2025-01-14

A Markov multiple state model for epidemic and insurance modelling

Research group on AIRM Lab, POSTECH

Keywoong Bae

POSTECH
POHANG UNIVERSITY OF SCIENCE AND TECHNOLOGY

Information

- Title: A Markov multiple state model for epidemic and insurance modelling
- Author: Minh-Hoang Tran
- Journal: Astin Bulletin
- Year: 2024

Introduction and Basic theoretical background

1. Introduction

Framework of this study

Modelling the financial impacts of epidemic

Deterministic SIR model

- Ross (1910)
- Kermack and McKendrick (1927,1932)
- Brauer et al. (2019)

Markov multiple state model

- Fend and Garrido (2011)
- Billard and Dayananda (2014)

Actuarial application

- Hua and Cox (2009)
- Feng and Garrido (2011)
- Hillairet and Lopez (2021)
- Hillairet et al. (2022)

Empirical studies

- Estimate the parameters
- Transition probabilities and intensities
- Premium and Reserve

Simulation studies

- Duration of the Epidemic
- Final susceptible size of the Epidemic

2. The Deterministic SIR model

Deterministic SIR model

- The population is divided into three compartments, Susceptible(S), Infected(I), and Removed(R).

Figure 1. *The deterministic SIR model.*

- Let $S(t), I(t),$ and $R(t)$ be the number of susceptible, infected, and removed individuals at time t ($s(t), i(t),$ and $r(t)$ as the corresponding proportions in the population, respectively).

$$\boldsymbol{s'(t) = -\beta s(t) i(t)}$$
$$\boldsymbol{i'(t) = \beta s(t) i(t) - \alpha i(t)}$$
$$\boldsymbol{r'(t) = \alpha i(t)}$$

where α and β are positive constants.

3. The SIR Markov multiple state model

Transition Matrix

- Construct a Markov multiple state model that can capture the dynamics of the deterministic SIR model.
- Three states 0, 1, and 2 which correspond to the three compartments $(S), (I)$, and (R).

- For $0 \leq z \leq t$, $P(z, t)$ is the matrix whose entries are given by

$$\boldsymbol{P^{ij}(z,t) = P(X_t = j \,|X_z = i) = \frac{P(X_t = j, X_z = i)}{P(X_z = i)}}$$

  where X_t is the state that this individual is in at time t, P is the transition matrix for X_t, for all $i, j \in \{0, 1, 2\}$.

- For the observed number of individuals in each compartment of the deterministic SIR model to match the expected number of individuals in the SIR Markov multiple state model, the transition matrix $P(z, t)$ needs to satisfy the following equations:

$$\boldsymbol{s(t) = s(z)P^{00}(z,t)}$$
$$\boldsymbol{i(t) = s(z)P^{01}(z,t) + i(z)P^{11}(z,t)}$$
$$\boldsymbol{r(t) = s(z)P^{02}(z,t) + i(z)P^{12}(z,t) + r(z)}$$

3. The SIR Markov multiple state model

Transition Probabilities

- For the SIR multiple state model, the transition probabilities are given by

$$\boldsymbol{P^{11}(z,t) = e^{-\alpha(t-z)}, P^{12}(z,t) = 1 - e^{-\alpha(t-z)}}.$$

- In addition,

$$\boldsymbol{P^{00}(z,t) = \frac{s(t)}{s(z)}}$$

$$\boldsymbol{P^{01}(z,t) = \frac{i(t) - i(z)e^{-\alpha(t-z)}}{s(z)}}$$

$$\boldsymbol{P^{02}(z,t) = \frac{r(t) - r(z) - i(z)(1 - e^{-\alpha(t-z)})}{s(z)}}$$

- Let $P^{ij}(z,\infty) = \lim_{t\to\infty} P^{ij}(z,t)$. Then,

$$\boldsymbol{P^{00}(z,\infty) = \frac{s(\infty)}{s(z)}, \qquad P^{01}(z,\infty) = 0, \qquad P^{02}(z,\infty) = 1 - \frac{s(\infty)}{s(z)}}.$$

3. The SIR Markov multiple state model

Transition Intensity

- For a Markov multiple state model with transition matrix P and transition intensity matrix μ, the Kolmogorov forward differential equation is given by

$$\boldsymbol{P_t(z,t) := \frac{\partial}{\partial t} P(z,t) = P(z,t)\mu_t}$$

- Suppose for $t \geq 0$,

$$\boldsymbol{P(X_t = 0) = s(t), \qquad P(X_t = 1) = i(t), \qquad P(X_t = 2) = r(t).}$$

