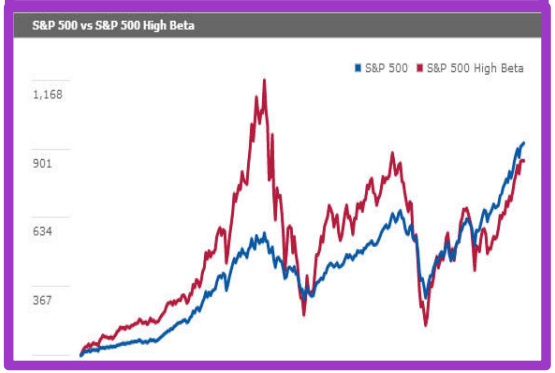
Cours de Finance

Bêtas (définitions, décomposition du risque), droite caractéristique (estimation des bêtas)

S&P500 vs high beta stocks



$$\boldsymbol{\beta_i} = \frac{\mathrm{Cov}(\boldsymbol{r_i}, \boldsymbol{r_P})}{\mathbf{Var}[\boldsymbol{r_P}]}$$

$$\boldsymbol{\beta_i} = \rho_{iP} \frac{\sigma_i}{\sigma_P}$$

Bêta et risque de marché

- Objectifs pédagogiques de la séance
  - *Connaître les définitions et les principaux résultats*
  - *Savoir les appliquer pour les exercices à venir*
  - *Développer les intuitions économiques*
  - *Démonstrations et éléments mathématiques à voir ensuite, sous forme de travail personnel.*

Bêta et risque

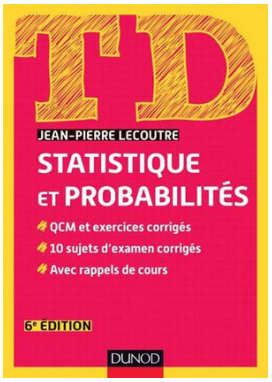
- Notion de Bêta
  - *Intuition*
  - *Décomposition du risque total d'un titre en risque de marché et risque idiosyncratique*
  - *Relation avec le coefficient de corrélation linéaire*
  - *Bêtas d'un actif sans risque, du portefeuille de référence, d'un portefeuille de deux titres*
- Droite « caractéristique » d'un titre
  - *Calcul des bêtas à partir d'historiques de cours*
  - *Bêtas calculés de manière glissante (rolling regressions)*
- Dynamique des bêtas
  - *Retour à la moyenne ?*

Rappels : loi conditionnelle, espérance conditionnelle

Rappel : cours de probabilité niveau L (Jean-Pierre Lecoutre)

Cet ouvrage présente de façon claire et pédagogique les principaux outils de la statistique et des probabilités. Chaque chapitre s'organise en quatre temps forts :une introduction présentant la problématique abordée, assortie d'objectifs de connaissances et des notions à maîtriser ;un cours proposant de nombreux théorèmes, applications et définitions ;une page L'essentiel", mentionnant les points clés à retenir dans chaque chapitre ;des exercices de difficulté progressive et leurs corrigés détaillés.Avec, en fin d'ouvrage, les principales tables statistiques et un index des notions clés."

Ce manuel de cours est destiné principalement aux étudiants de la Licence économie et gestion mais peut être utile à toute personne souhaitant connaître et surtout utiliser les principales méthodes de la statistique inférentielle. Il correspond au programme de probabilités et statistique généralement enseigné dans les deux premières années de Licence (L1 et L2). Cette 7e édition s'est enrichie d'exercices nouveaux. Le niveau mathématique requis est celui de la première année de Licence, avec quelques notions (séries, intégrales multiples...) souvent enseignées seulement en deuxième année.



- https://univ-scholarvox-com.ezpaarse.univ-paris1.fr/book/88941530#
- Disponible également sur Kindle
- Manuel de TD avec corrigés disponible
- https://univ-scholarvox-com.ezpaarse.univ-paris1.fr/catalog/book/docid/88826941?searchterm=Lecoutre,%20Jean-Pierre

Rappel : cours de probabilité niveau L (Jean-Pierre Lecoutre)

- À voir : les treize pages ci-dessous, qui traitent du cas bivarié
  - *En priorité (si besoin), les trois pages associées aux sections 1.1 à 1.4 qui traitent du cas discret (le plus simple)*
  - *Cela suppose que les notions antérieures sur le cas univarié aient été assimilées*

Rappel : cours de probabilité niveau L (Jean-Pierre Lecoutre)

- Loi d'un couple de v.a. discrètes, loi marginale

1.2 Lois marginales

À la loi d'un couple sont associées deux lois marginales qui sont les lois de chacun des éléments du couple pris séparément, définies par l'ensemble des valeurs possibles et les probabilités associées obtenues par sommation, soit :

$$P_X(X = x_i) = \sum_{j \in J} P_{(X,Y)}(X = x_i, Y = y_j) = \sum_{j \in J} p_{ij} = p_{i.}$$

$$P_Y(Y = y_j) = \sum_{i \in I} P_{(X,Y)}(X = x_i, Y = y_j) = \sum_{i \in I} p_{ij} = p_{.j}$$

Si la loi du couple est présentée dans un tableau, ces lois sont obtenues dans les marges, par sommation de ligne ou de colonne.

Y \ X	x_i	
y_j	$p_{ij} \rightarrow$	$p_{.j}$ ↓
	$p_{i.} \rightarrow$	1

Rappel : cours de probabilité niveau L (Lecoutre)

- Loi conditionnelle de X sachant Y : cas discret

1.3 Lois conditionnelles

On peut également associer deux lois conditionnelles à la loi d'un couple, c'est-à-dire la loi d'une variable, l'autre ayant une valeur fixée (loi dans une ligne ou dans une colonne donnée). Par exemple, pour $Y = y_j$ fixé, la loi conditionnelle de X est définie par l'ensemble des valeurs possibles et les probabilités associées :

$$P\left(X = x_i | Y = y_j\right) = \frac{P\left(X = x_i, Y = y_j\right)}{P\left(Y = y_j\right)} = \frac{p_{ij}}{p_{.j}} = p_i^j$$

on vérifie que c'est bien une loi de probabilité sur $\Omega_X = \{x_i; i \in I\}$:

$$\sum_{i \in I} p_i^j = \frac{1}{p_{.j}} \sum_{i \in I} p_{ij} = 1$$

- Dans cette présentation, il y a en fait autant de lois de probabilité conditionnelles que de valeurs y_j
- On peut de même définir les lois de Y sachant X

Rappel : cours de probabilité niveau L (Lecoutre)

- Espérance conditionnelle (cas discret)

1.4 Moments conditionnels

Aux lois conditionnelles sont associés des moments conditionnels, comme par exemple l'espérance conditionnelle de Y pour $X = x_i$ fixé, qui est l'espérance de la loi définie par les couples $\left\{\left(y_j, p_j^i\right); j \in J\right\}$, soit :

$$E\left(Y | X = x_i\right) = \sum_{j \in J} y_j P\left(Y = y_j | X = x_i\right) = \sum_{j \in J} y_j p_j^i$$

- L'espérance conditionnelle (sachant X) d'une v.a. Y est l'espérance de Y pour « la » loi conditionnelle de Y sachant X
- Dans le cas qui nous intéresse (X, Y) sont deux rentabilités
- $E[Y|X = x_i] = a + \beta_{Y|X} x_i$: Relation affine

Rappel : cours de probabilité niveau L (Lecoutre)

- Si $E[Y|X = x_i] = a + \beta_{Y|X} x_i$, on peut définir $E[Y|X]$ comme $a + \beta_{Y|X} X$
- Ci-dessous, on montre que $E\left[E[Y|X]\right] = E[Y]$
  - *« L'espérance de l'espérance conditionnelle est l'espérance »*
- Dans notre cas (linéaire), $E[Y] = E\left[a + \beta_{Y|X} X\right] = a + \beta_{Y|X} E[X]$

Notons que $E(Y|X)$ est une fonction de X, donc une variable aléatoire discrète dont la loi de probabilité est définie par l'ensemble des valeurs possibles, soit ici $\left\{E\left(Y|X = x_i\right); i \in I\right\}$, et les probabilités associées $p_{i.} = P(X = x_i)$. On peut donc calculer la valeur moyenne de cette v.a., soit :

$$\begin{aligned} E\left[E\left(Y|X\right)\right] &= \sum_{i \in I} p_{i.} E\left(Y|X = x_i\right) \\ &= \sum_{i \in I} p_{i.} \sum_{j \in J} y_j P\left(Y = y_j | X = x_i\right) \\ &= \sum_{i \in I} p_{i.} \sum_{j \in J} y_j p_j^i = \sum_{i \in I} \sum_{j \in J} y_j p_{i.} \frac{p_{ij}}{p_{i.}} = \sum_{j \in J} y_j \sum_{i \in I} p_{ij} \\ &= \sum_{j \in J} p_{.j} y_j \\ &= E\left(Y\right) \end{aligned}$$

Exercices de probabilité niveau L (Jean-Pierre Lecoutre)

Loi conditionnelle discrète

15. On jette deux dés à six faces numérotées de 1 à 6 et on note X la v.a. qui représente la somme des deux chiffres obtenus et Y la v.a. qui est égale au plus grand de ces deux chiffres.

a) Déterminer la loi de probabilité du couple (X, Y) puis calculer $E(X)$. Les v.a. X et Y sont-elles indépendantes?

b) Déterminer la loi conditionnelle de Y pour $X = 6$, puis son espérance $E(Y|X = 6)$.

Analyse de l'énoncé et conseils. Les couples ordonnés (x_1, x_2) de résultats associés aux deux dés ont la même probabilité 1/36. Pour chacun d'eux il faut calculer la somme $x_1 + x_2$, qui donnera la valeur de X, et $\max\{x_1, x_2\}$ qui sera celle de Y, puis regrouper les couples de valeurs identiques ; leur nombre sera la probabilité, en 1/36-ème du couple obtenu. La loi conditionnelle est obtenue à partir de la colonne du tableau associée à $X = 6$, en divisant les probabilités de cette colonne par la probabilité marginale $P(X = 6)$.

Exercices de probabilité niveau L (Jean-Pierre Lecoutre)

Espérance conditionnelle

16. La loi de probabilité du couple (X, Y) est indiquée dans le tableau suivant, les valeurs étant exprimées en dizièmes :

Y \ X	1	2	3	4
1	1	1	2	1
2	0	1	1	0
3	1	0	0	2

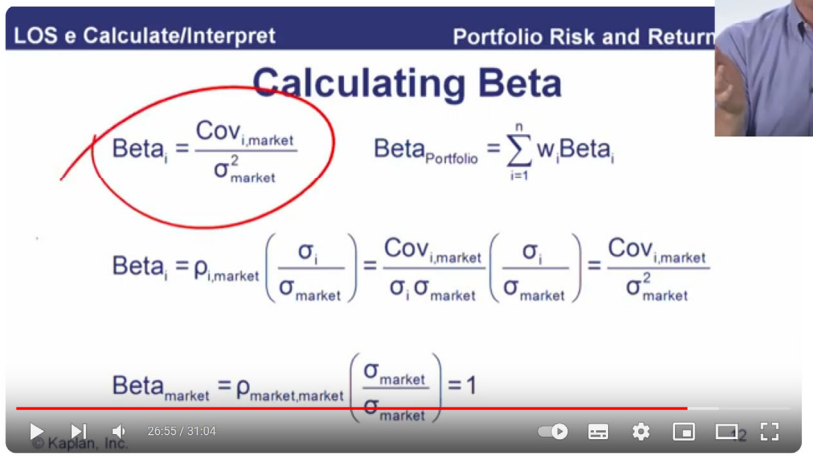
Déterminer la loi de probabilité de la v.a. $E(Y|X)$, puis calculer son espérance et la comparer avec la valeur de $E(Y)$.

Analyse de l'énoncé et conseils. L'expression $E(Y|X)$ est bien celle d'une variable aléatoire puisqu'il s'agit d'une fonction de X, qui est elle-même une v.a. Pour une réalisation ω telle que X prend la valeur $X(\omega) = x$, cette variable prend la valeur de la régression en x, c'est-à-dire que $E(Y|X)(\omega) = E(Y|X = x)$. Il faut donc calculer les différentes valeurs de cette fonction de régression, qui seront les valeurs possibles de cette v.a. On peut ensuite calculer l'espérance de cette v.a. par rapport à la loi de X.

Bêtas et formation des analystes financiers

Bêta : le coin des « apprentis praticiens »

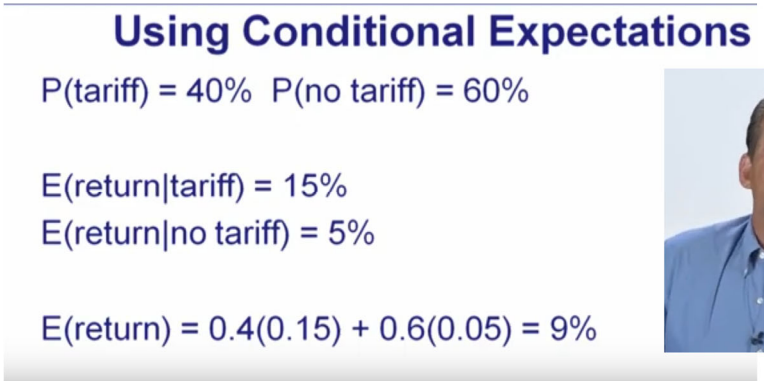
- https://www.youtube.com/watch?v=7nxKY_3eSvA



CFa Level I. Reading 40-1. Module 40 1 Systematic Risk and Beta

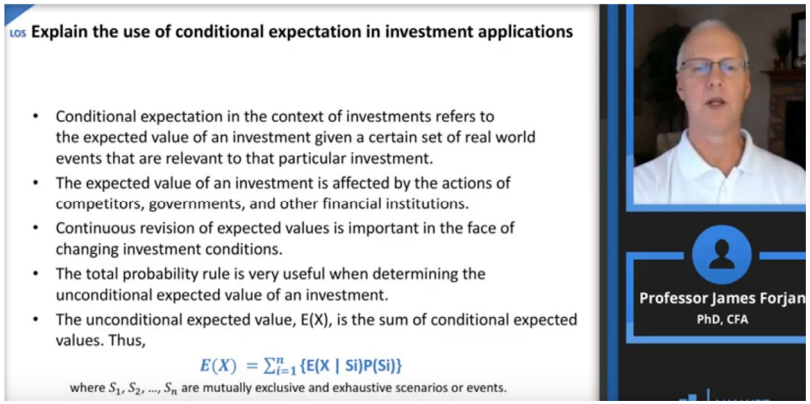
Lien avec le monde professionnel ?

