

CHAPTER 8 ANALYSIS EXAMPLES REPLICATION SAS V9.2

GENERAL NOTES ABOUT ANALYSIS EXAMPLES REPLICATION

These examples are intended to provide guidance on how to use the commands/procedures for analysis of complex sample survey data and assume all data management and other preliminary work is done. The relevant syntax for the procedure of interest is shown first along with the associated output for that procedure(s). In some examples, there may be more than one block of syntax and in this case all syntax is first presented followed by the output produced.

In some software packages certain procedures or options are not available but we have made every attempt to demonstrate how to match the output produced by Stata 10+ in the textbook. Check the ASDA website for updates to the various software tools we cover.

GENERAL NOTES ON LOGISTIC REGRESSION AND PROC SURVEYLOGISTIC

PROC SURVEYLOGISTIC is the general purpose tool for logistic regression with complex sample survey data. This procedure is a multi-purpose tool that can do correct subpopulation analyses and offers a number of output options such as a class statement for categorical predictors, and a test statement for custom hypothesis testing of parameters.

This chapter first focuses on simple logistic regression with a binary outcome and therefore uses the default link of logit. It then includes a comparison of the logit link with the probit and Clog-log links for Probit and Complementary Log-Log models. One thing to note when comparing SAS PROC SURVEYLOGISTIC SE's to svy: logit: the SAS SE's do not match exactly due to differences in how Stata handles the SE calculations between Stata svy: logit and svy: logistic. See the Stata documentation for further details.

Some options to note: use of the class statement requires the / param=ref specification if you want to use a reference group parameterization instead of the default effects coding approach, use of the (ref=first) allows specification of the omitted category for the class variables, and use of the (event='1') syntax declares the probability modeled for the outcome variable. There are examples of these options in this chapter. PROC SURVEYLOGISTIC also allows the use of the test statement and use of the crossing operator for interaction variables in the model statement.

```

title "Analysis Example 8.1: Contingency Tables: NCSR" ;
proc surveyfreq data=ncsr ;
strata sestrat ;
cluster seclustr ;
weight ncsrwtlg ;
tables (ag4cat sexm ald ed4cat mar3cat)*mde / row chisq ;
format ag4cat agf. sexm sexmf. ald mdef. ed4cat edf. mar3cat marf. mde mdef. ;
run ;

```

Analysis Example 8.1: Contingency Tables: NCSR

The SURVEYFREQ Procedure

Data Summary

```

Number of Strata              42
Number of Clusters            84
Number of Observations        9282
Number of Observations Used   5692
Number of Obs with Nonpositive Weights 3590
Sum of Weights                5692.00048

```

Table of ag4cat by mde

ag4cat	mde	Frequency	Weighted Frequency	Std Dev of Wgt Freq	Percent	Std Err of Percent	Row Percent	Std Err of Row Percent
18-29	No	974	1091	76.02329	19.1743	0.9132	81.6012	0.8854
	Yes	397	246.08116	21.73234	4.3233	0.3073	18.3988	0.8854
	Total	1371	1337	93.35782	23.4976	1.1167	100.000	
30-44	No	1176	1267	64.54216	22.2561	0.8483	77.1234	1.1045
	Yes	650	375.76746	22.57099	6.6017	0.3006	22.8766	1.1045
	Total	1826	1643	75.91359	28.8578	0.8673	100.000	
45-59	No	997	1169	77.92918	20.5443	0.9923	77.6673	1.2607
	Yes	524	336.24696	29.16619	5.9074	0.3410	22.3327	1.2607
	Total	1521	1506	97.73017	26.4516	1.0712	100.000	
60+	No	749	1073	68.78690	18.8488	1.0286	88.9391	0.9563
	Yes	225	133.42778	12.10988	2.3441	0.1614	11.0609	0.9563
	Total	974	1206	72.75483	21.1930	1.0038	100.000	
Total	No	3896	4600	194.66294	80.8236	0.6407		
	Yes	1796	1092	68.88569	19.1764	0.6407		
	Total	5692	5692	251.09560	100.000			

Rao-Scott Chi-Square Test

```

Pearson Chi-Square  75.9697
Design Correction   0.9423

```

```

Rao-Scott Chi-Square 80.6232
DF                    3
Pr > ChiSq           <.0001

```

```

F Value            26.8744
Num DF             3
Den DF            126
Pr > F             <.0001

