Multivariate time series analysis of ICU mortality

Student: Danning He, DHSI, JHSOM
Mentors: Dr. Jim Fackler, Dr. Harold Lehmann

1. Project Summary

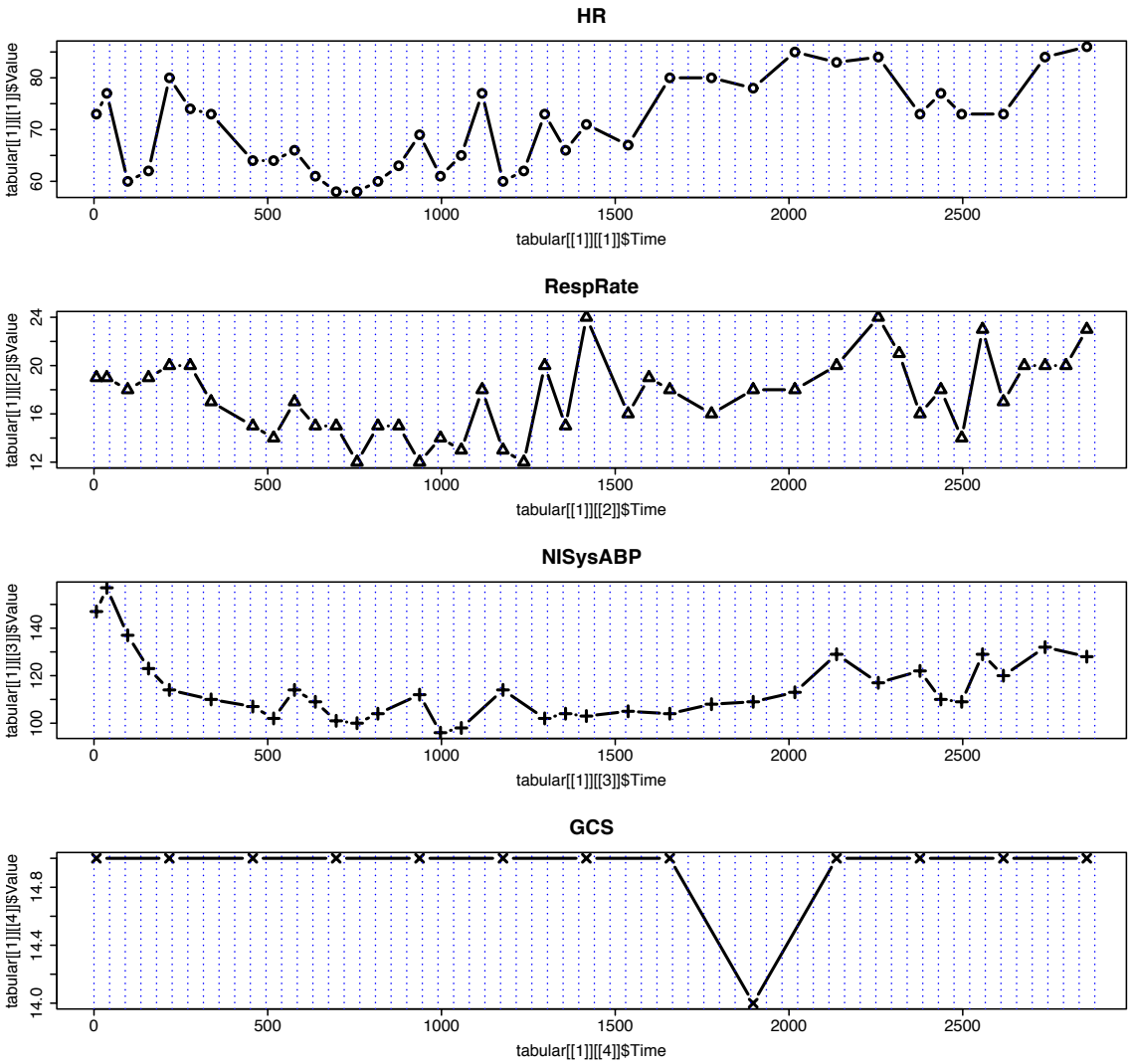
The modern ICU is a complex, expensive and data-intensive environments, acuity assessment of the patients are based on multiple temporal observations and trends produced by various monitoring systems and laboratory tests. Tools that quickly and automatically interpret patterns in the data can greatly facilitate the clinical decision-making by physicians and healthcare givers.

