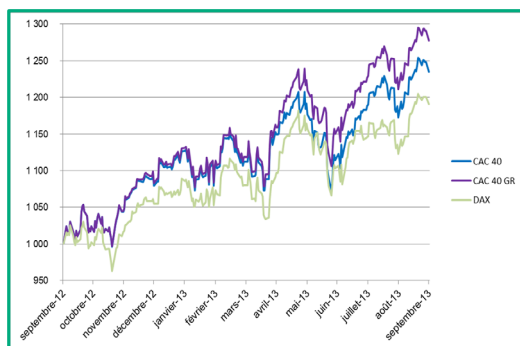


Négociation des actions et information



CAC 40 GR (gross return / dividendes réinvestis)

<https://indices.nyx.com/fr/products/indices/QS0011131834-XPAP/quotes>

1

Négociation et information

- Prix d'actions et calculs de rentabilité
- Données et marchés boursiers
 - Horodatage, jet lag
 - High Frequency Trading, relocalisation des ordres
 - Fragmentation du flux d'ordres
 - Régulation du HFT, flash krachs, opacité des ordres
- Stratégies boursières
- Arbitrage statistique, trading algorithmique
- Prédiction boursière : méthode des plus proches voisins
- Le problème de l'induction
- 2^e énigme de l'induction : encodage des informations financières

2

La forme standard d'une entreprise (rappel)

- Entreprise : personne morale, détenant un actif
- Bailleurs de fonds : actionnaires et créanciers
- Fonds apportés par les actionnaires : capital, par les créanciers : dette
- Capital + dette = passif de l'entreprise (ce qui est dû aux actionnaires et aux créanciers)
- Valeur des actifs = valeurs du passif
- Autres parties prenantes : salariés, clients, fournisseur
- Actions : titres de propriété : division du capital, cessibilité des titres sur des marchés organisés ou de gré à gré

3

Droits des actionnaires et des créanciers (rappel)

- Les deux droits attachés à une action
 - Droit de recevoir des dividendes et de revendre ses titres
 - = droit à recevoir des cashs-flows
 - Droit de participer aux votes en AG, de désigner le conseil d'administration
 - = droit de contrôler et gouverner l'entreprise
- Pour les créanciers : dette bancaire, obligations
 - Droit à recevoir coupons (intérêt) et principal
 - En cas de détresse financière, contrôle de l'entreprise,
 - Via une administration provisoire, éventuellement un liquidateur

4

Evolution des marchés d'actions

- Pour les actions des sociétés cotées en Bourse, les cours cotés sont une source importante d'information sur la valeur des actions émises par l'entreprise
 - *Approche externe à l'entreprise (via les marchés)*
 - *Meilleure approche de l'évaluation si efficacité informationnelle*
 - *Notamment pour les actifs incorporels et les actifs intangibles*
- Bourses : lieux de rencontre entre acheteurs et vendeurs
 - *Lieux physiques : marchés à la criée*
 - *Transformation en marchés électroniques*
 - *Marchés centralisés : carnet d'ordres*
 - *Passage du fixing aux transactions « en continu »*
 - *Diminution des temps de latence entre l'envoi d'un ordre et son exécution : transactions à haute fréquence*

5

Finance d'entreprise et de marché

- Finance d'entreprise
 - *Rentabilité moyenne, écarts-types des rentabilités supposés constants*
 - *Outils simples d'aide à la décision*
 - *Choix d'investissement : coût du capital, rentabilité cible*
 - *Évaluation d'entreprise dans le cadre de fusions-acquisitions*
- Finance de marché
 - *Rentabilités moyennes et risque fluctuent au cours du temps*
 - *Plus réaliste, mais demande plus de capacités de conceptualisation*
 - *Excès de complexité : modélisation financière mal maîtrisée*

6

7

8

Goodman on Induction

- Pour une action
- Sur une période : rentabilité simple, rentabilité logarithmique
- Pour un portefeuille
- cas multi-périodique : différents modes de calcul
 - Moyenne arithmétique, géométrique
- Composition des rentabilités
- Taux de rendement interne (TRI)

9

Rentabilité des titres

- Achat et vente de titres
 - Exemple
 - Plus value ou moins value
 - Rentabilité positive ou négative
- Vente à découvert
 - Emprunt d'un titre, vendu sur le marché, puis racheté
 - Rentabilité = opposé de la rentabilité de l'achat d'un titre
- Dividendes réinvestis : calcul de la rentabilité

10

La rentabilité d'un titre



- Titres : **actions (stocks)** et **obligations (bonds)**
- L'acheteur d'une action paye le prix d'achat P_0 à un vendeur à la date d'achat t_0
 - Cette transaction peut être réalisée sur une « Bourse » (exchange)
 - Un marché organisé où se rencontrent acheteurs et vendeurs comme Euronext Paris
 - Ou sur un marché de « gré à gré » (over-the-counter)
- Il peut revendre cette action à une date future t_1 au cours qui prévaut à ce moment sur le marché boursier, P_1
 - $P_1 - P_0$ représente la plus-value réalisée par l'investisseur

11

Taux de rentabilité



- À la date t_1 , juste avant la revente du titre, le détenteur d'une action peut percevoir un dividende versé par l'entreprise, soit D_1
 - Ce dividende est un revenu lié à la détention de l'action
 - Il est en général versé une fois par an
- Le montant encaissé par l'investisseur à la date t_1 est la somme du prix de vente et du dividende $P_1 + D_1$
 - Le gain net total lié à l'achat de l'action en t_0 et à sa revente en t_1 est $P_1 + D_1 - P_0$
 - Le **taux de rentabilité (rate of return)** lié à l'opération financière précédente est le gain net rapporté à l'investissement initial P_0
 - $r = \frac{P_1 + D_1 - P_0}{P_0}$

12

Taux de rentabilité



- Le **taux de rentabilité** $r = \frac{P_1 + D_1 - P_0}{P_0}$ correspond au taux d'accroissement de la richesse de l'actionnaire entre les dates t_0 et t_1
 - Pour l'opération financière consistant en l'achat d'une action en t_0 , revendue en t_1
 - Si l'écart entre t_1 et t_0 est d'une journée, on parle de taux de rentabilité quotidien, si l'écart est d'une semaine, on parle de taux de rentabilité hebdomadaire
 - On parle aussi de **taux de rentabilité simple**
 - C'est un pourcentage (quantité sans dimension)
 - Le prix d'une action ne pouvant être négatif, le taux de rentabilité est au pire de -100%
 - D_1/P_0 **taux de rendement (dividend yield)**

13

Non agrégation temporelle des rentabilités simples

- Pour simplifier l'analyse, le titre ne verse pas de dividende
- $r_{t,t+2} = \frac{P_{t+2} - P_t}{P_t} \neq (r_{t+1,t+2} + r_{t,t+1})$
 - Exemple numérique : $P_t = 100, P_{t+1} = 150, P_{t+2} = 100$
 - $r_{t,t+1} = \frac{150-100}{100} = 50\%$
 - $r_{t+1,t+2} = \frac{100-150}{150} = -33,3\%$
 - $r_{t,t+2} = 0\% \neq 50\% - 33,3\% = 16,6\%$
 - $r_{t,t+2} = \frac{P_{t+2} - P_t}{P_t} = \frac{P_{t+2} - P_{t+1}}{P_t} + \frac{P_{t+1} - P_t}{P_t}$
 - $r_{t+1,t+2} + r_{t,t+1} = \frac{P_{t+2} - P_{t+1}}{P_{t+1}} + \frac{P_{t+1} - P_t}{P_t}$
 - Conclusion $r_{t,t+2} = r_{t+1,t+2} + r_{t,t+1} \Leftrightarrow P_{t+1} = P_t$ ou $P_{t+2} = P_{t+1}$

14

Rentabilités simple et logarithmique

- y_t : rentabilité logarithmique
- $y_t = \ln\left(\frac{P_t + d_t}{P_{t-1}}\right) \approx \frac{P_t + d_t - P_{t-1}}{P_{t-1}} = r_{t-1,t}$
 - $P_{t-1} = 100, P_t + d_t = 104, r_{t-1,t} = 4\%, y_t = 3,92\%$
 - $P_{t-1} = 100, P_t + d_t = 30, r_{t-1,t} = -70\%, y_t = -120\%$
 - Rentabilités simples et logarithmiques proches pour des valeurs usuelles des rentabilités, mais gros écarts si on considère des cas extrêmes (krachs boursiers, stress tests)
- Rentabilité logarithmique (pas de dividende)
 - $y_t + y_{t+1} = \ln\left(\frac{P_t}{P_{t-1}}\right) + \ln\left(\frac{P_{t+1}}{P_t}\right) = \ln\left(\frac{P_{t+1}}{P_{t-1}}\right)$
 - Composition temporelles des rentabilités plus facile
 - Mais difficultés pour l'écriture des rentabilités des portefeuilles

15

16

Calculs de rentabilité : Attention aux « données »

- Beaucoup de bases de données de prix d'actifs financiers
- Pour les données quotidiennes, souvent cours de clôture
 - Et cours d'ouverture, plus haut, plus bas
- Auxquels les investisseurs ne peuvent pas traiter
- Tant que l'on n'a pas traité, un « prix » (committed quote, RFQ – Request For Quote) reste virtuel
- Qualité des données financières : enjeu important pour les professionnels :
 - Données intraday : utiles pour le THF (trading haute fréquence), le suivi de la volatilité
 - Données d'information financière, données sur l'état d'esprit des investisseurs
- Accès à la base de données Factset pour les mémoires en finance

17

L'accès aux données financières en temps réel ou aux données passées est stratégique



Bloomberg



Michael Bloomberg

18

Taux de rentabilité



EURONEXT



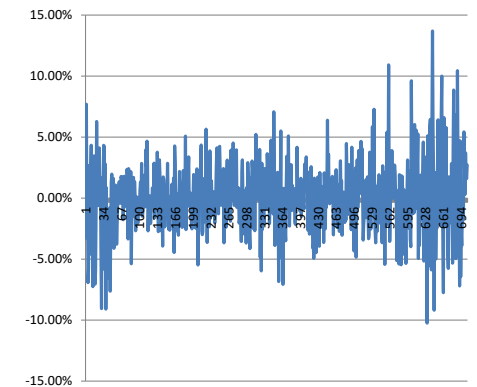
- Les rentabilités **quotidiennes** sont souvent calculées à partir des **cours de clôture**
 - Dernier cours auquel ont lieu des transactions pendant la journée
- Souvent sur **Euronext** pour les actions françaises
 - Anciennement la « Bourse de Paris »
 - Euronext est un marché réglementé par l'AMF, Autorité des Marchés Financiers :
 - Transparence
 - Publication du carnet d'ordres
 - Protection des petits investisseurs
 - Carnet d'ordres centralisé
 - Les gros ordres ne passent pas avant les petits



19

La rentabilité d'un titre

- Considérons l'action Peugeot
 - Code ISIN : FR0000121501
 - ISIN : International Securities Identification Number
 - Permet d'identifier un titre négociable sur un marché organisé
 - Environ 2 millions de titres
 - et de codes ISIN correspondants
 - Mnémo ou « ticker » : UG pour Peugeot
 - Le graphique de droite représente l'évolution des rentabilités quotidiennes



Historique des rentabilités quotidiennes Janvier 2009 – septembre 2011

20

La rentabilité d'un titre



DERNIÈRES TRANSACTIONS		
Heures	Cours	Qté
11:08:00	16.535	6
11:07:34	16.535	5
11:07:27	16.535	122
11:07:27	16.535	215
11:07:27	16.535	200
11:07:27	16.535	10
11:07:13	16.545	120
11:07:06	16.535	28
11:07:06	16.535	15
11:06:25	16.535	100

- Actions Peugeot négociables sur Euronext
- Historique des dernières transactions
 - Cours, quantités négociées
- Le carnet d'ordres indique les ordres d'achat et de vente à cours limité à un moment donné
 - Voir Cours et livre d'E. Le Saout

CARNET D'ORDRES					
Ordres	Qté	Achat	Vente	Qté	Ordres
3	697	16.590	16.600	509	1
4	1 756	16.585	16.610	1 873	2
1	464	16.580	16.615	1 065	3
6	1 325	16.575	16.620	5 352	7
2	756	16.570	16.625	1 285	5
5	4 164	16.565	16.630	2 125	3
3	1 186	16.560	16.635	3 527	2
2	600	16.555	16.650	482	1
2	800	16.550	16.655	4 475	1
4	3 954	16.545	16.660	400	1
32	15 702	TOTAL	TOTAL	21 093	26

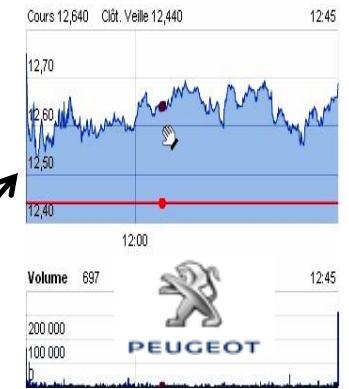
Le carnet d'ordres indique les ordres d'achat et de vente à cours limité à un moment donné

21

La rentabilité d'un titre

- Day-trading, exécution des ordres

- *Il n'y pas de raison pour que le prix d'achat dans la journée soit égal au cours de clôture*
- *Évolution des cours de l'action Peugeot le 19 septembre*
- *Pas de versement de dividende*
- *Rentabilité* $\frac{12,690 - 12,440}{12,440} = 2,01\%$
- *Mais si l'on avait revendu au plus haut, la rentabilité aurait été de* $\frac{12,760 - 12,440}{12,440} = 2,57\%$
 - Différence non négligeable



Variation	2.01%
Dernier échange	19/09/13 17:35:42
Ouverture	12.740
Clôture	12.690
+ Haut	12.760
+ Bas	12.525
Clôture veille	12.440

22

On ne peut pas forcément traiter au cours de clôture



- Ni même à aucun cours réalisé dans la journée ..
- Cours de clôture : dernier cours **traité** de la journée
- Quand on passe un ordre pour un montant, son exécution n'est pas garantie
 - Cela dépend de la nature de l'ordre (à cours limité, etc.), du carnet d'ordres et du mécanisme d'appariement.
 - « fixing » de clôture à 17h35 sur Euronext
 - Passer un ordre peut changer le cours de clôture
 - Beaucoup d'études...
 - Mesure de risques, évaluation de stratégies d'investissement, analyse de l'efficacité des marchés
 - font « comme si » on pouvait exécuter des ordres au cours de clôture.












23

Heures de clôture et marché

- Les marchés boursiers européens fonctionnent de manière (à peu près) synchrone, les cotations ont lieu de 9h à 17h30 (heure de Paris !).
- Il y a maintenant des « post trading hours »
- Mais décalages importants d'heures de clôture selon que l'on considère des bourses européennes, asiatiques ou américaines.
- Pas loin d'une demi-journée entre Paris et Tokyo.
- On ne peut pas calculer des risques sur des portefeuilles à partir de données non synchrones, sans correctifs statistiques appropriés

24

Dates de clôture (document non mis à jour)

Exchange Name 	Time Zone 	Opening Time (local) 	Opening Time (UTC) 	Closing Time (local) 	Closed for Lunch 
 Australian Securities Exchange (ASX)	AEST (UTC+10, DST)	10:00	00:00	16:10	No
 Vienna Stock Exchange	CET (UTC+1, DST)	8:55	07:55	17:35	No
 Euronext Amsterdam (AMS)	CET (UTC+1, DST)	9:00	08:00	17:40	No
 Euronext Paris (EPA)	CET (UTC+1, DST)	9:00	08:00	17:30	No
 Frankfurt Stock Exchange (FSX) / Xetra / Eurex	CET (UTC+1, DST)	8:00 (Eurex) 9:00 (floor) 9:00 (Xetra)	07:00	22:00 (Eurex) 20:00 (floor)	No

Prévisions boursières ?

- Objectif : peut-on « prévoir » aujourd'hui le prix d'une action (d'un titre négociable) demain ?
 - Notons P_u le prix à la date u .
 - On fait des opérations régulièrement espacées dans le temps : $u = 0, 1, 2, \dots, t-1, t, t+1, \dots$
 - Unité de temps : journalière ou plus haute fréquence
 - Exemple : toutes les minutes.
 - Le « pas de temps » est conventionnel, la constance de ce pas également (simplification de présentation)
 - Hausse notée H ou 1)
 - Baisse, (notée B ou 0)
 - $I_u = 0, 1$ variable indicatrice de hausse ou de baisse
 - $I_u = 1$ si $P_u - P_{u-1} \geq 0$, $I_u = 0$ si $P_u - P_{u-1} < 0$

29

Prévisions boursières ?

- Stratégies boursières : formalisation
 - n_{u-1} : nombre de titres achetés ($n_{u-1} \geq 0$) ou vendus ($n_{u-1} < 0$) à la date $u-1$
 - Gain à la date u : $n_{u-1} \times (P_u - P_{u-1})$
 - On ne prend pas en compte les versements de dividendes
 - Gain cumulé entre les dates 0 et t
 - $n_0 \times (P_1 - P_0) + n_1 \times (P_2 - P_1) + \dots + n_{t-1} \times (P_t - P_{t-1})$
 - Pour simplifier la présentation, on ne prend pas ici en compte les gains liés au réinvestissement des gains
 - Ou de manière équivalente, on suppose le taux d'intérêt nul
 - On se limite maintenant à soit l'achat d'une action, soit à la vente d'une action à une date u donnée : $n_u = \pm 1$
 - Gain maximal (prévision parfaite)
 - $|P_1 - P_0| + |P_2 - P_1| + \dots + |P_t - P_{t-1}|$: variation totale

30

Stratégies boursières

- On va s'intéresser à deux problèmes importants pour la finance et la gestion de portefeuilles
 - Peut-on prévoir les cours boursiers ?
 - Peut-on construire des stratégies de placement gagnantes de manière certaine ?
- Nous allons voir que ces deux problèmes sont liés
- Pour simplifier l'exposé, nous ferons les hypothèses fortes suivantes
 - Absence de coûts de transaction (fixes ou proportionnels)
 - **Absence d'impact des ordres d'achat ou de vente sur le prix de marché**
 - Prix par titre acheté ou vendu indépendant de la quantité
 - Ces hypothèses correspondent à celles d'un « marché parfait »

31

Stratégies boursières : formalisation

- P_t représente le prix d'un actif financier à la date t
 - $t = 0, 1, \dots, T$
 - Pour simplifier les notations, on suppose que l'on ne traite que d'un titre risqué et que les dividendes sont inclus dans le prix.
 - Il peut éventuellement s'agir d'un ETF dupliquant la rentabilité d'un indice boursier
- θ_t représente la quantité de titres détenues à la date t
 - On pourra acheter ou vendre des titres aux dates $t = 0, 1, \dots, T$
- V_t représente la valeur du portefeuille à la date t
- Il existe un placement sans risque au taux (constant) r .
 - L'hypothèse de constance du taux sans risque est faite pour simplifier les notations
 - On pourra éventuellement supposer $r = 0$

32

Stratégies boursières : formalisation

- Le portefeuille est donc composé d'une poche risquée dont la valeur à la date t est $\theta_t P_t$ et d'une poche investie en placement sans risque.
- **Hypothèse d'autofinancement** : l'évolution de la valeur du portefeuille ne résulte que des plus-values sur les titres et de la capitalisation au taux sans risque de la poche investie en placement sans risque
- $V_t = \theta_{t-1} P_t + (V_{t-1} - \theta_{t-1} P_{t-1}) \times (1 + r)$
 - $\theta_{t-1} P_t$ est la valeur en t , des θ_{t-1} titres acquis en $t - 1$, pour un montant $\theta_{t-1} P_{t-1}$
 - $V_{t-1} - \theta_{t-1} P_{t-1}$ est le montant à la date $t - 1$ de la poche investie en placement sans risque

33

Stratégies boursières : formalisation

- $V_t = \theta_{t-1} P_t + (V_{t-1} - \theta_{t-1} P_{t-1}) \times (1 + r)$
 - Cas $r = 0$; l'équation précédente devient
 - $V_t = \theta_{t-1} P_t + V_{t-1} - \theta_{t-1} P_{t-1}$
- $V_t - V_{t-1} = \theta_{t-1} (P_t - P_{t-1})$
 - Plus ou moins-values sur les titres \Rightarrow Variations de richesse
 - Cas général $r \neq 0$
 - On définit $\tilde{V}_t = \frac{V_t}{(1+r)^t}$, $\tilde{P}_t = \frac{P_t}{(1+r)^t}$: richesse et prix actualisés
 - L'équation donnant V_t s'écrit comme
 - $V_t - (1+r)V_{t-1} = \theta_{t-1} (P_t - (1+r)P_{t-1})$
 - Soit $\tilde{V}_t - \tilde{V}_{t-1} = \theta_{t-1} \times (\tilde{P}_t - \tilde{P}_{t-1})$
- $\tilde{V}_T = \tilde{V}_0 + (\tilde{V}_1 - \tilde{V}_0) + \dots + (\tilde{V}_T - \tilde{V}_{T-1}) = V_0 + \sum_{t=0}^{T-1} \theta_{t-1} \times (\tilde{P}_t - \tilde{P}_{t-1})$

34

Stratégies boursières : Remarques

- $\tilde{P}_t - \tilde{P}_{t-1} > 0 \Leftrightarrow P_t - (1+r)P_{t-1} > 0 \Leftrightarrow \frac{P_t - P_{t-1}}{P_{t-1}} > r$
- De même, $\tilde{V}_t - \tilde{V}_{t-1} > 0 \Leftrightarrow \frac{V_t - V_{t-1}}{V_{t-1}} > r$
- Comme $\tilde{V}_t - \tilde{V}_{t-1} = \theta_{t-1} \times (\tilde{P}_t - \tilde{P}_{t-1})$
- Si $\theta_{t-1} > 0$, $\tilde{V}_t - \tilde{V}_{t-1} > 0$ si et seulement si $\tilde{P}_t > \tilde{P}_{t-1}$
- On vient d'écrire
 - Un achat de titres ($\theta_{t-1} > 0$) est associé à une rentabilité du portefeuille supérieure au taux sans risque, si et seulement si la rentabilité du titre est supérieure au taux sans risque
 - Ce qui est conforme à l'intuition financière
- Dans la suite, on supposera $r = 0$, d'où $\tilde{V}_t = V_t$, $\tilde{P}_t = P_t$

35

Stratégies boursières : exemples numériques

- Deux dates ($T = 2$), $V_0 = 100$, $r = 0$
 - $P_0 = 100$, $P_1 = 110$, $P_2 = 80$
- $\theta_0 = \theta_1 = 1$
 - $V_2 = V_0 + (V_1 - V_0) + (V_2 - V_1) = 100 + 10 - 30 = 80$
- Si $\theta_0 = \theta_1$, stratégie buy and hold
- Si $\text{signe}(\theta_1) = \text{signe}(P_1 - P_0)$, stratégie momentum
 - Si hausse des cours à la date précédente, achat, si baisse, vente
- Si $\text{signe}(\theta_1) = -\text{signe}(P_1 - P_0)$, stratégie contrarian
- $\theta_0 = 1$, $\theta_1 = -1$
 - $V_2 = V_0 + (V_1 - V_0) - (V_2 - V_1) = 100 + 10 + 30 = 140$

36

Stratégies boursières gagnantes

- On rappelle que par simplification $r = 0$
- Pour qu'une stratégie boursière soit gagnante à toute date, il faut :
 - Acheter des titres quand le cours augmente à la date suivante
 - Vendre des titres quand le cours baisse à la date suivante
 - Ceci correspond à un « market timing » parfait
- Il faut donc déterminer à la date $t - 1$ le signe de $P_t - P_{t-1}$
 - Il faut établir une prévision à la date $t - 1$ sur le prix futur P_t
 - Quelles informations peuvent-elles être utilisées pour prévoir ?
 - Les prix passés du titres (ou de l'ensemble des titres)
 - D'autres informations financières et économiques publiques
 - Des informations connues seulement d'insiders ou d'agents disposant d'un avantage informationnel (informations dites privées)

37

Stratégies boursières gagnantes

- **Opportunité d'arbitrage : c'est une stratégie autofinancée, telle que $V_0 = 0$ et $V_T > 0$**
 - Il existe des définitions légèrement différentes, mais l'idée reste que cela correspond à gagner de l'argent de manière certaine à partir de rien. On parle aussi de « free lunch ».
 - Si l'on peut déterminer dès aujourd'hui si le prix du titre va augmenter ou baisser demain ...
 - C'est-à-dire si l'on peut prévoir le signe de $P_t - P_{t-1}$ à la date $t - 1$
 - ... On peut construire une opportunité d'arbitrage : acheter si le prix monte, vendre si le prix baisse.
- Remarque : sous les hypothèses précédentes (marchés sans frictions) et si les quantités achetées ou vendues ne sont pas limitées, les gains associés à une opportunité d'arbitrage ne sont pas limités non plus.

38

Stratégies boursières gagnantes

- D'où l'intérêt de la prévision des cours de Bourse pour le trading et l'investissement en titres financiers.
- Cela dépend de l'**information** dont dispose le spéculateur.
- S'il connaît un **devin** qui lui souffle à l'oreille aujourd'hui (et à lui seul) si le prix de l'action demain sera supérieur ou inférieur à celui aujourd'hui, c'est parfait.
- Mais la possibilité de la prémisse semble problématique pour un esprit rationnel.

