

```

* SAS Analysis Examples Replication for ASDA 3rd Edition
* Berglund Winter 2025
* Chapter 9 ;
title ;
libname ncsr "P:\ASDA3\Data Sets for Analysis Examples and Stata R Code" ;
data c9_ncsr ;
  set ncsr.ncsr ;
run ;
proc contents ;
run ;

proc format ;
  value af 1='18-29' 2='30-44' 3='45-59' 4='60+' ;
  value sf 1='M' 2='F' ;
  value edf 1='0-11' 2='12' 3='13-15' 4='16+' ;
  value mf 1='Currently Married' 2='Previously Married' 3='Never Married' ;
  value yn 1='Yes' 0='No' ;
  value wkf 1='Employed' 2='Unemployed' 3='NLF' ;
run ;

ods rtf style=minimal bodytitle ;
* Example 9.2.6 ;
* Vertical Bar Chart ;
ods graphics on ;
title "Figure 9.2: Bar Chart of Marital Status, Weighted by NCSRWTLG" ;
proc freq data=c9_ncsr ;
  format wkstat3c wkf. ;
  tables wkstat3c / plots=freqplot (scale=percent) ;
  weight ncsrwtlg ;
run ;

* Check bivariate relationships using macro with surveylogistic (note: this could also be done using PROC SURVEYFREQ) ;
title "Bivariate Relationships of Work Status and Selected Predictors" ;
%macro biv (pred,ref) ;
proc surveylogistic data=c9_ncsr ;
  strata sestrat ; cluster seclustr ; weight ncsrwtlg ;
  class wkstat3c (ref=first) &pred (ref=&ref) / param=ref ;
  model wkstat3c = &pred / link=glogit ;
run ;
%mend biv ;
%biv(ag4cat,first) ;
%biv(sex,last) ;
%biv(ald,first) ;
%biv(mde,first) ;
%biv(ed4cat,first) ;
%biv(mar3cat,first) ;

title "Example 9.2.6 Multinomial logistic regression using NCSR data" ;
proc surveylogistic data =c9_ncsr ;
  strata sestrat ; cluster seclustr ; weight ncsrwtlg ;
  class wkstat3c (ref=first) sex (ref=last) ed4cat (ref=first) ag4cat (ref=first) mar3cat (ref=first) / param=ref ;
  model wkstat3c=sex ald mde ed4cat ag4cat mar3cat / link=glogit ;
  testwk2_3_education: test ed4cat2_2 - ed4cat2_3, ed4cat3_2-ed4cat3_3 , ed4cat4_2-ed4cat4_3 ; * refer to output est
data set to find correct names ;
run ;

ods text="Figure 9.3 and 9.4, Margins Plot: Not Available Directly in PROC SURVEYLOGISTIC, see
https://support.sas.com/kb/22/604.html for information on calculating marginal effects manually";
ods text="GOF test not available in SAS SURVEYLOGISTIC" ;
ods text="Bayesian Analysis not Available in SAS SURVEYLOGISTIC" ;

libname d 'P:\ASDA3\Data Sets for Analysis Examples and Stata R Code' ;

title " 9.3.6 Example: Fitting a Cumulative Logit Regression Model to Complex Sample Survey Data" ;
* prepare for analysis, data management ;
data c9_ess6_russia ;
  set d.ess6_russia ;
  if stflife =. then stflife2=. ;
  else if stflife =0 or stflife =1 then stflife2=1 ;
  else if stflife <=4 then stflife2=2 ;
  else if stflife <=5 then stflife2=3 ;
  else if stflife <=8 then stflife2=4 ;
  else if stflife <=10 then stflife2=5 ;
run ;

proc format ; value stf 1='0-1' 2='2-4' 3='5' 4='6-8' 5='9-10' ; run ;

title "Figure 9.6, Bar Chart of Satisfaction with Life, Weighted by PSPWGHT" ;
ods graphics on ;
proc freq data=c9_ess6_russia ;

