
Supplementary Document for Construction of Dependent Dirichlet Processes based on Poisson Processes

Dahua Lin
CSAIL, MIT
dhlin@mit.edu

Eric Grimson
CSAIL, MIT
welg@csail.mit.edu

John Fisher
CSAIL, MIT
fisher@csail.mit.edu

Abstract

This document comprises two sections. The first section is a brief introduction of the mathematical background. The second section gives the proof of main theorems presented in the paper.

1 Mathematical Background

In this section, we review mathematical concepts and important results that are closely related to the paper. In particular, the emphasis is placed onto Poisson processes, Gamma processes, Dirichlet processes, as well as completely random measures. We assume that the readers are familiarized with the basic knowledge in measure theory, probability theory, and stochastic processes.

1.1 Random Point Processes and Random Measures

Definition 1 (Random point process). *Let (Ω, \mathcal{F}) be a measurable space, a **random point process** (or simply **point process**) on Ω is a random variable, whose value is a countable subset of Ω .*

Given a point process Π , each sample π of Π is a set of points in Ω , which induces a σ -finite counting measure over Ω , denoted by N_π . The measure N_π is defined to be

$$N_\pi(A) = \#\{\Omega \cap \pi\}.$$

Intuitively, N_π counts the number of points in π that are also in A . Then, we can introduce a random variable N_Π that depends on Π , such that when $\Pi = \pi$, $N_\Pi = N_\pi$. It means that the value of N_Π is a counting measure over Ω . This kind of random variable is called a *random measure*, which is defined formally, as follows

Definition 2 (Random measure). *Let (Ω, \mathcal{F}) be a measurable space, a **random measure** on Ω is a random variable, whose value is a measure over Ω .*

According to the analysis above, we can see that each point process Π corresponds uniquely to a random counting measure N_Π .

Definition 3 (Completely random measure). *Let M be a random measure over (Ω, \mathcal{F}) , then M is called a **completely random measure**, if for any collection of disjoint measurable subsets $A_1, A_2, \dots \in \mathcal{F}$, the random variables $M(A_1), M(A_2), \dots$ are independent.*

Definition 4 (Completely random point process). *Let Π be a point process on Ω , then Π is called a **completely random point process** if N_Π is a completely random measure.*

1.2 Poisson Distributions and Poisson Processes

Definition 5 (Poisson distribution). A discrete random variable X whose values are non-negative integers has a **Poisson distribution**, denoted by $X \sim \mathcal{P}(\lambda)$, if

$$\mathbb{P}(X = k) = \frac{\lambda^k e^{-\lambda}}{k!}, \quad k = 0, 1, \dots$$

Here, λ is called the mean parameter of the distribution.

Let $X \sim \mathcal{P}(\lambda)$, then

$$\mathbb{E}(X) = \lambda, \quad \text{and} \quad \text{Var}(X) = \lambda. \quad (1)$$

For Poisson distribution, we have the following important results.

Proposition 1 (Countable Additivity). Let X_1, X_2, \dots be a countable collection of independent poisson distributed variables as $X_k \sim \mathcal{P}(\lambda_k)$. Then

$$\sum_{k=1}^{\infty} X_k \sim \mathcal{P}\left(\sum_{k=1}^{\infty} \lambda_k\right). \quad (2)$$

We note that as a special case, poisson distribution also satisfies *finite additivity*.

Proposition 2. Let $X \sim \mathcal{P}(\lambda)$, and $(Y_1, \dots, Y_K) \sim \text{Mult}(p_1, \dots, p_K; X)$, then Y_1, \dots, Y_K are independent, and each k has $Y_k \sim \mathcal{P}(p_k \lambda)$.

Next, we extend Poisson distributions to Poisson processes.

Definition 6 (Poisson Process). A point process Π on Ω is called a *Poisson process with mean measure μ* , denoted by $\Pi \sim \text{PoissonP}(\mu)$, if it satisfies

1. for each measurable subset $A \in \mathcal{F}_\Omega$, $N_\Pi(A)$ has a Poisson distribution as $N_\Pi(A) \sim \mathcal{P}(\mu(A))$; and
2. Π is completely random, i.e. for any collection of disjoint measurable subsets $A_1, A_2, \dots \in \mathcal{F}$, $N_\Pi(A_1), N_\Pi(A_2), \dots$ are independent.

