

BIOS 6612 Lecture 8

Multiple Logistic Regression Deviance

Review (Lecture 7)/ Current (Lecture 8)/ Preview (Lecture 9)

- Lecture 7: Logistic Regression VI
 - Goodness of fit
 - Deviance Chi- Square
 - Compares the model of interest to the “saturated model”
 - Homer-Lemshow statistic (not preferred)
 - Predictive Power
 - Generalized R^2 , Max-rescaled R^2 (based on log-likelihood)
 - Somer's D, Gamma, Tau-a, c-index (based on predicted probabilities)
- Lecture 8: Logistic Regression VII
 - Multiple logistic regression
 - Grouped data in multiple logistic regression
 - Deviance
 - Models and submodels
- Lecture 9: Categorical Outcome
 - Horseshoe crab data

Multiple Logistic Regression

- Assume $\pi_i = \Pr[Y_i = 1]$ and $\mathbf{X}_i\boldsymbol{\beta} = \beta_0 + \beta_1 X_{1i} + \cdots + \beta_p X_{pi}$

- Log-odds form:

$$\log\left(\frac{\pi_i}{1 - \pi_i}\right) = \mathbf{X}_i\boldsymbol{\beta}$$

- Odds form:

$$\frac{\pi_i}{1 - \pi_i} = \exp(\mathbf{X}_i\boldsymbol{\beta})$$

- $\Pr[Y_i=1]$

$$\pi_i = \Pr[Y_i = 1] = \frac{\exp(\mathbf{X}_i\boldsymbol{\beta})}{1 + \exp(\mathbf{X}_i\boldsymbol{\beta})}$$

- Interpretation, as before:

- A unit increase in X_j multiplies the odds by $\exp(\beta_j)$ for $j=1,\dots,p$
- A unit increase in X_j increases the log-odds by β_j for $j=1,\dots,p$

Multiple Logistic Regression: Grouped and Ungrouped Data

- To group individuals in multiple logistic regression, the individuals must have the same values for all the covariates
 - i.e. dose=2, gender=female
- Each distinct set of covariates is called a covariate pattern
- If there are m distinct covariate patterns, we record for each pattern the number of individuals having that pattern (n_j) and the number of “successes” (s_j) for $j=1, \dots, m$
- For ungrouped data, the log-likelihood where $z_i = \mathbf{X}_i \boldsymbol{\beta} = \beta_0 + \beta_1 X_{1i} + \dots + \beta_p X_{pi}$

$$\begin{aligned}
 l &= \sum_{i=1}^n \left[Y_i \log \left(\frac{e^{z_i}}{1 + e^{z_i}} \right) + (1 - Y_i) \log \left(1 - \frac{e^{z_i}}{1 + e^{z_i}} \right) \right] = \sum_{i=1}^n \left[Y_i \log \left(\frac{e^{z_i}}{1 + e^{z_i}} \right) + (1 - Y_i) \log \left(\frac{1}{1 + e^{z_i}} \right) \right] \\
 &= \sum_{i=1}^n \left[Y_i \log(e^{z_i}) - Y_i \log(1 + e^{z_i}) + (1 - Y_i) \log(1) - (1 - Y_i) \log(1 + e^{z_i}) \right] = \sum_{i=1}^n \left[Y_i z_i - \log(1 + e^{z_i}) \right]
 \end{aligned}$$

- For grouped data, the log-likelihood=

$$l = \sum_{j=1}^m \left[s_j z_j - n_j \log(1 + e^{z_j}) \right]$$

- The j^{th} covariate pattern is (X_{1j}, \dots, X_{pj}) for z_j

Likelihood for Grouped Data

- For m distinct covariate patterns, assuming the success $s_j \sim \text{Binomial}(n_j, \pi_j)$
- The likelihood L is a function of each covariate pattern is

$$L(\pi_1, \dots, \pi_m) = \prod_{j=1}^m \binom{n_j}{s_j} \pi_j^{s_j} (1 - \pi_j)^{n_j - s_j} \propto \prod_{j=1}^m \pi_j^{s_j} (1 - \pi_j)^{n_j - s_j}$$

- Ignoring the sections no depending on π_j , the log-likelihood is

$$l(\pi_1, \dots, \pi_m) = \sum_{j=1}^m \left[s_j \log(\pi_j) - (n_j - s_j) \log(1 - \pi_j) \right]$$

