

CIRJE-F-445

**Thermodynamic Limits of Macroeconomic or
Financial Models: One-and Two-Parameter
Poisson-Dirichlet Models**

Masanao Aoki
University of California, Los Angeles

October 2006

CIRJE Discussion Papers can be downloaded without charge from:

<http://www.e.u-tokyo.ac.jp/cirje/research/03research02dp.html>

Discussion Papers are a series of manuscripts in their draft form. They are not intended for circulation or distribution except as indicated by the author. For that reason Discussion Papers may not be reproduced or distributed without the written consent of the author.

Thermodynamic Limits of Macroeconomic or Financial Models: One- and Two-Parameter Poisson-Dirichlet Models

Masanao Aoki*

Department of Economics

University of California, Los Angeles

Fax Number 310-825-9528, e-mail aoki@econ.ucla.edu

June, 2005; last revision July,2006

Abstract

This paper examines asymptotic behavior of two types of economic or financial models with many interacting heterogeneous agents. They are one-parameter Poisson-Dirichlet models, also called Ewens models, and its extension to two-parameter Poisson-Dirichlet models.

The total number of clusters, and the components of partition vectors (the number of clusters of specified sizes), both suitably normalized by some powers of model sizes, of these classes of models are shown to be related to the Mittag-Leffler distributions.

Their behavior as the model sizes tend to infinity (thermodynamic limits) are qualitatively very different. In the one-parameter models, the number of clusters, and components of partition vectors are both self-averaging, that is, their coefficients of variations tend to zero as the model sizes become very large, while in the two-parameter models they are not self-averaging, that is, their coefficients of variations do not tend to zero as model sizes becomes large.

Key Words: Poisson-Dirichlet distributions; Mittag-Leffler distributions; Thermodynamic limits; Non-self averaging phenomena, Power laws.

*The author is grateful for many helps he received from M. Sibuya. Comments by Ted Theodosopoulos, Thomas Lux and an anonymous referee's comments have been useful in revising the original draft. I thank them all.

Introduction

In old industrial organization literature, several measures of industrial concentration have been used to detect if a given firm or industry is monopolistic or not.¹ One such measure is Herfindahl, or Herfindahl-Hirschman index of concentration. It is defined by

$$H = \sum_i x_i^2,$$

where x 's are the fractions of "shares" of markets or sales by sectors or firms. By definition, the x 's are non-negative and sum to one, $\sum_i x_i = 1$.

As we discuss shortly, this literature also used a rudimentary version of the size-biased sampling scheme as a test of oligopoly. This measure of concentration is used in both the domestic and foreign trade context. It is sometimes (mistakenly) called Gini-index.²

In the physics literature, the same measure has been used by Derrida-Flyvbjerg (1987) in discussing relative sizes of basins of attractions of Kaufman random maps and random dynamics in statistics and physics. Sornette (2000, 370) states "By now, the standard method to detect non self-averaging (i.e., breaking of ergodicity) is to study the quantity $Y = \sum_\alpha W_\alpha^2$," where $\sum_\alpha W_\alpha = 1$ " where W_α is the weight, i.e., fraction. Note that Sornette's Y is the same as H above.

Sornette uses the ratio of sample variance normalized by the square of the sample mean of an extensive random variable X as the measure of its sample-to-sample fluctuations, which he denotes by $D_N(X)$ where N is the "size" of the model. Letting the model size N tend to infinity is called thermodynamic limit in the physics literature, Sornette (2000, 369).

Note that his $D_N(X)$ is the same as the square of the coefficient of variation, that is the sample standard deviation divided by the sample mean in the statistics and econometric literature. In this paper we use the coefficients of variations as the measure of sample-to-sample variability. Namely, if the coefficient of variation of X does not go to zero as model size increases, X is not self-averaging, and if the coefficient tends to zero, then X is self-averaging.

If a random variable X is self-averaging, then, its coefficient of variation approaches zero as the sample size tends to infinity. If not self-averaging, its coefficient of variation does not approach zero, no matter how large the sample size becomes. In such cases, a sample, however large is never a good representative of the whole ensemble. In other words, the behavior of X is sample-dependent.

This paper examines two types of models and find one type self-averaging but the other not self-averaging. More specifically, this paper compares the thermodynamic limits of two classes of Poisson-Diarrichlet models, that is,

¹See for example Scherer (1980) which describes many case studies.

²Sometimes it is called Gini-Simpson index of diversity. See Hirschman (1960) about the origin and mis-attribution of this notion to Herfindahl. In the population genetics literature H is called homozygosity, see Ewens (1972).

one-parameter Poisson-Dirichlet model, $PD(\theta)$,³ and its extension to the two-parameter Poisson-Dirichlet model, $PD(\alpha, \theta)$, where α is between 0 and 1, and θ is positive.⁴ In these two classes of models we focus on the number of clusters and components of the partition vectors as defined in Aoki (2002), for example.

We find that the number of clusters and the components of partition vectors of $PD(\theta)$ models are self-averaging but those of $PD(\alpha, \theta)$ models are not. In the third type of model which does not belong to $PD(\theta)$ or $PD(\alpha, \theta)$, we examine asymptotic properties as time tends to infinity. In this model, cluster sizes remain bounded. Their coefficients of variations, however, do not tend to zero.

In the physics literature, Mekjian and Chase (1997) have used two-parameter models. They refer to the work by Pitman (1996). There are other works in the physics literature, in particular the papers by Derrida-Flyvbjerg (1987), which is a one parameter model. Random maps mentioned in these references are further elaborated in Samuelsson and Trevein (2005). See also Derrida (1994).⁵ There are other papers in the physics literature that deal with random partitions. Higgs (1995) have noted the similarities of some physical distributions and power laws, and mention population genetics papers by Ewens in particular. There are many papers on stick-breeding version of the residual allocation processes, such as Krapivvsky, Grosse, and B. Nadin (2002). They do not touch on connections with the two-parameter Poisson-Dirichlet distributions, however.

In macroeconomics and finance, agents of different characteristics or strategies are of different types and form separate clusters and affect aggregate behavior. In this paper, we therefore explore more broadly economic implications of long-run relations, more specifically their thermodynamic limits in the sense of (econ)physics, that may exist among non-self averaging economic or financial variables. See Sornette (2000,240).

Alternative characterization of one- and two-parameter Ewens distribution is given by Garibaldi, Costantini, Viarengo, and Donadio, (2004, 2006).

