+

Ministère de l'Enseignement Supérieur et de la Recherche Scientifique

Université 8 Mai 1945 Guelma

Faculté des Mathématiques et de l'Informatique et des Sciences de la Matière

Département de Mathématiques

Polycopié du cours

Dr : SEGNI Sami

Statistique et Probabilités

Deuxième Année Licence SNV

2015/2016

Table des matières

Chapitre 1

Rappels sur les statistiques descriptives

1.1 Statistiques à une variable

1.1.1 Vocabulaire de la statistique

Un ensemble d'objets ou de personnes d'une étude statistique est appelé **population**. Un élément de cette population est appelé **individu**.

L'étude statistique porte sur un caractère. Si le caractère est **quantitatif**, les mesures sont alors les valeurs d'une **variable statistique** (un age, une taille...).

La variable est dite **discrète** si elle ne prend que des valeurs isolées (entières). Elle est **continue** si elle peut prendre toutes les valeurs d'un intervalle.

L'**effectif** d'une population est le nombre d'individus total de cette population. La **fréquence absolue** d'un caractère est le nombre d'individus possédant ce caractère. La **fréquence relative** d'un caractère est le nombre d'individus possédant ce caractère divisé par l'effectif total de la population.

1.1.2 Les variables discrètes

Représentation

On représente les variables discrètes sous forme d'histogramme grâce aux différentes fréquences.

Crarctéristiques

(1) La moyenne

Soit n valeurs distinctes ou non d'une variable. Si cette variable prend p valeurs disinctes $(p = n)$, $x_1, \ ..., \ x_p$, d'effectifs respectifs $n_1, \ ..., \ n_p$ alors la moyenne est donnée par

$$\overline{X} = \frac{1}{n}\sum_{i=1}^{p} n_i x_i.$$

(2) Propriété

Si pour tout i on peut opérer un changement de variable affine du type : $y_i = ax_i + b$ alors :

$$\overline{Y} = a\overline{X} + b.$$

(3) La variance

Elle est donnée par la formule

$$var(X) = \frac{1}{n}\sum_{i=1}^{p} n_i \left(x_i - \overline{X}\right)^2.$$

(4) L'écart-type

il est donné par la formule

$$\sigma_x = \sqrt{var\left(X\right)}.$$

(5) Propriétés

La formule suivante est plus pratique pour le calcul de la variance :

$$var(X) = \left(\frac{1}{n} \sum_{i=1}^{n} n_i x_i^2 \right) - \overline{X}^2.$$

De plus si pour tout i, $y_i = ax_i + b$ alors $var\,(y) = a^2 var\,(x)$ et $\sigma_y = |a|\,\sigma_x$.

1.1.3 Les variables continues

Représentation Pour leur représentation, on regroupe en général dans des classes adjacentes d'amplitudes pas forcément égales. Ceci est représenté dans le tableau ci-dessous :

classes	$[x_0, x_1[$	$[x_1, x_2[$	...	$[x_{p-1}, x_p[$
centre des classes	c_1	c_2	...	c_p
effectifs	n_1	n_2	...	n_p
fréquences	n_1/n	n_2/n	...	n_p/n

La représentation s'effectue alors grâce à un histogramme dont les rectangles sont de largeur l'amplitude de la classe et dont **l'aire est proportionnelle à l'effectif.**

Caractéristiques Pour calculer moyenne et écart-type, on prend les formules connues avec les c_i centres des classes c'est à dire :

$$c_i = \frac{x_{i-1} + x_i}{2}.$$

Exercice

Un tour automatique produit des axes cylindriques. Les diamètres en $\frac{1}{10}mm$ mesurées sur un lot de 1000 pièces ont donné les résultats suivants :

Classes	[244;246[	[246;248[	[248;249[	[249;250[	[250;251[	[251;252[	[252;254[	[254;258[
Effectifs	11	132	152	200	194	158	139	14

1- Tracer un histogramme du caractère X.

2- Calculer la moyenne ainsi que l'écart-type.

1.2 Statistiques à deux variables

1.2.1 Tableau des données, nuage de points

On observe que dans certains cas, il semble exister un lien entre deux caractères d'une population (entre poids et taille, entre l'épaisseur d'un mur et sa résistance thermique, ...).

On définit alors une série statistisque à deux variables x et y, prenant des valeurs $x_1,\ x_2, ...,\ x_n$ et $y_1,\ y_2, ..., y_n$.

Tableau de données

Le nom est explicite. Il représente les différentes valeurs de x et y dans un tableau à deux entrées.

x en mm	2	4	6	8	10	12	15	20
y en $m^2 {}^\circ C$	0.83	1.34	1.63	2.29	2.44	2.93	4.06	4.48

Nuage de points

Le plan P étant muni d'un repère *orthogonal*, on peut associer au couple (x_i, y_i) de la série statistique double, le point M_i de coordonnées x_i et y_i. L'ensemble des points M_i obtenus constitue le nuage de points représentant la séris statistique.

1.2.2 Ajustement affine par la méthode des moindres carrées

La droite de régression

Soit D une droite d'ajustement. Soit $M_i = (x_i, y_i)$ un point du nuage. $P_i$est le point de même abscisse x_i que M_i situé sur la droite D d'équation $y = ax + b$. Q_i est le point de même coordonnée y_i que M_i situé sur la droite D' d'équation $x = a'y + b'$.

On appelle droite de régression de y en x, la droite D telle que $\sum_{i=1}^{n} M_iP_i^2 = \sum_{i=1}^{n} [y_i - (ax_i + b)]^2$ soit minimale.