- The maximum value of the log-likelihood occurs when $\hat{\pi}_j = \frac{s_j}{n_j}$
 - Most software use 0 for the log L if $s_j = n_j$ or 0
- This is the log-likelihood for the saturated model

Deviance

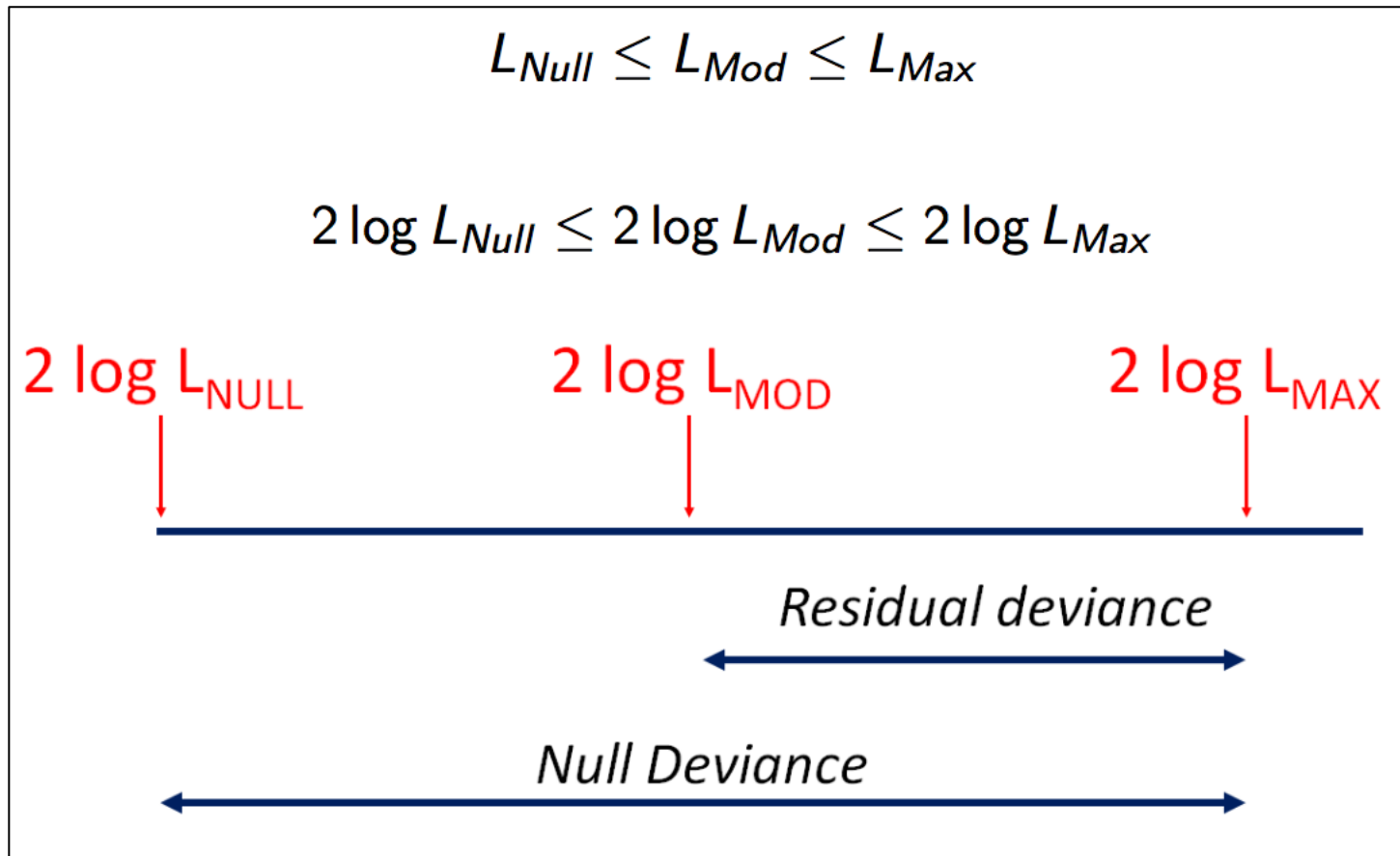
- Compares the model of interest to the “saturated” model
 - $2 * [\log L(\text{saturated model}) - \log L(\text{model of interest})] \sim \chi^2$
 - Number of degrees of freedom depends on the difference between the number of parameters
- Intuitively, the better the logistic model of interest
 - The closer $\log L(\text{saturated model})$ is to $\log L(\text{model of interest})$
 - Then, the smaller the deviance should be
- How small is small?
 - If m (# distinct covariate patterns) is small and the n_i 's are large
 - Then when the logistic model is true
 - The deviance has approximately a chi-squared distribution with $m - k - 1$ degrees of freedom where
 - m : number of covariate patterns
 - k : number of covariates in the model of interest
 - Thus, if the deviance is less than the upper 95% percentage point of the appropriate chi-square distribution, the logistic model fits well
- In this sense, the deviance is the analogue of R^2

