MKTG776 HW5

Jordan Farrer

2017-02-24

Contents

1 Question 1

Hazard Function Shape	Business Story
Monotonically increasing	Probability of response from customer service at a tech company. All young tech companies have great customer support and you will hear back from them, with certainty. However, the probability you hear back in the next minute is always greater than the probability you hear back in the current minute.
Monotonically decreasing	Probability of reordering from a food delivery service after your first order. The highest probability occurs in the following day and the next highest day is second day. Likewise, this probability of reordering falls monotonically to some long-run probability.
U-shaped	Probability of responding to an email from a coworker with a significant ask. You may immediately respond saying it's best to meet or that you'll follow up with a longer response or you can not respond immediately. If so eventually the probability that you will respond only goes up over time after bottoming out. The assumption is that you will eventually respond to all such emails.
Upside-down U-shaped	Probability of being promoted from the time you enter a new role. You are extremely unlikely of being promoted immediately after starting and the probability increases over time. However, at some point there is a chnage where if you didn't get promoted yet, the likelihood of getting promoted levels off at a lower state.

2 Question 2

```
pacman::p_load(tidyverse, pander)
panderOptions('round', 4)
panderOptions('keep.trailing.zeros', TRUE)
options(scipen = 10, digits = 5)

kb <-
  readxl::read_excel("Krunchy Bits.xlsx", skip = 5, col_names = c("week", "num_households")) %>%
  filter(complete.cases(.))
```

```
fn_eg <- function(r, alpha, t) {
  p_x <- 1 - (alpha / (alpha + t))^r
  return(p_x)
```

```
}

fn_eg_ll <- function(par, t, x, x_total) {
  p_x <- fn_eg(par[1], par[2], t)
  p_x_diff <- p_x - dplyr::lag(p_x, default = 0)
  incremental <- x - dplyr::lag(x, default = 0)
  ll <- sum(incremental * log(p_x_diff)) + (x_total - x[length(x)])*log(1 - p_x[length(p_x)])
  return(-ll)
}

params_eg <- nlminb(c(1,1), fn_eg_ll, lower = c(0,0),
                    upper = c(Inf,Inf), t = kb$week, x = kb$num_households, x_total = 1499)$par

kb_r <- params_eg[1]
kb_alpha <- params_eg[2]
```

We estimate the model parameters of the EG model to be $r = 0.05025$ and $\alpha = 7.97338$.

```
kb_probs <-
  kb %>%
    mutate(
      `P(T <= t)` = map_dbl(week, fn_eg, r = kb_r, alpha = kb_alpha)
      , `P(T = t)` = `P(T <= t)` - lag(`P(T <= t)`, default = 0)
    )
```

Table 2: Exponential-Gamma Model on KB Dataset

week	num_households	P(T <= t)	P(T = t)
1	8	0.0059	0.0059
2	14	0.0112	0.0053
3	16	0.0159	0.0047
4	32	0.0202	0.0043
5	40	0.0242	0.0039
6	47	0.0278	0.0036
7	50	0.0312	0.0034
8	52	0.0343	0.0031
9	57	0.0373	0.0029
10	60	0.0400	0.0028

```
two_years <-
  data_frame(
    week = c(52, 52*2)
    , `P(T <= t)` = map_dbl(week, fn_eg, r = kb_r, alpha = kb_alpha)
    , `P(T = t)` = `P(T <= t)` - lag(`P(T <= t)`, default = 0)
  )
```

Table 3: Exponential-Gamma Model on KB Dataset for Year 1 and Year 2

week	P(T <= t)	P(T = t)
52	0.0964	0.0964
104	0.1243	0.0279

We know that the probability of "failing" (in our case buying KB) in the next *small* interval of time given "survival" to time t is

