Communications of the Korean Statistical Society
2011, Vol. 18, No. 3, 295–300

DOI: 10.5351/CKSS.2011.18.3.295

Bayesian Estimation of Uniformly Stochastically Ordered Distributions with Square Loss

Myongsik Oh[1,a]

[a]Department of Statistics, Pusan University of Foreign Studies

Abstract

The Bayesian nonparametric estimation of two uniformly stochastically ordered distributions is studied. We propose a restricted Dirichlet Process. Among many types of restriction we consider only uniformly stochastic ordering in this paper since the computation of integrals is relatively easy. An explicit expression of the posterior distribution is given. When square loss function is used the posterior distribution can be obtained by easy integration using some computer program such as Mathematica.

Keywords: Bayesian estimation, Dirichlet Process, square loss, uniform stochastic ordering.

1. Introduction

Stochastic ordering has been widely used to explain and solve many statistical problems such as association of random variables as well as comparing two or more survival functions. Although the three types of stochastic ordering (usual stochastic ordering, uniform stochastic ordering and likelihood ratio ordering) have been widely used, the estimation of distribution functions under stochastic ordering constraint is not easy except for the uniform stochastic ordering case. Note that likelihood ratio ordering is the strongest and the usual stochastic ordering is the weakest among three types of stochastic ordering.

The X distribution is said to be uniformly stochastically larger than the Y distribution if

$$\frac{1 - F(x)}{1 - G(x)} \text{ is nondecreasing for } x \text{ in } \left(-\infty, G^{-1}(x)\right),$$

where F and G are cdf's of X and Y, respectively. The conditional distributions (given that the random variables are at least of a certain size) are all stochastically ordered (in the usual sense) in the same direction.

This type of ordering is certainly of interest when populations correspond to survival times for different medical treatments. Even if the better of two treatments (better in the sense that its survival time stochastically dominates that of the other treatment) is administered initially, it may not be the better treatment when patients are examined at a later point in time. However, there is no doubt which treatment is preferred at any point in time if the treatment populations are ordered in this stronger sense.

Statistical inference including estimation and testing under uniform stochastic ordering has received a substantial amount of interest recently. Dykstra *et al.* (1991) studied nonparametric maximum likelihood estimation and testing for discrete population when right-censored are present. Rojo

[1] Professor, Department of Statistics, Pusan University of Foreign Studies, 15 SukPo-Ro, Nam-Gu, Pusan 608-738, Korea.
E-mail: moh@pufs.ac.kr

and Samaniego (1991, 1993) show that nonparametric maximum likelihood estimators do not have the consistence property in either one or two sample cases. They proposed an estimator that is consistent but lacks some desirable properties. Mukerjee (1996) proposed a one-parameter family of estimator. On the other hand, limited research has emerged from the Bayesian point of view. Dykstra and Laud (1981) studied the Bayesian estimation problem arising in reliability. Proschan and Singpurwalla (1980) proposed a Bayesian estimation procedure which uses a pooling adjacent violators algorithm(PAVA).

Bayesian nonparametric estimation has been widely used since the seminal papers by Ferguson (1973) and Doksum (1974), as well as Antoniak (1974). Ferguson introduced the Dirichlet process as a prior distribution which became a main tool in Bayesian nonparametric analysis. Ferguson and others studied many properties of the Dirichlet process. Antoniak (1974) studied the properties of a mixture of Dirichlet processes. Another important property was discovered by Doksum, which is known as "neutral to the right" that enables an analysis of incomplete data, especially censored data. Dykstra and Laud (1981) studied Bayesian nonparametric estimation of a hazard rate function when it is assumed to be increasing. The so-called Gamma process is used as a prior distribution that is not neutral to the right.

Recently Bayesian nonparametric estimation under stochastic ordering has been of interest among many researchers. Dunson and Peddada (2008) proposed the classes of a restricted dependent Dirichlet process prior for Bayesian nonparametric inference on stochastic ordering. Evans *et al.* (1997) studied Bayesian analysis of stochastically ordered distributions of categorical variables. Hoff (2003) studied Bayesian methods for partial stochastic orderings. More recently, Karabatsos and Walker (2009) studied Bayesian nonparametric inference of stochastically ordered distributions.

