

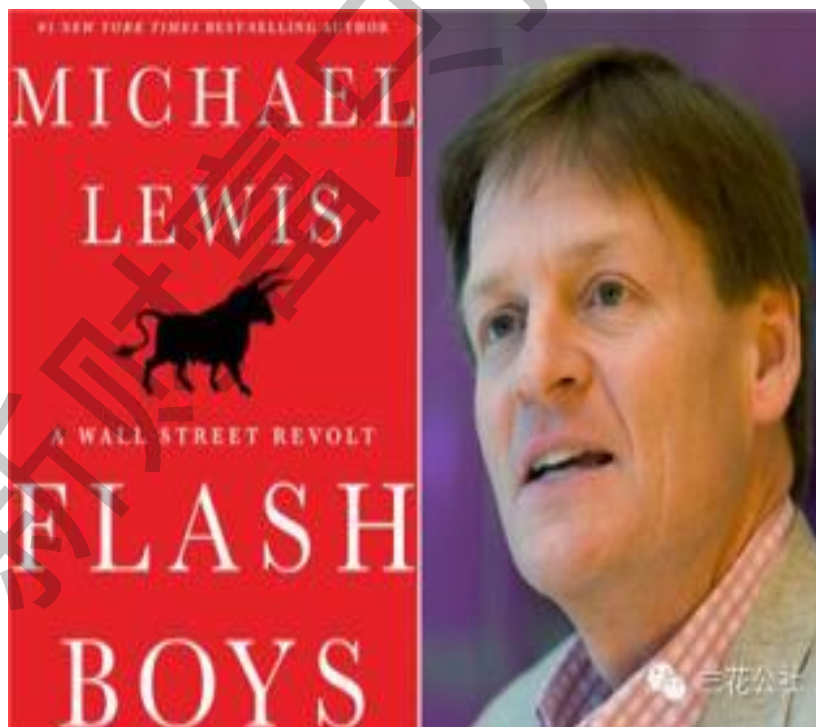
会议主题：方正金工Q培训第三期  
会议内容：期权做市与高频交易  
主讲人：中粮期货董事总经理 黄敏  
时间：2017年6月10日

## 期权做市与高频交易

黄敏

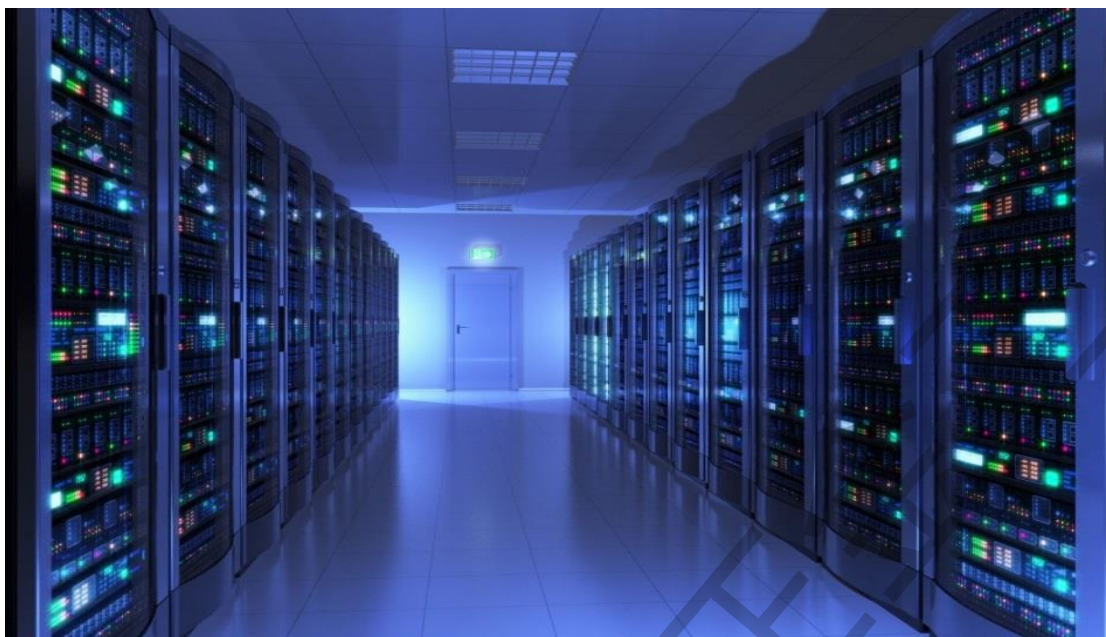
各位早上好，今天我的题目是期权做市与高频交易，我之前是在美国从事个股期权的做市商。先介绍一下美国和国内期权的对比，国内的期权品种目前有三个，一个是2015年2月9日上市的50ETF期权，第二个是今年3月31号大商所的豆粕期权，这是国内第一个商品期权，第二个商品期权是4月19号的白糖期权。这些日子我都记得比较清楚。因为是中国期权里程碑式的日子。期权本身是量化交易非常细分的领域，而期权做市商是在期权交易里面更细分的领域。大家了解一下期权做市商是怎么做交易的，了解自己交易的对手方交易的模式，可能对在座的各位有所帮助。

先介绍一下这两本书，第一本是《Flash Boys》，这本书引起了监管机构和普通民众的关注。第二本书是业内人士写的：《Flash Boys :Not So Fast》，他们就像三国演义和三国志的区别，大部分人都知道是三国志是真实的，但是大家更愿意看有更多虚构的情节的《三国演义》。其实我们国内不用说高频交易了，程序化交易也被限制的非常严重，很大程度上来源于国内过于严厉的监管，就像我们说15年国内股灾和股指期货可能有一定的关系，但是并没有那么紧密，或者说股指期货导致了股灾，那就更没有必然的联系了，但目前监管的一些限制股指期货的措施导致现在交易股指期货有了较大的困难，也同时限制了量化交易和衍生出来的程序化交易。



介绍一下我在美国的经验，我05年去的美国，在当时我们的脑海里，虽然05年也不算非常古老的年代了，但是之前华尔街交易的场景应该是电影里这般，交易场内人声嘈杂喧哗，大家都在喊价格做交易，后来我去参观了CBOE，和第二张图差不多，其实人并不是非常多，那时候交易场所里面基本上没有什么工作人员，所有的工作人员都到了电脑后面，这个才是我们真正的交易公司，全都是机房了，一台台交换机、电脑、服务器，你基本见不到人声嘈杂的环节，自营交易的交易室和图书馆差不多(投行面向客户的flow trading交易室会是另一番景观)，大家都在埋头算东西，写策略，盯着屏幕做交易。如果大家去芝加哥参观交易所的话，CBOE其实没有什么好参观的。现在唯一可以参观的交易所只有CME了，CME有一两个pit保留了原来的模式，还是有人在里面交易，其实没有太大的意义了，因为大部分的交易还是转移到电脑上完成的，保留一两个pit更多是为了保留一些传统，或者老一代的交易员还在里面，估计再过5年也会被取消了。所以看美国金融交易的发展，高频交易或者自动化交易应该是二级市场中国未来的方向，但是具体要走多久还不好估计，美国也走了一段时间。美国单从股票来说的话70%的交易都是电脑自动完成的，如果你有一个美国的个股帐户你要做交易的话，完成一笔交易你可以大概知道你的对手方不是按鼠标的人而是一台电脑。





高频交易和速度非常有关系，中间有一个误解，很多人会认为高频交易就是高频率的交易。我们工作的地方是上海期货大厦，有很多期货炒单手每天会和我们一起上下电梯，他们可能一天完成的交易量比电脑还要大，一整天都要点鼠标，买进卖出，忙碌的一天可以完成上万笔交易，频率确实非常高，但是这并不是我们定义的高频交易。其实在古书也记载了，我们经常说的一刹那，一瞬到底代表多少时间，一瞬间的话，360毫秒，在往下是一刹那，一刹那是18毫秒，再进一步就是到微秒级别，但国内交易所的基础环境和硬件环境还没有达到那么快的速度，可能会达到微秒，但是再往下到纳秒就不行了。但在美国就可以看到一些技术比较先进的交易所，他们处理一些行情的速度可以精确到纳秒级别了。所以在技术方面和交易所来说我们和美国还是有不小或者说很多年的差距。

## 刹那与须臾

- “一刹那”，在古代印度梵典《僧祇律》（又有资料是《倡祇律》）有这样的记载：
  - 一刹那者为一念，二十念为一瞬，二十瞬为一弹指，二十弹指为一罗预，二十罗预为一须臾，一日一昼为三十须臾。
- 一天一夜24小时有
  - 480万个“刹那”
  - 24万个“瞬”
  - 1万2千个“弹指”
- 一昼夜有86400秒
  - 一“须臾”等于2880秒（48分钟）
  - 一“弹指”为7.2秒
  - 一“瞬”为0.36秒
  - 一“刹那”却只有0.018秒。



1秒等于1000毫秒等于100万微秒。在美国，如果交易员一直在发订单，如果10微秒完成一笔交易，那么一秒钟就可以完成10万笔交易了。而在国内有一定的限制。在美国，订单簿上有任何一笔更新都会推送给我们，比如说你看到买单数量变化，有人加入买价，看到买单的量变大了，市场上并没有交易。或者说非最优报价上的卖单增加了，它都会推送给你，另外任何订单删除或者增加了都会推送给你。国内的话，期货交易所比证券交易所技术更先进一些，证券交易所的推送更慢一些，期货交易所每半秒推送一次的切片数据。这半秒可能在不同价格上成交了很多手，但它只推送给你半秒切片时间内最后一笔成交价和这半秒成交了多少手。其实你也很难看到半秒的中间发生了什么事，打个比方，比如说李小龙和对手武打的电影，你看电影的时候可以很连续很连贯地判断谁赢谁输，谁被打了几拳，踢了几脚之类的。但是如果对方不给你看这部电影，他把每分钟末尾的画面切一个画面给你看，你用这个画面来判断谁赢谁输就比较难了，比如说那一分钟正好是李小龙被对方打了一拳踢了一脚，这个比喻可以很类似的判断切片数据的影响。切片数据是500毫秒一次，大家收到数据都是500毫秒的，国内的环境，大家如果要做高频交易的话，每500毫秒那一刻大家都收到了最新数据，大家都在同一时刻更新模型的数据判断下一时刻的市场趋势，开始发订单和做交易，所以在那个时刻网络非常拥挤，能不能抢到订单就看谁速度更快了。如果国内提供tick数据画图的话，就能看到一个很明显的峰值和谷值的出现，每隔500毫秒会有一个下单成交的峰值。如何探测500毫秒中间发生了什么不同的交易员会有不同的手段，多掌握信息就会比只知道末尾切片数据的交易员占有优势。比如交易员A和B都收到T0和T1的切片数据价格都是100，看似500毫秒的时间段没有变化，但交易员A如果在T0到T1的时间段做了额外的操作，发出比如一手的不同价位99.9, 99.8, 99.7..的买单和100.1, 100.2, 100.3的卖单并被成交了数手买单，他就知道那一时段市场交易的倾向性和短暂方向。这当然是个简化的例子，另外可以做一下对比，一般来说眨一下眼是300毫秒，但是对高频交易来说，300毫秒是属于非常缓慢的速度了。但国内我们经常遇到如此缓慢的速度，比如说我下订单的时间，和交易所收到订单回馈给我的时间，用这个时间差来衡量你的网络延时有多少。在美国这个是在微秒级别，在国内网络延迟比较大的话，经常是几百毫秒甚至数秒了。可能过了一两秒，你还没有收到回馈但是忙碌的市场环境下期货的价格已经跳了10几个点了，我肯定已经不想卖了，于是再发一个取消的订单。结果就是后来收到了两个订单确认，一个是你之前发出的订单成交了，第二个是第二次发出的订单被打回因为要取消的订单已经被成交了。高频交易其实很多地方会讨论到延时，有各方面的延时，刚刚说到一个是我发订单到交易所收到订单的网络延时，后来为了减少网络延时就把服务器放在网络机房里面，这样就不需要那么长的时间了。另外有本身系统软件的延时和另一边交易所的延时，交易所系统自己处理订单也需要一段时间，并不是收到就能马上处理完了。有些地方是我们可以控制的，但有些时间我们没有办法控制，只能靠交易所提高网络软件环境和基础硬件设施。我们可以控制的，比如有些人的行情就比你快一点，不管是网络高效还是因为他有专门的更快的行情提供商。在高频交易的领域里差个10毫秒或者100毫秒的感觉就是瞬间和永恒的区别了。国内也有单独提供行情的服务商，你不能通过它做交易，但是可以通过它看到行情，有些人会专门订阅它的行情。可能行情的速度会快100微秒，这样对你的策略可能会有天翻地覆的区别。打个比方，如果你做中低频交易，比如一天交易一次，如果你看到的价格是昨天的价格的话，这样肯定就非常劣势。现在讨论交易频率快一些极致一点的话，如果你能将看到行情的速度比别人快100微秒或者50微秒，那么你的系统更新决策和下单也会更快一些，更有优势一点。

# 速度

- 1秒=1000毫秒=100万微秒 (second/millisecond/microsecond)
- 1微秒=1000纳秒(nanosecond)
  - 美国证券/期权交易所处理速度精确级别nanosecond
  - 高频交易执行时间在微秒级别甚至纳秒级别讨论
  - 国内的技术瓶颈和机制差异(毫秒级别, 至多微秒)
  - 美国交易所订单处理速度10微秒,国内每秒报单数量和全天报单数量限制
- 一次闪电过程 = 250毫秒, 眨一次眼睛 = 300 毫秒
- 假设一次高频交易完成时间100微秒
  - 一次闪电过程可以完成2500笔交易
  - 眨一次眼睛可以完成3000笔交易
- 延时(latency)
  - 传输, 处理
  - 行情, 交易
  - 硬件, 软件

