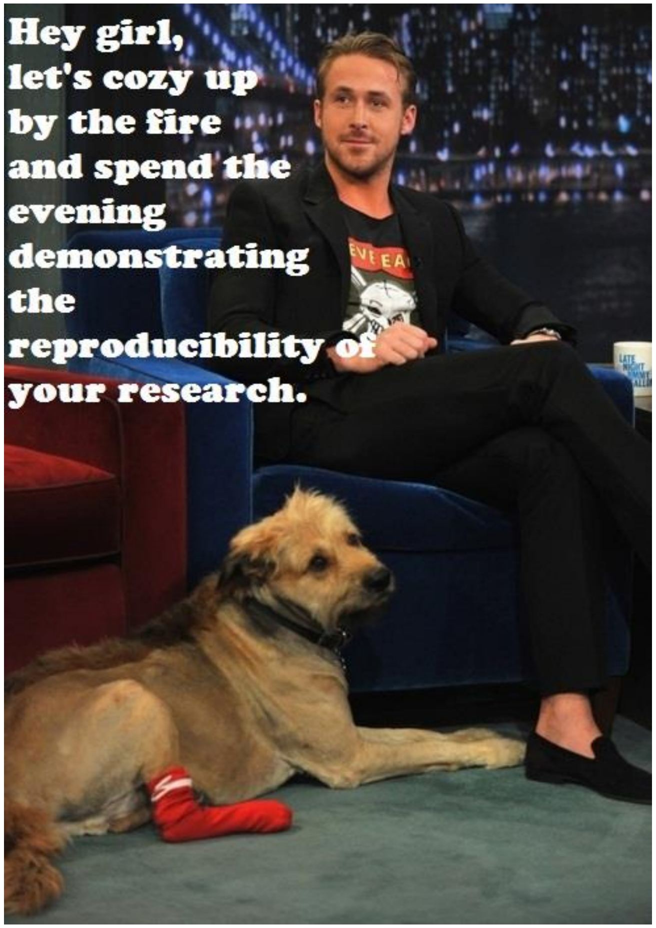


BIOS 6612

Lecture 7

Goodness of Fit, Predictive Power, And Diagnostics for Logistic Regression

Review (Lecture 6) / Current (Lecture 7)/ Preview (Lecture 8)

- Lecture 6: Covariate Adjustment in Logistic Regression
    - Confounding
        - Operational vs Classical Criteria (NOT the same)

- Lecture 7
    - 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: Multiple Logistic Regression
    - Diagnostics
    - Examples
        - Code in R and SAS

Goodness-of-Fit Statistics

- How well does the model "fit" the data?
  - **Goodness-of-fit statistics** implicitly involve a comparison between a model of interest and a "maximal" model that is more complex.
  - The maximal model will always fit better, but is it "significantly" better?
  - The **global chi-square test** compares the model of interest and a "minimal" model that contains only an intercept.

- **Deviance Chi-Square** (available in PROC GENMOD)
  - Compares the model of interest to the "saturated" model.
    - 2* [log L(saturated model) - log L(model of interest)] ~ χ^2
  - Number of degrees of freedom depends on the difference between the number of parameters in the models to compare
  - The saturated model has one parameter for every predicted probability and therefore produces a perfect fit to the data.
  - SAS's choice of the "saturated" model depends on how the data are arranged.
  - Analogous to multiple linear regression lack-of-fit F-test

Goodness of Fit: Hosmer-Lemeshow statistic

The ***Hosmer-Lemeshow statistic*** is based on comparing observed versus expected (from the model) frequencies in categories defined by the fitted values for the outcome.

It is used to examine goodness of fit for individual level data
-recall the Deviance Chi-Square is used for grouped data

- The observations are first divided into approximately ten groups (deciles).

- The number of groups, *g*, may be smaller than 10 if there are fewer than 10 distinct patterns of explanatory variables (but *g* must be ≥ 3).

- The Hosmer-Lemeshow goodness-of-fit statistic is obtained by calculating the Pearson chi-square statistic from the 2×*g* table of observed and expected frequencies, where *g* is the number of groups:

$$\chi^2 = \sum_g \frac{(O_g - E_g)^2}{E_g(1 - E_g / n_g)}$$

Goodness of Fit: Hosmer-Lemeshow statistic

- The Hosmer-Lemeshow statistic is then compared to a chi-square distribution with (g-n) degrees of freedom, where the value of n can be specified in the LACKFIT option in the MODEL statement.
  - The default is n=2.
- Large values of the test statistic (and *small p*-values) indicate a lack of fit of the model.
  - Ideally the p-value for this test will be >>>0.05.
- ***Note: The HL statistic is an ad-hoc method and does not have very good properties (including low power).***
  - ***You'll see it reported widely but it is not necessarily a good statistic to report.***

Predictive Power

- Sometimes the main interest of a logistic regression analysis is prediction.
  - That is, how well can we predict the outcome (dependent variable) based on the values of the independent variables.

- This is a very different criterion from goodness-of-fit measures.
  - A model can have good prediction but poor fit or good fit but poor prediction.