Deviance

- ONLY applies to grouped data
 - When m (# distinct covariate patterns) is small
 - And the n_i (the number of individuals having that pattern) are large
- Other names for deviance: model deviance, residual deviance
- At the other extreme, the most restrictive model is one where all the probabilities π_i are the same
 - i.e. they don't depend on the covariates
 - Intercept only model
 - The deviance for this model is called the null deviance
- Intuitively
 - If none of the covariates is related to the binary response
 - Then, the model deviance won't be much greater than the null deviance

Deviance: Graphical Interpretation

- $\text{Log } L_{\text{null}}$ refers to the log-likelihood for the intercept only model
- $\text{Log } L_{\text{mod}}$ refers to the log-likelihood for the model of interest
- $\text{Log } L_{\text{max}}$ refers to the log-likelihood for the saturated model



Example: Budworm Data

- The data come from an experiment on the toxicity to the tobacco budworm *Heliothis virescens* of doses of the pyrethroid trans-cypermethrin to which the moths were beginning to show resistance
- Batches of 20 moths of each sex were exposed for three days to the pyrethroid and the number in each batch that were dead or knocked down was recorded
 - Collette, D. (1991) Modelling Binary Data. Chapman and Hall, London. p 75
- Batches of 20 moths subjected to increasing doses of a poison
 - “event” or “success” = death
- Data is grouped: for each of 6 doses (1.0, 2.0, 4.0, 8.0, 16.0, 32.0 mg) and sex
 - m=12 covariate patterns, 20 moths in each covariate pattern
- Dataset is in the R330 package in R
- R commands denoted by >
 - > install.packages("R330") #only need to install the package once
 - > library(R330) #load the library
 - > data(budworm.df) #dataset for budworm

Example: Budworm Data

- R commands denoted by >

```
> budworm.df
```

	sex	dose	s	n
1	0	1	1	20
2	0	2	4	20
3	0	4	9	20
4	0	8	13	20
5	0	16	18	20
6	0	32	20	20
7	1	1	0	20
8	1	2	2	20
9	1	4	6	20
10	1	8	10	20
11	1	16	12	20
12	1	32	16	20

sex= the sex of the budworm

dose= amount of cypermethrin exposed to

s= number of budworms affected

n= total number of budworms

Example: Budworm Data

```
# number of events over total number in each group
```

```
> max.mod.probs<-budworm.df$s/budworm.df$n
```

```
# model of interest
```

```
> budworm.glm<-glm( cbind(s, n-s) ~ sex + dose, family=binomial, data = budworm.df)
```

```
# fitted probabilities for model of interest
```

```
> logist.mod.probs<-predict(budworm.glm, type="response")
```

```
# intercept only model
```

```
> null.mod.probs<-sum(budworm.df$s)/sum(budworm.df$n)
```

