TEORI KREDIBILITAS UNTUK AKTUARIS
DENGAN PUSTAKA ACTUAR DALAM R

IKHSAN MAULIDI

Konsep Dasar Statistika

Review Beberapa Konsep Dasar Peluang

Misalkan diberikan fungsi sebaran bersyarat dari dua peubah acak X dan Y . Fungsi peluang bersama dinotasikan sebagai

$$f_{X,Y}(x, y) = P(X = x, Y = y).$$

- Jika X dan Y peubah acak yang kontinu, maka

$$f_{X,Y}(x, y) = \frac{\partial^2}{\partial x \partial y} F_{X,Y}(x, y).$$

- Jika X adalah peubah acak diskret, maka fungsi sebaran marginal dari X diberikan oleh

$$p_X(x) = P(X = x) = \sum_{\forall y} P_{X,Y}(x, y).$$

- Jika X adalah peubah acak kontinu, maka fungsi sebaran marginal dari X diberikan oleh

$$f_X(x) = \int_{\forall y} f_{X,Y}(x, y) dy.$$

- Fungsi peluang bersyarat diberikan oleh

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

- Fungsi peluang bersama dan fungsi peluang bersyarat memiliki hubungan sebagai berikut

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

- Jika X dan Y adalah peubah acak yang saling bebas, maka berlaku

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

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

$$f_{Y|X}(y|x) = f_Y(y).$$

Contoh:

Misalkan $X|(\Theta = \theta) \sim \text{Poisson}(\theta)$ dan $\Theta \sim \text{Gamma}(\alpha, \beta)$. Tentukan sebaran untuk X .

Jawab:

$$\begin{aligned}
 f_X(x) &= \int_{\forall \theta} f_{X|\theta}(x, \theta) d\theta \\
 &= \int_{\forall \theta} f_{X|\theta}(x|\theta) f_{\theta}(\theta) d\theta = \int_0^{\infty} \left(\frac{\theta^x e^{-\theta}}{x!} \right) \left(\frac{\theta^{\alpha-1} e^{-\frac{\theta}{\beta}}}{\Gamma(\alpha) \beta^{\alpha}} \right) d\theta \\
 &= \frac{\Gamma(x + \alpha) \left(\frac{1}{1 + \beta} \right)^{\alpha} \left(\frac{\beta}{1 + \beta} \right)^x}{x! \Gamma(\alpha)} \int_0^{\infty} \frac{\theta^{x+\alpha-1} e^{-\frac{\theta}{\beta}}}{\Gamma(x + \alpha) \left(\frac{\beta}{1 + \beta} \right)^{x+\alpha}} d\theta \\
 &= \frac{\Gamma(x + \alpha)}{x! \Gamma(\alpha)} \left(\frac{1}{1 + \beta} \right)^{\alpha} \left(\frac{\beta}{1 + \beta} \right)^x.
 \end{aligned}$$

Dengan demikian, X mengikuti sebaran Binomial negatif dengan parameter α dan β . Sebagai fakta tambahan, misalkan $\alpha = 1$, maka $X \sim \text{Geo}(\beta)$.

Fungsi Sebaran Campuran (*Mixture Distribution Function*)

Asumsikan beberapa peubah acak Θ yang berpengaruh terhadap peubah acak kerugian X , memiliki fungsi sebaran peluang $\pi(\theta)$ dan fungsi sebaran bersyarat atau fungsi sebaran dari X dapat dievaluasi sebagai berikut.

$$\begin{aligned}
 F_X(x) &= P(X \leq x) = \int P(X \leq x | \Theta = \theta) \pi(\theta) d\theta = \int F_{X|\theta}(x|\theta) \pi(\theta) d\theta, \\
 f_X(x) &= \frac{\partial}{\partial x} F_X(x) = \frac{\partial}{\partial x} \int F_{X|\theta}(x|\theta) \pi(\theta) d\theta = \int f_{X|\theta}(x|\theta) \pi(\theta) d\theta.
 \end{aligned}$$

Misalkan didefinisikan fungsi sebaran campuran $F_X(x) = \alpha_1 F_1(x) + \dots + \alpha_n F_n(x)$, jika diberikan sekumpulan n fungsi kepekatatan peluang atau fungsi sebaran $F_1(x), F_2(x), \dots, F_n(x)$ dan berkorespondensi dengan bobot $\alpha_1, \alpha_2, \dots, \alpha_n$ sedemikian sehingga $\alpha_i \geq 0$ dan $\sum_{i=1}^n \alpha_i = 1$ dengan $i \in \mathbb{Z}$, $0 \leq 1 \leq n$.

Jika kita asumsikan $F_i(x)$ masing masing fungsi sebaran untuk X_i , $i = 1, \dots, n$ dan $F_X(x)$ fungsi sebaran untuk X , maka

$$E[X^k] = \alpha_1 E[X_1^k] + \dots + \alpha_n E[X_n^k].$$

Nilai Harapan Bersyarat (*Conditional Expectation*)

$$E[X|Y = y] = \int_{\forall x} x f_{X|Y}(x|y) dx.$$

$$E[X|Y] = g(Y)$$

$$Var(X|Y) = E[X^2|Y] - (E[X|Y])^2 = h(Y).$$

Teorema

$$E[E[X|Y]] = E[X]$$

$$Var(X) = Var(E[X|Y]) + E(Var(X|Y))$$

(Bukti lihat Ross 2010).

Contoh:

Misalkan $X|(\Theta = \theta > 0)$ memiliki fungsi sebaran eksponensial bersyarat dengan fungsi $F_{X|\theta}(x|\theta) = 1 - e^{-\theta x}$, $x > 0$. $\Theta \sim \text{Gamma}(\alpha, \frac{1}{\lambda})$ dengan fungsi $\pi(\theta) = \frac{\lambda^\alpha \theta^{\alpha-1} e^{-\lambda\theta}}{\Gamma(\alpha)}$, $\theta > 0$. Tentukan nilai harapan dan ragam X .

Jawab:

$$E[X] = E[E[X|\Theta]] = E\left[\frac{1}{\Theta}\right] = \int_0^\infty \frac{1}{\theta} \frac{\lambda^\alpha \theta^{\alpha-1} e^{-\lambda\theta}}{\Gamma(\alpha)} d\theta = \frac{\lambda \Gamma(\alpha - 1)}{\Gamma(\alpha)} = \frac{\lambda}{\alpha - 1}.$$

$$Var(X) = Var(E[X|\Theta]) + E(Var(X|\Theta))$$

$$= Var\left(\frac{1}{\Theta}\right) + E\left[\frac{1}{\Theta^2}\right] = \frac{\alpha \lambda^2}{(\alpha - 1)^2 (\alpha - 2)}.$$

Mean Square Error

Mean square error dari suatu penduga didefinisikan sebagai berikut

$$MSE_{\hat{\theta}}(\theta) = E \left[(\hat{\theta} - \theta)^2 \right] = Var(\hat{\theta}) + (E[\hat{\theta}] - \theta)^2$$

Pendugaan Bayes (*Bayes Estimation*)

- Sebaran prior merupakan sebaran peluang berdasarkan seluruh kemungkinan nilai parameter θ . Fungsi sebaran prior dinotasikan dengan $\pi(\theta)$.
- Sebaran Posterior merupakan sebaran peluang dari parameter θ jika diberikan data amatan X , dinotasikan dengan $\pi_{\theta|X}(\theta|x)$. Sebaran posterior dapat ditentukan dengan cara menentukan

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

- Sebaran prediksi (*predictive distribution*) merupakan sebaran peluang bersyarat dari amatan baru jika diberikan data lampau X dinotasikan dengan $f_{Y|X}(y|x)$. Sebaran prediksi dapat ditentukan dengan

$$f_{Y|X}(y|x) = \int f_{Y|\theta}(y|\theta) \pi_{\theta|X}(\theta|x) d\theta.$$

Contoh:

1. Berikut ini merupakan data besarnya pembayaran polis untuk suatu rumah sakit:

125 132 141 107 133 319 126 104 145 223

Besarnya pembayaran tunggal merupakan sebaran pareto dengan $\theta = 100$ dan α tidak diketahui. Sebaran prior mengikuti sebaran Gamma dengan $\alpha = 2, \theta = 1$. Tentukan sebaran posteriornya.

Jawab:

Sebaran prior $\pi(\alpha) = \alpha e^{-\alpha}, \alpha > 0$.

Sebaran untuk data mengikuti fungsi pareto sebagai berikut

$$\begin{aligned}
f_{X|A}(X|\alpha) &= \frac{\alpha^{10}(100)^{10\alpha}}{(\prod_{j=1}^{10} X_j^{\alpha+1})} \\
&= \frac{\alpha^{10}(100)^{10\alpha}}{(4,475132601 \times 10^{21})^{\alpha+1}} \\
&= \alpha^{10} \frac{10^{20\alpha}}{(4,475132601)^{\alpha+1} \cdot 10^{21\alpha+21}} \\
&= \frac{\alpha^{10} 10^{20\alpha}}{(4,475132601)^\alpha (4,475132601) \cdot 10^{21\alpha} \cdot 10^{21}} \\
&= \frac{\alpha^{10}}{(44,75132601)^\alpha (4,475132601) \cdot 10^{21}} \\
&= \frac{\alpha^{10}}{\exp(\alpha \ln(44,75132601)) \cdot 4,475132601 \cdot 10^{21}} \\
&= \frac{\alpha^{10}}{\exp(3,801121\alpha) + 49,852823} \\
&= \alpha^{10} e^{-3,801121\alpha - 49,852823}
\end{aligned}$$

Note: sebaran pareto

$$F(x) = \frac{\alpha \cdot \theta^\alpha}{x^{\alpha+1}}$$

Fungsi peluang bersama untuk X dan A

$$\begin{aligned}
f_{X,A}(X, \alpha) &= f_{X|A}(x|\alpha)\pi(\alpha) \\
&= \alpha^{11} e^{-4,801121\alpha - 49,852823},
\end{aligned}$$

sehingga diperoleh sebaran posterior untuk A yaitu

$$\pi_{A|X}(\alpha|x) = \frac{\alpha^{11} e^{-4,801121\alpha - 49,852823}}{\int_0^\infty \alpha^{11} e^{-4,801121\alpha - 49,852823} d\alpha} = \frac{\alpha^{11} e^{-4,801121\alpha}}{(11!) \left(\frac{1}{4,801121}\right)^{12}}.$$

Terlihat bahwa $A|X \sim \text{gamma} \left(12, \frac{1}{4,801121}\right)$.

Selanjutnya kita dapat menentukan sebaran prediksi, yaitu

$$\begin{aligned}
f_{Y|X}(y|x) &= \int_0^\infty \frac{\alpha \cdot 100^\alpha}{y^{\alpha+1}} \cdot \frac{\alpha^{11} e^{-4,801121\alpha}}{(11!)(1/4,801121)^{12}} d\alpha \\
&= \frac{12 \cdot (4,801121)^{12}}{y(0,195951 + \ln y)^{13}}, y > 100.
\end{aligned}$$

Sebaran ini merupakan sebaran yang tidak familiar. Akan tetapi, dapat dibuktikan $\ln y - \ln 100$ memiliki sebaran pareto.

2. Terdapat dua tipe pengendara sepeda motor, yaitu tipe pengendara baik dan buruk, misalkan θ menyatakan tipe pengendara tersebut. $\theta = \{G, B\}$, dengan G menyatakan baik dan B menyatakan buruk. Misalkan diketahui 75% pengendara adalah termasuk tipe G . Peluang pengendara mengalami sebanyak X kecelakaan diberikan oleh tabel berikut :

Tipe Pengendara	$X = 0$	$X = 1$	$X = 2$
G	0.7	0.2	0.1
B	0.5	0.3	0.2

Misalkan kita punya amatan $X_1 = 0, X_2 = 1$,

- Tentukan sebaran prediksi (*predictive distribution*) dari $X_3 | X_1 = 0, X_2 = 1$.
- Tentukan sebaran posterior $\theta | X_1 = 0, X_2 = 1$.

