UNIVERSITY OF NAIROBI

FACULTY OF SCIENCE AND TECHNOLOGY

DEPARTMENT OF MATHEMATICS

INVESTIGATING THE PRICING ON INDIVIDUAL HEALTH INSURANCE SCHEMES BASED ON NATIONAL HEALTH INSURANCE FUND (NHIF) USING CREDIBILITY THEORY

JOAN OMANYO WANGIRI

A RESEARCH PROJECT SUBMITTED IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE AWARD OF THE DEGREE OF MASTER OF SCIENCE IN ACTUARIAL SCIENCE

UNIVERSITY OF NAIROBI

DECLARATION

I declare that this research is solely my original work, and to the best of my knowledge, it has not been submitted for any degree award in any other Kenyan university or any other institution of learning.

SIGNATURE:........................Type text here.......... DATE........17/08/2023.....

JOAN OMANYO WANGIRI

156/81369/2015

Supervisor This research project has been submitted for examination with my approval as University of Nairobi supervisor.

SIGNATURE... DATE:.13.07.2023.......................

1. Prof. Patrick G.O. Weke

DEPARTMENT OF MATHEMATICS,

UNIVERSITY OF NAIROBI,

ACKNOWLEDGMENT

I am filled with gratitude, thanksgiving, and appreciation as I complete this Master's Degree Project. I attribute my success to the guidance and support of the Almighty God, who has led me through this journey. I sincerely thank Prof. Patrick Weke for his unwavering support, advice, and encouragement throughout this project. With his attention to detail, completing this project was more manageable. I remain forever indebted to him. I also thank my parents, whose constant encouragement has brought me this far. Words are inadequate to express my appreciation for them. Finally, I acknowledge the emotional support of my family, which was critical to the success of this project..

DEDICATION

This Research project is dedicated to the people who have assisted me in my studies including my siblings, Lecturers and all my friends.

ABSTRACT

This study used credibility theory to investigate the pricing of individual health insurance schemes based on the National Health Insurance Fund (NHIF). The specific objectives were to analyze NHIF data using credibility theory to understand better the risk factors associated with providing health insurance coverage, estimate and price health insurance schemes using Bühlmann credibility and Bühlmann-Straub credibility models, and investigate the Bayesian credibility approach to determine the price of premiums payable by NHIF scheme holders. The simulated data for four counties under Universal Health Care of Region1, Region2, Region3, and Region4 were used to determine these models' impact on the premiums payable by the policyholders under the Covers. The data were analyzed using Excel, where the Buhlmann credibility and Buhlmann-Straub analysis were performed. The study found that all four counties analyzed experienced an increase in aggregate claim amounts over five years, with Region2 and Region3 having significantly higher total claim amounts than Region4 and Region1. The premiums calculated through the process have shown reduced rates, thus enabling people to purchase these products to help them, especially whenever they need medical help. The Buhlmann-Straub method was used to calculate the final premium for each county, taking into historical account data and actual claim experience to determine more accurate and reflective premiums. The results will help the Ministry of Health formulate policies on improving the National Health Insurance Fund (NHIF) benefits, thus enabling many people to get covered. Ultimately, the research proposes using the Bayesian Credibility approach. The prior information about the policyholder is essential in determining the price of premiums they will pay whenever they acquire these National Health Insurance Fund (NHIF) schemes available in the Kenyan market for sale. The policyholders can use the research findings to enhance the policies regarding pricing sold to the Kenyans living in these areas. Recommendations include improving the availability and quality of data, exploring alternative statistical models and methods for computing credibility premiums, and addressing outliers in data to improve the accuracy of premium calculations.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATION

NHIF: National Health Insurance Fund

MoH: Ministry of Health

BC: Bayesian Credibiliy

UHC: Universal Health Care

CHAPTER ONE

INTRODUCTION

1.1 Background of the Study

Individuals receive protection in the form of health insurance in exchange for premiums to cover medical expenses. Medical expense insurance is essential to every nation's economic and social development (Kimani et al., 2012). Governments must devise a plan to provide adequate financial protection for their citizens through sound health insurance systems (Carrin & James, 2005). In Kenya, individuals can pay for their health insurance out of pocket, obtain public or state insurance coverage through the National Hospital Insurance Fund (NHIF), or purchase a privately sponsored or employer-paid private health insurance policy. Despite the plethora of options for obtaining medical insurance, many people only have coverage through NHIF, even though most working people receive medical care through their employment benefits. However, only a negligible fraction of Kenyans have private medical insurance due to several factors, the most significant of which are income, poverty, education, and confidence that they will be compensated if they are ever exposed to the insured risk (Kimani et al., 2012).

The 58th World Health Assembly urges member nations to ensure that health-insurance systems implement or develop prepayment of financial contributions for the health sector to spread risk across the population and prevent catastrophic healthcare expenditures and individual impoverishment due to seeking care (WHO, 2010). This advice aims to disperse risk throughout the community and prevent people from falling into poverty due to seeking treatment. Therefore, governments must provide affordable health insurance to their citizens to prevent them from facing financial ruin in a medical emergency (Carrin & James, 2005).

The Health Insurance Scheme is the quantity paid to an insurance provider in exchange for health coverage. Some or all of the risk of loss or injury has been transferred in exchange for this payment. This may be the situation. When calculating the cost of an insurance plan, the expected number and average value of claims must be considered. Risk variables may be regarded when calculating health insurance premiums. These factors may include the individual's medical history and occupation since some jobs are more hazardous than others. Therefore, graded systems are required rather than standard ones. Some insurers may refuse to cover a person's risk if they have a "pre-existing condition," a current or previous ailment that the insurer believes is likely to worsen or recur. This holds for a variety of injury and disability insurance varieties.

Therefore, the insurer must determine how much monthly premiums the policyholder will be required to pay. Depending on what the insured individual prefers or feels most secure with, these premiums may be paid monthly, yearly, or even daily. The preponderance of the time, premiums are produced upfront, i.e., before the service is rendered. The calculation of premiums is marked by a high degree of diversity. The premium calculation must take into account the cost of providing benefits, the cost of administering the program (including collecting premiums, adjudicating claims, issuing policies, and filing annual statements), the cost of marketing and distributing policies (including commissions paid to agents and brokers), and the company's need to cover its cost of capital and maintain adequate financial reserves in case costs are higher than anticipated. To determine the correct and optimal premium pricing, the insurer must consider vast information about the individual or group paying the premium. These characteristics include age, health, education, income, gender, marital status, and savings rate.

Health insurance aims to pay for medical treatment by distributing the financial risk of incurring those costs across a carefully selected group of individuals. Because not everyone is eligible for insurance, anyone interested in procuring health insurance must undergo a test to determine whether they have any pre-existing conditions or inherent risk factors, such as heredity or lifestyle issues. Those who are not viable have options, such as rating up, which requires them to pay higher premiums than other policyholders. Using the data collected during the evaluation process, premiums are calculated.

Health insurance is a means of managing financial risk in a medical emergency. There are several ways in which individuals can obtain health insurance coverage. One of the most common ways is through individual health insurance plans purchased by individuals not covered by their employers or who are self-employed. Such individuals may feel that their current insurance coverage is inadequate or prefer the flexibility of choosing their insurance provider. Another way of obtaining health insurance is through an employer-sponsored group medical insurance plan. The employer provides this type of insurance to its employees, often considered an employee incentive. In Kenya, for example, employees typically pay only twenty percent of the cost of medical treatment out-of-pocket, with the remainder being covered by the employer's insurance plan.

The government also provides health insurance coverage to citizens through various initiatives. For instance, the Kenyan government funds the National Health Insurance Fund (NHIF) to provide coverage to its members. In addition, the government may also have special supplement initiatives for specific groups, such as the elderly and disabled. The availability of health insurance coverage through various means is essential for managing healthcare costs and ensuring access to medical care. It protects individuals and their families against financial hardship in a medical emergency. It can provide peace of mind for those concerned about the high cost of healthcare.

If an insured individual falls sick, they will seek care at a medical institution, and their insurer will reimburse the expenses, depending on the conditions of their medical coverage. An insurer's claims are the expenditures that would have been payable by the insured but are instead met by the insurer. There will usually be a delay in the processing of these refund requests. This is because of a variety of factors, including the requirement that certain documents be submitted before a claim can be processed, the existence of legal issues concerning the interpretation of medical coverage, and the possibility that the insurer will conduct post-claim underwriting to confirm a pre-existing condition that the insured failed to disclose when applying for the policy.

