

2017 年 04 月 09 日

## 事件驱动系列专题报告——高送转预测(逻辑回归 VS 随机森林)

■高送转事件投资逻辑：大比例的送转股票表明了企业对自身业绩持续增长的信心，同时高股价导致股票流动性差的问题也可以得以解决。因此高送转行情每逢年报或者中报披露期前后总会引起市场强烈关注，期间概念股走势明显跑赢大盘。高送转投资关键词：高送转比例+高业绩支撑

■影响高送转实施因子分析：选股时应规避无业绩支撑的“滥送转”股票，报告将从估值类因子、规模类因子等维度分析因子有效性。每股资本公积、每股未分配利润、每股收益(EPS)、每股净资产(BPS)、股票价格这 5 个因子与高送转实施概率呈显著正相关，即该值越大实施高送转概率越大；而总股本和上市天数与实施概率呈显著负相关。

■高送转预测模型构建及回测：报告使用了逻辑回归和随机森林两种方法去分别构建预测模型，选择上述 7 个因子作为输入变量。

逻辑回归：将预测概率在 0.5 及以上的公司判定为发生高送转，则模型预测为高送转并实际发生的平均准确率大概在 60%左右。同时前 30 名股票组合预测结果表现最好，准确率可以达到 70%。

随机森林：参数选择方面选取预测时点过去三年的分红实施数据作为训练集，同时设决策树数量为 1000，特征数目为 3。随机森林预测模型表现较优于逻辑回归预测模型，整体平均预测精度保持在 70%，同时不同排名水平的预测准确度表现也比较稳定。

■高送转填权行情分析：“填权”行情一般持续 2 周左右，2 周之后的股票收益开始出现回撤。报告从前期行情、送转比例以及年报业绩三方面来分析高送转填权股特征：

1. 预案公告日到分红实施日期间跑出大行情的个股在除权日之后大概率出现回撤。
2. 高送转比例是布局“填权”行情时市场关注焦点。
3. 涨幅靠前的高送转填权股年报业绩表现较好，净利润同比均值在 70%左右。

■风险提示：市场风格变化可能导致模型失效风险、系统性风险、流动性风险

## 金融工程主题报告

证券研究报告

刘帅路

分析师

SAC 执业证书编号：S1450517030003  
liusl1@essence.com.cn

刘潇

报告联系人

liuxiao1@essence.com.cn  
010-83321032

## 相关报告

(场内基金和量化配置周报)	
四月增配 A 股美股减配商品	2017-04-05
安信金工次新股周报：关注超跌反弹次新股	2017-04-04
安信金工主题和强势股周报：看好军民融合概念股	2017-04-02
事件驱动系列专题报告——从业绩预告捕捉摘帽行情	2017-04-01
多家 LOF 募集火热，一带一路主题领涨	2017-03-26

## 内容目录

<b>1. 高送转事件投资逻辑</b>	<b>3</b>
1.1. 送转比例与市场关注度成正比	3
1.2. 高比例 $\neq$ 高业绩增长	3
<b>2. 影响高送转实施的因子分析</b>	<b>4</b>
2.1. 每股资本公积	4
2.2. 每股未分配利润	4
2.3. 股票价格	5
2.4. 每股收益(EPS)	5
2.5. 每股净资产(BPS)	6
2.6. 总股本	6
2.7. 上市时间	7
<b>3. 高送转预测模型构建</b>	<b>7</b>
3.1. 基于逻辑回归的预测模型	7
3.2. 基于随机森林的预测模型	8
<b>4. 高送转填权行情分析</b>	<b>9</b>
4.1. 前期行情	9
4.2. 送转比例	10
4.3. 年报业绩	10
4.4. 最新高送转“填权”组合	11
<b>5. 风险提示</b>	<b>11</b>

## 图表目录

图 1: 预案公告日后不同送转比例下个股平均收益情况	3
图 2: 历年分红实施方案中送转比例情况	3
图 3: 不同每股资本公积下股票实施高送转概率统计	4
图 4: 不同每股未分配利润下股票实施高送转概率统计	5
图 5: 不同股价下个股实施高送转概率统计	5
图 6: 不同每股收益下个股实施高送转概率统计	6
图 7: 不同每股净资产下个股实施高送转概率统计	6
图 8: 不同上市时间下个股实施高送转概率统计	7
图 9: 不同排名水平下逻辑回归模型预测准确率	8
图 10: 不同排名水平下随机森林模型预测准确率	9
图 11: 前期行情 vs 填权行情	10
图 12: 送转比例 vs 填权行情	10
图 13: 年报业绩 vs 填权行情	11
表 1: 不同股本区间下实施高送转公司数量分布情况	6
表 2: 逻辑回归模型中主要解释变量	8
表 3: 变量重要性度量结果	9
表 4: 最新高送转“填权”组合	11

