Coordinator
Andrea Gabrielli

Non-Life Insurance: Mathematics and Statistics

Exercise sheet 1

Exercise 1.1 Discrete Distribution
Suppose that N follows a geometric distribution with parameter $p \in (0,1)$, i.e.

$$\mathbb{P}[N = k] \;=\; \begin{cases} (1-p)^{k-1}p, & \text{if } k \in \mathbb{N}_{>0}, \\ 0, & \text{else.} \end{cases}$$

(a) Show that the geometric distribution indeed defines a probability distribution on $\mathbb{R}$.

(b) Let $n \in \mathbb{N}_{>0}$. Calculate $\mathbb{P}[N \geq n]$.

(c) Calculate $\mathbb{E}[N]$.

(d) Let $r < -\log(1-p)$. Calculate the moment generating function $M_N(r) = \mathbb{E}[\exp\{rN\}]$ of N.

(e) Calculate $\frac{d}{dr} M_N(r)\big|_{r=0}$. What do you observe?

Exercise 1.2 Absolutely Continuous Distribution
Suppose that Y follows an exponential distribution with parameter $\lambda > 0$, i.e. the density f_Y of Y is given by

$$f_Y(x) \;=\; \begin{cases} \lambda \exp\{-\lambda x\}, & \text{if } x \geq 0, \\ 0, & \text{else.} \end{cases}$$

(a) Show that the exponential distribution indeed defines a probability distribution on $\mathbb{R}$.

(b) Let $0 < y_1 < y_2$. Calculate $\mathbb{P}[y_1 \leq Y \leq y_2]$.

(c) Calculate $\mathbb{E}[Y]$ and $\mathrm{Var}(Y)$.

(d) Let $r < \lambda$. Calculate the cumulant generating function $\log M_Y(r) = \log \mathbb{E}[\exp\{rY\}]$ of Y.

(e) Calculate $\frac{d^2}{dr^2} \log M_Y(r)\big|_{r=0}$. What do you observe?

Exercise 1.3 Gaussian Distribution
For a random variable X we write $X \sim \mathcal{N}(\mu, \sigma^2)$ if X follows a Gaussian distribution with mean $\mu \in \mathbb{R}$ and variance $\sigma^2 > 0$. The density f_X of $X \sim \mathcal{N}(\mu, \sigma^2)$ is given by

$$f_X(x) \;=\; \frac{1}{\sqrt{2\pi}\sigma} \exp\left\{-\frac{1}{2}\frac{(x-\mu)^2}{\sigma^2}\right\}, \qquad \text{for all } x \in \mathbb{R}.$$

(a) Show that the moment generating function M_X of $X \sim \mathcal{N}(\mu, \sigma^2)$ is given by

$$M_X(r) \;=\; \exp\left\{r\mu + \frac{r^2\sigma^2}{2}\right\}, \qquad \text{for all } r \in \mathbb{R}.$$

(b) Let $X \sim \mathcal{N}(\mu, \sigma^2)$ and $a, b \in \mathbb{R}$. Show that

$$a + bX \sim \mathcal{N}(a + b\mu, b^2\sigma^2).$$

(c) Let $X_1, \ldots, X_n$ be independent with $X_i \sim \mathcal{N}(\mu_i, \sigma_i^2)$ for all $i \in \{1, \ldots, n\}$. Show that

$$\sum_{i=1}^n X_i \sim \mathcal{N}\left(\sum_{i=1}^n \mu_i, \sum_{i=1}^n \sigma_i^2\right).$$

Exercise 1.4 χ^2-Distribution

For all $k \in \mathbb{N}_{>0}$ we assume that X_k has a χ^2-distribution with k degrees of freedom, i.e. X_k has density

$$f_{X_k}(x) = \begin{cases} \frac{1}{2^{k/2}\Gamma(k/2)} x^{k/2-1} \exp\{-x/2\}, & \text{if } x \geq 0, \\ 0, & \text{else.} \end{cases}$$

(a) Let M_{X_k} be the moment generating function of X_k. Show that

$$M_{X_k}(r) = \frac{1}{(1-2r)^{k/2}}, \qquad \text{for } r < 1/2.$$

(b) Let $Z \sim \mathcal{N}(0,1)$. Show that $Z^2 \overset{(d)}{=} X_1$.

(c) Let $Z_1, \ldots, Z_k \overset{\text{i.i.d.}}{\sim} \mathcal{N}(0,1)$. Show that $\sum_{i=1}^k Z_i^2 \overset{(d)}{=} X_k$.

Non-Life Insurance: Mathematics and Statistics

Exercise sheet 2

Exercise 2.1 Maximum Likelihood and Hypothesis Test
Let $Y_1, \ldots, Y_n$ be claim amounts in CHF that an insurance company has to pay. We assume that $Y_1, \ldots, Y_n$ are independent and identically distributed (i.i.d.) random variables following a log-normal distribution with unknown parameters $\mu \in \mathbb{R}$ and $\sigma^2 > 0$. Then, by definition, $\log Y_1, \ldots, \log Y_n$ are i.i.d. Gaussian random variables with mean $\mu \in \mathbb{R}$ and variance $\sigma^2 > 0$. Let $n = 8$ and suppose that we have the following observations $x_1, \ldots, x_8$ for $\log Y_1, \ldots, \log Y_8$:

$$x_1 = 9, \qquad x_2 = 4, \qquad x_3 = 6, \qquad x_4 = 7, \qquad x_5 = 3, \qquad x_6 = 11, \qquad x_7 = 6, \qquad x_8 = 10.$$

(a) Write down the joint density $f_{\mu,\sigma^2}(x_1, \ldots, x_8)$ of $\log Y_1, \ldots, \log Y_8$.

(b) Calculate $\log f_{\mu,\sigma^2}(x_1, \ldots, x_8)$.

(c) Calculate the maximum likelihood estimates (MLEs)

$$(\widehat{\mu}, \widehat{\sigma}^2) = \arg \max_{(\mu,\sigma^2)\in\mathbb{R}\times\mathbb{R}_{>0}} \log f_{\mu,\sigma^2}(x_1, \ldots, x_8).$$

(d) Now suppose that we are interested in the mean μ of the logarithms of the claim amounts. An expert claims that $\mu = 6$. Perform a statistical test to test the null hypothesis H_0: $\mu = 6$ against the (two-sided) alternative hypothesis H_1: $\mu \neq 6$.

Exercise 2.2 Chebychev's Inequality and Law of Large Numbers
Suppose that an insurance company provides insurance against bike theft. In our model a bike gets stolen with a probability of 0.1, and in case of a theft the insurance company has to pay 1'000 CHF. We assume that we have n i.i.d. risks $X_1, \ldots, X_n$ with

$$X_i = \begin{cases} 1'000, & \text{with probability } 0.1, \\ 0, & \text{with probability } 0.9, \end{cases}$$

for all $i = 1, \ldots, n$. In this exercise we are interested in the probability

$$p(n) \stackrel{\text{def}}{=} \mathbb{P}\left[\left|\frac{1}{n}\sum_{i=1}^{n} X_i - \mu\right| \geq 0.1\mu\right]$$

of a deviation of the sample mean $\frac{1}{n}\sum_{i=1}^{n} X_i$ to the mean claim size $\mu = \mathbb{E}[X_1]$ of at least 10%, and how diversification effects this probability.

(a) Calculate μ.

(b) Suppose that $n = 1$. Calculate $p(1)$.

(c) Suppose that $n = 1'000$. Calculate $p(1'000)$.

(d) Apply Chebychev's inequality to derive a minimum number n of risks such that $p(n) < 0.01$.

(e) What can you say about $\lim_{n\to\infty} \frac{1}{n}\sum_{i=1}^{n} X_i$?

Exercise 2.3 Central Limit Theorem

Let n be the number of claims and $Y_1, \dots, Y_n$ the corresponding claim sizes, where we assume that $Y_1, \dots, Y_n$ are i.i.d. random variables with expectation $\mathbb{E}[Y_1] = \mu$ and coefficient of variation $\mathrm{Vco}(Y_1) = 4$.

