Introduction à la biostatistique

Probabilités - notions

MASTER INFORMATIQUE MEDICALE

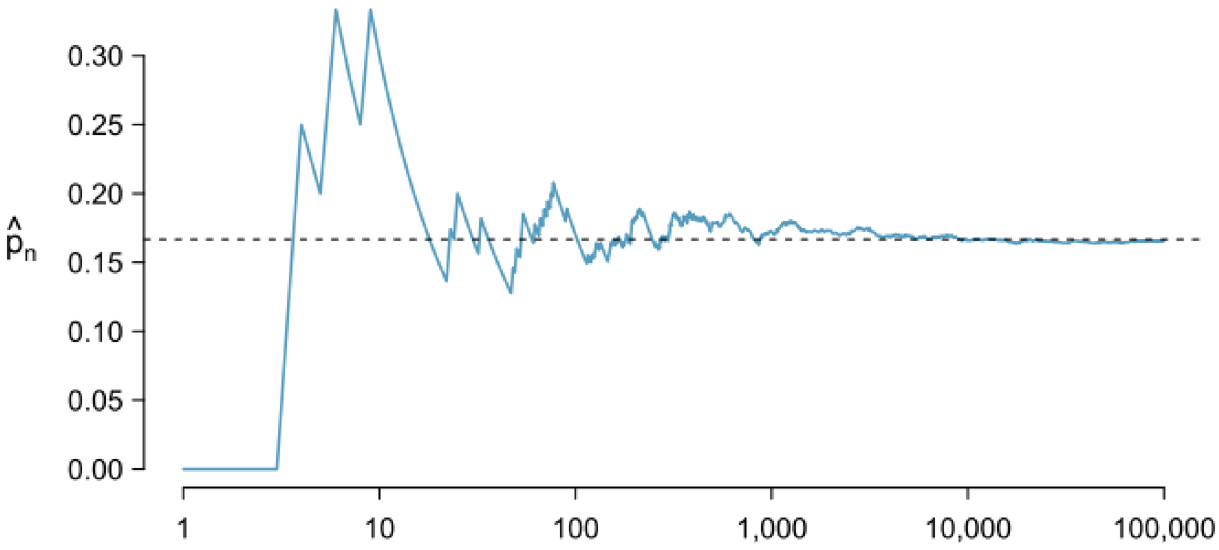
Kankoé SALLAH MD, PhD

kankoe.sallah@univ-amu.fr
kankoe@skml.fr

Nov 2021

Bases de probabilités

- **Expérience :** on lance un dé n fois. La variable représentée ci-dessous est la proportion (p_n) de fois où on obtient la face 1



Quand le nombre de lancers est très grand, cette proportion tend vers 1/6=0,167
→ 0,167 est la probabilité de l’évènement : « obtenir Face 1 à un lancer de dé à 6 faces »

Bases de probabilités

Définition de la probabilité d'un évènement

La probabilité d'un évènement est la proportion de fois où cet évènement s'accomplirait si on observait son processus aléatoire un nombre infini de fois.

Bases de probabilités

- **Probabilité :** discipline permettant l'évaluation de l'incertitude en statistiques
- Phénomène aléatoire : un phénomène est dit aléatoire si son issue ne peut être connu avant sa réalisation
- **Une variable aléatoire** peut prendre différentes valeurs dépendamment de la chance.
- On appelle **évènement** une réalisation de la variable (Ex face 1)
- On appelle probabilité la **proportion attendue** de réalisation d'un évènement dans une population
- **Espace de probabilité** : L'ensemble des réalisations possibles d'un phénomène constitue son espace de probabilité S

Exemples

Lancer de pièce S={Pile, Face}

Lancer de dé S={1,2,3,4,5,6}

Génotype S={AA, aa, Aa}

La somme des probabilités de l'espace de probabilité est 1

Un évènement est un sous-ensemble de l'espace de probabilité.

Bases de probabilités

- **Evènements disjoints :** ne peuvent se produire simultanément.