On appelle droite de régression de x en y, la droite D' telle que $\sum_{i=1}^{n} M_iQ_i^2 = \sum_{i=1}^{n} [x_i - (a'y_i + b')]^2$ soit minimale.

Covariance d'une série statistique double

C'est le nombre $cov\,(x, y) = \sigma_{xy} = \frac{1}{n}\sum_{i=1}^{n} \left(x_i - \overline{X}\right)\left(y_i - \overline{Y}\right)$. Pratiquement on utilise la formule $\sigma_{xy} = \frac{1}{n}\sum_{i=1}^{n} x_iy_i - \overline{XY}$.

Equations des droites de régression

On montre que la droite de régression D de y en x pour équation $y = ax + b$ avec $a = \frac{\sigma_{xy}}{\sigma_x^2}$ et $b = \overline{Y} - a\overline{X}$.

De même on montre que la droite de régression D' de x en y a pour équation $x = a'y+b'$ avec $a' = \frac{\sigma_{xy}}{\sigma_y^2}$ et $b' = \overline{X} - a'\overline{Y}$.

1.2.3 Coefficient de corrélation

C'est le nombre $r = \frac{\sigma_{xy}}{\sigma_x\sigma_y}$ et vérifie $r \in [-1, 1]$.

Les nombres $r,\ \sigma_{xy},\ a,\ a',\ b,\ b'$ sont du même signe.

Si $r^2 = 1$ alors $aa' = 1$. Les droites D et D' sont alors confondues ; on dit alors que l'ajustement affine est parfait.

Si $0.7 < |r| < 1$ alors les deux droites sont proches l'une de l'autre (on fait l'angle entre les deux est inférieur à $45°$) ; on dit que l'ajustement affine est justifié.

Si $|r| < 0.7$, alors l'angle entre les deux droites est supérieur à $45°$. L'ajustement affine ne se justifie pas.

Chapitre 2

Introduction au calcul des probabilités

2.1 L'analyse combinatoire

Définition

L'analyse combinatoire a pour objet de dénombrer les manières distinctes de grouper tout ou partie des objets d'un ensemble, suivant des lois déterminées.

2.1.1 Notions élémentaires

Définition

Le nombre d'éléments d'un ensemble E, $n < +\infty$ *est appelé cardinal de* E, *qu'on note* $card\,(E)$ *ou* $|E|$.

Il est évident que le nombre d'éléments de l'ensemble vide est égal à 0 $(card\,(\phi) = 0)$.

Propriétés

1. Soit E ensemble tel que $card\,(E) = n$, alors il existe 2^n sous-ensembles de E.

2. Soit A et B deux sous-ensembles d'un ensemble fini E. Ainsi si :

$$A \cap B = \phi \Longrightarrow card\,(A \cup B) = card\,(A) + card\,(B)\,.$$

En particulier pour $\overline{A} = \complement_E^A$ on sait que $\overline{A} \cap A = \phi$, d'où :

$$card\,(E) = card\left(A \cup \overline{A}\right) = card\,(A) + card\left(\overline{A}\right),$$

donc

$$card\left(\overline{A}\right) = card\,(E) - card\,(A)\,.$$

Et généralement :

$$card\,(A \cup B) = card\,(A) + card\,(B) - card\,(A \cap B)\,.$$

On accepte que :

$$card\,(A \times B) = card\,(A)\,.card\,(B)\,.$$

Et en particulier si $A = B$:

$$card\left(A^k\right) = (card\,(A))^k\,,\ k \geq 1.$$

3. Soient $E,\ F$ deux ensembles tels que $card\,(E) = n$ et $card\,(F) = p$. Une application f de E vers F, notée $f : E \longrightarrow F$ est ene relation, telle que tout élément x de E possède une image unique $f\,(x)$ de F.

Ainsi on peut dire que le nombre d'ablications possibles définies de E vers F est de p^n.

Arrangements et permutations

Définition (arrangement)

Etant donnés n objets distincts, on appelle arrangement de ces n objets pris p à p, *tout groupe formé par p objets choisis dans un certain ordre parmi les n objets.*

Conséquence

Le nombre A_n^p d'arrangements de p objets pris parmi n, qui est le nombre d'applications injectives d'un ensemble de p éléments dans un ensemble de n éléments, s'exprime pour $p \leq n$ par :

$$A_n^p = n(n-1)(n-2)...(n-(p-1)) = n(n-1)(n-2)...(n-p+1).$$

Exemple

Un pari à l'arriver d'un tiercé de 10 chevaux est A_{10}^3.

Définition (Permutation)

Les permutations de n objets sont définies comme toutes les suites ordonnées possibles obtenues en prenant ces objets.

Le nombre P_n des permutations de n objets sera A_n^n d'où :

$$P_n = n(n-1)(n-2)...3.2.1 = n!$$

En convenant que $0! = 1$.

Pour les grandes valeurs de n, une bonne approximation de $n!$ est donnée par la formule de Stirling :

$$n! \sim \sqrt{2\pi n}\left(\frac{n}{e}\right)^n.$$

P_n est le nombre de bijections d'un ensemble de n éléments dans lui-même. D'où

$$A_n^p = n(n-1)(n-2)...(n-p+1) = \frac{n!}{(n-p)!}.$$

Exemple

Trouver le nombre d'anagrammes du mot NOEL.

Les anagrammes sont les mots formés avec les mêmes lettres (qu'ils aient ou non un sens).

Ainsi il y a $P_4 = 4!$ façons de former des mots à partir de NOEL.

