I. Introduction to the variable annuity

1. Product design

A variable annuity is a contract between the policyholder and an insurance company. The insurer makes periodic payments to the policyholder, starting immediately or at a specific time. The policyholder can purchase a variable annuity contract with a single payment or a series of payments. A variable annuity has two phases: an accumulation phase and a payout phase.

During the accumulation phases, the policyholder makes payment to their annuity account, which can be allocated to several investment options, including mutual funds, stocks, bonds, or money market funds. The policyholder can allocate their payments among these options and tailor their investment strategy to their requirements and risk tolerance. Depending on the fund's performance, the money assigned to each investment option will increase or decrease over time.

The policyholder will receive purchase payments, investment income, and gains during the payout phase. They can receive them as a lump-sum payment or as a stream of payments at regular intervals. If they choose to receive a stream payment, they may have several choices regarding how long the payment will last or for the rest of the policyholder's life. The policyholder can also select either receiving payments that are fixed in amount or payments that vary based on the performance of mutual fund investment options. (U.S. Securities and Exchange Commission’s website, 2011)

One unique feature of the variable annuity is the guaranteed minimum income benefits. It ensures a particular minimum level of an annuity payment, even if insufficient money is in your account to support this payment (due to the investment losses). This payment protects the policyholder's investment, as they are guaranteed to receive a certain amount of money even if the fund performs poorly. One other common feature of a variable annuity is the death benefits. If the policyholder dies, a person selected as a beneficiary will receive more than all the money in the policyholder’s account or some guaranteed minimum payments. Long-term care insurance is also a benefit of the variable annuity, which pays for nursing home care or home health care if the policyholder becomes seriously ill.

2. Risk appetites

The variable annuities allow the policyholder to invest in different types of funds, providing a wide range of investment options to suit different risk appetites. For instance, policyholders with a high-risk appetite may invest in high-risk subaccounts such as stocks.

Policyholders with a lower risk appetite might invest their money in a more conservative investment strategy, such as bonds or the money market. However, research has shown that households with greater risk tolerance are more likely to purchase variable annuities than those with lower risk tolerance. (Brown & Poterba 2006)

The reason is variable annuities take a long time to maturity, and policyholders may need clarification on whether the insurance company will remain solvent when their contract reaches maturity. Moreover, even though there is a minimum guarantee payment, some policyholders still need to be more confident about the possibility of losing their principal investment. Policyholders are also concerned that the fees associated with variable annuities are higher than other investment products. The higher fee will reduce the value of their annuity and impact their overall return. Hence, those with higher-risk appetites are more likely to purchase variable annuities.

3. Pricing methods

There are two standard methods for evaluating the price of a variable annuity: no-arbitrage pricing and traditional actuarial pricing method. The no-arbitrage pricing states that two investments with the same outcomes should have the same initial cost. This approach determines the price of a complex financial instrument by using the prices of a portfolio of other assets that replicate the economic outcome of the instrument. This approach extends the classical equivalent premium principle: the expected present value of incomes equals the expected present value of the outgoes. The significant difference is that the expectation in the classical premium principle is computed based on the physical probability measures. In contrast, the anticipation in the no-arbitrage pricing is calculated based on the risk-neutral probability measure. (Feng et al. 2022)

The second approach is the extension of the traditional actuarial pricing method, whereby the insurance product prices are determined to adequately cover the insurance benefits with a level of assurance. To implement this approach, the insurers need to forecast the cash flows (inflows and outflows) throughout the life cycle of the insurance product and assess the profitability under different economic circumstances. This approach is the most popular in the insurance practice. (Feng et al. 2022)

II. Summary methods for estimating investment returns options and the healthy-ill health-death model.

The estimation of investment options returns and health-illness-death was done using the Maximum Likelihood transition matrix estimation. The transition matrix was then used to generate sample paths, from which the probability distribution was calculated. This approach has two main advantages. Firstly, generating transition matrices enables the probability of transitioning between different states to be captured. Secondly, accurate probability distributions and potential outcomes can be obtained by generating many sample paths.

The maximum likelihood estimation was utilised to estimate the transition matrices for the investment returns and health- ill health- death model. This method used the observed data to estimate the probabilities of moving from one state to another.

In the investment return, there are two options: defensive and growth. Each option has 150 data on the rate return each year. Therefore, the transition matrix for each option can be obtained from two data sets. Besides, demographic data can provide information on the living states (illness, health, death) of individuals aged 65 to 84 over 40 years. There are typically 50 to 60 data points for each age group, allowing for calculating a transition matrix for that specific age group.

