

法律声明

□ 本课件包括：演示文稿，示例，代码，题库，视频和声音等，量邦科技与小象学院拥有完全知识产权的权利；只限于善意学习者在本课程使用，不得在课程范围外向任何第三方散播。任何其他人或机构不得盗版、复制、仿造其中的创意，我们将保留一切通过法律手段追究违反者的权利。

□ 课程详情请咨询

■ 微信公众号：小象

■ 新浪微博：ChinaHadoop



量化投资概述

量化投资与机器学习系列课程

目录

- 量化投资哲学
- 细分类别
- 经典案例
- 现状与小结

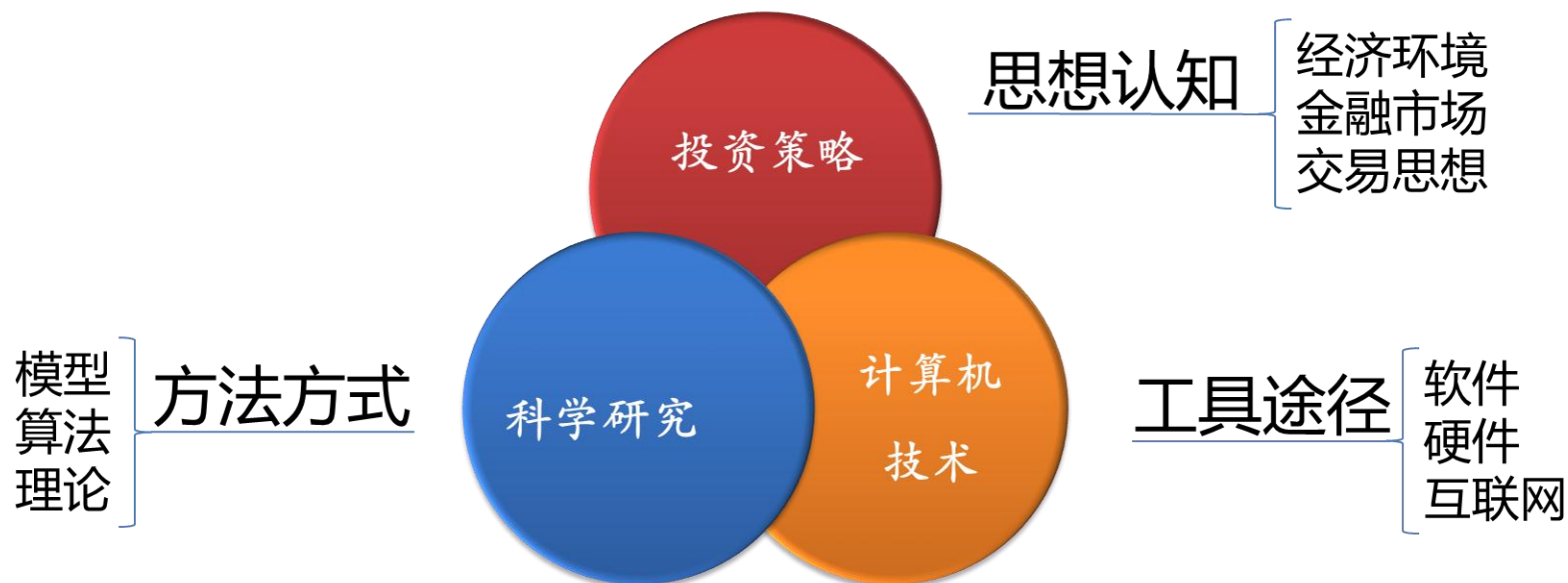


量化投资概述

量化投资哲学

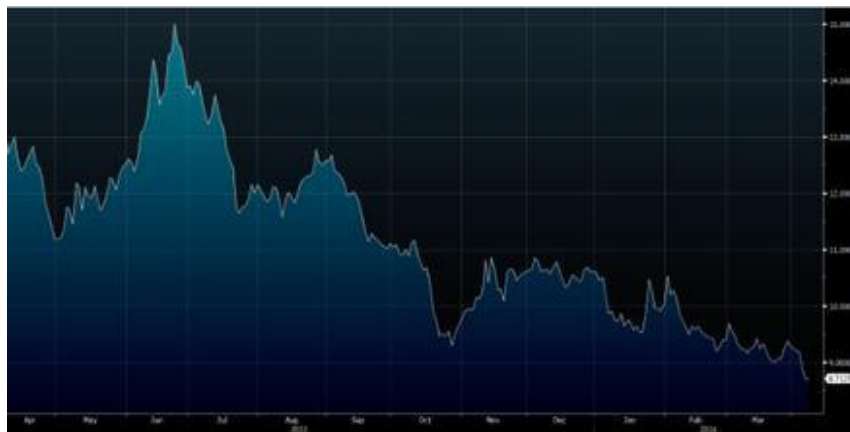


量化投资定义



- 量化投资就是将人的投资思想规则化、变量化、模型化，形成一整套完整、可量化的操作思路，这套操作思路可以用历史数据加以分析验证，并在交易的执行阶段可以选择使用计算机自动执行