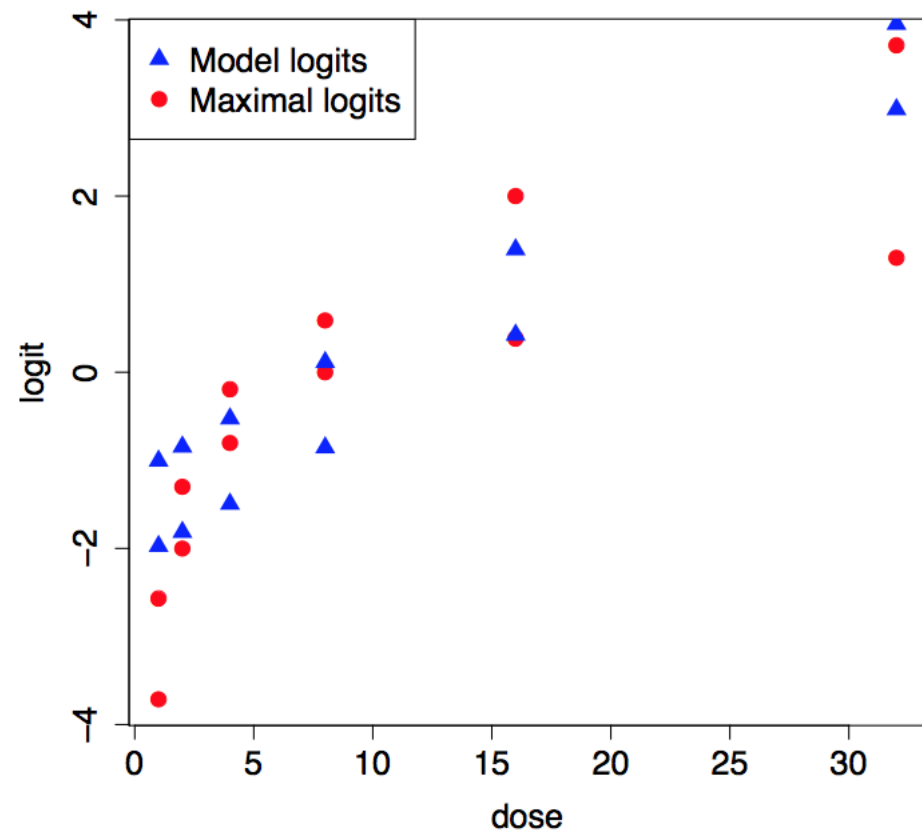
```
> cbind(max.mod.probs,logist.mod.probs,null.mod.probs)
```

	max.mod.probs	logist.mod.probs	null.mod.probs
1	0.05	0.2677414	0.4625
2	0.20	0.3002398	0.4625
3	0.45	0.3713931	0.4625
4	0.65	0.5283639	0.4625
5	0.90	0.8011063	0.4625
6	1.00	0.9811556	0.4625
7	0.00	0.1218892	0.4625
8	0.10	0.1400705	0.4625
9	0.30	0.1832034	0.4625
10	0.50	0.2983912	0.4625
11	0.60	0.6046013	0.4625
12	0.80	0.9518445	0.4625

Example: Log-Odds for Model of Interest and Saturated Model

```
> max.logit = log((budworm.df$s+0.5)/(budworm.df$n -budworm.df$s+0.5))
> model.logit = predict(budworm.glm)
> cbind(max.logit,model.logit)
```

	max.logit	model.logit
1	-2.5649494	-1.0061121
2	-1.2992830	-0.8461564
3	-0.1910552	-0.5262451
4	0.5877867	0.1135776
5	2.0014800	1.3932230
6	3.7135721	3.9525137
7	-3.7135721	-1.9746604
8	-2.0014800	-1.8147047
9	-0.8023465	-1.4947934
10	0.0000000	-0.8549707
11	0.3856625	0.4246747
12	1.2992830	2.9839654



- Poor fit: maximal logits are not linear

Calculating the likelihoods

Likelihood is

$$L(\pi_1, \dots, \pi_M) = \prod_{i=1}^M \binom{n_i}{r_i} \pi_i^{r_i} (1 - \pi_i)^{n_i - r_i}$$

$$\begin{aligned} L_{Max} &= 2.8947 \times 10^{-7}, \quad 2 \log L_{Max} = -30.1104 \\ L_{Mod} &= 2.4459 \times 10^{-13}, \quad 2 \log L_{Mod} = -58.0783 \\ L_{Null} &= 2.2142 \times 10^{-34}, \quad 2 \log L_{Null} = -154.9860 \end{aligned}$$

- Where $r_i = s_i$ is the number of events of success
- Calculating the log-likelihoods via R
 - Fit the intercept only model
 - Fit the saturated model
 - Fit the model of interest
 - Deviance by hand or via R

Null Model: Intercept Only

```
budwormN.glm<-glm( cbind(s, n-s) ~ 1, family=binomial, data = budworm.df)
```

```
summary( budwormN.glm)
```

