International Futures at the Pardee Center

INTERNATIONAL FUTURES HELP SYSTEM

Using Lognormal Income Distributions

LOGNORMAL DISTRIBUTION OF INCOME

A variable X is said to be lognormally distributed when its log has a normal distribution. To be lognormally distributed X always has to be positive. Let us assume that income x has a lognormal distribution, such that has a normal distribution with mean μx and standard deviation σx.

The probability density function (pdf) *fx(x)* of the lognormal distribution is given by

$$f_x(x) = \frac{1}{x\sigma_x\sqrt{2\pi}} e^{-\frac{1}{2}\left[\frac{\ln x - \mu_x}{\sigma_x}\right]^2} \qquad (1a)$$

Let us denote the lognormal distribution by Λ(μx, σx) and the normal distribution by *N*(μx, σx). The pdf of N is given by

$$f_y = \frac{1}{\sqrt{2\pi}} e^{-\frac{y^2}{2}} \qquad (1b)$$

DERIVATION OF THE PARAMETERS OF LOGNORMAL DISTRIBUTION FROM AVAILABLE DATA

Aitchison and Brown (1963: 8) note that, the mean, μ of a variable X (e.g., income or consumption), when X has a lognormal distribution, Λ(μx, σx) can be found from the following:

$$\mu = \exp\left[\mu_x + \frac{1}{2}\sigma_x^2\right] \qquad (2)$$

From the Theorem 2.7 of Aitchison and Brown (1963: 13) the Gini co-efficient, G for lognormal distribution can be derived as (see Chotikapanich, Valenzuela and Prasada Rao, 1997):

$$G = 2\Phi\left(\frac{\sigma_x}{\sqrt{2}}\right) - 1 \qquad (3)$$

where Φ is the standard normal distribution.

From the above equation, we can calculate one of the parameters of Λ,

$$\sigma_x = \sqrt{2}\Phi^{-1}\left(\frac{G+1}{2}\right) \qquad (4)$$

Given the mean income, m, we can use equations (4) and (2) to calculate the other parameter of the lognormal distribution:

$$\mu_x = \ln(\mu) - \frac{1}{2}\sigma_x^2 \qquad (5)$$

CALCULATING POPULATION AND INCOME SHARES

Once we find mean μx and standard deviation σx, we can construct the distribution equation and integrate it for any cut off of income.

The proportion of the population with incomes less than or equal to a given level x is given by the distribution function:

$$\pi(x) = \int_0^x f_x(x)dx \qquad (6)$$

The integral $\int_0^x f_x(x)dx$, is the lognormal cumulative distribution function (cdf) at x, i.e.:

$$\text{Population Fraction below income } x = \Lambda(x \mid \mu_x, \sigma_x) \qquad (7)$$

The corresponding income shares (at x) can be obtained from the following first moment distribution (Chotikapanich, Valenzuela and Prasada Rao, 1997; Aitchison and Brown, 1963):

$$\eta(x) = \frac{1}{\mu}\int_0^x x f_x(x)dx \qquad (8)$$

and according to the fundamental theorem of the moment distribution (Aitchinson and Brown 1963: 12), the first moment distribution with parameters (μx, σx) is the same as the lognormal distribution with parameters (μx+0.5σx ^{2}, σx) , i.e. income fraction held by people earning below income x,

$$\eta(x) = \Lambda(x \mid \mu_x + 0.5\sigma_x^2, \sigma_x) \quad (9)$$

POVERTY MEASURE: POVERTY HEADCOUNT

Replacing x with the poverty line income, z (e.g. $1 PPP a day, i.e., $365 PPP per year) in eqn (7), we obtain the percentage of people living below dollar a day, i.e., the headcount index, H/P where H is the number of poor and P the total population:

$$\text{Poverty Headcount Index, } H/P = \Lambda(z \mid \mu_x, \sigma_x) \quad (10)$$

POVERTY MEASURE: POVERTY GAP

Poverty Gap is obtained from the following generalized class of Foster-Gear-Thorbeck (FGT) poverty measures,

$$P_\alpha = \int_0^z \left[\frac{z-x}{z}\right]^\alpha f_z(x)dx \quad (11)$$

where, $\alpha \geq 0$, $fx(x)$ is the density function at income x and z is the income at the poverty line.

The above equation returns poverty headcount index for α=0. When α=1, we get the Poverty Gap, PG which can be interpreted as the shortfall from the poverty line or the depth of poverty below the line. The poverty gap, expressed as a percentage, can be further simplified to:

$$PG = \int_0^z \left[\frac{z-x}{z}\right] f_z(x)dx \quad (12)$$

$$PG = \int_0^z f_z(x)dx - \frac{1}{z}\int_0^z x f_z(x)dx \quad (13)$$

$$PG = \frac{H}{P} - \frac{\mu}{z}\eta(z) \quad (14)$$

where μ is the mean income (or consumption), using equation (8).

$$PG = \frac{H}{P} - \frac{\mu}{z}\Lambda\left(z \mid \mu_z + 0.5\sigma_z^2, \sigma_z\right) \quad (15)$$

using (9) from above.

RECONCILIATION BETWEEN NATIONAL ACCOUNTS AND SURVEY DATA

To reconcile the discrepancy between national accounts (NA) and household survey (HS) figures, International Futures converts its NA mean income (GDP per capita in PPP dollars) to an equivalent HS mean consumption. This is done by a reverse calculation of the mean consumption from the available data on Gini index and the population share with consumption below a dollar PPP a day, both calculated (at the source) by using the HS data.

We know (from the definition section on lognormal distribution above),

$$\Lambda(x \mid \mu_z, \sigma_z) = \Phi\left(\frac{\ln(x) - \mu_z}{\sigma_z}\right) \quad (16)$$

therefore,

$$\Lambda(365 \mid \mu_z, \sigma_z) = \Phi\left(\frac{\ln(365) - \mu_z}{\sigma_z}\right) \quad (17)$$

or, Population Fraction below income dollar PPP a day or n$365/year, H/P

$$= \Phi\left(\frac{\ln(365) - \mu_z}{\sigma_z}\right) \quad (18)$$

or, $$\Phi^{-1}\left(\frac{H}{P}\right) = \frac{\ln(365) - \mu_z}{\sigma_z} \quad (19)$$

or, $$\mu_z = \ln(365) - \Phi^{-1}\left(\frac{H}{P}\right) * \sigma_z \quad (20)$$

In the above equation $\Phi^{-1}\left(\frac{H}{P}\right), \sigma_z$ is available from World Bank where it is calculated using HS data on (mostly) consumption. We can calculate μx from this equation and obtain a (HS equivalent) mean consumption using equation (2) above.