THE DEFINITION OF A DIRICHLET PROCESS

H.LUO

Abstract. The proof ans motive of the Dirichlet process is scattered around in [1] and is not very accessible via [4], this short note explains explicitly and rigidly how a Dirichlet process can be defined on $M(\mathbb{R})$.

We must assume the following result, whose proof can be founded in [3]. This result told us that if we specified "reasonable" marginal distributions to subsequence of an infinite random variable sequence, we can expect there exists some sort of distribution of the distribution***s*** of X_n as $n \to \infty$. This is a special case of Ergodic Theorem.

Among all those specifications of marginal distributions, Ferguson [4] investigate a most natural marginals in nonparametric settings.

Theorem 1. *(Kolmogorov's Extension Theorem) Let $T \subset \mathbb{R}$, for a string $\omega = (t_1, t_2, \cdots, t_k), t_i \in T$ we define a probability measure ν_ω on $((\mathbb{R}^n, \mathcal{B}^n, P))^k$ such that v_ω satisfy following "**consistency conditions**":*

(i) $v_\omega(B_1, B_2, \cdots B_k) = \nu_{\pi(\omega)}(B_{\pi(1)}, B_{\pi(2)}, \cdots B_{\pi(k)}) \forall \pi \in Perm(k), B_i \in \mathcal{B}^n$

(ii)$v_\omega(B_1, B_2, \cdots B_k) = \nu_{\bar{\omega}}(B_1, B_2, \cdots B_k, \mathbb{R}^n, \cdots \mathbb{R}^n) \forall \bar{\omega} = (\omega, t_{k+1}, \cdots t_N), B_i \in \mathcal{B}^n$

Then there exists a stochastic process $X(t, \bullet)$ defined on a probability space $(\mathcal{X}, \mathcal{A}, P)$with values in $\mathbb{R}^n$ that $X : T \times \mathcal{X} \to \mathbb{R}^n$ and

$$v_\omega(B_1, B_2, \cdots B_k) \overset{d}{=} P(X_{t_1} \in F_1, \cdots X_{t_k} \in F_k), B_i \in \mathcal{B}^n$$

.

We knew from Egorov's Theorem that once we define a real valued (Lebesgue) measurable function on a dense set of $\mathbb{R}^n$ or complete separable metric space $\mathcal{X}$, then we can determine it up uniformly to a zero (Lebesgue) measure set. Therefore to define a probability measure on $M(\mathcal{X})$ it suffices to define a probability measure on a dense set $\mathcal{F}^* = [0,1]^{\mathcal{X}}$ of $M(\mathcal{X})$.

This extension theorem told us that we can always find a probability space to "adapt" a finite sequence as a marginal. Thus, we yield Theorem 2.3.2 in [1].

In classical statistic literature, we usually assume $X \sim P_X$ as random variable, and $X_1, \cdots X_n$ are i.i.d. samples from a fixed probability measure P_X. Then we can talk about the distribution of $(x_1, \cdots x_n | P_X)$ as a random vector because each X_i is of stochastic nature. Now we fix the value of $X_i = x_i$, and suppose that we randomly choose a P_X from a collection of probability measures, say $M(\mathcal{X})$. Then we can talk about the distribution of $(P(x_1), \cdots P(x_n))$ because P is now randomly chosen. If we know the random rule Π of choosing a P, then an appropriate notation should be $((P(x_1), \cdots P(x_n)|\Pi))$. In order to distinguish from the usual notation, since $\Pi \in M(M(\mathcal{X}))$, we use $\mathcal{L}((P(x_1), \cdots P(x_n)|\Pi))$.

This notation is not chosen arbitrarily. $\mathcal{L}$ stands for a linear functional of a linear functional in functional analysis. Please refer to the following diagram for relation and clarification.

original space	*dual space*	*double dual space*
$x \in S$	$f \in S^*$	$f_{x_0} : f \mapsto f(x_0) \in S^{**}$
$Events \in \mathcal{X}$	$Probability\, measures \in M(\mathcal{X})$	$Priors \in M(M(\mathcal{X}))$
		$Posteriors \in$

Please compare the following theorem with the Kolmogorov's Extension Theorem and see what is the difference between these conditions.

