LISBON SCHOOL OF ECONOMICS & MANAGEMENT
UNIVERSIDADE DE LISBOA

Master in Actuarial Science
Rate Making and Experience Rating

Exam 2, 04/02/2019, 9:00, Room 106F1
Time allowed: 2:30

Instructions:

1. This paper contains 4 groups of questions and comprises 4 pages including the title page;
2. Enter all requested details on the cover sheet;
3. You have 10 minutes of reading time. You must not start writing your answers until instructed to do so;
4. Number the pages of the paper where you are going to write your answers;
5. Attempt all questions;
6. Begin your answer to each of the 4 question groups on a new page;
7. Marks are shown in brackets. Total marks: 200;
8. Show calculations where appropriate;
9. An approved calculator may be used. No mobile phones or other communication devices are permitted;
10. The distributed Formulary and the Formulae and Tables for Actuarial Examinations (the 2002 edition) may be used. Note that the parametrization used for the different distributions is that of the distributed Formulary.

1. Consider an insurance portfolio where risks are classified into three types, say 1, 2, and 3. Each individual risk in the portfolio can produce annually an aggregate amount of one of the following quantities (in €10^3 units): 0, 5 or 10, only. **[85]**

   The composition of the portfolio as well as the individual probabilities are given in the following table:

Type	1	2	3
Composition	50%	40%	10%
Claim size			
0	0.85	0.75	0.65
5	0.10	0.15	0.20
10	0.05	0.10	0.15

   Risks has been observed for two consecutive years, let X denote the annual aggregate claim amount for a risk taken out at random from the portfolio, and let θ denotes its risk characteristic (a particular outcome from a random variable Θ, it represents the risk type).

   Consider that the usual hypothesis (H_1 and H_2) in credibility theory are applicable to the risk group under study (with the additional given by Bühlmann's model, where appropriate).

   Bühlmann's credibility formula for a given risk X in a homogeneous portfolio, for a coming year of exposure, is given by

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

   where $z = n/(n+k)$, $k = \upsilon/a$, $\mu = E[\mu(\theta)]$, $\upsilon = E[\upsilon(\theta)]$, $a = Var[\mu(\theta)]$, $\mu(\theta)$ 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) Calculate $\mu(\theta)$, $\upsilon(\theta)$, μ, υ, and a. [15]

   (b) Determine the probability function for the annual aggregate claims X. [10]

   Let X_1 and X_2 be the annual aggregate claim amounts for two consecutive years. Let $S = X_1 + X_2$.

   (c) Calculate the probability of having been observed $S = 10$. [15]

   (d) Compute the posterior distribution $\pi_{\Theta|S}(\theta|10)$. [10]

   (e) Determine the conditional distribution of next year aggregate amount X_3, given $S = 10$: $f_{X_3|S}(x|10)$. [15]

   (f) Calculate the Bayesian premium for next year, year 3. [7.5]

   (g) Find Bühlmann's credibility premium. [7.5]

   (h) Compare the figures found for the Bayesian and Bühlmann's premia. Suppose you had to choose between one of them, which one would you choose? Comment briefly both situations. [5]

2. From a given portfolio, suppose that you have observed a certain risk X for n years. Consider Bühlmann's H_1 and H_2 hypothesis and that given $\Theta = \theta$, the n risk observations, X_j, $j = 1, 2, \ldots, n$, are i.i.d. with f.p. **[35]**

   $$f_{X|\theta}(x|\theta) = \binom{k}{x}\theta^x (1-\theta)^{k-x}\,, \quad x = 0, 1, 2, \cdots, k$$

   Also, $\Theta \frown Beta(\alpha, \beta)$ with p.d.f.

   $$\pi(\theta) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}\theta^{\alpha-1}(1-\theta)^{\beta-1}\,, \quad 0 < \theta < 1\,,$$

   where $\Gamma(\cdot)$ is the Gamma function.

