# Non-Life Insurance — Assignment 4

AJ*

Autumn 2021

For questions 1 - 4 we will be using the data of De Vijlder (1978):

```
Xij <- scan(text = "      0      0      0     0     0  4627
                          0      0      0     0 15140 13343
                          0      0      0 43465 19018 12476
                          0      0 116531 42390 23505 14371
                          0 346807 118035 43784 12750 12284
                     308580 407117 132247 37086 27744     0
                     358211 426329 157415 68219     0     0
                     327996 436744 147154     0     0     0
                     377369 561699      0     0     0     0
                     333827      0      0     0     0     0")
```

## Question 1

```
## filling in the dots
i <- rep(1:10, each = 6)
j <- rep(1:6, times = 10)
k <- i + j - 1
future <- k > 10
valid <- as.integer(Xij!= 0)

## check
> xtabs(Xij~i+j)
    j
i         1      2      3      4      5      6
  1       0      0      0      0      0   4627
  2       0      0      0      0  15140  13343
  3       0      0      0  43465  19018  12476
  4       0      0 116531  42390  23505  14371
  5       0 346807 118035  43784  12750  12284
  6  308580 407117 132247  37086  27744      0
  7  358211 426329 157415  68219      0      0
  8  327996 436744 147154      0      0      0
  9  377369 561699      0      0      0      0
  10 333827      0      0      0      0      0
```

*Student number: ∞

## Question 2

```
## (a)
alpha <- alpha*sum(beta); beta <- beta/sum(beta)

## (b)
> round(alpha); round(beta,3)

    fi1     fi2     fi3     fi4     fi5     fi6     fi7     fi8     fi9    fi10
 270638  664133  790749  796639  798643  939137 1032577 1009003 1249258 1033618
  fj1   fj2   fj3   fj4   fj5   fj6
0.323 0.434 0.147 0.054 0.025 0.017
```

Comparing the output to De Vijlder (1978), **Table 3**, p. 253., we have exactly the same beta values. The alpha values are exact with the caveat that we have used slightly different units ($\times 1000$).

```
## (c)
> xtabs(round(fitted(gg))*future~i+j)[6:10,2:6]
    j
i         2      3      4      5      6
  6       0      0      0      0  16056
  7       0      0      0  25666  17654
  8       0      0  54669  25080  17251
  9       0 183413  67686  31052  21358
  10 448672 151753  56003  25692  17671
```

Now, comparing the output to De Vijlder (1978), **Table 2**, p. 253., we have been able to reproduce the numbers exactly but we have zeros above the diagonal whereas his table is blank.

## Question 3

```
## add inflation term k
ggg <- glm(Xij~fi+fj+k, gaussian(link=log), weights=valid, mustart=Xij+0.1)

## model parameters
cq3 <- exp(coef(ggg))
alpha.prime <- cq3[1] * c(1,cq3[2:10]);
beta.prime <- c(1,cq3[11:15]);

## comparison
> all.equal(alpha, alpha.prime, check.attributes = FALSE)
[1] TRUE
> all.equal(beta, beta.prime, check.attributes = FALSE)
[1] TRUE

## visually
> exp(coef(gg))
 (Intercept)          fi2          fi3          fi4          fi5          fi6
8.740792e+04 2.453952e+00 2.921791e+00 2.943555e+00 2.950961e+00 3.470083e+00
         fi7          fi8          fi9         fi10          fj2          fj3
3.815340e+00 3.728235e+00 4.615971e+00 3.819185e+00 1.344025e+00 4.545855e-01
         fj4          fj5          fj6
1.677592e-01 7.696118e-02 5.293571e-02

> exp(coef(ggg))
 (Intercept)          fi2          fi3          fi4          fi5          fi6
8.740792e+04 2.453952e+00 2.921791e+00 2.943555e+00 2.950961e+00 3.470083e+00
         fi7          fi8          fi9         fi10          fj2          fj3
3.815340e+00 3.728235e+00 4.615971e+00 3.819185e+00 1.344025e+00 4.545855e-01
         fj4          fj5          fj6            k
1.677592e-01 7.696118e-02 5.293571e-02           NA

## deviance
> gg$deviance; ggg$deviance
[1] 2685831336
[1] 2685831336
```

The models are identical but ggg has `k = NA` *i.e.* `R` has detected linear dependence and dropped the parameter in the fitting process.