- CFA Level I. Reading 9. Module 9 2 Conditional Expectations, Correlation
  - *Remarque : ici, on conditionne par rapport à des états du monde des affaires (scénarios) et pas par le niveau de rentabilité d'un indice*
  - *https://www.youtube.com/watch?v=jrNVQYzyJps*



Lien avec le monde professionnel ?

- Application de la propriété $E[Y] = E\big[E[Y|X]\big]$ pour candidats analystes financiers (CFA level 1)
  - *Remarque : Le conditionnement par rapport au niveau de rentabilité d'un indice est traité dans la partie « régression linéaire et Bêtas)*



Lien avec le monde professionnel ?

- Question CFA level 1
  - *https://analystprep.com/cfa-level-1-exam/quantitative-methods/conditional-expectations-investments/*

There is a 20% chance that the government will impose a tariff on imported cars. A company that assembles cars locally expects returns of 14% if the tariff is imposed and returns of 11% if the tariff is not imposed. Determine the (unconditional) expected return.

A. 11.6%

B. 12.8%

C. 12.5%

Solution

The correct answer is A.

The unconditional expected return will be the sum of:

(1) The expected return **given** *no tariff times the probability that a tariff will not be imposed, and*

(2) The expected return **given** *tariff times the probability that the tariff will be imposed. Thus,*

$$\begin{aligned} E(X) &= \sum \{E(X|S_i)\,P(S_i)\} \\ &= 0.11(0.8) + 0.14(0.2) \\ &= 0.116 \text{ or } 11.6\% \end{aligned}$$

Lien avec le monde professionnel ?

- Programme CFA level 2, partie méthodes quantitatives

Quantitative Methods

Lien avec le monde professionnel ?

- Programme CFA level 2, partie méthodes quantitatives

Lien avec le monde professionnel ?

- Programme CFA level 2, partie méthodes quantitatives

Lien avec le monde professionnel ?

- Programme CFA level 2, partie méthodes quantitatives

Lien avec le monde professionnel ?

- Programme CFA level 2, partie méthodes quantitatives

- + partie « big data » (learning module 7)

Le concept de Bêta d'un titre

La notion de Bêta

- r_i : rentabilité (aléatoire) d'un titre i (ou d'un portefeuille de titres)
- r_P : rentabilité d'un portefeuille de titres P (ou d'un indice boursier évoluant comme un portefeuille de titres)
- Pour simplifier l'exposé, on va supposer que l'on considère des **rentabilités r_i et r_p centrées**
- C'est-à-dire que l'on va considérer par la suite $r_i - E[r_i]$ et $r_P - E[r_P]$
  - *On remarque que* $E[r_i - E[r_i]] = E[r_i] - E[r_i] = 0$
  - *D'où le terme « centré »*
  - *Par abus de langage, on continue à noter* r_i *et* r_p *les rentabilités centrées*

La notion de Bêta

- Bêta du titre i :
  - *Noté* β_i
  - ***Nombre** qui indique de combien, **en moyenne**, la rentabilité du titre i augmente quand la rentabilité du portefeuille augmente :*
  - *Si le Bêta du titre i est égal à* 2 *et si la rentabilité du portefeuille P augmente de* 1%*, la rentabilité du titre i augmente **en moyenne** de* $\mathbf{2} \times 1\% = 2\%$
  - *Si le Bêta du titre i est égal à* 0,8*, rentabilité du titre i augmente **en moyenne** de* $\mathbf{0,8} \times 1\% = 0{,}8\%$
  - *En toute rigueur, on aurait dû indiquer qu'il s'agit du Bêta du titre i par rapport au portefeuille P*

La notion de bêta (suite)

- Le concept de bêta implique une **proportionnalité** de la relation (en moyenne) entre les rentabilités
  - *Exemple : Bêta du titre i égal à* 2
  - *Si la rentabilité du portefeuille P augmente de* 1%*, la rentabilité du titre i augmente **en moyenne** de* $\mathbf{2} \times 1\% = 2\%$
  - *Si la rentabilité du portefeuille P augmente de* 5%*, la rentabilité du titre i augmente **en moyenne** de* $\mathbf{2} \times 5\% = 10\%$
  - *Si la rentabilité du portefeuille P diminue de* 1%*, la rentabilité du titre i diminue **en moyenne** de* $\mathbf{2} \times 1\% = 2\%$

La notion de bêta (suite)

- Comment expliciter les concepts de « en moyenne » et de proportionnalité ? $\boldsymbol{r_i = \beta_i \times r_P + \varepsilon_i}$
- Proportionnalité liée au terme $\beta_i \times r_P$
  - *Si on laisse de côté le terme ε_i qui sera nul en moyenne, $r_i = \beta_i \times r_P$ (fonction linéaire de r_P)*
  - *Et non pas, par exemple, $r_i = \beta_i \times \left((1 + r_P)^{12} - 1\right)$*
- La rentabilité r_i augmente en moyenne de $\beta_i \times r_P$, si « en moyenne » ε_i est égal à zéro :
  - *$E[\varepsilon_i] = 0$, où $E[\varepsilon_i]$ est l'espérance de la variable aléatoire ε_i (parfois appelée « bruit »), est nulle.*
  - *Il faut que **pour un niveau quelconque de $\boldsymbol{r_P}$**, r_i augmente en moyenne de $\beta_i r_P$*
  - *Donc, l'espérance (moyenne) de ε_i doit être nulle **pour tout $\boldsymbol{r_P}$***

La notion de bêta (suite)

- L'espérance (moyenne) de ε_i doit être nulle **pour tout $\boldsymbol{r_P}$**
  - *Cela fait appel à la notion de moyenne (ou espérance) conditionnelle de ε_i « étant donné » (« sachant » par abus de langage) r_P*
- $E[\varepsilon_i|r_P]$ **: moyenne conditionnelle**
  - voir rappels de probabilités (transparents suivants)
  - *La condition s'écrit alors $E[\varepsilon_i|r_P] = 0$*
  - *On peut **démontrer** que cela **implique** que la covariance entre ε_i et r_P est nulle : $\mathrm{Cov}(\varepsilon_i, r_P) = 0$*
  - *Et que le coefficient de corrélation linéaire entre ε_i et r_P est nul car $\rho_{iP} = \frac{\mathrm{Cov}(\varepsilon_i, r_P)}{\sigma(\varepsilon_i) \times \sigma(r_P)} = 0$*

La notion de bêta (suite)

- Décomposition des rentabilités $\boldsymbol{r_i = \beta_i \times r_P + \varepsilon_i}$
  - *Remarque : si les variables aléatoires r_P et ε_i sont **indépendantes**, alors la condition $E[\varepsilon_i|r_P] = 0$ est vérifiée*
  - *Mais la réciproque n'est pas vraie*
- Repartons de la condition $\mathrm{Cov}(\varepsilon_i, r_P) = 0$
  - $\mathrm{Cov}(\varepsilon_i, r_P) = \mathrm{Cov}(r_i - \beta_i r_P, r_P) = \mathrm{Cov}(r_i, r_P) - \beta_i \mathrm{Cov}(r_P, r_P) = \mathrm{Cov}(r_i, r_P) - \beta_i \mathrm{Var}[r_P] = 0$
- $\boldsymbol{\beta_i = \frac{\mathrm{Cov}(r_i, r_P)}{\mathrm{Var}[r_P]}}$
- $\boldsymbol{\beta_i = \rho_{iP} \frac{\sigma_i}{\sigma_P}}$

Deux expressions à connaître du Bêta d'un titre

  - *Car $\mathrm{Cov}(r_i, r_P) = \rho_{iP}\sigma_i\sigma_P$ et $\mathrm{Var}[r_P] = \sigma_P^2$*

Bêta d'un actif sans risque, du portefeuille de marché P, d'un portefeuille de deux titres

- *Actif sans risque* $\boldsymbol{\beta_f = \frac{\mathrm{Cov}(r_f, r_P)}{\sigma_P^2} = 0}$
- *Portefeuille de marché P :* $\boldsymbol{\beta_P = \frac{\mathrm{Cov}(r_P, r_P)}{\sigma_P^2} = \frac{\sigma_P^2}{\sigma_P^2} = 1}$
- *Bêta d'un portefeuille de titres*
- $\boldsymbol{r = x_1 r_1 + (1 - x_1) r_2}$
- $\boldsymbol{\beta = \frac{\mathrm{Cov}(x_1 r_1 + (1 - x_1) r_2, r_P)}{\sigma_P^2}}$
- $\boldsymbol{\beta = \frac{x_1 \times \mathrm{Cov}(r_1, r_P) + (1 - x_1)\mathrm{Cov}(r_2, r_P)}{\sigma_P^2}}$

$$\beta = x_1\beta_1 + (1 - x_1)\beta_2$$

Bêta d'un portefeuille : moyenne pondérée des bêtas des titres le constituant.

Coefficients de pondération : fraction de la richesse investie dans chaque titre

Décomposition du risque total

Décomposition du risque total

- Décomposition de la rentabilité du titre i

$$r_i = E_i + \beta_i \times (r_P - E_P) + \varepsilon_i$$

- Décomposition du risque associé au titre i

$$\underbrace{\sigma_i^2}_{\text{risque total}} = \underbrace{\beta_i^2 \times \sigma_P^2}_{\text{risque du marché } P} + \underbrace{\text{Var}[\varepsilon_i]}_{\text{risque spécifique}}$$

Décomposition du risque total

- Décomposition de la rentabilité du titre i

$$r_i = E_i + \beta_i \times (r_P - E_P) + \varepsilon_i$$

- Décomposition du risque associé au titre i

$$\underbrace{\sigma_i^2}_{\text{risque total}} = \underbrace{\beta_i^2 \times \sigma_P^2}_{\text{risque du marché } P} + \underbrace{\text{Var}[\varepsilon_i]}_{\text{risque spécifique}}$$

σ_i écart-type de r_i, σ_P écart-type de r_P, β_i Bêta du titre i

Décomposition du risque total

- Décomposition de la rentabilité du titre i

$$r_i = E_i + \beta_i \times (r_P - E_P) + \varepsilon_i$$

- Décomposition du risque associé au titre i

$$\underbrace{\sigma_i^2}_{\text{risque total}} = \underbrace{\beta_i^2 \times \sigma_P^2}_{\text{risque du marché } P} + \underbrace{\text{Var}[\varepsilon_i]}_{\text{risque spécifique}}$$

σ_i écart-type de R_i, σ_P écart-type de r_P,
β_i Bêta du titre i (par rapport au marché P)

Lien entre β_i et coefficient de corrélation linéaire :

$$\boldsymbol{\beta_i = \frac{\rho_{iP} \times \sigma_i}{\sigma_P}}$$

Décomposition du risque total

- Décomposition de la rentabilité et du risque du titre i

$$r_i = E_i + \beta_i \times (r_P - E_P) + \varepsilon_i$$

- Décomposition du risque associé au titre i

$$\underbrace{\sigma_i^2}_{\text{risque total}} = \underbrace{\beta_i^2 \times \sigma_P^2}_{\text{risque du marché } P} + \underbrace{\text{Var}[\varepsilon_i]}_{\text{risque spécifique}}$$

Risque spécifique ε_i non corrélé au risque du portefeuille P

Décomposition du risque total

$$r_i = E_i + \beta_i \times (r_P - E_P) + \varepsilon_i$$

- D'où :

$$\text{Var}[r_i] = \text{Var}[E_i + \beta_i \times (r_P - E_P) + \varepsilon_i] = \text{Var}[\beta_i r_P + \varepsilon_i]$$

- On a utilisé $\mathbf{Var}[\boldsymbol{X} + \boldsymbol{a}] = \mathbf{Var}[\boldsymbol{X}]$ si $\boldsymbol{a}$ est constant.

$$\text{Var}[\beta_i r_P + \varepsilon_i] = \underbrace{\text{Var}[\beta_i r_P]}_{=\beta_i^2 \times \text{Var}[r_P]} + 2 \underbrace{\text{Cov}(\beta_i r_P, \varepsilon_i)}_{=\beta_i \times \text{Cov}(r_P, \varepsilon_i) = 0} + \text{Var}[\varepsilon_i]$$

$$\underbrace{\sigma_i^2}_{\text{risque total}} = \underbrace{\beta_i^2 \times \sigma_P^2}_{\text{risque du marché } P} + \underbrace{\text{Var}[\varepsilon_i]}_{\text{risque spécifique}}$$

Décomposition du risque total

- Décomposition du risque total σ_i^2 : $\sigma_i^2 = \beta_i^2 \sigma_P^2 + \sigma_{\varepsilon_i}^2$
- Pourcentage du risque total associé au risque du portefeuille de référence P : $\frac{\beta_i^2 \sigma_P^2}{\sigma_i^2} = R^2$
- R^2 : coefficient de détermination linéaire de Pearson
- Montrons que : $-1 \leq \rho_{iP} \leq 1$
  - *Comme* $\sigma_{\varepsilon_i}^2 \geq 0$, $R^2 = \frac{\beta_i^2 \sigma_P^2}{\beta_i^2 \sigma_P^2 + \sigma_{\varepsilon_i}^2} \leq 1$
  - $\beta_i = \rho_{iP} \frac{\sigma_i}{\sigma_P}$ *et* $R^2 = \frac{\beta_i^2 \sigma_P^2}{\sigma_i^2} \Rightarrow R^2 = \left(\rho_{iP} \frac{\sigma_i}{\sigma_P}\right)^2 \times \frac{\sigma_P^2}{\sigma_i^2} = \rho_{iP}^2$
  - *Comme* $R^2 = \rho_{iP}^2 \leq 1$, $-1 \leq \rho_{iP} \leq 1$
  - *Ce qu'on voulait démontrer*

Décomposition du risque et trigonométrie

- Soit E un espace vectoriel. Un produit scalaire $\langle .,. \rangle$ est une forme bilinéaire, symétrique positive de $E \times E$ dans $\mathbb{R}$
  - $\forall u, v \in E, \langle u, v \rangle = \langle v, u \rangle$ *(symétrie)*
  - $\forall p, u, v \in E, \forall \alpha, \beta \in \mathbb{R}\ \langle \alpha p + \beta u, v \rangle = \alpha \langle p, v \rangle + \beta \langle u, v \rangle$ *(linéarité)*
  - $\forall v \in E, \langle v, v \rangle \geq 0$ *et* $\langle v, v \rangle = 0 \Rightarrow v = 0$ *(positivité)*
- On définit alors la norme de $v \in E$ comme $\|v\| = \langle v, v \rangle^{1/2}$
- Considérons le plan formé à partir de deux vecteurs u, v
- On peut alors appliquer les concepts de géométrie plane
  - *Programme spé maths, 1er ci-dessous*

- **Calcul vectoriel et produit scalaire**

Contenus

- Produit scalaire à partir de la projection orthogonale et de la formule avec le cosinus. Caractérisation de l'orthogonalité.
- Bilinéarité, symétrie. En base orthonormée, expression du produit scalaire et de la norme, critère d'orthogonalité.