量化投资哲学



□ 收益来源

- 所有的投资收益来自市场不同参与方对标的价值看法的时空错位
- 估值的认知差异引起价格波动
- 价格波动创造盈利空间

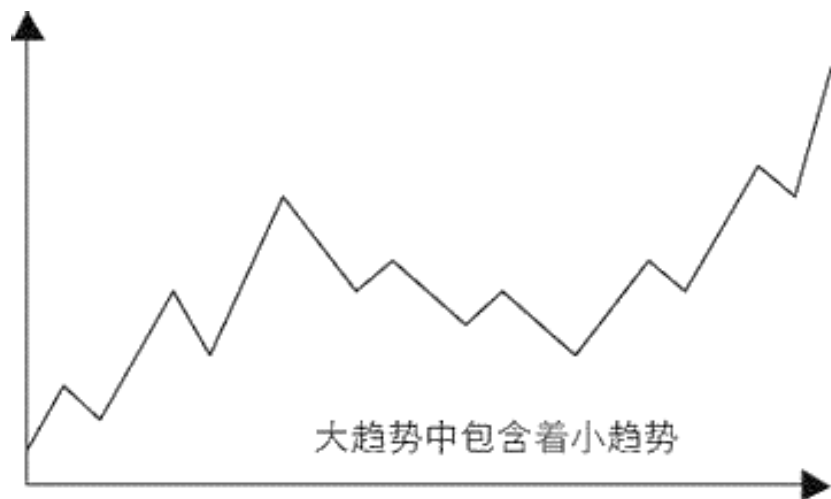
□ 波动分类

- 系统性波动：系统性波动
- 相对性波动：
- 交易性波动：

□ 交易策略开发

- 分类策略研发
- 策略环境研究
- 策略配置优化

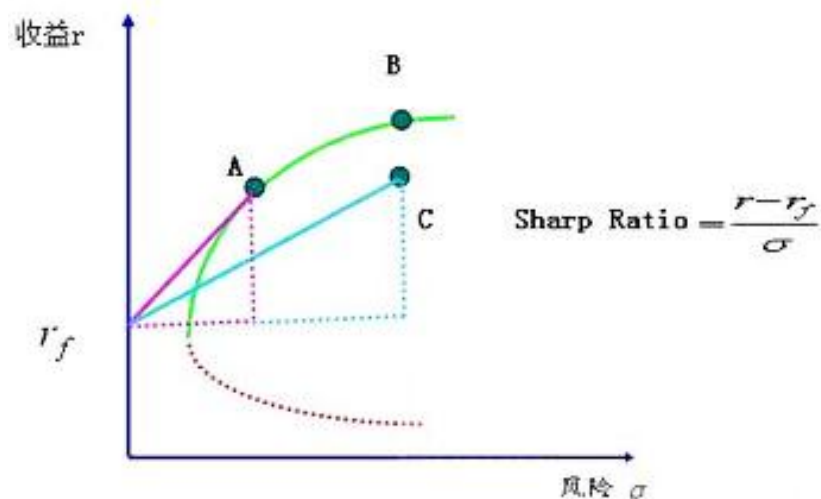
量化投资哲学



□ 效率

- 信息效率【有效市场假说】
 - 价格反映一切市场信息，价格不可预测
- 价格运动效率【技术分析】
 - 价格反映一切信息
 - 历史会重复
 - 价格以趋势的方式运动
- 极化分型效率

量化投资哲学



- ❑ 风险预算
- ❑ 夏普比率：年化收益率除以标准差
- ❑ 定性投资者觉得年化收益率最实在，量化投资者看夏普比率
- ❑ 奥妙在于，如果A策略的收益率低于B，但夏普比率高于B，A策略可以通过杠杆承担和B策略同样大的风险，但收益率会大大提高
- ❑ 夏普比率决定举债能力！

量化投资哲学



□ 胜率

- 出手赚钱的次数与总出手次数相比

□ 赔率

- 平均每次出手赚到的钱除以平均每次出手赔的钱

量化投资哲学



□ 交易圣杯

- 没有包赚不赔的策略，分散是唯一免费的午餐
- 仓位管理和资金管理无比重要

量化投资概述

细分类别



量化投资分类

- 量化投资的门类很多：如同金融包含银行、证券、保险、信托、期货、基金等不同分支一样，量化投资也包括阿尔法套利、低风险套利、统计套利、程序化CTA、高频交易、算法交易等
- 阿尔法套利：做多股票，做空股指期货，获取选股超额收益
- 低风险套利：ETF折溢价套利、股指期货期现套利、期货跨期套利
- 统计套利：包括股票的配对交易、期货的跨期套利、跨品种套利、跨市场套利
- 程序化CTA：用程序化的方法交易商品
- 高频交易：针对盘口机会的及极高频率交易商品
- 算法交易：为了降低冲击成本的算法自动交易

