Modifying the Chain-Ladder model as a basis for estimating mortality rates in the Egyptian insurance companies.

تعديل نموذج Chain-Ladder كأساس لتقدير معدلات الوفاة في شركات التأمين المصرية.

Pr. Mohammed T. El-bolkiny

Professor of Actuarial Sciences and Mathematics, Department of Applied Statistics and Insurance, Faculty of Commerce - Mansoura University

Pr. Jamal A. Wasif

Professor of Actuarial Sciences and Mathematics, Department of Applied Statistics and Insurance, Faculty of Commerce - Mansoura University

Dr. Mohammed A. El-barqawy

Lecturer of Insurance, Department of Applied statistics and Insurance,

Faculty of Commerce - Mansoura University

Maha M. Elzekey

Assistant lecturer in Statistics & Insurance Department

Faculty of Commerce - Mansoura University

Abstract:

Estimating the price of life insurance policies is the most significant challenge facing insurance companies. Egyptian insurance companies use mortality tables that are inconsistent with the Egyptian reality in terms of the data used and the time period. Medical and scientific advances have led to the phenomenon of longevity, which is considered a risk for life insurance companies due to the increase in the value of the company's liabilities.

Therefore, this research will apply the chain-ladder method to predict mortality rates. This is the first time this method has been applied to data specific to the Egyptian market (data from insurance companies). After obtaining the rates using this method, these rates will be refined to become smoother.

الملخص:

إن تقدير سعر وثائق تأمينات الحياة يعد أهم تحديًا يواجه شركات التأمين. حيث أن شركات التأمين المصرية تستخدم جداول وفاة غير متوافقة مع الواقع المصري من حيث البيانات المستخدمة والفترة الزمنية أيضًا. إن التقدم الطبي والعلمي تسبب في ظاهرة طول العمر، والتي تعتبر خطر بالنسبة لشركات التأمين على الحياة لما تتسبب فيه من زيادة في قيمة التزامات الشركة.

لذلك سوف يتم في هذا البحث تطبيق طريقة Chain-ladder في التنبؤ بمعدلات الوفيات. وتعد هذه المرة الأولي لتطبيق هذه الطريقة على البيانات الخاصة بالسوق المصري (بيانات من واقع شركات التأمين). وبعد الحصول على المعدلات من خلال هذه الطريقة سيتم تهذيب هذه المعدلات حتى تصبح أكثر سلاسة.

Keywords: Chain ladder method – Mortality rate - Run-off Triangle - Longevity.

Introduction:

Understanding and predicting mortality rates are fundamental elements in many fields, including social planning, economics, and risk assessment in the insurance industry. Central mortality rates (m_x) are widely used in actuarial and demographic analyses to evaluate and forecast mortality trends, helping to inform decisions related to public policies and risk management.

The science of creating and forecasting mortality patterns has strong origins in actuarial and demographic studies. Since the introduction of one of the earliest models of mortality- the life table - by John Graunt in 1662, models have been used to depict death patterns at various ages and times. One of the most significant and traditional probabilistic demographic tables, it is frequently employed as a special and efficient demographic tool to solve population issues. For millennia, Life tables have been used in the study of mortality.

According to some researchers, "the mortality rate is the most important indicator of health that is easily and reliably measured" [(1)]. Actuaries and anybody else interested in life insurance use mortality statistics as a foundation for many actuarial computations, whether they are involved in predicting technical provisions or premiums. Mortality tables offer crucial information for insurers, policyholders, and society at large.

Since the early 1900s, when English actuaries started calculating the potential financial burden of longevity due to higher survival rates and its impact on insurance reserves and pensions, one of the most significant applications of mortality tables has been to estimate mortality rates and forecast future deaths.

As a result of significant advancements in medicine, society, and the economy, survival rates have increased, raised people's life expectancy and, in turn, lower mortality rates. According to some researchers [(2)], there are two primary phases that account for this improvement:

The first: is the notable drop in the number of deaths at an early age.

Second: There has been a slowdown in the rise of mortality rates among individuals aged 65 and older.

---

1) James W. Vaupel, "Biodemography of Human Ageing", *Nature 464*, *Issue* 7288 (2010): 536-542.

2) Jim Oeppen et al., "Broken Limits to Life Expectancy". *Science* 296, Issue 5570 (2002): 1029-1031.

Researchers Oeppen and Vaupel (2002) ,have demonstrated that life expectancy surpasses historical projections, and that individuals should expect to live longer as time goes on, particularly as health care quality improves.

Over the last 200 years (since the early 19th century), life expectancy has more than doubled throughout the entire world. As a result, People live longer and better lives overall thanks to this change in life expectancy. It has significantly increased both the size of the population and economic output, including the number of elderly people. Knowing the dynamics of human mortality is crucial given the fast aging and longer lifespans that most populations now experience. Academics and professionals agree that appropriately estimating longevity risk is crucial to coping with these rapid changes and preventing unfavorable outcomes. A great deal of work has gone into creating more accurate techniques and mathematical models for future prediction. Academics and professionals agree that appropriately estimating longevity risk is crucial to coping with these rapid changes and preventing unfavorable outcomes. A great deal of work has gone into creating more accurate techniques and mathematical models for future prediction. For example, some researchers [3] identified a variety of predictive models of age-specific death rates using:

- extrapolation of transformation of death-rates and mortality probabilities.
- projection using a life tables model .
- projection by cause of death.
- projection by reference to another population.
- projection by reference to a law of mortality .
- combinations of these techniques .

The current research presents novel and cutting-edge techniques for projecting and calculating human mortality rates in the future.

This research aims to:

We use the popular chain-ladder approach to simulate mortality rates in life insurance, inspired by the booking techniques used by property and liability insurers. Mortality evolution patterns observed from past data can be used to predict future mortality rates. We can better predict future mortality rates by modeling mortality trends using the chain-ladder technique, which incorporates evolving patterns.

