Les algorithmes probabilistes

Johanne Cohen

Les algorithmes probabilistes

Un **algorithme probabiliste** est un algorithme dans lequel certaines de ces instructions sont de la forme :

tirer un entier selon la loi uniforme entre 1 et n

Son utilisation permet de

1. Concevoir un algorithme plus simple que l'algorithme déterministe pour le même problème.

   Par exemple : Vérification d'identités

2. concevoir des algorithmes on-line

3. ...

Plan

Plus précisément

Espace de probabilité

- Un **espace de probabilité** est donné par un triplet $(\Omega, \mathcal{A}, Pr)$, où
  - Ω est un ensemble (de toutes les issues/résultats possibles d'une expérence aléatoire).
  - $\mathcal{A}$ est une **tribu** :
    - $\mathcal{A}$ est une famille de parties de Ω qui contient l'ensemble vide, qui est close par union dénombrable, et qui est close par passage au complémentaire.
    - Les éléments de $\mathcal{A}$ sont appelés des **événements**.
  - $\Pr : \mathcal{A} \to [0, 1]$ est une fonction de probabilité :

Espace de probabilité

- Un **espace de probabilité** est donné par un triplet $(\Omega, \mathcal{A}, Pr)$, où
  - Ω est un ensemble (de toutes les issues/résultats possibles d'une expérence aléatoire).
    - **Lancé de dé :** $\{1, 2, 3, 4, 5, 6\}$
    - **Naissance :** $\{G, F\}$
  - $\mathcal{A}$ est une **tribu** :
    - $\mathcal{A}$ est une famille de parties de Ω qui contient l'ensemble vide, qui est close par union dénombrable, et qui est close par passage au complémentaire.
    - Les éléments de $\mathcal{A}$ sont appelés des **événements**.
  - $\Pr : \mathcal{A} \to [0, 1]$ est une fonction de probabilité :

En résumé

- **Univers Ω :** ensemble de toutes les issues/résultats possibles d'une expérence aléatoire
  - Lancé de dé : $\{1, 2, 3, 4, 5, 6\}$
  - Naissance : $\{G, F\}$

  En informatique, souvent Ω est soit fini, soit dénombrable.

- **Evénement :** sous-ensemble de l'univers
  - Lancé de dé : le résultat impair$= \mathcal{E} = \{1, 3, 5\}$
  - Naissance : est une fille $= \mathcal{E} = \{F\}$

La fonction de probabilité

- $\Pr : \mathcal{A} \to [0, 1]$ est une fonction de probabilité :
  - pour chaque évènement A, $1 \geq Pr(A) \geq 0$
  - $\Pr(\Omega) = 1$,
  - et pour toute suite d'éléments $A_1, A_2, \cdots A_n \in \mathcal{A}$ deux à deux disjoints,

$$\Pr(\bigcup_{i=1}^{n} A_i) = \sum_{i=1}^{n} \Pr(A_i). \tag{1}$$

La fonction de probabilité

- $\Pr : \mathcal{A} \to [0,1]$ est une fonction de probabilité :
  - pour chaque évènement A, $1 \geq Pr(A) \geq 0$
  - $\Pr(\Omega) = 1$,
  - et pour toute suite d'éléments $A_1, A_2, \cdots A_n \in \mathcal{A}$ deux à deux disjoints,

$$\Pr(\bigcup_{i=1}^{n} A_i) = \sum_{i=1}^{n} \Pr(A_i). \quad (1)$$

- **Exemple : pour modéliser le tirage uniforme d'un dé,**
  - $\Omega = \{1, 2, \cdots, 6\}$
  - $\mathcal{A} = \mathcal{P}(\Omega)$
  - $\Pr(\{i\}) = \frac{1}{6}$.
  - $\Pr(\{1,3,5\}) = \Pr(\{1\}) + \Pr(\{3\}) + \Pr(\{5\}) = \frac{1}{2}$.
  - Pour $U \in \mathcal{A}$, $\Pr(U) = \frac{|U|}{6}$.

Plus précisément

Paradoxe de Monty Hall

Lors d'un jeu télévisé, un candidat est placé devant 3 portes dont derrière l'une d'elles se trouve une voiture.

Paradoxe de Monty Hall

Lors d'un jeu télévisé, un candidat est placé devant 3 portes dont derrière l'une d'elles se trouve une voiture.

1. L'emplacement du cadeau a été choisi de façon uniforme.

Paradoxe de Monty Hall

Lors d'un jeu télévisé, un candidat est placé devant 3 portes dont derrière l'une d'elles se trouve une voiture.

1. L'emplacement du cadeau a été choisi de façon uniforme.

2. Ensuite le présentateur systématique ouvre l'une des deux portes autre que celle qui a été choisie et autre que celle qui cache la voiture.

Paradoxe de Monty Hall

Lors d'un jeu télévisé, un candidat est placé devant 3 portes dont derrière l'une d'elles se trouve une voiture.

1. L'emplacement du cadeau a été choisi de façon uniforme.

2. Ensuite le présentateur systématique ouvre l'une des deux portes autre que celle qui a été choisie et autre que celle qui cache la voiture.

3. Le candidat a le choix entre maintenir son premier choix ou le modifier.

Paradoxe de Monty Hall

Lors d'un jeu télévisé, un candidat est placé devant 3 portes dont derrière l'une d'elles se trouve une voiture.

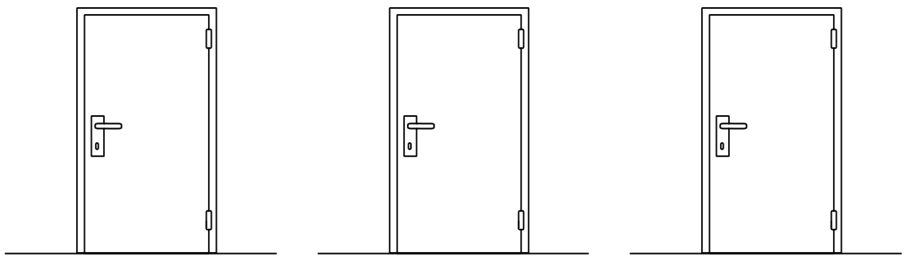
1. L'emplacement du cadeau a été choisi de façon uniforme.

2. Ensuite le présentateur systématique ouvre l'une des deux portes autre que celle qui a été choisie et autre que celle qui cache la voiture.

3. Le candidat a le choix entre maintenir son premier choix ou le modifier.

Que lui conseillez-vous de faire ?

Paradoxe de Monty Hall



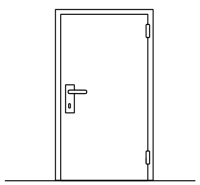
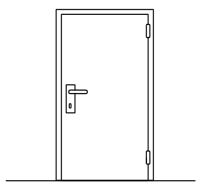
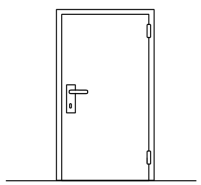
- Lorsque le candidat maintient son choix,

  sa probabilité de gagner est $1/3$.

- Cette probabilité ne dépend pas des actions du présentateur.

Paradoxe de Monty Hall

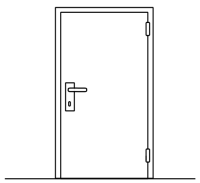
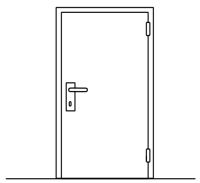
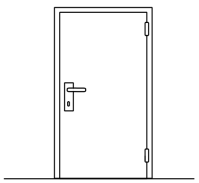
- Le candidat choisit la porte (1).

 (1)	 (2)	 (3)
		
		
		

Paradoxe de Monty Hall

- Le candidat choisit la porte (1).

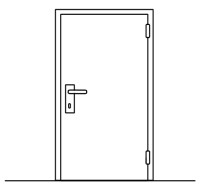
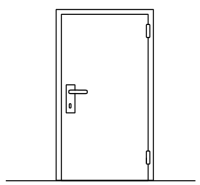
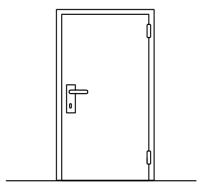
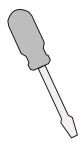
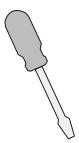
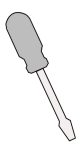
  Le présentateur lui ouvre une porte.

 (1)	 (2)	 (3)
		
		
		

Paradoxe de Monty Hall

- Le candidat choisit la porte (1).

  Le présentateur lui ouvre une porte.