Theorem 1. A point process Π on a regular measure space is a Poisson process if and only if N_Π is completely random. If this is true, the base measure is given by $\mu(A) = \mathbb{E}(N_\Pi(A))$.

Note: For the convenience of the reader, all theorems presented in the paper are re-stated in this supplemental document with the same indices.

1.3 Gamma Distributions and Gamma Processes

Definition 7 (Gamma Distribution). A non-negative real-valued random variable X is said to have a **Gamma distribution** with shape parameter u and scale parameter λ , denoted by $X \sim \text{Gamma}(u, \lambda)$, if its probability density function is given by

$$f(x; u, \lambda) = \frac{x^{u-1} e^{-x/\lambda}}{\lambda^u \Gamma(u)}.$$

For $X \sim \text{Gamma}(u, \lambda)$, we have

$$\mathbb{E}(X) = u\lambda, \quad \text{and} \quad \text{Var}(X) = u\lambda^2.$$

In particular when $\lambda = 1$, $\mathbb{E}(X) = \text{Var}(X) = u$.

Like Poisson distribution, Gamma distribution also satisfies countable additivity.

Proposition 3. Let X_1, X_2, \dots be a countable collection of independent Gamma distributed variables as $X_k \sim \text{Gamma}(u_k, \lambda)$. Then

$$\sum_{k=1}^{\infty} X_k \sim \text{Gamma}\left(\sum_{k=1}^{\infty} u_k, \lambda\right). \quad (3)$$

Definition 8 (Gamma Process). *A random measure G on Ω is called a **Gamma process** with base measure μ and scale parameter λ , denoted by $G \sim \text{GP}(\mu, \lambda)$, if it satisfies*

1. *for each measurable subset $A \in \mathcal{F}_\Omega$, $G(A)$ has a Gamma distribution as $G(A) \sim \text{Gamma}(\mu(A), \lambda)$; and*
2. *G is completely random.*

For the purpose of studying Dirichlet processes, the scale parameter λ does not affect the results, and we therefore assume $\lambda = 1$ in both the paper and this document. For conciseness, we write $\text{GP}(\mu)$ in place of $\text{GP}(\mu, 1)$, and $\text{Gamma}(\mu(A))$ in place of $\text{Gamma}(\mu(A), 1)$, without making the scale parameter explicit.

1.4 Dirichlet Distributions and Dirichlet Processes

Let \mathbb{S}_d denote the probability simplex in the d -dimensional real vector space \mathbb{R}^d , as

$$\mathbb{S}_d = \{(x_1, \dots, x_d) \in \mathbb{R}^d : x_i \geq 0, i = 1, \dots, d, \text{ and } x_1 + \dots + x_d = 1\}. \quad (4)$$

Definition 9 (Dirichlet Distribution). *An \mathbb{S}_d -valued random variable X is said to have a Dirichlet distribution, denoted by $X \sim \text{Dir}(\alpha_1, \dots, \alpha_d)$ with $\alpha_1, \dots, \alpha_d > 0$, if it has a probability density function with respect to the Lebesgue measure over \mathbb{S}_d given by*

$$f(x_1, \dots, x_d; \alpha_1, \dots, \alpha_d) = \frac{\Gamma\left(\sum_{i=1}^d \alpha_i\right)}{\prod_{i=1}^d \Gamma(\alpha_i)} \prod_{i=1}^d x_i^{\alpha_i-1}. \quad (5)$$

For $X = (X_1, \dots, X_d) \sim \text{Dir}(\alpha_1, \dots, \alpha_d)$, we have

$$\mathbb{E}(X_i) = \frac{\alpha_i}{\alpha_*}, \quad \text{Var}(X_i) = \frac{\mathbb{E}(X_i)(1 - \mathbb{E}(X_i))}{\alpha_* + 1}, \quad \text{and} \quad \text{Cov}(X_i, X_j) = -\frac{\mathbb{E}(X_i)\mathbb{E}(X_j)}{\alpha_* + 1}. \quad (6)$$

Here, $\alpha_* = \sum_{i=1}^d \alpha_i$.