然后再说一下期权交易, 期权交易是量化交易里面非常小的一个子集。华尔街最早的quant就是从期权定价发扬光大的, 那时候需要大量做物理、数学的人才, 期权的quant主要是把目光集中在定价这块领域上。但是期货交易和股票交易等, 这些quant可能集中在统计中发觉一些规律, 这是不同类型的两种quant。高频交易在期权里面又是非常小的子集, 做市商肯定需要高频交易, 因为我们有这个义务给全场的期权提供报价, 根据标的物的变化不停地更新我们的价格, 我们相当于是其他客户的天然对手方。如果不是做高频交易的话完全不可能做做市商交易, 我们看到在80、90年代美国期权交易所那边还有很多交易所floor上的场内期权交易员, 他们不持仓, 或者中性持仓。他们买进卖出一个期权马上对冲。在那个年代你反应要快还是可以做做市商交易的, 但是到了现在这个时代, 像美国有几千支股票都有期权, 如果一个人管一个品种的话, 脱离计算机你至少需要招千个期权交易员来帮你做这个业务。实际上我们当时在美国是4个人管理1000支期权产品的交易, 所以全部都是电子化的。美国不管是主观交易还是量化交易, 基本都在一个量化框架下, 不管是做股票期货期权各种其他衍生产品, 还是风控, 选股, 都有一个量化的框架帮助你衡量。量化交易在美国是非常成熟的体系, 在国内起步时间比较晚。并没有像美国那么成熟。

下面简单比较下量化交易, 算法交易, 程序化交易, 高频交易。算法交易是我们投行经常听到的, 比如一个客户希望将10万股票卖出去, 一下子同时卖出对市场的冲击太大, 所以需要算法交易的帮助, 用不同的方法在不同的时间投放不同的量, 帮助你在一天之内完成10万股股票的卖出, 是一种衡量对市场影响, 优化下单达到最优平均成交价格的方法。程序化交易如果用来下单的是交易程序而不是手工都可以称之为程序化交易, 比如用客户VBA写了个股票下单的程序并能通过券商接口下单那也是程序化交易的一种, 程序化交易相当于是高频交易上面一层, 高频交易是程序化交易下面的一个小的子集。高频交易未必需要成交高频率, 比如说国内市场豆粕期权和白糖期权的成交频率非常低。但是我们更新订单的频率是非常高的, 我们不停发新的订单给交易所, 不停的撤回老订单, 虽然这些动作不会产生交易, 但还是高频交易。但刚刚说的期货炒单手, 他成交非常高, 但是他不是高

高频交易，高频交易必须是和电脑和程序化交易是有联系的。在美国做期权做市商高频交易原来最主要的利润来源就是买卖价差，低价买进，高价卖出，后来发展到08年股灾之后，11年、12年，已经没有办法靠价差赚钱了，买卖价差变得非常非常小了，像美国交易最活跃的期权，标普500SPY等，价差只有1分钱，你哪怕买量和卖量是均衡的，你能抢到这一分钱的价差，但是扣除人员成本、服务费之类的，你肯定不会赚钱。所以现在美国做市商赚钱的模式变成提供流动性赚钱，在很多交易所上我给交易所挂单提供流动性成交是不需要支付手续费的，并且交易所会给我一个返还，比如当前的买价是100元，我提供100元的买价，当前的卖价是101元，我提供101元的卖价。如果再当前价格由成交手续费由主动成交的一方来支付，因为他从市场上取走了流动性，而我提供流动性不需要支付手续费，而且我可以得到对方付的手续费当中的一部分作为rebate补偿。所以后来交易所最大的收入来源也不是收取手续费，它最大的收入来源是服务器共置的那一块，各家交易股票的公司都要把服务器放在交易所，所以交易所的机房非常非常大，相当于几个足球场大小，而且设计都是抗八级地震，有非常多的安全措施。美国交易所是非常分散化的，不像国内交易一个股票必须在上交所或者深交所，美国不一样，你可以在股票挂牌的交易所比如NYSE交易，也可以在纳斯达克等等其它的股票交易所，甚至可以在券商提供的暗池交易。还有一些OTC的交易中心，所以股票交易不像国内的单一交易所环境这么垄断。

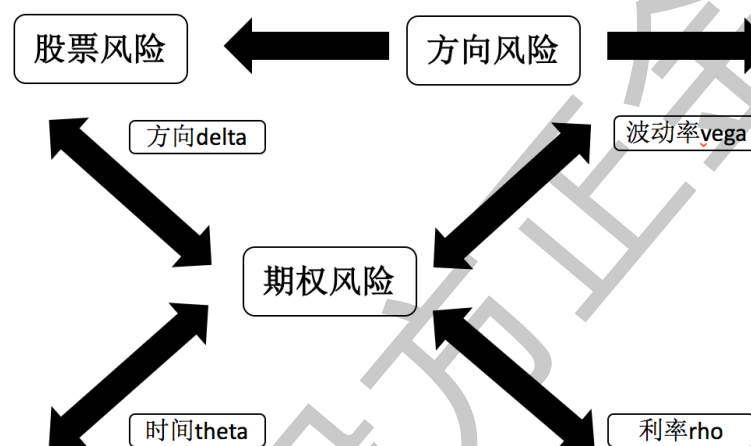
## 量化交易，算法交易，程序化交易，高频交易

- 市场价格/量，技术指标，财务数据，财务指标，风险指标，etc
- 数学量化模型
- 计算机程序实现
- 策略：高频率，低频率，主动交易，被动交易，短期，中长期，etc
- 买卖价差趋动盈利-> 流动性驱动盈利
  - 交易所
  - OTC交易中心ECN
  - 黑池

期权交易所也是这样，我离开美国的时候有12家期权交易所，现在应该有更多了。同样一个期权，比如IBM或者微软的期权，你可以在一个交易所买入，在另外一个交易所卖出，我们做市商会很关注这件事，因为不同交易所的手续费不一样，连接速度不一样。但是其它客户不用关心从哪买从哪卖这些事，经纪商会管理好，客户只要知道手中净持仓和风险就好。为什么我们要做期权呢？期权相对股票或者期货提供了不同的风险结构。做股票只有两个方向，看多或者看空，国内不能融券还只能看多。做股票说拔河一样两个方向。但是国内只能像赛跑一样，选对正确的股票一路往前。但是期权就不一样，它就像足球或者篮球，如果你有方向观点的话你可以交易方向，如果没有方向观点的话，可以交易股票的隐含波动率或者实现波动率，下注将来一段时间股票价格的波动是变大变小。还可以卖保险，就像巴菲特在股灾的时候卖了标普500二十年后的期权，毕竟他有充分的资金来支付保证金。20年后理性的人也会判断，健全的股市大概率会上涨，虽然20年后巴菲特大

概率看不到那一天。至少美国股市是这样。国内的股市可能未必，10年前买入指数十年后也许没有什么变化。另外期权可以交易利率，美联储在08年股灾之前的基准利率是4~5%的水平上，但是聪明的交易者已经意识到了股灾发生后，美联储会不停地降低利率，你可以通过利率期货去交易这个东西，也可以通过期权，因为期权有一个变量是要把利率，如果做市商放入了一个当前的利率，或者他认为的较高的利率在里面，你可以通过买卖期权来构成一个你对未来利率会降低的观点和看法来获利。

## 期权的多维风险



### • 赛跑，拔河，足球

那么我们简单介绍一下美国期权市场，从73年开始CBOE正式交易场内个股期权，比国内早了40多年。所以国内目前的发展速度相对美国还是有后发优势的。我们50ETF期权日交易量最高的时候到了170万手，这是非常大的数量，放在全球来看的话，单个品种170万手，可以排到前10名的水平。后面我会对比一下。亚洲除了日本，还有几个大的市场，比如印度、韩国、台湾，台湾排全球第6大市场，亚洲主要交易的期权都是指数期权，没有什么个股期权期权，和全球比例正好相反。台湾的数据是平均60万手日期权交易量。50ETF期权应该不用太久就能超过日均60万手了。然后我们就会变成全球第6大期权市场。所以这样看来有期权的市场并有有意义成交量的全球也并不是很多。

这是刚刚提到的美国的12个交易所，相互竞争，国外期权交易所和客户之间是非常友好平等的关系，经常有什么需求、交易或者想法，给交易所打电话，交易所会积极地和会员探讨。ISE是在纽约的期权交易所，期权上ISE单独的一个订单本叫组合订单本。可以让交易几个期权leg交易在一起的组合，而且这个组合订单本交易占了ISE期权交易量的30~40%，这是非常大的交易量。我们作为做市商之前一直没有参与后来想了解这个部分，ISE知道了这个需求第二天就派了一组人来我们芝加哥的办公室介绍各项细节，提供了非常详尽的数据，告诉我们怎么参与，需要技术支持的联系是谁。国内相反，我们经常会接到这样的电话，明天来我们这里开个会吧，特别是商品交易所都是在郑州或大连这些和上海相隔千里的地方，过去一次势必影响一天的交易。国内的交易所在和会员沟通和满足需

求方面和美国交易所还是不一样的，这个也是缺乏了交易所间竞争的结果。再看一下交易公司的技术差别。比如说微波(Microwave)和激光传送信号(LaserRadio)，可能国内都没有实际运用过。美国某些交易机构觉得光纤传输的速度太慢了，我们国内现在的网络速度都是4M、8M的管理专线或者100、200M的互联网专线。在很多年前我们芝加哥连接到纽约交易所机房的网络已经是1G的互联网光纤了。但是在美国交易我们觉得1G的互联网光纤还是太慢了，有财大气粗的高频交易公司会把军方退役的微波发射塔买下来，这样就不用走光线线路了，因为光线辗转的速度会比光速衰减1/3，而微波的话直接通过信号传播，速度会更快一点，在常人的感官来说其实这个速度差不了多少，但是对高频交易公司来说是非常大的竞争优势。后来在14年的时候，美国交易公司觉得微波受天气影响过大，开始用激光来传送了，在交易技术美国比我们还是领先至少5~10年。

## 美国期权市场

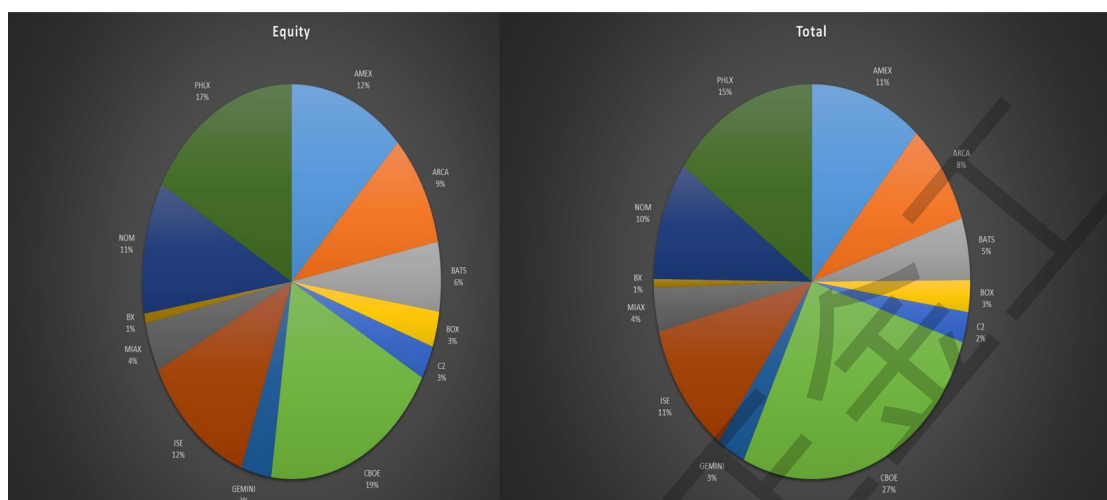
- 1973年CBOE和OCC建立
- 第一个全电子化交易所Eurex
- 2000年ISE，美国第一个全电子化期权交易所
- 目前美国12个互相竞争的期权交易所,OCC提供最终担保: AMEX (NYSE),ARCA (NYSE),BATS,BOX,BX (NASDAQ),C2 (CBOE), CBOE,ISE,GEMINI (ISE),MIAX,NOM (NASDAQ),PHLX (NASDAQ)
- size(1,10), pro-rata/price-time, NBBO/OPRA, Microwave/LaserRadio
- 期货期权: CME, ICE

期权市场在美国是交易非常大的市场，期权交易量和期货交易各占半壁江山。在国内的话，期货交易量和股票交易量相比差很多，期权和期货就差的更多了，例如豆粕期权和白糖期权大概一天交易量每个品种单边是1~2万手。而期货交易差很多，豆粕期货交易量一天大概2~3百万手，所以发展空间还是很大的。目前国内交易所各种各样的限制把期权产品控制的很严格，任何一个散户要开户要通过非常严格的考试，机构要开户也要通过考试。考试是一个门槛，之后还有很多的交易所限制和门槛，所以交易客户的数量非常少，每天相当部分都是做市商和做市商之间交易，特别是国内这两个商品（豆粕和白糖）都有夜盘，做夜盘的话大家会发现这个电脑3~4分钟都没有动静，往往以为电脑屏幕坏了，或者出去喝杯茶，上个洗手间回来看到屏幕上还是一笔交易都没有。但确实交易就是这么少，并不是电脑坏了。当时韩国的期权KOSP期权是全世界最大的交易品种，在美国做市交易股票期权我们还可以一笔笔看过来，但是看韩国的KOSP期权是看不过来的，滚动速度一秒钟之后前面100笔交易都滚动掉了，那个就没有办法用人眼去交易了，只能靠程序化交易。