Jawab:

Sebaran Prediksi (*predictive distribution*)

Misalkan $X = (X_1, X_2)$. Dengan menggunakan fungsi peluang marginal

$$\begin{aligned}
 f_X(0,1) &= \sum f_{X_1|\theta}(0|\theta) f_{X_2|\theta}(1|\theta) \pi(\theta) \\
 &= (0,7)(0,2)(0,75) + (0,5)(0,3)(0,25) \\
 &= 0,1425.
 \end{aligned}$$

Dengan cara yang sama:

$$\begin{aligned}
 f_{XX_3}(0,1,x_3) &= \sum_{\forall \theta} f_{X_1|\theta}(0|\theta) f_{X_2|\theta}(1|\theta) f_{X_3|\theta}(x_3|\theta) \pi(\theta) \\
 f_{XX_3}(0,1,0) &= (0,7)(0,2)(0,7)(0,75) + (0,5)(0,3)(0,5)(0,25) \\
 &= 0,09225
 \end{aligned}$$

$$f_{XX3}(0,1,1) = (0,7)(0,2)(0,2)(0,75) + (0,5)(0,3)(0,3)(0,25) \\ = 0,03225$$

$$f_{XX3}(0,1,2) = (0,7)(0,2)(0,1)(0,75) + (0,5)(0,3)(0,2)(0,25) \\ = 0,01800$$

Sehingga *predictive distribution*:

$$f_{X3|X}(0|0,1) = \frac{0,09225}{0,1425} = 0,6474$$

$$f_{X3|X}(1|0,1) = \frac{0,03225}{0,1425} = 0,2263$$

$$f_{X3|X}(2|0,1) = \frac{0,01800}{0,1425} = 0,1263$$

Sebaran posterior

$$\pi_{\theta|X} = \frac{f_{X,\theta}(x, \theta)}{f_x(x)} = \frac{1}{f_x(x)} \left[\prod_{j=1}^n f_{X_j|\theta}(x_j|\theta) \right] \pi(\theta)$$

$$\pi(G|0,1) = \frac{f(0|G) \cdot f(1|G)\pi(G)}{f(0,1)} = \frac{(0,7)(0,2)(0,25)}{0,1425} = 0,7368$$

$$\pi(B|0,1) = \frac{f(0|B) \cdot f(1|B)\pi(B)}{f(0,1)} = \frac{(0,5)(0,3)(0,25)}{0,1425} = 0,2632$$

Cara lain menentukan sebaran prediksi:

$$f_{X_{n+1}|X}(x_{n+1}|x) = \int f_{X_{n+1}|\theta}(x_{n+1}|\theta) \pi_{\theta|X}(\theta|x) d\theta$$

$$f(0|0,1) = \sum f(0|\theta) \pi(\theta|0,1) \\ = (0,7)(0,7368) + (0,5)(0,2632) = 0,6473$$

$$f(1|0,1) = \sum f(1|\theta) \pi(\theta|0,1) = 0,2263$$

$$f(2|0,1) = (0,2)(0,7368) + (0,3)(0,2632) = 0,1263$$

TEORI KREDIBILITAS (*CREDIBILITY THEORY*)

Pengantar

Credibility theory merupakan sekumpulan analisis kuantitatif yang diakses oleh pemberi asuransi untuk menilai suatu pengalaman prospektif (penentuan premi di masa depan berdasarkan pengalaman di masa lalu) dari sekumpulan risiko atau kelompok risiko. Jika pengalaman dari pemegang polis konsisten, maka penentuan ini akan lebih baik bila dibandingkan dengan penentuan premi secara manual (biasanya disebut *pure premium*).

Seorang pemegang polis biasanya berargumen sebagai berikut : Manual rate (premi manual) biasanya didesain untuk merefleksikan rata-rata pengalaman (masa lalu dan masa depan) dari seluruh tingkat kelas dan secara implisit mengasumsikan bahwa risiko bersifat homogen. Akan tetapi, tidak ada sistem penilaian yang sempurna, dan selalu terdapat ketidakhomogenan dalam tingkat risiko setelah semua kriteria tercatat. Sebagai konsekuensi, beberapa pemegang polis akan lebih berisiko daripada yang diasumsikan dalam sistem premi manual. Tentu ide yang sama menyatakan bahwa tingkat perbedaan premi seharusnya diaplikasikan pada yang berisiko kecil, akan tetapi pemegang polis dalam situasi ini yakin tidak akan meminta kenaikan premi, sehingga suatu kenaikan mungkin diperlukan karena pertimbangan ekuitas dan kondisi ekonomi.

Full Credibility

Misalkan $X_1, \dots, X_n$ menyatakan besarnya kerugian pada waktu ke- j atau banyaknya klaim pada waktu ke- j . Secara umum, ini merupakan n unit *exposure* yang diamati.

Asumsi

$$E[X_j] = \xi,$$

$$\text{Var}(X_j) = \sigma^2,$$

$$\text{Var}(\bar{X}) = \frac{\sigma^2}{n},$$

untuk $j = 1, \dots, n$ dan $\bar{X} = \frac{1}{n} \sum_{j=1}^n (X_j)$.

Berdasarkan prinsip *credibility* klasik, jika $\bar{X}$ “*credible*”, kita dapat menyatakan suatu *full credibility* untuk $\bar{X}$ atau menggunakan $\bar{X}$ untuk menduga total kehilangan untuk amatan ke $n + 1$.

Adapun untuk menyatakan *credibility* untuk $\bar{X}$ kita harus mendefinisikan error relatif terlebih dahulu. Jika untuk suatu nilai peluang p , $0 \leq p \leq 1$ dan untuk tingkat error ϵ , jika dipenuhi

$$P\left(\left|\frac{\bar{X} - \xi}{\xi}\right| \leq \epsilon\right) \geq p,$$

maka kita katakan $\bar{X}$ adalah *credible* untuk ξ .

Tanpa kehilangan asumsi-asumsi dari sebaran X_j , $\bar{X}$ dapat diaproksimasi dengan Teorema Limit Pusat sebagai sebaran normal.

$$\bar{X} \sim \text{Normal}\left(\xi, \frac{\sigma^2}{n}\right),$$

sehingga $\frac{\bar{X} - \xi}{\frac{\sigma}{\sqrt{n}}} \sim \text{Normal}(0,1)$.

Diberikan asumsi untuk n yang besar kita dapat menurunkan beberapa kondisi yang ekuivalen untuk *full credibility*.

$$P\left(\left|\frac{\bar{X} - \xi}{\xi}\right| \leq \epsilon\right) \geq p \tag{1}$$

$$\Leftrightarrow P\left(\left|\frac{\bar{X} - \xi}{\frac{\sigma}{\sqrt{n}}}\right| \leq \frac{\epsilon \xi \sqrt{n}}{\sigma}\right) \geq p. \tag{2}$$

Sekarang kita menggunakan fakta bahwa :

$$\frac{\bar{X} - \xi}{\frac{\sigma}{\sqrt{n}}} \sim \text{Normal}(0,1),$$

Untuk penyederhanaan, kita misalkan :

$$Y = \frac{\bar{X} - \xi}{\frac{\sigma}{\sqrt{n}}} \sim \text{Normal}(0,1)$$

Dan kita turunkan kondisi ekuivalen dari persamaan (2) sebagai berikut :

$$P\left(\left|\frac{\bar{X} - \xi}{\frac{\sigma}{\sqrt{n}}}\right| \leq \frac{\varepsilon \xi \sqrt{n}}{\sigma}\right) \geq p$$

Misalkan $P(|Y| \leq y_p) = p$. (3)

Persamaan (1) dan (2) akan dipenuhi jika :

$$\frac{\varepsilon \xi \sqrt{n}}{\sigma} \geq y_p. \quad (4)$$

Berdasarkan pertidaksamaan-pertidaksamaan di atas kita dapat menurunkan hubungan berikut :

$$\begin{aligned} \frac{\varepsilon \xi \sqrt{n}}{\sigma} &\geq y_p \\ \Leftrightarrow \frac{\sigma}{\xi} &\leq \frac{\varepsilon \sqrt{n}}{y_p} = \sqrt{\frac{n}{\lambda_0}} \end{aligned} \quad (5)$$

$$\Leftrightarrow Var(\bar{X}) \leq \frac{\xi^2}{\lambda_0} \quad (6)$$

$$\Leftrightarrow n \geq \lambda_0 \left(\frac{\sigma}{\xi}\right)^2, \quad (7)$$

di mana $\lambda_0 = \left(\frac{y_p}{\varepsilon}\right)^2$ dan $\frac{\sigma}{\xi}$ merupakan *coefficient of variation* dari X_j .

Misalkan Φ adalah notasi untuk fungsi sebaran kumulatif dari $N(0,1)$, maka nilai y_p dapat ditentukan sebagai berikut:

$$P(|Y| \leq y_p) = p \Leftrightarrow \Phi(y_p) - \Phi(-y_p) = p$$

$$\Leftrightarrow 2\Phi(y_p) - 1 = p$$

$$\Leftrightarrow \Phi(y_p) = \frac{1+p}{2}$$

$$\Leftrightarrow y_p = \Phi^{-1}\left(\frac{1+p}{2}\right).$$

Dengan demikian y_p adalah persentil $100\left(\frac{1+p}{2}\right)\%$ dari sebaran $N(0,1)$.

Contoh Soal:

Misalkan $X_1, \dots, X_n$ merupakan sampel dengan $E(X_j) = \xi$. Jika nilai amatan dari 10 data dengan 6 data bernilai nol, sedangkan yang lainnya 353, 394, 439, 756, maka tentukan apakah kondisi *full credibility* dipenuhi, dengan $\epsilon = 0.05$, $p = 0.9$?

Jawab:

Syarat full credibility: $n \geq \lambda_0 \left(\frac{\sigma}{\xi} \right)^2$.

Dari soal diperoleh $\xi = 184.6$

$$\sigma = 267.89$$

$$\lambda_0 = \left(\frac{y_p}{\epsilon} \right)^2 = \left(\frac{y_{0.9}}{0.05} \right)^2 = \left(\frac{1.645}{0.05} \right)^2 = 1082.41,$$

sehingga $n \geq 1082.41 \left(\frac{267.89}{184.6} \right)^2 = 2279.51$. Jadi, kondisi *full credibility* tidak dipenuhi.

