Ratemaking and Experience Rating

Master on Actuarial Science

Alfredo D. Egídio dos Reis

LISBON SCHOOL OF ECONOMICS & MANAGEMENT
UNIVERSIDADE DE LISBOA

Programme

- Denuit, M.; Maréchal, X.; Pitrebois, S. & Walhin, J-F. (2007). *Actuarial Modelling of Claim Counts: Risk Classification, Credibility and Bonus-malus Systems*, John Wiley & Sons, Chichester, England.
- Klugman, S.A.; Panjer, H.H. & Willmot, G.E. (2008 or 2012). *Loss Models, From Data to Decisions*, 3rd or 4th editions, John Wiley, Hoboken NJ.
- Kaas, R.,Goovaerts, M., Dhaene, J. & Denuit, M. (2008). *Modern Actuarial Risk Theory: Using R*, 2nd edition, Springer.
- Ohlsson, E. & Johansson, B. (2010). *Non-Life Insurance Pricing with Generalized Linear Models*, EAA series/EAA Lecture Notes, Springer.

1. **Ratemaking**:
    - "Pricing" insurance, calculation of Insurance *Premia*
    - Building a **tariff** for a portfolio, or portfolios somehow connected
2. **Experience rating**: adjust future premiums based on past experience
3. **Prior** and **Posterior** Ratemaking

Insurance **Premium**: Price for buying insurance (for a period). Two components:

1. Economic criteria: *market price*, admin costs
2. Actuarial criteria:
    - based on technical aspects of the risk
    - Meant to cover future claims
    - We only consider this here

- **Tariff**:
  - It's a *list of prices*
  - System of premiums for the risks of a portfolio (homogeneous)
  - Sets a base premium (homogeneous)
  - plus a set of bonus/malus (heterogeneous)
- **Exposure**: Risk volume, in risk units, no.
- **Risk unit**: Commonly, a policy; sometimes a set of policies
- **Claim**: an accident generates a claim, monetary amount
- **Claim frequency**: number of claims, distribution
- **Severity**: amount of the claim
- Loss reserving
- **Pure premium**: Risk mean, loss mean
- **Loss ratio**: paid claims/premiums

Let X be a given risk in a portfolio, with Pure Premium $E(X)$, unknown:

- If the risk is has been sufficiently observed

$$E(X) \simeq \overline{X} \quad (\textit{Full Credibility})$$

- If not, use *Partial Credibility*, **Credibility Formula**:

$$\begin{aligned} E(X) &\simeq z\overline{X} + (1-z)M \\ z &= \frac{n}{n+k} \end{aligned}$$

- Credibility factor: z, $0 \leq z < 1$
- n: No. observations; k: some positive constant
- M: Externally obtained mean (*Manual rate*).

Example

A given risk $X|\theta \frown Bin(1;\theta)$, obs'd 10 yrs, 20 risks. $\bar{X} = 0.0145$.

Ano i	1	2	3	4	5	6	7	8	9	10	11	12	13	14	15	16	17	18	19	20
1			1								1						1	1		
2							1		1		1						1			
3			1				1		1	1										
4																	1			
5									1		1									
6						1			1			1								
7									1					1			1			
8												1								
9						1					1		1							
10									1			1					1		1	
$\hat{\theta}_j$	0,0	0,0	0,2	0,0	0,0	0,2	0,2	0,0	0,6	0,1	0,4	0,3	0,1	0,1	0,0	0,0	0,5	0,1	0,1	0,0

1. *Limited Fluctuation Theory*: Classical approach
   1. From some computed $n : n > n_0$ use *Full credibility*;
   2. Otherwise: Use *Partial credibility*. But what M, k?
2. *Greatest Accuracy Theory*: Bayesian approach.

Example (Ex. 20.1, Classical, Full credibility)

Past losses: $X_1, X_2, \dots X_n$, estimate $\xi = E[X_j]$. (Normal) Find n:

$$
\begin{aligned}
Pr\left\{-r\xi \leq \bar{X} - \xi \leq r\xi\right\} &\geq p \\
Pr\left\{\left|\frac{\bar{X} - \xi}{\sigma/\sqrt{n}}\right| \leq \frac{r\xi\sqrt{n}}{\sigma}\right\} &\geq p
\end{aligned}
$$

Suppose 10 obs: 6 "0's" and 253, 398, 439, 756, $r = 0.05, p = 0.9$

$$
n \geq \left(\frac{z_\alpha}{r}\right)^2 \left(\frac{\sigma}{\xi}\right)^2 = 1082.41 \left(\frac{267.89}{184.6}\right)^2 = 2279.51
$$

Example (Ex. 20.1 cont'd, Classical, Partial credibility)

10 obs: 6 "0's" and 253, 398, 439, 756, $r = 0.05$, $p = 0.9$

$$n \geq 2279.51$$

$n = 10$ does not deserve full credibility. **Credibility Formula**:

$$E(X) \quad \simeq \quad z\overline{X} + (1 - z)M\,. \qquad (z = ?)$$

$$\begin{aligned} z &= \frac{n}{n+k} \\ z &= \min\left\{ \frac{\xi}{\sigma}\sqrt{\frac{n}{\lambda_0}}\,;\ 1 \right\} \\ z &= 0.06623 \\ P_c &= 0.06623(184.6) + 0.93377(225) = 222.32 \end{aligned}$$

Exercises 20.1, 20.3, pg 565

Theory, outgrowth of Buhlman's (1967) paper

Example (Ex. 20.9, Bayesian approach)

Two types of drivers: *Good* and *Bad*. Good are 75% of the population and in one year have have 0 claims w.p. 0.7, 1 w.p. 0.2 and 2 w.p. 0.1. Bad drivers, respectively, 25%, 0.5, 0.3, 0.2. when a driver buys insurance insurer does not know it's category. We assign an unknown risk parameter, θ.

Example (Ex. 20.9 cont.)

x	$P(X = x \mid \theta = G)$	$P(X = x \mid \theta = B)$	θ	$P(\Theta = \theta) = \pi(\theta)$
0	0.7	0.5	G	0.75
1	0.2	0.3	B	0.25
2	0.1	0.2		

Some basic concepts:

Recap Joint & conditional distr. & expectation

Bivariate random variable: (X, Y). D.f. $F_{X,Y}$, pdf or pf $f_{X,Y}$

- $f_{X,Y}(x, y)$, marginals f_X, f_y. If independent: $f_{X,Y} = f_X f_Y$.
- Conditional (Conditional ind.: $f_{X,Y|Z} = f_{X|Z} f_{Y|Z}$):

$$f_{X|Y}(x) = \frac{f_{X,Y}(x,y)}{f_Y(y)} \qquad f_{Y|X}(y) = \frac{f_{X,Y}(x,y)}{f_X(x)}$$

$$f_{X,Y}(x, y) = f_{X|Y}(x) f_Y(y) \qquad f_{X,Y}(x, y) = f_{Y|X}(y) f_X(x)$$

- Marginals

$$f_X(x) = \int f_{X,Y}(x,y)dy\,; \qquad f_Y(y) = \int f_{X,Y}(x,y)dx$$
$$f_X(x) = \int f_{X|Y}(x)f_Y(y)dy\,; \qquad f_Y(x) = \int f_{Y|x}(x)f_X(y)dx$$

- Expectations, Iterated expectation

$$\begin{aligned} E[E(X|Y)] &= E[X];\;\; E[E(Y|X)] = E[Y] \\ V[X] &= E[V(X|Y)] + V[E(X|Y)] \end{aligned}$$

$$Cov[X,Y] = E[Cov(X,Y|Z)] + Cov[E(X|Z);E(Y|Z)]$$

Example (Ex. 20.9 cont'd)

Suppose we observed for a particular risk: $\mathbf{X} = (X_1, X_2) = (0; 1)$.
Given θ obs are independent.

$$\begin{aligned}
f_{\mathbf{X}}(0,1) &= \sum_{\theta} f_{\mathbf{X}|,}(0,1|\theta)\pi(\theta) = \sum_{\theta} f_{X_1|\theta}(0|\theta) f_{X_2|\theta}(1|\theta)\pi(\theta) \\
&= 0.7(0.2)(0.75) + 0.5(0.3)(0.25) = 0.1425 \\
f_{\mathbf{X}}(0,1,x_3) &= \sum_{\theta} f_{\mathbf{X},\mathbf{X}_3|,}(0,1,x_3|\theta)\pi(\theta) \\
&= \sum_{\theta} f_{X_1|\theta}(0|\theta) f_{X_2|\theta}(1|\theta) f_{X_3|\theta}(x_3|\theta)\pi(\theta) \\
f(0,1,0) &= 0.09995;\ f(0,1,1) = 0.003225;\ f(0,1,2) = 0.01800
\end{aligned}$$

Predictive and Posterior distribution

$$\begin{aligned}
f(0|0,1) &= 0.647368;\ f(1|0,1) = 0.226316;\ f(2|0,1) = 0.126316 \\
\pi(G|0,1) &= 0.736842;\ \pi(B|0,1) = 0.263158
\end{aligned}$$

Example (Ex. 20.11)

Let $X|\theta \frown Poisson(\theta)$ and
$\Theta \frown Gamma(\alpha, \beta) \Rightarrow X \frown NBinomial(\alpha, \beta)$

$$\begin{aligned} E(X|\theta) &= \theta \Rightarrow \\ E(X) &= E(E(X|\Theta)) = E(\Theta) = \alpha\beta \\ V(X|\theta) &= \theta \Rightarrow \\ V(X) &= V(E(X|\Theta)) + E(V(X|\Theta)) = \alpha\beta\,(1+\beta) \end{aligned}$$

Example (Ex. 20.10)

Let $X|\theta \frown \exp(1/\theta)$, mean $1/\theta$, and $\Theta \frown Gamma(4, 0.001)$.

$$\begin{aligned} f(x|\theta) &= \theta e^{-\theta x},\ x,\theta > 0 \\ \pi(\theta) &= \theta^3 e^{-1000\theta} 1000^4/6,\ \theta > 0 \end{aligned}$$

Example (Ex. 20.10)

Suppose a risk had 3 claims of 100, 950, 450.

$$\begin{aligned} f(100, 950, 450) &= \int_0^\infty f(100, 950, 450|\theta)\pi(\theta)d\theta \\ &= \int_0^\infty f(100|\theta)f(950|\theta)f(450|\theta)\pi(\theta)d\theta \\ &= \frac{1,000^4}{6}\frac{6!}{2,500^7} \end{aligned}$$

Similarly,

$$f(100, 950, 450, x_4) = \frac{1,000^4}{6}\frac{7!}{(2,500 + x_4)^8}$$

Example (Ex. 20.10)

Predictive density, posterior density

$$
\begin{aligned}
f(x_4|100, 950, 450) &= \frac{7\,(2500)^7}{(2,500 + x_4)^8} \rightarrow Pareto(7; 2500) \\
\pi(\theta|100, 950, 450) &= \theta^6 e^{-2500\theta} 2500^7/\Gamma(7) \\
&\rightarrow Gamma(7; 1/2500)
\end{aligned}
$$

(Conjugate distributions) Risk premium and *potential* estimates:

$$
\begin{aligned}
\mu_4(\theta) &= E(X_4|\theta) = ? \\
E(X_4|100, 950, 450) &= 416,67 \\
\mu &= E(X_4) = E(1/\Theta) = 1000/3 = 333.3(3) \\
\bar{X} &= 500 \\
\mu &< E(X_4|100, 950, 450) < \bar{X}
\end{aligned}
$$

Exercices 20.20, 20.23, p. 605.

Bayesian approach

From now onwards, assume a Bayesian approach:

Let a portfolio of risks, homogeneous, but "different":

- Homogeneous: risks follow the same distribution family
- Heterogeneous: distribution parameter is different.

A **given risk** comes attached with a paramenter θ:

- Fixed, but unknown, not observable;
- Only claims are observed: $(X_1, X_2, ..., X_n) = \mathbf{X}$;
- θ is the hidden aspects of the risk, which differs from others;
- Like classical statistics: Use past data $\mathbf{X}$ to predict X_{n+1}
- **Risk (pure) Premium**: $E(X_{n+1}|\theta) = \mu_{n+1}(\theta)$.
- Opposed to **Collective (pure) Premium**: $E(X_{n+1}) = \mu_{n+1}$.

Hypothesis

- H1 Given θ, $X_1|\theta, X_2|\theta, \ldots, X_n|\theta, X_{n+1}|\theta$ are (conditionally) independent.
  θ is realization of a random variable: $\Theta \frown \pi(\theta)$
- H2 The different risks in the portfolio are independent.

