

知乎

首发于
川流不息

估值过高能否预测市场下行风险？



石川

量化交易 话题的优秀回答者

已关注

32 人赞同了该文章

摘要

基于 A 股的实证显示，使用市场估值指标后，发生市场下行的条件概率显著高于非条件概率。市场整体估值对预测股灾有一定的作用。

1 引言

预测市场下行风险不仅是量化投资中的一个热门课题，对于高波动的 A 股市场有更重要的实践意义。学术界的预测模型主要包括基本面模型、统计模型和行为模型三大类（下表）。



基本面模型	Price-to-earnings (P/E) ratio	N/A
	Shiller's CAPE	Lleo and Ziemba (2017)
统计模型	Local martingale model	Jarrow, Kchia, and Protter (2011)
	Stopping time disorder detection model	Shiryaev and Zhitlukhin (2012)
	Earthquake model	Gresnigt, Kole, and Franses (2015)
行为模型	Market crash vs Behavioral biases	Goetzmann, Kim, and Shiller (2016)

出处：Lleo and Ziemba (2018)

知乎 @石川

无论采用哪种模型，其目标是给出未来一段时间市场发生给定程度以上跌幅的条件概率。如果预测准确，它可以帮我们有效的规避下行风险、降低投资损失。

近三十年来，学术界使用以上模型针对海外股市进行了大量的研究。然而由于 A 股的年头较短，这方面的研究相对欠缺。近日，一篇发表在 Journal of Portfolio Management 的文章对于 A 股进行了相关研究（Lleo and Ziemba 2018）。该文使用基本面模型中的估值指标——BSEYD、P/E、CAPE——对沪、深两市的下行风险构建了预测模型和准确性检验。从结果来看，使用估值指标可以显著提高市场下行风险的条件概率，对预测股灾有一定作用。