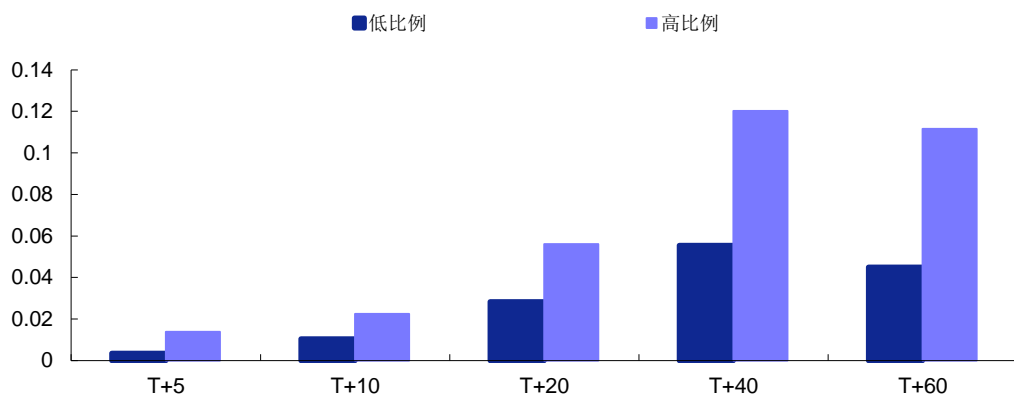
## 1. 高送转事件投资逻辑

高送转，顾名思义就是指高比例的送股或转增股本，通常送转比例达到 10 股送或转 5 股及以上。对于上市公司而言，大比例的送转股票表明了其对自身业绩持续增长的信心，同时高股价导致的股票流动性较差，也可以通过高送转的方式得以解决。因此，高送转释放的利好信号一直受到市场的强烈关注，每逢年报或者中报披露期前后总会爆发一轮行情，期间高送转概念股走势明显跑赢大盘。所以如何选对标的成为了把握高送转收益的关键。由于年度分红无论从样本数量还是信息完整度方面对预测模型来说都更具统计价值，因此本篇报告的后续分析主要集中在年报高送转上面。

### 1.1. 送转比例与市场关注度成正比

市场对于高送转比例的个股反响如何呢？以 10 股送转 5 股为基准将分红样本分为两组，统计分红预案公告日后个股涨跌幅情况。结果表明，高送转个股在公告日后 40 天的平均累计收益为 12.03%，表现明显优于低比例个股的 5.54%。

图 1：预案公告日后不同送转比例下个股平均收益情况

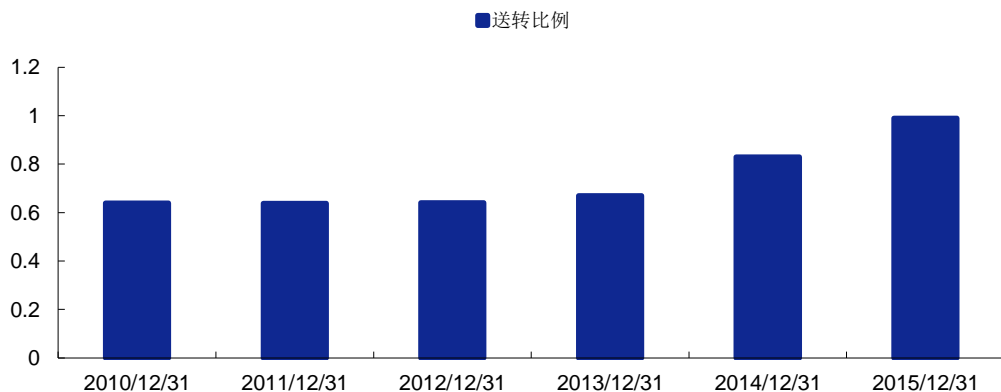


资料来源：wind，安信证券研究中心

### 1.2. 高比例 ≠ 高业绩增长

我们统计了报告期为 2010-2015 年年报、分红实施方式为送股或转增股本的全部分红方案，结果发现平均送转比例基本呈现逐年递增的趋势，6 年间这一数字从 0.64 上升到 0.99。趋势上来看，上市公司为了吸引眼球开始争相提高转送力度。

图 2：历年分红实施方案中送转比例情况



资料来源：wind，安信证券研究中心

然而是否每一个高送转比例背后都有相应的业绩作为支撑呢？近年来，高送转脱离基本面的情况时有发生，像去年的南威软件，公司在净利润连年下滑的情况下依然拟 10 股转 30 股派息 1.6 元。因此在选股方面，投资者不应盲目追求送转比例，需结合公司基本面等因素理性投资。

## 2. 影响高送转实施的因子分析

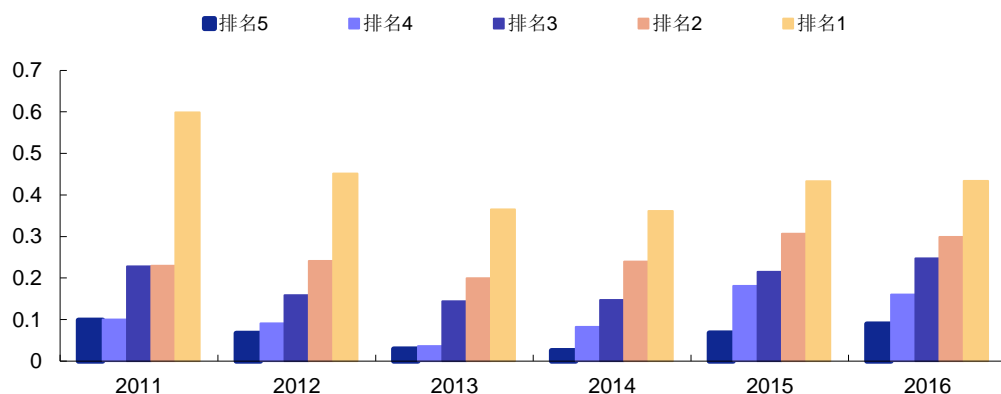
上文我们提到选股时应规避无业绩支撑的“滥送转”股票，下面我们将主要从估值类因子、规模类因子等维度分析因子有效性。

