**Task 6: Predictive and Statistical Modelling**

Task Overview

**What you'll learn**

* Predictive modeling for insurance reserves.
* How to calculate aggregate claims and expected costs.
* How to find Value at Risk (VaR) for reserves.

**What you'll do**

* Create a presentation about reserves and data analysis.
* Review assumptions, fit distributions, and calculate losses.
* Determine VaR and consider modeling enhancements.

Here is the background information on your task

You and Jakob hit the road again. Your last stop is Lucerne, where our client, a mutual insurance association, is domiciled. They offer insurance coverage for a wide range of insurance risks relating to waste incinerator stations and other companies in the waste management industry. Two types of sites are covered:

* Waste incinerator stations (WIS)
* Landfill sites (non-WIS)

The company insures four WIS sites and two 2 non-WIS or “dump” sites in Switzerland. The insurance provides coverage for the following perils:

* Property damage for conventional installations
* Fire
* Machinery breakdown

At WIS sites, the majority of the property is conventional, making these property damage risks comparable to those in other industrial sites and include e.g. turbines. Fire but also small risks such as surveillance cameras are covered. Most of the claims are a result of these small types of risks. However, medium and large losses – up to a few 100 million CHF due to large fire – are possible.

**Here’s where predictive modelling comes in:**

The insurer needs to set up equalisation reserves for annual aggregate claim amounts that may occur in the future. All historically reported claims until the end of 2020 were immediately settled in full or paid out to the insured; no case or IBNR reserves were needed. The purpose of the equalisation reserves is to provide a cushion for fluctuation in the claims experience over the next year.

As a part of the actuarial services team, your job is to derive the aggregate claims distribution over a 1-year time horizon. Based on this distribution, derive the expected annual aggregate claim costs as well as the 80% VaR (Value at Risk). The value of 80% is not untypical for a Swiss insurance company setting up equalisation reserves.

You will need to know several assumptions around claims modelling and distribution, as well as parameters of the claim distribution. Learn about these in the resources provided.

Here is your task

**As part of the Actuarial Services team, you and Jakob prepare a PowerPoint slide deck or a Word document to tackle the task. If needed, provide the R script in your slide deck to explain your assumptions.**

The amount to be booked for the equalisation reserves shall be determined as of 31 December 2020 for claims that potentially occur in the year 2021.

**a)** Do you have all the information available (e.g. regarding claims data, assumptions for distributions or the collective risk model) that is required to fit the data to the collective risk model of each risk type (i.e. WIS, non-WIS) and, thus, to derive the aggregate loss distributions? If not, which information is potentially missing? If any, apply expert judgment and explain your decisions you made.

**b)** Are the assumptions made for each of the claims’ frequency and severity distributions reasonable? How could you validate it?

**c)** What challenges are you most likely facing given the data? How would you characterise the dataset of each type of site?

**d)** How could you validate the provided data set regarding its accuracy and completeness?

**e)** Which parameters do you obtain when you are fitting the claims frequency and severity distributions for each site to the provided data set? Check out Jakob’s hints to get you started.

**f)**What is the expected value of the claims for the current year 2021 at the individual (i.e. for each site) as well as at the aggregate level? Do you need a simulation tool or can the expected values be derived analytically? Jakob has your back, check out his hint.

**g)** Approximate the Value at Risk at level 80% of the (total) aggregate loss distribution for both risk types using the standard normal distribution (or “CLT”) to obtain an estimate of the equalisation reserves to be set up by the insurer. What now? Jakob tells you in some hints.

**h)**Quantify the impact on the expected aggregate loss for year 2021 if the claims frequency for both sites immediately doubles?

**i)**How could the modelling approach be further improved?

**j)** Could you apply Machine Learning methods to this example? Explain.

**Bonus question**: Without performing any simulations, is it possible to estimate the confidence interval at the 80% level for the expected aggregate loss (sound statistical approximations can be applied)?

Jakob notes: for e)–g) all calculations should be rounded to one decimal place. The slides can be simple as we’ll use them internally first. Focus on correct solutions to set up the team with the right insights.

This experience is self-paced. However, Jakob is expected to send the presentation to the client in 1.5 hours and needs your slides and analyses. We recommend you spend no more than that completing this task.

Created the last slide? Well done this concludes your “Tour de Suisse”. Jakob left you a voice message on your way back to Zurich.

*“Hey intern, you’re done. Time to unpack and do your laundry. And post pics from all the Swiss cities you worked in!*

*We hope you’ve enjoyed the Job Simulation with our PwC Digital Intelligence teams.”*

Do you want to learn more? Level up by listening to a portion (or all!) of our podcasts linked below and read our blogs to level up on predictive and statistical modelling to help you complete the task.

**Podcasts**

* [Tech Talk series: Episode 6 - And then there was data](https://www.pwc.lu/en/podcasts-tech-talks/podcasts-tech-talks-1st-season.html)
* [Tech Talk series: Episode 13 - Data - to govern, or not to govern: that is (never) the question](https://www.pwc.lu/en/podcasts-tech-talks/podcasts-tech-talks-1st-season.html)

**Blog posts**

* [Data: Are we speaking the same language?](https://pwc.blogs.com/technology-insights/2016/10/data-are-we-speaking-the-same-language-.html)
* [Big Data = Big opportunity? How does it work?](https://pwc.blogs.com/technology-insights/2016/03/big-data-big-opportunity-how-does-it-work.html)
* [Data analytics: The new win-win-win by Peter Kasahara](https://www.pwc.ch/en/insights/digital/data-analytics-new-win-win-win.html)

**On top**: You must be curious about predictive and statistical modelling, so check out our websites.

* [Predictive analytics](https://www.pwc.ch/en/services/digital/data-analytics/apply-advanced-analytics.html)
* [Actuarial Services & Insurance analytics](https://www.pwc.ch/en/services/assurance/actuarial-services-insurance-analytics.html)
* [Experience analytics](https://www.pwc.ch/en/services/assurance/actuarial-services-insurance-analytics/experience-analysis.html)