The goal of this project is to develop methods for patient-specific prediction of in-hospital morality/survival based on 5 general descriptors collected on admission and 37 time-series physiological variables during the first 48 hours of an ICU stay.

2. Background Overview

I) Physiological time series

There's rich content in time series modeling using either the time domain approach to predict future value of a series as a parametric function of current and past observations, or the frequency domain approach to characterize periodic variations of interest. Real-world processes produce series of measureable observations as a function of underlying hidden states. Similar to clinical diagnosis, which is inferred from several observations with significant degree of uncertainty, generation of multivariate physiologic profiles by latent disease status or signatures can help reveal the manifestation of disease. For the task of feature discovery or latent signature detection in univariate, continuous time series, various unsupervised learning approaches are available [1,2], with the underlying assumption that there's a fixed set of disease topics common to the collection of time series (such as physiologic heart rate, HR) distributed among patient samples. And the disease topic is again a distribution over the vocabulary of all 'words' in the corpus. The topic proportions can be used as features for explanatory grading task. However, in many cases, a common difficulty from multivariate physiological data is its irregular measurement in terms of time and frequency from patient to patient (the variables are measured from once 30 minutes to once several hours, and not all of them are taken for each individual). Figure 1 show 4 out of 37 physiological variables extracted from a patient in the data set. Therefore, for the task of outcome prediction in ICU, majority of existing acuity models and severity scoring systems are based on such supervised algorithms as logistic regression or artificial neural networks, trained with static variables on admission [3], sequential assessment of organ dysfunctions [4], daily adverse events [5], 24h acuity score [6] and log odds ratio [7].



II) Logistic regression model

To solve the problem of two-class classification, the probability of class 'mortality' (C_1) given original variables (X) can be written as a logistic sigmoid acting on the linear combination of the feature vector $\phi = \phi(X)$ so that

$$p(C_1 \mid X) = \sigma\left(-\mathbf{w}^T\phi\right) = \frac{1}{1+\exp\left(-\mathbf{w}^T\phi\right)},$$

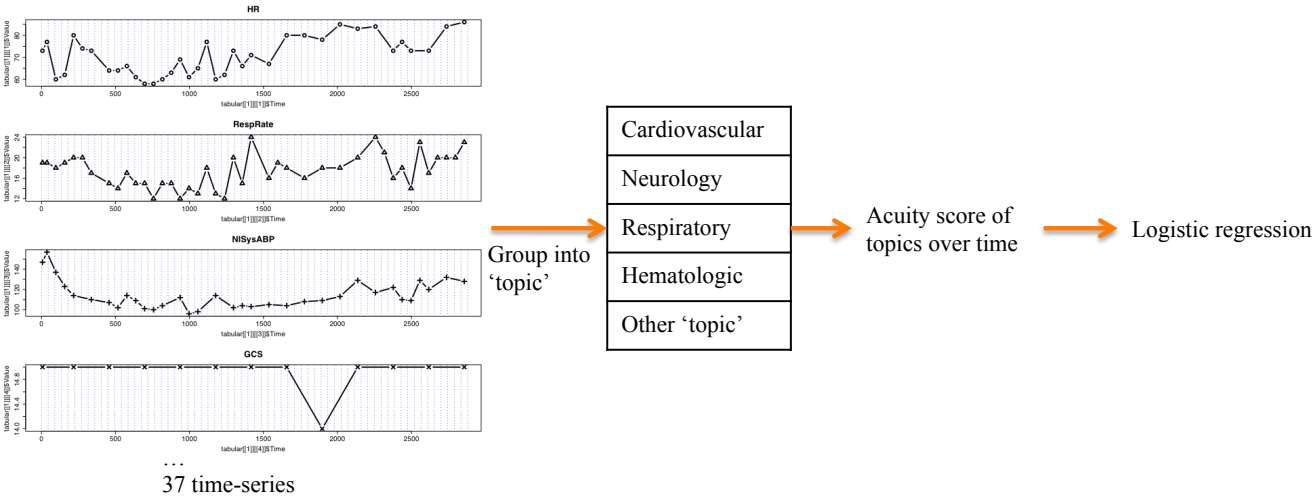
where $\phi(\cdot)$ is the basis function that transforms original variables X into feature vectors, $\mathbf{w} = \left(w_1, w_2, ..., w_M\right)^T$ is model parameter, M is the total number of features. $\mathbf{w}$ is solved via minimizing the error function or maximizing the likelihood function.