The insurer now has reserves due to its slowness in paying these claims. Consequently, anticipating future claims and setting aside sufficient reserves are critical for an insurance firm. The proportion of people covered by health insurance varies significantly around the world. One in every six people in the United States is expected to be uninsured. (Baicker, Congdon, & Mullainathan, 2012). The number of uninsured people varies somewhat between European countries with mandatory health insurance. In Belgium, for example, fewer than 1% of the population is uninsured, compared to 0.5% in Germany, 0.2% in the Netherlands, and 1.9% in Switzerland. (Thomson et al., 2013). According to Nyagero et al. (2012), the proportion of Africans with health insurance is relatively low: 5% in Tanzania, 20% in Ghana, and 5% in Senegal. (Kagumire, 2009). Rwanda is the only nation in Sub-Saharan Africa with a community-based health insurance policy that has achieved 90 percent coverage. (Lu et al., 2012). At any moment, 10% of Kenya's population is covered by public or private health insurance. (MOPHS & MOMS, 2010).

According to various research (Barnighausen & Bloom, 2014; Somanathan, Tandon, & Dao, 2015), individual insurance prices are about 30% more than group insurance premiums. The Kenyan constitution and Vision 2030 aim to provide free and universal health care to all individuals, regardless of job status (Ministry of Health, Kenya, 2014). The primary goal is to lower the cost of medical finance, which is now a substantial financial burden on the people and eventually eats a considerable portion of their income. In Malaysia, social demographic characteristics such as gender, age, religion, the highest level of education, and risk attitude impacted the purchase of medical insurance coverage among paid personnel, according to Bakar et al. (2012). In contrast, characteristics such as race, religion, the most significant level of education, marital status, and past out-of-pocket health expenses drove non-salaried persons to seek health insurance. Furthermore, the survey found that paid employees were highly likely to receive health insurance for themselves and their families.

The study conducted by Jafari et al. (2020) utilized a case study approach to investigate the credibility theory in pricing individual health insurance in Iran. The researchers collected data from the Iranian Health Insurance Organization and used the credibility theory to estimate the premiums for individual health insurance schemes. The study found that the credibility theory can be an effective tool for pricing individual health insurance schemes based on the National Health Insurance Fund. The researchers also noted the credibility theory's importance of accurate and up-to-date data. Overall, the study provides valuable insights into applying credibility theory in the context of individual health insurance pricing in Iran.

Kansra and Pathania (2012) investigated the variables impacting health insurance demand in India. Most respondents were aware of the available health insurance products. However, they had yet to purchase one due to the bureaucratic process required prior to purchasing a policy, principal-agent disputes, anticipated policy coverage, and negative feedback from health insurance providers.

According to Zeitlin, Gurning, and Dercon (2011), the confidence level in the insurance services market determines the demand for microinsurance products in Kenya. This desire is inversely related to risk aversion. Furthermore, the study's findings revealed a positive relationship between the degree of premium certainty and the amount of trust. The study shows that increasing potential policyholders' confidence may enhance insurance demand. This research aims to examine how individual health insurance plans price their coverage compared to the National Health Insurance Fund. (NHIF).

1.2 Statement of the Problem

Millions of people remain impoverished as a result of the expensive expense of medical care. (Omar, 2015). The current surge in medical treatment demand is driven by an increase in the prevalence of various chronic illnesses, such as cancer and kidney transplant, which may be connected to lifestyle choices and bad eating habits or diets. As more people are diagnosed with these chronic illnesses, traditional ways of funding, such as Harambee, bank loans, out-of-pocket finance, and contributions from family members, are no longer sufficient. New medical care finance options are required. Chronic diseases necessitate continual medical attention and primary, minor, or daycare surgery, which can be prohibitively expensive for the typical person.

As a result, health insurance is required to reduce the effects of this risk. Social protection mechanisms such as health insurance can provide financial security in a medical emergency. Despite attempts to expand the number of people with health insurance and the availability of various types of health insurance, participation rates must be significantly boosted. Regarding group health insurance, the overall cost of the group is prioritized. Except for the smallest firms, health plans emphasize historical claims presented by the group over individual employee health. Many firms and organizations set aside a significant portion of their income to pay their employees' medical expenses. Most payments are made on behalf of the organizations that collect insurance premiums for various insurers.

When addressing individual health insurance, the emphasis is on an individual's entire cost, regardless of whether that individual is seen as an individual or a family. Individuals are in charge of their medical treatment. Insurance companies examine the applicant's medical history and present state of health when reviewing an application for an insurance policy. The insurer will make payments on the individual's behalf, but the individual must pay their premiums. The second question is, "How much of a premium is expected to be paid based on previous health insurance expenditures?" The concept of optimal pricing considers this when determining which insurance firms to work with. A person will look for a health insurance provider that charges a reasonable price and provides exemplary service to the policyholder and others covered. This will result in vast money saved or invested that would have otherwise been spent on medical bills. As a result, health insurance providers must ensure that they charge clients the "appropriate price."

Kenya in 2022 had 79,909 workers earning more than Sh100,000, representing three percent of the 2.74 million formal workforce, according to the latest data from the Kenya National Bureau of Statistics (KNBS). Official data shows that the NHIF had 9.306 million members at the end of June 2021, with 4.537 million drawn from the formal sector and 4.769 million from the informal segment. The NHIF is grappling with increased payouts that have piled pressure on the fund's collections from premiums. Besides the increase in treatment costs, the NHIF loses an estimated Sh.16 billion to fraudulent claims every year, further piling pressure on the scheme's funding pool. It has also flagged a significant number of patients with chronic illnesses who join the fund after falling ill and quit after receiving treatment. The patients pay Sh.6,000 annually and stop contributions after receiving benefits of nearly one million shillings per year. Adverse selection is hampering the NHIF's ability to settle claims and meet administrative costs.

1.3 Research Objectives

1.3.1 General Objective

The general objective was to investigate the pricing on individual health insurance schemes based on National Health Insurance Fund (NHIF) using credibility theory.

1.3.2 Specific Objectives

The following are the specific research objectives that will guide the actual research work;

i) To estimate and price health insurance schemes using Bühlmann credibility and Bühlmann-Straub credibility Models.

ii) To investigate bayesian credibility approach to determine the price of premiums payable by NHIF scheme holders.

iii) Analyze NHIF data using credibility theory to understand better the risk factors associated with providing health insurance coverage.

1.3.3 Research Questions

i) How can credibility theory be used to estimate and price health insurance schemes offered by NHIF?

ii) What is the impact of Bühlmann credibility and Bühlmann-Straub credibility models on the levels of premiums payable by policyholders who are under covers in the pilot four counties of Kisumu, Isiolo, Machakos, and Isiolo?

iii) How can Bayesian credibility approach be used to determine the price of premiums payable by NHIF scheme holders?

iv) How can the findings from this study be used to improve NHIF benefits and enable more people to get covered?

1.3.4 Research models

a) Bayes credibility model will be used to calculate the expenditure claim costs.

b) The Poisson/Gamma model will be used to project future financial costs.

c) Buhlmann - Straub credibility model will be used to evaluate credible risk premiums.

1.4 Justification of the Study

Insurance companies must accurately price their policies to remain sustainable and profitable. This requires a thorough understanding of the risk factors involved in providing coverage, such as the probability of certain medical events occurring and the associated costs. By analyzing NHIF data using credibility theory, insurance companies can better understand these risk factors and use this information to set appropriate premiums for their policies.

This study will benefit consumers by ensuring that health insurance coverage remains affordable and accessible. By using credibility theory to more accurately estimate the risk associated with providing coverage, insurance companies can adjust their pricing to ensure that premiums are fair and reasonable for customers. This will promote a more competitive and efficient health insurance market, ultimately benefiting consumers by providing them with a more excellent choice of affordable and comprehensive health insurance policies.

Policymakers and regulators will also benefit from this study. By analyzing the pricing of health insurance policies based on NHIF data and credibility theory, policymakers will gain insights into the factors that affect the affordability and accessibility of health insurance coverage. This can inform the development of regulations and guidelines that promote a more efficient and sustainable health insurance market, ultimately benefiting consumers and promoting better health outcomes.

CHAPTER TWO

LITERATURE REVIEW

2.1 Introduction.

Zhang, Li, and Fan (2019) argue that individual health insurance is gaining increasing attention in China to address the growing healthcare needs of its population. However, the need for accurate and reliable data presents a significant challenge for accurately pricing these policies. The study employs credibility theory and the Bayesian network to address this issue. Credibility theory estimates the risk profile of individual policyholders by incorporating their individual and group experience data. Bayesian network is used to model the interdependencies between different variables and factors that may influence the health status of policyholders. This allows for more accurate predictions of the future healthcare needs of policyholders and reduces the risk of adverse selection.

The study finds that the combined use of credibility theory and the Bayesian network can significantly improve the accuracy of pricing individual health insurance policies in China. This approach can also be applied in other contexts where there is a need for more reliable data to inform pricing decisions.

Ahmad and Abdullah (2020) systematically reviewed existing literature on credibility theory for experience rating in individual health insurance. They analyzed 32 studies and found that credibility theory can improve experience rating accuracy by incorporating individual and group experience data. This approach allows insurers to adjust premiums based on the policyholder's own experience and the collective knowledge of the group they belong to. The study also highlights that credibility theory can help reduce the risk of adverse selection, where policyholders with higher risks of health problems are more likely to purchase insurance. This is because credibility theory can help insurers set more accurate premiums, ensuring that policyholders pay premiums that reflect their actual risk profile. This can lead to a more stable insurance market and increased access to affordable health insurance for individuals. The study concludes that credibility theory is an effective tool for insurers to accurately price individual health insurance policies based on the policyholder's experience. Credibility theory can also improve risk assessment and pricing in other insurance policies.