Définition (Permutation avec répétition)

Soient n objets parmi lesquels k objets sont identiques, les $n-k$ étant distincts l'un de l'autre, et des k premiers objets, $P_{n,k}$ le nombre de permutations de ces n objets est la répartition quelconque de n objets moins le nombre de permutations de k objets.

Ainsi :

$$P_{n,k} = \frac{n!}{k!}.$$

D'une manière générale si parmi les n objets on peut distinguer :

k_1 objets d'un premier type et identiques entres eux,

k_2 objets d'un second type et identiques entres eux, etc...

.

.

.

k_r objets d'un $r^{ième}$ type et identiques entres eux, avec $n = k_1 + k_2 + ... + k_r$, le nombre de permutation de ces n objets est

$$P_{n,k_1,...,k_r} = \frac{n!}{k_1!k_2!...k_r!}.$$

Exemple

Trouver le nombre d'anagrammes du mot LENTILLE

Définition (Combinaisons)

Une combinaison de n objets p à p est tout groupe qu'on peut former en choisissant p objets parmi les n objets, sans considération d'ordre.

$$A_n^P = C_n^p.p! \Longrightarrow C_n^p = \frac{A_n^P}{p!}.$$

Propriétés (du C_n^P)

$$C_n^n = 1, C_n^0 = 1, C_n^1 = 1, C_n^{n-p} = \frac{n!}{(n-p)!p!} = C_n^p, C_{n-1}^p + C_{n-1}^{p-1} = C_n^p.$$

Exemple

Combien de comités différents peut-ils être formés à partir de 3 personnes sur la base de 8?

Binôme de Newton

Les identités remarquables : $(a+b)^2 = a^2 + 2ab + b^2$; $(a+b)^3 = a^3 + 3ab^2 + 3a^2b + a^3$, peuvent être généralisée grâce à la notion de combinaison où

$$\begin{aligned}(a+b)^p &= C_n^0 a^0 b^n + C_n^1 a^1 b^{n-1} + C_n^2 a^2 b^{n-2} + ... + C_n^p a^p b^{n-p} + ... + C_n^n a^n b^0 \\ &= \sum_{p=0}^{n} C_n^p a^p b^{n-p.}\end{aligned}$$

2.2 Notion de calcul des probabilités

Définition (Expérience aléatoire)

Toute expérience dont on ne peut pas prévoir les résultats est dite expérience (épreuve) aléatoire.

Exemples

Jet d'une pièce de monnaie.

Durée de vie d'une ampoule.

Définition (Univers des éventualités)

C'est l'ensemble des résultats possibles d'une épreuve qu'on note Ω.

Exemples

Le jet d'une pièce de monnaie

On jet successivement un dé jusqu'à l'obtention de la face N°6.

Définitions

1- Un événement élémentaire est toute issue possible d'une épreuve.

2- Un événement est tout fait dont la réalisation dépend exclusivement de l'issue d'une épreuve aléatoire.

Exemple

Si on lance deux dés non truqués et on s'intéresse à la somme des points obtenus, alors

$$\Omega = \{1,2,3,4,5,6\} \times \{1,2,3,4,5,6\}.$$

Il y a 36 possibilités, se sont des événements élémentaires.

L'obtention d'une somme ≤ 7 est un événement (composé).

Thérminologie des probabilités

Théorie des ensembles	Théorie des probabilités
Référentiel Ω	Evénement sûr
Ensemble vide ϕ	Evénement impossible
$A \subset \Omega$	Evénement
$\overline{A} = \complement_{\Omega}^{A}$	Evénement contraire de A
$A \cap B = \phi$	Evénements incompatibles
$A \cup B$	Réalisation d'au moins d'un de ces deux événements
$A \cap B$	Réalisation simultanée
$A \subset B$	La réalisation de A entraîne celle de B

Définition (Probabilités)

Une probabilité P sur Ω est une application : $P : \mathcal{P}(\Omega) \longrightarrow [0,1]$ *vérifiant :*

1- $P(\Omega) = 1$, 2- Si A et B sont incompatibles alors : $P(A \cup B) = P(A) + P(B)$.

Résultats

1- Puisque A et $\overline{A}$ sont incompatibles on a :

$$\begin{cases} P\left(A \cup \overline{A}\right) = P(A) + P\left(\overline{A}\right) \\ P\left(A \cup \overline{A}\right) = P(\Omega) = 1 \end{cases} \Longrightarrow P\left(\overline{A}\right) = 1 - P(A).$$

2- En particulier :

$$P(\phi) = 1 - P(\Omega) = 1 - 1 = 0.$$

3- Si A et B sont des événements quelconque alors :

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

Définition (Equiprobabilités)

On dit qu'on a équiprobabilité, chaque fois que, lors d'une épreuve toutes les éventualités ont la même probabilité de survenir.

Exemple

Lors du jet d'une pièce de monnaie, chaque côté à 1/2 de chance.