Theorem 2. *([1] p.65) Let $Q \subset \mathbb{R}$ be a* ***countable dense*** *subset, for a string $\omega = (t_1, t_2, \cdots, t_k), t_i \in Q$ we define a probability measure ν_ω on $\left(\left([0,1]^n, \mathcal{B}^n_{[0,1]}, P_{[0,1]}\right)\right)^k$ and a c.d.f. F associated with ν_ω such that:*

(i)$t_i < t_j \Rightarrow \nu_{t_i,t_j}\{F(t_i) \le F(t_j)\} = 1$

(ii)$v_\omega(B_1, B_2, \cdots B_k) = \nu_{\bar\omega}(B_1, B_2, \cdots B_k, \mathbb{R}^n, \cdots \mathbb{R}^n) \forall \bar\omega = (\omega, t_{k+1}, \cdots t_N), B_i \in \mathcal{B}^n_{[0,1]}$

(iii)$lim_{n\to\infty}\left(t_1^{(n)}, t_2^{(n)}, \cdots t_k^{(n)}\right) \downarrow (t_1, t_2, \cdots t_k) \Rightarrow lim_{n\to\infty}\nu_{\left(t_1^{(n)}, t_2^{(n)}, \cdots t_k^{(n)}\right)} \overset{d}{\to} \nu_{(t_1,t_2,\cdots t_k)}$

(iv)$lim_{n\to\infty}t_n \downarrow -\infty \Rightarrow lim_{n\to\infty}\nu_{t_n} \overset{d}{\to} 0$;$lim_{n\to\infty}t_n \uparrow \infty \Rightarrow lim_{n\to\infty}\nu_{t_n} \overset{d}{\to} 1$

Then there exists a $\Pi \in M(M(\mathbb{R}))$ defined on $M(\mathbb{R})$ such that

$$\mathcal{L}\left((F(t_1), \cdots F(t_n)|\Pi)\right) \overset{d}{=} \nu_\omega, \forall t_i < t_{i+1}, (t_1, t_2, \cdots, t_k), t_i \in Q$$

.

Proof. Theorem 1 ensures that there exists Π on $[0,1]^Q$ with ν_ω as marginals.

Let $\mathcal{F}^*$ be the collection of all c.d.f.s defined on Q., following steps lead to conclusion. □

(1) Π is a probability measure on $\mathcal{F}^*$.
   (a) Since Π is already a stochastic process over $[0,1]^Q$, the definition of a stochastic process ensures that it is countably additive.
   (b) We prove $\Pi(\mathcal{F}^*) = 1$ by proving that it is a closure of a dense set of measure 1 over $[0,1]^Q$.
      (i) $\mathcal{F}_1^* := \cap_{t_i<t_j}\left\{F \in [0,1]^Q | F(t_i) \le F(t_j)\right\}$, by (ii), $\nu_{t_i,t_j}\{F(t_i) \le F(t_j)\} = 1$ and $\Pi(\mathcal{F}_1^*) = 1$ by filling the Q with countable sequences ω. For example if we use the binaries, we can fill the Q using $0, 1, 00, 01, 11, 10\cdots$
      (ii) For a fixed $t \in Q$, choose a fixed descending sequence $\{t_n\}$ which converges to t descendingly. By (iii), $lim_{n\to\infty}t_n \downarrow t, t \in Q \Rightarrow lim_{n\to\infty}\nu_{t_n} \overset{d}{\to} \nu_t$. And $lim_{n\to\infty}F(t_n) \to F^*(t)$ for some $F^* \in \mathcal{F}_1^*$due to the monotonic convergence theorem.
      (iii) Consequently, $F^*(t) \ge F(t)$ since $F^* \in \mathcal{F}_1^*$ implies it is right-continuous. And $\mathbb{E}_\Pi F^*(t) = \mathbb{E}_\Pi F(t)$ by (iii), then $F^*(t) = F(t) a.e. \Pi, \forall t \in Q$ because they are c.d.f.s.

(iv) Therefore

$$\Pi\left\{F \in \mathcal{F}_1^* : F\, is\, right\, continuous\, at\, t\right\} = 1$$

by previous two lines.

(v) The countability of Q mentioned above leads to

$$\Pi\left\{F \in \mathcal{F}_1^* : F\, is\, right\, continuous\, AND\, monotone\, at\, \forall t \in Q\right\} = 1$$

since we can always find a descending sequence for $\forall t \in Q$ and repeat the arguments in previous three lines.

(vi) A symmetric argument of choosing an ascending/descending sequences to $\pm\infty$ will lead to the

$$\Pi\left\{F \in \mathcal{F}_1^* : lim_{t\to\infty} F(t) = 1\, AND\, lim_{t\to-\infty} F(t) = 0\, at\, \forall t \in Q\right\} = 1$$

.

