RESEARCH PAPERS OF WROCŁAW UNIVERSITY OF ECONOMICS
No. 66 — 2009
Towards Information-based Welfare Society

Stanisława Ostasiewicz
Wrocław University of Economics, Wrocław, Poland

DIVERSITY AND HETEROGENEITY OF POPULATION

Abstract

The paper addresses some problems concerning the diversification of populations.

The problem of economic diversification, particularly with respect to income inequality, is reviewed very briefly. The main purpose, however, is to present the recently developed theory of heterogeneity of population with respect to the mortality. The core of this theory has been initiated in 1979, the exposition of it is illustrated by a series of simple examples.

1. Introductory remarks about the unequal world

Diversity concerns both income and health. To see the income gap let us notice that the ratio of income of richest to poorest in 1960 [8] was as 30 : 1, in 1991 this ratio was as 61 : 1, and in 1997 the income ratio of richest to poorest was as 74 : 1.

In 1960 the poorest 20% have only 1% of world's GDP, while the middle 60% have 13% of world's GDP, and the richest 20% have 86% of world's GDP.

According to the HDR published in 2005 the world's richest 50 individuals have a combined income greater than that of the poorest 416 000 000.

Polarization is increasing dramatically. Recently, of the world's 6 billion people, 2.8 billion live on less than $ 2 a day, 1.5 billion live on less than $ 1 a day. And on the other hand, according to American Obesity Association, 127 000 000 people are overweighed, 60 000 000 people suffer from overweight and 9 000 000 Americans are dramatically obese.

2. Measurement of inequality

One distinguishes two classes of measures of financial inequality: mechanistic and statistical. Among the mechanistic measures (see [2]) the well-known is the relative mean deviation (Schutz's coefficient). This measure is defined as follows:

$$D = \frac{\frac{1}{n}\sum_{i=1}^{n}\left|x_i - \mu\right|}{2\mu} \quad , \tag{1}$$

where x_1, x_2,…, x_3 denote annual income of n persons and μ is an average annual income. The other measure of this kind is the famous Gini index which is the most commonly used measure of inequality. This index is usually defined in terms of the Lorenz curve. The following equivalent definition makes it clear that the Gini index is a measure of dispersion divided by twice the mean. The numerator of this expression known as Gini coefficient is the average absolute difference between all pairs of individuals

$$G = \frac{\frac{1}{n^2}\sum_{i=1}^{n}\sum_{j=1}^{n}\left|x_i - x_j\right|}{2\mu} \quad . \tag{2}$$

As the third measure it should be mentioned the Teils' measure

$$T = \frac{1}{n}\sum_{i=1}^{n}\left(\frac{x_i}{\mu}\right)\log\left(\frac{x_i}{\mu}\right) \quad . \tag{3}$$

The most often natural logarithms are used.

General family of inequality measures that includes, among others, both the Gini index and the coefficient of variation, is given by the following formula:

$$I = \frac{\left(\frac{1}{2n^2}\sum_{i=1}^{n}\sum_{j=1}^{n}\left|x_i - x_j\right|^r\right)^{\frac{1}{r}}}{\mu} \quad . \tag{4}$$

When $r = 1$, then I is the Gini index.

Statistical measure of inequality is defined as a mapping (see [3])

$$K\colon D_{[0,\,\alpha]} \rightarrow R \ ,$$

where $D_{[0,\,\alpha]}$ denotes the set of all distribution functions satisfying

$$F_X(y) = \begin{cases} 0 & \text{for } y < 0, \\ 1 & \text{for } y \geq a \quad (a \in R_{++}) \end{cases} \qquad (\mu_F > 0) \quad . \tag{5}$$

The statistical measures most commonly used in theoretical as well as in empirical work include the coefficients listed below.

1. Coefficient of variation

$$C(F_X) = \frac{V(F_X)}{\mu_F} \ , \tag{6}$$

where $V(F_X)$ is a variation random variable X, and μ_F is a mean random variable X.

2. Relative mean deviation

$$R(F_X) = \int_0^a \left| \frac{x}{\mu_F} - 1 \right| dF_X(x)\,. \tag{7}$$

3. Gini coefficient

$$G(F_X) = \frac{1}{\mu_F} \int_0^a (F_X - F_X^2)\,dx\,. \tag{8}$$

4. Area of concentration

$$A(F_X) = \frac{1}{2} G(F_X)\,, \tag{9}$$

where $G(F_x)$ is Gini coefficient.

