Aix–Marseille Université	2012–2013

Maths pour la bio

COURS : PROBABLITÉS

1 Probabilité

1.1 Loi de probabilité

Définition 1. On appelle *expérience aléatoire* tout processus dont on sait que le résultat appartient à un ensemble donné mais sans que l'on puisse pour autant déterlminer avec certitude quel sera ce résultat.
On appelle *univers* l'ensemble Ω des résultats possibles

Exemples 2.

Expérience	Univers Ω
lancé de dé	$\{1, 2, 3, 4, 5, 6\}$
couleur des yeux d'un individu	$\{\text{bleu}, \text{gris-vert}, \text{noir}\}$
taux d'hémoglobine d'un individu	$\mathbb{R}_+$

Définition 3. On appelle *événement* tout sous-ensemble $A \subset \Omega$ de l'univers. Cela correspond à une potientialité de résultat.

Exemples 4.

Expérience	Evénement	A
lancé de dé	faire un 5	$\{5\}$
lancé de dé	ne pas faire un 5	$\{1, 2, 3, 4, 6\}$
lancé de dé	faire un nombre pair	$\{2, 4, 6\}$
taux d'hémoglobine (g/l)	taux normal pour un homme	$[140, 180]$

Définition 5. A toute expérience aléatoire, on associe une variable aléatoire X, modélisant le résultat de l'expérience, munie d'une loi de probabilité

$$\mathbb{P} : \begin{array}{ccc} \{\text{événements } A \subset \Omega\} & \longrightarrow & [0, 1] \\ A & \longmapsto & \mathbb{P}(A) \text{ ou } \mathbb{P}(X \in A) \end{array}$$

qui associe à tout événement la probabilité que le résultat satisfasse cet événement.

Exemples 6.

Expérience	Evénement	Probabilité $\mathbb{P}(X \in A)$
lancé de dé	faire un 5	$\frac{1}{6}$
lancé de dé	ne pas faire un 5	$\frac{5}{6}$
lancé de dé	faire un nombre pair	$\frac{1}{2}$
taux d'hémoglobine (g/l)	taux normal pour un homme	$0,91 = 91\%$

Proposition 7.

i. $\mathbb{P}(\Omega) = 1$.

ii. $\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B)$ *si* $A \cap B = \emptyset$ *(c'est à dire si les événements* A *et* B *sont incompatibles).*

Proposition 8.

i. $\mathbb{P}(\emptyset) = 0$.

ii. $\mathbb{P}(\overline{A}) = 1 - \mathbb{P}(A)$ *où* $\overline{A}$ *est l'événement contraire de* A.

iii. $\mathbb{P}(A) \leq \mathbb{P}(B)$ *si* $A \subset B$.

iv. $\mathbb{P}(A \cup B) = \mathbb{P}(A) + \mathbb{P}(B) - \mathbb{P}(A \cap B)$.

1.2 Probabilités conditionelles

Imposer une condition, c'est ne plus regarder tout l'univers des possibles, mais seulement le sous-monde des résultats vérifiant cette condition. On se limite donc aux expériences dont le résultat faire partie d'un événement A (la condition).

L'expérience aléatoire est donc modifiée dans le sens ou l'univers Ω est remplacé par A. La variable aléatoire ainsi que sa loi de probabilité sont modifiées en conséquence.

Définition 9. Si X est une variable aléatoire, $\mathbb{P}$ sa loi de probabilité et A un événement, on note $\mathbb{P}_A$ (probabilité sachant A) la nouvelle loi de probabilité sur l'univers A obtenue en imposant la condition "le résultat est dans A".
On note aussi parfois $\mathbb{P}(B|A)$ pour $\mathbb{P}_A(B)$. Cela se lit "probabilité d'avoir B sachant A".

Exemple 10. X est la variable associée à un lancer de dé. On a $\mathbb{P}(X = 1) = \mathbb{P}(X = 2) = \mathbb{P}(X = 3) = \mathbb{P}(X = 4) = \mathbb{P}(X = 5) = \mathbb{P}(X = 6) = \frac{1}{6}$. Si on impose la condition "le résultat est pair" correspondant à l'événement $A = \{2, 4, 6\}$, on obtient une nouvelle loi de probabilité $\mathbb{P}_A X = 2 = \mathbb{P}_A X = 4 = \mathbb{P}_A X = 6 = \frac{1}{3}$.