Premium for the next year:

- **Risk Premium**: $E(X_{n+1}|\theta) = \mu_{n+1}(\theta)$. Unknown.
- **Collective Premium**: $E(E(X_{n+1}|\theta)) = \mu_{n+1}$. In general $\mu_{n+1}(\theta) \neq \mu_{n+1}$
- **Bayesian premium** (mean of the predictive dist. and Bayes estimate for the *squared-error loss*):

$$
\begin{aligned}
E(X_{n+1}|\mathbf{X}) &= \int x f_{X_{n+1}|\mathbf{X}}(x|\mathbf{x})dx \\
&= \int \mu_{n+1}(\theta)\pi_{\Theta|\mathbf{X}}(\theta|\mathbf{x})d\theta
\end{aligned}
$$

Some Basic concepts:

$\mathbf{X} = (X_1, X_2 \ldots, X_n)$; **Predictive** distribution: $f_{X_{n+1}|\mathbf{X}}(x|\mathbf{x})$; **Prior** distr.: $\pi_\Theta(\theta)$; and **Posterior** dist.: $\pi_{\Theta|\mathbf{X}}(\theta|\mathbf{x})$

- Posterior dist.:

$$\pi_{\Theta|\mathbf{X}}(\theta|\mathbf{x}) = \frac{f_{\Theta,\mathbf{X}}(\theta,\mathbf{x})}{f_{\mathbf{X}}(\mathbf{x})} = \frac{f_{\mathbf{X}|\Theta}(\mathbf{x}|\theta)\pi(\theta)}{\int f_{\mathbf{X}|\Theta}(\mathbf{x}|\theta)\pi(\theta)d\theta}$$

- Preditive dist.:

$$\begin{aligned} f_{X_{n+1}|\mathbf{X}}(x|\mathbf{x})dx &= \frac{f_{X_{n+1};\mathbf{X}}(x;\mathbf{x})}{f_{\mathbf{X}}(\mathbf{x})} = \frac{\int f_{X_{n+1},\mathbf{X}|\Theta}(x,\mathbf{x}|\theta)\pi_\Theta(\theta)d\theta}{f_{\mathbf{X}}(\mathbf{x})} \\ &= \int f_{X_{n+1}|\Theta}(x|\theta)\frac{f_{\mathbf{X}|\Theta}(\mathbf{x}|\theta)\pi_\Theta(\theta)d\theta}{f_{\mathbf{X}}(\mathbf{x})} \\ &= \int f_{X_{n+1}|\Theta}(x|\theta)\pi_{\Theta|\mathbf{X}}(\theta|\mathbf{x})d\theta \end{aligned}$$

Definition (Credibility Premium)

The Credibility (pure) Premium $\widetilde{\mu_{n+1}}(\theta) = \alpha_0 + \sum_{j=1}^{n} \alpha_j X_j$ is an estimator of linear form, such that:

$$\min Q = E\left\{\left[\mu_{n+1}(\Theta) - \left(\alpha_0 + \sum_{j=1}^{n} \alpha_j X_j\right)\right]^2\right\}$$

Solution: Find $\alpha_0, \alpha_1, \ldots, \alpha_n$:

$$\begin{aligned}
\frac{\partial}{\partial \alpha_0} Q &= -2E\left\{\mu_{n+1}(\Theta) - \left(\alpha_0 + \sum_{j=1}^{n} \alpha_j X_j\right)\right\} = 0 \\
\frac{\partial}{\partial \alpha_i} Q &= -2E\left\{\left[\mu_{n+1}(\Theta) - \left(\alpha_0 + \sum_{j=1}^{n} \alpha_j X_j\right)\right] X_i\right\} = 0,\ i = 1, \ldots, n
\end{aligned}$$

θ, $X_1, X_2, \ldots, X_n, X_{n+1}$ are all random variables.

Equivalent to

$$E[\mu_{n+1}(\Theta)] = \widetilde{\alpha}_0 + \sum_{j=1}^{n} \widetilde{\alpha}_j E\left[X_j\right] = E\left(\widetilde{\mu_{n+1}}(\theta)\right);$$

$$E[\mu_{n+1}(\Theta)X_i] = \widetilde{\alpha}_0 E[X_i] + \sum_{j=1}^{n} \widetilde{\alpha}_j E\left[X_i, X_j\right],\ i = 1, ..., n.$$

Or,

Normal equations

$\widetilde{\alpha}_0, \widetilde{\alpha}_1, \ldots, \widetilde{\alpha}_n$ such that:

$$E(X_{n+1}) = \widetilde{\alpha}_0 + \sum_{j=1}^{n} \widetilde{\alpha}_j E\left[X_j\right] = E\left(\widetilde{\mu_{n+1}}(\theta)\right);$$

(unbiasedness equation)

$$Cov(X_i, X_{n+1}) = \sum_{j=1}^{n} \widetilde{\alpha}_j Cov\left[X_i, X_j\right],\ i = 1, ..., n.$$

We know that
$E\left[X_{n+1}\right] = E\left[E\left[X_{n+1}|\mathbf{X}\right]\right] = E\left[E\left[X_{n+1}|\Theta\right]\right] = E\left[\mu_{n+1}(\Theta)\right];$
$\mu_{n+1}(\theta) = E\left[X_{n+1}|\theta\right].$

$\widetilde{\mu_{n+1}}(\theta)$ also minimises, $\mathbf{X} = (X_1, \ldots, X_n)$,

$$
\begin{aligned}
\min Q &= \min E\left\{\left[\mu_{n+1}(\Theta) - \left(\alpha_0 + \sum_{j=1}^{n} \alpha_j X_j\right)\right]^2\right\} \\
&= \min E\left\{\left[E\left[X_{n+1}|\mathbf{X}\right] - \left(\alpha_0 + \sum_{j=1}^{n} \alpha_j X_j\right)\right]^2\right\} \\
&= \min E\left\{\left[X_{n+1} - \left(\alpha_0 + \sum_{j=1}^{n} \alpha_j X_j\right)\right]^2\right\}
\end{aligned}
$$

Bühlmann's model

Initial hypothesis

1. Given θ, $X_1|\theta, X_2|\theta, ..., X_n|\theta, X_{n+1}|\theta$ are (conditionally) independent.
   θ is realization of a random variable: $\Theta \frown \pi(\theta)$
2. The different risks in the portfolio are independent.

Addition to **H1**

1. Given θ, $X_1|\theta, X_2|\theta, ..., X_n|\theta, X_{n+1}|\theta$ have the same mean and variance:

$$\begin{aligned} \mu(\theta) &= E\left(X_j|\theta\right) \\ v(\theta) &= Var\left(X_j|\theta\right). \end{aligned}$$

Let

$$\mu = E\left[\mu(\theta)\right],\ v = E\left[v(\theta)\right],\ a = Var\left[\mu(\theta)\right]$$

Solution:

$$
\begin{aligned}
\widetilde{\mu_{n+1}}(\theta) &= \widetilde{\alpha}_0 + \sum_{j=1}^{n} \widetilde{\alpha}_j X_j = z\overline{X} + (1-z)\mu \\
z &= \frac{n}{n+k} \\
k &= v/a
\end{aligned}
$$

1. z: called Bühlmann's credibility factor
2. Credibility premium is a weighted average from $\overline{X}$ and μ.
3. $z \to 1$ when $n \to \infty$, more credit to sample mean
4. If portfolio is fairly homogeneous w.r.t. Θ, then $\mu(\Theta)$ does not vary much, hence small variability.
   Thus a is small relative to $v \to k$ is large, z is closer to 0
5. Conversely, if the portfolio is heterogeneous, z is closer to 1
6. Bühlmann's model is the simplest credibility model, no change over time

Proof

Estimator proposed for given risk, say j : $\hat{m}_j = \alpha + \beta \bar{X}_{.j}$, so that

$$\min R = \min \mathrm{E}\left[(\mu(\theta_j) - \hat{m}_j)^2\right] = \min \mathrm{E}\left[(\mu(\theta_j) - \alpha - \beta \bar{X}_{.j})^2\right].$$

Set

$$\begin{aligned} \mathrm{E}\left[((\mu(\theta_j) - \beta \bar{X}_j]) - \alpha)^2\right] &= \mathrm{V}[\mu(\theta_j) - \beta \bar{X}_j] \\ &\quad + (\mathrm{E}\left[\mu(\theta_j) - \beta \bar{X}_{.j}\right] - \alpha)^2 \end{aligned}$$

Minimizing α, such that:

$$\begin{aligned} \alpha^* &= \mathrm{E}[\mu(\theta_j) - \beta^* \bar{X}_{.j}] = \mathrm{E}[\mu(\theta_j)] - \beta^* \,\mathrm{E}[\bar{X}_{.j}]. \\ \alpha^* &= (1 - \beta^*)\,\mathrm{E}[\mu(\theta_j)], \text{ since} \\ \mathrm{E}[\bar{X}_{.j}] &= \mathrm{E}[\mathrm{E}[\bar{X}_{.j} | \theta_j]] = \mathrm{E}[\mu(\theta_j)] \end{aligned}$$

Proof (cont'd)

2nd part

$$
\begin{aligned}
\mathrm{V}[\mu(\theta_j) - \beta\,\bar{X}_{.j}] &= \mathrm{E}[\mathrm{V}[\mu(\theta_j) - \beta\,\bar{X}_{.j}|\theta_j]] + \mathrm{V}[\mathrm{E}[\mu(\theta_j) - \beta\,\bar{X}_{.j}|\theta_j]] \\
&= \frac{\beta^2}{n}\mathrm{E}[v(\theta)] + (1-\beta)^2\mathrm{V}[\mu(\theta_j)]. \\
&= \frac{\beta^2}{n}v + (1-\beta)^2 a. \\
\mathrm{V}[\bar{X}_{.j}|\theta_j] &= \frac{1}{n}\mathrm{V}[X_{ij}|\theta_j]
\end{aligned}
$$

Differentiating w.r.t. β and equating,

$$
\begin{aligned}
\frac{2\,\beta}{n}v - 2(1-\beta)a = 0\;, \\
\beta^* \;=\; \frac{a}{a+\frac{1}{n}v} = \frac{n}{n+v/a}
\end{aligned}
$$

Theorem

Let $P_{c,n+1}$ denote Bühlmann's credibility premium for year $n+1$, $n = 1, 2, \dots,$ based on the n previous annual observations. $P_{c,n+1}$ can be recursively calculated as the weighted average

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

with weight $\alpha_n = z/n$, where z is Bühlmann's credibility factor.

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

Example (Ex.20.9 cont'd)

$$\mu_3(G) = 0.4 \qquad \mu_3(B) = 0.7$$
$$E[X_3|0,1] = 0.478948 \qquad \mu_3 = 0.475 \quad \bar{X} = 0.5$$
$$a = V[\mu(\theta)] = 0.016875 \qquad v = E[v(\theta)] = 0.4825$$
$$k = v/a = 28.5926 \qquad z = 2(2+k)^{-1} = 0.0654$$
$$z\bar{X} + (1-z)\mu = 0.0654(0.5) + 0.9346(0.475) = 0.4766$$

Example (Ex. 20.10. Exact credibility example)

$$\begin{aligned} E(X_4|100,950,450) &= 416,67; \quad \bar{X} = 500 \\ \mu = E(X_4) = E(1/\Theta) &= 1000/3 = 333.3(3) \\ z\bar{X} + (1-z)\mu &= E(X_4|100,950,450). \end{aligned}$$

Exercises 20.24-27, p. 606.