**统计口径：**定义 10 股送转 5 股及以上的个股为高送转股票，提取报告期为 2010-2015 年所有年报分红实施方案来分析下述单因子表现。由于高送转预期行情需在分红预案公告日之前布局，而几乎所有分红预案日与年报披露日为同一天，因此考虑到预测模型，后续分析均选择公司三季报财务数据，同时剔除每年三季报披露日之后上市的公司。

### 2.1. 每股资本公积

可以说，每股资本公积越高带来的高送转预期信号就越强。2011 到 2016 年，实际发生高送转的个股平均每股资本公积金为 2.8 元，是未发生高送转个股的 2 倍。同时，如果每股资本公积在 10 元以上的话，公司实施高送转的概率超过 80%。我们将每股资本公积从高到低分成 5 组去统计股票发生高送转的概率，结果表明每股资本公积与实施概率呈现明显单调性，资本公积最高的组实施概率明显高于低排名组。

图 3：不同每股资本公积下股票实施高送转概率统计

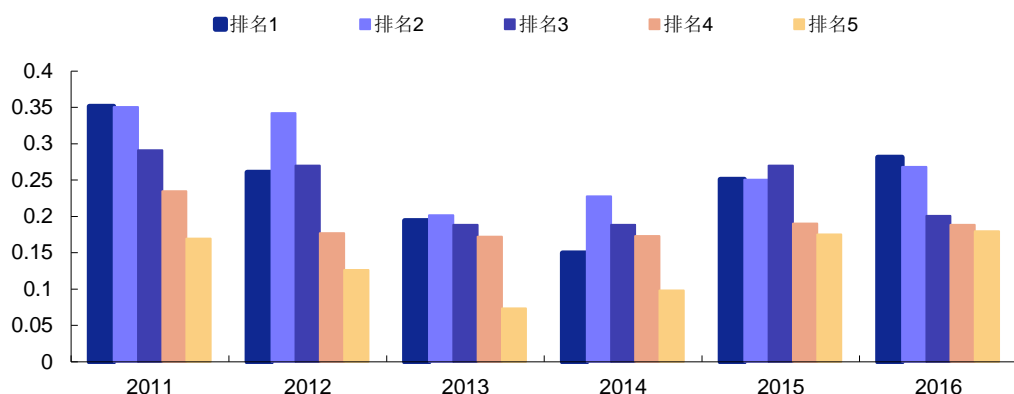


资料来源：wind，安信证券研究中心

### 2.2. 每股未分配利润

每股未分配利润也是衡量企业的分红和盈利能力的重要指标。我们同样按照上面做法将未分配利润从高到低分 5 组，排名最高的组表现优于低排名组，该因子在实施高送转概率上也体现了较好的单调性。

图 4：不同每股未分配利润下股票实施高送转概率统计

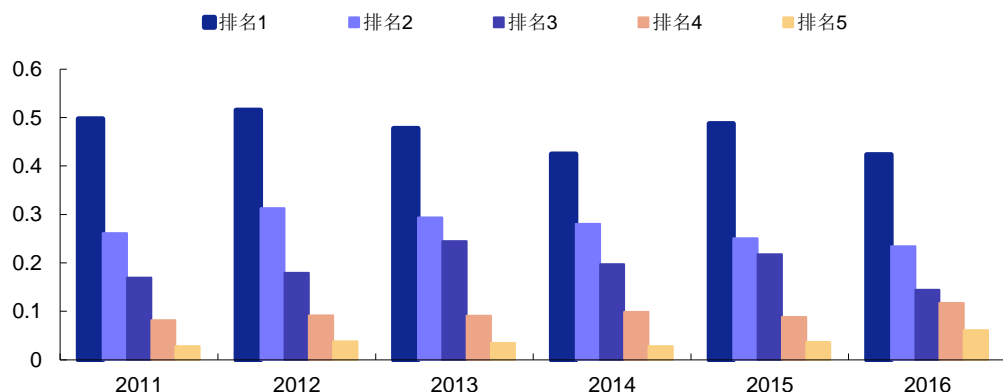


资料来源：wind，安信证券研究中心

### 2.3. 股票价格

高价格股票为了改善其流动性，公司一般实施高送转的意愿比较强。根据统计结果，2011 到 2016 年实施高送转的股票均价在 25 元左右。股票价格高于 50 元的企业实施高送转的比例接近 6 成。股票价格与送转概率呈现显著正相关性，排名最高的组发生高送转概率达到 47%，而低于 10 元的个股仅有不到 10% 的概率实施高送转。