本文简要介绍 Lleo and Ziemba (2018)。

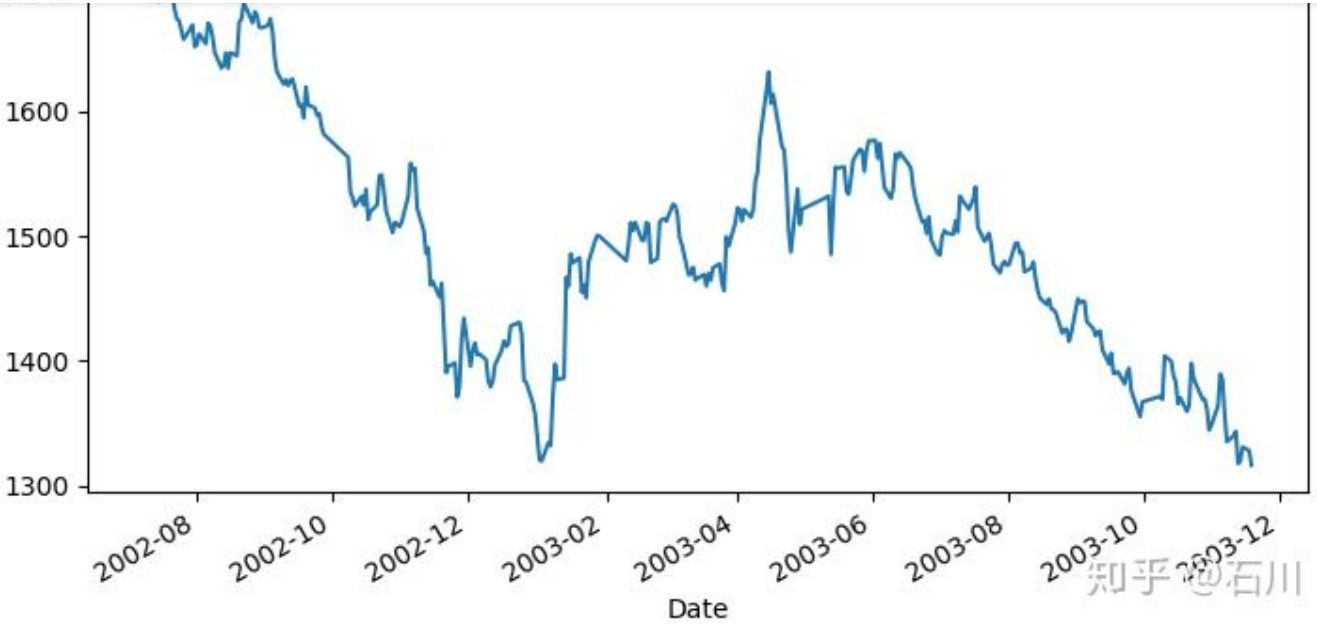
2 下行样本识别

研究的第一步是识别下行样本。根据 Lleo and Ziemba (2018) 的定义，一个下行样本始于某局部高点（该交易日称为 Peak Date）、止于某局部低点（该交易日称为 Trough Date），且两者之间的跌幅不小于 10%。此外，对于一个下行样本，还有一个非常重要的概念——Crash Identification Date，即下行（或股灾）识别日。

从任意局部高点开始，如果在后续的某个交易日的收盘价满足下行识别日条件，该交易日被选做下行识别日，从而确定了下行样本，该下行样本直到再后面的某个局部低点结束。Lleo and Ziemba (2018) 对于下行识别日需要满足的条件限制十分苛刻，目的是为了保证算法找出的下行样本相互独立，否则在市场连续长时间阴跌的时期，挑出的下行样本可能出现相互重叠的现象。

下行识别日在 Lleo and Ziemba (2018) 的模型中的作用十分关键。本文第四节中我们会看到，在评价预测模型准确性时也会用到它。

下图是上证指数上的一个下行样本。根据Lleo and Ziemba (2018) 的识别算法，它从2002/07/08 开始（收盘价 1732.93），然而它并没有在图中第一个“深沟”即 2003/01/03 结束。其原因是算法确认的下行识别日是 2003/04/23，晚于第一个“深沟”的日期。最终，这行样本止于 2003/11/18（收盘价 1316.56），跌幅超过 24%。



实证中，Lleo and Ziemba (2018) 分别考虑了 1990 年到 2016 年的上证指数和深证综指（399106）。按照上述说明，他们在沪、深两市分别找到了 22 和 21 个下行区间，分别汇总于下表中。

Downturns on Shanghai's SHCOMP (1990–2016)

	Crash Identification Date	Peak Date	SHCOMP Index at Peak	Trough Date	SHCOMP Index at Trough	Peak-to-Trough Decline	Peak-to-Trough Duration (days)
1	May 27, 1992	May 25, 1992	1,421.57	Nov. 17, 1992	393.52	72.30%	176
2	Feb. 23, 1993	Feb. 15, 1993	1,536.82	Mar. 31, 1993	925.91	39.80%	44
3	Sep. 19, 1994	Sep. 13, 1994	1,033.47	Feb. 7, 1995	532.49	48.50%	147
4	Aug. 26, 1996	Jul. 24, 1996	887.60	Sep. 12, 1996	757.087	14.70%	50
5	Nov. 6, 1996	Oct. 28, 1996	1,022.86	Dec. 24, 1996	865.577	15.40%	57
6	May 16, 1997	May 12, 1997	1,500.40	Sep. 23, 1997	1,041.97	30.60%	134
7	Aug. 7, 1998	Jun. 3, 1998	1,420.00	Aug. 17, 1998	1,070.41	24.60%	75
8	Jul. 1, 1999	Jun. 29, 1999	1,739.21	Dec. 27, 1999	1,345.352	22.60%	181
9	Sep. 22, 2000	Aug. 21, 2000	2,108.69	Sep. 25, 2000	1,875.911	11.00%	35
10	Feb. 21, 2001	Jan. 10, 2001	2,125.62	Feb. 22, 2001	1,907.26	10.30%	43
11	Jul. 30, 2001	Jun. 13, 2001	2,242.42	Jan. 22, 2002	1,358.691	39.40%	223
12	Apr. 23, 2003	Jul. 8, 2002	1,732.93	Nov. 18, 2003	1,316.562	24.00%	498
13	Apr. 29, 2004	Apr. 6, 2004	1,777.52	Sep. 13, 2004	1,260.316	29.10%	160
14	Aug. 4, 2006	Jul. 11, 2006	1,745.81	Aug. 7, 2006	1,547.436	11.40%	27
15	Feb. 2, 2007	Jan. 24, 2007	2,975.13	Feb. 5, 2007	2,612.537	12.20%	12
16	Jun. 4, 2007	May 29, 2007	4,334.92	Jul. 5, 2007	3,615.872	16.60%	37
17	Nov. 8, 2007	Oct. 16, 2007	6,092.06	Nov. 4, 2008	1,706.703	72.00%	385
18	Aug. 12, 2009	Apr. 4, 2009	3,471.44	Aug. 31, 2009	2,667.745	23.20%	27
19	Oct. 27, 2010	Nov. 23, 2009	3,338.66	Jan. 25, 2011	2,677.432	19.80%	428
20	Dec 27, 2012	Mar. 2, 2012	2,460.69	Jun. 27, 2013	1,950.012	20.80%	482
21	Jun. 25, 2014	Sep. 12, 2013	2,255.60	Jun. 25, 2014	2,025.502	10.20%	286
22	Jun. 19, 2015	Jun. 12, 2015	5,166.35	Aug. 26, 2015	2,927.288	43.30%	75

