Mathématiques en technologie de l’information 1

Support du cours

Enseignant : Niklaus Eggenberg

Email : niklaus.eggenberg@hesge.ch

Support du cours : www.eswys.ch/hepia

Probabilités et statistiques

La seule certitude, c’est que rien n’est certain.

Pline L’Ancien (Italie, 1er siècle après J.-C.)

Probabilités et Statistiques

- Théorie des probabilités

  étude et modélisation mathématique
  d'évènements à traits aléatoires ou hasardeux

- Théorie des Statistiques

  mesure et analyse d'un ensemble de données

Où trouve-t-on les probabilités

- Jeux de hasard

- Prévision financières ou météorologiques

- Modèles physiques
  - Physique des particules

Pourquoi étudier l’incertain

De nombreux évènements semblent dépendre du hasard.

Pourtant, il existe des différences – un certain «ordre dans le désordre» - dans certains cas, parfois pas.

Le probabilités modélisent certaines de ces situations et permettent de mieux en comprendre le comportement.

Les statistiques mesurent dés réalisations d’un évènement pour tenter d’en déduire le modèle probabiliste.

Quelle différence

Probabilités :

Supposons un dé à 6 faces de 1 à 6. Nous avons alors, à chaque lancé, une probabilité de 1 chance sur 6 que le dé donne un 6.

Statistiques :

Nous avons un dé à 6 faces et le lançons 1000 fois, et comptons le nombre de fois que nous obtenons 1, 2, …, 6.
Le dé est-il pipé ?

Probabilités - formalisation

Un modèle probabiliste est basé sur les notions suivantes :

Ω	l'***ensemble réalisable***, espace des possibles ou encore ***univers***
$\omega \in \Omega$	une réalisation de l'ensemble réalisable
$\mathbb{P}$	la ***fonction de probabilité***
	$\mathbb{P} : \Omega \to [0, 1]$.

Probabilités discrètes vs continues

Dans le cas général, il n'y a pas de contraintes sur l'univers Ω.

Nous distinguerons un cas très particuliers :

Les ***probabilités discrètes*** se basent sur des univers Ω finis ou dénombrables.

Les probabilités sur des univers non-dénombrables (dits ***continus***) existent, mais dépassent le cadre de ce cours, bien que quelques exemples illustratifs soient possibles.

Exemples

Probabilités Discrètes	Probabilités Continues
Lancé de dés à N faces numérotées de 1 à N	Durée de vie d'une ampoule
Tirage du loto (6 boules parmi 40)	Taille d'une personne

<u>Fait intéressant :</u> dans les probabilités continues, la probabilité qu'un élément précis de l'univers se réalise est toujours 0 !
Le contenu du cours devrait vous permettre de comprendre pourquoi !

Exercice

Définissez l'ensemble des réalisables, un ensemble de réalisations ainsi qu'une fonction de probabilités pour les deux cas suivants.

Note : dans les deux cas, la définition ne suffit pas à déterminer uniquement la fonction de probabilité. Il vous suffit d'en proposer une !

Probabilités Discrètes	**Probabilités Continues**
Lancé de dés à N faces numérotées de 1 à N	Durée de vie d'une ampoule

Corrigé – Lancé de dé à *N* faces

$\Omega = \{1, 2, \dots, N\}$ est l'ensemble réalisable.

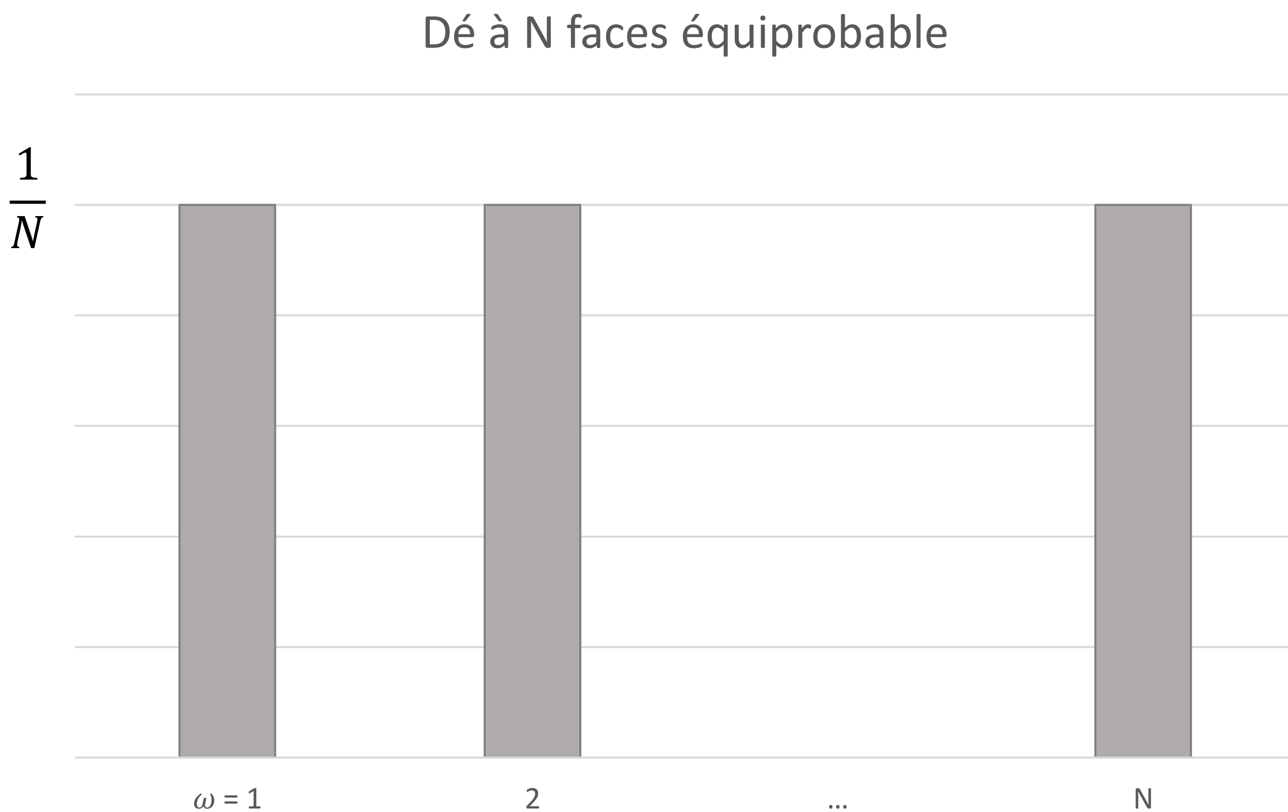
Une réalisation est un lancé, donc une valeur de 1 à N :

$\omega = 3$ est une réalisation (pour $N \geq 3$).

Si le dé n'est pas pipé, chaque réalisation est ***équiprobable***, donc la fonction de probabilité est constante :

$$\mathbb{P} : \Omega \to [0, 1] \text{ et } \mathbb{P}(\omega) = \frac{1}{N}, \forall \omega \in \Omega\,.$$

Corrigé – Lancé de dé à *N* faces



Corrigé – Durée de vie d'une ampoule

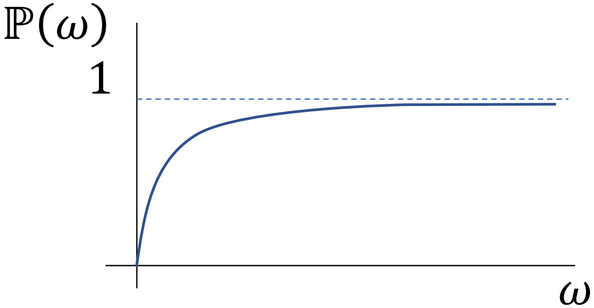
$\Omega = \mathbb{R}^+$ est l'ensemble réalisable.

Une réalisation le temps écoulé jusqu'à ce que l'ampoule casse (nous dirons ici en secondes) soit :

$\omega = 3600s.$ ou $\omega = \sqrt{2}\, s.$ sont des réalisations possibles.

Voici une fonction de probabilités possible

$$\mathbb{P} : \Omega = \mathbb{R}^+ \to [0,1] \text{ et}$$
$$\mathbb{P}(\omega) = 1 - \frac{1}{\omega+1}, \forall \omega \in \Omega\ .$$



Discret vs continu – nuance !

Notez la différence PRIMORDIALE entre le cas discret et le cas continu.

Dans le cas discret, $\mathbb{P}(\omega)$ donne la probabilité que le lancé de dé donne la valeur ω.

Dans le cas continu, $\mathbb{P}(\omega)$ n'est pas la probabilité que l'ampoule se casse exactement au temps ω, mais donne en fait la probabilité que l'ampoule soit cassée après ω secondes.

Nous comprendrons plus tard ce qu'est cette différence !

L’importance de l’univers

L’univers Ω est primordial pour l’étude d’un modèle probabiliste.

Il doit impérativement contenir ***TOUTES*** les issues possibles une et ***UNE ET UNE SEULE FOIS***.

Si le dé contenait 2 fois le chiffre 1 et pas le chiffre N, alors il faudra impérativement que l’on définisse

$$\Omega = \{1,2,\dots,N-1\}$$

Et la fonction de probabilités en sera changée.

Probabilité nulle

Dans l'exemple du dé pipé (2 fois la face 1 et pas la face N), si toutes les faces restent équiprobables, nous aurons alors une fonction de probabilité différente

$$\mathbb{P}(\omega) = \begin{cases} \dfrac{2}{N} & \omega = 1 \\ \dfrac{1}{N} & \omega \in [2, N-1] \end{cases}$$

Probabilité nulle

Notez qu'il est possible de garder l'univers du dé non pipé $\Omega = \{1,2,\dots,N\}$.

Quelle sera alors la fonction de probabilités ?

Réponse :

$$\mathbb{P}(\omega) = \begin{cases} \dfrac{2}{N} & \omega = 1 \\ \dfrac{1}{N} & \omega \in [2, N-1] \\ 0 & \omega = N \end{cases}$$

Probabilité d'un sous-ensemble

Si $\omega \in \Omega$ décrit la probabilité d'une seule réalisation (ou évènement), il est également possible de mesurer la probabilité d'un sous-ensemble de réalisations.

Par exemple :

Pour un lancé de dé, quelle est la probabilité d'obtenir un nombre pair.

Ici, $\Omega = \{1,2,3,4,5,6\}$ et nous cherchons quelle est la probabilité de $A = \{2,4,6\} \subseteq \Omega$.

Probabilité d’un sous-ensemble

Proposition 1:

Pour tout sous-ensemble $A \subseteq \Omega$, nous calculons la probabilité $\mathbb{P}(A)$ comme la somme des probabilités de chaque évènement de A, donc

$$\mathbb{P}(A) = \sum_{\omega \in A} \mathbb{P}(\omega) \,.$$

Probabilité d'un sous-ensemble

Proposition 2 :

Pour deux sous-ensembles $A, B \subseteq \Omega$ disjoints, donc tels que $A \cap B = \emptyset$

$$\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B).$$

Exercice

Prouvez la proposition 2

Pour deux sous-ensembles $A, B \subseteq \Omega$ disjoints, donc tels que $A \cap B = \emptyset$, alors

$$\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B).$$

Corrigé

Soient deux sous-ensembles $A, B \subseteq \Omega$ disjoints ,donc tels que $A \cap B = \emptyset$

Selon la proposition 1, nous avons que

$\mathbb{P}(A) = \sum_{\omega \in A} \mathbb{P}(\omega)$, et $\mathbb{P}(B) = \sum_{\omega \in B} \mathbb{P}(\omega)$

Donc

$$\mathbb{P}(A) + \mathbb{P}(B) = \sum_{\omega \in A} \mathbb{P}(\omega) + \sum_{\omega \in B} \mathbb{P}(\omega)$$

$$\underset{1}{=} \sum_{\omega \in (A \cup B)} \mathbb{P}(\omega) \underset{2}{=} \mathbb{P}(A \cup B).$$

Corrigé - suite

L’égalité 1 s’explique par la théorie des ensembles : les deux sommes assurent que nous prenons une et une seule fois chaque élément de A et de B. Les deux sous-ensembles étant disjoints, nous prendrons alors exactement une fois chaque élément de $A \cup B$.

L’égalité 2 est, à nouveau, l’application de la proposition 1 pour $A \cup B$.

Propriétés de $\mathbb{P}: \Omega \rightarrow [0,1]$

- $0 \leq \ \mathbb{P}(\Omega) \leq 1$,
- $\mathbb{P}(\Omega) = 1$ si Ω contient toutes les réalisations possibles,
- $\mathbb{P}(\Omega) < 1$ implique que Ω n’est pas exhaustif,
- Pour tout $A \subseteq \Omega, \mathbb{P}(A) = \mathbb{P}(\Omega \backslash A) = 1 - \mathbb{P}(A)$,
- $\mathbb{P}(\emptyset) = 0$,
- $\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B)$.

Exercices

1. Nous lançons un dé équilibré 2 fois et comptons la somme des deux lancés. Ecrivez l'espace réalisable ainsi que la fonction de probabilité.

