Teoria da Credibilidade

Alfredo D. Egídio dos Reis

Dexembro de 2001

Conteúdo

Lista de Tabelas

Lista de Figuras

1 Introdução

Na actividade seguradora, a Teoria da Credibilidade pretende desenvolver um método de tarifação baseado na experiência de um conjunto de riscos, aqui significando um conjunto de contratos de seguro de certo tipo. Muitas vezes as companhias seguradoras têm que estabelecer um prémio para um grupo ou uma carteira de contratos de seguro, basicamente porque os diferentes contractos se referem a riscos semelhantes. A tarifação torna-se mais simples de implementar. Grande parte das vezes o estabelecimento de um mesmo prémio a um conjunto de riscos é um recurso *obrigatório* já que não dispomos de informação ou experiência suficiente para um conjunto particular de apólices e, em contraste, *muita* informação referente a outros contratos de alguma forma relacionados. A solução será portanto agrupá-los a todos e atribuir um prémio comum. Este será estimado através da informação disponível sobre todos os riscos. Esta solução será boa se a carteira for homogénea. Se existir um grau de heterogeneidade significativa então será claramente mais apropriado estabelecer um esquema de tarifação baseado na experiência individual, ou seja um prémio para cada risco da carteira. Note-se que numa situação destas fosse atribuído o prémio coletivo, ou prémio comum a todos os riscos, este seria influenciado pelos maus riscos, podendo levar à *fuga* dos *bons* para a concorrência, restando os outros, *menos bons*. Esta não é concerteza boa estratégia para a seguradora. No entanto esta solução (extrema) de tarifar individualmente, pode não ser solução porque muitas vezes não se dispõe de informação suficiente e, por outro lado, é a negação completa de que os riscos são à partida semelhantes. Porque não então encontrarmos uma solução de compromisso?

É essa solução de compromisso que pretendemos alcançar com a teoria da credibilidade. Concretizemos: Pretendemos estimar o prémio puro de um risco, representado pela variável aleatória X_j, para o próximo período a tarifar, risco este pertencente a uma determinada carteira, através da informação obtida pela observações dos riscos da carteira em períodos passados. O prémio puro é o valor esperado do risco. Sabemos da teoria do risco e da ruína que um prémio deve ser superior ao valor esperado do risco, sob pena da ruína da carteira ser *quase certa. Mais tarde ou mais cedo.* Isto é, um prémio deve ser igual ao valor esperado do risco mais uma quantia, a carga, para fazer face a desvios aleatórios relativamente à sua média. Não consideramos aqui cargas de índole administrativa e/ou comerciais. Seja $\bar{X}_j$ a média de indemnizações em períodos passados do risco j e seja $\bar{X}$ a média global de indemnizações de todos os riscos nos mesmos períodos na carteira. A teoria da credibilidade. propõe estimar o prémio puro para o risco j através da média ponderada:

$$z_j\bar{X}_j + (1 - z_j)\bar{X}$$

O factor de ponderação z_j ($0 \leq z_j \leq 1$) expressa o grau de credibilidade. que se atribui à experiência particular do risco j e por isso chamamos Factor de Credibilidade e o prémio representado por esta média ponderada chamamos Prémio de Credibilidade. Esta é a chamada fórmula básica de credibilidade.

No presente texto apresentamos alguns modelos importantes desenvolvidos pela da Teoria da Credibilidade como aplicação ao cálculo/estimação de prémios de seguros, baseado na experiência. O conceito de credibilidade, como perspectiva de aplicação prática ao cálculo de prémios de seguros, foi introduzido pelos actuários americanos no início do século XX, mais concretamente por alturas da Grande Guerra, relacionado com o problema do reajustamento dos prémios no seguro de acidentes de trabalho. Pretendia-se que tivesse em conta o seu registo de indemnizações.

O problema do cálculo do prémio é um problema fulcral. Pretende-se que seja equilibrado. Que proporcione satisfação mútua às entidades envolvidas, o segurador e o segurado. Por um lado, o segurado vê transferida para a companhia, uma situação de risco, aleatória, e para a qual não se sente vocacionado a suportar. Estará então disposto a pagar uma determinada quantia que considere justa, o prémio. Por outro lado, a seguradora está disposta assumir o risco de pagamento de eventuais indemnizações, actividade para a qual está vocacionada, mediante o recebimento do prémio, desde que não lhe traga desvantagem do ponto de vista financeiro.

O cálculo do prémio assume um papel fundamental na relação segurador-segurado. Por um lado, que não seja excessivo de forma a levar o segurado a afastar-se e a procurar outra entidade (assumindo uma situação de concorrência) e que não seja demasiado baixo, que leve a companhia a sentir dificuldades no assumir do seu compromisso de pagamento das indemnizações. Os prémios são pois quantias monetárias, pagas por período, normalmente um ano. O prémio deve reflectir o mais possível as características do risco a que se refere. No entanto, deve referir-se que é um objectivo difícil de alcançar em absoluto, pois existem em geral características que são difíceis de concretizar à partida. Numa tentativa de minorar este problema examina-se a experiência obtida pelo risco ao longo de vários períodos de tarifação, nos casos em que exista.

De forma a diversificar o risco (e não só!), as companhias agrupam os riscos em classes mais ou menos homogéneas, de acordo com as suas semelhanças, a cuja classe é inicialmente atribuído o mesmo prémio, partindo do princípio de que os riscos são semelhantes. Este prémio é designado por Prémio Coletivo. Outro motivo deste agrupamento reside na simplificação prática do cálculo do prémio, porque as companhias são normalmente confrontadas com o cálculo de milhares de prémios, tornando a sua individualização mais difícil de efectuar, sob pena de incorrerem em custos administrativos desnecessários. No entanto, cada risco dentro de cada classe tem uma individualidade própria que os distingue entre si dentro da classe. Esta individualidade é na maior parte dos casos apenas compreensível quando se analizam os registos de indemnizações, referentes à experiência de cada risco dentro da classe. Será então desejável que o prémio seja ajustado tendo em conta esta individualidade. O prémio, individual, que tem em conta as características específicas do risco, é designado por Prémio de Risco. Este prémio permanece em geral desconhecido, pelas razões apresentadas.

Ao desenvolvimento das várias fórmulas heurísticas de estimação do prémio, baseadas na experiência individual das indemnizações em períodos passados, deram os actuários americanos o nome de Teoria da Credibilidade. Mais concretamente, chamaram Fórmula de Credibilidade à média ponderada entre a média de indemnizações passadas sobre o risco em análise e o Prémio Puro Colectivo. Ao factor de ponderação convencionou-se chamar Factor de Credibilidade.

Os resultados obtidos pela teoria da credibilidade têm hoje em dia aplicação prática em vários tipos de seguro, como sejam por exemplo em sistemas de *bonus-malus* no seguro automóvel [ver Andrade e Silva (1991)], nos contratos de resseguro, em seguros de grupo, nomeadamente seguros de acidentes de trabalho em que cada grupo é constituído por indivíduos que trabalham na mesma firma, no cálculo/constituição das provisões para sinistros, em que se pretende fazer uma avaliação futura e localizada no tempo de responsabilidades de pagamentos de indemnizações por parte das seguradoras. Isto inclui a avaliação das indemnizações IBNR (*Incurred But Not Reported*), sinistros ocorridos e ainda não participados ou encerrados [ver Cordeiro (1991)].

2 Princípios de cálculo do prémio

Como complemento, neste capítulo apresentamos alguns princípios de cálculo do prémio, bem como algumas das propriedades a que um princípio deve satisfazer. Não se pretendemos fazer um estudo exaustivo, mas apenas uma pequena apresentação para sublinhar alguns aspectos importantes. Algumas propriedades vão ser discutidas consoante o princípio. Para o estudo da Teoria da Credibilidade interessa-nos apenas a reter que no cálculo do prémio terá que se calcular o prémio puro mais uma componente que é a *carga*. O cálculo de prémios é baseado no princípio de que um conjunto de indemnizações possa ser compensado por pagamentos fixos, habitualmente de forma periódica e anual.

2.1 Alguns princípios

Consideremos um determinado risco, que representa uma apólice ou um conjunto de apólices e designe-se por X a variável aleatória (v.a.) não negativa, que representa o montante das indemnizações agregadas ocorridas num determinado período de tarifação. Seja $F(x)$ a função de distribuição da v.a. X e admita-se a existência dos momentos de X envolvidos, particularmente designe-se por μ_X e σ_X^2 o valor esperado e a variância da v. a. X, respectivamente. Eventualmente também nalguns casos se admite uma hipótese mais restritiva, ou seja a existência da função geradora de momentos de X, o que se tornará claro no texto. Seja P_X, ou mais simplesmente P, o prémio cobrado, líquido de cargas para despesas de carácter administrativo relativo a este risco X, para o mesmo período.

Alguns exemplos de princípios de cálculo do prémio P são a seguir apresentados:

1. Princípio do Valor Esperado:
$$P = (1+\alpha)\mu_X \ \ , \ \alpha > 0 \, ;$$
2. Princípio da Variância:
$$P = \mu_X + \alpha\sigma_X^2 \ \ , \ \alpha > 0 \, ;$$
3. Princípio do Desvio Padrão:
$$P = \mu_X + \alpha\sigma_X \ \ , \ \alpha > 0 \, ,$$
em que α é uma constante positiva designada por Coeficiente de Carga.
4. Princípio da Utilidade Nula:

   Suponha que a seguradora adopta uma função de utilidade $u(x)$, com $u'(x) > 0$ e $u''(x) < 0$ (aversão ao risco). Seja w a riqueza actual do segurador, então P_X é calculado a partir de
$$u(w) = \mathrm{E}[u(w + P_X - X)]$$
   Em geral P_X depende de w. Se $u(x) = -\exp\{-\beta x\}$, $\beta > 0$, então tem-se $P_X = \beta^{-1}\log E[e^{-\beta X}]$. Este caso designa-se por Princípio Exponencial. Neste caso o prémio não depende da riqueza inicial.
5. Princípio de Esscher
$$P_X = \frac{\mathrm{E}[Xe^{hX}]}{\mathrm{E}[e^{hX}]}, \ h > 0.$$
   em que $\mathrm{E}[e^{hX}] = M_X(h)$ é a função geradora de momentos de X no ponto $h > 0$ e supomos a sua existência. Para interpretar a expressão acima, consideremos
$$g(x) = \frac{e^{hx}f(x)}{\int_0^\infty e^{hx}f(x)dx}, \ h > 0$$
   ou seja $g(x)$ pode ser vista como uma função de densidade de uma variável aleatória não-negativa Y com função de distribuição $G(x) = \int_0^x e^{hy}f(y)dy/M_X(h)$. $G(x)$ é a Transformada de Esscher de $F(x)$ com parâmetro h.

   É fácil verificar que a função geradora de momentos de Y, $M_Y(t) = M_X(t+h)/M_X(h)$ e que $\mathrm{E}[Y] = P_X$.
6. Princípio do Prémio de Risco Ajustado (*risk adjusted premium principle*):
$$P = \int_0^\infty [1 - F(x)]^{1/\rho}\, dx$$
   onde $\rho \geq 1$ é o índice de risco. Note-se que se se considerar $\rho = 1$ então $P_X = \mathrm{E}[X]$. O princípio é baseado numa transformada da distribuição $F(x)$. Seja uma v.a. Z com função distribuição $H(x)$ onde $1 - H(x) = [1 - F(x)]^{1/\rho}$, então a respectiva função densidade
$$h(x) = H'(x) = \frac{1}{\rho}\,[1 - F(x)]^{1/\rho - 1}\, f(x).$$

- **Exemplo 1** *Seja* $X \sim \mathrm{Pareto}(\alpha, \beta)$ *com f.d.* $F(x) = 1 - \left(\frac{\beta}{\beta+x}\right)^\alpha$. *Então* $1 - H(x) = \left(\frac{\beta}{\beta+x}\right)^{\alpha/\rho}$, *donde* $Z \sim \mathrm{Pareto}(\alpha/\rho, \beta)$ *e portanto* $P = E[Z] = \frac{\beta}{\alpha/\rho - 1}$ *desde que* $\alpha > \rho$. *Note-se que* P *aumenta com* ρ. ◇

Relativamente aos três primeiros princípios apresentados, princípio do valor esperado, da variância e do desvio padrão, temos que em cada um deles o prémio é composto por duas parcelas: uma – o valor esperado do risco $\mathrm{E}(X)$– comum com os três princípios e a qual é designada por Prémio Puro, e a outra – uma carga para compensar desvios de sinistralidade – calculada de maneira diferente de princípio para princípio, mais exactamente, proporcional ao parâmetro que dá o nome ao princípio. α é uma constante positiva designada por Coeficiente de Carga. É evidente que para estes princípios $P > \mathrm{E}[X]$. Relativamente aos outros princípios também se mostra mais à frente, a propósito das propriedades que P é não inferior ao valor esperado do risco, Prémio Puro.

De uma forma mais geral, um princípio de cálculo de prémios pode ser representado por uma funcional $\mathcal{H}$ do conjunto das funções em $\Re^{+}$.

2.2 Propriedades

Seja o prémio P, calculado duma forma geral através da funcional $\mathcal{H}$ aplicada à variável aleatória X. Exemplos de Propriedades consideradas desejáveis para um prémio, de um conjunto mais vasto, são as seguintes:

1. $\mathcal{H}(X) \geq \mathrm{E}(X)$.
2. $\mathcal{H}(X) \leq M$, com $M = \inf\{x : F_X(x) = 1\}$.
3. Aditividade: Sejam dois riscos, X_1 e X_2, independentes: $\mathcal{H}(X_1 + X_2) = \mathcal{H}(X_1) + \mathcal{H}(X_2)$.
4. Sejam dois riscos, X_1 e X_2: $\mathcal{H}(X_1 + X_2) \leq \mathcal{H}(X_1) + \mathcal{H}(X_2)$.

   Se adicionalmente se considerar que X_1 e X_2 são independentes, então a propriedade é designada por Sub-aditividade.
5. Sejam dois riscos X_1 e X_2: $\mathcal{H}(X_1) \leq \mathcal{H}(X_1 + X_2)$.
6. Iteratividade: Sejam dois riscos X e Y: $\mathcal{H}(X) = \mathcal{H}[\mathcal{H}(X|Y)]$.

A Propriedade 1 pode ser justificada do seguinte modo: Considere-se uma sucessão $\{X_n\}$ de n variáveis aleatórias independentes e identicamente distribuídas, com média μ e desvio padrão σ, e a aplicação do princípio $\mathcal{H}$ para o cálculo do prémio P associado a cada um dos riscos.

Se aplicarmos o Teorema do Limite Central, obtemos

$$\lim \ \text{Prob} \ (\sum_{i=1}^{n} X_i > nP) = 1$$

se considerássemos que $P < \mu$. Isto significa que a companhia obteria um prejuízo *quase certo* com a operação (no longo prazo).

A Propriedade 2 diz que o prémio P deve ser menor que a indemnização máxima, o que garante a possibilidade de o segurado não incorrer num prejuízo quase certo.

A Propriedade 3 diz que o prémio da soma é igual à soma dos prémios. Assim, a junção de riscos independentes não deve afectar o prémio total.

Quanto à Propriedade 4, caso ela não se verificasse o segurado ganharia com a subdivisão do risco. Claramente, a aditividade implica a sub-aditividade.

A Propriedade 5 diz que o prémio aplicado a um risco não será superior ao prémio aplicado caso se juntarem outros riscos.

A Propriedade 6 significa que o prémio para o risco X pode ser calculado em dois passos. Em primeiro lugar calcula-se o prémio condicionado (dado Y) para X, $\mathcal{H}(X|Y)$, aplicando o princípio $\mathcal{H}$ à distribuição condicionada de X. Este prémio é uma função de Y e portanto é uma variável aleatória com determinada distribuição. Aplicando por sua vez o mesmo princípio $\mathcal{H}$ à distribuição de $\mathcal{H}(X|Y)$, obtém-se $\mathcal{H}[\mathcal{H}(X|Y)]$. O princípio é iterativo se $\mathcal{H}(X) = \mathcal{H}[\mathcal{H}(X|Y)]$. Esta é uma propriedade desejável pois é muito conveniente do ponto de vista matemático. A verificar-se seria particularmente útil para a Teoria da Credibilidade.

Iremos seguidamente examinar algumas propriedades relativamente aos princípios de cálculo do prémio atrás apresentados. Alguns casos que se omitem ficam fora do âmbito deste trabalho.

Constata-se imediatamente que os princípios do valor esperado, da variância e do desvio padrão verificam a Propriedade 1.

- Para o princípio da utilidade nula, recorrendo a desigualdade de Jansen[1] tem-se que

$$u(w) = \mathrm{E}[u(w+P-X)] \leq u(\mathrm{E}[w+P-X]) = u(w+P-\mathrm{E}[X])$$

donde se obtém $P \geq \mathrm{E}[X]$, já que $u'(x) > 0$.

- Para o princípio de Esscher note-se que se $h = 0$ (i.e. não há transformação) então $P_X = \mathrm{E}[Y] = \mathrm{E}[X]$. Considere-se agora $h \geq 0$. Tem-se

$$E[Y^r] = \frac{d^r}{dt^r} M_Y(t)\bigg|_{t=0} = \frac{d^r}{dt^r} \frac{M_X(t+h)}{M_X(h)}\bigg|_{t=0} = \frac{M_X^{(r)}(h)}{M_X(h)}$$

portanto

$$\begin{aligned} \frac{d}{dh} P_X &= \frac{d}{dh} E[Y] = \frac{d}{dh} \frac{M_X'(h)}{M_X(h)} \\ &= \frac{1}{M_X(h)^2} \left[M_X^{(2)}(h) M_X(h) - M_X'(h)^2 \right] = E[Y^2] - E[Y]^2 \geq 0 \end{aligned}$$

ou seja P_X é uma função não decrescente de h e portanto $P_X \geq \mathrm{E}[X]\ \forall h \geq 0$.

- Relativamente ao último princípio note-se que para $\rho \geq 1$ se tem $1 - F(x) \leq [1-F(x)]^{1/\rho}$, então $\mathrm{E}[X] = \int_0^\infty [1-F(x)]\, dx \leq \int_0^\infty [1-F(x)]^{1/\rho}\, dx = P$.

A conclusão a retirar da verificação desta propriedade por todos os princípios apresentados é a de que o Prémio Puro $\mathrm{E}[X]$ é a componente de referência no cálculo do prémio. O Prémio Puro é em geral o objecto da teoria da credibilidade.

Quanto à Propriedade 2, nenhum dos três primeiros princípios apresentados a verificam em geral:

- Para o Princípio do Valor Esperado tomemos como contra-exemplo o seguinte:
Seja a v.a. X tal que $\Pr(X = M) = p$ e $\Pr(X = 0) = 1 - p$. Assim, tem-se

$$\mathrm{E}(X) = Mp, \quad \mathrm{V}(X) = M^2 p(1-p), \quad \mathcal{H}(X) = M\left[p + \alpha\sqrt{p(1-p)}\right];$$

Derivando $\mathcal{H}(X)$ em ordem a p, temos

$$\frac{d\mathcal{H}}{dp} = M\left(1 + \alpha \frac{1-2p}{2\sqrt{p(1-p)}}\right)$$

calculando o limite quando $p \to 1$, tem-se

$$\lim_{p\to 1} \frac{d\mathcal{H}}{dp} = -\infty, \quad \lim_{p\to 1} \mathcal{H}(X) = M.$$

Deve pois existir um $p < 1$ tal que:

$$\mathcal{H}_p(X) > M = \lim_{p\to 1} \mathcal{H}(X).$$

- Para o Princípio da Variância e tomando o exemplo anterior, temos $\mathcal{H}(X) = Mp + \alpha M^2 p(1-p)$. Por outro lado a Propriedade 2 impõe que $Mp + \alpha M^2 p(1-p) \leq M$, condição que só é verificada para $\alpha \leq 1/Mp$.

[1] Dada uma função côncava ϕ e uma v.a. X tem-se que $\mathrm{E}[\phi(X)] \leq \phi(\mathrm{E}[X])$.

- Para o Princípio da Utilidade Nula notemos que $w+P-X \geq w+P-M$, então $\mathrm{E}[u(w+P-X)] \geq \mathrm{E}[u(w+P-M)] \Rightarrow u(w) \geq u(w+P-M) \Rightarrow P \leq M$ porque $u'(x)>0$.

- Para o princípio de Esscher tem-se que

$$P = \frac{\mathrm{E}[Xe^{hX}]}{\mathrm{E}[e^{hX}]} \leq \frac{\mathrm{E}[Me^{hX}]}{\mathrm{E}[e^{hX}]} = M$$

- Relativamente ao último princípio seja $M < \infty$ tal que $F(M)=1$. Tem-se

$$P = \int_0^\infty [1-F(x)]^{1/\rho}\, dx = \int_0^M [1-F(x)]^{1/\rho}\, dx \leq \int_0^M dx = M$$

Relativamente à Propriedade 3, facilmente concluímos que tanto o Princípio do Valor Esperado como o Princípio da Variância a verificam. Em relação a este último notemos que $\mathrm{V}(X_1+X_2) = \mathrm{V}(X_1)+\mathrm{V}(X_2)$, em virtude da independência das variáveis.

- Para o Princípio do Desvio Padrão, temos

$$\begin{aligned}\mathcal{H}(X_1+X_2) &= \mathrm{E}(X_1)+\mathrm{E}(X_2)+\alpha\sqrt{\mathrm{V}(X_1)+\mathrm{V}/X_2)}\\ \mathcal{H}(X_1)+\mathcal{H}(X_2) &= \mathrm{E}(X_1)+\mathrm{E}(X_2)+\alpha\left(\sqrt{\mathrm{V}(X_1)}+\sqrt{\mathrm{V}(X_2)}\right),\end{aligned}$$

donde

$$\mathcal{H}(X_1+X_2) \neq \mathcal{H}(X_1)+\mathcal{H}(X_2).$$

- Para o princípio da utilidade nula esta propriedade não se verifica em geral. Verifica-se no entanto para o princípio exponencial: $u(x) = -\exp(-\beta x)$. É um bom exercício.

- Para o princípio de Esscher, seja

$$\mathcal{H}(X_1+X_2) = \frac{\mathrm{E}\left[(X_1+X_2)e^{h(X_1+X_2)}\right]}{\mathrm{E}\left[e^{h(X_1+X_2)}\right]} \leq \frac{\mathrm{E}[X_1e^{hX_1}]\mathrm{E}[e^{hX_2}]+\mathrm{E}[X_2e^{hX_2}]\mathrm{E}[e^{hX_1}]}{\mathrm{E}\left[e^{hX_1}\right]\mathrm{E}\left[e^{hX_2}\right]} = \mathcal{H}(X_1)+\mathcal{H}(X_2)$$

- Relativamente ao último princípio esta propriedade não se verifica em geral. Consideremos o exemplo:

Exemplo 2 *Sejam X e Y i.i.d. com* $\Pr[X=0]=\Pr[Y=1]=1/2$ *e seja* $\rho=2$. *Então*

$$\mathcal{H}(X)+\mathcal{H}(Y) = \int_0^1 0,5^{1/2} dx = 0,5^{1/2}$$

por outro lado, $\Pr[X+Y=0] = 0,5^2 = 0,25$, $\Pr[X+Y=1] = 2\times 0,5^2 = 0,5$, $\Pr[X+Y=2] = 0,5^2 = 0,25$. *Então* $\mathcal{H}(X+Y) = \int_0^1 0,75^{1/2}dx + \int_1^2 0,25^{1/2}dx = 0,5(1+\sqrt{3}) < \mathcal{H}(X)+\mathcal{H}(Y)$ $\diamondsuit$

Relativamente à Propriedade 4, tanto o Princípio do Valor Esperado como o Princípio do Desvio Padrão a verificam.

- Em relação ao primeiro princípio a sua verificação é imediata.

- Quanto ao segundo, notando que a propriedade impõe

$$\mathrm{E}(X_1)+\mathrm{E}(X_2)+\alpha\sqrt{\mathrm{V}(X_1+X_2)} \leq \mathrm{E}(X_1)+\mathrm{E}(X_2)+\alpha\left[\sqrt{\mathrm{V}(X_1)}+\sqrt{\mathrm{V}(X_2)}\right]$$

equivalente a

$$\sqrt{\mathrm{V}(X_1)+\mathrm{V}(X_2)+2\,\mathrm{Cov}(X_1;X_2)} \leq \sqrt{\mathrm{V}(X_1)}+\sqrt{\mathrm{V}(X_2)},$$

levantando ao quadrado e simplificando, tem-se a condição equivalente

$$\mathrm{Cov}(X_1; X_2) \leq \sqrt{\mathrm{V}(X_1).\mathrm{V}(X_2)},$$

ou seja

$$\frac{\mathrm{Cov}(X_1; X_2)}{\sqrt{\mathrm{V}(X_1)\mathrm{V}(X_2)}} \leq 1,$$

o que se verifica sempre, uma vez que o primeiro membro é o coeficiente de correlação de X_1 e X_2. Conclui-se também que o Princípio do Desvio Padrão é sub-aditivo. No entanto note-se que este princípio não é aditivo, como atrás se mostrou.

- Quanto ao Princípio da Variância, a Propriedade 4 não se verifica em geral. Só está verificada se e só se

$$\mathrm{V}(X_1 + X_2) \leq \mathrm{V}(X_1) + \mathrm{V}(X_2),$$

isto é, se e só se Cov $(X_1; X_2) \leq 0$. No entanto o Princípio da Variância é sub-aditivo, como facilmente se conclui da condição acima ou tendo em conta que a aditividade implica a sub-aditividade.

Relativamente à Propriedade 5:

- O Princípio do Valor Esperado verifica-a, como facilmente se conclui.
- Quanto ao Princípio do Desvio Padrão deve-se ter

$$\mathrm{E}(X_1) + \alpha\sqrt{\mathrm{V}(X_1)} \leq \mathrm{E}(X_1 + X_2) + \alpha\sqrt{\mathrm{V}(X_1 + X_2)},$$

o que é equivalente a

$$\alpha \leq \frac{\mathrm{E}(X_2)}{\sqrt{\mathrm{V}(X_1)} - \sqrt{\mathrm{V}(X_1 + X_2)}},$$

e como $\alpha > 0$, então deveria ter-se

$$\mathrm{V}(X_1) > \mathrm{V}(X_1 + X_2),$$

ou seja,

$$\mathrm{V}(X_2) + 2\mathrm{Cov}(X_1; X_2) < 0$$

e como $\mathrm{V}(X_2) \geq 0$, então deveria ter-se, Cov $(X_1; X_2) < 0$, o que não se verifica em geral.

- Quanto ao Princípio da Variância e se usarmos o mesmo raciocínio atrás, deveríamos ter

$$\alpha \leq \frac{\mathrm{E}(X_2)}{\mathrm{V}(X_1) - \mathrm{V}(X_1 + X_2)}.$$

Concluímos, portanto, que a Propriedade 5 não se verifica em geral.

Relativamente à Propriedade 6, verifica-se facilmente que nenhum dos três primeiros princípios apresentados é iterativo:

- Para o Princípio do Valor Esperado temos

$$\mathcal{H}(X) = (1+\alpha)\mu_X \quad \text{e} \quad \mathcal{H}(X|y) = (1+\alpha)\mu_X(y)$$

donde $\mathcal{H}[\mathcal{H}(X|Y)] = (1+\alpha)^2\mu_X = (1+\alpha)\mathcal{H}(X)$.

- Para o Princípio da Variância temos.

$$\begin{aligned}
\mathcal{H}(X) &= \mu_X + \alpha\sigma_X^2 = \mathrm{E}[\mu_X(Y)] + \alpha\left(\mathrm{E}[\sigma_X^2(Y)] + \mathrm{V}[\mu_X(Y)]\right) \\
\mathcal{H}(X|y) &= \mu_X(y) + \alpha\mathrm{V}(X|y) \\
\mathcal{H}[\mathcal{H}(X|Y)] &= \mathrm{E}[\mu_X(Y)] + \alpha(\mathrm{E}[\sigma_X^2(Y)] + \mathrm{V}[\mu_X(Y) + \alpha\sigma_X^2(Y)]),
\end{aligned}$$

donde $\mathcal{H}(X) \neq \mathcal{H}[\mathcal{H}(X|Y)]$.

- Para o Princípio do Desvio Padrão temos:

$$\begin{aligned}\mathcal{H}(X) &= \mu_X + \alpha\sigma_X = \mu_X + \alpha\sqrt{\mathrm{E}[\sigma_X^2(Y)] + \mathrm{V}[\mu_X(Y)]} \\ \mathcal{H}(X|y) &= \mu_X(y) + \alpha\sigma_X(y) \\ \mathcal{H}[\mathcal{H}(X|Y)] &= \mu_X + \alpha\left(\mathrm{E}[\sigma_X(Y)] + \sqrt{\mathrm{V}[\mu_X(Y) + \alpha\sigma_X(Y)]}\right),\end{aligned}$$

  donde concluímos $\mathcal{H}(X) \neq \mathcal{H}[\mathcal{H}(X|Y)]$.

Uma nota importante relativamente a esta propriedade e ao desenvolvimento que se iremos fazer relativamente à Teoria da Credibilidade, é o facto de que o Prémio Puro goza da propriedade iterativa, já que o valor esperado é um operador iterativo.

Nenhum dos princípios apresentados verificam todas as propriedades, uns verificarão mais que outros, no entanto em termos absolutos não existe nenhum que possa ser considerado melhor que os outros.

Sublinhemos que o Princípio do Valor Esperado verifica mais propriedades. Atribui no entanto o mesmo prémio a riscos com a mesma média, podendo ter uma variância diferente. Nalguns casos pode ser claramente desadequado. Suponhamos por exemplo dois riscos de características bem distintas e de certa forma extremas:

Exemplo 3 *Sejam dois riscos X e Y com média 1, $X \sim \mathrm{Exponencial}(1)$ e $Y \sim \mathrm{Pareto}(2,1)$. Intuitivamente P_Y deveria ser maior que P_X, Y não tem variância finita.* ♢

Outro aspecto a colocar será como atribuir valores ao parâmetro, Coeficiente de Carga α? Por exemplo:

Exemplo 4 *Suponha-se $X \sim \mathrm{Normal}(200, 100)$. Se se considerar $P_X = (1+\alpha)\mathrm{E}[X]$, pode-se escolher α tal que*

$$\Pr[X \leq P_X] = 0,99 \Leftrightarrow \Phi(20\alpha) = 0,99$$

$\Rightarrow 20\alpha = 2,326 \Rightarrow \alpha = 0,1163$ *ou, em termos percentuais, obtemos um coeficiente de carga de 11,63%.* ♢

Para um estudo mais aprofundado, encontramos um tratamento extensivo de princípios de cálculo do prémio em Goovaerts, de Vylder & Hazendonck (1984). Veja-se também Centeno (2001).

3 Fórmula básica de credibilidade

3.1 Prémio de risco, Prémio coletivo

As companhias seguradoras agrupam os riscos em classes mais ou menos homogéneas, com o objectivo de diversificar, *distribuir*, o risco ou de forma a simplificar o esquema de tarifação, obtendo com isto uma diminuição nos seus custos. O benefício é mútuo, tanto para o segurador como para o segurado, que poderá ver o seu prémio reduzido pela diminuição de custos e mesmo pela diversificação, *disseminação* do risco.

Os riscos agrupados numa determinada classe devem ser semelhantes em relação às suas características básicas, embora se admita um certo grau de heterogeneidade interior, que não seja dominante. Por exemplo, no seguro de automóvel ligeiro, as apólices são classificadas quanto à cilindrada do automóvel, idade do condutor, idade da carta de condução, idade do veículo.

Cada classe assim formada, contém características básicas comuns, que lhe confere um grau de homogeneidade preponderante dentro do grupo. Existem no entanto determinados factores que, ou são inobserváveis ou são difíceis de traçar ou quantificar, ou ainda, a sua especificação é política ou socialmente desaconselhável. Cite-se como exemplo, a habilidade do condutor, as condições de tráfego em que o automobilista circula, o sexo do condutor. Este tipo de factores atribuem a cada risco um certo grau de heterogeneidade dentro do grupo.

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	0,2	0	0	0,2	0,2	0	0,6	0,1	0,4	0,3	0,1	0,1	0	0	0,5	0,1	0,1	0

Tabela 1: Registo de indemnizações de um conjunto de 20 riscos

Cada conjunto de riscos assim formado é designado por Coletivo. O Coletivo é um conjunto de riscos, semelhantes em relação às suas características básicas. Uma vez que os riscos são semelhantes em relação às suas características básicas, não é difícil de admitir que à partida os riscos possam ser tarifados de forma idêntica, i.e., a cada risco será atribuído o mesmo prémio, o Prémio Coletivo. No entanto, as características específicas *escondidas* de cada risco podem dar origem a uma heterogeneidade muitas vezes não desprezível dentro do coletivo. Esta heterogeneidade só se torna clara quando se analizam os registos individuais de indemnizações. Esta característica específica pode ser quantificada por um parâmetro, seja θ, o qual permanece em geral desconhecido.

Consideremos o seguinte exemplo muito sugestivo (de Norberg (1979)).

Exemplo 5 *Suponha-se um coletivo formado por* 20 *riscos, semelhantes em relação às suas características básicas. A cada risco é atribuído anualmente o mesmo prémio puro. Para simplificação, suponha-se que cada risco pode produzir no máximo uma indemnização por ano, de montante igual a uma unidade monetária. Assim, a probabilidade do risco j ($j = 1, 2, ..., 20$) produzir uma indemnização em determinado período é dado pelo parâmetro:*

$$\theta_j = \mathrm{Prob}(X_j = 1),$$

em que X_j representa o montante das indemnizações em determinado período. Claramente, $X_j = 0, 1$ ($j = 1, 2, \ldots, 20$). Suponha-se ainda qua a característica específica de cada risco, aqui representada pelo parâmetro θ_j, se mantém constante ao longo do tempo. Ao admitir-se que o parâmetro θ dependente de j significa admitir-se que os riscos possam ter comportamento algo diferente.

