Biostat 682 Project Proposal

Lulu Shang, Zheng Xu, Daiwei Zhang

Introduction

In this project, we use Bayesian hierarchical modeling on prostate cancer progression. Patients who have been diagnosed with prostate cancer and undergone therapy can be divided into two categories: those whose cancer progress (as measured by the level of prostate-specific antigen (PSA)) 1) didn't change or immediately after therapy (ono-responding), or 2) decreased for a latent period after the therapy then perhaps increased again (responding). For the former, we assume that PSA has constant change rate. For the latter, we assume that PSA has piecewise linear pattern with knot at the unknown phase-change point. Our main interest is to see given observations 1) the probability of a patient is responding to the treatment (phase-change exists), or is not, 2) the covariance structure of baseline level, decreasing rate, phase-change time and growth rate.

Model formulation

The model described above is formulated as

$$Y_{0ij} = a_i + b_{1i} t_{ij} + \varepsilon_{ij}, \; for \; z_i = 0$$

$$Y_{1ij} = a_i + b_i \, t_{ij} + b_{1i}(t_{ij} - r_{i,z_i})_+ + \varepsilon_{ij}, \; for \; z_i = 1$$

Here for subject i at observation j :

- Y_{ij} : PSA level
- t_{ij} : time after the therapy (measured in months)
- z_i : the indicator function that for whether the subject has a latent period ($z_i = 1$) or not ($z_i = 0$)
- r_{i,z_i} : the phase- change point
- a_i : baseline level
- b_i : decreasing rate, restrained to be negative
- b_{1i} : change of growth rate after phase-change point, restrained to be non-negative
- $\varepsilon_{ij} \sim Normal(0, \sigma^2)$

We assign a MVN prior to a_i , b_i, b_{1i} , and r_i with unstructured covariance matrix, in order to investigate how the baseline, phase-change point and growth rate are correlated with each other. Indicator function z is assigned with a Bernoulli prior with hyperparameter p , which is assigned with the uninformative prior Beta(1,1).

Prior:

- $(a_i\, b_i\, b_{1i}\ r_i)^T \sim MVN((\mu_a, \mu_b, \mu_{b_1}\ \mu_r)^T, \Sigma),\ \pi(\mu, \Sigma) \propto |\Sigma|^{-\frac{d+1}{2}},\ d = 4$
    - If the unstructured model has too many parameter to estimate, we focus our modeling on the correlation between latent point and growth rate
      $(b_{1i}\ r_i)^T \sim MVN(\mu_{b_1}, \mu_r)^T, \Sigma),\ \pi(\mu, \Sigma) \propto |\Sigma|^{-\frac{d+1}{2}},\ d = 2$
- $z_i \sim Bernoulli\ (p),\ p \sim Betat(1, 1)$
- $\sigma^2 \sim IG(0.001,\ 1000)$

MCMC

Markov chain Monte Carlo (MCMC) procedure will be used to estimate posterior distributions by repeatedly drawing samples from the full conditional distributions of the parameters for $a_i,\ b_i,\ b_{1i}$. Slicing sampling will be used for r_i. For z_i the reversible-jump procedure will be used to jump between parameter spaces.

Data

We use a dataset combined from two multi-center clinical trials that includes 485 patients with metastatic prostate cancer. An experimental treatment agent was compared with another drug. The sizes of the treatment and control groups are 249 and 236, respectively. PSA is measured weekly or monthly until death or treatment termination. The plots below summarize the dataset we will use.

References

Zhao, L., & Banerjee, M. (2012). Bayesian piecewise mixture model for racial disparity in prostate cancer progression. *Computational statistics & data analysis, 56*(2), 362-369.

Zhao, L., Feng, D., Neelon, B., & Buyse, M. (2015). Evaluation of treatment efficacy using a Bayesian mixture piecewise linear model of longitudinal biomarkers. *Statistics in medicine*, *34*(10), 1733-1746.

Appendix: Dataset Plots