Master in Actuarial Science, Exam 1. **2h30m**
Ratemaking and Experience Rating, 2nd year, 1st semester

1. Consider a group of 291 insured households with residence in some middle class (urban) neighbourhood. The area has became recently a target of burglary and theft. In the past year there were 179 claim cases of robbery distributed along with the following table:

No. of Claims	0	1	2	3	4
No. of households	158	96	29	7	1

For a given household (risk unit) in that population group, let X be the number of claims due to robbery per year, let $X \frown$Poisson(θ) and consider that the distribution mean may vary among the different group units. Consider that the usual hypothesis in credibility theory are applicable to the risk group under study, θ is the associated risk parameter and you are a promising actuary, expert on experience rating.

Bühlmann's credibility (pure) premium for the given risk, for the next year, is given by formula

$$P_c = z\bar{X} + (1-z)\mu,$$

where $z = n/(n + \upsilon/a)$, $\mu = E[\mu(\theta)]$, $\upsilon = E[\upsilon(\theta)]$, $a = Var[\mu(\theta)]$, $\mu(\theta)$ and $\upsilon(\theta)$ are the risk mean and variance, respectively, n is the number of years in force of that risk, and $\bar{X}$ is its sample mean.

(a) i. Compute $E[P_c]$. Comment briefly on the result.
    ii. Show that $Var[P_c] = za$.

(b) Explain briefly the behaviour of the credibility factor z (and then, of P_c) as function of n, υ and a.

(c) i. Calculate estimates for the structural parameters μ, υ and a.
    ii. Compute the (*empirical)* Bühlmann's estimate for the given risk for the next period, if she has produced three claims.
    [**Remark**: Take care of particularities of the Poisson distribution]

(d) Do you find appropriate that the number of claims is Poisson distributed?

(e) Assume from now on that the prior distribution of the risk parameter Θ is known to be Exponential with mean β and that the given risk has produced $X = x$ claims in the last year.
    i. Show that the posterior distribution is a Gamma distribution with parameters $\alpha_* = x + 1$ and $\beta_* = \left(1 + \beta^{-1}\right)^{-1}$.
    ii. Can we talk here on *Exact Credibility*? Explain.
    iii. Set $x = 3$ and $\beta = 3/5$.Compute the Bayesian premium $E(\Theta|X = 3)$. Does it should coincide with the estimate asked in item (1c)?

2. Consider a *bonus* system based on the claims frequency to rate the risk of some given motor insurance portfolio. The system has simply three classes, numbered 1, 2, and 3 and ranked increasingly from low to higher premium. Transition rules are the following: A policy with no claims in one year goes to the previous lower class in the next year unless it is already Class 1, where it stays. In the case of a claim, goes to Class 3, if it is already there no change is made. Entry class is Class 2 and premia vector (indeces) is given by $\mathbf{b} = (75, 100, 150)$.

Let $\exp(-\lambda)$ be the probability of a certain policy in the portfolio no claims in the year. Suppose that the number of claims per year of that policy can be explained by a Poisson distribution with mean λ. The portfolio of risks is considered to be (basically) homogeneous.

For a certain risk in the portfolio consider the following:

(a) Write the transitions rules and the one-step transition probability matrices, T and P respectively.

(b) Discuss briefly the existence of a limiting distribution. In the case of its existence compute that distribution and the corresponding average premium.

(c) Calculate the probability of a policy being ranked in Class 2 after two years of entering the system.

(d) Lemaire's (1985) *Transient Elasticity* evaluation measure is based on the following equation

$$V_\lambda(j) = b(j) + \beta_j \sum_{k=1}^{s} p_{T,\lambda}(j,k)V_\lambda(k), \; j = 1, ..., s \; ,$$

where$V_\lambda(j)$ is the expected present value to be paid by a policy from class C_j, $b(j)$ is the premium for Class j, β_j (< 1) is a discount rate, $p_{T,\lambda}(j,k)$ is a one-step transition probability and s is the number of classes.

Explain fully the formula and its components.

Consider now that λ is a realization of a random variable Λ following an exponential distribution with mean one.

(e) Compute the resulting stationary distribution of the premia. Compute the average premium.

(f) Suppose now that the insurer decided to make a slight change in the *bonus* system, as follows: For a policy to move to a lower a class (apart form those already in Class 1) it is now necessary to have two years (in a row) without any claim. All other rules remain unchanged.

Make appropriate changes, if any, in the class set so that the system can be workable with the usual Markovian approach.