Généralisation

Dans la situation courante d'équiprobabilité, on trouve en général la probabilité d'un événement A par la formule :

$$P(A) = \frac{card(A)}{card(\Omega)} = \frac{\text{nombre de cas favorables à l'événement } A}{\text{nombre de cas possibles en tout}}.$$

Définition (probabilité conditionnelle)

La probabilité d'un événement B, conditionnée par l'événement A, ou encore la probabilité de B sachant que A est réalisé, est par définition :

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

Remarque

$$P(A \cap B) = P(B/A)\,P(A) = P(A/B)\,P(B).$$

Définition (Indépendance de deux événements)

Soient A et B deux événements de probabilité non nulle. On dit que A est indépendant de B si :

$$P(A/B) = P(A) \;\; ou \; P(B/A) = P(B).$$

Théorème de Bayes

Soit $A \subset \Omega$ et $E_1, E_2, ..., E_n$ une partition de Ω, alors :

$$P(E_i/A) = \frac{P(A/E_i)\,P(E_i)}{\sum_{i=1}^{n} P(A/E_i)\,P(E_i)}.$$

Chapitre 3

Les variables aléatoires

Définition

Une variable aléatoire (v.a), fait correspondre à chacun des événements élémentaires de Ω un nombre.

Remarques

1- La v.a est discrète si elle ne peut prendre que des valeurs numériques isolées. Dans le cas contraire, on dit qu'elle est continue.

2- Une v.a discrète finie est donc une application :

$$X : \Omega \longrightarrow \{x_{1,}x_2, ..., x_n\},$$

par exemple la somme des numéros des faces lors du lancés de deux dés, le nombre d'enfants d'une famille, ...

Définition

La loi de probabilité (distribution de probabilité) d'une v.a X, est une fonction qui associe à chacune des valeurs possibles de X, la probabilité de l'événement correspondant.

Remarques

1- La loi de probabilité d'une v.a $X : \Omega \longrightarrow \mathbb{R}$ est la fonction $P_X : \mathbb{R} \longrightarrow [0, 1]$ définie par :

$$\forall x \in \mathbb{R} : P_X(x) = P\left(X^{-1}(x)\right) = P\left(\{\omega \in \Omega / X(\omega) = x\}\right) \text{ où}$$

on note par abréviation : $P_X(x) = P(X = x)$, c'est la probabilité de l'événement qui consiste à avoir la valeur x.

2- Donc l'ensemble des couples (x_i, p_i) constitue la loi de probabilité de la v.a X, et dans le cas finie on a toujours $\sum_{i=1}^{n} P(X = x_i) = 1$.

3.1 Les variables aléatoires discrètes

Exemples

1- On lance successivement deux fois une pièce de monnaie. X est la v.a qui dénombre les faces obtenues en ces deux jets. Déterminer la loi de probabilité de X.

2- Une urne contient 5 boules blanches et 3 boules noires. Si deux boules sont tirées au hasard sans remplacement. Et si X désigne le nombre de boules blanches. Déterminer la loi de probabilité de X.

Définition (Fonction de répartition)

La fonction de répartition F associée à une v.a X est la fonction définie sur $\mathbb{R}$ à valeurs dans $[0, 1]$ par :

$$F_X(x) = F(x) = P(X \leq x).$$

Propriétés

1- $\lim_{x \longrightarrow -\infty} F_X(x) = 0$; $\lim_{x \longrightarrow +\infty} F_X(x) = 1$.

2- F_X est une fonction croissante, continue à droite, i. e. $\lim_{x \longrightarrow < x_0} F_X(x) = F_X(x_0)$.

3- Lorsque l'ensemble des valeurs possibles de la v.a X est finie $X : \Omega \longrightarrow \{x_1, x_2, ..., x_n\}$ on a :

$$
\begin{aligned}
F_X(x) &= 0 \text{ si } x \in \left]-\infty, x_1\right[\\
F_X(x) &= 0 \text{ si } x \in \left[x_n, +\infty\right[\\
F_X(x) &= F_X(x_i) \text{ si } x \in \left[x_i, x_{i+1}\right[.
\end{aligned}
$$

Remarques

1- $F_X(x) = \sum_{x_i \leq x < x_{i+1}} P(X = x_i)$;

2- $P(X = x_i) = F_X(x_i) - F_X(x_{i-1})$;

3- La représentation graphique de la fonction de répartition est la courbe cumulative, qu'on appelle courbe en escalier.

Définition (Espèrence mathématique)

L'espérance mathématique $E(X)$ d'une v.a discrète est la moyenne arithmétique des valeurs possibles pondérée par les probabilités correspondantes.

Remarques

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

2- L'espérance d'une v.a discrète est une forme linéaire, i.e. si X et Y sont deux v.a discrètes sur le même univers Ω et $a, b \in \mathbb{R}$: $E(aX + bY) = aE(X) + bE(Y)$.

3- Une v.a discrète X est dite centrée si $E(X) = 0$.

Définition (variance d'une v.a)

Pour une v.a X prenant les valeurs x_i avec les probabilités p_i, son espèrance étant $E(X) = \sum_{i=1}^{n} p_i x_i$, on définit sa variance :

$$
\begin{aligned}
var(X) &= E\left((X - E(X))^2\right). \\
&= E\left(X^2\right) - E^2(X) \\
&= \sum_{i=1}^{n} p_i x_i^2 - \left(\sum_{i=1}^{n} p_i x_i\right)^2.
\end{aligned}
$$

Définition (l'écart-type d'une v.a)

L'écart-type d'une v.a notée σ *est* $\sigma(X) = \sqrt{var(X)}$.

Propriété

$$\forall a, b \in \mathbb{R} : \sigma(aX + b) = |a| \sigma(X).$$

3.2 Les variables aléatoires continues

Une v.a réelle est continue lorsque l'ensemble de ses valeurs est constitué par des intervalles de $\mathbb{R}$.

Fonction de répartition

Pour caractériser une v.a continue X, on peut utiliser sa fonction de répartition F, qui est :

$$F : \mathbb{R} \longrightarrow [0,1] ; x \longmapsto P(X \leq x).$$

On a donc : $F_X(x) = P(X \leq x)$. Et pour $X \in]a,b]$ on a : $\mathrm{P}(X \leq x) = F(b) - F(a)$.

