The Aleatory Distribution of a Human Aptitude or a Talent in a Finite Population

Edmund H. Mantell
Professor of Finance and Economics
Lubin School of Business
Pace University
New York, N.Y.

Abstract

There are numerous economic and sociological and theories addressing the question of how specific aptitudes and talents come to be apportioned among individuals in a population. This paper considers how an arbitrary aptitude or a talent will be distributed among the individuals in a population of a fixed size of it is distributed purely by chance. The main result demonstrates that even if an aptitude or a talent is not objectively measurable, if it is apportioned randomly among the individual in a finite population the specific distribution function can be derived. If the aptitude or talent is empirically measurable, the distribution function derived in this paper can be tested for goodness of fit.

Key words: aptitude, random distribution, human population, empirical tests.

1. Introduction

One of the questions that interests social scientists and humanists is the question of how human aptitudes or talents are distributed in a finite population. It is obvious that individual aptitudes and talents for activities like performing or composing music, an aptitude for mathematics, creating art, balletic dancing, etc. are distributed unevenly throughout the population of a society.

This paper demonstrates that if an aptitude is distributed purely by chance, it is possible to derive a specific function governing the distribution of that aptitude in a finite population.

The analysis of the statistical distribution of income among individuals in a population has a long history in the field of economics.[1] Some of the methods that have been developed in that field are applied in this paper. The theory developed here is amenable to many different kinds of human attributes distributed among the individuals in a finite population.

2. The Theoretical Model

At the threshold it is necessary to define what I mean by an aptitude or a talent in the context of this paper. An aptitude is an individual’s natural or acquired capacity to accomplish specific objectives that are considered desirable by himself and other members of the population. Thus defined, a specific aptitude is obviously apportioned among individuals in the population heterogeneously.

Moreover, the kind of aptitude considered in this paper is additive among the individuals in the population. Additivity means that it is plausible to describe the total quantity of the aptitude that is distributed in the population. It is not necessary that the aptitude should be objectively measurable, although many are. However, it is meaningful to describe it as an attribute of individuals apportioned differently among the individuals in a population. For example, it is reasonable to describe the distribution of mathematical aptitude among a large population of children of the same age.

[1] Among a huge list of published studies on the statistical properties of the distribution of income, some modern compendia are references

The question addressed by this paper is: How will the total quantity of a specific aptitude be distributed among the individuals in a finite population if it is distributed purely by chance.

Consider a finite population of N individuals. Each individual can be characterized by his or her aptitude for the specific activity. The aptitude is symbolized by x. For example, x may represent the aptitude of an individual for multilingualism. The aptitude is distributed heterogeneously among the members of the population. I assume the aptitude is bounded from below: $0 \leq x < \infty$

Suppose the population is partitioned with respect to x . Let symbol N_i represent the number of individuals who are homogeneous with respect to their aptitude x. That quantity of is x symbolized by x_i, where $i = 1, 2, \ldots$ Efthimiou and Wearne (2016, p. 3) define N_i as the *occupation number*. It is obvious that the sum of the occupation numbers must be equal to the size of the population:

$$\sum_i N_i = N \tag{1}$$

Suppose the distribution function governing the number of individuals in each occupation number with an aptitude between x and $x + dx$ is symbolized by $A(x)$. The analytical task is to establish the functional form of that distribution if the apportionment of x among the occupation numbers is due to pure chance.

The total number of ways all the individuals in the population can be partitioned into the occupation numbers $\{N_1, N_2, \ldots.\}$ can be calculated as:[2]

$$M = \frac{N!}{N_1!\, N_2! \ldots\ldots} \tag{2}$$

If the aptitude is distributed among the occupation numbers randomly, the largest value of M is the value that is most likely to occur by chance. Thus, the mathematical problem devolves into finding the distribution corresponding to the maximum value of M. This can be accomplished by transforming M to its logarithmic function and finding the maximization of that function.

Taking the logarithm of equation (2) we have:

$$\ln M = \ln N! - \sum_i lnN_i! \tag{3}$$

Appendix A demonstrates the application of Stirling's approximation formula for the factorials in equation (3). That application results in equation (4):

$$\ln M = N\, lnN - \sum_i N_i\, ln\, N_i \tag{4}$$

Define the subset of weights as:

$$w_i = \frac{N_i}{N} \qquad (i = 1, 2, \ldots) \tag{5}$$

Obviously,

$$\sum_i w_i = 1 \tag{6}$$

The set of weights is symbolized by $W \equiv \{w_1, w_2, \ldots\}$ The elements of W represent the relative frequencies of the aptitude corresponding to each occupation number.