相对价值投资

□ 相对性波动

- 不同标的价格之间的相对波动

□ 相对性价格波动产生相对价值投资

- 各种低风险套利均属于相对价值投资
- 阿尔法套利是相对价值投资的典型手法
- 统计套利是相对价值投资的最直接应用

阿尔法投资概述

量化选股

多Alpha源选择好的股票

对冲

用期指对现货进行对冲

风险管理

仓位、基差、回撤、迁仓

策略逻辑

- 标的选择：确定在哪个池子中选股，沪深300、中证500、中证800、全市场具有截然不同的含义
- 选股因子：用来选择具有超额收益股票的选股因子
- 对冲结构：是否行业中性，是否对冲掉规模、估值等多因素风险



[illegible]

□

-

阿尔法投资概述

对冲套保

行业中性对冲策略净值走势

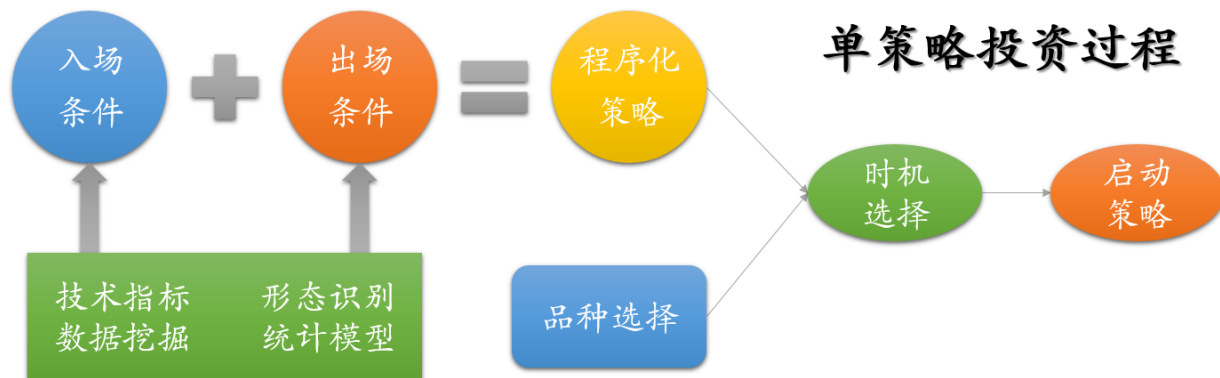


套期保值 (Hedging) 的目的是对冲现货价格波动的风险。传统工业生产、经营和加工企业的套期保值，主要是为了规避原材料或产品价格的波动，利用和自己现货持仓反方向的期货操作，来锁定价格波动的风险，从而获取稳定和合理的企业经营利润

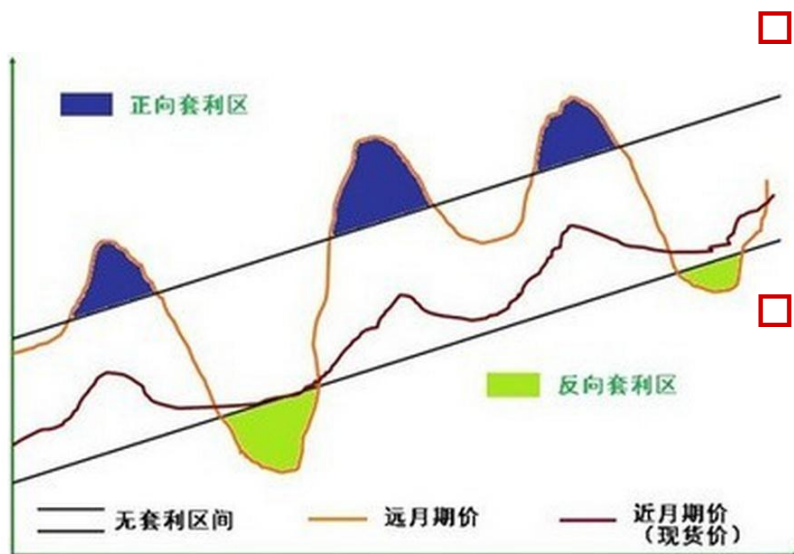
套期保值策略指建构套保组合时不同的优化目标和不同的方法，包括静态套保策略、基差套保策略、动态套保策略以及期权套保策略。为了区分不同策略，首先明确两个不同概念，套期保值比例 (Hedged Ratio)，套期保值比率 (Hedge Ratio)。其中，套期保值比例指被套保的现货价值和现货总价值的比例，即现货被套保的比例；套期保值比率指期货合约价值与被套保的现货价值的比率

程序化交易

- 程序化交易，也叫做系统交易（Systematic Trading），是量化投资大家族中的一员，具体是指将交易策略的思想设计成完整的逻辑运行体系，然后使用合适的计算机语言编写成计算机程序，设置特定的标的和参数之后，使用程序化交易平台运行程序，由计算机进行自动交易。程序化交易的优点是，将交易模式系统化、制度化，排除人性的心理障碍，确保交易方法和操作的一致性。完全由计算机进行操作突破了人力手动交易的诸多限制，保证了交易的及时和高效，使得进行高频率、操作多标的、使用多策略的交易成为可能。



