

## 证券市场

# 实证研究中股票投资 收益率估算的若干问题

## Some Considerations on the calculation of Return on Stock Investment in Empirical Studies

陈灯塔 陈浪南

厦门大学财政金融系, 厦门市, 361005

(Chen Max, Chen Langnan, Dept. of finance and int. Banking, Xiamen Univ., Xiamen, 361005)

**摘要:** 讨论了股票除权除息时投资收益率的4种计算方法及其等价性, 分析了错误计算方法的原因, 比较了一个时期内有多次除权除息时投资收益率的3种计算方法的异同, 论述了数据缺失时几种处理方法的区别与联系。

**关键词:** 投资收益; 除权除息; 数据缺失

**Abstract:** In this paper, the details of the four methods of calculating returns on investments when ex-rights and ex-dividend happen is discussed with the explanation of their equivalence and the reason of some of the incorrect methods. The difference of the three methods for computing returns when multiple exs arisen is compared with example. Furthermore, we treat the situation when data missing, reveal the relationship among the usual ways.

**Key words:** Returns on investments, Ex-dividend and ex-right, Data missing

中图分类号: F830.91

文献标识码: A

文章编号: 1002-2252(2000)03-0034-05

### 1 投资收益率的计算是基础性工作

在证券市场的实证研究中, 股票投资收益率的计算是基础性工作。特别是我国的咨询业中, 系统地提供上市公司的股票价格、分红、配股、送股和拆细资料的服务还没有完全开展, 现成的股票收益率资料更是无处购买, 对于我国证券市场的实证研究工作来说, 股票投资收益率的计算又是一件繁重而艰难的任务, 常常历时两三年进行收集和整理, 耗费大量宝贵的时间和精力。而且, 有一点让

人感到痛苦的是, 因为没有统一的资料来源或者数据加工的方法不同, 一位研究者的研究结论很难或者根本不能被另一位研究者所验证, 甚至得出相反的结论, 违背了科学研究的可验证性基本原则。

我国证券业的实践, 极大地促进了市场经济的发展和巩固, 也吸引了大量的学者来研究我国的股票市场。无论是对市场有效性的检验, 还是对市场理性的检验, 股票投资收益的计算总是一件逃避不了的基础工作。但是, 有些文献记载, 在计算股票投

收稿日期: 2000-04-03

作者简介: 陈灯塔(1973-), 男(汉族), 现为金融学博士研究生, 获自动控制工程学士学位、系统工程硕士学位。研究方向: 组合投资理论、金融市场计量经济学、金融工程。

本文为国家社科基金项目(79800010和79870039)和中加教育合作课题成果之一

资收益率的过程中, 进行除息除权处理时, 有些处理不妥当; 其次, 我们在进行数据处理时, 遇上了一个期间内有多次分红、送配股的情况, 手头又缺少可供参考的资料; 再次, 缺失数据的处理也因不同的处理方式而带来一定的差异。

## 2 投资收益率计算方法

股票投资收益率计算: 在大多数情况下, 即在没有分红、送配股和拆细的情况下, 可以很简单地计算股票投资的简单收益率:

$$r_t = \frac{p_t - p_{t-1}}{p_{t-1}} = \frac{p_t}{p_{t-1}} - 1 \quad (1)$$

其中,  $p_t$  为  $t$  时期的股票价格, 一般取收盘价;  $p_{t-1}$  为  $t-1$  时期的股票价格, 一般取收盘价;  $r_t$  为  $t$  时期股票的收益率(已经实现的)。

注: 一般的收益率为  $r_t = \frac{p_{t+1} - p_t}{p_t}$ , 式(1)的定义是为了和除权除息日对应。

股票分割(拆细):

$$r_t = \frac{np_t - p_{t-1}}{p_{t-1}} = n \frac{p_t}{p_{t-1}} - 1 \quad (2)$$

其中,  $n$  为分割比率, 即原来一股拆细为  $n$  股。

同时有送股、配股和现金股息:

$$r_t = \frac{(1 + n_s + n_p)p_t + c - n_p p}{p_{t-1}} - 1 \quad (3)$$

其中,  $n_s$  为送股比率, 持有原来一股送多少红股;  $n_p$  为配股比率, 持有原来一股配多少股;  $p$  为配股价格;  $c$  为现金股息, 持有原来一股的现金股息(是除权除息前的); 其他符号与前面同。

