

TUGAS #1 STK731 – MODEL LINIER TERAMPAT

Tugas ini mengolah data *Beetle Mortality* seperti yang tercantum pada contoh 7.3.1 pada buku Dobson (2001) sebagai berikut:

Dose, x_i ($\log_{10}CS2mg l^{-1}$)	Number of beetles, n_i	Number killed, y_i
1.6907	59	6
1.7242	60	13
1.7552	62	18
1.7842	56	28
1.8113	63	52
1.8369	59	53
1.8610	62	61
1.8839	60	60

Model Fitting and Link Function

Pada kasus ini ingin dijelaskan hubungan antara proporsi kematian kumbang ($P_i=Y_i/n_i$) setelah lima jam diberi gas *carbon disulphide* pada beberapa dosis (x_i), sehingga $E(P_i)=\pi_i$ dan model peluang π_i adalah

$$g(\pi_i) = x_i^T \beta$$

yang dapat disederhanakan menjadi

$$\pi = x^T \beta = \beta_1 + \beta_2 x, \text{ dimana } \pi \in [0, 1]$$

Terdapat empat link function yang mungkin dapat digunakan, yaitu:

1. Logit function (LINK=LOGIT)

$$g(\pi) = \log\left(\frac{\pi}{1-\pi}\right)$$

yang merupakan invers dari fungsi sebaran logistic kumulatif

$$F(x) = \frac{1}{1+e^{-x}} = \frac{e^x}{1+e^x}$$

2. Probit (normit) function (LINK=PROBIT)

$$g(\pi) = \Phi^{-1}(\pi)$$

yang merupakan invers dari fungsi sebaran normal baku kumulatif.

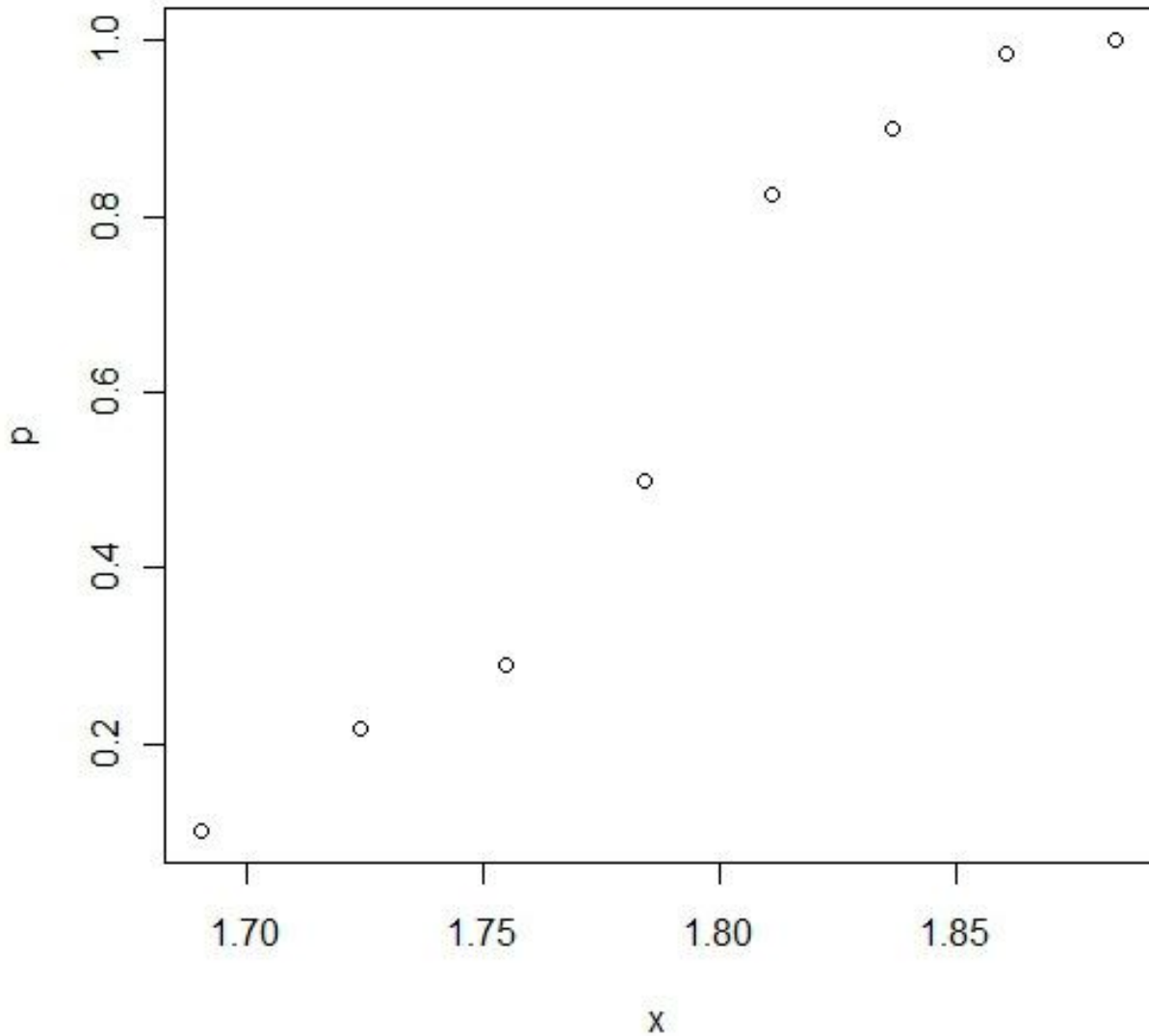
3. Complementary log-log function (LINK=CLOGLOG)

$$g(\pi) = \log(-\log(1 - \pi))$$

4. Log function (LINK=LOG)

$$g(\pi) = \log(\pi)$$

Plot data antara $p_i=y_i/n_i$ dengan x_i dapat ditunjukkan seperti gambar di bawah ini.



