Description of models in bayesDCA

Giuliano Netto Flores Cruz   `cruz.gnf@gmail.com`

Introduction

In Decision Curve Analysis, we are typically interested in estimating the net benefit of adopting a given clinical decision strategy [1]. A decision strategy may be simply treating all patients under suspicion of a given condition, because the condition is so deadly that any costs of intervention are outweighed – e.g., treating aggressive cancer with surgery. That is the "Treat all" strategy, which alludes to intervening in all patients regardless of the nature of the intervention. Conversely, intervention may be risky and the underlying condition, not so much – e.g., surgery for non-aggressive cancer. In this case, a reasonable decision strategy is not to treat any patient – the "Treat none" strategy. A better decision strategy is, perhaps, to decide based on a predictive model which estimates a patient's likelihood of having aggressive disease. If the patient's likelihood is above a decision threshold t, then we intervene. The same idea serves for binary tests, even though the test result is independent of our decision threshold t. We must then be able to estimate NB for these four types of decision strategies: treat all, treat none, model-based decisions, and decisions based on binary tests.

The {`bayesDCA`} R package helps with the above task while properly propagating uncertainty in the NB parameters. We can ultimately infer what's the best decision strategy, which strategies are useful at all, make pairwise comparisons between strategies, and estimate the expected cost of uncertainty imposed by the available data. Here we briefly describe the models behind the Bayesian Decision Curve Analysis implemented in the {`bayesDCA`} R package. We start with binary outcomes, which is the application {`bayesDCA`} was primarily designed for. We then describe its extension to survival (time-to-event) outcomes.

1 Bayesian Decision Curve Analysis for binary outcomes

Suppose we have a validation dataset with a binary outcome. For convenience, we refer to the outcome as disease status or diagnosis (diseased v.s. non-diseased), but the model generalizes for any binary endpoint (e.g., short-term death status or even disease subtype). For each patient, we also have the outcome prediction for models $m = 1, 2, \ldots, M$. This prediction can be a probability prediction (i.e., as in risk prediction models) or a binary prediction (i.e., as in binary tests). For a range of decision thresholds $t \in (0, 1)$, the Net Benefit for model m at threshold t is:

$$\mathrm{NB}_{m,t} = \mathrm{Se}_{m,t} \cdot p - (1 - \mathrm{Sp}_{m,t}) \cdot (1 - p) \cdot \mathrm{odds}(t) \tag{1}$$

where $\mathrm{odds}(t) = t/(1-t)$, Se represents sensitivity, Sp represents specificity, and p represents the disease prevalence or outcome proportion. The equation (1) has three unknown parameters, namely $\mathrm{Se}_{m,t}$, $\mathrm{Sp}_{m,t}$, and p. The first first two are model- and threshold-specific, while the p is the same for all models and thresholds.

To estimate these from cross-sectional data, the {bayesDCA} R package implements the following model:

$$\text{Diseased} \sim \text{Binomial}(\text{N} = n, p)$$
$$\text{True Positives}_{m,t} \sim \text{Binomial}(\text{Diseased} = d, \text{Se}_{m,t}) \tag{2}$$
$$\text{True Negatives}_{m,t} \sim \text{Binomial}(\text{Non-diseased} = n - d, \text{Sp}_{m,t})$$

$$p, \text{Se}_{m,t}, \text{Sp}_{m,t} \sim \text{Beta}(\alpha_0, \beta_0) \tag{3}$$

Together, model (2) and prior (3) represent a series of beta-binomial models with closed-form solutions given by

$$p \mid \text{Diseased} = d, \text{N} = n \sim \text{Beta}(\alpha_0 + d, n - d + \beta_0)$$

$$\text{Se}_{m,t} \mid \text{True Positives}_{m,t} = \text{tp}_{m,t}, \text{Diseased} = d \sim \text{Beta}(\alpha_0 + \text{tp}_{m,t}, d - \text{tp}_{m,t} + \beta_0) \tag{4}$$

$$\text{Sp}_{m,t} \mid \text{True Negatives}_{m,t} = \text{tn}_{m,t}, \text{Non-diseased} = n - d \sim \text{Beta}(\alpha_0 + \text{tn}_{m,t}, n - d - \text{tn}_{m,t} + \beta_0)$$

By default, {bayesDCA} fixes the shared prior parameters to be $\alpha_0 = \beta_0 = 1$, though the user may provide different values – which may be model- and threshold-specific. For binary tests, where sensitivity and specificity are fixed for all thresholds, it may be beneficial to elicitate more informative priors based on previous diagnostic test accuracy studies, for instance. However, prior elicitation in the case of risk prediction models may be substantially more challenging due to the threshold-varying parameters. Since here the unknown quantities are relatively easy to estimate, using vague priors is unlikely to be an issue.

