Group Coursework Submission Form

Specialist Masters Programme

Please list all names of group members: (Surname, first name) 1. Raghuvanshi, Pawni 2. Heijliger-Krogulski, Boris 3. Wu, Zhuojun	4. 5. **GROUP NUMBER:**	**F**
MSc in: Actuarial Science		
Module Code: SMM048		
Module Title: Insurance Risk Modelling (Subject CS2)		
Lecturer: Prof. Michail Chronopoulos	**Submission Date: 29/03/2024**	

Declaration:

By submitting this work, we declare that this work is entirely our own except those parts duly identified and referenced in my submission. It complies with any specified word limits and the requirements and regulations detailed in the coursework instructions and any other relevant programme and module documentation. In submitting this work we acknowledge that we have read and understood the regulations and code regarding academic misconduct, including that relating to plagiarism, as specified in the Programme Handbook. We also acknowledge that this work will be subject to a variety of checks for academic misconduct.

We acknowledge that work submitted late without a granted extension will be subject to penalties, as outlined in the Programme Handbook. Penalties will be applied for a maximum of five days lateness, after which a mark of zero will be awarded.

Marker's Comments (if not being marked on-line):

Deduction for Late Submission of assignment:

For Students:
Once marked please refer to Moodle for your final coursework grade, including your Peer Assessment grade.

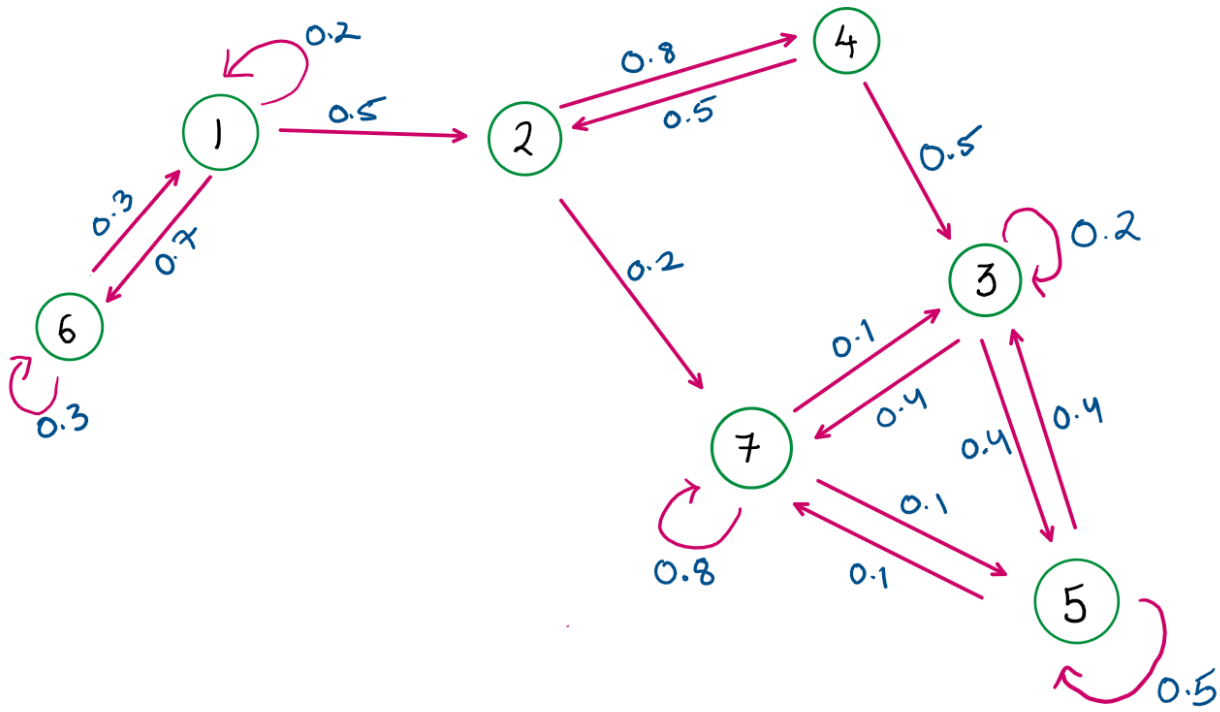
SMM048 Coursework 2
Group F

Group F

Question 1

i.

