Estimasi Cadangan Klaim dengan Pendekatan Over-dispersed Poisson

1. Latar Belakang

Cadangan klaim merupakan sejumlah dana yang dihimpun oleh suatu perusahaan asuransi yang digunakan untuk memenuhi kewajiban untuk membayar klaim yang terjadi di masa mendatang. Perhitungan cadangan klaim menjadi hal yang krusial bagi perusahaan karena jika estimasi cadangan klaim terlalu rendah dari nilai sebenarnya, maka perusahaan asuransi tidak memiliki dana yang cukup untuk membayar klaim sehingga dapat mengakibatkan kesulitan keuangan dan mengancam keberlangsungan perusahaan. Sedangkan jika estimasi cadangan klaim terlalu tinggi, maka penggunaan sumber daya menjadi tidak efisien dan perusahaan asuransi tidak dapat menggunakan dana tersebut untuk berinvestasi pada pertumbuhan perusahaan. Untuk itu diperlukan suatu metode yang dapat digunakan sebagai dasar perhitungan estimasi cadangan klaim sehingga estimasi yang diperoleh sesuai dengan kebutuhan perusahaan.

Secara umum, metode estimasi cadangan klaim terbagi menjadi dua, yaitu metode deterministik dan metode stokastik. Metode Chain Ladder merupakan metode deterministic yang paling popular karena mudah diterapkan dan bersifat bebas distribusi. Dalam metode Chain Ladder, perhitungan estimasi cadangan klaim dilakukan berdasar data klaim historis. Sedangkan salah satu metode yang paling popular dalam metode stokastik adalah *Generalized Linear Model* (GLM). Dalam GLM terdapat asumsi-asumsi awal yang diberikan seperti distribusi dari data dengan parameter yang diestimasi menggunakan *Maximum Likelihood Estiamtion* (MLE).

Salah satu keunggulan metode stokastik yaitu keakuratan dari estimasi dapat diukur dengan *prediction error* yang dapat diperoleh dari metode ini. Estimasi yang dihasilkan dari metode GLM juga menjadi lebih informatif karena dapat memberikan *Confidence Interval* dengan tingkat keyakinan tertentu. Dalam kasus ini, akan digunakan metode stokastik *Generalized Linear Model* (GLM) dengan pendekatan Over-dispersed Poisson (ODP) untuk menghitung estimasi cadangan klaim.

2. Landasan Teori

1) *Generalized Linear Model* (GLM)

   Analisis regresi linear merupakan metode yang digunakan untuk mengetahui hubungan linear antara variable respons dan variable predictor. Dalam regresi linear terdapat asumsi yang harus dipenuhi yaitu eror yang berdistribusi normal dan varians yang konstan. Tetapi, asumsi tersebut seringkali tidak dapat terpenuhi sehingga metode ini kemudian dikembangkan menjadi *Generalized Linear Model* (GLM). Dalam GLM, distribusi variable respons dipilih dari distribusi yang termasuk dalam keluarga eksponensial. GLM merupakan salah satu metode stokastik populer yang digunakan dalam perhitungan cadangan klaim dengan kelebihan yaitu mampu memberikan nilai *prediction error* dan *confidence interval* dari hasil estimasi.

2) Distribusi Keluarga Eksponensial

   Variable respons y dalam GLM yang dipilih dari keluarga eksponensial memiliki bentuk umum yaitu:

$$f(y|\theta,\phi) = \exp\{\frac{y\theta - a(\theta)}{\phi} + c(y,\theta)\}$$

   Terdapat beberapa distribusi yang termasuk dalam keluarga eksponensial, diantaranya

Distribusi	θ	$a(\theta)$	ϕ	E(Y)	V(μ)
Normal (μ,σ^2)	μ	$\frac{1}{2}\theta^2$	σ^2	μ	1
Poisson (λ)	λ	e^θ	1	μ	μ
Gamma (μ, v)	$-\frac{1}{\mu}$	$-ln(-\theta)$	$\frac{1}{v}$	μ	μ^2
Binomial (n,π)	$ln\frac{\pi}{1-\pi}$	$nln(1+e^\theta)$	1	$n\pi$	$n\pi(1-\pi)$

   Dalam GLM, terdapat fungsi link yang menjelaskan hubungan yang linier antara fungsi dan transformasi dari mean dengan variabel respon dan prediktor, dinotasikan dengan $g(\mu)$.

