UNIVERSITÉ PARIS 1
PANTHÉON SORBONNE

Projet de gestion des risques

Auteurs : Yanis Ahdjoudj, Xenia Zaricinii, El Mehdi Agunaou, Lucas Diaz

Résumé

Notre devoir se veut de proposer un portefeuille d'actions intégrant progressivement divers indicateurs de risque dans le critère de sélection. Nous commençons ainsi avec une sélection naïve utilisant uniquement une espérance de rendement. Puis ce rendement est ajusté pour prendre en compte successivement un risque de modèle, le risque de marché et le risque de liquidité.

L'ensemble de notre travail est accessible via le repository GitHub suivant : `https://github.com/YanisAhdjoudj/Projet_RiskManagement`

Gestion des risques
Master 2 MoSEF
Avril 2021

Table des matières

Chapitre 1

Sélection des actions

Pour notre projet nous devons choisir 20 actions , nous décidons de prendre uniquement des actions américaines car cela nous évite d'avoir à géré le risque de change (en nous considérant comme des investisseurs américains) de plus elles sont plus nombreuses et diversifiées ce qui nous offre une plus grande liberté de choix

Afin d'avoir un portefeuille diversifié nous décidons de choisir des actions de différents secteurs :

— Information Technology : Google(GOOGL), Facebook (FB), Apple(AAPL), Amazon(AMZN)
— Retail : Walmart(WMT), Costco (COST)
— Food : The Coca-Cola Company (KO), McDonald's(MCD)
— Financials : BlackRock(BLK), JPMorgan Chase Co(JPM), Berkshire Hathaway(BRK.A)
— Industrials : Builders FirstSource Inc (BLDR), General Motors(GM), Tesla(TSLA), Union Pacific(UNP)
— Healthcare and Biotech : Pfizer(PFE), Heron Therapeutics(HRTX)
— Energy : ExxonMobil(XOM), Chevron(CVX), Plug Power(PLUG)

Les prix récupéré sont les prix à l'ouverture du 1 janvier 2017 au 1 er janvier 2021. L'implémentation est faite sous python en récupérant les données de yahoo finance.

Chapitre 2

Rendement espéré

Question 1 . *A partir d'un historique d'observations, comment peut-on déterminer le paramètre de tendance le plus vraisemblable ?*

Réponse 1 : Pour estimer le paramètre de tendance le plus vraisemblable il nous semble pertinent d'utiliser la moyenne des rendement comme estimateur de μ. Ainsi pour chaque action i le paramètre de tendance est déterminé par :

$$\mu = \frac{1}{T}\sum_{t=1}^{T}(r_t)$$

Question 2 . *Quel est le paramètre de tendance de chacune des 20 actions ? Indiquez par ailleurs la significativité de chacun de ces paramètres (l'hypothèse nulle du test statistique demandé est une tendance nulle (vous pouvez estimer la p-value de chaque tendance trouvée à l'aide d'un test bien choisi et dont vous justifiez le choix, ou bien à l'aide d'un bootstrap)*

Réponse 2 : Pour déterminé la significativité du paramètre de tendance nous avons utiliser un test de student. Sous l'hypothèse nulle la tendance vaut zero. Nous calculons ainsi la statistique de student avec la formule usuelle et determinons ensuite la P-value en utilisant une loi normale (0,1) (le nombre d'observation le permettant).

Action	**Tendance**	**Volatilité**	**Tstat**	**Pvalue**
GOOGL	0.000919994	0.0172729	0.0532623	0.957523
FB	0.00108409	0.0216971	0.0499645	0.960151
AAPL	0.0017154	0.0195188	0.0878843	0.929969
AMZN	0.00166847	0.0206373	0.0808473	0.935563
WMT	0.000834348	0.0145386	0.0573886	0.954236
COST	0.000924898	0.0127727	0.0724121	0.942274
KO	0.000363445	0.0135883	0.0267468	0.978662
MCD	0.000658575	0.0149411	0.0440779	0.964842
BLK	0.000771733	0.0181469	0.0425269	0.966079
JPM	0.000529671	0.0185608	0.0285371	0.977234
BRK-A	0.000416049	0.0128925	0.0322707	0.974256
BLDR	0.00191034	0.034085	0.0560465	0.955305
GM	0.000479619	0.024704	0.0194146	0.98451
TSLA	0.00360342	0.0409205	0.0880592	0.92983
UNP	0.000828914	0.017464	0.047464	0.962143
PFE	0.000276185	0.0148866	0.0185526	0.985198
HRTX	0.00127347	0.0395791	0.0321753	0.974332
XOM	-0.000582319	0.0198455	-0.0293427	0.976591
CVX	-0.000136382	0.01964	-0.00694411	0.994459
PLUG	0.00480522	0.0573679	0.0837615	0.933246

Table (2.1) Tendance avec Tstats et Pvalues associées

Question 3 . *Déduisez-en le portefeuille maximisant le rendement espéré, en supposant que la même dynamique décrit aussi ce qui se passera dans l'avenir et en veillant à bien respecter la contrainte énoncée au début du projet (10% maximum par titre)*

Réponse 3 : Le portefeuille maximisant le rendement espéré dans notre cadre correspond au portefeuille équipondéré des 10 actions présentant les 10 tendances les plus élevées. Ce qui correspond au portefeuille suivant :

Action	**tendance**	**Poids**
GOOGL	**0.000919994**	**10%**
FB	**0.00108409**	**10%**
AAPL	**0.0017154**	**10%**
AMZN	**0.00166847**	**10%**
WMT	**0.000834348**	**10%**
COST	**0.000924898**	**10%**
KO	0.000363445	0%
MCD	0.000658575	0%
BLK	0.000771733	0%
JPM	0.000529671	0%
BRK-A	0.000416049	0%
BLDR	**0.00191034**	**10%**
GM	0.000479619	0%
TSLA	**0.00360342**	**10%**
UNP	0.000828914	0%
PFE	0.000276185	0%
HRTX	**0.00127347**	**10%**
XOM	-0.000582319	0%
CVX	-0.000136382	0%
PLUG	**0.00480522**	**10%**

Table (2.2) Premier portefeuille naïf

Question 4 . *D'une manière générale, hormis la mauvaise spécification d'un modèle, quelle peuvent être les différentes natures de risque de modèle ?*

Réponse 4 : Le risque de modèle peut provenir de différentes sources. Premièrement comme mentionné la mauvaise spécification d'un modèle , c'est à dire le fait d'utiliser un modèle qui représente mal la réalité, entraîne un premier risque de modèle car les estimations seront forcément faussées.
Dans un second temps , même si un modèle est bien spécifié il peut y avoir des incertitudes dans l'estimation. Par exemple le manque de données peut entraîner des paramètres mal estimés, pareil si les données comportent beaucoup de bruit. Ce détail met en lumière le fait de devoir suivre l'évolution des paramètres en fonction des données disponibles, en évaluant si les données de références sont stables dans le temps par exemple, en observant aussi si les variables sont toujours pertinentes dans le temps.
Enfin le risque de modèle présente une dernière nature lié a l'incertitude des résultats. En effet il est important de pouvoir quantifier et de crée des indicateurs sur les résultats, on peut chercher à crée des intervalles de confiances, à préciser l'espérance mais aussi la variance des résultats produits.

