

我国城镇居民收入与消费支出的典型相关分析

安徽农业大学管理科学学院 黄婷婷 安徽农业大学经济与贸易学院 刘鹏凌

[摘要]在市场经济条件下,收入是决定居民消费的最主要因素,城镇居民的消费性支出与其收入之间有着内在的联系。本文运用典型相关分析法,借助 SAS 软件对我国 31 个省、市、自治区城镇居民收入与消费性支出之间的相关关系进行了实证性分析判断。结论认为:工薪收入对城镇居民消费结构有决定性影响,财产性收入促进消费高级化。建议多渠道增加城镇居民工薪收入,开拓多元化收入渠道,稳定居民消费预期。

[关键词]典型相关分析 城镇居民 居民收入 消费支出

在市场经济条件下,影响居民消费支出的因素很多,诸如现期收入、预期收入、利率水平、借贷限制、财富存量、持久收入、暂时收入、消费者所在群体的平均消费水平、消费者曾经达到的最高消费水平等。但对于我国的城镇居民而言,收入仍然是决定居民消费的最主要因素。因此本文认为有必要就城镇居民收入与消费支出之间的关系进行典型相关分析,目的在于寻求影响城镇居民消费结构的深层次因素,为相关部门决策提供参考。

一、典型相关模型与指标选择

(一)典型相关模型。典型相关分析(Canonical Correlation Analysis)是近年来开始使用的一种新型多元统计分析方法。它最早由 H. Hotelling 于 1936 年在《生物统计》期刊上提出,经过长期应用发展,到 70 年代臻于成熟。特别是飞速发展的计算机技术基本解决典型相关分析中涉及的较大矩阵计算问题之后,它在各个领域的应用越来越广泛。典型相关分析所揭示的是两组多元随机变量之间的关系。其具体做法是:在第一组变量中提出一个典型变量,在第二组变量中也提出一个典型变量,并使这一对典型变量具有最大的相关;然后又在每一组变量中提出第二个典型变量,使得在与第一个典型变量不相关的典型变量中,这两个典型变量线性组合之间的相关性最大,如此下去,直到两组变量间的相关性被提取完毕为止。可见,典型相关分析是把原来两组变量之间的相关,转化为研究从各组中提出的少数几个典型变量之间的典型相关,从而减少研究变量的个数。下面给出具体表达式:设有两组随机变量 $X = (X_1, X_2, \dots, X_p)$, $Y = (Y_1, Y_2, \dots, Y_q)$, X 、 Y 的协方差阵为:

$$\text{cov} \begin{pmatrix} X \\ Y \end{pmatrix} = \begin{pmatrix} \sum_{11} & \sum_{12} \\ \sum_{21} & \sum_{22} \end{pmatrix}$$

$\sum_{11}$ 其中是第一组变量的协方差, $\sum_{12}$ 是第一组与第二组的协方差, $\sum_{22}$ 是第二组的协方差。

要研究 X 、 Y 两组变量的典型相关关系,需作两组变量的线性组合,即:

$$U = \alpha'X = \alpha_1 X_1 + \alpha_2 X_2 + \dots + \alpha_p X_p; V = b'Y = b_1 Y_1 + b_2 Y_2 + \dots + b_q Y_q$$

在 X 、 Y 及 Σ 给定条件下,即是求 α 、 b ,使得 U 、 V 之间的相关系数 $r = \frac{\text{cov}(U, V)}{\sqrt{\text{var}(U) \text{var}(V)}}$

3.2 公路工程项目中环境质量的定量核算。公路工程中具体的环境指标可以落实到如:

自然资源利用率 = 利用某自然资源的产出数量 / 实践消耗的该自然资源数量

环境污染程度 = 公路某影响区空气内某种污染物质的含量 / 正常空气中某种污染物质的含量

农田占用率 = 影响区内用于公路建设的农田数 / 该地区原有农田总数

.....

很多指标都可以反映公路建设中的环境质量。

通过上面的数学分析,我们了解到增加单位环境质量水平是通过环境质量获得的机会成本而获得的。但是对于不同的公路建设中投入的资源其质量机会成本的计算又都是不一样的。所以要想在公路项目决策中获得不同质量指标的提高需要分别计算不同的质量机会成本。在此以公路建设中占用农田的数量为例介绍数量关系的表达。

假设农田占用率环境质量指标是我们的决策指标,在满足帕累托最优标准的前提下,根据数学分析,要想获得环境质量的单位提高,必须通过得到农田占用机会成本而换得。而公路工程中农田的机会成本就是核算农田的最好可行替代用途而获得的,举例如下:

在江苏某地修建某港口项目,开工建设期为 2005 年,征用耕地 1133.771 亩,项目寿命期为 30 年。经测算该地区农田最好可行替代用途测算值 2004 年是 1205.75 元/亩且年增长率为 2%,社会折现率为 12%,该项目中土地机会成本的计算值:

$$OC = 1133.771 \times 1205.75 \times (1 + 0.02) \times \frac{1 - (1 + 0.02)^{30} (1 + 0.10)^{-30}}{0.10 - 0.02} = 1562.05 \text{ 万元}$$

通过计算出农田的机会成本,我们可以估算出获得农田占用率环境质量指标单位数量的提高需要在公路建设过程中多投入资金数量是多少才能达到要求,此结果可以为决策者参考。

[参考文献]

- 1.《绿色经济发展和管理》赵弘志等编 东北大学出版社 2003 年
 - 2.《资源与环境经济学》鲁传一编 清华大学出版社 2004 年
 - 3.《可行性研究原理与方法》李红锬编 电子科大出版社 2005 年 10
- 本文是重庆市教委人文社会科学项目资助项目,项目编号 04JWSK030

达到最大。标准化随机变量 U 与 V , 则: $\text{var}(U) = \text{var}(\alpha'X) = \alpha' \sum_{11} \alpha = 1$ $\text{var}(V) = \text{var}(b'Y) = b' \sum_{22} b = 1$
故: $r = \text{cov}(U, V) = \alpha' \text{cov}(X, Y) b = \alpha' \sum_{12} b$

于是问题转化为在 $\alpha' \sum_{11} \alpha = 1, b' \sum_{22} b = 1$ 的条件下, α, b 取何值, 使得 $r = \text{cov}(U, V) = \alpha' \sum_{12} b$ 达到最大。
为解决这个问题, 引入拉格朗日乘数 $\lambda_1/2, \lambda_2/2$ 构造一个目标函数:

$$f(\alpha, b) = \alpha' \sum_{12} b - \frac{\lambda_1}{2} \{\alpha' \sum_{11} \alpha - 1\} - \frac{\lambda_2}{2} \{b' \sum_{22} b - 1\}$$

求 $f(\alpha, b)$ 的一阶偏导数, 并令其为零, 得方程组:
$$\begin{cases} \frac{\partial f}{\partial \alpha} = \sum_{12} b - \lambda_1 \sum_{11} \alpha = 0 \\ \frac{\partial f}{\partial b} = \sum_{21} \alpha - \lambda_2 \sum_{22} b = 0 \end{cases}$$

将上面两方程分别左乘 α', b' , 并由前面条件可知: $\lambda_1 = \lambda_2 = r = \lambda$

于是, 解方程组:
$$\begin{cases} \sum_{12} b - \lambda \sum_{11} \alpha = 0 \\ \sum_{21} \alpha - \lambda \sum_{22} b = 0 \end{cases}$$

即可得 α_{jk} 和 b_{jk} , 并写出各典型变量 U_k 和 V_k 。

对所求得的典型变量, 还需对其显著性予以检验, 它是运用“去掉前 k 个典型相关系数的影响”之后, 剩下的 $p - k$ 个典型相关系数是否可达到显著水平而进行的。检验统计量:

$$Q_k = -[n - j - \frac{1}{2}(P + Q + 1)] \ln \left[\prod_{i=j+1}^p (1 - \lambda_i^2) \right] \text{ 服从 } \chi^2 \text{ 分布, 自由度为 } (p - k + 1)(q - k + 1)。$$

典型相关分析在经济实证分析研究中有着广阔的应用前景, 因为许多经济现象之间都表现出多个变量对多个变量的关系。下面就用典型相关分析法来分析我国城镇居民收入与消费支出之间的关系。

(二) 数据来源及指标选择。本文所用的数据为 2003 年我国 31 个省市自治区城镇居民家庭收入和消费性支出数据, 数据来自《2004 年中国统计年鉴》。两组指标变量, 第一组为反映城镇居民家庭收入的变量: X_1 - 工薪收入, X_2 - 经营净收入, X_3 - 财产性收入, X_4 - 转移性收入; 第二组为反映城镇居民消费性支出的变量: Y_1 - 食品支出, Y_2 - 衣着支出, Y_3 - 居住支出, Y_4 - 家庭设备用品及服务支出, Y_5 - 医疗保健支出, Y_6 - 交通通讯支出, Y_7 - 文教支出, Y_8 - 杂项商品及服务支出。本文计算结果都是通过 SAS 软件运算得出的。

