COURS

2 lois de probabilité

Les lois de probabilité sont essentielles en biologie pour quantifier et prédire la variabilité dans divers processus biologiques. Elles permettent aux biologistes d'analyser des données, de formuler des hypothèses et de prendre des décisions éclairées dans des domaines tels que la génétique, l'évolution et l'écologie. Ces lois mathématiques contribuent ainsi à une meilleure compréhension des phénomènes aléatoires dans le monde vivant.

1 Généralités

Définition

Une probabilité est une application $P : \Omega \to [0,1]$ telle que : pour tout $A \in \Omega$, on a

1. $P(A) \geq 0$,
2. $P(\Omega) = 1$,
3. Si $A \cap B = \varnothing$, alors $P(A \cup B) = P(A) + P(B)$.

Remarque

On appele espace probabilisé le triplet $(\Omega, P(\Omega), P)$

Propriétés élémentaires

1. $P(A) = 1 - P(A)$
2. $P(\varnothing) = 0$
3. $A \subset B \Rightarrow P(A) \leq P(B)$
4. $0 \leq P(A) \leq 1$
5. $P(A \cup B) = P(A) + P(B) - P(A \cap B)$

Remarque

P(A) = 0 ne signifie pas que A est un événement impossible.

Probabilité conditionnelle

Soit A et B deux événements tels que $P(B) \neq 0$. La probabilité conditionnelle de A par rapport à B, notée $P(A \backslash B)$ ou $P_B(A)$ (probabilité de A sachant B), est donnée par :
$P(A \backslash B) = \dfrac{P(A \cap B)}{P(B)}$

Définition

Soit E un ensemble. $A_1, A_2, ..., A_n$ constituent une partition de E si :

1. $\forall\ i \in 1, ..., n;\ A_i \neq \varnothing$
2. $\forall\ i \neq j;\ A_i \cap A_j = \varnothing$
3. $A_1 \cup A_2 \cup ... \cup A_n = E$

Formule des probabilités totales

Si les événements $B_1, B_2, ..., B_n$ forment une partition de Ω

$$P(A) = P(B_1)P(A \backslash B_1) + P(B_2)P(A \backslash B_2) + ... + P(B_n)P(A \backslash B_n) = \sum_{i=1}^{n} P(B_i)P(A \backslash B_i)$$

Formule de Bayes

Si $P(A) \neq 0$ et $P(B) \neq 0$ alors

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

Si les événements $A_1, A_2, ..., A_n$ forment une partition de Ω alors

$$P(A_i|B) = \frac{P(A_i)P(B|A_i)}{\sum_{i=1}^{n} P(A_i)P(B|A_i)}$$

Exemple

Un examen systématique de dépistage est institué pour détecter une maladie M . On sait que le risque d'avoir cette maladie est de 0.001. L'examen donne des faux positifs avec probabilité 0.1 et des faux négatifs avec une probabilité de 0.3. Un individu subit un examen qui se révèle négatif. Quelle est la probabilité qu'il soit malade ?

Indépendance

Deux événements A et B sont dits indépendants si et seulement si

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

Remarques

Si A et B sont indépendants, alors

$$P(A|B) = P(A) \Leftrightarrow P(B|A) = P(B)$$

Attention à ne pas confondre indépendance et incompatibilité.

2 lois de probabilité discrète

2.2.1 Loi de Bernoulli

Définition

La loi de Bernoulli de paramètre p est la loi d'une variable aléatoire discrète X qui prend la valeur 1 avec probabilité p et la valeur 0 avec probabilité $1-p$. L'expérience associée est appelé une **épreuve de Bernoulli**.

Notation

$$X \sim \mathcal{B}(p)$$

Fonction de probabilité

$$P(X=x) = \begin{cases} p & si \quad x=1 \\ 1-p & si \quad x=0 \\ 0 & sinon \end{cases}$$

Espérance et variance

$$E(X) = \sum_{i=1}^{2} x_i p_i = p$$

$$Var(X) = \sum_{i=1}^{2} x_i^2 p_i - E(X)^2 = p(1-p)$$

Exemple

Pile ou Face

2.2.2 Loi binomiale

Définition

La loi binomiale de paramètres n et p est la loi de la somme X de n variables aléatoires Y_i indépendantes telles que $Y_i \sim \mathcal{B}(p)$.

