```
1          ************************************************************************;
2          *** Logistic regression example -                                     ***;
3          *** Data is from Statistical Methods II classes in recent years      ***;
4          *** The objective is to determine the probability of getting an "A"   ***;
5          ***    in the class from the grade on the first exam.                ***;
6          ************************************************************************;
7
8          options ps=256 ls=88 nocenter nodate nonumber;
9
10         data grades; infile
10       ! "C:\Geaghan\EXST\EXST7015New\Fall2002\SAS\05-LogisticReg.DAT" missover;
11              TITLE1 'EXST7015: Probability  of A grade in EXST7015';
12           input Semester $ Exam1 Grade_A $;
13              if exam1 eq . then delete;
14              interval = 5; Score1 = int(exam1/interval)*interval + (interval/2);
15              if score1 gt 100 then score1=100;
16              indicator = 0; if Grade_A eq 'TRUE' then indicator = 1;
17         cards;
NOTE: The infile "C:\Geaghan\EXST\EXST7015New\Fall2002\SAS\05-LogisticReg.DAT" is:
      File Name=C:\Geaghan\EXST\EXST7015New\Fall2002\SAS\05-LogisticReg.DAT,
      RECFM=V,LRECL=256
NOTE: 424 records were read from the infile
      "C:\Geaghan\EXST\EXST7015New\Fall2002\SAS\05-LogisticReg.DAT".
      The minimum record length was 0.
      The maximum record length was 14.
NOTE: The data set WORK.GRADES has 423 observations and 6 variables.
NOTE: DATA statement used:
      real time           0.06 seconds
      cpu time            0.06 seconds
17       !        run;
18         ;
19         proc sort data=grades; by exam1; run;
NOTE: There were 423 observations read from the data set WORK.GRADES.
NOTE: The data set WORK.GRADES has 423 observations and 6 variables.
NOTE: PROCEDURE SORT used:
      real time           0.05 seconds
      cpu time            0.05 seconds
21         proc freq data=grades; table score1*Grade_A / norow nocol nopercent;
22              TITLE2 'Simple frequencies by 5 point groupings';
23         run;
NOTE: There were 423 observations read from the data set WORK.GRADES.
NOTE: The PROCEDURE FREQ printed page 1.
NOTE: PROCEDURE FREQ used:
      real time           0.09 seconds
      cpu time            0.09 seconds

EXST7015: Probability  of A grade in EXST7015
Simple frequencies by 5 point groupings

The FREQ Procedure
Table of Score1 by Grade_A
Score1     Grade_A
Frequency|FALSE   |TRUE    |  Total
---------+--------+--------+
    52.5 |      1 |      0 |      1
---------+--------+--------+
    57.5 |      4 |      1 |      5
---------+--------+--------+
    62.5 |      5 |      0 |      5
---------+--------+--------+
    67.5 |     12 |      1 |     13
---------+--------+--------+
    72.5 |     25 |      1 |     26
---------+--------+--------+
    77.5 |     40 |      7 |     47
---------+--------+--------+
    82.5 |     51 |     14 |     65
---------+--------+--------+
    87.5 |     41 |     45 |     86
---------+--------+--------+
    92.5 |     23 |     88 |    111
---------+--------+--------+
    97.5 |      7 |     51 |     58
---------+--------+--------+
     100 |      2 |      4 |      6
---------+--------+--------+
Total         211      212      423
```

```
25         proc means data=grades mean max min std stderr print; var exam1;
26              TITLE2 'Raw data mean';
27         run;
NOTE: There were 423 observations read from the data set WORK.GRADES.
NOTE: The PROCEDURE MEANS printed page 2.
NOTE: PROCEDURE MEANS used:
      real time           0.02 seconds
      cpu time            0.02 seconds

EXST7015: Probability  of A grade in EXST7015
Raw data mean

The MEANS Procedure
                          Analysis Variable : Exam1
        Mean         Maximum         Minimum         Std Dev       Std Error
-------------------------------------------------------------------------------
  85.8628842     100.0000000      52.0000000       9.0178926       0.4384649
-------------------------------------------------------------------------------


29         proc logistic data=grades DESCENDING; TITLE2 'Logistic
regression';
30            model Grade_A = exam1;
31            output out=next1 PREDICTED=yhat Lower=lcl Upper=ucl;
32         run;
NOTE: PROC LOGISTIC is modeling the probability that Grade_A='TRUE'.
NOTE: Convergence criterion (GCONV=1E-8) satisfied.
NOTE: There were 423 observations read from the data set WORK.GRADES.
NOTE: The data set WORK.NEXT1 has 423 observations and 10 variables.
NOTE: The PROCEDURE LOGISTIC printed page 3.
NOTE: PROCEDURE LOGISTIC used:
      real time           0.08 seconds
      cpu time            0.08 seconds


EXST7015: Probability  of A grade in EXST7015
Logistic regression

The LOGISTIC Procedure

              Model Information
Data Set                          WORK.GRADES
Response Variable                 Grade_A
Number of Response Levels         2
Number of Observations            423
Model                             binary logit
Optimization Technique            Fisher's scoring

          Response Profile
 Ordered                          Total
   Value     Grade_A      Frequency
       1     TRUE               212
       2     FALSE              211

Probability modeled is Grade_A='TRUE'.

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

Model Fit Statistics

Criterion	Intercept Only	Intercept and Covariates
AIC	588.400	425.407
SC	592.448	433.502
-2 Log L	586.400	421.407

Testing Global Null Hypothesis: BETA=0

Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	164.9934	1	<.0001
Score	132.7164	1	<.0001
Wald	96.1179	1	<.0001

Analysis of Maximum Likelihood Estimates

Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	-16.9098	1.7443	93.9760	<.0001
Exam1	1	0.1952	0.0199	96.1179	<.0001

Odds Ratio Estimates

Effect	Point Estimate	95% Wald Confidence Limits	
Exam1	1.216	1.169	1.264

Association of Predicted Probabilities and Observed Responses

Percent Concordant	82.8	Somers' D	0.681
Percent Discordant	14.7	Gamma	0.698
Percent Tied	2.4	Tau-a	0.341
Pairs	44732	c	0.841