(a) Use the Central Limit Theorem to determine an approximate minimum number of claims n^{CLT} such that

$$\mathbb{P}\left[\left|\frac{1}{n^{\mathrm{CLT}}}\sum_{i=1}^{n^{\mathrm{CLT}}} Y_i - \mu\right| < 0.01\mu\right] \geq 0.95,$$

i.e. with probability of at least 95% the deviation of the sample mean $\frac{1}{n}\sum_{i=1}^{n} Y_i$ from the mean claim size μ is less than 1%.

(b) Compare the resulting minimum number of claims n^{CLT} to the corresponding minimum number of claims n^{Che} when using Chebychev's inequality instead of the Central Limit Theorem in part (a). What do you observe?

Exercise 2.4 Conditional Distribution and Variance Decomposition

Suppose that Θ follows an exponential distribution with parameter $\lambda > 0$. We assume that, conditionally given Θ, the number of claims N in a particular line of business of an insurance company is modeled by a Poisson distribution with frequency parameter Θv, where $v > 0$ denotes the volume, i.e. we have

$$\mathbb{P}[N = k|\Theta] = \begin{cases} e^{-\Theta v}\frac{(\Theta v)^k}{k!}, & \text{if } k \in \mathbb{N}, \\ 0, & \text{else.} \end{cases}$$

We remark that the expectation and the variance of a Poisson distribution are equal to its frequency parameter, i.e. here we have $\mathbb{E}[N|\Theta] = \mathrm{Var}(N|\Theta) = \Theta v$.

(a) Calculate $\mathbb{P}[N = 0]$.

(b) Calculate $\mathbb{E}[N]$.

(c) Show that $\mathrm{Var}(N) = \mathbb{E}[\mathrm{Var}(N|\Theta)] + \mathrm{Var}(\mathbb{E}[N|\Theta])$ and use this result to calculate $\mathrm{Var}(N)$.

Non-Life Insurance: Mathematics and Statistics

Exercise sheet 3

Exercise 3.1 No-Claims Bonus
An insurance company decides to offer a no-claims bonus to good car drivers, namely

- a 10% discount on the premium after three years of no claim, and
- a 20% discount on the premium after six years of no claim.

How does the premium need to be adjusted such that the premium income is expected to remain the same as before the grant of the no-claims bonus? For simplicity, we consider one car driver who has been insured for at least six years. Answer the question in the following two situations:

(a) The claim counts of the individual years of the considered car driver are i.i.d. Poisson distributed random variables with frequency parameter $\lambda = 0.2$.

(b) Suppose Θ follows an exponential distribution with parameter $c = 1$. Conditionally given Θ, the claim counts of the individual years of the considered car driver are i.i.d. Poisson distributed random variables with frequency parameter $\Theta\lambda$, where $\lambda = 0.2$ as above.

Exercise 3.2 Compound Poisson Distribution
For the total claim amount S of an insurance company we assume $S \sim \mathrm{CompPoi}(\lambda v, G)$, 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 1: Claim size distribution $Y \sim G$.

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

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

Exercise 3.3 Compound Distribution
Assume that

$$S = \sum_{i=1}^{N} Y_i$$

has a compound distribution with N having a geometric distribution with parameter $p \in (0,1)$ and $Y_1, Y_2, \ldots$ being i.i.d. exponentially distributed with parameter $\lambda > 0$. Show that S is exponentially distributed with parameter λp.

Exercise 3.4 Compound Binomial Distribution
Assume $S \sim \mathrm{CompBinom}(v, p, G)$ for given $v \in \mathbb{N}$, $p \in (0,1)$ and individual claim size distribution G. Let $M > 0$ such that $G(M) \in (0,1)$. Define the compound distribution S_{sc} with individual claims Y_i that are at most of size M and the compound distribution S_{lc} with individual claims Y_i that exceed threshold M by, respectively,

$$S_{\mathrm{sc}} = \sum_{i=1}^{N} Y_i \, 1_{\{Y_i \le M\}} \qquad \text{and} \qquad S_{\mathrm{lc}} = \sum_{i=1}^{N} Y_i \, 1_{\{Y_i > M\}}.$$

(a) Show that $S_{\mathrm{lc}} \sim \mathrm{CompBinom}(v, p[1 - G(M)], G_{\mathrm{lc}})$, where the large claims size distribution function G_{lc} satisfies $G_{\mathrm{lc}}(y) = \mathbb{P}[Y_1 \le y | Y_1 > M]$.

(b) Give a short argument which shows that S_{sc} and S_{lc} are not independent.

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 claim counts N_t and corresponding volumes v_t.

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

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

(c) Estimate the claim frequency parameter $\lambda > 0$ and the dispersion parameter $\gamma > 0$ of the negative-binomial model. Moreover, calculate a prediction interval which should contain roughly 70% of the observed claim 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 claim count distribution) in situations where the claim counts are simulated from a Poisson distribution and a negative binomial distribution, respectively.

(a) Write an R code that generates $n = 10'000$ times claim counts $N_1, \ldots, N_T \stackrel{\text{i.i.d.}}{\sim} \text{Poi}(\lambda v)$ 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 at significance level of 5% of having a Poisson distribution as claim count distribution. 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 claim count distribution?

(b) Write an R code that generates $n = 10'000$ times claim counts $N_1, \ldots, N_T \stackrel{\text{i.i.d.}}{\sim} \text{NegBin}(\lambda v, \gamma)$ 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 at significance level of 5% of having a Poisson distribution as claim count distribution. Answer the following questions:

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

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

Exercise 4.3 Claim Count Distribution
Suppose that in a given line of business of an insurance company the numbers of claims of the last ten years are modeled by random variables $N_1, \ldots, N_{10}$. We assume that $N_1, \ldots, N_{10}$ are i.i.d. and that we have collected the following observations:

t	1	2	3	4	5	6	7	8	9	10
N_t	7	21	19	18	25	17	33	6	39	28

Table 2: Observed numbers of claims N_t over the last ten years.

Which claim count distribution would you prefer in this situation? Give a short argument.

Exercise 4.4 Method of Moments
The i.i.d. claim sizes $Y_1, \ldots, Y_8$ are supposed to follow a Gamma distribution with unknown shape parameter $\gamma > 0$ and unknown scale parameter $c > 0$. We assume that we have the following observations for $Y_1, \ldots, Y_8$:

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

Use the method of moments to estimate the parameters γ and c.

Non-Life Insurance: Mathematics and Statistics

Exercise sheet 5

Exercise 5.1 Large Claims

In this exercise we are interested in storm and flood events with claim amounts exceeding CHF 50 million as an example of large claims modeling. Assume that for the total yearly claim amount S of storm and flood events with claim amounts exceeding CHF 50 million we model $S \sim \text{CompPoi}(\lambda, G)$, where λ is unknown and G is the distribution function of a Pareto distribution with threshold $\theta = 50$ and unknown tail index $\alpha > 0$. Note that we set $\theta = 50$ as we will work in units of CHF 1 million. Moreover, all numbers of claims and claim sizes and, thus, also the total claim amounts are assumed to be independent across different years. During the years $1986 - 2005$ we observed the following 15 storm and flood events with corresponding claim amounts in CHF millions:

date	amount in millions	date	amount in millions
20.06.1986	52.8	18.05.1994	78.5
18.08.1986	135.2	18.02.1999	75.3
18.07.1987	55.9	12.05.1999	178.3
23.08.1987	138.6	26.12.1999	182.8
26.02.1990	122.9	04.07.2000	54.4
21.08.1992	55.8	13.10.2000	365.3
24.09.1993	368.2	20.08.2005	1'051.1
08.10.1993	83.8		

Table 1: Dates and claim amounts in CHF millions of the 15 storm and flood events observed during the years 1986-2005.

(a) Show that the MLE $\widehat{\alpha}_n^{\text{MLE}}$ of α for n i.i.d. claim sizes $Y_1, \ldots, Y_n \sim \text{Pareto}(\theta, \alpha)$ is given by

$$\widehat{\alpha}_n^{\text{MLE}} = \left(\frac{1}{n} \sum_{i=1}^{n} \log Y_i - \log \theta \right)^{-1}.$$

(b) According to Lemma 3.8 of the lecture notes (version of March 20, 2019),

$$\frac{n-1}{n} \, \widehat{\alpha}_n^{\text{MLE}}$$

is an unbiased version of the MLE. Estimate α using the unbiased version of the MLE for the storm and flood data given in Table 1.