Question 5 . *Quel est l'exposant de Hurst de chacune des 20 actions ? (utilisez l'estimateur vu en cours) Indiquez par ailleurs la significativité de chacun de ces exposants de Hurst (l'hypothèse nulle du test statistique demandé est un exposant de Hurst égal à 0.5, ce qui correspond au cas d'un mouvement brownien classique ; vous pouvez estimer la p-value de chaque tendance trouvée à l'aide d'un test bien choisi et dont vous justifierez le choix, ou bien à l'aide d'un bootstrap paramétrique)*

Réponse 5 : Pour calculer l'exposant de Hurst nous utilisons la méthode évoqué en cours qui consiste à estimer la pente de la droite de régression entre la log variance des incréments et la log durée des incréments. Cette pente vaut par définition $2 \times$ H ce qui nous permet d'estimer notre exposant de Hurst.

Pour une fenêtre temporelle d'un an (n = 252) les résultats sont les suivants :

Action	**Exposant de hurst**
GOOGL	0.240723
FB	0.352325
AAPL	0.466365
AMZN	0.470221
WMT	0.280654
COST	0.29624
KO	0.317601
MCD	0.339446
BLK	0.425528
JPM	0.363877
BRK-A	0.280405
BLDR	0.404738
GM	0.346677
TSLA	0.554028
UNP	0.231902
PFE	0.357115
HRTX	0.51338
XOM	0.410407
CVX	0.375164
PLUG	0.472966

TABLE (2.3) Exposants de Hurst sur une fenêtre d'un an

Question 6 . *Comment peut-on interpréter la valeur d'un exposant de Hurst ? Commentez les valeurs obtenues pour les 20 actions.*

Réponse 6 : Un exposant de Hurst égal à 1/2 indique que les incréments des prix sont indépendants et que nous sommes donc dans le cadre d'un mouvement brownien géométrique classique .

Si l'exposant de Hurst est supérieur à 1/2 cela signifie que les incréments sont positivement corrélées, on note alors la présence d'autocorrélation. Ainsi plus cet exposant est élever et plus on efface les singularitées, permettant d'observer des tendance relativement lisses.

Si l'exposant de Hurst est en revanche inférieur à 1/2 cela signifie que les incréments sont négativement corrélés, chaque incrément prend donc généralement le contre pied du précédent.

Globalement sur nos données nous observons majoritairement des exposants de hurst inférieurs à 0.5 (seulement 2 au dessus de ce seuil). Cela tendrait à prouver que les rendements des actions américaines ne sont pas modélisables par un mouvement brownien géométrique classique, ce qui porte atteinte aux travaux effectué sur la base de l'équation de Black and Scholes.Cela remet également au cause l'hypothèse d'efficience des marchés financiers. De plus cela nous indique qu'on observe bien sur les marchés financiers des tendances sur les rendements, ce qui dans un but d'investissement peut se montrer intéressant car cela voudrait dire qu'on peu justement suivre ces tendances, l'investisseur de long terme aurait alors plutôt intérêt à choisir des actions présentant un exposant de hurst faible.

Question 7 . *Proposez un indicateur proche d'un rendement espéré (dans le sens où vous pouvez le voir comme un estimateur de rendement espéré, estimateur dont les propriétés de convergence ne nous intèresseront pas ici) et utilisant l'exposant de Hurst.*

Réponse 7 : En nous positionnant dans le cadre d'un modèle de mouvement brownien fractionnaire géométrique avec drift nous pouvons proposé comme indicateur de rendement espéré basé sur l'exposant de hurst l'estimateur suivant qui correspond au paramètre de drift du modèle :

$$\hat{\mu} = \frac{\overline{\ln(x_i)}}{t-s} + \frac{\hat{\sigma}^2}{2}$$

Avec :

$\hat{\sigma} = \sqrt{\dfrac{S^2_{\ln(x_i)}}{|t-s|^{2H}}} = \dfrac{S_{\ln(x_i)}}{|t-s|^H}$ Le paramètre de volatilité du modèle (où est présent l'exposant de hurst).

$\overline{\ln(x_i)} = \dfrac{\sum_{i=1}^{N} \ln(x_i)}{N}$ La moyenne des rendements logarithmiques.

$S^2_{\ln(x_i)} = \dfrac{\sum_{i=1}^{N} \left(\ln(x_i) - \overline{\ln(x_i)}\right)^2}{N-1}$ La variance des rendements

Les résultats ainsi obtenus sont les suivants :

Action	**drift**
GOOGL	0.00664593
FB	0.00719863
AAPL	0.00597033
AMZN	0.00621101
WMT	0.00534348
COST	0.00469789
KO	0.00455616
MCD	0.00499574
BLK	0.005416
JPM	0.00586873
BRK-A	0.00457806
BLDR	0.0106794
GM	0.00787846
TSLA	0.0112936
UNP	0.0067458
PFE	0.00467502
HRTX	0.0103497
XOM	0.00532632
CVX	0.00576949
PLUG	0.0172923

TABLE (2.4) Drifts du gbfm obtenus à partir de nos exposants de hurst

Question 8 . *Déduisez-en le portefeuille maximisant ce nouvel estimateur de rendement espéré, en supposant que la même dynamique décrit aussi ce qui se passera dans l'avenir et en veillant à bien respecter la contrainte énoncée au début du projet (10% maximum par titre). Comparez ce portefeuille avec celui obtenu dans la première méthode.*

Réponse 8 : Comme pour la question 3, le portefeuille maximisant le rendement espéré dans notre cadre correspond au portefeuille équipondéré des 10 actions présentant les 10 paramètres de drift les plus élevées. Ce qui correspond au portefeuille suivant :

Action	**drift**	**Poids**
GOOGL	**0.00664593**	**10%**
FB	**0.00719863**	**10%**
AAPL	**0.00597033**	**10%**
AMZN	**0.00621101**	**10%**
WMT	0.00534348	0%
COST	0.00469789	0%
KO	0.00455616	0%
MCD	0.00499574	0%
BLK	0.005416	0%
JPM	0.00586873	0%
BRK-A	0.00457806	0%
BLDR	**0.0106794**	**10%**
GM	**0.00787846**	**10%**
TSLA	**0.0112936**	**10%**
UNP	**0.0067458**	**10%**
PFE	0.00467502	0%
HRTX	**0.0103497**	**10%**
XOM	0.00532632	0%
CVX	0.00576949	0%
PLUG	**0.0172923**	**10%**

Table (2.5) Choix de portefeuille basé sur le drift du gbfm

Nous nous retrouvons avec un portefeuille ressemblant beaucoup au précédent avec 8 titres sur 10 toujours sélectionné.