39

Stratégies boursières gagnantes

- Voici un exemple concret de « divinisation »
- Supposons qu'il y ait deux places de marché, distantes géographiquement, disons New-York et Amsterdam.
- Où l'on traite du même actif financier
- Les contraintes de la physique font qu'on ne peut connaître au même moment les prix traités sur chacun des marchés.
- Supposons qu'un fonds dispose d'une technologie de transmission privée et plus rapide que celle de ses concurrents.
- Elle sera informée pendant un laps de temps, même faible des prix sur chacun des marchés

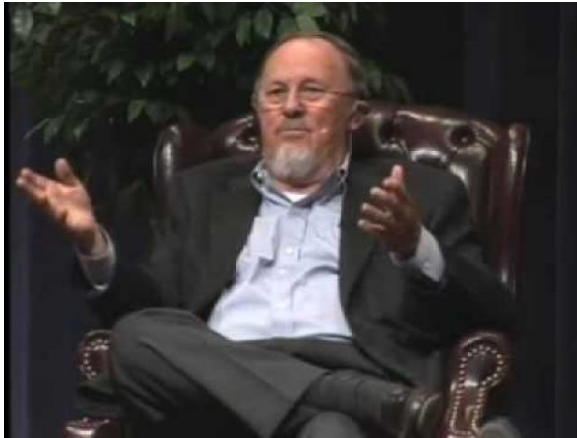
40

Stratégies boursières gagnantes

- Un **avantage informationnel** peut induire des **opportunités d'arbitrage**.
- Il faut néanmoins prendre en compte l'**impact des ordres sur le marché**
 - *Il faut trouver des contreparties prêtes à exécuter les ordres : fournisseurs de liquidité plaçant des ordres à cours limités ou market-makers (teneurs de marché)*
 - *Il faut que ces market-markers ne puissent pas inférer du montant des ordres ou de la nature du donneur d'ordre qu'il dispose d'une **information privilégiée***
 - *Car dans ce cas, les teneurs de marché vont réviser les prix proposés et/ou restreindre les quantités offertes*

La notion d'efficience informationnelle

- There are no shortcuts in investing
 - <http://www.youtube.com/watch?v=pGIzygsvqck>
 - *Fonds indiciels, stratégies systématiques (risk parity)*



William Sharpe
Stanford University
1h35mn

45

La notion d'efficience informationnelle

- Deux types d'informations
- Informations publiques, données de cours boursiers, documents financiers, conférence call auprès des investisseurs



- Informations transmises par les ordres d'achat et de vente d'actions

46

La notion d'efficience informationnelle

- Pour la théorie des « marchés informationnellement efficients » (voir transparent suivant), ce sont les informations financières (exogènes au marché)
- Qui gouvernent les variations des prix des actions
- Pour un courant marginal, mais significatif, un examen attentif de la microstructure des marchés conclut à une dynamique endogène des prix
 - *Par endogène, on entend mécanismes de rencontre des ordres d'achat et de vente dans des marchés où la liquidité est limitée*
 - *Où des manipulations des ordres et de l'information transmise par les ordres impactent les prix de marché*
- Où placer le curseur ?

47

La notion d'efficience informationnelle

SESSION TOPIC: STOCK MARKET PRICE BEHAVIOR

SESSION CHAIRMAN: BURTON G. MALKIEL

EFFICIENT CAPITAL MARKETS: A REVIEW OF THEORY AND EMPIRICAL WORK*

EUGENE F. FAMA**

Market Efficiency (Fama 1)

- A market in which prices always "fully reflect" all available information is called "efficient." (383-384)
- The definitional statement that in an efficient market prices "fully reflect" available information is so general that it has no empirically testable implications.

Fama, Eugene F. "Efficient Capital Markets: A Review of Theory and Empirical Work," J. of Finance, V. 25, May 1970



Eugene Fama

48



La notion d'efficience informationnelle

Handbook of Financial Markets: Dynamics and Evolution

Handbooks in Finance

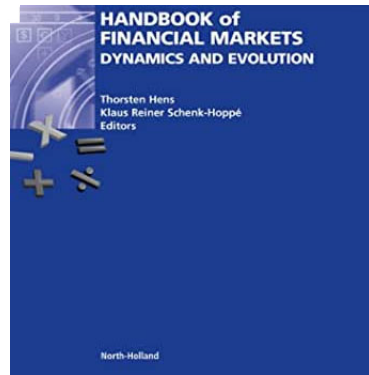
2009, Pages 57-160

CHAPTER 2 - How Markets Slowly Digest Changes in Supply and Demand

Jean-Philippe Bouchaud, J. Doyne Farmer, Fabrizio Lillo

Most of the processed information appears to come from supply and demand itself, rather than from external news.

J-P. Bouchaud



49

La notion d'efficience informationnelle

- Qu'est-ce qui est à l'origine des changements de prix des actions ?
- Informations nouvelles sur la performance financière d'une entreprise
 - *Associées à une anticipation de hausse des cash-flows (dividendes)*
 - *Valeur fondamentale : valeur actualisée des cash-flows futurs anticipés*
 - *Information positive → hausse des prix*
 - *Information négative → baisse des prix*
 - *Informations nouvelles relatives au risque d'une entreprise*
 - *Augmentation du risque → baisse du prix des actions*

50

La notion d'efficience informationnelle

- Qu'est-ce qui est à l'origine des changements de prix des actions ?
- Marchés informationnellement efficients
- Informations correctement traitées (processed) par les investisseurs ?
 - *Rôle des analystes financiers*
 - *Biais cognitifs et comportementaux*
- Les prix de marché correspondent aux valeurs fondamentales (économiques)
- Les prix de marché suivent des « marches aléatoires »
 - *Pas de possibilité de « free lunch » : « repas gratuits »*

51

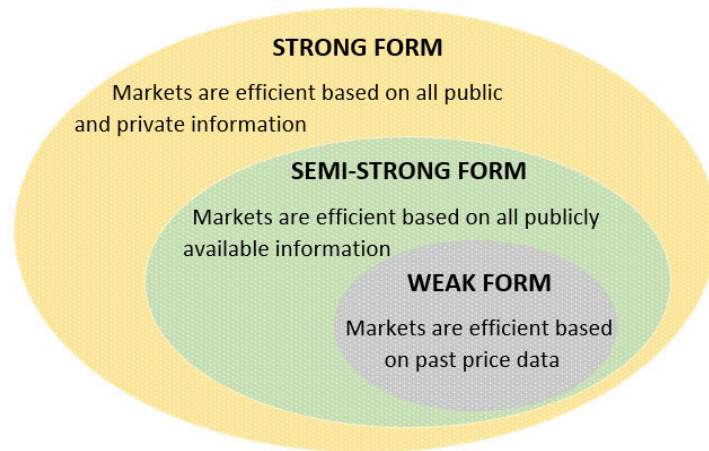
La notion d'efficience informationnelle

- Fama a distingué trois ensembles d'information :
 - *Le prix courant et l'ensemble des prix passés*
 - Ceci est lié à la notion d'efficience faible
 - Pour Fama, il s'agissait de critiquer l'analyse technique
 - Sa critique se transpose aujourd'hui aux techniques de trading algorithmiques inductives ou « data driven »
 - *Toutes les informations publiques*
 - informations comptables ou qualitatives, réseaux sociaux, Bloomberg, commentaires d'analystes financiers, ...
 - Efficience semi-forte
 - *Toutes les informations, y compris privées (initiés)*
 - Efficience forte
- **Concept d'efficience de marché : pas de stratégie spéculative gagnante et sans risque**

52

La notion d'efficience informationnelle

- Les trois ensembles d'information considérés par Fama



53

La notion d'efficience informationnelle

- **Pour Fama**, le marché est efficace sur le plan informationnel si les prix des actions reflètent de manière rationnelle toute l'information disponible.
 - *Il peut s'agir de l'information faible, semi-forte ou forte*
- Il en résulte que le prix de marché est la meilleure évaluation possible d'une action :
 - *Valeur fondamentale et prix de marché doivent coïncider*
 - *Il n'est pas possible pour les analystes financiers de déterminer des actions sous ou surévaluées*
- Seules les informations nouvelles, non anticipées sont susceptibles d'entraîner des variations de prix
 - *Les historiques de prix doivent présenter des caractéristiques des suites (ou marches) aléatoires*

54

La notion d'efficience informationnelle

- Il résulte du principe d'efficience informationnelle qu'il n'est pas possible de « battre le marché »
 - *Par des stratégies de stock-picking : achat de titres sous-évalués, vente de titres surévalués (rôle des analystes financiers)*
 - *De market timing, achat quand le marché est trop bas, vente quand il est trop élevé (stratégies global macro des hedge funds)*
 - *Par des stratégies de trading algorithmique (analyse de motifs)*
- Recherche d'« anomalies de marché » ?
 - *Effet du lundi, des petites capitalisations boursières*
 - *Mais il s'est avéré que les méthodologies statistiques étaient le plus souvent erronées*
 - Data snooping : optimisation de paramètres des stratégies dans l'échantillon
 - Biais de sélection : seules les études présentant des anomalies sont publiées.

55

La notion d'efficience informationnelle

- Qu'est-ce qui est à l'origine des changements de prix des actions ?
- Microstructure des marchés – Market impact
 - *Ordre d'achat d'important entraîne mécaniquement une augmentation des cours par "vidange" du carnet d'ordres*
 - Et vice-versa
 - Liquidité limitée : searching costs, limites de risque, information transmise par les ordres
 - *Market impact : Manipulations boursières*
 - Volonté d'un acteur de modifier les prix,
 - par une action directe d'achat ou de vente
 - et/ou par la manipulation d'informations relatives à la santé financière de la société cotée;

56

La notion d'efficacité informationnelle

- Qu'est-ce qui est à l'origine des changements de prix des actions ?
- Microstructure des marchés – Market impact
 - *Ordre d'achat d'important entraîne mécaniquement une augmentation des cours par "vidange" du carnet d'ordres*
 - Et vice-versa
 - Liquidité limitée : searching costs, limites de risque, information transmise par les ordres
 - *Market impact : Manipulations boursières*
 - Volonté d'un acteur de modifier les prix
 - Par une action directe d'achat ou de vente
 - Et/ou par la manipulation d'informations relatives à la santé financière de la société cotée

57

Efficacité informationnelle et mécanismes de négociation

- Mimétisme de comportement
 - *Un investisseur qui cherche à suivre les mouvements à l'achat ou à la vente d'un autre investisseur informé*
 - *Passager clandestin : LTCM, Deutsche Bank (enchères de bunds)*
 - *L'informé cherche à rester anonyme*
- Qui suivre ?
 - *Toute transaction sur un marché est la rencontre d'un acheteur et d'un vendeur*
- Le marché est toujours à l'équilibre ...
 - *Toujours autant d'acheteurs que de vendeurs*
 - *Fisher Black : une TTF (taxe sur les transactions financières) n'entraînerait pas une baisse des actions, car elle pénalise tout autant les vendeurs que les acheteurs.*

58

Efficacité informationnelle et mécanismes de négociation

- Rothschild et la bataille de Waterloo
- La bourse de Londres à l'arrêt
- Le système d'information des Rothschild
 - « *High Frequency Trading* » !
- La transmission d'information par les ordres
 - *Une erreur dans le raisonnement récursif*
 - *Keynes et le concours de beauté*
 - *Initié et manipulation boursière*



THE HOUSE OF ROTHSCHILD Money's Prophets, 1798-1848

- Cette histoire de manipulation est romancée

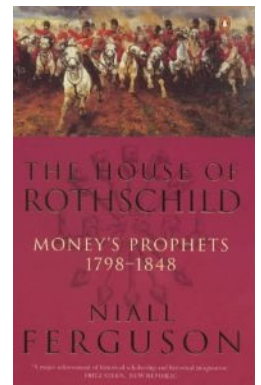
- *Et non dénuée d'arrière-pensées ...*

*Il jouait à la baisse, et montait à mesure
Que notre chute était plus profonde et plus sûre ; (...)*

*Un million joyeux sortit de Waterloo ;
Si bien que du désastre il a fait sa victoire.*

Victor Hugo, Contemplations, Melancholia, 1838

- *Montant allégué des gains non compatible avec la taille du marché des obligations à l'époque*
- *Selon l'historien N. Ferguson, N. Rothschild n'aurait cherché qu'à se protéger contre la baisse d'un important stock d'or*
 - Motivation de couverture du risque



60

Efficienc e informationnelle et mécanismes de négociation

Qui ne sait dissimuler
ne sait pas régner.



Louis XI

61

Efficienc e informationnelle et paris

On dit quelquefois, avec une évidente exagération que, sur deux parieurs, il y a toujours un voleur et un imbécile. On entend par là que l'un des deux parieurs est mieux renseigné que l'autre et sait presque sûrement si l'objet du pari se réalisera ou non. En fait, il arrive souvent que chacun des deux parieurs a des raisons de croire que c'est lui le mieux renseigné, de sorte que chacun d'eux pense qu'il fait un pari avantageux, ce qui ne veut pas dire qu'il mérite d'être traité de voleur s'il a raison ni d'imbécile s'il se trompe.

Emile Borel : Probabilités et Certitude

62

Efficienc e informationnelle et mécanismes de négociation

- Dans l'exemple précédent, les investisseurs ont réagi de manière « stupide ».
- Ils auraient dû anticiper que la stratégie optimale aurait été d'acheter et non de vendre
 - *Ce qui aurait annulé les gains des Rothschild*
 - *En revanche, les Rothschild avaient bien compris à qui ils avaient à faire.*
 - *S'ils avaient pensé que les autres investisseurs auraient anticipé leur manipulation, les Rothschild aurait dû acheter...*
 - *A quel niveau, faut-il arrêter le raisonnement récursif?*
 - Il faut être capable d'aller un cran plus loin que l'adversaire, qui cherche à faire la même chose que vous.

63

Évaluation des actions : le concours de beauté keynésien



Mécanismes de négociation : raisonnement récursif de second ordre

« Je sais que tu crois que je pense à toi » exprime une pensée (en partie contradictoire) que seuls les humains peuvent concevoir. L'anthropologue Robin Dunbar a parfaitement résumé ce qui est en jeu³. L'intentionnalité dite du premier ordre se définit comme la capacité de réfléchir au contenu de son propre esprit, comme en témoigne l'utilisation des verbes supposer, penser, s'interroger, croire, etc. La plupart des mammifères et des oiseaux entrent probablement dans cette catégorie. Plus intéressants sont les cas dans lesquels l'individu est capable de se représenter l'état mental de quelqu'un d'autre, de dire : « Je sais que vous aimez l'abricot. » Cette capacité définit un niveau plus élevé d'intentionnalité, conventionnellement appelé du second ordre. C'est l'équivalent du stade que les enfants atteignent vers l'âge de 6 ans lorsqu'ils acquièrent pour la première fois ce que les spécialistes de sciences cognitives appellent la « théorie de l'esprit ». Ils comprennent que les autres personnes peuvent avoir des idées différentes des leurs.

65

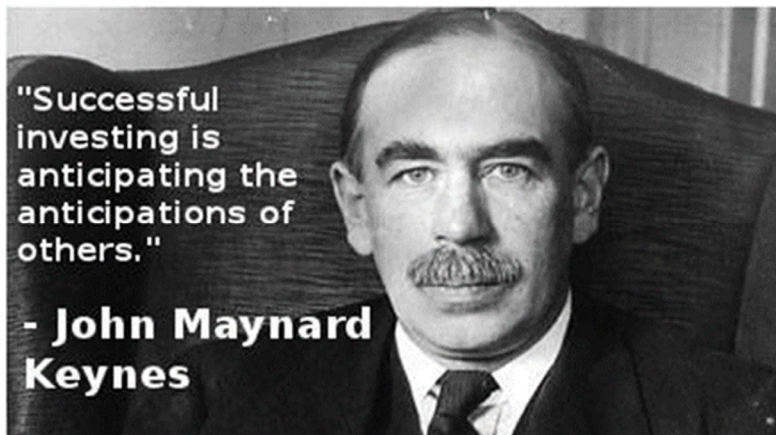
Raisonnement récursif d'ordre supérieur : jeux de miroir avec la pensée d'autrui

« Je sais que tu crois que je pense à toi » caractérise une intentionnalité d'ordre trois. Jusqu'où peut-on aller comme ça ? L'économiste George

Dunbar défend l'idée que les humains peuvent aspirer à une intentionnalité d'ordre cinq. Le cinquième ordre équivaut à pouvoir dire : je suppose [1] que vous croyez [2] que je veux [3] que vous pensiez [4] que j'ai l'intention de vous menacer... [5].

Le génie de Shakespeare nous fait atteindre ces sommets. Dans *Othello*, Shakespeare utilise quatre états d'esprit : Iago veut qu'Othello croie que Desdémone aime Cassio et que ce dernier aime celle-ci. Mais Shakespeare lui-même doit persuader le public de croire en tout cela. Et de plus, ce qui n'est pas la moindre des choses, il doit imaginer tout lui-même, il doit être capable de travailler – au minimum – avec une intentionnalité du sixième ordre : il veut que le public comprenne que Iago veut qu'Othello etc. Seul un humain (pas n'importe lequel) est capable d'un tel exploit.

66



67

Beauty contest selon Keynes

Chaque concurrent doit donc choisir non les visages qu'il juge lui-même les plus jolis, mais ceux qu'il estime les plus propres à obtenir le suffrage des autres concurrents, lesquels examinent tous le problème sous le même angle.

Il ne s'agit pas pour chacun de choisir les visages qui, autant qu'il en peut juger, sont réellement les plus jolis ni même ceux que l'opinion moyenne considérera réellement comme tels.

Au troisième degré où nous sommes déjà rendus, on emploie ses facultés à découvrir l'idée que l'opinion moyenne se fera à l'avance de son propre jugement.

Et il y a des personnes, croyons-nous, qui vont jusqu'au quatrième ou au cinquième degré ou plus loin encore.

68

Efficiency informationnelle et mécanismes de négociation

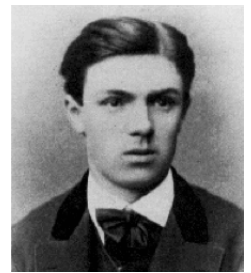


■ Conformisme en matière d'opinion

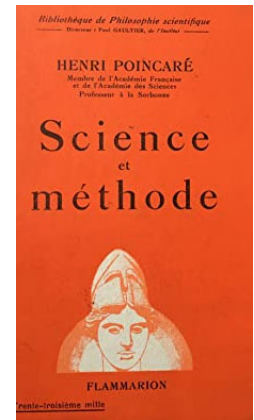
- Mimétisme, « herding », instinct grégaire (moutons de Panurge)
- Si tous les investisseurs s'accordent sur le niveau de la prime de risque ou l'évolution des bénéfices futurs
- « Bulle internet » des années 2000

Information, mimétisme, instinct grégaire

- Sur le mimétisme et les interactions humaines
 - *Quand des hommes sont rapprochés, ils ne se décident plus au hasard et indépendamment les uns des autres ; ils réagissent les uns sur les autres.*
 - *Des causes multiples entrent en action, elles troublent les hommes, les entraînent à droite et à gauche, mais il y a une chose qu'elles ne peuvent détruire, ce sont leurs habitudes de moutons de Panurge.*
 - *Et c'est cela qui se conserve.*
 - Henri Poincaré dans Science et Méthode



Henri Poincaré



Just a normal day at the nations most important financial institution...



Efficiency informationnelle et mécanismes de négociation

- Admati & Hellwig (2014). The bankers' new clothes.
- Le titre de cet ouvrage consacré à la crise financière de 2008 est issu du conte d'Andersen, “les habits neufs de l'empereur”
- Deux tailleurs proposent à l'empereur un habit si merveilleux qu'il ne pouvait être vu des personnes stupides ou incompetentes.
- En fait, le roi est nu, mais personne ne peut le dire, *ni même le voir*, par peur pour sa réputation...
- Illusions collectives systématiques étudiées en psychologie sociale (conformisme)

THE BANKERS' NEW CLOTHES



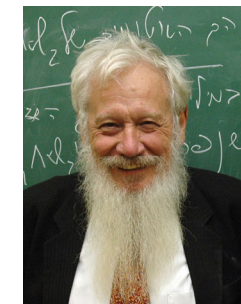
What's Wrong with Banking and What to Do about It
With a new preface by the authors
ANAT ADMATI & MARTIN HELLWIG

Efficiency informationnelle et mécanismes de négociation

- Vous pensez que le cours d'une action est sur-évalué
 - Vente ? Attention à l'excès de confiance en soi (overconfidence). Que vaut votre appréciation ?
 - Et même si vous aviez raison, vous pourriez être perdant
- Ce qui compte n'est pas ce que vous pensez, ce que pensent les autres et l'évolution de leurs croyances
 - Le marché est autoréférentiel
 - Supposons par exemple que vous pensez, disons à juste titre, que le niveau du CAC est trop élevé
 - Vous vendez à découvert
 - Si le marché continue à monter, vous devrez payer des appels de marge et quand vous serez à court de liquidité, vous devrez racheter votre position, alimentant la hausse...

73

Efficiency informationnelle et mécanismes de négociation



R. Aumann
Théorie des jeux

- Agreeing to disagree? (Aumann, 1976)
 - Le processus de négociation révèle de l'information.
 - Megabank dispose d'information privée (non connue d'Alphaplus) sur la valeur de AAA et que cette asymétrie d'information est connue
 - Supposons que Megabank dise au fonds de placement Alphaplus « je suis prêt à vendre un actif financier AAA à 100 ».
 - Informations transmises :
 - Alphaplus sait que Megabank veut vendre (le vrai prix est plus bas)
 - Megabank sait qu'Alphaplus sait que Megabank veut vendre (Megabank devrait proposer un prix plus élevé que le vrai prix)
 - Donc, Alphaplus sait que Megabank sait qu'Alphaplus sait que Megabank veut vendre (Alphaplus s'attend à ce que Megabank propose un prix élevé)
 - Et ainsi de suite...

74

Efficiency informationnelle et mécanismes de négociation

- Aumann a formalisé ce concept de “common knowledge”
 - « Nous savons qu'ils mentent, ils savent aussi qu'ils mentent, ils savent que nous savons qu'ils mentent, nous savons aussi qu'ils savent que nous savons qu'ils mentent et pourtant ils persistent à mentir ».
 - Alexandre Soljenitsyne



75

Efficiency informationnelle et mécanismes de négociation

- Aumann's agreement theorem
 - Si Alphaplus et Megabank disposent de la même information publique et que Megabank dispose d'une information privée et qu'il s'agit d'une connaissance commune, alors ils aboutiront à la même évaluation.
 - On peut appréhender ce résultat par une négociation séquentielle où Megabank est amenée à réviser son prix à la baisse jusqu'au **niveau du prix en information complète**.
 - Ce prix est celui au-delà duquel, elle refuse de vendre
- Pour Milgrom et Stokey (1976), cela implique **qu'il n'y a plus de transactions** dans un marché efficient, **même quand des agents acquièrent de l'information privée**.
 - Milgrom & Stokey (1982). Information, trade and common knowledge. *Journal of economic theory*.

76

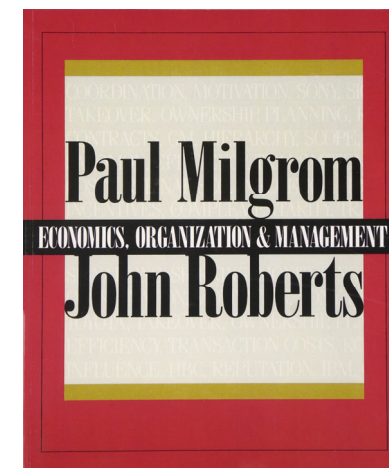
Efficiency informationnelle et mécanismes de négociation

- On peut appréhender ce résultat par une négociation séquentielle où Megabank est amenée à réviser son prix à la baisse jusqu'au **niveau du prix en information complète**.
- Ce prix est celui au-delà duquel, elle refuse de vendre
- Supposons que la « vraie valeur », compte-tenu de l'information privée de Megabank soit 90.
- Tant que Megabank propose de vendre à un prix supérieur à 90, Alphaplus sait que Megabank fait un profit, refuse l'offre et attend une enchère plus basse.
- Quand on atteint le prix de 90, Megabank n'a plus intérêt à vendre.
- La négociation a révélée l'information privée de Megabank.

77



Paul Milgrom, Prix Nobel d'économie (2020) pour notamment ses travaux sur les enchères: le trading peut être vu comme un processus d'enchères bilatérales



Des modèles micro-économiques pour la gestion des entreprises

78

Efficiency informationnelle et mécanismes de négociation

Buyer: How much is it?

Seller: £1.50.

Buyer: OK, I'll take it.

Seller: It's £1.60.

Buyer: What? You just said £1.50.

Seller: That was before I knew you wanted it.

Buyer: You cannot do that!

Seller: It's my stuff.

Buyer: But I need a hundred of those!

Seller: A hundred? It's £1.70 apiece.

Buyer: This is insane!

Seller: It's the law of supply and demand, buddy. You want it or not?