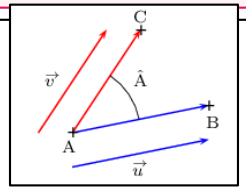
Décomposition du risque et trigonométrie

- Nombreux tutos pour le cours de première
  - $\langle u, v\rangle = \|u\| \times \|v\| \times \cos(u, v)$

PROPRIÉTÉ

Soient $\vec{u}$ et $\vec{v}$ deux vecteurs non nuls du plan. Alors on a :

$$\vec{u} \cdot \vec{v} = \|\vec{u}\| \times \|\vec{v}\| \times \cos(\vec{u}, \vec{v})$$



- Soit X, Y deux variables aléatoires. On définit $\langle X, Y\rangle = E[XY]$
  - *On peut vérifier qu'il s'agit bien d'un produit scalaire*
  - *On a supposé que l'espérance $E[XY]$ est bien définie*

Décomposition du risque et trigonométrie

- On prend $u = r_i - E_i$, $v = r_p - E_p$
- $\|u\| = (E[(r_i - E_i)^2])^{1/2} = \sigma_i$
- $\|v\| = (E[(r_P - E_P)^2])^{1/2} = \sigma_P$
- $\langle u, v\rangle = E\big[(r_i - E_i)(r_p - E_p)\big] = \mathrm{Cov}(r_i, r_P)$
- On rappelle que $\langle u, v\rangle = \|u\| \times \|v\| \times \cos(u, v)$
  - *Ce qui donne* $\mathrm{Cov}(r_i, r_P) = \sigma_i \times \sigma_P \times \cos(u, v)$
  - *Et comme* $\rho_{iP} = \frac{\mathrm{Cov}(r_i, r_P)}{\sigma_i \times \sigma_P}$
- $\rho_{iP} = \cos(u, v)$

Spécificités des risques spécifiques

Risque spécifique non corrélé au risque du marché P :

$$\mathrm{Cov}(\varepsilon_i, r_P) = 0$$

Les risques spécifiques peuvent-ils être corrélés ?

$$\boldsymbol{i \neq j, \mathrm{Cov}(\varepsilon_i, \varepsilon_j) = 0\ ?}$$

- *Prenons l'exemple de Peugeot et de Renault*
- *Du fait des risques spécifiques à l'industrie automobile …*
- ***Ils sont positivement corrélés***

- Ne pas confondre avec l'absence de corrélation entre risque spécifique et risque du marché P

Régression, corrélation et prévision

Rappels : Régression, corrélation linéaire, espérance conditionnelle

- r_i, r_P rentabilités centrées $(r_i - E_i, r_P - E_P)$
- Décomposition des rentabilités : $r_i = \beta_i \times r_P + \varepsilon_i$
  - ε_i : *risque spécifique*
- $\beta_i = \frac{\mathrm{Cov}(r_i, r_P)}{\mathrm{Var}[r_P]}$
- $\beta_i = \rho_{iP} \frac{\sigma_i}{\sigma_P}$

Deux expressions à connaître du Bêta d'un titre

- En particulier, si $\sigma_i = \sigma_P$, $\beta_i = \rho_{iP}$
  - *Lien étroit entre Bêta et coefficient de corrélation linéaire* ρ_{iP}
- $E[\varepsilon_i | r_P] = 0$: espérance conditionnelle nulle
- D'où $E[r_i | r_P] = \beta_i \times r_P$
  - *Meilleure prévision de* r_i *sachant* r_P

Régression et meilleure prévision

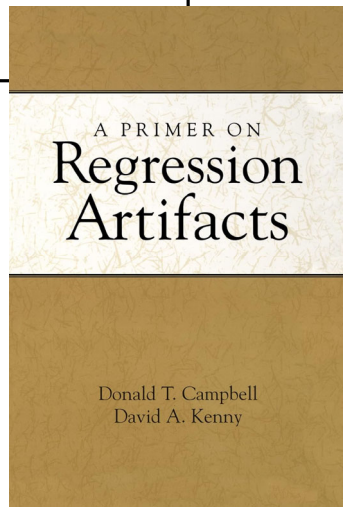
- $r_{i,t} = \beta_i \times r_{P,t} + \varepsilon_{i,t}$
- ε_i est l'erreur de prévision de r_i connaissant r_P
- MSE (Mean Square Error) : $\frac{1}{T}\sum_{t=1}^{T}(Predicted_t - Actual_t)^2$
- Predicted : $\beta_i \times r_{P,t}$,Actual : $r_{i,t}$
- $Predicted_t - Actual_t = \varepsilon_{i,t}$: erreur de prévision
- MSE $= E[\varepsilon_i^2]$
- Critère le plus couramment utilisé pour quantifier l'erreur de prévision : écart quadratique moyen
- RMSE (Root Mean Square Error) : $\sqrt{E[\varepsilon_i^2]}$
- Propriété : β_i minimise l'erreur de prévision

Régression et meilleure prévision

- Propriété : β_i minimise l'erreur de prévision $E[\varepsilon_i^2]$
- Démonstration : $\varepsilon_i = r_i - \beta r_P$
- $E[\varepsilon_i^2] = E[(r_i - \beta r_P)^2] = E[r_i^2 - 2\beta r_i r_P + \beta^2 r_P^2]$
  - *On rappelle que* r_i, r_P *rentabilités centrées* $(r_i - E_i, r_P - E_P)$
  - *Le critère s'écrit* $E[r_i^2] - 2\beta E[r_i r_P] + \beta^2 E[r_P^2]$
  - *Polynôme de degré deux en* β
  - *Le minimum s'obtient en dérivant par rapport à* β
  - *Ce qui donne* $-2E[r_i r_P] + 2\beta E[r_P^2] = 0$
- Soit $\beta = \frac{\mathrm{Cov}(r_i, r_P)}{\mathrm{Var}[r_P]}$
- Ce qu'on voulait démontrer

Comment le concept de corrélation (co-relation) a-t-il été découvert ?

> Regression to the mean is as inevitable as death and taxes.
> —Campbell and Kenny (1999, p. ix)



Definition

A person's score on a variable that is ***extreme*** (in the sense of being far away from the mean) ***tends*** to be less extreme ***in standard deviation units*** when that person is measured on ***another variable***.

Comment le concept de corrélation (co-relation) a-t-il été découvert ?

- Par Francis Galton, puis formalisation par Karl Pearson
  - Galton (1890). Kinship and correlation. *The North American Review*.
  - Stigler (1989). Francis Galton's account of the invention of correlation. *Statistical Science*.
- Cette découverte est l'aboutissement de plusieurs recherches, notamment
  - Galton (1886). Regression towards mediocrity in hereditary stature. *The Journal of the Anthropological Institute of Great Britain and Ireland*.
  - Galton (1889). Co-relations and their measurement, chiefly from anthropometric data. *Proceedings of the Royal Society of London*.
- Stanton (2001). Galton, Pearson, and the peas: A brief history of linear regression for statistics instructors. *Journal of Statistics Education*.
- Campbell & Kenny (2002). *A primer on regression artifacts*. Guilford Press.

Régression et corrélation linéaire : la notion de Bêta apparait dans un article de Galton

- Galton (1890). Kinship and correlation. *The North American Review*.

> Suppose there are three commercial ventures *a*, *b*, and *c*, whose daily profits vary independently of one another, and that a certain investor, whom we will call R, has one share in *a* and another in *b*, while a second investor, S, has several shares in *a* and others in *c*. The total profits, day by day, of R and of S will be related together because they are partly due to an investment in a common concern, but they will vary on different scales, because the ups and downs of the profits of R, who has only one share in *a*, must be less wide than those of S, who has many shares. The estimate that we may (and shall) find it possible to take of the probable profit of S on any particular day, from a knowledge of the profit of R on that day, would not work backwards without modification. There is not that *reciprocal* relation between them which is conveyed by the word correlation.
>
> 426

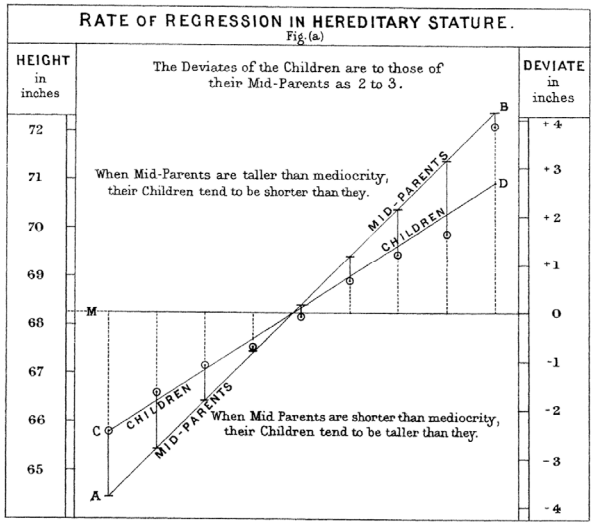
Régression et prévision

- Galton (1886). Regression towards mediocrity in hereditary stature. *The Journal of the Anthropological Institute of Great Britain and Ireland*.

- Quelle est la taille des enfants sachant celle des parents ?
  - *Galton a utilisé un échantillon de 930 fils (adultes) pour lesquels on connaissait la taille des parents.*
  - *En ce qui concerne la taille des parents :* $(1{,}08 \times t_m + t_f)/2$ *où* t_m, t_f *taille du père et de la mère et* $\bar{t}_f = 1{,}08 \times \bar{t}_m$
  - *Galton a considéré des écarts à la moyenne (à la médiane en fait)*
  - $\delta_p = (1{,}08 \times (t_m - \bar{t}_m) + t_f - \bar{t}_f)/2$, *cad à des variables centrées.*
  - *Les parents ont été classés en 9 catégories de taille*
  - *Pour chaque catégorie de taille, Galton a calculé l'écart de taille des enfants à la moyenne* $\delta_e = t_e - \bar{t}_e$
  - *Voir graphique suivant :* δ_p *en abscisses,* δ_e *en ordonnées*

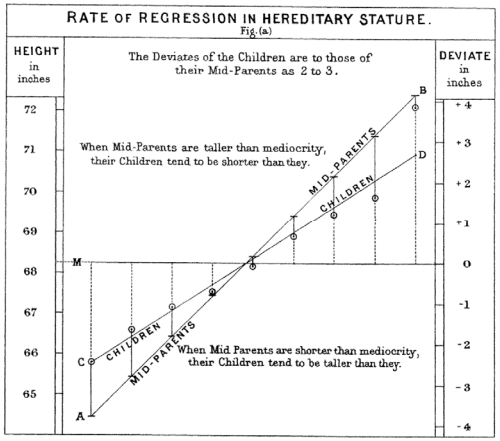
Régression et prévision

- Taille des parents δ_p en abscisses, δ_e des enfants en ordonnées



Régression et prévision

- Ratio des écarts (à la moyenne) de taille entre enfants et parents : indépendant de la taille des parents et est égal à $\frac{2}{3}$



- D'où l'origine du terme « régression vers la moyenne »

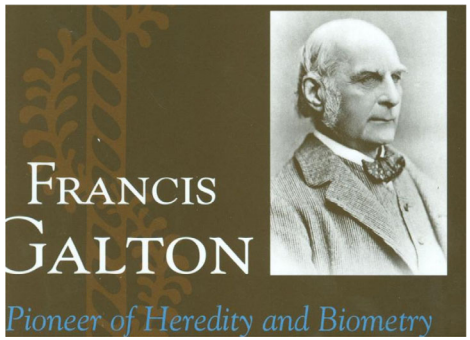
Régression et prévision

- Galton (1886). Regression towards mediocrity in hereditary stature. *The Journal of the Anthropological Institute of Great Britain and Ireland*.

> Elles donnèrent des résultats qui semblaient dignes d'intérêt, et je m'en servis pour étayer une conférence devant la Royal Institution le 9 février 1877. À la lueur de ces expériences, il apparaissait que les rejetons n'avaient *pas* tendance à ressembler à leurs parents par la taille, mais semblaient toujours être plus médiocres qu'eux – plus petits qu'eux si les parents étaient grands ; plus grands si les parents étaient très petits […]. Les expériences montrèrent en outre que la régression filiale moyenne vers la médiocrité était directement proportionnelle à la déviation parentale par rapport à ladite moyenne.

Régression et prévision

- Bulmer (2003). *Francis Galton: pioneer of heredity and biometry*.