## Question 4

**(a)** Consider the expression:

$$\sum_{i,j} w_{ij}\,(x_{ij} - \alpha_i\,\beta_j)^2\,.$$

Fix $i$, then:

$$\begin{aligned}
\frac{\partial}{\partial\alpha_i}\{\sum_{i,j} w_{ij}\,(x_{ij} - \alpha_i\,\beta_j)^2\} &= \frac{\partial}{\partial\alpha_i}\{\sum_{j} w_{ij}\,(x_{ij} - \alpha_i\,\beta_j)^2\} \\
&= \sum_j \frac{\partial}{\partial\alpha_i}\{w_{ij}\,(x_{ij} - \alpha_i\,\beta_j)^2\} \\
&= \sum_j 2\,w_{ij}\,(x_{ij} - \alpha_i\,\beta_j)\cdot -\beta_j \\
&= -2\sum_j w_{ij}\,\beta_j\;(x_{ij} - \alpha_i\,\beta_j)\,.
\end{aligned}$$

Similarly, fix $j$, then:

$$\begin{aligned}
\frac{\partial}{\partial\beta_j}\{\sum_{i,j} w_{ij}\,(x_{ij} - \alpha_i\,\beta_j)^2\} &= \frac{\partial}{\partial\beta_j}\{\sum_{i} w_{ij}\,(x_{ij} - \alpha_i\,\beta_j)^2\} \\
&= \sum_i \frac{\partial}{\partial\beta_j}\{w_{ij}\,(x_{ij} - \alpha_i\,\beta_j)^2\} \\
&= \sum_i 2\,w_{ij}\,(x_{ij} - \alpha_i\,\beta_j)\cdot -\alpha_i \\
&= -2\sum_i w_{ij}\,\alpha_i\;(x_{ij} - \alpha_i\,\beta_j)\,.
\end{aligned}$$

**(b)** Now, setting both equations equal to $0$:

$$\begin{aligned}
\sum_i w_{ij}\,\alpha_i\;(x_{ij} - \alpha_i\,\beta_j) = 0 \\
\sum_j w_{ij}\,\beta_j\;(x_{ij} - \alpha_i\,\beta_j) = 0
\end{aligned}$$

Consider solving for $\alpha_i$.

Expanding the summation, we have:

$$\begin{aligned}
&\sum_j w_{ij}\,\beta_j\,x_{ij} - \sum_j w_{ij}\,\beta_j^2\,\alpha_i = 0 \\
&\implies \alpha_i\cdot\sum_j w_{ij}\,\beta_j^2 = \sum_j w_{ij}\,\beta_j\,x_{ij} \\
&\implies \alpha_i = \frac{\sum_j w_{ij}\,\beta_j\,x_{ij}}{\sum_j w_{ij}\,\beta_j^2}.
\end{aligned}$$

The same method yields the required equation for $\beta_j$.

```
## Question 4 continued

## (c)
beta <- rep(1,6)
repeat {
    beta.old <- beta
    alpha <- tapply(valid*Xij*beta[j],i,sum)/tapply(valid*(beta[j]^2),i,sum)
    beta <- tapply(valid*Xij*alpha[i],j,sum)/tapply(valid*(alpha[i]^2),j,sum)
    if (sum(abs((beta.old-beta)/beta)) < 1e-7) break  ## finish loop
    cat(beta,"\n")  ## monitor the iteration process
}

## check
> round(xtabs(alpha[i]*beta[j]*future~i+j)[6:10,2:6])
    j
i         2      3      4      5      6
  6       0      0      0      0  16056
  7       0      0      0  25666  17654
  8       0      0  54669  25080  17251
  9       0 183413  67686  31052  21358
  10 448672 151753  56003  25692  17671
```

## Question 5

**(a)** Consider $\delta > 0\,$,

$$\beta(x) = e^{-\gamma(x-1)} \cdot x^{\delta}.$$

Therefore,

$$\begin{aligned}
\frac{\mathrm{d}\beta}{\mathrm{d}x} &= e^{-\gamma(x-1)} \cdot \; -\gamma\, x^{\delta} + e^{-\gamma(x-1)} \cdot \delta\, x^{\delta-1} \\
&= e^{-\gamma} \left( e^{-\gamma x} \cdot - \gamma\, x^{\delta} + e^{-\gamma x} \cdot \delta\, x^{\delta-1} \right) \\
&\overset{!}{=} 0.
\end{aligned}$$