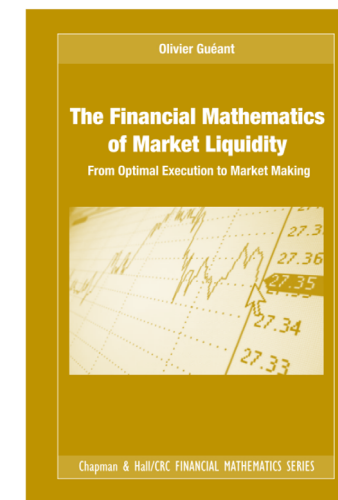
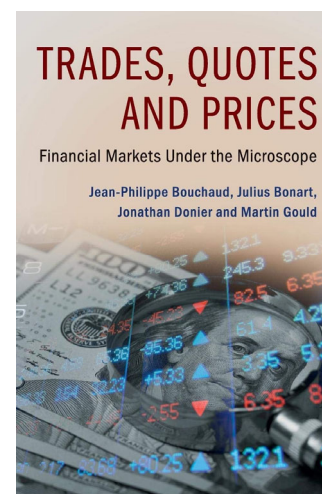
(Translated from "6", by Alexandre Laumonier)

Illustration de l'information transmise par les ordres, tirée du livre de Bouchaud et al

79

Efficiency informationnelle et mécanismes de négociation

- Exécution optimale des ordres, microstructure des marchés



80

Efficiences informationnelles et mécanismes de négociation

■ Le paradoxe de Grossman et Stiglitz

- Investisseurs rationnels et initiés
- Autres investisseurs : « passagers clandestins », bénéficiaires de l'information privée produite par les initiés
- Les initiés ne peuvent bénéficier de leur information privée
- Personne n'accepte de payer pour l'analyse financière
- Les marchés ne peuvent donc être efficaces
- Noise traders (idiots) sont indispensables...
- Car ils permettent aux initiés de se cacher
 - Grossman & Stiglitz (1980). On the impossibility of informationally efficient markets. *The American economic review*



S. Grossman
Anticipations rationnelles

81

Stratégies optimales et aléatoires

- Le guépard court plus vite que la gazelle de Thomson
- Mais celle-ci peut néanmoins s'en sortir



82



Alain Pavé
Vers une théorie synthétique de la biodiversité



83

Stratégies optimales et aléatoires

- Chiappori, Levitt, & Grogg (2002). Testing mixed-strategy equilibria when players are heterogeneous: The case of penalty kicks in soccer. *American Economic Review*.
- Les tireurs de penalty sont plus performants d'un côté, en général quand ils tirent à gauche
- Les gardiens de but le savent et les tireurs de penalties savent qu'ils savent



P-A. Chiappori

84

Stratégies optimales et aléatoire

- Voici les statistiques de succès du tireur, selon qu'il tire à gauche ou à droite et que le gardien plonge ou pas du bon côté

TABLE 1—OBSERVED SCORING PROBABILITIES, BY FOOT AND SIDE

Kicker	Goalie	
	Correct side	Middle or wrong side
Natural side ("left")	63.6 percent	94.4 percent
Opposite side ("right")	43.7 percent	89.3 percent

85

Stratégies optimales et aléatoire

- Quelle est la meilleure stratégie pour le tireur ?

TABLE 1—OBSERVED SCORING PROBABILITIES, BY FOOT AND SIDE

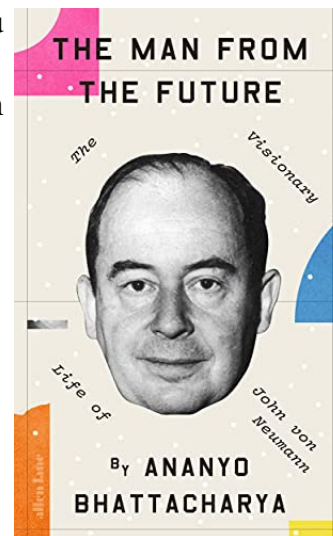
Kicker	Goalie	
	Correct side	Middle or wrong side
Natural side ("left")	63.6 percent	94.4 percent
Opposite side ("right")	43.7 percent	89.3 percent

- S'il tire toujours de son meilleur côté, il devient trop prévisible pour le gardien
- Meilleure stratégie : tirer de manière aléatoire à gauche ou à droite avec une probabilité un peu plus élevée du bon côté
- Stratégie dite mixte \neq stratégie pure (déterministe)

86

Stratégies optimales et aléatoire

- Le gardien sachant la stratégie optimale du tireur va lui aussi plonger d'un côté aléatoire, un peu plus fréquemment du bon côté du tireur
- La stratégie du tireur est également optimale sachant la stratégie du gardien
- Equilibre de Nash en stratégies mixtes
- La notion de stratégie mixte et son optimalité est due à Von Neumann (article de 1928) – minimax theorem
- L'article de Chiappori et al valident les principes de l'optimalité des stratégies aléatoires



87

Stratégies optimales et aléatoire

- Existence d'un équilibre de Nash en stratégies mixtes
 - Von Neumann



- Extrait du livre « Le hasard: Une approche mathématique » d'Ivar Ekeland

88

Efficiencia informacionnelle et mecanismes de negociation

- « Inventory model » : un market maker qui est en position longue va chercher à attirer des ordres d'achat
- Pour réduire son stock de titres et donc son risque
- Toutes choses égales par ailleurs, il devrait donc diminuer le « cours vendeur » (ask price)



- Mais en faisant ceci, il communique de l'information aux acheteurs potentiels

89

Efficiencia informacionnelle et mecanismes de negociation

- Est-ce que les markets-makers ont intérêt à brouiller les pistes en faisant semblant d'avoir une position courte ?
- Par exemple, en augmentant son bid price sur de petites quantités
 - *Hansch, Naik, & Viswanathan (1998). Do inventories matter in dealership markets? Evidence from the London Stock Exchange. The Journal of Finance.*

Do Inventories Matter in Dealership Markets? Evidence from the London Stock Exchange

OLIVER HANSCH, NARAYAN Y. NAIK, and S. VISWANATHAN*

- On retombe sur l'idée que le renforcement learning devrait encourager les techniques de leurres et de brouillage aléatoire (quote stuffing, smoking), voire de manipulation (pump and dump)

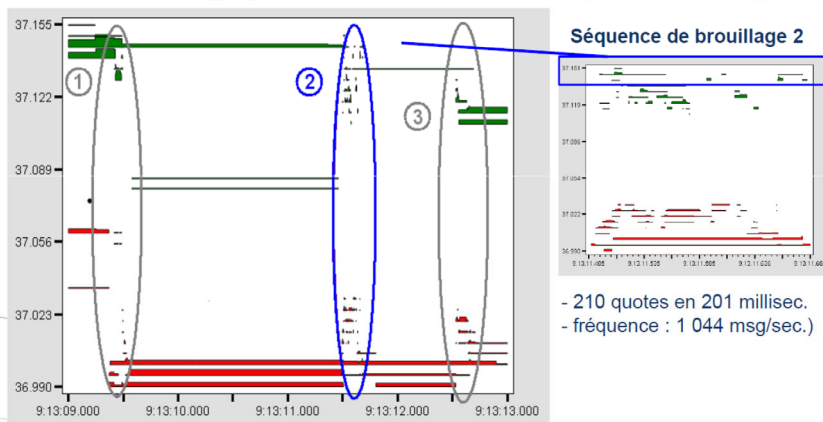
90

Efficiencia informacionnelle et mecanismes de negociation



II. Risques pour l'intégrité/la stabilité des marchés 2. Quote stuffing - Focus (3)

Avant de décaler ses prix, ce market maker HFT lance une séquence de brouillage



Trading Haute Fréquence - 27 mars 2012

<https://www.eifr.eu/files/file9435613.pdf>

91

Efficiencia informacionnelle et mecanismes de negociation



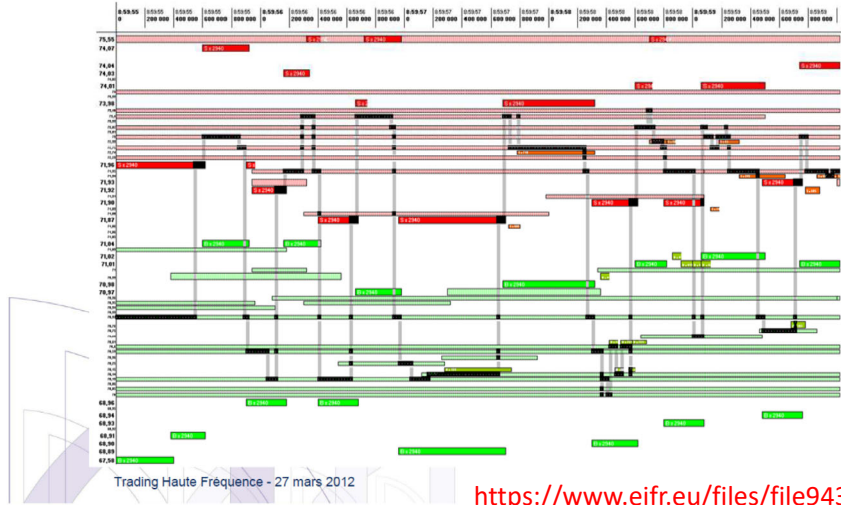
Airbus A400 M effectuant un largage de leurres

92

II. Risques pour l'intégrité/la stabilité des marchés

3. Interaction entre algorithmes

L'interaction d'algorithmes peut induire des dynamiques explosives ("feedback loops") Ex. : fixing d'ouverture d'une action du CAC40



Modes d'échange des titres et révélation de l'information

■ Marché décentralisé

- Rencontres séparées des couples d'acheteurs et de vendeurs
- Un titre peut être négocié à différents prix au même instant.
- Marché immobilier, marché des changes



Belles demeures de France
Fine Residences

CHRISTIE'S
GREAT ESTATES

Daniel FEAU
CONSEIL IMMOBILIER



Modes d'échange des titres et révélation de l'information

■ Marchés gouvernés par les prix

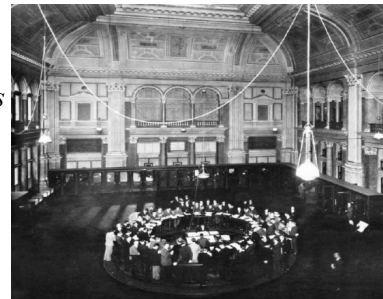
- Quote driven market
- Teneurs de marché ou « Market-makers »
- Le plus souvent de grandes banques
- Affichent des fourchettes de prix (bid-ask)
 - À l'achat et à la vente, pour des quantités données
 - Les prix évoluent selon les affichages des teneurs de marchés
- Marchés de « gros » (« wholesale banking »)
 - pour les investisseurs institutionnels, compagnies d'assurance, banques, grandes entreprises états
 - Blocs d'actions de grandes sociétés, marché interbancaire des changes, obligations d'État

Currency	Buying		Selling
	Notes	T/C	
USA USD	35.03	35.80	36.10
EURO EUR	45.72	45.97	46.88
ENGLAND GBP	68.98	69.66	70.90
JAPAN JPY	0.2879	0.2912	0.2974
SINGAPORE SGD	22.89	22.99	23.59
HONG KONG HKD	4.47	4.50	4.69
AUSTRALIA AUD	27.08	27.36	27.97
NEW ZEALAND NZD	24.18	24.32	25.31
SWITZERLAND CHF	27.99	28.32	28.81
SWEDEN SEK	4.89	5.03	5.19
DENMARK DKK	5.94	6.13	6.29
CANADA CAD	29.65	29.98	30.68
NORWAY NOK	5.43	5.60	5.76
BRUNEI BND	22.10	-	23.50
INDONESIA IDR	0.0025	-	0.009
MALAYSIA MYR	8.53	-	10.76
CHINA CNY	3.70	-	4.97
KOREA KRW	0.029	-	0.043
TAIWAN TWD	0.88	-	1.19
U A E AED	7.75	-	10.20
BAHRAIN BHD	69.63	-	96.33
OMAN OMR	69.41	-	94.21
QATAR QAR	7.67	-	10.22
SAUDI ARABIA SAR	7.58	-	10.19
SOUTH AFRIC ZAR	3.65	-	5.45

Modes d'échange des titres et révélation de l'information

■ Marché électronique

- Acheminement et appariement des ordres d'achat et de vente de titres informatisés.
 - Les risques du « presse-bouton »
- Système en évolution constante
- Disparition de la criée et du parquet depuis 1986
- Loi sur la dématérialisation des titres de 1984
 - Les titres sont conservés chez les intermédiaires dans des registres informatiques
 - En cas d'achat ou de vente, inscription au compte du client le jour ou le lendemain de la négociation.



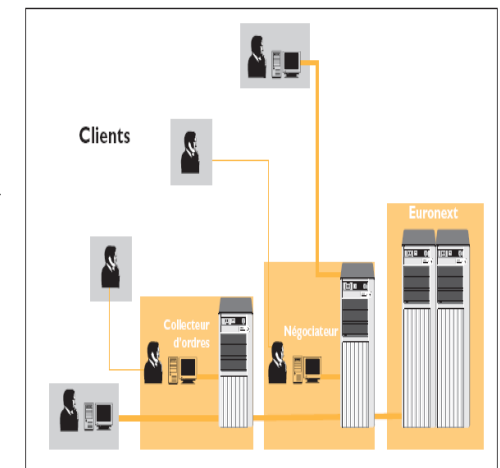
©2002 Shannon Burns www.shannonburns.com



"Oh, I thought you wanted to buy at 64 and sell when it hit 46. Sorry, I never was very good with numbers."

Modes d'échange des titres et révélation de l'information

- Principe de concentration des ordres
 - Aujourd'hui remis en question
- Transmission des ordres des particuliers essentiellement via les sites Internet des établissements financiers
- Les ordres d'achat et de vente sont automatiquement transmis à l'ordinateur central de la Bourse



Carnet d'ordres : principes

■ L'ordre à cours limité

- *Le plus répandu*
- *Exemple à l'achat*
 - ordre d'achat de 100 actions AA à cours limité à 61,05 euros
 - L'acheteur souhaite acheter 100 actions à un prix inférieur ou égal à 61,05 euros
 - Tant qu'il n'existe pas de vendeur à un prix inférieur ou égal à 61,05, l'ordre n'est pas exécuté.
- *Priorité*
 - Un ordre d'achat à un cours plus élevé est exécuté avant un ordre d'achat à un cours plus bas
 - A cours limite égal, l'ordre d'achat le plus ancien est exécuté en premier.

101

Carnet d'ordres : principes

■ Feuille de marché (ou carnet d'ordres)

- *Regroupe tous les ordres d'achat et de vente pour un titre*

Carnet d'Ordres de l'action Olympique Lyonnais

Quantité	Limite	Quantité	Limite
40	62,00	20	62,05
10	61,95	30	62,10
20	61,90	20	62,15
30	61,85	20	62,25
10	61,80	40	62,30

- *Supposons qu'arrive un nouvel ordre d'achat de 100 titres à cours limité à 62,20 euros*
- *Appariement avec les ordres de vente à 62,05 (20 titres), 62,10 (30 titres), 62,15 (20 titres)*
 - Les vendeurs vendent à des cours différents

102

Carnet d'ordres : principes

- *Feuille de marché après l'exécution de l'ordre d'achat*

Carnet d'Ordres de l'action Olympique Lyonnais

Quantité	Limite	Quantité	Limite
30	62,20	20	62,25
40	62,00	40	62,30
10	61,95		
20	61,90		
30	61,85		
10	61,80		

- *Il reste 30 titres à l'achat à 62,20 euros*
 - « Fractionnement » de l'ordre



103

Carnet d'ordres : principes

■ Priorité par le temps

- *Supposons qu'arrive un ordre de vente à cours limité de 30 actions à 62,25 euros*
- *Nouveau carnet d'ordres*

Carnet d'Ordres de l'action OL

Quantité	Limite	Quantité	Limite
30	62,20	50	62,25
40	62,00	40	62,30
10	61,95		
20	61,90		
30	61,85		
10	61,80		

- *Arrive un ordre d'achat de 20 actions à 62,25 euros*

104

Carnet d'ordres : principes

- L'ordre d'achat rencontre des vendeurs à 62,25 euros
 - Problème : il y a 50 actions à la vente à 62,25 euros
 - Mais l'ordre d'achat à 62,25 ne porte que sur 20 actions
 - Qui vend ?
 - Les premiers ordres de vente pour 20 actions sont exécutés
 - L'ordre de vente le plus récent pour 30 actions n'est pas exécuté du tout
 - Priorité par le temps.

Quantité	Limite	Quantité	Limite
30	62,20	30	62,25
40	62,00	40	62,30
10	61,95		
20	61,90		
30	61,85		
10	61,80		

105

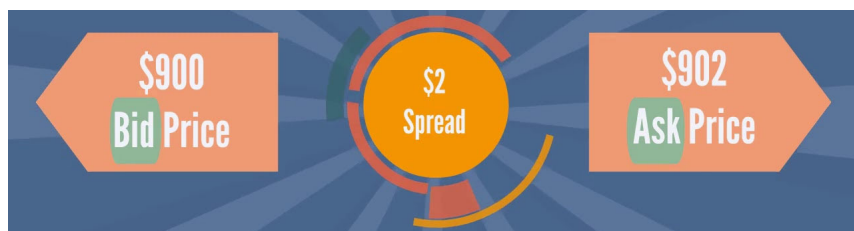
Sondage du carnet d'ordres - transparence



- Turquoise operates a Multilateral Trading Facility (MTF) majority owned by London Stock Exchange Group in partnership with the user community.
- Turquoise Order Book combines **simple limit and iceberg orders** with Large in Scale **hidden orders** and is fully compliant with pre-trade and post-trade transparency obligations.

106

Sondage du carnet d'ordres - transparence

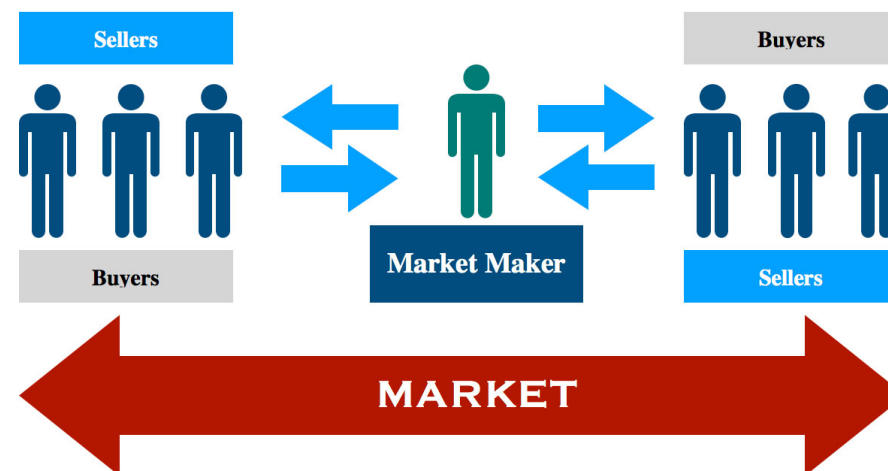


- Ordre à cours limité : apporte de la liquidité au marché centralisé (liquidity provider)
- Tout comme les réponses aux RFQ (request for quote) sur les marchés de gré à gré
- Ordre au marché ou RFQ : « prend » de la liquidité

107

Sondage du carnet d'ordres - transparence

- Market Maker : liquidity provider

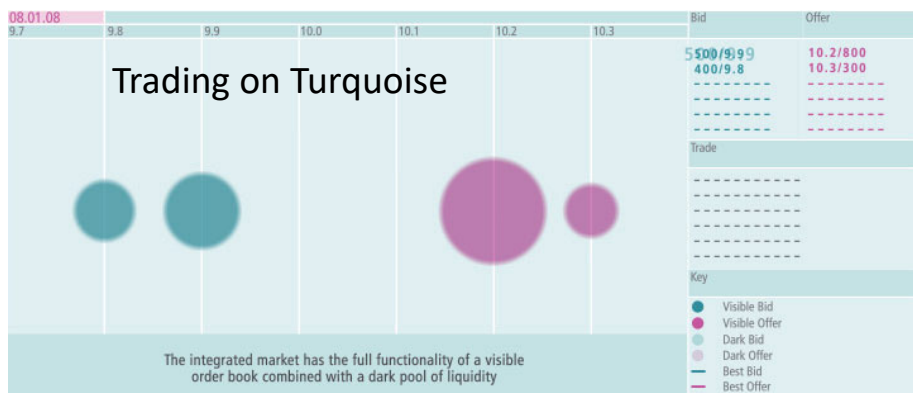


108

Sondage du carnet d'ordres - transparence

■ Transparence des échanges ?

- Dark pools (ordres cachés)
- Iceberg orders
- Flash orders
- Paiement pour placement d'ordres

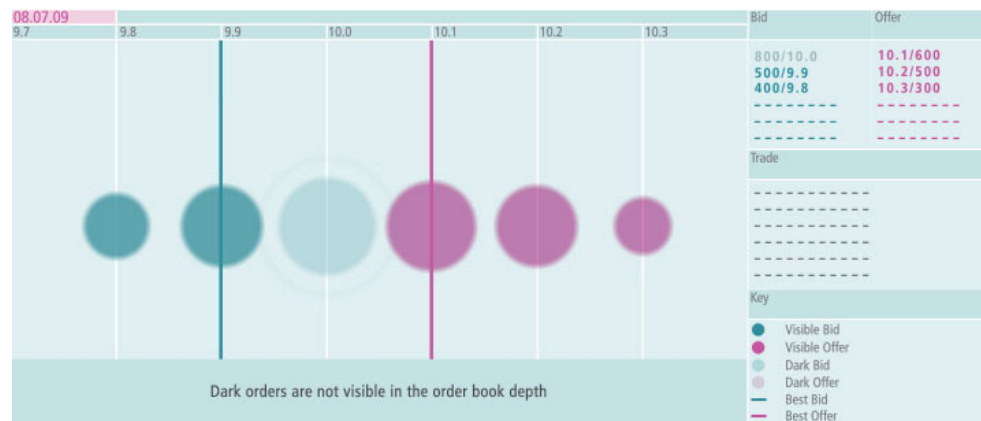


109

Sondage du carnet d'ordres - transparence

■ Transparence des échanges ?

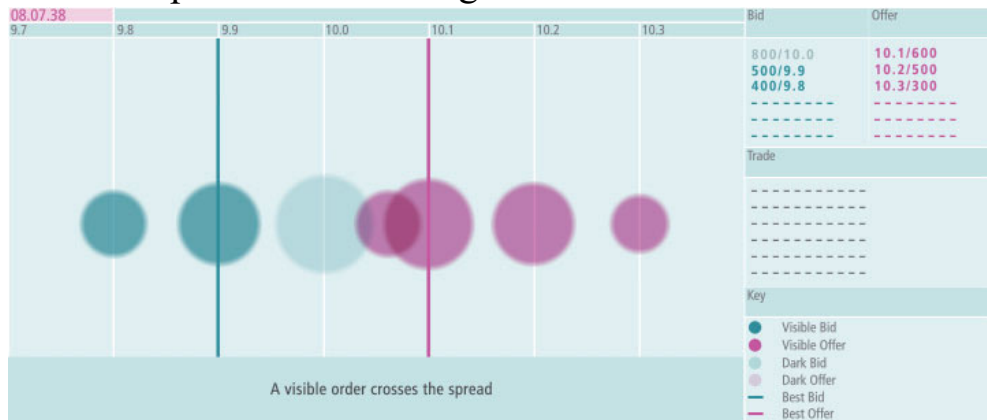
- Fourchette « visible 9,9-10,1
- Fourchette « réelle 10-10,1



110

Sondage du carnet d'ordres - transparence

■ Transparence des échanges ?



- On peut sonder le marché avec un ordre fill or kill pour une petite quantité
 - Soit l'ordre est exécuté immédiatement en totalité ou il est éliminé

111

Sondage du carnet d'ordres - transparence

■ Transparence des échanges ?



- Il n'est plus possible de détecter le « dark order » par un « tatonnement » (passage d'un petit ordre à cours limité)

112

Sondage du carnet d'ordres - transparence

■ Transparence des échanges ?

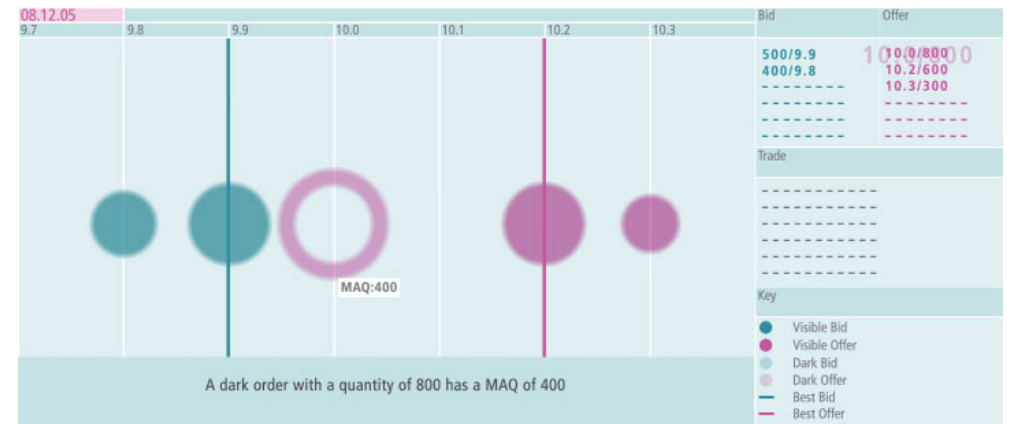


113

Sondage du carnet d'ordres - transparence

■ Transparence des échanges ?