Chapitre 3

Ajustement au risque de marché

3.1 Volatilité comme mesure de risque

Question 9 . *Comment interpréter la valeur de λ ? Choisissez alors un λ pour la suite du projet et justifiez votre choix.*

Réponse 9 . En intégrant le risque de marché par la volatilité des rendement on devra construire un portefeuille de sorte à maximiser la fonction suivante :

$$\mathrm{E}\left(\sum_{i=1}^{20}\omega_i r_i\right) - \lambda\sqrt{\mathrm{Var}\left(\sum_{i=1}^{20}\omega_i r_i\right)} \tag{3.1}$$

Ici λ est l'aversion au risque, et sa valeur s'interprète de la façon suivante :
Plus λ est grand plus l'investisseur est averse au risque. Cela signifie que si $\lambda = 0$ l'investisseur n'a aucune aversion au risque et construit son portefeuille en se basant uniquement sur le rendement espéré.
Dans la littérature, en fonction des ordres de grandeurs habituels des rendements espérés et de leur volatilité, λ est souvent choisi entre 1 et 10 ($1 <= \lambda <= 10$) selon l'aversion au risque de l'investisseur.

Nous avons choisi $\lambda = 2$ car il s'agit de nos préférences quant à l'aversion au risque mais aussi d'une valeur souvent choisi pour cet exercice.
Par ailleurs, nous avons choisi un horizon d'investissement annuel $h = 252$ (Nombre de jours ouvrés).

Question 10 . *Le résultat de l'optimisation va-t-il dépendre de h ? Justifiez.*

Réponse 10 . Le résultat de l'optimisation dépendra de l'horizon d'investissement h.
En effet, mathématiquement, le problème d'optimisation n'est pas fonction de l'horizon h, mais l'horizon d'investissement influe directement sur le rendement espéré et donc sur la répartition des actions ω_i qui est le résultat de l'optimisation.

Question 11 . *Après avoir estimé la matrice de covariance de votre univers d'investissement, déduisez-en le portefeuille maximisant le rendement ajusté défini à l'equation (1). Comparez ce portefeuille avec celui obtenu sans ajustement au risque de marché.*

Réponse 11 . Pour avoir le portefeuille maximisant le rendement il faut résoudre le problème d'optimisation suivant :

$$\max_{\sum_{i=1}^{20}\omega_i=1,\omega_i<=0.1\forall i} E\left(\sum_{i=1}^{20}\omega_i r_i\right) - \lambda\sqrt{\mathrm{Var}\left(\sum_{i=1}^{20}\omega_i r_i\right)} \tag{3.2}$$

Les contraintes sont donc : $\sum_{i=1}^{20} \omega_i = 1$ et $\omega_i \in [0; 0.1] \forall i$
La matrice de variance covariance est la suivante :
Nous avons résolu le problème d'optimisation sous contrainte avec python, le code se trouve dans l'annexe. Les résultats obtenus sont les suivants :

Stock	**Poids**
GOOGL	8.549%
FB	8.64%
AAPL	10%
AMZN	10%
WMT	8.43%
COST	10%
KO	0%
MCD	1.17%
BLK	3.75%
JPM	0%
BRK-A	0%
BLDR	10%
GM	0%
TSLA	10%
UNP	5.7%
PFE	0%
HRTX	3.76%
XOM	0%
CVX	0%
PLUG	10%

Table (3.1) Maximum portfolio

La différence majeure entre les résultats de l'optimisation Espérance-Variance et ceux basés uniquement sur l'espérance se manifestent sur le nombre d'actions utilisées dans le portefeuille.
En utilisant uniquement l'espérance et sans contrainte, on aurait tout mis (100%) sur l'action au rendement le plus élevé et (0%) sur le reste, sauf que, avec la contrainte des 10% maximale par action, le meilleur portefeuille serait construit en sélectionnant les 10 actions les plus rentables et en mettant 10% chacune.
En rajoutant le critère de variance, nous ne pouvons plus beaucoup investir dans un groupe d'actions et négliger le reste, il faudra diversifier pour minimiser le risque. L'optimisation Espérance-Variance comme décrite en haut donne des résultats ou l'on sélectionne plusieurs actions dans le portefeuille.

Question 12 . *Vous avez déjà déterminé l'exposant de Hurst de chacun des titres. Utilisez les estimateurs de l'exposant de Hurst vus en cours pour d ́eterminer le paramètre d'échelle de chacun des titres (ce paramètre d'échelle n'est pas affecté par le fait que l'on a un mBf multivarié : chaque série peut être considérée indépendamment des autres).*

Réponse 12 . Selon le travail mené par Siti Nur Iqmal Ibrahim, Masnita Misiran et Mohamed Faris Laham en 2020 nous calculons notre paramètre d'échelle σ de la manière suivante

$$\hat{\sigma} = \sqrt{\frac{S^2_{\ln(x_i)}}{|t-s|^{2H}}} = \frac{S_{\ln(x_i)}}{|t-s|^H} \tag{3.3}$$

avec la variance de l'échantillon :

$$S^2_{\ln(x_i)} = \frac{\sum_{i=1}^{N}\left(\ln(x_i) - \overline{\ln(x_i)}\right)^2}{N-1}$$

ou

$$\overline{\ln(x_i)} = \frac{\sum_{i=1}^{N}\ln(x_i)}{N}$$

Les résultats ainsi obtenus sont les suivants :

Action	**paramètre d'échelle**
GOOGL	0.014618413
FB	0.016995808
AAPL	0.014127457
AMZN	0.014897138
WMT	0.011968403
COST	0.010401726
KO	0.01090333
MCD	0.011808657
BLK	0.013511584
JPM	0.014423071
BRK-A	0.010615137
BLDR	0.025746897
GM	0.026601242
TSLA	0.027871575
UNP	0.014870828
PFE	0.011622343
HRTX	0.027728284
XOM	0.014931941
CVX	0.015142799
PLUG	0.041332502