We can jointly sample from the posterior distributions in (4) to compute $\text{NB}_{m,t}$ as defined in equation (1). The {bayesDCA} R package employs Stan for this task, primarily because of Stan's rich model and sampling diagnostics. For instance, if the user specifies some really bad priors, Stan may issue multiple warnings about low effective (posterior) sample size and other convergence issues [2].

1.1   Using external prevalence information for case-control data

The model (2) has a single parameter for the outcome proportion or disease prevalence, p. If we perform Decision Curve Analysis in a case-control study, where the observed outcome proportion is determined by investigators, the observed data will not be suitable to estimate p. Instead, we can use external prevalence information and estimate p. Its posterior distribution is then given by

$$p \mid \text{Diseased} = d_{\text{ext.}}, \text{N} = n_{\text{ext.}} \sim \text{Beta}(\alpha_0 + d_{\text{ext.}}, n_{\text{ext.}} - d_{\text{ext.}} + \beta_0)$$

where $d_{\text{ext.}}$ and $n_{\text{ext.}}$ come from an external prevalence study (or any other source of prevalence information). This differs from common frequentist approaches that simply plug-in a point estimate for p. Here, we propagate uncertainty around p regardless of the source of prevalence information. The {bayesDCA} R package allows the user to provide $d_{\text{ext.}}$ and $n_{\text{ext.}}$, while still using the observed data to estimate the sensitivity and specificity terms in equation (1).

2   Bayesian Decision Curve Analysis for survival outcomes

Decision Curve Analysis has been previously extended to survival (time-to-events) outcomes [3]. Because of right-censoring, we cannot condition on "having the event" or "not having the event". We then use Bayes theorem to rewrite the Net Benefit as

$$\text{NB}_{m,t}^{\tau} = [1 - S(\tau \mid \widehat{r}_{m,\tau} > t)] \cdot \mathbb{P}[\widehat{r}_{m,\tau} > t] - S(\tau \mid \widehat{r}_{m,\tau} > t) \cdot \mathbb{P}[\widehat{r}_{m,\tau} > t] \cdot \text{odds}(t) \tag{5}$$

where $\widehat{r}_{m,\tau}$ is the predicted risk of event from model m at time horizon τ (e.g., the predicted one-year mortality according to model m), $S\left(\tau \mid \widehat{r}_{m,\tau} > t\right)$ is the survival at time τ among patients with a positive prediction (i.e., for which $\widehat{r}_{m,\tau} > t$), and $\mathbb{P}\left[\hat{r}_{m,\tau} > t\right]$ is the probability of a positive prediction.

There are two unknown quantities in (5). The first one, $\mathbb{P}\left[\hat{r}_{m,\tau} > t\right]$, is estimated using another beta-binomial model:

$$\text{Positives}_{m,t}^{\tau} \sim \text{Binomial}(N = n, \gamma_{m,\tau}^{(t)})$$

where $\gamma_{m,\tau}^{(t)} = \mathbb{P}\left[\hat{r}_{m,\tau} > t\right]$. Notice that the outcome in this case is not censored – we have predictions for all $N = n$ patients from each model m.

The second unknown in (5) is $S\left(\tau \mid \widehat{r}_{m,\tau} > t\right)$, the survival among patients with a positive prediction ("positive patients"). In a frequentist framework, we could simply plug-in the Kaplan-Meier point estimates as previously suggested [3]. However, in order to properly propagate uncertainty throughout equation (5), we need a Bayesian model for $S\left(\tau \mid \widehat{r}_{m,\tau} > t\right)$. In the following section, we explain such a model. We employ the Poisson trick for a piecewise-constant hazards model, in a similar approach to the one implemented in the spBayesSurv R package [4]. The main difference is that, here, we don't have any covariates.

2.1   Bayesian estimation of survival among positive patients

To estimate $S\left(\tau \mid \widehat{r}_{m,\tau} > t\right)$, first consider partitioning time into K intervals:

$$0 = \tau_0 < \tau_1 < \tau_2 < \cdots < \tau_{K-1} < \tau_K = \infty \tag{6}$$

Define λ_k as the constant hazard in the kth interval given by $I_k = [\tau_{k-1}, \tau_k)$, for $k = 1, 2, \ldots, K$. Usually λ_k is attached to a linear predictor term $\exp\left(x_i^T \beta\right)$ with characteristics x_i from the ith patient (under a proportional hazards assumption). We don't need that term here since we are only interested in the baseline hazards. Then, the hazard for patient i at time t is modeled as:

$$\lambda_i(t) = \lambda_k \quad \text{for} \quad t \in [\tau_{k-1}, \tau_k)$$

Since (6) is a time partition the above can be further simplified to

$$\lambda_{ik} = \lambda_k \tag{7}$$

Now, suppose we observe the exposure time t_i and death indicator d_i for patients $i = 1, 2, \ldots, n$ (or similarly for any event other than death). The expression (7) only refers to patient- and interval-level information, so we may define corresponding pseudo-observations for each individual at each time interval:

$$d_{ik} = \mathbb{1}\left\{t_i \in [\tau_{k-1}, \tau_k) \cap d_i = 1\right\} = \begin{cases} 1 & \text{if patient } i \text{ died in the } k^{th} \text{ interval } I_k \\ 0 & \text{if patient } i \text{ died in some other interval or was censored} \end{cases} \tag{8}$$

$$t_{ik} = \begin{cases} \tau_k - \tau_{k-1} & \text{if } t_i > \tau_k \\ t_i - \tau_{k-1} & \text{if } t_i \in I_k \\ 0 & \text{if } t_i < \tau_{k-1} \end{cases} = \text{exposure time of patient } i \text{ within } k^{th} \text{ interval}$$

Notice that, if a patient i died at time $t \in I_k$, then $d_{ik} = 1$ and $d_{im} = 0$ for all $m \neq k$. If a patient was censored, then $d_{ik} = 0$ for all $k = 1, 2, \ldots, K$. The Poisson trick comes from recognizing that we can treat the interval-specific death indicators d_{ik} as if

$$d_{ik} \sim \text{Poisson}(t_{ik} \cdot \lambda_{ik}) \tag{9}$$

Since we are assuming independence across individuals $i = 1, 2, \ldots, n$, we further simplify (9) to:

$$d_k \sim \text{Poisson}(t_k \cdot \lambda_k) \tag{10}$$

where

$$t_k \cdot \lambda_k = \sum_{i=1}^{n} t_{ik} \cdot \lambda_{ik} = \sum_{i=1}^{n} t_{ik} \cdot \lambda_k = \lambda_k \sum_{i=1}^{n} t_{ik}$$

$$d_k = \sum_{i=1}^{n} d_{ik} = \text{total deaths within } k^{th} \text{ interval}$$

$$t_k = \sum_{i=1}^{n} t_{ik} = \text{total exposure time within } k^{th} \text{ interval}$$

Placing a conjugate Gamma prior[1] $\Gamma(\alpha_k, \beta_k)$ for the parameters λ_k, we have the posterior distributions in closed form:

$$\lambda_k \mid d_k, t_k \ \sim \ \Gamma(\alpha_k + d_k, \beta_k + t_k) \tag{11}$$

With draws from the posterior distributions for the hazard terms, we can then compute the cumulative hazard and survival probability at any time t by:

$$S(t) = \exp(-\Lambda(t))$$

$$\Lambda(t) = \sum_{k=1}^{K} \lambda_k \cdot T_k(t) \tag{12}$$

$$T_k(t) = \begin{cases} \tau_k - \tau_{k-1} & \text{if } t > \tau_k \\ t - \tau_{k-1} & \text{if } t \in I_k \\ 0 & \text{if } t < \tau_{k-1} \end{cases}$$

where $T_k(t)$ is analogous to the interval-specific exposure time as defined in (8).

Finally, to estimate the survival among positive patients, we fit the model defined in (10 —12) using data for patients whose model predictions $\widehat{r}_{m,\tau}$ fall above the decision threshold t.

2.1.1 Choosing cutpoints and prior parameters

References

[1] A. J. Vickers and E. B. Elkin, "Decision curve analysis: A novel method for evaluating prediction models," *Medical Decision Making*, vol. 26, pp. 565–574, nov 2006.

[2] B. Carpenter, A. Gelman, M. D. Hoffman, D. Lee, B. Goodrich, M. Betancourt, M. Brubaker, J. Guo, P. Li, and A. Riddell, "Stan: A probabilistic programming language," *Journal of Statistical Software*, vol. 76, no. 1, p. 1–32, 2017.

[3] A. J. Vickers, A. M. Cronin, E. B. Elkin, and M. Gonen, "Extensions to decision curve analysis, a novel method for evaluating diagnostic tests, prediction models and molecular markers," *BMC Medical Informatics and Decision Making*, vol. 8, nov 2008.

[4] H. Zhou, T. Hanson, and J. Zhang, "spbayessurv: Fitting bayesian spatial survival models using r," *Journal of Statistical Software*, vol. 92, no. 9, 2020.

---

[1] α_k and β_k are the shape and rate parameters, respectively.