To our knowledge no Bayesian nonparametric inference under uniform stochastic ordering have been studied yet. In this paper we introduce a restricted Dirichlet process which will serve as a prior distribution for the Bayesian nonparametric estimation of the distribution functions when restriction of uniform stochastic ordering between two functions is imposed. This approach has a significant advantage over the existing Bayesian approaches. In most Bayesian inference cases we use Gibbs sampler or other numerical computation methods (such as the Metropolis-Hasting algorithm) to obtain the results. However, in this approach we do not need any numerical methods but need only the evaluation of multiple integrations analytically or numerically. In Section 3, Bayesian nonparametric estimator under a square loss function is studied. The estimators are expressed in terms of the ratios of two double integrals that can be approximated easily. In Section 4, we briefly discuss future research topics then include handling censored data.

2. Restricted Dirichlet Process and Posterior Distribution

Let $\alpha_i(\cdot)$ be non-null finitely additive measure on $(R^+, \mathcal{B})$ where $R^+ = (0, \infty)$ and $\mathcal{B}$ is the Borel σ-field on R^+. Consider a random measure $(P_1, P_2, \ldots, P_g)$. Suppose for every $k = 1, 2, \ldots,$ and for every measurable partition $B_1, B_2, \ldots, B_k$ of R^+ the probability density function of

$$\Big(P_1(B_1), P_1(B_2), \ldots, P_1(B_k), \ldots, P_g(B_1), P_g(B_2), \ldots, P_g(B_k)\Big)$$

is defined as follows; let $p_{ij} = P_i(B_j)$, $\mathbf{p}_i = (p_{i1}, p_{i2}, \ldots, p_{ik})$, and $\alpha_{ij} = \alpha_i(B_j)$ for $i = 1, \ldots, g$, then

$$f(\mathbf{p}_1, \ldots, \mathbf{p}_g | \alpha_1, \ldots, \alpha_g) = \begin{cases} K \cdot \prod_{i=1}^{g} \prod_{j=1}^{k} p_{ij}^{\alpha_{ij}-1}, & \text{if } (\mathbf{p}_1, \ldots, \mathbf{p}_g) \in \Big\{(\mathbf{p}_1, \ldots, \mathbf{p}_g) \in \mathbf{R}^{gk} : \mathbf{p}_1 \preceq \mathbf{p}_2 \preceq \cdots \preceq \mathbf{p}_g\Big\}, \\ 0, & \text{otherwise}, \end{cases}$$

where $\preceq$ is a quasi order, which requires *reflexive* and *transitive*, defined on an index set $\{1, 2, \ldots, g\}$ and K is determined so that $f(\mathbf{p}_1, \ldots, \mathbf{p}_g | \alpha_1, \ldots, \alpha_g)$ be a joint probability density function. Then we say that the random measure $(P_1, P_2, \ldots, P_g)$ form a restricted Dirichlet process on $(R^+, \mathcal{B})$ with parameters $\alpha_1, \alpha_2, \ldots, \alpha_g$.

Many types of restrictions may be associated to each corresponding quasi order or more specifically a partial order that requires *antisymmetry* in addition to quasi order. Among such restrictions stochastic ordering has received considerable interests. There are several types of stochastic orderings including stochastic ordering of usual sense, uniform stochastic ordering, and likelihood ratio ordering. We mainly focus on uniform stochastic ordering next.

To avoid the notational complexity we consider only the case of $g = 2$. The extension to the case of general g is quite straightforward. Now we determine K so that $f(\mathbf{p}_1, \mathbf{p}_2 | \alpha_1, \alpha_2)$ be a density function under uniform stochastic ordering. It is not difficult to show that uniform stochastic ordering between two multinomial distributions $\mathbf{p}_1$ and $\mathbf{p}_2$, for example, $\mathbf{p}_1 \preceq \mathbf{p}_2$, can be expressed as

$$\frac{\sum_{j=i}^k p_{2j}}{\sum_{j=i}^k p_{1j}} \text{ is nondecrasing in } i$$

which is equivalent to

$$\frac{\sum_{j=i}^k p_{2j}}{\sum_{j=i}^k p_{1j}} \leq \frac{\sum_{j=i+1}^k p_{2j}}{\sum_{j=i+1}^k p_{1j}}, \quad \text{for } i = 1, \ldots, k-1.$$

