Exploring Pricing Strategies using Regime Switching Stochastic Models and Machine Learning Techniques for Insurance Contracts

MOHAMMAD SARFRAZ

DEPARTMENT OF MATHEMATICS
INDIAN INSTITUTE OF TECHNOLOGY DELHI

FEBRUARY 2025

Exploring Pricing Strategies using Regime Switching Stochastic Models and Machine Learning Techniques for Insurance Contracts

by

MOHAMMAD SARFRAZ

Department of Mathematics

submitted
in fulfillment of the requirements of the degree of
Doctor of Philosophy
to the

Indian Institute of Technology Delhi

FEBRUARY 2025

Dedicated To My Parents

Certificate

This is to certify that the thesis entitled **Exploring Pricing Strategies using Regime Switching Stochastic Models and Machine Learning Techniques for Insurance Contracts** submitted by **Mr. MOHAMMAD SARFRAZ** to the **Indian Institute of Technology Delhi**, for the award of the degree of **Doctor of Philosophy**, is a record of the original bonafide research work carried out by him under my supervision and guidance. The thesis has reached the standards fulfilling the requirements of the regulations relating to the degree. The results contained in this thesis have not been submitted in part or full to any other university or institute for the award of any degree or diploma.

New Delhi
February 2025

Prof. S. Dharmaraja
(Thesis Supervisor)
Department of Mathematics
Indian Institute of Technology Delhi

Acknowledgments

This thesis marks the culmination of a significant journey in my academic and personal life, and it would not have been possible without the support, guidance, and encouragement of many individuals to whom I am profoundly grateful.

First and foremost, I would like to express my deepest gratitude to my supervisor, Prof. S. Dharmaraja, for their invaluable guidance, constructive feedback, and unwavering support throughout the course of my PhD. Their expertise and encouragement have been instrumental in shaping this research and inspiring my growth as a researcher.

I extend my heartfelt gratitude to my Student Research Committee (SRC) members—Prof. Aparna Mehra and Prof. Vikas Vikram Singh from the Department of Mathematics, IIT Delhi, and Prof. Smita Kashiramka from the Department of Management Studies, IIT Delhi—for their valuable insights and critical feedback, which significantly enhanced the quality of this research. I am also deeply appreciative of the faculty and staff of the Department of Mathematics, IIT Delhi, for fostering a supportive and conducive environment. My sincere thanks go to the department for providing essential resources, including access to financial databases, which greatly facilitated the completion of this work.

I would like to express my sincere gratitude to Prof. Viswanathan Arunachalam (National University of Colombia, Bogota, Colombia) and Prof. Guglielmo D'Amico (University G. d'Annunzio of Chieti-Pescara, Pescara, Italy) for their invaluable collaboration and insightful discussions, which greatly contributed to this research. I am also deeply thankful to Prof. N. Selvaraju (IIT Guwahati) for providing me with the opportunity to visit IIT Guwahati and work under his guidance.

I would like to take this opportunity to express my heartfelt gratitude to the amazing individuals I had the privilege of meeting during my Ph.D. journey, who made my time at IIT Delhi both enriching and enjoyable.

I am deeply thankful to my friends, Dr. Nitin Kumar, Dr. Shakti Singh, and Dr. R. Kavin for their invaluable support and camaraderie. My sincere gratitude also goes to my friends Priya, Samiksha, and Utpal, as well as my senior colleagues, Dr. Nitu Sharma, Dr. Raina Raj and Dr. Anisha, for their friendly assistance and encouragement throughout this journey.

I am deeply grateful to my dear friends, Dr. Ishtiaq Ahmed, Dr. J. I. Khaki, and Dr. I. Ahmed Shauq from JNU New Delhi, for their unwavering motivation and support during challenging times. I also thanks to my dear friend Mr. Yasser Bashir from Civil Engineering Department of IIT Delhi, for his moral support and helpful discussions during my research.

I am profoundly grateful to my family for their unwavering love, support, and patience. To my beloved mother, Shamim Akhter, for her endless care and encouragement, and to my late father, Mohammad Khaliq, whose values of perseverance and dedication continue to guide me every day. I also extend my heartfelt thanks to my siblings and extended family for their constant motivation and belief in me throughout this journey. A special thanks to Zakiya and Mua'z, whose support and understanding have been my anchor throughout this demanding journey.