Ao fim de 10 *períodos de tarifação compilaram-se os registos de indemnizações apresentados na Tabela* 1*, correspondentes ao resultado de* 10 *experiências de Bernoulli.*

Uma estimativa para o prémio puro coletivo a atribuir a cada risco, para o período seguinte é dada pela média empírica global $\overline{\hat{\theta}} = 0,145$. Isto pressupõe no entanto que os riscos são homogéneos, i.e., $\theta_1 = \theta_2 = \cdots = \theta_{20}$. Examinando a Tabela 1 *esta hipótese não parece nada óbvia já que, por exemplo, os riscos números* 9, 11 *e* 17 *destacam-se negativamente dos restantes. Aliás, a simples elaboração dum teste estatístico, a hipótese de homogeneidade é rejeitada (ver Norberg (1979)).*

Assim, o Prémio Puro correcto a tarifar para cada risco será a média individual, que se designa Prémio Puro de Risco, dada pelos parâmetros θ_j ($j = 1, 2, \cdots, 20$) e que permanece desconhecida, já que não será justo a atribuição dum prémio comum a todos os riscos. A estimação óbvia neste caso para o prémio individual seria através da média empírica individual, $\hat{\theta}_j$. No entanto o número de observações dos riscos individualizados não é suficientemente grande para que o estimador possa ser considerado 100% credível. Se se considerar esta alternativa, no caso do exemplo presente, o prémio puro estimado vinha nulo para os riscos números 1*,* 2*,* 4*,* 5*,* 8*,* 15*,* 16 *e* 20*! Além disso, a adopção desta alternativa ignora o facto de que os riscos são semelhantes em certo sentido.* ◇

Consideramos neste trabalho basicamente dois tipos de prémios: o **Prémio Coletivo** e o **Prémio de Risco**. O Prémio Coletivo é o prémio aplicado a um risco X tirado ao acaso do Colectivo. Cada risco dentro do

Colectivo é identificado pela sua característica específica, representada pelo parâmetro θ, que permanece desconhecido em geral. O Prémio de Risco é entendido como o *verdadeiro* prémio a atribuir a um risco dentro do coletivo, tendo em atenção a sua característica específica *escondida* θ. Por este facto, o Prémio de Risco permanece desconhecido em geral.

O problema fundamental da tarifação reside na atribuição, estimação, do prémio de risco. Uma vez que é desconhecido em geral, tem de se estimar. No entanto, a sua estimação, baseada na experiência individual, traz outro problema. É que em grande parte dos casos os actuários não dispõem de experiência individual em quantidade considerada suficiente para que a sua estimação possa ser considerada *credível*.

Apenas em condições excepcionais o prémio de risco pode ser conhecido: se o coletivo for perfeitamente homogéneo (nesse caso o Prémio de Risco é igual ao Prémio Coletivo) e se o risco puder ser observado durante um período de tempo bastante longo, mantendo-se inalteradas as condições de risco. Estas situações são muito particulares.

Definição 1 *Seja P o Prémio Coletivo, o prémio obtido pela aplicação do princípio $\mathcal{H}$, definido no Capítulo 2, à distribuição $F_X(x)$ das indemnizações agregadas X do risco no coletivo, referente a determinado período. $F_X(x)$ supõe-se conhecida.*

Definição 2 *Seja $P(\theta)$ o Prémio de Risco, o prémio obtido pela aplicação do mesmo princípio $\mathcal{H}$ à distribuição $F_X(x|\theta)$ das indemnizações agregadas do risco, referente a determinado período. $F_X(x)$ supõe-se conhecida assim como o parâmetro de risco θ.*

Temos então o prémio de risco $P(\theta) = \mathcal{H}(X|\theta)$ e o prémio coletivo $P = \mathcal{H}(X)$. Exemplificando com o Princípio da Variância (ver Capítulo 2) temos $P(\theta) = \mu_X(\theta) + \alpha\sigma_X^2(\theta)$ e $P = \mu_X + \alpha\sigma_X^2$, com μ_X =E(X), σ_X^2 =V(X), $\mu_X(\theta)$ =E$(X|\theta)$ e $\sigma_X^2(\theta)$ =V$(X|\theta)$, cuja existência assumimos. Designaremos sempre daqui em diante $\mu_X(\theta)$ como o Prémio Puro de Risco e μ_X como o Prémio Puro Coletivo.

3.2 Fórmula básica de credibilidade

A atribuição do prémio de risco desconhecido $P(\theta)$ será, em geral, através da sua estimação. Pretendemos pelas razões atrás enunciadas que o estimador reflita não só a informação sobre o coletivo de riscos, mas também a(s) característica(s) individual(ais) específica(s). Teremos que encontrar um estimador cuja expressão incorpore estes dois aspectos.

Quase todo o desenvolvimento da teoria da credibilidade se centra em torno do Prémio Puro. A carga seria calculada posteriormente, com base no risco coletivo ou individual, conforme o caso. O problema da tarifação do prémio puro de risco $\mu_X(\theta)$ foi resolvido numa fase inicial pela atribuição do prémio puro coletivo μ_X a todos os riscos pertencentes a esse mesmo coletivo. Simplesmente isto não era solução. Em 1918 o actuário americano A.W. Whitney (ver Longley-Cook (1962) ou Norberg (1979)[2]) propôs pela primeira vez que a estimativa do prémio puro fosse uma média ponderada entre a experiência do coletivo e a experiência particular do risco.

Concretizemos: Seja $\tilde{m}$ o estimador de $\mu(\theta)$ e X_i $(i = 1, 2, \cdots, n)$ o montante das indemnizações agregadas no período i, e n o número de períodos para o qual existem disponíveis, estatísticas individuais do risco e $\bar{X} = \frac{1}{n}\sum_{i=1}^{n} X_i$. A fórmula proposta foi a seguinte:

$$\tilde{m} = z\bar{X} + (1 - z)\mu_X \tag{1}$$

em que z é uma constante compreendida entre 0 e 1 e é designada por **Factor de Credibilidade**. A Fórmula 1 ficou conhecida como **Fórmula de Credibilidade**.

Vendo a Fórmula 1, verifica-se que é uma média ponderada entre o prémio puro coletivo μ_X e um possível estimador para o verdadeiro prémio puro $\mu(\theta)$. A constante z foi designada por Factor de Credibilidade em virtude de pretender medir o nível de *crédito* a atribuir à influência individual, através da estatística $\bar{X}$ como

[2] Norberg (1979) é um trabalho notável de síntese.

estimador de $\mu(\theta)$, comparativamente com a hipótese inicial de atribuição do Prémio Coletivo. Com esse objectivo deveria atribuir-se a z um valor entre 0 e 1.

Se z fosse um valor próximo de 0, então seria equivalente considerar-se que a experiência individual não trazia grande informação sobre a característica θ desconhecida, ou por outro lado, a informação era relevante mas o risco não se diferenciava do resto do coletivo. No primeiro caso por insuficiência de informação, no segundo admitindo a homogeneidade preponderante do coletivo de riscos.

Se a z se atribuísse um valor próximo de 1, então significaria que a informação disponível descrevia quase ou mesmo completamente o risco (no caso limite), então $\overline{X}$ seria um bom estimador de $\mu(\theta)$. Dizia-se então que era atribuída credibilidade total à informação.

O grande problema passava a ser a concretização de z para cada caso. Ficava um pouco à intuição do actuário, parecendo óbvio que z deveria ser crescente com n, i.e., com a quantidade de informação disponível sobre o risco. No caso de ausência de informação, ou seja, para um risco que integrava o coletivo pela primeira vez, então z tomava o valor 0, e era-lhe atribuído o prémio coletivo μ.

Se a informação fosse numerosa de tal forma que se sentisse que descrevia completamente o risco, então $z = 1$. Além disso era sentido que Z deveria reflectir flutuações aleatórias de sinistralidade de período para período.

Sugestivamente, Witney (ver Longley-Cook (1962)) sugeriu mesmo a seguinte fórmula:

$$z = \frac{n}{n+k} \tag{2}$$

em que k era uma constante heurísticamente determinada e tinha em conta o tipo de seguro ao qual o prémio se referia.

Na Fórmula 1, μ era normalmente substituída por um seu estimador, o que não se afigurava difícil pela possibilidade de existência de um grande número de registos sobre o colectivo, como referimos atrás.

O desenvolvimento da teoria da credibilidade, que é basicamente um problema de estimação e de inferência estatística, depende do tipo de inferência que se tem por base: a Inferência Estatística Clássica ou a Inferência Bayesiana, que tem por base a interpretação do parâmetro θ aqui em causa. Nós vamos apenas tomar em consideração a estimação do prémio de risco do ponto de vista Bayesiano. Para alguma discussão sobre esta questão consulte-se Norberg (1979) ou Egídio dos Reis (1987).

4 Modelos Bayesianos puros

Consideremos uma carteira com N riscos numerados de 1 a N que produziram indemnizações num conjunto de períodos numerados de 1 a n. Sejam as variáveis:

X_{ij} - Representando o montante agregado de indemnizações geradas pelo risco j $(j = 1, 2, \cdots, N)$ no ano i $(i = 1, 2, \cdots, n)$, ou mais simplesmente, risco j no período i.

θ_j - Parâmetro de risco associado ao risco j, independente de i.

$\bar{X}_{.j} = \frac{1}{n}\sum_{i=1}^{n} X_{ij}$ - Média das indemnizações do risco j nos n períodos.

$\overline{\overline{X}}_{..} = \frac{1}{N}\sum_{j=1}^{N} \bar{X}_{.j}$ - Média global dos N riscos nos n períodos.

$\underline{X}_j = (X_{1j}, X_{2j}, \cdots, X_{nj})$ - Vector das variáveis observáveis do risco j nos n períodos.

Consideremos que os N riscos fazem parte dum coletivo em virtude de terem uma característica fundamental comum. A característica individual que diferencia cada risco j no coletivo é quantificada pelo parâmetro θ_j, e é independente de i. A hipótese é simplista pois o risco mantém a mesma característica ao longo do tempo. Isto significa, por exemplo, que a destreza de um condutor não evolui com o decorrer do tempo...

Conhecendo o passado (n períodos passados) pretendemos calcular o Prémio Puro para o período a seguir $(n+1)$, para o risco j. De acordo com o que foi referido no Capítulo 2, o Prémio Puro do risco j é

$$\mu_{n+1}(\theta_j) = \mathrm{E}[X_{n+1,j}|\theta_j]\,.$$

Consideremos agora que $\theta_1, \theta_2, \cdots, \theta_N$, são variáveis aleatórias independentes e identicamente distribuídas, geradas pela distribuição *a priori* $U(\theta)$. À função $U(\theta)$ chamamos Função Estrutural (os parâmetros θ_j são os Parâmetros de Estrutura).

Fica assim resolvido o *aparente paradoxo* em que os riscos são diferentes mas semelhantes em certo sentido. São semelhantes no sentido de que $\theta_1, \theta_2, \cdots, \theta_N$, são v.a.'s identicamente distribuídas, e diferentes no sentido em que podem tomar valores distintos.

Resumimos as hipóteses, apresentadas por Bühlmann (1967) a propósito do seu modelo linear, nas seguintes:

B1 - Dado θ_j, as v.a.'s $X_{1j}, X_{2j}, \cdots, X_{nj}, X_{n+1,j}$ são i.i.d.

B2 - Os pares $(\underline{X}_1, \theta_1), (\underline{X}_2, \theta_2), \cdots, (\underline{X}_N, \theta_N)$ são i.i.d. $\theta_1, \theta_2, \cdots, \theta_N$, são v.a.'s i.i.d. com f.d. $U(\theta)$.

Estas são as hipóteses que correspondem ao modelo de Bühlmann que veremos mais à frente. Veremos no entanto que este aparece na sequência dos modelos puramente Bayesianos e vão servir para a apresentação e explicação destes.

A hipótese B1 expressa a independência condicionada dentro do mesmo risco. Isto é, dada uma característica θ_j, as indemnizações agregadas do risco j são independentes de período para período. Expressa também homogeneidade temporal: a distribuição de X_{ij}, condicionada por θ_j, não depende de i. A hipótese B2 expressa independência entre riscos e *equivalência exterior* dos riscos.

Uma vez que a distribuição X_{ij} condicionada por θ_j não depende de i, então o prémio puro a tarifar para o período $(n+1)$ vem

$$\mu_{n+1}(\theta_j) = \mu(\theta_j).$$

O prémio $\mu(\theta_j)$ vai ser ser estimado em função das observações nos n períodos passados. Uma vez que pela hipótese B2 os riscos são independentes, pode desprezar-se a informação colateral e consideramos apenas as observações do risco j.

Seja $m(\underline{X}_j)$ uma função do vector $\underline{X}_j = (X_{1j}, X_{2j}, \cdots, X_{nj})$, um estimador baseado na experiência do risco j nos n períodos anteriores, para os quais se possui experiência de indemnizações. Como o Prémio Puro $\mu(\theta_j)$ é em geral desconhecido, os Bayesianos tentaram a *aproximação* de $\mu(\theta_j)$ através de $m(\underline{X}_j)$, utilizando o critério do Erro Quadrático Médio,

$$\mathrm{E}([\mu(\theta_j) - m(\underline{X}_j)]^2)\,. \tag{3}$$

O problema pode ser apresentado da forma que segue.

Definição 3 *Chamemos Prémio Puro de Credibilidade ao prémio* $\overset{*}{m}(\underline{X}_j)$ *que minimiza o erro quadrático médio dado por (3).*

Usando a propriedade iterativa do valor esperado, de (3) temos

$$\mathrm{E}([\mu(\theta_j) - m(\underline{X}_j)]^2) = \mathrm{E}[\mathrm{E}([\mu(\theta_j) - m(\underline{X}_j)]^2|\underline{X}_j)]. \tag{4}$$

A expressão (4) corresponde a

$$\int_{\mathcal{X}} \left(\int_{\Theta} [\mu(\theta_j) - m(\underline{x}_j)]^2 dU(\theta|\underline{x}_j) \right) dF(\underline{x}_j).$$

Como a densidade $dF(\underline{x}_j) \geq 0$, a minimização de (3) equivale a minimizar

$$\mathrm{E}[(\mu(\theta_j) - m(\underline{X}_j))^2|\underline{X}_j] = \mathrm{V}[\mu(\theta_j)|\underline{X}_j] + (\mathrm{E}[\mu(\theta_j)|\underline{X}_j] - m(\underline{X}_j))^2. \tag{5}$$

Donde o valor que minimiza (3) é

$$\overset{*}{m}(\underline{X}_j) = \mathrm{E}[\mu(\theta_j)|\underline{X}_j], \tag{6}$$

ou seja, $\overset{*}{m}(\underline{X}_j)$ é o Estimador Bayes relativamente à função de Perca Quadrática e função *a priori* $U(\theta)$.

A expressão (6) vem explicitamente aplicando o teorema de Bayes

$$\mathrm{E}[\mu(\theta_j)|\underline{X}_j] = \int_\Theta \mu(\theta_j) dU(\theta|\underline{x}_j) = \int_\Theta \mu(\theta_j) \frac{dF(\underline{x}_j|\theta_j)dU(\theta)}{\int_\Theta dF(\underline{x}_j|\theta_j)dU(\theta)} = \frac{\int_\Theta \mu(\theta_j)\prod_{i=1}^n dF(x_{ij}|\theta_j)dU(\theta)}{\int_\Theta \prod_{i=1}^n dF(x_{ij}|\theta_j)dU(\theta)}. \tag{7}$$

Na última expressão atentendemos à hipótese de independência condicionada de B1.

Em teoria o prémio $\overset{*}{m}(\underline{X}_j)$ será um bom substituto do *verdadeiro* Prémio Puro $\mu(\theta_j)$. Na prática não será bem assim, já que muitas vezes (a maior parte!) chegamos a expressões complicadas ou não explícitas e portanto *intratáveis*. O seguinte exemplo retirado de Norberg (1979) é sugestivo:

Exemplo 6 *Considere-se o risco j. Suponha-se que X_{ij} ~Binomial$(1;\ \theta_j)$ $(i = 1, 2, \cdots, n, n+1)$ com função de probabilidade*

$$\mathrm{f}(x_{ij}|\theta_j) = \theta_j^{x_{ij}}(1-\theta_j)^{1-x_{ij}}, \quad x_{ij} = 0,\ 1; \quad 0 < \theta_j < 1.$$

O Prémio Puro a tarifar para o período $(n+1)$ é $\mu(\theta_j) = \theta_j$. Tome-se para distribuição a priori $U(\theta)$, uma distribuição Uniforme de parâmetros α e β, com função densidade:

$$\mathrm{u}(\theta) = \begin{cases} \frac{1}{\beta-\alpha} & 0 < \alpha < \theta < \beta < 1 \quad (\beta > \alpha) \\ 0 & \text{outros} \end{cases}$$

Neste caso o estimador $\overset{*}{m}(\underline{x}_j) = \mathrm{E}[\theta_j|\underline{x}_j]$ vem,

$$\mathrm{E}[\theta_j|\underline{x}_j] = \frac{\sum_{k=1}^{n-n\hat{\theta}_j}(-1)^k \frac{\beta^{n\hat{\theta}_j+k+2}-\alpha^{n\hat{\theta}_j+k+2}}{(n-n\hat{\theta}_j-k)!k!(n\hat{\theta}_j+k+2)}}{\sum_{k=1}^{n-n\hat{\theta}_j}(-1)^k \frac{\beta^{n\hat{\theta}_j+k+1}-\alpha^{n\hat{\theta}_j+k+1}}{(n-n\hat{\theta}_j-k)!k!(n\hat{\theta}_j+k+1)}},$$

sendo $\hat{\theta}_j = \frac{1}{n}\sum_{i=1}^n X_{ij}$. ◊

Utilizando distribuições simples obtém-se para a média *a posteriori* uma expressão que poderá ser chamada de tudo menos simples, o que leva com certeza a que os práticos evitem a sua utilização. Imagine-se a existência de um colectivo composto com grande número de riscos, o que acontece frequentemente.

Se para cada risco se tivesse de calcular um prémio $\mathrm{E}[\theta_j|\underline{x}_j]$ semelhante ao do exemplo acima, o custo e a morosidade da tarifação poderia ser significativo, apesar das grandes potencialidades existentes hoje em dia pelos meios informáticos.

O ganho obtido na minimização do erro poderia perder-se pela *impraticabilidade*. Ora, para uma companhia seguradora tem todo o interesse que o estimador seja *bom* teoricamente, mas também simples de calcular, interpretar e explicar.

Uma maneira de ultrapassar esta dificuldade consiste em utilizar para distribuição *a priori* uma distribuição conjugada da função de verosimilhança, o que permite uma passagem *suave* da distribuição *a priori* para a distribuição *a posteriori*. Será no entanto *forçar*, eventualmente sem explicação evidente, certas distribuições. Concretize-se o exposto nos seguintes exemplos.

Exemplo 7 *Modelo Bernoulli-Beta*

Seja X_{ij} ~Binomial$(1;\ \theta_j)$, $(i = 1, 2, \cdots, n)$ com f.p.

$$f(x_{ij}|\theta_j) = \theta_j^{x_{ij}}(1-\theta_j)^{1-x_{ij}}; \quad 0 < \theta_j < 1,$$

com média $\mu(\theta_j) = \theta_j$.

Suponha-se que $U(\theta) \equiv$*Beta*(α, β) *com f.d.p.*

$$u(\theta) = \frac{1}{B(\alpha,\beta)}\theta^{\alpha-1}(1-\theta)^{\beta-1}; \quad 0<\theta<1 \quad \alpha,\,\beta>0,$$

sendo

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

Obtém-se a função de verosimilhança

$$L(\theta_j) = \prod_{i=1}^{n} f(X_{ij}|\theta_j) = \theta_j^{\sum_{i=1}^{n} x_{ij}}(1-\theta_j)^{n-\sum_{i=1}^{n} x_{ij}}.$$

A respectiva distribuição a posteriori vem

$$u(\theta|\underline{x}_j) = \frac{L(\theta)u(\theta)}{\int_0^1 L(\theta)u(\theta)d\theta} = \frac{1}{B(\sum_i x_{ij}+\alpha;\ n+\alpha-\sum_i x_{ij})}\theta^{\sum_i x_{ij}+\alpha-1}(1-\theta)^{n+\beta-\sum_i x_{ij}-1},$$

ou seja $U(\theta|\underline{x}_j) \equiv$*Beta*$(\sum_i x_{ij}+\alpha;\ n+\beta-\sum_i x_{ij})$. *A média a posteriori é dada pela expressão*

$$\mathrm{E}[\mu(\theta)|\underline{x}_j] = \mathrm{E}[\theta|\underline{x}_j] = \frac{\sum_i x_{ij}+\alpha}{\alpha+\beta+n}.$$

A expressão da média a posteriori é simples porque a distribuição Beta é conjugada da Binomial. Se se fizer um pequeno arranjo na expressão acima $\mathrm{E}[\mu(\theta)|\underline{x}_j]$ *pode ser escrito na forma:*

$$\mathrm{E}[\mu(\theta_j)|\underline{x}_j] = \frac{n}{\alpha+\beta+n}\bar{x}_{.j} + \frac{\alpha}{\alpha+\beta+n} = \frac{n}{\alpha+\beta+n}\bar{x}_{.j} + \frac{\alpha+\beta}{\alpha+\beta+n}\mu$$

em que $\bar{x}_{.j} = \frac{1}{n}\sum_{i=1}^{n} x_{ij}$ *e a média colectiva* $\mu = \mathrm{E}[X_{ij}] = \mathrm{E}[\theta] = \frac{\alpha}{\alpha+\beta}$. *Ou seja,*

$$\mathrm{E}[\mu(\theta_j)|\underline{x}_j] = z\bar{x}_{.j} + (1-z)\mu$$

com

$$z = \frac{n}{n+k} \ \text{ e } \ k = \alpha+\beta.$$

A média a posteriori tem, a Forma de Credibilidade (1) e o Factor de Credibilidade (2), o que torna este resultado interessante.

No entanto tem limitações óbvias, pois apesar da família de distribuições Beta ser bastante larga, existem inúmeras distribuições que não podem ser representadas por uma Beta. ♢

Exemplo 8 *Modelo Poisson-Gama*

Se $X_{ij} \sim$*Poisson*(θ_j), $i = 1, 2, \cdots, n$, *com função de probabilidade*

$$f(x_{ij}|\theta_j) = \frac{e^{-\theta_j}\theta_j^{x_{ij}}}{x_{ij}!}, \quad x_i = 0, 1, 2, \cdots;\ \theta_j > 0$$

com média $\mu(\theta_j) = \theta_j$. *Suponha-se que* $U(\theta) \equiv$*Gama*(α, β) *com f.d.p.*

$$u(\theta) = \frac{\beta^{\alpha}}{\Gamma(\alpha)}e^{-\beta\theta}\theta^{\alpha-1}; \quad \theta>0 \ \text{ e } \ \alpha,\beta>0\,.$$

A função de verosimilhança vem,

$$L(\theta_j) = \prod_{i=1}^{n} f(x_{ij}|\theta_j) = \frac{e^{-n\theta_j}\theta_j^{\sum x_{ij}}}{\prod_{i=1}^{n} x_{ij}!}\ .$$

Obtém-se a distribuição a posteriori

$$
\begin{aligned}
u(\theta|\underline{x}_j) &= \frac{L(\theta)u(\theta)}{\int_0^{+\infty} L(\theta)u(\theta)d\theta} = \frac{(n+\beta)^{\sum x_{ij}+\alpha}}{\Gamma(\sum x_{ij}+\alpha)} \exp\{-(n+\beta)\theta\}\theta^{\sum x_{ij}+\alpha-1} \\
&\equiv \mathit{Gama}\left(\sum_i x_{ij}+\alpha;\, n+\beta\right)
\end{aligned}
$$

Obtém-se então a média a posteriori

$$
E[\mu(\theta)|\underline{x}_j] = E[\theta|\underline{x}_j] = \tfrac{\sum_i x_{ij}+\alpha}{\beta+n} = \tfrac{n}{n+\beta}\bar{x}_{.j} + \tfrac{\alpha}{n+\beta} = \tfrac{n}{n+\beta}\bar{x}_{.j} + \tfrac{\beta}{n+\beta}\mu\,,
$$

sendo o prémio colectivo $\mu = \mathrm{E}[X_{ij}] = \alpha/\beta$.

A média a posteriori tem também a Forma de Credibilidade (1) e o Factor de Credibilidade z de acordo com (2):

$$
\mathrm{E}[\mu(\theta)|\underline{x}_j] = z\overline{x}_{.j} + (1-z)\mu
$$

com $z = \frac{n}{n+k}$ *e* $k = \beta$. ♢

Exemplo 9 *Modelo Exponencial-Gama*

Seja $X_{ij} \sim$*Gama*$(1;\, \theta_j)$, $(i = 1, 2, \cdots, n)$ *com f.d.p.*

$$
f(x_{ij}|\theta_j) = \theta_j e^{\theta_j x_{ij}}; \quad x_{ij} > 0 \quad (\theta_j > 0),
$$

e média $\mu(\theta_j) = 1/\theta_j$.

Suponha-se que $U(\theta) \equiv$*Gama*(α, β), *cuja f.d.p. é*

$$
u(\theta) = \frac{\beta^\alpha}{\Gamma(\alpha)} e^{-\beta\theta}\theta^{\alpha-1}; \qquad \theta > 0 \;\; e \;\; \alpha, \beta > 0
$$

obtendo-se para a função de verosimilhança

$$
L(\theta_j) = \prod_{i=1}^n f(x_{ij}|\theta_j) = \theta_j^n \exp\{-\theta_j \sum x_{ij}\}.
$$

Daqui se obtém a distribuição a posteriori

$$
\begin{aligned}
u(\theta|\underline{x}_j) &= \frac{L(\theta)u(\theta)}{\int_0^{+\infty} L(\theta)u(\theta)d\theta} = \frac{(\beta+\sum_i x_{ij})^{n+\alpha}}{\Gamma(n+\alpha)} \exp\{-\theta(\beta+\sum_i x_{ij})\}\theta^{n+\alpha-1} \\
&\equiv \mathit{Gama}(n+\alpha;\, \beta+\sum_i x_{ij})\,.
\end{aligned}
$$

Por outro lado, a média colectiva $\mu = \mathrm{E}[X_{ij}] = \mathrm{E}[1/\theta]$:

$$
\mu = \frac{\beta^\alpha}{\Gamma(\alpha)} \int_0^{+\infty} e^{-\beta\theta}\theta^{\alpha-2}d\theta = \beta\frac{\Gamma(\alpha-1)}{\Gamma(\alpha)} = \frac{\beta}{\alpha-1}\,, \;\; \mathit{para} \;\; \alpha > 1\,,
$$

atendendo a que $\Gamma(\alpha-1)/\Gamma(\alpha) = 1/(\alpha-1)$. *A média a posteriori vem*

$$
\begin{aligned}
\mathrm{E}[\mu(\theta)|\underline{x}_j] &= \frac{(\beta+\sum_{i=1}^n x_{ij})^{n+\alpha}}{\Gamma(n+\alpha)} \int_0^{+\infty} e^{-(\beta+\sum_i x_{ij})\theta}\theta^{n+\alpha-2}d\theta \\
&= \frac{(\beta+\sum_i x_{ij})\Gamma(n+\alpha-1)}{\Gamma(n+\alpha)} = \frac{\beta+\sum_i x_{ij}}{n+\alpha-1}
\end{aligned}
$$

atendendo a que $\Gamma(n-1+\alpha)/\Gamma(n+\alpha) = (n+\alpha)^{-1}$. *Tem-se ainda*

$$\mathrm{E}[\mu(\theta)|\underline{x}_j] = \frac{n}{n+\alpha-1}\bar{x}_{\cdot j} + \frac{\beta}{n+\alpha-1} = \frac{n}{n+\alpha-1}\bar{x}_{\cdot j} + \frac{\alpha-1}{n+\alpha-1}\mu$$

A média a posteriori tem a Forma de Credibilidade (1) e o Factor de Credibilidade z de acordo com (2)

$$\mathrm{E}[\mu(\theta)|\underline{x}_j] = z\bar{x}_{\cdot j} + (1-z)\mu$$

com $z = \frac{n}{n+k}$ *e* $k = \alpha - 1$. $\diamondsuit$

Exemplo 10 *Modelo Normal-Normal, com variância conhecida*

Seja X_{ij} $\sim$*Normal*(θ_j, σ), *com* $\sigma > 0$, *conhecido e* θ $\sim$*Normal*(μ, τ), μ *e* τ *conhecidos. Obtém-se para distribuição a posteriori* $U(\theta|\underline{x}_j)$ *uma distribuição*

$$Normal\left(\frac{1}{\rho_n}\left[\frac{1}{\tau^2}\mu + \frac{n}{\sigma^2}\bar{x}_{\cdot j}\right];\ \rho_n^{-1}\right)$$

com $\rho_n = \frac{1}{\tau^2} + \frac{n}{\sigma^2} = \frac{\sigma^2 + n\tau^2}{(\sigma\tau)^2}$, *(veja-se Murteira (1988)).*

Sendo assim tem-se $\mu(\theta_j) = \theta_j$ *e*

$$\mathrm{E}[\mu(\theta_j)|\underline{x}_j] = \mathrm{E}[\theta_j|\underline{x}_j] = \frac{1}{\rho_n}\left(\frac{1}{\tau^2}\mu + \frac{n}{\sigma^2}\bar{x}_{\cdot j}\right) = \frac{n}{\rho_n\sigma^2}\bar{x}_{\cdot j} + \frac{1}{\rho_n\sigma^2}\mu = \frac{n\tau^2}{\sigma^2+n\tau^2}\bar{x}_{\cdot j} + \frac{\sigma^2}{\sigma^2+n\tau^2}\mu$$

A média a posteriori tem a Forma de Credibilidade (1) e o Factor de Credibilidade z da forma (2):

$$\mathrm{E}[\mu(\theta_j)|\underline{x}_j] = z\bar{x}_{\cdot j} + (1-z)\mu$$

com $z = n/(n+k)$ *e* $k = \sigma^2/\tau^2$. $\diamondsuit$

Nos Exemplos 7-10 escolheram-se para distribuições *a priori*, distribuições conjugadas das funções de verosimilhança, em cada caso. Em todas os casos obtiveram-se expressões para a média *a posteriori* bastante simples e lineares nas observações. O outro aspecto interessante do ponto de vista da Teoria da Credibilidade, é que todas estas expressões apresentam a forma de credibilidade de Whitney, que carecia de suporte teórico.

A média *a posteriori* é uma média ponderada entre a média empírica do registo de indemnizações do risco, e a média colectiva. O factor de ponderação z tem em conta a quantidade de informação disponível n, e as características específicas do risco através da constante k.

Em todos os casos, o factor de ponderação z não depende do risco em questão, em virtude de se ter considerado para todos os riscos do colectivo um n igual. A individualização do Prémio Puro a tarifar é obtido através da média empírica $\bar{X}_{\cdot j}$. Note-se ainda que é dado igual peso de credibilidade às indemnizações passadas em cada período, em virtude de se ter considerado que a característica específica não evolui ao longo do tempo, i.e., não depende de i. Daí dar-se a mesma importância às indemnizações, tenham elas acontecido há 1 ou 20 anos, por exemplo.

Jewell (1974) mostrou que se chega à forma de credibilidade se se escolher para X_{ij} uma distribuição da *Família Exponencial* dotada de parametrização natural, e para a qual $\bar{X}_{\cdot j}$ é estatística suficiente, i.e., se se escolher uma distribuição da forma

$$f(x_{ij}|\theta_j) = c(\theta_j)^{-1}e^{-\theta_j x_{ij}}H(x_{ij}), \quad i = 1, 2, \cdots, n$$

sendo $c(\theta_j)$ uma constante de normalização, i.e.

$$c(\theta_j) = \int_{\mathcal{A}} e^{-\theta_j x}H(x)dx \text{ ou } c(\theta_j) = \sum_{\mathcal{A}} e^{-\theta_j x}H(x),$$

consoante se trate do caso contínuo ou discreto, respectivamente.

Então a função de verosimilhança vem

$$\begin{aligned} L(\theta_j) &= \prod_{i=1}^{n} f(x_{ij}|\theta_j) = [c(\theta_j)]^{-n} \exp\{-\theta_j \sum_i x_{ij}\} \prod_{i=1}^{n} H(x_{ij}) \\ &= \exp\{-n \log c(\theta_j) - \theta_j \sum_i x_{ij}\} \prod_{i=1}^{n} H(x) \\ &\propto \exp\{-n \ln c(\theta_j) - \theta_j \sum_i x_{ij}\}. \end{aligned}$$

Tomando para distribuição *a priori* $u(\theta)$

$$u(\theta) = u(\theta|\alpha, \beta) \propto \exp\{-\alpha \ln c(\theta) - \beta\theta\} = [c(\theta)]^{-\alpha} e^{-\beta\theta}$$

i.e.

$$u(\theta) = \frac{[c(\theta)]^{-\alpha} e^{-\beta\theta}}{d(\alpha, \beta)}$$

obtém-se uma distribuição *a posteriori* do mesmo tipo:

$$u(\theta|\underline{x}_j) = u(\theta|\alpha + n, \beta + \sum_i x_{ij}) \propto [c(\theta)]^{-\alpha+n} e^{-(\beta+\sum_i x_{ij})\theta}$$

(para mais detalhes veja-se Murteira (1988)).

Alternativamente, podemos considerar considerar a reformulação proposta por Gerber (1995) a este modelo geral, que permite uma *identificação* mais simples dos casos particulares apresentados nos Exemplos 7-10. Isto é considerar que $f(x_{ij}|\theta_j)$ é da forma

$$f(x_{ij}|\theta_j) = \frac{a(x_{ij}) b(\theta_j)^{x_{ij}}}{c\,(\theta_j)}, \quad x_{ij} \in \mathcal{A}, \quad i = 1, 2, \cdots, n,$$

onde $\mathcal{A}$ é o conjunto de valores possíveis das indemnizações (contínuas ou discretas) e $c\,(\theta_j)$ é uma constante de normalização:

$$c(\theta_j) = \int_{\mathcal{A}} a(x) b(\theta_j)^x dx \; .$$

Se escolhermos para distribuição *a priori*

$$u(\theta; \alpha, \beta) = \frac{c(\theta)^{-\alpha} b(\theta)^{\beta} b'(\theta)}{d(\alpha, \beta)},$$

onde θ varia num determinado intervalo Θ, $d(\alpha, \beta)$ é a constante de normalização, e α e β são dois parâmetros. Obtemos facilmente que a distribuição *a posteriori* é da mesma família (ver acima): $u(\theta|\underline{x}_j) = u(\theta|\alpha+n, \beta + \sum_i x_{ij})$.

