DM - Master Pro

Tran Viet Chi, `chi.tran@math.univ-lille1.fr`, bureau 316 (bâtiment M3).

Rque : j'attends des commentaires pertinents et pas une simple obtention des chiffres demandés. Toute remarque bienvenue sera appréciée, même si elle n'était pas exigée dans la question !

Exercice 1 (Etude d'une série financière avec SAS)
Le CAC40, qui prend son nom du système de *Cotation Assistée en Continu*, est le principal indice boursier de la place de Paris. Créé avec 1 000 points de base au 31 décembre 1987 par la Compagnie des agents de change, il est déterminé à partir des cours de quarante actions cotées en continu parmi les cent sociétés dont les échanges sont les plus abondants sur Euronext Paris. Ces actions, représentatives des différentes branches d'activités, reflètent en principe la tendance globale de l'économie des grandes entreprises françaises et leur liste est revue régulièrement pour maintenir cette représentativité. Le CAC40 est publié du lundi au vendredi de 9 h 00 à 17 h 30 et mis à jour toutes les 15 secondes.
Les données du fichier `cac40.sas7bdat` sont disponibles à l'adresse
`http://math.univ-lille1.fr/~tran/enseignements.html`.

Partie A : Description générale des données

1. Tracer graphiquement, et sur un même graphique les évolutions des cours en ouverture, en cloture, les maxima et minima. On pourra préalablement créer une variable t indiquant le numéro de l'observation.

2. A l'aide d'une `PROC UNIVARIATE`, sortir des statistiques résumant la distribution de l'indice. La réponse sera présentée sous forme d'un paragraphe de quelques lignes où seules les informations pertinentes seront mises en valeur.

3. Tracer l'histogramme de la variable `ouverture`.

4. Peut-on supposer que cette variable suit une loi log-normale ? Détaillez le test et expliquez comment votre conclusion est tirée.

Partie B : Etude des rendements

Pour un cours $(x_t; t \in \{1, \dots, T\})$, on définit les rendements par :

$$r_t = \log\big(\frac{x_t}{x_{t-1}}\big), \qquad \text{rendements logarithmiques}$$
$$\rho_t = \frac{x_t - x_{t-1}}{x_{t-1}}, \qquad \text{rendements arithmétiques.}$$

5. Exprimer pour chacun de ces rendements le cours x_t au temps t en fonction de x_0 et des valeurs des rendements aux temps intermédiaires $(r_s; s \in \{0, \dots, t\})$ et $(\rho_s; s \in \{0, \dots, t\})$.

6. Calculer les rendements logarithmiques et arithmétiques pour le cours `ouverture`.

7. Tracer l'évolution de ces rendements au cours du temps. Que pensez-vous de ces séries en terme de stationnarité ?

8. Décrire les distributions de ces rendements. Encore une fois, la réponse sera sous forme d'un paragraphe de quelques lignes où seules les informations pertinentes seront mises en valeur.

9. Les rendements sont-ils gaussiens ? Ceci était-il prévisible, en particulier pour le rendement logarithmique ?

10. Comparez les deux séries de rendements.

Partie C : Autres indicateurs financiers

11. Pour un cours $(x_t; t \in \{1, \dots, T\})$, on définit les indicateurs financiers suivants qu'on vous demande de calculer et d'analyser (en quelques brèves lignes) :

1. le rendement annualisé est obtenu à partir des rendements arithmétiques mensuels ρ_t, en utilisant le caractère géométrique obtenu à la question **5.**
$$\rho_{an} = (1 + \bar{\rho})^{12} - 1$$
où $\bar{\rho}$ est le rendement mensuel moyen.
2. le *drawdown* (perte maximale historique) donne une indication sur la perte depuis le dernier pic : au temps t, il est donné par
$$D_t = \frac{\max_{0 \le s \le t} x_s}{x_t}.$$
3. Soit $\alpha \in [0, 1]$. La *Value at Risk* (VaR) est le quantile $1 - \alpha$ de la distribution des rendements arithmétiques, c'est-à-dire la valeur telle que : $\alpha = \mathbb{P}(r > \text{VaR}(\alpha))$.

□

Exercice 2 (Tests du générateur uniforme de R, d'après le partiel 2008)
Le but de cet exercice est d'étudier un test du générateur de nombres pseudo-aléatoires de **R**. On attend de ce générateur qu'il retourne des suites $X_1, \dots X_N$ de variables aléatoires uniformes statistiquement indépendantes.

1. Générer un vecteur de $N = 1000$ variables aléatoires uniformes indépendantes. Tracer l'histogramme associé aux simulations $X_1, \dots X_N$.

2. Compter le nombre n_0 de valeurs dans l'intervalle $[0, 0.1[$, le nombre n_1 de valeurs dans l'intervalle $[0.1, 0.2[$, ..., le nombre n_9 de valeurs dans l'intervalle $[0.9, 1]$. Faire un graphique et commenter.

3. Test des paires en série : On veut s'assurer qu'il n'y a pas de lien statistique entre deux nombres pseudo-aléatoires générés consécutivement. Représenter le nuage de points (X_i, X_{i+1}) pour $i \in \{1, \dots, 999\}$. Commenter.

4. Nous nous proposons de tester l'absence de liaison statistique entre X_i et X_{i+1}. Pour cela, on sépare l'intervalle $[0, 1]$ en trois : $C_1 = [0, 0.25[$, $C_2 = [0.25, 0.5[$, $C_3 = [0.5, 1]$. On note par $n_{k\ell}$ (pour $k, \ell \in \{1, 2, 3\}$) le nombre de points (X_i, X_{i+1}) tels que $X_i \in C_k$ et $X_{i+1} \in C_\ell$. Etablir le tableau de contingence $T = (n_{k\ell})_{k,\ell \in \{1,2,3\}}$.

4.1. Enoncer l'hypothèse nulle H_0 que l'on cherche à tester.

4.2. Calculer les fréquences marginales.

4.3. Calculer la distribution empirique de X_{i+1} conditionnellement à $X_i \in C_k$, pour $k \in \{1, 2, 3\}$ et tracer les histogrammes de ces distributions. Comparer aux résultats de la question **4.2** et commenter.

4.4. Réaliser le test du χ^2 avec la fonction `chisq.test`.

4.5. Recalculer sans la fonction `chisq.test`, à partir du tableau de contingence T, la statistique de test ξ.

4.6. Quelle est la loi de ξ quand $n \to +\infty$ lorsque H_0 est vraie ? lorsque H_0 est fausse ?

4.7. Quelle est la valeur de `pchisq(xi,4)` ? Justifier comment on peut trouver la conclusion du test à partir de cette valeur, si l'on veut que sous H_0 la probabilité d'erreur soit inférieure à $\alpha = 5\%$?

5. A partir des fréquences marginales pour X_i obtenues à la question **4.2**, calculer une approximation de l'espérance et de la variance des X_i. □