I would like to express my heartfelt gratitude to my younger brother, Mohammad Sajjad, for his unwavering support throughout my PhD journey. His thoughtful gestures, especially preparing delicious meals that kept me energized during long working hours, meant so much to me. Beyond that, his willingness to shoulder other responsibilities at home provided me with the peace of mind and focus needed to complete this thesis. I am deeply thankful for his care, patience, and encouragement, which made this challenging journey much smoother.

Lastly, I acknowledge IIT Delhi for the financial support in the form of fellowship and Department of Science and Technology (DST) Delhi for research grant to carry out this work.

Above all, I express my deepest gratitude to the Almighty for His blessings, guidance, and strength throughout this journey.

This work is dedicated to everyone who believed in me and contributed to this journey in ways big and small.

December, 2024

New Delhi MOHAMMAD SARFRAZ

Abstract

This thesis presents a comprehensive study on the pricing of insurance contracts, including variable annuities with embedded guarantees and participating life insurance contracts such as pure endowment options, whole life deferred annuities, and guarantee annuity options. A particular emphasis is placed on integrating regime switching dynamics into various models. Additionally, the thesis explores the application of machine learning techniques to enhance the valuation of these contracts. Variable annuities are unit-linked managed funds that provide several guarantees to policyholders including death benefit protection, living benefit protection, and lifetime income protection.

Similarly participating life insurance contracts are a key product in the life insurance sector, characterized by a minimum interest rate guarantee and participation in the insurer's annual asset portfolio returns. These contracts often include additional features such as bonus participation rules based on the insurer's surplus reserves and terminal bonus payments. Moreover, participating contracts encompass products like pure endowment options, guaranteed annuity options, and deferred whole life annuities. The interest rate guarantee may apply either only at maturity or accrue incrementally throughout the contract's term, offering policyholders both guaranteed benefits and exposure to investment returns.

This thesis addresses the pricing challenges and computational efficiency of variable annuities and liabilities of participating contracts, such as deferred whole life annuities and pure endowment options. It focuses on key market risks, including interest rate risk, investment risk, mortality risk, and volatility risk. To address these challenges, the thesis introduces various regime switching stochastic models designed to effectively capture these risks.

A continuous-time model is proposed that integrates regime switching dynamics into the local volatility framework, providing a robust approach to modeling volatility risk and developing a comprehensive pricing model for variable annuities. The calibration and valuation of the model are conducted using Monte Carlo simulations and machine learning techniques, such as Gaussian process regression. The practical implications of the proposed models are demonstrated through an empirical test on historical S&P 500 Index data, showcasing their effectiveness in addressing real-world pricing scenarios.

A key challenge in the pricing of insurance contracts, lies in the computational intensity of traditional numerical methods such as Monte Carlo simulations and partial differential equations. Moreover, the local volatility model, while effective in capturing implied volatility, fails to account for stochastic volatility dynamics observed in real markets, particularly during periods of market uncertainty.

To address these limitations, machine learning techniques, particularly Gaussian process regression and neural networks, have been implemented for efficient pricing. A regime-switching Heston model is proposed, which incorporates stochastic volatility and regime changes to enhance both the accuracy and computational efficiency of the valuation process. This thesis is the first to explore the potential of machine learning techniques for pricing variable annuity products with embedded guaranteed minimum benefits under the Markov regime switching Heston stochastic volatility model framework.

The majority of existing literature on the valuation of financial assets using regime switching models assumes that model parameters are governed by an observable Markov chain. However, relatively little attention has been devoted to the use of regime switching hidden Markov models in this domain. This thesis bridges this gap by implementing regime switching hidden Markov models for the pricing of variable annuity guarantees, offering a more robust framework for identifying and capturing hidden market regimes. Furthermore, the thesis examines the impact of the COVID-19 pandemic on variable annuity pricing using Indian stock market data, specifically the Nifty50 index data. The analysis divides the data into pre- and post-COVID-19 periods to investigate the influence of market dynamics during and after the pandemic.