Bühlmann-Straub's model

Bühlmann's H1 is changed:

1. Given θ, $X_1|\theta, X_2|\theta, \ldots, X_n|\theta, X_{n+1}|\theta$ have the same mean, variance:

$$\begin{aligned} E\left(X_j|\theta\right) &= \mu(\theta)\ (same) \\ Var\left(X_j|\theta\right) &= \frac{v(\theta)}{m_j}. \end{aligned}$$

- m_j is some known constant measuring exposure
- Ex: group insurance where its size changes
- Initially, the model was first presented for reinsurance.
- $Var\left(X_j\right) = E\left[Var\left(X_j|\theta\right)\right] + Var\left[E\left(X_j|\theta\right)\right] = \frac{v}{m_j} + a$

Solution:

$$P_c = \widetilde{\alpha}_0 + \sum_{j=1}^{n} \widetilde{\alpha}_j X_j = z\bar{X} + (1-z)\mu$$

$$z = \frac{m}{m+k} \qquad k = \upsilon / a$$

$$\bar{X} = \sum_{j=1}^{n} \frac{m_j}{m} X_j \qquad m = \sum_{j=1}^{n} m_j \text{ (total exposure)}$$

Obs.:

- Factor z depends on m (total exposure)
- $\bar{X}$ is a weighted average, m_j / m is the weight
- $m_j X_j$ is the total loss of the group in year j
- (Total) Credibility premium for the group, next year:

$$m_{n+1}\left[z\bar{X} + (1-z)\mu\right]$$

Example (Ex.20.19)

N_j: No. of claims in year j for a group policy holder with risk parameter and m_j individuals. $N_j \frown Poisson(m_j\theta)$. Let $X_j = N_j / m_j$. $\Theta \frown Gamma(\alpha, \beta)$.

$$\begin{aligned}
\mathbb{E}(X_j|\theta) &= \mu(\theta) = \theta;\ \mathbb{V}(X_j|\theta) = \mathbb{V}(N_j/m_j|\theta) = \frac{v(\theta)}{m_j} = \frac{\theta}{m_j} \\
\mu &= \mathbb{E}(\Theta) = \alpha\beta;\ a = \mathbb{V}(\Theta) = \alpha\beta^2;\ v = \mathbb{E}(\Theta) = \alpha\beta. \\
k &= v/a = 1/\beta;\ z = \frac{m\beta}{m\beta + 1} \\
P_c &= \frac{m\beta}{m\beta + 1}\bar{X} + \frac{1}{m\beta + 1}\alpha\beta
\end{aligned}$$

Example (Ex.20.19)

N_j: No. of claims in year j for a group policy holder with risk parameter θ and m_j individuals, $j = 1, ..., n$. $N_j \frown Poisson(m_j\theta)$. Let $X_j = N_j/m_j$. $\Theta \frown Gamma(\alpha, \beta)$. Bayesian premium (mean of the preditive dist.):

$$\begin{aligned}
\mathbb{E}(X_{n+1}|\mathbf{X}) &= \mathbb{E}(\mathbb{E}(X_{n+1}(\theta)|\theta, \mathbf{X})) = \mathbb{E}(\mu_{n+1}(\theta)|\mathbf{X}) \\
&= \mathbb{E}(\theta|\mathbf{X})
\end{aligned}$$

$$\begin{aligned}
\Pr\left[N_j = n|\theta\right] &= \Pr\left[X_j m_j = n|\theta\right] = \Pr\left[X_j = n/m_j|\theta\right], \; n \in \mathbb{N}_0 \\
&= (m_j\theta)^n e^{-m_j\theta}/n!; \; \pi(\theta) = \frac{\theta^{\alpha-1}e^{-\theta/\beta}}{\Gamma(\alpha)\beta^\alpha} \\
\pi_{\Theta|\mathbf{X}}(\theta|\mathbf{x}) &\propto \left[\prod_{i=1}^n f_{X_j|\theta}(x_j|\theta)\right]\pi(\theta); \\
& \qquad f_{X_j|\theta}(x_j|\theta) = \Pr\left[X_j = x|\theta\right]
\end{aligned}$$

Example (Ex.20.19)

N_j: No. of claims in year j for a group policy holder with risk parameter and m_j individuals, $j = 1, ..., n$. $N_j \frown Poisson(m_j\theta)$. Let $X_j = N_j/m_j$. $\Theta \frown Gamma(\alpha, \beta)$.

$$\Theta|\mathbf{x} \frown Gamma\left(\alpha_* = \alpha + \sum_{j=1}^{n} m_j x_j; \beta_* = (1/\beta + m)^{-1}\right)$$

$$\begin{aligned} \mathbb{E}(X_{n+1}|\mathbf{X}=\mathbf{x}) &= \alpha_*\beta_* = \frac{\alpha + \sum_{j=1}^{n} m_j x_j}{(1/\beta + m)} \\ &= \frac{m\beta}{m\beta+1}\bar{X} + \frac{1}{m\beta+1}\alpha\beta = P_c \end{aligned}$$

Exercises 20.28, 29, p. 608

- *Recap* Credibility Premium,

$$\widetilde{\mu_{n+1}}(\theta)\text{: } \min\left\{ Q = E\left\{\left[\mu_{n+1}(\theta) - \left(\alpha_0 + \sum_{j=1}^{n}\alpha_j X_j\right)\right]^2\right\}\right\}.$$

- Now, don't impose a linear estimator. Let $m(\mathbf{X})$, some function of $\mathbf{X}$, and find estimator $\overset{*}{m}(\mathbf{X})$ such that:

$$\min\left(E\left\{[\mu_{n+1}(\theta) - m(\mathbf{X})]^2\right\} = E\left[E\left\{[\mu_{n+1}(\theta) - m(\mathbf{X})]^2 \,|\mathbf{X}\right\}\right]\right),$$

$$\text{or minimize } \; E\left\{[\mu_{n+1}(\theta) - m(\mathbf{X})]^2 \,|\mathbf{X}\right\} =$$

$$= \; V\left[\mu_{n+1}(\theta)|\mathbf{X}\right] + \left(E\left[\mu_{n+1}(\theta)|\mathbf{X}\right] - m(\mathbf{X})\right)^2$$

$$\Rightarrow \overset{*}{m}(\mathbf{X}) = E\left[\mu_{n+1}(\theta)|\mathbf{X}\right]$$

Bayes estimator, relative to Square Loss function and prior $\pi(\theta)$.

Exact Credibility: When $\widetilde{\mu_{n+1}}(\theta) = \overset{*}{m}(\mathbf{X}) = E\left[\mu_{n+1}(\theta)|\mathbf{X}\right]$, i.e., **Credibility Premium=Bayesian Premium**.

Stronger Bühlmann's H1

Change Bühlmann's **H1**, in addition, to:

H1: $f_{X_j}(.|\theta) = f_X(.|\theta)\,, \forall j = 1, \ldots, n, n+1.$

$$\begin{aligned}
\mathrm{E}[\mu(\theta)|\mathbf{X}] &= \int \mu(\theta)\pi(\theta|\mathbf{x})d\theta = \int \mu(\theta)\frac{f(\theta,\mathbf{x})}{f(\mathbf{x})}d\theta \\
&= \int \mu(\theta)\frac{f(\mathbf{x}|\theta)\pi(\theta)}{\int f(\mathbf{x}|\theta)\pi(\theta)}d\theta = \frac{\int \mu(\theta)\prod_{j=1}^n f(x_j|\theta)\pi(\theta)d\theta}{\int_\Theta \prod_{j=1}^n f(x_j|\theta)\pi(\theta)d\theta} \\
&= \frac{\int \mu(\theta)L(\theta)\pi(\theta)d\theta}{\int_\Theta L(\theta)\pi(\theta)d\theta}; \\
\pi(\theta|\mathbf{x}) &= \frac{L(\theta)\pi(\theta)}{\int_\Theta L(\theta)\pi(\theta)d\theta}
\end{aligned}$$

Example (Norberg [1979])

For a given risk $X|\theta \frown Bin(1;\theta)$, $\Theta \frown U(\alpha,\beta)$, obs'd for 10 yrs, 20 risks. $\bar{X} = 0.0145$, $\mu_{n+1}(\theta) = \mu(\theta) = \theta$.

$$f(x|\theta) = \theta^x (1-\theta)^{1-x}, \quad x = 0,\ 1; \quad 0 < \theta < 1.$$

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

$$\overset{*}{m}(\mathbf{x}) = \mathrm{E}[\theta|\mathbf{x}] = \frac{\sum_{k=1}^{n-n\bar{x}} (-1)^k \frac{\beta^{n\bar{x}+k+2} - \alpha^{n\bar{x}+k+2}}{(n-n\bar{x}-k)!k!(n\bar{x}+k+2)}}{\sum_{k=1}^{n-n\bar{x}} (-1)^k \frac{\beta^{n\bar{x}+k+1} - \alpha^{n\bar{x}+k+1}}{(n-n\bar{x}-k)!k!(n\bar{x}+k+1)}},$$

Example (*Beta-Binomial* model)

Let risk $X|\theta \frown Bin(1;\theta)$, $\Theta \frown Beta(\alpha,\beta)$, $\alpha,\ \beta > 0$, $\bar{X} = 1.45$.

$$\pi(\theta) = \frac{\theta^{\alpha-1}(1-\theta)^{\beta-1}}{B(\alpha,\beta)};\ \theta\epsilon(0;1);$$

$$B(\alpha,\beta) = \int_0^1 x^{\alpha-1}(1-x)^{\beta-1}dx$$

$$L(\theta) = \prod_{j=1}^{n} f(x_j|\theta) = \theta^{\sum_{j=1}^n x_j}(1-\theta)^{n-\sum_{j=1}^n x_j};$$

$$\pi(\theta|\mathbf{x}) = \frac{L(\theta)\pi(\theta)}{\int_0^1 L(\theta)\pi(\theta)d\theta} = \frac{\theta^{\sum_j x_j+\alpha-1}(1-\theta)^{n+\beta-\sum_j x_j-1}}{B(\sum_j x_j+\alpha;\ n+\alpha-\sum_j x_j)},$$

$$\pi(\theta|\mathbf{x}) \equiv Beta(\sum_j x_j+\alpha;\ n+\beta-\sum_j x_j)$$

$$\mathrm{E}[\theta|\mathbf{x}] = \frac{\sum_j x_j+\alpha}{\alpha+\beta+n} = \frac{n}{\alpha+\beta+n}\bar{x} + \frac{\alpha+\beta}{\alpha+\beta+n}\mu.$$

Example (Gamma-exponential model)

$X|\theta \sim \text{Exp}(\theta), \mu(\theta) = 1/\theta,\ f(x|\theta) = \theta e^{-\theta x}, x > 0;$
$\Theta \frown Gamma(\alpha, \beta = 1/\beta^*),$

$$\begin{aligned}
\pi(\theta) &= \frac{\beta^\alpha}{\Gamma(\alpha)} e^{-\beta\theta} \theta^{\alpha-1};\ \theta > 0; \\
L(\theta) &= \prod_{j=1}^{n} f(x_j|\theta) = \theta^n \exp\{-\theta \sum x_j\}; \\
\pi(\theta|\mathbf{x}) &= \frac{L(\theta)\pi(\theta)}{\int_0^\infty L(\theta)\pi(\theta)d\theta} \\
&= \frac{(\beta + \sum_j x_j)^{n+\alpha}}{\Gamma(n+\alpha)} \exp\{-\theta(\beta + \sum_j x_j)\}\theta^{n+\alpha-1}, \\
\pi(\theta|\mathbf{x}) &\equiv \text{Gama}(n+\alpha;\ \beta + \sum_j x_j);\ \ \mu = \mathrm{E}[X_{ij}] = \mathrm{E}[1/\theta]
\end{aligned}$$

Example (*Gamma-exponential* model cont'd)

$$\begin{aligned}
\mu &= \frac{\beta^\alpha}{\Gamma(\alpha)}\int_0^{+\infty} e^{-\beta\theta}\theta^{\alpha-2}d\theta = \beta\frac{\Gamma(\alpha-1)}{\Gamma(\alpha)} = \frac{\beta}{\alpha-1} \\
\mathrm{E}[1/\theta|\mathbf{x}] &= \frac{(\beta+\sum_{j=1}^n x_j)^{n+\alpha}}{\Gamma(n+\alpha)}\int_0^{+\infty} e^{-(\beta+\sum_j x_j)\theta}\theta^{n+\alpha-2}d\theta \\
&= \frac{(\beta+\sum_j x_j)\Gamma(n+\alpha-1)}{\Gamma(n+\alpha)} = \frac{\beta+\sum_j x_j}{n+\alpha-1} \\
&= \frac{n}{n+\alpha-1}\bar{x}_{.j} + \frac{\alpha-1}{n+\alpha-1}\mu
\end{aligned}$$

Bühlmann's Empirical Bayes.. Unbiased and consistent estimators.

$$\begin{aligned} \mu &= \mathrm{E}[X] = \mathrm{E}[\mathrm{E}[X|\theta]] = \mathrm{E}[\mu(\theta)]. \\ \hat{\mu} &= \bar{X} = \frac{1}{r}\sum_{i=1}^{r}\bar{X}_i = \frac{1}{nr}\sum_{i=1}^{r}\sum_{j=1}^{n}X_{ij} \end{aligned}$$

$$\begin{aligned} \mathrm{V}[X] &= \mathrm{V}[\mu(\theta)] + \mathrm{E}[v(\theta)] = a + v \\ \mathrm{V}[\bar{X}_i] &= a + \frac{1}{n}v \\ \hat{v} &= \frac{1}{r}\sum_{i=1}^{r}S_i'^2 = \frac{1}{r}\sum_{i=1}^{r}\sum_{j=1}^{n}\frac{(X_{ij}-\bar{X}_i)^2}{n-1} \\ \hat{a} &= \max\left\{\frac{1}{r-1}\sum_{i=1}^{r}(\bar{X}_i-\bar{X})^2 - \frac{1}{n}\hat{v}\,;\ 0\right\}. \end{aligned}$$

Bühlmann-Straub's Empirical Bayes.

