

Les limites de la gestion du risque financier : Big-data et risque de modélisation et d'analyse

Financial risk management limits: Big-data and Risk of modelling and analysis

Wiam ZAIMI, (PhD student)

*Ecole Nationale de Commerce et de Gestion de Kénitra,
Université Ibn Tofail de Kénitra, Maroc*

Abdeslam El MOUDDEN, (PhD professeur)

*Ecole Nationale de Commerce et de Gestion de Kénitra,
Université Ibn Tofail de Kénitra, Maroc*

Adresse de correspondance :	Ecole Nationale de Commerce et de Gestion Campus universitaire B.P : 1420. Kenitra - 14000 Université Ibn Tofail Maroc (Kenitra) 14000 0682588079 et/ou 0661765545
Déclaration de divulgation :	Les auteurs n'ont pas connaissance de quelconque financement qui pourrait affecter l'objectivité de cette étude.
Conflit d'intérêts :	Les auteurs ne signalent aucun conflit d'intérêts.
Citer cet article	ZAIMI, W., & El MOUDDEN, A. (2023). Les Limites de la Gestion du Risque Financier : Big-data et Risque de Modélisation et d'analyse. International Journal of Accounting, Finance, Auditing, Management and Economics, 4(3-1), 98-110. https://doi.org/10.5281/zenodo.7977473
Licence	Cet article est publié en open Access sous licence CC BY-NC-ND

Received: April 25, 2023

Accepted: May 30, 2023

International Journal of Accounting, Finance, Auditing, Management and Economics - IJAFAME

ISSN: 2658-8455

Volume 4, Issue 3-1 (2023)

Les limites de la gestion du risque financier : Big-data et risque de modélisation et d'analyse

Résumé

L'explosion de l'information dans le paysage financier, couplée à une hausse en termes de capacités de stockage, de traitement et d'analyse, justifie la passion de l'industrie financière pour les données massives (Big Data). Cette technologie permet aux organismes financiers d'utiliser plus de données et de réaliser un travail d'analyse méthodique beaucoup plus vaste et rapide, pour minimiser les risques et d'accompagner les futures fluctuations engendrées par les différentes conjonctures économiques nationales et internationales. Or, cette évolution est susceptible non pas seulement de nuire à la sécurité opérationnelle, elle pose également un défi majeur en matière d'analyse et de modélisation des risques financiers lors de la prévision de la fluctuation d'un produit financier ou d'une prise de décision au niveau managérial. Nous soulignons, dans ce papier certains défis liés au traitement et à la modélisation de ces méga-données, en s'appuyant sur une revue de littérature critique, de façon à pouvoir critiquer les modèles traditionnels de présentation des fluctuations des cours boursiers (normalité des marchés financiers). Nous discutons également quelques résolutions quant à des méthodes et des approches statistiques qui peuvent mener à une gestion quantitative adéquate des risques financiers. Nous proposons ainsi le bootstrap, qui présente une méthodologie à mettre en œuvre précisément dans les cas où les méthodes classiques ne parviennent pas à couvrir le champ du problème traité, ou lorsque les modalités de leur application ne sont plus valides.

Mots clés : Big Data ; modélisation financière ; gestion du risque financier ; méthodes asymptotiques ; le bootstrap.

Classification JEL : G17

Type de l'article : Article théorique

Abstract

The explosion of information in the financial industry, combined with an increase in storage, calculation and analysis capacities, explains the financial industry's infatuation with Big Data. This technology allows financial organizations to use more data and to perform a much larger and faster methodical analysis, to minimize risks and to accompany future fluctuations generated by the different national and international economic conditions. Nevertheless, this development could not only reduce the security of operations, but it also represents a major challenge in the analysis and modeling of financial risks when predicting the fluctuation of a financial product or making decisions at the managerial level. In this paper, we highlight some of the challenges related to the processing and modeling of these mega-data, relying on a critical literature review, so as to be able to criticize traditional models of stock price fluctuations (financial market normality). We also discuss some resolutions to statistical methods and approaches that can lead to an adequate quantitative management of financial risks. We propose the use of the bootstrap methodology in the financial area, which presents a technique that intervenes precisely when the scope of the problem considered is not covered by classical methods or when the conditions of application of the previous methods are no longer valid.

Keywords: Big Data; financial modeling; financial risk management; asymptotic methods; the bootstrap.

JEL Classification: G17

Paper type: Theoretical Research

1. Introduction

D'après le Data Driven Marketing et Bosten Consulting Groupe, le marché des données vaut 315 billions de dollars en 2016 rien qu'en Europe, ce chiffre sera multiplier par 3 en 2020.

