Nonparametric Bayesian inference under stochastic order constraints, with an application in epidemiology

Athanasios Kottas
Department of Applied Mathematics and Statistics
Baskin School of Engineering
University of California, Santa Cruz

Econometrics and Statistics Colloquium
The University of Chicago Booth School of Business
February 18, 2010

Outline

1. Introduction and background
2. A mixture modeling approach for stochastic order
3. Application to ROC analysis
4. Modeling under stochastic precedence restrictions
5. Summary

1. Introduction and background

- *Probability order* restrictions often appropriate/desirable when comparing two or more populations — different types of probability orders

- $\mathbb{R}$-valued random variables Y_1, Y_2 with respective distribution functions F_1, F_2, density functions f_1, f_2, and hazard functions h_1, h_2
- **Stochastic order**: $Y_1 \leq_{st} Y_2$ (or $F_1 \leq_{st} F_2$) if, by definition,

$$F_1(u) \geq F_2(u),\ \forall u \in \mathbb{R} \ \Leftrightarrow\ \Pr(Y_1 > u) \leq \Pr(Y_2 > u),\ \forall u \in \mathbb{R}$$

$\rightarrow$ characterization: $Y_1 \leq_{st} Y_2$ if-f there exist r.v.s Y_1' and Y_2', defined on the same probability space, such that Y_1 and Y_2 have the same distribution with Y_1' and Y_2', and $\Pr(Y_1' \leq Y_2') = 1$

- **Hazard rate order**: For $\mathbb{R}^+$-valued r.v.s Y_1, Y_2, $Y_1 \leq_{hr} Y_2$ if, by definition, the function $(1 - F_1(t))/(1 - F_2(t))$ decreases in t
  $\rightarrow$ equivalent definition for continuous r.v.s Y_1, Y_2:

$$Y_1 \leq_{hr} Y_2 \quad \text{if-f} \quad h_1(t) \geq h_2(t), \ \forall t \in \mathbb{R}^+$$

  $\rightarrow$ stronger restriction than stochastic order ($Y_1 \leq_{hr} Y_2$ implies $Y_1 \leq_{st} Y_2$)

- **Stochastic precedence order**: $Y_1 \leq_{sp} Y_2$ (or $F_1 \leq_{sp} F_2$) if, by definition, $\Pr(Y_1 \leq Y_2) \geq 0.5$
  $\rightarrow$ weaker restriction than stochastic order ($Y_1 \leq_{st} Y_2$ implies $Y_1 \leq_{sp} Y_2$)

Introduction and background

- Substantial literature on properties of distributions ordered according to one of these constraints, and on other probability orders (e.g., Shaked & Shanthikumar, 1994)

- Also, extensive literature on classical estimation (typically, maximum likelihood estimation) and distribution-free testing for various types of probability orderings (e.g., Robertson, Wright & Dykstra, 1988)

- Arguments for forcing order restriction in the distribution function estimates:
  → order constraint of interest may not hold for the empirical distribution functions (especially for small or moderate sample sizes)
  → incorporating the order restriction can improve predictive accuracy (Bayesian framework)

Introduction and background

- Bayesian work:
  → Inference for stochastically ordered survival functions (Arjas & Gasbarra, 1996)

  → Stochastic order analysis for categorical variables (Evans, Gilula, Guttman & Swartz, 1997)

  → Stochastic order and partial stochastic orders (Gelfand & Kottas, 2001; Hoff, 2003; Karabatsos & Walker, 2007; Dunson & Peddada, 2008; Hanson, Kottas & Branscum, 2008)

  → Stochastic precedence order (Chen & Dunson, 2004; Kottas, 2009)

  → Variability orders (Kottas & Gelfand, 2001)

Introduction and background

- **Objective:** Bayesian modeling and inference for distributions subject to stochastic order or stochastic precedence constraints
  → working within the Bayesian framework is attractive: any order restriction in the prior model is preserved to the posterior analysis
  → Bayesian *nonparametric* (or *semiparametric*) specifications to avoid restricting the distribution functions to specified parametric forms

- With two distribution functions F_1 and F_2, we seek nonparametric prior models over the space

$$\mathcal{P}_{st} = \{(F_1, F_2) : \ F_1 \leq_{st} F_2\} \quad \text{or} \quad \mathcal{P}_{sp} = \{(F_1, F_2) : \ F_1 \leq_{sp} F_2\}$$

- The modeling approach is developed using Dirichlet process (DP) mixtures (Antoniak, 1974; Escobar & West, 1995)