- Measures of Predictive Power

(1) Measures based on the log-likelihood:

- Generalized R^2
- Max-Rescaled R^2 (Nagelkerke R^2)

RSQ option in PROC LOGISTIC

(2) Measures based on predicted probabilities and observed responses:

- Somer's D
- Gamma
- Tau-a
- c-index

SAS Defaults in PROC LOGISTIC

- These should <u>not</u> be interpreted as the proportion of variance explained by the independent variables.
  - NOT the same thing as R^2 in linear regression

Measures of Predictive Power
Based on the log-likelihood

- **<u>Generalized R^2</u>** $$R^2 = 1 - \exp\left[-\frac{L^2}{n}\right]$$

  - where L^2 is the likelihood ratio chi-square for testing the null hypothesis that all coefficients are 0 and n is the sample size.

- **<u>Max-Rescaled R^2 (Nagelkerke R^2)</u>**

  - Divides the generalized R^2 by its upper bound.

Measures of Predictive Power
Based on predicted probabilities and observed responses

- The following measures use information based on the number of pairs of observations in which one observation had the response and the other did not.

- For each pair, if the observation with the response has the higher predicted probability, it is "**concordant**."

- If the observation with the response has the lower predicted probability, it is "**discordant**."

- If both observations have the same predicted probability, it is a "**tie**".

  - *Somer's D*: $\frac{C-D}{C+D+T}$

  - *Gamma*: $\frac{C-D}{C+D}$

  - *Tau-a*: $\frac{C-D}{N}$

  - *c-index*: $0.5\,(1+Somer's\ D)$

C = Number of Concordant Pairs

D = Number of Discordant Pairs

T = Number of Ties

N = Total Number of Pairs
(including those with identical responses – pairs where both have the response or both don't have the response)

NOTE: The c-index is equivalent to the area under the ROC curve; ROC curves are measures of predictive ability based on sensitivity and specificity.

Example: Predictors of attrition from weight-loss program

gender	Gender (1=female; 0=male)
age	Age (years)
bmi	BMI (m/kg^2)
steps	Baseline Physical Activity (Steps Per Day)
complete	Program Completer (0=No; 1=Yes)

Question: Is baseline BMI or baseline physical activity (steps/day) a better predictor of attrition?

Obs	gender	age	bmi	steps	complete
1	1	30.0395	25.1523	6242.87	0
2	1	39.5773	27.9006	5039.61	1
3	1	34.2372	28.1800	6224.94	1
4	1	39.4810	31.1162	6190.36	1
5	1	29.5557	32.4988	4487.81	0
6	1	40.4104	22.7115	10047.78	1
7	1	27.9558	28.3043	6872.79	1
8	1	35.3751	27.1835	8016.53	1
9	1	31.7030	44.1499	560.53	0
10	1	26.1246	25.4449	7683.27	1
11	0	35.7991	29.7877	7652.74	1
12	0	34.3726	24.5994	6637.54	1
13	0	26.8996	32.7637	4367.06	0
14	0	37.3319	34.6787	4199.86	0
15	0	30.0353	25.1707	9349.20	1
16	0	32.9544	37.1013	939.48	0
17	0	34.9997	26.3737	6820.67	1
18	0	19.1743	20.7199	11765.55	1
19	0	25.8094	31.1911	4415.38	0
20	0	29.8202	25.3290	5994.75	1

bmi	steps
Mean= 29.0	Mean= 6175.4
SD= 5.4	SD= 2701.6

13 participants completed the program

7 dropped out

Small n: LRT is more appropriate than Wald

Steps as predictor (Model 1):

```
PROC LOGISTIC;
  MODEL complete = steps /RSQ LACKFIT CLODDS=WALD;
  UNITS steps=100;
RUN;
```

Requests R-square measures

Requests Hosmer-Lemeshow

```
The LOGISTIC Procedure

          Response Profile

 Ordered                        Total
   Value      complete      Frequency
       1             0              7
       2             1             13
Probability modeled is complete=0.
```

Modeling Attrition (drop-out)

```
         Model Fit Statistics
                                 Intercept
              Intercept                and
Criterion          Only         Covariates
AIC              27.898             12.554
SC               28.894             14.546
-2 Log L         25.898              8.554
```

LL = 7*LN(7/20)+13*log(13/20) = -12.9489

-2LL = 25.8979

```
  R-Square     0.5799     Max-rescaled R-Square     0.7986

        Testing Global Null Hypothesis: BETA=0

Test                  Chi-Square       DF     Pr > ChiSq
Likelihood Ratio         17.3434        1         <.0001
Score                    10.2874        1         0.0013
Wald                      4.8496        1         0.0277
```

We will compare to model with BMI as predictor (Model 2).

Steps as predictor (Model 1)

Analysis of Maximum Likelihood Estimates

Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	11.7273	5.5413	4.4789	0.0343
steps	1	-0.00219	0.000993	4.8496	0.0277

Odds Ratio Estimates

Effect	Point Estimate	95% Wald Confidence Limits	
steps	0.998	0.996	1.000

Note that the point estimate is significant, but close to 1.0

The odds of attrition decrease 0.2% (95% CI: 0.02% to 0.4%) for each 1 step/day higher baseline steps.

What are the relative odds for 100 steps/day?

OR = EXP(-0.00219 × 100) = 0.803

The odds of attrition decrease 19.6% for each 100 step/day higher baseline steps.

