SUPERPOSITION OF BETA PROCESSES

Pierpaolo De Blasi[1], Stefano Favaro[1] and Pietro Muliere[2]

[1] Dipartimento di Statistica e Matematica Applicata,
Università degli Studi di Torino
Corso Unione Sovietica 218/bis, 10134 Torino, Italy
(e-mail: pierpaolo.deblasi@unito.it, stefano.favaro@unito.it)

[2] Dipartimento di Scienze delle Decisioni,
Università L. Bocconi
Via Roentgen 1, 20135 Milano, Italy
(e-mail: pietro.muliere@unibocconi.it)

ABSTRACT. We consider a neutral to the right process that corresponds to the superposition of independent beta processes at the cumulative hazard level. It places a prior distribution on the survival distribution resulting from independent competing failure times. It can be derived as the infinitesimal weak limit of a discrete time process which has the conditional probability of an event at time t given survival up to t defined as the result of a series of m independent Bernoulli experiments. The continuous time version of the process, termed *m-fold beta NTR process*, is described in terms of completely random measures. We discuss prior specification and illustrate posterior inference on a real data example.

1 INTRODUCTION

In this paper we review some results of De Blasi, Favaro and Muliere (2009), where a new family of neutral to the right priors is introduced. A random distribution function F on $\mathbb{R}_+$ is neutral to the right (NTR) if, for any $0 \le t_1 < t_2 < \ldots < t_k < \infty$ and for any $k \ge 1$,

$$F_{t_1},\ \frac{F_{t_2}-F_{t_1}}{1-F_{t_1}},\ \ldots\ ,\frac{F_{t_k}-F_{t_{k-1}}}{1-F_{t_{k-1}}} \tag{1}$$

are independent random variables, see Doksum (1974). NTR priors have well developed theoretical properties. The form of the posterior distribution and its large sample properties are well known (see, e.g., Ferguson and Phadia (1979), Kim and Lee (2001, 2004)). These results are based on the representation of NTR priors as suitable transformations of completely random measures (CRMs), see Lijoi and Prünster (2009) for an account on their connections with Bayesian nonparametrics. We recall here that a CRM $\tilde{\mu}$ on $\mathbb{R}_+$ is a purely atomic random measure which admits *Lévy-Khintchine* representation

$$\mathrm{E}\Big[\mathrm{e}^{-\int_{\mathbb{R}_+} f(x)\tilde{\mu}(\mathrm{d}x)}\Big] = \exp\Big\{-\int_{\mathbb{R}_+\times\mathbb{R}_+}\Big[1-\mathrm{e}^{-sf(x)}\Big]\nu(\mathrm{d}s,\mathrm{d}x)\Big\} \tag{2}$$

for any $\tilde{\mu}$-integrable (almost surely) real-valued function f, where ν, referred to as the *Lévy intensity*, is a diffuse measure on $\mathbb{R}_+\times\mathbb{R}_+$ such that $\int_B\int_{\mathbb{R}_+}\min\{s,1\}\nu(\mathrm{d}s,\mathrm{d}x) < \infty$ for any bounded B in $\mathbb{R}_+$. Doksum (1974) showed that F is NTR if and only if $F_t = 1-\mathrm{e}^{-\tilde{\mu}((0,t])}$ for some CRM $\tilde{\mu}$ on $\mathbb{R}_+$ such that $\mathrm{P}[\lim_{t\to\infty}\tilde{\mu}((0,t]) = \infty] = 1$. Moreover, if F is NTR for a CRM

$\tilde{\mu}$, then the posterior distribution of F, given (possibly) right-censored data, is described by an NTR process for a CRM $\tilde{\mu}^*$ with fixed atoms at uncensored observations. This is of great importance for statistical inference, since one can resort to the simulation algorithm suggested in Ferguson and Klass (1972) to sample the trajectories of the posterior CRM, thus obtaining approximated evaluations of posterior distribution of F.

