LISBON SCHOOL OF ECONOMICS & MANAGEMENT
UNIVERSIDADE DE LISBOA

Master in Actuarial Science
Rate Making and Experience Rating

Exam 2, 06/02/2020, 9:00-11:30, Room 116 F1
Time allowed: 2:30

Instructions:

1. This paper contains 4 groups of questions and comprises 3 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 a portfolio of a certain insurance business in which each individual risk can produce in one year claims of two types, say Type 1 and Type 2. In Type 1 claims, the annual total amount can be 0 (no claims) or €5,000; In Type 2 claims, similarly, annual total amounts can be 0 or €10 000. For each individual risk, Type 1 and Type 2 claims occur independently and are denoted as X_1 and X_2, so that total annual amount of claims per individual risk is denoted by $S = X_1 + X_2$. Each individual risk is taken from a collective, which contains three risk kinds (unknown for the insurer): Good (60% of the portfolio), Medium (30%), Bad (10%). For each risk kind the probability of having claims (5 000 or 10 000) in one year is given in the following table: [**95**]

Total claim size for each "Type"	Conditional probabilities Good	Medium	Bad
Type 1: 5 000	0.10	0.20	0.40
Type 2: 10 000	0.05	0.15	0.30

The risk kind (Good, Medium or Bad) is not observed by the insurer, the composition of the portfolio by risk kind is a market estimation, let the parameter θ denote the risk kind in the portfolio, θ is a realization of a random variable Θ, taking outcomes G ,M or B. Consider that the usual hypothesis (H_1 and H_2) in credibility theory are applicable to the risk group under study (with the addition to H_1 given by Bühlmann's model, where appropriate).

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

$$P_c = z\bar{Y} + (1 - z)\mu_Y, \tag{1}$$

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) Determine the conditional probability function $f_{S|\theta}(x|\theta)$ for the total amount S of a given risk. (10)

(b) Calculate the risk premia and the collective premium. (7.5)

(c) Calculate $\upsilon(\theta)$, $\theta = G, M, B$, υ and a. (7.5)

(d) Calculate the probability function of the annual total amount of claims, S, from a risk taken out at random from the portfolio. (10)

In the following consider that in two consecutive years the following values of S have been observed $(S_1, S_2) = (10\,000, 0)$, where appropriate.

(e) Calculate the probability of getting such an outcome. (10)

(f) Find the posterior distribution $\pi_{\Theta|\mathbf{s}}(\theta|(10\,000, 0))$. (10)

(g) Calculate the predictive probability function for the third year annual claim amount S_3. (10)

(h) Determine the Bayesian premium for year 3. (10)

(i) Compute Bühlmann's credibility premium and look for exact credibility. (10)

(j) On what condition(s) can we talk on *exact credibility model*? Which of the two premia would you choose? Comment appropriately. (10)

2. Let $P_{c,n+1}$ be Bühlmann's credibility premium for year $n + 1$, $n = 1, 2, \ldots$, based on the n previous annual observations [it is given by Formula (1)]. [**25**]

(a) Show that $P_{c,n+1}$ can be recursively calculated as (15)

$$P_{c,n+1} = \alpha_n X_n + (1 - \alpha_n)P_{c,n}\,, \qquad n = 1, 2, \ldots\,,$$

where weight α_n is some function of the credibility factor z [z is given in (1)].