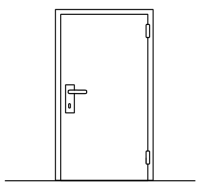
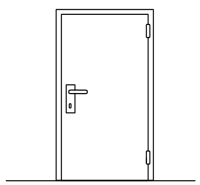
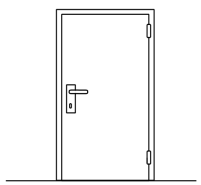
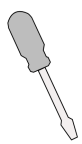
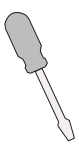
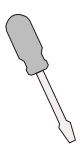
 (1)	 (2)	 (3)
		
		
		

Paradoxe de Monty Hall

- Le candidat choisit la porte (1).

  Le présentateur lui ouvre une porte.

- Lorsque le candidat change de porte,

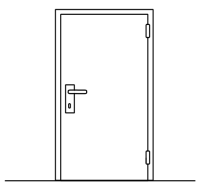
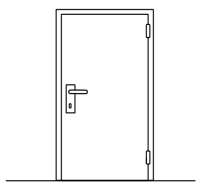
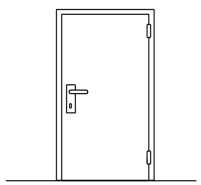
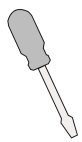
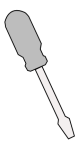
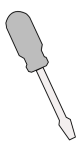
 (1)	 (2)	 (3)
		
		
		

Paradoxe de Monty Hall

- Le candidat choisit la porte (1).

  Le présentateur lui ouvre une porte.

- Lorsque le candidat change de porte,

 (1)	 (2)	 (3)	
			**Le candidat gagne**
			**Le candidat gagne**
			**Le candidat perd**

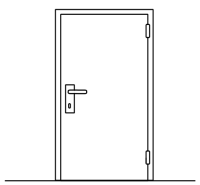
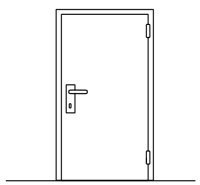
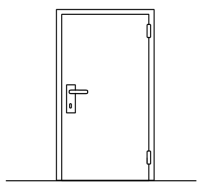
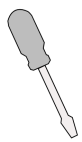
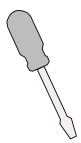
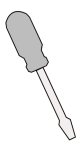
Paradoxe de Monty Hall

- Le candidat choisit la porte (1).

  Le présentateur lui ouvre une porte.

- Lorsque le candidat change de porte,

  La probabilité de gagner est donc 2/3.

 (1)	 (2)	 (3)	
			**Le candidat gagne**
			**Le candidat gagne**
			**Le candidat perd**

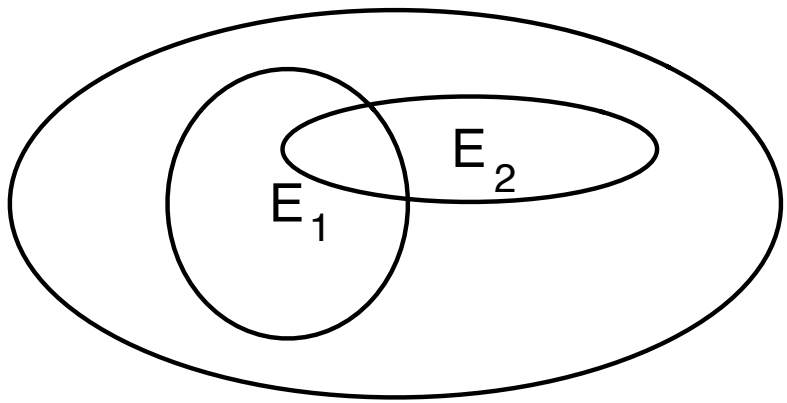
Plus précisément

Union-bound

Proposition [Union bound]

Pour toute suite finie ou dénombrablement infinie d'événements $E_1, E_2, \cdots$

$$\Pr(\bigcup_{i \geq 1} E_i) \leq \sum_{i \geq 1} \Pr(E_i).$$



Indépendance

- Deux événements E_1 et E_2 sont dit **indépendants** si

$$\Pr(E_1 \cap E_2) = \Pr(E_1)\Pr(E_2)$$

- Plus généralement, les événements $E_1, E_2, \cdots E_k$ sont dits **mutuellement indépendants** si et seulement si pour tout $I \subset \{1, 2, \cdots, k\}$,

$$\Pr(\bigcap_{i \in I} E_i) = \prod_{i \in I} \Pr(E_i).$$

Principe de décision différée

Principe

selon une loi uniforme $\mathbf{x} = (x_1, x_2, \cdots, x_n)$ dans $\{0,1\}^n$ est équivalent à choisir chaque x_i de façon indépendante et uniforme dans $\{0,1\}$.

Preuve :

- Dans les deux cas, la probabilité de choisir chacun des 2^n vecteurs possibles est 2^{-n}.

Plan

Variable aléatoire et moyenne

- Une **variable aléatoire** sur un espace de probabilité Ω discret est une fonction $X : \Omega \to \mathbb{R}$.

- Une **variable aléatoire discrète** est une variable aléatoire qui prend un nombre fini ou dénombrable de valeurs.

Variable aléatoire et moyenne

- Une **variable aléatoire** sur un espace de probabilité Ω discret est une fonction $X : \Omega \to \mathbb{R}$.
- Une **variable aléatoire discrète** est une variable aléatoire qui prend un nombre fini ou dénombrable de valeurs.
- La moyenne d'une variable aléatoire discrète X, notée $E[X]$, est définie par

$$E[X] = \sum_i i \Pr(X = i).$$

Quelques lois

- Loi de Bernoulli : $\Pr(X = 1) = p$, $\Pr(X = 0) = 1 - p$.

$$\underbrace{\checkmark}_{\textit{probabilite } p \textit{ d'avoir un succes}}$$

  - la variable aléatoire qui code le résultat d'une épreuve :
    1 pour " succès " ✓avec la probabilité p,
    0 pour " échec" ✗ avec la probabilité $1 - p$

Quelques lois

- Loi de Bernoulli : $\Pr(X = 1) = p$, $\Pr(X = 0) = 1 - p$.

$$\underbrace{\checkmark}_{\text{probabilite } p \text{ d'avoir un succes}}$$

- Loi Géométrique : $\Pr(X = n) = (1 - p)^{n-1} p$.

$$\underbrace{\text{✗ ✗ ✗ ✗ ✗ ✗ } \checkmark}_{n \text{ tirages avant le premier succes}}$$

  - On renouvelle une épreuve de Bernoulli de manière indépendante jusqu'au premier succès. Cette loi correspond à la variable aléatoire donnant le rang du premier succès.

Quelques lois

- Loi de Bernoulli : $\Pr(X = 1) = p$, $\Pr(X = 0) = 1 - p$.

$$\underbrace{\checkmark}_{\text{probabilite } p \text{ d'avoir un succes}}$$

- Loi Géométrique : $\Pr(X = n) = (1 - p)^{n-1} p$.

$$\underbrace{\text{✗ ✗ ✗ ✗ ✗ ✗ } \checkmark}_{n \text{ tirages avant le premier succes}}$$

- Loi Binomiale de paramètres n et p :

$$\Pr(X = j) = C_n^j p^j (1 - p)^{n-j}.$$

$$\underbrace{\text{✗ } \checkmark \text{ ✗ } \checkmark \text{ ✗ ✗ } \checkmark}_{\text{combien de succes parmi } n \text{ tirages}}$$

  - On renouvelle n fois de manière indépendante une épreuve de Bernoulli de paramètre p. Cette loi correspond à la variable aléatoire donnant le nombre de succès obtenus à l'issue des n épreuves.

Linéarité de la moyenne

Theorem (Linéarité de la moyenne)

Pour toute famille finie de variables aléatoires $X_1, X_2, \cdots, X_n$ discrète de moyennes finies

$$E[\sum_{i=1}^{n} X_n] = \sum_{i=1}^{n} E[X_i].$$

Linéarité de la moyenne

Theorem (Linéarité de la moyenne)

Pour toute famille finie de variables aléatoires $X_1, X_2, \cdots, X_n$ discrète de moyennes finies

$$E[\sum_{i=1}^{n} X_n] = \sum_{i=1}^{n} E[X_i].$$

- Remarque importante :
  - aucune hypothèse sur l'indépendance des variables aléatoires.