二、收入与消费的典型相关分析

(一) 总体典型相关系数及其检验。由表 1 知, 前两个典型相关系数较高, 表明相应典型变量之间密切相关, 并且前两对典型相关通过似然率卡方检验, 而第 3、4 对特征值小于 1, 且典型相关统计检验也不显著。再对 X 变量与 Y 变量之间的多变量进行检验, 如表 2 所示, Wilks' Lambda、Pillai's Trace、Hotelling - Lawley Trace、Roy's Greatest Root 统计量的概率值均小于 0.0001, 因此可以说明 X 变量与 Y 变量之间有显著的交互效应。

表 1 典型相关系数、特征根及其检验

序号	典型相关系数	特征根	F 值	P 值
1	0.990137	49.9475	8.23	< 0.0001
2	0.795644	1.7252	2.48	0.0037
3	0.663318	0.7857	1.91	0.0624
4	0.528106	0.3868	1.62	0.1972

表 2 X 变量与 Y 变量之间的多变量检验

统计量	值	F 值	自由度	P 值
Wilks' Lambda	0.00290856	8.23	32	< 0.0001
Pillai's Trace	2.33230784	3.67	32	< 0.0001
Hotelling - Lawley Trace	52.84515390	27.91	32	< 0.0001
Roy's Greatest Root	49.94754457	131.11	8	< 0.0001

(二) 线性结合模型与结构分析

根据典型变量重要程度及系数大小, 从建立的典型相关模型可看出, 中国城镇居民消费性支出情况受收入各因素变动的影响程度可用两对典型相关变量予以综合描述。

第一对典型变量将食品、杂项商品劳务从其他类型支出中分离出来 (典型载荷分别为 0.401 和 0.5608), 与工薪收入呈现最大相关 (相应典型载荷为 0.7140), 次之是转移性收入 (典型载荷为 0.3298)。综合表 3 和表 4, 一些典型相关系数与典型系数符号不同, 如 U_1 和 X_2 的相关系数为负值 (-0.0221), 而典型相关系数却为正值 (0.1813), 这说明经营净收入为抑制变量, 也就是说该变量降低了该组典型变量的相关系数。

表 3 典型相关模型

序号	典型相关模型
1	$U_1 = 0.7140X_1 - 0.0221X_2 + 0.0597X_3 + 0.3298X_4$ $V_1 = 0.401Y_1 + 0.0601Y_2 + 0.1889Y_3 + 0.0355Y_4 + 0.0724Y_5 + 0.1821Y_6 + 0.0954Y_7 + 0.5608Y_8$
2	$U_2 = 1.1264X_1 + 0.7102X_2 - 1.1891X_3 - 0.6188X_4$ $V_2 = -0.6443Y_1 - 0.7662Y_2 - 0.6534Y_3 - 0.9424Y_4 - 0.077Y_5 + 1.1174Y_6 + 0.6661Y_7 + 0.5201Y_8$

表 4 结构分析表(相关系数)

	U ₁	U ₂	V ₁	V ₂
X ₁	0.9704	0.1356	0.9608	0.1079
X ₂	0.1813	0.3231	0.1795	0.2571
X ₃	0.5923	- 0.4173	0.5865	- 0.3320
X ₄	0.8362	- 0.1964	0.8279	- 0.1563
	U ₁	U ₂	V ₁	V ₂
Y ₁	0.7310	0.1333	0.7238	0.1061
Y ₂	0.2647	0.8050	0.2621	0.6405
Y ₃	0.8452	0.1762	0.8368	0.1402
Y ₄	0.7534	- 0.2534	0.7460	- 0.2016
Y ₅	0.8623	0.1525	0.8538	0.1213
Y ₆	0.8078	0.3760	0.7998	0.2992
Y ₇	0.7281	0.2211	0.7209	0.1759
Y ₈	0.4023	- 0.4043	0.3984	- 0.3216

第二对典型变量与消费支出中的 Y₁、Y₂、Y₃、Y₄、Y₆、Y₇ 和 Y₈ 均高度相关,其中交通通讯支出和家庭设备用品及服务支出最为突出(典型载荷分别为 1.1174、- 0.9424),与之相对应的“收入组”与 X₁、X₂、X₃ 和 X₄ 均呈高度相关,其中最主要是财产性收入和工薪收入(典型载荷分别为 - 1.1891 和 1.1264)。相应的抑制变量分别为衣着、家庭设备、交通通讯和杂项商品支出。

从表 4 我们可以看出,收入组的第一典型变量 U₁ 和 X₁、X₄ 的相关系数较大(分别为 0.9704 和 0.8362),也就是说具有较强的预测能力,同时,U₁ 还对消费支出的 Y₁、Y₃、Y₄、Y₅、Y₆ 和 Y₇ 具有较强的预测能力,因为收入的第一典型变量与工薪收入和转移性收入密切相关,而这两部分占据了城镇居民收入的绝大部分,这也就不奇怪它们将对大部分的消费支出起决定作用。收入组的第二典型变量 U₂ 对消费性支出的 Y₂ 具有较强的预测能力。

(三)典型冗余分析。典型冗余分析(Canonical Redundancy Analysis)就是分析每一个典型变量解释或预测跟它相应的那组变量的能力大小。收入变量通过它的第一个典型变量解释的共享方差的比例是 50.

