

对铁路统计分析的因素分析法的意见

甘肃职工财经学院铁路分部 郭自诚

摘 要

本化提出用和差式的综合指数体系来评价铁路运输经营指标及经济效果,并论证它能解决用乘积式综合指数法的不足之处。

北方交大丁振京、陈齐洪编著的《铁路统计学》,在旧有的铁路统计教材基础上,广征博采,集众家之长,并加入许多经济学的理论与方法,内容丰富,文笔简炼,颇有特色。特别是“铁路统计分析”一章,总结了铁路统计分析的一些经验和方法,且有精辟的见解。在“计划完成情况的分析”一节中,关于“分解各因素,说明指标计划完成的原因”部分,笔者有些不同看法,特提出来和丁、陈两位老师商榷。

教科书P226关于货运量完成情况的分析,举了一个范例:

设某路局某年货物发送吨数、装车数及静载重资料如下:

	货物发送吨数 (万吨)	装 车 数 (车)	货车平均载重 (吨)
计 划	3,255	775,000	42.0
实 际	3,408	800,000	42.6
计划完成%	104.69%	103.22%	101.43%

教材书用的是乘积式综合指数体系分析法。

$$\frac{\sum q_1 m_1}{\sum q_0 m_0} = \frac{\sum q_1 m_0}{\sum q_0 m_0} \times \frac{\sum q_1 m_1}{\sum q_1 m_0}$$

$$\sum q_1 m_1 - \sum q_0 m_0 = (\sum q_1 m_0 - \sum q_0 m_0) + (\sum q_1 m_1 - \sum q_1 m_0)$$

研究数量指标因素变动时,将同度量因素固定在基期。装车数是数量指标,其指数为

$$\frac{\sum q_1 m_0}{\sum q_0 m_0} = \frac{800,000 \times 42}{775,000 \times 42} \times 100\% = 103.22\%$$

本文于1985年2月18日收到

绝对数增加: $(800,000 - 775,000) \times 42\% = 105$ 万吨

研究质量指标因素影响时,将同度量因素固定在报告期。货车平均静载重为质量指标,

其指数为 $\frac{\sum q_1 m_1}{\sum q_1 m_0} = \frac{42.6 \times 800,000}{42 \times 800,000} \times 100\% = 101.43\%$

绝对数增加: $(42.6 - 42) \times 800,000 = 48$ 万吨

货物发送吨数计划完成百分数为

$$\frac{\sum q_1 m_1}{\sum q_0 m_0} = \frac{3408}{3255} \times 100\% = 104.69\%$$

绝对数增加: $3408 - 3255 = 153$ 万吨

从相对数看,货物发送吨数超额完成计划4.69%,由于装车数超额完成计划3.22%,由于货车平均静载重超额完成计划1.43%,故能超额完成。 $103.22\% \times 101.43\% = 104.69\%$

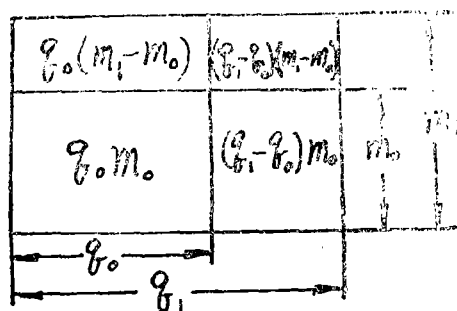
从绝对数看,货物发送吨数超额153万吨,其中由于装车数超额完成计划,增加了105万吨,由于货车平均静载重超额完成计划,增加了48万吨。 $105 + 48 = 153$ 万吨

用乘积式综合指数法,分析多因素的影响,这种方法是过去常用的。但这种方法有三个问题。

1. 区别某些指标为质量指标或数量指标比较困难。如判断不准确,同度量因素固定的时期也就不同,计算的结果也随着有所不同。特别是多因素分析,各因素置换次序不同,计算的结果迥异。

2. 从相对数来看,由于乘积式综合指数体系指数间是乘积关系,不能具体说明在货物发送吨数超额完成计划4.69%中,由于装车数超额完成计划,增加了百分之几,由于货车平均静载重超额完成计划,增加了百分之几。

3. 受两因素影响的指标增长的数学模型为矩形面积增长图形,如下图:



$$\begin{aligned} & [q_0 + (q_1 - q_0)] \cdot [m_0 + (m_1 - m_0)] \\ &= q_0 m_0 + q_0 (m_1 - m_0) + (q_1 - q_0) m_0 + \\ & \quad (q_1 - q_0) (m_1 - m_0) \end{aligned}$$