Then the normalizing constant K must satisfy

$$K^{-1} = \int_A \prod_{j=1}^k p_{1j}^{\alpha_{1j}-1} p_{2j}^{\alpha_{2j}-1} d(\mathbf{p}_1, \mathbf{p}_2),$$

where

$$A = \left\{ (\mathbf{p}_1, \mathbf{p}_2) \in \mathbf{R}^{2k} : \frac{\sum_{j=i}^k p_{2j}}{\sum_{j=i}^k p_{1j}} \leq \frac{\sum_{j=i+1}^k p_{2j}}{\sum_{j=i+1}^k p_{1j}}, \text{ for } i = 1, \ldots, k-1 \right\}.$$

The integration seems to be intractable. We can simplify the integration by introducing a one-to-one transformation of variables. Let $u_i = \sum_{\ell=i+1}^k p_{1\ell} / \sum_{\ell=i}^k p_{1\ell}$, $v_i = \sum_{\ell=i+1}^k p_{2\ell} / \sum_{\ell=i}^k p_{2\ell}$. Then we have

$$p_{11} = 1 - u_1, \quad p_{1i} = (1 - u_i) \prod_{\ell=1}^{i-1} u_\ell, \quad \text{for } i = 1, \ldots, k-1, \; p_{1k} = \prod_{\ell=1}^{k-1} u_\ell,$$
$$p_{21} = 1 - v_1, \quad p_{2i} = (1 - v_i) \prod_{\ell=1}^{i-1} v_\ell, \quad \text{for } i = 1, \ldots, k-1, \; p_{2k} = \prod_{\ell=1}^{k-1} v_\ell$$

and

$$|J| = \left| \frac{\partial(\mathbf{p}_1, \mathbf{p}_2)}{\partial(\mathbf{u}, \mathbf{v})} \right| = \prod_{i=1}^{k-2} u_i^{k-1-i} v_i^{k-1-i}.$$

By some algebra we can show that

$$K^{-1} = \prod_{i=1}^{k-1} \iint_{0\le u_i \le v_i \le 1} u_i^{\sum_{\ell=i+1}^{k} \alpha_{1\ell}-1}(1-u_i)^{\alpha_{1i}-1} v_i^{\sum_{\ell=i+1}^{k} \alpha_{2\ell}-1}(1-v_i)^{\alpha_{2i}-1} du_i dv_i.$$

The integration is greatly simplified; however, we still need to evaluate the double integral that cannot be obtained analytically. This will be discussed later.

Next we discuss about the posterior distribution given data. Let $X_1, \ldots, X_m$ and $Y_1, \ldots, Y_n$ be two independent random samples from two populations. We assume that all observations are complete, that is, there are no censored observations. Let $\delta_x(B) = 1$ if $x \in B$, zero otherwise. Then

$$\left(\sum_{j=1}^{m} \delta_{X_j}(B_1), \ldots, \sum_{j=1}^{m} \delta_{X_j}(B_k)\right) \quad \text{and} \quad \left(\sum_{j=1}^{n} \delta_{Y_j}(B_1), \ldots, \sum_{j=1}^{n} \delta_{Y_j}(B_k)\right)$$

have multinomial distributions with parameters $(m, p_1, \ldots, p_k)$ and $(n, q_1, \ldots, q_k)$, respectively. Following the similar steps in the proof of Theorem 1 of Ferguson (1973) we can easily show that the posterior distribution of $(\mathbf{p}_1, \mathbf{p}_2)$ is also a restricted Dirichlet process. The probability density function is, for $(\mathbf{p}_1, \mathbf{p}_2) \in A$,

$$f(\mathbf{p}_1, \mathbf{p}_2) = K' \cdot \prod_{i=1}^{k} p_{1i}^{\alpha_{1i}+\sum_{j=1}^{m} \delta_{X_j}(B_i)-1} p_{2i}^{\alpha_{2i}+\sum_{j=1}^{n} \delta_{Y_j}(B_i)-1},$$

