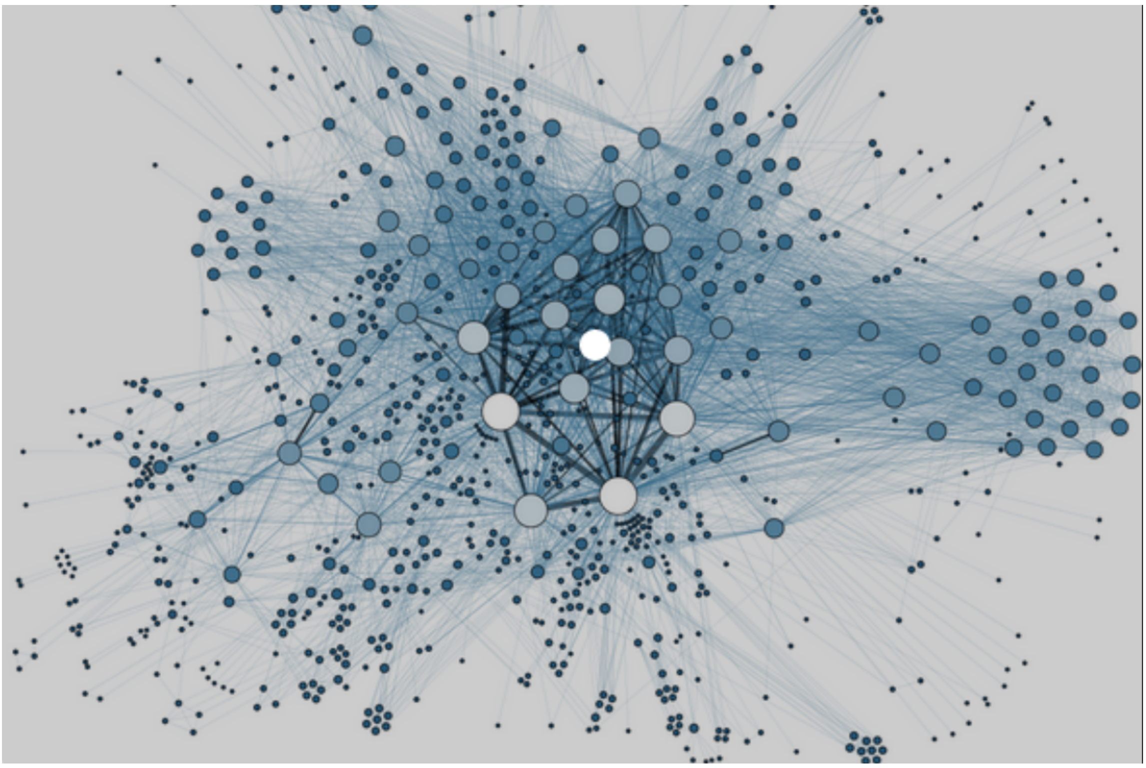
L2 - MIASHS - Statistiques



Ces notes ont pour vocation d'assurer une base complète et uniformisée du contenu du cours "statistiques".
Elles ne contiennent pas la totalité des dessins, exemples, remarques et commentaires donnés au cours des séances : pour compléter, la prise de notes individuelle est donc vivement recommandée !

Table des matières

Chapitre 1

Statistiques descriptives

I Introduction

Dans une étude statistique, on considère un certain ensemble Ω que l'on nomme population. Celle-ci peut aussi bien être constituée d'objets matériels que d'êtres vivants. Les éléments ω de cette population s'appellent des individus ou unités statistiques. En pratique, la définition d'une population doit être suffisamment précise pour que l'on puisse décider sans ambiguïté ce qui en fait partie et ce qui n'en fait pas partie. En démographie, par exemple, on devra préciser, à propos de la fécondité, si l'on considère toutes les femmes, ou les femmes en âge d'avoir des enfants et ce que l'on entend par là etc.
Le terme "caractère" désigne la propriété que l'on étudie sur les individus de cette population (taille, poids, âge, durée de vie, nombre annuel de défaillances dans le cas de machines, nombre de lettres pour les mots d'un texte etc.). On utilise aussi fréquemment le terme "variable" que nous réserverons plutôt au cas des variables aléatoires et qui représentent en quelque sorte la modélisation idéale de la réalité représentée par les caractères. Plus précisément, un caractère X est une application de la population Ω dans un ensemble E. L'ensemble $\{X(\omega), \omega \in \Omega\} \subset$ E, des valeurs du caractère X pour chaque individu ω de la population constitue une série statistique.
Un caractère peut être quantitatif ou qualitatif. Un caractère quantitatif est soit discret, soit continu. Il est discret dans le cas où l'ensemble $\mathrm{E} = \mathrm{X}(\Omega)$, des valeurs possibles du caractère, est intrinsèquement fini ou dénombrable. Il est au contraire continu si cet ensemble, $E = X(\Omega)$, est un intervalle ou une réunion d'intervalles de $\mathbb{R}$

Remarque

Un caractère quantitatif est aussi appelé caractère numérique, mais il ne suffit pas que E soit un ensemble de nombre pour que le caractère soit quantitatif : le numéro du département de naissance n'est pas un caractère quantitatif. Il s'agit, en fait, du codage d'un caractère qualitatif.

La représentation graphique et le calcul de coefficients résumant l'information contenue dans une série statistique sont du ressort de la statistique descriptive. Une étude élémentaire en est faite au chapitre 1. Le prolongement naturel de la statistique descriptive élémentaire est l'analyse des données qui exploite des méthodes géométriques (Cours d'algèbre de deuxième année) lorsque l'on traite un grand nombre de caractères simultanément ($n \geqslant 3$) et qui sera étudié en 3-ième et 4-ième année.
Souvent, on ne dispose pas des valeurs du caractère pour tous les individus de la population, mais seulement pour une petite partie d'entre eux, ce que l'on appelle échantillon [1] (sample). On cherche alors à obtenir des informations sur la population totale à partir de celles contenues dans l'échantillon (Induction statistique). Cela ressortit à la statistique mathématique qui repose sur des hypothèses probabilistes. Les notions indispensables du cours de probabilités sont rappelées au chapitre 2. Trois

aspects de la statistique mathématique sont abordés ensuite : l'estimation ponctuelle au chapitre 3, l'estimation par intervalles au chapitre 4, la théorie des tests envisagés sous le double point de vue : paramétrique et non paramétrique aux chapitres 5 et 6 .
Une définition plus précise et plus technique du terme échantillon sera vue au chapitre 3.

II Présentation des données

1 Caractère discret

La série statistique (data) est présentée sous forme de tableau où à chaque valeur x_i du caractère, on fait correspondre le nombre n_i des individus possédant cette valeur ; n_i est l'effectif (ou fréquence absolue, frequency). Les valeurs x_i sont classés par ordre croissant.

Valeurs du caractère	effectif
x_1	n_1
x_2	n_2
...	...
x_i	n_i
...	...
x_k	n_k

Naturellement l'effectif total, n, est égal à la somme, $n = \displaystyle\sum_{i=1}^{k} n_i$, des effectifs de chaque classe.

Définitions

- Les fréquences relatives (fréquences, relative frequency) :
  $f_i = \dfrac{n_i}{n}$ est la proportion des individus présentant le caractère x_i par rapport à l'effectif total. Ce nombre peut être exprimé en pourcentage (percentage) en étant multiplié par cent.
- Les effectifs cumulés croissants (cumulative frequency) :
  $T_i = \sum_{p=1}^{i} n_p$ est le total des effectifs de la i-ème classe et des classes dont le caractère est inférieur à x_i. Sur la i-ème ligne du tableau, l'effectif cumulé représente le nombre d'individus ayant une valeur du caractère au plus égale à x_i. On peut naturellement définir les effectifs cumulés décroissants etc..
- Les fréquences relatives cumulées (fréquences cumulées, relative cumulative frequency) :
  t_i est le rapport de l'effectif cumulé d'une classe à l'effectif total (on peut aussi l'exprimer en pourcentage).

2 Caractère continu

On procède comme précédemment mais en subdivisant $X(\Omega)$ en classes $[e_i, e_{i+1}[$ (ou intervalle de classe) ; l'effectif n_i représente le nombre des individus dont la valeur du caractère est dans la i-ème classe (i.e. compris entre e_i et e_{i+1}).

classes	centre de classe	effectifs
$e_1 \leqslant x < e_2$	x_1	n_1
$e_2 \leqslant x < e_3$	x_2	n_2
...	...	...
$e_i \leqslant x < e_{i+1}$	x_i	n_i
...	...	...
$e_k \leqslant x < e_{k+1}$	x_k	n_k

Les autres quantités (fréquences etc.) sont définies de la même manière que dans le cas discret.

III Représentations graphiques

1 Caractère discret

Les effectifs et fréquences relatives se représentent à l'aide d'un diagramme en bâtons (barchart). le diagramme est d'ailleurs le même pour les fréquence absolues et relatives par simple changement d'échelle sur l'axe des ordonnées.

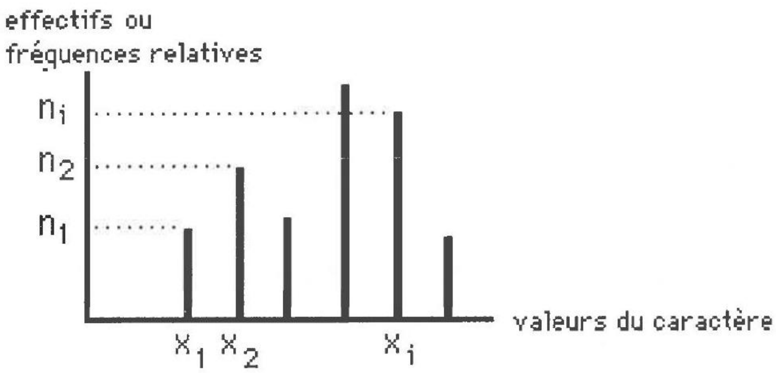


fig 1 Diagramme en bâtons

Les effectifs cumulés et fréquences cumulées se représentent à l'aide d'un diagramme en marches d'escalier. (même remarque concernant l'échelle)

Il s'agit dans ce cas de la représentation d'une fonction qui est définie pour tous les réels et que l'on appelle fonction de répartition. Ainsi la fonction de répartition des effectifs est définie par :

$$F_X(x) = \frac{\operatorname{card}\{\omega \in \Omega, X(\omega) \leq x\}}{\operatorname{card} \Omega}$$

Tous les points d'une ligne horizontale ont une interprétation. On vérifie en effet que même si x n'est l'une des extrémités d'un des intervalles de classe, l'expression précédente a un sens. Notons que cela n'était pas le cas des diagrammes en bâton, où les points d'un segment reliant deux bâtons consécutifs n'ont pas de sens, même s'il est courant de les relier pour des raisons esthétique.

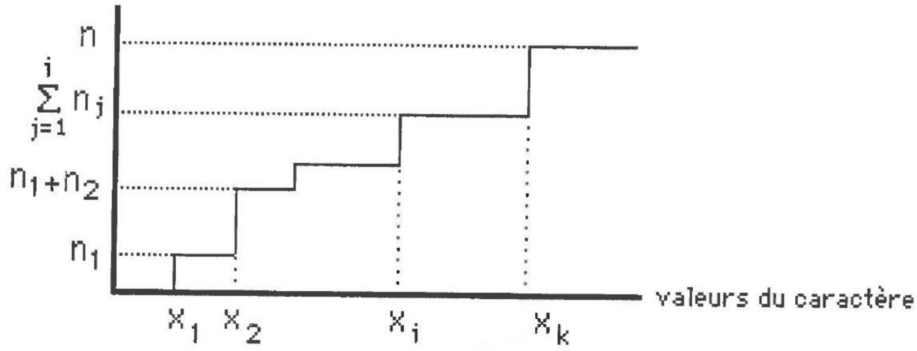
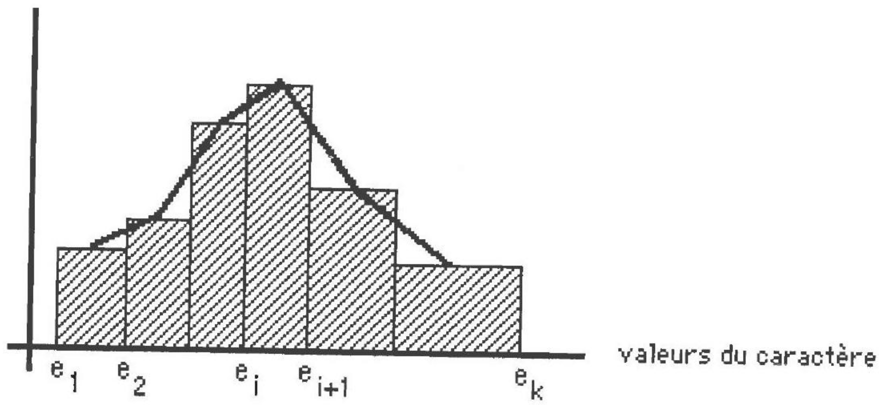


fig 2 Courbe des fréquences cumulées.

2 Caractère continu

Les effectifs et les fréquences relatives se représentent à l'aide d'un histogramme (frequency histogram). Attention l'histogramme doit assurer la proportionnalité des aires des rectangles aux effectifs (ou aux fréquences) et ceci afin de respecter la cohérence des représentations dans le cas de subdivisions d'intervalles de classes en sous-intervalles.
On peut éventuellement tracer une ligne brisée joignant les milieux des segments horizontaux de l'histogramme. On obtient alors, le polygone des effectifs (ou des fréquences).
effectifs ou fréquences relatives



L'histogramme correspond à l'hypothèse que les effectifs sont uniformément répartis à l'intérieur d'une même classe, alors que le polygone des fréquences (frequency polygon) correspond à l'hypothèse que la densité de répartition augmente (ou diminue) progressivement en passant d'une classe à l'autre.
Remarque : en augmentant indéfiniment le nombre de classes (et en diminuant l'amplitude de chaque classe) on obtient une courbe limite, f, qui représente la densité de la distribution des individus. Elle correspondra a une densité de probabilité. En effet le produit $f(x)dx$ peut s'interpréter comme la probabilité que la valeur du caractère d'un individu pris au hasard dans la population, soit contenue dans l'intervalle $[x, x+dx[$. Notons cependant qu'il s'agit là, de ce qu'on peut appeler, une expérience par la pensée, car pour augmenter indéfiniment le nombre de classes il faudrait que la population soit de taille infinie ! Cette interprétation sera pourtant très utile puisqu'elle fait le lien entre la statistique descriptive et la théorie des probabilités et qu'elle ouvre la voie à la statistique mathématique.
Pour les fréquence relatives cumulées, on construit les points de la fonction de répartition

$$F_X(x) = \frac{\text{card}\{\omega \in \Omega, X(\omega) \leq x\}}{\text{card}\,\Omega}$$

au points d'abscisse $e_1, \ldots, e_k$, et on rejoint ces points par des segments de droite, ce qui correspond à l'hypothèse d'une répartition uniforme à l'intérieur de la chaque classe.

