```
*******************************************************;
*// MAT8414- CATEGORICAL DATA ANALYSIS                 *;
*// PROJECT CODE for Reference                         *;
*// By ADHIRAJ ROKA                                    *;
*******************************************************;

FILENAME REFFILE '/folders/myshortcuts/myfolder/CMC.csv';

PROC IMPORT DATAFILE=REFFILE replace
    DBMS=CSV
    OUT=cmc;
    GETNAMES=YES;
RUN;
proc print data=cmc;

*Initial;
*Test ordinal or nominal for education and index;
*Education;
proc logistic data= cmc;
class contraceptive Education/ param=ref;
model contraceptive(order=data)=Education/ link=glogit;
run;
proc logistic data= cmc;
class contraceptive/ param=ref;
model contraceptive(order=data)=Education/ link=glogit;
run;
*Index;
proc logistic data= cmc;
class contraceptive index/ param=ref;
model contraceptive(order=data)=index/ link=glogit;
run;
proc logistic data= cmc;
class contraceptive/ param=ref;
model contraceptive(order=data)=index/ link=glogit;
run;
*Treat both these variables ordinally;
/*****************************************************************************/
*Now test for each predictor individually to see significant or not;
/*****************************************************************************/
proc logistic data= cmc;
class contraceptive/ param=ref;
model contraceptive(order=data)=Age/ link=glogit;
run;
proc logistic data= cmc;
class contraceptive/ param=ref;
model contraceptive(order=data)=Education/ link=glogit;
run;
proc logistic data= cmc;
class contraceptive / param=ref;
model contraceptive(order=data)=Children/ link=glogit;
run;
proc logistic data= cmc;
class contraceptive Religion/ param=ref;
model contraceptive(order=data)=Religion/ link=glogit;
run;
proc logistic data= cmc;
class contraceptive Employed/ param=ref;
model contraceptive(order=data)=Employed/ link=glogit;
run;
proc logistic data= cmc;
class contraceptive / param=ref;
model contraceptive(order=data)=index/link=glogit;
run;
proc logistic data= cmc;
class contraceptive Media_Exposure/ param=ref;
model contraceptive(order=data)=Media_Exposure/ link=glogit;
run;
/*****************************************************************************/
*Remove none, all have p-values<0.2 ;
/*****************************************************************************/
/* *start by looking for odinal or not; */
/* proc logistic data= cmc ; */
/* class Religion Education Employed Index Media_Exposure Contraceptive/ param=ref; */
/* model Contraceptive(order=data)= Age Education Religion Employed Index Media_Exposure/ */
```

```
/* link=glogit; */
/* run; */
/* *index=2911.823-2910.308 =1.515 df=4 ordinal;  */
/* proc logistic data= cmc ; */
/* class Religion Education Employed Media_Exposure Contraceptive/ param=ref; */
/* model Contraceptive(order=data)= Age Education Religion Employed Index Media_Exposure/ */
/* link=glogit; */
/* run; */
/* *education=2910.567-2910.308=0.259 df=4 ordinal;  */
/* proc logistic data= cmc ; */
/* class Religion Employed Media_Exposure Contraceptive/ param=ref; */
/* model Contraceptive(order=data)= Age Education Religion Employed Index Media_Exposure/ */
/* link=glogit; */
/* run; */
/*****************************************************************************/
*Model selection for baseline-category logit;
/*****************************************************************************/
*forward selction;
proc logistic data= cmc ;
class Religion Employed Media_Exposure Contraceptive/ param=ref;
model Contraceptive(order=data)= Age Education Children Religion Employed Index Media_Exposure/
        link=glogit selection=stepwise slentry=0.1 slstay=0.1 details;
run;

*Employed,Media_exposure removed. p-value>0.1;
/*****************************************************************************/
*fit baseline model with significant predictors;
proc logistic data= cmc ;
class Religion Contraceptive / param=ref;
model Contraceptive(order=data)= Age Education Children Religion Index/link=glogit
    aggregate=(Age Education Children Religion Index) scale=none lackfit;
run;

/*****************************************************************************/
/* Consider adding in any variables that were not included in the model after  */
/* Step 1 or Step 2. A predictor can be added in even if p-value>0.1 if the least coefficient change by atleast 10%*/
/*****************************************************************************/
proc logistic data= cmc ;
class Religion Contraceptive Employed / param=ref;
model Contraceptive(order=data)= Age Education Children Religion Index Employed/link=glogit
    aggregate=(Age Education Children Religion Index) scale=none;
run;
proc logistic data= cmc ;
class Religion Contraceptive Media_Exposure/ param=ref;
model Contraceptive(order=data)= Age Education Children Religion Index Media_Exposure/link=glogit
    aggregate=(Age Education Children Religion Index) scale=none;
run;
/*****************************************************************************/
*NONE. proceed with following model- check for interactions;
/*****************************************************************************/
*model:;
proc logistic data= cmc ;
class / param=ref;
model Contraceptive(order=data)= Age Education Children Religion Index
     /link=glogit
    aggregate=(Age Education Children Religion Index) scale=none ;
run;
****************;
*test for interaction;
proc logistic data= cmc ;
class / param=ref;
model Contraceptive(order=data)= Age Education Children Religion Index
     Age*Age
     /link=glogit
    aggregate=(Age Education Children Religion Index) scale=none ;
run;

proc logistic data= cmc ;
class Religion Contraceptive/ param=ref;
model Contraceptive(order=data)= Age Education Children Religion Index
     Age*Education
     /link=glogit
    aggregate=(Age Education Children Religion Index) scale=none  ;
run;

proc logistic data= cmc ;
```

```
class Religion Contraceptive/ param=ref;
model Contraceptive(order=data)= Age Education Children Religion Index
     Age*Children