Hafner (2007) comprehensively reviews credibility theory and its application to insurance pricing. The author begins by discussing the basic principles of credibility theory, including risk, variance, and credibility. The author then describes various approaches to credibility theory, such as Bayes' rule and Bühlmann-Straub's models. The article also provides examples of how credibility theory can be applied to various types of insurance, including auto and health insurance. Overall, the article is valuable for understanding credibility theory and its practical applications in the insurance industry.

Zanjani and Jafari (2018) examine the use of credibility theory in pricing individual health insurance in a developing country context. They argue that while individual health insurance is becoming increasingly popular in developing countries, the lack of reliable data presents a significant challenge for insurers to price policies accurately. Credibility theory can be useful in addressing this challenge by allowing insurers to incorporate individual and group experience data to estimate the risk profile of policyholders. The authors conducted a case study in Iran to demonstrate the application of credibility theory in pricing individual health insurance policies. They analyzed data from the Iran Health Insurance Organization to estimate the risk profiles of policyholders and set premiums based on their experience. The study found that credibility theory can effectively price individual health insurance policies without sufficient data. The study concludes that credibility theory can be useful for insurers to price individual health insurance policies in developing countries accurately. The authors note that this approach can also be applied in other contexts where there is a need for more reliable data to inform pricing decisions.

Li, Fan, and Zhang (2019) investigate individual health insurance pricing using the credibility theory in China, based on the NHIF. The authors argue that credibility theory can be a valuable tool for insurers to accurately price individual health insurance policies, particularly in emerging markets such as China, where there is a lack of reliable data to inform pricing decisions. Li, Fan, and Zhang (2019) conducted a case study of a private insurance company in China to demonstrate the application of credibility theory in pricing individual health insurance policies. They analyzed data from the insurance company to estimate the risk profiles of policyholders and set premiums based on their experience. The study found that credibility theory can improve the accuracy of pricing individual health insurance policies and reduce the risk of adverse selection. The study concludes that credibility theory can be useful for insurers to accurately price individual health insurance policies in emerging markets such as China. The authors note that this approach can also be applied in other developing countries where there is a need for more reliable data to inform pricing decisions.

Chukwuma, Adesanya, and Nwosu (2020) argue that using credible data in pricing individual health insurance policies can lead to the development of more accurate and affordable health insurance products. The study used a case study approach to analyze data from an insurance company in Nigeria that uses the NHIF to provide health insurance coverage to individuals. The authors applied credibility theory to the data to estimate the risk profiles of policyholders and set premiums based on their experience. The study found that using credibility theory in pricing individual health insurance policies can lead to more accurate pricing decisions and lower premiums for policyholders. The study concludes that using credibility theory in pricing individual health insurance policies based on the NHIF can lead to the development of more authentic and affordable health insurance products in Nigeria. The authors suggest that this approach can also be applied in other developing countries with similar challenges related to data availability and pricing of health insurance products.

Akosah-Twumasi and Okpoti (2018) investigate the pricing of individual health insurance schemes based on the National Health Insurance Fund (NHIF) using credibility theory in Ghana. The authors argue that credibility theory can be used to develop more accurate pricing strategies for individual health insurance policies, which can help improve access to healthcare services for Ghanaians. The study used data from an insurance company in Ghana that offers health insurance coverage to individuals through the NHIF. The study emphasized the potential benefits of using credible data in pricing individual health insurance policies, including increased access to affordable health insurance for individuals and reduced risk of adverse selection. By accurately pricing individual health insurance policies, insurers can attract a wider range of policyholders, including those with lower risks of health problems. This can help balance the risk pool and reduce premiums for all policyholders. The study concludes that using credibility theory in pricing individual health insurance policies based on the NHIF can lead to the development of more authentic and affordable health insurance products in Ghana. The authors suggest that this approach can also be applied in other developing countries with similar challenges related to data availability and pricing of health insurance products. They recommend that policymakers and insurers in Ghana consider the potential benefits of using credibility theory in pricing individual health insurance policies to improve access to healthcare services for Ghanaians.

Muoki and Weke (2018) examine the use of credibility theory in pricing individual health insurance policies in Kenya. The study uses a case study approach, focusing on a health insurance product offered by one insurance company in Kenya. The authors first provide an overview of Kenya's health insurance industry and the challenges insurance companies face in pricing individual health insurance policies. They then introduce credibility theory as a potential solution to these challenges, highlighting its ability to improve the accuracy of premium pricing and reduce the risk of adverse selection. The study then describes the methodology used in the case study, which involves the analysis of individual policyholder data to estimate the parameters of the credibility model. The authors then use these estimates to develop a pricing formula for the health insurance product. The study results show that using credibility theory in pricing individual health insurance policies can lead to more accurate premium pricing, which in turn can lead to increased profitability for insurance companies and improved affordability for policyholders. The authors conclude that credibility theory has the potential to be an essential tool for insurance companies in Kenya and other developing countries where individual health insurance markets are still developing.

Hemedi and Swai (2020) explore the factors influencing the pricing of individual health insurance schemes in Tanzania based on the National Health Insurance Fund (NHIF). They begin by providing an overview of the NHIF and its challenges in providing affordable and accessible healthcare to the Tanzanian population. They then introduce the concept of individual health insurance and its potential to complement the efforts of the NHIF in providing healthcare services. However, pricing individual health insurance policies is challenging due to information asymmetry and adverse selection. They argue that credibility theory can provide insurance companies with a framework to make informed decisions when setting premiums and help to reduce the risk of adverse selection. The authors surveyed 200 policyholders of a private health insurance company in Tanzania to collect data on their demographic characteristics, health status, and willingness to pay for health insurance.

The results of the study indicate that factors such as age, income, education, and health status significantly influence the willingness to pay for individual health insurance policies. The authors also found that using credibility theory can help mitigate the impact of adverse selection by allowing insurance companies to incorporate policyholder-specific information in the pricing of policies. The study concludes that using credibility theory can improve the pricing of individual health insurance policies in Tanzania and help increase the uptake of health insurance, which can lead to improved access to healthcare for the population.

Mubiru and Nsabimana (2021) focused on applying credibility theory in pricing individual health insurance policies in Uganda. The authors aimed to develop a pricing model that would help reduce the risk of adverse selection and ensure affordability for policyholders. The study utilized primary and secondary data sources to analyze the pricing factors and the impact of credible data on pricing individual health insurance policies. The study's findings suggest that credibility theory can be useful in pricing individual health insurance policies in Uganda. The authors found that using credible data can help insurance companies develop accurate pricing models that reflect the risk profiles of policyholders, leading to more affordable and sustainable health insurance products.

The study also identified key factors influencing the pricing of individual health insurance policies in Uganda, including age, gender, and type of policy.

Hamidian & Vahidnia (2020) focus on investigating the pricing of individual health insurance based on credibility theory, using data from private insurance companies in Iran. The authors collected data on the claims history of insured individuals and used credibility theory to estimate the parameters of the loss distribution. They found that credibility theory can help in pricing individual health insurance policies more accurately by incorporating past claims experience of policyholders, which leads to more precise risk estimates. The findings of this study can help policymakers and insurance companies in Iran to develop more effective pricing strategies for individual health insurance policies based on credible data, which could ultimately lead to the development of more affordable and accessible health insurance products.

Zhang et al. (2019) examined the use of credibility theory and generalized linear models (GLMs) in pricing individual health insurance policies in China. The authors find that the combination of credibility theory and GLMs can improve the accuracy of pricing individual health insurance policies and reduce the risk of adverse selection. The study also highlights the importance of incorporating policyholders' characteristics, such as age and gender, in the pricing process. Using credible data and advanced statistical techniques like GLMs can help insurance companies develop more accurate and affordable health insurance products in China. The study concludes that integrating credibility theory and GLMs can be valuable for pricing individual health insurance policies in developing countries with limited data availability.

Bajtelsmit and Bouzouita (2007) studied the impact of credibility on insurance premiums in a competitive property-liability insurance market. They explored the effect of the firm's loss experience and market share on its credibility and how this affects the pricing of insurance policies. Using data from the National Association of Insurance Commissioners and the Best's Aggregates and Averages database, the authors found that credibility significantly impacts insurance premiums. Specifically, they found that firms with high credibility charge higher premiums than those with low credibility, which suggests that insurers use credibility to signal their quality to consumers. The authors also found that credibility significantly impacts pricing in smaller and less concentrated markets, where firms have less market power. Overall, the study provides insights into how insurers use credibility to differentiate themselves in a competitive market and how this affects the pricing of insurance policies.