低风险套利交易

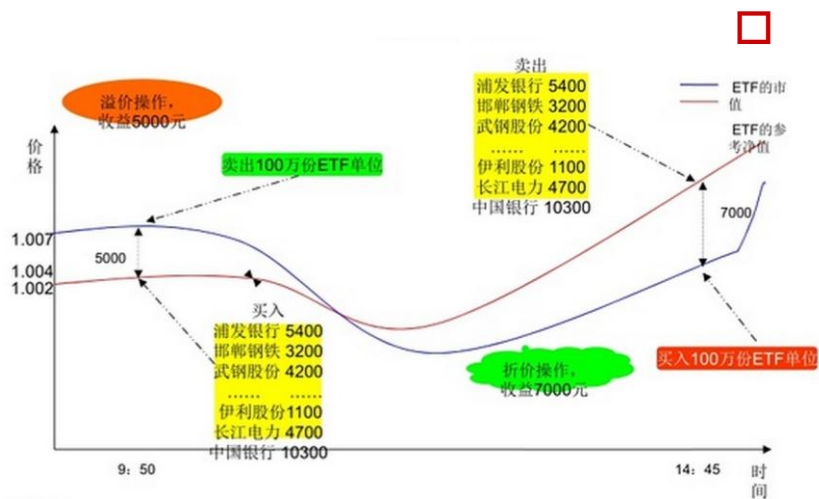


无风险套利模型归纳为：当实际价差>套利成本时，跨期利润=实际价差-套利成本。当利润达到一定程度的时候，进行获利平仓。当两合约价差逆向走高时，到期可进行交割，获取稳定套利利润，达到无风险套利目的。

当然，无风险套利的机会不可能经常出现，但一旦出现，将是一种最稳健的获取收益的方法。期现套利是一种无风险套利，只要收益能够弥补基本的交易成本，就可以在不用关心市场走势的同时获得无风险收益。所以，期现套利的焦点在于对股指期货合约准确地定价从而确定套利的上下边界。

那么股指期货的定价模型为： $F_t = S_t e^{r(T-t)} - D(t, T)$ 。其中， S_t 、 F_t 分别表示股票现货和股指期货合约在时刻 t 的价格， r 表示借款利率， $D(t, T)$ 表示持有时间 $T-t$ 内的股利。理论上，当期指价格偏离模型的定价时便可执行套利。

低风险套利交易-ETF套利



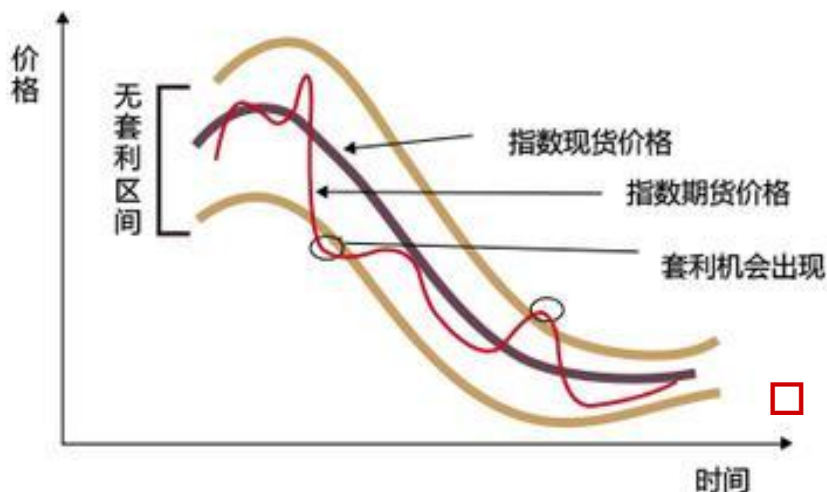
ETF的叫做交易型开放式指数基金，在国外通常被称为交易所交易基金（Exchange Traded Funds，简称“ETF”），是一种在交易所上市交易的、基金份额可变的一种开放式基金。ETF属于开放式基金的一种特殊类型，它综合了封闭式基金和开放式基金的优点，投资者既可以向基金管理公司申购或赎回基金份额，同时，又可以像封闭式基金一样在证券市场上按市场价格买卖ETF份额



由于ETF既可以在一级市场申购赎回，又可以在二级市场买卖，因此当ETF净值和价格出现不合理价差是，投资者可以在一、二级市场间进行套利。这也使得ETF在一定程度上克服了传统的封闭式基金普遍存在的折价问题。ETF套利的原理归根结底源于折溢价的存在，由此衍生出折价套利和溢价套利，其他的套利一、二级市场套利策略本质上都是折溢价套利的发展



低风险套利交易-股指期货期现套利



期现套利过程中，构建能够精确拟合现货指数的投资组合是至关重要的步骤之一。实际的指数复制过程中包括以指数基金、ETF基金和成分股复制为手段的办法。在指数复制问题上，从上世纪八十年代中外研究者都有过许多的研究结论，用指数基金来替代其标的现货指数是最为简单的办法之一，可选的办法包括LOF基金和ETF基金，