Appendix B demonstrates that equation (4) can be expressed in terms of the subset of relative frequencies:

$$\ln M = -N \sum_i w_i \ln w_i \tag{7}$$

As remarked above, the distribution function $A(x)$ is the function generating the maximum value of $\ln M$. Appendix C carries out the maximization of $\ln M$ by finding its total differential, setting it to zero and then solving the resulting equation. Appendix C demonstrates that $A(x)$ can be written as a continuous distribution:

[2] Expression 2 is the coefficient of the multinomial probability distribution. See Feller (1968, p. 168).

$$A(x) = N\lambda e^{-\theta x} \tag{8}$$

where θ is the scale parameter of the distribution. The value of the coefficient λ is determined by the equation:

$$\lambda \int_0^\infty e^{-\theta x}\, dx = 1 \tag{9}$$

Solving equation (9) gives $\lambda = \theta$. Substituting this value into (8) we have:

$$A(x) = N\theta e^{-\theta x} \tag{10}$$

The average value $\bar{x}$ for the distribution is given by:

$$\bar{x} = \frac{1}{N} \int_0^\infty x\, A(x)\, dx \tag{11}$$

Substituting (10) into (11) and carrying out the integration, the average value of $\bar{x}$ is found to be

$$\bar{x} = \frac{1}{\theta} \tag{12}$$

If the total quantity of aptitude x in the population is a constant, independent of population size. It is symbolized by C.

$$\sum_i x_i = C \tag{13}$$

Then the value of the average person's talent will decrease if the population increases because $\bar{x} = \frac{C}{N}$. This means that a fixed quantity of the aptitude is being distributed more thinly over a larger population.

The analytical findings can be expressed as a proposition:

> *Proposition: If a fixed quantity of an aptitude or a talent x is distributed by pure chance among individuals in a finite population of size N, the most probable distribution is a declining exponential distribution:*
>
> $$A(x) = N\theta e^{-\theta x}$$

3. Empirical Implications of the Propostion

The Proposition can be applied in econometric studies of the distributions of income and wealth across national boundaries. Figure 1 and Figure 2 are graphic displays of the empirical distributions of household income in the United States as well as the United Kingdom in recent years. The shapes of those empirical distributions suggest that the Proposition derived in this paper should provide statistical explanatory power.

However, the Proposition is not limited to macroeconomic studies of income and wealth. The large compensation packages given to chief executive officers of public corporations in recent years are major controversial compensation practices. See Cremers (2013) The Proposition can be applied to studies of the pay of the chief executive officers in various industries. Consider, for example, two theories how the compensation of chief executive offers in the private sector is determined: (1) benchmarking and (2) "pay for luck." Payment for being "lucky" is a colloquial description of compensation payments distributed by pure chance. The Proposition can be applied to conduct statistical tests as to whether the hugely inflated compensation packages of CEOs reflect demonstrable merit, or whether those chief executive officers are being paid for being lucky.

The Proposition can be applied to measure the distributions of congenital aptitude or talents. See, for example, the studies by Lee (1989) and Vinkhuyzen (2009.) Figure 3 displays a good fit of the Proposition to the statistical data measuring the distribution of mathematical aptitude among students in school.

APPENDIX A

If every subset of the partition of N includes a large number of individuals, the value of M can be approximated well by an application of Stirling's formula to estimate the values of the factorials in equation (3). The formula is:[3]

$$\ln n! = n \ln n - n + \mathcal{O}(\ln n) \qquad (A1)$$

Applying Stirling's formula to equation (3)

$$\ln M = N\, ln\, N - N - \sum_i (N_i\, lnN_i - N_i) \qquad (A2)$$

Equation (A2) can be decomposed to read:

$$\ln M = N\, ln\, N - N - \sum_i N_i\, lnN_i + \sum_i N_i \qquad (A3)$$

Substituting equation (1) into equation (A2) and cancelling terms we have equation (4).

APPENDIX B

The definition of the subset weights implies:

$$ln\, N_i = ln\, w_i + ln\, N \qquad (B1)$$

Substituting (B1) into equation (4) we have:

$$\ln M = N\, lnN - \sum_i N_i (ln\, w_i + ln\, N) \qquad (B2)$$

$$= N\, ln\, N - \sum_i N_i\, ln\, w_i - ln\, N \sum_i N_i$$

Substituting equation (1) into (B2) and cancelling terms we have:

$$\ln M = -\sum_i N_i \ln w_i \qquad (B3)$$

Substituting into (B3) the definition of the weights in expression (5), we have equation (7).

