STATISTIQUE APPLIQUEE

(Outils d'aide à la décision)

Rachid MCHICH

Chap. 1: Rappels mathématiques

III- Probabilités

Définition : La probabilité est une évaluation du caractère probable d'un événement.

Elle donne lieu à des applications nombreuses dans plusieurs domaines comme : la météorologie, les statistiques, la finance, la théorie des jeux (maths, économie, micro-économie…), les mathématiques financières (les cours de bourse…)…etc.

III-1 Dénombrement - Théorie des ensembles

- Deux ensembles sont **égaux** :

$$A = B \Leftrightarrow (x \in B \Leftrightarrow x \in A)$$

- Un ensemble A est une **partie** ou un **sous-ensemble** d'un ensemble B si :

$$A \subset B \Leftrightarrow (\forall x \in A \Rightarrow x \in B)$$

- L'ensemble vide, noté $\varnothing$ est le seul ensemble qui ne contient aucun élément.

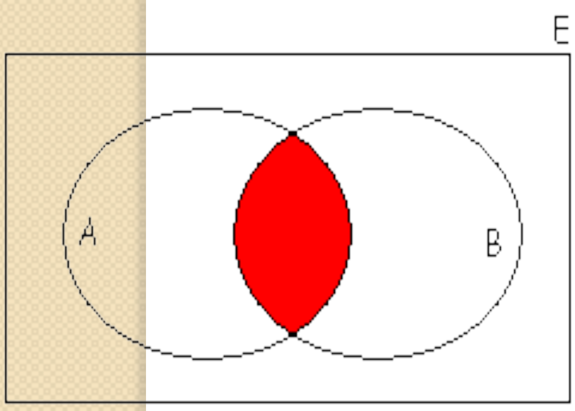
- $A \cap B = \{x \,/\, (x \in A) \text{ et } (x \in B)\}$

- Deux ensembles sont disjoints si leur intersection est l'ensemble vide $\varnothing$.

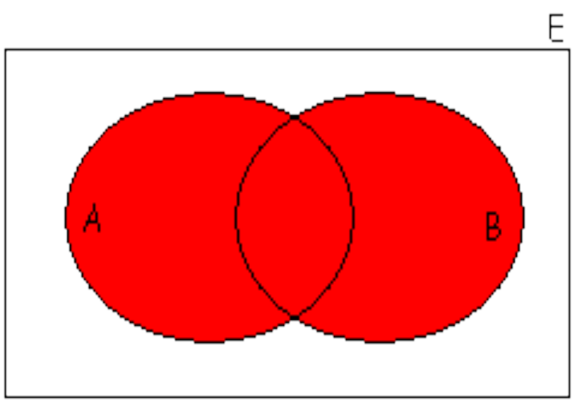
- $A \cup B = \{x \,/\, ((x \in A) \text{ ou } (x \in B)\}$

- $A \setminus B = \{(x \in A) \text{ et } (x \notin B)\}$.

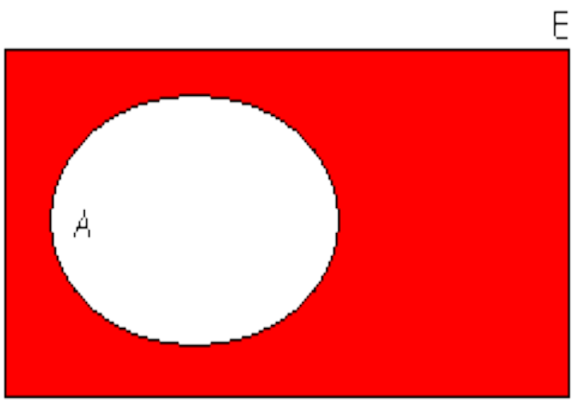
- $\bar{A} = \{x \,/\, x \notin A\}$. On note aussi ${}^{C}A$ ou $C_E{}^{A}$.



