Appendix A
Poisson Process and Poisson Random Measure

Reference [130] is the main source for the material on Poisson process and Poisson random measure.

A.1 Definitions

Definition A.1. Let $(\Omega, \mathscr{F}, \mathbb{P})$ be a probability space, and S be a locally compact, separable metric space with Borel σ-algebra $\mathscr{B}$. The set $\mathscr{S}$ denotes the collection of all countable subsets of S. A *Poisson process* with state space S, defined on $(\Omega, \mathscr{F}, \mathbb{P})$, is a map $\digamma$ from $(\Omega, \mathscr{F}, \mathbb{P})$ to $\mathscr{S}$ satisfying:

(a) for each B in $\mathscr{B}$,

$$N(B) = \#\{\digamma \cap B\}$$

is a Poisson random variable with parameter

$$\mu(B) = \mathbb{E}[N(B)];$$

(b) for disjoint sets $B_1, \ldots, B_n$ in $\mathscr{B}$, $N(B_1), \ldots, N(B_n)$ are independent.

If $B_1, B_2, \ldots$ are disjoint, then, by definition, we have

$$N(\cup_{i=1}^{\infty} B_i) = \sum_{i=1}^{\infty} N(B_i),$$

and

$$\mu(\cup_{i=1}^{\infty} B_i) = \sum_{i=1}^{\infty} \mu(B_i).$$

Hence, N is a random measure, and μ is a measure, both on $(S, \mathscr{B})$. The random measure N can also be written in the following form:

S. Feng, *The Poisson–Dirichlet Distribution and Related Topics*,
Probability and its Applications, DOI 10.1007/978-3-642-11194-5,

$$N = \sum_{\varsigma \in \digamma} \delta_\varsigma.$$

Definition A.2. The measure μ is called the *mean measure* of the Poisson process $\digamma$, and N is called a *Poisson random measure* associated with the Poisson process $\digamma$. The measure μ is also called the mean measure of N.

Remark: Let $S = [0, \infty)$ and μ be the Lebesgue measure on S. Then the Poisson random measure associated with the Poisson process $\digamma$ with state space S and mean measure μ is just the one-dimensional time-homogeneous Poisson process that is defined as a pure-birth Markov chain with birth rate one. The random set $\digamma$ is composed of all jump times of the process.

Definition A.3. Assume that $S = \mathbb{R}^d$ for some $d \geq 1$. Then the mean measure is also called the *intensity measure*. If there exists a positive constant c such that for any measurable set B,

$$\mu(B) = c|B|, \ |B| = \text{Lebesgue measure of } \ B,$$

then the Poisson process $\digamma$ is said to be *homogeneous* with *intensity* c.

A.2 Properties

Theorem A.1. *Let μ be the mean measure of a Poisson process $\digamma$ with state space S. Then μ is diffuse; i.e., for every x in S,*

$$\mu(\{x\}) = 0.$$

Proof. For any fixed x in S, set $a = \mu(\{x\})$. Then, by definition,

$$\mathbb{P}\{N(\{x\}) = 2\} = \frac{a^2}{2} e^{-a} = 0,$$

which leads to the result.

□

The next theorem describes the close relation between a Poisson process and the multinomial distribution.

Theorem A.2. *Let $\digamma$ be a Poisson process with state space S and mean measure μ. Assume that the total mass $\mu(S)$ is finite. Then, for any $n \geq 1, 1 \leq m \leq n$, and any set partition $B_1, \ldots, B_m$ of S, the conditional distribution of the random vector $(N(B_1), \ldots, N(B_m))$ given $N(S) = n$ is a multinomial distribution with parameters n and*

$$\left(\frac{\mu(B_1)}{\mu(S)}, \ldots, \frac{\mu(B_m)}{\mu(S)} \right).$$

Proof. For any partitions $n_1,\ldots,n_m$ of n,

$$
\begin{aligned}
&\mathbb{P}\{N(B_1)=n_1,\ldots,N(B_m)=n_m \mid N(S)=n\}\\
&\quad=\frac{\mathbb{P}\{N(B_1)=n_1,\ldots,N(B_m)=n_m\}}{\mathbb{P}\{N(S)=n\}}\\
&\quad=\frac{\prod_{i=1}^{m}\frac{\mu(B_i)_i^{n}e^{-\mu(B_i)}}{n_i!}}{\frac{\mu(S)^n e^{-\mu(S)}}{n!}}\\
&\quad=\binom{n}{n_1,\ldots,n_m}\prod_{i=1}^{m}\left(\frac{\mu(B_i)}{\mu(S)}\right)^{n_i}.
\end{aligned}
$$

□

Theorem A.3. (Restriction and union)

(1) *Let $\digamma$ be a Poisson process with state space S. Then, for every B in $\mathscr{B}$, $\digamma\cap B$ is a Poisson process with state space S and mean measure*

$$\mu_B(\cdot)=\mu(\cdot\cap B).$$

Equivalently, $\digamma\cap B$ can also be viewed as a Poisson process with state space B with mean measure given by the restriction of μ on B.

(2) *Let $\digamma_1$ and $\digamma_2$ be two independent Poisson processes with state space S, and respective mean measures μ_1 and μ_2. Then $\digamma_1\cup\digamma_2$ is a Poisson process with state space S and mean measure*

$$\mu=\mu_1+\mu_2.$$

Proof. Direct verification of the definition of Poisson process.

□

The remaining theorems in this section are stated without proof. The details can be found in [130].

Theorem A.4. (Mapping) *Let $\digamma$ be a Poisson process with state space S and σ-finite mean measure $\mu(\cdot)$. Consider a measurable map h from S to another locally compact, separable metric space S'. If the measure*

$$\mu'(\cdot)=\mu(f^{-1}(\cdot))$$

is diffuse, then $f(\digamma)=\{f(\varsigma):\varsigma\in\digamma\}$ is a Poisson process with state space S' and mean measure μ'.

Theorem A.5. (Marking) *Let $\digamma$ be a Poisson process with state space S and mean measure μ. The mark of each point ς in $\digamma$, denoted by m_ς, is a random variable, taking values in a locally compact, separable metric space S', with distribution $q(z,\cdot)$. Assume that:*

(1) *for every measurable set B' in S', $q(\cdot,B')$ is a measurable function on $(S,\mathscr{B})$;*
(2) *given $\digamma$, the random variables $\{m_\varsigma:\varsigma\in\digamma\}$ are independent;*

then $\tilde{F} = \{(\varsigma, m_\varsigma) : \varsigma \in F\}$ *is a Poisson process on with state space* $S \times S'$ *and mean measure*

$$\tilde{\mu}(dx, dm) = \mu(dx)q(x, dm).$$

The Poisson process $\tilde{F}$ *is aptly called a marked Poisson process.*

Theorem A.6. (Campbell) *Let* F *be a Poisson process on space* $(S, \mathscr{B})$ *with mean measure* μ. *Then for any non-negative measurable function* f,

$$\mathbb{E}\left[\exp\left\{-\sum_{\varsigma \in F} f(\varsigma)\right\}\right] = \exp\left\{\int_S (e^{-f(x)} - 1)\mu(dx)\right\}.$$