Exemple

Considérons :

$$\begin{array}{rcl} F & : & \mathbb{R} \longrightarrow [0,1] \\ x & \longmapsto & F_X(x) = \left\{ \begin{array}{c} 0 \text{ si } x < 0 \\ \frac{3}{4}x^2 - \frac{1}{4}x^3 \text{ si } 0 \leq x \leq 2 \\ 1 \text{ si } x > 2 \end{array} \right. . \end{array}$$

Ainsi : $P(x \leq 1.5) = ?, P(0.5 < x \leq 1.5) = ?$

Remarque

On considérera par la suite des v.a continues dont les fonctions de répartition sont dérivables sur $\mathbb{R}$.

Définition (Densité de probabilité)

La densité de probabilité f d'une v.a continue X est la dérivée de la fonction de répartition F :

$$\forall x \in \mathbb{R} : f(x) = F'_X(x).$$

Propriété

$$P(a < X \leq b) = F(b) - F(a) = \int_a^b f(t)\, dt.$$

Remarque

Du moment que la fonction de répartition est croissante, sa dérivée (la densité de probabilité) est une fonction positive.

Propriétés de la fonction de répartition

$$\begin{aligned}
1- \int_a^{+\infty} f(t)\, dt &= \lim_{b \longrightarrow +\infty} \int_a^b f(t)\, dt, \text{ et } \int_{-\infty}^b f(t)\, dt = \lim_{a \longrightarrow -\infty} \int_a^b f(t)\, dt. \\
2- \text{ Puisque } &: P(-\infty < X < +\infty) = 1 \text{ on a donc : } \int_{-\infty}^{+\infty} f(t)\, dt = 1. \\
3- \ F_X(x) &= P(X \leq x) = \int_{-\infty}^{x} f(t)\, dt.
\end{aligned}$$

Théorème

Pour qu'une fonction continue f soit une densité de probabilité, il faut et il suffit qu'elle soit positive et :

$$\int_{-\infty}^{+\infty} f(t)\, dt = 1.$$

Formules utiles

$$\begin{aligned}
E(X) &= \int_{-\infty}^{+\infty} t f(t)\, dt, \\
var(X) &= \int_{-\infty}^{+\infty} (t - E(X))^2 f(t)\, dt = \int_{-\infty}^{+\infty} t^2 f(t)\, dt - (E(X))^2, \\
\sigma(X) &= \sqrt{var(X)}.
\end{aligned}$$

Opérations sur les v.a continues

Les v.a continues ont des propriétés analogues aux v.a discrètes :

si X, Y sont des v.a continues définies sur Ω, alors les fonctions :

$$X + Y, X.Y, \lambda.X \ \ (\lambda \in \mathbb{R}) \ \text{ sont des v.a et :}$$

$$\begin{aligned} E\left(X \pm Y\right) &= E\left(X\right) \pm E\left(Y\right); \ E\left(\lambda.X\right) = \lambda.E\left(X\right) \ \text{ (linéarité)} \\ \sigma^2\left(\lambda.X\right) &= \lambda^2\sigma^2\left(X\right); \ \sigma^2\left(X + \mu\right) = \sigma^2\left(X\right), \ \text{ d'où } \sigma^2\left(\lambda.X + \mu\right) = \lambda^2\sigma^2\left(X\right) \end{aligned}$$

Variables aléatoire centrée réduite

Dans le calcul des problèmes liés à une v.a continue X, on peut utiliser la v.a

$$Y = \frac{X - E\left(X\right)}{\sigma\left(X\right)} \Longrightarrow E\left(Y\right) = 0 \text{ et } \sigma\left(Y\right) = 1. \text{ La v.a } Y \text{ est centrée réduite.}$$

Chapitre 4

Les lois de distribution statistique

4.1 Les modèles discrets

4.1.1 La loi de Bernoulli $B(p)$

On réalise une expérience aléatoire qui a deux résultats possibles : soit le succès qui a une probabilité p de se réaliser, soit l'échec qui a une probabilité $q = 1 - p$. La v. a $X =$ nombre de succès obtenus, suit la loi de Bernoulli notée $B(1, p)$ et définie par :

$$\begin{aligned} P &: \{0,1\} \longrightarrow [0,1] \\ P(X=0) &= 1-p \text{ et } P(X=1)=p. \end{aligned}$$

Propriétés

$$\text{Si } X\tilde{\ }B(1,p) \Longrightarrow E(X) = p \text{ et } var(X) = pq.$$

4.1.2 La loi binomiale $B(n, p)$

On réalise n fois successivement et d'une manière indépendante une expérience aléatoire qui a deux résultats possibles, le succès (associé au résultat pour lequel nous voulons déterminer la probabilité) qui a une probabilité p de se réaliser et l'échec qui a une probabilité $q = 1 - p$ de se réaliser. La v. a X = nombre de succès obtenus au cours des n épreuves suit la loi binomiale notée $B(n, p)$ définie par :

$$\begin{array}{ccl} P & : & \{0, 1, ..., n\} \longrightarrow [0, 1] \\ k & \longmapsto & P(X = k) = C_n^p p^k (1 - p)^{n-k} \end{array}$$ (qui représente la probabilité d'avoir k succès en n essais).

Exemples

Lancement d'une pièce de monnaie (pile ou face) ; qualité d'un produit (bon ou défectueux) ; sondage électoral (pour ou contre) ;...