图 5：不同股价下个股实施高送转概率统计

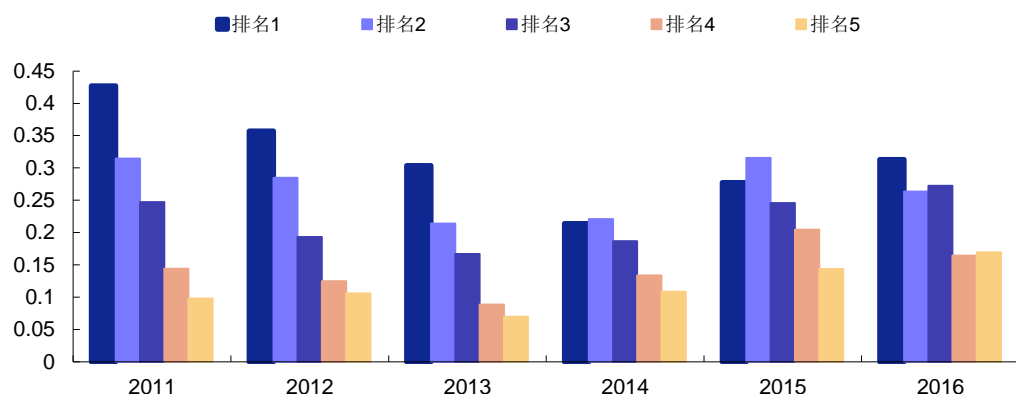


资料来源：wind，安信证券研究中心

### 2.4. 每股收益(EPS)

每股收益直接反应了公司的盈利能力，该指标越大代表企业分红资源越充足。总体来看，EPS 排名最高的组发生高送转的概率几乎是排名最低组的 3 倍。

图 6：不同每股收益下个股实施高送转概率统计

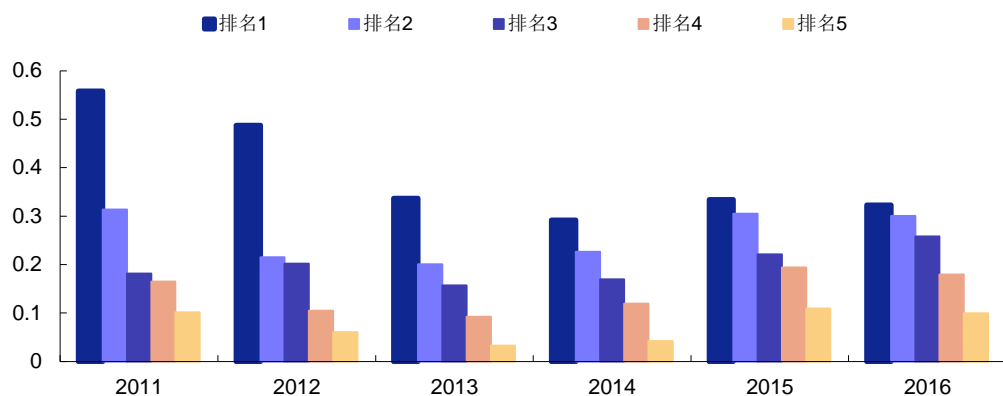


资料来源：wind，安信证券研究中心

## 2.5. 每股净资产(BPS)

从每股净资产的角度来看，最高排名的企业实施高送转的平均概率将近 40%，而最低排名组的发生概率仅为 7.44%。

图 7：不同每股净资产下个股实施高送转概率统计



资料来源：wind，安信证券研究中心

## 2.6. 总股本

股本越低的公司实施高送转意愿越强。2011-2016 年实施高送转公司的平均总股本为 4.1 亿股，远小于未实施高送转的 24.7 亿股。此外，总股本小于等于 1 亿股的公司实施高送转的概率为 62%。

表 1：不同股本区间下实施高送转公司数量分布情况

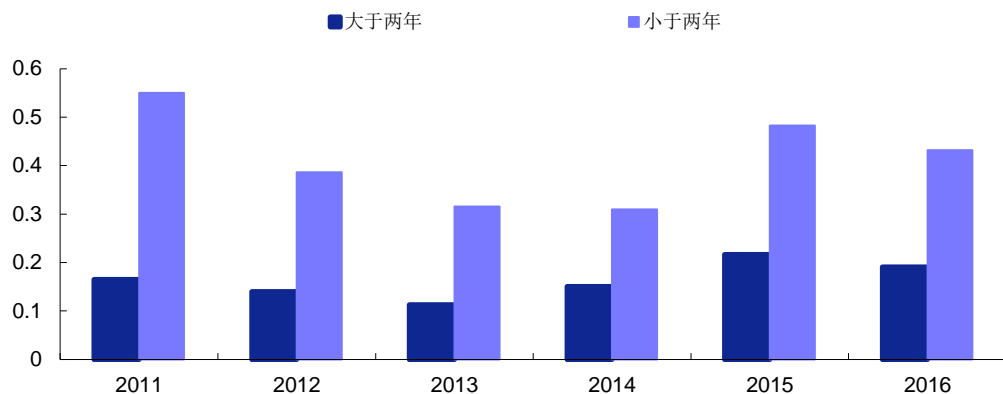
区间	数量	占比
[0, 10]	2017	92.78%
(10, 20]	96	4.42%
(20, 30]	30	1.38%
(30, 50]	20	0.92%
(50,150]	11	0.51%