Limited Fluctuation Credibility

Misalkan X adalah peubah acak Poisson di mana

$$X = Y_1 + Y_2 + \cdots + Y_n = \sum_{i=1}^N Y_i,$$

di mana N menyatakan banyaknya klaim dan Y_j menyatakan besarnya klaim j . Asumsi yang biasa digunakan adalah N, Y_j saling bebas dan Y_j memiliki sebaran yang identik. Dengan asumsi ini dapat dibuat formula bersyarat untuk menghitung ξ dan σ^2 yang dinyatakan dalam N dan Y_j .

$$E[X] = \xi = E\left[\sum_{i=1}^N Y_i\right] = E[E(\sum_{i=1}^N Y_i | N)] = E[NE[Y_i]] = E[N]E[Y_i].$$

$$\begin{aligned} Var(X) = \sigma^2 &= Var\left(E(\sum_{i=1}^N Y_i | N)\right) + E[Var(\sum_{i=1}^N Y_i | N)] \\ &= Var(NE[Y_i]) + E[NVar(Y_i)] \\ &= (E[Y_i])^2 Var(N) + Var(Y_i)E[N]. \end{aligned}$$

Contoh Soal

Misalkan banyaknya klaim di masa lalu $X_1, \dots, X_n$ adalah peubah acak Poisson dengan parameter λ dan besarnya klaim menyebar dengan suatu sebaran tertentu dengan nilai harapan θ_Y dan ragam σ_Y^2 , maka

$$\xi = E[X_j] = \lambda\theta_Y$$

$$\sigma^2 = Var(X_j) = \lambda(\sigma_Y^2 + \theta_Y^2).$$

Kondisi yang harus dipenuhi agar terjadi *full credibility* ialah

$$n \geq \lambda_0 \left(\frac{\sigma}{\xi}\right)^2 = \lambda_0 \frac{\lambda(\sigma_Y^2 + \theta_Y^2)}{(\lambda\theta_Y)^2} = \frac{\lambda_0}{\lambda} \left(1 + \left(\frac{\sigma_Y}{\theta_Y}\right)^2\right).$$

Kondisi *full credibility* untuk rata-rata total banyak klaim di masa lalu ialah

$$n\lambda \geq \lambda_0 \left(1 + \left(\frac{\sigma_Y}{\theta_Y}\right)^2\right).$$

Kondisi *full credibility* untuk rata-rata total besar klaim di masa lalu ialah

$$n\lambda\theta_Y \geq \lambda_0 \left(\theta_Y + \frac{\sigma_Y^2}{\theta_Y} \right).$$

Partial Credibility

Jika kondisi *full credibility* tidak dipenuhi, maka *partial credibility* dapat digunakan untuk data klaim di masa lalu. Didefinisikan *partial credibility* sebagai berikut

$$P_c = Z\bar{X} + (1 - Z)M.$$

Definisi ini merupakan rata-rata terbobot dari besarnya klaim $\bar{X}$ dan *manual premium* yang ditentukan dengan menduga kerugian untuk data ke $n + 1$. Nilai Z ialah $0 \leq Z \leq 1$ dan $M > 0$.

Selanjutnya bagaimana kita dapat menentukan *credibility factor* Z ? Bagaimana kita memilih Z sedemikian sehingga meminimumkan ragam dari peubah acak *partial credibility*:

$$\text{Var}(P_c) = Z^2 \text{Var}(\bar{X}) = Z^2 \frac{\sigma^2}{n}.$$

Dari konsep *full credibility* kita punya kondisi ekuivalen berikut

$$\begin{aligned} \text{Var}(\bar{X}) &= \frac{\sigma^2}{n} \leq \frac{\xi^2}{\lambda_0}, \\ \frac{\xi^2}{\lambda_0} &\geq \frac{\sigma^2}{n} \Leftrightarrow \frac{\xi^2}{\lambda_0} \geq Z^2 \frac{\sigma^2}{n} \Leftrightarrow Z^2 \frac{\sigma^2}{n} \leq \frac{\xi^2}{\lambda_0} \end{aligned}$$

Kita bisa menyatakan

$$\text{Var}(P_c) = Z^2 \frac{\sigma^2}{n} = \frac{\xi^2}{\lambda_0},$$

sehingga diperoleh $Z = \frac{\xi}{\sigma} \sqrt{\frac{n}{\lambda_0}}$. Karena nilai $Z \leq 1$, maka kita perlu membatasi nilai Z ini dengan

$$Z = \min \left\{ \frac{\xi}{\sigma} \sqrt{\frac{n}{\lambda_0}}, 1 \right\}.$$

Contoh soal

1. Misalkan data klaim $X_1, \dots, X_n$ menyebar bebas dan identik dan merupakan compound Poisson dengan besarnya klaim menyebar eksponensial. Jika *credibility factor* berdasarkan banyaknya klaim ialah 0.8. Tentukan *credibility factor* berdasarkan total klaim jika diberikan $\lambda_0 = 1082.41$.

Jawab :

Diketahui $X_j = \sum_{i=1}^N Y_i$ adalah peubah acak compound Poisson dengan besar klaim $Y_j \sim \text{eksponensial}(\lambda)$ dan banyaknya klaim $N \sim \text{Poisson}(\lambda)$.

Berdasarkan banyaknya klaim kita punya

$$\begin{aligned}\xi &= E[N] = \lambda \\ \sigma^2 &= \text{Var}(N) = \lambda \\ Z &= \frac{\xi}{\sigma} \sqrt{\frac{n}{\lambda_0}} = 0.8\end{aligned}$$

Berdasarkan besar total klaim kita punya

$$\begin{aligned}\xi &= E[X_j] = \lambda \theta_Y \\ \sigma^2 &= \text{Var}(X_j) = \lambda(\sigma_Y^2 + \theta_Y^2) \\ Z &= \frac{\xi}{\sigma} \sqrt{\frac{n}{\lambda_0}} = 0.566.\end{aligned}$$

Catatan :

- Untuk *credibility* dari banyaknya klaim kita punya

$$\begin{aligned}\xi &= E[N] \\ \sigma^2 &= \text{Var}(N).\end{aligned}$$

- Untuk *credibility* dari besarnya klaim kita punya

$$\begin{aligned}\xi &= E[X_j] \\ \sigma^2 &= \text{Var}(X_j).\end{aligned}$$

2. Dari soal sebelumnya (contoh soal pada *full credibility*) jika manual premium M ialah 225. Tentukan dugaan *premi credibility*.

Jawab:

$$\xi = 184.6$$

$$\lambda_0 = 1082.41$$

$$\sigma = 267.89$$

$$Z = \left(\frac{\xi}{\sigma}\right) \sqrt{\frac{10}{\lambda_0}} = 0.06623$$

$$P_c = Z\bar{X} + (1 - z)M$$

$$= 0,06623(104,6) + 0,93377(225) = 222,32.$$

Greatest Accuracy Credibility

Bayesian premium adalah total kehilangan di masa lalu dari amatan $X_1, \dots, X_n$. Dalam aplikasinya, kita akan lebih tertarik untuk menentukan besar kehilangan di amatan berikutnya X_{n+1} . Asumsi yang digunakan adalah bahwa $\Theta = \theta$ menjadikan $X_1, \dots, X_n, X_{n+1}$ *conditional independent* dan memiliki sebaran peluang $f(x|\theta)$ yang sama. Θ dikatakan memiliki sebaran prior $\pi(\theta)$.

Suatu peubah acak dikatakan *conditional independent* jika

$$f(x_1 x_2 | \theta) = f(x_1 | \theta) f(x_2 | \theta),$$

dan memenuhi

$$E[X_1 X_2 | \theta] = E[X_1 | \theta] E[X_2 | \theta].$$

Berikut ini beberapa konsep penting dalam menentukan *Bayesian premium*

- (a) Sebaran posterior dari Θ yang diberikan $X_1 = x_1, \dots, X_n = x_n$ adalah

$$\begin{aligned} \pi(\theta | x_1, \dots, x_n) &= \frac{(\prod_{i=1}^n f(x_i | \theta)) \pi(\theta)}{\int_{\forall \theta} (\prod_{i=1}^n f(x_i | \theta)) \pi(\theta) d\theta} \\ &\propto \left(\prod_{i=1}^n f(x_i | \theta) \right) \pi(\theta). \end{aligned}$$

- (b) Fungsi sebaran peluang bersyarat atau fungsi peluang dari X_{n+1} jika diberikan $X_1 = x_1, \dots, X_n = x_n$ adalah:

$$\begin{aligned} f(x_{n+1} | x_1, \dots, x_n) &= \frac{\int_{\forall \theta} (\prod_{i=1}^{n+1} f(x_i | \theta)) \pi(\theta) d\theta}{\int_{\forall \theta} (\prod_{i=1}^n f(x_i | \theta)) \pi(\theta) d\theta} \\ &= \int_{\forall \theta} f(x_{n+1} | \theta) \pi(\theta | x_1, \dots, x_n) d\theta. \end{aligned}$$

Persamaan ini dikenal sebagai *predictive distribution* dari X_{n+1} jika diberikan $X_1 = x_1, \dots, X_n = x_n$.

- (c) *Hypothetical mean* atau premi individu (*individual premium*) untuk amatan ke $n + 1$ adalah:

$$\mu_{n+1}(\theta) = E[X_{n+1} | \Theta = \theta]$$

- (d) *Pure premium* atau *collective premium* untuk amatan ke $n + 1$ adalah

$$\mu_{n+1} = E[X_{n+1}] = E[E[X_{n+1} | \Theta = \theta]] = E[\mu_{n+1}(\Theta)]$$

- (e) *Bayesian premium* untuk amatan ke $n + 1$ adalah

$$\begin{aligned} E[X_{n+1} | x_1, \dots, x_n] &= \int_{\forall x_{n+1}} x_{n+1} f(x_{n+1} | x_1, \dots, x_n) dx_{n+1} \\ &= \int_{\forall x_{n+1}} x_{n+1} \int_{\forall \theta} f(x_{n+1} | \theta) \pi(\theta | x_1, \dots, x_n) d\theta dx_{n+1} \\ &= \int_{\forall \theta} \left(\int_{\forall x_{n+1}} x_{n+1} f(x_{n+1} | \theta) dx_{n+1} \right) \pi(\theta | x_1, \dots, x_n) d\theta \\ &= \int_{\forall \theta} \mu_{n+1}(\theta) \pi(\theta | x_1, \dots, x_n) d\theta. \end{aligned}$$

Contoh soal

- Misalkan banyaknya klaim dari seorang pemegang polis pada tahun j adalah X_j , $j = 1, 2$. Diberikan $\Theta = \theta$, X_j saling bebas dan memiliki fungsi peluang

$$P(X = 1 | \Theta = \theta) = 1 - \theta$$

$$P(X = 2 | \Theta = \theta) = \theta.$$

Sebaran prior $\Theta \sim \text{Beta}(2, 3)$, $0 < \theta < 1$, tentukan fungsi peluang untuk $X_2 | X_1 = 2$. Tentukan pula *Bayesian premium* nya.

Jawab

$$\begin{aligned}
 P(X_2 = 2|X_1 = 2) &= \frac{P(X_1 = 2, X_2 = 2)}{P(X_1 = 2)} \\
 &= \frac{\int_0^1 P(X_1 = 2|\theta)P(X_2 = 2|\theta)\pi(\theta)d\theta}{\int_0^1 P(X_1 = 2|\theta)\pi(\theta)d\theta} \\
 &= \frac{E[\theta^2]}{\theta} = \frac{1}{2}.
 \end{aligned}$$

Selanjutnya karena hanya ada dua kemungkinan untuk nilai X_2 , maka kita peroleh

$$P(X_2 = x|X_1 = 2) = \begin{cases} \frac{1}{2} & x = 2 \\ 1 - \frac{1}{2} & x = 1 \end{cases}$$

sehingga Bayesian premium

$$E[X_2|X_1 = 2] = 1 \cdot \frac{1}{2} + 2 \cdot \frac{1}{2} = \frac{3}{2}.$$

2. Dari contoh soal pada Bab 1, *Hypothetical means* untuk $\theta = \{G, B\}$ adalah

$$\mu_3(G) = 0(0.7) + 1(0.2) + 2(0.1) = 0.4$$

$$\mu_3(B) = 0(0.5) + 1(0.3) + 2(0.2) = 0.7.$$

Pada contoh ini telah diketahui bahwa $X_1 = 0$, $X_2 = 1$, sehingga kita dapat menentukan *Bayesian Premium* dengan menggunakan fungsi peluang $f(x_{n+1}|x_1, \dots, x_n)$ sebagai berikut:

$$\begin{aligned}
 E(X_3|0,1) &= 0(0,6474) + 1(0,2263) + 2(0,1263) \\
 &= 0,4790.
 \end{aligned}$$

Selain dengan cara ini, *Bayesian Premium* dapat juga ditentukan dengan menggunakan *hypothetical mean*.

$$E(X_3|X_1 = 0, X_2 = 1) = (0,4)(0,7368) + 0,7(0,2632) = 0,4790$$

Adapun *pure premium* nya adalah

$$\begin{aligned}\mu_3 &= E(X_3) = \sum_{\forall \theta} \mu_3(\theta)\pi(\theta) \\ &= (0,4)(0,75) + (0,7)(0,25) = 0,475\end{aligned}$$

3. Sebaran agregat kehilangan untuk 2 risiko dalam satu periode amatan adalah sebagai berikut

Risiko	Aggregate loss		
	\$0	\$50	\$1000
A	0.80	0.16	0.04
B	0.60	0.24	0.16

Sebuah risiko dipilih pada amatan acak yang memiliki 0 kehilangan di tahun pertama dan kedua periode amatan. Tentukan Bayesian premium dari nilai harapan dari *aggregate losses* untuk risiko yang sama pada tahun ketiga periode tersebut.