a. We are required to make a transition diagram for this part of the question:



b. The Commutating Classes that we observe are the follows:

1. {1,6}
   This forms a commutating class as both 1 and 6 can access each other.
   This forms an aperiodic class since if we start at 1, 6 comes at all following stages of the trajectory.
   This is a transient class since if we leave it once we cannot enter again.
2. {2,4}
   This forms a commutating class as both 2 and 4 can access each other.
   This forms a periodic class since if we start at 2, 4 and 2 come at following stages at a period of two in the trajectory.
   This is a transient class since if we leave it once we cannot enter again.
3. {3,5,7}
   This forms a commutating class all three both 3, 5, and 7 can access each other.
   This forms an aperiodic class since if we start at any of 3, 5, or 7, all of them come at all following stages of the trajectory.
   This is a recurrent class since if we enter it once we cannot leave again.

c. 1. To Prove this, we'll use the definition of a stationary distribution. Given that the chain converges to states (3,5,7):
Base Case (n=1):
Since X_0 follows the stationary distribution π,

$$P(X_1 = i) = \sum_{j\in\{3,5,7\}} P(X_1 = i|X_0 = j)P(X_0 = j) = \sum_{j\in\{3,5,7\}} P_{ij}\pi_j = \pi_i$$

Now assume that $P(X_n = i) = \pi_i$ holds for some $n \geq 1$
Then

$$P(X_{n+1} = i) = \sum_{j \in \{3,5,7\}} P(X_{n+1} = i | X_n = j) P(X_n = j) = \sum_{j \in \{3,5,7\}} P_{ij}\pi_j = \pi_i$$

Therefore, by induction, $(X_n = i) = \pi_i \forall\, n > 0$

2. Since the entire Markov chain is not ergodic, we cannot guarantee that P$(X_n = j | X_0 = i)$ approaches π_j as $n \to \infty$ for every value of j. Specifically, if j outside the subset {3, 5, 7}, it may never reach these states, and the limiting behaviour will not necessarily match the stationary distribution. While the subset {3, 5, 7} converges to its stationary distribution according to the ergodic theorem which states that:
If the discrete-time Markov chain X is ergodic, with transition matrix P, then there is exactly one probability vector π, called the equilibrium probability vector, which satisfies π^T P = π^T . In addition, for each i and j, lim n→∞ P[Xn = j|X0 = i] = π_j and π_j represents the long-run proportion of time spent in state j, when the chain satisfies the condition of irreducibility and aperiodicity.
Even though the subset {3, 5, 7} satisfies these conditions the entire Markov chain does not. So the stationary distribution we can possibly have is where $\pi_j = 0$ for values of j outside the subset {3, 5, 7}.

d. Here we are supposed to find a probability vector π such that we get $\pi^T = P\pi^T$
By ergodic theorem we have that states 3, 5, and 7 are irreducible and aperiod therefore our vector $\pi^T = (\pi_3, \pi_5, \pi_7)$

The probability matrix is P = $\begin{matrix} 0.2 & 0.4 & 0.4 \\ 0.4 & 0.5 & 0.1 \\ 0.1 & 0.1 & 0.8 \end{matrix}$

So the equation we have is as follows:

$$(\pi_3, \pi_5, \pi_7) = (\pi_3, \pi_5, \pi_7) * \begin{matrix} 0.2 & 0.4 & 0.4 \\ 0.4 & 0.5 & 0.1 \\ 0.1 & 0.1 & 0.8 \end{matrix}$$

Using this we get three equations:

$$0.2\pi_3 + 0.4\pi_5 + 0.1\pi_7 = \pi_3$$
$$0.4\pi_3 + 0.5\pi_5 + 0.1\pi_7 = \pi_5$$
$$0.4\pi_3 + 0.1\pi_5 + 0.8\pi_7 = \pi_7$$

