ACTL3162 General Insurance Techniques Assignment

Due time: **Monday 9th October 2017 5 pm**

1 Learning outcomes

The assignment aims particularly at developing the course learning outcomes associated with Aim A and Aim B. It specifically assesses the program goals "Knowledge", "Problem solving and critical thinking", as well as "Communication". You are expected to demonstrate your ability to analyse an actuarial problem, apply appropriate theories and logic to interpret the problem, and develop solutions and conclusions. The communication of those will also be assessed.

2 Assignment task

Task 1

You are an actuarial analyst for a general insurer who introduced a new motor insurance product to the market just over one year ago. During this time, the company has received over 1,000 claims and you now believe the claims experience is significant enough for you to investigate the form of the accident severity distribution. Some policy details are given below:

1. Every policy has a standard excess of $700. (A policyholder may have multiple accidents in the year and may have reported more than one claim. The excess is applied to each accident (or loss), not the total loss in the year.)
2. In addition, policyholders can choose to add on an additional excess in order to lower their premium. There are multiple levels for the policyholder to choose from.

The claims data are stored in data.csv. This file contains claim amounts paid by the insurer, along with the amount of any additional deductible elected by the policyholder.

Your task is to use Maximum Likelihood Estimation (MLE) to fit an appropriate accident severity distribution (i.e. the total loss before deducting the excess) for individual claims. You are required to fit the Pareto, Log-normal and Gamma distributions to the claims data and use appropriate goodness-of-fit tests to decide and subsequently justify which of the three distributions is the most appropriate to use for modelling the claim severity distribution.

In addition, you must briefly describe your methodology in reaching your MLE estimates of your parameters. However, providing detailed mathematical formulas and code snippets is not necessary (but the entire R code (or the code of other software if you are not using R) must be provided separately). (Hint: The MLE estimates for parameters may not be analytically tractable in this case! Consider using a numerical package in R to find the parameters. Also, try various initial conditions for the optimisation in case you're getting errors.)

Task 2

Let C_t be the surplus of an insurer at time t, measured in months. Recall that C_t is defined as

$$C_t = c_0 + \sum_{i=1}^{t}(\pi_i - S_i), \quad t = 0, 1, 2, \ldots \tag{1}$$

where $S_i \sim \text{Gamma}(\alpha, \beta)$ is the total loss for month i and $\pi_i = \pi$ is the monthly premium income, equal to the expected loss plus a loading of 20%. This implies that the premium income is fixed month to month. Let $\psi_t(c_0)$ denote the probability that ruin occurs within time t given initial surplus c_0.

a) Derive an expression for $\psi_1(c_0)$ in terms of the Gamma distribution function G.

b) Provide an explanation for the expression

$$\psi_2(c_0) = \int_0^{c_0+\pi} \psi_1(c_0 + \pi - y)g(y; \alpha, \beta)dy + 1 - G(c_0 + \pi; \alpha, \beta) \tag{2}$$

where g and G are the Gamma density and distribution functions, respectively.

In addition, evaluate this expression for $\alpha = 20$ and $\beta = 0.2$ where $c_0 = 0, 10, 20, 50$. (Hint: You may try to use the 'integrate' function in R!)

c) Explain the difficulty involved in calculating $\psi_t(c_0)$ for larger values of t.

d) Suppose S_i is replaced by a discrete distribution over the non-negative integers $0, 1, 2, \ldots$. We call this new distribution S_i^* with probability mass function g^* and distribution function G^*. We can now calculate the surplus at time t as

$$C_t^* = c_0 + \sum_{i=1}^{t} (\pi_i - S_i^*), \quad t = 0, 1, 2, \ldots \tag{3}$$

Let $\psi_t^*(c_0)$ denote the probability that ruin occurs within time t given initial surplus c_0 in the above model.

- Give an expression for $\psi_1^*(c_0)$.
- Provide a recursive expression for $\psi_t^*(c_0)$.
- Comment on the implementation of this expression in comparison to the original case when S_i was continuous.

e) Your task is now to approximate $\psi_2(c_0)$ by $\psi_2^*(c_0)$ by replacing S with S^* with various methods of discretisation discussed in the lecture slides of Module 4. Use the same values of c_0 as in part b).

(A) Method of rounding

(1) $h = 1, m = 150$
(2) $h = 1, m = 300$
(3) $h = 5, m = 30$
(4) $h = 5, m = 60$

(B) Method of upper and lower bounds

(1) $d = 1$
(2) $d = 5$

Above, d and h refers to the span while m refers to the number of discrete intervals. For (B), you are required to calculate S^* based on **both** the method of upper and lower bounds.

Compare and comment on differences in your results relative to part b). (Hint: You may try and find suitable R packages to help you do the discretisation.)

Task 3

Let $S_i \sim \text{Gamma}(\alpha, \beta)$ where $\alpha = 5$ and $\beta = 0.8$.

a)
- What is the distribution of $Y_5 = \sum_{i=1}^{5} S_i$?
- Compute the distribution function $F_{Y_5}(x)$ at $x = 10, 20, 30, 40, 50$.

b) Now, you will need to discretise S_i and perform convolution using Panjer recursion to approximate the distribution function of Y_5.