资料来源：wind，安信证券研究中心

## 2.7. 上市时间

上市越晚的公司发生高送转的概率越大。结合往年数据，我们可以看到实施高送转个股中次新股所占的比重也比较大。上市时间在两年内的公司发生高送转的概率超过 40%，明显高于上市时间大于两年的公司。

图 8：不同上市时间下个股实施高送转概率统计



资料来源：wind，安信证券研究中心

## 3. 高送转预测模型构建

统计来看，平均每年实施高送转的股票数量约为 380 只。虽然 15-16 年这一数字提升到了 450 左右，但在全部 A 股当中的占比还是很低的。因此通过预测模型进一步筛选出高送转预期个股是准确把握高送转行情的必要前提。下面我们将使用逻辑回归和随机森林两种方法分别构建高送转事件预测模型。

### 3.1. 基于逻辑回归的预测模型

Logit 模型是常用的一种二元离散模型，主要用来预测离散因变量与一组解释变量之间的关系以及因变量事件发生的概率。最常用的是二值选择模型，即因变量的取值只包含两个类别，如事件发生、事件不发生，常用  $Y=1$  或  $Y=0$  表示。

Logit 模型的核心公式如下：

$$\ln\left(\frac{P_i}{1-P_i}\right) = \beta_0 + \beta_1 x_{1i} + \dots + \beta_n x_{ni} + \varepsilon_i$$

其中  $P_i$  为因变量事件发生的概率， $x_{ni}$  为选定的解释变量， $\varepsilon_i$  为残差项。

构建 Logit 预测模型时，我们选取了 2011-2016 年发布分红实施方案的全部 A 股进入样本池，定义 10 股送转 5 股及以上比例为高送转事件，将上述分析的每股资本公积、每股未分配利润等 7 个因子作为解释变量，最后使用极大似然估计(MLE)对模型参数进行求解。结果显示所有解释变量都在 5% 的显著性水平下显著。预测模型中我们选定的解释变量和影响方向如下表所示：

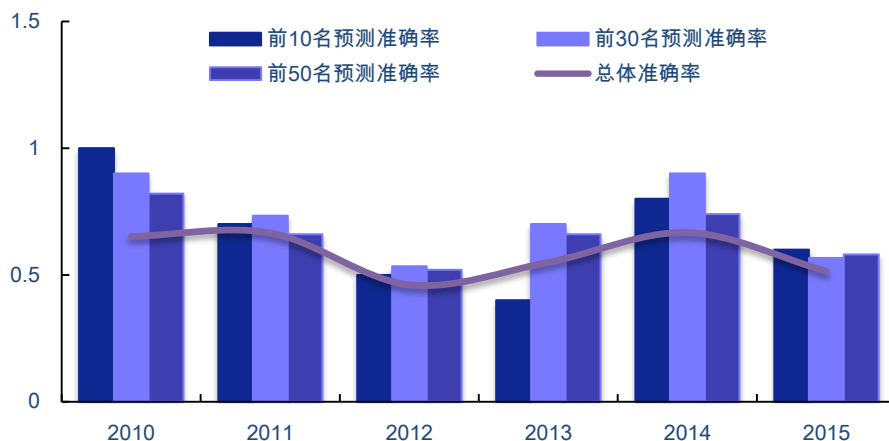


**表 2：逻辑回归模型中主要解释变量**

解释变量	具体选取指标	影响方向
$x_1$	上一年度年末总股本	负向
$x_2$	上一年度三季报每股资本公积	正向
$x_3$	上一年度三季报每股未分配利润	正向
$x_4$	上一年度年末股价	正向
$x_5$	上市天数	负向
$x_6$	上一年度三季度每股收益	正向
$x_7$	上一年度三季度每股净资产	正向

资料来源：wind，安信证券研究中心

我们将模型预测概率在 0.5 及以上的公司判定为发生高送转，0.5 以下判定为不发生，模型预测为高送转并实际发生了的平均准确率大概在 60% 左右。此外我们还选取了每年模型预测高送转概率最高的前 10 名、30 名和 50 名股票分别检验预测准确率，其中发现前 30 名股票组合预测结果表现最好，准确率可以达到 70%。具体各组合预测准确率如下图所示：

**图 9：不同排名水平下逻辑回归模型预测准确率**


资料来源：wind，安信证券研究中心

### 3.2. 基于随机森林的预测模型

随机森林，是用随机的方式构建一个由许多决策树组成的分类器，每一颗决策树之间的相关性很低。在构建每一个决策树时，都是一个随机取样和完全分裂的过程。假设有  $N$  个输入样本和  $M$  个输入变量，首先对输入样本进行  $N$  次 bootstrap 抽样(采用有放回的重重复抽样方式)生成决策树的训练集，之后在决策树的每个节点随机选择  $m(m < M)$  个特定的变量进行分析以确定最佳分裂点，尽最大可能地完全分裂下去。由于过程的随机性，一般不进行“剪枝”操作。最后将一个低相关性的决策树结果组合起来形成最终的预测结果。由于单一决策树的结构对随机森林模型影响较小，比较不容易出现过拟合的现象。

