BAB II

TINJAUAN PUSTAKA

2.1 Pendahuluan

Pada bab ini akan diuraikan mengenai beberapa teori dan metode yang mendukung serta mempermudah dalam melakukan perhitungan dan dapat membantu di dalam pembahasan bab-bab selanjutnya. Adapun teori dan metode tersebut yaitu asuransi kendaraan bermotor, metode penaksiran, distribusi untuk frekuensi klaim, distribusi untuk besar klaim, uji kecocokan distribusi, metode penaksiran *Linear Empirical Bayesian*, taksiran premi untuk asuransi kendaraan bermotor, dan transformasi Box-Cox.

2.2 Asuransi Kendaraan Bermotor

Dalam setiap kegiatan yang dilakukan oleh suatu kelompok atau perorangan pasti ada risiko yang harus ditanggung. Risiko merupakan kemungkinan terjadinya suatu kerugian yang tidak diduga atau tidak diinginkan. Ketidakpastian akan terjadinya risiko menjadi hal yang perlu diperhatikan dalam melakukan suatu kegiatan, karena hidup penuh dengan ketidakpastian dan manusia selalu berusaha memperkecil atau meminimumkan ketidakpastian tersebut. Risiko dapat terjadi dimanapun dan kapanpun, dimana jika risiko terjadi maka akan mengakibatkan kerugian. Salah satu upaya untuk menghindari risiko yaitu dengan memindahkan risiko kepada pihak lain. Asuransi merupakan bentuk perlindungan yang dapat menanggung kerugian dengan mengganti kerugian tersebut dengan cara klaim. Asuransi merupakan langkah yang tepat untuk mengalihkan kerugian tersebut dengan membuat perlindungan yang diatur dalam polis. Salah satu produk asuransi yang paling banyak diminati di Indonesia adalah asuransi kendaraan bermotor.

Asuransi kendaraan bermotor adalah produk asuransi kerugian yang melindungi tertanggung dari risiko kerugian yang mungkin timbul sehubungan dengan kepemilikan dan pemakaian kendaraan bermotor. Ada dua jenis perlindungan untuk asuransi kendaraan bermotor, yaitu *Total Loss Only* (TLO) dan *Comprehensive* (Komprehensif).

a. *Total Loss Only* (TLO)

Jaminan ganti rugi atas kehilangan atau kerusakan total pada kendaraan akibat dari kejatuhan benda, kebakaran, perbuatan jahat, pencurian, perampasan, tabrakan, benturan atau kecelakaan lalu lintas lainnya. Dalam produk asuransi TLO, jenis klaim yang diajukan adalah *total loss*.

b. *Comprehensive* (Komprehensif)

Jaminan ganti rugi atau biaya perbaikan atas kehilangan atau kerusakan sebagian maupun keseluruhan pada kendaraan akibat kejatuhan benda, kebakaran, perbuatan jahat, pencurian, perampasan, tabrakan, benturan atau kecelakaan lalu lintas lainnya. Dalam produk asuransi *Comprehensive,* jenis klaim yang diajukan ada dua kemungkinan yaitu, *total loss* dan *partial loss.*

Perlu diketahui bahwa klaim *total loss* hanya dapat diajukan sekali oleh tertanggung ke penanggung. Sedangkan klaim *partial loss* dapat diajukan lebih dari sekali oleh tertanggung ke penanggung.

Peraturan Ketua Bapepam dan Lembaga keuangan (LK) pada tahun 2011 membagi kategori kendaraan bermotor menjadi delapan kategori jenis kendaraan seperti tertera pada Tabel 2.1.

Tabel 2.1 Kategori Kendaraan Bermotor

Kategori	Harga Pertanggungan
(1)	(2)
Jenis Kendaraan Non Bus dan Non Truk	
Kategori 1	0 s.d Rp. 150.000.000
Kategori 2	Rp. 150.000.001,00 s.d. Rp. 300.000.000,00
Kategori 3	Rp. 300.000.001,00 s.d. Rp. 500.000.000,00
Kategori 4	Rp. 500.000.001,00 s.d. Rp. 800.000.000,00
Kategori 5	Lebih dari Rp.800.000.000,00
Jenis Kendaraan Bus dan Truk	
Kategori 6	Truk, untuk semua uang pertanggungan
Kategori 7	Bus, untuk semua uang pertanggungan
Jenis Kendaraan Roda 2 (dua)	
Kategori 8	Semua uang pertanggungan

Sumber: Peraturan Ketua Bapepam dan LK 2011

Harga pertanggungan adalah harga kendaraan saat baru atau taksiran harga kendaraan apabila diteliti pada saat pertanggungan dimulai dengan kondisi yang sama.

2.3 Metode Penaksiran

Metode penaksiran parameter secara umum dibagi dua yaitu metode penaksiran klasik dan metode penaksiran Bayes (Walpole dan Myers, 1995). Dalam metode klasik, pengambilan kesimpulan didasarkan sepenuhnya pada informasi yang diperoleh melalui sampel yang diambil dari populasi. Metode penaksiran klasik diantaranya metode penaksiran momen dan penaksiran kemungkinan maksimum. Salah satu dari metode penaksiran klasik yang akan dibahas dalam skripsi ini adalah penaksiran kemungkinan maksimum. Sedangkan metode Bayes menggunakan pengetahuan subjektif terdahulu mengenai distribusi peluang parameter yang tidak diketahui bersama dengan informasi yang diberikan oleh data sampel.