出处：Lleo and Ziemba (2018)

	Identification Date	Peak Date	Index at Peak	Trough Date	Index at Trough	Decline	Duration (days)
1	Jun. 3, 1992	May 26, 1992	312.21	Jun. 16, 1992	233.73	25.10%	21
2	Mar 5, 1993	Feb. 22, 1993	359.44	Jul. 21, 1993	203.91	43.30%	149
3	May 10, 1996	May 22, 1995	169.66	Aug. 26, 1996	152.55	10.10%	462
4	Sep. 10, 1996	Sep. 4, 1996	274.56	Dec. 24, 1996	242.01	11.90%	111
5	May 16, 1997	May 12, 1997	517.91	Sep. 23, 1997	312.73	39.60%	134
6	Jul. 6, 1998	Jun. 3, 1998	441.04	Aug. 18, 1998	317.1	28.10%	76
7	Jul 1, 1999	Jun. 29, 1999	525.14	Dec. 27, 1999	395.69	24.70%	181
8	Sep. 25, 2000	Aug. 21, 2000	643.77	Sep. 25, 2000	578.76	10.10%	35
9	Feb. 8, 2001	Nov. 23, 2000	654.37	Feb. 22, 2001	568.26	13.20%	91
10	Jul. 30, 2001	Jun. 13, 2001	664.85	Jan. 22, 2002	371.79	44.10%	223
11	Apr 26, 2004	Apr. 7, 2004	470.55	Sep. 13, 2004	315.17	33.00%	159
12	Aug. 2, 2006	Jul. 12, 2006	446.61	Aug. 7, 2006	380.26	14.90%	26
13	Jun. 1, 2007	May 29, 2007	1,292.44	Jul. 5, 2007	1,015.85	21.40%	37
14	Oct. 25, 2007	Oct. 9, 2007	1,551.19	Nov. 28, 2007	1,219.98	21.40%	50
15	Jan. 22, 2008	Jan. 15, 2008	1,576.50	Nov. 4, 2008	456.97	71.00%	294
16	Aug. 14, 2009	Aug. 4, 2009	1,149.27	Sep. 1, 2009	900.53	21.60%	28
17	Dec. 22, 2009	Dec. 3, 2009	1,234.17	Jul. 5, 2010	921.34	25.30%	214
18	Nov. 17, 2010	Nov. 10, 2010	1,389.54	Jan. 25, 2011	1,136.58	18.20%	76
19	Jun. 24, 2013	May 30, 2013	1,043.47	Jul. 25, 2013	879.93	15.70%	26
20	Mar. 28, 2014	Feb. 17, 2014	1,160.39	Apr. 28, 2014	1,007.27	13.20%	70
21	Jun. 19, 2015	Jun. 12, 2015	3,140.66	Sep. 15, 2015	1,580.26	49.70%	95

