Webappendix for the letter entitled novel way of quantifying the long-term benefit of immunotherapy using quantile regression

Bassirou Mboup, Christophe Le Tourneau, Aurélien Latouche

7/10/2019

This document makes an easier access to the supplementary material of the letter entitled **novel way of quantifying the long-term benefit of immunotherapy using quantile regression**.

1) Data set

We use the algorithm of guyot. al 2012 to reconstructing individual-level time-to-event data based on the published Kaplan–Meier curves of the randomized controlled trial (Rittmeyer et al. 2017). This algorithm is available in code R based on the inverse of the Kaplan Meier equation.

After reconstruction, we get in this dataset the following variables.

- time : vector of observed failure times.
- event: vector of indicator of status (0 for censoring, 1 for type of event).
- tmt.arm.number: vector of treatment indicator (1 if treated and 0 otherwise).
- treatment.type: the type of treatment immonotherapy or chemotherapy

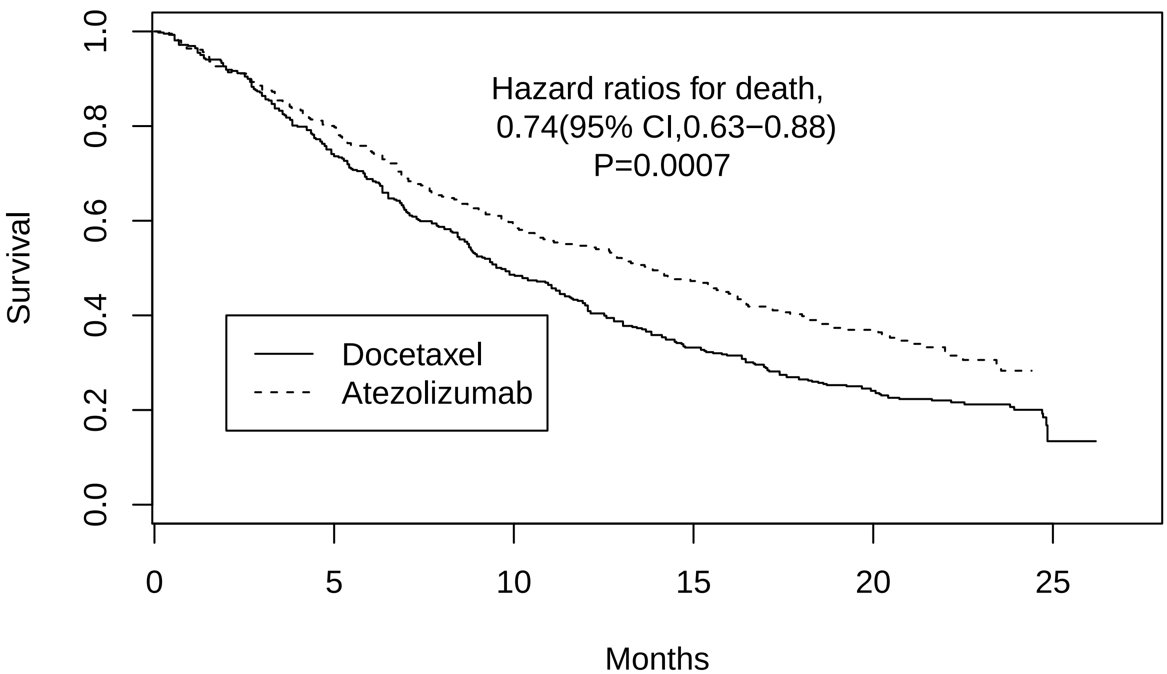
```
##        time event tmt.arm.number treatment.type
## 1 0.4059140     1              1   Atezolizumab
## 2 0.4059140     1              1   Atezolizumab
## 3 0.4059140     1              1   Atezolizumab
## 4 0.5599768     1              1   Atezolizumab
## 5 0.5599768     1              1   Atezolizumab
## 6 0.5599768     1              1   Atezolizumab
```

```r
library(survival)
fit_cox <- coxph(Surv(time, event)~tmt.arm.number,data=data_ICI_Rittmeyer)
summary(fit_cox)
```

```
## Call:
## coxph(formula = Surv(time, event) ~ tmt.arm.number, data = data_ICI_Rittmeyer)
##
##   n= 850, number of events= 555
##
##                    coef exp(coef) se(coef)     z Pr(>|z|)
## tmt.arm.number -0.29602   0.74377  0.08733 -3.39    7e-04 ***
## ---
## Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
##
##                exp(coef) exp(-coef) lower .95 upper .95
## tmt.arm.number    0.7438      1.344    0.6268    0.8826
##
## Concordance= 0.535  (se = 0.012 )
## Likelihood ratio test= 11.68  on 1 df,   p=6e-04
```

```
## Wald test            = 11.49  on 1 df,   p=7e-04
## Score (logrank) test = 11.57  on 1 df,   p=7e-04
```

```r
fit_KM <- survfit(Surv(time,event)~tmt.arm.number,data=data_ICI_Rittmeyer)

plot(fit_KM, xlab=c("Months"),
     ylab=c("Survival"),xlim=c(0.98,27),
     mark.time=FALSE, lty=c(1, 2))

legend(2, 0.4, legend=c("Docetaxel","Atezolizumab"), lty=c(1, 2))
text(14, 0.8, c("Hazard ratios for death,\n  0.74(95% CI,0.63-0.88)\n P=0.0007"))
```



