07	0.05	0.47
Tennis	0.02	0.03	0.06	0.07	0.03	0.21
						1

La marginale extraite ici correspond à $P(\text{Sport})$. La nouvelle loi somme toujours à 1.

Définition

La marginalisation consiste à projeter une loi jointe sur l'une des variables aléatoires.

Par exemple, extraire $P(A)$ à partir de $P(A, B)$.

$$P(A) = \sum_i P(A, B = b_i)$$

Sport \ Age	< 20	[20, 30[[30, 40[,	[40, 50[≥ 50
Natation	0.02	0.05	0.09	0.08	0.08
Jogging	0.10	0.15	0.10	0.07	0.05
Tennis	0.02	0.03	0.06	0.07	0.03
Marginale	0.14	0.23	0.25	0.22	0.16

Définition

la probabilité d'un événement A conditionnellement à un événement B , que l'on note $P(A|B)$, est la probabilité que A se produise sachant que B s'est ou va se produire.

Rem : $P(A|\Omega) = P(A)$ puisqu'on sait que Ω sera réalisé

Problème : comment calculer $P(A|B)$?

Exemple

- Tirer une carte parmi un jeu de 32 cartes
- $\Omega = \{32 \text{ cartes}\}$
- événements : $A = \text{tirer un roi}$ $B = \text{tirer un cœur}$
- $P(A|B) = ?$

Interprétation de $P(A|B)$

Dans l'univers réduit B ($\Omega' = B$), quelle est la probabilité de A ?

Exemple

- Tirer une carte parmi un jeu de 32 cartes
- $\Omega = \{32 \text{ cartes}\}$
- événements : $A = \text{tirer un roi}$ $B = \text{tirer un cœur}$
- $P(A|B) = ?$

Interprétation de $P(A|B)$

Dans l'univers réduit B ($\Omega' = B$), quelle est la probabilité de A ?

- $\Omega' = B = \text{cœur}$ (8 cartes)
- $P(A|B) = \text{un roi parmi les cœur...}$

Exemple

- Tirer une carte parmi un jeu de 32 cartes
- $\Omega = \{32 \text{ cartes}\}$
- événements : $A = \text{tirer un roi}$ $B = \text{tirer un cœur}$
- $P(A|B) = ?$

Interprétation de $P(A|B)$

Dans l'univers réduit B ($\Omega' = B$), quelle est la probabilité de A ?

- $\Omega' = B = \text{cœur}$ (8 cartes)
- $P(A|B) = \text{un roi parmi les cœur...}$

$$P(A|B) = \frac{1}{8}$$

Théorème des probabilités totales :

En partant de la loi jointe

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

Interprétation : l'observation conjointe de A et B ($P(A, B)$) correspond à l'observation de B ET à l'observation de A dans l'univers restreint B .

Théorème des probabilités totales :

En partant de la loi jointe

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

Interprétation : l'observation conjointe de A et B ($P(A, B)$) correspond à l'observation de B ET à l'observation de A dans l'univers restreint B .

Exemple : Roi de coeur = Observer un coeur ET observer un roi dans l'univers des coeurs

Propriétés

- Réversible : $P(A, B) = P(A|B)P(B) = P(B|A)P(A)$
- Théorème de Bayes :

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

- Intégration des probabilités totales :

$$P(A) = \sum_i P(A, B = b_i) = \sum_i P(A|B = b_i)P(B_i)$$

Tableau de probabilité conditionnelle :

- Sport : $P(S) =$

Natation	0.32
Jogging	0.47
Tennis	0.21

- Répartition des ages pour chaque sport :
 $P(A|S) =$

Sport \ Age	< 20	[20, 30[[30, 40[,	[40, 50[≥ 50
Natation	0.06	0.16	0.28	0.25	0.25
Jogging	0.21	0.32	0.21	0.15	0.11
Tennis	0.10	0.14	0.29	0.33	0.14

- **Propriété** : chaque ligne somme à 1 (=chaque ligne est un univers à part)
- **Questions** : comment extraire la distribution des ages ?
Comment obtenir la distribution jointe ?

