

# cours 2

## Modèle probabiliste



Master SID — Raisonnement dans l'incertain

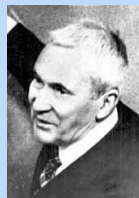
## *Définition des probabilités (Kolmogorov)*

- ▶  $\Omega$  = ensemble fini ou dénombrable  
d'événements élémentaires  $e_k, k \in K \subseteq \mathbb{N}$



## *Définition des probabilités (Kolmogorov)*

- ▶  $\Omega$  = ensemble fini ou dénombrable d'événements élémentaires  $e_k, k \in K \subseteq \mathbb{N}$
- ▶  $\mathcal{A} = 2^\Omega$  = ensemble des événements



## *Définition des probabilités (Kolmogorov)*

- ▶  $\Omega$  = ensemble fini ou dénombrable d'événements élémentaires  $e_k, k \in K \subseteq \mathbb{N}$
- ▶  $\mathcal{A} = 2^\Omega$  = ensemble des événements
- ▶ pour tout  $A \in \mathcal{A}, 0 \leq P(A) \leq 1$
- ▶  $P(\Omega) = 1$



## Définition des probabilités (Kolmogorov)

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



$\implies$  Les probabilités des événements élémentaires déterminent entièrement  $P$

## Définition des probabilités (Kolmogorov)

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



$\implies$  Les probabilités des événements élémentaires déterminent entièrement  $P$

- ▶ **Exemple** : Jet de dé à 6 faces :  $e_k = \ll \text{le dé atterrit sur la face } k \gg$

## *Variable aléatoire*

Variable dont la valeur est déterminée après la réalisation d'une expérience aléatoire.

► **Exemple** :  $D = \ll \text{résultat d'un jet de dé} \gg$

Domaine de  $D = \Omega_D = \{1, \dots, 6\}$

$$P(D = i) = \frac{1}{6} \quad \forall i \in \{1, \dots, 6\}$$

$P(A)$  = fonction( $P(A,B)$ ) ?



$$P(A) = \text{fonction}(P(A,B)) ?$$

► **Exemple :**

	Hommes	Femmes
Petits	20	35
Grands	40	5

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

► **Exemple :**

	Hommes	Femmes
Petits	20	35
Grands	40	5

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

► **Exemple :**

	Hommes	Femmes
Petits	20	35
Grands	40	5

$$P(A|B) = ?$$

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

► **Exemple :**

	Hommes	Femmes
Petits	20	35
Grands	40	5

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

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

► **Exemple :**

	Hommes	Femmes
Petits	20	35
Grands	40	5

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



Interprétation :  $P(A|B)$  = la proba de  $A$  sachant la valeur de  $B$   
 $\implies$  à droite du signe de conditionnement : connaissances

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

► **Exemple :**

	Hommes	Femmes
Petits	20	35
Grands	40	5

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



Interprétation :  $P(A|B)$  = la proba de  $A$  sachant la valeur de  $B$   
 $\implies$  à droite du signe de conditionnement : connaissances

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

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

► **Exemple :**

	Hommes	Femmes
Petits	20	35
Grands	40	5

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



Interprétation :  $P(A|B)$  = la proba de  $A$  sachant la valeur de  $B$   
 $\implies$  à droite du signe de conditionnement : connaissances

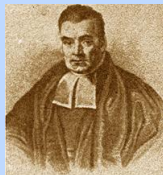
$$P(A, B) = P(A|B) \times P(B)$$

$$\sum_{a \in \Omega_A} P(A = a) = 1$$

## *Théorème de Bayes*

Pour tout  $A, B$  tels que  $P(A) \neq 0$  et  $P(B) \neq 0$  :

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

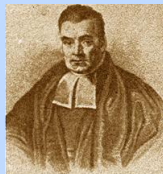




## *Théorème de Bayes*

Pour tout  $A, B$  tels que  $P(A) \neq 0$  et  $P(B) \neq 0$  :