Hence,

$$\begin{aligned}
&-e^{-\gamma x} \cdot \; \gamma\, x^{\delta} + e^{-\gamma x} \cdot \delta\, x^{\delta-1} = 0 \\
&\implies e^{-\gamma x} \cdot \delta\, x^{\delta-1} = e^{-\gamma x} \cdot \; \gamma\, x^{\delta} \\
&\implies x = \frac{\delta}{\gamma}\,, \quad \gamma > 0.
\end{aligned}$$

**(b)** Now,

$$\begin{aligned}
\beta(x+1) &= e^{-\gamma((x+1)-1)} \cdot (x+1)^{\delta} \\
&= e^{-\gamma x} \cdot (x+1)^{\delta}.
\end{aligned}$$

Therefore,

$$\begin{aligned}
\frac{\beta(x+1)}{\beta(x)} &= \frac{e^{-\gamma x} \cdot (x+1)^{\delta}}{e^{-\gamma(x-1)} \cdot x^{\delta}} \\
&= \frac{e^{-\gamma x} \cdot (x+1)^{\delta}}{e^{-\gamma x} \cdot e^{\gamma} \cdot x^{\delta}} \\
&= \frac{1}{e^{\gamma}} \cdot \frac{(x+1)^{\delta}}{x^{\delta}} \\
&= \frac{1}{e^{\gamma}} \cdot \left( 1 + \frac{1}{x} \right)^{\delta}.
\end{aligned}$$

Hence,

$$\begin{aligned}
\lim_{x\to\infty} \frac{\beta(x+1)}{\beta(x)} &= \lim_{x\to\infty} \frac{1}{e^{\gamma}} \cdot \left( 1 + \frac{1}{x} \right)^{\delta} \\
&= \frac{1}{e^{\gamma}} \cdot \lim_{x\to\infty} \left( 1 + \frac{1}{x} \right)^{\delta} \\
&= \frac{1^{\delta+1}}{e^{\gamma}}.
\end{aligned}$$

For $d > 0$:

$$\begin{aligned}
&d = \frac{1^{\delta+1}}{e^{\gamma}} \\
&\implies e^{\gamma} = \frac{1^{\delta+1}}{d} \\
&\implies \gamma = \log\left( \frac{1^{\delta+1}}{d} \right) \\
&\implies \gamma = -\log d\,, \qquad 0 < d < 1
\end{aligned}$$

$$\delta \in \mathbb{R}.$$

### Question 6

```
TT <- 10; x.top <- 2; d <- 1/2
gamma <- -log(d); delta <- x.top*gamma
j <- 1:TT
beta <- exp(-gamma*(j-1) + delta*log(j))
```

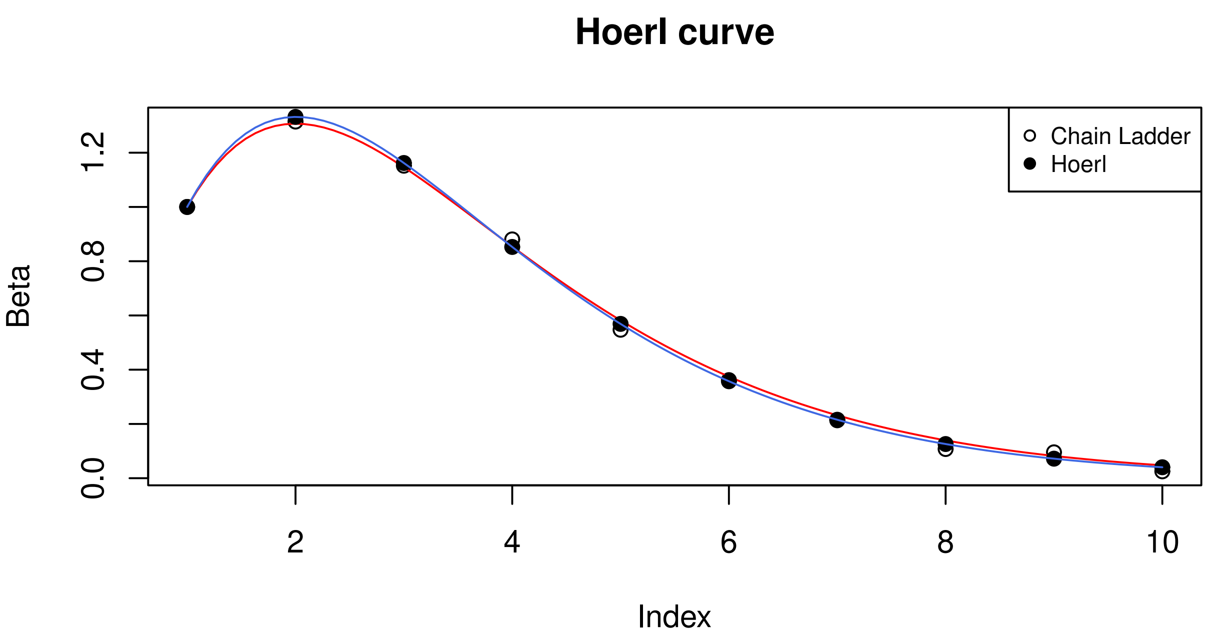
### Question 7

```
## Chain Ladder (CL model)
CL <- glm(Xij~0+fi+fj, poisson)
exp(coef(CL))
beta.CL <- c(1, exp(coef(CL))[(TT+1):length(coef(CL))])
```