Com esta formulação é simples verificar que

$$\begin{aligned} \mathrm{E}[X_{ij}|\theta_j] = \mu(\theta_j) &= \int_{\mathcal{A}} x a(x) b(\theta_j)^x dx = \frac{b(\theta_j)}{b'(\theta_j)} \frac{c'(\theta_j)}{c(\theta_j)} \\ \mu = \mathrm{E}[\mu(\theta)] &= \int_{\Theta} \mu(\theta) u(\theta) d\theta = d(\alpha, \beta)^{-1} \int_{\Theta} c(\theta)^{-(\alpha+1)} c'(\theta) b(\theta)^{\beta+1} d\theta \\ &= \frac{\beta+1}{\alpha d(\alpha, \beta)} \int_{\Theta} c(\theta)^{-\alpha} b(\theta)^{\beta} b'(\theta) d\theta = \frac{\beta+1}{\alpha}, \end{aligned}$$

integrando por partes e assumindo que $c(\theta)^{-\alpha} b(\theta)^{\beta+1} \to 0$ nos limites de integração (o que acontece mediante certas condições de regularidade, ver Jewell (1975b)). Então o prémio para o período $n+1$ vem

$$\mathrm{E}[\mu(\theta)|\underline{x}_j] = \frac{\beta + \sum_i x_{ij} + 1}{\alpha + n} = \frac{\beta+1}{\alpha+n} + \frac{\sum_i x_{ij}}{\alpha+n} = \frac{n}{\alpha+n} \overline{x}_{.j} + \frac{\alpha}{\alpha+n} \mu \; ,$$

ou seja,

$$\mathrm{E}[\mu(\theta)|\underline{x}_j] = z\,\overline{x}_{.j} + (1-z)\mu$$

com $z = n/(n+\alpha)$, ou seja, obtemos a Forma de Credibilidade (1).

Exemplo 11 *A partir a formulação acima os exemplos clássicos referidos atrás (Exemplos 7-10) podem ser obtidos com facilidade:*

Bernoulli-Beta: $\mathcal{A} = \{0,1\}$, $0 < \theta < 1$, $a(x) = 1$, $b(\theta) = \theta\,(1-\theta)^{-1}$, $c(\theta) = (1-\theta)^{-1}$ *e*

$$u(\theta;\alpha,\beta) = \frac{\Gamma(\beta)}{\Gamma(\beta+1)\,\Gamma(\alpha-\beta-1)}\theta^{\beta}\,(1-\theta)^{\alpha-\beta-2}\ .$$

Poisson-Gama: $\mathcal{A} = \{0,1,2,...\}$, $0 < \theta < \infty$, $a(x) = 1/x!$, $b(\theta) = \theta$, $c(\theta) = \exp\{\theta\}$ *e*

$$u(\theta;\alpha,\beta) = \frac{\alpha^{\beta+1}}{\Gamma(\beta+1)}\theta^{\beta}e^{-\alpha\theta}\ .$$

Exponencial-Gama: $\mathcal{A} = (0,\infty)$, $0 < \theta < \infty$, $a(x) = 1$, $b(\theta) = \exp\{\theta\}$, $c(\theta) = \theta^{-1}$ *e*

$$u(\theta;\alpha,\beta) = \frac{\alpha^{\beta+1}}{\Gamma(\beta+1)}\theta^{\beta}e^{-\alpha\theta}\ .$$

Normal-Normal: $\mathcal{A} = (-\infty,\infty)$, $-\infty < \theta < \infty$, $a(x) = \exp\left\{-x^2/2\sigma^2\right\}$, $b(\theta) = \exp\left\{\theta/\sigma^2\right\}$, $c(\theta) = \sqrt{2\pi}\sigma\exp\left\{\theta^2/2\sigma^2\right\}$ *e*

$$u(\theta;\mu,\tau^2) = \frac{1}{\sqrt{2\pi}\tau}\exp\left\{-\left(\frac{(\theta-\mu)^2}{2\tau^2}\right)\right\}\ ,\ com\ \mu = \frac{\beta+1}{\alpha},\ \tau^2 = \frac{\sigma^2}{\alpha}\ .$$

Beta-Geométrica: $\mathcal{A} = \{0,1,2,...\}$, $0 < \theta < 1$, $a(x) = 1$, $b(\theta) = (1-\theta)$, $c(\theta) = \theta^{-1}$ e

$$u(\theta;\alpha,\beta) = \frac{\Gamma(\alpha+\beta+2)}{\Gamma(\alpha+1)\,\Gamma(\beta+1)}\theta^{\alpha}\,(1-\theta)^{\beta}\ .$$

. ♢

5 Estimador de Credibilidade

As fórmulas para a média *a posteriori* dos modelos contidos nos Exemplos 7-10 e 11, assim como no trabalho de Jewell (1974) apresentados na Secção 4 têm a vantagem de serem simples de calcular e de interpretar. No entanto isto acontece com escolhas especiais para as distribuições *a priori*. De uma forma geral, a média *a posteriori* $\mathrm{E}[\mu(\theta_j)|\underline{x_j}]$, tem expressões pouco *simpáticas* e de difícil interpretação. Os exemplos referidos acima mostram uma característica comum e fundamental. Os estimadores obtidos são lineares (nas observações). Este é um aspecto fundamental no desenvolvimento da teoria da credibilidade. Como em geral a partir dos modelos bayesianos puros os estimadores encontrados não são lineares, surgiu então a hipótese de arranjar soluções dentro da classe dos estimadores lineares para qualquer tipo de modelo. Foi o que fez Bühlmann (1967) para um modelo particular e sob as hipóteses da Secção 4. É um modelo pioneiro, trouxe de facto um grande avanço à Teoria da Credibilidade, e ao problema da tarifação, pois veio permitir, com o seu método, o tratamento de modelos mais sofisticados e de fácil aplicação. A análise puramente Bayesiana dificilmente o poderia conseguir. O aspecto relevante é a escolha de estimadores lineares para o verdadeiro prémio puro em determinado período, utilizando o critério do Erro Quadrático Médio Mínimo.

Apresenta-se seguidamente uma definição geral de estimador de credibilidade, que vai permitir o tratamento de modelos mais sofisticados, que não necessariamente aqueles desenvolvidos sob as hipóteses relativamente restritivas apresentadas no capítulo anterior. Seguem-se alguns resultados a serem utilizados posteriormente.

Sejam uma variável aleatória m e $\underline{X} = (X_1, \cdots, X_t)$ um vector aleatório observável, conjuntamente distribuídos. Chame-se $\hat{m}$ um estimador linear de m baseado em $\underline{X}$, se $\hat{m}$ pode ser escrito na forma:

$$\hat{m} = a_0 + \underline{a}'\underline{X} = a_0 + \sum_{i=1}^{t} a_i X_i \tag{8}$$

onde a_0 é uma constante e $\underline{a} = (a_1, \cdots, a_t)$, um vector t-dimensional, também de constantes.

Definição 4 *Suponha-se a existência dos momentos de primeira e segunda ordens. Define-se $\tilde{m}$, Estimador de Credibilidade da variável aleatória m, como sendo o estimador linear da forma (8), que minimiza a função de perca quadrática esperada,* $\mathrm{E}\left[(m - \hat{m})^2\right]$.

Uma vez definido estimador de credibilidade vai estudar-se a sua obtenção de forma geral, e a unicidade da sua solução. Considere-se então o seguinte teorema.

Teorema 1 *a) Um estimador linear da forma (8) é um Estimador de Credibilidade se satisfaz as seguintes Condições de Normalidade,*

$$\begin{aligned} \mathrm{E}[m] &= \mathrm{E}[\hat{m}] & (9)\\ \mathrm{Cov}[m, \underline{X}'] &= \mathrm{Cov}[\hat{m}, \underline{X}'] & (10) \end{aligned}$$

b) Se os estimadores $\dot{m}$ e $\ddot{m}$ da forma (8) verificam (9) e (10), então eles são iguais quase certamente.

c) O estimador de credibilidade $\tilde{m}$ satisfaz então

$$\mathrm{Cov}[m, \tilde{m}] = \mathrm{V}[m] - \mathrm{V}[m - \tilde{m}] \tag{11}$$

Dem.

a) Seja
$$R = \mathrm{E}\left[(m - \hat{m})^2\right] = \mathrm{E}\left[(m - a_0 - \underline{a}'\underline{X})^2\right].$$
Derivando R em ordem a a_0 e a $\underline{a}$ e igualando a zero, obtém-se
$$\begin{aligned} \frac{\partial R}{\partial a_0} &= -2\,\mathrm{E}[m - \hat{m}] = 0 & (12)\\ \frac{\partial R}{\partial \underline{a}} &= -2\,\mathrm{E}[(m - \hat{m})\underline{X}'] = \underline{0}'. & (13) \end{aligned}$$
De (12) obtém-se imediatamente (9). Pós-multiplicando (12) por $\mathrm{E}[\underline{X}']$ e subtraindo a (13) obtém-se
$$\mathrm{E}[(m - \hat{m})\underline{X}'] - \mathrm{E}[m]\,\mathrm{E}[\underline{X}'] + \mathrm{E}[\hat{m}]\,\mathrm{E}[\underline{X}'] = \underline{0}'\ ,$$
desenvolvendo o primeiro valor esperado verifica-se facilmente que a igualdade acima é equivalente a $\mathrm{Cov}[m, \underline{X}'] = \mathrm{Cov}[\hat{m}, \underline{X}']$, ficando completa a demonstração de a).

b) Considerem-se dois estimadores $\dot{m}$ e $\ddot{m}$ da forma (8) e suponha-se que $\dot{m}$ verifica (9) e (10). Calcule-se
$$\begin{aligned} \mathrm{E}[(m - \ddot{m})^2] &= \mathrm{E}\left[\left(m - \dot{m} + \dot{m} - \ddot{m}\right)^2\right] \\ &= \mathrm{E}[(m - \dot{m})^2] + \mathrm{E}[(\dot{m} - \ddot{m})^2] + 2\,\mathrm{E}[(m - \dot{m})(\dot{m} - \ddot{m})] \end{aligned}$$

Da terceira parcela tem-se

$$\mathrm{E}[(m - \dot{m})(\dot{m} - \ddot{m})] = \mathrm{Cov}[(m - \dot{m}), (\dot{m} - \ddot{m})] \tag{14}$$

porque por (9) tem-se $\mathrm{E}[m - \dot{m}] = 0$. Como $\dot{m} - \ddot{m}$ por hipótese são da forma (8), então $\dot{m} - \ddot{m}$ pode escrever-se na forma $\dot{m} - \ddot{m} = d_0 + \underline{d}'\underline{X}$, em que d_0 é uma constante e $\underline{d}$ é um vector de constantes, e nesse caso (14) é igual a $\mathrm{Cov}[m - \dot{m}, \underline{X}']\,\underline{d}$. Por (10) tem-se

$$\mathrm{Cov}[m - \dot{m}, \underline{X}']\underline{d} = \left(\mathrm{Cov}[m, \underline{X}'] - \mathrm{Cov}[\dot{m}, \underline{X}']\right)\underline{d} = \underline{0}'.$$

Então $\mathrm{E}[(m-\ddot{m})^2] = \mathrm{E}[(m-\dot{m})^2] + \mathrm{E}[(\dot{m}-\ddot{m})^2]$, ou seja $\mathrm{E}[(m-\ddot{m})^2] \geq \mathrm{E}[(m-\dot{m})^2$ e $\mathrm{E}[(m-\ddot{m})^2] = \mathrm{E}[(m-\dot{m})^2]$ se $\mathrm{E}[(\dot{m}-\ddot{m})^2] = 0$. Isto é, se e só se $\dot{m}$ e $\ddot{m}$ forem iguais quase certamente. Fica provada a aliena b) do teorema.

c) Para provar a última alínea, tem-se que

$$\mathrm{V}[\tilde{m}] = \mathrm{Cov}[\tilde{m}, \tilde{m}] = \mathrm{Cov}[\tilde{m}, a_0 + \underline{X}'\underline{a}] = \mathrm{Cov}[\tilde{m}, \underline{X}']\underline{a}.$$

como por (10), $\mathrm{Cov}[m, \underline{X}'] = \mathrm{Cov}[\tilde{m}, \underline{X}']$, então $\mathrm{V}[\tilde{m}] = \mathrm{Cov}[m, a_0 + \underline{X}'\underline{a}] = \mathrm{Cov}[m, \tilde{m}]$. Fica provada a primeira parte de (11).

Por outro lado, $\mathrm{V}[m - \tilde{m}] = \mathrm{V}[m] - 2\,\mathrm{Cov}[m, \tilde{m}] + \mathrm{V}[\tilde{m}]$, e pela primeira igualdade de (11), obtém-se finalmente $\mathrm{V}[\tilde{m}] = \mathrm{V}[m] - \mathrm{V}[m - \tilde{m}]$. □

O Teorema 1 mostra que o valor das constantes a_0 e a_k $(k = 1, 2, \cdots, t)$ que tornam o erro quadrático médio mínimo de estimação, são obtidos através das condições de normalidade (9) e (10). Da condição (8) vê-se que o estimador de credibilidade assim calculado é em média igual à média da variável estimada, característica que aparece *naturalmente* em virtude de se considerar no estimador um termo independente de $\underline{X}$.

Seja τ o Erro Quadrático Médio de Estimação, ou mais simplesmente Erro de Estimação, i.e.

$$\tau = \mathrm{E}[(\tilde{m} - m)^2].$$

Aproveitando o resultado (11) do Teorema 1, o Erro de Estimação pode ser apresentado na seguinte forma equivalente:

$$\tau = \mathrm{V}[m] - \mathrm{V}[\tilde{m}] = \mathrm{V}[m - \tilde{m}], \tag{15}$$

porque $\mathrm{E}[(\tilde{m} - m)^2] = \mathrm{E}[(\tilde{m} - \mathrm{E}[\tilde{m}] + \mathrm{E}[m] - m)^2]$. Introduzindo (9) tem-se

$$\begin{aligned} \tau &= \mathrm{V}[\tilde{m}] + \mathrm{V}[m] - 2\,\mathrm{Cov}[\tilde{m}, m] \\ &= \mathrm{V}[m] - \mathrm{V}[\tilde{m}] \\ &= \mathrm{V}[m - \tilde{m}], \end{aligned}$$

introduzindo (11).

Nos modelos de credibilidade, é comum supor-se que as indemnizações agregadas dentro do mesmo risco são condicionalmente independentes período a período, uma vez conhecido o parâmetro de risco. Esta hipótese adicional permite apresentar uma visão um pouco diferente no problema da estimação nos modelos de credibilidade. O teorema que a seguir se apresenta contempla esta hipótese adicional. Tem no entanto uma aplicação mais geral. Seja então o seguinte teorema:

Teorema 2 *Seja o vector aleatório* $\underline{X} = (X_1, \cdots, X_t)$ *e a variável aleatória* m. $\underline{X}$ *e* m *são condicionalmente independentes, dado um parâmetro, seja* θ *(que pode ser um vector). Sejam,* $\overset{*}{m}$ *um estimador de* m *baseado em* $\underline{X}$ *e o valor esperado* $\mathrm{E}[m|\theta]$*, e suponha-se a sua existência. Então,*

$$\mathrm{E}[(\overset{*}{m} - m)^2] = \mathrm{E}[(\overset{*}{m} - \mathrm{E}[m|\theta])^2] + \mathrm{E}[\mathrm{V}[m|\theta]] \tag{16}$$

e $\overset{*}{m}$ *é um estimador óptimo para* m*, se e só se é um estimador óptimo para* $\mathrm{E}[m|\theta]$.

Dem.

$$\begin{aligned}
\mathrm{E}[(\overset{*}{m}-m)^2] &= \mathrm{E}[\mathrm{E}[(\overset{*}{m}-m)^2|\theta]] \\
&= \mathrm{E}[(\overset{*}{m}-\mathrm{E}[m|\theta]+\mathrm{E}[m|\theta]-m)^2] \\
&= \mathrm{E}[(\overset{*}{m}-\mathrm{E}[m|\theta])^2]+\mathrm{E}[(\mathrm{E}[m|\theta]-m)^2]-2\,\mathrm{E}[(\overset{*}{m}-\mathrm{E}[m|\theta])(m-\mathrm{E}[m|\theta])].
\end{aligned}$$

Da terceira parcela da igualdade acima tem-se

$$\mathrm{E}[(\overset{*}{m}-\mathrm{E}[m|\theta])(m-\mathrm{E}[m|\theta])] = \mathrm{E}[\mathrm{E}[(\overset{*}{m}-\mathrm{E}[m|\theta])(m-\mathrm{E}[m|\theta])|\theta]],$$

e o valor esperado condicionado, dado θ, vem

$$\mathrm{E}[(\overset{*}{m}-\mathrm{E}[m|\theta])(m-\mathrm{E}[m|\theta])|\theta] = \mathrm{E}[\overset{*}{m}-\mathrm{E}[m|\theta]|\theta] \times \mathrm{E}[m-\mathrm{E}[m|\theta]|\theta] = 0$$

porque m e $\underline{X}$ são condicionalmente independentes, dado θ. Então,

$$\mathrm{E}[(\overset{*}{m}-m)^2] = \mathrm{E}[(\overset{*}{m}-\mathrm{E}[m|\theta])^2]+\mathrm{E}[\mathrm{V}[m|\theta]].$$

Uma vez que $\mathrm{E}[\mathrm{V}[m|\theta]] \geq 0$ não depende do estimador $\overset{*}{m}$, tem-se imediatamente que $\overset{*}{m}$ é um estimador óptimo para m se só se é um estimador óptimo para $\mathrm{E}[m|\theta]$. □

O teorema acima permite concluir que o estimador de credibilidade $\tilde{m}$ obtido pelas condições do Teorema 1, e que minimiza $\mathrm{E}[(m-\overset{*}{m})^2]$, também minimiza $\mathrm{E}[(m-\mathrm{E}[m|\theta])^2]$. Relacionando este resultado com o modelo de Bühlmann, verifica-se que o estimador de credibilidade (34) (estimador óptimo dentro da classe dos estimadores lineares da forma (24), para o verdadeiro Prémio Puro, para o período $(n+1)$, $\mu(\theta)=\mathrm{E}[X_{n+1}|\theta]$) é também estimador óptimo para X_{n+1}. No entanto, o resultado do Teorema 2 é um resultado mais geral, pois permite o tratamento de modelos mais complexos, por exemplo, modelos que não imponham como o modelo de Bühlmann, $\mathrm{E}[X_i|\theta]=\mu(\theta)$, $(i=1,2,\cdots,n,n+1)$.

Além disso, sempre que se verifiquem as condições iniciais do Teorema 2, situação comum na teoria da credibilidade, permite substituir a estimação de $\mathrm{E}[m|\theta]$ por m, de notação mais simples, considerando por exemplo que $\mathrm{E}[m|\theta]$ representa o verdadeiro Prémio Puro relativamente ao período a tarifar. Além do mais, na literatura aparece com frequência a referência a estimação de $\mathrm{E}[m|\theta]$ ou m. Por outro lado, o facto de se ter considerado que θ pode ser eventualmente um vector, vai permitir a análise de modelos em que se aceite a possibilidade de evolução com o tempo da característica específica representada no parâmetro, i.e., θ_{ij} depende de i $(i=1,2,\cdots,n,n+1)$.

Jewell (1975*a*) afirma que nalgumas aplicações pode-se estar interessado em encontrar um estimador de credibilidade, considerando a constante $a_0=0$. Isto é, um estimador $\hat{m}$ da forma

$$\hat{m} = \underline{b}'\underline{X} = \sum_{k=1}^{t} b_t X_t, \tag{17}$$

que minimize a função de perca quadrática esperada

$$\mathrm{E}[(m-\hat{m})^2]. \tag{18}$$

Só que em geral o estimador é enviesado e o erro quadrático médio será maior. Pode-se remover o enviesamento à custa de um erro adicional, impondo a restrição

$$\mathrm{E}[\hat{m}] = \mathrm{E}[m]. \tag{19}$$

Considere-se então o seguinte teorema.

Teorema 3 *Um estimador linear da forma (17), que minimiza (18), sujeito à restrição (19), satisfaz as seguintes condições de normalidade,*

$$\begin{aligned}
\mathrm{Cov}[\hat{\hat{m}},\underline{X}'] - \mathrm{Cov}[m,\underline{X}'] &= \mathrm{Cov}[\hat{\hat{m}},\underline{X}'] - \mathrm{Cov}[\tilde{m},\underline{X}'] = \lambda\mathrm{E}[\underline{X}'] && (20)\\
\mathrm{E}[\hat{m}] &= \mathrm{E}[m] && (21)
\end{aligned}$$

onde $\tilde{m}$ é o estimador de credibilidade definido pelo Teorema 1 e λ é o multiplicador de Lagrange.

Dem.

A Lagrangeana vem,

$$L = \mathrm{E}[(m - \hat{m})^2] + 2\lambda\left(\mathrm{E}[m] - \mathrm{E}[\hat{m}]\right)$$

Derivando em ordem a $\underline{b}$ e λ, e igualando a 0, tem-se

$$\frac{\partial L}{\partial \underline{b}} = -2\,\mathrm{E}[(m - \hat{m})\underline{X}'] - 2\lambda\,\mathrm{E}[\underline{X}'] = \underline{0}' \tag{22}$$

$$\frac{\partial L}{\partial \lambda} = 2\left(\mathrm{E}[m] - \mathrm{E}[\hat{m}]\right) = 0 \tag{23}$$

De (22) sai,

$$\begin{aligned}\mathrm{Cov}[\hat{m}, \underline{X}'] + \mathrm{E}[\hat{m}]\,\mathrm{E}[\underline{X}'] &= \mathrm{Cov}[m, \underline{X}'] + \mathrm{E}[m]\,\mathrm{E}[\underline{X}'] + \lambda\,\mathrm{E}[\underline{X}'] \\ &= \mathrm{Cov}[m, \underline{X}'] + \mathrm{E}[m]\,\mathrm{E}[\underline{X}'] + \lambda\,\mathrm{E}[\underline{X}']\end{aligned}$$

introduzindo (23). A igualdade acima é equivalente a

$$\mathrm{Cov}[\hat{m}, \underline{X}'] = \mathrm{Cov}[m, \underline{X}'] + \lambda\,\mathrm{E}[\underline{X}'] \Leftrightarrow \underline{b}'\mathrm{Cov}[\underline{X}, \underline{X}'] = \mathrm{Cov}[m, \underline{X}'] + \lambda\,\mathrm{E}[\underline{X}'],$$

e finalmente substituindo (10), obtém-se (20). □

Na Secção 6.3.3, a propósito do modelo de regressão de Hachemeister apresentar-se-á um teorema que generaliza o Teorema 1, ou seja o conceito de Estimador de Credibilidade.

6 Modelos empírico-Bayesianos

6.1 Modelo de Bühlmann

6.1.1 Estimador de credibilidade

Como se disse na 4, Bühlmann (1967) trouxe um avanço histórico à Teoria da Credibilidade e ao problema da tarifação, propondo a estimação do prémio puro através duma função linear das indemnizações passadas para cada risco do colectivo. O seu trabalho esclareceu as fórmulas de credibilidade apresentadas por Whitney, e abriu caminho a um grande avanço posterior na Teoria da Credibilidade. Vai-se fazer a dedução do estimador de credibilidade tal como esta é apresentada no trabalho original (ou alternativamente em Bühlmann (1970)) e não aplicando directamente o Teorema 1 da Secção 4.

Concretamente, formulando as hipóteses B1 e B2 apresentadas na Secção 4, propôs estimar $\mu(\theta_j)$ através duma função linear do vector observável $\underline{X}_j = (X_{1j}, X_{2j}, \cdots, X_{nj})$:

$$\hat{m}_j = a + b\overline{X}_{.j} \tag{24}$$

com $\overline{X}_{.j} = \frac{1}{n}\sum_{i=1}^{n} X_{ij}$, com as constantes a e b calculadas de forma a minimizar a função de perca quadrática esperada:

$$R = \mathrm{E}\left[\left(\mu(\theta_j) - \hat{m}_j\right)^2\right] = \mathrm{E}\left[\left(\mu(\theta_j) - a - b\overline{X}_{.j}\right)^2\right] \tag{25}$$

O estimador $\tilde{m}_j$ assim calculado chama-se Estimador de Credibilidade do verdadeiro prémio puro ou Fórmula de Credibilidade, tal como Bühlmann apelida.

Procedendo à dedução do estimador de credibilidade, a expressão (25) pode ser escrita da forma

$$\mathrm{E}\left[\left((\mu(\theta_j) - b\overline{X}_j]) - a\right)^2\right] = \mathrm{V}[\mu(\theta_j) - b\overline{X}_j] + \left(\mathrm{E}\left[\mu(\theta_j) - b\overline{X}_{.j}\right] - a\right)^2 \tag{26}$$

atendendo à relação $\mathrm{E}[(X-c)^2]=\mathrm{V}[X]+(\mathrm{E}[X]-c)^2$, válida para qualquer v.a. X e uma constante c.

Na relação (26) a segunda parcela é minimizada para qualquer b fazendo,

$$a=\mathrm{E}[\mu(\theta_j)-b\overline{X}_{.j}]=\mathrm{E}[\mu(\theta_j)]-b\,\mathrm{E}[\overline{X}_{.j}].$$

Atendendo à hipótese de equidistribuição condicionada de B1 e à propriedade iterativa do valor esperado,

$$\mathrm{E}[\overline{X}_{.j}]=\mathrm{E}[\mathrm{E}[\overline{X}_{.j}|\theta_j]]=\mathrm{E}[\mu(\theta_j)] \tag{27}$$

obtém-se

$$a=(1-b)\,\mathrm{E}[\mu(\theta_j)]. \tag{28}$$

Por outro lado, a primeira parcela do segundo membro de (26) pode ser escrita na forma

$$\mathrm{V}[\mu(\theta_j)-b\,\overline{X}_{.j}]=\mathrm{E}[\mathrm{V}[\mu(\theta_j)-b\,\overline{X}_{.j}|\theta_j]]+\mathrm{V}[\mathrm{E}[\mu(\theta_j)-b\,\overline{X}_{.j}|\theta_j]],$$

ou

$$\mathrm{V}[\mu(\theta_j)-b\,\overline{X}_{.j}]=b^2\mathrm{E}[\mathrm{V}[\overline{X}_{.j}|\theta_j]]+(1-b)^2\mathrm{V}[\mu(\theta_j)],$$

uma vez que dado θ_j, $\mu(\theta_j)$ é uma constante e $\mathrm{E}[\overline{X}_{.j}]=\mathrm{E}[\mu(\theta_j)]$. Além disso, pela hipótese de independência condicionada de B1, tem-se

$$\mathrm{V}[\overline{X}_{.j}|\theta_j]=\frac{1}{n}\mathrm{V}[X_{ij}|\theta_j]=\frac{1}{n}\sigma(\theta_j)^2, \tag{29}$$

pois, dado θ_j, a distribuição condicionada de X_{ij} não depende de i.

Tem-se então,

$$\mathrm{V}[\mu(\theta_j)-b\,\overline{X}_{.j}]=\frac{b^2}{n}\mathrm{E}[\sigma(\theta_j)^2]+(1-b)^2\mathrm{V}[\mu(\theta_j)]. \tag{30}$$

Derivando a expressão acima em ordem a b e igualando a zero,

$$\frac{2\,b}{n}\mathrm{E}[\sigma(\theta_j)^2]-2(1-b)\mathrm{V}[\mu(\theta_j)]=0,$$

donde

$$b^*=\frac{\mathrm{V}[\mu(\theta_j)]}{\mathrm{V}[\mu(\theta_j)]+\frac{1}{n}\mathrm{E}[\sigma(\theta_j)^2]}. \tag{31}$$

Por outro lado a segunda derivada de (30) em ordem a b é positiva, pois $(1/n)\mathrm{E}[\sigma(\theta_j)^2]+\mathrm{V}[\mu(\theta_j)]>0$, pelo que (31) é um minimizante. Substituindo (31) em (28) tem-se

$$a^*=\frac{\frac{1}{n}\mathrm{E}[\sigma(\theta_j)^2]}{\mathrm{V}[\mu(\theta_j)]+\frac{1}{n}\mathrm{E}[\sigma(\theta_j)^2]}\,\mathrm{E}[\mu(\theta_j)]. \tag{32}$$

Aplicando a notação

$$\mu=\mathrm{E}[\mu(\theta_j)],\ \phi=\mathrm{E}[\sigma(\theta_j)^2]\ \text{e}\ \psi=\mathrm{V}[\mu(\theta_j)] \tag{33}$$

(31) e (32) vêm

$$\begin{aligned} b^* &= \frac{\psi}{\psi+\frac{1}{n}\phi}=\frac{n}{n+\phi/\psi} \\ a^* &= \frac{\phi/\psi}{n+\phi/\psi}\mu \end{aligned}$$

e substituindo em (24), obtém-se o estimador de credibilidade

$$\boxed{\tilde{m}_j \ = \ z\overline{X}_{.j}+(1-z)\mu,\ \text{com } z=\tfrac{n}{n+k}\ \text{ e }\ k=\phi/\psi} \tag{34}$$

Note-se que os momentos (33) não dependem de j em virtude da hipótese de equidistribuição dos parâmetros de risco B2. Bühlmann designou estes parâmetros por Parâmetros Estruturais. De referir ainda que se supõe a existência dos momentos envolvidos. Tecendo algumas observações sobre o estimador (34), vê-se que tem a fórmula de credibilidade proposta por Whitney.

O trabalho de Bühlmann é um trabalho pioneiro, tendo esclarecido o emprego da Fórmula de Credibilidade (1), concretizando, no Factor de Credibilidade (2), o valor constante k.

A interpretação do factor de credibilidade z é bastante simples. É um número compreendido entre 0 e 1. No caso de ausência de informação, i.e. $n = 0$, então $z = 0$, o que permite pressupor, que o Prémio Puro a tarifar será o prémio coletivo μ. Por outro lado, quando n aumenta então z tende para 1, pois k é uma constante, e é atribuído à experiência um peso preponderante sobre o risco em questão. Além disso o factor de credibilidade z é decrescente com k. Um grande valor k significa que a média de variações dentro do mesmo risco é muito superior às variações das médias dos vários riscos. Nesse caso atribui-se um peso menor à experiência individual em virtude da sua irregularidade comparativa.

O prémio coletivo μ é a média das indemnizações anuais dum risco tirado ao acaso do coletivo.

O Prémio de Credibilidade é em média igual ao prémio coletivo pois $\mathrm{E}[\tilde{m}_j] = \mu$. Aproxima-se do verdadeiro Prémio Puro $\mu(\theta_j)$, à medida que n tende para o infinito.

Na expressão (24), dá-se o mesmo peso às variáveis observáveis X_{ij} $(i = 1, 2, \cdots, n)$ o que salta imediatamente à vista pelo facto do modelo ser estacionário. Isto é, não há evolução na característica particular θ_j com o tempo, mantendo um *status quo* ao longo da vida do risco, hipótese simplista e um pouco irreal.

Note-se ainda que sendo o historial em quantidade igual para cada risco da carteira, i.e., registo de indemnizações igual em número de períodos de tarifação, hipótese inicialmente considerada na Secção 4, o Factor de Credibilidade z vem igual para todos os riscos j $(j = 1, 2, \cdots, N)$. No entanto a fórmula de credibilidade é igualmente utilizável no caso desta hipótese não se verificar, apenas se substituindo n por um n_j $(j = 1, 2, \cdots, N)$, e, claro, $\overline{X}_{\cdot j} = (1/n_j)\sum_{i=1}^{n_j} X_{ij}$.

Nesse caso a carteira pode admitir novos riscos, $N+1$, $N+2$, etc., aos quais será atribuído o prémio coletivo no primeiro ano de tarifação. A Forma de Credibilidade toma assim a forma

$$\boxed{\tilde{m}_j \quad = \quad z_j\overline{X}_{\cdot j} + (1 - z_j)\mu, \text{ com } z_j = \tfrac{n_j}{n_j+k} \text{ e } k = \phi/\psi} \tag{35}$$

Na Secção 6.3.2 a propósito do modelo de Hachemeister, far-se-á referência a esta expressão. Note-se ainda que nesse caso o factor de credibilidade é em geral diferente de risco para risco, pois depende de j através de n_j.

Tanto (34) como (35) dependem dos parâmetros ϕ, ψ e μ que se assumiram conhecidos. No entanto na prática não o são em geral, tornando-se necessária a sua estimação, o que se vai fazer na secção seguinte.

6.1.2 Estimação dos parâmetros estruturais e propriedades

Para resolver o problema da estimação dos parâmetros estruturais note-se em primeiro lugar que os dados disponíveis dizem respeito ao colectivo. Para tornar claro a sua estimação é necessário que se mostre que os parâmetros estruturais podem ser escritos em função de momentos do coletivo.

O parâmetro μ facilmente se apresenta em função de $\mathrm{E}[X_{ij}]$, pois

$$\mu = \mathrm{E}[\mu(\theta_j)] = \mathrm{E}[\mathrm{E}[X_{ij}|\theta_j]] = \mathrm{E}[X_{ij}].$$

Além disso sabe-se que $\mathrm{V}[X_{ij}] = \mathrm{V}[\mu(\theta_j)] + \mathrm{E}[\sigma(\theta_j)^2]$ e utilizando (32) e (33), vem

$$\mathrm{V}[X_{ij}] = \psi + \phi. \tag{36}$$

Por outro lado, o denominador de (31),

$$\psi + \frac{1}{n}\phi = \mathrm{V}[\overline{X}_{\cdot j}], \tag{37}$$

pois, $\mathrm{V}[\overline{X}_{\cdot j}] = \mathrm{E}[\mathrm{V}[\overline{X}_{\cdot j}|\theta_j] + \mathrm{V}[\mathrm{E}[\overline{X}_{\cdot j}|\theta_j]]$, e tendo em conta (27), (29) e (33), obtém-se (37).

