

基于主成分回归法的新疆卫生总费用现状及影响因素分析

卢芸潇

伊犁师范大学 教育科学研究所, 新疆 伊宁

收稿日期: 2024年3月19日; 录用日期: 2024年4月9日; 发布日期: 2024年4月18日

摘要

本文使用主成分回归分析方法对影响新疆卫生总费用的主要因素进行实证分析, 结果表明城镇化率对人均卫生费用的影响最大, 卫生机构人员数(人)、住院病人日均费用(元)、卫生机构床位数(张)、人均国内生产总值(亿元)等卫生资源、经费与社会经济因素均正向影响新疆卫生总费用, 出生率负向影响新疆卫生总费用。新时代中国在经济增速放缓的趋势下, 新疆卫生费用的增长受多方面因素共同影响, 城镇化率的进程是推动新疆卫生费用增长的重要原因, 卫生资源与社会经济因素将在推动新疆卫生总费用增长中发挥巨大作用。

关键词

新疆卫生总费用, 主成分分析, 多元线性回归分析

Analysis of the Status Quo and Influencing Factors of Total Health Expenditure in Xinjiang Based on Principal Component Regression Method

Yunxiao Lu

Institute of Education Science, Yili Normal University, Yining Xinjiang

Received: Mar. 19th, 2024; accepted: Apr. 9th, 2024; published: Apr. 18th, 2024

Abstract

In this paper, principal component regression analysis is used to make an empirical analysis of the

main factors affecting the total health expenditure in Xinjiang. The results show that the urbanization rate has the greatest impact on per capita health expenditure, and health resources and expenditures such as the number of personnel in health institutions (people), the average cost of inpatients (yuan), the number of beds in health institutions (Zhang), and per capita GDP (100 million yuan) are positive with socio-economic factors. The birth rate negatively affects the total health cost in Xinjiang. Under the trend of slowing down economic growth in China in the new era, the increase in health expenses in Xinjiang is jointly affected by many factors. The urbanization rate is an important reason for the increase of health expenses in Xinjiang, and health resources and socio-economic factors will play a huge role in promoting the increase in total health expenses in Xinjiang.

Keywords

Total Health Expenditure in Xinjiang, Principal Component Analysis, Multiple Linear Regression Analysis

Copyright © 2024 by author(s) and Hans Publishers Inc.

This work is licensed under the Creative Commons Attribution International License (CC BY 4.0).

<http://creativecommons.org/licenses/by/4.0/>



Open Access

1. 引言

近 20 年尤其是近 10 年,我国政府不断加大对卫生健康的投入力度,卫生健康事业呈现高质量发展态势[1]。“健康中国”是实现人民健康与经济社会协调发展的国家战略,而卫生总费用是与两大产业的总产出、增加值一起作为衡量“健康中国”规模的重要经济指标[2]。

卫生总费用(Total Health Expenditure, THE)是指以货币形式作为综合计量手段,全面反映一个国家或地区在一定时期内(通常指一年)全社会用于医疗卫生的资金总额[3]。卫生总费用是衡量一个国家或地区的卫生投入和卫生筹资状况的指标,代表了卫生事业发展与经济、社会之间的关系,反映一个国家或地区卫生医疗筹资水平和利用程度,是政府制定和调整卫生政策的重要依据[4]。

卫生总费用持续上涨是包括中国在内的大部分国家共同关注的社会问题,而新疆作为我国西北地区的主要省份,在调整经济结构、促进经济发展的同时,也面临着卫生总费用持续上涨带来的社会经济负担。在新医改的时代背景下,新疆政府卫生投入持续增加,医保覆盖范围进一步扩大,居民的卫生健康需求随着生活质量的提高而增加,新疆卫生总费用稳步增长。分析新疆卫生总费用的发展趋势及其影响因素,探索新疆卫生总费用的发展趋势、各项影响因素及其与卫生总费用的相关性,为控制新疆卫生总费用增长以及为新疆制定卫生改革措施具有重要意义。

2. 资料来源

研究数据来源于 2012~2017 年《中国卫生统计年鉴》[5] 2018~2022 年《中国卫生健康统计年鉴》[6] 及 2012~2022 的《新疆统计年鉴》[7]。卫生总费用包含政府、社会及个人卫生支出。分别从经济社会、人口情况和卫生资源等角度,选取影响新疆卫生总费用的 9 个变量,见下表:

维度	变量名称	变量标号
经济社会	人均国内生产总值(亿元)	X_1
	城镇化率	X_2

续表

卫生资源	卫生机构数(个)	X_3
	卫生机构床位数(张)	X_4
	卫生机构人员数(人)	X_5
卫生经费	门诊病人次均医疗费用(元)	X_6
	住院病人次均费用(元)	X_7
卫生服务	诊疗人次数(万人次)	X_8
人口	出生率	X_9
	卫生总费用(亿元)	Y