$$
\begin{aligned}
\hat{\mu} &= \bar{X} = \frac{1}{m}\sum_{i=1}^{r} m_i \bar{X}_i = \frac{1}{m}\sum_{i=1}^{r}\sum_{j=1}^{n_i} m_{ij} X_{ij} \\
m &= \sum_{i=1}^{r} m_i = \sum_{i=1}^{r}\sum_{j=1}^{n_i} m_{ij}; \qquad \hat{\mu} = \frac{\sum_{i=1}^{r} \hat{Z}_i \bar{X}_i}{\sum_{i=1}^{r} \hat{Z}_i}
\end{aligned}
$$

$$
\begin{aligned}
\hat{\upsilon} &= \frac{\sum_{i=1}^{r}\sum_{j=1}^{n_i} m_{ij}\left(X_{ij} - \bar{X}_i\right)^2}{\sum_{i=1}^{r}(n_i - 1)} \\
\hat{a} &= \max\left\{\left(m - m^{-1}\sum_{i=1}^{r} m_i^2\right)^{-1}\left[\sum_{i=1}^{r} m_i\left(\bar{X}_i - \bar{X}\right)^2 - \hat{\upsilon}\,(r-1)\right]; 0\right\}
\end{aligned}
$$

Example (A Bonus-Malus system)

Let X_j: claims in year j, $X_j \frown Poisson(\theta)$, $\mu(\theta) = v(\theta) = \theta$

$$\tilde{\theta} = \frac{n}{n + \mathrm{E}[\theta]/\mathrm{V}[\theta]}\bar{X} + \frac{\mathrm{E}[\theta]/\mathrm{V}[\theta]}{n + \mathrm{E}[\theta]/\mathrm{V}[\theta]}\mathrm{E}[\theta]$$

Data: Portfolio of 106974 policies in one year (stable period):

x	0	1	2	3	4	≥ 5
n_x	96 978	9 240	704	43	9	0

- $\hat{E}[\theta] = \hat{E}[X] = \bar{X} = (1/106974)\sum_{k=0}^{4} x_k n_{x_k} = 0.1011.$
- $\hat{V}[X] = s^2 = (1/106974)\sum_{k=0}^{4} {x_k}^2 n_{x_k} - \bar{x}^2 = 0.1074.$
- $\mathrm{V}[X] = \mathrm{E}[\theta] + \mathrm{V}[\theta]$. $\hat{V}[\theta] = 0.1074 - 0.1011 = 0.0063.$

Example (A Bonus-Malus system cont'd)

$P^*_{n+1}(\mathbf{X}_i)$: $100\times$ *Risk premium/Collective premium*

$$\begin{aligned}\tilde{\theta} &= \frac{n}{n+0.1011/0,0063}\bar{X} + \frac{0.1011/0.0063}{n+0.1011/0.0063}\times 0.1011\\ &= \left(\sum_{j=1}^{n} x_j + 16,047\,(0.1011)\right) / \,(n+16.0476)\end{aligned}$$

$$P^*_{n+1}(\mathbf{X}_i) = 100\times\frac{\sum_{j=1}^{n} X_{ij}+1.6224}{0.1011(n+16.0476)} = 100\times\frac{\sum_{i=1}^{n} X_{ij}+1.6224}{0.1011\,n+1.6224}$$

	No. of claims				
no years	0	1	2	3	4
0	100	-	-	-	-
1	94,13	152,16	210,18	268,20	326,22
2	88,92	143,72	198,53	253,34	308,14
3	84,25	136,18	188,11	240,04	291,97
4	80,05	129,39	178,73	228,06	277,40
5	76,24	123,24	170,23	217,23	264,22
6	72,79	117,65	162,51	207,38	252,24
7	69,63	112,54	155,46	198,38	241,29
8	66,73	107,86	149,00	190,13	231,26
9	64,07	103,56	143,05	182,54	222,03
10	61,61	99,58	137,56	175,53	213,50

Table: Relative premium for a *Bonus-malus* system

Example (Life group insurance)

N_{ksij}: No. people dying, with ins. capital x_k, age s, group j, year i.
$N_{ij} = \sum_{k,s} N_{ksij}$ - ...in group j year i
x_k: insured capital
q_s: mortality rate, age s, known.
$q_s\theta_j$: mortality, age s, group j (unknown)
n_{ksij}: No. people group j, capital x_k, age s, year i.
$S_{ij} = \sum_k \left(x_k \sum_s N_{ksij}\right)$: aggregate claims, group j, year i

$$\begin{aligned} N_{ksij}|\theta &\frown \text{Poisson}(n_{ksij} \times q_s \times \theta_j) \Rightarrow \\ \sum_s N_{ksij}|\theta &\sim \text{Poisson}\left(\theta_j \sum_s q_s n_{ksij}|\theta_j\right) \end{aligned}$$

Example (Life group insurance, cont'd)

$$S_{ij}|\theta = \sum_k \left(x_k \sum_s N_{ksij} \right)$$

$$S_{ij}|\theta \frown \text{CPoisson}\left(\theta_j \sum_{k,s} n_{ksij} q_s;\; f_{ij}(x) = \frac{\sum_s q_s n_{ksij}}{\sum_{k,s} q_s n_{ksij}} \right)$$

$$\mathrm{E}[S_{n+1,j}|\theta_j] = \sum_k x_k \sum_s \mathrm{E}[N_{ks(n+1)j}|\theta_j] = \theta_j \sum_{k,s} x_k q_s n_{ks(n+1)j}$$

$$P_c = \tilde{\theta}_j \sum_{k,s} x_k q_s n_{ks(n+1)j},$$

$$\tilde{\theta}_j = \frac{m_j}{m_j + \mathrm{E}[\theta_j]/\mathrm{V}[\theta_j]} \bar{X}_{\cdot j} + \frac{\mathrm{E}[\theta_j]/\mathrm{V}[\theta_j]}{m_j + \mathrm{E}[\theta_j]/\mathrm{V}[\theta_j]} \mathrm{E}[\theta_j]$$

Example (Life group insurance, cont'd)

$$
\begin{aligned}
\mathrm{E}[S_{n+1,j}|\theta_j] &= \sum_k x_k \sum_s \mathrm{E}[N_{ks(n+1)j}|\theta_j] = \theta_j \sum_{k,s} x_k q_s n_{ks(n+1)j} \\
P_c &= \tilde{\theta}_j \sum_{k,s} x_k q_s n_{ks(n+1)j},
\end{aligned}
$$

$$
\begin{aligned}
\tilde{\theta}_j &= \frac{m_j}{m_j + \mathrm{E}[\theta_j]/\mathrm{V}[\theta_j]} \bar{X}_{.j} + \frac{\mathrm{E}[\theta_j]/\mathrm{V}[\theta_j]}{m_j + \mathrm{E}[\theta_j]/\mathrm{V}[\theta_j]} \mathrm{E}[\theta_j] \\
X_{ij} &= N_{ij}/m_{ij}; \quad m_{ij} = \sum_{k,s} q_s n_{ksij}
\end{aligned}
$$

Problem 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 $X_1, X_2, \ldots$, 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.

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

1. Calculate the mean and variance of X.
2. Compute the probability function of X.

Problem 1 (cont'd)

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

3. 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.
4. Compute $Cov[X_1, X_2]$. [Note: For r.v.'s X, Y and Z, $Cov[X, Y] = E[Cov[X, Y|Z]] + Cov[E[X|Z]; E[Y|Z]]$]
5. Compute the posterior probability function of Θ given $(X_1 = 0, X_2 = 2)$.
6. 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.

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

7. Compute the collective pure premium.
8. Compute the Bayes premium $E[X_3|X = (0, 2)] = E(\mu(\Theta)|X = (0, 2))$.
9. Compute Bühlmann's credibility premium, say, $\tilde{E}(X_3|\theta)$.
10. Can we talk here on Exact Credibility? Comment appropriately.

Ratemaking and Experience Rating concepts, *Recap...*

Ratemaking portfolios/groups:

- Similar risks grouping in collectives of risks for ratemaking.

Tariff:

- Set of premia, for each risk in a (homogeneous) portfolio. A basic premium plus a system of *bonus* or *malus*.

Tariff structure:

- System of bonus/malus applied to a basic premium.

"**Prior**" **and** "**Posterior**" **ratemaking:**

- First rate following given *prior* variables, then make a *posterior re-evaluation/readjustment*, according to the reported accidents/claims by the risk/policy.

Bonus-malus systems, use of **GLM**'s, ..

- Bonus systems are in general based on **claim counts**, not amounts. This is explained by the usual assumption of independence between **number** and **severity** of claims. The base model is Markovian.

Bonus-malus (or bonus) systems

- Common tariff in motor insurance;
- Usually based on a counting variable, not the amounts
- A Markov chain model (discrete time) is often used:
- Basic idea:
    - year(s) with no claim: *bonus*
    - year with 1 claims: *malus;* 2 claims: + **malus**...
- Study *Long Term* behaviour

Bonus-Malus Systems

- A priori classification variables: age, sex, type and use of car, territory
- A posteriori variables: deductibles, credibility, bonus-malus
- Bonus malus:
  - Answer to heterogeneity of behavior of drivers in each cell
  - Answer to adverse selection
  - Inducement to drive more carefully
- Strongly influenced by regulatory environment and culture

BMS as they should be: **Bayesian analysis**

Example (Lemaire (1995, p. 37))

Observed distribution of third-party liability motor insurance claims

Mean: $\bar{x} = 0.1011$

Variance: $s^2 = 0.1074$

Number of claims	Observed policies
0	96,978
1	9,240
2	704
3	43
4	9
5+	0
Total	106,974

Example (cont'd)

Non-contagious model: Poisson fit

Number of claims	Observed policies	Poisson fit
0	96,978	96,689.6
1	9,240	9,773.5
2	704	493.9
3	43	16.6
4	9	0.4
5+	0	0
Total	106,974	106,974

Contagious model: Negative Binomial fit

Example (cont'd)

Number of claims	Observed policies	Poisson fit	Negative Binomial fit
0	96,978	96,689.6	96,985.5
1	9,240	9,773.5	9,222.5
2	704	493.9	711.7
3	43	16.6	50.7
4	9	0.4	3.6
5+	0	0	0
Total	106,974	106,974	106,974

Example (Lemaire 91995, p. 44; Deaths by horse kicks in the ten corps of the Prussian Army, 1875-1894)

N	Observed	Poisson	Neg Bin
0	109	108.67	111.99
1	65	66.29	61.80
2	22	20.22	20.00
3	3	4.11	4.95
4	1	0.72	1.04
5+	0	0.00	0.22
Total – Chi-Square	200	0.33	1.24

Shows total randomness of accidents. Clearly, Poisson fit is better.

Example (Optimal BMS with Negative Binomial model)

Year	Claims				
	0	1	2	3	4
0	100				
1	94	153	211	269	329
2	89	144	199	255	310
3	84	137	189	241	294
4	80	130	179	229	279
5	76	123	171	218	266
6	73	118	163	208	253
7	69	113	156	199	242

Link with Credibility theory, **Credibility idea**:

$$Premium \quad = \quad (1-z)(\text{Population } Pr.) + z(\text{Individual } Pr.)$$

Credibility is an exact rating formula for the Poisson-Gamma mix

- This **optimal BMS** is:
  - Fair (as it results from the application of Bayes theorem)
  - Financially balanced (the average income of the insurer stays at 100, year after year)
- BUT, It is not acceptable to regulators and managers, as the harsh penalties:
  - Encourage uninsured driving
  - Suggest *hit-and-run* behavior
  - Induce policyholders to leave the company after one accident

$\Rightarrow$ In practice, another approach, based on Markov Chains, is used

BMS as they are: definition of Markov Chain (MC) $\{Z_n\}$ is a discrete-time, non-homogeneous Markov Chain when Z is an infinite sequence of random variables $Z_0, Z_1, \ldots$ such that:

1. Z_n denotes the state at time n, $n = 0, 1, 2, \ldots$
2. Each Z_n is a discrete random variable that can take s values (s is the number of states)
3. All transition probabilities are history-independent:

$$\begin{aligned} P_{(n)}(i,j) &= Pr[z_{n+1} = j \mid z_n = i, Z_{n-1} = i_{n-1}, \ldots, , M_0 = i_0] \\ &= Pr[z_{n+1} = j \mid Z_n = i] \end{aligned}$$

For all BMS applications, MC are homogeneous: $\mathbf{P}_n = \mathbf{P}$. We can have MC of order higher than 1. See Next example

Example (Centeno [2003])

A *Bonus* system in motor insurance, 3rd party liability (directly, the system is not pure Markovian, Markov of Order 2)

- 30% discount, no claim for 2 yrs.
- 15% malus, 1 claim
- 30% malus, 2 claims
- 45% malus, 3 claims
- 100% malus, 4 claims
- $>$ 4, case by case...