从上图可以看出 $q_1 m_1$ 可分四个部分,除 $q_0 m_0$ 外,还有 $q_0 (m_1 - m_0)$,为货车平均静载重变动影响货物发送吨数的增长额; $(q_1 - q_0) m_0$ 为装车数的变动影响货物发送吨数的增长额; $(q_1 - q_0) (m_1 - m_0)$ 为两因素共同变动影响货物发送吨数的增长额。

按乘积式综合指数体系进行分析,研究质量指标因素的变动,将数量指标固定在报告期, $(m_1 - m_0) q_1 = (m_1 - m_0) [q_0 + (q_1 - q_0)] = q_0 (m_1 - m_0) + (m_1 - m_0) (q_1 - q_0)$,实际上是将两因素共同变动的影响也算作质量指标变动的影响了,这样就夸大了质量指标变动的影响,未能真实地反映两因素变动的影响情况。

当前我国统计学界也注意到以上三个问题,为了弥补其不足之处,在《经济统计电视讲座》中,人民大学的徐前教授提出了共变指数,其形式为:

$$\left(\frac{\sum q_1 m_1}{\sum q_0 m_0} \div \frac{\sum q_0 m_1}{\sum q_0 m_0} \div \frac{\sum q_1 m_0}{\sum q_0 m_0} \right) \text{保持了乘积式综合指数体系形式。}$$

$$\frac{\sum q_1 m_1}{\sum q_0 m_0} = \frac{\sum q_0 m_1}{\sum q_0 m_0} \times \frac{\sum q_1 m_0}{\sum q_0 m_0} \times \left(\frac{\sum q_1 m_1}{\sum q_0 m_0} \div \frac{\sum q_0 m_1}{\sum q_0 m_0} \div \frac{\sum q_1 m_0}{\sum q_0 m_0} \right)$$

总变动指数 = 质量指标指数 × 数量指标指数 × 两因素共变指数。

这样将质量的指标因素变动影响与两因素共变的影响区分开来,也能从相对数、绝对数来分析三类因素变动的影响,但是共变因素指数形式太复杂,并且三个以上的多因素变动分析,就难于进行了。

指数体系是指若干个指数由于数量上的联系而形成的一个整体。既然乘积式综合指数体系进行因素分析有一定缺点,能不能用其它方法,例如和差法将指数联系起来,建立新的指数体系呢?笔者是从下式开始探索的。

$$\begin{aligned} \frac{\sum q_1 m_1}{\sum q_0 m_0} &= \frac{\sum [q_0 + (q_1 - q_0)] [m_0 + (m_1 - m_0)]}{\sum q_0 m_0} \\ &= \frac{\sum q_0 m_0 + \sum q_0 (m_1 - m_0) + \sum (q_1 - q_0) m_0 + \sum (q_1 - q_0) (m_1 - m_0)}{\sum q_0 m_0} \\ &= \frac{\sum q_0 m_0}{\sum q_0 m_0} + \frac{\sum q_0 (m_1 - m_0)}{\sum q_0 m_0} + \frac{\sum (q_1 - q_0) m_0}{\sum q_0 m_0} + \frac{\sum (q_1 - q_0) (m_1 - m_0)}{\sum q_0 m_0} \\ \left(\frac{\sum q_1 m_1}{\sum q_0 m_0} - 1 \right) &= \left(\frac{\sum q_0 m_1}{\sum q_0 m_0} - 1 \right) + \left(\frac{\sum q_1 m_0}{\sum q_0 m_0} - 1 \right) + \frac{\sum (q_1 - q_0) (m_1 - m_0)}{\sum q_0 m_0} \end{aligned}$$

这是用和差法将几个指数联系成一个整体,构成和差式的综合指数体系。从经济意义上来说,上式表示:

总变动指数增长率 = 质量指标因素影响增长率 + 数量指标因素影响增长率 + 共变因素影响增长率。

由于上式等号两边分母相同,通分后,可得下面等式:(左边分子等于右边分子):

$$\sum q_1 m_1 - \sum q_0 m_0 = (\sum q_0 m_1 - \sum q_0 m_0) + (\sum q_1 m_0 - \sum q_0 m_0) + \sum (q_1 - q_0) (m_1 - m_0)$$

