Statistica & Applicazioni - Vol. VIII, n. 1, 2010, pp. 105-116

A NON PARAMETRIC MODEL FOR ECOLOGICAL RELATIVE ABUNDANCE

Mariateresa Cuoccio*

SUMMARY

In this paper we derive nonparametric Bayesian estimate for a real parameter of a measure which is, in its turn, the function parameter of a mixture of a Dirichlet process and credibility interval for the predictive value. We apply this approach to data of plant species of relative abundance as given by Alodat, Odat, Muhaidat and Beldjillali (2008).

Keywords: *Power Function Distribution, Bayes Estimator, Prior Distribution, Plant Relatives abundance.*

1. INTRODUCTION

In the paper by Alodat Odat, Muhaidat and Beldjillali (2008), the problem of species of relative abundance is faced in a Bayesian and non Bayesian settings using a parametric model for observations. In that paper Authors analyzed data gathered from 71 grasslands communities in central Germany. They provided data for *Dactylis glomerata* as a representative species from 19 sites that are comparable with regard to elevation and climatic conditions.

The Authors consider observations $X_1, X_2, ..., X_n$, conditional given a real parameter θ, i.i.d. with distribution function

$$F(x|\theta) = \begin{cases} 0, & x < 0, \\ x^{\frac{\theta}{1-\theta}}, & 0 \leq x < 1\,, \qquad 0 < \theta < 1, \\ 1, & 1 \leq x, \end{cases}$$

with density (ordinary) $f(x|\theta) = \dfrac{\theta}{1-\theta} x^{\frac{2\theta-1}{1-\theta}} I_{(0,1)}(x)$ and give the Bayesian estimate of θ (by numeric integration for the data at hand), when the prior for θ is assumed uniform on $[0, 1]$ or Jeffrey's non informative, and the predictive distribution. They also give various credibility intervals.

After having reminded the notion of Dirichlet process and mixture of Dirichlet process, our aim is to face the same problem from a nonparametric Bayesian point of view using a mixture of Dirichlet process.

---

* Facoltà di Economia - Università Lum Jean Monnet - S.S. 100, Km. 18 - 70020 CASAMASSIMA (e-mail: cuoccio@lum.it).

Work supported by University Lum Jean Monnet.

We shall consider observations $X_1, X_2, ..., X_n$ conditional given θ and a probability measure P, i.i.d. $P$and P a random process. As we can see the obtained results are generally different from those obtained by Alodat *et al*. in a parametric setting that is when the statistical model is assumed non random (given θ). Our results coincide with their ones in the particular case in which $\bar{\alpha}$ (a parameter which will be explained in a forthcoming section) tends to infinite.

2. A PRIOR ON A RANDOM PROBABILITY MEASURE

In this section we want to state the definitions and main properties of a process (mixture of Dirichlet process) due to Antoniak (1974), which we can use to deal with our problem from a non-parametric viewpoint.

Let B be the $\sigma-$algebra generated by the intervals of R, $[0,1]^{\mathrm{B}}$ the class of functions $P : \mathrm{B} \to [0,1]$, $\Im^{\mathrm{B}}$ the $\sigma-$algebra generated by the algebra of cylinders of $[0,1]^{\mathrm{B}}$ and $B_1, B_2, ..., B_m$ a measurable partition of R. To denote that the random vector

$$(P(B_1), P(B_2), ..., P(B_m))$$

has a $(m-1)$- dimensional density given by the Dirichlet density:

$$\frac{\Gamma\left(\sum_1^m \alpha_i\right)}{\Gamma(\alpha_1)\Gamma(\alpha_2).....\Gamma(\alpha_m)} \left(\prod_{j=1}^{m-1} y_j^{\alpha_j - 1}\right)\left(1 - \sum_{j=1}^{m-1} y_j\right)^{\alpha_m - 1} \qquad \alpha_j > 0; \;\; j = 1, ..., m$$

$$\text{on} \left\{(y_1, ..., y_{m-1}) \in R^{m-1}, \;\; y_j \geq 0, \quad \sum_{i=1}^{m-1} y_j \leq 1\right\}, \tag{1}$$

we write:

$$(P(B_1), ..., P(B_m)) \in D(\cdot|\alpha_1, ..., \alpha_m).$$

Definition 1 (Ferguson, 1973)
Let α be a finite, finitely additive non negative measure on (R, B). If for every measurable partition $(B_1, ..., B_m)$of R it holds that

$$(P(B_1), ..., P(B_m)) \in D(\cdot|\alpha(B_1), ..., \alpha(B_m)) \; ; \;\; \alpha(R) = \sum_{j=1}^{m} \alpha(B_j)$$

then there exists a probability space $\left([0,1]^{\mathrm{B}}, \;\; \Im^{\mathrm{B}}, \mathrm{P}\right)$ in which the probability P is compatible with $D(\cdot|\alpha(B_1), ..., \alpha(B_m))$. We say then $(P(B), \;\; B \in \mathrm{B})$ is a Dirichlet process with parameter $\alpha(\cdot)$.

Important properties of Dirichlet process are:
if P is a Dirichlet process with parameter $\alpha(\cdot)$, then $E(P(A_1)) = \alpha(A_1)/\alpha(R)$, if P is a Dirichlet process and conditional given P, the observations $X_1, X_2, ..., X_n$ are i.i.d. P, then $P|X_1, X_2, ..., X_n$ is again a Dirichlet process with parameter $\alpha(\cdot) + \sum_i^n \delta_{X_i}$ where δ_X denotes the measure giving a mass one at the point X.

Moreover it follows from the previous properties that

$$E\big(P((-\infty, x])|X_1, X_2, ..., X_n\big) = \frac{\alpha(R)}{\alpha(R)+n}\left(\frac{\alpha((-\infty, x])}{\alpha(R)}\right) + \frac{n}{\alpha(R)+n}\ \hat{F}_n(x), \quad (2)$$

where $\hat{F}_n(\cdot)$ is the empirical distribution function.

The magnitude of $\alpha(R)$represents ‘‘the degree of faith’’ in the prior guess. One can chose its magnitude to represent the strength of his conviction about the ‘‘shape’’ of the distribution.

Useful readings about the Dirichlet process are Schervish (1995) and Rodriguez, Dunson, Gelfand (2008).

Now we are able to state.

DEFINITION 2 (Antoniak, 1974)
If, given a real parameter θ, the probability measure P is a Dirichlet process with parameter $\alpha(\theta, \cdot) : R \times \mathrm{B} \to [0, \infty)$, we say that P is a mixture of Dirichlet process if

$$\mathrm{P}(P(B_1) \le y_1, ..., P(B_m) \le y_m) = \int_R D(y_1, ..., y_m|\alpha(\theta, B_1), ..., \alpha(\theta, B_m))d\varphi(\theta),$$

where P *is a probability measure on* $\left([0,1]^{\mathrm{B}},\ \Im^{\mathrm{B}}\right)$; $(B_1, ..., B_m)$ *a measurable partition of R, and* $D(y_1, ..., y_{m-1}, |\alpha_1, ..., \alpha_m)$ *the distribution function of Dirichlet density (1) with parameters* $\alpha_1, ..., \alpha_m$. *All that means that we consider* θ *a random variable with distribution* $\varphi(\theta)$ *and conditional given* θ, P *is a Dirichlet process with parameter* $\alpha(\theta, \cdot)$.

DEFINITION 3 (Antoniak, 1974)
We say that $X_1, ..., X_n$ *is a sample of size* $n \ge 1$ *drawn from the mixture of Dirichlet process P if for any given* $m \ge 1$ *and for each couple of measurable subsets of R* $(C_1, ..., C_n)$, $(A_1, ..., A_m)$ *it holds that*

$$\mathrm{P}(X_1 \in C_1, ..., X_n \in C_n|\theta, P(A_1), ..., P(A_m),\ \ P(C_1), ..., P(C_n)) = \prod_{j=1}^{n} P(C_j).$$

The meaning of this Definition is clear: conditional given θ, if a realization of P is known, then the random variables $X_1, ..., X_n$ are independent.

