

Modélisation de l'incertitude — Master 2 MIAGE IA²

Travaux dirigés N° 1 : Probabilités

Andrea G. B. Tettamanzi
Université côte d'Azur
andrea.tettamanzi@univ-cotedazur.fr

Année universitaire 2020/2021

Résumé

On va calculer, en utilisant le langage Python, des probabilités simples, conjointes et conditionnelles. On basera ces calculs sur des données réelles issues des marchés financiers.

1 Introduction

On ne pourrait pas penser à quelque chose de plus incertain que les marchés financiers. C'est la raison pour laquelle on va étudier les mouvements relatifs d'un certain nombre d'indices de ces marchés, notamment :

- CAC 40 (`^FCHI`), l'indice de la Bourse de Paris ;
- DAX 30 (`^GDAXI`), l'indice de la Bourse de Francfort ;
- FTSE 100 (`^FTSE`), l'indice de la Bourse de Londres ;
- Dow Jones Industrial Average (`^DJI`), l'indice de la bourse de New York ;
- NASDAQ Composite (`^IXIC`), l'indice du marché des actions des sociétés *hi-tech*.

Les codes entre parenthèses sont appelés *tickers* et permettent de consulter la valeur de ces indices auprès des fournisseurs de données financières.

2 Consignes

1. On commence par récupérer les données brutes en ligne. Les séries historiques des indices qui nous intéressent peuvent être téléchargées au format CSV depuis Yahoo! Finance. Comme vous le pourrez constater, les données sont ordonnées de la date la plus récente à celle la plus éloignée.

Pour chaque jour d'ouverture du marché où l'action est traitée, vous avez six champs (ou colonnes), qui sont appelés, dans le fichier CSV : Date, Open, High, Low, Close, Volume et Adj Close. Ce dernier contient le prix de clôture (le même que Close), ajusté pour les dividendes, les fractionnements ou les regroupements d'actions.

Téléchargez les données brutes pour la dernière année an par an (depuis la même date qu'aujourd'hui l'année dernière jusqu'à hier : c'est l'option par défaut).

2. Ne prenez en compte que la colonne Adj Close ; vous pouvez ignorer les autres. Remplissez les données manquantes en recopiant à chaque fois la donnée disponible la plus récente.
3. Définissez une série discrète sur la base de chacune des séries ainsi obtenues en utilisant la transformation ci-dessous. Soit x_t la valeur d'Adj Close pour la période $t = \{0, 1, \dots\}$;

on va calculer le terme de la série discrète pour la période $t = \{1, \dots\}$, s_t , comme suit :

$$s_t = \begin{cases} +1, & \ln \frac{x_t}{x_{t-1}} > 0.005; \\ 0, & \left| \ln \frac{x_t}{x_{t-1}} \right| \leq 0.005; \\ -1, & \ln \frac{x_t}{x_{t-1}} < -0.005; \end{cases} \quad (1)$$

En gros, +1 signifie que la valeur de l'indice a augmenté, -1 qu'elle a diminué et 0 qu'il n'y a pas eu de mouvement digne de note, juste un petit ajustement.

4. Obtenez la table des probabilités des s_t pour chacun des indices considérés.
5. Calculez, pour chaque indice : $\Pr(s_t, s_{t+1})$, $\Pr(s_t | s_{t-1})$ et $\Pr(s_{t+1} | s_t)$.
6. Calculez, pour chaque couple d'indices $s^{(i)}$, $s^{(j)}$, $\Pr(s_t^{(i)}, s_t^{(j)})$ et $\Pr(s_t^{(i)} | s_t^{(j)})$.
7. Calculez, pour chaque indice $s^{(i)}$,
 - (a) $\Pr(s_t^{(i)} | s_t^{(1)}, \dots, s_t^{(i-1)}, s_t^{(i+1)}, \dots, s_t^{(n)})$,
 - (b) $\Pr(s_t^{(i)} | s_{t-1}^{(1)}, \dots, s_{t-1}^{(n)})$.
8. Qu'est-ce que vous pouvez déduire en comparant ces probabilités ?
9. Comment pensez-vous qu'un agent intelligent pourrait les utiliser pour gagner de l'argent en achetant et vendant ces indices ?

Rendez votre code et vos observations dans un archive zippé par courriel.