其经济意义仍为:总变动指数增长额 = 质量指标因素影响增长额 + 数量指标因素影响增长额 + 共变因素影响增长额。

用这种和差式综合指数体系分析货物发送吨数各因素影响计划完成情况的方法如下。

装车数(车)		静载重(吨)		货物发送吨数(万吨)				共变因素		
计划 q_0	实际 q_1	计划 m_0	实际 m_1	计划 $q_0 m_0$	实际 $q_1 m_1$	$q_0 m_1$	$q_1 m_0$	$q_1 - q_0$	$m_1 - m_0$	$(q_1 - q_0)(m_1 - m_0)$
775,000	800,000	42.0	42.6	3255	3408	3301.5	3360	25,000	0.6	1.5(万吨)

货物发送吨数超额完成计划 $\left(\frac{3408}{3255} - 1\right) = 4.69\%$ ，其中装车量超额完成计划影响

$$\left(\frac{3360}{3255} - 1\right) \times 100\% = 3.22\%$$

$$\text{静载重超额完成计划影响} \left(\frac{3301.5}{3255} - 1\right) \times 100\% = 1.43\%$$

$$\text{两因素共变影响} \frac{1.5}{3255} \times 100\% = 0.04\%$$

$$3.22\% + 1.43\% + 0.04\% = 4.69\%$$

从绝对数来看，货物发送吨数增加绝对数为 $3408 - 3255 = 153$ 万吨

其中，由于装车量超额完成计划，增加 $(3360 - 3255) = 105$ 万吨，由于静载重超额完成计划，增加了 $3301.5 - 3255 = 46.5$ 万吨，由于两因素共变影响，增加了 1.5 万吨。

$$105 + 46.5 + 1.5 = 153 \text{ (万吨)}$$

三因素变动的分析，用乘积式综合指数体系一般形式为：

$$\frac{\sum q_1 m_1 p_1}{\sum q_0 m_0 p_0} = \frac{\sum q_1 m_0 p_0}{\sum q_0 m_0 p_0} \times \frac{\sum q_1 m_1 p_0}{\sum q_1 m_0 p_0} \times \frac{\sum q_1 m_1 p_1}{\sum q_1 m_1 p_0}$$

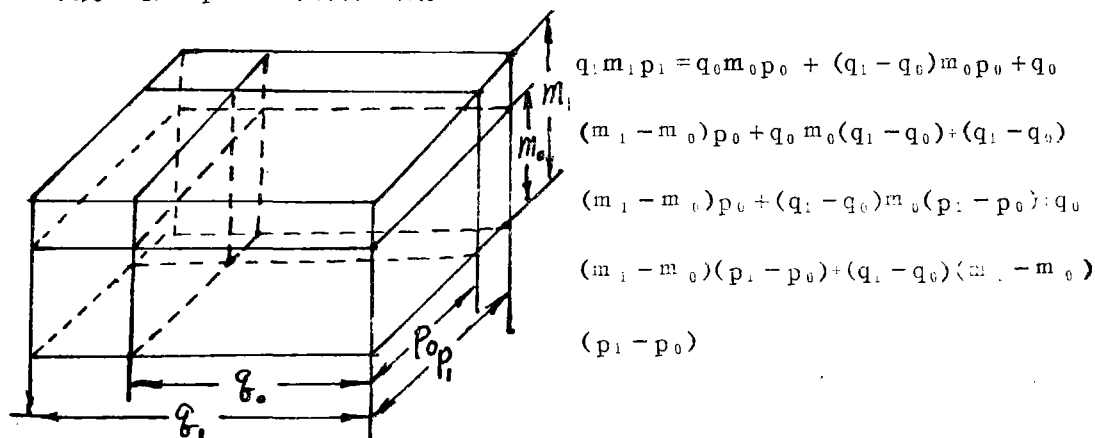
这种形式虽然整齐、好看，但问题不少。

首先，判断质量指标、数量指标，排列 q 、 m 、 p 的顺序就相当困难，三因素顺序颠倒，计算出来的结果相差很大。

其次，不能用和的形式说明三因素影响的增长率。

第三，不符合三因素变动影响的数学模型。三因素变动情况，其模型如长方体体积的变动，如下图：

长方体 $q_1 m_1 p_1$ 由 8 个部分组成。



即长方体 $q_1 m_1 p_1$ 可以分为 8 部分，其中一个因素单纯变动的 3 部分，两个因素共同变动的 3 部分，三个因素共同变动的一部分，以及原有的长方体 $q_0 m_0 p_0$ 。