美国虽然分12个交易所，但是CBOE还是有先发优势的，占据了25%以上的市场份额。CBOE还有自己的专利的期权，比如VIX。这些专利期权其他的交易所是不能交易的。

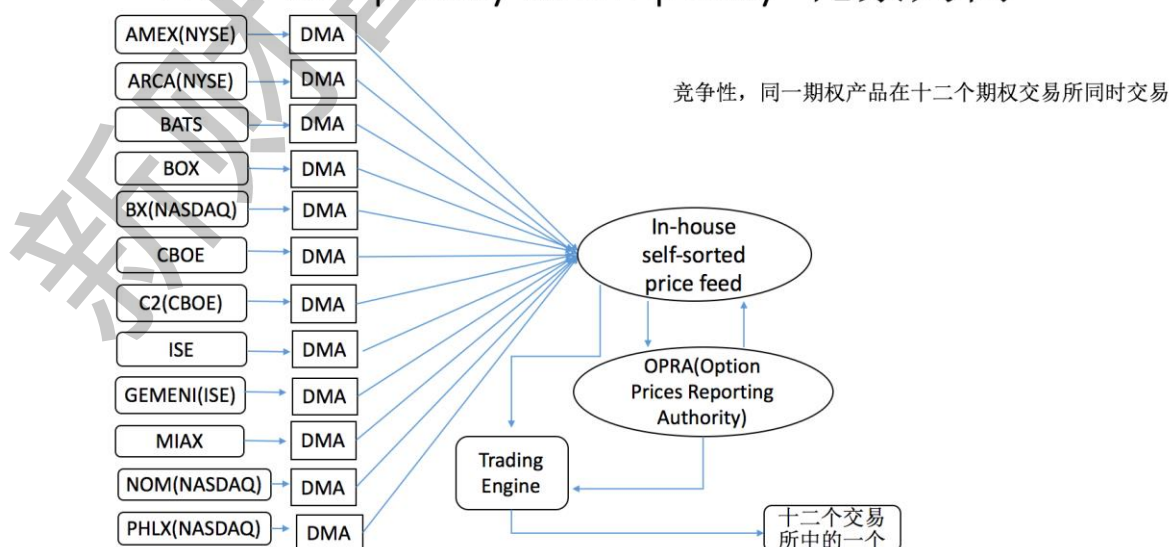


## 个股/ETP期权总成交量占比 CBOE包含指数期权总成交量占比

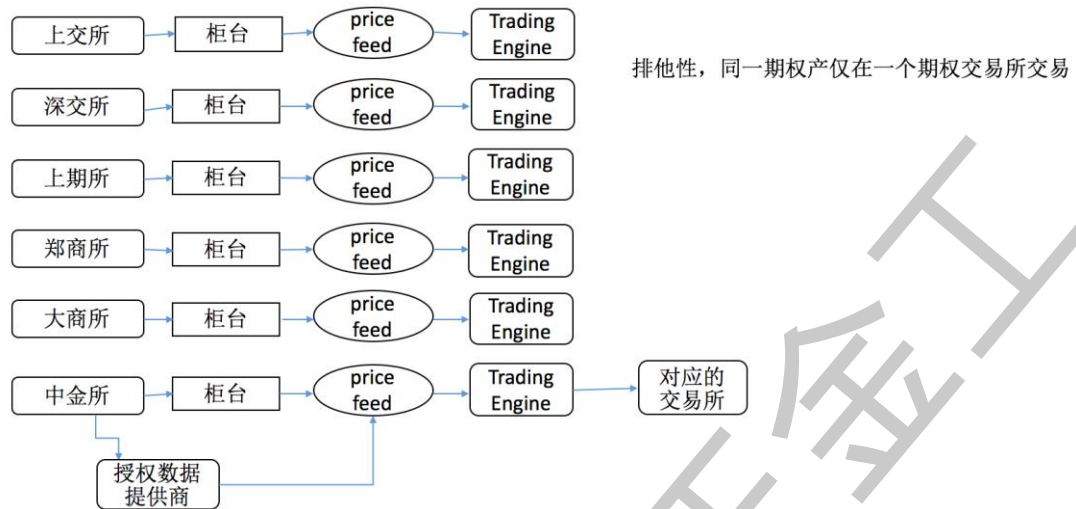


在美国作为做市商所需要的系统要求其实非常复杂，要直连12家交易所，之后把数据拿过来自己排序，形成一个自己的市场订单簿，然后经过模型策略环节开始做交易。美国还有OPRA这样的期权行情提供商，会直接把12家交易所的价格帮你排序发给你，但是一般做市商都不会用这个东西，因为速度比我们自己排序慢上几十微秒，我们只是作为备用核对价格的后备措施。有些交易所技术故障可能暂停交易或者某些期权传送了明显错误的买卖价格，OPRA可以作为一个效验。但是在国内的交易所是垄断的，上交所的50ETF期权是不能通过其他交易所交易的，只能在上交所交易，并且中间必须要通过一个柜台不允许直连，你作为交易机构做数据清理的时候也不用做价格排序，一共只有一个交易所，提供的价格就是NBBO (National Best Bid Offer)。但是我之前提到授权的数据提供商的概念，它把数据不通过柜台直接传给你，这样会快一点。如果做高频交易的话这个也有一定的优势。

## From complexity to simplicity 化繁为简



## From complexity to simplicity 化繁为简



我们再看一下期权交易量的全球分布，期权在美国、欧洲都是非常大的交易品种，在国内才刚刚起步，还没有发展到那个水平。但是如果按照成熟金融市场的发展历程来看，未来国内期权市场肯定也会是一个非常大的市场。美国期权市场现在一天两千万手的个股期权成交量。全球期权交易量大部分都集中在股票和ETF期权上，除去印度和韩国两个股指期权异常大成交量的outlier，这两个品种统称的个股期权占了70%的交易量，然后还有一些股指期权、商品期权，当然这些交易量都比较小。从全球的角度来看，大家经常忽略的一个个股期权市场就是巴西市场，巴西的股票期权交易量其实占了全球个股期权交易量的15%，经常有人问全球最大的股票期权交易所在哪里？很多人会回答CBOE，但实际上是巴西的交易所。美国的个股期权交易量分散在12个期权交易所，而巴西的期权交易量总量低于美国但集中在两个交易所上，单一期权交易所的交易量还是超过了CBOE交易量。股票期权对于股票交易者来说是一个非常好的对冲风险的方式。内幕交易是大家最关注的，即便在惩罚力度苛刻的美国我们做股票期权这么多年还是遇到了不少很明显的内幕交易。可能平时都没有什么交易量流动性很差的一个期权，突然某一天从平均交易量几百手一下子交易了1万手的期权。而且大部分都是看涨期权。一周之后公开市场信息这家公司被另外一家公司收购了。这就是做市商最大的风险，因为我们都是被动交易者。在美国出现内幕交易可以向监管机构申诉，但大部分没有太大的回应，有几个确实是把责任查的很明确，记忆中最严重的一次就是中海油有一次收购加拿大的一家主营石油的公司，ticker是NXY，周五临近收盘突然市场价成交了一大批价外看涨期权，做市商的对冲让股票价格又飙升不少，周末公布消息NXY要被大幅溢价收购，我们因为这个事件一下子赔了上百万。我们向交易所申诉之后，交易所和监管机构查到并冻结了对方的帐户，对方是从注册在香港的交易公司，控制人又和中海油有千丝万缕的关系，但冻结了之后过了1、2年钱也没有回到我们帐上，所以这个事情对公司损失还是比较大的。

之前说过，期权是可以全方位交易，你可以交易标的价格方向、交易股息，在美国一般都是一年发4次股息，股息的预期数量可以去交易，如果交易员判断这个公司今年利润非常好，要增加股息了，但是市场并没有反应出来，比如期权隐含的股息只有3毛钱，交易员认为会增加到5毛钱，都可以通过期权市场做一些交易。期权交易员通常还会做一些事件型的交易。美国的事件比较有规律，比如一年4次发季报，或者是医药类的公司，FDA批准或者不批准会导致这个药价是上涨50%或者是下跌50%。另外可以交易波动率，我们见到最高

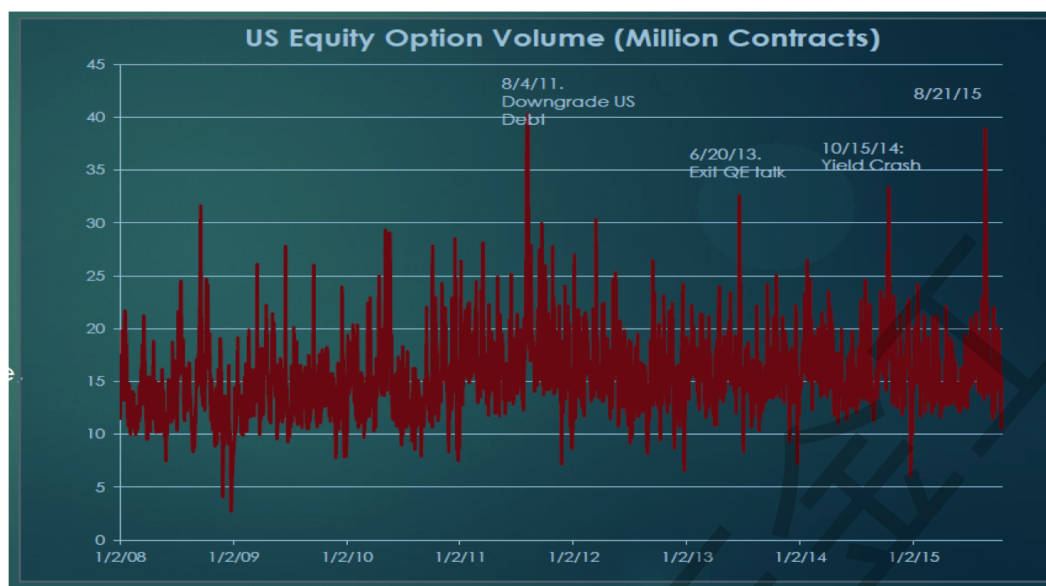
的股票波动率，也是在08年股灾的时候遇到的，300~400%，我们本来系统的预设最大值是500%，正常情况下一般不会遇到，但是那时候已经没有办法处理这些非常疯狂的波动率了。其实那时候这些期权并没有太多的做法，这个公司要么破产归零，要么不破产股价大幅反弹。就这两种结果，大家都只做极端的价外期权对赌二元的结果而不像正常做期权你会有各种各样的价格区间结果。

说完股票期权接下来讲一下股指期货，当时国内的交易所2012年去芝加哥的公司做宣传，当时他们推荐了两个产品，股指期货和原油期货。原油期货经过那么多年今年好像终于要推出来了，而股指期货因为股灾的原因无限期的延后了。当然如果能推出股指期货的话，这个品种有极大概率是全球交易最大的期权品种了。这对市场的风险对冲，或者对高频交易者来说都是非常好的交易工具。亚洲人非常喜欢赌博，韩国这么小的国家当年一度是全球最大的期权交易市场，它一个单独的交易品种KOSPI，一天的交易量超过香港全年的各种股票和指数期权交易量的总和。直到后来监管机构采取一些措施把交易量降下来了，降下来了他还是全球最大的几个期权交易市场之一。现在除去最大的期权市场美国，印度的Nifty股指期货，是全球最大的品种，但它的合约size非常小，很多印度人都把它当做交易彩票的方式了，花一点钱赌一赌，赌不中也没有什么损失。所以散户基本上还是买为主，卖期权的还是机构。巴西期权市场之前介绍过。欧洲主要是德国比较大一点。接下来还是韩国，韩国就靠它的KOSPI股指期货，台湾排在其后第六，我们大陆马上就要超过台湾了，现在我们50ETF大概已经一天的平均成交规模是50万手。

## 期权交易量分布

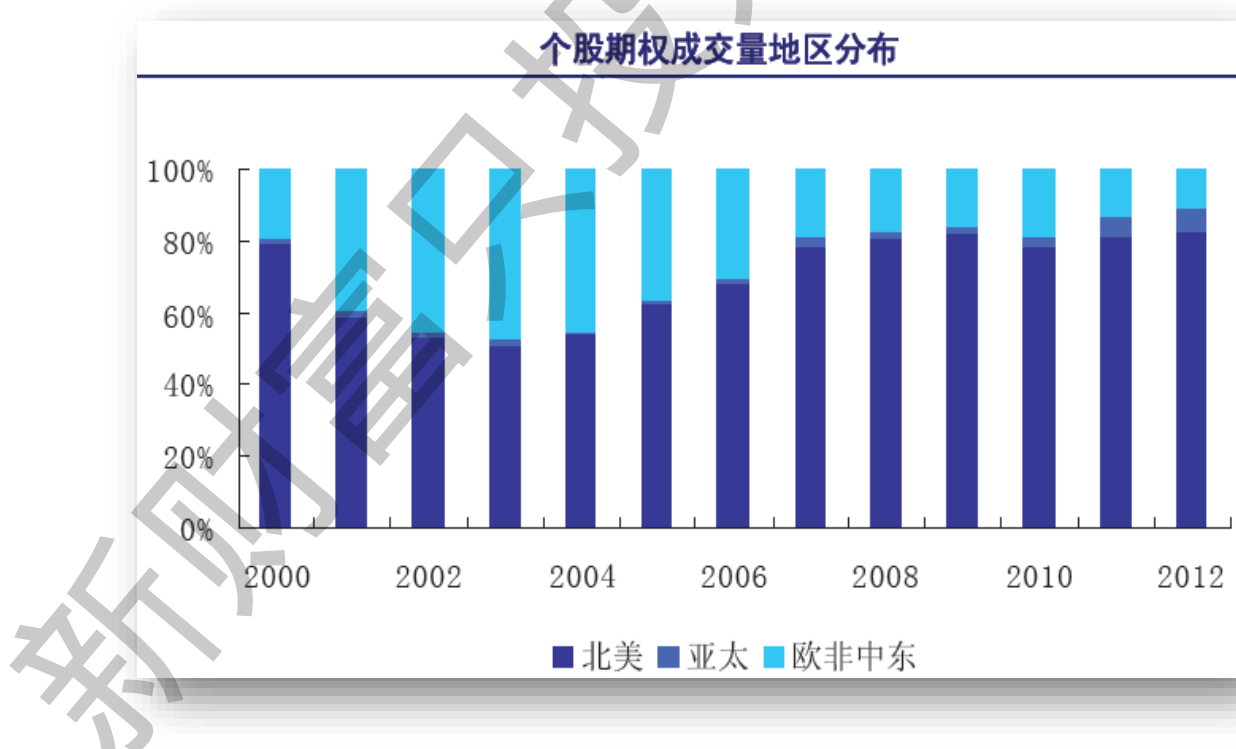
- 美国期权市场: 15-20 million 日交易量
- 股票/ETF期权占有所有期权交易量份额: 70%+
  - 股票/ETF option: 70%+ 交易量在美国
- 股票/ETF期权交易量:
  - 美国 70%+, 欧洲 7%, 亚太 7%(澳大利亚, 香港, 印度), 巴西 15%,
  - 最大的股票期权交易所: BM&F BOVESPA
- 个股期权的多样化风险: 内幕交易, dividend, borrow rate, earnings, corporate actions(merger, acquisition, rights issue, split, etc), special events(investor day, FDA, corporate announcements, exposition, product introduction, etc)
- Index/futures options: Nifty options, KOSPI options
  - 韩国2003年KOSPI日交易量2800万手, 超过当时香港全年期权(个股+指数)交易量(1455万全年)
  - KOSPI期权一度占据全球70%股指期货交易量
  - 机构个人交易量占比1: 1
- 最大的期权市场排名: 美国(1500万手日交易量), 印度(合约size小), 巴西, 德国, 韩国, 台湾(60万手, 指数期权only)

