1. Statistique et probabilité

Statistique et probabilité sont deux côtés d'une même pièce. Ces deux domaines des mathématiques s'occupent de décrire le résultat d'expériences ou d'observations faisant intervenir le hasard. Par exemple, tirer un bulletin dans une urne, choisir des arbres au hasard dans une forêt, sélectionner des bactéries au hasard dans un environnement, etc, ne peuvent se décrire que par la statistique et les probabilités. Pourquoi avoir deux noms ? Statistique et probabilité ne sont pas identiques dans la démarche bien que traitant des mêmes sujets. La statistique sert à observer et décrire le monde alors que les probabilités, aussi appelés théorie de la probabilité, tentent de l'expliquer théoriquement, mathématiquement. Ces deux approches se complètent mutuellement comme l'observation physique et le modèle mathématique se complètent.

1.1. Expérience aléatoire

Les expériences aléatoires le lancer un ou plusieurs dés, le tirage de cartes ou de numéros sont: on connaît les résultats possibles, mais on ne sait pas lequel va se produire avant l'expérience réalisée.

EXEMPLE 1

L'expérience de 150 lancers du même dé servira à illustrer la différence entre statistique et probabilité. Lors de cette expérience, la statistique décrit les propriétés statistiques de ces 150 lancers. Par exemple, la statistique montre que chaque face du dé est presque apparue autant de fois. La valeur moyenne se rapproche de 3,5. De manière complémentaire à cet aspect observationnel, la théorie de la probabilité prouve que si vous avez un dé non pipé, il y a autant de chance de tomber sur une face que sur une autre. Pour un nombre de 150 lancers, la théorie de la probabilité indique quelle est la chance d'obtenir la face 1 par exemple. Les prévisions de la théorie de la probabilité doivent se confirmer dans les résultats observationnels de la statistique. Si tel n'est pas le cas pour les lancers de dé, cela veut dire que le dé est pipé et qu'il faut changer les lois de probabilité (passer de la loi "chaque valeur a une chance sur 6 de sortir" à une loi plus compliquée qui va définir le dé pipé), à l'instar d'une expérience mettant en échec une théorie physique.

A) Théorème.

Dans une expérience dont chaque issue a la même probabilité, la probabilité d'un évènement est égale au quotient suivant :

$$\boxed{P = \frac{\text{Nombre d'issues favorables à l'évènement}}{\text{nombre total d'issues}}}$$

Exemple.

Lors d'un lancer de dés il y a 6 issues en tout et chacune a la même probabilité soit 1/6. On calcule la probabilité de l'évènement : « obtenir un multiple de 3 ». Il y a deux issues favorables à l'évènement : 3 et 6. La probabilité de cet évènement est égale à 2/6 soit 1/3

B) Probabilité expérimental

Dans des domaines très différents comme les domaines scientifique, sociologique ou médical, on s'intéresse à de nombreux phénomènes dans lesquels apparaît l'effet du hasard. Ces phénomènes sont caractérisés par le fait que les résultats des observations varient d'une expérience à l'autre. Une expérience est appelée "aléatoire" s'il est impossible de prévoir à l'avance son résultat et si, elle est répétée dans des conditions identiques, elle peut donner des résultats différents :

– jeu de pile ou face.

– observation de la durée de vie d'un individu anonyme dans une population ;

– observation de la durée de fonctionnement sans panne d'appareil ;

1.2. Notion de probabilité

1.2.1 univers

On appelle univers Ω associé à une expérience aléatoire, l'ensemble de tous les résultats possibles de cette expérience. Donnons quelques exemples :

1. On lance une pièce de monnaie. Pour l'ensemble Ω, on peut choisir soit Ω = { pile, face }, soit Ω = { pile, face, tranche }.

2. On s'intéresse à l'état de fonctionnement d'un système. Dans ce cas Ω = {0, 1} avec la convention 0 si le système est en panne et 1 s'il fonctionne.

5. On considère l'expérience aléatoire "durée de vie d'un individu". L'ensemble Ω est soit l'ensemble N, soit R + selon le procédé discontinu ou continu de cette mesure. Nous constatons que Ω peut être fini (exemples 1 et 2), dénombrable (exemples 3 et 5) ou non dénombrable (exemples 4 et 5). Lorsque Ω est fini ou dénombrable, on parle d'univers discret. Sinon on parle d'univers continu.