> Many words in our statistical lexicon were coined by Galton. For example, *correlation* and *deviate* are due to him, as is *regression*, and he was the originator of terms and concepts such as *quartile*, *decile* and *percentile*, and of the use of *median* as the midpoint of a distribution[2]. Of

Régression et prévision

- Régression vers la moyenne
  - *Des parents très grands ont des enfants qui sont plus grands (que la moyenne) mais pas autant que les parents*
  - *Et pour les « petits parents », leurs enfants ne sont pas aussi petits que leurs parents*
- Que faut-il retenir à ce stade ?
  - *Pourquoi un ratio inférieur à 1 ?*
  - *Comment expliquer la relation entre la taille des enfants et des parents ?*
- Notons $u_p = \delta_p/\sigma_p$, l'écart de taille des parents, normalisé par l'écart-type de leur taille, σ_p (variable centrée et réduite)
  - u_p *est de moyenne nulle et d'écart-type égal à 1.*
- De même pour les enfants $u_e = \delta_e/\sigma_e$

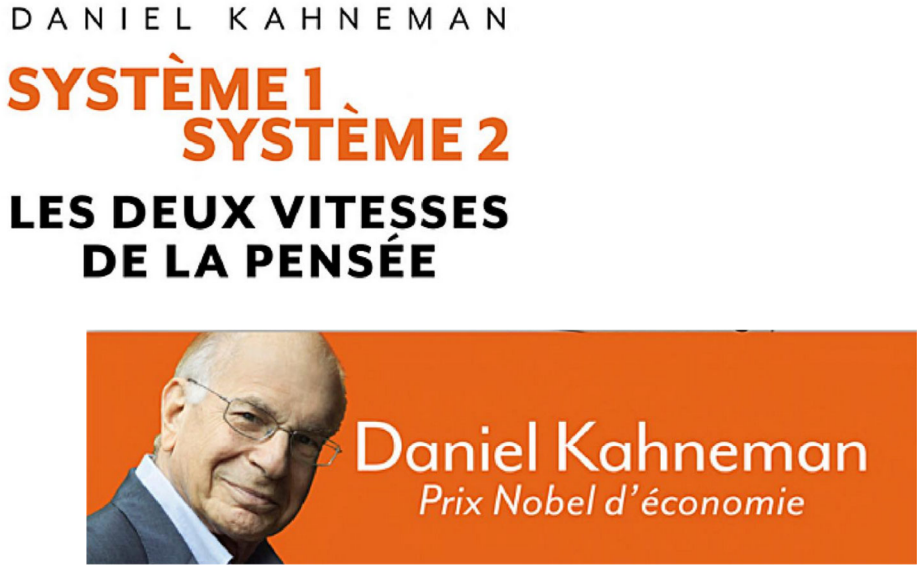
Régression, corrélation et prévision

- $u_p = \delta_p/\sigma_p\,, u_e = \delta_e/\sigma_e$ écarts de taille normalisés
  - $Var(u_p) = Var(u_e) = 1$
- Supposons les tailles déterminées par un facteur génétique g transmis par les parents et par un aléa, ε spécifique aux parents ou à l'enfant : $u_p = g + \varepsilon_p$, $u_e = g + \varepsilon_e$
  - *Le modèle proposé ici a un but purement illustratif.*
- Meilleure prévision de u_e sachant u_p ?
  - β *qui minimise l'erreur de prévision* $E\left[\left(u_e - \beta u_p\right)^2\right]$ *?*
  - $\beta = \mathrm{Cov}(u_e, u_p)/\mathrm{Var}(u_p) = \mathrm{Cov}(u_e, u_p)$
  - $\rho_{ep} = \mathrm{Cov}(u_e, u_p)/\sqrt{\mathrm{Var}(u_p)\mathrm{Var}(u_e)} = \mathrm{Cov}(u_e, u_p) = \beta$
- Meilleure prévision de u_e sachant u_p : $\rho_{ep}u_p$

Régression, corrélation et prévision

- Meilleure prévision de u_e sachant u_p : $\rho_{ep}u_p$
  - $\rho_{ep} = 1$. $u_p = u_e$. *Les écarts de taille normalisés entre parents et enfants sont égaux. Aucun aléa spécifique – tout génétique. Dans ce cas la meilleure prévision de* u_e *sachant* u_p *est* u_p
  - $\rho_{ep} = 0$. *Aucun effet de la génétique. La meilleure prévision de* u_e *sachant* u_p *est* 0.
  - $-1 \le \rho_{ep} \le 1$. *Meilleure prévision* u_e *sachant* u_p *:* $\left|\rho_{ep}u_p\right| \le \left|u_p\right|$.
- $\left|\rho_{ep}u_p\right| \le \left|u_p\right|$: retour à la moyenne lié au paramètre de corrélation ρ_{ep} et au problème de meilleure prévision.
- Biais cognitif et statistique : penser que la meilleure prévision de prévision u_e sachant u_p est u_p
  - *Les enfants de parents très grands ne sont pas aussi grands que leurs parents car la taille est en partie liée à « la chance »*

Régression, corrélation et prévision



Comment le concept de corrélation (co-relation) a-t-il été découvert ?

Régression, corrélation et prévision

Vous éprouvez sans doute de la sympathie pour les difficultés que rencontra Galton avec le concept de régression. Vous n'êtes pas le seul : le statisticien David Freedman avait coutume de dire que, quand la question de la régression est évoquée dans un procès pénal ou civil, la partie qui doit l'expliquer au jury est sûre de perdre. Pourquoi est-ce si difficile ? La principale raison de cette difficulté est un thème

- Kahneman, D (2016). *Système 1, système 2: les deux vitesses de la pensée*.
- Freedman, D. A. (2010). *Statistical models and causal inference: a dialogue with the social sciences*.

Régression, corrélation et prévision

- Pourquoi est-il difficile d'expliquer le concept de régression ?

récurrent de ce livre : notre esprit est profondément biaisé en faveur d'explications causales et gère mal les « simples statistiques ». Quand on attire notre attention sur un événement, la mémoire associative va en rechercher la cause – plus précisément, l'activation s'étendra automatiquement à toute cause déjà stockée dans la mémoire. Des explications causales seront évoquées quand une régression est détectée, mais elles seront fausses parce que la vérité, c'est que la régression vers la moyenne a une explication, mais elle n'a pas de cause. L'événement qui attire

Régression, corrélation et prévision

resterait sans doute pas longtemps sur nos écrans. Nos difficultés avec le concept de la régression trouvent leur origine autant dans le Système 1 que dans le Système 2. Sans instruction spécialisée, et dans bien des cas, même après avoir reçu une formation en statistiques, la relation entre la corrélation et la régression reste obscure. Le Système 2 peine à la comprendre et à l'apprendre. Cela est dû en partie au besoin insistant d'interprétations causales, une caractéristique du Système 1.

Les interprétations causales incorrectes des effets de la régression ne concernent pas que les lecteurs de la presse populaire. Le statisticien Howard Wainer a dressé une longue liste de chercheurs éminents qui ont commis la même erreur – ils ont confondu une simple corrélation avec la relation causale[7]. Les effets de régression sont une source courante de problèmes dans la recherche, et les scientifiques chevronnés nourrissent une saine méfiance face au piège de la déduction causale infondée.

Do scientists get fooled by Galton's regression to the mean? All the time!

- Senn (2011). Francis Galton and regression to the mean. *Significance*.

ject under consideration. I suspect that the regression fallacy is the most common fallacy in the statistical analysis of economic data, allevi-

- Friedman (1992). Do old fallacies ever die? *Journal of economic literature*.

Les interprétations causales incorrectes des effets de la régression ne concernent pas que les lecteurs de la presse populaire. Le statisticien Howard Wainer a dressé une longue liste de chercheurs éminents qui ont commis la même erreur – ils ont confondu une simple corrélation avec la relation causale[7]. Les effets de régression sont une source courante de problèmes dans la recherche, et les scientifiques chevronnés nourrissent une saine méfiance face au piège de la déduction causale infondée.

- Wainer (2000). Visual revelations: Kelley's paradox. *Chance*.
- Wainer (2007). The most dangerous equation. *American Scientist*.

- La formule de Tanner donne comme taille cible pour les enfants la moyenne de la taille des parents.

Régression, corrélation et prévision

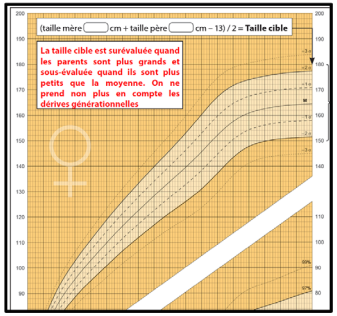
- La formule de Tanner donne comme taille cible pour les enfants la moyenne de la taille des parents (taille cible parentale).

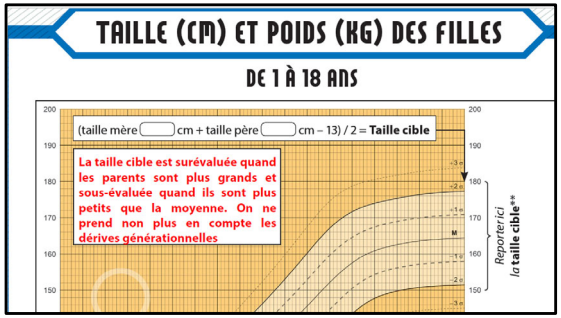
La taille cible (pages 84 et 86)

L'interprétation des mesures de taille tient compte de celles des parents. Pour cela, la formule de calcul de la taille cible parentale (en cm) retenue dans le référentiel national du Collège des Enseignants de Pédiatrie est proposée en haut des courbes de taille de un à 18 ans. Une flèche guide son report vers la fin de la courbe de l'enfant ce qui permet d'établir une distance (qui doit être exprimée en écarts-types) entre le couloir de croissance de l'enfant et la taille cible parentale. Le comité d'expertise a souhaité rappeler que 80% des enfants en bonne santé auront une taille finale comprise entre cette taille cible parentale – 6 cm et + 6 cm.

Régression, corrélation et prévision

- On dispose de données de corrélation entre la taille des enfants et celle des parents
  - *Ordres de grandeur compatibles avec les résultats de Galton*

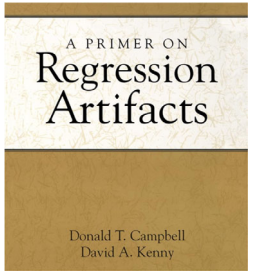




- C'est à confirmer, mais il semblerait que la méthodologie utilisée aboutisse à des biais pour les enfants de parents plus grands ou plus petits que la moyenne.

Régression, corrélation et prévision

- "Regression to the mean is an artifact that as easily fools statistical experts as lay people".



- "The universal phenomenon of regression toward the mean is just as universally misunderstood".
  - *Campbell and Kenny*

Corrélation, causalité

Corrélation, causalité

- Reprenons l'équation $r_i = E_i + \beta_{iP} \times (r_P - E_P) + \varepsilon_i$
  - *Ceci permet de déterminer r_i à partir de r_P et de ε_i*
- Il est erroné d'y voir une interprétation causale
  - *Où r_P serait la « cause » de r_i, affectée d'un bruit ε_i*
  - *Selon la terminologie de « variables expliquée et explicative »*
  - *Ou selon une approche du global au local (top-down)*
  - *Comme $r_P = \sum_i x_i r_i$, ce serait plutôt les r_i qui déterminent la rentabilité du portefeuille*
  - *On peut aussi considérer la relation $r_P = E_P + \beta_{Pi} \times (r_i - E_i) + \eta_i$*
  - *Dans ce cas, $\beta_{Pi} = \frac{\mathrm{Cov}(r_i, r_P)}{\mathrm{Var}(r_i)}$ et $\beta_{Pi} \times \beta_{iP} = \rho_{iP}^2$*
- Le caractère non causal de la régression apparait quand on régresse la taille des parents sur celle des enfants
  - *Comment la taille des enfants pourrait-elle causer celle des parents ?*

Corrélation, causalité et modèles à facteurs

Régression linéaire et prévision

- Y variable à prévoir
- X prédicteur (variable d'entrée)
- $a, b \in \mathbb{R}$: paramètres
- $a + bX$: règle de prévision linéaire, valeur prédite
- $Y - (a + bX) = \varepsilon$: erreur de prévision, résidu
- Erreur quadratique (MSE : Mean Square Error) : $E[\varepsilon^2]$
- Prévision optimale : trouver a, b qui minimisent la MSE

Régression linéaire et prévision

- Régression linéaire théorique
- ε, X, Y variables aléatoires
- $Y = a + bX + \varepsilon$
- $\min_{a,b} E[\varepsilon^2]$
- Ajustement par la méthode des moindres carrés
- $t = 1, \dots, T$ dates
- $X_t, t = 1, \dots, T$ valeurs observées du prédicteur
- $Y_t, t = 1, \dots, T$ valeurs observées de la variable à prévoir
- $\min_{a,b}(\varepsilon_1^2 + \cdots + \varepsilon_T^2)$

Régression linéaire et prévision

- Régression linéaire théorique
- ε, X, Y variables aléatoires
- $Y = a + bX + \varepsilon$
- $\min_{a,b} E[\varepsilon^2]$
- $a = E[X] - bE[Y]$
- $b = \frac{\mathrm{Cov}(X,Y)}{\sigma_X^2} = \rho_{XY}\frac{\sigma_X}{\sigma_Y}$
- $\mathrm{Cov}(X, \varepsilon) = 0$
- Analyse de la variance :
- $\sigma_Y^2 = b^2\sigma_X^2 + \sigma_\varepsilon^2$
- Ajustement par la méthode des moindres carrés
- $t = 1, \dots, T$ dates
- $X_t, t = 1, \dots, T$ valeurs observées du prédicteur
- $Y_t, t = 1, \dots, T$ valeurs observées de la variable à prévoir
- X_t, Y_t : réalisations des variables aléatoires X, Y
- $\min_{a,b}(\varepsilon_1^2 + \cdots + \varepsilon_T^2)$
- Minimisation de l'erreur de prévision dans l'échantillon

Régression linéaire et prévision

- Régression linéaire théorique
- ε, X, Y variables aléatoires
- $Y = a + bX + \varepsilon$
- $\min_{a,b} E[\varepsilon^2]$
- $a = E[X] - bE[Y]$
- $b = \frac{\mathrm{Cov}(X,Y)}{\sigma_X^2} = \rho_{XY}\frac{\sigma_X}{\sigma_Y}$
  - *Si* $\sigma_X = \sigma_Y$, $b = \rho_{XY}$
- $\mathrm{Cov}(X, \varepsilon) = 0$
- Analyse de la variance :
- $\sigma_Y^2 = b^2\sigma_X^2 + \sigma_\varepsilon^2$
- $R^2 = \frac{b^2\sigma_X^2}{\sigma_Y^2} = \rho_{XY}^2$
- Ajustement par la méthode des moindres carrés
- $t = 1, \dots, T$ dates
- $X_t, t = 1, \dots, T$ valeurs observées du prédicteur
- $Y_t, t = 1, \dots, T$ valeurs observées de la variable à prévoir
- X_t, Y_t : réalisations des variables aléatoires X, Y
- $\min_{a,b}(\varepsilon_1^2 + \cdots + \varepsilon_T^2)$
- Minimisation de l'erreur de prévision dans l'échantillon