CRMs arise also when one considers the distribution induced by a NTR process on the space of cumulative hazard functions, i.e. the stochastic process $\{H_t, t > 0\}$ defined as $H_t = H_t(F) = \int_0^t (1-F_{x^-})^{-1} \mathrm{d}F_x$. Let $\{F_t, t > 0\}$ be an NTR process defined according to some CRM $\tilde{\mu}$ and let $\nu(\mathrm{d}s, \mathrm{d}x) = \nu(s,x)\mathrm{d}s\,\mathrm{d}x$ (with a little abuse of notation) be the corresponding Lévy intensity. Then $H_t = \tilde{\eta}((0,t])$, where $\tilde{\eta}$ is a CRM with Lévy intensity $\nu_H(\mathrm{d}v, \mathrm{d}x) = \nu_H(v,x)\mathrm{d}v\,\mathrm{d}x$ such that $\nu_H(v,x) = 0$ for any $v > 1$. The conversion formula for deriving the Lévy intensity of $\tilde{\eta}$ from that of $\tilde{\mu}$ is

$$\nu_H(v,x) = \tfrac{1}{1-v}\,\nu(-\log(1-v), x), \quad (v,x) \in [0,1] \times \mathbb{R}_+ \tag{3}$$

that is $\nu_H(\mathrm{d}v, \mathrm{d}x)$ is the distribution of $(s,x) \mapsto (1-\mathrm{e}^{-s}, x)$ under ν. Let us consider the beta-Stacy process of Walker and Muliere (1997), which is a NTR prior with a parametrization well suited for prior specification. To this aim, let α be a diffuse measure on $\mathbb{R}_+$ and $\beta : \mathbb{R}_+ \to \mathbb{R}_+$ a piecewise continuous function such that $\int_0^t \beta(x)^{-1}\alpha(\mathrm{d}x) \to \infty$ as $t \to \infty$. A beta-Stacy process $\{F_t, t > 0\}$ with parameters (α, β) is NTR for a CRM $\tilde{\mu}$ whose Lévy intensity is defined by $\nu(\mathrm{d}s, \mathrm{d}x) = \frac{\mathrm{d}s}{1-\mathrm{e}^{-s}}\,\mathrm{e}^{-s\beta(x)}\,\alpha(\mathrm{d}x)$. Using (3), one finds that the associated random cumulative hazard $\{H_t(F), t > 0\}$ is the cumulative of a CRM with Lévy intensity $\nu_H(\mathrm{d}v, \mathrm{d}x) = v^{-1}(1-v)^{\beta(x)-1}\mathrm{d}v\,\alpha(\mathrm{d}x)$, which corresponds to the Lévy intensity of the beta process of parameter (c, H_0) where $c(x) = \beta(x)$ and $H_0(t) = \int_0^t \beta(x)^{-1}\alpha(\mathrm{d}x)$, see Hjort (1990). Since the beta process has $\mathrm{E}(H_t) = H_0(t)$ and $\mathrm{Var}(H_t) = \int_0^t [\beta(x)+1]^{-1}\mathrm{d}H_0(x)$, the quantity α/β takes the role of prior guess of the hazard rate, while β acts as concentration parameter.

In the sequel we consider a new class of NTR priors which allows a more flexible specification of prior beliefs still retaining a simple interpretation in terms of prior parameters. It consists of a NTR process F that has $H_t(F)$ given by the superposition of m independent beta processes, say $H_{1,t}, \ldots, H_{m,t}$, according to

$$F_t \overset{d}{=} 1 - \prod\nolimits_{[0,t]} \left\{1 - \textstyle\sum_{i=1}^m \mathrm{d}H_{i,x}\right\}, \tag{4}$$

where $\prod_{[0,t]}$ denotes the product integral operator. Such F can be derived as the weak limit of a sequence of discrete time NTR processes. We discuss prior specification and posterior inference on a sample of right-censored survival data.