出处：Leo and Ziemba (2018)

知乎 @石川

接下来就是针对这些下行样本构建预测模型。

3 预测模型

下行风险预测模型使用选定的基本面估值指标 —— 如 BSEYD、P/E ratio 或 CAPE —— 计算信号：信号等于 1 说明在未来有跌幅超过 10% 的下行风险；信号为 0 说明没有该风险。此外，Leo and Ziemba (2018) 给出了评价该模型准确性的算法。本节和下节分别介绍预测模型和评价模型准确性的算法。

预测模型的原理非常简单。假设基本面指标在交易日 t 的取值为 M_t ，使用该指标在过去一段长度为 d 的窗口内的均值和标准差计算出一个阈值 K_t ；如果 $M_t > K_t$ ，则认为市场估值过高，因而下行风险信号 $S_t = 1$ 、反之则认为下行风险信号 $S_t = 0$ 。

在计算阈值 K_t 时，Leo and Ziemba (2018) 选择的滚动窗口长度 d 为 252 个交易日，从而动态计算不同市场环境下估值的阈值。在具体计算 K_t 时，他们采用了两种方法：（1）正态分布假设；（2）Cantelli 不等式。

由于仅考虑估值过高风险，故采用单边 95% 的置信区间。在正态分布的假设下，该阈值 K_t 对应均值加上 1.645 个标准差：

知乎

首发于
川流不息

$$\text{where } \mu_t^d = \frac{1}{d} \sum_{i=0}^d M_{t-i}$$

$$\sigma_t^d = \sqrt{\frac{1}{d-1} \sum_{i=0}^{d-1} (M_{t-i} - \mu_t^d)^2}$$

考虑到市场估值指标在任意一段给定长度为 d 的窗口内不一定很好的满足正态分布假设, Lleo and Ziemba (2018) 同时使用了Cantelli 不等式作为另外一种计算阈值 K_t 的方法。该不等式是我们熟悉的 Chebyshev' s inequality (切比雪夫不等式) 的单尾形式。对于未知分布的随机变量 M , 假设其均值和标准差为 μ 和 σ , 由 Cantelli 不等式有:

$$\text{prob}(M - \mu \geq \gamma\sigma) \leq \frac{1}{1 + \gamma^2}$$

上式中, 另 $\beta = 1/(1+\gamma^2)$ 可得:

$$\text{prob} \left(M - \mu \geq \sigma \sqrt{\frac{1}{\beta} - 1} \right) \leq \beta$$

在模型中, Lleo and Ziemba (2018) 选择 $\beta = 0.25$, 对应 σ 的乘数为 1.732, 比正态分布下的 1.645 略高。

采用上述正态分布模型或 Cantelli 不等式即可计算出给定交易日 t 的估值阈值上限 K_t 。通过比较当日实际估值指标 M_t 和 K_t 的关系, 就能发出市场是否在未来出现下行风险的信号 S_t 。在这个判断中, 必须要回答的另一个问题就是: **这个信号管多久 —— 市场在未来多长时间窗口内 (记为 H) 会出现 10% 以上的下行风险。** H 的选择是预测模型设定问题, 在 Lleo and Ziemba (2018) 中其取值为 252 个交易日。

4 模型准确性评价

接下来介绍如何评价预测模型的准确性。

当模型给出 $S_t = 1$ 的信号后, 如果在 t 之后的 H 个交易日包含了某个下行识别日 (回想一下第二节的定义), 那么则认为发生了市场下行 (记为 $C_{t,H} = 1$)、模型信号准确。根据这个定义, 预测模型的准确性可以由下面这个条件概率度量:

$$\text{prob}(C_{t,H} = 1 | S_t = 1)$$



知乎

首发于
川流不息

识别日，即可计算出上述条件概率。数学推导可知，使用样本数据计算的条件概率也是模型准确性的 maximum likelihood estimate（具体推导请见 Lleo and Ziemba 2018，这里不再赘述）。

