Chapitre 5
Couple de variables aléatoires

Définitions

1. On appelle **couple de variables aléatoires** (discrètes) l'application:

$$\begin{array}{ccc} \Omega & \rightarrow & \mathbb{R} \\ \omega & \mapsto & (X(\omega), Y(\omega)) \end{array}$$

2. La distribution d'un couple de v.a. est définie par
   - les ensembles $X(\Omega)$ et $Y(\Omega)$
   - $\forall x_i \in X(\Omega)$ et $y_j \in Y(\Omega)$, la probabilité p_{ij} de l'évenement $[(X = x_i) \cap (Y = y_j)]$
3. De la loi conjointe, on tire les lois marginales

$$\begin{cases} \mathbb{P}(X = x_i) = \sum_{y_j \in Y(\Omega)} \mathbb{P}\left[(X = x_i) \cap (Y = y_j)\right] = \sum_{y_j \in Y(\Omega)} p_{ij} = p_{i\bullet} \\ \mathbb{P}(Y = y_j) = \sum_{x_i \in X(\Omega)} \mathbb{P}\left[(X = x_i) \cap (Y = y_j)\right] = \sum_{x_i \in X(\Omega)} p_{ij} = p_{\bullet j} \end{cases}$$

Exemple

Si X et Y prennent un nb fini de valeurs, on représente graphiquement la loi conjointe dans un tableau.

Exemple : On tire 3 cartes d'un jeu de 32. Soit X le nombre de cœurs et Y le nombre de brelan (3 cartes de même "type").

$X(\Omega) = \{0, 1, 2, 3\}$; $Y(\Omega) = \{0, 1\}$; $C_{32}^3 = 4960$ tirages possibles

Cardinaux :

$Y \backslash X$	0	1	2	3
0	$C_{24}^3 - 8$	$C_8^1 C_{24}^2 - 24$	$C_8^2 C_{24}^1$	C_8^3
1	8	24	0	0

Distribution de probabilités :

$Y \backslash X$	0	1	2	3	
0	$p_{00} = \frac{252}{620}$	$p_{10} = \frac{273}{620}$	$p_{20} = \frac{84}{620}$	$p_{30} = \frac{7}{620}$	$p_{\bullet 0} = \frac{616}{620}$
1	$p_{01} = \frac{1}{620}$	$p_{11} = \frac{3}{620}$	$p_{21} = 0$	$p_{31} = 0$	$p_{\bullet 1} = \frac{4}{620}$
	$p_{0\bullet} = \frac{253}{620}$	$p_{1\bullet} = \frac{276}{620}$	$p_{2\bullet} = \frac{84}{620}$	$p_{3\bullet} = \frac{7}{620}$	$p_{\bullet\bullet} = 1$

Indépendance

Définition

Deux v.a. X et Y sont dites indépendantes si: $\forall x_i \in X(\Omega)$ et $\forall y_j \in Y(\Omega)$, on a $\mathbb{P}\left[(X = x_i) \cap (Y = y_j)\right] = \mathbb{P}(X = x_i)\mathbb{P}(Y = y_j)$

Dans l'exemple, on remarque
$\mathbb{P}\left[(X = 3) \cap (Y = 1)\right] = 0 \neq \mathbb{P}(X = 3)\mathbb{P}(Y = 1)$,
donc X et Y ne sont pas indépendants

Généralisation

n v.a. $X_1, ..., X_n$ sont dites indépendantes si:
$\forall x_1 \in X_1(\Omega)$,..., $\forall x_n \in X_n(\Omega)$, on a
$\mathbb{P}\left[(X_1 = x_1) \cap ... \cap (X_n = x_n)\right] = \mathbb{P}(X_1 = x_1)...\mathbb{P}(X_n = x_n)$

Typiquement, si on répête une même expérience plusieurs fois dans les mêmes conditions, et que X_i se réfère à la $i^{\text{ème}}$ expérience, on a indépendance des X_i

Covariance

Obejctif : Étudier le "lien" entre deux v.a.