Link Function	$g(\mu)$	*Canonical Link*
identity	μ	Normal
log	$\ln(\mu)$	Poisson
power	μ^p	Gamma (p=-1)
logit	$\ln(\frac{\mu}{1-\mu})$	Binomial

3) Distribusi Tweedie

Dalam distribusi keluarga eksponensial, terdapat distribusi *tweedie* yang memiliki fungsi varians sebanding dengan μ^p dengan p menjadi paremeter tambahan. Distribusi yang merupakan anggota dari distribusi Tweedie yaitu antara lain

Distribusi	Nilai p
Normal	0
Over-dispersed Poisson	1
Gamma	2
Inverse Gaussian	3

4) Over –dispersed Poisson Model

Distribusi ODP merupakan anggota distribusi *twiddie* dengan nilai p sebesar 1 yang dapat dimodelkan ke dalam bentuk generalized linear model (GLM). Misalkan *incremental claim* C_{ij} berdistribusi ODP dengan parameter m_{ij}, diperoleh nilai mean dan variance dari distribusi ODP:

$$E[C_{ij}] = m_{ij} = e^{c+\alpha_i+\beta_j} \text{ dan } Var[C_{ij}] = \phi m_{ij} = \phi e^{c+\alpha_i+\beta_j}$$

Diketahui $m_{ij} = e^{c+\alpha_i+\beta_j}$, jika persamaan tersebut diubah dalam bentuk transformasi logaritma natural, maka diperoleh

$$\ln(m_{ij}) = \eta = e^{c+\alpha_i+\beta_j}$$

Perbedaan antara distribusi Poisson dengan ODP adalah pada ODP terjadi overdispersi yaitu kondisi dimana varians lebih besar daripada mean. Hal ini terjadi karena adanya parameter ϕ yang merupakan suatu parameter dispersi bernilai lebih dari 1. Kemudian, parameter parameter yang terdapat dalam fungsi GLM tersebut diestimasi menggunakan metode Maximum Likelihood Estimator (MLE) dengan langkah pertama menentukan ln likelihood dari model yaitu:

$$\ln L(C_{ij}) = \sum_{i=1}^{n} \sum_{j=1}^{n-i+1} \{\frac{C_{ij}(c+\alpha_i+\beta_j) - e^{c+\alpha_i+\beta_j}}{\phi} + \ln(C_{ij}\,!)\}$$

Kemudian fungsi ln $L(C_{ij})$ dideferensialkan terhadap parameter dan disama denganka n 0 untuk mencari nilai maksimum dari fungsi

$$\frac{\partial ln\, L(C_{i,j})}{\partial c} = \sum_{i=1}^{n} \sum_{j=1}^{n-1+1} (\frac{C_{i,j} - e^{c+\alpha_i+\beta_i}}{\phi}) = 0$$

$$\frac{\partial ln\, L(C_{i,j})}{\partial \alpha_i} = \sum_{j=1}^{n-1+1} (\frac{C_{i,j} - e^{c+\alpha_i+\beta_i}}{\phi}) = 0$$

$$\frac{\partial ln\, L(C_{i,j})}{\partial \beta_i} = \sum_{i=1}^{n} (\frac{C_{i,j} - e^{c+\alpha_i+\beta_i}}{\phi}) = 0$$

Sehingga diperoleh persamaan dari setiap parameter sebagai berikut

$$\hat{c} = ln\left[\sum_{i=1}^{n} \sum_{j=1}^{n-i+1} \left(\frac{C_{i,j}}{e^{\alpha_i+\beta_i}}\right)\right]$$

$$\hat{\alpha}_i = ln\left[\sum_{j=1}^{n-1+1} (\frac{C_{i,j}}{e^{c+\beta_i}})\right]$$

$$\hat{\beta}_j = ln\left[\sum_{i=1}^{n} (\frac{C_{i,j}}{e^{c+\alpha_i}})\right]$$

Kemudian estimasi cadangan klaim dapat dicari dengan persamaan

$$\hat{C}_{i,y} = \hat{m}_{iy} = \exp(\eta) = \exp(c + \alpha_i + \beta_j)$$

Estimasi total cadangan klaim selanjutnya juga dapat diperoleh berdasarkan persamaan

$$\hat{R} = \sum_{i=2}^{n} \hat{R}_i$$

5) *Prediction Error*

Untuk mengukur keakuratan dari hasil prediksi, akan digunakan *prediction error* yang merupakan akar dari mean square error of prediction (MSEP).

$$MSEP(\hat{C}_{i,j}) = \phi \hat{m}_{i,j} + \hat{m}_{i,j}^2 var(\hat{\eta}_{i,j})$$

Sehingga dapat diperoleh nilai *prediction error* yaitu

$$prediction\ error = \sqrt{\phi \hat{m}_{iy} + \hat{m}_{iy}^2 var(\hat{\eta}_{i,y})}$$

6) *Confidence Interval*

Pada estimasi cadangan klaim dengan menggunakan metode stokastik dapat diperoleh batas atas dan batas bawah dari estimasi atau nilai *confidence interval* dengan persamaan yaitu

$$CI = \hat{R} \pm \alpha_{0.05} \sqrt{\frac{\hat{R}}{n}}$$

3. Asumsi yang Mendasari Over-dispersed Poisson

Berikut asumsi asumsi yang mendasari analisis data dengan pendekatan Over-dispersed Poisson, diantaranya yaitu:

