Statistique et Informatique (3I005)

2018-2019

Nicolas Baskiotis - Pierre-Henri Wuillemin

prenom.nom@lip6.fr

Sorbonne Université)

Laboratoire d'Informatique de Paris 6 (LIP6)

Cours 1 :
Probabilités sur des ensembles discrets et dénombrements

Description de l'UE

Objectifs du cours

- Introduction aux domaines :
  - de la théorie des probabilités,
  - de la statistique,
- donner des exemples de leurs applications (en informatique),
- pratiquer les concepts introduits sur des exemples $\longrightarrow$ mini-projets.

Organisation

- Calcul des probabilités (Nicolas Baskiotis – cours 1 à 5) :
  - introduction aux probabilités, conditionnement, marginalisation
  - probabilités discrètes, continues
  - Loi des grands nombres et applications
  - études de différentes lois
- L'inférence statistique (Pierre-Henri Wuillemin – cours 6 à 11) :
  - recueil et analyse des données,
  - estimation, tests et validation,
  - Processus séquentiels.

Description de l'UE (2)

Informations pratiques

- Site Web : `http://3i005.lip6.fr`
- Organisation en *mini-projets* (en python) :
  - TME 1-4 : (intro +) projet 1,
  - TME 5-7 : projet 2,
  - TME 8-11 : projet 3 (+ révisions).

Evaluation

- Les trois mini-projets sont notés
- les mini-projets comptent dans la note finale *dans tous les cas* !
- un partiel et un examen.

Évaluation

Calcul de la note finale

- Note de Contrôle Continu : le partiel (40$\%$),
- note d'écrit : sur 60$\%$, Examen (70$\%$, soit $\sim$ 40$\%$ du total) et note des 3 projets (30$\%$)

Exemple

- un étudiant a eu 12, 13, et 14 aux mini-projets, 12 au partiel et 11 à l'examen. Alors :
- note de CC : $24/40$,
- note écrit : $11 * 0.7 + (13 + 14 + 15)/3 * 0.3 = 11.9/20$, donc $35.7/60$,
- note finale : $59.7/100$.

Plan

De quoi parle ce cours . . .

- Qu'est ce que la chance ? le hasard ? le aléas ?
- Comment mesurer le hasard ?
- Peut on l'utiliser ? Comment ?
- Comment modéliser des phénomènes aléatoires ?
- Comment les étudier ?
- Comment les caractériser ?
- Que peut-on prédire ?
- . . .

Définitions

Probabilités

- domaine des mathématiques qui étudie des phénomènes *aléatoires*,
- fournit des outils pour étudier les *expériences aléatoires* :
  des expériences qui, répétées dans les mêmes conditions, ne donnent pas nécessairement le même résultat,
- modélise à l'aide de ces outils les processus aléatoires pour en étudier le fonctionnement et les résultats.

Exemple : modéliser un lancé de dé

Statistique

- domaine des mathématiques qui étudie la collecte, l'analyse, et l'interprétation de données
- permet d'établir des protocoles expérimentaux et d'analyser les résultats
- permet d'inférer des conclusions sur les processus aléatoires.

Exemple : à partir d'un certain nombre de lancés, le dé est-il biaisé ?

Une (très) petite histoire des probabilités et statistiques

- $-XVI^e$ siècle : préhistoire, (Cardan 1501-1576, Galilée 1564-1642)
- XVI^e-$XVII^e$: la découverte du domaine
  - Fermat (161x-1665), Pascal (1623-1662)
  - Huyghens (1629-1695)
- $XVIII^e - XIX^e$: developpement et premières applications scientifiques
  - Montmort (1678-1719), de Moivre (1667-1754)
  - la dynastie Bernoulli : Jacob (1657-1705), Jean (1667-1748), Daniel (1700-1782), Nicolas (1687-1759)
  - Bayes (1700-1761)
  - Buffon (1707-1788), Simpson (1710-1761), D'alembert (1717-1783)
  - Lagrange (1736-1813), Laplace (1749-1827), Poisson (1781-1840)
- $XIX^e - XX^e$: théorie de la mesure, axiomatisation, applications multiples
  - Tchebychev (1821-1894), Emile Borel (1871-1956), Johann Radon (1887-1956), Paul Lévy (1886-1971), Andreï Kolmogorov (1903-1987)
  - Gibbs (1839-1903), Boltzmann (1844-1906), Poincaré (1854-1912) , Pearson (1857-1936), Markov (1856-1922)
- XX^e- : vraie reconnaissance du domaine, consolidations théoriques et développement de multiples applications.

Exemples d’application : en informatique fondamentale

- Algorithmique : *tri rapide*
  $\longrightarrow$ meilleure performance “en moyenne” que les autres tris ;
  “en moyenne” $\approx$ les valeurs dans le tableau initial sont aléatoires.