```

```

format stflife2 stf. ;
tables stflife2 / plots=freqplot ;
weight pspwght ;
run ;
ods graphics off ;

title "Numbers for Table 9.6" ;
* note use of (desc) in model statement so direction of betas for predictors match Stata;
proc surveylogistic data=c9_ess6_russia ;
  strata stratify ; cluster psu ; weight pspwght ;
  class agecat (ref=first) marcat (ref=first) / param=ref ;
  model stflife2 (desc) = agecat marcat male ;
run ;

ods text="No Design-Adjusted GOF test for Ordinal Logistic Regression in SAS" ;
ods text="Bayesian Analysis not Available in SAS SURVEYLOGISTIC" ;
ods text="Section 9.4.7 Example: Fitting Poisson and Negative Binomial Regression Models to Complex Sample Survey
Data" ;
ods text="SURVEY PROCEDURES do not include POISSON, NEGATIVE BINOMIAL OR ZERO-INFLATED NB MODELS, see ASDA website for
user written SAS macros for these models. We do show how to prepare weighted plots below" ;

* Prepare weighted frequency plot for Number of Falls Past 24 Months ;
data c9_hrs ;
  set d.hrs12 ;
run ;

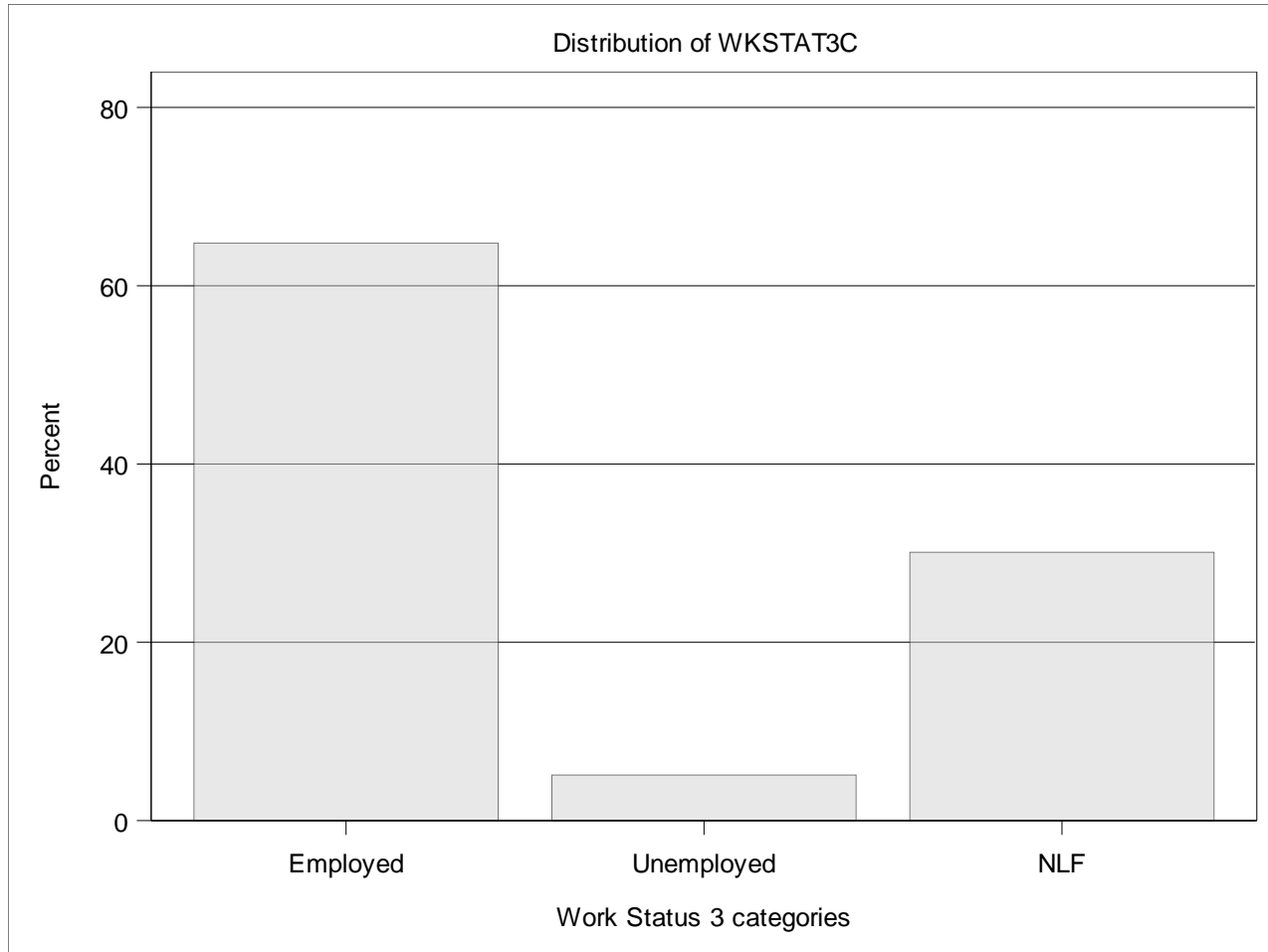
ods graphics on ;
*tables Hair*Region / plots=freqplot(type=dotplot scale=percent);
title "Weighted Frequency plot of Number of Falls Past 24 Months" ;
proc freq data=c9_hrs ;
  table numfalls24 / plots=freqplot (type=bar scale=percent) ;
  where age65p=1 ;
  weight nwgtr ;
  run ;
ods graphics off ;
ods rtf close ;

```

Figure 9.2: Bar Chart of Marital Status, Weighted by NCSRWTLG

The FREQ Procedure

Work Status 3 categories				
WKSTAT3C	Frequency	Percent	Cumulative Frequency	Cumulative Percent
Employed	3671.472	64.78	3671.472	64.78
Unemployed	289.8166	5.11	3961.289	69.90
NLF	1705.896	30.10	5667.185	100.00
Frequency Missing = 24.81548				



Bivariate Relationships of Work Status and Selected Predictors

The SURVEYLOGISTIC Procedure

Model Information		
Data Set	WORK.C9_NCSR	
Response Variable	WKSTAT3C	Work Status 3 categories
Number of Response Levels	3	
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	Generalized Logit	
Optimization Technique	Newton-Raphson	
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	5679
Sum of Weights Read	5692
Sum of Weights Used	5667.185

Response Profile			
Ordered Value	WKSTAT3C	Total Frequency	Total Weight
1	1	3766	3671.4725
2	2	283	289.8166
3	3	1630	1705.8959

Logits modeled use WKSTAT3C=1 as the reference category.