LOF在二级市场的交易价格通常会面临升贴水变化的风险，使得套利策略产生很多不确定性；ETF基金复制指数是另外一种选择，但当资金较大的时候会在市场产生较大的冲击成本。成分股复制指数可以容纳较大的资金，成分股复制又分为完全复制和非完全复制。完全复制就是购买标的指数的全部成分股，并且按照其权重配置购买的一种复制方法；当成分股太多，实际操作有难度时，可以采用非完全复制，比如分层加权法、分层抽样法以及优化复制法等。

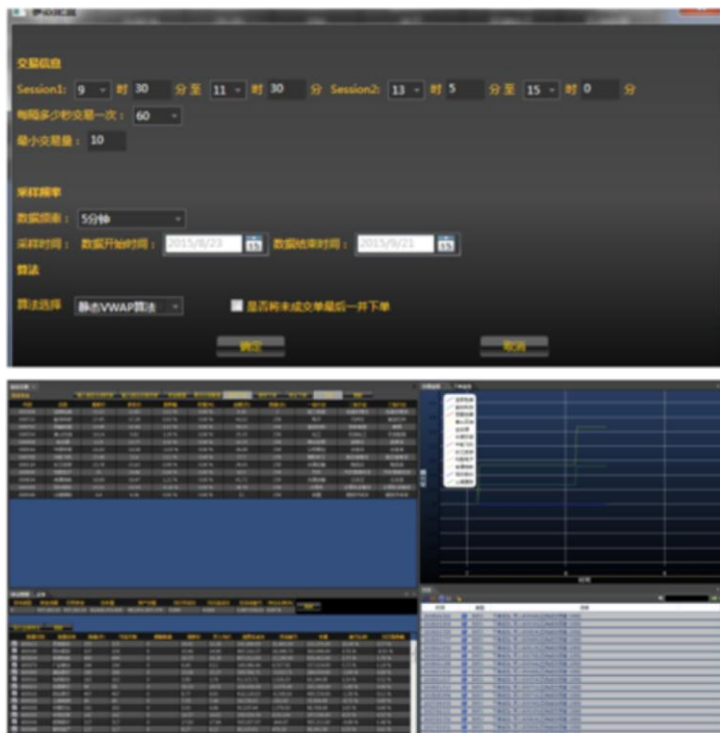
统计套利

- 统计套利交易是当金融产品或产品组合在短期内偏离较长期的期望值时进行相应的交易活动以期望在此金融产品或产品组合回复到长期期望值时获得利润。当一个金融产品的价格序列具有稳定的均值回复特性的时，我们可以在其价格走低到均值之下时买入等待它回升到均值；反之亦然。但是在金融市场中，一般单个金融产品很难具备有效的均值回复特性，所以统计套利交易更多的是应用于包含两个或以上金融产品的组合中，其中配对交易是最常见也是应用最广的方式。在进行配对交易时，最关键的地方是找到两个金融产品，它们之间某种线性关系的价差在“有效时间”内存在均值回复的特性。
- 当一个“价格”序列具有平稳性的时候，它会表现出均值回复的特性，从而创造出统计套利的机会。然而在金融市场中，原始的价格序列往往是非平稳的。所以我们能做的是找到两个（或者两个以上）的价格序列，虽然它们每一个都是非平稳的，但是它们的某种线性组合表现出平稳性。这就是2003年诺贝尔经济学奖获得者Enger和Granger提出的协整（cointegration）的概念，这样的序列对（组）称为协整序列对（组）。

高频交易

- 高频交易（High Frequency Trading, HFT），是指利用计算机系统处理数据和进行量化分析，尤其针对分笔等高级别数据，然后对市场变化做出迅速反映，最终实现资金高周转率的一种日内交易方式。高频交易一般不持有隔夜仓位。其盈利机会来源于市场微观机构层面的不均衡性，往往交易次数多，持仓时间短，每笔交易平均利润小但稳定，相对于传统的买入持有交易方式有更高的夏普比率。高频交易的优点是，收益率极高，当日平仓降低隔夜风险，隔夜资金利息收入降低资金成本，绩效评估周期短，和其它类型策略相关度比较低等。与高频交易相关的另外一个概念是日内交易，即不持有隔夜仓的交易（Day Trading）
- 高频交易是量化投资的一种，虽然必然使用程序进行执行，但从策略逻辑而言，和上面提到以趋势投机和系统交易为特征的程序化交易不同。广义的高频交易可能包含的特征有使用超级计算机生成、发送和执行交易指令；使用服务器托管和特别网络缩短信息处理时间；建立和结清交易头寸的时间非常短，尤其不持有隔夜仓；可能会送大量交易指令，又快速撤单。常见的高频交易策略包括自动做市商交易（Automated Market Trading, AMMs），流动性回扣交易（Liquidity Rebate Trading），闪电订单（Flash Orders）和暗池（Dark Pool）。