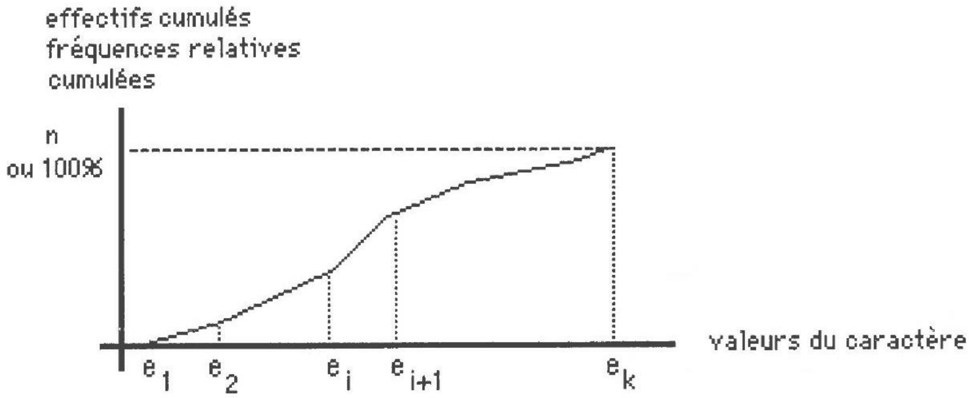


fig 4 Courte des fréquences cumulées : fonction de répartition

IV Paramètres de position.

On appelle paramètre de position, un nombre, qui permet de situer les valeurs de la série statistique. Il est difficile de chercher à en donner une définition plus précise, mais on peut remarquer que tous les paramètres qui seront définis ont en commun la propriété suivante : Si toutes les valeurs de la série subissent une même translation, alors le paramètre de position subit également cette translation. Cette propriété pourrait servir de définition tant elle parait naturelle pour un nombre qui doit résumer la position de la série

Définition

La médiane est la valeur du caractère telle que l'effectif des valeurs inférieures et celui des valeurs supérieures soient égaux. Cette définition, dans certains cas, est ambiguë et ne permet pas de calculer la médiane, aussi lui substitue-t-on la définition suivante :
La médiane (median) est la valeur μ du caractère, X, telle que l'effectif des valeurs inférieures et celui des valeurs supérieures soient au moins la moitié de l'effectif total.
La médiane doit donc vérifier la double condition suivante :

$$\operatorname{Card}\{\omega \in \Omega / \mathrm{X}(\omega) \leq \mu\} \geq \frac{1}{2} \operatorname{Card} \Omega, \operatorname{Card}\{\omega \in \Omega / \mathrm{X}(\omega) \geq \mu\} \geq \frac{1}{2} \operatorname{Card} \Omega$$

Remarque

Dans certain cas, un intervalle entier vérifie la définition on parle d'intervalle médian ou alors on prend le milieu de l'intervalle.

1 Détermination de la médiane :

1 .1 caractère discret :

Si l'effectif total de la population est impair égal à $2p + 1$, la médiane est la valeur du caractère de $(p + 1)^{\text{ème}}$ individu (les individus étant classés par valeurs croissantes du caractère).
Si l'effectif total de la population est pair égal à $2p$, toute valeur du caractère comprise entre celle du $p^{\text{ème}}$ individu et celle du $(p + 1)^{\text{ème}}$ individu répond à la définition.

1 .2 Caractère continu :

La médiane est la valeur du caractère correspondant à la fréquence relative cumulée de 0,5 . On peut la déterminer graphiquement ou par interpolation linéaire en faisant à chaque fois l'hypothèse de répartition uniforme à l'intérieur de chaque classe.
L'avantage majeur de la médiane réside dans ce qu'elle est peu sensible aux valeurs aberrantes (outfiers). Son inconvénient majeur vient de ce qu'elle n'a pas d'expression analytique.

Définition

On appelle quartiles les trois nombres qui partagent l'effectif total dans les rapports respectifs $\frac{1}{4}; \frac{2}{4}; \frac{3}{4}$. Plus précisément le 1er quartile Q_1 (Lower quartile) et le troisième quartile Q_3 (upper quartife) sont définis par les relations :

$$\begin{array}{ll} \operatorname{Card}\{\omega \in \Omega / \mathrm{X}(\omega) \leq \mathrm{Q}_1\} \geq \frac{1}{4} \operatorname{Card}\Omega & \operatorname{Card}\{\omega \in \Omega / \mathrm{X}(\omega) \geq \mathrm{Q}_1\} \geq \frac{3}{4} \operatorname{Card}\Omega \\ \operatorname{Card}\{\omega \in \Omega / \mathrm{X}(\omega) \leq \mathrm{Q}_3\} \geq \frac{3}{4} \operatorname{Card}\Omega & \operatorname{Card}\{\omega \in \Omega / \mathrm{X}(\omega) \geq \mathrm{Q}_3\} \geq \frac{1}{4} \operatorname{Card}\Omega \end{array}$$

Remarque

Le second quartile Q_2 n'est autre que la médiane. On définit aussi les déciles et plus généralement les fractiles (percentiles).

Définition

On appelle moyenne arithmétique (average, mean) $\overline{\mathrm{x}}$ d'une série statistique associée au caractère X, le nombre définit par :

$$\overline{x} = \frac{1}{\operatorname{Card}\Omega} \sum_{\omega \in \Omega} \mathrm{X}(\omega).$$

Remarque

Dans le cas ou plusieurs valeurs sont groupées (cf. présentation des données), on a

$$\overline{x} = \frac{1}{n} \sum_{1}^{k} n_i x_i, \quad n = \sum_{1}^{k} n_i$$

Les $\mathrm{x_i}$ représentent les valeurs du caractère dans le cas discret et les centres de classes dans le cas continu. Mais dans ce dernier cas on obtient en réalité une approximation de la moyenne.

Remarques

La moyenne arithmétique a grosso modo les mêmes propriétés que l'espérance.
On définit aussi la moyenne géométrique (geometric mean) et la moyenne harmonique.
La moyenne géométrique intervient en finance (notons que la moyenne géométrique ne satisfait pas au critère lié à la translation, énoncé au début).
La moyenne harmonique intervient par exemple pour le calcul de vitesse moyenne.
Ces deux dernières sont néanmoins peu utilisées dans le cadre des statistiques.

Définition

On appelle mode (resp. classe modale) la valeur du caractère (resp. la classe) correspondant à l'effectif le plus grand.

Remarque

Une série peut être unimodale ou plurimodale selon qu'il y a un ou plusieurs maximums relatifs.

V Paramètres de dispersion.

Deux séries statistiques peuvent avoir la même moyenne et deux distributions tout à fait différentes. La figure 6 montre le cas de deux séries, dont l'une à des valeurs très regroupées autour de la moyenne, alors que les valeurs de l'autre sont en général très éloignées de la valeur moyenne. Un paramètre de dispersion est un nombre attaché à la série, dont la valeur est d'autant plus grande que la série est plus dispersée. Il est naturel que ce nombre soit multiplié par K, lorsque chaque valeur de la série est elle même multipliée par K. En revanche un paramètre de dispersion doit rester invariable lorsque toutes les valeurs du caractère subissent une même translation. Ces propriétés sont partagées par les paramètres que nous allons définir

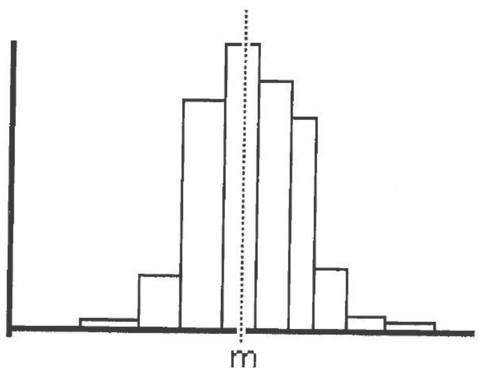
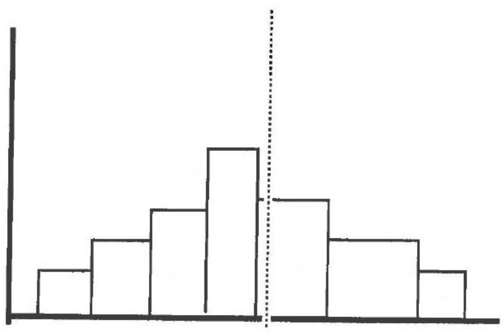


fig 6 Deux distributions de dispersion différentes et de même moyenne.



Définition

On appelle étendue (range) la distance entre le maximum et le minimum de la série.

Définition

L'intervalle interquartile (interquartile range) est la distance $\mathrm{Q}_3 - \mathrm{Q}_1$.

Remarque

Elle correspond à un intervalle qui contient 50% de la population et laisse 25% à gauche et 25% à droite. L'intervalle interquartile n'est pas affecté par les valeurs extrêmes éventuellement aberrantes.

Définition

L'écart absolu moyen par rapport à la médiane μ, est la moyenne arithmétique des écarts à la médiane.

$$\mathrm{e}_\mu = \frac{1}{n}\sum_{i=1}^{n} |x_i - \mu|$$

Remarque

on peut définir un écart absolu par rapport à n'importe quel nombre u en posant

$$\mathrm{e}_u = \frac{1}{n}\sum_{\mathrm{i}=1}^{n} |x_i - u|$$

On peut alors montrer que la médiane est la valeur de u pour laquelle e_u est minimale.

Définitions

On appelle écart quadratique moyen ou écart type (standard deviation) le nombre σ défini par la relation

$$\sigma^2 = \frac{1}{n}\sum_{i=1}^{n} (x_i - \bar{x})^2 = \left(\frac{1}{n}\sum_{i=1}^{n} x_i^2\right) - \bar{x}^2.$$

Le nombre σ^2 s'appelle la variance (variance), noté V ou Var, d'où :

$$\sigma = \sqrt{V}$$

Remarque

On peut montrer que $\bar{x}$ est la valeur de a pour laquelle $\frac{1}{n}\sum_{i=1}^{n} (x_i - a)^2$ est minimale.

Remarque

Si des valeurs ont été regroupées, on utilisera des moyennes pondérées :

$$\sigma^2 = \frac{1}{n}\sum_{i=1}^{k} n_i (x_i - \bar{x})^2 = \left(\frac{1}{n}\sum_{i=1}^{k} n_i x_i^2\right) - \bar{x}^2, \quad n = \sum_{1}^{k} n_i.$$

On ne confondra pas l'écart type avec l'erreur standard (standard error) qui vaut $\frac{\sigma}{\sqrt{n}}$.

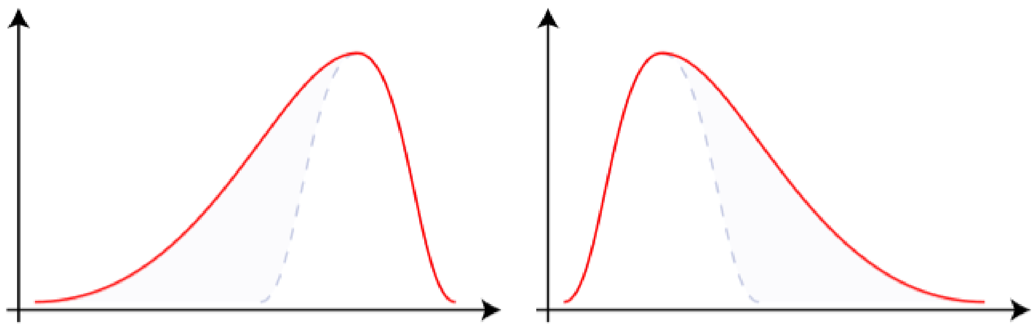
VI Coefficient d'asymétrie et d'aplatissement

Définition

On appelle coefficient d'asymétrie (skewness en anglais) d'une série statistique le réel γ défini par :

$$\gamma = \frac{1}{n}\sum_{i=1}^{n}\left(\frac{x_i - \overline{x}}{\sigma}\right)^3$$

Exemple



La distribution de gauche aura un coefficient asymétrie négatif alors que celle de droite aura un coefficient d'asymétrie positif.

Remarques

Deux série peuvent avoir les même moyennes, les mêmes écart type et pourtant des coefficients d'asymétries différents.
Karl Pearson a proposé d'autres estimations de l'asymétrie par des calculs plus simples1, ne faisant pas appel aux moments mais à d'autres paramètres statistiques :

- Avec la moyenne et la médiane :

$$\frac{\overline{x} - \text{med}}{\sigma}$$

- Avec les quartiles :

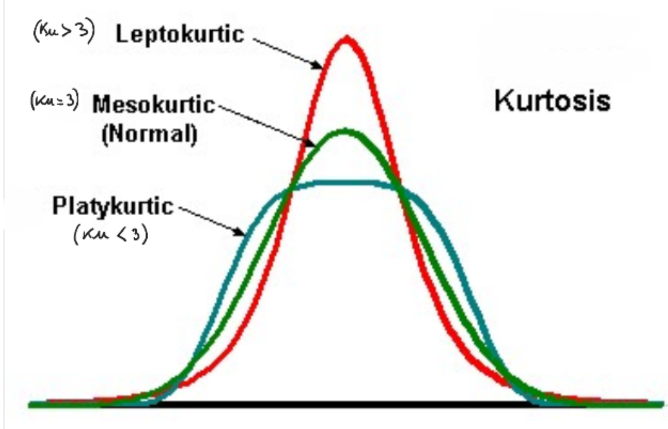
$$\frac{\frac{Q_3 - Q_1}{2} - Q_2}{\frac{Q_3 - Q_1}{2}}$$

Définition

On appelle coefficient d'aplatissement de kurtosis d'une série statistique le réel Ku défini par :

$$Ku = \frac{1}{n}\sum_{i=1}^{n}\left(\frac{x_i - \overline{x}}{\sigma}\right)^4$$

Exemple



Remarque

Certain auteur préfère parler du coefficient d'excès de kurtosis, qui va comparer la distribution avc la distribution normale.
Dans ce cas, une distribution pointue aura un coefficient d'excès de kurtosis positif, la distribution normale, un coefficient d'excès de kurtosis nul et une distribution très plate un coefficient d'excès de kurtosis négatif.