```

Sample Size = 5692

Analysis Example 8.1: Contingency Tables: NCSR

The SURVEYFREQ Procedure

Table of sexm by mde

sexm	mde	Frequency	Weighted Frequency	Std Dev of Wgt Freq	Percent	Std Err of Percent	Row Percent	Std Err of Row Percent
Female	No	2117	2337	123.28903	41.0494	0.9079	77.3830	0.6728
	Yes	1193	682.90913	38.66449	11.9977	0.3932	22.6170	0.6728
	Total	3310	3019	154.57970	53.0471	1.0105	100.000	
Male	No	1779	2264	100.28695	39.7741	1.1391	84.7107	0.9138
	Yes	603	408.61423	33.68576	7.1787	0.3731	15.2893	0.9138
	Total	2382	2673	121.20178	46.9529	1.0105	100.000	
Total	No	3896	4600	194.66294	80.8236	0.6407		
	Yes	1796	1092	68.88569	19.1764	0.6407		
	Total	5692	5692	251.09560	100.000			

Rao-Scott Chi-Square Test

Pearson Chi-Square 49.1166
 Design Correction 0.9961

Rao-Scott Chi-Square 49.3081
 DF 1
 Pr > ChiSq <.0001

F Value 49.3081
 Num DF 1
 Den DF 42
 Pr > F <.0001

Sample Size = 5692

Table of ald by mde

ald	mde	Frequency	Weighted Frequency	Std Dev of Wgt Freq	Percent	Std Err of Percent	Row Percent	Std Err of Row Percent
No	No	3664	4432	183.07460	77.8586	0.7366	82.3086	0.6507
	Yes	1585	952.55085	63.26164	16.7349	0.6019	17.6914	0.6507
	Total	5249	5384	234.33376	94.5935	0.3248	100.000	
Yes	No	232	168.76731	19.21233	2.9650	0.2706	54.8409	2.9025
	Yes	211	138.97251	11.25575	2.4415	0.1796	45.1591	2.9025
	Total	443	307.73983	25.12875	5.4065	0.3248	100.000	
Total	No	3896	4600	194.66294	80.8236	0.6407		
	Yes	1796	1092	68.88569	19.1764	0.6407		
	Total	5692	5692	251.09560	100.000			

Analysis Example 8.1: Contingency Tables: NCSR

The SURVEYFREQ Procedure

Table of ald by mde

Rao-Scott Chi-Square Test

Pearson Chi-Square 141.7044
 Design Correction 1.1047

Rao-Scott Chi-Square 128.2716
 DF 1
 Pr > ChiSq <.0001

F Value 128.2716
 Num DF 1
 Den DF 42
 Pr > F <.0001

Sample Size = 5692

Table of ED4CAT by mde

ED4CAT	mde	Frequency	Weighted Frequency	Std Dev of Wgt Freq	Percent	Std Err of Percent	Row Percent	Std Err of Row Percent
0-11 Yrs	No	613	798.37379	51.04296	14.0262	0.8281	83.6916	1.2121
	Yes	236	155.57319	14.85368	2.7332	0.1917	16.3084	1.2121
	Total	849	953.94698	59.04526	16.7594	0.8617	100.000	
12 Yrs	No	1177	1508	92.13927	26.4916	0.9349	81.4503	0.8272
	Yes	535	343.41297	24.89564	6.0333	0.3480	18.5497	0.8272
	Total	1712	1851	110.50869	32.5249	1.1167	100.000	
11-13 Yrs	No	1139	1235	66.33923	21.6922	0.7670	78.7515	1.0439
	Yes	570	333.14937	22.24957	5.8529	0.2640	21.2485	1.0439
	Total	1709	1568	79.48027	27.5452	0.7622	100.000	
16+ Yrs	No	967	1059	65.88639	18.6135	0.8864	80.3325	1.0876
	Yes	455	259.38783	24.27390	4.5571	0.3183	19.6675	1.0876
	Total	1422	1319	84.24656	23.1706	1.0446	100.000	
Total	No	3896	4600	194.66294	80.8236	0.6407		
	Yes	1796	1092	68.88569	19.1764	0.6407		
	Total	5692	5692	251.09560	100.000			