**参数选择：**选取预测时点过去三年的分红实施数据作为训练集，主要是因为过少的数据作为训练集会丢失信息，导致预测稳定性较差；同时如果选取过多的数据又会由于较长时间跨度使预测和训练的相关性下降，进而降低预测精确度。此外，设决策树的数量为 1000，特征数目为 3，即在每一个节点上随机从上述 7 个因子中抽取 3 个去进行分裂。

训练得到预测模型之后，我们首先来评估下 7 个因子在模型中的有效性。表 3 为因子有效性检验相关参数。MeanDecreaseAccuracy 是一种变量重要性度量方法，简单来讲就是将数据自变量进行轻微扰动后再次预测，对比之前结果以计算出预测精度的下降程度。该值越大表示对应因子的重要性越大。

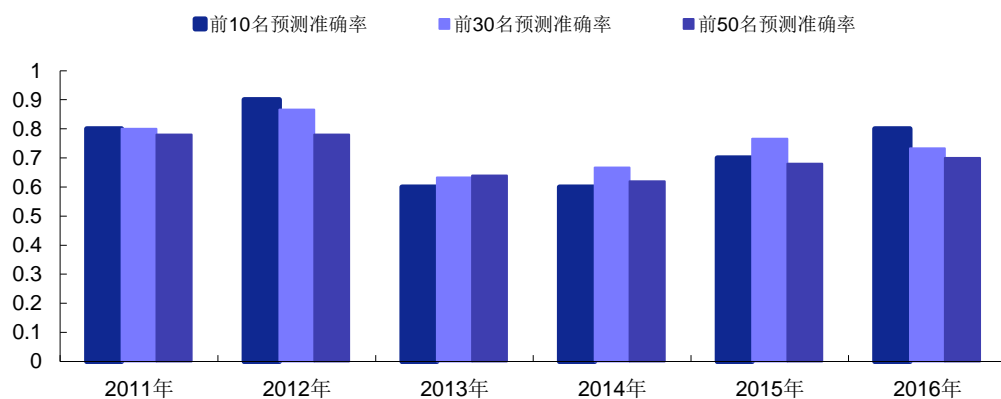


**表 3：变量重要性度量结果**

	MeanDecreaseAccuracy
总股本	48.00649
股价	80.05599
每股资本公积	50.69527
BPS	47.72420
EPS	48.08982
未分配利润	40.07316
上市天数	37.18121

资料来源：wind，安信证券研究中心

通过对模型的三年轮动测试，得到 2011 年至 2016 年的高送转概率预测结果。从预测准确率方面来看，随机森林预测模型表现较优于逻辑回归预测模型，整体平均预测精度保持在 70% 左右，同时不同排名水平的预测准确度表现也比较稳定。

**图 10：不同排名水平下随机森林模型预测准确率**


资料来源：wind，安信证券研究中心

## 4. 高送转填权行情分析

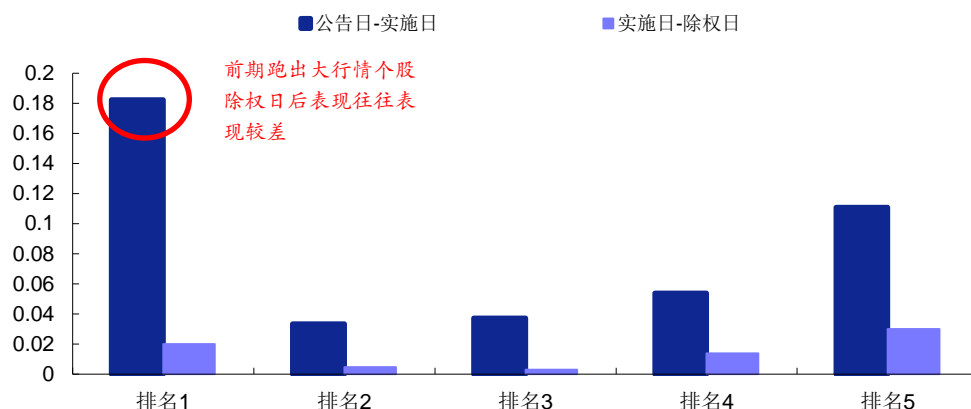
根据上述预测模型可以帮助我们提前布局高送转潜力股，把握“实施”预期行情。后面我们再来简单分析下后续的高送转“填权”行情。通过分析高送转股票除权除息日后的涨跌幅情况，我们发现高送转填权股仅占整体高送转股票数量的 45% 左右。同时，所谓的“填权”行情一般持续 2 周左右，2 周之后的股票收益开始出现回撤。

