Reinforced urns and the subdistribution beta-Stacy process prior for competing risks analysis

Andrea Arfè $^1$ Stefano Peluso $^2$ Pietro Muliere $^1$

$^1$Università Commerciale Luigi Bocconi
$^2$Università Cattolica del Sacro Cuore

SISBAYES 2017 workshop, Rome, Italy
In clinical prognostic research with a time-to-event outcome, occurrence of competing risks may preclude the occurrence of another event of interest.

Example: Amsterdam Cohort Study of AIDS progression in HIV-infected men [Geskus et al., 2003]
Subdistribution functions

Subdistribution function [Kalbfleisch and Prentice, 2002]

\[ F(t, d) = P(T \leq t, C = c) \]

where:

- \( T > 0 \) is the time to event onset
- \( C \in \{1, \ldots, K\} \) encodes the type of occurring event

- Can consider covariates: \( F(t, c|\mathbf{x}) \)
- **Note:** will assume that \( T \in \{1, 2, 3, \ldots\} \)
Outline

1 Literature review

2 The subdistribution beta-Stacy process

3 Predictive characterization via reinforced urn processes

4 Theoretical properties

5 A semiparametric competing risks regression model

6 Application: analysis of data from the Amsterdam Cohort Studies

7 Future work
Competing risks data has received widespread attention in the classical nonparametric literature, c.f. Kalbfleisch and Prentice [2002], Aalen et al. [2008], Andersen et al. [2012], Lawless [2011], Crowder [2012], and Pintilie [2006].

Classical Kalbfleisch and Prentice [2002] estimator of the subdistribution function:

\[
\hat{F}(t, c) = \sum_{u=1}^{t} \hat{S}(u - 1) \Delta \hat{A}_c(u)
\]

where:

- \(\hat{S}(t)\) is the Kaplan-Meier estimator of the survival function of \(T\);
- \(\hat{A}_c(t)\) is the Nelson-Aalen estimator of the cause-specific cumulative hazard function for the event \(C = c\).
• In contrast with the frequentist literature, the Bayesian nonparametric literature on competing risks is still sparse.
• Nonparametric Bayesian inference for competing risks involves a process prior on the space of subdistribution functions.
• Current proposals are based on either the gamma process [Ge and Chen, 2012] or the beta process [Hjort, 1990, De Blasi and Hjort, 2007] and its extensions, such as the beta-Dirichlet process [Kim et al., 2012, Chae et al., 2013], for subdistribution or cause-specific hazard functions.
Let $\alpha_{t,c} > 0$ for all $t = 1, 2, \ldots$ and $c = 1, \ldots, K$ and suppose that

$$\lim_{\tau \to +\infty} \prod_{t=1}^\tau \frac{\alpha_{t,0}}{\sum_{d=0}^K \alpha_{t,d}} = 0.$$  \hfill (1)

**Subdistribution beta-Stacy process**

A random subdistribution function $F$ is *subdistribution beta-Stacy* if:

1. $F(0, c) = 0$ with probability 1 for all $c = 1, \ldots, k$;
2. for all $c = 1, \ldots, k$ and all $t \geq 1$,

$$\Delta F(t, c) = W_{t,c} \prod_{u=1}^{t-1} \left(1 - \sum_{d=1}^k W_{u,d}\right),$$

where $\Delta F(t, c) = F(t, c) - F(t - 1, c)$ and

$$W_t = (W_{t,0}, \ldots, W_{t,k}) \overset{\text{ind}}{\sim} \text{Dirichlet}(\alpha_{t,0}, \ldots, \alpha_{t,k}).$$
**Reinforced Urn Processes**

- A **reinforced urn process (RUP)** is a stochastic process \((X_n)_{n\geq 0}\) with countable state-space \(S\) and \(X_0 = x_0 \in S\).
  - Each state \(x \in S\) is associated with a Pólya urn \(U(x)\) with balls of colors represented by the elements of the finite set \(E\);
  - If \(X_n = x\) and a ball of color \(c \in E\) is extracted from \(U(x)\), this is replaced in \(U(x)\) together with another ball of the same color (**reinforcement**). Then, \(X_{n+1} = q(x, c) \in S\), where \(q(x, c)\) is the **law** of motion.

**Theorem (Muliere et al. [2000])**

*If \((X_n)_{n\geq 0}\) is recurrent, then there exists a random transition matrix \(Q\) on \(S\) conditionally on which \((X_n)_{n\geq 0}\) is a Markov Chain with transition matrix \(Q\). The rows of \(Q\) are independent Dirichlet processes with base measure determined by the initial composition of the urns.*

- RUPs provide a predictive characterization of Pólya trees and discrete-time neutral-to-the-right processes [Muliere et al., 2000].
Construction via reinforced urn processes I

- $X_0 = (0, 0), X_1 = (1, 0), X_2 = (2, 0), X_3 = (3, 2), X_4 = (0, 0), X_5 = (1, 0), X_6 = (2, 1)$
- $(T_1, C_1) = (3, 2), (T_2, C_2) = (2, 1)$
The parameters $\alpha_{t,0}, \alpha_{t,1}, \ldots, \alpha_{t,K}$ determine the initial composition of the urn $U((t-1,0))$.

