Cross-Sectional Distributions and Power Law with Incomplete Markets

Alexis Akira Toda*

November 24, 2014

Abstract

This note explains how to study cross-sectional distributions in incomplete-market dynamic general equilibrium models.

1 Introduction

Consider the Saito (1998) model explained in the previous lecture note. For simplicity, assume that there is no stock market and no aggregate shock (only private equity with idiosyncratic shock). Then the equilibrium risk-free rate is

$$r = \mu - \gamma\sigma^2,$$

where μ is the expected return on private equity, $\gamma > 0$ is the relative risk aversion, and $\sigma > 0$ is the idiosyncratic volatility. The optimal consumption rate is then

$$m = \beta\varepsilon + (1-\varepsilon)\left(\mu - \frac{\gamma\sigma^2}{2}\right),$$

where ε is the elasticity of intertemporal substitution. Substituting into the budget constraint, we get

$$\mathrm{d}w_{it}/w_{it} = g\mathrm{d}t + v\mathrm{d}B_{it},$$

where $g = \mu - \beta\varepsilon - (1-\varepsilon)\left(\mu - \frac{\gamma\sigma^2}{2}\right)$ and $v = \sigma$. Therefore, in the Saito (1998) model, individual wealth satisfies Gibrat (1931)'s law of proportionate growth.

We can characterize the cross-sectional wealth distribution as follows. Applying Itô's formula to $f(x) = \log x$, we get

$$\begin{aligned}
\mathrm{d}(\log w_{it}) &= \frac{1}{w_{it}}\mathrm{d}w_{it} + \frac{1}{2}\left(-\frac{1}{w_{it}^2}\right)(\mathrm{d}w_{it})^2 \\
&= (g\mathrm{d}t + v\mathrm{d}B_{it}) - \frac{1}{2}\frac{1}{w_{it}^2}(vw_{it})^2\mathrm{d}t \\
&= \left(g - \frac{1}{2}v^2\right)\mathrm{d}t + v\mathrm{d}B_{it}.
\end{aligned}$$

*atoda@ucsd.edu

Therefore log wealth follows the Brownian motion with drift $g-\frac{1}{2}v^2$ and volatility v. Assuming every agent starts with initial wealth w_0, the cross-sectional wealth distribution at time t is log normal, with

$$\log w_{it} \sim N\left(\log w_0 + (g - v^2/2)t, v^2 t\right).$$

Thus both the log mean and variance increases linearly over time. In this model, consumption is proportional to wealth, so the same holds for consumption. This model is consistent with the empirical findings of Deaton and Paxson (1994).

2 Fokker-Planck equation

In the above example, we were lucky because we were able to show that the log wealth follows the Brownian motion with drift, and therefore the calculation of the cross-sectional distribution was straightforward. In general, we cannot use such tricks. The *Fokker-Planck equation*, also known as the *Kolmogorov forward equation*, lets us calculate the cross-sectional distribution in general settings.

2.1 Derivation of Fokker-Planck equation

Consider the diffusion

$$\mathrm{d}X_t = g(t, X_t)\mathrm{d}t + v(t, X_t)\mathrm{d}B_t, \tag{1}$$

where B_t is standard Brownian motion. Let $p(x,t)$ be the density of $X(t)$ at time t. To derive a partial differential equation that $p(x,t)$ satisfies, we do the following (unintuitive) calculation.

First, fix $t_1 < t_2$ and let $F(t,x)$ be a smooth function such that $F(t_1,x) = F(t_2,x) = 0$ and $F(t,x), F_x(t,x) \to 0$ as $x \to \pm\infty$. There are plenty of such functions, for example

$$F(t,x) = (t-t_1)(t-t_2)f(x)$$

with $f(x) > 0$ and $f(x), f'(x) \to 0$ as $x \to \pm\infty$.