In the first part of this paper and in appendices some basic notions on random partitions is presented from the literature of combinatorial stochastic processes, J. Pitman (1996, 2002) and Yamato and Sibuya (2000), including notions of size-biased permutation, residual allocation models, of frequency spectrum and structure distribution, Mittag-Leffler probability density and power-laws.

³A simple application to shares of a market by two types of agents, using one-parameter Poisson-Dirichlet distribution (also called Ewens distribution, Ewens (1972, 1979, 1990)) has been made by Aoki (2000a, 2000b). See also Aldous (1985).

⁴See, for example, Kingman (1993), Carlton (1999), Holst (2001), Pitman (1999, 2002), and Pitman and Yor (1996), among others.

⁵Derrida (1994) has later added some material on residual allocation models.

Distributional Stability: Invariance under Size-biased Permutation

We introduce the notion of invariance under size-biased sampling or permutation in the statistics literature as a proper concept of the size distribution of types in statistical equilibrium.

Heuristically this notion may arise in the following way: Suppose that fractions of "shares" are arranged in decreasing order, $x_1 > x_2 > \dots$. We may be interested in the question of how large is the share of the second type, excluding the presence of the first, that is the largest type. This is the fraction $x_2/(1 - x_1)$. Analogously, we may be interested in the share of the i -th largest type excluding or correcting for the effects of the first through the $(i - 1)$ th shares, given by $x_i/(1 - x_1 - \dots - x_{i-1})$. Actually, this is one of the ways industrial organization economists measured the concentration of industries, see Scherer 1980), even though they did not know of the notion of the size-biased sampling or permutation. This is precisely what is involved in size-biased sampling.

More formally, we consider the set of all possible fractions (p_1, p_2, \dots) where p_i , the fraction of type i agents, is positive, and the fractions sum to 1, $\sum_i p_i = 1$. Suppose that one agent is sampled. The probability that the first sampled agent is of type j is

$$\Pr(\hat{p}_1 = p_j | p_1, p_2, \dots, p_n) = p_j, : j = 1, 2, \dots, n.$$

This first pick is called the size-biased pick, because types of agents with larger fraction are more likely to be sampled. This equation says that the sample is taken in proportion to the sizes of various types. Having picked $\hat{p}_1, \dots, \hat{p}_k$, the next sampled agent is of type n with probability given by

$$\Pr(\hat{p}_{k+1} = p_n | \hat{p}_i, i = 1, 2, \dots, k; p_1, p_2, \dots) = \frac{p_n}{1 - \hat{p}_1 - \hat{p}_2 - \dots - \hat{p}_k},$$

provided that $p_n \neq \hat{p}_i, i = 1, 2, \dots, k$. The collection, $\{\hat{p}_j\}$, is called size-biased sampling or permutation abbreviated as SBP.

Since distributions of agents by types are more useful when they are in statistical equilibrium, we define that the set of fractions is invariant under size biased permutation (abbreviated as ISBP) if

$$\{\hat{p}_n\} =^d \{p_n\},$$

where $=^d$ means equality in distribution.

Pitman (1996) considered $\{p_n\}, p_n > 0, a.s.,$ for all $n, \sum_n p_n = 1$, such that $\{p_n\}$ are distributed as RAM (residual allocation model) for independent random variables $W_i, i = 1, 2, \dots,$ that is, ps are generated by the following formula

$$p_1 = W_1, p_2 = W_2(1 - W_1), \dots, p_n = W_n(1 - W_1)(1 - W_2) \dots (1 - W_{n-1}).$$

Note that $p_1 = W_1, p_2/(1 - p_1) = W_2, \dots, p_n/(1 - p_1 - \dots - p_{n-1}) = W_n$ are independent.

Let α and θ be such that $0 \leq \alpha < 1$, and $\theta > -\alpha > 0$. Let W_i be Beta distributed random variable, $Be(1 - \alpha, \theta + i\alpha)$, where random variable X has density $Be(a, b)$ when the density is given by

$$f_X(x) = \frac{1}{B(a, b)} x^{a-1} (1-x)^{b-1},$$

for $0 < x < 1$, where $B(a, b) = \Gamma(a)\Gamma(b)/\Gamma(a+b)$.

Then the collection of fractions, $\{p_n\}$, is said to have a $GEM(\alpha, \theta)$ distribution.⁶

Then he showed that $\{p_n\}$ is invariant under size-biased permutation if and only if $\{p_n\}$ is distributed as $GEM(\alpha, \theta)$.

Next, arrange samples by order statistics, i.e., we reorder \hat{p}_i , $i = 1, 2, \dots$ as

$$p_{(1)} > p_{(2)} > \dots$$

When $\{p_n\}$ is distributed as $GEM(\alpha, \theta)$, then the ranked sequence $\{p_{(n)}\}$ is said to have the two-parameter Poisson Dirichlet distribution, $PD(\alpha, \theta)$.

To summarize, if fractions of agents of type n are given by $\{p_n\}$, $p_n > 0$, a.s., and $\sum_n p_n = 1$, the size-biased permutation of $PD(\alpha, \theta)$ is a $GEM(\alpha, \theta)$, and the ranked sequence of a $GEM(\alpha, \theta)$ is a $PD(\alpha, \theta)$. Furthermore, $GEM(\alpha, \theta)$ is ISBP. See Carlton (1999) for more details.

With $\alpha = 0$, $PD(\alpha, \theta)$ reduces to the Ewens distribution, denoted from now on by $PD(\theta)$. For earlier works see Kingman (1978), Perman, Pitman, and Yor (1992), and Pitman and Yor (1997) on earlier works.

A short description of structural distribution and frequency spectrum is found in Appendix.

Number of Clusters in two-parameter Poisson-Dirichlet Distributions

The probabilities of new types entering models in $PD(\theta)$, and the number of clusters have been applied for example in Aoki (2002, p.176, App. A.5). In the two-parameter Poisson-Dirichlet distribution the conditional probabilities for the number of clusters in a sample of size n , K_n is given by

$$\Pr(K_{n+1} = k + 1 | K_1, \dots, K_n = k) = \frac{k\alpha + \theta}{n + \theta}, \quad (1)$$

and

$$\Pr(K_{n+1} = k | K_1, \dots, K_n = k) = \frac{n - k\alpha}{n + \theta}. \quad (2)$$

In other words, the random variable K_n is the number of different types of agents present in a sample of size n . Eq.(1) means that the $(n+1)$ th entrant is a new type. Eq.(2) means that it is one of the previously existing types. Hence the number of clusters does not change.

⁶The name GEM was given by Ewens to honor the pioneers, Griffiths, Engen, and McCloskey.

