Indépendance

- Soit les variables A et B, elles sont **indépendantes** si et seulement si
  - $P(A|B) = P(A)$ ou
  - $P(B|A) = P(B)$ ou
  - $P(A, B) = P(A)\ P(B)$

- Exemple : $P(Pluie, Pourriel) = P(Pluie)\ P(Pourriel)$

P(Pluie = vrai) = 0.3

P(Pourriel = vrai) = 0.1

Pluie	***Pourriel***	**Probabilité**	
vrai	*vrai*	0.03	$= P(Pluie{=}\mathrm{V})\ P(Pourriel{=}\mathrm{V}) = 0.3 * 0.1$
vrai	*faux*	0.27	$= P(Pluie{=}\mathrm{V})\ P(Pourriel{=}\mathrm{F}) = 0.3 * 0.9$
faux	*vrai*	0.07	$= P(Pluie{=}\mathrm{F})\ P(Pourriel{=}\mathrm{V}) = 0.7 * 0.1$
faux	*faux*	0.63	$= P(Pluie{=}\mathrm{F})\ P(Pourriel{=}\mathrm{F}) = 0.7 * 0.9$

Indépendance

- L’indépendance totale est puissante mais rare
- L’indépendance entre les variables permet de réduire la taille de la distribution de probabilités et rendre les inférences plus efficaces
  - dans l’exemple précédent, on n’a qu’à stocker en mémoire $P(Pluie = vrai) = 0.3$ et $P(Pourriel = vrai) = 0.1$, plutôt que la table au complet
- Mais il est rare d’être dans une situation où toutes les variables sont réellement indépendantes

Indépendance conditionnelle

- Si j'ai une carie, la probabilité que la sonde accroche dans la dent ne dépend pas du fait que j'aie mal à la dent ou non :
  - $P(Croche \mid MalDeDents, Carie{=}vrai) = P(Croche \mid Carie{=}vrai)$
- Même chose si je n'ai pas la carie :
  - $P(Croche \mid MalDeDents, Carie{=}faux) = P(Croche \mid Carie{=}faux)$

- On dit que *Croche* est **conditionnellement indépendante** de *MalDeDents* étant donnée *Carie*, puisque :
  - $P(Croche \mid MalDeDents, Carie) = P(Croche \mid Carie)$
- Formulations équivalentes :
  - $P(MalDeDents \mid Croche , Carie) = P(MalDeDents \mid Carie)$
  - $P(MalDeDents, Croche \mid Carie) = P(MalDeDents \mid Carie)\, P(Croche \mid Carie)$

Indépendance conditionnelle

- Réécrivons la distribution conjointe en utilisant la **règle de chaînage** (***chain rule***) :

  $P(MalDeDents, Croche, Carie)$

  $= P(MalDeDents \mid Croche, Carie)\, P(Croche, Carie)$

  $= P(MalDeDents \mid Croche, Carie)\, P(Croche \mid Carie)\, P(Carie)$

  $= P(MalDeDents \mid Carie)\, P(Croche \mid Carie)\, P(Carie)$
- C-à-d., 2 + 2 + 1 = 5 **paramètres individuels/distincts**

- Dans des cas idéals, l'exploitation de l'indépendance conditionnelle réduit la complexité de représentation de la distribution conjointe de exponentielle (O(2^n)) en linéaire (O(n))
- En raisonnement probabiliste, l'indépendance conditionnelle est le concept de représentation des connaissances le plus basique et utile