The importance of this research stems from the fact that:

Mortality rates are used in many different disciplines, particularly those in which future benefits (such as pension payments and social pension systems) depend on mortality rates. There are differences in the importance of

---

3) John H. Pollard, "Projection of Age-Specific Mortality Rates". *Population Bulletin of the United Nations* 21, Issue *22* (1987): 55-69.

mortality rates between the insurance and social insurance industries, as well as between the insurance industry and demographic studies of society. This importance can be briefly described as follows:

For life insurance:

- Survival rates are the foundation for creating life tables, and if they stay up to date with changes, they will help lower the cost of insurance (annuities, pension funds) and boost demand for its purchase, which will benefit insurance companies that have an impact on the national economy as an investment.
- Achieving fairness in the insurance premium calculation process that supports insurance firms' financial stability entails avoiding inflating the rise in premiums received over actual commitments, which would be detrimental to the insured's interests.
- The capacity to accurately compute mathematical allowances prevents the business from suffering financial losses if financial insurance amounts are paid that were not accurately projected, whether they were very high or significantly lower than anticipated.
- Bolstering the financial standing of insurance firms by increasing the size of their investment and insurance portfolios because of higher premium collections and, consequently, larger sports allocations, as well as improving the precision of their technical operations to meet the law of large numbers, which is regarded as one of the most crucial technical principles of insurance.

For social insurance:

- It contributes to the realization of the fairness principle by calculating old-age pensions in a manner that ensures equitable distribution among successive generations.
- It aids in creating long-term financial plans for pension distribution or determining the anticipated contribution amounts, which differ according to the demographic makeup of the community.
- The capacity to determine the price of full pension insurance for men and women independently, since it is known that the rates for men and women differ.
- Because of the study of the extent of the projected change in these rates from the actual rates, it is vital to know the shape of the mortality curve in the society in order to establish the best possible allocation of financial resources.

Data used in this research:

It is important to consider relevant data when creating and constructing life tables. The most important of these data are mortality rates, which form the basis of any mortality table. This data can be found through life insurance company records, as they are the largest source of mortality rates for life insurance experts. This data needs to be large and uniform in scope. In order to collect the statistical data required to create a life table that these companies can use to determine insurance rates, multiple insurance companies frequently collaborate.

Therefore, data on life insurance policies was collected from several Egyptian insurance companies for the period (2014-2022).

The research is divided into the following sections:

First: Longevity risk and mortality rates.

Second: The chain-ladder method and how it is used for mortality rates.

Third: Application of the chain-ladder model, results, and research recommendations.

1. Longevity risk and mortality rates.

1.1 The concept of longevity risk:

Longevity risk is defined as the risk that people live longer than expected, and it is a significant concern for modern civilizations. The ongoing increases in longevity that are being seen are giving rise to a great deal of new problems and difficulties for society on many different levels. Millions of people live longer and are healthier thanks to longevity advancements, but the cost of defined benefit and pay-as-you-go pension plans is rising as well, endangering the long-term viability of financial institutions by raising unanticipated future liabilities. Public health expenses are also impacted if healthier life expectancy is increased as a result of reductions in mortality rates at older ages.

The longevity of the population presents significant long-term challenges for pension funds and life insurance firms that hold large amounts of annuities. The companies faced a new difficulty in pricing annuities as individuals began to focus more on the longer life expectancy of the population. The need for models that accurately predict changes in insured individuals' death rates has become even more apparent, considering the recent financial crisis, as investment gains are no longer sufficient to offset undervalued longevity concerns. These models also need to produce realistic projections for the future because we anticipate that longevity will continue to rise.

1.2 Mortality rate:

The insurance premiums collected must be commensurate with the responsibilities assumed by the insurance company in order to fulfill its role as a regulator or broker. Insurance companies seek to determine optimal pricing and therefore employ a variety of methods and approaches (as actuarial science is one of the most important mechanisms that help solve problems related to future risks), including calculating probabilities based on available statistics to formulate insurance strategies, make pricing decisions, and design insurance and pension plans. To achieve this goal, probabilities are calculated using statistics that are now accessible. In short, actuarial science is the application of mathematics to the insurance industry. With its help, insurance companies can provide appropriate solutions that ensure the safety of the policyholder and their family in the event of an emergency arising from an unsafe life, especially given the variety of life insurance policies available, each with varying levels of risk, premiums, and compensation.

Life insurance is primarily based on the probabilities of life or death, which are considered experimental probabilities that are determined by observing and testing many individuals to minimize the effect of chance and achieve results that are as accurate as possible. From this, it is evident how important the theory of probabilities and the law of large numbers are to the life insurance industry, as life tables are thought to be how these probabilities for various life years are clarified

The mortality rate has a significant role in determining the cost of life insurance and annuity policies offered by life insurance firms as well as government-provided social security insurance. Based on anticipated mortality rates, they must appropriately hedge mortality and longevity risks and set policies prices. The fact that life insurance and annuity Policies are long-term contracts is one of their primary features. Annuities and whole life insurance, for instance, offer rewards for the insured person's whole lifetime (as long as he was alive). Consequently, it's critical to estimate mortality rates accurately.

Mortality measures are essential tools for understanding the health and demographic dynamics of a population. These measures provide insights into the frequency, causes, and distribution of deaths within a given population, enabling policymakers, actuaries, and public health officials to make informed decisions. Mortality rates serve as key indicators of population health, reflecting the impact of factors such as healthcare access, disease prevalence, and socioeconomic conditions. By analyzing these rates, researchers can identify trends, assess risks, and develop strategies to improve public health outcomes. Below, we will review the two most important types of mortality rates [(4)]:

[4)] Nathan Keyfitz et al., “Applied mathematical demography*”, New York: Springer* 47, 2005.

- ✓ Annual mortality rate (q_x): The probability that an individual of a given age x will die between the age (x) and (x+1).
- ✓ Central annual mortality rate (m_x) : determined by dividing the number of people who died during this period while aged x (after they had reached the age x but before reached $x+1$) by the average number who were living in that age group during the period.

Although both measures are used to determine the probability of death within a specific age group, they differ in their calculation and interpretation.

Despite their differences, they are fundamentally related, and understanding the relationship between them is essential for constructing life tables, estimating survival probabilities, and analyzing mortality trends. We will explore the mathematical and conceptual connections between these two measures, highlighting their importance in actuarial and demographic studies.

$$q_x = \frac{\text{Number of deaths between age } x \; and (x+1)}{\text{Number of individuals alive at age } x} \qquad q_x = \frac{d_x}{l_x}$$

$$m_x = \frac{\text{Number of deaths between age } x \; and (x+1)}{Average\ population\ aged\ x\ during\ the\ year} \qquad m_x = \frac{d_x}{L_x}$$

$$f_x = \frac{L_x - l_{x+1}}{d_x}$$

Since:

Average fraction (f_x): The average number of years lived in (x,x+1] by those who die in that interval. This average number is necessarily less than one.