$A \cap B$



$A \cup B$



$\bar{A}$

Propriétés :

- $A \cup B = B \cup A$; $A \cup \emptyset = A$
- $A \cap B = B \cap A$; $A \cap \emptyset = \emptyset$
- $A \cup (B \cap C) = (A \cup B) \cap (A \cup C)$
- $A \cap (B \cup C) = (A \cap B) \cup (A \cap C)$
- $\overline{A \cap B} = \overline{A} \cup \overline{B}$
- $\overline{A \cup B} = \overline{A} \cap \overline{B}$

Notion de cardinal

Si E a un nombre fini d'éléments, alors pour tout sous-ensemble A de E, A a également un nombre fini d'éléments, noté: **card(A)**. On a alors les propriétés suivantes:

- $Card(\overline{A}) = card(E) - card(A)$
- $Card(A \cup B) = card(A) + card(B) - card(A \cap B)$
- $Card(A \setminus B) = card(A) - card(A \cap B)$
- $card \emptyset = 0$

III-2 Analyse combinatoire

<u>Définition:</u>

L'**analyse combinatoire** est un ensemble de méthodes permettant de donner le nombre de tous les résultats possibles d'une expérience donnée.

En d'autre termes, l'analyse combinatoire permet de **dénombrer** les différentes dispositions (ou sous-ensembles) que l'on peut former à partir d'un ensemble fini d'éléments.

Calcul pratique

Dispositions	Sans répétition	Avec répétition
Arrangements (ordonnées)	$A_n^p = \dfrac{n!}{(n-p)!}$	n^p
Combinaisons (non ordonnées)	$C_n^p = \dfrac{n!}{p!(n-p)!}$	$K_n^p = \dfrac{(n-1+p)!}{p!(n-1)!}$

III-3 Calcul des probabilités

Vocabulaire :

- Une ***expérience aléatoire*** est une action ou une expérience dont le résultat ne peut être prévu avec certitude (lancement d'un dé, tirage d'une carte…).
- Pour définir une expérience aléatoire, il faudra recenser l'ensemble de ses résultats possibles, appelés : ***éventualités*** ou ***événements élémentaires***.
- L'ensemble de ces éventualités est appelé : ***univers*** et sera noté en général : Ω.

- Un ***événement*** est un ensemble d'éventualités noté A, c'est donc aussi une partie de Ω (Ω est l'événement certain et $\varnothing$ est l'événement impossible).

Probabilité d'un événement :

Si Ω est un univers et ***A*** un événement, la question que l'on peut se poser est de **définir** ou de **mesurer** la chance qu'a l'événement ***A*** de se réaliser : on parlera alors de « **probabilité de *A*** ».

La complexité de cette définition dépendra de celle de Ω : **fini**, **infini dénombrable** ou **infini non dénombrable**.

Probabilités conditionnelles – Notion d'indépendance :

Souvent, la probabilité d'un événement est influencée par le fait qu'un événement, lié au premier, se soit produit. On parle alors de **probabilité conditionnelle**.

Définition :

Soient Ω un univers et A et B deux événements avec $p(A) \succ 0$.

La probabilité que l'événement B soit réalisé sachant que A l'est déjà, est appelée: **probabilité conditionnelle de *B* sachant *A***, et est définie par :

$$p(B / A) = \frac{p(A \cap B)}{p(A)}$$

Exemple :

Considérons les possibilités de promotion professionnelle dans une grande multinationale qui compte 1200 cadres et où, au cours des deux dernières années, 36 cadres femmes ont été promues sur un total de 240 et 288 cadres hommes ont été promus sur un total de 960.

On voudrait savoir s'il y a eu discrimination lors des différentes promotions.

- Un responsable d'une grande surface sait que 80% des clients payent par carte de crédit. Quelle est la probabilité que les deux prochains clients payent par carte de crédit?