```
34         proc sort data=next1 nodupkey; by exam1; run;
NOTE: 380 observations with duplicate key values were deleted.
NOTE: There were 423 observations read from the data set WORK.NEXT1.
NOTE: The data set WORK.NEXT1 has 43 observations and 10 variables.
NOTE: PROCEDURE SORT used:
      real time           0.04 seconds
      cpu time            0.04 seconds
35         proc print data=next1; var yhat lcl ucl;
36              TITLE2 'Listing of one kept value for each value of exam1';
37         run;
NOTE: There were 43 observations read from the data set WORK.NEXT1.
NOTE: The PROCEDURE PRINT printed page 4.
NOTE: PROCEDURE PRINT used:
      real time           0.02 seconds
      cpu time            0.02 seconds
```

EXST7015: Probability of A grade in EXST7015
Listing of one kept value for each value of exam1

Obs	yhat	lcl	ucl		Obs	yhat	lcl	ucl
1	0.00116	0.00029	0.00469		22	0.18403	0.13095	0.25238
2	0.00208	0.00058	0.00748		23	0.21516	0.15891	0.28459
3	0.00307	0.00092	0.01021		24	0.24995	0.19128	0.31951
4	0.00373	0.00116	0.01192		25	0.28829	0.22807	0.35706
5	0.00453	0.00146	0.01392		26	0.32993	0.26906	0.39710
6	0.00550	0.00185	0.01625		27	0.37442	0.31367	0.43940
7	0.00983	0.00371	0.02579		28	0.42114	0.36104	0.48368
8	0.01192	0.00468	0.03005		29	0.46932	0.41001	0.52950
9	0.01752	0.00743	0.04073		30	0.51807	0.45935	0.57629
10	0.02121	0.00936	0.04736		31	0.56648	0.50791	0.62325
11	0.02567	0.01178	0.05501		32	0.61365	0.55474	0.66941
12	0.03103	0.01482	0.06383		33	0.65879	0.59923	0.71372
13	0.03747	0.01862	0.07396		34	0.70121	0.64098	0.75520
14	0.04518	0.02337	0.08557		35	0.74045	0.67981	0.79309
15	0.05439	0.02928	0.09883		36	0.77617	0.71563	0.82693
16	0.06535	0.03663	0.11392		37	0.80825	0.74845	0.85656
17	0.07833	0.04571	0.13103		38	0.83670	0.77830	0.88205
18	0.09363	0.05689	0.15032		39	0.86165	0.80529	0.90365
19	0.11156	0.07057	0.17197		40	0.88332	0.82954	0.92174
20	0.13243	0.08717	0.19613		41	0.90198	0.85120	0.93673
21	0.15651	0.10716	0.22290		42	0.91794	0.87045	0.94904
					43	0.93149	0.88747	0.95909