Régression linéaire et prévision

- Régression linéaire théorique
- ε, X, Y variables aléatoires
- $Y = a + bX + \varepsilon$
- $\min_{a,b} E[\varepsilon^2]$
- $a = E[X] - bE[Y]$
- $b = \frac{\mathrm{Cov}(X,Y)}{\sigma_X^2} = \rho_{XY}\frac{\sigma_X}{\sigma_Y}$
  - *Si* $\sigma_X = \sigma_Y$, $b = \rho_{XY}$
- Analyse de la variance :
  - $\mathrm{Cov}(X, \varepsilon) = 0$
  - $\sigma_Y^2 = b^2\sigma_X^2 + \sigma_\varepsilon^2$
  - $R^2 = \frac{b^2\sigma_X^2}{\sigma_Y^2} = \rho_{XY}^2$
- Ajustement par la méthode des moindres carrés
- $t = 1, \ldots, T$ dates
- $X_t, t = 1, \ldots, T$ valeurs observées du prédicteur
- $Y_t, t = 1, \ldots, T$ valeurs observées de la variable à prévoir
- X_t, Y_t : réalisations des variables aléatoires X, Y
- $\min_{a,b}(\varepsilon_1^2 + \cdots + \varepsilon_T^2)$
- Minimisation de l'erreur de prévision dans l'échantillon

Régression linéaire et prévision

- Régression linéaire théorique
- ε, X, Y variables aléatoires
- $Y = a + bX + \varepsilon$
- $b = \frac{\mathrm{Cov}(X,Y)}{\sigma_X^2} = \rho_{XY}\frac{\sigma_X}{\sigma_Y}$
- Analyse de la variance :
  - $\mathrm{Cov}(X, \varepsilon) = 0$
  - $\sigma_Y^2 = b^2\sigma_X^2 + \sigma_\varepsilon^2$
  - $R^2 = \frac{b^2\sigma_X^2}{\sigma_Y^2} = \rho_{XY}^2$
- Estimateur des moindres carrés
- $\min_{a,b} \frac{1}{T}(\varepsilon_1^2 + \cdots + \varepsilon_T^2)$
- Probabilité empirique : équiprobable sur les dates
- $\frac{1}{T}(\varepsilon_1^2 + \cdots + \varepsilon_T^2) = E[\varepsilon^2]$
- $b = \frac{\mathrm{Cov}(X,Y)}{\sigma_X^2}$
- $\mathrm{Cov}(X,Y) = \frac{1}{T}\sum_{t=1}^{T}(X_t - \bar{X})(Y_t - \bar{Y})$
- $\bar{X} = \frac{1}{T}\sum_{t=1}^{T} X_t = E[X]$
- $\bar{Y} = \frac{1}{T}\sum_{t=1}^{T} Y_t = E[Y]$
- $\sigma_X^2 = \mathrm{Cov}(X,Y)$

Calculer des Bêtas à partir d'historiques de rentabilités

- $\boldsymbol{\beta_i} = \frac{\mathrm{Cov}(\boldsymbol{r_i}, \boldsymbol{r_P})}{\mathbf{Var}[\boldsymbol{r_P}]} = \rho_{iP}\frac{\sigma_i}{\sigma_P}$. Comment calculer $\rho_{iP}, \sigma_i, \sigma_P$?
- Supposons que nous ayons des séries chronologiques de rentabilités du titre i et du portefeuille P
- $\boldsymbol{r_{i,t}}$ représente la rentabilité du titre i à la date t
  - Il peut aussi s'agir de la rentabilité d'un portefeuille de titres
- $\boldsymbol{r_{P,t}}$ représente la rentabilité du portefeuille P à la date t
  - *P: Typiquement un indice boursier pondéré par la capitalisation boursière : CAC40, indice S&P 500, Russell 3000 Index aux États-Unis*
  - *Russell 3000 : Indice large, 98% de la capitalisation boursière US*

Calculer des Bêtas à partir d'historiques de rentabilités

- Probabilité empirique : équipondérée sur les dates passées
- Espérance (sous la mesure empirique) = moyenne des rentabilités passées
  - $E[\tilde{r}_i] = \bar{r}_i = \frac{1}{252} \times (r_{i,t-1} + \cdots + r_{i,t-252})$
  - $E[\tilde{r}_P] = \bar{r}_P = \frac{1}{252} \times (r_{P,t-1} + \cdots + r_{P,t-252})$
  - $\mathrm{Cov}[\tilde{r}_i, \tilde{r}_P] = E[(\tilde{r}_i - \bar{r}_i)(\tilde{r}_P - \bar{r}_P)]$
  - $\mathrm{Var}[\tilde{r}_P] = \mathrm{Cov}[\tilde{r}_P, \tilde{r}_P]$
  - $\beta_{iP} = \mathrm{Cov}[\tilde{r}_i, \tilde{r}_P]/\mathrm{Var}[\tilde{r}_P]$

Calculer des Bêtas à partir d'historiques de rentabilités

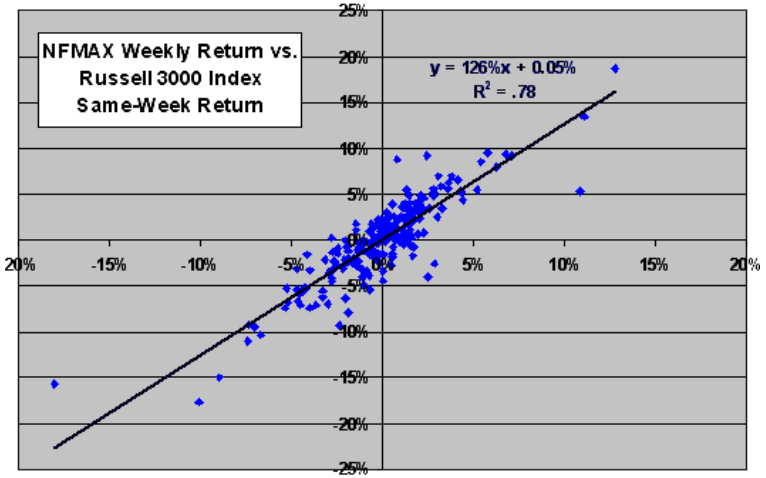
La droite caractéristique d'un titre

- On considère un plan où les valeurs de r_P sont en abscisses et les valeurs de r_i en ordonnées.
- Un point dans le plan est associé à chaque date t
- Coordonnées des points dans le plan : $(r_{P,t}, r_{i,t})$
- Dans le graphique qui suit :
  - $r_{i,t}$: *rentabilités* ***hebdomadaires*** *du fonds Navellier Fundamental A (NFMAX, fonds géré par Navellier)*
  - $r_{P,t}$: *Russell 3000 Index*
  - *Données de 2005 à 2009, soit 60 points*

La notion de Bêta

NAVELLIER
Calculated Investing

- Rentabilités du Russell 3000 Index Navellier en abscisses et de Fundamental A (NFMAX, fonds géré par Navellier) en ordonnées

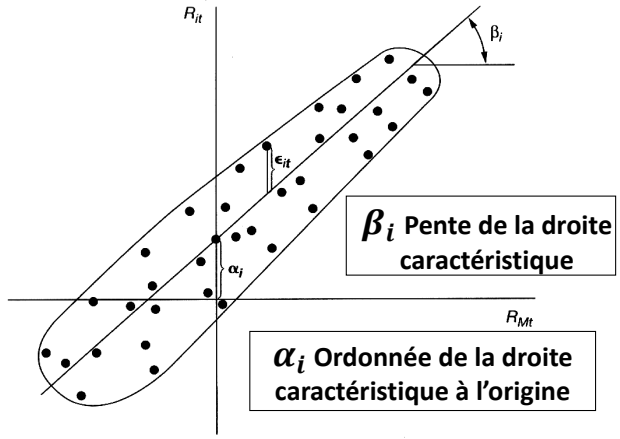


Droite caractéristique d'un titre, méthode des moindres carrés

- Droite caractéristique du titre $\boldsymbol{i}$: $\boldsymbol{y_{i,t} = \alpha_i + \beta_i r_{P,t}}$

Graphique 3.6 – *Droite caractéristique d'un titre* i

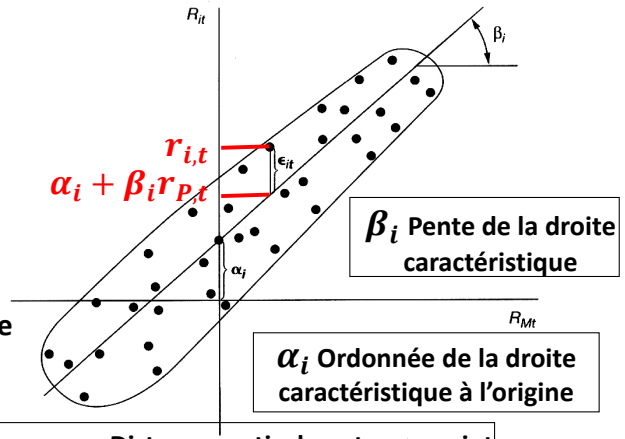
À chaque point est associé une date *t*
En abscisse, la rentabilité du portefeuille *P* à cette date *t* : $r_{P,t}$
En ordonnée, la rentabilité du titre *i* à cette date *t* : $r_{i,t}$



Droite caractéristique d'un titre, méthode des moindres carrés

- Droite caractéristique du titre $\boldsymbol{i}$: $\boldsymbol{y_{i,t} = \alpha_i + \beta_i r_{P,t}}$

Graphique 3.6 – *Droite caractéristique d'un titre* i



À chaque point est associé une date t

En abscisse, la rentabilité du portefeuille de marché à cette date t : $\boldsymbol{r_{M,t}}$

En ordonnée, la rentabilité du titre i à cette date t : $\boldsymbol{r_{i,t}}$

La droite caractéristique du titre est obtenu par un ajustement linéaire au nuage de points selon la méthode des moindres carrés

β_i Pente de la droite caractéristique

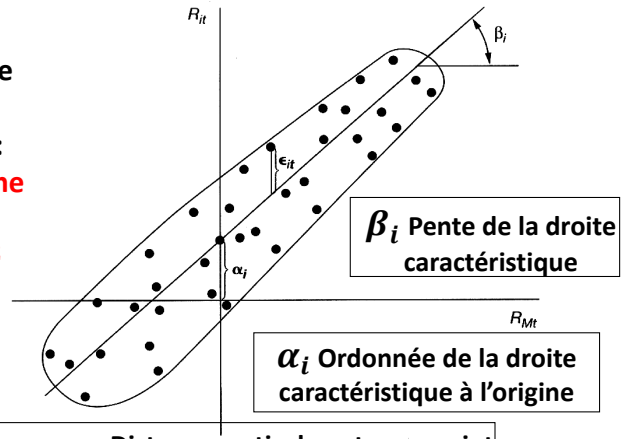
α_i Ordonnée de la droite caractéristique à l'origine

$\varepsilon_{i,t} = r_{i,t} - (\alpha_i + \beta_i r_{P,t})$ **Distance verticale entre un point et la droite caractéristique**

Droite caractéristique d'un titre, méthode des moindres carrés

- Droite caractéristique du titre $\boldsymbol{i}$: $\boldsymbol{y_{i,t} = \alpha_i + \beta_i r_{P,t}}$

Graphique 3.6 – *Droite caractéristique d'un titre* i



La droite caractéristique du titre est obtenu par un ajustement linéaire au nuage de points selon la méthode des moindres carrés (MCO) :

α_i et β_i minimisent la somme des carrés des écarts $\varepsilon_{i,t}$:

$\sum_t \left(r_{i,t} - (\alpha_i + \beta_i r_{P,t})\right)^2$

β_i Pente de la droite caractéristique

α_i Ordonnée de la droite caractéristique à l'origine

$\varepsilon_{i,t} = r_{i,t} - (\alpha_i + \beta_i r_{P,t})$ **Distance verticale entre un point et la droite caractéristique**

Estimation par la méthode des moindres carrés

- $r_{i,t} = \alpha_i + \beta_i r_{P,t} + \varepsilon_{i,t}$
- α_i, β_i constantes, β_i : bêta du titre i
- Méthode des moindres carrés ordinaires (MCO)
- On cherche α_i, β_i tels que la somme des écarts quadratiques $\left(r_{i,t-h} - \alpha_i - \beta_i r_{P,t-h}\right)^2$ soit minimale
- $\alpha_i + \beta_i r_{P,t}$ est une prévision linéaire de $r_{i,t}$ sachant $r_{P,t}$
- $\varepsilon_{i,t-h} = r_{i,t} - \alpha_i - \beta_i r_{P,t}$: erreur de prévision à la date t
- $\min_{\alpha_i,\beta_i} \sum_h \left(r_{i,t-h} - \alpha_i - \beta_i r_{P,t-h}\right)^2$
  - *Périodes glissantes* : $h = 1, \ldots, 252$

Estimation par la méthode des moindres carrés

- $\min_{\alpha_i,\beta_i} \sum_h \left(r_{i,t-h} - \alpha_i + \beta_i r_{P,t-h}\right)^2$
- Conditions du premier ordre : dérivées du critère par rapport à $\alpha_i, \beta_i = 0$
  - *Dérivée par rapport à* $\alpha_i = 0$
  - $\sum_h \left(r_{i,t-h} - \alpha_i - \beta_i r_{P,t-h}\right) = 0 \Rightarrow \alpha_i = \bar{r}_i - \beta_i \bar{r}_P$
  - *Il faut alors minimiser* $\sum_h \left(r_{i,t-h} - \bar{r}_i + \beta_i \left(r_{P,t-h} - \bar{r}_P\right)\right)^2$
  - *Dérivée par rapport à* $\beta_i = 0 \Rightarrow$
  - $\beta_i = \left(\frac{1}{252}\sum_{h=1}^{252}\left(r_{i,t-h} - \bar{r}_i\right)\left(r_{P,t-h} - \bar{r}_P\right)\right) / \left(\frac{1}{252}\sum_{h=1}^{252}\left(r_{P,t-h} - \bar{r}_P\right)^2\right)$