(vii) With these properties, we assert that $\Pi(\mathcal{F}^*) = 1$ because the closure of $\mathcal{F}_1^*$ is exactly $\mathcal{F}^*$ and $\Pi(\mathcal{F}^*) = \Pi(\overline{\mathcal{F}^*}) \geq \Pi(\mathcal{F}_1^*) = 1$.

(2) By Egorov's theorem, any defined $F \in \mathcal{F}^*$ can be extended to $\mathcal{F}$, the collection of all c.d.f.s defined over $\mathbb{R}$ except for a zero $P_{[0,1]}$ measure set.

(3) By duality, any real-valued c.d.f. on $\mathbb{R}$ corresponds to an element in $M(\mathbb{R})$. So we have derived a distribution for any element class in $M(\mathbb{R})$, such an element class is the equivalent class in $M(\mathbb{R})$ defined by $P_1 \sim P_2 \Leftrightarrow \int_{-\infty}^{t} \frac{dP_1}{dx}(dx) = \int_{-\infty}^{t} \frac{dP_2}{dx}(dx), \forall t \in \mathbb{R}$.

(4) By the definition of such an equivalence, $\mathcal{L}\left((F(t_1), \cdots F(t_n)|\Pi)\right) \overset{d}{=} \nu_\omega$. Q.E.D.

This theorem told us that there exists a specific type of prior, or elements of distributions on $M(\mathbb{R})$, such that if we prescribe such prior then we may yield a posterior of ν_ω. A limitation of this result is that $t_i \in Q$, if $Q = \mathbb{Q}$ is the rational number sets, then a possibility is that we cannot track the behavior of the "time gaps" between t_i, t_{i+1} although we can approximate it as close as we want. Following result overcomes this limitation.

Theorem 3. *([1] p.67) Let $\mathbb{R}$ be the real line, for a string $\omega = (t_1, t_2, \cdots, t_k), t_i \in \mathbb{R}$ we define a probability measure ν_ω on $S^k := \{(p_1, \cdots p_k) : p_i \geq 0, \sum_i p_1 \leq 1\}$ such that:*

(i)$v_\omega(B_1, B_2, \cdots B_k) = \nu_{\bar{\omega}}(B_1, B_2, \cdots B_k, \mathbb{R}^n, \cdots \mathbb{R}^n) \forall \bar{\omega} = (\omega, t_{k+1}, \cdots t_N), B_i \in \mathcal{B}_{S_k}^n$

(ii)$lim_{n\to\infty}\left(t_1^{(n)}, t_2^{(n)}, \cdots t_k^{(n)}\right) \downarrow (t_1, t_2, \cdots t_k) \Rightarrow lim_{n\to\infty} \nu_{\left(t_1^{(n)}, t_2^{(n)}, \cdots t_k^{(n)}\right)} \overset{d}{\to} \nu_{(t_1, t_2, \cdots t_k)}$

(iii)$lim_{n\to\infty} t_n \downarrow -\infty \Rightarrow lim_{n\to\infty} \nu_{t_n} \overset{d}{\to} 0$;$lim_{n\to\infty} t_n \uparrow \infty \Rightarrow lim_{n\to\infty} \nu_{t_n} \overset{d}{\to} 1$

Then there exists a $\Pi \in M(M(\mathbb{R}))$ defined on $M(\mathbb{R})$ such that

$$\mathcal{L}\left((F(t_1), F(t_2) - F(t_1), \cdots F(t_n) - F(t_{n-1})|\Pi)\right) \overset{d}{=} \nu_\omega, \forall t_i < t_{i+1}, (t_1, t_2, \cdots, t_k), t_i \in \mathbb{R}$$

.

Proof. Because the "gaps" can be regarded as half-open intervals $(-\infty, t_1], (t_1, t_2], \cdots, (t_k, \infty)$, so instead of specifying ν_ω using sequences we can equivalently specify ν_ω using a collection of intervals $(-\infty, t_1], (t_1, t_2], \cdots, (t_k, \infty)$. Thinking of the 1-1 correspondence $(F(t_1), \cdots, F(t_n)) \longleftrightarrow (F(t_1), F(t_2) - F(t_1), \cdots, F(t_n) - F(t_{n-1}))$ and consider the probability measure $\nu_{(-\infty, t_1], (t_1, t_2], \cdots, (t_k, \infty)}$. For these $(k+1)$

cells consider a $(k+1)$ probability vector $(p_1, p_2, \cdots p_{k+1})$ on S_k and construct the a multinomial distribution ν_ω whose parameter is given by $(p_1, p_2, \cdots p_{k+1})$, we can see that □