VII Remarques sur les logiciels statistiques.

Toutes ces statistiques, c'est à dire les différentes valeurs numériques associées à une population (max., min., moyenne, etc.), sont disponibles dans la plupart des logiciels de traitement statistique. En particulier les valeurs maximales et minimales ainsi que les quartiles sont matérialisés par des schémas que l'on appelle boite à moustaches (Box-and-Whiskers plot). Il faut prendre garde que dans Statgraphics plus, par exemple, l'écart-type fourni est en fait l'écart type estimé par l'estimateur sans biais s^2 de σ^2 (cf. p 26).
Il faut noter que dans le contexte des logiciels, on appelle souvent observation, ce que nous avons appelé individu, et variable, ce que nous avons appelé caractère. Les données sont présentées sous formes de tableau où chaque ligne correspond à une observation (i.e. un individu) et chaque colonne à une variable. Pour des variables numérotées $\mathrm{X}_1, \mathrm{X}_2, \ldots$, on trouvera à l'intersection de la i-ème ligne et de la colonne X_j, la valeur du j-ème caractère (ou j-ème variable) X_j pour le i-ème individu c'est à dire $\mathrm{X_j(i)}$.
I1 ne faut pas confondre ce genre de tableau avec ceux que l'on a construit aux $\S 1-1$ et $\S 1-2$: les logiciels ne "travaillent" pas en général sur des données regroupées. Ceci vient de ce que les données sont saisies individuellement et qu'il coûte en général peu de les transmettre entièrement

telles quelles. En d'autres termes les logiciels "travaillent" sur les données brutes. Mais ils existe des procédures qui construisent ces tableaux synthétiques.
Il faut aussi noter que les variables qualitatives doivent être "codées" et que les "camemberts" (pie-chart) qui sont des disques découpés en secteurs, n'ont de sens que pour les variables qualitatives. Chaque secteur représente la proportion d'individus ayant une modalité donnée du caractère qualitatif. Mais si l'on tente d'obtenir un camembert avec une variable quantitative continue on obtient une jolie figure qui a peu de sens et strictement aucun intérêt ! Le cas d'un caractère discret est intermédiaire et peut éventuellement faire l'objet d'une "camemberisation".

VIII Exercices

Exercice 1

On considère la série suivantes :

classes	effectifs
$120 \leqslant x < 150$	3300
$150 \leqslant x < 160$	2000
$160 \leqslant x < 170$	3500
$170 \leqslant x < 190$	4000
$190 \leqslant x < 230$	3000

1. Déterminer la moyenne de cette série.
2. Tracer l'histogramme de cette série.
3. Déterminer le premier quartile, la médiane ainsi que le troisième quartile par interpolation linéaire.
4. Déterminer l'écart type.

Exercice 2

oyenne harmonique.
Un cycliste fait l'entraînement suivant sur une boucle de x km :

- 1 boucle à 20 km/h
- 1 boucle à 40 km/h
- 1 boucle à 10 km/h

Déterminer la vitesse moyenne de ce cycliste.

Exercice 3

Soient $n \in \mathbb{N}^*$ et $x_1, x_2, ..., x_n$, n nombres réels.
Soit f la fonction définie par :

$$f : \begin{array}{rcl} \mathbb{R} & \to & \mathbb{R} \\ x & \mapsto & \dfrac{1}{n}\displaystyle\sum_{k=1}^{n} |x - x_k| \end{array}$$

Déterminer le minimum de f sur $\mathbb{R}$ et en quelles valeurs ce minimum est atteint.

Exercice 4

Soient $n \in \mathbb{N}^*$ et $x_1, x_2, ..., x_n$, n nombres réels.
Soit f la fonction définie par :

$$f : \begin{array}{rcl} \mathbb{R} & \to & \mathbb{R} \\ x & \mapsto & \sqrt{\dfrac{1}{n}\displaystyle\sum_{k=1}^{n}(x - x_k)^2} \end{array}$$

Déterminer le minimum de f sur $\mathbb{R}$ et en quelles valeurs ce minimum est atteint.

Exercice 5

Soient $x_1, x_2, ..., x_n$ les n valeurs d'une série statistique et $\overline{x}$ la moyenne arithmétique de cette série.

1. Démontrer que :
$$\frac{1}{n}\sum_{k=1}^{n}(\overline{x} - x_k)^2 = \frac{1}{n}\sum_{k=1}^{n}x_k^2 - \overline{x}^2$$
2. Application.
Un test dans 10 lycées différents donne les résultats suivants :
   - Moyenne : 11
   - Ecart type : 4
   - Nombre total d'élève ayant participé au test : 900

   100 élèves d'un 11$^{\text{ième}}$ lycée souhaite aussi participer au test. Leurs résultats sont les suivants :
   - Moyenne : 16
   - Ecart type : 2

   Déterminer la moyenne et l'écart type de l'ensemble des élèves des 11 lycée.

Exercice 6

On considère la série suivantes :

classes	effectifs
$0 \leqslant x < 10$	100
$10 \leqslant x < 30$	100
$30 \leqslant x < 35$	50
$35 \leqslant x < 50$	300

1. Déterminer la moyenne de cette série.
2. Tracer l'histogramme de cette série.
3. Déterminer le premier quartile, la médiane ainsi que le troisième quartile par interpolation linéaire.
4. Déterminer l'écart type.

Exercice 7

On considère les séries statistiques $x_1, x_2, \ldots, x_n$ et $y_1, y_2, \ldots, y_n$ donnant les résultats de deux caractères X et Y d'une population de 30 individus tels que pour tout individu ω_i, on a : $X(\omega_i) = x_i$ et $Y(\omega_i) = y_i$.
De plus, on donne :

$$\sum_{i=1}^{30} x_i = -30 \quad ; \quad \sum_{i=1}^{30} y_i = 60 \quad ; \quad \sum_{i=1}^{30} x_i^2 = 300 \quad ; \quad \sum_{i=1}^{30} y_i^2 = 600$$

$$\sum_{i=1}^{30} (y_i - \overline{y})^3 = -180 \quad ; \quad \sum_{i=1}^{30} (y_i - \overline{y})^4 = 3000$$

1. Calculer $\overline{x}$, $\overline{y}$, l'écart type relatif au caractère X, noté σ_x et l'écart type relatif au caractère Y noté σ_y.
2. Calculer les coefficients d'asymétrie et d'aplatissement de kurtosis de la série $(y_i)_{1 \leqslant i \leqslant n}$.
3. faire un schéma simplifier de la distribution du caractère Y par rapport à une distribution normale.

Chapitre 2

Régression linéaire

I Introduction

Dans cette partie, on considère une population Ω et deux caractères X et Y définis sur ω.
C'est à dire que l'on associe à chaque individu ω les valeurs de leurs caractères $X(\omega$ et $Y(\omega)$.
Par exemple le poids et la taille dans la population des pays de la Loire. de personne.
L'objectif est alors de déterminer s'il existe un lien "linéaire" entre ses deux caractères, et de connaître la "force" de ce lien. Autrement dit, existe-t-il deux réels a et b tels que $Y(\omega) \approx aX(\omega) + b$.
En considérant une population finie de n individus, on notera :

$$X(\Omega) = \{x_1, x_2, \dots, x_n\} \quad \text{et } Y(\Omega) = \{y_1, y_2, \dots, y_n\}$$

de tel que pour tout individu ω_i de Ω, on ait :

$$X(\omega_i) = x_i \quad \text{et} \quad Y(\omega_i) = y_i$$

II Méthode des moindres carrés

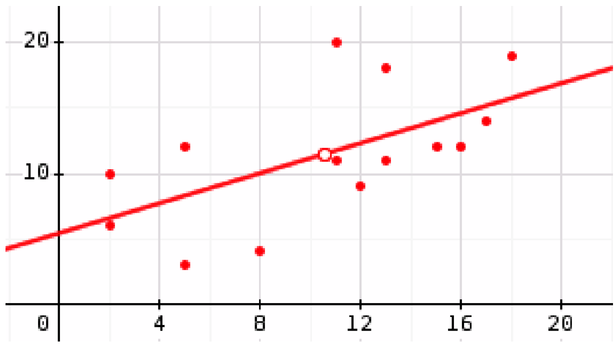
Par exemple, peut-on obtenir une bonne approximation de l'évolution du chiffre d'affaire en connaissant l'investissement publicitaire d'une entreprise ?

L'idée est de déterminer la "droite moyenne" d'un nuage de points.
Comme on l'a vu précédemment, la moyenne est la valeur qui minimise la somme des distances au carré.
C'est à dire que $\overline{x}$ est la valeur qui minimise la fonction :

$$\begin{aligned} f : \mathbb{R} &\to \mathbb{R} \\ x &\mapsto \sum_{w \in \Omega} \big(x - X(\omega)\big)^2 \end{aligned}$$

Dans la même idée, on cherche à trouver les coefficients $\widehat{a}$ et $\widehat{b}$ de la droite d'équation $y = \widehat{a}x + \widehat{b}$ qui va minimiser la somme :

$$\sum_{w \in \Omega} \left(Y(\omega) - \widehat{a}X(\omega) - \widehat{b}\right)^2 = \sum_{i=1}^{n} \left(y_i - \widehat{a}x_i - \widehat{b}\right)^2$$



Définition

On considère une population Ω d'individus avec deux caractères X et Y.
On appelle droite de régression linéaire de y en fonction de x la droite d'équation :

$$y = \widehat{a}x + \widehat{b}$$

avec :

$$\widehat{a} = \frac{\dfrac{1}{\mathrm{card}(\Omega)}\left(\displaystyle\sum_{w\in\Omega} X(\omega)Y(\omega)\right) - \overline{X}.\overline{Y}}{V(X)} \quad \text{et} \quad \widehat{b} = \overline{Y} - \widehat{a}\overline{X}$$

ou encore, dans le cas d'une population de n individus :

$$\widehat{a} = \frac{\dfrac{1}{n}\left(\displaystyle\sum_{i=1}^{n} x_i y_i\right) - \overline{x}.\overline{y}}{V_x} \quad \text{et} \quad \widehat{b} = \overline{y} - \widehat{a}\overline{x}$$

Démonstration

Soit f la fonction définie par :

$$\begin{aligned} f : \mathbb{R}^2 &\to \mathbb{R} \\ (a,b) &\mapsto \sum_{i=1}^{n} (y_i - ax_i - b)^2 \end{aligned}$$

$$\begin{aligned} \frac{\partial f}{\partial b}(a,b) = 0 &\Leftrightarrow \sum_{i=1}^{n} -2\,(y_i - ax_i - b) = 0 \\ &\Leftrightarrow -2\left(\sum_{i=1}^{n} y_i - a\sum_{i=1}^{n} x_i - nb\right) = 0 \\ &\Leftrightarrow \overline{y} - a\overline{x} = b \end{aligned}$$

D'autre part :

$$\begin{aligned} \frac{\partial f}{\partial a}(a,b) = 0 &\Leftrightarrow \sum_{i=1}^{n} -2x_i\,(y_i - ax_i - b) = 0 \\ &\Leftrightarrow -2\left(\sum_{i=1}^{n} x_i y_i - a\sum_{i=1}^{n} x_i^2 - b\sum_{i=1}^{n} x_i\right) = 0 \end{aligned}$$

$$\Leftrightarrow \left(\sum_{i=1}^{n} x_i y_i - a\sum_{i=1}^{n} x_i^2 - b\sum_{i=1}^{n} x_i\right) = 0$$
$$\Leftrightarrow a\sum_{i=1}^{n} x_i^2 + b\sum_{i=1}^{n} x_i = \sum_{i=1}^{n} x_i y_i$$

d'où

$$\begin{cases} \dfrac{\partial f}{\partial b}(a,b) = 0 \\ \dfrac{\partial f}{\partial a}(a,b) = 0 \end{cases} \Leftrightarrow \begin{cases} b = \overline{y} - a\overline{x} \\ a\displaystyle\sum_{i=1}^{n} x_i^2 + b\sum_{i=1}^{n} x_i = \sum_{i=1}^{n} x_i y_i \end{cases}$$
$$\Leftrightarrow \begin{cases} b = \overline{y} - a\overline{x} \\ a\displaystyle\sum_{i=1}^{n} x_i^2 + (\overline{y} - a\overline{x})\sum_{i=1}^{n} x_i = \sum_{i=1}^{n} x_i y_i \end{cases}$$
$$\Leftrightarrow \begin{cases} b = \overline{y} - a\overline{x} \\ a\left(\displaystyle\sum_{i=1}^{n} x_i^2 - \overline{x}\sum_{i=1}^{n} x_i\right) = -\overline{y}\sum_{i=1}^{n} x_i + \sum_{i=1}^{n} x_i y_i \end{cases}$$
$$\Leftrightarrow \begin{cases} b = \overline{y} - a\overline{x} \\ a\left(\dfrac{1}{n}\displaystyle\sum_{i=1}^{n} x_i^2 - \overline{x}^2\right) = -\overline{x}.\overline{y} + \dfrac{1}{n}\sum_{i=1}^{n} x_i y_i \end{cases}$$
$$\Leftrightarrow \begin{cases} b = \overline{y} - a\overline{x} \\ a = \dfrac{\dfrac{1}{n}\displaystyle\sum_{i=1}^{n} x_i y_i - \overline{x}.\overline{y}}{V_x} \end{cases}$$

Pour compléter cette démonstration, il faut que la forme quadratique associée à la matrice hessienne de f soit strictement négative sur $\mathbb{R}^2$ ou strictement négative sur $\mathbb{R}^2$, c'est à dire que les valeurs propres de cette matrice hessienne soient toutes strictement négatives, ou toutes strictement positives. Mais ceci est complètement hors programme et pourra éventuellement être vus au second semestre.