Jawab :

Mula-mula kita tentukan peluang tak syarat :

$$\begin{aligned}P[A \cap (0 \text{ klaim di thn 1 dan 0 klaim di tahun 2})] \\ &= P[0 \text{ klaim di thn 1 dan 0 klaim di tahun 2}|A]. P[A] \\ &= (0.8)(0.8)(0.5) = 0.32.\end{aligned}$$

$$\begin{aligned}P[B \cap (0 \text{ klaim di thn 1 dan 0 klaim di tahun 2})] \\ &= P[0 \text{ klaim di thn 1 dan 0 klaim di tahun 2}|B]. P[B] \\ &= (0.6)(0.6)(0.5) = 0.18,\end{aligned}$$

$$\text{sehingga } P[0,0] = 0.32 + 0.18 = 0.50$$

Selanjutnya dapat kita tentukan peluang bersyarat:

$$P[A|0,0] = \frac{P[A \cap (0,0)]}{P[0,0]} = \frac{0.32}{0.5} = 0.64,$$

$$P[B|0,0] = \frac{P[B \cap (0,0)]}{P[0,0]} = \frac{0.18}{0.5} = 0.36.$$

Nilai harapan bersyarat dari agregat klaim per periode untuk masing masing risiko:

$$E[X|A] = (0)(0.8) + (50)(0.16) + (1000)(0.04) = 48,$$

$$E[X|B] = (0)(0.6) + (50)(0.24) + (1000)(0.16) = 172,$$

Nilai harapan bersyarat dari agregat klaim per periode diberikan tanpa klaim pada dua periode pertama ditemukan menggunakan peluang Bayesian yang diperbarui dari A dan B :

$$\begin{aligned} E[X|0,0] &= E[X|A].P[A|0,0] + E[X|B].P[B|0,0] \\ &= (48)(0.64) + (172)(0.36) = 92.64. \end{aligned}$$

4. Misalkan N_j menyatakan total banyak klaim di tahun ke j untuk suatu grup pemegang polis yang memiliki m_j anggota di tahun $j, j = 1, 2, \dots, n, n+1$. $N_j | (\theta = \theta) \sim \text{Poisson}(m_j \theta)$ dan *conditional independent*, $\theta \sim \text{Gamma}(\alpha, \beta)$. Selanjutnya $X_j = \frac{N_j}{m_j}$ adalah rata-rata banyaknya klaim per tahun. Jika diberikan $X_1 = x_1, \dots, X_n = x_n$, gunakan *Bayesian premium* untuk menduga rata-rata banyaknya klaim untuk setiap anggota pada tahun ke $n+1$.

Jawab:

$$\begin{aligned} f_j(x_j|\theta) &= P(X_j = x_j|\theta = \theta) = P(N_j = x_j m_j|\theta = \theta) \\ &= \frac{e^{-m_j\theta} (m_j\theta)^{m_j x_j}}{(m_j x_j)!}. \end{aligned}$$

Nilai yang mungkin untuk X_j adalah $\frac{N_j}{m_j} = \left\{ \frac{0}{m_j}, \frac{1}{m_j}, \frac{2}{m_j}, \dots \right\}$.

Selanjutnya dapat ditentukan sebaran posterior :

$$\begin{aligned} \pi(\theta|x_1, \dots, x_n) &\propto \left(\prod_{i=1}^n f(x_i|\theta) \right) \pi(\theta) \\ &\propto \left(\prod_{i=1}^n e^{-m_i\theta} (m_i\theta)^{m_i x_i} \right) \left(e^{\alpha-1} e^{-\frac{\theta}{\beta}} \right) \\ &\propto e^{-\theta \left(\frac{1}{\beta} + \sum_{i=1}^n m_i \right)} \theta^{\alpha-1 + \sum_{i=1}^n m_i x_i}. \end{aligned}$$

Terlihat bahwa sebaran posterior mengikuti sebaran Gamma:

$$\theta|X_1 = x_1, \dots, X_n = x_n \sim \text{Gamma} \left(\alpha + \sum_{i=1}^n m_i x_i, \frac{1}{\frac{1}{\beta} + \sum_{i=1}^n m_i} \right).$$

Selanjutnya kita akan mencari *hypothetical mean* atau *individual premium*:

$$\mu_{m+1}(\theta) = E[X_{n+1}|\theta = \theta] = \frac{1}{m_{n+1}} E[N_{n+1}|\theta = \theta] = \theta.$$

Karena *hypothetical mean* hanya θ , yang menjadi nilai harapan sebaran posterior, sehingga *Bayesian premium* nya adalah

$$\begin{aligned}
 E[X_{n+1}|x_1, \dots, x_n] &= \int_{\forall \theta} \theta \pi(\theta|x_1, \dots, x_n) d\theta \\
 &= \left(\alpha + \sum_{i=1}^n m_i x_i \right) \frac{1}{\frac{1}{\beta} + \sum_{i=1}^n m_i}
 \end{aligned}$$

Agar bentuk ini lebih menarik untuk dilihat, misalkan

$$\begin{aligned}
 m &= \sum_{i=1}^n m_i \\
 \bar{X}^* &= \sum_{j=1}^n \frac{m_j}{m} X_j \\
 Z &= \frac{m}{m + \beta^{-1}},
 \end{aligned}$$

maka *Bayesian premium* dapat dinyatakan sebagai berikut:

$$E[X_{n+1}|x_1, \dots, x_n] = \alpha\beta(1 - Z) + Z\bar{X}^*.$$

Bentuk $\alpha\beta$ merupakan nilai harapan dari sebaran Θ ,

$$E[\Theta] = E[\mu_{n+1}(\Theta)] = \alpha\beta = E[X_{n+1}].$$

Credibility Premium

Metode lain yang digunakan untuk menduga X_{n+1} adalah dengan menggunakan kombinasi linear dari amatan di masa lampau

$$\alpha_0 + \alpha_1 X_1 + \cdots + \alpha_n X_n.$$

Untuk memperoleh model yang baik maka kita perlu meminimumkan *mean square error* antara model yang sebenarnya dengan penduga.

$$Q = Q(\alpha_0, \alpha_1, \dots, \alpha_n) = E \left[(X_{n+1} - (\alpha_0 + \alpha_1 X_1 + \cdots + \alpha_n X_n))^2 \right].$$

Selanjutnya kita perlu mencari penduga $\hat{\alpha}_0, \hat{\alpha}_1, \dots, \hat{\alpha}_n$ sedemikian sehingga

$$Q(\hat{\alpha}_0, \hat{\alpha}_1, \dots, \hat{\alpha}_n) = \min_{\forall \alpha_i, i=0, \dots, n} Q(\alpha_0, \alpha_1, \dots, \alpha_n).$$

Kita bisa mengambil $n + 1$ turunan parsial untuk mengoptimasi ukuran tersebut.

Untuk $i = 0$

$$\begin{aligned} \frac{\partial}{\partial \alpha_0} Q &= E[2(X_{n+1} - (\alpha_0 + \alpha_1 X_1 + \cdots + \alpha_n X_n))(-1)] = 0 \\ &\rightarrow E[X_{n+1}] = \alpha_0 + \alpha_1 E[X_1] + \cdots + \alpha_n E[X_n] \\ &= \alpha_0 + \sum_{i=1}^n \alpha_i E[X_i] \end{aligned} \tag{1}$$

Untuk $i = 1, 2, \dots, n$

$$\frac{\partial}{\partial \alpha_i} Q = E[2(X_{n+1} - \alpha_0 - \alpha_1 X_1 - \cdots - \alpha_n X_n)(-X_i)] = 0$$

$$\begin{aligned}\rightarrow E[X_{n+1}X_i] &= \alpha_0 E[X_i] + \alpha_1 E[X_1X_i] + \cdots + \alpha_i E[X_i^2] + \cdots + \alpha_n E[X_nX_i] \\ &= \alpha_0 E[X_i] + \sum_{j=1}^n \alpha_j E[X_jX_i].\end{aligned}\quad (2)$$

Jika kita kurangkan persamaan (2) dengan persamaan (1) dikali $E[X_i]$, maka diperoleh

$$Cov(X_{n+1}, X_i) = \sum_{j=1}^n \alpha_j Cov(X_i, X_j) \quad (3)$$

Persamaan (3) disebut juga persamaan normal (*normal equations*). Solusi untuk $\hat{\alpha}_0, \hat{\alpha}_1, \dots, \hat{\alpha}_n$ dan $\hat{\alpha}_0 + \sum_{j=1}^n \hat{\alpha}_j X_j$ merupakan premi untuk amatan ke $n + 1$.

Catatan : *credibility premium* $\hat{\alpha}_0 + \sum_{j=1}^n \hat{\alpha}_j X_j$ merupakan penduga linear terbaik untuk $E[X_{n+1}|X_1, \dots, X_n]$ dan $\mu_{n+1}(\theta)$.

Teorema

Jika $X_j|\theta, j = 1, 2, \dots, n, n + 1$ saling bebas dan $E[X_j] = \mu$, $Var(X_j) = \sigma^2$ dan $Cov(X_i, X_j) = \rho\sigma^2, i \neq j$ dan $-1 < \rho < 1$, maka *credibility premium* untuk periode $n + 1$ adalah

$$Z\bar{X} + (1 - Z)\mu \text{ dengan } Z = \frac{n\rho}{1 - \rho + n\rho}.$$

Bukti :

Dari persamaan (1) kita punya

$$\begin{aligned}
 E[X_{n+1}] &= \alpha_0 + \sum_{i=1}^n \alpha_i E[X_j] \\
 \rightarrow \mu &= \alpha_0 + \sum_{i=1}^n \alpha_i \mu \\
 \rightarrow \mu &= \alpha_0 + \mu \sum_{i=1}^n \alpha_i \\
 \rightarrow \sum_{i=1}^n \alpha_i &= 1 - \frac{\alpha_0}{\mu}
 \end{aligned} \tag{4}$$

Dari persamaan (3) kita punya

$$\begin{aligned}
 Cov(X_{n+1}, X_i) &= \sum_{j=1}^n \alpha_j Cov(X_i, X_j); \quad i = 1, 2, \dots, n \\
 \rightarrow \rho \sigma^2 &= \sum_{j \neq i} \alpha_j \rho \sigma^2 + \alpha_i \sigma^2 \\
 \rightarrow \rho &= \rho \sum_{j \neq i} \alpha_j + \alpha_i \\
 \rightarrow \rho &= \rho \sum_{j=1}^n \alpha_j - \rho \alpha_i + \alpha_i = \rho \left(1 - \frac{\alpha_0}{\mu}\right) + \alpha_i (1 - \rho) \\
 \rightarrow \alpha_i &= \frac{\rho \alpha_0}{\mu(1 - \rho)} \quad i = 1, 2, \dots, n
 \end{aligned} \tag{5}$$

Selanjutnya dengan menjumlahkan persamaan (5) untuk semua nilai i , $i \leq 1 \leq n$, diperoleh

$$\sum_{i=1}^n \alpha_i = \frac{n \rho \alpha_0}{\mu(1 - \rho)} = 1 - \frac{\alpha_0}{\mu}. \tag{6}$$

Dengan menyubstitusikan persamaan (6) ke persamaan (4) diperoleh

$$\alpha_0 = \frac{(1 - \rho)\mu}{1 - \rho + n\rho}.$$

Dengan menyubstitusikan α_0 yang diperoleh tersebut ke persamaan (5) diperoleh

$$\alpha_i = \frac{\rho}{1 - \rho + n\rho} ; i = 1, 2, \dots, n$$

Selanjutnya kita bisa memperoleh *credibility premium* sebagai berikut

$$\begin{aligned} \alpha_0 + \sum_{j=1}^n \alpha_j X_j &= \frac{(1 - \rho)\mu}{1 - \rho + n\rho} + \frac{\rho}{1 - \rho + n\rho} \sum_{j=1}^n X_j \\ &= (1 - Z)\mu + Z\bar{X}, \end{aligned}$$

dengan $Z = \frac{n\rho}{1 - \rho + n\rho}$.