Définition

Soient X et Y deux v.a. On appelle covariance X et Y le réél $\sigma_{XY} = \text{cov}(XY) = \mathbb{E}\left[(X - \mathbb{E}(X))(Y - \mathbb{E}(Y))\right]$

Propriétés

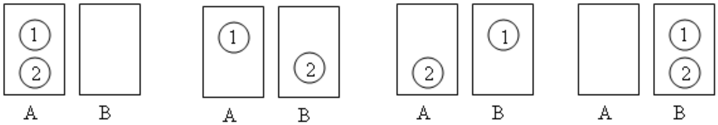
1. $\text{cov}(X, Y) = \mathbb{E}(XY) - \mathbb{E}(X)\mathbb{E}(Y)$
2. $\text{cov}(X, Y) = \text{cov}(Y, X)$ (symétrie)
3. $\text{cov}(X, X) = \sigma_X^2 = \text{Var}(X)$
4. $\text{cov}(aX + bY, Z) = a\text{cov}(X, Z) + b\text{cov}(Y, Z)$
5. $\text{Var}(X + Y) = \text{Var}(X) + 2\text{cov}(X, Y) + \text{Var}(Y)$
6. Si X et Y sont indépendants alors $\text{cov}(X, Y) = 0$
   ($\text{cov}(X, Y) = 0$ n'implique pas nécessairement X et Y indpts)
7. $|\text{cov}(X, Y)| \leq \sigma_x \sigma_y$ (égalité ssi $Y = aX + b$)

Exemple

On a deux boîtes A et B et 2 billes ① et ②. Chaque bille est placée au hasard dans une des deux boîtes.

Soit X=nb de billes dans la boite A ; Y=nb de boîtes vides.
On a $X(\Omega) = \{0, 1, 2\}$ et $Y(\Omega) = \{0, 1\}$

4 cas équiprobables



$Y \backslash X$	0	1	2	
0	0	1/2	0	1/2
1	1/4	0	1/4	1/2
	1/4	1/2	1/4	1

$\mathbb{E}(X) = 0x1/4 + 1x1/2 + 2x1/4 = 1$;
$\mathbb{E}(Y) = 0x1/2 + 1x1/2 = 1/2$
$\mathbb{E}(XY) =?$, $XY(\Omega) = \{0, 1, 2\}$
$\mathbb{P}(XY = 0) = 3/4$, $\mathbb{P}(XY = 1) = 0$, $\mathbb{P}(XY = 2) = 1/4$
$\Rightarrow \mathbb{E}(XY) = 2/4 = 1/2$
$\Rightarrow \text{cov}(X, Y) = 1x1/2 - 1/2 = 0$

On a $\text{cov}(X, Y) = 0$, or X et Y ne sont pas indépendants, par ex $\mathbb{P}(X = 1 \cap Y = 1) = 0 \neq \mathbb{P}(X = 1)\mathbb{P}(Y = 1) = 1/2x1/2 = 1/4$

Corrélation

Définition

Soient X et Y des v.a. de moyennes finies et de variances non nulles. Le **coefficient de correlation** de X et Y est alors

$$\rho_{X,Y} = \frac{\mathrm{cov}(X,Y)}{\sigma_X \sigma_Y}$$

Propriété : D'après la propriété 7 de la covariance :

$$-1 \leq \rho_{X,Y} \leq 1$$

Par ailleurs $\mid \rho_{X,Y} \mid = 1$ ssi $Y = aX + b$ avec $a > 0$ si $\rho = 1$ et $a < 1$ si $\rho = -1$
$\Rightarrow$ si $\rho_{X,Y} = 1$, X et Y évoluent dans la même direction alors que si $\rho_{X,Y} = -1$ Y diminue quand X augmente et vice-versa

Vocabulaire

$\rho_{X,Y} = 0 \Rightarrow$ X et Y sont "non corrélés"

indépendants $\Rightarrow$ non corrélés (la réciproque n'est pas vraie)

$\rho_{X,Y} > 0 \Rightarrow$ X et Y sont positivement corrélés

$\rho_{X,Y} = 1 \Rightarrow$ X et Y sont parfaitement positivement corrélés

$\rho_{X,Y} < 0 \Rightarrow$ X et Y sont négativement corrélés

$\rho_{X,Y} = -1 \Rightarrow$ X et Y sont parfaitement négativement corrélés

Meilleur prédicteur linéaire

- Le coefficient de corrélation est souvent présenté comme une mesure de la relation linéaire entre deux variables
- On a vu précédemment qu'une correlation parfaite correspond à une relation linéaire (parfaite)
- Qu'en est-il pour les correlations "imparfaites"?