By Itô's formula, we get

$$\begin{aligned}
\mathrm{d}F(t, X(t)) &= F_t\mathrm{d}t + F_x\mathrm{d}X_t + \frac{1}{2}F_{xx}(\mathrm{d}X_t)^2 \\
&= F_t\mathrm{d}t + F_x(g\mathrm{d}t + v\mathrm{d}B) + \frac{1}{2}F_{xx}v^2\mathrm{d}t \\
&= \left(F_t + F_x g + \frac{1}{2}F_{xx}v^2\right)\mathrm{d}t + F_x v\mathrm{d}B.
\end{aligned}$$

Taking expectations and using the martingale property of the Brownian motion, we get

$$\begin{aligned}
\mathrm{E}[\mathrm{d}F(t, X(t))] &= \mathrm{E}\left[\left(F_t + F_x g + \frac{1}{2}F_{xx}v^2\right)\mathrm{d}t\right] \\
&= \int_{-\infty}^{\infty} \left(F_t + F_x g + \frac{1}{2}F_{xx}v^2\right) p(x,t)\mathrm{d}t\mathrm{d}x.
\end{aligned}$$

Integrating from $t = t_1$ to t_2 and using $F(t_1, x) = F(t_2, x) = 0$, we get

$$\begin{aligned}
0 &= \mathrm{E}[F(t_2, X(t_2)) - F(t_1, X(t_1))] \\
&= \int_{-\infty}^{\infty} \int_{t_1}^{t_2} \left(F_t + F_x g + \frac{1}{2} F_{xx} v^2 \right) p(x,t) \mathrm{d}t \mathrm{d}x =: I_1 + I_2 + I_3.
\end{aligned}$$

Integrating by parts, we get

$$\begin{aligned}
I_1 &= \int_{-\infty}^{\infty} \int_{t_1}^{t_2} \frac{\partial F}{\partial t} p(x,t) \mathrm{d}t \mathrm{d}x \\
&= \int_{-\infty}^{\infty} \left(F(t_2, x) - F(t_1, x) - \int_{t_1}^{t_2} F \frac{\partial}{\partial t} p(x,t) \mathrm{d}t \right) \mathrm{d}x \\
&= - \int_{t_1}^{t_2} \int_{-\infty}^{\infty} F \frac{\partial}{\partial t} p(x,t) \mathrm{d}x \mathrm{d}t,
\end{aligned}$$

where I have used $F(t_1, x) = F(t_2, x) = 0$ and Fubini's theorem. By similar calculations, we get

$$\begin{aligned}
I_2 &= - \int_{t_1}^{t_2} \int_{-\infty}^{\infty} F \frac{\partial}{\partial x} (g p(x,t)) \mathrm{d}x \mathrm{d}t, \\
I_3 &= \int_{t_1}^{t_2} \int_{-\infty}^{\infty} F \frac{\partial^2}{\partial x^2} \left(\frac{1}{2} v^2 p(x,t) \right) \mathrm{d}x \mathrm{d}t.
\end{aligned}$$

Putting all the pieces together, we get

$$0 = I_1 + I_2 + I_3 = \int_{t_1}^{t_2} \int_{-\infty}^{\infty} F \left[-\frac{\partial p}{\partial t} - \frac{\partial}{\partial x}(gp) + \frac{\partial^2}{\partial x^2} \left(\frac{1}{2} v^2 p \right) \right] \mathrm{d}x \mathrm{d}t.$$

Since F is (nearly) arbitrary, the integrand must be zero.[1] Therefore we obtain the (parabolic) partial differential equation (PDE)

$$\frac{\partial p}{\partial t} = -\frac{\partial}{\partial x}(gp) + \frac{1}{2} \frac{\partial^2}{\partial x^2}(v^2 p), \tag{2}$$

which is known as the *Fokker-Planck (Kolmogorov forward) equation.*

The Fokker-Planck equation (2) holds if the diffusion (1) holds at all times. However, we can consider situations in which the process is occasionally reset. For example, if $X(t)$ in (1) describe individual wealth, since the individual will die eventually, we need to specify what happens when an individual dies. If there is influx $j_+(x,t)$ and outflux $j_-(x,t)$ per unit of time at location x at time t, then the Fokker-Planck equation (2) must be modified as

$$\frac{\partial p}{\partial t} = -\frac{\partial}{\partial x}(gp) + \frac{1}{2} \frac{\partial^2}{\partial x^2}(v^2 p) + j_+ - j_-.$$

For example, if the unit dies at constant probability d per unit of time (Poisson rate d) and is reborn at location x_0, then the FPE becomes

$$\frac{\partial p}{\partial t} = -\frac{\partial}{\partial x}(gp) + \frac{1}{2} \frac{\partial^2}{\partial x^2}(v^2 p) + d\delta(x - x_0) - dp,$$

where $\delta(x - x_0)$ is the Dirac delta function located at x_0.

