CIRJE-F-388

Cluster Size Distributions of Heterogeneous Economic Agents: Are there non-self-averaging phenomena in economics?

Masanao Aoki
University of California

November 2005

Cluster Size Distributions of Heterogeneous Economic Agents: Are there non-self-averaging phenomena in economics?

Masanao Aoki
Department of Economics
University of California, Los Angeles
Fax Number 310-825-9528, e-mail aoki@econ.ucla.edu

June, 2005; revised Nov.,2005

Abstract

This paper outlines the applications of one- and two-parameter Poisson-Dirichlet distributions to describe stationary statistical distributions of clusters of agents by types. We discuss how the notion of residudal allocation processes in statistics and population genetics literature also arises as stick-breaking processes in the physics literature. The phenomena of self-(non-) averaging in the physics literature are analogous to long-run non-vanishing of profits or variances of capital sizes in some disequilibrium economic dynamics. We offer an economic interpretation of the physical notion of non-self-averaging as something that refers to the existence of long-run disequilibrium phenomena in economics, rather than thermodynamic limits in statistical physics, since both involve non-vanishing of variances as the size or the time goes to infinity.

Introduction

In old literature on industrial market structure and economic performance, several measures of industrial concentration have been used to design some tests to answer questions if a given industry is monopolistic or not. See for example Scherer (1980). One example of the index is called Herfindahl, or Herfindahl-Hirschman index of concentration. It is defined as the sum of squares of fractions of shares, i.e.,

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

where x_i is the fraction of "share" of markets or sales by sector i or firm i. By definition x_i is positive, and

$$\sum_i x_i = 1.$$

This meassure of concentration is used in both domestic and foregin trade context. It is sometimes (mistakenly) called Gini-index.[1]

The question of concentration is that of distribution of cluster sizes of agents by types. A simple application of shares of 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).

This paper develops further the original ideas in these papers by applying some of the results in the recent combinatorial stochastic process literature, in particular results by Pitman (2002). It also makes connection with the physics literature, in particular the papers by Derrida-Flyvbjerg (1987), and Derrida (1994, 1997)[2] They also deal with the questions of relative sizes of basins of attraction in random dynamics in the former and the question of residual allocation processes in the latter. There are other papers in the physics literature that deal with random partitions. Mekjian and Chase (1997)and Higgs (1995) have discussed cluster size distributions and power laws, and mention population genetics papers by Ewens in particular. Frontera, Goicoechea, Rafols, and Vivies (1995), and Krapivsky, Grosse, and Ben Nadin (2002) discuss partitions and fragmentations, that is residual allocation processes explicitly.

Physicists' concern in the context of this paper appears to be focussed on the presence of non-self-averaging phenomena in models. In this paper, the question is whether variances of some economic concentration measure asymptotically go to zero as the sizes (number of agents) goes to infinity, or better yet in economic applications, as time goes to infinity.

In the first part of this paper we introduce the reader to some basic notions on random partitions from the literature of combinatorial stochastic processes, in particular the works by statisticians, J. Pitman (1996, 2002) and M. Carlton (1999), such as invariance of distributions of agents by types or categories under size-biased permutation, notions of frequency spectrum and structure distribution. We describe the notions of distributions of sizes of clusters of agents by types, in particular one- and two-parameter Poisson-Dirichlet distributions. We then show a remarkable fact that the so-called residual allocation models (RAM) arise in combinatorial stochastic processes and in physics, namely, Kingman's model of partition structures of agents by types, Kingman (1978) and Derrida's derivation in physics, based on power laws.

[1] Sometimes it is called Gini-Simpson index of divesity. 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). This expression also arises in random maps and ramdom dynamics in statistics and physics, see Aldous (1985), or Derrida-Flyvbjerg (1987)

[2] Derrida has added some material on residual allocation models in his unpublished version.