Quelques lois

Espérance

- Loi de Bernoulli :

$$\Pr(X = 1) = p, \quad \Pr(X = 0) = 1 - p \qquad E[X] = p.$$

- Loi Géométrique :

$$\Pr(X = n) = (1 - p)^{n-1} p. \qquad E[X] = 1/p.$$

- Loi Binomiale :

$$\Pr(X = j) = C_n^j p^j (1 - p)^{n-j}. \qquad E[X] = np.$$

La variance

La **variance** d'une variable aléatoire X est définie par

$$\mathrm{Var}[X] = E[(X - E[X])^2] = E[X^2] - E[X]^2.$$

Quelques lois

Espérance **Variance**

- Loi de Bernoulli :

$\Pr(X = 1) = p$,
$\Pr(X = 0) = 1 - p$ $\quad E[X] = p.$ $\quad \mathrm{Var}[X] = p(1-p).$

- Loi Géométrique :

$\Pr(X = n) = (1-p)^{n-1}p.$ $\quad E[X] = 1/p.$ $\quad \mathrm{Var}[X] = (1-p)/p^2.$

- Loi Binomiale :

$\Pr(X = j) = C_n^j p^j (1-p)^{n-j}.$ $\quad E[X] = np.$ $\quad \mathrm{Var}[X] = np(1-p).$

Plan

Plus précisément

Application : Tri

- Application : Tri Bucket Sort
- Application : Temps moyen du tri rapide

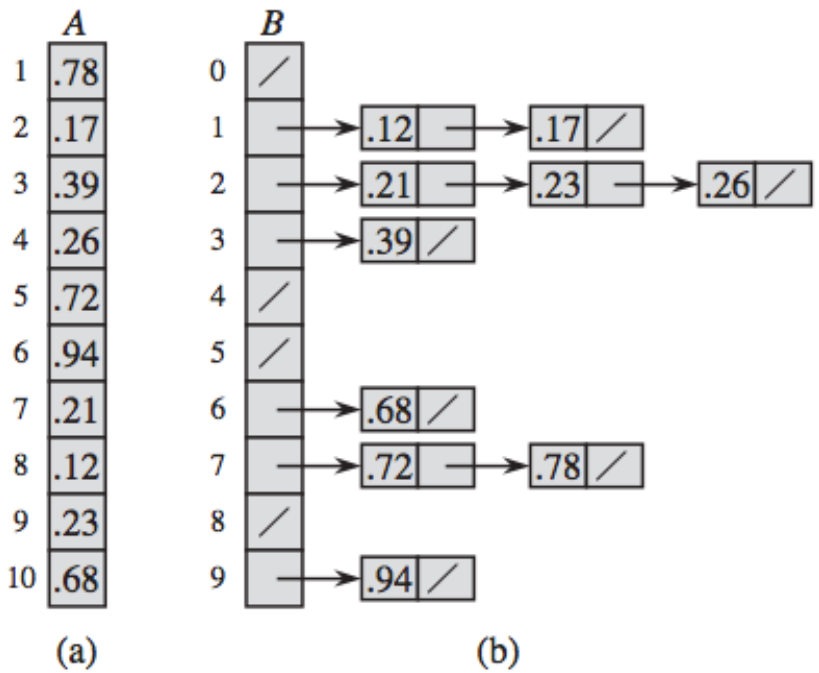
Tri Bucket Sort

Le tri **Bucket Sort** est un exemple de méthode de tri qui brise la borne inférieure de $\Omega(n \log(n))$ opérations par comparaisons :

- supposons que l'on ait n éléments à trier, et que chaque élément est un entier choisit uniformément dans l'intervalle 0 à 1.
- Avec ce tri, on peut trier en temps moyen $\mathcal{O}(n)$.

Tri Bucket Sort

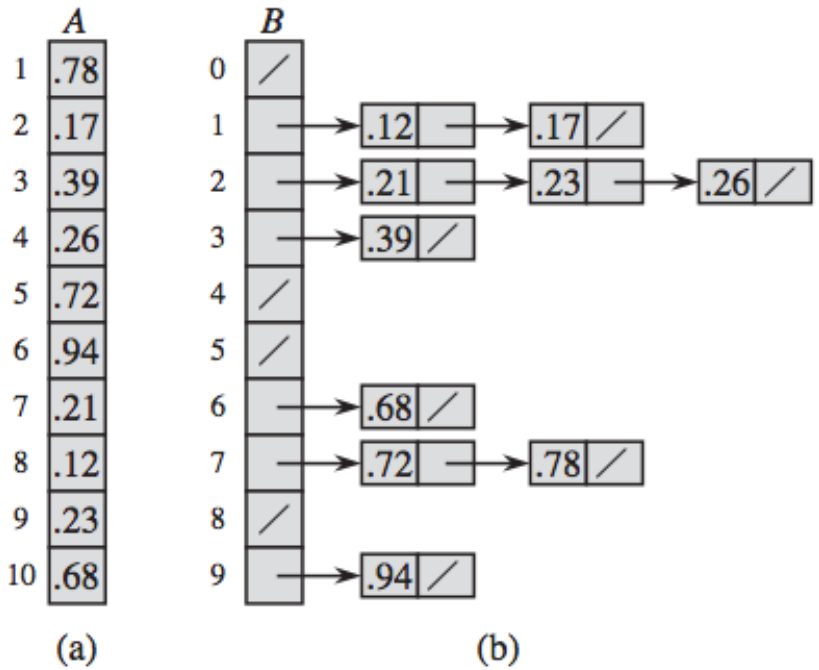
- Dans une première étape,
  - on partage l'intervalle $[0, 1]$ en n intervalles de même longueur.
  - on place les éléments dans n emplacements : l'emplacement j contient tous les éléments dont les m premiers bits correspondent au nombre j.
  - En supposant que placer un élément dans un emplacement se fait en temps $\mathcal{O}(1)$, cette étape nécessite un temps $\mathcal{O}(n)$.



Tri Bucket Sort

- Dans une première étape, $\mathcal{O}(n)$ comparaisons
  - on partage l'intervalle $[0, 1]$ en n intervalles de même longueur.
  - on place les éléments dans n emplacements : l'emplacement j contient tous les éléments dont les m premiers bits correspondent au nombre j.
  - En supposant que placer un élément dans un emplacement se fait en temps $\mathcal{O}(1)$, cette étape nécessite un temps $\mathcal{O}(n)$.
- Dans une deuxième étape,
  - l'algorithme trie chaque emplacement, avec un algorithme de complexité quadratique.
- L'algorithme produit alors en sortie le résultat de la concaténation des listes triées.

Tri Bucket Sort : Illustration



$$\text{Nombre de comparaisons} = \mathcal{O}(n) + \sum_{i=0}^{n-1} \mathcal{O}(X_i^2)$$

avec X_i la variable aléatoire correspondant au nombre d'éléments dans l'emplacement j.

Analyse

- Le nombre d'éléments qui tombent dans chaque emplacement suit une loi binomiale $B(n, 1/n)$. Les emplacements peuvent s'implémenter par des listes par exemple.

Analyse

- Le nombre d'éléments qui tombent dans chaque emplacement suit une loi binomiale $B(n, 1/n)$. Les emplacements peuvent s'implémenter par des listes par exemple.
- X_j : le nombre d'éléments qui tombent dans l'emplacement j.

Analyse

- Le nombre d'éléments qui tombent dans chaque emplacement suit une loi binomiale $B(n, 1/n)$. Les emplacements peuvent s'implémenter par des listes par exemple.
- X_j : le nombre d'éléments qui tombent dans l'emplacement j.
- Le temps pour trier chaque emplacement est de la forme $c(X_j)^2$ pour une constante c.

Analyse

- Le nombre d'éléments qui tombent dans chaque emplacement suit une loi binomiale $B(n, 1/n)$. Les emplacements peuvent s'implémenter par des listes par exemple.
- X_j : le nombre d'éléments qui tombent dans l'emplacement j.
- Le temps pour trier chaque emplacement est de la forme $c(X_j)^2$ pour une constante c.
- Le temps total pour la deuxième phase est donné par

$$E[\sum_{j=1}^{n} c(X_j)^2] = c \sum_{j=1}^{n} E[X_j^2] = cnE[X_1^2],$$

où l'on a utilisé la linéarité de la moyenne et le fait que chaque emplacement joue un rôle symétrique.

Analyse

- Le nombre d'éléments qui tombent dans chaque emplacement suit une loi binomiale $B(n, 1/n)$. Les emplacements peuvent s'implémenter par des listes par exemple.
- X_j : le nombre d'éléments qui tombent dans l'emplacement j.
- Le temps pour trier chaque emplacement est de la forme $c(X_j)^2$ pour une constante c.
- Le temps total pour la deuxième phase est donné par

$$E[\sum_{j=1}^{n} c(X_j)^2] = c \sum_{j=1}^{n} E[X_j^2] = cnE[X_1^2],$$

où l'on a utilisé la linéarité de la moyenne et le fait que chaque emplacement joue un rôle symétrique.