However, some assumptions need to be made to obtain the transition matrix. For the investment return options data, there are three assumptions. Firstly, the data are collected annually, and there is no lapse. Secondly, the data are assumed to follow the Markov Chain. It means that the probability of transitioning to the next state depends on the current state and not any states in history. Thirdly, the Markov Chain is assumed to be time-homogeneous, which means the transition matrix does not change over time.

For the demography data, besides three assumptions similar to the investment return options (data were collected on an annual basis and no lapse, time-homogenous Markov Chain), there are three more assumptions. Fourthly, the absence of generation effects implies that each age group exhibits patterns or trends that are not shared with other groups. Fifthly, to use the MLE method, the demography Markov Chain is assumed to be ergodic, even in an absorbing state (death). Sixthly, it is assumed that there is no surrender during the 40-year data collection period.

From the transition matrix, 1000 sample paths will be generated for each investment option. The initial distribution for the investment return is assumed to be uniform. This decision is justified by the data from the "Defensive" and "Growth" options. The probability of getting -0.01 return rate from the "Defensive" option is 0.61133, while the probability of getting 0.05 is 0.38667. Similarly, the "Growth" option has a probability of 0.5333 for getting -0.05 and 0.4667 for getting 0.2. Hence, choosing a uniform distribution for the initial distribution can help avoid a bias towards a specific initial state.

One thousand sample paths will be generated for each age group using their respective transition matrices, with the initial state assumed to be healthy for all individuals. This decision is motivated by two factors related to health status. Firstly, the premium charges for insurance policies differ between health and ill policyholders, requiring separate pricing algorithms for each group. In this case, only healthy policyholders are considered. Secondly, as there is an absorbing state (death) in the model, setting the initial distribution to be uniform will cause some sample paths to immediately transition to the absorbing state, resulting in poor quality and non-representative sample paths.

III. Summary of the result obtained in the previous exercise.

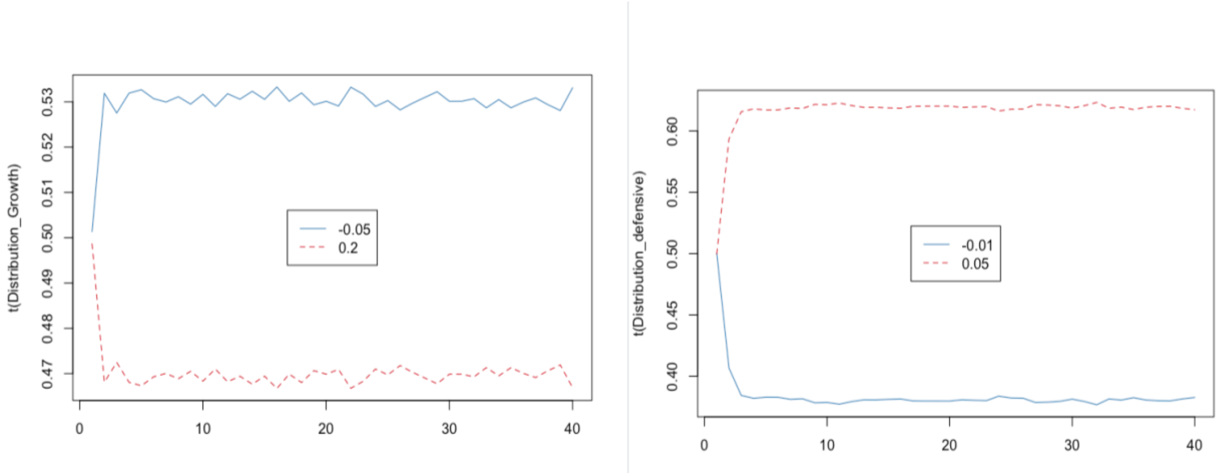

Figure 1: The probability distribution of investment return options for 40 year-period.

According to the graph, the growth option has a higher probability of negative returns (approximately 0.53) and a lower probability of positive returns (approximately 0.47). Conversely, the defensive option has a higher probability of positive returns (approximately 0.66) and a lower probability of negative returns (approximately 0.34). Additionally, the growth option has a higher positive rate of return and a lower negative rate than the defensive options.