- Then the transition intensities are given by

$$\boldsymbol{\mu_t^{01} = \beta i(t), \qquad \mu_t^{12} = \alpha.}$$

- For the SIR multiple state model, the forward Kolmogorov equations are

$$\boldsymbol{\begin{aligned} P_t^{00}(z,t) &= -\beta i(t) P^{00}(z,t), \\ P_t^{01}(z,t) &= \beta i(t) P^{00}(z,t) - \alpha P^{01}(z,t), \\ P_t^{02}(z,t) &= \alpha P^{01}(z,t). \end{aligned}}$$

Empirical studies

Empirical studies

- **Checkpoint**
  - Estimate the parameters $(\hat{\alpha}, \hat{\beta})$ by conditional loglikelihood.
  - Transition probabilities of the epidemic.
  - Calculate the premium and reserve.

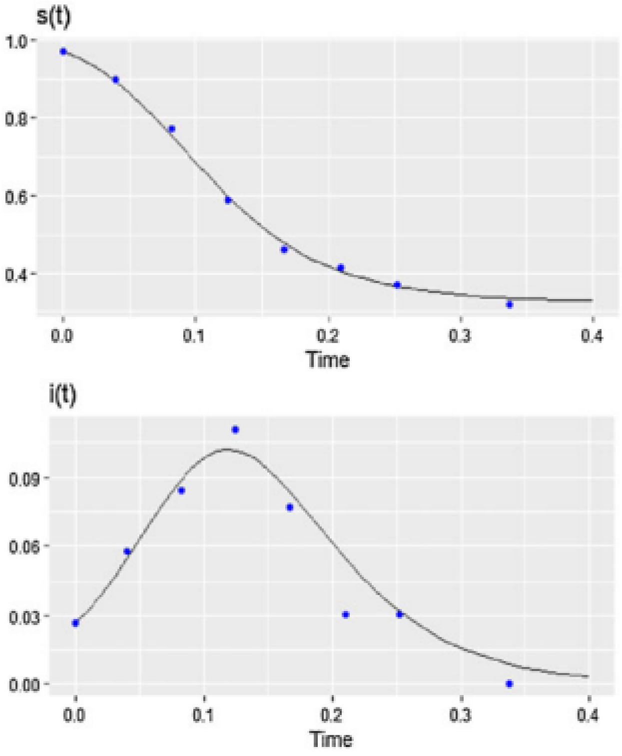
Tran (2024)

Empirical studies

Data set

- For numerical illustration, let us consider the Eyam data set from Raggett (1982).
- Describes the evolution of the bubonic plague in the village of Eyam in **1665-1666**.
- From **a population of 261** in June 1666, there were only **83 survivors in October 1666** when the plague ended.

Date	Time (years)	Susceptible $S(t)$	Infected $I(t)$
June 18	0.0000	254	7
July 3–4	0.0397	235	14
July 19	0.0822	201	22
August 3–4	0.1247	153	29
August 19	0.1671	121	21
September 3–4	0.2096	108	8
September 19	0.2521	97	8
October 20	0.3370	83	0



Parameter Estimation

Conditional log-likelihood and Sum of squares

- We can form the **conditional log-likelihood function** as follows:

$$\boldsymbol{l(\alpha,\beta)=\sum_{i=1}^{M-1}\log\left(P\left(\left((\tilde{S}(t_{i+1})),(\tilde{I}(t_{i+1}))\right)=(S_{t_{i+1}},I_{t_{i+1}})\mid\left((\tilde{S}(t_i)),(\tilde{I}(t_i))\right)=(S_{t_i},I_{t_i})\right)\right)},$$

  given the values (S_{t_i}, I_{t_i}) of the number of susceptible and infected individuals $(\tilde{S}(t_i), \tilde{I}(t_i))$ at time t for $i = 1, \dots, M$ where $t_i = 0$.

- Given the initial values $s(0) = \frac{254}{261}, i(0) = \frac{7}{261}$, we can solve the differential Equations for $(s(t_i), i(t_i), r(t_i))$.

Parameter Estimation

Conditional log-likelihood and Sum of squares

- $l(\alpha, \beta)$ can be maximized by minimizing the sum of squares,

$$\sum_{k=1}^{M} (\hat{s}(t_k) - s(t_k))^2 + \sum_{k=1}^{M} (\hat{i}(i_k) - i(t_k))^2 \,,$$

  where $s(t_k), i(t_k)$ are the values obtained from differential Equations and $\hat{s}(t_k), \hat{i}(t_k)$ are the values calculated from data.

- The initial estimates are $\hat{\alpha}_0 = 34.739, \hat{\beta}_0 = 56.441$, and the **maximum likelihood estimates** of α and β are given by $\boldsymbol{\hat{\alpha} = 34.150,\ \hat{\beta} = 55.437}$.