3. 研究方法与实证分析

3.1. 因素确定

根据《新疆统计年鉴》《中国卫生统计年鉴》《中国卫生健康统计年鉴》[5] [6] [7]获取原始数据如表 1。

Table 1. Raw data

表 1. 原始数据

年份	X_1	X_2	X_3	X_4	X_5
2011	30,087	43.73	7587	125,391	155,701
2012	33,796	44.22	7820	131,542	150,083
2013	37,553	44.94	8096	137,325	159,513
2014	40,648	46.79	8285	142,956	167,598
2015	40,036	48.78	8359	150,263	178,063
2016	40,564	50.42	8330	165,925	188,399
2017	44,941	51.90	8294	169,636	208,204
2018	49,475	54.01	8250	173,670	212,718
2019	54,280	55.51	8285	181,175	221,297
2020	53,593	56.53	8175	182,884	228,800
年份	X_6	X_7	X_8	X_9	
2011	153.7	5159.5	7680.39	14.99	
2012	106.3	4353.0	8386.56	15.32	
2013	117.9	4646.6	9381.95	15.84	
2014	126.4	4873.8	9978.97	16.44	
2015	133.0	5168.8	10333.27	15.60	
2016	136.2	5479.8	9584.68	15.34	
2017	144.5	5786.1	9851.08	15.88	
2018	158.6	6353.1	9392.27	10.69	
2019	164.9	6779.9	10619.02	8.14	
2020	178.9	6932.4	9509.16	6.90	

2011~2020 年的新疆卫生费用相关的数据如表 1。根据表 1 中数据对新疆卫生总费用的相关影响因素进行主成分分析。

3.2. 卫生总费用的主成分分析

运用主成分分析法从经济社会、卫生资源、人口三个方面对新疆卫生总费用的主要因素进行分析，得出影响新疆卫生总费用的首要因素为相关部门卫生战略决策提供参考。

3.2.1. 主成分分析方法的步骤

利用主成分分析的方法做综合评价的原理和步骤如下[8] [9]:

假设有 n 个样本，每个样本有 p 个指标： x_1, x_2, \dots, x_p ，得到原始数据观测矩阵：

$$X_{n \times p} = \begin{pmatrix} x_{11} & x_{12} & \cdots & x_{1p} \\ x_{21} & x_{22} & \cdots & x_{2p} \\ \vdots & \vdots & & \vdots \\ x_{n1} & x_{n2} & \cdots & x_{np} \end{pmatrix} \quad (2.1)$$

综合指标为 $F_1, F_2, \dots, F_m (m \leq p)$ ，即：

$$\begin{cases} F_1 = a_{11}x_1 + a_{21}x_2 + \cdots + a_{p1}x_p \\ F_2 = a_{12}x_1 + a_{22}x_2 + \cdots + a_{p2}x_p \\ \vdots \\ F_m = a_{1m}x_1 + a_{2m}x_2 + \cdots + a_{pm}x_p \end{cases} \quad (2.2)$$

1) 为了消除量纲在实际应用中的影响，在计算前首先进行标准化处理，使得各项指标更加集中，得到的标准化形式为：

$$x'_{ij} = (x_{ij} - \bar{x}_i) / \sigma_i (i = 1, 2, \dots, n; j = 1, 2, \dots, p) \quad (2.3)$$

2) 相关系数矩阵 $R_{n \times n}$ 的计算公式为：

$$r_{ij} = \frac{1}{n} \sum_{k=1}^n (x_{ki} - \bar{x}_i)(x_{kj} - \bar{x}_j) / \sigma_i \sigma_j \quad (2.4)$$

计算 $R_{n \times n}$ 的特征值和特征向量。通过计算 $R - \lambda I = 0$ ，得出特征值 λ_i 满足条件 $\lambda_1 \geq \lambda_2 \geq \dots \geq \lambda_n$ ，且对应与其标准正交的特征向量 u_1, u_2, \dots, u_n 。

3) 计算主成分：

$$F_j = \sum_{i=1}^n \sum_{j=1}^p u_{ij} x'_{ij} \quad (2.5)$$

4) 综合分析。使用前 m 个主成分来代替总方差的绝大部分，通过累计贡献率进行判断，通过累计贡献率达到 80% 以上来确定 m ，接着对前 m 个主成分进行综合分析。

5) 计算综合得分。第 i 个影响因子的第 k 个主成分得分： $s_{ik} = \sum_{j=1}^p u_{jk} x'_{ij}$ ； $k = 1, 2, \dots, m$ ； $i = 1, 2, \dots, n$ 。