[1] To see this more rigorously, set

$$F = (t - t_1)(t - t_2) \left[-\frac{\partial p}{\partial t} - \frac{\partial}{\partial x}(gp) + \frac{\partial^2}{\partial x^2} \left(\frac{1}{2} v^2 p \right) \right].$$

2.2 Stationary density

If the diffusion has time-independent drift $g(x)$ and variance $v(x)$ and admits a stationary distribution $p(x)$, then we get

$$0 = -\frac{\mathrm{d}}{\mathrm{d}x}(gp) + \frac{1}{2}\frac{\mathrm{d}^2}{\mathrm{d}x^2}(v^2 p).$$

Integrating with respect to x and using the boundary condition $p(x), p'(x) \to 0$ as $x \to \pm\infty$, we get

$$0 = -g(x)p(x) + \frac{1}{2}(v(x)^2 p(x))'.$$

Letting $q(x) = v(x)^2 p(x)$ and solving the ODE, we get

$$\begin{aligned} q' = \frac{2g}{v^2}q \iff \frac{q'}{q} &= \frac{2g}{v^2} \\ \iff \log q(x) &= \int \frac{q'(x)}{q(x)}\mathrm{d}x = \int \frac{2g(x)}{v(x)^2}\mathrm{d}x \\ \iff q(x) &= \exp\left(\int \frac{2g(x)}{v(x)^2}\mathrm{d}x\right). \end{aligned}$$

Therefore the stationary density is

$$p(x) = \frac{q(x)}{v(x)^2} = \frac{1}{v(x)^2}\exp\left(\int \frac{2g(x)}{v(x)^2}\mathrm{d}x\right), \tag{3}$$

where the constant of integration is determined by the condition $\int_{-\infty}^{\infty} p(x)\mathrm{d}x = 1$ since $p(x)$ is a density.

If there is a constant probability of death d, the stationary density is the solution of the second-order ODE

$$0 = -\frac{\mathrm{d}}{\mathrm{d}x}(gp) + \frac{1}{2}\frac{\mathrm{d}^2}{\mathrm{d}x^2}(v^2 p) - dp,$$

which holds at every point except x_0.

A natural question is, does $p(x,t)$ always converge to the stationary density? The answer is yes, under some assumptions. For a proof, see pp. 61–62 of Gardiner (2009).

3 Examples

Let us compute the stationary density for typical diffusion processes.

Example 1 (Ornstein-Uhlenbeck process)**.** $X(t)$ is called the Ornstein-Uhlenbeck process if it satisfies the stochastic differential equation

$$\mathrm{d}X(t) = -\kappa(X(t) - \mu)\mathrm{d}t + v\mathrm{d}B_t,$$

where $\kappa > 0$. This process is the continuous-time analogue of the AR(1) process in discrete time. Using the formula (3) with $g(x) = -\kappa(x - \mu)$ and $v(x) = v$, the stationary density is proportional to

$$\exp\left(\int -\frac{2\kappa(x-\mu)}{v^2}\mathrm{d}x\right) = \exp\left(-\frac{\kappa(x-\mu)^2}{v^2}\right),$$

so the stationary distribution is $N(\mu, \frac{v^2}{2\kappa})$.

Example 2. Consider the stochastic differential equation

$$\mathrm{d}X(t) = -\kappa\,\mathrm{sgn}(X(t)-\mu)\mathrm{d}t + v\mathrm{d}B_t,$$

where $\mathrm{sgn}(x)$ is the sign function, *i.e.*, $\mathrm{sgn}(x) = -1, 0, 1$ according as $x < 0$, $x = 0$, and $x > 0$. This process has been applied in Alfarano et al. (2012) and Toda (2012). The stationary density is proportional to

$$\exp\left(\int -\frac{2\kappa\,\mathrm{sgn}(x-\mu)}{v^2}\mathrm{d}x\right) = \exp\left(-\frac{2\kappa\,|x-\mu|}{v^2}\right),$$

which is the (symmetric) Laplace distribution, explained in the Appendix. Toda (2012) explicitly solves the Fokker-Planck equation.

