Non-Life Insurance: Mathematics and Statistics

Exercise sheet 4

Exercise 4.1 Poisson Model and Negative-Binomial Model
Suppose that we are given the following claim count data of ten years:

t	1	2	3	4	5	6	7	8	9	10
N_t	1'000	997	985	989	1'056	1'070	994	986	1'093	1'054
v_t	10'000	10'000	10'000	10'000	10'000	10'000	10'000	10'000	10'000	10'000

Table 1: Observed claims counts N_t and corresponding volumes v_t.

(a) Estimate the claims frequency parameter $\lambda > 0$ of the Poisson model. Moreover, calculate an estimated interval which should contain roughly 70% of the observed claims frequencies N_t/v_t. What do you observe?

(b) Perform a χ^2-goodness-of-fit test at the significance level of 5% to test the null hypothesis of having Poisson distributions.

(c) Estimate the claims frequency parameter $\lambda > 0$ and the dispersion parameter $\gamma > 0$ of the negative-binomial model. Moreover, calculate an estimated interval which should contain roughly 70% of the observed claims frequencies N_t/v_t. What do you observe?

Exercise 4.2 χ^2-Goodness-of-Fit-Analysis (R Exercise)
In this exercise we analyze the sensitivity of the χ^2-goodness-of-fit test (of having a Poisson distribution as claims count distribution) in the situations where the claims counts are simulated from a Poisson distribution and a negative binomial distribution, respectively.

(a) Write an R Code that generates $n = 10'000$ times claims counts $N_1, \ldots, N_T \overset{\text{i.i.d.}}{\sim} \text{Poi}(\lambda v)$ with $T = 10$, $\lambda = 10\%$ and $v = 10'000$. Apply for each of these n replications of $N_1, \ldots, N_T$ a χ^2-goodness-of-fit test of having a Poisson distribution as claims count distribution. Use a significance level of $\alpha = 5\%$. Answer the following questions:

   (i) What can you say about the distribution of the n MLEs of λ?

   (ii) Consider a QQ plot to analyze whether the n values of the test statistic may indeed come from a χ^2-distribution with $T - 1 = 9$ degrees of freedom.

   (iii) How often do we wrongly reject the null hypothesis H_0 of having a Poisson distribution as claims count distribution?

(b) Write an R Code that generates $n = 10'000$ times claims count data $N_1, \ldots, N_T \overset{\text{i.i.d.}}{\sim} \text{NegBin}(\lambda v, \gamma)$ with $T = 10$, $\lambda = 10\%$, $v = 10'000$ and $\gamma \in \{100, 1'000, 10'000\}$. Apply for each of these n replications of $N_1, \ldots, N_T$ a χ^2-goodness-of-fit test of having a Poisson distribution as claims count distribution. Use a significance level of $\alpha = 5\%$. Answer the following questions:

   (i) How often are we able to reject the null hypothesis H_0 of having a Poisson distribution as claims count distribution?

   (ii) Does the size of γ influence this percentage?

Exercise 4.3 Compound Poisson Distribution
For the total claim amount S of an insurance company we assume $S \sim \text{CompPoi}(\lambda v, G)$, where $\lambda = 0.06$, $v = 10$ and for a random variable Y with distribution function G we have

k	100	300	500	6'000	100'000	500'000	2'000'000	5'000'000	10'000'000
$\mathbb{P}[Y = k]$	3/20	4/20	3/20	2/15	2/15	1/15	1/12	1/24	1/24

Table 2: Distribution of $Y \sim G$.

Suppose that the insurance company wants to distinguish between

- small claims: claim size $\leq$ 1'000,
- medium claims: 1'000 $<$ claim size $\leq$ 1'000'000 and
- large claims: claim size $>$ 1'000'000.

Let S_{sc}, S_{mc} and S_{lc} be the total claims in the small claims layer, in the medium claims layer and in the large claims layer, respectively.

(a) Give definitions of S_{sc}, S_{mc} and S_{lc} in terms of mathematical formulas.

(b) Determine the distributions of S_{sc}, S_{mc} and S_{lc}.

(c) What is the dependence structure between S_{sc}, S_{mc} and S_{lc}?

(d) Calculate $\mathbb{E}[S_{\text{sc}}]$, $\mathbb{E}[S_{\text{mc}}]$ and $\mathbb{E}[S_{\text{lc}}]$ as well as $\text{Var}(S_{\text{sc}})$, $\text{Var}(S_{\text{mc}})$ and $\text{Var}(S_{\text{lc}})$. Use these values to calculate $\mathbb{E}[S]$ and $\sqrt{\text{Var}(S)}$.

(e) Calculate the probability that the total claim in the large claims layer exceeds 5 million.

Exercise 4.4 Method of Moments
We assume that the independent claim sizes $Y_1, \ldots, Y_8$ all follow a Gamma distribution with the same unknown shape parameter $\gamma > 0$ and the same unknown scale parameter $c > 0$, and that we have the following observations for $Y_1, \ldots, Y_8$:

$$y_1 = 7, \quad y_2 = 8, \quad y_3 = 10, \quad y_4 = 9, \quad y_5 = 5, \quad y_6 = 11, \quad y_7 = 6, \quad y_8 = 8.$$

Calculate the method of moments estimates of γ and c.