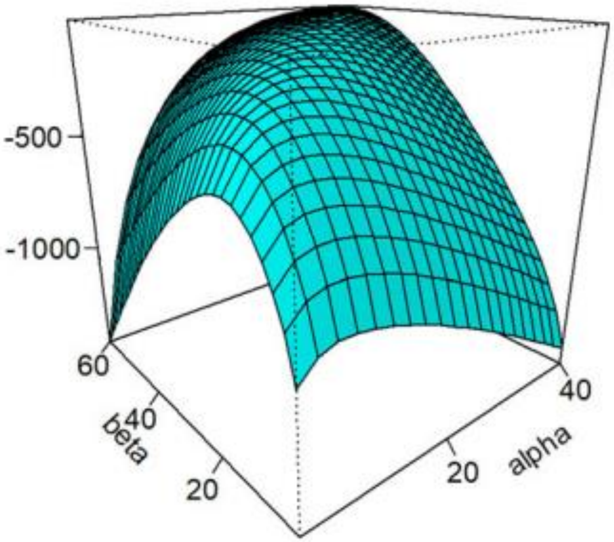


Figure 2. Log-likelihood function.

Transition probabilities

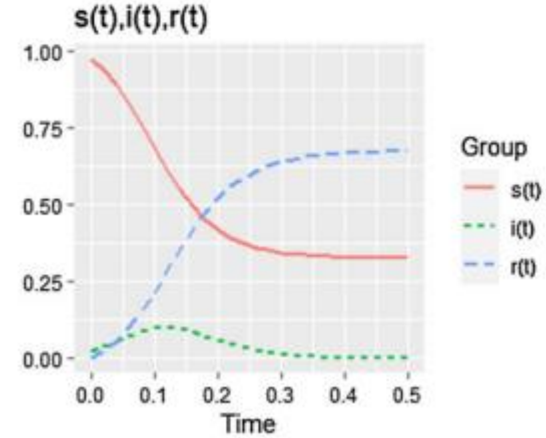
Calculation of transition probabilities, $P^{0i}(z, z+t)$

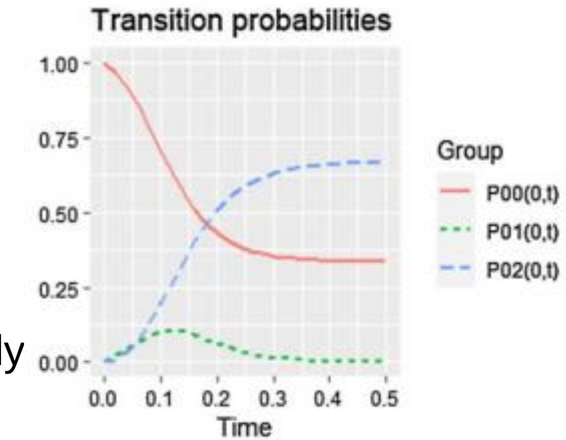
- Substituting the values of $\hat{\alpha}$ **and** $\hat{\beta}$, we can calculate $P^{00}(z,t), P^{01}(z,t), P^{02}(z,t)$ for all t.

$$\boldsymbol{P^{00}(z,t) = \frac{s(t)}{s(z)}}$$

$$\boldsymbol{p^{01}(z,t) = \frac{i(t) - i(z)e^{-\alpha(t-z)}}{s(z)}},$$

$$\boldsymbol{P^{02}(z,t) = \frac{r(t) - r(z) - i(z)(1 - e^{-\alpha(t-z)})}{s(z)}}$$





- Using this formula, we can calculate the transition probabilities $P(z, z+t)$ for any z, t.

- Through the calculation, we can observe that calculated $P^{00}(0,t), P^{01}(0,t), P^{02}(0,t)$ have remarkably **similar shapes** to $s(t), i(t), r(t)$.