Wald Confidence Interval for Adjusted Odds Ratios

Effect	Unit	Estimate	95% Confidence Limits	
steps	100.0	0.804	0.661	0.976

Steps as predictor (Model 1)

Association of Predicted Probabilities and Observed Responses

Percent Concordant	**95.6**	**Somers' D**	**0.912**
Percent Discordant	**4.4**	**Gamma**	**0.912**
Percent Tied	**0.0**	**Tau-a**	**0.437**
Pairs	**91**	**c**	**0.956**

Partition for the Hosmer and Lemeshow Test

		complete = 0		complete = 1	
Group	Total	Observed	Expected	Observed	Expected
1	2	0	0.00	2	2.00
2	2	0	0.00	2	2.00
3	2	0	0.01	2	1.99
4	2	0	0.07	2	1.93
5	2	1	0.18	1	1.82
6	2	0	0.27	2	1.73
7	2	0	0.87	2	1.13
8	2	2	1.76	0	0.24
9	2	2	1.82	0	0.18
10	2	2	2.00	0	0.00

Hosmer and Lemeshow Goodness-of-Fit Test

Chi-Square	DF	Pr > ChiSq
6.3757	**8**	**0.6052**

P >>> 0.05; No significant lack of fit. The current model is adequate.

Predicted Values from Model 1 (baseline steps model):

Predicted probability for complete = 0

Obs	gender	age	bmi	steps	complete	dropped	phat	ll
1	**1**	**30.0395**	**25.1523**	**6242.87**	**0**	**1**	**0.12694**	**-2.06404**
2	1	39.5773	27.9006	5039.61	1	0	0.66901	-1.10566
3	1	34.2372	28.1800	6224.94	1	0	0.13135	-0.14082
4	1	39.4810	31.1162	6190.36	1	0	0.14022	-0.15108
5	1	29.5557	32.4988	4487.81	0	1	0.87110	-0.13800
6	1	40.4104	22.7115	10047.78	1	0	0.00004	-0.00004
7	1	27.9558	28.3043	6872.79	1	0	0.03536	-0.03600
8	1	35.3751	27.1835	8016.53	1	0	0.00299	-0.00300
9	1	31.7030	44.1499	560.53	0	1	0.99997	-0.00003
10	1	26.1246	25.4449	7683.27	1	0	0.00619	-0.00621
11	0	35.7991	29.7877	7652.74	1	0	0.00661	-0.00663
12	0	34.3726	24.5994	6637.54	1	0	0.05778	-0.05952
13	0	26.8996	32.7637	4367.06	0	1	0.89797	-0.10762
14	0	37.3319	34.6787	4199.86	0	1	0.92694	-0.07587
15	0	30.0353	25.1707	9349.20	1	0	0.00016	-0.00016
16	0	32.9544	37.1013	939.48	0	1	0.99994	-0.00006
17	0	34.9997	26.3737	6820.67	1	0	0.03946	-0.04026
18	0	19.1743	20.7199	11765.55	1	0	0.00000	-0.00000
19	0	25.8094	31.1911	4415.38	0	1	0.88787	-0.11893
20	0	29.8202	25.3290	5994.75	1	0	0.20012	-0.22329

Obs #1: phat = 1/(1+EXP[-(11.7273-.0.0021874*6242.87)]) = 0.12694
LL = LN[0.12694×(1-0.66901)×…×(0.88787)×(1-0.20012)] = -4.42772238
-2LL = -2*(-4.2772238) = 8.55445

$$\frac{\exp\left(\beta_0+\beta_1 X_{1i}+\ldots\right)}{1+\exp\left(\beta_0+\beta_1 X_{1i}+\ldots\right)}=\frac{1}{1+\exp\left(-\left[\beta_0+\beta_1 X_{1i}+\ldots\right]\right)}$$

Predicted Values from Model 1 (baseline steps model): sorted by phat

Obs	id	gender	age	bmi	steps	complete	dropped	phat
1	18	0	19.1743	20.7199	11765.55	1	0	0.00000
2	6	1	40.4104	22.7115	10047.78	1	0	0.00004
3	15	0	30.0353	25.1707	9349.20	1	0	0.00016
4	8	1	35.3751	27.1835	8016.53	1	0	0.00299
5	10	1	26.1246	25.4449	7683.27	1	0	0.00619
6	11	0	35.7991	29.7877	7652.74	1	0	0.00661
7	7	1	27.9558	28.3043	6872.79	1	0	0.03536
8	17	0	34.9997	26.3737	6820.67	1	0	0.03946
9	12	0	34.3726	24.5994	6637.54	1	0	0.05778
10	1	1	30.0395	25.1523	6242.87	0	1	0.12694
11	3	1	34.2372	28.1800	6224.94	1	0	0.13135
12	4	1	39.4810	31.1162	6190.36	1	0	0.14022
13	20	0	29.8202	25.3290	5994.75	1	0	0.20012
14	2	1	39.5773	27.9006	5039.61	1	0	0.66901
15	5	1	29.5557	32.4988	4487.81	0	1	0.87110
16	19	0	25.8094	31.1911	4415.38	0	1	0.88787
17	13	0	26.8996	32.7637	4367.06	0	1	0.89797
18	14	0	37.3319	34.6787	4199.86	0	1	0.92694
19	16	0	32.9544	37.1013	939.48	0	1	0.99994
20	9	1	31.7030	44.1499	560.53	0	1	0.99997

Observed	Expected (sum phat for 2 Obs)
0	.00004
0	.00315
0	.01280
0	.07482
1	.18472
0	.27157
0	.86913
2	1.75897
2	1.82491
2	1.99991

Hosmer-Lemeshow

$$\chi^2 = \sum_g \frac{(O_g - E_g)^2}{E_g(1 - E_g / n_g)}$$

Is BMI or baseline steps a better predictor of attrition from the weight-loss program?

