ETH Zürich, D-MATH
HS 2023
Prof. Dr. Mario V. Wüthrich

Coordinator
Selim Gatti

Non-Life Insurance: Mathematics and Statistics

Solution sheet 4

Solution 4.1 Poisson Model and Negative-Binomial Model

(a) In the Poisson model, we assume that $N_1, \dots, N_{10}$ are independent with $N_t \sim \mathrm{Poi}(\lambda v_t)$ for all $t \in \{1, \dots, 10\}$. We use Estimator 2.32 of the lecture notes (version of January 9, 2023) to estimate the claim frequency parameter λ by

$$\widehat{\lambda}_{10}^{\mathrm{MLE}} = \frac{\sum_{t=1}^{10} N_t}{\sum_{t=1}^{10} v_t} = \frac{10\text{'}224}{100\text{'}000} \approx 10.22\%.$$

Let $t \in \{1, \dots, 10\}$. We have

$$\mathbb{E}\left[\frac{N_t}{v_t}\right] = \frac{\mathbb{E}[N_t]}{v_t} = \frac{\lambda v_t}{v_t} = \lambda \qquad \text{and} \qquad \mathrm{Var}\left(\frac{N_t}{v_t}\right) = \frac{\mathrm{Var}(N_t)}{v_t^2} = \frac{\lambda v_t}{v_t^2} = \frac{\lambda}{v_t}.$$

Note that for the random variable $N_t \sim \mathrm{Poi}(\lambda v_t)$ we can write

$$N_t \overset{(\mathrm{d})}{=} \sum_{i=1}^{v_t} \widetilde{N}_i,$$

where $\widetilde{N}_1, \dots, \widetilde{N}_{v_t}$ are i.i.d. random variables following a $\mathrm{Poi}(\lambda)$-distribution. Thus, we can use the Central Limit Theorem to get

$$\frac{N_t/v_t - \mathbb{E}\left[N_t/v_t\right]}{\sqrt{\mathrm{Var}\left(N_t/v_t\right)}} = \frac{N_t/v_t - \lambda}{\sqrt{\lambda/v_t}} \Longrightarrow Z,$$

as $v_t \to \infty$, where Z is a random variable following a standard normal distribution. This leads to the approximation

$$\mathbb{P}\left[\lambda - \sqrt{\lambda/v_t} \le N_t/v_t \le \lambda + \sqrt{\lambda/v_t}\right] = \mathbb{P}\left[-1 \le \frac{N_t/v_t - \lambda}{\sqrt{\lambda/v_t}} \le 1\right] \approx \mathbb{P}(-1 \le Z \le 1) \approx 0.7,$$

i.e. with a probability of roughly 70%, N_t/v_t lies in the interval $[\lambda - \sqrt{\lambda/v_t}, \lambda + \sqrt{\lambda/v_t}]$. Since λ is unknown, we replace it by the estimator $\widehat{\lambda}_{10}^{\mathrm{MLE}}$ to get the approximate prediction interval

$$\left[\widehat{\lambda}_{10}^{\mathrm{MLE}} - \sqrt{\widehat{\lambda}_{10}^{\mathrm{MLE}}/v_t}, \widehat{\lambda}_{10}^{\mathrm{MLE}} + \sqrt{\widehat{\lambda}_{10}^{\mathrm{MLE}}/v_t}\right] \approx [9.90\%, 10.54\%],$$

which should contain roughly 70% of the observed claim frequencies N_t/v_t. We have the following observations of the claim frequencies:

t	1	2	3	4	5	6	7	8	9	10
N_t/v_t	10%	9.97%	9.85%	9.89%	10.56%	10.70%	9.94%	9.86%	10.93%	10.54%

Table 1: Observed claim frequencies N_t/v_t.

We observe that instead of the expected seven observations, only four observations lie in the estimated interval. We conclude that the assumption of having Poisson distributions might not be reasonable.

