

COMPLEMENTARY LOG-LOG MODEL

Under the assumption of binary response, there are two alternatives to logit model: probit model and complementary-log-log model.

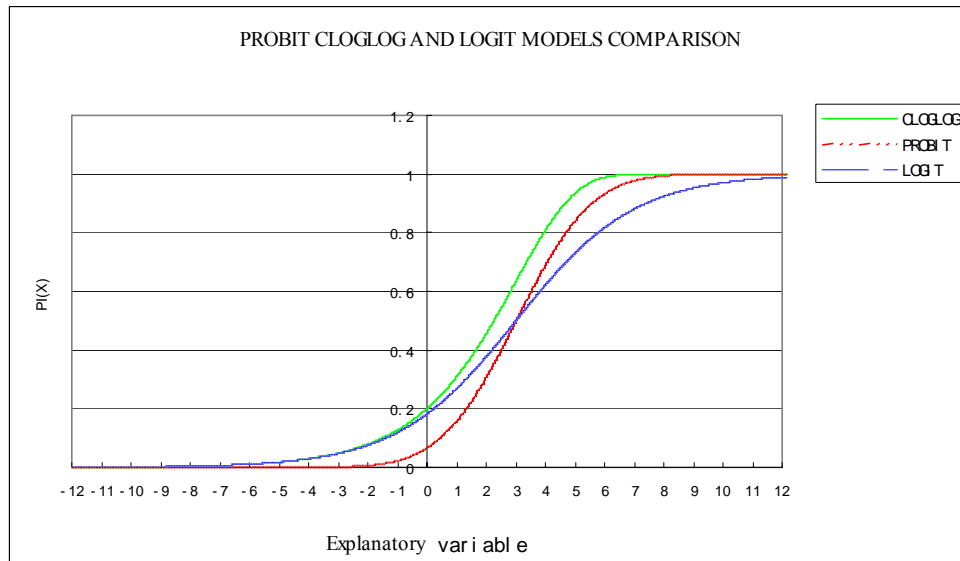
They all follow the same form

$$\pi(x) = \Phi(\alpha + \beta x) \tag{1}$$

for a continuous cdf Φ .

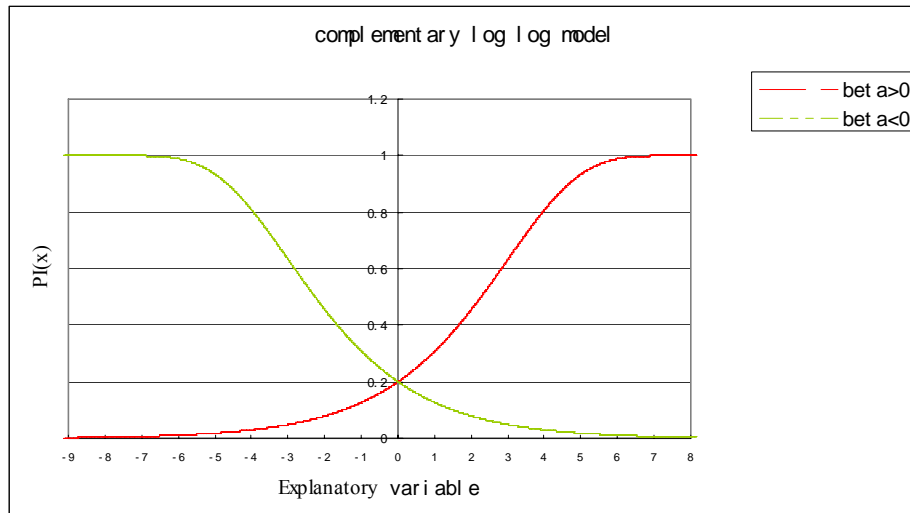
Complementary log-log model says $\log\{-\log[1 - \pi(x)]\} = X_{p \times n}^T \beta_{p \times 1}$. The expression on the left-hand side is called Complementary Log-Log transformation. Like the logit and the probit transformation, the complementary log-log transformation takes a response restricted to the (0,1) interval and converts it into something in $(-\infty, +\infty)$ interval. Here, we need mentioned that the log of $1 - \pi(x)$ is always a negative number. This is changed to a positive number before taking the log a second time. We can also write the model down like form (1), as $\pi(x) = 1 - \exp[-\exp(X_{p \times n}^T \beta_{p \times 1})]$.

Both logit and probit links have the same property, which is $\text{link}[\pi(x)] = -\text{link}[1 - \pi(x)]$. This means that the response curve for $\pi(x)$ has a symmetric appearance about the point $\pi(x) = 0.5$ and so $\pi(x)$ has the same rate for approaching 0 as well as for approaching 1. When the data given is not symmetric in the [0,1] interval and increase slowly at small to moderate value but increases sharply near 1. The logit and probit models are inappropriate. However, in this situation, the complementary log-log model might give a satisfied answer.