1) Distribusi Poisson Terdispersi Tinggi

   Data diasumsikan memiliki distribusi poisson yang terdispersi tinggi dimana varians dari data lebih besar dari yang diperkirakan model poisson standard. Varians bersifat proporsional terhadap mean

2) Independensi

   Incremental claims (C_{ij}) didistribusikan sebagai variabel acak poisson yang independen dan over-dispersed dengan mean dan variance:

   $$E[C_{ij}] = m_{ij} = x_i y_j \text{ dan } Var[C_{ij}] = \phi x_i y_j$$

3) Pembatasan Terhadap y_j

   y_j yang merupakan proporsi ultimate claims yang muncul pada setiap tahun pengembangan harus bernilai positif karena muncul dalam varians sehingga jumlah *incremental claims* dalam kolom j juga harus bernilai positif. Terdapat kasus dimana *incremental claims* negatif diperbolehkan selama jumlah dari setiap kolom tidak bernilai negatif.

4) Rata Rata (μ_{ij}) memiliki Struktur M*ultiplicative*

   Yaitu merupakan hasil dari efek baris x_i (klaim ultimate yang diharapkan dalam *origin year)* dan efek kolom y_i (proporsi klaim ultimate yang akan muncul dalam setiap *development year*)

5) Reparametrisasi Model

   Reparametrisasi model dilakukan agar rata rata memiliki bentuk linear sehingga dapat memberikan fleksibilitas ketika terdapat variable respon yang memiliki rentang nilai luas atau ketika hubungan antar variable tidak linear. Digunakan log link function dengan demikian:

   $$\log(m_{ij}) = c + \alpha_i + \beta_i$$

4. Algoritma Pemodelan Over-dispersed Poisson

Terdapat beberapa langkah yang perlu dilakukan dalam estimasi cadangan klaim menggunakan GLM dengan pendekatan Over-dispersed Poisson, diantaranya yaitu:

1) Pembentukan data klaim asuransi dalam bentuk *run-off triangle*

2) Fitting distribution pada data incremental claim

Data incremental klaim dimodelkan menggunakan metode glm dengan asumsi data berdistribusi poisson. Kemudian akan diuji asumsi data berdistribusi poisson untuk mengetahui kecocokan dataset dengan model poisson.

3) Uji Over-dispersi pada model

Pada distribusi poisson diasumsikan mean bernilai sama dengan varians. Namun asumsi ini umumnya tidak terpenuhi sehingga terjadi overdispersi dimana nilai varians lebih besar dari nilai means. Uji overdispersi dilakukan untuk mengetahui apakah model mengalami overdispersi, sehingga nantinya akan dilakukan estimasi cadangan klaim menggunakan glm dengan pendekatan over-dispersed poisson.

4) Estimasi parameter pada Model Over-Dispersed Poisson

Dilakukan estimasi parameter-parameter variable independen yaitu tahun kejadian (α) dan Tahun Penundaan (β) dengan *Maximum Likelihood Estiamtion* (MLE) dan dibentuk model regresi.

5) Estimasi total Cadangan Klaim dengan Pendekatan ODP

Model regresi yang terbentuk dari parameter-parameter yang diperoleh sebelumnya, digunakan untuk estimasi klaim-klaim pada bagian bawah run-off triangle. Kemudian dilakukan perhitungan estimasi cadangan klaim pada setiap tahun kejadian yang lalu dijumlahkan seluruhnya untuk memperoleh estimasi total cadangan klaim.

6) Menghitung *prediction error* dan *confidence interval*

Digunakan untuk mengetahui keakuratan dari hasil prediksi cadangan klaim menggunakan pendekatan Over-Dispersed Poisson dan memberikan batas atas dan batas bawah dari hasil estimasi.