是否可以发现一些简单因子去进一步筛选高送转股票，以提高捕捉填权股的成功率呢？下面我们主要从前期行情、送转比例以及年报业绩三方面来进行分析：

### 4.1. 前期行情

我们分别统计了从预案公告日到分红实施日以及实施日到除权日这两段时间内的高送转个股收益情况，结果发现一般从预案公告日到分红实施日跑出大行情的个股在除权日之后大概率出现回撤，主要可能因为前期市场已经对于股票的高送转预期充分反应，后期就算除权后价格变低，但上方涨幅空间有限，投资风险较高。此外，从分红实施日到除权除息日的期间涨跌幅与除权后个股走势呈现正相关，市场可能已开始提前布局“填权”行情。我们将高送转个股除权日后两周涨跌幅从高到低分为 5 组，下表统计了这 5 组高送转个股在前期行情的表现情况。

图 11：前期行情 vs 填权行情

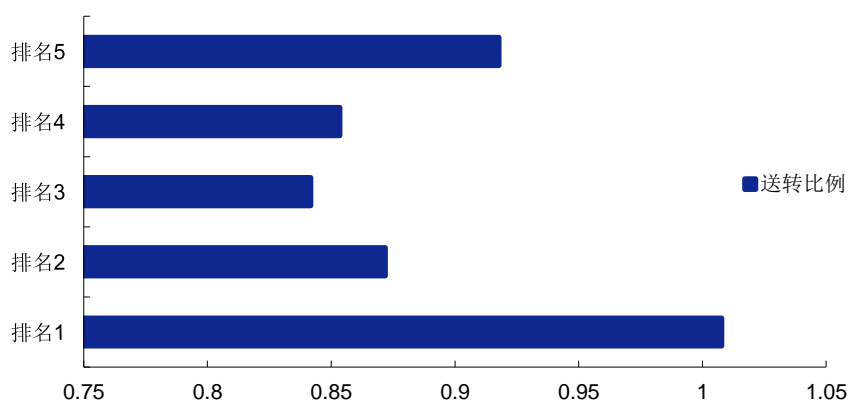


资料来源：wind，安信证券研究中心

## 4.2. 送转比例

我们发现最受市场关注的高送转股票往往也是送转比例最高的一组，行情也多爆发在公告日披露日之后的那段时间。然而送转比例次高的一组在填权时期表现最好。可见在剔除掉前期涨势很猛的高送转个股后重新布局填权行情时，送转比例也一直是大家关注的焦点。毕竟高送转比例背后往往伴随着公司对自身成长性很有信心。

图 12：送转比例 vs 填权行情

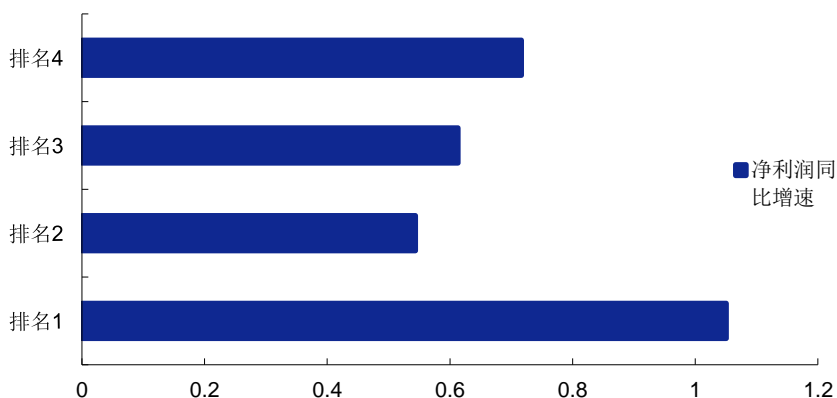


资料来源：wind，安信证券研究中心

## 4.3. 年报业绩

通常公司年报和分红预案在同一天发布，年报业绩也成为市场精选高送转个股的重要指标之一。统计结果发现近几年前期(预案-实施)涨幅较大的高送转个股的年度净利润同比均值在100%以上，高送转有了基本面的支撑才能更加吸引眼球。同时涨幅靠前的填权股也在业绩方面表现较好，净利润同比均值在70%左右。

图 13: 年报业绩 vs 填权行情



资料来源: wind, 安信证券研究中心

#### 4.4. 最新高送转“填权”组合

我们按照下述几个条件整理出最新高送转股票“填权”组合, 仅供参考。

- 剔除掉前期(预案日-实施日)涨幅超过 20% 的个股
- 剔除掉年报净利润同比增速小于 50% 的个股
- 选择送转比例超过 2 的个股

表 4: 最新高送转“填权”组合

代码	名称	预案公告日	送转比例	净利润同比
300304.SZ	云意电气	2017-01-25	2.8	57%
002383.SZ	合众思壮	2017-03-23	3.0	55%
300379.SZ	东方通	2017-03-31	3.0	58%
300317.SZ	珈伟股份	2017-03-24	2.8	129%
002131.SZ	利欧股份	2017-04-01	2.5	145%
300562.SZ	乐心医疗	2017-03-03	2.2	60%
002354.SZ	天神娱乐	2017-03-31	2.0	53%
300008.SZ	天海防务	2017-03-29	2.0	179%
300017.SZ	网宿科技	2017-03-14	2.0	50%
300062.SZ	中能电气	2017-03-31	2.0	250%
300194.SZ	福安药业	2017-03-31	2.0	239%
300319.SZ	麦捷科技	2017-03-28	2.0	99%
603239.SH	浙江仙通	2017-03-31	2.0	71%