Se se subtrair membro a membro (36) de (37) depois de se multiplicar nesta última ambos os membros por n, obtém-se $n\mathrm{V}[\overline{X}_{\cdot j}] - \mathrm{V}[X_{ij}] = (n-1)\psi$, obtendo-se imediatamente ψ em função dos momentos do coletivo

$$\psi = \frac{n\,\mathrm{V}[\overline{X}_{\cdot j}] - \mathrm{V}[X_{ij}]}{n-1}.$$

Substituindo a expressão atrás em (36) e resolvendo em ϕ tem-se

$$\phi = \frac{n\left(\mathrm{V}[X_{ij}] - \overline{X}_{\cdot j}]\right)}{n-1},$$

que é função de momentos do coletivo.

O estimador natural para μ, que expressa a média da população de riscos tomada sobre as médias de cada risco nos n períodos, é o estimador

$$\hat{\mu} = \overline{\overline{X}}_{\cdot\cdot} = \frac{1}{nN}\sum_{j=1}^{N}\sum_{i=1}^{n} X_{ij}. \tag{38}$$

O parâmetro ϕ expressa a média das variações de cada risco de período para período. Portanto, o estimador natural é a correspondente média das variâncias (corrigidas) da amostra

$$\hat{\phi} = \frac{1}{N}\sum_{j=1}^{N} {S'_j}^2 = \frac{1}{N}\sum_{j=1}^{N}\sum_{i=1}^{n} \frac{\left(X_{ij} - \overline{X}_{\cdot j}\right)^2}{n-1}. \tag{39}$$

Toma-se a variância corrigida de forma a ter-se um estimador centrado.

O parâmetro $\psi = \mathrm{V}[\overline{X}_{\cdot j}] - (1/n)\phi$ expressa a variação no verdadeiro Prémio Puro entre os riscos da população. O estimador natural é a variância empírica (corrigida) da média da amostra. No entanto, a média $\overline{X}_{\cdot j}$ tem uma variação superior à variação de $\mu(\theta_j)$ pois é afectada por variações de período para período, dentro do mesmo risco, o que torna este estimador enviezado. Assim, propõe-se o estimador

$$\hat{\psi} = \frac{1}{N-1}\sum_{j=1}^{N}\left(\overline{X}_{\cdot j} - \overline{\overline{X}}_{\cdot\cdot}\right)^2 - \frac{1}{n}\hat{\phi}. \tag{40}$$

Os estimadores (39) e (40) são os estimadores propostos por Bühlmann & Straub (1970), e particularizados a este modelo. Em relação a (40) verifica-se que este pode ser negativo, se

$$\frac{1}{n}\hat{\phi} > \frac{1}{N-1}\sum\left(\overline{X}_{\cdot j} - \overline{\overline{X}}_{\cdot\cdot}\right)^2, \tag{41}$$

o que não faz sentido em virtude de $\psi \geq 0$. Para tornear este problema Bühlmann & Straub (1970) propõem substituir $\hat{\psi}$ pelo estimador

$$\hat{\psi}^* = \max\left(0, \hat{\psi}\right). \tag{42}$$

estes autores afirmam que a escolha de $\hat{\psi}^*$ em vez de $\hat{\psi}$, pode ser intuitivamente explicada pelo facto de quando (41) se verifica, então as estatísticas suportam a hipótese que os riscos pertencem a um coletivo homogéneo (nesse caso $\psi = 0$ pois o prémio de risco não depende de θ).

De referir ainda que os estimadores apresentados podem ser *melhorados* através de algum conhecimento da distribuição condicionada de X_{ij}, que podem trazer algum conhecimento adicional sobre os seus parâmetros. Por exemplo, no caso da distribuição de Poisson tem-se $\mu(\theta_j) = \sigma(\theta_j)^2$, então $\hat{\phi} = \hat{\mu} = \overline{\overline{X}}_{\cdot\cdot}$.

Quanto a propriedades, os estimadores (38)-(40) são centrados e consistentes.

Verifique-se o não-enviesamento. Para o estimador $\hat{\mu}$, a sua verificação é imediata

$$\mathrm{E}[\hat{\mu}] = \mathrm{E}[\overline{\overline{X}}_{\cdot\cdot}] = \frac{1}{N}\sum_{j=1}^{N}\mathrm{E}[\overline{X}_{\cdot j}],$$

e atendendo a (27) e à propriedade iterativa do valor esperado tem-se

$$\frac{1}{N}\sum_{j=1}^{N}\mathrm{E}[\overline{X}_{\cdot j}] = \frac{1}{N}\sum_{j=1}^{N}\mathrm{E}[\mu(\theta_j)] = \mu.$$

Relativamente ao estimador $\hat{\phi}$, operando o valor esperado, tem-se

$$\mathrm{E}[\hat{\phi}] = \frac{1}{N}\sum_{j=1}^{N}\mathrm{E}[{S'_j}^2 = \frac{1}{N}\sum_{j=1}^{N}\mathrm{E}[\mathrm{E}[{S'_j}^2|\theta_j].$$

Atendendo a que $\mathrm{E}[{S'_j}^2|\theta_j] = \sigma(\theta_j)^2$, (veja-se por exemplo Murteira (1990*a*)), tem-se

$$\mathrm{E}[\hat{\phi}] = \frac{1}{N}\sum_{j=1}^{N}\mathrm{E}[\sigma(\theta_j)^2] = \frac{1}{N}\sum_{i=1}^{N}\phi = \phi\ ,$$

por (33).

Em relação ao estimador $\hat{\psi}$, opere-se o valor esperado à estatística $\sum_{j=1}^{N}\left(\overline{X}_{\cdot j} - \overline{\overline{X}}_{\cdot\cdot}\right)^2/(N-1)$, que representa a variância (corrigida) da média empírica dentro do mesmo risco j, e cujo valor esperado é igual à variância da média empírica $\overline{X}_{\cdot j}$. Este valor esperado é igual a $\mathrm{V}[\overline{X}_{\cdot j}]$, e por (37)

$$\mathrm{V}[\overline{X}_{\cdot j}] = \psi + \frac{1}{n}\phi\ ,$$

donde se obtém que este valor é centrado.

Os estimadores apresentados são também consistentes, i.e.

$$(\hat{\mu}, \hat{\phi}, \hat{\psi}) \xrightarrow{P} (\mu, \phi, \psi) \text{ quando } N \to \infty.$$

O símbolo $\xrightarrow{P}$ significa convergência em probabilidade, consequência da Lei Fraca dos Grandes Números, que a seguir se apresenta:

Definição 5 *Dada uma sucessão de variáveis aleatórias* $\{Y_N\}$*, forme-se a sucessão* $\{Z_N\}$

$$Z_N = \frac{1}{N}\sum_{j=1}^{N}Y_j - \frac{1}{N}\sum_{j=1}^{N}\mathrm{E}[Y_j]$$

onde $\mathrm{E}[Y_N]$ *é o valor esperado de* Y_N*, cuja existência se admite. Diz-se que* $\{Y_N\}$ *obedece à Lei Fraca dos Grandes Números quando* $Z \xrightarrow{P} 0$.

Para provar a consistência do estimador $\hat{\mu}$, note-se que $\{\overline{X}_{\cdot N}\}$ é uma sucessão de v.a.'s aleatórias independentes com valor esperado μ e variância $\psi + (1/n)\phi$, para $N = 1, 2, \cdots$, donde pelo Corolário 5.5 de Murteira (1990*b*), a sucessão $\{\overline{X}_{\cdot N}\}$ obedece à lei fraca dos grandes números. Então,

$$\hat{\mu} \xrightarrow{P} \mu$$

porque $\mu = (1/N)\sum_{j=1}^{N}\mathrm{E}[\overline{X}_{\cdot j}]$.

Pelas mesmas razões a sucessão $\{{S'_N}^2\}$ obedece à Lei Fraca dos Grandes Números, já que as v.a.'s ${S'_j}^2$, $(j = 1, 2, \cdots, N)$ são i.i.d.(resulta da hipótese B2) com média $\mathrm{E}[{S'_j}^2] = \phi$ e variância

$$\begin{aligned}
\mathrm{V}[{S'_j}^2] &= \mathrm{V}[\mathrm{E}[{S'_j}^2|\theta_j]] + \mathrm{E}[\mathrm{V}[{S'_j}^2|\theta_j]] \\
&= \mathrm{V}[\sigma(\theta_j)^2] + \mathrm{E}\left[(1/n)\left(\mu_4(\theta_j) - \frac{n-3}{n-1}\sigma(\theta_j)^4\right)\right]
\end{aligned}$$

com $n > 1$, e $\mu_4(\theta_j)$ é o quarto momento centrado de X_{ij}, dado θ_j, cuja existência se admite. Donde

$$\hat{\phi} \xrightarrow{P} \phi.$$

Para o estimador $\hat{\psi}$, tem-se que a primeira parcela é a variância da amostra das médias $\overline{X}_{\cdot j}$, que são i.i.d. Como tal, é estimador consistente da variância do universo, $\mathrm{V}[\overline{X}_{\cdot j}]$, ou seja,

$$\hat{\psi} + \frac{1}{n}\hat{\phi} \xrightarrow{P} \psi + \frac{1}{n}\phi$$

donde se conclui que

$$\hat{\psi} \xrightarrow{P} \psi$$

pois como se mostrou atrás $\hat{\phi}$ é estimador consistente de ϕ.

Ainda em relação aos estimadores (38)-(40) e (42). Estes podem ser generalizados para n_j $(j = 1, 2, \cdots, N)$ períodos genéricos (Sundt (1982)). Neste caso os estimadores tomam a seguinte forma:

$$\hat{\mu} = \overline{\overline{X}}_{\cdot\cdot} = \frac{1}{\sum_{j=1}^{N} n_j} \sum_{j=1}^{N} \sum_{i=1}^{n_j} X_{ij} \tag{43}$$

$$\hat{\phi} = \frac{1}{\sum_{j=1}^{N} n_j - N} \sum_{j=1}^{N} \sum_{i=1}^{n_j} \left(X_{ij} - \overline{X}_{\cdot j}\right)^2 \tag{44}$$

$$\hat{\psi} = \frac{\sum_{j=1}^{N} \left(\overline{X}_{\cdot j} - \overline{\overline{X}}_{\cdot\cdot}\right)^2 - (N-1)\hat{\phi}}{\sum_{j=1}^{N} n_j \left(1 - \frac{n_j}{\sum_{k=1}^{N} n_k}\right)} \tag{45}$$

sendo $\overline{X}_{\cdot j} = (1/n_j) \sum_{i=1}^{n_j} X_{ij}$. Para evitar que seja negativo, em vez de (45) toma-se

$$\hat{\psi}^* = \max\left(0, \hat{\psi}\right).$$

Os estimadores (43)-(45) são centrados e consistentes. Far-se-á referência a propósito do modelo de Hachemeister (1975) na Secção 6.3.4.

6.1.3 Prémio Empírico de Credibilidade

Se no Estimador de Credibilidade (34) se substituir os parâmetros (33) por suas estimativas obtém-se

$$\tilde{m}_{j,N} = z_N \overline{X}_{\cdot j} + (1 - z_N)\hat{\mu} \tag{46}$$

com $z_N = n/(n + \hat{\phi}/\hat{\psi})$, que Norberg (1979) chama **Prémio Empírico de Credibilidade**, realçando as seguintes propriedades:

- Fácil de calcular
- Converge em probabilidade para o Estimador de Credibilidade (34) quando $N \to \infty$
- O erro quadrático médio
$$\mathrm{E}\left[(\mu(\theta_j) - \tilde{m}_{j,N})^2\right] \to \mathrm{E}\left[(\mu(\theta_j) - \tilde{m}_j)^2\right]$$
quando $N \to \infty$
- $\tilde{m}_{j,N}$ é constituído inteiramente através das observações e a sua equivalência assintótica para $\tilde{m}_j$ é válida para qualquer que seja o espaço parâmetro Θ e da forma da família das distribuições condicionadas $\mathcal{F} = \{F(\cdot|\theta); \theta \epsilon \Theta\}$.

O modelo de Bühlmann permite a construção dum Sistema *Bonus-Malus*, baseado no número de indemnizações, aplicável ao seguro automóvel. Suponha-se o prémio puro equacionado em termos da frequência de indemnizações (equivale a estabelecer a unidade monetária como custo médio de uma indemnização). No exemplo que se apresenta, retirado de Lemaire (1995), o conceito de risco tem o significado de uma apólice do ramo automóvel duma determinada companhia.

Exemplo 12 *Aplicação a um Sistema Bonus-Malus no seguro automóvel.*

Suponha-se que para um determinado risco j $(j = 1, 2, \cdots, N)$, X_{ij} representa o número de indemnizações ocorridas no período i. X_{ij} tem distribuição de Poisson com parâmetro θ_j e que dado θ_j as variáveis X_{ij} $(i = 1, 2, \cdots, n)$ são independentes. $\theta_1, \theta_2, \cdots, \theta_N$ são i.i.d. Então o estimador de credibilidade de $\mu(\theta_j) = \sigma(\theta_j)^2 = \theta_j$, baseado em $\underline{X}_j = (X_{1j}, \cdots, X_{nj})$ para o período $(n+1)$ é dado pela seguinte expressão

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

No quadro seguinte são apresentados dados sobre a frequência de indemnizações, referentes a um período de um ano, retiradas de Lemaire (1995). Estes dados referem-se a uma carteira constituída por 106974 *apólices de determinada companhia belga. Suponha-se ainda que o período em observação corresponde a um período de estabilidade em relação ao comportamento dos segurados. Assim vai-se considerar que o período em observação pode ser tomado como representativo de períodos seguintes e vão-se estimar os momentos do coletivo com base nestes dados. No quadro seguinte x é o número de indemnizações e n_x é o número de riscos com x indemnizações.*

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

Como estimador para $\mathrm{E}[X] = \mathrm{E}[\theta_j]$ considera-se $\overline{X} = (1/106974)\sum_{k=0}^{4} x_k n_{x_k} = 0.1011$. Como estimador para $\mathrm{V}[X] = \mathrm{E}[\theta_j] + \mathrm{E}[\theta_j]$ considera-se $s^2 = (1/106974)\sum_{k=0}^{4} x_k n_{x_k} - \overline{X}^2 = 0.1074$. Então o estimador para $\mathrm{V}[\theta_j]$ vem $\hat{\mathrm{V}}[\theta_j] = 0,1074 - 0,1011 = 0,0063$. O estimador empírico de credibilidade vem

$$\begin{aligned}\tilde{\theta}_{j,N} &= \frac{n}{n + 0,1011/0,0063}\overline{X}_{\cdot j} + \frac{0,1011/0,0063}{n + 0,1011/0,0063} \times 0,1011 \\ &= \left(\sum_{i=1}^{n} x_{ij} + 16,047 \times 0,1011\right) / \left(n + 16,0476\right)\end{aligned}$$

Para uma melhor interpretação do efeito da ocorrência ou não de sinistros para uma determinada apólice j, vai-se tomar o quociente Prémio Puro de Risco/Prémio Puro Coletivo. Para tomar a unidade 100, como a unidade a atribuir no primeiro período de tarifação, multiplica-se este quociente por 100. Note-se que o prémio a atribuir no primeiro período de tarifação para determinado risco, corresponde ao prémio coletivo. Assim, tomando o estimador $\tilde{\theta}_{j,N}$, o prémio relativo para o período $(n+1)$ vem

$$P^*_{n+1}(\underline{X}_j) = 100 \times \frac{\sum_{i=1}^{n} X_{ij} + 1,6224}{0,1011(n + 16,0476)} = 100 \times \frac{\sum_{i=1}^{n} X_{ij} + 1,6224}{0,1011\, n + 1,6224}$$

A Tabela 2 seguinte mostra os valores de $P^*_{n+1}(\underline{X}_j)$ correspondente a determinada apólice j que causa um total de $\sum_{i=1}^{n} X_{ij}$ acidentes em n períodos.

Como se pode ver na tabela, o prémio é reduzido como compensação de ausência de sinistros e aumentado no caso da sua ocorrência. Note-se a forma como o risco é fortemente penalizado pela ocorrência de sinistros. $\diamondsuit$

Apenas algumas notas adicionais para finalizar a secção. Nas secções anteriores, em que foram tratados os modelos puramente Bayesianos e o modelo de Bühlmann, foi realçado a vantagem prática deste último em virtude das dificuldades de implementação na prática da análise puramente Bayesiana. Todo o raciocínio é elaborado pressupondo que o risco a tarifar se integra num determinado grupo ou coletivo de riscos do mesmo tipo, situação comum nalguns ramos de seguro.

	número de acidentes				
nº periodos	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

Tabela 2: Prémio relativo para um sistema *Bonus-malus*

No entanto existem situações em que não é possível integrar um determinado risco em qualquer coletivo. Isto acontece por exemplo quando uma companhia introduz um novo tipo de cobertura, para a qual não existem estatísticas fiáveis disponíveis (por exemplo um seguro de responsabilidades sobre acidentes nucleares, ou uma cobertura tradicional sob circunstâncias nunca antes encontradas (por exemplo considerando uma situação pouco comum, um seguro de acidentes para Limpa-Chaminés ou um seguro para cobrir os prémios a atribuir *aos jogadores do Benfica pelo título de campeão nacional de futebol!*).

Em tais situações o actuário não tem à sua disposição estatísticas sobre o coletivo, mas a companhia tendo em conta a sua vocação comercial não vai à partida abandonar o negócio por esta dificuldade, nem vai dizer ao cliente que faz o negócio no futuro assim que o *acaso tenha feito das suas*. Em princípio, recorrendo à sua intuição e à sua experiência adquirida noutro tipo de riscos, vai tentar atribuir um prémio possível, inicial, ao risco em questão. Este prémio é, no início, um palpite, na convicção de que será posteriormente ajustado à medida que a experiência se desenrola. Ora, este tipo de raciocínio ajusta-se perfeitamente à análise Bayesiana. Concretizando, o actuário representa o seu conhecimento subjectivo numa distribuição *a priori* do parâmetro θ. A distribuição *a posteriori* de θ, dadas as observações $\underline{X} = (X_1, \cdots, X_n)$, expressa o conhecimento subjectivo do investigador após o julgamento motivado pelas observações.

De notar ainda que o modelo de Bühlmann é um modelo de inspiração Bayesiana, não exigindo no entanto o conhecimento da Distribuição Estrutural. Apenas é exigido o conhecimento dos parâmetros estruturais. Quando se desconhecem estes parâmetros estruturais, o que obriga à sua estimação através de estimadores baseados em estatísticas sobre o coletivo, o modelo é designado por Empírico-Bayesiano.

6.2 Modelo de Bühlmann-Straub

6.2.1 Definições e hipóteses

O modelo de Bühlmann-Straub segue na sequência do modelo de Bühlmann, pondo pela primeira vez de lado a hipótese de homogeneidade temporal proposta na hipótese B1 daquele modelo. Por outro lado, abriu caminho a um tratamento mais geral dos modelos Empírico-Bayesianos, como se verá mais adiante.

O modelo foi originalmente apresentado na óptica seguradora-resseguradora, no entanto é aplicável ao seguro directo, já que este caso está contido no modelo mais geral considerado. Assim, o conceito de risco é tomado como sinónimo de contrato de resseguro para o período a tarifar. A apresentação e desenvolvimento do problema são feitos no seguimento do raciocínio do modelo de Bühlmann do Capítulo 4.

Considerem-se um conjunto de N riscos (contractos de resseguro) numerados de 1 a N (número de contractos da carteira do ressegurador. Sejam as variáveis:

S_{ij} - O montante global das indemnizações cedidas, do contracto j ($j = 1, 2, \cdots, N$) no ano i ($i = 1, 2, \cdots, n$).

P_{ij} - Volume de prémios do seguro directo do contracto j no ano i, ou volume de exposição de risco.

$X_{ij} = \frac{S_{ij}}{P_{ij}}$ - *loss ratio* ou valor ressegurado por unidade de risco.

θ_j - Representa o parâmetro desconhecido associado ao risco j, independente de i.

As hipóteses do modelo são as seguintes:

BS1 - Dado θ_j, as variáveis $X_{1j}, X_{2j}, \cdots, X_{nj}, X_{n+1,j}$, são independentes e

$$\mathrm{E}[X_{ij}|\theta_j] = \mu(\theta_j) \tag{47}$$

$$\mathrm{V}[X_{ij}|\theta_j] = \frac{\sigma(\theta_j)^2}{P_{ij}} \tag{48}$$

$(i = 1, 2, \cdots, n)$, $(j = 1, 2, \cdots, N)$.

BS2 := B2.

Tecendo alguns comentários em relação às hipóteses aqui propostas e relacionando-as com as hipóteses do modelo de Bühlmann, verifica-se que a hipótese de homogeneidade temporal é retirada. A expressão (48) pode ser justificada pelo seguinte raciocínio:

Caso P_{ij} fosse o número de apólices do risco j no ano i e definindo Y_{ijk} como as indemnizações relativas à apólice k do seguro directo, do risco j no ano i, então

$$X_{ij} = \frac{S_{ij}}{P_{ij}} = \sum_{k=1}^{P_{ij}} Y_{ijk},$$

e a variância condicionada de X_{ij} dado θ_j vem

$$\mathrm{V}[X_{ij}|\theta_j] = \frac{1}{P_{ij}^2} \mathrm{V}\left[\sum_{k=1}^{P_{ij}} Y_{ijk}|\theta_j\right] = \frac{1}{P_{ij}^2} \sum_{k=1}^{P_{ij}} \mathrm{V}\left[Y_{ijk}|\theta_j\right]$$

supondo a independência condicionada das P_{ij} apólices. Se além disso se admitir $\mathrm{V}[Y_{ijk}|\theta_j] = \sigma(\theta_j)^2$, $(k = 1, 2, \cdots, P_{ij})$, então tem-se $\mathrm{V}[X_{ij}|\theta_j] = \sigma(\theta_j)^2/P_{ij}$.

6.2.2 Estimador de credibilidade

O Prémio Puro do Risco j para o período $n+1$ é o valor esperado $\mu(\theta_j)$, dado por (47), a estimar com base em em X_{ij} ($i = 1, 2, \cdots, n$). O modelo vai ser desenvolvido a partir de uma óptica que difere um pouco do método proposto originalmente por Bühlmann & Straub (1970).

Propõe-se um estimador linear da forma

$$\hat{X}_{n+1,j} = a_0 + \sum_{i=1}^{n} a_i X_{ij} \tag{49}$$

Uma vez que por BS1 os riscos são independentes pode-se desprezar a informação colateral. Além disso, como não se supõe a hipótese de equidistribuição (condicional) das variáveis X_{ij} ($i = 1, 2, \cdots, n$), não há razões para se propor um estimador da forma (24). Note-se que por (16) é equivalente estimar $X_{n+1,j}$ ou $\mu(\theta_j)$.

De forma a simplificar o desenvolvimento que se vai fazer a seguir, considere-se a seguinte notação:

$$P_{\cdot j} = \sum_{i=1}^{n} P_{ij} \tag{50}$$

$$P_{\cdot\cdot} = \sum_{j=1}^{N} P_{\cdot j} \tag{51}$$

$$\overline{X}_{\cdot j} = \frac{1}{P_{\cdot j}} \sum_{i=1}^{n} P_{ij} X_{ij} \tag{52}$$

$$\overline{\overline{X}}_{\cdot\cdot} = \frac{1}{P_{\cdot\cdot}} \sum_{j=1}^{N} P_{\cdot j} \overline{X}_{\cdot j} \tag{53}$$

Tenha-se igualmente presente a notação (33). O estimador de Credibilidade para este modelo de Bühlmann-Straub, tendo em atenção as hipóteses propostas, é apresentado no seguinte teorema.

Teorema 4 *O estimador de credibilidade de $X_{n+1.j}$ ($j = 1, 2, \cdots, N$) com base em $\underline{X}_j = (X_{1j}, \cdots, X_{nj})$ vem*

$$\boxed{\tilde{X}_{n+1,j} = z\overline{X}_{\cdot j} + (1-z)\mu, \quad \text{com } z_j = \frac{P_{\cdot j}}{P_{\cdot j}+k} \quad \text{e} \quad k = \phi/\psi} \tag{54}$$

em que $P_{\cdot j}$, $\overline{X}_{\cdot j}$, μ, ϕ e ψ são definidos respectivamente por (50), (52) e (33).

Dem.

Seja $\hat{X}_{n+1,j}$ o estimador de $X_{n+1,j}$ da forma (49). É um estimador de credibilidade se e só se satisfaz as condições de normalidade (9) e (10) do Teorema 1, i.e.

$$\mathrm{E}\left[\hat{X}_{n+1,j}\right] = X_{n+1,j} \tag{55}$$

$$\mathrm{Cov}\left[\hat{X}_{n+1,j}, X_{ij}\right] = \mathrm{Cov}\left[X_{n+1,j}, X_{ij}\right] \quad i = 1, 2, \cdots, n. \tag{56}$$

De (56) tem-se

$$\mathrm{Cov}\left[X_{n+1,j}, X_{ij}\right] = \mathrm{Cov}\left[a_0 + \sum_{k=1}^{n} a_k X_{kj}, X_{ij}\right] = \sum_{k=1}^{n} a_k \mathrm{Cov}\left[X_{kj}, X_{ij}\right]$$

equivalente a

$$\mathrm{Cov}\left[\mathrm{E}\left[X_{n+1,j}|\theta_j\right], \mathrm{E}\left[X_{ij}|\theta_j\right]\right] = \sum_{k=1}^{n} a_k \left(\mathrm{Cov}\left[\mathrm{E}\left[X_{kj}|\theta_j\right], \mathrm{E}\left[X_{ij}|\theta_j\right]\right] + \mathrm{E}\left[\mathrm{Cov}\left[X_{kj}, X_{ij}|\theta_j\right]\right]\right)$$

pela hipótese de independência condicional, dado θ_j, donde

$$\mathrm{V}\left[\mu(\theta_j)\right] = \sum_{k=1}^{n} a_k \left(\mathrm{V}\left[\mu(\theta_j)\right] + \delta_{ik} \frac{\mathrm{E}\left[\sigma(\theta_j)^2\right]}{P_{ij}}\right),$$

sendo δ_{ik} o índice de Kronecker, i.e.

$$\delta_{ik} = \begin{cases} 1 & i = k \\ 0 & i \neq k \end{cases}, \tag{57}$$

e porque pela hipótese de independência condicionada de $\mathsf{BS1}$ tem-se

$$\mathrm{Cov}\left[X_{kj}, X_{ij}|\theta_j\right] = \begin{cases} \mathrm{V}\left[X_{ij}|\theta_j\right] & i = k \\ 0 & i \neq k \end{cases}.$$

Tem-se, introduzindo a notação (33),

$$\psi = \psi \sum_{k=1}^{n} a_k + \frac{a_i}{P_{ij}} \phi,$$

equivalente a

$$a_i\phi = \psi P_{ij}\left(1-\sum_{k=1}^{n} a_k\right). \tag{58}$$

Somando as equações em i obtém-se

$$\sum_{i=1}^{n} a_i = \frac{\psi P_{\cdot j}}{\phi + \psi P_{\cdot j}}. \tag{59}$$

Substituindo (59) em (58) e resolvendo a equação em ordem a a_i obtém-se o valor óptimo

$$a_i^* = \frac{\psi}{\phi + \psi P_{\cdot j}} P_{ij}.$$

Por outro lado da condição (55) obtém-se

$$a_0 = \mathrm{E}\left[X_{n+1,j}\right] - \sum_{i=1}^{n} a_i \mathrm{E}\left[X_{ij}\right] = \left(1-\sum_{i=1}^{n} a_i\right)\mu,$$

por BS1. Introduzindo (59) obtém-se o valor óptimo,

$$a_0{}^* = \frac{\phi}{\phi + \psi P_{\cdot j}}\mu. \tag{60}$$

Introduzindo a_0^* e a_i^* em (49) obtém-se (54), como se pretendia demonstrar. □

O estimador dado pela expressão (54) tem ainda a forma de credibilidade (1) proposta por Whitney, no entanto o factor de credibilidade z_j pode variar consoante o risco em virtude da constante $P_{\cdot j}$. Refira-se um caso particular. Se $P_{ij} = 1\ \forall i$, então $P_{\cdot j} = n$ e o modelo reduz-se ao modelo de Bühlmann.

O estimador (54) foi deduzido considerando a hipótese $n_1 = n_2 = \cdots = n_N = n$. No entanto facilmente se generaliza para diferentes n_j, tomando o estimador a forma:

$$\boxed{\tilde{X}_{n_j+1,j} \;=\; z_j\overline{X}_{\cdot j} + (1-z_j)\,\mu, \quad \text{com } z_j = \tfrac{P_{\cdot j}}{P_{\cdot j}+k},\ k=\phi/\psi,\ P_{\cdot j} = \textstyle\sum_{i=1}^{n_j} P_{ij} \ \text{ e } \ \overline{X}_{\cdot j} = \tfrac{1}{P_{\cdot j}}\textstyle\sum_{i=1}^{n_j} P_{ij}X_{ij}}. \tag{61}$$

Voltar-se-á a abordar ainda este caso na Secção 6.3.2, a propósito de casos particulares do modelo de Hachemeister.

O problema da estimação dos parâmetros estruturais μ, ϕ e ψ será tratado na secção seguinte.

6.2.3 Estimação dos parâmetros estruturais

O estimador natural para μ, que representa a média coletiva teórica, é a média empírica ponderada

$$\hat{\mu} = \overline{\overline{X}}_{\cdot\cdot} = \frac{1}{P_{\cdot\cdot}}\sum_{j=1}^{N}\sum_{i=1}^{n} P_{ij}X_{ij}. \tag{62}$$

Para ϕ, Bühlmann & Straub (1970) propõem a correspondente média das variâncias (corrigidas) da amostra dentro do mesmo risco, mas multiplicada por $P_{\cdot\cdot}$ de forma a tornar o estimador centrado:

$$\hat{\phi} = P_{\cdot\cdot}\frac{1}{N}\sum_{j=1}^{N}\left(\frac{1}{n-1}\sum_{i=1}^{n}\frac{P_{ij}}{P_{\cdot\cdot}}\left(X_{ij}-\overline{X}_{\cdot j}\right)^2\right) = \frac{P_{\cdot\cdot}}{N(n-1)}\left(\sum_{j=1}^{N}\sum_{i=1}^{n}\left(\frac{P_{ij}}{P_{\cdot\cdot}}X_{ij}^2\right) - \sum_{j=1}^{N}\left(\frac{P_{ij}}{P_{\cdot\cdot}}\overline{X}_{\cdot j}^2\right)\right). \tag{63}$$

Para estimar o parâmetro ψ, que representa a variância das médias dentro do mesmo risco, Bühlmann & Straub (1970) propõem uma função do correspondente momento empírico (assim considerado por forma que o estimador seja centrado):

$$\begin{aligned}\hat{\psi} &= \left(\frac{1}{nN-1}\sum_{j=1}^{N}\sum_{i=1}^{n}\frac{P_{ij}}{P_{\cdot\cdot}}\left(X_{ij}-\overline{\overline{X}}_{\cdot\cdot}\right)^2-\frac{\hat{\phi}}{P_{\cdot\cdot}}\right)\frac{1}{\Pi} \\ &= \left(\frac{1}{nN-1}\left(\sum_{j=1}^{N}\sum_{i=1}^{n}\frac{P_{ij}}{P_{\cdot\cdot}}X_{ij}^2-\overline{\overline{X}}_{\cdot\cdot}^2\right)-\frac{\hat{\phi}}{P_{\cdot\cdot}}\right)\frac{1}{\Pi}\end{aligned} \tag{64}$$

com

$$\Pi=\frac{1}{nN-1}\left(\sum_{j=1}^{N}\frac{P_{\cdot j}}{P_{\cdot\cdot}}\left(1-\frac{P_{\cdot j}}{P_{\cdot\cdot}}\right)\right)$$

e $\hat{\phi}$ dado por (63). Por forma a evitar que o estimador para ψ possa ser negativo, propõe-se substituir o estimador (64) por

$$\hat{\psi}^*=\max\left(0,\hat{\psi}\right). \tag{65}$$

O estimador $\hat{\psi}$ pode ainda ser escrito numa forma mais simples:

$$\hat{\psi}=\frac{\sum_{j=1}^{N}P_{\cdot j}\left(\overline{X}_{\cdot j}-\overline{\overline{X}}_{\cdot\cdot}\right)^2-(n-1)\,\hat{\phi}}{\sum_{j=1}^{N}P_{\cdot j}\left(1-\frac{P_{\cdot j}}{P_{\cdot\cdot}}\right)}\ , \tag{66}$$

porque se se desenvolver a seguinte expressão contida em (64)

$$\begin{aligned}\sum_{j=1}^{N}\sum_{i=1}^{n}\tfrac{P_{ij}}{P_{\cdot\cdot}}\left(X_{ij}-\overline{\overline{X}}_{\cdot\cdot}\right)^2 &= \sum_{j=1}^{N}\sum_{i=1}^{n}\tfrac{P_{ij}}{P_{\cdot\cdot}}\left(X_{ij}-\overline{X}_{\cdot j}+\overline{X}_{\cdot j}-\overline{\overline{X}}_{\cdot\cdot}\right)^2 \\ &= \sum_{j=1}^{N}\sum_{i=1}^{n}\tfrac{P_{ij}}{P_{\cdot\cdot}}\left(X_{ij}-\overline{X}_{\cdot j}\right)^2+\sum_{j=1}^{N}\sum_{i=1}^{n}\tfrac{P_{ij}}{P_{\cdot\cdot}}\left(\overline{X}_{\cdot j}-\overline{\overline{X}}_{\cdot\cdot}\right)^2\end{aligned}\ , \tag{67}$$

porque $\sum_{j=1}^{N}\sum_{i=1}^{n}(P_{ij}/P_{\cdot\cdot})\,(\overline{X}_{\cdot j}-\overline{\overline{X}}_{\cdot\cdot})(X_{ij}-\overline{X}_{\cdot j})=0$. Então (67) vem ainda

$$\frac{1}{P_{\cdot\cdot}}\left(\sum_{j=1}^{N}P_{\cdot j}\left(\overline{X}_{\cdot j}-\overline{\overline{X}}_{\cdot\cdot}\right)^2+N\,(n-1)\,\hat{\phi}\right),$$

somando a segunda parcela de (67) em i e substituindo a primeira parcela por $N\,(n-1)\,\hat{\phi}$, a partir da expressão (63). Então $\hat{\psi}$ vem

$$\begin{aligned}\hat{\psi} &= \tfrac{1}{(nN-1)P_{\cdot\cdot}\Pi}\left(\sum_{j=1}^{N}P_{\cdot j}\left(\overline{X}_{\cdot j}-\overline{\overline{X}}_{\cdot\cdot}\right)^2+N\,(n-1)\,\hat{\phi}-(nN-1)\,\hat{\phi}\right) \\ &= \tfrac{1}{(nN-1)P_{\cdot\cdot}\Pi}\left[\sum_{j=1}^{N}P_{\cdot j}\left(\overline{X}_{\cdot j}-\overline{\overline{X}}_{\cdot\cdot}\right)^2-(N-1)\,\hat{\phi}\right]\end{aligned}\ ,$$

que é a expressão (66).