(b) By equation (2.11) of the lecture notes (version of January 9, 2023), the test statistic

$$\widehat{\chi}^* = \sum_{t=1}^{10} v_t \frac{\left(N_t/v_t - \widehat{\lambda}_{10}^{\mathrm{MLE}}\right)^2}{\widehat{\lambda}_{10}^{\mathrm{MLE}}}$$

is approximately χ^2-distributed with $10-1=9$ degrees of freedom. By inserting the numbers, we get $\widehat{\chi}^* \approx 14.84$. The probability that a random variable with a χ^2-distribution with 9 degrees of freedom is greater than 14.84 is approximately equal to 9.55%. Hence we can reject the null hypothesis of having Poisson distributions only at significance levels that are higher than 9.55%. In particular, we can not reject the null hypothesis at significance level of 5%.

(c) In the negative-binomial model, we assume that $N_1, \ldots, N_{10}$ are independent with, conditionally given Θ_t, $N_t \sim \mathrm{Poi}(\Theta_t \lambda v_t)$ for all $t \in \{1, \ldots, 10\}$, where $\Theta_1, \ldots, \Theta_{10} \overset{\text{i.i.d.}}{\sim} \Gamma(\gamma, \gamma)$ for some $\gamma > 0$. We use Estimator 2.28 of the lecture notes (version of January 9, 2023) to estimate the claim frequency parameter λ by

$$\widehat{\lambda}_{10}^{\mathrm{NB}} = \frac{\sum_{t=1}^{10} N_t}{\sum_{t=1}^{10} v_t} = \frac{10'224}{100'000} \approx 10.22\%.$$

As in equation (2.8) of the lecture notes (version of January 9, 2023), we define

$$\widehat{V}_{10}^2 = \frac{1}{9}\sum_{t=1}^{10} v_t \left(\frac{N_t}{v_t} - \widehat{\lambda}_{10}^{\mathrm{NB}}\right)^2 \approx 0.17 > \widehat{\lambda}_{10}^{\mathrm{NB}}.$$

Let $v = v_1 = \cdots = v_{10} = 10'000$. Now we can use Estimator 2.30 of the lecture notes (version of January 9, 2023) to estimate the dispersion parameter γ by

$$\widehat{\gamma}_{10}^{\mathrm{NB}} = \frac{\left(\widehat{\lambda}_{10}^{\mathrm{NB}}\right)^2}{\widehat{V}_{10}^2 - \widehat{\lambda}_{10}^{\mathrm{NB}}} \frac{1}{9}\left(\sum_{t=1}^{10} v_t - \frac{\sum_{t=1}^{10} v_t^2}{\sum_{t=1}^{10} v_t}\right) = \frac{\left(\widehat{\lambda}_{10}^{\mathrm{NB}}\right)^2}{\widehat{V}_{10}^2 - \widehat{\lambda}_{10}^{\mathrm{NB}}} \frac{\left(10v - \frac{10v^2}{10v}\right)}{9} = \frac{\left(\widehat{\lambda}_{10}^{\mathrm{NB}}\right)^2 v}{\widehat{V}_{10}^2 - \widehat{\lambda}_{10}^{\mathrm{NB}}}$$
$$\approx 1576.15.$$

For all $t \in \{1, \ldots, 10\}$, we have

$$\mathbb{E}\left[\frac{N_t}{v_t}\right] = \frac{\mathbb{E}[N_t]}{v_t} = \frac{\mathbb{E}[\mathbb{E}[N_t|\Theta_t]]}{v_t} = \frac{\mathbb{E}[\Theta_t \lambda v_t]}{v_t} = \frac{\lambda v_t}{v_t} = \lambda,$$

since $\mathbb{E}[\Theta_t] = \gamma/\gamma = 1$, and

$$\mathrm{Var}\left(\frac{N_t}{v_t}\right) = \frac{\mathbb{E}[\mathrm{Var}(N_t|\Theta_t)] + \mathrm{Var}(\mathbb{E}[N_t|\Theta_t])}{v_t^2} = \frac{\mathbb{E}[\Theta_t \lambda v_t] + \mathrm{Var}(\Theta_t \lambda v_t)}{v_t^2} = \frac{\lambda + \lambda^2 v_t/\gamma}{v_t},$$