第 i 个影响因子的综合得分：

$$SOCRE_i = \sum_{k=1}^m e_k s_{ik} ; i = 1, 2, \dots, n \quad (2.6)$$

3.2.2. 主成分分析的实证结果

运用主成分分析法先消除量纲影响变量进行标准化，得到标准化变量相关系数矩阵，用于判断选取的 9 个指标是否能进行因子分析，结果见表 2。

Table 2. Correlation coefficient matrix

表 2. 相关系数矩阵

	X_1	X_2	X_3	X_4	X_5	X_6	X_7	X_8	X_9
Z score X_1	1.000	0.964	0.652	0.948	0.950	0.707	0.900	0.681	-0.817
Z score X_2	0.964	1.000	0.621	0.988	0.992	0.787	0.949	0.598	-0.808
Z score X_3	0.652	0.621	1.000	0.685	0.582	0.147	0.393	0.929	-0.152
Z score X_4	0.948	0.988	0.685	1.000	0.979	0.716	0.907	0.638	-0.732
Z score X_5	0.950	0.992	0.582	0.979	1.000	0.819	0.960	0.561	-0.790
Z score X_6	0.707	0.787	0.147	0.716	0.819	1.000	0.933	0.175	-0.820
Z score X_7	0.900	0.949	0.393	0.907	0.960	0.933	1.000	0.413	-0.887
Z score X_8	0.681	0.598	0.929	0.638	0.561	0.175	0.413	1.000	-0.237
Z score X_9	-0.817	-0.808	-0.152	-0.732	-0.790	-0.820	-0.887	-0.237	1.000

从表 2 中可见这 9 个变量的相关性程度较高，检验后得到，从 9 个变量中提取出主成分公因子，进行主成分分析可行。

为了检验各个自变量是否适宜进行因子分析，首先对原始数据进行 KMO 和 Bartlett 检验，结果见表 3。

Table 3. The KMO and Bartlett test

表 3. KMO 和 Bartlett 的检验

取样足够度的 Kaiser-Meyer-Olkin 度量	0.641
近似卡方	140.492
Bartlett 的球形度检验	df
	36
	Sig.
	0.000

表 3 中 KMO 检验的取值 $0.641 \geq 0.5$ ，Bartlett 球度检验统计量近似值 χ^2 为 140.492，其伴随概率的 Sig 值为 0.000，小于显著性水平 0.05，检验结果表明数据能够进行主成分分析。

表 4 输出的是两个主成分与 9 个自变量的共同度，它显示运用主成分分析法提取了 9 个自变量的特征值。

Table 4. Table of changes in variance of common factors

表 4. 公因子方差变化表

	初始	提取
X_1	1.000	0.959
X_2	1.000	0.986
X_3	1.000	0.975
X_4	1.000	0.959
X_5	1.000	0.975

续表

X_6	1.000	0.896
X_7	1.000	0.990
X_8	1.000	0.929
X_9	1.000	0.877

表 4 中主成分分析得到的 9 个原始变量标准化后的方差是 1，通过因子分析的方法提取的主成分因子可以用来很好的解释该因子的方差。

表 5 为主成分分析法按照累计方差贡献率大于 85% 得到的提取结果。

Table 5. Explains the total variance

表 5. 解释的总方差

成分	初始特征值			提取平方和载入		
	合计	方差的%	累积%	合计	方差的%	累积%
1	6.899	76.653	76.653	6.899	76.653	76.653
2	1.648	18.311	94.964	1.648	18.311	94.964
3	0.232	2.580	97.544			
4	0.159	1.772	99.316			
5	0.035	0.385	99.701			
6	0.020	0.224	99.925			
7	0.004	0.050	99.974			
8	0.002	0.017	99.991			
9	0.001	0.009	100.000			

由表 5 知，前 2 个主成分的累积贡献率为 94.964%，代表了绝大部分信息，由此可选用前两个新变量作为主成分以代替原来的 9 个原始变量。

为了锁定公因子数，依照公因子特征值大于 1 的法则，从统计软件获取各成分特征值图表，输出如图 1 的碎石图。

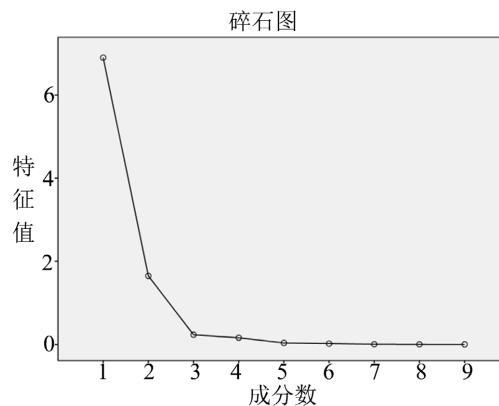


Figure 1. Gravel diagram

图 1. 碎石图

图 1 中的特征值变化情况在前 2~3 个公因子位置变化较明显, 从第 3 个因子起, 特征值变化逐渐趋于平缓, 这说明截取前 2 个公因子作为主成分是适宜的, 能够将绝大部分的信息囊括在内。

