Obtaining Measures from Subjective Probabilistic Expectations

By Paul Brimble

Measures of expectations can provide important insights into future economic behaviour, reflecting the forward-looking nature of many economic decisions. This post explains how to obtain measures, such as quantiles and moments of interest, derived from subjective probabilistic expectations. The core methodology behind this approach is to obtain subjective probabilities about certain events, and then fit these probabilities to a distribution, resulting in an expectations distribution (Dominitz and Manski (1997); Dominitz (1998); Manski (2001); Delavande et al. (2011a, b); McKenzie et al. (2013), Attanasio and Britta (2016)).

The simplest method of eliciting an expectation would be to ask for a single (non-probabilistic) point estimate such as: "what do you think". The main drawbacks of using this approach is that it is unclear what the individual is actually reporting (perhaps a measure of central tendency such as the mean or something else entirely) and it limits researchers from learning about other statistical concepts such as measures of dispersion and relevant quantiles. These can be addressed through fitting an expectations distribution based on both probabilistic expectations and distributional assumptions, from which many interesting statistical concepts can be obtained (see Figure 1). There are a wide range of concerns about eliciting subjective probabilities such as the design of the support (predetermined or self-provided), number of intervals, and the relevant distributional form, in addition to more context dependent issues. The purpose of this post is not to cover these topics, but rather to assist researchers with the techniques required to obtain the measures of interest after data collection.

The post is structured into four sections. First, I comment on subjective probabilities in more detail. Second, I provide an example of fitting a distribution to data. Third, I demonstrate how to obtain measures from the distribution. Finally, I briefly discuss possible extensions.

Subjective Probabilities

There are several ways to elicit subjective probabilities, but for continuous variables, the resulting data should contain a measure of probability between certain bounds (either open or closed). For example, consider a variable such as income where the researcher can ask an individual for a probabilistic statement on the likelihood that they will earn between two values. More generally, a researcher can define N intervals I_n for $n = \{1, \dots, N\}$, where each interval is defined using lower and upper thresholds T_{n-1} and T_n and T_0 and T_N are the minimum and maximum thresholds respectively. For a random variable X, the intervals and their respective subjective probabilities $\mathbf{P}_n$ can be written as:

$$\mathbf{P}_n\,(X \in I_n) = \mathbf{P}_n\,(T_{n-1} < X \leq T_n) \qquad \text{for } n = \{1, \dots, N\}.$$

Based on this information, researchers can already make several probabilistic statements. Consider in cases where the number of intervals is $N = 2$, then a researcher can state the probability that an individual expects to be below or above a certain threshold T_1. Combining these subjective probabilities together through summation yields cumulative probabilities $\mathbf{F}_n$:

$$\mathbf{F}_n\,(X \leq T_n) = \sum_{k=1}^{n} \mathbf{P}_k\,(X \in I_k) \qquad \text{for } n = \{1, \dots, N\}.$$

Using the N pairs of thresholds T_n and cumulative probabilities $\mathbf{F}_n$, a researcher can fit the individual-level observations to a distribution.

Distributional Assumptions and Estimation

Assumptions of model to estimate model parameters. Example of log-normal distribution.

piecewise uniform, triangle distributions or log-normal distributions. The number of distributional parameters will inform which distributions can be fit, depending on the number of intervals from data collection.