The mixture of Dirichlet process possesses the reproducibility property that is: if $X_1, ..., X_n$ is a sample from P, then the process $P|X_1, ..., X_n$ is still a mixture of Dirichlet process. Precisely, if $x_1, ..., x_n$ are the sample values from P then

$$\mathrm{P}(P(B_1) \leq y_1, ..., P(B_{m-1}) \leq y_{m-1}|x_1, ..., x_n) =$$

$$\int_R D(y_1, ..., y_{m-1}|\alpha(\theta, B_1) + n_1, ..., \alpha(\theta, B_m) + n_m) d\varphi(\theta|x_1, ..., x_n),$$

where n_i is the number of observations among $x_1, ..., x_n$, which belong to $B_i$and $\varphi(\theta|x_1, ..., x_n)$ represents the distribution of θ given $x_1, ..., x_n$.

For the expression of $\varphi(\theta|x_1, ..., x_n)$, let us suppose that the observations have been rearranged (note that because of a characteristic feature of the process, more observations may coincide with positive probability):

$$\begin{cases} x_1 & x_2 & \ldots & x_r \\ n_1 & n_2 & \ldots & n_r \end{cases} \qquad \sum_{i=1}^{r} n_i = n, \; r = \text{distinct values of } x_1, ..., x_n,$$

and assume μ a measure equals to the Lebesgue measure on R except in those points where $\alpha(\theta, \cdot)$ has a atom where μ concentrates a unit mass.

One has

$$d\varphi_\mu\ (\theta|x_1, ..., x_n)\ \propto \frac{1}{(\alpha(\theta, R))_n} \prod_{j=1}^{r} \alpha'_*(\theta, x_j)\ \left(m(\theta, x_j) + 1\right)_{n_j - 1}\ d\varphi_\mu(\theta),$$

where $(a)_n = a\,(a+1)\ldots(a+n-1)$, $\alpha'_*\left(\theta, x_j\right) = \alpha'\left(\theta, (-\infty, x_j]\right)$ is the Radon-Nikodym derivative of $\alpha$with respect to μ and

$$m(\theta, x_j) = \begin{cases} \alpha'_*(\theta, x_j) & \text{if } x_j \text{ is an atom of } \alpha \\ 0 & \text{elsewhere} \end{cases}.$$

A special case arises when $\alpha_*(\theta, x)$ and $\varphi(\theta)$ are absolutely continuous with respect to the Lebesgue measure. In this last case

$$d\ \varphi(\theta|x_1, ..., x_n)\ \propto \frac{1}{(\alpha(\theta, R))_n} \prod_{j=1}^{r} \alpha'_*(\theta, x_j)\ \left(n_j - 1\right)\ \varphi'(\theta) d\theta. \qquad (3)$$

It is well known that the predictive distribution (the probability law of a new observation given $X_1, ..., X_n$ can be viewed as a fitting of the empirical distribution $\hat{F}_n(\cdot)$ (De Finetti, 1935; Cifarelli, 1986) and it is the only inferential toolkit when parameters have no objective meaning.

In our case, from (2) we have, for each $t = 1, 2, 3, ...$

$$\mathrm{P}(X_{n+t} \leq x|x_1, ..., x_n) = E(P((-\infty, x])\,|x_1, ..., x_n) =$$

$$\int_R \left\{ \frac{\alpha(\theta, R)}{\alpha(\theta, R) + n}\ \frac{\alpha_*(\theta, x)}{\alpha(\theta, R)} + \frac{n}{\alpha(\theta, R) + n}\ \hat{F}_n(x) \right\} d\varphi_n(\theta|x_1, ..., x_n) \qquad (4)$$

where

$$\alpha(\theta, R) = \alpha(\theta, (-\infty, +\infty),$$
$$\alpha_*(\theta, x) = \alpha(\theta, (-\infty, x]), \text{ and}$$
$$\hat{F}_n(x) = \frac{1}{n}\sum_{j=1}^{n} c(x - x_j) = \text{empirical distribution function of } X_1, ..., X_n.$$

(being $c(x) = 1, \; x \geq 0, \quad c(x) = 0, \; x < 0$).