2. Le résultat du point 2 est-il le même que si on lance une fois 2 dés équilibrés simultanément ?

3. Nous lançons une pièce de monnaye avec un côté pile et un côté face 3 fois. Ecrivez l'espace réalisable ainsi que la fonction de probabilité sachant que pour 1 lancé, la pièce tombe sur pile avec probabilité 2/3 et que l'ordre importe !

NOTE: lors de répétitions, les probabilités se ***MULTIPLIENT*** !

Exercices – Corrigé 1

1. Avec 2 lancés, nous faisons au moins 1 à chaque lancé et au plus 6. Nous en déduisons que
$$\Omega = \{2,4, \dots, 12\}$$

Pour calculer la fonction de probabilité, nous utiliserons le tableau suivant

Valeur du 2e lancé

Valeur du 1er lancé	1	2	3	4	5	6
1	2	3	4	5	6	7
2	3	4	5	6	7	8
3	4	5	6	7	8	9
4	5	6	7	8	9	10
5	6	7	8	9	10	11
6	7	8	9	10	11	12

Exercices – Corrigé 1

Selon la table, nous voyons qu'il existe au total 36 possibilités mais, vu que nous ne nous intéressons qu'à la somme, certaines réalisations donnent le même résultat :

$\omega \in \Omega$	2	3	4	5	6	7	8	9	10	11	12
# d'occurrences	1	2	3	4	5	6	5	4	3	2	1
$\mathbb{P}(\omega)$	$\frac{1}{36}$	$\frac{2}{36}$	$\frac{3}{36}$	$\frac{4}{36}$	$\frac{5}{36}$	$\frac{6}{36}$	$\frac{5}{36}$	$\frac{4}{36}$	$\frac{3}{36}$	$\frac{2}{36}$	$\frac{1}{36}$

Il est facile à vérifier que $\mathbb{P}(\omega) \in [0,1], \forall \omega \in \Omega$.

Exercices – Corrigé 2

En refaisant la même table avec 2 lancés simultanés, nous constaterons que la distribution de probabilité est la même.

La raison est que chaque lancé de dé est indépendant des autres, donc la probabilité est la même pour chaque dé et chaque lancé.

On appelle ceci une situation sans mémoire, pour lequel la réalisation du premier lancé n’apporte aucune information sur le second.

Exercices – Corrigé 2

Les deux situations (lancés simultanés et lancés itératifs) ne sont plus les mêmes si on ajoute des règles liées à l'ordre :

- Le second lancé sera répété jusqu'à avoir un autre résultat que le premier

Dans ce cas, il n'est plus possible d'assimiler le lancé simultané à plusieurs dés à cette situation !

Exercices – Corrigé 3

Notons Pile = p et Face = f.
Nous avons au total 8 résultats possibles

$$\Omega = \{ppp, ppf, pfp, pff, fpp, fpf, ffp, fff\}$$

Notons que pour chaque lancé

$\mathbb{P}(p) = \frac{2}{3}$ et $\mathbb{P}(f) = \frac{1}{3}$.

Lors d'une répétition, les probabilités se multiplient.

Par conséquent

$$\mathbb{P}(ppp) = \mathbb{P}(p) \times \mathbb{P}(p) \times \mathbb{P}(p)$$
$$\mathbb{P}(pfp) = \mathbb{P}(p) \times \mathbb{P}(f) \times \mathbb{P}(p)$$

Exercices – Corrigé 3

Selon la table, nous voyons qu'il existe au total 8 possibilités, chacune étant unique :

$\omega \in \Omega$	ppp	ppf	pfp	pff	fpp	fpf	ffp	fff
$\mathbb{P}(\omega)$	$\frac{8}{27}$	$\frac{4}{27}$	$\frac{4}{27}$	$\frac{2}{27}$	$\frac{4}{27}$	$\frac{2}{27}$	$\frac{2}{27}$	$\frac{1}{27}$

Il est facile à vérifier que $\mathbb{P}(\omega) \in [0,1], \forall \omega \in \Omega$.

Pourquoi multiplier

Reprenons l'exemple du lancé de pièces précédent.

A chaque lancé, nous avons certes 2 réalisations possibles, mais celles-ci ont des probabilités de $\frac{1}{3}$ et $\frac{2}{3}$.

C'est donc comme si nous avions 3 réalisations équiprobables, p_1, p_2 ou f ayant chacune probabilité $\frac{1}{3}$.

Pourquoi multiplier

Ainsi, effectuons deux lancés et examinons les cas possibles avec l'univers équiprobable devient :

$\omega \in \Omega'$	p_1p_1	p_1p_2	p_1f	p_2p_1	p_2p_2	p_2f	fp_1	fp_2	ff
$\mathbb{P}(\omega)$	$\frac{1}{9}$	$\frac{1}{9}$	$\frac{1}{9}$	$\frac{1}{9}$	$\frac{1}{9}$	$\frac{1}{9}$	$\frac{1}{9}$	$\frac{1}{9}$	$\frac{1}{9}$

Or, nous savons que $p_1 = p_2$, il nous suffit alors de regrouper les cas équivalents et remarquer que nous aurons 4 cas

$\omega \in \Omega'$	pp	pp	pf	pp	pp	pf	fp	fp	ff
$\mathbb{P}(\omega)$	$\frac{1}{9}$	$\frac{1}{9}$	$\frac{1}{9}$	$\frac{1}{9}$	$\frac{1}{9}$	$\frac{1}{9}$	$\frac{1}{9}$	$\frac{1}{9}$	$\frac{1}{9}$

Pourquoi multiplier

Puis en regroupant nous obtenons

$$\mathbb{P}(pp) = \frac{4}{9} = \frac{2}{3} \times \frac{2}{3} = \mathbb{P}(p) \times \mathbb{P}(p),$$

$$\mathbb{P}(pf) = \frac{2}{9} = \frac{2}{3} \times \frac{1}{3} = \mathbb{P}(p) \times \mathbb{P}(f),$$

$$\mathbb{P}(fp) = \frac{2}{9} = \frac{1}{3} \times \frac{2}{3} = \mathbb{P}(f) \times \mathbb{P}(p),$$

$$\mathbb{P}(ff) = \frac{1}{9} = \frac{1}{3} \times \frac{1}{3} = \mathbb{P}(f) \times \mathbb{P}(f).$$

Pourquoi multiplier

Puis en regroupant nous obtenons

$$\mathbb{P}(pp) = \frac{4}{9} = \frac{2}{3} \times \frac{2}{3} = \mathbb{P}(p) \times \mathbb{P}(p),$$

$$\mathbb{P}(pf) = \frac{2}{9} = \frac{2}{3} \times \frac{1}{3} = \mathbb{P}(p) \times \mathbb{P}(f),$$

$$\mathbb{P}(fp) = \frac{2}{9} = \frac{1}{3} \times \frac{2}{3} = \mathbb{P}(f) \times \mathbb{P}(p),$$

$$\mathbb{P}(ff) = \frac{1}{9} = \frac{1}{3} \times \frac{1}{3} = \mathbb{P}(f) \times \mathbb{P}(f).$$

Fonction de répartition - motivation

Nous avons vu que la fonction de probabilité associe une probabilité à chaque réalisation :

$$\mathbb{P}: \Omega \to [0,1]$$

Or, dans le cas d'un univers continu (non-dénombrable), nous avons vu également que la probabilité d'un élément précis de chaque évènement est toujours nulle :

$$\mathbb{P}(\omega) = 0$$

Si Ω n'est pas dénombrable. Comment faire alors ?

Fonction de répartition – le principe

Le cas d'un univers continu non-dénombrable montre que regarder la probabilité d'une réalisation une à une ne fonctionne pas toujours.

L'idée est alors de regarder la probabilité ***<u>cumulative</u>*** des évènements.

Prérequis : il faut que Ω ait un opérateur $\leq: \Omega \rightarrow \{0,1\}$ permettant de trier l'univers.

Mesure : c'est le cumul des probabilités qui est mesuré !

Fonction de répartition – définition

$$F_{\mathbb{P}}(\omega) = \mathbb{P}(\{x \mid x \leq \omega)$$

La fonction de répartition mesure donc, pour tout $\omega \in \Omega$, la probabilité du sous-ensemble $X = \{x \mid x \leq \omega)$.

Fonction de répartition – propriétés

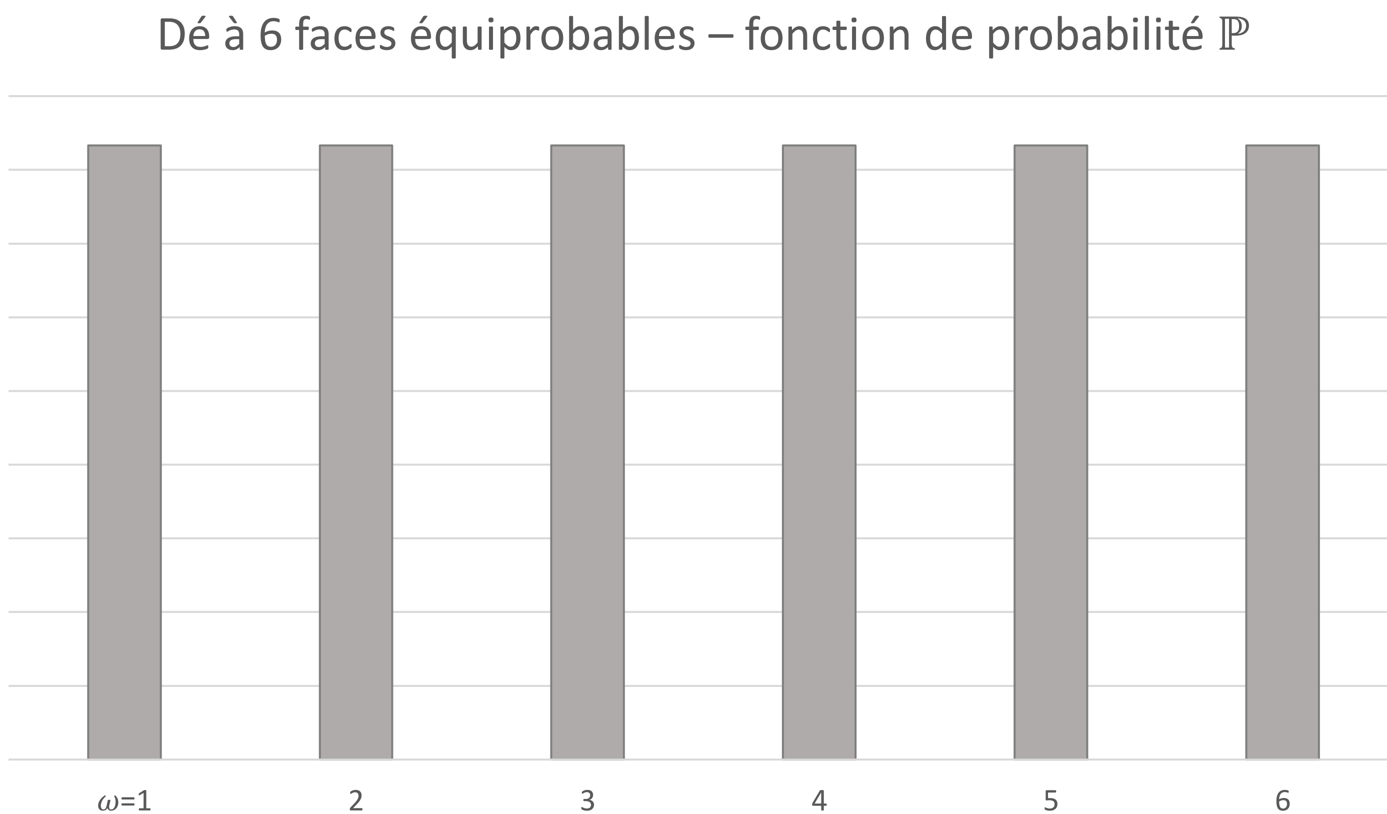
- $F_{\mathbb{P}}(\omega) \in [0,1], \forall \omega \in \Omega$,
- $F_{\mathbb{P}}(\omega)$ est une fonction croissante,
- $F_{\mathbb{P}}(\omega) = \mathbb{P}(]-\infty, \omega]$,
- Dans le cas d'un univers discret (ou dénombrable)