别着急，还没完。

由于 A 股牛短熊长、波动巨大的特点，从任何一个天往后看 252（H 的取值）个交易日，出现下行风险（跌 10% 以上）的先验概率都很大（实证数据显示，这个先验概率大概是 70% 左右）。
只有当预测模型给出的条件概率显著高于这个非条件先验概率时，我们才能说使用估值指标进行预测能带来一定的价值。

出于这个考虑，Lleo and Ziemba (2018) 使用 Likelihood Ratio Test 对此进行了假设检验。该检验的 null hypothesis 是 $p = p_0$ ，这里 p_0 是下行先验概率——即在实证期内任选一天，在其后 252 个交易日内包含下行识别日的概率。该假设的 Likelihood Ratio Test 如下：

$$\Lambda = \frac{L(p = p_0)}{L(p = \hat{p})}$$

其中 L 为 likelihood function， \hat{p} 是预测模型给出的条件概率（ $p = \hat{p}$ 为 alternative hypothesis）。使用 Λ 定义 test statistic: $Y = -2\log\Lambda$ ，它满足自由度为 1 的 chi-squared 分布，其 10%、5%、1% 显著性水平下对应的阈值分别为 2.71、3.83 和 6.63。如果实证数据给出的 Y 值超过上述某个阈值，则我们就可以在相应的显著性水平下拒绝原假设。

5 实证结果

下面来看实证结果。

Lleo and Ziemba (2018) 首先考察了 P/E 和 $\log(P/E)$ 两个市场估值指标。下表总结了模型的预测结果以及准确性分析结果，其中对于沪、深两市各有四行，这是因为对于每个指标，都分别有两种计算阈值的方法——正态分布和 Cantelli 不等式。



Signal Model	Total Number of Signals	of Correct Predictions	Likelihood Estimate \hat{p}	$L(\hat{p})$	Likelihood Ratio Λ	Test Statistics $-2 \log \Lambda$	p -value
SHCOMP							
PE0 (confidence)	19	17	89.47%	1.67e-3	0.1159	4.31*	3.79%
PE0 (Cantelli)	18	16	88.89%	1.88e-3	0.1486	3.81†	5.09%
logPE0 (confidence)	18	16	88.89%	1.88e-3	0.1486	3.81†	5.09%
logPE0 (Cantelli)	19	17	89.47%	1.67e-3	0.1159	4.31*	3.79%
SZECOMP							
PE0 (confidence)	9	8	88.89%	4.33e-2	0.1313	4.06*	4.39%
PE0 (Cantelli)	9	8	88.89%	4.33e-2	0.1313	4.06*	4.39%
logPE0 (confidence)	9	8	88.89%	4.33e-2	0.1313	4.06*	4.39%
logPE0 (Cantelli)	8	7	87.5%	4.91e-2	0.1980	3.24†	7.19%

[†]Significant at the 10% level. *Significant at the 5% level. **Significant at the 1% level. ***Significant at the 0.5% level.

出处：Leo and Ziemba (2018)

知乎 @石川

以上证指数为例，P/E 指标在这使用这两种计算阈值方法时，分别给出了 19 和 18 次下行风险信号，其中正确的次数分别为 17 和 16 次，因此条件概率高达 89% 上下。另一方面，对于上证指数，出现下行风险的非条件先验概率为 70%。

通过假设检验可知，test statistic Y 的取值在这两种方法下分别为 4.31（5% 的显著性水平）和 3.81（10% 的显著性水平），因此可以说基于 P/E 估值指标的预测模型对于鉴别市场下行风险提供了显著的增量信息。我们可以对上表中其他各行的结果做类似的解读。

除了预测信号的准确性外，我们也关注召回率。对于沪深两市，在整段实证期内分别由 22 和 21 次下行样本。使用 P/E 及 log(P/E)，在上证指数上正确找到 17 个，召回率为 77%；而在深证综指上正确找出 8 个，召回率仅有可怜的 38%。**该模型的召回率在沪深两市上的巨大差异值得进一步的研究。**