since $\mathrm{Var}(\Theta_t) = \gamma/\gamma^2 = 1/\gamma$. Similarly as in part (a), we get the prediction interval

$$\left[\widehat{\lambda}_{10}^{\mathrm{NB}} - \sqrt{\frac{\widehat{\lambda}_{10}^{\mathrm{NB}} + \left(\widehat{\lambda}_{10}^{\mathrm{NB}}\right)^2 v_t/\widehat{\gamma}_{10}^{\mathrm{NB}}}{v_t}}, \widehat{\lambda}_{10}^{\mathrm{NB}} + \sqrt{\frac{\widehat{\lambda}_{10}^{\mathrm{NB}} + \left(\widehat{\lambda}_{10}^{\mathrm{NB}}\right)^2 v_t/\widehat{\gamma}_{10}^{\mathrm{NB}}}{v_t}}\right] \approx [9.81\%, 10.63\%],$$

which should contain roughly 70% of the observed claim frequencies N_t/v_t. Looking at the observations given in Table 1 above, we see that eight of them lie in the estimated interval, which is clearly better than in the Poisson case in part (a). In conclusion, here, the negative-binomial model seems more reasonable than the Poisson model.

Solution 4.2 χ^2-Goodness-of-Fit-Analysis

(a) The R code used in part (a) is provided in Listing 1.

- (i) In Figure 1 (left), we can see that the n MLEs of λ approximately have a Gaussian distribution with mean equal to the true value of $\lambda = 10\%$. On the one hand, this is due to the fact that (under regularity assumptions) the MLE is consistent and asymptotically Gaussian distributed (as $T \to \infty$). For more details we refer to Chapter 6 of the textbook "Theory of Point Estimation" by E.L. Lehmann and G. Casella (2nd edition, 1998). On the other hand, in the Poisson case, we directly have an approximate Gaussian distribution of the MLE, independently of the value of T, provided that the volume v is large enough, see also Exercise 4.1.
- (ii) From the QQ plot, see Figure 1 (right), we deduce that the test statistic indeed has approximately a χ^2-distribution with $T - 1 = 9$ degrees of freedom. We only observe slightly heavier tails in the observations, compared to a χ^2-distribution with $T - 1 = 9$ degrees of freedom. By increasing the values for n and v, we get even closer to a χ^2-distribution with $T - 1 = 9$ degrees of freedom.
- (iii) We observe that we wrongly reject the null hypothesis H_0 of having a Poisson distribution as claim count distribution in 5.16% of the cases. This corresponds almost perfectly to the chosen significance level (indicating the probability of rejecting H_0 even though it is true) of 5%.

Listing 1: R code for Exercise 4.2 (a).

```r
### Function generating the data and applying the chi-squared goodness-of-fit test
chi.squared.test.1 <- function(seed1, n, t, lambda, v, alpha){

  ### Generate the claim counts
  set.seed(seed1)
  claim.counts <- array(rpois(n*t,lambda*v), dim=c(t,n))

  ### Distribution of the MLEs of lambda
  lambda_MLE <- colSums(claim.counts)/(t*v)
  plot(density(lambda_MLE), main="Distribution of the MLEs", xlab="Values of the MLEs",
       cex.lab=1.25, cex.main=1.25, cex.axis=1.25)
  abline(v=mean(lambda_MLE), col="red")
  legend("topleft", lty=1, col="red", legend="mean")
  print("1: See plot for the distribution of the MLEs")

  ### Distribution of the test statistic
  lambda_MLE_array <- array(rep(lambda_MLE,each=t), dim=c(t,n))
  test.statistic <- colSums(v*(claim.counts/v-lambda_MLE_array)^2/lambda_MLE_array)
  theoretical.quantiles <- qchisq(p=(1:n)/(n+1), df=t-1)
  empirical.quantiles <- test.statistic[order(test.statistic)]
  lim <- c(min(theoretical.quantiles,empirical.quantiles),
           max(theoretical.quantiles,empirical.quantiles))
  plot(theoretical.quantiles, empirical.quantiles, xlim=lim, ylim=lim,
       xlab="Theoretical Quantiles", ylab="Empirical Quantiles", main="QQ plot", cex.lab=1.25,
       cex.main=1.25, cex.axis=1.25)
  abline(a=0, b=1, col="red")
  print("2: See the QQ plot for a comparison between the empirical quantiles of the test
         statistic and the theoretical quantiles of a chi-squared distribution with t-1
         degrees of freedom")

  ### Result of the hypothesis test
  print(paste("3: How often we wrongly reject the null hypothesis: ",
              sum(test.statistic > qchisq(p=1-alpha, df=t-1))/n,sep=""))
}

### Apply the function with the desired parameters
chi.squared.test.1(seed1=100, n=10000, t=10, lambda=0.1, v=10000, alpha=0.05)
```