BMI as predictor (Model 2):

```
PROC LOGISTIC;
  MODEL complete = bmi /RSQ LACKFIT;
RUN;
```

```
         The LOGISTIC Procedure

          Response Profile

 Ordered                        Total
   Value       complete     Frequency
       1              0             7
       2              1            13
Probability modeled is complete=0.
```

Modeling Attrition (drop-out)

```
         Model Fit Statistics
                                Intercept
              Intercept               and
Criterion          Only        Covariates
AIC              27.898            17.521
SC               28.894            19.512
-2 Log L         25.898            13.521

R-Square    0.4614    Max-rescaled R-Square    0.6355

        Testing Global Null Hypothesis: BETA=0

Test                    Chi-Square       DF     Pr > ChiSq
Likelihood Ratio           12.3770        1         0.0004
Score                       9.3057        1         0.0023
Wald                        4.9650        1         0.0259
```

Compared to model 1 with steps:
Larger LR p-value for model 2
Smaller Wald p-value for model 2

BMI as predictor (Model 2)

Analysis of Maximum Likelihood Estimates

Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	-17.1372	7.4997	5.2214	0.0223
bmi	1	0.5601	0.2514	4.9650	0.0259

Odds Ratio Estimates

Effect	Point Estimate	95% Wald Confidence Limits	
bmi	1.751	1.070	2.866

Wald Confidence Interval for Adjusted Odds Ratios

Effect	Unit	Estimate	95% Confidence Limits	
bmi	1.0000	1.751	1.070	2.866

Association of Predicted Probabilities and Observed Responses

Percent Concordant	**89.0**	**Somers' D**	**0.791**
Percent Discordant	**9.9**	**Gamma**	**0.800**
Percent Tied	**1.1**	**Tau-a**	**0.379**
Pairs	**91**	**c**	**0.896**

Compared to model 1 with steps: Larger Somer's D, Gamma, Tau-A and C-index for model 1

BMI as predictor (Model 2)

Partition for the Hosmer and Lemeshow Test

		complete = 0		complete = 1	
Group	Total	Observed	Expected	Observed	Expected
1	2	0	0.02	2	1.98
2	2	1	0.08	1	1.92
3	2	0	0.10	2	1.90
4	2	0	0.14	2	1.86
5	2	0	0.31	2	1.69
6	2	0	0.42	2	1.58
7	2	0	0.96	2	1.04
8	2	2	1.33	0	0.67
9	2	2	1.68	0	0.32
10	2	2	1.97	0	0.03

Hosmer and Lemeshow Goodness-of-Fit Test

Chi-Square	DF	Pr > ChiSq
15.6474	**8**	**0.0477**

P < 0.05; Significant lack of fit. The current model is not adequate.

Mode 1: p=0.60521
i.e. Model 1 is adequate.

Model Comparison Summary:

	Model 1: Steps	Model 2: BMI	Choose	Model 1 or 2?
Model Comparison:				
-2LL	**8.554**	13.521		NOT appropriate
AIC	**12.554**	17.521	Smaller (diff of 2)	1
BIC	**14.546**	19.512	Smaller (diff of 2)	1
Measures of Predictive Power (based on log-likelihood):				
R-Square	**0.5799**	0.4614	Larger	1
Max-Rescaled R-Square	**0.7986**	0.6355	Larger	1
Test of H_0: Beta=0				
Likelihood Ratio p-value	**<.0001**	0.0004	Smaller	1
Wald p-value (Not appropriate for small n)	0.0277	**0.0259**	Smaller	Not Wald (n=20)
Measures of Predictive Power:				
Somers	**.912**	.791	Larger	1
Gamma	**.912**	.800	Larger	1
Tau-A	**.437**	.379	Larger	1
C-index	**.956**	.896	Larger	1
Goodness of Fit:				
Hosmer-Lemeshow Chi-Square	**6.3757**	15.6474	Smaller	1
Hosmer-Lemeshow p-value	**0.6052**	0.0477	Larger (fail to reject H0)	1

```
PROC LOGISTIC DATA=steps;
  MODEL complete = bmi steps / CLPARM=BOTH CLODDS=BOTH;
  UNIT steps=100;
RUN;
```


Residuals

Residuals are useful for identifying potential outliers or miss-specified models. Two types of residuals (reported by SAS) are Deviance residuals and Pearson residuals.

- **Deviance residuals:** Measures the contribution of each observation to the Deviance chi-square.
- **Pearson residuals:** Measures the contribution of each observation to the Pearson chi-square.

Influential observations

As in linear regression, we can use DFFITS and DFBETAS to identify influential observations. The statistics produced by SAS LOGISTIC include:

- **DFBETAS:** These statistics tell you how much each regression coefficient changes when a particular observation is deleted. The actual change is divided by the standard error of the coefficient.
- **DIFDEV:** Measures the change in the deviance with deletion of the observation.
- **DIFCHISQ:** Measures the change in the Pearson chi-square with deletion of the observation.
- **C and CBAR:** Measures the overall change in the regression coefficients (analogous to Cook's D in linear regression).
- **Hat Matrix Diagonal:** A measure of leverage or potential influence (how extreme the observation is in the "predictor space.")

```
PROC LOGISTIC;
  MODEL complete =  steps100 / INFLUENCE IPLOTS;
RUN;
```



```
        The LOGISTIC Procedure

              Model Information

Data Set                     WORK.STEPS
Response Variable            complete
Number of Response Levels    2
Model                        binary logit
Optimization Technique       Fisher's scoring

Number of Observations Read         20
Number of Observations Used         20

          Response Profile
 Ordered                       Total
   Value     complete      Frequency
       1            0              7
       2            1             13
Probability modeled is complete=0.

         Model Fit Statistics
                                 Intercept
              Intercept                and
Criterion          Only         Covariates
AIC              27.898             12.554
SC               28.894             14.546
-2 Log L         25.898              8.554
```