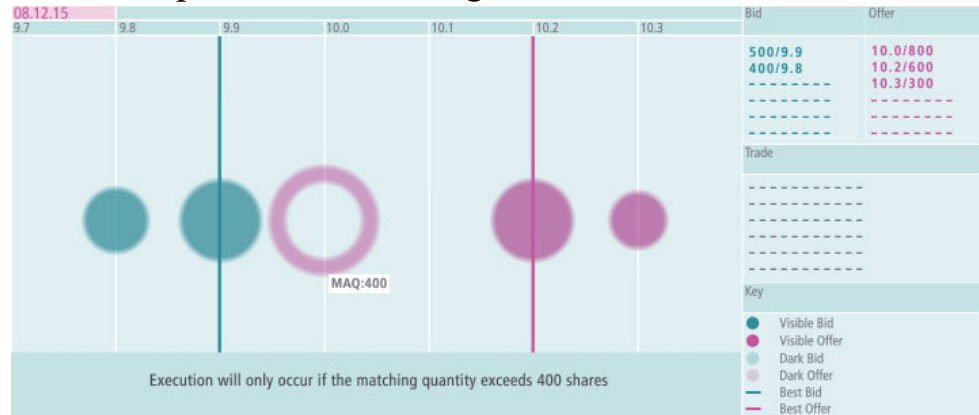
- A minimum acceptable quantity (MAQ) may be added to control the execution of dark order



114

Sondage du carnet d'ordres - transparence

■ Transparence des échanges ?

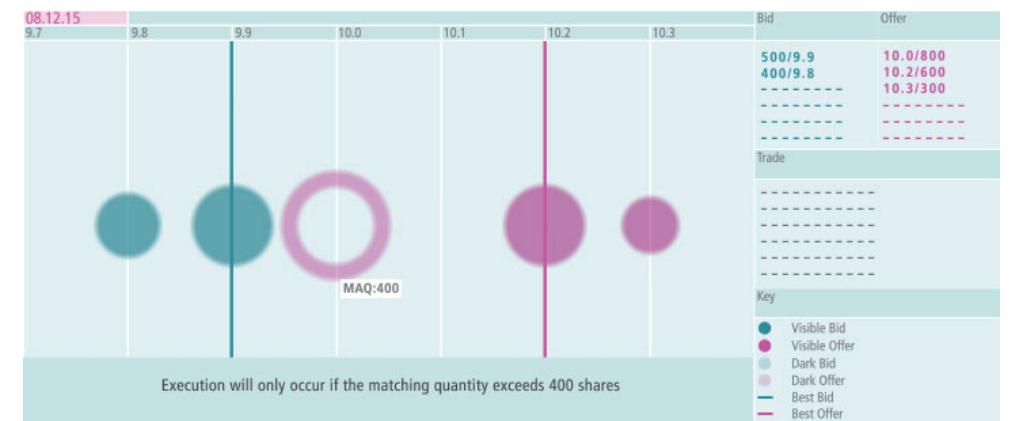


- Ceci permet au dark order de rester invisible face à un ordre à cours limité de montant inférieur à 400

115

Sondage du carnet d'ordres - transparence

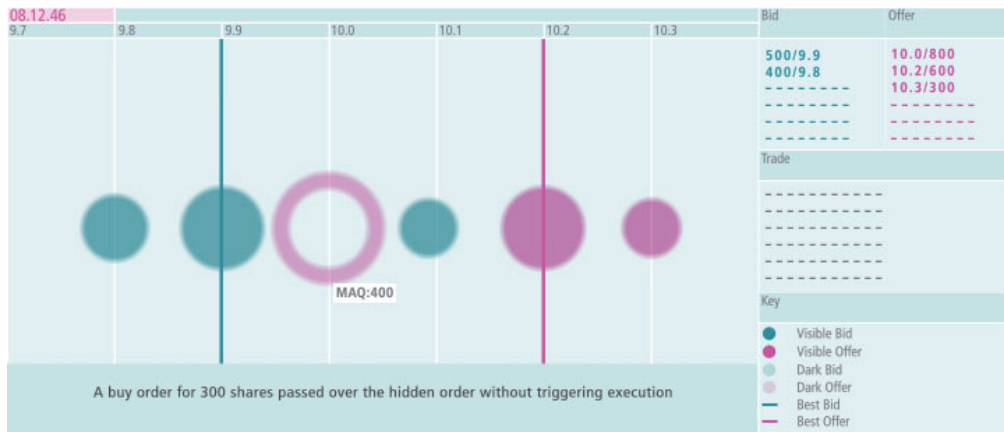
■ Transparence des échanges ?



116

Sondage du carnet d'ordres - transparence

■ Transparence des échanges ?



117

Sondage du carnet d'ordres - transparence

■ Transparence des échanges ?

- On a des ordres d'achat à 10,1 € et des ordres de vente cachés à 10 € sans appariement



118

Sondage du carnet d'ordres - transparence

■ Transparence des échanges ?

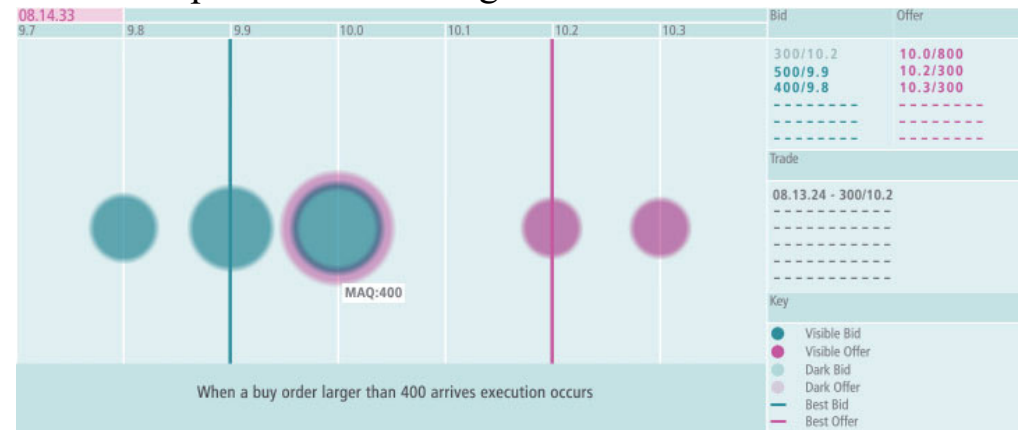
- L'écart visible entre meilleurs bid et ask ne veut rien dire
- Un nouvel ordre d'achat peut arriver et être exécuté en dessous du meilleur prix à l'achat



119

Sondage du carnet d'ordres - transparence

■ Transparence des échanges ?

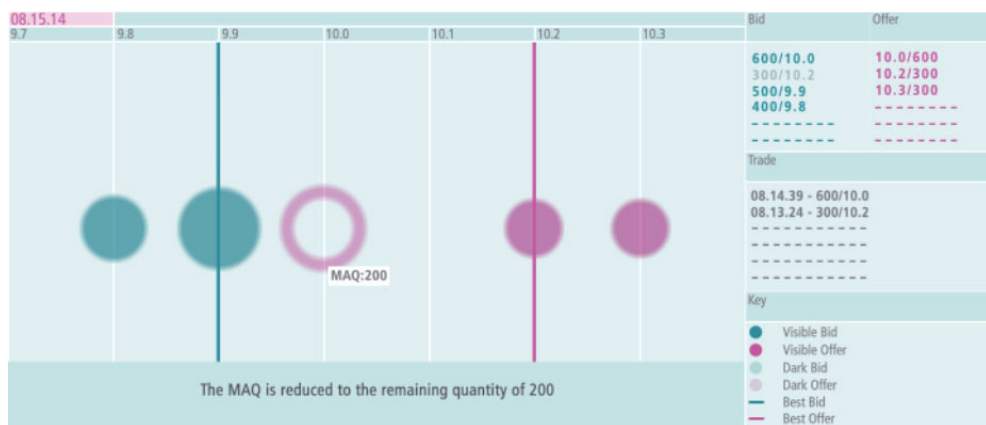


- Gestion de la priorité par les prix
- Utilisation d'ordres fill or kill pour sonder le marché

120

Sondage du carnet d'ordres - transparence

■ Transparence des échanges ?



121

Sondage du carnet d'ordres - transparence

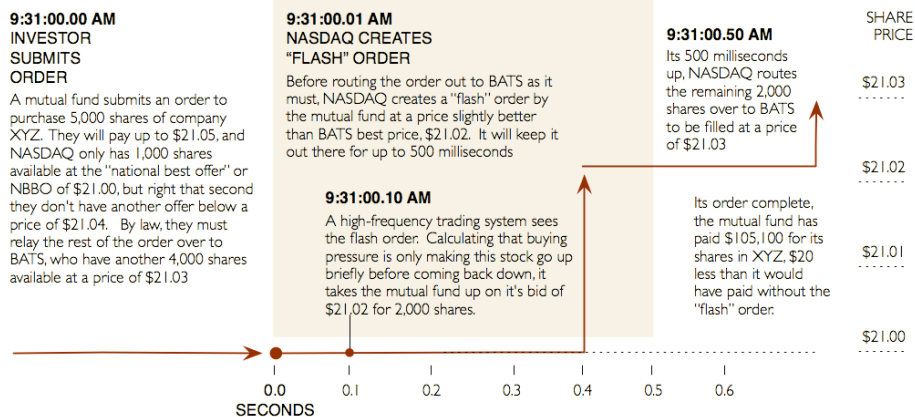
- Iceberg order (ou « hidden size order »)
 - Seule une quantité partielle de l'ordre est affichée
 - Utilisables sur Euronext
 - Quand la quantité affichée est exécutée, une nouvelle quantité affichée (de même montant apparaît)
 - Perd la priorité par le temps
 - peut être réalisé par un automate de trading
- Flash order (Nasdaq)
 - Certains participants au marché ont le droit (moyennant finances) à voir les ordres arriver sur le marché 500 ms avant qu'ils ne soient connus des autres participants
- Rémunération des gros donneurs d'ordres
 - Pourvoyeurs de liquidité

122

Sondage du carnet d'ordres - transparence

The 500-Millisecond Advantage

Under the "Regulation NMS" rules, exchanges are forced to route participants orders out to other exchanges if the other exchange is briefly offering a better price on a stock the customer wants to buy, even if that price would be bad for the customer – and bad for the exchange too, who lose the custom. To avoid this, some exchanges are offering an order option called "flash" that displays for just 500ms, that might help get the customer filled at a better price without breaking the Regulation NMS rules about not holding up orders for longer than a second. Each customer saving might just be pennies, but it can sometimes happen millions of times a day.



123

124

Fragmentation du flux d'ordres



- La fragmentation des marchés
 - Il n'y a plus de monopole de la « Bourse de Paris »
 - On n'est pas obligé d'acheter ou vendre son action sur Euronext
 - Plates-formes alternatives pour exécuter des ordres de bourse : Blink MTF, BATS Europe Chi-X, Equiduct, Turquoise.
 - On a intérêt à acheter là où c'est le moins cher
 - Concurrence entre les places boursières
 - Principe de « best execution » : orientation des ordres vers la meilleure plate-forme de négociation
 - Arbitrage entre les marchés
 - Pour profiter des décalages de prix (acheter bas, vendre haut)
 - HFT (High Frequency Trading)
 - Initiés technologiques: fibre optique transatlantique, microondes

125

Fragmentation du flux d'ordres



<http://www.agefi.fr/articles/marches-boursiers-chi-x-bats-talonne-les-marches-reglementes-1252510.html>

- Plates formes de négociation : parts de marché en France et en Angleterre

Importance de Bats Europe Chi-X et des marchés de gré à gré (OTC « Over The Counter »)

Agefi 17 janvier 2013

Dans le tableau de l'Agefi, les marchés « opaques » incluent les « dark pools, les SI et les marchés OTC

FTSE 100				CAC 40			
Lieu d'échanges	Montant en mds €	Part de marché en %		Lieu d'échanges	Montant en mds €	Part de marché en %	
		2012	2011			2012	2011
Marchés transparents				Marchés transparents			
LSE	842,52	31,88	29,75	Nyse Euronext	681,06	32,30	33,25
Bats Europe Chi-X*	484,30	18,33	4,81	Bats Europe Chi-X*	283,35	13,24	2,64
Chi-X*	-	-	15,73	Chi-X*	-	-	11,16
Turquoise	95,41	3,61	4,11	Turquoise	63,08	2,95	3,03
Nyse Arca	1,13	0,04	0,10	Equiduct	18,24	0,85	0,76
Equiduct	3,20	0,12	0,10	Autres	-	-	-
Marchés opaques*				Marchés opaques*			
Bats Dark + Chi Delta*	36,77	1,39	0,70	Bats Dark + Chi Delta*	15,71	0,73	0,23
Chi-Delta*	-	-	0,67	Chi-Delta*	-	-	0,32
UBS MTF	19,77	0,75	0,20	UBS MTF	8,49	0,40	0,14
Posit	12,22	0,46	0,23	Sigma X	6,14	0,29	0,09
Sigma X	12,16	0,46	0,10	Posit	4,45	0,21	0,07
Turquoise Dark	7,73	0,29	0,58	Turquoise Dark	3,64	0,17	0,29
Liquidnet	4,67	0,18	0,21	Instinet	2,28	0,11	0,04
Instinet	4,55	0,17	0,07	Nomura NX	2,13	0,10	0,10
Nomura NX	4,02	0,15	0,16	Smartpool	1,20	0,06	0,14
Smartpool	1,61	0,06	0,28	Liquidnet	1,14	0,05	0,05
Blink	0,14	0,01	-	Blink	0,32	0,01	-
SI	79,11	2,99	3,07	SI	23,34	1,09	2,89
Marchés OTC	1.033,21	39,10	39,12	Marchés OTC	1.014,77	47,43	44,80

126

Fragmentation du flux d'ordres



- La montée en puissance des nouvelles plates-formes d'échange : BATS en Europe et aux États-Unis

U.S. Equities (BATS Exchanges – BZX, BYX, EDGA, EDGX) February 2014	
Overall Industry – Average Daily Consolidated Volume	7.0 billion shares
BATS Exchanges – Matched Market Share	20.4%
BATS Exchanges – Average Daily Matched Volume	1.4 billion shares
U.S. Equity Options (BATS Options)	
BATS Options – Matched Market Share	3.6%
BATS Options – Average Daily Matched Volume	633,237 contracts
European Equities (BATS Chi-X Europe)	
Pan-European Consolidated Volume ¹ – Average Daily Notional Value	€40.6 billion
BATS Chi-X Europe – Overall European Market Share	21.9%
BATS Chi-X Europe – Total Notional Value Traded	€177.7 billion
BATS Chi-X Europe – Average Daily Notional Value	€8.9 billion

127

Velocity - Managed Colocation
Confidential and Proprietary

Service d'exécution des ordres sur

Europe	Equities	Commodities	Derivatives	Source	Latency (microseconds one way)					
					PRDC	LD4	Interxion	Telehouse	FR2	Basilidon
Bats	✓			LD4	166	0	295	342	4,369	762
BME	✓		✓	TH / IXN	126	295	0	0	4,326	454
Borsa Italiana	✓			LSE	166	335	292	25	4,366	479
Euronext	✓			Basilidon	593	762	467	454	4,793	0
LSE UK Level 2	✓			LSE	166	335	292	25	4,366	479
NASDAQ OMX	✓			Lunda						
Oslo	✓			LSE	166	335	292	25	4,366	479
SWX	✓			CH	4,200	4,369	4,326	4,373	0	4,793
Turquoise	✓			LSE	166	335	292	25	4,366	479
Deutsche Boerse	✓			FR2	4,200	4,369	4,326	4,373	0	4,793
ICE	✓		✓	Basilidon	593	762	467	454	4,793	0

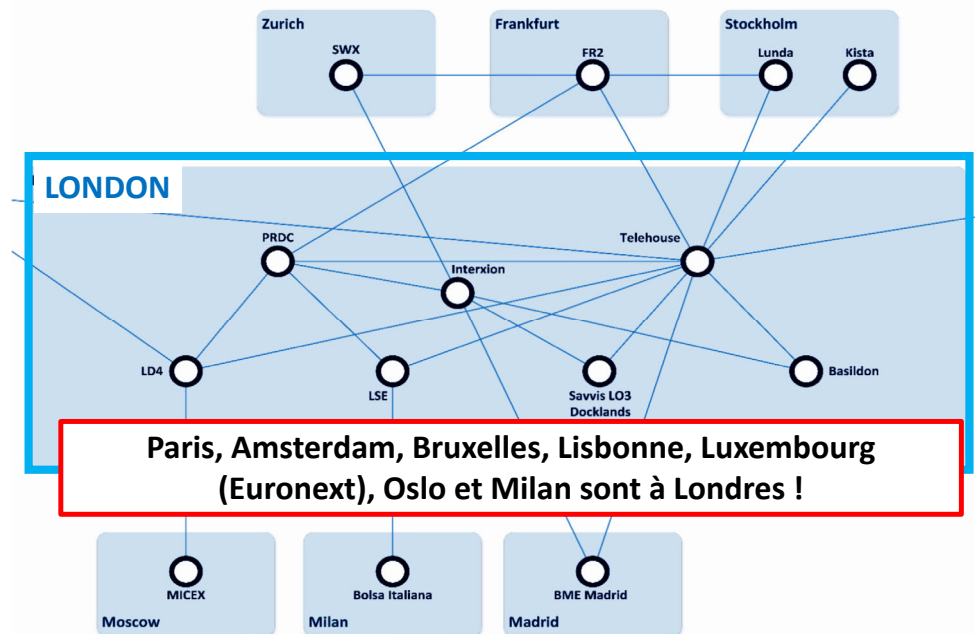
128

En Rouge les principaux marchés organisés en Europe pour le négoce d'actions

	Equities	Commodities	Derivatives	Source	PRDC	LD4	Interxion
Europe							
Bats	✓			LD4	166	0	295
BME	✓		✓	TH / IXN	126	295	0
Borsa Italiana	✓			LSE	166	335	292
Euronext	✓			Basildon	593	762	467
LSE UK Level 2	✓			LSE	166	335	292
NASDAQ OMX	✓			Lunda			
Oslo	✓			LSE	166	335	292
SWX	✓			CH	4,200	4,369	4,326
Turquoise	✓			LSE	166	335	292
Deutsche Boerse	✓			FR2	4,200	4,369	4,326
ICE	✓		✓	Basildon	593	762	467

129

Maillage des data centers européens : réseau interconnecté de data centers de Velocity et centralité de Londres



130

Temps d'interconnexion entre marchés : BATS utilise le data center LD4 (temps de latence nul). Allemagne et Suisse sont isolés (pré-Brexit)

	Equities	Commodities	Derivatives	Source	PRDC	LD4	Interxion	Telehouse	FR2	Basildon
Latency (microseconds one way)										
Europe										
Bats	✓			LD4	166	0	295	342	4,369	762
BME	✓		✓	TH / IXN	126	295	0	0	4,326	454
Borsa Italiana	✓			LSE	166	335	292	25	4,366	479
Euronext	✓			Basildon	593	762	467	454	4,793	0
LSE UK Level 2	✓			LSE	166	335	292	25	4,366	479
NASDAQ OMX	✓			Lunda						
Oslo	✓			LSE	166	335	292	25	4,366	479
SWX	✓			CH	4,200	4,369	4,326	4,373	0	4,793
Turquoise	✓			LSE	166	335	292	25	4,366	479
Deutsche Boerse	✓			FR2	4,200	4,369	4,326	4,373	0	4,793
ICE	✓		✓	Basildon	593	762	467	454	4,793	0

131

Fragmentation du flux d'ordres



- Fragmentation des marchés
 - Possibilité d'exécuter un ordre de bourse sur plusieurs plates-formes de négociation
 - Des prix différents au même moment ?
 - Fractionnement du flux d'ordre
 - Moindre liquidité du fait de la diminution des contreparties ?
 - Impact sur l'efficacité des marchés financiers
 - Pour aller plus loin à propos de l'organisation des marchés
 - « Système Multilatéral de Négociation »
 - Voir le site de l'Autorité des Marchés Financiers
 - <http://www.amf-france.org/Acteurs-et-produits/Marches-financiers-et-infrastructures/Autres-lieux-de-negociation/Systemes-multilateraux-de-negociation.html>
 - « Internalisateurs systématiques »
 - <http://www.agefi.fr/articles/boursorama-optimise-ses-couts-d-execution-avec-l-appui-de-sg-cib-1249485.html>

132

Trading à haute-fréquence, automatisation et relocalisation des négociations

- Les données de prix d'actifs financiers *disponibles* dépendent de l'**organisation des marchés** où se négocient les titres (**infrastructures de marché**)
- Deux cas extrêmes : le **carnet d'ordres centralisé** et le marché de gré à gré décentralisé
- Marché **centralisé** : les **ordres d'achat et de vente** convergent vers un lieu unique
 - Anciennement un lieu physique, une **place de marché**, comme le Palais Brongniart à Paris où se rencontraient les agents de change et leurs courtiers
 - Aujourd'hui, les places de marché sont **virtuelles**. C'est un ordinateur qui reçoit les ordres de bourses, les **apparie** et stockent ceux qui sont en **attente d'exécution**

133



Palais Brongniart : bâtiment et « Corbeille »



New York Stock Exchange; Wall Street et « trading pits » (1937)



134



Maurice Allais, prix Nobel d'Économie, Professeur à l'École des Mines, défenseur du « fixing » quotidien

I think a single daily price quotation in each place for each stock would be by far preferable and would benefit both small and large investors

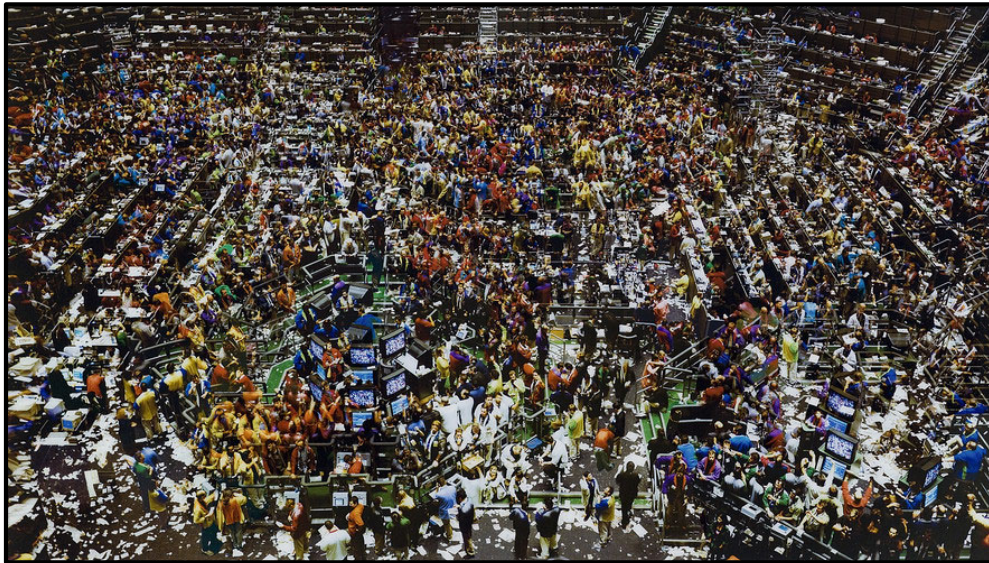
135

Avantages et inconvénients du « fixing »

- Cotation en continu :
 - Diminue le temps de latence et répond mieux aux besoins instantanés de liquidité
- Inconvénients et avantages du fixing
 - Temps de latence
 - mais moins d'effets indésirables de microstructure : price impact, manque de profondeur du marché, risque de manipulations boursières
- Pour les petites valeurs, le fixing peut-être utilisé.
- En cas d'arrivée massive d'informations nouvelles, on peut suspendre la cotation en continu (circuit breakers)

136

L'ancien pit floor du CME (Chicago Mercantile Exchange)



137

l'ancien lieu central de négociation du CME (Chicago Mercantile Exchange)



138

Une visualisation
du concept de
« corbeille » :
l'organisation
circulaire facilite
l'acheminement des
ordres et leur
exécution



139



Unité de mesure pour le High Frequency Trading : un milliardième de seconde. Que signifie le passage de la fréquence quotidienne ? Une seconde à cette échelle représente 4 millions d'années à l'échelle de la journée

140

Relocalisation physique et automatisation

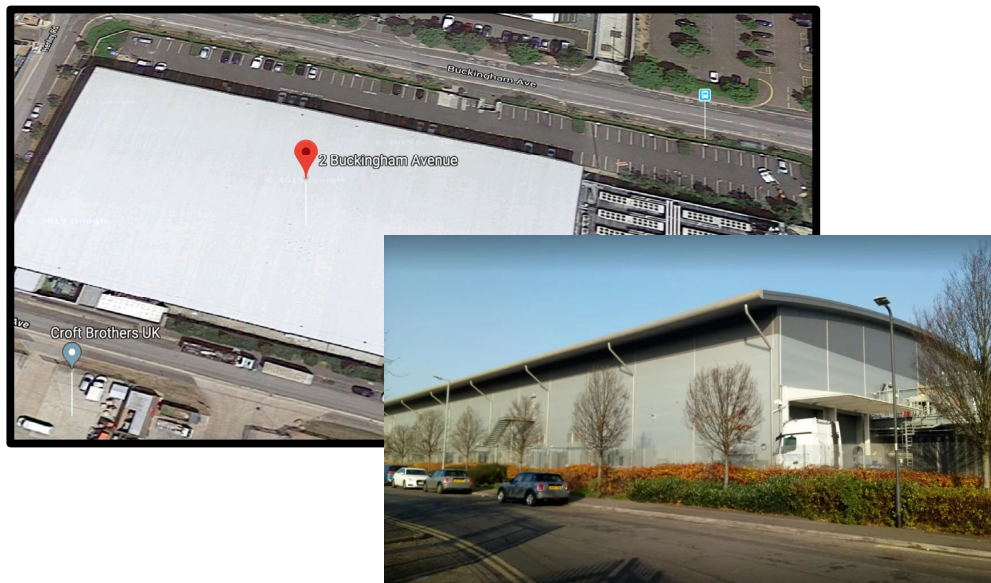
- Il faut être très proche de l'ordinateur qui apparie les ordres d'achat et de vente
- Les donneurs d'ordres louent des serveurs à proximité immédiate de l'ordinateur.
- **Relocalisation autour de « data centers »**
- Une majorité d'ordres est exécutée selon des ordinateurs exécutant des algorithmes, sans intervention humaine.
- Disparition des traders sur actions, recrutement massifs d'ingénieurs en IT, télécommunications, machine learning, etc.