[3] See Wilks (p. 177)

APPENDIX C

Taking the total differential of equation (7) results in the following expression:

$$\delta(ln\ M) = \sum_i \ln w_i\, \delta w_i + \sum_i w_i \delta(ln\ w_i) \qquad (C1)$$

Suppose the average value of x is symbolized by $\bar{x}$. The differential of $ln\ M$ must satisfy the conditions:

$$\sum_i N_i x_i = N \sum_i w_i\, x_i = N\bar{x} = a\ constant \qquad (C2)$$

Constraints (C2) can be substituted into the total differential of $ln\ M$. It is then set to zero to find the set of values of the weights $\{w_1, w_2 \ldots.\}$ corresponding to its maximum value. This gives the equation

$$\delta(ln\ M) = \sum_i \ln w_i\, \delta w_i + \sum_i \delta w_i = 0 \qquad (C3)$$

The constraint on the sum of the weights in (6) implies:

$$\sum_i \delta w_i = 0 \qquad (C4)$$

Substituting (C4) into (C3), the general solution to the resulting differential equation can be expressed as:

$$\ln w_i = a\ constant - \theta x_i \qquad (C5)$$

where θ is a characteristic parameter. If the symbol λ is assigned as an arbitrary coefficient, a solution to the equation (C5) is:

$$w_i = \lambda e^{-\theta x_i} \qquad (C6)$$

Equation (C6) can be rewritten as (C7):

$$N_i = N\lambda e^{-\theta x_i} \qquad (C7)$$

The continuous functional form of (C7) appears as equation (8) in the text.

REFERENCES

Atkinson, A. B. and F. Bourguignon. Handbook of Income Distribution, Vol. 1, Elsevier, Amsterdam, The Netherlands, 2000.

________________. Handbook of Income Distribution, Vol. 2, Elsevier, Amsterdam, The Netherlands, 2015.

Cremers, K. J. M and Yaniv Grinstein. *Does the Market for CEO Talent Explain Controversial CEO Pay Practices?* Review of Finance, 18 (3), July 2013, pp 921-960 doi:10.1093/rof/rft024

Chakrabarti, B. K., A. Chakrabarti and A. Chatterjee (eds.) Econophysics and Sociophysics: Trends and Perspectives, Wiley-VCH, Weinheim, Germany, 2000

Efthimiou, C. J. and A. Wearne. Household Income Distribution in the USA. arXiv:1602.06234v1 [q-fin.GN] 19 Feb 2016

Feller, W. An Introduction to Probability Theory and Its Applications. Vol. 1 (3rd ed.) John Wiley & Sons, Inc. 1968.

Lee, Valerie E. and A. S. Brick. *A Multilevel Model of the Social Distribution of High School Achievement.* Sociology of Education, Vol. 62, No. 3 (July 1989) pp. 172-192, DOI: 10.2307/2112866

Vinkhuyzen, A.A.E., van der Sluis, S., Posthuma, D. et al. *The Heritability of Aptitude and Exceptional Talent Across Different Domains in Adolescents and Young Adults.* Behavioral Genetics Vol. 39, No. 4 : pp. 380 - 392. (July 2009) doi:10.1007/s10519-009-9260-5

Wilks, S. Mathematical Statistics. John Wiley & Sons. New York, 1962.