算法交易



算法交易 (Algorithmic Trading) 是把一个具有指定交易量、交易时间区间和价格区间约束的交易指令通过计算机程序来完成的交易行为，整个过程中用算法决定交易指令的递交时间、价格、数量和指令的类型。算法交易的目的是为了优化交易的执行，管理交易成本和风险，从而在交易中不落下风。算法交易通过科学的成本估算模型和交易实施算法，将一个大额的交易拆分成系列小额交易在合理的时间点分别执行，以此来尽量减少对市场价格造成的冲击，降低交易成本，而且还能帮助机构投资者快速增加交易量。在国际衍生品市场上，这种交易方式主要运用在各类金融衍生品之间的对冲。在国内，适合的对象包括大小非减持者，大宗交易接盘出货者，“大宗交易-融券卖出”套利者，Alpha套利者，套期保值者，日以上级别程序化交易者等。

大数据基金

- 自2014年10月中证百度百发策略100指数型基金正式发布后，“大数据基金”一词便在基金行业内流行起来
- 大数据基金是利用“大数据”独特的洞察发现力和海量的信息资产建立量化模型，用以选择股票进行投资，合作模式上为业内大中型基金公司+互联网商业巨头，所以诞生时就受到投资者青睐。
- 近日，百度与国金证券联合宣布，共同推出国内首只大数据量化基金——“国金百度大数据基金”，一年多的时间内，市场上已产生10多只大数据基金，版本从最初的跟踪大数据指数的1.0版本上升为现在采用灵活配置策略的大数据2.0版本，未来还将针对高净值用户的对冲专户产品——3.0版本。从发展历程上看，相信伴随科技及市场的发展，未来1-2年的时间大数据基金会全面升级现有的量化基金，促进发展新的量化基金，3-5年的时间内，会成为整个股票型证券投资基金的主流。



量化投资概述

经典案例



相对价值投资方法之殇：LTCM基金

- LTCM基金（Long-Term Capital Management，简称LTCM）由《说谎者的扑克牌》中史诗般的债券交易员John Meriwether于1994年2月建立，巅峰时期与量子基金、老虎基金、欧米伽基金一起被称为国际四大"对冲基金"。在1994-1997年间，LTCM业绩辉煌骄人。成立之初，资产净值为12.5亿美元，到1997年末，上升为48亿美元，净增长2.84倍。
- LTCM基金的创始人早在1986年在所罗门兄弟工作时，就将MIT的物理学博士引入债券分析，为利率期限结构建模，卖出高估的债券，买入低估的债券，进行后来被业内广泛模仿的债券相对价值投资。LTCM基金延续了John早期的这一做法，以寻找各种证券之间的相对价值为投资目标。
- 由于LTCM的合伙人中包括了期权定价BS公式创始人、诺贝尔经济学奖得主Robert Merton和Myron Scholes，增加了其中蕴含的学术色彩。实际上，虽然相对价值投资思想属于量化投资大体系的重要分支，LTCM并未重度使用量化投资模型，他们寻找低估和高估资产主要方法是从基本面出发获取方向性判断，然后根据价差的历史数据进行简单的正态分布建模，并用来测算资金管理。



相对价值投资方法之殇：LTCM基金

- LTCM的结局是于1998年完败，2000年破产清算，并被华尔街银团接管。其失败的主要原因是：
 - 笃信相对价值的走势符合正态分布
 - 相对价值，即价差的判断往往需要对其随机性作出假设，LTCM采用简单频率统计的方式假设其符合正态分布，其结果是严重低估价差朝持仓反方向运行的概率。1997年-1998年发生的各种国际突发事件验证了金融资产价格走势的“肥尾”特征，也就是正态分布假设下的“小概率”事件，其实具有很大的现实概率。
 - 价差走势的误判是相对价值投资最大的死穴。
 - 无限度使用杠杆
 - LTCM基金在投资标的（以应用金融衍生品和融资融券为主）、基金资产、基金公司股权上面分别使用了巨额杠杆，导致公司整体杠杆无比之高，这将资产价格不利走势带来的影响放大了数千倍甚至上万倍。
- 总结一下，LTCM基金失败的主要原因是投资方法有缺陷，进行相对价值投资的时候对价差走势假设过于自信，没有必要的风控和止损设定，过度使用杠杆。

1987年股灾

- 1987年股灾是人类历史上最著名的股灾之一。1987年10月19日，星期一，华尔街上的纽约股票市场刮起了股票暴跌的风潮，爆发了历史上最大的一次崩盘事件。道·琼斯指数一天之内重挫22.6%，创下自1941年以来单日跌幅最高纪录。6.5小时之内，纽约股指损失5000亿美元，其价值相当于美国全年国民生产总值的1/8。这次股市暴跌震惊了整个金融世界，并在全世界股票市场产生“多米诺骨牌”效应，全球市场股票跌幅多达10%以上。这一天被金融界称为“黑色星期一”，《纽约时报》称其为“华尔街历史上最坏的日子”。
- 股灾的产生的根本原因，可能有投资者的“羊群”效应，集体止损引起的“多杀多”，股市泡沫累积到一定程度是自然的价值引力等。从量化投资相关角度，主要有两种策略争议较大。