交易量什么时候会有大幅增长呢？基本上就是有事件发生的时候，交易量会大幅的跳跃，比如说美债在11年被降级的时候期权的交易量一下子跳到了4、5千万的交易量。国内50ETF期权创新高也就是1~2个礼拜前吧，也是因为股票大涨。和美国不太一样的是，我们一般波动率大涨是股票大涨的时候，美国一般是出现金融危机股票大跌的时候波动率大涨。和我们的数据规律是反过来的。



• Source: Bell Curve Capital/Derek Wang

这是分地区的个股期权交易量分布。最近几年印度和俄罗斯期权市场发展的非常快。。



这是香港前几年的期权市场规模，它的权证规模更大一些，权证在我们国内也试点交易过，那时候大家遇到最疯狂的是股票今天大涨了，认沽期权也涨停板了，散户们看到同一名字的证券就一窝蜂一起去买了，正常的理性市场不应该会出现这样的情况。后来权证就被叫停了，期权和权证相比在抑制市场疯狂这方面还是有诸多优势的。。



# 香港期权市场

- 1993年恒生指数期权
- 1995年首只股票期权
- 2000年首只ETF期权
- 截至2013年9月，香港交易所有5只指数期权，65只股票期权和5只ETF期权
- 个股期权(含ETF期权)2005年成交量900万，2006年1800万，2007年4600万

这是美国的发展历程，从floor trading到屏幕交易。当然我们国内直接跳到屏幕交易了。在美国老的交易员一定要长的高长的壮，喊得响亮，这样才能容易被看到听到，把你报价被broker记录进去。现在看的是软件和硬件的速度。虽然国内存在不公平公正的地方，但是美国曾经也走过不少的弯路，70、80年代惠普计算器可以输入公式计算期权价格。最早是被很多交易所禁止的入场使用的，于是较早运用计算器和量化技术的采取了一个修正措施，每天半夜把股票标的物价格和对应的不同波动率的价格算好写在纸上，第二天开市后对比一下，然后按照前一天计算的理论价报价。现在高频交易环境下其实也是这么用的，很多数值不是实时去计算的，也是事先计算好的，股票对应不同波动率的时候应该是什么价格，这样就可以直接从预先算好的表格中拿出来，直接套上去使用了。

## floor trading到screen trading

- 长得高和壮-> 软硬件和策略的速度
- 可编程计算器禁止入交易所floor，BS公式理论值单-> 全电子化
- 纯粹的MM交易风格-> 多种策略的并行
- 单一期权定价的精确-> 搜寻alpha
- BS期权定价本身的缺陷
  - geometric Brownian motion (jumps)
  - normal distribution of daily log returns (fat tails)
  - stable volatility (stochastic)
  - all strikes same vol (skew)
- 期货定价 vs 指数定价
  - 可对冲标的 (VIX options)

说到期权大家都会提到Black-Scholes，虽然存在不少的漏洞和缺陷，但是如果全市场都用Black-Scholes在定价的话，那做市商也只能接受和按照它来报价如果不想单向积累头寸的话。这些缺陷其中会给期权交易员提供很多机会，比如说它的一些假设，股价遵循几何布朗运动的，我们知道股价并不是连续的运动过程，可能会有跳空，高开。BS发布之后有很多修正的措施和改善的模型。但其实大部分的期权做市商还是在用Black-Scholes或稍加参数改动，因为比较简单方便。

这会给我们提供很多机会。举个例子，豆粕、白糖这些商品期权有一个询价机制，在50ETF期权刚刚出来的时候规章里也有询价的制度，美国倒比较少。作为客户可以发一个询价指令，做市商收到这个指令之后必须要报一个价格出来，并且买卖价符合交易所的价差要求，维持10~20秒。大家都用Black-Scholes来定价的话，有两个必要的变量，一个是标的物价格，一个是波动率，代入得到期权价格，根据这个理论价格在各加一定区间形成买价和卖价进行报价。期货期权和股票期权不太一样的地方就是每个月都对应的是不同标的物，不同月份到期的期货，股指期货稍微好一点，股指期货可以通过当前指数，利率，还有预估红利来估算出每个月对应的公允的理论价格是多少，商品期货就不一样了，每个月期货价格和相互间的基差变化很大。我们现在主力月份就是1、5、9三月份，除了主力月份外的非主力合约可能一整天都没有什么交易。有时候你会看到非主力合约50个点的价差，那么你取中间价来作为Black-Scholes参数中标的物价格还是选按照模型算出的其他价格，对算出来的期货价格都会有非常巨大的影响。另外我们说隐含波动率，这个期权市场上标的物这么宽，期权没有人报价，你很难倒推出波动率的，这也会对你的价格产生很大的影响。最早的时候如果你要影响这个期权定价的话，你可以在期货市场挂一个单，假设做市商都按期货的中间价来报价的话，比如说现在期货价差是100，你可以挂一个买单上去，比如说比当前卖价低50个点，你一下子就把中间价压低了25个点，你可以放心的是这个月份就算你把价差压到10个点5个点都没有人成交。这样的话你交易完之后可能过了一分钟，在期货市场上再挂一个买价上去，一下子又把中间价格往上拉上去了，对方如果还是非常愚蠢的做市商的话，他按照新的期货中间价又报一个新的价格出来，你可以把刚才那笔期权订单完美的平仓平掉。我们发觉做市商也在不断的更新报价技巧，当然上面只是一个假想的例子，实际操作中会有更多的可以操纵的漏洞，这也是我们当时和交易所提过的非主力合约，一天只有10手20手成交的月份，没有必要挂一个期权合约，这样将来不管是平仓、交割期权或转成期货都会造成很大的麻烦。交易所都是这样，积极听取大家的意见，但是不太愿意去采纳。我印象非常深刻的是股灾之后交易所说要推熔断，其实大家觉得没有必要。当然，美国也有熔断，但是从06年到现在这么多年只出现过一两次，因为他们没有涨跌停版，所以有一定的意义。在国内都有涨跌停板，所以意义并不大。反而容易把市场造成更大的恐慌。以前巴基斯坦搞了一个1%的跌停板，后来连续一个月天天都是跌停板。当然对于期权是有意义，如果每天只能跌1%，那么到期也就只能跌这么一定的量，如果是现金交割的期权的话那还是可以赚取只用black-scholes不做修改定价的交易员的钱的。

这是美国监管机构对高频交易的定义，需要满足下面几条，一个是用超高速的复杂计算机系统下单。所以期货炒单手的手工下单不属于范畴之内。第二个是co-location，这是为了减少网络延时。现在我们做市商绝大部分都是把机房放在交易所机房的。第三是持仓时间极短，怎么定义呢？有些是1秒前买入，1秒后卖出。期权不是这样，因为期权不像股票的流动性这么好。期权一个品种你要把流动性散布在衍生出来的几百个品种的流动性上面，只能尽量做到一段时间内交易风险尽量中性。第四是下单与撤单频率极高。程序化交易的话，现在交易所不愿意听到程序化交易，于是程序化交易往往说自己是半手工交易，想当然另一半是什么。普通程序化交易一般也不会把服务器放在交易所机房里面。然后我们另外说到，真的要严格按照这个定义来说的话，国内交易所行情传送方式500毫秒一次也是满足不了高频交易的。然后交易所还有一个报撤单限制。每天不能超过500次。后来大家也找到了应对措施。比如说收集很多身份证开户到达499次就换一个帐户等等。做市商没有这样的限制，是有天然优势的，期权这么多合约也不可能按500次的标准，我们报撤单一天都在10万次以上。

# CFTC/SEC高频交易定义

- HFT指什么
  - 硬件技术 vs 交易策略
- 程序化交易 vs 高频交易
- 定义:
  - 使用超高速的复杂计算机系统下单
  - co-location直连交易所
  - 持仓时间极短
  - 下单与撤单频率极高
  - 日间交易，风险中性过夜
- 国内行情传送方式和速度的限制
- 交易所硬件技术和撤单频率规则的限制(500次)

我们对比一下国外和国内对高频交易的监管。国外还是以促进发展为主，不管对市场发展还是产品流动性都有帮助。高频交易要占用很多资源，提供流动性的时候要占交易所大部分的带宽资源，所以交易所会给高频交易者一个限额，超过的话会有一个罚款作为限制，当然这个限额随着交易所技术环境进步也在逐年提高，并不是为了限制而限制，相反是为了在有限的资源内尽量公平满足各方需求。国内就不太一样了，国内股灾后把股指期货限制到10手和对程序化交易的严苛监管，还是以抑制为主。但程序交易化交易或者高频交易还是偷偷摸摸在做。期权有很多券商是不让你接入程序化的，有几家可以开，比如说交易量最大的华宝证券，很多期权程序化交易最初都在它那边开户，当初华宝占了全市场交易量的20%。但是客户量很少。但现在竞争更激烈，期货公司把交易费用降得非常低。期货公司期权程序化方面的接入措施也更激进。如果要做期权的程序化交易，可以考虑去期货公司开户。交易所经常会把公平放在监管程序化交易和高频交易的政策上，用程序化和其它交易给予的同样监管措施作为公平的比较杠杆，但我们经常说的公平和公正两个概念，交易所会说对做程序化交易和高频交易的交易者在监管上和全市场其余参与者对比都是公平的，比如说500次撤单，大家都同样的标准，或者交易所现在给所有做市商的流控都是一秒不能超过多少笔，如果超过多少笔就把订单打回去。这个对市场参与者也是统一或者类似标准。在美国如果交易者有这个需求可以申请更高的流量，但是国内是一刀切的规则，交易所认为是公平的，从交易者的角度觉得未必，有点限制技术发展了，比如说参加奥运会比赛，每个选手脚上都加了20公斤的铁链，对每个人的限制都是一样的，虽然比较公平，但是未必公正。就像读小学的时候，我被人打了，老师把我们叫到前面去每个人罚站一小时，大家惩罚一样时间一样但不分辨原因和限制公平，未必公正。

## Fairness vs Equality



### • 对高频交易的监管

- Eurex: HFT交易者备案登记
- order-to-trade ratio (OTR): MM更优惠的阈值
- excessive bandwidth usage: 每月发出超过限额的message数量(极大)罚款
- flagging of HFT orders(algorithmically generated orders)
- 以上措施为保证现有系统技术环境下能维护市场正常运行，交易所会在确保系统稳定和市场有序的前提下提高或取消上述阈值以发展HFT

### • 平等未必公平

- 抑制 vs 发展: 委托数量收费和OTR阶梯性罚款的区别
- 苛刻限制后的同一化，劣币驱逐良币
  - 每日最大委托笔数过小
  - 每秒最大下单数过低
  - 缺乏对MM需要的技术保护，交易豁免和更宽的通道以满足义务需要

再看看美国的高频做市商，Virtu Financial。在《Flash Boys》这本书出版以后本来这家公司要上市也因为书中的批评和引起的争议延迟上市了，一般高频交易公司不会公布它的财务报表，因为Virtu Financial要上市才公布。可以看到2014年的时候平均每人创造利润128万美元，然后它参与全球市场的交易，没有参与中国市场。从他的报表来看，在美国股市每天赚44万美元。此收益占整个公司的30%。15年国内股灾时在国内也有一家类似的进入中国市场的高频交易公司Citadel，在股灾后被媒体大量报道，当然后来他被作为恶意做空的典型，其实人家没有恶意做空。我买了股票肯定有对应的卖空的股票和卖空期货作为对冲，整个组合其实是非常中性的。在国内也没有办法做成美国那样高频交易。不过他后来拖累了我们两家大的券商，它是股灾时这两家券商的最大的融券客户。我们看一下高频交易的次数，其实未必每笔都赚钱，一半赚一半不赚。你可能会自己止损，去掉手续费之后就盈亏平衡或者赔一点钱，另外一些交易是赚钱的。如果一天要赚那么多钱，你可以看到高频交易公司每天的交易量是多少，一天16亿股的这么一个交易量。因为它参与全球市场，然后可以看到在全球交易的次数。