Proposition 11. $\mathbb{P}(B \cap A) = \mathbb{P}(A)\mathbb{P}_A(B)$.

Définition 12. On dit qu'un événement B est indépendant d'un autre événement A si $\mathbb{P}_A(B) = \mathbb{P}(B)$.

Proposition 13. *B est indépendant de A ssi $\mathbb{P}(A \cap B) = \mathbb{P}(A)\mathbb{P}(B)$.*

Corollaire 14. *B est indépendant de A ssi A est indépendant de B. On dit alors que A et B sont indépendants (sans se soucier de l'ordre).*

2 Caractérisation d'une variable aléatoire

2.1 Variable qualitative

L'univers Ω d'une variable aléatoire X est constitué d'un nombre fini (discret) de résultats non numérique. La seule façon de décrire entièrement la loi sur X est de donner $\mathbb{P}(X = a)$ pour tout élément a de Ω.

Exemple 15. On tire à pile ou face avec une pièce baisée : $\mathbb{P}(X = pile) = 60\%$ et $\mathbb{P}(X = face) = 40\%$.

Remarque 16. Il suffit de connaitre la probabilité de tout élément sauf un.

2.2 Variable quantitative discrète

L'univers Ω d'une variable aléatoire X est constitué d'un nombre fini (discret) de valeurs numériques. On peut toujours décrire la loi en donnant la probabilité $\mathbb{P}(X = a)$ pour tout événement élémentaire a de Ω.

Exemple 17. On regarde le nombre d'enfant dans une famille : $\mathbb{P}(X = 0) = 30\%$, $\mathbb{P}(X = 1) = 25\%$, $P(X = 2) = 30\%$, $P(X = 3) = 10\%$, $P(X = 4) = 5\%$.

Mais la variable X est également entièrement décrite par sa fonction de répartition F qui donne la probabilité d'avoir un résultat inférieur à une valeur donnée : $F(a) = \mathbb{P}(x \leq a)$. C'est une fonction croissante qui va de 0 à 1.

Exemple 18. On regarde toujours le nombre d'enfant dans une famille : $\mathbb{F}(0) = 30\%$, $F(1) = 55\%$, $F(2) = 85\%$, $F(3) = 95\%$, $F(4) = 100\%$.

Remarque 19. On récupère $\mathbb{P}(X = a)$ en faisant $F(a) - F(a - 1)$ et $\mathbb{P}(a < X \leq b)$ en faisant $F(b) - F(a)$.

2.3 Variable quantitative continue

L'univers Ω d'une variable aléatoire X est un intervalle entier.

Remarque 20. Une variable quantitative discrète pourra parfois être considérée comme continue si il y a un nombre très important de valeurs possibles.

Cela n'a pas de sens de considérer $\mathbb{P}(X = a)$ car cela vaudra toujours zero (ou presque). Par contre, la variable aléatoire est toujours caractérisée par sa fonction de répartition.

Exemple 21. Longueur (en cm) des saumons en Arctique : $\mathbb{P}(X = 63,23) = 0$ mais $F(50) = 25\%$, $F(65) = 50\%$, $F(80) = 95\%$.

Il y a toutefois un "équivalent" de la donnée de $\mathbb{P}(X = a)$, c'est la fonction de densité : une fonction g telle que $\mathbb{P}(a < X \leq b) = F(b) - F(a) = \int_a^b g(x)dx$, c'est à dire une fonction telle que l'aire sous la courbe entre a et b est égale à la probabilité que X soit entre a et b.

Remarque 22. Si on approche une variable continue par des variables discrètes en découpant en intervalles de plus en plus petit, la suite des histogrammes obtenus va converger vers le graphe de la fonction de densité. Cela justifie de faire des histogrammes dont la surface (et non la hauteur) des barres est proportionnelle à l'effectif.

3 Variables aléatoires quantitatives

Dans cette section, on ne considère que des variables quantitatives.