Let the probability for $K_n = k$ be denoted by $q_{\alpha\theta}(n, k)$. From (1) and (2) it can be recursively computed using the two conditional probability equations above

$$q_{(\alpha\theta)}(n+1, k) = \frac{(n-k\alpha)}{(n+\theta)}q_{(\alpha\theta)}(n, k) + \frac{\theta+(k-1)\alpha}{n+\theta}q_{(\alpha\theta)}(n, k-1), \quad (3)$$

for $1 \leq k \leq n$. The expressions for the boundary $K_n = 1$ for all n , and that of $K_n = n$ are given by the expression

$$q_{(\alpha\theta)}(n, 1) = \frac{(1-\alpha)(2-\alpha)\cdots(n-1-\alpha)}{(\theta+1)(\theta+2)\cdots(\theta+n-1)},$$

and

$$q_{(\alpha\theta)}(n, n) = \frac{(\theta+\alpha)(\theta+2\alpha)\cdots(\theta+(n-1)\alpha)}{(\theta+1)(\theta+2)\cdots(\theta+n-1)}.$$

These expressions generalize the recurrence relation for the one-parameter $PD(\theta)$. In the one-parameter case, $\theta/(\theta+n)$ is a probability that the $(n+1)$ th agent that enters the model is of a new type, and $n/(\theta+n)$ is the probability that the next agent is one of the types already in the model.

In the one-parameter case, $q_\theta(n, k) := P(K_n = k)$ is governed by the recurrence relation

$$q_\theta(n+1, k) = \frac{n}{n+\theta}q_{n,k} + \frac{\theta}{\theta+n}q_{n,k-1}.$$

The solution of this recurrence equation is expressible as

$$q_{n,k} = \frac{c(n, k)\theta^k}{\theta^{[n]}},$$

where $\theta^{[n]} := \theta(\theta+1)\cdots(\theta+n-1) = \frac{\Gamma(\theta+n)}{\Gamma(\theta)}$, and $c(n, k)$ is the unsigned (signless) Stirling number of the first kind. It satisfies the recursion

$$c(n+1, k) = nc(n, k) + c(n, k-1).$$

Since $q_{n,k}$ sums to one with respect to k we have

$$\theta^{[n]} = \sum_{k=1}^n c(n, k)\theta^k. \quad (4)$$

See Aoki (2002, p.208) for example on the Stirling numbers, and their combinatorial interpretations.

In the two-parameter $PD(\alpha, \theta)$ case, the probability of the number of clusters is given by

$$P_{\alpha,\theta}(K_n = k) = \frac{\theta^{[k,\alpha]}}{\alpha^k \theta^{[n]}} c(n, k; \alpha), \quad (5)$$

where

$$\theta^{[k,\alpha]} := \theta(\theta+\alpha)(\theta+2\alpha)\cdots(\theta+(k-1)\alpha),$$

and the expression $c(n, k; \alpha)$ generalizes the signless Stirling number of the first kind of one-parameter situation. This is called generalized Stirling number of the first kind. See Charalambides (2002).

Let $S_\alpha(n, k) := \frac{1}{\alpha^k} c(n, k; \alpha)$. It satisfies the recursion

$$S_\alpha(n+1, k) = (n - k\alpha)S_\alpha(n, k) + S_\alpha(n, k-1).$$

Instead of (4) we have

$$\theta^{[n]} = \sum_{k=1}^n S_\alpha(n, k) \theta^{[k, \alpha]}. \quad (6)$$

Pitman (1999) obtained its asymptotic expression as

$$S_\alpha(n, k) \sim \frac{\Gamma(n)}{\Gamma(k)} n^{-\alpha} \alpha^{1-k} g_\alpha(x),$$

where $k \sim xn^\alpha$. Here, g_α is the Mittag-Leffler (α) function. This function is discussed in the next section.

Asymptotic Behavior of Cluster Sizes

We collect here some known asymptotic facts about cluster sizes as $n \rightarrow \infty$.

The number of clusters K_n

$$EK_n = \frac{\theta}{\alpha} \left[\frac{(\theta + \alpha)^{[n]}}{\theta^{[n]}} - 1 \right],$$

where we note that

$$\frac{(\theta + \alpha)^{[n]}}{\theta^{[n]}} = \frac{\Gamma(\theta)}{\Gamma(\theta + \alpha)} \frac{\Gamma(\theta + \alpha + n)}{\Gamma(\theta + n)}.$$

Applying the asymptotic expression for the Gamma function for large n

$$\frac{\Gamma(n+a)}{\Gamma(n)} \sim n^a,$$

to the above expression, we have an asymptotic expression,

$$E\left(\frac{K_n}{n^\alpha}\right) \sim \frac{\Gamma(\theta + 1)}{\alpha \Gamma(\theta + \alpha)}. \quad (7)$$

Yamato and Sibuya (2000) obtained the asymptotic value of the variance of K_n/n^α ,

$$\text{var}(K_n/n^\alpha) \sim \frac{\Gamma(\theta + 1)}{\alpha^2} \gamma_{\alpha, \theta} \geq 0, \quad (8)$$

where

$$\gamma_{\alpha, \theta} := \frac{\theta + \alpha}{\Gamma(\theta + 2\alpha)} - \frac{\Gamma(\theta + 1)}{[\Gamma(\theta + \alpha)]^2}. \quad (9)$$

Note that

Fact: $\gamma_{0,\theta} = 0$.

This fact is important in the long-run behavior of components of the partition vectors, to be discussed in the next subsection.

Actually they calculate more generally

$$\lim E\left(\frac{K_n}{n^\alpha}\right)^r = \mu'_r,$$

where μ'_r is the r -th moment of the generalized Mittag-Leffler distribution with density

$$g_{\alpha,\theta} := \frac{\Gamma(\theta + 1)}{\Gamma(\theta/\alpha + 1)} x^{\frac{\theta}{\alpha}} g_\alpha(x),$$

where $\theta/\alpha > -1$, and where $g_\alpha(x)$ is the Mittag-Leffler (α) density function. It is known that this function is uniquely determined by the moment conditions

$$\int_0^\infty x^p g_\alpha(x) dx = \frac{\Gamma(p + 1)}{\Gamma(p\alpha + 1)},$$

for all $p > -1$. The moments of this density satisfy the sufficient condition for the density to be uniquely determined by the set of all moments so that the method of moments applies. Note that the integral of $g_{\alpha,\theta}$ over the interval from zero to infinity is 1, as it should be. See Pollard (1946), for example, for the expression of the density. See also Blumenfeld and Mandelbrot (1997) who credit Feller (1949) as the original source.