Jika $n \rightarrow \infty$ maka $Z \rightarrow 1$, sebagaimana kita ketahui bahwa semakin besar ukuran amatan, maka bobot yang diberikan untuk data akan semakin besar. Jika $\rho = 0$, yang mengindikasikan bahwa data tidak berhubungan maka bobot yang diberikan adalah $Z = 0$.

Buhlmann Credibility

Model ini merupakan model kredibilitas yang paling sederhana. Yang spesifik pada masing masing pemegang polis (bergantung pada θ), kehilangan dari masa lalu $X_1, \dots, X_n$ memiliki rata-rata dan ragam yang sama dan i.i.d bergantung pada θ .

Didefinisikan

$$\mu(\theta) = E(X_j | \theta = \theta),$$

dan

$$V(\theta) = \text{var}(X_j | \theta = \theta)$$

$\mu(\theta)$ merupakan *hypothetical mean*.

$V(\theta)$ merupakan *process variance*.

Didefinisikan pula

$$\mu = E(\mu(\theta))$$

merupakan *Expected value of the hypothetical means*,

$$v = E(V(\theta)),$$

dan

$$a = \text{var}(\mu(\theta)).$$

μ disebut juga *collective premium*.

Selanjutnya kita dapat menentukan rata-rata, ragam dan koragam X_j sebagai berikut:

$$E(X_j) = E[E(X_j | \theta)] = E[\mu(\theta)] = \mu$$

$$\text{var}(X_j) = E[\text{var}(X_j | \theta)] + \text{var}[E(X_j | \theta)] = \mu$$

$$= E(v(\theta)) + \text{var}(\mu(\theta))$$

$$= v + a$$

$$\text{cov}(X_i X_j) = E(X_i X_j) - E(X_i)E(X_j), \quad i \neq j$$

$$= E[E(X_i X_j | \theta)] - \mu^2 \dots \text{identik}$$

$$\begin{aligned}
 &= E(E(X_i|\theta) \cdot E(X_j|\theta)) - [E(\mu(\theta))]^2 \\
 &= E((\mu(\theta))^2) - [E(\mu(\theta))]^2 \\
 &= \text{var}(\mu(\theta)) = a.
 \end{aligned}$$

Dari Teorema di atas kita dapat menentukan Z sebagai berikut:

$$\begin{aligned}
 Z &= \frac{\rho}{1 - \rho + n\rho} = \frac{\frac{\text{cov}(X_i X_j)}{\text{var}(X_j)}}{1 - \frac{\text{cov}(X_i X_j)}{\text{var}(X_j)} + n \frac{\text{cov}(X_i X_j)}{\text{var}(X_j)}} \\
 &= \frac{\frac{a}{v + a}}{1 - \frac{a}{v + a} + n \frac{a}{v + a}} = \frac{na}{v + na} = \frac{n}{k + n}
 \end{aligned}$$

dengan $k = \frac{v}{a} = \frac{E[\text{var}(X_j|\theta)]}{\text{var}[E(X_j|\theta)]}$.

Sehingga premi kredibilitasnya adalah

$$\alpha_o + \sum_{j=1}^n \alpha_j X_j = Z\bar{X} + (1 + Z)\mu.$$

Contoh Soal

1. (Melanjutkan contoh sebelumnya pada Pertemuan 1) Tentukan *Buhlman estimate* dari $E(X_3|X_1 = 0, X_2 = 1)$

Jawab:

Sebelumnya sudah kita peroleh

$$\mu(G) = 0,4$$

$$\mu(B) = 0,7$$

$$\pi(G) = 0,75$$

$$\pi(B) = 0,25$$

$$\rightarrow \mu = E(\mu(\theta))$$

$$= (0,4)(0,75) + (0,7)(0,25) = 0,475$$

$$\begin{aligned} \alpha &= \text{var}(\mu(\theta)) = E(\mu(\theta)^2) - [E(\mu(\theta))]^2 \\ &= (0,16)0,75 + 0,49(0,25) - (0,475)^2 \\ &= 0,0169 \end{aligned}$$

Untuk *process variance*

$$v = E(v(\theta)) = \sum_{\sigma} v(\sigma)\Pi(\sigma)$$

dengan

$$\begin{aligned} v(G) &= \text{var}(x_j|G) \\ &= 0^2(0,7) + 1^2(0,2) + 2^2(0,1) - 0,4^2 = 0,44 \end{aligned}$$

$$\begin{aligned} v(B) &= \text{var}(x_j|B) \\ &= 0^2(0,5) + 1^2(0,3) + 2^2(0,2) - 0,7^2 = 0,61 \end{aligned}$$

$$\rightarrow v = 0,44(0,75) + 0,61(0,25) = 0,4825.$$

Jadi diperoleh

$$k = \frac{v}{\alpha} = \frac{0,4825}{0,0169} = 28,5926$$

$$\rightarrow Z = \frac{n}{n+k} = \frac{2}{2+28,8926} = 0,0654$$

Buhlmann Credibility

$$0,0654(0,5) + 0,9346(0,475) = 0,4766.$$

2. Asumsikan banyaknya klaim tahun untuk seorang individu yang diasuransikan memiliki sebaran Poisson. Nilai harapan frekuensi klaim

tahunan (parameter poisson Δ) dari member populasi yang diasuransikan menyebar seragam (0,1). Rataan frekuensi klaim tahunan individu adalah konstan sepanjang waktu. Tentukan *buhlman credibility factor* Z jika ada n amatan,

Jawab:

Rumus: $Z = \frac{n}{n + \frac{v}{a}}$

Hypothetical mean $= \mu(\lambda) = E(X|\lambda) = \Delta \quad \dots \quad X|\lambda \sim \text{poisson}(\Delta)$

$$v(\lambda) = \text{var}(X|\lambda) = \Delta$$

$$\pi = EX = E(E(X|\lambda)) = E(\mu(\lambda)) = E(\Delta) = 0,5$$

$$v = E(V(\lambda)) = E(\Delta) = 0,5$$

$$\alpha = \text{var}(\mu(A)) = \text{var}(\Delta) = \frac{1}{12}$$

$$Z = \frac{n}{n + \frac{1/2}{1/12}} = \frac{n}{n + 6}.$$

Premi Kredibilitas Buhlmann-Starub (*Buhlmann-Straub Credibility Premium*)

Kita akan memperkenalkan suatu model *credibility premium* baru, yaitu model Buhlmann-Straub. Model ini memiliki beberapa asumsi tambahan daripada model Buhlmann. Diberikan $\theta, X_1, \dots, X_{n+1}$ saling bebas untuk $j = 1, 2, \dots, n + 1$ berlaku:

$$E[X_j|\theta] = \mu(\theta) \quad \text{Var}(X_j|\theta) = \frac{v(\theta)}{m_j}.$$

Selanjutnya kita notasikan kuantitas berikut:

$$\mu = E[\mu(\theta)] \quad a = \text{Var}(\mu(\theta))$$

$$v = E[v(\theta)] \quad k = \frac{v}{a}$$

Dalam model Buhlmann-Straub kita dapat membuktikan bahwa kuantitas di atas dapat digunakan untuk menentukan *credibility premium*.

$$E[X_j] = E[E[X_j|\theta]] = E[\mu(\theta)] = \mu.$$

Untuk $i \neq j$

$$\begin{aligned} \text{Cov}(X_i, X_j) &= E[X_i X_j] - E[X_i]E[X_j] = E[E[X_i X_j|\theta]] - \mu^2 \\ &= E[E[X_i|\theta]E[X_j|\theta]] - \mu^2 \\ &= E[\mu(\theta)^2] - (E[\mu(\theta)])^2 = \text{Var}(\mu(\theta)) = a. \end{aligned}$$

$$\begin{aligned} \text{Var}(X_j) &= \text{Var}(E[X_j|\theta]) + E[\text{Var}(X_j|\theta)] \\ &= \text{Var}(\mu(\theta)) + E\left[\frac{v(\theta)}{m_j}\right] = \frac{v}{m_j} + a. \end{aligned}$$

Selanjutnya kita dapat menurunkan persamaan normal berikut:

$$E[X_{n+1}] = \alpha_0 + \sum_{i=1}^n \alpha_i E[X_i]$$

$$\text{Cov}(X_{n+1}, X_i) = \sum_{j=1}^n \alpha_j \text{Cov}(X_i, X_j).$$

Untuk penyederhanaan kita misalkan $\sum_{j=1}^n m_j = m$.

Diperoleh

$$\mu = \alpha_0 + \sum_{i=1}^n \alpha_i \mu$$

$$\sum_{i=1}^n \alpha_i = 1 - \frac{\alpha_0}{\mu} \quad (1)$$

$$a = \sum_{j=1}^n \alpha_j a - \alpha_i a + \alpha_i \left(\frac{v}{m_i} + a \right) = \sum_{j=1}^n \alpha_j a + \frac{v \alpha_i}{m_i}, \quad i = 1, 2, \dots, n$$

$$\rightarrow \alpha_i = \frac{a m_i}{v} \left(1 - \sum_{j=1}^n \alpha_j \right)$$

Dengan menggunakan persamaan (1) diperoleh

$$\alpha_i = \frac{a m_i}{v} \cdot \frac{\alpha_0}{\mu} = \frac{\alpha_0 m_i}{k \mu} \quad (2)$$

$$\rightarrow \sum_{i=1}^n \alpha_i = \sum_{i=1}^n \frac{\alpha_0 m_i}{k \mu} = \frac{\alpha_0}{k \mu} \sum_{i=1}^n m_i = \frac{\alpha_0}{k \mu} m$$

$$\rightarrow \frac{\alpha_0}{k \mu} m = 1 - \frac{\alpha_0}{\mu}$$

$$\rightarrow \alpha_0 = \frac{k}{m+k} \mu \quad (3)$$

Dengan menggunakan (2) dan (3) diperoleh

$$\alpha_i = \frac{m_i}{m+k} \quad i = 1, 2, \dots, n$$

Telah kita ketahui bahwa *credibility premium* berbentuk

$$\alpha_0 + \sum_{j=1}^n \alpha_j X_j.$$

Dengan menyubstitusikan nilai α_0 dan α_i di atas diperoleh

$$\alpha_0 + \sum_{j=1}^n \alpha_j X_j = \frac{k}{m+k} \mu + \sum_{j=1}^n \frac{m_j}{m+k} X_j.$$

Terlihat bahwa $Z = \frac{m}{m+k}$, sehingga diperoleh *credibility premium*

$$(1 - Z)\mu + Z\bar{X},$$

dengan $\bar{X} = \sum_{j=1}^n \frac{m_j X_j}{m}$.