1.2.2 événement

Un événement aléatoire est une partie de l'ensemble des résultats possibles de l'expérience, c'est donc un sous-ensemble A de l'univers Ω. On dit que l'événement A est réalisé si le résultat ω de l'expérience appartient à A.

On sait que l'événement A est réalisé seulement une fois l'expérience aléatoire réalisée.

Exemples :

– Si l'on s'intéresse à l'événement suivant : "on a obtenu un chiffre pair lors d'un lancer d'un dé à 6 faces", on introduit A = {2, 4, 6}, qui est un sous-ensemble de Ω = {1, 2, 3, 4, 5, 6}.

– Si l'on s'intéresse à l'événement suivant : "la durée de vie d'une lampe est supérieure ou égale à 1000 heures", A = [1000, +∞[est un sous ensemble de Ω = R +. L'ensemble Ø est appelé l'événement impossible et Ω est appelé l'événement certain.

1.2.3 Opérations sur les événements

Les événements aléatoires étant des ensembles, introduisons les opérations ensemblistes classiques de la théorie des ensembles.

1° — On appelle événement contraire de A, noté A^C , le complémentaire de A dans Ω :

$A^C = \{\omega \in \Omega : \omega /\in A\}$. L'événement contraire A^C est réalisé si et seulement si A n'est pas réalisé.
Exemple : Si A est l'événement "la durée de vie du composant est supérieure ou égale à 1000 heures" : A = [1000, +∞[, l'événement contraire est l'événement "la durée de vie du composant est strictement inférieure à 1000 heures" : A^C = [0, 1000[.

2° — Soient A et B deux événements d'un univers Ω.

- L'événement "A et B" est celui qui est réalisé si A et B sont réalisés. C'est l'intersection $A \cap B = \{\omega \in \Omega : \omega \in A \text{ et } \omega \in B\}$.
- L'événement "A ou B" est celui qui est réalisé si l'un des deux est réalisé ou si les deux sont réalisés. C'est l'union $A \cup B = \{\omega \in \Omega : \omega \in A \text{ ou } \omega \in B\}$.
- L'inclusion $A \subset B$ signifie que l'événement A ne peut être réalisé sans que B le soit.

2. Rappel de quelques lois de probabilités

Une loi de probabilité est une distribution théorique de fréquences, se base sur la modélisation mathématiques qui permettent aux statisticiens de fabriquer des modèles pour décrire des phénomènes où le hasard intervient..
Soit Ω un ensemble muni d'une probabilité P. Une variable aléatoire X est une application définie sur Ω dans ℝ. X permet de transporter la loi P en la loi P' définie sur $\Omega'=X(\Omega)$: on a $P'(x_j)=P(X^{-1}(x_j))=P(X=x_j)$. La loi P' est appelée loi de X.

2.1. Variable aléatoire discréte

Une variable aléatoire discréte X est une application dont la valeur est la valeur du caractère étudié, c'est à dire le résultat d'une épreuve.
Si X prend n valeurs $x_1,...x–n$, on définit :

- $\bar{x}$ la moyenne ou, $E(X)$ l'espérance de X par :
  $\bar{x} = E(X) = \sum_{i=1}^{n} x_i P(X=x_i)$

  $\bar{x}$

- μ_2 le moment d'ordre 2 par :
  $\mu_2 = E(X^2) = \sum_{i=1}^{n} x_i^2 P(X=x_i))$
- $var(X)$ la variance de X par :
  $var(X) = \sum_{i=1}^{n} (x_i - \bar{x})^2 P(X=x_i) = E(X^2) - E(X)^2$
- σ l'écart type par : $\sigma = \sqrt{var(X)}$

Les n valeurs observées du caractère forment un échantillon de X d'ordre n : on dira que ces n valeurs sont les valeurs de n variables aléatoires $X_1,X_2,...,X_n$ qui suivent la même loi que X. Par exemple, lorsqu'on lance un dé, on peut définir la variable aléatoire X qui est égale à la valeur de la face visible, donc X vaut 1 ou 2 ou ... 6.
Il y a trop de paramètres en jeu pour pouvoir déterminer le résultat du lancer d'un dé, mais à chaque lancer la valeur de X est définie.