Unlike logit and probit the complementary log-log model is asymmetrical, it is frequently used when the probability of an event is very small or very large. Under the assumption

that the general features are not lost, let us only consider a simple complementary log-log model with one predictor $\pi(x)=1-\exp[-\exp(\alpha+\beta x)]$ in the report left. The response has an S-shaped curve, it approach 0 fairly slowly but approaching 1 quite sharply, when $\beta > 0$.



Since the log-log applies to the complement of $\pi(x)$, the link for this GLM is called the complementary log-log link.

From the three-model comparison plot above, we also can get more information. All link function are increasing, continuous, and differentiable over $0 < p < 1$, the logit and probit are almost linearly related over the interval $p \in [0.1 \ 0.9]$. For small p , the complementary log-log close to logit. The complementary log-log approaches 0 infinitely slower than any other link function.

Because logit and probit models are symmetrical, reversing the coding of the dependent variable (from 0,1 to 1,0) only change the sign of the coefficients ($\text{link}[\pi(x)] = -\text{link}[1-\pi(x)]$). For the complementary log-log model, on the other hand, reversing the coding can give us completely [different results](#). It is crucial to setup the model to predict the probability of an event, not the absence of the event.

Now let us talk more details about complementary log-log model $\pi(x)=1-\exp[-\exp(\alpha+\beta x)]$. To interpret it, we note that at x_1 and x_2 two predictors' points, we have $\log[-\log(1-\pi(x_2))] - \log[-\log(1-\pi(x_1))] = \beta(x_2 - x_1)$, so that

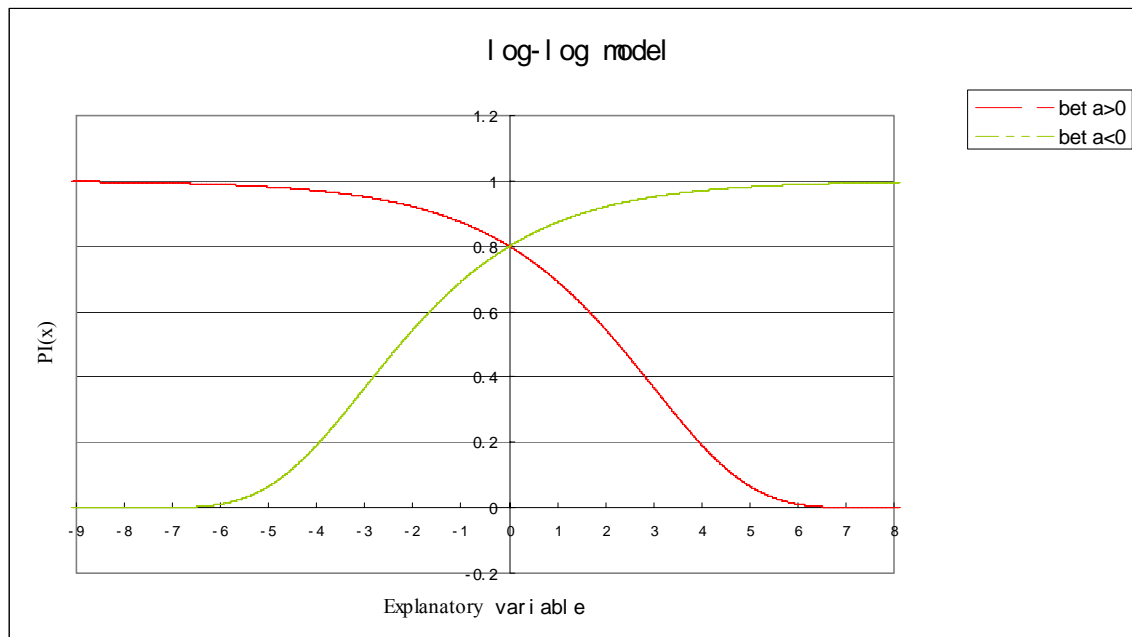
$$\frac{\log[1-\pi(x_2)]}{\log[1-\pi(x_1)]} = \exp[\beta(x_2 - x_1)]$$

and

$$1-\pi(x_2) = [1-\pi(x_1)]^{\exp[\beta(x_2-x_1)]}$$

For $x_2 - x_1 = 1$, the complement probability at x_2 equals the complement probability at x_1 raised to the power $\exp(\beta)$.

Here, we will give out another related model $\pi(x) = \exp[-\exp(\alpha + \beta x)]$, it is called log-log model. For it, $\pi(x)$ approaches 0 sharply but approaches 1 slowly. As x increases, the curve is monotone decreasing when $\beta > 0$, and monotone increasing when $\beta < 0$. In GLM form it uses the log-log link $\log[-\log(\pi(x))] = \alpha + \beta x$. When the complementary log-log model holds for the probability of a success, the log-log model holds for the probability of a failure.



