Linear Regression Model is Used to Analyze Regional Development Differences and Influencing Factors in China's Insurance Market

Zitong Yin*

Actuarial Science Xi'an Jiaotong-Liverpool University Suzhou, Jiangsu, China 215123
* Corresponding author: 1306442735@qq.com

Abstract

Insurance is one of the three drivers of the financial industry and an essential part of the national economy. Its growth and expansion are of great significance to the healthy development of the national economy. Since the reform and opening-up, China's insurance industry has developed rapidly. Global premium growth is expected to stabilize at 4.4 percent over the next 10 years, while growth in the Chinese market is expected to average 9.5 percent[1]. However, unbalanced regional development has appeared in China's insurance market. Therefore, this paper will study the regional differences in China's insurance market growth and its influencing factors.

At first, per capita disposable income, per capita pension of enterprises, per capita savings deposit, juvenile and child dependency ratio, elderly population dependency ratio, and the proportion of the tertiary industry were selected as explanatory variables to study the factors affecting premium income. The least-square method is used to build a regression model to calculate its correlation.

The data are the cross-sectional data of provincial administrative units in China from 2018 to 2020. After the model establishment and test, the conclusion is drawn that per capita disposable income, per capita pension and per capita savings are the important factors that lead to the unbalanced development of regional premium income.

Keywords-premium income; personal insurance; per capita disposable income; enterprise retirement; per capita pension; per capita savings deposit

1. INTRODUCTION

Since the reform and opening up, the living standard of the Chinese people has been gradually improved, and people's awareness of insurance has been enhanced. On this basis, China's insurance industry has developed rapidly and become one of the fastest developing industries in China. At the same time, the regional differences in the development of China's insurance industry have been gradually reflected. This paper will summarize the current situation of regional differences in China and find the influencing factors. As insurance is divided into personal insurance and property insurance, this paper chooses personal insurance as the research object. In studying the factors influencing the regional development of the personal insurance market, the sample data set is the cross-sectional data of provincial administrative units throughout the country from 2018 to 2020[2]. The influencing factors are respectively: per capita disposable income, per capita pension of enterprises, per capita savings deposit, juvenile and child dependency ratio, elderly population dependency ratio, and the proportion of the tertiary industry. The explanatory variable is the per capita premium income of personal insurance. Since the data of the dependency ratio of minors and the dependency ratio of the elderly population were not found until 2020, the data of 2019 were used. This paper will be helpful for individual insurance companies to formulate development plans in different regions of China according to local conditions and help insurance practitioners choose suitable employment areas, which is conducive to the stable development of China's insurance industry.

2. REGIONAL DIFFERENCES IN THE INSURANCE MARKET

In 2020, the insurance industry achieved a cumulative original premium income of 4.53 trillion yuan, up 6.12% year on year, and a year-on-year growth of 12.17% in the

same period of 2019. In terms of personal insurance, China's original insurance premium income reached 3.33 trillion yuan in 2020, up 7.53% year on year. From the property insurance business perspective, the accumulative actual insurance premium income will reach 1.19 trillion yuan in 2020, with a year-on-year growth of 2.4%[3]. According to the per capita income of personal insurance premiums, the regions can be divided into three levels.

The first level includes Beijing, Shanghai, Tianjin, Jiangsu, Guangdong, Zhejiang, Shandong. In 2020, the sum of the total premium income of these regions accounted for 46.15% of the total national premium income, the personal insurance premium income accounted for 46.88% of the national personal insurance premium income, and the property insurance premium income accounted for 48.26% of the national property insurance premium income[4]. These regions have a significant market share and high per capita premium, so they belong to the developed areas of the insurance market. The number of state-owned insurance companies accounted for 81.59 percent of the total[5]. The number of foreign insurance companies accounted for 93.62 percent of the total[6], far ahead of other regions.

The second level includes Henan, Sichuan, Hebei, Hubei, Hunan, Liaoning, Fujian, Shaanxi, Chongqing, Heilongjiang, Shanxi, Inner Mongolia, Jilin, Xinjiang, Ningxia. In these regions, there is a gap between per capita premium income and level 1, but the level of insurance depth and the first par, especially in Heilongjiang province is the country's insurance depth areas[7], growth is relatively stable, Belong to insurance market medium developed area.

The third level includes Anhui, Jiangxi, Gansu, Hainan, Qinghai, Guangxi, Yunnan, Guizhou, and Tibet[8]. Per capita premium income of these regions is relatively low, but the growth rate is fast. The premium growth rate of Jiangxi, Guangxi, and Gansu has ranked among the top five in terms of the premium growth rate in 2020[8]. Thus, it can be seen the market potential is relatively large.