**Lemma (Recurrency condition)**

$(X_n)_{n \geq 0}$ is recurrent if and only if

$$\lim_{\tau \to +\infty} \prod_{t=1}^{\tau} \frac{\alpha_{t,0}}{\sum_{d=0}^{k} \alpha_{t,d}} = 0$$

(2)

**Theorem (Predictive characterization)**

Assume (2) holds. The sequence $((T_i, C_i))_{i \geq 1}$ generated by $(X_n)_{n \geq 0}$ is exchangeable. Its de Finetti measure is the law of a subdistribution beta-Stacy process.
Let $F_0$ be a fixed subdistribution function and $\omega_t > 0$ for all $t \geq 1$. Write

$$F \sim s\mathcal{BS}(\omega, F_0)$$

if $F$ has a subdistribution beta-Stacy process with parameters

$$\alpha_{t,c} = \omega_t \Delta F_0(t, c), \quad \alpha_{t,0} = \omega_t \left(1 - \sum_{d=1}^{k} F_0(t, d)\right)$$

Moments of $F \sim s\mathcal{BS}(\omega, F_0)$

- $E[F(t, c)] = F_0(t, c)$
- $\text{Var}(F(t, c))$ is a decreasing function of $\omega_t$
- $\text{Var}(F(t, c)) \to 0$ as $\omega_t \to +\infty$
Relation with other prior processes

- If $F(t, c) \sim s\mathcal{BS}(\omega, F_0)$, then $\sum_{d=1}^{K} F(t, d)$ is a random distribution function distributed according to the discrete-time beta-Stacy process of Walker and Muliere [1997].

- The cumulative hazards $A_c(t) = \Delta F(t, c)/(1 - \sum_{d=1}^{k} F(t - 1, d))$ are independent discrete-time beta processes [Hjort, 1990].

- The subdistribution beta-Stacy process is therefore also related to the beta-Dirichlet process of Kim et al. [2012], a generalization of Hjort’s beta process.
Posterior computations with censored data

- **Right-censored data**: $T_i^* = \min(T_i, R_i)$, $C_i^* = C_i \cdot I \{T_i < R_i\}$, where $R_i$ is a random censoring time, $i = 1, \ldots, n$.

- Censoring is **ignorable** if i) $R_1, \ldots, R_n$ are independent with common distribution function $H$, ii) conditional on $F$ and $H$, $(T_1, C_1), \ldots, (T_n, C_n)$ and $R_1, \ldots, R_n$ are independent [Heitjan and Rubin, 1991].

**Theorem**

1. If censoring is ignorable and $F \sim sBS(\omega, F_0)$ a priori, conditionally on $(T_1^*, C_1^*), \ldots, (T_n^*, C_n^*)$ the posterior distribution of $F$ is $sBS(\omega^*, F_0^*)$.

2. $F_0^*(t, c) \to \hat{F}(t, c)$ as $\max(\omega_t) \to 0$, where $\hat{F}$ is the classical Kalbfleisch-Prentice estimator of $F$. 
Semiparametric regression for competing risks

- A subdistribution beta-Stacy mixture regression model:

\[(T_i, C_i) \overset{\text{ind}}{\sim} F_i \quad (i = 1, \ldots, n)\]

\[F_i \overset{\text{ind}}{\sim} sBS(\omega(\theta, x_i), F_0(\cdot|\theta, x_i))\]

\[F_0(t, c|\theta, x_i) = F_0^{(1)}(c|\theta)F_0^{(2)}(t|c, \theta, x_i)\]

\[\omega_t(\theta, x_i) = \frac{1}{\sum_{d=1}^{K} \Delta F_0(t, d|\theta, x_i)}\]

\[\theta \sim \pi(\theta)\]

- We choose: \(F_0^{(1)} = \) multinomial logistic model; \(F_0^{(2)} = \) Weibull regression model; regression coefficients are assigned diffuse normal priors; Weibull scale parameters are assigned non-informative \(Gamma(\epsilon, \epsilon)\) distributions.

- Inference can be easily carried out via a Metropolis-Hastings algorithm after marginalizing away the \(F_i\) from the likelihood.
Objective: assessing the long-term prognosis of HIV infected men with respect to the risk of non-SI AIDS onset or SI onset as a function of CCR5 genotype: WW (wild type allele on both chromosomes) or WM (mutant allele on one chromosome).

A total of 324 male patients were followed-up from the date of HIV infection to the earliest among the dates of AIDSs onset, SI phenotype onset, or right censoring, i.e. death, study drop-out, or end of the study period.

A total of 65 (20.1\%) patients had a WM genotype, while mean age at HIV infection was about 34.6 years (s.d., 7.2 years). Overall, the 324 patients accumulated 2262.2 person-years of follow up (minimum - maximum follow-up: 0.1 years - 13.9 years), generating 117 cases of AIDS onset and 107 cases of SI onset.
Future developments

• Other reinforced urn schemes could be contemplated. For example, each extracted ball may be reinforced by a general amount \( m > 0 \) of new similar balls, or \( m \) may be random variable depending on the color of the extracted balls, as in Muliere et al. [2006]

• A continuous-time generalization of the subdistribution beta-Stacy process could be considered. We are currently developing a characterization of such process from a predictive perspective by means of the continuous-time urn models of Muliere et al. [2003] and Bulla and Muliere [2007]

• A generalization of the reinforced urn process considered in this work could be attempted to characterize a process prior on the space of transition kernels of a Markovian multistate process. Such process could be useful for the predictive Bayesian nonparametric analysis of event-history data [Aalen et al., 2008]