   (a) Show that the posterior is a Beta distribution with parameters α_* and β_*. Identify parameters α_* and β_*. [15]

   (b) Determine the risk premium and the collective premium. [5]

   (c) Calculate the Bayesian premium. [5]

   (d) Calulate the credibility premium, express in credibility form. [10]

3. A certain insurer is considering a *bonus-malus* system (BMS) based on the individual's annual claim count record to rate each individual risk in a given motor insurance portfolio. **[50]**

Class number	Premium level %	New step after claims 0	1	2	3	4	5	6+
1	250	2	1	1	1	1	1	1
2	200	3	1	1	1	1	1	1
3	150	4	1	1	1	1	1	1
4	125	5	2	1	1	1	1	1
5	100	6	3	1	1	1	1	1
6	90	7	4	2	1	1	1	1
7	80	8	5	3	1	1	1	1
8	70	9	6	4	2	1	1	1
9	60	10	7	5	3	1	1	1
10	50	11	8	6	4	2	1	1
11	40	12	9	7	5	3	1	1
12	37.5	12	10	8	6	4	2	1

Table 1: Rules and premium percentages

(a) For motor insurance, BMS based on Markov chains have been chosen by most ratemakers, in preference to using credibility models. Describe briefly reasons for that choice.

(b) Consider a *bonus* system that evolves according to what is shown in Table 1, and that a Poisson($\lambda = 0.1$) distribution models the claim counts appropriately. Define $t_{ij}(k)$, such that

$$t_{ij}(k) = \begin{cases} 1 & \text{if policy transfers from } i \text{ to } j\,, \\ 0 & \text{otherwise,} \end{cases}$$

if k claims are reported. Write the Transition Rule Matrix $\mathbf{T}(k) = [t_{ij}(k)]$, for $k = 0$ and $k = 2$, only.

(c) Refer to the system above and Table 1. Consider a driver that originally entered the scale at level 100%, drove with no claims for six years, then filed one claim during the seventh year, and has been driving claim-free for two years since then. Would the total of the premiums the driver paid (as percentage of the basic premium) have been different if the one claim filed had occurred in the second year?

(d) Determine the value of θ such that the transition probability matrix $\mathbf{P}$ has vector $(2/3, 1/3)$ as its steady state vector, if is given by

$$\mathbf{P} = \begin{pmatrix} 1-\theta & \theta \\ 0.5 & 0.5 \end{pmatrix}.$$

4. Suppose you work as an actuary for an insurer and you need to define a new tariff for the third party liability in motor insurance to put in place for the coming year 2020. As done in the past we model separately the claim frequency (denoted as N) and the expected cost per unit (denoted as Y). For the expected cost we decided to consider the following rating factors: [**30**]

- Driver's age (Age 1: ≤ 20 years of age, 2: 21-25, 3: 26-35, 4: 36-50 and 5: > 50)
- Experience (1: ≤ 2 license years , 2: > 2)
- Region (1: Greater Lisbon, 2: Greater Oporto, 3: South, 4: Centre and 5: North)
- Horse power (HP 1: ≤ 75, 2: $75 - 120$ and 3: > 120)

Consider the model output shown in the Annex.

(a) If you wanted to estimate directly the annual aggregate mean, denoted as $E[S]$ $(= E[N] \times E[Y])$, what sort of model would you use?

(b) Compute the ratio between the largest and the smallest claim cost expectation.

(c) Comment briefly the statement:

> ...levels 'Age1', 'Experience1', 'Region1' and 'HP1' are not shown in the output because they were taken out after being rejected in a significance test.

The model estimate does not show figures for Levels

(d) Assume that the proposed estimation is valid (or last). Suppose you select a driver with 25 years of age, 2 years of Experience, from greater Lisbon area, driving a vehicle with 120HP, and expected frequency of 0.1. What would be the corresponding pure premium?