2649 observations were deleted due to missing values for the response or explanatory variables.

954 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	0	0	0
	2	1	0	0
	3	0	1	0
	4	0	0	1

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

Model Fit Statistics		
Criterion	Intercept Only	Intercept and Covariates
AIC	9011.140	7804.927
SC	9024.425	7858.067
-2 Log L	9007.140	7788.927

Testing Global Null Hypothesis: BETA=0				
Test	F Value	Num DF	Den DF	Pr > F
Likelihood Ratio	204.56	4.8088	201.97	<.0001
Score	47.97	6	37	<.0001
Wald	106.76	6	37	<.0001
NOTE: Second-order Rao-Scott design correction 0.2477 applied to the Likelihood Ratio test.				

Type 3 Analysis of Effects				
Effect	F Value	Num DF	Den DF	Pr > F
ag4cat	106.76	6	37	<.0001

Analysis of Maximum Likelihood Estimates						
Parameter		WKSTAT3C	Estimate	Standard Error	t Value	Pr >  t
Intercept		2	-3.1803	0.2121	-14.99	<.0001
Intercept		3	-0.9554	0.0865	-11.04	<.0001
ag4cat	2	2	-0.1362	0.2770	-0.49	0.6255
ag4cat	2	3	-0.6444	0.1020	-6.32	<.0001
ag4cat	3	2	-0.1243	0.2311	-0.54	0.5936
ag4cat	3	3	-0.3455	0.1342	-2.57	0.0136
ag4cat	4	2	2.6528	0.2483	10.68	<.0001
ag4cat	4	3	1.9902	0.1243	16.01	<.0001
NOTE: The degrees of freedom for the t tests is 42.						

Odds Ratio Estimates				
Effect		WKSTAT3C	Point Estimate	95% Confidence Limits
ag4cat 2 vs 1		2	0.873	0.499 1.526
ag4cat 2 vs 1		3	0.525	0.427 0.645
ag4cat 3 vs 1		2	0.883	0.554 1.408
ag4cat 3 vs 1		3	0.708	0.540 0.928
ag4cat 4 vs 1		2	14.193	8.598 23.428
ag4cat 4 vs 1		3	7.317	5.693 9.403

Odds Ratio Estimates			
Effect	WKSTAT3C	Point Estimate	95% Confidence Limits
NOTE: The degrees of freedom in computing the confidence limits is 42.			

Bivariate Relationships of Work Status and Selected Predictors

The SURVEYLOGISTIC Procedure

Model Information		
Data Set	WORK.C9_NCSR	
Response Variable	WKSTAT3C	Work Status 3 categories
Number of Response Levels	3	
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	Generalized Logit	
Optimization Technique	Newton-Raphson	
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	5679
Sum of Weights Read	5692
Sum of Weights Used	5667.185

Response Profile			
Ordered Value	WKSTAT3C	Total Frequency	Total Weight
1	1	3766	3671.4725
2	2	283	289.8166
3	3	1630	1705.8959

Logits modeled use WKSTAT3C=1 as the reference category.

2649 observations were deleted due to missing values for the response or explanatory variables.

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

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 Only	Intercept and Covariates
AIC	9011.140	8878.892
SC	9024.425	8905.462
-2 Log L	9007.140	8870.892

Testing Global Null Hypothesis: BETA=0				
Test	F Value	Num DF	Den DF	Pr > F
Likelihood Ratio	58.01	1.6457	69.1184	<.0001
Score	28.21	2	41	<.0001
Wald	21.22	2	41	<.0001
NOTE: Second-order Rao-Scott design correction 0.2153 applied to the Likelihood Ratio test.				

Type 3 Analysis of Effects				
Effect	F Value	Num DF	Den DF	Pr > F
sex	21.22	2	41	<.0001

Analysis of Maximum Likelihood Estimates						
Parameter		WKSTAT3C	Estimate	Standard Error	t Value	Pr >  t
Intercept		2	-2.0950	0.1106	-18.94	<.0001
Intercept		3	-0.5291	0.0677	-7.81	<.0001
sex	1	2	-1.1591	0.2060	-5.63	<.0001
sex	1	3	-0.5183	0.0998	-5.19	<.0001
NOTE: The degrees of freedom for the t tests is 42.						

Odds Ratio Estimates				
Effect	WKSTAT3C	Point Estimate	95% Confidence Limits	
sex 1 vs 2	2	0.314	0.207	0.475
sex 1 vs 2	3	0.596	0.487	0.728
NOTE: The degrees of freedom in computing the confidence limits is 42.				

Bivariate Relationships of Work Status and Selected Predictors

The SURVEYLOGISTIC Procedure

Model Information		
Data Set	WORK.C9_NCSR	
Response Variable	WKSTAT3C	Work Status 3 categories
Number of Response Levels	3	
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	Generalized Logit	
Optimization Technique	Newton-Raphson	
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	5679
Sum of Weights Read	5692
Sum of Weights Used	5667.185

Response Profile			
Ordered Value	WKSTAT3C	Total Frequency	Total Weight
1	1	3766	3671.4725
2	2	283	289.8166
3	3	1630	1705.8959

Logits modeled use WKSTAT3C=1 as the reference category.