If f *is a real-valued measurable function on* $(S, \mathscr{B})$ *satisfying*

$$\int_S \min(|f(\mathbf{x})|, 1)\mu(d\mathbf{x}) < \infty,$$

then for any complex number λ *such that the integral*

$$\int_S (e^{\lambda f(x)} - 1)\mu(dx)$$

converges, we have

$$\mathbb{E}\left[\exp\left\{\lambda \sum_{\varsigma \in F} f(\varsigma)\right\}\right] = \exp\left\{\int_S (e^{\lambda f(x)} - 1)\mu(dx)\right\}.$$

Moreover, if

$$\int_S |f(x)|\mu(dx) < \infty, \tag{A.1}$$

then

$$\begin{aligned} \mathbb{E}\left[\sum_{\varsigma \in F} f(\varsigma)\right] &= \int_S f(x)\mu(dx), \\ Var\left[\sum_{\varsigma \in F} f(\varsigma)\right] &= \int_S f^2(x)\mu(dx). \end{aligned}$$

In general, for any $n \geq 1$, *and any real-valued measurable functions* $f_1, \ldots, f_n$ *satisfying* (*A.1*), *we have*

$$\mathbb{E}\left[\sum_{distinct\ \varsigma_1, \ldots, \varsigma_n \in F} f_1(\varsigma_1) \cdots f_n(\varsigma_n)\right] = \prod_{i=1}^{n} \mathbb{E}\left[\sum_{\varsigma_i \in F} f_i(\varsigma_i)\right]. \tag{A.2}$$

Appendix B
Basics of Large Deviations

In probability theory, the law of large numbers describes the limiting average or mean behavior of a random population. The fluctuations around the average are characterized by a fluctuation theorem such as the central limit theorem. The theory of large deviations is concerned with the rare event of deviations from the average. Here we give a brief account of the basic definitions and results of large deviations. Everything will be stated in a form that will be sufficient for our needs. All proofs will be omitted. Classical references on large deviations include [30], [50], [168], and [175]. More recent developments can be found in [46], [28], and [69]. The formulations here follow mainly Dembo and Zeitouni [28]. Theorem B.6 is from [157].

Let E be a complete, separable metric space with metric ρ. Generic elements of E are denoted by x, y, etc.

Definition B.1. A function I on E is called a *rate function* if it takes values in $[0,+\infty]$ and is lower semicontinuous. For each c in $[0,+\infty)$, the set

$$\{x \in E : I(x) \leq c\}$$

is called a level set. The *effective domain* of I is defined as

$$\{x \in E : I(x) < \infty\}.$$

If all level sets are compact, the rate function is said to be *good*.

Rate functions will be denoted by other symbols as the need arises. Let $\{X_\varepsilon : \varepsilon > 0\}$ be a family of E-valued random variables with distributions $\{\mathbb{P}_\varepsilon : \varepsilon > 0\}$, defined on the Borel σ-algebra $\mathscr{B}$ of E.

Definition B.2. The family $\{X_\varepsilon : \varepsilon > 0\}$ or the family $\{\mathbb{P}_\varepsilon : \varepsilon > 0\}$ is said to satisfy a large deviation principle (LDP) on E as ε converges to zero, with speed ε and a good rate function I if

S. Feng, *The Poisson–Dirichlet Distribution and Related Topics*,
Probability and its Applications, DOI 10.1007/978-3-642-11194-5,

$$\text{for any closed set } F,\ \limsup_{\varepsilon\to 0}\ \varepsilon \log \mathbb{P}_\varepsilon\{F\} \le -\inf_{x\in F} I(x), \tag{B.1}$$

$$\text{for any open set } G,\ \liminf_{\varepsilon\to 0}\ \varepsilon \log \mathbb{P}_\varepsilon\{G\} \ge -\inf_{x\in G} I(x). \tag{B.2}$$

Estimates (B.1) and (B.2) are called the upper bound and lower bound, respectively. Let $a(\varepsilon)$ be a function of ε satisfying

$$a(\varepsilon) > 0,\ \lim_{\varepsilon\to 0} a(\varepsilon) = 0.$$

If the multiplication factor ε in front of the logarithm is replaced by $a(\varepsilon)$, then the LDP has speed $a(\varepsilon)$.

It is clear that the upper and lower bounds are equivalent to the following statement: for all $B \in \mathscr{B}$,

$$-\inf_{x\in B^\circ} I(x) \le \liminf_{\varepsilon\to 0} \varepsilon \log \mathbb{P}_\varepsilon\{B\}$$
$$\limsup_{\varepsilon\to 0} \varepsilon \log \mathbb{P}_\varepsilon\{B\} \le -\inf_{x\in \overline{B}} I(x),$$

where B° and $\overline{B}$ denote the interior and closure of B respectively. An event $B \in \mathscr{B}$ satisfying

$$\inf_{x\in B^\circ} I(x) = \inf_{x\in \overline{B}} I(x)$$

is called a I-continuity set. Thus for a I-continuity set B, we have that

$$\lim_{\varepsilon\to 0} \varepsilon \log \mathbb{P}_\varepsilon\{B\} = -\inf_{x\in B} I(x).$$

If the values for ε are only $\{1/n : n \ge 1\}$, we will write P_n instead of $P_{1/n}$.

If the upper bound (B.1) holds only for compact sets, then we say the family $\{\mathbb{P}_\varepsilon : \varepsilon > 0\}$ satisfies the *weak LDP*. To establish an LDP from the weak LDP, one needs to check the following condition which is known as *exponential tightness*: For any $M > 0$, there is a compact set K such that on the complement K^c of K we have

$$\limsup_{\varepsilon\to 0}\ \varepsilon \log \mathbb{P}_\varepsilon\{K^c\} \le -M. \tag{B.3}$$

Definition B.3. The family $\{\mathbb{P}_\varepsilon : \varepsilon > 0\}$ is said to be *exponentially tight* if (B.3) holds.

An interesting consequence of an LDP is the following theorem.

Theorem B.1. (Varadhan's lemma) *Assume that the family $\{\mathbb{P}_\varepsilon : \varepsilon > 0\}$ satisfies an LDP with speed ε and good rate function I. Let f and the family $\{f_\varepsilon : \varepsilon \ge 1\}$ be bounded continuous functions on E satisfying*

$$\lim_{\varepsilon\to 0} \sup_{x\in E} \rho(f_\varepsilon(x), f(x)) = 0.$$

Then

$$\lim_{\varepsilon\to 0}\varepsilon\log E^{\mathbb{P}_\varepsilon}[e^{\frac{f_\varepsilon(x)}{\varepsilon}}]=\sup_{x\in E}\{f(x)-I(x)\}.$$

Remark: Without knowing the existence of an LDP, one can guess the form of the rate function by calculating the left-hand side of the above equation.

The next result shows that an LDP can be transformed by a continuous function from one space to another.

