Academia Open
Vol. 10 No. 2 (2025): December
DOI: 10.21070/acopen.10.2025.12422

Statistical Modeling of Regional Insurance Market Dynamics in Uzbekistan: Model Statistik Dinamika Pasar Asuransi Regional di Uzbekistan

Bekberganova Mokhira Department of Econometrics and Economic Modelling
National University of Uzbekistan

General Background: The insurance sector significantly supports economic growth, requiring accurate statistical modeling. **Specific Background:** In Uzbekistan, despite rapid market expansion, regional disparities and weak analytical tools limit development. **Knowledge Gap:** Previous global studies have applied advanced models, but integrated statistical and econometric analysis of Uzbekistan's insurance market remains scarce. **Aims:** This research analyzes regional insurance premiums, tests data distribution, and builds econometric forecasting models. **Results:** Shapiro-Wilk tests confirmed non-normality, leading to Gamma distribution modeling, while region-specific ARIMA(p,d,q) models were validated through AIC, AICc, BIC, and log likelihood. Regression, factor, and cluster analyses highlighted structural differences between regions. **Novelty:** The study combines distribution testing, time-series forecasting, and multivariate analysis to create region-specific models, accounting for market shifts in areas such as Karakalpakstan and Tashkent. **Implications:** Findings provide a basis for optimizing insurance services, improving tariff policies, and supporting inclusive and sustainable regional insurance development in Uzbekistan.
Highlight :

1. Statistical analysis identifies non-normal distribution in insurance premiums.
2. ARIMA models provide accurate regional forecasting of market trends.
3. Results guide policy and support sustainable insurance development.

Keywords : Insurance Market, Uzbekistan, ARIMA Model, Gamma Distribution, Shapiro-Wilk Test

Introduction

In the countries of the world the development of the insurance market and the growth rate of the national economy helps speed up. Such as USA, Germany, France, Japan, China Insurance payments in the G20 countries account for 4% of the country's GDP each yearis 17.4 percent. "The

world average in this field is 7.3 equal to percent"1. This, in turn, "causes a sharp increase in insurance premiums due to this, in 2021 it increased by 3.3% and totaled 6.9 trillion dollars [1]. Insurance premium growth varies in different parts of the world. in particular, 6.3% in China, 1.7% in the USA, 2.8% in Western European countries and it was 5.6% in developing countries"2. of the world insurance market

positive trends in the number of clients and activities of insurance companies in connection with the expansion of its geography, it develops new products new to improve quality of output and customer service includes the use of technologies. This is the conclusion It can be concluded that the management of the world insurance market is a continuous improvement and requires adaptation to changing conditions, which provides quality to customers provides service and competitiveness of companies increases [2][3].

It is important to understand and forecast the risks and insurance claims in the market when making effective decisions and developing insurance tariffs. to date, various statistical analyzes and models aimed at solving various problems in the insurance market have been used in scientific research. For example, one of the traditional methods, generalized linear models (GLM) helps insurers to optimize prices, manage risks, retain customers, and maintain their position in the market. In addition, time series models can help predict future trends by identifying patterns in previous periods [4][5]. Decision tree, stochastic and copula models are used to determine uncertainties and relationships between different risk indicators [6][7]. Machine learning algorithms are used to analyze large volumes of data and identify hidden patterns [8]. The use of Markov models and panel data models to model the dynamics of insurance policy holders over time and to analyze several indicators over a long period of time allows further strengthening of the analysis [9].

The analysis of the existing methodology for assessing the development of the insurance market shows that it is evaluated primarily on the basis of absolute and relative indicators in the assessment and research at the regional level.

The most important absolute indicators:

1. The number of registered and operating insurance companies;

2. Insurance payments (for all industries and types);

3. Insurance compensation (for all industries).

By solving the following main problems in the classical commercial insurance work process;

a. to find (correct) the ratio between payment and reward;

b. calculating the probability of a crash to make the right decision;

c. calculation of risk probability;

d. calculation of the cross section of insurance companies.

This field is known as actuarial mathematics it is characterized as the use of mathematical modeling, mathematical statistics and probability theory methods.The important features and problems of the insurance market in the process of marketing research of insurance services are as follows:

a. determining the term of the life insurance product, determining the validity period of the insurance contract, the real value of which is measured only when the insurance event occurs

(extending the term to obtain reliable information);

b. priority impact on customers of refund and compensation as well as extension of insurance period.

c. non-separability of insurance services from insurance loss insurers,

d. regulation of the activities of insurance companies by state and local bodies;

e. failure of most consumers to be provided with information about various types and rare types of insurance;

f. a large number of clients of insurance companies for all types of insurance (a large amount of research);

g. the need to use statistical analysis methods [10].