需要说明, 送红股等价于股票分割, 即有:

$$n = n_s + 1$$

式(3)等价于下面 3 种计算方法:

方法一: 简单收益率调整法

$$r_t = (1 + n_s + n_p)(1 + r) + \frac{c - n_p p}{p_{t-1}} - 1 \quad (4)$$

其中,  $r = \frac{p_t}{p_{t-1}} - 1$  是我们没有考虑除权除息下的简单收益率, 方法一可以看成是对简单收益率进行调整。

方法二: 现金股息调整法

$$r_t = \frac{p_t - p_{t-1} + d_t}{p_{t-1}} = \frac{p_t + d_t}{p_{t-1}} - 1 \quad (5)$$

其中,  $d_t = c + p_t(n_s + n_p) - n_p p$ , 即将分红和送配股信息看成现金股息, 大部分的实证研究中使用现金股息调整法进行计算。

方法三: 股价调整法

在第  $m$  个交易期间, 发生了分红送配股信息, 需要对  $m$  时期以前的股价作调整, 以剔除非交易因素带来的价格变化, 调整公式如下:

$$p_t^1 = P \cdot p_t \quad t = 1, 2, \dots, m - 1$$

其中

$$P = \frac{P_m}{(1 + n_s + n_p)p_m + c - n_p p} \quad (6)$$

除权除息期间的收益率计算公式为:

$$r_t = r_m = \frac{P_m}{P_{m-1}} - 1 \quad (7)$$

股价调整法虽然应用不多, 但如果能得到调整好的股价资料, 公式(7)计算股票投资收益率就再简单不过了。

上述式(3)、式(5)和式(7)的计算方法, 如果非除权除息日的分红现金股息、配股比率和送股比率都取零, 则 3 种方法中除权除息日与非除权除息日投资收益率的计算具有一致性。但是, 我们也发现一些错误的计算方法。

错误方法一:

假设  $p_{t-1}$  的除息除权理论价为  $p_1$ , 则:

$$p_{t-1} = p_1(1 + n_s + n_p) + c - n_p p$$

即

$$p_1 = \frac{p_{t-1} + n_p p - c}{1 + n_s + n_p} \quad (8)$$

$$r_t = \frac{p_t - p_1}{p_1} = \frac{p_t}{p_1} - 1 = \frac{(1 + n_s + n_p)p_t}{p_{t-1} + n_p p - c} - 1 \quad (9)$$

用(9)式计算, 分子  $p_t - p_1$  是投资增值部分, (9)式的错误在于分母  $p_1$  是参考价, 不是“本金”, 除权除息后的每股因为除权除息我们还投入了

$$\frac{c - n_p p}{1 + n_s + n_p} \text{ 资金, 正确的“本金”是 } p_1 + \frac{c - n_p p}{1 + n_s + n_p} = \frac{p_{t-1}}{1 + n_s + n_p}$$

错误方法二:

$$r_t = \frac{(1 + n_s + n_p)p_t + c - n_p p}{p_{t-1} + n_p p} - 1 \quad (10)$$

显然, 式(10)的错误是把配股缴纳的金额记入“本金”了。

例 1 1995 年 7 月 10 日是大众科创(股票代码: 600635)的除息除权日, 其中: 每 10 股送 2 元现金股息, 每 10 股送 1 股红股, 每 10 股配股 3 股, 配股价为 5.20 元。当日收盘价为 16.17 元, 上一交易日的收盘价为 20.88 元(资料来源: 康熙胜券系统)。

显然,  $c = 0.2$ ,  $n_s = 0.1$ ,  $n_p = 0.3$ ,  $p = 5.2$ ,  $p_t = 16.17$ ,  $p_{t-1} = 20.88$ , 计算结果如表 1 所示, 虽然绝对误差不是很大, 但相对误差都在 6% 以上。

表 1 正确与错误的收益率比较

方法	收益率	
正确	0.0190613	
错误	式(9)	0.0178957
	式(10)	0.0177362

### 3 一个期间内多次除权除息的处理

如果我们考虑的时期为一个月、一个季度或者更长,计算月收益率或者季度收益率时就有可能碰上一个期间内有多次除权除息的情况。一个现实的例子是:山东焦化,股票代码600740,1998年5月18日每10股现金股息分红2.5元,当月的28日,又实施了每10股配股3股,配股价为4元,类似的情况1997年在代码为600810、600873和600887等的股票中也发生过(资料来源:康熙胜券系统)。这种情况如何处理,我国现有实证研究的书籍和论文中还没有该方面材料公开以供参考,下面将期间选取为一月为例,尝试其解决方法。