Apart from these 3 equations we also have:

$\pi_3 + \pi_5 + \pi_7 = 1$ (which comes from the fact that sum of all probabilities is = 1)

Solving these 4 equations we get our vector:

$$\pi^T = \left(\frac{1}{5}, \frac{4}{15}, \frac{8}{15}\right)$$

ii.

a. $$p_{ss}(t) = \int_0^t p_{\overline{ss}}(u) \cdot (3.9) \cdot p_{Hs}(t-u) \cdot du + p_{\overline{ss}} \cdot dt$$

$$\Rightarrow \qquad p_{ss}(t) = \int_0^t e^{-4u} \cdot (3.9) \cdot p_{Hs}(t-u) \cdot du + e^{-4t}$$

b. Let the probability of a person staying healthy for at least 10 years and then experiencing sickness and eventually leaving be denoted by p_{HR}, then we have:

$$p_{HR}(t) = \int_{10}^{\infty} p_{\overline{HH}}(u) \cdot (0.12) \cdot p_{Hs} \ \cdot \ p_{sR} \cdot du$$

$$= \int_0^t e^{-0.12u} \cdot (0.12) \cdot \frac{0.1}{0.12} \cdot \frac{0.1}{4} du$$

$$= \int_0^t e^{-0.12u} \cdot 0.0025 \, du$$

$$= 0.0025 \ \cdot \ -\frac{e^{-0.12u}}{-0.12}\Big| 10 - \ \infty$$

= 0.006272916

c. The benefit is paid when the person is sick, so there is a money outflow that happens during the time of sickness. There is inflow of money when the person is healthy. To ensure that the company remains solvent, using probabilities $p_{Hs}(t)$ *and* p_{HH}(t) we can calculate the expected value of the time person would spend in a healthy state E[H_T] and the time a person would spend in a sick state E[S_t].
Now we equate the expected cash inflow and the expected cash outflow as:

E[H_t]*c = E[S_t]*b

Where LHS is expected inflow and RHS is expected outflow.
We use this equation to get a ratio of b and c which can be used by the company to cut costs equally. To be prudent we take a higher value of c to b.

Question 2

i) We are given the Hazard Function $\mu_x = 0.00018(1.06^x)$, representing the cumulative probability for a part to fail. To represent this as a survival function $S(x)$, we realise that $S(x) = e^{-\int_0^x \mu_u \, du}$. To answer this question, we use $x = 60$, to give us

$S(60) = e^{-\int_0^{60} \mu_u \, du} = e^{-\int_0^{60} 0.00018(1.06^u)\, du}$.

We isolate $\int_0^{60} 0.00018(1.06^u)\, du = 0.00018 \int_0^{60} 1.06^u \, du = 0.00018 \left[\frac{1.06^u}{\ln(1.06)}\right]_{u=0}^{u=60} =$

$0.101903 - 0.003089 = 0.098814$.

Therefore, $S(60) = e^{-0.098814} = 0.905911$, so there is a 90.59% that the chosen part will survive after 60 hours of testing.

ii) Let $S(50|20)$ be the survival function that a part works after 50 hours given that it has already worked for 20 hours.

Using probability, $S(50|20) = \frac{S(50)}{S(20)}$.

$$S(50) = e^{-0.00018\left[\frac{1.06^u}{\ln(1.06)}\right]_{u=0}^{u=50}} = e^{-(0.056902-0.003089)} = e^{-0.053813} = 0.947609.$$

$$S(20) = e^{-0.00018\left[\frac{1.06^u}{\ln(1.06)}\right]_{u=0}^{u=20}} = e^{-(0.009907-0.003089)} = e^{-0.006818} = 0.993205.$$

So $S(50|20) = \frac{S(50)}{S(20)} = \frac{0.947609}{0.993205} = 0.954092$. So, there is a 95.4092% chance that a chosen part will work after 50 hours of testing, given that it works after 20 hours.