From the literature, it is evident that, alongside volatility and investment risks, interest rate risk and biometric risk significantly influence the price of insurance contracts. Biometric risk, encompassing uncertainties related to mortality and longevity, directly impacts the expected payouts of life insurance and annuity products. To address these risks, in the end, this thesis proposed a regime switching stochastic interest rate model that extends the contingent claim-based framework for participating contracts, such as endowment policies and deferred life-long annuities. By integrating regime switching dynamics, the model evaluates the combined effects of interest rate fluctuations and biometric

uncertainties on the market value of life insurance liabilities. Numerical results and sensitivity analysis highlight the critical role of these risks, providing comprehensive insights into liability valuations across varying economic and demographic conditions.

This thesis contributes to the existing literature by introducing various models that integrate regime switching stochastic processes and application of machine learning techniques for the pricing of variable annuities and participating life insurance contracts. It offers valuable insights into capturing evolving market conditions and risks, thereby enhancing the resilience and sustainability of financial products. Furthermore, this thesis establishes a foundation for advancing model calibration methodologies, improving risk management strategies, and developing insurance products that are better aligned with the complexities of dynamic market environments.

सारांश

यह शोध प्रबंध बीमा अनुबंधों की मूल्य निर्धारण प्रक्रिया का एक व्यापक अध्ययन प्रस्तुत करता है, जिसमें अंतर्निहित गारंटियों वाले परिवर्तनीय वार्षिकी (Variable Annuities) और भागीदारी जीवन बीमा अनुबंध (Participating Life Insurance Contracts) जैसे कि शुद्ध एंडॉमेंट विकल्प (Pure Endowment Options), संपूर्ण जीवन स्थगित वार्षिकियां (Whole Life Deferred Annuities) और गारंटी वार्षिकी विकल्प (Guaranteed Annuity Options) शामिल हैं। विशेष रूप से, विभिन्न मॉडलों में शासन-स्विचिंग (Regime Switching) गतिशीलता को एकीकृत करने पर जोर दिया गया है। इसके अतिरिक्त, इस शोध प्रबंध में इन अनुबंधों के मूल्यांकन को बेहतर बनाने के लिए मशीन लर्निंग तकनीकों के अनुप्रयोग का भी विश्लेषण किया गया है।

परिवर्तनीय वार्षिकियां यूनिट-लिंक्ड प्रबंधित फंड होते हैं, जो पॉलिसीधारकों को कई गारंटियां प्रदान करते हैं, जिनमें मृत्यु लाभ सुरक्षा (Death Benefit Protection), जीवन सुरक्षा लाभ (Living Benefit Protection) और आजीवन आय सुरक्षा (Lifetime Income Protection) शामिल हैं। इन गारंटियों को अक्सर GMxB के रूप में संदर्भित किया जाता है, जहां x लाभ के प्रकार को इंगित करता है। सामान्य उदाहरणों में गारंटीकृत न्यूनतम मृत्यु लाभ (GMDB), गारंटीकृत न्यूनतम संचय लाभ (GMAB), गारंटीकृत न्यूनतम निकासी लाभ (GMWB), और गारंटीकृत न्यूनतम आय लाभ (GMIB) शामिल हैं।

इसी प्रकार, भागीदारी जीवन बीमा अनुबंध (Participating Life Insurance Contracts) जीवन बीमा क्षेत्र में एक प्रमुख उत्पाद हैं, जो न्यूनतम ब्याज दर गारंटी और बीमाकर्ता की वार्षिक परिसंपत्ति पोर्टफोलियो रिटर्न में भागीदारी की विशेषता रखते हैं। इन अनुबंधों में आमतौर पर बोनस भागीदारी नियम (Bonus Participation Rules) और टर्मिनल बोनस भुगतान जैसी अतिरिक्त विशेषताएँ शामिल होती हैं। भागीदारी अनुबंधों में शुद्ध एंडॉमेंट विकल्प (Pure Endowment Options), गारंटीकृत वार्षिकी विकल्प (Guaranteed Annuity Options) और स्थगित संपूर्ण जीवन वार्षिकियां (Deferred Whole Life Annuities) शामिल होते हैं। ब्याज दर गारंटी या तो केवल परिपक्वता पर लागू हो सकती है या अनुबंध की अवधि के दौरान धीरे-धीरे संचित हो सकती है, जिससे पॉलिसीधारकों को गारंटीकृत लाभ और निवेश रिटर्न दोनों का लाभ मिलता है।