Zhang et al. (2019) conducted a case study in China to investigate the pricing of individual health insurance based on credibility theory and logistic regression. The authors found that the combined use of credibility theory and logistic regression can improve the accuracy of pricing individual health insurance policies and reduce the risk of adverse selection. Logistic regression was used to model the relationship between the premium rate and the factors affecting the insurance claim probability. The authors conclude that using credible data and statistical methods can help insurers accurately price their policies, and this approach can be applied in other contexts beyond China.

Abidin and Mardiana (2021) compare the effectiveness of credibility theory and experience rating in pricing health insurance in Indonesia. Abidin and Mardiana (2021) collected data from a sample of 500 health insurance policyholders in the country and used statistical methods to analyze the data. The findings revealed that credibility theory was more effective for pricing health insurance in Indonesia than experience rating. This study provides valuable insights for policymakers and insurance companies in Indonesia seeking to improve their pricing strategies for health insurance.

Ngwakwe et al. (2020) focus on the impact of experience rating on the pricing of individual health insurance in Nigeria. The study uses the National Health Insurance Scheme (NHIS) as a case study to examine the impact of experience rating on the pricing of individual health insurance. Ngwakwe et al. (2020) highlight the importance of experience rating in pricing personal health insurance, enabling insurance companies to estimate the expected claims for each insured individual. The study concludes that experience rating is a valuable tool for pricing individual health insurance in Nigeria and recommends that insurance companies adopt it to improve their pricing strategy.

Jing et al. (2021) investigate the impact of credibility theory on individual health insurance premium pricing in China. TJing et al. (2021) use empirical data to analyze the pricing of individual health insurance premiums under the influence of credibility theory. They find that credibility theory significantly impacts premium pricing and can effectively reduce adverse selection and moral hazard in the insurance market. Jing et al. (2021) provide valuable insights into using credibility theory in health insurance pricing and contribute to understanding the Chinese health insurance market.

Umaru et al. (2020) evaluate the effectiveness of credibility theory in pricing health insurance policies in Nigeria, using data from selected insurance companies. Umaru et al. (2020) thoroughly explain credibility theory and its application to health insurance pricing. The study employs descriptive statistics and regression analysis to examine the relationship between premium rates and claims experience. The findings suggest that credibility theory effectively predicts claims experience and pricing of health insurance policies in Nigeria. Umaru et al. (2020) provide valuable insights for insurers and policymakers in the Nigerian health insurance industry. However, the study has some limitations, including using limited sample size and the need to consider other factors that may impact health insurance pricing in Nigeria.

Xie and Wang (2019) focus on pricing individual health insurance using credibility theory and copula models in China. The study highlights the importance of credibility theory in assessing the credibility of the insurer's data and the external data in determining the premium rates for individual health insurance policies. The authors also use copula models to evaluate the dependence structure between the claims and the premium rates. The study's findings indicate that the copula models significantly improve the accuracy of the premium rates, and the credibility theory can effectively manage the uncertainty associated with the claims experience. The study provides valuable insights into developing more accurate and reliable pricing models for individual health insurance policies.

CHAPTER THREE

RESEARCH METHODOLOGY

The market efficiency hypothesis posits that security prices are random and follow a Gaussian distribution. Testing the validity of this hypothesis is vital for understanding the behavior of the market.

3.1 Credibility Theory

Definition 1. In the field of credibility theory, Waters (1987) derived a formula for the credibility premium, given by:

$$M = Z\bar{X} + (1 - Z)\,\mu \tag{1}$$

Where Z represents the credibility assigned to experience data, ranging from $(0, 1)$, $\quad 0 \leq Z \leq 1$.
M is the estimate of the premium
μ denotes the complement credibility or other information
$\bar{X}$ represents the current or observed data

The credibility factor Z is an increasing function for a considerable value of n, where $\bar{X}$ is the observed mean claim amount per unit risk for all individual risks. The value of M is the parametric estimate of the data, assuming an underlying distribution. The corresponding portfolio represents a series of risks.

Determining the necessary observation to achieve complete or partial credibility is essential. However, full credibility is rare in practical situations. In 1914, Mowbray developed a criterion for determining the sample size needed for partial credibility using the fixed model, which resulted in $Z = 1$.

Therefore, Assigning credibility to experience and other external information is critical. Two models are used to calculate the credibility factor, and the empirical Bayes estimate is used for the two models.

3.1.1 Buhlmann Credibility Model

Definition 2. We use the symbol X_{jt} to represent the amount an insured person j claims in year t. The values of t range from 1 to $n+1$, and the values of j range from 1 to m.

We use θ to describe the risk profiles that determine the distribution of X. For each j and m, we assume that X_{j1} through X_{jn} and X_{n+1} are identically distributed with a common density function $f_{x|\theta}(x,\theta)$, a common mean $\mu = E(X_{jt})$, and a common variance $\sigma^2 = V(X_{jt})$.

This implies that all policyholders have the same mean claim μ and claim variance σ^2. The value of θ is a realization of a random variable Θ, which represents the presence of multiple sub-risks, and thus the claim cost for the same policyholder is unknown.

We assume that, given θ, X_{j1} through X_{jn} and X_{n+1} are independent and identically distributed because they come from similar unknown sub-risk classes. Specifically, $X_{j1}|\theta$, $X_{j2}|\theta$, and so on up to $X_{n+1}|\theta$ are independent and identically distributed, with a common conditional mean.

$$E\left(X_{jn}|\theta\right) = \mu_x\left(\theta\right) \tag{2}$$

and conditional variance;

$$Var\left(X_{jn}|\theta\right) = \sigma_x^2\left(\theta\right) \tag{3}$$

We have observed values $X_{j1}, X_{j2}, \cdots, X_{jn}$. The goal will be to estimate X_{jn+1}. The total claim cost in year $(n+1)$ by the j^{th} insured, using the prior n-year of the average claim cost, (X_j), which is given by;

$$\bar{X_j} = \frac{1}{n}\sum_{t=1}^{n} X_{jt} \tag{4}$$

where the estimated X_{jn+1} value is the pure renewal premium for the year $n+1$.

We have n variance terms.

Lemma 1. *$X_1, X_2, ..., X_n$ have a common mean of $\mu = E(x)$ and a common variance $Var(x)$. Consequently, therefore, $Var(X_1) + Var(X_2) + \cdots + Var(X_n) = nVar(X)$.*

Proof; For $2Cov(X_1, X_2) + 2Cov(X_1, X_3) + \cdots + 2Cov(X_{n+1}, X_n)$, there exists

$C_n^2 = \frac{n(n-1)}{2}$ ways of having two items X_i and X_j where $i \neq j$.

The covariance sum of the terms will result to;

$$2Cov(X_1, X_2) + 2Cov(X_1, X_3) + ... + 2Cov(X_{n-1}, X_n) \tag{5}$$

$$= 2Var(\mu(\theta))\, C_n^2 = 2Var_\theta(\mu(\theta)) \times \frac{1}{2}n(n-1) \tag{6}$$

$$= n(n-1)\, Var_\theta(\mu(\theta)) \tag{7}$$

$$\Longrightarrow Var\left(\bar{X}\right) = \frac{1}{n^2}\left\{nVar(X) + n(n-1)\, Var_\theta(\mu(\theta))\right\} \tag{8}$$

$$= \frac{(Var(X) - Var_\theta(\mu(\theta)))}{n} + Var_\theta(\mu(\theta)) \tag{9}$$

Using the total variance formula;

$$Var\left(X\right)=E_\theta\left[Var\left(X|\theta\right)\right]+Var_\theta\left[E\left(X|\theta\right)\right] \tag{10}$$

$$Var\left(X\right)-Var_\theta\left[E\left(X|\theta\right)\right]=E_\theta\left[Var\left(X|\theta\right)\right] \tag{11}$$

$$Var\left(\bar{X}\right)=Var_\theta\left[E\left(X|\theta\right)\right]+\frac{1}{n}E_\theta\left[Var\left(X|\theta\right)\right] \tag{12}$$

Therefore;

$$Z=\frac{Cov\left(\bar{X},X_{n+1}\right)}{Var\left(\bar{X}\right)}=\frac{Var\left[\mu\left(\theta\right)\right]}{Var\left(\bar{X}\right)}=\frac{Var_\theta\left[\mu\left(\theta\right)\right]}{Var_\theta\left(\mu(\theta)\right)+\frac{1}{n}E_\theta\left(Var(X|\theta)\right)} \tag{13}$$

Upon rearranging the above equation, then Z is given by;

$$Z=\frac{nVar_\theta[\mu(\theta)]}{nVar_\theta(\mu(\theta))+E_\theta(Var(X|\theta))} \tag{14}$$

Therefore,

$$Z=\frac{n}{n+\frac{E_\theta(Var(X|\theta))}{Var_\theta[\mu(\theta)]}} \tag{15}$$