- Puisque chaque X_i est une variable binomiale,

$$E[X_1^2] = 1 + \frac{n(n-1)}{n^2} = 2 - \frac{1}{n} < 2.$$

Analyse : comment calculer $E[X_1^2]$

- X_1 correspond à une variable binomiale $B(n, \frac{1}{n})$.

Analyse : comment calculer $E[X_1^2]$

- X_1 correspond à une variable binomiale $B(n, \frac{1}{n})$.
- Y_i : la variable aléatoire de Bernouilli correspondant au fait que le $i$ème élément soit dans le premier interval.

$$E[Y_i] = \frac{1}{n} = 1 \cdot \frac{1}{n} + 0 \cdot (1 - \frac{1}{n})$$

Analyse : comment calculer $E[X_1^2]$

- X_1 correspond à une variable binomiale $B(n, \frac{1}{n})$.
- Y_i : la variable aléatoire de Bernouilli correspondant au fait que le $i$ème élément soit dans le premier interval.

$$E[Y_i] = \frac{1}{n} = 1 \cdot \frac{1}{n} + 0 \cdot (1 - \frac{1}{n})$$

- $X_1 = \sum_{i=1}^{n} Y_i$

Analyse : comment calculer $E[X_1^2]$

- X_1 correspond à une variable binomiale $B(n, \frac{1}{n})$.
- Y_i : la variable aléatoire de Bernouilli correspondant au fait que le $i$ème élément soit dans le premier interval.

$$E[Y_i] = \frac{1}{n} = 1 \cdot \frac{1}{n} + 0 \cdot (1 - \frac{1}{n})$$

- $X_1 = \sum_{i=1}^{n} Y_i$
- Par linéarité de l'espérance, $E[X_1] = \sum_{i=1}^{n} E[Y_i] = 1$

Analyse : comment calculer $E[X_1^2]$

- X_1 correspond à une variable binomiale $B(n, \frac{1}{n})$.

- Y_i : la variable aléatoire de Bernouilli correspondant au fait que le $i$ème élément soit dans le premier interval.

$$E[Y_i] = \frac{1}{n} = 1 \cdot \frac{1}{n} + 0 \cdot (1 - \frac{1}{n})$$

- $X_1 = \sum_{i=1}^{n} Y_i$
- Par linéarité de l'espérance, $E[X_1] = \sum_{i=1}^{n} E[Y_i] = 1$
- Par définition de la variance, $Var[X_1] = E[X_1^2] - E[X_1]^2$

$$Var[X_1] = np(1-p) = E[X_1^2] - E[X_1]^2$$

$$n\frac{1}{n}(1 - \frac{1}{n}) = E[X_1^2] - 1$$

Analyse : comment calculer $E[X_1^2]$

- X_1 correspond à une variable binomiale $B(n, \frac{1}{n})$.

- Y_i : la variable aléatoire de Bernouilli correspondant au fait que le $i$ème élément soit dans le premier interval.

$$E[Y_i] = \frac{1}{n} = 1 \cdot \frac{1}{n} + 0 \cdot (1 - \frac{1}{n})$$

- $X_1 = \sum_{i=1}^{n} Y_i$
- Par linéarité de l'espérance, $E[X_1] = \sum_{i=1}^{n} E[Y_i] = 1$
- Par définition de la variance, $Var[X_1] = E[X_1^2] - E[X_1]^2$

$$Var[X_1] = np(1-p) = E[X_1^2] - E[X_1]^2$$

$$E[X_1^2] = 2 - \frac{1}{n} < 2.$$

Plus précisément

Application : Tri

- Application : Tri Bucket Sort
- Application : Temps moyen du tri rapide

Algorithme du tri rapide

L'algorithme **Quicksort** est un algorithme de tri récursif, qui consiste, étant donné une liste $S = \{y_1, \cdots, y_n\}$ d'éléments distincts aux opérations suivantes :

- retourner S si S ne possède qu'un ou zéro élément ;
- choisir sinon un élément y de S, appelé **pivot** ;
  - comparer chaque élément à y, pour diviser S en 2 sous-listes :

    S_1, ceux qui sont inférieurs à y,
    S_2 ceux qui sont plus grands.
  - utiliser récursivement Quicksort sur chacune des listes S_1 et S_2 pour les trier ;
  - retourner le résultat S_1 concaténé avec y concaténé avec le résultat S_2.

Analyse

- Dans le pire des cas : par exemple
  - si la liste est dans l'ordre décroissant
  - et si l'on prend comme pivot systématiquement le premier élément,

  l'algorithme nécessite $\mathcal{O}(n^2)$ comparaisons.

- Supposons que
  - dans Quicksort on choisisse le pivot systématiquement selon une loi uniforme
  - et des tirages indépendants parmi les possibilités.

  Alors pour toute entrée, l'algorithme effectue un nombre moyen de comparaisons donné par $2n \log n + \mathcal{O}(n)$.

Analyse

- Dans le pire des cas : par exemple
  - si la liste est dans l'ordre décroissant
  - et si l'on prend comme pivot systématiquement le premier élément,

  l'algorithme nécessite $\mathcal{O}(n^2)$ comparaisons.

- Supposons que
  - dans Quicksort on choisisse le pivot systématiquement selon une loi uniforme
  - et des tirages indépendants parmi les possibilités.

  Alors pour toute entrée, l'algorithme effectue un nombre moyen de comparaisons donné par $2n \log n + \mathcal{O}(n)$.

- Supposons que
  - dans Quicksort on choisisse systématiquement le 1er élément.
  - les entrées sont choisies de façon uniforme parmi les permutations de $\{1, 2, \cdots, n\}$,

  Alors l'algorithme effectue un nombre moyen de comparaisons de l'ordre de $2n \log n + \mathcal{O}(n)$.

Preuve du premier résultat (1/2)

- Soient $y_1, y_2, \cdots, y_n$ les mêmes valeurs que les valeurs en entrée $x_1, x_2, \cdots, x_n$ mais dans l'ordre trié.

- Pour $i < j$, soit X_{ij} la variable aléatoire qui prend la valeur :
$$\begin{cases} 1 & \text{si } x_i \text{ et } x_j \text{ sont comparés par l'algorithme} \\ 0 & \text{sinon} \end{cases}$$

- Le nombre total de comparaisons est donné par
$$X = \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} X_{ij}.$$

- On a donc par linéarité de la moyenne $E[X] = \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} E[X_{ij}]$.

- Puisque X_{ij} prend les valeurs 0 ou 1, $E[X_{ij}]$ est la probabilité que x_i soit comparé à x_j.

Preuve du premier résultat (2/2)

- x_i est comparé à x_j si et seulement si x_i ou x_j est le premier pivot choisit parmi l'ensemble $Y^{ij} = \{x_i, x_{i+1}, \cdots, x_j\}$.
  - Si x_i (ou x_j) est le premier pivot choisit dans cette liste, alors x_i et x_j seront dans la même liste et donc seront comparés.
  - Symétriquement, si aucun des deux n'est le premier pivot choisit dans cette liste, alors x_i et x_j seront séparés dans deux listes distinctes et donc ne seront jamais comparés.
- Comme les pivots sont choisis de façon uniforme et indépendante, la probabilité que cela se produise est $2/(j-i+1)$.
- En posant $k = j - i + 1$, on obtient

$$\begin{aligned} E[X] &= \sum_{i=1}^{n-1} \sum_{j=i+1}^{n} \frac{2}{j-i+1} = \sum_{i=1}^{n-1} \sum_{k=2}^{n-i+1} \frac{2}{k} \\ &= \sum_{k=2}^{n} \sum_{i=1}^{n+1-k} \frac{2}{k} = \sum_{k=2}^{n} (n+1-k)\frac{2}{k} \\ &= (2n+2)\sum_{k=1}^{n} \tfrac{1}{k} - 4n = 2n\log(n) + \Theta(n). \end{aligned}$$

Plan

Principe de décision différée

Principe

selon une loi uniforme $\mathbf{x} = (x_1, x_2, \cdots, x_n)$ dans $\{0,1\}^n$ est équivalent à choisir chaque x_i de façon indépendante et uniforme dans $\{0,1\}$.