Calcul des bêtas à partir d'historiques de cours

- $\beta_i = \left(\frac{1}{252}\sum_{h=1}^{252}(r_{i,t-h} - \bar{r}_i)(r_{P,t-h} - \bar{r}_P)\right) / \left(\frac{1}{252}\sum_{h=1}^{252}(r_{P,t-h} - \bar{r}_P)^2\right)$
  - *$\bar{r}_i$ et $\bar{r}_P$ sont les rentabilités moyenne du titre i et du marché sur la période d'estimation, ici de longueur T*
  - $\bar{r}_i = \frac{1}{T}\sum_{h=1}^{252} r_{i,t-h}$
  - *β_i, Bêta ex-post (a posteriori) du titre i, déterminé à partir des rentabilités passées sur la période* $[t-1; t-252]$
    - $\frac{1}{252}\sum_h \varepsilon_{i,t-h} = 0$ le terme résiduel a une moyenne nulle
    - $\frac{1}{252}\sum_h \varepsilon_{i,t-h}(r_{M,t-h} - \bar{r}_P) = 0$ pas de corrélation entre résidu et rentabilité du portefeuille de marché

Calcul des bêtas à partir d'historiques de cours

- $\beta_{iP} = \frac{\mathrm{Cov}(r_i, r_P)}{Var[r_P]}$
  - *Se calcule à partir de la loi de probabilité de* (r_i, r_P)
- $\beta_{iP} = \left(\frac{1}{252}\sum_{h=1}^{252}(r_{i,t-h} - \bar{r}_i)(r_{P,t-h} - \bar{r}_P)\right) / \left(\frac{1}{252}\sum_{h=1}^{252}(r_{P,t-h} - \bar{r}_P)^2\right)$
  - *Se calcule à partir des données passées*
- Le choix de périodes glissantes et de rentabilités quotidiennes est fréquent en finance
  - *Notamment quand on s'intéresse à la dynamique des Bêtas.*
  - *On peut choisir des périodes d'estimation plus longues*
  - *Dans l'approche économétrique standard , on préfère utiliser toutes les données disponibles*
  - *Ce qui permet d'utiliser des théorèmes limites (loi des grands nombres) dans leur domaine de validité.*

Calcul des bêtas à partir d'historiques de cours

- <u>Rappel</u> : $\frac{1}{252}\sum_{h=1}^{252}(r_{i,t-h} - \bar{r}_i)(r_{P,t-h} - \bar{r}_P)$ correspond à la covariance « dans l'échantillon des données historiques » ou covariance empirique.
  - *covariance (dite empirique) entre les rentabilités r_i et r_P pour une loi de probabilité discrète, dont les valeurs sont les couples de rentabilité observées $(r_{P,t}, r_{i,t})$ et où ces valeurs sont équipondérées (on donne une probabilité de $\frac{1}{251}$ à chaque valeur.*
  - *Cette loi de probabilité discrète s'appelle la mesure empirique (ou loi empirique)*

Droite caractéristique d'un titre

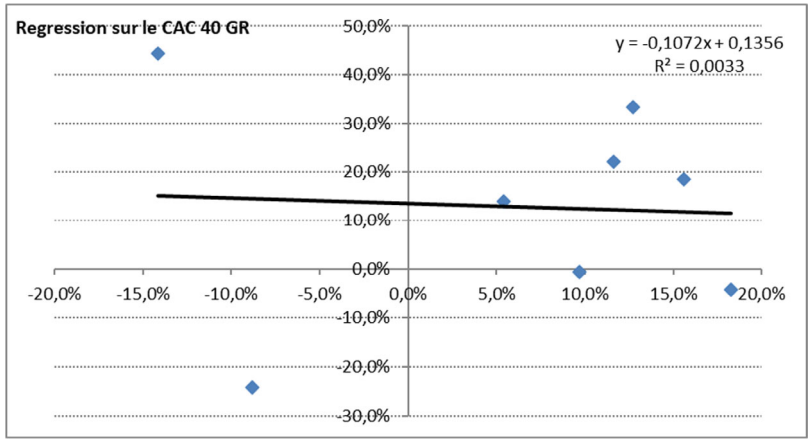
- Quelles données pour déterminer les Bêtas ?
  - *Quel marché de référence retenir?*
    - Quand l'action est cotée sur différents marchés
  - *Les rentabilités de l'action American Express ne sont pas les mêmes en $, £, €, etc. Quelle devise de référence*
  - *Quels cours boursiers: pertinence des cours de clôture?*
    - Pour les actions peu fréquemment traitées, le dernier cours peut être antérieur d'une heure ou deux à l'heure de clôture
    - Y a-t-il un fixing pour déterminer le cours de clôture ?
    - Différence entre heures de clôture pour différents marchés
  - *Disponibilité des données : suspension de cotation, actions nouvellement introduites en Bourse*
    - Il existe des algorithmes pour "compléter les données manquantes"
    - Brownian bridge, algorithme EM, Singular Spectrum Analysis

Droite caractéristique d'un titre

- Quelles données pour déterminer les Bêtas ?
  - *Période d'observation : date de début, date de fin ?*
  - *Périodicité des données: quotidienne, hebdomadaire,… Fréquence de mise à jour pour le calcul: mensuelle, trimestrielle, annuelle, jamais ?*
  - *Choix du portefeuille de référence P*
    - Indice national ou international ? Sectoriel ?
    - Large caps ?
    - Benchmark pour un investisseur, portefeuille efficient ?
- Risque de choix opportunistes pour influer sur les décisions d'investissement
  - *Augmenter le Bêta d'une branche d'un groupe industriel revient à augmenter les exigences de rentabilité*

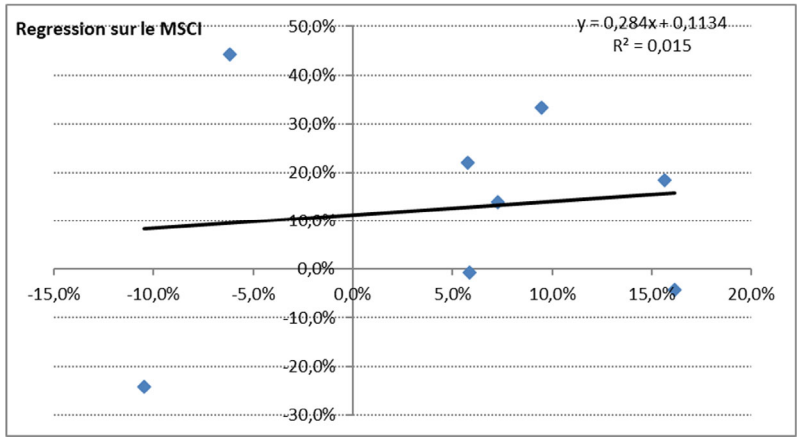
Estimation par la méthode des moindres carrés : choix de l'indice de référence

- Bêta de Michelin par rapport au CAC40 : -0,107



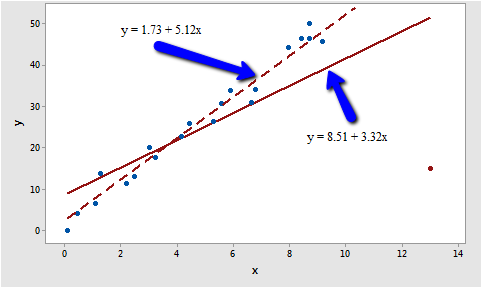
Estimation par la méthode des moindres carrés

- Bêta de Michelin par rapport au MSCI : 0,284



Calcul des Bêtas : le problème des « outliers »

- Crise covid : rentabilités hebdomadaires extrêmes (négatives) en 2020 à la fois pour les titres et l'indice.
- L'ajout d'une seule donnée (influential point), ici point rouge peut complètement modifier le Bêta



- Estimation par moindres carrés des Bêtas non robuste à la présence de certains « outliers »

Bêtas conditionnels

THE JOURNAL OF FINANCE • VOL. LI, NO. 1 • MARCH 1996

The Conditional CAPM and the Cross-Section of Expected Returns

RAVI JAGANNATHAN and ZHENYU WANG*

ABSTRACT

Most empirical studies of the static CAPM assume that betas remain constant over time and that the return on the value-weighted portfolio of all stocks is a proxy for the return on aggregate wealth. The general consensus is that the static CAPM is unable to explain satisfactorily the cross-section of average returns on stocks. We assume that the CAPM holds in a conditional sense, i.e., betas and the market risk premium vary over time. We include the return on human capital when measuring the return on aggregate wealth. Our specification performs well in explaining the cross-section of average returns.

Médaf et bêtas conditionnels

- On remplace tous les moments (espérances, variances, Bêtas) par les moments conditionnels
- La relation rentabilité risque reste de la même forme que dans le cas statique.

THE CONDITIONAL CAPM. *For each asset i and in each period t,*

$$E[R_{it}|I_{t-1}] = \gamma_{0t-1} + \gamma_{1t-1}\beta_{it-1}, \quad (2)$$

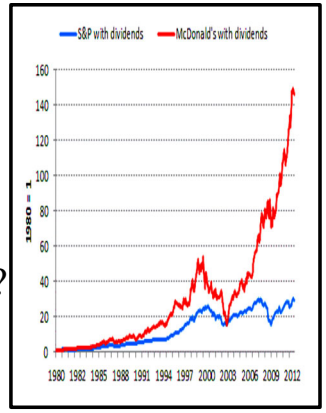
where β_{it-1} is the conditional beta of asset i defined as

$$\beta_{it-1} = \mathrm{Cov}(R_{it}, R_{mt}|I_{t-1})/\mathrm{Var}(R_{mt}|I_{t-1}), \quad (3)$$

γ_{0t-1} is the conditional expected return on a "zero-beta" portfolio, and γ_{1t-1} is the conditional market risk premium.

Droite caractéristique d'un titre : Estimation des betas

- Le cas McDonald's Corp (MCD)
  - *Yahoo Finance fournit un Beta de 0,45*
  - *Google Finance nous indique un Beta de 0,34*
  - *Site du Nasdaq : Beta de 0,56*
- Comment expliquer ces différences ?
  - *Indice utilisé pour le portefeuille de marché ?*
  - *Quelle est la période d'estimation ?*
  - *Fréquence des rentabilités observées ?*



Action McDo vs Indice S&P500 dividendes réinvestis

Droite caractéristique d'un titre : Estimation des betas

- Le cas McDonald's Corp (MCD)
  - ***Rentabilité historique de 17%**, celle de l'indice S&P500 de 11,2%*
  - *r_f = 2,37% taux 10 ans US Treasuries 10/2014*
  - *Choisissons une prime de risque de 6% (cf étude JP Morgan, 2008)*
- **On a alors une rentabilité attendue de 4,4%, 5,1% ou 5,7% selon le Beta retenu**
  - *8,4% pour l'indice*

- Le cas McDonald's Corp (MCD)
  - ***Rentabilité historique de 17%**, celle de l'indice S&P500 de 11,2%*
  - *r_f = 2,37% taux 10 ans US Treasuries 10/2014*
  - *Choisissons une prime de risque de 6% (cf étude JP Morgan, 2008)*
- **On a alors une rentabilité attendue de 4,4%, 5,1% ou 5,7% selon le Beta retenu**
  - *8,4% pour l'indice*

Action McDo vs Indice S&P500 dividendes réinvestis

Décorrélation des actions constituant l'indice Dow Jones et l'indice.

Comparaison entre l'été 2015 et l'été 2016 (calculs sur 50 jours glissants antérieurs au 9 septembre) On remarque aussi la distribution des Bêtas.

$$\beta_i = \rho_{iP}\frac{\sigma_i}{\sigma_P}$$

Comme les Bêtas restent en moyenne égaux à 1, il faut que les ratios $\frac{\sigma_i}{\sigma_P}$ augmentent.

Ce qui veut dire que les risques idiosyncratiques ont augmenté

Out of step

Last year, two-thirds of the Dow 30 were close (r >= 0.8) to the market. Now, it's just one.

Stock	Correlation 2015	Correlation **2016**	β
AXP	0.73	**0.81**	1.38
MMM	0.85	**0.79**	0.90
IBM	0.80	**0.79**	1.10
UTX	0.74	**0.72**	1.15
HD	0.88	**0.71**	1.09
GE	0.91	**0.71**	0.98
GS	0.92	**0.70**	1.39
BA	0.87	**0.70**	1.29
INTC	0.78	**0.70**	1.34
CSCO	0.83	**0.70**	0.97
JPM	0.93	**0.64**	1.00
DD	0.73	**0.63**	1.12
TRV	0.83	**0.61**	0.75
CAT	0.70	**0.60**	1.45
MSFT	0.86	**0.60**	1.07
KO	0.81	**0.59**	0.88
CVX	0.72	**0.58**	1.08
DIS	0.61	**0.58**	0.61
VZ	0.86	**0.57**	0.91
PFE	0.91	**0.49**	0.73
XOM	0.82	**0.47**	0.86
JNJ	0.90	**0.46**	0.46
PG	0.81	**0.42**	0.45
NKE	0.93	**0.41**	0.83
AAPL	0.79	**0.41**	0.95
MRK	0.83	**0.40**	1.15
WMT	0.77	**0.39**	0.54
UNH	0.84	**0.38**	0.44
V	0.80	**0.34**	0.55
MCD	0.89	**0.28**	0.49

Note: Correlation is 50-day rolling correlation with S&P 500 as of Sept. 9.
Source: Yahoo Finance; CNBC calculations

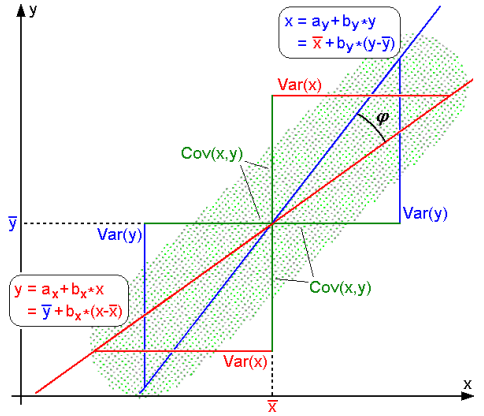
CNBC

Bêtas : prévision de $r_{i,t}$ sachant $r_{P,t}$ et de $r_{P,t}$ sachant $r_{i,t}$

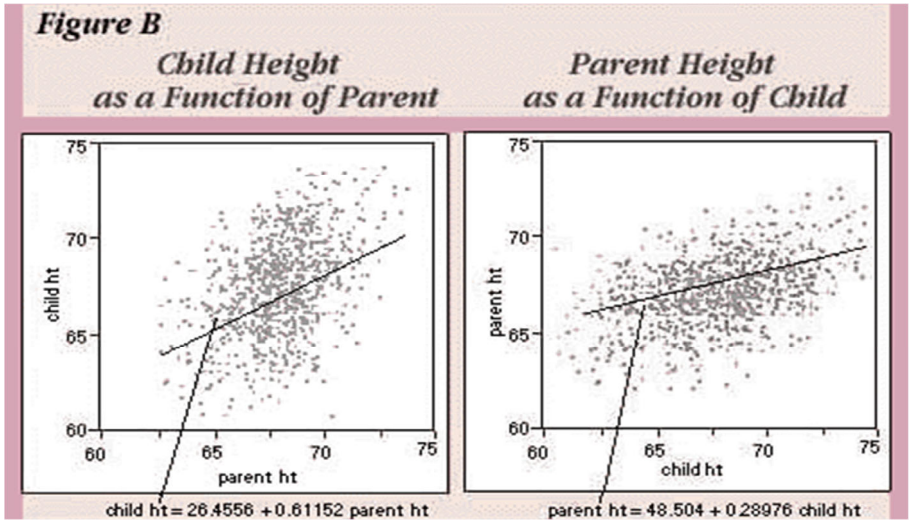
- $\beta_{iP} = \frac{\mathrm{Cov}(r_i, r_P)}{\mathrm{Var}[r_P]} = \rho_{iP}\frac{\sigma_i}{\sigma_P}$, $\beta_{Pi} = \frac{\mathrm{Cov}(r_i, r_P)}{\mathrm{Var}[r_i]} = \rho_{iP}\frac{\sigma_P}{\sigma_i}$
- D'où $\beta_{iP} \times \beta_{Pi} = \rho_{iP}^2 \leq 1$

Construction géométrique des deux droites de régression.