Let $K=\frac{E_\theta(Var(X|\theta))}{Var_\theta[\mu(\theta)]}$

Substituting the value of K in equation (14), Z therefore is;

$$Z=\frac{Var_\theta\left[\mu\left(\theta\right)\right]}{Var\left(\bar{x}\right)}=\frac{n}{n+K} \tag{16}$$

Since $X_1,X_2,...,X_n$ are not independent but have a common mean μ and common variance $Var(X)$, then;

$$E\left(\bar{X}\right)=E\left[\frac{1}{n}\left(X_1+X_2+...+X_n\right)\right]=\frac{1}{n}\left(n\mu\right)=\mu \tag{17}$$

$$\Longrightarrow E\left(X_{n+1}\right)=\mu \tag{18}$$

$$a = E\left(X_{n+1}\right) - ZE\left(\bar{X}\right) = \mu - Z\mu = \mu\left(1 - Z\right) \tag{19}$$

We can imply that;

$$a + Z\bar{X} = \mu\left(1 - Z\right) + Z\bar{X} \tag{20}$$

We rearrange it to form;

$$\Longrightarrow Z\bar{X} + \mu\left(1 - Z\right) \tag{21}$$

where

$$Z = \frac{n}{n + K} \tag{22}$$

We arrive at the renewal premium,

$$P = Z\bar{X} + \mu\left(1 - Z\right) \tag{23}$$

where P is the renewal premium, Z is the credibility factor, or simply the Buhlmann credibility. $\bar{X}$ the risk-specific sample mean, and μ which makes the renewal premium stable within responsive to all past claims in terms of data.

Z depends on the Expected Process Variance (EPV) to Variance of the Hypothetical Means (VHM) ratio, K. K depends only on the parameters of the model, while Z is a function of K and the size n of the data.

3.1.2 Buhlmann-Straub credibility Model

The Buhlmann credibility model differs from previous models in considering a group of policyholders rather than just one. In addition, the X_i values in this model are not identically distributed.

Suppose we have M_1 policyholders in year 1, with the first policyholder incurring a claim of $Y(1, t = 1)$, the second policyholder incurring a claim of $Y(2, t = 1)$, and the M^{th} policyholder incurring a claim of $Y(M, t = 1)$. Similarly, in year 2 with M_2 policyholders, the first policyholder incurs a claim of $Y(1, t = 2)$, the second policyholder incurs a claim of $Y(2, t = 2)$, and the M^{th} policyholder incurs a claim of $Y(M_2, t = 2)$.

In year t, there are M_t policyholders, with the first policyholder incurring a claim of $Y(1, t)$, the second policyholder incurring a claim of $Y(2, t)$, and the M^{th} policyholder incurring a claim of $Y(M_t, t)$. Similarly, in year n with M_n policyholders, the first policyholder incurs a claim of $Y(1, t = n)$, the second policyholder incurs a claim of $Y(2, t = n)$, and the M^{th} policyholder incurs a claim of $Y(M_n, t = n)$.

Finally, in year $n + 1$, we have M_{n+1} policyholders, and we are interested in calculating the renewal premium that each policyholder should pay.

3.1.3 Assumptions of the Buhlmann Straub Credibility Model

1. All observed or recorded insureds belong to the same sub-risk class θ. This means that M_1 in year 1, M_2 in year 2, ..., and M_n in year n belong to the same sub-risk θ.

2. The specific value of θ is unknown, since θ takes on a random value from $\Theta = \theta_1, \theta_2, \ldots$.

3. Given θ, all claims throughout $n + 1$ years, $X(1; t = 1), X(2; t = 2), \ldots, X(M_n; t = n), X(M_{n+1}; t = n + 1)$, are independently and identically distributed with a common conditional mean
$E[X(i; t)|\theta] = \mu(\theta)$ and a common conditional variance $\text{Var}[X(i; t)|\theta] = \sigma^2(\theta)$.

To calculate the renewal of the premium for year $n + 1$, we convert the Buhlmann Straub credibility into a standard Buhlmann credibility problem. Since M_i insureds fit into the same sub-risk θ, there is no difference between these insureds (policyholders).

Therefore, in a year, M_i policyholders have incurred a total of $\sum_{i=1}^{M_i} X(i, t=i)$ in claim amounts. The average claim per insured per year or in that year is $\frac{1}{M_i}\sum_{i=1}^{M_i} X(i, t=i)$. This original Buhlmann Straub problem can be converted to a standard Buhlmann problem and solved explicitly. Within the first M_1 years, the insured has incurred a total of $\sum_{i=1}^{M_1} X(i, t=1)$.

Within the first M_2 years, the insured has incurred a total of $\sum_{i=1}^{M_2} X(i, t=2)$, and within the first M_n years, the insured has incurred a total of $\sum_{i=1}^{M_n} X(i, t=n)$.

In $M = M_1 + M_2 + \cdots + M_n$ years, one insured has incurred a total of $\sum_{t=1}^{n}\sum_{i=1}^{M_t} X(i, t)$ claims. By applying the Buhlmann credibility formula, we can determine the expected claim cost in year $m+1$ for one insured.

$$P = Z\left(\bar{X}\right) + (1 - Z)\,\mu \tag{24}$$

where;

$$\bar{X} = \frac{TotalobservedClaims}{TotalNumberofobservations} = \frac{1}{m}\sum_{t=1}^{n}\sum_{i=1}^{m_i} x\,(i,t) \tag{25}$$

which are the unified formulas for both the Buhlmann Straub and Buhlmann credibility approaches models.

Definition 3. Our goal is to estimate $E\left(x_{n+1}\right)$, that gives the average/mean claim in year $n+1$ where the application of $a + Z\bar{Y}$ when approximating the values of $E\left(x_{n+1}\right)$ and $\bar{X} = \frac{1}{m}\sum\limits_{i=1}^{n} m_i X_i$. The results will minimizes the value of $E\left[a + Z\bar{X} - E(X_{n+1})\right]^2$ where $Z = \frac{Cov\left(\bar{X}, X_{n+1}\right)}{Var\left(\bar{X}\right)} a = (1 - Z)\,\mu$.
The resulting formula will be as follows;

$$E\left(X_t|\theta\right) = E\left(\frac{1}{m_t}\sum_{i=1}^{m_t}[X\left(i,t=i\right)|\theta]\right) = \frac{1}{m_t}\sum_{i=1}^{m_t}\mu\left(\theta\right) = \mu\left(\theta\right) \tag{26}$$

$$Var\left(X_t|\theta\right) = Var\left(\frac{1}{m_t}\sum_{i=1}^{m_t}[X\left(i,t=i\right)|\theta]\right) = \frac{1}{m_t}\left[m_t\sigma^2\left(\theta\right)\right] = \frac{\sigma^2\left(\theta\right)}{m_t} \tag{27}$$

$$E\left(\bar{X}|\theta\right) = E[\sum_{i=1}^{n}\frac{m_i}{m}\left(X_i|\theta\right)] = \frac{\mu\left(\theta\right)}{m}\sum_{i=1}^{n}m_i = \mu\left(\theta\right) \tag{28}$$

$$Var\left(\bar{X}|\theta\right) = Var\sum_{i=1}^{n}\frac{m_i}{m}\left(X_i|\theta\right)] = \frac{1}{m^2}\sum_{i=1}^{n}m_i^2\frac{(\sigma^2\theta)}{m_i} \tag{29}$$

$$Var\left(\bar{X}|\theta\right) = \frac{\sigma^2}{m^2}\sum_{i=1}^{n}m_i = \frac{\sigma^2\left(\theta\right)}{m} \tag{30}$$

Given θ has mean $\mu(\theta)$ and variance $\frac{\sigma^2(\theta)}{m}$, then the covariance is given by;

$$Cov\left(\bar{X},X_{n+1}\right) = E\left(\bar{X}X_{n+1}\right) - E\left(\bar{X}\right)E\left(X_{n+1}\right) \tag{31}$$

$$E\left(\bar{X},X_{n+1}\right) = E_\theta\left[E\left(\bar{X},X_{n+1}\right)|\theta\right] = E_\theta\left[\mu\left(\theta\right)\right]^2 \tag{32}$$

$$E\left(\bar{X}\right) = E_\theta\left[E\left(\bar{X}\right)|\theta\right] = E_\theta\left[\mu\left(\theta\right)\right] \tag{33}$$

$$Z = \frac{Cov\left(\bar{X},X_{n+1}\right)}{Var\left(\bar{X}\right)} = \frac{Var\left[\mu\left(\theta\right)\right]}{Var\left(\bar{X}\right)} \tag{34}$$

$$Z = \frac{Var_\theta\left[\mu\left(\theta\right)\right]}{Var_\theta\left(\mu(\theta)\right) + \frac{1}{n}E_\theta\left(Var(X|\theta)\right)} = \frac{n}{n + \frac{E_\theta(Var(X|\theta))}{Var_\theta[\mu(\theta)]}} \tag{35}$$

In the Buhlmann Straub credibility model, the total exposure denoted by m and the historical mean claim per exposure denoted by $\bar{x}$ are the crucial factors. In this model,

Risk, θ	**Year 1**	**Year 2**	**...**	**Year n**
1	X_{11}	X_{12}	...	X_{1n}
2	X_{21}	X_{22}	...	X_{2n}
...	....	....	....	...
n	X_{r1}	X_{r1}	....	X_{rn}

Table 1: **Risk Profiles Attributes**

the amount of individual claims denoted by $X(i;t)$ does not matter.

