Interplay between Epidemiology and Actuarial Modeling

Runhuan Feng, Ph.D., FSA, CERA; Longhao Jin, MS;
and Sooie-Hoe Loke, Ph.D.

Abstract

In an era where rapidly evolving situations is the new normal, collaboration between multiple disciplines offers a concerted effort and provides a comprehensive perspective in tackling problems and challenges. In this essay, we illustrate various modeling tools and key ideas used in epidemiology that can be applied to the insurance framework. Specifically, we give an overview of the compartmental models, network models, and agent-based models, and discuss their applications to epidemic and cyber insurance coverages.

Keywords. Compartmental model, network models, agent-based models, epidemic, reserve level, cyber insurance

---

1. INTRODUCTION

The actuarial profession is constantly evolving. The age of big data has prompted actuaries to learn and apply the state-of-the-art predictive analytics methods and machine learning techniques from non-traditional fields. The current COVID-19 pandemic has changed the world in many ways, and in the insurance industry, has prompted us to think about the challenges posed in the profession and ways to address these problems.

Written from an academic perspective, the goal of this essay is to discuss several modeling tools from epidemiology and how they can be applied to an insurance setting. Rather than reinventing the wheel, actuarial researchers and practitioners could take advantage of a wide variety of epidemic models developed in the medical literature to estimate the evolution of mortality and morbidity rates during a pandemic. We believe that many of these models can be easily adapted for actuarial applications. As both actuaries and epidemiologists share a common interest in using advanced mathematical tools to model risk, we hope that this essay motivates actuarial researchers to utilize these models as building blocks to study insurance coverages targeting infectious diseases.

In the following sections, we present three distinct levels of modeling in epidemiology and describe their applications to insurance modeling. We begin with widely used compartmental models, including both deterministic and stochastic versions. This first level of modeling assumes homogeneous mixing, that is, everyone in the population is equally likely to be susceptible and infectious regardless of age, sex, social structure, etc. The next level of modeling is called contact network models, where the individuals are connected by links to describe interactions between them. Finally, we discuss agent-based models which are microscopic models used to simulate real-world complex patterns. Within each model, we provide a list of references in epidemiology as well as in actuarial science.

2. COMPARTMENTAL MODELS

Consider a population of size $N(t)$ which is indexed by time t. The basic framework of compartmental models is to classify the entire population into several distinct compartments or categories. We illustrate this idea via the well-known SIR model based on the seminal work of Kermack and McKendrick (1927). Consider the following three compartments: **S**usceptible, **I**nfected, and **R**emoved. There are many variations of the SIR model, so in what follows we lay out the assumptions for the most basic SIR model. First, there are no births or immigration, which implies that $N(t) = N$, and the population mixes homogeneously. Second, there is an infection rate β and a recovery rate α. In other words, $1/\beta$ is the average time between contacts and $1/\alpha$ is the average time until removal. Finally, once recovered (or removed), an individual is immune and can no longer spread the disease. The following figure summarizes some of the assumptions in the model:

This gives rise to the following system of differential equations:

$$S'(t) = -\beta\, S(t)\, I(t) \tag{2.1}$$

$$I'(t) = \beta\, S(t)\, I(t) - \alpha\, I(t) \tag{2.2}$$

where $S(t)$ and $I(t)$ are the number of susceptible and infected individuals, respectively. The number of removed individuals is thus $N - S(t) - I(t)$.

A main concern in epidemiology modeling is whether a disease will spread upon its introduction. Observe from (2.2) that the disease will spread (i.e., $I'(t)$ is positive) if $\beta S(t) - \alpha > 0$ and the disease will die out if $\beta S(t) - \alpha < 0$. This motivates the relevance of a key quantity called the effective (or general) reproductive number, given by $R_t = \beta S(t)/\alpha$, which represents the average number of secondary infections due to a single infectious individual at a given time t. It is also worth noting that β/α is the average number of contacts by an infected person with others before removal. When $t = 0$, R_0 (read as R-naught) is known as the basic reproduction number, which is a common measure used to determine if a disease will spread out during the early phase of the outbreak. If $R_0 > 1$, the disease will start to spread, but not if $R_0 < 1$. Zhao et al. (2020) gave a preliminary estimate of R_0 of the coronavirus to be between 2.24 and 3.58.

More details regarding the SIR model as well as other compartmental models like SIS (Susceptible-Infected-Susceptible) and SEIR (Susceptible-Exposed-Infected-Removed) can be found in classic monographs such as Anderson and May (1992), Brauer and Castillo-Chavez (2012), and Diekmann et al. (2012). For a more recent account of developments in mathematical epidemiology, see Brauer (2017). The aforementioned monographs also discussed heterogeneous mixing models (incorporating age, social structure, etc.) as well as cross-population models.