De même, l'information est considérée comme étant la ressource la plus précieuse du 21^{ème} siècle, ce gisement d'information ou ce qu'on appelle le Big Data est devenu la nouvelle révolution industrielle. Ainsi, et en vertu de la méga-crise du Covid'19, la digitalisation des différents opérations en particulier les opérations financières, est devenue une nécessité incontournable pour les managers et les parties prenantes des différents organismes. Donc, le traitement massif et le stockage de données est utilisé dans tous les secteurs : commerce, finance, politique, santé, sport... L'objectif est de traquer précisément le comportement des agents pour anticiper le futur, prévenir et influencer leurs comportements prochains. En finance, cette révolution a eu lieu au début des années 90 avec le démarrage de la mondialisation et les débuts de la cotation électronique et le suivi en temps réel des données, telles que : Les comptes courants et des crédits permanents, Les indices des valeurs, ...

Littéralement, les termes Big Data signifient « mégadonnées, grosses données ou encore données massives. Ils désignent un ensemble très volumineux de données qu'aucun outil classique de gestion et d'analyse de base de données ou de gestion de l'information ne peut vraiment travailler. Il s'agit d'un concept permettant le stockage et le suivi d'un nombre indicible d'informations sur une base numérique ». (Bremme, 2016)

La motivation de notre recherche vient du fait que l'application du big data est fortement demandée dans la gestion et l'analyse de données avec un grand risque de manipulation et d'exploitation. En effet, selon une étude Ernst & Young (2014), si 72 % des entreprises sondées pensent que les mégadonnées contribuent grandement à la gestion des risques, seulement 7 % d'entre elles savent quelles sont les solutions disponibles et seulement 2 % les appliquent. (Eddye Dibar, 2015). Nous avons également constaté que si de nombreuses entreprises ont des ambitions élevées en matière de données, peu d'entre elles réalisent ces ambitions. En 2019, seuls environ 10 % des entreprises ont déclaré avoir atteint les objectifs en matière de données qu'elles avaient fixés en 2016. En outre, la plupart étaient loin d'atteindre leurs ambitions pour 2021, qu'elles avaient fixées en 2018. (Close & all, 2020) Ainsi, l'hypothèse de cet essai est que Le big data ne se résume pas à des besoins accrus en matière de stockage ou à une collecte de données basée sur des systèmes informatiques, mais aussi une adaptation et une évolution des outils et des approches de gestion de risque suite à leurs manipulations.

D'après Flood & all (2016) : « le système financier étant un organisme en évolution, les autorités de surveillance macroprudentielle doivent elles aussi innover pour garder une série de modèles adaptés aux institutions et aux pratiques de marché existantes. Le résultat est que le système financier et la boîte à outils macroprudentielle doivent évoluer en parallèle dans le temps, chacun répondant aux innovations de l'autre ». Ceci reflète alors, l'importance de la modélisation et de l'analyse comme composantes centrales de la supervision macroprudentielle de la stabilité financière, qui doivent être traitées minutieusement et attentivement pour une gestion fiable des risques financiers loin des considérations classiques et triviales.

La première section est dédiée à l'analyse des opportunités et des risques du big data selon plusieurs points de vue (académiques et professionnels) dans le domaine de la finance, tout en évoquant la revue de la littérature. La section suivante présente le bootstrap en tant une approche permettant de mener à bien la modélisation et l'analyse de données massives pour

une gestion efficace des risques financiers. Cette section est subdivisée en deux petites rubriques : la première décrit la méthodologie du bootstrap et la seconde explique comment utiliser le bootstrap en finance en tirant parti des travaux antérieurs.

2. Les théories fondamentales de la gestion du risque financier

L'histoire de la gestion de risque n'a commencé qu'après la Deuxième Guerre mondiale, à cette époque ; la gestion de risque s'est limitée à l'utilisation des outils très restreints comme l'assurance de marché pour protéger les entreprises et les individus contre quelques pertes liées à des accidents. Jusqu'à les années 1950, d'autres formes de gestion du risque sont apparues, mais qui sont seulement associées aux risques purs, ces formes viennent pour alterner l'assurance de marché qui était incomplète et coûteuse.

Selon Dionne (2013), « Sinder (1956) observera à cette époque qu'il n'y avait pas de livre sur la gestion des risque et qu'aucune université n'offrait de cours sur le sujet. Les deux premiers livres académiques ont été publiés par Mehr et Hedges (1963) et Williams et Heins (1964) et leur contenu portait sur la gestion des risques purs. Ce qui excluait les risques financiers des entreprises ».

La naissance de la théorie de la modélisation financière est généralement associée aux travaux fondateurs de Louis Bachelier, le mathématicien français qui, en 1900, supposait que les variations des cours boursiers sont des variables aléatoires indépendantes qui suivent une loi gaussienne, avec cette proposition plutôt révolutionnaire à une époque où la science baignait dans le déterminisme laplacien : "Le marché n'obéit qu'à une seule loi : la loi du hasard".