Contoh soal

Misalkan N_j menyatakan banyaknya klaim di tahun j , dimana terdapat m_j anggota di tahun j tersebut. Diberikan $\theta = \theta, N_j | (\theta = \theta) \sim \text{Poisson}(m_j \theta)$ dan $\theta \sim \text{Gamma}(\alpha, \beta)$.

Selanjutnya dari informasi tersebut kita dapat menentukan Buhlmann-Straub *credibility premium* dari rata rata banyaknya klaim pada tahun $n + 1$ untuk setiap anggota.

Misalkan $X_j = \frac{N_j}{m_j}$, maka

$$\mu(\Theta) = E[X_j|\Theta] = \frac{1}{m_j} E[N_j|\Theta] = \frac{1}{m_j} m_j \Theta = \Theta,$$

$$v(\Theta) = \text{Var}(X_j|\Theta) = \frac{1}{m_j^2} \text{Var}(N_j|\Theta) = \frac{1}{m_j^2} m_j \Theta = \frac{\Theta}{m_j},$$

$$\mu = E[\mu(\Theta)] = E[\Theta] = \alpha\beta,$$

$$v = m_j E[v(\Theta)] = m_j \frac{1}{m_j} E[\Theta] = \alpha\beta,$$

$$a = \text{Var}(\mu(\Theta)) = \text{Var}(\Theta) = \alpha\beta^2,$$

$$k = \frac{v}{a} = \frac{1}{\beta},$$

sehingga diperoleh *credibility factor* $Z = \frac{m}{m+k} = \frac{m\beta}{m\beta+1}$, dengan demikian

Buhlmann-Straub *Credibility* untuk model ini adalah

$$\frac{m\beta}{m\beta+1} \bar{X} + \frac{1}{m\beta+1} \alpha\beta.$$

Exact Credibility

Jika Bayesian premium sama dengan kriteria penentuan premi kredibilitas lainnya (Model Buhlmann atau Model Buhlmann-Straub) maka kita katakan memenuhi konsep *exact credibility*.

Misalkan premi

$$E[X_{n+1}|X_1, \dots, X_n],$$

$$\hat{\alpha}_0 + \sum_{i=1}^n \hat{\alpha}_i X_i.$$

Terdapat dua kondisi yang menyebabkan terjadinya *exact credibility*, yaitu :

1. Karena dalam penentuan *credibility premium* kita meminimalkan

$$Q_1 = E \left[\left(X_{n+1} - \hat{\alpha}_0 - \sum_{i=1}^n \hat{\alpha}_i X_i \right)^2 \right]$$

$$Q_2 = E \left[\left(E[X_{n+1}|X_1, \dots, X_n] - \hat{\alpha}_0 - \sum_{i=1}^n \hat{\alpha}_i X_i \right)^2 \right],$$

Jika *Bayesian premium* berupa fungsi linear maka Bayesian premium akan sama dengan credibility premium yang menyebabkan terjadinya *exact credibility*.

2. Misalkan $X_j|\theta$ peubah acak yang saling bebas $j = 1, \dots, n, n+1$ dengan sebaran peluang yang sama $F(x_j|\theta) = \frac{p(x_j)e^{-\theta x_j}}{q(\theta)}$. Jika sebaran prior $\pi(\theta) = \frac{(q(\theta))^{-k}e^{-\mu k\theta}}{c(\mu, k)}$, maka $E[X_{n+1}|X_1, \dots, X_n]$ adalah linear exponential family.

Berikut tabel yang menyebabkan terjadinya *exact credibility*

$f(x_j \theta)$	$\pi(\theta)$	$\pi(\theta x_1, \dots, x_n)$	$f(x_{n+1})$	$E[X_{n+1} X_1, \dots, X_n]$
Poisson	Gamma	Gamma	Binomial negatif	$Z\bar{X} + (1 - Z)\mu$
Exponensial	Gamma	Gamma	Pareto	$Z\bar{X} + (1 - Z)\mu$
Normal	Normal	Normal	Normal	$Z\bar{X} + (1 - Z)\mu$
Bernouli	Beta	Beta	Bernouli	$Z\bar{X} + (1 - Z)\mu$

Contoh Soal

Diberikan $\theta, X_j | (\theta = \theta) \sim \text{Binomial}(1, \theta)$ dan $\theta \sim \text{Beta}(\alpha, \beta)$.

- Tentukan Buhlmann Credibility premium
- Tentukan Bayesian premium

c. Apakah memenuhi kondisi exact credibility?

Jawab

a. Buhlmann Credibility premium

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

$$\begin{aligned} v = E[v(\theta)] &= E[Var[X_j|\theta]] = E[\theta(1 - \theta)] = E[\theta] - E[\theta^2] \\ &= E[\theta] - Var(\theta) + (E[\theta])^2 = \frac{\alpha\beta}{(\alpha + \beta)(1 + \alpha + \beta)}. \end{aligned}$$

$$a = Var(\mu(\theta)) = Var(E[X_j|\theta]) = Var(\theta) = \frac{\alpha\beta}{(\alpha + \beta)^2(1 + \alpha + \beta)}.$$

$$k = \frac{v}{a} = \alpha + \beta$$

sehingga $Z = \frac{n}{n+k} = \frac{n}{n+\alpha+\beta}$, diperoleh *Buhlmann credibility premium*

$$\begin{aligned} Z\bar{X} + (1 - Z)\mu &= \frac{n}{n + \alpha + \beta} \bar{x} + \left(\frac{\alpha + \beta}{n + \alpha + \beta} \right) \frac{\alpha}{\alpha + \beta} \\ &= \frac{n\bar{x} + \alpha}{n + \alpha + \beta}. \end{aligned}$$

b. *Bayesian premium*

Mula mula kita tentukan dahulu sebaran posterior

$$\begin{aligned} \pi(\theta|x_1, \dots, x_n) &\propto \left(\prod_{i=1}^n f(x_i|\theta) \right) \pi(\theta) \\ &= \prod_{j=1}^n \theta^{x_j} (1 - \theta)^{1-x_j} \theta^{\alpha-1} (1 - \theta)^{\beta-1} \\ &= \theta^{\alpha-1+\sum_{j=1}^n x_j} (1 - \theta)^{\beta-1+\sum_{j=1}^n 1-x_j} \end{aligned}$$

Terlihat bahwa sebaran posterior tersebut merupakan sebaran

Beta($\alpha + \sum_{j=1}^n x_j, \beta + n - \sum_{j=1}^n x_j$). Selanjutnya

$$\mu_{n+1}(\theta) = E[X_{n+1}|\Theta = \theta] = \theta,$$

sehingga diperoleh *Bayesian premium*

$$\begin{aligned} E[X_{n+1}|x_1, \dots, x_n] &= \int_0^1 \theta \pi(\theta|x_1, \dots, x_n) d\theta = E[\Theta|x_1, \dots, x_n] = \frac{\alpha^*}{\alpha^* + \beta^*} \\ &= \frac{\alpha + \sum_{j=1}^n x_j}{\alpha + \sum_{j=1}^n x_j + \beta + n - \sum_{j=1}^n x_j} = \frac{n + \alpha}{n + \alpha + \beta}, \end{aligned}$$

dengan $\bar{x} = \frac{\sum_{j=1}^n x_j}{n}$.

c. Ya, karena *Buhlmann premium* = *Bayesian premium*

Pendugaan Empiris Parameter Bayes Untuk *Credibility*

Ketika kita menyederhanakan asumsi sebaran klaim, maka metode berikut dapat kita gunakan. Kita cukup membutuhkan beberapa data dan selanjutnya kita dapat mendefinisikan strukturnya.

Misalkan terdapat r (grup) pemegang polis. Untuk pemegang polis $i = 1, 2, \dots, r$ terdapat n_i tahun dari pengalaman klaim/unit amatan.

Misalkan X_{ij} merupakan rata-rata banyaknya kehilangan/klaim untuk pemegang polis i di tahun j , dan m_{ij} merupakan banyaknya anggota untuk pemegang polis i pada tahun j . Ilustrasi ini akan lebih mudah apabila ditampilkan dalam tabel berikut:

Observed Risk (Year) unit	(Group) policyholder	1	...	i	...	r
1		x_{11}	...	x_{1i}	...	x_{1r}
2		x_{12}	...	x_{2i}	...	x_{2r}
$\vdots$		$\vdots$	$\ddots$	$\ddots$	$\ddots$	$\vdots$
n_i		x_{1n_i}	...	x_{in_i}	...	x_{1r}

Secara serupa, tabel ini dapat juga dibentuk sebagai berikut:

Observed Risk Policy holder (Group)	Unit (Year)	1	...	i	...	j
1		x_{11}	...	x_{1i}		x_{1j}
$\vdots$		$\vdots$	...	...		
i		x_{i1}	...	x_{ii}		x_{ij}
$\vdots$		$\vdots$	...			
r		x_{r1}	...	x_{ri}		x_{rj}

Kita perlu mendefinisikan beberapa notasi untuk membuat perhitungan lebih jelas. Misalkan vektor klaim untuk rata-rata banyaknya kehilangan/klaim untuk pemegang polis i sepanjang tahun,

$$\underline{X}_i = (x_{i1}, \dots, x_{in_i}).$$

Vektor yang menyatakan banyaknya anggota untuk pemegang polis i dinotasikan dengan

$$\underline{m}_i = (m_{i1}, \dots, m_{in_i}).$$

Selanjutnya kita memiliki hubungan hubungan berikut:

$$m_i = \sum_{j=1}^{n_i} m_{ij}$$

$$m = \sum_{i=1}^r m_i = \sum_{i=1}^r \sum_{j=1}^{n_i} m_{ij}$$

$$\bar{X}_i = \frac{1}{m_i} \sum_{j=1}^{n_i} m_{ij} x_{ij}$$

$$\bar{X} = \frac{1}{m} \sum_{i=1}^r m_i \bar{X}_i = \frac{1}{m} \sum_{i=1}^r \sum_{j=1}^{n_i} m_{ij} x_{ij}$$

Dalam hal ini, kita mengasumsikan bahwa $X_1, \dots, X_r$ adalah saling bebas. Alasan ini cukup *reasonable* karena untuk grup klaim yang berbeda akan saling bebas satu sama lain. Untuk setiap i , diberikan $\theta_i, X_{i1}, \dots, X_{in_i}$, yang saling bebas. θ_i merupakan peubah acak yang bebas dan identik, $i = 1, 2, \dots, r$. Secara serupa pada konsep credibility sebelumnya kita dapat mendefinisikan “Parameter strukur” μ, v , dan a .

Tujuan kita dari menentukan μ, v , dan a adalah untuk menentukan Buhlmann credibility dan Buhlmann-Straub credibility dari data yang diberikan.