- Dans les deux cas, la probabilité de choisir chacun des 2^n vecteurs possibles est 2^{-n}.

$$x_1 \quad x_2 \quad x_3 \quad x_4 \quad \cdots \quad x_{n-1} \quad x_n$$

Variables aléatoires indicatrices

Soit S un univers et A un évènement dans S.
La **variable aléatoire indicatrice** de l'évènement A est :

$$\mathbb{1}_A = \begin{cases} 1 & si \quad A \ se \ ralise \\ 0 & sinon \end{cases}$$

Lemme

Soit A un évènement. Soit $X_A = \mathbb{1}_A$ sa variable aléatoire indicatrice. Alors,

$$E[X_A] = Pr(A)$$

Preuve :

- Soit $\neg A$ le complémentaire de A
-

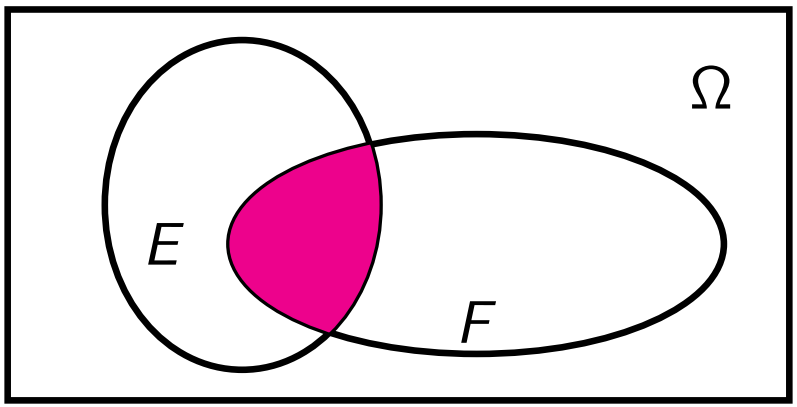
$$\begin{array}{lcll} E[X_A] & = & E[\mathbb{1}_A] & par \ definition \\ E[X_A] & = & 1 * Pr(A) + 0 * Pr(\neg A) & par \ definition \ de \ l' \ esperance \\ E[X_A] & = & Pr(A) & \end{array}$$

Plan

Loi conditionnelle

- Rappelons la notion de **probabilité conditionnelle** :
- La probabilité de E sachant F est définie par

$$\Pr(E|F) = \frac{\Pr(E \cap F)}{\Pr(F)}.$$



Paradoxe des deux enfants

Le roi a deux enfants. $\Omega = \{FF, FG, GF, GG\}$.

Paradoxe des deux enfants

Le roi a deux enfants. $\Omega = \{FF, FG, GF, GG\}$.

Quelle est la probabilité que les 2 enfants soient des garçons ?

Paradoxe des deux enfants

Le roi a deux enfants. $\Omega = \{FF, FG, GF, GG\}$.

Quelle est la probabilité que les 2 enfants soient des garçons ?

- La probabilité vaut $\mathbf{P(GG)} = \dfrac{|A|}{|\Omega|} = \dfrac{1}{4}$

Paradoxe des deux enfants

Le roi a deux enfants. $\Omega = \{FF, FG, GF, GG\}$.

Quelle est la probabilité que les 2 enfants soient des garçons ?

- La probabilité vaut $\mathbf{P(GG)} = \dfrac{|A|}{|\Omega|} = \dfrac{1}{4}$
- sachant que l'ainé est un garçon.

Paradoxe des deux enfants

Le roi a deux enfants. $\Omega = \{FF, FG, GF, GG\}$.

Quelle est la probabilité que les 2 enfants soient des garçons ?

- La probabilité vaut $\mathbf{P(GG)} = \dfrac{|A|}{|\Omega|} = \dfrac{1}{4}$
- sachant que l'ainé est un garçon.
  - les seuls cas possibles sont : GG et GF.

Paradoxe des deux enfants

Le roi a deux enfants. $\Omega = \{FF, FG, GF, GG\}$.

Quelle est la probabilité que les 2 enfants soient des garçons ?

- La probabilité vaut $\mathbf{P(GG)} = \dfrac{|A|}{|\Omega|} = \dfrac{1}{4}$
- sachant que l'ainé est un garçon.
  - les seuls cas possibles sont : GG et GF.
  - La probabilité vaut : 1/2.

Paradoxe des deux enfants

Le roi a deux enfants. $\Omega = \{FF, FG, GF, GG\}$.

Quelle est la probabilité que les 2 enfants soient des garçons ?

- La probabilité vaut $\mathbf{P(GG)} = \dfrac{|A|}{|\Omega|} = \dfrac{1}{4}$
- sachant que l'ainé est un garçon.
  - les seuls cas possibles sont : GG et GF.
  - La probabilité vaut : 1/2.
- sachant qu'au moins l'un des deux est un garçon.

Paradoxe des deux enfants

Le roi a deux enfants. $\Omega = \{FF, FG, GF, GG\}$.

FF FG

GF GG

Quelle est la probabilité que les 2 enfants soient des garçons ?

- La probabilité vaut $\mathbf{P(GG)} = \dfrac{|A|}{|\Omega|} = \dfrac{1}{4}$
- sachant que l'ainé est un garçon.
  - les seuls cas possibles sont : GG et GF.
  - La probabilité vaut : 1/2.
- sachant qu'au moins l'un des deux est un garçon.

Paradoxe des deux enfants

Le roi a deux enfants. $\Omega = \{FF, FG, GF, GG\}$.

FF FG

GF $\textcircled{GG}$

Quelle est la probabilité que les 2 enfants soient des garçons ?

- La probabilité vaut $\mathbf{P(GG)} = \dfrac{|A|}{|\Omega|} = \dfrac{1}{4}$
- sachant que l'ainé est un garçon.
  - les seuls cas possibles sont : GG et GF.
  - La probabilité vaut : 1/2.
- sachant qu'au moins l'un des deux est un garçon.

Paradoxe des deux enfants

Le roi a deux enfants. $\Omega = \{FF, FG, GF, GG\}$.

Quelle est la probabilité que les 2 enfants soient des garçons ?

- La probabilité vaut $\mathbf{P(GG)} = \dfrac{|A|}{|\Omega|} = \dfrac{1}{4}$
- sachant que l'ainé est un garçon.
  - les seuls cas possibles sont : GG et GF.
  - La probabilité vaut : 1/2.
- sachant qu'au moins l'un des deux est un garçon.

Paradoxe des deux enfants

Le roi a deux enfants. $\Omega = \{FF, FG, GF, GG\}$.

FF FG

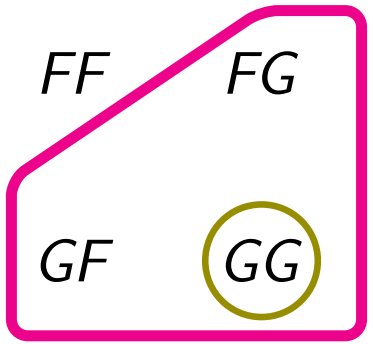
GF $\textcircled{GG}$

Quelle est la probabilité que les 2 enfants soient des garçons ?

- La probabilité vaut $\mathbf{P(GG)} = \dfrac{|A|}{|\Omega|} = \dfrac{1}{4}$
- sachant que l'ainé est un garçon.
  - les seuls cas possibles sont : GG et GF.
  - La probabilité vaut : 1/2.
- sachant qu'au moins l'un des deux est un garçon.

Paradoxe des deux enfants

Le roi a deux enfants. $\Omega = \{FF, FG, GF, GG\}$.

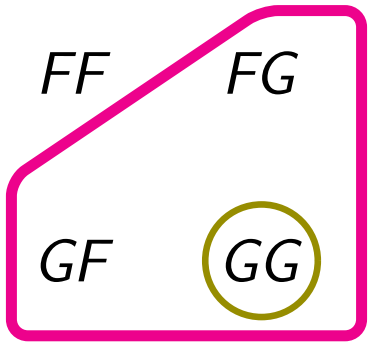


Quelle est la probabilité que les 2 enfants soient des garçons ?

- La probabilité vaut $\mathbf{P(GG)} = \dfrac{|A|}{|\Omega|} = \dfrac{1}{4}$
- sachant que l'ainé est un garçon.
  - les seuls cas possibles sont : GG et GF.
  - La probabilité vaut : 1/2.
- sachant qu'au moins l'un des deux est un garçon.
  - les seuls cas possibles sont : FG, GF et GG.

Paradoxe des deux enfants

Le roi a deux enfants. $\Omega = \{FF, FG, GF, GG\}$.



Quelle est la probabilité que les 2 enfants soient des garçons ?