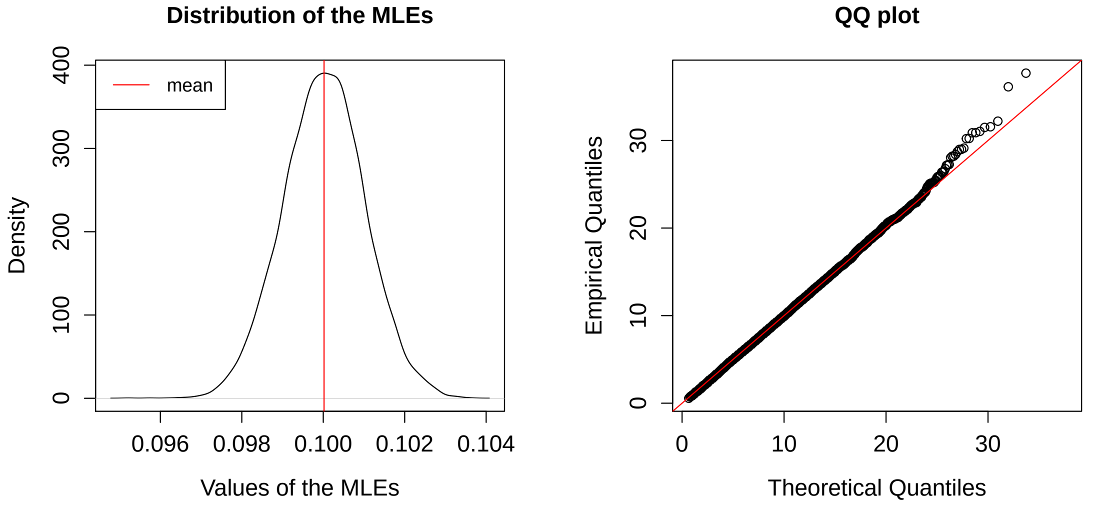

Figure 1: Left: Density plot of the distribution of the MLEs. Right: QQ plot of the theoretical quantiles of a χ^2-distribution with $T - 1 = 9$ degrees of freedom against the empirical quantiles of the values of the test statistic.

Listing 2: R code for Exercise 4.2 (b).

```
### Function generating the data and applying the chi-squared goodness-of-fit test
chi.squared.test.2 <- function(seed1, n, t, lambda, v, alpha, gamma){
  ### Generate the claim counts
  set.seed(seed1)
  claim.counts <- array(rnbinom(n*t, size=gamma, mu=lambda*v), dim=c(t,n))

  ### Calculate the MLEs
  lambda_MLE <- colSums(claim.counts)/(t*v)

  ### Calculate the test statistic
  lambda_MLE_array <- array(rep(lambda_MLE,each=t), dim=c(t,n))
  test.statistic <- colSums(v*(claim.counts/v-lambda_MLE_array)^2/lambda_MLE_array)

  ### Result of the hypothesis test
  print(paste("How often we correctly reject the null hypothesis: ",
              sum(test.statistic > qchisq(p=1-alpha, df=t-1))/n,sep=""))
}

### Apply the function with the desired parameters
chi.squared.test.2(seed1=100, n=10000, t=10, lambda=0.1, v=10000, alpha=0.05, gamma=100)
chi.squared.test.2(seed1=100, n=10000, t=10, lambda=0.1, v=10000, alpha=0.05, gamma=1000)
chi.squared.test.2(seed1=100, n=10000, t=10, lambda=0.1, v=10000, alpha=0.05, gamma=10000)
```

(b) The R code used in part (b) is provided in Listing 2.