Annex

```
Call:
glm(formula = Y ~ Age + Experience + Region + HP, family =
Gamma(link = "log"), data = dados)

Deviance Residuals:
    Min       1Q   Median       3Q      Max
-2.0741  -0.4282  -0.0868   0.2428   4.8993

Coefficients:
              Estimate Std. Error t value Pr(>|t|)
(Intercept)   7.287927   0.018583 392.181  < 2e-16 ***
Age2         -0.072685   0.015306  -4.749 2.07e-06 ***
Age3         -0.045600   0.016796  -2.715  0.00664 **
Age4         -0.234246   0.018794 -12.464  < 2e-16 ***
Age5         -0.254591   0.022173 -11.482  < 2e-16 ***
Experience2   0.087757   0.013683   6.414 1.46e-10 ***
Region2       0.063669   0.015509   4.105 4.06e-05 ***
Region3       0.030505   0.017394   1.754  0.07949 .
Region4       0.128912   0.019063   6.763 1.40e-11 ***
Region5       0.104366   0.021465   4.862 1.17e-06 ***
HP2           0.050938   0.010551   4.828 1.39e-06 ***
HP3           0.148721   0.013310  11.174  < 2e-16 ***
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

(Dispersion parameter for Gamma family taken to be 0.302994)

    Null deviance: 4590.4  on 16353  degrees of freedom
Residual deviance: 4422.0  on 16340  degrees of freedom
AIC: 260880

Number of Fisher Scoring iterations: 5
```

Solutions:

1.

(a) $\mu = 1.45$, $v = 8.9$, $a = 0.2475$, and

θ	1	2	3
$\mu(\theta)$	1	1.75	2.5
$v(\theta)$	6.5	10.6875	13.75

(b) $f(x) = \sum_{i=1}^{3} f(x|i)\pi(i)$, $x = 0, 5, 10$:

x	0	5	10
$f(x)$	0.79	0.13	0.08

(c) $Pr[S = 10]Pr[X1 + X_2 = 10] = 2\left(\sum_{i=1}^{3} f(0|i)f(10|i)\pi(i)\right) + \sum_{i=1}^{3} f(5|i)^2\pi(i) = 0.14.$

(d)

$$\pi(\theta|S = 10) = \frac{Pr(S = 10|\theta)\pi(\theta)}{Pr(S = 10)}\,, \quad \theta = 1, 2, 3\,.$$

θ	1	2	3	
$\pi(\theta\|10)$	0.339285714	0.492857143	0.167857143	1

(e)

$$Pr(X_3 = x|S = 10) = \sum_{\theta=1,2,3} Pr(X_3 = x|\theta)\pi(\theta|S = 10)\,,$$

$X_3 = x$	0	5	10
$f(x\|S = 10)$	0.767142857	0.141428571	0.091428571

(f) $E(X_3|S = 10) = 1.621428571.$

(g)

$$n = 2,\ \bar{X} = 5,\ k = 35.95959596,\ z = 0.0526876,\ P_c = 1.637040979\,.$$

(h) They are different, we don't have credibility exact. In this case Bayesian premium is not a linear function on the observations, Bühlmann's is the best linear estimator under the same criteria, is an optimal estimator under more restrictive conditions, this leads us to choose the Bayesian premium estimate.

2. From a given portfolio, suppose that you have observed a certain risk X for n years. Consider Bühlmann's H_1 and H_2 hypothesis and that given $\Theta = \theta$, the n risk observations, X_j, $j = 1, 2, \ldots, n$, are i.i.d. with f.p.

$$f_{X|\theta}(x|\theta) = \binom{k}{x}\theta^x\,(1-\theta)^{k-x}\,, \quad x = 0, 1, 2, \cdots, k$$

Also, $\Theta \frown Beta(\alpha, \beta)$ with p.d.f.

$$\pi(\theta) = \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}\theta^{\alpha-1}\,(1-\theta)^{\beta-1}\,, \quad 0 < \theta < 1\,,$$

where $\Gamma(\cdot)$ is the Gamma function. **[15]**

(a) Let $\underline{x} = (x_1, \ldots, x_n)$, and let $L(\theta)$ be the Likelihood fucntion:

$$\begin{aligned}\pi(\theta|\underline{x}) &\propto L(\theta)\pi(\theta) = \left[\prod_{j=1}^{n} f(x_j|\theta)\right]\pi(\theta)\\ &\propto \left[\prod_{j=1}^{n} \theta^{x_j-1}(1-\theta)^{k-x_j}\right]\theta^{\alpha-1}(1-\theta)^{\beta-1}\\ &= \theta^{\alpha_*-1}(1-\theta)^{\beta_*-1}\,,\end{aligned}$$

with $\alpha_* = \alpha + \sum_{j=1}^{n} x_j$ and $\beta_* = \beta + n\,k - \sum_{j=1}^{n} x_j$. It's a $Beta(\alpha_*, \beta_*)$ distribution.