La théorie de Bachelier stipule que le prix d'une action ou d'un indice boursier suit un processus brownien, analogue au mouvement erratique de particules microscopiques qui subissent des chocs aléatoires des molécules environnantes. Mais ce n'est que dans les années 1970 qu'une légère modification du modèle de Bachelier est apparue. Les Américains Fischer Black et Myron Scholes ont en effet représenté les rendements (ou variations logarithmiques) d'une action comme des variables aléatoires indépendantes et gaussiennes. Le point commun de ces modèles est de quantifier l'incertitude sur la valeur future d'un portefeuille d'actions à travers la notion de volatilité ou de variance. Dès lors que ce risque a été quantifié, des outils mathématiques peuvent être déployés pour tenter de le minimiser. En fait, en 1954, Harry Markowitz (prix Nobel d'économie en 1990) avait notamment proposé une méthodologie permettant de minimiser le risque d'un portefeuille en imposant un rendement donné. C'est ainsi qu'est née la théorie de l'optimisation de portefeuille (CONT, 2009).

Avant les années 1970, les produits dérivés étaient seulement utilisés pour couvrir des produits agricoles, mais après cette année, l'utilisation des produits dérivés s'est développée comme instruments de gestion différents de risques assurables et non assurables surtout durant les années 1980 (Dionne 2013).

La gestion financière des risques et la gestion des risques de marché et de crédit ont commencé à entrer en vigueur pour les entreprises et pour les institutions financières, dont les banques et les compagnies d'assurances au début des années 1980, la période de l'instauration de la réglementation internationale des risques.

En effet, après les différents scandales et faillites dus à une gestion des risques déficiente, la réglementation Sarbanes-Oxley est entrée en vigueur en 2002 aux États-Unis en vue d'instaurer des normes de gouvernance d'entreprise. Les bourses de valeurs, dont le NYSE en 2002, ont aussi adopté des règles de gouvernance en matière de gestion des risques

destinées aux entreprises cotées sur ces marchés. Cependant, toutes ces réglementations, règles et méthodes de gestion des risques se sont avérés insuffisants pour prévenir la crise financière de 2007. Ce qui faisait défaut, ce n'était pas forcément la réglementation des risques et les normes de gouvernance, mais plutôt leur mise en œuvre ou leur respect. Il est généralement admis que les réglementations et les règles peuvent être contournées par certains acteurs sur de nombreux marchés. Or, ces comportements de contournement semblent être devenus des pratiques courantes avant la crise financière, sans être anticipés, vus ou réprimandés par les régulateurs. (Dionne 2013).

3. Le Big Data financier, une masse d'opportunités et de risques :

Selon Talend (2022), « Le big data dans la finance fait référence aux pétaoctets de données structurées et non structurées qui peuvent être utilisées pour anticiper le comportement des clients et créer des stratégies pour les banques et les institutions financières ».

Le secteur financier est un grand générateur de données. Les données structurées représentent les informations qui sont traitées au sein d'une entreprise en vue de dégager des informations clés pour la prise de décision. Les données non structurées proviennent de sources multiples, en quantités croissantes, et offrent d'importantes possibilités d'analyse. Structurées ou non, les mégadonnées sont analysées en permanence par toutes les entités à des fins multiples. Cependant, seuls ceux qui savent comment les utiliser peuvent réussir à en tirer profit.

Le big data dispose d'une source colossale de données pour les acteurs financiers. C'est à eux de choisir les plus appropriées. L'analyse des différents indicateurs et rapports a le pouvoir d'influencer les futures décisions. Or, les données émotionnelles, démographiques, ou de localisation sont également à prendre en compte. Il en découle que certaines entreprises engagent des experts, pendant que d'autres font appel à l'intelligence artificielle. On trouve actuellement des robots chargés d'analyser les mégadonnées à la place des humains.

Cela permet de réduire le coût de l'investissement, mais aussi d'augmenter le risque financier classique ainsi qu'engendrer de nouveaux risques. C'est pourquoi les entreprises devraient revoir leur politique de gestion de risques financiers.

En raison de la numérisation croissante des processus et des services et compte tenu de la fluctuation de la conjoncture financière (crise financière de 2008), sanitaire (Covid'19) et la crise de l'inflation au niveau international (guerre de l'Ukraine 2022), les sociétés peuvent collecter et générer davantage de données avec plus de prudence et de méfiance. En plus de l'évolution de leurs activités, il est possible d'exploiter cette masse d'informations afin de renforcer leur gestion de risques (financiers, de marché, opérationnels, de conformité...) et d'améliorer la performance au niveau du processus décisionnel. Dans cette optique, certains auteurs justifient l'utilisation de big data comme outil d'aide pour les entreprises afin d'apprécier les risques de façon plus précise et plus exhaustive, conformément à des réglementations, explique Reda Gomery, associé responsable data et analytics chez Deloitte, tout en utilisant leurs modèles qui sont basés essentiellement sur des échantillons et des sous-ensembles qui ne fournissent pas une vision globale du risque (Eddye Dibar, 2015).