2649 observations were deleted due to missing values for the response or explanatory variables.

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

Class Level Information		
Class	Value	Design Variables
ald	0	0
	1	1

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

Model Fit Statistics		
Criterion	Intercept Only	Intercept and Covariates
AIC	9011.140	9008.606
SC	9024.425	9035.176
-2 Log L	9007.140	9000.606

Testing Global Null Hypothesis: BETA=0				
Test	F Value	Num DF	Den DF	Pr > F
Likelihood Ratio	3.76	1.5260	64.0913	0.0394
Score	8.89	2	41	0.0006
Wald	3.23	2	41	0.0496
NOTE: Second-order Rao-Scott design correction 0.3106 applied to the Likelihood Ratio test.				

Type 3 Analysis of Effects				
Effect	F Value	Num DF	Den DF	Pr > F
ald	3.23	2	41	0.0496

Analysis of Maximum Likelihood Estimates						
Parameter		WKSTAT3C	Estimate	Standard Error	t Value	Pr >  t
Intercept		2	-2.5058	0.1087	-23.04	<.0001
Intercept		3	-0.7640	0.0459	-16.66	<.0001
ald	1	2	-0.8629	0.3578	-2.41	0.0203
ald	1	3	-0.0459	0.1230	-0.37	0.7112
NOTE: The degrees of freedom for the t tests is 42.						

Odds Ratio Estimates				
Effect	WKSTAT3C	Point Estimate	95% Confidence Limits	
ald 1 vs 0	2	0.422	0.205	0.869
ald 1 vs 0	3	0.955	0.745	1.224
NOTE: The degrees of freedom in computing the confidence limits is 42.				

Bivariate Relationships of Work Status and Selected Predictors

The SURVEYLOGISTIC Procedure

Model Information		
Data Set	WORK.C9_NCSR	
Response Variable	WKSTAT3C	Work Status 3 categories
Number of Response Levels	3	
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	Generalized Logit	
Optimization Technique	Newton-Raphson	
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	5679
Sum of Weights Read	5692
Sum of Weights Used	5667.185

Response Profile			
Ordered Value	WKSTAT3C	Total Frequency	Total Weight
1	1	3766	3671.4725
2	2	283	289.8166
3	3	1630	1705.8959

Logits modeled use WKSTAT3C=1 as the reference category.

2649 observations were deleted due to missing values for the response or explanatory variables.

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

Class Level Information		
Class	Value	Design Variables
mde	0	0
	1	1

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

Model Fit Statistics		
Criterion	Intercept Only	Intercept and Covariates
AIC	9011.140	9007.284
SC	9024.425	9033.854
-2 Log L	9007.140	8999.284

Testing Global Null Hypothesis: BETA=0				
Test	F Value	Num DF	Den DF	Pr > F
Likelihood Ratio	4.96	1.6203	68.0532	0.0147
Score	7.17	2	41	0.0021
Wald	6.72	2	41	0.0030
NOTE: Second-order Rao-Scott design correction 0.2343 applied to the Likelihood Ratio test.				

Type 3 Analysis of Effects				
Effect	F Value	Num DF	Den DF	Pr > F
mde	6.72	2	41	0.0030

Analysis of Maximum Likelihood Estimates						
Parameter		WKSTAT3C	Estimate	Standard Error	t Value	Pr >  t
Intercept		2	-2.4692	0.1125	-21.94	<.0001
Intercept		3	-0.7426	0.0527	-14.10	<.0001
mde	1	2	-0.4082	0.1217	-3.35	0.0017
mde	1	3	-0.1246	0.0782	-1.59	0.1188
NOTE: The degrees of freedom for the t tests is 42.						

Odds Ratio Estimates				
Effect	WKSTAT3C	Point Estimate	95% Confidence Limits	
mde 1 vs 0	2	0.665	0.520	0.850
mde 1 vs 0	3	0.883	0.754	1.034
NOTE: The degrees of freedom in computing the confidence limits is 42.				

Bivariate Relationships of Work Status and Selected Predictors

The SURVEYLOGISTIC Procedure

Model Information		
Data Set	WORK.C9_NCSR	
Response Variable	WKSTAT3C	Work Status 3 categories
Number of Response Levels	3	
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	Generalized Logit	
Optimization Technique	Newton-Raphson	
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	5679
Sum of Weights Read	5692
Sum of Weights Used	5667.185

Response Profile			
Ordered Value	WKSTAT3C	Total Frequency	Total Weight
1	1	3766	3671.4725
2	2	283	289.8166
3	3	1630	1705.8959

Logits modeled use WKSTAT3C=1 as the reference category.

2649 observations were deleted due to missing values for the response or explanatory variables.

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

Class Level Information				
Class	Value	Design Variables		
ED4CAT	1	0	0	0
	2	1	0	0
	3	0	1	0
	4	0	0	1

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

Model Fit Statistics		
Criterion	Intercept Only	Intercept and Covariates
AIC	9011.140	8698.885
SC	9024.425	8752.024
-2 Log L	9007.140	8682.885

Testing Global Null Hypothesis: BETA=0				
Test	F Value	Num DF	Den DF	Pr > F
Likelihood Ratio	51.12	5.0579	212.43	<.0001
Score	23.71	6	37	<.0001
Wald	29.98	6	37	<.0001
NOTE: Second-order Rao-Scott design correction 0.1863 applied to the Likelihood Ratio test.				