2) Model

Let $Treat$ the treatment indicator ($Treat = 1$ if therapy is immuno, $Treat = 0$ otherwise), T the time of event and C the censored time. Let $\widetilde{T} = \min(T, C)$ the observed time, $\delta = \mathbb{1}_{(T \leq C)}$ the indicator of event ($\delta = 1$ if event and $\delta = 0$ if censored). The observed data consist of n iid replicates of $(\widetilde{T},\ \delta\ ,\ Treat)$, denoted by $\{\widetilde{T}_i,\ \delta_i\ ,\ Treat_i),\ i = 1, \ldots, n\}$. Without loss of generality the conditional quantiles of a random variable is defined as $Q_T(\tau|Treat) = \inf\{t : \mathrm{P}(\widetilde{T} \leq t|Treat) \geq \tau\}, 0 \leq \tau \leq 1$. We consider the following quantile regression model

$$Q_T(\tau|Treat) = \exp\{\beta_0(\tau) + \beta_1(\tau)Treat\}$$

where $\boldsymbol{\beta}(\tau) = (\beta_0(\tau), \beta_1(\tau))$ is the vector of unknown quantile regression coefficients describing the effects of the treatment on the τth quantile of $\log(T)$.

To estimates the coefficients $\boldsymbol{\beta}(\tau)$ of censored quantiles regression, several methods are proposed in the literature Powell et al. 1986, Portnoy et al. 2003, Peng et al. 2008. We use the method of Peng et al. 2008 that is based on the Nelson-Aalen estimator of the cumulative hazard function. This method is similar from let $F_i(t|x_i) = \mathrm{P}(T_i \leq t|Treat_i)$, $\Lambda_i(t|Treat_i) = -\log(1 - F_i(t|Treat_i))$ the cumulative hazard function $N_i(t) = \mathbb{1}\{\widetilde{T}_i \leq t, \delta_i = 1)\}$ a counting process et $M_i(t) = N_i(t) - \Lambda_i\{t\Lambda\widetilde{T}_i)|Treat_i\}$ the martingale process associated with the counting process $N_i(t)$. We know that $\mathbb{E}(M_i(t)|Treat_i) = 0$. They derive the following estimating equation:

$$\mathbb{E}(n^{-1/2}\sum_{i=1}^{n} x_i\big[N_i(\exp(\boldsymbol{x_i'}\boldsymbol{\beta}(\tau))) - \int_0^{\tau} I(\widetilde{T}_i \geq \exp(\boldsymbol{x_i'}\boldsymbol{\beta}(\tau)))dH(u)\big]) = 0 \tag{1}$$

where $\boldsymbol{x_i'} = (1, Treat_i)$ and $H(u) = -\log(1-u)$ for $u \in [0,1)$. The integral in the estimating equation 1 is approximated on a grid $0 = \tau_0 < \tau_1 < \cdots < \tau_L < 1$. The regression quantiles $\beta(\tau_k)$, $k = 1, \ldots, L$, can be estimated sequentially by solving a linear program see (Peng et al. 2008).

We use this approach developed in the R package **quantreg**.

3) Application

```
library(quantreg)
```

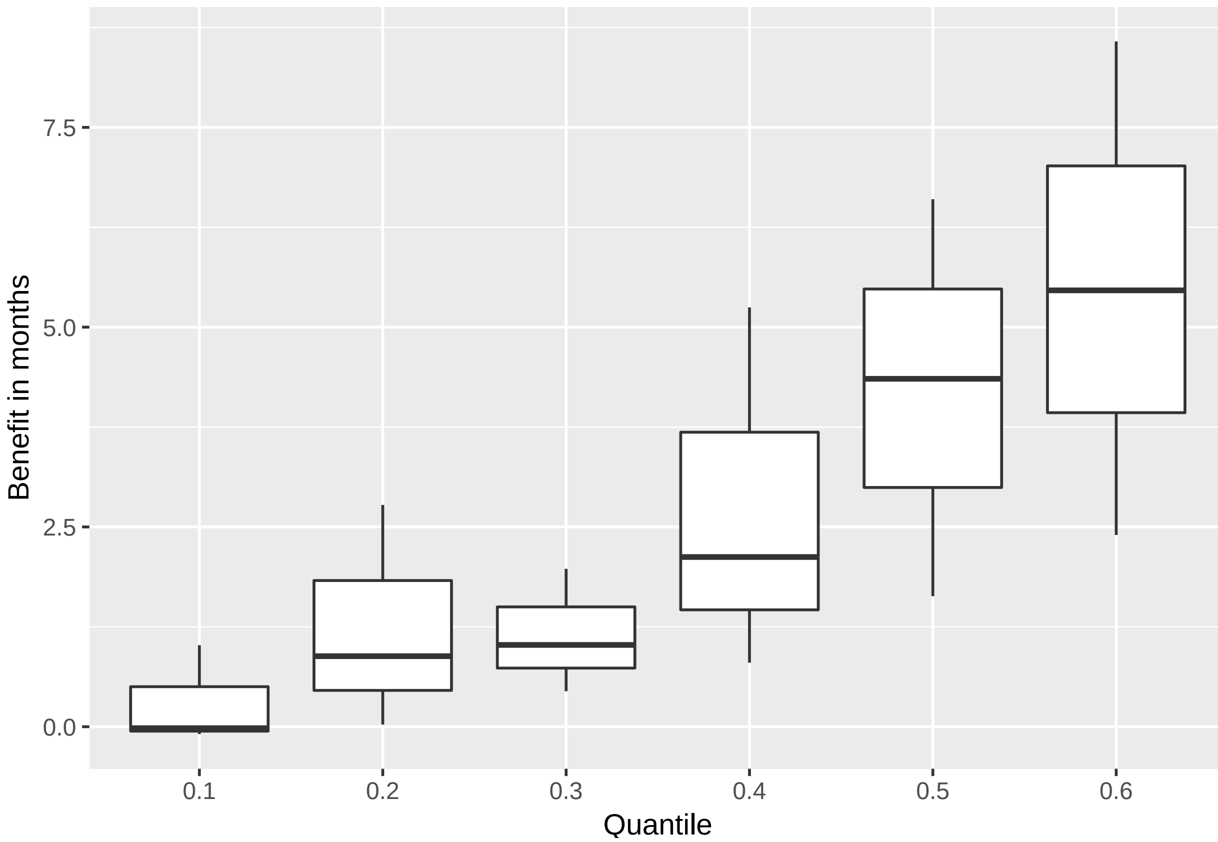
```
## Loading required package: SparseM

##
## Attaching package: 'SparseM'

## The following object is masked from 'package:base':
##
##     backsolve

##
## Attaching package: 'quantreg'

## The following object is masked from 'package:survival':
##
##     untangle.specials
```

```
x <- c( 0.1, 0.2, 0.3, 0.4,0.5,0.6)

Rq <- crq(Surv(time,event)~tmt.arm.number,data=data_ICI_Rittmeyer,method="Pen")
result <- summary(Rq,taus=x)
result
```

```
##
## tau: [1] 0.1
##
## Coefficients:
##                 Value    Lower Bd Upper Bd Std Error T Value  Pr(>|t|)
## (Intercept)      2.70243  2.58645  2.70243  0.02959  91.33789  0.00000
## tmt.arm.number -0.01844 -0.09236  1.01991  0.28375  -0.06498  0.94819
##
## tau: [1] 0.2
##
## Coefficients:
##                 Value   Lower Bd Upper Bd Std Error T Value Pr(>|t|)
## (Intercept)     4.15397 1.73010  4.69080  0.75529   5.49982 0.00000
## tmt.arm.number 0.88255 0.02660  2.77559  0.70129   1.25847 0.20822
##
## tau: [1] 0.3
##
## Coefficients:
##                 Value    Lower Bd Upper Bd Std Error T Value  Pr(>|t|)
```