3.1 Opérations sur les variables aléatoires

Une variable aléatoire X correspond au résultat d'une expérience aléatoire. Si c'est une variable quantitative, on peut appliquer une fonction f à X. Cela définit une nouvelle expérience, où le résultat observé n'est plus X mais $f(X)$.

Exemples 23.

1. Si X est la variable aléatoire du prix d'une maison en euros, la variable aléatoire $Y = 655,957.X$ sera le prix d'une maison en ancien francs.
2. Si X est la variable aléatoire du rayon d'une tache circulaire, $Y = \pi X^2$ sera celle de sa surface.

La fonction peut éventuellement prendre plusieurs variables en argument.

Exemples 24.

1. Si X et Y sont chacune la variable aléatoire du lancé d'un dé, la variable $Z = X + Y$ sera la variable aléatoire du résultat obtenu en lançant les deux dés simultanément.
2. Si X est la variable aléatoire de la taille d'une personne et Y celle de son poids, alors $Z = \frac{Y}{X^2}$ sera la variable aléatoire de l'IMC de cette personne.

3.2 Espérance

L'espérance est à une variable aléatoire ce que la moyenne est à un jeu de donnée statistique.

Définition 25. Si X discrète : $\mathbb{E}(X) = \sum_{\alpha \in \Omega} \alpha p_\alpha$ avec $p_\alpha = \mathbb{P}(X = \alpha)$.

Si X continue : $\mathbb{E}(X) = \int_a^b xg(x)$ où $\Omega = [a, b]$ et g est la fonction de densité.

Proposition 26. *Soit X, Y deux variables aléatoires quelconques et $a \in \mathbb{R}$.*

1. $\mathbb{E}(X+Y) = \mathbb{E}(X) + \mathbb{E}(Y)$;
2. $\mathbb{E}(aX) = a\mathbb{E}(X)$.

Proposition 27. *Soit X et Y deux variables aléatoires indépendantes, alors* $\mathbb{E}(XY) = \mathbb{E}(X)\mathbb{E}(Y)$.

Remarque 28. Ce dernier est faux si les variables ne sont pas indépendantes. En particulier $\mathbb{E}(X^2) \neq \big(\mathbb{E}(X)\big)^2$.

3.3 Variance

La variance évalue l'étendu probable des résultats d'une expérience aléatoire.

Définition 29. Dans tous les cas, $\mathrm{Var}(X) = \mathbb{E}\Big(\big(X - \mathbb{E}(X)\big)^2\Big) = \mathbb{E}(X^2) - \big(\mathbb{E}(X)\big)^2$.

Si X discrète : $\mathrm{Var}(X) = \displaystyle\sum_{\alpha \in \Omega} \big(\alpha - \mathbb{E}(X)\big)^2 p_\alpha = \sum_{\alpha \in \Omega} \alpha^2 p_\alpha - \Big(\sum_{\alpha \in \Omega} \alpha p_\alpha\Big)^2$.

Si X continue : $\mathrm{Var}(X) = \int_a^b \big(x - \mathbb{E}(X)\big)^2 g(x) = \int_a^b x^2 g(x) - \Big(\int_a^b x g(x)\Big)^2$.

Proposition 30. *Soit X une variable aléatoire et $a \in \mathbb{R}$.*

1. $\mathrm{Var}(X+a) = \mathrm{Var}(a)$;
2. $\mathrm{Var}(aX) = a^2\mathrm{Var}(X)$.

Proposition 31. *Soit X et Y deux variables aléatoires indépendantes, alors* $\mathrm{Var}(X+Y) = \mathrm{Var}(X) + \mathrm{Var}(Y)$.

4 Exemples de lois classiques

4.1 Lois discrètes

4.1.1 Loi uniforme

Modélisation :
Une expérience ayant un nombre fini de résultats possibles équiprobables. Par exemple un lancé de dé.

Paramètre :
Le nombre n de résultats possibles.

Définition :
$\mathbb{P}(X = k) = \frac{1}{n}$ pour tout $k \in \{1, \cdots, n\}$.

Caractéristiques :
– $\mathbb{E}(X) = \frac{n+1}{2}$;
– $\mathrm{Var}(X) = \frac{n^2-1}{12}$.