3. For tariff modelling purposes regarding some insurance large portfolio, we studied some factors suposedly with impact in the claims frequency mean. We have first selected 5 factors labeled as F1 to F5. All these factors are qualitative, rank variables, ranked from 1 on, where Class 1 is the lower one. There are available information about Claims Counts (CC), Number of Policies (NP) for each cell. LNP is the logarithm of NP. We considered a Poisson model with logarithm link. For decisions consider a significance level of 5%. The Annex shows an estimated model, Model 1.

   (a) After a brief analysis on the estimated model, argue about the quality of the model? Can it be improved? If so, clarify.

   (b) Conclude about the importance of level 3 of Factor 4 (F43).

   (c) Comment briefly the statement: *Level 1 of all factors are missing from the estimated table in the Annex because they were removed for having no significance.*

   (d) How would you calculate the expected frequency for the standard risk?

   (e) Would you accept this example to come from some sort of vehicle insurance? Explain briefly.

Marks (out of 200):

1.*a*)*i*.	*ii*.	*b*)	*c*)*i*.	*ii*.	*d*)	*e*)*i*.	*ii*.	*iii*.	2.*a*)	*b*)	*c*)	*d*)	*e*)	*f*)	3.*a*)	*b*)	*c*)	*d*)	*e*)
10	10	15	15	15	10	20	10	15	10	10	5	7.5	20	7.5	5	5	2.5	5	2.5
(10	20	35	50	65	75	95	105	120	130	140	145	152.5	172.5	180	185	190	192.5	197.5	200)

ANNEX

Model 1

```
Call:
glm(formula = CC ~ F1 + F2 + F3 + F4 + F5, family = poisson,
    data = freq2, offset = LNP)
Deviance Residuals:
    Min       1Q   Median       3Q      Max
-5.0031  -0.9156  -0.1510   0.6291   4.0551

Coefficients:
            Estimate Std. Error z value Pr(>|z|)
(Intercept) -1.80413    0.03545 -50.891  < 2e-16 ***
F12         -0.15887    0.02861  -5.554 2.80e-08 ***
F13         -0.12305    0.03140  -3.919 8.88e-05 ***
F14         -0.30068    0.03632  -8.278  < 2e-16 ***
F15         -0.32130    0.04339  -7.406 1.31e-13 ***
F22          0.03211    0.02116   1.517 0.129142
F23          0.12177    0.02596   4.691 2.71e-06 ***
F32          0.09465    0.02807   3.372 0.000746 ***
F42          0.09466    0.03191   2.966 0.003018 **
F43          0.06811    0.03604   1.890 0.058762 .
F44          0.16338    0.03907   4.182 2.89e-05 ***
F45          0.15807    0.04347   3.636 0.000277 ***
F52         -0.02656    0.01992  -1.334 0.182367
F53         -0.02626    0.02365  -1.110 0.266855
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for poisson family taken to be 1)

    Null deviance: 1622.0  on 1077  degrees of freedom
Residual deviance: 1468.2  on 1064  degrees of freedom
AIC: 5079.3

Number of Fisher Scoring iterations: 4
```

Solutions

1.

(a) i. Compute $E[P_c]$. Comment briefly on the result. **R**:

$$\begin{aligned} E[P_c] &= zE[E[\bar{X}|\Theta]] + (1-z)\mu \\ &= zE[\mu(\Theta)] + (1-z)\mu = \mu. \end{aligned}$$

P_c is an unbiased estimator, on average P_c coincides with the collective (pure) premium..

ii. Show that $Var[P_c] = za$. **R**:

$$\begin{aligned} Var[P_c] &= z^2 Var[\bar{X}] = z^2 \left(Var[\mu(\Theta)] + E[\upsilon(\theta)]/n \right) \\ &= z^2 \left(a + \upsilon/n \right) = za. \end{aligned}$$

(b) Explain briefly the behaviour of the credibility factor z (and then, of P_c) as function of n, υ and a.

z increases with n, approaches 1 when n increases, is 0 when $n = 0$, increases when the ratio υ/a decreases, decreases otherwise. The ratio increases with υ, decreases with a, this means that credibility for the risk is lower with higher variability within the risk and crebibility is higher if there is higher variability among the different risks in the portfolio, respectively. On average P_c equals the collective premium.