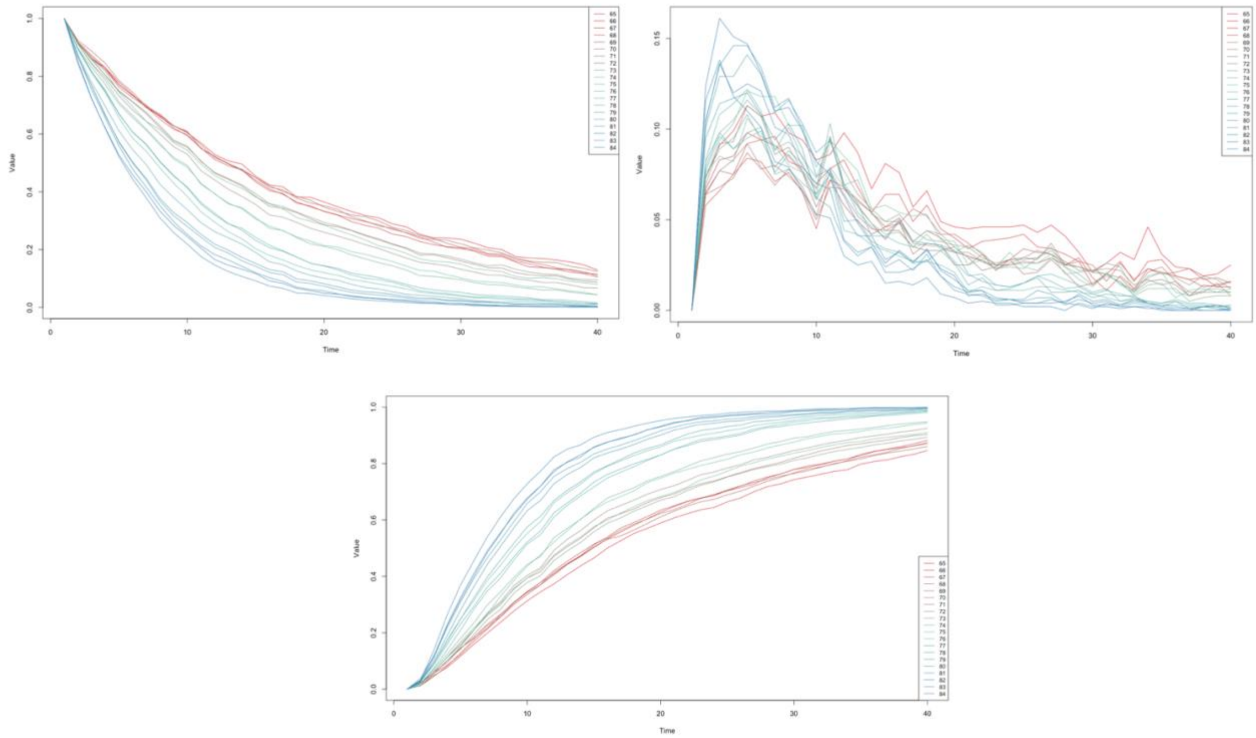

Figure 2a, b, c. The probability distribution of health, illness, and death for ages 65 to 84 years old for 40 years.

The graphs confirm the common understanding that younger policyholders are less likely to transfer to an ill state and tend to live longer. In comparison, older policyholders are more likely to move to the ill states and die sooner. The graphs also illustrate the generational effect across age groups. For instance, from age 65 to 68 (the red lines), there is a similar trend and probability magnitude across all three states (health, illness, and death).

IV. Pricing algorithm

a. A pricing algorithm for the variable annuity for retirees of different ages

After generating 1000 sample paths for each investment option and age group in the demographic data, the Monte Carlo method was used to price the variable annuity. The basic idea behind Monte Carlo methods is to generate many random samples from a probability distribution, which can then be used to approximate complex mathematical problems that are difficult or impossible to solve analytically. Monte Carlo methods are often used in pricing because they can handle complex and nonlinear models where it is challenging to obtain closed-form solutions using traditional mathematical methods. In addition, compared with a probability distribution, the Monte Carlo method allows the estimation of the probability of various outcomes. It assesses the potential risks and rewards associated with different investment strategies.

To apply the Monte Carlo method, each sample path for a particular age group was paired with one sample path for a specific investment option. This resulted in 1000 fund calculations for each investment option and 1000 benefit payments for each individual in each age cohort.

To determine the variable annuity's price for each cohort, the benefit payments for each year were discounted to their present value. The average of all possible values was then taken, resulting in the variable annuities for each cohort. Below is the algorithm for pricing the variable annuity for retirees at different age

$rate$ refers to the 1000 investment growth rate sample paths generated for each investment option over 40 years (defensive and growth option).

age refers to the 1000 health-ill-death sample paths generated for each age group (from 65 to 84 years old) over 40 years.

$fund$ refers to the 1000 fund value over 40 years after deducting for the retirement benefit or the first-time illness benefit paid to policyholder.

$benefit$ refers to the 1000 investment growth rate sample paths generated for each investment option over 40 years (defensive and growth option).