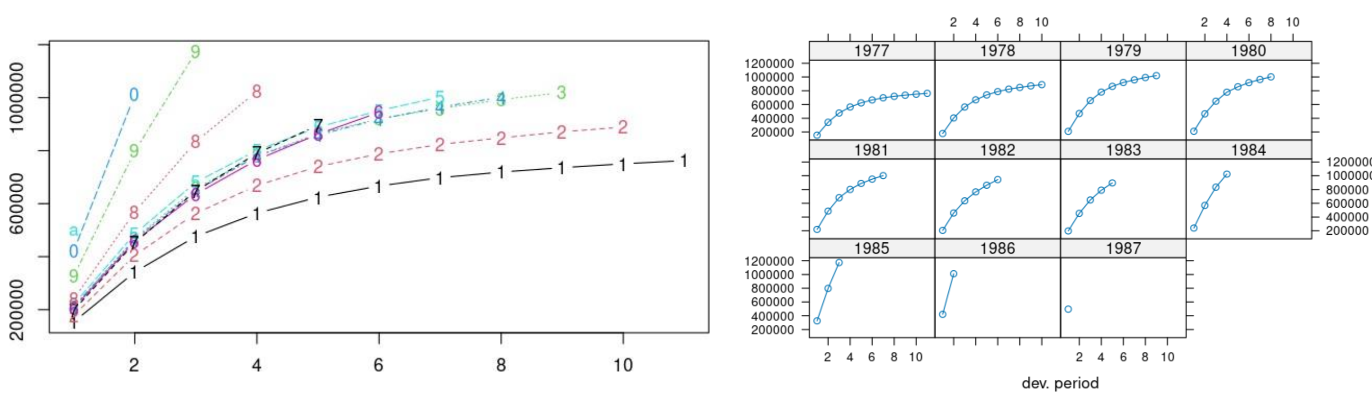
5. Studi Kasus

Data yang akan digunakan dalam studi kasus ini yaitu data ABC yang tersedia dalam library ChainLadder pada R yang merupakan data *cumulative claims* dari portfolio kompensasi pekerja di sebuah perusahaan besar yang berbentuk *run off triangle.* Data diketahui terdiri dari 11 *accident years* yaitu dari tahun 1977 hingga 1978 dan 11 *development years*.

able 1 Run-off Triangle Cumulative Claims

	1	2	3	4	5	6	7	8	9	10	11
1977	153638	342050	476584	564040	624388	666792	698030	719282	735904	750344	762544
1978	178536	404948	563842	668528	739976	787966	823542	848360	871022	889022	
1979	210172	469340	657728	780802	864182	920268	958764	992532	1019932		
1980	211448	464930	648300	779340	858334	918566	964134	1002134			
1981	219810	486114	680764	800862	888444	951194	1002194				
1982	205654	458400	635906	765428	862214	944614					
1983	197716	453124	647772	790100	895700						
1984	239784	569026	833828	1024228							
1985	326304	798048	1173448								
1986	420778	1011178									
1987	496200										

Selanjutnya akan ditampilkan visualisasi dari data *cumulative claims* untuk melihat karakteristik dari data dan trend dari besar klaim yang terjadi pada setiap tahun kejadian.



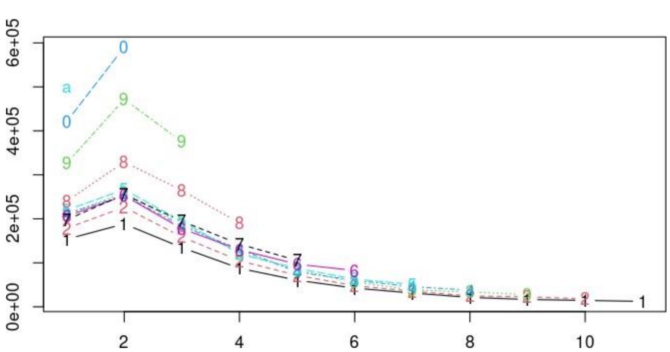
Terlihat pada grafik bahwa besar klaim terbesar yang telah dibayarkan oleh perusahaan asuransi adalah klaim yang terjadi pada tahun 1986 yang ditunjukkan dengan garis bernomor 0. Garis bernomor 0 cenderung berada di posisi paling atas dibanding lainnya. Sedangkan besar klaim terkecil yang dibayarkan perusahaan berada pada tahun 1977 yang ditunjukkan dengan garis bernomor 1. Hal yang sama dapat terlihat pada grafik bagian kanan yang merupakan besar *cumulative claims* yang disajikan setiap tahun kejadian.

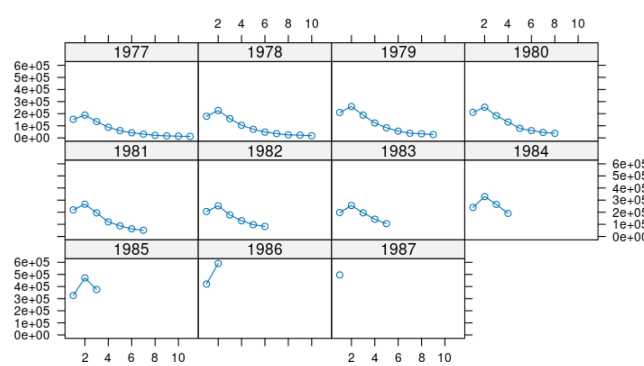
Selanjutnya data *cumulative claims* akan ditransformasikan menjadi bentuk *incremental claim* yang kemudian akan digunakan dalam pemodelan dengan glm.

Table 2 Run-off Triangle Incremental Claims

	1	2	3	4	5	6	7	8	9	10	11
1977	153638	188412	134534	87456	60348	42404	31238	21252	16622	14440	12200
1978	178536	226412	158894	104686	71448	47990	35576	24818	22662	18000	
1979	210172	259168	188388	123074	83380	56086	38496	33768	27400		
1980	211448	253482	183370	131040	78994	60232	45568	38000			
1981	219810	266304	194650	120098	87582	62750	51000				
1982	205654	252746	177506	129522	96786	82400					
1983	197716	255408	194648	142328	105600						
1984	239784	329242	264802	190400							
1985	326304	471744	375400								
1986	420778	590400									
1987	496200										

Kemudian akan ditampilkan visualisasi dari data *cumulative claims* untuk melihat karakteristik dari data.