Theorem B.2. (Contraction principle) *Let E,F be complete, separable spaces, and h be a measurable function from E to F. If the family of probability measures $\{\mathbb{P}_\varepsilon:\varepsilon>0\}$ on E satisfies an LDP with speed ε and good rate function I and the function h is continuous at every point in the effective domain of I, then the family of probability measures $\{P_\varepsilon\circ h^{-1}:\varepsilon>0\}$ on F satisfies an LDP with speed ε and good rate function, I', where*

$$I'(y)=\inf\{I(x):x\in E,y=h(x)\}.$$

Theorem B.3. *Let $\{Y_\varepsilon:\varepsilon>0\}$ be a family of random variables satisfying an LDP on space E with speed ε and rate function I. If E_0 is a closed subset of E, and*

$$\mathbb{P}\{Y_\varepsilon\in E_0\}=1,\ \{x\in E:I(x)<\infty\}\subset E_0$$

then the LDP for $\{Y_\varepsilon:\varepsilon>0\}$ holds on E_0.

The next concept describes the situation when two families of random variables are indistinguishable exponentially.

Definition B.4. Let

$$\{X_\varepsilon:\varepsilon>0\},\ \{Y_\varepsilon:\varepsilon>0\}$$

be two families of E-valued random variables on a probability space $(\Omega,\mathscr{F},\mathbb{P})$. If for any $\delta>0$ the family $\{\mathscr{P}_\varepsilon:\varepsilon>0\}$ of joint distributions of $(X_\varepsilon,Y_\varepsilon)$ satisfies

$$\limsup_{\varepsilon\to 0}\varepsilon\log\mathscr{P}_\varepsilon\{\{(x,y):\rho(x,y)>\delta\}\}=-\infty,$$

then we say that $\{X_\varepsilon:\varepsilon>0\}$ and $\{Y_\varepsilon:\varepsilon>0\}$ are *exponentially equivalent* with speed ε.

The following theorem shows that the LDPs for exponentially equivalent families of random variables are the same.

Theorem B.4. *Let $\{X_\varepsilon:\varepsilon>0\}$ and $\{Y_\varepsilon:\varepsilon>0\}$ be two exponentially equivalent families of $E-$valued random variables. If an LDP holds for $\{X_\varepsilon:\varepsilon>0\}$, then the same LDP holds for $\{Y_\varepsilon:\varepsilon>0\}$ and vice versa.*

To generalize the notion of exponential equivalence, we introduce the concept of exponential approximation next.

Definition B.5. Consider a family of random variables $\{X_\varepsilon : \varepsilon > 0\}$ and a sequence of families of random variables $\{Y_\varepsilon^n : \varepsilon > 0\}, n = 1,2,\ldots,$ all defined on the same probability space. Denote the joint distribution of $(X_\varepsilon, Y_\varepsilon^n)$ by $\mathscr{P}_\varepsilon^n$. Assume that for any $\delta > 0$,

$$\lim_{n\to\infty} \limsup_{\varepsilon\to 0} \varepsilon \log \mathscr{P}_\varepsilon^n\{\{(x,y) : \rho(x,y) > \delta\}\} = -\infty, \tag{B.4}$$

then the sequence $\{Y_\varepsilon^n : \varepsilon > 0\}$ is called an *exponentially good approximation* of $\{X_\varepsilon : \varepsilon > 0\}$.

Theorem B.5. *Let the sequence of families $\{Y_\varepsilon^n : \varepsilon > 0\}, n = 1,2,\ldots,$ be an exponentially good approximation to the family $\{X_\varepsilon : \varepsilon > 0\}$. Assume that for each $n \geq 1$, the family $\{Y_\varepsilon^n : \varepsilon > 0\}$ satisfies an LDP with speed ε and good rate function I_n. Set*

$$I(x) = \sup_{\delta>0} \liminf_{n\to\infty} \inf_{\{y:\rho(y,x)<\delta\}} I_n(y). \tag{B.5}$$

If I is a good rate function, and for any closed set F,

$$\inf_{x\in F} I(x) \leq \limsup_{n\to\infty} \inf_{y\in F} I_n(y), \tag{B.6}$$

then the family $\{X_\varepsilon : \varepsilon > 0\}$ satisfies an LDP with speed ε and good rate function I.

The basic theory of convergence of sequences of probability measures on a metric space have an analog in the theory of large deviations. Prohorov's theorem, relating compactness to tightness, has the following parallel that links exponential tightness to a partial LDP, defined below.

Definition B.6. A family of probability measures $\{\mathbb{P}_\varepsilon : \varepsilon > 0\}$ is said to satisfy the *partial LDP* if for every sequence ε_n converging to zero there is a subsequence ε_n' such that the family $\{\mathbb{P}_{\varepsilon_n'} : \varepsilon_n' > 0\}$ satisfies an LDP with speed ε_n' and a good rate function I'.

Remark: The partial LDP becomes an LDP if the rate functions associated with different subsequences are the same.

Theorem B.6. (Pukhalskii)
(1) *The partial LDP is equivalent to exponential tightness. Thus the partial LDP always holds on a compact space E.*
(2) *Assume that $\{\mathbb{P}_\varepsilon : \varepsilon > 0\}$ satisfies the partial LDP with speed ε, and for every x in E*

$$\begin{aligned}\lim_{\delta\to 0} \limsup_{\varepsilon\to 0} \varepsilon \log \mathbb{P}_\varepsilon\{\rho(y,x) \leq \delta\} \\ = \lim_{\delta\to 0} \liminf_{\varepsilon\to 0} \varepsilon \log \mathbb{P}_\varepsilon\{\rho(y,x) < \delta\} = -I(x).\end{aligned} \tag{B.7}$$

Then $\{\mathbb{P}_\varepsilon : \varepsilon > 0\}$ satisfies an LDP with speed ε and good rate function I.

For an $\mathbb{R}^d$-valued random variable Y, we define the *logarithmic moment generating function* of Y or its law μ as

$$\Lambda(\lambda) = \log E[e^{\langle \lambda, Y \rangle}] \text{ for all } \lambda \in \mathbb{R}^d \tag{B.8}$$

where $\langle , \rangle$ denotes the usual inner product in $\mathbb{R}^d$. $\Lambda(\cdot)$ is also called the *cumulant generating function* of Y. The Fenchel–Legendre transformation of $\Lambda(\lambda)$ is defined as

$$\Lambda^*(x) := \sup_{\lambda \in \mathbb{R}^d} \{\langle \lambda, x \rangle - \Lambda(\lambda)\}. \tag{B.9}$$

Theorem B.7. (Cramér) *Let $\{X_n : n \geq 1\}$ be a sequence of i.i.d. random variables in $\mathbb{R}^d$. Denote the law of $\frac{1}{n}\sum_{k=1}^n X_k$ by $\mathbb{P}_n$. Assume that*

$$\Lambda(\lambda) = \log E[e^{\langle \lambda, X_1 \rangle}] < \infty \textit{ for all } \lambda \in \mathbb{R}^d.$$

Then the family $\{\mathbb{P}_n : n \geq 1\}$ satisfies an LDP with speed $1/n$ and good rate function $I(x) = \Lambda^*(x)$.