3. A BAYESIAN NON PARAMETRIC MODEL

Let a population that has a distribution function G and $X_1, ..., X_n$ a sample from G. Conditional given a parameter θ and G we suppose therefore that $X_1, ..., X_n$ are i.i.d. with distribution function G

$$\mathrm{P}(X_1 \leq x_1, ..., X_n < x_n | \theta, G) = \prod_{j=1}^{n} G(x_j).$$

Assume now that G is not known but random according to a Dirichlet process with parameter $\alpha_*(\theta, x) = \alpha(\theta, (-\infty, x])$ and θ random with distribution function $\varphi(\theta)$.

To make the realization of G close to the power function distribution we choose

$$\alpha_*(\theta, x) = \alpha(\theta, R) \; x^{\frac{\theta}{1-\theta}} \qquad 0 < x < 1, \quad 0 < \theta < 1,$$

and $\alpha(\theta, R) = \bar{\alpha}$ = constant independent of θ.

These hypothesis imply that

$$E(G(x)|\theta) = \frac{\alpha_*(\theta, x)}{\bar{\alpha}} = x^{\frac{\theta}{1-\theta}} = \mathrm{P}(X_j \leq x|\theta) \quad j = 1, 2, ..., n,$$

and therefore that the distribution of each X_j conditional given θ, has the power function distribution. But, note that even conditional given θ, $X_1, ..., X_n$, are not independent. Indeed one has:

$$\mathrm{P}(X_1 \leq x_1, ..., X_n < x_n|\theta) = \int_G \mathrm{P}(X_1 \leq x_1, ..., X_n < x_n|\theta, G) \; \mathrm{P}(dG|\theta) =$$

$$E\left(\prod_{j=1}^{n} G(x_j|\theta)\right) = \prod_{j=1}^{n} \frac{\alpha_*(\theta, x_{(j)}) + j - 1}{\bar{\alpha} + j - 1} = \prod_{j=1}^{n} \frac{\bar{\alpha} x_{(j)}^{\frac{\theta}{1-\theta}} + j - 1}{\bar{\alpha} + j - 1}, \tag{5}$$

where $0 \leq x_{(1)} \leq x_{(2)} \leq ... \leq x_{(n)} \leq 1$ are given by $x_1, ..., x_n$ in non decreasing order. $X_1, ..., X_n$ are therefore exchangeable (and not independent) given θ.

From (5) we also have:

$$\mathrm{P}(X_1 \leq x_1, ..., X_n < x_n) \prod_{j=1}^{n} \frac{1}{\bar{\alpha} + j - 1} \int_0^1 \bar{\alpha} x_{(j)}^{\frac{\theta}{1-\theta}} + j - 1 \ \ d\varphi(\theta).$$

Note that when the parameter $\bar{\alpha} \rightarrow +\infty$ then $X_1, ..., X_n$, given θ, are independent with the same probability law given by the power distribution function.

The magnitude of $\bar{\alpha}$ represents the strength of our conviction about the power function distribution and its nature is exquisitely subjective. Assume now that θ has a Beta density (a rather flexible prior) given by

$$\varphi'(\theta) \propto \theta^{p-1}(1-\theta)^{q-1} \qquad 0 < \theta < 1, \qquad p, q > 0.$$

Then the posterior of θ is given by (3) which becomes

$$\varphi'(\theta|x_1, ..., x_n) \propto \prod_{j=1}^{r} \left(\frac{\theta}{1-\theta}\right) x_j^{\frac{2\theta-1}{1-\theta}} \ \theta^{p-1}(1-\theta)^{q-1} \qquad 0 < \theta < 1,$$

$$\propto \left(\frac{\theta}{1-\theta}\right)^r \theta^{p-1}(1-\theta)^{q-1} \prod_{j=1}^{r} x_j^{\frac{2\theta-1}{1-\theta}} \qquad 0 < \theta < 1,$$

where $x_1, ..., x_r$ are the distinct values of $x_1, ..., x_n$.