Most data can be modeled well by either logit or probit model. Why should we ever want an asymmetrical model? Because the complementary log-log model is closely related to continuous-time models for the occurrence of events, so it has an important application in the area of survival analysis. Here, we just talk about a specific case with binary response.

Survival analysis with binary response model

In toxicology, binary response models describe the effect of dosage of a toxin on whether a subject dies. The tolerance distribution provides justification for model

$\pi(x) = \Phi(\alpha + \beta x)$. Let x denote the dosage level. For a randomly selected subject, let $Y=1$ if the subject dies. Suppose that the subject has tolerance T for the dosage, with $(Y=1)$ equivalent to $(T \leq x)$. For instance, an insect survive if the dosage x is less than T and Dies if the dosage is at least T . Tolerances vary among subjects, and let

$F(t)=P(T \leq t)$. For fixed dosage x , the probability a randomly selected subject dies is $\pi(x)=P(Y=1| X=x)=P(T \leq x)=F(x)$

That is the appropriate binary model is the one having the shape of the cdf F of the tolerance distribution. Let Φ denote the standard cdf for the family to which F belongs. A common standardization uses the mean and standard deviation of T , so that $\pi(x)=F(x)=\Phi[(x - \mu) / \sigma]$ Then, the model has form $\pi(x)=\Phi(\alpha+\beta x)$.

Let us use the beetle data as an example.

Beetles Killed after Exposure to Carbon Disulfide

Log Dose	Number of Beetles	Number Killed	Fitted Values		
			Comp. Log-Log	Probit	Logit
1.691	59	6	5.7	3.4	35
1.724	60	13	11.3	10.7	9.8
1.755	62	18	20.9	23.4	22.4
1.784	56	28	30.3	33.8	33.9
1.811	63	52	47.7	49.6	50.0
1.837	59	53	54.2	53.4	53.3
1.861	62	61	61.1	59.7	59.2
1.884	60	60	59.9	59.2	58.8

In the table, we find that the underlying cdf of number killed increases moderately before $x=1.811$, then there is a big jump on the number of beetles killed.

SAS result:

For logistic proc:

the logit model for beetles data 31
21:41 Sunday, March 14, 2004

The LOGISTIC Procedure

Model Fit Statistics

Criterion	Intercept and Covariates	
	Intercept Only	Intercept and Covariates
AIC	647.441	376.354
SC	651.617	384.706
-2 Log L	645.441	372.354

Testing Global Null Hypothesis: BETA=0

Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	273.0869	1	<.0001
Score	227.5973	1	<.0001
Wald	138.5112	1	<.0001

The LOGISTIC Procedure

Analysis of Maximum Likelihood Estimates

Parameter	DF	Standard Estimate	Wald Error	Chi-Square	Pr > ChiSq
-----------	----	-------------------	------------	------------	------------

Intercept	1	-60.7339	5.1814	137.3964	<.0001
logdose	1	34.2824	2.9129	138.5112	<.0001

the complementary log log model for beetles data

33

21:41 Sunday, March 14, 2004

The LOGISTIC Procedure

Model Fit Statistics

Criterion	Intercept and Covariates	
	Intercept Only	Intercept and Covariates
AIC	647.441	368.753
SC	651.617	377.105
-2 Log L	645.441	364.753

Testing Global Null Hypothesis: BETA=0

Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	280.6881	1	<.0001
Score	227.5973	1	<.0001
Wald	150.0854	1	<.0001

The LOGISTIC Procedure

Analysis of Maximum Likelihood Estimates

Parameter	DF	Standard Estimate	Wald Error	Chi-Square	Pr > ChiSq
logdose	1	22.0148	1.7970	150.0854	<.0001

the probit model for beetles data

35

21:41 Sunday, March 14, 2004

The LOGISTIC Procedure

Model Fit Statistics

Criterion	Intercept and Covariates	
	Intercept Only	Intercept and Covariates
AIC	647.441	375.226
SC	651.617	383.577
-2 Log L	645.441	371.226

Testing Global Null Hypothesis: BETA=0

Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	274.2155	1	<.0001
Score	227.5973	1	<.0001
Wald	175.9925	1	<.0001

The LOGISTIC Procedure

Analysis of Maximum Likelihood Estimates

Parameter	DF	Standard Estimate	Wald Error	Chi-Square	Pr > ChiSq

logdose 1 19.7408 1.4880 175.9925 <.0001

For logistic proc:

The SAS System 23:18 Monday, March 15, 2004 4

The GENMOD Procedure

Model Information

Data Set WORK.BEETLES1
Distribution Binomial
Link Function Probit
Response Variable (Events) nkilled
Response Variable (Trials) nbeetles
Observations Used 8
Number Of Events 291
Number Of Trials 481

Criteria For Assessing Goodness Of Fit

Criterion	DF	Value	Value/DF
Deviance	6	9.9870	1.6645
Scaled Deviance	6	9.9870	1.6645
Pearson Chi-Square	6	9.3690	1.5615
Scaled Pearson X2	6	9.3690	1.5615
Log Likelihood		-185.6128	

Algorithm converged.