The i.i.d. assumption plays a crucial role in Cramér's theorem. For general situations one has the following Gärtner–Ellis theorem.

Theorem B.8. (Gärtner–Ellis) *Let $\{Y_\varepsilon : \varepsilon > 0\}$ be a family of random vectors in $\mathbb{R}^d$. Denote the law of Y_ε by $\mathbb{P}_\varepsilon$. Define*

$$\Lambda_\varepsilon(\lambda) = \log E[e^{\langle \lambda, Y_\varepsilon \rangle}].$$

Assume that the limit

$$\Lambda(\lambda) = \lim_{n \to \infty} \varepsilon \Lambda_\varepsilon(\lambda/\varepsilon),$$

exists, and is lower semicontinuous. Set

$$\mathscr{D} = \{\lambda \in R^d : \Lambda(\lambda) < \infty\}.$$

If $\mathscr{D}$ has an nonempty interior $\mathscr{D}^\circ$ on which Λ is differentiable, and the norm of the gradient of $\Lambda(\lambda_n)$ converges to infinity, whenever λ_n in $\mathscr{D}^\circ$ converges to a boundary point of $\mathscr{D}^\circ$ (Λ satisfying these conditions is said to be essentially smooth)*, then the family $\{\mathbb{P}_\varepsilon : \varepsilon > 0\}$ satisfies an LDP with speed ε and good rate function $I = \Lambda^*$.*

The next result can be derived from the Gärtner–Ellis theorem.

Corollary B.9 *Assume that*

$$\{X_\varepsilon : \varepsilon > 0\}, \{Y_\varepsilon : \varepsilon > 0\}, \{Z_\varepsilon : \varepsilon > 0\}$$

are three families of real-valued random variables, all defined on the same probability space with respective laws

$$\{\mathbb{P}^1_\varepsilon : \varepsilon > 0\}, \{\mathbb{P}^2_\varepsilon : \varepsilon > 0\}, \{\mathbb{P}^3_\varepsilon : \varepsilon > 0\}.$$

If both $\{\mathbb{P}^1_\varepsilon : \varepsilon > 0\}$ *and* $\{\mathbb{P}^3_\varepsilon : \varepsilon > 0\}$ *satisfy the assumptions in Theorem B.8 with the same* $\Lambda(\cdot)$*, and with probability one*

$$X \le Y \le Z,$$

then $\{\mathbb{P}^2_\varepsilon : \varepsilon > 0\}$ *satisfies an LDP with speed* ε *and a good rate function given by*

$$I(x) = \sup_{\lambda \in \mathbb{R}} \{\lambda x - \Lambda(\lambda)\}.$$

Infinite-dimensional generalizations of Cramér's theorem are also available. Here we only mention one particular case: Sanov's theorem.

Let $\{X_k : k \ge 1\}$ be a sequence of i.i.d. random variables in $\mathbb{R}^d$ with common distribution μ. For any $n \ge 1$, define

$$\eta_n = \frac{1}{n} \sum_{k=1}^{n} \delta_{X_k}$$

where δ_x is the Dirac measure concentrated at x. The empirical distribution η_n belong to the space $M_1(\mathbb{R}^d)$ of all probability measures on $\mathbb{R}^d$ equipped with the weak topology. A well-known result from statistics says that when n becomes large one will recover the true distribution μ from η_n. Clearly $M_1(\mathbb{R}^d)$ is an infinite dimensional space. Denote the law of η_n, on $M_1(\mathbb{R}^d)$, by $\mathbb{Q}_n$. Then we have:

Theorem B.10. (Sanov) *The family* $\{\mathbb{Q}_n : n \ge 1\}$ *satisfies an LDP with speed* $1/n$ *and good rate function*

$$H(\nu|\mu) = \begin{cases} \int_{R^d} \log \frac{d\nu}{d\mu} d\nu, & \textit{if } \nu \ll \mu \\ \infty, & \textit{otherwise,} \end{cases} \tag{B.10}$$

where $\nu \ll \mu$ *means that* ν *is absolutely continuous with respect to* μ *and* $H(\nu|\mu)$ *is called the relative entropy of* ν *with respect to* μ*.*

References

1. M. Abramowitz and I.A. Stegun (1965). *Handbook of Mathematical Functions*. Dover Publ. Inc., New York.
2. D.J. Aldous (1985). *Exchangeability and Related Topics*, Ecole d'Été de Probabilités de Saint Flour, Lecture Notes in Math., Vol. 1117, 1–198, Springer-Verlag.
3. C. Antoniak (1974). Mixtures of Dirichlet processes with applications to Bayesian nonparameteric problems. *Ann. Statist.* **2**, 1152–1174.
4. M. Aoki (2008). Thermodynamic limit of macroeconomic or financial models: one- and two-parameter Poisson–Dirichlet models. *J. Econom. Dynam. Control* **32**, No. 1, 66–84.
5. R. Arratia (1996). Independence of prime factors: total variation and Wasserstein metrics, insertions and deletions, and the Poisson–Dirichlet process. Preprint.
6. R. Arratia, A.D. Barbour and S. Tavaré (1992). Poisson process approximations for the Ewens sampling formula. *Ann. Appl. Probab.* **2**, No. 3, 519–535.
7. R. Arratia, A.D. Barbour and S. Tavaré (1997). Random combinatorial structures and prime factorizations. *Notices of AMS* **44**, No. 8, 903–910.
8. R. Arratia, A.D. Barbour and S. Tavaré (1999). The Poisson–Dirichlet distribution and the scale invariant Poisson process. *Combin. Probab. Comput.* **8**, 407–416.
9. R. Arratia, A.D. Barbour and S. Tavaré (2003). *Logarithmic Combinatorial Structures: A Probabilistic Approach.* EMS Monographs in Mathematics, European Mathematical Society.
10. R. Arratia, A.D. Barbour and S. Tavaré (2006). A tale of three couplings: Poisson–Dirichlet and GEM approximations for random permutations. *Combin. Probab. Comput.* **15**, 31–62.
11. A.D. Barbour, L. Holst and S. Janson (1992). *Poisson Approximation*. Oxford University Press, New York.
12. J. Bertoin (2006). *Random Fragmentation and Coagulation Processes*. Cambridge Studies in Advanced Mathematics, 102. Cambridge University Press, Cambridge.
13. J. Bertoin (2008). Two-parameter Poisson–Dirichlet measures and reversible exchangeable fragmentation-coalescence processes. *Combin. Probab. Comput.* **17**, No. 3, 329–337.
14. P. Billingsley (1968). *Convergence of Probability Measures*, Wiley, New York.
15. P. Billingsley (1972). On the distribution of large prime divisors. *Period. Math. Hungar*. **2**, 283–289.
16. D. Blackwell and J.B. MacQueen (1973). Ferguson distribution via Pólya urn scheme. *Ann. Statist.* **1**, 353–355.
17. M.A. Carlton (1999). Applications of the two-parameter Poisson–Dirichlet distribution. Unpublished Ph.D. thesis, Dept. of Statistics, University of California, Los Angeles.
18. M.F Chen (2005). *Eigenvalues, Inequalities, and Ergodic Theory*. Probability and its Applications (New York). Springer-Verlag, London.
19. S. Coles (2001). *An Introduction to Statistical Modeling of Extreme Values*. Springer-Verlag, London.