Invariance under Size-biased Permutation

We introduce the notion of invariance under size-biazed sampling or permutation in the statistics literature as a proper concepts of distribution of sizes of types in statisstical equilibrium. Heuristically this notion may arise in the following way:

Suppose that fractions of "shares" are ordered in decreasing order, $x_1 > x_2 > \cdots$. We may be interested in the question of how large is the share of the second type, exclusive of the first, that is the largest type. This is the fraction $x_2/(1-x_1)$. Analogously, we may be interested in the i-th largest type excluding or correcting for the effects of the first through the $(i-1)$th shares, given by $x_i/(1-x_1-\cdots x_{i-1})$. Actually, this is one of the ways economists measured the concentration of industries, 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, suppose that there are N types of agents with fractions p_i, $i = 1, 2, \ldots, N$, with a large N. Suppose that one agent is sampled. The probability that the first sampled agent is of type j is

$$\Pr(\hat{p}_1 = p_j | p_1, p_2, \ldots, p_N) = p_j.$$

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

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

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

Since distributions of agents by types are more useful when they are in statistical equilibrium, at least for models with infinitely many types, we define the fractions are invariant under size biased permutation (abbreviated as ISBP) when

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

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, $i = 1, 2, \ldots$, that is ps are generated by the following formula

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

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

Let α and θ be such that 0 □ $\alpha < 1$, and $\theta + \alpha > 0$. Let W_i be Beta distriuted random variable, $Be(1-\alpha, \theta+i\alpha)$, where we say a 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 $\{p_n\}$ is said to have a $GEM(\alpha, \theta)$ distribution.[3] With $\alpha = 0$, the above reduces to the one-parameter Poisson-Dirichlet distribution, due to Kingman (1978), which formalizes what Ewens conjectured (1972).

Then he showed that $\{p_n\}$ is ISBP if and only if $\{p_n\}$ is distributed as $GEM(\alpha, \theta)$.

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

$$p_{(1)} > p_{(2)} > \cdots .$$

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 is 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, Carlton (1999).

With $\alpha = 0$, $PD(\alpha, \theta)$ reduces to the Ewens distribution, denoted from now on by $PD(0, \theta)$ or by $PD(\theta)$.

Frequency Spectrum

In population genetics literature, there is a measure of cluster size distribution called frequency spectrum. 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 $PD(\theta)$, $x_1 > x_2 > \cdots$, 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.[4] This density behaves like x_1^{-1} for small x. 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 $(x, x + dx)$.

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

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

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

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

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). [5]

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.

When $\{p_n\}$ is distriubted as a two parameter Poisson-Dirichlet distribution $PD(\alpha, \theta)$, let W be distributed as $Be(1 - \alpha, \theta + \alpha)$ (Beta distribution). Then, the first size-biased pick is $\hat{p_1} = W$ as we have shown above. Then, we have a result due to Pitman and Yor (1997):

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

$$E[g(W)/W] = E[g(\hat{p_1}/\hat{p_1}]$$

$$= E\{E \sum_i \frac{g(p_n)}{p_n} Pr(\hat{p_1} = p_n | p_1, p_2, \ldots)\}$$

$$= E(\sum \frac{g(p_n)}{p_n} p_n) = E[\sum g(p_n)].$$

Structural Distribution

The structural distribution, F, of $\{p_n\}$, is defined by Engen to be the distribution on $(0, 1]$ of the first term of a size-biased permutation of the distribution of agents by type, $\{p_n\}$, denoted as $\hat{p_1}$.

Pitman (1996) pointed out that $v^{-1}F(dv)$ is the frequency spectrum. By lemma above, the expected value of any positive measurable function g is expressible in terms of the structural distribution as

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

[5]Karlin (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 $L(.)$ is some slowly varying function.

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.

Let K_n be the number of types of clusters in this random population of $[n] := \{1, 2, \ldots, n\}$. In other words, K_n is the number of distinct values observed in the sequence of random partitions of $[n]$, $\{A_1, A_2, \cdots, A_n\}$, where A_j is the number of types in the sample with exactly j agents of the same type. Clearly, A_i is non-negative, and sums to K_n.