Parameter Struktur

Dari contoh $(Y_1, \dots, Y_n)$ kita peroleh nilai amatan $(y_1, \dots, y_n)$. Dari data ini kita dapat mengkontruksi penduga $\hat{\theta}(Y_1, \dots, Y_n)$ dari $\hat{\theta}(y_1, \dots, y_n)$, sedemikian sehingga $\hat{\theta}$ takbias. Jika menambahkan asumsi tentang $Y_i, i = 1, \dots, n$ selanjutnya kita dapat memperoleh hasil tambahan berikut:

- (a) Jika $Y_1, \dots, Y_n$ saling bebas dan memiliki nilai harapan yang sama $\mu = E[Y_j]$ dan ragam yang sama $\sigma^2 = Var(Y_j)$ untuk $j = 1, 2, \dots, n$, maka :

$$E[\bar{Y}] = \mu$$

$$E\left[\frac{1}{n-1}\sum_{j=1}^n(Y_j - \bar{Y})^2\right] = \frac{n}{n-1}\sigma^2.$$

- (b) Jika diberikan θ , $Y_1, \dots, Y_n$ saling bebas dan memiliki nilai harapan bersyarat $E[Y_j|\theta] = \mu(\theta)$ dan ragam bersyarat yang sama $Var(Y_j|\theta) = \sigma^2(\theta)$, maka nilai harapan bersyarat dan ragam bagi $\bar{Y}$:

$$E[\bar{Y}|\theta] = \mu(\theta)$$

$$E\left[\frac{1}{n-1}\sum_{j=1}^n(y_j - \bar{Y})^2 \mid \theta\right] = \sigma^2(\theta)$$

Pendugaan Non-Parametrik untuk Model Buhlmann

- Asumsikan $m_{ij} = 1, n_i = n > 1, i = 1, \dots, r, m = nr$.

Selanjutnya kita peroleh hasil berikut:

$$\bar{X}_i = \frac{1}{n}\sum_{j=1}^n x_{ij}$$

$$\bar{X} = \frac{1}{rn}\sum_{i=1}^r\sum_{j=1}^n X_{ij}$$

- Diberikan $\theta_i, X_{i1}, \dots, X_{in}$ yang saling bebas, $i = 1, \dots, r$, sedangkan $\underline{X}_1, \dots, \underline{X}_r$ takbebas bersyarat, dimana $\underline{X}_i = (X_{i1}, \dots, X_{in})$.

- Karena $\theta_1, \theta_2, \dots, \theta_r$ menyebar bebas dan identik.

$$\mu(\theta_i) = E[X_{ij}|\theta_i]$$

$$v(\theta_i) = \text{Var}(X_{ij}|\theta_i)$$

$$\mu = E[\mu(\theta_i)]$$

$$v = E[v(\theta_i)]$$

$$a_i = \text{Var}(\mu(\theta_i)) \quad i = 1, 2, \dots, r.$$

- Kita dapat memperoleh penduga yang takbias untuk parameter μ , v , dan a .

1. $\hat{\mu} = \bar{X}$ takbias untuk μ

$$E[\hat{\mu}] = \frac{1}{nr} \sum_{i=1}^r \sum_{j=1}^n E[E[X_{ij}|\theta_i]] = \frac{1}{nr} \sum_{i=1}^r \sum_{j=1}^n \mu = \mu.$$

Catatan : $\bar{X}_i$ juga takbias untuk μ .

2. Diberikan i dan $\theta_i, X_{i1}, \dots, X_{in}$ yang saling bebas $\mu(\theta_i)$ dan ragam bersyarat $v(\theta_i)$

$$\hat{v}_i = \frac{1}{n-1} \sum_{j=1}^n (X_{ij} - \bar{X}_i)^2$$

$$E[\hat{v}_i|\theta_i] = v(\theta_i)$$

$$\rightarrow E[\hat{v}_i] = E[v(\theta_i)] = v,$$

Selanjutnya $\hat{v}_i$ takbias untuk v , $i = 1, 2, \dots, r$. Catatan bahwa:

$$\hat{v} = \frac{1}{r} \sum_{i=1}^r \hat{v}_i$$

tak bias untuk v .

Misalkan untuk pendugaan non-parametrik model Buhlmann kita memiliki informasi sebagai berikut:

1. $\hat{\mu} = \bar{X}$
2. $\hat{v} = \frac{1}{r} \sum_{i=1}^r \hat{v}_i$, di mana $\hat{v}_i = \frac{1}{n-1} \sum_{j=1}^n (X_{ij} - \bar{X}_i)^2$
3. $E[\bar{X}_i | \theta_i] = \frac{1}{n} \sum_{j=1}^k \mu(\theta_j)$

Sehingga $E[\bar{X}_i] = \frac{1}{n} \sum_{j=1}^k E[\mu(\theta_j)] = \mu$.

$$\begin{aligned} \text{Var}(\bar{X}_i) &= \text{Var}(E[\bar{X}_i | \theta_i]) + E[\text{Var}[\bar{X}_i | \theta_i]] \\ &= \text{Var}(\mu(\theta_i)) + E\left[\frac{v[\theta_i]}{n}\right] = a + \frac{v}{n}. \end{aligned}$$

4. Jika kita lihat $Y_1 = \bar{X}_1, \dots, Y_r = \bar{X}_r$, kita tahu bahwa kehilangan antar grup saling bebas berdasarkan asumsi kita sebelumnya, sehingga $\bar{X}_1, \dots, \bar{X}_r$ saling bebas.

Selanjutnya kita menginginkan sebuah penduga tak bias untuk a , sehingga kita akan melakukan beberapa manipulasi untuk mendapatkannya. Berdasarkan sebelumnya kita telah membuktikan bahwa $\frac{1}{r-1} \sum_{i=1}^r (\bar{X}_i - \bar{X})^2$ dibuktikan sebagai penduga tak bias bagi $\text{Var}(\bar{X}_i) = a + \frac{v}{n}$. Jadi sekarang kita punya

$$E\left[\frac{1}{r-1} \sum_{i=1}^r (\bar{X}_i - \bar{X})^2\right] = a + \frac{v}{n},$$

Diperoleh

$$\hat{a} = \frac{1}{r-1} \sum_{i=1}^r (\bar{X}_i - \bar{X})^2 - \frac{\hat{v}}{n}.$$

Telah kita ketahui bahwa Buhlman credibility premium untuk pemegang polis i dirumuskan:

$$Z\bar{X}_i + (1 - Z)\mu,$$

Di mana $Z = \frac{n}{n+k}$ dan $k = \frac{v}{a}$.

Sekarang dugaan Buhlmann premium untuk pemegang polis i adalah:

$$\hat{Z}\bar{X}_i + (1 - \hat{Z})\mu,$$

Dimana $\hat{Z} = \frac{n}{n+k}$ dan $\hat{k} = \frac{v}{a}$.

Perlu dicatat bahwa $\hat{Z}$ dan $\hat{k}$ tidak selalu merupakan penduga yang tak bias. Beberapa kasus jika $\hat{a} \leq 0$ maka partial credibility akan terlihat aneh, sehingga dalam kasus ini nilai $\hat{Z} = 0$.

Contoh Soal

Dalam model Buhlmann, 2 pemegang polis dengan 3 tahun pengalaman memiliki rekam kehilangan sebagai berikut:

Observed Risk	(Group) policyholders	1	2
(Year) Unit			
1		3	6
2		5	12
3		7	9
$\bar{X}_i$		5	9

Kita punya $r = 2$ dan $n = 3$, kita dapat menentukan dengan mudah nilai nilai berikut:

$$\bar{X}_1 = 5$$

$$\bar{X}_2 = 9$$

$$\hat{\mu} = \bar{X} = \frac{5+9}{2} = 7$$

$$\hat{v}_1 = \frac{1}{3-1} [(3-5)^2 + (5-5)^2 + (7-5)^2] = 4$$

$$\hat{v}_2 = \frac{1}{3-1} [(6-9)^2 + (12-9)^2 + (9-9)^2] = 9$$

$$\hat{v} = \frac{4+9}{2} = \frac{13}{2}$$

$$\hat{a} = \frac{1}{r-1} \sum_{i=1}^r (\bar{X}_i - \bar{X})^2 - \frac{\hat{v}}{n} = \frac{1}{2-1} \sum_{i=1}^2 (\bar{X}_i - 7)^2 - \frac{\frac{13}{2}}{3} = \frac{35}{36}$$

$$\hat{k} = \frac{\hat{v}}{\hat{a}} = \frac{\frac{13}{2}}{\frac{35}{36}} = \frac{39}{35}$$

$$\hat{Z} = \frac{n}{n + \hat{k}} = \frac{3}{3 + \frac{39}{35}} = \frac{35}{48},$$

Jadi dugaan Bulhmann credibility premium untuk pemegang polis 1 di tahun 4 adalah :

$$\hat{Z}\bar{X}_1 + (1 - \hat{Z})\mu = \frac{133}{24}.$$

dugaan Bulhmann credibility premium untuk pemegang polis 2 di tahun 4 adalah :

$$\hat{Z}\bar{X}_2 + (1 - \hat{Z})\mu = \frac{203}{24}.$$

PENGHITUNGAN KREDIBILITAS DENGAN PUSTAKA ACTUAR DALAM R

I. MAULIDI¹, W. ERLIANA², A.D. GARNADI², S. NURDIATI², I G. P. PURNABA²

Abstrak

Teori Kredibilitas (*Credibility Theory*) merupakan perangkat penting dalam pekerjaan aktuaria. Dengan menggunakan Kredibilitas dapat diduga besarnya pembayaran premi atau banyaknya klaim yang akan terjadi di masa mendatang secara kredibel. Dalam tulisan ini akan diperkenalkan konsep dalam teori kredibilitas dan aplikasinya dengan menggunakan paket Actuar yang ditulis menggunakan software R.

PENDAHULUAN

Kredibilitas adalah sebuah ukuran penilaian seorang aktuaris terhadap dugaan premi di masa mendatang. Data di masa lampau tidak sepenuhnya dapat digunakan untuk menduga di masa mendatang, sehingga revisi dari penduga pun sering dilakukan. Selain itu risiko untuk setiap kelompok individu sangat dimungkinkan berbeda beda. Oleh karena itu kita perlu mengembangkan suatu analisis dari data di masa lampau untuk menduga data di masa mendatang. Sering kali data di masa lampau tidak stabil apabila digunakan untuk memprediksi premi di masa mendatang, sehingga revisi pendugaan sering kali dilakukan. Oleh karena itu diperlukan suatu tool untuk menggunakan data di masa lampau sehingga dapat memprediksi besarnya tingkat premi dan banyaknya klaim di masa mendatang.

Misalkan $X_1, X_2, \dots, X_n$ merupakan besarnya kehilangan di n tahun yang lalu atau n pemegang polis yang lalu. Secara umum ini merupakan n unit exposure yang diamati.

Asumsi

$$\begin{aligned}E[X_j] &= \xi, \\ \text{Var}(X_j) &= \sigma^2, \\ \text{Var}(\bar{X}) &= \frac{\sigma^2}{n},\end{aligned}$$

untuk $j=1, \dots, n$ dan

¹ Program Studi Matematika Universitas Syiah Kuala, Banda Aceh. Email: ikhsanmath47@gmail.com.

² Departemen Matematika, Fakultas Ilmu Matematika dan Pengetahuan Alam, Jalan Meranti Kampus IPB, Dramaga Bogor, 16680.

$$\bar{X} = \frac{1}{n} \sum_{j=1}^n X_j$$

Ada pun untuk menyatakan kredibilitas untuk X kita harus mendefinisikan error relatif terlebih dahulu. Untuk suatu nilai peluang $p, p \in R, 0 \leq p \leq 1$ dan untuk tingkat error ε , jika dipenuhi

$$P\left(\left|\frac{\bar{X} - \xi}{\xi}\right| \leq \varepsilon\right) \geq p,$$

maka kita katakan X adalah credible untuk ξ .

BEBERAPA MODEL KREDIBILITAS

Model Buhlmann

Model ini merupakan model kredibilitas yang paling sederhana. Pada masing-masing pemegang polis (bergantung pada Θ), besarnya klaim dari masa lalu $X_1, \dots, X_n$ memiliki rata-rata dan ragam yang sama dan bergantung pada Θ .

Didefinisikan

$$\mu(\theta) = E(X_j | \Theta = \theta),$$

dan

$$v(\theta) = \text{Var}(X_j | \Theta = \theta).$$

Dalam hal ini $\mu(\theta)$ kita disebut sebagai *hypothetical mean* dan $v(\theta)$ disebut *process variance*.

