ECON 245 PROJECT: Notes on Life Span Study (LSS) of atomic bomb survivors

This set of notes supplements the `readme_lung_2017.pdf`.

The dataset was created as a person-year table. Person-years and numbers of cases were stratified on city, sex, attained age, age at the time of the bombing, calendar year, ground distance from the hypocenter, smoking status, DS02R1 weighted absorbed lung dose, and whether an individual's unweighted total shielded kerma was above 4 Gy.

The number of strata defined by each of the above variables are:

`city`: 2
`sex`: 2
`agxcat`: 15
`agecat`: 18
`time`: 13
`distcat`: 3
`scat`: 4
`dcat`: 24
`un4gy`: 2

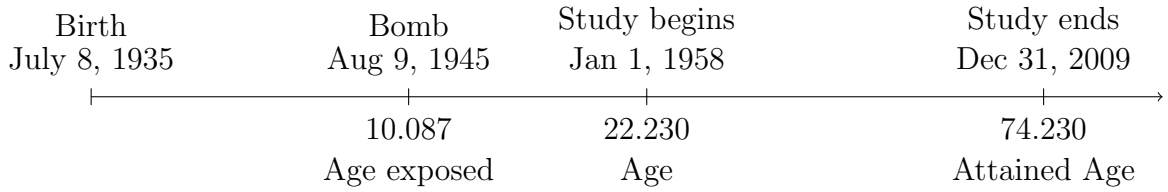
The cross-classifications of these variables creates a potential (2)(2)(15)(18)(13)(3)(4)(24)(2) = 8087040 possible strata combinations. However, not all these strata are represented in the survivors in the LSS.

Each row in the file `lun_smk_naga.csv` represents data from one stratum, not one individual (subject). The number of cases are (`lung` or `larynx` or `othresp`) incidences within each stratum. The exposure are *total* person-years (`upyr` or `pyr` or `pyr92`) between all cases in the same stratum. Since data from multiple individuals have been combined to create strata, it would not be possible de-construct the stratum data to recover data at the individual level.

In a cohort study, the analysis usually involves estimation of incidence (rates) of events during a defined period of observation. Incidence is often defined as the rate of newly diagnosed cases per 100,000 person years of observation time. Therefore it is important to understand the meaning of person years.

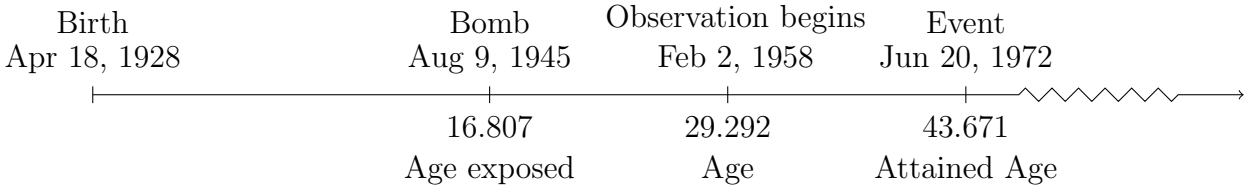
The following are two hypothetical individuals in the LSS.

Person 1 is present for the entire study (i.e. from Jan 1, 1958 to Dec 31, 2009).



Total person years at risk = 74.230 - 22.230 = 52 years.

Person 2 is present on Feb 2, 1958 but follow-up ends on Jun 20, 1972 when the event of interest takes place (cancer occurrence, death, loss of follow-up etc).



Total person years at risk = 43.671 - 29.292 = 14.379