iii) For this question, we need to find the value for x for which

$$S(x) = 0.5$$

$$\Rightarrow S(x) = e^{-0.00018\left[\frac{1.06^u}{\ln(1.06)}\right]_{u=0}^{u=x}} = e^{-0.00018\left(\frac{1.06^x-1}{\ln(1.06)}\right)} = e^{-\frac{0.00018}{\ln(1.06)}(1.06^x-1)}$$

$$= e^{-\frac{0.00018}{\ln(1.06)}(1.06^x)+\frac{0.00018}{\ln(1.06)}} = e^{-\frac{0.00018}{\ln(1.06)}(1.06^x)} \times e^{\frac{0.00018}{\ln(1.06)}}.$$

Rearranging this, we get

$$e^{-\frac{0.00018}{\ln(1.06)}(1.06^x)} = S(x) \times e^{-\frac{0.00018}{\ln(1.06)}} = 0.5 \times e^{-\frac{0.00018}{\ln(1.06)}} = 0.5 \times e^{-0.003089}.$$

We can take the natural log of each side, we get

$$\ln\left(e^{-0.003089(1.06^{x})}\right) = \ln(0.5 \times e^{-0.003089})$$

So $-0.003089(1.06^{x}) = \ln(0.5) - 0.003089$

$\Leftrightarrow 1.06^{x} = \frac{\ln(0.5)}{-0.003089} + 1 = 225.392095 \Leftrightarrow x = \frac{\ln(225.392095)}{\ln(1.06)} = 92.98.$

So it would take 92.98 hours for roughly 50% of the new parts to survive testing, which ultimately means that 50% of the new parts malfunction after this amount of time.

iv) A possible form of the hazard function that accounts for real-world factors could be: $\mu_x = 0.00018 \times 1.06^{\beta x + \gamma D + \delta Q}$, where β is the coefficient for the time-dependent effect, γ is the coefficient for the driving style factor D, and δ is the coefficient for the road quality factor Q. The Factor D would be increased if the driving style was worse, such as being more aggressive. The Factor Q would increase when the road conditions are poorer, like cracks or potholes.

v) The baseline function, denoted $h_0(t)$, is defined as being the function when all the covariates are 0, namely $x_i = 3, g_i = c_i = 0$.

Therefore $h_0(t) = e^{-0.12(3-3)+0.3(0)-0.2(0)} = 1$.

Note that the baseline function is defined to be a driver who drove their car only in the city, serviced at designated time intervals, and drives for 3 hours in each individual session.

vi) In the first part, $x_i = 2$, $g_i = 0$, $c_i = 1$, giving us $e^{-0.12(2-3)+0.3(0)-0.2(1)} = e^{-0.08}$.

For the second part, $x_j = 4$, $g_j = 1$, $c_j = 1$, $e^{-0.12(4-3)+0.3(1)-0.2(1)} = e^{-0.02}$.

We can now compare the models by comparing the ratio.

$$\frac{h_i(t)}{h_j(t)} = \frac{h_0(t)e^{-0.08}}{h_0(t)e^{-0.02}} = e^{-0.06} = 0.9418,$$

where $h_i(t)$ is the first driver and $h_j(t)$ is the second.

Therefore $S_i(t) = \left(S_j(t)\right)^{0.9418} \Leftrightarrow S_i(t) > S_j(t)$ for all $t > 0$. The survival function for the first driver is always greater than the second driver.

vii) The high level of standard error could be caused by the variability in the data that is not accounted for by the current parameters. This could be due to the dataset collected being too small, or having lots of variability which the model cannot account for, such as outliers.

Including driving style as a parameter could reduce the standard error by accounting for more variability in how different driving behaviours impact the parts. Driving style is a significant factor that can affect the wear and tear of car parts, and its inclusion could lead to a more accurate and predictive model. However, Quantifying driving style can be challenging because it's a qualitative factor and can vary greatly among individuals. Since the data is self-reported by drivers, it can lead to subjective and biased answers. Also, fitting the parameter could be difficult due to the multidimensional nature of driving style, which includes speed, acceleration patterns, braking habits, and more.

viii) **Road Conditions**: This covariate could account for the impact of road conditions on car part longevity. Poor road conditions can lead to increased wear and tear, potentially reducing the part's lifespan. However, quantifying road conditions may require extensive data collection on regional infrastructure, which would be incredibly time-consuming.

