MKT 500T

Spring 2018 HW5

Sami Cheong

Feb 25 2018

Q1 Hazard Function Business Examples

Monotonically Increasing

Machine components An obvious example here is the failure rate of a machine component in a manufacturing process, which will have a higher failure rate over time due to wear and tear, so it is reasonable to correpsond a monotonicallu increasing hazard function to this process.

Monotonically Decreasing

Service industry Consider Tom the electrician, who owns his own business and trains his own apprentices. In order to stay competitive, he has to make sure everyone can complete an assignment (usually very specific tasks) in one hour or less. The harzard function can be used to describe the rate of failing to complete a job within an hour, and t is the number of years of experience. This is clearly a monotonically decraseing function, as we know
the more experience an electrician has, the less time it should take him/her to perform a task.

U-shaped

Car insurance example In this example, let us consider $h(t)$ to be the risk of clarims for a car insurance policy, and t here stands for the number of years a policy holder has been driving. A U-shape curve tells that the hazard rate is higher when t is small and when t is large. So, in our example, this corresponds well with new drivers (small t) and senior drives (large t).

Upside-down U-shaped

An example on graduate school drop-out risk: Consider Mary, a graduate student who just began a program for biomedical research. She is very optimistic and excited her future in the field, so her risk of dropping out (failure rate) of the graduate program in her first year is relatively low. However, after a few semesters of heavy course work, unsuccessful research experiments, and meager stipends, she might begin to wonder if she is better off getting a job instead of staying in the academic world. At this point in time, the risk of dropping out will start to climb and eventually peaks. However, by year 4, Mary realizes that she is almost done, so her risk of dropping out late in the grad program will start to decrease again.

Q2 KB dataset

The exponential-gamma model is defined as $P(T \leq t|r, \alpha) = 1 - (\frac{\alpha}{\alpha+t})^r$, when fitted on the KB dataset the parameters are optimized at $r = 0.0503$ and $\alpha = 7.973$. In here, we'd like to find $P(T \leq 2|T > 1)$. Notice that

$$P(T > t + s|T > s) = \frac{P(T > t + s, T > s)}{P(T > s)} = (\frac{\frac{\alpha}{\alpha+t+s}}{\frac{\alpha}{\alpha+s}})^r = (\frac{\alpha + s}{\alpha + t + s})^r,$$

which means that $P(T \le t+s|T>s) = 1-(\frac{\alpha+s}{\alpha+t+s})^r$. Since the data is recorded weekly,$t=s=52$, so $P(T \le 104|T>52) = (\frac{\alpha+52}{\alpha+104})^r = 0.3088$. So there is around 3% chance that someone will purchase KB by the end of year 2 given they haven't made a purchase yet in their first year. This is almost 3 times less likely compared to the probability that someone would make an initial purchase by the end of year 1 (i.e.$P(T \le 1) = 1-(\frac{\alpha}{\alpha+52})^r$), the difference here can be attributed to the Bayesian update based on information learned from previous year.

Q3 comScore data

To fit the 'exponential with never triers' model, we need to solve for p and θ by setting:

$$p(1-e^{7\theta}) - 0.716 = 0$$

and

$$p(1-e^{30\theta}) - 0.808 = 0$$

```r
library(nleqslv)
EN<-function(par){
  p <-par[1]
  theta <- par[2]

  y<-numeric(2)
  y[1]<-p*(1-exp(-7*theta)) - 0.716
  y[2]<-p*(1-exp(-30*theta)) - 0.808

  return(y)
}

par_EN_solv<-nleqslv(c(0.1,0.5),
                 EN,
                 control = list(ftol = 1e-10, allowSingular = TRUE),
                 jacobian = TRUE,method = 'Newton')
p<-par_EN_solv$x[1]
theta<-par_EN_solv$x[2]
yearly_reach<-p*(1-exp(-365*theta))
print(yearly_reach)
```

```
## [1] 0.8080732
```

Interestingly, the yearly reach forecasted based on these 2 data points shows a very close result to the monthly reach. Now, using the exponential-gamma, again we need to solve for a system of equations:

$$1-(\frac{\alpha}{\alpha+7})^r = 0.716,$$

and

$$1-(\frac{\alpha}{\alpha+30})^r = 0.808$$

```r
library(nleqslv)
EG<-function(par){
  alpha <-par[1]
  r <- par[2]

  y<-numeric(2)
  y[1]<-1-(alpha/(alpha+7))^r - 0.716
  y[2]<-1-(alpha/(alpha+30))^r - 0.808
```

```r
  return(y)
}

par_EG_solv<-nleqslv(c(0.5,0.5),
                  EG,
                  control = list(ftol = 1e-10, allowSingular = TRUE),
                  jacobian = TRUE,method = 'Newton')
alpha<-par_EG_solv$x[1]
r<-par_EG_solv$x[2]
yearly_reach_EG<-1-(alpha/(alpha+365))^r
print(yearly_reach_EG)
```

```
## [1] 0.9022422
```

This time, the yearly-reach is projected to be closer to 90.2%. Compared to the first estimate, this number seems a lot more optimistic, as it does not account for the never-trier group. Without more knowledge about the customer behavior and more data points, our best guess would be somewhere around 81% to 90%.