Markovian, if classes are split (see later)

Classical "Markovian" BMS consider (long term) stable behaviour. See next examples

Example (Markov chain, T&K, p.102, Ex. 2.2)

A particle travels through states $\{0, 1, 2\}$ according to a Markov chain

$$P = \begin{array}{c} \\ 0 \\ 1 \\ 2 \end{array} \begin{array}{c} \begin{array}{ccc} 0 & 1 & 2 \end{array} \\ \begin{pmatrix} 0 & 1/2 & 1/2 \\ 1/2 & 0 & 1/2 \\ 1/2 & 1/2 & 0 \end{pmatrix} \end{array}$$

$$P^2 = \begin{bmatrix} \frac{1}{2} & \frac{1}{4} & \frac{1}{4} \\ \frac{1}{4} & \frac{1}{2} & \frac{1}{4} \\ \frac{1}{4} & \frac{1}{4} & \frac{1}{2} \end{bmatrix}; P^3 = \begin{bmatrix} \frac{1}{4} & \frac{3}{8} & \frac{3}{8} \\ \frac{3}{8} & \frac{1}{4} & \frac{3}{8} \\ \frac{3}{8} & \frac{3}{8} & \frac{1}{4} \end{bmatrix}; P^4 = \begin{bmatrix} \frac{3}{8} & \frac{5}{16} & \frac{5}{16} \\ \frac{5}{16} & \frac{3}{8} & \frac{5}{16} \\ \frac{5}{16} & \frac{5}{16} & \frac{3}{8} \end{bmatrix}$$

$$P^5 = \begin{bmatrix} \frac{5}{16} & \frac{11}{32} & \frac{11}{32} \\ \frac{11}{32} & \frac{5}{16} & \frac{11}{32} \\ \frac{11}{32} & \frac{11}{32} & \frac{5}{16} \end{bmatrix}; P^{10} = \begin{bmatrix} \frac{171}{512} & \frac{341}{1024} & \frac{341}{1024} \\ \frac{341}{1024} & \frac{171}{512} & \frac{341}{1024} \\ \frac{341}{1024} & \frac{341}{1024} & \frac{171}{512} \end{bmatrix}$$

Example (cont.)

$$P^{100} =$$

$$\begin{bmatrix} \frac{211\,275\,100\,038\,038\,233\,582\,783\,867\,563}{633\,825\,300\,114\,114\,700\,748\,351\,602\,688} & \frac{422\,550\,200\,076\,076\,467\,165\,567\,735\,125}{1267\,650\,600\,228\,229\,401\,496\,703\,205\,376} & P_{31} \\ \frac{422\,550\,200\,076\,076\,467\,165\,567\,735\,125}{1267\,650\,600\,228\,229\,401\,496\,703\,205\,376} & \frac{211\,275\,100\,038\,038\,233\,582\,783\,867\,563}{633\,825\,300\,114\,114\,700\,748\,351\,602\,688} & P_{21} \\ \frac{422\,550\,200\,076\,076\,467\,165\,567\,735\,125}{1267\,650\,600\,228\,229\,401\,496\,703\,205\,376} & \frac{422\,550\,200\,076\,076\,467\,165\,567\,735\,125}{1267\,650\,600\,228\,229\,401\,496\,703\,205\,376} & P_{22} \end{bmatrix}$$

$$\approx \begin{bmatrix} 0.33333 & 0.33333 & 0.33333 \\ 0.33333 & 0.33333 & 0.33333 \\ 0.33333 & 0.33333 & 0.33333 \end{bmatrix}$$

Example

Let a Markov chain with transition matrix:

$$P = \begin{array}{c} \\ 0 \\ 1 \\ 2 \\ 3 \\ 4 \\ 5 \\ 6 \end{array} \begin{array}{c} \begin{array}{ccccccc} 0 & 1 & 2 & 3 & 4 & 5 & 6 \end{array} \\ \begin{pmatrix} 0.9 & 0.1 & 0 & 0 & 0 & 0 & 0 \\ 0.9 & 0 & 0.1 & 0 & 0 & 0 & 0 \\ 0.9 & 0 & 0 & 0.1 & 0 & 0 & 0 \\ 0.9 & 0 & 0 & 0 & 0.1 & 0 & 0 \\ 0.9 & 0 & 0 & 0 & 0 & 0.1 & 0 \\ 0.9 & 0 & 0 & 0 & 0 & 0 & 0.1 \\ 0.9 & 0 & 0 & 0 & 0 & 0 & 0.1 \end{pmatrix} \end{array}$$

Example

Long term: $P^8 =$

$$\begin{bmatrix} .9 & .09 & .009 & .0009 & .00009 & 9.0\times10^{-6} & 1.0\times10^{-6} \\ .9 & .09 & .009 & .0009 & .00009 & 9.0\times10^{-6} & 1.0\times10^{-6} \\ .9 & .09 & .009 & .0009 & .00009 & 9.0\times10^{-6} & 1.0\times10^{-6} \\ .9 & .09 & .009 & .0009 & .00009 & 9.0\times10^{-6} & 1.0\times10^{-6} \\ .9 & .09 & .009 & .0009 & .00009 & 9.0\times10^{-6} & 1.0\times10^{-6} \\ .9 & .09 & .009 & .0009 & .00009 & 9.0\times10^{-6} & 1.0\times10^{-6} \\ .9 & .09 & .009 & .0009 & .00009 & 9.0\times10^{-6} & 1.0\times10^{-6} \end{bmatrix}$$

Example (Entry class: 5.)

Table 4.1 Transition rules for the scale -1/top.

Starting level	Level occupied if 0 claim is reported	Level occupied if ≥ 1 claim is reported
0	0	5
1	0	5
2	1	5
3	2	5
4	3	5
5	4	5

Table 4.2 Transition rules for the scale $-1/+2$.

Starting level	Level occupied if 0 claim(s) is/are reported	1	2	≥ 3
5	4	5	5	5
4	3	5	5	5
3	2	5	5	5
2	1	4	5	5
1	0	3	5	5
0	0	2	4	5

A posterior ratemaking system, experience rating, is a *Bonus-malus* sytem if

- The rating periods are equal (1 year)
- The risks, policies, are divided into (finite) classes:

$$C_1, C_2, \dots, C_s; \quad \cup_i C_i = C; \quad C_i \cap C_j = \varnothing.$$

- No transitions within the year
- Position in Class in the year n depends on:
  - Position in $n-1$, and
  - The year claim counts.

Composition of the B-S system:

1. A vector of *premia* (or multiplying factor, index)

$$\mathbf{b} = (b(1), b(2), ..., b(s))$$

2. Transition rules among classes, in matrix:

$$\mathbf{T} \quad = \quad [T_{ij}]\text{, each entry } T_{ij} \text{ is a set of integers...}$$

$$\mathbf{T} \quad : \quad \cup_{j=1}^{s} T_{ij} = \{0, 1, 2, ...\}\,,\ T_{ij} \cap T_{ij'} = \varnothing,\, j \neq j'$$

3. Entry class, C_{i_0} is the same for all policies.

If k claims are reported

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

The $t_{ij}(k)$s are put in matrix form $\boldsymbol{T}(k)$, i.e.

$$\boldsymbol{T}(k) = \begin{pmatrix} t_{00}(k) & t_{01}(k) & \cdots & t_{0s}(k) \\ t_{10}(k) & t_{11}(k) & \cdots & t_{1s}(k) \\ \vdots & \vdots & \ddots & \vdots \\ t_{s0}(k) & t_{s1}(k) & \cdots & t_{ss}(k) \end{pmatrix}$$

Example 4.3 (−1/Top Scale) In this case, we have

$$\boldsymbol{T}(0) = \begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \end{pmatrix}, \quad \boldsymbol{T}(1) = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

and $\boldsymbol{T}(k) = \boldsymbol{T}(1)$ for all $k \geq 2$.

Example 4.4 (−1/+2 Scale) In this case, we have

$$\boldsymbol{T}(0)=\begin{pmatrix} 1 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \end{pmatrix}, \quad \boldsymbol{T}(1)=\begin{pmatrix} 0 & 0 & 1 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix}$$

$$\boldsymbol{T}(2)=\begin{pmatrix} 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} \text{ and } \boldsymbol{T}(k)=\begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 & 0 & 1 \end{pmatrix} \text{ for all } k \geq 3.$$

- Symbolically, a B-M S can be written as a triplet: $\Delta = (C_{i_0}, \mathbf{T}, \mathbf{b})$.
- Bonus Class in year n: $Z_{\Delta,n}$, defined by set of rules $\mathbf{T}$ and entry class C_{i_0}.
- The system is supposed to be a Markov chain

$$\{Z_{\Delta,n}, n = 0, 1, 2, \ldots\}$$

- Transition probability matrix: $P_T = [p_T(i,j)]$
- Transition rules is based on claim counts, often
  - Poisson distributed (usually bad), or
  - mixed Poisson (much better), $i, j = 1, 2, \ldots, s$,

$$\begin{aligned} p_T(i,j) &= \Pr\left(Z_{\Delta,n+1} = j | Z_{\Delta,n} = i\right) \\ p_T^{(n)}(i,j) &= \Pr\left(Z_{\Delta,n} = j | Z_{\Delta,0} = i\right) \\ p_T^{(n)}(j) &= \Pr\left(Z_{\Delta,n} = j\right) \end{aligned}$$

Further, $\boldsymbol{P}(\vartheta)$ is the one-step transition matrix, i.e.

$$\boldsymbol{P}(\vartheta) = \begin{pmatrix} p_{00}(\vartheta) & p_{01}(\vartheta) & \cdots & p_{0s}(\vartheta) \\ p_{10}(\vartheta) & p_{11}(\vartheta) & \cdots & p_{1s}(\vartheta) \\ \vdots & \vdots & \ddots & \vdots \\ p_{s0}(\vartheta) & p_{s1}(\vartheta) & \cdots & p_{ss}(\vartheta) \end{pmatrix}$$

$$\begin{aligned} p_{(i,j)}(\lambda) &= \sum_{k=0}^{\infty} p_k(\lambda)\, t_{ij}(k)\ , \ \ i,j = 1,\ldots,S\,, \\ \mathsf{P}_{T,\lambda} &= \left[p_{(i,j)}(\lambda)\right]_{S\times S} = \sum_{k=0}^{\infty} p_k(\lambda)\mathsf{T}_k \\ &= \sum_{k=0}^{\infty} e^{-\lambda}\frac{\lambda^k}{k!}\mathsf{T}_k\ . \ \text{(if Poisson)} \end{aligned}$$

Example 4.5 (−1/Top Scale) The transition matrix $\boldsymbol{P}(\vartheta)$ associated with this bonus-malus system is given by

$$\boldsymbol{P}(\vartheta)=\begin{pmatrix} \exp(-\vartheta) & 0 & 0 & 0 & 0 & 1-\exp(-\vartheta) \\ \exp(-\vartheta) & 0 & 0 & 0 & 0 & 1-\exp(-\vartheta) \\ 0 & \exp(-\vartheta) & 0 & 0 & 0 & 1-\exp(-\vartheta) \\ 0 & 0 & \exp(-\vartheta) & 0 & 0 & 1-\exp(-\vartheta) \\ 0 & 0 & 0 & \exp(-\vartheta) & 0 & 1-\exp(-\vartheta) \\ 0 & 0 & 0 & 0 & \exp(-\vartheta) & 1-\exp(-\vartheta) \end{pmatrix}.$$

Example 4.6 (−1/+2 Scale) The transition matrix $\boldsymbol{P}(\vartheta)$ associated with this bonus-malus system is given by

$$\boldsymbol{P}(\vartheta)=\begin{pmatrix} \exp(-\vartheta) & 0 & \vartheta\exp(-\vartheta) & 0 & \frac{\vartheta^2}{2}\exp(-\vartheta) & 1-\Sigma_1 \\ \exp(-\vartheta) & 0 & 0 & \vartheta\exp(-\vartheta) & 0 & 1-\Sigma_2 \\ 0 & \exp(-\vartheta) & 0 & 0 & \vartheta\exp(-\vartheta) & 1-\Sigma_3 \\ 0 & 0 & \exp(-\vartheta) & 0 & 0 & 1-\exp(-\vartheta) \\ 0 & 0 & 0 & \exp(-\vartheta) & 0 & 1-\exp(-\vartheta) \\ 0 & 0 & 0 & 0 & \exp(-\vartheta) & 1-\exp(-\vartheta) \end{pmatrix}.$$

where Σ_i represents the sum of the elements in columns 1 to 5 in row i, $i=1,2,3$, that is,

$$\Sigma_1=\exp(-\vartheta)\left(1+\vartheta+\frac{\vartheta^2}{2}\right)$$

Transition rules is based on claim counts, often

- Poisson distributed (usually bad), $i, j = 1, 2, ..., s$, $n = 0, 1, ...$

$$\begin{aligned} p_{T,\lambda}(i,j) &= \Pr\left(Z_{\Delta,n+1} = j | Z_{\Delta,n} = i, \Lambda = \lambda\right) \\ p^{(n)}_{T,\lambda}(i,j) &= \Pr\left(Z_{\Delta,n} = j | Z_{\Delta,0} = i, \Lambda = \lambda\right) \\ p^{(n)}_{T,\lambda}(j) &= \Pr\left(Z_{\Delta,n} = j | \Lambda = \lambda\right) \ . \end{aligned}$$

- Mixed Poisson (much better), 1st compute the conditional $p^{(n)}_{T,\lambda}(i,j)$, $i, j = 1, 2, ..., s$, then

$$\begin{aligned} p_T(i,j) &= \int_0^\infty p_{T,\lambda}(i,j) d\pi(\lambda) \\ p^{(n)}_T(i,j) &= \int_0^\infty p^{(n)}_{T,\lambda}(i,j) d\pi(\lambda) = E\left[p^{(n)}_{T,\lambda}(i,j)\right] \\ p^{(n)}_T(j) &= \int_0^\infty p^{(n)}_{T,\lambda}(j) d\pi(\lambda) = E\left[p^{(n)}_{T,\lambda}(j)\right] . \end{aligned}$$

Remark: neither $p^{(n)}_T(i,j)$ nor $p^{(n)}_T(j)$ are obtained from the initial mixed Poisson distribution.

