



DAMSS
DATA ANALYSIS METHODS FOR SOFTWARE SYSTEMS

12th Conference on

DATA ANALYSIS METHODS FOR SOFTWARE SYSTEMS

Druskininkai, Lithuania, Hotel "Europa Royale"
<http://www.mii.lt/DAMSS>

December 2-4, 2021

Several Stochastic Models for Non-Life Insurance Business

Jonas Šiaulyš

Institute of Mathematics
Vilnius University
jonas.siaulyš@mif.vu.lt

The so-called risk renewal model with certain supplements describing the investment environment is commonly used to describe non-life insurance business. The main part of such models is described by the flow of claims. If claims are considered to be identically distributed random variables, then we get so-called homogeneous risk renewal model. If we suppose that the claims are not necessarily identically distributed, then we obtain the inhomogeneous risk renewal model. The report will review several results on the critical characteristics of the inhomogeneous risk renewal model.

On the Computed Tomography Image Data to Diagnose Pancreatic Cancer Using Machine Learning

Aušra Šubonienė¹, Olga Kurasova¹, Viktorija
Aistė Kielaitė-Gulla², Artūras Samuilis³, Džiugaitis
Kęstutis Strupas², Gintautas Dzemyda¹

¹ Institute of Data Science and Digital Technologies
Vilnius University

² Institute of Clinical Medicine, Faculty of Medicine
Vilnius University

³ Institute of Biomedical Sciences, Department of
Nuclear Medicine and Medical Physics, Faculty of
Vilnius University

⁴ Faculty of Medicine
Vilnius University

ausra.suboniene@mif.vu.lt

Medical imaging data, which is suitable for some applications, are difficult to obtain due to data sensitivity issues that is required to make ground truth segmentation. To improve the robustness of results by including more data, different datasets are often combined. However, challenges arise when combine such datasets from different sources. Conditions that differ by age and other conditions could affect the results. There might be different approaches to segmentation. Some methods can segment medical images as true to anatomical structures, or they might include some surrounding tissue. Manual segmentation. Also, rough region boundaries are used for segmentation. Lastly, there can be different preprocessing steps which might result in different pre-processing of the data.

Data sources. Due to data sensitivity and privacy concerns, it is difficult to anonymise images, there is a lack of publicly available pancreatic cancer data. Most publicly available medical data are annotated by experts, and images that do have some annotations are scarce. Currently, the largest publicly available dataset of computed tomography (CT) images of pancreatic cancer is