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



- ▶ Permet d'invertir causes et conséquences

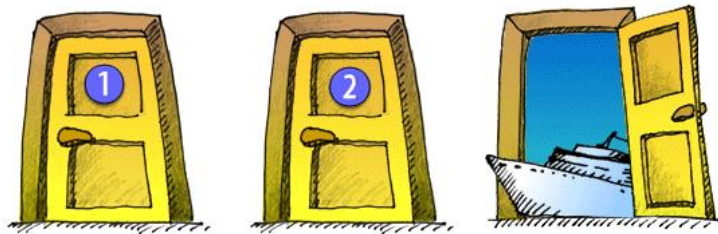
# Monty Hall

<http://www.apprendre-en-ligne.net/random/monty/>



# Monty Hall

<http://www.apprendre-en-ligne.net/random/monty/>



- ▶ derrière une des portes = 1 bateau

# Monty Hall

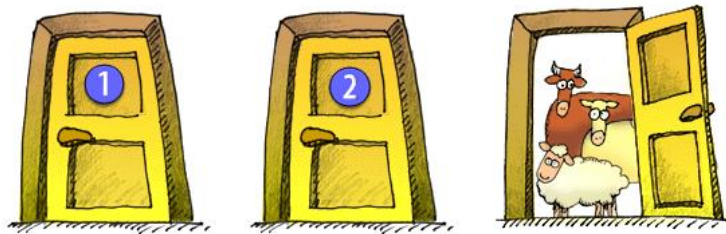
<http://www.apprendre-en-ligne.net/random/monty/>



- ▶ derrière une des portes = 1 bateau

# Monty Hall

<http://www.apprendre-en-ligne.net/random/monty/>



- ▶ derrière une des portes = 1 bateau
- ▶ derrière les autres portes : des moutons

# Monty Hall

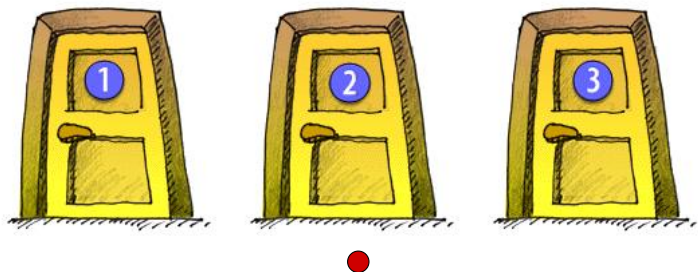
<http://www.apprendre-en-ligne.net/random/monty/>



- ▶ derrière une des portes = 1 bateau
- ▶ derrière les autres portes : des moutons

# Monty Hall

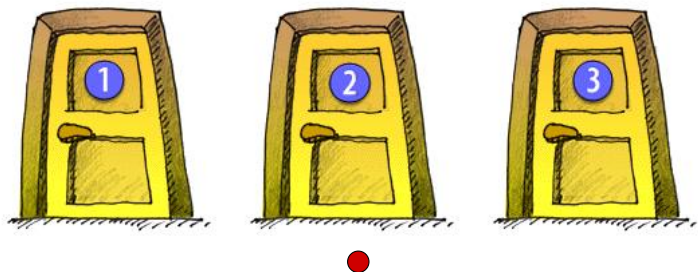
<http://www.apprendre-en-ligne.net/random/monty/>



- ▶ derrière une des portes = 1 bateau
- ▶ derrière les autres portes : des moutons

# Monty Hall

<http://www.apprendre-en-ligne.net/random/monty/>

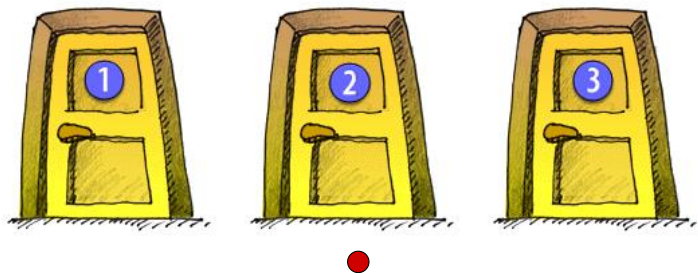


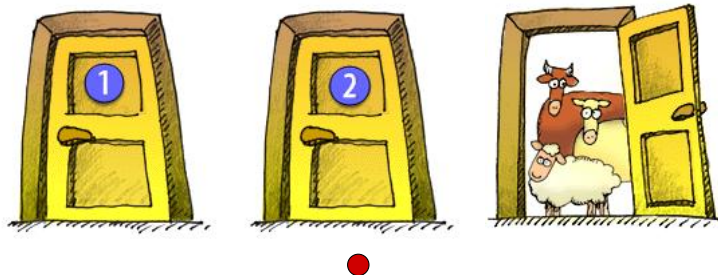
- ▶ derrière une des portes = 1 bateau
- ▶ derrière les autres portes : des moutons

proba que le bateau se trouve derrière la porte ?