表 6 为计算得到的因子载荷矩阵, 能显示 9 个指标在公因子上的重要度。

Table 6. Factor loading matrix

表 6. 因子载荷矩阵

标准化后的自变量	成份	
	1	2
X_1	0.977	0.073
X_2	0.993	-0.010
X_3	0.628	0.762
X_4	0.975	0.087
X_5	0.986	-0.052
X_6	0.805	-0.498
X_7	0.956	-0.276
X_8	0.635	0.725
X_9	-0.822	0.449

为了规避各因子变量不突出的问题, 明确的解释出 2 个公因子的实际意义, 利用方差最大化旋转因子载荷的结构见表 7。

Table 7. Rotation factor load matrix

表 7. 旋转因子载荷矩阵

	成份	
	1	2
X_1	0.796	0.570
X_2	0.853	0.508
X_3	0.140	0.977
X_4	0.788	0.582
X_5	0.869	0.469
X_6	0.946	-0.007
X_7	0.960	0.262
X_8	0.165	0.950
X_9	-0.936	-0.044

由表 7 可以看出, 第 1 主成分在门诊病人均次医疗费用、住院病人均次费用上有着较大的负荷, 载荷分别为 0.946、0.960, 记为 F_1 成分主要反映了卫生经费因素; 第 2 主成分在卫生机构数、诊疗人次数上有着较大的负荷, 载荷分别为 0.977、0.950, 记为 F_2 成分, 反映了卫生资源与服务因素。

由统计软件分析得到 2 个主成分的成分得分系数矩阵以计算每个指标所对应的系数。见表 8。

Table 8. Matrix of component score coefficients
表 8. 成分得分系数矩阵

	成份	
	1	2
X_1	0.098	0.111
X_2	0.126	0.070
X_3	-0.163	0.442
X_4	0.093	0.119
X_5	0.138	0.047
X_6	0.257	-0.197
X_7	0.206	-0.071
X_8	-0.150	0.424
X_9	-0.244	0.171

由表 8 得到两个主成分表达式为：

$$F_1 = 0.098X_1 + 0.126X_2 - 0.163X_3 + 0.093X_4 + 0.138X_5 + 0.257X_6 + 0.206X_7 - 0.150X_8 - 0.244X_9,$$

$$F_2 = 0.111X_1 + 0.070X_2 + 0.442X_3 + 0.119X_4 + 0.047X_5 - 0.197X_6 - 0.071X_7 + 0.424X_8 + 0.171X_9.$$

综合得分计算公式如下：

$$F = \frac{6.899}{6.899 + 1.648} F_1 + \frac{1.648}{6.899 + 1.648} F_2$$

3.3. 基于主成分的新疆卫生总费用的回归分析

3.3.1. 多元线性回归模型的求解

多元线性回归模型数学表达式为：

$$y_i = a + b_1x_{1i} + b_2x_{2i} + \dots + b_jx_{ji} + \varepsilon (j = 1, 2, \dots, n) \tag{2.7}$$

式中 y 叫做因变量， ε 、 a 、 b_1 、 b_2 、 \dots 、 b_j 是模型中的未知参数，分别称作回归常数和偏回归系数， ε 称作误差，是一个随机变量。用最小二乘法求解 a 和 b_j ， $j = 1, 2, \dots, p$ 的值。

$$SSE = \sum_{i=1}^n (y_i - \bar{y}_i)^2 \tag{2.8}$$

$$\bar{y}_i = a + b_1x_{1i} + b_2x_{2i} + \dots + b_jx_{ji}, (i = 1, 2, \dots, n) \tag{2.9}$$

令 a 和 b_j 的偏导数为 0 得：

$$\begin{cases} \frac{\partial SSE}{\partial a} = 0 \\ \frac{\partial SSE}{\partial b_j} = 0 \end{cases}, j = 1, 2, \dots, p \tag{2.10}$$