2 THE m-FOLD BETA NTR PROCESS

We start with a discrete time process which satisfies the independence condition in (1). Following the idea of Walker and Muliere (1997), we adopt a stick breaking construction: let $0 < t_1 < t_2 < \ldots$ be a sequence of time points indexed by $k = 1, 2, \ldots$ and define $F_{t_k} = \sum_{j=1}^k V_j \prod_{l=1}^{j-1}(1-V_l)$ for $V_1, V_2, \ldots$ a sequence of independent random variables with values in $(0,1)$ such that $\lim_{k\to\infty} \mathrm{E}(1-F_{t_k}) = \prod_{k\geq 1}\mathrm{E}(1-V_k) = 0$. For any $k \geq 1$, V_k is defined as

$$V_k \overset{d}{=} 1 - \textstyle\prod_{i=1}^m (1 - Y_{i,k}) \tag{5}$$

for m sequences of positive real numbers $(\alpha_{1,\bullet},\beta_{1,\bullet}) := \{(\alpha_{1,k},\beta_{1,k}), k \geq 1\}, \ldots, (\alpha_{m,\bullet},\beta_{m,\bullet}) := \{(\alpha_{m,k},\beta_{m,k}), k \geq 1\}$ and m independent sequences of random variables (r.v.s) $\{Y_{1,k}, k \geq 1\}, \ldots, \{Y_{m,k}, k \geq 1\}$ such that $\{Y_{i,k}, k \geq 1\}$ is a sequence of independent r.v.s with $Y_{i,k} \sim \mathrm{beta}(\alpha_{i,k},\beta_{i,k})$. If we define $X_k = V_j \prod_{l=1}^{j-1}(1-V_l)$ so that $F_t = \sum_{t_k \leq t} X_k$, then

$$X_1 \stackrel{d}{=} V_1, \quad X_k | X_1, \ldots, X_{k-1} \stackrel{d}{=} (1 - F_{t_{k-1}})V_k, \; k \geq 2 \tag{6}$$

Note that $X_k < 1 - F_{k-1}$ almost surely (a.s.), so that $F_k \leq 1$ a.s.. Moreover, the condition $\prod_{k\geq 1}\prod_{i=1}^{m}[1 - \alpha_{i,k}/(\alpha_{i,k}+\beta_{i,k})] = 0$ ensures that $F_{t_k} \to 1$ a.s. as $k \to \infty$ so that $\{F_t, t \geq 0\}$ is a discrete time NTR process.

This construction includes as particular case the discrete time version of the beta-Stacy process, see Walker and Muliere (1997, Secion 3), which has $V_k \sim \mathrm{beta}(\alpha_k,\beta_k)$. In fact, using some known properties for the product of independent beta distributed r.v.s, for the parameter configuration $(\alpha_{1,\bullet}, \beta_\bullet), (\alpha_{2,\bullet}, \beta_\bullet + \alpha_{1,\bullet}), \ldots, (\alpha_{m,\bullet}, \beta_\bullet + \sum_{i=1}^{m-1}\alpha_{i,\bullet})$, $\{F_t, t \geq 0\}$, is a discrete time beta-Stacy process with parameter $(\sum_{i=1}^{m}\alpha_{i,\bullet}, \beta_\bullet)$. Note that $V_k = X_k/(1-F_{k-1})$ is the conditional probability of observing the event at time t_k given survival up to t_k. According to (5), V_k corresponds to the minimum of m independent r.v.s $Y_{i,k} \sim \mathrm{beta}(\alpha_{i,k},\beta_{i,k})$ and can be interpreted as the result of m independent Bernoulli experiments: we observe the event if at least one of the m experiment has given a positive result. When $\beta_{1,k} = \beta_k$ and $\beta_{i,k} = \beta_k + \sum_{j=1}^{i-1}\alpha_{j,k}$, $2 \leq i \leq m$, the minimum of independent beta r.v.s is beta distributed, hence we recover the construction in Walker and Muliere (1997).

The next theorem establishes the existence of a continuous version of the process and can be proved by taking the infinitesimal weak limit of a sequence of discrete time processes defined in (5)-(6), see the proof of Theorem 2 in Walker and Muliere (1997).