```
39         proc sort data=grades; by score1; run;
NOTE: There were 423 observations read from the data set WORK.GRADES.
NOTE: The data set WORK.GRADES has 423 observations and 6 variables.
NOTE: PROCEDURE SORT used:
      real time           0.04 seconds
      cpu time            0.04 seconds
40         proc sort data=next1; by score1; run;
NOTE: There were 43 observations read from the data set WORK.NEXT1.
NOTE: The data set WORK.NEXT1 has 43 observations and 10 variables.
NOTE: PROCEDURE SORT used:
      real time           0.03 seconds
      cpu time            0.03 seconds
41         proc means data=grades noprint; by score1; var indicator;
42              output out=next2 n=n mean=mean var=var; run;
NOTE: There were 423 observations read from the data set WORK.GRADES.
NOTE: The data set WORK.NEXT2 has 11 observations and 6 variables.
NOTE: PROCEDURE MEANS used:
      real time           0.04 seconds
      cpu time            0.04 seconds
43
44         data two; set next1 next2; run;
NOTE: There were 43 observations read from the data set WORK.NEXT1.
NOTE: There were 11 observations read from the data set WORK.NEXT2.
NOTE: The data set WORK.TWO has 54 observations and 15 variables.
NOTE: DATA statement used:
      real time           0.05 seconds
      cpu time            0.05 seconds
45         options ps=56 ls=111;
46         proc plot data=two;  plot yhat*exam1='x' mean*score1='o' /
overlay;
47              TITLE2 'Plot of observed means (o) and predicted values (p)';
48         run;
```

```
EXST7015: Probability  of A grade in EXST7015
Plot of observed means (o) and predicted values (p)

                          Plot of yhat*Exam1.   Symbol used is 'x'.
                          Plot of mean*Score1.  Symbol used is 'o'.

                                                                                                    x
                                                                                                  x x
                                                                                                 xo
                                                                                               x
                                                                                             x
                                                                                           x
                                                                                         ox
                                                                                        x
                                                                                      x
E                                                                                   x                 o
s                                                                                 x
t                                                                               x
i                                                                             xo
m                                                                           x
a                                                                         x
t                                                                       x
e                                                                     x
d                                                                   x
                                                                  x
P                                                               x
r                                                             x
o                                                           x
b                                                         x
a                                                       x
b                                                     x
i                                                   x
l                                                 x
i                                               x
t                                             x   o
y                                           x
                                          xox
                                        x
                                      x x
                                   o      x x x
                                x x x x x    o
                             xo    x   x x x x    ox x
```

```
     +---------+---------+---------+---------+---------+---------+---------+---------+---------+---------+---
     50        55        60        65        70        75        80        85        90        95        100
                                               Exam1
```

```
NOTE: 54 obs had missing values.
```