En revanche, une telle évolution serait non seulement capable de compromettre la sécurité des interventions et/ou de favoriser le phénomène de la cybercriminalité, voire même de multiplier les périls traditionnels du système financier, comme lors de la crise financière de 2008. En outre et avec l'introduction de la monnaie digitale, le volume de données

électroniquement lisibles qui sont enregistrées ou diffusées dans le secteur financier connaît une croissance exponentielle, ce qui a des conséquences importantes pour la surveillance macroprudentielle (FLOOD & all, 2016). Ce dernier précise également que même si les données sont importantes et quasi omniprésentes au sein des marchés financiers, il convient d'adapter les procédures de gestion de ces données. Ceci est flagrant pour assurer la stabilité financière et macroprudentielle.

Nombreux sont les auteurs qui ont discuté et analysé les opportunités et les défis du big data, par exemple Fan & all (2014) fournit un aperçu des principales caractéristiques du Big Data et de l'impact de ces caractéristiques sur le changement de paradigme dans les méthodes de calcul et les architectures statistiques et informatiques. Il présente également diverses nouvelles perspectives sur l'analyse et le calcul du méga-données. D'une part, le Big Data est extrêmement prometteur en ce qui concerne la détection de modèles de la base de données et d'hétérogénéités subtiles inaccessibles avec des données à faible échelle. D'autre part, la taille énorme de l'échantillon et la haute dimensionnalité des Big Data entraînent des défis particuliers sur le plan informatique et statistique, notamment l'extensibilité et le goulot d'étranglement du stockage, l'accumulation de bruit, la corrélation parasite, l'endogénéité aléatoire et les erreurs de mesure.

VISNOVSKY (2016), le Secrétaire général adjoint de La Banque de France a donné également son point de vue en tant qu'un superviseur, sur la manière de piloter les risques à l'ère de la transformation digitale. En considérant que le Big Data est un choc technologique multiforme/multiservice qui s'est manifesté sous forme d'opportunités dans les institutions financières à savoir dans leurs infrastructures, automatisation d'activité à haute valeur ajoutée, réduction de l'intermédiation, rôle stratégique de la donnée, niches et produits très spécifiques... Cette technologie met en évolution les activités principales de ces institutions :

Tableau 1: Evolution des activités principales des institutions financières

Paie ment			Financement	Investissement
Nouvelle solution de paiement : Paie ment via mobile, porte-monnaie électronique... Nouveaux services de paiement : Compte bancaire digital, initiateurs de paiements, agrégations de paiement. Monnaies virtuelles Big data – open data Nouveaux usages des données clients Développement de risques et de scores commerciaux Cloud computing Blockchain et smart contrats Authentification client (identité numérique, biométrie...) Comparateurs de tarifs.			Crowdfunding : Intermédiation de dons, prêts, titres financiers	Trading électronique : Trading haute fréquence
			Autres plateformes d’intermédiation : Affacturage, bons de caisse...	Conseil financier automatisé : « Robot advisory »

Source : Banque de France (2016), « Le Pilotage des Risques à L'ère de la Transformation Digitale »

Cependant, les atouts ne devraient pas occulter les principaux points d'attention, VISNOVSKY (2016), présente dans le même papier un tableau comparatif où il compare les opportunités et les points d'attention qu'on doit surveiller :

Tableau 2: Les opportunités et les points d'attention de la transformation digitale

Opportunités	Points d'attention
• De nouveaux services pour les utilisateurs • une baisse des coûts de transaction pour les utilisateurs et de traitement interne des intermédiaires • une résilience accrue de certains process • un impact positif pour l'économie réelle et son financement • Un enjeu de compétitivité	• Une application plus difficile de la supervision • Un impact sur le risque opérationnel engendré par une augmentation des interdépendances entre acteurs et une évolution des risques (risques cyber en hausse) • des préoccupations en matière de LCB-FT et de protection de la clientèle • Des risques accrus pour la stabilité financière liés à l'automatisation de certains processus • Risque de délocalisation/passeports entrants en cas d'inadéquation du cadre de régulation et de contrôle

Source : Banque de France (2016), « Le pilotage des risques à l'ère de la transformation digitale »

En outre, la vitesse à laquelle l'information est exploitée grâce à cette technologie exacerbe les phénomènes de volatilité des marchés et de propagation inter-classes d'actifs, éléments fondamentaux dans la gestion des risques financiers. De même, Le haut niveau de corrélation existant entre les diverses stratégies de trading à haute fréquence a tendance à amplifier les chocs (Galhau, 2016). Les aspects des risques du trading à haute fréquence étaient traités au regard des constats académiques par Foucault (2016), Professeur de finance à HEC Paris, en repérant leurs sources, et qui nécessitent d'être abordées par les autorités de réglementation. Selon lui, le gain réalisé par les THF sur les opérations d'achat, suite à "l'annonce de bonnes nouvelles macroéconomiques" et juste avant une augmentation des prix liée à ces annonces, constitue une forme de perte du côté de leurs contreparties, celles-ci étant en effet victimes de la "sélection adverse" pratiquée par les intervenants les plus informés. Ainsi, Brogaard & all (2014) indiquent : « nous montrons que le trading haute fréquence entraîne des coûts de sélection adverse pour d'autres investisseurs ». Donc ce risque augmente en parallèle avec la technologie de THF, surtout que Hibernia Networks a installé une nouvelle fibre optique sous-marine transatlantique, le "Project Express", qui permettra d'accélérer de 5 millisecondes le temps d'interconnexion entre les bourses de New York et de Londres. Les 300 millions de dollars investis dans l'infrastructure proviennent des frais payés par les traders pour un accès rapide aux informations. Selon les estimations du Tabb Group, pour la seule année 2013, le volume investi en faveur des solutions de négociation à haute fréquence a avoisiné 1,5 milliard de dollars, deux fois plus qu'en 2012 (Patterson, 2014).