Rao-Scott Chi-Square Test

Pearson Chi-Square 10.0806
 Design Correction 0.7887

Rao-Scott Chi-Square 12.7821
 DF 3
 Pr > ChiSq 0.0051

F Value 4.2607
 Num DF 3
 Den DF 126
 Pr > F 0.0067

Sample Size = 5692

Analysis Example 8.1: Contingency Tables: NCSR

The SURVEYFREQ Procedure

Table of MAR3CAT by mde

MAR3CAT	mde	Frequency	Weighted Frequency	Std Dev of Wgt Freq	Percent	Std Err of Percent	Row Percent	Std Err of Row Percent
Married	No	2316	2633	129.78766	46.2527	1.1889	82.6738	0.7420
	Yes	920	551.74202	41.57012	9.6933	0.4206	17.3262	0.7420
	Total	3236	3184	162.63940	55.9460	1.2360	100.000	
Previously Married	No	750	901.32010	51.92644	15.8349	0.7256	76.0979	1.4494
	Yes	489	283.10112	19.65880	4.9737	0.2597	23.9021	1.4494
	Total	1239	1184	59.49737	20.8085	0.6883	100.000	
Never Married	No	830	1066	75.59104	18.7360	1.0212	80.6006	1.1549
	Yes	387	256.68022	20.56723	4.5095	0.3091	19.3994	1.1549
	Total	1217	1323	88.18366	23.2455	1.1512	100.000	
Total	No	3896	4600	194.66294	80.8236	0.6407		
	Yes	1796	1092	68.88569	19.1764	0.6407		
	Total	5692	5692	251.09560	100.000			

Rao-Scott Chi-Square Test

Pearson Chi-Square 24.1420
Design Correction 1.1063

Rao-Scott Chi-Square 21.8218
DF 2
Pr > ChiSq <.0001

F Value 10.9109
Num DF 2
Den DF 84
Pr > F <.0001

Sample Size = 5692

```

title "Analysis Example 8.1: Logistic Regression with Binary Outcome: NCSR" ;
*NOTE: No deff or Rao Scott F test options in PROC SURVEYLOGISTIC ;

```

```

proc surveylogistic data=ncsr order=internal ;
strata sestrat ;
cluster seclustr ;
weight ncsrwtlg ;
class ag4cat (ref=first) ed4cat (ref=first) mar3cat (ref=first) / param=ref ;
model mde (event='Yes')= ag4cat sexm ald ed4cat mar3cat;
format ag4cat agf. sexm sexmf. ald mdef. ed4cat edf. mar3cat marf. mde mdef. ;
run ;

```

Analysis Example 8.1: Logistic Regression with Binary Outcome: NCSR
The SURVEYLOGISTIC Procedure

Model Information		
Data Set	WORK.NCSR	
Response Variable	mde	
Number of Response Levels	2	
Stratum Variable	SESTRAT	SAMPLING ERROR STRATUM
Number of Strata	42	
Cluster Variable	SECLUSTR	SAMPLING ERROR CLUSTER
Number of Clusters	84	
Weight Variable	NCSRWTLG	NCSR sample part 2 weight
Model	Binary Logit	
Optimization Technique	Fisher's Scoring	
Variance Adjustment	Degrees of Freedom (DF)	

Variance Estimation		
Method	Taylor Series	
Variance Adjustment	Degrees of Freedom (DF)	
Number of Observations Read	9282	
Number of Observations Used	5692	
Sum of Weights Read	5692	
Sum of Weights Used	5692	

Response Profile			
Ordered Value	mde	Total Frequency	Total Weight
1	No	3896	4600.4771
2	Yes	1796	1091.5234

Probability modeled is mde='Yes'.

NOTE: 3590 observations having nonpositive frequencies or weights were excluded since they do not contribute to the analysis.

Class Level Information				
Class	Value	Design Variables		
ag4cat	1=18-29	0	0	0
	2=30-44	1	0	0
	3=45-59	0	1	0
	4=60+	0	0	1
ED4CAT	1=0-11 Yrs	0	0	0
	2=12 Yrs	1	0	0
	3=11-13 Yrs	0	1	0
	4=16+ Yrs	0	0	1
MAR3CAT	1=Married	0	0	
	2=Previously Married	1	0	
	3=Never Married	0	1	

Analysis Example 8.1: Logistic Regression with Binary Outcome: NCSR
The SURVEYLOGISTIC Procedure

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

Criterion	Intercept and Covariates	
	Intercept Only	
AIC	5566.174	5290.526
SC	5572.820	5363.641
-2 Log L	5564.174	5268.526

Testing Global Null Hypothesis: BETA=0

Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	295.6471	10	<.0001
Score	307.6436	10	<.0001
Wald	356.6201	10	<.0001