(b) Risk Premium: $E[X|\theta] = k\theta$. Collective premium:

$$\mu = E[\mu(\theta)] = k\, E[\theta] = k\, \frac{\alpha}{\alpha + \beta}\, .$$

(c) Bayesian premium.

$$E[X_{n+1}|\underline{x}] = E[\mu(\theta)] = k\, E[\theta] = k\, \frac{\alpha_*}{\alpha_* + \beta_*}\, .$$

(d) Credibility premium, expressed in credibility form: Since the Bayesian premium is a linear function of $\underline{x}$, also, we are in the presence of conjugate distributions, then the Bayesian premium is equal to the credibility premium:

$$\tilde{\mu}(\theta) = k\, \left(\frac{\sum x_j}{\alpha_* + \beta_*} + \frac{\alpha}{\alpha_* + \beta_*} \right) = z\, \bar{x} + (1 - z)\mu\, ,$$

with $z = n\, k/(\alpha + \beta + n\, k)$.

3.

(a) A BMS based on credibility is not favoured by regulators and/or managers, as the harsh penalties tend to: Encourage uninsured driving; Suggest hit-and-run behavior; Induce policyholders to leave the company after one accident...

(b)

$$T(0) = \begin{array}{cc} & \begin{matrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 & 11 & 12 \end{matrix} \\ \begin{matrix} 1\\2\\3\\4\\5\\6\\7\\8\\9\\10\\11\\12 \end{matrix} & \begin{pmatrix} & 1 & & & & & & & & & & \\ & & 1 & & & & & & & & & \\ & & & 1 & & & & & & & & \\ & & & & 1 & & & & & & & \\ & & & & & 1 & & & & & & \\ & & & & & & 1 & & & & & \\ & & & & & & & 1 & & & & \\ & & & & & & & & 1 & & & \\ & & & & & & & & & 1 & & \\ & & & & & & & & & & 1 & \\ & & & & & & & & & & & 1 \\ & & & & & & & & & & & 1 \end{pmatrix} \end{array}$$

$$T(2) = \begin{array}{cc} & \begin{matrix} 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 & 11 & 12 \end{matrix} \\ \begin{matrix} 1\\2\\3\\4\\5\\6\\7\\8\\9\\10\\11\\12 \end{matrix} & \begin{pmatrix} 1 & & & & & & & & & & & \\ 1 & & & & & & & & & & & \\ 1 & & & & & & & & & & & \\ 1 & & & & & & & & & & & \\ 1 & & & & & & & & & & & \\ & 1 & & & & & & & & & & \\ & & 1 & & & & & & & & & \\ & & & 1 & & & & & & & & \\ & & & & 1 & & & & & & & \\ & & & & & 1 & & & & & & \\ & & & & & & 1 & & & & & \\ & & & & & & & 1 & & & & \end{pmatrix} \end{array}$$

Empty cells are filled with "0's".

(c) 640 *vs* 805.

(d) 1/4.

4.

(a) In that case we can consider that S is a compound Poisson model, and we can use a Tweedie model estimation with parameter $1 < p < 2$, Poisson counts with Gamma severities. Tweedie models are scale invariant, they appear when the variance function is $v(\mu) = \mu^p$.

(b)

Max: $\exp\{7.287927 + 0.087757 + 0.128912 + 0.148721\} \simeq 2107.625\,;$

Min: $\exp\{7.287927 - 0.254591\} \simeq 1133.807\,.$ Ratio: $2107.625/1133.807 \simeq 1.86\,.$

(c) No. They are not there explicitly because of multicollinearity reasons, since we are dealing with categorical variables, they are *hidden* in the intercept estimate.

(d)

$$0.1\,\exp\{7.287927 - 0.072685 + 0.050938\} \simeq 143.11\,.$$