- Calcul d’une coupe minimale dans un graphe : algorithme de Karger
- structure de données : *Table de Hashage*
  $\longrightarrow$ propriété souhaitée de `hashCode` : donner des valeurs différentes aux différentes chaînes de caractères stockées dans la table de hachage.
  $\longrightarrow$ nécessite un modèle (probabiliste) des chaînes de caractères qui seront stockées.
- Compression de données : codage d’Huffman

Cryptographie et cryptanalyse

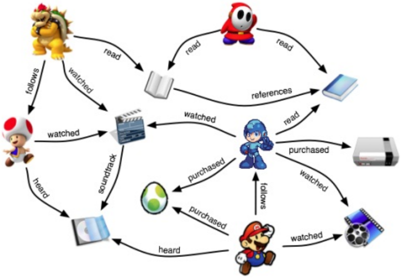
Enigma : machine de cryptage allemande pendant la Seconde Guerre mondiale.

Le décryptage des messages par les alliés a été facilité par un mauvais algorithme de génération de *permutations* aléatoires.

La sécurité des communications sur Internet est gérée par des algorithmes de cryptographie.

Les algorithmes de cryptographie utilisent des générateurs de nombres aléatoires.

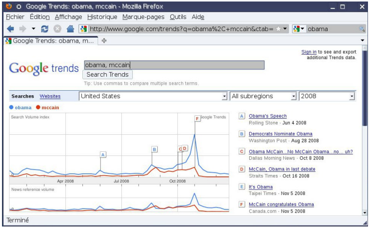
Réciproquement : les cryptanalystes cherchent les *régularités* (déviations par rapport à l'aléatoire) dans les textes cryptés.



Systèmes de reco. (Netflix, Amazon, . . .) :
Les clients qui ont acheté/vu ...
ont aussi acheté/vu ...

Analyses statistiques des achats/recherches des différents produits

Ciblage publicitaire

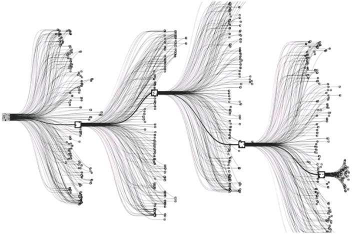


Google Trends : analyse des requêtes effectuées par les utilisateurs de Google.

Applications possibles : suivi des intérêts dans une population, détection des épidémies, ...

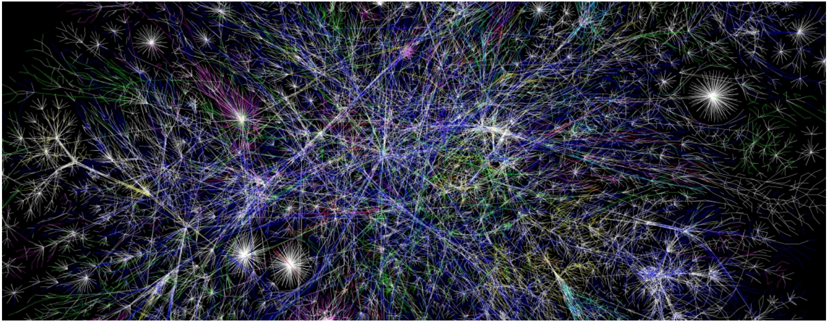
IA et jeux

- Exploration efficiente des possibilités dans une combinatoire élevée,
- Modélisation de l'adversaire (Poker par exemple)
- Matchmaking

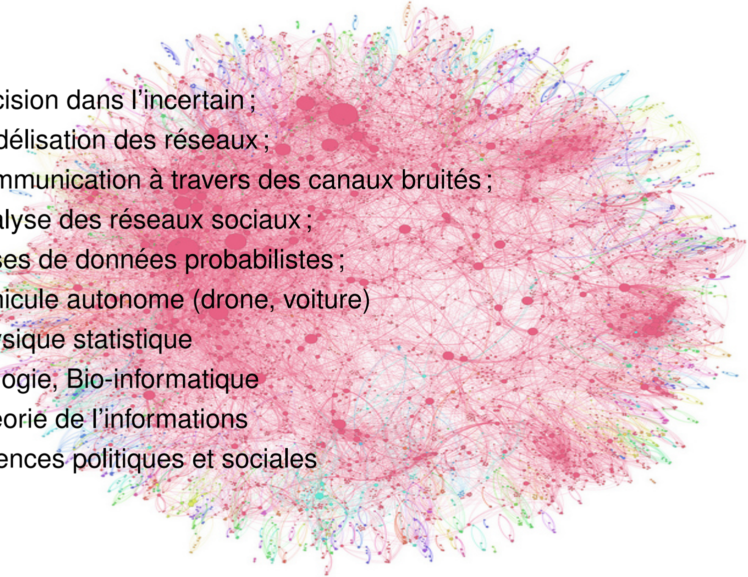


Machine Learning (Apprentissage statistique)

- Apprentissage bayésien
- Réseaux de neurones
- Applications en
  - Classification (image, texte, ...)
  - traduction automatique,
  - génération automatique (musique, textes)
  - Moteur de recherche
  - Interface cerveau-machine (BCI)
  - Recommendation



Et bien d'autres...