where

$$\begin{aligned}(K')^{-1} = \prod_{i=1}^{k-1} \iint_{0\le u_i \le v_i \le 1} & u_i^{\sum_{j=i+1}^{k}\left(\alpha_{1j}+\sum_{\ell=1}^{m} \delta_{X_\ell}(B_j)\right)-1}(1-u_i)^{\alpha_{1i}+\sum_{j=1}^{m} \delta_{X_j}(B_i)-1} \\ & \times\ v_i^{\sum_{j=i+1}^{k}\left(\alpha_{2j}+\sum_{\ell=1}^{n} \delta_{Y_\ell}(B_j)\right)-1}(1-v_i)^{\alpha_{2i}+\sum_{j=1}^{n} \delta_{Y_j}(B_i)-1} du_i dv_i.\end{aligned}$$

3. Estimation under Square Loss

Consider the space $(\mathbf{R}, \mathcal{B})$, where $\mathbf{R}$ is the real line and $\mathcal{B}$ is the associated σ-field of Borel sets. Let $X_1, X_2, \ldots, X_m$ be *iid* with distribution function F and $Y_1, Y_2, \ldots, Y_n$ be *iid* with distribution function G. We are interested in finding F and G subject to uniform stochastic ordering. Let Ω be the space of all distribution functions and the loss function be

$$L\Big((F,G), \big(\hat{F}, \hat{G}\big)\Big) = \int_{-\infty}^{\infty} \left\{\left(F(t) - \hat{F}(t)\right)^2 + \left(G(t) - \hat{G}(t)\right)^2\right\} dw(t),$$

where w is a given finite measure on $(\mathbf{R}, \mathcal{B})$,

$$F(t) = P_F[(-\infty, t]], \qquad G(t) = P_G[(-\infty, t]]$$

and $\hat{F}(t)$ and $\hat{G}(t)$ are restricted estimates of $F(t)$ and $G(t)$, respectively.

Suppose $B_1 = (-\infty, t]$ and $B_2 = (t, \infty)$. Then it is well known that $\hat{F}(t) = E(P(B_1))$ and $\hat{G}(t) = E(Q(B_1))$. More specifically

$$\hat{F}(t) = \frac{K(a_2, a_1, b_2, b_1)}{K(a_2, a_1+1, b_2, b_1)}, \qquad \hat{G}(t) = \frac{K(a_2, a_1, b_2, b_1)}{K(a_2, a_1, b_2, b_1+1)},$$

where, for $i = 1, 2$,

$$a_i = \alpha_{1i} + \sum_{j=1}^{m} \delta_{X_j}(B_i) - 1,$$

$$b_i = \alpha_{2i} + \sum_{j=1}^{n} \delta_{Y_j}(B_i) - 1,$$

$$K^{-1}(a, b, c, d) = \iint_{0 \le u \le v \le 1} u^a (1-u)^b v^c (1-v)^d du dv.$$

It is important to evaluate the double integral. First suppose b is an integer. Using binomial formula we have $(1-u)^b = \sum_{i=0}^{b} \binom{b}{i} (-1)^{b-i} u^{b-i}$ and

$$\iint_{0 \le u \le v \le 1} u^a (1-u)^b v^c (1-v)^d du dv = \sum_{i=0}^{b} \binom{b}{i} (-1)^{b-i} \int_0^1 \left[\int_0^v u^{a+b-i} du \right] v^c (1-v)^d dv$$

$$= \sum_{i=0}^{b} \binom{b}{i} (-1)^{b-i} \frac{1}{a+b-i+1} \int_0^1 v^{a+b-i+1+c} (1-v)^d dv$$

and hence

$$K^{-1}(a, b, c, d) = \sum_{i=0}^{b} \frac{\binom{b}{i}(-1)^{b-i}}{a+b-i+1} \frac{\Gamma(a+b+c-i+2)\Gamma(d+1)}{\Gamma(a+b+c+d-i+3)},$$

where $\Gamma(\cdot)$ is a gamma function.

