Lois conjointes

- Loi conjointe des va discrètes
- Propriétés de l'espérance, covariance, coefficient de corrélation
- Loi faible des grands nombres
- Théorème central limite
- Loi forte des grands nombres

Exemple

- On tire au hasard 3 boules d'une urne contenant 3 boules rouges , 4 blanches et 5 bleues. On note X et Y le nombre de boules rouges et blanches tirées.
- Déterminer la probabilité du couple de v.a. (X,Y)

Calcul des P(X=i,Y=j)

P(X=i,Y=j) est nulle dès que i+j>3.

Pour i+j ≤ 3, on a P(X=i,Y=j)= $\frac{C_3^i C_4^j C_5^{3-i-j}}{C_{12}^3}$

On peut donc construire le tableau suivant

Probabilité conjointe de X et Y

i \ j	0	1	2	3	P(X=i)
0	10/220	40/220	30/220	4/220	84/220
1	30/220	60/220	18/220	0	108/220
2	15/220	12/220	0	0	27/220
3	1/220	0	0	0	1/220
P(Y=j)	56/220	112/220	48/220	4/220	1

Loi discrète conjointe

Lorsque X et Y sont des v.a. discrètes, on peut définir une loi de probabilité simultanée ou loi conjointe

$P_{(X,Y)}(x,y)=P(X=x,Y=y)$

Loi de probabilité marginale

On peut déduire les loi de probabilité de X et Y de leur loi conjointe :

$$P_X(x) = \sum_{y,\, p(x,y)>0} p(x,y)$$

$$P_Y(y) = \sum_{x,\, p(x,y)>0} p(x,y)$$

Exercice 5.1

Dans un certain village

- 15% des familles n'ont pas d'enfants
- 20% des familles ont un enfant
- 35% des familles ont deux enfants
- 30% des familles ont trois enfants
- Chaque enfant a une chance sur deux d'être une fille indépendamment du sexe de ses éventuels frères et sœurs

- Soit F et G les deux variables aléatoires représentant le nombre de filles et de garçons dans une famille de ce village.
- Déterminer la loi conjointe de F et G.

Variables indépendantes

Définition :

Deux variables aléatoires **discrètes** sont indépendantes si et seulement si leur loi de probabilité conjointe est égale au produit de leur lois de probabilité.

$$P(X=i\ ,\ Y=j)=P(X=i)P(Y=j)$$

Exercice 5.2

On réalise n+m épreuves indépendantes, chacune ayant probabilité de succès n.

X= nombre de succès pendant les m premières épreuves

Y= nombre de succès pendant les n dernières épreuves

Z= nombre total de succès

X et Y sont elles indépendantes ?

X et Z sont elles indépendantes ?

Cas continu

On ne peut utiliser la même définition d'indépendance pour des variables continues.

Le cas continu ne sera pas traité dans le cadre de ce cours.

Espérance d'un produit de variables indépendantes

Exercice 5.3 : Montrez que si X et Y deux variables aléatoires discrètes indépendantes alors on a

E[XY]=E[X]E[Y]

Covariance

- La covariance de deux variables aléatoires quelconques X et Y est notée Cov(X,Y) et est définie par
- Cov(X,Y)=E[(X-E[X])(Y-E[Y])]

  =E[XY]-E[X]E[Y]

Remarque : La covariance de deux variables indépendantes est nulle.

Mais la réciproque est fausse.

Exercice 5.4 Montrez que

Cov(X,Y)=Cov(Y,X)

Cov(X,X)=Var(X)

Cov(aX,Y)=aCov(X,Y)

Cov(X+Y,Z)=Cov(X,Z)+Cov(Y,Z)

Variance de la somme de v.a. indépendantes

- Exercice 5.5 : Montrez que

  Si les Xi sont deux à deux indépendantes, alors alors la variance de leur somme est égale à la somme de leur variance.

Application :variance d'une loi binomiale

Soit X une variable aléatoire binomiale de paramètre n et p. On peut la considérer comme la somme de n variables indépendantes de Bernouilli.

Comme la variance d 'une variable de Bernouilli est égale à p(1-p), on retrouve que la variance de la loi binomiale est

np(1-p).