Terlihat pada grafik diketahui bahwa besar klaim-klaim yang diselesaikan pembayarannya pada umumnya besar pada tahun pertama dan kedua kemudian semakin menurun pada tahun-tahun berikutnya. Hal ini terjadi karena pada umumnya perusahaan asuransi cenderung akan membayarka klaim pada tahun pertama kejadian. Kemudian, besar klaim akan menurun dikarenakan hanya menyisakan klaim-klaim yang pembayarannya tertunda.

Selanjutnya akan dilakukan estimasi cadangan estimasi menggunakan GLM dengan tahapan sebagai berikut:

1) Fitting Distribution pada Data *Incremental Claim*

Pada studi kasus ini, akan dilakukan estimasi cadangan klaim menggunakan GLM dengan pendekatan over-dispersed poisson. Dikarenakan pada model ODP harus memenuhi asumsi bahwa incremental claims berdistribusi ODP, maka perlu diketahui bahwa data yang digunakan dalam studi kasus ini memiliki distribusi poisson yang mengalami overdispersi. Data incremental klaim pertama dimodelkan menggunakan metode glm dengan asumsi data berdistribusi poisson.

```
Call:
glm(formula = value ~ factor(origin) + factor(dev), family = poisson,
    data = data3)

Coefficients:
                     Estimate Std. Error z value Pr(>|z|)
(Intercept)        11.8544785  0.0013221 8966.57   <2e-16 ***
factor(origin)1978  0.1695902  0.0015677  108.18   <2e-16 ***
factor(origin)1979  0.3269460  0.0015250  214.40   <2e-16 ***
factor(origin)1980  0.3350550  0.0015339  218.44   <2e-16 ***
factor(origin)1981  0.3687535  0.0015370  239.92   <2e-16 ***
factor(origin)1982  0.3560457  0.0015601  228.22   <2e-16 ***
factor(origin)1983  0.3730871  0.0015829  235.69   <2e-16 ***
factor(origin)1984  0.6146345  0.0015445  397.96   <2e-16 ***
factor(origin)1985  0.9329081  0.0015145  616.00   <2e-16 ***
factor(origin)1986  1.1355073  0.0015802  718.58   <2e-16 ***
factor(origin)1987  1.2602558  0.0019399  649.65   <2e-16 ***
factor(dev)2        0.2689569  0.0008639  311.33   <2e-16 ***
factor(dev)3       -0.0282497  0.0009994  -28.27   <2e-16 ***
factor(dev)4       -0.4216988  0.0012137 -347.46   <2e-16 ***

factor(dev)5       -0.8061033  0.0014983 -538.00   <2e-16 ***
factor(dev)6       -1.1431269  0.0018459 -619.28   <2e-16 ***
factor(dev)7       -1.4977757  0.0023583 -635.10   <2e-16 ***
factor(dev)8       -1.7807444  0.0030265 -588.38   <2e-16 ***
factor(dev)9       -2.0197567  0.0039724 -508.45   <2e-16 ***
factor(dev)10      -2.2488641  0.0056437 -398.47   <2e-16 ***
factor(dev)11      -2.4452873  0.0091496 -267.26   <2e-16 ***
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1

(Dispersion parameter for poisson family taken to be 1)

    Null deviance: 6238163  on 65  degrees of freedom
Residual deviance:   36594  on 45  degrees of freedom
  (55 observations deleted due to missingness)
AIC: 37521

Number of Fisher Scoring iterations: 4
```

Kemudian asumsi data berdistribusi poisson diuji untuk mengetahui kecocokan dataset dengan model poisson.

```
scaled.deviance df chisquare.95 p.value
       36593.86 45     61.65623       0
```

Dapat diketahui bahwa nilai chisquare yaitu sebesar 61.65623 dengan df sebesar 45 yang mana cukup kecil sehingga dapat disimpulkan dataset sesuai dengan model distribusi poisson.