Les deux droites passent par $(\bar{x}, \bar{y})$
Pour obtenir la droite rouge (prévision de y sachant x, on se dirige vers le haut de $\mathrm{Cov}(x,y)$, puis vers la droite de $\mathrm{Var}[x]$, ce qui donne une pente de $\frac{\mathrm{Cov}(x,y)}{\mathrm{Var}[x]} = \beta_{y|x}$



Prévoir la taille des parents sachant celle des enfants est un problème différent de prévoir la taille des enfants sachant celle des parents

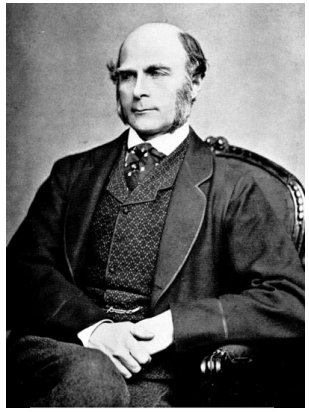


Galton : Bêta enfant/parent (0,61) très différent de l'inverse du Bêta parent/enfant (0,29), le ratio entre les deux Bêtas est le ratio entre la variance de la taille des parents et celle des enfants

Bêtas et corrélation

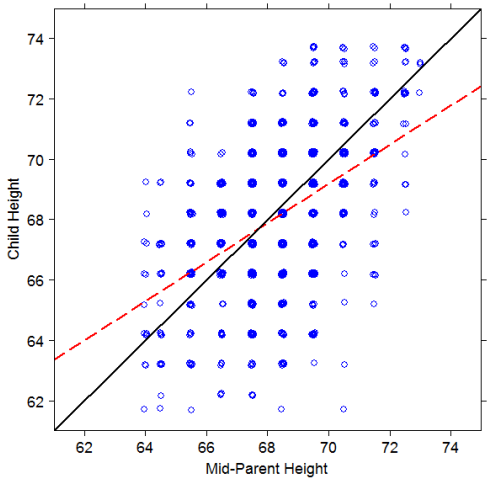
Pour aller plus loin, voir Kahneman, p. 219-225

- Regression to the mean (retour à la moyenne)
  - *Analyse du phénomène due à Francis Galton (1886)*
    - Selon les observations de Galton, si les parents mesurent 6 cm de plus que la moyenne, les enfants ne mesurent plus que **2/3** x 6 cm de plus que la moyenne
    - D'où l'origine du terme "**regression**"
      - On parle parfois de "régression vers la médiocrité"
  - *Ceci est exact, mais on a $\bar{y}_i = \beta_{y|x}\bar{x}_i + \varepsilon_i$.*
  - *Supposons $\sigma_x \approx \sigma_y$ (même variabilité de taille pour parents et enfants)*
    - On reviendra sur cette hypothèse ultérieurement
  - $\beta_{y|x} = \rho_{xy}\frac{\sigma_y}{\sigma_x} \approx \rho_{xy} \leq 1$



Francis Galton

Galton s'attendait à ce que la meilleure prévision soit associée à une pente de 1 (en noir) et pas à une pente de 0,61 (en rouge) : « les enfants héritent des caractéristiques biologiques de leurs parents »



En effet, le nuage de points forme une ellipse. On peut montrer sous certaines conditions (même variabilité des tailles chez les parents et les enfants) que le grand axe de l'ellipse a une pente de 1 (45°)

Bêtas et corrélation

- Pourquoi la meilleure prévision de la taille des enfants n'est-elle pas la taille des parents ?
  - $\bar{y}_i = y_i - m$, *taille des enfants,* $\bar{x}_i = x_i - m$ *taille des parents*
  - *Galton pensait que l'on aurait une relation du type* $\bar{y}_i = 1 \times \bar{x}_i +$ terme résiduel, $\beta_{y|x} = 1$
  - *Or, l'estimation de* $\beta_{y|x}$ *par la méthode des moindres carrés donne* $\beta_{y|x} = 0{,}61 = \rho_{xy}\frac{\sigma_y}{\sigma_x}$, *d'où le résultat « l'écart à la moyenne de la taille des enfants est environ 2/3 de l'écart des parents à la moyenne (de leur génération)*
  - $\bar{y}_i = \beta_{y|x}\bar{x}_i +$ terme résiduel
- Si la taille des parents est grande, ce pourrait être en partie dû « à la chance » et pour le reste aux gènes qu'ils transmettront à leurs enfants : la chance ne se transmet pas

Bêtas et corrélation

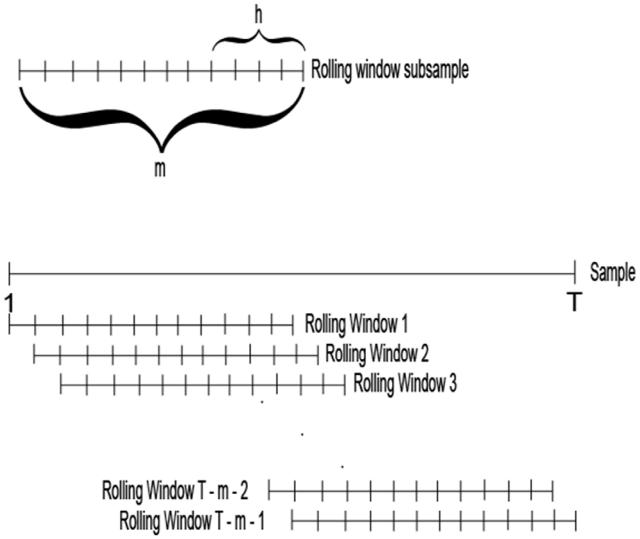
- Aurait-on $\beta_{x|y} = 1/\beta_{y|x} = 1{,}64$ (analyse de sensibilité) ?
  - *Or* $\beta_{x|y} = 0{,}29 < 1$!
  - $\beta_{x|y} = \rho_{xy}\frac{\sigma_x}{\sigma_y}$, $\beta_{y|x} = \rho_{xy}\frac{\sigma_y}{\sigma_x}$
  - *D'où* $\beta_{x|y} \times \beta_{y|x} = \rho_{xy}^2 < 1$ *(sauf si* $\rho_{xy} = 1$, *cas où la liaison entre x et y devient déterministe), ce qui donne* $\rho_{xy} = 42\%$
  - *Remarque :* $\frac{\sigma_y^2}{\sigma_x^2} = \beta_{y|x}/\beta_{x|y} = 0{,}61/0{,}29 \approx 2$,
  - $\sigma_x^2 \approx \frac{1}{2} \times \sigma_y^2$ *? x_i est la moyenne de la taille des deux parents (pondérée par un facteur 1,08 pour la taille des mères)*
  - $\bar{x}_i \approx \frac{1}{2} \times \left(\bar{x}_i^P + \bar{x}_i^M\right) \Rightarrow \sigma_{\bar{x}}^2 \approx \frac{1}{4} \times \left(\sigma_{\bar{x}^P}^2 + \sigma_{\bar{x}^M}^2\right)$ *si non corrélation entre les tailles du père et de la mère*

La dynamique des Bêtas

Bêta (ex-ante) d'un titre, d'un portefeuille

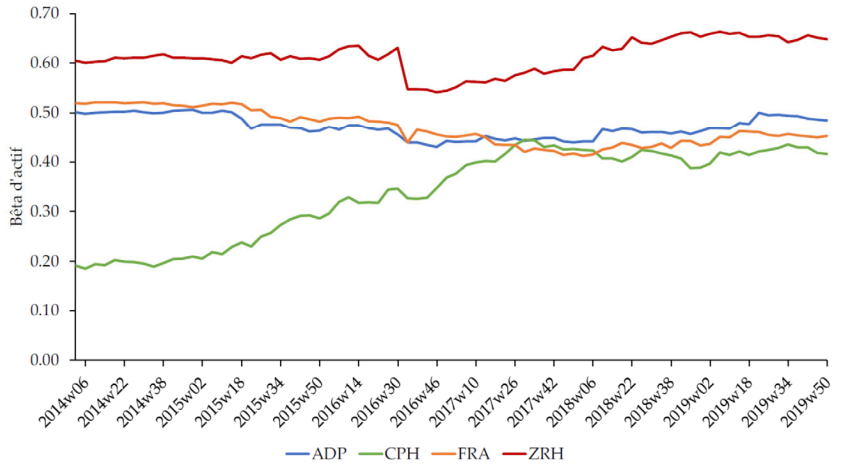
- On s'intéresse aux rentabilités **<u>à venir</u>** r_i et r_P
  - r_i *et* r_P *variables aléatoires*
- $r_i = E_i + \beta_i \times (r_P - E_P) + \varepsilon_i$
  - ε_i *<u>risque spécifique</u> ou <u>risque idiosyncratique</u> du titre i*
  - *De moyenne nulle, non corrélé avec le risque de marché* r_P
  - E_i *: espérance de rentabilité du titre i*
  - E_P *: espérance de rentabilité du portefeuille de marché*
- Si r_P augmente de 1%, r_i augmente <u>en moyenne</u> de β_i%
- β_i Bêta (ex-ante ou prospectif) du titre *i*
  - E_i, E_P, β_i *paramètres (non aléatoires)*

Estimation des Bêtas avec des fenêtres glissantes de longueur donnée, « rolling window », rolling regressions : les bêtas vont être amenés à fluctuer au cours du temps



Dynamique des Bêtas

Figure 2 : Bêtas des actifs des aéroports comparables sur une période glissante de 5 ans



Trinkner, U., Binz, T., & Mattmann, M. (2020). Bêtas des aéroports français à partir de l'observation des marchés boursiers et de précédents de régulation. Note établie pour l'autorité de régulation des transports.

Robert Engle

Nested Asynchronous Model

- Combining the constant beta and dynamic conditional beta into one regression:

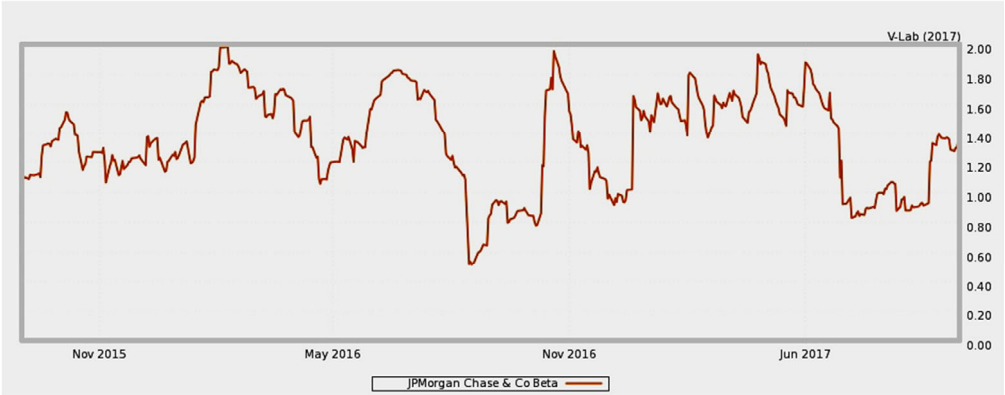
$$R_{i,t} = \left(\phi_1 + \lambda_1 \beta_{i,t}\right) R_{m,t} + \left(\phi_2 + \lambda_2 \gamma_{i,t}\right) R_{m,t-1} + u_t$$

- Where u will be an MA(1) GARCH
- Updated weekly for 1200 Global financial institutions
- Results on V-LAB

Dynamic Conditional Betas in V-Lab

https://vlab.stern.nyu.edu/

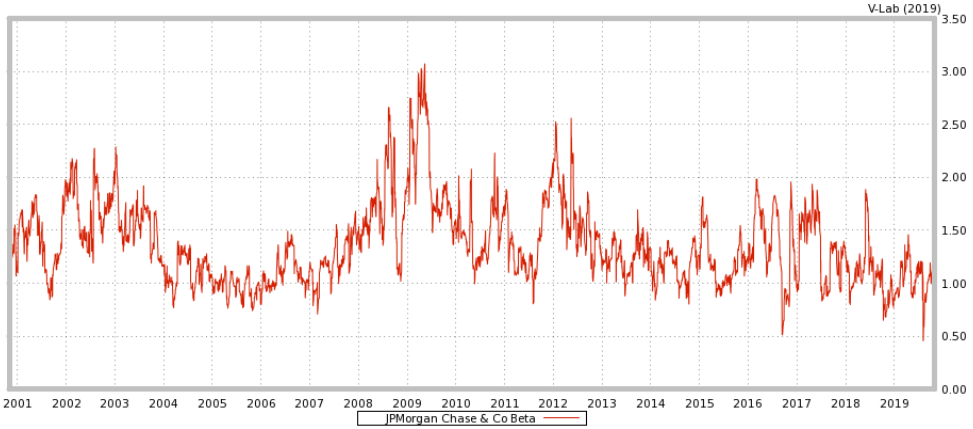
Évolution du Bêta de l'action JP Morgan au cours des deux dernières années (octobre 2015 – octobre 2017)



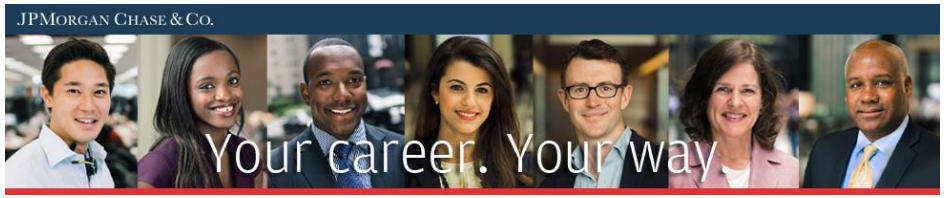
Le Bêta varie entre 0,5 et 2 avec des fluctuations extrêmement rapides. C'est tout sauf une constante caractéristique de l'action.