Définition

On considère une population Ω d'individus avec deux caractères X et Y.
On appelle covariance des deux caractère X et Y le réel, note $\operatorname{cov}(X,Y)$ défini par :

$$\operatorname{cov}(X,Y) = \frac{1}{\operatorname{card}(\Omega)}\left(\sum_{w\in\Omega} X(\omega)Y(\omega)\right) - \overline{X}.\overline{Y}$$

ou encore, dans le cas d'une population de n individus :

$$\operatorname{cov}(X,Y) = \frac{1}{n}\left(\sum_{i=1}^{n} x_i y_i\right) - \overline{x}.\overline{y}$$

Remarque

Dans le cas de la définition précédente, on a :

$$\widehat{a} = \frac{\mathrm{cov}\,(X, Y)}{V(X)}$$

III Coefficient de corrélation linéaire

Définition

On appelle coefficient de corrélation linéaire de deux caractères X et Y d'une population le réel noté r défini par :

$$r = \frac{\mathrm{cov}(X, Y)}{\sqrt{V(X) \times V(Y)}}$$

Remarque

Le coefficient de corrélation linéaire (ou coefficient de Bravais-Pearson) permet de déterminer le sens de la relation entre deux caractères d'une même population ainsi que le degré de la relation

Propriété

On considère une population Ω d'individus avec deux caractères X et Y. (avec X non constant sur Ω)
Soit r le coefficient de corrélation linéaire de X et Y.
$|r| = 1$ si et seulement si il existe $(a, b) \in \mathbb{R}^2$ tel que $Y = aX + b$.

Définition

On considère une population Ω d'individus avec deux caractères X et Y.
Soit $y = \widehat{a}x + \widehat{b}$ l'équation de la droite de régression linéaire de y en fonction de x.

- On appelle résidu de l'observation i le réel e_i défini par :

$$e_i = y_i - (\widehat{a}x_i + \widehat{b})$$

- On appelle Somme des carrés des résidus le réel SCR défini par :

$$SCR = \sum_{i=1}^{n} \left(y_i - \widehat{a}x_i - \widehat{b}\right)^2$$

IV Exercices

Exercice 8

On considère les séries statistiques $x_1, x_2, \ldots, x_n$ et $y_1, y_2, \ldots, y_n$ donnant les résultats de deux caractères X et Y d'une population de 30 individus tels que pour tout individu ω_i, on a : $X(\omega_i) = x_i$ et $Y(\omega_i) = y_i$.
De plus, on donne :

$$\sum_{i=1}^{30} x_i = -30 \quad ; \quad \sum_{i=1}^{30} y_i = 60 \quad ; \quad \sum_{i=1}^{30} x_i^2 = 300 \quad ; \quad \sum_{i=1}^{30} y_i^2 = 600$$

$$\sum_{i=1}^{30} x_i y_i = -240 \quad ; \quad \sum_{i=1}^{30} (y_i - \overline{y})^3 = -180 \quad ; \quad \sum_{i=1}^{30} (y_i - \overline{y})^4 = 3000$$

1. Calculer $\overline{x}$, $\overline{y}$, l'écart type relatif au caractère X, noté σ_x et l'écart type relatif au caractère Y noté σ_y.
2. Déterminer la covariance $\mathrm{Cov}(X, Y)$ puis le coefficient de corrélation linéaire des deux caractère X et Y.
3. Donner, en fonction de $\mathrm{Cov}(X, Y)$ et $\mathrm{Var}(X, Y)$ les coefficients $\widehat{a}$ et $\widehat{b}$ de la droite de régression linéaire de y en fonction de x.
   Donner des valeurs approchées à 10^{-2} de $\widehat{a}$ et $\widehat{b}$.
4. Donner une approximation à 1% du pourcentage de variation du caractère Y par rapport au caractère X.

Exercice 9

Un hypermarché dispose de 20 caisses. On s'intéresse au temps moyen d'attente en fonction du nombre de caisses ouvertes un jour de semaine.
Soient X et Y les caractères associant respectivement à chaque jour d'ouverture i le nombre x_i de caisse ouverte, et temps d'attente y_i. Le tableau ci-dessous donne le nombre x_i de caisses ouvertes et le temps moyen d ?attente y_i correspondant.

Jour	1	2	3	4	5	6	7
X	3	4	5	6	8	10	12
Y	16	12	9,6	7,9	6	4,7	4

1. Déterminer le coefficient de corrélation linéaire r.
2. Montrer que le coefficient de régression linéaire de Y en X est $\widehat{a} = -0.7067$.
3. En déduire une approximation de la réduction du temps d'attente à chaque ouverture d'une caisse supplémentaire.
4. Déterminer les coefficient de la droite de régression linéaire.
5. Selon vous, y a-t-il causalité entre le nombre de caisses et le temps d'attente.
6. Faites une estimation du temps d'attente si on ouvre qu'une caisse, 7 caisses, les 20 caisses.

Exercice 10

On a relevé pour différents pays le PIB par habitant en 2004, X (en dollars) et le taux brut de scolarisation des moins de 24 ans la même année Y (en pourcentage).
Les résultats sont les suivants :

Pays	PIBX	Taux de scolarisation$Y$
Pays en d´eveloppement	4775	63
Pays les plus pauvres	1350	45
Pays arabes	5680	62
Asie de l ?Est et Pacifique	5872	69
Amérique latine et Caraïbes	7964	81
Asie du Sud	3072	56
Afrique Sub-saharienne	1942	50
Europe centrale,orientale et CEI	8802	83

$$\sum x_i = 39457 \quad ; \quad \sum y_i = 509 \quad ; \quad \sum x_i^2 = 245474957 \quad ; \quad \sum y_i^2 = 33685$$
$$\sum x_i y_i = 2763685$$

1. On cherche à expliquer le taux de scolarisation en fonction du PIB. Identifier la variable à expliquer et la variable variable explicative. Pour chaque variable calculer la moyenne observée et la variance observée.
2. Expliquer l'objectif de la régression linéaire simple.
3. Donner l'équation du modèle théorique.

Exercice 11

Soit X et Y deux caractères définis sur une même population.

1. Démontrer que s'il existe deux réels a et b tels que $Y = aX + b$, alors le coefficient de corrélation linéaire r vérifie : $|r| = 1$
2. Démontrer la réciproque.

Chapitre 3

Rappel de probabilité.

I Loi de probabilité, variable aléatoire.

Définition

Soit E un ensemble, on appelle tribu sur E un sous ensemble $\mathscr{B}$ de l' ensemble des parties de E tel que :

- $\varnothing \in \mathscr{B}$ et $\mathrm{E} \in \mathscr{B}$.
- si $\mathrm{A} \in \mathscr{B}$ alors $\mathrm{A}^{c} \in \mathscr{B}$
- si $(A_n)_{n\in IN}$ est une suite d'éléments de $\mathscr{B}$ alors on a : $\underset{n\in\mathbb{N}}{\cup} A_n \in \mathscr{B}$ et $\underset{n\in IN}{\cap} A_n \in \mathscr{B}$.

Définitions

On appelle espace de probabilité, un triplet $(\mathrm{E}, \mathcal{A}, \mathrm{P})$ où E est un ensemble, $\mathcal{A}$ une tribu sur E et P une probabilité sur $(\mathrm{E}, \mathcal{A})$ c.a.d. une application de $\mathcal{A}$ dans $[0,1]$ vérifiant les axiomes suivants.

- $\mathrm{P}(\emptyset) = 0$.
- Si $(A_n)_{n\in IN}$ est une suite d'éléments de $\mathcal{A}$ tels que $\mathrm{A_i} \cap \mathrm{A_j} = \varnothing$, pour $i \neq j$ alors :

$$P\left(\bigcup_{n\in\mathbb{N}} A_n\right) = \sum_{n\in\mathbb{N}} P\left(A_n\right)$$

Un événement est un ensemble **A** appartenant à la tribu $\mathcal{A}$.

Définitions

Soit $(\mathrm{E}, \mathcal{A}, \mathrm{P})$ un espace de probabilité, on appelle variable aléatoire réelle (v.a.r.) une application X, de E dans $\mathbb{R}$, telle que pour tout intervalle I contenu dans $\mathbb{R}$, on ait :

$$\mathrm{X}^{-1}(I) \in \mathcal{A}$$

- On dira que X est une variable aléatoire discrète si $X\left(\Omega\right)$ est fini ou dénombrable.
- On dira que X est une variable aléatoire continue si $X\left(\Omega\right)$ est infini et non dénombrable.

Remarque

La condition imposée est une condition technique qui relève de la théorie de la mesure En pratique il suffit de retenir qu'une variable aléatoire réelle est une application définie sur E. Nous rencontrerons aussi des applications définies sur E valeurs dans $\mathbb{R}^n$ (par opposition à $\mathbb{R}$). Dans ce cas on parle de vecteur aléatoire. Les variables aléatoires définies sur un espace $(\mathrm{E}, \mathcal{A}, \mathrm{P})$ permettent de construire des événements.

Exemple

Supposons par exemple que l'on veuille modéliser le lancer de deux dés. On pourra considérer que E est l'ensemble des couples d'entiers compris entre 1 et 6 :

$$\mathrm{E} = \{1, 2, 3, 4, 5, 6\} \times \{1, 2, 3, 4, 5, 6\} = \{(1, 1), (1, 2), \ldots, (6, 6)\}.$$

Si X associe à chaque issue le résultat du premier lancer et Y le résultat du second, alors X et Y sont deux v.a. définies sur E. Par exemple les événements "le résultat du second lancer est pair" et "les deux lancers sont identiques" s'écrivent $\{\mathrm{Y} \in \{2, 4, 6\}\}, \{\mathrm{X} = \mathrm{Y}\}$. Ainsi quand on voudra désigner la probabilité de réalisation de ces événements, on écrira $\mathrm{P}(\mathrm{Y} \in \{2, 4, 6\})$ et $\mathrm{P}(\mathrm{X} = \mathrm{Y})$.

Définition

Soit $(\mathrm{E}, \mathcal{A}, \mathrm{P}\ )$ un espace de probabilité et X une v.a.r..
La famille des réels $\mathrm{P}(\mathrm{X} \in \mathrm{A})$ où A parcourt (la tribu engendrée par) les intervalles de $\mathbb{R}$, définit une probabilité sur $\mathbb{R}$ notée μ_{X}. Ainsi on appelle loi de X, et on note μ_{X}, la probabilité définie sur $\mathbb{R}$ par :

$$\mu_{\mathrm{X}}(\mathrm{A}) = \mathrm{P}(\mathrm{X} \in \mathrm{A})$$

Remarque

Deux cas se présentent en pratique :

1. 1^{er} cas : L'ensemble des valeurs prises par X est dénombrable (i.e. on peut l'indexer par $\mathbb{N}$), alors connaître la loi de probabilité de X, c'est connaître la famille des nombres :

$$P(X = x) \text{ pour tout } x \in X(\mathrm{E})$$

2. $2^{\text{ième}}$ cas : X peut prendre toutes les valeurs de $\mathbb{R}$. Il se peut alors qu'il existe une fonction f définie sur $\mathbb{R}$ telle que :

$$P(a \leq X \leq b) = \int_a^b f(x).dx$$

On dit alors que la loi de X admet f pour densité. L'existence d'une densité exprime que la probabilité pour que la variable aléatoire prenne ses valeurs dans un petit intervalle autour de x de longueur Δx est approximativement proportionnel à la longueur de cet intervalle et que le coefficient de proportionnalité est précisément $f(x)$.
Penser au cas d'une densité f constante et au cas où Δx est assez petit pour que les variations de f soit négligeables. On peut exprimer la relation précédente sous la forme plus intuitive

mais moins rigoureuse :

$$\mathrm{P}(x \leqslant \mathrm{X} \leqslant x + \Delta x) \approx f(x)\Delta x$$

Exemples

1. Supposons par exemple, que l'on jette plusieurs fois une pièce de monnaie équilibrée jusqu'à ce qu'elle tombe sur face. Soit T la variable aléatoire qui associe à chaque issue le nombre de lancer. T est donc une v.a. à valeurs dans $\mathbb{N}^*$ dont la loi est définie par :

$$\forall t \in \mathbb{N}^*, \mathrm{P}(\mathrm{T} = \mathrm{t}) = \left(\frac{1}{2}\right)^{\mathrm{t}}.$$

2. La durée de vie d'une ampoule électrique peut être modélisée par une v.a.r. ayant la densité

$$\begin{aligned} f : \quad & \mathbb{R} \to \mathbb{R} \\ x \mapsto & \begin{cases} 0 & \text{si } x < 0 \\ \theta e^{-\theta x} & \text{si } x \geqslant 0 \end{cases} \end{aligned}$$

pour un paramètre θ donné. Ce qui signifie que la probabilité pour que la durée de vie de l'ampoule soit comprise, par exemple, entre 2,3 et 7,8 unités de temps est donnée par :

$$\mathrm{P}(2,3 \leq \mathrm{T} \leq 7,8) = \int_{2.3}^{7.8} \theta \mathrm{e}^{-\theta x} dx.$$

Remarque

Il est possible qu'une v.a. prenant toutes les valeurs de $\mathbb{R}$ n'admette pas de densité. Nous n'en utiliserons pas cette année.

II Fonction de répartition

Définition

Soit $(\Omega, \mathcal{A}, P)$ un espace probabilisé et X une variable aléatoire définie sur Ω.
On appelle fonction de répartition de X la fonction F définie sur $\mathbb{R}$ par :

$$F(x) = P\left(X < x\right) = P\left(X^{-1}\big(\,]-\infty, x[\,\big)\right) = P\left(\{\omega \in \Omega | X(\omega) < x\}\right)$$

Remarques

- Si X est **discrète**, alors F est représentée par une fonction en escalier.
- Si X est **continue** et a pour densité f, alors :

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

Exemples

1. En reprenant l'exemple précédent, où l'on jette plusieurs fois une pièce de monnaie équilibrée jusqu'à ce qu'elle tombe sur face.
   T étant la variable aléatoire qui associe à chaque issue le nombre de lancer effectués.
   La fonction de répartition associée à T est définie par :
$$\begin{cases} \forall t \in]-\infty,1], F(t)=0 \\ \forall t \in]1,+\infty[, F(t)=\mathrm{P}\left(\mathrm{T}<t\right)=\displaystyle\sum_{k=1}^{\lceil t-1\rceil}\left(\frac{1}{2}\right)^{t} \end{cases}$$
2. Pour la durée de vie d'une ampoule, X étant la variable aléatoire qui associe à chaque ampoule sa durée de vie.
   X suit la loi exponentielle de paramètre λ, c'est à dire que sa fonction de répartition est la suivante :
$$\begin{cases} \forall t \in]-\infty,0], F(t)=0 \\ \forall t \in]0,+\infty[, F(t)=1-\mathrm{e}^{-\lambda t} \end{cases}$$

Propriétés

Soit F la fonction de répartition associée à une variable aléatoire X.