5. Standard deviation of logarithm

$$S(F_X) = \int_0^a \left[\log\left(\frac{x}{\mu_F} \right) \right]^2 . \tag{10}$$

6. Atkinson index

Atkinson's measure of inequality is based on the social welfare function (see [2; 3]). The problem is that the choice of a social welfare function is normative, not empirical, and it is difficult to achieve any agreement on what that function should be. Perhaps the most widely acceptable class of function is the additive concave welfare function. For this class, Atkinson showed that the following inequality measure was especially appropriate:

$$A(F_X) = 1 - \frac{1}{\mu_F} \left[\int_0^a x^{\varepsilon} F_X(dx) \right]^{\frac{1}{\varepsilon}} . \tag{11}$$

Atkinson's index in case of discrete finite random variable is defined in the following way:

$$A(F_X) = 1 - \left[\sum_{i=1}^{N} \left(\frac{x_i}{\mu_F} \right)^{\varepsilon} f_X(x_i) \right]^{\frac{1}{\varepsilon}} , \tag{12}$$

where $f_X(x_i)$ is the density function random variable X in the point x_i.

Yitzhaki index is a generalization of an absolute parametric Gini index of equality.

It is defined as follows:

$$\delta_F(\upsilon) = \int_0^b A^{\upsilon}(y)dy, \qquad \upsilon \geq 0 \quad , \tag{13}$$

where $A(x) = 1 - F(x)$ and $F(y)$ represent cumulative income distribution and υ is a parameter ranging from 0 to ∞.

3. Longevity diversification

Benjamin Gompertz proposed in 1825 that the force of mortality increased exponentially with age for humans (see [7]). It has been discovered that Gompertz's formula describes the mortality at younger adult ages. This formula does not capture the rise of mortality for advanced ages. There is an evidence that the mortality decelerated at highest ages. For the explanation of this phenomena there were proposed several theories.

In 1990 Medvedev reviewed more than 300 theories of aging. Mortality deceleration came as a big surprise (see [5]). All populations are heterogeneous. Some individuals are frailer than others, and they tend to die first. Let T be a positive random variable representing the life span of an individual drawn at random from G population of newborn. Random variable T is conveniently characterized by a survival function:

$$S(t)=P(T>t)+1-F(t), \quad t \geq 0,$$

where

$$F(t)=P(T \leq t).$$

Force of mortality is defined as follows

$$\mu(t) = \frac{f(t)}{S(t)}, \tag{14}$$

where $f(t)$ is the probability density function of T.

Force of mortality is related to survival function:

$$\mu(t) = -\frac{S'(t)}{S(t)} = -\frac{d}{dt}\big(1 - F(t)\big).$$

In the case of Gompertz law functions characterizing life span are the following:

$$f(t) = \exp\{-\frac{\beta}{\gamma}\exp((\gamma t) - 1)\exp(\gamma t)), \tag{15}$$

$$\mu(t) = \beta c^t, \tag{16}$$

$$S(t) = e^x g^{c^x - 1}, \tag{17}$$

where $g = \exp(-\beta/\ln c)$.
Suppose that there is some non-observable characteristic Z which affects the mortality multiplicatively (see [8]):

$$\mu(t,z) = z\mu(t,1) = z\mu(t). \tag{18}$$

From this one has

$$\mu(t) = \int_0^\infty \mu(t,z)dF(z). \tag{19}$$

Assume that Z can be treated as a random variable with a cumulative density function $F(z)$. The probability that an individual at the age t given the value z of variable Z will die in an interval from t to $t+dt$ is denoted as $f(t,z)dt$. Conditional density function of mortality will be denoted as $f(t/z)$ The probability density function $f(t)$ can be now considered as a following mixture:

$$f(t) = \int f(t,z)dF(z)\ . \tag{20}$$

It is convenient to observe that

$$f(t)=E(f(t,Z)).$$

Let $f_x(z/T \geq x)$ denote the frailty distribution in a population survived till x, then the mean of Z is calculated as (see [5; 10]):

$$\bar{z}(x) = \int_0^\infty z \cdot f_x(z/T \geq x)dz. \tag{21}$$

Hence

$$\overline{\mu}(x) = \int_0^\infty \mu(x,z)\cdot f_x(z)dz = \int_0^\infty z\cdot\mu(x,1)\cdot f_x(z)dz = \mu(x,1)\int_0^\infty z\cdot f_x(z)dz = \mu(x,1)\cdot\bar{z}(x). \tag{22}$$