S. Feng, *The Poisson–Dirichlet Distribution and Related Topics*,
Probability and its Applications, DOI 10.1007/978-3-642-11194-5,

20. D.A. Dawson (1993). *Measure-valued Markov Processes*, Ecole d'Été de Probabilités de Saint Flour, Lecture Notes in Math., Vol. 1541, 1–260, Springer-Verlag.
21. D.A. Dawson and S. Feng (1998). Large deviations for the Fleming–Viot process with neutral mutation and selection. *Stoch. Proc. Appl.* **77**, 207–232.
22. D.A. Dawson and S. Feng (2001). Large deviations for the Fleming–Viot process with neutral mutation and selection, II. *Stoch. Proc. Appl.* **92**, 131–162.
23. D.A. Dawson and S. Feng (2006). Asymptotic behavior of Poisson–Dirichlet distribution for large mutation rate. *Ann. Appl. Probab.* **16**, No. 2, 562–582.
24. D.A. Dawson and K.J. Hochberg (1979). The carrying dimension of a stochastic measure diffusion. *Ann. Probab.* **7**, 693–703.
25. D.A. Dawson and K.J. Hochberg (1982). Wandering random measures in the Fleming–Viot model. *Ann. Probab.* **10**, 554–580.
26. J.M. DeLaurentis and B.G. Pittel (1985). Random permutations and Brownian motion. *Pacific J. Math.* **119**, 287–301.
27. J. Delmas, J. Dhersin, and A. Siri-Jegousse (2008). Asymptotic results on the length of coalescent trees. *Ann. Appl. Probab.* **18**, No. 3, 997–1025.
28. A. Dembo and O. Zeitouni (1998). *Large Deviations Techniques and Applications*. Second edition. Applications of Mathematics, Vol. 38, Springer-Verlag, New York.
29. B. Derrida (1997). From random walks to spin glasses. *Physica D* **107**, 186–198.
30. J.D. Deuschel and D.W. Stroock (1989). *Large Deviations*. Academic Press, Boston.
31. K. Dickman (1930). On the frequency of numbers containing prime factors of a certain relative magnitude. *Arkiv för Matematik, Astronomi och Fysik* **22**, 1–14.
32. R. Dong, C. Goldschmidt and J. B. Martin (2006). Coagulation–fragmentation duality, Poisson–Dirichlet distributions and random recursive trees. *Ann. Appl. Probab.* **16**, No. 4, 1733–1750.
33. P. Donnelly (1986). Partition structures, Polya urns, the Ewens sampling formula, and the age of alleles. *Theor. Pop. Biol.* **30**, 271–288.
34. P. Donnelly and G. Grimmett (1993). On the asymptotic distribution of large prime factors. *J. London Math. Soc.* **47**, No. 3, 395–404.
35. P. Donnelly and P. Joyce (1989). Continuity and weak convergence of ranked and size-biased permutations on the infinite simplex. *Stoc. Proc. Appl*. **31**, 89–104.
36. P. Donnelly and P. Joyce (1991). Weak convergence of population genealogical processes to the coalescent with ages. *Ann. Prob.* **20**, No. 1, 322–341.
37. P. Donnelly, T.G. Kurtz, and S. Tavaré (1991). On the functional central limit theorem for the Ewens sampling formula. *Ann. Appl. Probab.* **1**, No. 4, 539–545.
38. P. Donnelly and T.G. Kurtz (1996a). A countable representation of the Fleming–Viot measure-valued diffusion. *Ann. Prob.* **24**, No. 2, 698–742.
39. P. Donnelly and T.G. Kurtz (1996b). The asymptotic behavior of an urn model arising in population genetics. *Stoc. Proc. Appl*. **64**, 1–16.
40. P. Donnelly and T.G. Kurtz (1999a). Genealogical processes for the Fleming–Viot models with selection and recombination. *Ann. Appl. Prob.* **9**, No. 4, 1091–1148.
41. P. Donnelly and T.G. Kurtz (1999b). Particle representations for measure-valued population models. *Ann. Prob.* **27**, No. 1, 166–205.
42. P. Donnelly and S. Tavaré (1986). The ages of alleles and a coalescent. *Adv. Appl. Prob.* **12**, 1–19.
43. P. Donnelly and S. Tavaré (1987). The population genealogy of the infinitely-many-neutral alleles model. *J. Math. Biol.* **25**, 381–391.
44. M.D. Donsker and S.R.S. Varadhan (1975). Asymptotic evaluation of certain Markov process expectations for large time, I. *Comm. Pure Appl. Math.* **28**, 1–47.
45. M. Döring and W. Stannat (2009). The logarithmic Sobolev inequality for the Wasserstein diffusion. *Probab. Theory Relat. Fields* **145**, No. 1-2, 189–209.
46. P. Dupuis and R.S. Ellis (1997). *A Weak Convergence Approach to the Theory of Large Deviations*. Wiley, New York.
47. R. Durrett (1996). *Probability: Theory and Examples*. Second edition. Duxbury Press.