1. F est croissante
2. $\forall x \in \mathbb{R}, 0 \leqslant F(x) \leqslant 1$
3. $\lim\limits_{x\to-\infty} F(x)=0$ et $\lim\limits_{x\to+\infty} F(x)=1$

Démonstration

1. Pour tout réels x et y tels que $x<y$, on a $]-\infty,x[\,\subset\,]-\infty,y[$, donc :
   $\{X^{-1}(]-\infty,x[)\}\subset\{X^{-1}(]-\infty,y[)\}$, que l'on peut aussi écrire $\{(X<x)\}\subset\{(X<y)\}$, d'où :
$$P(X<x)\leqslant P(X<y)$$
2. En effet, pour tout réel x, $F(x)$ est une probabilité, donc
$$0\leqslant F(x)\leqslant 1$$
3. $\lim\limits_{x\to-\infty}]-\infty,x[=\emptyset$, donc $\lim\limits_{x\to-\infty} F(x)=0$.
   de même $\lim\limits_{x\to+\infty}]-\infty,x[\,=\mathbb{R}$, donc $\lim\limits_{x\to+\infty} X^{-1}(]-\infty,x[)=X^{-1}(\mathbb{R})=\Omega$, et par conséquent
$$\lim_{x\to+\infty} F(x)=P(\Omega)=1$$

Définition

Soient X une variable aléatoire définie sur $(\Omega,\mathcal{A})$ et F la fonction de répartition associée à la variable aléatoire X.
On appelle **densité de probabilité** de la variable aléatoire X toute fonction f, définie et

intégrable sur $\mathbb{R}$ telle que :

$$\forall x \in \mathbb{R}, F(x) = \int_{-\infty}^{x} f(t)dt$$

Remarque

Il existe des variables aléatoires qui n'ont pas de densité de probabilité

Propriété

Soit X une variable aléatoire définie sur $(\Omega, \mathcal{A})$ ayant pour densité de probabilité la fonction f.
Soit F la fonction de répartition associée à la variable aléatoire X.
Dans ce cas, si F est dérivable sur un intervalle I, alors :

$$\forall x \in I, f(x) = F'(x)$$

Exemple

Dans le cas de l'ampoule de l'exemple 2.
On admet que la variable aléatoire X associant à chaque ampoule sa durée de vie a une densité de probabilité f.
La fonction de répartition de X est la fonction F définie par :

$$\begin{cases} \forall t \in]-\infty, 0], F(t) = 0 \\ \forall t \in]0, +\infty[, F(t) = 1 - \mathrm{e}^{-\lambda t} \end{cases}$$

Donc F est dérivable sur $]-\infty, 0[$ et sur $]0, +\infty[$.
Sachant que X a pour densité de probabilité la fonction f, on a :

$$\forall x \in]-\infty, 0[, f(x) = F'(x) = 0$$
$$\forall x \in]0, +\infty[, f(x) = F'(x) = \lambda \mathrm{e}^{-\lambda x}$$

Propriétés

1. Une densité de probabilité est toujours positives
2. Une densité de probabilité est intégrable sur $\mathbb{R}$ et on a :

$$\int_{-\infty}^{+\infty} f(x)dx = 1$$

3. Pour tout réels a et b tels que $a \leqslant b$, on a :

$$P(X \in ([a, b]) = \int_{a}^{b} f(x)dx$$

Démonstration

1. En effet, f est la dérivée d'une fonction croissante
2. En effet, $\lim\limits_{x \to +\infty} F(x) = 1$
3. Soient a et b deux réels tels que $a \leqslant b$, on a :
   $F(b) = \int_{-\infty}^{b} f(x)dx$
   $\quad = \int_{-\infty}^{a} f(x)dx + \int_{a}^{b} f(x)dx$
   $\quad = F(a) + \int_{a}^{b} f(x)dx$
   donc
   $F(b) - F(a) = \int_{a}^{b} f(x)dx$
   et par conséquent
   $P(X \in [a,b]) = F(b) - F(a)$
   $\quad = \int_{a}^{b} f(x)dx$

III Indépendance, vecteur aléatoire.

Définition

n variables aléatoires réelles $X_1, X_2, \ldots, X_n$ sont indépendantes si et seulement si pour toutes les familles d'intervalles $\mathbf{I}_1, \ldots, \mathbf{I}_n$ de $\mathbb{R}$, on a :

$$\mathrm{P}(\mathrm{X}_1 \in \mathbf{I}_1, \ldots, \mathrm{X}_n \in \mathbf{I}_n) = \mathrm{P}(\mathrm{X}_1 \in \mathbf{I}_1) \times \ldots \times (\mathrm{X}_n \in \mathbf{I}_n)$$

Remarque

En pratique on sait a priori que des variables aléatoires sont indépendantes et on applique cette égalité.
Les n composantes $\mathrm{X}_1, \mathrm{X}_2, \ldots, \mathrm{X}_n$ d'un vecteur aléatoire sont, par définition, des variables aléatoires réelles. La notion de loi d'une variable aléatoire réelle se généralise aux vecteurs aléatoires. Mais il faut utiliser la notion d'intégrales multiples lorsque l'on veut généraliser la définition de la densité.

Exemple

1. 1^{er} cas :
   L'ensemble des valeurs prises par le vecteur aléatoire est dénombrable, par exemple $(\mathrm{X}_1, \mathrm{X}_2)$ de dimension deux pour fixer les idées !
   La loi du vecteur $(\mathrm{X}_1, \mathrm{X}_2)$ est alors la famille dénombrable des probabilités :
   $$\{\mathrm{P}(\mathrm{X}_1 = \mathrm{x}_1, \mathrm{X}_2 = \mathrm{x}_2)\,;\mathrm{x}_1 \in \mathrm{X}_1(\mathrm{E}), \mathrm{x}_2 \in \mathrm{X}_2(\mathrm{E})\}$$
   Il est ainsi possible de calculer la probabilité d'un événement quelconque A en effectuant la somme des probabilités des événements élémentaires qui le réalisent :
   $$\mathrm{P}\big((\mathrm{X}_1, \mathrm{X}_2) \in A\big) = \sum_{(x_1;x_2) \in A} \mathrm{P}(\mathrm{X}_1 = x_1, \mathrm{X}_2 = x_2)$$
2. $2^{\text{èm}}$ cas :

Le vecteur (X_1, X_2) prend toutes les valeurs de $\mathbb{R}^2$ et sa loi admet une densité f.
Cela signifie qu'il existe une fonction f de $\mathbb{R}^2$ dans $\mathbb{R}$ telle que :

$$P\big((X_1, X_2) \in A\big) = \iint_A f(x_1, x_2)\, dx_1 dx_2$$

Définition

Soient $X_1, X_2, \ldots, X_n$ n variables aléatoires admettant les densités $f_1, f_2, \ldots, f_n$.
Ces variables aléatoires sont indépendantes si et seulement si la loi du vecteur $(X_1, X_2, \ldots, X_n)$, admet la densité

$$f(x_1, x_2, \ldots, x_n) = f_1(x_1)\, f_2(x_2) \ldots f_n(x_n)$$

Exemple

IV Espérance, variance, covariance et coefficient de corrélation.

Nous donnons plusieurs définitions de l'espérance mathématique d'une variable aléatoire selon que cette variable est discrète ou continue. Mais on verra en 3 ème année que ce sont des cas particuliers d'une définition générale qui relève de la théorie de la mesure.

1 Espérance et variance

Définitions

1. Soit X une variable aléatoire discrète.
   On appelle espérance de X le nombre réel :
   $$\mathrm{E}(\mathrm{X}) = \sum_{x \in \mathrm{X}(\Omega)} x P(\mathrm{X} = x)$$
2. Soit X une variable aléatoire ayant une densité f.
   On appelle espérance de X le nombre réel :
   $$E(X) = \int_{X(\Omega)} x f(x) dx$$

Propriété : Loi de transfert

Soient X une variable aléatoire et h une application de $X(\Omega)$ dans $\mathbb{R}$.
$Y = h(X)$ définie donc une nouvelle variable aléatoire, et on a :

$$\mathrm{E}\big(h(\mathrm{X})\big) = \sum_{x \in \mathrm{X}(\Omega)} h(x) \mathrm{P}(\mathrm{X} = x), \quad \text{si X est une v.a. discrète,}$$

ou

$$E(h(X)) = \int_{X(\Omega)} h(x)f(x)dx, \text{ si X admet une densité } f.$$

Exemple

Si X est, par exemple, une variable aléatoire à valeurs dans $\mathbb{R}^+$ ayant une densité f, on obtient :

$$E(\sqrt{X}) = \int_{[0,\infty[} \sqrt{x}f(x)dx$$

Puisque X est à valeurs dans $\mathbb{R}^+$, la densité est nulle sur $\mathbb{R}^-$. Les bornes d'intégration sont 0 et $+\infty$.

Définition

Soit X une variable aléatoire.
La variance de la variable aléatoire X est le réel positif défini par :

$$\mathrm{V(X) = E\big((X - E(X))^2\big)}$$

Propriétés

Soient X et Y deux variables aléatoires définies sur un même univers Ω et a un réel. On a alors :

- $\mathrm{E}(a\mathrm{X} + \mathrm{Y}) = a\mathrm{E}(\mathrm{X}) + \mathrm{E}(\mathrm{Y})$
- $\mathrm{V}(\mathrm{X}) = \mathrm{E}\left(\mathrm{X}^2\right) - \mathrm{E}(\mathrm{X})^2$
- $\mathrm{V}(a\mathrm{X}) = a^2\mathrm{V}(\mathrm{X})$

Démonstration

2 Covariance et coefficient de corrélation linéaire

Définition

oient X et Y deux variables aléatoires définies sur un même univers Ω.
La covariance du couple (X, Y) est alors le réel défini par :

$$\mathbf{Cov}\,(X,Y) = \mathrm{E}\left(\big(X - \mathrm{E}(X)\big) \times \big(Y - \mathrm{E}(Y)\big)\right)$$

Exemple

Considérons le couple de variables aléatoires dont la loi est définie par le tableau suivant :

$Y \backslash X$	-1	0	1
-1	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{8}$
0	$\frac{1}{16}$	$\frac{1}{8}$	$\frac{1}{16}$
1	$\frac{1}{8}$	$\frac{1}{8}$	$\frac{1}{8}$

On a $E(X) = 0$ et $E(Y) = 0$.

$$\begin{aligned}
\text{Cov}\,(X,Y) &= \text{E}\left(\left(X - \text{E}(X)\right) \times \left(Y - \text{E}(Y)\right)\right)\\
&= \text{E}\left(XY\right)\\
&= \left((-1)\times(-1)\times\frac{1}{8}\right) + \left((-1)\times 0\times\frac{1}{8}\right) + \left((-1)\times 1\times\frac{1}{8}\right) + \left(0\times(-1)\times\frac{1}{16}\right) + \ldots\\
&= \left(1\times\frac{1}{8}\right) + \left((-1)\times\frac{1}{8}\right) + \left((-1)\times\frac{1}{8}\right) + \left(1\times\frac{1}{8}\right)\\
&= 0
\end{aligned}$$

Propriété

Soient X et Y deux variables aléatoires définies sur un même univers Ω.

$$\mathbf{Cov}\,(X,Y) = \text{E}(XY) - \text{E}(X)\text{E}(Y)$$

Démonstration

Propriété

Soient X et Y deux variables aléatoires définies sur un même univers Ω.

$$V(X+Y) = V(X) + V(Y) + 2\times\mathbf{Cov}\,(X,Y)$$

Démonstration

Propriété

Soient X et Y deux variables aléatoires **indépendantes** définies sur un même univers Ω. On a alors :

- $\text{E(XY)} = \text{E(X)E(Y)}$
- $\mathbf{Cov}\,(X,Y) = 0$
- $\text{V(X+Y)} = \text{V(X)} + \text{V(Y)}$

Démonstration

Soient X et Y deux variables aléatoires discrète et **indépendantes** définies sur un même univers Ω. On a alors :

- $$\begin{aligned}\text{E}(XY) &= \sum_{\omega\in\Omega} XY(w).P(w)\\ &= \sum_{\omega\in\Omega} X(w)Y(w).P(w)\end{aligned}$$

  or $\left(X^{-1}(x)\cap Y^{-1}(y)\right)_{(x,y)\in X(\Omega)\times Y(\Omega)}$ est une partition de l'univers, donc :

  $$\begin{aligned}\text{E}(XY) &= \sum_{(x,y)\in X(\Omega)\times Y(\Omega)} xy.P\left(\{\omega\in\Omega | X(\omega) = x \text{ et } Y(\omega) = y\}\right)\\ &= \sum_{(x,y)\in X(\Omega)\times Y(\Omega)} xy.P\left(\{X^{-1}(x)\cap Y^{-1}(y)\}\right)\end{aligned}$$

$$= \sum_{(x,y)\in X(\Omega)\times Y(\Omega)} xy.P\left(\{(X=x)\cap(Y=y)\}\right)$$
$$= \sum_{(x,y)\in X(\Omega)\times Y(\Omega)} xy.P\left(X=x\right).P\left(Y=y\right) \quad \text{car } X \text{ et } Y \text{ sont indépendantes.}$$
$$= \sum_{x\in X(\Omega)} x.P\left(X=x\right) \times \sum_{y\in Y(\Omega)} y.P\left(Y=y\right)$$
$$= \mathrm{E(X)E(Y)}$$

- D'après la propriété précédente :

$$\mathrm{Cov}\left(X,Y\right) = \mathrm{E}(XY) - \mathrm{E(X)E(Y)} = 0$$

- $V(X+Y) = V(X) + V(Y) + 2.\mathrm{Cov}\left(X,Y\right) = V(X) + V(Y)$

Danger

La réciproque de cette propriété est fausse, comme le montre l'exemple suivant :
...