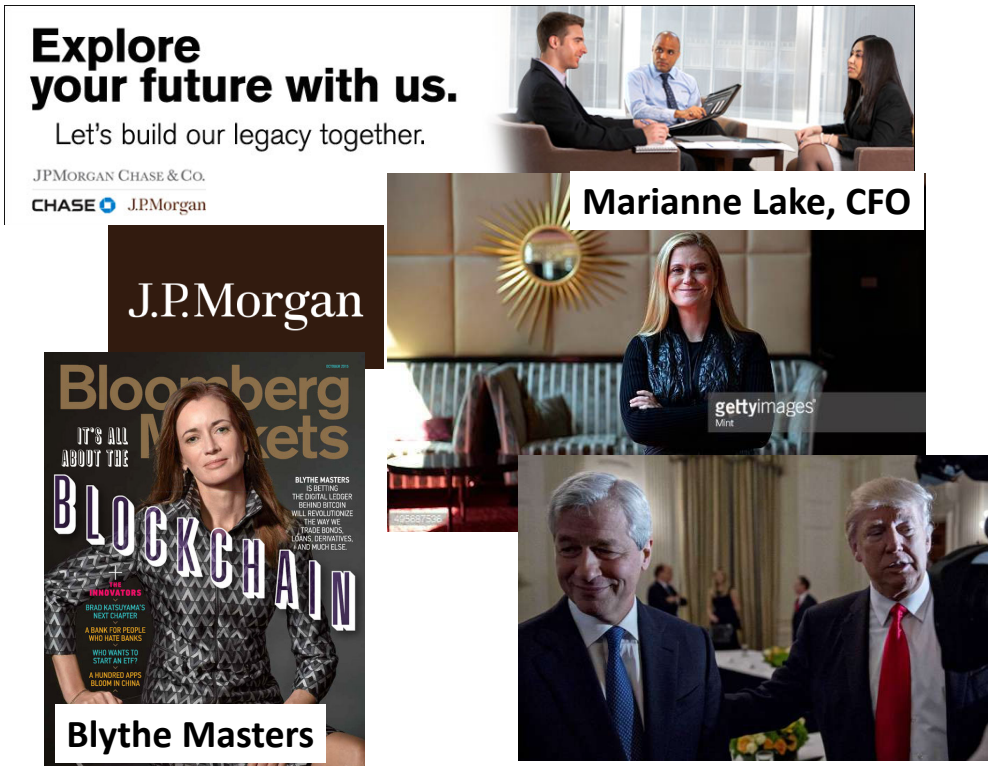
Engle (2016). Dynamic conditional bêta. Journal of Financial Econometrics.
https://vlab.stern.nyu.edu/analysis/RISK.USFIN-MR.MES

Dynamic Conditional Beta de l'action JP Morgan : retour à la moyenne ?



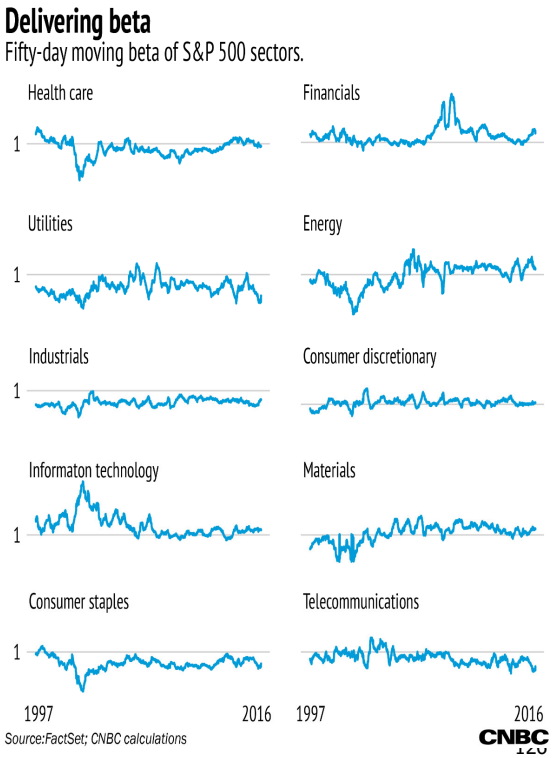
https://vlab.stern.nyu.edu/analysis/RISK.USFIN-MR.MES





Évolution des Bêtas par secteur d'activité de 1997 à 2016. Ils sont calculés à partir de rentabilités quotidiennes avec des périodes d'estimation glissantes de 50 jours

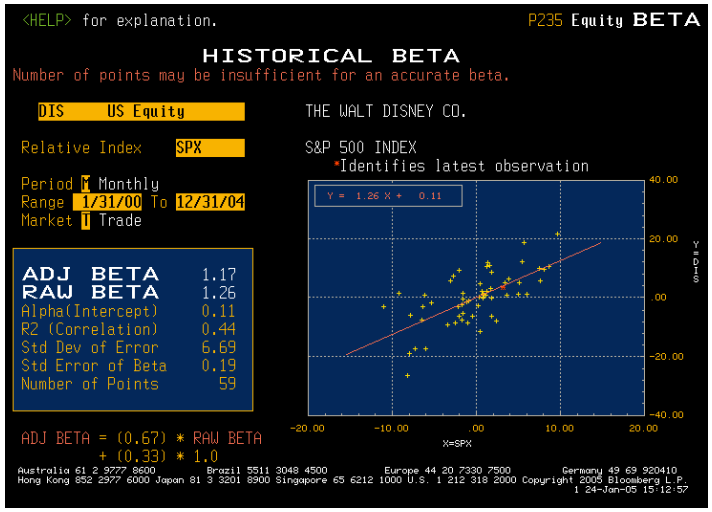
Il n'existe pas de secteurs caractérisés de manière évidente par la permanence de bêtas supérieurs ou inférieurs à 1. L'augmentation des Bêtas liés au secteur financier ou de l'informatique est conjoncturelle (bulle internet des années 2000, crise financière de 2008)



Etude de cas : utilisation de Bloomberg

- Bêta de l'action Disney par rapport à l'indice S&P500 (source Bloomberg)

Rentabilités mensuelles sur une période de 5 ans (31/1/00 – 31/12/04
Raw Beta = 1,26» correspond au Bêta historique
Adjusted Beta =1,17
2/3 raw bêta + 1/3 x 1



Grands et petits Bêtas …

- Adjusted Beta?
- Bloomberg s'appuie sur les travaux de Blume et de Vasicke
- Blume part de la constatation d'un retour à la moyenne des bêtas estimés sur des périodes consécutives

> A PREVIOUS STUDY [3] showed that estimated beta coefficients, at least in the context of a portfolio of a large number of securities, were relatively stationary over time. Nonetheless, there was a consistent tendency for a portfolio with either an extremely low or high estimated beta in one period to have a less extreme beta as estimated in the next period. In other words, estimated betas exhibited in that article a tendency to regress towards the grand mean of all betas, namely one. This study will examine in further detail this regression tendency.[1]

- Blume (1975). Betas and their regression tendencies. The Journal of Finance.
- Blume (1979). Betas and their regression tendencies: some further evidence. Journal of Finance

Grands et petits Bêtas …

- Le tableau ci-dessous considère des Bêtas estimés sur une période de 7 ans (1926-1933) et regroupés en 4 groupes
  - *Le groupe 1 est constitué des actions avec le Bêta le plus faible et ainsi de suite*
  - *La dernière colonne représente les Bêtas pour ces titres sur la période de 7 ans qui suit*

Portfolio	Grouping Period	First Subsequent Period
	7/26-6/33	7/33-6/40
1	0.50	0.61
2	0.85	0.96
3	1.15	1.24
4	1.53	1.42

  - *On constate un retour à la moyenne (sauf pour la catégorie 3)*

Grands et petits Bêtas …

- Retour à la moyenne en partie dû à une illusion statistique
  - *Pour l'illustrer, considérons des tirages indépendants d'un dé*
  - *Indépendance des valeurs observées pour 2 tirages consécutifs*
    - Pourtant, si le premier tirage est 6, le second tirage est plus petit
    - Si le premier tirage est 1, le second tirage est plus élevé.
  - *Le résultat du premier tirage n'apporte pourtant aucune information sur la valeur du second tirage.*
  - *Considérons maintenant un questionnaire d'évaluation administré à des étudiants.*
    - 100 questions et trois réponses proposées à chaque question
    - Les étudiants répondent « au hasard »
    - Le nombre de réponses justes suit une loi binomiale $B(100, 1/3)$. L'espérance du nombre de réponses justes est $\frac{100}{3}$

Grands et petits Bêtas …

  - *Les « meilleurs » étudiants à la première évaluation (notes supérieures à 100/3) auront des notes en moyenne égales à 100/3 à la seconde évaluation*
  - *Les « moins bons » étudiants vont voir leurs notes augmenter et se rapprocher de la moyenne.*
- Système cognitif automatique et recherche de causalité :
  - *Les meilleurs étudiants se sont relâchés entre la première évaluation et la seconde.*
  - *Les « moins bons » étudiants se sont accrochés et ont amélioré leur performance*
- Le système **déductif** permet de comprendre qu'il n'y aucun lien entre les résultats des deux évaluations
  - *Pour Kahneman, il y a **une explication mais pas de cause***
  - *Le système 1 est inadapté pour prendre en compte la chance*

Grands et petits Bêtas

- Daniel Kahneman, prix Nobel d'économie, père fondateur de la finance comportementale, psychologue cognitif et statisticien

"Our comforting conviction that the world makes sense rests on a secure foundation: our almost unlimited ability to ignore our ignorance."

Estimation des Bêtas / biais de sélection

- Revenons aux Bêtas ajustés et à l'étude de Blume
  - Le Bêta moyen étant de 1, la formule du Bêta ajustée: **2/3 bêta estimé sur la période précédente + 1/3 x 1**
  - Ceci pour prendre en compte le « retour à la moyenne » (ici = 1) pour les Bêtas
- *Blume avait conscience des problématiques statistiques*
  - On parle d'un problème d'erreur sur les variables, les Bêtas estimées étant entachées d'une erreur d'estimation.
  - Dans le cas du dé, la moyenne est constante à chaque tirage (3,5), le coefficient de corrélation entre deux tirages est nul
- *Le phénomène de retour à la moyenne subsiste même après les correctifs statistiques adaptés*
- *Explications économiques ?*

Estimation des Bêtas / erreur sur les variables

- *Supposons que l'on ait une liaison du type :*
- $\beta_{i,t} = a\beta_{i,t-1} + 1 - a + \varepsilon_{i,t}$ *(avec par exemple* $a = 2/3$
- *On cherche à vérifier si* $a < 1$ *(retour à la moyenne)*
- *Mais on ne connaît qu'un estimateur de* $\beta_{i,t-1}$*, le vrai Bêta n'est pas observé.*
- $\hat{\beta}_{i,t-1} = \beta_{i,t-1} + z_{i,t-1}$
  - $z_{i,t-1}$ erreur d'estimation indépendante de $\beta_{i,t-1}$
- $\beta_{i,t} + z_{i,t} = a \times (\beta_{i,t-1} + z_{i,t-1} - 1) + \varepsilon_{i,t}$
  - $\hat{a} = \frac{\mathrm{cov}(\beta_{i,t}+z_{i,t},\beta_{i,t-1}+z_{i,t-1})}{\mathrm{var}(\beta_{i,t-1}+z_{i,t-1})} = \frac{\mathrm{cov}(\beta_{i,t},\beta_{i,t-1})}{\mathrm{var}(\beta_{i,t-1})+\mathrm{var}(z_{i,t-1})} < a = \frac{\mathrm{cov}(\beta_{i,t},\beta_{i,t-1})}{\mathrm{var}(\beta_{i,t-1})}$
- *Le problème de Galton : on peut avoir* $a = 1$
  - Relation entre caractéristiques biologiques des parents et des enfants
- *et* $\hat{a} < 1$,
  - L'inné ne détermine pas tout

Estimation des bêtas / approche de Vasicek

O. Vasicek

- « Adjusted Beta » (Vasicek)
  - Vasicek (1973). A note on using cross-sectional information in Bayesian estimation of security Betas. *The Journal of Finance.*

    L'idée est différente de celle de Blume
  - *Vasicek propose une approche bayésienne*
  - *L'estimateur du bêta va être une moyenne pondérée du bêta estimé à partir de données historiques et d'un bêta déterminé a priori égal à un*
- Comme pour Blume, le choix des pondération 2/3 et 1/3 est arbitraire …
  - http://www.stat.ucla.edu/~nchristo/statistics_c183_c283/vasicek_betas.pdf
  - http://guides.lib.byu.edu/content.php?pid=53518&sid=401576
- Validité des Bloomberg "adjusted betas" ?