Remarque 32. Si les résultats vont de $p = 1 + (p-1)$ à $n + p - 1$, on considérera $Y = X + p - 1$.

4.1.2 Loi de Bernouilli

Modélisation :
Une expérience ayant deux résultats possibles *a priori* non équiprobables. Par exemple un pile ou face éventuellement truqué.

Paramètre :
La propbabilité p d'un des résultats.

Définition :
$\mathbb{P}(X = 1) = p$ et $\mathbb{P}(X = 0) = 1 - p$.

Caractéristiques :
– $\mathbb{E}(X) = p$;
– $\mathrm{Var}(X) = p(1 - p)$.

Remarque 33. Si les deux résultats sont a et b, on considérera $Y = (a - b)X + b$.

4.1.3 Loi binomiale

Modélisation :
Le nombre d'expériences ayant un résultat donné lorsque l'on répéte plusieurs fois une expérience ayant deux résultats possibles. Par exemple le nombre de pile obtenu en lançant n pièces. Il importe que le résultat d'une des expériences soit indépendant des résultats des autres.

Paramètres :
La propbabilité p du résultat dénombré et le nombre n de fois où l'expérience est répétée.

Définition :
$\mathbb{P}(X = k) = \binom{n}{k} p^k (1 - p)^k$ pour tout $k \in \{0, \cdots, n\}$.

Caractéristiques :
– $\mathbb{E}(X) = np$;
– $\mathrm{Var}(X) = np(1 - p)$.

4.1.4 Loi de Poisson

Modélisation :
On considère un événement susceptible d'arriver à tout moment avec une probabilité constante et proportionnel à tout laps de temps (raisonablement court) considéré. On veut modéliser le nombre de fois où cet événement est intervenu depuis un intant t_0. Par exemple le nombre de clients dans une file d'attente à un guichet fermé après un temps donné.

Paramètres :
Un coefficient λ lié au rapport enre la probabilité de réalisation de l'événement pendant un laps de temps et ce laps de temps.

Définition :
$\mathbb{P}(X = k) = \frac{e^{-\lambda}\lambda^k}{k!}$ pour tout $k \in \mathbb{N}$.

Caractéristiques :
– $\mathbb{E}(X) = \lambda$;
– $\mathrm{Var}(X) = \lambda$.

4.2 Lois continues

4.2.1 Loi uniforme

Modélisation :
Cette loi correspond au cas discret uniforme lorsqu'il y a un nombre n de résultats possibles de plus en plus en grand, mais toujours entre deux valeurs réels a et b.

Paramètres :
Les deux bornes a et b des valeurs possibles.

Définition :
La fonction de densité vaut $g(x) = \frac{1}{b-a}$ et l'on a donc $\mathbb{P}(\alpha \leq X \leq \beta) = \int_\alpha^\beta \frac{dx}{b-a} = \frac{\beta-\alpha}{b-a}$ pour tout α et β tels que $a \leq \alpha \leq \beta \leq b$.

Caractéristiques :
- $\mathbb{E}(X) = \frac{a+b}{2}$;
- $\mathrm{Var}(X) = \frac{(b-a)^2}{12}$.

4.2.2 Loi normale

Modélisation :
Cette loi est ce vers quoi converge une loi binomial lorsque l'on répète l'expérience un très grand nombre de fois.

Paramètres :
Il existe deux paramètres souvent noté $\mu \in \mathbb{R}$ et $\sigma \in \mathbb{R}_+^*$ en raison de ce qui suit.

Définition :
La fonction de densité vaut $g(x) = \frac{1}{\sigma\sqrt{2\pi}} e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2}$ et l'on a donc $\mathbb{P}(\alpha \leq X \leq \beta) = \frac{1}{\sigma\sqrt{2\pi}} \int_\alpha^\beta e^{-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2} dx$.

Caractéristiques :
- $\mathbb{E}(X) = \mu$;
- $\mathrm{Var}(X) = \sigma^2$.

Définition 34. Si $\mu = 0$ et $\sigma = 1$, on parle de loi normale centrée réduite.

Proposition 35. *Si X est une loi normale de paramètres μ et σ, alors $Y = \frac{X-\mu}{\sigma}$ est une loi normale centrée réduite.*