Moreover, $d_x = l_x - l_{x+1}$

$$\therefore\ l_{x+1} = l_x - d_x$$

$$f_x . d_x = L_x - l_{x+1}$$

$$f_x . d_x = L_x - l_x + d_x$$

Or $L_x = l_x - (1 - f_x)\, d_x$

Then $m_x = \frac{d_x}{L_x} = \frac{d_x}{l_x - (1-f_x)\, d_x} = \frac{q_x}{1-(1-f_x) q_x}$

Alternatively, $l_x = L_x + (1 - f_x)\, d_x$

So $q_x = \frac{d_x}{l_x} = \frac{d_x}{L_x + (1-f_x)\, d_x} = \frac{m_x}{1+(1-f_x) m_x}$

2. The model of the research.

The insurer needs sufficient reserves to ensure that all claims are paid. For the policies that are now active, an insurance firm must have sufficient funds set aside to cover all claims, both current and future. Insolvency may result from inadequate reserves, and uncompetitive premium rates may result from excessive reserves. We refer to the money as a claim reserve.

An insurance firm guarantees that its policyholders will get benefits in the event of specific occurrences, such as a vehicle accident or medical issues. When this occurs, the insurance provider is obligated to pay the claim through claim reserving. Care must be taken when calculating claim reserving so as not to result in losses for the business.

The claim reserve is utilized to pay out claims that have been reported and are qualified to be paid. Payment of the claim may be made immediately after it is reported, but occasionally it may take longer, resulting in delays. The outstanding claim is the relationship between the incident and the related delays.

There are two categories of unresolved claims: Reported but Not Settled (RBNS) and Incurred but Not Reported (IBNR). IBNR is an event that has already occurred but has not yet been documented. Conversely, RBNS is a phenomenon that has previously been documented, but all associated payments have not yet been settled.

Because many policyholders are sometimes unsure of whether or not to submit a claim, IBNR claims reserve calculations are typically more challenging than RBNS. In the event of an accident, for example, the policyholder files a claim a few weeks after the incident. There will be a lag between the occurrence date and the date the insured was informed of the claim as a result. The same thing occurs when claims are paid. Although the process is lengthy, a claim may be paid multiple times before being finalized. The reporting date and the time frame during which the claim was deemed completed are likewise delayed.

Claims reserves are frequently estimated using the chain ladder approach (CL). where this method is one of the most straightforward ways to determine claim reserving.

The chain-ladder (CL) approach is one of the most important techniques in loss-reserving and claim setting in both theory and practice. It is considered one of the popular techniques for estimating the quantity of reserves for IBNRs (incurred but unreported claims). The chain-ladder method's main premise is that past loss development patterns from prior claims experience are reliable predictors of future loss development patterns. This approach makes use of payment information that is totaled in relation to the accident time and the interval between the accident and the payment.

The chain ladder method is an algorithm-driven approach used to calculate loss reserves in insurance. Professional insurance companies can leverage the data generated by this method to enhance their enterprise risk management by calculating various risk metrics. Typically, insurers estimate the necessary surplus to allocate to stakeholders at the end of a financial period. To determine this surplus, underwriters must assess the size of a buffer fund required to cover all outstanding claims for which full premiums have been collected [5]. There is often a significant delay between the occurrence of a claim and its final settlement, as closed claims may be reopened for legal disputes or the reporting of the claim to the insurer may be delayed. Underwriters usually investigate all reported claims through loss adjusters to verify their legitimacy, which can further extend the settlement process.

This approach uses the link ratios between the cumulative claims for two successive development years to anticipate future IBNR losses using the loss development triangle (historical experience data). The CL method's basic assumption is that historical claims data may be used to forecast future claim trends.

These claims naturally form an upper left triangle of actual incurred losses and a lower right triangle of future losses which referred to future triangle. Row in the form of run-off triangle stating the time period in which the accident occurred and the column states that the period during the claims settlement. Generally, run-off triangle data contains several claims or amounts of claims. The CL method estimates the empty cells of future triangle so the overall claims reserves needed to pay claims will occur in the future can be met.

The Chain Ladder method is the first deterministic technique created to calculate technical provisions for incurred but unreported claims (IBNR). The actuary uses previously paid or reported claims to predict future expected claims, assuming that the time series of claims is stable over time. We need a run-off triangle for the data input, which gathers cumulative data over the incurred claims with respect to the accident year and development year. The accident year is the year that the accident occurred, and the development year specifies how claims are paid in years gradually following the accident. Typically, there may be a delay between the occurrence of a claim and its settlement.

2.1 The original chain-ladder method (a distribution-free model):

The loss development triangle—also known as the run-off triangle—typically appears in property and liabilities insurance, where it could take a while following a loss to determine the full scope of the claims that need to be paid. It is crucial that the claims are attributed to the year in which the policy was created.

The claims must be analyzed by cohorts, and it's critical that each claim is assigned to the appropriate cohort. Cohorts can be defined in a variety

---

[5] Desnu Anggara Suwardi et al., "The analysis of motor vehicle insurance claim reserve using robust chain ladder", In *5th Global Conference on Business, Management and Entrepreneurship (GCBME 2020)*, Atlantis Press 187 (2021): 153–158.

of ways. Claims can be categorized, for instance, by the year the insurance was first issued, the year the accident happened, or a variety of other criteria.

The insurance firm must know how much it is obligated to pay in claims, to determine how much surplus it has created. Though, it might take several years before it is aware of the precise numbers of claims. The finalization of claim totals has been delayed for a variety of reasons. The delay could take place prior to the claim being notified or between notification and the ultimate settlement.

It is obvious that the insurance company must make every effort to estimate the total number of claims each year as accurately and confidently as possible, even while it is unsure of the precise number.

2.2 Run-off Triangle

The loss development triangle is a commonly used method for forecasting incurred but not reported (IBNR) losses based on historical data. We can forecast the final settlement date of insurance liabilities based on the trends in the data that have been noticed over time. When analyzing mortality data, we examine the trends in mortality rates by cohort, which reveal shifts in the mortality trend over time for different cohorts.