     /link=glogit
    aggregate=(Age Education Children Religion Index) scale=none ;
run;

proc logistic data= cmc ;
class Religion Contraceptive/ param=ref;
model Contraceptive(order=data)= Age Education Children Religion Index
     Age*Index
     /link=glogit
    aggregate=(Age Education Children Religion Index) scale=none ;
run;

proc logistic data= cmc ;
class Religion Contraceptive/ param=ref;
model Contraceptive(order=data)= Age Education Children Religion Index
     Education*Children
     /link=glogit
    aggregate=(Age Education Children Religion Index) scale=none ;
run;

proc logistic data= cmc ;
class Religion Contraceptive/ param=ref;
model Contraceptive(order=data)= Age Education Children Religion Index
     Children*Children
     /link=glogit
    aggregate=(Age Education Children Religion Index) scale=none ;
run;

proc logistic data= cmc ;
class Religion Contraceptive/ param=ref;
model Contraceptive(order=data)= Age Education Children Religion Index
     Children*Religion
     /link=glogit
    aggregate=(Age Education Children Religion Index) scale=none ;
run;

proc logistic data= cmc ;
class Religion Contraceptive/ param=ref;
model Contraceptive(order=data)= Age Education Children Religion Index
     Children*Index
     /link=glogit
    aggregate=(Age Education Children Religion Index) scale=none ;
run;

/*****************************************************************************/
*age*age,age*education,age*children, age*index,education*children
children*children,children*religon, children*index significant;
*every interaction above significant. those not listed not significant: 8 interactions;
/*****************************************************************************/
*****************************************************************************;

*original baseline model;
proc logistic data= cmc ;
class Religion Contraceptive / param=ref;
model Contraceptive(order=data)= Age Education Children Religion Index/link=glogit
    aggregate=(Age Education Children Religion Index) scale=none lackfit rsquare;
run;

*fit model with interactions and compare, interactions better but lof;
proc logistic data= cmc;
class Religion Contraceptive/ param=ref;
model Contraceptive(order=data)= Age Education Children Religion Index
     Age*Age Age*Education Age*Children Age*Index Education*Children
     Children*Children Children*Religion Children*Index
     /link=glogit aggregate=(Age Education Children Religion Index) scale=none lackfit rsquare;
run;

*****************************************************************************;
*****************************************************************************;

*stepwise again;
proc logistic data= cmc;
class Religion Contraceptive/ param=ref;
```

```
model Contraceptive(order=data)= Age Education Children Religion Index
     Age*Age Age*Education Age*Children Age*Index Education*Children
     Children*Children Children*Religion Children*Index
     /link=glogit selection=stepwise slentry=0.1 slstay=0.1 details;
run;

*we get:;
proc logistic data= cmc;
class Contraceptive/ param=ref;
model Contraceptive(order=data)= Age Education Children Index
 Age*Age Age*Children Age*Index Children*Children /link=glogit
     aggregate=(Age Education Children Index) scale=none lackfit rsquare;
run;
*****************************************************************************;*******************************
*just check if educ. index can still be treated ordinally;=>yes;
proc logistic data= cmc;
class Contraceptive Education/ param=ref;
model Contraceptive(order=data)= Age Education Children Index
     Age*Age Age*Children Age*Index Children*Children
     /link=glogit aggregate=(Age Education Children Index) scale=none lackfit rsquare;
run;
proc logistic data= cmc;
class Contraceptive Index/ param=ref;
model Contraceptive(order=data)= Age Education Children Index
     Age*Age Age*Children Age*Index Children*Children
     /link=glogit aggregate=(Age Education Children Index) scale=none lackfit rsquare;
run;

*****************************************************************************;*******************************
*Predicted probabilities
*****************************************************************************;*******************************
ods graphics on;
proc logistic data= cmc plots= (influence effectplot);
class Contraceptive/ param=ref;
model Contraceptive(order=data)=Age Education Children Index
     Age*Age Age*Index
     /link=glogit aggregate scale=none lackfit rsquare  ;
     output out = results predprobs=I ;
run;

proc print data=results ;
ods graphics off;
*****************************************************************************;*******************************
*Check for ordinality of education and index again;
proc logistic data= cmc plots= (influence effectplot);
class Contraceptive/ param=ref;
model Contraceptive(order=data)=Age Education Children Index
     Age*Age Age*Index
     /link=glogit aggregate scale=none lackfit rsquare  ;
     output out = results predprobs=I ;
run;
*education;
proc logistic data= cmc plots= (influence effectplot);
class Contraceptive Education/ param=ref;
model Contraceptive(order=data)=Age Education Children Index
     Age*Age Age*Index
     /link=glogit aggregate scale=none lackfit rsquare  ;
     output out = results predprobs=I ;
run;
*index;
proc logistic data= cmc plots= (influence effectplot);
class Contraceptive Index/ param=ref;
model Contraceptive(order=data)=Age Education Children Index
     Age*Age Age*Index
     /link=glogit aggregate scale=none lackfit rsquare  ;
     output out = results predprobs=I ;
run;
 *****************************************************************************;***************;
*Cumulative logit propotional odds;
proc logistic data= cmc ;
class Religion Employed Media_Exposure Contraceptive/ param=ref;
model Contraceptive(order=data)= Age Education Children Religion Employed Index Media_Exposure;
run;
*****************************************************************************;*******************************
```