- 1 Fonctionnement de l'UE MAPSI
- 2 Variables aléatoires & probabilités : vocabulaire, définitions
- 3 Description d'une population, d'un échantillon
- 4 Variables multiples, loi jointe, conditionnelle
- 5 Indépendance probabiliste**
- 6 Conclusion

Définition de l'indépendance

deux événements A et B sont indépendants si :

$$P(A, B) = P(A) \times P(B)$$

Corrolaire : deux événements A et B sont indépendants si :

$$P(A|B) = P(A) \text{ (avec } P(B) > 0)$$

l'indépendance n'est pas une propriété du couple (A, B) mais du couple $(\{A, A^c\}, \{B, B^c\})$:

A et B sont indépendants $\implies A$ et B^c indépendants
 $\implies A^c$ et B indépendants
 $\implies A^c$ et B^c indépendants

Démonstration :

A et B sont indépendants $\implies P(A, B) = P(A) \times P(B)$

$$P(A) = P(A, B) + P(A, B^c)$$

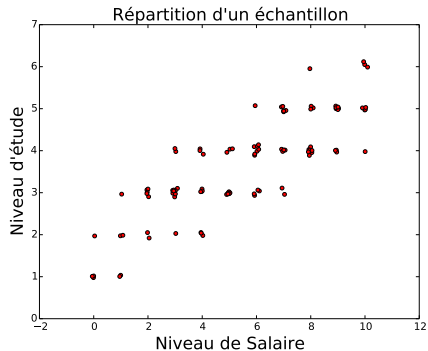
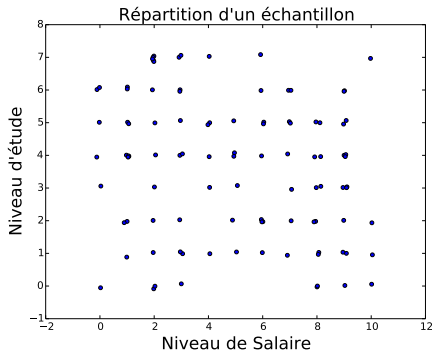
$$\implies P(A, B^c) = P(A) - P(A, B)$$

$$= P(A) - P(A) \times P(B)$$

$$= P(A) \times [1 - P(B)]$$

$$= P(A) \times P(B^c)$$

Exemple



Qu'est ce qui est dépendant ou indépendant ?

Indépendance (exemple)

Indépendance ou pas entre X et Y ?

Cas 1 :

	y_1	y_2
x_1	0.04	0.06
x_2	0.36	0.54

Cas 3 :

	y_1	y_2	y_3
x_1	0.07	0.24	0.16
x_2	0.07	0.30	0.16

Cas 2 :

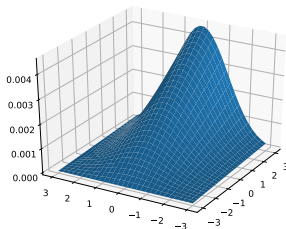
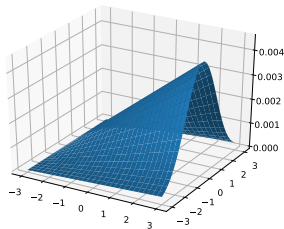
	y_1	y_2
x_1	0.1	0.2
x_2	0.3	0.4

Cas 4 :

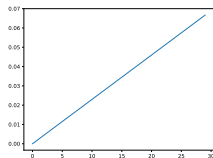
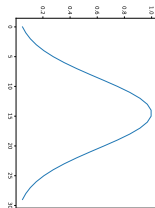
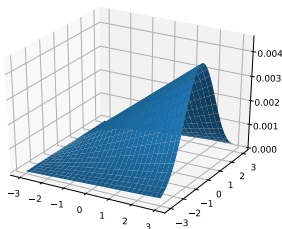
	y_1	y_2	y_3
x_1	0.01	0.02	0.07
x_2	0.09	0.18	0.63

Indépendance (exemple graphique)

Représentation d'une loi jointe $P(X_1, X_2)$

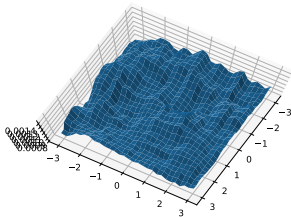
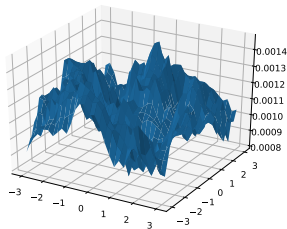


Représentation d'une loi jointe $P(X_1, X_2)$



Indépendance (exemple graphique)

Représentation d'une loi jointe $P(X_1, X_2)$



- Combien de variable pour modéliser la probabilité de voir son toit s'envoler ?
 - Mois de l'année (12)
 - Catégorie des ouragans (6)
 - Type de construction (10)

- Combien de variable pour modéliser la probabilité de voir son toit s'envoler ? $12 \times 6 \times 10 = 720$
 - Mois de l'année (12)
 - Catégorie des ouragans (6)
 - Type de construction (10)

- Combien de variable pour modéliser la probabilité de voir son toit s'envoler ? $12 \times 6 \times 10 = 720$
 - Mois de l'année (12)
 - Catégorie des ouragans (6)
 - Type de construction (10)
- Combien de variable pour modéliser les probabilités de tirage de 3 dés (cumul) ?