Notation

$$X \sim \mathcal{B}(n, p)$$

Fonction de probabilité

$$P(X = k) = C_n^k p^k (1 - p)^{(n-k)}; k = 1, ..., n.$$

Espérance et variance

$$E(X) = \sum_{i=1}^{n} E(Y_i) = np$$

$$Var(X) = \sum_{i=1}^{n} Var(Y_i) = np(1 - p)$$

Exemple

Comptage du nombre de succès sur n épreuves de Bernoulli.

3 lois de probabilité continues

2.3.1 Loi normale

Définition

Une variable aléatoire X suit une loi normale (ou loi de Gauss-Laplace) de paramètres μ et σ^2 si :

$$f(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2}\frac{(x-\mu)^2}{\sigma^2}}$$

On note :

$$X \sim \mathcal{N}(\mu, \sigma^2)$$

Fonction de probabilité

$$F(x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi\sigma^2}} e^{-\frac{1}{2}\frac{(x-\mu)^2}{\sigma^2}}$$

Il n'existe pas de forme analytique de la fonction de répartition F .

Espérance et variance

Si X est une variable aléatoire réelle telle que $X \sim \mathcal{N}(\mu, \sigma^2)$, alors :

$$E(X) = \mu$$

$$Var(X) = \sigma^2$$

Stabilité par combinaisons linéaires

1. Si X est une variable aléatoire telle que $X \sim \mathcal{N}(\mu, \sigma^2)$, alors :

$$aX + b \sim \mathcal{N}(a\mu + b, a^2\sigma^2)$$

2. Si X et Y sont deux variables aléatoires indépendantes telles que $X \sim \mathcal{N}(\mu_1, \sigma_1^2)$ et $Y \sim \mathcal{N}(\mu_2, \sigma_2^2)$, alors :

$$X + Y \sim \mathcal{N}(\mu_1 + \mu_2, \sigma_1^2 + \sigma_2^2)$$

2.3.2 Loi normale centrée réduite

Définition

On appelle loi normale centrée réduite la loi $\mathcal{N}(0, 1)$.

Propriété

Si $X \sim \mathcal{N}(\mu, \sigma^2)$, alors : $Y = \dfrac{X - \mu}{\sqrt{\sigma^2}} \sim \mathcal{N}(0, 1)$

Notations

Par convention, on note ϕ la densité d'une $\mathcal{N}(0, 1)$ et Φ sa fonction de répartition.

Propriétés

1. ϕ est une fonction paire
2. $\Phi(x) = 1 - \Phi(-x)$
3. $P(|X| \leq x) = P(-x \leq X \leq x) = 2(\Phi(x) - 1/2)$
4. $P(|X| \geq x) = P((X \leq -x) \cap (X \geq x)) = 2(1 - \Phi(x))$

4 lois déduites de la loi normale

2.4.1 Loi du khi-deux χ_n^2

Définition

Soient $X_1, ..., X_n$ n variables aléatoires indépendantes et identiquement distribuées de loi normale centrée réduite. La variable aléatoire $Y = X_1^2 + ... + X_n^2$ n suit une loi continue appelée loi du χ^2 à n degrés de liberté :

$$Y = \sum_{i=1}^{n} X_i^2 \sim \chi_n^2$$

Propriétés

1. Si $Y_1 \sim \chi^2_{n_1}$ et $Y_2 \sim \chi^2_{n_2}$ avec $Y_1 \perp\perp Y_2$, alors $Y = Y_1 + Y_2 \sim \chi^2_{n_1+n_2}$
2. Si $Y \sim \chi^2_n$, alors $E(Y) = n$ et $Var(Y) = 2n$

Densité

$$f(y) = \frac{1}{2^{n/2}\Gamma(n/2)} x^{(n-2)/2} e^{-x/2}$$

2.4.2 Loi de student T

Définition

Soient X et Y deux variables aléatoires indépendantes telles que $X \sim \mathcal{N}(0,1)$ et $Y \sim \chi_n^2$. La variable aléatoire $T = X/\sqrt{Y/n}$ suit une loi continue appelée loi de Student à n degrés de liberté :

$$T = X\sqrt{Y/n} \sim t_n$$

df

1. $E(T) = 0$
2. $Var(T) = \dfrac{n}{n-2}$ si $n > 2$

Densité

$$f(t) = \frac{\Gamma(\frac{n+1}{2})}{\sqrt{n\pi}\Gamma(\frac{n}{2})(1 + \frac{x^2}{n})^{\frac{n+1}{2}}}$$