TABLE (3.2) Paramètres d'échelle du gbfm obtenus à partir de nos exposants de hurst

Question 13 . *A l'aide de l'équation 2, proposez un estimateur pour ρ_{ij}, sachant que σ_i, σ_j, Hi et Hj sont déjà estimés. Comparez les valeurs obtenues pour votre univers d'investissement à celles du cas d'un mouvement brownien standard géométrique.*

Réponse 13 . A partir de l'équation donnée dans le sujet nous calculons la covariance entre série i et série j à la date t de la manière suivante :

$$\begin{aligned}&\text{Cov } = E\left[\ln(x_i(t))\ln(x_j(t))\right] = \\ &\frac{\rho_{ij}\sigma_i\sigma_j}{2}\left(|t|^{H_i+H_j} + |t|^{\mu_i+H_j} + |t-t|^{H_i+H_j}\right) = \\ &= \frac{\rho_{ij}\sigma_i\sigma_j}{2}\left(2|t|^{H_i+H_j}\right) = \rho_{ij}\sigma_i\sigma_j|t|^{H_i+H_j}\end{aligned}$$

D'où nous obtenons la formule d'estimation de la corrélation :

$$\rho_{ij} = \frac{E\left[\ln(x_i(t))\ln(x_j(t))\right]}{\sigma_i\sigma_j|t|^{H_i+H_j}} = \frac{\text{cov}\left(ln(x_i)\,ln(x_j)\right)}{\sigma_i\sigma_j|t|^{H_i+H_j}} \tag{3.4}$$

avec :

$$\text{cov}\left(\ln(x_i)\ln(x_j)\right) = \frac{\sum_{i=1}^{N}\left(\ln(x_i) - \overline{\ln(x_i)}\right)\left(\ln(x_j) - \overline{\ln(x_j)}\right)}{N-1}$$

Question 14 . *Avec la même aversion au risque λ que précédemment, déterminez le portefeuille qui maximise l'estimateur de rendement espéré (selon la méthode du mBf) sous contrainte de risque.Comparez avec le portefeuille précédent.*

Réponse 14 . Suite de l'optimisation du portefeuille selon la méthode de mBf nous obtenons les poids suivants :

Stock	Poids
GOOGL	10.0%
FB	10.0%
AAPL	10.0%
AMZN	3.6%
WMT	10.0%
COST	10.0%
KO	0.0%
MCD	10.0%
BLK	0.0%
JPM	0.0%
BRK-A	0.0%
BLDR	10.0%
GM	3.0%
TSLA	3.9%
UNP	10.0%
PFE	0.0%
HRTX	0.0%
XOM	0.0%
CVX	0.0%
PLUG	10.0%

Table (3.3) Maximum portfolio mBf

Avec la méthode de mBf nous constatons que par rapport à la méthode précédente nous avons une tendance similaire. Cependant, nous investissons plus dans les actions GOOGL, FB,WMT,MCD,UNP,GM. Au même niveau nous investissons dans les actions COST, BLDR, PLUG, KO,JPM,BRK-A,PFE,XOM et CVX. Nous investissons moins dans les actions TSLA et HRTX.

Avec ces deux méthodes les résultats sont assez similaires, mais ils ne sont pas identiques.

Question 15 . *Le résultat de l'optimisation va-t-il dépendre de h ? Justifiez*

Réponse 15 . Le résultat de l'optimisation dépendra de l'horizon d'investissement h.
Comme nous avons évoqué précédemment, mathématiquement, le problème d'optimisation n'est pas fonction de l'horizon h, mais l'horizon d'investissement influe directement sur le rendement espéré et donc sur la répartition des actions ω_i qui est le résultat de l'optimisation.

3.2 VaR comme mesure du risque

Question 16 . *A l'aide de l'estimateur de Hill ou de Pickands, vus en cours, estimez les paramètres de la GPD de chacun des titres. Comment interprétez-vous ces paramètres ?.*

Réponse 16 . Cet estimateur a pour objectif de décrire les queues de distribution de nos rendements. On va donc commencer par ordonner les rendements de nos titres. Il est donc possible en fonction d'une ordination croissante ou décroissante, d'analyser les gains maximaux ou pertes maximales pour chacun des titres. On s'interesse ici au risque et allons donc modéliser les queues de distributions des pertes de nos rendements.
Dans cette question nous allons utiliser la méthode d'estimation de Pickands. On utilise donc la formule suivante :

$$\xi^P_{k(n),n} = \frac{1}{\log(2)} \log\left(\frac{X_{n-k(n)+1:n} - X_{n-2k(n)+1:n}}{X_{n-2k(n)+1:n} - X_{n-4k(n)+1:n}}\right)$$

On choisit k tel que :

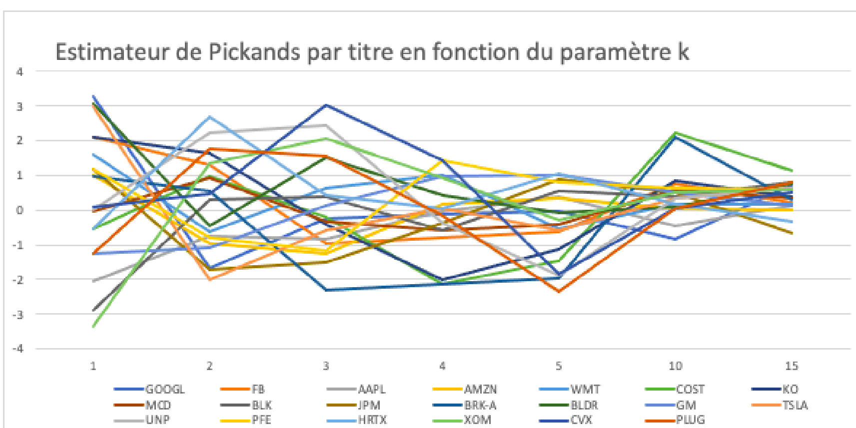
$$\lim_{n\to\infty} k(n) = \infty$$
$$\lim_{n\to\infty} \frac{k(n)}{n} = 0$$

On obtient les résultats suivant pour nos 20 titres.