Claims data can be presented in a variety of ways that highlight various data points. The most popular approach will be used in this paper, where they will be displayed as a triangle. The accident year is the year that the occurrence occurred, and the insurer was at risk. The delay, also known as the development period, is the number of years that pass before payment is made. The accident year and the development year are used to separate the claims data. An example of claims data by accident year and development year can be found in the following table. Examining the evolution of claims by month or quarter may be pertinent in certain insurance kinds, but the fundamentals remain the same.

Each row in the triangle represents an origin year which establishes a cohort of claims. A cohort of accident years is used in this scenario. because a large portion of statistical theory in general insurance was established about automobile insurance, The term "accident year" is used to refer to circumstances where the claim event is obviously not an accident, such as auto thefts, arson, and burglaries. All claims pertaining to accidents that happened in 2010 are included in the 2010 row.

Most general insurers used an accounting year that begins on January 1, so the rows represent calendar years. Development years are represented by the columns, which illustrate how the group of claims associated with a specific origin year "develops" over time. The year of the accident is indicated in column 0. The year following the accident is shown in Column 1, etc.

Table (1): The development triangle for IBNR losses with I = J = 9

	Development Year j									
Accident year i	0	1	2	3	4	5	6	7	8	9
0(2010)	C_{00}	C_{01}	C_{02}	C_{03}	C_{04}	C_{05}	C_{06}	C_{07}	C_{08}	C_{09}
1(2011)	C_{10}	C_{11}	C_{12}	C_{13}	C_{14}	C_{15}	C_{16}	C_{17}	C_{18}	
2(2012)	C_{20}	C_{21}	C_{22}	C_{23}	C_{24}	C_{25}	C_{26}	C_{27}		
3(2013)	C_{30}	C_{31}	C_{32}	C_{33}	C_{34}	C_{35}	C_{36}			
4(2014)	C_{40}	C_{41}	C_{42}	C_{43}	C_{44}	C_{45}				
5(2015)	C_{50}	C_{51}	C_{52}	C_{53}	C_{54}					
6(2016)	C_{60}	C_{61}	C_{62}	C_{63}						
7(2017)	C_{70}	C_{71}	C_{72}							
8(2018)	C_{80}	C_{81}								
9(2019)	C_{90}									

Indicate $C_{\mathrm{i,j}}$ is the total number of claims or payments made in accident year i after a j-year reporting delay. The total of all the individual reported claims can be used to calculate cumulative losses $C_{\mathrm{i,j}}$. For $0 \leq i \leq I$ and $0 \leq j \leq J$, that is, $X_{\mathrm{i,j}}s$ occurred in a particular year i with delayed reports for j years as follows:

$$C_{i,j} = \sum_{k=0}^{j} x_{i,k}$$

We use I = J for a generic presentation even though I and J don't have to be the same .With two-time axes, i and j, we arrange $\mathrm{C_{i,j}}$s to form an upper left triangle (Table (1) for I = J = 9) given an observed data set

$$\mathrm{DI} = \{\mathrm{C_{i,j}} | \mathrm{i} + \mathrm{j} \leq \mathrm{I}, 0 \leq \mathrm{i}, \mathrm{j} \leq \mathrm{I}\}.$$

where:

$C_{i,j}$: cumulative payments for accident year I follow j development years.

i: accident year, year of occurrence (vertical axis).

j: development year, development term (horizontal axis).

Under the following Model Assumptions I:

- $C_{i,j}s$: losses or claims are independent for every different year i.
- There exist development factors $\mathrm{f}_0, \mathrm{f}_1, \dots, \mathrm{f}_{\mathrm{I}-1} > 0$ such that for every $0 \leq \mathrm{i} \leq \mathrm{I}$ and every $0 \leq \mathrm{j} \leq \mathrm{I}$ we have :
  $E\left[C_{i,j} \middle| C_{i,0}, \dots, C_{i,j-1}\right] = E\left[C_{i,j} \middle| C_{i,j-1}\right] = f_{j-1} \times C_{i,j-1}$

The chain ladder method assumes that the trend of claim development is consistent throughout accident years. Additionally, a weighted average of previous levels where inflation from prior years is likely to recur in the next years because of the possibility that inflation is a significant fiscal variable that is being carried forward into the future by development factors.

2.3 Forecasting Future Claims:

After having the cumulative claims, a run-off loss triangle with incremental claims must be applied. The unknown part of the loss triangle is estimated by applying the development factors.

We use link ratios to forecast consecutive cumulative claims. We expect the link ratios to capture the growth patterns of losses over time. The link ratio, also known as the age-to-age factor or loss development factor (LDF), shows the ratio of two total loss amounts for two consecutive years, from a specific development year to its preceding development year. The final amount of losses is also projected using these criteria.

For each development period, we can calculate the ratio of cumulative claims from one period to the previous period (e.g., claims at development period *j+1* divided by claims at period *j*).

2.4 Mortality Data

Before going over how mortality data are reorganized for use in the (CL)method, let's remember various forms of mortality data[6] :

$q_{x,t}$: The probability that a person at age x dies between age x + t and age $x + t + 1$.

$p_{x,t}$: $(1 - q_{x,t})$ The probability that a person at age x will be alive between age $x + t$ and age $x + t + 1$.

${}_{s}q_{x,t}$: The probability that a person at age x dies between age $x + t$ and age $x + t + s$, i.e. lives t years and then dies during the following s years.

${}_{s}p_{x,t}$: $(1 - {}_{s}q_{x,t})$ The probability of surviving for s years for a person at age x at time t.

$\mu_{x,t}$: The instantaneous death rate (force of mortality- hazard rate) for a person at age *x* at time *t*; specifically:

$$\mu_{x,t} = \frac{f_{x,t}}{1-F_{x,t}}$$

where:

- $f_{x,t}$: The probability density function (pdf),
- $F_{x,t}$: The cumulative distribution function (cdf), of the future lifetime of a person aged x at time t.

$m_{x,t}$: The central death rate for an individual aged x at time t.