We note that

$$\Pr(K_n = 1) = \int_0^1 v^{n-1} F(dv).$$

Engen derived that $E(K_n)$ is a polynomial moment of F of degree $n - 1$

$$EK_n = \int_0^1 \frac{(1 - (1 - v)^n)}{v} F(dv).$$

Number of Clusters in two-parameter Poisson-Dirichlet Distribution

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 entries and exits are given by

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

and

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

Pitman (2002) shows that the probability for $K_n = k$, $q(n, k)$, can be recursively computed from the forward equation

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

for $1 \leq k \leq n$, given the boundary formula

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

and

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

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

In this case, $q_{n,i} := P(K_n = i)$ is governed by the recurrence relation

$$q_{n+1,i} = \frac{n}{n + \theta} q_{n,i} + \frac{\theta}{\theta + n} q_{n,i-1}.$$

The solution of this recurrence equation is expressible as

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

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

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

See Aoki (2002, p.208) for example.

Markov Chains

We can construct Markov chains using the transition probabilities of (1) and (2). Some special cases of these equations for the case $\alpha = 0$ have been simulated by Aoki (2002, Sec. 8.6). We give some details later in this paper as **Example 2**. More extensive examples are to be found in the forthcoming book by Aoki and Yoshikawa (2006).

Asymptotic Behavior of Cluster Sizes

We collect some known facts about cluster sizes as $n \to \infty$ in this section. The number of clusters is given by

$$K_n = \sum_{j=1}^{\infty} I(\text{type j is in our sample of size n}).$$

Usinf the one-parameter distribution Ewens obtained

$$EK_n = \int_0^1 [1-(1-x)^n]\theta x^{-1}(1-x)^{\theta-1}dx = \sum_0^{n-1} \frac{\theta}{\theta+n-j},$$

which is about $1+\theta[\gamma + ln(n-1)]$ for a small positive value of θ, where γ is the Euler's constant, Aoki (2002, p.185). Conditional on $\{p_n\} \sim PD(\alpha,\theta)$, we have

$$E(K_n|\{p_n\}) = \sum_j [1-(1-p_j)^n],$$

from which

$$E(K_n) = E\sum_j \frac{1-(1-w)^n}{w} = \frac{\Gamma(\theta+1)}{\Gamma(1-\alpha)\Gamma(\theta+\alpha)} \int_0^1 \frac{1-(1-w)^n}{w} w^{-\alpha}(1-w)^{\theta+\alpha-1}dw.$$

Carrying out the integral successively for large n

$$EK_n \sim \frac{\Gamma(\theta+1)}{\alpha\Gamma(\theta+\alpha)n^{\alpha}},$$

by using Stirling's formula, Carlton (1999, p.69).

Pitman (2002, Sec. 3) has a stronger result

$$K_n/n^\alpha \to S_\alpha, a.s.,$$

where the expression S_α has the density

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

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

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

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

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

Denote by $A_j = a_j(n)$ the number of clusters of size j when there are n agents in the model. We note that $\sum_{j=1}^n j a_j(n) = n$, and $K_n := \sum_j^n a_j(n)$ is the total number of clusters formed by the total of n agents. Pitman (2002) shows also that

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

for every $j = 1, 2, \ldots$ a.s. as n goes to infinity. and that $a_j(n) \sim P_{\alpha,j} S_\alpha n^\alpha$ in a two-parameter Poisson-Dirichlet case, where S_α is a random variable with the Mittag-Leffler density, and

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

See also Blumenfeld and Mandelbrot (1997) who credit Feller (1949) as the original source. Calculating its characteristic function we derive the p-th moment of g_α to be $\Gamma(p+1)/\Gamma(p\alpha+1)$. For example, $E_{\alpha,\theta}(S_\alpha) = \Gamma(\theta+1)/\Gamma(\theta+\alpha+1)$, and $E_{\alpha,\theta}(S_\alpha^2) = \Gamma(\theta+1)(\theta+\alpha)/\alpha^2\Gamma(\theta+\alpha)$.

Yamato and Sibuya (2000) has shown that the cluster sizes asymptotically approaches Sibuya distribution.

Mekjian and Chase (1997) connect their work to those by Ewens and its two-parameter extensio by Pitman.