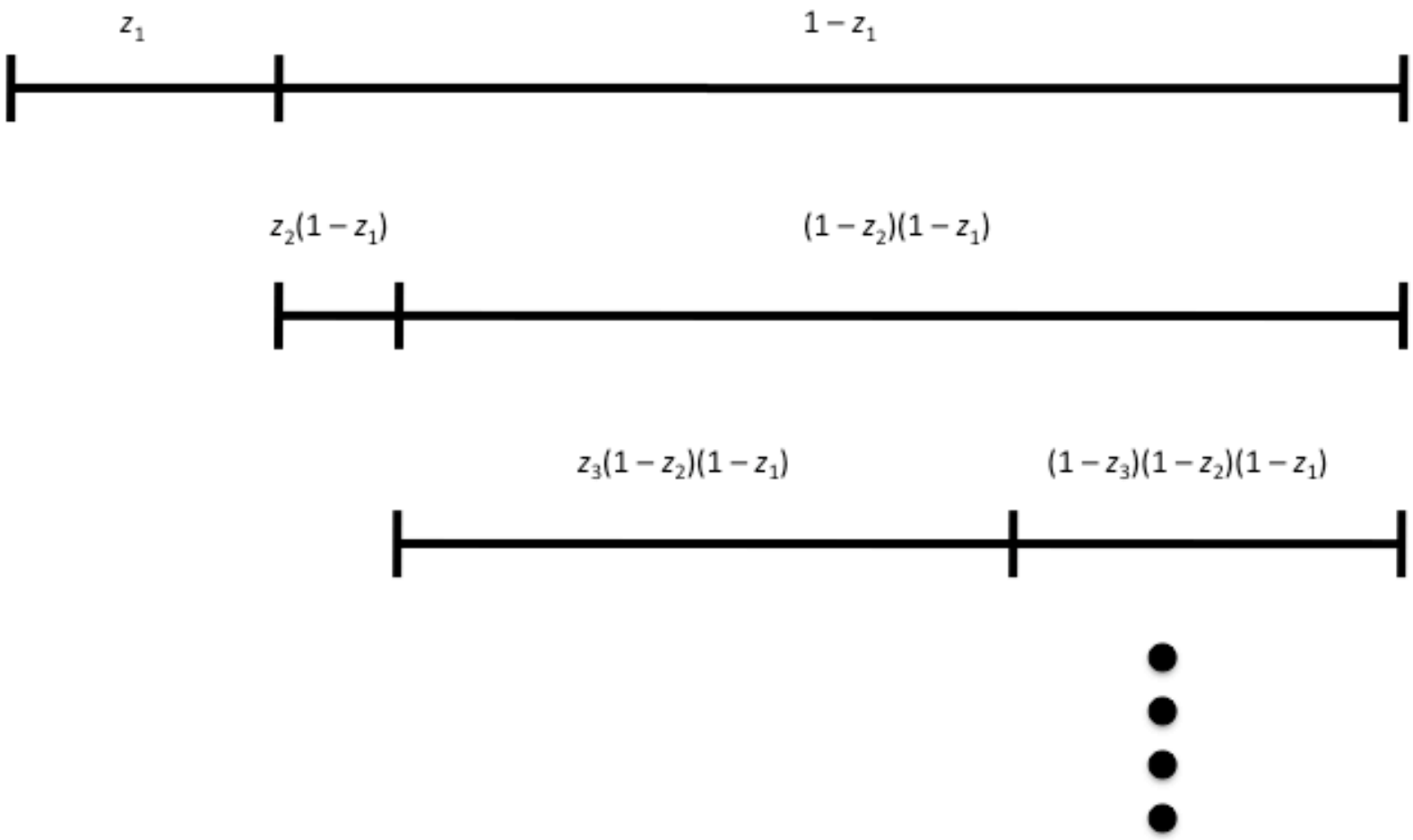
Dirichlet process and Dirichlet process mixtures

- The DP (Ferguson, 1973) is a random probability measure on distributions characterized by two parameters: a base distribution G_0 (the center of the process) and a (precision) parameter $\alpha > 0$
- DP constructive definition (Sethuraman, 1994)
  $\rightarrow$ let $\{z_s, s = 1, 2, ...\}$ and $\{\phi_j, j = 1, 2, ...\}$ be independent sequences of random variables, with z_s i.i.d. Beta$(1, \alpha)$, and ϕ_j i.i.d. G_0
  $\rightarrow$ define $\omega_1 = z_1$, $\omega_j = z_j \prod_{s=1}^{j-1}(1 - z_s)$, $j \geq 2$ (*stick-breaking* construction)
  $\rightarrow$ then, a realization G from DP(α, G_0) is (almost surely) of the form

$$G(\cdot) = \sum_{j=1}^{\infty} \omega_j \delta_{\phi_j}(\cdot)$$

  i.e., a discrete distribution that can be represented as a countable mixture of point masses

Introduction and background



- **Dirichlet process mixing:** for a parametric family of distributions $K(\cdot;\boldsymbol{\theta})$, $\boldsymbol{\theta} \in \Theta \subseteq \mathbb{R}^q$, define

$$F(\cdot;G) = \int K(\cdot;\boldsymbol{\theta})\mathrm{d}G(\boldsymbol{\theta}), \quad G \sim \mathrm{DP}(\alpha, G_0)$$

- **Hierarchical model:** for $y_1, ..., y_n$ i.i.d., given G, from $F(\cdot;G)$,

$$\begin{aligned} y_i \mid \boldsymbol{\theta}_i &\overset{ind.}{\sim} K(\cdot;\boldsymbol{\theta}_i), \;\; i = 1, ..., n \\ \boldsymbol{\theta}_i \mid G &\overset{i.i.d.}{\sim} G, \;\; i = 1, ..., n \\ G &\sim \mathrm{DP}(\alpha, G_0) \end{aligned}$$

$\rightarrow$ further stages can be added by placing hyperpriors on α and/or the parameters of G_0

2. A mixture modeling approach for stochastic order

- Focus on modeling for two stochastically ordered distribution functions $F_1 \leq_{st} F_2$ (corresponding to distributions supported on $\mathbb{R}$)
- Constructive approach to building the restriction $F_1(u) \geq F_2(u)$, $u \in \mathbb{R}$, through latent distribution functions G_1 and G_2 (on $\mathbb{R}$) such that

$$F_1(u) = G_1(u), \quad F_2(u) = G_1(u)G_2(u)$$

(Note: with $\theta \sim G_1$ and independently $\delta \sim G_2$, F_1 and F_2 are the distributions of θ and $\max\{\theta, \delta\}$, respectively)

- Now, work with (independent) nonparametric priors for G_1 and G_2 to induce a prior over $\mathcal{P}'_{st} = \{(F_1, F_2) : F_1 = G_1,\ F_2 = G_1G_2\}$, and hence over $\mathcal{P}_{st} = \{(F_1, F_2) :\ F_1 \leq_{st} F_2\}$

A mixture modeling approach for stochastic order

- How about using DP priors for G_1 and G_2
  $\rightarrow$ discreteness? simulation-based model fitting?