Transition probabilities

Calculation of transition probabilities, $P^{0i}(z, z+t)$

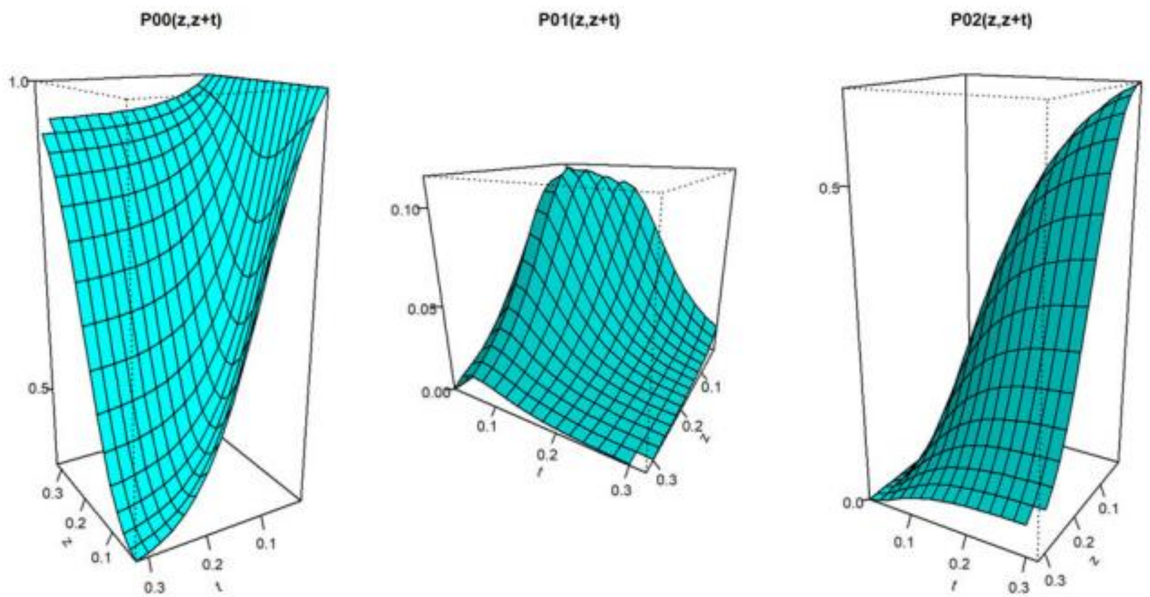


Figure 5. *Transition probabilities* $P^{0i}(z, z+t)$ *for* $i = 0, 1, 2$.

Transition probabilities

Calculation of transition probabilities, $\boldsymbol{P^{0i}(z, z+t)}$ with $\boldsymbol{t \to \infty}$

- To find the value of $s(\infty)$ and $r(\infty)$, we have to solve

$$\boldsymbol{s(\infty) = 1 + \frac{\alpha}{\beta}\log\left(\frac{s(\infty)}{s(0)}\right)}\ ^{*}$$

which yields $\boldsymbol{s(\infty) = 0.3257,\ i(\infty) = 0,\ r(\infty) = 1 - s(\infty) = 0.674.}$

- Also, we can deduce

$$\begin{aligned}
\boldsymbol{P^{00}(0,\infty)} &= \boldsymbol{\frac{s(\infty)}{s(0)} = 0.3346,} \\
\boldsymbol{P^{01}(0,\infty)} &= \boldsymbol{\frac{i(\infty) - i(0)e^{-\alpha\cdot\infty}}{s(0)} = 0,} \\
\boldsymbol{P^{02}(0,\infty)} &= \boldsymbol{\frac{r(\infty) - r(0) - i(0)(1 - e^{-\alpha\cdot\infty})}{s(0)} = 0.6654}
\end{aligned}$$

- Hence, at the start of the epidemic, a susceptible individual has a **33.46%** chance of not getting infected and **66.54%** chance of getting removed eventually.

* Proposition 2.1

Premium and reserve estimation

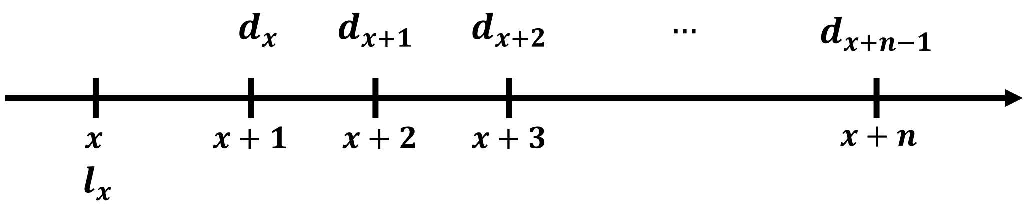
EPV Calculation

- We derive formulas for the **expected present values (EPV)** of insurance and annuity benefits in the SIR multiple state model.
- Consider a Markov multiple state model with $M+1$ states $0, 1, \ldots, M$.
- The transition matrix P and the force of transition matrix is μ_t.
- Let δ be the constant force of interest and $v = e^{-\delta}$.