Untuk melihat hasil masing-masing link function, data diolah dengan PROC GENMOD pada SAS sebagai berikut:

```
data mortality;
  input x n y;
  datalines;
  1.6907 59 6
  1.7242 60 13
  1.7552 62 18
  1.7842 56 28
  1.8113 63 52
  1.8369 59 53
  1.8610 62 61
  1.8839 60 60
  ;
run;
proc genmod data=mortality;
  model y/n=x / dist=binomial
              link=logit;
  title 'Model with LOGIT';
run;
proc genmod data=mortality;
  model y/n=x / dist=binomial
              link=probit;
  title 'Model with PROBIT';
run;
proc genmod data=mortality;
  model y/n=x / dist=binomial
              link=log;
  title 'Model with LOG';
run;
proc genmod data=mortality;
  model y/n=x / dist=binomial
              link=cloglog;
  title 'Model with COMPLEMENTARY LOG-LOG';
run;
```

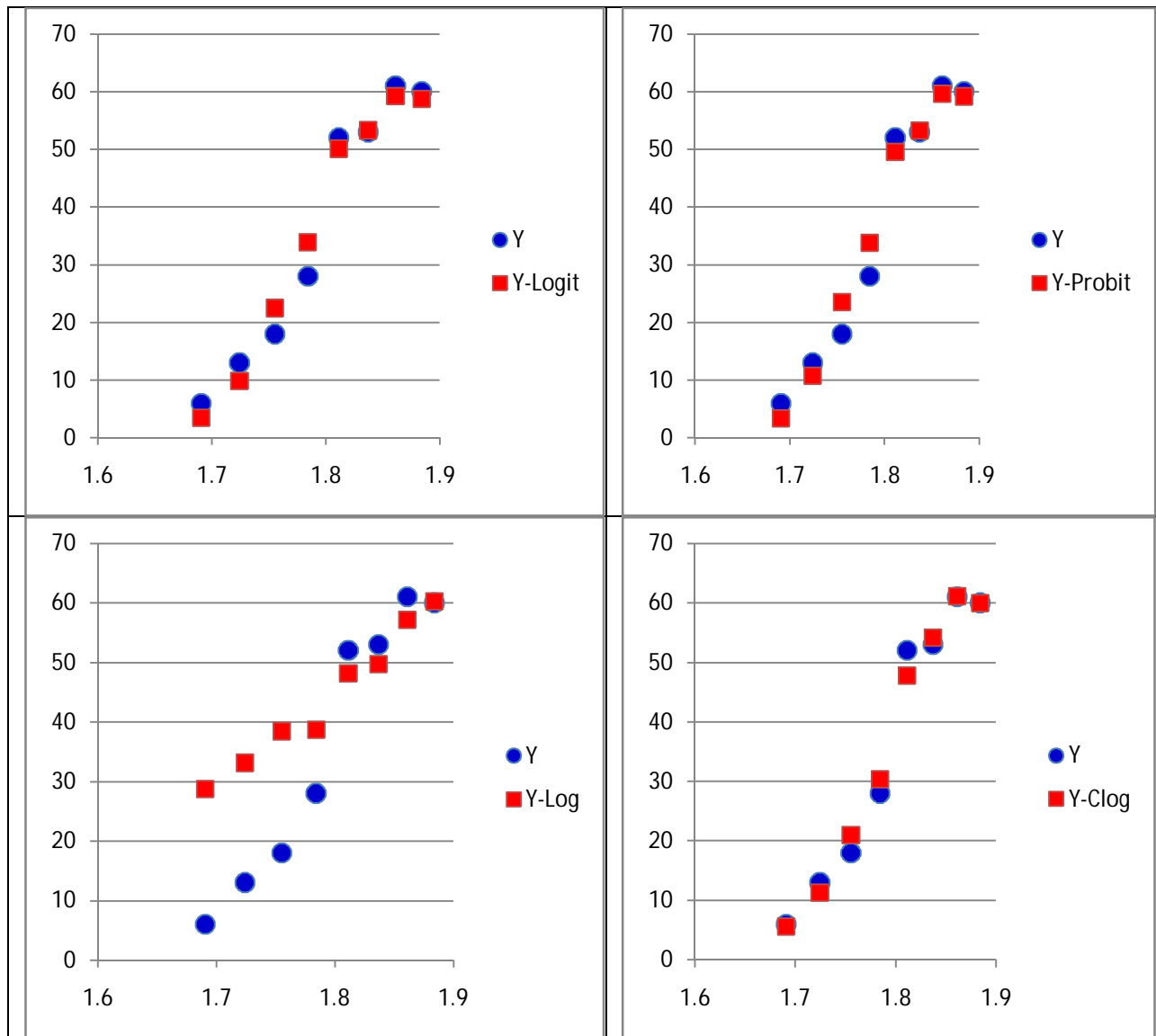
Penduga dan standard error untuk model dengan masing-masing link function adalah:

Link Function	Parameter	DF	Estimate	Std Error	Wald 95% Conf. Limits		Chi-Square	Pr>ChiSq
Logit	Intercept	1	-60.7175	5.1807	-70.8715	-50.5634	137.36	<.0001
	x	1	34.2703	2.9121	28.5626	39.9780	138.49	<.0001
Probit	Intercept	1	-34.9353	2.6395	-40.1086	-29.7619	175.18	<.0001
	x	1	19.7279	1.4841	16.8192	22.6366	176.71	<.0001
Log	Intercept	0	-7.0550	0.0000	-7.0550	-7.0550	.	.
	x	0	3.7475	0.0000	3.7475	3.7475	.	.
CLog	Intercept	1	-39.5723	3.2290	-45.9012	-33.2435	150.19	<.0001
	x	1	22.0412	1.7931	18.5268	25.5556	151.10	<.0001