- Introduce DP mixing to overcome both difficulties

- **Key result:** for a parametric family of distributions $K(\cdot;\theta)$, $\theta \in (\underline{\theta}, \overline{\theta})$, strictly decreasing in θ, and H_1, H_2 two distribution functions on $(\underline{\theta}, \overline{\theta})$ with $H_1 \leq_{st} H_2$, defining

$$F(\cdot; H_i) = \int_{\underline{\theta}}^{\overline{\theta}} K(\cdot;\theta)\, \mathrm{d}H_i(\theta), \quad i = 1, 2$$

  we have $F(\cdot; H_1) \leq_{st} F(\cdot; H_2)$
  $\rightarrow$ result valid, e.g., for normal kernel mixing on the mean
  $\rightarrow$ add a dispersion parameter σ^2 to the model, to conclude that $F\left(\cdot; H_1, \sigma^2\right) \leq_{st} F\left(\cdot; H_2, \sigma^2\right)$ (semiparametric specification)

A mixture modeling approach for stochastic order

- Next, setting $H_1 = G_1$ and $H_2 = G_1 G_2$, we obtain the stochastically ordered DP mixture of normals model:

$$F_1(\cdot) \equiv F(\cdot; G_1, \sigma^2) = \int \mathrm{N}(\cdot; \theta, \sigma^2)\, \mathrm{d}G_1(\theta)$$

$$F_2(\cdot) \equiv F(\cdot; G_1, G_2, \sigma^2) = \iint \mathrm{N}(\cdot; \max\{\theta, \delta\}, \sigma^2)\, \mathrm{d}G_1(\theta)\mathrm{d}G_2(\delta)$$

  $\rightarrow$ with independent $\mathrm{DP}(\alpha_\ell, \mathrm{N}(\mu_\ell, \tau_\ell^2))$ priors for G_ℓ, $\ell = 1, 2$
  $\rightarrow$ an inverse gamma prior for σ^2, and priors for hyperparameters $\boldsymbol{\psi} = \{\alpha_\ell, \mu_\ell, \tau_\ell^2 : \ell = 1, 2\}$

- Consider data $= \{y_{1i} : i = 1, ..., n_1;\ y_{2j} : j = 1, ..., n_2\}$ where the y_{1i} (given G_1, σ^2) are ind. from F_1 and the y_{2j} (given G_1, G_2, σ^2) are ind. from $F_2(\cdot)$

A mixture modeling approach for stochastic order

- Hierarchical formulation of the model:

$$
\begin{aligned}
y_{1i} \mid \theta_i, \sigma^2 &\overset{\text{ind.}}{\sim} \mathrm{N}(\theta_i, \sigma^2), \;\; i = 1, ..., n_1 \\
y_{2j} \mid \theta_{n_1+j}, \delta_j, \sigma^2 &\overset{\text{ind.}}{\sim} \mathrm{N}(\max\{\theta_{n_1+j}, \delta_j\}, \sigma^2), \;\; j = 1, ..., n_2 \\
\theta_i \mid G_1 &\overset{\text{i.i.d.}}{\sim} G_1, \;\; i = 1, ..., n_1 + n_2 \\
\delta_j \mid G_2 &\overset{\text{i.i.d.}}{\sim} G_2, \;\; j = 1, ..., n_2 \\
G_1, G_2 \mid \boldsymbol{\psi} &\sim \mathrm{DP}(\alpha_1, \mathrm{N}(\mu_1, \tau_1^2)) \times \mathrm{DP}(\alpha_2, \mathrm{N}(\mu_2, \tau_2^2))
\end{aligned}
$$

- Through the introduction of the additional mixing parameters θ_{n_1+j}, $j = 1, ..., n_2$, the first stage conditionally independent specification in the hierarchical model is retained after marginalizing the random distribution functions G_1 and G_2 over their DP priors

A mixture modeling approach for stochastic order

Posterior inference

- Marginalizing G_1, G_2 yields priors $p(\boldsymbol{\theta} \mid \mu_1, \tau_1^2)$ and $p(\boldsymbol{\delta} \mid \mu_2, \tau_2^2)$ for $\boldsymbol{\theta} = \{\theta_i : i = 1, ..., n_1 + n_2\}$ and $\boldsymbol{\delta} = \{\delta_j : j = 1, ..., n_2\}$ that can be defined according to a generalized Pólya urn scheme (Blackwell & MacQueen, 1973)
- Marginal posterior $p(\boldsymbol{\theta}, \boldsymbol{\delta}, \sigma^2, \boldsymbol{\psi} \mid \text{data})$ is proportional to

$$\prod_{i=1}^{n_1} \mathrm{N}(y_{1i}; \theta_i, \sigma^2) \prod_{j=1}^{n_2} \mathrm{N}(y_{2j}; \max\{\theta_{n_1+j}, \delta_j\}, \sigma^2) p(\sigma^2) p(\boldsymbol{\theta} \mid \mu_1, \tau_1^2) p(\boldsymbol{\delta} \mid \mu_2, \tau_2^2) p(\boldsymbol{\psi})$$

- Posterior simulation from $p(\boldsymbol{\theta}, \boldsymbol{\delta}, \sigma^2, \boldsymbol{\psi} \mid \text{data})$ combining techniques from Esobar & West (1995) and Neal (2000)
- For more general inference involving the posteriors of G_1 and G_2,

$$p(G_1, G_2, \boldsymbol{\theta}, \boldsymbol{\delta}, \sigma^2, \boldsymbol{\psi} \mid \text{data}) = p(G_1 \mid \boldsymbol{\theta}, \mu_1, \tau_1^2) p(G_2 \mid \boldsymbol{\delta}, \mu_2, \tau_2^2) p(\boldsymbol{\theta}, \boldsymbol{\delta}, \sigma^2, \boldsymbol{\psi} \mid \text{data})$$

