Utility Functions for Competing Risk

zhedong liu

2022-11-25

Model

We observe $\boldsymbol{D} = (t_i, \delta_i)$, where t_i is the occurrence time of a event, δ_i is the event type of individual i, where $\delta_i = 1, \dots J$ and $i = 1, \dots n$. If only one event can happen, we can use a competing risk model to model this data.

We define t_i is a realization of the random variable T_i and δ_i is a realization of the random variable Z_i. For clarity, in this section we drop the notation i. We define $T = \min(T|Z = 1, \dots, T|Z = J)$ and $Z = \arg\min_{\delta}(T|Z = \delta)$, and $T|Z = j$ being the random variable describing

The probability of not experiencing any event before s is

$$P(T > s) = \prod_{k=1}^{J} P(T_k > s).$$

The density of jth event happens at time t is

$$\pi(T = t, \delta = j) = \pi(T_j = t) \prod_{k \neq j} P(T_k > t).$$

In prediction points of view, the density of jth event happens at time t given nothing happening before time s is more interesting, which is

$$\pi(T = t, \delta = j | T > s) = \frac{\pi(T_j = t) \prod_{k \neq j} P(T_k > t)}{P(T > s)}.$$

The density of observing an event at t is

$$\pi(T = t) = \sum_{j=1}^{J} \pi(T_j = t) \prod_{k \neq j} P(T_k > t).$$

The conditional density of observing an event at t given nothing happening before time s is

$$\pi(T = t | T > s) = \frac{\sum_{j=1}^{J} \pi(T_j = t) \prod_{k \neq j} P(T_k > t)}{P(T > s)} \tag{1}$$

We would also happy to know

$$P(\delta = j) = \int_0^\infty \pi(T_j = t) \prod_{k \neq j} P(T_k > t) dt.$$

This will correspond to the relative frequency of each events.

The conditional version is

$$P(\delta=j|T>s)=\frac{\int_s^\infty \pi(T_j=t)\prod_{k\neq j}P(T_k>t)dt}{P(T>s)}. \tag{2}$$

People are particularly interest in the probability of experiencing jth event within a time interval (s,s+t] given the whole information accumulated till the landmark time s. The probability is

$$P(T<s+t,\delta=j|T>s)=\frac{\int_s^{s+t}\pi(T=x,\delta=j)dx}{P(T>s)}. \tag{3}$$

To summerise, the conditional version is more general since we simply set $s=0$ to get the unconditional version. The above probability can be checked using data.

Prediction Tasks

There are some potentially interesting questions related to prediction:

1. When will the next event (whatever event) happen given no event has happened yet?

With probability .95, we can observe a event happen before t, where $P(T<s+t|T>s)=0.95$. (1)

2. What's the next event happen given no event has happened yet?

The most likely happened event is $\delta=j$ which maximize $P(\delta=j|T>s)$. (2)

3. Will event j happen within t unit of time given no event has happened yet?

The probability that j will happen within t unit of time is $P(T<s+t,\delta=j|T>s)$. (3)

4. When will event j happen given no event has happened yet?

With probability $P(\delta\neq j|T>s)=1-P(\delta=j|T>s)$, j will not happen. (2)

With probability $.95*P(\delta=j|T>s)$, we can observe j happens before t, where $P(T<s+t,\delta=j|T>s)=.95*P(\delta=j|T>s)$. (3)

Scores

It seems question 2 - 4 are more interesting. We may focus on compute scores to check (2) and (3).

Brier Score

We have observed $(T_1,\delta_1),\ldots(T_n,\delta_n)$.

We compute

$$S_B(s)=\frac{1}{n(s)}\sum_{i=1}^{n}\sum_{j=1}^{J}(P(\delta=j|T>s,\boldsymbol{D})-\mathbf{1}_{\delta_i=j})^2$$

to check (2). This is Brier score.

We compute

$$S_B(s,t)=\frac{1}{n(s)}\sum_{i=1}^{n}\sum_{j=1}^{J}(P(T\leq s+t,\delta=j|T>s,\boldsymbol{D})-\mathbf{1}_{T_i\leq s+t,\delta_i=j})^2$$

to check (3). This is Brier score for conditional prediction. (Blanche et al. (2015))

Logarithmic Score

We compute

$$S_L(s) = \frac{1}{n(s)} \sum_{i=1}^{n} \mathbf{1}_{T_i > s} \log \pi(T = T_i, \delta = \delta_i | T > s, \boldsymbol{D})$$

to check (3). This is logarithmic scores or expected cross-entropy, which can check (3) indirectly. (Commenges, Liquet, and Proust-Lima (2012))

Receiver Operating Characteristic Curve and Area Under the Curve

We predict an individual will encounter j within a time interval $(s, s+t]$ when $P(T < s+t, \delta = j | T > s, \boldsymbol{D}) > c$, $\tilde{P} > c$ in short. $\tilde{P}$ is different for each i because of the covariates.