Type 3 Analysis of Effects				
Effect	F Value	Num DF	Den DF	Pr > F
ED4CAT	29.98	6	37	<.0001

Analysis of Maximum Likelihood Estimates						
Parameter		WKSTAT3C	Estimate	Standard Error	t Value	Pr >  t
Intercept		2	-1.5019	0.2195	-6.84	<.0001
Intercept		3	0.1078	0.0960	1.12	0.2675
ED4CAT	2	2	-0.8458	0.2229	-3.79	0.0005
ED4CAT	2	3	-0.7553	0.1284	-5.88	<.0001
ED4CAT	3	2	-1.4620	0.2456	-5.95	<.0001
ED4CAT	3	3	-1.1379	0.1302	-8.74	<.0001
ED4CAT	4	2	-1.8935	0.2701	-7.01	<.0001
ED4CAT	4	3	-1.4515	0.1508	-9.62	<.0001
NOTE: The degrees of freedom for the t tests is 42.						

Odds Ratio Estimates					
Effect		WKSTAT3C	Point Estimate	95% Confidence Limits	
ED4CAT 2 vs 1		2	0.429	0.274	0.673
ED4CAT 2 vs 1		3	0.470	0.363	0.609
ED4CAT 3 vs 1		2	0.232	0.141	0.380
ED4CAT 3 vs 1		3	0.320	0.246	0.417
ED4CAT 4 vs 1		2	0.151	0.087	0.260
ED4CAT 4 vs 1		3	0.234	0.173	0.318

Odds Ratio Estimates			
Effect	WKSTAT3C	Point Estimate	95% Confidence Limits
NOTE: The degrees of freedom in computing the confidence limits is 42.			

Bivariate Relationships of Work Status and Selected Predictors

The SURVEYLOGISTIC Procedure

Model Information		
Data Set	WORK.C9_NCSR	
Response Variable	WKSTAT3C	Work Status 3 categories
Number of Response Levels	3	
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	Generalized Logit	
Optimization Technique	Newton-Raphson	
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	5679
Sum of Weights Read	5692
Sum of Weights Used	5667.185

Response Profile			
Ordered Value	WKSTAT3C	Total Frequency	Total Weight
1	1	3766	3671.4725
2	2	283	289.8166
3	3	1630	1705.8959

Logits modeled use WKSTAT3C=1 as the reference category.

2649 observations were deleted due to missing values for the response or explanatory variables.

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

Class Level Information			
Class	Value	Design Variables	
MAR3CAT	1	0	0
	2	1	0
	3	0	1

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

Model Fit Statistics		
Criterion	Intercept Only	Intercept and Covariates
AIC	9011.140	8833.583
SC	9024.425	8873.437
-2 Log L	9007.140	8821.583

Testing Global Null Hypothesis: BETA=0				
Test	F Value	Num DF	Den DF	Pr > F
Likelihood Ratio	51.19	3.3356	140.10	<.0001
Score	43.38	4	39	<.0001
Wald	36.72	4	39	<.0001
NOTE: Second-order Rao-Scott design correction 0.1992 applied to the Likelihood Ratio test.				

Type 3 Analysis of Effects				
Effect	F Value	Num DF	Den DF	Pr > F
MAR3CAT	36.72	4	39	<.0001

Analysis of Maximum Likelihood Estimates						
Parameter		WKSTAT3C	Estimate	Standard Error	t Value	Pr >  t
Intercept		2	-2.3396	0.1274	-18.36	<.0001
Intercept		3	-0.9195	0.0561	-16.38	<.0001
MAR3CAT	2	2	0.2656	0.1905	1.39	0.1706
MAR3CAT	2	3	0.6045	0.0997	6.06	<.0001
MAR3CAT	3	2	-2.8180	0.3434	-8.21	<.0001
MAR3CAT	3	3	0.0843	0.1069	0.79	0.4345
NOTE: The degrees of freedom for the t tests is 42.						

Odds Ratio Estimates					
Effect		WKSTAT3C	Point Estimate	95% Confidence Limits	
MAR3CAT 2 vs 1		2	1.304	0.888	1.916
MAR3CAT 2 vs 1		3	1.830	1.497	2.238
MAR3CAT 3 vs 1		2	0.060	0.030	0.119
MAR3CAT 3 vs 1		3	1.088	0.877	1.350
NOTE: The degrees of freedom in computing the confidence limits is 42.					

Example 9.2.6 Multinomial logistic regression using NCSR data

The SURVEYLOGISTIC Procedure

Model Information		
Data Set	WORK.C9_NCSR	
Response Variable	WKSTAT3C	Work Status 3 categories
Number of Response Levels	3	
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	Generalized Logit	
Optimization Technique	Newton-Raphson	
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	5679
Sum of Weights Read	5692
Sum of Weights Used	5667.185

Response Profile			
Ordered Value	WKSTAT3C	Total Frequency	Total Weight
1	1	3766	3671.4725
2	2	283	289.8166
3	3	1630	1705.8959

Logits modeled use WKSTAT3C=1 as the reference category.