NPV refers to net present value of the benefit payment including retirement benefit, first-time illness benefit and death benefit for each observation for each age group.

Algorithm 1: Calculating fund value for 1000 observations for each age group

$fund \leftarrow matrix(0, nrow = 1000, ncol = 40)$
$for(i\, in\, 1 : 1000)$
$fund[i,\ 1] \leftarrow 0.95 \times P \times (1 + rate[i, 1])$
$for(j\, in\, 2 : 40)$
if $age[i,\ j] = 1$ **then**
  $fund[i,\ j] \leftarrow 0.95 \times fund[i, j-1] \times (1+rate[i, j]) - 0.03 \times fund[i, j-1]$
else
  if $age[i,\ j] = 2\, and\, all(age[i, 1 : (j-1)]! = 2))$ **then**
    $fund[i, j] \leftarrow 0.95 \times fund[i, j-1] \times (1 + rate[i, j]) - max(0.5 \times fund[i, j-1]; 0.5 * P)$
  else
    if $age[i,\ j] = 2\, and\, any(age[i,\ 1 : (j-1)] = 2))$ **then**
      $fund[i, j] \leftarrow 0.95 * fund[i, j-1] \times (1 + rate[i, j] - 0.03 \times fund[i.j-1]$
    else
      if $age[i, j] = 3$ **then**
        $break$
      end
    end
  end
end

Algorithm 2: Calculating benefit payment for 1000 observations for each age group

$benefit \leftarrow matrix(0, nrow = 1000, ncol = 40)$
$for(i\, in\, 1 : 1000)$
$for(j\, in\, 1 : 40)$
if $age[i,\ j] = 1$ **then**
  $benefit[i,\ j] \leftarrow 0.03 \times fund[i,\ j]$
else
  if $age[i,\ j] = 2\, and\, all(age[i, 1 : (j-1)]! = 2))$ **then**
    $benefit[i, j] \leftarrow max(0.5 \times fund[i, j-1],\ 0.5 \times P)$
  else
    if $age[i,\ j] = 2\, and\, any(age[i,\ 1 : (j-1)] = 2))$ **then**
      $benefit[i, j] \leftarrow 0.03 * fund[i, j]$
    else
      if $age[i, j] = 3$ **then**
        $benefit[i, j] \leftarrow max(fund[i, j-1], 0.3 * P)$
        $break$
      end
    end
  end
end

Algorithm 3: Calculating variable annuity price each age group

$NPV \leftarrow list()$
$for(i\, in\, 1 : 1000)$
  $for(j\, in\, 1 : 40)$
    $NPV[i] \leftarrow sum((benefit[i,]) * (1 + interest\, rate)^{(-j)})$
$price \leftarrow sum(unlist(NPV))/1000$

b. Justifications for the above calculation

Let P be the initial price and $r_{t,i}$ be the return rate of the investment option. The fund value in the first year can be calculated as follows:

$$S_1 = 0.95 \times P \times (1 + r_{1,i})$$

If the individual is healthy or has received their first-time illness benefit and gets sick again, the retirement benefit payment will be paid and deducted from the fund:

$$S_t = 0.95 \times S_{t-1} \times \left(1 + r_{1,i}\right) - 0.03 \times S_{t-1}$$

If the individual gets sick for the first time, the illness benefit will be paid and deducted from the fund.

$$S_t = 0.95 \times S_{t-1} \times \left(1 + r_{1,i}\right) - \max\left(0.5 \times S_{t-1}, 0.5 \times P\right)$$

Then the fund will be compared to the initial deposit to pay the benefit to the policyholder. The retirement benefit is:

$$Retirement\ benefit = 0.03 \times S_t$$

Let τ_1 is the first-time individual is ill; the first-time illness benefit is:

$$Illness\ benefit = \max\left(0.5 \times S_{\tau 1}, 0.5 \times P\right)$$

Let τ_2 is the time the individual dies, the death benefit is:

$$Death\ benefit = \max\left(S_{\tau 2}, 0.3 \times P\right)$$

To calculate the price of the variable annuity, the benefits paid each year are multiplied by the discount factor and then summed up. This calculation is done for individuals, and then the average value is taken to obtain the price of the variable annuity for each age group.

c. Assumption for the above calculation and justification

In addition to the assumptions made for estimating the investment options and the healthy-illness-death model, certain assumptions must be made for the pricing algorithm used above.

1. There are only two investment options: defensive and growth options and each option has only two growth rates. The defensive option has a return of {-0.01, 0.05} while the growth option has a return of {-0.05, 0.2}.