- *(Une expérience aléatoire est déterminée par l'expérience que l'on effectue et donc l'univers aussi, c.à.d. que si l'on change d'expérience aléatoire, on change d'univers aussi).* Exemple des tirages successifs de deux cartes d'un jeu de cartes de 32: quelle est la probabilité que la deuxième carte soit un *Roi* sachant que la première carte est un *Roi* ?

Théorème des probabilités totales :

Soient $A_1, A_2, \ldots, A_n$ des événements tels que :

$$A_1 \cup A_2 \cup \ldots \cup A_n = \Omega$$

$$et$$

$$A_i \cap A_j = \varnothing$$

Alors on a :

$$p(B) = \sum_{i=1}^{n} p(B / A_i).p(A_i)$$

Exemple : Deux machines M_1 et M_2 fabriquent des chemises. Elles produisent respectivement 1/3 et 2/3 de la production totale. M_1 sort 5% de chemises avec défaut. M_2 en sort 6%. Soient les événements suivants :

A : la chemise est fabriquée par M_1

B : la chemise est fabriquée par M_2

C : la chemise contient un défaut

1. Quelle est la probabilité que la chemise soit fabriquée par M_1 ?
2. On tire une chemise de la production de M_1. Quelle est la probabilité qu'elle contienne un défaut ?
3. On tire une chemise de la production. Quelle est la probabilité qu'elle provienne de M_1 et qu'elle soit défectueuse ?
4. On tire une chemise de la production. Quelle est la probabilité qu'elle soit défectueuse ?
5. Quelle est la probabilité qu'une pièce défectueuse soit fabriquée par M_1?

Théorème de Bayes :

Soient $A_1, A_2, \ldots, A_n$ des événements tels que :

$$A_1 \cup A_2 \cup \ldots \cup A_n = \Omega$$

$$et$$

$$A_i \cap A_j = \varnothing$$

Alors on a :

$$p(A_i / B) = \frac{p(A_i)p(B / A_i)}{\sum_{i=1}^{n} p(A_i)p(B / A_i)}$$

Exemple : Une entreprise utilise 3 machines différentes A, B et C pour fabriquer des pièces. 15 000 pièces sont fabriquée par A, 35 000 par B et 50 000 par C. La machine A produit 5% de pièces défectueuses, la machine B 2% et la machine C 3%. Soient les événements suivants :

A : pièce produite par A

B : pièce produite par B

C : pièce produite par C

D : pièce défectueuse

1. On prélève une pièce au hasard. Quelle est la probabilité qu'elle soit défectueuse ?
2. On prélève une pièce défectueuse. Quelle est la probabilité qu'elle soit fabriquée par A? (ou par B ou par C?)

IV - Variables aléatoire et lois de probabilité

IV.1 Notion de variable aléatoire :

Une notion importante en probabilités est la notion de **variable aléatoire** : c'est une fonction qui dépend d'une expérience aléatoire; c.à.d. : $X : \Omega \longrightarrow IR$

Ainsi, à chaque événement, on peut associer une valeur valeur x_i dans *IR*.

- Une variable aléatoire est **discrète** si l'ensemble de ses valeurs est fini ou infini dénombrable.

- Une variable aléatoire **continue** est une v. a. *X* dont l'ensemble des valeurs est *IR* ou un intervalle de *IR*.

IV.2 - V. A. D. et lois de probabilités discrètes :

Distributions (lois) de probabilité discrètes

Définition : Soit X une v. a. d. X. On peut associer à chaque x_i de X, une probabilité p_i ($x_i \rightarrow p_i$) . Cette correspondance s'appelle : **loi de probabilité** de *X*.

- C'est une fonction *f(x)* qui décrit comment sont distribuées les probabilités en fonction des valeurs de la v. a. d.