Type 3 Analysis of Effects

Effect	DF	Wald	
		Chi-Square	Pr > ChiSq
ag4cat	3	59.8356	<.0001
sexm	1	55.8073	<.0001
ald	1	85.1358	<.0001
ED4CAT	3	6.6969	0.0822
MAR3CAT	2	33.9545	<.0001

Analysis of Maximum Likelihood Estimates

Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
ag4cat 2=30-44	1	0.2556	0.0945	7.3219	0.0068
ag4cat 3=45-59	1	0.2064	0.0916	5.0785	0.0242
ag4cat 4=60+	1	-0.6757	0.1414	22.8275	<.0001
sexm	1	-0.5773	0.0773	55.8073	<.0001
ald	1	1.4237	0.1543	85.1358	<.0001
ED4CAT 2=12 Yrs	1	0.0792	0.0970	0.6678	0.4138
ED4CAT 3=11-13 Yrs	1	0.2305	0.0932	6.1230	0.0133
ED4CAT 4=16+ Yrs	1	0.1629	0.1107	2.1655	0.1411
MAR3CAT 2=Previously Married	1	0.4864	0.0855	32.3693	<.0001
MAR3CAT 3=Never Married	1	0.1156	0.1080	1.1459	0.2844

Odds Ratio Estimates

Effect	Point Estimate	95% Wald Confidence Limits	
ag4cat 2=30-44 vs 1=18-29	1.291	1.073	1.554
ag4cat 3=45-59 vs 1=18-29	1.229	1.027	1.471
ag4cat 4=60+ vs 1=18-29	0.509	0.386	0.671
sexm	0.561	0.482	0.653
ald	4.152	3.069	5.619
ED4CAT 2=12 Yrs vs 1=0-11 Yrs	1.082	0.895	1.309
ED4CAT 3=11-13 Yrs vs 1=0-11 Yrs	1.259	1.049	1.511
ED4CAT 4=16+ Yrs vs 1=0-11 Yrs	1.177	0.947	1.462
MAR3CAT 2=Previously Married vs 1=Married	1.626	1.376	1.923
MAR3CAT 3=Never Married vs 1=Married	1.123	0.908	1.387

Association of Predicted Probabilities and Observed Responses

Percent Concordant	61.5	Somers' D	0.247
Percent Discordant	36.8	Gamma	0.252
Percent Tied	1.7	Tau-a	0.107
Pairs	6997216	c	0.624

```

title "Analysis Example 8.1: Logistic Regression with Binary Outcome and Interaction Variables: NCSR" ;
proc surveylogistic data=ncsr ;
strata sestrat ;
cluster seclustr ;
weight ncsrwtlg ;
class ag4cat (ref=first) ed4cat (ref=first) mar3cat (ref=first) / param=ref ;
model mde (event='1')= ag4cat sexM ald ed4cat mar3cat ag4cat*sexM ald*sexM ed4cat*sexM mar3cat*sexM ;
format ag4cat agf. sexm sexmf. ald mdef. ed4cat edf. mar3cat marf. ;
run ;

```

Analysis Example 8.1: Logistic Regression with Binary Outcome and Interaction Variables: NCSR
The SURVEYLOGISTIC Procedure

		Model Information	
Data Set	WORK.NCSR		
Response Variable	mde		
Number of Response Levels	2		
Stratum Variable	SESTRAT		SAMPLING ERROR STRATUM
Number of Strata	42		
Cluster Variable	SECLUSTR		SAMPLING ERROR CLUSTER
Number of Clusters	84		
Weight Variable	NCSRWTLG		NCSR sample part 2 weight
Model	Binary Logit		
Optimization Technique	Fisher's Scoring		
Variance Adjustment	Degrees of Freedom (DF)		

Variance Estimation		
Method	Taylor Series	
Variance Adjustment	Degrees of Freedom (DF)	
Number of Observations Read	9282	
Number of Observations Used	5692	
Sum of Weights Read	5692	
Sum of Weights Used	5692	

Response Profile			
Ordered Value	mde	Total Frequency	Total Weight
1	0	3896	4600.4771
2	1	1796	1091.5234

Probability modeled is mde=1.

NOTE: 3590 observations having nonpositive frequencies or weights were excluded since they do not contribute to the analysis.