Définition

Soient X et Y deux variables aléatoires définies sur un même univers Ω.
On appelle coefficient de corrélation linéaire le réel défini par :

$$\mathrm{Corr}\left(X,Y\right) = \frac{\mathrm{Cov}\left(X,Y\right)}{\sqrt{V(X)}\sqrt{V(Y)}}$$

Propriétés

- $-1 \leqslant \mathrm{Corr}\left(X,Y\right) \leqslant 1$
- $|\mathrm{Corr}\left(X,Y\right)| = 1 \Leftrightarrow \exists(a,b)\in\mathbb{R}^*\times\mathbb{R} \;\mid\; Y = aX+b$

V Lois usuelles

1 Loi de binômiale

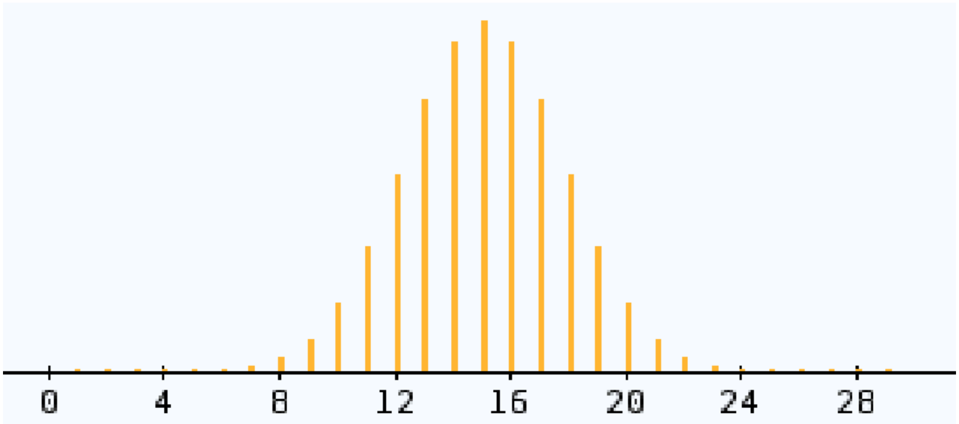
Définition

On renouvelle n fois de manière **indépendante** une épreuve de a deux issues appelées succès et échecs, et dont la probabilité d'un succès est p.
Soit X la variable aléatoire associant à chaque issue le nombre de succès obtenus après les n épreuves.
On dit alors que X suit la loi binomiale $\mathcal{B}(n,p)$.

Propriété

Soit X une variable aléatoire qui suit la loi binomiale $\mathcal{B}(n,p)$.
Pour tout entier naturel $k \in [\![0,n]\!]$ on a :

$$P(X=k) = \binom{n}{k} p^k(1-p)^{n-k}$$



Démonstration

On considère une expérience aléatoire constitué de n épreuves indépendantes à deux issues.
Soient E l'ensemble de cardinal n des n épreuves et k un entier naturel tel que $k \in [0;n]$
Chaque issue de l'expérience aléatoire peut être représentée par la combinaison de E constitué des succès. Par exemple $\{1;5;6\}$ signifie que seules la première, la cinquième et la sixième épreuves de Bernoulli ont été couronnées d'un succès avec bien sur, dans ce cas, $n \geq 6$.
Les n épreuves de Bernoulli étant indépendantes, la probabilité de n'importe quelle combinaison constitué de k succès est $p^k(1-p)^{n-k}$
Or le nombre de combinaisons à k éléments de E est $\binom{n}{k}$, donc la probabilité d'avoir k succès est :

$$P(X=k) = \binom{n}{k} p^k(1-p)^{n-k}$$

X étant une variable aléatoire suivant la loi binomiale $\mathcal{B}(n;p)$.

Propriété

Soit X une variable aléatoire suivant la loi binomiale de paramètre n et p, noté $\mathcal{B}(n,p)$. On a alors :

1. $E(X) = np$
2. $V(X) = np(1-p)$

Démonstration

2 Loi hypergéométrique

Définition

On effectue un tirage simultané de n objets, dans un urne contenant N objets, dont N_A objet A.
Soit X la variable aléatoire qui associe à chaque issue le nombre d'objets A obtenus.
On dit alors que X suit la loi hypergéométrique $\mathcal{H}\left(N, n, N_A\right)$ de paramètre N, n et N_A.

Propriété

Soit X une variable aléatoire qui suit la loi hypergéométrique $\mathcal{H}(N, n, N_A)$.
Pour tout entier naturel $k \in [\![0, n]\!]$ on a :

$$P(X = k) = \frac{\binom{N_A}{k} \times \binom{N-N_A}{n-k}}{\binom{N}{k}}$$

Propriété

Soit X une variable aléatoire suivant la loi hypergéométrique $\mathcal{H}(N, n, N_A)$. On a alors :

1. $E(X) = n\dfrac{N_A}{N} = np$ avec $p = \dfrac{N_A}{N}$
2. $V(X) = np(1-p)\dfrac{N-n}{N-1}$

Démonstration

3 Loi de Poisson

Définition

On dit que la variable aléatoire X suit la loi de Poisson de paramètre $\lambda \in]0, +\infty[$ si $X\left(\Omega\right) = \mathbb{N}$, et que pour tout $k \in \mathbb{N}$, on a :

$$P(X = k) = \mathrm{e}^{-\lambda}\frac{\lambda^k}{k!}$$

On écrit que X suit la loi $\mathcal{P}(\lambda)$

Propriétés

Soit X une variable aléatoire qui suit $\mathcal{P}(\lambda)$, alors on a :

- $E(X) = \lambda^2$
- $V(X) = \lambda$

Démonstration

4 Loi géométrique

5 Loi binômiale négative

6 Loi uniforme

7 Loi exponentielle

8 Loi normale

VI Exercices

Exercice 12

Les pièces fabriquées dans une usine proviennent de deux machines. La machine A produit 1% de pièces défectueuses et la machine B en produit 5%. La production de l'usine provient à 70% de la machine A. On choisit au hasard une pièce produite par l'usine.

1. Quelle est la probabilité qu'elle soit défectueuse ?
2. On constate que la pièce est défectueuse. quelle est la probabilité qu'elle provienne de la machine A.
3. Même question dans le cas ou l'on constate que la pièce n'est pas défectueuse.

Exercice 13

Soit A et B deux événements indépendants d'un espace probabilisé. Démontrer que $\overline{A}$ et $\overline{B}$ sont indépendants.

Exercice 14

Soit $(n, k) \in \mathbb{N}^2$ tel que $2 \leq k \leq n$. Une urne contient n boules numérotés de 1 à n.

1. On tire simultanément k boules de l'urne.
   Soit $p \in [\![k, n]\!]$, calculer la probabilité que p soit le plus grand numéro tiré ?
2. On tire successivement et avec remise k boules de l'urne. Calculer la probabilité que deux numéros identiques et seulement deux apparaissent ?

Exercice 15

Soit X une variable aléatoire dont la fonction de répartition F est définie par :

$$F(x) = \begin{cases} 0 & \text{si} \quad x \leqslant 1 \\ 1 - \dfrac{1}{1 \times 2} & \text{si} \quad 1 < x \leqslant 2 \\ 1 - \dfrac{1}{2 \times 3} & \text{si} \quad 2 < x \leqslant 3 \\ \ldots & \ldots \quad \ldots \\ 1 - \dfrac{1}{n \times (n+1)} & \text{si} \quad n < x \leqslant n+1 \\ \ldots & \ldots \quad \ldots \end{cases}$$

1. Calculer les probabilités $P(X = n)$ pour tout $n \in \mathbb{N}^*$.
2. Calculer $E(X)$ et $V(X)$.

Exercice 16

Soit X une variable aléatoire de densité $f(x) = x$ pour $x \in [0; 1]$, $f(x) = 2 - x$ pour $x \in [1; 2]$, et est nulle en dehors de ces intervalles.

1. Vérifier que f définie bien une densité de probabilité.
2. Déterminez la fonction de répartition de X et en déduire la médiane de cette loi.
3. Calculer $E(X)$ et $V(X)$.
4. Soit $x \in \mathbb{R}$, calculer $P\{|X - 1| < x\}$.

Exercice 17

Soit X une variable aléatoire de densité $f(x) = x + 2$ pour $x \in [-2; -1]$, $f(x) = -x + 2$ pour $x \in [1; 2]$, et est nulle en dehors de ces intervalles.

1. Vérifier que f définie bien une densité de probabilité.
2. Déterminez la fonction de répartition de X et en déduire la médiane de cette loi.
3. Calculer $E(X)$ et $V(X)$.

Exercice 18

On fixe un entier N Un athlète saute successivement par dessus des barres numéroté de 1 à N.
Il s'arrête au premier échec ou bien quand il a passé la barre numéro N.
Lorsqu'il tente la barre numéro i, il a une chance sur i de réussir.
On défini pour $i \in \{1; 2; ...; N\}$ l'événement A_i :"l'athlète a franchi la barre numéro i" et l'événement B_i :"la dernière barre réussit par l'athlète est la barre numéro i".
Il faut comprendre que si l'athlète ne franchi pas une barre, il ne franchi pas les suivantes, puisqu'il n'a alors même pas le droit de les tenter.

1. Calculer $P\big(A_i\big)$ pour tout $i \in \{1; 2; ...; N\}$.
2. Démontrer que pour tout $i \in \{1; 2; ...; N-1\}$, $P\big(B_i\big) = \dfrac{1}{i!} - \dfrac{1}{(i+1)!}$.

3. Que vaut $P\big(B_N\big)$

Exercice 19

Le cinéma de la commune propose différents tarifs :

- Un tarif plein à 12 euros.
- Un tarif étudiant à 7 euros.
- Un tarif enfant (moins de 12 ans) à 5 euros.

Une étude portant sur la clientèle a montré que 48% des clients paient un tarif plein, 22% des clients un tarifs enfants et les autres un tarif étudiants.
Par ailleurs, sur l'ensemble des clients, 12% achètent uniquement un paquet de bonbons a 4 euros, 23% achète uniquement un paquet de pop-corn taille standard à 5 euros et 15% achètent uniquement un paquet de pop-corn grande taille à 7 euros. Les autres clients n'achètent rien.
On choisit au hasard un client du cinéma et on appelle respectivement X_1 et X_2 les variables aléatoires qui associe à chaque issue respectivement le prix payé pour la place de cinéma et pour les confiseries.

1. Déterminer les lois de probabilité de X_1 et de X_2.
2. Déterminer $E(X_1)$ et $E(X_2)$.
3. Soit X la variable aléatoire qui associe à chaque issue le prix total dépensé pas le client.
   (a) Exprimer X en fonction de X_1 et X_2.
   (b) En déduire $E(X)$ et interpréter ce résultat dans le contexte de l'exercice.

Exercice 20

On considère une urne dans laquelle se trouve différentes boules de couleur : des rouges, des vertes et des noires.
On tire avec remise trois boules de l'urne.
a chaque étape, obtenir une boule noire rapporte 5 points, obtenir une boule verte rapporte 2 points et obtenir une boule rouge fait perdre 10 points.
Pour tout entier naturel i compris entre 1 et 3, on note respectivement R_i, V_i et N_i l'événement "obtenir une boule rouge (respectivement une boule verte ou une boule noire) au $i^{\text{ième}}$ tirage.
On note enfin X_1, X_2 et X_3 les variables qui associe à chaque issue le nombre de points obtenu au premier (respectivement deuxième et troisième) tirage.

1. Calculer les valeurs suivantes :
   $X_1\left(V_1 \cap R_2 \cap N_3\right)$, $X_2\left(V_1 \cap R_2 \cap N_3\right)$, $X_1\left(V_1 \cap R_2 \cap R_3\right)$, $X_2\left(V_1 \cap R_2 \cap R_3\right)$
   et $X_3\left(V_1 \cap R_2 \cap R_3\right)$.
2. Déterminer les lois de probabilité de X_1, X_2 et X_3.
3. Déterminer $E(X_1)$, $E(X_2)$ et $E(X_3)$.
4. Soit X la variable aléatoire qui associe à chaque issue le nombre total de points.
   (a) Exprimer X en fonction de X_1, X_2 et X_3
   (b) En déduire $E(X)$

Exercice 21

Deux lycées sont situés dans la même commune.
Le premier lycée, noté lycée A, réalise de très bons résultats aux examen : 75% des élèves de l'établissement obtiennent le bac avec mention.
Dans le lycée B, seulement 55% des élèves l'obtiennent avec mention.
On choisit 12 élèves du lycée A et 20 du lycée B.
Le nombre important d'élèves dans chaque lycée permet d'assimiler ces expériences à des tirage avec remises.
On note respectivement X et Y les variables aléatoires associant à chaque issue le nombre d'élèves du lycée A (respectivement du lycée B) ayant eu le bac avec mention.

1. Donner les lois suivies par X et Y.
2. Soit Z la variable aléatoire qui associe à chaque issue le nombre d'élèves ayant eu le bac avec mention indépendamment de leur lycée d'origine.
   (a) Exprimer Z en fonction de X et de Y.
   (b) En déduire $E(Z)$ et interpréter le résultat dans le contexte de l'exercice.
   (c) Calculer $V(Z)$ et en déduire $\sigma(Z)$

Exercice 22

On considère une urne contenant trois boules indiscernables au toucher et numéroté de 1 à 3.
On tire successivement et sans remise 3 boules de l'urne.
Pour $k \in \{1;2;3\}$, on dit qu'il y a "rencontre" au $k^{\text{ième}}$ tirage lorsque l'on tire la boule numéroté k au $k^{\text{ième}}$ tirage.
On cherche a connaître le nombre moyen de rencontres.
Pour tout $(m,k) \in \{1;2;3\}^2$, on note $B_{m,k}$ l'événement "On a obtenu la boule m au $k^{\text{ième}}$ tirage".

1. On note X_1, X_2 et X_3 les variables aléatoires définies par :
   - $X_1 = 1$ si la boule numéroté 1 est tirée en premier et 0 sinon.
   - $X_2 = 1$ si la boule numéroté 2 est tirée en deuxième et 0 sinon.
   - $X_3 = 1$ si la boule numéroté 3 est tirée en troisième et 0 sinon.

   Donner alors $E(X_1)$, $E(X_2)$ et $E(X_3)$.
2. Soit X la variable aléatoire qui associe à chaque issue le nombre de rencontre.
   (a) Exprimer X en fonction de X_1, X_2 et X_3.
   (b) En déduire le nombre moyen de rencontre.
3. On considère une urne contenant n boules numéroté de 1 à n. On tire successivement et sans remise n boules de l'urne.
   Quel est théoriquement le nombre moyen de rencontre ?

Exercice 23

Soit n un entier naturel non nul.
Une urne contient une unique boule, qui est blanche.
On dispose d'une pièce dont la probabilité de donner pile est $p \in \left]0;1\right[$. Les différents lancers de la pièces sont indépendants. On note $q = 1 - p$.
On lance n fois de suite la pièce. on ajoute des boules noire dans l'urne a chaque fois que l'on obtient pile : deux pour le premier pile, 3 pour le deuxième pile, etc. Ainsi on ajoute $k+1$ boules

noire dans l'urne pour la $k^{\text{ième}}$ obtention de pile.
On note X la variable aléatoire associant à chaque issue le nombre de piles obtenus, et N la variable aléatoire associant à chaque issue le nombre de total de boules dans l'urne.