$\rightarrow$ where $p(G_1 \mid \boldsymbol{\theta}, \mu_1, \tau_1^2)$ and $p(G_2 \mid \boldsymbol{\delta}, \mu_2, \tau_2^2)$ are DPs with updated parameters (Antoniak, 1974)

A mixture modeling approach for stochastic order

- Sample from these two DPs using the DP stick-breaking representation with a truncation approximation

- Posterior samples G_{1b}, G_{2b} (and σ_b^2), $b = 1, ..., B$, for G_1, G_2 (and σ^2) yield, for any set of grid points u, samples:

  $\rightarrow \{F_{1b}(u) = \int \mathrm{N}(u; \theta, \sigma_b^2) \mathrm{d}G_{1b}(\theta) : b = 1, ..., B\}$ from the posterior of $F_1(u; G_1, \sigma^2)$, and

  $\rightarrow \{F_{2b}(u) = \int \int \mathrm{N}(u; \max\{\theta, \delta\}, \sigma_b^2) \mathrm{d}G_{1b}(\theta) \mathrm{d}G_{2b}(\delta) : b = 1, ..., B\}$ from the posterior of $F_2(u; G_1, G_2, \sigma^2)$

  $\rightarrow$ analogously for the posteriors of the mixture densities $f_1(u; G_1, \sigma^2)$ and $f_2(u; G_1, G_2, \sigma^2)$

3. Application to ROC analysis

(Joint work with Tim Hanson and Adam Branscum)

- A commonly encountered task in epidemiologic research (both human and veterinary) involves the characterization of the discriminatory ability of a continuous diagnostic test
- In particular, *serologic scores* measure the concentration of antigen-specific antibodies in serum
- Commonly used continuous diagnostic measures result in an optical density value or a serum-to-positive ratio for an enzyme linked immunosorbent assay (ELISA) serological test — a relatively large serologic score is suggestive of disease or infection presence
- Data illustrations with commercially available ELISAs designed to detect antibodies to Johne's disease in dairy cows — Johne's disease is endemic throughout the US affecting multiple species of animals

Application to ROC analysis

- *Gold-standard* data setting: disease (infection) status is assumed known

- F_1 and F_2 are the distribution functions associated with serologic scores for the noninfected and infected populations, respectively — typically, F_1 and F_2 are modelled independently

- Incorporate stochastic order constraint $F_1 \leq_{st} F_2$

- Biologically such a constraint is appropriate in many applications because serologic values for infected individuals tend to be larger than serologic values for noninfected individuals (provided the diagnostic test has reasonable discriminatory ability)

- Employ the stochastically ordered DP mixture model $F_1(\cdot) = F(\cdot; G_1, \sigma^2)$; $F_2(\cdot) = F(\cdot; G_1, G_2, \sigma^2)$

Application to ROC analysis

- Receiver operating characteristic (ROC) curve provides a commonly used graphical measure of the accuracy of the diagnostic test
- A cutoff value z can be used to dichotomize the serologic data into test positive (score $> z$) or test negative (score $< z$) categories
- ROC curve plots all possible pairs of true positive probability of infection $(1 - F_2(z))$ versus false positive probability $(1 - F_1(z))$ across all cutoff values z

$$\mathrm{ROC}(u) = 1 - F_2(F_1^{-1}(1-u)),\ u \in [0,1]$$

- Area under the curve, $\mathrm{AUC} = \int_0^1 \mathrm{ROC}(u)\,\mathrm{d}u$
  (probability that a randomly selected infected individual has a serologic score that is greater than that for a randomly selected noninfected individual)

- Posterior inference for $\mathrm{ROC}(\cdot)$ and AUC through the posteriors of $F(\cdot; G_1, \sigma^2)$ and $F(\cdot; G_1, G_2, \sigma^2)$