Corrélation

- La corrélation entre deux variables aléatoires X et Y de variance non nulle est notée ρ(X,Y) et est définie par

$$\rho(X,Y)=\frac{Cov(X,Y)}{\sqrt{Var(X)Var(Y)}}$$

Théorème : -1≤ρ(X,Y) ≤1

- Démonstration

$$0\leq Var(\frac{X}{\sqrt{Var(X)}}+\frac{Y}{\sqrt{Var(Y)}})$$

$$=\frac{Var(X)}{Var(X)}+\frac{Var(Y)}{Var(Y)}+2\frac{Cov(X,Y)}{\sqrt{Var(X)Var(Y)}}$$

$$=2(1+\rho(X,Y))$$

Donc -1≤ρ(X,Y)

- D'autre part

$$0 \le Var\left(\frac{X}{\sqrt{Var(X)}} - \frac{Y}{\sqrt{Var(Y)}}\right)$$
$$= \frac{Var(X)}{Var(X)} + \frac{Var(Y)}{Var(Y)} - 2\frac{Cov(X,Y)}{\sqrt{Var(X)Var(Y)}}$$
$$= 2(1-\rho(X,Y))$$

Donc $\rho(X,Y) \le 1$

Théorèmes limites

Théorèmes limites

- **Loi faible des grands nombres**
  - Inégalités de Markov et de Tchebychev
  - Loi faible des grands nombres
- **Théorème central limite**
  - Convergence en probabilité
  - Théorème central limite
  - Applications
- **Loi forte des grands nombres**
  - Énoncé
  - Applications

Probabilité qu'une v.a. positive soit plus grande qu'une valeur donnée : Inégalité de Markov

Théorème

Soit X une variable aléatoire à valeurs positives ou nulles. Pour tout réel a positif

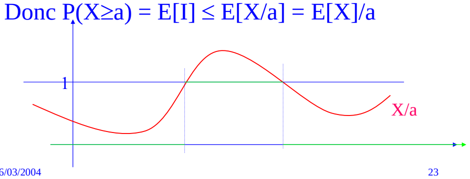
$$P(X \ge a) \le E[X]/a$$

Ceci quelque soit la loi de X

Remarque : aucun intérêt si $a<E[X]$

Preuve

- Soit I la variable aléatoire qui vaut 1 quand X est supérieure où égale à a et 0 sinon.
- I est inférieure ou égale à X/a, puisque X est à valeurs positives ou nulle.
- Donc $P(X \ge a) = E[I] \le E[X/a] = E[X]/a$



Probabilité d'écart entre une v.a. et sa moyenne : inégalité de Tchebychev

Conséquence de l'inégalité de Markov

Théorème : inégalité de Tchébychev

Soit X une variable aléatoire d'espérance m et de variance σ^2 finie. Pour tout réel k>0

$$P(|X-m| \ge k) \le \sigma^2/k^2$$

Ceci quelque soit la loi de X

Remarque : aucun intérêt si $k < \sigma$

Exercice 5.6

- Prouver l'inégalité de Tchebychev

Exercice 5.7

Supposons que je produise en moyenne 50 diapositives par jour, quelle est, au plus, la probabilité pour que demain je produise plus de 75 diapositives ?

Si je sais de plus que la variance de ma production journalière est de 25, quelle est, au moins, la probabilité pour que demain je produise entre 40 et 60 diapositives ?

Exercice 5.8

- Soit X une variable aléatoire de moyenne m et d'écart-type σ. Le rapport $r=|m|/\sigma$ est appelé rapport signal-bruit. On définit $D=|(X-m)/m|$, l'erreur relative de X par rapport à son signal m.

  Montrez que pour tout $\alpha > 0$

  $$P(D < \alpha) \geq 1- 1/(r\alpha)^2$$

(im)précision de cette borne

Exercice 5.9

Supposons que X suive une loi normale d'espérance m et de variance σ^2, calculez la borne donnée par l'inégalité de Tchebychev sur $P(|X-m| \geq 2\sigma)$ ainsi que la valeur exacte de cette probabilité, comparez.