Propriétes

Si $X\tilde{\ }B(n, p) \Longrightarrow E(X) = np; var(X) = npq.$

Si $X_1\tilde{\ }B(n_1, p)$ et $X_2\tilde{\ }B(n_2, p)$ alors ces v. a sont indépendantes et $Y = X_1 + X_2\tilde{\ }B(n_1 + n_2, p)$.

Exercice

On jette une pièce de monnaie bien équilibrée 10 fois de suite. On désigne par X la v. a. comptant le nombre de piles obtenus.

1- Quelle est la loi de probabilté de la v. a. X ?

2- Quelle est la probabilité d'obtenir exactement 3 piles ?

4.1.3 La loi de Poisson $P(\lambda)$

Une v. a. X suivant la loi de Poisson admet la probabilité :

$$P(X = k) = \frac{e^{-\lambda}\lambda^k}{k!} \text{ où } \lambda \text{ est un paramètre positif.}$$

Symboliquement on note par $X\tilde{\ }P(\lambda)$ une v. a. suivant la loi de Poisson d'un paramètre λ.

Conditions d'application

La loi de Poisson peut être introduite soit :

1- comme la résultante d'un processus aléatoire particulier ; le processus de Poisson,

2- pendant un intervalle de temps donné ; comme :

- le nombre d'appelles téléphoniques reçus par un standard ;

- le nombre de voitures traversant un certain pont ;

-le nombre d'arrivées à un guichet...

Propriétés

Si $X\tilde{}P(\lambda) \Longrightarrow E(X) = var(X) = \lambda$.

Si $X_1\tilde{}P(\lambda_1)$ et $X_2\tilde{}P(\lambda_2)$, et X_1 et X_2 sont indépendantes, alors $Y = X_1 + X_2\tilde{}P(\lambda_1 + \lambda_2)$.

4.1.4 Convergence de la loi Binomiale vers la loi de Poisson

Théorème

Si $X\tilde{}B(n,p)$ alors $n \longrightarrow +\infty$, p tend vers 0 et np tend vers une limite finie λ. Alors on peut approcher cette loi par une loi de Poisson de moyenne λ.

Explications

1- Le terme approcher signifie ici que pour chaque valeur entière k, là $P(X = k)$ aura pour limite $\frac{e^{-\lambda}\lambda^k}{k!}$.

2- Si $X\tilde{}B(n,p)$ où $n > 50$ et $p < 10\%$ alors la loi Binomiale sera approximée par la loi de Poisson où $\lambda = np = E(X)$.

Exercice

Le nombre moyen de communications téléphoniques reçues par un standard entre 8H15 et 8H16 est de 2. Quelle est la probabilité qu'entre 10H53 et 10H54, on reçoive :

1- un seul appel ; 2- trois appels.

4.1.5 La loi Hypergéométrique $H(N, n, p)$

Dans une population de taille N, on a deux types d'éléments, N_1 éléments de type I et N_2 éléments de type II. On effectue n tirages sans remise (=prélèvement d'un seul coup de n éléments). La v. a. discrète X =nombre d'éléments de type I obtenus après les n tirages suit la loi hypergéométrique notée $H(N, n, p)$ avec $p = \frac{N_1}{N}$, définie par :

$$\begin{array}{rcl} P & : & \{0, 1, ..., n\} \longrightarrow [0, 1] \\ k & \longmapsto & P(X = k) = \dfrac{C_{N_1}^{k} C_{N_2}^{n-k}}{C_N^n} \text{ avec } N_1 = N_p, N_2 = N_q. \end{array}$$

Propriété

Si $X\tilde{}H(N, n, p)$ alors $E(X) = np, var(X) = \frac{N-n}{N-1}npq.$

Convergence de la loi Hypergéométrique vers la loi Binomiale

Si $N \longrightarrow \infty$ avec $\frac{N_1}{N}$ et $\frac{N_2}{N}$ restant finis alors : $H(N, n, p) \longrightarrow B(n, p)$. (en pratique $\frac{n}{N} < 10\%$).

Exercice

Une urne contient 10 boules dont 6 sont blanches. On extrait au hasard 5 boules de cette urne. En notant par X la v. a. comptant le nombre de boules blanches parmi les 5 tirées, déterminer la loi de probabilité de la v. a. X.

4.2 Les modèles continus

4.2.1 La loi normale $N(\mu, \sigma)$

- $\mu \in IR, \ \sigma \in IR_+^*$

C'est la plus importante des lois de probabilité continue. Des questions tant théoriques que pratiques font appel à cette loi (souvent loi limite). Historiquement elle apparaît vers 1773 comme la forme limite de la loi binomiale (Abraham de Moivre). Gauss en 1812 lui donnèrent sa forme définitive.

Définition (la loi Normale)

La loi normale de paramètres μ *et* σ*, notée* $N(\mu,\sigma)$*, est la loi de probabilité de la v. a. continue* X *dont la densité de probabilité est :*

$$f(x) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\frac{1}{2}\left(\frac{x-\mu}{\sigma}\right)^2\right]; \mu \in \mathbb{R} \text{ } et \text{ } \sigma \in \mathbb{R}_+^*.$$

On note alors $X\tilde{}N(\mu,\sigma)$.