Premium and reserve estimation

Net Single Premium (NSP) on Life insurance

$$NSP = \frac{v^n l_{x+n}}{l_x} = \frac{v^{x+n} l_{x+n}}{v^x l_x}$$



$$NSP = \frac{v^1 d_x + v^2 d_{x+1} + \cdots + v^n d_{x+n-1}}{l_x} = \frac{v^{x+1} d_x + v^{x+2} d_{x+2} + \cdots + v^{x+n} d_{x+n-1}}{v^x l_x}$$

Premium and reserve estimation

EPV Calculation (Individual)

- **For any i, j,** let $\bar{a}^{ij}(z,n)$ be the EPV of an insurance benefit that pays a **continuous annuity** at the rate of \$1 per annum between time z and n to an individual, who is at state i at time z, as long as he is in state j.

$$\bar{a}^{ij}(z,n) = \int_z^n v^t P^{ij}(z,t)\,dt$$

- **For $i \neq j$,** let $\bar{A}^{ij}(z,n)$ be the EPV of an insurance benefit that pays \$1 at time t for $z \leq t \leq n$ if the policyholder who is at state i at time z **moves to state j at time t.**

$$\bar{A}^{ij}(z,n) = \sum_{k:k\neq j}^{M} \int_z^n v^t P^{ik}(z,t)\mu_t^{kj}\,dt$$

- Let $\bar{B}^{ij}(z,n)$ be the EPV of an insurance benefit that pays \$1 at time t for $z \leq t \leq n$ if the policyholder who was at state i at time z **transfer out of state j at time t.**

$$\bar{B}^{ij}(z,n) = \sum_{k:k\neq j}^{M} \int_z^n v^t P^{ij}(z,t)\mu_t^{jk}\,dt$$

Premium and reserve estimation

EPV Calculation (Aggregate)

- Define the aggregate actuarial functions and calculate premiums and reserves considering the all population including those who are already infected or removed.

- If each individual is entitled to a **continuous annuity** at the rate of \$1 per annum between time z and n, then the EPV of the cost to the insurer per individual is given by

$$\bar{a}^s(z,n) = \int_z^n v^t s(t)dt\,, \qquad \bar{a}^i(z,n) = \int_z^n v^t i(t)dt, \qquad \bar{a}^r(z,n) = \int_z^n v^t r(t)dt\,.$$

- If each individual is **compensated \$1 immediately** upon getting infected (or removed) between time z and n, then the EPV of the cost to the insurer per individual is given by

$$\bar{A}^i(z,n) \coloneqq \int_z^n v^t s(t)\mu_t^{01} dt = \beta \int_z^n v^t s(t) i(t) dt\,,$$

$$\bar{A}^r(z,n) \coloneqq \int_z^n v^t i(t)\mu_t^{12} dt = \alpha \int_z^n v^t i(t) dt\,.$$

Premium and reserve estimation

Premiums

- Consider the following insurance policy which is in force for n years:

  ① If the policyholder is susceptible, then he pays a **continuous premium at the rate of P** per annum;

  ② If the policyholder is infected, then he receives **immediately a payment of S^1** and a **continuous hospitalization benefit at the rate of H** per annum;

  ③ If the policyholder is removed, then he receives **immediately a payment of S^2**.

- Under the equivalence principle, the premium rate P at the individual level is

$$\boldsymbol{P} = \frac{\boldsymbol{S^1\,\overline{A}^{01}(0,n) + H\,\overline{a}^{01}(0,n) + S^2\,\overline{A}^{02}(0,n)}}{\boldsymbol{\overline{a}^{00}(0,n)}}.$$

- The premium rate per individual $\boldsymbol{\pi}$ at the aggregate level is given by

$$\boldsymbol{\pi} = \frac{\boldsymbol{S^1\,\overline{A}^{i}(0,n) + H\,\overline{a}^{i}(0,n) + S^2\,\overline{A}^{r}(0,n)}}{\boldsymbol{\overline{a}^{s}(0,n)}}.$$

Premium and reserve estimation

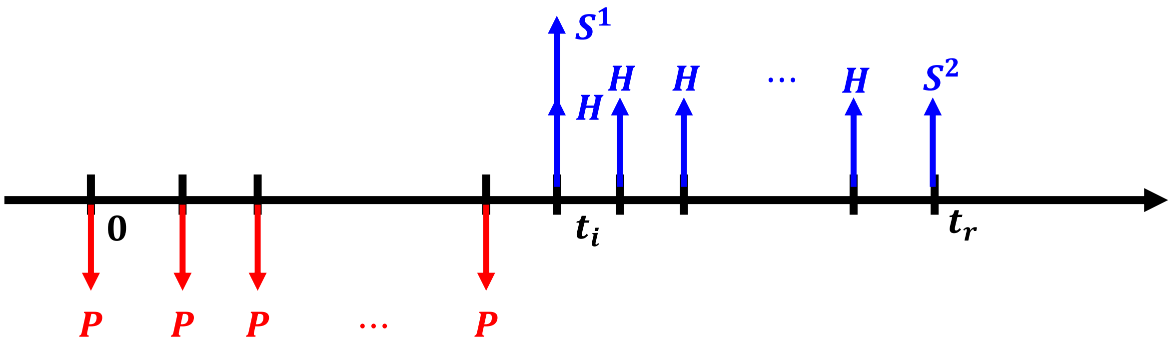
Premiums

- Consider the following insurance policy which is in force for n years:

  ① If the policyholder is susceptible, then he pays a **continuous premium at the rate of P** per annum;

  ② If the policyholder is infected, then he receives **immediately a payment of S^1** and a **continuous hospitalization benefit at the rate of H** per annum;

  ③ If the policyholder is removed, then he receives **immediately a payment of S^2**.



Premium and reserve estimation

Reserves

- For reserves, we differentiate 4 different types
  - prospective reserves at the individual or aggregate level,
  - retrospective reserves at the individual or aggregate level.
- **Prospective** means we are taking the EPV of the difference **between future benefits and future premium payments** (between time t and n).
- **Retrospective** means we are taking the EPV of the difference between the accumulating value of **benefits and premium payments, which are already paid and received,** respectively (between time 0 and t).

Premium and reserve estimation

Reserves (Individual)

- The **prospective** reserves **at the individual level** at time t for a susceptible policyholder and an infected policyholder are respectively,

$$
{}_tV^{(0)} = \boldsymbol{v}^{\boldsymbol{t}} \left(\boldsymbol{S}^{\mathbf{1}} \, \overline{\boldsymbol{A}}^{\mathbf{01}}(\boldsymbol{t}, \boldsymbol{n}) + \boldsymbol{H} \, \overline{\boldsymbol{a}}^{\mathbf{01}}(\boldsymbol{t}, \boldsymbol{n}) + \boldsymbol{S}^{\mathbf{2}} \, \overline{\boldsymbol{A}}^{\mathbf{02}}(\boldsymbol{t}, \boldsymbol{n}) - \boldsymbol{P} \, \overline{\boldsymbol{a}}^{\mathbf{00}}(\boldsymbol{t}, \boldsymbol{n}) \right),
$$

$$
{}_tV^{(1)} = \boldsymbol{v}^{\boldsymbol{t}} \left(\boldsymbol{H} \, \overline{\boldsymbol{a}}^{\mathbf{11}}(\boldsymbol{t}, \boldsymbol{n}) + \boldsymbol{S}^{\mathbf{2}} \, \overline{\boldsymbol{A}}^{\mathbf{12}}(\boldsymbol{t}, \boldsymbol{n}) \right).
$$

- The **retrospective** reserves **at the individual level** at time t for a susceptible policyholder and an infected policyholder are respectively,

$$
{}_tV_R^{(0)} = \left(\boldsymbol{P} \, \overline{\boldsymbol{a}}^{\mathbf{00}}(\mathbf{0}, \boldsymbol{t}) - \boldsymbol{S}^{\mathbf{1}} \, \overline{\boldsymbol{A}}^{\mathbf{01}}(\mathbf{0}, \boldsymbol{t}) - \boldsymbol{H} \, \overline{\boldsymbol{a}}^{\mathbf{01}}(\mathbf{0}, \boldsymbol{t}) - \boldsymbol{S}^{\mathbf{2}} \, \overline{\boldsymbol{A}}^{\mathbf{02}}(\mathbf{0}, \boldsymbol{t}) \right),
$$

$$
{}_tV_R^{(1)} = \left(-\boldsymbol{H} \, \overline{\boldsymbol{a}}^{\mathbf{11}}(\mathbf{0}, \boldsymbol{t}) - \boldsymbol{S}^{\mathbf{2}} \, \overline{\boldsymbol{A}}^{\mathbf{12}}(\mathbf{0}, \boldsymbol{t}) \right).
$$

Premium and reserve estimation

Reserves (Aggregate)

- Let $W^p(t)$ and $W^R(t)$ be the prospective and retrospective reserves **per individual at the aggregate level**. Then,

$$\boldsymbol{W^P(t) = v^t\left(S^1\,\overline{A}^i(t,n) + H\,\overline{a}^i(t,n) + S^2\,\overline{A}^r(t,n) - \pi\,\overline{a}^s(t,n)\right)},$$
$$\boldsymbol{W^R(t) = \left(\pi\,\overline{a}^s(0,t) - S^1\,\overline{A}^i(0,t) - H\,\overline{a}^i(0,t) - S^2\,\overline{A}^r(0,t)\right)}.$$

Premium and reserve estimation

Premium and Reserve calculation

- Consider an insurance policy that pays a **continuous annual hospitalization benefit of $1,000** per annum for a maximum of one year, with the constant force of interest, 5%.