File ou Pace

Exercice 5.10:

On effectue *n* parties de pile ou face. Trouver *n* pour que l'on puisse affirmer que la fréquence d'apparition de pile durant ces *n* lancers soit comprise entre 0.45 et 0.55 avec une probabilité supérieure à 0.90.

Loi faible des grands nombres

Soient X_1, X_2, … une suite de variables aléatoires indépendantes de même loi de probabilité. On suppose qu'elles admettent toutes la même espérance finie m, alors pour tout $\varepsilon > 0$,
$P(|(X_1+\ldots+X_n)/n - m| \geq \varepsilon)$ tend vers 0 quand n tend vers l'infini

Preuve

Preuve dans le cas où les X_i admettent de plus une variance finie

- Notons $Y_n = (X_1+\ldots+ X_n)/n$
- $E[Y_n] = m$, et $var(Y_n) = \sigma^2/n$
- Donc d'après l'inégalité de Tchebychev,
- $P(|Y_n - m| > \varepsilon) \leq \sigma^2/n\varepsilon^2$.

Interprétation de la loi faible des grands nombres

Si on répète n fois une expérience et que l'on calcule la moyenne des résultats d'une variable aléatoire X, alors quand n tend vers l'infini, la probabilité que cette moyenne soit *loin* de E[X] tend vers 0.

On parle alors de convergence en probabilité.

Définition : convergence en probabilité

La suite X_n converge en probabilité vers la variable aléatoire X, et on note $X_n \overset{P}{\to} X$

si

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

Théorème central limite

- Théorème central limite
- Applications

Théorème central limite

Ce théorème établit que la somme d'un *grand nombre* de variables aléatoires indépendantes de même loi de probabilité suit approximativement une loi normale.

Énoncé du théorème central limite

Soit $X_1, X_2, \ldots$ une suite de variables aléatoires indépendantes de même loi de probabilité, d'espérance m et de variance σ^2.

Alors la variable aléatoire $(X_1+X_2+\ldots+X_n)$ tend vers la loi normale de moyenne nm et de variance $n\sigma^2$ quand n tend vers l'infini.

Autrement dit

$((X_1+X_2+\ldots+X_n) - nm)/\sigma\, n^{1/2}$ tend vers la loi normale standard quand n tend vers l'infini.

Utilisation du théorème central limite

Utilisation :

on ramène les probabilités cherchées à des probas

concernant $Z_n = \dfrac{X_1 + ... + X_n - nm}{\sigma\sqrt{n}}$

et on approxime la probabilité demandée sur Z_n par

la probabilité demandée sur la loi normale standard N.

Exemple

- On lance 10 dés équilibrés. On cherche à donner une approximation de la probabilité que la somme des 10 résultats soit comprise entre 30 et 40.

- Notons X_i le résultat du i-ème dé, on recherche $P(30 \leq \Sigma X_i \leq 40)$.
  $E[X_i] = 7/2$ et
  $Var(X_i) = E[(X_i)^2] - (E[X_i])^2 = 91/6 - 49/4$
  $Var(X_i) = 35/12$
  On applique le théorème central limite :

  On ramène les probabilités cherchées à des probas

  concernant $Z_{10} = \dfrac{X_1 + ... + X_{10} - 10m}{\sqrt{10 Var(X)}}$

  et on remplace Z_{10} la loi normale standard.

- $(S - 10x7/2)/ (350/12)^{1/2}$ se comporte comme la loi normale standard.
- Il faut ramener $P(30 \leq S \leq 40)$ à une probabilité concernant $(S - 35)/ (350/12)^{1/2}$

$$P(30 \leq S \leq 40) = P\left(\frac{30-35}{\sqrt{\frac{350}{12}}} \leq \frac{S-35}{\sqrt{\frac{350}{12}}} \leq \frac{40-35}{\sqrt{\frac{350}{12}}}\right)$$

$$\approx P\left(\frac{30-35}{\sqrt{\frac{350}{12}}} \leq N \leq \frac{40-35}{\sqrt{\frac{350}{12}}}\right) = P\left(-\sqrt{\frac{6}{7}} \leq N \leq \sqrt{\frac{6}{7}}\right)$$