Par ailleurs, les organismes demeurent confinés aux méthodes classiques d'analyse des données issues d'approches asymptotiques. En effet, une modélisation des risques accompagnée d'une augmentation des capacités de stockage, de calculs et d'analyse, mène à une modélisation inadéquate et par conséquent une mauvaise gestion des risques financiers, une des principales raisons de la crise financière de 2008. En effet, la crise a fait la preuve que les vulnérabilités sont susceptibles d'apparaître dans des angles morts échappant à la collecte formelle de données. Par exemple, au mois de septembre 2008, « les régulateurs étaient loin d'en savoir assez sur les activités de produits dérivés de gré à gré de Lehman Brothers et d'autres banques d'investissement, qui étaient pourtant des opérateurs de premier plan dans ce domaine » (FCIC, 2011). Pour remédier à cette situation, il a fallu se concentrer sur l'identification et la correction de ces lacunes en matière de données. (FLOOD, 2016). De ce fait, on peut dire que la robustesse d'un système financier est

fortement fragile et sensible à toute fausse de main-œuvre des données disponibles ou à collecter !

De surcroît, le petit nombre d'événements systémiques dévoile le degré de robustesse des outils d'analyse statistiques, ce qui oblige à réfléchir plus finement en termes de scénarios susceptibles de conduire à des événements systémiques probables. Ainsi, les mesures du risque systémique et les instruments macroprudentiels n'ont pas de correspondance précise, ce qui nécessite un suivi simultané de nombreux éléments avec la prise en considération du volet de l'analyse qualitative et descriptive de tels événements systémique (Bandt & all, 2015).

Dans le cadre de l'étude des limites, de la gestion des risques financiers à l'ère de big data, nous avons effectué une étude analytique de la modélisation de l'indice MASI (Zaimi, 2022), l'indice principal de la bourse de Casablanca, via deux approches classiques. Approche paramétrique en utilisant deux processus stochastique (la marche aléatoire et le processus autorégressif AR(1)), et approche non paramétrique basée sur le lissage linéaire ou la Moyenne mobile. Les résultats ont montré par ailleurs que le recours à modèles classiques de l'analyse chronologique des données de cet indice entraînent des prévisions biaisées de la fluctuation associée. Ce qui nécessite l'introduction de nouvelles méthodologies d'analyse de données permettant une bonne maîtrise de risque financier et une efficacité des prévisions de la fluctuation future des données.

4. La nature de la modélisation et de l'analyse de big data pour une gestion efficace du risque financier

Les statisticiens ont appris depuis toujours à établir des hypothèses, sélectionner le bon modèle en fonction de quelques paramètres, procéder à l'estimation de ces derniers et valider le modèle en effectuant des tests sur les paramètres. Pour effectuer toutes ces opérations, il existe plusieurs méthodes d'estimation qui ont été mises au point. Il est notamment possible de penser à la méthode des moindres carrées ordinaire, la méthode de vraisemblance, les intervalles de confiance, les tests d'hypothèse, les erreurs standards... . Cependant, ces méthodes échouent promptement si certains postulats techniques cessent d'être vérifiés - supposition intolérable de normalité, manque de données d'observation - ce qui limite par conséquent le champ d'application.

Il s'agit donc de penser à des techniques qui répondent à la fois au problème de quantité de données (insuffisantes ou massives), et aux hypothèses de normalité non acceptable.

Le recours à la méthodologie bootstrap permet de combler les lacunes des méthodes classiques ou de s'affranchir des conditions d'application de ces dernières.

4.1. La méthodologie du bootstrap :

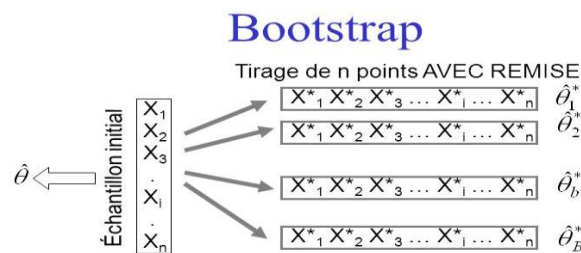
La motivation du bootstrap (Efron, 1982) et (Efron & Tibshyrani, 1993) est « d'approcher par simulation (Monte Carlo) la distribution statistique d'un caractère étudié lorsque l'on ne connaît pas la loi exacte de sa population, ou de même lorsque l'on ne peut pas supposer qu'elle est gaussienne. L'objectif est de remplacer des hypothèses probabilistes pas toujours vérifiées ou même invérifiables par des simulations numériques déduites à partir de l'échantillon observé ». Ce qui permet une amélioration considérable de plusieurs outils d'analyse utilisés dans divers domaines ; en particulier dans le domaine d'un marché financier. Nous signalons dans ce contexte la possibilité d'une amélioration importante de la formule de Blach & scholes qui modélise la fluctuation d'un titre financier, mais malheureusement en se basant sur la normalité présumée du marché !

