Master in Actuarial Science, Exam 12/01/2013. **2h30m**
Ratemaking and Experience Rating, 2nd year, 1st semester

1. Consider a group of 340 households with residence in some particular urban area which has become an increasing target of burglary and theft. In the last year there were 210 claim cases of robbery distributed along with the following table:

No. of Claims	0	1	2	3
No. of households	200	80	50	10

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 credibility.

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 $z = n/(n + \upsilon/a)$, $\mu = E[\mu(\theta)]$, $\upsilon = E[\upsilon(\theta)]$, $a = V[\mu(\theta)$, $\mu(\theta)$ 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) Write down formulae for $E[P_c]$ and $Var[P_c]$ in terms of the structural parameters.
   (b) Calculate proper estimates for the structural parameters μ, υ and a.
   (c) Explain briefly the behaviour of the credibility factor z (and then, of P_c) as function of n, υ and a.
   (d) Suppose that the given risk produced two claims in the last year.
      i. Compute the (*empirical)* Bühlmann's estimate of the risk claims mean for the next period.
      ii. Assume from now on that the prior distribution of the risk parameter Θ is a Gamma(α, β), with mean $\alpha\beta$, show that the posterior distribution belongs to same family with the *new* parameters $\alpha_* = 2 + \alpha$ and $\beta_* = \left(1 + \beta^{-1}\right)^{-1}$.
      iii. Compute the Bayesian premium $E(\Theta|X = 2)$.
      iv. Compare the premiums P_c and $E(\Theta|X = 2)$ and explain similarities.
      v. How could you use the data above to estimate $E(\Theta|X = 2)$, we mean, to compute the *empirical* Bayes premium? Explain.

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 $\mathbf{b} = (80, 100, 150)$.

Let θ be the probability of a certain policy in the portfolio generates at least one claim in the year. Suppose that the number of claims per year of that policy can be explained by 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.
   (b) Compute the limiting distribution. Compute the average premium.
   (c) Calculate the elasticity of the average premium in stationary conditions for a risk with a claim frequency of 10%. Comment briefly.

   Consider now that λ is a realization of a random variable Λ following an exponential distribution with mean 0.01.

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

(e) 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 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 in the class set so that the system can be workable with the usual Markovian approach.

3. For tariff modelling purposes, we studied different factors 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 Class1 is the lower one. There are available information about the Total Number of Claims (TNC), Number of Policies (NP) for each cell. LTNC and LNP is the logarithm of TNC and NP, respectively. We considered a Poisson model with logarithm link. For decisions consider a significance level of 5%. The Annex shows there estimated models, models 1, 2 and 3.

(a) After a brief analysis on the three estimated models, choose the one that better fits the study purpose. Besides, would you consider the improvement of the model chosen? If so, clarify.

(b) Conclude about the importance of level 2 of Factor 2 (F22).

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

(d) Supose you don't reject the hypothesis of the levels 2 and 3 of factor F4 being equal. What would be your procedure?

(e) What would be the expected claims frequency of this portfolio?

Marks (out of 200):

1.*a*)	*b*)	*c*)	*d*)*i*.	*ii*.	*iii*.	*iv*.	*v*.	2.*a*)	*b*)	*c*)	*d*)	*e*)	3.*a*)	*b*)	*c*)	*d*)	*e*)
15	15	15	15	20	10	10	10	10	10	15	20	10	5	5	5	5	5
(15	30	45	60	80	90	100	110	120	130	145	165	175	180	185	190	195	200)

ANNEX

Model 1

```
Call:
glm(formula = TNC ~ 1, family = poisson, data = freq2, offset = LNP)
Deviance Residuals:
    Min       1Q   Median       3Q      Max
-5.6605  -0.9802  -0.1937   0.6576   4.7653

Coefficients:
             Estimate Std. Error z value Pr(>|z|)
(Intercept) -1.831725   0.008448  -216.8   <2e-16 ***
---
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: 1622.0  on 1077  degrees of freedom
AIC: 5207.1
Number of Fisher Scoring iterations: 4
```

Model 2

```
Call:
glm(formula = TNC ~ 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
```

Model 3

```
Call:
glm(formula = TNC ~ F1 + F2 + F3 + F4 + F5 , family = quasi(link =
"log", var = "mu"), 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 t value Pr(>|t|)
(Intercept) -1.80413    0.04147 -43.509  < 2e-16 ***
F12         -0.15887    0.03346  -4.748 2.33e-06 ***
F13         -0.12305    0.03672  -3.351 0.000834 ***
F14         -0.30068    0.04249  -7.077 2.67e-12 ***
F15         -0.32130    0.05075  -6.331 3.57e-10 ***
F22          0.03211    0.02475   1.297 0.194780
F23          0.12177    0.03036   4.011 6.47e-05 ***
F32          0.09465    0.03283   2.883 0.004018 **
F42          0.09466    0.03733   2.536 0.011364 *
F43          0.06811    0.04215   1.616 0.106427
F44          0.16338    0.04569   3.576 0.000365 ***
F45          0.15807    0.05085   3.109 0.001929 **
F52         -0.02656    0.02329  -1.140 0.254510
F53         -0.02626    0.02766  -0.949 0.342693
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for quasi family taken to be 1.368097)

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

Number of Fisher Scoring iterations: 4
```