Testing Global Null Hypothesis: BETA=0

Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	17.3434	1	<.0001
Score	10.2874	1	0.0013
Wald	4.8496	1	0.0277

Analysis of Maximum Likelihood Estimates

Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	11.7273	5.5413	4.4789	0.0343
steps100	1	-0.2187	0.0993	4.8496	0.0277

The LOGISTIC Procedure

Odds Ratio Estimates

Effect	Point Estimate	95% Wald Confidence Limits	
steps100	0.804	0.661	0.976

Alternative way to obtain OR per 100 steps.

```
steps100 = steps/100;
```

Association of Predicted Probabilities and Observed Responses

Percent Concordant	95.6	Somers' D	0.912
Percent Discordant	4.4	Gamma	0.912
Percent Tied	0.0	Tau-a	0.437
Pairs	91	c	0.956

Residuals for Model 1(steps)

Output from INFLUENCE option

```
              Regression Diagnostics

                            Pearson Residual
         Covariates
     Case                                  (1 unit = 0.33)
    Number   steps100        Value        -8   -4   0  2  4  6  8

       1     62.4287        2.6225       |        |        *|
       2     50.3961       -1.4217       |    *   |         |
       3     62.2494       -0.3889       |       *|         |
       4     61.9036       -0.4038       |       *|         |
       5     44.8781        0.3847       |        |*        |
       6       100.5      -0.00594       |        *         |
       7     68.7279       -0.1915       |       *|         |
       8     80.1653       -0.0548       |        *         |
       9      5.6053       0.00524       |        *         |
      10     76.8327       -0.0789       |        *         |
      11     76.5274       -0.0816       |        *         |
      12     66.3754       -0.2476       |       *|         |
      13     43.6706        0.3371       |        |*        |
      14     41.9986        0.2808       |        |*        |
      15     93.4920       -0.0128       |        *         |
      16      9.3948       0.00794       |        *         |
      17     68.2067       -0.2027       |       *|         |
      18       117.7      -0.00091       |        *         |
      19     44.1538        0.3554       |        |*        |
      20     59.9475       -0.5002       |      * |         |
```

The first two observations are suspect in all diagnostic plots

- **Pearson residuals:** Measures the contribution of each observation to the Pearson chi-square.

Residuals for Model 1 (steps)

```
                            Regression Diagnostics

              Deviance Residual             Hat Matrix Diagonal
  Case              (1 unit = 0.25)                (1 unit = 0.01)
Number     Value   -8  -4  0 2 4 6 8       Value  0 2 4 6 8  12  16
     1    2.0318  |        |       *|     0.1489  |           *     |
     2   -1.4871  |  *     |        |     0.2258  |                *|
     3   -0.5307  |      * |        |     0.1503  |           *     |
     4   -0.5497  |      * |        |     0.1530  |           *     |
     5    0.5254  |        | *      |     0.2051  |               * |
     6  -0.00840  |        *        |   0.000748  |*                |
     7   -0.2683  |       *|        |     0.0904  |      *          |
     8   -0.0774  |        *        |     0.0210  | *               |
     9   0.00742  |        *        |   0.000686  |*                |
    10   -0.1114  |        *        |     0.0337  |  *              |
    11   -0.1152  |        *        |     0.0352  |  *              |
    12   -0.3450  |       *|        |     0.1128  |        *        |
    13    0.4639  |        | *      |     0.1909  |              *  |
    14    0.3895  |        | *      |     0.1684  |            *    |
    15   -0.0180  |        *        |    0.00250  |*                |
    16    0.0112  |        *        |    0.00135  |*                |
    17   -0.2838  |       *|        |     0.0953  |       *         |
    18  -0.00128  |        *        |   0.000033  |*                |
    19    0.4877  |        | *      |     0.1969  |              *  |
    20   -0.6683  |     *  |        |     0.1669  |            *    |
```

- **Deviance residuals:** Measures the contribution of each observation to the Deviance chi-square.

- **Hat Matrix Diagonal:** A measure of leverage or potential influence (how extreme the observation is in the "predictor space.")

Influential Observations for Model 1 (steps)

```
        Intercept
  Case     DfBeta        (1 unit = 0.07)
Number      Value       -8  -4  0 2 4 6 8
     1    -0.5926      |*       |        |
     2    -0.5199      | *      |        |
     3     0.0866      |        |*       |
     4     0.0871      |        |*       |
     5     0.1853      |        |  *     |
     6   0.000152      |        *        |
     7     0.0465      |        |*       |
     8    0.00708      |        *        |
     9   0.000137      |        *        |
    10     0.0127      |        *        |
    11     0.0134      |        *        |
    12     0.0628      |        |*       |
    13     0.1590      |        | *      |
    14     0.1252      |        | *      |
    15   0.000591      |        *        |
    16   0.000291      |        *        |
    17     0.0499      |        |*       |
    18   4.939E-6      |        *        |
    19     0.1694      |        | *      |
    20     0.0804      |        |*       |
```

- **DFBETAS for the intercept:** These statistics tell you how much each regression coefficient changes when a particular observation is deleted. The actual change is divided by the standard error of the coefficient.