2) Uji Overdispersi Pada Model

Pada distribusi poisson diasumsikan mean dan varians memiliki nilai yang sama (*equidispersi*). Namun asumsi ini umumnya tidak terpenuhi sehingga terjadi kondisi overdispersi dimana nilai varians lebih besar daripada nilai means. Uji overdispersi kemudian dilakukan untuk mengetahui apakah terjadi overdispersi pada model atau tidak.

```
        Overdispersion test

data:  poi
z = 5.2749, p-value = 6.64e-08
alternative hypothesis: true dispersion is greater than 1
sample estimates:
dispersion
  562.3924
```

- Hipotesis

  H_0 : Tidak terdapat overdispersi pada model regresi Poisson

  H_1 : Terdapat overdispersi pada model regresi Poisson

- Tingkat signifikansi

  $\alpha = 0.05$

- Daerah kritik

  H_0 ditolak jika p-value $< \alpha = 0.05$

- Statistik uji

  p-value = 6.64e-08

- Kesimpulan

  Dari hasil uji overdispersi diketahui p-value = 6.64e-08 $< \alpha = 0.05$ sehingga H_0 ditolak maka dapat disimpulkan bahwa terjadi overdispersi pada model regresi poisson

Karena diketahui dari hasil uji overdispersi bahwa terjadi overdispersi pada model regresi poisson maka selanjutnya akan dilakukan perhitungan estimasi cadangan klaim dengan menggunakan GLM dengan pendekatan over-dispersed poisson.

3) Estimasi Parameter pada Model Over-Dispersed Poisson

Langkah pertama yang perlu dilakukan dalam metode GLM adalah melakukan estimasi parameter variable independen pada data, yaitu variable *accident years* (α) dan *development years* (β) dengan *Maximum Likelihood Estiamtion* (MLE). Diperoleh hasil estimasi parameter sebagai berikut

Table 3 Estimasi Parameter Model GLM

Parameter	Estimasi	P-value
C	11.8545	< 2e-16
α_{1978}	0.16959	0.000478
α_{1979}	0.32695	2.09e-09
α_{1980}	0.33506	1.30e-09
α_{1981}	0.36875	1.06e-10
α_{1982}	0.35605	4.13e-10
α_{1983}	0.37309	1.73e-10
α_{1984}	0.61463	< 2e-16
α_{1985}	0.93291	< 2e-16
α_{1986}	1.13551	< 2e-16
α_{1987}	1.26026	< 2e-16
β_2	0.26896	3.91e-14
B_3	-0.02825	0.330285
B_4	-0.42170	9.63e-16
B_5	-0.80610	< 2e-16
B_6	-1.14313	< 2e-16
B_7	-1.49778	< 2e-16
B_8	-1.7807	< 2e-16
B_9	-2.01976	< 2e-16
B_{10}	-2.24886	< 2e-16
B_{11}	-2.44529	4.72e-12

Dari table di atas diperoleh estimasi untuk variabel prediktor *accident years* (α) dan variabel prediktor *development years* (β). Selanjutnya akan dilakukan uji hipotesis untuk mengetahui parameter-parameter yang berpengaruh signifikan terhadap model regresi. Hipotesis untuk pengujian parameter *accident years* pada persamaan regresi adalah sebagai berikut.

- Hipotesis
  $H_0 : \alpha_i = 0$ (variabel *accident year* ke-i tidak berpengaruh signifikan terhadap estimasi cadangan klaim)
  $H_1 : \alpha_i \neq 0$ (variabel *accident year* ke-i berpengaruh signifikan terhadap estimasi cadangan klaim)
- Tingkat signifikansi
  $\alpha = 0.05$
- Daerah kritik
  H_0 ditolak jika p-value < $\alpha = 0.05$
- Statistik uji

Parameter	P-Value	Kesimpulan
α_{1978}	0.000478	H_0 ditolak
α_{1979}	2.09e-09	H_0 ditolak
α_{1980}	1.30e-09	H_0 ditolak
α_{1981}	1.06e-10	H_0 ditolak
α_{1982}	4.13e-10	H_0 ditolak
α_{1983}	1.73e-10	H_0 ditolak
α_{1984}	< 2e-16	H_0 ditolak
α_{1985}	< 2e-16	H_0 ditolak
α_{1986}	< 2e-16	H_0 ditolak
α_{1987}	< 2e-16	H_0 ditolak

- Kesimpulan

  Diketahui bahwa semua parameter memiliki nilai p-value < $\alpha = 0.05$ sehingga H_0 ditolak dan dapat disimpulkan bahwa variable *accident year* dari tahun 1978 sampai 1987 berpengaruh signifikan terhadap estimasi cadangan klaim.

Selanjutnya dilakukan uji hipotesis untuk pengujian parameter *developmet year* pada persamaan regresi dan didapat hasil sebagai berikut sebagai berikut.