Résultats avec les données ordonnées de marnière croissante :

K	1	2	3	5	10	15
COOGL	$3,2849$	$-1,6878$	$-0,2350$	$-0,0336$	$-0,8193$	$0,7094$
FB	$2,0965$	$1,3064$	$-0,9435$	$-0,6024$	$0,7723$	$0,2302$
AAPL	$-2,0615$	$-0,7533$	$-0,8163$	$0,3880$	$-0,4334$	$0,1227$
AMZN	$1,0061$	$-0,9369$	$-1,2464$	$0,3515$	$0,0505$	$-0,0014$
WMT	$1,6166$	$-0,6092$	$0,6371$	$-0,5426$	$0,1796$	$0,1647$
COST	$-0,5298$	$0,9814$	$-0,1979$	$-1,4715$	$2,2272$	$1,1620$
KO	$2,1056$	$1,6437$	$-0,4268$	$-1,1326$	$0,8342$	$0,3771$
MCD	$-0,0156$	$0,9408$	$-0,3111$	$-0,3918$	$0,6165$	$0,3466$
BLK	$-2,9044$	$0,3219$	$0,4031$	$0,5468$	$0,3730$	$0,7918$
JPM	$1,1751$	$-1,7231$	$-1,5204$	$0,9089$	$0,5004$	$-0,6607$
BRK-A	$0,9617$	$0,5379$	$-2,3044$	$-1,9545$	$2,1048$	$0,3183$
BLDR	$3,0923$	$-0,4345$	$1,5242$	$-0,0659$	$0,1135$	$0,7819$
GM	$-1,2460$	$-1,0932$	$0,1162$	$1,0102$	$0,5022$	$0,1579$
TSLA	$2,9989$	$-2,0007$	$-0,5868$	$-0,5880$	$0,3474$	$0,6571$
UNP	$0,0274$	$2,2159$	$2,4351$	$-1,8742$	$0,3275$	$0,6225$
PFE	$1,1884$	$-0,7911$	$-1,1602$	$0,8172$	$0,6476$	$0,6065$
HRTX	$-0,5381$	$2,7084$	$0,4298$	$1,0511$	$0,1762$	$-0,3172$
XOM	$-3,3374$	$1,3347$	$2,0485$	$-0,2817$	$0,5316$	$0,6050$
CVX	$0,0942$	$0,4721$	$3,0245$	$-1,8403$	$0,0425$	$0,5146$
PLUG	$-1,2515$	$1,7930$	$1,5628$	$-2,3289$	$0,0600$	$0,8255$

En voici la représentation graphique par titre.



En testnt nos différents k, on décide de fixer notre k à k=15 car nos estimateurs explosent moins.

On peut donc interpréter la nature du comportement du maximum grâce à cet estimateur. Ces distribution des queues de distribution s'apparente à des lois paramétrique en fonction des valeurs de ξ.

si $\xi > 0$, la loi est de type Fréchet ;

si $\xi = 0$, la loi est de type Gumbel ;

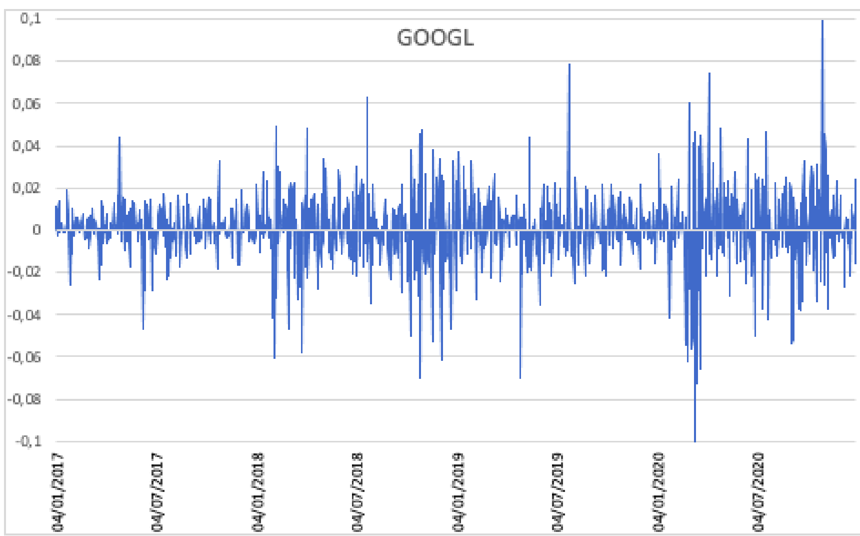
si $\xi < 0$, la loi est de type Weibull.

Pour une loi de type Weibull, cela correspond à une variable aléatoire bornée. La loi de Fréchet correspond à une queue de distribution épaisse. La loi de Gumbel est le cas de queues fines, il y a peu d'évènement extrême.

Question 17 . *Quelles sont les hypothèses sur les rendements qui permettent d'utiliser la théorie des valeurs extrêmes*

Réponse 17 . La théorie des valeurs extrêmes nécessite de fortes hypothèses sur les rendements. Il faut que les variables aléatoires soient indépendantes et identiquement distribuées. En réalité ce n'est pas le cas pour les rendements de nos actifs. Le caractère d'indépendance peut être revu dans l'extension de la théorie des valeurs extrêmes. L'autre hypothèse sur le caractère identiquement distribué peut aussi être contourné en fixant une dynamique sur le rendement,et en appliquant la théorie des valeurs extrêmes sur les résidus des rendements suivant un mouvement AR par exemple.

Voici les rendements pour les actions de GOOGLE.



On peut constater que les rendements ne sont pas identiquement et indépendemment distribués.

Question 18 . *Déterminez le portefeuille optimal en remplaçant la volatilitée par la VaR, pour les deux types de rendements calculées. La VaR sera calculée pour une probabilité de 90%, 95%, 99%, 99.9% et 99.99% et un horizon d'un jour. Expliquez l'éevolution de votre portefeuille en fonction du niveau de probabilitée, notamment à partir des paramètres de la GPD de chaque titre.*

Réponse 18 . Comme nous avons utilisé l'estimateur de Pickands, nous utilons la formule vue en cours de la Value at Risk utilisant cet estimateur. Cette formule est la suivante :

$$\mathrm{VaR}(p) = \frac{\left(\frac{k}{n(1-p)}\right)^{\xi^P} - 1}{1 - 2 - \xi^p} \left(X_{n-k+1:n} - X_{n-2k+1:n}\right) + X_{n-k+1:n}$$

Où ξ^p est l'estimateur de Pickands. On obtient donc les résultats suivants :