Next suppose b is not an integer. From a well-known formula

$$\int_0^v u^a (1-u)^b du = v^{a+1} \sum_{k=0}^{\infty} \frac{(-b)_k}{k!(a+1+k)} v^k,$$

where $(a)_k = a(a-1)\cdots(a-k+1)$ with convention $0/0 = 0$. By similar manner and some algebra we have the following;

$$K^{-1}(a, b, c, d) = \frac{\Gamma(a+b+2)\Gamma(d+1)}{(1+a)\Gamma(a+c+d+3)} \sum_{k=0}^{\infty} \frac{(1+a)_k (-b)_k (2+a+c)_k}{k!(2+a)_k (3+a+c+d)_k}.$$

On the other hand, we may use some computational package such as *Mathematica*.

4. Further Research

We have discussed about the Bayesian nonparametric estimation of two distribution functions under uniform stochastic ordering with square loss. In estimating two distributions we only use complete data or uncensored data. In a real problem, such as the survival analysis of cancer patients, we need to be able to handle incomplete data, *i.e.*, censored one. We note that the Dirichlet process is "neutral to the right" that enables the estimation of distribution functions with censored data, refer to Ferguson

(1973) and Doksum (1974). We are, however, not aware that the restricted Dirichlet process has the same property; however, we may be able to estimate distribution functions under square loss. For estimation problem of a distribution function, refer to Ferguson and Phadia (1979) and Susarla and Van Ryzin (1976).

We can extend this problem to the estimation of g distribution functions under various types of uniform stochastic ordering, for instance, simple order and simple tree order. The solution to those problems may involve a multiple integral like other problems in Bayesian analysis.

References

Antoniak, C. E. (1974). Mixtures of Dirichlet processes with application to Bayesian nonparametric problems, *Annals of Statistics*, **2**, 1152–1174.

Doksum, K. (1974). Tail free and neutral random probabilities and their posterior distributions, *Annals of Probability*, **2**, 183–201.

Dunson, D. B. and Peddada, S. D. (2008). Bayesian nonparametric inference on stochastic ordering, *Biometrika*, **95**, 859–874.

Dykstra, R. L., Kochar, S. C. and Robertson, T. (1991). Statistical inference for uniform stochastic ordering in several populations, *Annals of Statistics*, **19**, 870–880.

Dykstra, R. L. and Laud, P. (1981). A Bayesian nonparametric approach to reliability, *Annals of Statistics*, **9**, 356–367.

Evans, M., Gilula, Z., Guttman, I. and Swartz, T. (1997). Bayesian analysis of stochastically ordered distributions of categorical variables, *Journal of the American Statistical Association*, **92**, 208–214.

Ferguson, T. S. (1973). A Bayesian analysis of some nonparametric problems, *Annals of Statistics*, **1**, 209–230.

Ferguson, T. S. and Phadia, E. G. (1979). Bayesian nonparametric estimation based on censored data, *Annals of Statistics*, **7**, 163–186.

Hoff, P. D. (2003). Bayesian methods for partial stochastic orderings, *Biometrika*, **90**, 203–317.

Karabatsos, G. and Walker, S. G. (2009). Bayesian nonparametric inference of stochastically ordered distributions, with Pólya trees and Bernstein polynomials, *Statistics and Probability Letters*, **77**, 907–913.

Mukerjee, H. (1996). Estimation of survival functions under uniform stochastic ordering, *Journal of the American Statistical Association*, **91**, 1684–1689.

Proschan, F. and Singpurwalla, N. (1980). A new approach to inference from accerlated life tests, *IEEE Transaction on Reliability*, **29**, 98–102.

Rojo, J. and Samaniego, F. J. (1991). On nonparametric maximum likelihood estimation of a distribution uniformly stochastically smaller than a standard, *Statistics and Probability Letters*, **11**, 267–271.

Rojo, J. and Samaniego, F. J. (1993). On estimating a survival curve subject to a uniform stochastic ordering constraint, *Journal of the American Statistical Association*, **88**, 566–572.

Susarla, V. and Van Ryzin, J. (1976). Nonparametric Bayesian estimation of survival curves from incomplete observations, *Journal of the American Statistical Association*, **71**, 897–902.

Received November 2010; Accepted March 2011