Definition 10 (Dirichlet Process). *A random measure D on Ω is called a Dirichlet process with base measure μ , denoted by $D \sim \text{DP}(\mu)$, if for any finite measurable partition $\{A_1, \dots, A_n\}$ of Ω ,*

$$(D(A_1), \dots, D(A_n)) \sim \text{Dir}(\mu(A_1), \dots, \mu(A_n)). \quad (7)$$

2 Proofs of Theorems

There are seven theorems in the paper. Theorem 1 is a deep result established based on Lévy-Khinchin representation. Interested readers can refer to [1](see chapter 1.4 and chapter 8) for details. Theorem 2 is an immediate corollary of the superposition theorem (page 16 of [1]). We derived the remaining theorems (theorem 3, 4, 5, 6, and 7) in developing our approach. In this section, we prove all theorems except theorem 1, as well as the formulas for computing expectation and covariance (those below theorem 3 and theorem 5 in the paper).

Section 2 in the paper briefly characterizes the relations between Poisson processes, compound Poisson processes, Gamma processes, and Dirichlet processes. These relations are crucial for deriving the theorems. In particular, we will repeatedly use the following two facts in our proofs:

1. Let γ denote a measure over \mathbb{R}^+ given by $\gamma(dw) = w^{-1}e^{-w}dw$. If G is a compound Poisson process whose underlying Poisson process has a mean measure $\mu \times \gamma$, then G is a Gamma process with base measure μ , i.e. $G \sim \text{GP}(\mu)$.
2. Let $G \sim \text{GP}(\mu)$, then $D := G/G(\Omega) \sim \text{Dir}(\mu)$, $G(\Omega) \sim \text{Gamma}(\mu(\Omega))$, and D is independent of $G(\Omega)$.

2.1 The Proofs for Theorem 2 and Theorem 3

Lemma 1 (Disjointness Lemma [1]). *Let Π_1, \dots, Π_m be independent Poisson processes on Ω , and $\Pi_k \sim \text{PoissonP}(\mu_k)$. Suppose each μ_k is a σ -finite non-atomic measure, then Π_1, \dots, Π_m are disjoint almost surely.*

The proof of this lemma can be found in pages 15 to 16 in [1]. The implication of this lemma is that for any $A \in \mathcal{F}_\Omega$, we have

$$N_{\Pi_1 \cup \dots \cup \Pi_m}(A) = \Pi_1(A) + \dots + \Pi_m(A), \text{ a.s.} \quad (8)$$

In addition, σ -finiteness and being non-atomic are two properties that we assume implicitly for all the Poisson processes discussed in this document, such that the disjointness lemma can apply.

Theorem 2 (Superposition Theorem for Poisson Process). *Let Π_1, \dots, Π_m be independent Poisson processes on Ω with $\Pi_k \sim \text{PoissonP}(\mu_k)$, then*

$$\bigcup_{k=1}^m \Pi_k \sim \text{PoissonP}\left(\sum_{k=1}^m \mu_k\right). \quad (9)$$

Proof. Let $\Pi' := \Pi_1 \cup \dots \cup \Pi_m$. Then according to the disjointness lemma and the additivity of Poisson distribution, for each $A \in \mathcal{F}_\Omega$, we have

$$N_{\Pi}(A) = \sum_{i=1}^n N_{\Pi_k}(A) \sim \mathcal{P}\left(\sum_{k=1}^m \mu_k(A)\right). \quad (10)$$

Moreover, N_{Π} is completely random due to the complete randomness of N_{Π_k} . Hence, $\Pi \sim \text{PoissonP}\left(\sum_{k=1}^m \mu_k\right)$. \square

Note that [1] gives a more general version of this theorem (in page 16) that considers the sum of a countable collection of Poisson processes. Nonetheless, the finite sum is sufficient for the development of our approach.

