Chapter 2
The Poisson–Dirichlet Distribution

The focus of this chapter is the Poisson–Dirichlet distribution, the central topic of this book. We introduce this distribution and discuss various models that give rise to it. Following Kingman [125], the distribution is constructed through the gamma process. An alternative construction in [8] is also included, where a scale-invariant Poisson process is used. The density functions of the marginal distributions are derived through Perman’s formula. Closely related topics such as the GEM distribution, the Ewens sampling formula, and the Dirichlet process are investigated in detail through the study of urn models. The required terminology and properties of Poisson processes and Poisson random measures can be found in Appendix A.

2.1 Definition and Poisson Process Representation

For each $K \geq 2$, set

$$
\nabla_K = \left\{ (p_1, \ldots, p_K) : p_1 \geq p_2 \geq \cdots \geq p_K \geq 0, \sum_{j=1}^{K} p_j = 1 \right\},
$$

$$
\nabla_\infty = \left\{ (p_1, p_2, \ldots) : p_1 \geq p_2 \geq \cdots \geq 0, \sum_{j=1}^{\infty} p_j = 1 \right\},
$$

$$
\nabla = \left\{ (p_1, p_2, \ldots) : p_1 \geq p_2 \geq \cdots \geq 0, \sum_{j=1}^{\infty} p_j \leq 1 \right\}. \tag{2.1}
$$

The space $\nabla_K$ can be embedded naturally into $\nabla_\infty$ and thus viewed as a subset of $\nabla_\infty$. The space $\nabla$ is the closure of $\nabla_\infty$ in $[0,1]^\infty$, and the topology on each of $\nabla_\infty$ and $\nabla$ is the subspace topology inherited from $[0,1]^\infty$. For $\theta > 0$, let $(X_1, \ldots, X_K)$ have a Dirichlet($\frac{\theta}{K-1}, \ldots, \frac{\theta}{K-1}$) distribution. Let $Y^K = (Y^K_1, \ldots, Y^K_K)$ be the decreasing order statistics of $(X_1, \ldots, X_K)$. 

Theorem 2.1. Let $M_1(\nabla)$ denote the space of probability measures on $\nabla$. Then the sequence $\{\mu_K : K \geq 2\}$ of laws of $Y^K$ converges weakly in $M_1(\nabla)$.

Proof. Let $\gamma(t)$ be a gamma process; i.e., a process with stationary independent increments such that each increment, $\gamma(t) - \gamma(s)$ for $0 \leq s < t$, follows a Gamma$(\theta(t - s), 1)$ distribution. We sometimes write $\gamma_t$ instead of $\gamma(\theta)$ for notational convenience. Set

$$I_l = \left( \frac{l}{K - 1}, \frac{l + 1}{K - 1} \right), \quad l = 0, \ldots, K - 1,$$

$$\tilde{X}_l = \frac{\gamma(l+1) - \gamma(l)}{\gamma(l)},$$

and $Y^K = (\tilde{Y}^K_1, \ldots, \tilde{Y}^K_K)$ denote the descending order statistics of $(\tilde{X}_1, \ldots, \tilde{X}_K)$. It follows from Theorem 1.1 that $(\tilde{X}_1, \ldots, \tilde{X}_K)$ has the same distribution as $(X_1, \ldots, X_K)$. For $l = 0, \ldots, K - 1$, denote the $l$th highest jump of $\gamma(t)$ over the interval $(0, 1)$ by $J_l$. Since the process $\gamma(t)$ does not have any fixed jump point, for almost all sample paths the jumps occur in $I_1, \ldots, I_K - 1$. Thus with probability one, we have

$$\liminf_{K \to \infty} \tilde{Y}^K_l \gamma \left( \frac{K}{K - 1} \right) \geq J_l$$

which implies that

$$\liminf_{K \to \infty} \tilde{Y}^K_l \gamma \left( \frac{K}{K - 1} \right) \geq J_l \frac{1}{\gamma(\theta)}.$$  \hspace{1cm} (2.2)

Fatou’s lemma guarantees that the summation over $l$ of the left-hand side of equation (2.2) is no more than one. This, combined with the fact that the right-hand side adds up to one, implies that the inequality in (2.2) is actually an equality for all $l$ and the lower limit is actually the limit. Let $P_l(\theta) = \frac{J_l}{\gamma(\theta)}$ for any $l \geq 0$, $P(\theta) = (P_1(\theta), P_2(\theta), \ldots)$, and let $\Pi_\theta$ denote the law of $P(\theta)$. Then $\Pi_\theta$ belongs to $M_1(\nabla)$, and we have shown that for any fixed $r \geq 1$ and any bounded continuous function $g$ on $\nabla_r$,

$$\lim_{K \to \infty} \int_{\nabla} g(p_1, \ldots, p_r) d\mu_K = \int_{\nabla} g(p_1, \ldots, p_r) d\Pi_\theta. \hspace{1cm} (2.3)$$

It now follows from the Stone–Weierstrass theorem that (2.3) holds for every real-valued continuous function $g$ on $\nabla$.

Definition 2.1. The law $\Pi_\theta$ of $P(\theta)$ in Theorem 2.1 is called the Poisson–Dirichlet distribution with parameter $\theta$.

Let $\xi_1 \geq \xi_2 \geq \ldots$ be the random points of a Poisson process with intensity measure

$$\mu(dx) = \theta e^{-\theta x} dx, x > 0.$$  

Set