假设由于某种原因,比如说日交易数据购买太昂贵,我们只拥有股票的月交易数据,即月的交易基本数据的时间、开盘、最高、最低、收盘、成交量和成交金额。另外,我们还有分红、送配股的数据,有一点要说明的是,每个交易日最多可以有一次的除权除息,现实的证券市场是这样的。

本文的第二部分中,已经处理了一个时期仅有一次除权除息的情况。事实上,一个时期的多次除权除息可以合并为一次综合的除权除息,然后用第二部分的方法进行计算。下面是合并迭代算法。

假设一个时期里,有 $m$ 次( $m \geq 2$ )的除权除息,记合并后的总股本扩张倍数为 $sh$ ,总分红金额为 $sfh$ ,总配股金额为 $spj$ 。

迭代算法:

初始值:  $sh = 1, sfh = 0, spj = 0$

迭代公式:

$$\begin{cases} sfh(j+1) = sfh(j) + c(j) \cdot sh(j) \\ spj(j+1) = spj(j) + n_p(j) \cdot p(j) \cdot sh(j) \\ sh(j+1) = (1 + n_s(j) + j_p(j)) \cdot sh(j) \end{cases} \quad (11)$$

最后,取总配股比率:  $n_p = 1$ ; 总送股比率:  $n_s = sh - 2$  即可。

迭代过程(11)将分红在期间内进行汇总,同时也将期间的所有配股应缴纳的金额进行汇总,对总配股比率和总送股比率的处理,不仅正确地计算了总股本的扩张倍数,而且合并后的综合除权除息数据可以用式(3)、式(4)、式(5)或者式(7)进行计算。

有一点必须说明的是,式(11)中期间内的下一次(第 $j+1$ 次)除息除权的送股和配股比率必须以

当时的股本作为基数,即第 $j$ 次除权除息后的总股本为基数,否则,应先进行转化。

另外,如果我们能得到日交易数据,并用式(6)进行价格调整后,提取出月交易数据,只要简单计算股票投资收益率即可。但是,价格调整法不仅要求我们有该股票完整的日交易数据以及分红、送配股库,而且每一次除权除息的调整都要上溯到最初的日交易价格,考虑的时间越长,计算量越大;除权除息日越靠后,计算量越大;除权除息次数越多,调整次数越多,带来的数值计算误差越大。

如果我们有正确的股票投资的日收益率,我们还可以用几何累积法计算股票投资的月收益率,即:

$$r_t = \prod_{j=1}^m (1 + r_{jt}) - 1 \quad (12)$$

其中,  $r_{jt}$  为第 $t$ 月里第 $j$ 个交易日的收益率,第 $t$ 月里共有 $m$ 个交易日。

一个月一般有20个左右的交易日,用式(12)计算,数值计算误差显然不可避免,特别是考虑期间为季度甚至半年等。如果没法得到日收益率,几何累积法就无法计算了。

自此,当期间内有多次分红送配股需要进行除息除权的考虑时,我们有了3种方法:除权除息合并法、价格调整法和几何累积收益法。从理论上讲,价格调整法和几何累积法是一致的,但数值计算时会产生微小的差别;价格调整法和几何累积法与除权除息合并法的结果显然不同,相对来说结果偏小,因为除权除息合并法将只是简单地将期间的分红和应缴纳的配股款进行汇总,忽略其短期的投资效果。由于一般计算时,期间内发生的资本运动认为是发生在期间末,比如当银行活期存款考虑的期间(时间单位)是天,一天内的多次存取款显然不会因为时刻的不同而带来利息上的好处或者损失,而是被简单的汇总,当成同一时刻发生。因此除权除息合并法具有一定的现实基础,而且当我们没有日交易数据或者日收益率时,3种方法中只能选择除权除息合并法。

例2 1997年6月,伊利股份(股票代码600887)有两次除权除息,如表2所示,伊利股份的日交易收盘价如表3所示,该只股票在当月有20个交易日数据(资料来源:康熙胜券系统)。

表2 分红的配股

除权日期	分红数量	送股比率	配股比率	配股价	股价调整系统
1997/06/09	1.25/10	0	0	0	0.9940996
1997/06/27	0	0	3.00/10	6.80	0.8320327
综合	0.125	-0.7	1	2.04	