141



En bleu, antennes d'une filiale du NYSE (août 2019), la flèche sur le toit : nouvelle localisation. En jaune, antennes utilisées par les réseaux concurrents. 300 mètres supplémentaires de câblage en fibre optique : retard d'exécution de **2 millionnièmes de secondes**

142



Data Center LD4, Equinix, 2 Buckingham Avenue
Slough Trading Estate (utilisable en **colocation**)

143

Plates-formes de négociations

- Si une action est achetée via la plate-forme Euronext Paris, la transaction n'est pas pour autant effectuée à Paris
 - Après avoir été installés à Aubervilliers
 - Suresnes pour les serveurs de sauvegarde
 - Les ordinateurs d'Euronext Paris sont installés dans le « liquidity center » de Basildon dans la banlieue de Londres
 - <http://www.youtube.com/watch?v=3b821UX-Xw0>



144



Foires de Lübeck et de Leipzig

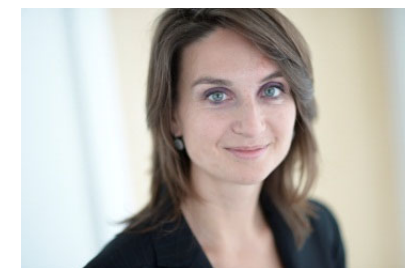


15

146

Plates-formes de négociation (trading venues) moins réglementées (transparence pre-post trade) : brokers, dark pools, OTC markets (anglo-saxon + suisse, brokers de données : Bloomberg, Thomson Reuters)

Brokers		
Alliance Bernstein	Morgan Stanley	Morgan Stanley Night Vision
ANZ	Neovest	ITG Posit
Bank of America Merrill Lynch	NewEdge	Bloomberg Tradebook
Barclays	UBS	IEX
ICAP BrokerTec	Westpac	OTC Markets
Citi	Dark Pools	Currenex
Credit Suisse	KCG Knightmatch	eSpeed (Nasdaq)
Deutsche Bank	UBS Pin	FX SpotStream
Goldman Sachs	Credit Suisse Crossfinder	FXAll
ICAP EBS	Barclays LX	GainFX
Instinet (Nomura)	JP Morgan JPMX	Hotspot/Knight
ITG	Citi Match	Liquidnet
JP Morgan Chase	Deutsche Bank SuperX	LMAX
Lime	Goldman Sachs SigmaX	Thomson Reuters Matching
	Bank of America Merrill Lynch	
	Instinct	



Natasha Cazenave et Verena Ross



Robert Ophèle, Président de l'AMF

En Europe, les régulateurs nationaux (AMF, ...) et européen (ESMA, European Securities and Market Authority) doivent faire face à des enjeux technologiques et politiques importants.

Difficultés pour mettre en place une supervision coordonnée et efficace des marchés de capitaux : projet ambitieux de CMU, marché européen unifié des capitaux

Le flash krach du 6 mai 2010

Le 6 mai 2010, Wall Street connut, entre 14 h 41 et 14 h 45 et 28 secondes, les 4 minutes 30 les plus longues de son histoire.

Les actions américaines s'effondrèrent de près de 10 %, le Dow Jones perdit 1.000 points dans des volumes élevés.

Dans le quart d'heure qui suivit, les actions remontèrent mais non sans avoir atteint parfois des valeurs aberrantes : une action Apple valut jusqu'à 100.000 dollars alors que titre Accenture chuta un moment à 1 cent.

Neuf ans après le flash krach du 6 mai 2010, on n'en a que des explications parcellaires !



<https://www.bloomberg.com/view/articles/2015-04-21/guy-trading-at-home-caused-the-flash-crash>

153

Pathologies possibles du trading haute fréquence

Vues de l'AMF (Autorité des Marchés Financiers)

- **Risque systémique et perturbations du mécanisme de formation des prix**
 - Le flash crash du 6 mai : rôle perturbateur, mais non moteur du HFT
 - Sur le marché français : quelques exemples vécus de pertes de contrôle d'algos, ou d'interactions défavorables entre algorithmes, notamment en pre-opening ou pre-closing
- **Abus de marché :**
 - Layering /spoofing : déséquilibrer le carnet par ses ordres, pour réaliser une transaction en sens inverse. Nombreuses variantes. Sanctions aux Etats-Unis (Trillium), travaux en Europe.
 - Incertitude juridique sur certains schémas « limite » à expertiser :
 - Momentum ignition : accompagner / déclencher une bulle de très court terme en espérant attirer des suiveurs, puis déboucler la position
 - Fixation indirecte du prix importé sur un dark pool
 - Sondages de carnet
 - Quote-stuffing : envoi d'ordres en grand nombre, souvent répétitifs, sans logique économique, pour perturber ou freiner la lecture du carnet par les autres participants, sonder de manière répétée leurs intentions, ou masquer ses propres anticipations
 - « Smoking » : envoi d'ordres passifs attractifs en espérant déclencher des ordres « au marché », et repositionnement à un niveau plus élevé avant d'être « tapé »

4

LesEchos.fr

Trading haute fréquence : l'AMF sanctionne lourdement Euronext et Virtu

LAURENCE BOISSEAU | LE 08/12/15 À 08H39



Virtu Financial, one of the world's top electronic trading firms, yesterday switched its European headquarters from London to Dublin

155

Matt Levine, Bloomberg : de la difficulté à prouver la manipulation boursière



<https://www.bloomberg.com/view/articles/2014-10-16/high-speed-traders-put-a-bit-too-much-gravy-on-their-meat>

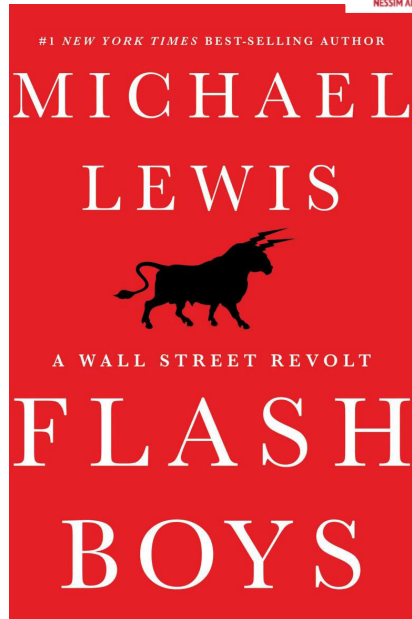
One good general rule is that it's harder than you think it is to figure out what's market manipulation and what isn't.

Trading a lot, cancelling a lot of orders, putting in orders or doing trades on both sides of the market, trading a lot right before a close or fixing -- all of those things could be signs of nefarious manipulation, or just normal risk management.

156

« Flash Boys » : le putsch de Wall Street par les « golden boys » 2.0

NESSIM AÏT-KACIMI - LES ECHOS | LE 29/01/2016



Livre vivant et captivant, procès à charge du High Frequency Trading

Des débats complexes sur la réglementation :

- 1) Order matching randomisation (Latency Floor / Order Batching)**
- 2) Minimum Quote Life**

157

158

159

160

Prévoir le futur à partir du passé



La question est bien sûr la suivante : est-ce ces **relations entre données passées et futures** sont *fortuites* ou sont-elles *significatives* ?

161

Prévoir le futur à partir du passé

- Fonction indicatrice de hausse du prix :
 - $I_t = 1$ si $P_t - P_{t-1} > 0$
 - $I_t = 0$ si $P_t - P_{t-1} \leq 0$
- Une chronique de hausses et de baisses est donc représentée par une suite binaire du type 1010001011
 - Ici, pour simplifier l'analyse, on ne prend pas en compte l'amplitude des hausses ou des baisses
 - Dans le cadre du HFT (High Frequency Trading), les prix ne varient pas plus que d'un échelon de cotation (tick)
- Supposons que les 10 dernières variations de prix à la hausse ou la baisse soient représentées par 1010001011
- Prévoir à partir de ces observations revient à compléter cette suite : 1010001011**1** ou 1010001011**0**

162

Prévoir le futur à partir du passé

- Prévoir à partir de ces observations revient à compléter cette suite : 1010001011**1** ou 1010001011**0**
- Dans cet exemple, ce qui sert à prévoir (information disponible, inputs pour la prévision) se limite à la connaissance des hausses ou baisses du titre risqué pour les dix dernières transactions.
- Si l'on pense que les variations de prix d'un autre actif précèdent celui qui nous intéresse (phénomène lead-lag), on peut inclure ces variations dans l'ensemble des informations pertinentes
- **1010001011**
0100010110 la seconde ligne représente ici les variations de prix de l'actif lead

163

Prévoir le futur à partir du passé

- Considérons la représentation binaire (en base 2) d'un nombre : 11010 ... (suite infinie de 1 et de 0)
 - En finance, 1 (hit) si hausse des prix, 0 si baisse
- « Machine learning » : prévoir à partir de données en repérant des « **patterns** » (**motifs**) **répétitifs**
 - Exemple : les suites périodiques répètent un même motif.
 - La suite 1010101010 est une répétition du motif 10.
 - La fréquence des 0 et des 1 est égale à $\frac{1}{2}$: **il y a autant de hausses que de baisses**
 - **Mais on peut gagner à tous les coups, en vendant après une hausse et en achetant après une baisse : stratégie « contrarian »**

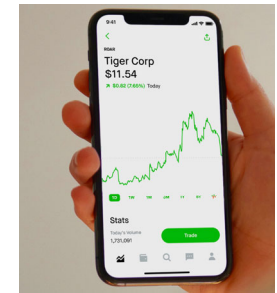
164

Prévoir le futur à partir du passé

- L'information qui peut être utilisée pour la prévision peut être de nature financière et comptable, annonces de résultats, notes d'analystes financiers, ...
 - Voir la suite des transparents à propos du rôle des analystes financiers.
- L'information peut provenir de l'analyse textuelle - des tweets ou des discussions sur les forums en ligne, notamment reddit
 - <https://www.reddit.com/r/RobinHood/>
 - RobinHood est un acteur majeur du brokerage en ligne à prix cassés

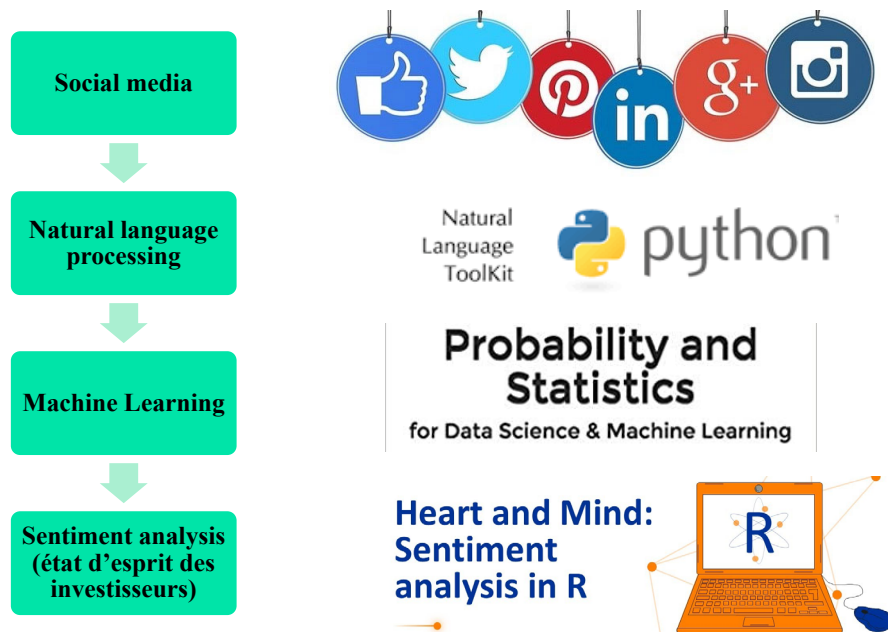


165



166

Utilisation d'informations textuelles



167

Comment prévoir le futur à partir du passé

- Idée de base du machine learning
- Soit le motif (pattern, feature) 1010001011, on recherche une **fonction** qui dont l'entrée (input) est ce motif
 - Ici un nombre codé sur 10 bits
- Et la sortie (output) est soit 1 soit 0.
 - On a utilisé le terme « fonction » – l'entrée est une **variable**.
 - On aurait pu utiliser le terme **programme** ou **algorithme**.
- Dans notre exemple, il y a $2^{10} = 1024$ entrées possibles
- Pour **calibrer** la fonction, on a une base d'apprentissage, constituée de motifs sur 11 jours.
- On va rechercher les motifs identiques (ou proches) de 1010001011

168

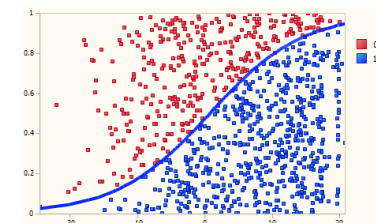
Comment prévoir le futur à partir du passé

- Pour les motifs identiques à 1010001011, on va examiner si le successeur est un 1 ou si c'est un 0
- On peut aussi voir ce problème de **prévision**, comme une **classification** (voir infra)
 - Pour un motif donné, on veut savoir s'il est suivi d'une hausse ou d'une baisse
 - **Deux classes** : les antécédents d'une hausse et les antécédents d'une baisse
- Il s'agit d'**apprentissage supervisé** : on indique ce qui doit être prévu (hausse ou baisse) dans l'échantillon de données « historique » (base d'apprentissage)

169

Comment prévoir le futur à partir du passé ?

- Classification binaire : chaque point est associé à un motif de variation des cours observé



- Si un motif est en bleu, il est suivi d'une hausse (attribution à la classe des hausses)
- En rouge, le motif est suivi d'une baisse (attribution à la classe des baisses)
- Ici tout fonctionne bien, la zone des points bleus et celle des points rouges sont bien dissociées : prévision parfaite
 - **Décision optimale** : achat si motif bleu, vente si motif rouge

170

Comment prévoir le futur à partir du passé : difficultés théoriques et pratiques

- Cela a à voir notamment avec (la nature du problème considéré) le nombre d'observations dans la base d'apprentissage et le nombre possible d'entrées
- Exemple précédent : on se sert des dix dernières observations de hausse ou baisse pour la prévision, soit $2^{10} = 1024 \approx 10^3$ antécédents (entrées) possibles
- Supposons que l'on ait 10^5 éléments dans la base d'apprentissage.
- On aura (en moyenne) $10^5 / 10^3 = 100$ observations / cas.
- Pour un motif donné, si toutes ces 100 observations sont suivies d'une hausse (ou bien d'une baisse), on prédira une hausse (pour ce motif)

171

Comment prévoir le futur à partir du passé ?

- Considérons à nouveau ce qui suit le motif 1010001011
 - Supposons qu'il y a 500 occurrences de ce motif dans la base d'apprentissage
 - Si elles sont suivies d'un 1 (hausse), on sera conduit à prédire qu'une nouvelle occurrence du motif, sera suivie d'une hausse.
 - Si elles sont toutes suivies d'un 0 (baisse), on prédira une baisse
 - Nous discuterons de la validité de ces inférences dites par **induction** ...
 - généralisation de la base d'apprentissage à l'ensemble des situations non encore observées
 - **Induction probable** : si 495 occurrences sur 500 sont suivies d'une hausse, on aura tendance à prévoir également une hausse :
 - $\frac{495}{500} = 99\%$: fréquence conditionnelle d'apparition de 1 après 1010001011.
 - 99% est proche de 1
 - 1% : fréquence conditionnelle d'apparition de 0 après 1010001011.

172

Comment prévoir le futur à partir du passé ?

- Considérons à nouveau ce qui suit le motif 1010001011
 - Il y a 500 occurrences de ce motif dans la base d'apprentissage
- Prédications imparfaites dans l'échantillon
 - *Quelle prédiction faire si la fréquence conditionnelle d'apparition de 1, après le motif précédent est de 56% ?*
 - Dans l'échantillon, le motif 1010001011 est plus souvent suivi d'une hausse 1 que d'une baisse.
 - Mais est-ce que l'écart 56% contre 44% est suffisant pour l'on décide d'acheter (anticipant une hausse) ?
 - *Quel serait le seuil (en termes de fréquences) à partir duquel on devrait acheter ?*
 - *Est-ce qu'adopter la **règle majoritaire** (acheter dès que la fréquence conditionnelle est > 50%) est une stratégie pertinente ?*

173

Comment prévoir le futur à partir du passé ?

- **Distance de Hamming** entre deux motifs (voir exercices corrigés)
 - Hamming, Intro to The Art of Doing Science and Engineering: Learning to Learn (March 28, 1995)
 - <https://www.youtube.com/watch?v=AD4b-52jtos&list=PL2FF649D0C4407B30>
 - *Considérons deux motifs de hausses et de baisses*
 - 1010001011
 - 111100011
- La distance de Hamming est ici de 3 : Nombre de cas (dates) où hausses et baisses ne coïncident pas
- Distance de Hamming nulle : deux motifs identiques
- Distance de Hamming faible : deux motifs proches.

174

Comment prévoir le futur à partir du passé ? méthode des k – plus proches voisins (introduction)

- Le problème des petits échantillons
 - *Supposons maintenant qu'il y ait 2 occurrences 1010001011 suivies de 1 et aucune suivie de 0, dans la base d'apprentissage*
 - *La fréquence conditionnelle d'apparition d'une hausse est de 100%, mais la robustesse de la prévision est faible*
- Que se passe-t-il dans le cas précédent ou si un nouveau motif n'a pas été observé dans la base d'apprentissage ?
 - *Une idée est de considérer des motifs proches, par exemple, des motifs dont la distance de Hamming avec le nouveau motif est inférieure ou égale à 1*
 - C'est-à-dire des motifs qui ne diffèrent du motif qui apparaît que d'un plus pour une date.
 - *L'hypothèse qui est faite ici est que des prémisses proches ont les mêmes conséquences (pas de disruption)*

175

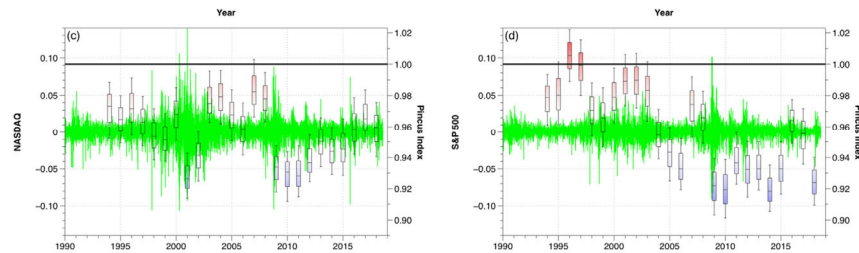
Comment prévoir le futur à partir du passé ? méthode des k – plus proches voisins (introduction)

- Le problème des petits échantillons (suite)
 - *Quand la taille de l'échantillon (nombre de motifs observés) diminue, il y aura davantage de motifs qui n'auront que pas ou trop peu de voisins*
 - *Pour que l'on fasse une prédiction quelconque et donc que l'on puisse décider d'acheter ou de vendre.*
- La taille de l'échantillon est à mettre en rapport avec la longueur des motifs
 - *Supposons que l'on ait 1000 observations*
 - *Si motifs de longueur 3, en moyenne 125 observations par motif*
 - *Si motifs de longueur 10, en moyenne 1 observation par motif*
 - *Si motifs de longueur 15, en moyenne 30 motifs pour 1 observation*
 - *Nécessité de compresser l'information : encodage*

176

Comment prévoir le futur à partir du passé ? Évolution du degré de prévisibilité

- L'indice de Pincus a été considéré dans l'article suivant :
 - Delgado-Bonal (2019). Quantifying the randomness of the stock markets. Scientific reports, Nature.



- Une faible valeur de l'indice est associée à des répétitions plus fréquentes des motifs : marchés plus prévisibles
- Évolution de l'indice pour le Nasdaq et le S&P500 : graphique ba – chart échelle de droite
- On remarque des fluctuations de la prévisibilité (mesurée par l'indice de Pincus).

177

prévisibilité : un peu de vocabulaire

what's the
opposite of
predictable?



unpredictable, unforeseen,
unexpected, improbable,
unlikely, capricious, freakish,
unforeseeable, uncertain



Un peu de vocabulaire anglais : si l'on examine les termes de droite, on se rend compte qu'ils recouvrent des concepts différents. Par exemple, imprévisible, improbable, incertain et inattendu n'ont pas le même sens.

178

prévisibilité : un peu de vocabulaire

what's the
opposite of
unpredictable?



predictable, certain,
unchanging, stable, steady,
reliable, unvarying, constant,
sure, foreseeable

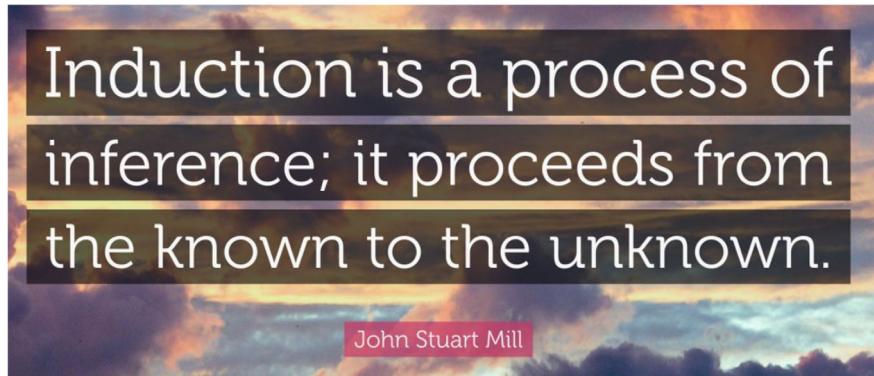


Même remarque que dans le transparent précédent. Par ailleurs, la forme littéraire permet de connoter de manière subtile les termes employés, mais n'a pas la précision du formalisme mathématique.

179

180

Prévision en finance et méthode inductive



181

Induction : Pourquoi est-ce important en finance ?

- L'induction est liée à l'**apprentissage** à partir des **données en vue de faire des prévisions**
 - Données : observations de cours boursiers, communications financières, informations textuelles.
 - Données massives (big data) associées au développement du trading à haute fréquence
- De passer de l'**échantillon** à la **population**.
 - Pour prévoir des cours boursiers futurs (market timing), de repérer des sociétés sous-évaluées (stock-picking).
 - **Évaluer des sociétés non-cotées par comparaison**
- Concevoir des algorithmes de trading
 - Stratégies boursières à partir de l'analyse des données
 - Apprentissage statistique (machine learning).

182



La finance utilise de plus en plus de données, notamment via l'émergence du trading à haute fréquence, la multiplicité des plates formes de négociation, l'utilisation de données comportementales (sentiment de marché), souvent issues des réseaux sociaux, l'analyse textuelle ... Ces données font l'objet de traitements automatisés, économétrie et statistiques « classiques », mais aussi apprentissage profond.

183

Induction : Pourquoi est-ce important en finance ?

- La « logique inductive » sous-tend le problème de la prévision des cours boursiers, des ratios financiers.
- Deux approches pour les cours boursiers
 - *Les cours boursiers ont un certain degré de prévisibilité*
 - Et par conséquent, on peut envisager d'utiliser et construire certains algorithmes en vue de construire des stratégies boursières gagnantes
 - C'est la situation où la « logique inductive » trouve son application
 - *Les cours boursiers ne sont pas prévisibles*
 - Et par conséquent, aucune méthode inductive, aucun algorithme de trading, même utilisant des méthodes avancées de deep learning ne peut fonctionner de manière robuste
 - Problème (plus ou moins) caché de validation (back-testing) statistique des pseudo-stratégies gagnantes.

184

Qu'est-ce que l'induction (par énumération) ?