जीवन बीमा योजनाएं, विशेष रूप से भागीदारी अनुबंध, मृत्यु लाभ प्रदान करने के साथ-साथ बीमाकर्ता के प्रदर्शन से जुड़े संभावित प्रोत्साहन भी प्रदान करते हैं। हालांकि, इन अनुबंधों में ब्याज दर जोखिम (Interest Rate Risk), निवेश जोखिम (Investment Risk) और दीर्घायु जोखिम (Longevity Risk) जैसे अतिरिक्त जोखिम भी शामिल होते हैं। जीवन बीमा कंपनियों के पोर्टफोलियो में ब्याज दरों में उतार-चढ़ाव के कारण महत्वपूर्ण जोखिम उत्पन्न होता है, जो परिसंपत्तियों की वापसी और दायित्वों के मूल्यांकन दोनों को प्रभावित करता है।

यह शोध प्रबंध परिवर्तनीय वार्षिकियों और भागीदारी अनुबंधों की देनदारियों (Liabilities) के मूल्य निर्धारण और कम्प्यूटेशनल दक्षता से संबंधित चुनौतियों को संबोधित करता है, जिसमें स्थगित संपूर्ण जीवन वार्षिकियां और शुद्ध एंडॉमेंट विकल्प शामिल हैं। इसमें ब्याज दर जोखिम, निवेश जोखिम, मृत्यु दर जोखिम (Mortality Risk) और अस्थिरता जोखिम (Volatility Risk) जैसे प्रमुख बाजार जोखिमों पर ध्यान केंद्रित किया गया है। इन चुनौतियों से निपटने के लिए, यह शोध प्रबंध विभिन्न शासन-स्विचिंग यादृच्छिक मॉडल (Regime-Switching Stochastic Models) प्रस्तुत करता है, जो इन जोखिमों को प्रभावी ढंग से कैप्चर करने के लिए डिज़ाइन किए गए हैं।

एक सतत-समय मॉडल (Continuous-Time Model) प्रस्तावित किया गया है, जो शासन-स्विचिंग गतिशीलता को स्थानीय अस्थिरता रूपरेखा (Local Volatility Framework) में एकीकृत करता है। यह अस्थिरता जोखिम को मॉडल करने और परिवर्तनीय वार्षिकियों के लिए एक व्यापक मूल्य निर्धारण मॉडल विकसित करने के लिए एक सशक्त दृष्टिकोण प्रदान करता है। मॉडल का अंशांकन (Calibration) और मूल्यांकन मोंटे कार्लो सिमुलेशन (Monte Carlo Simulations) और मशीन लर्निंग तकनीकों जैसे कि गॉसियन प्रोसेस रिग्रेशन (Gaussian Process Regression) का उपयोग करके किया गया है। प्रस्तावित मॉडलों के व्यावहारिक प्रभावों को ऐतिहासिक S\&P 500 इंडेक्स डेटा पर एक अनुभवजन्य परीक्षण के माध्यम से प्रदर्शित किया गया है, जिससे वास्तविक दुनिया की मूल्य निर्धारण स्थितियों को संबोधित करने में उनकी प्रभावशीलता स्पष्ट होती है।

बीमा अनुबंधों के मूल्य निर्धारण में एक प्रमुख चुनौती पारंपरिक संख्यात्मक विधियों जैसे मोंटे कार्लो सिमुलेशन और आंशिक अंतर समीकरणों (Partial Differential Equations - PDEs) की कम्प्यूटेशनल तीव्रता है। इसके अतिरिक्त, स्थानीय अस्थिरता मॉडल निहित अस्थिरता को प्रभावी ढंग से कैप्चर करने में सक्षम होते हुए भी वास्तविक बाजारों में देखी गई यादृच्छिक अस्थिरता गतिशीलता (Stochastic Volatility Dynamics) को पकड़ने में विफल रहता है, विशेष रूप से बाजार अनिश्चितता की अवधि के दौरान।