经过计算得到如下方程式：

$$\begin{cases} \sum_{i=1}^n y_i = na + b_1 \sum_{i=1}^n x_{1i} + b_2 \sum_{i=1}^n x_{2i} + \cdots + b_p \sum_{i=1}^n x_{pi} \\ \sum_{i=1}^n x_{1i} y_i = a \sum_{i=1}^n x_{1i} + b_1 \sum_{i=1}^n x_{1i}^2 + b_2 \sum_{i=1}^n x_{1i} x_{2i} + \cdots + b_p \sum_{i=1}^n x_{1i} x_{pi} \\ \vdots \\ \sum_{i=1}^n x_{pi} y_i = a \sum_{i=1}^n x_{pi} + b_1 \sum_{i=1}^n x_{pi} x_{1i} + b_2 \sum_{i=1}^n x_{pi} x_{2i} + \cdots + b_p \sum_{i=1}^n x_{pi}^2 \end{cases} \quad (2.11)$$

利用矩阵表示法化简这些表达式。令

$$Y = \begin{pmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{pmatrix}, \quad X = \begin{pmatrix} 1 & x_{11} & x_{21} & \cdots & x_{p1} \\ 1 & x_{12} & x_{22} & \cdots & x_{p2} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_{1n} & x_{2n} & \cdots & x_{pn} \end{pmatrix}, \quad B = \begin{pmatrix} a \\ b_1 \\ \vdots \\ b_p \end{pmatrix}, \quad \varepsilon = \begin{pmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{pmatrix} \quad (2.12)$$

线性回归模型可表示为：

$$Y = XB + \varepsilon \quad (2.13)$$

$$X'XB = X'Y \quad (2.14)$$

最后得到：

$$B = (X'X)^{-1} X'Y \quad (2.15)$$

3.3.2. 拟合程度的测定

为了确定模型的合理性，回归模型由最小二乘法的估计值得到参数值，进而检验与评价。

因变量的值“总离差平方和” SST 分为两部分：

$$SST = \sum (y_i - \bar{y})^2 = \sum (\hat{y}_i - \bar{y})^2 + \sum (y_i - \hat{y}_i)^2 \quad (2.16)$$

分别为：

$$SSR = \sum (\hat{y}_i - \bar{y})^2, \quad SSE = \sum (y_i - \hat{y}_i)^2 \quad (2.17)$$

SSR 为回归平方和， SSE 为残差平方和。

多元线性回归中决定系数 r^2 的计算公式为：

$$r^2 = \frac{SSR}{SST} = 1 - \frac{SSE}{SST} \quad (2.18)$$

在因变量 y 的变化中， r^2 越大，回归的样本数据的拟合程度越强，因变量 Y 与自变量 X 的关系越密切。

复相关系数 r 定义为：

$$r = \sqrt{\frac{SSR}{SST}} = \sqrt{1 - \frac{SSE}{SST}} \quad (2.19)$$

复相关系数是因变量 Y 与所有自变量 X 构成的整体相关系数。复相关系数 $r \geq 0$ ，通过检验复相关系数的显著性对回归方程进行方差分析。

3.3.3. 回归模型的统计量检验

回归方程的显著性检验通常用 F 检验。多元线性回归方程显著性检验的零假设为

$H_0: b_1 = b_2 = \dots = b_j = 0$ ，用于检验的 F 统计量计算公式为：

$$F = \frac{\frac{\sum(\hat{y}_i - \bar{y})^2}{p}}{\frac{\sum(y_i - \hat{y}_i)^2}{n-p-1}} = \frac{r^2}{\frac{1-r^2}{n-p-1}} \sim F(p, n-p-1) \quad (2.20)$$

式中 p 表示自变量个数， n 是样本的观测个数。当给定显著性水平 α 时，若 P 值小于给定显著性水平即 $F > F_\alpha$ 时，拒绝零假设。

当回归方程检验的显著性通过时，继续进行回归系数的检验，回归系数的显著性检验零假设： $H_0: b_1 = b_2 = \dots = b_i = 0 (i = 1, 2, \dots, p)$ ，检验 b_i 的显著性统计量为 t 统计量。计算公式为：

$$t_i = \frac{\hat{b}_i}{S_{\hat{b}_i}} \quad (2.21)$$

式中，

$$S_{\hat{b}_i} = \sqrt{\frac{\sum(y_i - \hat{y}_i)^2}{(n-p-1)\sum(x_{ji} - \bar{x}_i)^2}} \quad (2.22)$$