2.3.1 Metode Penaksiran Kemungkinan Maksimum

Metode penaksiran kemungkinan maksimum merupakan salah satu metode terpenting pada penaksiran. Dalam menentukan penaksir parameter dengan menggunakan metode kemungkinan maksimum harus diketahui fungsi densitas peluangnya untuk kasus kontinu atau fungsi massa peluangnya untuk kasus diskrit. Pada uraian berikutnya hanya akan digunakan istilah fungsi peluang. Jika $X_1, X_2, \ldots, X_n$ adalah sampel acak berukuran n dari distribusi dengan fungsi peluang $f(x;\theta)$ dan nilai dari sampel acak tersebut adalah $x_1, x_2, \ldots, x_n$, maka fungsi kemungkinannya adalah:

$$L(\theta) = \prod_{i=1}^{n} f(x_i;\theta).$$

Jika $\hat{\theta}$ merupakan nilai taksiran dari parameter θ (jumlah parameternya mungkin saja lebih dari satu) yang memaksimumkan fungsi kemungkinan $L(\theta)$, maka $\hat{\theta} = u(x_1, x_2, \ldots, x_n)$ merupakan penaksir kemungkinan maksimum untuk θ.

Memaksimumkan fungsi kemungkinan $L(\theta)$ pada persamaan di atas adalah sama saja dengan memaksimumkan logaritma natural dari $L(\theta)$, $\ln L(\theta)$. Biasanya lebih mudah menggunakan $\ln L(\theta)$ daripada menggunakan $L(\theta)$ untuk memperoleh penaksir kemungkinan maksimum. Salah satu cara untuk memperoleh penaksir kemungkinan maksimumnya adalah dengan menggunakan turunan (diferensial). Taksiran parameternya adalah nilai θ yang merupakan solusi dari turunan pertama berikut:

$$\frac{\partial \ln L(\theta)}{\partial \theta} = 0.$$

Berdasarkan persamaan di atas akan diperoleh satu atau beberapa persamaan parameter yang ditaksir. Kemudian dari persamaan-persamaan tersebut dieliminasi dan atau disubstitusi untuk mendapatkan taksiran parameter. Turunan kedua pada saat $\theta = \hat{\theta}$ adalah:

$$\frac{\partial^2 \ln L(\theta)}{\partial \theta^2}\Big|_{\theta=\hat{\theta}} < 0.$$

2.3.2 Metode Penaksiran Bayes

Pandanglah suatu masalah mencari taksiran titik parameter θ menggunakan teori penaksiran Bayes dari populasi dengan distribusi peluang $f(x|\theta), \theta \in \Omega$. Misalkan diambil sampel acak dari populasi tersebut berukuran *n*, yaitu $X_1, X_2, \dots, X_n$. Dengan demikian fungsi peluang bersama peubah acak $X_1, X_2, \dots, X_n$ adalah:

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

$$= \prod_{i=1}^{n} f(x_i|\theta) \qquad \dots (2.1)$$

Misalkan θ diketahui berubah sesuai dengan distribusi peluang $h(\theta)$. Dengan kata lain dianggap bahwa θ merupakan suatu nilai peubah acak Θ dengan distribusi peluang $h(\theta)$ dan ingin ditaksir nilai θ tertentu untuk populasi yang diambil sampelnya. $h(\theta)$ didefinisikan sebagai distribusi prior untuk parameter peubah acak Θ. Teknik penaksiran Bayes menggunakan distribusi prior $h(\theta)$ dan distribusi bersama sampel, seperti pada Persamaan (2.1), untuk menghitung distribusi posterior $f(\theta|x_1, x_2, \dots, x_n)$. Distribusi bersama sampel $X_1, X_2, \dots, X_n$ dan peubah acak Θ adalah

$$f(x_1, x_2, \dots, x_n; \theta) = f(x_1, x_2, \dots, x_n|\theta)h(\theta) \qquad \dots (2.2)$$

Berdasarkan pada Persamaan (2.2), dapat diperoleh distribusi marginal dari $X_1, X_2, \dots, X_n$, yaitu:

$$g(x_1, x_2, \dots, x_n) = \int_{-\infty}^{\infty} f(x_1, x_2, \dots, x_n; \theta)d\,\theta$$

$$= \int_{-\infty}^{\infty} f(x_1, x_2, \dots, x_n|\theta)h(\theta)d\,\theta$$

Dengan menggunakan distribusi peluang bersyarat menurut Walpole dan Myers (1995), maka distribusi posteriornya dapat ditulis sebagai berikut:

$$f(\theta|x_1, x_2, \dots, x_n) = \frac{f(x_1, x_2, \dots, x_n; \theta)}{g(x_1, x_2, \dots, x_n)}$$

Adapun prinsip dari penaksiran parameter melalui metode Bayes adalah mencari suatu fungsi dari $X_1, X_2, \dots, X_n$ misalkan saja $w(X_1, X_2, \dots, X_n)$, sedemikian sehingga meminimumkan ekspektasi bersyarat dari fungsi rugi berikut:

$$E\{L[\theta, w(X_1, X_2, \dots, X_n)]|X_1, X_2, \dots, X_n\}$$

$$= \int_{-\infty}^{\infty} L[\theta, w(X_1, X_2, \dots, X_n)]\, f(\theta|X_1, X_2, \dots, X_n)d\theta \qquad \dots (2.3)$$

Dengan demikian penaksir parameter bergantung pada bentuk fungsi ruginya. Untuk fungsi rugi yang berbeda akan menghasilkan penaksir $w(X_1, X_2, \dots, X_n)$ yang berbeda pula. Dalam penaksiran Bayes ada beberapa fungsi rugi, diantaranya adalah:

1. Fungsi rugi kuadratik, dengan bentuk fungsinya

$$L[\theta, w(X_1, X_2, \dots, X_n)] = [\theta - w(X_1, X_2, \dots, X_n)]^2$$

2. Fungsi rugi absolut, dengan bentuk fungsinya

$$L[\theta, w(X_1, X_2, \dots, X_n)] = |\theta - w(X_1, X_2, \dots, X_n)|$$

3. Fungi rugi *zero-one*, dengan bentuk fungsinya

$$L[\theta, w(X_1, X_2, \dots, X_n)] = \begin{cases} 0 & \text{, jika } w(X_1, X_2, \dots, X_n) = \theta \\ 1 & \text{, lainnya} \end{cases}$$

Dalam bagian ini penaksiran Bayes yang akan dibahas didasarkan pada fungsi rugi kuadratik. Jika fungsi ruginya kuadratik, maka Persamaan (2.3) akan menjadi

$E\{L[\theta, w(X_1, X_2, \dots, X_n)]|X_1, X_2, \dots, X_n\}$

$$= \int_{-\infty}^{\infty} [\theta - w(X_1, X_2, \dots, X_n)]^2 f(\theta|X_1, X_2, \dots, X_n) d\theta$$

$$= \int_{-\infty}^{\infty} [\theta^2 - 2\theta w(X_1, X_2, \dots, X_n) + w^2(X_1, X_2, \dots, X_n)] f(\theta|X_1, X_2, \dots, X_n) d\theta$$

$$= \int_{-\infty}^{\infty} \theta^2 f(\theta|X_1, X_2, \dots, X_n) d\theta - 2w(X_1, X_2, \dots, X_n) \int_{-\infty}^{\infty} \theta f(\theta|X_1, X_2, \dots, X_n) d\theta$$

$$+ w^2(X_1, X_2, \dots, X_n) \int_{-\infty}^{\infty} f(\theta|X_1, X_2, \dots, X_n) d\theta$$

$$= \int_{-\infty}^{\infty} \theta^2 f(\theta|X_1, X_2, \dots, X_n) d\theta) - 2w(X_1, X_2, \dots, X_n) \int_{-\infty}^{\infty} \theta f(\theta|X_1, X_2, \dots, X_n) d\theta$$

$$+ w^2(X_1, X_2, \dots, X_n)$$

Untuk meminimumkan ekspektasi bersyarat dari fungsi rugi di atas bisa didapat dengan mencari:

$$\frac{\partial E\{L[\theta, w(X_1, X_2, \dots, X_n)]|X_1, X_2, \dots, X_n\}}{\partial w(X_1, X_2, \dots, X_n)} = 0$$

$$-2 \int_{-\infty}^{\infty} \theta f(\theta|X_1, X_2, \dots, X_n) d\theta + 2w(X_1, X_2, \dots, X_n) + 0 = 0$$

$$2w(X_1, X_2, \dots, X_n) = 2 \int_{-\infty}^{\infty} \theta f(\theta|X_1, X_2, \dots, X_n) d\theta$$

$$w(X_1, X_2, \dots, X_n) = \int_{-\infty}^{\infty} \theta f(\theta|X_1, X_2, \dots, X_n) d\theta.$$

Dengan demikian akan dihasilkan penaksir parameter θ adalah

$$w(X_1, X_2, \dots, X_n) = E(\theta|X_1, X_2, \dots, X_n) \qquad \dots (2.5)$$

atau dengan kata lain untuk fungsi rugi kuadratik, taksiran titik parameter θ adalah rata-rata dari distribusi posteriornya. Sementara itu, taksiran titik parameter θ dengan fungsi rugi absolut adalah median dari distribusi posteriornya. Sedangkan taksiran titik parameter θ dengan fungsi rugi *zero-one* adalah modus dari distribusi posteriornya.

2.4 Distribusi untuk Frekuensi Klaim

Frekuensi klaim adalah banyaknya klaim yang dilakukan oleh seorang pemegang polis selama masa asuransinya. Misalkan pemegang polis A mengajukan klaim sebanyak 3 kali selama masa asuransinya. Dengan demikian frekuensi klaim pemegang polis A adalah 3. Distribusi yang sering digunakan untuk memodelkan data frekuensi klaim diantaranya adalah distribusi Poisson dan distribusi binomial negatif (Klugman dkk., 2004).

2.4.1 Distribusi Poisson

Peubah acak diskrit K dikatakan berdistribusi Poisson dengan fungsi peluangnya sebagai berikut:

$$P(K=k)=p_k=\frac{e^{-\lambda}\lambda^k}{k!}, \qquad \text{untuk } k = 0,1,2,\cdots \qquad \text{... (2.6)}$$

Rata-rata dan varians dari distribusi Poisson adalah:

$$E(K)=\lambda,$$
$$Var(K)=\lambda.$$

Misalkan $K_1, K_2, \dots, K_n$ adalah suatu sampel acak berukuran *n* dari distribusi Poisson dengan parameter λ, dengan nilai dari sampel acak tersebut adalah $k_1, k_2, \dots, k_n$. Penaksir kemungkinan maksimum untuk parameter λ dari distribusi Poisson (penurunannya pada Lampiran 1) adalah rata-rata sampelnya:

$$\hat{\lambda}=\bar{k}=\frac{1}{n}\sum_{i=1}^{n} k_i \qquad \text{... (2.7)}$$

Dalam uji kecocokan distribusi Poisson, selain rata-rata sampel biasanya dibutuhkan juga varians sampelnya untuk mengetahui apakah terdapat overdispersi atau tidak di dalam data. Rumus dari varians sampel tersebut adalah:

$$s_k^2 = \frac{\sum_{i=1}^{n}(k_i - \bar{k})^2}{n-1} \qquad \text{... (2.8)}$$

2.4.2 Distribusi Binomial Negatif

Misalkan peubah acak K berdistribusi Poisson dengan parameter λ, atau disingkat Poisson (λ). Akan tetapi λ itu sendiri merupakan peubah acak dan diasumsikan berdistribusi gamma dengan parameter a dan τ yaitu:

$K|\lambda \sim$ Poisson (λ)

$\lambda \sim$ Gamma (a, τ)

dimana Gamma (a, τ) adalah distribusi gamma sebagai distribusi prior dengan fungsi peluang sebagai berikut:

$$u(\lambda) = \frac{\tau^a e^{-\tau\lambda}\lambda^{a-1}}{\Gamma(a)} \quad ; a > 0, \tau > 0 \text{ dan } \lambda > 0 \qquad \text{... (2.9)}$$

dimana $\Gamma(a)$ adalah fungsi gamma dari a adalah:

$$\Gamma(a) = \int_0^\infty t^{a-1}e^{-t}dt.$$

Fungsi peluang marginal bagi K diberikan oleh:

$$P(K = k) = p_k = \int_0^\infty f(k)(\lambda)u(\lambda)d\lambda = \int_0^\infty \frac{e^{-\lambda}\lambda^k}{k!}\frac{\tau^a e^{-\tau\lambda}\lambda^{a-1}}{\Gamma(a)}d\lambda$$

$$= \frac{\tau^a}{k!\,\Gamma(a)}\int_0^\infty \lambda^{k+a-1}e^{-(1+\tau)\lambda}d\lambda$$

Misalkan $y = (1 + \tau)\lambda$, maka:

$$p_k = \frac{\tau^a}{k!\,\Gamma(a)} \int_0^\infty \left(\frac{y}{1+\tau}\right)^{k+a-1} e^{-y} \frac{dy}{1+\tau}$$

$$= \frac{\tau^a}{k!\,\Gamma(a)(1+\tau)^{k+a}} \int_0^\infty y^{k+a-1} e^{-y} dy$$

$$= \frac{\tau^a \Gamma(\mathrm{k}+\mathrm{a})}{k!\,\Gamma(a)(1+\tau)^{k+a}}$$

$$= \frac{(k+a-1)!\,\tau^a}{k!\,(a-1)!\,(1+\tau)^{k+a}}$$

$$p_k = \binom{k+a-1}{k} \left(\frac{\tau}{1+\tau}\right)^a \left(\frac{1}{1+\tau}\right)^k \qquad \text{... (2.10)}$$

Jika $k = 0$ maka:

$$p_0 = \binom{0+a-1}{0} \left(\frac{\tau}{1+\tau}\right)^a \left(\frac{1}{1+\tau}\right)^0$$

$$= \binom{a-1}{0} \left(\frac{\tau}{1+\tau}\right)^a$$

$$= \left(\frac{\tau}{1+\tau}\right)^a \qquad \text{... (2.11)}$$

Sedangkan jika $k = k + 1$, maka:

$$p_{k+1} = \binom{k+1+a-1}{k+1} \left(\frac{\tau}{1+\tau}\right)^a \left(\frac{1}{1+\tau}\right)^{k+1}$$

$$p_{k+1} = \binom{k+a}{k+1} \left(\frac{\tau}{1+\tau}\right)^a \left(\frac{1}{1+\tau}\right)^k \left(\frac{1}{1+\tau}\right)$$

$$= \frac{k+a}{k+1} \frac{k+a-1}{k!\,(a-1)!} \left(\frac{\tau}{1+\tau}\right)^a \left(\frac{1}{1+\tau}\right)^k \left(\frac{1}{1+\tau}\right)$$

$$= \frac{k+a}{k+1} \left(\frac{1}{1+\tau}\right) \binom{k+a-1}{k} \left(\frac{\tau}{1+\tau}\right)^a \left(\frac{1}{1+\tau}\right)^k$$

$$= \frac{k+a}{(k+1)(1+\tau)} p_k \qquad \text{... (2.12)}$$

Persamaan (2.10) adalah fungsi peluang dari distribusi binomial negatif (Lemaire, 1995). Rata-rata dan varians dari distribusi binomial negatif adalah:

$$E(K) = \frac{a}{\tau} \quad \text{... (2.13)}$$

$$V(K) = \frac{a}{\tau}\left(1 + \frac{1}{\tau}\right).$$

Misalkan $K_1, K_2, \dots, K_n$ adalah suatu sampel acak berukuran n dari distribusi binomial negatif dengan parameter a dan τ, dengan nilai dari sampel acak tersebut adalah $k_1, k_2, \dots, k_n$. Penaksir kemungkinan maksimum untuk a dan τ dari distribusi binomial negatif tidak dapat diselesaikan secara eksplisit, namun hanya dapat ditaksir dengan menggunakan metode numerik. Salah satu metode numerik tersebut adalah metode Newton Raphson. Taksiran parameter a dan τ adalah solusi dari 2 persamaan (penurunannya pada Lampiran 2) berikut:

$$n \ln\left(\frac{1+\tau}{\tau}\right) = \sum_{i=1}^{n} \sum_{j=0}^{k_i - 1} \frac{1}{a+j} \quad \text{... (2.14)}$$

$$\tau = \frac{a}{\bar{k}}. \quad \text{... (2.15)}$$

Perangkat lunak Matlab menyediakan fungsi untuk menghitung taksiran kemungkinan maksimum dari parameter distribusi binomial negatif di atas.

2.5 Distribusi untuk Besar Klaim

Selain peubah acak frekuensi klaim, peubah acak besar klaim sering dimodelkan dalam asuransi kerugian. Besar klaim adalah besarnya pembayaran yang diberikan oleh perusahaan asuransi untuk menggantikan kerugian yang diklaim oleh pemegang polis. Misalkan pemegang polis A mengajukan klaim sebanyak 3 kali selama masa asuransinya. Dengan besar klaim masing-masing sebesar Rp. 300.000, Rp. 450.000, dan Rp. 500.000.