After some algebra we get

$$\varphi'(\theta|x_1, ..., x_n) \propto \left(\frac{\theta^{r+p-1}}{(1-\theta)^{r-q+1}}\right) e^{-\frac{\theta}{1-\theta}t} \qquad t = -\sum_{j=1}^{r} \log x_j > 0,$$

and finally

$$\varphi'(\theta|x_1, ..., x_n) = \frac{\left(\frac{\theta^{r+p-1}}{(1-\theta)^{r-q+1}}\right) e^{-\frac{\theta}{1-\theta}t}}{\int_0^1 \frac{\theta^{r+p-1}}{(1-\theta)^{r-q+1}} \ e^{-\frac{\theta}{1-\theta}t} d\theta} \qquad 0 < \theta < 1,$$

$t > 0 \quad p, q > 0 \quad r \geq 1.$

Remark is in order about φ'. If $t = 0$ and therefore $r = 1$ one has

$$\varphi(z|x_1, ..., x_n) = 1 \quad z \geq 1,$$

$$\varphi(z|x_1, ..., x_n) = 0 \quad z < 1 \ , q \leq 1,$$

$$\varphi'(\theta|x_1, ..., x_n) = \frac{\theta^p(1-\theta)^{q-2}}{\mathrm{B}(p+1, q-1)} \quad 0 < \theta < 1 \quad q > 1.$$

So when $t = 0$ and $q \leq 1$, no density exists for θ and the posterior concentrates all its unity mass at the unique point $\theta = 1$.

The Bayes estimate of θ, with respect to squared error loss function, is the posterior mean which is given by

$$ {}^{p,q}\hat{\theta} := {}^{p,q}\hat{\theta}_r(t) = E(\theta|x_1, \dots, x_n) = \int_0^1 \theta\varphi'(\theta|x_1, \dots, x_n)\, d\theta = $$

$$ = \frac{\int_0^1 \frac{\theta^{r+p}}{(1-\theta)^{r-q+1}} e^{-\frac{\theta}{1-\theta}t}\, d\theta}{\int_0^1 \frac{\theta^{r+p-1}}{(1-\theta)^{r-q+1}} e^{-\frac{\theta}{1-\theta}t} d\theta} \qquad t > 0\ p, q > 0 r \geq 1. $$

If $t = 0$

$$ \hat{\theta}_1(0) = 1 \qquad q \leq 1, $$
$$ \hat{\theta}_1(0) = \frac{p+1}{p+q} \qquad q > 1. $$

If we make the change of variable $x = \frac{\theta}{1-\theta} t$ in both of the integrals we obtain

$$ {}^{p,q}\hat{\theta} = E(\theta|x_1, \dots, x_n) = \frac{\int_0^\infty \frac{x^{r+p}}{(t+x)^{r+q+1}} e^{-x}\, dx}{\int_0^\infty \frac{x^{r+p-1}}{(t+x)^{r+q}} e^{-x}\, dx}, $$

$$ = 1 - t\frac{\int_0^\infty \frac{x^{r+p-1}}{(t+x)^{r+q+1}} e^{-x}\, dx}{\int_0^\infty \frac{x^{r+p-1}}{(t+x)^{r+q}} e^{-x}\, dx}, $$

$$ = 1 + t\frac{d}{dt}\log\left(\int_0^\infty \frac{x^{r+p-1}}{(t+x)^{p+q}} e^{-x}\, dx\right); \quad t > 0, \quad p, q > 0. \tag{6} $$

The integral may be computed by means of a confluent hypergeometric function

$$ \int_0^\infty \frac{x^{r+p-1}}{(t+x)^{p+q}} e^{-x}\, dx = \Gamma(r+p)t^{r-q}\psi(r+p; r-q+1; t). $$

A few important cases arise when one takes some particular values for p and q.

1) The prior for θ is Jeffrey's non informative that is $p = q = 0$.
In such a case

$$ {}^{0,0}\hat{\theta}_r(t) = \frac{1}{\Gamma(r)} \int_0^\infty \frac{x^r}{t+x} e^{-x} dx = rt^r e^t \Gamma(-r, t) = re^t E_{r+1}(t) r \geq 1 t > 0, $$

where $\Gamma(-r, t)$ and $E_{r+1}(t)$ are the incomplete gamma function and the exponential integral. To evaluate such an expression we may use the recurrence relation (Abramowitz, Stegun, 1965, p. 228)