48. R. Durrett (2008). *Probability Models for DNA Sequence Evolution*. Second edition. Springer-Verlag, New York.
49. E.B. Dynkin (1994). *An Introduction to Branching Measure-Valued Processes*. Vol. 6, CRM Monographs. Amer Math. Soc., Providence.
50. R. S. Ellis (1985). *Entropy, Large Deviations, and Statistical Mechanics*. Springer, Berlin.
51. S. Engen (1978). *Abundance Models with Emphasis on Biological Communities and Species Diversity*. Chapman-Hall, London.
52. A.M. Etheridge (1991). *An Introduction to Superprocesses*. University Lecture Series, Vol. 20, Amer. Math. Soc., Providence, RI.
53. A.M. Etheridge and P. March (1991). A note on superprocesses. *Probab. Theory Relat. Fields* **89**, 141–147.
54. A. Etheridge, P. Pfaffelhuber, and A. Wakolbinger (2006). An approximate sampling formula under genetic hitchhiking. *Ann. Appl. Probab.* **16**, No. 2, 685–729.
55. S.N. Ethier (1976). A class of degenerate diffusions processes occurring in population genetics. *Comm. Pure. Appl. Math.* **29**, 483–493.
56. S.N. Ethier (1990). The infinitely-many-neutral-alleles diffusion model with ages. *Adv. Appl. Probab.* **22**, 1–24.
57. S.N. Ethier (1992). Eigenstructure of the infinitely-many-neutral-alleles diffusion model. *J. Appl. Probab.* **29**, 487–498.
58. S.N. Ethier and R.C. Griffiths (1987). The infinitely-many-sites model as a measure-valued diffusion. *Ann. Probab.* **15**, No. 2, 515–545.
59. S.N. Ethier and R.C. Griffiths (1993). The transition function of a Fleming–Viot process. *Ann. Probab.* **21**, No. 3, 1571–1590.
60. S.N. Ethier and R.C. Griffiths (1993). The transition function of a measure-valued branching diffusion with immigration. In *Stochastic Processes. A Festschrift in Honour of Gopinath Kallianpur* (S. Cambanis, J. Ghosh, R.L. Karandikar and P.K. Sen, eds.), 71–79. Springer, New York.
61. S.N. Ethier and T.G. Kurtz (1981). The infinitely-many-neutral-alleles diffusion model. *Adv. Appl. Probab.* **13**, 429–452.
62. S.N. Ethier and T.G. Kurtz (1986). *Markov Processes: Characterization and Convergence*, John Wiley, New York.
63. S.N. Ethier and T.G. Kurtz (1992). On the stationary distribution of the neutral diffusion model in population genetics. *Ann. Appl. Probab.* **2**, 24–35.
64. S.N. Ethier and T.G. Kurtz (1993). Fleming–Viot processes in population genetics. *SIAM J. Control and Optimization* **31**, No. 2, 345–386.
65. S.N. Ethier and T.G. Kurtz (1994). Convergence to Fleming–Viot processes in the weak atomic topology. *Stoch. Proc. Appl.* **54**, 1–27.
66. W.J. Ewens (1972). The sampling theory of selectively neutral alleles. *Theor. Pop. Biol.* **3**, 87–112.
67. W.J. Ewens (2004). *Mathematical Population Genetics, Vol. I*, Springer-Verlag, New York.
68. J. Feng (1999). Martingale problems for large deviations of Markov processes. *Stoch. Proc. Appl.* **81**, 165–216.
69. J. Feng and T.G. Kurtz (2006). *Large Deviations for Stochastic Processes*. Amer. Math. Soc., Providence, RI.
70. S. Feng (2007). Large deviations associated with Poisson–Dirichlet distribution and Ewens sampling formula. *Ann. Appl. Probab.* **17**, Nos. 5/6, 1570–1595.
71. S. Feng (2007). Large deviations for Dirichlet processes and Poisson–Dirichlet distribution with two parameters. *Electron. J. Probab.* **12**, 787–807.
72. S. Feng (2009). Poisson–Dirichlet distribution with small mutation rate. *Stoch. Proc. Appl.* **119**, 2082–2094.
73. S. Feng and F.Q. Gao (2008). Moderate deviations for Poisson–Dirichlet distribution. *Ann. Appl. Probab.* **18**, No. 5, 1794–1824.
74. S. Feng and F.Q. Gao (2010). Asymptotic results for the two-parameter Poisson–Dirichlet distribution. *Stoch. Proc. Appl.* **120**, 1159–1177.

75. S. Feng and F.M. Hoppe (1998). Large deviation principles for some random combinatorial structures in population genetics and Brownian motion. *Ann. Appl. Probab.* **8**, No. 4, 975–994.
76. S. Feng and W. Sun (2009). Some diffusion processes associated with two parameter Poisson–Dirichlet distribution and Dirichlet process. *Probab. Theory Relat. Fields*, DOI 10.1007/s00440-009-0238-2.
77. S. Feng, W. Sun, F.Y. Wang, and F. Xu (2009). Functional inequalities for the unlabeled two-parameter infinite-alleles diffusion. Preprint.
78. S. Feng and F.Y. Wang (2007). A class of infinite-dimensional diffusion processes with connection to population genetics. *J. Appl. Prob.* **44**, 938–949.
79. S. Feng and J. Xiong (2002). Large deviation and quasi-potential of a Fleming–Viot process. *Elect. Comm. Probab.* **7**, 13–25.
80. T.S. Ferguson (1973). A Baysian analysis of some nonparametric problems. *Ann. Stat.* **1**, 209–230.
81. R.A. Fisher (1922). On the dominance ratio. *Proc. Roy. Soc. Edin.* **42**, 321–341.
82. W.H. Fleming and M. Viot (1979). Some measure-valued Markov processes in population genetics theory. *Indiana Univ. Math. J.* **28**, 817–843.
83. A. Gamburd (2006). Poisson–Dirichlet distribution for random Belyi surfaces. *Ann. Probab.* **34**, No. 5, 1827–1848.
84. A.J. Ganesh and N. O'Connell (1999). An inverse of Sanov's theorem. *Statist. Probab. Lett.* **42**, No. 2, 201–206.
85. A.J. Ganesh and N. O'Connell (2000). A large-deviation principle for Dirichlet posteriors. *Bernoulli* **6**, No. 6, 1021–1034.
86. J.H. Gillespie (1998). *Population Genetics: A Concise Guide*, The Johns Hopkins University Press, Baltimore and London.
87. J.H. Gillespie (1999). The role of population size in molecular evolution. *Theor. Pop. Biol.* **55**, 145–156.
88. A.V. Gnedin (1998). On convergence and extensions of size-biased permutations. *J. Appl. Prob.* **35**, 642–650.
89. A.V. Gnedin (2004). Three sampling formulas. *Combin. Probab. Comput.* **13**, No. 2, 185–193.
90. V.L. Goncharov (1944). Some facts from combinatorics. *Izvestia Akad. Nauk. SSSR, Ser. Mat.* **8**, 3–48.
91. R.C. Griffiths (1979a). On the distribution of allele frequencies in a diffusion model. *Theor. Pop. Biol.* **15**, 140–158.
92. R.C. Griffiths (1979b). A transition density expansion for a multi-allele diffusion model. *Adv. Appl. Probab.* **11**, 310–325.
93. R.C. Griffiths (1980a). Lines of descent in the diffusion approximation of neutral Wright–Fisher models. *Theor. Pop. Biol.* **17**, 35–50.
94. R.C. Griffiths (1980b). Unpublished note.
95. R.C. Griffiths (1988). On the distribution of points in a Poisson process. *J. Appl. Prob.* **25**, 336–345.
96. L. Gross (1993). Logarithmic Sobolev inequalities and contractivity properties of semigroups. *Dirichlet forms* (Varenna, 1992), 54–88, Lecture Notes in Math., Vol. 1563, Springer-Verlag, Berlin.
97. A. Guionnet and B. Zegarlinski (2003). Lectures on logarithmic Sobolev inequalities. *Séminaire de Probabilités, XXXVI*, 1–134, Lecture Notes in Math., Vol. 1801, Springer-Verlag, Berlin.
98. P.R. Halmos (1944). Random alms. *Ann. Math. Stat.* **15**, 182–189.
99. K. Handa (2005). Sampling formulae for symmetric selection, *Elect. Comm. Probab.* **10**, 223–234.
100. K. Handa (2009). The two-parameter Poisson–Dirichlet point process. *Bernoulli*, **15**, No. 4, 1082–1116.
101. J.C. Hansen (1990). A functional central limit theorem for the Ewens sampling formula. *J. Appl. Probab.* **27**, 28–43.