No caso de diferentes n_j $(j=1,2,\cdots,N)$ os estimadores tomam a forma [Sundt (1982)],

$$\hat{\mu} = \overline{\overline{X}}_{\cdot\cdot}=\frac{1}{P_{\cdot\cdot}}\sum_{j=1}^{N}P_{\cdot j}\overline{X}_{\cdot j} \tag{68}$$

$$\hat{\phi} = \frac{1}{\sum_{n=1}^{N}n_j-N}\sum_{j=1}^{N}\sum_{i=1}^{n_j}P_{ij}\left(X_{ij}-\overline{X}_{\cdot j}\right)^2 \tag{69}$$

$$\hat{\psi}^* = \max\left(0,\hat{\psi}\right) \tag{70}$$

com

$$\hat{\psi} = \frac{1}{\sum_{j=1}^{N} P_{\cdot j}\left(1 - \frac{P_{\cdot j}}{P_{\cdot\cdot}}\right)} \left(\sum_{j=1}^{N} P_{\cdot j} \left(\overline{X}_{\cdot j} - \overline{\overline{X}}_{\cdot\cdot} \right)^2 - (N-1)\,\hat{\phi} \right) \tag{71}$$

e $P_{\cdot j} = \sum_{j=1}^{n_j} P_{ij}$, $P_{\cdot\cdot} = \sum_{j=1}^{N} P_{\cdot j}$ e $\overline{X}_{\cdot j} = (1/P_{\cdot j}) \sum_{i=1}^{n_j} P_{ij} X_{ij}$.

Os estimadores (62)-(64), (68), (69) e (71) são centrados. A sua demonstração far-se-á na Secção 6.3.4 a propósito do modelo de Hachemeister. Substituindo os estimadores para os parâmetros estruturais na fórmula (54) ou (61), conforme os casos, obtém-se o Prémio Empírico de Credibilidade.

6.2.4 Modelo original de Bühlmann-Straub

De facto Bühlmann-Straub tratam o modelo de forma algo diferente. Originalmente consideraram um estimador linear sem termo independente e com base em todas as observações de todos os riscos do coletivo. O modelo original tem um aspecto interessante que merece referência. Norberg (1979) afirma que Bühlmann-Straub tiveram um objectivo mais ambicioso, que era essencialmente o da estimação do prémio coletivo μ.

Estes autores propuseram estimar o prémio puro $\mu(\theta_j)$, ou, de forma equivalente, a variável aleatória $X_{n+1,j}$ através dum estimador linear da forma

$$\hat{X}_{n+1,j} = \sum_{j=1}^{N} \sum_{i=1}^{n} b_{ij} X_{ij}, \quad j = 1, 2, \cdots, N\,, \tag{72}$$

sujeito à restrição de não-enviesamento

$$\mathrm{E}\left[\hat{X}_{n+1,j}\right] = \mathrm{E}\left[X_{n+1,j}\right].$$

O critério proposto é o da minimização do erro quadrático médio (ver Teorema 3). Este estimador não admite a existência duma constante independente de X_{ij} e utiliza toda a informação sobre o coletivo. Sem a utilização da constante independente, o estimador vem em geral enviesado, por isso a restrição. O estimador obtido foi o seguinte:

$$\boxed{\begin{array}{rcl} \tilde{X}_{n+1,j} & = & z_j \overline{X}_{\cdot j} + (1 - z_j) \\ \overline{X}_{\cdot l} & = & \sum_{k=1}^{n} \frac{P_{kl}}{P_{\cdot l}} X_{kl}\,, \quad l = 1, 2, \cdots, N \\ \overline{\overline{X}}_{\cdot\cdot}^{*} & = & \sum_{l=1}^{N} \frac{Z_l}{Z_{\cdot}} \overline{X}_l \\ z_l & = & \frac{P_{\cdot l}\psi}{\phi + P_{\cdot l}\psi} \\ z_{\cdot} & = & \sum_{l=1}^{N} z_l \end{array}} \tag{73}$$

Este estimador pode ser obtido através das condições de normalidade (20) e (21) do Teorema 3. Examinando a expressão (73), verifica-se que tem a forma de credibilidade e é semelhante à fórmula (54), substituindo μ por $\overline{\overline{X}}_{\cdot\cdot}^{*}$. A utilização de toda a informação disponível sobre o coletivo e restringindo ao não enviesamento, produziu um estimador para μ, bastando quando necessário a estimação dos parâmetros estruturais ϕ e ψ. Voltar-se-á a falar deste estimador $\overline{\overline{X}}_{\cdot\cdot}^{*}$ na Secção 6.3.6, onde se prova que é o melhor estimador centrado para μ. No entanto, $\overline{\overline{X}}_{\cdot\cdot}^{*}$ depende de ϕ e ψ, a serem estimados em geral. De Vylder (1981a) chama-lhe pseudo-estimador por este facto.

Assim, o estimador (62) para μ pode ser substituído por pelo estimador $\overline{\overline{X}}_{\cdot\cdot}^{*}$, faltando estimar ϕ e ψ. Sundt (1982) propõe a substituição por $\hat{\phi}$ e $\hat{\psi}^*$, dados por (63) e (65), respectivamente, considerando o estimador $\hat{\mu}^*$ assim obtido, como sendo preferível a (62). Isto é, o estimador $\hat{\mu}^*$ para μ vem

$$\hat{\mu}^* = \frac{\sum_{j=1}^{N} \frac{P_{\cdot j}}{P_{\cdot j}\hat{\psi}^* + \hat{\phi}} \overline{X}_{\cdot j}}{\sum_{j=1}^{N} \frac{P_{\cdot j}}{P_{\cdot j}\hat{\psi}^* + \hat{\phi}}}\,. \tag{74}$$

O estimador (74) facilmente se generaliza para diferentes n_j, apenas substituindo $\overline{X}_{\cdot j}$, $P_{\cdot j}$, e, obviamente, $\hat{\phi}$ e $\hat{\psi}^*$ pelos estimadores dados por (69) e (70).

Ainda uma referência ao estimador $\overline{\overline{X}}^*_{..}$. Se se considerar $P_{ij} = 1$, $\forall j$, em que se obtém o estimador de Bühlmann então $P_{.j} = n$, $z_j = z$ e $\overline{\overline{X}}^*_{..} = (1/nN) \sum_{j=1}^{N} \sum_{i=1}^{n} X_{ij}$. Este é o estimador (38), donde se conclui que é o melhor estimador não enviesado para μ, para o caso do modelo de Bühlmann.

Apresenta-se a seguir um exemplo de aplicação a um seguro de grupo, ramo vida, Neste ramo os prémios são calculados com base em tabelas de mortalidade construidas para determinado país. No modelo que se vai construir vai-se assumir que a taxa de mortalidade possa variar, de grupo para grupo. Os grupos deste tipo de seguro são normalmente construidos por empresa. Ora, é natural que as condições de risco variem de grupo para grupo. Por exemplo, as condições de risco de trabalhadores de uma empresa mineira ou de uma empresa de construção civil são totalmente diferentes de uma empresa de serviços, um banco por exemplo, citando casos extremos. Estas diferenças nas condições de risco são no entanto difíceis de quantificar. Vão-se assim representar através de um parâmetro desconhecido. Vai-se assumir que a taxa de mortalidade de um indivíduo com determinada idade pertence a determinado grupo é igual à taxa de mortalidade do país vezes o parâmetro específico.

Exemplo 13 *Seguro de Grupo, Ramo Vida*

Considere-se a seguinte notação:

N_{ksij} - *Variável aleatória que representa o número de pessoas que morrem, com capital seguro x_k, idade s, no grupo j, no ano i.*

x_k - *Capital seguro, que pode variar de ano para ano.*

q_s - *Taxa de mortalidade anual de uma pessoa de idade s, referente a determinado país. Esta taxa é conhecida.*

$q_s\theta_j$ - *Taxa de mortalidade (desconhecida) de uma pessoa com idade s no grupo j.*

n_{ksij} - *Número de pessoas existentes no grupo j com capital seguro x_k, com idade s no ano i.*

$N_{ij} = \sum_{k,s} N_{ksij}$ - *Número de pessoas que morrem no grupo j no ano i.*

$S_{ij} = \sum_k \left(x_k \sum_s N_{ksij}\right)$ - *Indemnizações agregadas do grupo j no ano i.*

Considerem-se as seguintes hipóteses ao modelo.

- *Dado θ_j, N_{ksij} tem distribuição de Poisson de parâmetro $n_{ksij} \times q_s \times \theta_j$.*
- *Dado θ_j, N_{ksij} são independentes para todo o k, s, i.*
- *$\theta_1, \theta_2, \cdots, \theta_j, \cdots$ são independentes e identicamente distribuídos.*

Perante as hipóteses consideradas tem-se que, dado θ_j

$$\sum_s N_{ksij} \sim Poisson\left(\theta_j \sum_s q_s n_{ksij} | \theta_j\right).$$

A função geradora de momentos condicional de S_{ij}, dado θ_j vem

$$\begin{aligned} M_{S_{ij}|\theta_j}(t) &= \prod_k \exp\left\{\theta_j \sum_s q_s n_{ksij}\left(e^{t\,x_k}-1\right)\right\} \\ &= \exp\left\{\theta_j \sum_{s,k} q_s n_{ksij}\left(e^{t\,x_k}-1\right)\right\} \\ &= \exp\left\{\theta_j \sum_s q_s n_{ksij}\left(\frac{\sum_{s,k} q_s n_{ksij}\left(e^{t\,x_k}-1\right)}{\sum_{s,k} q_s n_{ksij}}\right)\right\} \\ &= \exp\left\{\theta_j \sum_s q_s n_{ksij}\left(\sum_k \left(e^{t\,x_k}\frac{\sum_s q_s n_{ksij}}{\sum_{k,s} q_s n_{ksij}}\right)-1\right)\right\}. \end{aligned}$$

Ou seja, conhecido θ_j, S_{ij} tem distribuição de Poisson composta (em que N_{ij} segue uma distribuição de Poisson de parâmetro $\theta_j \sum_{k,s} n_{ksij} q_s$), e função de probabilidade $f_{ij}(x)$:

$$f_{ij}(x) = \frac{\sum_s q_s n_{ksij}}{\sum_{k,s} q_s n_{ksij}} \quad \textit{para} \quad x = x_k; \quad k = 1, 2, \cdots .$$

Note-se que $f_{ij}(x)$ é independente de θ_j, será então natural que para determinar o Prémio de Credibilidade se use apenas os dados referentes a N_{ij} e não a S_{ij}.

O Prémio Puro do grupo j para o ano $(n+1)$ é dado pela expressão

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

Vai propor-se como Estimador de Credibilidade para o Prémio Puro referente ao grupo j para o ano $(n+1)$

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

em que $\tilde{\theta}_j$ é o estimador de credibilidade de θ_j baseado em $N_{1j}, N_{2j}, \cdots, N_{nj}$. Para tal defina-se $X_{ij} = N_{ij}/P_{ij}$ com $P_{ij} = \sum_{k,s} q_s n_{ksij}$. Tem-se portanto $\mathrm{E}[X_{ij}|\theta_j] = \theta_j$ e $\mathrm{V}[X_{ij}|\theta_j] = \theta_j/P_{ij}$. As hipóteses do modelo de Bühlmann-Straub estão verificadas, com $\mu(\theta_j) = \sigma(\theta_j)^2 = \theta_j$. O estimador de credibilidade vem então

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

com $P_{\cdot j} = \sum_i P_{ij}$ e $\overline{X}_{\cdot j} = (1/P_{\cdot j})\sum_i P_{ij} X_{ij} = (1/P_{\cdot j})\sum_i N_{ij}$. Para estimar os parâmetros estruturais $\mu = \mathrm{E}[\theta_j]$ e $\psi = \mathrm{V}[\theta_j]$, considere-se a Tabela 3 com os dados sobre N_{ij} e P_{ij} referentes a 13 grupos:

N_{ij}

j / i	1	2	3	4	5	6	7	8	9	10	11	12	13	$N_{i\cdot}$
1	1	4	2	1	0	1	0	3	5	1	0	0	1	19
2	0	2	1	3	2	5	0	3	5	0	0	1	1	23
3	0	2	3	0	3	2	1	3	2	1	0	1	3	21
$N_{\cdot j}$	1	8	6	4	5	8	1	9	12	2	0	2	5	63

Tabela 3: Dados referentes a N_{ij} e P_{ij}.

Os estimadores $\hat{\mu}$ e $\hat{\psi}$ vêm

$$\begin{aligned}
\hat{\mu} &= \overline{\overline{X}}_{\cdot\cdot} = \sum_{j=1}^{13} \frac{P_{\cdot j}}{P_{\cdot\cdot}} \overline{X}_{\cdot j} = \frac{\sum_{i,j} N_{ij}}{P_{\cdot\cdot}} = \frac{63}{68,82} = 0,915 \\
\hat{\psi} &= \frac{1}{\sum_{j=1}^{N} P_{\cdot j}\left(1 - \frac{P_{\cdot j}}{P_{\cdot\cdot}}\right)} \left(\sum_{j=1}^{N} P_{\cdot j} \left(\overline{X}_{\cdot j} - \overline{\overline{X}}_{\cdot\cdot} \right)^2 - (N-1)\, \overline{\overline{X}}_{\cdot\cdot} \right) = 0,164 \\
\sum_{j=1}^{N} P_{\cdot j} \left(\overline{X}_{\cdot j} - \overline{\overline{X}}_{\cdot\cdot} \right)^2 &= \sum_{j} \frac{1}{P_{\cdot j}} (N_{ij})^2 - \frac{1}{P_{\cdot\cdot}} \left(\sum_{i,j} N_{ij} \right)^2 = 78,734 - \frac{1}{68,82} \times 63^2 = 21,062 \\
\sum_{j=1}^{N} P_{\cdot j} \left(1 - \frac{P_{\cdot j}}{P_{\cdot\cdot}} \right) &= P_{\cdot\cdot} - \frac{\sum_j {P_{\cdot j}}^2}{P_{\cdot\cdot}} = 61,471 \\
\hat{\psi} &= \frac{1}{61,471} (21,062 - 12 \times 0,915) = 0,164.
\end{aligned}$$

Os valores de $\overline{X}_{\cdot j}$ e os valores estimados $z_{j,n}$ e $\tilde{\theta}_{j,N}$ constam da Tabela 5, sendo o Estimador de Credibilidade dado pela formula seguinte,

$$\tilde{\theta}_{j,N} = z_{j,N} \overline{X}_{\cdot j} + (1 - z_{j,N})\, \overline{\overline{X}}_{\cdot\cdot}$$

com $z_{j,N} = P_{\cdot j} / \left(P_{\cdot j} + \overline{\overline{X}}_{\cdot\cdot} / \hat{\psi} \right)$. Substituindo $\tilde{\theta}_{j,N}$ no estimador $\tilde{m}_j$ atrás referido, obtém-se o estimador para o prémio puro a tarifar no 4º período, depois de substituídas com os dados as outras variáveis envolvidas.

P_{ij}

j / i	1	2	3	4	5	6	7	8	9	10	11	12	13	$P_{i\cdot}$
1	1,57	4,13	1,45	1,29	1,31	3,94	0,72	2,45	1,73	1,18	0,32	1,77	1,39	23,25
2	1,58	4,38	1,55	1,39	1,30	3,93	0,73	2,44	1,61	1,10	0,35	1,66	1,43	23,45
3	1,68	4,42	1,73	1,07	1,36	3,75	0,61	2,20	1,47	0,99	0,30	1,16	1,38	22,12
$P_{\cdot j}$	4,83	12,93	4,73	3,75	3,97	11,62	2,06	7,09	4,81	3,27	0,97	4,59	4,20	68,82

Tabela 4: Valores de $P_{\cdot j}$ e $P_{i\cdot}$

j	1	2	3	4	5	6	7	8	9	10	11	12	13
$\overline{X}_{\cdot j}$	0,207	0,619	1,268	1,067	1,259	0,688	0,485	1,269	2,495	0,612	0,000	0,436	1,190
$Z_{j,N}$	0,464	0,699	0,459	0,402	0,416	0,676	0,270	0,560	0,463	0,370	0,148	0,451	0,429
$\hat{\theta}_{j,N}$	0,586	0,708	1,077	0,976	1,058	0,762	0,799	1,113	1,646	0,803	0,780	0,699	1,033

Tabela 5: Valores de $\overline{X}_{\cdot j}$, $Z_{j,N}$ e $\hat{\theta}_{j,N}$.

$\diamondsuit$

6.3 Modelo de regressão

6.3.1 Definições e hipóteses

O modelo de regressão introduzido por Hachemeister (1975) é fortemente inspirado no modelo de Bühlmann & Straub (1970). Neste, foi permitida a variação de E[V[$X_{ij}|\theta_j$]] com o tempo. No modelo de Hachemeister é além disso também permitida a variação de E[X_{ij}].

Hachemeister (1975) afirma a necessidade de estabelecer estimativas de tendência (*trends*) na frequência e montante dos sinistros do seguro automóvel por estado (EUA), porque, afirma, a inflação se tinha tornado um elemento importante na tarifação. Dispondo de dados sobre estes valores em cada estado e por período de tarifação, propôs estimar o valor esperado do montante de indemnizações no período i para o estado j, estabelecendo o seguinte modelo de regressão linear:

$$\mathrm{E}[X_{ij}] = \underline{y}_{ij}\underline{\beta}_j \ . \tag{75}$$

X_{ij} representa a gravidade média das indemnizações no período i no estado j, $\underline{y}_{ij}$ é um vector linha e $\underline{\beta}_j$ um vector coluna, ambos com r componentes e não aleatórios.

Hachemeister fez uma escolha particular para a regressão dada por (75), ao considerar que

$$\underline{y}_{ij}{}' = \begin{bmatrix} 1 \\ i \end{bmatrix} \quad \underline{\beta}_j = \begin{bmatrix} \beta_{0j} \\ \beta_{1j} \end{bmatrix}$$

isto é, considerou que $\mathrm{E}[X_{ij}] = \beta_{0j} + \beta_{1j} i$.

O problema da regressão proposta por (75) situa-se na estimação do vector de constantes $\underline{\beta}_j$, partindo de dados sobre as variáveis observáveis X_{ij} $(i = 1, 2, \cdots, n)$, sendo n o número de períodos, tal como considerado nos modelos tratados anteriormente. Fixando o estado j e considerando as n variáveis conjuntamente, o modelo de regressão para este estado toma a forma

$$\mathrm{E}[\underline{X}_j] = \mathbf{Y}_j \underline{\beta}_j \tag{76}$$

com

$$\underline{X}_j{}' = \begin{bmatrix} X_{1j} \\ X_{2j} \\ \cdots \\ X_{nj} \end{bmatrix} \quad \mathbf{Y}_j = \begin{bmatrix} \underline{y}_{1j} \\ \underline{y}_{2j} \\ \cdots \\ \underline{y}_{nj} \end{bmatrix},$$

em que $\mathbf{Y}_j$ é uma matriz $(n \times r)$. Tendo como ponto de partida a formulação do modelo de Bühlmann & Straub (1970), em que a variância de X_{ij} é proporcional a P_{ij}, que neste caso representa o número de indemnizações:

$$\mathrm{V}[X_{ij}] = \frac{\sigma_j{}^2}{P_{ij}},$$

o estimador para $\underline{\beta}_j$ segundo o modelo clássico GLS (*Generalized Least Squares*) é dado pela expressão

$$\hat{\underline{\beta}}_j = \left(\mathbf{Y}_j{}'\mathbf{P}_j\mathbf{Y}_j\right)^{-1}\mathbf{Y}_j{}'\mathbf{P}_j\mathbf{Y}_j\,, \tag{77}$$

em que $\mathrm{Cov}[\underline{X}_j, \underline{X}_j{}'] = \sigma_j{}^2\mathbf{P}_j{}^{-1}$, $\underline{X}_j = (X_{1j}, \cdots, X_{nj})$ e

$$\begin{bmatrix} P_{1j} & 0 & \ldots & 0 \\ 0 & P_{2j} & \ldots & 0 \\ & & \cdots & \\ 0 & 0 & \ldots & 0 \\ 0 & 0 & \ldots & P_{nj} \end{bmatrix},$$

matriz diagonal $(n \times n)$. Para ver sobre o modelo GLS pode consultar-se um manual de Econometria, por exemplo Johnston & DiNardo (2001) ou Johnston & DiNardo (1997).

A alternativa que se punha era considerar que os dados de todos os estados provinham da mesma população e portanto considerar um modelo de regressão da forma,

$$\mathrm{E}\left[\underline{X}\right] = \mathbf{Y}\underline{\beta} \tag{78}$$

com $\mathbf{Y}$ matriz $(nN \times r)$, $\underline{X}$ matriz $(nN \times 1)$ e $\underline{\beta}$ vector r-dimensional, sendo N o número de estados:

$$\mathbf{Y} = \begin{bmatrix} \mathbf{Y}_1 \\ --- \\ \cdots \\ --- \\ \mathbf{Y}_j \\ --- \\ \cdots \\ --- \\ \mathbf{Y}_N \end{bmatrix}, \quad \mathbf{X} = \begin{bmatrix} \underline{X}_1 \\ --- \\ \cdots \\ --- \\ \underline{X}_j \\ --- \\ \cdots \\ --- \\ \underline{X}_N \end{bmatrix}.$$

O correspondente estimador GLS para $\underline{\beta}$ vem

$$\hat{\underline{\beta}} = (\mathbf{Y}'\mathbf{P}\mathbf{Y})^{-1}\mathbf{Y}'\mathbf{P}\underline{X} \tag{79}$$

em que Cov$[\underline{X}, \underline{X}']$ é uma matriz de ordem $(nN \times nN)$, diagonal com elementos $\sigma^2 {P_{ij}}^{-1}$ $(i = 1, 2, \cdots, n;$ $j = 1, 2, \cdots, N)$.

No entanto ao implementar na prática estas duas soluções, Hachemeister verificou que as linhas de tendência para cada estado vinham suficientemente deslocadas da linha do conjunto do país para se assumir que esta última servia de aplicação para cada estado. Por outro lado, considerando cada linha separadamente, os pontos estavam demasiado dispersos para que a estimativa obtida pudesse ser considerada *fiável*.

Então Hachemeister propôs uma solução de compromisso para cada estado: utilizar um estimador que fosse uma *mistura* entre o estimador particular do estado e o estimador para o conjunto. Isto podia ser resolvido à luz do que tinha sido desenvolvido até então pela Teoria da Credibilidade. No fundo, pressupunha a existência dum coletivo formado pelo conjunto dos estados e em que os dados sobre cada estado não se podiam dissociar pura e simplesmente do conjunto, não se confundindo no entanto entre si.

Vai-se considerar genericamente que X_{ij} representa o risco j $(j = 1, 2, \cdots, N)$ dum conjunto de N riscos, no período i $(i = 1, 2, \cdots, n_j)$, em que o número de observações é de uma forma genérica n_j. A cada risco j está associado um parâmetro θ_j que representa a característica específica do risco dentro do coletivo.

Norberg (1979) formulou as seguintes hipóteses do modelo de Hachemeister (1975):

H1 a) Existe uma função vectorial $\underline{b}(\theta_j)$ definida em Θ e matrizes $\mathbf{Y}_j$ $(n_j \times r)$ conhecidas, com característica $r \leq n_j$, tal que

$$\mathrm{E}\left[\underline{X}_j|\theta_j\right] = \mathbf{Y}_j\underline{b}(\theta_j), \quad j = 1, 2, \cdots, N \tag{80}$$

com

$\underline{X}_j$ $= (X_{1j}, X_{2j}, \cdots, X_{n_j,j})$
$\mathbf{Y}_j$ - matriz $(n_j \times r)$ das variáveis independentes
$\underline{b}(\theta_j)$ - vector r-dimensional
n_j - número de observações para o risco j.

b) Existe uma função positiva $\sigma(\theta_j)^2$ definida em Θ tal que:

$$\mathrm{Cov}\left[\underline{X}_j|\theta_j\right] = \sigma(\theta_j)^2{\mathbf{P}_j}^{-1} \tag{81}$$

em que $\mathbf{P}_j$ é uma matriz diagonal $(n_j \times n_j)$ conhecida,

$$\mathbf{P}_j = \mathrm{diag}(P_{1j}, P_{2j}, ..., P_{n_j,j}),$$

e $\sigma(\theta_j)^2$ é um escalar.

H2 $\equiv$ BS2.

Os parâmetros estruturais que nos modelos de Bühlmann (1967) e Bühlmann & Straub (1970) eram dados pelas expressões (33), são agora dados pelas seguintes expressões:

$$\begin{aligned} \underline{\beta} &= \mathrm{E}\left[\underline{b}(\theta_j)\right] && (82)\\ \mathbf{\Lambda} &= \mathrm{Cov}\left[\underline{b}(\theta_j)\right] && (83)\\ \phi &= \mathrm{E}\left[\sigma(\theta_j)^2\right] && (84) \end{aligned}$$

$\underline{\beta}$ é um vector r-dimensional e $\mathbf{\Lambda}$ uma matriz $(r \times r)$ definida positiva.

6.3.2 Estimador de Credibilidade

Hachemeister (1975) propôs estimar o valor esperado do risco j para determinado período, $\mu(\theta_j)$, ou seja, o prémio puro do risco j dado através da expressão

$$\mu(\theta_j) = \underline{a}'\underline{b}(\theta_j)\,, \tag{85}$$

sendo $\underline{a}$ um vector com r componentes. O estimador proposto é uma função linear das observações sobre o risco j, $\underline{X}_j$:

$$\hat{m}_j = \alpha_{0j} + \sum_{i=1}^{n_j} \alpha_{ij} X_{ij} = \alpha_{0j} + \underline{\alpha}_j{}'\underline{X}_j \tag{86}$$

com as constantes α_{0j} e $\underline{\alpha}_j$ (vector de constantes) obtidas através da minimização de $\mathrm{E}\left[(\mu(\theta_j) - \hat{m}_j)^2\right]$. A solução deste problema é apresentada no Teorema 5.

Teorema 5 *O Estimador de Credibilidade do prémio puro $\mu(\theta_j)$, dado por (85), com base em $\underline{X}_j = (X_{1j}, \cdots, X_{n_j,j})$ é dado pela expressão*

$$\begin{aligned} \tilde{m}_j &= \underline{a}'\left(\mathbf{Z}_j\hat{\underline{b}}_j + (\mathbf{I} - \mathbf{Z}_j)\,\underline{\beta}\right) && (87)\\ \mathbf{Z}_j &= \mathbf{\Lambda}\mathbf{Y}_j{}'\mathbf{P}_j\mathbf{Y}_j\left(\phi\mathbf{I} + \mathbf{\Lambda}\mathbf{Y}_j{}'\mathbf{P}_j\mathbf{Y}_j\right)^{-1} = \mathbf{\Lambda}\mathbf{Y}_j{}'\mathbf{\Phi}_j{}^{-1}\mathbf{Y}_j\left(\mathbf{I} + \mathbf{\Lambda}\mathbf{Y}_j{}'\mathbf{\Phi}_j{}^{-1}\mathbf{Y}_j\right)^{-1} \\ \hat{\underline{b}}_j &= \left(\mathbf{Y}_j{}'\mathbf{P}_j\mathbf{Y}_j\right)^{-1}\mathbf{Y}_j{}'\mathbf{P}_j\underline{X}_j = \left(\mathbf{Y}_j{}'\mathbf{\Phi}_j{}^{-1}\mathbf{Y}_j\right)^{-1}\mathbf{Y}_j{}'\mathbf{\Phi}_j{}^{-1}\underline{X}_j \\ \mathbf{\Phi}_j &= \phi\mathbf{P}_j{}^{-1} \end{aligned}$$

sendo $\underline{\beta}$, $\mathbf{\Lambda}$, ϕ dados por (82)-(84), $\mathbf{Y}$, $\mathbf{P}$ dados através de H1 e H2, e $\mathbf{I}$ a matriz identidade $(r \times r)$.

Dem.

Usando o Teorema 1 o estimador $\hat{m}_j$ dado por (86) é um estimador de credibilidade se e só se satisfaz as condições de normalidade

$$\begin{aligned} \mathrm{E}\left[\hat{m}_j\right] &= \mathrm{E}\left[\mu(\theta_j)\right] && (88)\\ \mathrm{Cov}\left[\hat{m}_j, \underline{X}'\right] &= \mathrm{Cov}\left[\mu(\theta_j), \underline{X}'\right]. && (89) \end{aligned}$$

Partindo de (89) obtém-se

$$\begin{aligned} &\mathrm{Cov}\left[\alpha_{0j} + \underline{\alpha}_j{}'\underline{X}_j{}', \underline{X}_j{}'\right] = \mathrm{Cov}\left[\mu(\theta_j), \underline{X}_j{}'\right] \\ &\Leftrightarrow \underline{\alpha}_j{}'\mathrm{Cov}\left[\underline{X}_j, \underline{X}'\right] = \mathrm{Cov}\left[\mu(\theta_j), \underline{X}_j{}'\right]. && (90) \end{aligned}$$

Mas por outro lado

$$\mathrm{Cov}\left[\underline{X}_j, \underline{X}_j{}'\right] = \mathrm{E}\left[\mathrm{Cov}\left[\underline{X}_j, \underline{X}_j{}'|\theta_j\right]\right] + \mathrm{Cov}\left[\mathrm{E}\left[\underline{X}_j|\theta_j\right], \mathrm{E}\left[\underline{X}_j{}'|\theta_j\right]\right],$$

e aplicando (80), (81) e (87) vem

$$\mathrm{Cov}\left[\underline{X}_j, \underline{X}_j{}'\right] = \phi\mathbf{P}_j{}^{-1} + \mathbf{Y}_j\mathbf{\Lambda}\mathbf{Y}_j{}' = \mathbf{\Phi}_j + \mathbf{Y}_j\mathbf{\Lambda}\mathbf{Y}_j{}'. \tag{91}$$

Por outro lado, pelo mesmo processo tem-se

$$\mathrm{Cov}\left[\mu(\theta_j), \underline{X}_j{}'\right] = \mathrm{Cov}\left[\mathrm{E}\left[\mu(\theta_j)|\theta_j\right], \mathrm{E}\left[\underline{X}_j{}'|\theta_j\right]\right]$$

notando que $\mathrm{Cov}\left[\mu(\theta_j), \underline{X}_j{}'|\theta_j\right] = \underline{0}'$, já que dado θ_j, $\mu(\theta_j)$ é uma constante. Aplicando sucessivamente (80), (85) e (83) obtém-se

$$\mathrm{Cov}\left[\mu(\theta_j), \underline{X}_j{}'\right] = \mathrm{Cov}\left[\underline{a}'\underline{b}(\theta_j), \underline{b}(\theta_j)'\mathbf{Y}_j{}'\right] = \underline{a}'\mathbf{\Lambda}\mathbf{Y}_j{}' . \tag{92}$$

Introduzindo (91) e (92) em (90) obtém-se

$$\underline{\alpha}_j{}'\left(\mathbf{\Phi}_j + \mathbf{Y}_j\mathbf{\Lambda}\mathbf{Y}_j{}'\right) = \underline{a}'\mathbf{\Lambda}\mathbf{Y}_j{}' . \tag{93}$$

Pós-multiplicando ambos os membros da igualdade acima por $\mathbf{\Phi}_j{}^{-1}\mathbf{Y}_j$ [matriz $(n_j \times r)$] tem-se

$$\underline{\alpha}_j{}'\left(\mathbf{\Phi}_j + \mathbf{Y}_j\mathbf{\Lambda}\mathbf{Y}_j{}'\right)\mathbf{\Phi}_j{}^{-1}\mathbf{Y}_j = \underline{\alpha}_j{}'\mathbf{Y}_j\left(\mathbf{I} + \mathbf{\Lambda}\mathbf{Y}_j{}'\mathbf{\Phi}_j{}^{-1}\mathbf{Y}_j\right) = \underline{a}'\mathbf{\Lambda}\mathbf{Y}_j{}'\mathbf{\Phi}_j{}^{-1}\mathbf{Y}_j ,$$

pós-multiplicando por $\left(\mathbf{I} + \mathbf{\Lambda}\mathbf{Y}_j{}'\mathbf{\Phi}^{-1}\mathbf{Y}_j\right)^{-1}$ (esta inversa existe, veja-se Lema 1 de Sundt (1979)), tem-se

$$\underline{\alpha}_j{}'\mathbf{Y}_j = \underline{a}'\mathbf{\Lambda}\mathbf{Y}_j{}'\mathbf{\Phi}_j{}^{-1}\mathbf{Y}_j\left(\mathbf{I} + \mathbf{\Lambda}\mathbf{Y}_j{}'\mathbf{\Phi}_j{}^{-1}\mathbf{Y}_j\right)^{-1} .$$

Fazendo $\mathbf{Z}_j = \mathbf{\Lambda}\mathbf{Y}_j{}'\mathbf{\Phi}_j{}^{-1}\mathbf{Y}_j\left(\mathbf{I} + \mathbf{\Lambda}\mathbf{Y}_j{}'\mathbf{\Phi}_j{}^{-1}\mathbf{Y}_j\right)^{-1}$ [matriz $(r \times r)$], tem-se

$$\underline{\alpha}_j{}'\mathbf{Y}_j = \underline{a}'\mathbf{Z}_j . \tag{94}$$

Por outro lado, de (93) resolvendo a equação em ordem a $\underline{\alpha}_j{}'\mathbf{\Phi}_j$ e introduzindo (94) tem-se

$$\underline{\alpha}_j{}'\mathbf{\Phi}_j = \underline{a}'\mathbf{\Lambda}\mathbf{Y}_j{}' - \underline{\alpha}_j{}'\mathbf{Y}_j\mathbf{\Lambda}\mathbf{Y}_j{}' = \underline{a}'\left(\mathbf{I} - \mathbf{Z}_j\right)\mathbf{\Lambda}\mathbf{Y}_j{}' ,$$

pós-multiplicando sucessivamente por $\mathbf{\Phi}^{-1}$, $\underline{X}_j{}'$ e fazendo $\hat{\underline{b}}_j = \left(\mathbf{Y}_j\mathbf{\Phi}_j{}^{-1}\mathbf{Y}_j\right)^{-1}\mathbf{Y}_j{}'\mathbf{\Phi}_j{}^{-1}\underline{X}_j$ obtém-se

$$\begin{aligned} \underline{\alpha}_j{}' &= \underline{a}'\left(\mathbf{I} - \mathbf{Z}_j\right)\mathbf{\Lambda}\mathbf{Y}_j{}'\mathbf{\Phi}_j{}^{-1} & (95) \\ \underline{\alpha}_j{}'\underline{X}_j &= \underline{a}'\left(\mathbf{I} - \mathbf{Z}_j\right)\mathbf{\Lambda}\mathbf{Y}_j{}'\mathbf{\Phi}_j{}^{-1}\underline{X}_j = \underline{a}'\left(\mathbf{I} - \mathbf{Z}_j\right)\mathbf{\Lambda}\mathbf{Y}_j{}'\mathbf{\Phi}_j{}^{-1}\mathbf{Y}_j\hat{\underline{b}}_j . & (96) \end{aligned}$$

Por outro lado, pós-multiplicando (95) por $\mathbf{Y}_j$ e atendendo a (94), conclui-se que $\left(\mathbf{I} - \mathbf{Z}_j\right)\mathbf{\Lambda}\mathbf{Y}_j{}'\mathbf{\Phi}_j{}^{-1}\mathbf{Y}_j = \mathbf{Z}_j$, donde (96) vem

$$\underline{\alpha}_j{}'\underline{X}_j = \underline{a}'\mathbf{Z}_j\hat{\underline{b}}_j . \tag{97}$$

Da condição (88) obtém-se

$$\alpha_{0j} + \alpha_j{}'\mathrm{E}\left[\underline{X}_j\right] = \mathrm{E}\left[\mu(\theta_j)\right] ,$$

atendendo à propriedade iterativa do valor esperado, a (80) e (85), tem-se

$$\alpha_{0j} + \alpha_j{}'\mathrm{E}\left[\underline{b}(\theta_j)\right] = \underline{a}'\mathbf{Y}_j\mathrm{E}\left[\underline{b}(\theta_j)\right] ,$$

resolvendo em α_{0j}, factorizando, atendendo a (82) e (94), tem-se

$$\alpha_{0j} = \left(\underline{a}' - \underline{\alpha}_j{}'\mathbf{Y}_j\right)\underline{\beta} = \underline{a}'\left(\mathbf{I} - \mathbf{Z}_j\right)\underline{\beta} . \tag{98}$$

Somando membro a membro (97) e (98) obtém-se finalmente o estimador de credibilidade

$$\tilde{m}_j = \underline{a}'\left(\mathbf{Z}_J\hat{\underline{b}}_j + \left(\mathbf{I} - \mathbf{Z}_j\right)\underline{\beta}\right) ,$$

como se pretendia demonstrar. □

Note-se que o estimador de credibilidade (87) foi encontrado tendo em conta uma escolha especial da matriz $\mathbf{\Phi}_j = \phi\mathbf{P}_j{}^{-1}$. Contudo, as expressões (87) são válidas para uma matriz $\mathbf{\Phi}_j$ arbitrária definida positiva, como facilmente se verifica.