Mittag-Leffler distributions

As we discuss more fully later, Pitman (2002, Sec. 3) has an even stronger result:

$$K_n/n^\alpha \rightarrow \mathcal{L}, a.s.,$$

where the expression \mathcal{L} has the density

$$\frac{d}{ds} P_{\alpha,\theta}(\mathcal{L} \in ds) = g_{\alpha,\theta}$$

where letting $\eta = \frac{\theta}{\alpha}$ we define

$$g_{\alpha,\theta}(s) := \frac{\Gamma(\theta + 1)}{\Gamma(\eta + 1)} s^\eta g_\alpha(s),$$

where $s > 0$, and where $g_\alpha = g_{\alpha,0}$ is the Mittag-Leffler density

$$g_\alpha(s) = \frac{1}{\pi} \sum_{k=1}^{\infty} \left[\frac{\Gamma(k\alpha)}{\Gamma(k)} \sin(k\pi\alpha) (-s)^{k-1} \right].$$

We note that

$$\mu'_1 = E_{\alpha,\theta}(\mathcal{L}) = \Gamma(\theta + 1)/\alpha\Gamma(\theta + \alpha),$$

and

$$\mu'_2 = E_{\alpha,\theta}(\mathcal{L}^2) = \Gamma(\theta + 1)(\theta + \alpha)/\alpha^2\Gamma(\theta + 2\alpha).$$

Hence the variance of \mathcal{L} is given as $\mu'_2 - (\mu'_1)^2$.

For the record we have **Fact**

$$E\left(\frac{K_n}{n^\alpha} = \frac{\Gamma(\theta + 1)}{\alpha\Gamma(\alpha + \theta)}, \text{var}_{\alpha,\theta}\left\{\frac{K_n}{n^\alpha}\right\} = \text{var}_{\alpha,\theta}\mathcal{L} = \frac{\Gamma(\theta + 1)}{\alpha^2}\gamma_{\alpha,\theta}. \quad (10)$$

The partition vector \mathbf{a}

Denote the partition vector by $\mathbf{a} = (a_1, a_2, \dots)$, where we recall that a_i is the number of distinct clusters of size i , hence $\sum_i a_i = K_n$, and $\sum_i i a_i = n$. Yamato and Sibuya obtain the limit of the first component, a_1

$$\lim E\left[\frac{a_1}{n^\alpha}\right] = \frac{\Gamma(\theta + 1)}{\Gamma(\theta + \alpha)},$$

and

$$\lim \text{var}\left(\frac{a_1}{n^\alpha}\right) = \Gamma(\theta + 1)\gamma_{\alpha,\theta} \geq 0.$$

In fact a_j/n^α are all non-self averaging, as well as $j a_j/n^\alpha$, where $j a_j$ is the total number of agents in the clusters of size j . Note that their variances are all zero with $\alpha = 0$, that is the asymptotic variance of a_j/n^α are all zero in $PD(\theta)$ models.

Fact Combining the above with (7) we have

$$\frac{E a_1}{E K_n} \rightarrow \alpha.$$

The expression a_i/n^α , $i \geq 1$ are all non-self averaging with $0 < \alpha < 1$.

Sibuya (2005) used Formula 6.1.41 in Abramovitz and Stegun (1965) to obtain the asymptotic expression

$$E\left(\frac{a_j}{n^\alpha}\right) \approx \frac{(1 - \alpha)^{[j-1]}}{j!} \frac{\Gamma(\theta + 1)}{\Gamma(\theta + \alpha)} + O(n^{-1}).$$

We state the asymptotic behavior of K_n/n^α and a_j/n^α as

Proposition: As in (11)

$$\lim \text{var}_{\alpha,\theta}(K_n/n^\alpha) = \text{var}_{\alpha,\theta}(\mathcal{L}),$$

and

$$\lim \text{var}_{\alpha,\theta}(a_j/n^\alpha) = \alpha^2 \text{var}_{\alpha,\theta}(\mathcal{L}) = \Gamma(\theta + 1)\gamma_{\alpha,\theta}.$$

They also show that covariances of components of the partition vectors are non-self averaging with positive α values:

Fact:

$$\lim \text{Cov}\left(\frac{a_i}{n^\alpha}, \frac{a_j}{n^\alpha}\right) = \Gamma(\theta + 1)\gamma_{\alpha,\theta} \times \frac{(1 - \alpha)^{[i-1]}}{i!} \frac{(1 - \alpha)^{[j-1]}}{j!} > 0, \quad : \alpha > 0.$$

It is also known that

$$\frac{j! a_j/n^\alpha}{\alpha(1 - \alpha)^{[j-1]}} \rightarrow^d \mathcal{L}. \quad (11)$$

We have

$$E\left(\frac{a_j}{n^\alpha} | K_n = k\right) \sim \frac{(1-\alpha)^{[j-1]}}{j!} (1-j/n)^{-(1+\alpha)} \times \xi,$$

where ξ depends on $g(\alpha, \theta)$.

The number of clusters, K_n , is spread among the components of the partition vector, $a_i, i = 1, 2, \dots, n$ at the proportion $\alpha(1-\alpha)^{[j-1]}/j!, 0 < \alpha < 1$. Devroye (1993) calls this Sibuya distribution.

We also note that

$$\text{Lim} \frac{E(a_i)}{E(K_n)} = \frac{\alpha^2}{\Gamma(\theta + \alpha)\gamma_{\alpha, \theta}}.$$

We note that a_j/K_n is self-averaging for all $j = 1, \dots, n$. Yamato and Sibuya also examined the clusters of size k or less

$$K[1, k] := a_1 + a_2 + \dots + a_k,$$

and the number of agents in $K[1, k]$, denoted by $N[1, k]$ and obtained their limiting expressions as

$$\frac{K[1, k]}{n^\alpha} \rightarrow^d \left\{1 - \frac{(1-\alpha)^{[k]}}{k!}\right\} \mathcal{L},$$

and

$$\frac{N[1, k]}{n^\alpha} \rightarrow^d \alpha \frac{(2-\alpha)^{[k-1]}}{(k-1)!} \mathcal{L},$$

Sibuya also notes that

$$\left\{\frac{a_1}{n^\alpha}, \frac{2a_2}{n^\alpha} \dots \frac{ka_k}{n^\alpha}\right\}$$

converges in distribution to a sequence of random variables depending on \mathcal{L} as

$$\left\{1, \frac{(1-\alpha)}{1!}, \dots, \frac{(1-\alpha)^{[k-1]}}{(k-1)!}\right\}.$$

In $PD(\alpha)$ it is known that

$$\frac{K_n - \theta \ln(n)}{\sqrt{\theta \ln(n)}} \rightarrow N(0, 1).$$

Hence $(K_n/\ln(n))$ is self-averaging.