- The annual individual/aggregate premium rate is given by

$$\boldsymbol{P = \frac{1000\,\bar{a}^{01}(0,1)}{\bar{a}^{00}(0,1)} = 47.5408}$$

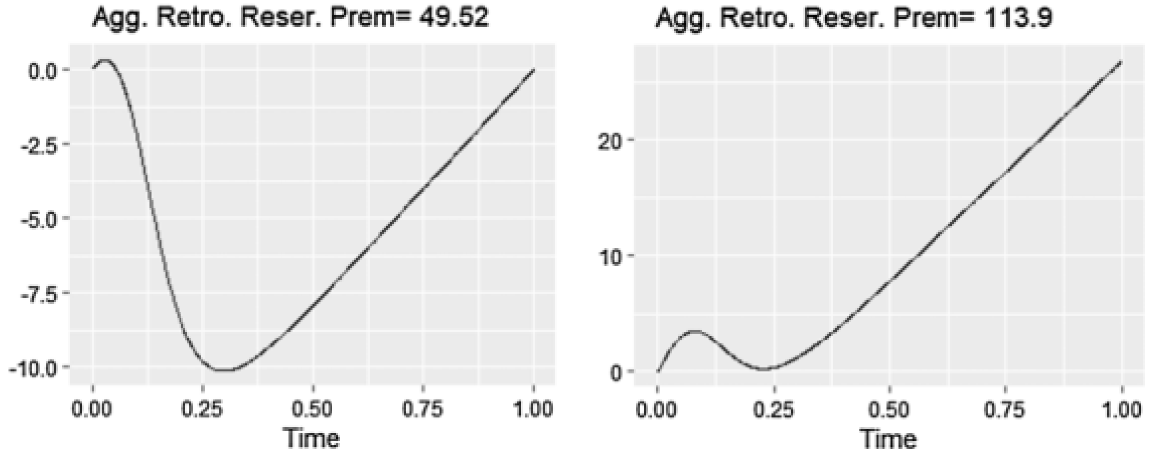
$$\boldsymbol{\pi = \frac{1000\,\bar{a}^{i}(0,1)}{\bar{a}^{s}(0,1)} = 49.5219}$$

where $\boldsymbol{S^1 = S^2 = 0, H = 1000}, \bar{a}^{00}(0,1) = 0.01934, \bar{a}^{01}(0,1) = 0.4068$

Premium and reserve estimation

Premium and Reserve calculation

- In the end of the epidemic, the **aggregate retrospective reserves are usually negative** since the infection rate decreases.
- This leads to an early surrender of policyholders at the peak and a deficit of reserves.
- So, insurance providers should charge premiums at the higher rate than that determined by the equivalence principle.
- We can obtain a higher premium rate of $113.90 based on the algorithm in Feng and Garrido (2011).



Feng, Runhuan, and Jose Garrido. "Actuarial applications of epidemiological models." North American Actuarial Journal 15.1 (2011): 112-136.

Simulation studies

Simulation studies

Simulations

- Describe the simulation procedure for the process $(X_t)_{t\geq 0}$.
- Observe the **duration and final susceptible size** of the epidemic.
- Assume the populations of 20,000 with 261 individuals, consisting of 254 susceptible and 7 infected at in the Eyam data set.

Simulation studies

The distribution of waiting times

- Let $T^{(0)}$ and $T^{(1)}$ be the waiting times in states 0 and 1, respectively.
- Let $T = T^{(0)} + T^{(1)}$ be the time until removal or time to absorption.

$$\boldsymbol{P\left(T^{(0)} > t\right) = \frac{s(t)}{s(0)}}, \qquad \boldsymbol{P\left(T^{(0)} = \infty\right) = \frac{s(\infty)}{s(0)}},$$

$$\boldsymbol{P\left(T^{(1)} = 0\right) = \frac{s(\infty)}{s(0)}}, \qquad \boldsymbol{P\left(T^{(1)} > t\right) = e^{-\alpha t}\left(1 - \frac{s(\infty)}{s(0)}\right)}.$$

- The distribution of waiting times $\boldsymbol{T}(= \boldsymbol{T^{(0)}} + \boldsymbol{T^{(1)}})$ is given by

$$\boldsymbol{P(T < t) = \frac{r(t) - i(0)(1 - e^{-\alpha t})}{s(0)}}, \qquad \boldsymbol{P(T = \infty) = \frac{s(\infty)}{s(0)}}.$$