Derrida sketched a derivation that the expected values of $Y_k = \sum_i x_i^k, k = 2, 3, \ldots$ 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.

Examples

Example 1 Instead of treating all possible configurations equi-probably, we weight 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 $\sum_j ja_j(n) = n$.

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

$$\{\frac{a_1(n)}{n^\alpha}, \frac{a_2(n)}{n^\alpha}, \dots, \frac{a_n(n)}{n^\alpha}\} \to^d S_\alpha\{\alpha, (1-\alpha)^{[1]}/2!,$$

$$\cdots, (1-\alpha)^{[n-1]}/n!\}.$$

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 $Z^{-\alpha} = \Gamma(1-\alpha)S_\alpha$.

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 deman signals they receive. Sectors interpret positive excess demans 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 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 sclae 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.[6] 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

[6]Their focus is on financial sector, not real sector.

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. 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 that $e^{1/q}$, 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 $K >> e^4 > 55$. 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

$$p(r) = Cexp(-a|r - r_m|),$$

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 (2006) for detail.

Local Limit Theorem

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

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

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 (2002) shows that the number of clusters with n agents, K_n with $0 < \alpha < 1$, is such that K_n/n^α converges a.s to S_α which is distributed with Mittag-Leffler desnity

$$g_{\alpha,\theta}(s) = \frac{\Gamma(\theta+1)}{\Gamma(\theta/\alpha+1)} s^{\theta/\alpha} g_\alpha(s), s > 0.$$

His 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 Ys are non-self averaging applies to this case as well.

Estimating the Parameters

Carlton is 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. Here the Mittag-Leffler random variable S makes its appearance again.

Lemma

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

$$\psi(1 + \hat{\theta}_n/\alpha) - \alpha\psi(1 + \hat{\theta}_n) \to logS, as.$$

Here ψ is the digamma function.

With θ known, and α unknown, Carlton proves *Lemma*

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

$$\hat{\alpha}_n = \frac{logK_n}{logn} \to \alpha \; a.s.$$

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

Non-Self Averaging

In the physics literature, a random variable X_n, where n indicates the size of a model is said to be non-self averaging when its variance does not vanish as n goes to infinity. This means that variability or fluctuations of samples persist even in the so-called thermodynamic limit of n going to infinity. Derrida has shown that models with power law where x are all positive and has density which behaves as

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

with $0 < \mu < 1$ for large x, the variance does not go to zero as n approaches infinity, i.e., non-self averaging effects exist. With the Herfindahl index, using the notion of spectrum density with $\alpha = 0$, moments can be easily calculated. This notion is also found in Aldous (1985) for random maps, which has $\theta = 1/2$. Given the spectral density $f(x) = \theta x^{-1}(1-x)^{\theta-1}$, we obtain

$$E(Y_1) = \frac{1}{1+\theta},$$

$$E(Y_1^2) = \frac{6}{(1+\theta)(2+\theta)(3+\theta)},$$

and

$$Var(Y_1) = \frac{2\theta}{(1+\theta)^2(2+\theta)(3+\theta)} > 0.$$

Since we can calculate the moments of the Mittag-Leffler density we can calculate the variance. The result is that as n goes to infinity

$$lim_n Var(\frac{K_n}{n^\alpha}) = \frac{\Gamma(\theta+1)}{\alpha^2}[\frac{\theta+\alpha}{\Gamma(\theta+2\alpha)} - \frac{\Gamma(\theta+1)}{\Gamma(\theta+\alpha)^2(1+\theta/\alpha)^2}].$$

The variance of K_n/n^α does not go to zero as n goes to infinity. We return to this point in the last section of this paper.

Derrida's development in our notation is $\alpha = -n$ which is between 0 and 1. His random numbers z_k is distributed as $Be(\alpha, k\alpha+\alpha)$ or using the parameter μ in the power law density expression , $Be(1-\mu, k\mu+\alpha)$. Derrida's construction of weights W_α (in his notation) generates GEM distribution.