The main potential of the insurance market is of great importance in assessing and analyzing the market. Because in practice, the entire population can be potential customers and the size of the analyzed sample can be large enough, so it is important to widely use various methods of insurance market analysis.

In this case, one of the important tasks to facilitate the work of analysts is to systematize existing programs and adapt them to the market of insurance services [11][12].

The research of marketing mechanisms is carried out using two interrelated parts, one of which is the study of external variables, which is not controlled by the management of insurance companies, because the company has elastic flexibility to operate effectively in the market of insurance services. should be. The second is the analysis of the situation within the organization and the impact of changes in the environment on the enterprise.

The results of the conducted marketing research allow to develop a strategic plan of insurance companies.

Analytical data collected in the process of marketing research can be widely used in the methodological complex.

The following quantitative methods of statistical analysis are available:

1. multifactor methods, used in the analysis of dependence on multiple changes;

2. the statistical method of decision-making (the theory of public service, etc.) is used to describe the probability of the effect of changes in market conditions on consumers;

3. use of the determinant method in process research (linear and non-linear programming). It is used to make optimal decisions in the case of interdependent changes;

4. simulation model and method. These elements are used in determining the marketing situation, in cases where analytical solutions are not provided [13].

5. the network planning and management model is used to sequentially determine the solution to various marketing problems.

Among the most commonly used quantitative methods in the analysis of the insurance market, we recommend the following:

- correlational and regression analysis;

- discriminant analysis;

- factor analysis;

- cluster analysis;

- multi-edge scaling.

The most common among quantitative methods in marketing research of the insurance market; trend extrapolation and correlation-regression analysis [14][15].

Methodology

This research employed a combination of quantitative statistical methods and econometric modeling to evaluate the development trends and optimization possibilities in Uzbekistan's regional insurance markets. Data spanning a 21-year period were collected and analyzed to assess the distributional characteristics of insurance premium indicators. Initially, normality of data was tested using the Shapiro-Wilk test, with results indicating a non-normal distribution across variables. Consequently, gamma distribution assumptions were applied to model claim amounts, which are positive continuous variables. To determine the predictive capacity of selected variables and to forecast regional insurance premium growth, autoregressive integrated moving average (ARIMA) models were employed. Region-specific ARIMA(p,d,q) models were estimated, adjusting parameters based on the time series characteristics of each locality. The quality of the forecasting models was assessed using log-likelihood values, Akaike Information Criterion (AIC), corrected AIC (AICc), and Bayesian Information Criterion (BIC), allowing for the selection of the most statistically appropriate model for each region. Additionally, correlation-regression analysis was used to examine the relationships between insurance premiums and regional development indicators. Complementary to this, discriminant, factor, and cluster analyses were utilized to categorize and interpret the structural differences across regional insurance behaviors. The methodology integrates elements of actuarial mathematics, probability theory, and statistical modeling to ensure accuracy in projecting future market dynamics and optimizing insurance services. This comprehensive analytical approach allowed for the identification of key regional disparities, estimation of future insurance market capacity, and generation of strategic recommendations for inclusive insurance development across Uzbekistan [16].

Results and Discussion

A. Problem Formulation

To forecast the insurance market in the regions of the Republic of Uzbekistan in the future, ARIMA(p,d,q) and trend models were used. In this study, the development of the insurance market in the regions is expressed by the insurance premium.

ARIMA(0,2,0) math appearance :

This model is only second level to differentiate based on . Time row general mathematician equation as follows will be :

Academia Open
Vol. 10 No. 2 (2025): December
DOI: 10.21070/acopen.10.2025.12422

$$\nabla^2 y_t = y_t - 2y_{t-1} + y_{t-2} + e_t \quad (1)$$

Here
y_t: observed value at time t (e.g., YIM)
y_{t-1}: the observed value at time t-1.
y_{t-2}: the observed value at time t-2.
∇^2
: second level differential operator.
e_t: residual (random error or noise)

Figure 1.

In the model autoregressive (p=0) and moving average (q=0) components no . Model only time row two times differentiation through trend take throws and the rest random noise as This is math . equation with forecasts future y_t It allows us to calculate the values based on previous observations .

B. Solution of the Problem

The following tests are used to determine the distribution of selected statistical data:

1. Kolmogorov-Smirnov test.

2. Shapiro-Wilk test .

3. Anderson-Darling test.

The Kolmogorov-Smirnov test is considered to be less powerful than other tests . The Anderson-Darling test is used to test whether multiple sample sets can be modeled from a single parent set. It does not require a distribution function to be specified. The Shapiro-Wilk test is a common test of normality in statistics .