$$F_{\mathbb{P}}(\omega) = \sum_{x \leq \omega} \mathbb{P}(x) \, .$$

Exercices

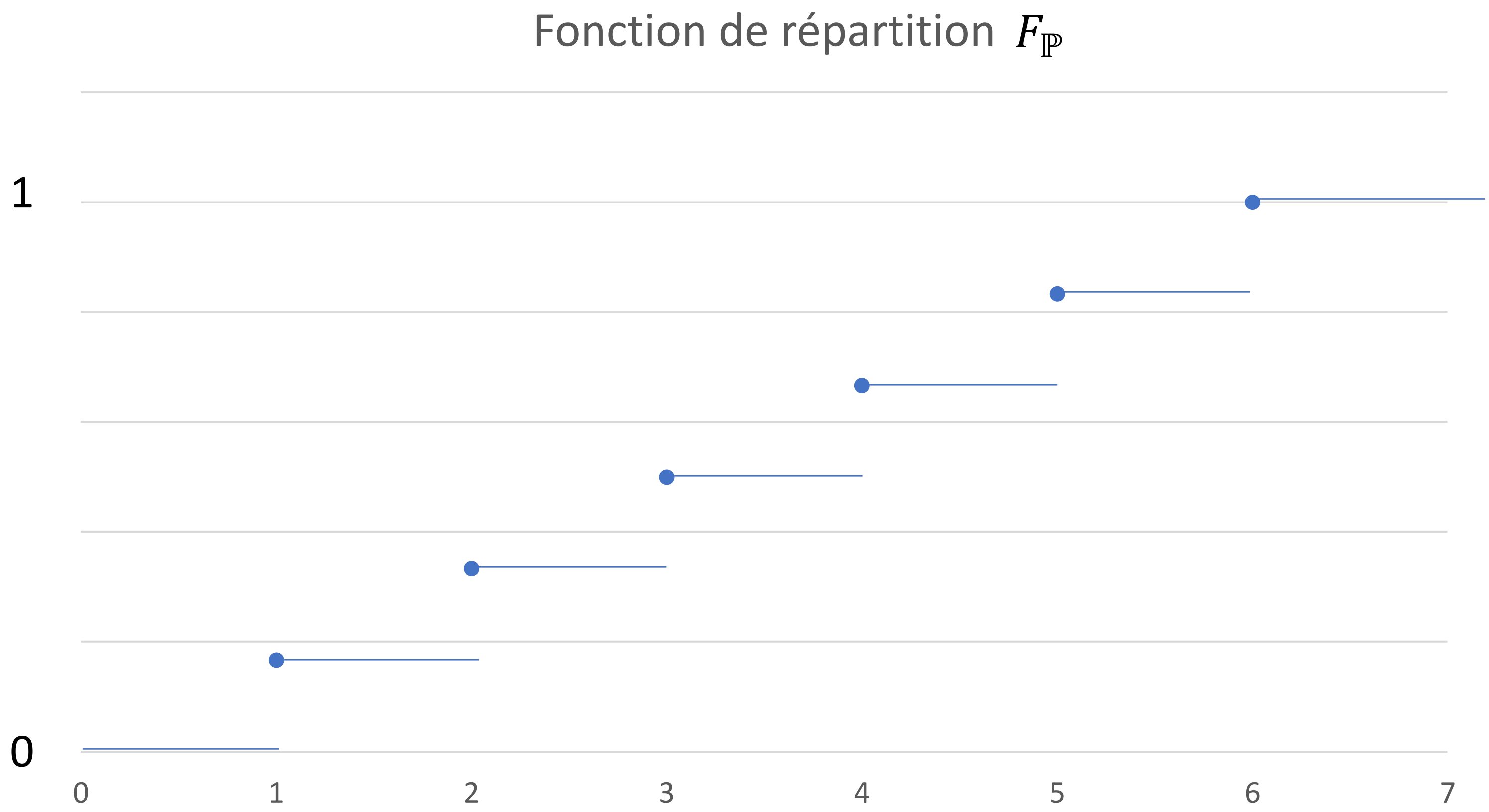
Définissez et dessinez la fonction de répartition pour le résultat d’un lancé de dé à 6 faces équiprobables.

Exercices



Dé à 6 faces équiprobables – fonction de probabilité ℙ
ω=1
2
3
4
5
6

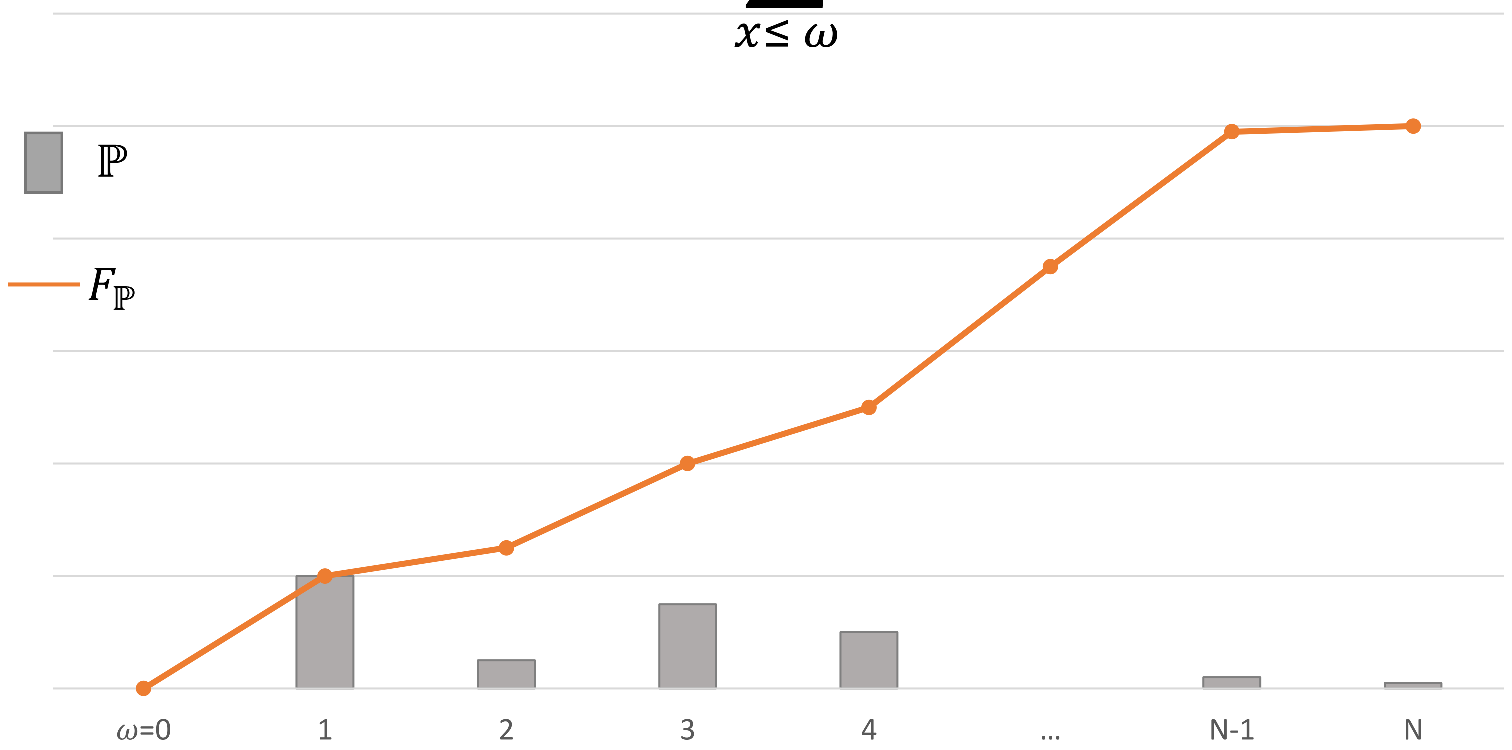
Exercices



Interprétation géométrique

Soit $\Omega = \{1, \ldots, \mathrm{N}\}$ un ensemble discret fini (numérotons les – alors

$$F_{\mathbb{P}}(\omega) = \sum_{x \leq \omega} \mathbb{P}(x).$$



Interprétations géométrique

Supposons que la largeur de chaque rectangle de la fonction $\mathbb{P}$ soit égale à 1.

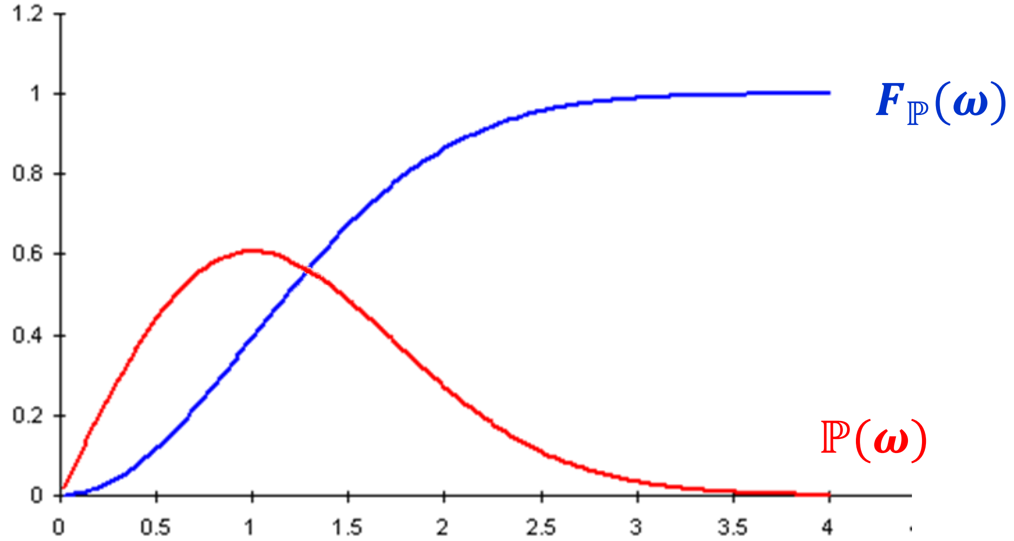
Alors la probabilité d'un évènement $\omega \in \Omega$ est égal à l'aire du rectangle correspondant.

Donc, la probabilité cumulée correspond à la sommes des aires.

Interprétations géométrique – cas continu

Supposons maintenant que $|\Omega| = \infty$ (que Ω soit dénombrable ou non) !

Dans ce cas, $\mathbb{P}$ est une fonction continue et $F_{\mathbb{P}}$ correspond à l'aire cumulée sous la courbe de $\mathbb{P}$.



Division en intervalles

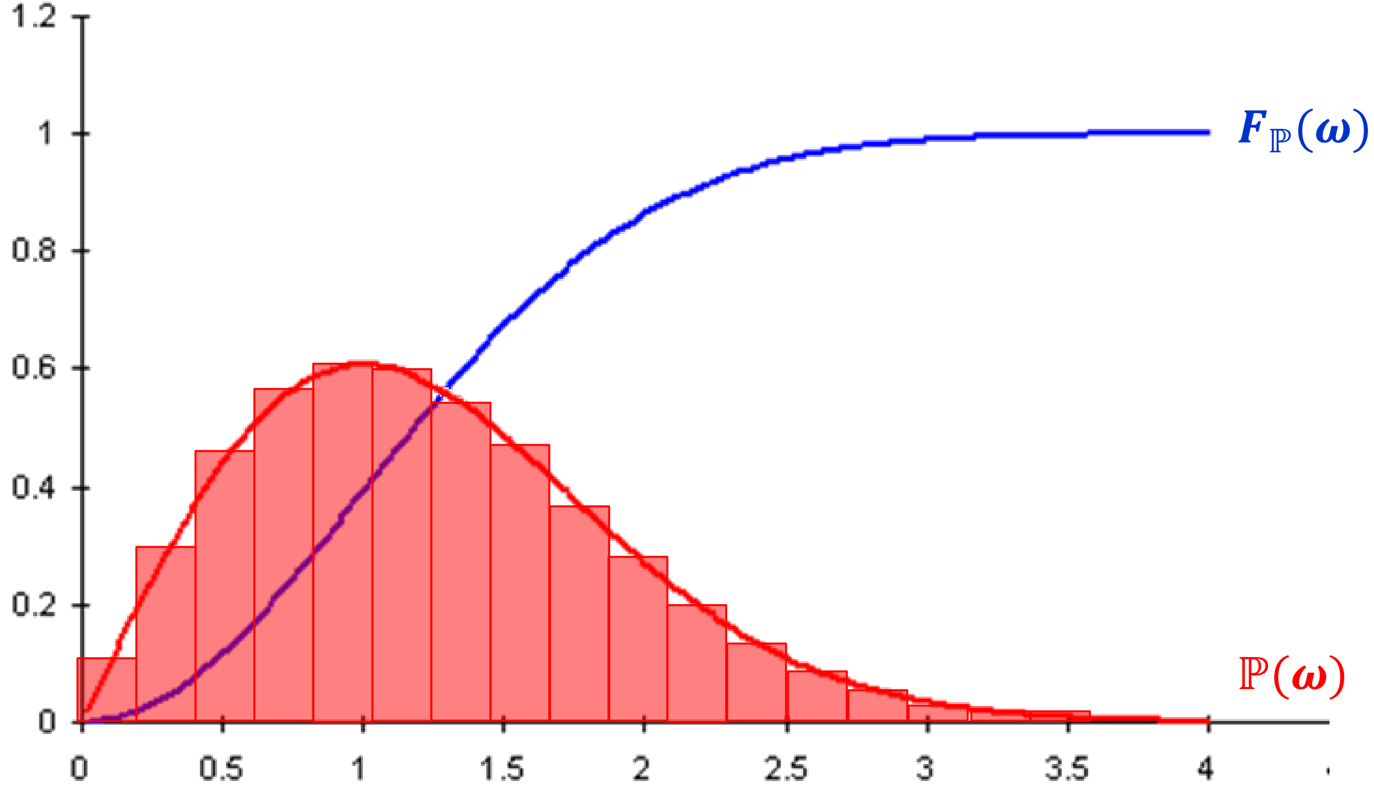
Dans le cas continu,
Pour «simuler» le cas continu, nous pouvons diviser Ω en intervalles de largeur $\frac{1}{N}$.