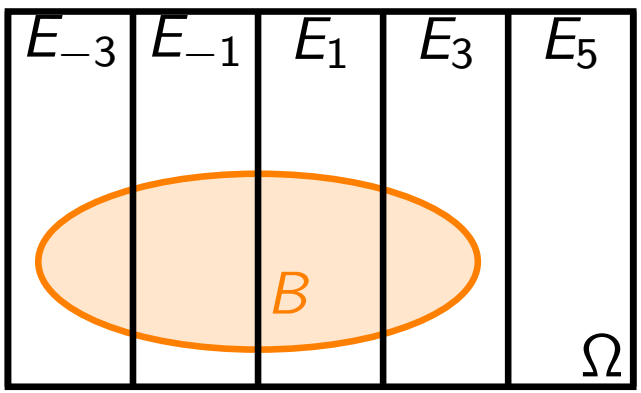
- La probabilité vaut $\mathbf{P(GG)} = \dfrac{|A|}{|\Omega|} = \dfrac{1}{4}$
- sachant que l'ainé est un garçon.
  - les seuls cas possibles sont : GG et GF.
  - La probabilité vaut : 1/2.
- sachant qu'au moins l'un des deux est un garçon.
  - les seuls cas possibles sont : FG, GF et GG.
  - La probabilité vaut : 1/3.

Law of Total Probability

Proposition [Law of Total Probability]

Pour toute suite finie d'événements $E_1, E_2, \cdots E_n$ deux-à-deux disjoints tel que $\cup_{i=1}^{n} E_i = \Omega$. Alors,

$$\Pr(B) = \sum_{i=1}^{n} \Pr(B \cap E_i) = \sum_{i=1}^{n} \Pr(B|E_i)\Pr(E_i)$$



Plan

Vérification d'identités

- Supposons que l'on se donne trois matrices $\mathbf{A}, \mathbf{B}$ et $\mathbf{C}$ de taille $n \times n$.
- Supposons que l'on veuille vérifier si $\mathbf{AB} = \mathbf{C}$.

Vérification d'identités

- Supposons que l'on se donne trois matrices $\mathbf{A}, \mathbf{B}$ et $\mathbf{C}$ de taille $n \times n$.
- Supposons que l'on veuille vérifier si $\mathbf{AB} = \mathbf{C}$.
- Pour simplifier la discussion, nous supposerons que ces matrices sont à coefficients dans $\mathbb{Z}_2$.

Vérification d'identités

- Supposons que l'on se donne trois matrices $\mathbf{A}, \mathbf{B}$ et $\mathbf{C}$ de taille $n \times n$.
- Supposons que l'on veuille vérifier si $\mathbf{AB} = \mathbf{C}$.
- Pour simplifier la discussion, nous supposerons que ces matrices sont à coefficients dans $\mathbb{Z}_2$.
- Une façon de faire consiste à multiplier $\mathbf{A}$ par $\mathbf{B}$, puis à tester l'égalité avec $\mathbf{C}$.
  - L'algorithme de multiplication le plus simple nécessite $\mathcal{O}(n^3)$ opérations.
  - Il existe des algorithmes plus efficaces en $\mathcal{O}(n^{2.37})$ opérations.

Algorithme randomisé

- Algorithme randomisé :
  - On choisit un vecteur aléatoire $\mathbf{x} \in \{0,1\}^n$.
  - On calcule alors $\mathbf{ABx}$ en calculant $\mathbf{Bx}$ puis $\mathbf{A}(\mathbf{Bx})$.
  - On calcule ensuite $\mathbf{Cx}$.
  - Si $\mathbf{ABx} \neq \mathbf{Cx}$ on répond que $\mathbf{AB} \neq \mathbf{C}$.
  - Sinon, on répond que $\mathbf{AB} = \mathbf{C}$.

Algorithme randomisé

- Algorithme randomisé :
  - On choisit un vecteur aléatoire $\mathbf{x} \in \{0,1\}^n$.
  - On calcule alors **ABx** en calculant **Bx** puis **A**(**Bx**).
  - On calcule ensuite **Cx**.
  - Si **ABx** $\neq$ **Cx** on répond que **AB** $\neq$ **C**.
  - Sinon, on répond que **AB** $=$ **C**.
- On a besoin uniquement de trois multiplications matrice/vecteur :
  - temps $\mathcal{O}(n^2)$ par l'algorithme évident.

Probabilité d'erreur

Lemma

Si $\mathbf{AB} \neq \mathbf{C}$, *et si* $\mathbf{x}$ *est choisi uniformément dans* $\{0,1\}^n$ *alors*

$$\Pr(\mathbf{ABx} = \mathbf{Cx}) \leq \frac{1}{2}.$$

Preuve (1/2)

- Soit $\mathbf{D} = \mathbf{AB} - \mathbf{C}$.
- $\mathbf{ABx} \neq \mathbf{Cx}$ implique que $\mathbf{Dx} \neq 0$, et puisque $\mathbf{D}$ n'est pas la matrice nulle, $\mathbf{D}$ doit avoir au moins un coefficient non nul.

Preuve (1/2)

- Soit $\mathbf{D} = \mathbf{AB} - \mathbf{C}$.
- $\mathbf{ABx} \neq \mathbf{Cx}$ implique que $\mathbf{Dx} \neq 0$, et puisque $\mathbf{D}$ n'est pas la matrice nulle, $\mathbf{D}$ doit avoir au moins un coefficient non nul.
- Supposons sans perte de généralité que cela soit $d_{1,1}$.

Preuve (1/2)

- Soit $\mathbf{D} = \mathbf{AB} - \mathbf{C}$.
- $\mathbf{ABx} \neq \mathbf{Cx}$ implique que $\mathbf{Dx} \neq 0$, et puisque $\mathbf{D}$ n'est pas la matrice nulle, $\mathbf{D}$ doit avoir au moins un coefficient non nul.

- Supposons sans perte de généralité que cela soit $d_{1,1}$.
- Pour que $\mathbf{Dx} = 0$, on doit avoir $\sum_{j=1}^{n} d_{1,j} x_j = 0$, et donc

$$x_1 = -(\sum_{j=2}^{N} d_{1,j} x_j)/d_{1,1}. \tag{2}$$

Preuve (1/2)

- Soit $\mathbf{D} = \mathbf{AB} - \mathbf{C}$.
- $\mathbf{ABx} \neq \mathbf{Cx}$ implique que $\mathbf{Dx} \neq 0$, et puisque $\mathbf{D}$ n'est pas la matrice nulle, $\mathbf{D}$ doit avoir au moins un coefficient non nul.

- Supposons sans perte de généralité que cela soit $d_{1,1}$.
- Pour que $\mathbf{Dx} = 0$, on doit avoir $\sum_{j=1}^{n} d_{1,j} x_j = 0$, et donc

$$x_1 = -(\sum_{j=2}^{N} d_{1,j} x_j)/d_{1,1}. \tag{2}$$

- Selon le **principe de la décision différée**, on peut voir choisir $\mathbf{x} \in \{0,1\}^n$ comme choisir chacune de ses composantes.

Preuve (1/2)

- Soit $\mathbf{D} = \mathbf{AB} - \mathbf{C}$.
- $\mathbf{ABx} \neq \mathbf{Cx}$ implique que $\mathbf{Dx} \neq 0$, et puisque $\mathbf{D}$ n'est pas la matrice nulle, $\mathbf{D}$ doit avoir au moins un coefficient non nul.

- Supposons sans perte de généralité que cela soit $d_{1,1}$.
- Pour que $\mathbf{Dx} = 0$, on doit avoir $\sum_{j=1}^{n} d_{1,j} x_j = 0$, et donc

$$x_1 = -(\sum_{j=2}^{N} d_{1,j} x_j)/d_{1,1}. \tag{2}$$

- Selon le **principe de la décision différée**, on peut voir choisir $\mathbf{x} \in \{0,1\}^n$ comme choisir chacune de ses composantes.

- Considérons la situation juste avant que x_1 soit choisi :
  - à ce moment là le membre droit est fixé, et il y a au plus une possibilité pour x_1 qui rend l'égalité vraie.

Preuve (1/2)

- Soit $\mathbf{D} = \mathbf{AB} - \mathbf{C}$.
- $\mathbf{ABx} \neq \mathbf{Cx}$ implique que $\mathbf{Dx} \neq 0$, et puisque $\mathbf{D}$ n'est pas la matrice nulle, $\mathbf{D}$ doit avoir au moins un coefficient non nul.

- Supposons sans perte de généralité que cela soit $d_{1,1}$.
- Pour que $\mathbf{Dx} = 0$, on doit avoir $\sum_{j=1}^{n} d_{1,j} x_j = 0$, et donc

$$x_1 = -(\sum_{j=2}^{N} d_{1,j} x_j)/d_{1,1}. \tag{2}$$

- Selon le **principe de la décision différée**, on peut voir choisir $\mathbf{x} \in \{0,1\}^n$ comme choisir chacune de ses composantes.