There is statistics distribution determination for From the Shapiro-Wilk test used . This of the test zero l hypothesis from that is that the population is assumed to be normally distributed is done . So so , if the p (probability) value chosen α (significance) level less if , then null the hypothesis is rejected and selection the set is normally distributed has from the main set that was that it is not means . Other from the side , if p value chosen α from the level bigger if so , then zero l hypothesis (that the data are normally distributed) has denying that For example , α = 0.05 level for p value less than 0.05 when selection the set is normally distributed from the head set that about n u l l rejects the hypothesis and on the contrary , at α =0.05 p value greater than 0.05 when selection set to the normally distributed prime set relevant about the fact that zero l We cannot reject the hypothesis [17].

Research for taken selection the sets Shapiro-Wilk normality test to a normal distribution using has or has that it is not We check . In this case, 2 H_(0) and H_1a hypothesis is put forward.

H_0: the sample set is drawn from a normally distributed parent set ;

H_1: the sample set is taken from a non-normally distributed population . We check whether each of the selected indicators has a normal distribution based on 21 years of data (Table 1).

No.	Indicator	W	p-value
1.	Insurance premium	0.74245	0.0001347

Table 1. *Shapiro-Wilk Test of Insurance Premium Inspection Results*

Academia Open
Vol. 10 No. 2 (2025): December
DOI: 10.21070/acopen.10.2025.12422

The results obtained show that none of the indicators in the sample set have a normal distribution. Because all of them have a p-value less than 0.05 at the α = 0.05 significance level. Therefore, the null hypothesis (H_(0)) that the sample sets are drawn from a normally distributed parent set H_(0)is rejected. The alternative hypothesis (H_(1)) that the sample sets are drawn from a non-normally distributed parent set H_1is accepted.

In order to determine the distribution law of factors, additional Gamma tests were performed. According to the results of the analysis, it was found that the above factors obey the Gamma distribution law.

In order to pay its customers, an insurance company must be able to calculate the amount of insurance coverage. Therefore, the insurance company must be able to estimate the number of claims and the amount of the claim. The number of claims or claim frequency is a measure of how often claims occur. Claims are a measure of the amount or severity of an insured event and its scope. The count data is assumed to be the number of claims, while the amount of the claim is continuous data. Since the value is always positive, the amount of the claim is assumed to have a Gamma distribution (Figure 1,2).