### Question 8

```
## (a) compute beta.Hoerl
k <- 1:TT
beta.Hoerl <- exp(-gamma.Hoerl*(k-1) + delta.Hoerl*log(k))

## (b) plot
plot(beta.CL,  ylab = "Beta", main = "Hoerl curve")
curve(exp(-gamma*(x-1) + delta*log(x)), col="red", add=TRUE)
points(beta.Hoerl, pch = 19)
curve(exp(-gamma.Hoerl*(x-1) + delta.Hoerl*log(x)), col="royalblue", add=TRUE)
legend("topright", c("Chain Ladder", "Hoerl"), pch = c(1,19), cex = .75)
```



## Question 9

```
> anova(Hoerl,CL)
Analysis of Deviance Table

Model 1: Xij ~ 0 + fi + I(j - 1) + log(j)
Model 2: Xij ~ 0 + fi + fj
  Resid. Df Resid. Dev Df Deviance
1        43     44.106
2        36     27.292  7   16.814
```

The decrease in deviance is 16.814. For a Poisson model, the deviance is the same as the scaled deviance, so we compare this value with the critical value of $\chi^2_7 = 14.06714$ (at the 95$^{\text{th}}$ percentile).

Comparing these values, we note that the change in deviance is significant at this confidence level *i.e.* the Chain Ladder model in preferred.

## Question 10

```
set.seed(1)

for(a in 1:10) {
    Xij <- rpois(length(mu.ij), mu.ij)
    CL <- glm(Xij~0+fi+fj, poisson)
    Hoerl <- glm(Xij~0+fi+I(j-1)+log(j), poisson)
    print(Hoerl$deviance - CL$deviance)
}

## output
[1] 5.009555
[1] 5.812632
[1] 7.945802
[1] 11.91367
[1] 5.824499
[1] 8.578764
[1] 8.240614
[1] 12.48348
[1] 6.680106
[1] 3.35949
```

None of the models are now significant at the 95$^{\text{th}}$ percentile *i.e.* the Chain Ladder model is not preferred for any of the data-sets we have generated.

### Question 11

```
## (a)
Xij <- c(232,106,35,16,2,258,115,56,27,221,82,4,359,71,349)
i <- c(1,1,1,1,1,2,2,2,2,3,3,3,4,4,5)
j <- c(1,2,3,4,5,1,2,3,4,1,2,3,1,2,1)
TT <- 5

## generate the run-off triangle
fi <- as.factor(i); fj <- as.factor(j)
> xtabs(Xij~i+j)
   j
i     1   2   3   4   5
  1 232 106  35  16   2
  2 258 115  56  27   0
  3 221  82   4   0   0
  4 359  71   0   0   0
  5 349   0   0   0   0

## Model 1 - Hoerl
Hoerl <- glm(Xij~0+fi+I(j-1)+log(j), poisson)

## Model 2 - Chain Ladder
CL <- glm(Xij~0+fi+fj, poisson)

## model comparison
> anova(Hoerl,CL)
Analysis of Deviance Table

Model 1: Xij ~ 0 + fi + I(j - 1) + log(j)
Model 2: Xij ~ 0 + fi + fj
  Resid. Df Resid. Dev Df Deviance
1         8     82.633
2         6     74.172  2   8.4608

> qchisq(.95,2)
[1] 5.991465
```

The decrease in deviance is 8.4608, and we again compare this with the critical value of $\chi^2_2 = 5.991465$ (at the $95^{\text{th}}$ percentile).