高频做市商：Virtu Financial。

1) 2009年由Vincent Viola在纽约创建。Viola在2001-2004年担任纽约商品交易所主席，也是Florida州冰棍球(Florida Panthers)老板。

2) 是全世界34个国家225市场中股票，固定收入，外汇和商品交易的做市商。

3) 2014年3月Virtu计划IPO，公告在其过去四年1278天交易中仅有一天亏损。当时公司估值为30亿美元，巧逢Lewis《Flash Boys》书的出版，引起公众注视。

4) 到2014年12月底，共有员工148人。2014年纯利润1亿9千万美元，平均每人创造利润128万美元。平均每天做530多万次交易。

✓从Virtu IPO的报表，Virtu在美国股市每天赚44万美元。此收益占整个公司收益的30%。Virtu CEO Cifu称公司51%的交易是盈利的。其余49%有些仅抵手续费甚至亏损

✓假设24%的交易是亏钱的，每股交易平均的盈利是 $0.01 \times 0.51 - 0.01 \times 0.24 = \$0.0027$

✓每天交易量： $440000 / 0.0027 = 160,000,000$ 股

✓按当时交易数据，平均一次交易200股，Virtu每天交易次数 $160,000,000 / 200 = 800,000$ 次

✓按美国股市交易占30%，Virtu每天总共交易次数为 $800,000 / 0.3 = 2,933,333$ 次。

为什么高频交易可以赚钱呢？从数学的概率上可以看到，交易次数足够多的话，一次是0.51，100次就是51%的赚钱概率了，一天交易10万笔的话，那就是稳定赚钱了，99.99%的概率了，按照这家公司80万笔的话，赢钱的概率是1。这符合我们数学概率上的估算。

- 设 $p$ 为每天交易的胜算率， $p=0.51$ 。设 $k$ 为 $n$ 次交易中所有盈利的交易次数， $0 < k < n$ ,  $k > n/2$
- 净盈利的概率是（见Greg Laughin文章）

$$P(k > n/2) = \sum_{k > n/2}^n \frac{n!}{k!(n-k)!} p^k (1-p)^{n-k}$$

$$\xrightarrow{n \rightarrow \infty} \int_{n/2}^{\infty} \Phi(x, np, np(1-p)) dx$$

$$= Q\left(\frac{n(0.5-p)}{\sqrt{np(1-p)}}\right) = Q\left(\frac{-0.01\sqrt{n}}{\sqrt{0.51 \times 0.49}}\right)$$

- $\Phi(x, \mu, \sigma^2)$  是平均值为  $\mu$ ，标准差为  $\sigma$  的正态分布密度函数， $Q(x)$  是  $Q$ -函数。这里当 $n$ 充分大时，二项式分布函数收敛于正态分布函数。
- 记  $P_d$  为一交易日是盈利的概率。我们估算Virtu的  $P_d$ 。从下表，

每天交易次数	$P_d$
1	51%
100	57.97%
1,000	73.5%
10,000	97.8%
100,000	99.99%
800,000	1

举一些高频交易的简单例子。首先NBBO（全国最优买卖价，National Best Bid Offer）的概念。其实国内没有这个概念，国内任何一个股票期货交易所的最优买卖价就是NBBO。在美国比如说三个交易所同时交易同一个品种，交易所1的价格是1.00/1.08，交易所2的价格是1.01/1.1，交易所3的价格是0.99/1.06。如果有个客户，他在交易所3上挂一个市价卖，那么他会成交在1.01，他始终会成交在市场最优价上。这是按照美国监管Reg NMS的要求。必须把订单转送Route到当时的最优价交易所。但是接受订单的交易所非常想把订单留到自己这边交易，最开始的时候他会给所有的会员客户提供非常短暂的时间，比如说20毫秒，现在告诉你了有这个订单了，你愿不愿意把价格更新上去match市场最优价？

比如单子下在交易所3，但它的bid 0.99其实并不是市场的最优价格，有没有人愿意更新到最优价格？20毫秒以后如果没有人愿意，那交易所3就把单子发到市场最优价格的交易所了，但是如果有人说我可以把我的价格提到1.01块，那么就把这个订单留在交易所，当时的目的是为了如此，看起来并没有太大的漏洞，大家皆大欢喜，但任何一个规则总会被发现漏洞来钻空子。举个例子，一个客户在交易所1买一个期权，按照交易所3的最优价格，应该是1.06，交易所为了留住这笔订单，允许做市商或其他高频交易者match当时的NBBO价格并使订单在本交易所成交而非route到其他提供NBBO的交易所。交易所1闪现客户订单20毫秒，如没有做市商或其他交易者愿意match，route此订单到NBBO的交易所。假设交易所3下一level的offer是1.09，高频交易者看到flash order后迅速将交易所3上的所有1.06的offer成交并挂出1.07的offer，此时客户订单和该交易员的1.07 offer(当时的NBBO)成交。这个相当于占了规则的一个漏洞的便宜。当然这个对于高频交易者还是要付出一定成本的，他必须直连到所有交易所并实时知道所有交易所的价格。

## 高频交易简例1

	Best Bid	Best Offer
Exchange 1	1.00	1.08
Exchange 2	1.01	1.1
Exchange 3	0.99	1.06

- NBBO: 1.01 / 1.06
- NBBO保护:客户不能在低于NBBO的价格上成交
  - Exchange 1上发出买一手期权，成交价格在1.06或者更优(有些交易所提供价格改善/拍卖机制)而不是当时Exchange 1上的最优卖价1.08
  - 交易所为了留下交易量，允许做市商或其他高频交易者match当时的NBBO价格并使订单在本交易所成交而非route到其他提供NBBO的交易所
- Flash Order: Exchange 1闪现客户订单20毫秒，如没有做市商或其他交易者愿意match，route此订单到NBBO的交易所
  - 假设exchange 3下一level的offer是1.09，高频交易者看到flash order后迅速将Exchange 3上的所有1.06的offer成交并挂出1.07的offer，此时客户订单和该交易员的1.07 offer(当时的NBBO)成交
- 直连所有交易所，硬件和速度的比赛

另外如果你收到的行情比别人快的话，在高频交易的市场上你也占了一定的优势。有这么一个订单，客户看到的是1.01/1.08，但实际上他看到的时候已经滞后了，别人比他早100微秒看到并且已经把1.01成交掉了，成交之后他又挂了一个1.02的卖价出来。实际现在最优价格是0.95/1.02，这个交易员看到的行情还是1.01/1.08，按照这个价格输入他的模型，他可能想挂一个1.08去买。结果就和比他快的交易员1.02成交掉了。这里可以顺便说下商品交易所一个漏洞叫做三价取中：最后成交价格、买价、卖价，三个取一个中间价就是最优价格。作为期权的话这个漏洞就比较明显了，不活跃的合约，可能做市商挂了一个100元在上面，来了一个客户想100块钱卖给做市商，客户可能手一抖写了10块钱卖，在美国的话，虽然10块钱卖，但是你成交还是100块钱。并不会因为你挂了10块钱的卖单就成交在不好的价格了，除非你错单错的比较严重不仅价格输错了还输了个10000手的巨大手数消耗完了所有买单的流动性可能直接卖到跌停板上，前不久中金所发生过几次股指期货被交易到涨停板和跌停板事故，也主要是由于下单时没有考虑期货手数被限制后的流动性和下了过大的量市价单。在国内，三价取中，买价是100，卖价是10，最后成交价可能是10天前成交价格或者是1个月前的成交价格，可能是20、30块钱，这就是交易所的规则。我们很早就发现了这个漏洞，但是交易所觉得现在的系统结构不容易修改一直保留在那。

Spooing在美国有过案例宣判，就是说我本来想买但挂一个卖单或者本来想卖但挂了

一个买单，我吸引别人过来，然后利用速度优势撤单再和别人成交。比如市场比较宽，1.00的买价和1.90的卖价，交易员想买，但挂出offer的单子1.45卖，有人加入继续改成1.44，1.43...并根据其他参与者加入量逐渐减少的情况判定何时中止并利用速度的优势撤单同时挂上买单成交在更优的价位上。因为美国交易所不同的机制，一个叫Pro-rata，就是按照比例来挂单和成交，和国内价格优先，时间优先不一样。在Pro-rata的机制下交易者并没有这个意愿去改善价格而是更改挂单的量去获得需要的成交量，在pro-rata的交易所spoofing的交易员需要成交大单时，通过逐步增大size观察跟单的size来停止自己的更新，达到平均成交价格的最优化，例如挂1.38卖后无人跟单，撤到1.39，挂10手时有10手的跟单，提高到20手时有总size 40，但提高到30手时总size 45，即可知此价位的最大提供量到达了，在次优的价格1.4, 1.41...也如法炮制，假设需要成交100手，可以通过此方法吸引出流动性提供者并最快地撤单并挂出方向的买单达到最优的平均成交价。当然，如果市场情况突然变化挂出的卖单未撤但有成交时需要能够最快地撤单防止被成交了相反的方向，因为我们需要买而非卖。

## 高频交易简例2

- “你看到的行情是我的成交”
  - 行情慢的交易员往往成为他人的猎物
  - 交易员A看到的行情滞后，想改善市场最优买价挂单1.02/1.08
  - 实际市场最优价并非A所见的1.01/1.08，1.01的价格100毫秒前已被成交，现在最优价是0.95/1.01，A刚挂上去的报价即被成交
  - OPRA比DMA慢，柜台比行情直连慢
- Spoofing钓鱼
  - 市场过宽1.00/1.90，交易员想买，但挂出offer的单子1.45卖，有人加入继续改善成1.44，1.43.....并根据其他参与者加入量逐渐减少的情况判定何时中止并利用速度的优势撤单同时挂上买单成交在更优的价位上
  - Pro-rata: 需要成交大单时，通过逐步增大size观察跟单的size，达到平均成交价格的最优化，例如挂1.38卖后无人跟单，撤到1.39，挂10手时有10手的跟单，提高到20手时有总size 40，但提高到30手时总size 45，即可知此价位的最大提供量，在次优的价格1.4, 1.41...也如法炮制，假设需要成交100手，可以通过此方法吸引出流动性提供者并最快地撤单并挂出方向的买单达到最优的平均成交价。当然，如果市场情况突然变化挂出的卖单未撤但有成交时需要能够最快地撤单，因为我们需要买而非卖。

bid	offer
	1.11
	1.08
1.01	
0.95	
0.93	

冰山订单国内没有，比如说客户想买1000手，他直接挂了1000手上去，市场上和其他交易者都看到了，所以每次呈现的是100手，成交完100手之后再跳出新的100手，不让你知道我到底要成交多少手。Trade for Rebate之前也有提及，交易者不是靠价差赚钱的，而是靠流动性赚钱的。靠流动性赚钱，那么策略不一样，我之前是1块钱买进来，马上通过1块01分卖，一买一卖虽然不赚钱，但是我可以从中赚取手续费，如果一天量非常大的话还是有可观的利润的。



## 高频交易简例3

size	bid	offer	size
		1.04	200
		1.03	200
		1.02	200
100(900)	1.01		
50	1.00		
20	0.99		

### • Iceberg order:

- 交易所只显示一部分的委托单
- 1.01的买价上只显示100手，剩余900手不显示，成交完100手后再补充显示100手，直到储蓄的900手消耗完毕。
- 探测冰山并调整报价,偏向买如果知道冰山存在

### • Trade for rebate:

- SPY, QQQ, Citi等penny option market, 高频交易员愿意同样价位买入卖出
  - 交易员A 报价1.01(10)/1.04(10), 市场成交1.01价格20手, 交易员A愿意马上更新卖价从1.04至1.01, 等待同样价位卖出。
- 不少期权交易所提供流动性不收取手续费且有rebate

size	bid	offer	size
		1.04	20
		1.03	20
		1.02	20
20	1.01		
20	1.00		
20	0.99		

下面我们介绍移动冰山，这个需要通过市场的历史数据计算概率判断出冰山订单是否存在，如果存在的话，最大概率是有多少量，根据目前挂单量，你要推测出他到底隐藏了多少单。假设你看到1块06分的订单成交掉400手之后又有400手挂出来了，通过你的判断，假设价位上有2000手订单在上面，这时候你可以知道有一个大单想在最优市场价上卖2000手，你就可以把市场上的买单全部成交掉，比如1.01你先卖给他，成交之后你挂1.05的成交上去，冰山订单会跟着你把价格调到市场最优价1.05上，然后你再故伎重演，不停的按这样的方式操作下去，一直操作到0.98，这时候你把单撤掉。这样一路上你可以发现你卖的平均价格是0.99，如果对方的冰山订单是如此地愚蠢一直跟着你设定的市场最优卖价走，最后你能在0.98全买回来，你就能赚到差价。但冰山订单不是这么愚蠢，不一定是每次都成交了100手就挂100手，可能他成交了100手，但过几秒再挂，或者随机地挂200手或者50手等不规则数字使得对手的判断更加困难。

## 高频交易简例4

### • 移动冰山

- 需要观察市场的行情成交变动做猜测判断和实施策略
- 市场报价1.01/1.06, 当1.06的400手被成交后市场又出现同样卖单1.06(400), 继续成交400手后同样状况依然发生, 猜测卖方上是iceberg order, 假设(预估) iceberg order有2000的size
- 1.01上成交所有买单, 并挂上1.05(300)的卖单改善市场最优卖价。Iceberg可能加入我们, 使得1.05的size变成700
- 继续在1.00上成交所有买单, 并挂上1.03(300)的卖单改善市场最优卖价。Iceberg继续加入我们, 使得1.03的size变成700
- 类似在0.99和0.98的买价上继续操作并改善卖价, 最后撤去自己0.98的卖价并和iceberg order在0.98价位上买入累积卖出的1200手均价0.991的合约

size	bid	offer	size	⇒	size	bid	offer	size	⇒	size	bid	offer	size
300	1.01	1.06	400		200	1.00	1.05	300+400		200	0.99	1.03	300+400
size	bid	offer	size	⇐	size	bid	offer	size	⇐	size	bid	offer	size
200	0.96	0.98	300+400		200	0.97	0.99	300+400		300	0.98	1.01	300+400