1. (a) Exprimer N en fonction de X.
   (b) Quelle est la loi de X.
   (c) En déduire l'espérance de N.
2. On tire une boule de l'urne et on pose B :"La boule tirée est blanche".
   (a) Démontrer que
   $$P(B)=\sum_{k=0}^{n}\frac{2}{(k+1)(k+2)}\binom{n}{k}p^kq^{n-k}$$
   (b) Calculer cette somme.
3. On change la règle : cette fois, on ajoute dans l'urne 2^{k-1} boules noires lors de l'obtention du $k^{\text{ième}}$ pile.

Exercice 24

Soient $n \in \mathbb{N}^*$ et $r \in \mathbb{N}$ tel que $1 \leqslant r \leqslant n$.
Un placard contient n paires de chaussures.
On tire au hasard $2r$ chaussures du placard et on note X la variable aléatoire qui associe à chaque issue le nombre de paires complètes parmi les chaussures tirées.
Les paires du placard sont numérotés de 1 à n et pour tout $i \in [\![1;n]\!]$, on note X_i la variable aléatoire qui associe 1 à chaque issue si les deux chaussures de la paire i se trouvent dans les chaussures tirées, et 0 sinon.

1. Démontrer que $\forall i \in [\![1;n]\!], E(X_i)=\dfrac{r(2r-1)}{n(2n-1)}$.
2. En déduire $E(X)$

Exercice 25

Soit (X,Y) un couple de variables aléatoires dont la loi de probabilité est donnée ci-dessous :

$X\backslash Y$	1	2	3	4
1	0,08	0,04	0,16	0,12
2	0,04	0,02	0,08	0,06
3	0,08	0,04	0,16	0,12

1. Déterminer les lois de probabilités des variables aléatoires X et Y
2. X et Y sont-elles indépendantes ?
3. Déterminer la loi du couple $(\max\{X,Y\},\min\{X,Y\})$

Exercice 26

On lance successivement trois pièces de monnaie.
Soient X la variables aléatoire qui à chaque issue associe le nombre de faces apparues sur les deux premières pièces et Y celle qui associe à chaque issue le nombre de piles apparues sur les deux dernières.

1. Déterminer la loi du couple (X, Y)
2. Déterminer la loi de X, puis celle de Y.
3. X et Y sont-elles indépendantes ?

Exercice 27

Soient a un réel et X une variable aléatoire de densité f définie par :

$$f(x) = \begin{cases} \dfrac{2}{a}\left(1 - \dfrac{x}{a}\right) & \text{si } 0 \leqslant x \leqslant a \\ 0 & \text{sinon} \end{cases}$$

1. Vérifier que f définie bien une densité de probabilité.
2. Déterminez la fonction de répartition de X et en déduire la médiane de cette loi.
3. Calculer $E(X)$ et $V(X)$.
4. Soient $X_1, X_2, \ldots, X_n$ des variables indépendantes de même loi que X.
   Déterminer la fonction de répartition puis la densité de la variable aléatoire m_n définie pour tout $n \in \mathbb{N}^*$ par :
   $$M_n = \max\{X_1, \ldots, X_n\}$$

Exercice 28

Soient X et Y deux variables aléatoires indépendantes de loi de Poisson de paramètres respectifs λ et μ.

1. Démontrer que la variable aléatoire $S = X + Y$ suit une loi de poisson de paramètre $\lambda + \mu$.
2. Démontrer que la loi conditionnelle de X lorsque la somme $S = X + Y$ a une valeur fixé s est une loi binomiale dont on donnera les paramètres.

Exercice 29

1. Soient $X_1, \ldots, X_n$, n variables aléatoires indépendantes, de même loi, de moyenne m et de variance σ^2.
   En justifiant soigneusement, calculer l'espérance et la variance de la variable aléatoire $\overline{X}_n = \dfrac{1}{n}\sum_{i=1}^{n} X_i$ en fonction de m et de σ.
2. Soient $X_1, \ldots, X_n$, n variables aléatoires qui ne sont **pas indépendantes**.
   (a) Exprimer $VAR\left(\sum_{i=1}^{n} X_i\right)$ en fonction de $VAR(X_i)$ et $\text{Cov}(X_i, X_j)$ pour $(i, j) \in$

$[\![1, 2, ..., n]\!]$.

(b) Que devient ce résultat quand les couples sont de même lois.

Exercice 30

Soient $X_1, \ldots, X_n$, n variables aléatoires indépendantes, de même loi de densité f_X définie par :

$$f : \begin{array}{l} \mathbb{R} \to \mathbb{R} \\ x \mapsto Kx^4 \mathrm{e}^{-\theta x} 1_{[0;+\infty[}(x) \end{array}$$

où K est une constante positive et θ un réel strictement positif.

1. Déterminer K en fonction de θ.
   (On admettra l'égalité $\displaystyle\int_0^{+\infty} u^n e^{-u} du = n!$
2. Représenter f_X.
3. Calculer $E(X_1)$ et $\mathrm{Var}(X_1)$.
4. Soient $n \in \mathbb{N}$ et $\overline{X}_n = \dfrac{1}{n} \displaystyle\sum_{i=1}^{n} X_i$
   Déterminer $E\left(\overline{X}_n\right)$ et $\mathrm{Var}\left(\overline{X}_n\right)$

Chapitre 4

Convergence

I Inégalités de Markov et de Bienaymé-Tchebychev

Théorème - Inégalité de Markov
Soit X une variable aléatoire réelle positive ou nulle d'espérance μ.

$$\text{pout tout } \delta > 0, P\big(X \geqslant \delta\big) \leqslant \frac{\mu}{\delta}$$

Démonstration
Soit X une variable aléatoire réelle positive ou nulle d'espérance μ.
Soient $n \in \mathbb{N}^*$ et $(x_i)_{i\in[\![1,n]\!]}$.
On a donc $\mu = \sum_{i=1}^{n} x_i P\left(X = x_i\right)$, d'où pour tout $\delta > 0$:

$$\mu = \sum_{x_i \geqslant \delta} x_i P\left(X = x_i\right) + \sum_{x_i < \delta} x_i P\left(X = x_i\right)$$

donc, X étant une variable aléatoire à valeur positive, on a :

$$\begin{aligned}
\mu &\geqslant \sum_{x_i \geqslant \delta} x_i P\left(X = x_i\right) \\
&\geqslant \sum_{x_i \geqslant \delta} \delta P\left(X = x_i\right) \qquad \text{car } x_i \geqslant \delta \\
&\geqslant \delta \times \sum_{x_i \geqslant \delta} P\left(X = x_i\right) \quad \text{en factorisant par } \delta \\
&\geqslant \delta \times P\left(X \geqslant \delta\right)
\end{aligned}$$

donc

$$P\left(X \geqslant \delta\right) \leqslant \frac{\mu}{\delta}$$

Théorème - Inégalité de Bienaymé-Tchebychev
Soit X une variable aléatoire d'espérance μ et de variance V.

$$\text{pout tout } \delta > 0, P\big(|X - \mu| \geqslant \delta\big) \leqslant \frac{V}{\delta^2}$$

Interprétation La probabilité que les valeurs prisent par X s'écarte de plus de δ de l'espérance $\mu = E(X)$ est d'autant plus petite que δ est grand.

Exemple

On lance une pièce équilibré 100 fois de suite. Soit X la variable aléatoire qui associé a chaque issue de l'expérience aléatoire constituée de cent lancers le nombre de piles obtenus.
X suit la loi binomiale de paramètre $n = 100$ et $p = \frac{1}{2}$.
On a donc $E(X) = 100 \times \frac{1}{2} = 50$ et $V(x) = 100 \times \frac{1}{2}\frac{1}{2} = 25$

D'après l'inégalité de Bienaymé-Tchebychev, pour tout réel $\delta > 0$, on a :

$$P\big(|X - 50| \geqslant \delta\big) \leqslant \frac{25}{\delta^2}$$

Par exemple, pour $\delta = 10$

$$P\big(|X - 50| \geqslant 10\big) \leqslant \frac{25}{100}$$

La probabilité que l'écart de X à 50 soit supérieur à 10 inférieur à 0,25.
D'autre part $P\big(|X - 50| < 10\big) = 1 - P\big(|X - 50| \geqslant 10\big)$, donc $P\big(|X - 50| < 10\big) \geqslant 0,75$ La probabilité que l'écart de X à 50 soit inférieur à 10 est supérieur à 0,75.

Démonstration

Soient X une variable aléatoire d'espérance μ et de variance V et δ un réel strictement positif.
$\delta > 0$, donc $|X - \mu| \geqslant \delta \Leftrightarrow (X - \mu)^2 \geqslant \delta^2$
La variable aléatoire $(X - \mu)^2$ étant positive ou nulle, on a, d'après l'inégalité de Markov,

$$P\big((X - \mu)^2 \geqslant \delta^2\big) \leqslant \frac{E\big((X - \mu)^2\big)}{\delta^2}$$

or $E\big((X - \mu)^2\big) = V$, donc $P\big((X - \mu)^2 \geqslant \delta^2\big) \leqslant \frac{V}{\delta^2}$ et par conséquent :

$$P\big(|X - \mu| \geqslant \delta\big) \leqslant \frac{V}{\delta^2}$$

II Loi des grands nombres

Théorème - Inégalité de concentration

Soit n variables aléatoires (avec $n \geqslant 1$) $X_1, X_2, ..., X_n$ indépendantes, de même espérance μ et de même variance V.
On note $M_n = \frac{X_1 + X_2 + ... + X_n}{n}$ la variable aléatoire moyenne.

$$\text{pout tout } \delta > 0, P\big(|M_n - \mu| \geqslant \delta\big) \leqslant \frac{V}{n\delta^2}$$

En effet, $E(M_n) = \mu$ et $V(M_n) = \dfrac{V}{n}$

Exemple

Dans une société de démarchage par téléphone, on estime que 40% des personnes appelées répondent effectivement.
On appelle n personnes. Soit X_k la variable aléatoire qui associe 1 à toute issue dont la k^{ime} personne appelée répond et 0 sinon.
les variables aléatoires $X_1, X_2, ..., X_n$ sont indépendantes et identiquement distribuées d'espérance $\mu = 0,6 \times 0 + 0,4 \times 1 = 0,4$ et de variance $V = 0,6 \times 0,4 = 0,24$
On a alors, pour tout réel $\delta > 0$, $P\big(|M_n - 0,4| \geqslant \delta\big) \leqslant \dfrac{0,24}{n\delta^2}$
Par exemple pour $\delta = 0,1$ et $n = 1000$
$P\big(|M_n - 0,4| \geqslant \delta\big) \leqslant \dfrac{0,24}{1000 \times 0,1^2}$
Donc $P\big(|M_n - 0,4| \geqslant \delta\big) \leqslant 0,024$. on dit que l'on obtient pour M_n une précision de $0,1$ avec un risque de 0,024.
Lorsque l'on appelle 1000 personnes, la probabilité que le nombre de personne qui répondent soit en dehors de l'intervalle $]300; 500[$ est inférieur à $0,024 = 2,4\%$

Démonstration

Soit n variables aléatoires (avec $n \geqslant 2$) $X_1, X_2, ..., X_n$ mutuellement indépendantes, de même espérance μ et de même variance V.
On a donc, $E(M_n) = \mu$ et $V(M_n) = \dfrac{V}{n}$ et par conséquent, l'inégalité de Bienaymé-Tchebychev donne pour tout $\delta > 0$:

$$P\big(|M_n - \mu| \geqslant \delta\big) \leqslant \frac{V}{n\delta^2}$$

Propriété - Loi faible des grands nombres

Soient $n \in \mathbb{N}$ et $(X_i)_{1\leqslant i\leqslant n}$, n variables aléatoires indépendantes d'espérance μ et de variance V.
Soit $M_n = \dfrac{1}{n}\displaystyle\sum_{i=1}^{n} X_i$ la variable aléatoire moyenne.
Pour tout réel $\delta > 0$

$$\lim_{n\to+\infty} P(|M_n - \mu| \geqslant \delta\big) = 0$$

Démonstration

Soient $n \in \mathbb{N}$ et $(X_i)_{1\leqslant i\leqslant n}$ un échantillon de taille n d'une variable aléatoire X d'espérance μ et de variance V.
Soit $M_n = \dfrac{1}{n}\displaystyle\sum_{i=1}^{n} X_i$ la variable aléatoire moyenne.
Pour tout réel $\delta > 0$ D'après l'inégalité de Bienaymé-Tchebychef,

$$0 \leqslant P\big(|M_n - \mu| \geqslant \delta\big) \leqslant \frac{V}{n\delta^2}$$

d'où, par encadrement et par passage à la limite quand n tend vers $+\infty$,

$$\lim_{n\to+\infty} P\big(|M_n-\mu|\geqslant\delta\big)=0$$

III Convergence en probabilité

Remarque

Si (X_n) est une suite de variables aléatoires qui converge vers une variable aléatoire X, cela signifie que X_n se "rapproche" de X quand n augmente.
On mesure la distance entre X_n et X par $|X_n-X|$ qui sera d'autant plus petite que n sera grand ; mais s'agissant de variables aléatoires, il faut considérer l'événement $|X_n-X|<\epsilon$ qui sera réalisé avec une probabilité d'autant plus élevé que n sera grand.
On va donc associer à la suite de variables aléatoires la suite numérique des probabilités de ces événements, qui devra converger vers 1.
On obtient donc la définition suivante :

Définition

On dit que la suite de variables aléatoires X_n converge en probabilités vers la variable aléatoire X, si :

$$\forall\epsilon>0,\ \lim_{n\to+\infty} P\left(|X_n-X|<\epsilon\right)=1$$

ou, de manière équivalente :

$$\forall\epsilon>0,\ \lim_{n\to+\infty} P\left(|X_n-X|>\epsilon\right)=0$$

On écrit :

$$X_n \underset{p}{\to} X$$

Exemple

Soit $(U_n)_{n\in\mathbb{N}^*}$ une suite de variables aléatoires indépendantes, de même loi uniforme sur $[0;1]$.
Soit $(X_n)_{n\in\mathbb{N}^*}$ la suite de variables aléatoires définie par :pour tout $n\in\mathbb{N}^*$ par :

$$\forall n\in\mathbb{N}^*, X_n=\min\{U_1,...,U_n\}$$

Démontrons alors que :

$$X_n \underset{p}{\to} 0$$

On doit montrer que pour tout $\epsilon>0$, on à $\lim\limits_{n\to+\infty} P\left(|X_n|>\epsilon\right)=0$ Tout d'abord, $\forall\epsilon\geqslant 1$, $\forall n\in\mathbb{N}^*$, on a $P\left(|X_n|>\epsilon\right)=0$. Il suffit donc de considérer le cas $\epsilon<1$.
Soit $\epsilon<1$
Les variables aléatoires (U_n) étant indépendantes et de même loi, on a :
$P\left(|X_n|>\epsilon\right)=P\left(\bigcap\limits_{i=1}^{n}(U_i>\epsilon)\right)$

$$= \prod_{i=1}^{n} P\left(U_i > \epsilon\right)$$
$$= \prod_{i=1}^{n} \left(1 - P\left(U_i > \epsilon\right)\right)$$
$$= (1-\epsilon)^n$$

or

$$\lim_{n\to+\infty} (1-\epsilon)^n = 0$$

Donc $\boxed{X_n \underset{p}{\to} 0}$

Remarque

Dans l'exemple précédent, on a pu établir la convergence en probabilité à partir de la définition. Cependant, pour des lois plus complexes, cette convergence peut être difficile à obtenir en utilisant cette définition.