表 3 股票数据

时间	收盘价	股价调整	简单收益率	真实收益率
97/05/30	<b>21.50</b>	17.78315		
97/06/02	22.33	18.46966	0.0386047	0.0386047
97/06/03	21.70	17.94858	-0.0282132	-0.0282132
97/06/04	21.70	17.94858	0.0000000	0.0000000
97.06/05	22.00	18.19671	0.0138249	0.0138249
97/06/06	21.10	7.45230	-0.0409091	-0.0409091
97/06/09	21.06	7.52261	-0.0018957	0.0040284
97/06/10	20.40	16.97347	-0.0313390	-0.0313390
97/06/11	20.70	17.22308	0.0147059	0.0147059
97/06/12	20.75	7.26468	0.0024155	0.0024155
97/06/13	21.50	17.88870	0.0361446	0.0361446
97/06/16	22.72	18.90378	0.0567442	0.0567442
97/06/17	22.59	18.79562	-0.0057218	-0.0057218
97/06/18	22.30	18.55433	-0.0128375	-0.0128375
97/06/19	21.98	18.28808	-0.0143498	-0.0143498
97/06/20	23.20	19.30316	0.0555050	0.0555050
97/06/23	23.65	19.67757	0.0193966	0.0193966
97/06/24	24.20	0.13519	0.0232558	0.0232558
97/06/25	24.19	0.12687	-0.0004132	-0.0004132
97/06/26	23.65	19.67757	-0.0223233	-0.0223233
97/06/27	<b>20.79</b>	20.79	-0.1209302	0.0565328

表 4 收益率比较

方法	收益率
除权除息合并法	0.16800
股价调整法	0.16908
几何累积法	0.16908

通过除权除息合并算法计算得到的综合分红、送配股数据列在表 2 的最后一行中, 根据表 2 和表 3, 用式(6)计算的股价调整系数列在表 2 的最后一列里。虽然在表 2 和表 3 中, 中间数据和结果仅表示为 7 位有效数字, 但在数据处理过程中, 我们使用双精度浮点数进行计算和存储, 以减少计算误差。

分别用 3 种方法计算得到的收益率如表 4, 正如

我们前面的分析, 股价调整法与几何累积法有相同的收益率, 除权除息合并法的计算结果则相对较小。

#### 4 缺失数据的处理

由于某种原因, 某只股票在正常交易日里停牌而没有交易数据, 这就造成数据的缺失。这里说的数据缺失是指其他股票正常交易, 而该股票因某种原因, 如公告或者发生重大事件而造成停牌。数据缺失将造成如下的困难: 在分析若干种证券的投资组合时, 样本数不一致的情况下, 如何计算方差协方差矩阵? 如代码为 600849 的股票, 竟然暂停交易达 3 个多月, 停了 103 天, 从 1998 年 5 月 29 日收盘后, 一直到 9 月 9 日才复开(资料来源: 康熙胜券系统), 停牌两个月以上的也有好几只, 如代码为 600626、600667 和 600689 的股票, 分别停了 63 天、70 天和

84天。对这样长时间停牌所带来的数据缺失如何妥善处理?

当数据缺失仅一个样本点时,常见有如下几种方法。假设股票A在考察的时间里共有 $n$ 个样本点,则其他正常的股票有 $n+1$ 个样本点,则

舍弃法:不考虑该缺失的样本点,样本数为 $n$ ,则股票A的期望收益率为:

$$r_A^{(1)} = \frac{1}{n} \sum_{i=1}^n r_i$$

方差为:

$$v_A^{(1)} = \frac{1}{n-1} \sum_{i=1}^n (r_i - r_A^{(1)})^2$$

补零法:就是认为缺失日的收益率为零。则样本数为 $n+1$ ,其期望收益率和方差分别为: $r_A^{(2)} = \frac{1}{n+1} \sum_{i=1}^n r_i = \frac{n}{n+1} r_A^{(1)}$ 和 $v_A^{(2)} = \frac{1}{n} [ \sum_{i=1}^n (r_i - r_A^{(2)})^2 + (r_A^{(2)})^2 ]$ ,下面比较舍弃法与补零法的期望收益与方差:

$$r_A^{(1)} - r_A^{(2)} = \frac{1}{n+1} r_A^{(1)} = \frac{1}{n(n+1)} \sum_{i=1}^n r_i \quad (13)$$

$$v_A^{(1)} - v_A^{(2)} = v_A^{(1)} - \frac{1}{n} [ \sum_{i=1}^n (r_i - r_A^{(1)} + r_A^{(1)} - r_A^{(2)})^2 + (r_A^{(2)})^2 ] = \frac{1}{n} v_A^{(1)} - \frac{1}{n+1} (r_A^{(1)})^2 \quad (14)$$