---

[6] Peter D. England, "Stochastic Claims Reserving in General Insurance." *British Actuarial Journal* 8, Issue 3 (2002): 443-518.

$$m_{x,t} = \frac{q_{x,t}}{\int_0^1 {}_s p_{x,t} ds}$$

We suppose that for every integer age x and integer year t, $\mu_{x+r,t+s,} r, s \in [0,1)$, is constant. In this instance, the mortality data transformation between $m_{x,t}$ and $q_{x,t}$: is provided by

$$m_{x,t} = \mu_{x,t} = -\ln(1 - q_{x,t})$$

Next, we must move the observed central death rates $m_{x,t}$ from the upper right triangle of the original mortality table (shown in Table 2) with two axes (year t on the horizontal axis and age x on the vertical axis) to the upper left triangle (shown in Table 3) with two axes (cohort group k on the horizontal axis and age x on the vertical axis).

For example, people aged 25 in 2010 and people aged 25 + k in 2010 + k,

k = 1, 2, . . . , are in the same cohort (called cohort group). The mortality sequence in the first cohort group, which is the main diagonal in the mortality table (the red group in the upper right triangle of Table 2), is rearranged to the first column in the upper left triangle of Table (3). With the same way, we rearrange the upper right triangle mortality data (Table 2) to upper left triangle data (Table 3), so that we can apply the classical chain-ladder method used for property and casualty reserving to predict the mortality rates in the lower right triangle in gray (Table 3).

Keep in mind that in Table (2), row entry j below year t means age $t_L + i$ and column entry i below age x indicate age $x_L + i$. With x_L and t_L representing the beginning study age and year, respectively. An example of predicting mortality rates over the ensuing nine years using the upper right triangle of a 10 by 10 matrix is provided in Table (3). Similarly, we can use the upper right triangle of an I by I matrix to forecast death rates for the upcoming P(P ≤ I) years. It is necessary to move the anticipated mortality rates for the upcoming years from the lower right triangle in Table (3) to the lower left triangle in Table (4). Finally, using the traditional chain-ladder method once more, the mortality rates in the upper right triangle (blank region) of Table (4) are predicted using the mortality rates in the gray area. We can ultimately forecast mortality rates for the full nine years (t = 10, 11, ..., 18) by combining the expected mortality rates in the two triangles.

Table (2): Original mortality table with I = J = 9

	Year t									
Age x	0	1	2	3	4	5	6	7	8	9
0	m_{00}	m_{01}	m_{02}	m_{03}	m_{04}	m_{05}	m_{06}	m_{07}	m_{08}	m_{09}
1	m_{10}	m_{11}	m_{12}	m_{13}	m_{14}	m_{15}	m_{16}	m_{17}	m_{18}	m_{19}

2	m_{20}	m_{21}	m_{22}	m_{23}	m_{24}	m_{25}	m_{26}	m_{27}	m_{28}	m_{29}
3	m_{30}	m_{31}	m_{32}	m_{33}	m_{34}	m_{35}	m_{36}	m_{37}	m_{38}	m_{39}
4	m_{40}	m_{41}	m_{42}	m_{43}	m_{44}	m_{45}	m_{46}	m_{47}	m_{48}	m_{49}
5	m_{50}	m_{51}	m_{52}	m_{53}	m_{54}	m_{55}	m_{56}	m_{57}	m_{58}	m_{59}
6	m_{60}	m_{61}	m_{62}	m_{63}	m_{64}	m_{65}	m_{66}	m_{67}	m_{68}	m_{69}
7	m_{70}	m_{71}	m_{72}	m_{73}	m_{74}	m_{75}	m_{76}	m_{77}	m_{78}	m_{79}
8	m_{80}	m_{81}	m_{82}	m_{83}	m_{84}	m_{85}	m_{86}	m_{87}	m_{88}	m_{89}
9	m_{90}	m_{91}	m_{92}	m_{93}	m_{94}	m_{95}	m_{96}	m_{97}	m_{98}	m_{99}

Table (3): Rearranged mortality table with I = J = P = 9

	Cohort group k									
Age x	0	1	2	3	4	5	6	7	8	9
0	m_{00}	m_{01}	m_{02}	m_{03}	m_{04}	m_{05}	m_{06}	m_{07}	m_{08}	m_{09}
1	m_{11}	m_{12}	m_{13}	m_{14}	m_{15}	m_{16}	m_{17}	m_{18}	m_{19}	
2	m_{22}	m_{23}	m_{24}	m_{25}	m_{26}	m_{27}	m_{28}	m_{29}		
3	m_{33}	m_{34}	m_{35}	m_{36}	m_{37}	m_{38}	m_{39}			
4	m_{44}	m_{45}	m_{46}	m_{47}	m_{48}	m_{49}				
5	m_{55}	m_{56}	m_{57}	m_{58}	m_{59}					
6	m_{66}	m_{67}	m_{68}	m_{69}						
7	m_{77}	m_{78}	m_{79}							
8	m_{88}	m_{89}								
9	m_{99}									

The lower right area indicates "unknown" future mortality rates, whereas the upper left corner represents "known" past mortality rates. To finish the lower right triangle, we must examine ways to estimate these unknown rates. In other words, we can estimate the values in the lower right triangle using the top left triangular data provided above.

Table (4): Predicted lower left mortality table with I = J = P = 9 (in gray).

	Year t									
Age x	9	10	11	12	13	14	15	16	17	18
0	m_{09}									
1	m_{19}									
2	m_{29}									

3	m_{39}									
4	m_{49}									
5	m_{59}									
6	m_{69}									
7	m_{79}									
8	m_{89}									
9	m_{99}									

2.5 Forecasting Future Mortality Rates

The development factor is computed as the ratio of mortality rates between successive periods, typically using an average (such as the arithmetic or volume-weighted average) of observed ratios across age groups. These factors are then applied to project future mortality rates, assuming that past trends will continue. The average method smooths out fluctuations by considering the central tendency of historical development patterns, providing a stable and reliable estimate. This approach is particularly useful for completing incomplete mortality triangles and forecasting future mortality rates for life insurance and pension planning.

Simple Arithmetic Mean (Unweighted Average) is useful for short-term projections or preliminary analyses where mortality data shows low volatility and minimal year-to-year fluctuations, The arithmetic mean provides a quick, transparent estimate by averaging past mortality rates without complex modeling. This approach may work well for stable populations with consistent health trends, segments without extreme events (e.g., pandemics).

Calculating development factors using the chain ladder method (average method) for central mortality rates:

- After organizing historical data.
- The link ratios (Development Factors) must be computed:

  For each period t, calculate the ratio of mortality rates between consecutive ages:

  $$r_{x,t} = \frac{m_{x,t+1}}{m_{x,t}}$$

  These ratios measure how mortality rates develop from one year to the next.
- Calculate average development factors:

  For each age x, take the arithmetic mean (or volume-weighted average) of the observed ratios across all available years:

$$f_x = \frac{1}{n}\sum_{t=1}^{n} r_{x,t}$$

This gives a stable estimate of how mortality progresses with age.