(b) Discuss briefly the behaviour of weight α_n. (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**]

(a) Do you find adequate that *classical* Experience Rating in motor insurance do not consider the annual aggregate claims record to adjust future premia, with claim counts and sizes, since it should reflect better the risk behaviour than just the claim record? Discuss briefly the pros and *cons.* (12.5)

(b) Consider a *bonus* system that considers the following situations: 25% discount in case of no claims for 2 consecutive years; a 20% penalty if the policy has one claim in the last year; 40% penalty in the case of two or more claims in the last year. Entry class corresponds to a class with no bonus and no penalty. Premia, indices, are set on an annual basis.

i. The system is not directly Markovian. Explain briefly, and discuss change procedures, namely on the scale as well as in the premium vector. (12.5)

ii. Rearrange and define the set of classes under a Markov framework, as well as the corresponding premia vector. (12.5)

(c) Suppose that for some 3-state *bonus-malus* system and some given θ, the steady state premium distribution is given by vector $\big(\theta, (1-\theta)^2, \theta(1-\theta)\big)$, where θ is the probability of at least one claim in one year. Number of claims is mixed Poisson distributed with parameter taking values according to a mean one exponential distribution. (12.5)

Compute the resulting limiting distribution.

4. A working party is modelling a tariff for a given large portfolio of some line of business. The study group is re-evaluating an existing tariff, evaluating a wide variety of studied and non-studied risk factors supposed to bring impact in both the claim frequency and the claim size means. [**30**]

(a) Discuss briefly the obejctives of Prior and Posterior Ratemaking, in both ratings where risk behaviours are estimated/adjusted in order to be reflected in the premium calculation. (7.5)

(b) Generalised Linear Models (GLM) are common tools in ratemaking, where different kinds of risk factors are studied.

i. What is intended by "Risk factors", explaind briefly. (7.5)

ii. Endogenous and Exogenous variables can be of different nature, special care should be taken when defining the different levels for each exogenous variables. Explain, give examples. (7.5)

(c) When estimating the annual aggregate mean, denoted as $E[S]$ $(= E[N] \times E[Y])$, we can use either an additive or a multiplicative model, model either separately claims counts and claim sizes or the aggregate $E[S]$. Explain choices briefly. (7.5)

Solutions:

1. **R**:

   (a) Let $S = X_1 + X_2$, $f_{S|\theta}(x|\theta)$,

S	G	M	B
	$0.9(0.95)$	$0.8(0.85)$	$0.6(0.7)$
0	$= 0.855$	$= 0.680$	$= 0.420$
5 000	0.095	0.170	0.280
10 000	0.045	0.120	0.180
15 000	0, 005	0.030	0.120
	1	1	1

   (b) and

   (c)

Θ	G	M	B
$\mu(\theta)$	1 000	2 500	5 000
$\upsilon(\theta)$	7 000 000	16 750 000	27 000 000
μ		1 850	
υ		10 925 000	
a		1 552 500	

   (d) We want $f_S(x) = \sum_{i\in\{G,M,B\}} f_S(x|i)\pi(i)$, $x = 0, 5\,000, 10\,000, 15\,000$:

x	0	5 000	10 000	15 000
$f_S(x)$	0.759	0.136	0.081	0.024

   (e) We want $f_{\mathbf{S}}(10\,000, 0) = \sum_{\theta\in\{G,M,B\}} f_S(10\,000|\theta)f_S(0|\theta)\pi(\theta) = 0.055125$:

θ	G	M	B
$f_S(10\,000, 0|\theta)$	0.038475	0.0816	0.0756

   (f)

$$\pi(\theta|(10\,000, 0)) = \frac{Pr(S_1 = 10\,000|\theta)Pr(S_2 = 0|\theta)\pi(\theta)}{f_S(10\,000, 0)},$$

	$f_S(10\,000, 0|\theta)$	$\pi(\theta)$		$\pi(\theta|(10\,000, 0))$
G	0.038475	0.6	0.023085	0.41877551
M	0.0816	0.3	0.02448	0.444081633
B	0.0756	0.1	0.00756	0.137142857
		1	0.055125	1

   (g)

$$Pr(S_3 = x|(10\,000, 0)) = \sum_{\theta\in\{G,M,B\}} Pr(S = x|\theta)\pi(\theta|(10\,000, 0)),$$

$S_3 = x$	0	5 000	10 000	15 000
$f(x|(10\,000, 0))$	0.717628571	0.153677551	0.096820408	0.031873469

   (h) $E(S_3|(10\,000, 0)) \simeq 2214.69$.