Lemma 2. *Let G_1, \dots, G_m be independent Gamma processes on Ω with $G_k \sim \Gamma\text{P}(\mu_k)$, then*

$$\sum_{k=1}^m G_k \sim \Gamma\text{P}\left(\sum_{k=1}^m \mu_k\right). \quad (11)$$

Proof. Let Π_k^* be the Poisson processes on the product space $\Omega \times \mathbb{R}^+$ that underlies G_k , which has $\Pi_k^* \sim \text{PoissonP}(\mu_k \times \gamma)$. According to the disjointness lemma, $\Pi' := \bigcup_{k=1}^m \Pi_k^*$ is the Poisson process that underlies $G' := \sum_{k=1}^m G_k$. By theorem 2, we have

$$\Pi' \sim \text{PoissonP}\left(\sum_{k=1}^m (\mu_k \times \gamma)\right) = \text{PoissonP}\left(\left(\sum_{k=1}^m \mu_k\right) \times \gamma\right). \quad (12)$$

Based on the relation between Gamma processes and Poisson processes, we can conclude that $G' \sim \Gamma\text{P}\left(\sum_{k=1}^m \mu_k\right)$. \square

Theorem 3. *Let D_1, \dots, D_m be independent Dirichlet processes on Ω with $D_k \sim \text{DP}(\mu_k)$, and $(c_1, \dots, c_m) \sim \text{Dir}(\mu_1(\Omega), \dots, \mu_m(\Omega))$ be independent of D_1, \dots, D_k , then*

$$\sum_{k=1}^m c_k D_k \sim \text{DP}\left(\sum_{k=1}^m \mu_k\right). \quad (13)$$

Proof. For each k , we draw $g_k \sim \text{Gamma}(\mu_k(\Omega))$ independently, and let $G_k = g_k D_k$. Then from the relation between Dirichlet process and Gamma process, we know that G_1, \dots, G_m are independent Gamma processes with $G_k \sim \Gamma\text{P}(\mu_k)$. Let $G' := \sum_{k=1}^m G_k = \sum_{k=1}^m g_k D_k$, then by the lemma above, we have $G' \sim \Gamma\text{P}\left(\sum_{k=1}^m \mu_k\right)$. Let $g' = \sum_{k=1}^m g_k$, then by normalizing G' , we get

$$D' := G'/G'(\Omega) = G'/g' = \sum_{k=1}^m c_k D_k. \quad (14)$$

Here, $c_k = g_k/g' = g_k/\sum_{l=1}^m g_l$. Hence, $(c_1, \dots, c_m) \sim \text{Dir}(\mu_1(\Omega), \dots, \mu_m(\Omega))$, and c_1, \dots, c_m are independent of D_1, \dots, D_m due to the fact that g_1, \dots, g_m are independent of D_1, \dots, D_m . In addition, since D' is obtained by normalizing G' , we have $D' \sim \text{DP}\left(\sum_{k=1}^m \mu_k\right)$. Combining this with Eq.(14) completes the proof. \square

Next, we prove the formulas of the expectation of D' and the covariance between D' and D_k .

Lemma 3. *Let X, Y be independent real-valued random variables, then*

$$\text{Cov}(XY, Y) = \mathbb{E}(X)\text{Var}(Y). \quad (15)$$

Proof. We can see this by

$$\begin{aligned} \text{Cov}(XY, Y) &= \mathbb{E}(XY^2) - \mathbb{E}(XY)\mathbb{E}(Y) = \mathbb{E}(X)\mathbb{E}(Y^2) - \mathbb{E}(X)\mathbb{E}(Y)^2 \\ &= \mathbb{E}(X)(\mathbb{E}(Y^2) - \mathbb{E}(Y)^2) = \mathbb{E}(X)\text{Var}(Y). \end{aligned} \quad (16)$$

□

Proposition 4. *Let D_1, \dots, D_m be independent Dirichlet processes on Ω with $D_k \sim \text{DP}(\mu_k)$, $(c_1, \dots, c_m) \sim \text{Dir}(\mu_1(\Omega), \dots, \mu_m(\Omega))$ be independent of D_1, \dots, D_m , and $D' = \sum_{k=1}^m c_k D_k$. Then for each $A \in \mathcal{F}_\Omega$,*

$$\mathbb{E}(D'(A)) = \sum_{k=1}^m \frac{\alpha_k}{\alpha'} \mathbb{E}(D_k(A)), \quad \text{and} \quad \text{Cov}(D'(A), D_k(A)) = \frac{\alpha_k}{\alpha'} \text{Var}(D_k(A)), \quad (17)$$

where $\alpha_k = \mu_k(\Omega)$ and $\alpha' = \sum_{k=1}^m \alpha_k$.

