Descriptive Statistics

The data set that we used in our project has 167 observations. Each one represents a survival patient with advanced lung cancer from the North Central Cancer Treatment Group. There are 10 independent variables. We put more attention on ”ph.ecog”, ” ph.karno” and ”pat.karno” because they are hard to understand for non-medical people.“ph.ecog” is a index of performance status and can be used to evaluate the effectiveness of therapies. It has 6 levels from 0 to 5. “0” represents fully active, able to carry on all performance without restriction. “1” represents restricted in physically strenuous activities but able to carry out work of a light nature such as light house work. “2” represent ambulatory and capable of all selfcare but unable to carry out any work activities. “3” is even worse, capable of only limited selfcare and confined to bed or chair. “4” means completely disabled. “5” is dead. So the degree of severity is deeper from “0” to “5”. “ph.larno” is another widely used method to assess the functional status of a patient. But the difference is the range of “ph.larno” is from 0 to 100. 0 represents dead and 100 represents normal. So higher scores of “ph.larno” means higher healthy status. They are rated by physician. ”pat.karno” is the same karnofsky performance score as rated by patient. The only categorical variable in the dataset is “sex”, it has two levels that 1 represents male and 2 represents female. The other variables are “age”(ages in years), “meal.cal”(calories consumed at meals), “wt.loss”(weight loss in last six months), “time”(survival times in days) and “status” (1=censored, 2=dead).

In the dataset 63% of observation is male and 37% of observation is female.

We analysis the variable “ph.ecog”. 28.14% of patient in level 0, 48.5% in level1, 22.75% in level 2, 0.6% in level 3 and nobody in level 4 or 5. Most of patients are in level 2 and 3.

Of those who was in level 0, 42.55% of patient was censored and 57.44% of patient was dead. Of those who was in level 1, 27.16% of patient was censored and 72.84% of patient was dead. Of those who was in level 2, 13.16% of patient was censored and 86.84% of patient was dead. Because only 1 patient in level 3. So the death rate is 100%. It shows people in high level of ECOG has a higher death rate.