- Induction de type (i) et de type (ii)
 - Prémisse : tous les cygnes observés sont blancs



- (i) : Le prochain cygne observé sera blanc
 - Intéresse le spéculateur : en remplaçant blanc par hausse du marché et cygne par mouvement du marché boursier.
 - Inférence singulière ou par énumération simple ou prédictive
- (ii) : tous les cygnes sont blancs
 - Inférence généralisante ou par énumération ou nomologique : loi qui intéresse le théoricien.

185

Qu'est-ce que l'induction ?

- Soit un ensemble d'observations ou « faits empiriques »
 - Exemple : motif (pattern) de trois hausses consécutives (codé en binaire 111) suivi d'une baisse (codée en binaire 0)
 - Supposons que toutes les observations dans l'ensemble donné (appelé échantillon en statistique) vérifient cette propriété.
 - L'induction énumérative ou de type (i) consiste à affirmer que la prochaine fois que l'observera 111, 0 (baisse des cours)
 - L'induction (i) correspond à une prévision (utile au financier)
 - L'induction de type (ii) consiste alors à affirmer que cette propriété est vraie pour tous les motifs, y compris ceux qui n'ont pas été observés, on n'observera jamais 1111
 - Cad une hausse ne suivra jamais trois hausses consécutives
 - L'induction (ii) est nomologique : on établit une loi (utile au scientifique)

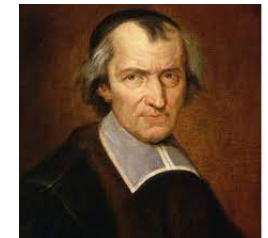
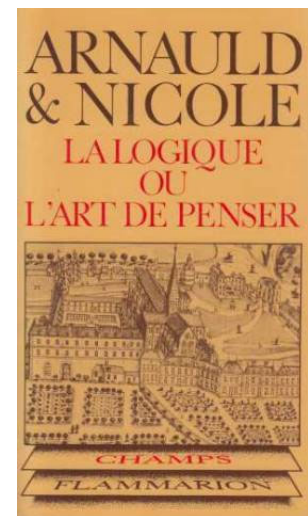
186

Induction
=
passage du particulier au général

"On appelle *induction* lorsque la recherche de plusieurs choses particulières nous mène à la connaissance d'une vérité générale. Ainsi, lorsqu'on a éprouvé sur beaucoup de mers que l'eau est salée, et sur beaucoup de rivières que l'eau en est douce, on conclut généralement que l'eau de la mer est salée et celle des rivières douce." [Arnauld & Nicole], p. 321

Antoine Arnauld - Pierre Nicole : la logique ou l'art de penser

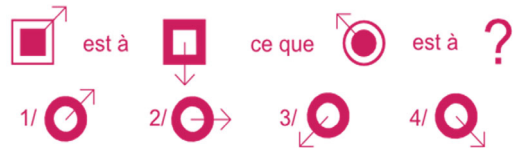
187



188

Qu'est-ce que l'induction ?

- Dans sa forme la plus générale (ii), l'induction vise à établir des lois (ou régularités)
- Mais on est amené à procéder très souvent par **analogie** : dans l'exemple ci-dessous de test de QI, on ne dispose que d'une observation et il faut inférer une règle de prédiction



189

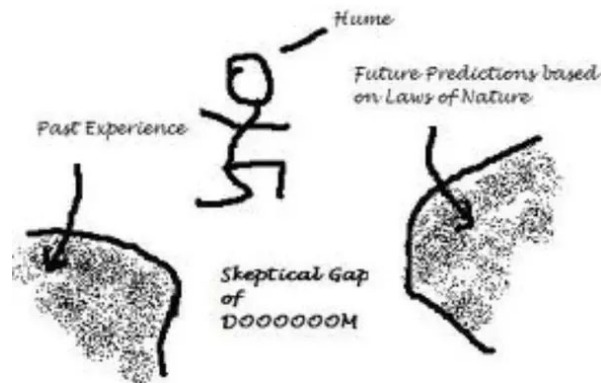
Qu'est-ce que l'induction ?

- Les raisonnements **analogiques** (par **association**) peuvent très souvent véhiculer des stéréotypes et leur **valeur prédictive** peut être **très douteuse** !
 - Exemple : types de personnalité de Kretschmer (1925)

Type	Physique	Personnalité	Acteur
Pycnique	Petit et rond	Expansif, gai, spontané, réaliste	Gérard Jugnot Dany de Vito
Leptosome	Grand et mince	Réservé, froid, rêveur	Jean Rochefort Clint Eastwood
Athlétique	Large carrure et musclé	Impulsif, coléreux	Lino Ventura Harvey Keitel
Dysplasique	Mal développé, anomalie	Asthénique, se sent inférieur	N'a pas réussi dans le cinéma

- L'intérêt pédagogique de cet étrange tableau est de montrer le lien entre **prévision** et **classification**

190



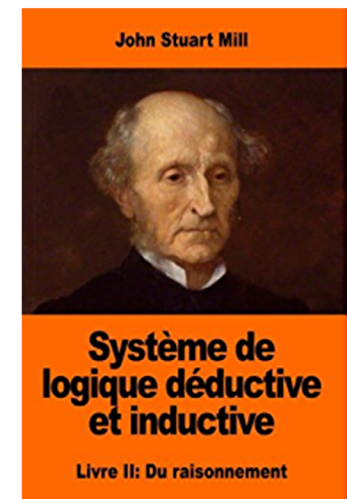
The Inductive Leap

Saut qualitatif non justifié logiquement entre données d'observations et données à prévoir

191

Qu'est-ce que l'induction ?

John Stuart Mill est l'un des premiers auteurs à avoir examiné l'utilisation de l'induction en vue de produire de nouvelles connaissances scientifiques



192

Les limites des connaissances inductives sont connues (citation de Mill)

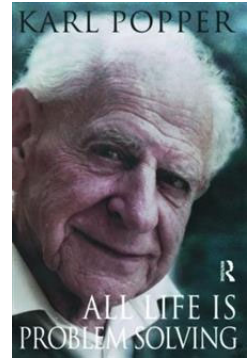
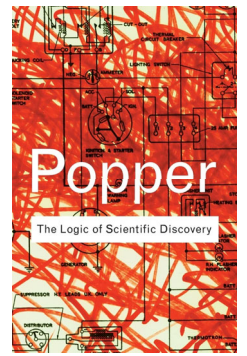
- Quelques unes, (...) crues bonnes pendant des siècles étaient, en réalité, mauvaises.
- L'induction que tous les cygnes sont blancs n'a pas pu être bonne puisque la conclusion a fini par se trouver fausse ».
- Cependant, l'expérience sur laquelle reposait la conclusion était de bon aloi.
- De temps immémorial, **le témoignage de tous les habitants du monde connu, concordante en un résultant commun, sans un seul exemple de déviation dans ce résultat, n'est pas toujours suffisante pour établir une conclusion générale.**



193

Les limites des connaissances inductives sont connues (Karl Popper)

- “Now it is far from obvious, from a logical point of view, that we are justified in inferring universal statements from singular ones, no matter how numerous; for any conclusion drawn in this way may always turn out to be false:
- No matter how many instances of white swans we may have observed, this does not justify the conclusion that all swans are white.”
 - *Popper, Karl, The Logic of Scientific Discovery*



194

Qu'est-ce que l'induction ?

- Approche empirique
 - *établir des régularités entre des observations (données, faits)*
- Loi (ou propriété)
 - *relation à caractère général entre les données*
 - *Données observées et non encore observées*
 - *Loi : passage de l'échantillon (statistique descriptive) à la population (statistique inférentielle)*
 - *Une loi portant sur des données non encore observées peut donc servir à faire des prédictions*

195

Qu'est-ce que l'induction ?

- Induction ou inférence inductive : forme de raisonnement qui permet de passer du **particulier** au **général**
 - *Le raisonnement est **ampliatif** (qui accroit, complète) : La conclusion contient davantage d'information que les prémisses*
 - *Prémisses : les données déjà observées*
 - *Conclusion : les propriétés des « données » non encore observées*
 - *Inférence factuelle ampliative : raisonnement par lequel, à partir de prémisses fondées sur des observations effectives (« factuelle »), on dépasse l'expérience finie pour atteindre une conclusion nouvelle (« ampliative »), sur ce qui n'a pas (encore) été expérimenté*

196

Déduction, induction, abduction

- Trois types de raisonnement
 - **Induction** : *prévision des valeurs boursières*
 - **Abduction** : *recherche des déterminants (variables explicatives) des valeurs boursières (Pierce)*
 - **Déduction** : *relations logiques entre valeurs boursières : actions vs dette, analyse de la diversification des risques, relation entre risque et rentabilité*

- Pour résumer, la **déduction**, qui repose sur des causes et des effets certains, aboutit à des énoncés certains ; l'**induction**, qui propose des causes certaines à des effets probables, aboutit à des énoncés probables ; et l'**abduction**, qui recherche des causes probables à des effets certains, aboutit à des énoncés plausibles.
 - Nicolas Chevassus-au-Louis

Validité du raisonnement inductif

- Mill invoque le principe d'**uniformité des lois de la nature**
 - « Lorsque ces diverses uniformités sont constatées par une induction jugée suffisante, on les appelle communément des Lois de la Nature »
 - Pour justifier ce principe d'uniformité, Mill fait appel à un principe de causalité : toutes choses égales par ailleurs, **les mêmes causes produisent les mêmes effets**.
- Causes des mouvements des prix des actifs financiers ?
Passages d'ordre (market impact) et **nouvelles** (non anticipées) **informations économiques et financières**
 - **Abduction** : recherche (difficile !) des **informations pertinentes** impactant les cours boursiers
 - Voir les transparents consacrés au biais de confirmation et au biais rétrospectif
- Mis à part les financiers formés par la physique, l'hypothèse d'uniformité rencontre peu de succès en finance.

201

« Le » problème de l'induction (Hume)

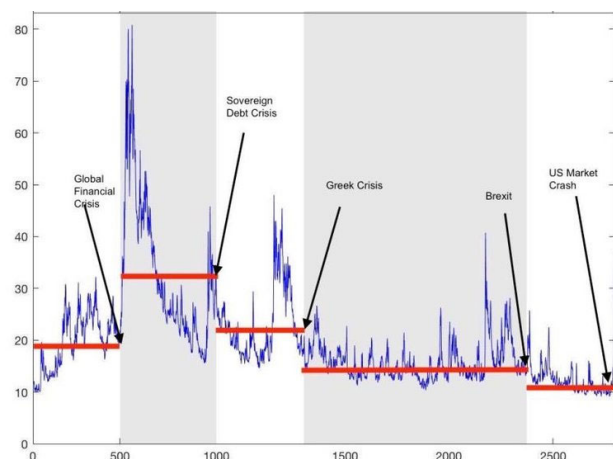
- Pour Nelson Goodman, ce qui ressort d'une **prétendue uniformité de la nature** s'apparente à une « **hypothèse non confirmée et peut-être même douteuse** » et il lui paraît « insensé » de justifier ainsi des prévisions
 - *Faits, fictions et prédiction* (page 78)
- Les économètres de la finance parlent de **stationnarité** (invariance des lois de probabilité des rentabilités) et de « structural breaks » (changement dans ces lois)



202

Validité du raisonnement inductif

- Stationnarité ou « structural breaks » ? évolution du risque sur les actions



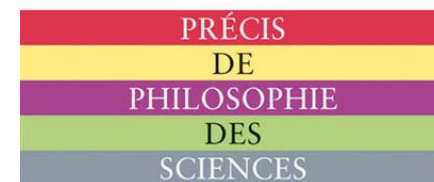
203

Validité du raisonnement inductif

- Est-ce qu'une augmentation d'observations dans le sens attendu augmente la plausibilité de l'hypothèse ?
 - Si l'on a vu beaucoup de cygnes blancs, est-ce qu'il est plus vraisemblable que le prochain cygne observé soit blanc ?
 - Ceci va dans le sens de ce que l'on appelle la **théorie bayésienne de la confirmation (TBC)**

M. Cozic (2010), "Confirmation et induction" in Barberousse & al., eds, *Précis de Philosophie des Sciences*, vol.1, Paris: Vuibert

Sous la direction de
Anouk BARBEROUSSE, Denis BONNAY & Mikael COZIC



204

Validité du raisonnement inductif : analogie, coïncidence ?

- Ian Fleming dans le roman Goldfinger, dont est issu le film homonyme

MR BOND, THEY HAVE A SAYING IN CHICAGO: 'ONCE IS HAPPENSTANCE. TWICE IS COINCIDENCE. THE THIRD TIME IT'S ENEMY ACTION.'

- Selon le proverbe français, jamais deux sans trois...Et pour Agatha Christie (Nemesis)

Any coincidence is worth noticing. You can throw it away later if it is only a coincidence.

205

Validité du raisonnement inductif

- Il y a même des situations où **une seule expérience** nous permet une prévision fiable.



- Supposons que votre niveau d'anglais soit un peu limité : vous ne comprenez pas le contenu du panneau.
- Combien de fois allez-vous plonger votre main dans l'eau bouillante pour en conclure que ça brûle ?!

206

Validité du raisonnement inductif

- Mais une seule expérience (ou expérience singulière) n'est pas forcément la meilleure manière de faire une prévision...

Le chat qui s'assoit sur le couvercle d'un poêle brûlant [...] ne s'assiera plus jamais sur un poêle chaud et c'est très bien ainsi ; mais il ne s'assiera plus jamais non plus sur un poêle froid.

– Mark Twain¹

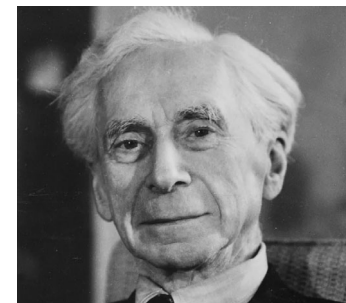
- On remarquera également qu'il suffit d'une seule expérience pour l'apprentissage, ce qui remet en cause la justification de l'induction par une « habitude » (Hume)

207

Validité du raisonnement inductif

- L'abus du raisonnement inductif est problématique pour la prise de décision...
 - *“The man who has fed the chicken every day throughout its life at last wrings its neck instead, showing that **more refined views as to the uniformity of nature** would have been useful to the chicken.”*

- Bertrand Russell



The Demand For Certainty Is One Which Is Natural To Man, But Is Nevertheless An Intellectual Vice.

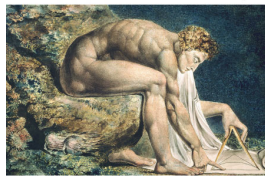
208

Validité du raisonnement inductif

- Dans le même esprit que Russel, le philosophe (des sciences) Quine considère le contre-exemple suivant
- Chaque jour qui passe et où l'on constate que nous sommes vivants devrait augmenter notre croyance en l'immortalité ...



Willard V. O. Quine
et Joseph S. Ullian
LA FABRIQUE DES
CROYANCES
Une introduction au
raisonnement scientifique

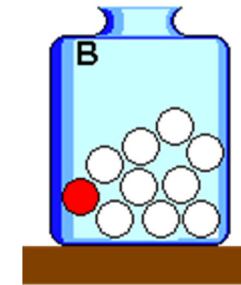


éditions
marcus halter

209

Validité du raisonnement inductif

- Supposons que nous considérons une urne avec une boule rouge et des boules blanches
- Chaque tirage (sans remise) d'une boule blanche diminue et non pas augmente la probabilité de tirer une boule blanche au coup suivant



210

Validité du raisonnement inductif

- Si l'on retient l'hypothèse d'efficience des marchés sous sa forme initiale ...
 - *Les cours boursiers suivent des marches aléatoires*
- Une longue succession de hausses n'est en aucun cas un indicateur de hausse future
 - *Pas de tendances*
 - *Voir la partie du cours sur les biais cognitifs (biais de représentativité, gambler fallacy, hot hand fallacy)*
- Une hausse des cours supplémentaires

211

Validité du raisonnement inductif

- Théorie bayésienne de la confirmation et paradoxe des corbeaux de Hempel



212

Validité du raisonnement inductif

- 1) "tous les corbeaux sont noirs"
- 2) "si quelque chose est un corbeau, alors ce truc est noir"
- 3) "si quelque chose n'est pas noir, alors ce n'est pas un corbeau"
- 4) "tous ce qui n'est pas noir, n'est pas un corbeau"

■ Source : Maël Cavan

- La proposition 1) est une reformulation de 2)
- La proposition 3) est la contraposée de la proposition 2)
- Elle est logiquement équivalente à 2)
- Enfin 4) est une reformulation de 3)

213

Validité du raisonnement inductif

- Le paradoxe ...

Observer un parapluie jaune augmenterait la croyance envers le fait que tous les corbeaux sont noirs !

- La solution ...

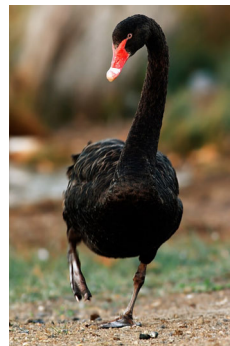
Il y a beaucoup plus d'objets non noirs que de corbeaux donc observer un objet non noir qui n'est pas un corbeau augmente la crédence de manière beaucoup plus négligeable qu'observer un corbeau noir

- Mais on ne connaît pas le nombre respectif d'objets non noirs et de corbeaux ...

214

Validité du raisonnement inductif

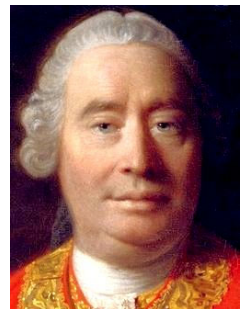
- Pour la petite histoire, tout comme il y a des cygnes noirs, il y a des corbeaux blancs !
- L'induction ne fonctionne donc pas dans ce cas
 - *Même s'il est plus probable de rencontrer un corbeau noir*



215

« Le » problème de l'induction (Hume)

- Hume a remarqué que l'induction n'a pas de fondement logique.
 - *Justifier la méthode inductive par l'uniformité de lois de la nature ?*
 - *Comment peut-on justifier le principe de l'uniformité des lois de la nature ?*
 - *En répétant des expériences ...*
 - *C'est-à-dire par la méthode inductive...*
 - **Raisonnement circulaire**
 - *On ne peut pas établir avec certitude une connaissance générale à partir de cas particuliers.*



David Hume,
philosophe écossais
(1711-1776)

216

« Le » problème de l'induction (Hume)

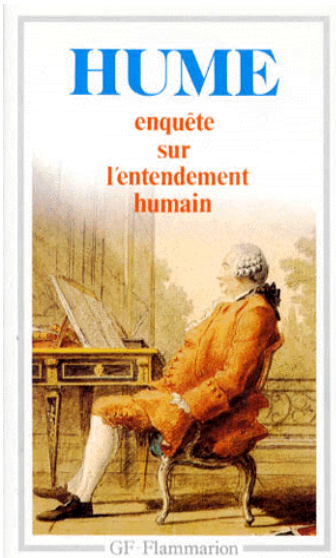
Dire qu'elle (la connaissance) vient de l'expérience, c'est une pétition de principe.

Car toutes les inférences tirées de l'expérience supposent, comme fondement, que le futur ressemblera au passé et que des pouvoirs semblables seront conjoints à de semblables qualités sensibles.

S'il y a quelque doute que le cours de la nature puisse changer et que le passé ne puisse être la règle pour l'avenir, toutes les expériences deviennent inutiles et ne peuvent engendrer d'inférence ou de conclusion.

Il est donc impossible qu'aucun argument tiré de l'expérience puisse prouver cette ressemblance du passé au futur, car tous les arguments se fondent sur la supposition de cette ressemblance.

David Hume, 1748

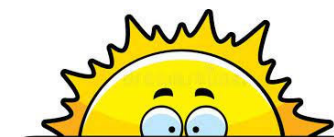
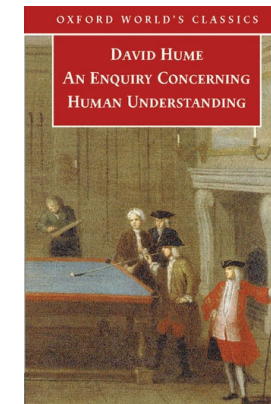


217

« Le » problème de l'induction (Hume)

■ “That the sun will not rise tomorrow is no less intelligible a proposition, and implies no more contradiction than the affirmation, that it will rise. We should in vain, therefore, attempt to demonstrate its falsehood”.

■ “Mr. Locke divides all arguments into demonstrative and probable. In this view, we must say, that it is only probable (...) that the sun will rise tomorrow”.



218

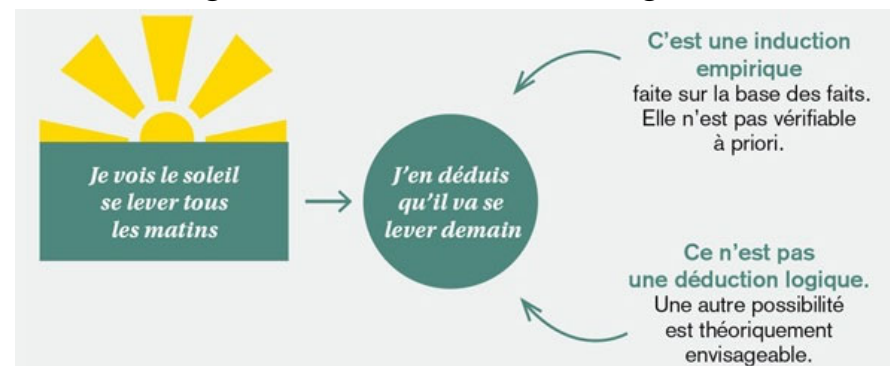
« Le » problème de l'induction (Hume)

- Rappelons que la prévision des cours boursiers repose sur l'idée de répétition de « motifs » dans l'échantillon
 - Supposée persister dans le futur et permettre des gains boursiers.
- Hume (Traité de la Nature Humaine) ?
 - « Persuadez une bonne fois pour toutes les hommes de ce principe : même après l'observation d'une fréquente ou constante conjonction d'objets, nous n'avons aucune raison de tirer aucune inférence au sujet d'aucun objet autre que ceux dont nous avons eu l'expérience »
 - On a pu observer certaines régularités dans l'échantillon des observations, sans que celles-ci puissent s'étendre aux quantités non observées, cad à l'ensemble de la population.
 - Scepticisme quant à la possibilité de prévisions boursières.

219

Notre connaissance du futur vient-elle de l'expérience ?

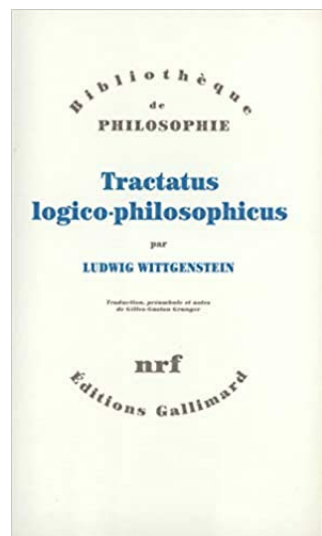
- Le problème de Hume : au cœur de la prévision boursière via les algorithmes de machine learning



- Limite à l'approche empirique (data driven), qui met en avant l'expérience comme base de la connaissance

220

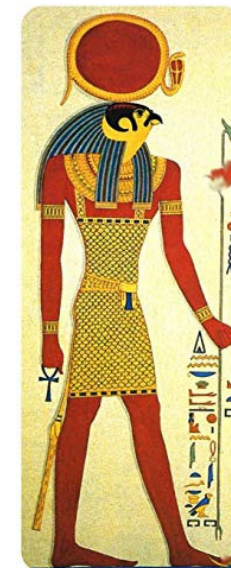
« Que le soleil se lèvera demain est une hypothèse », citation extraite du Tractatus et clin d'œil à la problématique de l'induction



221

Notre connaissance du futur vient-elle de l'expérience ?

- Le soleil se lèvera demain !
- Pour les Egyptiens de l'Antiquité, Rê voyage chaque jour à travers le ciel à bord de sa barque sacrée (parcours du Soleil), et chaque nuit au travers des mondes souterrains (la Douât).
- Chaque lever de soleil est une victoire remportée par Rê sur les « forces des ténèbres ».



222

Notre connaissance du futur vient-elle de l'expérience ?

- Le soleil se lèvera demain (David Hume)
 - *Traité de la nature humaine (1739-1740)*
- Tous les jours et depuis toujours, le soleil se lève. Il en sera donc de même demain.
- Laskar (2013) : « Le Système solaire est-il stable ? »
 - *le mouvement des planètes dans le Système solaire est chaotique, ce qui interdit toute prédiction précise de leur mouvement au-delà de quelques dizaines de millions d'années. Les simulations montrent que les collisions planétaires ou les éjections sont possibles.*
 - <https://www.franceculture.fr/emissions/la-methode-scientifique/la-methode-scientifique-emission-du-mardi-23-juin-2020>

223

Notre connaissance du futur vient-elle de l'expérience ?