- **Evènements indépendants :** la réalisation de l'un n'a pas d'incidence sur la réalisation de l'autre.

- **2 évènements disjoints de probabilités non nulles ne sont jamais indépendants**

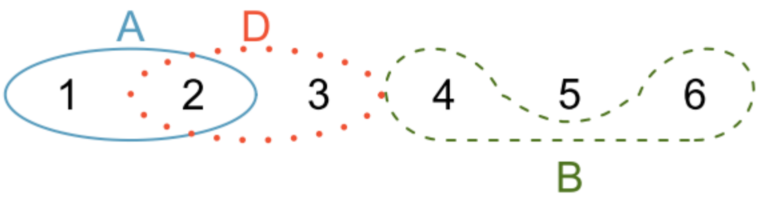
Bases de probabilités

Exemple

On considère le lancer d'un dé

1. Quelle est la probabilité de l'évènement A {1,2} ?
2. Quelle est la probabilité de l'évènement D {2,3} ?
3. Quelle est la probabilité de l'évènement B {4,6} ?
4. Quelle est la probabilité de l'évènement X {5} ?
5. Quelle est la probabilité de l'évènement AUB ?

P(A)=1/6+1/6



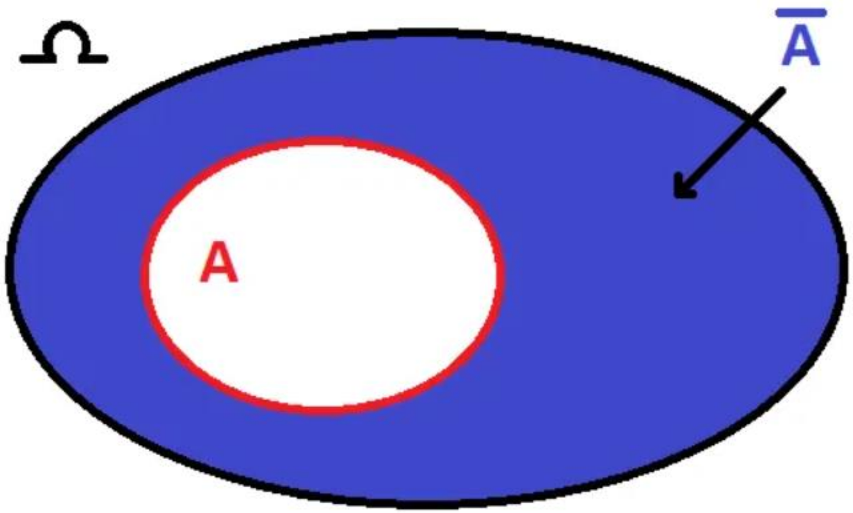
P(B)=1/6+1/6

P(AUB)=P (A or B) = P (A) + P (B) = 1/3 + 1/3 = 2/3

P(AUD)=P (A or D) = P (A) + P (D) - P (A et D) = 1/3 + 1/3-1/3 = 1/3

Note importante. 2 évènements disjoints de probabilités non nulles ne sont jamais indépendants

Bases de probabilités



- **L’évènement complémentaire Ā ou A^c** d’un évènement A est l’ensemble des issues de l’espace de probabilité qui ne font pas parti de A. C’est la non réalisation de A

$$P\left(\overline{A}\right) = 1 - P\left(A\right)$$

Bases de probabilités

- **L'union des évènements A et B est l'évènement :** A se réalise ou B se réalise ou A et B se réalisent simultanément. Il est noté AUB

- **L'intersection des évènements A et B est l'évènement :** A se réalise et B se réalise. Il est noté $A \cap B$

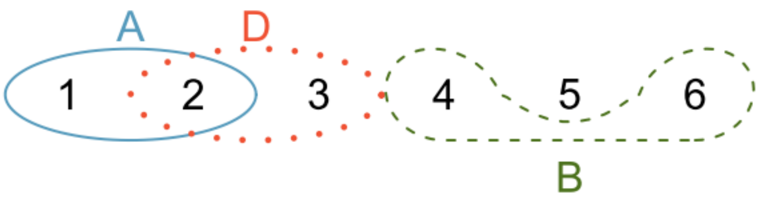
Bases de probabilités

Exemple

On considère le lancer d'un dé

1. Quelle est la probabilité de l'évènement A {1,2} ?
2. Quelle est la probabilité de l'évènement D {2,3} ?
3. Quelle est la probabilité de l'évènement B {4,6} ?
4. Quelle est la probabilité de l'évènement X {5} ?
5. Quelle est la probabilité de l'évènement AUB ?