3. Technical approach

I) General framework

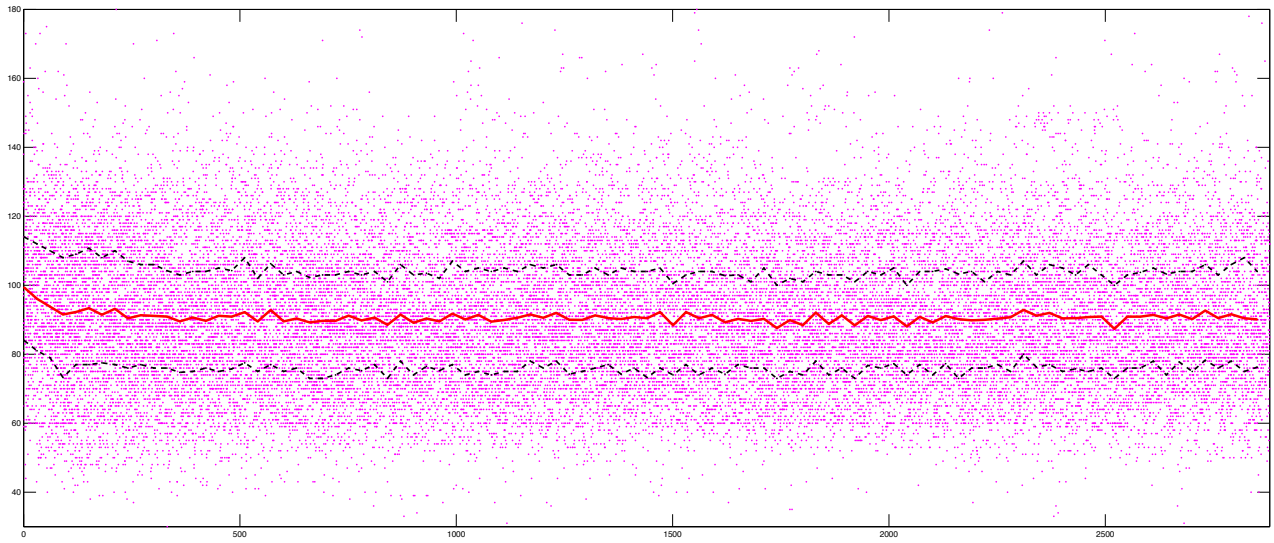
The 37 time series (although only 1/3 variables have more than 10 measurements within the first 48 hour) can be roughly grouped into several relatively independent topics: Neurologic (such as the GCS), Cardiovascular (such as Heart Rate, systolic blood pressure),

Respiratory (such as O_2 and CO_2), Hematologic (such as WBC) and etc. And it's possible to use PCA (principal component analysis) to identify the most representative variations over time, which is a linear combination of several correlated observations within a topic. The 1st principal component from each topic will be used to construct predictive features. In later work, the topics can also be conceptualized as a fixed set of hidden variables, which control what we've observed.