Influential Observations for Model 1 (steps)

```
                                         Confidence Interval
                                            Displacement C
          steps100
   Case     DfBeta    (1 unit = 0.09)                    (1 unit = 0.09)
 Number      Value   -8  -4  0 2 4 6 8      Value   0 2 4 6 8  12  16
      1     0.7517  |        |       *|    1.4131  |                *|
      2     0.3995  |        |   *    |    0.7615  |         *       |
      3    -0.1105  |       *|        |    0.0315  |*                |
      4    -0.1127  |       *|        |    0.0348  |*                |
      5    -0.1639  |      * |        |    0.0480  | *               |
      6   -0.00016  |        *        |  2.644E-8  |*                |
      7    -0.0528  |       *|        |   0.00401  |*                |
      8   -0.00763  |        *        |  0.000066  |*                |
      9   -0.00014  |        *        |  1.888E-8  |*                |
     10    -0.0139  |        *        |  0.000225  |*                |
     11    -0.0146  |        *        |  0.000252  |*                |
     12    -0.0733  |       *|        |   0.00879  |*                |
     13    -0.1426  |      * |        |    0.0331  |*                |
     14    -0.1139  |       *|        |    0.0192  |*                |
     15   -0.00062  |        *        |  4.097E-7  |*                |
     16   -0.00029  |        *        |  8.506E-8  |*                |
     17    -0.0570  |       *|        |   0.00478  |*                |
     18   -5.13E-6  |        *        |  2.68E-11  |*                |
     19    -0.1511  |      * |        |    0.0386  |*                |
     20    -0.1168  |       *|        |    0.0602  | *               |
```

- **DFBETAS for the steps:** These statistics tell you how much each regression coefficient changes when a particular observation is deleted. The actual change is divided by the standard error of the coefficient.

- **C:** Measures the overall change in the regression coefficients (analogous to Cook's D in linear regression).

Influential Observations for Model 1 (steps)

```
Confidence Interval Displacement CBar          Delta Deviance
  Case              (1 unit = 0.08)                (1 unit = 0.33)
Number      Value  0 2 4 6 8  12  16       Value  0 2 4 6 8  12  16
      1    1.2028 |                *|     5.3309 |                *|
      2    0.5896 |        *        |     2.8009 |        *        |
      3    0.0267 |*                |     0.3084 | *               |
      4    0.0295 |*                |     0.3316 | *               |
      5    0.0382 | *               |     0.3142 | *               |
      6  2.642E-8 |*                |   0.000071 |*                |
      7   0.00365 |*                |     0.0756 |*                |
      8  0.000064 |*                |    0.00606 |*                |
      9  1.887E-8 |*                |   0.000055 |*                |
     10  0.000217 |*                |     0.0126 |*                |
     11  0.000243 |*                |     0.0135 |*                |
     12   0.00779 |*                |     0.1268 |*                |
     13    0.0268 |*                |     0.2421 | *               |
     14    0.0160 |*                |     0.1677 | *               |
     15  4.086E-7 |*                |   0.000326 |*                |
     16  8.495E-8 |*                |   0.000126 |*                |
     17   0.00433 |*                |     0.0848 |*                |
     18  2.68E-11 |*                |   1.649E-6 |*                |
     19    0.0310 |*                |     0.2688 | *               |
     20    0.0501 | *               |     0.4967 | *               |
```

- **CBAR:** Measures the overall change in the regression coefficients (analogous to Cook’s D in linear regression).