3.1.4 Empirical Bayes Estimation for The Buhlmann Model

Suppose we have n-year claim data for r risks. For each year j $(j = 1, 2, \ldots, n)$ and for the i-th policyholder, let X_{ij} be the incurred claim amount. However, we don't know the distribution of the conditional claim random variable $X|\theta$ or the risk variable θ distribution, so we cannot calculate the two inputs of the credibility factor Z.

The expected process variance and variance of the hypothetical mean, abbreviated as EV and VE, are denoted as $E_\theta(\mathrm{Var}(X|\theta))$ and $\mathrm{Var}_\theta[E(X|\theta)]$, respectively. Therefore, the values of EV and VE can be estimated empirically from past claim data.

To get EV, we take the average of the variance of each risk, which is tabulated below in Table (2).

Risk,θ	**Year 1**	**Year 2**	**...Year n**	**Sample Mean,$\bar{X}_i$**	$\widehat{\sigma}_i^2 = \frac{1}{n-1}\sum_{t=1}^{n}\left(X_{it} - \bar{X}_t\right)^2$
1	X_{11}	X_{12}	...X_{1n}	$\bar{X}_1 = \frac{1}{n}\sum_{t=1}^{n} X_{1t}$	$\widehat{\sigma}_1^2 = \frac{1}{n-1}\sum_{t=1}^{n}\left(X_{1t} - \bar{X}_1\right)^2$
2	X_{21}	X_{22}	...X_{2n}	$\bar{X}_2 = \frac{1}{n}\sum_{t=1}^{n} X_{2t}$	$\widehat{\sigma}_2^2 = \frac{1}{n-1}\sum_{t=1}^{n}\left(X_{2t} - \bar{X}_2\right)^2$
...	....	....	....	...	....
r	X_{r1}	X_{r2}	....X_{rn}	$\bar{X}_r = \frac{1}{n}\sum_{t=1}^{n} X_{rt}$	$\widehat{\sigma}_r^2 = \frac{1}{n-1}\sum_{t=1}^{n}\left(X_{rt} - \bar{X}_r\right)^2$

Table 2: **Expectation Values of Empirical Bayes Estimates**

Risk, θ	Year 1	Year 2	Year n_1	Year n_2	Year n_r
1	X_{11}	X_{12}	X_{1n1}	X_{1n2}	X_{1nr}
	M_{11}	M_{12}	M_{1n_1}	M_{1n_1}	M_{1n_1}
2	X_{21}	X_{22}	X_{2n1}	X_{2n2}	X_{2nr}
	M_{21}	M_{22}	M_{2n1}	M_{2n_1}	M_{2n_1}
...	....	....	....	....	....
r	X_{r1}	X_{r2}	X_{rn1}	X_{1r2}	X_{rnr}
	M_{r1}	M_{r2}	M_{rn1}	M_{1n_1}	M_{rnr}

Table 3: **Expected Estimates**

We can then calculate the expected process variance as;

$$\widehat{EV} = \frac{1}{r}\sum_{i=1}^{r}\widehat{\sigma}_i^2 \tag{36}$$

3.1.5 Empirical Bayes Estimate for The Buhlmann –Straub Model

Here the number of policyholders varies from risk to risk and year to year.

For risk 1;

M_{11} policyholders incur X_{11} claim amount within year 1. In year 2, M_{12} insureds incur X_{12} claim amounts and in year n, M_{1n} insureds incur X_{1n} claim amount.

For risk 2;

M_{21} insureds incur X_{21} claim amount within year 1. In year 2, M_{22} policyholders incur X_{22} claim amount and in year n, M_{2n} insureds incur X_{2n} claim amount.

To estimate;

We first start by calculating the sample variance for every risk profile and the expected process variance for the combination of all the risks as;

Risk, i	Period, t	Total Exposures	Sample Mean	Sample Variance
1	n_1	$M_1 = \sum_{t=1}^{n_1} M_{1t}$	$\bar{X}_1 = \frac{1}{m_1} \sum_{t=1}^{n_1} M_{1t} X_{1t}$	$\hat{\sigma}_1^2 = \frac{1}{n_1 - 1} \sum_{t=1}^{n_1} M_{1t} \left(X_{1t} - \bar{X}_1 \right)^2$
2	n_2	$M_2 = \sum_{t=1}^{n_2} M_{2t}$	$\bar{X}_2 = \frac{1}{m_2} \sum_{t=1}^{n_2} M_{2t} X_{2t}$	$\hat{\sigma}_2^2 = \frac{1}{n_2 - 1} \sum_{t=1}^{n_2} M_{2t} \left(X_{2t} - \bar{X}_2 \right)^2$
...	...	...	....	....
r	n_r	$M_r = \sum_{t=1}^{n_t} M_{rt}$	$\bar{X}_r = \frac{1}{m_r} \sum_{t=1}^{n_r} M_{rt} X_{rt}$	$\hat{\sigma}_r^2 = \frac{1}{n_r - 1} \sum_{t=1}^{n_r} M_{rt} \left(X_{rt} - \bar{X}_r \right)^2$
Total		$M = \sum_{t=1}^{n} M_t$	$\bar{X} = \frac{1}{m} \sum_{t=1}^{n} M_t X_t$	$\widehat{EV} = \frac{\sum_{i=1}^{r} (n_i - 1)\, \sigma_i^2}{\sum_{i=1}^{r} (n_i - 1)}$

Table 4: **The Given Estimates of Sample Variance Vs Sample Mean**

$$\hat{\sigma}_i^2 = \frac{1}{n_i - 1} \sum_{i=1}^{n_i} M_{it} \left(X_{it} - \bar{X}_i \right)^2$$

$$\widehat{EV} = \frac{\sum_{i=1}^{r} (n_i - 1)\, \sigma_i^2}{\sum_{i=1}^{r} (n_i - 1)} = \frac{\sum_{t=1}^{r} \sum_{i=1}^{n_i} M_{it} \left(X_{it} - \bar{X}_i \right)^2}{\sum_{i=1}^{r} (n_i - 1)} \tag{37}$$

Then we calculate:

$$\widehat{VE} = \frac{\sum_{i=1}^{r} m_i \left(\bar{X}_i - \bar{X} \right)^2 - (r - 1)\, \widehat{EV}}{m - \frac{1}{m} \sum_{i=1}^{r} m_i^2} \tag{38}$$

The credibility weighted average premium approach uses loss models to prove the total loss as equal to the cumulative total premium whenever the value of μ is given by;

$$\mu = \frac{\sum_{i=1}^{r} \bar{Z}_i \bar{X}}{\sum_{i=1}^{r} \bar{Z}_i} \tag{39}$$

3.2 Parameter Estimation of the Credibility Premium

The parameters are as follows:

$$m = E\left(Y_{it}|\Theta_i\right) = \mu\left(\Theta_i\right) \text{ and}$$
$$s^2 = E\left(\sigma^2\left(\Theta_i\right)\right).$$

This means that we denote:

$$a = Var\left(\mu\left(\Theta_i\right)\right) \text{ and}$$
$$w_i = \sum_{t=1}^{T_i} w_{it} \text{ with } w = \sum_{i=1}^{n}\sum_{t=1}^{T_i} w_{it}$$
$$X_i = \sum_{t=1}^{T_i} \frac{w_{it}}{w_i}$$
$$\bar{Z} = \sum_{i=1}^{m} Z_i$$

The credibility premium is calculated through the minimization of mean square error. This is calculated using credibility premium as follows;

$$P_i = Z_i\bar{X}_i + m\left(1 - Z_i\right) \tag{40}$$

where $Z_i = \frac{w_i}{w_i+k}$ with $k = \frac{s^2}{a}$. The parameters are estimated using the value of the NHIF data.

It is important to note that parameters using NHIF claims data as follows with $m = Z_i$, which are the estimates of $s^2$and a functions for the credibility estimate.

We define the value of s^2 as the unbiased estimator of the variance of the expected means and the value, a as the unbiased mean of the hypothetical variance. The two values of s^2 and a are the calculated using NHIF data to calculate the credibility estimate of Z.

CHAPTER FOUR

RESULTS AND DISCUSSION

4.1 Introduction

In this chapter, we find the valuations of premiums as per counties for the Universal Health Care (UHC) and find the expected premiums according to the existing risks for the respective counties. We simulate Data for the four counties and find the estimate premiums.

4.2 Model Application

For model application, number of claims are generated with an assumption of Poisson parameter random variables of λ. The risks acquired by the insurance companies for the portfolios are assumed to be S, which aggregate claims.