$$ E_{r+1}(t) = \frac{1}{r}(e^{-t} - tE_r(t)), $$

from which

$$ {}^{0,0}\hat{\theta}_r(t) = 1 - \frac{t}{r-1}\theta_{r-1}(t), r \geq 2; $$

$$ \hat{\theta}_1(t) = 1 - te^t E_1(t). $$

And so

$$ {}^{0,0}\hat{\theta}_r(t) = \sum_{k=0}^{r-1} \frac{(-1)t^k}{(r-1)^{(k)}} + \frac{(-1)^r t^r e^t}{(r-1)} E_1(t), \tag{7} $$

where E_1 may be approximated by

$$ E_1(t) = \sum_{q=0}^{5} a_q t^q - \log t + \varepsilon(t); \ |\varepsilon(t)| < 2 \cdot 10^{-7} \qquad 0 \leq t \leq 1, $$

$$ a_0 = -0.57721,\ a_1 = 0.99999,\ a_2 = -0.24991, $$

$$ a_3 = 0.5519,\ a_4 = -0.00936,\ a_5 = 0.00107, $$

$$ te^t E_1(t) = \frac{t^2 + b_1 t + b_2}{t^2 + c_1 t + c_2} + \varepsilon(t); \qquad |\varepsilon(t)| < 5 \cdot 10^{-5} \qquad 1 \leq t < +\infty, $$

$b_1 = 2.33473$, $b_2 = 0.25062$, $c_1 = 3.33065$, $c_2 = 1.68153$, (Abramowitz, Stegun, 1965, p. 231).

A good approximation also is

$$ \frac{1}{t+r+1} < e^t E_{r+1} \leq \frac{1}{t+r}, \qquad t > 0, \qquad r \geq 1, $$

and

$$ \frac{r}{t+r+1} < {}^{0,0}\theta_r(t) \leq \frac{r}{t+r}, \qquad t > 0,\ r \geq 1 \tag{8} $$

2) The prior for θ is uniform on $[0,1]$ that is $p=q=1$

$$ {}^{1,1}\hat{\theta}_r(t) = \frac{\int_0^\infty \frac{x^{r+1}}{(t+x)^3} e^{-x}\,dx}{\int_0^\infty \frac{x^r}{(t+x)^2} e^{-x}\,dx} = 1 + t\frac{d}{dt}\log\int_0^\infty \frac{x^r}{(t+x)^2} e^{-x}\,dx = -r\frac{{}^{00}\hat{\theta}''_{r+1}(t)}{{}^{00}\hat{\theta}'_r(t)}, \quad (9) $$

where ${}^{00}\hat{\theta}_r(t)$ is given by (7).

As regards to these estimates we have to do a remark. It concerns the role of $\bar{\alpha}$. It seems that $\hat{\theta}$ does not depend on $\bar{\alpha}$, the weight assigned to the power function distribution. But $\bar{\alpha}$ affects the number of distinct observations in the sense that if $\bar{\alpha} \to +\infty$, then, $r \to n$ almost surely (Korwar, Hollander, 1973) and in such a case our estimates coincide with those given by Alodat *et al*. (2008).

4. THE PREDICTIVE DISTRIBUTION

The predictive distribution is a very useful tool for many purposes especially from a theoretical viewpoint and as it was said it also can be viewed as a fitting of the empirical distribution $\hat{F}_n(\cdot)$.

In our case, from (4) we have

$$ \mathrm{P}(X_{n+k} \le x | x_1, \ldots, x_n) = \int_0^1 \left\{ \frac{\bar{\alpha}}{\bar{\alpha}+n}\left(\frac{\alpha_*(\theta,x)}{\bar{\alpha}}\right) + \frac{n}{\bar{\alpha}+n}\hat{F}_n(x)\right\}\varphi'(\theta|x_1,\ldots,x_n), $$

where φ' is one of the previous densities:

$$ \mathrm{P}(X_{n+k} \le x | x_1, \ldots, x_n) = \frac{\bar{\alpha}}{\bar{\alpha}+n}\int_0^1 \frac{\alpha_*(\theta,x)}{\bar{\alpha}}\varphi'(\theta|x_1,\ldots,x_n) + \frac{n}{\bar{\alpha}+n}\hat{F}_n(x), \quad k \ge 1. $$

It appears as a mixture with weights $\frac{\bar{\alpha}}{\bar{\alpha}+n}$ and $\frac{n}{\bar{\alpha}+n}$ between the mixture of the prior guess and the empirical evidence.