- Evaluate this at the same values of x as in part a).
- Explain your methodology and justify any assumptions/decisions you have made in the process of finding an adequate approximation. (Hint: Try using the R package 'actuar'.)

c) Compare the above results in a) and b) and comment on any differences.

Additional Instructions

- Answers are to be provided in Word or pdf format.
- Intermediate steps for questions involving any form of derivation are required. Your comments and conclusions should be well justified and charts should be used to support your conclusions where applicable.
- You are **strongly recommended to use the software R for programming**, although the use of other software will also be accepted. Some sample R codes for fitting are available on the course web site which may be of use. In addition, feel free find packages online to perform your computations (but always check that your answer is sensible!).
- When making a comment or conclusion based on R outputs (or other software outputs), you should include the relevant outputs in the main body of your report. You should make sure that the marker can read and understand your arguments and statements without referring to the appendix.
- Your R codes (or codes of other software) should be included in the appendix. The marker will choose a substantial portion of the reports to check the code. He/she will need to copy the code, run it and check whether it is correct, implementable and consistent with the output presented in your answer. **Students will risk failing the assignment if the code cannot be run or the output provided in the answer is not consistent with the output generated by the code.**

- There is no page limit. However, you should think of a clear and effective structure for your responses. Responses provided of excessive lengths are explicitly penalised in the communications criteria.
- You should not
  - Include programming codes in the main body of your report
  - Have figures or tables that are not referred to or analysed in the main body of your report
  - Include materials that are not highly relevant in the main body of your report

2.1 Communication skills

We recommend students to seek feedback from the EDU (although you need to give them time to read your report) - connect to the EDU website on Moodle “Write well; Learn deeply”. The student enrolment key is “`ASB_LTP`”.

2.2 Assignment submission procedure

Your assignment must be uploaded as a **unique document** (either pdf or Word document) and all parts must be in **portrait format**. As long as the due date is still future, you can resubmit your work; the previous version of your assignment will be replaced by the new version.

Assignments must be submitted via the Turnitin submission box that is available on the course Moodle website. Turnitin reports on any similarities between their own cohort's assignments, and also with regard to other sources (such as the internet or all assignments submitted all around the world via Turnitin). More information is available at: [click]. Please read this page, as we will assume that you are familiar with its content.

Please note that the School of Risk and Actuarial Studies will apply the following policy on **late assignments**. A penalty of 25% of the mark the student would otherwise have obtained, for each full (or part) day of lateness (e.g., 0 day 1 minute = 25% penalty, 2 days 21 hours = 75% penalty). Students who are late must submit their assessment item to the LIC via e-mail. The LIC will then upload documents to the relevant submission boxes. The date and time of reception of the e-mail determines the submission time for the purposes of calculating the penalty.

You need to check your document once it is submitted (check it on-screen). **We will not mark assignments that cannot be read on screen.**

Students are reminded of the risk that technical issues may delay or even prevent their submission (such as internet connection and/or computer breakdowns). Students should then consider either submitting their assignment from the university computer rooms or **allow enough time (at least 24 hours is recommended) between their submission and the due time.** The Turnitin module will not let you submit a late report. **No paper copy will be either accepted or graded.**

In case of a technical problem, the full document must be submitted to the LIC before the due time by e-mail, with explanations about why the student was not able to submit on time. In principle, this assignment will not be marked. It is only in exceptional circumstances where the assignment was submitted before the due time by e-mail that it may be marked—and this only if a valid reason is established (and the LIC has the discretion in deciding whether a given reason is valid).

2.3 Plagiarism awareness

Students are reminded that the work they submit must be their own. While we have no problem with students discussing assignment problems if they wish, the material students submit for assessment must be their own. In particular, this means that any code you present are from your own computer, which you **yourself** developed, without any reference to any other student's work.

While some small elements of code are likely to be similar, big patches of identical code (even with different variable names, layout, or comments—Turnitin picks this up) *will* be considered as plagiarism. The best strategy to avoid any problem is *not* to share bits and pieces of code with other student outside your group.

Note however that you are allowed to use any R code that was made available during the course (either with the lectures or developed in the tutorial exercises). You don't need to reference them formally, and this will *not* be considered as plagiarism.

Students should make sure they understand what plagiarism is—cases of plagiarism have a very high probability of being discovered. For issues of collective work, having different persons marking the assignment does not decrease this probability.

Students should consult the "Write well; Learn deeply" website and consult the resources provided there. In particular, all students should do the quiz about plagiarism to make sure they know how to avoid any issue. For instance, did you know that sharing any part of your work with other students (outside your group) before the deadline is already considered as plagiarism?[1]

3 Assessment criteria

Please see the file, "Rubric".

[1] Yes, that's right, just sending it, even if the third party promises not to copy, is already plagiarism in the UNSW policy!