The aggregate claims, $S = X_1 + X_2 + \cdots + X_N$ for a given portfolio of policies are as follows. The mean of the hypothetical variance and variance of mean must be equal, in such a way that $E(\mu(\Theta_i)) = Var(\Theta_i) = \Theta_i$.

The claims distribution is skewed from the left to the right considering the recent project of Universal Health Care launch by the government on the NHIF program. By considering the data from the health records from the four counties that were under pilot test of the UHC program, we will use NHIF program to determine the projected premiums paid by the policyholders.

A classification rating plan is a system that groups policyholders based on their risk characteristics. While policyholders within a class share similarities, they are not identical, and their expected losses will differ. To produce a more accurate rate for individual policyholders, an experience rating plan can supplement a class rating plan by crediting their loss of experience with the class rate.

When no loss information about the risk is available, the expected loss for a randomly chosen policyholder from the class is given by $\mu = E(\mu(\theta))$, where the expectation is taken over all θ's in the class. In this situation, $Z = 0$ and the expected loss for the risk is $\hat{\mu}(\theta) = \mu$. The quantity μ, which can also be expressed as $\mu = E(X_j)$ or $\mu = E(X)$, is known as the overall mean or collective mean. Note that $E(X_j)$ is evaluated using the law of total expectation: $E(X_j) = E(E[X_j \mid \theta])$.

4.3 Data Analysis

The simulation of data might not be a simple process, but through the convolution process, the amounts will result into a compound Poisson distribution. We used a discrete form for the convolution of claims that resulted into a convolution of claims thus resulting into a compound Poisson distribution. This lead to calculation of credibility premiums based on the individual Health Insurance Schemes Based on National Health Insurance Fund (NHIF).

The generation of weights follows a uniform distribution (α, β), $0 \leq \alpha \leq \beta$. These are total weights demonstrated from the calculated values of total number of claims as per the year of a function of time. The risk functions are generated as Poisson distribution functions.

4.4 Aggregate claim amounts

The data represents the aggregate claim amounts for four different counties (Region1, Region2, Region3, and Region4) over a period of five years (year 1 to year 5). The total claim amount for each county is also included.

County	year1	year 2	year 3	year 4	year 5	Total
Region1	106,872	123,565	132,896	145,767	150,000	659,100
Region2	789,658	796,927	823,892	883,923	923,945	4,218,345
Region3	682,824	692,364	710,274	724,827	762,912	3,573,201
Region4	572,281	589,883	601,283	610,235	622,812	2,996,494

Table 5: **Aggregate claim amounts**

From the table above we can see that Region2 has the highest total claim amount of 4,218,345 followed by Region3 with 3,573,201. Region4 and Region1 have lower total claim amounts of 2,996,494 and 659,100, respectively.

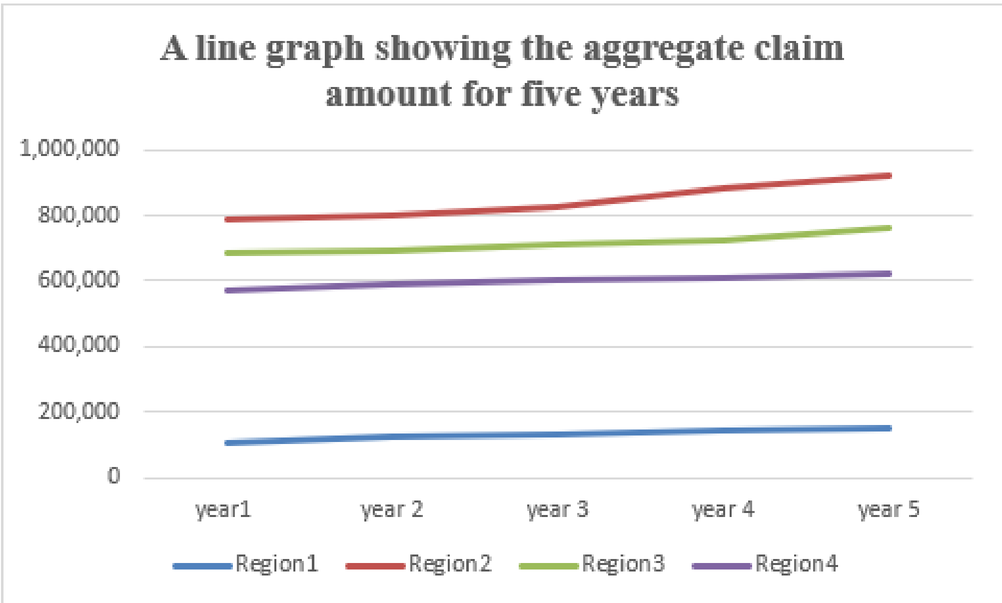

Figure 1: A line graph showing the aggregate claim amount for five years

Examining the data year by year, we can observe from the above line graph that all counties experienced an increase in the aggregate claim amounts from year 1 to year 5. In Region2, the aggregate claim amount increased from 789,658 in year 1 to 923,945 in year 5. Similarly, Region3 saw an increase from 682,824 to 762,912 during the same period. Region4 and Region1 also experienced an upward trend in the aggregate claim amounts over the years.

The highest claim amount in any given year was reported in Region2, with a claim amount of 923,945 in year 5. On the other hand, Region4 had the lowest claim amount in every year, with a claim amount of 106,872 in year 1 and 150,000 in year 5. The data suggests that the total claim amount increased steadily over the five-year period for all four counties. It is worth noting that Region2 and Region3 had significantly higher total claim amounts compared to Region4 and Region1. This may indicate a higher prevalence of health issues or higher healthcare costs in Region2 and Region3.

4.5 Mean and variance

The sample mean ($m(\theta)$) represents the average aggregate claim amount for each county over the five-year period. The sample variance ($S^2(\theta)$) measures the variability of the aggregate claim amounts for each county over the five-year period.

County	year1	year 2	year 3	year 4	year 5	Total	m(Θ)	S^2 (Θ)
Region1	106,872	123,565	132,896	145,767	150,000	659,100	131,820.0	304,184,178.5
Region2	945,900	908,903	823,892	883,923	923,945	4,486,563	897,312.6	2,193,555,858.3
Region3	682,824	692,364	710,274	724,827	762,912	3,573,201	714,640.2	990,375,235.2
Region4	572,281	589,883	601,283	610,235	622,812	2,996,494	599,298.8	373,756,725.2
						Average	585,767.9	965,467,999.3

Table 6: **Mean and variance**

$$Z = \frac{n}{n + \frac{E_\theta(Var(X/\theta))}{Var_\theta[\mu(\theta)]}} \tag{41}$$

$$Z = 0.999977367$$

We can see that the average claim amount for all four counties is 585,767.9, the county with the highest variance is Region$_2$, indicating that the claim amounts for Region2 have been more variable over the five years than the other counties. On the other hand, Region4 has the lowest variance, indicating that the claim amounts for Region$_4$ have been more consistent over the five years.

4.6 Buhlmann's Results

Collective premium: 585,767.9

Within county variance: 965,467,999.3

Between county variance: 106,640,000,000.

The Buhlmann credibility model results table below, summarizes the application of the model to the provided data.

County	$m(\Theta)$	Z_i	Credibility Premium
Region1	131820	0.999977367	131830.27
Region2	897312.6	0.999977367	897292.29
Region3	714640.2	0.999977367	714624.03
Region4	599298.8	0.999977367	599285.24

Table 7: **Bühlmann's Results**

It includes three columns: $m(\theta)$, representing the estimated mean claim amounts for each county; Z_i, indicating the credibility factor assigned to each county based on historical data and variability; and the credibility premium, which is the weighted average of the estimated mean claim amount and actual claim experience for each county.

4.7 Buhlmann-Straub Data

This table below shows the Buhlmann-Straub data for four counties, where each row represents a county, and each column represents a year. The second column (p1-p5) represents the number of policies sold each year, and the corresponding claim amounts for each year are given in columns Y1-Y5.

County	year1		year 2		year 3		year 4		year 5	
	Y1	p1	Y2	p2	Y3	p3	Y4	p4	Y5	p5
Region1	106,872	9	123,565	10	132,896	8	145,767	12	150,000	6
Region2	945,900	34	908,903	41	823,892	34	883,923	48	923,945	23
Region3	682,824	26	692,364	30	710,274	25	724,827	36	762,912	17
Region4	572,281	17	589,883	20	601,283	17	610,235	24	622,812	12

Table 8: **Buhlmann-Straub Data**

Using this model, a county's estimated mean claim amount in a given year is calculated as a weighted average of the individual policyholders' mean claim amounts. The weights are determined by the number of policies sold each year, with more weight given to more recent years.

