Master in Actuarial Science, Exam 06/01/2012. **2h30m**
Ratemaking and Experience Rating, 2nd year, 1st semester

1. Consider a motor insurance portfolio where the population is classified into categories A, B and C, respectively, where A is Good drivers, B is Bad drivers and C is *Sports* drivers. The population of drivers is split as follows: 70% is in category A, 25% in B and 5% in C. For each driver in category A, there is a probability of 0.75 of having no claims in a year, a probability of 0.2 of having one claim and a probability of 0.05 of having two or more claims in a year. For each driver in category B these probabilities are 0.25, 0.4 and 0.35, respectively. For each driver in category C these probabilities are 0.3, 0.4 and 0.3, respectively.

   Risk parameter representing the kind of driver is denoted by θ, which is a realization of the random variable Θ. The insurer does not know the value of that parameter. Let X be the (observable) number of claims per year for a risk taken out at random from the whole portfolio. For a given $\Theta = \theta$ yearly observations $X_1, X_2, ...,$ make a random sample from risk X. The insurer finds crucial that the annual premium for a given risk might be adjusted by its claim record.

   (a) Consider a risk X taken out at random from the portfolio.

	A	B	C
	$\pi(A) = 0.7$	$\pi(B) = 0.25$	$\pi(C) = 0.05$
X	$Pr(X = i \mid \theta = A)$	$Pr(X = i \mid \theta = B)$	$Pr(X = i \mid \theta = C)$
0	0.75	0.25	0.3
1	0.20	0.40	0.4
2	0.05	0.35	0.3

   i. Calculate the mean and variance of X. (**15 Marks out of 200**)

$$\begin{aligned}
E(X) &= \mu = E(E(X|\theta)) = 0.7(0.3) + 0.25(1.1) + 0.05(1) = 0.535 \\
v(A) &= V(X|A) = 0.2 + 4(0.05) - 0.09 = 0.31 \\
v(B) &= V(X|B) = 0.4 + 4(0.35) - 1.1^2 = 0.59 \\
v(C) &= V(X|C) = 0.4 + 4(0.3) - 1.0^2 = 0.6 \\
v &= E(V(X|\theta)) = 0.7\,(0.31) + 0.25\,(0.59) + 0.05\,(0.6) = 0.394\,5 \\
a &= V(E(X|\theta)) = V(\mu(\theta)) = E\left(\mu(\theta)^2\right) - \left[E\left(\mu(\theta)\right)\right]^2 \\
&= 0.3^2\,(0.7) + 1.1^2\,(0.25) + 1\,(0.05) - 0.535^2 \\
&= 0.415\,5 - 0.535^2 \; = 0.129\,275 \\
V(X) &= a + v = 0.394\,5 + 0.129\,28 = 0.523\,78
\end{aligned}$$

   ii. Compute the probability function of X. (**10 Marks/200**)

$$\begin{aligned}
\Pr(X = 0) &= 0.75\,(0.7) + 0.25(0.25) + 0.3(0.05) = 0.602\,5 \\
\Pr(X = 1) &= 0.2(0.7) + 0.4(0.25) + 0.4(0.05) = 0.26 \\
\Pr(X = 2) &= 0.05(0.7) + 0.35(0.25) + 0.3(0.05) = 0.137\,5 \\
1.0 &= 0.602\,5 + 0.26 + 0.137
\end{aligned}$$

   (b) For a particular risk of the portfolio we observed in the last two years $X_1 = x_1 = 0$ and $X_2 = x_2 = 2$.

   i. For a given $\Theta = \theta$ of risk X observations, $X_1, X_2, \ldots$, are a random sample but X_1 and X_2 are not independent. Comment briefly. (**10/200**)
   Conditional independence does not imply indepencence (unconditional). In the first instance we can work like in *classical* statistics sampling theory. See next.

   ii. Compute $Cov[X_1, X_2]$. (**10/200**)

$$\begin{aligned}
Cov[X_1, X_2] &= E[Cov[X_1, X_2|\theta]] + Cov[E[X_1|\theta]; E[X_2|\theta]] = 0.129\,275 \\
Cov[X_1, X_2|\theta] &= 0 \text{ and } E[X_1|\theta] = E[X_2|\theta] \Rightarrow \\
Cov[E[X_1|\theta]; E[X_2|\theta]] &= V[E[X_1|\theta]] = 0.129\,275
\end{aligned}$$

iii. Compute the posterior probability function of Θ given $(X_1 = 0, X_2 = 2)$. (**15/200**)

$$\begin{aligned} f((0,2)|A) &= 0.75(0.05) = 0.0375 \\ f((0,2)|B) &= 0.25(0.35) = 0.087\,5 \\ f((0,2)|C) &= 0.3(0.3) = 0.09 \end{aligned}$$