Untuk memilih model "terbaik", dilakukan dengan melihat nilai deviance dan hasil nilai dugaan dari setiap observasi (\hat{Y}_i) seperti tercantum pada tabel berikut:

Y	\hat{Y}			
	Logit	Probit	Log	Cloglog
6	3.4572	3.3571	28.7442	5.5898
13	9.8409	10.7201	33.1414	11.2813
18	22.4501	23.4796	38.4649	20.9553
28	33.8964	33.8133	38.7310	30.3707
52	50.0949	49.6138	48.2300	47.7778
53	53.2904	53.3178	49.7157	54.1435
61	59.2219	59.6641	57.1816	61.1136
60	58.7428	59.2278	60.2956	59.9473
Deviance	11.2322	10.1198	.	3.4464

Berdasarkan nilai deviance dan nilai dugaan Y, maka model yang sesuai adalah model yang menggunakan link function complementary log-log (CLOGLOG). Hal ini juga dapat dilihat dari hasil plot antara X dengan Y dan \hat{Y} sebagai berikut:



Overdispersion

Overdispersion adalah munculnya keragaman yang lebih besar pada sekumpulan data dibandingkan dengan ragam yang diharapkan berdasarkan model. Overdispersion sering terjadi ketika melakukan model fitting berdasarkan sebaran Binomial atau Poisson. Implikasinya, untuk model yang benar, nilai statistic Chi-square Pearson dibagi dengan derajat bebasnya akan bernilai sama dengan 1. Overdispersion terjadi jika nilai tersebut melebihi dari 1, dan underdispersion terjadi jika nilai tersebut kurang dari 1.

Dalam regresi logistik biner, peubah respon yang diamati diasumsikan hanya memiliki dua macam kejadian, misalnya "sukses" dan "gagal". Peubah respon seperti ini sering dinamakan sebagai peubah biner. Dalam pemodelan, diasumsikan bahwa peubah biner ini saling bebas satu dengan yang lainnya, sehingga jumlah dari peubah biner akan memiliki sebaran binom. Akan tetapi dalam beberapa kasus, seringkali asumsi ini tidak terpenuhi. Secara teori permasalahan ini tidak akan mengubah nilai harapan dari sebaran binom, tetapi akan mempengaruhi keragaman dari peubah respon tersebut. Oleh karena itu, asumsi sebaran binom terhadap peubah respon mungkin tidak akan terpenuhi. Jika peubah biner tersebut berkorelasi positif, maka keragamannya akan meningkat sebesar $\sum \sum cov(y_i, y_j)$. Masalah ini sering disebut sebagai overdispersi dalam data binom. Overdispersi dapat disebabkan oleh keragaman peluang respon di dalam suatu kelompok atau korelasi antara peubah biner. Dalam prakteknya dua kejadian ini terjadi secara simultan, artinya jika terdapat korelasi antara peubah biner, maka hal ini akan membawa pada keragaman peluang respon, begitu juga sebaliknya. Overdispersi dapat terjadi dalam dua kemungkinan, yaitu pengelompokkan di dalam populasi dan pengukuran atau percobaan secara berulang pada objek yang sama.

Ada dua statistik yang digunakan untuk menguji kelayakkan model yaitu khi-kuadrat Pearson dan devians. Kedua statistik ini merupakan fungsi dari sisaan, yaitu selisih dari nilai aktual dengan nilai dugaan. Untuk suatu peubah bebas tertentu, nilai sisaan Pearson untuk amatan ke-i

didefinisikan sebagai berikut:

$$r(y_j, \hat{p}_j) = \frac{y_j - n_j \hat{p}_j}{\sqrt{n_j \hat{p}_j (1 - n_j \hat{p}_j)}}$$

sehingga khi-kuadrat Pearson dapat dinyatakan sebagai berikut:

$$X^2 = \sum_{i=1}^n r(y_j, \hat{p}_j)^2$$

selanjutnya, nilai sisaan devians untuk amatan ke-i dinyatakan sebagai berikut:

$$d(y_j, \hat{p}_j) = \pm \left\{ 2 \left[y_j \ln \left(\frac{y_j}{n_j} \right) + (n_j - y_j) \ln \left(\frac{n_j - y_j}{n_j (1 - n_j \hat{p}_j)} \right) \right] \right\}^{1/2}$$

sehingga devians dapat dinyatakan sebagai berikut:

$$D = \sum_{i=1}^n d(y_j, \hat{p}_j)^2$$

Khi-kuadrat Pearson dan devians akan mengikuti sebaran χ^2 dengan derajat bebas (n-p), dengan p adalah banyaknya parameter dalam model yang diduga. Jika model regresi logistik yang digunakan terhadap data layak, maka nilai khi-kuadrat Pearson dan devians akan mendekati nilai derajat bebasnya. Hal ini dapat dijelaskan karena nilai harapan dari sebaran χ^2 sama dengan derajat bebasnya. Jika nilai khi-kuadrat Pearson dan devians jauh lebih besar dari derajat bebasnya, maka asumsi dari keragaman binom tidak terpenuhi dan data menunjukkan overdispersi.