(c) Calculate the MLE of λ for the storm and flood data given in Table 1.

(d) Suppose that we introduce a maximal claims cover of CHF $M = 2$ billion per storm and flood event, i.e. the individual claims are given by $\min\{Y_i, M\}$. Using the estimates of α and λ found in parts (b) and (c), calculate the estimated expected total yearly claim amount.

(e) Using the estimates of α and λ found in parts (b) and (c), calculate the estimated probability that we observe at least one storm and flood event next year which exceeds the level of CHF $M = 2$ billion.

Exercise 5.2 Claim Size Distributions (R Exercise)
Write an R code that generates i.i.d. samples of size $n = 10'000$ from each of the following distributions:

- $\Gamma(\gamma, c)$ with shape parameter $\gamma = \frac{1}{4}$ and scale parameter $c = \frac{1}{40'000}$,
- Weibull(τ, c) with shape parameter $\tau = 0.54$ and scale parameter $c = 0.000175$,
- LN(μ, σ^2) with mean parameter $\mu = \log\left(2000\sqrt{5}\right)$ and variance parameter $\sigma^2 = \log(5)$,
- Pareto(θ, α) with threshold $\theta = 10'000\frac{\sqrt{5}}{2+\sqrt{5}}$ and tail index $\alpha = 1 + \frac{\sqrt{5}}{2}$.

Note that the parameters are chosen such that the theoretical expectations and standard deviations are approximately equal to 10'000 and 20'000, respectively, for all the distributions listed above. For each of these i.i.d. samples consider

- the density plot (on the log scale),
- the box plot (on the log scale),
- the plot of the empirical distribution function (on the log scale),
- the plot of the empirical loss size index function,
- the empirical log-log plot,
- the plot of the empirical mean excess function.

Comment on your results.

Exercise 5.3 Hill Estimator (R Exercise)
Write an R code that samples 300 i.i.d. observations from a Pareto distribution with threshold $\theta = 10$ and tail index $\alpha = 2$. Create a Hill plot and a log-log plot. What do you observe?

Exercise 5.4 Pareto Distribution
Suppose the random variable Y follows a Pareto distribution with threshold $\theta > 0$ and tail index $\alpha > 0$.

(a) Show that the survival function of Y is regularly varying at infinity with tail index α.

(b) Show that for $\theta \leq u_1 < u_2$ the expected value of Y within the layer $(u_1, u_2]$ is given by

$$\mathbb{E}\left[Y 1_{\{u_1 < Y \leq u_2\}}\right] = \begin{cases} \theta \frac{\alpha}{\alpha-1}\left[\left(\frac{u_1}{\theta}\right)^{-\alpha+1} - \left(\frac{u_2}{\theta}\right)^{-\alpha+1}\right], & \text{if } \alpha \neq 1, \\ \theta \log\left(\frac{u_2}{u_1}\right), & \text{if } \alpha = 1. \end{cases}$$

(c) Show that for $\alpha > 1$ and $y > \theta$ the loss size index function for level y is given by

$$\mathcal{I}[G(y)] = 1 - \left(\frac{y}{\theta}\right)^{-\alpha+1}.$$

(d) Show that for $\alpha > 1$ and $u > \theta$ the mean excess function of Y above u is given by

$$e(u) = \frac{1}{\alpha - 1}\, u.$$

Coordinator
Andrea Gabrielli

Non-Life Insurance: Mathematics and Statistics

Exercise sheet 6

Exercise 6.1 Goodness-of-Fit Test
Suppose we are given the following claim size data (in increasing order) coming from independent realizations of an unknown claim size distribution:

210, 215, 228, 232, 303, 327, 344, 360, 365, 379, 402, 413, 437, 481, 521, 593, 611, 677, 910, 1623.

(a) Use the intervals

$$I_1 = [200, 239), \quad I_2 = [239, 301), \quad I_3 = [301, 416), \quad I_4 = [416, 725), \quad I_5 = [725, +\infty)$$

to perform a χ^2-goodness-of-fit test at significance level of 5% to test the null hypothesis of having a Pareto distribution with threshold $\theta = 200$ and tail index $\alpha = 1.25$ as claim size distribution.

(b) In goodness-of-fit tests with K disjoint intervals and a total of n observations we use the test statistic

$$X^2_{n,K} = \sum_{k=1}^{K} \frac{(O_k - E_k)^2}{E_k},$$

where O_k denotes the actual number of observations and E_k the expected number of observations in the k-th interval. We assume that the parameters of the null hypothesis distribution function are given and that the K disjoint intervals are chosen such that $E_k > 0$, for all $k = 1, \ldots, K$. Show that in case of $K = 2$ disjoint intervals, the test statistic $X^2_{n,2}$ converges to a χ^2-distribution with one degree of freedom, as $n \to \infty$.

Exercise 6.2 Log-Normal Distribution and Deductible
Assume that the total claim amount

$$S = \sum_{i=1}^{N} Y_i$$

in a given line of business has a compound distribution with $\mathbb{E}[N] = \lambda v$, where $\lambda > 0$ denotes the claim frequency and $v > 0$ the volume, and with a log-normal distribution with mean parameter $\mu \in \mathbb{R}$ and variance parameter $\sigma^2 > 0$ as claim size distribution.

(a) Show that

$$\begin{aligned} \mathbb{E}[Y_1] &= \exp\left\{\mu + \frac{\sigma^2}{2}\right\}, \\ \mathrm{Var}(Y_1) &= \exp\left\{2\mu + \sigma^2\right\}\left(\exp\left\{\sigma^2\right\} - 1\right) \quad \text{and} \\ \mathrm{Vco}(Y_1) &= \sqrt{\exp\left\{\sigma^2\right\} - 1}. \end{aligned}$$

(b) Suppose that $\mathbb{E}[Y_1] = 3\text{'}000$ and $\mathrm{Vco}(Y_1) = 4$. Up to now, the insurance company was not offering contracts with deductibles. Now it wants to offer a deductible of $d = 500$. Answer the following questions:

(i) How does the claim frequency λ change by the introduction of the deductible?
(ii) How does the expected claim size $\mathbb{E}[Y_1]$ change by the introduction of the deductible?
(iii) How does the expected total claim amount $\mathbb{E}[S]$ change by the introduction of the deductible?

Exercise 6.3 Kolmogorov-Smirnov Test
Suppose we are given the following data (in increasing order) coming from independent realizations of an unknown distribution:

$$x_1 = \left(-\log\frac{38}{40}\right)^2, x_2 = \left(-\log\frac{37}{40}\right)^2, x_3 = \left(-\log\frac{35}{40}\right)^2, x_4 = \left(-\log\frac{34}{40}\right)^2, x_5 = \left(-\log\frac{10}{40}\right)^2.$$

Perform a Kolmogorov-Smirnov test at significance level of 5% to test the null hypothesis that the data given above comes from a Weibull distribution with shape parameter $\tau = \frac{1}{2}$ and scale parameter $c = 1$. Moreover, explain why the Kolmogorov-Smirnov test is applicable in this example.

Exercise 6.4 Akaike Information Criterion and Bayesian Information Criterion
Assume that we fit a gamma distribution to a set of $n = 1\text{'}000$ i.i.d. claim sizes and that we obtain the following method of moments (MM) estimates and maximum likelihood estimates (MLE):

$$\begin{aligned} \widehat{\gamma}^{\text{MM}} &= 0.9794 \quad \text{and} \quad \widehat{c}^{\text{MM}} = 9.4249, \\ \widehat{\gamma}^{\text{MLE}} &= 1.0013 \quad \text{and} \quad \widehat{c}^{\text{MLE}} = 9.6360. \end{aligned}$$

The corresponding log-likelihoods are given by

$$\ell_{\mathbf{Y}}\left(\widehat{\gamma}^{\text{MM}}, \widehat{c}^{\text{MM}}\right) = 1\text{'}264.013 \quad \text{and} \quad \ell_{\mathbf{Y}}\left(\widehat{\gamma}^{\text{MLE}}, \widehat{c}^{\text{MLE}}\right) = 1\text{'}264.171.$$

(a) Why is $\ell_{\mathbf{Y}}\left(\widehat{\gamma}^{\text{MLE}}, \widehat{c}^{\text{MLE}}\right) > \ell_{\mathbf{Y}}\left(\widehat{\gamma}^{\text{MM}}, \widehat{c}^{\text{MM}}\right)$? Which fit should be preferred according to the Akaike Information Criterion (AIC)?