Example 3 (Brownian motion evaluated at exponential time)**.** Consider the diffusion

$$\mathrm{d}X(t) = g\mathrm{d}t + v\mathrm{d}B_t,$$

which is just the Brownian motion with constant drift g and volatility v. Assume that units die at Poisson rate $\delta > 0$ and are reborn at x_0. The FPE at steady state is

$$0 = -gp' + \frac{1}{2}v^2p'' - \delta p$$

except at x_0, where I used the fact that g, v are constant. Since this is a linear second-order ODE with constant coefficients, the general solution is

$$p(x) = C_1\mathrm{e}^{\lambda_1 x} + C_2\mathrm{e}^{\lambda_2 x},$$

where λ_1, λ_2 are solutions to the quadratic equation

$$\frac{1}{2}v^2\xi^2 - g\xi - \delta = 0. \tag{4}$$

After some algebra, we can show that the stationary density is asymmetric Laplace with model x_0 and exponents α, β, where $-\alpha, \beta > 0$ are solutions to the above quadratic equation.

Example 4 (geometric Brorwnian motion with drift)**.** Consider the diffusion

$$\mathrm{d}X(t) = (gX(t) + q)\mathrm{d}t + vX(t)\mathrm{d}B_t,$$

which is the geometric Brownian motion with drift $q > 0$. The FPE at steady state is

$$0 = -\frac{\mathrm{d}}{\mathrm{d}x}[(gx+q)p(x)] + \frac{1}{2}\frac{\mathrm{d}^2}{\mathrm{d}x^2}[v^2x^2p(x)].$$

Solving this second-order ODE with boundary conditions $p(x), p'(x) \to 0$ as $x \to \infty$ and $\int p(x) = 1$, we can show that the stationary density is the inverse gamma distribution[2] with density

$$p(x) = \frac{\beta^\alpha}{\Gamma(\alpha)}x^{-\alpha-1}\mathrm{e}^{-\frac{\beta}{x}},$$

where $\alpha = 1 - 2g/v^2$ and $\beta = 2q/v^2 > 0$.

[2] If X is gamma distributed, then the distribution of $1/X$ is inverse gamma.

The economic interpretation of this example is as follows. If agents have infinite lives, earn a constant labor income, and invest wealth into private equity (which is subject to idiosyncratic multiplicative shocks), then the stationary wealth distribution is inverse gamma. Furthermore, since $\beta/x \to 0$ as $x \to \infty$, we have

$$P(X > x) \sim \int_x^\infty y^{-\alpha-1}\mathrm{d}y \sim x^{-\alpha}$$

as $x \to \infty$, so the cross-sectional wealth distribution obeys the *power law* with exponent $\alpha = 1 - 2g/v^2$. Benhabib et al. (2011) study a discrete-time version of this model with optimizing agents.

4 Power law and the Laplace distribution

In this section I introduce the notion of the double power law and present parametric distributions that satisfy it.

4.1 Power law

A nonnegative random variable X obeys the *power law* (in the upper tail) with exponent $\alpha > 0$ if

$$\lim_{x\to\infty} x^\alpha P(X > x) > 0$$

exists (Pareto, 1896, 1897; Mandelbrot, 1960, 1961).[3] Recently, many economic variables have been shown to obey the power law also in the lower tail, meaning that

$$\lim_{x\to 0} x^{-\beta} P(X < x) > 0$$

exists for some exponent $\beta > 0$. A random variable obeys the *double power law* if the power law holds in both the upper and the lower tails. The double power law has been documented in city size (Reed, 2002; Giesen et al., 2010), income (Reed and Wu, 2008; Toda, 2011, 2012), and consumption and consumption growth (Toda and Walsh, 2014). If X obeys the double power law with exponents (α, β), then X^η obeys the double power law with exponents $(\alpha/\eta, \beta/\eta)$ if $\eta > 0$ and $(-\beta/\eta, -\alpha/\eta)$ if $\eta < 0$. To see this, for example if $\eta > 0$ we have

$$P(X^\eta > x) = P(X > x^{\frac{1}{\eta}}) \sim x^{-\frac{\alpha}{\eta}}$$

as $x \to \infty$, and other cases are similar. An important implication of this fact is that the η-th moment $\mathrm{E}[X^\eta]$ exists if and only if $-\beta < \eta < \alpha$. Since many econometric techniques rely on the existence of some moments, recognizing a power law is important.[4]

[3] See Gabaix (1999, 2009) for reviews of mechanisms generating the power law.