$$f((0,2)) = 0.0375(0.7) + 0.087\,5(0.25) + 0.09(0.05) = 5.\,262\,5 \times 10^{-2}$$

$$\begin{aligned} \pi(A|(0,2)) &= \frac{0.0375\,(0.70)}{5.\,262\,5 \times 10^{-2}} = 0.498\,81 \\ \pi(B|(0,2)) &= \frac{0.087\,5(0.25)}{5.\,262\,5 \times 10^{-2}} = 0.415\,68 \\ \pi(C|(0,2)) &= \frac{0.09(0.05)}{5.\,262\,5 \times 10^{-2}} = 8.\,551\,1 \times 10^{-2} \\ 1.0 &= 0.498\,81 + 0.415\,68 + 8.\,551\,1 \times 10^{-2} \end{aligned}$$

iv. You do not know from which risk category the above sample comes. Carry out appropriate calculations to determine from which category the sample is most likely to have come. (**10/200**)
We can compare the probabilities $\Pr(X_1 = 0, X_2 = 2)|\theta = i)$, $i = A, B, C$ and see that the sample is more likely to come from category C.

v. We need to compute a (pure) premium for the next year:

A. Compute the collective pure premium. (**5/200**)
R: $E(X) = 0.535$, already calculated, above. (**10/200**)

B. Compute the Bayes premium $E[X_3|\mathbf{X} = (0,2)] = E(\mu(\Theta)|\mathbf{X}) =$

$$0.3\,(0.498\,81) + 1.1\,(0.415\,68) + 1\left(8.\,551\,1 \times 10^{-2}\right) = 0.692\,4$$

C. Compute Bühlmann's credibility premium, say, $\tilde{E}(X_3|\theta)$. (**15/200**)

$$\begin{aligned} k &= \frac{v}{a} = \frac{0.394\,5}{0.129\,275} = 3.\,051\,634 \\ z &= \frac{2}{2 + 3.\,051\,634} = 0.395\,911\,5; \ \bar{x} = 1 \\ \tilde{E}(X_3|\theta) &= 0.395\,911\,5\,(1) + (1 - 0.395\,911\,5)(0.535) = 0.719\,098\,8 \end{aligned}$$

D. Can we talk here on *Exact Credibility*? Comment appropriately. (**10/200**)
No, $\tilde{E}(X_3|\theta) \neq E(\mu(\Theta)|\mathbf{X} = (\mathbf{0}, \mathbf{2}))$.

2. Retrieve the problem and data of Exercise 1 above. Suppose now that the insurer uses a *Bonus-malus* system based on the claims frequency to rate the risks of that portfolio. The system has simply three classes, numbered 1, 2 and 3 and ranked increasingly from low to higher risk. 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.

Let $\alpha(\theta)$ be the probability of not having any claim in one year for a policy in with risk parameter θ. Entry class is Class 2 and premia vector is $\mathbf{b} = (70, 100, 150)$.

(a) Consider a policy with risk parameter θ.

i. Write the transition rules matrix and compute the one year transition probability. (**10/200**)

$$T = \begin{bmatrix} \{0\} & & \{1,2,...\} \\ \{0\} & & \{1,2,...\} \\ & \{0\} & \{1,2,...\} \end{bmatrix}; \ P = \begin{bmatrix} \alpha & & 1-\alpha \\ \alpha & & 1-\alpha \\ & \alpha & 1-\alpha \end{bmatrix}$$

ii. Comment on the existence of the stationary distribution. (**10/200**)
R: It has stationary distribution, state space is finite and $P_{ii}^{(n)} > 0$ for some $n = 1, 2, ...$

iii. Calculate the probability of a policy being ranked in Class 1 two years after entering the system. (**10/200**)
Let Y_i be the state of the process in year i, then

$$\begin{aligned}\Pr[Y_2=1] &= \sum_1^3 \Pr[Y_2=1|Y_1=i]\Pr[Y_1=i] \\ &= \sum_1^3 \Pr[Y_2=1|Y_1=i]\Pr[Y_1=i|Y_0=2] \\ &= \alpha^2=\alpha(\theta)^2, \text{ (depends on } \theta).\end{aligned}$$

iv. Calculate the probability function of the premium for a type A driver after two years of stay in the portfolio. Compute the average premium. (**10/200**)
Let P^* be the premium

$$\begin{aligned}\Pr[P^*=70|\theta=A] &= \alpha^2=0.75^2=0.5625 \\ \Pr[P^*=100|\theta=A] &= 0.75(0.25)=0.1875 \\ \Pr[P^*=1300|\theta=A] &= 0.75(0.25)+0.25^2=0.25 \\ 1.0 &= 0.5625+0.1875+0.25 \\ E[P^*|\theta=A] &= 70(0.5625)+100(0.1875)+150(0.25)=95.625\end{aligned}$$