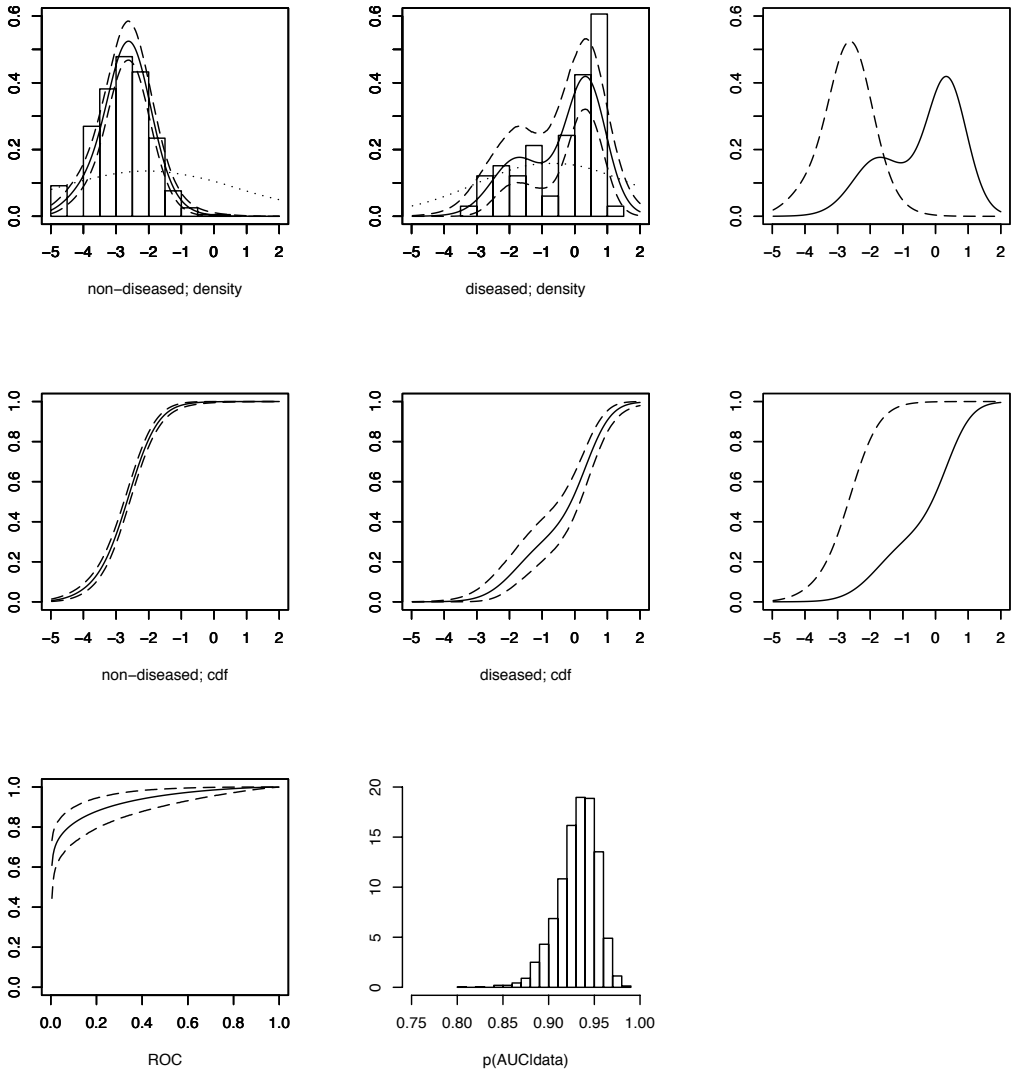
Application to ROC analysis

Data Examples

- Two data sets involving ELISAs designed to detect antibodies to Johne's disease (*Mycobacterium avium subspecies paratuberculosis*) in dairy cattle (Collins et al., 2005)
  → HerdChek ELISA test, manufactured by IDEXX Laboratories Inc., Westbrook, Maine: serology scores for $n_1 = 393$ noninfected and $n_2 = 66$ infected cows
  → ELISA test, developed by the Institut Pourquier, Montpellier, France: serology scores for $n_1 = 345$ noninfected and $n_2 = 258$ infected cows

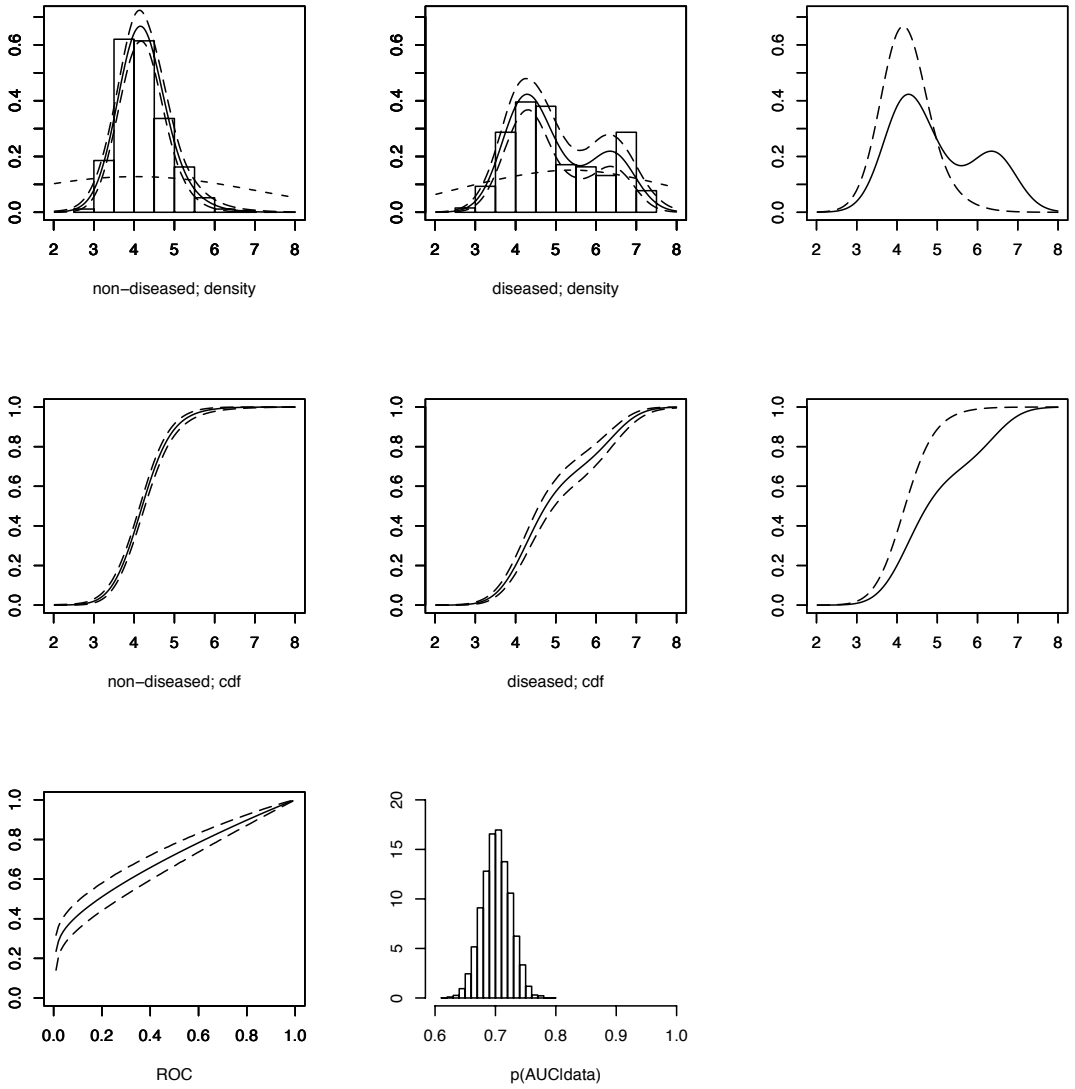
- Gold-standard data setting realistic in the data examples studied here, given surveillance measures taken to ensure absence of infection for the cows in the noninfected group, and cultures performed to identify infected cows

Application to ROC analysis



HerdChek ELISA test. Serologic scores for $n_1 = 393$ noninfected and $n_2 = 66$ infected cows.