Simulation studies

Procedure of simulation

20,000 populations

254 susceptible

① Generate $u \sim \mathrm{Uniform}(0,1)$ and $\mathrm{w} \sim \mathrm{Exponential}(\alpha)$

② Simulate the duration of the epidemic, T $(= T^{(0)} + T^{(1)})$.

$$\boldsymbol{T^{(0)}} = \begin{cases} \boldsymbol{s^{-1}(s(0)u)}, & \boldsymbol{u} > \dfrac{\boldsymbol{s(\infty)}}{\boldsymbol{s(0)}} \\ \infty, & \boldsymbol{u} \leq \dfrac{\boldsymbol{s(\infty)}}{\boldsymbol{s(0)}} \end{cases} \quad \& \quad \boldsymbol{T^{(1)}} = \begin{cases} \boldsymbol{0}, & \boldsymbol{T^{0}} = \infty \\ \boldsymbol{w}, & \boldsymbol{T^{0}} < \infty \end{cases}$$

③ Calculate the time $T\left(= T^{(0)} + T^{(1)}\right)$ of that individual.

④ Calculate the duration of the epidemic and compare the estimates from simulations and approximation.

$$\boldsymbol{D} = \mathbf{inf}\{\boldsymbol{t} \geq \boldsymbol{0}, \tilde{\boldsymbol{I}}(\boldsymbol{t}) = \boldsymbol{0} \text{ and there are no further infections after } \boldsymbol{t}\}$$

$$\boldsymbol{D} = \max_{\boldsymbol{1 \leq i \leq N}} \{\boldsymbol{T_i} : \boldsymbol{T_i} < \infty\}$$

⑤ Calculate the number of $s(\infty) = S(0) \cdot P(T^{(0)} = \infty)$.

Simulation studies

Results of the simulation

- The average duration of the epidemic from the simulations is **0.4749 years** compared with the expected value of 0.4751 years and standard deviation 0.0798.

- Suppose $z > t^*$ and $i(z)$ is sufficiently small. Let D_z be the duration of the epidemic beyond time z. That is $D_z = D - z$ if $D > z$ and 0 otherwise. Then the expected value of D_z given the state of the population at time z can be approximated by

$$\boldsymbol{E[D_z \mid \left(\tilde{S}(z), \tilde{I}(z)\right) = (S_z, I_z)] = \frac{S_z \beta i(z)}{\alpha - \beta s(\infty)} \left(\frac{1}{\alpha - \beta s(\infty)} + \frac{1}{\alpha} \right) + \frac{I_z}{\alpha} + o\big(i(z)\big)}.$$

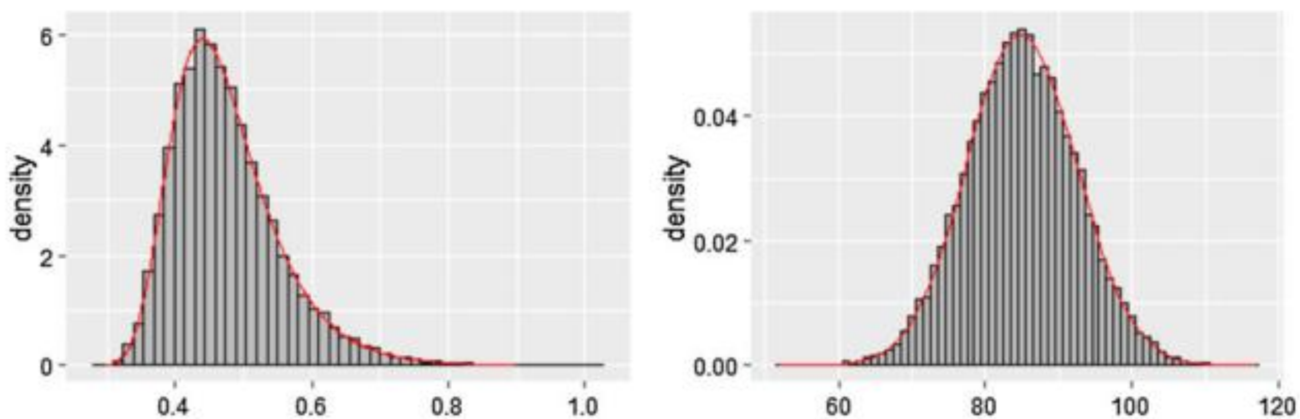


***Figure 8.** Histograms of the duration and the final susceptible size.*

* Proposition 5.14

Thank you

AIRM Research Group

Actuarial modeling, Insurance and Risk Management Lab.
Dept. of Industrial and Management Engineering, POSTECH