Distribusi yang sering digunakan untuk memodelkan data besar klaim diantaranya adalah distribusi lognormal, distribusi gamma, dan distribusi eksponensial (Klugman dkk., 2004). Dalam skripsi ini pembahasan akan dibatasi untuk distribusi lognormal. Misalkan X merupakan peubah acak yang berdistribusi lognormal dengan fungsi peluang sebagai berikut:

$$f(x) = \frac{1}{x\sigma\sqrt{2\pi}} exp\left[-\frac{1}{2}\left(\frac{\ln x - \theta}{\sigma}\right)^2\right] \; ; \; 0 < x < \infty \; ; \; -\infty < \theta < \infty \; ; \sigma > 0$$

Fungsi distribusi kumulatif dari distribusi lognormal adalah:

$$F(x) = \Phi\left(\frac{\ln x - \theta}{\sigma}\right); \; 0 < x < \infty. \quad \text{... (2.16)}$$

dimana $\Phi(\cdot)$ adalah notasi untuk fungsi distribusi kumulatif normal baku. Rata-rata dan varians dari distribusi lognormal masing-masing adalah:

$$E(X) = \exp\left(\theta + \frac{1}{2}\sigma^2\right),$$

$$Var(X) = \exp(2\theta + \sigma^2)\,[\exp(\sigma^2) - 1].$$

Misalkan $X_1, X_2, \dots, X_n$ adalah suatu sampel acak berukuran n dari distribusi lognormal dengan parameter θ dan σ^2, dengan nilai dari sampel acak tersebut adalah $x_1, x_2, \dots, x_n$. Penaksiran kemungkinan maksimum untuk θ dan σ^2 dari distribusi lognormal (penurunannya ada Lampiran 3) adalah:

$$\hat{\theta} = \frac{1}{n}\sum_{i=1}^{n} \ln x_i, \quad \text{... (2.17)}$$

$$\hat{\sigma}^2 = \frac{1}{n}\sum_{i=1}^{n} \left(\ln x_i - \hat{\theta}\right)^2 \quad \text{... (2.18)}$$

2.6 Uji Kecocokan Distribusi

Uji kecocokan distribusi adalah suatu pengujian hipotesis statistik yang digunakan untuk mengetahui apakah $x_1, x_2, \dots, x_n$ adalah nilai dari sampel acak $X_1, X_2, \dots, X_n$ yang berdistribusi dengan fungsi distribusi $F(\cdot)$. Uji kecocokan distribusi dapat digunakan untuk menguji hipotesis berikut:

H_0 : $x_1, x_2, \ldots, x_n$ merupakan nilai dari sampel acak yang berdistribusi dengan fungsi distribusi $F(\cdot)$.

H_1 : $x_1, x_2, \ldots, x_n$ merupakan nilai dari sampel acak yang berdistribusi dengan fungsi distribusi bukan $F(\cdot)$.

Dalam skripsi ini uji kecocokan distribusi yang akan digunakan adalah uji kecocokan chi-kuadrat untuk frekuensi klaim dan uji kecocokan Anderson-Darling untuk besar klaim.

2.6.1 Uji Kecocokan Chi-Kuadrat

Salah satu uji kecocokan distribusi untuk kasus data diskrit adalah uji kecocokan chi-kuadrat. Statistik uji untuk menguji kecocokan chi-kuadrat untuk kasus data diskrit adalah:

$$\chi^2 = \sum_{k=1}^{l} \frac{(n_k - np_k)^2}{np_k} \qquad \text{... (2.19)}$$

dimana n_k adalah banyaknya pengamatan untuk kategori k, np_k adalah nilai harapan untuk kategori k, dan l adalah banyaknya kategori. Statistik uji di atas berdistribusi chi-kuadrat dengan derajat bebas $l - r - 1$, dimana r menyatakan banyaknya parameter yang ditaksir dari distribusi. Dengan demikian kriteria pengujiannya adalah tolak hipotesis nol jika statistik uji chi-kuadrat lebih besar dari nilai kuantil distribusi chi-kuadrat pada taraf nyata α dan derajat bebas $l - r - 1$ atau $\chi^2 \geq \chi^2_{(l-r-1),(1-\alpha)}$.

2.6.2 Uji Anderson-Darling

Pendekatan yang awal untuk menguji kecocokan distribusi pada data adalah uji visual atau grafik. Uji visual tersebut diantaranya adalah normal PP plot, normal QQ plot, Boxplot dan Stem-and-leaf. Salah satu pengujian secara formal adalah Uji Anderson-Darling yang digunakan untuk menguji kecocokan distribusi untuk suatu kumpulan data (Law dan Kelton, 1991). Misalkan $x_1, x_2, \dots, x_n$ adalah nilai dari sampel acak berukuran *n*. Misalkan juga bahwa hasil dari pengurutan data dari kecil ke besar adalah $x_{(1)}, x_{(2)}, \dots, x_{(n)}$. Statistik uji Anderson-Darling adalah:

$$A_n{}^2 = -\sum_{i=1}^{n}\left(\frac{2i-1}{n}\right)\left[ln\left(\hat{F}\left(x_{(i)}\right)\right) + ln\left(1-\hat{F}\left(x_{(n+1-i)}\right)\right)\right] - n \qquad \text{... (2.20)}$$

dimana $\hat{F}(\cdot)$ adalah taksiran fungsi distribusi kumulatif untuk distribusi yang dihipotesiskan. Nilai kritis untuk Anderson-Darling adalah 1,933 untuk taraf nyata α =10%, 2,492 untuk taraf nyata α =5% dan 3,857 untuk taraf nyata α =1%. (Klugman dkk., 2004). Sedangkan kriteria pengujiannya adalah tolak hipotesis nol jika nilai statistik uji $A_n{}^2$ lebih besar dari nilai kritis pada taraf nyata yang ditetapkan.