Salah satu cara untuk mengatasi overdispersion adalah mengalikan matrik covariance dengan parameter dispersi. Statistik Chi-square Pearson χ^2_p dan simpangannya χ^2_D adalah

$$\chi^2_p = \sum_{i=1}^m \sum_{j=1}^{k+1} \frac{(r_{ij} - n_i \hat{\pi}_{ij})^2}{n_i \hat{\pi}_{ij}}$$

$$\chi^2_D = 2 \sum_{i=1}^m \sum_{j=1}^{k+1} r_{ij} \log \left(\frac{r_{ij}}{n_i \hat{\pi}_{ij}} \right)$$

dimana m adalah banyaknya profil subpopulasi, k+1 adalah banyaknya level respon, r_{ij} adalah jumlah hasil kali frekuensi dan bobot yang terkait dengan level respon ke-j pada profil ke-i,

$$n_i = \sum_{j=1}^{k+1} r_{ij}$$

dan $\hat{\pi}_{ij}$ adalah penduga peluang level ke-j pada profil ke-i. Derajat bebas statistic tersebut adalah mk-p, dimana p adalah banyaknya parameter yang diduga. Sedangkan parameter dispersi diduga dengan:

$$\hat{\sigma}^2 = \begin{cases} \chi^2_p / (mk - p) & \text{SCALE} = \text{PEARSON} \\ \chi^2_D / (mk - p) & \text{SCALE} = \text{DEVIANANCE} \\ (\text{const})^2 & \text{SCALE} = \text{const} \end{cases}$$

Misalkan data terdiri dari n pengamatan binomial, dimana y_i/n_i adalah proporsi pengamatan ke-i, dan x_i adalah variabel penjelas. Misalkan P_i adalah peluang untuk pengamatan ke-i dengan nilai tengah dan ragam sebagai

$$E(P_i) = \pi_i \text{ dan } V(P_i) = \phi \pi_i (1 - \pi_i)$$

Williams (1982) menduga parameter skala ϕ yang tidak diketahui dengan nilai persamaan dari statistic Chi-square Pearson untuk model penuh. Misalkan w_i^* adalah bobot pengamatan ke-i, maka statistic Chi-square Pearson adalah

$$\chi^2 = \sum \frac{\omega_i^* (r_i - n_i \hat{\pi}_i)^2}{n_i \hat{\pi}_i (1 - \hat{\pi}_i)}$$

Nilai harapan dari χ^2 adalah

$$E_{\chi^2} = \sum_{i=1}^n \omega_i^* (1 - \omega_i^* v_i d_i) [1 + \phi(n_i - 1)]$$

dimana $v_i = n_i / (\pi_i (1 - \pi_i) [g'(\pi_i)]^2)$ dan d_i adalah ragam dari penduga $\hat{\alpha}_i + x_i' \hat{\beta}$. Parameter skala ϕ diduga dengan prosedur iterasi.

Pada hasil analisis data, kemunculan overdispersion terlihat pada nilai statistik dalam Criteria For Assessing Goodness Of Fit untuk masing-masing link function sebagai berikut:

Link	Criterion	DF	Value	Value/DF
Logit	Deviance	6	11.2322	1.8720
	Scaled Deviance	6	11.2322	1.8720
	Pearson Chi-Square	6	10.0268	1.6711
	Scaled Pearson X2	6	10.0268	1.6711
	Log Likelihood		-186.2354	
Probit	Deviance	6	10.1198	1.6866
	Scaled Deviance	6	10.1198	1.6866
	Pearson Chi-Square	6	9.5134	1.5856
	Scaled Pearson X2	6	9.5134	1.5856
	Log Likelihood		-185.6792	
Log	Deviance	6	0.0000	0.0000
	Scaled Deviance	6	0.0000	0.0000
	Pearson Chi-Square	6	157.4775	26.2463
	Scaled Pearson X2	6	157.4775	26.2463
	Log Likelihood		-1.79769E308	
Cloglog	Deviance	6	3.4464	0.5744
	Scaled Deviance	6	3.4464	0.5744
	Pearson Chi-Square	6	3.2947	0.5491
	Scaled Pearson X2	6	3.2947	0.5491
	Log Likelihood		-182.3425	

Pada link function Logit dan Probit terjadi overdispersion (nilai deviance dan Pearson Chi-Square lebih dari 1), sedangkan pada link function Cloglog terjadi underdispersion. Overdispersion akan menyebabkan presisi statistik uji terlalu tinggi sehingga cenderung signifikan padahal sebenarnya tidak. Sebaliknya underdispersion menyebabkan presisi statistik uji sangat rendah sehingga cenderung tidak signifikan padahal sebenarnya signifikan.

Untuk mengatasi overdispersion maupun underdispersion, digunakan pendugaan parameter skala. Oleh karena itu, analisis data dilakukan dengan memberikan pilihan SCALE pada PROC GENMOD sebagai berikut:

```
proc genmod data=mortality;
  model y/n=x / dist=binomial
            link=logit
            scale=DEVIANCE
            scale=PEARSON;
  title 'Model with LOGIT';
run;
proc genmod data=mortality;
  model y/n=x / dist=binomial
            link=probit
            scale=DEVIANCE
            scale=PEARSON;
  title 'Model with PROBIT';
run;
proc genmod data=mortality;
  model y/n=x / dist=binomial
            link=log
            scale=DEVIANCE
            scale=PEARSON;
  title 'Model with LOG';
run;
proc genmod data=mortality;
  model y/n=x / dist=binomial
            link=cloglog
            scale=DEVIANCE
            scale=PEARSON;
  title 'Model with COMPLEMENTARY LOG-LOG';
run;
```

Overdispersion dapat diatasi dengan menggunakan nilai skala tersebut di atas, yaitu nilai Scaled Deviance dan Scaled Pearson X2 yang mendekati 1 kecuali pada link function Log yang memang model menjadi tidak konvergen, seperti terlihat pada output dalam Criteria For Assessing Goodness Of Fit untuk masing-masing link function sebagai berikut:

Link	Criterion	DF	Value	Value/DF
Logit	Deviance	6	11.2322	1.8720
	Scaled Deviance	6	6.0000	1.0000
	Pearson Chi-Square	6	10.0268	1.6711
	Scaled Pearson X2	6	5.3561	0.8927
	Log Likelihood		-99.4827	
Probit	Deviance	6	10.1198	1.6866
	Scaled Deviance	6	6.0000	1.0000
	Pearson Chi-Square	6	9.5134	1.5856
	Scaled Pearson X2	6	5.6405	0.9401
	Log Likelihood		-110.0891	
Log	Deviance	6	0.0000	0.0000
	Scaled Deviance	6	0.0000	0.0000
	Pearson Chi-Square	6	157.4775	26.2463
	Scaled Pearson X2	6	157.4775	26.2463
	Log Likelihood		-1.79769E308	
Cloglog	Deviance	6	3.4464	0.5744
	Scaled Deviance	6	6.0000	1.0000
	Pearson Chi-Square	6	3.2947	0.5491
	Scaled Pearson X2	6	5.7358	0.9560
	Log Likelihood		-317.4451	

Penduga dan standard error untuk model dengan masing-masing link function setelah dilakukan pendugaan parameter skala adalah:

Link Function	Parameter	DF	Estimate	Std Error	Wald 95% Conf. Limits		Chi-Square	Pr>ChiSq
Logit	Interce	1	-60.7175	7.0884	-74.6104	-46.8245	73.37	<.0001
	x	1	34.2703	3.9845	26.4609	42.0797	73.98	<.0001
	Scale	0	1.3682	0.0000	1.3682	1.3682		
Probit	Intercept	1	-34.9353	3.4279	-41.6539	-28.2166	103.86	<.0001
	x	1	19.7279	1.9273	15.9504	23.5055	104.77	<.0001
	Scale	0	1.2987	0.0000	1.2987	1.2987		
Log	Intercept	0	-7.0550	0.0000	-7.0550	-7.0550	.	.
	x	0	3.7475	0.0000	3.7475	3.7475	.	.
	Scale	0	1.0000	0.0000	1.0000	1.0000		
Cloglog	Intercept	1	-39.5723	2.4473	-44.3689	-34.7757	261.47	<.0001
	x	1	22.0412	1.3590	19.3776	24.7047	263.06	<.0001
	Scale	0	0.7579	0.0000	0.7579	0.7579		

Algoritme untuk Fit Model

Algoritme maximum likelihood antara lain adalah menggunakan metode Fisher-scoring dan Newton-Raphson. Keduanya menduga parameter yang sama, tetapi metode Fisher-scoring berdasarkan matrik nilai harapan, sedangkan metode Newton-Raphson berdasarkan matrik pengamatan atau observasi. Namun demikian, pada kasus model binary logit, kedua matrik tersebut identik sehingga menghasilkan matrik covariance penduga yang identik pula.

Fisher-scoring

Misalkan terdapat variabel ganda $\mathbf{Z}_j = (Z_{1j}, \dots, Z_{kj})'$ sedemikian sehingga

$$Z_{ij} = \begin{cases} 1 & Y_j = i \\ 0 & \text{selainnya} \end{cases}$$

Jika π_{ij} melambangkan peluang bahwa observasi ke-j mempunyai nilai respon ke-i, maka

$$E(\mathbf{Z}_j) = \boldsymbol{\pi}_j = (\pi_{1j}, \dots, \pi_{kj})' \text{ dan } \pi_{(k+1)j} = 1 - \sum_{i=1}^k \pi_{ij}$$

Misalkan matrik covariance $\mathbf{Z}_j = \mathbf{V}_j$, dan $\boldsymbol{\theta} = (\alpha_1, \dots, \alpha_k, \beta)'$, serta \mathbf{D}_j adalah turunan bagian dari π_j terhadap $\boldsymbol{\theta}$, maka penduga parameter $\boldsymbol{\theta}$ adalah

$$\sum_j \mathbf{D}_j' \mathbf{W}_j (\mathbf{Z}_j - \boldsymbol{\pi}_j) = 0$$

dimana $W_j = w_j f_j V_j$, w_j dan f_j adalah nilai bobot dan frekuensi dari observasi ke-j. Dengan nilai awal θ_0 penduga maksimum likelihood bagi θ diperoleh melalui proses iterasi, yaitu

$$\theta_{m+1} = \theta_m + \left(\sum_j D_j' W_j D_j \right)^{-1} \sum_j D_j' W_j (Z_j - \pi_j)$$

dimana D_j , W_j , dan π_j dihitung pada nilai θ_m , sedangkan ekspresi setelah tanda plus (+) merupakan ukuran step pada proses iterasi. Proses iterasi dilakukan hingga nilai θ_{m+1} yang diperoleh konvergen ke θ_m . Penduga maksimum likelihood bagi θ adalah $\hat{\theta} = \theta_m$ dan matrik covariance dari $\hat{\theta}$ diduga dengan

$$\text{cov}(\hat{\theta}) = \left(\sum_j \hat{D}_j' \hat{W}_j \hat{D}_j \right)^{-1}$$