Proof. Let $\mu' = \sum_{k=1}^m \mu_k$, then $D' \sim \text{DP}(\mu')$. It follows that

$$\mathbb{E}(D'(A)) = \frac{\mu'(A)}{\mu'(\Omega)} = \frac{1}{\alpha'} \sum_{k=1}^m \mu_k(A) = \frac{1}{\alpha'} \sum_{k=1}^m \alpha_k \frac{\mu_k(A)}{\mu_k(\Omega)} = \sum_{k=1}^m \frac{\alpha_k}{\alpha'} \mathbb{E}(D_k(A)). \quad (18)$$

For the covariance, we have

$$\begin{aligned} \text{Cov}\left(\sum_{i=1}^m c_i D_i(A), D_k(A)\right) &= \sum_{i=1}^m \text{Cov}(c_i D_i(A), D_k(A)), \\ &= \text{Cov}(c_k D_k(A), D_k(A)), \\ &= \mathbb{E}(c_k) \text{Var}(D_k(A)) = \frac{\alpha_k}{\alpha'} \text{Var}(D_k(A)). \end{aligned}$$

The proof is completed. □

2.2 The Proofs for Theorem 4 and Theorem 5

Consider a marking process described as follows. Let Π be a Poisson process on Ω with $\Pi \sim \text{PoissonP}(\mu)$; and for each point $\theta \in \Pi$, we randomly draw a mark $m_\theta \sim p(\theta, \cdot)$ in M . Here, $p(\theta, \cdot)$ is a probability measure over the mark space M , which may or may not depend on θ . Collecting all the pairs (θ, m_θ) leads to a point process on the product space $\Omega \times M$. The following theorem shows that such a point process is also a Poisson process.

Lemma 4 (Marking Theorem [1]). *Suppose (M, \mathcal{F}_M) is a measurable space, and for each $\theta \in \Omega$, there exists a probability measure $p(\theta, \cdot)$ over M , such that for each $B \in \mathcal{F}_M$, $p(\theta, B)$, as a function of θ , is measurable with respect to \mathcal{F}_Ω . Let $\Pi \sim \text{PoissonP}(\mu)$ be a Poisson process on Ω , and for each $\theta \in \Pi$, we independently draw $m_\theta \sim p(\theta, \cdot)$, then the point process given by $\Pi^* = \{(\theta, m_\theta) : \theta \in \Pi\}$ is a Poisson process on $\Omega \times M$ as $\Pi^* \sim \text{PoissonP}(\mu^*)$, where*

$$\mu^*(C) = \int_C \mu(\theta) p(\theta, dm). \quad (19)$$

Based on the marking theorem, we derive the coloring theorem as below.

Lemma 5 (Generalized Coloring Theorem). *Let $\Pi \sim \text{PoissonP}(\mu)$ be a Poisson process on Ω , and $q_k : \Omega \rightarrow [0, 1]$ be a non-negative measurable function for $k = 1, \dots, m$, such that $\sum_{k=1}^m q_k(\theta) = 1$ for all $\theta \in \Omega$. For each $\theta \in \Pi$, we independently draw $c_\theta \in \{1, \dots, m\}$ with $\mathbb{P}(c_\theta = k) = q_k(\theta)$. Let $\Pi_k = \{\theta \in \Pi : c_\theta = k\}$ for each k , then Π_k is a Poisson process on Ω , as*

$$\Pi_k \sim \text{PoissonP}(q_k \mu). \quad (20)$$

Here, $q_k\mu$ is defined by

$$(q_k\mu)(A) := \int_A q_k d\mu, \quad \text{for } A \in \mathcal{F}_\Omega. \quad (21)$$

This can also be written as $(q_k\mu)(d\theta) = q_k(\theta)\mu(d\theta)$. Moreover, Π_1, \dots, Π_n are independent.