- Hipotesis

H_0 : $\alpha_i = 0$ (variabel *accident year* ke-i tidak berpengaruh signifikan terhadap estimasi cadangan klaim)

H_1 : $\alpha_i \neq 0$ (variabel *accident year* ke-i berpengaruh signifikan terhadap estimasi cadangan klaim)

- Tingkat signifikansi

  $\alpha = 0.05$

- Daerah kritik

  H_0 ditolak jika p-value < $\alpha = 0.05$

- Statistik uji

Parameter	P-Value	Kesimpulan
β_2	3.91e-14	H_0 ditolak
B_3	0.330285	H_0 tidak ditolak
B_4	9.63e-16	H_0 ditolak
B_5	< 2e-16	H_0 ditolak
B_6	< 2e-16	H_0 ditolak
B_7	< 2e-16	H_0 ditolak
B_8	< 2e-16	H_0 ditolak
B_9	< 2e-16	H_0 ditolak
B_{10}	< 2e-16	H_0 ditolak
B_{11}	4.72e-12	H_0 ditolak

- Kesimpulan

  Diketahui bahwa semua parameter kecuali j_3 memiliki nilai p-value < $\alpha = 0.05$ sehingga H_0 ditolak dan dapat disimpulkan bahwa semua paraemeter signifikan terhadap estimasi cadangan klaim, kecuali parameter *development year* kedua yang tidak berpengaruh signifikan terhadap estimasi cadangan klaim.

Kemudian, dari hasil estimasi parameter konstanta, parameter *accident years* (α) dan *development years* (β) dapat dibentuk model regresi sebagai berikut

$$\begin{aligned}\hat{C}_{iy} = 11{,}8545 &+ 0.16959X_{1,1978} + 0.32695X_{2,1979} + 0.33506X_{3,1980} + 0.36875X_{4,1981} \\ &+ 0.35605X_{5,1982} + 0.37309X_{6,1983} + 0.61463X_{7,1984} + 0.93291X_{8,1985} \\ &+ 1.13551X_{9,1986} + 1.26026X_{10,1987} + 0.26896X_{2,2} - 0.02825X_{2,3} \\ &- 0.42170X_{2,4} - 0.80610X_{2,5} - 1.14313X_{2,6} - 1.49778X_{2,7} - 1.7807X_{2,8} \\ &- 2.01976X_{2,9} - 2.24886X_{2,10} - 2.44529X_{2,11}\end{aligned}$$

4) Estimasi total Cadangan Klaim dengan Pendekatan ODP

Model regresi yang terbentuk dari parameter-parameter yang diperoleh sebelumnya, kemudian digunakan untuk estimasi klaim-klaim pada bagian bawah run-off triangle. Sebagai contoh akan dihitung estimasi klaim untuk tahun kejadian1979.

$$\hat{C}_{1979,10} = \exp(11{,}8545 + 0.32695 - 2.24886) = 20590.1$$

$$\hat{C}_{1979,11} = \exp(11{,}8545 + 0.32695 - 2.44529) = 16918$$

Berikut ini hasil estimasi klaim-klaim pada bagian bawah run-off triangle secara lengkap

Table 4 Hasil Estimasi Cadangan Klaim

	1	2	3	4	5	6	7	8	9	10	11
1977	153638	342050	476584	564040	624388	666792	698030	719282	735904	750344	762544
1978	178536	404948	563842	668528	739976	787966	823542	848360	871022	889022	14454.77
1979	210172	469340	657728	780802	864182	920268	958764	992532	1019932	20590.19	16918.11
1980	211448	464930	648300	779340	858334	918566	964134	1002134	26102.33	20757.85	17055.88
1981	219810	486114	680764	800862	888444	951194	1002194	34285.92	26996.70	21469.10	17640.28
1982	205654	458400	635906	765428	862214	944614	44925.01	33853.24	26656.01	21198.17	17417.66
1983	197716	453124	647772	790100	895700	65149.51	45697.09	34435.04	27114.12	21562.48	17717.00
1984	239784	569026	833828	1024228	116193.2	82948.90	58181.92	43842.98	34521.93	27453.53	22557.43
1985	326304	798048	1173448	234612.4	159738	114034.92	79986.25	60273.63	47459.41	37742.04	31011.08
1986	420778	1011178	425806.4	287302.2	195612.3	139645.17	97949.76	73810.04	58117.96	46218.25	37975.63
1987	496200	649332.4	482381.2	325474.7	221602.4	158199.16	110963.88	83616.82	65839.81	52359.05	43021.27

Setelah diketahui jumlah klaim-klaim yang akan datang, kemudian dapat dihitung estimasi cadangan klaim pada setiap *accident years*. Berikut ini contoh perhitungan dari estimasi cadangan klaim

$$\hat{R}_{1978} = 14454.77$$

$$\hat{R}_{1979} = 20590.19 + 16918.11 = 37508.30$$

Berikut ini hasil estimasi cadangan klaim pada setiap *accident years* yang kemudian dijumlahkan seluruhnya untuk memperoleh estimasi total cadangan klaim.

Accident year	Cadangan Klaim (USD)
1978	14454.77
1979	37508.30
1980	63916.06
1981	100392.00
1982	144050.09

1983	211675.25
1984	385699.94
1985	764857.68
1986	1362437.73
1987	2192790.67
Total	5277782

Dari hasil perhitungan, diketahui bahwa total cadangan klaim adalah sebesar 5277782 USD. Hal ini dapat diartikan bahwa perusahaan asuransi pada tahun 1988 perlu menyiapkan dana sebesar 5277782USD untuk dapat memenuhi kewajibannya membayar klaim-klaim yang terjadi di masa depan.