当 $|t| > t_{\alpha/2}(n-k-1)$ 时，拒绝零假设。

3.3.4. 多元线性回归的实证分析

将因子分析中的两个公因子作为回归分析中的自变量，将新疆卫生总费用作为回归分析中的因变量进行因子分析。

多元线性回归输出结果如表 9。

Table 9. Model summary table

表 9. 模型汇总表

模型	R	R 方	调整 R 方	标准估计的 误差	更改统计量					Durbin-Watson
					R 方更改	F 更改	df1	df2	Sig.0F 更改	
1	0.992a	0.983	0.978	0.14705135	0.983	204.601	2	7	0.000	2.179

a. 预测变量：(常量)， F_1 。b. 预测变量：(常量)， F_1 ， F_2 。c. 因变量：Z score (Y)。

线性回归结果在表 9 表现出 R^2 值达到 0.983 说明模型与数据拟合程度较好，方程的 DW 检验值为 2.179，可见残差存在一定的正相关。

表 10 是多元线性回归分析的最终方程输出与对变量所建立方程的检验 (F 检验)。

Table 10. ANOVA table

表 10. 方差分析表

模型	平方和	df	均方	F	Sig.
1 回归	8.849	2	4.424	204.601	0.000b
残差	0.151	7	0.022		
总计	9.000	9			

a: 因变量：Z score (Y)。b: 预测变量：(常量)， F_1 。c: 预测变量：(常量)， F_1 ， F_2 。

从表 10 中可以看出回归模型中观测到回归方差是大于残差的，且 F 值为 204.601，显著水平为 0.000 小于显著性水平 $\alpha = 0.05$ ，证明变量间的线性关系显著、建立线性回归模型合理。

表 11 中输出的是自变量的回归系数以及回归系数显著性检验的情况。

Table 11. Coefficients output table
表 11. 系数输出表

模型	非标准化系数			t	Sig.	共线性统计量
	B	标准误差	试用版			
(常量)	-3.911E-017	0.047	0.000	1.000		
2 F_1	0.327	0.019	0.860	17.544	0.000	1.000
F_2	0.385	0.038	0.494	10.070	0.000	1.000

a: 因变量: Z score (Y)。

由表 11 可见显著性水平 $\alpha = 0.05$ ，表中回归系数显著性检验的概率 $P < \alpha$ ，故得到的模型是合理的。表中最后两列为共线性统计，容差和膨胀系数(VIF)互为倒数关系，容差越大发生共线性的可能就越大。由表中结果可知解释变量 F_1 、 F_2 的容差和膨胀系数(VIF)均等于 1，都小于 5，因此解释变量间无共线性情况。

图 2 为数据的检验结果，从残差的直方图(左)可以看出残差分布比较均匀，近似正态分布，说明被解释变量服从正态分布，从残差的正态 P-P 图(右)发现散点基本呈直线趋势，说明模型拟合较好，该数据样本满足回归分析的条件。

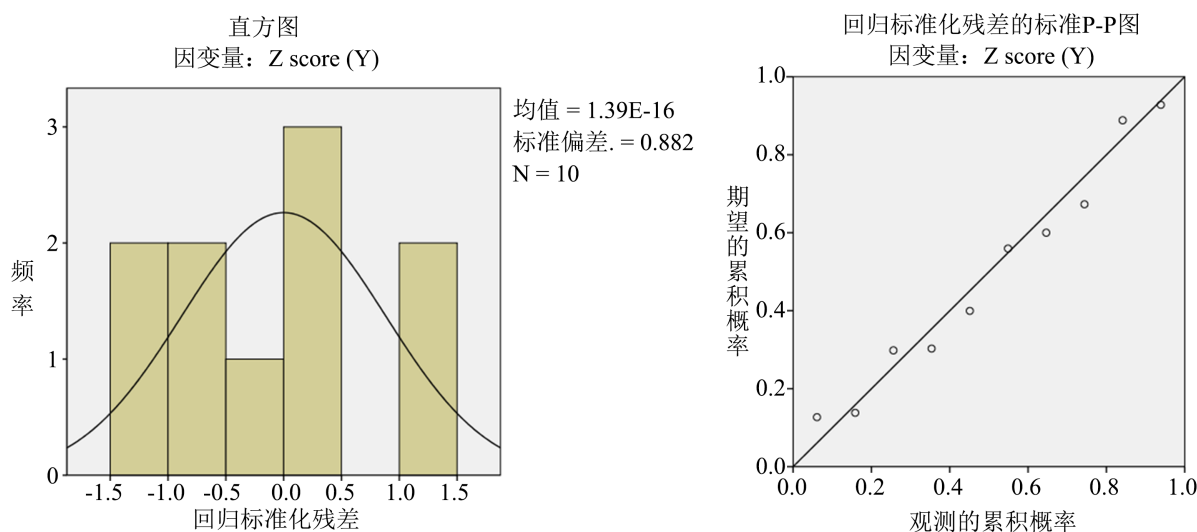


Figure 2. Graph of data test results
图 2. 数据检验结果图

通过检验发现解释变量的选取对新疆卫生总费用的影响较为直接，可以作为分析新疆卫生总费用的理论依据，模型建立合理，有实际应用价值。故根据输出结果得到多元线性回归方程为：

$$Y = 0.860X_{F_1} + 0.494X_{F_2}$$

将因子分析中 F_1 和 F_2 的表达式带入回归方程得:

$$Y = 0.139X_1 + 0.143X_2 + 0.078X_3 + 0.139X_4 + 0.142X_5 + 0.124X_6 + 0.142X_7 + 0.080X_8 - 0.125X_9.$$

从回归方程式可得, 城镇化率、卫生机构人员数(人)和住院病人均费用(元)系数较大, 可以说这三个变量对于新疆卫生总费用的影响是较大的, 是新疆卫生总费用的重要影响因素。