## 高频交易简例5

size	bid	offer	size
300	0.98	1.07	400

- 探测iceberg(PT/AT)的最低意愿价

- IOC order: 0.99(100)买, 未触发卖单, 1.00, 1.01, 1.02买均未触发成交, 直到1.03(100)买触发了成交, 大体知道对方愿意承受的最低价位在1.03
- 利用此信息做市商愿意挂单在1.06, 1.05, 1.04等价位卖并有大概率可以卖出后继续反手低价买入并赢利

- 变聪明的iceberg:

- 随机分割时间等待后再显示储藏的订单
- 探索牵引订单的动向并不加入对方以防市场朝逆向移动
- 放上与意图相反的大size订单在市场上引诱做市商, 实际意图需要卖, 但放上1.01(3000)的买单, 如果做市商如前例故伎重演则慢慢落入了对方的陷阱

size	bid	offer	size	⇒	size	bid	offer	size
3000	1.01	1.06	200		200	1.02	1.07	200

然后我们说一下高频交易和做市商的区别, 虽然硬件环境和交易策略都是一样的, 但是高频交易员可以有选择性的只做某几个期权的报价或者套利。因为大家都有流控, 做市商要给几百个期权报价, 万一遇到市场不利的情况, 同时要撤几百手单, 考虑一下流控只有100笔, 真的要全部撤掉的话那就需要好几秒了。这时候就有非常大的风险把自己裸露在外面, 在美国发生过类似的情况flash crash的时候, 几秒时间市场一下子下挫了好几个百分点。高频交易者没有这个风险, 因为没有报价的义务同时挂几百几千个订单。打个比方, 同样的带宽或者流控, 做市商相当于把十车道都排满了, 所有的车都在上面排着队开, 高频交易者可能也有十车道的环境, 但是只有两辆车在上面开, 所以想掉头转弯什么都非常方便。价外期权相对风险小, 放着的危险并不大, 实值期权先撤虚值后撤, 然后市场下跌的时候, 做市商先把看涨期权的买单和看跌期权的卖单优先撤掉。这些是做市商要考虑的在流控限制下的优化。但越来越多的优化条件和买卖单如何撤的复杂性一旦提高, 出错的概率就越来越大。可以确保大家将来有非常多的机会由于流控的因素, 会出现某些风险事件, 大家可以薅做市商的羊毛。国内高频交易还在慢慢发展, 另外我们的交易所也没有竞争, 这样会带来一些坏处, 交易所不会有那么大的动力或者激励机制去搞各种创新或者和别的交易所做竞争。

# 高频交易 vs 做市商

- 高频交易员(HFT, market taker) vs 做市商(market maker)
  - 有选择的建立头寸
  - 同样网络带宽条件下的速度优势(较少的messages和更新)
    - 逆势环境下无需更新或者撤消大量订单
  - 订单排序的优势
    - 时间优先(time priority)
    - 比例分配(size pro-rata)
  - 系统复杂性降低, 出错概率减少
  - 美国市场: fragmented market, loosely regulated, mature HFT market
  - 欧洲市场: several separated markets, developing HFT market
  - 中国市场: central, tightly regulated market, no real HFT

这是高频交易的发展情况, 美国其实没有发展太多年, 05年之后才是高速发展的阶段, 10来年基本上绝大部分交易都是自动化交易, 股票上是这样, 期权就比别的再高一点了。而co-location的流行, 最早服务器放在交易所的时候, 交易所机房占地几个足球场, 有些人的服务器离撮合主机非常遥远, 有些是就在撮合主机隔壁。当然会有客户提意见, 他的速度比我快! 最后交易所做出的改善措施, 不管你是在撮合主机的隔壁还是在机房的角落, 大家都保证在天花板上把网线绕几圈再接入确保所有人的连接网线一样长度。这样导致所有人的速度都是一样的。Flash crash之后美国也研究过是否要限制高频交易, 87年美国股灾的时候也研究过是否要限制股指期货, 但并没有把股指期货禁掉, 美国的高频交易也在继续发展。美国当时研究得出结论是高频交易不是导致股市崩盘的原因, 反而在flash crash中提供了更多的流动性并最晚撤出市场, 帮助稳定市场和恢复市场。

## HFT现状

- 2005后高速发展, 2014年时
  - 美国70%+的trades自动化交易
  - 欧洲22%的value, 29%的trade, 57%的order(大部分小订单, 撤单频率高)
  - Eurex期权: 36%, 44%, **75%** (期权产品特性和做市商因素)
- 交易所数据传送和处理: 纳秒级别 (国内微秒和毫秒级别)
  - co-location的流行: 订单到交易所 100微秒
  - 对比没有co-location: 5毫秒, 德国到美国: 50毫秒
- HFT导致股市崩盘/波动 or 促进价格发现
  - 2010 flash crash
    - HFT在最初阶段依然提供流动性
    - 期权做市商最早离开市场, 导致期权市场完全没有流动性

有人指责高频交易后最优价格上报单量变少, 这是对高频交易的误解, 高频交易肯定把挂单量变小的减小成交后的风险, 但极大缩小了价差, 在现在价差非常小的情况下, 两边的size只有5手、10手。但把不同价格的量加在一起肯定会比原来的价差较大的买卖价

的挂单量更大一些。从流动性角度来说的话，高频交易就可以增加了流动性。然后还可以看到一些场外OTC市场没有高频交易，因为没有硬件环境，不存在这些巨大的流动性和交易量，这些市场的波动率比场内市场更高。高频交易和波动率之间并没有明确的联通关系也并不导致市场变得更加动荡。Flash Crash最后调查下来也没有发现两者的联系。隔了很多年美国找了伦敦的一名交易员，把当时的罪名都压在他身上了。这个有点替罪羊的感觉了。

## 对HFT的误解

- 降低了spread，但也减少了size，增加了交易成本

size	bid	offer	size
		1.1	55
		1.09	35
		1.08	25
		1.06	15
15	1.04		
25	1.03		
35	1.02		
55	1.00		

size	bid	offer	size
		1.1	100
100	1.00		

- HFT的特性决定了在多个layer挂小size的单，但更优价格的总量高于原先
- 研究表明过去的15年美国市场交易成本降低了60%
- 中国用实际市场做了一个遏制程序化交易的实验：流动性枯竭，波动率激增
- 增加了市场波动率，下跌时导致市场崩溃
  - flash crash的事后报告并没有证据是HFT导致股市的崩盘，相反HFT在下跌过程中始终在提供流动性
  - CME的研究表明HFT的发展伴随而来的是所在市场波动率的逐步下降
  - OTC市场(CDS, IR swap,etc)没有HFT，更高的波动率

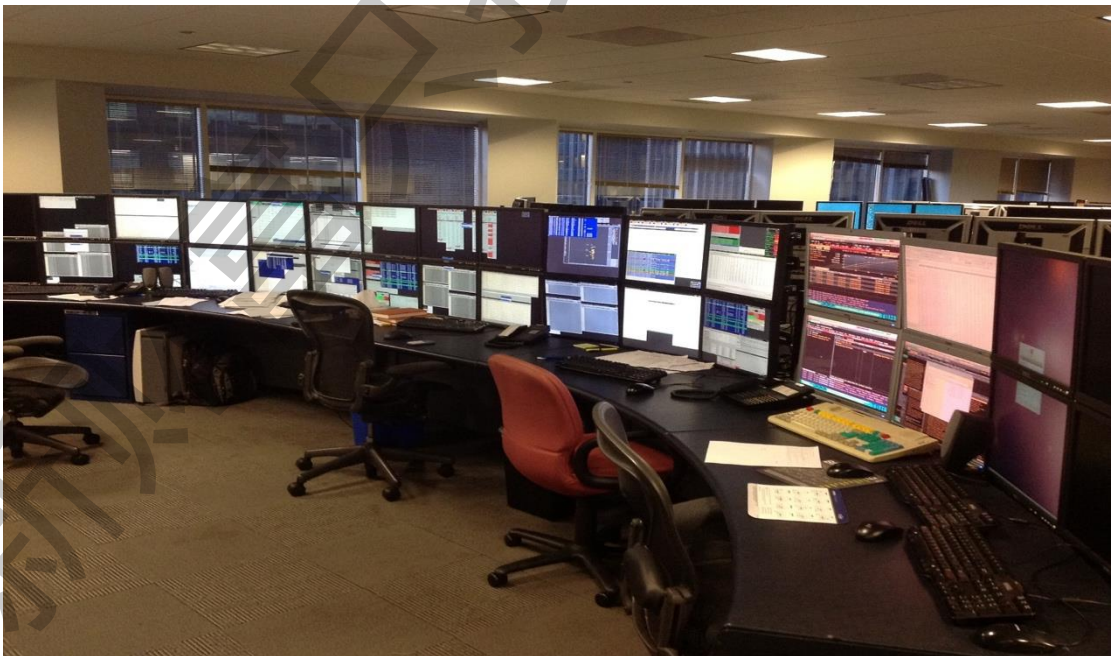
然后我们再介绍一下场内期权做市商，场内期权在国内的推出是一个漫长的过程，让我们等待了很久。15年原本传闻股指期货上市，后来发生股灾推迟了，郑商所经常说自己是国内最早研究期权的交易所，90年代就开始研究了，到现在已经20多年了，然后15年最早研究期权的一位郑商所的老师退休了，研究了20多年就是没有看到商品期权在郑商所上市，令人惋惜。美国的做市商团队大致上有这么一些人员，交易员、风控、量化分析、程序开发、系统管理员、数据管理员、后台结算人员，其实和任何一个私募或者公募基金做交易的团队也类似。然后在美国交易主体，国内是券商投行做为做市商主体。国外的话投行做市商那时候也只有3, 4家比较大一些，像高盛、摩根斯坦利、巴克莱、瑞银等等。有一家期权做市商需要特别提及，IB公司，它最早是一家期权做市商，它有很多的有趣轶事包括早年的使用机械手臂下单回避交易所的程序下单限制，现在美国做期权的公司上溯本源，大致就是IB或者O'Connor两家最早的期权做市商公司。IB后来演变成美国最大的网上经纪业务公司，今年IB自己也步之前UBS的后尘宣布逐步退出美国期权做市商市场，发现美国期权做市商的稳定盈利之路在高度竞争和市场结构变化的大环境下也走到尽头了。现在只有几家最大做市商的瓜分了市场份额，国内未来发展也应该会是这样，几家最大的期权做市商占领了绝大部分市场。



# 期权做市商

- 做市商团队结构:
  - 交易经理, 交易员(trader, mm, liquidity provider)
  - 风险控制(risk manager)
  - 量化分析(quant)
  - 程序开发(developer)
  - 系统网络管理员(system/network admin)
  - 数据库管理员(database admin)
  - 后台结算和支持人员(back office)
- 做市商主体:
  - 交易公司(Trading Firm)
  - 投行的自营交易部门(GS, MS, BCS, DB)
  - 经纪商的做市商部门(IB/Timber Hill)
  - 对冲基金(Citadel)
- 是否需要做市商?
  - 全球前十大的股指期货产品, 除了韩国Kospi, 都是采用做市商系统
- 国内大的产品(股指期货)可以不需要做市商
  - 但有做市商则能让市场更有效和波动率更平稳
- 相对较小的产品个股股票期权
  - 没有做市商会遇到流动性的问题

这是典型的期权做市商的工作环境, 4个人面对60个屏幕, 团队要完成对1000多支股票的做市和监控。国内的品种少, 但国内商品期权有夜盘, 其实夜盘没有太大的必要, 美国股票也有盘后交易, 但是没有夜盘期权, 期权4点收市就结束了, 第二天再上班。国内的夜盘制度让交易员非常的疲惫。早上、晚上都要上班, 而且晚上目前也没有交易量, 收益弥补电费空调费尚有困难。



接下来说高估/低估的概率与期权定价, 其实政策或者监管对期权定价的影响还是很大的, 如果你只是单纯用BS公式代入参数而不把这些影响价格的条件概率事件考虑在内的

话，算出来的期权价格和实际的期权价格会出现一个比较大的偏离。比如说在股灾的时候，交易所会打电话给期权交易机构，日末持仓所有期权累加在一起不能是卖看涨期权或者买看空期权这么一个结构。很多人会把看跌期权卖掉或者把看涨期权买回来使得自己总量保持在交易所的规定范围之内。这时交易所的“窗口指导”会导致本来较为平坦skew不明显的曲线两边都会高高翘起来。

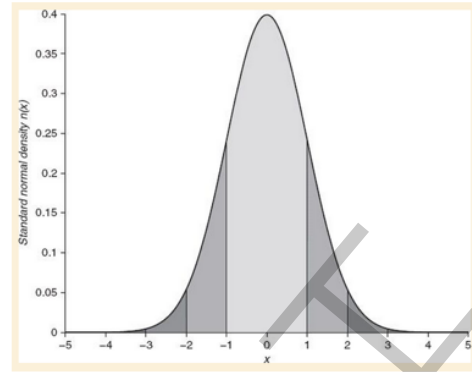
## 高估/低估的概率与期权定价

- 随机事件的概率
  - 同生日概率：
  - 需要选取多少人能找到两人生日相差k天的概率大于50%？：
  - 每天生日概率分布如果不均等：
    - 1-6月25%，7-12月75%
- 条件概率与无条件概率
  - 两羊一车
  - 熟人偶遇
  - 涨跌停板和熔断
  - 窗口指导（内幕交易?）
  - 政策指导
- 个股期权与内幕交易

人数	两人同生日	2人生日相差k天	需要的人数
10	12%	0	23
20	41%	1	14
23	50.70%	2	11
30	70%	3	9
50	97%	4	8
100	99.9996%	5	7
		6	6
		7	?