2. The illness benefit is only paid for the first time the individual gets ill, and the payment is 50% of the fund in the year the event happens, regardless of whether the payment is enough to cover all the individual's medical expenses.

3. The fund was calculated separately in two options, meaning no mixed options exist.

There are several reasons why the fund was calculated separately with two options. Firstly, the growth and defensive options have different levels of risk. The defensive option is more conservative, while the growth option is more aggressive. Therefore, keeping them separate makes it easier to compare their performance based on their respective risk profiles. Additionally, policyholders who choose the defensive option will always receive a first-time illness benefit of 0.5P. In contrast, those who prefer the growth option will receive a benefit that depends on their investment performance.

Secondly, calculating the performance of each investment option separately provides greater transparency for the overall fund. This can be helpful for investors who want to know exactly how their money is allocated and how each investment option is performing.

Thirdly, knowing how each investment option performs separately allows the marketing team to explain the fund's benefits more clearly to customers with different goals. For example, they can tailor their explanations to customers who want to invest in an insurance product or those primarily interested in ensuring their financial security.

4. The investment growth rate sample paths are independent and identically distributed (i.i.d) for each investment option over 40 years.

This means that the performance of the investment in one year does not depend on the performance of the investment in previous years and that the performance of the investment is not affected by external factors. This assumption is made to simplify the mathematical calculations needed to model investment behaviour. Another reason is that it helps to isolate the effects of changes in investment performance from external factors such as changes in the economic environment or political instability. By assuming that investment returns are independent of external factors, the effects of changes in investment performance on the fund's overall performance can be focused.

5. The health-ill-death sample paths are independent and identically distributed (i.i.d) for each age group (65 to 84 years old) over 40 years.

This means that an individual's health status in one year does not depend on their health status in previous years and that external factors do not affect their health status. In addition, this simplifies the modelling process because it allows us to use a single probability distribution to model the health-ill-death outcomes for each age group rather than each year separately.

6. All the benefits are assumed to be paid at the end of the year, which means there are no mid-year withdrawals or payouts.

The assumption that all benefits are paid at the end of the year simplifies the calculations and makes them easier to manage. It allows for a clear and consistent time frame for evaluating the fund's performance and projecting future payouts. It also assumes that the fund has access to all necessary resources at the end of each year to fulfil its obligations.

While mid-year withdrawals and payouts can certainly occur, assuming that all benefits are paid at the end of the year provides a reasonable baseline for calculating the fund's performance and projecting future payouts.

7. The interest rate used to calculate the net present value is constant across all ages and observations. This means the interest rate is assumed to remain the same throughout the entire investment period. Taking a constant interest rate across all ages and observations simplifies the net present value calculation. In addition, it allows for a consistent discounting of future cash flows to their current values, making it easier to compare the present value of cash flows across different age groups and investment options.

V. Calculation of the variable annuity price based on the specific data provided and the proposed algorithm

Assume the initial deposit is $1, and the interest rate is 1%. The obtained result is below:

	Growth	Defensive
65	1.517037	0.9469966
66	1.488261	0.9357471
67	1.457239	0.9073534
68	1.500749	0.9275139
69	1.481621	0.9452913
70	1.430559	0.9261717
71	1.442092	0.9484743
72	1.476865	0.9512319
73	1.413097	0.9537271
74	1.466501	0.9658479
75	1.453044	0.9570309
76	1.309069	0.9233107
77	1.385018	0.9632665
78	1.342323	0.9348702
79	1.314892	0.9497804
80	1.22227	0.9101848
81	1.28611	0.9499238
82	1.258983	0.9485395
83	1.201908	0.9279594
84	1.238382	0.9402878
Price	1.384301	0.94067546

VI. A summary of sensitivity analysis findings

1. The accuracy and amount of data used to generate the probability transition matrix and investment return sample paths significantly impact the accuracy of the resulting price. Different data sets can lead to different prices.

2. The investment option the policyholder selects also affects the product's price. The growth option is more volatile, leading to potentially higher returns and risk, whereas the defensive option is more stable, with more modest returns but lower risk.

3. The interest rate used to calculate the net present value of all benefits is also a significant factor in determining the price. A higher interest rate leads to a lower net present value of all benefits, and vice versa.

4. The policyholder's age is another important factor in determining the price. As the policyholder ages, the expected lifespan decreases, leading to a lower expected payout and price. This effect is more pronounced in the growth option, where the payouts are more dependent on investment returns than in the defensive option, where the payouts are more fixed.