5) Menghitung P*rediction Error* dan C*onfidence Interval*

Nilai prediction error dari Metode GLMs dengan pendekatan ODP digunakan untuk mengetahui keakuratan dari hasil estimasi cadangan klaim. Dari hasil perhitungan didapat hasilnya yaitu sebagai berikut

Accident Year	*Prediction Error* (USD)	*% Prediction Error*
1978	5145.180	35.5945%
1979	7943.521	21.178%
1980	9844.203	15.401%
1981	12016.838	11.969%
1982	14119.379	9.801%
1983	17164.510	8.108%
1984	24455.473	6.340%
1985	38014.244	4.970%
1986	59598.641	4.374%
1987	111376.672	5.079%
Total	173177.855	3.281%

Berdasarkan hasil perhitungan diketahui bahwa nilai *prediction error* cenderung menurun secara signifikan tiap tahunnya. Diketahui total *prediction error* untuk Metode GLM dengan pendekatan ODP adalah sebesar 173177.855 atau 3.281%. Sehingga dapat disimpulkan bahwa estimasi cadangan klaim yang dihasilkan baik dan memberikan hasil yang cukup akurat karena memiliki nilai total prediction error yang cukup kecil. Kemudian dilakukan perhitungan confidence interval total cadangan klaim dengan tingkat signifikansi 5% dan diperoleh hasil sebagai berikut.

$$batas\ atas = 5277782 + 1.96\sqrt{\frac{5277782}{66}} = 5278336.25$$

$$batas\ bawah = 5277782 - 1.96\sqrt{\frac{5277782}{66}} = 5277227.745$$

Dari hasil perhitungan dapat diketahui bahwa cadangan klaim yang harus disiapkan oleh perusahaan asuransi yaitu berada pada rentang 5277227.745 USD hingga 5278336.25 USD. Dengan *confidence interval* ini, perusahaan asuransi dapat memiliki fleksibilitas menentukan total cadangan klaim selama masih berada dalam selang interval tersebut.

6) Perbandingan dengan Metode Chain Ladder

Jika dibandingkan dengan metode Chain Ladder, estimasi cadangan klaim yang diperoleh menggunakan metode Generalized Linear Model memiliki hasil yang hampir sama dengan metode Chain Ladder. Diketahui hasil estimasi dengan menggunakan metode Chain Ladder yaitu sebesar 5277760 USD sedangkan dengan menggunakan metode GLM diperoleh estimasi cadangan klaim sebesar 5277782 USD. Pada metode GLM dengan pendekatan Over-dispersed Poisson juga dapat diperoleh selang interval antara 5277227.745 USD hingga 5278336.25 USD yang mana tidak dapat diperoleh saat perhitungan cadangan klaim dengan metode Chain Ladder.

6. Kesimpulan

Berdasarkan penerapan metode *Generalized Linear Model* (GLM) dengan pendekatan Over-dispersed Poisson (ODP) dalam estimasi cadangan klaim diperoleh kesimpulan sebagai berikut:

- Besar inkremental klaim pada setiap *accident year* cenderung menurun seiring dengan bertambahnya waktu. Hal ini terjadi karena perusahaan asuransi cenderung membayar sebagian besar klaim-klaim pemegang polis pada saat tahun pertama dan membayar sisa besar klaim pada tahun selanjutnya.
- Hasil estimasi cadangan klaim dengan pendekatan ODP memberikan hasil yang cukup sama dengan estimasi cadangan klain dengan metode Chain Ladder
- Estimasi cadangan klaim dengan pendekatan ODP dapat menyempurnakan metode Chain Ladder yang tidak dapat menghasilkan nilai prediction error dan juga nilai *confidence interval* pada total estimasi cadangan klaim.

Daftar Pustaka

England, P. D., & Verrall, R. J. (2002). Stochastic claims reserving in general insurance. *British Actuarial Journal*, *8*(3), 443-518.

Charpentier, A. (Ed.). (2014). *Computational actuarial science with R*. CRC press.

Ogutu, J. A. (2011). *Claims reserving using overdispersed poisson model* (Doctoral dissertation).

Muharam, R. (2017). Perbandingan Estimasi Cadangan Klaim Menggunakan Metode Chain Ladder dan Generalized Linear Models (GLMs) dengan Pendekatan Over-dispersed Poisson (ODP) Pada Asuransi Umum.

Prastiwi, A. (2018). Estimasi Cadangan Klaim Incurred But Not Reported (IBNR) Menggunakan Metode Chain Ladder dan Pendekatan Over-Dispersed Poisson.

McCullangh, P., & Nelder, J. (1989). Generalized Linear Models 2nd Edition. Chapman and Hall.

Taylor, G., & McGuire, G. (2016). Stochastic Loss Reserving Using Generalized Linear Models. Virginia: Casualty Actuarial Society.