(i) We observe the following results:

dispersion parameter γ	100	1'000	10'000
Percentage with which we reject H_0	99.78%	48.38%	7.96%

Table 2: Percentage with which we reject H_0 for different values of γ.

(ii) We see that in the case of a negative binomial distribution with a comparably small parameter ($\gamma = 100$) for the latent gamma distribution, we are almost always able to reject the null hypothesis H_0 of having a Poisson distribution as claim count distribution. The bigger γ, the less we are able to reject H_0. This is because for very large values of γ, the corresponding gamma density does not vary a lot, i.e. the random variable Θ is almost constantly equal to 1. Thus, for increasing γ, we move back to the Poisson model and, consequently, the χ^2-goodness-of-fit test does not detect the latent variable anymore.

Solution 4.3 Claim Count Distribution

The sample mean $\widehat{\mu}$ and the sample variance $\widehat{\sigma}^2$ of the observed numbers of claims $N_1, \ldots, N_{10}$ are given by

$$\widehat{\mu} = \frac{1}{10}\sum_{t=1}^{10} N_t = 21.3 \qquad \text{and} \qquad \widehat{\sigma}^2 = \frac{1}{9}\sum_{t=1}^{10}(N_t - \widehat{\mu})^2 \approx 109.1.$$

We have

$$\widehat{\sigma}^2 \approx 5\widehat{\mu},$$

which suggests $\operatorname{Var}(N_1) > \mathbb{E}[N_1]$. In such a case, we would choose a negative binomial distribution (mixed Poisson distribution), as it allows the variance to exceed the expectation.

Solution 4.4 Method of Moments

If $Y \sim \Gamma(\gamma, c)$, we have

$$\mathbb{E}[Y] = \frac{\gamma}{c} \quad \text{and} \quad \operatorname{Var}(Y) = \frac{\gamma}{c^2}.$$

The sample mean $\widehat{\mu}_8$ and the sample variance $\widehat{\sigma}_8^2$ of the eight observations $y_1, \ldots, y_8$ are given by

$$\widehat{\mu}_8 = \frac{1}{8}\sum_{i=1}^{8} y_i = \frac{64}{8} = 8 \quad \text{and} \quad \widehat{\sigma}_8^2 = \frac{1}{7}\sum_{i=1}^{8}(y_i - \widehat{\mu}_8)^2 = \frac{28}{7} = 4.$$

The method of moments estimates the pair $(\widehat{\gamma}, \widehat{c})$ by solving the equations

$$\widehat{\mu}_8 = \frac{\widehat{\gamma}}{\widehat{c}} \quad \text{and} \quad \widehat{\sigma}_8^2 = \frac{\widehat{\gamma}}{\widehat{c}^2}.$$

We see that $\widehat{\gamma} = \widehat{\mu}_8\widehat{c}$ and, thus,

$$\widehat{\sigma}_8^2 = \frac{\widehat{\mu}_8\widehat{c}}{\widehat{c}^2} = \frac{\widehat{\mu}_8}{\widehat{c}},$$

which is equivalent to

$$\widehat{c} = \frac{\widehat{\mu}_8}{\widehat{\sigma}_8^2} = \frac{8}{4} = 2.$$

Moreover, we get

$$\widehat{\gamma} = \widehat{\mu}_8\widehat{c} = \frac{\widehat{\mu}_8^2}{\widehat{\sigma}_8^2} = \frac{64}{4} = 16.$$

We conclude that the method of moments estimates are given by $(\widehat{\gamma}, \widehat{c}) = (16, 2)$.