3. INFLUENCING FACTORS OF PREMIUM INCOME

3.1. Data collection

The factors affecting premium income may be the level of economic development, the level of social security, the financial assets of residents, the age structure of the population, and the level of economic growth. Therefore, per capita disposable income ($x1$/yuan), enterprise per capita pension ($x2$/yuan), per capita savings deposit ($x3$/ten thousand yuan), juvenile and child dependency ratio ($x4$/%), the elderly population dependency ratio ($x5$/%), proportion of the tertiary industry ($x6$/%), and per capita income of personal insurance premium (y/yuan) are selected as reference data[9]. This paper collected cross-sectional data of provincial administrative units across the country from 2018 to 2020. Data are from the China Statistical Yearbook[2], in which per capita personal insurance premium income = total personal insurance premium income of each region/number of permanent residents.

3.1.1. Establishment of the model

3.1.1.1. Correlation between data

For intuitive purposes, scatter plots of each explanatory variable and the explained variable are successively made below.
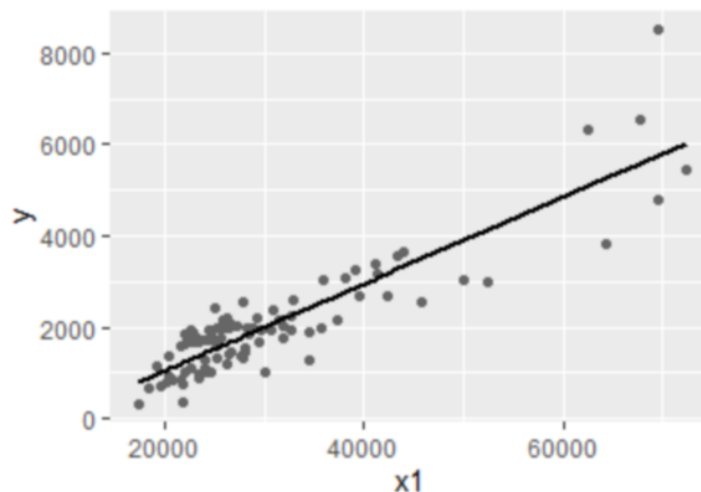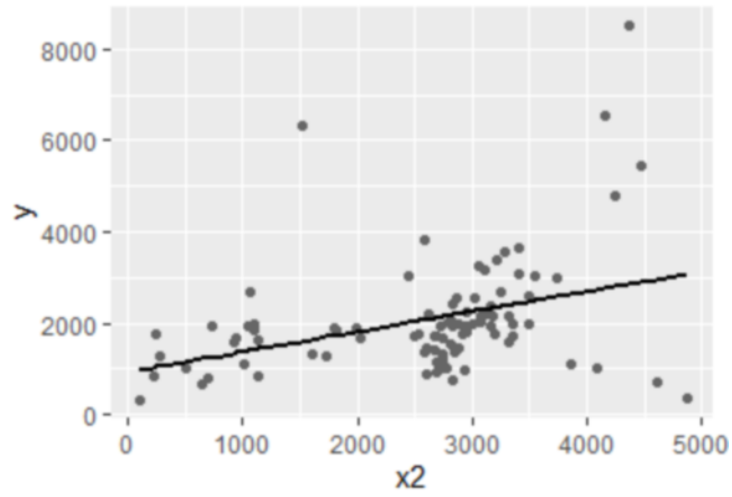

Figure1 Scatterplot of variables x1 and y

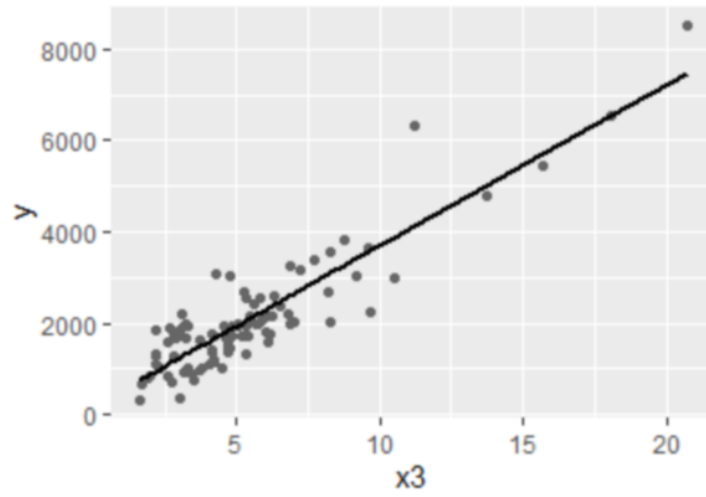**Figure2** Scatterplot of variables x2 and y