2649 observations were deleted due to missing values for the response or explanatory variables.

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

Class Level Information				
Class	Value	Design Variables		
sex	1	1		
	2	0		
ED4CAT	1	0	0	0
	2	1	0	0
	3	0	1	0
	4	0	0	1
ag4cat	1	0	0	0

Class Level Information				
Class	Value	Design Variables		
	2	1	0	0
	3	0	1	0
	4	0	0	1
MAR3CAT	1	0	0	
	2	1	0	
	3	0	1	

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

Model Fit Statistics		
Criterion	Intercept Only	Intercept and Covariates
AIC	9011.140	7399.903
SC	9024.425	7559.322
-2 Log L	9007.140	7351.903

Testing Global Null Hypothesis: BETA=0				
Test	F Value	Num DF	Den DF	Pr > F
Likelihood Ratio	78.38	12.6078	529.53	<.0001
Score	29.89	22	21	<.0001
Wald	73.62	22	21	<.0001

NOTE: Second-order Rao-Scott design correction 0.7450 applied to the Likelihood Ratio test.

Type 3 Analysis of Effects				
Effect	F Value	Num DF	Den DF	Pr > F
sex	35.61	2	41	<.0001
ald	5.03	2	41	0.0111
mde	1.13	2	41	0.3316
ED4CAT	13.63	6	37	<.0001
ag4cat	83.25	6	37	<.0001
MAR3CAT	24.72	4	39	<.0001

Analysis of Maximum Likelihood Estimates						
Parameter		WKSTAT3C	Estimate	Standard Error	t Value	Pr >  t
Intercept		2	-0.6438	0.2967	-2.17	0.0357
Intercept		3	-0.3795	0.1733	-2.19	0.0341
sex	1	2	-1.3932	0.1980	-7.04	<.0001
sex	1	3	-0.6403	0.1103	-5.81	<.0001

Analysis of Maximum Likelihood Estimates						
Parameter		WKSTAT3C	Estimate	Standard Error	t Value	Pr >  t
ald		2	-0.1638	0.3577	-0.46	0.6494
ald		3	0.3332	0.1305	2.55	0.0144
mde		2	-0.1398	0.1576	-0.89	0.3801
mde		3	0.0985	0.0882	1.12	0.2701
ED4CAT	2	2	-0.8470	0.2359	-3.59	0.0009
ED4CAT	2	3	-0.6514	0.1413	-4.61	<.0001
ED4CAT	3	2	-1.3653	0.2580	-5.29	<.0001
ED4CAT	3	3	-0.9169	0.1468	-6.25	<.0001
ED4CAT	4	2	-1.7310	0.3111	-5.56	<.0001
ED4CAT	4	3	-1.2295	0.1599	-7.69	<.0001
ag4cat	2	2	-0.8524	0.2951	-2.89	0.0061
ag4cat	2	3	-0.3164	0.1290	-2.45	0.0184
ag4cat	3	2	-0.8377	0.2586	-3.24	0.0023
ag4cat	3	3	0.0650	0.1712	0.38	0.7062
ag4cat	4	2	1.8284	0.2953	6.19	<.0001
ag4cat	4	3	2.3806	0.1738	13.70	<.0001
MAR3CAT	2	2	-0.5899	0.2257	-2.61	0.0124
MAR3CAT	2	3	-0.0523	0.1052	-0.50	0.6220
MAR3CAT	3	2	-2.7834	0.3807	-7.31	<.0001
MAR3CAT	3	3	0.5528	0.1326	4.17	0.0002
NOTE: The degrees of freedom for the t tests is 42.						

Odds Ratio Estimates					
Effect		WKSTAT3C	Point Estimate	95% Confidence Limits	
sex	1 vs 2	2	0.248	0.166	0.370
sex	1 vs 2	3	0.527	0.422	0.659
ald		2	0.849	0.412	1.747
ald		3	1.395	1.072	1.816
mde		2	0.870	0.633	1.195
mde		3	1.104	0.924	1.318
ED4CAT	2 vs 1	2	0.429	0.266	0.690
ED4CAT	2 vs 1	3	0.521	0.392	0.693
ED4CAT	3 vs 1	2	0.255	0.152	0.430
ED4CAT	3 vs 1	3	0.400	0.297	0.538
ED4CAT	4 vs 1	2	0.177	0.095	0.332
ED4CAT	4 vs 1	3	0.292	0.212	0.404
ag4cat	2 vs 1	2	0.426	0.235	0.774
ag4cat	2 vs 1	3	0.729	0.562	0.946
ag4cat	3 vs 1	2	0.433	0.257	0.729
ag4cat	3 vs 1	3	1.067	0.755	1.508

Odds Ratio Estimates					
Effect	WKSTAT3C	Point Estimate	95% Confidence Limits		
ag4cat 4 vs 1	2	6.224	3.430	11.294	
ag4cat 4 vs 1	3	10.811	7.614	15.352	
MAR3CAT 2 vs 1	2	0.554	0.352	0.874	
MAR3CAT 2 vs 1	3	0.949	0.768	1.174	
MAR3CAT 3 vs 1	2	0.062	0.029	0.133	
MAR3CAT 3 vs 1	3	1.738	1.330	2.272	
NOTE: The degrees of freedom in computing the confidence limits is 42.					