Propriétés

Si $X\tilde{}N(\mu,\sigma)$ alors :

1- La fonction de répartition

$$F_X(x) = \int_{-\infty}^{x} \frac{1}{\sigma\sqrt{2\pi}} \exp\left[-\frac{1}{2}\left(\frac{t-\mu}{\sigma}\right)^2\right] dt.$$

2-

$$E(X) = \mu \text{ et } var(X) = \sigma^2.$$

Définition (la loi normale centrée réduite)

Une v. a. continue X est dite normale centrée réduite si sa densité de probabilité s'écrit :

$$f(x) = \frac{1}{\sqrt{2\pi}} \exp\left(\frac{-x^2}{2}\right); \; x \in \mathbb{R}.$$

On note $X\tilde{}N(0,1)$.

Propriétés

Si $X\tilde{}N(0,1)$ alors :

1- La fonction de répartition

$$F_X(x) = \int_{-\infty}^{x} \frac{1}{\sqrt{2\pi}} \exp\left[-\frac{t^2}{2}\right] dt.$$

2-

$$E(X) = 0 \text{ et } var(X) = 1.$$

Remarques

1- La primitive de la fonction $f(x)$ n'est pas calculable par les méthodes classiques. Les valeurs de cette intégrale sont alors obtenues par approximation. Pour t donnée la valeur de cette intégrale sera notée $\Phi(x)$. Une table nous donne ces valeurs pour $x > 0$.

Par exemple $\Phi(0.15) = 0.5596$, $\Phi(1.24) = 0.8925$.

Pour les valeurs négatives de x on utilise la transformation :

$$\Phi(-x) = 1 - \Phi(x).$$

2-

$$Y\tilde{\ }N(\mu, \sigma) \Longleftrightarrow X = \frac{Y - \mu}{\sigma}\tilde{\ }N(0, 1).$$

Et $P(a \leq Y \leq b) = \Phi\left(\frac{b-\mu}{\sigma}\right) - \Phi\left(\frac{a-\mu}{\sigma}\right)$.

4.2.2 Utilisation de la table

1er cas

On connaît la loi de X, la valeur de x et on veut calculer $P(X \leq x)$: on calcule $t = \frac{X-E(X)}{\sigma(X)}$, on cherche $\Phi(t) = \pi(t)$ dans la table.

2ème cas

On connaît la loi de X, $P(X \leq x)$ et on veut calculer la valeur de x. On cherche dans la table les deux valeurs encadrant la probabilité connue, on calcule par interpolation linéaire la valeur de t corréspondant à cette probabilité. On en déduit $x = \sigma(X)\, t + E(X)$.

Exemples

1- Soit $X\tilde{\ }N(60, 3)$; $P(57 < X \leq 61) =?$

2- Si $X\tilde{\ }N(60, 3)$ et on a $P(X \leq x) \simeq 0.8686$; $x =?$

3ème cas

On sait que X suit une loi normale, mais on ne connaît ni la moyenne μ, ni l'écart-type σ. On peut les calculer à partir de deux donées qui nous permettent de former un système de deux équations à deux inconnues.

Exemple

$P(X \leq 0.5) = 0.5517;\ P(X \leq 2.6) = 0.9515.$

4.2.3 Approximation par une loi normale

Loi Binomiale : Une loi binomiale $B(n,p)$ tend vers une loi normale de moyenne $\mu = np$ et de variance $\sigma^2 = npq$ lorsque $n \longrightarrow +\infty$. La convergence est rapide lorsque p est proche de 0.5. On prend une loi normale lorsque le produit $np(1-p)$ est suffisamment important.

Exemple $X\tilde{}B(50, 0.4)$.

La loi de Poisson : une loi de Poisson de paramètre λ tend vers une loi normale de moyenne λ et d'écart-type $\sqrt{\lambda}$ lorsque $\lambda \longrightarrow +\infty$. L'approximation est justifiée lorsque μ est suffisamment grand.

Remarque

Dans les deux cas précédents, les valeurs limites des paramètres pour justifier l'approximation par une loi normale sont liées à la précision que l'on exige des calculs numériques. Les normes usuelles sont :

Loi exacte	Conditions	Approximation
Loi Binomiale $B(n,p)$	$np \leq 5$ et $n \geq 30(50)$	loi de $P(np)$
	$np > 5$ et $n(1-p) > 5$ et $n \geq 30$ ou $p \simeq 0.5$ et $n \geq 20$	loi normale $N\left(np, \sqrt{npq}\right)$
Loi de poisson	$\lambda > 20$	loi normale $N\left(\lambda, \sqrt{\lambda}\right)$

Correction de continuité

L'utilisation de la loi normale comme approximation de la loi binomiale et de la loi de Poisson pose le problème du passage du cas discret au cas continue. Il faut donc associer à chaque valeur isolée un intervalle de $\mathbb{R}$. Pour les valeurs entières on pose n tel que : $\left]n-\frac{1}{2}, n+\frac{1}{2}\right[$.

4.3 Autres lois de probabilité continues

4.3.1 La loi Log-Normale de Galton

Définition (la loi Log-Normale)

La loi log-normale de Galton d'une v. a. X continue de paramètres μ et σ est définie par sa densité de probabilité :

$$f(x) = \frac{1}{\sigma x\sqrt{2\pi}} \exp\left[\frac{-1}{2}\left(\frac{\ln x-\mu}{\sigma}\right)^2\right]; \quad \mu \in \mathbb{R}, \sigma \in \mathbb{R}_+^*, \ x > 0.$$

Où $E(X) = \exp\left(\mu + \frac{\sigma^2}{2}\right)$; $var(X) = \exp(2\mu + 2\sigma^2) - \exp(2\mu + \sigma^2)$.