- All B-M systems have (at least) a *bonus* class where a policy:
  - stays if keeps with no claims
  - goes, transits to, if has no claims
  - goes out, transits from (to another)
- That class is a periodic state
- If the Markov chain is irreducible, finite number of states, it will be aperiodic and stationary;
- Then, it exists a limit distribution, for a given λ

$$p_{T,\lambda}^{(\infty)}(j) = \lim_{n\uparrow\infty} p_{T,\lambda}^{(n)}(i,j).$$

If λ is considered to be the outcome of a r.v. with dist. $\pi(\lambda)$, usually

$$p_{T}^{(\infty)}(j) = \int_0^{\infty} p_{T,\lambda}^{(\infty)}(j) d\pi(\lambda) = E\left[p_{T,\lambda}^{(\infty)}(j)\right]$$

Remark: $p_{T}^{(\infty)}(j)$ is not got from the initial "mixed Poisson".

Example 4.7 (−1/Top Scale) Starting from

$$
\boldsymbol{P}(0.1) = \begin{pmatrix}
0.904837 & 0 & 0 & 0 & 0 & 0.095163 \\
0.904837 & 0 & 0 & 0 & 0 & 0.095163 \\
0 & 0.904837 & 0 & 0 & 0 & 0.095163 \\
0 & 0 & 0.904837 & 0 & 0 & 0.095163 \\
0 & 0 & 0 & 0.904837 & 0 & 0.095163 \\
0 & 0 & 0 & 0 & 0.904837 & 0.095163
\end{pmatrix}
$$

$$
\boldsymbol{P}^5(0.1) = \begin{pmatrix}
0.606531 & 0.063789 & 0.070498 & 0.077913 & 0.086107 & 0.095163 \\
0.606531 & 0.063789 & 0.070498 & 0.077913 & 0.086107 & 0.095163 \\
0.606531 & 0.063789 & 0.070498 & 0.077913 & 0.086107 & 0.095163 \\
0.606531 & 0.063789 & 0.070498 & 0.077913 & 0.086107 & 0.095163 \\
0.606531 & 0.063789 & 0.070498 & 0.077913 & 0.086107 & 0.095163 \\
0.606531 & 0.063789 & 0.070498 & 0.077913 & 0.086107 & 0.095163
\end{pmatrix}
$$

Example 4.8 (−1/+2 Scale) In this case,

$$
\boldsymbol{P}(0.1) = \begin{pmatrix}
0.904837 & 0 & 0.090484 & 0 & 0.004524 & 0.000155 \\
0.904837 & 0 & 0 & 0.090484 & 0 & 0.004679 \\
0 & 0.904837 & 0 & 0 & 0.090484 & 0.004679 \\
0 & 0 & 0.904837 & 0 & 0 & 0.095163 \\
0 & 0 & 0 & 0.904837 & 0 & 0.095163 \\
0 & 0 & 0 & 0 & 0.904837 & 0.095163
\end{pmatrix}
$$

$$
\boldsymbol{P}^{20}(0.1) = \begin{pmatrix}
0.782907 & 0.082338 & 0.090996 & 0.022276 & 0.016387 & 0.005096 \\
0.782903 & 0.082332 & 0.091006 & 0.022275 & 0.016387 & 0.005097 \\
0.782902 & 0.082326 & 0.090993 & 0.022295 & 0.016386 & 0.005098 \\
0.782803 & 0.082424 & 0.090984 & 0.022285 & 0.016406 & 0.005098 \\
0.782776 & 0.082352 & 0.091082 & 0.022278 & 0.016403 & 0.005108 \\
0.782774 & 0.082327 & 0.091011 & 0.022376 & 0.016399 & 0.005113
\end{pmatrix}
$$

which slowly converges to

$$
\boldsymbol{\Pi}(0.1) = \begin{pmatrix}
0.782901 & 0.082338 & 0.090998 & 0.022278 & 0.016387 & 0.005097 \\
0.782901 & 0.082338 & 0.090998 & 0.022278 & 0.016387 & 0.005097 \\
0.782901 & 0.082338 & 0.090998 & 0.022278 & 0.016387 & 0.005097 \\
0.782901 & 0.082338 & 0.090998 & 0.022278 & 0.016387 & 0.005097 \\
0.782901 & 0.082338 & 0.090998 & 0.022278 & 0.016387 & 0.005097 \\
0.782901 & 0.082338 & 0.090998 & 0.022278 & 0.016387 & 0.005097
\end{pmatrix}.
$$

Problem 2 (Problem 1 cont'd)

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 $X_1, X_2, \dots,$ 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.

Suppose 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.

Problem 2 (cont'd)

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 $b = (70, 100, 150)$.
– Consider a policy with risk parameter θ.

1. Write the transition rules matrix and compute the one year transition probability.
2. Comment on the existence of the of the stationary distribution.
3. Calculate the probability of a policy being ranked in Class 1 two years after entering the system.
4. Calculate the probability function of the premium for a type A driver after two years os stay in the portfolio. Compute the average premium.
5. 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, B^*. For a driver in this new class, compute the probability funcion of the premium after one year of staying in the system (since his entry).

– Stationary distr. for a given θ is given by vector $\left(\alpha(\theta)^2;\ [1-\alpha(\theta)]\,\alpha(\theta);\ 1-\alpha(\theta)\right)$.

6. Compute the probability function of the premium for a policy taken out at random from the portfolio. Calculate the average premium.

Example (Cont'd, Centeno [2003])

A *Bonus* system in motor insurance, 3rd party liability (directly, the system is not Markovian)

- 30% discount, no claim for 2 yrs.
- 15% malus, 1 claim
- 30% malus, 2 claims
- 45% malus, 3 claims
- 100% malus, 4 claims
- >4, case by case...

This is **not** Markovian, unless... Classes are split.

Example (Centeno [2003]. Class splitting:)

- C_1 Policies with 30% *bonus*
- C_2 Policies with neither *bonus* nor *malus* for the 2nd consecutive year
- C_3 Policies with neither *bonus* nor *malus* for the 1st yr
- C_4 Policies with 15% *penalty* and no claims last yr
- C_5 Policies with 15% *penalty* and claims last yr
- C_6 Policies with 30% *penalty* and no claims last yr
- C_7 Policies with 30% *penalty* and claims last yr
- C_8 Policies with 45% *penalty* and no claims last yr
- C_9 Policies with 45% *penalty* and claims last yr
- C_{10} Policies with 100% *penalty* and no claims last yr
- C_{11} Policies with 100% *penalty* and claims last yr.

Now is Markovian.

Example (Cont'd)

$$\mathbf{b} = (70, 100, 100, 115, 115, 130, 130, 145, 145, 200, 200)$$

$$\mathbf{T} =$$

$$\begin{array}{c|ccccccccccc|}
 & 1 & 2 & 3 & 4 & 5 & 6 & 7 & 8 & 9 & 10 & 11 \\
1 & \{0\} & & & & \{1\} & & \{2\} & & \{3\} & & \{4,\ldots\} \\
2 & \{0\} & & & & \{1\} & & \{2\} & & \{3\} & & \{4,\ldots\} \\
3 & & \{0\} & & & \{1\} & & \{2\} & & \{3\} & & \{4,\ldots\} \\
4 & \{0\} & & & & & & \{1\} & & \{2\} & & \{3,\ldots\} \\
5 & & & & \{0\} & & & \{1\} & & \{2\} & & \{3,\ldots\} \\
6 & \{0\} & & & & & & & & \{1\} & & \{2,\ldots\} \\
7 & & & & & & \{0\} & & & \{1\} & & \{2,\ldots\} \\
8 & \{0\} & & & & & & & & & & \{1,\ldots\} \\
9 & & & & & & & & \{0\} & & & \{1,\ldots\} \\
10 & \{0\} & & & & & & & & & & \{1,\ldots\} \\
11 & & & & & & & & & & \{0\} & \{1,\ldots\}
\end{array}$$

Example (cont'd)

Class j	b_j	New Class after step, with				
		0	1	2	3	4+
1	70	1	5	7	9	11
2	100	1	5	7	9	11
3	100	2	5	7	9	11
4	115	1	7	9	11	11
5	115	4	7	9	11	11
6	130	1	9	11	11	11
7	130	6	9	11	11	11
8	145	1	11	11	11	11
9	145	8	9	11	11	11
10	200	1	11	11	11	11
11	200	10	11	11	11	11

Example (cont'd)

If claim counts follow a Poisson(λ), $\mathbf{P}_{\Delta,\lambda}$:

	1	2	3	4	5	6	7	8	9	10	11
1	$e^{-\lambda}$				$\lambda e^{-\lambda}$		$\lambda^2 e^{-\lambda}/2$		$\lambda^3 e^{-\lambda}/6$		$1-e^{-\lambda}\sum_{i=0}^{3}\lambda^i/i!$
2	$e^{-\lambda}$				$\lambda e^{-\lambda}$		$\lambda^2 e^{-\lambda}/2$		$\lambda^3 e^{-\lambda}/6$		$1-e^{-\lambda}\sum_{i=0}^{3}\lambda^i/i!$
3		$e^{-\lambda}$			$\lambda e^{-\lambda}$		$\lambda^2 e^{-\lambda}/2$		$\lambda^3 e^{-\lambda}/6$		$1-e^{-\lambda}\sum_{i=0}^{3}\lambda^i/i!$
4	$e^{-\lambda}$						$\lambda e^{-\lambda}$		$\lambda^2 e^{-\lambda}/2$		$1-e^{-\lambda}\sum_{i=0}^{2}\lambda^i/i!$
5				$e^{-\lambda}$			$\lambda e^{-\lambda}$		$\lambda^2 e^{-\lambda}/2$		$1-e^{-\lambda}\sum_{i=0}^{2}\lambda^i/i!$
6	$e^{-\lambda}$								$\lambda e^{-\lambda}$		$1-e^{-\lambda}\sum_{i=0}^{1}\lambda^i/i!$
7	0					$e^{-\lambda}$			$\lambda e^{-\lambda}$		$1-e^{-\lambda}\sum_{i=0}^{1}\lambda^i/i!$
8	$e^{-\lambda}$										$1-e^{-\lambda}$
9								$e^{-\lambda}$			$1-e^{-\lambda}$
10	$e^{-\lambda}$										$1-e^{-\lambda}$
11										$e^{-\lambda}$	$1-e^{-\lambda}$

- The Markov chain is not irreducible.
- You cannot go to Class/State 3.
- Class of states $\{C_2, C_3\}$ is transient.
- Class, $\{C_1, C_4, C_5, C_6, C_7, C_8, C_9, C_{10}, C_{11}\}$ is a class of positive recurrent aperiodic states.

Re-order states in two classes of states:

- Class 1: $\{C_2, C_3\}$
- Class 2: $\{C_1, C_4, C_5, C_6, C_7, C_{8,} C_9, C_{10}, C_{11}\}$

So that $\mathbf{P}_{\Delta,\lambda}$ is split into 4 blocks:

$$\mathbf{P}_{\Delta,\lambda} = \begin{bmatrix} \mathbf{P}_{1,(\Delta,\lambda)} & \mathbf{P}_{3,(\Delta,\lambda)} \\ \mathbf{0} & \mathbf{P}_{2,\Delta,\lambda} \end{bmatrix}$$

- $\mathbf{P}_{1,\Delta,\lambda}$: Transition Prob'ty block inside Class 1, $\{C_2, C_3\}$;
- $\mathbf{P}_{3,\Delta,\lambda}$: Transition Prob'ty block between Class of states 1 & 2,

$$\{C_2, C_3\} \text{ and } \{C_1, C_4, C_5, C_6, C_7, C_{8,} C_9, C_{10}, C_{11}\}$$

- $\mathbf{P}_{2,\Delta,\lambda}$: Transition Prob'ty block among states $\{C_1, C_4, C_5, C_6, C_7, C_{8,} C_9, C_{10}, C_{11}\}$.