	Estimate	Std. Error	z value	Pr(> z)
(Intercept)	-0.1503	0.1295	-1.161	0.246

Null deviance: 124.88 on 11 degrees of freedom

Residual deviance: 124.88 on 11 degrees of freedom

AIC: 156.99

- $AIC = -2 \log L + 2p$
- Then
 - $2 \log L = (AIC - 2p) / (-1) = (156.99 - 2) / (-1) = -154.99$
 - Then $\log L = -77.495$
 - Same as previous slide for $2 \log L$

Example: Saturated Model

```
budwormF.glm<-glm( cbind(s, n-s) ~ sex + as.factor(dose)+sex* as.factor(dose), family=binomial,
data = budworm.df)
```

```
summary(budwormF.glm)
```

	Estimate	Std. Error	z value	Pr(> z)
(Intercept)	-2.944	1.026	-2.870	0.00411 **
sex	-22.807	52998.328	0.000	0.99966
as.factor(dose)2	1.558	1.168	1.334	0.18234
as.factor(dose)4	2.744	1.120	2.450	0.01430 *
as.factor(dose)8	3.563	1.128	3.159	0.00158 **
as.factor(dose)16	5.142	1.268	4.054	5.02e-05 ***
as.factor(dose)32	28.696	52998.283	0.001	0.99957
sex:as.factor(dose)2	21.996	52998.328	0.000	0.99967
sex:as.factor(dose)4	22.161	52998.328	0.000	0.99967
sex:as.factor(dose)8	22.188	52998.328	0.000	0.99967
sex:as.factor(dose)16	21.016	52998.328	0.000	0.99968
sex:as.factor(dose)32	-1.558	74950.923	0.000	0.99998

Null deviance: 1.2488e+02 on 11 degrees of freedom

Residual deviance: 5.2391e-10 on 0 degrees of freedom

AIC: 54.11 then $2 \cdot \log L = (AIC - 2 \cdot p) / (-1) = (54.11 - 2 \cdot 12) / (-1) = -30.11$ the $\log L = -15.055$

Null Deviance for Saturated Model

- Compare null model to the saturated model using a LRT

```
library(lmtest)
```

```
lrtest(budwormN.glm,budwormF.glm) #LRT
```

Likelihood ratio test

Model 1: $\text{cbind}(s, n - s) \sim 1$

Model 2: $\text{cbind}(s, n - s) \sim \text{sex} + \text{as.factor}(\text{dose}) + \text{sex} * \text{as.factor}(\text{dose})$

	#Df	LogLik	Df	Chisq	Pr(>Chisq)
1	1	-77.493			
2	12	-15.055	11	124.88	< 2.2e-16 ***

```
> 1-pchisq(124.88,11)
0
```

Note: $-2(-77.493--15.055)= 124.88$ and $DF=12-1=11$

Model of Interest

```
> summary(budworm.glm)
```

Residual deviance = $-30.1104 - (-58.0783) = 27.9679$

Null deviance = $-30.1104 - (-154.9860) = 124.8756$

Coefficients:

	Estimate	Std. Error	z value	Pr(> z)	
(Intercept)	-1.16607	0.26155	-4.458	8.26e-06	***
sex	-0.96855	0.32954	-2.939	0.00329	**
dose	0.15996	0.02341	6.832	8.39e-12	***

Null deviance: 124.876 on 11 degrees of freedom

Residual deviance: 27.968 on 9 degrees of freedom

AIC: 64.078

Deviance: Goodness of Fit

- Grouped data
 - n_j 's reasonably large (20), m small (12)
- Can interpret residual deviance as a measure of fit


```
> 1-pchisq(27.968,9)
[1] 0.0009656815
```

- Not a good fit!! As suspected from the plot

- $\log(\text{dose})$ works better

```
> logdose.glm<-glm(cbind(s, n - s) ~ sex + log(dose), family = binomial,data = budworm.df)
> summary(logdose.glm)
```

	Estimate	Std. Error	z value	Pr(> z)
(Intercept)	-2.3724	0.3855	-6.154	7.56e-10 ***
sex	-1.1007	0.3558	-3.093	0.00198 **
log(dose)	1.5353	0.1891	8.119	4.70e-16 ***