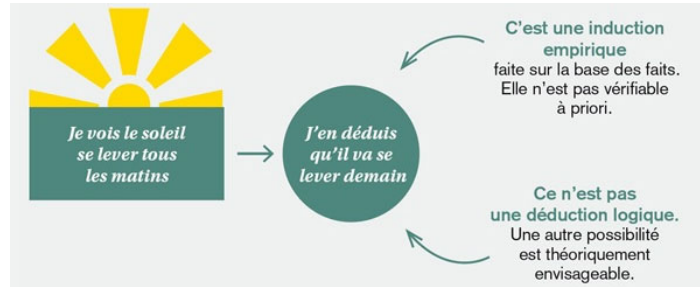
- La « provocation » de Laplace
 - *Il suppose qu'on ne sait pas si le soleil se lèvera demain*
 - *Il postule une distribution uniforme a priori sur la probabilité correspondante (mémoire, 1774) : principe d'indifférence*
 - *Laplace suppose que l'on sait que le soleil se lève chaque matin depuis environ 5000 ans (l'invention de l'écriture)*
 - Soit environ $n = 1826200$ jours
 - *La probabilité que le soleil se lève demain est égale à $\frac{n+1}{n+2}$...*
 - (règle de succession)
 - Ou qu'il ne se lève pas demain à $\frac{1}{n+2}$, soit une chance sur deux millions...
 - A titre de comparaison, il y 139 millions de combinaisons à l'EuroMillions
 - Emile Borel (Valeur pratique et philosophie des probabilités, 26-27), fait remarquer que l'hiver, au-delà du cercle polaire, le soleil ne se lève pas tous les jours ; d'où l'importance de bien spécifier une hypothèse ...

... tirée de l'événement observé. On trouve ainsi qu'un événement étant arrivé de suite un nombre quelconque de fois, la probabilité qu'il arrivera encore la fois suivante est égale à ce nombre augmenté de l'unité, divisé par le même nombre augmenté de deux unités. En faisant, par exemple, remonter la plus ancienne époque

224

Notre connaissance du futur vient-elle de l'expérience ?

- C'est notre compréhension de la mécanique céleste qui nous permet de savoir que le soleil se lèvera demain



- Limite à l'approche « data driven », qui met en avant uniquement l'expérience comme base de la connaissance
- Pas de théorie dans un algorithme de machine learning

D'où vient le raisonnement inductif ?

- Pourquoi est-il si couramment utilisé ?
- D'après Hume, c'est l'**habitude** qui guide notre pensée, pas la **logique**.
- Habitude selon Hume :
 - « transition coutumière de l'imagination d'un objet à celui qui l'accompagne habituellement ».
 - C'est une disposition habituelle de l'esprit humain
 - **Mécanisme associatif** et « psychologique »

229

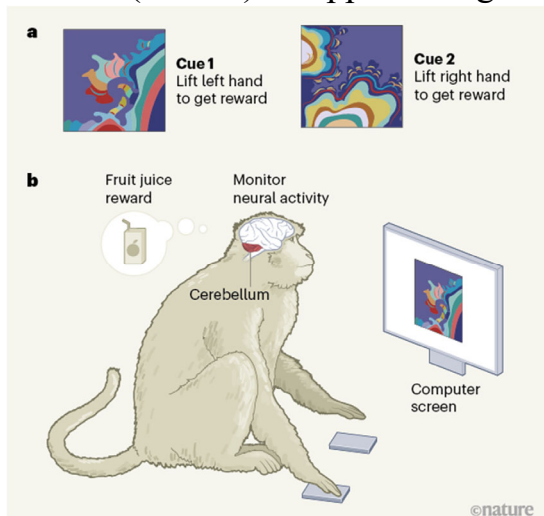
D'où vient le raisonnement inductif ?

- Pour Hume, l'animal, comme l'homme, apprend à connaître les choses extérieures par l'**habitude**, c'est-à-dire par la perception répétée des mêmes successions de phénomènes.
- Mais le principe inductif n'est pas (uniquement) de l'ordre de la pensée ou du langage,
- Même si Daniel Kahneman a mis en évidence le problème de l'ancrage en psychologie cognitive.
 - A quoi pensez-vous après le mot « ancre » ?
 - Normalement, bateau ; c'est le principe associatif
 - De manière plus étonnante, si on demande à quelqu'un de faire une prévision, elle sera plus élevée si elle est précédée de la vision d'un nombre élevé...

230

D'où vient le raisonnement inductif ?

- Bases biologiques de l'induction : *réflexe conditionné* (Pavlov) ou apprentissage



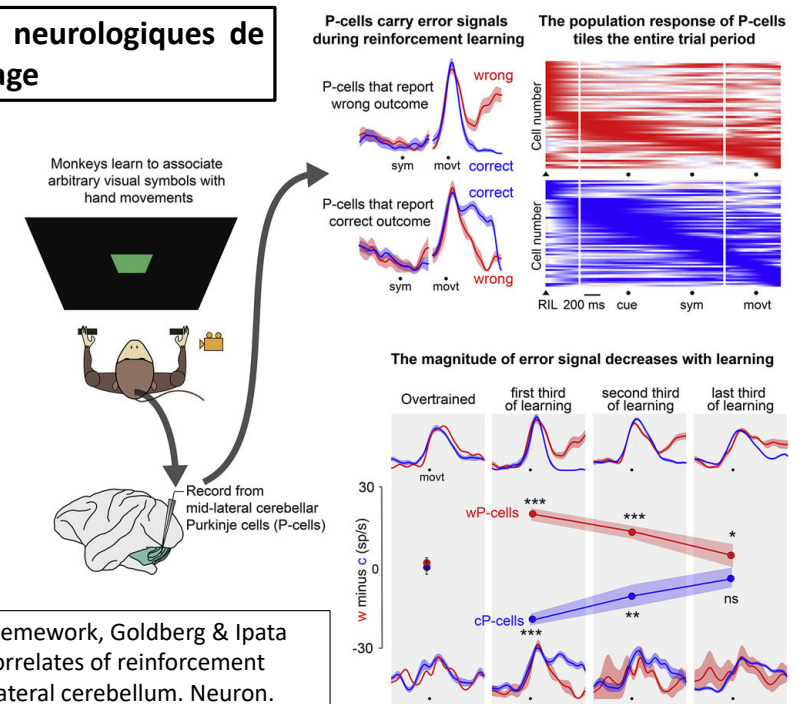
Reconnaissance d'image : Selon l'image qui apparaît à l'écran, le singe doit lever le bras droit ou le bras gauche à l'écran pour obtenir une récompense.

Le cervelet (cerebellum) est impliqué dans l'apprentissage par renforcement (reinforcement learning)

Ce que nous voudrions faire, c'est montrer un graphique des cours passés sur un écran Bloomberg et savoir si c'est un prédicteur de hausse ou de baisse des cours ...

231

Implications neurologiques de l'apprentissage



Sendhilnathan, Semework, Goldberg & Ipata (2020). Neural correlates of reinforcement learning in mid-lateral cerebellum. Neuron.

232

D'où vient le raisonnement inductif ?

- On retrouve ces bases animales et biologiques de l'induction dans l'analyse du « conditionnement » de Pavlov
 - Pavlov. Conditioned reflexes: an investigation of the physiological activity of the cerebral cortex. *Annals of neurosciences*
 - *Pavlov s'intéresse aux « réflexes conditionnels »*
 - *Après un temps d'apprentissage, un stimuli donné ou information va provoquer une réponse comportementale donnée*
 - *De même qu'un enfant qui aurait par mégarde approché sa main d'une flamme sait instinctivement*
- L'habitude de Hume, bien qu'elle puisse être vue sous un angle psychologique relève aussi de cette base animale

Quelles inductions (prévisions) faisons-nous ?

- L'habitude ou la propension de l'esprit à repérer des régularités par association n'épuise pas la question méthodologique
- **Reprenons Quine** : supposons que nous fassions la prévision (funèbre) de notre décès le 30 septembre 2023
 - *Nous sommes le 29 septembre 2023*
- Depuis le jour de notre naissance, chaque jour augmente la « crédence » de notre prévision.
 - *Si elle s'avère exacte, ce sera notre dernière prévision ...*
- Mais chaque jour augmente également la « crédence » entre notre immortalité
- Pourquoi devrions-nous induire notre décès demain ou jamais ?
 - *« Hume néglige le fait que certaines régularités établissent de telles régularités et d'autres pas » (Goodman)*

237

Quelles inductions (prévisions) faisons-nous ?

- Revenons sur la prévision des cours boursiers à partir des motifs
- Quelle est la longueur pertinente des motifs ?
- Plus l'on considère des motifs de grande taille, moins on aura d'observations (voire pas d'observations)
- Mais chaque motif va apporter davantage d'information (être plus caractéristique d'une trajectoire spécifique)
- Quelle est la longueur pertinente d'un motif pour la prévision ?
 - *Par exemple, HHH peut prévoir H et BHH peut prévoir B*
 - *Auquel cas, HH ne permet pas de prévoir*
- Pour gérer la double contrainte précédente, on peut être tenté de considérer des résumés de l'information passée

238

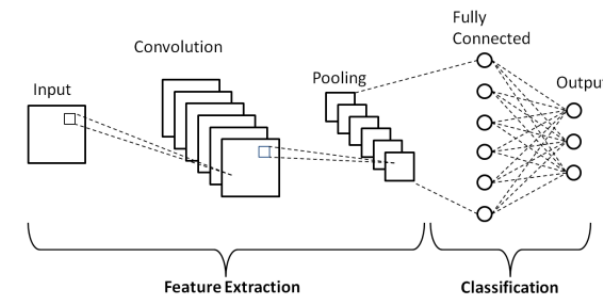
Quelles inductions (prévisions) faisons-nous ?

- Exemple simple : momentum - tendance
 - *On considère les 9 derniers cours boursiers*
 - *Si le nombre de hausses est au moins de 5, le marché est haussier et l'on prédit une hausse (approche d'élan ou momentum)*
 - *Si le nombre de baisses est au moins de 5, le marché est baissier et l'on prédit une baisse (toujours l'approche momentum)*
- Exemple simple : reversal
 - *On peut s'intéresser à la longueur moyenne des séquences de hausse ou de baisse (runs) ou au nombre de cas où une hausse est suivie d'une baisse pour caractériser un marché cyclique.*
- En faisant cela, on va compresser l'information contenue dans les cours passés

239

Prévisions boursières : encodage = angle mort de l'IA

- L'encodage consiste à compresser l'information disponible
- Avec le minimum de pertes
- Pour pouvoir la traiter par des méthodes de prévision (réseaux de neurones) ou de classification

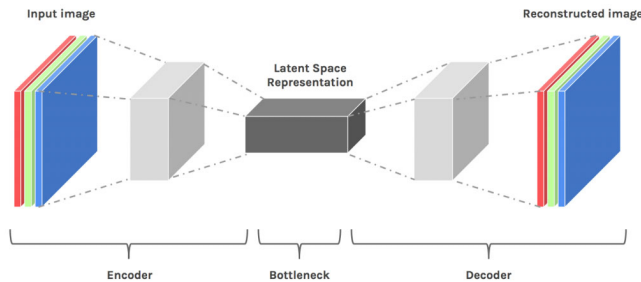


- L'encodage (feature extraction) permet une réduction de la dimensionalité des inputs

240

Prévisions boursières : encodage = angle mort de l'IA

- Encodage : réduction de la dimensionalité des inputs



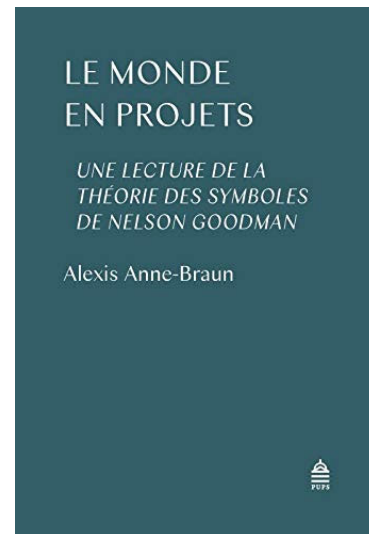
- Un encodage efficace permet non seulement de réduire la complexité
- Mais aussi de faciliter le traitement des données, une fois encodées : réduction des non-linéarités

241

Quelles inductions (prévisions) faisons-nous ?

- Notre manière de caractériser l'histoire des cours par des mots, « haussier » (bull), « baissier » (bear), de tendance (momentum), cyclique (reversal) n'est pas neutre
- Dire que le marché est haussier, c'est dire qu'il a une disposition à la hausse.
- C'est donc déjà faire une prédiction
 - Pour Ian Hacking (le plus pur nominalisme), « *les êtres humains classent et généralisent à la fois* »
 - « *Utiliser un nom pour une espèce, c'est vouloir réaliser des généralisations et former des anticipations concernant des individus de cette espèce* ».
 - « *Utiliser un prédicat, c'est lui faire jouer un rôle dans la formation des anticipations et l'établissement des généralisations* »

242



243

Quelles inductions (prévisions) faisons-nous ?

- Nous devons caractériser les représentations pertinentes des patterns boursiers
 - *Inputs de la prévision*
- Nous devons aussi nous intéresser à ce que l'on doit prévoir
 - *Supposons que le motif HHHH précède les motifs suivants : BHH, HBH.*
 - *Le motif HHHH ne peut pas être associé à une hausse ou bien à une baisse le jour suivant.*
 - *En revanche, on sait qu'il y aura deux hausses parmi les trois jours suivants, ce qui est suffisant pour gagner avec une stratégie buy and hold sur un horizon de trois jours.*

244

La « nouvelle énigme de l'induction » : Nelson Goodman



<https://www.youtube.com/watch?v=xZVVarMDso8>

245

Comment former des prévisions boursières ?

- Nelson Goodman apporte un éclairage nouveau au « problème de l'induction »
 - *Rôle fondamental du langage dans nos descriptions des phénomènes*
- Ce qui est régulier dans un langage ne l'est pas dans une autre et vice versa
 - *Comment compléter la suite 012345789 ?*
 - *101112131415 en système décimal... ABCDEF en hexadécimal*

246

Comment former des prévisions boursières ?

- Induction : 12345 ... peut se prolonger par ... 6789
 - *Réurrence $x_{n+1} = x_n + 1$ ou $x_{n+1} - x_n = 1$*
 - *On applique ici induction par énumération à la suite des accroissements*
 - *régularité généralisable ?*
- La notion de régularité généralisables pose un problème d'ordre logique
 - *Qu'induire à partir de 1235 ?*
 - *1235 peut se prolonger ou par 81321*
 - *Réurrence $x_{n+1} = x_n + x_{n-1}$ (suite de Fibonacci)*
 - *Ou par : 71113 (concaténation de la suite croissante des nombres premiers ou constante de Copeland-Erdős)*
 - *Ou par une suite croissante quelconque d'entiers ...*

247

Comment former des prévisions boursières ?

- Il y a donc plusieurs règles qui permettent d'extrapoler les données observées
 - *Laquelle choisir ? (problème soulevé par Saul Kripke)*
 - *On retrouve la question de la multiplicité des inférences inductives*
 - *Ici, plusieurs règles (donc plusieurs théories prédictives) : sous-détermination des théories par l'expérience (Quine)*
- Remarque : La notion de correspondances fortuites pose quant à elle des problèmes d'ordre psychologique.
 - *Ce qui est fortuit n'a pas d'interprétation logique ou causale.*
 - *Ceci se heurte aux biais du système I (machine explicative de Kahneman) et biais rétrospectif de Bischof*
 - *Voir séances sur la finance comportementale*

248

Comment former des prévisions boursières ?

- Autre exemple : Que dites-vous après 123 ?
 - (a) 123456 ...
 - Énumération de la liste des entiers naturels
 - $x_n = x_{n-1} + 1$
 - (b) 12357 ...
 - Début de la suite des nombres premiers
 - (c) 123123 ...
 - « Répéter 123 » ou bien $x_n = x_{n-1}[3] + 1$
 - (d) 123☀️😊
 - Grandir avec Nathan <https://grandiravec.nathan.fr/activites-famille/jeux/regles-jeux/1-2-3-soleil.html>

249

Comment former des prévisions boursières ?

- Observation de la série de nombres : 0010010101001
- Quelle est la prochaine valeur ?
 - La plupart des individus (et des algorithmes programmés par des individus) vont chercher si cela peut être 0 ou 1
- Univers des possibles donné par le contexte
 - Si on joue à pile ou face
 - et si on exclut que la pièce retombe sur la tranche
 - Et que l'on code par 0, pile et 1 face, la valeur suivante est 0 ou 1
 - En l'absence de contexte, on peut prolonger 0010010101001 par 2, a, &, π, « je ne sais pas » selon que l'on considère tel ou tel ensemble alphanumérique.

250

Comment former des prévisions boursières ?

- Comment compléter la suite 0123456789 ?
 - Comme c'est le début de la suite croissante des entiers naturels, on s'attend à la réponse 1011121314 ...
- Codage en hexadécimal (base 16) des nombres entiers.
 - Les « chiffres » en hexadécimal sont 0123456789ABCDEF
 - Codage standard en informatique car on passe facilement de la base $2^4 = 16$ à la base 2 utilisée par les ordinateurs
 - L'hexadécimal a une écriture plus compacte
 - Il est aussi utilisé dans le codage ASCII des caractères
 - 0000 correspond à 0, 0001 à 1, ..., 1110 correspond à E, 1111 correspond à F

251

Comment former des prévisions boursières ?

- Codage en hexadécimal

Écriture décimale	1763
Écriture binaire	11011100011
Écriture binaire en groupes de 4	0110 1110 0011
Écriture hexadécimale	6E3

- Pour un informaticien, dont la langue est l'hexadécimal, la suite 0123456789 se complète par ABCDEF101112131415 et non par 101112131415
- Celui qui égrène 0123456789 101112131415 a oublié ABCDEF
 - Dans le langage hexadécimal : beaucoup plus utilisé que le décimal, car qui fait encore des multiplications à la main ?!
 - C'est comme si en langage décimal, on faisait suivre 0123456789 par 16171819

252

Comment former des prévisions boursières ?

- L'irrégularité dans la prédiction est appelée une disjonction ou en langage managérial une disruption
 - *Ce qui est disjonctif dépend du langage utilisé*
 - 0123456789 101112131415 est disjonctif en hexadécimal
 - *Ce qui est régulier dépend du langage utilisé*
 - 0123456789 101112131415 est régulier en décimal
- Qu'en est-il de la vérification d'une hypothèse
 - *Les observations arrivent de manière séquentielle 0, 01, 012..., 0123456,*
 - *Ces observations confirment aussi bien l'hypothèse de la régularité décimale que de la régularité hexadécimale*
 - *Goodman (FFP p88) « n'importe quoi confirme n'importe quoi »*
 - *Problème méthodologique majeur*

253

Comment former des prévisions boursières ?

- Considérons la suite de caractères suivants :
 - *C, CC, CCC*
 - *On a envie de compléter le début de cette suite par CCCC*
 - *Or il serait tout aussi légitime de considérer la suite*
 - *C, CC, CCC, CD, D, DC, DCC, DCCC, CM, M*
 - *On ne voit pas de régularité apparaître dans les signes.*
 - *Si l'on se rend compte qu'il s'agit de chiffres romains et si on les exprime en chiffres arabes*
 - *100,200,300 précèdent 400,500,600,700,800,900,1000*

254

Comment former des prévisions boursières ?

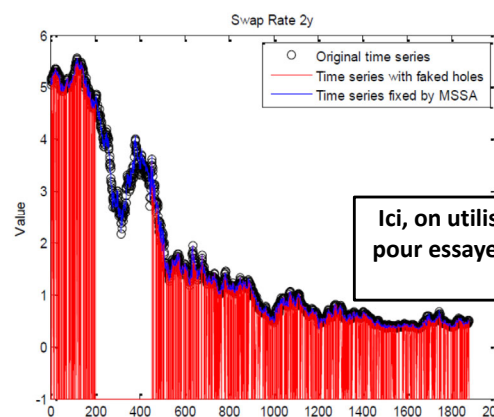
- Considérons la série
00
- *On aimerait la prolonger par un 0, car on n'a observé que des 0*
- 11111111101111111111111111101111011111111111111
- *On aurait tendance à la prolonger par un 1 car la fréquence d'apparition des 1 est plus élevée*
- 00001100111100110011111100110000011001111
- *On aurait tendance à la prolonger par un 1 car la probabilité d'une répétition est supérieure à $\frac{1}{2}$*
- Pour Goodman, on pourrait prolonger toutes ces séries par 0 ou 1, car il n'y a pas de contexte culturel ou de langage penchant vers une généralisation particulière

255

Comment former des prévisions boursières ?

- Capacité de l'algorithme utilisé par Bloomberg LP (MSSA: Multichannel Singular Spectrum Analysis) à compléter des données

■ Source P. Dos Santos



$$A = (1 \ 2 \ 3 \ 4 \ 5 \ 6 \ 7 \ 8.81 \ 10.83)$$

$$C = [1 \ 0 \ 1 \ 0 \ 1 \ 0 \ 1 \ 0 \ 1]$$

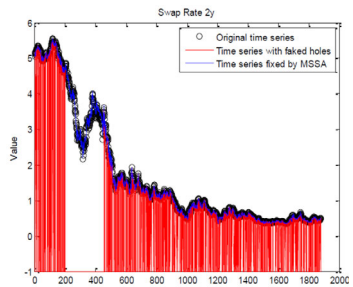
$$D = [1 \ 1 \ 0 \ 0 \ 1 \ 1 \ 0 \ 0.15 \ 0.07]$$

Ici, on utilise aussi des données futures, pour essayer de déterminer des données manquantes

256

Comment former des prévisions boursières ?

- Goodman établit qu'il n'y a aucune règle logique qui permette de choisir une généralisation d'une propriété dans un échantillon à la population, plutôt qu'une autre.
- Toutes les règles d'inférence inductive sont logiquement équivalentes.
- Reprenons le cas des données manquantes :



Une approche « raisonnable » et couramment utilisée est de recourir à une interpolation linéaire ou d'autres méthodes qui visent à avoir une trajectoire régulière (voir l'article, Goodman, Poincaré et la nouvelle énigme de l'induction). Mais dans un contexte de mesure des risques, une interpolation lisse va sous-estimer le risque. Il n'y a pas d'implantation « naturelle ».

257

258

259

260

Induction et langage



- Induction et ambiguïté du langage
- Le double sens de l'énoncé « le cygne est blanc »
 - « ce cygne est blanc » a un caractère singulier, blanc étant un prédicat (attribut du sujet) de ce cygne.
 - Il peut aussi énoncer une loi « tous les cygnes sont blancs », qui établit une relation permanente entre la blancheur et le cygne. On est passé de la constatation à la prédiction

261

Induction : ambiguïté du langage et cygnes blancs

- Si l'induction, ou passage du particulier/spécifique au général, a des bases animales, ses caractéristiques humaines sont spécifiques
- Du fait de l'**utilisation du langage**.
- Nous allons travailler à partir d'un exemple simple.
- Considérons la proposition « le cygne est blanc ».



262

Induction et langage

- Remarque : « cygne » est un concept qui dénote un objet particulier ou un ensemble délimité (ou pas) d'objets (l'espèce cygne).
 - A propos de la perception de l'objet cygne et de sa réalité ou non (approche nominaliste), nous suivons l'approche de W. James et de H. Putman, appelée « réalisme naturel »
- « est blanc » exprime une propriété du cygne : c'est un **prédicat nominal**
- La proposition « le cygne est blanc » attribue la propriété d'être blanc au cygne.
- « est » un verbe attributif ou substantif, « blanc » est un adjectif (attributif).

263

Possibles et propositions réfutables

- Remarque : même si l'on n'a observé que des cygnes blancs, la proposition « le cygne est bleu » est logiquement et syntaxiquement valide.
 - Mais elle devient fausse dès que l'on a observé un cygne blanc.
- L'observation d'un cygne non-blanc rend la proposition « les cygnes sont blancs » fausse.
 - On dit que cette proposition est **réfutable**.
 - Remarque : toute proposition à caractère général est réfutable, ne peut donc avoir le caractère de loi universelle.
- Le langage permettant de construire des propositions « les cygnes sont blancs ou » « les cygnes sont bleus » permet de concevoir un ensemble de possibles.
 - La pluralité des possibles est le germe de l'aléatoire.

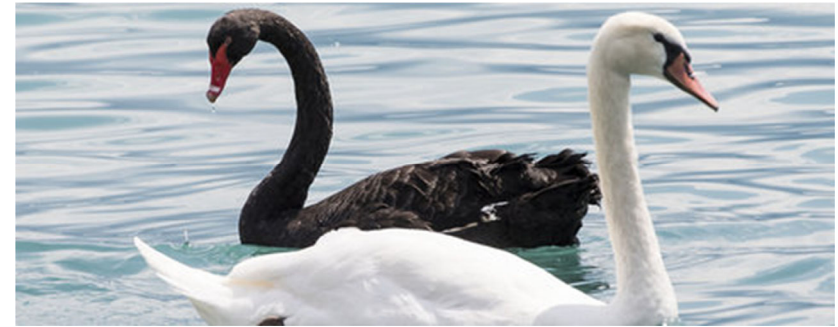
264

Induction et langage

- La proposition « le cygne est blanc » peut avoir deux sens
 - *Sens 1 : le cygne particulier (ce cygne) devant lequel je me trouve est blanc*
 - *Sens 2 : les cygnes en général (catégorie d'oiseaux) sont blancs.*
- La première proposition est de nature descriptive.
- La seconde proposition a une portée générale.
- Selon le contexte, on saura (en principe !) si le locuteur utilise le sens 1 ou 2.
 - *Mais l'énoncé est identique et l'on peut passer insensiblement du premier au second sens.*
 - *L'expérience ne peut être un moyen de preuve de la proposition générale : ce n'est pas parce que l'on a observé que des cygnes blancs que tous les cygnes sont blancs*

265

cygne noir et cygne blanc



- Pour mémoire, jusqu'au 18^e siècle, on ignorait l'existence de cygnes noirs, « apparus » en Australie
- *Un oiseau rare dans le pays, rare comme un **cygne noir***
 - « *rara avis in terris nigroque simillima cygno* »
 - *Le cygne noir, expression qui remonte à l'auteur latin Juvénal (fin du I^{er} siècle) est un oxymore*

266

Induction : ambiguïté du langage et cygnes blancs

- Considérons la proposition « le cygne est blanc ».