Application to ROC analysis



Institut Pourquier ELISA test. Scores for $n_1 = 345$ noninfected and $n_2 = 258$ infected cows.

4. Modeling under stochastic precedence restrictions

- How about applications where a stochastic relationship between two distributions is anticipated, but the ordering of the respective distribution functions over their entire support is too restrictive?

- Weaker constraint of **stochastic precedence order** that builds the restriction through $\Pr(Y_1 \leq Y_2)$ for two independent random variables Y_1 and Y_2 with distribution functions F_1 and F_2: $Y_1 \leq_{sp} Y_2$ if $\Pr(Y_1 \leq Y_2) \geq 0.5$ (Arcones, Kvam & Samaniego, 2002)

- Using the (conditional) independence of Y_1 and Y_2, the stochastic precedence assumption implies $\mathrm{E}^{F_2}\{F_1(Y_2)\} \geq 0.5$ and $\mathrm{E}^{F_1}\{F_2(Y_1)\} \leq 0.5$

- For example, the normal distribution is stochastic precedence constrained with respect to its mean, $\mathrm{N}(\theta_1, \sigma_1^2) \leq_{sp} \mathrm{N}(\theta_2, \sigma_2^2)$ if-f $\theta_1 \leq \theta_2$, with no further restriction on σ_1^2 and σ_2^2

Modeling under stochastic precedence restrictions

- Semiparametric prior model for stochastic precedence using another mixture representation result: if

$$F_\ell(x) \equiv F_\ell(x; H_\ell, \sigma_\ell^2) = \int \mathrm{N}(x; \theta, \sigma_\ell^2)\, \mathrm{d}H_\ell(\theta), \quad \ell = 1, 2$$

  with stochastically ordered mixing distributions H_1 and H_2 (that is, $H_1 \leq_{st} H_2$), then $F_1 \leq_{sp} F_2$

- **DP mixture model**:

$$F_1(\cdot) \equiv F(\cdot; G_1, \sigma_1^2) = \int \mathrm{N}(\cdot; \theta, \sigma_1^2)\, \mathrm{d}G_1(\theta)$$

$$F_2(\cdot) \equiv F(\cdot; G_1, G_2, \sigma_2^2) = \iint \mathrm{N}(\cdot; \max\{\theta, \delta\}, \sigma_2^2)\, \mathrm{d}G_1(\theta)\mathrm{d}G_2(\delta)$$

  with independent DP priors for G_1 and G_2

Modeling under stochastic precedence restrictions

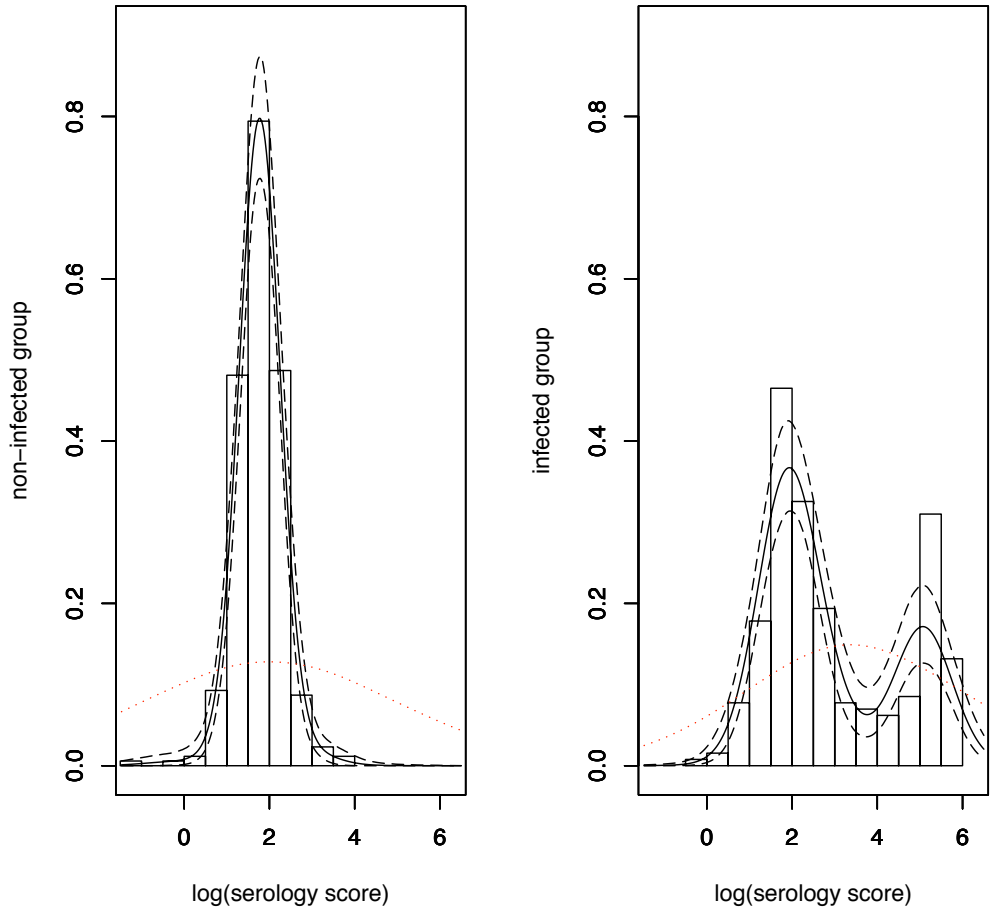
Application to ROC modeling

- Again, gold-standard data setting, with F_1 and F_2 representing the distributions of serologic scores for the noninfected and infected populations