1987年股灾和投资组合保险技术

- ❑ 投资组合保险技术是Hayne Leland, John O'Brien和Mark Rubinstein于1981年2月创立的一种投资策略，核心思路是让投资组合在风险可控的前提下具有大幅上升潜力，具体手段是对用一部分资产做固定收益投资产生安全垫，作为风险资产的保护。如果对标的价格的随机性做出假设，就可以建立金融工程模型，根据资产价格的走势，使用股指期货动态复制一个看跌股指期货，保护自己投资组合的下行风险。
- ❑ 复制看跌期权呈现的操作就是典型的“追涨杀跌”，价格下跌时要求以加速度的方式迅速减仓，获取头寸安全。

1987年股灾和程序化交易

- 这里所说的程序化交易主要是系统化交易，即趋势投机策略。当系统判定交易信号发生时，进行买卖操作，由于趋势投机的基本入场设定，策略必然跟随市场的上涨或下跌进行相应的做多或者做空。
- 虽然投资组合保险策略和程序化交易技术都是典型的“追涨杀跌”，但是考虑细节，绝非股灾元凶。投资组合保险技术需要动态复制期权，根据检查的频率动态调整仓位，除非是绝对连贯性的下跌才会导致复制策略持续性卖空，如果下跌过程有反弹发生，复制策略同样会积极做多建仓，放大上涨。
- 程序化交易一方面具有组合保险策略的类似功能，一方面不同程序化交易者的策略差别非常大，区别在于K线周期，入场信号和出场信号，除非在绝对连贯下跌中所有策略基本趋同，稍微有波动的下跌中，策略会大相径庭。

Virtue，闪击者和高频交易之争

- 2014年3月11日，美国高频交易公司Virtu Financial向美国证监会SEC提交IPO申请文件，数据显示该公司2013年营收约为6.65亿美元，同比增长8%；净利润为1.82亿美元，同比增长一倍以上。文件中，由于实时的风险管理策略和技术，从2009年初到2013年底，在总共1238个交易日里只有1天出现亏损。
- 无独有偶，前文提到《说谎者的扑克牌》的作者Michael Lewis于2014年3月中旬，出版了一本新书《闪击者》（Flash Boys），书中的基本对高频交易持批评态度。作者主要强调抓住高频交易捕捉微观价差这一事实，为私人交易所鸣不平。但没有考虑高频交易具有不同目的多种策略，另外作为连续交易的市场，高频交易整体上为市场提供了流动性，和其对应的风险补偿。
- 2014年4月15日，欧洲议会通过包含一系列限制高频交易措施的《金融工具市场指令II》，内容包括限制报价货币单位过小，强制对交易算法进行测试，要求做市商每个交易日每小时上报交易额，以及当价格波动超过一定限制时的熔断机制（价格增量规则标准）。法令所涵盖管理范围包括股票市场、衍生品交易和各类新型交易平台。根据欧盟的立法机制，欧盟层面通过法令后，其将会被下放到各个主权国家等待签署，当时预计法令实际生效时间将在2016年底，个别条款会给予做市商更长的过渡期。
- 随后，Virtual公司宣传无限延期IPO。



高频交易的天敌：幌骗交易

- 幌骗交易（Spoofing）采用和高频交易类似的技术手段，但并非利用公开信息进行策略性盈利，而是以哄骗交易对手、操作市场为目的获取利润。常见的做法是，比低于市场卖价的价格挂出卖单，这样其它卖家就会被迫挂出更低的价格以寻求快速成交，当发现更低的卖价后，幌骗策略迅速撤单，反手做多，这样通过诱骗交易对手以更低的价格获取头寸。用高于市场买价的价格挂出买单诱骗买家，以更高的价格结清头寸同理。
- 幌骗交易完成的时间极短，对于大部分手动交易者是不敏感的。但是对于高度依赖盘口信息捕捉市场微观机构的高频交易策略而言，幌骗交易几乎是他们的天敌。2010年美国奥巴马签署多德-弗兰克（Dodd Frank）金融改革法案后，“幌骗”被明确为违法行为。但在美国市场中这种行为依然猖獗。

幌骗交易案例

- 2014年11月，美国商品期货交易委员会（CFTC）发布公告，芝加哥投资公司3 Red Trading LLC以及交易员Igor B. Oystacher涉嫌利用“幌骗”（spoofing）手段及欺诈设备操纵市场，对其发起诉讼。据CFTC的指控文件，Oystacher在359790份交易合约中累计进行1316次“幌骗”交易。同期，CME向Oystacher处以15万美元的罚款和一个月的市场禁入，Oystacher同意支付罚款并接受处罚，但对于违规行为既不承认也未否认。后来据金融博客Zero Hedge评论，Oystacher通过发现僵化的高频交易市场漏洞，从而操纵了其他的高频交易者。
- 2015年11月3日，美国联邦法院裁定Panther Energy Trading公司的负责人Michael Coscia在商品交易欺诈以及幌骗(spoofing)罪名成立，这是美国2011年多德-弗兰克规定出台以来，关于其中“防欺诈法规”的首个案例，也是全球的首宗刑事起诉。
- 在庭审中，来自美国证券交易委员会（SEC）和美国商品期货交易委员会（CFTC）的证人提供的相关数据表明，嫌疑人自2011年以来在期货市场挂出大量买卖单，而事实上这些买卖单的目标不是执行，而是制造需求假象，诱使其他交易员入市，从而让自己从中获利。比如嫌疑人常常在挂出大单之后撤单，但对小单撤单的概率较小。检方则指控嫌疑人在三个月的时间里通过“诱饵调包阴谋（bait-and-switches cheme）”非法获利140万美元。
- 最终法院裁定Michael Coscia六项商品欺诈和六项幌骗罪名全部成立。据悉每项欺诈罪名的最高刑期都是25年，外加25万美元罚款；而幌骗罪名的最高刑期是10年，外加100万美元罚款。