De plus, pour $\omega \in \Omega$, notons I_{ω} tous les sous-ensembles de largeur $\frac{1}{N}$ strictement plus petits que ω.

Ainsi

$$F_{\mathbb{P}}(\omega) \cong \sum_{I \in I_{\omega}} \mathbb{P}(I)\,.$$

Division en intervalles



Division en intervalles

Plus les intervalles seront petits, plus l'approximation sera exacte.

Le cas continu est en fait le passage à la limite de cette décomposition de sous-intervalles de largeur infinitésimale.

A quoi cela correspond-il ?

Division en intervalles

C’est la définition de l’intégrale :

$$F_{\mathbb{P}}(\omega) = \int_{x=-\infty}^{\omega} \mathbb{P}(x)\,.$$

Probabilités conditionnelles

Lorsque nous mesurons une probabilité, y a-t-il des informations qui influencent la probabilité ?

<u>**Exemple :**</u>
Quelle est la probabilité d'avoir un cancer ?

Quelle est la probabilité d'avoir un cancer des poumons ***sachant que*** la personne fume 30 cigarettes par jour ?

Notation

Soit une Ω avec une fonction de probabilité $\mathbb{P}$. Notons une condition quelconque A.

Alors nous noterons la probabilité de l'évènement $\omega \in \Omega$

$$\mathbb{P}(\omega|A) = \text{probabilité de } \omega \text{ sachant } A.$$

Exercices

Considérons un lancé d'un dé à 6 faces équilibrées.

$\Omega = \{1,2,3,4,5,6\}$ avec $\mathbb{P}(\omega) = \frac{1}{6}$.

1. Quelle est la probabilité de tirer 6 sachant qu'on a tiré un nombre pair ?

2. Quelle est la probabilité de tirer ω sachant que nous n'avons pas tiré 1

3. Quelle est la probabilité de retirer 5 sachant que le précédent lancé était déjà un 5 ?

Corrigé

1. Ici, $\mathbb{P}(\omega = 6 | \omega \in \{2,4,6\}) = \frac{1}{3}$. L'information sur ω réduit les issues possibles et donc la probabilité !

2. Nous n'avons plus que 5 issues possibles, donc

ω	1	2	3	4	5	6
$\mathbb{P}(\omega \mid \omega \neq 1)$	0	$\frac{1}{5}$	$\frac{1}{5}$	$\frac{1}{5}$	$\frac{1}{5}$	$\frac{1}{5}$

Corrigé - suite

3. Chaque lancé de dé est indépendant du précédent. Pour nous en convaincre, listons toutes les possibilités de 2 lancés avec leur probabilités

Lancé 1 \ Lancé 2	$\omega_2 = 1$	2	3	4	5	6
$\omega_1 = 1$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$
2	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$
3	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$
4	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$
5	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$
6	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$	$\frac{1}{36}$

Corrigé - suite

La condition «sachant que le premier lancé était un 5» correspond à sélectionner la ligne $\omega = 5$ dans le tableau ci-dessus.

La probabilité est calculée par

$$\mathbb{P}(\omega_2 = 5|\omega_1 = 5) = \frac{\mathbb{P}(5\&5)}{\mathbb{P}(x \in \{x|\omega_1 = 5\})} =$$

$$\frac{\frac{1}{36}}{\mathbb{P}(5\&1) + \mathbb{P}(5\&2) + \cdots \mathbb{P}(5\&6)} = \frac{\frac{1}{36}}{\frac{6}{36}} = \frac{1}{6}.$$

Calcul de la probabilité conditionnelle

La probabilité conditionnelle d'un évènement est calculée comme suit :

$$\mathbb{P}(\omega|A) = \begin{cases} \dfrac{\mathbb{P}(\omega)}{\mathbb{P}\left(A\right)} si\ \omega \in A \\ 0\ si\ \omega \notin A \end{cases}$$

Propriété

- $\mathbb{P}(A|A) = 1$
- $\mathbb{P}(B|A) = \mathbb{P}(B \cap A)/\mathbb{P}(A)$ (Formule de Bayes)

Evènements indépendants

L'évènement ω et la condition A sont dits ***indépendants*** si

$$\mathbb{P}(\omega|A) = \mathbb{P}(\omega).$$

ATTENTION AUX INVERSIONS !!

Soient $A, B \in \Omega$ deux sous-ensembles de réalisations, alors dans le cas général

$$\mathbb{P}(A|B) \neq \mathbb{P}(B|A).$$

Attention aux abus !!!

Les probabilités conditionnelles peuvent être manipulées pour donner une fausse signification à certaines études.

<u>**Exemple:**</u>
Dans une prison, 7 personnes sur 10 ont des yeux verts.

<u>FAUX :</u> Une personne aux yeux verts a 70% de chances d'être incarcérée.

<u>JUSTE :</u> sachant qu'une personne est incarcérée, il y a 70% de chances qu'elle ait les yeux verts.

Note : cet exemple est purement hypothétique et ne saurait être interprété comme du racisme – j'ai les yeux verts !

Quelle différence

Il y a 70% de chances qu'une personne aux yeux verts soit incarcérée
est une formulation quelque peu floue – la condition implicite est «*sachant qu'une personne a les yeux verts, il y a 70% de chances qu'elle soit incarcérée*», ce qui correspond à $\mathbb{P}(incarc\text{é}r\text{é}|yeux\ verts) = 0.7$.

Or, dans notre cas, nous avons bien observé que $\mathbb{P}(yeux\ verts|incarc\text{é}r\text{é}) = 0.7$.

Pourquoi cette différence ?

En fait, nous avons deux sous-ensembles de la population totale ici. Soit Ω la population totale, et soient

$A = \{incarcérés\} \subseteq \Omega$ et
$B = \{yeux\ verts\} \subseteq \Omega$.

Donc $\mathbb{P}(B|A) = 0.7 = \frac{\mathbb{P}(B \cap A)}{\mathbb{P}(A)}$.

Exercice

Supposons que la population totale Ω est composée de 1Mio de personnes, dont 20% ont les yeux verts, 1000 sont emprisonnés et 700 des prisonniers ont les yeux verts, et soient :

$A = \{incarcérés\} \subseteq \Omega$ et $B = \{yeux\ verts\} \subseteq \Omega$.

Utilisez l'écriture probabiliste pour écrire et calculer:

1. La probabilité qu'une personne soit à la fois incarcérée et ait les yeux verts,

2. La probabilité qu'une personne ait les yeux verts sachant qu'elle est incarcérée,

3. La probabilité qu'une personne soit incarcérée sachant qu'elle a les yeux verts.

Exercice

$\Omega = \{1Mio\ de\ personnes\}$,

$A = \{incarc\text{é}r\text{é}s\} \subseteq \Omega$ et $\mathbb{P}(A) = \frac{1000}{1Mio} = 0.001 = 0.1\%$

$B = \{yeux\ verts\} \subseteq \Omega$ et $\mathbb{P}(B) = 0.2 = 20\%$

1. Cela correspond à $\mathbb{P}(A \cap B) = \frac{700}{1Mio} = 0.0007 = 0.07\%$.
2. Il s'agit de $\mathbb{P}(B|A) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(A)} = \frac{0.0007}{0.001} = 0.7 = 70\%$.
3. Il s'agit de $\mathbb{P}(A|B) = \frac{\mathbb{P}(A \cap B)}{\mathbb{P}(B)} = \frac{0.0007}{0.2} = 0.0035 = 0.35\%$.

Que conclure de cet exercice ?

Si 70% des prisonniers ont les yeux verts, cela ne signifie pas pour autant qu'une personne aux yeux verts ait 70% de chances d'être un criminel, ni même qu'un criminel a plus de chances d'avoir les yeux verts !

Pourquoi est-ce un abus ?

Si $B = \{incarcéré\}$ et $C = \{criminels\}$, alors (dans le cas idéal sans erreur judiciaire)

$$B = \{incarcéré\} \subseteq C = \{criminels\}$$

Mais est-ce que $B = C$? Est-ce que tous les criminels sont incarcérés ?

Très certainement pas. Nous ne mesurons aucunement la relation entre ces deux sous-ensembles !

<u>Moralité :</u>
Il est plus probable qu'une personne aux yeux verts soit condamnée…

Variable aléatoire

Une variable aléatoire X est une fonction associant une valeur numérique à chaque élément de l'ensemble réalisable :

$$X: \Omega \rightarrow \mathbb{R}$$
$$\omega \rightarrow X(\omega)$$

Il se peut que l'espace d'arrivée soit un ensemble fini, dénombrable ou non.

Variable aléatoire : exemple

Considérons le cas de 2 lancés de dés. Nous avons donc $\Omega = \{1,2,3,4,5,6\}^2$, soit l'ensemble des tous les couples $[\omega_1, \omega_2]$ composés des chiffres 1 à 6.

Alors nous pouvons définit $X: \Omega \to \mathbb{R}$ tel que

$$X([\omega_1, \omega_2]) = \omega_1 + \omega_2.$$

Un autre exemple est

$$X([\omega_1, \omega_2]) = \omega_2.$$

Ces deux exemples sont des variables aléatoires.

Variable aléatoire : propriétés

Soit $X: \Omega \to \mathbb{R}$ une variable aléatoire

1. Pour deux $\omega_1, \omega_2 \in \Omega$,si $X(\omega_1) \neq X(\omega_2)$, alors $\omega_1 \neq \omega_2$,

2. L'inverse n'est pas vrai : pour une valeur donnée de X, il peut exister plusieurs évènements tels que X,

3. Les valeurs atteignables de X, $Im_X(\Omega)$, forme une partition (couvrante ou complète) de Ω.

Cela s'écrit aussi

$$\cup_i (X = i) = \Omega.$$

Exemples – 2 lancés de dés

Reprenons la variable aléatoire $X([\omega_1, \omega_2]) = \omega_1 + \omega_2$, soit la somme des deux lancés de dés

Il y a deux évènements possibles donnant $X = 3$: $[1,2]$ ou $[2,1]$. Si $X = 4$ il y a 3 possibilités ($[1,3]$, $[2,2]$ ou $[3,1]$). Clairement, les deux sous-ensembles sont disjoints.

Cela illustre les propriétés 1 et 2.

$Im_X(\Omega) = \{2,3, \dots, 12\}$ car la somme de 2 lancés de dés est une valeur de 2 à 12.

Exercice: donnez la décomposition de Ω induite par

$$X([\omega_1, \omega_2]) = \omega_2.$$

Corrigé

$X([\omega_1, \omega_2]) = \omega_2$ a 6 valeurs possibles, et le tableau ci-dessous donne la partition :

X	1	2	3	4	5	6
$[\omega_1, \omega_2]$ tels que $X([\omega_1, \omega_2]) = X$	$[1,1]$ $[2,1]$ $[3,1]$ $[4,1]$ $[5,1]$ $[6,1]$	$[1,2]$ $[2,2]$ $[3,2]$ $[4,2]$ $[5,2]$ $[6,2]$	$[1,3]$ $[2,3]$ $[3,3]$ $[4,3]$ $[5,3]$ $[6,3]$	$[1,4]$ $[2,4]$ $[3,4]$ $[4,4]$ $[5,4]$ $[6,4]$	$[1,5]$ $[2,5]$ $[3,5]$ $[4,5]$ $[5,5]$ $[6,5]$	$[1,6]$ $[2,6]$ $[3,6]$ $[4,6]$ $[5,6]$ $[6,6]$

Observons que toutes les paires figurent une et une seule fois dans le tableau ci-dessous. C'est ce que nous appelons une partition couvrante de Ω.

A ne pas confondre

Il est important de ne pas confondre la notion de *variable* aléatoire et de réalisation.

La variable aléatoire est une fonction retournant une valeur pour toute réalisation de $\omega \in \Omega$ de l'espace réalisable.

Une observation, évènement ou réalisation est une valeur (ou un sous-ensemble) de $A \subseteq \Omega$.

Discret vs continu

Il existe deux grands familles de variables aléatoires :

Les ***variables aléatoires discrètes***, pour lesquelles l'ensemble des valeurs possibles est fini *ou* dénombrable,

- Les ***variables aléatoires continues***, pour lesquelles l'ensemble des valeurs possibles est infini et non-dénombrable.

Variable aléatoire et probabilités

Pour un même ensemble réalisable Ω, la fonction de probabilités $\mathbb{P}$ n'est pas explicitement liée à une variables aléatoire X sur Ω. En fait, il existe de nombreuses variables aléatoires pour un même couple $(\Omega, \mathbb{P})$.

En revanche, la fonction de probabilités $\mathbb{P}$ nous permet de mesurer les probabilités pour les valeurs de X:

La notation $X = i$ correspond en fait à un ensemble de réalisations qui prennent la valeur

$$(X = i) = \{\omega \in \Omega \mid X(\omega) = i\} \subseteq \Omega.$$

Variable aléatoire et probabilités

Ainsi, nous pouvons observer que

$$\mathbb{P}(X = i) = \sum_{\omega | X(\omega) = i} \mathbb{P}(\omega)\,.$$

Autrement dit, la probabilité d'observer une valeur *donnée* i de la variable aléatoire X est égale à la somme des probabilités de tous les évènements ω prenant la valeur $X(\omega) = i$.