We have the true positive counts,

$$TP_{s,t,j}(c) = \sum_{i=1}^{n} \mathbf{1}_{s < T_i \leq s+t, \delta_i = j} \mathbf{1}_{\tilde{P} > c},$$

false positive counts,

$$FP_{s,t,j}(c) = \sum_{i=1}^{n} \mathbf{1}_{s < T_i \leq s+t, \delta_i \neq j \cup T_i > s+t} \mathbf{1}_{\tilde{P} > c},$$

true negative counts,

$$TN_{s,t,j}(c) = \sum_{i=1}^{n} \mathbf{1}_{s < T_i \leq s+t, \delta_i \neq j \cup T_i > s+t} \mathbf{1}_{\tilde{P} \leq c},$$

and false negative counts

$$FN_{s,t,j}(c) = \sum_{i=1}^{n} \mathbf{1}_{s < T_i \leq s+t, \delta_i = j} \mathbf{1}_{\tilde{P} \leq c}.$$

Then the true positive rate, or sensitivity, is

$$TPR_{s,t,j}(c) = \frac{TP_{s,t}(c)}{TP_{s,t}(c) + FN_{s,t}(c)} = \frac{\sum_{i=1}^{n} \mathbf{1}_{s < T_i \leq s+t, \delta_i = j} \mathbf{1}_{\tilde{P} > c}}{\sum_{i=1}^{n} \mathbf{1}_{s < T_i \leq s+t, \delta_i = j}},$$

and the false positive rate, or specificity, is

$$FPR_{s,t,j}(c) = \frac{FP_{s,t}(c)}{FP_{s,t}(c) + TN_{s,t}(c)} = \frac{\sum_{i=1}^{n} \mathbf{1}_{s < T_i \leq s+t, \delta_i \neq j \cup T_i > s+t} \mathbf{1}_{\tilde{P} > c}}{\sum_{i=1}^{n} \mathbf{1}_{s < T_i \leq s+t, \delta_i \neq j \cup T_i > s+t}}.$$

Then the receiver operating characteristic curve (ROC) is defined by

$$ROC_{s,t,j}(p) = TPR_{s,t,j}(FPR_{s,t,j}^{-1}(p)),$$

and the area under the receiver operating characteristic curve (AUC) is

$$S_{AUC}(s,t) = \int_0^1 ROC_{s,t}(p) dp.$$

Using Bamber's Equivalence theorem, we can compute Wilcoxon statistic, equivalent to AUC,

$$S_{AUC}(s,t,j) = \frac{1}{K_1 K_2} \sum_{i_1=1}^{K_1} \sum_{i_2=1}^{K_2} \mathbf{1}_{\tilde{P}_{i_1} > \tilde{P}_{i_2}},$$

where i_1 are those $s < T_{i_1} < s+t, \delta_{i_1} = j$ and i_2 are the compliments.

Blanche et al. (2015) considers decision about if subject i_1 has higher risk than i_2. That is

$$S_{AUCb}(s,t,j) = \frac{\sum_{i_1=1}^{n} \sum_{i_2=1}^{n} \mathbf{1}_{\tilde{P}_{i_1} > \tilde{P}_{i_2}} \mathbf{1}_{s < T_{i_1} < s+t, \delta_{i_1} = j} (1 - \mathbf{1}_{s < T_{i_2} < s+t, \delta_{i_2} = j})}{\sum_{i_1=1}^{n} \sum_{i_2=1}^{n} \mathbf{1}_{s < T_{i_1} < s+t, \delta_{i_1} = j} (1 - \mathbf{1}_{s < T_{i_2} < s+t, \delta_{i_2} = j})}.$$

Cross Validation

We use cross validation to estimate those scores because we don't have future data. Leave-group-out cross-validation (LGOCV) will be involved when we have longitudinal data jointly modeled with our survival data depending on the definition of current time.

$$S_B(s,t) \approx \frac{1}{n(s)} \sum_{i=1}^{n} \sum_{j=1}^{J} (P(T \leq s+t, \delta = j | T > s, \boldsymbol{D}_{-I_i}) - \mathbf{1}_{T_i \leq s+t, \delta_i = j})^2$$

$$S_L(s) \approx \frac{1}{n(s)} \sum_{i=1}^{n} \mathbf{1}_{T_i > s} \log \pi(T = T_i, \delta = \delta_i | T > s, \boldsymbol{D}_{-I_i})$$

$$S_{AUC}(s,t,j) \approx \frac{1}{K_1 K_2} \sum_{i_1=1}^{K_1} \sum_{i_2=1}^{K_2} \mathbf{1}_{\tilde{P}_{i_1} | \boldsymbol{D_{-i_1}} > \tilde{P}_{i_2} | \boldsymbol{D_{-i_2}}}$$

$$S_{AUCb}(s,t,j) \approx \frac{\sum_{i_1=1}^{n} \sum_{i_2=1}^{n} \mathbf{1}_{\tilde{P}_{i_1} | \boldsymbol{D_{-i_1}} > \tilde{P}_{i_2} | \boldsymbol{D_{-i_2}}} \mathbf{1}_{s < T_{i_1} < s+t, \delta_{i_1} = j} (1 - \mathbf{1}_{s < T_{i_2} < s+t, \delta_{i_2} = j})}{\sum_{i_1=1}^{n} \sum_{i_2=1}^{n} \mathbf{1}_{s < T_{i_1} < s+t, \delta_{i_1} = j} (1 - \mathbf{1}_{s < T_{i_2} < s+t, \delta_{i_2} = j})}.$$

References

Blanche, Paul, Cécile Proust-Lima, Lucie Loubère, Claudine Berr, Jean-François Dartigues, and Hélène Jacqmin-Gadda. 2015. "Quantifying and Comparing Dynamic Predictive Accuracy of Joint Models for Longitudinal Marker and Time-to-Event in Presence of Censoring and Competing Risks." *Biometrics* 71 (1): 102–13.

Commenges, Daniel, Benoit Liquet, and Cécile Proust-Lima. 2012. "Choice of Prognostic Estimators in Joint Models by Estimating Differences of Expected Conditional Kullback–Leibler Risks." *Biometrics* 68 (2): 380–87.