Explication

Si X suit la loi log-normale de paramètresμ et σ, la v. a. $Y = \ln(X)$ suit la loi normale $N(\mu, \sigma)$. Quand on utilise une loi de Galton, les calculs pratiques sont faits avec la v. a. $\ln(X)$ qui suit une loi normale. La loi log-normale trouve ses applications en géologie.

Exemple

La dimension X en millimètres d'un grain de verre abrasif suit une loi log-normale de paramètres $\mu = -0.69$ et $\sigma = 0.1$. Calculer le pourcentage de grains dont les dimensions sont comprises entre $0.3mm$ et $0.6mm$.

4.3.2 La loi du Khi-deux à n degrés de liberté $\left(\chi_n^2\right)$

Cette loi joue un rôle important dans les tests statistiques.

On obtient une valeur χ_n^2 en additionnant des nombres au carré, donc cette valeur ne peut pat être négative.

L'aspect de la courbe d'une distribution χ_n^2 variera selon le nombre de degrés n qui est le seul paramètre de cette distribution.

Définition

Soient $X_1, X_2, ..., X_n$ *n v. a. indépendantes telles que* $X_i \tilde{} N(0,1) \quad \forall i$. *Alors :*

$$X_1^2 + X_2^2 + ... + X_n^2 \tilde{} \chi_n^2.$$

Remarque

La fonction densité de probabilité de χ_n^2 est :

$$f_{\chi_n^2}(t) = c_n t^{\frac{n}{2}-1} \exp\left(\frac{-t}{2}\right),$$

où c_n sont telles que :

$$\int_{\mathbb{R}} f_{\chi_n^2}(t)\, dt = 1.$$

Si $n > 2$ alors le mode$= n-2$ (mode =valeur pour laquelle la courbe atteint son maximum).

Propriétés

Si $X \tilde{} \chi_n^2$ (mode$= n-2,\ n > 2$) alors $E(X) = n,\ var(X) = 2n$.

4.3.3 Convergence de la loi χ_n^2 vers la loi normale (approximation)

Soit $X \tilde{} \chi_n^2$, alors $\frac{X-n}{\sqrt{2n}} \longrightarrow N(0,1)$, quand $n \longrightarrow \infty$.

Ou bien : $\chi_n^2 \approx N\left(n, \sqrt{2n}\right),\ n \longrightarrow \infty$. En pratique $n > 30$.

4.3.4 La loi de Fisher-Snedecor $(F(n_1, n_2))$

Définition

Soient $Y_1\tilde{}\chi^2_{n_1}$ et $Y_2\tilde{}\chi^2_{n_2}$, deux v.a. indépendantes. Alors

$$F = \frac{Y_1/n_1}{Y_2/n_2}\tilde{}F(n_1, n_2) \text{ (loi de Fischer-Snedecor à } n_1 \text{ et } n_2 \text{ degrés de liberté)}$$

Remarque

La fonction densité de probabilité de $F(n_1, n_2)$ est :

$$f_F(t) = c_{n_1,n_2} t^{\frac{n_1}{2}-1} (n_1 t + n_2)^{-\frac{(n_1+n_2)}{2}}, \ t > 0.$$

Propriétés

Si $F\tilde{}F(n_1, n_2)$, alors $E(F) = \frac{n_1}{n_2-2}, n_2 > 2$ et $var(F) = \frac{2n_2^2(n_1+n_2-2)}{n_1(n_2-2)^2(n_2-4)};\ n_2 > 4.$

4.3.5 La loi de Student à n degrés de liberté (T_n)

Elle joue un rôle important dans l'estimation par intervalle de confiance. Elle est symétrique, de moyenne nulle et dépend d'un paramètre n appelé nombre de degrés de liberté.

L'aspect de la courbe variera selon le nombre de degrés de liberté n (de façon générale, elle est plus aplatie que $N(0,1)$ et quand n augmente $(n > 30)$ les deux courbes se confondent).

Définition

Soit $X\tilde{}N(0,1)$, $Y\tilde{}\chi^2_n$ v. a. indépendantes. Alors :

$$Z = \frac{X}{\sqrt{Y/n}}\tilde{}T_n.$$

Remarque

La fonction densité de probabilité de T_n est :

$$f_{t_n}(t) = c_n \left(1 + \frac{t^2}{n}\right)^{-\frac{(n+1)}{2}} ;$$

où c_n sont tels que : $\int_{\mathbb{R}} f_{t_n}(t)\, dt = 1$.

Propriétés

Si $X \tilde{} T_n$, alors : $E(X) = 0,\ n > 1$ et $var(X) = \frac{n}{n-2}, n > 2$.

4.3.6 Convergence de la loi de Student vers la loi normale (approximation)

Soit $X \tilde{} T_n$ alors : $X \longrightarrow N(0,1)$ quand $n \longrightarrow \infty$. En pratique $n > 30$.

Bibliographie

[1] P.Trousset., J.F. Morin, Mathématiques-pour-Les Sciences de la vie-2- probabilités et statistiques, Ediscience International ;1994.

[2] H.Carnec, J.M.Dagoury, R.Seroux, M.Thomas, Itinéraires en statistiques et probabilités, édtition . Ellipses.

[3] Avner Bar-Hen,Cours de DEUG Probabilités et Statistiques, Université Aix-Marseille III ; 2002-2003.

[4] J.M.Bernabotte, Cours de Statistiques ; D'IUT.

[5] D.Mouchiroud, Mathématiques : Outils pour la Biologie - Deug - SV1-UCBL .10/10/2002.

[6] http ://www.maths.tetras.org/ : une introduction sur les livres de Probabilités et Statistiques de Guillaume Laget.