Weather Conditions: Extreme temperatures and weather events can affect car part performance. Including this covariate could help model the environmental stressors that parts are subjected to. However, weather is highly variable and localised, making it difficult to correlate directly with part failure without comprehensive, geographically tagged data.

Question 3

i) In all urban areas, mortality rates are computed by each medical district and factored into the total population with particular consideration on demography, lifestyle, healthcare and environmental determinants. So, even though these may greatly vary, they can still affect either the rates of mortality or the expected years in the life of a person. Tailoring to each district death rate tables for better forecasting of the future population thanks to which it is possible to adjust hte resources allocated to healthcare and social services at the local level.

ii) The approach taken to create mortality tables is by applying graduation through parametric function for the groups of younger age to be able to properly model the mortality rates variation across the age groups to help in making fine adjustments in each of the districts. For the oldest age class, using of the country standard mortality table offers a viable option to the case of not enough data which is highly applicable for modeling trend or mortality in an individual district. This would ensure reliability and consistency somewhere the district data will be far from sufficient. Such a hybrid type makes fine detail showing inner insights and general broader knowledge on oldest of citizens stable.

iii)

a. Chi-squared Test: Assesses the goodness of fit of the graduated rates to the observed data.
Standardised Deviations Test: Evaluates the variance of the observed data from the expected data to check if it's within an acceptable range.
Signs Test: Determines if there's a systematic bias in the graduation, by counting the number of positive and negative deviations.
Grouping of Signs Test: Checks for the randomness in the sequence of positive and negative deviations, indicating whether deviations are clustered or evenly distributed.
Cumulative Deviation Test: Examines the cumulative effect of deviations across the whole range and in specified sub-ranges (half and quarter ranges) to identify patterns or trends in the data.

b. Chi-square statistics: We got the 24.05: we can infer that if the Chi-squared statistic exceeds the critical value, the fit of the graduated rates to the actual data is poor.

Standardised Deviations Test: The standard deviations equate the magnitude of actual observations to the effectiveness of the standard deviations against the expectations, while considering the variation that is expected. This test did not present a statistic that predetermined the following tests but was able to pave the way for them.

Signs Test: A normal valuation is 14 positive and 10 negative standardized deviations. Therefore, the result might mean that there is no decisive systematic competition among the tracked accident trends either to overestimate or underestimate the mortality rates being modified, as there is a relatively equal number of positive and negative deviations.

Grouping of Signs Test:

The following examination will be based on a more detailed signal analysis for the sign of randomness, which is not precisely contained in this segment. Yet, the study indicated unanimous mixture of both positive and negative signs which may then be taken as a club gathering.

Cumulative Deviation Test:

Whole Range: The total differentiation gives you the figure of 1.50, this indicating relatively an underestimation of mortality rates for all age groups.

Half Range:

First Half: Precision, in case of age group, substantially deteriorates when the randomisation is carried out on the seamless basis starting at -1.11.

Second Half: 2.62 proved that the bias exists in that younger ones perceive that the recalled event happened faster in 1980 than in 1920.

Quarter Ranges:

First Quarter: -1.49, the undershooting in the former trimester.

Second Quarter: 0.38, meanwhile, is far from balance but less on the underestimation side.

Third Quarter: Again, Entry which was also a mistake.

Fourth Quarter: One outstanding result was the most accurate approval degree 3.14 which exaggerated too much in the last quarter.

iv) Driving tests on graduated rates:

A gradient in mortality rates graph is an example of a disorderly test in which the rates are inspected to confirm that they are not sudden from one age to another. The true second differencing can be obtained by calculating the second differences between the successive rates of reduction by different methods. Then you analyze the acquired data. The large absolute values of the second differences might demonstrate inflection' points or increase fluctuation which can be pointing that the rate is not gradual increase.