- Here, the stochastic precedence constraint $F_1 \leq_{sp} F_2$ implies AUC ≥ 0.5, which is a natural constraint on ROC curves that effectively any diagnostic test must satisfy

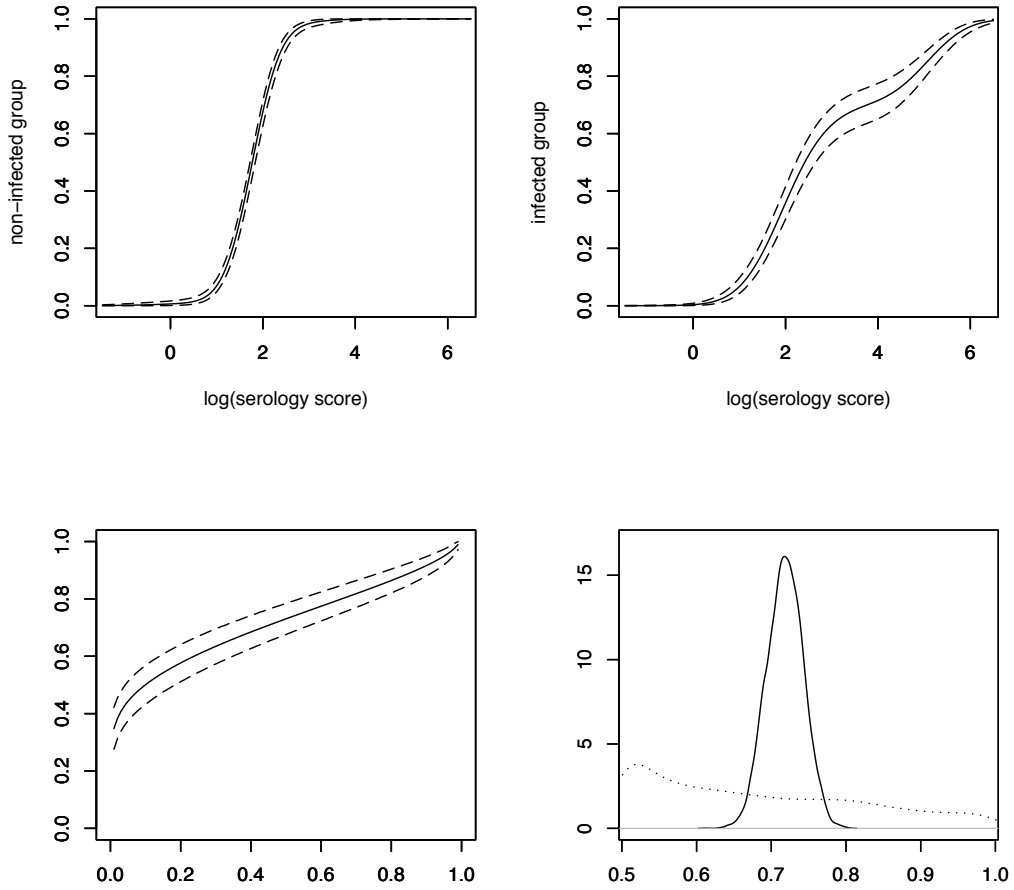
- **Data illustration:** one more ELISA test for Johne's disease, developed by the Synbiotic Corp. in San Diego, California (serology scores for $n_1 = 345$ noninfected and $n_2 = 258$ infected cows)

Modeling under stochastic precedence restrictions



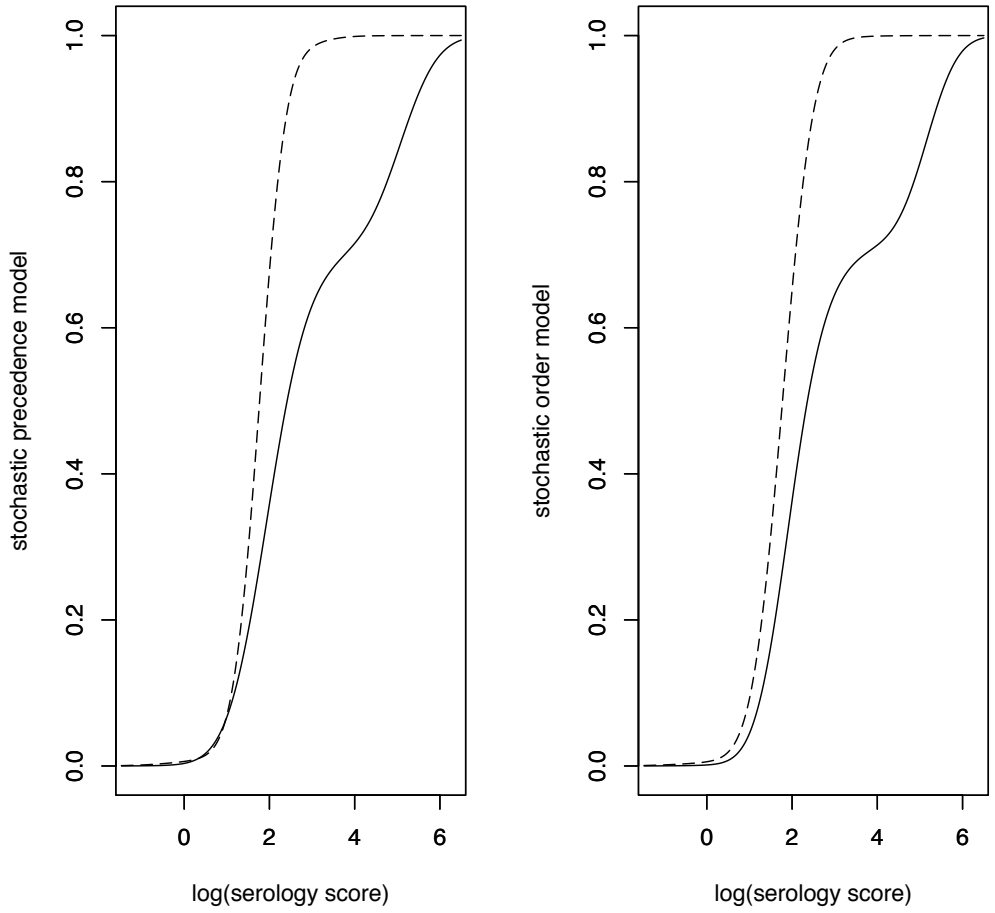
Synbiotic ELISA test. Posterior mean and 95% interval estimates, and prior predictive densities.

