Mandatory assignment STK 4540 Autumn 2014

In Norway all buildings and all accessories with insurance policies that cover fire are also automatically insured against natural catastrophes. This is a consequence of Norwegian legislation. The arrangement is managed by 'Norsk Naturskadepool' (Norwegian Pool for Natural Catastrophes) where all Norwegian insurance companies are members. This arrangement guarantees that the consumer receives compensation in the event of a natural catastrophe.

All Norwegian insurance companies pay premium to Norsk Naturskadepool to ensure that its funds are adequate in the event of future natural catastrophes. But what should the premium be? As a recently employed analyst in Norsk Naturskadepool the manager asks YOU to prepare a report addressing this issue.

Your strategy is to calculate the premium using the definition of risk premium, which is claim frequency multiplied by claim severity. The data set contains all Norwegian natural catastrophes since 1980. The catastrophe types include storm, stormflood, flood, avalanche, earthquake and volcanic eruptions. Your task is limited to calculate the premium for storms.

About the assigment: Each assignment should contain a written report as well as R code. The preferred format for the report is computer written (word etc) but hand written is acceptable if the hand writing is neat. Students may discuss problems and solutions with each other but each student must deliver one package each (i.e., report and code).

The assignment is mandatory so you need an approved assignment to be able to participate on the exame. The assignment is delivered to Nils F. Haavardsson personally or to The Expedetition, Department of Mathematics, Nils Henrik Abel's Building 7th floor, no later than November 7th, 2014.

Problem 1

Import the datafiles StormCntPCD.txt and StormCntPEvD.txt into your computer and into R. The files are storm claims only. The structure of the files are

Column 1   number of claims

Column 2   index adjusted claim estimate

Column 3   claim year

Column 4   claim month

Column 5   claim day

Column 6   index counting the days from the start, January 1st, 1980

The file StormCntPCD.txt contains claims *per calendar day*. Natural catastrophes do not occur so often and when they do they occur in clusters. Thus, this file have a lot of entries without claims as well (i.e., the calendar days without natural catastrophes). The file StormCntPEvD.txt contains claims *per event day*. Thus, it only contains entries for those days in the past 34 years when natural catastrophes have occurred.

Problem 2

Find percentiles for the number of claims and claim estimate from StormCntPCD.txt. Use percentile intervals of 5 from 5% up to 90%, then intervals of 1 from 90% to 99% and then 0.1 from 99% to 100%. Comment on the table – what have you learnt about the distribution of the number of claims and the claim estimate?

Problem 3

Try to fit a model for claim frequency. Select the data up to the 95th percentile. Try non-parametric approach, try a Poisson model and try a negative binomial distribution. To estimate lambda in the Poisson model you may for example use total number of event days divided by total number of calendar days. Compare the three models. Create a table that ranges from 0-20, 21-150 and >150 and calculate the number of claims in each category. Calculate the proportion in each category and comment. If you only manage to create a good model up to the 95th percentile what share of the total number of claims have you treated?

Problem 4

Since you are most interested in the consequence of a natural catastrophe from the insurance company's perspective you abandon the idea of developing an advanced frequency model and focus on developing a claim size model. For frequency model purposes you simply assume that the number of natural catastrophes is constant per year (divide the total number of storms by the total number of years).

In the following use the file StormCntPEvD.txt. Remove records with claims estimate that are less than or equal to 0. Try first to fit a model for small and medium sized claims (say up to the 98th percentile). Try some parametric models like log-normal, gamma, Pareto and Weibull

and try the non-parametric. For the parametric models use the maximum likelihood to estimate parameters. Evaluate the quality of the models by using for example qq plot, density plots and comparisons of percentiles in tables. *(Hint: try to write your own code, you may for example review the code from the exercises. If you have tried this for a while and don't know how to proceed, consult exercises 7.9, 7.10, 7.11, 9.13, 9.15, 9.16 from the book for hints concerning the R code for this problem).*

Problem 5

Now focus on large claims. Note that some claims are really extreme. For this reason focus on large claims between the threshold selected in problem 4 and say 150 000 000 NOK. Repeat the drill from problem 4. Pickand stated that if the treshold is high enough, all claims will follow a Pareto distribution. Is Pickand valid for your choice of large claims threshold? What happens if you select a really high large claims threshold like 99.5% or 99.9%? Does Pickand's result seem to be valid for the Norwegian storms?

Problem 6

Propose a model for claim size, for both small and medium sized claims and large events, using your experience from problem 5 and problem 6. For really extreme events calculate the average impact. Thus you have an aggregate model that is split in three regions, small and medium sized claims, large claims and extreme claims.

Sample say 10 000 or 100 000 simulations from the model and estimate the mean premium and the standard deviation from the Monte Carlo simulations. Use the formula for risk premium, claim frequency multiplied by claim severity to calculate the required natural catastrophe premium in each Monte Carlo simulation. What do you think of the uncertainty measured as the standard deviation compared to the mean? How should this concern Norsk Naturskadepool and what could they do to mitigate this effect? If a company has 10% market share, how much should the premium for the company be?