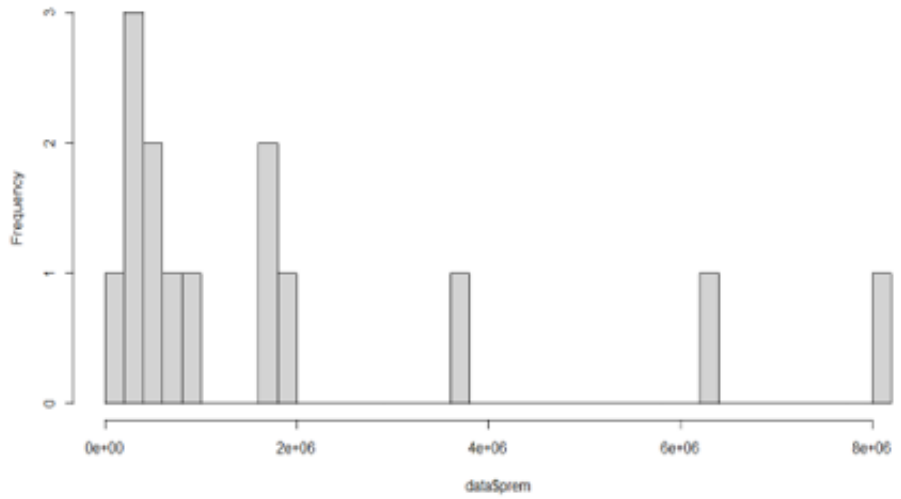

Figure 2. *Histogramm of Insurance Premium*

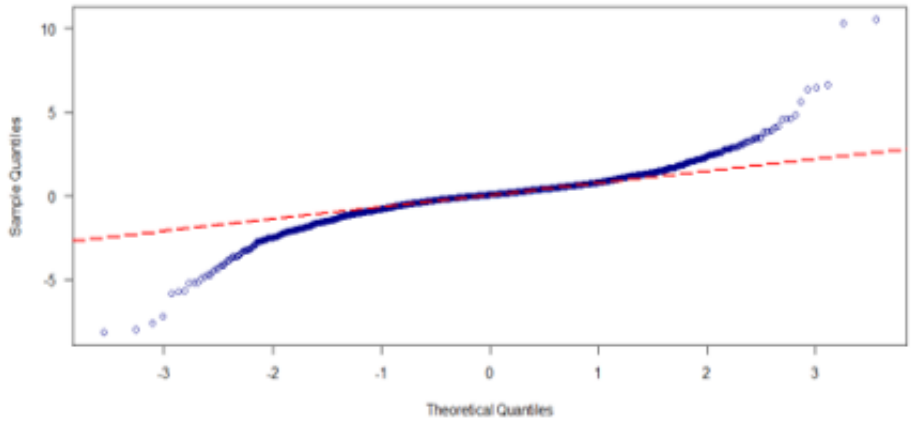

Figure 3. *Q-Q Plot of Insueance Premium Data*

Academia Open
Vol. 10 No. 2 (2025): December
DOI: 10.21070/acopen.10.2025.12422

When forecasting insurance premiums in the regions using the ARIMA(p,d,q) model, p represents the insurance premium in the past periods (autoregressive order), the changed value of the insurance premium in the regions (difference of the insurance premium or first difference order), and q represents the average rolling order. Since the trends in the development of the insurance market in the regions differ from each other, ARIMA(p,d,q) models of various forms were used. In addition, econometric models were used that take into account the structural shift in the development of the insurance market in the Republic of Karakalpakstan and Tashkent region.

The quality of the models developed for forecasting the insurance market in the regions was assessed using the log likelihood, Akaika (AIC) and Bayesian (BIC) criteria. These criteria were used to select models with the lowest error in forecasting insurance premiums in the regions (Table 2).

No.	Area	Model	log likelihood	AIC	AICc	BIC
1	Republic of Karakalpakstan	ARIMA(0,1,0)	-275,85	553,7	553,89	554,83
2	Andijan	ARIMA(0,2,2)	-241,09	488,19	489,52	491,46
3	Bukhara	ARIMA(0,2,0)	-245,42	492,85	493,05	493,94
4	Jizzakh	ARIMA(0,2,0)	-231,15	464,29	464,49	465,38
5	Kashkadarya	ARIMA(0,2,1)	-242,75	489,49	490,12	491,68
6	Navoi	ARIMA(0,1,1)	-252,05	508,1	508,7	510,37
7	Namangan	ARIMA(0,1,1)	-254,54	515,07	516,34	518,48
8	Samarkand	ARIMA(0,2,1)	-232,71	469,41	470,04	471,59
9	Surkhandarya	ARIMA(1,2,1)	-237,75	481,5	482,84	484,78
10	Syrdarya	ARIMA(0,2,2)	-232,54	471,09	472,42	474,54
11	Tashkent and	ARIMA(0,2,2)	-243,29	492,58	493,92	495,86
12	Fergana	ARIMA(0,2,1)	-247,5	499,01	499,64	501,19
13	Khorezm	ARIMA(0,2,0)	-242,44	486,88	487,08	487,98
14	Tashkent city	ARIMA(0,2,0)	-258.23	518.47	518.7	519.41

Table 2. *Statistical Criteria for Evaluating the Quality of Econometric Models Used in Regional Insurance Market Forecasting*

Conclusion

In conclusion, about 30 percent of insurance premiums are in Tashkent collected. This indicator is increasing dynamically, except for 2020, because at that time quarantine measures were introduced due to the COVID-19 pandemic was The rest of the regions received very little insurance premiums and this situation is similar to other economic indicators of regional development incomparable, especially for insurance premiums In Tashkent region, whose share in GDP is only 0.18 percent clearly visible.

The current stage of the development of the insurance market of Uzbekistan is population payment limited capacity, instability of insurance companies, sufficient lack of professional personnel, weak financial resources of companies characterized by technical and information base. The conducted research proved that joint stock companies are not only is the most effective form of organizing insurance and other commercial enterprises, but also to the socio-economic development of the country and its regions, the population it is also positive for raising living standards and encouraging innovation affects. Using chatbots as an innovative technology is insurance contributes to the growth and development of the portfolio, as well as to customers to offer new products that suit their needs and preferences gives.

Insurance market research, statistical analysis and modeling using insurance data help to find optimal solutions in the process of setting tariff rates and choosing a marketing policy in the insurance sector.