Attention
Ce n'est pas parce que deux variables aléatoires suivent la même loi qu'elles sont égales. Par exemple, je lance deux dés, un rouge et un vert : la variable X_1 égale à la face visible du dé rouge et la variable X_2 égale à la face visible du dé vert suivent toutes les deux une loi équirépartie de probabilité p=1/6 sur {1,2,3,4,5,6}.

2.1.1. La loi de Bernouilli

La variable aléatoire X suit une loi de Bernouilli de probabilité p, si X vaut 1 ou 0 avec les probabilités respectives p et $1-p$.
On a alors :

$E(X)=p$,

$E(X^2)=p$,

$\sigma(X)=\sqrt{p(1-p)}$.

2.1.2. La loi binomiale

Si la variable aléatoire X suit une loi binomiale $B(n,p)$, cela veut dire que X est égale au nombre de succès obtenus dans une série de n épreuves de Bernouilli de probabilité p. La variable aléatoire X peut donc prendre $n+1$ valeurs : $0,1,\ldots,n$.
La loi binomiale $B(n,p)$ est la somme de n variables de Bernouilli indépendantes.
On a :

$Proba(X=k)=C_n^k\ p^k(1-p)^{n-k}$, pour $0 \leq k \leq n$,

$E(X)=np$,

$\sigma(X)=\sqrt{np(1-p)}$.

Example 1
On fabrique des pièces et on suppose que la probabilité pour qu'une pièce soit défectueuse est p=0.05 et donc il y a un contrôle de ces pièces.
Soit X la variable alèatoire égale au nombre de défectueuses trouvées lors d'un contrôle de n=1000 pièces Déterminer la loi de X ainsi que son espérance et son écart-type.
Ici X suit une loi binomiale $B(n,p)$ de probabilité p.
$E(X)=np=50$
$\sigma(X)=\sqrt{np(1-p)}=6.89202437604$

2.1.3. La loi de Poisson

La variable aléatoire X suit une loi de Poisson $P(\lambda)$ de paramètre λ ($\lambda \geq 0$) si :
- X a pour valeurs les entiers naturels,
- Prob$(X=k)=e^{-\lambda}\lambda^k/k!$, pour $k\in \mathbb{N}$.

On a :
$E(X)=\lambda$
$\sigma(X)=\sqrt{\lambda}$

Exemple 2
Soit une variable aléatoire X qui vérifie pour $\lambda \geq 0$ et pour tout entier $n \geq 1$:
$\text{Prob}(X=n)=\lambda/n\text{Prob}(X=n-1)$
Montrer que X suit une loi de Poisson.
On cherche pour k entier strictement positif :
$\text{Prob}(X=k)=\lambda/k\text{Prob}(X=k-1)=\ldots\ \lambda^k/(k)!\text{Prob}(X=0)$.
Donc :

$$\text{Prob}(X=k)=\frac{\lambda^k}{(k)!}\text{Prob}(X=0)$$

On tape :
```
sum(lambda^k/k!,k=0..inf)
```
On obtient :
```
exp(lambda)
```
On doit avoir :
$\sum_{k=0}^{+\infty}\text{Prob}(X=k)=1$
Donc on a le relation :
$\text{Prob}(X=0)*\sum_{k=0}^{+\infty}\lambda^k/k!=\exp(\lambda)*\text{Prob}(X=0)=1$
c'est à dire :
$\text{Prob}(X=0)=\exp(-\lambda)$
Donc on a bien :
$\text{Prob}(X=k)=e^{-\lambda}\lambda^k/k!$, pour $k\in \mathbb{N}$

Exemple 3

Soient X et Y deux variables aléatoires indépendantes qui suivent une loi de Poisson :
X suit une loi de Poisson $P(\lambda_1)$ de paramètre λ_1 ($\lambda_1 \geq 0$) et
Y suit une loi de Poisson $P(\lambda_2)$ de paramètre λ_2 ($\lambda_2 \geq 0$).
Déterminer la loi de la variable aléatoire $Z=X+Y$.
On a :
$\text{Prob}(X=k)=e^{-\lambda_1}\lambda_1^k/k!$, pour $k\in \mathbb{N}$
$\text{Prob}(X=k)=e^{-\lambda_2}\lambda_2^k/k!$, pour $k\in \mathbb{N}$
Donc :
$\text{Prob}(Z=n)=\text{Prob}(X+Y=n)=\sum_{k=0}^{n}\text{Prob}(X=k)*\text{Prob}(Y=n-k)=\sum_{k=0}^{n}e^{-\lambda_1}\lambda_1^k/k!*e^{-\lambda_2}\lambda_2^{n-k}/(n-k)!$
On sait que : $(\lambda_1+\lambda_2)^n=\sum_{k=0}^{n}\lambda_1^k\lambda_2^{n-k}n!/k!(n-k)!$
Donc :
$\text{Prob}(Z=n)=e^{-(\lambda_1+\lambda_2)}(\lambda_1+\lambda_2)^n/n!$
Donc Z suit une loi de Poisson de paramètre $\lambda_1+\lambda_2$.