- Décision dans l'incertain ;
- Modélisation des réseaux ;
- Communication à travers des canaux bruités ;
- Analyse des réseaux sociaux ;
- Bases de données probabilistes ;
- Véhicule autonome (drone, voiture)
- Physique statistique
- Biologie, Bio-informatique
- Théorie de l'informations
- Sciences politiques et sociales
- ...

Plan

Une probabilité ?

Trois sachets, un croissant . . .

- Pourquoi a-t-on une chance sur 3 de trouver le croissant ?
- Est-ce toujours le cas ?
- Que veut dire *chance* dans ce contexte ?

Une probabilité ?

Trois sachets, un croissant . . .

- Pourquoi a-t-on une chance sur 3 de trouver le croissant ?
- Est-ce toujours le cas ?
- Que veut dire *chance* dans ce contexte ?

La probabilité d'un événement

- c'est la fréquence d'apparition de l'événement, le nombre de fois où il apparaît rapporté au nombre d'expériences.
- Notions (intuitives) :
  - d'expérience : un cadre bien défini, avec des conditions initiales et un ensemble de résultats déterminés
  - d'événement : un résultat de l'expérience
  - de répétition : l'expérience peut être reproduite dans les mêmes conditions !

Un peu plus compliqué

30 croissants, 30 pains au chocolat, 20 pains aux raisins, 10 pains au lait, 10 chaussons

Qu'est ce qui est équiprobable ?
Quelle est la probabilité :

- d'un pain au chocolat ?
- si 5 croissants ont été tirés avant ?
- qu'un pain soit tiré ?

Plusieurs tirages

Qu'elle est la probabilité :

- de tirer aucun croissants au bout de deux tirages ? au bout de trois ?
- de tirer au moins un croissant ?

Et le café ?

Une machine à café a un seul bouton

Elle peut faire un café court ou long, avec du sucre et/ou du lait.

- Quelle est la probabilité d'avoir un café court sucré ?

Et le café ?

Une machine à café a un seul bouton

Elle peut faire un café court ou long, avec du sucre et/ou du lait.

- Quelle est la probabilité d'avoir un café court sucré ?

$\Rightarrow$ Telle quelle, la question a autant de sens que "quel est l'âge du client" ...

Et le café ?

Une machine à café a un seul bouton

Elle peut faire un café court ou long, avec du sucre et/ou du lait.

- Quelle est la probabilité d'avoir un café court sucré ?

$\Rightarrow$ Telle quelle, la question a autant de sens que "quel est l'âge du client" ...

De quoi a-t-on besoin pour répondre à la question ?

Et le café ?

Une machine à café a un seul bouton

Elle peut faire un café court ou long, avec du sucre et/ou du lait.

- Quelle est la probabilité d'avoir un café court sucré ?

$\Rightarrow$ Telle quelle, la question a autant de sens que "quel est l'âge du client" . . .

De quoi a-t-on besoin pour répondre à la question ?

- Idéalement, la probabilité de chaque événement *(court, lait, sucré)*, *(court, lait, non sucré)*, *(long, lait, sucré)*, *(long, lait, non sucré)*, *(court,non lait, sucré)*, *(court, non lait, non sucré)*, *(long, non lait, sucré)*, *(long, non lait, non sucré)*
- Peut-on tout calculer ?

$\Rightarrow$ Oui, ce sont tous les événements , tous les résultats attendus

- Exemple : probabilité de *court lait* = *court, lait, sucré* + *court, lait, non sucré*

Et le café ?

Une machine à café a un seul bouton

Elle peut faire un café court ou long, avec du sucre et/ou du lait.

- Quelle est la probabilité d'avoir un café court sucré ?

$\Rightarrow$ Telle quelle, la question a autant de sens que "quel est l'âge du client" . . .

De quoi a-t-on besoin pour répondre à la question ?

- Probabilités des événement *lait*, *sucré* suffisent ?

Et le café ?

Une machine à café a un seul bouton

Elle peut faire un café court ou long, avec du sucre et/ou du lait.

- Quelle est la probabilité d'avoir un café court sucré ?

$\Rightarrow$ Telle quelle, la question a autant de sens que "quel est l'âge du client" ...

De quoi a-t-on besoin pour répondre à la question ?