(1) We can verify that if the specification of ν_ω satisfies (i)(ii)(iii) of Theorem 3, then ν_ω must satisfy (ii)(iii)(iv) of Theorem 2. And (i) of Theorem 2 is automatically satisfied due to our 1-1 representation. The arguments in Theorem 2 directly applies to $\mathbb{R}$ instead of Q because Step 2 in its proof. Egorov Theorem allows us to choose the values on a zero P measure set arbitrarily. Such a choice will not affect out conclusion since it will not be reflect in c.d.fs, whose random behavior is all we care about.

(2) The multinomial distribution $Multi\,(p_1, p_2 \cdots p_k, p_{k+1})$ with density $\frac{1}{\frac{\Gamma(n_1+n_2+\cdots+n_k+n_{k+1})}{\Gamma(n_1)\cdots\Gamma(n_k)\Gamma(n_{k+1})}} \left(\prod_{i=1}^k p_i^{n_i}\right)\left(\left[1-\sum_{i=1}^k p_i\right]\right)^{n_{k+1}}$ will be the Π indicated by the conclusion of Theorem 2, where $p_i = \begin{cases} F(t_1) & i=1 \\ F(t_i)-F(t_{i-1}) & i=2,\cdots k \\ 1-\sum_{i=1}^k p_i) & i=k+1 \end{cases}$. This can be derived either from Ergodic Decomposition Theorem[1][2] or from Polya's urn scheme.

The difference between this "continuous" case and the "discrete" case in Theorem 2 is analogous to the distribution of order statistic vector for continuous and discrete distributions. The condition (i) in Theorem 3 is sometimes referred as "consistency conditions" because it is sometimes possible that the marginal distribution of $\nu_{\overline{\omega}}$ will not correspond to ν_ω, an intuitive example is the degenerated Gaussian distribution in a higher dimension will project to a single point in a lower dimension instead of a Gaussian of lower dimension.

Now we extend to the following result which is necessary for defining a Dirichlet process in [1].

Theorem 4. *([1] p.68) Let $\mathbb{R}$ be the real line, for a string of* ***disjoint*** *Borel sets $\Omega = (B_1, B_2, \cdots, B_k), B_i \in \mathcal{B}$ we define a probability measure ν_Ω on $S^k := \{(p_1, \cdots p_k) : p_i \geq 0, \sum_i p_i \leq 1\}$ such that:*

(i) For any string of ***disjoint*** *Borel sets $\Omega_{sub} = (A_1, A_2, \cdots, A_l), A_j \in \mathcal{B}$ where A_j are union of sets from the collection $\{B_1, B_2, \cdots, B_k\}$,*

$$\nu_{\Omega_{sub}} \overset{d}{=} \mathcal{L}\left(\left(\sum_{B_i \subset A_1} P(B_i), \cdots \sum_{B_i \subset A_l} P(B_i) | \Pi\right)\right)$$

.

(ii)$lim_{n\to\infty} B_n \downarrow \emptyset \Rightarrow lim_{n\to\infty} \nu_{B_n} \overset{d}{\to} 0$ (zero measure);$lim_{n\to\infty} B_n \uparrow \mathbb{R} \Rightarrow lim_{n\to\infty} \nu_{B_n} \overset{d}{\to} 1$ (trivial measure)
Then there exists a $\Pi \in M(M(\mathbb{R}))$ defined on $M(\mathbb{R})$ such that
$\mathcal{L}\left((P(B_1), \cdots P(B_k)|\Pi)\right) \overset{d}{=} \nu_\Omega, \forall$string of ***disjoint*** *Borel sets*
$\Omega = (B_1, B_2, \cdots, B_k), B_i \in \mathcal{B}$.

Proof. The following steps lead to conclusion. □

[1] Note that we are not claiming $t_1, \cdots t_k$ are exchangeable but $F(t_1), \cdots, F(t_n)$ are exchangeable

[2] Actually, Ergodic Decomposition Theorem can be yielded from Kolmogorov's Extension Theorem, so they are somehow equivalent.