Economic interpretation

In economics, the notion of thermodynamic limits may not be too appropriate, even though in growth context, long-run in the time scale may corresponds to the number of agents going to infinity.

We propose, instead, that we define or intepret this term as the existence of long-run disequilibrium distributions, and consequently, non-vanishing of variances of economic variables such as long-run profits, effects of innovations or immitations and the like, that is, existence of their non-degenerate long-run distributions. In this section we sketch two models as examples. The first is a Schumperterian dynamics of a model with innovation and imitation, Aoki, Nakano, and Yoshida (2004). The second example is the model by Iwai (2001), which is also a Schumperterian dynamics of long-run profits.

Example 4: Long-run effects of innovation and imitation This is based on a two-sector model discussed in Aoki (2002, Sec. 7.4). There are two types of firms, innovators and imitators. The state vector is (n_1, n_2) where the size of the technically advanced sector is n_1, and n_2 is the size of the sector of imitators. We specify transition rates for growth or entry as

$$w\{(n_1, n_2), (n_1+1, n_2)\} = c_1 n_1 + f,$$

and

$$w\{(n_1, n_2), (n_1, n_2+1)\} = c_2 n_2,$$

where f is exogenous innovation that hits sector 1. Firms exit or die at the rate

$$w\{(n_1, n_2), (n_1 - 1, n_2)\} = d_1 n_1,$$

and

$$w\{(n_1, n_2), (n_1, n_2 - 1))\} = d_2 n_2.$$

The next two transition rates specify how firms change their types, that is, an imitating firm may succeed in upgrading the technical level to the level of innovator, or an innovator may drop down in its technical level to that of type 2 firms

$$w\{(n_1, n_2, (n_1 + 1, n_2 - 1)\} = \mu g_1 g_2 (n_1 + h),$$

and

$$w\{(n_1, n_2, (n_1 - 1, n_2 + 1)\} = \mu g_2 n_1 n_2,$$

where $g_2 = c_2/d_2$, and $h = f/c_1$, and $\mu = \lambda g_2 d - 1 d_2$.

With these transition rates, we write the master equation. We compute the probability generating function, and then convert it into the cumulant generating function, since we are interested in calculating only the first and second order moments, $k_1, k_2, k_{1,1}, k_{1,2}$, and $k_{2,2}$. Fortunately, this model is specified in such a way that the equations for the moment are closed at the second moments, that is no higher order moments appear in the equations for the first and second moments. We derive a coupled ordinary differential equations for these moments. With the help of Mathematica we calculate the stationary state values of these moments for varioous parameter values, and verify the positive definiteness of the second moment matrix.

To the knowledge of the author this is the first example of Schumpert-erian dynamics with innovations and immitation effects for which the first two moments have been analytically derived and numerically evaluated. The model allows us to examine parametrically the relative importance of net death rate and innovation rate, and draw conclusions about qualitative behavior of interacting two sectors. We show also that the means of stationary locally stable equilibria scale with parameters of the innovation rate, and death rate.

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 passsage 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 willl 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) 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) for a simple example, or Krpisvsky et al (2000). This paper is another attempt at exposing non-self-averaging phenomena in economics, in particular in problems involving combinatorial optimization. We also have mentioned 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 we cannot blindly try for larger size samples in the hope that we obtain better estimates.

The example above is 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. It would be an interesting application of the random partition theory and see if non-self-averaging phenomena exist in the sense of physics literture in this area.

References

Aldous, D.J., (1985), "Exchangeability and related topics" in Lecture notes in mathemaatics, No.1117, Springer-Verlarg, 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.

Aoki, M., T. Nakano, and G. Yoshida, (2004), "Two sector Schumpeterian model of Industry" Mimeo, Dept. Physics, Chuo University, Tokyo.

—–, H. Yoshikawa (2006), *Reconstructing Macroeconomics: A Perspective from Statistical Physics and Combinatorial Stochastic Processes* , forthcoming from Cambridge University Press, New York.

—– , and—-, (2006),"Stock prices and real economy: Exponentila 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.

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.

——, (1994) " 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.

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.

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

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

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

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

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

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

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

Yamoato 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.