s282

282. Calculating excess and 'avoidable' deaths from life tables.

(a) Load the Melanoma data, drop subjects diagnosed 1975-1984.

```
. //==================//
. // EXERCISE 282
. // REVISED MAY 2015
. //==================//
.
. use melanoma11, clear
(Skin melanoma, diagnosed 1975-94, follow-up to 1995)

. keep if year8594 == 1
(3,031 observations deleted)

. stset surv_mm, fail(status==1 2) id(id) scale(12)

                id:  id
     failure event:  status == 1 2
obs. time interval:  (surv_mm[_n-1], surv_mm]
 exit on or before:  failure
    t for analysis:  time/12
------------------------------------------------------------------------------
      4,744  total observations
          0  exclusions
------------------------------------------------------------------------------
      4,744  observations remaining, representing
      4,744  subjects
      1,404  failures in single-failure-per-subject data
   22,108.5  total analysis time at risk and under observation
                                                at risk from t =         0
                                     earliest observed entry t =         0
                                          last observed exit t =  10.95833

. set trace off

. strs using popmort, notables br(0(1)5) mergeby(_year sex _age) by(agegrp sex) save(replace)

         failure _d:  status == 1 2
   analysis time _t:  surv_mm/12
                 id:  id

No late entry detected - p is estimated using the actuarial method

.
. use grouped, clear
(Collapsed (or grouped) survival data)

. keep start end n cp cp_e2 cr_e2 sex agegrp
```

(b) What is the difference in five-year relative survival between males and females in each age group?

```
. list agegrp sex cr_e2 if end == 5, noobs sepby(agegrp)
```

agegrp	sex	cr_e2
0-44	Male	0.8236
0-44	Female	0.9233
45-59	Male	0.7969
45-59	Female	0.8740
60-74	Male	0.7413
60-74	Female	0.7958
75+	Male	0.6627
75+	Female	0.7006

(c) Reshape the data.

```
. bysort sex (agegrp start): gen j = _n
. gen sexlab =cond(sex==1,"_m","_f")
. drop sex
. reshape wide start end n cp cp_e2 cr_e2 agegrp, i(j) j(sexlab) string
(note: j = _f _m)
Data                               long   ->   wide
-----------------------------------------------------------------------------
Number of obs.                       40   ->      20
Number of variables                   9   ->      15
j variable (2 values)            sexlab   ->   (dropped)
xij variables:
                                  start   ->   start_f start_m
                                    end   ->   end_f end_m
                                      n   ->   n_f n_m
                                     cp   ->   cp_f cp_m
                                  cp_e2   ->   cp_e2_f cp_e2_m
                                  cr_e2   ->   cr_e2_f cr_e2_m
                                 agegrp   ->   agegrp_f agegrp_m
-----------------------------------------------------------------------------
. rename agegrp_m agegrp
. rename start_m start
. rename end_m end
. drop agegrp_f start_f end_f
```

(d) For males, calculate the expected number of all-cause deaths, Nd m, the expected number of deaths if the study population were free of cancer, NExp d m and the excess deaths associated with a diagnosis of cancer, ED m.

```
. bys agegrp: gen Nrisk_m = n_m[1]/10
. gen p_dead_m = 1 - cp_e2_m * cr_e2_m
. gen Nd_m = Nrisk_m*p_dead_m
. gen NExp_d_m = Nrisk_m*(1-cp_e2_m)
. gen ED_m = Nd_m - NExp_d_m
```

```
. format Nd_m NExp_d_m ED_m %4.1f
. list agegrp Nrisk_m p_dead_m Nd_m NExp_d_m ED_m if end==5, noobs
```

agegrp	Nrisk_m	p_dead_m	Nd_m	NExp_d_m	ED_m
0-44	53.7	.1889797	10.1	0.8	9.3
45-59	75.2	.2440302	18.4	3.9	14.5
60-74	70.9	.3905036	27.7	12.6	15.1
75+	33.7	.6542017	22.0	16.1	5.9

```
. table agegrp if end == 5, c(sum Nd_m sum NExp_d_m sum ED_m) row format(%4.1f)
```