- Project future mortality rates:
  Apply the development factors to the latest observed mortality rates.
- Complete the mortality triangle:
  Fill in missing values in the triangle using the projected rates.

2.6 Egyptian Mortality data

Steps to Apply Chain Ladder to Egyptian Life Insurance Data:

- **Collect and Organize the Data:**

Egyptian insurance companies' Data.

Collecting data from multiple insurance companies in Egypt can significantly enhance the accuracy and reliability of actuarial analyses, particularly for life insurance. By pooling data from various insurers, a more comprehensive picture of mortality rates, policyholder behavior, and claims patterns can be established, especially for niche segments or low-frequency events. This collaborative approach can also help identify emerging trends and regional variations that may not be evident in individual company data. However, challenges such as data standardization, privacy concerns, and competitive sensitivities must be addressed through frameworks like industry-wide databases or regulatory initiatives facilitated by the Egyptian Financial Regulatory Authority (FRA).

The available dataset includes the number of individuals living and the number of deaths, enabling the calculation of q_x (the probability of death at age *x*) using Equation $q_x = \frac{d_x}{l_x}$.

Subsequently, q_x is converted to m_x (the central death rate at age *x*) using Equation $\boldsymbol{m_{x,t} = \mu_{x,t} = -\ln(1 - q_{x,t})}$.

These m_x rates are then organized into **runoff triangles**, which serve as the basis for predicting future mortality rates. This structured approach allows for the projection of mortality trends over time, providing valuable insights for actuarial modeling and life insurance applications.

- **Calculate Development Factors:**

Compute age-to-age factors: For each development period, calculate the age-to-age factor (also called link ratio), which is the ratio of rates in one period to the previous period.

Calculate Average Development Factors: Compute the average (or weighted average) of the age-to-age factors for each development period.

- **Predicted future rates:**

Apply Development Factors: Use the average development factors to predict future rates for each year. Multiply the latest rate for each year by the appropriate development factor to estimate the next rates. Repeat this process until all future development periods are filled.

Complete the Triangle: Fill in the lower-right portion of the triangle (the "tail") with the predicted rates.

2.7 Insurance companies' data

Data for life insurance policies was compiled by several Egyptian insurance companies, and this data was provided in the form of annual individual records. The dataset included detailed information on the number of policies issued during the study period (2014–2022) and the number of policies that terminated due to the death of the policyholders. Importantly, the ages of individuals were recorded both at the time of policy issuance and at the time of death, allowing for precise age-specific analysis. This level of granularity enabled the direct calculation of q_x (the probability of death at age *x*) for each individual age without the need for additional adjustments or completion of the data, as is often required with census data.

The availability of individual-level data significantly streamlined the process of estimating mortality rates. Unlike aggregated or grouped data, which may require interpolation or extrapolation to fill gaps, the individual records provided a clear and accurate representation of mortality patterns across different ages. This allowed for the direct derivation of q_x for each age, ensuring that the mortality rates were both reliable and specific to the insured population. The elimination of the need to complete or adjust the data reduced potential sources of error and enhanced the accuracy of the results.

With q_x calculated for each age, the data was then used to derive other key metrics, such as the central death rate (m_x), which is essential for actuarial modeling and forecasting. The ability to work with individual ages rather than age groups provided a more detailed and nuanced understanding of mortality trends, which is critical for life insurance companies when setting premiums, calculating reserves, and managing risks. This approach also allowed for the analysis of gender-specific mortality rates, as the data could be segmented by male and female policyholders, reflecting the differences in mortality patterns between the two groups.

The use of individual-level data from Egyptian insurance companies not only improved the accuracy of the mortality rate calculations but also ensured that the results were directly applicable to the insured population. This is particularly important in the context of life insurance, where precise mortality estimates are essential for financial stability and regulatory compliance. By leveraging this detailed dataset, the study was able to provide valuable insights into mortality trends and support more informed decision-making for insurance companies operating in the Egyptian market.

Table (5): q_x for Males from insurance companies.

	c						
	2014	….	**2017**	….	2020	2021	2022
20	0.0014974	….	0.0011494	….	0.0011673	0.0011364	0.00106383
21	0.0015551	….	0.0012884	….	0.0012959	0.0012329	0.001156069
22	0.0016233	….	0.0013323	….	0.0014085	0.0013652	0.001304348
….	….	….	….	….	….	….	….
33	0.0028859	….	0.0020019	….	0.0017779	0.0016165	0.00159442
34	0.0029889	….	0.0020772	….	0.0018253	0.0016691	0.001612903
….	….	….	….	….	….	….	….
44	0.0045511	….	0.0031471	….	0.0029698	0.0028456	0.002870515
45	0.004788	….	0.0033421	….	0.0032124	0.0031165	0.003151907
….	….	….	….	….	….	….	….
59	0.0133147	….	0.0105949	….	0.0110736	0.0105758	0.010466369
60	0.0143727	….	0.0114485	….	0.0119527	0.0113605	0.011203575

After obtaining q_x from insurance companies, then m_x can be obtained through the equation $m_{x,t} = \mu_{x,t} = -\ln\left(1 - q_{x,t}\right)$

Table(6): $m_x * 10^{-3}$ for Males from insurance companies.

	m_x *10^{-3}						
	2014	….	**2017**	….	2020	2021	2022
21	1.5551386	….	1.2892307	….	1.2967404	1.2336606	1.156738
22	1.6233169	….	1.3331783	….	1.4094929	1.3661327	1.305199
….	….	….	….	….	….	….	….
32	2.791994	….	1.9341693	….	1.7417159	1.5921668	1.514213
33	2.8859002	….	2.0039265	….	1.7794823	1.6178079	1.595692
….	….	….	….	….	….	….	….
44	4.5511006	….	3.1520826	….	2.9742186	2.8496564	2.874643
45	4.7880143	….	3.3476772	….	3.2175708	3.1213664	3.156885
….	….	….	….	….	….	….	….
59	13.314698	….	10.651466	….	11.135369	10.632121	10.52153
60	14.372714	….	11.514518	….	12.024708	11.425523	11.26681

2.8 Estimation of death rate (m_x)using chain ladder from insurance companies' data:

We need to reorganize the data for m_x to make it suitable for the Chain Ladder method.