   (i)

$$n = 2,\ \bar{S} = 5\,000\ \ k = 7.68115942\ \ z = 0.206586826\ \ P_c = 2500.748503\,.$$

   Bayes premium is clearly different from credibility premium, there is no "Exact Credibility.

   (j) 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. (a)

$$\begin{aligned}
P_{c,n+1} &= \frac{1}{n+k}\left(\sum_{i=1}^{n-1} X_i + X_n\right) + \frac{k}{n+k}\mu \\
&= \frac{1}{n+k}X_n + \frac{n-1+k}{n+k}\left(\frac{n-1}{n-1+k}\bar{X}_{n-1} + \frac{k}{n-1+k}\mu\right) \\
&= \alpha_n X_n + (1-\alpha_n)P_{c,n}\,, \quad n+1,2,\ldots \\
\alpha_n &= \frac{1}{n+k}\,.
\end{aligned}$$

(b) $\alpha_n = z/n$, is a function of z and n. The behaviour of z is very well known, it is a number between zero and one, $z \to 1$ when $n \to \infty$, it is small when k is high when compared to n, k gives a “notion” of the heterogeneity of the portfolio.

α_n decreases when n increases fading away, we mean, more information makes the premium to stabilize, as expected.

3.

(a) Certainly it would better reflect the risk behaviour, however it would turn premium calculation more complicated and not understandable for insureds, increasing admin costs. On the other hand basing just on claim counts bring other problems like the *bonus-hunger* problem and lack of information of risk behaviour...

(b) i. We have to split classes,increasing the number of levels in the scale, not increasing premia. Premia are set in an annual basis and to get a bonus we need to wait for two years, we have to consider situations where there are no claims in the first year...

ii. Rearrange and define the set of classes under a Markov framework, as well as the corresponding premia vector.

$$\mathbf{b} = (75, 100, 100, 120, 120, 140, 140)$$

C_1 25% *bonus class*

C_2 Class with neither bonus nor penalty, with no claims in the last year

C_3 Entry class

C_4 20% *penalty* and no claims last yr

C_5 20% *penalty*

C_6 Policies with 40% *penalty* and no claims last yr

C_7 Policies with 40% *penalty* and claims last yr

(c)

$$\begin{aligned}
E[\theta] &= 1 - E[e^{-X}] = 1/2 \\
E[(1-\theta)^2] &= E[e^{-2X}] = 1/3 \\
E[\theta(1-\theta)] &= E[e^{-X}(1-e^{-X})] = 1/6.
\end{aligned}$$

Then we get the vector $(1/2\,, 1/3\,, 1/6)$

4.

(a) Prior ratemaking means building a tariff for a certain (large) portfolio. Actuaries often define a certain number of different variables representing risk factors, with the help of historical data study and estimate quantitatively their influence in the risk reflected in the premium.

Posterior ratemaking, or experience rating, adapts each individual premium according to his own claim experience, like a bonus-malus system in motor insurance for instance.

(b) i. Risk factors are “proxys” of risk particularities, represented by variables that are supposed to explain differences in the risk behaviour of differnt groups supposedly existing, the presence of these variables are statistically studied as well as their influence, quantatively speaking.

ii. Factors can be of Nominal (binary), Ordinal/Categorical, Discrete or Continuous nature. Very often, for instance in motor insurance, they are of binary, discrete (finite) or categoriacal, not continuous and we should take care of multicolinearity. We should choose levels to be one less of the set chosen to be present.

(c) Multiplicative or additive model is used according to whether we want to consider how the different risk factors are related.

Modelling separately, or not, claim counts or sizes is a choice of the actuary to decide if he wants to separate, or not, risk factors influence. They can work in the same direction, similarly, quantittively different, or not. If we want to model the aggregate mean 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 $\upsilon(\mu) = \mu^p$.