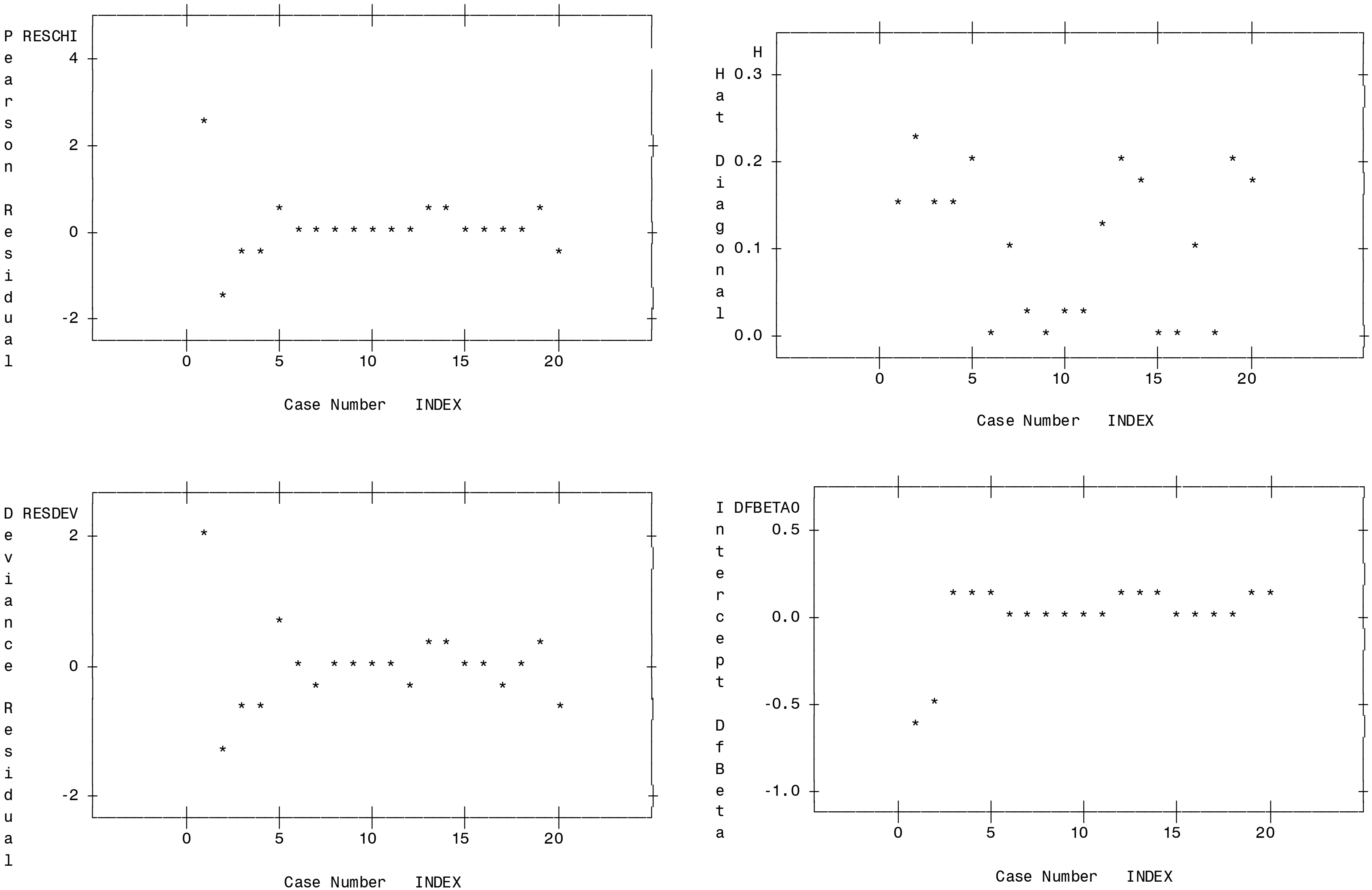
- **DIFDEV:** Measures the change in the deviance with deletion of the observation.

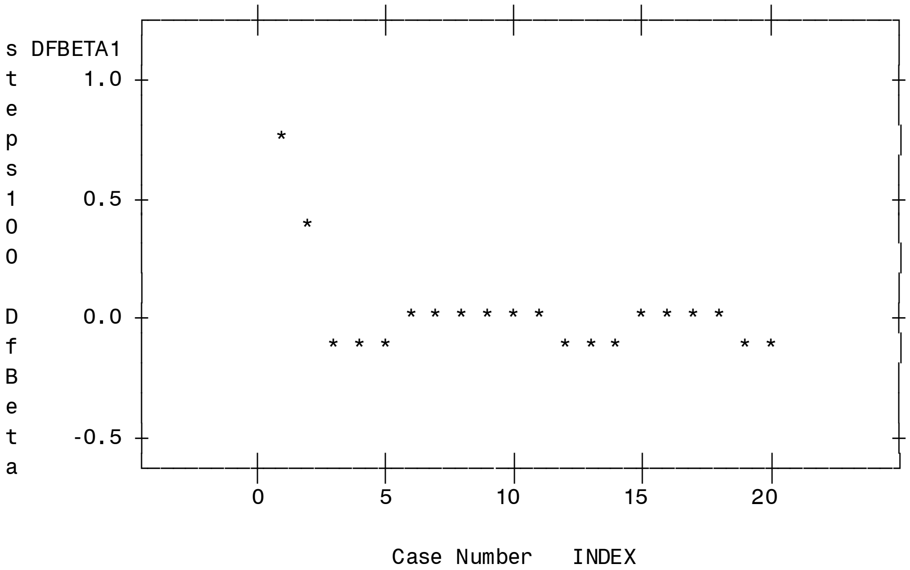
Influential Observations for Model 1 (steps)

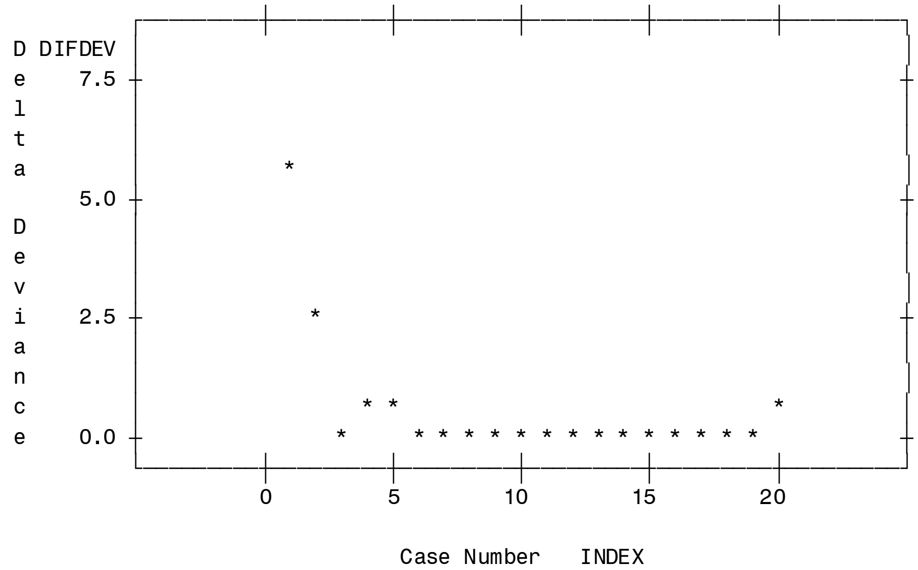
```
                  Delta Chi-Square

   Case                  (1 unit = 0.51)
 Number      Value      0 2 4 6 8  12  16
      1     8.0805     |                *|
      2     2.6108     |     *           |
      3     0.1780     |*                |
      4     0.1926     |*                |
      5     0.1862     |*                |
      6   0.000035     |*                |
      7     0.0403     |*                |
      8    0.00307     |*                |
      9   0.000028     |*                |
     10    0.00644     |*                |
     11    0.00690     |*                |
     12     0.0691     |*                |
     13     0.1404     |*                |
     14     0.0948     |*                |
     15   0.000163     |*                |
     16   0.000063     |*                |
     17     0.0454     |*                |
     18   8.245E-7     |*                |
     19     0.1573     |*                |
     20     0.3003     | *               |
```