- Combien de variable pour modéliser la probabilité de voir son toit s'envoler ? $12 \times 6 \times 10 = 720$
 - Mois de l'année (12)
 - Catégorie des ouragans (6)
 - Type de construction (10)
- Combien de variable pour modéliser les probabilités de tirage de 3 dés (cumul) ?
 - Indépendance ! Dés identiques = 6 valeurs
 - Dés différents = $3 \times 6 = 18$ valeurs
 - Dés non indépendants (?) $6 \times 6 \times 6 = 216$ valeurs

Indépendance : pourquoi c'est important

- Combien de variable pour modéliser la probabilité de voir son toit s'envoler ? $12 \times 6 \times 10 = 720$
 - Mois de l'année (12)
 - Catégorie des ouragans (6)
 - Type de construction (10)
- Combien de variable pour modéliser les probabilités de tirage de 3 dés (cumul) ?
 - Indépendance ! Dés identiques = 6 valeurs
 - Dés différents = $3 \times 6 = 18$ valeurs
 - Dés non indépendants (?) $6 \times 6 \times 6 = 216$ valeurs
- Combien de variables pour modéliser les probabilités d'apparition de groupes de 3 mots (tri-grammes) ? -Vocabulaire réduit à 10k mots-

Indépendance : pourquoi c'est important

- Combien de variable pour modéliser la probabilité de voir son toit s'envoler ? $12 \times 6 \times 10 = 720$
 - Mois de l'année (12)
 - Catégorie des ouragans (6)
 - Type de construction (10)
- Combien de variable pour modéliser les probabilités de tirage de 3 dés (cumul) ?
 - Indépendance ! Dés identiques = 6 valeurs
 - Dés différents = $3 \times 6 = 18$ valeurs
 - Dés non indépendants (?) $6 \times 6 \times 6 = 216$ valeurs
- Combien de variables pour modéliser les probabilités d'apparition de groupes de 3 mots (tri-grammes) ? -Vocabulaire réduit à 10k mots-
 - $10k^3 = 10^{12}$ valeurs (=4000 Go)

Propriétés de la variance

- $\forall X, Y, V(X + Y) = V(X) + V(Y) + 2cov(X, Y)$ où :
 - $cov(X, Y) = \text{covariance}$ de X et Y

$$COV(X, Y) = E[(X - E[X])(Y - E[Y])]$$

- si X et Y discrètes

$$COV(X, Y) = \sum_i \sum_j (x_i - E[X])(y_j - E[Y])P(X = x_i, \cap Y = y_j)$$

- si X et Y continues, de densité $p(x, y)$,

$$cov(X, Y) = \iint [x - E(X)][y - E(Y)]p(x, y)dx dy$$

Estimateur sur un échantillon : $\{(x_1, y_1), \dots, (x_k, y_k), \dots, (x_N, y_N)\}$

$$cov(X, Y) = \frac{1}{N} \sum_k (x_k - \bar{x})(y_k - \bar{y})$$

Définition : Coefficient de corrélation linéaire

Soit X, Y deux variables. Le coefficient de corrélation linéaire entre X et Y est :

$$r = \frac{\text{cov}(X, Y)}{\sigma_X \sigma_Y},$$

où σ_X et σ_Y sont les écart-types des variables X et Y .

- 1 Fonctionnement de l'UE MAPSI
- 2 Variables aléatoires & probabilités : vocabulaire, définitions
- 3 Description d'une population, d'un échantillon
- 4 Variables multiples, loi jointe, conditionnelle
- 5 Indépendance probabiliste
- 6 Conclusion

Les 4 règles qui vont vous sauver la vie.

- Probabilité :

$$\forall x \in \mathcal{X}, \quad 0 \leq P(x) \leq 1 \text{ et } \sum_{x \in \mathcal{X}} p(x) = 1$$

- Marginalisation

$$P(X_1 = x_1) = \sum_{x_2, \dots, x_n} P(X_1 = x_1, X_2 = x_2, \dots)$$

- Conditionnement :

$$P(X_1, X_2) = P(X_1|X_2)P(X_2) \text{ (et vice versa)}$$

- Indépendance :

Si X_1 et X_2 sont indépendantes : $P(X_1, X_2) = P(X_1)P(X_2)$