- Probabilités des événement *lait*, *sucré* suffisent ?

$\Rightarrow$ non, on ne sait pas court ou long ...

- Et avec *court* en plus ?

Et le café ?

Une machine à café a un seul bouton

Elle peut faire un café court ou long, avec du sucre et/ou du lait.

- Quelle est la probabilité d'avoir un café court sucré ?

$\Rightarrow$ Telle quelle, la question a autant de sens que "quel est l'âge du client" ...

De quoi a-t-on besoin pour répondre à la question ?

- Probabilités des événement *lait*, *sucré* suffisent ?

$\Rightarrow$ non, on ne sait pas court ou long ...

- Et avec *court* en plus ?

$\Rightarrow$ oui, *court* et *long* sont complémentaires !

- probabilité de *court* = 1 - *long*

Plan

Probabilités sur les ensembles discrets

Pour modéliser une expérience aléatoire :

Nous avons besoin des notions de :

- *événement élémentaire* : un résultat simple non composé de l'expérience
- *événement* : un résultat simple ou composé de plusieurs événements élémentaires de l'expérience
- *univers* : l'ensemble de tous les résultats possibles

Formalisation

- Soit Ω, un ensemble dénombrable, appelé univers,
  - Ω représente l'ensemble de tous les résultats possibles d'une expérience aléatoire
- un élément $\omega \in \Omega$ est *un événement élémentaire*,
- un sous-ensemble E de Ω est un *événement*.

Probabilités sur les ensembles discrets

Pour modéliser une expérience aléatoire :

Nous avons besoin des notions de :

- *événement élémentaire* : un résultat simple non composé de l'expérience
- *événement* : un résultat simple ou composé de plusieurs événements élémentaires de l'expérience
- *univers* : l'ensemble de tous les résultats possibles

Formalisation

- Soit Ω, un ensemble dénombrable, appelé univers,
  - Ω représente l'ensemble de tous les résultats possibles d'une expérience aléatoire
- un élément $\omega \in \Omega$ est *un événement élémentaire*,

$\Rightarrow$ Un croissant, un pain au chocolat, un pain aux raisins, un café court sucré sans lait

- un sous-ensemble E de Ω est un *événement*.

$\Rightarrow$ Un pain, quelque chose sans beurre, un café court

Probabilités sur des ensembles discrets (2)

Mesure de probabilité (ou distribution) : caractérise l'aléa

- elle définie une probabilité pour chaque événement, qui correspond à la fréquence d'apparition de l'événement, entre 0 et 1
- la probabilité de l'univers est de 1 : au moins un des événements de l'univers se réalise lors d'une expérience
- la probabilité de deux événements qui ne peuvent arriver en même temps (incompatibles) est la somme de leur probabilité.
- la probabilité qu'aucun événement de l'univers n'arrive est donc de 0.

Elle est entièrement définie par les probas des événements élémentaires.

Définitions

Soit l'univers Ω, ensemble discret.

Une mesure de probabilité sur Ω est une fonction $P : \mathcal{P}(\Omega) \rightarrow [0, 1]$ tq :

- $P(\Omega) = 1$ (Ω est l'événement certain),
- pour tout événement E, $P(E) \geq 0$,
- Pour toute suite $(E_i)_{i \in \mathbb{N}}$ d'événements deux à deux disjoints (*incompatibles*, $E_i \cap E_j = \emptyset, i \neq j$) : $P(\bigcup_i E_i) = \sum_i P(E_i)$.

avec Ω ensemble discret et $\mathcal{P}(\Omega)$ l'ensemble des sous-ensembles de Ω.

Remarques

- Ces 3 axiomes définissent le cadre des probabilités discrètes.
- Les événements élémentaires sont forcément incompatibles.
- La définition de la probabilité comme étant la limite de la fréquence du nombre d'apparition de l'événement en répétant à l'infini l'expérience n'est pas considérée comme axiome (pourquoi ?), mais est déduite de ces 3 axiomes.

Probabilités sur des ensembles discrets (3)

Définitions

- Fonction de masse p associée à P : $\forall\omega \in \Omega, p(\omega) = P(\{\omega\})$
(rappel : ω événement élémentaire !)
- Pour tout événement E :

$$P(E) = \sum_{\omega \in E} p(\omega)$$

- La probabilité uniforme sur un univers *fini* Ω est définie par la fonction de masse : $p(\omega) = \frac{1}{card(\Omega)}$. Dans ce cas, $\forall E \in \mathcal{P}(\Omega),\ P(E) = \frac{card(E)}{card(\Omega)}$

Interprétation : si on répète (indéfiniment) l'expérience aléatoire

- le résultat de l'expérience sera ω avec une fréquence de $P(\{\omega\})$,
- un événement E se produit avec une fréquence $P(E)$
$\rightarrow$ le résultat appartient à l'ensemble E avec une fréquence $P(E)$.