$\hat{\underline{b}}_j$ é o bem conhecido estimador GLS de $\underline{b}(\theta_j)$ obtido pela minimização de

$$\left(\underline{X}_j - \mathbf{Y}_j\underline{b}\right)' \mathbf{\Phi}_j{}^{-1} \left(\underline{X}_j - \mathbf{Y}_j\underline{b}\right) .$$

O estimador (87) de $\mu(\theta_j) = \underline{a}'\underline{b}(\theta_j)$ é obtido substituindo $\underline{b}(\theta_j)$ por uma *média ponderada* do estimador $\hat{\underline{b}}_j$ baseado na experiência individual e a *média coletiva* $\underline{\beta}$ e é semelhante à Formula de Credibilidade (1) proposta por Whitney. No entanto o Factor de Credibilidade é aqui uma matriz $(r \times r)$. No caso do modelo de Bühlmann-Straub ou de Bühlmann, o factor de credibilidade é um número positivo. No caso do modelo de regressão de Hachemeister a Matriz de Credibilidade pode conter elementos negativos (veja-se o Exemplo 14no final da Secção 6.3.4), mas é uma matriz definida positiva.

Ainda uma nota em relação à matriz de credibilidade $\mathbf{Z}_j$. Esta pode ainda ser escrita na forma

$$\mathbf{Z}_j = \mathbf{\Lambda}\mathbf{Y}_j{}'\mathbf{\Phi}_j{}^{-1}\mathbf{Y}_j \left(\mathbf{I} + \mathbf{\Lambda}\mathbf{Y}_j{}'\mathbf{\Phi}_j{}^{-1}\mathbf{Y}_j\right)^{-1} = \left[\left(\mathbf{I} + \mathbf{\Lambda}\mathbf{Y}_j{}'\mathbf{\Phi}_j{}^{-1}\mathbf{Y}_j\right)\left(\mathbf{\Lambda}\mathbf{Y}_j{}'\mathbf{\Phi}_j{}^{-1}\mathbf{Y}_j\right)^{-1}\right]^{-1} \tag{99}$$

atendendo à propriedade da inversa dum produto de duas matrizes quadradas $(AB)^{-1} = B^{-1}A^{-1}$ e notando que a inversa de $\mathbf{\Lambda}\mathbf{Y}_j{}'\mathbf{\Phi}_j{}^{-1}\mathbf{Y}_j$ existe, pois tem-se um produto de duas matrizes quadrada se o determinante $|\mathbf{\Lambda}\mathbf{Y}_j{}'\mathbf{\Phi}_j{}^{-1}\mathbf{Y}_j| = |\mathbf{\Lambda}||\mathbf{Y}_j{}'\mathbf{\Phi}_j{}^{-1}\mathbf{Y}_j| \neq 0$, porque $\mathbf{\Lambda}$ e $\mathbf{\Phi}_j{}^{-1}$ são definidas positivas e a característica de $\mathbf{Y}_j$, $c(\mathbf{Y}_j) = r \neq n_j$, por H1. Desenvolvendo o produto de matrizes dentro do parêntesis recto, (99) vem ainda

$$\mathbf{Z}_j = \left(\mathbf{I} + \left(\mathbf{\Lambda}\mathbf{Y}_j{}'\mathbf{\Phi}_j{}^{-1}\mathbf{Y}_j\right)^{-1}\right)^{-1} \tag{100}$$

ou introduzindo $P_{\cdot j} = \sum_{i=1}^{n_j} P_{ij}$ (traço da matriz $\mathbf{P}_j$) e fazendo $\mathbf{K}_j = P_{\cdot j}\left(\mathbf{\Lambda}\mathbf{Y}_j{}'\mathbf{\Phi}_j{}^{-1}\mathbf{Y}_j\right)^{-1}$, $\mathbf{Z}_j$ vem

$$\mathbf{Z}_j = P_{\cdot j}\left(P_{\cdot j}\mathbf{I} + \mathbf{K}_j\right)^{-1} ,$$

que é a correspondente forma do factor de credibilidade (54) do modelo de Bühlmann & Straub (1970), para o modelo de Hachemeister (1975).

Considerem-se agora alguns Casos Particulares ao modelo:

1. Se se fizer $r = 1$, $\mathbf{Y}_j{}' = \left[\begin{array}{cccc} 1 & 1 & \cdots & 1 \end{array}\right]$, $\underline{b}(\theta_j) = \mu(\theta_j)$ e $\underline{a} = 1$, então a expressão toma a forma

$$\mathrm{E}\left[\underline{X}_j|\theta_j\right] = \left[\begin{array}{c} 1 \\ 1 \\ \cdots \\ 1 \end{array}\right] \mu(\theta_j) .$$

   além disso os parâmetros estruturais (82)-(84) vêm $\beta = \mathrm{E}\left[\mu(\theta_j)\right] = \mu$, $\phi = \mathrm{E}\left[\sigma(\theta_j)^2\right]$ e $\mathbf{\Lambda} = \mathrm{Cov}\left[\mu(\theta_j)\right] = \mathrm{V}\left[\mu(\theta_j)\right] = \psi$. O estimador $\hat{\underline{b}}_j = \left(\mathbf{Y}_j{}'\mathbf{P}_j\mathbf{Y}\right)^{-1}\mathbf{Y}_j\mathbf{P}_j\underline{X}_j = (1/P_{\cdot j})\sum_{i=1}^{n_j} P_{ij}X_{ij} = \overline{X}_{\cdot j}$, a matriz de credibilidade $\mathbf{Z}_j$ é um escalar

$$z_j = \left(\psi\sum_{i=1}^{n_j} P_{ij}\right)\left(\phi + \sum_{i=1}^{n_j} P_{ij}\right)^{-1} = \frac{\psi P_{\cdot j}}{\phi + \psi P_{\cdot j}}$$

   e o Estimador de Credibilidade do Prémio Puro de Risco $\mu(\theta_j)$ vem

$$\tilde{m}_j = \frac{\psi P_{\cdot j}}{\phi + \psi P_{\cdot j}}\overline{X}_{\cdot j} + \frac{\phi}{\phi + \psi P_{\cdot j}}\mu$$

   que é o estimador (61) do modelo de Bühlmann & Straub (1970) para n_j períodos genéricos.

2. Se além disso se fizer $n_j = n$ então obtém-se (54).

3. Voltando a (1.) mas fazendo $\mathbf{P}_j = \mathbf{I}_{(n_j \times n_j)}$, então

$$\tilde{m}_j = \frac{n_j}{n_j + \phi/\psi}\overline{X}_{\cdot j} + \frac{\phi/\psi}{n_j + \phi/\psi}\mu$$

   com $\overline{X}_{\cdot j} = (1/n_j)\sum_{i=1}^{n_j} X_{ij}$, que é o estimador (35) do modelo de Bühlmann para n_j períodos.

4. Se adicionalmente a (3.) se fizer $n_j = n$, obtém-se o estimador (36) de Bühlmann.

6.3.3 Generalização do conceito de Estimador de Credibilidade

A propósito do modelo de regressão, o conceito de Estimador de Credibilidade apresentado no Capítulo 5 do presente texto para uma variável aleatória m, pode ser generalizado para um vector aleatório $\underline{m}$.

Considere-se um vector aleatório $\underline{m} = (m_1, m_2, \cdots, m_s)$ e considere-se dois estimadores de $\underline{m}$,

$$\underline{\dot{m}} = \left(\dot{m}_1, \dot{m}_2, \cdots, \dot{m}_s\right) \text{ e } \underline{\ddot{m}} = \left(\ddot{m}_1, \ddot{m}_2, \cdots, \ddot{m}_s\right).$$

Diz-se que $\underline{\dot{m}}$ é Melhor Estimador que $\ddot{m}$ relativamente à Função de Perca Quadrática Esperada, se

$$\mathrm{E}\left[\left(\dot{m}_i - m_i\right)^2\right] \leq \mathrm{E}\left[\left(\ddot{m}_i - m_i\right)^2\right] \qquad i = 1, 2, \cdots, s\,,$$

com desigualdade estrita para pelo menos um i [Sundt (1981)].

Sejam $\underline{X}$ e $\underline{m}$ genericamente definidos como dois vectores aleatórios conjuntamente distribuídos e $\underline{X}$ observável. Chame-se $\underline{\hat{m}}$ um estimador linear de $\underline{m}$ baseado em $\underline{X}$ se $\underline{\hat{m}}$ pode ser escrito na forma

$$\underline{\hat{m}} = \underline{a} + \mathbf{A}\underline{X} \tag{101}$$

onde $\underline{a}$ é um vector não aleatório e $\mathbf{A}$ é uma matriz também não aleatória.

Chame-se a $\underline{\tilde{m}}$ Estimador de Credibilidade baseado em $\underline{X}$, como sendo o melhor estimador linear de $\underline{m}$. As condições que permitem a obtenção do estimador $\underline{\tilde{m}}$, assim como algumas propriedades são apresentadas no seguinte teorema [Sundt (1981)] e que é uma generalização do Teorema 1.

Teorema 6 *a) Seja $\underline{\hat{m}}$ um estimador linear de $\underline{m}$ definido em por (101). Então $\underline{\hat{m}}$ é um estimador de credibilidade de $\underline{m}$ se satisfaz as seguintes condições de normalidade*

$$\mathrm{E}\left[\underline{\hat{m}}\right] = \mathrm{E}\left[\underline{m}\right] \tag{102}$$

$$\mathrm{Cov}\left[\underline{\hat{m}}, \underline{X}'\right] = \mathrm{Cov}\left[\underline{m}, \underline{X}'\right] \tag{103}$$

b) Se os estimadores $\underline{\hat{m}}$ e $\underline{\dot{m}}$, lineares da forma (101), verificam as condições de normalidade (102) e (103), então eles são iguais quase certamente.

c) O estimador de credibilidade $\underline{\tilde{m}}$ satisfaz então

$$\mathrm{Cov}\left[\underline{m}, \underline{\tilde{m}}'\right] = \mathrm{Cov}\left[\underline{\tilde{m}}\right] - \mathrm{Cov}\left[\underline{m} - \underline{\tilde{m}}\right]. \tag{104}$$

□

A definição de Estimador de Credibilidade vectorial atrás apresentada vai permitir apresentar outro aspecto interessante do Estimador de Credibilidade (87) para o Prémio Puro $\mu(\theta_j)$:

$$\tilde{m}_j = \underline{a}'\left[\mathbf{Z}_j\underline{\hat{b}}_j + (\mathbf{I} - \mathbf{Z}_j)\,\underline{\beta}\right]$$

Este aspecto vai ser apresentado no teorema seguinte.

Teorema 7 *Seja $\tilde{m}_j$ o estimador de credibilidade de $\mu(\theta_j)$ dado por (87). então,*

$$\tilde{m}_j = \underline{a}'\underline{\tilde{b}}_j$$

em que $\underline{\tilde{b}}_j$ é o estimador de credibilidade de $\underline{b}(\theta_j)$ com base em $\underline{X}_j$.

Dem.
$\underline{\tilde{b}}_j$ é o estimador de credibilidade de $\underline{b}(\theta_j)$ se e só se verifica as condições de normalidade (102) e (103) do Teorema 6:

$$\begin{aligned} \mathrm{E}\left[\underline{\tilde{b}}_j\right] &= \mathrm{E}\left[\underline{b}(\theta_j)\right] \\ \mathrm{Cov}\left[\underline{\tilde{b}}_j, \underline{X}_j{}'\right] &= \mathrm{Cov}\left[\underline{b}(\theta_j), \underline{X}_j{}'\right]. \end{aligned}$$

Em relação à primeira condição, tem-se

$$\mathrm{E}\left[\underline{\tilde{b}}_j\right] = \mathbf{Z}_j \mathrm{E}\left[\underline{\hat{b}}_j\right] + (\mathbf{I} - \mathbf{Z}_j)\,\underline{\beta}$$

pois a matriz $\mathbf{Z}_j$ é não aleatória. Em relação a $\mathrm{E}\left[\underline{\hat{b}}_j\right]$, sendo $\underline{\hat{b}}_j$ dado por (87), tem-se

$$\begin{aligned} \mathrm{E}\left[\underline{\hat{b}}_j\right] &= \left(\mathbf{Y}_j{}'\mathbf{P}_j\mathbf{Y}_j\right)^{-1}\mathbf{Y}_j{}'\mathbf{P}_j\mathrm{E}\left[\underline{X}_j\right] = \left(\mathbf{Y}_j{}'\mathbf{P}_j\mathbf{Y}_j\right)^{-1}\mathbf{Y}_j{}'\mathbf{P}_j\mathrm{E}\left[\underline{b}(\theta_j)\right] \\ &= \left(\mathbf{Y}_j{}'\mathbf{P}_j\mathbf{Y} - j\right)^{-1}\mathbf{Y}_j{}'\mathbf{P}_j\underline{\beta} = \underline{\beta}, \end{aligned} \tag{105}$$

atendendo à propriedade iterativa do valor esperado, a (80) e a (82). Donde

$$\mathrm{E}\left[\underline{\tilde{b}}_j\right] = \mathrm{E}\left[\underline{b}(\theta_j)\right] = \underline{\beta}.$$

Em relação à segunda condição tem-se que

$$\mathrm{Cov}\left[\underline{\tilde{b}}_j, \underline{X}_j{}'\right] = \mathrm{Cov}\left[\mathbf{Z}_j\underline{\hat{b}}_j + (\mathbf{I} - \mathbf{Z}_j)\,\underline{\beta}; \underline{X}_j{}'\right] = \mathbf{Z}_j\mathrm{Cov}\left[\underline{\hat{b}}_j, \underline{X}_j{}'\right]$$

pois $\mathbf{Z}_j$ e $\underline{\beta}$ são matrizes de constantes. Substituindo $\underline{\hat{b}}_j$ pela respectiva expressão (87) e atendendo a (91), tem-se sucessivamente

$$\begin{aligned} \mathrm{Cov}\left[\underline{\tilde{b}}_j, \underline{X}_j{}'\right] &= \mathbf{Z}_j\mathrm{Cov}\left[\left(\mathbf{Y}_j{}'\mathbf{\Phi}_j{}^{-1}\mathbf{Y}_j\right)^{-1}\mathbf{Y}_j{}'\mathbf{\Phi}_j{}^{-1}\underline{X}_j, \underline{X}_j{}'\right] \\ &= \mathbf{Z}_j\left(\mathbf{Y}_j{}'\mathbf{\Phi}_j{}^{-1}\mathbf{Y}_j\right)^{-1}\mathbf{Y}_j{}'\mathbf{\Phi}_j{}^{-1}\mathrm{Cov}\left[\underline{X}_j, \underline{X}_j{}'\right] \\ &= \mathbf{Z}_j\left(\mathbf{Y}_j{}'\mathbf{\Phi}_j{}^{-1}\mathbf{Y}_j\right)^{-1}\mathbf{Y}_j{}'\mathbf{\Phi}_j{}^{-1}\left(\mathbf{\Phi}_j + \mathbf{Y}_j\mathbf{\Lambda}\mathbf{Y}_j{}'\right). \end{aligned}$$

Simplificando e atendendo a (100) tem-se

$$\begin{aligned} \mathrm{Cov}\left[\underline{\tilde{b}}_j, \underline{X}_j{}'\right] &= \mathbf{Z}_j\left(\left(\mathbf{Y}_j{}'\mathbf{\Phi}_j{}^{-1}\mathbf{Y}_j\right)^{-1} + \mathbf{\Lambda}\right)\mathbf{Y}_j{}' \\ &= \mathbf{Z}_j\left(\left(\mathbf{Y}_j{}'\mathbf{\Phi}_j{}^{-1}\mathbf{Y}_j\right)^{-1}\mathbf{\Lambda}^{-1} + \mathbf{I}\right)\mathbf{\Lambda}\mathbf{Y}_j{}' \\ &= \mathbf{Z}_j\mathbf{Z}_j{}^{-1}\mathbf{\Lambda}\mathbf{Y}_j{}' = \mathbf{\Lambda}\mathbf{Y}_j{}'. \end{aligned}$$

Por outro lado,

$$\mathrm{Cov}\left[\underline{b}(\theta_j), \underline{X}_j{}'\right] = \mathrm{E}\left[\mathrm{Cov}\left[\underline{b}(\theta_j), \underline{X}_j{}'|\theta_j\right]\right] + \mathrm{Cov}\left[\mathrm{E}\left[\underline{b}(\theta_j)|\theta_j\right], \mathrm{E}\left[\underline{X}_j{}'|\theta_j\right]\right] = \mathrm{Cov}\left[\underline{b}(\theta_j), \underline{b}(\theta_j)'\right]\mathbf{Y}_j{}'$$

pois $\mathrm{Cov}\left[\underline{b}(\theta_j), \underline{X}_j{}'|\theta_j\right] = \mathbf{0}$ porque dado θ_j, $\underline{b}(\theta_j)$ é um vector não aleatório, donde usando (83), tem-se

$$\mathrm{Cov}\left[\underline{b}(\theta_j), \underline{X}_j{}'\right] = \mathbf{\Lambda}\mathbf{Y}' = \mathrm{Cov}\left[\underline{\tilde{b}}_j, \underline{X}_j{}'\right]\quad. \tag{106}$$

□

6.3.4 Estimação dos parâmetros estruturais

Os parâmetros $\underline{\beta}$, $\mathbf{\Lambda}$ e ϕ não são conhecidos em geral, pelo que será necessário a sua estimação através de estatísticas com base nos vectores observáveis, $\underline{X}_j$, $j = 1, 2, \cdots, N$.

Estimadores de Hachemeister. Prémio empírico de credibilidade Hachemeister propôs os seguintes estimadores, aqui generalizados para n_j períodos genéricos e de acordo com o modelo desenvolvido na Secção 6.3.1:

$$\underline{\hat{\beta}} = \left(\mathbf{Y}'\mathbf{P}\mathbf{Y}\right)^{-1}\mathbf{Y}'\mathbf{P}\underline{X} \tag{107}$$

com $\mathbf{Y}$, $\underline{X}$ e $\mathbf{P}$ definidos em (78) e (79).

$$\hat{\phi} = \left(\sum_{k=1}^{N} n_k - Nr\right)^{-1}\sum_{j=1}^{N} {S_j}'^2 \tag{108}$$

com

$$ {S_j}'^2 = \left(\underline{X}_j - \mathbf{Y}_j\underline{\hat{b}}_j\right)'\mathbf{P}_j\left(\underline{X}_j - \mathbf{Y}_j\underline{\hat{b}}_j\right) = \left(\underline{X}_j{}'\mathbf{P}_j\underline{X}_j - \underline{\hat{b}}_j'{\mathbf{Y}_j}'\mathbf{P}_j\mathbf{Y}_j\underline{\hat{b}}_j\right).$$

$$\hat{\mathbf{\Lambda}} = (1/2)\left(\mathbf{H}+\mathbf{H}'\right) \tag{109}$$

com

$$\begin{aligned}
\mathbf{H} &= \mathbf{\Pi}^{-1}\left(\mathbf{G} - (N-1)\left(\mathbf{Y}'\mathbf{P}\mathbf{Y}\right)^{-1}\hat{\phi}\right)\\
\mathbf{G} &= \sum_{j=1}^{N}\left(\mathbf{Y}'\mathbf{P}\mathbf{Y}\right)^{-1}\left({\mathbf{Y}_j}'\mathbf{P}_j\mathbf{Y}_j\right)\left(\underline{\hat{b}}_j - \underline{\hat{\beta}}\right)\left(\underline{\hat{b}}_j - \underline{\hat{\beta}}\right)'\\
\mathbf{\Pi} &= \mathbf{I} - \sum_{j=1}^{N}\left(\mathbf{Y}'\mathbf{P}\mathbf{Y}\right)^{-1}\left({\mathbf{Y}_j}'\mathbf{P}_j\mathbf{Y}_j\right)\left(\mathbf{Y}'\mathbf{P}\mathbf{Y}\right)^{-1}\left({\mathbf{Y}_j}'\mathbf{P}_j\mathbf{Y}_j\right)
\end{aligned} \tag{110}$$

Os estimadores (107)-(109) são centrados. A sua demonstração vai ser feita seguidamente.

Teorema 8 *Sejam os estimadores (107) e (108), respectivamente dos parâmetros estruturais* $\underline{\beta} = \mathrm{E}[\underline{b}(\theta_j)]$ *e* $\phi = \mathrm{E}[\sigma(\theta_j)^2]$. *Então,* $\mathrm{E}[\underline{\hat{\beta}}] = \underline{\beta}$ *e* $\mathrm{E}[\hat{\phi}] = \phi$.

Dem.

Seja o estimador $\underline{\hat{\beta}}$. Este estimador pode ser escrito na forma

$$\underline{\hat{\beta}} = \sum_{k=1}^{N}\left(\mathbf{Y}\mathbf{P}\mathbf{Y}\right)^{-1}{\mathbf{Y}_k}'\mathbf{P}_k\underline{X}_k$$

porque

$$\mathbf{Y}'\mathbf{P}\underline{X} = \left[{\mathbf{Y}_1}'|{\mathbf{Y}_2}'|\cdots|{\mathbf{Y}_N}'\right]\begin{bmatrix}\mathbf{P}_1 & \mathbf{0} & ... & \mathbf{0}\\ \mathbf{0} & \mathbf{P}_2 & ... & \mathbf{0}\\ & & \cdots & \\ \mathbf{0} & \mathbf{0} & ... & \mathbf{P}_N\end{bmatrix}\begin{bmatrix}\underline{X}_1\\ --\\ \underline{X}_2\\ --\\ \cdots\\ --\\ \underline{X}_N\end{bmatrix} = \sum_{k=1}^{N}{\mathbf{Y}_k}'\mathbf{P}_k\underline{X}_k. \tag{111}$$

$\underline{\hat{\beta}}$ pode ainda ser escrita na forma:

$$\underline{\hat{\beta}} = \sum_{k=1}^{N}\left(\mathbf{Y}'\mathbf{P}\mathbf{Y}\right)^{-1}\left({\mathbf{Y}_k}'\mathbf{P}_k\mathbf{Y}_k\right)\underline{\hat{b}}_k \tag{112}$$

introduzindo a matriz identidade $\left(\mathbf{Y}_k\mathbf{P}_k\mathbf{Y}_k\right)\left(\mathbf{Y}_k\mathbf{P}_k\mathbf{Y}_k\right)^{-1}$ e atendendo a (87). Então o valor esperado de $\underline{\hat{\beta}}$ vem, usando (112)

$$\mathrm{E}[\underline{\hat{\beta}}] = \sum_{k=1}^{N}\left(\mathbf{Y}'\mathbf{P}\mathbf{Y}\right)^{-1}\left(\mathbf{Y}_k\mathbf{P}_k\mathbf{Y}_k\right)\mathrm{E}[\underline{\hat{b}}_k].$$

Tendo em atenção o resultado (105) sabe-se que $\mathrm{E}[\underline{\hat{\beta}}_k] = \underline{\beta}$, e por outro lado, da mesma forma que (111) tem-se

$$\mathbf{Y}'\mathbf{P}\mathbf{Y} = \sum_{k=1}^{N} \mathbf{Y}_k{}'\mathbf{P}_k\mathbf{Y}_k\,, \tag{113}$$

donde se obtém que $\mathrm{E}[\underline{\hat{\beta}}] = \underline{\beta}$. Fica provado o não-enviesamento de $\underline{\hat{\beta}}$.

Para provar o não-enviezamento de $\hat{\phi}$ dado por (108), estimador do parâmetro estrutural $\phi = \mathrm{E}[\sigma^2(\theta_j)]$, note-se que ${S_j{}'}^2$ pode ser escrito da seguinte forma

$${S_j{}'}^2 = \left(\underline{X}_j{}'\mathbf{P}_j\underline{X}_j\right) - \underline{\hat{b}}_j\mathbf{Y}_j{}'\mathbf{P}_j\mathbf{Y}_j\underline{\hat{b}}_j$$

porque desenvolvendo, atendendo a que $\mathbf{P}$ é uma matriz simétrica e à propriedade da transposta dum produto de matrizes:

$$\left(\underline{X}_j - \mathbf{Y}_j\underline{\hat{b}}_j\right)'\mathbf{P}_j\left(\underline{X}_j - \mathbf{Y}_j\underline{\hat{b}}_j\right) = \left(\underline{X}_j{}'\mathbf{P}_j - \underline{\hat{b}}_j'\mathbf{Y}_j{}'\mathbf{P}_j{}'\right)\left(\underline{X}_j - \mathbf{Y}_j\underline{\hat{b}}_j\right)\,;$$

desenvolvendo o produto tem-se

$${S_j{}'}^2 = \underline{X}_j{}'\mathbf{P}_j\underline{X}_j - \underline{X}_j{}'\mathbf{P}_j\mathbf{Y}_j\underline{\hat{b}}_j - \underline{\hat{b}}_j{}'\mathbf{Y}_j'\mathbf{P}_j\underline{X}_j + \underline{\hat{b}}_j{}'\mathbf{Y}_j{}'\mathbf{P}_j\mathbf{Y}_j\underline{\hat{b}}_j\,,$$

introduzindo a matriz identidade $\mathbf{I} = \left(\mathbf{Y}_j{}'\mathbf{P}_j\mathbf{Y}_j\right)^{-1}\left(\mathbf{Y}_j{}'\mathbf{P}_j\mathbf{Y}_j\right)$ (matriz $r \times r$) na 2^a e 3^a parcelas obtém-se

$$\begin{aligned} {S_j{}'}^2 &= \underline{X}_j{}'\mathbf{P}_j\underline{X}_j - \underline{X}_j{}'\mathbf{P}_j\mathbf{Y}_j\left(\mathbf{Y}_j{}'\mathbf{P}_j\mathbf{Y}_j\right)^{-1}\left(\mathbf{Y}_j{}'\mathbf{P}_j\mathbf{Y}_j\right)\underline{\hat{b}}_j \\ &\quad -\underline{\hat{b}}_j'\left(\mathbf{Y}_j{}'\mathbf{P}_j\mathbf{Y}_j\right)\left(\mathbf{Y}_j{}'\mathbf{P}_j\mathbf{Y}_j\right)^{-1}\mathbf{Y}_j{}'\mathbf{P}_j\underline{X}_j + \underline{\hat{b}}_j'\mathbf{Y}_j{}'\mathbf{P}_j\mathbf{Y}_j\underline{\hat{b}}_j\,, \end{aligned}$$

atendendo a (87) e que a transposta de $\underline{\hat{b}}_j$

$$\underline{\hat{b}}_j' = \underline{X}_j{}'\mathbf{P}_j\mathbf{Y}_j\left(\mathbf{Y}_j{}'\mathbf{P}_j\mathbf{Y}_j\right)^{-1}\,, \tag{114}$$

tem-se

$$\begin{aligned} {S_j{}'}^2 &= \underline{X}_j{}'\mathbf{P}_j\underline{X}_j - \underline{\hat{b}}_j'\left(\mathbf{Y}_j{}'\mathbf{P}_j\mathbf{Y}_j\right)\underline{\hat{b}}_j - \underline{\hat{b}}_j'\left(\mathbf{Y}_j{}'\mathbf{P}_j\mathbf{Y}_j\right)\underline{\hat{b}}_j + \underline{\hat{b}}_j'\left(\mathbf{Y}_j{}'\mathbf{P}_j\mathbf{Y}_j\right)\underline{\hat{b}}_j \\ &= \underline{X}_j{}'\mathbf{P}_j\underline{X}_j - \underline{\hat{b}}_j'\left(\mathbf{Y}_j{}'\mathbf{P}_j\mathbf{Y}_j\right)\underline{\hat{b}}_j \end{aligned}$$

Desenvolvendo $\underline{\hat{b}}_j$ tem-se ainda

$${S_j{}'}^2 = \underline{X}_j{}'\mathbf{P}_j\underline{X}_j - \underline{X}_j{}'\mathbf{P}_j\mathbf{Y}_j\left(\mathbf{Y}_j{}'\mathbf{P}_j\mathbf{Y}_j\right)^{-1}\left(\mathbf{Y}_j{}'\mathbf{P}_j\mathbf{Y}_j\right)\left(\mathbf{Y}_j{}'\mathbf{P}_j\mathbf{Y}_j\right)^{-1}\mathbf{Y}_j{}'\mathbf{P}_j\underline{X}_j\,;$$

simplificando e fazendo

$$\mathbf{R} = \mathbf{P}_j - \mathbf{P}_j\mathbf{Y}_j(\mathbf{Y}_j\mathbf{P}_j\mathbf{Y}_j)^{-1}\mathbf{Y}_j{}'\mathbf{P}_j \qquad (n_j \times n_j) \tag{115}$$

obtém-se

$${S_j{}'}^2 = \underline{X}_j{}'\mathbf{R}\underline{X}_j\,. \tag{116}$$

A matriz $\mathbf{R}$ é uma matriz simétrica pois $\mathbf{R}' = \mathbf{R}$, já que $\mathbf{P}_j$ é uma matriz diagonal e portanto simétrica. A expressão (116) pode ser escrita na forma

$${S_j{}'}^2 \;=\; \sum_{k=1}^{n_j}\sum_{l=1}^{n_j} X_{kj}r_{kl}X_{lj} = \sum_{k=1}^{n_j}\sum_{l=1}^{n_j} r_{kl}X_{kj}X_{lj}\,,$$

em que r_{kl} é o elemento da linha k e coluna l de $\mathbf{R}$.