[4] For instance, Kocherlakota (1997) tests the representative agent, consumption-based capital asset pricing model (CAPM) by considering the possibility that the time series of aggregate consumption growth may have fat tails. Toda and Walsh (2014) documents that cross-sectional distributions of consumption and its growth rate have fat tails and the standard GMM estimation of heterogeneous-agent CAPM models is susceptible to type II errors.

4.2 Double Pareto and double Pareto-lognormal distributions

A canonical distribution that obeys the double power law is the *double Pareto* distribution (Reed, 2001), which has the probability density function (PDF)

$$f_{\mathrm{dP}}(x) = \begin{cases} \frac{\alpha\beta}{\alpha+\beta} M^\alpha x^{-\alpha-1}, & (x \geq M) \\ \frac{\alpha\beta}{\alpha+\beta} M^{-\beta} x^{\beta-1}, & (0 \leq x < M) \end{cases} \tag{5}$$

where $M > 0$ is a scale parameter (the mode if $\beta > 1$), and $\alpha, \beta > 0$ are shape parameters (power law exponents). The classical Pareto distribution with minimum size M is a special case of the double Pareto distribution by letting $\beta \to \infty$ in (5).

The density of the double Pareto distribution has a cusp at M. An example of a distribution with a smooth density that obeys the double power law is the *double Pareto-lognormal* distribution (Reed, 2003), abbreviated as dPlN. A dPlN random variable is defined as the product of independent double Pareto and lognormal random variables. Its density is

$$\begin{aligned} f_{\mathrm{dPlN}}(x) = \frac{\alpha\beta}{\alpha+\beta} \Bigg[& \mathrm{e}^{\frac{\alpha^2\sigma^2}{2}+\alpha\mu} \Phi\left(\frac{\log x - \mu}{\sigma} - \alpha\sigma\right) x^{-\alpha-1} \\ & + \mathrm{e}^{\frac{\beta^2\sigma^2}{2}-\beta\mu} \Phi\left(-\frac{\log x - \mu}{\sigma} - \beta\sigma\right) x^{\beta-1} \Bigg], \end{aligned}$$

where α, β are the power law exponents of the double Pareto variable (with $M = 1$), μ, σ are the mean and the standard deviation of the logarithm of the lognormal variable, and $\Phi(\cdot)$ denotes the cumulative distribution function (CDF) of the standard normal distribution. As is clear from the above density, the double Pareto-lognormal distribution obeys the double power law with exponents α, β.

The double Pareto distribution and the lognormal distribution are special cases of the double Pareto-lognormal distribution by letting $\sigma \to 0$ and $\alpha = \beta \to \infty$, respectively. This is an important point because it means the lognormal distribution, which is nested within dPlN, can be tested against dPlN by the likelihood ratio test.

4.3 Laplace and normal-Laplace distributions

Instead of working with double Pareto and dPlN random variables, it is often more convenient to work with their logarithms. The logarithm of a double Pareto random variable is called *Laplace*.[5] By changing variables in (5) and setting $m = \log M$, the density of the Laplace distribution is given by

$$f_{\mathrm{L}}(x) = \begin{cases} \frac{\alpha\beta}{\alpha+\beta} \mathrm{e}^{-\alpha|x-m|}, & (x \geq m) \\ \frac{\alpha\beta}{\alpha+\beta} \mathrm{e}^{-\beta|x-m|}, & (x < m) \end{cases} \tag{6}$$

[5] Hence the Laplace and the double Pareto distributions have the same relation as the normal and the lognormal distributions. As an interesting historical remark, Laplace discovered the Laplace distribution in 1774, which predates by a quarter of century the discovery of the normal distribution by himself and Gauss in the early 1800s. For more historical background, see Kotz et al. (2001).

where m is the mode and $\alpha, \beta > 0$ are scale parameters. A Laplace distribution is said to be *asymmetric* if $\alpha \neq \beta$. A comprehensive review of the Laplace distribution can be found in Kotz et al. (2001).