Class Level Information				
Class	Value	Design Variables		
ag4cat	1=18-29	0	0	0
	2=30-44	1	0	0
	3=45-59	0	1	0
	4=60+	0	0	1
ED4CAT	1=0-11 Yrs	0	0	0
	2=12 Yrs	1	0	0
	3=11-13 Yrs	0	1	0
	4=16+ Yrs	0	0	1
MAR3CAT	1=Married	0	0	
	2=Previously Married	1	0	
	3=Never Married	0	1	

Analysis Example 8.1: Logistic Regression with Binary Outcome and Interaction Variables: NCSR

The SURVEYLOGISTIC Procedure

Model Convergence Status

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

Model Fit Statistics

Criterion	Intercept and	
	Intercept Only	Covariates
AIC	5566.174	5304.858
SC	5572.820	5437.794
-2 Log L	5564.174	5264.858

Testing Global Null Hypothesis: BETA=0

Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	299.3160	19	<.0001
Score	317.6468	19	<.0001
Wald	568.3284	19	<.0001

Type 3 Analysis of Effects

Effect	DF	Wald	
		Chi-Square	Pr > ChiSq
ag4cat	3	40.0285	<.0001
sexm	1	2.3329	0.1267
ald	1	53.9894	<.0001
ED4CAT	3	7.0727	0.0696
MAR3CAT	2	14.3847	0.0008
sexm*ag4cat	3	0.7773	0.8549
sexm*ald	1	0.6822	0.4088
sexm*ED4CAT	3	0.3942	0.9414
sexm*MAR3CAT	2	1.5623	0.4579

Analysis of Maximum Likelihood Estimates

Parameter	DF	Estimate	Standard Error	Wald	
				Chi-Square	Pr > ChiSq
Intercept	1	-1.5999	0.1342	142.0548	<.0001
ag4cat 2=30-44	1	0.2204	0.1140	3.7411	0.0531
ag4cat 3=45-59	1	0.2146	0.1027	4.3709	0.0366
ag4cat 4=60+	1	-0.6456	0.1755	13.5327	0.0002
sexm	1	-0.5464	0.3578	2.3329	0.1267
ald	1	1.5531	0.2114	53.9894	<.0001
ED4CAT 2=12 Yrs	1	0.1305	0.0839	2.4229	0.1196
ED4CAT 3=11-13 Yrs	1	0.2973	0.1172	6.4317	0.0112
ED4CAT 4=16+ Yrs	1	0.2422	0.1521	2.5352	0.1113
MAR3CAT 2=Previously Married	1	0.4178	0.1107	14.2408	0.0002
MAR3CAT 3=Never Married	1	0.0173	0.1300	0.0178	0.8939
sexm*ag4cat 2=30-44	1	0.0967	0.2012	0.2313	0.6306
sexm*ag4cat 3=45-59	1	0.00264	0.2132	0.0002	0.9901
sexm*ag4cat 4=60+	1	-0.0378	0.3025	0.0156	0.9005
sexm*ald	1	-0.2004	0.2427	0.6822	0.4088
sexm*ED4CAT 2=12 Yrs	1	-0.1378	0.2715	0.2576	0.6118
sexm*ED4CAT 3=11-13 Yrs	1	-0.1688	0.2698	0.3914	0.5316
sexm*ED4CAT 4=16+ Yrs	1	-0.1940	0.3447	0.3167	0.5736
sexm*MAR3CAT 2=Previously Married	1	0.1825	0.2082	0.7682	0.3808
sexm*MAR3CAT 3=Never Married	1	0.2319	0.2123	1.1930	0.2747

Association of Predicted Probabilities and Observed Responses

Percent Concordant	61.4	Somers' D	0.250
Percent Discordant	36.4	Gamma	0.256
Percent Tied	2.2	Tau-a	0.108
Pairs	6997216	c	0.625

```

title "Analysis Example Comparison of Logistic, Probit and CLogLog Regression with Binary Outcome: NCSR" ;
proc surveyl logistic data=ncsr ;
title2 "Logit Link" ;
strata sestrat ;
cluster seclustr ;
weight ncsrwtlg ;
class ag4cat (ref=first) ed4cat (ref=first) mar3cat (ref=first) sex (ref=last) / param=ref ;
model ald (event='1')= ag4cat sex ed4cat mar3cat ;
format ag4cat agf. sexm sexmf. ed4cat edf. mar3cat marf. ;
run ;

```

Analysis Example Comparison of Logistic, Probit and CLogLog Regression with Binary Outcome: NCSR
Logit Link
The SURVEYLOGISTIC Procedure