Theorem 1. *Let $\alpha_1, \ldots, \alpha_m$, $m \geq 1$, be diffuse measures on $\mathbb{R}_+$ and let $\beta_1, \ldots, \beta_m$ be positive and piecewise continuous functions defined on $\mathbb{R}_+$ such that $\int_0^t \beta_i(x)^{-1}\alpha_i(\mathrm{d}x) \to \infty$ as $t \to \infty$ for $i = 1, \ldots, m$. Then, there exists a CRM $\tilde{\mu}$ with Lévy intensity*

$$\nu(\mathrm{d}s, \mathrm{d}x) = \frac{\mathrm{d}s}{1-\mathrm{e}^{-s}} \sum_{i=1}^{m} \mathrm{e}^{-s\beta_i(x)}\alpha_i(\mathrm{d}x). \tag{7}$$

The process $\{F_t = 1 - \mathrm{e}^{-\tilde{\mu}((0,t])}, t > 0\}$ is NTR so that, at the infinitesimal level, $\mathrm{d}F_t | F_t \stackrel{d}{=} (1 - F_t)[1 - \prod_{i=1}^{m}(1 - Y_{i,t})]$ where $(Y_{1,t}, \ldots, Y_{m,t})$ are independent with $Y_{i,t} \sim \mathrm{beta}(\alpha_i(\mathrm{d}t), \beta_i(t))$.

We name the process $\{F_t, t > 0\}$ in Theorem 1 an *m-fold beta NTR process* of parameter $(\alpha_1,\beta_1), \ldots, (\alpha_m,\beta_m)$. The definition extends to the case of α_i having point masses: let $\{t_k, k \geq 1\}$ be now the sequence obtained by collecting all t_k such that $\alpha_i\{t_k\} > 0$ for some $i = 1, \ldots, m$ and let $\alpha_{i,c}$ be the non-atomic part of α_i. Then $\tilde{\mu}$ has a fixed point at t_k with random masses V_k distributed according to

$$1 - \mathrm{e}^{-V_k} \stackrel{d}{=} 1 - \textstyle\prod_{i=1}^{m}(1 - Y_{i,t_k}), \qquad Y_{i,t_k} \sim \mathrm{beta}\big(\alpha_i\{t_k\}, \beta_i(t_k)\big), \tag{8}$$

while the intensity measure in (7) has α_i replaced by $\alpha_{i,c}$. The beta-Stacy process is recovered when $\beta_i(x) = \beta(x) + \sum_{j=1}^{i-1}\alpha_j\{x\}$ for some fixed function $\beta(\cdot)$. In fact, one obtains $\nu(\mathrm{d}s, \mathrm{d}x) = \frac{\mathrm{d}s}{1-\mathrm{e}^{-s}}\, \mathrm{e}^{-s\beta(x)}(\sum_{i=1}^{m}\alpha_{i,c})(\mathrm{d}x)$, and $1 - \mathrm{e}^{-V_k} \sim \mathrm{beta}(\sum_{i=1}^{m}\alpha_i\{t_k\}, \beta(t_k))$, that is a beta-Stacy process of parameter $(\sum_{i=1}^{m}\alpha_i, \beta)$, see Walker and Muliere (1997, Definition 3).

3 Superposition of beta processes

We consider now the prior distribution on the space of cumulative hazard functions induced by an m-fold beta NTR process. Let us consider the case of no fixed points of discontinuity, so take α_i absolutely continuous for any i. Then, by using equation (3),

$$\nu_H(\mathrm{d}v,\mathrm{d}x) = \frac{\mathrm{d}v}{v(1-v)}\sum_{i=1}^{m}(1-v)^{\beta_i(x)}\alpha_i(\mathrm{d}x) = \sum_{i=1}^{m}\beta_i(x)v^{-1}(1-v)^{\beta_i(x)-1}\mathrm{d}v\,\frac{\alpha_i(\mathrm{d}x)}{\beta_i(x)} \tag{9}$$