_m agegrp	sum(Nd_m)	sum(NExp_d_m)	sum(ED_m)
0-44	10.1	0.8	9.3
45-59	18.4	3.9	14.5
60-74	27.7	12.6	15.1
75+	22.0	16.1	5.9
Total	78.2	33.4	44.8

i. We would expect to see 10, 18, 28 and 22 all cause deaths in the (ascending) age groups.
ii. This is given by the excess deaths, ED m. In ascending age groups there are 9, 14, 15, and 6 excess deaths at 5 years post diagnosis when compared to a similar cancer free population. This is for a typical cohort diagnosed in one calendar year.
iii. There are 45 excess deaths when compared to the general population.

(e) Repeat calculations for females.

```
. bys agegrp: gen Nrisk_f = n_f[1]/10
. gen p_dead_f = 1 - cp_e2_f * cr_e2_f
. gen Nd_f = Nrisk_f*p_dead_f
. gen NExp_d_f = Nrisk_f*(1-cp_e2_f)
. gen ED_f = Nd_f - NExp_d_f
. format Nd_f NExp_d_f ED_f %4.1f
. list agegrp Nrisk_f p_dead_f Nd_f NExp_d_f ED_f if end==5, noobs
```

agegrp	Nrisk_f	p_dead_f	Nd_f	NExp_d_f	ED_f
0-44	62.4	.0814915	5.1	0.3	4.8
45-59	61.2	.1431934	8.8	1.2	7.6
60-74	66.1	.2800009	18.5	6.3	12.2
75+	51.2	.5766043	29.5	20.3	9.3

```
. table agegrp if end == 5, c(sum Nd_f sum NExp_d_f sum ED_f) row format(%4.1f)
```

_m agegrp	sum(Nd_f)	sum(NExp_d_f)	sum(ED_f)
0-44	5.1	0.3	4.8
45-59	8.8	1.2	7.6
60-74	18.5	6.3	12.2

75+	29.5	20.3	9.3
Total	61.9	28.1	33.8

(f) In terms of the total number of all cause deaths, females have fewer at all ages except the 70+ group. This is because they are more females diagnosed in this group 51 vs 34, so even though females have lower relative survival they have more deaths due to a number of women in the oldest age groups being diagnosed. This leads to there being more excess deaths in this age group for women when compared to men. As a whole there are more excess deaths in men. How many deaths would be 'avoided' if males could achieve the same relative survival as females for Melanoma?

```
. gen Nd_m_f = Nrisk_m*(1 - cp_e2_m * cr_e2_f)
. gen AD_m = Nd_m - Nd_m_f
. format Nd_m_f AD_m %4.1f
. list agegrp Nrisk_m p_dead_m Nd_m NExp_d_m ED_m Nd_m_f AD_m if end==5, noobs
```

agegrp	Nrisk_m	p_dead_m	Nd_m	NExp_d_m	ED_m	Nd_m_f	AD_m
0-44	53.7	.1889797	10.1	0.8	9.3	4.9	5.3
45-59	75.2	.2440302	18.4	3.9	14.5	12.9	5.5
60-74	70.9	.3905036	27.7	12.6	15.1	24.5	3.2
75+	33.7	.6542017	22.0	16.1	5.9	21.4	0.7

There would be about 15 deaths 'avoided'. The youngest two age groups contribute most to the avoidable deaths.

(g) List the avoidable deaths for the oldest age group over all follow-up times. Why are the number of avoidable deaths decreasing as follow-up time increases?

```
. list agegrp end AD_m if agegrp==3
```

	agegrp	end	AD_m
16.	75+	1	1.4
17.	75+	2	2.2
18.	75+	3	2.1
19.	75+	4	1.2
20.	75+	5	0.7

This is because we can not avoid deaths for ever. Remember that we are looking at all cause deaths. If we had unlimited follow-up we would avoid no deaths at all. In the oldest age group we can actually see that we are just postponing deaths.