Now, m_x has been derived from $\boldsymbol{q_x}$, the data will be reorganized into a run-off triangle format to apply the Chain Ladder technique. This reorganization is essential for predicting future mortality rates, as the Chain Ladder method relies on the development patterns observed in historical data. To facilitate this process, ages will be divided into groups, and separate tables will be created for each age group. This division allows for a more structured and manageable application of the Chain Ladder method, ensuring that predictions are accurate and reflective of the underlying mortality trends.

Each age group will be analyzed independently, following the same steps previously applied to census data. The Chain Ladder technique will be used to predict future central death rates (m_x) for each group, based on the historical development patterns observed in the run-off triangles. These projections will provide insights into how mortality rates are expected to evolve over time for different age cohorts.

After completing all the tables and applying the Chain Ladder method to each age group, the final table of central death rates will be compiled. This table will summarize the predicted m_x values for all age groups, providing a comprehensive overview of the expected mortality rates across the entire age range under study. The final table serves as a key output of the analysis, offering valuable insights for insurance companies to use in their pricing, reserving, and risk management processes. By leveraging these projected rates, insurers can make more informed decisions and ensure the financial sustainability of their products.

After completing all the tables, the final table of the central death rates will be as follows:

Table (7): Predicted Mortality Rates ($m_x * 10^{-3}$) from insurance companies (Male) (Age 21-60).

Age	2023	2024	2025	2026	2027	2028	2029	2030
21	1.08							
22	1.25	1.17						
23	1.29	1.23	1.15					
24	1.28	1.22	1.17	1.09				
25	1.31	1.28	1.23	1.17	1.10			
26	1.28	1.26	1.23	1.18	1.13	1.05		
27	1.36	1.31	1.30	1.27	1.22	1.16	1.08	
28	1.36	1.31	1.27	1.25	1.23	1.17	1.12	1.04
29	1.41							
30	1.39	1.36						

31	1.42	1.34	1.31					
32	1.43	1.34	1.26	1.24				
33	1.54	1.45	1.36	1.28	1.25			
34	1.57	1.51	1.43	1.33	1.26	1.24		
35	1.57	1.53	1.47	1.39	1.30	1.22	1.19	
36	1.61	1.52	1.48	1.42	1.34	1.25	1.18	1.15
37	1.71							
38	1.75	1.69						
39	1.83	1.71	1.65					
40	1.97	1.84	1.72	1.66				
41	2.13	2.02	1.88	1.75	1.69			
42	2.36	2.25	2.13	1.99	1.85	1.79		
43	2.49	2.41	2.30	2.17	2.03	1.89	1.82	
44	2.70	2.55	2.47	2.36	2.23	2.09	1.94	1.87
45	3.19							
46	3.41	3.44						
47	3.76	3.71	3.74					
48	4.15	4.11	4.04	4.08				
49	4.56	4.52	4.48	4.41	4.45			
50	5.01	4.96	4.92	4.87	4.79	4.84		
51	5.38	5.33	5.28	5.24	5.18	5.10	5.14	
52	5.80	5.64	5.59	5.54	5.49	5.43	5.34	5.39
53	6.66							
54	7.06	7.11						
55	7.85	7.69	7.75					
56	8.53	8.53	8.36	8.42				
57	9.25	9.28	9.28	9.10	9.16			
58	9.90	9.97	10.00	10.01	9.81	9.88		
59	10.29	10.34	10.41	10.44	10.45	10.24	10.31	
60	10.91	10.67	10.72	10.80	10.83	10.84	10.62	10.69

After obtaining the central mortality rates for males, we can use the same method to obtain the mortality rates for females.

Table (8): q_x for Females from insurance companies.

	q_x						
Age	**2014**	….	**2017**	….	2020	2021	2022
21	0.000615764	….	0.00050063	….	0.000488	0.000483	0.000472
22	0.000680851	….	0.0005618	….	0.000511	0.000544	0.000348
….	….	….	….	….	….	….	….
33	0.001389646	….	0.00099824	….	0.00089	0.000812	0.000813
34	0.001475038	….	0.00105784	….	0.000947	0.000878	0.00086
….	….	….	….	….	….	….	….

41	0.002203879	….	0.00150294	….	0.001454	0.001449	0.001423
42	0.0023549	….	0.00160464	….	0.001587	0.001573	0.001538
….	….	….	….	….	….	….	….
50	0.004015664		0.0029801		0.003206	0.003198	0.003196
51	0.004306908		0.00324301		0.003491	0.003506	0.003493
….	….		….		….	….	….
59	0.008034084	….	0.00635116	….	0.006334	0.006272	0.006153
60	0.0086901	….	0.00688239	….	0.006758	0.006694	0.006545

Table (9): m_x for Females from insurance companies.

	$m_x * 10^{-3}$						
Age	**2014**	….	**2017**	….	2020	2021	2022
21	***0.6159532***	….	***0.5007554***	….	***0.4881191***	***0.4831167***	***0.471811***
22	***0.6810829***	….	***0.5619579***	….	***0.5111306***	***0.544148***	***0.348491***
….	***….***	….	***….***	….	***….***	***….***	***….***
33	***1.3906122***	….	***0.9987386***	….	***0.8903963***	***0.8123299***	***0.813751***
34	***1.4761268***	….	***1.0583999***	….	***0.9474487***	***0.8783857***	***0.86038***
….	***….***	….	***….***	….	***….***	***….***	***….***
41	***2.2063116***	….	***1.5040705***	….	***1.4550581***	***1.4500508***	***1.423853***
42	***2.3576772***	….	***1.6059288***	….	***1.5882606***	***1.5742385***	***1.539645***
….	***….***	….	***….***	….	***….***	***….***	***….***
51	***4.3162095***		***3.24828***		***3.4971078***	***3.5121604***	***3.499125***
52	***4.6484686***		***3.539025***		***3.8172765***	***3.8152689***	***3.79457***
….	***….***		***….***		***….***	***….***	***….***
59	***8.0665312***	….	***6.3714144***	….	***6.3541449***	***6.2917516***	***6.171837***
60	***8.7280788***	….	***6.9061829***	….	***6.7809387***	***6.7165053***	***6.566965***

After applying all the previous steps of the Chain Ladder method to the female mortality data, the final table of predicted rates has been generated.