2.1 Insurance Application

Using the basic SIR model, Feng and Garrido (2011) proposed several insurance policies and models to quantify infection risk using actuarial methodologies. For example, susceptible individuals pay premiums at a constant rate π and once infected, the insurer will pay hospitalization benefits, say at a constant rate of 1. For this particular policy, the insurer's reserve level is

$$V(\pi,t) = \pi \int_0^t s(x)dx - \int_0^t i(x)dx, \tag{2.3}$$

where $s(t) = S(t)/N$ and $i(t) = I(t)/N$. In analyzing the reserve, there are four possible shapes of graph (as a function of time) and it was recently discovered that these shapes are closely linked to the effective reproduction number, R_t, as summarized in the following table:

Shape of $V(\pi,t)$	Interval for values of π
Increasing concave	At least $(1/R_T) - 1$
Increasing concave-then-convex	Between $(1/R_{tm}) - 1$ and $(1/R_T) - 1$
Nonmonotonic concave-then-convex	Between $(1/R_0) - 1$ and $(1/R_{tm}) - 1$
Nonmonotonic convex	Between 0 and $(1/R_0) - 1$

Here, T is the time of the disease-free state (i.e., $I(T) = 0$) and the exact expression of R_{tm} is provided in Feng and Garrido (2011). It is interesting to note that a related quantity from the table, namely $1 - 1/R_0$, is called the herd immunity threshold (see Fine et al. (2011)).

This deterministic insurance model was later extended to a stochastic one by Lefèvre et al. (2017), which provided many elegant results on integral functionals using martingale arguments. The work of Lefèvre et al. (2017) in stochastic SIR insurance model has since opened the door to many classical problems in probability. Lefèvre and Picard (2018a) generalized the SIR model to a controlled epidemic model, where the infectious will be isolated by health organizations to ease the severity of the disease and studied the representation of epidemic outcomes and path integrals in terms of pseudo-polynomials. Lefèvre and Simon (2018) considered cross-infection between two linked populations. A general approach to study Laplace transforms of these integral functionals was developed by Lefèvre and Picard (2018b). More recently, Lefèvre and Simon (2019) proposed a general block-structured Markov processes for epidemic modeling. For a thorough discussion of stochastic epidemic models and methods for their statistical analysis, see Andersson and Britton (2012).

In terms of other deterministic insurance models, Perera (2017) considered the control strategy in the simple SIR model as well as the variation of the premium with respect to the model parameters. Nkeki and Ekhaguere (2020) constructed the SIDRS model and studied its insurance applications. Billard and Dayananda (2014b) and Billard and Dayananda (2014a) developed a multi-stage HIV/AIDS model considering non-disease death in each compartment, where the waiting time distribution is used to measure the total amount of time one individual holds in one state. Not only that premiums are defined by different insurance functions, but health-care cost adjustments are also included. Shemendyuk et al. (2019) investigated the deterministic and stochastic SIR models with multiple centers and migration fluxes. The optimal health-care premium is determined by considering different vaccine allocation strategies. Optimal resource allocation and contingency planning were also addressed in Chen et al. (2020).

3. NETWORK MODELS

To capture the interactions among individuals while still preserving some aspects of the compartmental models, one can use a network model. Using the language of graph theory, the individuals are known as vertices or nodes and are connected by edges or links, which indicate relationships between various vertices. Some common metrics used to describe a network include shortest path length, degree distribution, and clustering coefficient. Epidemiologists have a long history of using network models to study diseases. An extensive survey by Pastor-Satorras et al. (2015) describes various types of network models used in epidemiology and highlights recent major results in the field. While network-based models have desirable and complex properties that cannot be replicated by equation-based models, network models are computationally challenging and often result in high-dimensional data analysis (see Pellis et al. (2015)). For general study of the structure and dynamics of network models, as well as their applications in various disciplines, Boccaletti et al. (2006) and Newman (2003) are excellent reads.

3.1 Risk Network and Cyber Insurance

Böhme et al. (2017) provided a framework of cyber risk modeling and assessment. In contrast with the frequency and severity analysis of conventional risks, cyber risk is analyzed in the cybersecurity literature by a chain of causality from cyber risk factors such as threats, vulnerabilities, controls, and assets to financial losses. Each cyber incident is viewed as a threat acting on an objective's vulnerabilities. Companies can place technical controls to remove or reduce vulnerabilities. An incident is turned into a loss when attacks hit assets. The contagion of cyber risk is characterized by a network model, under which each firm is represented by a node and all nodes are interconnected by physical and social links. An example of cyber risk dependency is the propagation of malware through a network of interconnected computers. The specific dependence structure can be modeled by graph topology. Such topological models originated in epidemiological literature to model the transmission of infectious diseases and are extended in the context of cybersecurity to the spread of computer virus.