The logarithm of a double Pareto-lognormal variable is called *normal-Laplace* (Reed and Jorgensen, 2004), which is simply the convolution of independent normal and Laplace random variables. The normal-Laplace distribution has four parameters, a location parameter μ and three scale parameters $\sigma, \alpha, \beta > 0$, with probability density function

$$\begin{aligned} &f_{\mathrm{NL}}(x) \\ &= \frac{\alpha\beta}{\alpha+\beta}\left[\mathrm{e}^{\frac{\alpha^2\sigma^2}{2}-\alpha(x-\mu)}\Phi\left(\tfrac{x-\mu}{\sigma}-\alpha\sigma\right)+\mathrm{e}^{\frac{\beta^2\sigma^2}{2}+\beta(x-\mu)}\Phi\left(-\tfrac{x-\mu}{\sigma}-\beta\sigma\right)\right] \end{aligned} \tag{7}$$

and cumulative distribution function

$$\begin{aligned} F_{\mathrm{NL}}(x) = \Phi\left(\frac{x-\mu}{\sigma}\right) - \frac{1}{\alpha+\beta}\Bigg[&\beta \mathrm{e}^{\frac{\alpha^2\sigma^2}{2}-\alpha(x-\mu)}\Phi\left(\tfrac{x-\mu}{\sigma}-\alpha\sigma\right) \\ &-\alpha \mathrm{e}^{\frac{\beta^2\sigma^2}{2}+\beta(x-\mu)}\Phi\left(-\tfrac{x-\mu}{\sigma}-\beta\sigma\right)\Bigg]. \end{aligned} \tag{8}$$

Again the Laplace and the normal distributions are special cases of the normal-Laplace distribution by letting $\sigma \to 0$ and $\alpha = \beta \to \infty$, respectively, and therefore can be tested against the normal-Laplace distribution by the likelihood ratio test.

Using (6), the characteristic function of a Laplace random variable X is

$$\begin{aligned} \phi_X(t) &= \int_{-\infty}^{m} \mathrm{e}^{itx}\frac{\alpha\beta}{\alpha+\beta}\mathrm{e}^{-\beta|x-m|}\mathrm{d}x + \int_{m}^{\infty} \mathrm{e}^{itx}\frac{\alpha\beta}{\alpha+\beta}\mathrm{e}^{-\alpha|x-m|}\mathrm{d}x \\ &= \frac{\mathrm{e}^{imt}}{1-i(\frac{1}{\alpha}-\frac{1}{\beta})t+\frac{t^2}{\alpha\beta}}, \end{aligned} \tag{9}$$

from which we obtain the mean $m+\frac{1}{\alpha}-\frac{1}{\beta}$ and the variance $\frac{1}{\alpha^2}+\frac{1}{\beta^2}$. It is often useful to parameterize the Laplace distribution in terms of its characteristic function. Letting $a = \frac{1}{\alpha}-\frac{1}{\beta}$ be an asymmetry parameter and $\sigma = \sqrt{\frac{2}{\alpha\beta}}$ be a scale parameter in (9), we write $X \sim \mathcal{AL}(m, a, \sigma)$ if

$$\phi_X(t) = \frac{\mathrm{e}^{imt}}{1-iat+\frac{\sigma^2t^2}{2}}. \tag{10}$$

The mean, mode, and variance of $\mathcal{AL}(m, a, \sigma)$ are $m+a$, m, and $a^2+\sigma^2$, respectively. Comparing (9) and (10), we obtain $1/\alpha - 1/\beta = a$ and $\alpha\beta = 2/\sigma^2$, so $-\alpha$ and β are the solutions to the quadratic equation

$$\frac{\sigma^2}{2}\zeta^2 - a\zeta - 1 = 0. \tag{11}$$

Perhaps the most important property of the Laplace distribution is that it is the only limit distribution of geometric sums. Theorem 1 below (which generalizes the i.i.d. case reviewed in Kotz et al. (2001)) shows that it is a robust property that the limit of a geometric sum is a Laplace distribution.

Theorem 1. *Let $\{X_n\}_{n=1}^\infty$ be a sequence of zero mean random variables such that the central limit theorem holds, $N^{-1/2}\sum_{n=1}^N X_n \xrightarrow{\mathrm{d}} N(0,\sigma^2)$; $\{a_n\}_{n=1}^\infty$ be a sequence such that $N^{-1}\sum_{n=1}^N a_n \to a$; and ν_p be a geometric random variable with mean $1/p$ independent from $\{X_n\}_{n=1}^\infty$. Then as $p \to 0$ we have*

$$p^{\frac{1}{2}}\sum_{n=1}^{\nu_p}(X_n + p^{\frac{1}{2}}a_n) \xrightarrow{\mathrm{d}} \mathcal{AL}(0,a,\sigma).$$