Proba	90	95	99	99,9	99,99
GOOGL	$-0,0123$	$-0,0121$	$-0,0120$	$-0,0119$	$-0,0119$
FB	$-0,2562$	$-0,2525$	$-0,2497$	$-0,2491$	$-0,2491$
AAPL	$-0,9947$	$-0,9873$	$-0,9817$	$-0,9805$	$-0,9804$
AMZN	$179,1042$	$179,1183$	$179,1290$	$179,1313$	$179,1315$
WMT	$-0,2895$	$-0,2867$	$-0,2845$	$-0,2840$	$-0,2840$
COST	$-0,0133$	$-0,0133$	$-0,0133$	$-0,0133$	$-0,0133$
KO	$-0,0432$	$-0,0420$	$-0,0411$	$-0,0409$	$-0,0409$
MCD	$-0,0620$	$-0,0605$	$-0,0594$	$-0,0592$	$-0,0591$
BLK	$-0,0143$	$-0,0142$	$-0,0141$	$-0,0141$	$-0,0141$
JPM	$15,7778$	$16,3535$	$16,8067$	$16,9078$	$16,9179$
BRK-A	$-0,0639$	$-0,0626$	$-0,0616$	$-0,0614$	$-0,0614$
BLDR	$-0,0246$	$-0,0244$	$-0,0243$	$-0,0243$	$-0,0243$
GM	$-0,6339$	$-0,6279$	$-0,6233$	$-0,6223$	$-0,6222$
TSLA	$-0,0316$	$-0,0311$	$-0,0308$	$-0,0307$	$-0,0307$
UNP	$-0,0152$	$-0,0149$	$-0,0147$	$-0,0146$	$-0,0146$
PFE	$-0,0164$	$-0,0162$	$-0,0161$	$-0,0160$	$-0,0160$
HRTX	$5,6705$	$5,7707$	$5,8482$	$5,8654$	$5,8671$
XOM	$-0,0195$	$-0,0192$	$-0,0190$	$-0,0189$	$-0,0189$
CVX	$-0,0244$	$-0,0238$	$-0,0234$	$-0,0233$	$-0,0233$
PLUG	$-0,0416$	$-0,0414$	$-0,0413$	$-0,0412$	$-0,0412$

Pour trois de nos titres (AMZN, JPM, et HRTX), on obtient des VaR explosives, les estimateurs de Pickands pour ces titres étaient négatifs. Le seuil de probabilité ne fait varier nos VaR que très légèrement. On constate que lorsque l'estimateur de Pickands est autour ou supérieur à 0.5, la VaR semble être beaucoup plus faible. Lorsque ξ^p est compris entre 0 et 0.5, la VaR est distinctement plus importante en valeur absolue, les pertes sont plus importantes. Concernant les poids des titres de notre portefeuille calculés à partir de ces VaR :

Proba	90	95	99	99,9	99,99
GOOGL	0,03%	0,02%	0,03%	0,01%	0,01%
FB	10,00%	10,00%	10,00%	10,00%	10,00%
AAPL	10,00%	10,00%	10,00%	10,00%	10,00%
AMZN	0,00%	0,00%	0,00%	0,00%	0,00%
WMT	10,00%	10,00%	10,00%	10,00%	10,00%
COST	0,02%	0,03%	0,03%	0,01%	0,01%
KO	9,98%	9,98%	9,98%	9,99%	9,99%
MCD	9,99%	9,99%	10,00%	10,00%	10,00%
BLK	0,03%	0,04%	0,04%	0,01%	0,01%
JPM	0,00%	0,00%	0,00%	0,00%	0,00%
BRK-A	9,99%	9,99%	9,99%	10,00%	10,00%
BLDR	9,42%	9,64%	9,65%	9,95%	9,94%
GM	10,00%	10,00%	10,00%	10,00%	10,00%
TSLA	9,96%	9,95%	9,97%	9,99%	9,99%
UNP	0,05%	0,03%	0,04%	0,01%	0,01%
PFE	0,06%	0,03%	0,04%	0,01%	0,01%
HRTX	0,00%	0,00%	0,00%	0,00%	0,00%
XOM	0,03%	0,04%	0,06%	0,02%	0,02%
CVX	0,46%	0,28%	0,18%	0,01%	0,01%
PLUG	9,98%	9,98%	9,99%	10,00%	10,00%

Chapitre 4

Codes

4.1 Sélection d'actions

```
import pandas as pd
from typing import Union
import yfinance as yf


# Nous devons choisir 20 actions pour notre projet
#
# Nous d cidons de prendre uniquement des actions am ricaines car cela nous   vite
# d'avoir   g r  le risque de change (en nous consid rant comme des investisseurs
# am ricains) de plus elles sont plus nombreuses et diversifi es ce qui nous offre
# une plus grande libert  de choix
#
# Afin d'avoir un portefeuille diversifi  nous d cidons de choisir des actions
# de diff rents secteurs :
#
#
# Information Technology : Google(GOOGL), Facebook (FB), Apple(AAPL), Amazon(AMZN)
#
# Retail : Walmart(WMT), Costco (COST)
#
# Food : The Coca-Cola Company (KO), McDonald's(MCD)
#
# Financials : BlackRock(BLK), JPMorgan Chase & Co(JPM), Berkshire Hathaway(BRK.A)
#
# Industrials : Builders FirstSource Inc (BLDR), General Motors(GM), Tesla(TSLA),
#               Union Pacific(UNP)
#
# Healthcare and Biotech: Pfizer(PFE), Heron Therapeutics(HRTX)
#
# Energy : ExxonMobil(XOM), Chevron(CVX), Plug Power(PLUG)


# On commence par initialiser la liste des tickers
ticker_list=["GOOGL","FB","AAPL","AMZN","WMT","COST","KO","MCD","BLK","JPM",
             "BRK-A","BLDR","GM","TSLA","UNP","PFE","HRTX","XOM","CVX","PLUG"]

# On choisi les dates pour d finir les donn es a r cup r es
start_date= "2017-01-01"
end_date="2021-01-01"


# On d fini les fonctions qui vont nous permettre de prendre les donn es

#########################################################################################

def get_stock(stock, start, end , value: Union[str , list], index_as_date=True):
    """
    Parameters
    ----------
```

```
    stock : str
        Stock abreviation
    start : str
        Start date in format yyyy-mm-dd
    end : str
        End date in format yyyy-mm-dd
    value : Union[str , list]
        Open, High, Low, Close, Adj Close or Volume
    index_as_date : bool
        Set True if you want dates as index and False if you don't want dates.

    Returns
    -------
    pd.DataFrame
        Dataframe with requested data.
    """

    serie = yf.download(stock , start = start , end = end)

    if index_as_date == True:
        serie = serie[value]
        serie.index = serie.index.to_period('d')
    else:
        serie = serie[value].reset_index(drop=True)


    return serie

#########################################################################################

def get_multiple_stock(tickers, start, end, value: Union[str , list], index_as_date=True):
    """

    Parameters
    ----------
    tickers : list
        The list with the stocks wanted (as strings)
    start : str
        Start date in format yyyy-mm-dd
    end : str
        End date in format yyyy-mm-dd
    value : Union[str , list]
        Open, High, Low, Close, Adj Close or Volume
    index_as_date : str, optional
        dafaut True if you want dates as index and False if you don't want dates.

    Returns
    -------
    stocks_df : pd.Dataframe
        Dataframe with requested data.

    """
    list_stock=[]

    for ticker in tickers:
        stock_serie = get_stock(stock=ticker, start=start, end=end,  value=value, index_as_date=index_as_date)
        list_stock.append(stock_serie)

    stocks_df=pd.concat(list_stock, axis=1)
    stocks_df.columns = tickers

    return stocks_df

##########################################################################################

def get_returns(df_stocks):
    """

    Parameters
    ----------
```

```
    df_stocks : pd.DataFrame
        Dataframe with stocks value

    Returns
    -------
    returns_df : pd.DataFrame
        Dataframe with stocks returns

    """

    list_returns=[]

    for ticker in df_stocks.columns:

        serie_return= df_stocks[ticker].pct_change()
        list_returns.append(serie_return)

    returns_df=pd.concat(list_returns, axis=1)
    returns_df.columns = df_stocks.columns

    returns_df = returns_df.iloc[1: , :]

    return returns_df

#########################################################################################


# Application pour les valeurs

df_stocks=get_multiple_stock(tickers=ticker_list,start=start_date, end=end_date , value = 'Open',
    index_as_date=True)

# Exportation des valeurs

df_stocks.to_csv(r"C:\Users\yanis\01 Projets\01 Python Projects\Projet_RiskManagment\
    Projet_RiskManagement\3_Donn es\stocks_data.csv",sep=';',header=True,index=True)



# Calcul des rendements

df_returns=get_returns(df_stocks)

# Exportation des rendements

df_returns.to_csv(r"C:\Users\yanis\01 Projets\01 Python Projects\Projet_RiskManagment\
    Projet_RiskManagement\3_Donn es\stocks_returns_data.csv",sep=';',header=True,index=True)
```