(1) It suffices to consider only B_i of the form $(t_i, t_{i+1}]$ since they are dense in $\mathcal{B}$ so the difference will only lie in a zero measure set, although we know that not all Borel sets are in the $\sigma-$algebra generated by these kind of intervals.
(2) Let $P(B_i) = F(t_{i+1}) - F(t_i)$ for some c.d.f. on $\mathbb{R}$, recalling that such a c.d.f. always exists, then the existence of Π such that is ensured by Theorem 3 as long as we verify the assumptions of Theorem 3.
   (a) (i) of Theorem 3 is guaranteed by (i) of Theorem 4, simply noticing that B_i can be chosen as $\mathbb{R}$.
   (b) (ii) and (iii) of Theorem 3 is a special case of (ii) of Theorem 4, by recognizing that sets of finitely many points are Borel on $\mathbb{R}$.
(3) Establish the monotonic class, i.e. prove the limit of probability measure sequence is still in this class.
   (a) Given $lim_{n\to\infty}\left(B_1^{(n)}, \cdots B_k^{(n)}\right) \downarrow (B_1, \cdots B_k)$ by previous step a correspondence $P(B_i) = F(t_{i+1}) - F(t_i)$ gives $lim_{n\to\infty}\left(P(B_1^{(n)}), \cdots P(B_k^{(n)})\right) \overset{d}{=} (P(B_1), \cdots P(B_k))$, we can rewrite it as

$$lim_{n\to\infty}\left(P(B_1^{(n)}) - P(B_1) + P(B_1), \cdots P(B_k^{(n)}) - P(B_k) + P(B_k)\right) \downarrow (P(B_1), \cdots P(B_k))$$

   by countable additivity.
   (b) For each i, $B_i^{(n)} - B_i \downarrow \emptyset$, and this means $P(B_i^{(n)}) - P(B_i) \to 0$ in distribution, by Slutsky Theorem, in probability too. Thus

$$lim_{n\to\infty}\left(P(B_1^{(n)}) - P(B_1), \cdots P(B_k^{(n)}) - P(B_k)\right) \downarrow (0, \cdots 0)$$

(4) Induction on the components of $(P(B_1), \cdots P(B_k))$ using monotonic class arguments.
   (a) Let $C_i := (a_i, b_i]\,, i = 2, \cdots k$, and denote that

$$\mathcal{B}_1 := \left\{B_1 : \mathcal{L}\left((P(B_1), P(C_2) \cdots P(C_k)|\Pi)\right) \overset{d}{=} \nu_{B_1, C_2, \cdots C_k}\right\}$$

   .
   (b) It is obvious that $\mathcal{B}_1$ contains all B_i of the form $(t_i, t_{i+1}]$ due to the arguments provided in Step 2. But the problem is we still do not know whether $\mathcal{B}_1$ contains all Borel sets because the representation $P(B_i) = F(t_{i+1}) - F(t_i)$ only holds for B_i of the form $(t_i, t_{i+1}]$.
   (c) Now notice that Step 3 already established that $\mathcal{B}_1$ is a monotonic class and that told us $\mathcal{B} \subset \mathcal{B}_1$ because $\mathcal{B}$ is the smallest σ-algebra that contains all intervals of the form $(a_i, b_i]$ by definition.
   (d) Try to consider

$$\mathcal{B}_2 := \left\{B_2 : \mathcal{L}\left((P(B_1), P(B_2), P(C_3) \cdots P(C_k)|\Pi)\right), \forall B_1 \in \mathcal{B} \overset{d}{=} \nu_{B_1, B_2, C_3 \cdots C_k}\right\}$$

   , using induction will formalize this argument in finite k case.

Finally, we can establish the existence theorem for the Dirichlet process. Loosely speaking, a Dirichlet process is a random measure whose marginals on finite Borel partition is *Dirichlet distribution* of finite dimension.

Theorem 5. *([1] p.96) Let α be a finite measure on $(\mathbb{R}, \mathcal{B})$. Then there* ***exists a unique*** *probability measure $D_\alpha \in M(M(\mathbb{R}))$ on $M(\mathbb{R})$ with parameter α such that For every partition $\{B_1, B_2, \cdots B_k\}, B_i \in \mathcal{B}$ of $\mathbb{R}$,*

$$\nu_\Omega = Dirichlet_k\left(\alpha(B_1), \cdots \alpha(B_k)\right) \overset{d}{=} \mathcal{L}\left((P(B_1), \cdots P(B_k)|D_\alpha)\right)$$