```
## (Intercept)     5.86294  5.75690  6.26086  0.12857  45.60287  0.00000
## tmt.arm.number  1.02312  0.44421  1.97604  0.39078   2.61815  0.00884
##
## tau: [1] 0.4
##
## Coefficients:
##                Value   Lower Bd Upper Bd Std Error T Value Pr(>|t|)
## (Intercept)    7.80078 4.60256  9.32782  1.20544   6.47129 0.00000
## tmt.arm.number 2.12340 0.80239  5.24705  1.13386   1.87271 0.06111
##
## tau: [1] 0.5
##
## Coefficients:
##                Value    Lower Bd Upper Bd Std Error T Value  Pr(>|t|)
## (Intercept)     9.78031  7.79325 10.55873  0.70549  13.86307  0.00000
## tmt.arm.number  4.35360  1.63372  6.60095  1.26717   3.43568  0.00059
##
## tau: [1] 0.6
##
## Coefficients:
##                Value    Lower Bd Upper Bd Std Error T Value  Pr(>|t|)
## (Intercept)    12.69400 10.75741 13.09069  0.59524  21.32596  0.00000
## tmt.arm.number  5.46031  2.39929  8.57559  1.57562   3.46551  0.00053

## jack

## jack

## jack

## jack

## jack

## jack
```

```r
x <- x[-length(x)]
valu <- c()
lower <- c()
supp <- c()
for(i in 1:length(result)){
  valu[i] <- result[[i]]$coefficients[2,1]
  lower[i] <- result[[i]]$coefficients[2,2]
  supp[i] <-result[[i]]$coefficients[2,3]
}

datafr <- rbind.data.frame(valu,lower,supp)
names(datafr) <- c( "0.1", "0.2", "0.3", "0.4", "0.5","0.6")

library(ggplot2)

p <- ggplot(stack(datafr), aes(x=factor(ind,levels=names(datafr)), y=values))+geom_boxplot()
p+ labs(x= "Quantile",y="Benefit in months")
```



Benefit in months
7.5
5.0
2.5
0.0
0.1
0.2
0.3
0.4
0.5
0.6
Quantile