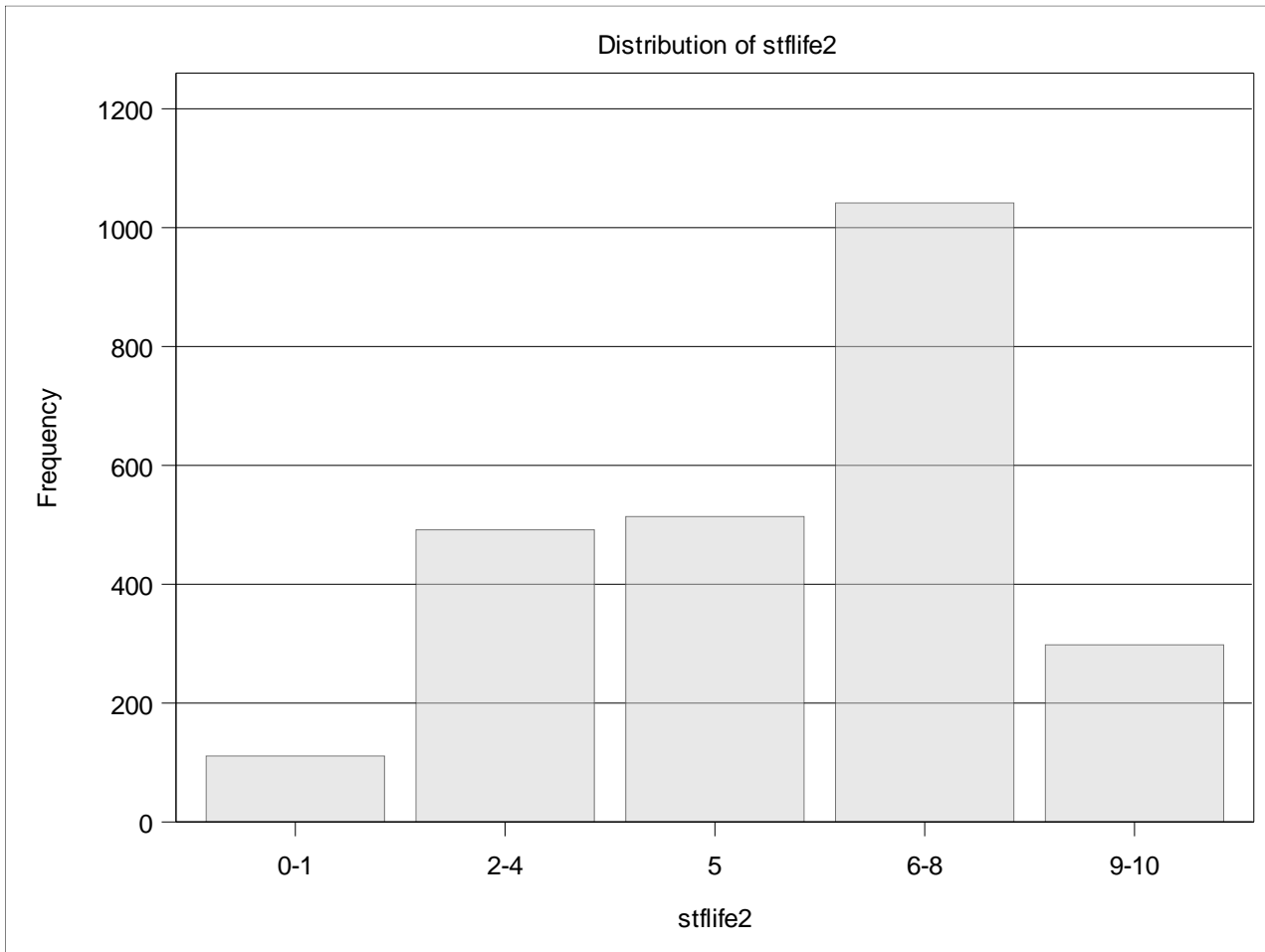
Linear Hypotheses Testing Results				
Label	F Value	Num DF	Den DF	Pr > F
testwk2_3_education	1.25	3	40	0.3047

Figure 9.3 and 9.4, Margins Plot: Not Available Directly in PROC SURVEYLOGISTIC, see <https://support.sas.com/kb/22/604.html> for information on calculating marginal effects manually  
 GOF test not available in SAS SURVEYLOGISTIC  
 Bayesian Analysis not Available in SAS SURVEYLOGISTIC

Figure 9.6, Bar Chart of Satisfaction with Life, Weighted by PSPWGHT

The FREQ Procedure

stflife2	Frequency	Percent	Cumulative Frequency	Cumulative Percent
0-1	111.1011	4.52	111.1011	4.52
2-4	491.5778	20.01	602.6789	24.54
5	513.8991	20.92	1116.578	45.46
6-8	1041.598	42.41	2158.176	87.86
9-10	298.0937	12.14	2456.27	100.00
Frequency Missing = 27.729974041				



Numbers for Table 9.6

The SURVEYLOGISTIC Procedure

Model Information		
Data Set	WORK.C9_ESS6_RUSSIA	
Response Variable	stflife2	
Number of Response Levels	5	
Stratum Variable	stratify	Stratification
Number of Strata	8	
Cluster Variable	psu	Primary Sampling Unit
Number of Clusters	184	
Weight Variable	pspwght	Post-stratification weight including design weight
Model	Cumulative 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	2484
Number of Observations Used	2415
Sum of Weights Read	2484
Sum of Weights Used	2422.138

Response Profile			
Ordered Value	stflife2	Total Frequency	Total Weight
1	5	283	294.8272
2	4	1036	1035.3251
3	3	492	502.8743
4	2	497	478.0106
5	1	107	111.1011

Probabilities modeled are cumulated over the lower Ordered Values.