Probabilités sur des ensembles discrets (4)

Propriétés

- $P(\emptyset) = 0$, ($\emptyset$ est l'événement impossible)
  ne pas avoir de café ...
- $P(\bar{E}) = 1 - P(E)$ ($\bar{E}$: complémentaire de E dans Ω)
  café court et café long
- $P(E \cup F) = P(E) + P(F) - P(E \cap F)$
  P(*café au lait* ou *café sucré*) = P(*café au lait*)+P(*café sucré*) - P(*café au lait et sucré*)
- $E \subset F \Rightarrow P(F) = P(F \backslash E) + P(E) \Rightarrow P(E) \leq P(F)$
  ($F \backslash E$: ensemble des éléments de F qui ne sont pas dans E),
  La probabilité d'un café au lait sucré est inférieure à celle d'un café au lait
- $P(\bigcup_i E_i) \leq \sum_i P(E_i)$

Comment résoudre un problème de probabilité discrète ?

4 étapes :

1. Bien définir l'univers ! une majorité d'erreurs/paradoxes proviennent d'une confusion sur les événements élémentaires.
2. Déterminer le(s) événement(s) d'intérêt(s).
3. Assigner les probabilités $\Rightarrow$ la plupart du temps, savoir compter ! (plus difficile qu'il n'y paraît, cf dénombrements).
4. Calculer la probabilité des événements d'intérêts (généralement simple, par addition/soustraction des probabilités définies précédemment).

Prochaine partie : les dénombrements, ou comment apprendre à compter

Plan

Principes élémentaires

Dénombrer les résultats de tuples d'expériences

- Soit 2 expériences telles que il y ait n_1 résultats possibles pour la première expérience et pour chacun de ces résultats n_2 résultats pour la deuxième. Alors il y a en tout $n_1 \times n_2$ résultats possibles.
- Généralisation : Soit r expériences et $(n_1, \ldots, n_r)$ le nombre de résultats possibles pour chacune de ces expériences indépendamment des autres. Alors le nombre de résultats possibles est $n_1 \times n_2 \times \ldots \times n_r$.

Exemples

- Nombre de résultats pour le jet d'un dé suivi du jet d'une pièce ?
- Nombre de plaques d'immatriculation formé de deux lettres, 3 chiffres puis deux lettres ?
- Nombre de configurations au jeu de go (plateau de 19×19) ?
- Combien de questions binaires (oui/non) sont nécessaires pour différencier 10 millions de personnes ?

Principes élémentaires

Dénombrer les résultats de tuples d'expériences

- Soit 2 expériences telles que il y ait n_1 résultats possibles pour la première expérience et pour chacun de ces résultats n_2 résultats pour la deuxième. Alors il y a en tout $n_1 \times n_2$ résultats possibles.
- Généralisation : Soit r expériences et $(n_1, \ldots, n_r)$ le nombre de résultats possibles pour chacune de ces expériences indépendamment des autres. Alors le nombre de résultats possibles est $n_1 \times n_2 \times \ldots \times n_r$.

Exemples

- Nombre de résultats pour le jet d'un dé suivi du jet d'une pièce ?
  6 résultats pour le dé, 2 pour la pièce, donc $6 * 2$.
- Nombre de plaques d'immatriculation formé de deux lettres, 3 chiffres puis deux lettres ?
- Nombre de configurations au jeu de go (plateau de 19×19) ?
- Combien de questions binaires (oui/non) sont nécessaires pour différencier 10 millions de personnes ?

Principes élémentaires

Dénombrer les résultats de tuples d'expériences

- Soit 2 expériences telles que il y ait n_1 résultats possibles pour la première expérience et pour chacun de ces résultats n_2 résultats pour la deuxième. Alors il y a en tout $n_1 \times n_2$ résultats possibles.
- Généralisation : Soit r expériences et $(n_1, \ldots, n_r)$ le nombre de résultats possibles pour chacune de ces expériences indépendamment des autres. Alors le nombre de résultats possibles est $n_1 \times n_2 \times \ldots \times n_r$.

Exemples

- Nombre de résultats pour le jet d'un dé suivi du jet d'une pièce ?
- Nombre de plaques d'immatriculation formé de deux lettres, 3 chiffres puis deux lettres ?
  $26 * 26 * 10 * 10 * 10 * 26 * 26$
- Nombre de configurations au jeu de go (plateau de 19×19) ?
- Combien de questions binaires (oui/non) sont nécessaires pour différencier 10 millions de personnes ?