Supposons que l'on veuille approximer une variable aléatoire Y par une fonction linéaire de X (autre v.a.) : $\alpha + \beta X$.

Comment choisir au mieux α et β? (on parlera alors de **meilleur prédicteur linéaire**)

L'erreur de l'approximation s'écrit

$$\epsilon = Y - \alpha - \beta X$$

et est appelée la **variable aléatoire résiduelle**.

L'idée est de choisir α et β de façon à minimiser $\mathbb{E}(\epsilon^2)$, l'**erreur quadratique moyenne**. On note MSE (Mean Squared Error) $= \mathbb{E}(\epsilon^2) = \mathbb{E}\left[(Y - \alpha - \beta X)^2\right]$

On a déjà vu que si $\rho_{X,Y} = \pm 1$, l'apprimation peut-être faite de manière exacte, i.e. avec $\mathbb{E}(\epsilon^2) = 0$.

Dans le cas général, l'erreur quadratique moyenne peut s'écrire:

$$\text{MSE} = \mathbb{E}(Y^2) - 2\alpha\mathbb{E}(Y) - 2\beta\mathbb{E}(XY) + \beta^2\mathbb{E}(X^2) + 2\alpha\beta\mathbb{E}(X) + \alpha^2$$

La valeur minimum de cette expression (si elle existe) est obtenue en égalisant les dérivées partielles (par rapport à α et à β) à 0. On obtient alors

$$\begin{cases} \beta\mathbb{E}(X^2) + \alpha\mathbb{E}(X) = \mathbb{E}(XY) \\ \beta\mathbb{E}(X) + \alpha = \mathbb{E}(Y) \end{cases}$$

C'est-à-dire

$$\begin{cases} \beta = \dfrac{\sigma_{XY}}{\sigma_x^2} \\ \alpha = \mathbb{E}(Y) - \beta\mathbb{E}(X) \end{cases}$$

Remarque : alors $\mathbb{E}(\epsilon) = 0$

Pour résumer, on a

Definition

Soient X et Y deux v.a. Soit l'équation $Y = \alpha + \beta X + \epsilon$ où α et β sont des constantes et ϵ une v.a. définie par $\epsilon = Y - \alpha - \beta X$.

Alors Y est approximée par la relation linéaire $\alpha + \beta X$ avec ϵ comme variable aléatoire résiduelle. Le **meilleur prédicteur linéaire** de Y par X est l'appliquation linéaire $\alpha + \beta X$ qui minimise l'**erreur quadratique moyenne** $\mathbb{E}(\epsilon^2)$.

Le coefficient β est appelé le **beta** de Y par rapport à X et la droite $y = \alpha + \beta X$ est appelée **droite de régression**.

Théroème

Le meilleur prédicteur linéaire de Y par X est donné par :

$$\frac{\sigma_{XY}}{\sigma_X^2}X + \mathbb{E}(Y) - \frac{\sigma_{XY}}{\sigma_X^2}\mathbb{E}(X)$$

Alors, l'erreur quadratique moyenne minimum est égale à

$$\mathbb{E}(\epsilon^2) = \sigma_Y^2(1 - \rho_{X,Y}^2)$$

On peut donc définir les propriétés suivantes du coefficient de correlation

- $\rho_{X,Y} = \pm 1$ ssi il existe une relation linéaire en X et Y
- Plus $\rho_{X,Y}$ est proche de ± 1, plus l'erreur quadratique moyenne (MSE) est faible quand on utilise le meilleur prédicteur linéaire

- Si $\rho_{X,Y}$ est positif, le meilleur prédicteur linéaire a une pente positive. Ainsi, quand X augmente (resp. diminue), le meilleur prédicteur linéaire de Y augmente (resp. diminue) aussi
- Si $\rho_{X,Y}$ est négatif, le meilleur prédicteur linéaire a une pente négative. Ainsi, quand X augmente, le meilleur prédicteur linéaire de Y diminue et vice-versa.