- Considérons la situation juste avant que x_1 soit choisi :
  - à ce moment là le membre droit est fixé, et il y a au plus une possibilité pour x_1 qui rend l'égalité vraie.
  - Puisqu'il y a deux choix pour x_1, cela se produit avec probabilité $\leq \frac{1}{2}$.

Preuve (2/2)

- Calculons $\Pr(\mathbf{AB}r = \mathbf{C}r) = \Pr(\mathbf{D}r = 0)$

$$
\begin{aligned}
\Pr(\mathbf{AB}r = \mathbf{C}r) &= \sum_{(x_2,\ldots,x_n)\in\{0,1\}^{n-1}} \Pr((\mathbf{AB}r = \mathbf{AB}r) \cap ((x_2,\ldots,x_n) = (r_2,\ldots,r_n)) \\
&\leq \sum_{(x_2,\ldots,x_n)\in\{0,1\}^{n-1}} \Pr((r_1 = -(\sum_{j=2}^{N} d_{1,j} r_j)/d_{1,1}) \cap ((x_2,\ldots,x_n) = (r_2,\ldots,r_n))) \\
&\leq \sum_{(x_2,\ldots,x_n)\in\{0,1\}^{n-1}} \Pr(r_1 = -(\sum_{j=2}^{N} d_{1,j} r_j)/d_{1,1}) \cdot \Pr((x_2,\ldots,x_n) = (r_2,\ldots,r_n)) \\
&\leq \sum_{(x_2,\ldots,x_n)\in\{0,1\}^{n-1}} \frac{1}{2} \cdot \Pr((x_2,\ldots,x_n) = (r_2,\ldots,r_n)) \\
&\leq \tfrac{1}{2}
\end{aligned}
$$

Plan

Algorithme de Monte Carlo.

- On appelle un tel algorithme un algorithme de **Monte Carlo** : l'algorithme est efficace, mais peut parfois se tromper.

Algorithme de Monte Carlo.

- On appelle un tel algorithme un algorithme de **Monte Carlo** : l'algorithme est efficace, mais peut parfois se tromper.
- Il s'agit d'un algorithme à erreur unilatérale :

Algorithme de Monte Carlo.

- On appelle un tel algorithme un algorithme de **Monte Carlo** : l'algorithme est efficace, mais peut parfois se tromper.
- Il s'agit d'un algorithme à erreur unilatérale :
  - Si l'on trouve un $\mathbf{x}$ tel que $\mathbf{ABx} \neq \mathbf{Cx}$, on est certain de la réponse : $\mathbf{AB} \neq \mathbf{C}$.

Algorithme de Monte Carlo.

- On appelle un tel algorithme un algorithme de **Monte Carlo** : l'algorithme est efficace, mais peut parfois se tromper.

- Il s'agit d'un algorithme à erreur unilatérale :

  - Si l'on trouve un $\mathbf{x}$ tel que $\mathbf{ABx} \neq \mathbf{Cx}$, on est certain de la réponse : $\mathbf{AB} \neq \mathbf{C}$.
  - Sinon, il y a une possibilité de se tromper.

Algorithme de Monte Carlo.

- On appelle un tel algorithme un algorithme de **Monte Carlo** : l'algorithme est efficace, mais peut parfois se tromper.
- Il s'agit d'un algorithme à erreur unilatérale :
  - Si l'on trouve un **x** tel que **ABx** $\neq$ **Cx**, on est certain de la réponse : **AB** $\neq$ **C**.
  - Sinon, il y a une possibilité de se tromper.
- La probabilité d'erreur est au pire cas de $1/2$.

Algorithme de Monte Carlo.

- On appelle un tel algorithme un algorithme de **Monte Carlo** : l'algorithme est efficace, mais peut parfois se tromper.
- Il s'agit d'un algorithme à erreur unilatérale :
  - Si l'on trouve un **x** tel que $\mathbf{ABx} \neq \mathbf{Cx}$, on est certain de la réponse : $\mathbf{AB} \neq \mathbf{C}$.
  - Sinon, il y a une possibilité de se tromper.
- La probabilité d'erreur est au pire cas de $1/2$.
- Cependant, on peut la rendre assez facilement aussi petite que l'on veut.

Réduction de l'erreur

- Supposons que l'on recommence alors k fois le test,

Réduction de l'erreur

- Supposons que l'on recommence alors k fois le test,
  - tant que **l'on n'est pas certain de la réponse**, en testant à chaque fois un nouveau $\mathbf{x}$ tiré uniformément dans $\{0,1\}^n$.

Réduction de l'erreur

- Supposons que l'on recommence alors k fois le test,
  - tant que **l'on n'est pas certain de la réponse**, en testant à chaque fois un nouveau $\mathbf{x}$ tiré uniformément dans $\{0,1\}^n$.

- La probabilité que l'on se trompe à l'issu de k tests successifs est en 2^{-k}.

Réduction de l'erreur

- Supposons que l'on recommence alors k fois le test,
  - tant que **l'on n'est pas certain de la réponse**, en testant à chaque fois un nouveau $\mathbf{x}$ tiré uniformément dans $\{0,1\}^n$.

- La probabilité que l'on se trompe à l'issu de k tests successifs est en 2^{-k}.

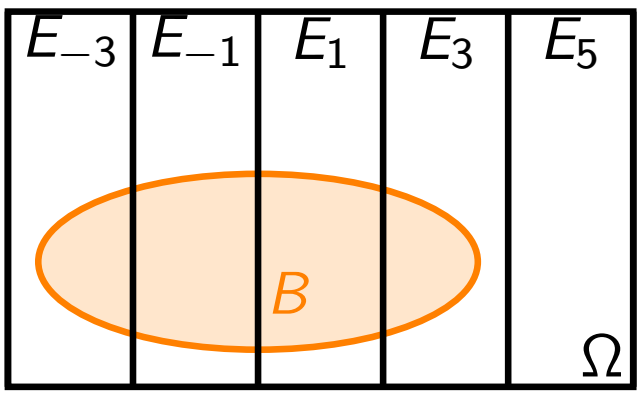
- Si l'on prend $k = 100$, cela donne une probabilité d'erreur 2^{-100} avec un temps $\mathcal{O}(kn^2) = \mathcal{O}(100n^2) = \mathcal{O}(n^2)$.

la loi de Bayes

Définition [loi de Bayes]

Pour toute suite finie d'événements $E_1, E_2, \cdots E_n$ deux-à-deux disjoints tel que $\cup_{i=1}^{n} E_i = \Omega$. Alors,

$$\Pr(E_j|B) = \frac{\Pr(E_j \cap B)}{\Pr(B)} = \frac{\Pr(B|E_j)\Pr(E_j)}{\sum_{i=1}^{n} \Pr(B|E_i)\Pr(E_i)}$$



Récapitulatif

- E est l'évènement que $\mathbf{AB} = \mathbf{C}$ soit vrai.

  Nous noterons le complément de E par E^c.
- B est l'évènement que l'algorithme retourne vrai.

On a :

$$\Pr(B|E) = 1 \text{ et } \Pr(B|E^c) \leq 1/2$$

1. Supposons que
   - $\Pr(E) = \Pr(E^c) = 1/2$.
   - Après la première itération de l'algorithme, il retourne *vrai*.
2. Après la première itération de l'algorithme, il retourne *vrai*.
3. Calculons $\Pr(E|B)$

$$\Pr(E|B) = \frac{\Pr(B|E)\Pr(E)}{\Pr(B|E)\Pr(E) + \Pr(B|E^c)\Pr(E^c)} \geq \frac{\frac{1}{2}}{\frac{1}{2} + \frac{1}{2}\frac{1}{2}} \geq \frac{2}{3}$$

Application du principe de Bayes

- Maintenant, notre croyance est la suivante :

$$\Pr(E) \geq 2/3 \text{ et } \Pr(E^c) \leq 1/3.$$

- Après la **deuxième** itération de l'algorithme, il retourne *vrai*.
- Calculons $\Pr(E|B)$

$$\Pr(E|B) = \frac{\Pr(B|E)\Pr(E)}{\Pr(B|E)\Pr(E) + \Pr(B|E^c)\Pr(E^c)} \geq \frac{\frac{2}{3}}{\frac{2}{3} + \frac{1}{3}\frac{1}{2}} \geq \frac{4}{5}$$

- Maintenant, notre croyance est la suivante :

  Avant la **(i)ème** itération de l'algorithme, $\Pr(E) \geq \frac{2^{i-1}}{2^{i-1}+1}$
- Si après cette itération de l'algorithme il retourne vrai, alors

$$\Pr(E) \geq \frac{2^i}{2^i + 1}$$

En conclusion

- Ainsi, si les 100 exécutions de cet algorithme retournent *vrai* alors la confiance que la vérification soit correcte est au moins de $1 - \frac{1}{2^{100}+1}$.