Newton-Raphson

Misalkan vektor parameter untuk model kumulatif adalah $\theta = (\alpha_1, \dots, \alpha_k, \beta)'$, dan untuk model logit terampat dilambangkan dengan $\theta = (\alpha_1, \dots, \alpha_k, \beta_1', \dots, \beta_k)'$. Diberikan vektor gradient (**g**) dan matrik Hessian (**H**) sebagai berikut:

$$g = \sum_j w_j f_j \frac{\partial l_j}{\partial \theta}$$

$$H = \sum_j -w_j f_j \frac{\partial^2 l_j}{\partial \theta^2}$$

dimana $l_j = \log L_j$ adalah log likelihood dari observasi ke-j. Dengan nilai awal θ_0 penduga maksimum likelihood bagi θ diperoleh melalui proses iterasi hingga konvergen, yaitu

$$\theta_{m+1} = \theta_m + H^{-1} g$$

dan matrik covariance dari $\hat{\theta}$ diduga dengan

$$\text{cov}(\hat{\theta}) = \hat{H}^{-1}$$

Untuk mengaplikasikan kedua algoritme tersebut pada data, maka dilakukan pengolahan dengan program SAS sebagai berikut:

```
proc logistic data=mortality;
  model y/n=x / link=logit ITPRINT LACKFIT TECH=NEWTON;
  title 'Model LOGIT with Newton Raphson Technique';
run;
proc logistic data=mortality;
  model y/n=x / link=logit ITPRINT LACKFIT TECH=FISHER;
  title 'Model LOGIT with Fisher Scoring';
run;
proc logistic data=mortality;
  model y/n=x / link=probit ITPRINT LACKFIT TECH=NEWTON;
  title 'Model PROBIT with Newton Raphson Technique';
run;
proc logistic data=mortality;
  model y/n=x / link=probit ITPRINT LACKFIT TECH=FISHER;
  title 'Model PROBIT with Fisher Scoring';
run;
proc logistic data=mortality;
  model y/n=x / link=cloglog ITPRINT LACKFIT TECH=NEWTON;
  title 'Model CLOGLOG with Newton Raphson Technique';
run;
proc logistic data=mortality;
  model y/n=x / link=cloglog ITPRINT LACKFIT TECH=FISHER;
  title 'Model CLOGLOG with Fisher Scoring';
run;
```

dan proses iterasi Maximum Likelihood yang dihasilkan adalah sebagai berikut:

```
METODE NEWTON RAPHSON DENGAN LINK=LOGIT
```

Iter	Ridge	-2 Log L	Intercept	x
0	0	645.441025	0.426299	0
1	0	395.942398	-39.615531	22.321622
2	0	374.092965	-54.667208	30.842343
3	0	372.485593	-60.122669	33.933246
4	0	372.470808	-60.711358	34.266871

Last Change in -2 Log L 0.0147846852

Last Evaluation of Gradient

Intercept	x
0.0028509146	0.0054776466

Convergence criterion (GCONV=1E-8) satisfied.

```
METODE FISHER SCORING DENGAN LINK=LOGIT
```

Iter	Ridge	-2 Log L	Intercept	x
0	0	645.441025	0.426299	0
1	0	395.942398	-39.615531	22.321622
2	0	374.092965	-54.667208	30.842343
3	0	372.485593	-60.122669	33.933246
4	0	372.470808	-60.711358	34.266871

Last Change in -2 Log L 0.0147846852

Last Evaluation of Gradient

Intercept	x
0.0028509146	0.0054776466

Convergence criterion (GCONV=1E-8) satisfied.

METODE NEWTON RAPHSON DENGAN LINK=PROBIT

Iter	Ridge	-2 Log L	Intercept	x
0	0	645.441025	0.266284	0
1	0	390.667213	-24.328341	13.740838
2	0	371.957157	-33.013589	18.638521
3	0	371.359432	-34.854132	19.681826
4	0	371.358334	-34.935102	19.727845

Last Change in -2 Log L 0.0010984738

Last Evaluation of Gradient

Intercept x

0.0003386506 0.0006425704

Convergence criterion (GCONV=1E-8) satisfied.

METODE FISHER SCORING DENGAN LINK=PROBIT

Iter	Ridge	-2 Log L	Intercept	x
0	0	645.441025	0.266284	0
1	0	391.812557	-24.585499	13.853815
2	0	372.758152	-32.175975	18.156667
3	0	371.376539	-34.648465	19.563453
4	0	371.358369	-34.925819	19.722434
5	0	371.358334	-34.934848	19.727698

Last Change in -2 Log L 0.0000349064

Last Evaluation of Gradient

Intercept x

0.0017097969 0.003147133

Convergence criterion (GCONV=1E-8) satisfied.

METODE NEWTON RAPHSON DENGAN LINK=CLOGLOG

Iter	Ridge	-2 Log L	Intercept	x
0	0	645.441025	-0.073815	0
1	0	396.104534	-27.475075	15.471120
2	0	366.681777	-35.782397	19.968250
3	0	364.700787	-39.204069	21.839028
4	0	364.685015	-39.568900	22.039294

Last Change in -2 Log L 0.0157718879

Last Evaluation of Gradient

Intercept x

-0.005229276 -0.008830572

Convergence criterion (GCONV=1E-8) satisfied.

METODE FISHER SCORING DENGAN LINK=CLOGLOG

Iter	Ridge	-2 Log L	Intercept	x
0	0	645.441025	-0.073815	0
1	0	387.981839	-26.154526	14.538890
2	0	367.248942	-35.115317	19.536025
3	0	364.761787	-38.818867	21.616440
4	0	364.685090	-39.550513	22.028828
5	0	364.685014	-39.572451	22.041247

Last Change in -2 Log L 0.0000755979

Last Evaluation of Gradient

Intercept x

0.000093289 0.0001438767

Convergence criterion (GCONV=1E-8) satisfied.