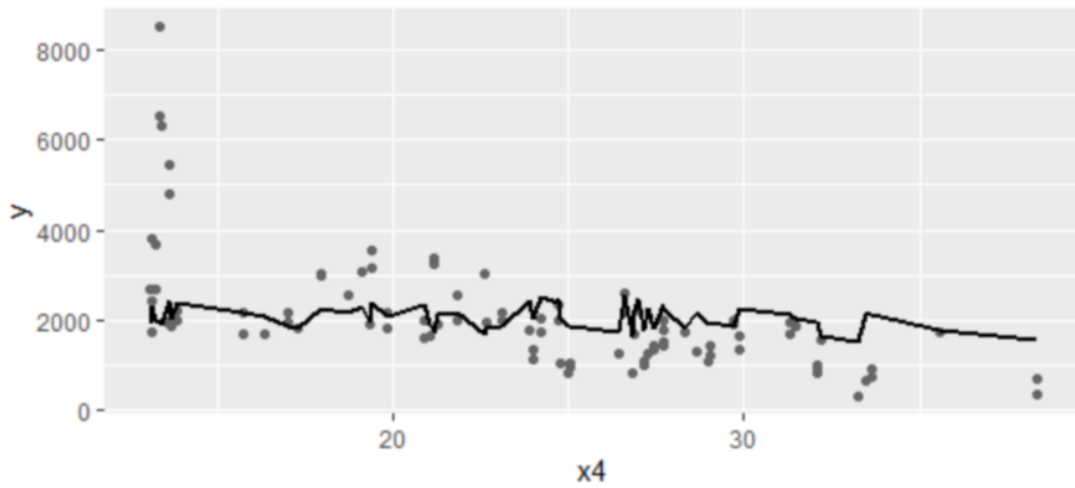**Figure3** Scatterplot of variables x3 and y

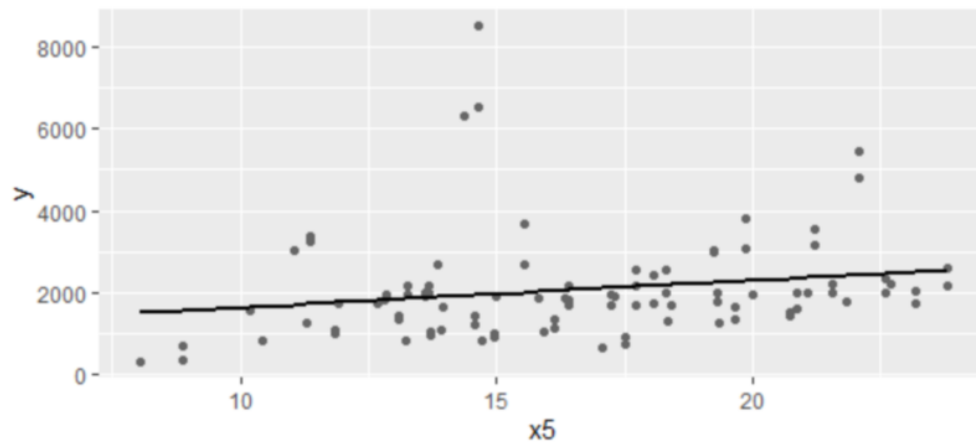
Figure4 Scatterplot of variables x4 and y

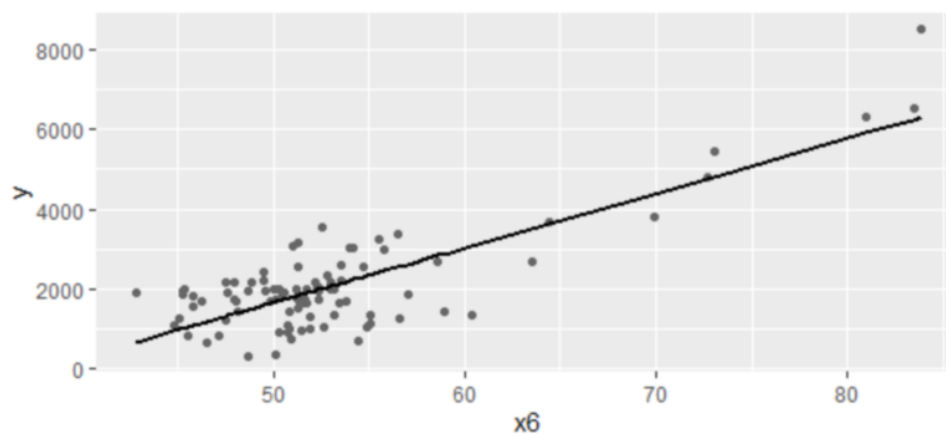**Figure5** Scatterplot of variables x5 and y