P(A)=1/6+1/6

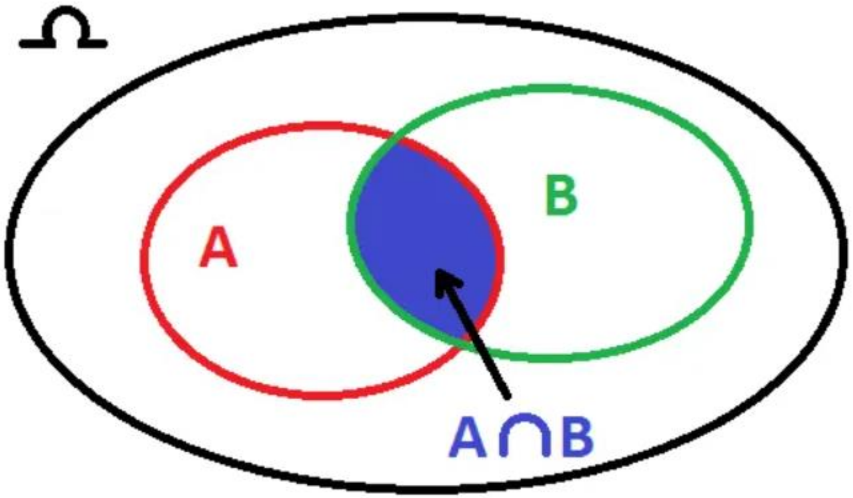


P(B)=1/6+1/6

P(AUB)=P (A or B) = P (A) + P (B) = 1/3 + 1/3 = 2/3

P(AUD)=P (A or D) = P (A) + P (D) - P (A et D) = 1/3 + 1/3-1/3 = 1/3

Note importante. 2 évènements disjoints de probabilités non nulles ne sont jamais indépendants



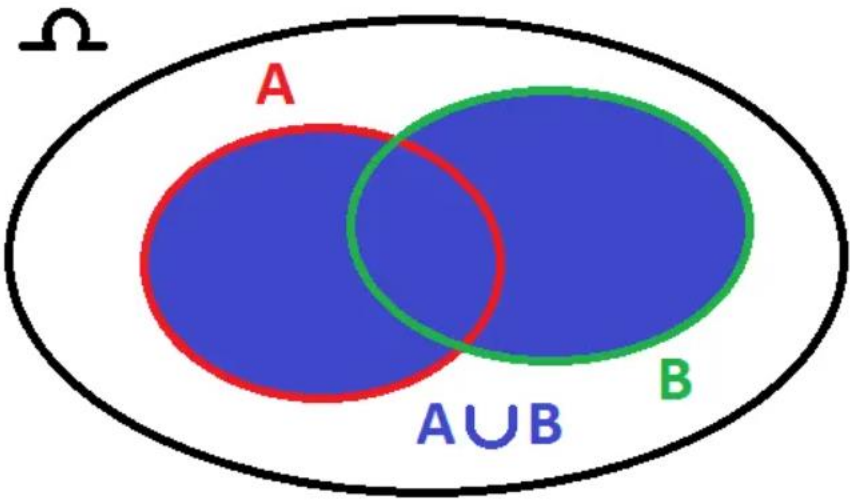
- **Probabilité conjointe de 2 évènements indépendants A et B**

$$P(A \cap B) = P(A) \times P(B)$$

Similairement, si k évènements dépendent de k processus aléatoires différents, on peut écrire :

$$P(A_1 \cap A_2 \ldots \cap A_k) = P(A_1) \times P(A_2) \times \ldots \times P(A_k)$$