Calculation Steps:
First Differences: Figure out the disparity between the mortality rate pairs that are sequentially incremented.
Second Differences: Choose first differences, subtract the first and the second and do the same one by one with the rest of the pairs.
Analysis: Check whether the second differences represent significant gaps for many of them. It shows that the function is jagged.
Slight test conditions are repeatedly applied to make a concrete smooth's surface finish.
Smooth-fit test is typically not needed in standard actuarial practice since the mortality table is created by a process of graduation. Graduation is a term for creating smooth transition between age classes. The shape of the curve is tending to increase based on age with algorithms (models) algebraic expressing this in the extremely and mild way.

Having been asked for it this time around, however, there is a clear rationale behind my eagerly awaited trip.
The request for a smoothness test in this scenario might stem from several reasons:The request for a smoothness test in this scenario might stem from several reasons:

Data Concerns: In situations where there was a confusion concerning data quality and patterns, it is paramount that one ascertains that the trend will not be erratic as a result of the same data quality and patterns.
Methodology Validation: The technique used to be included (for instance, the newly-developed, less popular method) may need to be validated again to convince it generates results consistent with population projections.
Application Sensitivity: Looking at mortality rates and assuming these would be applied to an "extremely sensitive" type of business cases (e.g., setting premiums on life insurance or pensions), even a minor malfunction could have a great impact (financially) for sure and, thus, the control for the whole process needs to be carried out properly.

v) The the analysis also highlighted areas for improvement:

Among the ways to Improve the Outcome:

- Review and Refine Graduation Method: With a respect to the method chosen (parametric, graphical, etc.), you may need to either modify or try the alternative techniques which best reflect real mortality trends while sparing sudden fluctuations.
- Increase Data Quality and Quantity: Vary the precision of mortality calculations by applying more comprehensive data, which include health patterns, socioeconomic indicators, and new records showing medical care trends, that may affect mortality.
- Regular Updating and Validation: Meaningful mortality rates should be frequently monitored, and verification against real outcomes should be mandatory to be sure that they remain relevant. Adjustments should be made in the graduation method accordingly in case a new evidence is discovered.
- The standard mortality rate is an indicator that contains a societal categorization which might disregard and even undermine the fact that there are individual factors such as gender and age that affect longevity.
- Using standard mortality rates to project demand for services (other than pensions) for the elderly could have several consequences:Using standard mortality rates to project demand for services (other than pensions) for the elderly could have several consequences:
- Resource Allocation: The inaccurate mortality rate could be the reason behind misallocation of resources, for example, primary healthcare, secondly housing and social support system lastly. Inexactness of forecasting may lead to using of resources more than needed or not enough to fill the gaps and meet the needs.
- Healthcare Planning: Accurate mortality rates hold the key to sorting out healthcare infrastructure problems, from planning and building of hospitals up to the institution of nursing homes for the elderly. As there could be mistakes while there is graduation, quality and communalism of healthcare services could be disturbed.
- Social Services: Much of these programs offered to elderly in the course of their daily lives to ensure cost effective include transportation, meal services and recreational activities, must be well supported by accurate population projections. A miscalculation on the part of differentiating the goals and timelines between the various teams will hinder the overall efficiency and success of the endeavour.

Question 4

i) ${}_tp_x^{\overline{11}}$ is defined to be the probability that a life in state 1 at age n remains in state 1 until age $n + t$.