如果大家都相信国家队的存在会一直干预市场，预期交易的结果是50ETF的隐含波动率也不会太大，然后因为有国家队的实际干预市场。实现波动率也不会太高，过去一年卖50ETF期权的人回报率是非常高的，当然往后看很难预测，你也不知道将来国家队会怎么做。你只可以看到他过去怎么做。未来如果有明确的内部消息，知道确定他们会这么干预下去，那么卖期权还是非常好的策略。用BS公式的时候一般是无条件概率，但市场上每个新闻都增加了一个信息或者增加了一个条件在期权里面，举个简单的例子，比如说城市里面85%的出租车是绿色的，15%的出租车是蓝色的。然后出了一个事故，目击者认为逃跑的车子是蓝色的，目击者作测试发现80%时间里识别车辆颜色正确。通过条件概率算出来，其实只有41%的概率是蓝色的。更大的概率应该是绿色的。通过期权我们看到的话，用正态分布的假设放进去的话，日回报率落在一个标准差内的频率，理论上应该是三天一次。但是实际上我们把标普500的数据拿出来，1个标准差出现的概率不是理论上的68.2%。三天一次，而是81.79%。两个标准差是95.45%。一个月一次。三个标准差是一年一次，99.73%，4个标准差是99.99%。一世纪一次，但实际概率出现在四个标准差之内的概率要小，也就意味着肥尾事件发生的频率比我们预期要高，黑天鹅期权的定价按照BS公式来定会被明显低估。

- 城市中85%的出租车是绿色的，15%的出租车是蓝色的
- 目击者识别车祸后逃逸的车辆为蓝色
- 目击者之后作测试发现80%时间里识别车辆颜色正确
- 59%概率逃逸车辆时绿色
- 某种疾病的感染率是1/1000
- 检测该疾病的测试有5%的错误率(false-positive rate)，95%正确率
- 某人通过该测试被检测出有该疾病
- 实际确实有该疾病的概率2%



标准差	理论概率	理论频率	实际概率(S&P500,2008-2013)
1 std	68.269%	三天一次	81.79%
2 std	95.450%	一月一次	94.64%
3 std	99.730%	一年一次	98.08%
4 std	99.994%	一世纪一次	99.21%

Edge的概念不管对做市商还是别的套利交易者都是类似的概念。简单来说Edge就是我买和卖期权的成交价和我自己理论价的一个差价，如果我比理论价低的价格买进来，那么我的Edge就是正的，如果比理论价高的价格买进来，那么就是负的。类似赌场的例子，轮盘赌的游戏，38个数字，1一直到36、0和00。赌注是1块钱，押对了回报是36元，错了分文不取。如果没有0和00，这其实是公平的游戏。简单计算下赌场的edge是5分钱，赌场是按照大数定律来获胜的，赌场只能作为玩家的对手方，赌场不能像做市商一样，又能买游戏又能卖游戏。所以赌场可能连输几次并赔钱，赌场所以会设置一个每次最大下注数量。这和做市商设置最大vega/gamma/delta数值一样。但是游客多了，来1万个或者10万个，赌场输钱的概率很低。盈利模式也是这样通过大量交易（库存）来盈利，希望0库存获取差价，我们既想1块钱卖给对方这个游戏，也想0.9毛参与这个游戏。通过大量买卖赚取价差。和赌场不一样的是赌场每次都对应新一轮游戏不同的结果，做市商买一个期权，卖一个期权，对应最后到期的都是是同一个结果来计算盈亏。

## Edge:

- 赌场: 轮盘赌roulette, 38个数字1, 2, 3, . . . . . , 35, 36, 0, 00, 赌注\$1,押对了回报\$36,错则分文不取。
- 期望值  $36/38=95$  cents
- 赌场edge:  $\$1-\$0.95 = 5$  cents
- 赌注如果降为 90 cents, 玩家有edge  $0.95-0.9 = 5$  cents
- 如果赌场也可以做买方 -> market maker
- 0.9 bid/1.0 offer
- 盈利模式: 1.大库存依靠大数定律长期盈利, 2.零库存获取差价。赌场游戏赔率固定, 期权市场赔率时刻动态变化->不希望库存
- 比理论价格低买, 比理论价格高卖, edge = theo-bid or ask-theo

## Edge continued:

- 假设利率 10%: 赌对了36元要一年后发放,  $95 \text{ cents}/(1+10\%)=86 \text{ cents}$
- 理论值变低后之前90 cents买入游戏的参与权变成了负的edge, 调整bid/ask
- Rho: 理论价值的变化量/利率没变化1%
- Market making过程: 理论价格, 报价, 库存管理, 调整报价。没有理论价格的时候根据交易情况调整报价, e.g. 0.86 theo, 0.9 bid/1.0 offer, 单边市场
- 信息不对称的套利: 市场大部分时间是有效的
  - e.g. 加拿大公司股票distribution, 期权定价偏离理论价值, 最后原因: 对美国机构延后发放并有额外税率。

当然期权还有很多其他的风险, 比如标的物价格、波动率、利率都会实时变化, 也可能增发减少股息, 国家队干预等等这些新的外部条件都可能加进来影响定价和交易, 如果赌场告诉你赌对了数字后的36块钱不是现在当场给你, 要过一年才能取, 你要把利率放进去, 重新算一个理论价, 赌场也要根据新的价格来确定参与的赔率, 所以任何一个新的信息都会影响到你这个期权应该有的定价, 在期权做市的过程当中, 你要根据新的市场情况算出新的理论价, 根据你的库存算出新的理论价, 增减edge确定报价。作为其它类型期权交易者也是一样的, 你要了解自己的库存, 了解自己的交易方向, 根据市场的新的情况调整你的期权头寸。不是固定的一成不变的过程, 而是一个动态变化的过程。期权的变量比股票或者期货多很多交易的复杂度也相应增加。

下面是做市商双边报价的流程, 可以看到做市商并不是操纵这个市场, 而是接受这个市场, 有了市场价格观察之后, 我们就有了变量, 输入模型里面, 拟合获得一条波动曲线, 综合考虑其它变量获得理论价格之后就可以根据库存和风险阈值参数进行市场报价, 但报价过程并不是依据单一期权或互相独立给单一期权报价, 而是所有的期权一起考虑, 整体的风险一起考虑最后得到整个系列的报价。我们在美国的做市商都把从事的行业称之为风险转移的行业。像股指期货, 交易量足够大没有做市商也完全可以交易起来, 因为所有的交易都在一个合约上, 就没有太大的问题。但是期权的话, 就会造成很大的困难, 同一个期货合约会衍生出几十个期权品种, 你要保持每个期权合约非常好的流动性还是比较难的, 所以这是后需要做市商提供报价, 使得没有流动性的时候还是有一个价格存在时刻可以交易。然后定价模型的变量某些可以直接观察到, 比如说股票价格、行权价格、到期时间, 还有一些有模糊区间的变量, 比如利率、股息等等, 还有间接推导的变量, 隐含波动率, 有了隐含波动率可以通过其他变量计算出价格。所以做市商的定价和交易有点递归循环的过程, 从市场价格倒推出隐含波动率再用隐含波动率计算理论价格回到市场交易。听上去有些类似做鸡汤的食谱告诉厨师需要准备的材料是鸡汤。



# 做市商双边报价

- 市场价格->变量输入->波动率曲线->定价引擎->理论价格->报价/库存管理->风险管理
  - 依据整体头寸而非孤立性地报价
- 供需曲线, 市场决定均衡价格, 做市商不决定或者操纵价格
  - 价格发现
- 流动性提供商, 风险转移(risk transferring)
- 新产品的报价
  - 和交易所共同设计和推广产品
- 定价模型的变量:
  - 直接观测:
    - 股票价格
    - 行权价格
    - 到期时间
  - 模糊区间:
    - 利率
    - 股息
  - 间接推导:
    - 隐含波动率

然后做市商赚取价差，国内还没有达到通过提供流动性来赚钱，通过价差赚钱的前提是交易量足够大。大家通过50ETF期权去做对冲的话，其实是非常好的品种，价差非常的小，但50ETF受到的国家队干扰过大，对冲往往并没有太大意义。期权交易者喜欢波动率高的环境，这样的话交易量比较大，流动性也比较大。所以经常有人把做市商比作赌场，客人越多越赚钱，但其实做市商本身还是靠天吃饭的客人，交易所是更好的类比对象，旱涝保收。

待会儿大家可以看到美国一家高频交易做市商的报表。我们国内是15年股灾的时候期权做市商是最赚钱的，那时候波动率都到了100的波动水平，这是非常夸张的水平了，在美国VIX代表的美国标普500历史最高的波动率是也就80几。现在VIX是到了接近历史最低水平，93年的时候是9点几，现在也在9点几，预计马上就要突破历史最低点了。在这10年中我们交易期权还是非常幸运的，既看到了历史最高点，也看到了历史最低点。

这是之前说的那家公司的案例，这家公司Interactive Brokers可能是高频交易的先驱吧，做高频交易这件事已经是在几十年前了。不过很遗憾最后还是选择退出期权做市了。但它非常早就在经纪业务上布局了，这是非常有前瞻的一步。07年时候还是做市商占了大头，08年的时候这块业务最大了，超过了10个亿了。之后就萎缩了，12~14年逐步缩减，现在已经不到1个亿的收入了，其实基本已经放弃做市商业了，主要收入来源变成了经纪业务，经纪业务同年稳定增长。

		Three Months Ended December 31, 2008		Twelve Months Ended December 31, 2008	
		2007		2007	
		(in millions)			
<b>Market Making</b>	Net revenues	\$308.8	\$279.0	\$1,343.5	\$1,031.2
	Non-interest expenses	<u>87.2</u>	<u>74.8</u>	<u>315.9</u>	<u>311.4</u>
	Income before income taxes	<u>\$221.6</u>	<u>\$204.2</u>	<u>\$1,027.6</u>	<u>\$719.8</u>
	Pre-tax profit margin	72%	73%	76%	70%
<b>Electronic Brokerage</b>	Net revenues	\$118.7	\$118.2	\$505.8	\$425.2
	Non-interest expenses	<u>75.7</u>	<u>56.1</u>	<u>281.8</u>	<u>227.3</u>
	Income before income taxes	<u>\$43.0</u>	<u>\$62.1</u>	<u>\$224.0</u>	<u>\$197.9</u>
	Pre-tax profit margin	36%	53%	44%	47%
<b>Corporate*</b>	Net revenues	\$1.8	\$0.3	\$0.8	\$11.8
	Non-interest expenses	<u>(2.6)</u>	<u>(2.1)</u>	<u>2.7</u>	<u>(2.1)</u>
	Income before income taxes	<u>\$4.4</u>	<u>\$2.4</u>	<u>(\$1.9)</u>	<u>\$13.9</u>
<b>Total</b>	Net revenues	\$429.3	\$397.5	\$1,850.1	\$1,468.2
	Non-interest expenses	<u>160.3</u>	<u>128.8</u>	<u>600.4</u>	<u>536.6</u>
	Income before income taxes and minority interest	<u>\$269.0</u>	<u>\$268.7</u>	<u>\$1,249.7</u>	<u>\$931.6</u>
	Pre-tax profit margin	63%	68%	68%	63%

\* Corporate includes corporate related activities as well as inter-segment eliminations.