इन सीमाओं को दूर करने के लिए, मशीन लर्निंग तकनीकों, विशेष रूप से गॉसियन प्रोसेस रिग्रेशन और न्यूरल नेटवर्क (Neural Networks) को कुशल मूल्य निर्धारण के लिए लागू किया गया है। एक शासन-स्विचिंग हेस्टन मॉडल (Regime-Switching Heston Model) प्रस्तावित किया गया है, जो यादृच्छिक अस्थिरता और शासन परिवर्तनों को एकीकृत करता है, जिससे मूल्यांकन प्रक्रिया की सटीकता और कम्प्यूटेशनल दक्षता दोनों में सुधार होता है। यह शोध प्रबंध मार्कोव शासन-स्विचिंग हेस्टन यादृच्छिक अस्थिरता मॉडल (Markov Regime-Switching Heston Stochastic Volatility Model) के तहत न्यूनतम गारंटियों वाले परिवर्तनीय वार्षिकी उत्पादों के मूल्य निर्धारण के लिए मशीन लर्निंग तकनीकों की संभावनाओं का पता लगाने वाला पहला अध्ययन है।

अधिकांश मौजूदा साहित्य शासन-स्विचिंग मॉडल का उपयोग करते हुए वित्तीय परिसंपत्तियों के मूल्यांकन में यह मानता है कि मॉडल पैरामीटर एक प्रेक्षणीय (Observable) मार्कोव श्रृंखला द्वारा संचालित होते हैं। हालाँकि, शासन-स्विचिंग छिपे हुए मार्कोव मॉडल (Regime-Switching Hidden Markov Models) के उपयोग पर अपेक्षाकृत कम ध्यान दिया गया है। यह शोध प्रबंध इस अंतर को पाटने के लिए शासन-स्विचिंग छिपे हुए मार्कोव मॉडल को परिवर्तनीय वार्षिकी गारंटियों के मूल्य निर्धारण में लागू करता है, जिससे छिपे हुए बाजार शासनों की पहचान और कैप्चर करने के लिए एक अधिक मजबूत रूपरेखा प्रस्तुत की जाती है।

इसके अतिरिक्त, यह शोध प्रबंध भारतीय स्टॉक मार्केट डेटा, विशेष रूप से Nifty50 इंडेक्स डेटा का उपयोग करके COVID-19 महामारी के परिवर्तनीय वार्षिकी मूल्य निर्धारण पर प्रभाव का विश्लेषण करता है। इस विश्लेषण में डेटा को पूर्व- और पश्चात- COVID-19 अवधियों में विभाजित किया गया है, ताकि महामारी के दौरान और बाद में बाजार गतिशीलता के प्रभावों की जांच की जा सके।

यह शोध प्रबंध शासन-स्विचिंग यादृच्छिक प्रक्रियाओं और मशीन लर्निंग तकनीकों के एकीकरण द्वारा मूल्य निर्धारण मॉडल को उन्नत करने में योगदान देता है। इसके निष्कर्ष वित्तीय उत्पादों की स्थिरता बढ़ाने, मॉडल अंशांकन पद्धतियों में सुधार, और जीवन बीमा एवं वार्षिकी अनुबंधों के जोखिम प्रबंधन रणनीतियों को परिष्कृत करने में सहायक हैं।

Contents

List of Figures

List of Tables

List of Symbols

Symbol	Meaning
y_0	Initial insured age [i.e., Insured age at period 0]
$r(t)$	Rate of interest at time t
$V(t)$	Volatility at time t
$F(t)$	Value of fund at time t
$F(0)$	Initial value of fund
$A(t)$	Account value at time t
$A(0)$	Initial account value
$S(t)$	Value of stock price at time t year
$S(0)$	Initial Value of stock price
g	Fee charged by the insurance company
$G(t)$	Guaranteed amount at time t
h_0	Single premium payment
τ	The future lifetime of an individual
c	Roll-up rate in %
r	Constant rate of interest
τ_d	The death time of an individual
${}_tp_{y_0}$	The probability of survival for t years for an individual currently aged y_0
${}_tq_{y_0}$	The probability of death for t years for an individual currently aged y_0
X_n	Discrete time Markov chain
$W(t)$	Standard Brownian motion
$X(t)$	Continuous time Markov chain