Proof. For $w>0$ and $0<p<1$, define $N_p(w) = \lceil -w/\log(1-p)\rceil$, the integer obtained by rounding up $-w/\log(1-p)>0$. If W is standard exponential, then $\nu_p = N_p(W)$ (in distribution). In fact,

$$\begin{aligned}&\Pr[N_p(W)=n] = \Pr[n-1 < -W/\log(1-p) \le n]\\ &= \int_{-(n-1)\log(1-p)}^{-n\log(1-p)} \mathrm{e}^{-w}\mathrm{d}w = (1-p)^{n-1}-(1-p)^n = (1-p)^{n-1}p.\end{aligned}$$

Since $-\log(1-p)\approx p$ for small p, it follows that $N_p(w)\to\infty$ and $pN_p(w)\to w$ as $p\to 0$. Conditioning on $W=w$, since by assumption the central limit theorem holds, as $p\to 0$ we obtain

$$p^{\frac{1}{2}}\sum_{n=1}^{N_p(w)}(X_n+p^{\frac{1}{2}}a_n) = \frac{\sqrt{pN_p(w)}}{\sqrt{N_p(w)}}\sum_{n=1}^{N_p(w)}X_n + \frac{pN_p(w)}{N_p(w)}\sum_{n=1}^{N_p(w)}a_n \xrightarrow{\mathrm{d}} \sqrt{w}\sigma Z + wa,$$

where Z is a standard normal variable. Therefore

$$p^{\frac{1}{2}}\sum_{n=1}^{\nu_p}(X_n+p^{\frac{1}{2}}a_n) = p^{\frac{1}{2}}\sum_{n=1}^{N_p(W)}(X_n+p^{\frac{1}{2}}a_n) \xrightarrow{\mathrm{d}} \sqrt{W}\sigma Z + aW,$$

where W is standard exponential that is independent of Z. The claim follows since its characteristic function is

$$\begin{aligned}&\mathrm{E}\left[\exp(it(\sqrt{W}\sigma Z + aW))\right] = \mathrm{E}\left[\mathrm{E}\left[\exp(it(\sqrt{W}\sigma Z + aW))\middle| W\right]\right]\\ &= \int_0^\infty \mathrm{e}^{iatw-\frac{1}{2}\sigma^2t^2w}\mathrm{e}^{-w}\mathrm{d}w = \frac{1}{1-iat+\frac{\sigma^2t^2}{2}},\end{aligned}$$

which is the characteristic function of $\mathcal{AL}(0,a,\sigma)$ in (10). □

4.4 Literature

Gabaix (2009) is a nice literature review of the power law. The definition of the double Pareto distribution and its generative mechanism (geometric Brownian motion evaluated at exponential time) are due to Reed (2001). The double Pareto distribution has been obtained in a general equilibrium model in Benhabib et al. (2014) and Toda (2014). Theorem 1 is due to Toda (2014).

Exercises

1. Following the hints below, show that the stationary distribution of the Brownian motion evaluated at exponential time (Example 3) is asymmetric Laplace. Let $-\alpha,\beta>0$ be the solution to the quadratic equation (4).

1. Using the general solution and the fact that $p(x) \to 0$ as $x \to \infty$, show that $p(x) = C_+ \mathrm{e}^{-\alpha x}$ for some $C_+ > 0$ when $x > x_0$.

2. Using the general solution and the fact that $p(x) \to 0$ as $x \to -\infty$, show that $p(x) = C_- \mathrm{e}^{\beta x}$ for some $C_- > 0$ when $x < x_0$.

3. Using the continuity of $p(x)$ at $x = x_0$, show that $p(x)$ is asymmetric Laplace with model x_0 and exponents α, β.

2. Fill in the details of the derivation of the stationary density in Example 4 (inverse gamma).

References

Simone Alfarano, Mishael Milaković, Albrecht Irle, and Jonas Kauschke. A statistical equilibrium model of competitive firms. *Journal of Economic Dynamics and Control*, 36(1):136–149, January 2012. doi:10.1016/j.jedc.2011.07.002.

Jess Benhabib, Alberto Bisin, and Shenghao Zhu. The distribution of wealth and fiscal policy in economies with finitely lived agents. *Econometrica*, 79(1): 123–157, January 2011. doi:10.3982/ECTA8416.