5. The insurance company's expense to the policyholder can affect the price.

VII. A discussion of the limitations of the model, including exogenous risks not covered in the structure

The first limitation of the model is its assumption of a Markov chain and time-homogeneous process. The returns on an investment option and an individual's health conditions can be inherited from the past. Furthermore, the probability of transitioning from one state to another can also change over time. For example, as people age, they may have a higher probability of getting ill or dying than when they were younger. In addition, investment returns can be affected by both systematic and idiosyncratic risks, which can change over time and are not constant. The returns may also vary over time and are not restricted to just two rates for each option. Therefore, the model's assumption of independence between periods may not always hold in the real world, which can result in inaccuracies in the predictions made by the model**.**

The second limitation of the model arises from the assumption of ergodicity when estimating the transition probability matrix for the health-illness-death model. The assumption assumes that the Markov chain can reach every state from every other state. However, the chain is not ergodic in the presence of a death state, which is an absorbing state. The transition probability matrix derived from this assumption is used to generate the sample paths, but many of them would reach the absorbing state too early, leading to low accuracy in the data and pricing results.

The third limitation of the model is that it does not capture the generation effect, which refers to the impact of being born in a particular time period on an individual's characteristics, behaviour, and outcomes. For instance, individuals born during the same period may share common experiences, such as exposure to the same historical events, cultural influences, and economic conditions, that can shape their attitudes, preferences, and expectations. However, the generation effect can be observed in the data, as shown in Figure 2, where people of the same age group exhibit similar trends in health status and mortality rates. Therefore, instead of charging the same price for all individuals with different ages, the model could be improved by capturing the generation effect and charging different prices for different age cohorts. This approach may lead to more accurate pricing and better risk management, as it would reflect each generation's unique characteristics and risks. Moreover, it could increase customer satisfaction by providing more personalised and fair pricing for everyone.

The fourth limitation of the model is that it assumes a constant interest rate and a fixed investment option for the entire investment period. However, interest rates and market conditions can fluctuate and change over time, leading to different investment returns and ultimately affecting the price of the annuity product. Besides, if it does not capture the effect of inflation, then it may underestimate the actual cost of the product and lead to lower-than-expected benefit payments for the policyholder. Additionally, this model only calculates the price of benefits paid at the end of each year, whereas, in reality, claims can be made at any time. This limitation may result in the model producing inaccurate prices that do not reflect the real-world scenario of claim payments.

The fifth limitation of the model is the policyholder will not make any mid-year withdrawals or payouts, which may not reflect reality. Policyholders may need to withdraw funds due to unforeseen circumstances, which could affect the payout structure and, ultimately, the price of the annuity product.

Sixthly, the model does not account for exogenous risks that could impact individuals' investment returns and health conditions. Examples of such risks include economic downturns, pandemics, and natural disasters. These risks can have significant impacts on the returns of investment options and the health of individuals and could lead to a higher-than-expected payout for the insurance company.

The seventh limitation of the model is the need for more consideration of competition from other companies and internal company factors. For example, if the price set for the insurance product is too high, it may deter potential customers from purchasing the product. On the other hand, if the price is low, the company may be able to manage the influx of customers. Therefore, it is essential to consider these external and internal factors when determining the pricing strategy for the product.

Lastly, the model does not consider the tax or the possibility of changes in government policies and regulations, which could impact the investment options and the payout structure of the variable annuity. Such changes could also result in unexpected outcomes and affect the accuracy of the model's predictions.

IIX. Declaration of using A.I. tools.

The student used chatGPT to address any unclear issues and to improve the grammar and wording of the text in the assignment. The student conducted all aspects of the assignment, including literature review, calculation methods, assumptions and results.

REFERENCE

Office of Investor Education and Advocacy 2011, ‘Variable Annuities: What You Should Know’, *U.S. Securities and Exchange Commision*, 04/18/2011, <https://www.sec.gov/investor/pubs/varannty.htm>

Brown, J.R. & Poterba, J.M. 2006, ‘Household Ownership of Variable Annuities’, *Tax Policy and Economy*, vol. 20, no. 36, pp. 163-191.

Feng, R., Gan, G. & Zhang, N. 2022,’Variable annuity pricing, valuation, and risk management: a survey, *SCANDINAVIAN ACTUARIAL JOURNAL*, vol. 2022, no. 10, 867-900, pp. 9-10.