Null deviance: 124.8756 on 11 degrees of freedom

Residual deviance: 6.7571 on 9 degrees of freedom

AIC: 42.867

```
> 1-pchisq(6.7571,9)
[1] 0.662392
```

Deviance: ONLY for group level data

```
sex<-c(rep(0,120),rep(1,120))
dose<-c(rep(1,20),rep(2,20),rep(4,20),rep(8,20),rep(16,20),rep(32,20))
dose<-c(dose,dose)
```

```
y<-c(rep(1,1),rep(0,20-1),rep(1,4),rep(0,20-4),rep(1,9),rep(0,20-9),rep(1,13),rep(0,20-13),rep(1,18),rep(0,20-18),rep(1,20),rep(0,20),rep(1,2),rep(0,20-2),rep(1,6),rep(0,20-6),rep(1,10),rep(0,20-10),rep(1,12),rep(0,20-12),rep(1,16),rep(0,20-16))
```

```
summary(glm(y~sex+dose,family=binomial()))
```

	Estimate	Std. Error	z value	Pr(> z)
(Intercept)	-1.16607	0.26155	-4.458	8.26e-06 ***
sex	-0.96855	0.32954	-2.939	0.00329 **
dose	0.15996	0.02341	6.832	8.39e-12 ***

Null deviance: 331.36 on 239 degrees of freedom

Residual deviance: 234.45 on 237 degrees of freedom

AIC: 240.45

LAYOUT #1**PROC LOGISTIC: SAS Example**

```

DATA smoke;
  INPUT smoke passive cancer n;
  DATALINES;
0 1 1 120
0 1 0 80
0 0 1 111
0 0 0 155
1 1 1 161
1 1 0 130
1 0 1 117
1 0 0 124
;
PROC LOGISTIC DESCENDING;
  MODEL cancer = passive;
  FREQ n;
RUN;

```

LAYOUT #2

```

DATA smoke2;
  INPUT smoke passive cancer_n tot_n;
  DATALINES;
0 1 120 200
0 0 111 266
1 1 161 291
1 0 117 241
;
PROC LOGISTIC;
  MODEL cancer_n/tot_n = passive;
RUN;

```

Smokers (1)

<i>Passive Smoke</i>	<i>Cancer</i>		
	Case (1)	Control (0)	
Yes (1)	161	130	291
No (0)	117	124	241
	278	254	532

Non-Smokers (0)

<i>Passive Smoke</i>	<i>Cancer</i>		
	Case (1)	Control (0)	
Yes (1)	120	80	200
No (0)	111	155	266
	231	235	466

LAYOUT #1 The LOGISTIC Procedure

Model Information

Data Set	WORK.SMOKE
Response Variable	cancer
Number of Response Levels	2
Frequency Variable	n
Model	binary logit
Optimization Technique	Fisher's scoring
Number of Observations Read	8
Number of Observations Used	8
Sum of Frequencies Read	998
Sum of Frequencies Used	998

```
PROC LOGISTIC DESCENDING;
  MODEL cancer = passive;
  FREQ n;
RUN;
```

<i>Passive Smoke</i>	<i>Cancer</i>		
	Case	Control	
Yes	281	210	491
No	228	279	507
	509	489	998

Response Profile

Ordered Value	cancer	Total Frequency
1	1	509
2	0	489

Probability modeled is cancer=1.

$$\text{AIC: } -2\text{LL} + 2p = 1383.121 + 2 \times 1 = 1385.121$$

$$\text{SC: } -2\text{LL} + p \cdot \log(n) = 1383.121 + 1 \cdot \text{LN}(998) = 1390.027$$

$$-2\text{LL} = -2[509 \cdot \log(509/998) + 489 \cdot \log(489/998)] = 1383.121$$

SAS counts the intercept for p!!!!