We have

$$P^2_{\Delta,\lambda} = \left[\begin{array}{c|c} P^2_{1,\Delta,\lambda} & P_{1,(\Delta,\lambda)}P_{3,(\Delta,\lambda)} + P_{3,(\Delta,\lambda)}P_{2,(\Delta,\lambda)} \\ - & - - - - - - - \\ 0 & P^2_{2,(\Delta,\lambda)} \end{array}\right]$$

$$= \left[\begin{array}{c|c} 0 & P_{1,(\Delta,\lambda)}P_{3,(\Delta,\lambda)} + P_{3,(\Delta,\lambda)}P_{2,(\Delta,\lambda)} \\ - & - - - - - - - \\ 0 & P^2_{2,(\Delta,\lambda)} \end{array}\right]$$

$$\text{with } P^2_{1,\Delta,\lambda} = \begin{bmatrix} 0 & 0 \\ a & 0 \end{bmatrix}^2 = \begin{bmatrix} 0 & 0 \\ 0 & 0 \end{bmatrix}.$$

Result

Recursively, $n \geq 2$,

$$P^n_{\Delta,\lambda} = \begin{bmatrix} 0 & \left(P_{1,(\Delta,\lambda)}P_{3,(\Delta,\lambda)} + P_{3,(\Delta,\lambda)}P_{2,(\Delta,\lambda)}\right) P^{n-2}_{2,(\Delta,\lambda)} \\ 0 & P^n_{2,(\Delta,\lambda)} \end{bmatrix}$$

Calculate the limit $\lim_{n\to\infty} \mathbf{P}^n_{\Delta,\lambda} = \mathbf{P}^\infty_{\Delta,\lambda}$

$$\mathbf{P}^\infty_{\Delta,\lambda} = \begin{bmatrix} \mathbf{0} & \left(\mathbf{P}_{1,(\Delta,\lambda)}\mathbf{P}_{3,(\Delta,\lambda)} + \mathbf{P}_{3,(\Delta,\lambda)}\mathbf{P}_{2,(\Delta,\lambda)}\right)\mathbf{P}^\infty_{2,(\Delta,\lambda)} \\ \mathbf{0} & \mathbf{P}^\infty_{2,(\Delta,\lambda)} \end{bmatrix}$$

with

$$\mathbf{P}^\infty_{2,(\Delta,\lambda)} = \lim_{n\to\infty} \mathbf{P}^{n-2}_{2,(\Delta,\lambda)} \quad \text{and}$$

$$\mathbf{P}^\infty_{2,(\Delta,\lambda)} = \mathbf{P}^\infty_{2,(\Delta,\lambda)}\mathbf{P}_{2,(\Delta,\lambda)} \Leftrightarrow \mathbf{0} = \mathbf{P}^\infty_2\left(\mathbf{I} - \mathbf{P}_2\right)$$

$\mathbf{P}^n_{\Delta,\lambda}$ tends for a matrix with all lines equal, of the form

$$\mathbf{P}^n_{\Delta,\lambda} \to \left[\mathbf{0} \mid \mathbf{P}^\infty_{2,(\Delta,\lambda)}\right]$$

Example (cont'd)

$$\mathbf{P}_{2,\Delta,\lambda} = \begin{bmatrix}
e^{-\lambda} & & \lambda e^{-\lambda} & & \lambda^2 e^{-\lambda}/2 & & \lambda^3 e^{-\lambda}/6 & & 1-e^{-\lambda}\sum_{i=0}^{3}\lambda^i/i! \\
e^{-\lambda} & & & & \lambda e^{-\lambda} & & \lambda^2 e^{-\lambda}/2 & & 1-e^{-\lambda}\sum_{i=0}^{2}\lambda^i/i! \\
 & e^{-\lambda} & & & \lambda e^{-\lambda} & & \lambda^2 e^{-\lambda}/2 & & 1-e^{-\lambda}\sum_{i=0}^{2}\lambda^i/i! \\
e^{-\lambda} & & & & & & \lambda e^{-\lambda} & & 1-e^{-\lambda}\sum_{i=0}^{1}\lambda^i/i! \\
0 & & & e^{-\lambda} & & & \lambda e^{-\lambda} & & 1-e^{-\lambda}\sum_{i=0}^{1}\lambda^i/i! \\
e^{-\lambda} & & & & & & & & 1-e^{-\lambda} \\
 & & & & & e^{-\lambda} & & & 1-e^{-\lambda} \\
e^{-\lambda} & & & & & & & & 1-e^{-\lambda} \\
 & & & & & & & e^{-\lambda} & 1-e^{-\lambda}
\end{bmatrix}$$

$\lambda = 0.1$

With $\lambda = 0.1$, we get $\mathbf{P}^{\infty}_{2,(\Delta,\lambda)}$ *as*

$$\begin{pmatrix} 0.81873 & 0.067032 & 0.074082 & 0.014905 & 0.016473 & 0.0032584 \\ & & 0.0036011 & 91126 \times 10^{-4} & 10071 \times 10^{-3} & \end{pmatrix}$$

In stationarity, Average Premium is 78.997% of entry Premium.

- Lemaire's (1995):
  - *Relative Stationary Average Level* (RSAL):

$$
\begin{aligned}
RSAL &= \frac{SAP - mP}{MP - mP} \\
SAP &= \sum_{j=1}^{s} b(j) p_T^{(\infty)}(j)
\end{aligned}
$$

    SAP: Stationary Average Premium, mP: minimum Premium, MP: Max Premium
  - Premium variation coefficient (VC):

$$
\begin{aligned}
VC &= SDP / SAP \\
SDP &= \sqrt{\sum_{j=1}^{s} b(j)^2 p_T^{(\infty)}(j) - SAP^2}
\end{aligned}
$$

- Loimaranta's (1972) Efficiency. Elasticity of the average premium (Response to changes in frequency mean)

$$\eta(\lambda) = \frac{\frac{d\, SAP(\lambda)}{SAP}}{\frac{d\lambda}{\lambda}} = \frac{d \ln SAP(\lambda)}{d \ln \lambda}$$

If

$$\begin{aligned} \lambda &\to \infty \Rightarrow SAP(\lambda) \to \max\{b(j)\} < \infty; \\ \lambda &\to \infty \Rightarrow \eta(\lambda) \to 0; \qquad \lambda \to 0 \Rightarrow \eta(\lambda) \to 0. \end{aligned}$$

- Lemaire's (1985) *Transient Elasticity* (1st step analysis)

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

  - $V_\lambda(j)$: Expected present value to be paid by policy from C_j;
  - β_j (< 1): Discount rate.

- Lemaire's (1985) *Transient Elasticity* (1st step analysis)

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

  - $V_\lambda(j)$: Expected present value to be paid by policy from C_j;
  - β_j (< 1): Discount rate.

The system has a unique solution and elasticity comes:

$$\begin{aligned} \mu_\lambda(j) &= \frac{dV_\lambda(j)/V_\lambda(j)}{d\lambda/\lambda} \\ \mu(j) &= \int_0^\infty \mu_\lambda(j) d\pi(\lambda) \end{aligned}$$

"Bonus hunger"

- Due to "Claims Frequency System"
- (Some?) Small accidents aren't reported;
  - It changes: the reported frequency and amonts dist's;
  - Decreases insurer's management costs;
  - "No-report" decision depends:
    - solely **on insured**, and
    - his bonus class C_j;
- Let x_j: Retention level (works like a "Franchise" not a "Deductible");
- It's possible to find an optimal retention point: x_j^* (under some assumptions).

Hypothesis

- (Unreal) Insured knows single amount distr. $F_X(\cdot)$, and x_j;
- $N \frown Poisson(\lambda)$; Single amount $X_i \frown F_X(\cdot)$;
  Let N^*: no. of accidents reported in C_j:

$$N^* = \sum_{i=0}^{N} Y_i, \quad Y_0 \equiv 0$$

$$Y_i \frown binomial(1; p); \qquad p = \Pr\left[X_i > x_j\right] = \bar{F}_X(x_j).$$

  Then

$$N^* \frown CPoisson(\lambda, F_y) \equiv Poisson(\lambda \bar{F}_X(x_j))$$

- Let D : Cost of unreported claim, with mean $E\left[D(x_j)\right]$:

$$D(x_j) = X\mathbb{1}_{\{X \le x_j\}}$$

Hypothesis (cont'd)

- $$E\left[D(x_j)\right] = 0 \times \lambda \bar{F}_X(x_j) + \lambda F_X(x_j)$$
- and payments are made in mid-year:

$$\begin{aligned} V_{\lambda,\mathbf{x}}(j) &= b(j) + \beta^{1/2} E\left[D(x_j)\right] + \beta \sum_{k=1}^{s} p_{T,\lambda,x_j}(j,k) V_{\lambda,\mathbf{x}}(k), \\ & j = 1, ..., s; \end{aligned}$$

Matrix form equation:

$$\begin{aligned} \mathbf{V}_{\lambda,\mathbf{x}} &= \mathbf{b}(\mathbf{x}) + \beta \mathbf{P}_{T,\lambda,\mathbf{x}}(j,k) \mathbf{V}_{\lambda,\mathbf{x}} \\ \mathbf{V}_{\lambda,\mathbf{x}} &= \left(\mathbf{I} - \beta \mathbf{P}_{T,\lambda,\mathbf{x}}\right)^{-1} \mathbf{b}(\mathbf{x}) \\ \mathbf{b}(\mathbf{x})' &= \left(..., b(j) + \beta^{1/2} E\left[D(x_j)\right], ...\right). \end{aligned}$$

Under those conditions it's possible to find optimums x_j^*, see Centeno (2003, pp 181-184), and for algorithms.

- **Norberg's (1976) model**. Once fixed Rules of a BMS, Efficiency Measure of premium $b_n(Z_{\Delta,n})$, as estimator of risk premium $E\left(S_n|\lambda\right)$

$$\begin{aligned} Q_n(\Delta) &= E\left[\left(E\left(S_n|\lambda\right) - b_n(Z_{\Delta,n})\right)^2\right] \\ &= \int_0^\infty \sum_{j=1}^{s} \left(E\left(S_n|\lambda\right) - b_n(j)\right)^2 p_{\Delta,\lambda}^{(n)}(j) d\Pi(\lambda) \end{aligned}$$

$$\begin{aligned} \text{Bonus class in } n &: \quad Z_{\Delta,n}, \quad n = 0, 1, 2, \ldots \\ S_n &: \quad \text{Aggregate claims of policy in } n \\ E\left(S_n|\lambda\right) &: \quad \text{Risk premium, unknown.} \end{aligned}$$

$$\begin{aligned} Q_n(\Delta) &= E\left[E\left[\left(E\left(S_n|\lambda\right) - b_n(Z_{\Delta,n})\right)^2\right]|Z_{\Delta,n}\right] \quad \text{(Like in credibility)} \\ &= E\left[V\left[E\left(S_n|\lambda\right)|Z_{\Delta,n}\right]\right] \\ &\quad + E\left[\left(E\left[b_n(Z_{\Delta,n}) - E(E\left(S_n|\lambda\right)\right]|Z_{\Delta,n})\right)^2\right] \end{aligned}$$

- Norberg's (1976) model (cont'd). Optimal Scale Efficiency Measure

$$Q_n(\Delta) = E\left[\left(E\left(S_n|\lambda\right) - b_n(Z_{\Delta,n})\right)^2\right]$$

Theorem

$$\begin{aligned} Q_n(\Delta) &\geq E\left[V\left[E\left(S_n|\lambda\right)|Z_{\Delta,n}\right]\right]. \\ Q_n(\Delta) &= E\left[V\left[E\left(S_n|\lambda\right)|Z_{\Delta,n}\right]\right] \\ &\quad \textit{iff} \quad \Pr\left[b_n(Z_{\Delta,n}) = \mu_n(Z_{\Delta,n})\right] = 1 \\ \mu_n(Z_{\Delta,n}) &= E\left[E\left(S_n|\lambda\right)|Z_{\Delta,n}\right], \textit{ credibility pr. for yr } n \end{aligned}$$

- Note: $E\left[\mu_n(Z_{\Delta,n})\right] = E\left[E\left(S_n|\lambda\right)\right] = E\left(S_n\right)$

Optimal scale for limiting situation: $Q_0(\Delta) = \lim Q_n(\Delta)$, as $n \to \infty$

$$Q_0(\Delta) = E\left[\left(E\left(S|\lambda\right) - b(Z_T)\right)^2\right],\ S \overset{d}{=} S_n$$

$$b_{\mathbf{T}}(j) = E\left[E\left(S|\lambda\right)|Z_{\mathbf{T}} = j\right] = \frac{\int_0^\infty E\left(S|\lambda\right) p_{T,\lambda}^{(\infty)}(j) d\Pi(\lambda)}{p_T^{(\infty)}(j)}$$

If S_n depends only of λ and use $E(X_i)$ as monetary unit

$$b_{\mathbf{T}}(j) = \frac{\int_0^\infty \lambda p_{T,\lambda}^{(\infty)}(j) d\Pi(\lambda)}{p_T^{(\infty)}(j)}$$