$$= 2\Phi\left(\sqrt{\frac{6}{7}}\right) - 1 \approx 0.65$$

Exercice 5.11

Un astronome désire mesurer la distance, en années-lumière entre la terre et une certaine étoile. Chaque fois qu'il réalise une mesure, il obtient une distance approchée du fait de petites erreurs dues entre autres, aux influences atmosphériques. Il prévoit donc de faire plusieurs mesures et de prendre la moyenne de ces mesures comme estimation de la distance réelle. Il suppose que les valeurs mesurées sont des variables aléatoires indépendantes suivant toutes la même loi, d'espérance commune d (la vraie valeur) et de variance commune 4. Combien de mesures doit-il effectuer pour être raisonnablement sûr (95%) que l'erreur soit inférieure à une demi-année-lumière ?

Exercice 5.12

On fait une suite de n tirages Pile ou Face avec une pièce non biaisée (p=1/2). En utilisant le théorème de la limite centrale, calculer une approximation de la probabilité

- que le nombre de Pile soit dans l'intervalle [45..55] si on fait 100 tirages
- que le nombre de Pile soit dans l'intervalle [450..550] si on fait 1000 tirages

Combien faut-il faire de tirages pour que le nombre de Pile se trouve dans l'intervalle [0.45n ,0.55n] avec une probabilité de 95% ?

Comparer avec la borne obtenue par Tchebychev.

Version plus générale du théorème central limite

Il existe des versions plus générales (mais plus compliquées) du théorème central limite, où l'on n'exige pas que toutes les variables aléatoires suivent la même loi.

Convergence presque sûre

- On dit que la suite de variables aléatoires réelles X_n converge presque sûrement vers la variable aléatoire X, si il existe N un sous-ensemble de R tel que

$$\begin{cases} \forall \omega \notin N, X_n(\omega) \to X(\omega) \\ P(N) = 0 \end{cases}$$

Loi forte des grands nombres

Soient X_1, X_2 , … une suite de variables indépendantes et identiquement distribuées d'espérance finie m, alors $Z_n=(X_1+X_2+\ldots+X_n)/n$ converge presque sûrement vers m.

(non démontré)

Comparaison entre loi faible et loi forte

- La loi faible assure qu'en moyenne Z_n est proche de m
- La loi forte assure que Z_n est presque partout proche de m.

Application de la loi forte des grands nombres

- Supposons que l'on réalise une série d'expériences indépendantes. Soit E un événement donné relatif à l'expérience et P(E) sa probabilité.
- On pose :
  - $X_i=1$ si E est réalisé lors de la ième expérience
  - $X_i=0$ sinon
- On a : $E[X_i] = P(E)$
- La loi forte des grands nombres établit que $(X_1+X_2+\ldots+X_n)/n$ tend vers E[X]=P(E).
- Autrement dit la fréquence relative de l'événement E tend vers la probabilité de l'événement E, P(E).

Processus de Poisson

On considère un processus dans lequel des événements se produisent à des instants aléatoires, par exemple :
- émissions de particules radioactives
- arrivées de voitures dans une station service
- requêtes dans un standard téléphonique

Sous certaines hypothèses on peut modéliser ces processus par des processus de Poisson

Si on a un processus dans lequel des événements se produisent à des instants aléatoires, avec les hypothèses suivantes, le processus après le temps t :

- est indépendant du processus avant le temps t (perte de mémoire)
- se comporte exactement comme le processus originel (propriété de régénération)

Alors

- les v.a. *temps entre 2 arrivées d'événements* suivent des lois exponentielles
- les v.a. *nombre d'arrivées avant t* suivent des lois de Poisson

Les variables aléatoires

- T_k : l'heure de la k-ième arrivée. Ce sont des v.a. réelles
- Pour k dans IN, T_k est associée à l'expérience aléatoire consistant à laisser le processus évoluer jusqu'à la k-ième arrivée.

Les variables aléatoires

- N_t : le nombre d'arrivées dans l'intervalle de temps [0,t], ce sont des v.a. discrètes
- Pour t réel, N_t est associée à l'expérience consistant à laisser le processus se dérouler jusqu'au temps t.
  C'est l'expérience de Poisson.