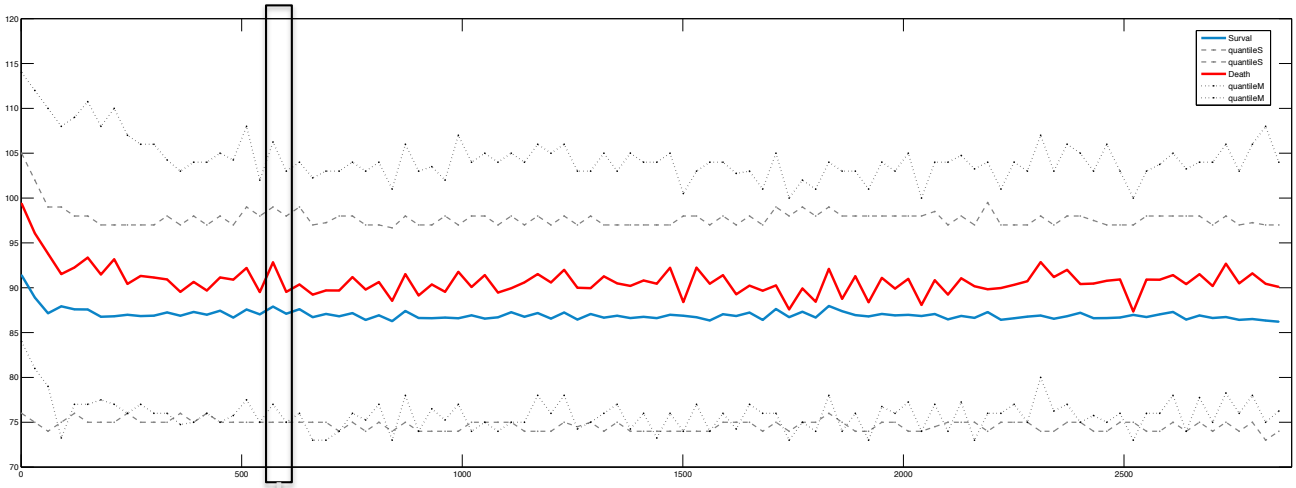


II) Log odds ratio as risk features

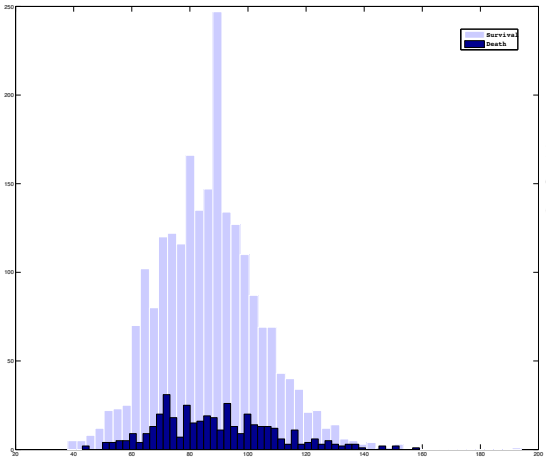
Although for each patient, physiological variables are recorded rather sparsely (Figure 1), given large sample size (~4000), the population behaves in a nearly continuous manner (Figure 2, solid red line represent population mean at a time slice, dashed black line represent 75% and 25% quantile at that time slice. Therefore, we get a time-dependent 'envelope' for this particular physiological variable).



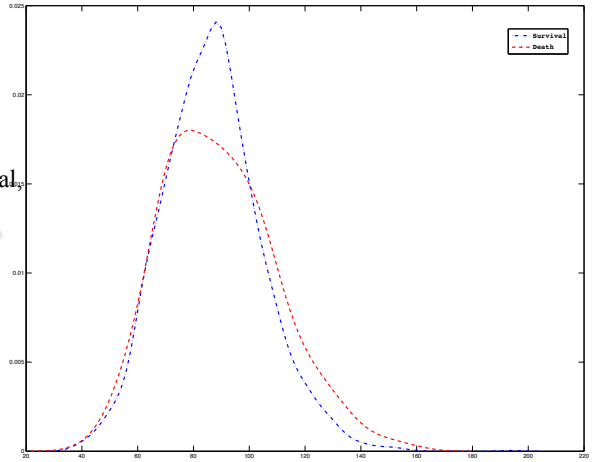
At each time slice, we can fit the distribution of observed values with parametric models in case and control separately (Figure 3), and get the log odds ratio defined as $\log\left(\frac{p(v_i \mid C_1)}{p(v_i \mid C_2)}\right)$.



For each risk factor: A time slice



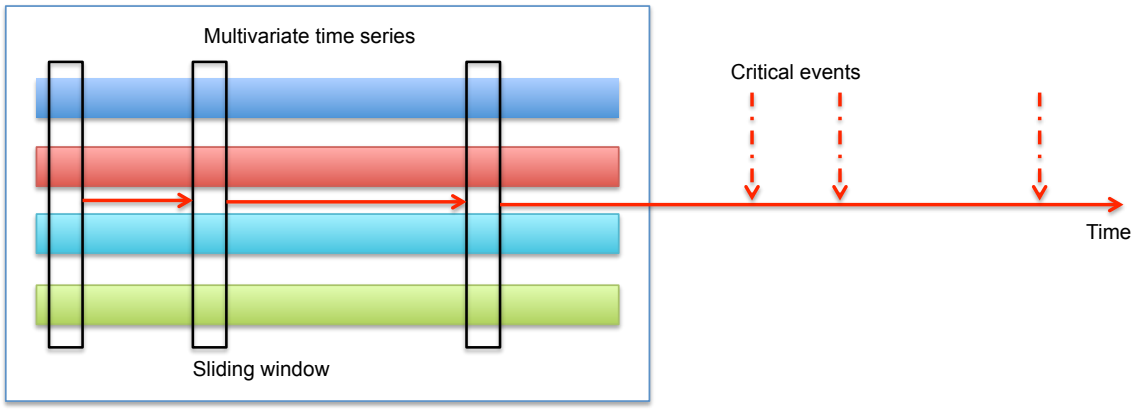
fit parametric distribution: normal, exponential…??