Analysis Of Parameter Estimates

Parameter	DF	Standard Estimate	Wald Error	95% Confidence Limits	Chi-Square	Pr > ChiSq
Intercept	1	-34.9561	2.6413	-40.1330 -29.7793	175.15	<.0001
logdose	1	19.7410	1.4853	16.8300 22.6521	176.66	<.0001
Scale	0	1.0000	0.0000	1.0000 1.0000		

NOTE: The scale parameter was held fixed.

The SAS System 23:18 Monday, March 15, 2004 5

The GENMOD Procedure

Model Information

Data Set WORK.BEETLES1
Distribution Binomial
Link Function CLL
Response Variable (Events) nkilled
Response Variable (Trials) nbeetles
Observations Used 8
Number Of Events 291
Number Of Trials 481

Criteria For Assessing Goodness Of Fit

Criterion	DF	Value	Value/DF
Deviance	6	3.5143	0.5857
Scaled Deviance	6	3.5143	0.5857
Pearson Chi-Square	6	3.3592	0.5599
Scaled Pearson X2	6	3.3592	0.5599
Log Likelihood		-182.3765	

Algorithm converged.

Analysis Of Parameter Estimates

Parameter	DF	Standard Estimate	Wald Error	95% Confidence Limits	Chi-Square	Pr > ChiSq
Intercept	1	-39.5223	3.2229	-45.8391 -33.2055	150.38	<.0001
logdose	1	22.0148	1.7899	18.5067 25.5228	151.28	<.0001
Scale	0	1.0000	0.0000	1.0000 1.0000		

NOTE: The scale parameter was held fixed.

The SAS System 23:18 Monday, March 15, 2004 6

The GENMOD Procedure

Model Information

Data Set WORK.BEETLES1
 Distribution Binomial
 Link Function Logit
 Response Variable (Events) nkilled
 Response Variable (Trials) nbeetles
 Observations Used 8
 Number Of Events 291
 Number Of Trials 481

Criteria For Assessing Goodness Of Fit

Criterion	DF	Value	Value/DF
Deviance	6	11.1156	1.8526
Scaled Deviance	6	11.1156	1.8526
Pearson Chi-Square	6	9.9067	1.6511
Scaled Pearson X2	6	9.9067	1.6511
Log Likelihood		-186.1771	

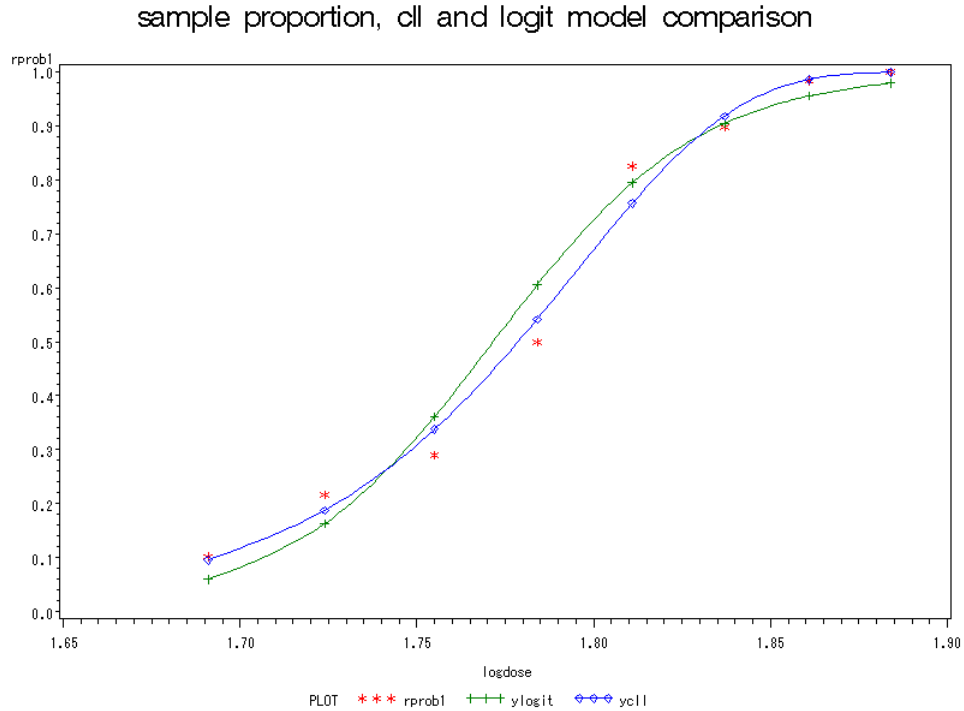
Algorithm converged.