(b) The estimates of γ are very close to 1 and, thus, we could also use an exponential distribution as claim size distribution. For the exponential distribution we obtain the MLE $\widehat{c}^{\text{MLE}} = 9.6231$ and the corresponding log-likelihood $\ell_{\mathbf{Y}}\left(\widehat{c}^{\text{MLE}}\right) = 1\text{'}264.169$. According to the AIC and the Bayesian Information Criterion (BIC), should we prefer the gamma model or the exponential model?

Non-Life Insurance: Mathematics and Statistics

Exercise sheet 7

Exercise 7.1 Re-Insurance Covers and Leverage Effect
In Figure 1 we compare the distribution function of a loss $Y \sim \Gamma(1, \frac{1}{400})$ (in black color) to the distribution function of the loss after applying different re-insurance covers to Y (in red color).

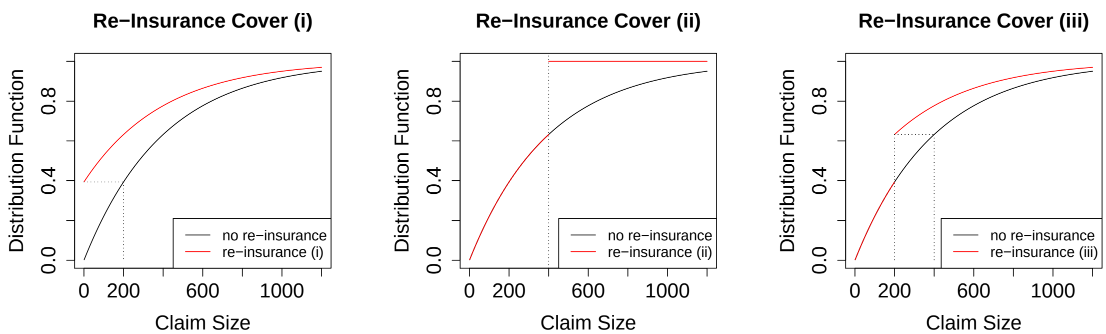

Figure 1: Distribution functions implied by re-insurance contracts (i), (ii) and (iii).

(a) Let $d > 0$. Show that Y satisfies

$$\mathbb{E}[(Y-d)_+] = \mathbb{P}[Y > d]\mathbb{E}[Y].$$

(b) Can you explicitly determine the re-insurance covers from the graphs in Figure 1?

(c) Calculate the expected values of these modified contracts.

(d) Assume that a first claim Y_0 has the same distribution as Y, and that a second claim Y_1 fulfills $Y_1 \overset{(d)}{=} (1+i)Y_0$, for a constant inflation rate $i > 0$. Let $d > 0$. Show the leverage effect

$$\mathbb{E}[(Y_1-d)_+] > (1+i)\mathbb{E}[(Y_0-d)_+].$$

Give an appropriate explanation for this leverage effect.

Exercise 7.2 Inflation and Deductible
We assume that this year's claims in a storm insurance portfolio have been modeled by a Pareto distribution with threshold $\theta > 0$ and tail index $\alpha > 1$. The threshold θ can be understood as deductible. Suppose that the inflation in the next year is expected to be $100 \cdot r\,\%$ for some $r > 0$. By how much do we have to increase the deductible θ next year such that the average claim payment remains unchanged?

Exercise 7.3 Normal Approximation
Assume that the total claim amount S has a compound Poisson distribution with expected number of claims $\lambda v = 1'000$ and claim sizes following a gamma distribution with shape parameter $\gamma = 100$ and scale parameter $c = \frac{1}{10}$. Use the normal approximation to estimate the 0.95-quantile $q_{0.95}$ and the 0.99-quantile $q_{0.99}$ of S.

Exercise 7.4 Translated Gamma and Translated Log-Normal Approximation
Consider the same setup as in Exercise 7.3. This time we use the translated gamma and the translated log-normal approximation to estimate the quantiles $q_{0.95}$ and $q_{0.99}$ of the total claim amount S.

(a) Use the translated gamma approximation to estimate $q_{0.95}$ and $q_{0.99}$.

(b) Use the translated log-normal approximation to estimate $q_{0.95}$ and $q_{0.99}$.

(c) Compare the results of (a) and (b) to the results found in Exercise 7.3, where we used the normal approximation.

Non-Life Insurance: Mathematics and Statistics

Exercise sheet 8

Exercise 8.1 Panjer Algorithm

In this exercise we use the Panjer algorithm to calculate monthly health insurance premiums for different franchises d. We assume that the yearly claim amount

$$S = \sum_{i=1}^{N} Y_i$$

of a given customer is compound Poisson distributed with $N \sim \mathrm{Poi}(1)$ and $Y_1 \overset{(d)}{=} k + Z$, where $k = 100$ CHF and $Z \sim \mathrm{LN}(\mu = 7.8, \sigma^2 = 1)$. In health insurance the policyholder can choose between different franchises $d \in \{300, 500, 1'000, 1'500, 2'000, 2'500\}$. The franchise d describes the threshold up to which the policyholder has to pay everything by himself. Moreover, the policyholder has to pay $\alpha = 10\%$ of the part of the total claim amount S that exceeds the franchise d, but only up to a maximal amount of $M = 700$ CHF. Thus, the yearly amount paid by the customer is given by

$$S_{\text{ins}} = \min\{S, d\} + \min\{\alpha \cdot (S - d)_+, M\}.$$

If we define $\pi_0 = \mathbb{E}[S]$ and $\pi_{\text{ins}} = \mathbb{E}[S_{\text{ins}}]$, the monthly pure risk premium π is given by

$$\pi = \frac{\pi_0 - \pi_{\text{ins}}}{12}.$$

Calculate π for the different franchises $d \in \{300, 500, 1'000, 1'500, 2'000, 2'500\}$ using the Panjer algorithm. In order to apply the Panjer algorithm, discretize the translated log-normal distribution using a span of $s = 10$ and putting all the probability mass to the upper end of the intervals.

Exercise 8.2 Monte Carlo Simulations (R Exercise)

Consider the same setup as in Exercise 7.3. This time we use Monte Carlo simulations to determine the distribution of the total claim amount S.

(a) Write an R code that simulates $n = 100'000$ times the total claim amount S. Compare the resulting distribution function of S to the approximate distribution functions found in Exercises 7.3 and 7.4, where we used the normal, the translated gamma and the translated log-normal approximation.

(b) Write an R code that simulates $n \in \{100, 1'000, 10'000\}$ times the total claim amount S and replicates these simulations 100 times. For each $n \in \{100, 1'000, 10'000\}$ discuss the distribution of the resulting 100 values of the quantiles $q_{0.95}$ and $q_{0.99}$ of S.

Exercise 8.3 Fast Fourier Transform (R Exercise)

Consider the same setup as in Exercise 7.3. Write an R code that applies the fast Fourier transform using a threshold of $n = 2'000'000$ in order to determine the distribution function of the total claim amount S. Compare the resulting distribution function to the distribution function found in Exercise 8.2, where we used Monte Carlo simulations. Moreover, determine the quantiles $q_{0.95}$ and $q_{0.99}$ of S and compare them to the values found in Exercises 7.3 and 7.4, where we used the normal, the translated gamma and the translated log-normal approximation.

Exercise 8.4 Panjer Distribution

Let N be a random variable that has a Panjer distribution with parameters $a, b \in \mathbb{R}$. Calculate $\mathbb{E}[N]$ and $\mathrm{Var}(N)$. What can you say about the ratio of $\mathrm{Var}(N)$ to $\mathbb{E}[N]$?

Non-Life Insurance: Mathematics and Statistics

Exercise sheet 9

Exercise 9.1 Utility Indifference Price
In this exercise we calculate the premium for the accident insurance of a given company COMP using the utility indifference price principle. We suppose that all employees of COMP have been divided into two groups, depending on their work, and that the total claim amounts S_1 and S_2 of the two groups are independent and compound Poisson distributed with volumes, claim frequencies and claim size distributions as given in Table 1.