- En tant que proposition logique, elle peut prendre deux valeurs de vérité : vrai ou faux.
- Ici, la proposition (ou énoncé d'observation) est vraie.
- La photographie dénote (renvoie à) un cygne qui est blanc.
- Il s'agit de ce cygne spécifique : cas particulier (singulier)

267

Induction : ambiguïté du langage et cygnes blancs

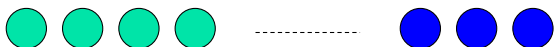
- Considérons à nouveau la proposition « le cygne est blanc ».
- Le sujet de la proposition « le cygne » dénote-t-il le cygne particulier dont nous venons de voir la photo (à un instant donné) ou l'ensemble de tous les cygnes ?
 - *La couleur blanche n'est plus seulement une caractéristique du cygne photographié, mais un **attribut** de l'espèce*
- Dans le second cas, la proposition s'applique à l'ensemble de tous les cygnes, ceux que l'on a observés et ceux que l'on n'a pas encore observés.
- Ce n'est plus l'énoncé d'une propriété particulière, mais une loi qui a une portée générale.

268

La « nouvelle énigme de l'induction »

■ Le paradoxe Vleu de Goodman

- Vleu : contraction de Vert et Bleu:
- On considère un ensemble de boules placées en file indienne.



- Les premières sont de couleur verte, les dernières sont de couleur bleue.
- Pour simplifier, on dira que toutes les boules sont vleues.
- Considérons un expérimentateur qui observe d'abord la première boule, puis la seconde.
- Il peut énoncer que les boules sont vertes. Il pourrait aussi dire que les boules sont vleues

269

La « nouvelle énigme de l'induction » Le paradoxe Vleu de Goodman

- Vleu est bien défini
- L'expérimentateur (qui n'est pas encore arrivé aux boules bleues) peut énoncer deux hypothèses :
 - H0 : les boules sont vertes
 - H1 : les boules sont vleues
- Les deux hypothèses sont compatibles avec les observations
 - Chaque nouvelle observation vient aussi bien confirmer H1 que H0
- « est vleu » est un prédicat bizarre par rapport à « est vert » !
- Goodman examine pourquoi cela nous paraît bizarre.
- Renversons l'ordre dans lequel on examine les boules



- En allant de la droite vers la gauche, on peut dire que les boules sont bleues ou qu'elles sont « blertes »

270

La « nouvelle énigme de l'induction » Le paradoxe Vleu de Goodman

- Analysons l'idée du « prédicat disjonctif »
- « est vert » ne serait-il pas disjonctif par rapport à « est vleu » ?
- En rupture avec l'idée d'uniformité de la nature souvent avancée pour justifier l'induction.
- En fait, si l'on considère comme couleur primaires vleu et blert, on peut dire qu'une série de boules vertes est de couleur vleue jusqu'à un certain rang, puis blerte ensuite
- Dans ce système, c'est « est vert » qui est disruptif
- Donc, la question que l'on doit se poser est : **Pourquoi utilisons-nous vert et bleu comme couleurs primaires dans le langage ordinaire, plutôt que « vleu » et « blert » ?**
- Pour Goodman, « **les racines de la validité inductive se trouvent dans notre façon d'utiliser le langage** » (FFP 124).

271

La « nouvelle énigme de l'induction » Le paradoxe Vleu de Goodman

- Ce qui est régulier dans un langage ne l'est pas dans un autre et vice-versa
- Goodman, contrairement à Hume, n'aborde pas le problème de l'induction comme une habitude psychologique contestable
 - Une « proposition reposante dans l'âme » selon le théologien franciscain Jean Duns Scot
- **C'est le concept même de régularité qui est remis en cause**
- **On utilise vert et bleu parce que ces termes sont ancrés (implantés) dans le français et qu'ils façonnent notre perception.**
- Goodman considère qu'une inférence inductive est **valide** si elle est couramment implantée
- Justification très différente de celle de Mill (uniformité de la nature)
- Approche de Goodman : « nominaliste » (en rapport avec les noms)

272

Régularités et impermanence

- L'illusion de la permanence est entretenue par l'**ambiguïté du langage**
- Si on énonce "le saphir est bleu", on ne prend pas en compte le changement du bleu au violet selon la lumière qui éclaire le saphir
- « Bleu » n'est pas un attribut général et permanent du saphir.
 - « Ce saphir est bleu si on l'éclaire à la lumière du jour ».
 - Existence par conditions : caractère relatif



273

La « nouvelle énigme de l'induction » Le paradoxe Vleu de Goodman

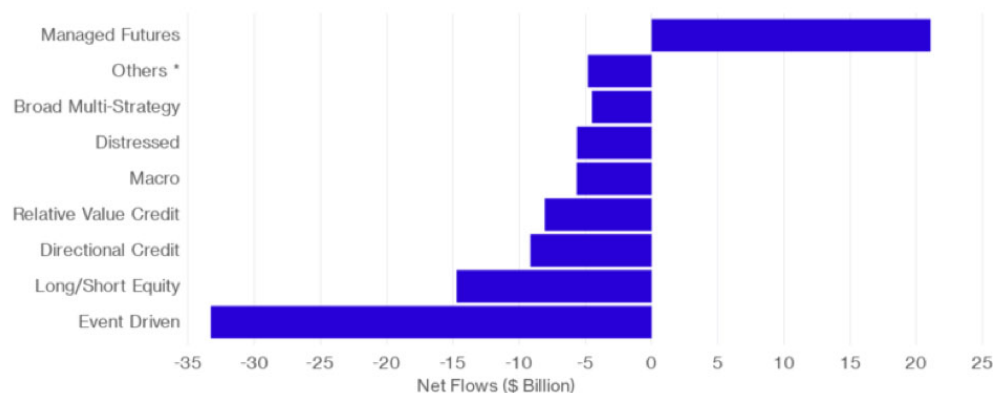
- Est-ce que les dénominations des couleurs sont uniquement affaire de langage ?
 - Connaissez-vous une langue qui utilise les couleurs vleues et blertes ?
 - Pouvons-nous apprendre à un singe le concept de couleur vleue ?
 - Goodman n'omet-il pas les bases animales de l'induction ?
- Mais les remarques de Goodman sont fondamentales pour une bonne méthodologie des prévisions boursières
 - Pas de langage naturellement pertinent de représentation (encodage) des données boursières et financières
 - Les données brutes (quantitatives, qualitatives, textuelles) sont « encodées », c'est-à-dire transformées en inputs (ratios financiers, ...) directement utilisables
 - Un réseau de neurones ne peut traiter qu'un nombre limité d'inputs

274

275

276

Computer-Driven Hedge Funds Rule



Source: eVestment

*Includes MBS strategies, market neutral, convertible arbitrage

Bloomberg

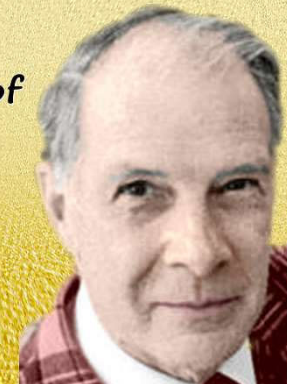
En 2016, les hedge funds gérés selon des stratégies “systématiques” (rule driven) ont collecté 21 milliards de \$. Les hedge funds gérés de manière discrétionnaire ont subi des retraits de 60 milliards de \$.

277

Peut-on investir avec succès à partir de règles « data-driven » ?

Mathematics is an interesting intellectual sport but it should not be allowed to stand in the way of obtaining sensible information about physical processes.

financial
— Richard Hamming



Richard Hamming, prix Turing 1968

278

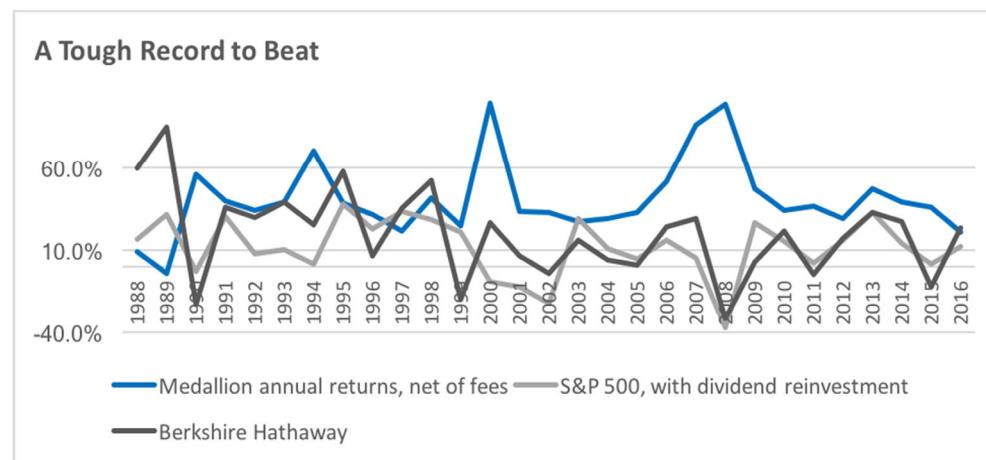


Jim Simons, mathématicien, patron de Renaissance Technologies : des stratégies d'investissement à partir d'analyses statistiques, de (très) bonnes données.

Une performance annuelle moyenne de 72% depuis le lancement du hedge fund, il y a 29 ans

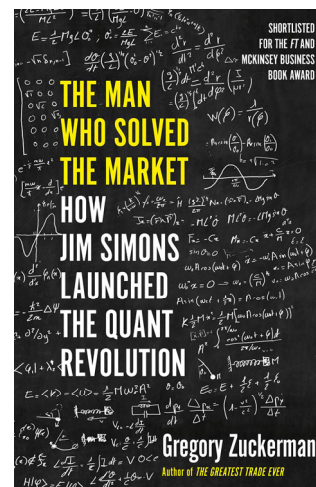
279

Best in class: rentabilités du fonds Medallion, de Berkshire Hathaway et du S&P500



280

Les bureaux de Renaissance Technologies à Stony Brook, près de l'Université



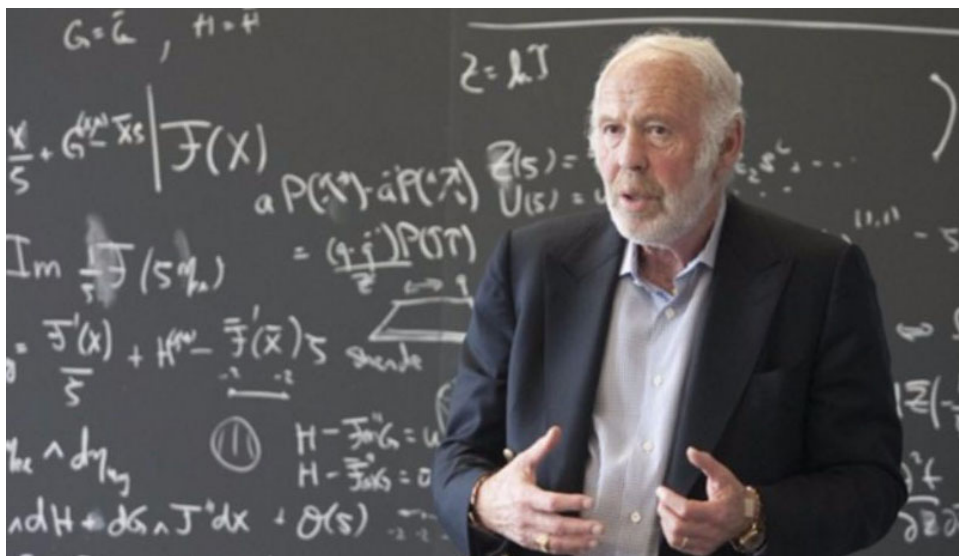
la probabilité que ce soit aléatoire n'est pas élevée.

Entretien avec Jim Simons (sous-titré en français !)
The mathematician who cracked Wall Street

<https://www.youtube.com/watch?v=U5kldtMJGc8>

In this fast-paced narrative, Zuckerman examines how Simons launched a quantitative revolution on Wall Street, and reveals the impact that Simons, the quiet billionaire king of the quants, has had on worlds well beyond finance.

282

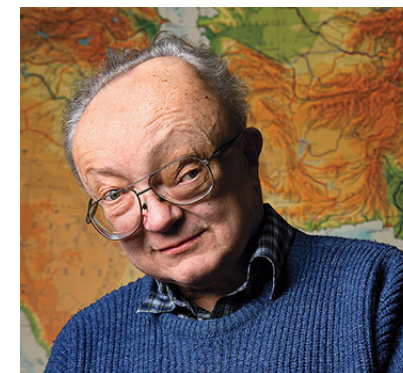


<https://awealthofcommonsense.com/2019/11/non-intuitive-lessons-from-the-man-who-solved-the-market/>

283

"LTCM's basic error was believing its models were truth, We never believed our models reflected reality — just some aspects of reality." Nick Patterson, RT partner.

<https://awealthofcommonsense.com/2019/11/non-intuitive-lessons-from-the-man-who-solved-the-market/>



[https://en.wikipedia.org/wiki/Nick_Patterson_\(scientist\)](https://en.wikipedia.org/wiki/Nick_Patterson_(scientist))

"Some signals that make no intuitive sense do indeed work". Robert Mercer, ancient co-PDG de RT et personnage par ailleurs très controversé...

Questions pertinentes, relatives à la modélisation statistique et économique



284

Long Term Capital Management LTCM

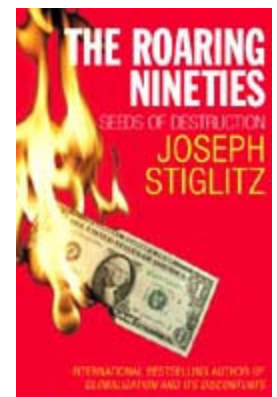
- Présentation du fonds
- Les protagonistes
- Les stratégies
 - *Convergence trades*
 - *Paris sur le retour à la moyenne des spreads*
- Communication financière
 - *probabilités de gains / pertes*
- 1998 : la chute
 - *Divergences : faillite de l'État Russe*
 - *Sauvetage par la Fed*



285

Long Term Capital Management LTCM

- Création en 1994
- Plus grand hedge fund des années 1990
- Investissements sur les marchés jusqu'à 1280 milliards de \$
 - *Surtout hors-bilan*
 - *Utilisation de produits dérivés*
 - *Très peu de fonds propres (fort effet de levier)*
- Performances les premières années : +21%, +43%, +41%



The Roaring Nineties
Seeds of Destruction
Joseph Stiglitz



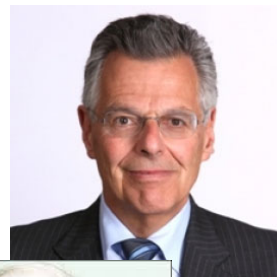
286

Long Term Capital Management LTCM

- Les principaux protagonistes



John Meriwether, Bob Merton, Myron Scholes, Alan Greenspan



287

La stratégie de base du fonds



- Jouer sur les écarts anormaux entre les cours de deux titres proches (basis risk) et parier sur un retour à la moyenne
 - *Relative value, spread betting, pair trading, convergence trades*
 - *Hypothèse de stationnarité des spreads*
- Communication financière auprès des investisseurs utilisant largement le recours aux probabilités
 - *Probabilité de perdre 5% sur un horizon de placement d'un an = 12%*
 - *Probabilité de perdre 20% sur un horizon de placement d'un an = 2%*
 - *Approche « scientifique »...*

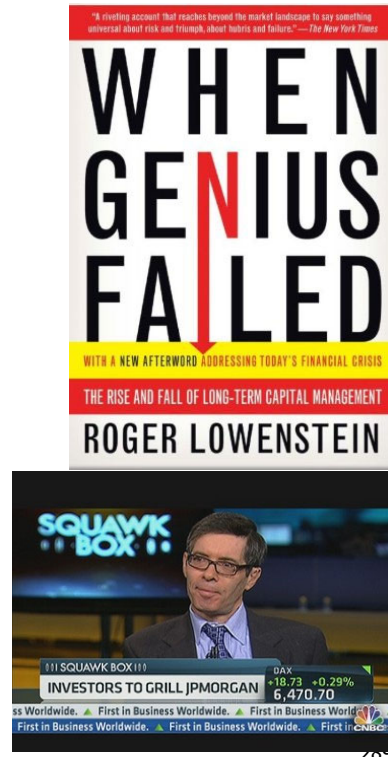
288

Résultat des courses

Après un début en fanfare, faillite rapide et retentissante ayant menacé la stabilité du système financier

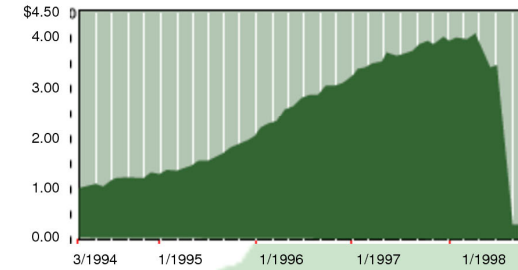
Organisation d'un sauvetage coûteux pour les grandes banques sous l'égide de la Fed

Au lieu de revenir vers leur niveau normal, les spreads avaient divergé



209

Long-Term Capital Management



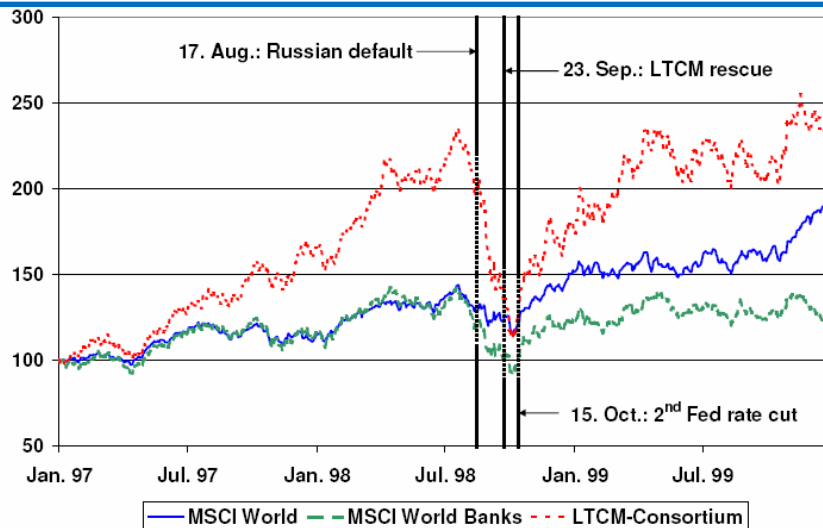
Gross value of \$1 invested March 1994 - October 1998 in LTCM
When Genius Failed pp. xiv

The company used complex mathematical models to take advantage of fixed income arbitrage deals (termed convergence trades) usually with U.S., Japanese, and European government bonds.

290

17 août 1998, défaut de la Russie, 23 Septembre bail-out de LTCM, 15 octobre nouvelle intervention de la Fed

Dans les « modèles statistiques », la baisse des prix des titres russes était anormale, et pourquoi le risque économique et financier, bien facile à appréhender



291

The Fed effect

Share prices
Jan 1st 1998=100

Fed funds rate %

Source: Primark Datastream

LTCM

The financial technology company

INVENTING MONEY

The story of Long-Term Capital Management and the legends behind it

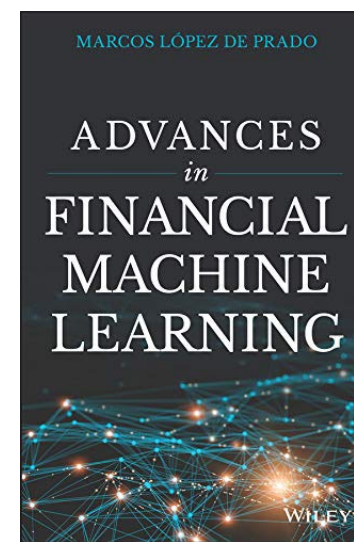
NICHOLAS DUNBAR

1998	2008
LTCM, JOHN MERIWETHER'S HEDGE FUND, COLLAPSES	INVESTORS IN MERIWETHER'S JWM JITTERY AFTER BOND FUND TUMBLES 28PC
THE PIONEER OF QUANTITATIVE INVESTMENT LOST \$4.6BN IN 4 MONTHS	REJECTS LTCM COMPARISONS - FUND 'HAS LESS RISK AND MORE LIQUIDITY'
TRIGGERED A FED-ORGANISED BAIL-OUT BY WALL STREET BANKS	SOME INVESTORS WANT OUT DESPITE THE FUND'S REASSURANCES
UNDONE BY MASSIVE LEVERAGE - UP TO \$50 FOR EVERY DOLLAR INVESTED	INVESTORS HAVE UNTIL MONDAY TO REQUEST WITHDRAWALS

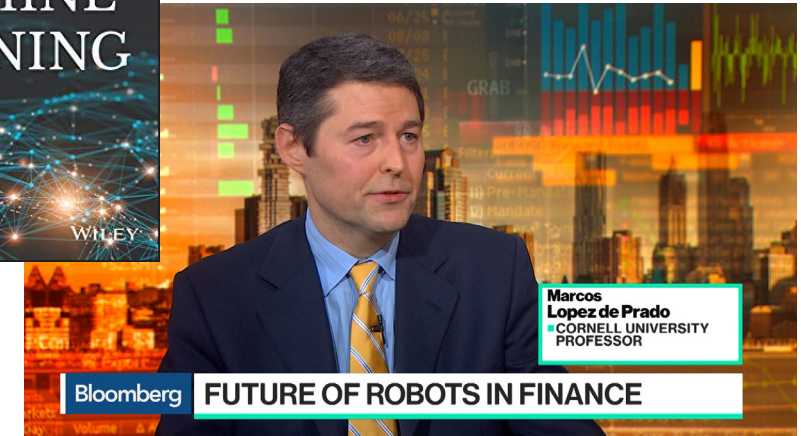
Statistique et stratégies d'investissement

- Quel est le principe des stratégies systématiques / quantitatives d'investissement boursier ?
- Quelles en sont les limites ?
 - Critiques de la méthode inductive (Hume - Goodman)
 - Encodage des entrants (inputs)
 - Validation des stratégies
- Quels sont les problèmes liés à la validation des stratégies d'investissement (backtesting) ?
 - Spurious correlations : corrélations fictives
 - Validation ex-post (data snooping)
 - Données aberrantes (outliers)
 - Pas de méthode expérimentale (! au price impact)

293



Marcos Lopez de Prado, Professeur de « finance professionnelle » (financial practice) – Cornell & responsable de la recherche ADIA - Abu Dhabi Investment Authority



294

Validation des stratégies : backtesting

- Principe d'un backtest : simulation d'une stratégie d'investissement sur un historique de données

The Dangers of Backtesting

11.1 MOTIVATION

Backtesting is one of the most essential, and yet least understood, techniques in the quant arsenal. A common misunderstanding is to think of backtesting as a research tool. Researching and backtesting is like drinking and driving. Do not research under the influence of a backtest.

295

Validation des stratégies : usual suspects, seven sins

1. **Survivorship bias:** Using as investment universe the current one, hence ignoring that some companies went bankrupt and securities were delisted along the way.
2. **Look-ahead bias:** Using information that was not public at the moment the simulated decision would have been made. Be certain about the timestamp for each data point. Take into account release dates, distribution delays, and backfill corrections.
3. **Storytelling:** Making up a story *ex-post* to justify some random pattern.
4. **Data mining and data snooping:** Training the model on the testing set.
5. **Transaction costs:** Simulating transaction costs is hard because the only way to be certain about that cost would have been to interact with the trading book (i.e., to do the actual trade).
6. **Outliers:** Basing a strategy on a few extreme outcomes that may never happen again as observed in the past.
7. **Shorting:** Taking a short position on cash products requires finding a lender. The cost of lending and the amount available is generally unknown, and depends on relations, inventory, relative demand, etc.

296

Seven sins of quant investing
Survivorship bias
Look-ahead bias
Story telling
Data mining and snooping
Signal decay and turnover
Outliers and data normalization
Asymmetric pattern



Deutsche Bank
Markets Research

Global

Quantitative Strategy
Signal Processing

Seven Sins of Quantitative Investing

- Survivorship bias
- Look-ahead bias
- The sin of story telling
- Data mining and data snooping bias
- Signal decay and turnover
- Outliers – the story of spectacular successes and failures
- The asymmetric payoff pattern and shorting cost

297

298

299

300