从式(13)可以看出,虽然式(13)右边的符号正负未定,但是补零法得到的期望收益率的绝对值比舍弃法的在非零时来得略小(两个期望值同号或者同为零);式(14)清楚地表明,补零法得到的方差与舍弃法的大小关系不能确定,依股票收益率分布的具体形式而定。必须说明的是当样本数 $n$ 较大时,两种方法的差别微乎其微。

补期望收益率法:用舍弃法的期望收益率作替代缺失的数据,样本数补充为 $n+1$ ,其期望收益率与舍弃法同: $r_A^{(3)} = r_A^{(1)}$ ,而方差略小: $v_A^{(3)} = \frac{n-1}{n} v_A^{(1)}$ 。

其他修补法:其它方法可以用指数收益或者收益率理论分布的随机值等代替。

综上所述,对缺失数据的处理可以归纳为舍弃和修补两大类。虽然当样本量大的情况下,几种方法对期望收益率和方差的影响甚微,但舍弃法使用的样本数量与其他股票不同,不能计算相关系数和协方差等,如果将其他股票该交易日的的数据也舍弃,以达到样本数量相一致,则其他股票都将损失一个有效样本点,可真是拣了芝麻,丢了西瓜;修补类中,缺失的样本被补齐,没有上述问题。此外,补零法比较符合实际,因为股票投资中股票停牌便没有收益。

以上讨论了缺失数据只有一个的情况,当缺失的数据比较少,上述的方法还是可以应用的,误差不会

太大。而当股票停牌的时间过长,缺失数据的处理按上述方法就行不通,该问题有待进一步研究。当然,该问题是可以回避的,只要我们把该只股票从样本中删除就是了。显然,实证研究时股票样本的选择也存在着很多现实的问题。

## 5 期望和建议

90年代以来,我国证券市场崛起并迅速扩大,证券市场在国民经济成长中的支持和辅助作用开始形成,并且构成了影响社会经济生活的重要因素。脚踏实地地对我国证券市场开展实证研究,研究市场对现代金融投资理论的吸收、消化和运用,完善和发展处在初级阶段的我国证券市场,借鉴现代投资理论,对其理论框架、模型结构及参数设定等进行必要的修正,使其适应我国证券市场的实际情况,规划和建立我国证券市场的监管体系,投资者结构、市场规模以及运行机制等,在理论和实践上都具有非常重要的现实意义,使证券市场在我国的经济发展中发挥更大的促进作用。

股票投资收益率的计算,是一件艰巨而繁重的任务,在实证研究中是基础性工作。因此,呼吁尽快地、系统地建立我国证券市场的交易数据库,建立以分钟为单位的交易数据库,为我国证券市场微观市场结构的研究提供宝贵的资料;建立日交易数据库和分红、送配股数据库,以方便开展市场有效性和理性等的研究;建立上市公司的财务数据库,为研究我国证券市场的市场行为服务;建立证券市场深加工数据的数据库,作为公司决策和国家经济政策的参考。

此外,由我国的信息咨询业提供以上数据库的服务,或者作为电子图书馆的一项服务在高校和科研机构中开展。这种专业化的分工,首先,将促进作为信息咨询业业务的发展,丰富其服务内容,更好地服务于社会经济的发展;其次,使实证研究学者能减少在获取原始数据和已加工数据上的精力,能专心所关心的问题;最后,使数据来源有据可查,一个学者在证券市场上的实证研究结果,能让其他研究者验证其真伪,符合科学研究的可验证原则。

## 参考文献

- [1] 陈共,周升业,吴晓求. 证券投资分析(第二版)[M]. 北京:中国人民大学出版社,1997
- [2] 刘波. 中国证券市场实证分析[M]. 上海:学林出版社,1997
- [3] 陆金海. 公开信息与股票市场行为实证研究[D]. 博士学位论文,1999
- [4] 汤羨祥,许耀钧. 现代证券理论[M]. 上海:中国纺织大学出版社,1996