69 observations were deleted due to missing values for the response or explanatory variables.

Class Level Information				
Class	Value	Design Variables		
agecat	1	0	0	0
	2	1	0	0
	3	0	1	0
	4	0	0	1
marcat	1	0	0	
	2	1	0	

Class Level Information				
Class	Value	Design Variables		
	3	0	1	

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

Model Fit Statistics		
Criterion	Intercept Only	Intercept and Covariates
AIC	6827.069	6766.264
SC	6850.238	6824.188
-2 Log L	6819.069	6746.264

Testing Global Null Hypothesis: BETA=0				
Test	F Value	Num DF	Den DF	Pr > F
Likelihood Ratio	11.22	5.6214	989.37	<.0001
Score	8.33	6	171	<.0001
Wald	7.60	6	171	<.0001

NOTE: Second-order Rao-Scott design correction 0.0673 applied to the Likelihood Ratio test.

Type 3 Analysis of Effects				
Effect	F Value	Num DF	Den DF	Pr > F
agecat	9.94	3	174	<.0001
marcat	2.18	2	175	0.1166
male	1.33	1	176	0.2509

Analysis of Maximum Likelihood Estimates					
Parameter		Estimate	Standard Error	t Value	Pr >  t
Intercept	5	-1.3841	0.1575	-8.79	<.0001
Intercept	4	0.8348	0.1643	5.08	<.0001
Intercept	3	1.7928	0.1713	10.47	<.0001
Intercept	2	3.7111	0.2176	17.05	<.0001
agecat	2	-0.5293	0.1392	-3.80	0.0002
agecat	3	-0.7455	0.1463	-5.10	<.0001
agecat	4	-0.8081	0.1661	-4.86	<.0001
marcat	2	-0.2089	0.1049	-1.99	0.0480
marcat	3	-0.1372	0.1349	-1.02	0.3105
male		-0.1096	0.0952	-1.15	0.2509

NOTE: The degrees of freedom for the t tests is 176.

Odds Ratio Estimates			
Effect	Point Estimate	95% Confidence Limits	
		agecat 2 vs 1	0.589
agecat 3 vs 1	0.474	0.356	0.633
agecat 4 vs 1	0.446	0.321	0.619
marcat 2 vs 1	0.811	0.660	0.998
marcat 3 vs 1	0.872	0.668	1.138
male	0.896	0.743	1.081
NOTE: The degrees of freedom in computing the confidence limits is 176.			

Association of Predicted Probabilities and Observed Responses			
Percent Concordant	54.1	Somers' D	0.149
Percent Discordant	39.2	Gamma	0.159
Percent Tied	6.7	Tau-a	0.106
Pairs	2089159	c	0.574

No Design-Adjusted GOF test for Ordinal Logistic Regression in SAS  
 Bayesian Analysis not Available in SAS SURVEYLOGISTIC  
 Section 9.4.7 Example: Fitting Poisson and Negative Binomial Regression Models to Complex Sample Survey Data  
 SURVEY PROCEDURES do not include POISSON, NEGATIVE BINOMIAL OR ZERO-INFLATED NB MODELS, see ASDA website for user  
 written SAS macros for these models. We do show how to prepare weighted plots below

Weighted Frequency plot of Number of Falls Past 24 Months

The FREQ Procedure

Number of Falls Past 2 Years				
numfalls24	Frequency	Percent	Cumulative Frequency	Cumulative Percent
0	27268099	65.22	27268099	65.22
1	5789028	13.85	33057127	79.06
2	3446787	8.24	36503914	87.31
3	2048115	4.90	38552029	92.20
4	990355	2.37	39542384	94.57
5	674539	1.61	40216923	96.19
6	471722	1.13	40688645	97.31
7	67152	0.16	40755797	97.48
8	149761	0.36	40905558	97.83
9	16081	0.04	40921639	97.87
10	248432	0.59	41170071	98.47
11	4260	0.01	41174331	98.48
12	235131	0.56	41409462	99.04
13	5095	0.01	41414557	99.05
15	71103	0.17	41485660	99.22
18	7744	0.02	41493404	99.24
19	3164	0.01	41496568	99.25
20	135477	0.32	41632045	99.57
24	5324	0.01	41637369	99.58
25	31386	0.08	41668755	99.66
30	63476	0.15	41732231	99.81
35	2052	0.00	41734283	99.82
40	16918	0.04	41751201	99.86
45	2763	0.01	41753964	99.86
50	57493	0.14	41811457	100.00
Frequency Missing = 544127				

Distribution of numfalls24