that is the sum of the Lévy intensities of m beta processes $H_{1,t},\dots,H_{m,t}$ with parameters $(c_i, H_{0,i})$ where $c_i(x) = \beta_i(x)$ and $H_{0,i}(t) = \int_0^t \beta_i(x)^{-1}\alpha_i(\mathrm{d}x)$, $i = 1,\dots,m$. It follows that $H(F)$ is the superposition of m independent beta processes, according to (4). In particular, F can be seen as the distribution function of the minimum of m independent failure times, $F_t = \mathrm{P}\{\min(X_1,\dots,X_m) \le t\}$ where $\mathrm{P}(X_i \le t) = 1-\prod_{[0,t]}\left(1-\mathrm{d}H_{i,x}\right)$ and $H_{i,x}$ takes the interpretation of the random cumulative hazard associated to the ith failure type (*ith failure-specific cumulative hazard*). It is interesting to see the similarity of (4) with the waiting time distribution in state 0 of a continuous time Markov chain $\{X_t, t>0\}$ in the state space $\{0,1,\dots,m\}$ where 0 is the initial state and $H_{i,x}$ is the cumulative intensity of the transition $0 \to i$, $i = 1,\dots,m$, cf. Andersen *et al* (1993, Section II.7). Then $\mathrm{P}(X_t = 0) = \prod_{[0,t]}\left\{1-\sum_{i=1}^m \mathrm{d}H_{i,x}\right\}$. The cumulative transition intensities are constrained to $\sum_{i=1}^m \mathrm{d}H_{i,x} \le 1$ since, conditionally on the past, the transition out of state 0 in an infinitesimal time interval is the result of a multinomial experiment. However, in (4) the transition is rather the result of a series of independent Bernoulli experiments, which is equivalent to considering a competing risks model generated by independent latent lifetimes, see Andersen *et al* (1993, Section III.1.2).

We can exploit (4) for prior elicitation and express different prior beliefs for the random failure-specific cumulative hazards $\{H_{i,t}, t>0\}$. Suppose we adopt the prior specification

$$\alpha_i(\mathrm{d}x) = k_i h_{0,i}(x)\mathrm{e}^{-H_{0,i}(t)}\mathrm{d}x, \qquad \beta_i(x) = k_i\mathrm{e}^{-H_{0,i}(t)} \tag{10}$$

for each of the beta processes $\{H_{i,t}, t>0\}$, where $h_{0,i}$ is the hazard rate corresponding to $H_{0,i}$. Then we center the ith failure-specific cumulative hazard on $H_{0,i}(t)$ and we set the concentration around $H_{0,i}$ according to k_i, which has a prior sample size interpretation: with i.i.d. survival times, $\beta_i(t)$ represents the number at risk at t in an imagined prior sample of uncensored survival times, with k_i the sample size , see Hjort (1990, Remark 2B). It follows that the prior mean of the cumulative hazard is set equal to the sum $\mathrm{E}[H_t(F)] = \sum_{i=1}^m H_{0,i}(x)$, while the k_i's allow to specify different degrees of prior beliefs on each of the m components $H_{0,i}$.

4 Illustration

We start with the derivation of the posterior distribution of a m-fold beta NTR process given a set of right censored observations. Consider data $(T_1,\Delta_1)\dots,(T_n,\Delta_n)$, where $\Delta_i = 1$ indicates an uncensored observation. We adopt a point process formulation, which is standard in survival analysis, by defining $N(t) = \sum_{i\le n}\mathbb{1}_{(T_i\le t,\Delta_i=1)}$ and $Y(t) = \sum_{i\le n}\mathbb{1}_{(T_i\ge t)}$. We know that the posterior NTR process is the exponential transform of a CRM, say $\tilde{\mu}*$, with fixed atoms at uncensored observations. By using Ferguson and Phadia (1979) (see also Lijoi and Prünster (2009)) we have $\tilde{\mu}^* = \tilde{\mu}_c^* + \sum_{t_k:N\{t_k\}>0} V_k^* \delta_{t_k}$, where $\tilde{\mu}_c^*$ has Lévy intensity

$$\nu^*(\mathrm{d}s, \mathrm{d}x) = \tfrac{\mathrm{d}s}{1-\mathrm{e}^{-s}} \textstyle\sum_{i=1}^m \mathrm{e}^{-s[\beta_i(x)+Y(x)]}\alpha_i(\mathrm{d}x), \tag{11}$$

while the density f_{t_k} of the jump V_k^* at time point t_k such that $N\{t_k\} > 0$ is given by