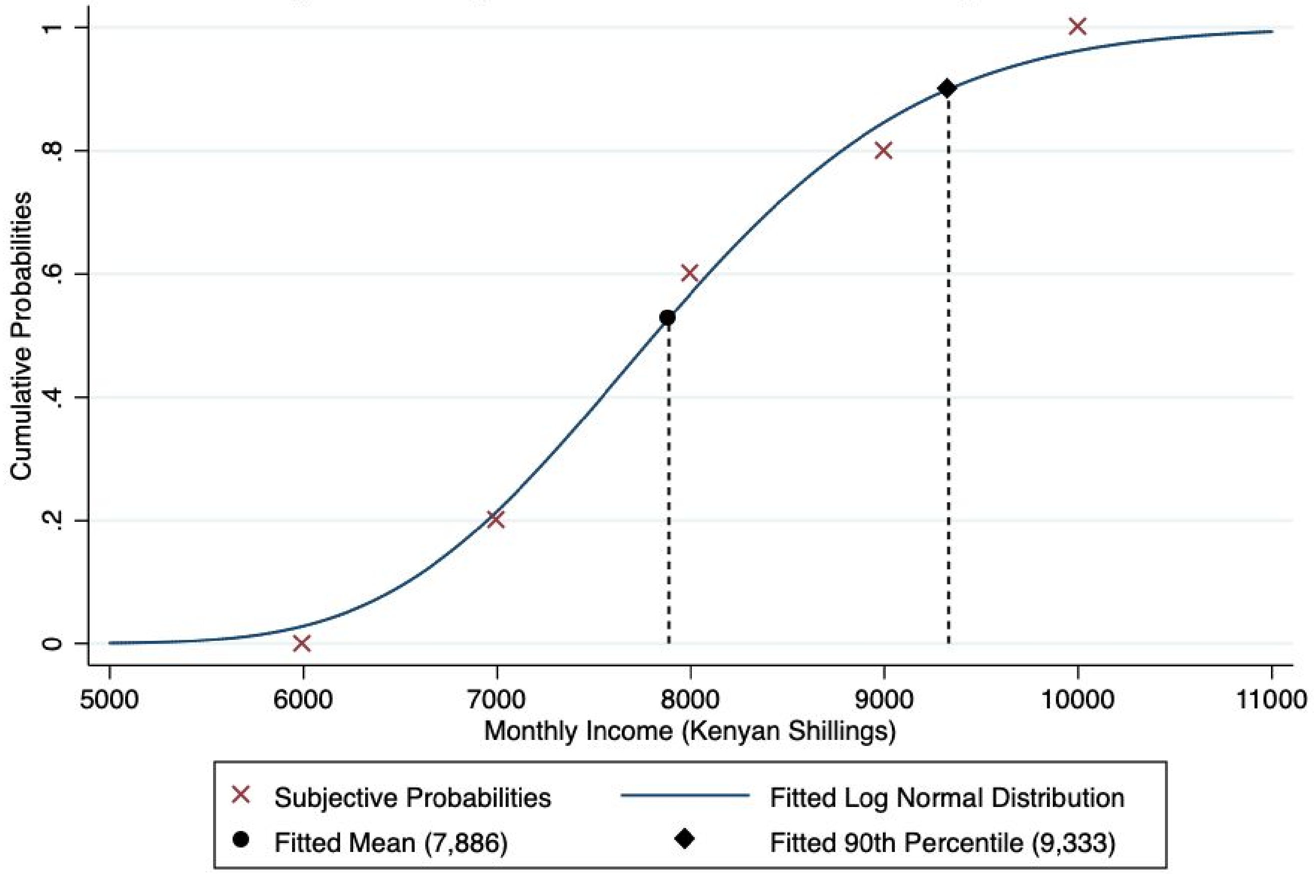
Code:

```
** Code to Fit Subjective Probabilities to a Log Normal Distribution in Stata
/*
The pairs of thresholds and cumulative probabilities are labelled as variable_t and
variable_f respectively. The data is long and for each individual, with a unique
identifier id, there are N rows, one for each interval. Each individual distribution
consists of two parameters: mu and sigma.
*/
gen parameter_mu     = .    // parameter mu
gen parameter_sigma  = .    // parameter sigma

gen ln_variable_t   = ln(variable_t)    // natural log of thresholds

sum id
forvalues i = 1/`r(max)' {                  // code iterates over all individuals
    sum ln_variable_t if id == `i'          // summarise log of variable for individual `i'
    local start_mu      = `r(mean)'         // starting point for mu
    local start_sigma   = `r(sd)'           // starting point for sigma

    ** Solution to the Least Squares Problem
    nl (variable_f = (norm((ln(variable_t) - {mu})/{sigma} ))) if id == `i', ///
        initial(mu `start_mu' sigma `start_sigma') nocons
    // nl fits an arbitrary nonlinear regression function by least squares
    replace parameter_mu     = _b[/mu]     if id == `i'   // update parameter mu
    replace parameter_sigma  = _b[/sigma]  if id == `i'   // update parameter sigma
}
```

Obtaining Measures

Obtaining measures once a distribution has been fitted is relatively straightforward and will depend entirely on the distribution of choice. Examples of measures such as the fitted mean and 90th percentile are provided in Figure 1.

Figure 1 - Subjective Distribution of Income Expectations



Extensions with Expectations Distribution

A natural extension of having an expectations distribution is to use it to positionalise other estimates. For example, consider a researcher who has elicited a point estimate for future aspirations. Comparisons with measures such as the mean or certain quantiles can be useful, but inputting this estimate into the cumulative distribution function would allow a researcher to understand how likely an individual believes their aspiration can be achieved. A similar approach using any other relevant estimate also works. Consider some poverty threshold which can be inputted into the distribution to provide a researcher with an individual's *subjective* expected probability of being in or out of poverty in the future.

References

Attanasio, Orazio, and Britta Augsburg. 2016. "Subjective expectations and income processes in rural India." *Economica* 83 (331): 416-442.

Delavande, Adeline, Xavier Giné, and David McKenzie. 2011a. "Measuring subjective expectations in developing countries: A critical review and new evidence." *Journal of Development Economics* 94 (2): 151-163.

Delavande, Adeline, Xavier Giné, and David McKenzie. 2011b. "Eliciting probabilistic expectations with visual aids in developing countries: how sensitive are answers to variations in elicitation design?." *Journal of Applied Econometrics* 26 (3): 479-497.

Dominitz, Jeff, and Charles F. Manski. 1997. "Using expectations data to study subjective income expectations." *Journal of the American Statistical Association* 92 (439): 855-867.

Dominitz, Jeff. 1998. "Earnings expectations, revisions, and realizations." *Review of Economics and Statistics* 80 (3): 374-388.

Dominitz, Jeff. 2001. "Estimation of income expectations models using expectations and realization data." *Journal of Econometrics* 102 (2): 165-195.

Manski, Charles F. 2004. "Measuring expectations." *Econometrica* 72 (5): 1329-1376.

McKenzie, David, John Gibson, and Steven Stillman. 2013. "A land of milk and honey with streets paved with gold: Do emigrants have over-optimistic expectations about incomes abroad?." *Journal of Development Economics* 102: 116-127.