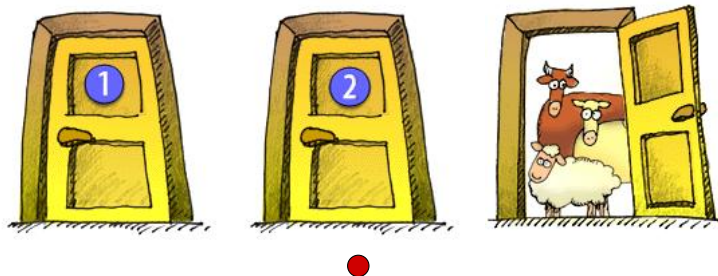


# Monty Hall



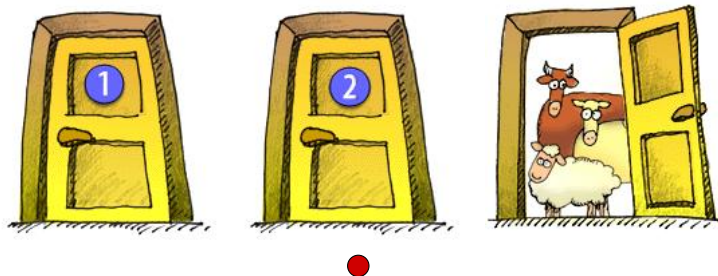


L'animateur choisit au hasard une des 2 autres portes qui contient des moutons et l'ouvre.



L'animateur choisit au hasard une des 2 autres portes qui contient des moutons et l'ouvre.

Choix : conserver la porte 2 ou choisir la porte 1 ?



L'animateur choisit au hasard une des 2 autres portes qui contient des moutons et l'ouvre.

Choix : conserver la porte 2 ou choisir la porte 1 ?

calculer la proba que le bateau soit derrière la porte 1 ou la 2

3 événements élémentaires :

- ▶ B1 : le bateau est derrière la porte 1
- ▶ B2 : le bateau est derrière la porte 2
- ▶ B3 : le bateau est derrière la porte 3

E = événement « l'animateur a choisi, parmi les portes 1 et 3, d'ouvrir la porte 3 »

3 événements élémentaires :

- ▶ B1 : le bateau est derrière la porte 1
- ▶ B2 : le bateau est derrière la porte 2
- ▶ B3 : le bateau est derrière la porte 3

E = événement « l'animateur a choisi, parmi les portes 1 et 3, d'ouvrir la porte 3 »

Choix  $\implies P(B1|E)$  et  $P(B2|E)$

$E$  = événement « l'animateur a choisi, parmi les portes 1 et 3, d'ouvrir la porte 3 »

$E$  = événement « l'animateur a choisi, parmi les portes 1 et 3, d'ouvrir la porte 3 »

$$\text{Formule de Bayes : } P(B1|E) = \frac{P(E|B1)P(B1)}{P(E)}$$



$E$  = événement « l'animateur a choisi, parmi les portes 1 et 3, d'ouvrir la porte 3 »

Formule de Bayes :  $P(B1|E) = \frac{P(E|B1)P(B1)}{P(E)}$

►  $P(E|B1) = 1$

$E$  = événement « l'animateur a choisi, parmi les portes 1 et 3, d'ouvrir la porte 3 »

$$\text{Formule de Bayes : } P(B1|E) = \frac{P(E|B1)P(B1)}{P(E)}$$

- ▶  $P(E|B1) = 1$
- ▶  $P(B1) = 1/3$

$E$  = événement « l'animateur a choisi, parmi les portes 1 et 3, d'ouvrir la porte 3 »

Formule de Bayes : 
$$P(B1|E) = \frac{P(E|B1)P(B1)}{P(E)}$$

▶  $P(E|B1) = 1$

▶  $P(B1) = 1/3$

▶ 
$$P(E) = P(E|B1)P(B1) + P(E|B2)P(B2) + P(E|B3)P(B3)$$
$$= 1 \times 1/3 + 1/2 \times 1/3 + 0 \times 1/3 = 1/2$$