Almost sure convergence

Denote by $a_j(n)$ the number of clusters of size j when there are n agents in the model. We noted earlier that $\sum_{j=1}^n j a_j(n) = n$, and $K_n := \sum_j a_j(n)$ is the total number of clusters formed by the total of n agents.

By Rouault (1976, 1978)

$$\frac{a_j(n)}{K_n} \rightarrow \frac{\alpha \Gamma(j - \alpha)}{\Gamma(1 - \alpha) j!}, a.s.$$

Recalling that $K_n/n^\alpha \rightarrow \mathcal{L}, a.s.$, we have

$$a_j(n)/n^\alpha \rightarrow \frac{\alpha \Gamma(j - \alpha)}{\Gamma(1 - \alpha) j!} \mathcal{L}, a.s.$$

where

$$\frac{a_j(n)}{K_n} \rightarrow \frac{\alpha}{j!} P_{\alpha, j},$$

where

$$P_{\alpha j} = \frac{\Gamma(j - \alpha)}{\Gamma(1 - \alpha)},$$

for every $j = 1, 2, \dots$ a.s. as n goes to infinity, and that $a_j(n) \sim P_{\alpha, j} \mathcal{L} n^\alpha$ in a two-parameter Poisson-Dirichlet case.

Local Limit Theorem

Suppose N independent positive random variables X_i , $i = 1, 2, \dots, N$ are normalized by their sum $S_N = X_1 + \dots + X_N$

$$x_i = X_i/S_N, i = 1, \dots, N,$$

so that

$$Y_1 := \sum_i x_i = 1.$$

Suppose that the probability density of X_i is such that it has a power-law tail,

$$\rho(x) \sim Ax^{-1-\mu},$$

with $0 < \mu < 1$. Then, $S_N/N^{1/\mu}$ has a stable distribution (called Lévy distribution).

Pitman's formula for the probability of $K_n = k$, with $k \sim sn^\alpha$ indicates that the power law n^α which is $2\alpha < 2$ or $2\alpha = 1 + \mu$ with $0 < \mu < 1$, the case in Derrida.

With the 2-parameter PD distribution satisfying the power law condition, Derrida's conclusion that the H s are non-self averaging applies to this case as well.

Estimating the Parameters

Carlton (1999) and Sibuya (2005) are the only systematic source on estimating the parameters of two-parameter Poisson-Dirichlet distributions.

With $\alpha = 0$, Ewens had shown that K_n is the sufficient statistics for θ . Carlton discusses the case where α is known and θ unknown. He derives the asymptotic distribution of the maximum likelihood estimate of θ , given n samples.

Lemma

Given α in $(0,1)$, the maximum-likelihood estimate of θ , $\hat{\theta}_n$ is given by

$$\psi(1 + \hat{\theta}_n/\alpha) - \alpha\psi(1 + \hat{\theta}_n) \rightarrow \log S, \text{ a.s.}$$

Here ψ is the digamma function.

With θ known, and α unknown, Carlton proves

Lemma

Let $\{A_1, \dots, A_n\}$ be distributed according to the two-parameter Ewens distribution of size n . (His Eq. (4.2) on page 55.) Then,

$$\hat{\alpha}_n = \frac{\log K_n}{\log n} \rightarrow \alpha \text{ a.s.}$$

Sibuya uses the conditional probability distribution of the partition vector components, given that $\sum_i a_i = k$, and expresses the distribution

$$P(\mathbf{a} | \sum a_j = k) = \frac{1}{S_\alpha(n, k)} \frac{n!}{\prod a_j!} \prod_j \left\{ \frac{(1-\alpha)^{[j-1]}}{j!} \right\}^{a_j}$$

which is proportional to

$$\exp\left\{-\sum \frac{j}{2(j-2)!} a_j\right\} \alpha + O(\alpha^2)$$

and test the hypothesis $\alpha = 0$, against the alternative hypothesis $\alpha < 0$.

Sibuya proposes the rejection region

$$\sum \frac{j}{2(j-2)!} a_j > \text{const.} \cdot k.$$

When both parameters are unknown, the estimation problem is apparently unsolved.

Some Potential Applications

In physics literature, Derrida (1994, 1997 a, b) sketched a derivation that the expected values of $Y_k = \sum_i x_i^k, k = 2, 3, \dots$ can be calculated for mean field spin glass models using the Parisi replica approach, and remarkably the formula is the same as the GEM model described above.

In the rest of this section we focus on economic examples.

Example 1 Instead of treating all possible configurations equi-probably, we weigh them by Poisson-Dirichlet distributions in this example. Consider a firm composed of total of n basic units. These units are organized into divisions or sections. The total number of divisions is K_n . The number of divisions of size j is denoted by $a_j(n)$. We observe that $\sum_j a_j(n) = K_n$, and $\sum_j j a_j(n) = n$.

The parameter $(\theta + (k-1)\alpha)/(n+\theta)$ is the probability that a new division (new product) is being introduced, as shown in the recursion equation (1). We note that, using his notation

$$\left\{ \frac{a_1(n)}{n^\alpha}, \frac{a_2(n)}{n^\alpha}, \dots, \frac{a_n(n)}{n^\alpha} \right\} \rightarrow^d \mathcal{L}\{\alpha, (1-\alpha)^{[1]}/2!, \dots, (1-\alpha)^{[n-1]}/n!\}.$$

Pitman shows that the largest division has fraction $P_{(1)}$ of agents, $P_{(1)} \sim Z$, the second largest division has fraction $P_{(2)} \sim Z/2^\alpha$, and so on. The random variable Z may be expressed as $Z^{-\alpha} = \Gamma(1-\alpha)\mathcal{L}$.

We note that the number of divisions of size j is a decreasing function of j

Example 2 Markov chains with transition rates of (1) and (2) for the case with $\alpha = 0$ have been simulated. The model in Sec. 8.6 of Aoki (2002) was constructed to examine the effects of demand managements. In this example, we strip the model of this aspect and merely show the effects of parameter θ which control the rates by which new sectors are created in the model. We can alternatively interpret θ as parameter which controls of sector size or introduction of new goods by a given firm or sector.