按照两因素变动的和差式综合指数体系的推导方法，可以推出三因素变动的和差式综合指数体系。得出：

$$\begin{aligned} \left(\frac{\sum q_1 m_1 p_1}{\sum q_0 m_0 p_0} - 1 \right) &= \left(\frac{\sum q_1 m_0 p_0}{\sum q_0 m_0 p_0} - 1 \right) + \left(\frac{\sum q_0 m_1 p_0}{\sum q_0 m_0 p_0} - 1 \right) + \left(\frac{\sum q_0 m_0 p_1}{\sum q_0 m_0 p_0} - 1 \right) \\ &+ \frac{\sum q_0 (m_1 - m_0) (p_1 - p_0)}{\sum q_0 m_0 p_0} + \frac{\sum (q_1 - q_0) (m_1 - m_0) p_0}{\sum q_0 m_0 p_0} + \frac{\sum (q_1 - q_0) m_0 (p_1 - p_0)}{\sum q_0 m_0 p_0} \\ &+ \frac{\sum (q_1 - q_0) (m_1 - m_0) (p_1 - p_0)}{\sum q_0 m_0 p_0} \end{aligned}$$

其经济意义为：总变动指数的增长率等于各类因素变动影响的增长率之和。

由于等式两边分母相同，通分后，左边的分子必等于右边的分子。这个等式经济意义为：总变动指数增长额等于各类因素变动影响的增长额之和。

如不需要详细分析两个或三个共变因素影响的增长率和增长额，可用总变动指数的增长率（或增长额）减去3个一因素单纯变动影响的增长率（或增长额），就可求得所有共变因素影响的增长率（或增长额）。

例、某局某季货物运送吨数、货物平均运程、运行吨公里与计费吨公里换算系数资料如下：

	货物运送吨数(万吨)	货物平均运程(公里)	运行吨公里与计费吨公里换算系数
计划	3255	146	1.02
实际	3408	150	1.04
计划完成%	104.69%	102.74%	101.96%

试分析三因素对运行吨公里计划完成情况的影响情况。

解：该局运行吨公里计划完成% = $\frac{3408 \times 150 \times 1.04}{3255 \times 146 \times 1.02} \times 100\% = 109.68\%$ ，即实际比计划超额9.68%。

绝对数增加：531648 - 484734.6 = 46913.4（万吨公里）

为了便于分析三因素对运行吨公里计划完成情况的影响，分别以q、s、k代表三因素。

$$\frac{(q_1 - q_0) s_0 k_0}{q_0 s_0 k_0} = \frac{22784.76}{484734.6} \times 100\% = 4.701\%$$

由于货物运送吨数超额完成计划，影响运行吨公里超额4.701%，绝对数增加22784.76（万吨公里）

$$\frac{q_0 (s_1 - s_0) k_0}{q_0 s_0 k_0} = \frac{13280.4}{484734.6} \times 100\% = 2.74\%$$

由于实际货物平均运程超过计划，影响运行吨公里超额2.74%，绝对数增加13280.4（万吨公里）

$$\frac{q_0 s_0 (k_1 - k_0)}{q_0 s_0 k_0} = \frac{9504.6}{484734.6} \times 100\% = 1.961\%$$

由于运行吨公里与计费吨公里的换算系数,实际超过计划,影响运行吨公里超额1.961%影响绝对数增加9504.6(万吨公里)

$$\frac{(q_1 - q_0)(s_1 - s_0)k_0}{q_1 s_1 k_0} = \frac{624.24}{484734.6} \times 100\% = 0.129\%$$

由于运送货物吨数、货物平均运程共同变动,影响运行吨公里超额0.129%,影响绝对数增加624.24(万吨公里)

$$\frac{(q_1 - q_0)s_0(k_1 - k_0)}{q_0 s_0 k_0} = \frac{446.76}{484734.6} \times 100\% = 0.092\%$$

由于运送货物吨数、换算系数共同变动,影响运行吨公里超额0.092%,影响绝对数增加446.76(万吨公里)

$$\frac{q_0(s_1 - s_0)(k_1 - k_0)}{q_0 s_0 k_0} = \frac{260.4}{484734.6} \times 100\% = 0.054\%$$