APPENDIX

```r
library(readxl)
library(dplyr)
library(ggplot2)
library(readxl)
set.seed(2023)
demography <- read_excel("~/Desktop/ETC3430/Assignment/data.xlsx",
                         sheet = "Demography data")
#1. TRANSITION PROBABILITY MATRIX
#Function to calculate the transition matrix at each age
transition_matrix_demography <- function(age) {
  #Extract the data from the data frame
  age_jan80 <- demography$`Age at Jan 1980`
  paths <- demography %>% filter(age_jan80 == age) %>% select(-1) %>% slice(-1)
  list_sample_paths <- list()

  for (i in 1:nrow(paths)) {
    list_sample_paths[i] <- list(as.vector(unname(unlist(paths[i,]))))
  }

  #Count values for calculating the transition matrix
  states <- sort(unique(unlist(list_sample_paths)))
  a <- length(list_sample_paths)
  n <- length(states)
  P <- matrix(NA, nrow = n, ncol = n)

  #Calculate the transition matrix
  for (j in 1:n) {
    for (k in 1:n) {
      num <- 0
      for (i in seq_along(list_sample_paths)) {
        for (l in 2:length(list_sample_paths[[i]])) {
          if (list_sample_paths[[i]][l - 1] == states[j] &
              list_sample_paths[[i]][l] == states[k])
            num <- num + 1
        }
        P[j, k] <- num
      }
    }
    P[j, ] <- P[j, ] / sum(P[j, ])
  }
  return(P)
}

transition_matrix_demography(84)
#2.GENERATE SAMPLE PATH + CALCULATE DISTRIBUTION
#Generate the path from the transition matrix
path_age <- function (age) {
  states_age <- 1:3
  set.seed(2023)
  length <- 40 # time periods
  N_age <- 1000 # number of paths
  path_age <- matrix(NA, nrow = N_age, ncol = length)
  path_age[, 1] <- 1 #state at state health
  P_age <- transition_matrix_demography(age)
  for (j in 1:N_age) {
    for (i in 2:length) {
      path_age[j, i] <-
        sample(x = states_age,
               size = 1,
               prob = P_age[path_age[j, i - 1], ])
    }
  }
  return(path_age)
}
```

```r
#1. TRANSITION MATRIX
#Function to calculate the transition matrix for two investment option
transition_matrix_investment <- function(option) {
  return_vector <- unname(unlist(return[, option]))
  states <- sort(unique(return_vector))
  n <- length(states)
  P <- matrix(NA, nrow = n, ncol = n)
  for (j in 1:n) {
    for (k in 1:n) {
      num <- 0
      for (i in 2:length(return_vector)) {
        if (return_vector[i - 1] == states[j] &
            return_vector[i] == states[k])
          num <- num + 1
      }
      P[j, k] <- num
    }
    P[j, ] <- P[j, ] / sum(P[j, ])
  }
  return(P)
}
transition_matrix_investment("Growth")
transition_matrix_investment("Defensive")

#2.GENERATE SAMPLE PATH
#Generate 100000 sample paths for GROWTH
return_vector_growth <- unname(unlist(return[, "Growth"]))
#Count value for calculating the transition matrix
states_growth <- c("-0.05", "0.2")
length <- 40 # time periods
N <- 1000 # number of paths
path_growth <- matrix(NA, nrow = N, ncol = length)
path_growth[, 1] <-
  sample(x = states_growth, N, replace = TRUE) #uniform distribution
P_Growth <- transition_matrix_investment("Growth")
row.names(P_Growth) <- c("-0.05", "0.2")
colnames(P_Growth) <- c("-0.05", "0.2")
for (j in 1:N) {
  for (i in 2:length) {
    path_growth[j, i] <-
      sample(x = states_growth,
             size = 1,
             prob = P_Growth[path_growth[j, i - 1], ])
  }
}
path_growth <- apply(path_growth, c(1, 2), as.numeric)


#Generate 10000 sample paths for DEFENSIVE
return_vector_defensive <- unname(unlist(return[, "Defensive"]))
#Count value for calculating the transition matrix
states_defensive <- c("-0.01", "0.05")
length <- 40 # time periods
N <- 1000 # number of paths
path_defensive <- matrix(NA, nrow = N, ncol = length)
path_defensive[, 1] <-
  sample(x = states_defensive, N, replace = TRUE) #uniform distribution