102. L. Holst (2001). The Poisson–Dirichlet distribution and its relatives revisited. Preprint.
103. F.M. Hoppe (1984). Polya-like urns and the Ewens sampling formula. *J. Math. Biol.* **20**, 91–94.
104. F.M. Hoppe (1987). The sampling theory of neutral alleles and an urn model in population genetics. *J. Math. Biol.* **25**, No. 2, 123–159.
105. R.R. Hudson (1983). Properties of a neutral allele model with intragenic recombination. *Theor. Pop. Biol.* **123**, 183–201.
106. W.N. Hudson and H.G. Tucker (1975). Equivalence of infinitely divisible distributions. *Ann. Probab.* **3**, No. 1, 70–79.
107. T. Huillet (2007). Ewens sampling formulae with and without selection. *J. Comput. Appl. Math.* **206**, No. 2, 755–773.
108. H. Ishwaran and L. F. James (2001). Gibbs sampling methods for stick-breaking priors. *J. Amer. Statist. Assoc.* **96**, No. 453, 161–173.
109. H. Ishwaran and L. F. James (2003). Generalized weighted Chinese restaurant processes for species sampling mixture models. *Statist. Sinica.* **13**, No. 4, 1211–1235.
110. J. Jacod, A. N. Shiryaev (1987). *Limit Theorems for Stochastic Processes*. Springer-Verlag, New York.
111. L.F. James (2003). Bayesian calculus for gamma processes with applications to semiparametric intensity models. *Sankhyā* **65**, No. 1, 179–206.
112. L.F. James (2005a). Bayesian Poisson process partition with an application to Bayesian Lévy moving averages. *Ann. Statist.* **33**, No. 4, 1771–1799.
113. L.F. James (2005b). Functionals of Dirichlet processes, the Cifarelli-Regazzini identity and beta-gamma processes. *Ann. Statist.* **33**, No. 2, 647–660.
114. L.F. James (2008). Large sample asymptotics for the two-parameter Poisson Dirichlet processes. In Bertrand Clarke and Subhashis Ghosal, eds, *Pushing the Limits of Contemporary Statistics: Contributions in Honor of Jayanta K. Ghosh*, 187–199, Institute of Mathematical Statistics, Collections, Vol. 3, Beachwood, Ohio.
115. L.F. James, A. Lijoi, and I. Prünster (2008). Distributions of linear functionals of the two parameter Poisson–Dirichlet random measures. *Ann. Appl. Probab.* **18**, No. 2, 521–551.
116. L.F. James, B. Roynette, and M. Yor (2008). Generalized gamma convolutions, Dirichlet means, Thorin measures, with explicit examples. *Probab. Surv.* **5**, 346–415.
117. P. Joyce and F. Gao (2005). An irrational constant separates selective under dominance from neutrality in the infinite alleles model, Preprint.
118. P. Joyce, S.M. Krone, and T.G. Kurtz (2002). Gaussian limits associated with the Poisson–Dirichlet distribution and the Ewens sampling formula. *Ann. Appl. Probab.* **12**, No. 1, 101–124.
119. P. Joyce, S.M. Krone, and T.G. Kurtz (2003). When can one detect overdominant selection in the infinite-alleles model? *Ann. Appl. Probab.* **13**, No. 1, 181–212.
120. P. Joyce and S. Tavaré (1987). Cycles, permutations and the structure of the Yule process with immigration. *Stoch. Proc. Appl.* **25**, 309–314.
121. S. Karlin and J. McGregor (1967). The number of mutant forms maintained in a population. *Proc. Fifth Berkeley Symp. Math. Statist. and Prob., L. LeCam and J. Neyman, eds*, 415–438.
122. S. Karlin and J. McGregor (1972). Addendum to a paper of W. Ewens. *Theor. Pop. Biol.* **3**, 113–116.
123. D.G. Kendall (1949). Stochastic processes and population growth. *J. Roy. Statist. Soc.* **11**, 230–264.
124. M. Kimura and J.F. Crow (1964). The number of alleles that can be maintained in a finite population. *Genetics* **49**, 725–738.
125. J.C.F. Kingman (1975). Random discrete distributions. *J. Roy. Statist. Soc. B* **37**, 1–22.
126. J.C.F. Kingman (1977). The population structure associated with the Ewens sampling formula. *Theor. Pop. Biol.* **11**, 274–283.
127. J.C.F. Kingman (1980). *Mathematics of Genetics Diversity*. CBMS-NSF Regional Conference Series in Appl. Math. Vol. 34, SIAM, Philadelphia.
128. J.C.F. Kingman (1982). The coalescent. *Stoch. Proc. Appl.* **13**, 235–248.

129. J.C.F. Kingman (1982). On the genealogy of large populations. *J. Appl. Prob.* **19A**, 27–43.
130. J.C.F. Kingman (1993). *Poisson Processes*. Oxford University Press.
131. N. Konno and T. Shiga (1988). Stochastic differential equations for some measure-valued diffusions. *Probab. Theory Relat. Fields* **79**, 201–225.
132. T.G. Kurtz (2000). Particle representations for measure-valued population processes with spatially varying birth rates. *Stochastic Models*, Luis G. Gorostiza and B. Gail Ivanoff, eds., CMS Conference Proceedings, Vol. 26, 299–317.
133. J.F. Le Gall (1999). *Spatial Branching Process, Random Snakes and Partial Differential Equations*. Lectures in Mathematics, ETH Zürich, Birkhäuser.
134. Z. Li, T. Shiga and L. Yao (1999). A reversible problem for Fleming–Viot processes. *Elect. Comm. Probab.* **4**, 71–82.
135. J. Lynch and J. Sethuraman (1987). Large deviations for processes with independent increments. *Ann. Probab.* **15**, 610–627.
136. E. Mayer-Wolf, O. Zeitouni, and M.P.W. Zerner (2002). Asymptotics of certain coagulation-fragmentation process and invariant Poisson–Dirichlet measures. *Electron. J. Probab.* **7**, 1–25.
137. J.W. McCloskey (1965). A model for the distribution of individuals by species in an environment. Ph.D. thesis, Michigan State University.
138. M. Möhle and S. Sagitov (2001). A classification of coalescent processes for haploid exchangeable population models. *Ann. Probab.* **29**, 1547–1562.
139. P.A.P. Moran (1958). Random processes in genetics. *Proc. Camb. Phil. Soc.* **54**, 60–71.
140. C. Neuhauser and S.M. Krone (1997). The genealogy of samples in models with selection. *Genetics* **145**, 519–534.
141. F. Papangelou (2000). The large deviations of a multi-allele Wright–Fisher process mapped on the sphere. *Ann. Appl. Probab.* **10**, No. 4, 1259–1273.
142. C.P. Patil and C. Taillie (1977). Diversity as a concept and its implications for random communities. *Bull. Int. Statist. Inst.* **47**, 497–515.
143. E.A. Perkins (1991). Conditional Dawson-Watanabe superprocesses and Fleming–Viot processes. *Seminar on Stochastic Processes*, Birkhäuser, 142–155.
144. E.A. Perkins (2002). *Dawson-Watanabe Superprocess and Measure-Valued Diffusions*, Ecole d'Été de Probabilités de Saint Flour, Lect. Notes in Math. Vol. 1781, 132–329.
145. M. Perman (1993). Order statistics for jumps of normalised subordinators. *Stoch. Proc. Appl.* **46**, 267–281.
146. M. Perman, J. Pitman and M. Yor (1992). Size-biased sampling of Poisson point processes and excursions. *Probab. Theory Relat. Fields* **92**, 21–39.
147. L.A. Petrov (2009). Two-parameter family of infinite-dimensional diffusions on the Kingman simplex. *Funct. Anal. Appl.* **43**, No. 4, 279–296.
148. J. Pitman (1992). The two-parameter generalization of Ewens' random partition structure. Technical Report 345, Dept. Statistics, University of California, Berkeley.
149. J. Pitman (1995). Exchangeable and partially exchangeable random partitions. *Probab. Theory Relat. Fields* **102**, 145–158.
150. J. Pitman (1996a). Partition structures derived from Brownian motion and stable subordinators. *Bernoulli* **3**, 79–96.
151. J. Pitman (1996b). Random discrete distributions invariant under size-biased permutation. *Adv. Appl. Probab.* **28**, 525–539.
152. J. Pitman (1996c). Some developments of the Blackwell-MacQueen urn scheme. *Statistics, Probability, and Game Theory*, 245–267, IMS Lecture Notes Monogr. Ser. 30, Inst. Math. Statist., Hayward, CA.
153. J. Pitman (1999). Coalescents with multiple collisions. *Ann. Probab.* **27**, 1870–1902.
154. J. Pitman (2002). Poisson–Dirichlet and GEM invariant distributions for split-and-merge transformations of an interval partition. *Combin. Probab. Comput.* **11**, 501–514.
155. J. Pitman (2006). *Combinatorial Stochastic Processes*, Ecole d'Été de Probabilités de Saint Flour, Lecture Notes in Math., Vol. 1875, Springer-Verlag, Berlin.
156. J. Pitman and M. Yor (1997). The two-parameter Poisson–Dirichlet distribution derived from a stable subordinator. *Ann. Probab.* **25**, No. 2, 855–900.