Let $f_0(z)$denote distribution at the birth, then for any age x we have:

$$f_x(z)=f_0(z)\cdot p(x,z),$$

where $p(x,z)$ is the probability survive till the age x. This probability is calculated according to the formula (see [9; 10]):

$$p(x,z) = \exp(-\int_0^x \mu(t,z)dt = \exp\left(-\int_0^x z\mu(t,1)\right)dt = \exp\left(-z\int_0^x \mu(t,1)\right)dt = \exp(-zH(x)). \tag{23}$$

Hence

$$f_x(z) = \frac{f_0(z)\cdot\exp(-zH(x))}{\int_0^\infty f_0(z)\cdot\exp(-zH(x))dz}. \tag{24}$$

The mean frailty can be now expressed as follows:

$$\bar{z}(x) = \int_0^\infty z \cdot f_x(z)dz = \frac{\int_0^\infty z \cdot f_0(z)\exp(-zH(x))dz}{\int_0^\infty f_0(z)\exp(-zH(x))dz}. \tag{25}$$

For the illustrative purpose suppose that mortality law is described by de Moivre's formula:

$$f(t) = \frac{1}{\omega},$$

where $\omega = 100$ and means upper limit of human life time.

Cumulative distribution is following:

$$F(t) = \int_0^t f(t)dt = \int_0^t \frac{1}{\omega}dt = \frac{t}{\omega}.$$

Survival function takes the following form:

$$S(t) = 1 - F(t) = \frac{\omega - t}{\omega} \quad .$$

Using the function of intensity of mortality $\mu(t)$ and the cumulative intensity $H(t)$, this function can be expressed as follows:

$$S(t) = \exp(-\int_0^t \mu(x)dx) = \exp(-H(t)) = \exp\left(-\ln\frac{\omega}{(\omega - t)}\right) = \exp\left(\frac{1}{\ln\frac{\omega}{(\omega - t)}}\right) = \frac{1}{\frac{\omega}{\omega - t}} = \frac{\omega - t}{\omega},$$

where

$$\mu(t) = \frac{f(t)}{S(t)} = \frac{\frac{1}{\omega}}{\frac{\omega - t}{\omega}} = \frac{1}{\omega - t}, \qquad 0 \le t < 100,$$

and

$$H(t) = \int_0^t \mu(x)dx = \int_0^t \frac{1}{\omega - x}dx = -\ln(\omega - x)\Big|_0^t = -\ln(\omega - t) + \ln\omega = \ln\frac{\omega}{\omega - t}, \quad 0 \le t < 100.$$

Furthermore, assume that mixing distribution is uniform on the interval (0, 3):

$$g(z) = \frac{1}{3} \quad \text{for} \quad 0 \le z \le 3,$$

then

$$f(t)=\int_0^\infty f(t/z)dG(z)=\int_0^\infty zf(t/1)dG(z)=\int_0^3 \frac{1}{3}zf(t/1)dz=\int_0^3 \frac{1}{3}z\frac{1}{100}dz=\frac{3}{200}.$$

The mean intensity of mortality is obtained as follows:

$$\bar{\mu}(t)=\int_0^3 \mu(t,z)g(z)dz=\int_0^3 z\mu(t/z=1)g(z)dz=\int_0^3 \frac{1}{3}z\frac{1}{100-t}dz=\frac{1{,}5}{100-t}.$$

The distribution of the frailty in the population for any age x has the following form:

$$f_x(z)=\frac{f_0(z)\cdot\exp(-zH(x))}{\int_0^\infty f_0(z)\cdot\exp(-zH(x))dz}=\frac{\frac{1}{3}\exp(-z\ln A(x))}{\int_0^3 \frac{1}{3\ln A(x)}\exp(-z\ln A(x))dz}=\frac{\ln A(x)\exp(-z\ln A(x))}{1-\exp(-3\ln A(x)},$$

where:

$$A(x)=\frac{\omega}{\omega-x},$$

and $f_0(x)$ means the density of frailty at the moment of birth.