Selanjutnya kita dapat menentukan rata-rata, ragam dan ragam X_j sebagai berikut:

$$E(X_j) = E[E(X_j | \Theta)] = E[\mu(\Theta)] = \mu,$$

$$\begin{aligned} \text{Var}(X_j) &= E[\text{var}(X_j | \Theta) + \text{var}[E(X_j | \Theta)]] = \mu \\ &= E(v(\Theta)) + \text{var}(\mu(\Theta)) \\ &= v + a, \end{aligned}$$

$$\begin{aligned} \text{Cov}(X_i, X_j) &= E(X_i X_j) - E(X_i)E(X_j), i \neq j \\ &= E[E(X_i X_j | \Theta)] - \mu^2 \\ &= E(E(X_i | \Theta) \cdot E(X_j | \Theta)) - [E(\mu(\Theta))]^2 \\ &= E((\mu(\Theta))^2) - [E(\mu(\Theta))]^2 \\ &= \text{Var}(\mu(\Theta)) = a. \end{aligned}$$

Model pedugaan premi kredibilitas Buhlmann dihipotesiskan dengan suatu fungsi linear berikut:

$$\alpha_o + \sum_{j=1}^n \alpha_j X_j = Z \bar{X} + (1 - Z) \mu,$$

dengan $Z = \frac{n}{k+n}$, $k = \frac{v}{a} = \frac{E[\text{var}(X_j|\theta)]}{\text{var}[E(X_j)|\theta]}$. Penurunan formula ini dapat dilihat pada Klugman (2012).

Model Buhlmann Straub

Model kredibilitas Buhlmann-Straub adalah pengembangan dari model kredibilitas Buhlmann dengan beberapa asumsi tambahan. Diberikan $\theta, X_1, \dots, X_{n+1}$ saling bebas untuk $j = 1, 2, \dots, n+1$, misalkan pula m_j menyatakan banyaknya individu pada tahun ke- j maka berlaku:

$$E[X_j|\theta] = \mu(\theta),$$

$$\text{Var}(X_j|\theta) = \frac{v(\theta)}{m_j}.$$

Selanjutnya kita notasikan sebagai berikut:

$$\mu = E[\mu(\theta)], a = \text{Var}(\mu(\theta))$$

$$v = E[v(\theta)], k = \frac{v}{a}.$$

Kuantitas-kuantitas di atas dapat digunakan untuk menentukan premi kredibilitas.

Premi kredibilitas untuk model Buhlmann-Straub ini adalah

$$(1 - Z)\mu + Z\bar{X},$$

dengan $\bar{X} = \sum_{j=1}^n \frac{m_j X_j}{m}$, $Z = \frac{m}{m+k}$, dan $\sum_{j=1}^n m_j = m$.

BEBERAPA CONTOH DAN APLIKASI

Contoh 1

Misalkan banyaknya klaim tahunan untuk seorang individu yang diasuransikan memiliki sebaran Poisson. Nilai harapan frekuensi klaim tahunan (parameter Poisson Δ) dari anggota populasi yang diasuransikan menyebar seragam (0,1). Rataan frekuensi klaim tahunan individu adalah konstan sepanjang waktu. Tentukan faktor kredibilitas Buhlman Z jika ada n amatan.

Jawab:

Hypothetical mean dapat ditentukan sebagai berikut:

$$\mu(\Delta) = E(X|\Delta) = \Delta,$$

hal ini karena $X|\Delta$ menyebar Poisson(Δ). Proses

variance dapat ditentukan sebagai berikut:

$$V(\Delta) = \text{var}(X|\Delta) = \Delta$$

sehinggadiperolehkuantitas:

$$\mu = E(X) = E(E(X|\Delta)) = E(\mu(\Delta)) = E(\Delta) = 0.5$$

$$v = E(V(\Delta)) = E(\Delta) = 0.5$$

$$\alpha = \text{var}(\mu(\Delta)) = \text{var}(\Delta) = \frac{1}{12}$$

sehinggadiperolehfaktorkredibilitasBuhlmann

$$Z = \frac{n}{n + \frac{1/2}{1/12}} = \frac{n}{n + 6}.$$

Contoh 2

Misalkan N_j menyatakan banyaknya klaim di tahun j , dimana terdapat m_j anggota di tahun j tersebut. Diberikan $\theta = \theta, N_j | (\theta = \theta) \sim \text{Poisson}(m_j \theta)$ dan $\theta \sim \text{Gamma}(\alpha, \beta)$. Selanjutnya dari informasi tersebut kita dapat menentukan Buhlmann-Straub *credibility premium* dari rata rata banyaknya klaim pada tahun $n + 1$ untuk setiap anggota.

Jawab:

Misalkan $X_j = \frac{N_j}{m_j}$, maka

$$\mu(\theta) = E[X_j | \theta] = \frac{1}{m_j} E[N_j | \theta] = \frac{1}{m_j} m_j \theta = \theta,$$

$$v(\theta) = \text{Var}(X_j | \theta) = \frac{1}{m_j^2} \text{Var}(N_j | \theta) = \frac{1}{m_j^2} m_j \theta = \frac{\theta}{m_j},$$

$$\mu = E[\mu(\theta)] = E[\theta] = \alpha\beta,$$

$$v = m_j E[v(\theta)] = m_j \frac{1}{m_j} E[\theta] = \alpha\beta,$$

$$a = \text{Var}(\mu(\theta)) = \text{Var}(\theta) = \alpha\beta^2,$$

$$k = \frac{v}{a} = \frac{1}{\beta},$$

sehinggadiperoleh *credibility factor* $Z = \frac{m}{m+k} = \frac{m\beta}{m\beta+1}$, dengan demikian Buhlmann-Straub *Credibility* untuk model ini adalah

$$\frac{m\beta}{m\beta+1} \bar{X} + \frac{1}{m\beta+1} \alpha\beta.$$

PENGUNAAN PACKAGES “ACTUAR” UNTUK KREDIBILITAS

Simulasi Data Hachemeister

Berikut ini kami sajikan sebuah contoh simulasi menentukan besarnya premi dengan menggunakan model kredibilitas. Simulasi dilakukan dengan menggunakan software R dan packages "actuar" yang diperoleh dari CRAN. Data yang digunakan dalam simulasi adalah data dari Hachemesiter (1975) yang memuat besarnya klaim dan banyaknya klaim dari pemegang polis asuransi. Data ini terdapat dalam packages berupa matriks dengan banyaknya baris 5 dan banyaknya kolom adalah 25. Kolom 1 merepresentasikan kelompok dari pemegang polis, kolom 2-13 merepresentasikan rata-rata besarnya klaim setiap periode dan kolom 14-25 merepresentasikan banyaknya klaim setiap periode.

```
> data(hachemeister)
> hachemeister
      state ratio.1 ratio.2 ratio.3 ratio.4 ratio.5 ratio.6 ratio.7 ratio.8
[1,]      1  1738   1642   1794   2051   2079   2234   2032   2035
[2,]      2  1364   1408   1597   1444   1342   1675   1470   1448
[3,]      3  1759   1685   1479   1763   1674   2103   1502   1622
[4,]      4  1223   1146   1010   1257   1426   1532   1953   1123
[5,]      5  1456   1499   1609   1741   1482   1572   1606   1735
      ratio.9 ratio.10 ratio.11 ratio.12 weight.1 weight.2 weight.3 weight.4
[1,]   2115    2262    2267    2517    7861    9251    8706    8575
[2,]   1464    1831    1612    1471    1622    1742    1523    1515
[3,]   1828    2155    2233    2059    1147    1357    1329    1204
[4,]   1343    1243    1762    1306    407    396    348    341
[5,]   1607    1573    1613    1690    2902    3172    3046    3068
      weight.5 weight.6 weight.7 weight.8 weight.9 weight.10 weight.11 weight.12
[1,]   7917    8263    9456    8003    7365    7832    7849    9077
[2,]   1622    1602    1964    1515    1527    1748    1654    1861
[3,]    998    1077    1277    1218    896    1003    1108    1121
[4,]    315    328    352    331    287    384    321    342
[5,]   2693    2910    3275    2697    2663    3017    3242    3425
> |
```

Fungsi yang digunakan untuk model kredibilitas adalah fungsi *icm*. Untuk model Buhlmann tidak mempertimbangkan adanya pembobotan, premi dapat ditentukan sebagai berikut:

```
> cm(~state, hachemeister, ratios=ratio.1:ratio.12)
Call:
cm(formula = ~state, data = hachemeister, ratios = ratio.1:ratio.12)

Structure Parameters Estimators

Collective premium: 1671.017

Between state variance: 72310.02
Within state variance: 46040.47
```

Berikut hasil simulasi untuk model Buhlmann-Straub (memperhitungkan adanya pembobotan) yang menggunakan model Bichsel Straub:

```
> cm(~state, hachemeister, ratios=ratio.1:ratio.12, weights=weight.1:weight.12)
Call:
cm(formula = ~state, data = hachemeister, ratios = ratio.1:ratio.12,
  weights = weight.1:weight.12)

Structure Parameters Estimators

Collective premium: 1683.713

Between state variance: 89638.73
Within state variance: 139120026
```

Dari simulasi di atas diperoleh nilai premi kredibilitas untuk model Buhlmann adalah 1671.017, sedangkan untuk model Buhlmann-Straub diperoleh premi kredibilitas sebesar 1683.713.

Simulasi Contoh 1

Untuk menguji hasil dari contoh 1, kami melakukan simulasi dengan membangkitkan data besarnya klaim. Data tersebut kami sajikan sebagai berikut:

```
> simulasi
periode bulan.1 bulan.2 bulan.3 bulan.4 bulan.5 bulan.6
1      1 0.6376503 0.6033292 0.06649485 0.06783176 0.0809631 0.07325076
2      2 0.9284553 0.8202483 0.50642640 0.80434591 0.7779573 0.79099336
```

Dari data tersebut diketahui $\bar{X} = 0.513$. Dengan menggunakan formula Z dalam contoh 1, diperoleh

$$Z = \frac{n}{n+6} = \frac{12}{12+6} = 0.67,$$

diperoleh hasil dugaan premi Buhlmann secara terpisah sebagai berikut:

$$(1-Z)\mu + Z\bar{X} = 0.5088.$$

Dengan menggunakan perintah *cm* dalam package *actuar* diperoleh hasil berikut:

```
> cm(~periode, simulasi, ratios=bulan.1:bulan.6)
Call:
cm(formula = ~periode, data = simulasi, ratios = bulan.1:bulan.6)

Structure Parameters Estimators

Collective premium: 0.5131622

Between periode variance: 0.1250378
Within periode variance: 0.0500418
```

Terlihat bahwa hasil dugaan premi yang diperoleh adalah 0.5132.

UCAPAN TERIMA KASIH

Penelitian ini didanai oleh PUPT-IPB *under contract* no: 079/SP2H/LT/DRPM/I I/2016.

DAFTAR PUSTAKA

- [1] Hachemeister CA. 1975. Credibility for regression models with application to trend. In *credibility, theory and applications*, Proceeding of the Berkeley Actuarial Research Conference on Credibility, pages 129-163. New York: Academic Press.
- [2] Buhlmann H, Gisler A. 2005. *A course in credibility theory and its applications*. Springer, ISBN 3-5402575-3-5.
- [3] Buhlmann H, Straub E. 1970. Glaubwürdigkeit für Schadensatz. *Bulletin of the Swiss Association of Actuaries*, 70: 111-133.
- [4] Klugman SA, et al. 2012. *Loss Models: From Data to Decisions*. Wiley, ISBN 1118315324.