Efficiency Measure

$e(T) = E\left[b_{\mathbf{T}}(Z_{\mathbf{T}})^2\right] = \sum_{j=1}^s b_{\mathbf{T}}(j)^2 p_{\mathbf{T}}^{(\infty)}(j)$

$$T \succ \tilde{T} \quad \text{iff} \quad e(T) > e(\tilde{T})$$

- **Borgan, Hoem & Norberg (1981)' scale**. Non asymptotic criterion and generalization of Norberg's (1976);
- **Linear scales by Gilde & Sundt (1989)**: Linear Norberg (1976) and Linear Borgan et al. (1981);
- **Geometric scales by Andrade & Centeno (2005)**: Geometric Norberg (1976) and Geometric Borgan et al. (1981);

- Ruin Probability criterion (Closed and Open systems): Afonso, Cardoso, Egidio & Guerreiro (2017, 2019)

Borgan *et al.***(1981) scale**: Introduces a system with weights, w_n:

$$\min Q(\Delta) = \sum_{n=0}^{\infty} w_n Q_n(\Delta)$$

$$Q_0(\Delta) = Q_\infty(\Delta)$$

$$p_{\Delta,\lambda}^{(0)}(j) = p_{T,\lambda}^{(\infty)}(j)$$

Solution

$$b_B(j) = \frac{\sum_{n=0}^{\infty} w_n E\left[E[S_n|\Lambda]p_{T,\Lambda}^{(n)}(j)\right]}{p_T(j)}$$

$$p_T(j) = \sum_{n=0}^{\infty} w_n E\left[p_{T,\Lambda}^{(n)}(j)\right]$$

Linear Borgan *et al.*

To turn scales more regular, Impose constraints

$$b(j) = a + j\,b\,, \quad j = 1, \ldots, s$$

- Statistical modelling
  - Model the pure premium
  - Model the Conditional Expected Value:

$$\begin{aligned} E(Y|x_1, x_2, ..., x_p) &= h(x_1, x_2, ..., x_p, \beta_1, \beta_2, ..., \beta_p) \\ Y &= h(x_1, x_2, ..., x_p, \beta_1, \beta_2, ..., \beta_p) + \varepsilon \end{aligned}$$

  Y: endogenous variable, x_i: factor, exogenous, β_j: parameter

- Identify risk factors;
- Different sorts of variables: **Nominal** (binary: gender, good/bad risk), **ordinal/Categorical** (ranks: age, power groups), **discrete** (age, experience yrs, claim counts...), **continuous** (income, claim amounts)
- Data, Information must be (always) reliable, as simple as possible, clean, neat...
- Y: Pure premium, Factors: risk factors influencing:
  - E.g motor insurance: kms, traffic, driver's ability, power, vehicle type, driver's experience, geographical factors...

Deal with the experts about the factors influencing, gather information, data, manageable data. E.g., in motor insurance we can consider

- Past accident record
- kms driven
- Car owner (company/private)
- Use (business or private)
- Vehicle value
- Power (cm^3)
- Weight
- Driver's age
- Driving region (usual, city/countryside...)
- Multiple driver's?
- Vehicle age
- Years fo driver's expereince
- Car brand and/or model

- Gender
- Sort of insurance (third party, own damages)
- Driver's profession
- etc,...
- ....

Then, we have to make choices, run/test models...

- Built classes of factors. Often Class aggregation is needed
- Often we have many binary or rank variables, qualitative data

If dependent variable Y is:

- Binary: Model a *Logit* or *Probit*
- Countig data: *Poisson* model. Ex: Number of claims in a Bonus system
- Continuous data: *Gamma* model. Ex: Amount of claims
- Compound Poisson data: Ex: *Poisson-Gamma Tweedie* model for Aggregate claims data.
  (*Tweedie* dist.family: $\text{Var}(Y) = a[\text{E}(Y)]^p$, $a, p > o$ const.)

Let S be the Aggregate claims in one year, N be the annual number of claims and X be the amount of each claim.

$$E(S) = E(N)E(X), \text{ is the } \textbf{pure premium}.$$

We can consider modeling the two expectations separately.
Or not... Jørgensen & de Souza (1994).

Explanatory variables may affect the expected cost by simultaneously increasing or decreasing both the claim frequency and the average claim size.

In practice, some explanatory factors will have a greater impact on the frequency of claims than on their size, or the opposite.

It is also possible for certain factors, e.g. no-claims bonus, to affect the frequency of claims and the claim size in opposite directions.

In a portfolio we can consider different level factors influencing each (conditional) expectation, building a tariff, such that:

$$E(Y|x_1, x_2, ..., x_p) = h(x_1, x_2, ..., x_p, \beta_1, \beta_2, ..., \beta_p)$$

Specifying $h(x_1, x_2, ..., x_p, \beta_1, \beta_2, ..., \beta_p)$ may not be an easy task, where the $x_1, x_2, ..., x_p$ are the factors.

A tariff analysis is based on insurer's own data.
Steps:

- Postulate a distribution of Y according to its nature, as well as the factors $(x_1, x_2, ..., x_p)$;
- Based on a sample for Y and $(x_1, x_2, ..., x_p)$ choose the *best* $h(.)$ and estimate $(\beta_1, \beta_2, ..., \beta_p)$;
- Hypothesis testing, for Y and $(x_1, x_2, ..., x_p)$.

We should consider:

- Existing information in the company;
- Used variables in other, previous, studies;
- Market used variables;
- Legal limitations.

Data:

- Must be reliable, objective;
- Number of variables must be adequate, no too long or too short;
- All information must cover an homogeneous period. Not too long periods, e.g.

Models:

- Additive models. ANOVA;
- Mutliplicative models, GLM, e.g. two rating factors:

$$\mu_{ij} = \gamma_0 \gamma_{1i} \gamma_{2j}$$

- Key ratio

$$Y_{ij} = X_{ij} / w_{ij}$$

- Mean of *key ratio*:

$$\mu_{ij} = E(Y_{ij}), \text{ with } w_{ij} = 1$$

- Mutliplicative models, extension to *many* rating factors, M:

$$
\begin{aligned}
\mu_{1i_1,i_2,\ldots,i_M} &= \gamma_0\gamma_{1i_1}\gamma_{2i_2}\times\ldots\times\gamma_{Mi_M} \\
\mu_{1i_1,i_2,\ldots,i_M}: &\quad \text{Mean of dependent var. with } M \text{ rating factors} \\
M: &\quad \text{Number of rating factors} \\
\gamma_{ij}: &\quad \text{Rating factor } i \text{ in Class } j
\end{aligned}
$$

- Exponential dispersion models (EDM's) of GLM's generalise the normal distribution used in the linear models.

$$\textit{Pure Premium} = \textit{Claim frequency} \times \textit{Claim severity}$$

For each of the two factors, we can have different rating factors, separately, since severity and frequency are independent.

Table 1.1 Rating factors in moped insurance

Rating factor	Class	Class description
Vehicle class	1	Weight over 60 kg and more than two gears
	2	Other
Vehicle age	1	At most 1 year
	2	2 years or more
Geographic zone	1	Central and semi-central parts of Sweden's three largest cities
	2	Suburbs and middle-sized towns
	3	Lesser towns, except those in 5 or 7
	4	Small towns and countryside, except 5–7
	5	Northern towns
	6	Northern countryside
	7	Gotland (Sweden's largest island)

Table 1.2 Key ratios in moped insurance (claim frequency per mille)

Tariff cell			Duration	No. claims	Claim frequency	Claim severity	Pure premium	Actual premium
Class	Age	Zone						
1	1	1	62.9	17	270	18 256	4 936	2 049
1	1	2	112.9	7	62	13 632	845	1 230
1	1	3	133.1	9	68	20 877	1 411	762
1	1	4	376.6	7	19	13 045	242	396
1	1	5	9.4	0	0	.	0	990
1	1	6	70.8	1	14	15 000	212	594
1	1	7	4.4	1	228	8 018	1 829	396
1	2	1	352.1	52	148	8 232	1 216	1 229
1	2	2	840.1	69	82	7 418	609	738
1	2	3	1 378.3	75	54	7 318	398	457
1	2	4	5 505.3	136	25	6 922	171	238
1	2	5	114.1	2	18	11 131	195	594
1	2	6	810.9	14	17	5 970	103	356
1	2	7	62.3	1	16	6 500	104	238
2	1	1	191.6	43	224	7 754	1 740	1 024
2	1	2	237.3	34	143	6 933	993	615
2	1	3	162.4	11	68	4 402	298	381
2	1	4	446.5	8	18	8 214	147	198
2	1	5	13.2	0	0	.	0	495
2	1	6	82.8	3	36	5 830	211	297
2	1	7	14.5	0	0	.	0	198
2	2	1	844.8	94	111	4 728	526	614
2	2	2	1 296.0	99	76	4 252	325	369
2	2	3	1 214.9	37	30	4 212	128	229
2	2	4	3 740.7	56	15	3 846	58	119
2	2	5	109.4	4	37	3 925	144	297
2	2	6	404.7	5	12	5 280	65	178
2	2	7	66.3	1	15	7 795	118	119

Table 1.3 Important key ratios

Exposure w	Response X	Key ratio $Y = X/w$
Duration	Number of claims	Claim frequency
Duration	Claim cost	Pure premium
Number of claims	Claim cost	(Average) Claim severity
Earned premium	Claim cost	Loss ratio
Number of claims	Number of large claims	Proportion of large claims

EDM's of GLM's

- Data, Key Ratios Obs org'zed in list form $(y_1, ...y_n)'$;
- Row i contains y_i, exposure weight w_i and rating factors ob's;

Tariff cell i	Covariates: Class x_{i1}	Age x_{i2}	Zone x_{i3}	Duration (exposure) w_i	Claim frequency y_i
1	1	1	1	62.9	270
2	1	1	2	112.9	62
3	1	1	3	133.1	68
4	1	1	4	376.6	19
5	1	1	5	9.4	0
6	1	1	6	70.8	14
7	1	1	7	4.4	228
8	1	2	1	352.1	148
9	1	2	2	840.1	82
⋮	⋮	⋮	⋮	⋮	⋮
21	2	1	7	14.5	0
22	2	2	1	844.8	111
23	2	2	2	1 296.0	76
24	2	2	3	1 214.9	30
25	2	2	4	3 740.7	15
26	2	2	5	109.4	37
27	2	2	6	404.7	12
28	2	2	7	66.3	15

- Prob'y dist of the Claim Frequency: Poisson, mixed Poisson. Let X_i in cell i with w_i,

$$X_i \frown Poisson(w_i \mu_i) \Rightarrow Y_i = X_i / w_i \frown relative\ Poisson$$

- Model for claim severity: Gamma, $X \frown Gamma(w\alpha, \beta)$

$$\Rightarrow Y = X / w \frown Gamma(w\alpha, w\beta)\,,\ \ E[X] = \alpha / \beta$$

- Tweedie models:
  - EDM's that are scale invariant, those with variance function $\nu(\mu) = \mu^p$.
  - If $1 < p < 2$ correspond to the Compound Poisson. Key ratio: Pure premium.
  - Model altogether the pure premium, not claim counts and size separately.

Table 2.7 Moped insurance: relativities from a multiplicative Poisson GLM for claim frequency and a gamma GLM for claim severity

Rating factor	Class	Duration	No. claims	Relativities, frequency	Relativities, severity	Relativities, pure premium
Vehicle class	1	9833	391	1.00	1.00	1.00
	2	8824	395	0.78	0.55	0.42
Vehicle age	1	1918	141	1.55	1.79	2.78
	2	16740	645	1.00	1.00	1.00
Zone	1	1451	206	7.10	1.21	8.62
	2	2486	209	4.17	1.07	4.48
	3	2889	132	2.23	1.07	2.38
	4	10069	207	1.00	1.00	1.00
	5	246	6	1.20	1.21	1.46
	6	1369	23	0.79	0.98	0.78
	7	147	3	1.00	1.20	1.20

36

Table 2.8 Motorcycle insurance: rating factors and relativities in current tariff

Rating factor	Class	Class description	Relativity
Geographic zone	1	Central and semi-central parts of Sweden's three largest cities	7.678
	2	Suburbs plus middle-sized cities	4.227
	3	Lesser towns, except those in 5 or 7	1.336
	4	Small towns and countryside, except 5–7	1.000
	5	Northern towns	1.734
	6	Northern countryside	1.402
	7	Gotland (Sweden's largest island)	1.402
MC class	1	EV ratio –5	0.625
	2	EV ratio 6–8	0.769
	3	EV ratio 9–12	1.000
	4	EV ratio 13–15	1.406
	5	EV ratio 16–19	1.875
	6	EV ratio 20–24	4.062
	7	EV ratio 25–	6.873
Vehicle age	1	0–1 years	2.000
	2	2–4 years	1.200
	3	5– years	1.000
Bonus class	1	1–2	1.250
	2	3–4	1.125
	3	5–7	1.000