La méthodologie de bootstrap présente en revanche une liberté de manipulation de données à l'égard des considérations et des hypothèses théoriques, tout en assurant une efficacité et une efficacité de l'analyse statistique suite à l'utilisation directe des données réellement observées qui peuvent être ré-échantillonnées autant de fois qu'on veut.

En effet, elle consiste à remplacer la distribution de probabilité inconnue F , sur laquelle repose l'échantillon initial de taille n , par la distribution empirique F_n qui attribue un poids de $1/n$ à chacune des réalisations. En conséquence, nous obtenons un ré-échantillon, dénommé échantillon bootstrap, conformément à la distribution F_n au moyen de n tirages aléatoires avec remise entre les n observations initiales, étant entendu que les données sont supposées être indépendantes.

Nous pouvons très simplement créer un nombre très important ré-échantillons qui serviront de base au calcul de l'estimateur en question. Par ailleurs, la distribution simulée de cet estimateur est une approximation asymptotiquement convergente sous des hypothèses raisonnables de la loi de l'estimateur. Elle aboutit donc à des estimations du biais et de la variance qui peuvent être utilisés dans l'interprétation de la volatilité de la fluctuation des actifs financiers ou non financiers, sans passer par la supposition de la normalité des caractères des statistiques utilisés.

Fig 1. Schéma démonstratif de la méthodologie du bootstrap



Par exemple, lorsque on cherche à étudier la qualité et le comportement d'un estimateur $T = \hat{\theta}$, on peut utiliser le bootstrap pour approcher le biais de cet estimateur indépendamment de la valeur inconnue θ du modèle comme suit :

$$\text{Biais}(T) = E(T) - \theta$$

$$\widehat{\text{Biais}}(T) \frac{1}{B} = \sum_{b=1}^B (\hat{\theta}_b^* - \hat{\theta})$$

De la même manière, on peut estimer la variance puis l'écart-type ou volatilité dans le cas financier en utilisant la méthodologie de bootstrap comme suit :

$$\widehat{V}(T) = \frac{1}{B} \sum_{b=1}^B (\hat{\theta}_b^* - \hat{\theta})^2$$

Plus précisément, supposons qu'on veut estimer l'espérance de la loi du caractère étudié (exemple : la valeur d'un actif financier), on peut alors procéder de la manière suivante :

Etape 1 : choix de l'échantillon de bootstrap

- On fixe d'abord le nombre B d'échantillons de taille n
- A partir de l'échantillon initial, on choisit aléatoirement et selon la distribution $\hat{F}$ les observations $X_1^*, X_2^*, \dots, X_n^*$ pour construire finalement un échantillon bootstrap. La distribution $\hat{F}$ vaut une probabilité équilibrée de $\frac{1}{n}$ dans le cas où les observations sont supposé non corrélés.

- Le ré-échantillonnage ci-dessus est répété B fois pour avoir B échantillon de bootstrap.

Etape 2 : l'estimation de bootstrap :

Dans le cas de la moyenne théorique de la population, on a :

- Calcul de la valeur estimée à partir de l'échantillon bootstrap :

$$\hat{\theta}_b = (X_1^* + X_2^* + \dots + X_n^*)/n ; \text{ avec } b = 1, 2, \dots, B$$

- Calcul de la moyenne empirique $\bar{\hat{\theta}}$ est la moyenne des moyennes empiriques des échantillons bootstraps $\hat{\theta}_b$

Etape 3 : vérification de l'efficacité de l'estimateur

Dans cette étape, on peut estimer la variance de l'estimateur comme suit :

$$\hat{\sigma}^2(\hat{\theta}) = \frac{1}{B} \sum_{b=1}^B (\hat{\theta}_b^* - \bar{\hat{\theta}})^2$$

Avec

$$\bar{\hat{\theta}} = \frac{1}{B} \sum_{b=1}^B \hat{\theta}_b$$

Qui constitue une alternative à l'estimateur classique $\hat{\theta} = \frac{1}{n} \sum_i X_i$ pour plus de consistance dans l'estimation des paramètres souhaités de la population.

Dans ce contexte, dans le domaine de la finance, on peut profiter de cette méthodologie de bootstrap pour pouvoir performer la qualité des prévisions de la fluctuation d'un certain nombre de produits financiers tout en remplaçant les paramètres analytiques par les paramètres provenant directement de la base de données observée. Ceci permet en outre de minimiser le risque de la modélisation non adéquate et de profiter davantage de la richesse en quantité d'informations de la base de données. Toutefois, l'application de cette méthodologie nécessite une connaissance à priori de la nature de la corrélation et d'interdépendance entre la succession des observations de point de vue structurel ou chronologique.