P_defensive <- transition_matrix_investment("Defensive")
rownames(P_defensive) <- c("-0.01", "0.05")
colnames(P_defensive) <- c("-0.01", "0.05")
for (j in 1:N) {
  for (i in 2:length) {
    path_defensive[j, i] <-
      sample(x = states_defensive,
             size = 1,
             prob = P_defensive[path_defensive[j, i - 1], ])
  }
}
path_defensive <- apply(path_defensive, c(1, 2), as.numeric)
```

```
#AGE
age <- path_age(65)

defensive_fund_value <- matrix(0, nrow = 1000, ncol = 40)
for (i in 1:1000) {
  defensive_fund_value[i, 1] <- 0.95 * 1 * (1 + path_defensive[i, 1])
  for (j in 2:40) {
    if (age[ i, j] == 1)  {
      defensive_fund_value[i, j] <- 0.95 * defensive_fund_value[i, j-1] * (1 + path_defensive[i, j])
                                    - 0.03 * defensive_fund_value[i, j-1]
    }
    if (age[i, j] == 2 && all(age[i, 1:(j - 1)] != 2)) {
       defensive_fund_value[i, j] <- 0.95 * defensive_fund_value[i, j-1] * (1 + path_defensive[i, j])
                                    - max(0.5 * defensive_fund_value[i, j-1], 0.5 * 1)
    }
    if (age[i, j] == 2 && any(age[i, 1:(j - 1)] == 2)) {
      defensive_fund_value[i, j] <- 0.95 * defensive_fund_value[i, j-1] * (1 + path_defensive[i, j])
                                     - 0.03 * defensive_fund_value[i, j-1]
    }
    if (age[i, j] == 3) {
      break
    }
  }
}
defensive_fund_value

growth_fund_value <- matrix(0, nrow = 1000, ncol = 40)
for (i in 1:1000) {
  growth_fund_value[i, 1] <- 0.95 * 1 * (1 + path_growth[i, 1])
  for (j in 2:40) {
    if (age[ i, j] == 1)  {
      growth_fund_value[i, j] <- 0.95 * growth_fund_value[i, j-1] * (1 + path_growth[i, j])
      - 0.03 * growth_fund_value[i, j-1]
    }
    if (age[i, j] == 2 && all(age[i, 1:(j - 1)] != 2)) {
      growth_fund_value[i, j] <- 0.95 * growth_fund_value[i, j-1] * (1 + path_growth[i, j])
      - max(0.5 * growth_fund_value[i, j-1], 0.5 * 1)
    }
    if (age[i, j] == 2 && any(age[i, 1:(j - 1)] == 2)) {
      growth_fund_value[i, j] <- 0.95 * growth_fund_value[i, j-1] * (1 + path_growth[i, j])
      - 0.03 * growth_fund_value[i, j-1]
    }
    if (age[i, j] == 3) {
      break
    }
  }
}
growth_fund_value
```

```
#BENEFIT
benefit_growth_matrix <- matrix(0, nrow = 1000, ncol = 40)
for (i in 1:1000) {
  for (j in 1:40) {
    if (age[ i, j] == 1)  {
      benefit_growth_matrix[i, j] <- 0.03*growth_fund_value[i, j]
    }
    if (age[i, j] == 2 && all(age[i, 1:(j - 1)] != 2)) {
      benefit_growth_matrix[i, j] <- max(0.5 * growth_fund_value[i, j], 0.5 * 1)
    }
    if (age[i, j] == 2 && any(age[i, 1:(j - 1)] == 2)) {
      benefit_growth_matrix[i, j] <- 0.03*growth_fund_value[i, j]
    }
    if (age[i, j] == 3) {
      benefit_growth_matrix[i, j] <- max(growth_fund_value[i, j-1], 0.3)
      break
    }
  }
}
benefit_growth_matrix

net_present_value <- list()
for (i in 1:1000) {
  for (j in 1:40) {
  net_present_value[i] <- sum((benefit_growth_matrix[i, ])*(1+0.01)^(-j))
  }
}
sum(unlist(net_present_value))/1000

benefit_defensive_matrix <- matrix(0, nrow = 1000, ncol = 40)
for (i in 1:1000) {
  for (j in 1:40) {
    if (age[ i, j] == 1)  {
      benefit_defensive_matrix[i, j] <- 0.03*defensive_fund_value[i, j]
    }
    if (age[i, j] == 2 && all(age[i, 1:(j - 1)] != 2)) {
      benefit_defensive_matrix[i, j] <- max(0.5 * defensive_fund_value[i, j], 0.5 * 1)
    }
    if (age[i, j] == 2 && any(age[i, 1:(j - 1)] == 2)) {
      benefit_defensive_matrix[i, j] <- 0.03*defensive_fund_value[i, j]
    }
    if (age[i, j] == 3) {
      benefit_defensive_matrix[i, j] <- max(defensive_fund_value[i, j-1], 0.3)
      break
    }
  }
}

net_present_value_def <- list()
for (i in 1:1000) {
  for (j in 1:40) {
    net_present_value_def[i] <- sum((benefit_defensive_matrix[i, ])*(1+0.01)^(-j))
  }
}
sum(unlist(net_present_value_def))/1000
```