Analysis Of Parameter Estimates

Parameter	DF	Standard Estimate	Wald Error	95% Confidence Limits	Chi-Square	Pr > ChiSq
Intercept	1	-60.7401	5.1819	-70.8964 -50.5838	137.40	<.0001
logdose	1	34.2859	2.9132	28.5761 39.9957	138.51	<.0001
Scale	0	1.0000	0.0000	1.0000 1.0000		

NOTE: The scale parameter was held fixed.

SAS plot:



For the beetles mortality data, the complementary log-log model has ML estimate model $\log[-\log(1-\pi(x))]=-39.5224+22.0148 \times \text{logdose}$. At dosage=1.7 the fitted probability

of survival is $1 - \hat{\pi}(x) = \exp[-\exp(-39.5224 + 22.0148 \times 1.7)] = 0.884445$ where at dosage =1.8 it is 0.3296029, and at dosage =1.9, it is 4.39966e-05, the probability of survival at dosage+0.1 equals the probability at dosage raised to the power

$\text{Exp}(22.0148 \times 0.1) = 9.03838$. For instance, $0.3296029 \approx (0.884445)^{9.03838}$
 $(0.3296027 = (0.884445)^{9.03838})$

Underlying the LOGISTIC proc:				
Model	intercept	logdose	Standard Error	AIC with Intercept and Covariates
logit	-60.7339	34.2824	5.1814/2.9129	376.354
probit	-34.9557	19.7408	2.6490/1.4880	375.226
Complementary log-log	-39.5224	22.0148	3.2356/1.7970	368.753

Underlying the GENMOD proc:					
Model	intercept	logdose	Standard Error	Deviance-G ²	DF
logit	-60.7401	34.2859	5.1819 /2.9132	11.1156	6
probit	-34.9561	19.7410	2.6413 /1.4853	9.9870	6
Complementary log-log	-39.5223	22.0148	3.2229/1.7899	3.5143	6

From the table we can find that under logistic proc, complementary log-log has the smallest AIC =368.753, under Genmod proc, it is still complementary log-log that has the smallest $G^2 = 3.5143$.

The last few things I need mentioned are there are

The reason is that GENMOD uses the Newton-Raphson algorithm to get the ML estimates, and LOGISTIC uses iteratively reweighed least squares(also called Fisher scoring), these two algorithms are equivalent for logit models but diverge for any other model.(That is because logit is the unique “canonical ” link function).

(With the coefficients, we can see with the one unit change on x, the logit will

????, the compare between coefficients and standard error)

Finally, we need mentioned that complementary log-log model is not only used for binary respond but also can be used for ordinal responses with cumulative link, the form is

$$\log\{-\log[1 - P(Y \leq j | x)]\} = \alpha_j + \beta^T x$$

and the ordinal model using this link is sometimes called a Proportional hazards model for survival data to handle grouped survival times.

model using cumulative link

Bookmark:

Proof for symmetric property of logit distribution and probit model:

$$\begin{aligned}\text{logit}[\pi(x)] &= \log[\pi(x)/(1-\pi(x))] \\ &= -\log[(1-\pi(x))/\pi(x)] \\ &= -\text{logit}[1-\pi(x)]\end{aligned}$$

$$\Phi^{-1}[\pi(x)] = \alpha + \beta x$$

$$\Phi^{-1}[1 - \pi(x)] = -(\alpha + \beta x)$$

$$\Phi^{-1}[\pi(x)] = -\Phi^{-1}[1 - \pi(x)]$$

$$\log\{-\log[\pi(x)]\} = \log\{-\log(1 - \exp[-\exp(\alpha + \beta x)])\}$$

The model says that

for large p, log-log close to logit

/*****storage plant*****/

$$F(x) = \exp\{-\exp[-(x-a)/b]\}$$

$$\log[-\log(1 - \pi(x_2))] - \log[-\log(1 - \pi(x_1))] = \beta(x_2 - x_1)$$

$$\frac{\log[1 - \pi(x_2)]}{\log[1 - \pi(x_1)]} = \exp[\beta(x_2 - x_1)]$$

$$1 - \pi(x_2) = [1 - \pi(x_1)]^{\exp[\beta(x_2 - x_1)]}$$

$$\text{link}[\pi(x)] = -\text{link}[1 - \pi(x)]$$

$$\pi(x) = \exp[-\exp(\alpha + \beta x)]$$

$$\log[-\log(1 - \pi(x))] = \alpha + \beta x$$

$$F(x) = \exp\{-\exp[-(x-a)/b]\}$$

$$x_2 - x_1 = 1$$

$$\pi(x) = \exp[-\exp(\alpha + \beta x)]$$

$$F(t) = P(T \leq t)$$

$$\pi(x) = P(Y=1 | X=x) = P(T \leq x) = F(x)$$

Φ

$$\pi(x) = F(x) = \Phi[(x - \mu) / \sigma]$$

$$\pi(x) = 1 - \exp[-\exp(\alpha + \beta x)]$$

$$\pi(x) = 1 - \exp[-\exp(\alpha + \beta x)]$$

$$\text{link}[\pi(x)] = -\text{link}[1 - \pi(x)]$$

$$\begin{aligned} \text{logit}[\pi(x)] &= \log[\pi(x)/(1 - \pi(x))] \\ &= -\log[(1 - \pi(x))/\pi(x)] \\ &= -\text{logit}[1 - \pi(x)] \end{aligned}$$

