```
1   ************************************************************
2   *Multinomial Logistic Regression - Breast Cancer Data
3   ************************************************************
4
5   set more off, permanently
6
7   log using multinom.log, name(multinom_reg) replace
8
9   insheet using Br_Cancer_New.csv, clear
10
11  ************************************************************
12
13  summ
14
15  ************************************************************
16  *Recoding variables
17  ********************
18
19  label define menlbl 0 "regular" 1 "irregular"
20  label val mencycl menlbl
21
22  label define menstatlbl 0 "premenopause" 1 "menopause"
23  label val menstat menstatlbl
24
25  label define partlbl 0 "parous" 1 "nulliparous"
26  label val parity partlbl
27
28  label define famhistlbl 0 "no" 1 "yes"
29  label val famhist famhistlbl
30
31  label define smokelbl 0 "no" 1 "yes"
32  label val smokhist smokelbl
33
34  label define alclbl 0 "no" 1 "yes"
35  label val alchist alclbl
36
37  label define exelbl 0 "no" 1 "yes"
38  label val exercise exelbl
39
40  label define brcanlbl 0 "benign" 1 "atypical" 2 "suspicious" 3 "malignant"
41  label val brcancer brcanlbl
42
43  label define ethniclbl 1 "others" 2 "kamba" 3 "kalenjin" 4 "luhya" 5 "luo" 6 "kikuyu"
44  label val ethniclev ethniclbl
45
46  label define educlbl 0 "none" 1 "primary" 2 "secondary" 3 "tertiary"
47  label val educlev educlbl
48
49  label define emplbl 0 "no" 1 "yes"
50  label val emplev emplbl
51
52  label define contralbl 0 "none" 1 "implant only" 2 "injection only" 3 "oral only" 4
    "injection+implant" 5 "oral+implant" 6 "oral+injection" 7 "oral+injection+implant"
53  label val contralev contralbl
54
55  ************************************************************************************
    ************************************************************************
56
```

```
*Descriptive stats
************************

tabstat age, stat(mean median min max)

foreach var of varlist mencycl-parity famhist-exercise brcancer-contralev {


    tab `var'


}

*******************************************************************************

*Univariable multinomial logistic regression (P<=0.20)
*****************************************************

mlogit brcancer age, nolog rrr // age is significant
test age

mlogit brcancer mencycl, nolog rrr // mencycl is not significant
test mencycl

mlogit brcancer menstat, nolog rrr // menstat is significant
test menstat

mlogit brcancer parity, nolog rrr // parity is significant
test parity

mlogit brcancer famhist, nolog rrr // famhist is significant
test famhist

mlogit brcancer smokhist, nolog rrr // smokhist is not significant
test smokhist

mlogit brcancer alchist, nolog rrr // alchist is not significant
test alchist

mlogit brcancer exercise, nolog rrr // exercise  is significant
test exercise

mlogit brcancer i.ethniclev, nolog rrr // ethniclev is significant
testparm i.ethniclev

mlogit brcancer ib3.educlev, nolog rrr // educlev is significant
testparm ib3.educlev

mlogit brcancer emplev, nolog rrr // emplev is significant
test emplev

*Recategorise 'contralev' owing to few observations in some cells
recode contralev (0=0 "no") (1/7=1 "yes"), gen(contralev_ct)
tab contralev_ct

mlogit brcancer contralev_ct, nolog rrr // contralev_ct is significant
test contralev_ct

```

```
115  ********************************************************************************
116
117  *Multivariable multinomial logistic regression (P<0.05)
118  ******************************************************
119
120  mlogit brcancer age menstat parity famhist exercise i.ethniclev ///
121   ib3.educlev emplev contralev_ct, nolog rrr
122
123  test age // not significant
124  test menstat // not significant
125  test parity // *significant
126  test famhist // not significant
127  test exercise // *significant
128  test emplev // not significant
129  test contralev_ct // *significant
130
131  testparm i.ethniclev // not significant
132  testparm ib3.educlev // not significant
133
134  ************************************************************
135  *Parsimonious Model
136  ********************
137
138  mlogit brcancer parity exercise contralev_ct, nolog rrr
139
140  test parity
141  test exercise
142  test contralev_ct
143
144  ********************************************************************
145  *Testing assumption of Independence of irrelevant alternatives (IAA)
146  ********************************************************************
147
148  xi:mlogit brcancer parity exercise contralev_ct, baseoutcome(3) nolog rrr
149  mlogtest, hausman // have to install the mlogtest command
150
151
152  log close multinom_reg
153
154  ********************************************************************************
     ***
155
```