Dénombrer les résultats de tuples d'expériences

- Soit 2 expériences telles que il y ait n_1 résultats possibles pour la première expérience et pour chacun de ces résultats n_2 résultats pour la deuxième. Alors il y a en tout $n_1 \times n_2$ résultats possibles.
- Généralisation : Soit r expériences et $(n_1, \ldots, n_r)$ le nombre de résultats possibles pour chacune de ces expériences indépendamment des autres. Alors le nombre de résultats possibles est $n_1 \times n_2 \times \ldots \times n_r$.

Exemples

- Nombre de résultats pour le jet d'un dé suivi du jet d'une pièce ?
- Nombre de plaques d'immatriculation formé de deux lettres, 3 chiffres puis deux lettres ?
- Nombre de configurations au jeu de go (plateau de 19×19) ?
  3 états possibles par case, donc $3^{19\times 19}$
- Combien de questions binaires (oui/non) sont nécessaires pour différencier 10 millions de personnes ?

Principes élémentaires

Dénombrer les résultats de tuples d'expériences

- Soit 2 expériences telles que il y ait n_1 résultats possibles pour la première expérience et pour chacun de ces résultats n_2 résultats pour la deuxième. Alors il y a en tout $n_1 \times n_2$ résultats possibles.
- Généralisation : Soit r expériences et $(n_1, \ldots, n_r)$ le nombre de résultats possibles pour chacune de ces expériences indépendamment des autres. Alors le nombre de résultats possibles est $n_1 \times n_2 \times \ldots \times n_r$.

Exemples

- Nombre de résultats pour le jet d'un dé suivi du jet d'une pièce ?
- Nombre de plaques d'immatriculation formé de deux lettres, 3 chiffres puis deux lettres ?
- Nombre de configurations au jeu de go (plateau de 19×19) ?
- Combien de questions binaires (oui/non) sont nécessaires pour différencier 10 millions de personnes ?
  Soit n ce nombre, on veut $2^n = 10^7$, donc $n = 10^7 = 23.2$, donc 24.

Exemples d'application

Lancé simultané de trois dés

- Univers ? Nombre d'événements élémentaires ?
- Probabilité de l'événement E : *La somme des 3 chiffres est inférieure stricte à* 5 ?
- Quelle est la probabilité :
  - d'obtenir aucun 6 ?
  - que $i = j = k$?

Exemples d'application

Lancé simultané de trois dés

- Univers ? Nombre d'événements élémentaires ?
  Evénements élémentaires : $(i, j, k) \in \{1, \ldots 6\}^3$. L'univers (l'ensemble des événements possibles) est :

$$
\begin{aligned}
\Omega &= \{(i, j, k) \mid i \in \{1, .., 6\}, j \in \{1, .., 6\},\ k \in \{1, .., 6\}\} \\
&= \{(1, 1, 1), (1, 1, 2), ..., (1, 2, 1), (1, 2, 2)..., (6, 6, 6)\}, \\
card(\Omega) &= 6^3 = 216
\end{aligned}
$$

  Si on considère chaque dé équilibré, alors chaque événement élémentaire est équiprobable :

$$\forall (i, j, k) \in \Omega, P\left((i, j, k)\right) = \frac{1}{216}$$

  $E = \{(1, 1, 1), (2, 2, 2), (3, 3, 3), (4, 4, 4), (5, 5, 5), (6, 6, 6)\}$ représente l'événement : *les 3 dés sont égaux*.

- Probabilité de l'événement E : *La somme des 3 chiffres est inférieure stricte à* 5 ?

Exemples d'application

Lancé simultané de trois dés

- Univers ? Nombre d'événements élémentaires ?
- Probabilité de l'événement E : *La somme des 3 chiffres est inférieure stricte à* 5 ?
  $P(E) = \frac{4}{256}$. En effet : $E = \{(1,1,1),(1,1,2),(1,2,1),(2,1,1)\}$.
- Quelle est la probabilité :
  - d'obtenir aucun 6 ?
  - que $i = j = k$?

Bien définir un problème

Problème du prince de Toscanne

Pourquoi en lançant trois dés, obtient-on plus souvent un total de 10 points qu'un total de 9 points, alors qu'il y a 6 façons d'obtenir ces deux totaux ?

9 pts	10 pts
$6+2+1$	$6+3+1$
$5+2+2$	$6+2+2$
$5+3+1$	$5+4+1$
$4+3+2$	$5+3+2$
$4+4+1$	$4+4+2$
$3+3+3$	$4+3+3$

Premier réflexe : bien définir les événements et l'univers !

Bien définir un problème

Problème du prince de Toscanne

Pourquoi en lançant trois dés, obtient-on plus souvent un total de 10 points qu'un total de 9 points, alors qu'il y a 6 façons d'obtenir ces deux totaux ?

Premier réflexe : bien définir les événements et l'univers !