Model Fit Statistics

Criterion	Intercept Only	Intercept and Covariates
AIC	1385.121	1372.080
SC	1390.027	1381.892
-2 Log L	1383.121	1368.080

$$\text{AIC: } -2\text{LL} + 2p = 1368.080 + 2 \times 2$$

$$\text{SC: } -2\text{LL} + p \cdot \log(n) = 1368.080 + 2 \cdot \text{LN}(998)$$

$$-2\text{LL} = -2[228 \cdot \log(228/507) + 279 \cdot \log(279/507) + 281 \cdot \log(281/491) + 210 \cdot \log(210/491)] = 1368.080$$

Testing Global Null Hypothesis: BETA=0

Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	15.0409	1	0.0001
Score	15.0022	1	0.0001
Wald	14.9262	1	0.0001

Analysis of Maximum Likelihood Estimates

Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	-0.2019	0.0893	5.1128	0.0238
passive	1	0.4931	0.1276	14.9262	0.0001

$$\log\left(\frac{p}{1-p}\right) = \beta_0 + \beta_{\text{passive}} \text{passive}$$

$$\log\left(\frac{\hat{p}_{\text{cancer}|\text{passive}=0}}{1 - \hat{p}_{\text{cancer}|\text{passive}=0}}\right) = \hat{\beta}_0 \Rightarrow \hat{\beta}_0 = \log\left(\frac{\frac{228}{507}}{1 - \frac{228}{507}}\right) = -0.2019$$

$$\log\left(\frac{\hat{p}_{\text{cancer}|\text{passive}=1}}{1 - \hat{p}_{\text{cancer}|\text{passive}=1}}\right) = \hat{\beta}_0 + \hat{\beta}_{\text{passive}} \Rightarrow \log\left(\frac{\frac{281}{491}}{1 - \frac{281}{491}}\right) = \hat{\beta}_0 + \hat{\beta}_{\text{passive}} \Rightarrow \hat{\beta}_{\text{passive}} = \log\left(\frac{\frac{281}{491}}{1 - \frac{281}{491}}\right) - \log\left(\frac{\frac{228}{507}}{1 - \frac{228}{507}}\right) = 0.4931$$

LAYOUT #2

The LOGISTIC Procedure

Model Information

Data Set	WORK.SMOKE2
Response Variable (Events)	cancer_n
Response Variable (Trials)	tot_n
Model	binary logit
Optimization Technique	Fisher's scoring

Number of Observations Read	4
Number of Observations Used	4
Sum of Frequencies Read	998
Sum of Frequencies Used	998

Response Profile

Ordered	Binary	Total
Value	Outcome	Frequency
1	Event	509
2	Nonevent	489

Model Fit Statistics

	Intercept	Intercept and
Criterion	Only	Covariates
AIC	1385.121	1372.080
SC	1390.027	1381.892
-2 Log L	1383.121	1368.080

```
PROC LOGISTIC;
  MODEL cancer_n/tot_n = passive;
RUN;
```

Results/Output are identical to Layout #1 other than the number of observations read.

Testing Global Null Hypothesis: $BETA=0$

Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	15.0409	1	0.0001
Score	15.0022	1	0.0001
Wald	14.9262	1	0.0001

Same as layout #1

Analysis of Maximum Likelihood Estimates

Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	-0.2019	0.0893	5.1128	0.0238
passive	1	0.4931	0.1276	14.9262	0.0001

PROC GENMOD

- You can also perform logistic regression using the PROC GENMOD.
- Note that the calculation of the Goodness of Fit statistics differ for layout #1 (A or B) and layout #2.

LAYOUT #1A

```
DATA smoke;
  INPUT smoke passive cancer;
  DATALINES;
  0 1 1
  0 1 1
  ...
  0 1 1
  0 1 0
  ...
  0 1 0
  0 0 1
  ...
  ...
  1 0 1
  1 0 0
  ...
  1 0 0
  ;
```

120 rows

80 rows

124 rows

```
PROC GENMOD DESCENDING DATA=smoke;
  MODEL cancer = passive /DIST=BINOMIAL TYPE3;
  RUN;
```

Smokers:

<i>Passive Smoke</i>	<i>Cancer</i>		
	Case	Control	
Yes	161	130	291
No	117	124	241
	278	254	532

Non-Smokers

<i>Passive Smoke</i>	<i>Cancer</i>		
	Case	Control	
Yes	120	80	200
No	111	155	266
	231	235	466

Request Likelihood Ratio Tests (Optional)

Request Binomial Distribution
(Required to fit Logistic Model)