- Une fonction de probabilité d'une v. a. d. doit satisfaire les deux conditions suivantes:

  - $f(x_i) = p_i \geq 0$

  - $\sum_{i=1}^{n} f(x_i) = \sum_{i=1}^{n} p_i = 1$

Définition: On appelle **fonction de répartition** de la v. a. X, l'application :

$$F : IR \to [0,1]$$
$$x \mapsto F(x) = P(X \leq x)$$

Caractéristiques d'une v. a. d. :

- Espérance : $E(X) = \mu = \sum_{i=1}^{n} p_i x_i$

- Variance : $V(X) = \sigma^2 = \sum_{i=1}^{n} p_i (x_i - \mu)^2$

$$= \sum_{i=1}^{n} p_i x_i^2 - \mu^2$$

- Ecart type : $\sigma(X) = \sqrt{V(X)}$

Exemple : Les données sur les ventes journalières chez un concessionnaire, pendant 300 jours, sont les suivantes:

Nombre de voitures vendues par jour	Nombre de jours
0	54
1	117
2	72
3	42
4	12
5	3

Expérience aléatoire : Choisir une journée parmi les 300 jours.

Variable aléatoire : Nombre de voitures vendues au cours de cette journée.

Lois de probabilité discrètes usuelles

Loi uniforme :

- Une distribution de probabilité suit une loi uniforme lorsque toutes les valeurs prises par la v. a. sont équiprobables.

- Si n est le nombre de valeurs différentes prises par la v. a., alors :

$$\forall i, \qquad P(X = x_i) = \frac{1}{n}$$

Loi de Bernoulli :

- On appelle variable de Bernoulli (ou variable indicatrice), la v. a. d. X tq :

$$X : \Omega \to IR \quad où : \quad X(\Omega) = \{0, 1\}$$

- La loi de probabilité associée à la variable de Bernoulli X tq :

$$P(X=1) = p \text{ et } P(X=0) = q \text{ (avec } p+q=1)$$

est appelée loi de Bernoulli, notée :

$$B(1, p)$$

Loi binomiale :

- La variable binomiale représente le nombre de succès obtenus lors de la répétition de n épreuves identiques et indépendantes, chaque épreuve ne pouvant donner que 2 résultats possibles. On note une telle variable S_n et elle est donnée par :

$$S_n = X_1 + X_2 + \ldots + X_n$$

Où X_i est une variable de Bernoulli.

- Ainsi, la loi de probabilité suivie par la somme de n variables de Bernoulli et où la probabilité associée au succès est p, est la loi binomiale de paramètres n et p, notée : $B(n,p)$.

Propriétés :

- Le nombre de résultats de *x* succès pendant *n* tirages :

$$C_n^x = \frac{n!}{x!(n-x)!}$$

- Fonction de probabilité binomiale:

$$f(x) = C_n^x p^x (1-p)^{(n-x)}$$

Caractéristiques (v. a. binomiale) :

Espérance : $E(X) = np$

Variance : $V(X) = np(1-p)$

Exemple : Considérons le comportement d'achat des 3 prochains clients d'un magasin. Si la probabilité qu'un client fasse un achat est de 0,30.

1. Quelle est la probabilité que deux des 3 prochains clients fassent un achat?
2. Quelle est la probabilité que, sur les 3 prochains clients, aucun client ne fasse un achat, qu'un client ou 3 fassent un achat?
3. Quelle est la probabilité que, sur les 10 prochains clients, 4 fassent un achat?

Loi de poisson :

- La variable aléatoire discrète de poisson est utilisée pour décrire le nombre d'occurrences d'un événement au cours d'un intervalle de temps ou d'espace bien défini.

- Elle est utilisée en général pour modéliser les taux d'arrivée dans une file d'attente.

- Propriétés :

  1. La probabilité d'occurrence est la même dans deux intervalles de même longueur

  2. L'occurrence ou non d'un événement dans un intervalle est indépendante de l'occurrence ou non de cet événement dans un autre intervalle

Exemple : Supposons qu'on s'intéresse au nombre d'arrivées au guichet d'une banque dans un intervalle de 15 minutes, le matin, en semaine. Si l'on suppose que la probabilité d'une arrivée est la même pour 2 intervalles de longueurs égales et que l'arrivée (ou non) est indépendante entre deux intervalles de temps.