资料来源: wind, 安信证券研究中心

#### 5. 风险提示

- 策略基于历史数据进行回测, 由于后期市场风格结构发生变化, 模型存在一定失效风险
- 市场系统性风险
- 流动性风险

## ■ 分析师声明

刘帅路声明，本人具有中国证券业协会授予的证券投资咨询执业资格，勤勉尽责、诚实守信。本人对本报告的内容和观点负责，保证信息来源合法合规、研究方法专业审慎、研究观点独立公正、分析结论具有合理依据，特此声明。

## ■ 本公司具备证券投资咨询业务资格的说明

安信证券股份有限公司（以下简称“本公司”）经中国证券监督管理委员会核准，取得证券投资咨询业务许可。本公司及其投资咨询人员可以为证券投资人或客户提供证券投资分析、预测或者建议等直接或间接的有偿咨询服务。发布证券研究报告，是证券投资咨询业务的一种基本形式，本公司可以对证券及证券相关产品的价值、市场走势或者相关影响因素进行分析，形成证券估值、投资评级等投资分析意见，制作证券研究报告，并向本公司的客户发布。

## ■ 免责声明

本报告仅供安信证券股份有限公司（以下简称“本公司”）的客户使用。本公司不会因为任何机构或个人接收到本报告而视其为本公司的当然客户。

本报告基于已公开的资料或信息撰写，但本公司不保证该等信息及资料的完整性、准确性。本报告所载的信息、资料、建议及推测仅反映本公司于本报告发布当日的判断，本报告中的证券或投资标的价格、价值及投资带来的收入可能会波动。在不同时期，本公司可能撰写并发布与本报告所载资料、建议及推测不一致的报告。本公司不保证本报告所含信息及资料保持在最新状态，本公司将随时补充、更新和修订有关信息及资料，但不保证及时公开发布。同时，本公司有权对本报告所含信息在不发出通知的情形下做出修改，投资者应当自行关注相应的更新或修改。任何有关本报告的摘要或节选都不代表本报告正式完整的观点，一切须以本公司向客户发布的本报告完整版本为准，如有需要，客户可以向本公司投资顾问进一步咨询。

在法律许可的情况下，本公司及所属关联机构可能会持有报告中提到的公司所发行的证券或期权并进行证券或期权交易，也可能为这些公司提供或者争取提供投资银行、财务顾问或者金融产品等相关服务，提请客户充分注意。客户不应将本报告为作出其投资决策的惟一参考因素，亦不应认为本报告可以取代客户自身的投资判断与决策。在任何情况下，本报告中的信息或所表述的意见均不构成对任何人的投资建议，无论是否已经明示或暗示，本报告不能作为道义的、责任的和法律的依据或者凭证。在任何情况下，本公司亦不对任何人因使用本报告中的任何内容所引致的任何损失负任何责任。

本报告版权仅为本公司所有，未经事先书面许可，任何机构和个人不得以任何形式翻版、复制、发表、转发或引用本报告的任何部分。如征得本公司同意进行引用、刊发的，需在允许的范围内使用，并注明出处为“安信证券股份有限公司研究中心”，且不得对本报告进行任何有悖原意的引用、删节和修改。

安信证券股份有限公司对本声明条款具有惟一修改权和最终解释权。

## ■ 销售联系人

上海联系人	葛娇妤	021-35082701	gejy@essence.com.cn
	朱贤	021-35082852	zhuxian@essence.com.cn
	许敏	021-35082953	xumin@essence.com.cn
	孟硕丰	021-35082788	mengsf@essence.com.cn
	李栋	021-35082821	lidong1@essence.com.cn
	侯海霞	021-35082870	houhx@essence.com.cn
北京联系人	潘艳	021-35082957	panyan@essence.com.cn
	原晨	010-83321361	yuanchen@essence.com.cn
	温鹏	010-83321350	wenpeng@essence.com.cn
	田星汉	010-83321362	tianxh@essence.com.cn
	王秋实	010-83321351	wangqs@essence.com.cn
	张莹	010-83321366	zhangying1@essence.com.cn
深圳联系人	李倩	010-83321355	liqian1@essence.com.cn
	周蓉	010-83321367	zhourong@essence.com.cn
	胡珍	0755-82558073	huzhen@essence.com.cn
	范洪群	0755-82558044	fanhq@essence.com.cn
	孟昊琳	0755-82558045	menghl@essence.com.cn

## 安信证券研究中心

### 深圳市

地 址： 深圳市福田区深南大道 2008 号中国凤凰大厦 1 栋 7 层

邮 编： 518026

### 上海市

地 址： 上海市虹口区东大名路638号国投大厦3层

邮 编： 200080

### 北京市

地 址： 北京市西城区阜成门北大街 2 号楼国投金融大厦 15 层

邮 编： 100034