61%,而通过配对的另一个典型变量解释消费性支出的方差比例是 98.04%,这样,冗余指数则为 49.62%(= 50.61% × 98.04%),说明收入变量的第一个典型函数可以解释消费支出(因变量组)的总方差的比例为 49.62%。第二个典型函数则分别解释了收入变量、消费性支出变量的 8.39%、63.6%,相应的收入变量的第二个典型函数可以解释消费支出的总方差的比例为 5.31%。

(四)典型变量的回归分析。由于第一对典型变量可以给我们提供大量的有效信息有助于解释收入组综合变量对消费组综合变量的影响,因而我们可建立 V₁ 对 U₁ 的线性回归方程,得 V₁ 对 U₁ 的样本回归方程为:

$$V_1 = 143.695 + 0.262U_1 \quad F = 501.654 \quad R^2 = 0.944 \\ (2.278)$$

此模型的回归拟合系数 R² = 0.944,也就是说消费性支出的综合变量 V₁ 的变动有 94.4%可以由收入来源综合变量 U₁ 的变动来解释,而收入组综合变量每增加一元钱将使消费性支出综合变量增加 0.262 元。

三、结论与政策建议

(一)工薪收入对城镇居民消费结构有决定性影响,财产性收入促进消费高级化。从全国平均水平来看,城镇居民收入的 77%来源于工薪收入,因而其对城镇居民的消费结构有着举足轻重的影响。我们运用典型相关方法分析的结果也验证了这一点,具有统计意义的第一组和第二组典型收入组变量都与工薪收入密切相关,两个典型变量可解释典型冗余共享方差的 56%。但我国地域广阔,各地区经济发展状况也不尽相同,因而各地区城镇居民的收入来源存在一些差异,这就会影响到消费结构,最为突出的就是财产性收入的影响。相对于食品支出影响因素而言,衣着、家庭设备、交通通讯等相对高级化消费的变动更受到财产性收入的影响。随着要素市场的放开,居民财产性收入将越来越多,高收入户拥有的股票、利息、出租房屋等的收入远高于低收入户,导致高、低收入户之间的收入差距拉大。尽管居民消费并不完全取决于收入,但是财产性收入较多的居民对收入的预期明显要比其他居民户要好,正是这些收入和心理因素使财产性收入表现出促进消费高级化的特征。

(二)多渠道增加城镇居民工薪收入,开拓多元化收入渠道,稳定居民消费预期。消费主要取决于收入,消费质量的进一步提高也主要取决于收入水平的提高,现阶段一个突出的问题就是人均可支配收入增长慢于经济增长。应该看到,工薪收入仍将是居民家庭收入增长的主要动力。因此,政府应采取相应的积极措施,将社会就业问题作为今后工作的重点,更好地解决经济发展中的社会问题,免费开展各种类型的岗位技能培训,提高贫困人口走向市场的能力,降低就业准入门槛,提高劳动者综合技能,扩大就业面。同时应进一步推进和完善收入分配政策,加大对低收入户教育扶贫力度和财政转移支付力度,加大社会保障力度,包括确保养老金按时发放、落实解除劳动合同的赔偿制度、扩大居民最低生活保障覆盖面和提高最低工资标准等措施,这对提高低收入家庭的收入,具有极其重要的作用。影响人们消费预期的一个重要因素就是物价,政府应通过宏观经济政策保持市场物价平衡,避免因物价上涨导致居民可支配收入下降从而影响消费质量。要缩小居民的收入差距,这样可以给居民消费一个较良好的未来预期,从而促进居民生活质量整体提高。

【参考文献】

[1]袁志发,周静芋.多元统计分析[M].科学出版社,2002
[2]阮桂海,等.SAS 统计分析实用大全[M].清华大学出版社,2003
[3]李谷成,冯中朝.中国农户消费-收入结构的实证分析[J].农业技术经济,2004(6)
[4]国家统计局.中国统计年鉴[Z].中国统计出版社,2004
本文是安徽农业大学校青年基金资助项目《我国居民消费行为研究》的部分研究成果。