4. 讨论

不同因素对卫生费用的增长具有不同作用。本研究使用主成分的多元线性回归模型, 研究影响2010~2020年新疆卫生总费用的影响因素。实证结果显示, 人均国内生产总值、城镇化率、卫生机构数、卫生机构床位数、卫生机构人员数、门诊病人均医疗费用、住院病人均费用诊疗人次数、出生率均显著影响新疆卫生总费用且影响效果显著。

随着经济水平和人民生活水平的日益提高, 居民越来越重视医疗卫生, 卫生总费用相应增多。其中, 卫生机构数、卫生机构床位数和卫生机构人员的增加会进一步推动医疗成本的上升, 从而带来卫生费用的上涨。出生率负向影响人均卫生费用, 出生人口的增加会提高医疗卫生的需求, 从而在一定程度上提高了新疆卫生总费用, 但不可否认的是, 随着社会进步和经济水平的提升, 特别是医疗保险政策的普及和使用减少了相关卫生支出, 降低了新疆卫生总费用。此外, 城镇化率正向影响人均卫生费用, 城镇化的推进促进了农民工收入水平的提升, 同时改进了农民工的医疗条件, 使得城乡居民都更加注重医疗健康, 提升新疆卫生总费用。

综上所述, 影响新疆卫生总费用的因素并非是单一的, 新疆卫生总费用的上涨是多个因素共同作用的结果。

4.1. 城镇化率对卫生总费用的影响

早有研究显示, 城乡居民的医疗服务利用情况存在严重不平等现象, 其中农村居民处于不利地位[10][11]。因此, 人口的城乡结构转变即城镇化, 将使卫生费用面临上升的压力[12]。纵观地区发展具体情况, 中国在过去几十年中经历了快速城镇化, 且一直存在较为明显的城乡二元结构, 但由于新疆地处西部偏远地区城镇化速度相较于内地仍较缓慢, 在城乡之间依旧存在医疗服务不平等现象。

从本文回归情况来看, 新疆城镇化率对卫生费用有显著正影响。新疆城镇化与卫生费用之间有着十分密切的联系, 城镇化可以促进经济发展, 缩小城乡差距, 改善居民生活水平[13], 进而提高居民对于卫生服务利用的地理可及性和经济可及性, 而卫生服务的可及性是影响卫生服务利用的重要因素之一[14], 所以城镇化水平促进了居民对于卫生服务的利用。随着新疆城镇化进程推进, 一方面表现在医疗资源的获取更加容易; 另一方面, 城镇化率也在一定程度上反映了新疆的经济社会发展, 会使偏远地区的农村居民被抑制的部分医疗服务需求得到满足, 从而导致新疆卫生总费用的增加。

4.2. 卫生机构人员数对卫生总费用的影响

近年来, 随着医疗技术、医疗信息化的快速发展, 居民卫生服务需求更加多样化, 追求更高质量的卫生服务。我国政府借助于医疗信息化平台, 一方面为居民提供了优质的医疗卫生资源配置服务, 另一方面也无疑是增加了卫生服务成本, 成为卫生总费用的增长的原因之一[15]。而在西部卫生资源配置的各指标里最重要的就是医疗机构人员数。新疆卫生机构人员可能会直接影响卫生服务消费者的卫生服务需求, 进而影响新疆卫生总费用。

从本文的回归结果也可发现, 新疆卫生机构人员数对新疆卫生总费用影响有显著正影响。新疆医疗

卫生机构人员数增加,对该地的卫生总费用起正向作用。新疆地处西部边远地区,大量高水平专业机构人员仍旧处于紧缺状态。因此近年来新疆相关部门大力重视医疗人才的培养与引进积极依托现有的医疗援疆政策,提升各地州医疗卫生水平,优化医疗卫生人才队伍的规模与结构[16],这也对该地卫生总费用的增加有所影响。

在相关性分析中同样得到,新疆卫生资源配置的相关指标与新疆卫生总费用相关性较大,说明在新疆卫生资源供给较为缺乏时,虽然资源供给水平较低,但是居民对医疗服务的需要并不低,一旦有资源供给,同样会产生更大的医疗费用。