To solve this, we use the following assumptions:

- The probabilities that a life at any given age will be in either state at any subsequent age depend only on the ages involved and on the state currently occupied. This is the Markov assumption.
- Consider any two distinct states g and h . Then, for t ≥ 0 and small time interval of length dt , we have:

$$_{dt}p_{x+t}^{gh} = \mu_{(x+t)}^{gh}\, dt + O(dt)$$

Using it here we get:

$$_{t+dt}p_x^{\overline{11}} = {}_tp_x^{\overline{11}} \times {}_{dt}p_{x+t}^{\overline{11}}$$

(Since the probability of staying in state 1 = 1 – probability of leaving state 1, we have)

$$= {}_tp_x^{\overline{11}}(1 - {}_{dt}p_{x+t}^{12} - {}_{dt}p_{x+t}^{14} - {}_{dt}p_{x+t}^{15})$$

(by the second assumption, we have)

$$= {}_tp_x^{\overline{11}}(1 - \mu_{x+t}^{12}dt - \mu_{x+t}^{14}dt - \mu_{x+t}^{15}dt + O(dt))$$

$$\Leftrightarrow \frac{{}_{t+dt}p_x^{\overline{11}} - {}_tp_x^{\overline{11}}}{dt} = -{}_tp_x^{\overline{11}}(\mu_{x+t}^{12} + \mu_{x+t}^{14} + \mu_{x+t}^{15}) + \frac{o(dt)}{dt}$$

Taking the limit $dt \longrightarrow 0$ we get:

$$\frac{\partial\, {}_tp_x^{\overline{11}}}{\partial t} = -{}_tp_x^{\overline{11}}(\mu_{x+t}^{12} + \mu_{x+t}^{14} + \mu_{x+t}^{15})$$

$$\Rightarrow \frac{1}{{}_tp_x^{\overline{11}}} \times \frac{\partial\, {}_tp_x^{\overline{11}}}{\partial t} = -(\mu_{x+t}^{12} + \mu_{x+t}^{14} + \mu_{x+t}^{15})$$

Integrating from 0 to t we obtain:

$$\int_0^t \frac{1}{{}_tp_x^{\overline{11}}} \times \frac{\partial\, {}_tp_x^{\overline{11}}}{\partial t} = \int_0^t -(\mu_{x+t}^{12} + \mu_{x+t}^{14} + \mu_{x+t}^{15})$$

$$\text{Now since } {}_0p_x^{\overline{11}} = 1, we\ have\ \ln\left({}_0p_x^{\overline{11}}\right) = 0$$

$$\ln\left({}_tp_x^{\overline{11}}\right) = -t(\mu_{x+t}^{12} + \mu_{x+t}^{14} + \mu_{x+t}^{15})$$

$$\boldsymbol{\Rightarrow {}_tp_x^{\overline{11}} = e^{-t(\mu_{x+t}^{12}+\mu_{x+t}^{14}+\mu_{x+t}^{15})}}$$

ii) ${}_tp_x^{22}$ is the probability of being in state 2 at time x and at time x+t.

We calculate it using the two assumptions above as follows:

$$_{t+dt}p_x^{22} = \sum_{j=1}^{5} {}_tp_x^{2j} \times {}_{dt}p_{x+t}^{j2}$$

$$= {}_tp_x^{21}\,{}_{dt}p_{x+t}^{12} + \cdots + {}_tp_x^{25}\,{}_{dt}p_{x+t}^{52}$$

$$(but\ {}_{dt}p_{x+t}^{j2} = 0\ for\ j = \{3,4,5\})$$

$$\therefore\quad {}_{t+dt}p_x^{22} = {}_tp_x^{21}\,{}_{dt}p_{x+t}^{12} + {}_tp_x^{22}\,{}_{dt}p_{x+t}^{22}$$

$$= {}_tp_x^{21}\,{}_{dt}p_{x+t}^{12} + {}_tp_x^{22}(1 - {}_{dt}p_{x+t}^{21} + {}_{dt}p_{x+t}^{23} + {}_{dt}p_{x+t}^{23})$$

$$= {}_tp_x^{21}(\mu_{x+t}^{12}dt + o(dt)) + {}_tp_x^{22}(1 - \mu_{x+t}^{21}.dt - \mu_{x+t}^{23}.dt - \mu_{x+t}^{25}.dt + O(dt))$$

$$\Rightarrow \frac{{}_{t+dt}p_x^{22} - {}_tp_x^{22}}{dt} = -{}_tp_x^{22}(\mu_{x+t}^{21} + \mu_{x+t}^{23} + \mu_{x+t}^{25}) + {}_tp_x^{21}(\mu_{x+t}^{12}) + \frac{o(dt)}{dt}$$