Model Information

Data Set	WORK.NCSR	
Response Variable	ald	
Number of Response Levels	2	
Stratum Variable	SESTRAT	SAMPLING ERROR STRATUM
Number of Strata	42	
Cluster Variable	SECLUSTR	SAMPLING ERROR CLUSTER
Number of Clusters	84	
Weight Variable	NCSRWTLG	NCSR sample part 2 weight
Model	Binary Logit	
Optimization Technique	Fisher's Scoring	
Variance Adjustment	Degrees of Freedom (DF)	

Variance Estimation

Method		Taylor Series
Variance Adjustment	Degrees of Freedom (DF)	
Number of Observations Read	9282	
Number of Observations Used	5692	
Sum of Weights Read	5692	
Sum of Weights Used	5692	

Response Profile

Ordered Value	ald	Total Frequency	Total Weight
1	0	5249	5384.2607
2	1	443	307.7398

Probability modeled is ald=1.

NOTE: 3590 observations having nonpositive frequencies or weights were excluded since they do not contribute to the analysis.

Class Level Information

Class	Value	Design Variables		
ag4cat	1=18-29	0	0	0
	2=30-44	1	0	0
	3=45-59	0	1	0
	4=60+	0	0	1
ED4CAT	1=0-11 Yrs	0	0	0
	2=12 Yrs	1	0	0
	3=11-13 Yrs	0	1	0
	4=16+ Yrs	0	0	1
MAR3CAT	1=Married	0	0	
	2=Previously Married	1	0	
	3=Never Married	0	1	

Analysis Example Comparison of Logistic, Probit and CLogLog Regression with Binary Outcome: NCSR
Logit Link

The SURVEYLOGISTIC Procedure

Class Level Information

Class	Value	Design Variables
SEX	1	1
	2	0

Model Convergence Status

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

Model Fit Statistics

Criterion	Intercept and	
	Intercept Only	Covariates
AIC	2396.234	2288.902
SC	2402.880	2355.371
-2 Log L	2394.234	2268.902

Testing Global Null Hypothesis: BETA=0

Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	125.3312	9	<.0001
Score	119.8419	9	<.0001
Wald	260.6371	9	<.0001

Type 3 Analysis of Effects

Effect	DF	Wald	
		Chi-Square	Pr > ChiSq
ag4cat	3	37.9225	<.0001
SEX	1	70.0985	<.0001
ED4CAT	3	15.0872	0.0017
MAR3CAT	2	13.3723	0.0012

Analysis of Maximum Likelihood Estimates

Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	-3.1243	0.2255	192.0471	<.0001
ag4cat 2=30-44	1	0.1463	0.1783	0.6733	0.4119
ag4cat 3=45-59	1	-0.0507	0.1440	0.1239	0.7248
ag4cat 4=60+	1	-1.1203	0.2127	27.7560	<.0001
SEX 1	1	0.9980	0.1192	70.0985	<.0001
ED4CAT 2=12 Yrs	1	-0.2684	0.1939	1.9170	0.1662
ED4CAT 3=11-13 Yrs	1	-0.2645	0.1763	2.2512	0.1335
ED4CAT 4=16+ Yrs	1	-0.7362	0.1973	13.9181	0.0002
MAR3CAT 2=Previously Married	1	0.5178	0.1422	13.2655	0.0003
MAR3CAT 3=Never Married	1	0.0653	0.1688	0.1496	0.6989

Odds Ratio Estimates

Effect	Point Estimate	95% Wald Confidence Limits	
ag4cat 2=30-44 vs 1=18-29	1.158	0.816	1.642
ag4cat 3=45-59 vs 1=18-29	0.951	0.717	1.261
ag4cat 4=60+ vs 1=18-29	0.326	0.215	0.495
SEX 1 vs 2	2.713	2.148	3.427
ED4CAT 2=12 Yrs vs 1=0-11 Yrs	0.765	0.523	1.118
ED4CAT 3=11-13 Yrs vs 1=0-11 Yrs	0.768	0.543	1.084
ED4CAT 4=16+ Yrs vs 1=0-11 Yrs	0.479	0.325	0.705
MAR3CAT 2=Previously Married vs 1=Married	1.678	1.270	2.218
MAR3CAT 3=Never Married vs 1=Married	1.067	0.767	1.486

Association of Predicted Probabilities and Observed Responses

Percent Concordant	65.0	Somers' D	0.334
Percent Discordant	31.5	Gamma	0.346
Percent Tied	3.5	Tau-a	0.048
Pairs	2325307	c	0.667