$$\log[-\log(\pi(x))] = \beta^T x$$

$$p \in [0.1 \ 0.9]$$

$$\pi(x) = 1 - \exp[-\exp(X_{p \times n}^T \beta_{p \times 1})]$$

$$(-\infty, +\infty)$$

$$1 - \hat{\pi}(x) = \exp[-\exp(-39.5224 + 22.0148 \times 1.7)]$$

$$\exp(\beta)$$

$$\Phi^{-1}[\pi(x)] = \alpha + \beta x$$

$$\Phi^{-1}[1 - \pi(x)] = -(\alpha + \beta x)$$

$$\Phi^{-1}[\pi(x)] = -\Phi^{-1}[1 - \pi(x)]$$

$$\log\{-\log[\pi(x)]\} = \log\{-\log(1 - \exp[-\exp(\alpha + \beta x)])\}$$

$$\log[-\log(1 - \pi(x))] = -39.5224 + 22.0148 \times \text{logdose}$$

$$\text{Exp}(22.0148 \times 0.1) = 9.03838$$

$$0.3296029 \approx (0.884445)^{9.03838}$$

$$0.3296027 = (0.884445)^{9.03838}$$

$$\text{Deviance-G}^2 = 3.5143$$

$$\log\{-\log[1 - P(Y \leq j | x)]\} = \alpha_j + \beta^T x$$

$$\log[-\log(1 - \pi(x_2))] - \log[-\log(1 - \pi(x_1))] = \beta(x_2 - x_1)$$

$$\frac{\log[1 - \pi(x_2)]}{\log[1 - \pi(x_1)]} = \exp[\beta(x_2 - x_1)]$$

$$1 - \pi(x_2) = [1 - \pi(x_1)]^{\exp[\beta(x_2 - x_1)]}$$

$$\text{link}[\pi(x)] = -\text{link}[1 - \pi(x)]$$

$$\pi(x) = \exp[-\exp(\alpha + \beta x)]$$

$$\log[-\log(\pi(x))] = \alpha + \beta x$$

$$F(x) = \exp\{-\exp[-(x-a)/b]\}$$

$$x_2 - x_1 = 1$$

$$\pi(x) = \exp[-\exp(\alpha + \beta x)]$$

$$F(t) = P(T \leq t)$$

$$\pi(x) = P(Y=1 | X=x) = P(T \leq x) = F(x)$$

Φ

$$\pi(x) = F(x) = \Phi[(x - \mu) / \sigma]$$

$$\pi(x) = 1 - \exp[-\exp(\alpha + \beta x)]$$

$$\text{link}[\pi(x)] = -\text{link}[1 - \pi(x)]$$

$$\begin{aligned} \text{logit}[\pi(x)] &= \log[\pi(x)/(1 - \pi(x))] \\ &= -\log[(1 - \pi(x))/\pi(x)] \\ &= -\text{logit}[1 - \pi(x)] \end{aligned}$$

$$\log[-\log(\pi(x))] = \beta^T x$$

$$p \in [0.1 \ 0.9]$$

$$\pi(x) = 1 - \exp[-\exp(X_{p \times n}^T \beta_{p \times 1})]$$

$$(-\infty, +\infty)$$

/***/

APPENDIX I

- For logistic proc

```
/*the begin of sas code
the code is used for two goal, so there is some result
that got from it and use it back to the code again fro plotting
*/
data beetles1;
input logdose nbeetles nkilled;
nsurvive=nbeetles-nkilled;
/*for sample proportion*/
rprobl=nkilled/nbeetles;
/*for logit model*/
ylogit=exp(-60.7339+34.2824*logdose)/(1+exp(-60.7339+34.2824*logdose));
/*for complementary log log*/
yc11= 1-exp(-exp(-39.5224+22.0148*logdose));
  datalines;
1.691 59 6
1.724 60 13
1.755 62 18
1.784 56 28
1.811 63 52
1.837 59 53
1.861 62 61
1.884 60 60
;
/**using logistic proc for test comparing the three kinds of model */
proc logistic data= beetles1; /*for logit*/
model nkilled/nbeetles = logdose ;
title 'the logit model for beetles data';

proc logistic data= beetles1; /*for cloglog*/
model nkilled/nbeetles = logdose / link=cloglog OUTROC=c11plotd;
title 'the complementary log log model for beetles data';

proc logistic data= beetles1; /*for probit*/
model nkilled/nbeetles = logdose / link=probit;
title 'the probit model for beetles data';
/*plot a curve*/

symbol1 color=red value=star interpol=NONE height=1 width=1;
symbol2 color=green value=plus interpol=spline height=1 width=1;
symbol3 color=blue value=DIAMOND interpol=spline height=1 width=1;
proc gplot data=beetles1;
plot rprobl*logdose ylogit*logdose yc11*logdose/haxis=1.65 to 1.90 by
.05 overlay legend=legend2;
title 'sample proportion, c11 and logit model comparison';
```

```
run;
/**the end of using logistic for test comparing the three kind of
model */
/*the end of sas code*/
```

```
/*result begin*/
/*for logistic proc*/
```

```
the logit model for beetles data 31
21:41 Sunday, March 14, 2004
```