4.2. Le Bootstrap en finance

Pour analyser le Bootstrap en Finance, Cogneau & Zakamouline (2010) ont fait appel à un sujet central dans leur article : « Il est souvent nécessaire d'estimer le risque et la récompense d'un actif à long terme, compte tenu d'un échantillon d'observations à court terme. Deux obstacles communs à ces estimations sont le manque de données suffisantes et l'incertitude quant à la nature du processus de génération de données ». Ils suggèrent alors, le facteur critique de la modélisation statistique de données et son impact sur l'ampleur du risque d'estimation de la fluctuation d'un actif financier à long terme.

Pour surmonter ces obstacles, leurs recherches s'appuient sur des méthodes bootstrap statistiques. Une réalisation pratique d'une telle méthode dépend essentiellement de la supposition que les données sont dépendantes de la série ou non : « Un bootstrap est une méthode informatique d'estimation des paramètres et des distributions en ré-échantillonnant les données d'origine. Cependant, la réalisation pratique d'une méthode bootstrap dépend essentiellement de la supposition que les rendements soient dépendants de la série ou non » (Cogneau & Zakamouline, 2010)

Si on suppose que les rendements sont indépendants, on utilise alors la méthode du **bootstrap standard**, proposée par Efron (1979), ce qui est assez simple dans la mise en

œuvre. Cependant, la plupart des praticiens pensent que les rendements boursiers sont en train de revenir à la moyenne. Cela signifie que les actions deviennent moins risquées à mesure que l'horizon d'investissement se prolonge.

Si tel est le cas, la méthode bootstrap standard ne peut pas être utilisée, car cette méthode détruit toute dépendance sérielle dans les données. Afin de capturer la structure de dépendance dans la série de retours tout en estimant le paramètre d'intérêt de la distribution de retour sur de longs horizons, beaucoup de recherches s'appuient sur les méthodes **bootstrap de blocs mobiles** introduites par Hall (1985) et Kunch (1989). Ces méthodes utilisent des blocs de données qui se chevauchent au lieu d'observations individuelles pour estimer les paramètres et les distributions.

Pour faciliter une large utilisation de la méthode bootstrap en finance, Chou, & Zhou, (2006) montrent par des arguments intuitifs et par des simulations comment elle peut améliorer les tests existants pour permettre des hypothèses distributionnelles moins restrictives sur les données et pour produire une inférence asymptotique plus fiable (précision d'ordre supérieur).

Dans le même contexte, Kung & Carverhill (2012) ont réalisé une étude visant à déterminer si le Nikkei 225, l'indicateur de référence du marché boursier japonais, évolue dans le temps selon certains processus aléatoires largement utilisés pour les cours boursiers, en utilisant la méthode bootstrap. Toutefois, le fait que ces chercheurs ne disposent que d'un seul échantillon historique les a limités. Par conséquent, les résultats empiriques antérieurs basés sur différentes séries chronologiques de données soient, dans une certaine mesure, différents en ce qui concerne la dynamique des prix des actifs. Leur étude a donc été motivée par l'utilisation de la méthodologie bootstrap afin de résoudre le problème de l'"échantillon unique". Le bootstrap les a permis donc de générer aléatoirement un grand nombre d'échantillons (2 000 échantillons dans cette étude) en les utilisant dans les processus aléatoires (séries de Nikkei artificielles), puisqu'ils possèdent les mêmes propriétés statistiques que les séries de Nikkei réelles. Avec ces 2 000 séries de Nikkei artificielles les auteurs ont construit un intervalle de percentile bootstrap à 95 %, pour examiner le niveau de correspondance entre les rendements de la série Nikkei réelle et les rendements générés par chacun des quatre processus.

5. Conclusion

Face à l'importance et à la généralisation des données sur les marchés financiers, il est nécessaire d'actualiser les processus de gestion de cette ressource fondamentale, en termes d'outils d'analyse et en termes de modèles à utiliser, afin de minimiser davantage le risque financier de point de vue managérial et de point de vue prévisionnel.

La motivation de cet article est de montrer, via les recherches scientifiques déjà effectuées et les avis des professionnels du domaine, que la technologie du Big data incite les agents financiers à adapter de plus en plus leur gamme d'outils d'analyse et de modélisation financière à ces transformations numériques excessives. Le Big Data devient un problème lorsqu'il impacte les conditions d'application des hypothèses technique et d'estimation des paramètres. La méthodologie bootstrap est utilisée dans le cas où l'étendue du problème considéré dépasse les normes des méthodes analytiques classiques, tel que la normalité des caractères financiers étudiés ; qui est le cas d'un marché financier arbitré. Ou de même ou lorsque les modalités d'application de ce dernier ont cessé d'être valides.

