

Discussion of: “History and Future of Bayesian Statistics”

Woncheol Jang

^aDepartment of Statistics, Seoul National University

먼저 베이저안 통계학의 대한 역사와 미래에 대해 살펴볼 수 있는 계기를 준 저자들에게 경의를 표하고 싶다. 본 토론에서는 저자들이 미처 다루지 못한 부분들. 특히 베이저안 비모수 방법에 대한 간략한 소개를 하고자 한다.

최근 들어 기계학습분야에서 베이저안 비모수 방법에 대한 연구가 활발히 진행되고 있다. 기계학습은 통계학과 전산학의 접목분야로 생각할 수 있으며 최근 빅데이터의 등장으로 각광을 받고 있다. 사실 국내 다수의 통계학과들이 전산통계, 또는 계산통계라는 학과로 시작하였으나 1990년대에 오면서 전산학과 통계학으로 분리되었다. 아이러니하게 전산학과 통계학은 2000년대에 들어와서 기계학습이라는 공통분야를 통해 학문적으로는 더욱 가까워지고 있다. 기계학습분야에서 특히 주목받고 있는 통계적 방법론은 베이저안 비모수 방법들이다. 예를 들면 텍스트마이닝 (text mining)분야에서 각광받고 있는 Latent Dirichlet Allocation (Blei, 2003)의 경우 인용 회수가 10,000번에 육박할 정도로 굉장히 인기 있는 통계적 분석방법이다. 기존의 베이저안 (모수) 방법들과 비교했을 때 베이저안 비모수 통계모형의 차이점은 자료의 갯수가 증가할 경우 모형의 복잡도도 같이 증가한다는 데 있다. 일반적으로 베이저안 비모수 모형에서는 모수공간들을 stochastic process를 이용하여 표현하는데 베이저안 비모수방법들은 이러한 stochastic process의 이름들로 불려지곤 한다. 이러한 대표적 베이저안 비모수 방법으로는 회귀 모형에서는 Gaussian process regression, 군집분석에서 Dirichlet process mixture model 등을 들 수 있다. 여기서는 군집모형에서 베이저안 비모수 모형이 어떻게 사용되는지 간략히 설명하고자 한다.

혼합모형은 군집분석에 많이 사용되는 통계적 방법들 중 하나로 다음과 같이 표현할 수 있다.

$$p(x) = \sum_{k=1}^K \pi_k p(x | \theta_k)$$

혼합모형 사용시 가장 큰 문제는 혼합의 갯수 K 를 정하는 것이라고 할 수 있다. 위의 모형은 적분을 이용하여 다음과 같이 나타낼 수 있다.

$$p(x) = \int p(x | \theta) G(\theta) d\theta, \quad G = \sum_{k=1}^K \pi_k \delta_{\theta_k}$$

베이저안 비모수 모형의 경우 $G = \sum_{k=1}^{\infty} \pi_k \delta_{\theta_k}$ 라 가정하고 G 의 사전분포로 흔히 Dirichlet process를 사용한다. 이 경우 $K = \infty$ 이기 때문에 Stick-Breaking을 사용하여 무한개의 혼합분포로 확장할 수 있다 (Dunson, 2010). 뿐만 아니라 이 경우 θ 의 주변분포는 Chinese restaurant process가 됨을 보일 수 있고 이 경우 새로운 자료가 관측될 경우 새 군집을 추가할 수 있다는 점에서 기존의 군집방법과 차이가 있다고 할 수 있다.

이 외에도 베이저안 비모수 방법은 잠재요인 모형(Latent factor model)등 다양한 방법과 많은 응용분야에서 널리 사용되고 있다. 관심이 있는 독자들은 Dunson (2010) 과 Müller and Quintana (2004) 를 참조하기를 권한다.

References

- Blei et al.(2003). Latent Dirichlet Allocation. *Journal of Machine Learning*. **3** 993–1022.
- Dunson(2010). Nonparametric Bayes Applications to Biostatistics. In *Bayesian Nonparametrics*. ed. Hjort, N. L., Homes, C., Müller, P. and Walker, S. G. Cambridge University Press.
- Gershman, S. J. and Blei, D. M. (2012). A tutorial on Bayesian nonparametric models. *Journal of Mathematical Psychology*. **56** 1–12.
- Müller, P. and Quintana, F. A. (2004). Nonparametric Bayesian Data Analysis. *Statistical Science*. **19** 95–110.
- Orbanz, P. and Teh, Y. W. (2010). Bayesian Nonparametric Models. *Encyclopedia of Machine Learning*. 81–89.