References

1. M. Goldburg, A. Khare, D. Tevet, and D. Guller, Generalized Linear Models for Insurance Rating, 2nd ed., Arlington, VA, USA: Casualty Actuarial Society, CAS Monograph Series No. 5, 2019.
2. R. R. Cerchiara, M. Edwards, and A. Gambini, “Generalized linear models in life insurance: Decrements and risk factor analysis under Solvency II,” in Proc. 18th Int. AFIR Colloquium, Rome, Italy, 2008.
3. L. Lenten, “A time-series analysis of the demand for life insurance companies in Australia: An unobserved components approach,” Australian Journal of Management, vol. 31, no. 2, pp. 159–184, Jul. 2006, doi: 10.1177/031289620603100201.
4. X. He and H. Zhang, “The trend analysis of life insurance premium income time series based on SARIMA model in underdeveloped areas,” Advances in Social Science, Education and Humanities Research (ASSEHR), vol. 156, Proc. 2nd Int. Seminar on Education Innovation and Economic Management (SEIEM), pp. 252–256, 2017, doi: 10.2991/seiem-17.2018.52.
5. R. Keng, H. Lian, Y. Wu, and H. C. Koh, “Decision tree analysis of terminated life insurance policies,” Journal of Actuarial Practice, vol. 11, pp. 229–249, 2004.
6. A. Omerasevic and J. Selimovic, “Classification ratemaking using decision tree in the insurance market of Bosnia and Herzegovina,” South East European Journal of Economics and Business, vol. 15, no. 2, pp. 124–139, 2020, doi: 10.2478/jeb-2020-0014.
7. Y. Chi, S. Jaimungal, and X. S. Lin, “An insurance risk model with stochastic volatility,” Insurance: Mathematics and Economics, vol. 46, no. 1, pp. 52–66, Feb. 2010, doi: 10.1016/j.insmatheco.2009.09.004.
8. P. R. Giaccotto, “Stochastic loss reserving in non-life insurance,” Lecture presented at KU Leuven Actuarial Contact Program (ACP), Leuven, Belgium, May 14, 2002.
9. J. M. Hendrickson, E. B. Mcellin, and A. M. Yambao, “Stochastic models in health insurance: A case study,” presented at the New Orleans Health/Pension Spring Meeting, New Orleans, LA, USA, Jun. 15–17, 2005.
10. T. Tsai and C. L. Lo, “Modeling underwriting risk: A copula regression analysis on U.S. property-casualty insurance byline loss ratios,” Pacific-Basin Finance Journal, vol. 83, Feb. 2024, Art. no. 102089, doi: 10.1016/j.pacfin.2023.102089.
11. M. Tian, F. Guo, and R. Niu, “Risk spillover analysis of China’s financial sectors based on a new GARCH copula quantile regression model,” The North American Journal of Economics and Finance, vol. 63, Nov. 2022, Art. no. 101849, doi: 10.1016/j.najef.2022.101849.
12. M.-P. Côté, C. Genest, and M. Omelka, “Rank-based inference tools for copula regression, with property and casualty insurance applications,” Insurance: Mathematics and Economics, vol. 89, pp. 1–15, Nov. 2019, doi: 10.1016/j.insmatheco.2019.07.001.
13. M. K. Severino and Y. Peng, “Machine learning algorithms for fraud prediction in property insurance: Empirical evidence using real-world microdata,” Machine Learning with Applications, vol. 5, Sep. 2021, Art. no. 100059, doi: 10.1016/j.mlwa.2021.100059.
14. U. Orji and E. Ukwandu, “Machine learning for an explainable cost prediction of medical insurance,” Machine Learning with Applications, vol. 15, Mar. 2024, Art. no. 100529, doi: 10.1016/j.mlwa.2023.100529.
15. Y. Shi, A. Punzo, H. Otneim, and A. Maruotti, “Hidden semi-Markov models for rainfall-related insurance claims,” Insurance: Mathematics and Economics, vol. 120, pp. 91–106, Jan. 2025, doi: 10.1016/j.insmatheco.2024.11.001.
16. G. D. Kaltchev, “Dynamic panel models with directors’ and officers’ liability insurance data,” in Contributions to Economic Analysis, vol. 274, Amsterdam, Netherlands: Elsevier, 2006, ch. 14, pp. 351–360, doi: 10.1016/S0573-8555(06)74014-0.
17. C.-C. Lee and Y.-B. Chiu, “The impact of real income on insurance premiums: Evidence from panel data,” International Review of Economics & Finance, vol. 21, no. 1, pp. 246–260, Jan.

2012, doi: 10.1016/j.iref.2011.06.002.