4.3. 人均次住院费用对卫生总费用的影响

本文回归结果表明,人次均住院费用的数目与卫生总费用呈现正相关关系,且新疆人均住院费用 X_t 每增长 1 单位,新疆卫生总费用 Y 增加 0.142 亿元。然而人均住院费用也受到多方面因素影响,人均住院费用构成有卫生材料费、其他费用、检查费、手术费、护理费、治疗费、床位费、药费,其中床位费和药费等。其中在住院费用中,体现医务人员劳务价值的手术费、治疗费占比近些年来不断上升。

由此可见住院病人均次费用的变动同样受诸多因素的长期影响。因此筹资和控费都需综合长期考虑。从长期看,继续推行药品零加成政策依旧十分必要,除了推行这一政策外,医院还需要规范医生诊疗行为、建立完善的监管体系,政府也需要加大卫生投入力度,健全按病种分值付费机制,逐步建立病种组合的标准体系,实现医院运行模式由粗放式管理向精细化管理的转变,降低不合理不必要的费用[17],此外医院本身要加强精细化管理和成本控制,推进现代化医院管理体系建设,利用医疗机构医用耗材使用数据进行监督并有效约束,充分利用新疆医疗监管的大数据平台,持续聚焦相关问题线索,加强对检查检验、卫生材料项目的监管[14]。

面对新医改带来的新形势、新要求,政府、卫生、医保、物价等部门应加强积极协同,高效推进医疗服务价格与医保支付方式改革[18]。只有通过全方面多层次的医政协同,才能从根本上帮助新疆居民减轻住院费用负担。

参考文献

- [1] 张加奇,李玲,毛宇凡,等. 新征程中我国卫生总费用核算的新特点,新问题及对策思考[J]. 卫生经济研究, 2023, 40(6): 7-9.
- [2] 刘伟,武瑞仙,许宪春.“健康中国”规模指标的关联研究[J]. 统计与信息论坛, 2022, 37(5): 2-11.
- [3] 程晓明,罗五金,刘国祥. 卫生经济学[M]. 北京:人民卫生出版社, 2012: 117.
- [4] 洪媛媛. 广东省卫生总费用影响因素主成分回归分析[J]. 卫生软科学, 2015, 29(12): 756-759.
- [5] 《中国卫生统计年鉴》编辑委员会. 中国卫生统计年鉴[M]. 北京:中国协和医科大学出版社, 2011.
- [6] 吴士勇. 中国卫生健康统计年鉴[M]. 北京:中国协和医科大学出版社, 2018-2022.
- [7] 新疆维吾尔自治区统计局. 新疆统计年鉴: 2012~2022 [M]. 北京:中国统计出版社, 2010-2022.
- [8] 任雪松,于秀林. 多元统计分析[M]. 北京:中国统计出版社, 2010.
- [9] 高晓红,李兴奇. 多元线性回归模型中无量纲化方法比较[J]. 统计与决策, 2022, 38(6): 5-9.
- [10] 林相森,艾春荣. 对中国医疗服务利用不平等问题的实证检验[J]. 中国人口科学, 2009(3): 86-95, 112.
- [11] 熊跃根,黄静. 我国城乡医疗服务利用的不平等研究——一项于 CHARLS 数据的实证分析[J]. 人口学刊, 2016(6): 62-76.
- [12] 徐长生,张泽栋. 城镇化、老龄化及经济发展对我国医疗费用影响回归分析[J]. 中国卫生经济, 2015(6): 54-55.
- [13] 周亮,车磊,孙东琪. 中国城镇化与经济增长的耦合协调发展及影响因素[J]. 经济地理, 2019, 39(6): 97-107.
- [14] 杨彤玲,汤少梁. 药改新政背景下公立医院医疗费用结构变动研究[J]. 卫生经济研究, 2023, 40(11): 57-60.

- [15] 刘丽丽, 周恭伟. 我国卫生总费用现状分析及影响因素研究——基于主成分回归分析法[J]. 卫生软科学, 2022, 36(8): 45-49.
- [16] 伏润得, 张新焕, 杨德刚. 新疆医疗卫生资源配置空间公平分析[J]. 中国科学院大学学报, 2023, 40(1): 69-78.
- [17] 吕国营, 周万里, 王超群. 人口老龄化、临近死亡时间与医疗费用支出——基于中国老年人健康影响因素跟踪调查的实证分析[J]. 中国卫生政策研究, 2020, 13(5): 1-9.
- [18] 陆婷婷, 谭坤. 新医改对四川省某公立医院医疗费用结构的影响[J]. 中国病案, 2023, 24(11): 41-44.