Exemple

Soit l'exemple du tiré d'une carte dans un jeu de cartes à 52 cartes (sans joker). Chaque carte a une probabilité de $\frac{1}{52}$ d'être tirée.

Soit la variable aléatoire $X: \Omega \rightarrow \{2,3, \dots, 10, 11\}$ correspondant à la valeur de la carte (nous dirons qu'un As vaut 11 et que les cartes à figures valent 10).

Quelle est la probabilité d'observer 7 en tirant une seule carte ?

Corrigé

Formellement, nous cherchons

$$\mathbb{P}(X = 7) = \sum_{\omega | X(\omega)=7} \mathbb{P}(\omega)\,,$$

Or nous savons qu'il n'y a que 4 possibilités (7 de cœur, 7 de carreau, 7 de trèfle et 7 de pic), donc

$$\mathbb{P}(X = 7) = \mathbb{P}(7\heartsuit) + \mathbb{P}(7\Diamond) + \mathbb{P}(7\clubsuit) + \mathbb{P}(7\spadesuit) = \frac{4}{52}.$$

Exercice

Dans l’exemple du jeu à 52 cartes, définissez la variable aléatoire pour 2 tirages de cartes dont on mesure la somme.

Quelle est la probabilité d’observer 18 ?

Corrigé

La variables aléatoire est

$$Y : \Omega \times \Omega \to \{4, \dots, 22\}$$

$$Y(\omega_1, \omega_2) = X(\omega_1) + X(\omega_2),$$

Où X est la variable aléatoire définie auparavant pour le tirage d'une seule carte.

Notons également que

$$\mathbb{P}(X = i) = \begin{cases} \frac{4}{52}, i \in \{2,3, \dots, 9,11\} \\ \frac{16}{52}, i = 10 \end{cases}$$

Et

$$\mathbb{P}(\omega_1, \omega_2) = \mathbb{P}(\omega_1) \times \mathbb{P}(\omega_2).$$

Corrigé - suite

Voici la liste des combinaisons donnant le résultat 18 ainsi que leurs probabilités
Rappelons que

(ω_1, ω_2)	$\mathbf{(11, 7)}$	$\mathbf{(10, 8)}$	$\mathbf{(9, 9)}$	$\mathbf{(8, 10)}$	$\mathbf{(7, 11)}$
$\mathbb{P}(\omega_1, \omega_2)$	$\mathbb{P}(11) \times \mathbb{P}(7)$	$\mathbb{P}(7) \times \mathbb{P}(8)$	$\mathbb{P}(9) \times \mathbb{P}(9)$	$\mathbb{P}(8) \times \mathbb{P}(10)$	$\mathbb{P}(7) \times \mathbb{P}(11)$
	$\frac{4}{52} \times \frac{4}{52} = \frac{16}{2704}$	$\frac{16}{52} \times \frac{4}{52} = \frac{64}{2704}$	$\frac{4}{52} \times \frac{4}{52} = \frac{16}{2704}$	$\frac{4}{52} \times \frac{16}{52} = \frac{64}{2704}$	$\frac{4}{52} \times \frac{4}{52} = \frac{16}{2704}$

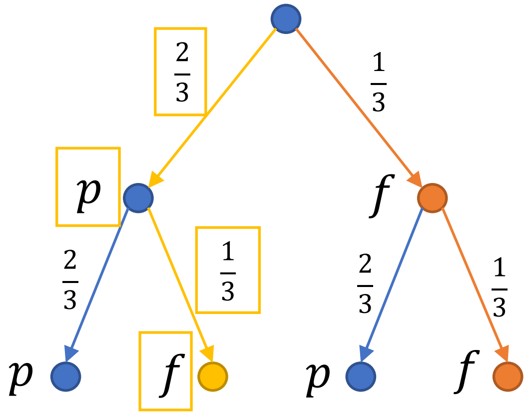
Corrigé

La probabilité de tirer 18 est donc égal à la somme des observations ci-dessus, soit

$$\mathbb{P}(Y = i) = \frac{176}{2704} = \frac{11}{169} \cong 0.065 = 6.5\%.$$

Pour plusieurs répétitions

Ici, nous tirons en fait plusieurs fois la même variable aléatoire. Cela peut se représenter sous une forme d'arbre dont les feuilles (nœuds sans parents) sont les réalisations et la probabilité de chaque réalisation est le produit des probabilités du chemin parcouru.



$$\boxed{\mathbb{P}(pf) = \frac{2}{3} \times \frac{1}{3} = \frac{2}{9}}$$

Probabilités et Combinatoire

Les probabilités discrètes sont très étroitement liées à la combinatoire, qui étudie les combinaisons possibles.

Nous n’aborderons pas la combinatoire en détail. Nous nous contenterons de quelques notions de base utiles dans les probabilités discrètes classiques.

Exemple : tirage du loto Suisse

Pour gagner au loto Suisse, il faut trouver la «combinaison gagnante», à savoir 6 bons numéros entre 1 et 42, puis un numéro «chance» compris entre 1 et 6.

Quelle est la probabilité de gagner ?

1. Il y a une chance sur 6 de trouver le bon numéro chance,
2. Calculons la probabilité de trouver les 6 bons numéros principaux – le résultat final sera le produit des deux probabilités

Exemple : tirage du loto Suisse

Supposons que nous cochons «1,2,3,4,5,6».
$\mathbb{P}$ (1^{er} numéro = 1) $= 1/42$,
$\mathbb{P}$ (2^{e} numéro = 2) $= 1/41$ (il n'y a plus que 41 boules),
$\mathbb{P}$ (3^{e} numéro = 3) $= 1/40$ (il n'y a plus que 40 boules),
$\mathbb{P}$ (4^{e} numéro = 4) $= 1/39$ (il n'y a plus que 39 boules),
$\mathbb{P}$ (5^{e} numéro = 5) $= 1/38$ (il n'y a plus que 38 boules),
$\mathbb{P}$ (6^{e} numéro = 6) $= 1/37$ (il n'y a plus que 37 boules).

Comme il s'agit de tirages indépendants, les probabilités se multiplient, donc

$$\mathbb{P}(\{1,2,3,4,5,6\}) = \frac{(42-6)!}{42!} = \frac{1}{42 \times 41 \times 40 \times 39 \times 38 \times 37} = \frac{1}{3'776'965'920}$$

Exemple : tirage du loto Suisse

Il y a donc 1 chance sur presque 4Mia de tirer dans l'ordre 1,2,3,4,5,6. Mais si les 6 mêmes numéros sont tirés dans un autre ordre, la combinaison reste gagnant.

Chacune des permutations de 6 nombres de la combinaison gagnante est, elle aussi, gagnante et a la même probabilité de sortir que «1,2,3,4,5,6».

Le nombre de permutations N de éléments est $N!$.

Donc, la probabilité de tirer les 6 bons numéros parmi 42 est $6! = 6 \times 5 \times \cdots \times 1 = 720$ fois la probabilité d'une combinaison dans l'ordre.

Exemple : tirage du loto Suisse

Il y a donc 1 chance sur presque 4Mia de tirer dans l'ordre 1,2,3,4,5,6. Mais si les 6 mêmes numéros sont tirés dans un autre ordre, la combinaison reste gagnant.

Chacune des permutations de 6 nombres de la combinaison gagnante est, elle aussi, gagnante et a la même probabilité de sortir que «1,2,3,4,5,6».

Le nombre de permutations N de éléments est $N!$.

Donc, il y a en fait $6! = 6 \times 5 \times \cdots \times 1 = 720$ «combinaison» différentes qui seront gagnantes (vu que l'ordre n'importe pas) !

Exemple : tirage du loto Suisse

La probabilité de tirer les 6 bons numéros est donc de

$$\mathbb{P}(tirer\ 6\ parmi\ 42) = 6!\,\frac{(42-6)!}{42!} = \frac{1}{5'245'786}$$

Et avec le numéro chance, cela donne

$$\mathbb{P}(gagner) = \frac{1}{5'245'786} \times \frac{1}{6} = \frac{1}{31'474'716}$$

Lien avec la combinatoire

L'exemple du loto donne la formule du tirage de p éléments parmi n.

Si l'ordre importe, il y a

$$A_p^N = \frac{n!}{(n-p)!} = n \times (n-1) \times \cdots \times (n-p+1)$$

Le nombre de permutations de p éléments est $\mathrm{p}!$

Le nombre de combinaisons non-ordonné de p éléments parmi n est alors

$$C_p^N = \frac{A_p^N}{p!} = \frac{n!}{p!\,(n-p)!} = \frac{n \times (n-1) \times \cdots \times (n-p+1)}{p \times (p-1) \times \cdots \times 2 \times 1}$$

Exercice :

Calculez les probabilités de gains des deux derniers rangs (3 numéros gagnants, et 3 numéros gagnants plus le numéro chance).

Exercice :

Nous avons coché 6 numéros, pas 3 ! Le nombre de triplets (sans répétition), sachant que nous choisissons 6 nombres parmi 42 est donné par la formule

$$C_{6-3}^{42-6} = \frac{36 \times 35 \times 34}{3 \times 2 \times 1} = 7'140.$$

De plus, parmi nos 6 numéros cochés, il y a $C_3^6 = 20$ triplets différents.

Il y a donc $C_{6-3}^{42-6} \times C_3^6 = 142'800$ combinaisons de 3 nombres «couvertes» par nos 6 nombres joués.

Exercice :

La probabilité de gagner le dernier rang est donc de

$$\mathbb{P}(3\,num\acute{e}ros) = \frac{142'800}{5'245'786} \cong \frac{1}{36.7}$$

La probabilité de gagner 3 numéros plus le complémentaire est donné par

$$\mathbb{P}(3\,num\acute{e}ros + compl) = \mathbb{P}(3\,num) \times \mathbb{P}(compl.)$$
$$\cong \frac{1}{36.7} \times \frac{1}{6} \cong \frac{1}{220}$$

Formule générale

La probabilité d'avoir n bons numéros parmi les 6 numéros du loto sont donc

$$\mathbb{P}(n\ bons\ numéros\ sur\ 6) = \frac{C_{6-n}^{42-6} \times C_n^6}{C_6^{42}}$$

Cela donne la table suivante

n	$\mathbb{P}(n)$	**1 Chance sur**
0	0.3713060	2.7
1	0.4311941	2.3
2	0.1684352	5.9
3	0.0272219	36.7
4	0.0018014	555
5	0.0000412	242'86
6	0.0000002	5'245'786

Autres calculs de probabilité :

1. Quelle est la probabilité d'avoir au moins $k \in [2, 6]$ nombres consécutifs parmi les 6 boules tirées ?

2. Quelle est la probabilité de gagner au loto au moins une fois en jouant toute votre vie ?

Corrigé 1 :

Notez qu'il y a $42 - k + 1$ possibilités d'avoir $k \in [1, 6]$ nombres consécutifs entre 1 et 42.

De plus, pour chaque k boules tirées, il existe C_{6-k}^{42-6+k} combinaisons différentes contenant ces k boules.
Ainsi,

$$\mathbb{P}(\geq k\ consécutifs) = \frac{C_1^{42-k+1} \times C_{6-k}^{42-6+k}}{C_6^{42}} = \frac{42-k+1 \times C_{6-k}^{42-6+k}}{5'245'786}$$

Par exemple,

$$\mathbb{P}(\geq 6\ consécutifs) = \frac{C_1^{37} \times C_0^{42}}{C_6^{42}} = \frac{37 \times 1}{5'245'786} = \frac{1}{141'778}$$

Corrigé 2 :

Pour effectuer ce calcul, supposons que vous jouez 1 grille à chaque tirage, toute votre vie dès 18 ans (âge légal) jusqu'à vos 85 ans. Cela fait environ 7000 tirages.

La probabilité de PERDRE avec 1 grille à 1 tirage

$$\mathbb{P}(perdre) = 1 - \mathbb{P}(gagner) = \frac{31'474'733}{31'474'734} = 0.99999996822$$

La probabilité de perdre 7000 fois de suite est donc égale à

$$\mathbb{P}(perdre\ 7000fois) = \mathbb{P}(perdre)^{7000}) = 0.99977762$$

Donc, la probabilité de gagner au moins une fois dans sa vie est de

$$\mathbb{P}(gagner \geq 1\ fois) = 1 - \mathbb{P}(jamais\ gagner) = 0.0002238 \approx \frac{1}{4'496}$$

Donc, en jouant toute votre vie à chaque tirage, vous avec ~1 chance sur 5'000 de gagner (pour un coût total de 17'500 Chf au prix actuel d'une grille) !