$$f_{t_k}(s) = \kappa(1-\mathrm{e}^{-s})^{N\{t_k\}-1} \textstyle\sum_{i=1}^m \mathrm{e}^{-s[\beta_i(t_k)+Y(t_k)-N\{t_k\}]}, \tag{12}$$

where κ is the appropriate normalizing constant. Note that $\tilde{\mu}_c^*$ can be described in terms of a m-fold beta NTR process with updated parameters $(\alpha_1, \beta_1 + Y), \ldots, (\alpha_m, \beta_m + Y)$. However, the densities f_{t_k} have not the form in (8). Upon definition of $\alpha_i^*(\mathrm{d}x) = \alpha_i(\mathrm{d}x) + \delta_{N\{x\}}(\mathrm{d}x)$ and $\beta_i^*(x) = \beta_i(x) + Y(x) - N\{x\}$, the distribution of V_k^* is a mixture of beta r.v.s:

$$1 - \mathrm{e}^{-V_k^*}|I = i \sim \mathrm{beta}(\alpha_i^*\{t_k\}, \beta_i^*(t_k)), \quad \mathrm{P}(I = i) = \frac{B(\alpha_i^*\{t_k\}, \beta_i^*(t_k))}{\sum_{j=1}^m B(\alpha_j^*\{t_k\}, \beta_j^*(t_k))} \tag{13}$$

where $B(a,b)$ is the beta function $B(a,b) = \int_0^1 v^{a-1}(1-v)^{b-1}\mathrm{d}v$.

As an illustrative example, we consider the Kaplan and Meier (1958) data set. We use a prior specification as in (10) with $m = 2$, $k_1 = k_2 = 1$ and $h_{0,i}(t) = \frac{\kappa_i}{\lambda_i}\left(\frac{t}{\lambda_i}\right)^{\kappa_i - 1}\mathrm{e}^{-(t/\lambda_i)^{\kappa_i}}$, $i = 1, 2$, that is two hazard rates of the Weibull type. We choose $\lambda_1 = \lambda_2 = 20$, $\kappa_1 = 1.5$ and $\kappa_2 = 0.5$, so that the prior process is centered on a survival distribution with non monotonic hazard rate, see Figure 1(a) and Figure 1(b). In order to simulate from the posterior, we implement a Ferguson and Klass algorithm (see also Walker and Damien (2000)) for the CRM $\tilde{\mu}_c^*$, while the jumps V_k^* are generated by using the mixture of beta densities in (13). We sample 5000 trajectories from the posterior process on the time interval $[0,T]$ for $T = 50$ and we evaluate the posterior distribution of two quantities: the probability F_t for $t = 5$ and the exponential-type functional $I_T(\tilde{\mu}^*) = [1 - \mathrm{e}^{-\tilde{\mu}^*((0,T])}]^{-1} \int_0^T \mathrm{e}^{-\tilde{\mu}^*((0,t])}\mathrm{d}t - T\mathrm{e}^{-\tilde{\mu}^*((0,T])}$, which corresponds to the random mean of the distribution obtained via normalization of F_t over $[0,T]$. $I_T(\tilde{\mu}^*)$ can be used to approximate, for T large, the random mean of the posterior NTR process which takes the interpretation of expected lifetimes. Figure 1(c) displays the histogram and a density estimate of the posterior distribution of F_t for $t = 5$, while, in Figure 1(d), we compare the density of the distribution of $I_T(\tilde{\mu}^*)$ with the density of $I_T(\tilde{\mu})$, the latter calculated over 5000 trajectories of the prior process.

REFERENCES

ANDEREN, P.K., BORGAN, Ø., GILL, R.D., KEIDING, N. (1993): *Statistical models based on counting processes*. Springer, New York.

DE BLASI, P., FAVARO, S., MULIERE, P. (2009): A class of neutral to the right priors induced by superposition of beta processes. *Technical Report*.

DOKSUM, K.A. (1974): Tailfree and neutral random probabilities and their posterior distributions. *Ann. Probab.*, *2*, 183–201.

FERGUSON, T.A., KLASS, M.J. (1972): A representation of independent increment processes without Gaussian components. *Ann. Math. Statist.*, *43*, 1634–1643.