Group i	v_i	λ_i	$Y_1^{(i)}$
1	2'000	50%	$\Gamma(\gamma = 20, c = 0.01)$
2	10'000	10%	$\text{expo}(\kappa = 0.005)$

Table 1: Volumes, claim frequencies and claim size distributions for the two groups of employees.

We write $S = S_1 + S_2$ for the total claim amount of COMP. Let c_0 be the initial capital of the insurance company that sells accident insurance to COMP.

(a) Let u be a risk-averse utility function. Show that if the utility indifference price $\pi = \pi(u, S, c_0)$ exists, then it is unique and satisfies $\pi > \mathbb{E}[S]$.

(b) Calculate $\mathbb{E}[S]$.

(c) Calculate π using the utility indifference price principle for the exponential utility function with parameter $\alpha = 1.5 \cdot 10^{-6}$.

(d) What happens to π if we replace the compound Poisson distributions of S_1 and S_2 by Gaussian distributions with the same corresponding first two moments?

(e) For this part we assume that S has a general compound Poisson distribution with expected number of claims $\lambda v \in \mathbb{N}$ and i.i.d. claim sizes $(Y_i)_{i\geq 1}$ for which the moment generating function M_{Y_1} exists at α for a given $\alpha > 0$. Moreover, let u be the exponential utility function with parameter α, and $c_0 > 0$ a given initial capital. We write $\pi = \pi(u, S, c_0)$ for the utility indifference price for S. Now define

$$\widetilde{S} = \sum_{i=1}^{\lambda v} Y_i,$$

i.e. for $\widetilde{S}$ the number of claims is exactly given by λv. Calculate the utility indifference price $\widetilde{\pi} = \widetilde{\pi}(u, \widetilde{S}, c_0)$ for $\widetilde{S}$ and compare $\widetilde{\pi}$ to π.

Exercise 9.2 Value-at-Risk and Expected Shortfall
Suppose that for the yearly claim amount S of an insurance company in a given line of business we have $S \sim \text{LN}(\mu, \sigma^2)$ with $\mu = 20$ and $\sigma^2 = 0.015$. Moreover, we set the cost-of-capital rate $r_{\text{CoC}} = 6\%$. Then, the premium π_{CoC} for the considered line of business using the cost-of-capital pricing principle with risk measure ρ is given by

$$\pi_{\text{CoC}} = \mathbb{E}[S] + r_{\text{CoC}} \cdot \rho(S - \mathbb{E}[S]).$$

(a) Calculate π_{CoC} using the value-at-risk (VaR) risk measure at security level $1-q = 99.5\%$.

(b) Calculate π_{CoC} using the expected shortfall risk measure at security level $1-q = 99\%$.

(c) Which security level is needed such that π_{CoC} using the VaR risk measure is equal to the price calculated in (b)?

(d) Let U and V be two independent copies of $\log S$. Show that on the one hand

$$\mathrm{VaR}_{1-q}(U+V) \; > \; \mathrm{VaR}_{1-q}(U) + \mathrm{VaR}_{1-q}(V)$$

for all $1-q \in \left(0, \frac{1}{2}\right)$, but on the other hand

$$\mathrm{VaR}_{1-q}(U+V) \; < \; \mathrm{VaR}_{1-q}(U) + \mathrm{VaR}_{1-q}(V)$$

for all $1-q \in \left(\frac{1}{2}, 1\right)$. In particular, the VaR is not subadditive, and hence not coherent.

Exercise 9.3 Variance Loading Principle

We would like to insure the car fleet given in Table 2 under the assumption that the total claim amounts for passenger cars, delivery vans and trucks can be modeled by independent compound Poisson distributions.

i	v_i	λ_i	$\mathbb{E}[Y_1^{(i)}]$	$\mathrm{Vco}(Y_1^{(i)})$
passenger car	40	25%	2'000	2.5
delivery van	30	23%	1'700	2.0
truck	10	19%	4'000	3.0

Table 2: Volumes, claim frequencies, expected claim sizes and coefficients of variation of the claim sizes for the three sections of the car fleet.

(a) Calculate the expected claim amount of the car fleet.

(b) Calculate the premium for the car fleet using the variance loading principle with $\alpha = 3 \cdot 10^{-6}$.

Exercise 9.4 Esscher Premium

Let S be a random variable with distribution function F and moment generating function M_S. Assume that there exists $r_0 > 0$ such that $M_S(r) < \infty$ for all $r \in (-r_0, r_0)$. For $\alpha \in (0, r_0)$, let π_α denote the Esscher premium of S.

(a) Show that if S is non-deterministic, then π_α is strictly increasing in α.

(b) Show that the Esscher premium for small values of α boils down to a variance loading principle.

(c) Suppose that $S \sim \mathrm{CompPoi}(\lambda v, G)$, where $\lambda v > 0$ and G is the distribution function of a gamma distribution with shape parameter $\gamma > 0$ and scale parameter $c > 0$. For which values of α does π_α exist? Calculate π_α where it is defined.

Non-Life Insurance: Mathematics and Statistics

Exercise sheet 10

Exercise 10.1 Method of Bailey & Simon
Suppose that a car insurance portfolio of an insurance company has been divided according to two tariff criteria

- vehicle type: {passenger car, delivery van, truck} = {1,2,3},
- driver age: {21-30 years, 31-40 years, 41-50 years, 51-60 years} = {1,2,3,4}.

For simplicity, we set the number of policies $v_{i,j} = 1$ for all risk classes (i, j), $1 \leq i \leq 3$, $1 \leq j \leq 4$. Moreover, we assume that we work with a multiplicative tariff structure and that we observed the following claim amounts:

	21-30y	31-40y	41-50y	51-60y
passenger car	2'000	1'800	1'500	1'600
delivery van	2'200	1'600	1'400	1'400
truck	2'500	2'000	1'700	1'600

Table 1: Observed claim amounts in the $3 \cdot 4 = 12$ risk classes.

Calculate the tariffs using the method of Bailey & Simon. Comment on the results.

Exercise 10.2 Method of Bailey & Jung
Consider the same setup as in Exercise 10.1. Calculate the tariffs using the method of Bailey & Jung (i.e. the method of total marginal sums). Compare the results to those found in Exercise 10.1, where we applied the method of Bailey & Simon.

Exercise 10.3 Log-Linear Gaussian Regression Model (R Exercise)
Consider the same setup as in Exercise 10.1. This time we calculate the tariffs using the log-linear Gaussian regression model.

(a) Determine the design matrix Z of the log-linear Gaussian regression model.

(b) Calculate the tariffs using the MLE method within the log-linear Gaussian regression model framework.

(c) Compare the results found in part (b) to the results found in Exercises 10.1 and 10.2, where we applied the methods of Bailey & Simon and Bailey & Jung.

(d) Is there statistical evidence that the classification into different types of vehicles could be omitted?

Exercise 10.4 Tweedie's Compound Poisson Model
Let $S \sim \text{CompPoi}(\lambda v, G)$, where $\lambda > 0$ is the unknown claim frequency parameter, $v > 0$ the known volume and G the distribution function of a gamma distribution with known shape parameter $\gamma > 0$ and unknown scale parameter $c > 0$. Then, S has a mixture distribution with a point mass of $\mathbb{P}[S = 0]$ in 0 and a density f_S on $(0, \infty)$.

(a) Calculate $\mathbb{P}[S = 0]$ and the density f_S of S on $(0, \infty)$.