Comparing these values, we note that the change in deviance is significant at this confidence level *i.e.* we reject the null hypothesis that the coarser model (`Hoerl`) is the better fit, hence the Chain Ladder model in preferred.

```
## 11(b)
## model 11b - alpha.Hoerl
alpha.Hoerl <- glm(Xij~0+fj+I(i-1)+log(i), poisson)

## model comparison
> anova(alpha.Hoerl,CL)
Analysis of Deviance Table

Model 1: Xij ~ 0 + fj + I(i - 1) + log(i)
Model 2: Xij ~ 0 + fi + fj
  Resid. Df Resid. Dev Df Deviance
1         8    108.442
2         6     74.172  2    34.27
```

Using the same reasoning as in part (a), we note that the change in deviance is significant at this confidence level *i.e.* we reject the null hypothesis and the Chain Ladder model is preferred.

## Question 12

```
## (a)
CL <- glm(Xij~fi+fj, poisson)
coefs.CL <- exp(coef(CL))
alpha.CL <- coefs.CL[1]*c(1, coefs.CL[2:TT])
beta.CL <- c(1, coefs.CL[(TT+1):length(coefs.CL)])

## (b)
> round(alpha.CL %o% beta.CL)
        fj2 fj3 fj4 fj5 fj6 fj7 fj8
    149  40  10   5   3   2   0   0
fi2 155  42  10   5   3   2   1   0
fi3 188  51  13   7   4   2   1   0
fi4 201  54  14   7   4   2   1   0
fi5 196  53  13   7   4   2   1   0
fi6 272  73  18   9   5   3   1   0
fi7 233  63  16   8   5   3   1   0
fi8 221  60  15   8   4   3   1   0
```

The last column contains only zeros because the fitted values in column 8 are of the form $\alpha_i \beta_8$ where the fitted value of $\beta_8$ is `beta.CL[8]` = 5.573023e-10, which is essentially zero.

### Question 13

```
## (a) arithmetic separation
AS <- glm(Xij~fj+fk, poisson)

## (b) intercept term vs. observation
> exp(coef(AS)[1])
(Intercept)
        156
> Xij[1]
[1] 156
```

The fitted value of this cell is given by $\beta_1 \cdot \gamma_1$, where $\gamma_1 = 1$. Hence, the value in this cell is determined by the intercept term $\beta_1$, which in this model has no element of time operating on it. Therefore, it is fitted as the observed value.

### Question 14

```
## (a)
l <- 1:15
gamma.fitted <- exp(ab[1]+l*ab[2])

## (b)
gammas <- c(gamma.AS, gamma.fitted[9:15])
```

### Question 15

```
## (a)
Threeway <- glm(Xij~fi+fj+fk, poisson)

## (b)
> anova(CL,Threeway)
Analysis of Deviance Table

Model 1: Xij ~ fi + fj
Model 2: Xij ~ fi + fj + fk
  Resid. Df Resid. Dev Df Deviance
1        21     34.158
2        15     21.252  6   12.907

## (c)
> anova(AS,Threeway)
Analysis of Deviance Table

Model 1: Xij ~ fj + fk
Model 2: Xij ~ fi + fj + fk
  Resid. Df Resid. Dev Df Deviance
1        21     40.210
2        15     21.252  6   18.958

## (d)
> AIC(CL, AS, Threeway)
         df      AIC
CL       15 227.4894
AS       15 233.5406
Threeway 21 226.5827
```

Analysing our output and noting that the critical value of $\chi^2_6 = 12.59159$ (at the $95^{\text{th}}$ percentile), we see that the change in deviance is significant in both instances, so we prefer the *three-way* model over both the *arithmetic separation* and *chain ladder* model.

Furthermore, the *three-way* model has performed best in terms of AIC.