由于货物平均运程、换算系数共同变动,影响运行吨公里超额0.054%,影响绝对数增加260.4(万吨公里)

$$\frac{(q_1 - q_0)(s_1 - s_0)(k_1 - k_0)}{q_0 s_0 k_0} = \frac{12.24}{484734.6} \times 100\% = 0.003\%$$

由于三因素的共同变动,影响运行吨公里超额0.003%,影响绝对数增加12.24(万吨公里)

$$4.701\% + 2.74\% + 1.961\% + 0.129\% + 0.092\% + 0.054\% + 0.003\% = 9.68\%$$

$22,784.76 + 13,280.4 + 9,504.6 + 624.24 + 446.76 + 260.4 + 12.24 = 46913.4$ (万吨公里)

和差式综合指数体系也可分析四、五、…… n 个因素变动的影响。其指数体系推导下如式:

$$[q_0 + (q_1 - q_0)] \cdot [q_0 + (q_1 - q_0)] \cdot [R_0 + (R_1 - R_0)] \cdots \\ = q_0 q_0 R_0 + \cdots (q_1 - q_0) q_0 R_0 \cdots + \cdots$$

每个中括号内有两项,其乘积项数必为 2^n 项。

故两因素变动可分为 $2^2 = 4$ 部分。

三因素变动可分为 $2^3 = 8$ 部分。

四因素变动可分为 $2^4 = 16$ 部分。

……

n 因素变动可分为 2^n 部分。

除去原来 $q_0 q_0 R_0 \cdots$ 基期水平部分后,

两因素变动要受 $2^2 - 1$ 类指数变动影响,

三因素变动要受 $2^3 - 1$ 类指数变动影响,

四因素变动要受 $2^4 - 1$ 类指数变动影响,

.....

n 因素变动要受 $2^n - 1$ 类指数变动影响。

在 $[q_0 + (q_1 - q_0)] [p_0 + (p_1 - p_0)] [R_0 + (R_1 - R_0)] \dots$ 式中,
 $q_0, p_0, R_0 \dots$ 为基期水平,是不变的。而 $(q_1 - q_0), (p_1 - p_0), (R_1 - R_0) \dots$
 是变动的因素,根据排列组合原理,选择一个变动的因素,其余选不变的因素的方法,必然
 有 C_n^1 种。选择两个变动的因素,其余选不变的因素的方法,必然有 C_n^2 种。……选择 k 个变
 动因素,其余选不变的因素的方法,必然有 C_n^k 种。也就是说一个因素单纯变动指数有 C_n^1 个,
 两因素共同变动指数有 C_n^2 个,三因素共同变动指数有 C_n^3 个,……, k 因素共变指数有 C_n^k 个
 ……, n 因素共同变动指数有 C_n^n 个。

当 $n = 2$ 时, $C_2^0 = 1, C_2^1 = 2, C_2^2 = 1$ $n = 3$ 时, $C_3^0 = 1, C_3^1 = 3, C_3^2 = 3, C_3^3 = 1$

$n = 4$ 时, $C_4^0 = 1, C_4^1 = 4, C_4^2 = 6, C_4^3 = 4, C_4^4 = 1$

$C_2^0, C_3^0, C_4^0 \dots C_n^0$ 表示没有因素变动,也就是原来的基期水平部分。

将 $n = 2, 3, 4 \dots n$ 排列起来,就组成有名的杨辉三角。故可以说,在研究 n 个因素变
 动时,其一个因素单纯变动、两因素共变、三因素共变…… n 因素共变,必然符合杨辉三角。
 其图形如下:

	1	2	1					$n = 2$
	1	3	3	1				$n = 3$
	1	4	6	4	1			$n = 4$
	1	5	10	10	5	1		$n = 5$
	1	6	15	20	15	6	1	$n = 6$
基 期 水 平	一 因 素 变 动	二 因 素 共 变	三 因 素 共 变	四 因 素 共 变	五 因 素 共 变	六 因 素 共 变		

多因素变动的分析是很有意义的,但计算上都相当麻烦。随着电子计算机的广泛应用,
 复杂的数学问题,电子计算机都能很快地计算出来,既节省人力,又能提高时效性,这对进
 行多因素变动分析将有很大帮助。今后多因素变动分析将能更好地开展。铁路统计指标中有
 许多受多因素的影响,如能用和差式综合指数体系进行分析,是比较方便的,不需要判断数
 量指标因素、质量指标因素,不需要绞尽脑汁安排置换顺序,用电子计算机计算,程序也较
 好编。这种方法如能补充到《铁路统计学》教材中,将是有益的。