Propriété

Soit $(X_n)_{n\in\mathbb{N}^*}$ un suite de variables aléatoires telle que :

$$\begin{cases} \lim\limits_{n\to+\infty} E(X_n) = a \\ \lim\limits_{n\to+\infty} V(X_n) = 0 \end{cases}$$

alors

$$X_n \underset{p}{\to} a$$

Démonstration

Soit $(X_n)_{n\in\mathbb{N}^*}$ un suite de variables aléatoires telle que :

$$\begin{cases} \lim\limits_{n\to+\infty} E(X_n) = a \\ \lim\limits_{n\to+\infty} V(X_n) = 0 \end{cases}$$

alors, d'après l'inégalité de Markov appliquée à la variable aléatoire $(X_n - a)^2$, pour tout $\delta > 0$, on a :

$$P\left(|X_n - a|^2 \geqslant \delta^2\right) \leqslant \frac{E\left((X_n-a)^2\right)}{\delta^2}$$

$$P\left(|X_n - a| \geqslant \delta\right) \leqslant \frac{V\left(X_n - a\right) + \left(E(X_n - a)\right)^2}{\delta^2}$$

donc

$$P\left(|X_n - a| \geqslant \delta\right) \leqslant \frac{V\left(X_n\right) + \left(E(X_n) - a\right)^2}{\delta^2}$$

d'où, par passage à la limite quand n tend vers $+\infty$, on a :

$$\lim_{x\to+\infty} P\left(|X_n - a| \geqslant \delta\right) = 0$$

et par conséquent

$$X_n \underset{p}{\to} a$$

Propriété

Soit $(X_n)_{n\in\mathbb{N}^*}$ un suite de variables aléatoires telle que :

$$\begin{cases} \lim\limits_{n\to+\infty} E(X_n - X) = 0 \\ \lim\limits_{n\to+\infty} V(X_n - X) = 0 \end{cases}$$

alors

$$X_n \underset{p}{\to} X$$

Démonstration

En effet, en appliquant la propriété précédente à la variable aléatoire $X_n - X$, on obtient $X_n - X \underset{p}{\to} 0$, c'est à dire que pour tout $\delta > 0$:

$$\lim_{n\to+\infty} P\big(|X_n - X| \geqslant \delta\big) = 0$$

c'est à dire :

$$X_n \underset{p}{\to} X$$

Théorème de Slutsky

Soit f une application réelle continue.

$$X_n \underset{p}{\to} X \Rightarrow f(X_n) \underset{p}{\to} f(X)$$

Remarque

Ce théorème exprime qu'une application continue conserve la convergence en probabilité, au sens où la limite de la suite des images est l'image de la limite de la suite.

Exemple

théorème de Bernoulli

On effectue n expériences successives indépendantes, où on s'intéresse à chaque fois à la réalisation d'un certain événement A. on associe donc à chaque expérience i, $1 \leqslant i \leqslant n$, un variable de Bernoulli :

$$\begin{cases} P(X_i = 1) = p \\ P(X_i = 0) = q = 1 - p \end{cases}$$

La fréquence empirique, c'est à dire le pourcentage de réalisation de l'événement A est :

$$f_n = \frac{1}{n}\sum_{i=1}^{n} X_i = \overline{X}_n$$

Avec $E(f_n) = p$ et $V(f_n) = \dfrac{pq}{n}$ et en appliquant l'inégalité de Bienaymé-Tchebychev, on obtient, pour tout $\delta > 0$:

$$P\big(|f_n - p| > \delta\big) \leqslant \frac{pq}{n\delta^2}$$

D'où, :

$$\lim_{n\to+\infty} P\big(|f_n - p| > \delta\big) = 0$$

Donc $\boxed{f_n \underset{p}{\rightarrow} p}$

C'est à dire que la fréquence empirique (observée) d'un événement converge, en probabilité, vers la fréquence théorique ou probabilité de réalisation de cet événement quand le nombre d'expériences augmente indéfiniment.
Ce théorème, connu sous le nom de théorème de Bernoulli, établi le lien entre l'expérience que l'on peut avoir des chances de réalisation d'un événement et la définition théorique que l'on a donnée de sa probabilité.

Définition

Soit $p \in \mathbb{N}^*$.
On dit que la suite de variables aléatoires X_n converge en moyenne d'ordre p vers la variable aléatoire X, si :

$$\lim_{n\to+\infty} E\left(|X_n - X|^p\right) = 0$$

On écrit :

$$X_n \underset{\mathcal{M}p}{\rightarrow} X$$

IV Convergence en Loi

1 Généralités

Définition

On dit que la suite de variables aléatoires X_n, de fonction de répartition F_n, converge en loi vers la variable aléatoire X de fonction de répartition F si la suite $\big(F_n(x)\big)$ converge vers $F(x)$ en tout point x où F est continue.
Donc, pour tout réel x où F est continue, on a :

$$\lim_{n\to+\infty} F_n(x) = F(x)$$

On écrit :

$$X_n \underset{loi}{\rightarrow} X$$

Propriété

La convergence en probabilité en probabilité d'une suite de variables aléatoires X_n implique sa convergence en loi.

$$X_n \underset{p}{\to} X \Rightarrow X_n \underset{loi}{\to} X$$

Remarque

Ce résultat est intuitif, car $F_n(x)$ représente la probabilité, associé à la variable aléatoire X_n, de l'événement $\{X_n < x\}$, et elle va converger vers la probabilité correspondante associé à la loi de X. La réciproque de ce résultat est fausse en général, mais vraie dans le cas particulier où la limite est est une variable aléatoire certaine $X = a$.

Démonstration

Propriétés

Soient (X_n) et (Y_n) deux suites de variables aléatoires telles que :

$$X_n \underset{loi}{\to} X \quad \text{et} \quad Y_n \underset{p}{\to} a$$

- $X_n + Y_n \underset{loi}{\to} X + a$
- $X_n Y_n \underset{loi}{\to} aX$
- $\dfrac{X_n}{Y_n} \underset{loi}{\to} \dfrac{X_n}{a}$ si $a \neq 0$

Démonstration

Théorème de Slutsky

Soit f une application réelle continue.

$$X_n \underset{loi}{\to} X \Rightarrow f(X_n) \underset{loi}{\to} f(X)$$

Démonstration

2 Théorème central limite

Remarque

La loi des grands nombres énonce la convergence de la suite $\left(\overline{X}_n\right)$ des moyennes empiriques vers la moyenne théorique m, c'est à dire une variable aléatoire certaine, ou loi de Dirac, loi dégénérée dont toute la masse est concentrée en un point. Si l'on équilibre la suite des variables aléatoires centrées réduites vers $\left(\dfrac{\overline{X}_n - m}{\sigma}\right)$, qui converge vers 0, par $\sqrt{n}$ qui converge vers $+\infty$, la forme indéterminée obtenue aura une limite qui sera celle d'une loi non dégénérée cette fois, la loi normale.
C'est ce qu'énonce le **théorème central limite** ci dessous.

Théorème central limite

Soit (X_n) une suite de variable aléatoire indépendantes et de même loi d'espérance $\mu = E(X_n)$ et de variance $V(X_n) = \sigma^2$ non nulle. On a alors :

$$\sqrt{n}\frac{\overline{X}_n - \mu}{\sigma} \underset{loi}{\rightarrow} \mathcal{N}(0,1)$$

Remarque

Le théorème central limite donne aussi :

$$\frac{\overline{X}_n - \mu}{\frac{\sigma}{\sqrt{n}}} \underset{loi}{\rightarrow} \mathcal{N}(0,1)$$

On peux approximer la loi suivie par la variable aléatoire $\overline{X}_n$ par la loi normale $\mathcal{N}\left(\mu, \dfrac{\sigma^2}{n}\right)$ dont l'espérance est μ et l'écart type $\dfrac{\sigma}{\sqrt{n}}$

Exemple

D'après le théorème central limite, on a :

$$\sqrt{n}\frac{\overline{X}_n - \mu}{\sigma} \underset{loi}{\rightarrow} \mathcal{N}(0,1)$$

C'est à dire, si Φ est la fonction de répartition associée à la loi normale centrée réduite, on a pour tout réel t :

$$P\left(\sqrt{n}\frac{\overline{X}_n - \mu}{\sigma} < t\right) \underset{n \to +\infty}{\rightarrow} \Phi(t)$$

On appelle q_α le quantile d'ordre $\alpha \in [0;1]$, c'est à dire la valeur pour laquelle on ait $\Phi(q_\alpha) = \alpha$.

On a alors :

$$P\left(\sqrt{n}\frac{\overline{X}_n-\mu}{\sigma}\in[q_\alpha,q_{1-\alpha}]\right)\underset{n\to+\infty}{\to}\Phi(1-q_\alpha)-\Phi(1-q_{1-\alpha})=1-2\alpha$$

On a donc, pour tout réel μ :

$$P\left(\mu\in\left[\overline{X}_n-\frac{\sigma q_{1-\alpha}}{\sqrt{n}},\overline{X}_n-\frac{\sigma q_\alpha}{\sqrt{n}}\right]\right)\underset{n\to+\infty}{\to}1-2\alpha$$

Avec α fixé, on obtient donc un intervalle de confiance sur la valeur de μ recherchée.
On cherche souvent un intervalle de confiance à $0,95=95\%$, on choisit donc $\alpha=0,0025$

V Convergence presque sûre

Définition

On dit que la suite (X_n) converge presque sûrement vers la variable aléatoire X si :

$$P\left\{\omega\in\Omega|\lim_{n\to+\infty}X_n(\omega)=X(\omega)\right\}=1$$

On écrit :

$$X_n\underset{p.s}{\to}X$$

Exercice 31

Etudier la convergence en probabilité, la convergence en moyenne puis la convergence en moyenne quadratique de la suite de variables aléatoires (X_n) dont la loi de probabilité est définie pour tout $n\in\mathbb{N}^*$ par :

$$P(X_n=-n)=\frac{1}{2n^2}\quad,\quad P(X_n=n)=\frac{1}{2n^2}\quad\text{et}\quad P(X_n=0)=1-\frac{1}{n^2}$$

Exercice 32

Soit (X_n) la suite de variables aléatoires dont la loi de probabilité est définie pour tout $n\in\mathbb{N}^*$ par :

$$P(X_n=n)=\frac{1}{n}\quad\text{et}\quad P(X_n=0)=1-\frac{1}{n}$$

1. Montrer que (X_n) converge en probabilité, mais pas en moyenne quadratique, vers zéro quand n tend vers $+\infty$.
2. Soit (Y_n) une suite de variables aléatoires de même loi $\mathcal{N}(0;1)$ et indépendantes des variables aléatoires (X_n).
   Etudiez la convergence en loi de la variable aléatoire $Z_n=X_n+Y_n$.
3. Déterminer $\lim\limits_{n\to+\infty}V(Z_n)$.

Exercice 33

Etudiez la convergence en loi de la suite de variables aléatoires (X_n) définies par :

$$\forall n \in \mathbb{N}^*, \quad P(X_n = n) = 1 \quad \text{et } P(X_n \neq n) = 0$$

Exercice 34

Soient $X_1, \ldots, X_n$, n variables aléatoires indépendantes, de même loi de densité f_X définie par :

$$f_X : \begin{array}{l} \mathbb{R} \to \mathbb{R} \\ x \mapsto Kx^4 \mathrm{e}^{-\theta x} 1_{[0;+\infty[}(x) \end{array}$$

où K est une constante positive et θ un réel strictement positif.

1. Déterminer K en fonction de θ.
   (On admettra l'égalité $\displaystyle\int_0^{+\infty} u^n e^{-u} du = n!$
2. Représenter f_X.
3. Calculer $E(X_1)$ et $\mathrm{Var}(X_1)$.
4. Soient $n \in \mathbb{N}$ et $\overline{X}_n = \dfrac{1}{n} \displaystyle\sum_{i=1}^{n} X_i$
   Déterminer $E\left(\overline{X}_n\right)$ et $\mathrm{Var}\left(\overline{X}_n\right)$
5. Calculer $\displaystyle\lim_{n \to +\infty} \overline{X_n}$ en justifiant votre calcul.
6. On pose, pour tout $n \in \mathbb{N}^*$, $Y_n = X_n^3$.
   Calculer $\displaystyle\lim_{n \to +\infty} Y_n$, en justifiant le calcul.

Exercice 35

Soit (X_n) une suite de variables aléatoires de loi géométrique de paramètre $p_n = \dfrac{\theta}{n}$, où θ est un nombre strictement positif.
Montrer que la suite $\left(\dfrac{X_n}{n}\right)$ converge vers la loi exponentielle de paramètre θ.

Exercice 36

Soient $p \in]0;1[$ et (U_n) une suite de variables aléatoires indépendantes et de même loi définie par :

$$P(U_n = 1) = p \quad \text{et} \quad P(U_n = -1) = q = 1 - p$$

Soient $n \in \mathbb{N}$ et V_n la variable aléatoire définie par :

$$V_n = \prod_{i=1}^{n} U_i$$

1. Déterminer la loi exacte de V_n.

2. Déterminer la loi limite de la suite (V_n)

Exercice 37

Soient $\theta > 0$ et (X_n) une suite de variables aléatoires de loi géométrique de paramètre $p_n = \dfrac{\theta}{n}$.
Montrer que la suites $\left(\dfrac{X_n}{n}\right)$ converge en loi vers la loi exponentielle de paramètre θ.