(b) Show that S belongs to the exponential dispersion family with

$$
\begin{aligned}
w &= v,\\
\phi &= \frac{\gamma+1}{\lambda\gamma}\left(\frac{\lambda v\gamma}{c}\right)^{\frac{\gamma}{\gamma+1}},\\
\theta &= -(\gamma+1)\left(\frac{\lambda v\gamma}{c}\right)^{-\frac{1}{\gamma+1}},\\
\boldsymbol{\Theta} &= (-\infty, 0),\\
b(\theta) &= \frac{\gamma+1}{\gamma}\left(\frac{-\theta}{\gamma+1}\right)^{-\gamma},\\
c(0,\phi,w) &= 0 \quad \text{and}\\
c(x,\phi,w) &= \log\left(\sum_{n=1}^{\infty}\left[\frac{(\gamma+1)^{\gamma+1}}{\gamma}\left(\frac{\phi}{w}\right)^{-\gamma-1}\right]^n \frac{1}{\Gamma(n\gamma)n!}\,x^{n\gamma-1}\right), \quad \text{if } x > 0.
\end{aligned}
$$

Non-Life Insurance: Mathematics and Statistics

Exercise sheet 11

Exercise 11.1 Claim Frequency Modeling with GLM (R Exercise)
Suppose that a motorbike insurance portfolio of an insurance company has been divided according to three tariff criteria

- vehicle class: {weight over 60 kg and more than two gears, other},
- vehicle age: {at most one year, more than one year},
- geographic zone: {large cities, middle-sized towns, smaller towns and countryside}.

Assume that we observed the following claim frequencies:

class	age	zone	volume	number of claims	claim frequency
1	1	1	100	25	0.250
1	1	2	200	15	0.075
1	1	3	500	15	0.030
1	2	1	400	60	0.150
1	2	2	900	90	0.100
1	2	3	7'000	210	0.030
2	1	1	200	45	0.225
2	1	2	300	45	0.150
2	1	3	600	30	0.050
2	2	1	800	80	0.100
2	2	2	1'500	120	0.080
2	2	3	5'000	90	0.018

Table 1: Observed volumes, numbers of claims and claim frequencies in the $2 \cdot 2 \cdot 3 = 12$ risk classes.

(a) Perform a GLM analysis for the claim frequencies using the Poisson model. Comment on the results.

(b) Plot the observed and the fitted claim frequencies against the vehicle class, the vehicle age and the geographic zone.

(c) Create a Tukey-Anscombe plot of the deviance residuals versus the fitted expected numbers of claims.

(d) Is there statistical evidence that the classification into the geographic zones could be omitted?

Exercise 11.2 Claim Frequency Modeling with Neural Networks (R Exercise)
In this exercise we consider the French motor third-party liability insurance data set prepared in Listing 1. We model the claim frequencies using the three continuous but categorized tariff criteria

power of the car, age of the car and age of the driver.

(a) Write an R code that performs a GLM analysis for the claim frequencies on the data `trainset` using the Poisson model. Calculate the deviance statistics of the resulting GLM model on both the data sets `trainset` and `testset`.

(b) Write an R code that models the claim frequencies on the data `trainset` using a neural network with two hidden layers with $(r_1, r_2) = (20, 10)$ hidden neurons. Choose the hyperbolic tangent activation function and 100 gradient descent steps. Calculate the deviance statistics of the resulting neural network model on both the data sets `trainset` and `testset`. Compare the results to the values obtained in part (a).

(c) Repeat the neural network fitting procedure of part (b) with 1'000 gradient descent steps instead of 100. Calculate the deviance statistics of the resulting neural network model on both the data sets `trainset` and `testset`. What do you observe?

Listing 1: R code for Exercise 11.2.

```
### Dataset preparation
install.packages("CASdatasets",repos="http://dutangc.free.fr/pub/RRepos/", type="source")
lapply(c("CASdatasets", "keras", "plyr"), require, character.only=TRUE)
data("freMTPL2freq")
data <- freMTPL2freq[,c(1:6)]
data$VehPower <- relevel(as.factor(data$VehPower), ref="6")
VehAgeCat <- cbind(c(0:100), c(1,rep(2,3),rep(3,2),rep(4,2),rep(5,3),rep(6,2),rep(7,2),rep(8,3),
                               rep(9,83)))
data$VehAge <- relevel(as.factor(VehAgeCat[data$VehAge+1,2]), ref="2")
DrivAgeCat <- cbind(c(18:100), c(rep(1,21-18),rep(2,26-21),rep(3,31-26),rep(4,41-31),
                                 rep(5,51-41),rep(6,71-51),rep(7,101-71)))
data$DrivAge <- relevel(as.factor(DrivAgeCat[data$DrivAge-17,2]), ref="6")

### Training set and test set
set.seed(100)
train <- sample(1:nrow(data),round(0.5*nrow(data)))
trainset <- ddply(data[train,], .(VehPower,VehAge,DrivAge), summarise, ClaimNb=sum(ClaimNb),
                  Exposure=sum(Exposure))[,c(4:5,1:3)]
testset <- ddply(data[-train,], .(VehPower,VehAge,DrivAge), summarise, ClaimNb=sum(ClaimNb),
                 Exposure = sum(Exposure))[,c(4:5,1:3)]
```

Exercise 11.3 Claim Severity Modeling with GLM (R Exercise)
In this exercise we consider the French motor third-party liability insurance data set prepared in Listing 2. This time we model the claim severities using the three (categorical) tariff criteria

- area code: {A, B, C, D, E, F},
- brand of the vehicle: {B1, B10, B11, B12, B13, B14, B2, B3, B4, B5, B6},
- diesel/fuel: {diesel, regular fuel}.

(a) Perform a GLM analysis for the claim severities using the gamma model with log-link function. Comment on the results.

(b) Is there statistical evidence that the area code could be omitted as tariff criterion?

Listing 2: R code for Exercise 11.3.

```
# install.packages("CASdatasets",repos="http://dutangc.free.fr/pub/RRepos/", type="source")
lapply(c("CASdatasets", "plyr"), require, character.only=TRUE)
data("freMTPL2freq")
data("freMTPL2sev")
data <- freMTPL2sev[is.element(freMTPL2sev$IDpol,freMTPL2freq$IDpol),]
data <- ddply(data, .(IDpol), summarize, ClaimAmount=sum(ClaimAmount))
data <- cbind(freMTPL2freq[is.element(freMTPL2freq$IDpol,data$IDpol),-3],data[,2])
colnames(data)[12]<- "ClaimAmount"
data <- ddply(data, .(Area, VehBrand, VehGas), summarize, ClaimNb=sum(ClaimNb),
              ClaimAmount=sum(ClaimAmount))
data$ClaimAmount <- data$ClaimAmount/data$ClaimNb
```

Exercise 11.4 Neural Networks and Gradient Descent
We assume that we have M independent large claims $\boldsymbol{Y} = (Y_1, \ldots, Y_M)$ with corresponding covariates $\boldsymbol{z}_1, \ldots, \boldsymbol{z}_M \in \mathcal{Z} \subset \mathbb{R}^{r_0+1}$, where $r_0 = 1$ and $\boldsymbol{z}_m = (1, z_m)$, for all $m = 1, \ldots, M$. Let $\alpha : \mathcal{Z} \to \mathbb{R}_+$ be a given (but unknown) regression function. We assume that Y_m is Pareto distributed with threshold $\theta > 0$ and tail index parameter $\alpha(\boldsymbol{z}_m) > 0$, for all $m = 1, \ldots, M$. The goal is to model the regression function $\alpha : \mathcal{Z} \to \mathbb{R}_+$ with a neural network.

(a) Set up a single hidden layer neural network with $r_1 \in \mathbb{N}$ hidden neurons for this regression problem using the hyperbolic tangent activation function. How many parameters does this model have?

(b) Calculate the deviance statistics for this regression problem.

(c) Assume that we have a large number of hidden neurons. Why are we in this situation in general not interested in finding the maximum likelihood estimator? What alternative solution do you propose?

(d) Calculate one step of the gradient descent optimization algorithm explicitly for the single hidden layer neural network defined in part (a) and the deviance statistics loss function derived in part (b).

Non-Life Insurance: Mathematics and Statistics

Exercise sheet 12

Exercise 12.1 (Inhomogeneous) Credibility Estimators for Claim Counts
Suppose that in Table 1 we are given the current year's claim counts data for 5 different regions, where, for all $i \in \{1, \dots, 5\}$, $v_{i,1}$ denotes the number of policies in region i and $N_{i,1}$ the number of claims in region i. We assume that we are in the Bühlmann-Straub model framework with $I = 5$, $T = 1$ and

$$N_{i,1}|\Theta_i \sim \text{Poi}(\mu(\Theta_i)v_{i,1}),$$

with $\mu(\Theta_i) = \Theta_i \lambda_0$ and $\lambda_0 = 0.088$, for all $i \in \{1, \dots, 5\}$. Moreover, we assume that the pairs $(\Theta_1, N_{1,1}), \dots, (\Theta_5, N_{5,1})$ are independent and $\Theta_1, \dots, \Theta_5$ are i.i.d. with $\mathbb{E}[\Theta_1] = 1$ and $\Theta_1 > 0$ a.s. Finally, we set $\tau^2 = \text{Var}(\mu(\Theta_1)) = 0.00024$.

region i	$v_{i,1}$	$N_{i,1}$
1	50'061	3'880
2	10'135	794
3	121'310	8'941
4	35'045	3'448
5	4'192	314

Table 1: Observed numbers of policies $v_{i,1}$ and numbers of claims $N_{i,1}$ in the 5 regions.