The LOGISTIC Procedure

Model Information

```
Data Set          WORK.BEETLES1
Response Variable (Events)  nKilled
Response Variable (Trials)  nbeetles
Number of Observations     8
Model                     binary logit
Optimization Technique     Fisher's scoring
```

Response Profile

Ordered Value	Binary Outcome	Total Frequency
1	Event	291
2	Nonevent	190

Model Convergence Status

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

Model Fit Statistics

Criterion	Intercept and Covariates	
	Intercept Only	Intercept and Covariates
AIC	647.441	376.354
SC	651.617	384.706
-2 Log L	645.441	372.354

Testing Global Null Hypothesis: BETA=0

Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	273.0869	1	<.0001
Score	227.5973	1	<.0001
Wald	138.5112	1	<.0001

```
the logit model for beetles data 32
21:41 Sunday, March 14, 2004
```

The LOGISTIC Procedure

Analysis of Maximum Likelihood Estimates

Parameter	DF	Standard Estimate	Standard Error	Wald	
				Chi-Square	Pr > ChiSq
Intercept	1	-60.7339	5.1814	137.3964	<.0001
logdose	1	34.2824	2.9129	138.5112	<.0001

Odds Ratio Estimates

Effect	Point Estimate	95% Wald Confidence Limits	
logdose	>999.999	>999.999	>999.999

Association of Predicted Probabilities and Observed Responses

Percent Concordant 87.0 Somers' D 0.802
 Percent Discordant 6.8 Gamma 0.856
 Percent Tied 6.3 Tau-a 0.384
 Pairs 55290 c 0.901

the complementary log log model for beetles data 33
 21:41 Sunday, March 14, 2004

The LOGISTIC Procedure

Model Information

Data Set WORK.BEETLES1
 Response Variable (Events) nKilled
 Response Variable (Trials) nbeetles
 Number of Observations 8
 Model binary cloglog
 Optimization Technique Fisher's scoring

Response Profile

Ordered Value	Binary Outcome	Total Frequency
1	Event	291
2	Nonevent	190

Model Convergence Status

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

Model Fit Statistics

Criterion	Intercept and Covariates	
	Intercept Only	Intercept and Covariates
AIC	647.441	368.753
SC	651.617	377.105
-2 Log L	645.441	364.753

Testing Global Null Hypothesis: BETA=0

Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	280.6881	1	<.0001
Score	227.5973	1	<.0001
Wald	150.0854	1	<.0001

the complementary log log model for beetles data 34
 21:41 Sunday, March 14, 2004

The LOGISTIC Procedure

Analysis of Maximum Likelihood Estimates

Parameter	DF	Standard Estimate	Wald Error	Chi-Square	Pr > ChiSq
Intercept	1	-39.5224	3.2356	149.1990	<.0001
logdose	1	22.0148	1.7970	150.0854	<.0001

Association of Predicted Probabilities and Observed Responses

Percent Concordant 87.0 Somers' D 0.802
 Percent Discordant 6.8 Gamma 0.856
 Percent Tied 6.3 Tau-a 0.384
 Pairs 55290 c 0.901

the probit model for beetles data 35
 21:41 Sunday, March 14, 2004

The LOGISTIC Procedure

Model Information

Data Set WORK.BEETLES1
Response Variable (Events) nkilled
Response Variable (Trials) nbeetles
Number of Observations 8
Model binary probit
Optimization Technique Fisher's scoring

Response Profile

Ordered Value	Binary Outcome	Total Frequency
1	Event	291
2	Nonevent	190

Model Convergence Status

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

Model Fit Statistics

Criterion	Intercept and Covariates	
	Intercept Only	Intercept and Covariates
AIC	647.441	375.226
SC	651.617	383.577
-2 Log L	645.441	371.226

Testing Global Null Hypothesis: BETA=0

Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	274.2155	1	<.0001
Score	227.5973	1	<.0001
Wald	175.9925	1	<.0001

the probit model for beetles data 36
21:41 Sunday, March 14, 2004

The LOGISTIC Procedure

Analysis of Maximum Likelihood Estimates

Parameter	DF	Standard Estimate	Wald Error	Chi-Square	Pr > ChiSq
Intercept	1	-34.9557	2.6490	174.1309	<.0001
logdose	1	19.7408	1.4880	175.9925	<.0001