L A M P I R A N
OUTPUT HASIL ANALISIS
LINK FUNCTIONS

The GENMOD Procedure

Model Information

Data Set WORK.MORTALITY
Distribution Binomial
Link Function Logit
Response Variable (Events) y
Response Variable (Trials) n

Number of Observations Read 8
Number of Observations Used 8
Number of Events 291
Number of Trials 481

Criteria For Assessing Goodness Of Fit

Criterion	DF	Value	Value/DF
Deviance	6	11.2322	1.8720
Scaled Deviance	6	11.2322	1.8720
Pearson Chi-Square	6	10.0268	1.6711
Scaled Pearson X2	6	10.0268	1.6711
Log Likelihood		-186.2354	

Algorithm converged.

Analysis Of Parameter Estimates

Parameter	DF	Estimate	Standard Error	Wald 95% Confidence Limits		Chi-Square	Pr > ChiSq
Intercept	1	-60.7175	5.1807	-70.8715	-50.5634	137.36	<.0001
x	1	34.2703	2.9121	28.5626	39.9780	138.49	<.0001
Scale	0	1.0000	0.0000	1.0000	1.0000		

The scale parameter was held fixed.

The GENMOD Procedure

Model Information

Data Set WORK.MORTALITY
Distribution Binomial
Link Function Probit
Response Variable (Events) y
Response Variable (Trials) n

Number of Observations Read 8
Number of Observations Used 8
Number of Events 291
Number of Trials 481

Criteria For Assessing Goodness Of Fit

Criterion	DF	Value	Value/DF
Deviance	6	10.1198	1.6866
Scaled Deviance	6	10.1198	1.6866
Pearson Chi-Square	6	9.5134	1.5856
Scaled Pearson X2	6	9.5134	1.5856
Log Likelihood		-185.6792	

Algorithm converged.

Analysis Of Parameter Estimates

Parameter	DF	Estimate	Standard Error	Wald 95% Confidence Limits		Chi-Square	Pr > ChiSq
Intercept	1	-34.9353	2.6395	-40.1086	-29.7619	175.18	<.0001
x	1	19.7279	1.4841	16.8192	22.6366	176.71	<.0001
Scale	0	1.0000	0.0000	1.0000	1.0000		

The scale parameter was held fixed.

The GENMOD Procedure

Model Information

Data Set WORK.MORTALITY
 Distribution Binomial
 Link Function Log
 Response Variable (Events) y
 Response Variable (Trials) n

Number of Observations Read 8
 Number of Observations Used 8
 Number of Events 291
 Number of Trials 481

Criteria For Assessing Goodness Of Fit

Criterion	DF	Value	Value/DF
Deviance	6	0.0000	0.0000
Scaled Deviance	6	0.0000	0.0000
Pearson Chi-Square	6	157.4775	26.2463
Scaled Pearson X2	6	157.4775	26.2463
Log Likelihood		-1.79769E308	

ERROR: The mean parameter is either invalid or at a limit of its range for some observations.

Analysis Of Parameter Estimates

Parameter	DF	Estimate	Standard Error	Wald 95% Confidence Limits		Chi-Square	Pr > ChiSq
Intercept	0	-7.0550	0.0000	-7.0550	-7.0550	.	.
x	0	3.7475	0.0000	3.7475	3.7475	.	.
Scale	0	1.0000	0.0000	1.0000	1.0000		

The scale parameter was held fixed.

The GENMOD Procedure

Model Information

Data Set WORK.MORTALITY
 Distribution Binomial
 Link Function CLL
 Response Variable (Events) y
 Response Variable (Trials) n

Number of Observations Read 8
 Number of Observations Used 8
 Number of Events 291
 Number of Trials 481

Criteria For Assessing Goodness Of Fit

Criterion	DF	Value	Value/DF
Deviance	6	3.4464	0.5744
Scaled Deviance	6	3.4464	0.5744
Pearson Chi-Square	6	3.2947	0.5491
Scaled Pearson X2	6	3.2947	0.5491
Log Likelihood		-182.3425	

Algorithm converged.

Analysis Of Parameter Estimates

Parameter	DF	Estimate	Standard Error	Wald 95% Confidence Limits		Chi-Square	Pr > ChiSq
Intercept	1	-39.5723	3.2290	-45.9012	-33.2435	150.19	<.0001
x	1	22.0412	1.7931	18.5268	25.5556	151.10	<.0001
Scale	0	1.0000	0.0000	1.0000	1.0000		

The scale parameter was held fixed.

L A M P I R A N
OUTPUT HASIL ANALISIS
OVERDISPERSION DAN ITERASI MLE

The GENMOD Procedure

Model Information

Data Set	WORK.MORTALITY
Distribution	Binomial
Link Function	Logit
Response Variable (Events)	y
Response Variable (Trials)	n

Number of Observations Read	8
Number of Observations Used	8
Number of Events	291
Number of Trials	481

Parameter Information

Parameter	Effect
Prm1	Intercept
Prm2	x

Iteration History For Parameter Estimates

Iter	Ridge	Log Likelihood	Prm1	Prm2
0	0	-186.62336	-56.50593	31.88326
1	0	-186.23723	-60.42209	34.102896
2	0	-186.2354	-60.71594	34.269468
3	0	-186.2354	-60.71745	34.270326

Criteria For Assessing Goodness Of Fit

Criterion	DF	Value	Value/DF
Deviance	6	11.2322	1.8720
Scaled Deviance	6	6.0000	1.0000
Pearson Chi-Square	6	10.0268	1.6711
Scaled Pearson X2	6	5.3561	0.8927
Log Likelihood		-99.4827	

The GENMOD Procedure

Last Evaluation Of The Negative
Of The Gradient and Hessian

	Prm1	Prm2
Gradient	-1.019E-8	-1.955E-8
Prm1	31.240911	55.560028
Prm2	55.560028	98.873061

Algorithm converged.