Listing (4.1) Sélection des actions : Importation des données

```
import pandas as pd
import numpy as np
from numpy import log, polyfit, var, subtract
from scipy import stats

# Importation des donn es

df_stocks=pd.read_csv(r"C:\Users\yanis\01 Projets\01 Python Projects\Projet_RiskManagment\
    Projet_RiskManagement\3_Donn es\stocks_data.csv",sep=';',index_col=0)

df_returns=pd.read_csv(r"C:\Users\yanis\01 Projets\01 Python Projects\Projet_RiskManagment\
    Projet_RiskManagement\3_Donn es\stocks_returns_data.csv",sep=';',index_col=0)

#########################################################################################

# Q1 : Calcul des param tres du moouvement brownien

def get_mu_and_sigma(df):
    '''

    Parameters
```

```
    ----------
    df : pd.DataFrame
        DataFrame with the series of interest
    Returns
    -------
    df_param : pd.DataFrame
        A DataFrame with the mean et the standard deviation of all the columns

    '''

    list_mu=[]
    list_sigma=[]

    for i in df:

        mu=df[i].mean()
        sigma=df[i].std()

        list_mu.append(mu)
        list_sigma.append(sigma)

    list_param=list(zip(list_mu, list_sigma))
    df_param=pd.DataFrame(list_param).T
    df_param.columns = df.columns

    return df_param

##########################################################################################

# Q2 : Test de student sur les param tres

def get_t_stats(df):
    '''

    Parameters
    ----------
    df : pd.DataFrame
        DataFrame with shape (2,n)
        mean on the first row
        std on the second row

    Returns
    -------
    df_tstats : pd.DataFrame
        A dataFrame with the tstat and the pvalue associated

    '''

    tstats_list=[]
    Pvalue_list=[]

    for i in range(0,len(df.T)):

        tstats=(df.iloc[0,i])/(df.iloc[1,i]/np.sqrt(len(df)))
        tstats_list.append(tstats)


    for p in range(0,len(tstats_list)):
        Pvalue=2*(stats.norm(0, 1).sf(abs(tstats_list[p])))
        Pvalue_list.append(Pvalue)

    list_tstat_pval=list(zip(tstats_list,Pvalue_list))
    df_tstats=pd.DataFrame(list_tstat_pval).T
    df_tstats.columns = df.columns

    return df_tstats




##########################################################################################
```

```

# Q5 : Calcul de l'exposant de hurst

def hurst_expo(serie_rdmt, n):

    serie_rdmt=serie_rdmt.to_numpy()

    retards = range(2, n)
    list_variances=[]

    for i in retards:

        returns = subtract(serie_rdmt[i:], serie_rdmt[:-i])
        list_variances.append(var(returns))

    m = polyfit(log(retards), log(list_variances), 1)

    hurst_expo = m[0] / 2.0

    return hurst_expo


def get_hursts(df,n_hurst):


    hurst_list=[]

    for col in df.columns:

        hurst=hurst_expo(df[col],n=n_hurst)
        hurst_list.append(hurst)


    df_hurst=pd.DataFrame(hurst_list).T
    df_hurst.columns = df.columns

    return df_hurst



#########################################################################################

# Q7 : Calcul des param tres de drift et de volatilit  pour le mod le gbfm

def get_drift_and_vol(df_param,df_hurst,delta_t):

    list_drift=[]
    list_vol=[]

    for col in df_param.columns:

        vol=np.sqrt(df_param[col][1]/delta_t**(2*df_hurst[col][0]))
        drift=(df_param[col][0]/delta_t)+(vol**2/2)

        list_drift.append(drift)
        list_vol.append(vol)

    list_gbfm_param=list(zip(list_drift, list_vol))
    df_drift_vol=pd.DataFrame(list_gbfm_param).T
    df_drift_vol.columns = df_param.columns


    return df_drift_vol


#########################################################################################

# Application

df_param=get_mu_and_sigma(df_returns)

```

```
df_tstats=get_t_stats(df_param)

df_hurst=get_hursts(df=df_stocks,n_hurst=252)

df_drift_vol=get_drift_and_vol(df_param=df_param,df_hurst=df_hurst,delta_t=2)
```

Listing (4.2) Calculs Partie 1 : Rendement espéré

4.2 Optimisation Espérance-Variance : Construction du portefeuille

```
from scipy.optimize import LinearConstraint , minimize , Bounds

avg_returns = df.mean().values
variances = np.diagonal(df.cov().values)
def objective(w):
    lambd = 2
    U = np.dot(avg_returns,w) - lambd * (np.dot(variances*252,w**2))**0.5
    return - U
bounds = Bounds([0]*20 , [0.1]*20)
linear_constraint = LinearConstraint([1]*20, 1, 1)

x0 = [0]*20
res = minimize(objective, x0, method='trust-constr',
               constraints=linear_constraint,
               options={'verbose': 1}, bounds=bounds)
X = res.x
final = pd.DataFrame({'stock':df.columns , 'weight':X})
```