Jess Benhabib, Alberto Bisin, and Shenghao Zhu. The distribution of wealth in the Blanchard-Yaari model. *Macroeconomic Dynamics*, 2014. doi:10.1017/S1365100514000066.

Angus Deaton and Christina Paxson. Intertemporal choice and inequality. *Journal of Political Economy*, 102(3):437–467, June 1994.

Xavier Gabaix. Zipf's law for cities: An explanation. *Quarterly Journal of Economics*, 114(3):739–767, 1999. doi:10.1162/003355399556133.

Xavier Gabaix. Power laws in economics and finance. *Annual Review of Economics*, 1:255–293, 2009. doi:10.1146/annurev.economics.050708.142940.

Crispin Gardiner. *Stochastic Methods: A Handbook for the Natural and Social Sciences*. Springer Series in Synergetics. Springer, fourth edition, 2009.

Robert Gibrat. *Les Inégalités économiques*. Librairie du Recueil Sirey, Paris, 1931.

Kristian Giesen, Arndt Zimmermann, and Jens Suedekum. The size distribution across all cities—double Pareto lognormal strikes. *Journal of Urban Economics*, 68(2):129–137, September 2010. doi:10.1016/j.jue.2010.03.007.

Narayana R. Kocherlakota. Testing the consumption CAPM with heavy-tailed pricing errors. *Macroeconomic Dynamics*, 1(3):551–567, September 1997. doi:10.1017/S136510059700401X.

Samuel Kotz, Tomasz J. Kozubowski, and Krzysztof Podgórski. *The Laplace Distribution and Generalizations*. Birkhäuser, Boston, 2001.

Benoît Mandelbrot. The Pareto-Lévy law and the distribution of income. *International Economic Review*, 1(2):79–106, May 1960. doi:10.2307/2525289.

Benoît Mandelbrot. Stable Paretian random functions and the multiplicative variation of income. *Econometrica*, 29(4):517–543, October 1961. doi:10.2307/1911802.

Vilfredo Pareto. *Cours d'économie politique*. F. Rouge, Lausanne, 1896.

Vilfredo Pareto. New theories of economics. *Journal of Political Economy*, 5 (4):485–502, September 1897.

William J. Reed. The Pareto, Zipf and other power laws. *Economics Letters*, 74(1):15–19, December 2001. doi:10.1016/S0165-1765(01)00524-9.

William J. Reed. On the rank-size distribution for human settlements. *Journal of Regional Science*, 42(1):1–17, February 2002. doi:10.1111/1467-9787.00247.

William J. Reed. The Pareto law of incomes—an explanation and an extension. *Physica A*, 319(1):469–486, March 2003. doi:10.1016/S0378-4371(02)01507-8.

William J. Reed and Murray Jorgensen. The double Pareto-lognormal distribution—a new parametric model for size distribution. *Communications in Statistics—Theory and Methods*, 33(8):1733–1753, 2004. doi:10.1081/STA-120037438.

William J. Reed and Fan Wu. New four- and five-parameter models for income distributions. In Duangkamon Chotikapanich, editor, *Modeling Income Distributions and Lorenz Curves*, chapter 11, pages 211–223. Springer, New York, 2008. doi:10.1007/978-0-387-72796-7_11.

Makoto Saito. A simple model of incomplete insurance: The case of permanent shocks. *Journal of Economic Dynamics and Control*, 22(5):763–777, May 1998. doi:10.1016/S0165-1889(97)00077-8.

Alexis Akira Toda. Income dynamics with a stationary double Pareto distribution. *Physical Review E*, 83(4):046122, 2011. doi:10.1103/PhysRevE.83.046122.

Alexis Akira Toda. The double power law in income distribution: Explanations and evidence. *Journal of Economic Behavior and Organization*, 84(1):364–381, September 2012. doi:10.1016/j.jebo.2012.04.012.

Alexis Akira Toda. Incomplete market dynamics and cross-sectional distributions. *Journal of Economic Theory*, 154:310–348, November 2014. doi:10.1016/j.jet.2014.09.015.

Alexis Akira Toda and Kieran Walsh. The double power law in consumption and implications for testing Euler equations. *Journal of Political Economy*, 2014. URL `http://ssrn.com/abstract=2319454`. Forthcoming.