Analysis Example Comparison of Logistic, Probit and CLogLog Regression with Binary Outcome: NCSR
 Probit Link

```
proc surveylogistic data=ncsr ;
title2 "Probit Link" ;
strata sestrat ;
cluster seclustr ;
weight ncsrwtlg ;
class ag4cat (ref=first) ed4cat (ref=first) mar3cat (ref=first) sex (ref=last) / param=ref ;
model ald (event='1')= ag4cat sex ed4cat mar3cat / link=probit;
format ag4cat agf. sexm sexmf. ed4cat edf. mar3cat marf. ;
run ;
```

The SURVEYLOGISTIC Procedure

Model Information

Data Set	WORK.NCSR	
Response Variable	ald	
Number of Response Levels	2	
Stratum Variable	SESTRAT	SAMPLING ERROR STRATUM
Number of Strata	42	
Cluster Variable	SECLUSTR	SAMPLING ERROR CLUSTER
Number of Clusters	84	
Weight Variable	NCSRWTLG	NCSR sample part 2 weight
Model	Binary Probit	
Optimization Technique	Fisher's Scoring	
Variance Adjustment	Degrees of Freedom (DF)	

Variance Estimation

Method	Taylor Series
Variance Adjustment	Degrees of Freedom (DF)

Number of Observations Read	9282
Number of Observations Used	5692
Sum of Weights Read	5692
Sum of Weights Used	5692

Response Profile

Ordered Value	ald	Total Frequency	Total Weight
1	0	5249	5384.2607
2	1	443	307.7398

Probability modeled is ald=1.

NOTE: 3590 observations having nonpositive frequencies or weights were excluded since they do not contribute to the analysis.

Class Level Information

Class	Value	Design Variables		
ag4cat	1=18-29	0	0	0
	2=30-44	1	0	0
	3=45-59	0	1	0
	4=60+	0	0	1
ED4CAT	1=0-11 Yrs	0	0	0
	2=12 Yrs	1	0	0
	3=11-13 Yrs	0	1	0
	4=16+ Yrs	0	0	1
MAR3CAT	1=Married	0	0	
	2=Previously Married	1	0	
	3=Never Married	0	1	

Analysis Example Comparison of Logistic, Probit and CLogLog Regression with Binary Outcome: NCSR
 Probit Link

The SURVEYLOGISTIC Procedure

Class Level Information

Class	Value	Design Variables
SEX	1	1
	2	0

Model Convergence Status

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

Model Fit Statistics

Criterion	Intercept and Covariates	
	Intercept Only	
AIC	2396.234	2287.418
SC	2402.880	2353.886
-2 Log L	2394.234	2267.418

Testing Global Null Hypothesis: BETA=0

Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	126.8157	9	<.0001
Score	119.8419	9	<.0001
Wald	241.0011	9	<.0001

Type 3 Analysis of Effects

Effect	DF	Wald Chi-Square	Pr > ChiSq
ag4cat	3	47.9370	<.0001
SEX	1	68.2417	<.0001
ED4CAT	3	14.9915	0.0018
MAR3CAT	2	13.3815	0.0012

Analysis of Maximum Likelihood Estimates

Parameter	DF	Estimate	Standard Error	Wald Chi-Square	Pr > ChiSq
Intercept	1	-1.7194	0.1066	260.3616	<.0001
ag4cat 2=30-44	1	0.0653	0.0855	0.5840	0.4448
ag4cat 3=45-59	1	-0.0345	0.0674	0.2615	0.6091
ag4cat 4=60+	1	-0.5313	0.0938	32.0811	<.0001
SEX 1	1	0.4708	0.0570	68.2417	<.0001
ED4CAT 2=12 Yrs	1	-0.1238	0.0950	1.6962	0.1928
ED4CAT 3=11-13 Yrs	1	-0.1244	0.0851	2.1359	0.1439
ED4CAT 4=16+ Yrs	1	-0.3396	0.0924	13.4909	0.0002
MAR3CAT 2=Previously Married	1	0.2548	0.0702	13.1751	0.0003
MAR3CAT 3=Never Married	1	0.0389	0.0780	0.2493	0.6176

Association of Predicted Probabilities and Observed Responses

Percent Concordant	64.9	Somers' D	0.335
Percent Discordant	31.4	Gamma	0.348
Percent Tied	3.8	Tau-a	0.048
Pairs	2325307	c	0.667