(a) Calculate the inhomogeneous credibility estimator $\widehat{\widehat{\mu(\Theta_i)}}$ for each region $i \in \{1, \dots, 5\}$ and comment on the results. What would we observe if we decreased the volatility τ^2 between the risk classes?

(b) We denote next year's numbers of policies by $v_{1,2}, \dots, v_{5,2}$ and next year's numbers of claims by $N_{1,2}, \dots, N_{5,2}$. Suppose that $N_{i,1}$ and $N_{i,2}$ are independent, conditionally given Θ_i, for all $i \in \{1, \dots, 5\}$, and that the number of policies grows 5% in each region. For all $i \in \{1, \dots, 5\}$, under the assumption $N_{i,2}|\Theta_i \sim \text{Poi}(\mu(\Theta_i)v_{i,2})$, calculate the mean square error of prediction

$$\mathbb{E}\left[\left(\frac{N_{i,2}}{v_{i,2}} - \widehat{\widehat{\mu(\Theta_i)}}\right)^2\right].$$

Exercise 12.2 (Homogeneous) Credibility Estimators for Claim Sizes
Suppose that in Table 2 we are given claim size data for two different years and four different risk classes, where $v_{i,t}$ denotes the number of claims in risk class i and year t and $Y_{i,t}$ the total claim size in risk class i and year t, for all $i \in \{1, 2, 3, 4\}$ and $t \in \{1, 2\}$. We assume that we are in the Bühlmann-Straub model framework with $I = 4$, $T = 2$ and

$$Y_{i,t}|\Theta_i \sim \Gamma(\mu(\Theta_i)cv_{i,t}, c),$$

with $\mu(\Theta_i) = \Theta_i$ and $c > 0$, for all $i \in \{1, 2, 3, 4\}$ and $t \in \{1, 2\}$. Moreover, we assume that $(\Theta_1, Y_{1,1}, Y_{1,2}), \dots, (\Theta_4, Y_{4,1}, Y_{4,2})$ are independent and that $\Theta_1, \dots, \Theta_4$ are i.i.d. with $\mathbb{E}[\Theta_1^2] < \infty$ and $\Theta_1 > 0$ a.s. Finally, we also assume that $Y_{i,1}$ and $Y_{i,2}$ are independent, conditionally given Θ_i, for all $i \in \{1, 2, 3, 4\}$.

risk class i	$v_{i,1}$	$Y_{i,1}$	$v_{i,2}$	$Y_{i,2}$
1	1'058	8'885'738	1'111	13'872'665
2	3'146	7'902'445	3'303	4'397'183
3	238	2'959'517	250	6'007'351
4	434	10'355'286	456	15'629'998

Table 2: Observed numbers of claims $v_{i,1}$ and $v_{i,2}$ and total claim sizes $Y_{i,1}$ and $Y_{i,2}$ in the 4 risk classes.

(a) Calculate the homogeneous credibility estimator $\widehat{\widehat{\mu(\Theta_i)}}^{\text{hom}}$ for each risk class $i \in \{1,2,3,4\}$ and comment on the results.

(b) We denote next year's numbers of claims by $v_{1,3}, \dots, v_{4,3}$ and next year's total claim sizes by $Y_{1,3}, \dots, Y_{4,3}$. Suppose that $Y_{i,1}, Y_{i,2}$ and $Y_{i,3}$ are independent, conditionally given Θ_i, for all $i \in \{1,2,3,4\}$, and that the number of claims grows 5% in each risk cell. For all $i \in \{1,2,3,4\}$, under the assumption $Y_{i,3}|\Theta_i \sim \Gamma(\mu(\Theta_i) c v_{i,3}, c)$, estimate the mean square error of prediction

$$\mathbb{E}\left[\left(\frac{Y_{i,3}}{v_{i,3}} - \widehat{\widehat{\mu(\Theta_i)}}^{\text{hom}}\right)^2\right].$$

Exercise 12.3 Degenerate MLE and the Poisson-Gamma Model

Suppose that in a given line of business we observed the following claim counts data for $T = 5$ years $t = 1, \dots, T$:

t	1	2	3	4	5
N_t	0	0	0	0	0
v_t	10	10	10	10	10

Table 3: Observed claim counts N_t and corresponding volumes v_t for $T = 5$ years $t = 1, \dots, T$.

(a) First, we assume a Poisson model for the claim counts, i.e. $N_1, \dots, N_T$ are independent with $N_t \sim \text{Poi}(\lambda v_t)$, $t = 1, \dots, T$, for an unknown claim frequency parameter $\lambda > 0$. Calculate the MLE $\widehat{\lambda}_T$ of λ. Does this estimate $\widehat{\lambda}_T$ make sense for premium calculation?

(b) Now we assume a Poisson-gamma model for the claim counts, i.e. $\Lambda \sim \Gamma(\gamma, c)$ with $\gamma = 1$ and $c = 50$, and, conditionally given Λ, $N_1, \dots, N_T$ are independent with $N_t \sim \text{Poi}(\Lambda v_t)$, $t = 1, \dots, T$.

  (i) Determine the prior estimator λ_0 and the posterior estimator $\widehat{\lambda}_T^{\text{post}}$, conditionally given data $(N_1, v_1), \dots, (N_T, v_T)$, of the unknown parameter Λ.

  (ii) Find the credibility weight $\alpha_T \in (0,1)$ such that

$$\widehat{\lambda}_T^{\text{post}} = \alpha_T \, \widehat{\lambda}_T + (1 - \alpha_T)\, \lambda_0.$$

  (iii) Suppose we have an additional observation $(N_{T+1}, v_{T+1}) = (1, 10)$ within the Poisson-gamma model framework and that $\widehat{\lambda}_{T+1}^{\text{post}}$ denotes the posterior estimator, conditionally given data $(N_1, v_1), \dots, (N_{T+1}, v_{T+1})$, of the unknown parameter Λ. Find the credibility weight $\beta_{T+1} \in (0,1)$ such that

$$\widehat{\lambda}_{T+1}^{\text{post}} = \beta_{T+1} \frac{N_{T+1}}{v_{t+1}} + (1 - \beta_{T+1})\, \widehat{\lambda}_T^{\text{post}}.$$

(c) Finally, we assume a Poisson-normal model for the claim counts, i.e. $\Lambda \sim \mathcal{N}(\mu, \sigma^2)$ with mean $\mu \in \mathbb{R}$ and variance $\sigma^2 > 0$, and, conditionally given Λ, $N_1, \dots, N_T$ are independent with $N_t \sim \text{Poi}(\Lambda v_t)$, $t = 1, \dots, T$. Is such a model reasonable?

Exercise 12.4 Pareto-Gamma Model
Suppose that $\Lambda \sim \Gamma(\gamma, c)$ with prior shape parameter $\gamma > 0$ and prior scale parameter $c > 0$ and, conditionally given Λ, the components of $\boldsymbol{Y} = (Y_1, \dots, Y_T)$ are independent with $Y_t \sim \text{Pareto}(\theta, \Lambda)$ for some threshold $\theta > 0$, for all $t \in \{1, \dots, T\}$.