量化投资概述

现状与小结

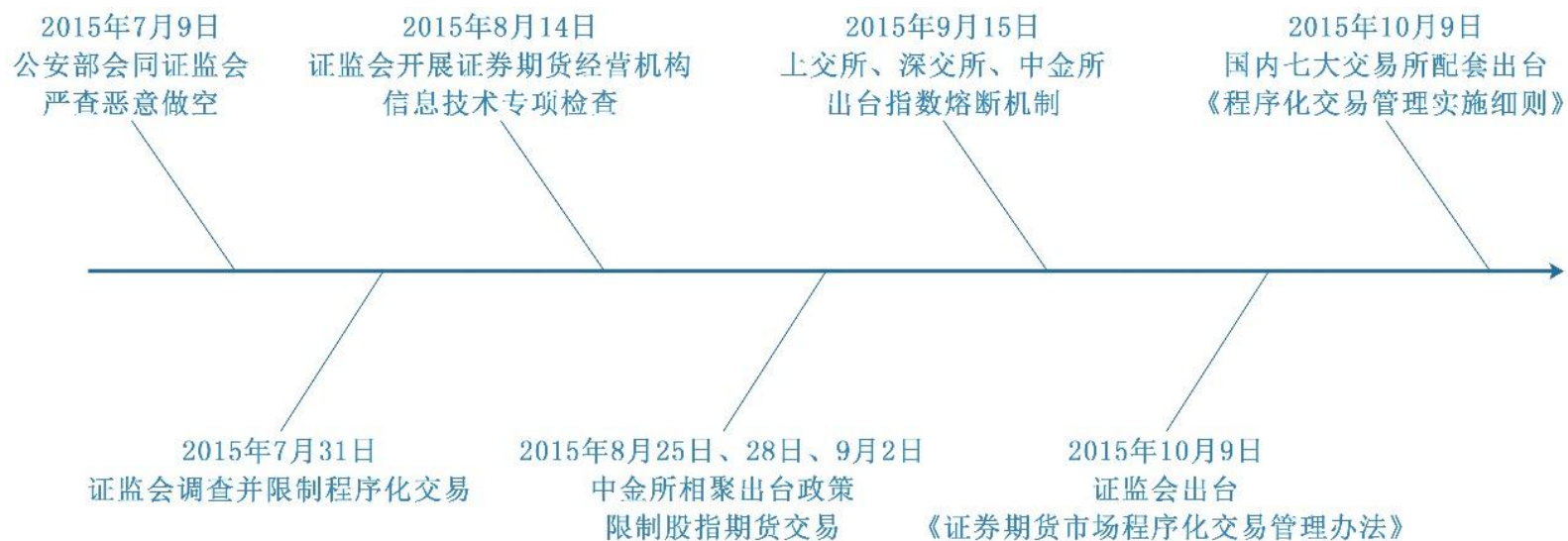


中国量化投资现状

□ 公募基金以指数增强和股票型大数据基金为主

- 截至2017年3月底，72家公募基金公司共设立591只指数型基金，总规模5693.53亿元，其中466只纯指数、63只增强指数及62只QDII指数型基金，资产净值分别为4767.64亿元、503.89亿元和422.00亿元。
- 26家公募基金公司共设立77只主动量化基金，总规模为989.99亿元，主要分为量化选股型和对冲策略两大类，两者规模比接近2:1，从基金公司分布来看，长信、博时、华泰柏瑞、摩根士丹利华鑫和华商基金位于前五名。5家公募基金公司共设立25只大数据基金。2017年3月31日，25只大数据基金资产净值的和为239.83亿元，长信基金独大，占比近24.23%。

“股灾”回顾



中国量化投资现状

- 私募对冲基金以阿尔法投资为主，CTA投资为辅助
 - 股指期货限仓造成流动性锐减，阿尔法和CTA投资均遭受较大打击
 - 高频交易受到较大影响
 - 《证券期货市场程序化交易管理办法(征求意见稿)》出台，程序化CTA未来会受到监管，喜忧参半
 - 量化选股和量化资产配置不受影响，成为私募对冲基金重点发展的对象

疑问

□ 问题答疑：<http://www.xxwenda.com/>

■ 可邀请老师或者其他人回答问题

□ 量化投资策略研发平台：<http://www.daquant.com>

相关微信号推荐



量化投资华山之巅



量化投资大家学



联系我们

小象学院：互联网新技术在线教育领航者

- 微信公众号：大数据分析挖掘
- 新浪微博：ChinaHadoop