The final table for females will be as follows:

Table (10): Predicted Mortality Rates ($m_x * 10^{-3}$) from insurance companies (Female) (Age 21-60).

Age	2023	2024	2025	2026	2027	2028	2029	2030
21	0.999							
22	0.303	0.641						
23	0.419	0.364	0.771					
24	0.562	0.477	0.414	0.877				
25	0.559	0.556	0.472	0.410	0.869			
26	0.581	0.548	0.546	0.463	0.402	0.852		
27	0.578	0.573	0.540	0.538	0.457	0.397	0.840	
28	0.578	0.572	0.566	0.533	0.531	0.451	0.392	0.830
29	0.603							

30	0.712	0.649						
31	0.678	0.668	0.609					
32	0.716	0.683	0.673	0.613				
33	0.795	0.750	0.715	0.705	0.642			
34	0.839	0.820	0.773	0.737	0.726	0.662		
35	0.864	0.843	0.823	0.777	0.741	0.730	0.665	
36	0.932	0.883	0.861	0.841	0.794	0.757	0.745	0.679
37	1.003							
38	1.113	1.078						
39	1.182	1.145	1.108					
40	1.267	1.217	1.178	1.141				
41	1.383	1.342	1.288	1.247	1.207			
42	1.512	1.469	1.425	1.368	1.324	1.282		
43	1.624	1.595	1.550	1.503	1.443	1.397	1.353	
44	1.762	1.702	1.672	1.624	1.575	1.512	1.464	1.418
45	2.029							
46	2.235	2.228						
47	2.412	2.418	2.411					
48	2.681	2.671	2.678	2.670				
49	2.961	2.959	2.948	2.955	2.947			
50	3.215	3.239	3.237	3.225	3.233	3.223		
51	3.428	3.444	3.469	3.467	3.454	3.462	3.452	
52	3.663	3.588	3.605	3.631	3.629	3.615	3.624	3.614
53	4.068							
54	4.368	4.344						
55	4.701	4.680	4.655					
56	5.081	5.062	5.039	5.012				
57	5.495	5.494	5.472	5.448	5.418			
58	5.800	5.855	5.854	5.831	5.805	5.773		
59	6.055	6.082	6.139	6.138	6.114	6.087	6.054	
60	6.373	6.252	6.280	6.339	6.338	6.314	6.285	6.251

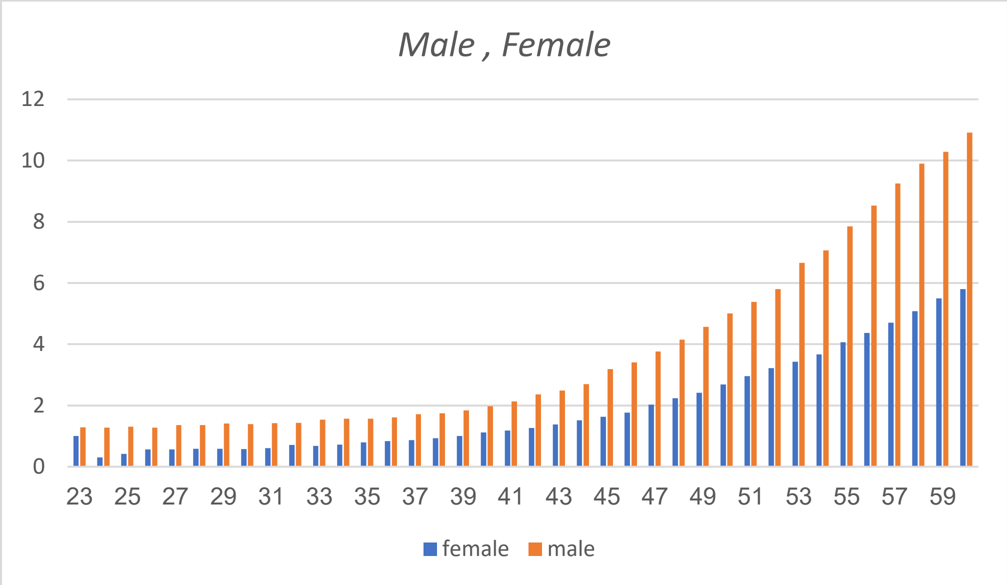

Figure (1): comparison between expected central death rate(m_x) for male and female in Egyptian companies.

Results of this research:

The results of Appling Chain ladder method of our estimation for m_x, using Insurance companies' data are:

1) *There is a significant improvement in mortality rates over the years (which is known as longevity).*
2) *Census data can't be relied upon for use in insurance companies.*
3) *From figure (1), Central mortality rates m_x for females are lower than central death rates $m_x$for males of all ages.*
4) *A combined table can't be used for both males and females, separate rates must be used for each because there are substantial differences between each of them.*

Insights and contributions for future research:

1) *Using the Bornhuetter-Ferguson (BF) method, the problem of outliers can be resolved, as predicted rates are based on the latest observations for each country.*
2) *By combining insights from insurance companies and census data, actuaries can develop more accurate and robust mortality models, ultimately improving risk assessment and decision-making.*
3) *Insurance companies should also consider incorporating census data to ensure their models account for underestimates of mortality rates.*

References:

- England, Peter D and Verrall, Richard J. "Stochastic Claims Reserving in General Insurance." *British Actuarial Journal* 8, Issue 3 (2002): 443-518.
- keyfitz, Nathan & Caswell, Hal ,“Applied mathematical demography*”, New York: Springer* 47, 2005.
- Oeppen, Jim and Vaupel, James. “Broken Limits to Life Expectancy”. *Science* 296, Issue 5570 (2002): 1029-1031.
- Pollard , John H. “Projection of Age-Specific Mortality Rates”. *Population Bulletin of the United Nations* 21, Issue *22* (1987): 55-69.
- Suwardi, Desnu Anggara and Purwono, Yogo. “The analysis of motor vehicle insurance claim reserve using robust chain ladder”, In *5th Global Conference on Business, Management and Entrepreneurship (GCBME 2020), Atlantis Press* 187 *(*2021*)*: 153–158.
- Vaupel, James W. “Biodemography of Human Ageing”, *Nature 464*, *Issue* 7288 (2010): 536-542.