The causality approach based on network models, which is well studied in the cybersecurity literature, is fundamentally different from the trend analysis based on historical claims data in the actuarial literature. It has been argued in Böhme et al. (2017) that historical data is of limited use due to the fast-changing nature of cyber technology development. Therefore, further research is needed to combine actuarial analysis with cybersecurity framework in cyber risk quantification and modeling.

4. AGENT-BASED MODELS

The decisions of one person in a group can vastly affect the final outcome of the whole group. Therefore, when it comes to a community-scale activity, it is important to consider the competing strategies among all individuals. For those scenarios, an agent-based model (ABM) performs exceptionally well. Formally, it is a data-driven, simulation-based model that captures individual (or agent) level actions and view the community as a complex adaptive system. Zhou (2013) pointed out that ABMs bring insight of how to model agent's behavior, how to understand the learning process of an agent, and how to quantify the complex interactions among all the agents. Palin et al. (2008) elaborated on the main features of ABMs, namely (i) there are finitely many types of agents following different rules, (ii) agents are allowed to learn, combine, and evolve during the process, (iii) there will be positive or negative feedback from the agents' action, (iv) agents make contact with the environment and the environment will also be affected by the actions of the agents. As noted in Niu (2016), there are three types of ABMs, namely homogeneous ABM (where the agents are uniform and follow the same decision rules but have different parameters), heterogeneous ABM (where there are different types of agents communicating within the same environment), and agent-based queuing model (ABM that applies queuing theory).

4.1 Insurance Application

In order to make a comparison among insurance companies with similar business volume, Taylor (2008) built a dynamic model based on the model framework that each insurer is an agent. Moreover, applying advanced machine learning technique, Parodi (2012a, b) categorized the actuarial problems into different machine learning frameworks and applied computational intelligence methods to insurance business.

In light of the current pandemic, we propose a possible case study using ABMs. Suppose the insurer is interested in the number of confirmed and death cases in the United States for the coming fall. To model complex human behaviors and social phenomenon such as stay-at-home orders, social distancing, and mask wearing culture, ABMs are robust enough to incorporate these features. Historical data can be used to estimate the parameters in the ABM and subsequently one can simulate multiple scenarios based on these fitted parameters. Circling back to the insurance example in Section 2.1, these results can then be used to analyze how much reserve the insurance company will need for future benefit payments and what is the fair premium amount for the policyholder to pay corresponding to each level of the reproduction number.

5. CONCLUSIONS

This essay highlights the interplay between actuarial science and epidemiology and provides literature review of recent results. The marriage of modeling techniques from the two areas may revolutionize the way we quantify and model epidemic insurance and the way we organize cyber insurance. We hope to point to new sources of inspiration for actuaries working in areas of risk modeling and assessment of dynamically evolving populations. We envision actuaries adding these epidemiological models and techniques into their growing toolkits for risk modeling and analysis.

Acknowledgment

This research is supported by a grant by the Canadian Institute of Actuaries in response to COVID-19 research and by an endowment from the State Farm Companies Foundation.

6. REFERENCES

[1] Anderson, R.M. and May, R.M., *Infectious diseases of humans: dynamics and control*, Oxford university press, 1992.
[2] Andersson, H. and Britton, T., *Stochastic epidemic models and their statistical analysis*, Springer Science & Business Media, 2012, Vol. 151.
[3] Billard, L. and Dayananda, P.W.A, "A multi-stage compartmental model for HIV-infected individuals: I–Waiting time approach," *Mathematical Biosciences*, 2014, Vol. 249, 92-101.
[4] Billard, L. and Dayananda, P.W.A, "A multi-stage compartmental model for HIV-infected individuals: II–Application to insurance functions and health-care costs," *Mathematical Biosciences*, 2014, Vol. 249, 102-109.
[5] Boccaletti, S., Latora, V., Moreno, Y., Chavez, M., and Hwang, D.U., "Complex networks: Structure and dynamics," *Physics reports*, 2006, Vol. 424, No. 4-5, 175-308.
[6] Böhme, R., Laube, S., Riek, M., "A fundamental approach to cyber risk analysis," *Variance*, 2017, Vol. 12, No. 2, 161-185.
[7] Brauer, F., and Castillo-Chavez, C., "*Mathematical models in population biology and epidemiology*," Springer, 2012, Vol. 2.
[8] Brauer, F., "Mathematical epidemiology: Past, present, and future," *Infectious Disease Modelling*, 2017, Vol. 2, No. 2, 113-127.
[9] Chen, X., Chong, W.F., Feng, R. and Zhang, L., "Pandemic risk management: resources contingency planning and allocation," 2020.
[10] Diekmann, O., Heesterbeek, H., and Britton, T., "*Mathematical tools for understanding infectious disease dynamics,*" Princeton University Press, 2012, Vol. 7.
[11] Feng, R., and Garrido, J., "Actuarial applications of epidemiological models," *North American Actuarial Journal*, 2011,

Vol. 15, No. 1, 112-136.