4.3 Optimisation Mbf : Construction du portefeuille

```
def objective(w):
    lambd = 2
    U = np.dot(avg_returns,w) - lambd * (w.T.dot(cov).dot(w))**0.5
    return - U

bounds = Bounds([0]*20 , [0.1]*20) # Sans vente  d couvert : de 0%  10%
linear_constraint = LinearConstraint([1]*20, 1, 1)


x0 = np.array([0.5]*20)
res = minimize(objective, x0, method='trust-constr',
               constraints=linear_constraint,
               options={'verbose': 1}, bounds=bounds)

X = res.x

final = pd.DataFrame({'stock':df.stock , 'weight':X})
```

4.4 Optimisation Espérance-VaR : Construction du portefeuille

```
def objective(w):
    lambd = 2
    U = np.dot(avg_returns,w) - lambd * np.dot(VaR,w)
    return - U

bounds = Bounds([0]*20 , [0.1]*20) # Sans vente a decouvert : de 0% a 10%
linear_constraint = LinearConstraint([1]*20, 1, 1)

avg_returns = df0.mean().values

final = pd.DataFrame({'stock':df.stock})

for proba in ['90' ,'95' , '99' , '99.9' , '99.99']:
    VaR = df[proba].values

    x0 = np.array([0.5]*20)
    res = minimize(objective, x0, method='trust-constr',
                   constraints=linear_constraint,
                   options={'verbose': 1}, bounds=bounds)

    X = res.x

    final['poids_'+proba] = X
```

4.5 Estimateur de Pickands et VaR

```
import pandas as pd
import numpy as np
import matplotlib.pyplot as plt

df=pd.read_csv('stocks_returns_data.csv',sep=';')


# ## Ordination des donn es

for col in df.columns:
    df[col]=sorted(df[col],reverse=True)

for col in df.columns:
    df[col]=sorted(df[col])

df.drop(columns='Date',inplace=True)


# ##  Value at risk
# ## Estimateurs de
# ### Estimateur de Pickands

# ### Pour k=1
k=1
for i in df.columns:
    e=(1/np.log(2))*np.log((df[i].iloc[-1+1-1*k]-df[i].iloc[-1+1-2*k])/(df[i].iloc[-1+1-2*k]-df[i
    ].iloc[-1+1-4*k]))
    e_p_1.append(e)

e_p_1=pd.DataFrame(e_p_1)
e_p_1.columns =['1']


# ### Pour k=2
k=2
e_p_2=[]

for i in df.columns:
    e=(1/np.log(2))*np.log((df[i].iloc[-1+1-1*k]-df[i].iloc[-1+1-2*k])/(df[i].iloc[-1+1-2*k]-df[i
    ].iloc[-1+1-4*k]))
```

```
    e_p_2.append(e)


e_p_2=pd.DataFrame(e_p_2)
e_p_2.columns =['2']


# ### Pour k=3

k=3
e_p_3=[]


for i in df.columns:
    e=(1/np.log(2))*np.log((df[i].iloc[-1+1-1*k]-df[i].iloc[-1+1-2*k])/(df[i].iloc[-1+1-2*k]-df[i
    ].iloc[-1+1-4*k]))
    e_p_3.append(e)

e_p_3=pd.DataFrame(e_p_3)
e_p_3.columns =['3']


# ### Pour k=4
k=4
e_p_4=[]

for i in df.columns:
    e=(1/np.log(2))*np.log((df[i].iloc[-1+1-1*k]-df[i].iloc[-1+1-2*k])/(df[i].iloc[-1+1-2*k]-df[i
    ].iloc[-1+1-4*k]))
    e_p_4.append(e)

e_p_4=pd.DataFrame(e_p_4)
e_p_4.columns =['4']


# ### Pour k=5
k=5
e_p_5=[]

for i in df.columns:
    e=(1/np.log(2))*np.log((df[i].iloc[-1+1-1*k]-df[i].iloc[-1+1-2*k])/(df[i].iloc[-1+1-2*k]-df[i
    ].iloc[-1+1-4*k]))
    e_p_5.append(e)

e_p_5=pd.DataFrame(e_p_5)
e_p_5.columns =['5']


# ### Pour k=7
k=7
e_p_7=[]

for i in df.columns:
    e=(1/np.log(2))*np.log((df[i].iloc[-1+1-1*k]-df[i].iloc[-1+1-2*k])/(df[i].iloc[-1+1-2*k]-df[i
    ].iloc[-1+1-4*k]))
    e_p_7.append(e)


e_p_7=pd.DataFrame(e_p_7)
e_p_7.columns =['7']


# ### Pour k=10
k=10

e_p_10=[]

for i in df.columns:
    e=(1/np.log(2))*np.log((df[i].iloc[-1+1-1*k]-df[i].iloc[-1+1-2*k])/(df[i].iloc[-1+1-2*k]-df[i
    ].iloc[-1+1-4*k]))
    e_p_10.append(e)
```

```

e_p_10=pd.DataFrame(e_p_10)
e_p_10.columns =['10']


# ### Pour k=15

k=15
e_p_15=[]


for i in df.columns:
    e=(1/np.log(2))*np.log((df[i].iloc[-1+1-1*k]-df[i].iloc[-1+1-2*k])/(df[i].iloc[-1+1-2*k]-df[i
    ].iloc[-1+1-4*k]))
    e_p_15.append(e)


e_p_15=pd.DataFrame(e_p_15)
e_p_15.columns =['15']

conc = pd.concat([e_p_1,e_p_2,e_p_3,e_p_4,e_p_5,e_p_10,e_p_15],axis=1,sort=False)
pickand=conc.transpose()

pickand[13].plot()

for i in range(0,20):
    pickand[i].plot()

pickand.columns = ['GOOGL', 'FB', 'AAPL', 'AMZN', 'WMT', 'COST', 'KO', 'MCD',
       'BLK', 'JPM', 'BRK-A', 'BLDR', 'GM', 'TSLA', 'UNP', 'PFE', 'HRTX',
       'XOM', 'CVX', 'PLUG']


# ## Calcul de la Value at Risk
p=0.05
k=15

var_df=[]
for i in df.columns:
    var=(((k/(len(df)*(1-p))**pickand[i][6])-1)/(1-2**-pickand[i][6]))*(df[i].iloc[-1+1-1*k]-df[i
    ].iloc[-1+1-2*k])+df[i].iloc[-1+1-1*k]
    var_df.append(var)
```