	Year ended December 31,		
	2014	2013	2012
	(in millions)		
<b>Net revenues</b>			
Electronic brokerage	\$ 952.3	\$ 818.5	\$ 672.2
Market making	284.4	361.1	490.5
Corporate and eliminations	(193.4)	(103.4)	(32.2)
Total net revenues	<u>\$1,043.3</u>	<u>\$1,076.2</u>	<u>\$1,130.5</u>
<b>Income before income taxes</b>			
Electronic brokerage	\$ 588.5	\$ 395.8	\$ 343.5
Market making	114.1	158.5	219.5
Corporate and eliminations	(196.5)	(103.0)	(36.0)
Total income before income taxes	<u>\$ 506.1</u>	<u>\$ 451.3</u>	<u>\$ 527.0</u>

	December 31, 2014	December 31, 2013	December 31, 2012
	(in millions)		
<b>Segment Assets</b>			
Electronic brokerage	\$38,280.1	\$31,333.5	\$25,741.5
Market making	12,172.4	12,139.5	12,730.8
Corporate and eliminations	<u>(7,067.5)</u>	<u>(5,602.3)</u>	<u>(5,272.7)</u>
Total assets	<u>\$43,385.0</u>	<u>\$37,870.7</u>	<u>\$33,199.6</u>

million	MM	Brokerage
2007	719.8	197.9
2008	1027.6	224.0
2012	219.5	343.5
2013	158.5	395.8
2014	114.1	588.5
	MM	Brokerage
2007	78.4%	21.6%
2008	82.1%	17.9%
2012	39.0%	61.0%
2013	28.6%	71.4%
2014	16.2%	83.8%

大家做期权的话，所有做期权的人都有市场的风险。我们一般用希腊字母来管理市场风险，比如说实时的风险指标更新，Greeks风险，隔夜突发事件，Pin risk，偏度，保证金、结算价盈亏风险。交易订单被取消，内幕交易，单边市场。后面的一些细分的风险，你刚刚进入国内市场也不一定马上会发现。交易所定的结算价和当前市场交易结算价格会偏差的非常远，不像期货，期货今天偏了，明天就回归回来了，期货常规下取平均价格也不会偏太远，期权的话有些一天都没有成交，或者有些价外期权市场波动率和交易所拟合的曲线会偏差很远，价格会差的很远。这是我们当前在市场遇到的情况。单边市场做市商

是比较痛苦的，作为普通的客户交易员如果你自己有一个观点的话，就可以通过单向交易获利了，内幕交易的话，通过期权会有加倍的杠杆效应不过比较容易被监管机构看出来。中间会有很多手段或者手法。做各种组合，让别人看得不是很明显。这里就不介绍了。

## 期权做市商的市场风险

- 连续的风险管理过程：
  - 实时的风险指标更新
  - Greeks(希腊字母)
  - 隔夜gap risks(突发事件，take over)
    - 情景分析矩阵
  - Pin risk
  - 偏度(skew)
  - 保证金,结算价盈亏风险
  - 交易订单被取消(order cancellation, obvious errors)
    - Flash crash
  - 内幕交易
  - 单边市场(one way market/toxic order flow)

当然做市商和高频交易者最大的风险是系统故障和人员风险。骑士资本的系统故障和法兴银行流氓交易员是很好的两个案例。风比如不要给期权交易员太大的特权，让他自己控制波动曲线和后台的理论结算曲线，他可以把自己赔钱的头寸mark到一个赚钱的头寸，风控根本看不出来他是赔还是赚。如果公司风控并不了解期权的话，会有比较大的影响。然后灾备系统也很重要，备用的服务器和网络。在美国对知识产权保护比较看重。特别是高频交易公司。内部监控严格，不能随便插USB接口，工作的时候不能登录外部网站，只能登录公司自己邮箱不能使用gmail等邮件系统。

## 期权做市商的运作风险

- 系统故障
  - 数据传输中断，数据错误
  - 网络故障
  - 交易界面不响应，交易程序停止工作
- 新软件release的流程与管理
  - 骑士资本，Knight Capital，出售给Getco
- 延迟(latency)，fast market
- 人员风险
  - 仓位核对，独立风控和结算部门
  - 里森，巴林银行，14亿美元
  - 克维尔，法兴银行 Kerviel，SocGen 70亿美元
- 灾备系统
  - 对帐核实交易(Drop copies)
  - 备用服务器，网路
- 知识产权保护
  - 内部监控
  - 非竞争条款
  - 保密协议

做市商或者高频交易，最后面临的的就是军备竞赛了，每年都要把CPU、内存更换一遍，如果网络升级我们肯定必须同步的，如果一家高频交易公司一年盈利只有2、3千万，你要花1、2千万更新硬件，那你只能退出市场了，所以美国高频交易能够做大的也就是这么几家。一家独大占领所有的市场。现在还有几家公司在逐步开展中国业务。有几家在国内做的也挺大的，特别是在期货市场。有相当大的交易量。大家如果要做期权的话，准备工作是非常多的门槛是相当高的，你要有非常好的交易系统，有完善的监控的界面。做市商如何避免亏损。你一旦头寸做大了，在国内你很大的风险就是不能及时撤单。因为有流控的限制。但是在国外就有一键撤单这个功能。这个功能我们也提了无数遍了，但是最终还是没有实现。还有交易所提供保护机制。以及不断提升的技术等等。。

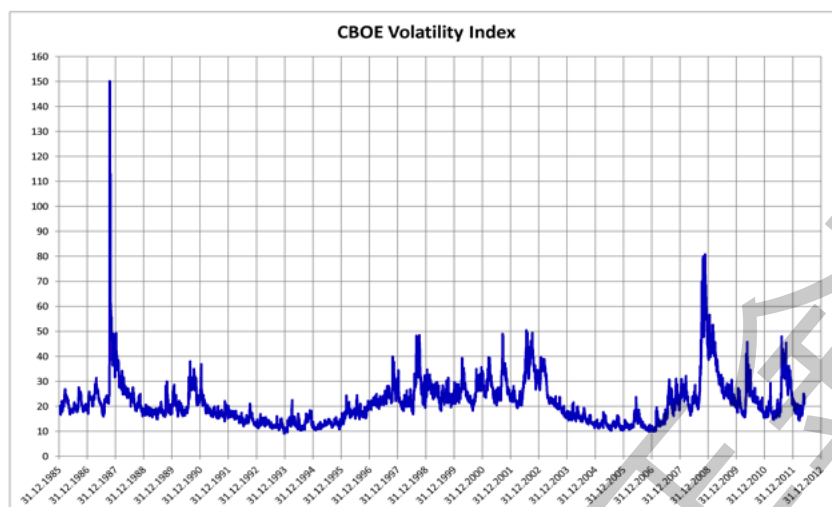
## 期权做市商如何避免亏损

- 操作风险管理
- 市场风险管理
- 交易所提供保护机制
- 不断提升技术

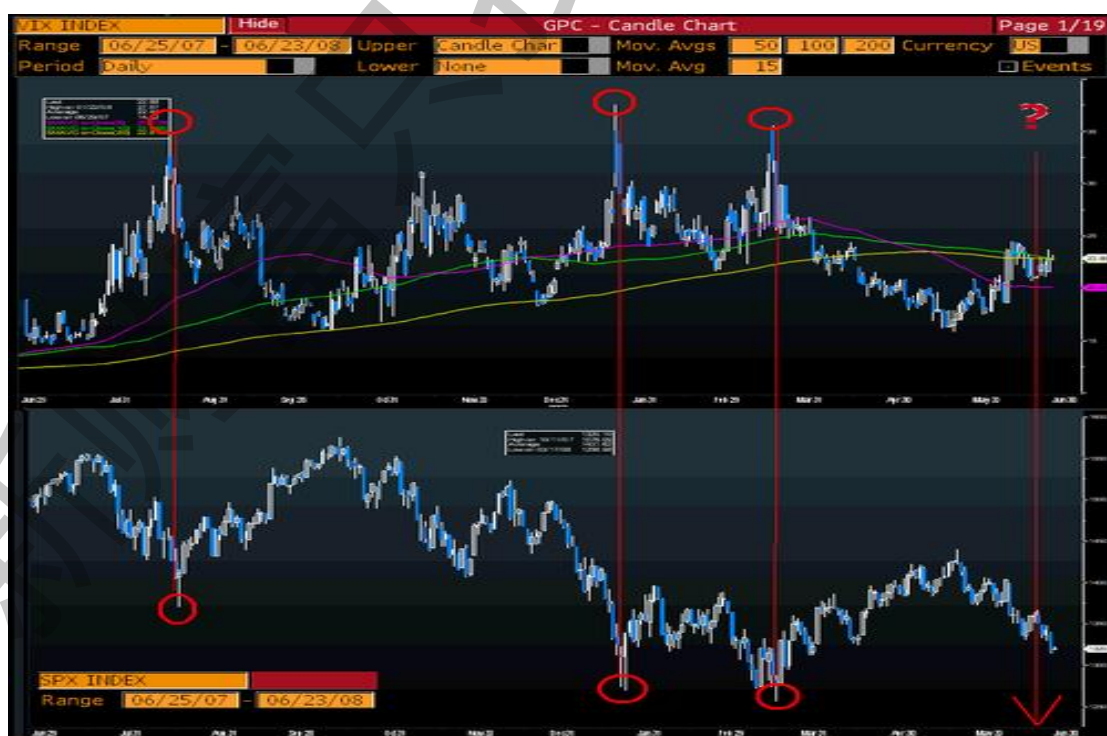
最后举两期权波动率的例子，我们要关注波动指数，隐含波动率，历史最高点是在08年，另外还有87年的股灾，那时还没有vix，是按照公式和数据倒推回去。但87年股灾之前做市商所有的期权都是按Black-Scholes来定价的，没有Skew。实际运用后有合理数据的VIX最高点出现在08年的80几，最低点是90年代的9，现在马上也要突破最低点了。中国的话有个类似的叫中国波指，用类似的方法算出的全市场的波动指数，用中国波指来衡量全市场波动指数还是有一定偏离的，因为美国的期权价格已经列到了国内标准几个涨跌停板之外了，中国话可以看到50ETF期权就上下4-4档，在股灾的时候常常跌破最后一个行权价格，这样就很难算出一个合理和公正的波动率指数出来。



## VIX – historical price



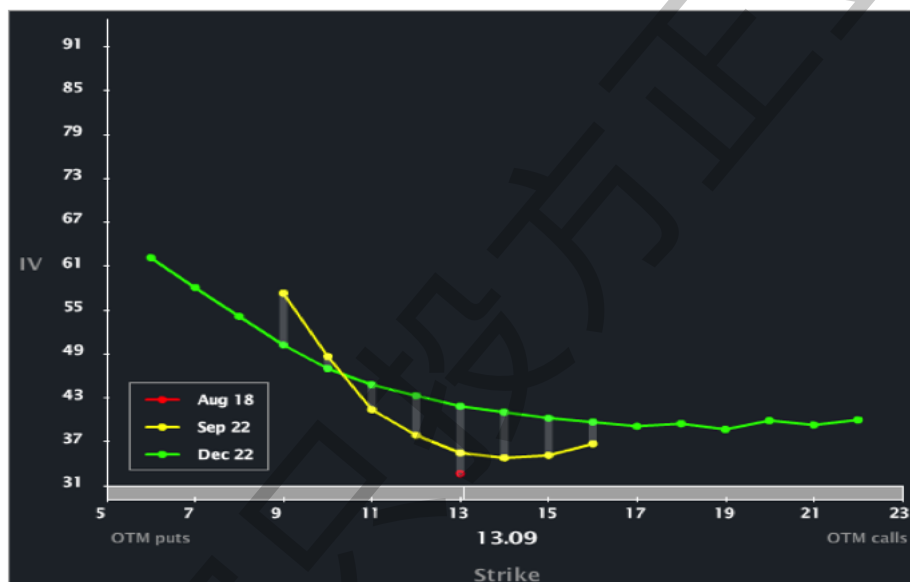
很多人都会做VIX和标普500的套利，在美国VIX被叫做恐慌指数，标普500下跌的时候VIX上涨，或者标普500上涨的时候VIX下跌，有较好的负相关。国内50ETF期权的数据表现正好相反，呈现正相关系数。部分是因为下跌时国家队干预造成的预期影响和股市上涨时投机心理加剧使得隐含波动率加速上涨。所以国内买put的对冲效果很有限，50指数缓慢下跌时隐含波动率也缓慢下跌。50指数快速下跌是国家队往往干预，使得隐含波动率曲折下跌。最后使得卖put的波动率消失殆尽往往得不偿失。



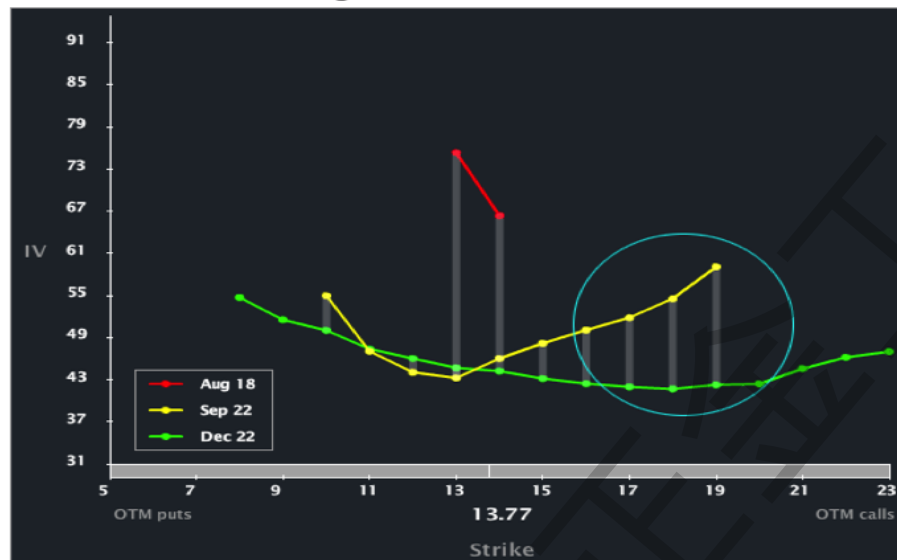
期权也需要关注一个期限结构。期权交易员当然有不同类型的，比如说作为高频交易，或者做市商，我们做波动率交易，还有做方向交易的，单纯买期权或者卖期权。如果做波动

率交易的话，你对波动率曲线需要有一个非常好的理解，这是亚马逊的一个数据，当时交易的时候亚马逊的平价期权在130块钱，2010年的时候，现在是1000块钱了，很难想象亚马逊从创立到现在只有2~3年是盈利，其他所有年份都在亏钱的公司，股票是一年比一年高，所以说美国是不以盈利来评价一家公司的。亚马逊做了很多创新的事情。这样一个波动率期限结构你可以做一下什么交易，举个例子，红色是最早到期的日子，市场公认的是下个月公布季报，如果这个公司宣布这个季报提前5天公布，提前5天的话最好会落在到期日之前，影响是这个期权的波动率一下子会升的很高，提早公布的话，对接下来的几个月份都会有影响，如果是2月份的话，对1月份是没有影响的，所以你要知道不同的信息带给期权的影响，当时如果你买了红色2月份的期权，可能你什么事都没有做，就已经赚了不少钱了。同样美国股息发布的时间和要派发的大小都是影响非常大的。国内的风险会稍微小一点，因为国内股息会调整行权价格不会影响50ETF期权的定价。

## EA – before rumor



## EA – rumor emerges



这是个简单的例子关于不同的事件会怎样影响我们期权的定价和盈亏，今天就介绍到这里，谢谢大家！