除了 P/E 和 log(P/E) 之外，Leo and Ziemba (2018) 还考虑了其他估值指标：BSEYD、log(BSEYD)、CAPE、log(CAPE)。由于计算 CAPE 需要用到十年的历史数据，因此预测模型的实证期为 2005 年到 2016 年。下表给出了不同估值指标的预测结果。

知乎

首发于
川流不息

Signal Model	Total Number of Signals	Number of Correct Predictions	Maximum Likelihood Estimate \hat{p}	$L(\hat{p})$	Likelihood Ratio Λ	Test Statistics $-2 \log \Lambda$	p -value
Panel A: SHCOMP							
BSEYD0	4	3	75.00%	1.05e-1	0.7170	0.67	41.47%
logBSEYD0	5	2	40.00%	3.46e-2	0.7901	0.47	49.24%
PE0	4	3	75.00%	1.05e-1	0.7170	0.67	41.47%
logPE0	4	3	75.00%	1.05e-1	0.7170	0.67	41.47%
BSEYD10	3	2	66.67%	1.48e-1	0.9228	0.16	68.86%
logBSEYD10	5	3	60.00%	3.46e-2	0.9778	0.04	83.23%
CAPE10	5	3	100.00%	—	—	∞	0.00%***
logCAPE10	5	3	100.00%	—	—	∞	0.00%***
Panel B: SZECOMP							
BSEYD0	6	6	100.00%	—	—	∞	0.00%***
logBSEYD0	7	7	100.00%	—	—	∞	0.00%***
PE0	6	6	100.00%	—	—	∞	0.00%***
logPE0	6	6	100.00%	—	—	∞	0.00%***
BSEYD10	7	6	85.71%	5.67e-2	0.5266	1.28	25.74%
logBSEYD10	7	7	100.00%	—	—	∞	0.00%***
CAPE10	6	6	100.00%	—	—	∞	0.00%***
logCAPE10	5	5	100.00%	—	—	∞	0.00%***

[†]Significant at the 10% level. *Significant at the 5% level. **Significant at the 1% level. ***Significant at the 0.5% level.

出处: Lleo and Ziemba (2018)

知乎 @石川

首先需要插一句的是，上表中特别用红框标出来的部分应该是 typo。根据 Lleo and Ziemba (2018) 上下文的描述，对于上证指数，CAPE 和 log(CAPE) 应该各自发出了 3 次信号、而非 5 次，且 3 次全部正确，因此准确率才有 100%。在此特别“鄙视”一下审稿人。

*For the SHCOMP, none of the measures produced more than five signals. The CAPE, logCAPE, and BSEYD10 generated **three** signals each.*

下面来讨论结果。非常有意思的是，对于 2005 年之后的市场，该模型——无论使用什么指标——在沪、深两市上的效果较整段实证期来说发生了反转。对于深证综指，这些指标的准确率和召回率都更高，而对于上证指数，仅 CAPE 和 log(CAPE) 较非条件先验概率显著的提升了准确性。然而从召回率来看，自 2005 年到 2016 年之间，上证指数一共出现 6 次下行样本，而上述两个指标仅找到三次，召回率仅有 50%。

Lleo and Ziemba (2018) 对于模型在沪、深两市上的差异也做了思考。除了样本量较少造成的误差之外，二位作者认为该分歧也和沪深两市上市公司的属性差异很大有关。上交所的上市公司多为国有企业，而深交所的上市公司多为民营企业。二位作者猜测，政府对国有企业支持而造成对市场的干预，有可能会推迟和减少下行风险的出现，从而影响预测模型的效果。

6 结语

预测市场什么时候要 crash 从来就不容易。将市场的 P/E 或者 CAPE 这类估值指标和未来收益关联起来的尝试并也不是什么新鲜事物。但今天之所以介绍 Lleo and Ziemba (2018) 是希望 A 股上的实证结果能给我们一些启发。

知乎

首发于
川流不息

The CAPE was never intended to indicate exactly when to buy and to sell. The market could remain at these valuations for years. But we should recognize that we are in an unusual period, and that it's time to ask some serious questions about it.