The above figures are just for illustrative purpose. In fact from the bottom right figure, there's not significant difference in case vs. control at this randomly selected time slice. In real work, it's the 1st principal component from each topic that will be used to construct features.

III) Discussion

At each time slice t_i, we observe a vector of observations $\mathbf{o}_{t_i}$. When the sliding window moves along the dimension of time, we have a sequence of observations $\mathbf{O}_j$ for patient j. A useful underlying assumption is that the observed physiological measurements were governed by hidden states, which evolve smoothly or jump over time. Another intriguing fact to consider is that critical events happens days or weeks after 48 hours' of observation. Associating the waiting time until critical events with patient outcome add to the practical usefulness the model. Approaches for this part will depend on the result from previous steps, and will be updated when progress is made. We've used the intuitive log odds ratio as predictive features, so a possible direction would be to consider the overall proportion of hidden states or the distribution of hidden states over time as predictive features.



4. Deliverables

Minimum

-- Logistic regression with log odds ratios as risk features

-- Performance evaluation: ROC, AUC

Expected

-- Minimum deliverables

-- Incorporating waiting time until the critical events

-- Try features constructed from standard HMM, Kalman Filter

Maximum

-- Expected deliverables

-- Optimize features to achieve better classification performance

5. Management Plan

--Regular weekly meeting/consult with Dr. Fackler or Dr. Lehmann

--Frequent consult with related experts when necessary

--Update wiki pages regularly at weekends, documentation of the work done in the past week and the work that will be done in the following week

--Report progress regularly to Dr. Fackler and Dr. Lehmann

6. Dependencies

--Confirm regular weekly meeting/consult with Dr. Fackler and Dr. Lehmann

--Data availability (resolved)

7. Timeline

Timeline		week 1	week 2	week 3	week 4	week 5	week 6	week 7	week 8	week 9	week 10	week 11	week 12
Milestone 1	reading list												
	project plan												
	preprocessing data												
Milestone 2 (Minimum)	Features as log odds ratio												
	Logistic regression												
	AUC and ROC												
Milestone 3 (Expected)	try HMM												
	try Kalman Filter												
	Optimization												
Milestone 4	Model comparison												
	Project report												

8. References

The list is by no means exhaustive, and it'll be updated when progress is made.

1. Saria S, Koller D, Penn AA. (2010) Learning individual and population level traits from Clinical Temporal data. Neural Information Processing Systems.
2. Imhoff M, Kuhls S. (2006) Alarm algorithms in critical care monitoring. Anesth Analg 102: 1525-1537.
3. Zimmerman JE, Kramer AA, McNair DS, Malila FM. (2006) Acute Physiology and Chronic Health Evaluation (APACHE) IV: hospital mortality assessment for today's critically ill patients. Crit Care Med 34: 1297-1310.
4. Ferreira FL, Bota DP, Bross A, Mélot C, Vincent JL. (2001) Serial Evaluation of the SOFA Score to Predict Outcome in Critically Ill Patients. JAMA 286: 1754-1758.
5. Silva A, Cortez P, Santos MF, Gomes L, Neves J. (2006) Mortality assessment in intensive care units via adverse events using artificial neural networks. Artif Intell Med 36: 223-234.
6. Hug CW, Szolovits P. (2009) ICU Acuity: Real-time Models versus Daily Models. AMIA Annu Symp Proc: 260-264.
7. Saria S, Rajani AK, Gould J, Koller D, Penn AA. (2010) Integration of Early Physiological Responses Predicts Later Illness Severity in Preterm Infants. Sci Transl Med 2: 48ra65.