

Pourquoi les analyses sur les performances des "hedge funds" sont souvent fausses ?

Supposons que vous ayez envie de regarder la performance moyenne des "hedge funds" sur une période donnée, afin par exemple de tester si, depuis 5 ans, ces derniers arrivent à battre le marché. Pour cela, vous allez donc aller extraire les informations sur la performance d'une liste de hedge funds ou d'un indice préconstruit dans une base de données (par exemple le Barclay Hedge Fund Index), et comparer les résultats à la performance du marché. Si vous trouvez que la performance des hedge funds, ajustée du risque, est meilleure que celle du marché, vous aurez tendance à vous dire "ah comme ils sont forts, je devrais vraiment leur confier tout mon argent...". Et c'est là que vous pouvez faire une belle erreur d'analyse.

Dans un monde idéal, "la construction d'un indice représentatif devrait normalement utiliser une base de données contenant des informations sur l'intégralité de l'univers des hedge funds" (source: AMF - Les indices de hedge funds doivent-ils être éligibles ou non aux fonds grand public ?). Cela signifie que si vous voulez regarder la performance moyenne des hedge funds, il vous faudrait donc avoir accès aux performances de l'ensemble des fonds existants... et morts. C'est là qu'intervient un potentiel premier biais dans votre analyse, ce que l'on appelle le biais du survivant. Le biais du survivant est l'erreur statistique que nous faisons quand nous considérons les performances d'une population, en ne considérant pas les résultats de ceux qui étaient présents au départ, mais qui ne sont pas restés jusqu'au bout de l'épreuve.

Mais en quoi le biais du survivant tend-il à surestimer les performances des hedge funds (et donc à biaiser votre analyse) ? Pour expliquer cela d'une manière très simple, le Captain' se permet de piquer un exemple publié sur Merkapt dans l'article "Le biais du survivant dans la création d'entreprise".

"Un exemple classique (et qui est probablement une légende urbaine, mais l'histoire est belle) est celui de la Royal Air Force qui souhaitait améliorer le taux de survie des avions bombardant l'Allemagne pendant la seconde guerre mondiale. Les ingénieurs de la RAF souhaitaient renforcer le blindage des avions sans trop les alourdir, pour qu'ils puissent continuer à transporter suffisamment de bombes. Pour ce faire, ils étudièrent la localisation des impacts de balles sur les avions qui revenaient de mission, pour ne renforcer que les parties les plus exposées. Très rapidement ils comprirent que c'étaient essentiellement les ailes et l'arrière du fuselage qui étaient les plus atteints. Ils prirent donc les mesures de renforcement adaptées. Mais après quelques semaines de raids, il fallut se rendre à l'évidence : il n'y avait pas moins d'avions abattus. En fait, il y en avait même un peu plus. Au vu du sujet de l'article, vous aurez donc probablement compris ce qui c'était passé :

1. Les ingénieurs britanniques avaient simplement fait l'impasse sur les avions qui ne revenaient pas et qui étaient essentiellement abattus par des tirs dans le cockpit ou les réservoirs.
2. En étudiant des symptômes biaisés, ils avaient apporté une solution inadéquate, et même dangereuse, puisqu'au final les avions renforcés à des endroits peu importants étaient moins manoeuvrants et plus facilement abattus par les chasseurs allemands."

Appliqué à notre exemple de hedge funds, le biais du survivant consiste à analyser uniquement les fonds existant actuellement en 2013, et à regarder leurs performances sur les 5 dernières années, sans

considérer les fonds ayant fermés durant cette période. Bien évidemment, cela va entraîner une surestimation de la performance des hedge funds, car cela consiste à supprimer de l'échantillon l'ensemble des fonds ayant fait faillite pour cause de mauvaises performances. Pour éviter ce biais, il faudrait donc avoir un échantillon sur la période 2008 à 2013, avec la performance de l'ensemble des fonds qui existaient en 2008, et non pas un échantillon basé sur les fonds existant encore en 2013). Mais il n'est pas toujours facile d'obtenir des informations détaillées de la part d'un mort...

Les recherches académiques ont estimé que le biais du survivant augmentait les performances de 0,16 % à 6,67 % par année ; le résultat dépendant de la période considérée, de la base de données utilisée ou même de la définition proprement dite de ce biais (source: AMF).

Voilà donc un premier biais, simple à comprendre, mais malheureusement rarement pris en compte. Continuons donc avec un autre biais, le "biais d'auto-sélection". La plupart des bases de données existantes couvrent les fonds dont les gérants ont accepté de communiquer les performances. Les petits fonds ayant des performances peu glorieuses auront tendance à ne pas divulguer leurs résultats. Selon l'AMF, "la performance des hedge funds acceptant de communiquer leurs performances peut donc être significativement différente de celle des fonds refusant de le faire. De ce fait, les bases de données ne constitueront pas d'une manière générale un échantillon non biaisé (au sens statistique) de la population entière des hedge funds, d'où des rendements qui seront également biaisés."

Vous n'êtes pas rassasié ? Vous voulez encore plus de biais statistiques ? En voici donc un autre. Le biais d'histoire instantanée provient du fait que les fonds entrant dans un indice ont le choix d'importer ou non leur historique de performance passée. Et très étrangement, seul les fonds ayant de bonnes performances vont vouloir importer leur historique ; les autres vont bizarrement choisir de le garder secret... Selon, la littérature académique a estimé, l'impact de ce biais sur les performances est compris entre 0,05 % et 4,35 % par année !

Allez un petit dernier pour la route... Le biais de sélection de la base de données. Chaque base de données n'acceptant que les fonds répondant à certains critères (montant minimum sous gestion, nombre d'années d'expérience...), l'ensemble de l'univers des hedge funds n'est jamais couvert et la base de donnée est donc non-aléatoire. "Ainsi, une base de données qui impose une exigence d'au moins deux années d'existence va ignorer systématiquement les jeunes fonds dont la performance serait inférieure à celle de leurs pairs, car ils ne survivront pas suffisamment longtemps pour atteindre le seuil critique des deux années et pouvoir entrer dans la base de données".

Conclusion: "Dites moi quels résultats vous souhaitez obtenir, et je vous dirai quel indice vous devez utiliser...". Cette phrase résume assez bien la possibilité de manipuler les chiffres pour arriver à la conclusion voulue. En choisissant un indice spécifique sur une période donnée et en introduisant volontairement des biais dans la construction de la base de données, le Captain' est capable sans aucun problème de vous prouver une théorie ... et son contraire. Les statistiques et les mathématiques sont essentiels à l'analyse économique, encore faut-il (1) être capable d'utiliser ces outils correctement et (2) ne pas faire la technique de l'autruche en se cachant les problèmes statistiques sous-jacent pour arriver au résultat et à la conclusion désirée.