Moyenne, Médiane, Mode

Moyenne (Espérance)

$$\overline{X} = E(X) = \sum_{i \in \Omega} p_i x_i \,,$$

Médiane

$$\mathrm{m}\ tel\ que\ \mathbb{P}(x \leq m) \geq_{i^*} 0.5 \text{ et } \mathbb{P}(x \geq\ m) \geq\ 0.5$$

Si la distribution de probabilité est SYMETRIQUE alors

$$m = E(X)$$

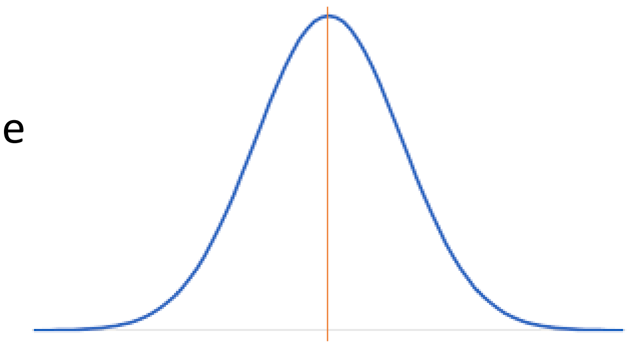
Mode

$$mod(X) = x_{i^*} \text{ où } i^* = \max_i\{p_i\}$$

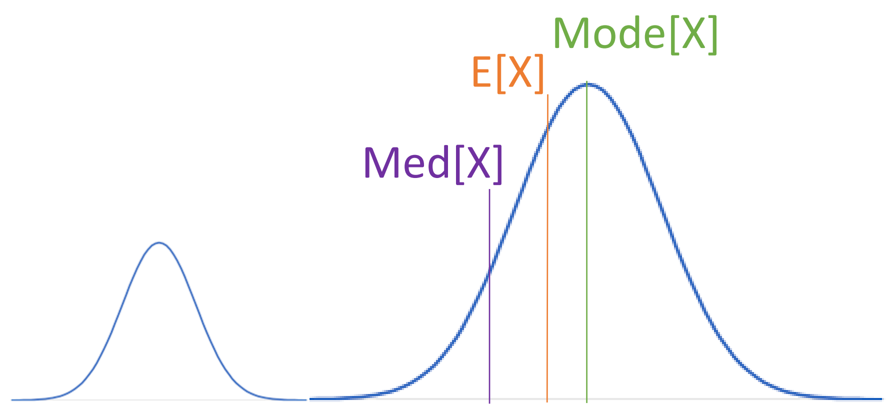
(élément avec la plus haute probabilité)

Moyenne, Médiane, Mode

Distribution symétrique
Moyenne = Médiane = Mode



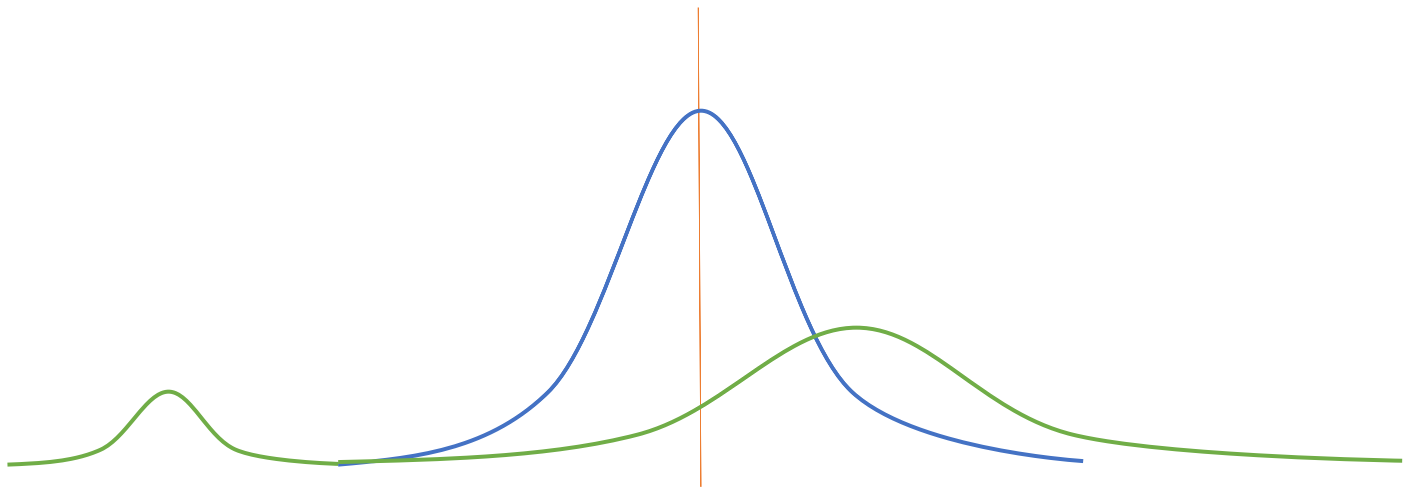
Distribution asymétrique
Tout change…



Limitation de la moyenne

La moyenne réduit toutes les données en une seule.

Deux distributions peuvent avoir la même moyenne, mais la situation peut être totalement différente :



Variance

Mesure l’écart moyen entre les données et la moyenne – mis au carré :

- Traite les élément en-dessous et en-dessus de la moyenne de la même manière
- Les données très éloignées de la moyenne engendrent une bien plus grande variance que beaucoup de données peu éloignées (à cause du carré !).

Variance

$$Var(X) = E(X - E[X])^2 = \sum_{i \in \Omega} (x_i - E[X])^2 p_i$$

Note : si X a une unité $[u]$ (p.ex. $, m., kg, …) alors Var(X) a l'unité $[u^2]$

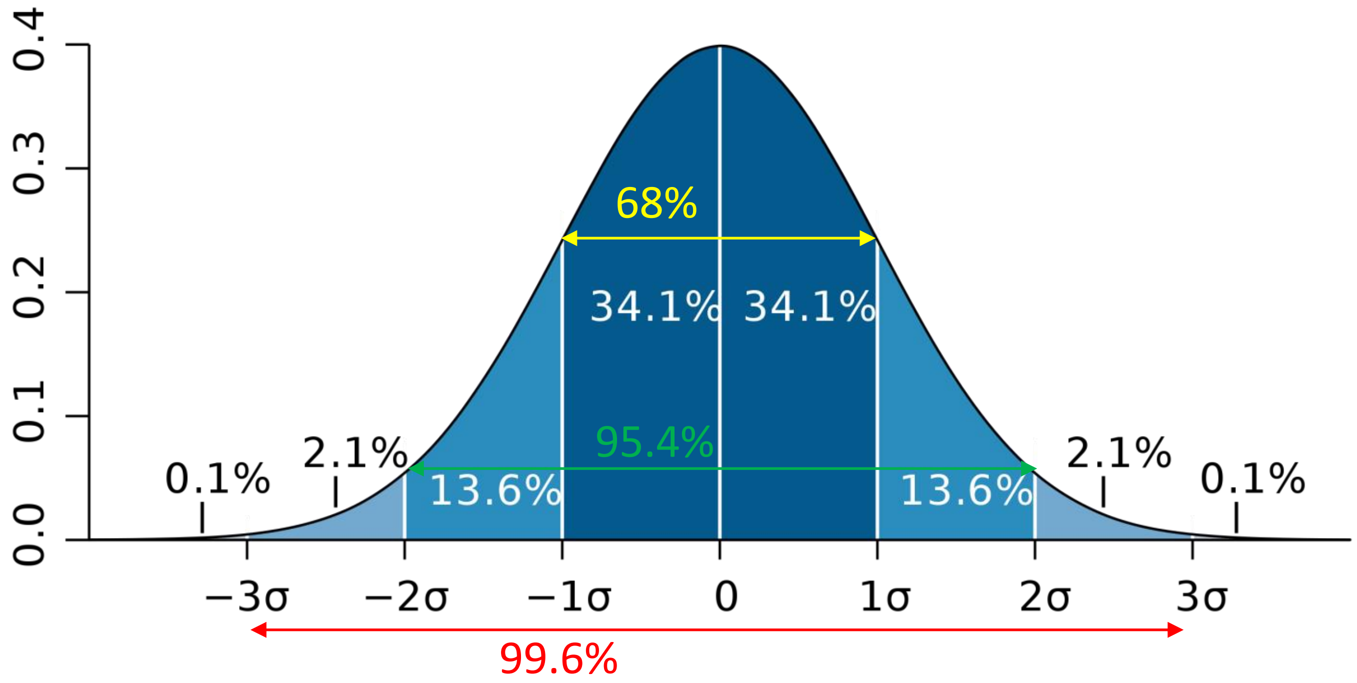
<u>L'écart-type</u> est la racine de la variance :

$$\sigma = \sqrt{Var(x)}$$

On notera parfois aussi la variance comme $\sigma^2 = Var(X)$.

Lien entre Moyenne et Ecart-Type

Dans le cas d'une distribution normale (ou Gaussienne), la relation suivante tient :



ATTENTION

Le lien entre écart-type et intervalles de confiance n’est pas TOUJOURS vrai !

Mesure de corrélation

La COVARIANCE mesure le lien entre deux variables aléatoires:

$$COV(X,Y) = E\left[\left(X - \overline{X}\right) \times \left(Y - \overline{Y}\right)\right]$$

$$= \sum_{i=1}^{N} \frac{\left(X_i - \overline{X}\right)\left(Y_i - \overline{Y}\right)}{N}$$

Mesure de corrélation

Si deux variables aléatoires sont indépendantes, alors $COV(X,Y) = 0$.

On définit le coefficient de corrélation des deux variables aléatoires comme

$$\rho(X,Y) = \frac{Cov(X,Y)}{\sigma_X \times \sigma_Y}$$

Où $\sigma_X = \sqrt{Var(X)}$.

Mesure de corrélation

$$\rho(X, Y) \in [-1, 1]$$

Si les deux variables sont non-corrélées (indépendantes), alors $\rho(X, Y) = \mathbf{0}$.

Si les variables sont fortement corrélées, alors $\rho(X, Y)$ sera proche de ± 1.

Application aux statistiques

Tous les éléments vus jusque-là sont des raisonnements probabilistes, donc théoriques.

En statistiques, nous effectuons le chemin «inverse» – nous observons des évènements et tentons de «deviner» la loi probabiliste qui se cache derrière !

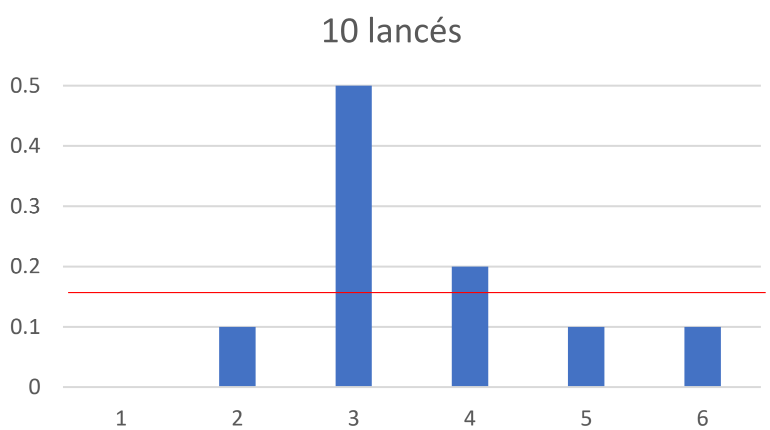
Loi des grands nombres

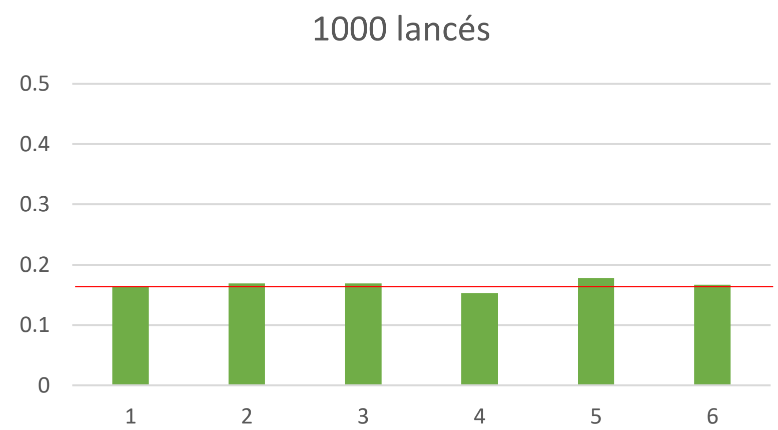
La «Loi des Grands nombres» dit que la répétition d’une même mesure nous permet d’approcher la distribution théorique.

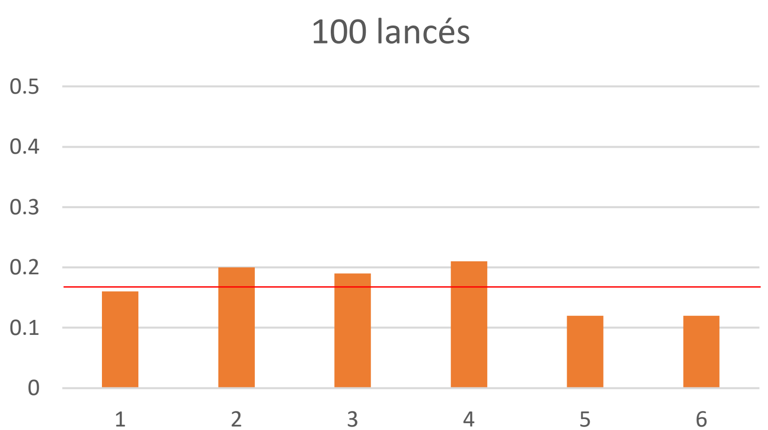
Autrement dit : si l’échantillon est «assez grand», on peut dire qu’il permet de décrire de manière «fiable» la distribution théorique d’une variable aléatoire.