$E$  = événement « l'animateur a choisi, parmi les portes 1 et 3, d'ouvrir la porte 3 »

$$\text{Formule de Bayes : } P(B1|E) = \frac{P(E|B1)P(B1)}{P(E)}$$

▶  $P(E|B1) = 1$

▶  $P(B1) = 1/3$

▶  $P(E) = P(E|B1)P(B1) + P(E|B2)P(B2) + P(E|B3)P(B3)$   
 $= 1 \times 1/3 + 1/2 \times 1/3 + 0 \times 1/3 = 1/2$

$$P(B1|E) = \frac{1 \times 1/3}{1/2} = 2/3$$

$E$  = événement « l'animateur a choisi, parmi les portes 1 et 3, d'ouvrir la porte 3 »

$$\text{Formule de Bayes : } P(B1|E) = \frac{P(E|B1)P(B1)}{P(E)}$$

$$\blacktriangleright P(E|B1) = 1$$

$$\blacktriangleright P(B1) = 1/3$$

$$\begin{aligned}\blacktriangleright P(E) &= P(E|B1)P(B1) + P(E|B2)P(B2) + P(E|B3)P(B3) \\ &= 1 \times 1/3 + 1/2 \times 1/3 + 0 \times 1/3 = 1/2\end{aligned}$$

$$P(B1|E) = \frac{1 \times 1/3}{1/2} = 2/3$$

$$P(B2|E) = 1 - P(B1|E) - P(B3|E) = 1/3$$

# Monty Hall



$$P(B1|E) = 2/3$$

$$P(B2|E) = 1/3$$

# Monty Hall



$$P(B1|E) = 2/3$$

$$P(B2|E) = 1/3$$

Problématique : consommation mémoire excessive



General Electric :

maintenance des moteurs CF6 :

350 variables aléatoires

>  $10^{105}$  événements élémentaires !



## Problématique : consommation mémoire excessive



### General Electric :

maintenance des moteurs CF6 :

350 variables aléatoires

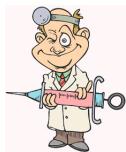
$> 10^{105}$  événements élémentaires !

### Monitoring de patients :

37 variables aléatoires

$> 10^{16}$  événements élémentaires !

## Problématique : consommation mémoire excessive



### General Electric :

maintenance des moteurs CF6 :

350 variables aléatoires

$> 10^{105}$  événements élémentaires !

### Monitoring de patients :

37 variables aléatoires

$> 10^{16}$  événements élémentaires !

### Liens entre gènes :

syndrome LQT – marqueur génétique

724 variables aléatoires

$> 10^{277}$  événements élémentaires !

# Solution : décomposabilité + indépendance

1 Décomposabilité

2 Indépendance

## 1 Décomposabilité

► **Définition** : Probabilités conditionnelles :  $P(A|B) = \frac{P(A, B)}{P(B)}$

## 2 Indépendance

## 1 Décomposabilité

► **Définition** : Probabilités conditionnelles :  $P(A|B) = \frac{P(A, B)}{P(B)}$

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

## 2 Indépendance

## 1 Décomposabilité

► **Définition** : Probabilités conditionnelles :  $P(A|B) = \frac{P(A, B)}{P(B)}$

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

$$\implies P(X_1, \dots, X_n) = P(X_1) \prod_{i=1}^n P(X_i | X_1, \dots, X_{i-1})$$

## 2 Indépendance

## 1 Décomposabilité

► **Définition** : Probabilités conditionnelles :  $P(A|B) = \frac{P(A, B)}{P(B)}$

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

$$\implies P(X_1, \dots, X_n) = P(X_1) \prod_{i=1}^n P(X_i | X_1, \dots, X_{i-1})$$

## 2 Indépendance

► **Définition** : Indépendance :  $A \perp\!\!\!\perp B \implies P(A|B) = P(A)$

## 1 Décomposabilité

► **Définition** : Probabilités conditionnelles :  $P(A|B) = \frac{P(A, B)}{P(B)}$

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

$$\implies P(X_1, \dots, X_n) = P(X_1) \prod_{i=1}^n P(X_i | X_1, \dots, X_{i-1})$$

## 2 Indépendance

► **Définition** : Indépendance :  $A \perp\!\!\!\perp B \implies P(A|B) = P(A)$

► **Définition** : Indép. conditionnelle :  $A \perp\!\!\!\perp B | C \implies P(A|B, C) = P(A|C)$



## 1 Décomposabilité

► **Définition** : Probabilités conditionnelles :  $P(A|B) = \frac{P(A, B)}{P(B)}$

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

$$\implies P(X_1, \dots, X_n) = P(X_1) \prod_{i=1}^n P(X_i | X_1, \dots, X_{i-1})$$

## 2 Indépendance

► **Définition** : Indépendance :  $A \perp\!\!\!\perp B \implies P(A|B) = P(A)$

► **Définition** : Indép. conditionnelle :  $A \perp\!\!\!\perp B | C \implies P(A|B, C) = P(A|C)$