Proof. Let $\Pi^* = \{(\theta, c_\theta) : \theta \in \Pi\}$ be the marked point process on the product space $\Omega \times \{1, \dots, n\}$. By the marking theorem, Π^* is a Poisson process as $\Pi^* \sim \text{PoissonP}(\mu^*)$. In particular, for any $A \in \mathcal{F}_\Omega$ and $k \in \{1, \dots, n\}$, we have

$$\mu^*(A \times \{k\}) = \int_A q_k(\theta)\mu(d\theta) = (q_k\mu)(A). \quad (22)$$

According to the construction of Π_k , we note that for any $A \in \mathcal{F}_\Omega$,

$$N_{\Pi_k}(A) = N_{\Pi^*}(A \times \{k\}) \sim \mathcal{P}((q_k\mu)(A)). \quad (23)$$

Moreover, N_{Π_k} inherits the completely randomness from Π^* . Hence, $\Pi_k \sim \text{PoissonP}(q_k\mu)$. Furthermore, we note that Π_1, \dots, Π_n are respectively based on $\Omega \times \{1\}, \dots, \Omega \times \{n\}$ in the product space, which are disjoint, implying that Π_1, \dots, Π_n are independent. The proof is completed. \square

Note that the coloring theorem that we show here is more general than the coloring theorem in page 53 of [1], which assumes that the ‘‘color distribution’’ for all θ is the same. While in this generalized coloring theorem, we allow different color distribution for different θ .

Theorem 4 (Subsampling Theorem). *Let $\Pi \sim \text{PoissonP}(\mu)$ be a Poisson process on the space Ω , and $q : \Omega \rightarrow [0, 1]$ be a non-negative measurable function. If we independently draw $z_\theta \in \{0, 1\}$ for each $\theta \in \Pi_0$ with $\mathbb{P}(z_\theta = 1) = q(\theta)$, and let $\Pi_k = \{\theta \in \Pi : z_\theta = k\}$ for $k = 0, 1$, then Π_0 and Π_1 are independent Poisson processes on Ω , with $\Pi_0 \sim \text{PoissonP}((1 - q)\mu)$ and $\Pi_1 \sim \text{PoissonP}(q\mu)$.*

Proof. This is just a special case of the generalized coloring theorem, where there are only two colors 0 and 1. \square

Lemma 6. *Let $G \sim \text{GP}(\mu)$ be a Gamma process on Ω given by $G = \sum_{i=1}^n w_i \delta_{\theta_i}$, and $q : \Omega \rightarrow [0, 1]$ be a non-negative measurable function. If for each i , we independently draw $z_i \in \{0, 1\}$ with $\mathbb{P}(z_i = 1) = q(\theta_i)$, and let*

$$G_0 = \sum_{i:z_i=0} w_i \delta_{\theta_i},$$

$$G_1 = \sum_{i:z_i=1} w_i \delta_{\theta_i}.$$

Then $G_0 \sim \text{GP}((1 - q)\mu)$, $G_1 \sim \text{GP}(q\mu)$, and G_0 is independent of G_1 .

Proof. Let $\Pi^* \sim \text{PoissonP}(\mu \times \gamma)$ be the Poisson process that underlies G . We randomly colors Π^* with the probability function q to get Π_0^* and Π_1^* respectively for color 0 and color 1. By the subsampling theorem for Poisson process, we have $\Pi_0^* \sim \text{PoissonP}((1 - q)\mu \times \gamma)$, $\Pi_1^* \sim \text{PoissonP}(q\mu \times \gamma)$, and Π_0^* is independent of Π_1^* . By the construction of G_0 and G_1 , we can easily see that Π_0^* and Π_1^* are the Poisson processes that underlie G_0 and G_1 respectively. Hence, we can conclude that $G_0 \sim \text{GP}((1 - q)\mu)$, $G_1 \sim \text{GP}(q\mu)$, and G_0 is independent of G_1 . \square

Theorem 5. *Let $D \sim \text{DP}(\mu)$ be represented by $D = \sum_{i=1}^n r_i \delta_{\theta_i}$ and $q : \Omega \rightarrow [0, 1]$ be a non-negative measurable function. For each i we independently draw z_i with $\mathbb{P}(z_i = 1) = q(\theta_i)$, then*

$$\sum_{i:z_i=1} r'_i \delta_{\theta_i} \sim \text{DP}(q\mu), \quad (24)$$

where $r'_i := r_i / \sum_{j:z_j=1} r_j$ is the re-normalized coefficients.