Le but de cette approche consiste à se servir de l'échantillon d'observations en vue d'une inférence statistique affinée. Nous créons un certain nombre d'échantillons - appelés

échantillons bootstrap - en tirant au hasard des observations de l'échantillon original. Nous estimons ensuite les différents paramètres du modèle pour chacun des ré-échantillons. On obtient ainsi une séquence de paramètres. La théorie prouve que, sous réserve de certaines conditions de régularité, la distribution de la séquence de paramètres estimés tend à converger vers la vraie loi de distribution du paramètre.

Par ailleurs, il est à signaler que l'approche bootstrap peut conduire à des inférences beaucoup plus précises par rapport aux approches traditionnelles. Cependant, comme tout outil en économétrie, il doit être utilisé avec précaution.

Les perspectives de ce papier est de tester l'efficacité de la méthodologie de bootstrap dans la modélisation et l'analyse de la fluctuation d'un produit financier dans le marché boursier marocain via les modèle économétriques ARCH et GARCH.

Références

- (1). Bandt, O. & all, (2015). La Mesure Du Risque Systémique Après La Crise Financière. *Revue économique*. 66(3). 481-500.
- (2). Bremme, L. (2016). <https://www.lebigdata.fr/definition-big-data>
- (3). Brogaard, J., Hendershott, T. & Riordan, R. (2014). High frequency trading and extreme price movements.
- (4). [https://www.rsm.nl/fileadmin/home/Department of Finance_ VG5 /LQ2015/Ryan_Riordan.pdf](https://www.rsm.nl/fileadmin/home/Department%20of%20Finance_VG5/LQ2015/Ryan_Riordan.pdf)
- (5). Chou, P. H., & Zhou, G. (2006). Using Bootstrap to Test Portfolio Efficiency. *Annals of Economics & Finance*, 7(2).
- (6). Close, K., Kimura, R., Laubier, R., Quarta, L., Charanya, T. & Evans, P. (2020). Data Dominates.
- (7). <https://www.bcg.com/publications/2020/how-data-can-create-competitive-advantage>
- (8). Cogneau, P. & Zakamouline, V. (2010). Bootstrap Methods for Finance: Review and Analysis.
- (9). Data Driven Marketing et Bosten Consulting Groupe; 2016
- (10). Dibar, E. (2015). Gérer Les Risques De Façon Plus Fine Grâce Au Big Data. BFM Business.
- (11). https://www.bfmtv.com/economie/entreprises/services/gerer-les-risques-de-facon-plus-fine-grace-au-big-data_AN-201506260169.html
- (12). Efron, B. (1979). Bootstrap methods : another look at the jackknife. *Annals of Statistics*. Vol. 7. pp. 1-26.
- (13). Efron, B. (1982). The jackknife, the bootstrap and other resampling plans. *Society for industrial and applied mathematics*.
- (14). Efron, B. & Tibshyrani, R., J. (1993). An introduction to the bootstrap. ed. Chapman and Hall, New-York
- (15). Fan, J., Han, F. & Liu, H., (2014). Challenges of Big Data analysis. *National Science Review*. 1(2). 293–314
- (16). Financial Crisis Inquiry Commission. The financial crisis inquiry report: final report of the national commission on the causes of the financial and economic crisis in the United States. U.S. Government Printing Office. (2011)

- (17). FLOOD, M., D., Louiqa, R., JAGADISH, H., V. (2016). Le big data : défis et opportunités pour la surveillance de la stabilité financière. *Revue de la stabilité financière*. N° 20. 145-160
- (18). Galhau, F., V. (2016). Construire le triangle de compatibilité de la finance numérique : innovations, stabilité, régulation. Banque de France. *Financial Stability Review*. No.20. 7-15
- (19). Hall, P. (1985). Resampling a coverage pattern. *Stochastic processes and their applications*, 20(2), 231–246.
- (20). Hannoun, H. (2010). Towards a global financial stability framework. Dans 45th seacen Governors' conference. 26-27.
- (21). Kung, J. J. & Carverhill, A. P. (2012). A Bootstrap Analysis of the Nikkei 225. *Journal of Economic Integration*. Vol. 27. No. 3. 487-504.
- (22). Kunsch, H. R. (1989). The jackknife and the bootstrap for general stationary observations. *The Annals of Statistics*, (pp. 1217–1241).
- (23). Patterson, S. (2014). High-speed stock traders turn to laser beams. *Wall Street Journal*, 11, 1-4.
- (24). Foucault, T. (2016). Quels sont les risques du trading haute fréquence ?. *La Stabilité Financière À L'ère Numérique*. Banque De France. *Revue De La Stabilité Financière*. N° 20. 63-78
- (25). Visnovsky, F. (2016). Le Pilotage Des Risques À L'ère De La Transformation Digitale. Banque de France. Paris. 2-13
- (26). Zaimi, W. (2022). An Empirical Analysis of a Stock Market Index of a Developing Country: Case of the Main Index of the Casablanca Stock Exchange MASI. *Global Business & Finance Review*, 27(4), 1-16.