For Jeffrey's prior we have

$$ \frac{\alpha_*(\theta,x)}{\bar{\alpha}} = x^{\frac{\theta}{1-\theta}}, $$

$$ \varphi'(\theta|x_1,\ldots,x_n) = \frac{\theta^{r-1}}{(1-\theta)^{r+1}} e^{-t\frac{\theta}{1-\theta}} \frac{t^r}{\Gamma(r)}, $$

$$ \mathrm{P}(X_{n+k} \le x | x_1, \ldots, x_n) = \frac{\bar{\alpha}}{\bar{\alpha}+n}\int_0^1 x^{\frac{\theta}{1-\theta}} \frac{t^r}{\Gamma(r)} \frac{\theta^{r-1}}{(1-\theta)^{r+1}} e^{-t\frac{\theta}{1-\theta}} d\theta + $$

$$ + \frac{n}{\bar{\alpha}+n}\hat{F}_n(x), \quad k \ge 1, $$

and after the integral calculation we have

$$P(X_{n+k} \leq x|x_1, ..., x_n) = \frac{\bar{\alpha}}{\bar{\alpha}+n}\left(\frac{t}{t-\log x}\right)^r + \frac{n}{\bar{\alpha}+n}\hat{F}_n(x),$$
$$0 < x < 1, \quad t > 0. \tag{10}$$

Of course no ordinary predictive density exists and, for $k = 1, 2, ...$

$$P(X_{n+k} \leq x|x_1, ..., x_n) = \left(\frac{t}{t-\log x}\right)^n \qquad \bar{\alpha} \to \infty,$$

$$P(X_{n+k} \leq x|x_1, ..., x_n) = \hat{F}_n(x) \qquad \bar{\alpha} \to 0.$$

To fix $\bar{\alpha}$ we may think of it as ‘‘a prior sample size’’ comparing it with the actual sample size. For instance, if we give the same ‘‘importance’’ to $\bar{\alpha}$ and n we may fix

$$\frac{\bar{\alpha}}{\bar{\alpha}+n} = \frac{1}{2}.$$

Similar results are obtained for uniform prior or for other priors of θ. We shall not pursue such a search.

5. APPLICATIONS

We use the data explained by Alodat *et al.* (2008) (Table 1).

TABLE 1. - *Data for Dactylis glomerata*

x_i	**DATA**	x_i	**DATA**	x_i	**DATA**	x_i	**DATA**
x_1	0.3710	x_7	0.0508	x_{13}	0.0380	x_{19}	0.0010
x_2	0.3100	x_8	0.0790	x_{14}	0.0250		
x_3	0.1460	x_9	0.0150	x_{15}	0.0440		
x_4	0.0280	x_{10}	0.0110	x_{16}	0.0630		
x_5	0.0620	x_{11}	0.0190	x_{17}	0.1580		
x_6	0.0230	x_{12}	0.0001	x_{18}	0.0020		

We have $r = n = 19$ and $t = -\sum_{i=1}^{19} \log x_i = 69.281822$.

Our estimates for θ in the case of Jeffrey’s non informative and uniform, given by (7) and (9) are respectively

$${}^{0,0}\hat{\theta}_{19}(t) \approx 0.213335; \qquad {}^{1,1}\hat{\theta}_{19}(t) \approx 0.218324.$$

Of course they do not differ from those of Alodat *et al.* (2008) (due to the equality $r = n = 19$).

As to the 95% credibility interval of X_{n+k}, $k = 1, 2, \ldots$ we have to solve (for x)

$$P(X_{n+k} \leq x | x_1, \ldots, x_n) = 0.025,$$
$$P(X_{n+k} \leq x | x_1, \ldots, x_n) = 0.975.$$

Of course it depends on $\bar{\alpha}$ which we are free to fix according to our degree of faith in the prior guess. If we set $\frac{\bar{\alpha}}{\bar{\alpha}+n} = 0.90$ we get the interval (with Jeffrey's prior) $(5.7033 \times 10^{-7},\ 0.9022)$.

Table 2 shows various credibility intervals for a few values of $\frac{\bar{\alpha}}{\bar{\alpha}+19}$.