Então o valor esperado de ${S_j{}'}^2$ dado θ_j vem

$$\begin{aligned} \mathrm{E}[{S_j{}'}^2|\theta_j] &= \sum_{k=1}^{n_j}\sum_{l=1}^{n_j} r_{kl}\mathrm{E}[X_{kj}X_{lj}|\theta_j] \\ &= \sum_{k=1}^{n_j}\sum_{l=1}^{n_j} r_{kl}\left(\mathrm{Cov}[X_{kj}, X_{lj}|\theta_j] + \mathrm{E}[X_{kj}|\theta_j]\mathrm{E}[X_{lj}|\theta_j]\right)\,. \end{aligned}$$

Como dado θ_j as variáveis em períodos diferentes não estão correlacionadas [por (81)] e na segunda parcela voltando à forma (116) obtém-se

$$\mathrm{E}[{S_j'}^2|\theta_j] = \sum_{k=1}^{n_j}\sum_{l=1}^{n_j} r_{kk}\mathrm{V}[X_{kj}|\theta_j] + \mathrm{E}[{X_j}'|\theta_j]\mathbf{R}\mathrm{E}[X_j|\theta_j]\,;$$

notando que a primeira parcela é o traço ou soma dos elementos da diagonal principal de $\sigma^2(\theta_j)\mathbf{R}{\mathbf{P}_j}^{-1}$ e usando (80) obtém-se

$$\mathrm{E}[{S_j'}^2|\theta_j] = \mathrm{tr}[\sigma^2(\theta_j)\mathbf{R}{\mathbf{P}_j}^{-1}] + \underline{b}(\theta_j)'{\mathbf{Y}_j}'\mathbf{R}\mathbf{Y}\underline{b}(\theta_j)$$

e desenvolvendo o 2^o membro através de (115), tem-se

$$\begin{aligned}\mathrm{E}[{S_j'}^2|\theta_j] \;&=\; \sigma^2(\theta_j)\mathrm{tr}\left[\left(\mathbf{P}_j - \mathbf{P}_j\mathbf{Y}_j\,({\mathbf{Y}_j}'\mathbf{P}_j\mathbf{Y}_j)^{-1}{\mathbf{Y}_j}'\mathbf{P}_j\right){\mathbf{P}_j}^{-1}\right]\\ &\quad + \underline{b}(\theta_j)'{\mathbf{Y}_j}'\left(\mathbf{P}_j - \mathbf{P}_j\mathbf{Y}_j\,({\mathbf{Y}_j}'\mathbf{P}_j\mathbf{Y}_j)^{-1}{\mathbf{Y}_j}'\mathbf{P}_j\right)\mathbf{Y}_j\underline{b}(\theta_j)\,;\end{aligned}$$

notando que o traço da soma é igual à soma dos traços de matrizes; simplificando a segunda parcela, tem-se

$$\mathrm{E}[{S_j'}^2|\theta_j] = \sigma^2(\theta_j)\left(\mathrm{tr}[\mathbf{I}_{n_j}] - \mathrm{tr}[\mathbf{P}_j\mathbf{Y}_j\,({\mathbf{Y}_j}'\mathbf{P}_j\mathbf{Y}_j)^{-1}{\mathbf{Y}_j}'] + \mathrm{tr}[\underline{b}(\theta_j)'({\mathbf{Y}_j}'\mathbf{P}_j\mathbf{Y}_j - {\mathbf{Y}_j}'\mathbf{P}_j\mathbf{Y}_j)\underline{b}(\theta_j)]\right)\;;$$

na segunda parcela obtém-se uma matriz nula e atendendo à propriedade do traço dum produto de matrizes, (tr[**ABC**] = tr[**BAC**] supondo que as matrizes continuam multiplicáveis) obtém-se

$$\begin{aligned}\mathrm{E}[{S_j'}^2|\theta_j] \;&=\; \sigma^2(\theta_j)\left(n_j - \mathrm{tr}\left[({\mathbf{Y}_j}'\mathbf{P}_j\mathbf{Y}_j)^{-1}({\mathbf{Y}_j}'\mathbf{P}_j\mathbf{Y}_j)\right]\right)\\ &=\; \sigma^2(\theta_j)\,(n_j - \mathrm{tr}[\mathbf{I}_r]) = \sigma^2(\theta_j)\,(n_j - r)\;.\end{aligned}$$

Então,

$$\mathrm{E}\left[\sum_{j=1}^{N}{S_j'}^2\right] = \sum_{j=1}^{N}\phi\,(n_j - r) = \phi\left(\sum_{j=1}^{N}n_j - Nr\right)\;,$$

donde $\hat{\phi}$ dado por (108) é centrado. □

De forma a mostrar que o estimador $\hat{\mathbf{\Lambda}}$ dado por (109) é centrado, considerem-se os seguintes resultados para utilização posterior:

Resultado 1

$$\mathbf{G} \;=\; \sum_{j=1}^{N}(\mathbf{Y}'\mathbf{P}\mathbf{Y})^{-1}\left({\mathbf{Y}_j}'\mathbf{P}_j\mathbf{Y}_j\right)\hat{\underline{b}}_j\hat{\underline{b}}_j' - \sum_{j=1}^{N}\sum_{k=1}^{N}(\mathbf{Y}'\mathbf{P}\mathbf{Y})^{-1}\left({\mathbf{Y}_j}'\mathbf{P}_j\mathbf{Y}_j\right)(\mathbf{Y}'\mathbf{P}\mathbf{Y})^{-1}\left({\mathbf{Y}_k}'\mathbf{P}_k\mathbf{Y}_k\right)\hat{\underline{b}}_k\hat{\underline{b}}_j' \quad (117)$$

$$\mathrm{E}[\hat{\underline{b}}_k\hat{\underline{b}}_j'] = \delta_{jk}\mathrm{Cov}[\hat{\underline{b}}_k,\hat{\underline{b}}_j'] + \underline{\beta}\underline{\beta}' \quad (118)$$

$$\mathrm{Cov}[\hat{\underline{b}}_j,\hat{\underline{b}}_j'] = \phi\left({\mathbf{Y}_j}'\mathbf{P}_j\mathbf{Y}_j\right)^{-1} + \mathbf{\Lambda} \quad (119)$$

onde δ_{jk} representa o índice de Kronecker.

Dem.

- A expressão (117) resulta do desenvolvimento de (110), i.e.

$$\begin{aligned}\mathbf{G} \;=\; &\sum_{j=1}^{N}(\mathbf{Y}'\mathbf{P}\mathbf{Y})^{-1}(\mathbf{Y}'\mathbf{P}\mathbf{Y})\,\hat{\underline{b}}_j\hat{\underline{b}}_j' - \sum_{j=1}^{N}(\mathbf{Y}'\mathbf{P}\mathbf{Y})^{-1}\left({\mathbf{Y}_j}'\mathbf{P}_j\mathbf{Y}_j\right)\hat{\underline{b}}_j\underline{\beta}'\\ &-\sum_{j=1}^{N}(\mathbf{Y}'\mathbf{P}\mathbf{Y})^{-1}\left({\mathbf{Y}_j}'\mathbf{P}_j\mathbf{Y}_j\right)\underline{\beta}_j\hat{\underline{b}}_j' + \sum_{j=1}^{N}(\mathbf{Y}'\mathbf{P}\mathbf{Y})^{-1}\left({\mathbf{Y}_j}'\mathbf{P}_j\mathbf{Y}_j\right)\underline{\beta}_j\underline{\beta}_j'\,,\end{aligned}$$

  usando os resultados (112) e (113) e simplificando.

- Para a expressão (118) tem-se

$$\begin{aligned}\mathrm{E}[\hat{\underline{b}}_k\hat{\underline{b}}_j'] &= \mathrm{Cov}[\hat{\underline{b}}_k,\underline{b}_j'] + \mathrm{E}[\hat{\underline{b}}_k]\mathrm{E}[\hat{\underline{b}}_j'] \\ &= \mathrm{Cov}[\left(\mathbf{Y}_k'\mathbf{P}_k\mathbf{Y}_k\right)^{-1}\mathbf{Y}_k'\mathbf{P}_k\underline{X}_k, \underline{X}_k'\mathbf{P}_j\mathbf{Y}_j\left(\mathbf{Y}_k'\mathbf{P}_k\mathbf{Y}_k\right)^{-1}] + \underline{\beta\beta'}\,,\end{aligned}$$

usando os resultados (87) e (105), e usando a hipótese de independência entre riscos de H2 obtém-se (118).

- A expressão (119) obtém-se a partir de

$$\mathrm{Cov}[\hat{\underline{b}}_j,\hat{\underline{b}}_j'] = \left(\mathbf{Y}_j'\mathbf{P}_j\mathbf{Y}_j\right)^{-1}\mathbf{Y}_j'\mathbf{P}_j\mathrm{Cov}[\underline{X}_j,\underline{X}_j']\mathbf{P}_j\mathbf{Y}_j\left(\mathbf{Y}_j'\mathbf{P}_j\mathbf{Y}_j\right)^{-1}$$

atendendo a (91) e simplificando. □

Teorema 9 *O valor esperado da estatística* $\mathbf{G}$ *dada através de (117) é dado pela expressão*

$$\mathrm{E}[\mathbf{G}] = (N-1)(\mathbf{Y}'\mathbf{P}\mathbf{Y})^{-1}\phi + \left(\mathbf{I} - \sum_{j=1}^{N}(\mathbf{Y}'\mathbf{P}\mathbf{Y})^{-1}\left(\mathbf{Y}_j'\mathbf{P}_j\mathbf{Y}_j\right)(\mathbf{Y}'\mathbf{P}\mathbf{Y})^{-1}\left(\mathbf{Y}_j'\mathbf{P}_j\mathbf{Y}_j\right)\right)$$

e $\mathrm{E}[\mathbf{H}] = \mathrm{E}[\hat{\mathbf{\Lambda}}] = \mathbf{\Lambda}$.

De (117) tem-se introduzindo (118) e (119):

$$\begin{aligned}\mathrm{E}[\mathbf{G}] &= \sum_{j=1}^{N}(\mathbf{Y}'\mathbf{P}\mathbf{Y})^{-1}\left(\mathbf{Y}_j'\mathbf{P}_j\mathbf{Y}_j\right)\left(\phi\left(\mathbf{Y}_j'\mathbf{P}_j\mathbf{Y}_j\right)^{-1} + \mathbf{\Lambda} + \underline{\beta\beta'}\right) \\ &\quad - \sum_{j=1}^{N}\sum_{k=1}^{N}(\mathbf{Y}'\mathbf{P}\mathbf{Y})^{-1}(\mathbf{Y}_j'\mathbf{P}_j\mathbf{Y}_j)(\mathbf{Y}'\mathbf{P}\mathbf{Y})^{-1}\left(\mathbf{Y}_k'\mathbf{P}_k\mathbf{Y}_k\right)\left(\delta_{jk}\left(\phi\left(\mathbf{Y}_j'\mathbf{P}_j\mathbf{Y}_j\right)^{-1} + \mathbf{\Lambda}\right) + \underline{\beta\beta'}\right)\,;\end{aligned}$$

somando em k na segunda parcela e desenvolvendo as parcelas, vem

$$\begin{aligned}\mathrm{E}[\mathbf{G}] &= \sum_{j=1}^{N}(\mathbf{Y}'\mathbf{P}\mathbf{Y})^{-1}\phi + \sum_{j=1}^{N}(\mathbf{Y}'\mathbf{P}\mathbf{Y})^{-1}(\mathbf{Y}_j'\mathbf{P}_j\mathbf{Y}_j)\left(\mathbf{\Lambda} + \underline{\beta\beta'}\right) \\ &\quad - \sum_{j=1}^{N}(\mathbf{Y}'\mathbf{P}\mathbf{Y})^{-1}(\mathbf{Y}_j'\mathbf{P}_j\mathbf{Y}_j)(\mathbf{Y}'\mathbf{P}\mathbf{Y})^{-1}(\mathbf{Y}_j'\mathbf{P}_j\mathbf{Y}_j)\left(\phi\left(\mathbf{Y}_j'\mathbf{P}_j\mathbf{Y}_j\right)^{-1} + \mathbf{\Lambda}\right) \\ &\quad - \sum_{j=1}^{N}\sum_{k=1}^{N}(\mathbf{Y}'\mathbf{P}\mathbf{Y})^{-1}(\mathbf{Y}_j'\mathbf{P}_j\mathbf{Y}_j)(\mathbf{Y}'\mathbf{P}\mathbf{Y})^{-1}\left(\mathbf{Y}_k'\mathbf{P}_k\mathbf{Y}_k\right)\underline{\beta\beta'}\,;\end{aligned}$$

simplificando obtém-se finalmente

$$\mathrm{E}[\mathbf{G}] = (N-1)(\mathbf{Y}'\mathbf{P}\mathbf{Y})^{-1}\phi + \left(\mathbf{I} - \sum_{j=1}^{N}(\mathbf{Y}'\mathbf{P}\mathbf{Y})^{-1}\left(\mathbf{Y}_j'\mathbf{P}_j\mathbf{Y}_j\right)(\mathbf{Y}'\mathbf{P}\mathbf{Y})^{-1}\left(\mathbf{Y}_j'\mathbf{P}_j\mathbf{Y}_j\right)\right)$$

Então tomando o estimador $\mathbf{H}$ (para $\mathbf{\Lambda}$) dado através de (110) é centrado e obviamente $\mathrm{E}[\hat{\mathbf{\Lambda}}] = \mathbf{\Lambda}$. □

Note-se que foi tomado para $\mathbf{\Lambda}$ o estimador $\hat{\mathbf{\Lambda}} = (\mathbf{H} + \mathbf{H}')/2$ para garantir a simetria, uma vez que $\mathbf{\Lambda}$ é uma matriz de covariâncias. Contudo, é preciso que seja no mínimo uma matriz semi-definida o que não é garantido. Caso não seja semi-definida poder-se-á adoptar o procedimento citado em Norberg (1982), substituindo-a pela sua projecção no espaço convexo de matrizes semi-definidas positivas e simétricas.

Se se substituir os estimadores (107)-(109) no Estimador de Credibilidade (87) obtém-se o Prémio Empírico de Credibilidade:

$$\boxed{\begin{array}{rcl}\tilde{m}_{j,N} & = & \underline{a}'\left(\mathbf{Z}_{j,N}\underline{\hat{b}}_j + (\mathbf{I}-\mathbf{Z}_{j,N})\,\underline{\hat{\beta}}\right) \\ \mathbf{Z}_{j,N} & = & \hat{\mathbf{\Lambda}}\mathbf{Y}_j{}'\mathbf{P}_j\mathbf{Y}_j\left(\hat{\phi}\mathbf{I} + \hat{\mathbf{\Lambda}}\mathbf{Y}_j{}'\mathbf{P}_j\mathbf{Y}\right)^{-1} \\ \underline{\hat{b}}_j & = & \left(\mathbf{Y}_j{}'\mathbf{P}_j\mathbf{Y}_j\right)^{-1}\mathbf{Y}_j\mathbf{Y}_j\underline{X}_j \\ \underline{\hat{\beta}} & = & (\mathbf{Y}'\mathbf{P}\mathbf{Y})^{-1}\mathbf{Y}\mathbf{P}\underline{X}\end{array}} \tag{120}$$

Comparando este estimador com as expressões (77) e (79) inicialmente apresentados por Hachemeister, verifica-se que o seu objectivo foi atingido pois o estimador acima é uma *média* ponderada entre (77) e (79), sendo os pesos atribuídos através da matriz de credibilidade $\mathbf{Z}_{j,N}$.

6.3.5 Alguns Casos Particulares

Considerem-se alguns casos particulares para os estimadores (107)-(109). Vão-se percorrer os casos particulares considerados nas alíneas (1.)-(4.) da Secção 6.3.2.

1. Para esta particularização tem-se para as seguintes expressões:

$$\begin{array}{rcl}\mathbf{Y}_j{}'\mathbf{P}_j\mathbf{Y}_j & = & \sum_{i=1}^{n_j} P_{ij} = P_{\cdot j} \\ \mathbf{Y}'\mathbf{P}\mathbf{Y} & = & \sum_{j=1}^{N}\mathbf{Y}_j{}'\mathbf{P}_j\mathbf{Y}_j = \sum_{j=1}^{N} P_{\cdot j} = P_{\cdot\cdot} \\ \underline{\hat{b}}_j & = & \frac{1}{P_{\cdot j}}\sum_{i=1}^{n_j} P_{ij}X_{ij} = \overline{X}_{\cdot j}\end{array}$$

Para o estimador $\underline{\hat{\beta}}$ obtém-se então

$$\underline{\hat{\beta}} = (\mathbf{Y}'\mathbf{P}\mathbf{Y})^{-1}\mathbf{Y}'\mathbf{P}\underline{X}_j = \sum_{k=1}^{N}(\mathbf{Y}'\mathbf{P}\mathbf{Y})^{-1}(\mathbf{Y}_k{}'\mathbf{P}_k\mathbf{Y}_k)\underline{\hat{b}}_k = \frac{1}{P_{\cdot\cdot}}\sum_{k=1}^{N} P_{\cdot k}\overline{X}_{\cdot k} = \overline{\overline{X}}_{\cdot\cdot},$$

que é o estimador (68) para μ do modelo de Bühlmann-Straub para n_j $(j = 1, 2, \cdots, N)$ períodos de observação. Para o estimador $\hat{\phi}$ obtendo o valor particular ${S_j{}'}^2$, tem-se

$$\begin{array}{rcl}{S_j{}'}^2 & = & \left(\underline{X}_j - \mathbf{Y}\underline{\hat{b}}_j\right)'\mathbf{P}\left(\underline{X}_j - \mathbf{Y}\underline{\hat{b}}_j\right) = \left(\underline{X}_j{}'\mathbf{P}_j\underline{X}_j - \underline{\hat{b}}_j'\mathbf{Y}_j{}'\mathbf{P}_j\underline{\hat{b}}_j\right) \\ & = & \sum_{i=1}^{n_j} P_{ij}\left(X_{ij} - \overline{X}_{\cdot j}\right)^2 = \sum_{i=1}^{n_j} P_{ij}{X_{ij}}^2 - P_{\cdot j}{\overline{X}_{\cdot j}}^2\end{array}$$

e $\hat{\phi}$ vem

$$\begin{array}{rcl}\hat{\phi} & = & \frac{1}{\sum_{j=1}^{N} n_j - N}\sum_{j=1}^{N}{S_j{}'}^2 = \frac{1}{\sum_{j=1}^{N} n_j - N}\sum_{j=1}^{N}\sum_{i=1}^{n_j} P_{ij}\left(X_{ij} - \overline{X}_{\cdot j}\right)^2 \\ & = & \frac{1}{\sum_{j=1}^{N} n_j - N}\left(\sum_{i=1}^{N}\sum_{i=1}^{n_j} P_{ij}{X_{ij}}^2 - \sum_{j=1}^{N} P_{\cdot j}{\overline{X}_{\cdot j}}^2\right)\end{array}$$

que é o estimador (69) do modelo de Bühlmann-Straub para n_j períodos.

Relativamente ao estimador $\hat{\mathbf{\Lambda}}$, os valores particulares das estatísticas $\mathbf{G}$ e $\mathbf{\Pi}$ vêm,

$$\mathbf{G}=\sum_{j=1}^{N}\left(\mathbf{Y}'\mathbf{P}\mathbf{Y}\right)^{-1}\left(\mathbf{Y}_j{}'\mathbf{P}_j\mathbf{Y}_j\right)\left(\underline{\hat{b}}_j-\underline{\hat{\beta}}\right)\left(\underline{b}_j{}'-\underline{\beta}_j{}'\right)=\frac{1}{P_{\cdot\cdot}}\sum_{j=1}^{N}P_{\cdot j}\left(\overline{X}_{\cdot j}-\overline{\overline{X}}_{\cdot\cdot}\right)^2,$$

ou alternativamente a partir de (117), tem-se a particularização

$$\begin{aligned}\mathbf{G} &= \sum_{j=1}^{N}\left(\mathbf{Y}'\mathbf{P}\mathbf{Y}\right)^{-1}\left(\mathbf{Y}_j{}'\mathbf{P}_j\mathbf{Y}_j\right)\underline{\hat{b}}_j\underline{\hat{b}}_j' - \sum_{j=1}^{N}\sum_{k=1}^{N}\left(\mathbf{Y}'\mathbf{P}\mathbf{Y}\right)^{-1}\left(\mathbf{Y}_j{}'\mathbf{P}_j\mathbf{Y}_j\right)\left(\mathbf{Y}'\mathbf{P}\mathbf{Y}\right)^{-1}\left(\mathbf{Y}_k{}'\mathbf{P}_k\mathbf{Y}_k\right)\underline{\hat{b}}_k\underline{\hat{b}}_j' \\ &= \frac{1}{P_{\cdot\cdot}}\sum_{j=1}^{N}P_{\cdot j}\overline{X}_{\cdot j}{}^2-\overline{\overline{X}}_{\cdot\cdot}{}^2 .\end{aligned}$$

$$\mathbf{\Pi}=\mathbf{I}_r-\sum_{j=1}^{N}\left(\mathbf{Y}'\mathbf{P}\mathbf{Y}\right)^{-1}\left(\mathbf{Y}_j{}'\mathbf{P}_j\mathbf{Y}_j\right)\left(\mathbf{Y}'\mathbf{P}\mathbf{Y}\right)^{-1}\left(\mathbf{Y}_j{}'\mathbf{P}_j\mathbf{Y}_j\right)=1-\sum_{j=1}^{N}\frac{P_{\cdot j}{}^2}{P_{\cdot\cdot}{}^2}.$$

Então $\hat{\mathbf{\Lambda}}$ vem,

$$\begin{aligned}\hat{\mathbf{\Lambda}} &= \mathbf{\Pi}^{-1}\left(\mathbf{G}-(N-1)\left(\mathbf{Y}'\mathbf{P}\mathbf{Y}\right)^{-1}\hat{\phi}\right) \\ &= \frac{1}{1-\sum_{j=1}^{N}\frac{P_{\cdot j}{}^2}{P_{\cdot\cdot}{}^2}}\left(\frac{1}{P_{\cdot\cdot}}\sum_{j=1}^{N}P_{\cdot j}\left(\overline{X}_{\cdot j}-\overline{\overline{X}}_{\cdot\cdot}\right)^2-\frac{1}{P_{\cdot\cdot}}(N-1)\hat{\phi}\right) \\ &= \frac{1}{P_{\cdot\cdot}-\sum_{j=1}^{N}\frac{P_{\cdot j}{}^2}{P_{\cdot\cdot}}}\left(\sum_{j=1}^{N}P_{\cdot j}\left(\overline{X}_{\cdot j}-\overline{\overline{X}}_{\cdot\cdot}\right)^2-(N-1)\hat{\phi}\right) \\ &= \frac{1}{\sum_{j=1}^{N}P_{\cdot j}-\sum_{j=1}^{N}\frac{P_{\cdot j}{}^2}{P_{\cdot\cdot}}}\left(\sum_{j=1}^{N}P_{\cdot j}\left(\overline{X}_{\cdot j}-\overline{\overline{X}}_{\cdot\cdot}\right)^2-(N-1)\hat{\phi}\right) \\ &= \frac{1}{\sum_{j=1}^{N}P_{\cdot j}\left(1-\frac{P_{\cdot j}}{P_{\cdot\cdot}}\right)}\left(\sum_{j=1}^{N}P_{\cdot j}\left(\overline{X}_{\cdot j}-\overline{\overline{X}}_{\cdot\cdot}\right)^2,-(N-1)\hat{\phi}\right)\end{aligned}$$

que é o estimador dado por (45) do modelo de Bühlmann-Straub para n_j $(j=1,2,\cdots,N)$ períodos.

2. Neste caso particular em que se consideraram os casos de (1.) mas com $n_j=n$ $(j=1,2,\cdots,n)$, obtêm-se imediatamente os estimadores (62)-(64) do modelo de Bühlmann-Straub.

3. Para esta alínea em que se considerou adicionalmente $\mathbf{P}_j=\mathbf{I}_{n_j}$, vão-se obter os estimadores (43)-(45) do modelo de Bühlmann para n_j períodos genéricos:

   - $$\underline{\hat{\beta}}=\frac{1}{\sum_{j=1}^{N}n_j}\sum_{j=1}^{N}n_j\overline{X}_{\cdot j}=\frac{1}{\sum_{j=1}^{N}n_j}\sum_{j=1}^{N}\sum_{i=1}^{n_j}X_{ij}=\overline{\overline{X}}_{\cdot\cdot},$$

     que é o estimador (43);

   - $$\hat{\phi}=\frac{1}{\sum_{j=1}^{N}n_j-N}\sum_{j=1}^{N}\sum_{i=1}^{n_j}\left(X_{ij}-\overline{X}_{\cdot j}\right)^2$$

     que é o estimador considerado em (44).

   - $$\hat{\mathbf{\Lambda}}=\frac{1}{\sum_{j=1}^{N}n_j\left(1-\frac{n_j}{\sum_{k=1}^{N}n_k}\right)}\left(\sum_{j=1}^{N}n_j\left(\overline{X}_{\cdot j}-\overline{\overline{X}}_{\cdot\cdot}\right)^2-(N-1)\hat{\phi}\right),$$

     que é a expressão (45).

4. Neste caso em que se consideraram a particularização de (3.) mas adicionalmente $n_j = n$, obtêm-se os estimadores (38)-(40) do modelo de Bühlmann.

6.3.6 Estimadores Mais Gerais

Retomando o problema de estimação dos parâmetros estruturais tratado na Secção 6.3.4, note-se ainda que Norberg (1979) e De Vylder (1978) propuseram estimar a média colectiva $\underline{\beta}$, dado por (82), através de

$$\hat{\underline{\beta}} = \sum_{j=1}^{N} \mathbf{F}_j \hat{\underline{b}}_j \tag{121}$$

em que $\mathbf{F}_j$ é uma matriz $(r \times r)$ sujeita à restrição $\sum_{j=1}^{N} \mathbf{F}_j = \mathbf{I}$ de forma a assegurar o não-enviesamento.

Norberg (1979) propõe estimar ϕ dado por (84) através de

$$\hat{\phi} = \sum_{j=1}^{N} S_{j,w}^2 \tag{122}$$

com

$$S_{j,w}^2 = \left(\underline{X}_j - \mathbf{Y}_j \hat{\underline{b}}_j\right)' \mathbf{W}_j \left(\underline{X}_j - \mathbf{Y}_j \hat{\underline{b}}_j\right)$$

e $\mathbf{W}_j$ matriz $(n_j \times n_j)$ positiva definida sujeita à restrição

$$\sum_{j=1}^{N} \left(\operatorname{tr}\left[\mathbf{P}_j{}^{-1}\mathbf{W}_j\right] - \operatorname{tr}\left[\left(\mathbf{Y}_j{}'\mathbf{P}_j\mathbf{Y}_j\right)^{-1}\left(\mathbf{Y}_j{}'\mathbf{P}_j\mathbf{Y}_j\right)\right]\right) = 1 \ .$$

Propõe ainda estimar $\mathbf{\Lambda}$, dado por (83), através do estimador

$$\hat{\hat{\mathbf{\Lambda}}} = \frac{1}{2}\sum_{j=1}^{N} \left(\mathbf{H}_j \left(\hat{\underline{b}}_j \hat{\underline{b}}_j' - \hat{\phi}\left(\mathbf{Y}_j{}'\mathbf{P}_j\mathbf{Y}_j\right)^{-1}\right) + \left(\hat{\underline{b}}_j \hat{\underline{b}}_j' - \hat{\phi}\left(\mathbf{Y}_j{}'\mathbf{P}_j\mathbf{Y}_j\right)^{-1}\right)\mathbf{H}_j{}'\right) - \underline{\hat{\beta}\hat{\beta}'} \tag{123}$$

em que $\mathbf{H}_j$ $(j = 1, 2, \cdots, N)$ são matrizes $(n_j \times n_j)$ sujeitas à restrição $\sum_{j=1}^{N} \mathbf{H}_j = \mathbf{I}$.

Segundo Norberg (1979) os estimadores (121)-(123) são consistentes sob restrições suaves, sendo os estimadores (107)-(109) obtidos através de escolhas particulares das matrizes $\mathbf{F}_j$, $\mathbf{W}_j$ e $\mathbf{H}_j$, respectivamente.

De Vylder (1978) estabelece a escolha óptima das matrizes $\mathbf{F}_j$ de (121), minimizando o traço de $\operatorname{Cov}[\hat{\hat{\underline{\beta}}}]$. Se $\hat{\hat{\underline{\beta}}}$ for um escalar, então o princípio invocado é o da variância do estimador. Calcule-se a $\operatorname{Cov}[\hat{\hat{\underline{\beta}}}]$:

$$\operatorname{Cov}[\hat{\hat{\underline{\beta}}}] = \operatorname{Cov}[\hat{\hat{\underline{\beta}}}, \hat{\hat{\underline{\beta}}}{}'] = \operatorname{Cov}[\sum_{j=1}^{N} \mathbf{F}_j \hat{\underline{b}}_j ; \sum_{j=1}^{N} \hat{\underline{b}}_j' \mathbf{F}_j{}'] = \sum_{j=1}^{N}\sum_{k=1}^{N} \mathbf{F}_j \operatorname{Cov}[\hat{\underline{b}}_j, \hat{\underline{b}}_j'] \mathbf{F}_k{}' \, ;$$

usando o resultados (118), somando em k e usando (119) tem-se

$$\operatorname{Cov}[\hat{\hat{\underline{\beta}}}] \quad = \quad \sum_{j=1}^{N}\sum_{k=1}^{N} \mathbf{F}_j \delta_{jk} \operatorname{Cov}[\hat{\underline{b}}_j, \hat{\underline{b}}_k'] \mathbf{F}_k{}' = \sum_{j=1}^{N} \mathbf{F}_j \operatorname{Cov}[\hat{\underline{b}}_j, \hat{\underline{b}}_j'] \mathbf{F}_j{}' = \sum_{j=1}^{N} \mathbf{F}_j \left(\phi\left(\mathbf{Y}_j{}'\mathbf{P}_j\mathbf{Y}_j\right)^{-1} + \mathbf{\Lambda}\right) \mathbf{F}_j{}' . \tag{124}$$

Considere-se o seguinte lema demonstrado em De Vylder (1978), que vai permitir a escolha óptima $\mathbf{F}_j{}'^{*}$:

Lema 1 *Sejam $\mathbf{M}_1, \mathbf{M}_2, \cdots, \mathbf{M}_k$ matrizes $(q \times q)$, simétricas, definidas positivas e sejam $\mathbf{X}_1, \mathbf{X}_2, \cdots, \mathbf{X}_k$ matrizes $(q \times q)$ de variáveis. Então o mínimo do traço $\operatorname{tr}\left[\sum_{i=1}^{k} \mathbf{X}_i \mathbf{M}_i \mathbf{X}_i{}'\right]$, sujeito à restrição $\sum_{i=1}^{N} \mathbf{X}_i = \mathbf{I}$ é atingido para*

$$\mathbf{X}_j = \left(\sum_{i=1}^{k} \mathbf{M}_i{}^{-1}\right)^{-1} (\mathbf{M}_j)^{-1} \quad (j = 1, 2, \cdots, k) \tag{125}$$

□

Partindo da expressão (124) fazendo

$$\mathbf{M}_j = \phi\big(\mathbf{Y}_j{}'\mathbf{P}_j\mathbf{Y}\big)^{-1} + \mathbf{\Lambda} = \mathbf{Z}_j{}^{-1} \quad (j = 1, 2, \cdots, N)$$

porque partindo da expressão (100) tem-se

$$\mathbf{Z}_j{}^{-1} = \phi\big(\mathbf{\Lambda}\mathbf{Y}_j{}'\mathbf{P}_j\mathbf{Y}_j\big)^{-1} + \mathbf{I} = \phi\big(\mathbf{Y}_j{}'\mathbf{P}_j\mathbf{Y}_j\big)^{-1}\mathbf{\Lambda}^{-1} + \mathbf{I}\,;$$

pós-multiplicando ambos os membros por $\mathbf{\Lambda}$ tem-se

$$\mathbf{Z}_j{}^{-1}\mathbf{\Lambda} = \phi\big(\mathbf{Y}_j{}'\mathbf{P}_j\mathbf{Y}_j\big)^{-1} + \mathbf{\Lambda}\,. \tag{126}$$

Então a escolha óptima de $\mathbf{F}_j$ vem, aplicando (125),

$$\mathbf{F}_j{}^* = \left(\sum_{j=1}^{N} \mathbf{\Lambda}^{-1}\mathbf{Z}_j\right)^{-1} \mathbf{\Lambda}^{-1}\mathbf{Z}_j = \left(\sum_{j=1}^{N} \mathbf{Z}_j\right)^{-1} \mathbf{Z}_j\ .$$

O estimador óptimo da forma (121) vem então

$$\underline{\hat{\hat{\beta}}}^* = \sum_{j=1}^{N} \mathbf{F}_j{}^*\underline{\hat{b}}_j = \left(\sum_{j=1}^{N} \mathbf{Z}_j\right)^{-1} \sum_{j=1}^{N} \mathbf{Z}_j\underline{\hat{b}}_j\,. \tag{127}$$

Se se considerar o caso particular da alínea (2.) da Secção 6.3.2, em que se obtém o modelo de Bühlmann-Straub, então (127) corresponde ao estimador $\overline{\overline{X}}_{\cdot\cdot}{}^*$ de (73), daí ter-se chamado o melhor estimador não-enviesado para a média coletiva μ:

$$\overline{\overline{X}}_{\cdot\cdot}{}^* = \left(\sum_{j=1}^{N} \frac{P_{\cdot j}}{P_{\cdot j} + \phi/\psi}\right)^{-1} \sum_{j=1}^{N} \frac{P_{\cdot j}}{P_{\cdot j} + \phi/\psi}\overline{X}_{\cdot j}\,.$$

De Vylder (1981*a*) desenvolveu o mesmo princípio de minimização do traço da covariância do estimador, para os outros parâmetros estruturais, assumindo a multi-normalidade condicionada por θ_j para o vector $\underline{X}_j$, sob hipóteses bastantes restritivas. Norberg (1982) simplifica o método e estende-o de forma a ser aplicável em situações mais gerais, sendo aplicável no caso de distribuições condicionadas, binomial, Poisson, Poisson Composta e Multinormal.