Attention : Une forte correlation n'implique pas de relation causale. Le fait qu'une variable Y prenne des valeurs qui peuvent être approximées linéairement par les valeurs d'une autre v.a. X, ne signifie pas qu'une évolution de X cause une évolution de Y (ex : augmentation des ventes d'ordinateur et d'automobile dans les années 90)

Application - Modèle de marché

Soit une action S. En utilisant le meilleur prédicteur linéaire, on veut approximer le rendement hebdomadaire de S (R_S) par une fonction linéaire du rendement hebdomadaire du marché (R_M).

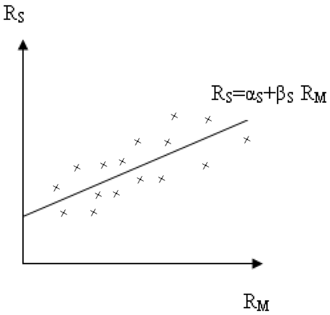
Le meilleur prédicteur linéaire donne

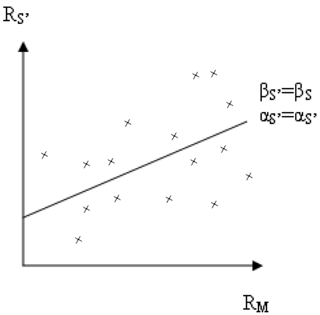
$$R_S = \alpha_S + \beta_S R_M + \epsilon$$

avec

$$\beta_S = \frac{cov(R_S, R_M)}{\sigma_M^2} \text{ et } \alpha_S = \mathbb{E}(R_S) - \beta \mathbb{E}(R_M)$$

β_S est le beta de l'action S et correspond à la pente de la régression linéaire de R_S sur R_M





Deux titres différents peuvent avoir le même beta

$\beta_S = 1$: les variations du titre suivent l'évolution du marché
$\beta_S > 1$: le titre amplifie les fluctuations du marché (titre volatil)
$\beta_S < 1$: le titre sous-réagit par rapport au marché (une variation de 1% du marché entraine une variation de moins de 1% du titre)

Calcul du risque du titre

$$\begin{aligned} Var(R_S) &= Var(\alpha_S + \beta_S R_M + \epsilon) = Var(\beta_S R_M + \epsilon) \\ &= Var(\beta_S R_M) + Var(\epsilon) + 2\beta_S cov(R_M, \epsilon) \end{aligned}$$

Or,

$$\begin{aligned} cov(R_M, \epsilon) &= cov(R_M, R_S - \alpha_S - \beta_S R_M) \\ &= cov(R_M, R_S) - \beta_S cov(R_M, R_M) = 0 \end{aligned}$$

par définition de β_S.
Ainsi,

$$Var(R_S) = \beta_S^2 Var(R_M) + Var(\epsilon) = \beta_S^2 \sigma_M^2 + \sigma_\epsilon^2$$

Le risque d'un titre peut donc être décomposé en :

- un **risque systématique** : $\beta_S^2 \sigma_M^2$ dépendant de l'état global (macroscopique) du marché (taux d'intérêt, offre de monnaie, guerre,...) et proportionnel à β_S (effet amplificateur)
- un **risque spécifique** : σ_ϵ^2 qui affecte uniquement le titre considéré et est indépendant des phénomènes macroscopiques. Ex : mauvaise gestion de l'entreprise, incendie qui détruit son usine ou invention technologique qui rend obsolète la gamme de produits

D'après la théorie économique, un portefeuille diversifié (cf. TD) permet de supprimer le risque spécifique. Seul le risque systématique devrait donc être pris en compte pour évaluer le rendement d'un titre. Le **beta** d'un titre est alors la caractéristique principale de l'analyse de son rendement.