Glossary of Terms

Abbreviation	Definition
ADAM	Adaptive Moment Estimation – An optimization algorithm that adjusts learning rates based on first and second moments of gradients.
AIC	Akaike Information Criterion – A metric for model comparison, balancing goodness of fit and complexity.
BIC	Bayesian Information Criterion – Similar to AIC but introduces a stronger penalty for model complexity.
CTMC	Continuous-Time Markov Chain – A stochastic model where transitions between states occur continuously.
DB	Death Benefits – Payouts from an insurance policy or annuity upon the policyholder's death.
DTMC	Discrete-Time Markov Chain – A stochastic model where transitions between states occur at fixed intervals.
EIAs	Equity-Indexed Annuities – Annuity contracts with returns linked to an equity index.
ELAs	Equity-Linked Annuities – Annuities that provide returns based on equity market performance.
ELRA	Exponential Learning Rate Adjustment – A technique for adjusting learning rates in optimization.
EM	Expectation-Maximization – An iterative algorithm for parameter estimation in probabilistic models.

Abbreviation	Definition
EPV	Expected Present Value – The discounted value of expected future cash flows.
ESG	Environmental, Social, and Governance – A framework assessing the sustainability and ethics of financial investments.
FP	Final Payoff – The final amount received from a financial contract.
FST	Fourier Space Time-stepping – A numerical method used in solving PDEs in finance.
GB	Guaranteed Benefits – The minimum benefits assured in an annuity or insurance contract.
GBM	Geometric Brownian Motion – A stochastic model used for financial asset pricing.
GB2	Generalized Beta of the Second Kind – A probability distribution capturing heavy-tailed data.
GMAB	Guaranteed Minimum Accumulation Benefit – Ensures a minimum account value in annuities.
GMDB	Guaranteed Minimum Death Benefit – Guarantees a minimum payout upon policyholder's death.
GMIB	Guaranteed Minimum Income Benefit – Provides a minimum level of income regardless of market performance.
GMWB	Guaranteed Minimum Withdrawal Benefit – Allows annuity holders to withdraw a fixed percentage annually.
GPR	Gaussian Process Regression – A non-parametric machine learning technique for function approximation.
GPs	Gaussian Processes – A class of stochastic processes used in statistical modeling.
GMB	Guaranteed Minimum Benefits – The category of guarantees in variable annuities.
GRS	Generalized Regime Switching – A framework allowing financial parameters to change dynamically.

Abbreviation	Definition
HMMs	Hidden Markov Models – A statistical model where states evolve as a Markov process, but observations are indirect.
LIME	Local Interpretable Model-Agnostic Explanations – A method for explaining black-box machine learning models.
MAE	Mean Absolute Error – Measures the average magnitude of errors between predicted and actual values.
MaxAE	Maximum Absolute Error – The largest absolute difference between predicted and actual values.
MaxRE	Maximum Relative Error – The largest relative error between predicted and actual values.
MC	Monte Carlo – A simulation-based numerical method for pricing financial instruments.
ML	Machine Learning – A field of AI that enables computers to learn from data.
MLT	Machine Learning Techniques – Various algorithms used in predictive modeling.
MRS	Markov Regime Switching – A model allowing regime-dependent changes in financial parameters.
NN	Neural Networks – Computational models inspired by the human brain, used in deep learning.
PDE	Partial Differential Equation – Mathematical equations involving multiple variables and their partial derivatives.
pdf	Probability Density Function – A function that represents the likelihood of different outcomes in a continuous distribution.
ReLU	Rectified Linear Unit – A common activation function in neural networks.
RMSE	Root-Mean-Squared Error – A metric measuring the square root of the average squared differences between predicted and actual values.
RMSRE	Root Mean Squared Relative Error – A variation of RMSE that accounts for relative differences.

Abbreviation	Definition
SDE	Stochastic Differential Equation – A differential equation incorporating randomness, widely used in finance.
SHAP	Shapley Additive Explanations – A method based on cooperative game theory for explaining machine learning model predictions.
SMP	Semi-Markov Process – A generalization of Markov processes allowing variable holding times.
VA	Variable Annuity – A type of annuity with investment-linked returns and embedded guarantees.
VIX	Volatility Index – Measures expected stock market volatility, often called the "fear index."
XAI	eXplainable Artificial Intelligence – A set of methods that improve the interpretability of AI models.