Figure6 Scatterplot of variables x6 and y

It can be seen roughly from the above figure that x1, x2, x3, x5, and x6 have significant effects on y.

3.1.1.2. The least-square regression model was preliminarily established

```
Call:
lm(formula = y ~ x1 + x2 + x3 + x5 + x6, data = X0713)

Residuals:
    Min      1Q  Median      3Q     Max
-981.85 -266.00  -33.84  267.06 1233.79

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept) -7.455e+02  5.840e+02  -1.276  0.20519
x1           4.047e-02  8.196e-03   4.938 3.78e-06 ***
x2          -1.660e-01  5.780e-02  -2.872  0.00513 **
x3           2.273e+02  3.664e+01   6.203 1.81e-08 ***
x5           9.708e-01  1.305e+01   0.074  0.94089
x6           1.446e+01  1.178e+01   1.227  0.22306
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 453.7 on 87 degrees of freedom
Multiple R-squared:  0.8814,     Adjusted R-squared:  0.8746
F-statistic: 129.3 on 5 and 87 DF,  p-value: < 2.2e-16
```
Figure7 Y fitting results of multiple linear regression with 6 explanatory variables

The above data are calculated by R studio[10].

3.1.1.3. *Preliminary statistical test*

The goodness of fit test:

From the results of regression estimation, the goodness of fit $R^2=0.87461$, This indicates a high degree of a good fit, indicating that 87.46% of the change in per capita personal insurance premium (y/yuan) can be explained by the difference in per capita disposable income (x1/yuan), per capita enterprise pension (x2/yuan), per capita savings deposit (x3/ten thousand yuan), the dependency ratio of the elderly population (x5/%), and the proportion of the tertiary industry (x6/%)

From the perspective of P-value:

Set $H_0 : \beta_0 = \beta_1 = \beta_2 = \beta_3 = \beta_5 = \beta_6 = 0$ $H_0 : \beta_j \neq 0$, j=0,1,2,3,5,6

The p values of per capita disposable income (x1/yuan), per capita enterprise pension (x2/yuan), per capita savings deposit (x3/ten thousand yuan) are all < 0.05, then reject the null hypothesis that H_0, x1, x2, and x3 have significant linear effects on y joint.

3.1.2. Model modification

```
Call:
lm(formula = y ~ x1 + x2 + x3, data = X0713)

Residuals:
     Min       1Q   Median       3Q      Max
-1094.54  -273.89   -45.61   246.82  1270.94

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept) -1.344e+02  1.837e+02  -0.732  0.46634
x1           4.341e-02  7.792e-03   5.572 2.66e-07 ***
x2          -1.773e-01  5.663e-02  -3.131  0.00236 **
x3           2.478e+02  3.241e+01   7.644 2.32e-11 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Residual standard error: 452.6 on 89 degrees of freedom
Multiple R-squared:  0.8793,    Adjusted R-squared:  0.8752
F-statistic: 216.2 on 3 and 89 DF,  p-value: < 2.2e-16
```

Figure8 Y fitting results of multiple linear regression with 3 explanatory variables

Y = 134.3576 0.0434 x1 + 0.1773 x2 + 247.7726 x3

The above data are calculated by R studio.

3.1.3. Statistical test of the modified model

3.1.3.1. *The goodness of fit test:*

From the results of regression estimation, the goodness of fit $R^2=0.8752$, closer to 1, indicating higher goodness of fit, indicating that 87.52% of the change in premium income y can be explained by per capita disposable income (x1 / yuan), per capita enterprise pension (x2 / yuan), per capita savings deposit (x3 / ten thousand yuan)

3.1.3.2. *From the perspective of P-value:*

Set H_0:$\beta_0 = \beta_1 = \beta_2 = \beta_3 = 0$ H_0:$\beta_j \neq 0$, j=0,1,2,3,

Because the p values of the three explanatory variables are all <0.05, the null hypothesis H_0 is rejected, indicating that the combined linear effect of each explanatory variable on Y is significant.