从 Lleo and Ziemba (2018) 的结果以及其他针对 A 股的实证分析来看，过高估值和未来的下行风险之间肯定存在着某种正相关性。然而，要想利用这种相关性、将其作为整套量化系统（特别是风险管理）的一部分，或者使用它进行市场或者其他风格因子择时，仍有很多的问题需要解答。

参考文献

- Goetzmann, W., D. Kim, and R. Shiller (2016). *Crash beliefs from investor surveys*. Working paper.
- Gresnigt, F., E. Kole, and P. Franses (2015). Interpreting financial market crashes as earthquakes: a new early warning system for medium-term crashes. *Journal of Banking & Finance*, Vol. 56, 123 – 139.
- Jarrow, R. A., Y. Kchia, and P. Protter (2011). How to detect an asset bubble. *SIAM Journal on Financial Mathematics*, Vol. 2(1), 839 – 865.
- Lleo, S. and W. T. Ziemba (2017). Does the bond-stock earnings yield differential model predict equity market corrections better than high P/E models? *Financial Markets, Institutions & Instruments*, Vol. 26(2), 61 – 123.
- Lleo, S. and W. T. Ziemba (2018). Predicting stock market crashes in China. *The Journal of Portfolio Management*, Vol. 44(5), 125 – 135.
- Shiryaev, A. N. and M. V. Zhitlukhin (2012). Bayesian disorder detection problems on filtered probability space. *Theory of Probability and Its Applications*, Vol. 57(3), 497 – 511.
- Ziemba, W. T. and S. L. Schwartz (1991). *Invest Japan*. Chicago: Probus.

免责声明：文章内容不可视为投资意见。市场有风险，入市需谨慎。

原创不易，请保护版权。如需转载，请联系获得授权，并注明出处，谢谢。已委托“维权骑士”（[维权骑士_免费版权监测/版权保护/版权分发](#)）为进行维权行动。

编辑于 2019-07-03

股灾 风险控制 量化交易

▲ 赞同 32 ▼ 3 条评论 分享 收藏 ...

知乎

首发于
川流不息

川流不息

北京量信投资管理有限公司是一家在中国基金业协会备案登记的专业私募基金管理人...

关注专栏

推荐阅读

那些年，做期货量化与高频的日子

读书的时候写了不少留学申请与就业的文章，现在就谈谈这几年做 futures quantitative and high frequency trading 的经历吧。2011年硕士毕业在美国开始了第一份工作。那一年办h1b还不用抽...
babyquant

利用简单指标，建立交易系统

<http://code.tradeclassroom.com>
交易策略源代码下载运用非对称随机指标，增加移动平均线和资金管理规则。对一些技术指标进行改进，或是利用一些非传统的方式，可以让交易者获利。在此，我...
insking

风险控制-何定义-20

关于这个问题，在专栏的文章里，重点不在细节，后来有朋友问起，由于事情复杂，所以在谋杀电视机

3 条评论

⇌ 切换为时间排序

写下你的评论...



SagittaireX sun

4 个月前

川总，有没有推荐的偏实践的风险量化相关的期刊？

👍 1



太昊悦和

3 个月前

拥有最大概率的预测是：上涨，或下跌。两个相背而行的预测，各占50%的概率，在各种算法预测的结论中，拥有极高的实用参考价值。

我们仅就上述结论提出商榷。涨跌二元性不能完全对股票趋势进行描述，我们通过对沪深二级市场所有可交易标的进行的实证研究，结果揭示了实际存在“不涨，就跌，或平”这一更加完善的趋势性描述，并揭露了上都特么属于废话这一赤裸裸的真相。这一结论因采样符合正态，除了有统计学意义，更具备极高的实用价值。

本年度糯耳朵扯淡学奖入围成果——基于沪深二级市场所有可交易标的任意两周期运行趋势

宁梓沁沁人，英江白史。

 赞  回复  踩  举报



太昊悦和 回复 太昊悦和

3 个月前

上文刊发时，因作者并不用心，错将第二自然段“我们谨就”误录为“我们仅就”，作者表示打死也绝不向阅读者致歉。

 赞