Plan

Problème du collectionneur

- Supposons que des boites de céréales contiennent chacune un coupon parmi n coupons possibles.
- Supposons que l'on veuille posséder au moins un exemplaire de chacun des coupons.
- Combien faut il acheter en moyenne de boites de céréales pour cela ?

Problème du collectionneur

- Supposons que des boites de céréales contiennent chacune un coupon parmi n coupons possibles.
- Supposons que l'on veuille posséder au moins un exemplaire de chacun des coupons.
- Combien faut il acheter en moyenne de boites de céréales pour cela ?

Ce problème apparaît dans de nombreux contextes en informatique.

Plus précisément

Application : Problème du collectionneur

- Calcul du nombre moyen de boîtes à acheter
- inégalité de Markov

Problème du collectionneur : méthode

- Soit X le nombre de boîtes achetées avant de posséder tous les n coupons. **On s'intéresse donc à $E[X]$.**

Problème du collectionneur : méthode

- Soit X le nombre de boîtes achetées avant de posséder tous les n coupons. **On s'intéresse donc à $E[X]$.**
- X_i désigne le nombre de boîtes achetées en ayant exactement $i-1$ coupons pour posséder un coupon supplémentaire, alors

$$X = \sum_{i=1}^{n} X_i.$$

Problème du collectionneur : méthode

- Soit X le nombre de boîtes achetées avant de posséder tous les n coupons. **On s'intéresse donc à $E[X]$.**
- X_i désigne le nombre de boîtes achetées en ayant exactement $i-1$ coupons pour posséder un coupon supplémentaire, alors

$$X = \sum_{i=1}^{n} X_i.$$

- Lorsque l'on a exactement $i-1$ coupons, la probabilité d'obtenir un nouveau coupon en achetant une boite est

$$p_i = 1 - \frac{i-1}{n}.$$

Problème du collectionneur : méthode

- Soit X le nombre de boîtes achetées avant de posséder tous les n coupons. **On s'intéresse donc à $E[X]$.**
- X_i désigne le nombre de boîtes achetées en ayant exactement $i-1$ coupons pour posséder un coupon supplémentaire, alors

$$X = \sum_{i=1}^{n} X_i.$$

- Lorsque l'on a exactement $i-1$ coupons, la probabilité d'obtenir un nouveau coupon en achetant une boite est

$$p_i = 1 - \frac{i-1}{n}.$$

- X_i est une variable aléatoire géométrique de paramètre p_i et

$$E[X_i] = \frac{1}{p_i} = \frac{n}{n-i+1}.$$

Problème du collectionneur : méthode

- Soit X le nombre de boîtes achetées avant de posséder tous les n coupons. **On s'intéresse donc à $E[X]$.**
- X_i désigne le nombre de boîtes achetées en ayant exactement $i-1$ coupons pour posséder un coupon supplémentaire, alors

$$X = \sum_{i=1}^{n} X_i.$$

- Lorsque l'on a exactement $i-1$ coupons, la probabilité d'obtenir un nouveau coupon en achetant une boite est

$$p_i = 1 - \frac{i-1}{n}.$$

- X_i est une variable aléatoire géométrique de paramètre p_i et

$$E[X_i] = \frac{1}{p_i} = \frac{n}{n-i+1}.$$

- En utilisant la linéarité de la moyenne, on obtient

$$E[X] = E[\sum_{i=1}^{n} X_i] = \sum_{i=1}^{n} E[X_i] = \sum_{i=1}^{n} \frac{n}{n-i+1} = n \sum_{i=1}^{n} \frac{1}{i}.$$

Problème du collectionneur : réponse

- Le résultat suivant est connu.
  - Le nombre $H(n) = \sum_{i=1}^{n} \frac{1}{i}$, connu sous le nom du $n$ème **nombre harmonique**, vérifie $H(n) = \log(n) + \Theta(1)$.
- Par conséquent, la réponse au problème du collectionneur de coupon est

$$E[X] = n \log(n) + \Theta(n).$$

Plus précisément

Application : Problème du collectionneur

- Calcul du nombre moyen de boîtes à acheter
- inégalité de Markov

Theorem (Inégalité de Markov)

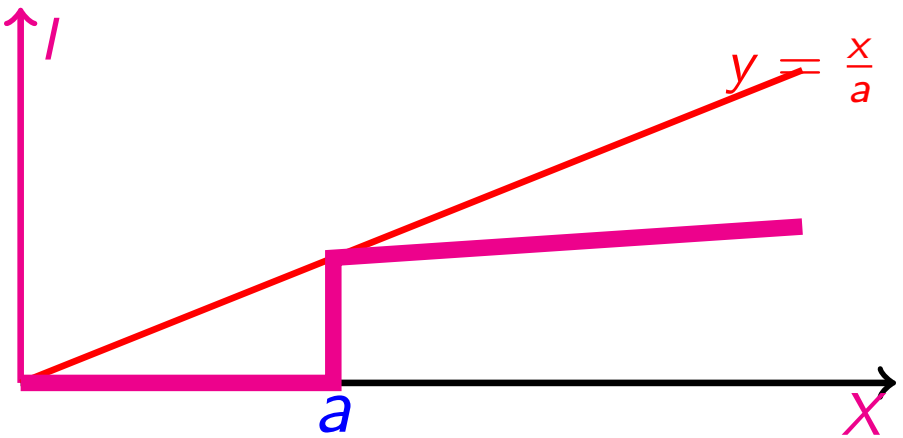
Soit X une variable aléatoire à valeurs positive ou nulles. Alors pour tout $a > 0$,

$$\Pr(X \geq a) \leq \frac{E[X]}{a}.$$

preuve

- Pour $a > 0$, posons $I = \begin{cases} 1 & \text{si } X \geq a, \\ 0 & \text{sinon.} \end{cases}$
- Puisque $X \geq 0$,

$$I \leq \frac{X}{a}. \tag{3}$$



- Puisque I est une variable à valeur dans $\{0, 1\}$,

$$E[I] = \Pr(I = 1) = \Pr(X \geq a)$$

- En passant à la moyenne dans (3), on obtient

$$E[I] \leq E\left[\frac{X}{a}\right] \Rightarrow \Pr(X \geq a) \leq \frac{E[X]}{a}.$$

Remarques

- Observons que l'on a égalité par exemple pour une loi telle que $\Pr(X = a) = 1$.

- Une façon qui peut être plus intuitive de comprendre l'inégalité est d'écrire

$$\Pr(X \geq \mu a) \leq \frac{1}{a},$$

pour tout $a > 0$, où $\mu = E[X]$.

Application de l'inégalité de Markov

- Rappelons $E[X] = n\log(n) + \Theta(n)$.
- Rappelons l'inégalité de Markov.

$$\Pr(X \geq a) \leq \frac{E[X]}{a} \text{ avec } a > 0.$$

- L'inégalité de Markov donne donc

$$\Pr(X \geq 2nH_n) \leq \frac{1}{2}.$$

en posant $a = 2nH_n$, c'est à dire $a = 2E[X]$

Plan

Récapitulatif : tirage de pièces

- Considérons une suite de n tirages **indépendants** de pièces.
- Soit la **variable aléatoire** $X_i = \begin{cases} 1 & \text{si le ième res. est pile} \\ 0 & \text{sinon.} \end{cases}$

$$X_i = \mathbb{1}_{\text{le ième résultat est pile}}$$

implique que

$$E[X_i] = Pr(\text{le ième résultat est pile}) = \frac{1}{2}$$

Récapitulatif : tirage de pièces

- Considérons une suite de n tirages **indépendants** de pièces.
- Soit la **variable aléatoire** $X_i = \begin{cases} 1 & \text{si le ième res. est pile} \\ 0 & \text{sinon.} \end{cases}$

$$E[X_i] = Pr(\text{le ième résultat est pile}) = \frac{1}{2}$$

- Notons par $X = \sum_{i=1}^{n} X_i$ le nombre de piles parmi les n tirages.
- On a $E[X] = \sum_{i=1}^{n} E[X_i]$ par la linéarité de l'espérance.
- On a $E[X] = \sum_{i=1}^{n} E[X_i] = \frac{n}{2}$.
- L'inégalité de Markov donne, pour $\lambda > 0$,
  $\Pr(X \geq \lambda n) \leq \frac{E[X]}{\lambda n} = \frac{n}{2\lambda n} = \frac{1}{2\lambda}$ ou encore $\Pr(X \geq \lambda \frac{n}{2}) \leq \frac{1}{\lambda}$.