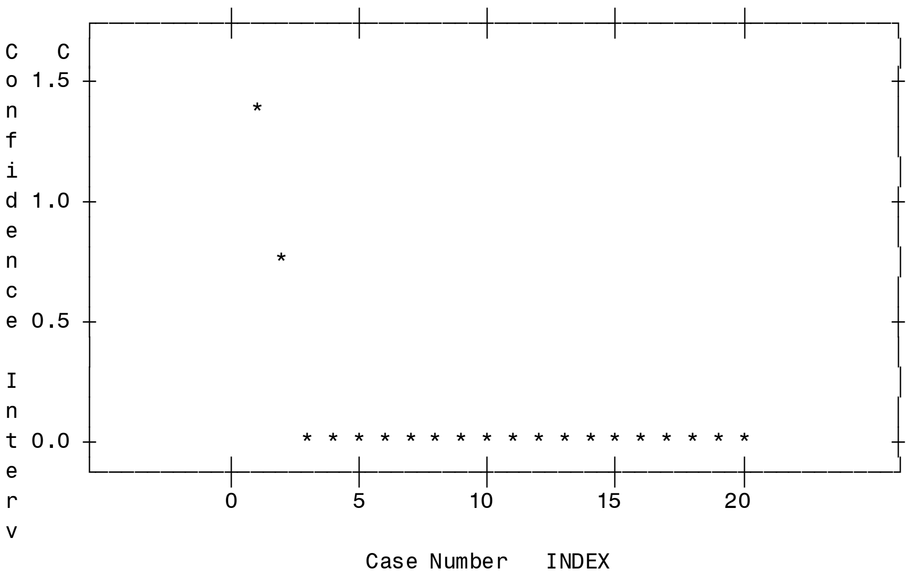
DIFCHISQ: Measures the change in the Pearson chi-square with deletion of the observation.

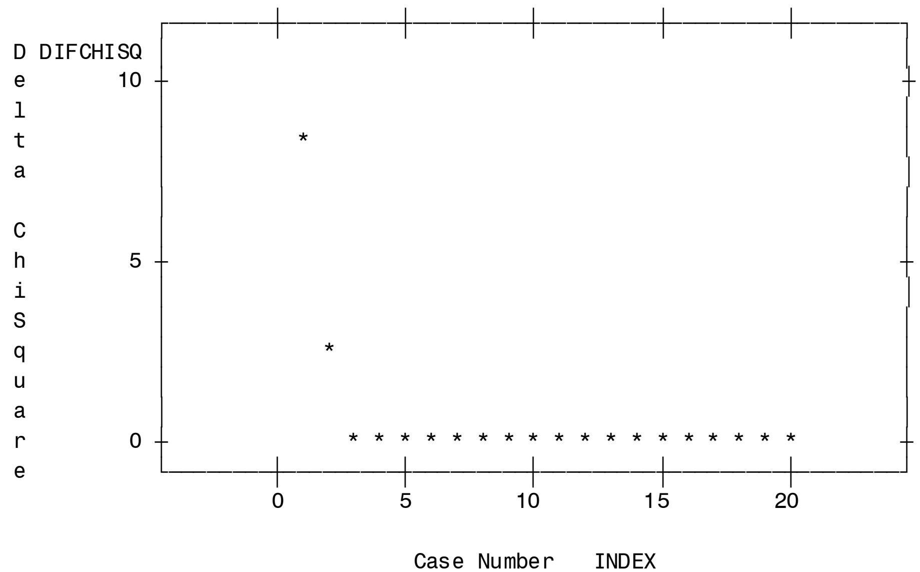
Output from IPLOTS option




DFBETA1
steps100 DfBeta
1.0
0.5
0.0
-0.5
0
5
10
15
20
Case Number INDEX


DIFDEV
Delta Deviance
7.5
5.0
2.5
0.0
0
5
10
15
20
Case Number INDEX


C
Confidence Interv
1.5
1.0
0.5
0.0
0
5
10
15
20
Case Number INDEX


DIFCHISQ
Delta ChiSquare
10
5
0
0
5
10
15
20
Case Number INDEX

Influential Observations for Model 1 (steps)

- Maybe case number 1 and 2?
    - First few observations are always suspect
    - May be true outliers or NOT
- Always report the results with and without those observations included
- Be VERY careful deleting outliers