v. After some time the insurer's chief actuary concluded that for ratemaking purposes it didn't make much difference to keep categories B and C apart, and merged them into, say, category B^*. For a driver in this new class, compute the probability function of the premium after one year of staying in the system (since his entry). (**10/200**)
New category B^* with $\pi(B^*)=0.3$

$$\begin{aligned}\alpha(B^*) &= \Pr[X=0|B^*]=\left(0.25^2+0.3(0.05)\right)/0.3=0.25833 \\ \Pr[X=1|B^*] &= (0.4(0.25)+0.4(0.05))/0.3=0.4 \\ \Pr[X=2|B^*] &= (0.35(0.25)+0.3(0.05))/0.3=0.34167 \\ 1.0 &= 0.25833+0.4+0.34167\end{aligned}$$

$$\begin{aligned}\Pr[Y_1=1|Y_0=2] &= \Pr[P^*=70]=\alpha=0.25833 \\ \Pr[Y_1=2|Y_0=2] &= \Pr[P^*=100]=0 \\ \Pr[Y_1=3|Y_0=2] &= \Pr[P^*=150]=1-\alpha=0.74167\end{aligned}$$

(b) Stationary distribution for a given θ is given by vector $\left(\alpha(\theta)^2;[1-\alpha(\theta)]\alpha(\theta);1-\alpha(\theta)\right)$. Compute the probability function of the premium for a policy taken out at random from the portfolio. Calculate the average premium. (**10/200**)

$$\begin{aligned}E[\alpha(\theta)] &= 0.75(0.7)+0.25(0.25)+0.3(0.05)=0.6025 \\ E[\alpha(\theta)^2] &= 0.75^2(0.7)+0.25^2(0.25)+0.3^2(0.05)=0.41388 \\ \Pr[P^*=70] &= E[\alpha(\theta)^2]=0.41388 \\ \Pr[P^*=100] &= E[(1-\alpha(\theta))\alpha(\theta)]=0.6025-0.41388=0.18862 \\ \Pr[P^*=150] &= E[(1-\alpha(\theta))]=1-0.6025=0.3975 \\ 1.0 &= 0.41388+0.18862+0.3975 \\ \text{Mean Premium} &= 0.41388(70)+0.18862(100)+0.3975(150)=107.46\end{aligned}$$

3. For tariff modelling purposes, we studied different factors with impact in the claim frequency mean and have found the model showed in the Annex, where

TVeic – 3 classes, in increasing power order of the vehicle

TProp – 2 classes (1=Private, 2=Company)

TUtil – 3 classes (1="male", 2="female", 3="retired")

Comb – 2 classes (1=petrol, 2=diesel)

Inexp–3 classes (0=driver's licence aged *five plus*, 1= from 2-5 years, 2=less than 2)

idSeg – "Dummy" variable valued 1 when policyholder is less than 25 years old

idveic – "Dummy" variable valued 1 when vehicle is aged more than 10 years

zona – Usual circulation geographical zone (4 zones in the country)

Expo – time period of the year the policy was in force (LExpo is the logarithm of Expo).

All the items below are based on the model presented. You can use an $\alpha = 0.05$ in all testes you make.

(a) Is it statistically acceptable to set a 30% aggravation for a type 2 vehicle (TVeic=2) when compared to the one of a type 1 (TVeic=1) (*ceteris paribus*)? (**7.5/200**)

R: Yes, corresponding $p - value \approx 0$, and relative estimated aggravation is

$$\frac{\exp\{-2.354489 + 0.28325\}}{\exp\{-2.354489\}} = 1.327\,437.$$

(b) Test if it is acceptable to consider the same rating factor for "female" and "retired". (**10/200**)

R: Test $H_0 : \beta_{TUtil2} = \beta_{TUtil3} \Leftrightarrow \beta_{TUtil2} - \beta_{TUtil3} = 0$... We need the value for $\hat{V}ar(\hat{\beta}_{TUtil2} - \hat{\beta}_{TUtil3})$, this can be calculated from values in the Var/Covar matrix attached.

(c) Set the situations that make the less and the greater claim frequency mean. (**7.5/200**)

$$\begin{array}{lll} Min & : & \text{Tveic1 “+”TProp1“+”TUtil3“+”Comb1“+” (Inexp0) “+” (idSeg0)”+”idVeic”+”Zona4} \\ Max & : & \text{Tveic3 “+”TProp2“+”TUtil0“+”Comb2“+”Inexp2 “+”idSeg”+” (idVeic0)”+” (Zona0)} \end{array}$$

(d) In the case you have chosen a Poisson model with logarithm link, what would be the effect in the output? (**2.5/200**)

R: Quasi-Poisson models allow to study over dispersion $Var(Y_i) = \phi E(Y_i)$.

(e) Why is the variable LExpo considered as "offset"? (**2.5/200**)

R: To consider risks with time exposure...