In the model, sectors want to respond to excess demand signals they receive. Sectors interpret positive excess demands as opportunities to expand their production, and negative excess demands as signals to contract their production. The model is constructed in such a way each sector is impacted by the changes in production by any other sector through externality of excess demands. Thus, production change by any single sector will impact the excess demand signals they observe. For this reason, only one sector which acts first realizes its desire to change its production, and the pattern starts all over again.

In short, only the sector with the shortest holding (sojourn) time acts according with the sign of excess demand. Parameter θ controls the rate of entries of new sectors. With larger values of θ , the model is expected to grow faster. Because of the construction of the model cyclical variations of output (GDP) is superimposed on the growth path. This is indeed what simulations show. See Aoki (2002, p.113-117).

Example 3: GDP Growth rate Scaling of GDP growth rates was considered by Canning, Amaral, Lee, Meyer, and Stanley (1998). They showed that the standard deviation of the GDP growth rate may scale as $Y^{-\beta}$, with β about 0.15. Here, we heuristically explain how their finding may be explained using a random partition framework.

We modify the model of Huang and Solomon (2001) and apply the same procedures to estimate the growth rate of real GDP.⁷ View the real economy as composed of K sectors of various sizes. Stochastically one or more of the sectors experience what we call elementary events, the aggregate of which yields the real growth of the economy, leading to its random growth rates. To be simple one may assume that the individual elementary growth of sectors is random $\lambda = 1 + g$, where $g = \pm\gamma$ randomly with some positive γ . Further, we adopt the mechanism of Huang and Solomon that a random number τ of this type of elementary events are experienced in a unit of calendar time.

⁷Their focus is on financial sector, not real sector. See Aoki and Yoshikawa (2006 a, b).

The random growth rate is the composite effects of these random elementary events.

We refer the detail of the mechanism to their paper, and mention only that the growth rate will be exponential only if the number of changes τ is less than some critical value τ_c , and change in GDP has a power law density with index $-(1 + \alpha)$.

The value of α is defined to be the ratio of minimum and average real consumption in the model $q = c_{min}/c_{average}$, and is tied to α by

$$\alpha \approx 1/(1 - q),$$

when K is sufficiently larger than $e^{1/q}$, due to inherent normalization conditions of densities involved.

For example, setting $q = 0.25$ leads to $\alpha = 1.33$, and K must be such that $K \gg e^4 > 55$. The value of τ_c is defined by $(N/2q)^\alpha$. With τ less than τ_c , the growth rate r can be shown to have the density

$$p(r) = C \exp(-a|r - r_m|),$$

for $r > r_m$, with a different constant for the case $r < r_m$.

The deviation of r is then related to variability of K and τ , among others. From this one can deduce that the average deviation in the growth rates is basically determined by percentage changes of the size of the largest cluster which can be related to the GDP when the productivity is assumed not to vary too much, and the conclusion follows that the standard deviation of the growth rate is $Y^{-\mu}$ with μ less than 1. See Aoki and Yoshikawa (2006a, b) for detail.

Example 4: Non-ergodic endogenous growth model Interpret the branching model of Feng and Hoppe (1998) as a model of an economy composed of several sectors. In this model innovations arrive either to one of the existing sectors or an innovation creates a new sector. There are two parameters θ and α as in the two-parameter *PD* model. Existing sector of size n_i grows at rate proportional to $(n_i - \alpha)/(n + \theta)$, while a new sector is created at the rate proportional to $(\theta + k\alpha)/(n + \theta)$ where $n = \sum_i n_i$, when there are k sectors in the economy.

Note that the larger sectors tend to have larger growth rates, and the rate of growth of the whole economy tends to accelerate. These could be interpreted as endogenous innovation activities by larger sectors and larger economy.

Example 5: Disequilibrium theory of long run profits. Iwai's model has more than two sectors with different productivity coefficients. His paper is too long and involved to give a thumb-nail sketch here. Instead we offer three quotes from his paper to explain what he does.

...while both the differential growth rates among different efficiency firms and the diffusion of better technologies through imitations push the state of technology towards uniformity, the punctuated appearance of technological innovations disrupts this equilibrating tendency.

... over a long passage of time these conflicting microscopic forces will balance each other in a statistical sense and give rise to a long-run

distribution of relative efficiencies across firms. This long-run distribution will in turn allow us to deduce an *upward-sloping* long-run supply curves...

This paper has challenged this long-held tradition in economics. It has introduced a simple evolutionary model which is capable of analyzing the development of the industry's state of technology as a dynamic interplay among many a firm's growth, imitation and innovation activities. And it has demonstrated that what the industry will approach over a long passage of time is not a classical or neoclassical equilibrium of uniform technology but a statistical equilibrium of technological disequilibria which maintains a relative dispersion of efficiencies in a statistically balanced form. Positive profits will never disappear from the economy nomatter how long it is run. 'Disequilibrium' theory of 'long-run profits' is by no means a condtradition in terms.

We see that our random partiton framework along the line of Aoki, Nakano, and Yoshida (2004) and Aoki, Nakano and Ono (2006) can be applied to at least three types of firms, and their tail distribution may satisfy power laws to substantiate Iwai's claim by using long-run in time rather than the thermodynamic limits.

Concluding Remarks

In physics non-self-averaging phenomena abound. In traditional microeconomic foundations of economics, one deals almost exclusively with well-posed optimization problems for the representative agents with well defined peaks and valleys of the cost functions. It is also taken for granted that as the number of agents goes to infinity, any unpleasant fluctuations vanish and well defined deterministic macroeconomic relations prevail. In other words, non-self-averaging phenomena are not in the mental pictures of average macro- or microeconomists.

However, we know that as we go to problems which require agents to solve some combinatorial optimization problems, this nice picture may disappear. In the limit of the number of agents going to infinity some results are sample-dependent and deterministic results will not follow. Some of this type of phenomena have been reported in Aoki (1996, Sec. 7.1.7) and also in Aoki (1996, p. 225) where Derrida's random energy model was introduced to the economic audience. Unfortunately it did not catch the attention of the economic audiences. See Mertens (2000). This paper is another attempt at exposing non-self-averaging phenomena in economics. We also mention a possibility of extending the phrase to cover existence of non-degenerate distributions with time going to infinity. What are the implications if some economic models have non-self averaging property? For one thing, it means that history matters. We cannot blindly try for larger size samples in the hope that we obtain better estimates.

The examples in this paper are just an indication of the potential of this approach of using exchangeable random partition methods. It is the opinion of this author that subjects such as in the papers by Fabritiis, Pammolli, and Riccaboni (2003), or by Amaral et al (1998) could be re-examined from

the random combinatorial partition approach with profit. Another example is Sutton (2002). He modeled independent business in which the business sizes vary by partitions of integers to discuss the dependence of variances of firm growth rates. He assumed each partition is equally likely, however. Use of random partitions discussed in this paper may provide more realistic or flexible framework for the question he examined.