(a) Show that the posterior distribution of Λ, conditional on $\boldsymbol{Y}$, is given by

$$\Lambda|\boldsymbol{Y} \sim \Gamma\left(\gamma + T, c + \sum_{t=1}^{T} \log \frac{Y_t}{\theta}\right).$$

(b) For the estimation of the unknown tail index parameter Λ of the Pareto distributions, we define the prior estimator $\lambda_0 = \mathbb{E}[\Lambda]$ and the observation based estimator (MLE of the Pareto tail index parameter)

$$\widehat{\lambda}_T = \frac{T}{\sum_{t=1}^{T} \log \frac{Y_t}{\theta}}.$$

Find the credibility weight $\alpha_T \in (0, 1)$ such that the posterior estimator $\widehat{\lambda}_T^{\text{post}} = \mathbb{E}[\Lambda|\boldsymbol{Y}]$ has the credibility form

$$\widehat{\lambda}_T^{\text{post}} = \alpha_T \, \widehat{\lambda}_T + (1 - \alpha_T) \, \lambda_0.$$

(c) Show that for the (conditional mean square error) uncertainty of the posterior estimator $\widehat{\lambda}_T^{\text{post}}$ we have

$$\mathbb{E}\left[\left(\Lambda - \widehat{\lambda}_T^{\text{post}}\right)^2 \middle| \boldsymbol{Y}\right] = (1 - \alpha_T) \, \frac{1}{c} \, \widehat{\lambda}_T^{\text{post}}.$$

(d) Let $\widehat{\lambda}_{T-1}^{\text{post}}$ denote the posterior estimator in the sub-model where we only have observed $(Y_1, \dots, Y_{T-1})$. Find the credibility weight $\beta_T \in (0, 1)$ such that the posterior estimator $\widehat{\lambda}_T^{\text{post}}$ has the recursive update structure

$$\widehat{\lambda}_T^{\text{post}} = \beta_T \, \frac{1}{\log \frac{Y_T}{\theta}} + (1 - \beta_T) \, \widehat{\lambda}_{T-1}^{\text{post}}.$$

Non-Life Insurance: Mathematics and Statistics

Exercise sheet 13

Exercise 13.1 Chain-Ladder Algorithm

We write $i = 1, \dots, I$ for the accident years denoting the years of claims occurrence. For every accident year we consider development years $j = 0, \dots, J$. For all $i = 1, \dots, I$ and $j = 0, \dots, J$ we write $C_{i,j}$ for the cumulative payments up to development year j for all claims that have occurred in accident year i. For simplicity, we set $I = J + 1 = 10$. Assume that we have observations

$$\mathcal{D}_I \;=\; \{C_{i,j} \,|\, i + j \leq I,\, 1 \leq i \leq I,\, 0 \leq j \leq J\}$$

given by the following upper claims reserving triangle:

accident year i	development year j									
	0	1	2	3	4	5	6	7	8	9
1	5'946'975	9'668'212	10'563'929	10'771'690	10'978'394	11'040'518	11'106'331	11'121'181	11'132'310	11'148'124
2	6'346'756	9'593'162	10'316'383	10'468'180	10'536'004	10'572'608	10'625'360	10'636'546	10'648'192	
3	6'269'090	9'245'313	10'092'366	10'355'134	10'507'837	10'573'282	10'626'827	10'635'751		
4	5'863'015	8'546'239	9'268'771	9'459'424	9'592'399	9'680'740	9'724'068			
5	5'778'885	8'524'114	9'178'009	9'451'404	9'681'692	9'786'916				
6	6'184'793	9'013'132	9'585'897	9'830'796	9'935'753					
7	5'600'184	8'493'391	9'056'505	9'282'022						
8	5'288'066	7'728'169	8'256'211							
9	5'290'793	7'648'729								
10	5'675'568									

Table 1: Upper claims reserving triangle $\mathcal{D}_I$.

This data set can be downloaded from https://people.math.ethz.ch/~wueth/exercises2.html by clicking on "Data to the Examples".

(a) Use the chain-ladder (CL) method to predict the lower triangle

$$\mathcal{D}_I^c \;=\; \{C_{i,j} \,|\, i + j > I,\, 1 \leq i \leq I,\, 0 \leq j \leq J\}.$$

(b) Calculate the CL reserves $\widehat{\mathcal{R}}_i^{\text{CL}}$ for all accident years $i = 1, \dots, I$.

Exercise 13.2 Bornhuetter-Ferguson Algorithm

Consider the same setup as in Exercise 13.1. We assume that we have prior informations $\widehat{\mu}_1, \dots, \widehat{\mu}_I$ for the expected ultimate claims $\mathbb{E}[C_{1,J}], \dots, \mathbb{E}[C_{I,J}]$ given by

accident year i	1	2	3	4	5
prior information $\widehat{\mu}_i$	11'653'101	11'367'306	10'962'965	10'616'762	11'044'881
accident year i	6	7	8	9	10
prior information $\widehat{\mu}_i$	11'480'700	11'413'572	11'126'527	10'986'548	11'618'437

Table 2: Prior informations $\widehat{\mu}_1, \dots, \widehat{\mu}_I$.

(a) Use the Bornhuetter-Ferguson (BF) method to calculate the BF reserves $\widehat{\mathcal{R}}_i^{\text{BF}}$ for all accident years $i = 1, \dots, I$.

(b) Explain why in this example we have $\widehat{\mathcal{R}}_i^{\text{CL}} < \widehat{\mathcal{R}}_i^{\text{BF}}$, for all accident years $i = 2, \dots, I$, where $\widehat{\mathcal{R}}_i^{\text{CL}}$ denotes the CL reserves for accident year i calculated in Exercise 13.1.

Exercise 13.3 Over-Dispersed Poisson Model
Consider the same setup as in Exercise 13.1. This time we apply the over-dispersed Poisson (ODP) model. To this end, for all $i = 1, \dots, I$ and $j = 0, \dots, J$, we write $X_{i,j}$ for all payments done in development year j for claims with accident year i. According to Model Assumptions 9.10 of the lecture notes (version of March 20, 2019), we assume that there exist positive parameters $\mu_1, \dots, \mu_I$, $\gamma_0, \dots, \gamma_J$ and ϕ such that all $X_{i,j}$ are independent (in i and j) with

$$\frac{X_{i,j}}{\phi} \sim \text{Poi}(\mu_i \gamma_j / \phi),$$

for all $i = 1, \dots, I$ and $j = 0, \dots, J$, and side constraint $\sum_{j=0}^{J} \gamma_j = 1$ holds.

(a) Determine the MLEs of $\mu_1, \dots, \mu_I$ and $\gamma_0, \dots, \gamma_J$.

(b) Calculate the ODP reserves $\widehat{\mathcal{R}}_i^{\text{ODP}}$ for all accident years $i = 1, \dots, I$. What do you observe?

(c) Perform a GLM analysis for the payments $X_{i,j}$ using the ODP model in order to check the results obtained in part (b).

Exercise 13.4 Mack's Formula and Merz-Wüthrich (MW) Formula (R Exercise)
Consider the same setup as in Exercise 13.1.

(a) Write an R code using the R package `ChainLadder` in order to determine the following quantities:

- the conditional mean square error of prediction

  $$\text{msep}^{\text{Mack}}_{C_{i,J}|\mathcal{D}_I}\left(\widehat{C}^{\text{CL}}_{i,J}\right),$$

  given in formula (9.21) of the lecture notes, for all accident years $i = 1, \dots, I$;
- the conditional mean square error of prediction for aggregated accident years

  $$\text{msep}^{\text{Mack}}_{\sum_{i=1}^{I} C_{i,J}|\mathcal{D}_I}\left(\sum_{i=1}^{I} \widehat{C}^{\text{CL}}_{i,J}\right),$$

  given in formula (9.22) of the lecture notes;
- the one-year (run-off) uncertainty

  $$\text{msep}^{\text{MW}}_{\text{CDR}_{i,I+1}|\mathcal{D}_I}(0),$$

  given in formula (9.34) of the lecture notes, for all accident years $i = 1, \dots, I$;
- the one-year (run-off) uncertainty for aggregated accident years

  $$\text{msep}^{\text{MW}}_{\sum_{i=1}^{I} \text{CDR}_{i,I+1}|\mathcal{D}_I}(0),$$

  given in formula (9.35) of the lecture notes.

The references for the four formulas above correspond to the version of the lecture notes of March 20, 2019.

(b) Interpret the square-rooted conditional mean square errors of prediction relative to the claims reserves calculated in Exercise 13.1.

(c) Interpret the square-rooted one-year (run-off) uncertainties relative to the square-rooted conditional mean square errors of prediction.