Les hypothèses de base

- Si on se fixe un temps t, alors le processus après le temps t
  - est indépendant du processus avant le temps t (perte de mémoire)
  - se comporte exactement comme le processus originel (propriété de régénération)

Temps inter-arrivées

- Considérons des nouvelles variables aléatoires, temps inter-arrivées, définies par $X_1=T_1$,
  et pour tout k >1, $X_k = T_k - T_{k-1.}$
- Nos hypothèses de bases se traduisent alors par
  - Les X_i sont des variables aléatoires indépendantes (processus sans mémoire)
  - $P(X_i>t+s|X_i>s)=P(X_i>t)$ (régénération)

Conséquences

- Soit $G(t) = P(X>t)$
- Montrez que G est une exponentielle :

Preuve

On a :

$P(X>t) = P(X>t+s|X>s) =$

$P(X>t+s \ \& \ X>s)/P(X>s)$

$= P(X>t+s)/P(X>s)$

Donc $G(t+s) = P(X>t+s) = G(t)G(s)$

$G(t+s) = G(t)G(s)$ d'où $G'(t+s) = G'(t)G(s)$

et pour $t = 0$: $G'(s) = G'(0)G(s)$
d'où G est une exponentielle.

$G(t)=c^t$.

Comme $G(\infty)=0$, $c<1$ et on peut pose :

$G(t) = e^{-rt}$

Fonction de répartition de X

$P(X > t) = P(X \le t) = F_X(t)$

On a donc $F_X(t) = 1-e^{-rt}$

$f(t) = F'(t)=re^{-rt}$

Donc les lois inter-arrivées suivent des lois exponentielles

Les T_k suivent une Gamma distribution

On a $T_k = X_1+X_2+\ldots+X_k$, et les X_k sont des v.a. exponentielles indépendantes.

Montrons par induction sur k que T_k est une Gamma distribution, c'est à dire que sa densité de probabilité est $(rt)^{k-1} re^{-rt}/(k-1)!$ pour tout $t>0$

- Pour k=1 c'est vrai
- Supposons le résultat vrai pour k-1
- La densité de probabilité de $T_k = T_{k-1} + X_k$ est égale à :

(en utilisant : $f_{X+Y}(a) = \int_{-\infty}^{+\infty} f_X(a-y)\, f_Y(y)dy$)

$$f_{T_k}(x) = \int_{-\infty}^{+\infty} f_{T_{k-1}}(y) f_{X_k}(x-y)dy$$

$$= \int_0^x \frac{(ry)^{k-2}}{(k-2)!} re^{-ry} re^{-r(x-y)} dy = \frac{r^k e^{-rx}}{(k-2)!} \int_0^x y^{k-2} dy$$

$$= \frac{r(rx)^{k-1}}{(k-1)!} e^{-rx}$$

Les N_t suivent une distribution de Poisson

On a :
plus de k arrivées avant t ssi
heure de la kième arrivée $\le$ t

Soit par définition des v.a. N_t et T_k, on a $N_t \ge k$ si et seulement si $T_k \le t$.

On a donc $P(N_t \ge k) = P(T_k \le t)$

Les N_t suivent une distribution de Poisson

Or $P(N_t \ge k) = P(T_k \le t)$

$$\int_0^t f_{T_k}(x)dx = \int_0^t (rx)^{k-1} \frac{re^{-rx}}{(k-1)!} dx = 1 - \sum_{j=0}^{k-1} e^{-rt} \frac{(rt)^j}{j!}$$

On peut vérifier la dernière égalité par induction via une intégration par parties.
Par ailleurs : $P(N_t = k) + P(N_t \ge k+1) = P(N_t \ge k)$

On a donc P(N_t= k) = $\dfrac{e^{-rt}(rt)^k}{k!}$

La densité de probabilité de N_t est donc une loi de Poisson de paramètre rt.

En résumé

Si on a un processus dans lequel des événements se produisent à des instants aléatoires, avec les hypothèses suivantes, le processus après le temps t :

- est indépendant du processus avant le temps t (perte de mémoire)
- se comporte exactement comme le processus originel (propriété de régénération)

Alors

- les v.a. *temps entre 2 arrivées d'événements* suivent des lois exponentielles
- les v.a. *nombre d'arrivées avant t* suivent des lois de Poisson