Proof. We independently draw $g \in \text{Gamma}(\mu(\Omega))$, and let $G := gD$, then

$$G = gD = \sum_{i=1}^n gr_i \delta_{\theta_i} \sim \text{GP}(\mu). \quad (25)$$

Let $G_k = \sum_{i:z_i=k} gr_i \delta_{\theta_i}$ for $k = 0, 1$. Then by lemma 6, we have $G_1 \sim \text{GP}(q\mu)$. Normalizing G_1 , we get

$$D_1 := G_1/G_1(\Omega) = \frac{\sum_{i:z_i=1} gr_i \delta_{\theta_i}}{\sum_{j:z_j=1} gr_j} = \sum_{i:z_i=1} \frac{r_i}{\sum_{j:z_j=1} r_j} \delta_{\theta_i} = \sum_{i:z_i=1} r'_i \delta_{\theta_i}. \quad (26)$$

By the relation between Gamma process and Dirichlet process, we have $D_1 \sim \text{DP}(q\mu)$, completing the proof. \square

Proposition 5. *With the setting in the proof of theorem 5, we have for each $A \sim \mathcal{F}_\Omega$,*

$$\mathbb{E}(D_1(A)) = \frac{(q\mu)(A)}{(q\mu)(\Omega)}, \quad (27)$$

and

$$\text{Cov}(D_1(A), D(A)) = \frac{(q\mu)(\Omega)}{\mu(\Omega)} \text{Var}(D_1(A)). \quad (28)$$

Proof. The formula for $\mathbb{E}(D_1(A))$ is an immediate consequence of the result $D_1(A) \sim \text{DP}(q\mu)$. In addition, we let $D_0 = G_0/G_0(\Omega)$. Note that D_0 and D_1 are independent due to the independence between G_0 and G_1 , and D is the superposition of D_0 and D_1 . Hence, the covariance formula immediately follows from proposition 4. \square

2.3 The Proofs for Theorem 6 and Theorem 7

A *probabilistic transition* is defined to be a function $T : \Omega \times \mathcal{F}_\Omega \rightarrow [0, 1]$ such that for each $\theta \in \mathcal{F}_\Omega$, $T(\theta, \cdot)$ is a probability measure over Ω , and for each $A \in \mathcal{F}_\Omega$, $T(\cdot, A)$ is integrable. T can be considered as a transformation of measures over Ω , as

$$(T\mu)(A) := \int_{\Omega} T(\theta, A) \mu(d\theta). \quad (29)$$

Theorem 6 (Transition Theorem). *Let $\Pi \sim \text{PoissonP}(\mu)$ and T be a probabilistic transition, then*

$$T(\Pi) := \{T(\theta) : \theta \in \Pi\} \sim \text{PoissonP}(T\mu). \quad (30)$$

Here, with abuse of notation, we use $T(\theta)$ to denote an independent sample from $T(\theta, \cdot)$.

Proof. Given a measurable partition $\{A_1, \dots, A_n\}$ of Ω . For each $k = 1, \dots, n$, we define $q_k : \Omega \rightarrow [0, 1]$ by $q_k(\theta) = T(\theta, A_k)$. By the definition of T , each q_k is integrable. For each $\theta \in \Pi$, we let $z_\theta = k$ when $T(\theta) \in A_k$, and $\Pi_k = \{\theta \in \Pi : z_\theta = k\}$. With such construction, we note that for each k ,

$$N_{T(\Pi)}(A_k) = N_{\Pi_k}(\Omega). \quad (31)$$

Moreover, this is equivalent to drawing z_θ independently for each θ with $\mathbb{P}(z_\theta = k) = q_k(\theta)$. By the generalized coloring theorem, we have

$$N_{\Pi_k}(\Omega) \sim \mathcal{P}((q_k\mu)(\Omega)) = \mathcal{P}\left(\int_{\Omega} q_k(\theta) \mu(d\theta)\right) = \mathcal{P}\left(\int_{\Omega} T(\theta, A_k) \mu(d\theta)\right) = \mathcal{P}((T\mu)(A_k)). \quad (32)$$