Soit $E_{i,j,k}$ l'évènement : $\{i, j, k\}$ sont obtenus sur les dés, sans tenir compte de l'ordre :
pour $i \leq j \leq k$, $E_{i,j,k} = \{(i,j,k),(j,i,k),(j,k,i),(k,j,i),(k,i,j),(i,k,j)\}$
Il y a 6 événements $E_{i,j,k}$ avec $i + j + k = 10$, et 6 qui donnent une somme à 9.

Bien définir un problème

Problème du prince de Toscanne

Pourquoi en lançant trois dés, obtient-on plus souvent un total de 10 points qu'un total de 9 points, alors qu'il y a 6 façons d'obtenir ces deux totaux ?

Premier réflexe : bien définir les événements et l'univers !

Les événements $\Omega_{i,j,k}$ ne sont pas "équiprobables" :

- si $i \neq j \neq k \neq i$, alors 6 manières d'obtenir le résultat, $P(\Omega_{i,j,k}) = \frac{6}{216}$,
- si $i = j \neq k$, alors 3 manières, $P(\Omega_{i,j,k}) = \frac{3}{216}$
- si $i = j = k$, alors une seule possibilité, $P(\Omega_{i,j,k}) = \frac{1}{216}$

$P(\{i+j+k=9\}) = \frac{6+3+6+6+3+1=25}{216}$ et $P(\{i+j+k=10\}) = \frac{6+3+6+6+3+3=27}{216}$

Permutation, arrangement, combinaison

Soit un ensemble $\Omega = \{e_1, \dots, e_n\}$, $card(\Omega) = |\Omega| = n$

Un tirage peut être

- ordonné : l'ordre des éléments compte (notation k-uplet (a, b, c, d))
- non ordonné : l'ordre ne compte pas (notation ensemble $\{a, b, c\}$)
- avec ou sans remise : un élément peut être tiré plusieurs fois ou non

Dénombrement de k-uplets

- nombre de k-uplets d'éléments de Ω : n^k (tirage avec remise, ordonné)
- nombre de k-uplets d'éléments distincts (tirage sans remise, ordonné, nombre d'arrangements de k parmi n) : $A_n^k = n \times (n-1) \times ... \times (n-k+1)$
- Nombre de permutations (cas $n = k$) : $n! = n \times (n-1) \times ... \times 1$
- Combinaison de k éléments (sous-ensembles distincts de k éléments) : $C_n^k = \frac{A_n^k}{k!} = \frac{n!}{k!(n-k)!}$ (tirage sans remise non ordonné).

Nombre de sous-ensembles

Nombre de sous-ensembles

- Le nombre de sous-ensembles distincts de cardinal k contenus dans E : $C_n^k = \frac{n!}{k!(n-k)!}$ (tirage sans remise non ordonné, C_n^k s'appelle aussi le *nombre de combinaisons de k parmi n éléments*
- Formule du binôme de Newton

$$(x+y)^n = \sum_{k=0}^{n} C_n^k x^{n-k} y^k \Rightarrow card\left(\mathcal{P}(\Omega)\right) = 2^n$$

- $C_n^k = \frac{A_n^k}{k!}$.

Dénombrements : exercices

Exemple du jeu de 52 cartes.

- On peut :
  - tirer avec ou sans remise
  - considérer la séquence des cartes (tirage ordonné) ou l'ensemble obtenu (non ordonné).
- Quelle est la probabilité d'obtenir :
  - d'obtenir une configuration donnée après un mélange aléatoire ?
  - 1,2,3,4 de la même couleur et dans l'ordre ?
  - que des dames sur 4 tirages ?
  - aucune dames sur 2 tirages ?

Exemple du jeu de 52 cartes.

- Quelle est la probabilité d'obtenir :
  - d'obtenir une configuration donnée après un mélange aléatoire ?
    Univers Ω ensemble des suites de 52 cartes : 52! possibilités, donc $\Rightarrow \frac{1}{52!}$
  - 1,2,3,4 de la même couleur et dans l'ordre ?
    Univers Ω ensemble de quadruplets de cartes, 4 couleurs et 1 configuration soit : $\Rightarrow \frac{4}{52*51*50*49}$
  - que des dames sur 4 tirages ?
    Univers Ω ensemble de 4 cartes (sans ordre) $\Rightarrow \frac{1}{C_{52}^4}$
  - aucune dames sur 2 tirages ?
    Univers Ω ensemble de 2 cartes $\Rightarrow \frac{C_{48}^2}{C_{52}^2}$

Dénombrements : exercices

Le singe savant

- Un singe organise au hasard 26 cartes représentant les lettres de l'alphabet. Quelle est la probabilité que le mot `SINGE` apparaîsse dans la chaîne ainsi formée ?
- On ne lui donne que les lettres du mot `SINGE`. Quelle est cette fois la probabilité ?
- Et si le mot était `OUISTITI` ?