2.2. Variable aléatoire absolument continue

Définitions

Variable aléatoire absolument continue
Une variable aléatoire X est absolument continue si il existe $f(x)$ telle que sa fonction de répartition $F(x)$ est égale à :

$$\mathrm{Prob}(X \leq x) = F(x) = \int_{-\infty}^{x} f(t)dt$$

Densité de probabilité
La fonction $f(x)$ est appelée densite de probabilité et on a :

$$f(x) = F'(x)$$

Espérance mathématique
L'espérance mathématique ou moyenne de x est :

$$E(X) = \overline{X} = \int_{-\infty}^{+\infty} t * f(t)dt$$

Variance et Ecart type
La variance de X est :

$$V(X) = \int_{-\infty}^{+\infty} (t - E(X))^2 f(t)dt$$

L'ecart type de X est :

$$\sigma(X) = \sqrt{V(X)}$$

2.2.1. La loi Normale ou loi de Gauss

La variable aléatoire X suit une loi Normale ou loi de Gauss de paramètres μ,σ ($\sigma > 0$) si :
- X a pour valeurs tous les réels,
- $Prob(a \leq X < b) = \int_a^b f(t)\, dt$ où $f(x) = 1/\sigma\sqrt{2\pi}e^{-1/2(x-\mu/\sigma)2}$ (f est la densité de probabilité et a comme représentation graphique une courbe en cloche).

On note cette loi $N(\mu,\sigma)$.
On a :
$E(X) = \mu$
$\sigma(X) = \sigma$.
On dit que $N(0,1)$ est la loi normale centrée réduite. Si X suit la loi $N(0,1)$ alors :

$Prob(a \leq X<b)=\int_a^b =1/\sqrt{2\pi}e^{-t2/2}dt$
Si X suit la loi $N(\mu,\sigma)$ alors $X-\mu/\sigma$ suit la loi $N(0,1)$.
On a des tables où on peut lire que :
$P(|X-\mu|/\sigma >1.96)=0.05$,
$P(|X-\mu|/\sigma >2.58)=0.01$,
$P(|X-\mu|/\sigma >3.1)=0.001$,
et on a $P(|X-\mu|/\sigma>t)=1-2\int_0^t f(x)dx$.

Théorèmes

On montre qu'une loi binomiale $B(n,p)$ peut être approchée :

- par la loi normale $N(np,\sqrt{np(1-p)})$ si $np>15$ et $n(1-p)>15$.
  Cela veut dire que pour tout entier k :
  $C_n^k p^k(1-p)^{n-k}$ est proche de $1/\sqrt{2np(1-p)\pi}e^{-(k-np)2/2np(1-p)}$ quand n est grand.
  Exemple :
  On a $p(k)=C_{100}^k 0.2^k 0.8^{100-k}$ est proche de $f(k)=1/4\sqrt{2\pi}e^{-(k-20)2/32}$.
  Ainsi si X suit la loi $B(n,p)$ et Y suit la loi $N(np,\sqrt{np(1-p)})$ alors :
  $Proba(X=x)$ sera approché par $Proba(x-0.5<Y<x+0.5)$
  $Proba(X<x)$ sera approché par $Proba(Y<x-0.5)$
  $Proba(X\leq x)$ sera approché par $Proba(Y<x+0.5)$

- par la loi de Poisson $P(np)$ si $np \leq 10$, $p\leq 0.1$ et $n \geq 15$.
  Exemple :
  On a $p(k)=C_{50}^k 0.04^k 0.96^{50-k}$ est proche de $e^{-2}2^k/k!$.