Therefore, $N_{T(\Pi)}(A_k) \sim \mathcal{P}((T\mu)(A_k))$. Again, by the generalized coloring theorem, Π_1, \dots, Π_n are independent, which implies that $N_{T(\Pi)}(A_1), \dots, N_{T(\Pi)}(A_n)$ are independent. Since the partition $\{A_1, \dots, A_n\}$ was arbitrarily given, the theorem is proved. \square

Lemma 7. *Let $G = \sum_{i=1}^{\infty} w_i \delta_{\theta_i} \sim \text{GP}(\mu)$ be a Gamma process on Ω , then*

$$T(G) := \sum_{i=1}^{\infty} w_i \delta_{T(\theta_i)} \sim \text{GP}(T\mu). \quad (33)$$

Proof. Let Π^* be the Poisson process on $\Omega \times \mathbb{R}^+$ that underlies G , which has $\Pi^* \sim \text{PoissonP}(\mu \times \gamma)$. Given a probabilistic transition T on Ω , let T^* be a probabilistic transition on $\Omega \times \mathbb{R}^+$ defined by

$$T^*((\theta, w), A \times B) = T(\theta, A)1_{\{w \in B\}}. \quad \text{for } A \in \mathcal{F}_\Omega, B \in \mathcal{B}_{\mathbb{R}^+}. \quad (34)$$

Here, $\mathcal{B}_{\mathbb{R}^+}$ is the Borel σ -algebra for \mathbb{R}^+ . Intuitively, we can think of T^* as a random transform that sends (θ, w_θ) to $(T(\theta), w_\theta)$. Note that a measure defined on a product space is uniquely determined by the measure values for all sets in form of $A \times B$. Hence, by the equation above, T^* is completely defined.

According to the construction of $T(G)$, we can see that its underlying Poisson process is $T^*(\Pi^*)$. By the transition theorem, $T^*(\Pi^*)$ is Poisson process on $\Omega \times \mathbb{R}^+$ whose base measure is given by

$$(T^*(\mu \times \gamma))(A \times B) = \int_{\Omega \times \mathbb{R}^+} T^*((\theta, w), A \times B)(\mu \times \gamma)(d\theta dw) \quad (35)$$

$$= \int_{\Omega \times \mathbb{R}^+} T(\theta, A)1_{\{w \in B\}}\mu(d\theta)\gamma(dw) \quad (36)$$

$$= \int_{\Omega} T(\theta, A)\mu(d\theta) \int_{\mathbb{R}^+} 1_{\{w \in B\}}\gamma(dw) \quad (37)$$

$$= (T\mu)(A)\gamma(B), \quad \text{for } A \in \mathcal{F}_\Omega, B \in \mathcal{B}(\mathbb{R}^+). \quad (38)$$

Since a measure over a product space is uniquely determined by the values of all the sets in form of $A \times B$, we can conclude that $T^*(\mu \times \gamma) = (T\mu) \times \gamma$. Therefore, $T(G)$ is a Gamma process on Ω with $T(G) \sim \text{GP}(T\mu)$. \square

Theorem 7. Let $D = \sum_{i=1}^{\infty} r_i \delta_{\theta_i} \sim \text{DP}(\mu)$ be a Dirichlet process on Ω , then

$$T(D) := \sum_{i=1}^{\infty} r_i \delta_{T(\theta_i)} \sim \text{DP}(T\mu). \quad (39)$$

Proof. Let $g \sim \text{Gamma}(\mu(\Omega))$, then $G := gD \sim \text{GP}(\mu)$. By lemma 7, we have $T(G) \sim \text{GP}(T\mu)$. It follows that

$$T(D) = T(G)/g = T(G)/(T(G)(\Omega)) \sim \text{DP}(T\mu). \quad (40)$$

Here, we use the fact $g = G(\Omega) = T(G)(\Omega)$. \square

References

- [1] J. F. C. Kingman. *Poisson Processes*. Oxford University Press, 1993.