

# Variational Inference for Dirichlet Process to Stratify Cancer Patients Using DNA Methylation

KEVIN YANG, XIAOXUAN LIANG

## 1 Abstract

Variational Inference for stratifying cancer patients based on DNA Methylation data from ICGC.

## 2 Background

### 2.1 Variational Inference

Optimization method of ELBO using variational parameters.

### 2.2 Dirichlet Process

Bayesian non-parametric method.

#### 2.2.1 DNA Methylation

Epigenetic marker that is correlated with cancer.

## 3 Methods

### 3.1 Probabilistic Programming Language

Pyro, PyProb, STAN?

### 3.2 VI for Dirichlet Process

Dirichlet Process is slow for high dimensional data. DNA methylation is high dimensional because there are many genes, so use VI.

## 4 Discussion

This project is cool and we want you to accept the proposal.

## References

- [1] Alqallaf, Fatemah, and Paul Gustafson. *On Cross-Validation of Bayesian Models*. Canadian Journal of Statistics, vol. 29, no. 2, June 2001, pp. 333–340, 10.2307/3316081. Accessed 16 Mar. 2020.