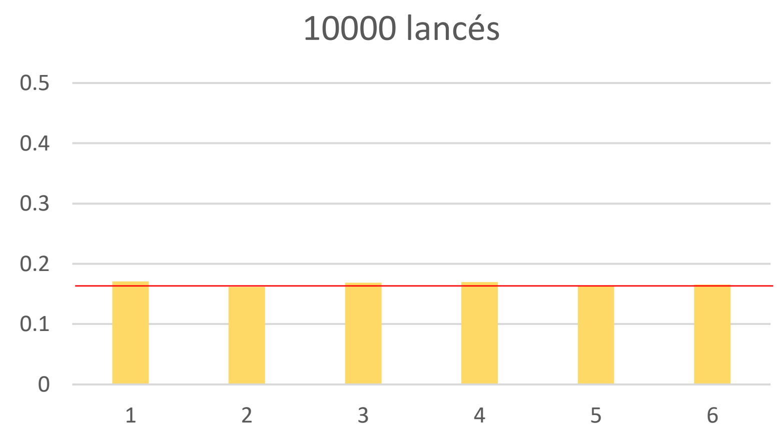
Exemple : lancé de dé

Voici le résultat de lancé de dés :



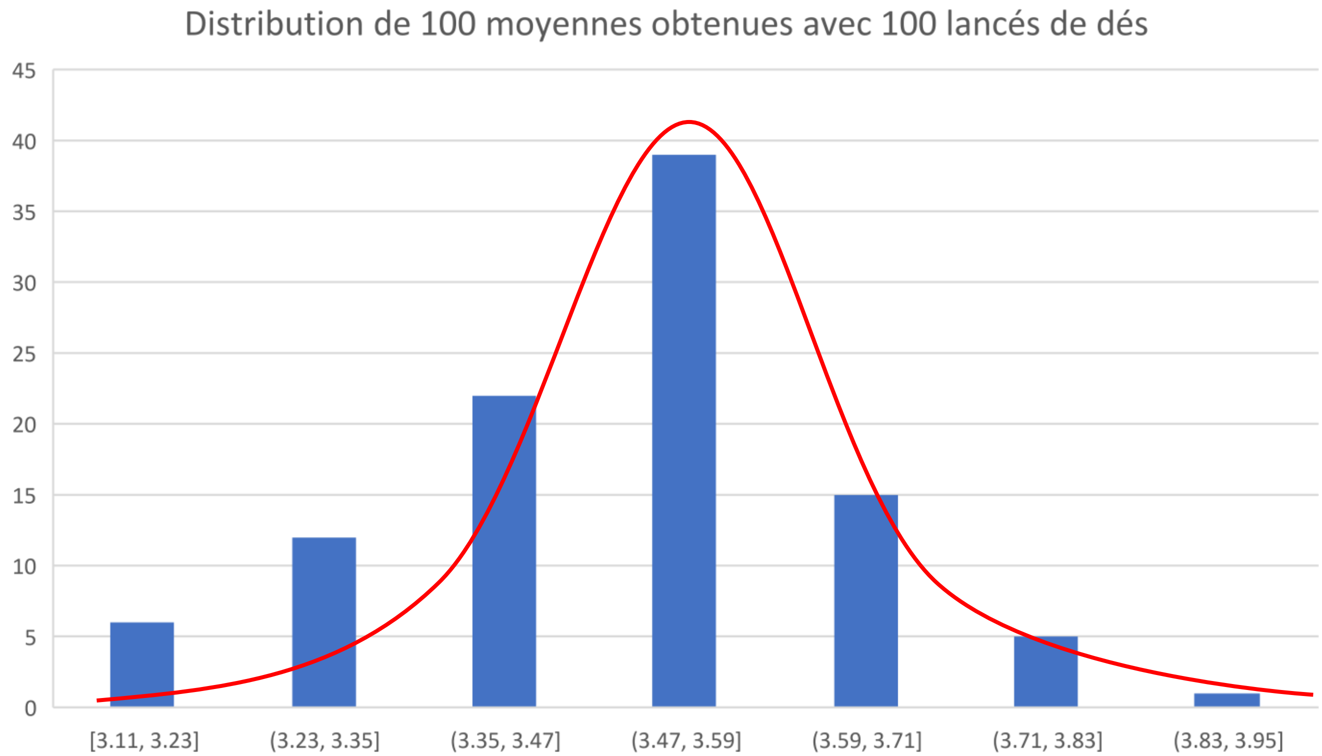






Théorème Central Limite

Si les mesures sont répétées, alors la distribution de la moyenne mesurée converge vers une distribution normale.



2 types de statistiques

- *Statistique Descriptive* a pour objectif de synthèse de données recueillies ;

- *Statistique inférentielle* a pour de faire des prévisions, en particulier comme outil d'aide à la décision;

Vocabulaire

Toute statistique se base sur un ensemble de mesure. On appelle

- *Variables* les caractéristiques à observer
- *Population* l'ensemble des mesures ;
- *Individu* un élément particulier de la population;
- *Echantillon* est un sous-ensemble représentatif des variables observées, s'il n'est pas possible d'étudier l'ensemble de la population.
- *Variabilité* le fait d'observer des différences de variables sur des individus semblables.

Objectif statistique

L'objectif des statistiques est d'étudier la variabilité des individus afin

- D'inférer des estimations (prévisions) sur la population entière, basée sur l'échantillon sur la sous-population,
- Expliquer la cause en lien avec des raisons externes

Exemple

Nous souhaitons observer étudier les effets climatiques sur les récoltes du vignoble lémanique. Nous avons recueilli des mesures sur une période de 100 ans, divisé par région (La Côte, Nord Genevois, Lavaux, ...)

<u>Population</u> => l'ensemble des combinaisons région-année

<u>Individu</u> => une région et une année bien spécifique (p.ex. Lavaux 2012)

<u>Variable</u> => la quantité de raisin, la quantité de vin produit, ...

Formalisation

Soient P notre population et Ω l'ensemble des observations.

Nous notons alors $X = \{x_1, x_2, \ldots\}$ l'échantillon des données ($X \subseteq \Omega$) sur une sous-population $P' = \{p_1, p_2, \ldots\} \subseteq P$.

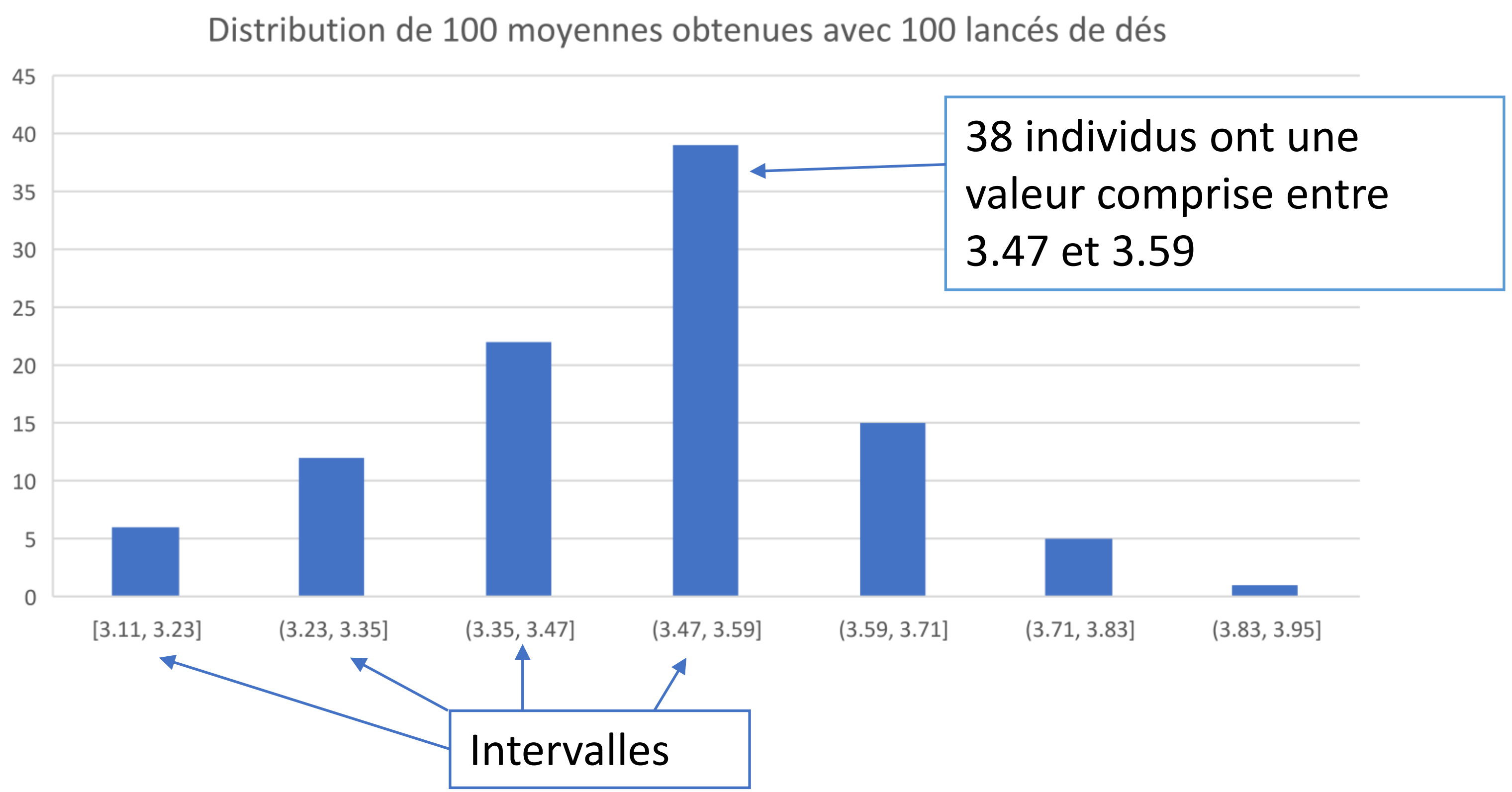
Exemple :
P est l'ensemble des combinaisons région-année possibles.

P' est le sous-ensemble des régions lémaniques sur les 100 dernières années.

X est l'ensembles de données relevées sur la sous-population P'.

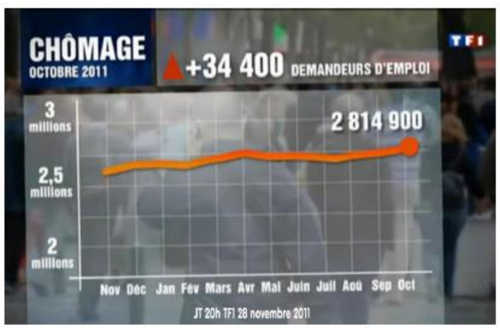
Présentation des données

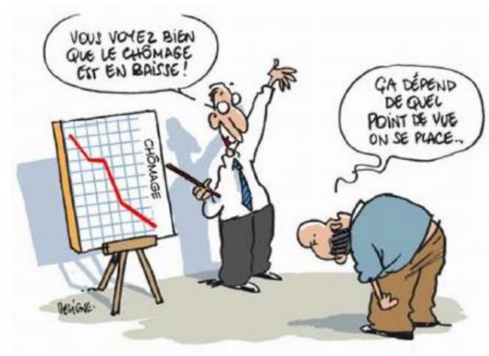
Les données sont souvent présentées sous forme d'histogrammes, par intervalles dans lesquels on compte le nombre d'individus figurant dans un intervalle.



ATTENTION aux échelles !

Tout graphique doit pouvoir être compris sans explications. Pensez à toujours indiquer les valeurs sur les axes ainsi qu'un titre explicatif !