Exemplo 14 *Este exemplo foi construido com base no exemplo apresentado em Hachemeister (1975). Como foi referido na Secção 6.3.1, Hachemeister estabeleceu a seguinte particularização ao seu modelo, para determinado risco j ($j = 1, 2, ...N$):*

$$\mu(\theta_j) = \mathrm{E}[X_{ij}|\theta_j] = b_0(\theta_j) + i\, b_1(\theta_j), \quad i = 1, 2, ..., n, n+1,$$

isto é, fazendo a particularização com $r = 2$ e $\underline{y}_{ij} = \begin{bmatrix} 1 & i \end{bmatrix}$, em que N corresponde ao número de estados (EUA) e i períodos trimestrais. Na Tabela 7 são apresentados valores sobre P_{ij} e X_{ij} , respectivamente a frequência e gravidade média de sinistros, referentes a 12 períodos e para 5 estados, assim como os valores globais do conjunto dos 5 estados (países). Na última linha do quadro são apresentados os valores globais $P_{\cdot j} = \sum_j P_{ij}$, e os valores médios em cada estado, $\overline{X}_{\cdot j} = (1/P_{\cdot j})\sum_i P_{ij}X_{ij}$. Na última coluna, são apresentados os valores globais dos cinco estados, da frequência de sinistros e as médias $\overline{X}_{i\cdot} = (1/P_{i\cdot})\sum_j P_{ij}X_{ij}$, para cada i ($P_{i\cdot} = \sum_j P_{ij}$).

O Prémio Empírico de Credibilidade para o 13º período é dado pela expressão:

$$\tilde{m}_{j,N} = \underline{a}'\left[\mathbf{Z}_{j,N}\underline{\hat{b}}_j + (\mathbf{I} - \mathbf{Z}_{j,N})\,\underline{\hat{\beta}}\right]$$

com

$$\begin{aligned}\mathbf{Z}_{j,N} &= P_{\cdot j}\left(P_{\cdot j}\mathbf{I}+\hat{\mathbf{K}}_j\right)\\ \hat{\mathbf{K}}_j &= \hat{\phi}P_{\cdot j}\left(\hat{\mathbf{\Lambda}}\mathbf{Y}_j'\mathbf{P}_j\mathbf{Y}_j\right)^{-1}\\ \underline{a}' &= [1 \quad 13],\end{aligned}$$

sendo $\underline{\hat{b}}_j$, $\underline{\hat{\beta}}$, $\hat{\phi}$ *e* $\hat{\mathbf{\Lambda}}$ *dados respectivamente por (87), (107) - (109).*
Tendo em conta que a matriz $\mathbf{Y}_j$ $(j=1,2,\cdots,12)$ *tem neste caso a forma*

$$\mathbf{Y}_j' = \begin{bmatrix} 1 & 1 & \cdots & 1 & \cdots & 1\\ 1 & 2 & \cdots & i & \cdots & 12\end{bmatrix}$$

Os estimadores obtidos para $\underline{\hat{b}}_j$ $(j=1,2,\cdots,5)$ *e* $\underline{\hat{\beta}}$*, foram os seguintes :*

$$\underline{\hat{b}}_1 = \begin{bmatrix} 1911,351624\\ -2,45394173\end{bmatrix},\quad \underline{\hat{b}}_2 = \begin{bmatrix} 1841,134053\\ 20,56195211\end{bmatrix},\quad \underline{\hat{b}}_3 = \begin{bmatrix} 1477,37852\\ 38,64578889\end{bmatrix}$$

$$\underline{\hat{b}}_4 = \begin{bmatrix} 1176,704065\\ 27,80701828\end{bmatrix},\quad \underline{\hat{b}}_5 = \begin{bmatrix} 1543,842627\\ 6,767537679\end{bmatrix},\quad \underline{\hat{\beta}} = \begin{bmatrix} 1645,506811\\ 16,90056574\end{bmatrix}$$

Na Figura 1 comparam-se as regressões particulares para cada estado e a regressão global do conjunto dos estados, propostas em (76) e (78), isto é, neste caso particular $\hat{\mathrm{E}}[X_{13,j}] = \hat{b}_{0j} + 13\hat{b}_{1j}$, $j = 1,2,...,5$ *e* $\hat{\mathrm{E}}[X_{13,\cdot}] = \hat{\beta}_0 + 13\hat{\beta}_1$. *Tem-se assim*

$$\begin{aligned}\hat{\mathrm{E}}[X_{13,1}] &= 1879,\quad \hat{\mathrm{E}}[X_{13,2}] = 2108,\quad \hat{\mathrm{E}}[X_{13,3}] = 1980\\ \hat{\mathrm{E}}[X_{13,4}] &= 1538,\quad \hat{\mathrm{E}}[X_{13,5}] = 1632,\quad \hat{\mathrm{E}}[X_{13\cdot}] = 1865\ .\end{aligned}$$

A linha mais carregada corresponde à do conjunto. Para o estimador $\hat{\phi}$ *tem-se, usando (108),*

$$\hat{\phi} = (1/50)\sum_{j=1}^{5} {S'}_j^2 = 74968355,08$$

com

j	1	2	3	4	5
${S'}_j^2$	1225150603	480099911,6	1287400485	487180106,7	268586647,8

Para a obtenção do estimador $\hat{\mathbf{\Lambda}}$*, dado por (109), obtiveram-se as matrizes* $\mathbf{G}$, $\mathbf{Y}'\mathbf{P}\mathbf{Y}$, $\mathbf{\Pi}$ *e* $\mathbf{H}$*:*

$$\mathbf{G} = \begin{bmatrix} 44736,37179 & -1768,987829\\ -1701,044438 & 231,2664546\end{bmatrix},\quad \mathbf{Y}'\mathbf{P}\mathbf{Y} = \begin{bmatrix} 150038 & 972451\\ 972451 & 8131303\end{bmatrix}$$

$$\mathbf{\Pi} = \begin{bmatrix} 0,772361572 & 0,01378235764\\ -0,000445225250 & 0,766512291\end{bmatrix}\quad \mathbf{H} = \begin{bmatrix} 46428,36583 & -915,6908019\\ -805,5027499 & 87,22429028\end{bmatrix}$$

Então $\hat{\mathbf{\Lambda}} = (1/2)(\mathbf{H}+\mathbf{H}')$*, com* $\mathbf{H} = \mathbf{\Pi}^{-1}\left[\mathbf{G} - 4\left(\mathbf{Y}'\mathbf{P}\mathbf{Y}\right)\hat{\phi}\right]$*, vem*

$$\hat{\mathbf{\Lambda}} = \begin{bmatrix} 46428,36583 & -860,596775\\ -860,5967759 & 87,22429028\end{bmatrix}$$

Obtiveram-se as seguintes matrizes $\hat{\mathbf{K}}_j$ *e* $\mathbf{Z}_{j,N}$ $(j=1,2,...,5)$*:*

$$\hat{\mathbf{K}}_1 = \begin{bmatrix} -1595,55337 & -475829,581\\ 550,3306188 & 76321,13267\end{bmatrix},\quad \hat{\mathbf{K}}_2 = \begin{bmatrix} -1547,129554 & -482691,5119\\ 534,7877323 & 7622651183\end{bmatrix},$$

$$\hat{\mathbf{K}}_3 = \begin{bmatrix} -1714,517182 & -467024,7366 \\ 585,7930014 & 77222,9304 \end{bmatrix} \quad \hat{\mathbf{K}}_4 = \begin{bmatrix} -1580,435472 & -454131,0636 \\ 561,048036 & 74715,22937 \end{bmatrix},$$

$$\hat{\mathbf{K}}_5 = \begin{bmatrix} -1501,359052 & -472632,8164 \\ 529,8760956 & 74987,82267 \end{bmatrix},$$

$$\hat{\mathbf{Z}}_{1,N} = \begin{bmatrix} 0,9723168441 & 4,341013536 \\ -0,005020689677 & 0,2614795985 \end{bmatrix}, \quad \hat{\mathbf{Z}}_{2,N} = \begin{bmatrix} 0,9769256117 & 4,28523466 \\ -0,004747734049 & 0,2864674299 \end{bmatrix},$$

$$\hat{\mathbf{Z}}_{3,N} = \begin{bmatrix} 0,9789614429 & 4,09834878 \\ -0,005140592874 & 0,2862503999 \end{bmatrix}, \quad \hat{\mathbf{Z}}_{4,N} = \begin{bmatrix} 0,8479868578 & 4,638650305 \\ -0,005730736899 & 0,06867679211 \end{bmatrix},$$

$$\hat{\mathbf{Z}}_{5,N} = \begin{bmatrix} 0,9859973772 & 3,938735385 \\ -0,004415778285 & 0,348566720 \end{bmatrix}.$$

O Estimador Empírico de Credibilidade de $\underline{b}(\theta_j)$*, dado por* $\underline{\tilde{b}}_{j,N} = \mathbf{Z}_{j,N}\underline{\hat{b}}_j + (\mathbf{I} - \mathbf{Z}_{j,N})\,\underline{\hat{\beta}}$*, vem para cada* j *:*

$$\underline{\tilde{b}}_{1,N} = \begin{bmatrix} 1819,974 \\ 10,505 \end{bmatrix}, \quad \underline{\tilde{b}}_{2,N} = \begin{bmatrix} 1852,310 \\ 17,021 \end{bmatrix}, \quad \underline{\tilde{b}}_{3,N} = \begin{bmatrix} 1570,035 \\ 23,989 \end{bmatrix},$$

$$\underline{\tilde{b}}_{4,N} = \begin{bmatrix} 1289,559 \\ 20,336 \end{bmatrix}, \quad \underline{\tilde{b}}_{5,N} = \begin{bmatrix} 1505,355 \\ 13,817 \end{bmatrix}.$$

As estimativas $\tilde{m}_{j,N}$ *vêm na Tabela 6: Nas Figuras 2-4 são comparadas para cada estado, as estimativas*

j	1	2	3	4	5
$\tilde{m}_{j,N}$	1957	2074	1882	1563	1685

Tabela 6: Estimativas de credibilidade para o 13o período

$\hat{m}_{j,N}$ *com as representadas na Figura 1. A linha C corresponde à estimativa de credibilidade, a linha P corresponde à do conjunto do país, e a linha E corresponde à linha particular de cada estado. Examinando os gráficos, pode ver-se que em geral a estimativa de credibilidade para o 13º período se situa entre a linha do conjunto e a particular do estado. Apenas para o Estado 1 isso não se verifica. No entanto, o comportamento deste estado é estranho comparativamente com os restantes. Repare-se que o declive da recta de credibilidade é influenciada pela do conjunto, o que se verifica de uma forma geral. Note-se ainda que para os estados 2, 4 e 5, a estimativa de credibilidade é próxima da particular do estado, enquanto que no estado esta vem próxima da do conjunto.* ◊

j	1		2		3		4		5		País	
i	P_{i1}	X_{i1}	P_{i2}	X_{i2}	P_{i3}	X_{i3}	P_{i4}	X_{i4}	P_{i5}	X_{i5}	$P_{i\cdot}$	$X_{i\cdot}$
1	2358	1738	2757	1789	2867	1702	814	1223	3482	1456	12278	1627
2	2775	1642	2961	1845	3392	1630	792	1146	3806	1499	13726	1615
3	2611	1794	2589	2064	3322	1363	696	1010	3655	1609	12873	1642
4	2572	2051	2575	1957	3010	1657	682	1257	3681	1741	12520	1803
5	2375	2079	2757	1776	2495	1599	630	1426	3231	1482	11488	1698
6	2478	2234	2723	2127	2692	2022	656	1532	3492	1572	12041	1932
7	2836	2032	3338	1939	3192	1546	704	1953	3930	1606	14000	1775
8	2400	2039	2575	1959	3045	1491	662	1123	3236	1735	11918	1748
9	2209	2115	2595	1934	2240	1729	574	1343	3195	1607	10813	1801
10	2349	1630	2971	2279	2507	2008	768	1243	3620	1573	12215	1824
11	2354	1721	2811	2079	2770	2194	642	1762	3890	1613	12467	1875
12	2940	1719	3163	1967	2802	1880	684	1306	4110	1568	13699	1743
	30257	1895	33815	1977	34334	1721	8304	1353	43328	1588	150038	1755
	$P_{\cdot 1}$	$X_{\cdot 1}$	$P_{\cdot 2}$	$X_{\cdot 2}$	$P_{\cdot 3}$	$X_{\cdot 3}$	$P_{\cdot 4}$	$X_{\cdot 4}$	$P_{\cdot 5}$	$X_{\cdot 6}$	$P_{\cdot\cdot}$	$X_{\cdot\cdot}$

Tabela 7: Frequência e gravidade dos sinistros.

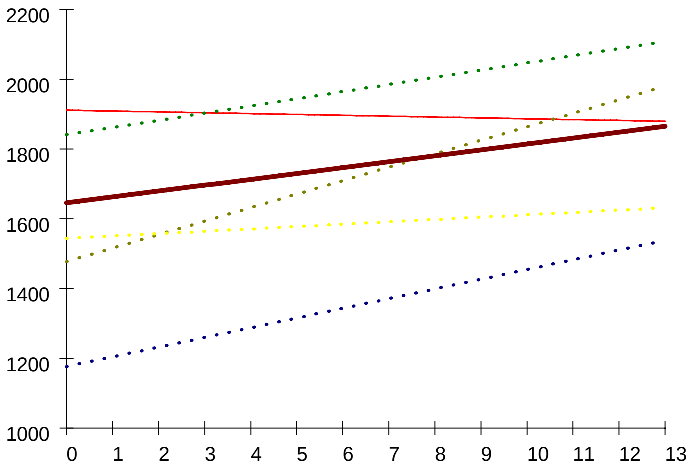
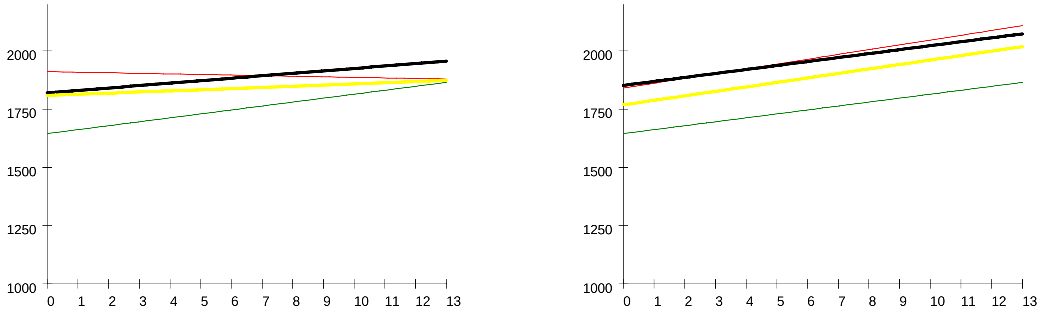
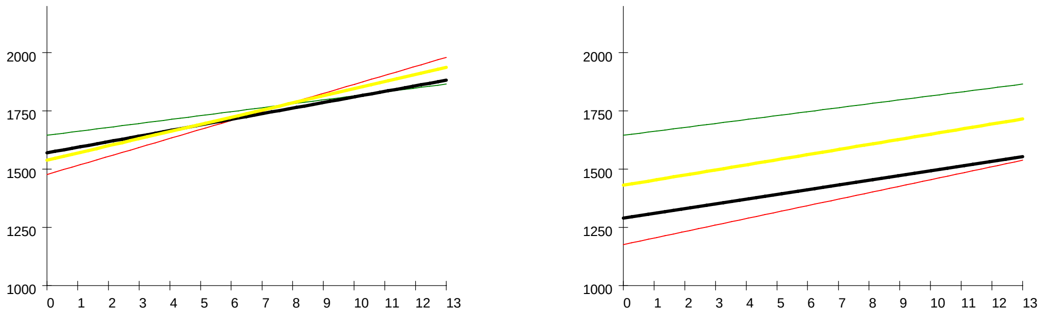


Figura 1: Rectas de regressão particulares e global



$j = 1$ $\qquad$ $j = 2$

Figura 2: Rectas de regressão por estado, $j = 1, 2$



$j = 3$ $\qquad$ $j = 4$

Figura 3: Rectas de regressão por estado, $j = 3, 4$

6.4 Modelo de regressão com factor de credibilidade escalar

De Vylder (1981b) propõe que no modelo de regressão de Hachemeister a matriz de credibilidade $\mathbf{Z}_j$ seja substituída por um escalar como factor de credibilidade . Em prol desta tese argumenta que as matrizes de credibilidade podem conter elementos negativos, levando a resultados que dificilmente podem ser interpretados nalguns casos particulares. Claro que a substituição da matriz de credibilidade traz vantagens do ponto de vista de calculatória. apoia-se ainda no facto de que o modelo confrontado em exemplos práticos ter dado bons resultados.

Propõe encontrar o peso de credibilidade Z_j escalar óptimo, obtido através da minimização da função de perca esperada

$$R = \mathrm{E}\left[\left(\tilde{\tilde{\underline{b}}}_j - \underline{b}(\theta_j)\right)' \mathbf{Q}_j \left(\tilde{\tilde{\underline{b}}}_j - \underline{b}(\theta_j)\right)\right] \tag{128}$$

em que $\tilde{\tilde{\underline{b}}}_j$ é o Estimador de Credibilidade do vector $\underline{b}(\theta_j)$ tal como foi definido em H1, mas substituindo a matriz de credibilidade $\mathbf{Z}_j$ do modelo de Hachemeister por um escalar, e $\mathbf{Q}_j$ uma matriz $(r \times r)$ definida positiva. Além disso, para ser um Factor de Credibilidade z_j (escalar) tem de ser um número compreendido entre 0 e 1.

O seguinte teorema [De Vylder (1981b)] dá solução ao problema proposto.

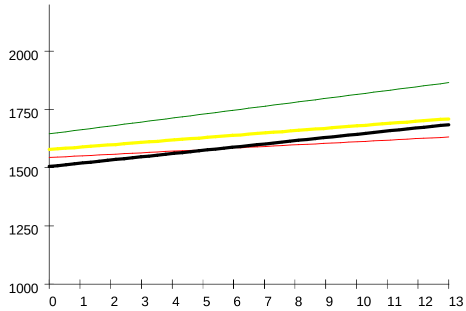


Figura 4: Rectas de regressão por estado, $j = 5$

Teorema 10 *O factor de credibilidade escalar z_j, obtido pela minimização de (128) vem,*

$$z_j = \frac{\mathrm{tr}\left[\mathbf{\Lambda}\mathbf{Q}_j\right]}{\mathrm{tr}\left[\left(\mathbf{\Lambda} + \phi\left(\mathbf{Y}_j'\mathbf{P}_j\mathbf{Y}\right)^{-1}\right)\mathbf{Q}_j\right]} \tag{129}$$

e é um número entre 0 e 1. O símbolo "tr" significa traço da matriz.

Dem.

Seja z_j escalar e $\tilde{\tilde{\underline{b}}}_j = z_j\hat{\underline{b}}_j + (1 - z_j)\underline{\beta}$, o estimador de credibilidade substituindo a matriz de credibilidade $\mathbf{Z}_j$ pelo escalar z_j, e seja

$$\tilde{\tilde{\underline{b}}}_j - \underline{b}(\theta_j) = z_j\hat{\underline{b}}_j - z_j\underline{\beta} + \underline{\beta} - \underline{b}(\theta_j),$$

como por (105) e (82) $\mathrm{E}[\hat{\underline{b}}_j] = \mathrm{E}[\underline{b}(\theta_j)] = \underline{\beta}$, então

$$\tilde{\tilde{\underline{b}}}_j - \underline{b}(\theta_j) = z_j(\hat{\underline{b}}_j - \mathrm{E}[\hat{\underline{b}}_j]) - (\underline{b}(\theta_j) - \mathrm{E}[\underline{b}(\theta_j)]) = z_j\hat{\underline{b}}_j^0 - \underline{b}^0(\theta_j) \tag{130}$$

fazendo

$$\hat{\underline{b}}_j^0 = \underline{b}(\theta_j) - \mathrm{E}[\hat{\underline{b}}_j] \tag{131}$$

e $\underline{b}^0(\theta_j) = \underline{b}(\theta_j) - \mathrm{E}[\underline{b}(\theta_j)]$. Como R é um escalar, então $R = \mathrm{tr}[R]$ e

$$R = \mathrm{tr}\left[\mathrm{E}\left[\left(\tilde{\tilde{\underline{b}}}_j - \underline{b}(\theta_j)\right)' \mathbf{Q}_j \left(\tilde{\tilde{\underline{b}}}_j - \underline{b}(\theta_j)\right)\right]\right],$$

em virtude da propriedade do traço de um produto de matrizes matriz, e usando (130) pode escrever-se (note-se que as matrizes continuam multiplicáveis),

$$R = \mathrm{tr}\left[\mathrm{E}\left[\left(z_j\hat{\underline{b}}_j^0 - \underline{b}^0(\theta_j)\right)\left(z_j\hat{\underline{b}}_j^0 - \underline{b}^0(\theta_j)\right)' \mathbf{Q}_j\right]\right].$$

Desenvolvendo a expressão obtém-se

$$\begin{aligned} R &= \mathrm{tr}\left[\mathrm{E}\left[{z_j}^2\hat{\underline{b}}_j^0\hat{\underline{b}}_j^{0\prime} - z_j\hat{\underline{b}}_j^0\underline{b}^0(\theta_j)' - z_j\underline{b}^0(\theta_j)\underline{b}_j^{0\prime} + \underline{b}^0(\theta_j)\underline{b}^0(\theta_j)'\right]\mathbf{Q}_j\right] \\ &= \mathrm{tr}\left[{z_j}^2\mathrm{E}\left[\hat{\underline{b}}_j^0\hat{\underline{b}}_j^{0\prime}\right]\mathbf{Q}_j - z_j\mathrm{E}\left[\hat{\underline{b}}_j^0\underline{b}^0(\theta_j)'\right]\mathbf{Q}_j - z_j\mathrm{E}\left[\underline{b}^0(\theta_j)\hat{\underline{b}}_j^{0\prime}\right]\mathbf{Q}_j + \mathrm{E}\left[\underline{b}^0(\theta_j)\underline{b}^0(\theta_j)'\right]\mathbf{Q}_j\right]. \end{aligned}$$

Tendo em atenção (130) e utilizando (119), tem-se

$$\mathrm{E}\left[\hat{\underline{b}}_j^0\hat{\underline{b}}_j^{0\prime}\right] = \mathrm{Cov}\left[\hat{\underline{b}}_j\right] = \phi(\mathbf{Y}_j'\mathbf{P}_j\mathbf{Y})^{-1} + \boldsymbol{\lambda}.$$

Por outro lado,

$$\begin{aligned}
\mathrm{E}[\underline{b}^0(\theta_j)\underline{b}^0(\theta_j)'] &= \mathrm{E}\left[\left(\underline{b}(\theta_j)-\underline{\beta}\right)\left(\hat{\underline{b}}_j-\underline{\beta}\right)'\right]\hat{\underline{b}}_j^{0\prime} \\
&= \mathrm{E}[\underline{b}(\theta_j)\,\hat{\underline{b}}_j{}'] - \mathrm{E}[\underline{b}(\theta_j)]\underline{\beta}' - \underline{\beta}\mathrm{E}[\hat{\underline{b}}_j'] + \underline{\beta\beta}'\,.
\end{aligned}$$

Substituindo $\hat{\underline{b}}_j$ pela respectiva expressão em (87) e $\mathrm{E}\left[\hat{\underline{b}}_j'\right] = \underline{\beta}'$ obtém-se

$$\begin{aligned}
\mathrm{E}[\underline{b}^0(\theta_j)\underline{b}^0(\theta_j)'] &= \mathrm{E}[\underline{b}(\theta_j)\,\underline{X}_j{}']\mathbf{P}_j\mathbf{Y}_j(\mathbf{Y}'\mathbf{P}_j\mathbf{Y})^{-1} - \underline{\beta\beta}' - \underline{\beta\beta}' + \underline{\beta\beta}' \\
&= \left(\mathrm{Cov}[\underline{b}(\theta_j),\underline{X}_j']\,\mathrm{E}[\underline{b}(\theta_j)]\,\mathrm{E}[\underline{X}_j']\right)\mathbf{PY}(\mathbf{Y}_j'\mathbf{P}_j\mathbf{Y})^{-1} - \underline{\beta\beta}' \\
&= (\mathbf{\Lambda}\mathbf{Y}_j' + \underline{\beta\beta}'\mathbf{Y}_j')\mathbf{P}_j\mathbf{Y}_j(\mathbf{Y}_j'\mathbf{P}_j\mathbf{Y}_j)^{-1} - \underline{\beta\beta}' = \mathbf{\Lambda},
\end{aligned}$$

na última expressão usando (106), (80) e (82). Além disso tem-se igualmente $\mathrm{E}\left[\hat{\underline{b}}_j^0\,\underline{b}^0(\theta_j)'\right]$ e através de (131) e (83) $\mathrm{E}\left[\underline{b}^0(\theta_j)\,\underline{b}^0(\theta_j)'\right] = \mathrm{Cov}\left[\underline{b}(\theta_j)\right] = \mathbf{\Lambda}$. Encontra-se assim uma expressão mais tratável para R:

$$R = {z_j}^2\,\mathrm{tr}\left[\left(\phi\left(\mathbf{Y}_j'\mathbf{P}_j\mathbf{Y}_j\right)^{-1} + \mathbf{\Lambda}\right)\mathbf{Q}_j\right] - 2\,z_j\,\mathrm{tr}\left[\mathbf{\Lambda Q}_j\right] + \mathrm{tr}\left[\mathbf{\Lambda Q}_j\right].$$

Derivando R em ordem a z_j e igualando a zero obtém-se

$$\frac{dR}{dz_j} = 2\,z_j\,\mathrm{tr}\left[\left(\phi\left(\mathbf{Y}_j'\mathbf{P}_j\mathbf{Y}-j\right)^{-1} + \mathbf{\Lambda}\right)\mathbf{Q}_j\right] - 2\,\mathrm{tr}\left[\mathbf{\Lambda Q}_j\right] = 0,$$

donde

$${z_j}^* = \frac{\mathrm{tr}(\mathbf{\Lambda Q}_j)}{\mathrm{tr}\left[\left(\phi(\mathbf{Y}_j'\mathbf{P}_j\mathbf{Y}-j)^{-1} + \mathbf{\Lambda}\right)\mathbf{Q}_j\right]}.$$

${Z_j}^*$ é um minimizante pois a segunda derivada

$$\frac{d^2R}{d{z_j}^2} = 2\,\mathrm{tr}\left[\left(\phi\left(\mathbf{Y}_j'\mathbf{P}_j\mathbf{Y}-j\right)^{-1} + \mathbf{\Lambda}\right)\mathbf{Q}_j\right] > 0,$$

já que o produto de duas matrizes definidas positivas tem traço positivo. Além disso ${z_j}^*$ é um número entre 0 e 1, primeiro porque o numerador é positivo e segundo porque ele é inferior ao denominador pois o traço da soma é igual à soma dos traços. □

No seu modelo, De Vylder propõe ainda fazer

$$\mathbf{Q}_j = \mathbf{Y}'\left(\mathrm{Cov}\left[\underline{X}_j,\right]\right)^{-1}\mathbf{Y}_j = \mathbf{Y}'\left(\phi{\mathbf{P}_j}^{-1} + \mathbf{Y}'\mathbf{\Lambda}\mathbf{Y}_j'\right)^{-1}\mathbf{Y}_j,$$

aplicando (91).

Para exemplificar , vai-se retomar o exemplo da secção anterior.

Exemplo 15 *Pretende-se calcular o estimador de credibilidade, substituindo a matriz de credibilidade pelo factor z_j escalar , calculado através de (129) , sendo a matriz $\mathbf{Q}$ calculada pela expressão acima. Na expressão vai-se substituir os parâmetros ϕ e $\mathbf{\Lambda}$ pelos respectivos estimadores obtidos no referido exemplo. As matrizes $\hat{\mathbf{Q}}_j$ obtidas são a seguir apresentadas.*

$$\hat{\mathbf{Q}}_1 = \begin{bmatrix} 0,00002432382379 & 0,0001824298553 \\ 0,0001824298553 & 0,004797724836 \end{bmatrix}, \quad \hat{\mathbf{Q}}_2 = \begin{bmatrix} 0,00002451629606 & 0,0001874582326 \\ 0,0001874582326 & 0,005133815123 \end{bmatrix},$$

$$\hat{\mathbf{Q}}_3 = \begin{bmatrix} 0,00002446778701 & 0,0001824756119 \\ 0,0001824756119 & 0,005082166008 \end{bmatrix}, \quad \hat{\mathbf{Q}}_4 = \begin{bmatrix} 0,00002086191063 & 0,0001401324801 \\ 0,0001401324801 & 0,002169972975 \end{bmatrix},$$

j	1	2	3	4	5
numerador	1,234	1,263	1,265	0,917	1,335
denominador	2	2	2	2	2
$Z_{j,N}$	0,617	0,632	0,633	0,458	0,667
$\tilde{\tilde{\underline{b}}}_{j,N}$	1809 4,961	1769 19,213	1539 30,657	1431 21,899	1578 10,139
$\tilde{\tilde{m}}_{j,N}$	1874	2019	1938	1715	1709

Tabela 8: Estimativas de credibilidade para o 13o período com factor escalar

$$\hat{\mathbf{Q}}_5 = \begin{bmatrix} 0,0000248417541 & 0,0001944751301 \\ 0,0001944751301 & 0,005914996713 \end{bmatrix}.$$

No quadro seguinte são apresentados os valores de z_j estimados, dado pela fórmula (129), bem como os respectivos valores do numerador e do denominador, e ainda as estimativas de $\underline{b}(\theta_j)$ e do prémio puro para o 13º período.

Nas Figuras 2-4 apresentam-se as regressões encontradas para este caso . A recta VC é a correspondente a este caso. Como se pode verificar as estimativas parecem razoáveis para os Estados 2, 3 e 5. Para o Estado 1 a estimativa sofre uma inflexão vindo a situar-se entre a E e P. Para o Estado 4, a recta VC vem bastante deslocada da recta C, a que não será alheio o facto de este estado ter pouco peso no conjunto.

Referências

Andrade e Silva, J. M. S. (1991), Estruturas Tarifárias No Ramos Reais Da Indústria Seguradora, PhD thesis, ISEG, Universidade Técnica de Lisboa.

Bühlmann, H. (1967), 'Experience rating and credibility', *Astin Bulletin* **4**, 199–207.

Bühlmann, H. (1970), *Mathematical Methods in Risk Theory*, Springer Verlag, Berlin.

Bühlmann, H. & Straub, E. (1970), 'Glaubwurdigkeit fur shadensatze', *Mitteilungen Vereinigung Shweizerischer Versicherungsmathematiker* **70**, 111–133. translated into English by C.E. Brooks, Zurique, "Credibility for loss ratios".

Centeno, L. (2001), *Teoria do Risco*, CEMAPRE-ISEG, Rua do Quelhas 6, 1200-781 Lisboa.

Cordeiro, I. M. F. (1991), Constituição de reservas para sinistros ocorridos e ainda não participados ou não encerrados. análise de alguns métodos de previsão, Master's thesis, ISEG, Universidade Técnica de Lisboa.

Dannenburg, D. R., Kaas, R. & Goovaerts, M. J. (1996), *Practical Actuarial Credibility Models*, IAE - University of Amsterdam.

De Vylder, F. (1978), 'Parameter estimation in credibility theory', *AB* **10**, 99–112.

De Vylder, F. (1981*a*), 'Practical credibility theory with emphasis on optimal parameter estimation', *Astin Bulletin* **12**, 115–131.

De Vylder, F. (1981*b*), 'Regression model with scalar credibility weigths', *Mitteilungen Vereinigung Shweizerischer Versicherungsmathematiker* **1**, 27–39.

Egídio dos Reis, A. D. (1987), Teoria da credibilidade. Uma síntese, Master's thesis, ISEG, Universidade Técnica de Lisboa.

Gerber, H. U. (1995), 'A teacher's remark on exact credibility', *Astin Bulletin* **25**(2), 189–192.

Goovaerts, M. J., de Vylder, F. & Hazendonck, J. (1984), *Insurance Premiums. Theory and Applications*, North-Holland, Amsterdam.

Hachemeister, C. A. (1975), Credibility for regression models with application to trend, *in* P. M. Kahn, ed., 'Credibility Theory and Applications', Academic Press.

Jewell, W. S. (1974), 'Credible means are exact Bayesian for exponential families', *Astin Bulletin* **8**, 77–90.

Jewell, W. S. (1975*a*), Model variatons in credibility theory, *in* P. M. Kahn, ed., 'Credibility Theory and Applications', Academic Press.

Jewell, W. S. (1975*b*), 'Regularity conditions for exact credibility', *Astin Bulletin* **8**, 336–341.

Johnston, J. & DiNardo, J. (1997), *Econometric Methods*, McGraw-Hill International.

Johnston, J. & DiNardo, J. (2001), *Métodos Econométricos*, McGraw-Hill, Lisboa.

Lemaire, J. (1995), *Bonus-Malus Systems in Automobile Insurance*, Hueber International Series on Risk, Insurance, and Economic Security, Kluwer Academic, Boston.

Longley-Cook, L. H. (1962), 'An introduction to credibility theory', *Proceedings of the Casualty Actuarial Society* **49**.

Murteira, B. J. F. (1988), *Estatística: Inferência e Decisão*, INCM-Imprensa Nacional Casa da Moeda, Lisboa.

Murteira, B. J. F. (1990*a*), *Probabilidades e Estatística*, Vol. II, 2ª edn, McGraw-Hill de Portugal.

Murteira, B. J. F. (1990*b*), *Probabilidades e Estatística*, Vol. I, 2a edn, McGraw-Hill de Portugal.

Norberg, R. (1979), 'The credibility approach to experience rating', *Scandinavian Actuarial Journal* pp. 181–221.

Norberg, R. (1982), 'On optimal parameter estimation in credibility', *Insurance: Mathematics and Economics* pp. 73–89.

Sundt, B. (1979), 'A hierarchical credibility regression model', *Scandinavian Actuarial Journal* pp. 107–114.

Sundt, B. (1981), 'Recursive credibility estimation', *Scandinavian Actuarial Journal* pp. 107–114.

Sundt, B. (1982), Parameter estimation in some credibility models, Technical report, University of Oslo.