*D_α is known as **Dirichlet process** with respect to α.*

Proof. The Dirichlet distribution of finite dimension k is defined by following density

$$\frac{1}{\frac{\Gamma(\alpha_1+\alpha_2+\cdots+\alpha_{k-1}+\alpha_k)}{\Gamma(\alpha_1)\cdots\Gamma(\alpha_{k-1})\Gamma(\alpha_k)}}\left(\prod_{i=1}^{k-1} p_i^{\alpha_i-1}\right)\left(\left[1-\sum_{i=1}^{k-1} p_i\right]\right)^{\alpha_k-1},$$
$$\forall (p_1, p_2 \cdots p_k) \in S^k := \left\{(p_1, \cdots p_k) : p_i \geq 0, \sum_i p_i \leq 1\right\}$$

□

(1) The consistency condition (i) of $\nu_\Omega := \mathcal{L}\left((P(B_1), P(B_2), \cdots, P(B_k))\,|D_\alpha\right)$ for the ***fixed*** α in Theorem 4 holds by finite additivity and tail-free property [4] of $\Pi = D_\alpha$.
The tail-free property can be understood as the events are independent over time variable yet possibly dependent at the same time point. Formally it can be expressed as:
If $P \sim Dirichlet_k\left(\alpha(B_1), \cdots \alpha(B_k)\right)$ is a ***random probability measure*** defined by $P(\bullet|B_i) := \begin{cases} \frac{P(B)}{P(B_i)} \forall B \in \mathcal{B} & \alpha(B_i) = 0 \\ P_0 \in M(\mathcal{X}) & \alpha(B_i) \neq 0 \end{cases}$ then $(P(B_1), \cdots P(B_k)) \perp P(\bullet|B_1) \perp \cdots \perp P(\bullet|B_k)$.
   (a) $P(\bullet|B_1) \perp \cdots \perp P(\bullet|B_k)$ comes from Gamma representation of $Dirichlet_k\left(\alpha(B_1), \cdots \alpha(B_k)\right)$and Factorization Theorem i.e. $\left(\frac{Y_1}{\sum_{j\in B_1} Y_j}, \cdots \frac{Y_k}{\sum_{j\in B_k} Y_j}\right) \sim Dirichlet_k\left(\alpha(B_1), \cdots \alpha(B_k)\right)$ if $Y_j \overset{i.i.d}{\sim} Gamma(\alpha(B_j), 1)$. This representation can be directly verified by change of variable formula.
   (b) $P(B_i) := \frac{Z_i}{\sum_i Z_i}, Z_i = \int_{x\in B_i \subset \mathcal{X}} Y_x, Y_x \overset{i.i.d}{\sim} Gamma(\alpha(B_i), 1)$.
   (c) Since $\{B_j\}$ is a partition, $\sum_i Z_i \perp P(\bullet|B_*)$ is proved. Simultaneously, the marginal distributions are also Dirichlet.

(2) The continuity condition (ii) in Theorem 4 follows from α is a finite measure so that $B_n \downarrow B \Rightarrow lim_{n\to\infty}\alpha(B_n) \to \alpha(B)$ and the limiting behavior of $Dirichlet_k\left(\alpha(B_1), \cdots \alpha(B_k)\right)$ [3]:
   (a) If $\alpha^{(n)} \to_w \alpha, c^{(n)} \to c \in (0, \infty)$ then $Dirichlet_k\left(c^{(n)}\alpha^{(n)}(B_1), \cdots c^{(n)}\alpha^{(n)}(B_k)\right) \to_w Dirichlet_k\left(c\alpha(B_1), \cdots c\alpha(B_k)\right)$
   (b) If $\alpha^{(n)} \to_w \alpha, c^{(n)} \to 0$ then $Dirichlet_k\left(c^{(n)}\alpha^{(n)}(B_1), \cdots c^{(n)}\alpha^{(n)}(B_k)\right) \to_w \sum_i \alpha(B_i)\delta_{B_i}$
   (c) If $\alpha^{(n)} \to_w \alpha, c^{(n)} \to \infty$ then $Dirichlet_k\left(c^{(n)}\alpha^{(n)}(B_1), \cdots c^{(n)}\alpha^{(n)}(B_k)\right) \to_w \delta_\alpha$

References

[1] Ghosh, Jayanta K., Ramamoorthi, R.V. Bayesian Nonparametrics. Springer-Verlag New York, 2003.
[2] Ghosh, Jayanta K., Mohan Delampady, and Tapas Samanta. An introduction to Bayesian analysis: theory and methods. Springer Science & Business Media, 2007.
[3] Doob, Joseph L. Measure theory, 1994.
[4] Ferguson, Thomas S. "A Bayesian analysis of some nonparametric problems." The annals of statistics (1973): 209-230.

[3] [1] p.93