►  $\{L_i, K_i\}$  = partition de  $\{X_1, \dots, X_{i-1}\}$  t.q.  $X_i \perp\!\!\!\perp L_i | K_i$

$$\implies P(X_1, \dots, X_n) = P(X_1) \prod_{i=1}^n P(X_i | K_i)$$

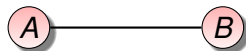
*A* : peinture

*B* : lecture



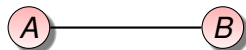
*A* : peinture

*B* : lecture

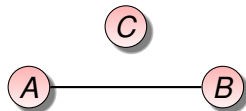


# Vers un modèle graphique pour les probas (1/3)

$A$  : peinture  
 $B$  : lecture

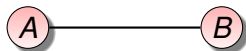


$A$  : peinture  
 $B$  : lecture  
 $C$  : âge

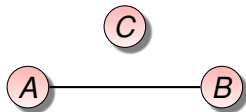


# Vers un modèle graphique pour les probas (1/3)

$A$  : peinture  
 $B$  : lecture



$A$  : peinture  
 $B$  : lecture  
 $C$  : âge



➔  $A$  et  $B$  sont dépendants

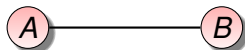
$\text{non}(A \perp\!\!\!\perp B)$

➔ sachant  $C$ ,  $A$  et  $B$  sont indépendants

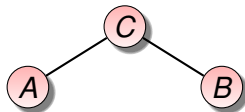
$A \perp\!\!\!\perp B \mid C$

# Vers un modèle graphique pour les probas (1/3)

$A$  : peinture  
 $B$  : lecture



$A$  : peinture  
 $B$  : lecture  
 $C$  : âge



➔  $A$  et  $B$  sont dépendants

$\text{non}(A \perp\!\!\!\perp B)$

➔ sachant  $C$ ,  $A$  et  $B$  sont indépendants

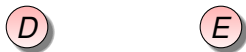
$A \perp\!\!\!\perp B \mid C$

$D$  : premier dé  
 $E$  : deuxième dé



# Vers un modèle graphique pour les probas (2/3)

$D$  : premier dé  
 $E$  : deuxième dé



$D$  : premier dé  
 $E$  : deuxième dé  
 $S$  :  $D + E$



➔  $D$  et  $E$  sont indépendants

$$D \perp\!\!\!\perp E$$

➔ sachant  $S$ ,  $D$  et  $E$  sont dépendants

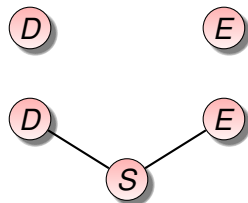
$$\text{non}(D \perp\!\!\!\perp E \mid S)$$



# Vers un modèle graphique pour les probas (2/3)

$D$  : premier dé  
 $E$  : deuxième dé

$D$  : premier dé  
 $E$  : deuxième dé  
 $S$  :  $D + E$



➔  $D$  et  $E$  sont indépendants

$$D \perp\!\!\!\perp E$$

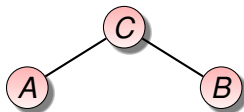
➔ sachant  $S$ ,  $D$  et  $E$  sont dépendants

$$\text{non}(D \perp\!\!\!\perp E \mid S)$$

# Vers un modèle graphique pour les probas (3/3)

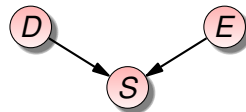
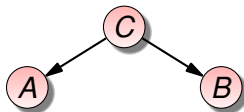
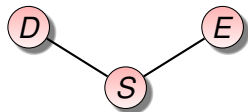
$\text{non}(A \perp\!\!\!\perp B)$