Le singe savant

- Un singe organise au hasard 26 cartes représentant les lettres de l'alphabet. Quelle est la probabilité que le mot `SINGE` apparaîsse dans la chaîne ainsi formée ?
- On ne lui donne que les lettres du mot `SINGE`. Quelle est cette fois la probabilité ?
- Et si le mot était `OUISTITI` ?

- Le mot `SINGE` apparaît dans A_{26}^{21} configurations, il y a $26!$ configurations possibles équi-probables, donc la probabilité est de $\frac{1}{21!}$.
- Cette fois, une configuration pour $5!$ possibles, donc $\frac{1}{5!}$.
- On peut numéroter les I : I_1, I_2, I_3, et les T : T_1, T_2. Il y a $3! * 2! * 1$ configurations possibles en échangeant les I et les T, donc la probabilité est $\frac{12}{8!}$.

Dénombrements : exercices

Tiroir à chaussettes

Un tiroir contient 20 chaussettes de 10 paires différentes. On en tire 4 au hasard. Quelle est la probabilité de :

- obtenir 2 paires ;
- d'obtenir au moins 1 paire.

Dénombrements : exercices

Tiroir à chaussettes

Un tiroir contient 20 chaussettes de 10 paires différentes. On en tire 4 au hasard. Quelle est la probabilité de :

- obtenir 2 paires ;
- d'obtenir au moins 1 paire.

$\Rightarrow$ On numérote les chaussettes de 1 à 20, tirage sans remise non ordonné,

- $\Omega = \{\{i, j, k, l\}, i \neq j \neq k \neq l\}$, $card(\Omega) = C_{20}^4$;
- il y a C_{10}^2 manières de faire 2 paires
- on calcul l'événement complémentaire, $\frac{20*18*16*14}{4!}$

Le problème des partis
Problème du Chevalier de Méré (Pascal, Fermat)

Deux joueurs jouent à pile ou face

- 128 euros sont en jeu, 64 de la poche de chacun ;
- le premier joueur marque 1 point si pile sort, sinon c'est le second ;
- le premier qui arrive à 7 points gagne tout.
- La partie s'arrête sur un score donné (x, y) subitement ;
- comment répartir équitablement l'argent ?

En particulier :

- si le score est de $(6, 6)$?
- si le score est de $(6, 5)$?
- si le score est $(1, 0)$?
- si le score est $(6, 0)$?

Le problème des partis
Problème du Chevalier de Méré (Pascal, Fermat)

Deux joueurs jouent à pile ou face

- 128 euros sont en jeu, 64 de la poche de chacun ;
- le premier joueur marque 1 point si pile sort, sinon c'est le second ;
- le premier qui arrive à 7 points gagne tout.
- La partie s'arrête sur un score donné (x, y) subitement ;
- comment répartir équitablement l'argent ?

Solution de Fermat

- On note *p* pour pile, *f* pour face ;
- soit p le nombre de points manquant au premier joueur, q au second pour gagner ;
- une succession de partie : une séquence *pffppf*..
- dénombrer les parties favorables à chaque joueur.

Dénombrements : exercices

Deux personnes dans une file d'attente de n personnes dans un ordre aléatoire

- avec quelle probabilité sont-ils les deux premiers ?
- avec quelle probabilité sont-ils distants de r places ?

Rencontres

- On considère n points distincts dans le plan, tous reliés 2 à 2 par des arêtes, coloriées soit en bleu soit en rouge.
  De combien de manière peut-on colorer les arêtes ?
- Soit n personnes. Montrer que si $n \geq 6$ il est toujours possible de trouver 3 personnes telles que soit elles se connaissent toutes, soit aucune des 3 ne connaît aucune des autres.
- Et si $n = 5$?

Dénombrements : exercices

Deux personnes dans une file d'attente de n personnes dans un ordre aléatoire

- avec quelle probabilité sont-ils les deux premiers ?
- avec quelle probabilité sont-ils distants de r places ?

$\Rightarrow$ Soit i et j les positions des deux personnes, $\Omega = \{\{i, j\}, i < j \in \{1, \cdots n\}\}$, $card(\Omega) = C_n^2$, équi-probables,

- il n'y a qu'une configuration
- il y a $n - r$ configurations

Rencontres

- On considère n points distincts dans le plan, tous reliés 2 à 2 par des arêtes, coloriées soit en bleu soit en rouge.
  De combien de manière peut-on colorer les arêtes ?
- Soit n personnes. Montrer que si $n \geq 6$ il est toujours possible de trouver 3 personnes telles que soit elles se connaissent toutes, soit aucune des 3 ne connaît aucune des autres.