In summary, the model passed the statistical test at a significance level of 5%.

3.1.4. Research conclusion

3.1.4.1. *Model explanation*

According to the above conjecture and verification, the final model can be concluded as follows:

y = 134.3576 0.0434 x1 + 0.1773 x2 + 247.7726 x3

In the setting of the model, due to the existence of multicollinearity, we eliminated the influence of three factors, namely, the juvenile and child dependency ratio (x4 /%), the elderly population dependency ratio (x5 /%), and the proportion of the tertiary industry (x6 /%), through P-value. However, it is undeniable that these factors directly influence the per capita personal insurance premium income.

According to our final model, the per capita savings deposit as a measure of residents' financial assets significantly impacts the per capita personal insurance premium income. Every increase of 10,000yuan per capita

savings deposit will increase the per capita personal insurance premium income of 247.7726 yuan.

Per capita disposable income and prices per capita pension back as economic development level and the level of social security measures, the personal insurance premium per capita income have a positive influence and negative influence respectively, every increase 1 yuan per capita disposable income, per capita personal insurance premium income increased 0.0434 yuan, companies per capita pension back every increase 1 yuan, The per capita personal insurance premium income decreased by 0.1773 yuan.

3.1.4.2. *Model application*

As an essential tool of risk transfer, insurance can obtain the most outstanding security at the lowest economic cost as far as possible. Insurance is a stabilizer for the country and society and can guarantee the regular progress of social production. The insurance industry is also a social and economic booster that can promote economic development and foreign trade. For individuals, insurance can ensure their timely recovery and transfer of risk when they suffer losses, which is conducive to the excellent result of individuals.

According to the model, to develop the insurance industry well, the government can vigorously develop the economy, create more jobs, relieve the pressure of unemployment, and thus raise the level of per capita disposable income.

For insurance companies facing the continuous improvement of social security levels, they should formulate corresponding countermeasures. They should actively develop new insurance products based on improving service quality, doing an excellent job in after-sales service, and maintaining the original customer group. It is suggested to first develop markets in regions with high per capita disposable income and savings deposits for emerging insurance companies.

4. CONCLUSION

This paper first divides China's insurance market into three classes. The analysis of premium income and premium income growth data summarizes the current situation of unbalanced regional development of China's insurance industry. Then, based on the cross-sectional data of provincial administrative units in China from 2018 to 2020, an econometric model was established to take personal insurance as an example to discuss the influencing factors that lead to different premium income: per capita disposable income/yuan (x1/yuan), per capita enterprise pension/yuan (x2/yuan) and per capita savings deposit/yuan (x3/ten thousand yuan). The influencing factors of premium income are not only affected by economic development indicators. The improvement of population quality, insurance awareness, and risk

occurrence probability, such as the increased incidence of cancer, natural disasters, the outbreak of novel coronavirus, and other factors, will affect the premium income. Because this paper only chose personal insurance as the research object, the lack of research on the influencing factors of property insurance, so there are certain limitations. Further studies can adopt the same research method to select the factors that affect the income of property insurance to establish the measurement model for research. The factors that affect the income of personal insurance and property insurance are summarized as the influencing factors of the total premium income.

REFERENCES

[1] Muyun Sun (2021) Deep reading! According to the analysis of the current situation and development trend of the global insurance industry in 2020, there is a big gap between China's insurance industry and the world [Online].

[2] National Bureau of Statistics of China (2020) China Statistical Yearbook [Online]. Chinese netizens (2021) The insurance industry's original premium income in 2020 will be 4.53 trillion yuan, up 6.12% year on year [Online]. Available from:

[3] No author (2021) Ranking of premium income scale of each province in 2020 [Online].

[4] Liyun Li & Yinan Chen(2021) Latest statistics: There are 239 insurance companies in China, accounting for more than 70% in Beijing, Shanghai and Guangzhou (list attached) [Online].

[5] Fishermen's pride477 (2019) List of domestic foreign-funded insurance companies [Online].

[6] Protect the world with wisdom（2021）Premiums of 36 provinces and cities are all ranked. Heilongjiang's insurance depth exceeds Beijing's, and there are even these value depressions [Online].

[7] Wei Huang (2006) Regional differences in the development of insurance industry deserve attention [Online].

[8] Qiujun Lu & Xiquan Shi (2008) Study on regional difference of insurance demand in China [Online].

[9] Robert I. Kabacoff (2011) R in Action: Data Analysis and Graphics with R.Manning Publications Co.