FERGUSON, T.A., PHADIA, E.G. (1979): Bayesian nonparametric estimation based on censored data. *Ann. Statist.*, *7*, 163–186.

HJORT, N.L. (1990): Nonparametric Bayes estimators based on Beta processes in models for life history data. *Ann. Statist.*, *18*, 1259–1294.

KAPLAN , E.L., MEIER, P. (1958): Nonparametric estimation from incomplete observations. *J. Am. Statist. Assoc.*, *53*, 457–481.

KIM, Y., LEE, J. (2001): On posterior consistency of survival models. *Ann. Statist.*, *29*, 666–686.

KIM, Y., LEE, J. (2004): A Bernstein-von Mises theorem in the nonparametric right-censoring model. *Ann. Statist.*, *32*, 1492–1512.

LIJOI, A., PRÜNSTER, I. (2009): Models beyond the Dirichlet process. In: N.L. Hjort, C.C. Holmes, P. Müller and S.G. Walker (Eds.), *Bayesian Nonparametrics*. Cambridge University Press, in press.

WALKER, S.G., DAMIEN, P. (2000): Representation of Lévy processes without Gaussian components. *Biometrika*, *87*, 477–483.

WALKER, S.G., MULIERE, P. (1997): Beta-Stacy processes and a generalization of the Polya-urn scheme. *Ann. Statist.*, *25*, 1762–1780.

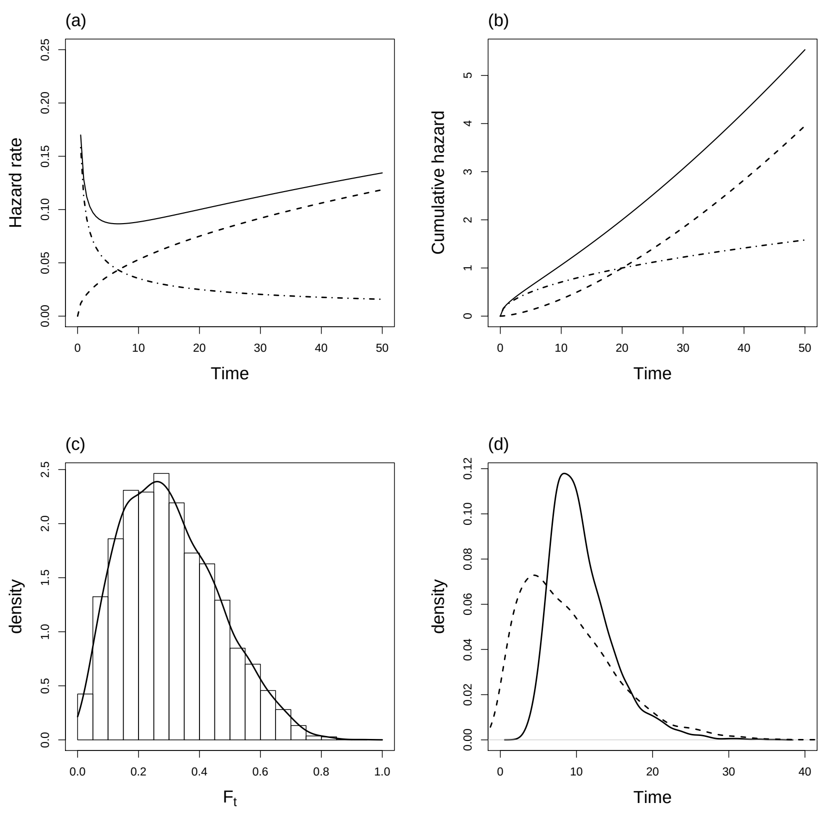

Figure 1. *(a) Hazard rate (—) and failure-specific hazard rates $h_{0,1}$ (- - -) and $h_{0,2}$ (-·-·-) in the prior. (b) The same for the cumulative hazard. (c) Histogram estimate of the posterior density of F_t for $t = 5$. (d) Estimate of the density of $I_T(\tilde{\mu}^*)$ (—) and $I_T(\tilde{\mu})$ (- - -) for $T = 50$.*