Finally, the key question in applications to macroeconomic or financial modelings of the random partition approach is "What are the most likely combinations of the values of $K_n = k$ and a_j ?" This question appears too complicated to answer analytically at this time. Some simulations would help.

References

- Aldous, D.J., (1985), "Exchangeability and related topics" in Lecture notes in mathematics, No.1117, Springer-Verlag, Berlin
- Amaral, Luis A. Nunes, S. V. Buldyrev, S. Havlin, M.A. Salinger and H.E. Stanley,(1998) "Power law scaling for a system of interacting units with complex internal structure, *Phys. Rev. Lett.*,**80**, 1385–1388.
- Aoki, M., (2000a), "Open models of share markets with two dominant types of participants, ", *J.Econ. Behav. Org.* **49** 199-216.
- , (2000b), "Cluster size distributions of economic agents of many types in a market", *J. Math Anal. Appl.*, **249**,32-52.
- , (2002), *Modeling Aggregate Behavior and Fluctuations in Economics: Stochastic Views of Interacting Agents* , Cambridge Univ. Press, New York.
- , (2003), "Models with Random Exchangeable Structures and Coexistence of Several Types of Agents in the Long-Run: New Implementations of Schumpeter's Dynamics" Tech. Report, Research Initiative and Development, Chuo Univ. Tokyo, Dec. 2003.
- ,(2006), "Non-ergodic endogenous growth models in macroeconomics," , working paper.
- Aoki, M., T. Nakano, and G. Yoshida, (2004), "Two sector Schumpeterian model of Industry" Mimeo, Dept. Physics, Chuo University, Tokyo.
- , —, and K. Ono, (2006), "Simulation Results of a Two-sector Model of Innovation and Immitation"
- Aoki, M., and H. Yoshikawa (2002). "Demand saturation-creation and economic growth " *J. Econ. Behav. Org.*,**48**, 127-154.
- , H. Yoshikawa (2006a), *Reconstructing Macroeconomics: A Perspective from Statistical Physics and Combinatorial Stochastic Processes* , forthcoming from Cambridge University Press, New York.
- , and—, (2006b), "Stock prices and real economy: Exponential and Power-Law Distributions", forthcoming invited paper *Journal of Interaction and Coordination of Heterogeneous Agents*, No.1, Volume 1. Springer-Verlag, New York.
- Arratia, R., and S. Tavaré (1992) "The cycle structure of random permutation", *Ann. prob.* **20**,1567-1591.

Blumenfeld, R., and B. B. Mandelbrot (1997), "Lévy dusts, Millag-Leffler statistics, mass fractal lacunarity, and perceived dimension," *Phy. Rev. E*, **56**, 112-118.

Charambides, A. C., *Enumerative Combinatorics*, Chapman Hall/CRC, 2002, london

Canning, D., L.A.N.Amaral, Y. Lee, M. Meyer, H. E. Stanley, (1998) "Scaling the volatility of GDP growth rates", *Econ. Lett.*, **60**, 335-341.

Carlton, M. A. (1999) *Applications of the Two-Parameter Poisson-Dirichlet Distribution* Ph.D. thesis, Dept. Math. Univ. California, Los Angeles

Derrida, B.,(1981) "Random energy model", *Phys. Rev. B*, **24**, 2613-2626.

Derrida, B., (1994), "From Random Walks to Spin Glasses", *Physica D*, **107**, 166-198.

—, (1997) " Non-self-averaging effects in sums of random variables, spin glasses, random maps and random walks" in *On Three Levels Micro-Meso- and Macro-Approaches in Physics*, M. Fannes, C. Maies, and A. Verberre (eds), Plenum Press, New York

—, and H. Flyvbjerg (1987), "The random map model: a disordered model with deterministic dynamics", *J. Physique*, **48**,971-978.

Devroye, (1993), " A triptych of discrete distributions related to the stable law", *Probab. Letters* **18**, 349–351.

Ewens, W. J. (1972), "The sampling theory of selectively neutral alleles," *Theor. Pop. Biol.*, **3**, 87-112.

—, (1979), *Mathematical Population Genetics*, Springer-Verlag, Berlin.

—, (1990) "Population genetics theory —The past and the future", in *Mathematical and statistical problems in evolution* ed. by S. Lessard, Kluwer Academic Pulbishers, Boston.

Fabritiis, G.de, F.Pammolli, and M. Riccaboni, (2003) "On size and growth of business firms," *Physica A***324**, 38–44.

Feller, W., (1949) "Fluctuation theory of recurrent events" *Trans. Am. Math. Soc.***67**, 98-119.

Feng, S., F.M. Hoppe (1998), "Large deviation principle for some random combinatorial structures in population genetics and Brownian motion," *Ann. Appl. Prob.* **6** 975–994.

Garibaldi,U, D. Costantini, and P. Viarengo (2004), "A finitary characterization of the Ewens sampling formula," *Advances in Complex Systems*, **7**,265–284.

—, —, S.Donadio, and P. Viarengo (2006), "Herding and Clustering in Economics: The Yule-Zipf-Simon Model," *Computational Econ.* **27** 115–134.

Griffiths, R., (2005), " Poisson Dirichlet Process" Version 0.1 Mimeo.

Higgs, P. (1995), "Frequency distributions in population genetics parallel those in statistical physics", *Phy. Rev. E* **51**, 95-101.

Hirschman, A.O. (1960)"The paternity of an index", *Amer. Econ. Review* bf 54 761.

Holst, L. (2001), "The Poisson-Dirichlet Distribution and Its Relatives Revisited," Tech. Report Dept. Math. Royal Inst. Technology, Stockholm

Huang, Z-F, and S.Solomon (2001), " Power, Lévy, Exponential and Gaussian Regimes in Autocatalytic Financial Systems," *Euro.Phys. Jou.*

B, **20**, 601–607.

Iwai, K. (1997). "A contribution to the evolutionary theory of innovation, imitation and growth," *J. Econ. Behav. Org.*, **43**, 167-198

— (2001), "Schumpeterian dynamics: A disequilibrium theory of long run profits" in L.Punzo (ed) *Cycles, Growth and Structural Change: Theories and empirical evidence*, Routledge, London and New York.

Karlin, S. (1967), "Central limit theorem for certain infinite urn schemes", *J. Math. Mech.*, **17**, 373-401.