Association of Predicted Probabilities and Observed Responses

Percent Concordant	87.0	Somers' D	0.802
Percent Discordant	6.8	Gamma	0.856
Percent Tied	6.3	Tau-a	0.384
Pairs	55290	c	0.901

• For Genmod Proc
/*the begin of sas code*/
data beetles1;
input logdose nbeetles nkilled;

datalines;

```

1.691 59 6
1.724 60 13
1.755 62 18
1.784 56 28
1.811 63 52
1.837 59 53
1.861 62 61
1.884 60 60
;

```

```

/*the end of sas code*/

```

```

proc genmod data= beetles1;
model nkilled/nbeetles= logdose /dist=bin link=probit;

```

```

proc genmod data= beetles1;
model nkilled/nbeetles= logdose /dist=bin link=cloglog;

```

```

proc genmod data= beetles1;
model nkilled/nbeetles= logdose /dist=bin link=logit;
quit;
run;

```

```

/**the end of using genmode for test comparing the three kind of
model */

```

```

*****

```

The SAS System 23:18 Monday, March 15, 2004 4

The GENMOD Procedure

Model Information

```

Data Set          WORK.BEETLES1
Distribution       Binomial
Link Function      Probit
Response Variable (Events) nkilled
Response Variable (Trials) nbeetles
Observations Used      8
Number Of Events      291
Number Of Trials      481

```

Criteria For Assessing Goodness Of Fit

Criterion	DF	Value	Value/DF
Deviance	6	9.9870	1.6645
Scaled Deviance	6	9.9870	1.6645
Pearson Chi-Square	6	9.3690	1.5615
Scaled Pearson X2	6	9.3690	1.5615
Log Likelihood		-185.6128	

Algorithm converged.

Analysis Of Parameter Estimates

Parameter	DF	Standard Estimate	Wald Error	95% Confidence Limits	Chi-Square	Pr > ChiSq
Intercept	1	-34.9561	2.6413	-40.1330 -29.7793	175.15	<.0001
logdose	1	19.7410	1.4853	16.8300 22.6521	176.66	<.0001
Scale	0	1.0000	0.0000	1.0000 1.0000		

NOTE: The scale parameter was held fixed.

The SAS System 23:18 Monday, March 15, 2004 5

The GENMOD Procedure

Model Information

```

Data Set          WORK.BEETLES1
Distribution       Binomial
Link Function      CLL
Response Variable (Events) nkilled

```


Response Variable (Trials) nbeetles
 Observations Used 8
 Number Of Events 291
 Number Of Trials 481

Criteria For Assessing Goodness Of Fit

Criterion	DF	Value	Value/DF
Deviance	6	3.5143	0.5857
Scaled Deviance	6	3.5143	0.5857
Pearson Chi-Square	6	3.3592	0.5599
Scaled Pearson X2	6	3.3592	0.5599
Log Likelihood		-182.3765	

Algorithm converged.

Analysis Of Parameter Estimates

Parameter	DF	Standard Estimate	Wald Error	95% Confidence Limits	Chi-Square	Pr > ChiSq
Intercept	1	-39.5223	3.2229	-45.8391 -33.2055	150.38	<.0001
logdose	1	22.0148	1.7899	18.5067 25.5228	151.28	<.0001
Scale	0	1.0000	0.0000	1.0000 1.0000		

NOTE: The scale parameter was held fixed.

The SAS System 23:18 Monday, March 15, 2004 6

The GENMOD Procedure

Model Information

Data Set WORK.BEETLES1
 Distribution Binomial
 Link Function Logit
 Response Variable (Events) nkilled
 Response Variable (Trials) nbeetles
 Observations Used 8
 Number Of Events 291
 Number Of Trials 481

Criteria For Assessing Goodness Of Fit

Criterion	DF	Value	Value/DF
Deviance	6	11.1156	1.8526
Scaled Deviance	6	11.1156	1.8526
Pearson Chi-Square	6	9.9067	1.6511
Scaled Pearson X2	6	9.9067	1.6511
Log Likelihood		-186.1771	

Algorithm converged.

Analysis Of Parameter Estimates

Parameter	DF	Standard Estimate	Wald Error	95% Confidence Limits	Chi-Square	Pr > ChiSq
Intercept	1	-60.7401	5.1819	-70.8964 -50.5838	137.40	<.0001
logdose	1	34.2859	2.9132	28.5761 39.9957	138.51	<.0001
Scale	0	1.0000	0.0000	1.0000 1.0000		

NOTE: The scale parameter was held fixed.

Reference:
 LOGISTIC REGRESSION USING THE SAS SYSTEM
 Theory and Application
 --Paul D. Allison

SAS ONLINE TUTOR SOFTWARE