Supposons qu'une analyse de données révèle que le nombre moyen d'arrivées pendant 15 minutes est de 10.

1. Quelle est la probabilité de 5 arrivées en 15 minutes?
2. Quel est l'écart type du nombre d'arrivées?
3. Quelle est la probabilité d'une arrivée en 3 minutes?

Loi hypergéométrique :

On considère *r* éléments de la population de taille *N* qui sont des succès et (*N* - *r*) éléments considérés comme des échecs.

La **fonction de probabilité hypergéométrique** est utilisée pour calculer la probabilité que, dans un échantillon de *n* éléments sélectionnés aléatoirement sans remise, nous obtenions *x* succès et $n - x$ échecs

Exemple : Problème de contrôle de qualité.

Des fusibles électriques sont conditionnés par boîtes de 12. On sélectionne aléatoirement 3 des 12 fusibles d'une boîte pour les tester. Supposons que la boîte contient 5 fusibles défectueux.

1. Quelle est la probabilité qu'on trouve exactement un fusible défectueux parmi les 3 sélectionnés au hasard?
2. Quelle est la probabilité qu'on trouve au moins un fusible défectueux parmi les 3 sélectionnés au hasard?
3. Calculer l'espérance et l'écart type.

III-3.2 V. A. C. et lois de probabilités continues :

Variables aléatoires continues :

- Une variable aléatoire **continue** est une v. a. *X* dont l'ensemble des valeurs est *IR* ou un intervalle de *IR*. Une telle variable est en général définie par sa **fonction de répartition** :

$$F(x) = P(X \leq x), \quad \forall x \in IR$$

- On désigne par fonction densité de probabilité *f*, la fonction dérivée de la fonction de répartition *F* :

$$f = F'$$

ou encore :

$$F(x) = \int_{-\infty}^{x} f(y)\,dy$$

- La fonction densité $f(x)$ pour une v. a. c. est l'équivalent de $f(x_i)$ pour une v. a. d.; sauf que la fonction densité ne fournit pas directement les probabilités.

- Cependant, l'aire située sous le graphique de $f(x)$ dans un intervalle particulier, donne la probabilité que la v. a. c. X prenne une valeur dans cet intervalle.

- La probabilité que la v. a. c. prenne une valeur particulière est nulle (car $\int_a^a f(x)dx = 0$).

- <u>Caractéristiques d'une v. a. c. :</u>

Espérance :

$$E(X)=\int_{-\infty}^{+\infty} xf(x)\,dx$$

Variance :

$$V(X)=\int_{-\infty}^{+\infty} x^2 f(x)\,dx-(E(X))^2$$

Ecart type :

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

Loi normale centrée réduite :

- Une v. a. c. *X* suit une loi normale centrée réduite si sa densité de probabilité est donnée par :

$$\forall x \in IR, \quad f(x) = \frac{1}{\sqrt{2\pi}} e^{-\frac{x^2}{2}} \quad ; (\mu = 0, \sigma = 1)$$

- On note une telle loi : $N(0,1)$.

Approximation normale des probabilités binomiales :

Pour une loi binomiale, lorsque le nombre de tirages devient grand, la fonction de probabilité binomiale devient difficile à calculer. Ainsi, dans le cas où :

$$np \geq 5 \quad et \quad n(1-p) \geq 5$$

la loi normale permet d'estimer facilement des probabilités binomiales, en posant:

$$\mu = np \quad et \quad \sigma = \sqrt{np(1-p)}$$

Loi exponentielle :

- Une v. a. c. X suit une loi exponentielle si sa densité de probabilité est donnée par :

$$f(x) = \frac{1}{\mu} e^{-\frac{x}{\mu}} \quad pour \quad x > 0 \, , \mu > 0$$