Bases de probabilités

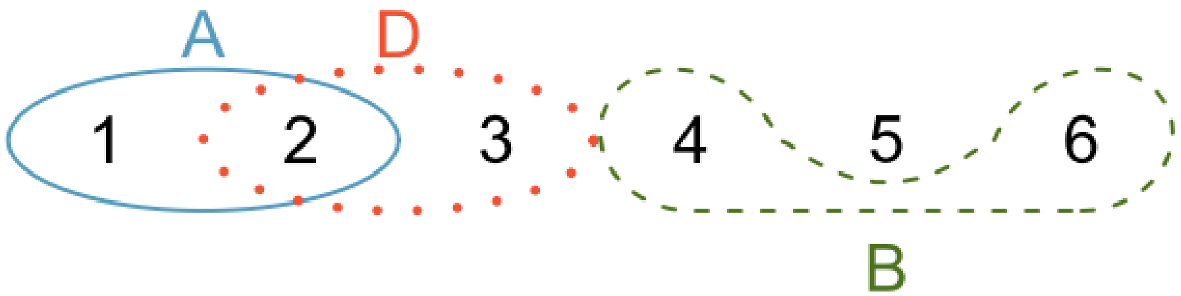


- **Probabilité marginale de A et B (de façon générale)**

$$P(A \cup B) = P(A) + P(B) - P(A) \cap P(B)$$

Bases de probabilités

Exemple Calculer P(A U B U D)



P(AUD)=P (A or D) = P (A) + P (D) - P (A et D) = 1/3 + 1/3-1/3 = 1/3

$$P(A \cup B \cup D)=P(A)+P(B)+P(D)-P(A\cap B)- P(A\cap D)- P(D\cap B)$$
$$=2/6+2/6+2/6-0-1/6-0$$
$$=5/6$$

Bases de probabilités

- **L'odd** ou **la côte** d'un évènement reflète la certitude qu'un événement puisse se produire plutôt que son complémentaire

$$Odd(E) = \frac{P(E)}{P(\overline{E})} = \frac{P(E)}{1 - P(E)}$$

- **L'odd ratio :** est le ratio des odds de 2 catégories

Exemple : Odd chez les exposés / Odd chez les non-exposés

Bases de probabilités

- **Probabilités conditionnelles.** Décrit une situation où la réalisation de A a une incidence sur a réalisation de B

Théorème de Bayes $$P(B|A) = \frac{P(A \cap B)}{P(A)}$$

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

Si A et B sont indépendants, alors P(B|A) = P(B)
d'où

$$P(A \cap B) = P(A) \times P(B)$$

Bases de probabilités

Exercice. 9% de la population burkinabè est gauchère. Supposons que 2 personnes sont sélectionnées au hasard. Les individus sont indépendants.

- Quelle est la probabilité que les 2 personnes soient gauchères ?

$$P=0.09 \times 0.09 = 0.0081.$$

- Quelle est la probabilité que les 2 personnes soient droitières ?

La probabilité qu'une personne soit droitière est de 1-0.09

Pour que les 2 soient droitières : $P=(1-0.09)*(1-0.09)=0.8281$

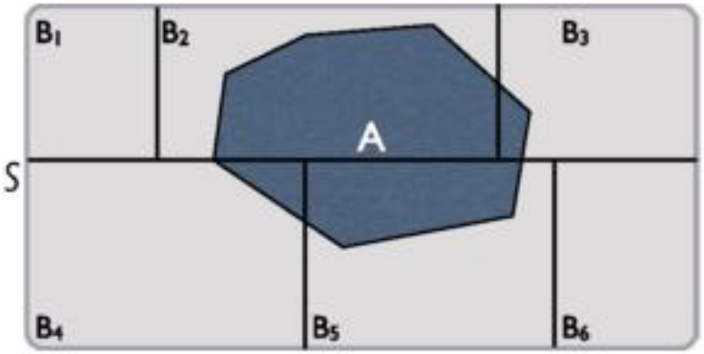
Bases de probabilités

- **Théorème des probabilités totales.** Si A1, A2…An représentent une partition de A alors

$$P(B) = P(A_1) \times P(B \mid A_1) + P(A_2) \times P(B \mid A_2) + \ldots + P(A_n) \times P(B \mid A_n)$$