Modeling under stochastic precedence restrictions



Synbiotic ELISA test. Posterior mean and 95% interval estimates for the distribution function of the non-infected and infected groups (upper panels), and for the ROC curve (lower left panel). The lower right panel shows the prior and posterior density of the AUC.

Modeling under stochastic precedence restrictions



Synbiotic ELISA test. Comparison of posterior mean estimates for the distribution functions of the non-infected group (dashed lines) and the infected group (solid lines), under the stochastic precedence and stochastic order models.

5. Summary

- Bayesian semiparametric modeling for distributions constrained according to stochastic order or stochastic precedence
- Applications to ROC data analysis

- Extensions to settings without gold-standard data: one set of serologic values $(y_1, ..., y_n)$ comprising both infected and noninfected subjects
- Data can be assumed to arise from the two-component mixture $(1 - \pi)F_1 + \pi F_2$, where π is the disease (infection) prevalence — model components F_1, F_2 and π are not identifiable without restrictions
- Stochastic order restrictions might prove important in providing identifiability results (along with other conditions)

References

- Antoniak, C.E. (1974). Mixtures of Dirichlet processes with applications to nonparametric problems. *Annals of Statistics*, 2, 1152-1174.
- Arcones, M.A., Kvam, P.H. & Samaniego, F.J. (2002). Nonparametric estimation of a distribution subject to a stochastic precedence constraint. *Journal of the American Statistical Association*, 97, 170-182.
- Arjas, E. & Gasbarra, D. (1996). Bayesian inference of survival probabilities, under stochastic ordering constraints. *Journal of the American Statistical Association*, 91, 1101-1109.
- Chen, Z. & Dunson, D.B. (2004). Bayesian estimation of survival functions under stochastic precedence. *Lifetime Data Analysis*, 10, 159-173.
- Blackwell, D. & MacQueen, J.B. (1973). Ferguson distributions via Pólya urn schemes. *Annals of Statistics*, 1, 353-355.
- Collins, M.T., Wells, S.J., Petrini, K.R., Collins, J.E., Schultz, R.D. & Whitlock, R.H. (2005). Evaluation of five antibody detection tests for diagnosis of bovine paratuberculosis. *Clinical and Diagnostic Laboratory Immunology*, 12, 685-692.
- Dunson, D.B. & Peddada, S.D. (2008). Bayesian nonparametric inference on stochastic ordering. *Biometrika*, 95, 859-874.
- Evans, M., Gilula, Z., Guttman, I. & Swartz, T. (1997). Bayesian analysis of stochastically ordered distributions of categorical variables. *Journal of the American Statistical Association*, 92, 208-214.
- Escobar, M.D. & West, M. (1995). Bayesian density estimation and inference using mixtures. *Journal of the American Statistical Association*, 90, 577-588.
- Ferguson, T.S. (1973). A Bayesian analysis of some nonparametric problems. *Annals of Statistics*, 1, 209-230.

References

- Gelfand, A.E. & Kottas, A. (2001). Nonparametric Bayesian modeling for stochastic order. *Annals of the Institute of Statistical Mathematics*, 53, 865-876.
- Hanson, T.E., Kottas, A. & Branscum, A.J. (2008). Modeling stochastic order in the analysis of receiver operating characteristic data: Bayesian non-parametric approaches. *Applied Statistics*, 57, 207-225.
- Hoff, P.D. (2003). Bayesian methods for partial stochastic orderings. *Biometrika*, 90, 303-317.
- Karabatsos, G. & Walker, S.G. (2007). Bayesian nonparametric inference of stochastically ordered distributions, with Pólya trees and Bernstein polynomials. *Statistics & Probability Letters*, 77, 907-913.
- Kottas, A. (2009). Bayesian semiparametric modeling for stochastic precedence, with applications in epidemiology and survival analysis. Technical Report SOE-09-27, Department of Applied Mathematics and Statistics, University of California, Santa Cruz.
- Kottas, A. & Gelfand, A.E. (2001). Modeling variability order: A semiparametric Bayesian approach. *Methodology and Computing in Applied Probability*, 3, 427-442.
- Neal, R.M. (2000). Markov chain sampling methods for Dirichlet process mixture models. *Journal of Computational and Graphical Statistics*, 9, 249-265.
- Robertson, T., Wright, F. & Dykstra, R. (1988). *Order Restricted Statistical Inference*. Chichester: Wiley.
- Sethuraman, J. (1994). A constructive definition of the Dirichlet prior. *Statistica Sinica*, 4, 639-650.
- Shaked, M. & Shanthikumar, J.G. (1994). *Stochastic Orders and Their Applications*. Boston: Academic Press.

- Contact info:
  e-mail: thanos@ams.ucsc.edu, web: http://www.ams.ucsc.edu/∼thanos

- UCSC Dept of Applied Math and Statistics website: www.ams.ucsc.edu

THANKS!!!