2.7 Metode Penaksiran *Linear Empirical Bayesian*

Dalam bagian ini akan diberikan teori penaksiran parameter menggunakan metode *Linear Empirical Bayesian* (LEB). Jika (θ, X) adalah suatu vektor acak. Parameter θ akan ditaksir oleh $\hat{\theta}$ sebagai fungsi dari X sedemikian sehingga meminimumkan $E\left(\hat{\theta} - \theta\right)^2$. Untuk kasus tersebut diperoleh penaksir $\hat{\theta} = E(\theta|X)$. Jika distribusi bersama dari (θ, X) tidak diketahui, maka $E(\theta|X)$ tidak dapat diperoleh secara eksplisit. Untuk mengatasi masalah tersebut Robbins (1985) memperkenalkan fungsi linier dari $\hat{\theta}$ yaitu $\hat{\theta}(X) = a + bX$. Misalkan

$$\begin{aligned}\phi(a,b) &= E\left(\hat{\theta}(X) - \theta\right)^2 \\ &= E(a + bX - \theta)^2 \\ &= a^2 + b^2E(X^2) + E(\theta^2) + 2abE(X) - 2aE(\theta) - 2bE(\theta X),\end{aligned}$$

dengan:

$$E(X^2) = Var(X) + [E(X)]^2$$
$$E(\theta^2) = Var(\theta) + [E(\theta)]^2$$
$$E(\theta X) = Cov(\theta, X) + E(\theta)E(X)$$

maka:

$$\begin{aligned}\phi(a,b) &= a^2 + b^2[Var(X) + \{E(X)\}^2] + [Var(\theta) + \{E(\theta)\}^2] + 2abE(X) \\ &\quad - 2aE(\theta) - 2b[Cov(\theta,X) + E(\theta)E(X)] \\ &= a^2 + b^2Var(X) + b^2[E(X)]^2 + Var(\theta) + [E(\theta)]^2 + 2abE(X) \\ &\quad - 2aE(\theta) - 2b\,Cov(\theta,X) - 2bE(\theta)E(X) \\ &= [a - E(\theta) + bE(X)]^2 + var(X)\left[b - \frac{cov(\theta,X)}{var(X)}\right]^2 \\ &\quad + \frac{var(\theta)var(X) - cov^2(\theta,X)}{var(X)}.\end{aligned}$$

Dari persamaan di atas akan dicari nilai a dan b yang meminimumkan fungsi $\phi(a,b)$.

$$\frac{\partial\phi(a,b)}{a} = 2[a - E(\theta) + bE(X)] = 0.$$

$$\frac{\partial\phi(a,b)}{b} = 2[a - E(\theta) + bE(X)]E(X) + 2Var(X)\left(b - \frac{cov(\theta,X)}{var(X)}\right) = 0.$$

Nilai a dan b yang meminimumkan fungsi $\phi(a,b)$ adalah solusi dari kedua persamaan di atas, yaitu

$$a = E(\theta) - bE(X), \quad \text{dan} \quad b = \frac{cov(\theta,X)}{var(X)}.$$

Dengan mensubstitusikan nilai a dan b, maka nilai taksiran parameter $\hat{\theta}$ adalah

$$\begin{aligned}\hat{\theta}(X) &= a + bX \\ &= E(\theta) + \frac{cov(\theta,X)}{var(X)}[X - E(X)]. \qquad \text{... (2.21)}\end{aligned}$$

Jika distribusi bersyarat dari $X|\theta$ adalah distribusi normal dengan rata-rata θ dan varians σ^2, maka:

$$E(X|\theta) = \theta, \quad \text{dan} \quad Var(X|\theta) = \sigma^2.$$

Dengan menggunakan Persamaan (2.20), taksiran parameter $\hat{\theta}$ dengan menggunakan teori penaksiran LEB adalah:

$$\theta = E(\theta) + \frac{cov(\theta, X)}{var(X)}[X - E(X)] \quad \text{... (2.22)}$$

dengan:

$$E(\theta) = E[E(X|\theta)] = E(X) \quad \text{... (2.23)}$$

$$\begin{aligned} Cov(\theta, X) &= E[\{\theta - E(\theta)\}\{X - E(X)\}] \\ &= E[E[\{\theta - E(\theta)\}\{X - E(X)\}]|\theta] \\ &= E\big[\{\theta - E(\theta)\}E[\{X - E(X)\}|\theta]\big] \\ &= E\big[\{\theta - E(\theta)\}E[\{X - E(\theta)\}|\theta]\big] \\ &= E[\{\theta - E(\theta)\}\{E(X|\theta) - E(\theta)\}] \\ &= E[\{\theta - E(\theta)\}\{\theta - E(\theta)\}] \\ &= Var(\theta) \end{aligned} \quad \text{... (2.24)}$$

$$\begin{aligned} Var(X) &= E[Var(X|Y)] + Var[E(X|Y)] \\ &= E[Var(X|Y)] + Var(\theta) \\ &= E(\sigma^2) + Var(\theta) \\ &= \sigma^2 + Var(\theta) \end{aligned}$$