Kingman, J.F.C. (1978), "The representation of partition structure", *J. London Math. Soc.*, **18**, 374–380.

—, (1993), *It Poisson Processes*, Clarendon Press, Oxford UK.

Krapivsky, P.L., I. Grosse, and E. Ben-Nadin, (2000). "Scale invariance and lack of self-averaging in fragmentation", *Phy. Rev. E*, **61**, R993-R996.

Mekjian, A.Z., and K.C. Chase (1997), "Disordered systems, power laws and random processes," *Phy. Lett A* **229**, 340-346.

Mertens, S. (2000) "Random costs in combinatorial optimization", *Phy. Rev. Lett.* **84**, 1347-1350.

Pitman, D (2002) "Sequential construction of random partitions" Lecture notes, St. Flour Summer Institute.

—, (1996) "Random discrete distributions invariant under size-biased permutation", *Adv. Appl. Prob.* **28**, 525–539.

—, (1999) "Brownian motion, bridge, excursion and meander characterized by sampling at independent uniform time," *Electron. J. Probab.* **4**, Paper 11, 1-33.

—, and M. Yor, (1992). "Arcsine laws and interval partition derived from a stable subordinator," *Proc. London Math. Soc.* (3) **65** 326–356.

—, and —. (1997), "The two-parameter Poisson-Dirichlet distribution derived from a stable subordinator," *Ann. prob.* **25**, 855-900.

Samuelsson, B., and C. Troein. (2005), "Random maps and attractors in random Boolean networks," arXiv:cond-mat/0505481 v1, 19 May 2005.

Sherer, F. M. (1980) *Industrial Market Structure and Economic Performance*, 2nd ed. Houghton Mifflin Co. Boston

Sornette, D., (2000). *Critical Phenomena in Natural Sciences*, Springer, New York.

Sutton, J., (2002) "The variance of firm growth rates: the "scaling" puzzle", *Physica A*, **312**, 577-590.

Yamamoto H., M. Sibuya, (2000), "Moments of some statistics of Pitman sampling formula", *Bull. Inform. Cybernet.* **32**, 1-10.

Watterson, G. A., (1974), "The Sampling Theory of Selectively Neutral Alleles," *Adv. Appl. Probab.* **6**, 463–488.

—, and H. A. Guess (1977), "Is the most frequent allele the oldest ?" *Theor. Pop. Biol.* **11**, 141-160.

Appendix: Structural distribution and frequency spectrum

The structural distribution, F , of $\{p_n\}$, is defined by Engen to be the distribution on $(0, 1]$ of the first size-biased pick, that is the first term of a size-biased permutation of the distribution of agents by type, $\{p_n\}$, that is \hat{p}_1 . The importance of this first pick is demonstrated by the lemma below of Pitman and Yor (1997).

When $\{p_n\}$ is distributed as a two parameter Poisson-Dirichlet distribution $PD(\alpha, \theta)$, let W_1 be distributed as $Be(1 - \alpha, \theta + \alpha)$ (Beta distribution). We drop subscript 1 from W_1 from now on. The first size-biased pick is $\hat{p}_1 = W$ as we have shown above. The structural distribution is important because it shows that \hat{p}_1 summarizes the distribution of $\{p_n\}$ as shown next.

Lemma: For any positive measurable function $g(t) \sim O(t)$ as t goes to zero,

$$\begin{aligned} E[g(W)/W] &= E[g(\hat{p}_1)/\hat{p}_1] \\ &= E\left\{E\sum_i \frac{g(p_n)}{p_n} Pr(\hat{p}_1 = p_n | p_1, p_2, \dots)\right\} \\ &= E\left(\sum \frac{g(p_n)}{p_n} p_n\right) = E\left[\sum g(p_n)\right]. \end{aligned}$$

By the above lemma, the expected value of any positive measurable function g is expressible in terms of the structural distribution as

$$E\left(\sum_n g(p_n)\right) = \int_0^1 \frac{g(v)}{v} F(dv).$$

Pitman (1996) pointed out that $v^{-1}F(dv)$ is the frequency spectrum. If one takes g to be $I(a < v < b)$, this expression gives the average number of n such that $a < p_n < b$, hence $v^{-1}F(dv)$ is the same as the frequency spectrum in population genetics literature which is a measure of cluster size distribution. See Ewens (1979). Aoki (2002, p.173, 2002a) has some elementary economic applications of this notion. In words, the frequency spectrum is the expected number of types with fraction in the interval $(x, x + dx)$.

Given order statistics of cluster sizes governed by $PD(\theta)$, $x_1 > x_2 > \dots$, the largest size x_1 has the density

$$f(x_1) = \theta x_1^{-1} (1 - x_1)^{\theta-1},$$

for x_1 in the range $1/2 < x_1 < 1$, that is when the largest cluster is more than 1/2 of the whole.⁸ This density behaves like x_1^{-1} for small x_1 . This indicates that there are many types with small fractions and $f(x)$ is not normalizable. However, $g(x) = xf(x) = \theta(1 - x)^{\theta-1}$ is normalizable. This function is interpreted as the probability that a randomly selected sample is of the type with fraction in $(x, x + dx)$.

⁸The expression is more complicated when x_1 is less than 1/2. See Watterson and Guess (1977).

The two largest fractions, x_1 and x_2 have the joint density

$$f(x_1, x_2) = \theta^2(x_1x_2)^{-1}(1 - x_1 - x_2)^{\theta-1},$$

when the two sizes are such that $0 < x_1 + x_2 < 1$, and more importantly when

$$\frac{x_2}{1 - x_1} > \frac{1}{2}.$$

Note that similar inequalities arise in size-biased permutation. See Aoki (2002, Sec. 10.6) for heuristic derivations based on Watterson and Guess (1977).⁹

In economic applications we are more interested in a few types with large shares, such as the ones discussed in Aoki (2000a).

For the one-parameter Poisson-Dirichlet process, the expected sizes of the three largest clusters are shown in the next table (see Griffiths (2005))

θ	largest	second	third
0.1	0.935	.059	.005
0.5	.758	.171	.049
1.0	.624	.210	.088

For example, with $\theta = 0.1$, the expected size of the largest and the second largest clusters sum to 99 per cent of the whole agents. With $\theta = 1/2$, the sum is about 93 per cent.

⁹Karlin (1967) focussed on the situation with many types of small probabilities such that $\beta(x) = x^{-\gamma}L(x)$, with $0 < \gamma < 1$, and where $\beta(x) = \sum_i^\infty I(p_n \geq x)$, and where $L(\cdot)$ is some slowly varying function.