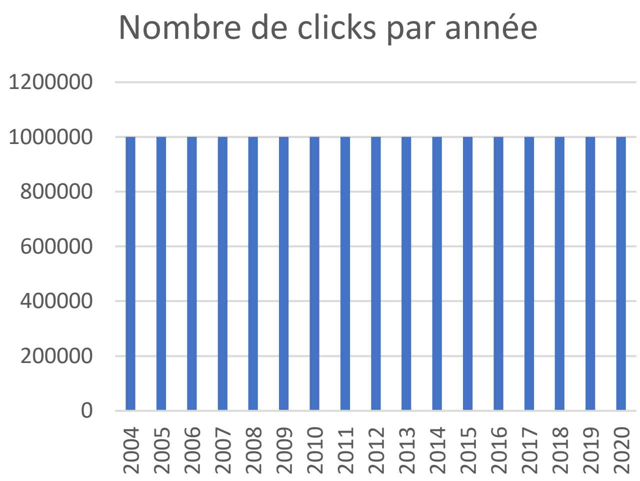


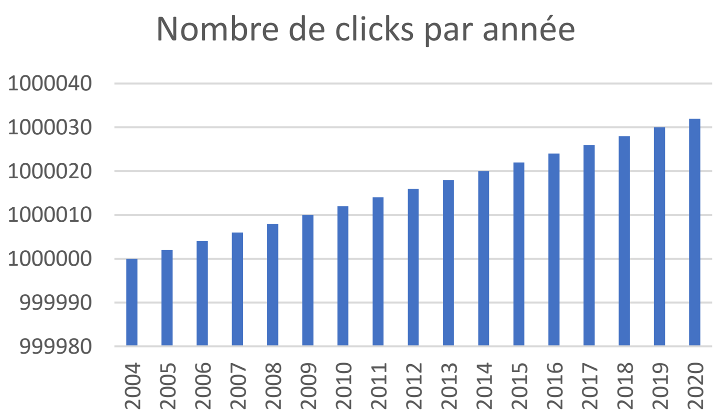


Source: https://hal.archives-ouvertes.fr/cel-01787365/file/Cours_Statistiques_Polisano.pdf

ATTENTION aux échelles !

Le choix des échelles peut être très important !

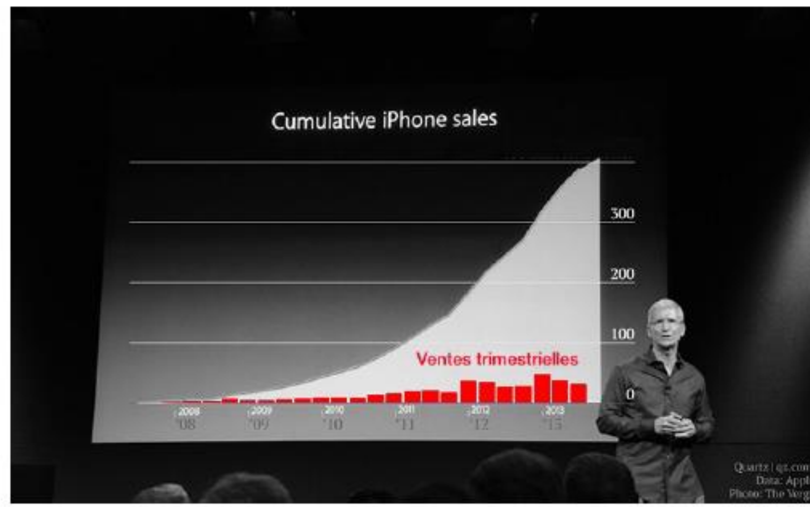




Ce sont bien les mêmes données !!!

Et attention aux interprétations !

Le choix du graphique influence fortement son interprétation ! Présentation de Tim Cook



Ventes cumulées par trimestre

Si on ajoute les ventes trimestrielles

Sur la courbe cumulative, il faut observer la pente !
Celle-ci tend à s'aplatir, donc les ventes diminuent !

Source: https://hal.archives-ouvertes.fr/cel-01787365/file/Cours_Statistiques_Polisano.pdf

Règles usuelles

Les variables sont triées dans un vecteur $\vec{x} \in \mathbb{R}^n$ trié par ordre croissant

- On créé environ $k \approx 1 + \log_2(n)$ intervalles (règle de Sturges) $I_i = [i \times h, (i+1) \times h[, i = 0, \dots, k$

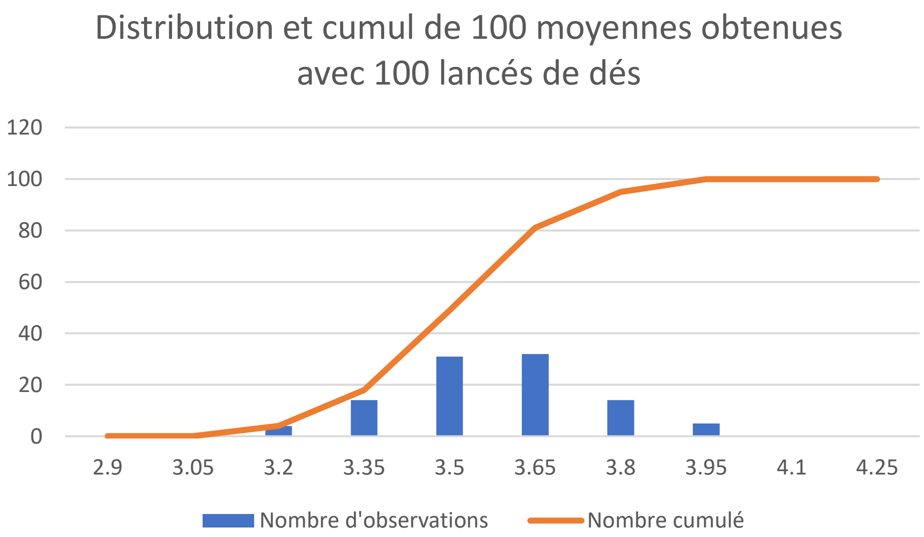
- La largeur d'un intervalle est $\frac{x_n - x_0}{k}$ (valeur max – valeur min)

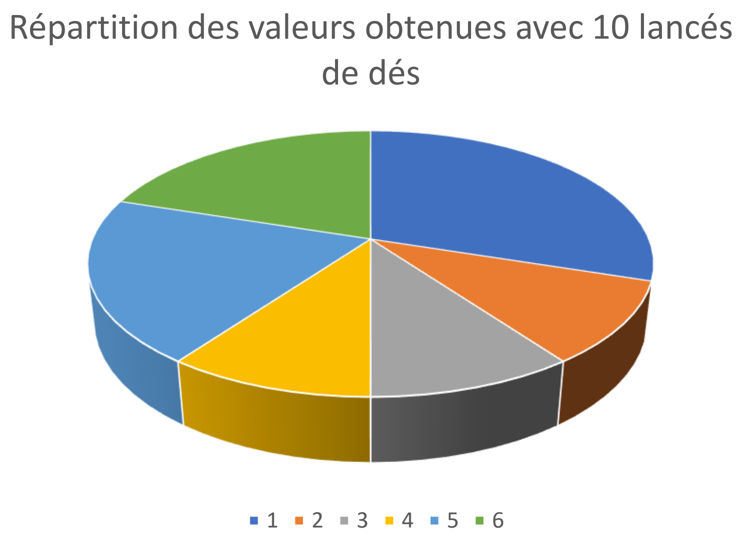
- c_i est le nombre d'individus dans chaque intervalle I_i

- On créé le graphique affichant I_i en abscisse et c_i en ordonnées.

Alternative

On peut également présenter les données cumulativement, ou via un camembert !





Statistiques empiriques

Les statistiques ***empiriques*** d'une série de données $\vec{x} \in \mathbb{R}^n$ triée se calculent de manière similaire aux variables aléatoires, mais ne connaissant pas la probabilité (c'est ce qui est à déterminer !):

<u>Moyenne Empirique</u> $\overline{x} = \frac{1}{N}\sum_{i=1}^{n} x_i$

<u>Médiane Empirique</u> $\tilde{x} = \begin{cases} x_{\frac{n+1}{2}} \; si \; n \; impair \\ \frac{1}{2}\left(x_{\frac{n}{2}} + x_{\frac{n}{2}+1}\right) si \; n \; pair \end{cases}$

<u>Variance Empirique</u> $\sigma_x^2 = \frac{1}{N}\sum_{i=1}^{n} (x_i - \overline{x})^2$

<u>Ecart-type Empirique</u> $\sigma_x = \sqrt{\sigma_x^2} = \sqrt{\frac{1}{N}\sum_{i=1}^{n} (x_i - \overline{x})^2}$

Les Quantiles

Les «***q-quantiles»*** sont des valeurs séparant les données $\vec{x} \in \mathbb{R}^n$ (triées par ordre croissant) en intervalles empiriquement équiprobables (avec la même population).
Notons $\tilde{q}_{\frac{i}{q}}$ le $i^{ème}$ q-quantile ($i = 1, \dots, q$) alors

$$\tilde{q}_{\frac{i}{q}} = \begin{cases} \frac{1}{2}\left(x_{\frac{n\times i}{q}} + x_{\frac{n\times i}{q}+1}\right) \; si \, \frac{n \times i}{q} \in \mathbb{N} \\ x_{\left\lfloor\frac{n\times i}{q}\right\rfloor+1} \; sinon \end{cases}$$

Notez que $\tilde{q}_{\frac{1}{2}} = \tilde{x}$ (la médiane).

Certains q-quantiles spéciaux ont des noms précis (quartiles pour $q = 4$, centiles pour $q = 100$, …)

Corrélation empirique

Soient deux observations $\vec{x}, \vec{y} \in \mathbb{R}^n$ (triées par ordre croissant).

La *covariance empirique* de $\vec{x}$ et $\vec{y}$ est donnée par

$$COV(\vec{x}, \vec{y}) = \sum_{i=1}^{N} \frac{(x_i - \overline{x})(y_i - \overline{y})}{N}$$

La *corrélation empirique* de $\vec{x}$ et $\vec{y}$ est donnée par

$$r_{xy} = \frac{COV(\vec{x}, \vec{y})}{\sigma_x \times \sigma_y}$$

Tests Statistiques

En statistiques, nous cherchons le plus souvent à mettre à l'épreuve une ***hypothèse de travail (*H_1*)*** qu'il s'agit de valider (ou infirmer) par les observations.

Cette hypothèse se fait toujours en opposition à ***l'hypothèse nulle (*H_0*)***, qui est l'hypothèse de référence.

Le test statistique permet de décider, avec un degré de confiance (probabilité), si l'hypothèse de travail est valide (H_0 est rejetée) ou non (H_0 est acceptée).

Exemple

Un tirage de pile ou face donne 8 piles pour 2 faces sur 10 tirages. Est-ce que la pièce est truquée ?

Hypothèses :

H_0 : la pièce n'est pas truquée, et donc N lancés suivent une loi Binomiale à probabilité 0.5 (*hypothèse nulle*)

Note: pour H_0, nous pouvons faire des calculs probabilistes ! Donc H0 est «testable».

H_1 : la pièce est truquée ou il y a triche, donc $\mathbb{P}(pile) > 0.5$ (hypothèse de travail).

Note: pour H_1, nous ne connaissons pas la probabilité de pile – l'hypothèse est alors dite «non-testable».

Exemple

Avec H_0 nous pouvons mesurer la probabilité que 10 lancés produisent un résultat de 8 piles, il s'agit de

$$\mathbb{P}(8) = C_8^{10} \times \left(\frac{1}{2}\right)^8 \left(\frac{1}{2}\right)^{10-8} = \frac{10!}{8! \times (10-8)!} \times \left(\frac{1}{2}\right)^{10} \cong 0.044.$$

Avec une pièce équilibrée, il y a donc 4.4% de chances que le résultat de 8 piles pour 2 faces se produisent.

Si ce pourcentage est trop faible, on dit alors qu'on rejette l'hypothèse H_0 en faveur de H_1. A l'inverse, on rejettera l'hypothèse H_1 en faveur de H_0 si la probabilité est très faible.

ATTENTION: rejeter H_0 ne signifie pas que la H_0 soit fausse et H_1 vrai : H_1 est ***probablement*** vrai !!!

Risques d’erreurs

Notons que nous nous avons 4 cas possibles, dont nous donnons une probabilité

- **ERREUR type I (condamner un innocent)**

  $$\alpha = \mathbb{P}(rejeter\ H_0 | H_0\ est\ vrai)$$

- **Vraisemblance**

  $$1 - \alpha = \mathbb{P}(accepter\ H_0 | H_0\ est\ vrai)$$

- **ERREUR type II (acquitter un coupable)**

  $$\beta = \mathbb{P}(accepter\ H_0 | H_1\ est\ vrai)$$

- Puissance

  $$1 - \beta = \mathbb{P}(rejeter\ H_0 | H_1\ est\ vrai)$$

Région de rejet

Nous définissons une ***région de rejet*** R_α de H_0 comme un ensemble de valeurs suffisamment peu probables sous H_0 pour rejeter H_0.

Autrement dit, la région de rejet comprend toutes les valeurs ayant une probabilité de réalisation $\leq \alpha$.

Cela s’écrit comme

$$R_\alpha = \{x \mid \mathbb{P}(x|H_0) \leq \alpha\}$$

Notez la probabilité conditionnelle $\mathbb{P}(x|H_0)$ => en effet, nous mesurons la probabilité «sous H_0», donc «sachant que H_0 est vraie» !

Exemple

Choisissons $\alpha = 0.05 = 5\%$.
Dans ce cas, l'intervalle de rejet $R_{0.05} = \{8,9,10\}$, donc on rejette H_0 si on observe 8 ou plus piles sur 10 lancés.

Dans notre exemple, ayant observé 8 piles, nous rejetons H_0 est conclurons donc que la pièce est ***probablement truquée.***

<u>ATTENTION:</u>
Ici, cela ne signifie pas que la pièce a 95% de chances d'être truquée, mais bien qu'il y a 95% de chances qu'avec une pièce équilibrée, ce résultat ne se produise pas !