4.7.1 Weights

County	X1	X2	X3	X4	X5
Region1	11874.7	12356.5	16612.0	12147.3	25000.0
Region2	27820.6	22168.4	24232.1	18415.1	40171.5
Region3	26262.5	23078.8	28411.0	20134.1	44877.2
Region4	33663.6	29494.2	35369.6	25426.5	51901.0

Table 9: **Buhlmann-straub Weights**

These weights (X1, X2, X3, X4, X5) were calculated by dividing the total claim amount for each year by the corresponding premium rate (p1, p2, p3, p4, p5) for each county. For example, for Region1 in year 1 (Y1), the weight (X1) was calculated as follows:

$$X1 = \frac{Y1}{p1} = \frac{106,872}{9} = 11,874.7$$

Similarly, for Region2 in year 5 (Y5), the weight (X5) was calculated as follows:

$$X5 = \frac{Y5}{p5} = \frac{923,945}{23} = 40,171.5$$

These weights are used in the Buhlmann-Straub method to calculate the final premium for each county. The procedure considers the historical data (weights) and the actual claim experience to determine premiums that are more accurate and reflective of the actual risk.

4.8 Buhlmann-Straub Credibility estimates

County	X1	X2	X3	X4	X5	m(Θ)	S^2 (Θ)
Region1	11874.66667	12356.5	16612	12147.25	25000	15598	31426002
Region2	27820.58824	22168.36585	24232.11765	18415.0625	40171.5	26562	69477022
Region3	26262.46154	23078.8	28410.96	20134.08333	44877.2	28553	93147626
Region4	33663.58824	29494.15	35369.58824	25426.45833	51901	35171	1.02E+08
					Average	26471	74099372

Table 10: **Buhlmann-Straub Credibility estimates**

From the above table, the estimated mean claim amounts for each county vary considerably, with Region1 having the lowest at 15,598.08 and Region4 having the highest at 35,170.96. This indicates that the risk of claims also varies across counties.

4.9 Buhlmann-Straub Results

Collective premium: 26470.81694

County	m(Θ)	Z_i	Credibility Premium
Region1	26470.82	9	238237.35
Region2	26470.82	50	1323540.85
Region3	26470.82	31	820595.33
Region4	26470.82	22	572907.97

Table 11: **Buhlmann-Straub Results**

The results show that the credibility premiums for each county vary based on the credibility factor assigned to them. The highest credibility factor was assigned to Region2, which resulted in the highest credibility premium of 1,323,540.85.

The Buhlmann-Straub credibility model provides a more accurate and fair way to determine insurance premiums by taking into account both historical data and recent claim experience.

CHAPTER FIVE

CONCLUSIONS AND RECOMMENDATIONS

5.1 Conclusions

In conclusion, this study aimed to analyze the credibility premiums of health insurance claims using the Buhlmann and Buhlmann-Straub credibility theory, with a focus on four counties in Kenya i.e Region1, Region2, Region3, and Region4 over a five-year period. The study found that all four counties experienced an increase in aggregate claim amounts over the five-year period, with Region2 and Region3 having significantly higher total claim amounts compared to Region4 and Region1. Region2 also had the highest claim amount in any given year, while Region1 consistently had the lowest.

The Buhlmann-Straub method was used to calculate the final premium for each county, taking into account historical data and actual claim experience to determine more accurate and reflective premiums. The estimated mean claim amounts for each county varied considerably, indicating that the risk of claims also varies.

However, the study also highlighted the challenges of obtaining detailed data in the medical sector, which can affect the accuracy of the credibility premium. Additionally, outliers in the data can distort mean and variance functions, further impacting the accuracy of the credibility premium.

5.2 Recommendations

As noted in the study, obtaining detailed and accurate data on health insurance claims can be challenging. To improve the accuracy and usefulness of future studies, efforts should be made to increase the availability and quality of data. This could involve partnering with insurance companies to obtain more comprehensive data sets, or investing in better data collection and management systems.

Buhlmann and Buhlmann-Straub credibility theory was effective for this study, it may not always be the best choice for all scenarios. Future studies could explore alternative statistical models and methods for computing credibility premiums. This could include machine learning algorithms or Bayesian methods, which may be better suited for certain types of data or contexts.

Outliers in the data can distort the mean and variance functions, leading to inaccuracies in credibility premium calculations. It is important for insurance companies to identify and address outliers in their data to improve the accuracy of premium calculations.

REFERENCES

1. Ahmad, M. N., & Abdullah, N. A. (2020). Credibility theory for experience rating in individual health insurance: A systematic review.

2. Hemedi, A. G., & Swai, E. (2020). Investigating factors influencing pricing of individual health insurance schemes: A study based on National Health Insurance Fund (NHIF) in Tanzania.

3. Zhang, Y., Fan, Y., & Li, S. (2019). Pricing individual health insurance based on credibility theory and generalized linear models: a case study in China.

4. Zhang, Y., Fan, Y., & Li, S. (2019). Pricing individual health insurance based on credibility theory and logistic regression: a case study in China.

5. Akosah-Twumasi, P., & Okpoti, C. A. (2018). Pricing individual health insurance schemes based on National Health Insurance Fund (NHIF) using credibility theory: Evidence from Ghana.

6. Zhang, Y., Li, S., & Fan, Y. (2019). Pricing individual health insurance using credibility theory and Bayesian network: A case study in China.

7. Li, S., Fan, Y., & Zhang, Y. (2019). Pricing individual health insurance using credibility theory in China: A case study of a private insurance company.

8. Muoki, E. M., & Weke, P. O. (2018). Pricing individual health insurance using credibility theory: A case study in Kenya.

9. Zanjani, G. S., & Jafari, M. (2018). Pricing individual health insurance with credible data in a developing country context.

10. Hamidian, M., & Vahidnia, F. (2020). Pricing of individual health insurance based on credibility theory: Evidence from private insurance companies in Iran.

11. Chukwuma, U. J., Adesanya, A. O., & Nwosu, I. A. (2020). Pricing of individual health insurance schemes based on National Health Insurance Fund (NHIF) using credibility theory: A case study of Nigeria.

12. Mubiru, K. S., & Nsabimana, A. (2021). Pricing of individual health insurance schemes in Uganda: An application of credibility theory.

13. Edwine Barasa, Khama Rogo, Njeri Mwaura, and Jane Chuma. Kenya national hospital insurance fund reforms: implications and lessons for universal health coverage. Health Systems & Reform, 4(4):346–361, 2018.

14. Guy Carrin and Chris James. Social health insurance: key factors affecting the transition towards universal coverage. International Social Security Review, 58(1):45–64, 2005.

15. James K Kimani, Remare Ettarh, Catherine Kyobutungi, Blessing Mberu, and Kanyiva Muindi. Determinants for participation in a public health insurance program among residents of urban slums in nairobi, kenya: results from a cross-sectional survey. BMC health services research, 12(1):1–11, 2012.

16. Mahiben Maruthappu, Johnathan Watkins, Aisyah Mohd Noor, Callum Williams, Raghib Ali, Richard Sullivan, Thomas Zeltner, and Rifat Atun. Economic downturns, universal health coverage, and cancer mortality in high-income and middleincome countries, 1990–2010: a longitudinal analysis. The Lancet, 388(10045): 684–695, 2016.

17. Rahab Mbau, Evelyn Kabia, Ayako Honda, Kara Hanson, and Edwine Barasa. Examining purchasing reforms towards universal health coverage by the national hospital insurance fund in kenya. International journal for equity in health, 19(1):1–18, 2020.

18. Ngwakwe, C. C., Asogwa, F. O., & Abasilim, U. D. (2020). The impact of experience rating on pricing of individual health insurance in Nigeria: a case study of National Health Insurance Scheme (NHIS). Journal of Economics and Sustainable Development, 11(15), 23-34.

19. Abidin, I. Z., & Mardiana, A. (2021). Pricing health insurance in Indonesia: A comparison of credibility theory and experience rating. Journal of Health Policy

and Management Research, 2(2), 74-82.

20. Bajtelsmit, V. L., & Bouzouita, M. (2007). The impact of credibility on insurance premiums in a competitive property-liability insurance market. Journal of Risk and Insurance, 74(2), 421-438.

21. Umaru, A. B., Umaru, J. A., & Ahmad, I. A. (2020). An evaluation of credibility theory in the pricing of health insurance policies in Nigeria: evidence from selected insurance companies. Journal of Business and Management Sciences, 8(2), 40-48.

22. Hafner, C. M. (2007). Credibility theory and its application to insurance pricing. North American Actuarial Journal, 11(3), 110-129.

23. Xie, B., & Wang, M. (2019). Pricing individual health insurance based on credibility theory and copula models: evidence from China. Asia-Pacific Journal of Risk and Insurance, 13(1), 1-19.

24. Jafari-Koshki, T., Mahdavi-Amiri, N., & Farsangi, E. N. (2020). An investigation of the credibility theory in pricing individual health insurance: Case study of Iran. Journal of Medical Systems, 44(2), 1-9.

25. Jing, X., Feng, J., & Xiaoling, G. (2021). The impact of credibility theory on individual health insurance premium pricing: an empirical study from China. International Journal of Health Economics and Policy, 6(2), 47-55.