(c) i. Calculate estimates for the structural parameters μ, υ and a. **R**:
$\mu = \upsilon$. $\hat{\mu} = \bar{x} = 0.61512$; $\hat{V}[X] = E(\Theta) + V(\Theta) = \hat{\mu} + \hat{a} = s^2 = 0.6216270$; $\hat{a} = 0.61512 - 0.6216270 = 0.00650677$.

ii. Compute the (*empirical)* Bühlmann's estimate for the given risk for the next period, if she has produced three claims.
[**Remark**: Take care of particularities of the Poisson distribution]
R: $= n/(n + \hat{\mu}/\hat{a}) = 1/\left(1 + 0.61512/0.00650677\right) = 1.0467 \times 10^{-2} =$; $\hat{P}_c = 1.0467 \times 10^{-2} \times 3 + (1 - 1.0467 \times 10^{-2})0.61512 = 0.64008$.

(d) Do you find appropriate that the number of claims is Poisson distributed?

R: We can start to compute both the sample mean and variance, we get values not too different, I may say to be consistent to the Poisson distribution. Then, a simple χ^2 goodness-of-fit test leads to *non rejection.*

(e) Assume from now on that the prior distribution of the risk parameter Θ is known to be Exponential with mean β and that the given risk has produced $X = x$ claims in the last year.

i. Show that the posterior distribution is a Gamma distribution with parameters $\alpha_* = x + 1$ and $\beta_* = \left(1 + \beta^{-1}\right)^{-1}$.
Compute, with $f_X(x|\theta) = e^{-\theta}\theta^x/x!$,

$$\pi\left(\theta | X = 2\right) = \frac{f_X(x|\theta)\pi(\theta)}{\int_0^\infty f_X(x|\theta)\pi(\theta)d\theta} \rightarrow Gamma(\alpha_*, \beta_*)$$

ii. Can we talk here on *Exact Credibility*? Explain.
R: Yes we can. This is a Poisson-Gamma, conjugate distribution families, and the posterior is of the same family as the prior (exponential).

iii. Set $x = 3$ and $\beta = 3/5$. Compute the Bayesian premium $E(\Theta|X = 3)$. Does it should coincide with the estimate asked in item (1c)?
R: $E(\Theta|X = 3) = \alpha_*, \beta_* = (3+1)\left(1 + 5/3\right)^{-1} = 1.5$. Exact credibility does not work with *empirical* credibility, we mean when you don't know the structural parameters and you have to estimate them. In this case, we have different environments as a mixed Poisson with a Gamma gives a negative binomial distribution and the variance is different from the mean, you would hardly have the same parameter values to fit the estimated Buhlmann's.

2. Consider a *bonus* system based on the claims frequency to rate the risk of some given motor insurance portfolio. The system has simply three classes, numbered 1, 2, and 3 and ranked increasingly from low to higher premium. Transition rules are the following: A policy with no claims in one year goes to the previous lower class in the next year unless it is already Class 1, where it stays. In the case of a claim, goes to Class

3, if it is already there no change is made. Entry class is Class 2 and premia vector (indeces) is given by $\mathbf{b} = (75, 100, 150)$.

Let $\exp(-\lambda)$ be the probability of a certain policy in the portfolio no claims in the year. Suppose that the number of claims per year of that policy can be explained by a Poisson distribution with mean λ. The portfolio of risks is considered to be (basically) homogeneous.

For a certain risk in the portfolio consider the following:

(a) Write the transitions rules and the one-step transition probability matrices, T and P respectively.

R: Set $\exp(-\lambda) = \theta$.

$$T = \begin{bmatrix} \{0\} & & \{1+\} \\ \{0\} & & \{1+\} \\ & \{0\} & \{1+\} \end{bmatrix}; \; P = \begin{bmatrix} \theta & 0 & 1-\theta \\ \theta & 0 & 1-\theta \\ 0 & \theta & 1-\theta \end{bmatrix}$$

(b) Discuss briefly the existence of a limiting distribution. In the case of its existence compute that distribution and the corresponding average premium.

R: After two periods we have

$$\begin{aligned} P^2 &= \begin{bmatrix} \theta^2 & \theta(1-\theta) & 1-\theta \\ \theta^2 & \theta(1-\theta) & 1-\theta \\ \theta^2 & \theta(1-\theta) & 1-\theta \end{bmatrix} \\ AP &= 75\theta^2 + 100\theta(1-\theta) + 150(1-\theta) = -25\theta^2 - 50\theta + 150 \end{aligned}$$

Alternatively, we could equate $\pi P = \pi$, with $\sum \pi_i = 1$.