$A \perp\!\!\!\perp B \mid C$

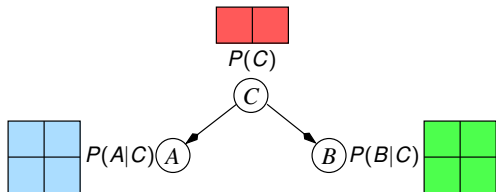


$D \perp\!\!\!\perp E$

$\text{non}(D \perp\!\!\!\perp E \mid S)$



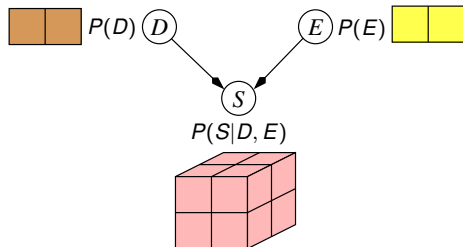
# Synthèse des exemples précédents



$$A \perp\!\!\!\perp B | C \quad \text{non}(A \perp\!\!\!\perp B)$$

$$P(A, B, C) = P(C)P(A|C)P(B|C, A)$$

$$P(A, B, C) = P(C)P(A|C)P(B|C)$$



$$D \perp\!\!\!\perp E \quad \text{non}(D \perp\!\!\!\perp E | S)$$

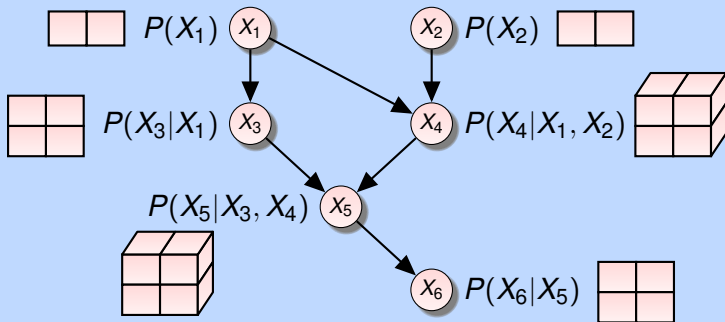
$$P(D, E, S) = P(D)P(E|D)P(S|D, E)$$

$$P(D, E, S) = P(D)P(E)P(S|D, E)$$

## Définition : réseau bayésien

[Pearl (1988)]

- 1 un graphe orienté acyclique (DAG) :



$$\text{distribution jointe : } P(X_1, \dots, X_6) = \prod_{i=1}^6 P(X_i | \mathbf{Pa}(X_i))$$

- 2 À chaque  $X_i$  est associé sa table  $P(X_i | \mathbf{Pa}(X_i))$

# Application pratique

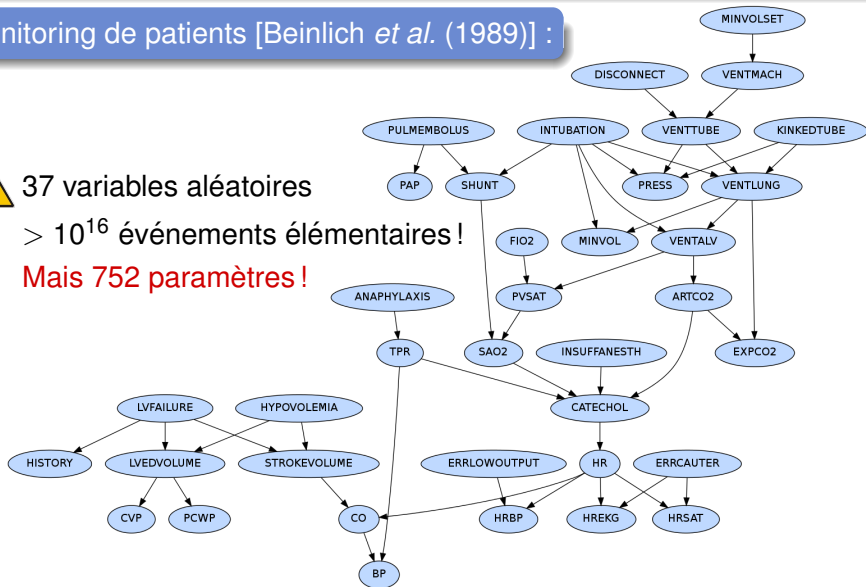
Monitoring de patients [Beinlich *et al.* (1989)] :



37 variables aléatoires

>  $10^{16}$  événements élémentaires !

Mais 752 paramètres !



- ▶ **Bayes, T., Price, R. (1763)** « An Essay towards Solving a Problem in the Doctrine of Chances », Philosophical Transactions of the Royal Society of London. 53 :370–418
- ▶ **Beinlich I., Suermondt H.J., Chavez R.M. et Cooper G.F. (1989)** « The ALARM Monitoring System : A Case Study with Two Probabilistic Inference Techniques for Belief Networks », Proceedings of the 2nd European Conference on Artificial Intelligence in Medicine, 247–256
- ▶ **Kolmogorov, A. (1933)** Foundations of the theory of probability. Chelsea Publishing Company.
- ▶ **Pearl, J. (1988)** Probabilistic reasoning in intelligent systems. Morgan Kaufmann