Analysis Of Parameter Estimates

Parameter	DF	Estimate	Standard Error	Wald 95% Confidence Limits		Chi-Square	Pr > ChiSq
Intercept	1	-60.7175	7.0884	-74.6104	-46.8245	73.37	<.0001
x	1	34.2703	3.9845	26.4609	42.0797	73.98	<.0001
Scale	0	1.3682	0.0000	1.3682	1.3682		

The scale parameter was estimated by the square root of DEVIANCE/DOF.

The GENMOD Procedure

Model Information

```

Data Set                WORK.MORTALITY
Distribution             Binomial
Link Function           Probit
Response Variable (Events)  y
Response Variable (Trials)  n
    
```

```

Number of Observations Read      8
Number of Observations Used      8
Number of Events                 291
Number of Trials                 481
    
```

Parameter Information

```

Parameter      Effect
Prm1           Intercept
Prm2           x
    
```

Iteration History For Parameter Estimates

Iter	Ridge	Log		
		Likelihood	Prm1	Prm2
0	0	-185.79552	-33.74422	19.052012
1	0	-185.67965	-34.87984	19.695837
2	0	-185.67917	-34.93319	19.726737
3	0	-185.67917	-34.93526	19.727934
4	0	-185.67917	-34.93526	19.727934

Criteria For Assessing Goodness Of Fit

Criterion	DF	Value	Value/DF
Deviance	6	10.1198	1.6866
Scaled Deviance	6	6.0000	1.0000
Pearson Chi-Square	6	9.5134	1.5856
Scaled Pearson X2	6	5.6405	0.9401
Log Likelihood		-110.0891	

The GENMOD Procedure

Last Evaluation Of The Negative
Of The Gradient and Hessian

	Prm1	Prm2
Gradient	-2.862E-7	-5.324E-7
Prm1	108.87473	193.56578
Prm2	193.56578	344.40519

Algorithm converged.

Analysis Of Parameter Estimates

Parameter	DF	Estimate	Standard Error	Wald 95% Confidence Limits		Chi-Square	Pr > ChiSq
Intercept	1	-34.9353	3.4279	-41.6539	-28.2166	103.86	<.0001
x	1	19.7279	1.9273	15.9504	23.5055	104.77	<.0001
Scale	0	1.2987	0.0000	1.2987	1.2987		

The scale parameter was estimated by the square root of DEVIANCE/DOF.

The GENMOD Procedure

Model Information

Data Set	WORK.MORTALITY
Distribution	Binomial
Link Function	Log
Response Variable (Events)	y
Response Variable (Trials)	n

Number of Observations Read	8
Number of Observations Used	8
Number of Events	291
Number of Trials	481

Parameter Information

Parameter	Effect
Prm1	Intercept
Prm2	x

Iteration History For Parameter Estimates

Iter	Ridge	Log Likelihood	Prm1	Prm2
0	0	-1.798E308	-7.055042	3.747525
1	0	-1.798E308	-7.055042	3.747525

Criteria For Assessing Goodness Of Fit

Criterion	DF	Value	Value/DF
Deviance	6	0.0000	0.0000
Scaled Deviance	6	0.0000	0.0000
Pearson Chi-Square	6	157.4775	26.2463
Scaled Pearson X2	6	157.4775	26.2463
Log Likelihood		-1.79769E308	

Last Evaluation Of The Negative
Of The Gradient and Hessian

	Prm1	Prm2
Gradient	0	0
Prm1	0	0
Prm2	0	0

The GENMOD Procedure

ERROR: The mean parameter is either invalid or at a limit of its range for some observations.

Analysis Of Parameter Estimates

Parameter	DF	Estimate	Standard Error	Wald 95% Confidence Limits		Chi-Square	Pr > ChiSq
Intercept	0	-7.0550	0.0000	-7.0550	-7.0550	.	.
x	0	3.7475	0.0000	3.7475	3.7475	.	.
Scale	0	1.0000	0.0000	1.0000	1.0000		

The scale parameter was held fixed.

The GENMOD Procedure

Model Information

Data Set WORK.MORTALITY
 Distribution Binomial
 Link Function CLL
 Response Variable (Events) y
 Response Variable (Trials) n

Number of Observations Read 8
 Number of Observations Used 8
 Number of Events 291
 Number of Trials 481

Parameter Information

Parameter	Effect
Prm1	Intercept
Prm2	x

Iteration History For Parameter Estimates

Iter	Ridge	Log Likelihood	Prm1	Prm2
0	0	-182.35441	-39.08608	21.772613
1	0	-182.34251	-39.56903	22.039346
2	0	-182.34251	-39.57233	22.041182

Criteria For Assessing Goodness Of Fit

Criterion	DF	Value	Value/DF
Deviance	6	3.4464	0.5744
Scaled Deviance	6	6.0000	1.0000
Pearson Chi-Square	6	3.2947	0.5491
Scaled Pearson X2	6	5.7358	0.9560
Log Likelihood		-317.4451	

The GENMOD Procedure

Last Evaluation Of The Negative
Of The Gradient and Hessian

	Prm1	Prm2
Gradient	0.0000114	0.0000275
Prm1	275.95324	496.79441
Prm2	496.79441	894.91285

Algorithm converged.

Analysis Of Parameter Estimates

Parameter	DF	Estimate	Standard Error	Wald 95% Confidence Limits		Chi-Square	Pr > ChiSq
Intercept	1	-39.5723	2.4473	-44.3689	-34.7757	261.47	<.0001
x	1	22.0412	1.3590	19.3776	24.7047	263.06	<.0001
Scale	0	0.7579	0.0000	0.7579	0.7579		

The scale parameter was estimated by the square root of DEVIANCE/DOF.