Sehingga

$$Var(\theta) = Var(X) - \sigma^2 \quad \text{... (2.25)}$$

Dengan mensubstitusikan Persamaan (2.23), (2.24) dan (2.25) ke Persamaan (2.22), akan diperoleh taksiran parameter $\hat{\theta}$, yaitu:

$$\begin{aligned} \theta &= E(\theta) + \frac{cov(\theta, X)}{var(X)}[X - E(X)] \\ &= E(X) + Var(\theta)Var^{-1}(X)[X - E(X)] \\ &= E(X) + \frac{Var(X) - \sigma^2}{Var(X)}[X - E(X)] \end{aligned}$$

$$\hat{\theta} = \left(1 - \frac{\sigma^2}{Var(X)}\right)X + \frac{\sigma^2}{Var(X)}E(X) \quad \text{... (2.26)}$$

Misalkan $x_1, x_2, \dots, x_n$ merupakan data pengamatan sejarah klaim, maka rata-rata dan varians dari sampel tersebut adalah:

$$\bar{x} = \frac{1}{n}\sum_{i=1}^{n} x_i \,, \qquad \text{dan } \; s^2 = \frac{1}{n-1}\sum_{i=1}^{n}(x_i - \bar{x})^2$$

Dengan mengganti $E(X)$ dan $Var(X)$ yang ada pada Persamaan (2.26) oleh $\bar{x}$ dan s^2, maka diperoleh:

$$\hat{\theta} = \left(1 - \frac{\sigma^2}{s^2}\right)x + \frac{\sigma^2}{s^2}\bar{x} \qquad \text{... (2.27)}$$

2.8 Taksiran Premi untuk Asuransi Kendaraan Bermotor

Premi dapat dihitung berdasarkan taksiran besar klaim dan frekuensi klaim. Taksiran besar klaim dapat diperoleh dengan menggunakan teori LEB yang dibahas pada subbab sebelumnya.

Asumsikan peubah acak X berdistribusi lognormal dengan parameter θ dan σ^2 dengan σ^2 diketahui dan fungsi peluangnya adalah:

$$f(x) = \frac{1}{x\sigma\sqrt{2\pi}} exp\left[-\frac{1}{2}\left(\frac{\ln x - \theta}{\sigma}\right)^2\right] \,; \qquad 0 < x < \infty \,; \, -\infty < \theta < \infty \,; \sigma > 0$$

Berdasarkan fungsi peluang di atas, ekspektasi besar klaim yang optimal adalah:

$$E(X) = \exp\left(\theta + \frac{1}{2}\sigma^2\right) \qquad \text{... (2.28)}$$

Jika diperoleh taksiran optimal dari parameter θ, maka dapat dihitung ekspektasi dari X.

Misalkan sejarah data besar klaim asuransinya adalah $x_1, x_2, \dots, x_n$, dan asumsikan bahwa pengamatan baru saat ini adalah x. Jika besar klaim X, berdistribusi lognormal dengan parameter θ dan σ^2, atau $\ln X$ berdistribusi normal dengan parameter θ dan σ^2, maka $E(\ln X|\theta) = \theta$ dan $Var(\ln X|\theta) = \sigma^2$. Ketika σ^2 diketahui, dapat diperoleh taksiran parameter θ berdasarkan Persamaan (2.27), yaitu:

$$\hat{\theta} = \left(1 - \frac{\sigma^2}{s^2}\right) x + \frac{\sigma^2}{s^2} \bar{x}$$

$$= \left(1 - \frac{\sigma^2}{s^2}\right) \ln x + \frac{\sigma^2}{s^2} \overline{\ln x} \qquad \text{... (2.29)}$$

dengan:

$$\overline{\ln x} = \frac{1}{n} \sum_{i=1}^{n} \ln x_i \qquad \text{... (2.30)}$$

$$s^2 = \frac{1}{n-1} \sum_{i=1}^{n} (\ln x_i - \overline{\ln x})^2 \qquad \text{... (2.31)}$$

dimana x_i adalah data besar klaim asuransi, dan n adalah banyaknya pengamtan data besar klaim.

Sedangkan nilai σ^2 dihitung berdasarkan data besar klaim pada periode sebelumnya. Misalkan sejarah data besar klaim asuransi pada periode sebelumnya adalah $x_1^p, x_2^p, \dots, x_m^p$, maka σ^2 dihitung dengan menggunakan rumus berikut:

$$\sigma^2 = \frac{1}{m-1} \sum_{i=1}^{m} (\ln x_i^p - \overline{\ln x^p})^2 \qquad \text{... (2.32)}$$

dengan:

$$\overline{\ln x^p} = \frac{1}{m} \sum_{i=1}^{m} \ln x_i^p \qquad \text{... (2.33)}$$

dimana x^p adalah data besar klaim periode sebelumnya, dan m adalah banyaknya pengamtan data besar klaim periode sebelumnya.

Dari Persamaan (2.28) dapat dihitung ekspektasi besar klaim yang optimal, yaitu:

$$E(X) = \exp\left\{\left(1 - \frac{\sigma^2}{s^2}\right) \ln x + \frac{\sigma^2}{s^2} \overline{\ln x} + \frac{\sigma^2}{2}\right\} \qquad \text{... (2.34)}$$

Berdasarkan taksiran besar klaim di masa yang akan datang maka dapat diperoleh taksiran premi bersih di masa yang akan datang yaitu:

$$P = E(K).E(X) \qquad \text{... (2.35)}$$

dimana $E(K)$ adalah ekspektasi frekuensi klaim di masa yang akan datang bagi seorang pemegang polis, dan $E(X)$ adalah ekspektasi besar klaim yang optimal di masa yang akan datang bagi seorang pemegang polis untuk sekali klaim.