(c) Calculate the probability of a policy being ranked in Class 2 after two years of entering the system.

R: $exp(-\lambda) - exp(-2\lambda)$.

(d) Lemaire's (1985) *Transient Elasticity* evaluation measure is based on the following equation

$$V_\lambda(j) = b(j) + \beta_j \sum_{k=1}^{s} p_{T,\lambda}(j,k) V_\lambda(k), \; j = 1, ..., s \; ,$$

where$V_\lambda(j)$ is the expected present value to be paid by a policy from class C_j, $b(j)$ is the premium for Class j, β_j (< 1) is a discount rate, $p_{T,\lambda}(j,k)$ is a one-step transition probability and s is the number of classes.

Explain fully the formula and its components.

R: This is a *typical* one-step analisys expectation calculation. Premium $b(j)$ is paid in the begining of the period, then you have to consider the one step.transition probability $(p_{T,\lambda}(j,k))$ for getting the next period discounted expectation $(V_\lambda(k))$, then you have to apply the discount rate to put it at the same initial period... You have a system of s equations to be solved.

Consider now that λ is a realization of a random variable Λ following an exponential distribution with mean one

(e) Compute the resulting stationary distribution of the premia. Compute the average premium.

R:

$$\begin{aligned} \theta^2 &= \exp(-2\lambda) \\ \theta(1-\theta) &= (1-\exp(-\lambda))\exp(-\lambda) \\ 1-\theta &= 1-\exp(-\lambda) \\ E[\exp(-2\lambda)] &= \int_0^\infty \exp(-2\lambda)\exp(-\lambda)\, d\lambda = \frac{1}{3} \\ E[\exp(-\lambda)(1-\exp(-\lambda))] &= \int_0^\infty \exp(-\lambda)(1-\exp(-\lambda))\exp(-\lambda)\, d\lambda = \frac{1}{6} \\ E[1-\exp(-\lambda)] &= \int_0^\infty (1-\exp(-\lambda))\exp(-\lambda)\, d\lambda = \frac{1}{2} \\ AP &= 75\left(\frac{1}{3}\right) + 100\left(\frac{1}{6}\right) + 150\left(\frac{1}{2}\right) = 116.67 \end{aligned}$$

(f) Suppose now that the insurer decided to make a slight change in the *bonus* system, as follows: For a policy to move to a lower a class (apart form those already in Class 1) it is now necessary to have two years (in a row) without any claim. All other rules remain unchanged.

Make appropriate changes, if any, in the class set so that the system can be workable with the usual Markovian approach.

R: It will be now necessary to split previous classes, creating a new class set: C_3: Entry Class or class with neither aggravation nor discount for the first year; C_1: Class with discount; C_2: Class with neither discount nor aggravation two consecutive years; C_4: Class with penalty and no claims in the last year; C_5: Class with penalty and claims in the last year

3. For tariff modelling purposes regarding some insurance large portfolio, we studied some factors suposedly with impact in the claims frequency mean. We have first selected 5 factors labeled as F1 to F5. All these factors are qualitative, rank variables, ranked from 1 on, where Class 1 is the lower one. There are available information about Claims Counts (CC), Number of Policies (NP) for each cell. LNP is the logarithm of NP. We considered a Poisson model with logarithm link. For decisions consider a significance level of 5%. The Annex shows an estimated model, Model 1.

(a) After a brief analysis on the estimated model, argue about the quality of the model? Can it be improved? If so, clarify.

R: The model needs some improvement. Two levels (F22 and F43) and Factor 5 need attention, this last may not be relevant. Join the two levels, or even remove them both. Then re-run and analyse again, may be the former two levels need to be re-analysed..

(b) Conclude about the importance of level 3 of Factor 4 (F43).

R: This level needs special care as it is near the frontier decision, may be we don't need to remove/join levels after the analysis of the other factors and levels.

(c) Comment briefly the statement: *Level 1 of all factors are missing from the estimated table in the Annex because they were removed for having no significance.*

R: The statement is wrong, level 1 of each factor is in the estimated intercept, this is to avoid a multicolinearity problem...

(d) How would you calculate the expected frequency for the standard risk?

R: $exp(-1.80413)$.

(e) Would you accept this example to come from some sort of vehicle insurance? Explain briefly.

R: Yes, it's usual a tariff of vehicles insurance depending on different rank factors, like power of the vehicle, age class of the driver, vehicle age class, region, for instance.