MTH5126 - Statistics for Insurance

Worksheet 3

Q1. Compound distribution

S has a compound distribution with Poisson parameter 4. The individual claim amounts are either 1, with probability 0.3, or 3, with probability 0.7.
Calculate the probability that $S = 4$.

Q2. Moments of compound distributions

An insurance portfolio contains policies for three categories of policyholder: A, B and C. The number of claims in a year, N, on an individual policy follows a Poisson distribution with mean λ. Individual claim sizes are assumed to be exponentially distributed with mean 4 and are independent from claim to claim. The distribution of λ, depending on the category of the policyholder, is:

Category	Value of λ	Proportion of policyholders
A	2	20%
B	3	60%
C	4	20%

Denote by S the total amount claimed by a policyholder in one year.

1. Prove that $E(S) = E\,[E(S|\lambda)]$
2. Show that $E(S|\lambda) = 4\lambda$ and $Var(S|\lambda) = 32\lambda$
3. Calculate $E(S)$
4. Calculate $Var(S)$

Q3. R

Before answering this question, generate the vector, X, in R using the following code:

```
set.seed(1027); X = rexp(n=1000, rate=0.01)
```

The vector X represents the gross claim sizes of 1,000 claims. The payments are to be split between an insurance company and its reinsurer under an Excess of Loss reinsurance arrangement with a retention level $M = 400$.

(i) Determine the proportion of the claims that are fully covered by the insurer. [2]

Hint: The following code might help.

```
length(X[X<=M])/length(X)
```

(ii) Generate an additional vector, *Y*, which is of the same length as *X*, such that *Y* represents the amounts to be paid by the insurer for each component of *X*. [1]

Hint: Use the `pmin` function.

(iii) Generate an additional vector, *Z*, which is of the same length as *X*, such that *Z* represents the amounts to be paid by the reinsurer for each component of *X*. [1]

An actuary assumes that the underlying gross claims distribution follows an exponential distribution of some unknown rate λ. The actuary needs to estimate λ using only the claim amounts recorded in vector *Y*.

(iv) Construct R code that calculates the log-likelihood, as a function of the parameter λ, given the claim amounts data in vector *Y*. [10]

Hint: This is estimation when sample is censored, see lecture slides.

(v) Using the function `nlm`, determine the value of λ at which the log-likelihood function reaches its maximum. [6]

Hint: The `nlm` function performs minimisation, not maximisation. However, maximising the log-likelihood function is the same as minimising the negative log-likelihood. So, we first define the function that we want to hand to `nlm` to be minimised.