$$P(t < T \leq t + \Delta T | T > t) \approx h(t) \times \Delta T \tag{1}$$

where the hazard function, $h(t)$ for the EG model is

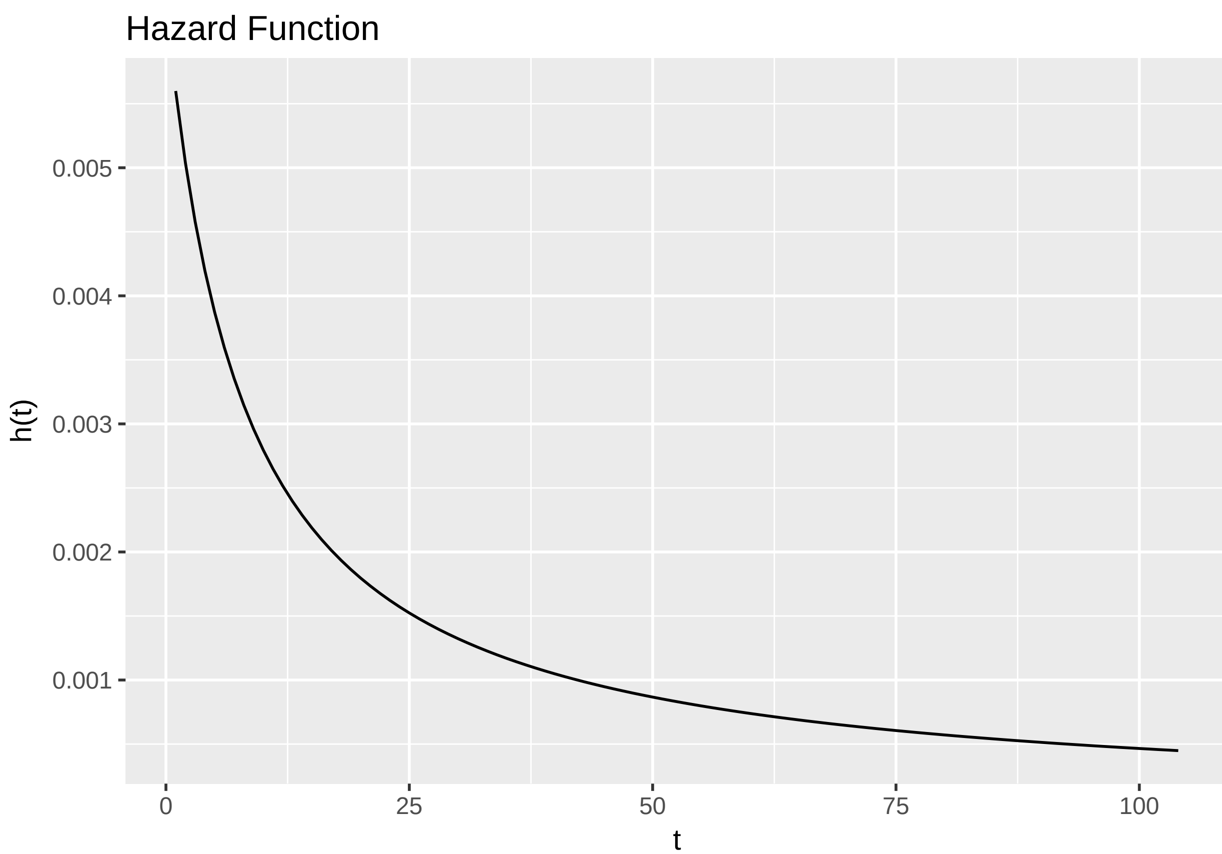
$$h(t|\alpha, r) = \frac{r}{\alpha + t} \tag{2}$$

So, the probability that someone who hasn't yet purchased KB by the end of the first year will make their initial purchase before the end of year 2 is

$$\begin{align}
P(t < T \leq t + \Delta T | T > t) &\approx h(t) \times \Delta T \tag{3} \\
&\approx \frac{r}{\alpha + t} \times \Delta T \tag{4} \\
&\approx \frac{0.05025}{7.97338 + 52} \times 52 \tag{5} \\
&\approx 0.04357 \tag{6}
\end{align}$$

On the other hand, the probability that a randomly chosen person makes an initial purchase within year 1 is simply **0.09642**. The reason for the drop-off in probability is that a customer who hasn't purchased KB for 52 weeks (1 year) in unlikely to suddenly purchase KB in the next subsequent 52 weeks (year 1 to year 2). There were many opporunities to purchase in year one. We can see this from the shape of the hazard function:

```r
data_frame(
  t = 1:104
) %>%
  mutate(
    hazard_function = kb_r / (kb_alpha + t)
    )  %>%
  ggplot(aes(x = t, y = hazard_function)) +
  geom_line() +
  labs(y = "h(t)", title = "Hazard Function")
```


Hazard Function
h(t)
0.005
0.004
0.003
0.002
0.001
0
25
50
75
100
t