TABLE 2. - *Credibility intervals*

$\frac{\bar{\alpha}}{\bar{\alpha}+19}$	*Credibility intervals*
$\to 1$	$(3.5693 \times 10^{-7},\ 0.9117)$
0.90	$(5.7033 \times 10^{-7},\ 0.9022)$
0.80	$(9.5310 \times 10^{-7},\ 0.8906)$
0.10	$(10^{-4},\ 0.371)$

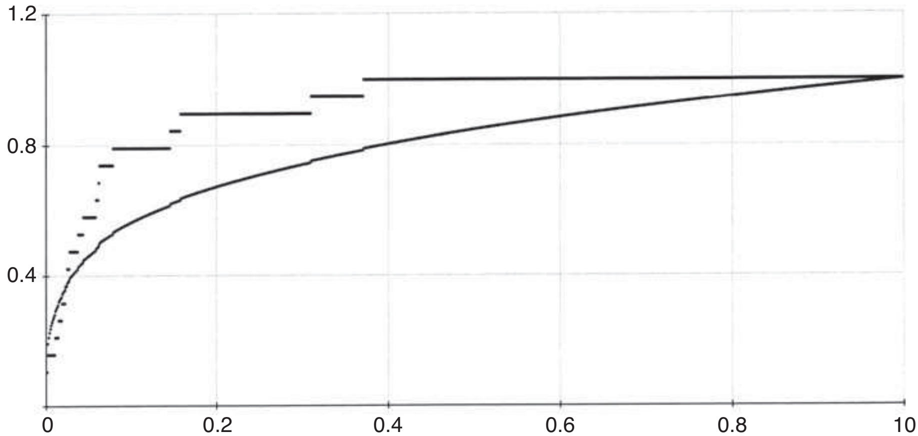

FIGURE 1. - *The step function is the empirical distribution and the other one is the predictive distribution*

Figure 1 shows the empirical distribution function and the predictive distribution function (as a fitting of the empirical distribution function) for $\frac{\bar{\alpha}}{\bar{\alpha}+19} = 0.90$. If we take $\bar{\alpha} \to 0$ then, of course, the predictive becomes closer and closer to the empiric distribution.

AKNOWLEDGEMENT

I am very grateful to D.M. Cifarelli and the reviewer for their helpful comments and suggestions.

RIASSUNTO

Nel presente lavoro viene dedotto la stima non parametrica bayesiana di un parametro reale relativa ad una misura che, a sua volta, costituisce la funzione parametro di una mistura di processi di Dirichlet, e l'intervallo di credibilità relativo alla distribuzione predittiva. L'approccio viene poi applicato ai dati riguardanti l'abbondanza delle specie di piante considerati da Alodat, Odat, Muhaidat, Beldjillali (2008).

REFERENCES

Alodat M.T., Odat N., Muhaidat R.A., Beldjillali H. (2008). Estimation of power function distribution with application to ecological relative abundance. *Statistica & Applicazioni*, **6**(2), 181-192.

Abramowitz M., Stegun I. (1965). *Handbook of mathematical function*. Dover Publications, New York.

Antoniak C.E. (1974). Mixture of Dirichlet process with applications to Bayesian nonparametric problems. *The Annals of Statistics*, **6**, 1152-1174.

Cifarelli D.M. (1986). Ancora sul problema della perequazione dal punto di vista induttivo, Scritti in onore di F. Brambilla. *Edizioni di "Bocconi Comunicazione*", **1**, 167-179.

De Finetti B. (1935). Il problema della perequazione. *Atti della Società Italiana per il Progresso delle Scienze*, **1**(2), 227-228.

Ferguson T.S. (1973). A Bayesian analysis of some non parametric problems. *The Annals of Statistics*, **1**, 209-230.

Korwar R.M., Hollander M. (1973). Contribution to the theory of Dirichlet process. *The Annals of Probability*, **1**, 705-711.

Rodriguez A., Dunson D.B., Gelfand A.E. (2008). The nested Dirichlet process. *Journal of the American Statistical Association,* **103**(483), 1131-1154.

Schervish M.J. (1995). *Theory of Statistics*. Springer-Verlag, New York.