2.9 Transformasi Box-Cox

Transformasi Box-Cox atau seringkali disebut transformasi pangkat Box-Cox adalah suatu metode mentransformasi data pengamatan dari yang asalnya tidak berdistribusi normal menjadi berdistribusi normal. Misalkan $x_1, x_2, \dots, x_n$, adalah nilai dari suatu sampel acak berukuran n dari suatu distribusi kontinu yang tidak berdistribusi normal. Transformasi pangkat Box-Cox untuk data pengamatan $x_i; i = 1, \cdots, n$ adalah (Asar dkk., 2013)

$$w_i = \begin{cases} \dfrac{{x_i}^{\lambda} - 1}{\lambda} & , \lambda \neq 0, \\ \ln x_i & , \lambda = 0, \end{cases}$$

dimana λ disebut sebagai parameter transformasi pangkat yang nilainya tidak diketahui dan perlu ditaksir. Nilai λ biasanya berada diantara -5 sampai 5. Tabel 2.2. menunjukkan berbagai nilai parameter λ dan bentuk transformasinya.

Tabel 2.2. Nilai λ dan Transformasinya

λ	Transformasi
2	X^2
1	X
0,5	$\sqrt{X}$
0	$\log(X)$ atau $\ln(X)$
-0,5	$1/\sqrt{X}$
-1,0	$1/X$
-2,0	$1/X^2$

Salah satu pendekatan yang dapat digunakan untuk menaksir parameter transformasi pangkat λ untuk data pengamatan $x_1, x_2, \dots, x_n$ adalah dengan menggunakan metode kemungkinan maksimum (Li, 2005). Misalkan fungsi peluang untuk peubah acak W_i adalah:

$$f(w_i) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left\{-\frac{1}{2}\left(\frac{w_i - \theta}{\sigma}\right)^2\right\} ; \quad -\infty < w_i < \infty ; -\infty < \theta < \infty ; \sigma > 0.$$

Melalui hubungan transformasi pangkat Box-Cox, dapat ditunjukkan bahwa fungsi peluang untuk peubah acak X_i adalah:

$$f(x_i) = \frac{1}{\sigma\sqrt{2\pi}} \exp\left\{\frac{({x_i}^{\lambda} - 1)/\lambda - \theta}{\sigma}\right\}^2 {x_i}^{\lambda-1} \; ; 0 < x_i < \infty ; -\infty < \theta < \infty ; \sigma > 0.$$

Dengan demikian fungsi kemungkinannya adalah:

$$L(\lambda, \theta, \sigma) = \prod_{i=1}^{n} \left[\frac{1}{\sigma\sqrt{2\pi}} \exp\left\{\frac{({x_i}^{\lambda} - 1)/\lambda - \theta}{\sigma}\right\}^2 {x_i}^{\lambda-1}\right].$$

Logaritma natural fungsi kemungkinannya adalah:

$$\ell(\lambda, \theta, \sigma) = \sum_{i=1}^{n} \left[-\ln\sigma - \ln\sqrt{2\pi} - \frac{1}{2}\left\{\frac{({x_i}^{\lambda} - 1)/\lambda - \theta}{\sigma}\right\}^2 + (\lambda - 1)\ln x_i\right]. \qquad \text{... (2.36)}$$

Untuk nilai λ tertentu, taksiran parameter θ yang memaksimumkan fungsi kemungkinan adalah:

$$\hat{\theta}(\lambda) = \frac{\sum_{i=1}^{n} ({x_i}^{\lambda} - 1)/\lambda}{n}.$$

Sedangkan taksiran parameter σ^2 yang memaksimumkan fungsi kemungkinan adalah

$$\hat{\sigma}^2(\lambda) = \frac{\sum_{i=1}^{n} \left\{({x_i}^{\lambda} - 1)/\lambda - \hat{\theta}(\lambda)\right\}^2}{n}.$$

Dengan mensubstitusikan taksiran parameter θ dan σ^2 ke Persamaan (2.36) akan diperoleh logaritma natural fungsi kemungkinan untuk parameter λ, yaitu

$$\ell(\lambda) = C - \frac{n}{2}\ln\hat{\sigma}^2(\lambda) + (\lambda - 1)\sum_{i=1}^{n} \ln x_i, \qquad \text{... (2.37)}$$

dimana $C = -n \ln \sqrt{2\pi} - \frac{n}{2}$. Nilai λ yang memaksimumkan logaritma natural fungsi kemungkinan pada Persamaan (2.37) adalah nilai λ yang akan dicari sebagai taksiran parameter transformasi pangkat. Nilai taksiran ini akan digunakan untuk mentransformasi data pengamatan $x_1, x_2, \ldots, x_n$ menjadi data pengamatan $w_1, w_2, \ldots, w_n$ yang berdistribusi normal.

Berdasarkan pada teori yang telah dijelaskan di atas, maka dapat dibangun suatu tahapan menentukan nilai taksiran parameter transformasi pangkat λ yang memaksimumkan logaritma natural fungsi kemungkinan pada Persamaan (2.37) untuk data pengamatan $x_1, x_2, \ldots, x_n$. Tahapan tersebut disajikan dalam Gambar 2.1.

Tahapan	**Aktifitas**
Tahap 1	Menentukan berbagai nilai parameter λ.
Tahap 2	Menghitung nilai logaritma natural fungsi kemungkinan $\ell(\lambda)$ menggunakan Persamaan (2.37) untuk setiap masing-masing nilai λ.
Tahap 3	Membuat plot antara logaritma natural fungsi kemungkinan $\ell(\lambda)$ dan λ kemudian memperhatikan titik kritis λ pada $\ell(\lambda)$ maksimum. Nilai λ yang membuat nilai $\ell(\lambda)$ maksimum itulah nilai taksiran parameter bagi λ.
Tahap 4	Menentukan jenis transformasi berdasarkan nilai taksiran parameter λ yang diperoleh menggunakan informasi yang ada pada Tabel 2.2.

Gambar 2.1. Tahapan Menentukan Nilai λ pada Transformasi Box-Cox

***