Or

$$P(Ai \mid B) = \frac{P(B \mid A_i) \times P(A_i)}{P(B)}$$



D'où

$$\boxed{P(Ai \mid B) = \frac{P(B \mid A_i) \times P(A_i)}{P(A_1) \times P(B \mid A_1) + P(A_2) \times P(B \mid A_2) + \ldots + P(A_n) \times P(B \mid A_n)}}$$

Probabilités en épidémiologie

- **Prévalence.** Proportion d’individus ayant une maladie dans le population en un instant t

- **Incidence.** Nouveaux cas / Δt

Probabilités en épidémiologie

		Réalité (gold-standard)	
		Malades	Sains
Test	+	*Vrais Positifs (VP)*	*Faux Positifs (FP)*
	-	*Faux Négatifs (FN)*	*Vrais Négatifs (VN)*
		Sensibilité = P(Test+/malades) = VP / (VP+FN)	**Spécificité** = P(Test-/sains) = VN / (VN+FP)

→ Qualité du test

Probabilités en épidémiologie

Réalité
(gold-standard)

		Malades	Sains	
Test	+	*Vrais Positifs (VP)*	*Faux Positifs (FP)*	**VPP** = P(Malades/Test+) = VP / (VP+FP)
	-	*Faux Négatifs (FN)*	*Vrais Négatifs (VN)*	**VPN** = P(Sains/Test-) = VN / (VN+FN)

→ Apport du test pour un patient donné

Indicateurs extrinsèques car dépendent du contexte épidémiologique (prévalence de la maladie)

Probabilités en épidémiologie

Maladie → Exposition ↓	M+ Malades	M- Non Malades	Total
E+ Exposés	O_{11}	O_{10}	L_1
E- Exposés	O_{01}	O_{00}	L_0
Total	C_1	C_0	n

$$\psi = \textit{Rapport} \text{ de côtes} = \frac{\textit{Côte} \text{ de l'exposition chez les malades}}{\textit{Côte} \text{ de l'exposition chez les non malades}} = \frac{\dfrac{O_{11}}{O_{01}}}{\dfrac{O_{10}}{O_{00}}} = \frac{O_{11}}{O_{01}} \times \frac{O_{00}}{O_{10}} = \frac{O_{11}O_{00}}{O_{10}O_{01}}$$

Plus l'exposition est spécifique des malades, plus Ψ est grand

Probabilités en épidémiologie

Exposition ↓ Maladie →	M+	M-	Total
E+	O_{11}	O_{10}	L_1
E-	O_{01}	O_{00}	L_0
Total	C_1	C_0	n

- Risque relatif

Valable si échantillon représentatif

$$RR = \frac{\text{Risque chez les exposés}}{\text{Risque chez les non exposés}} = \frac{\dfrac{O_{11}}{O_{11}+O_{10}}}{\dfrac{O_{01}}{O_{01}+O_{00}}} = \frac{R_1}{R_0}$$

- Odds ratio

$$OR = \psi = \mathit{Rapport}\ \text{de côtes} = \frac{\mathit{Côte}\ \text{de l'exposition chez les malades}}{\mathit{Côte}\ \text{de l'exposition chez les non malades}} = \frac{\dfrac{O_{11}}{O_{01}}}{\dfrac{O_{10}}{O_{00}}} = \frac{O_{11}}{O_{01}} \times \frac{O_{00}}{O_{10}} = \frac{O_{11}O_{00}}{O_{10}O_{01}}$$

On démontre :

$$OR = RR \times \frac{1-R_0}{1-R_1}$$

Si maladie rare [$R_0 << 1$ et $R_1 << 1$] alors OR≈RR

Merci