Figure 1

United States Bureau of the Census

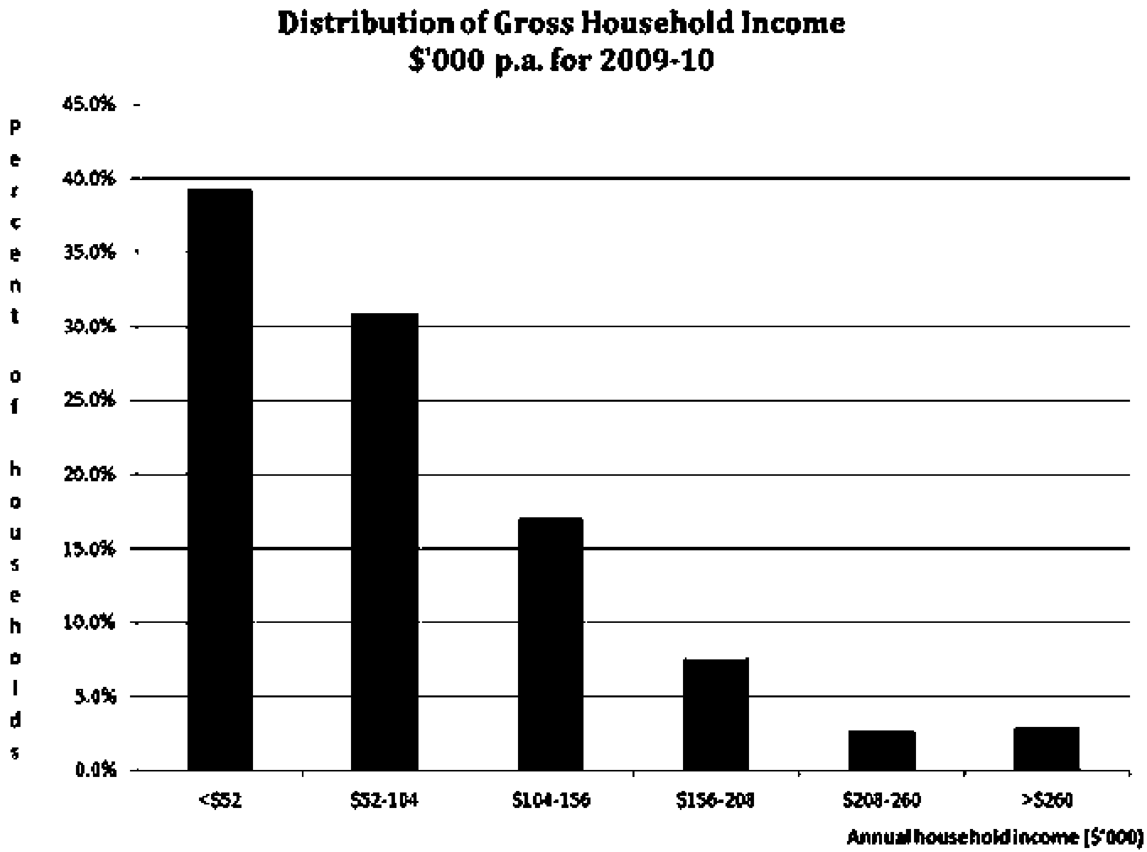

Figure 2

DISTRIBUTION OF INCOME IN THE UNITED KINGDOM

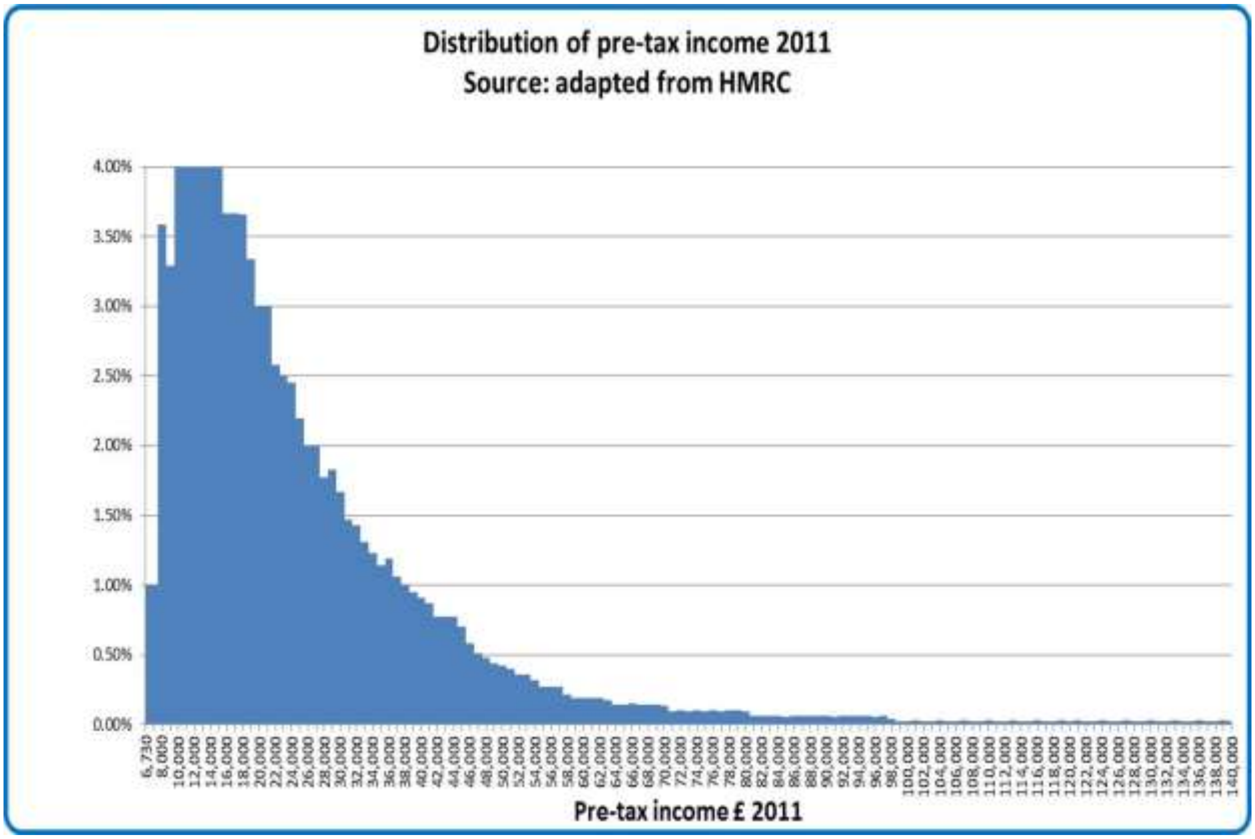

Figure 3

Comparison of the empirical distribution of mathematical aptitude x with the theoretical distribution of aptitude, A(x)