The mean frailty, in this case, is following:

$$\bar{z}(x)=\int_0^\infty zf_x(z)dz=\int_0^3 z\cdot\frac{\ln A(x)\exp(-z\ln A(x))}{1-\exp(-3\ln A(x)}=\frac{1}{B(x)^2}-\frac{1}{B(x)}\left[\exp(-3B(x))\left(\frac{1}{B(x)}-3\right)\right],$$

where

$$B(x)=\ln\frac{\omega}{\omega-x}.$$

In the case when mixing function has gamma distribution:

$$g(z)=\eta^{\eta}z^{\eta-1}e^{-\eta z}/\Gamma(z),$$

then

$$\bar{\mu}(t)=\frac{\mu(t)}{1+\sigma^2 H(t)},$$

where σ^2 is variation of mixing function and

$$H(t)=\int_0^t \frac{1}{\omega-x}dx=\ln\frac{\omega}{\omega-t},$$

$$\overline{\mu}(t)=\frac{\frac{1}{\omega-t}}{1+\sigma^{2}H(t)}=\frac{\frac{1}{\omega-t}}{1-\sigma^{2}\ln\frac{\omega}{\omega-t}}=\frac{1}{(\omega-\mathrm{t})(1-\sigma^{2})\ln\frac{\omega}{\omega-t}}\,,\qquad \text{for}\quad 0\le t<100.$$

Let us take a look at a few intensity functions and calculate the average intensity using gamma distribution function.

If $\mu(t)=ct,\ H(t)=\int_0^t cx\,dx=\frac{ct^2}{2}$, then the average intensity function is following:

$$\overline{\mu}(t)=\frac{ct}{1+c\sigma^{2}t^{2}/2}.$$

If $\mu(t)=ce^{bt}$, $H(t)=\frac{1}{b}[e^{a}-1]$

and mixing function is gamma, then the average intensity function has the following form:

$$\overline{\mu}(t)=\frac{ce^{(bt)}}{1+\sigma^{2}c[\exp(bt)-1]/b}.$$

If $\mu(t)=c\ (\exp(-ba))$, $H(t)=\frac{1}{b}[1-e^{-ba}]$

and mixing function is gamma, then

$$\overline{\mu}(t)=\frac{ce^{-bt}}{1+\frac{\sigma^{2}}{b}c[1-\exp(-bt)]}.$$

These examples demonstrate that individual aging can be different from aging in the entire cohort.

References

[1] Allison P.D., "Measures of population diversity", *American Sociological Review* 1978, vol. 43, pp. 865-880.
[2] Bürk R., Gehring W., "Indices of income inequality and societal income. An axiomatic approach", [in:] W. Eichhorn et. al. (Eds)., *Theory and Applications of Economic Indices*, Physica-Verlag, Würzburg 1978, pp. 309-356.
[3] Eichhorn W., Gehring W., "Measurement of inequality in economics", [in:] B. Korte (Ed.), *Modern Applied Mathematics*, North-Holland 1982.
[4] Gakidon E.E., Murray C.J.L., Frank J., "Measuring health inequality: An approach based on the distribution of health expectancy", *Bulletin of the World Health Organization* 2000, vol. 78, no. 1, pp. 42-54.
[5] Hougaard P., "Life table methods for heterogeneous populations: Distributions describing the heterogeneity", *Biometrica* 1984, vol. 71, no. 1, pp. 75-83.
[6] Liberson S., "Measuring population diversity", *American Sociological Review* 1969, pp. 951-862.
[7] Manton K.G., Stallard E., Vaupel J.Z., "Methods for comparing the mortality experience of heterogeneous populations", *Demography* 1981, vol. 18, no. 3, pp. 389-410.
[8] *World Health Report 2000. Health systems: Improving performance*, WHO, Geneva 2000.
[9] Vaupel J.Z., Manton K.G., Stallard E., "The impact of heterogeneity in individual frailty on the dynamics of mortality", *Demography* 1979, vol. 16, no. 3, pp. 439-454.
[10] Vaupel J.Z., Yashin A.I., "The deviant dynamics of death in heterogeneous populations", *Sociological Methodology* 1985, vol. 15, pp. 179-211.

DYWERSYFIKACJA I ZRÓŻNICOWANIE POPULACJI

Streszczenie

Artykuł dotyczy niektórych problemów zróżnicowania populacji. Bardzo skrótowo przedstawiony został problem dywersyfikacji ekonomicznej ze względu na nierówności dochodów. Głównym celem artykułu jest prezentacja zróżnicowania populacji ze względu na indywidualną witalność lub słabowitość. Istota teorii, która została zainicjowana w 1979 r., została zilustrowana prostymi przykładami.