[12] Fine, P., Eames, K., and Heymann, D.L., "“Herd immunity”: a rough guide," *Clinical infectious diseases*, 2011, Vol. 52, No. 7, 911-916.

[13] Kermack, W.O., and McKendrick, A.G., "A contribution to the mathematical theory of epidemics," *Proceedings of the Royal Society of London. Series A, Containing papers of a mathematical and physical character*, 1927, Vol. 115, No. 772, 700-721.

[14] Lefèvre, C. and Picard, P., "Final outcomes and disease insurance for a controlled epidemic model," *Applied Stochastic Models in Business and Industry*, 2018, Vol. 34, No. 6, 803-815.

[15] Lefèvre, C. and Picard, P., "A general approach to the integral functionals of epidemic processes," *Journal of Applied Probability*, 2018, Vol. 55, No. 2, 593-609.

[16] Lefèvre, C. and Picard, P., and Simon, M., "Epidemic risk and insurance coverage," *Journal of Applied Probability*, 2017, Vol. 54, No. 1, 286-303.

[17] Lefèvre, C. and Simon, M., "Cross-infection in epidemics spread by carriers," *Stochastic Models*, 2018, Vol. 34, No. 2, 166-185.

[18] Lefèvre, C. and Simon, M., "SIR-Type Epidemic Models as Block-Structured Markov Processes," *Methodology and Computing in Applied Probability*, 2019, 1-21.

[19] Newman, M.E., "The structure and function of complex networks," *SIAM review*, 2003, Vol. 45, No. 2, 167-256.

[20] Niu, G., "Actuarial Application of Agent Based Modeling," 2016, Doctoral dissertation, University of Connecticut.

[21] Nkeki, C.I. and Ekhaguere, G.O.S., "Some actuarial mathematical models for insuring the susceptibles of a communicable disease," *International Journal of Financial Engineering*, 2020, Vol. 7, No. 2, 2050014.

[22] Palin, J., Silver, N., Slater, A., and Smith, A.D., "Complexity Economics Application and Relevance to Actuarial Work," In *Report of a working party of the UK actuarial profession. Paper presented to the Finance Investment and Risk Management Conference*, 2008.

[23] Parodi, P., "Computational intelligence with applications to general insurance: a review-I. The role of statistical learning," *Annals of Actuarial Science*, 2012, Vol. 6, No. 2, 307.

[24] Parodi, P., "Computational intelligence with applications to general insurance: a review-II. Dealing with uncertain knowledge," *Annals of Actuarial Science*, 2012, Vol. 6, No. 2, 344-380.

[25] Pastor-Satorras, R., Castellano, C., Van Mieghem, P., and Vespignani, A., "Epidemic processes in complex networks," *Reviews of modern physics*, 2015, Vol. 87, No. 3, 925.

[26] Pellis, L., Ball, F., Bansal, S., Eames, K., House, T., Isham, V., and Trapman, P., "Eight challenges for network epidemic models," *Epidemics*, 2015, Vol. 10, 58-62.

[27] Perera, S.S.N., "An Insurance Based Model to Estimate the Direct Cost of General Epidemic Outbreaks," *Int. J. Pure Appl. Math*, 2017, Vol. 117, No. 14, 183-189.

[28] Shemendyuk, A.A., Chernov, A.A., Kelbert, M.Y., "Fair Insurance Premium Level in Connected SIR Model under Epidemic Outbreak," *arXiv preprint. arXiv:1910.04839*, 2019.

[29] Taylor, G., "A simple model of insurance market dynamics," *North American Actuarial Journal*, 2008, Vol. 12, No. 3, 242-262.

[30] Zhao, S., Lin, Q., Ran, J., Musa, S.S., Yang, G., Wang, W., Lou, Y., Gao, D., Yang, L., He, D., Wang, M.H., "Preliminary estimation of the basic reproduction number of novel coronavirus (2019-nCoV) in China, from 2019 to 2020: A data-driven analysis in the early phase of the outbreak," *International journal of infectious diseases*, 2020, Vol. 92, 214-217.

[31] Zhou, F., "*Application of Agent Based Modeling to Insurance Cycles*," 2013, Doctoral dissertation, City University London.