LAYOUT #1B

```

DATA smoke;
  INPUT smoke passive cancer n;
  DATALINES;
  0 1 1 120
  0 1 0 80
  0 0 1 111
  0 0 0 155
  1 1 1 161
  1 1 0 130
  1 0 1 117
  1 0 0 124
;

PROC GENMOD DESCENDING DATA=smoke;
  MODEL cancer = passive /DIST=BINOMIAL TYPE3;
  FREQ n;
RUN;

```

LAYOUT #2

You do not use
DESCENDING
option with
layout #2.

```

DATA smoke2;
  INPUT smoke passive cancer_n tot_n;
  DATALINES;
  0 1 120 200
  0 0 111 266
  1 1 161 291
  1 0 117 241
;

PROC GENMOD DATA=smoke2;
  MODEL cancer_n/tot_n = passive /DIST=BINOMIAL TYPE3;
RUN;

```

Model Information

Data Set	WORK.SMOKE
Distribution	Binomial
Link Function	Logit
Dependent Variable	cancer
Frequency Weight Variable	n

```
PROC GENMOD DESCENDING DATA=smoke2;
  MODEL cancer=passive DIST= B TYPE3;
  FREQ n;
  RUN;
```

Number of Observations Read	8
Number of Observations Used	8
Sum of Frequencies Read	998
Sum of Frequencies Used	998
Number of Events	4
Number of Trials	8

Response Profile

Ordered Value	cancer	Total Frequency
1	1	509
2	0	489

PROC GENMOD is modeling the probability that cancer='1'.

Deviance compared to model with a parameter fit for each individual.

Not distributed as a chi-square.

But is the -2LL

Criteria For Assessing Goodness Of Fit

Criterion	DF	Value	Value/DF
Deviance	996	1368.0800	1.3736
Scaled Deviance	996	1368.0800	1.3736
Pearson Chi-Square	996	998.0000	1.0020
Scaled Pearson X2	996	998.0000	1.0020
Log Likelihood		-684.0400	

GENMOD provides the Log Likelihood (not -2LL)

Analysis Of Parameter Estimates

Parameter	DF	Estimate	Standard Error	Wald 95% Confidence Limits	Chi-Square	Pr > ChiSq
Intercept	1	-0.2019	0.0893	-0.3768 -0.0269	5.11	0.0238
passive	1	0.4931	0.1276	0.2430 0.7433	14.93	0.0001
Scale	0	1.0000	0.0000	1.0000 1.0000		

NOTE: The scale parameter was held fixed.

LR Statistics For Type 3 Analysis

Source	DF	Chi-Square	Pr > ChiSq
passive	1	15.04	0.0001

Same as layout #1, #2

GENMOD provides the
Log Likelihood (not -2LL)

Layout #2

```
PROC GENMOD DESCENDING DATA=smoke2;
  MODEL cancer_n/tot_n =passive /DIST=BINOMIAL TYPE3;
RUN;
```

The GENMOD Procedure

Model Information

Data Set	WORK.SMOKE2
Distribution	Binomial
Link Function	Logit
Response Variable (Events)	cancer_n
Response Variable (Trials)	tot_n
Number of Observations Read	4
Number of Observations Used	4
Number of Events	509
Number of Trials	998

Criteria For Assessing Goodness Of Fit

Criterion	DF	Value	Value/DF
Deviance	2	3.4362	1.7181
Scaled Deviance	2	3.4362	1.7181
Pearson Chi-Square	2	3.4331	1.7165
Scaled Pearson X2	2	3.4331	1.7165
Log Likelihood		-684.0400	

Differs from Output for Layout #1!

Deviance compared to saturated model with a parameter fit for each group defined by passive×smoke.

IS distributed as a chi-square.

The deviance is only appropriate for group-level data!!!!

BE CAREFUL of the layout of the data when using deviance.

Analysis Of Parameter Estimates

Parameter	DF	Estimate	Standard Error	Wald 95% Confidence Limits	Chi-Square	Pr > ChiSq
Intercept	1	-0.2019	0.0893	-0.3768 -0.0269	5.11	0.0238
passive	1	0.4931	0.1276	0.2430 0.7433	14.93	0.0001
Scale	0	1.0000	0.0000	1.0000 1.0000		

NOTE: The scale parameter was held fixed.

Same as layout #1, #2, #1B

LR Statistics For Type 3 Analysis

Source	DF	Chi-Square	Pr > ChiSq
passive	1	15.04	0.0001