Analysis Example Comparison of Logistic, Probit and CLogLog Regression with Binary Outcome: NCSR
 CLogLog Link

```
proc surveylogistic data=ncsr ;
title2 "CLogLog Link" ;
strata sestrat ;
cluster seclustr ;
weight ncsrwtlg ;
class ag4cat (ref=first) ed4cat (ref=first) mar3cat (ref=first) sex (ref=last) / param=ref ;
model ald (event='1')= ag4cat sex ed4cat mar3cat / link=cloglog ;
format ag4cat agf. sexm sexmf. ed4cat edf. mar3cat marf. ;
run ;
```

The SURVEYLOGISTIC Procedure

		Model Information	
Data Set	WORK.NCSR		
Response Variable	ald		
Number of Response Levels	2		
Stratum Variable	SESTRAT		SAMPLING ERROR STRATUM
Number of Strata	42		
Cluster Variable	SECLUSTR		SAMPLING ERROR CLUSTER
Number of Clusters	84		
Weight Variable	NCSRWTLG		NCSR sample part 2 weight
Model	Binary Cloglog		
Optimization Technique	Fisher's Scoring		
Variance Adjustment	Degrees of Freedom (DF)		

Variance Estimation

Method	Taylor Series
Variance Adjustment	Degrees of Freedom (DF)
Number of Observations Read	9282
Number of Observations Used	5692
Sum of Weights Read	5692
Sum of Weights Used	5692

Response Profile

Ordered Value	ald	Total Frequency	Total Weight
1	0	5249	5384.2607
2	1	443	307.7398

Probability modeled is ald=1.

NOTE: 3590 observations having nonpositive frequencies or weights were excluded since they do not contribute to the analysis.

Class Level Information

Class	Value	Design Variables		
ag4cat	1=18-29	0	0	0
	2=30-44	1	0	0
	3=45-59	0	1	0
	4=60+	0	0	1
ED4CAT	1=0-11 Yrs	0	0	0
	2=12 Yrs	1	0	0
	3=11-13 Yrs	0	1	0
	4=16+ Yrs	0	0	1
MAR3CAT	1=Married	0	0	
	2=Previously Married	1	0	
	3=Never Married	0	1	

Analysis Example Comparison of Logistic, Probit and CLogLog Regression with Binary Outcome: NCSR
CLogLog Link

The SURVEYLOGISTIC Procedure

Class Level Information

Class	Value	Design Variables
SEX	1	1
	2	0

Model Convergence Status

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

Model Fit Statistics

Criterion	Intercept	Intercept
	Only	and Covariates
AIC	2396.234	2289.159
SC	2402.880	2355.627
-2 Log L	2394.234	2269.159

Testing Global Null Hypothesis: BETA=0

Test	Chi-Square	DF	Pr > ChiSq
Likelihood Ratio	125.0747	9	<.0001
Score	119.8419	9	<.0001
Wald	267.9978	9	<.0001

Type 3 Analysis of Effects

Effect	DF	Wald	Pr > ChiSq
		Chi-Square	
ag4cat	3	36.5061	<.0001
SEX	1	70.4151	<.0001
ED4CAT	3	15.0315	0.0018
MAR3CAT	2	13.4004	0.0012

Analysis of Maximum Likelihood Estimates

Parameter	DF	Estimate	Standard	Wald	Pr > ChiSq
			Error	Chi-Square	
Intercept	1	-3.1484	0.2173	209.8627	<.0001
ag4cat 2=30-44	1	0.1430	0.1714	0.6966	0.4039
ag4cat 3=45-59	1	-0.0452	0.1398	0.1048	0.7462
ag4cat 4=60+	1	-1.0829	0.2083	27.0173	<.0001
SEX 1	1	0.9652	0.1150	70.4151	<.0001
ED4CAT 2=12 Yrs	1	-0.2601	0.1848	1.9806	0.1593
ED4CAT 3=11-13 Yrs	1	-0.2556	0.1686	2.2988	0.1295
ED4CAT 4=16+ Yrs	1	-0.7127	0.1906	13.9847	0.0002
MAR3CAT 2=Previously Married	1	0.4935	0.1353	13.3016	0.0003
MAR3CAT 3=Never Married	1	0.0605	0.1638	0.1364	0.7119

Association of Predicted Probabilities and Observed Responses

Percent Concordant	64.9	Somers' D	0.335
Percent Discordant	31.4	Gamma	0.348
Percent Tied	3.7	Tau-a	0.048
Pairs	2325307	c	0.668