157. A.A. Puhalskii (1991). On functional principle of large deviations. In V. Sazonov and T. Shervashidze, eds, *New Trends in Probability and Statistics*, 198–218. VSP Moks'las, Moskva.
158. S. Sagitov (1999). The general coalescent with asynchronous mergers of ancestral lines *J. Appl. Probab.* **36**, 1116–1125.
159. A. Schied (1997). Geometric aspects of Fleming–Viot and Dawson-Watanabe processes. *Ann. Probab.* **25**, 1160–1179.
160. B. Schmuland (1991). A result on the infinitely many neutral alleles diffusion model. *J. Appl. Probab.* **28**, 253–267.
161. J. Schweinsberg (2000). Coalescents with simultaneous multiple collisions. *Electron. J. Probab.* **5**, 1–50.
162. J.A. Sethuraman (1994). A constructive definition of Dirichlet priors. *Statist. Sinica* **4**, No. 2, 639–650.
163. L.A. Shepp and S.P. Lloyd (1966). Ordered cycle length in a random permutation. *Trans. Amer. Math. Soc.* **121**, 340–351.
164. M. Sion (1958). On general minimax theorem. *Pacific J. Math.* **8**, 171–176.
165. W. Stannat (2000). On the validity of the log-Sobolev inequality for symmetric Fleming–Viot operators. *Ann. Probab.* **28**, No. 2, 667–684.
166. W. Stannat (2002). Long-time behaviour and regularity properties of transition semigroups of Fleming–Viot processes. *Probab. Theory Relat. Fields* **122**, 431–469.
167. W. Stannat (2003). On transition semigroup of (A,Ψ)-superprocess with immigration. *Ann. Probab.* **31**, No. 3, 1377–1412.
168. D.W. Stroock (1984). *An Introduction to the Theory of Large Deviations*. Springer-Verlag, New York.
169. F. Tajima (1983). Evolutionary relationship of DNA sequences in finite populations. *Genetics* **105**, 437–460.
170. S. Tavaré (1984). Line-of-descent and genealogical processes, and their applications in population genetics models. *Theor. Pop. Biol.* **26**, 119–164.
171. S. Tavaré (1987). The birth process with immigration, and the genealogical structure of large populations. *J. Math. Biol.* **25**, 161–168.
172. N.V. Tsilevich (2000). Stationary random partitions of positive integers. *Theor. Probab. Appl.* **44**, 60–74.
173. N.V. Tsilevich and A. Vershik (1999). Quasi-invariance of the gamma process and the multiplicative properties of the Poisson–Dirichlet measures. C. R. Acad. Sci. Paris, t. 329, Série I, 163–168.
174. N.V. Tsilevich, A. Vershik, and M. Yor (2001). An infinite-dimensional analogue of the Lebesgue measure and distinguished properties of the gamma process. *J. Funct. Anal.* **185**, No. 1, 274–296.
175. S.R.S. Varadhan (1984). *Large Deviations and Applications*. SIAM, Philadelphia.
176. A.M. Vershik and A.A. Schmidt (1977). Limit measures arising in the asymptotic theory of symmetric groups, I. *Theory Probab. Appl.* **22**, No. 1, 70–85.
177. A.M. Vershik and A.A. Schmidt (1978). Limit measures arising in the asymptotic theory of symmetric groups, II. *Theory Probab. Appl.* **23**, No. 1, 36–49.
178. J. Wakeley (2006). *An Introduction to Coalescent Theory*. Roberts and Company Publishers.
179. F.Y. Wang (2004). *Functional Inequalities, Markov Semigroups and Spectral Theory*. Science Press, Beijing.
180. G.A. Watterson (1974). The sampling theory of selectively neutral alleles. *Adv. Appl. Prob.* **6**, 463–488.
181. G.A. Watterson (1976). The stationary distribution of the infinitely-many neutral alleles diffusion model. *J. Appl. Prob.* **13**, 639–651.
182. G.A. Watterson (1977). The neutral alleles model, and some alternatives. *Bull. Int. Stat. Inst.* **47**, 171–186.
183. G.A. Watterson and H.A. Guess (1977). Is the most frequent allele the oldest? *Theor. Pop. Biol.* **11**, 141–160.
184. G.A. Watterson (1984). Lines of descent and the coalescent. *Theor. Pop. Biol.* **26**, 119–164.

185. S. Wright (1931). Evolution in Mendelian populations. *Genetics* **16**, 97–159.
186. S. Wright (1949). Adaption and selection. In *Genetics, Paleontology, and Evolution*, ed. G.L. Jepson, E. Mayr, G.G. Simpson, 365–389. Princeton University Press.
187. L. Wu (1995). Moderate deviations of dependent random variables related to CLT. *Ann. Prob.* **23**, No. 1, 420–445.
188. K.N. Xiang and T.S. Zhang (2005). Small time asymptotics for Fleming–Viot processes. *Infin. Dimens. Anal. Quantum Probab. Relat. Top.* **8**, No. 4, 605–630.

Index

S. Feng, *The Poisson–Dirichlet Distribution and Related Topics*,
Probability and its Applications, DOI 10.1007/978-3-642-11194-5,