Taking the limit $dt \longrightarrow 0$ we get:

$$\boldsymbol{\frac{\partial\, {}_tp_x^{22}}{\partial t} = -{}_tp_x^{22}\left(\mu_{x+t}^{21} + \mu_{x+t}^{23} + \mu_{x+t}^{25}\right)}$$

iii) The benefit scheme described pays out when an employee is long-term sick (State 2), when they retire (State 3), and upon death before retirement. There's no mention of a benefit for withdrawal from the company without retiring when in ill health, though the scheme explicitly mentions about early retirement due to ill health. Therefore, it's reasonable that there's no direct transition from being long-term sick to withdrawal.

iv) The suitability of the Markov assumption would depend on whether the future state of an employee's health, retirement status, or death can be considered independent of their past health history or employment status. If the transitions between are influenced only by the current state and not by how the individual arrived at that state, then the Markov assumption is appropriate. However, if there are lingering effects of an employee's history (for example, a long-term illness that increases the probability of future sickness or early retirement), then the Markov assumption might not be entirely suitable. The model would need to account for such history to accurately predict transitions, which could complicate the model but increase its accuracy.

v) A benefit to this assumption is that if forces of transition were constant for each year, it would simplify the model, making it easier to calculate probabilities and expected values. It would be particularly useful when data is limited or when the model needs to be computationally efficient. However, this may not accurately reflect reality, as transition rates may vary greatly within a year, this can be due to the fact that the company in involved in heavy industry and therefore the mortality rates are higher than average. Being involved in such an industry can also lead to drastic changes in mortality rates within just an year which will give us a table of mortality rates which is not smooth hence all out other results will not be smooth . Over-simplification can lead to incorrect predictions and potentially inadequate scheme funding.

vi) $\boldsymbol{L} =$

$$\boldsymbol{e^{-610(\mu_{50}^{12}+\mu_{50}^{14}+\mu_{50}^{15})}e^{-93(\mu_{50}^{21}+\mu_{50}^{23}+\mu_{50}^{25})}(\mu_{50}^{12})^{39}(\mu_{50}^{13})^{0}(\mu_{50}^{14})^{10}(\mu_{50}^{15})^{6}(\mu_{50}^{21})^{28}(\mu_{50}^{23})^{6}(\mu_{50}^{25})^{3}}$$

vii) By taking the log of the likelihood obtained above we get:

$$\ln(L) = -610(\mu_{50}^{12} + \mu_{50}^{14} + \mu_{50}^{15}) - 93(\mu_{50}^{21} + \mu_{50}^{23} + \mu_{50}^{25}) + 39\ln(\mu_{50}^{12}) + 10\ln(\mu_{50}^{14}) + 6\ln(\mu_{50}^{15}) + 28\ln(\mu_{50}^{21}) + 6\ln(\mu_{50}^{23}) + 3\ln(\mu_{50}^{25})$$

To maximize this log-likelihood we partially differentiate it with respect to μ_{50}^{15} and equate the result to 0

$$\frac{\partial \ln(L)}{\partial \mu_{50}^{15}} = -610 + \frac{6}{\mu_{50}^{15}} = 0$$

$$\Rightarrow \boldsymbol{\widehat{\mu_{50}^{15}} = \frac{6}{610} \cong 0.009836}$$

viii) Variance = -1/ $\frac{d^2 logL}{d{\mu_{50}^{15}}^2}$

$$\frac{d^2 logL}{d{\mu_{50}^{15}}^2} = \frac{-6}{\widehat{\mu_{50}^{15}}}$$

So, we get the variance as:

Variance = -1 / $\frac{-6}{\widehat{\mu_{50}^{15}}}$ = 0.000016125

So, **the standard error = standard deviation = $\sqrt{\boldsymbol{Variance}}$ = 0.0040156**