OPHI

OXFORD POVERTY & HUMAN DEVELOPMENT INITIATIVE
www.ophi.org.uk

UNIVERSITY OF OXFORD

Summer School on Multidimensional Poverty Analysis

11–23 August 2014

Oxford Department of International Development
Queen Elizabeth House, University of Oxford

Tabita, Kenya · Rabiya, India · Stéphanie, Madagascar · Agathe, Madagascar · Dalma, Kenya · Ann-Sophia, Kenya · Valérie, Madagascar

OPHI
OXFORD POVERTY & HUMAN DEVELOPMENT INITIATIVE
www.ophi.org.uk

UNIVERSITY OF OXFORD

Robustness Analysis and Statistical Inference

Suman Seth & Bouba Housseini

19 August 2014

Session I

Tabita, Kenya   Rabiya, India   Stéphanie, Madagascar   Agathe, Madagascar   Dalma, Kenya   Ann-Sophia, Kenya   Valérie, Madagascar

Main Sources of this Lecture

- Alkire S., J. E. Foster, S. Seth, S. Santos, J. M. Roche, P. Ballon, Multidimensional Poverty Measurement and Analysis, Oxford University Press, forthcoming, (Ch. 8).
- Alkire, S. and M. E. Santos (2014), Measuring Acute Poverty in the Developing World: Robustness and Scope of the Multidimensional Poverty Index. *World Development* Vol. 59, pp. 251–274, 2014.
- Batana, Y. M. (2013). Multidimensional Measurement of Poverty Among Women in Sub-Saharan Africa. Social Indicators Research, 1–26.

Focus of This Lecture

- How accurate are the estimates?
- If they are used for policy, what is the chance that they are mistaken?
  - How sensitive policy prescriptions are to choices of parameters used for designing the measure (**Robustness Analysis**)
  - How accurate policy prescriptions are subject to the sample from which the they are computed (**Statistical Inferences**).

Policy Prescriptions Often of Interest

- A central government wants to allocate budget to the poor according to the MPI in each **region** of the country
  - Need to test if the regional comparisons are robust and statistically significant

- A minister wants to show the steepest decrease in poverty in their region/dimension
  - Need to test if the inter-temporal comparisons are robust and statistically significant.

Importance of Robustness Analyses

Comparisons may alter when parameters vary

– An example with the Global MPI

For *k* = 1/3

– MPI for Zambia is 0.328 > MPI of Nigeria is 0.310

For *k* = 1/2

– MPI of Nigeria is 0.232 > MPI for Zambia is 0.214

k : The poverty cutoff. A person with a deprivation score equal to or greater than what is identified as poor.

How are Statistical Tests Important?

Differences in estimates may be of the same magnitude, but statistical inferences may not be the same

– An example comparing Indian states

State	Adjusted Headcount Ratio (M_0)	Difference	Statistically Significant?
Goa	0.057	0.31	Yes
Punjab	0.088		
Maharashtra	0.194	0.32	No
Tripura	0.226		

Source: Alkire and Seth (2013)

Tabita, Kenya Rabiya, India Stephanie, Madagascar Agathe, Madagascar Dalma, Kenya Ann-Sophia, Kenya Valérie, Madagascar

Robustness Analysis

OPHI Oxford Poverty & Human Development Initiative

UNIVERSITY OF OXFORD

Parameters of M_0

The M_0 measure and its partial indices are based on the following parameter values:

– Poverty cutoff (k)

– Weighting vector (w)

– Deprivation cutoffs (z)

An extreme form of robustness is *dominance.*

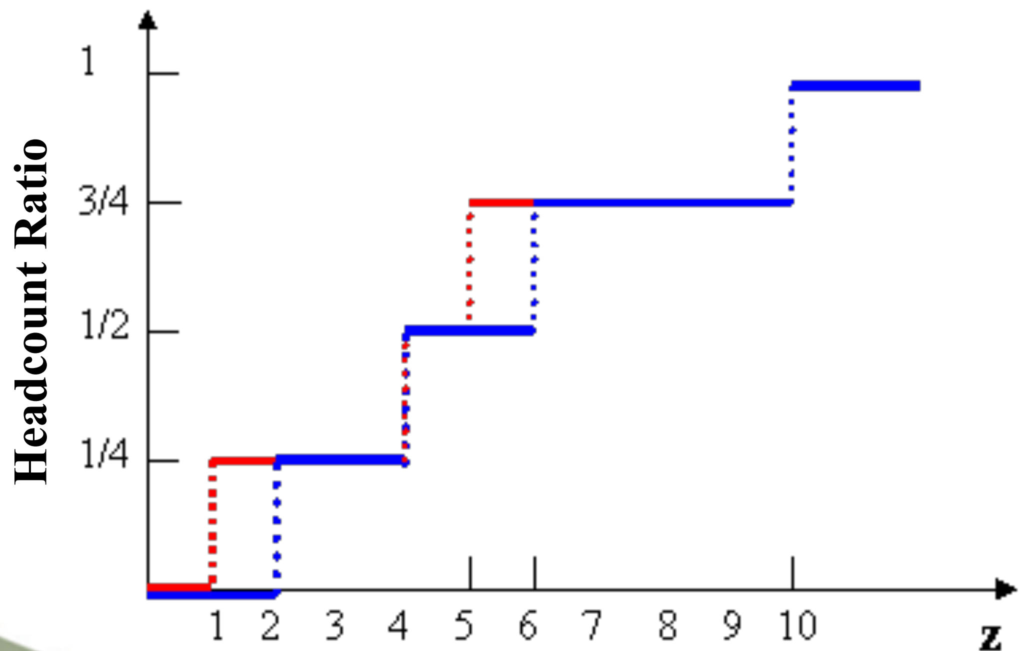
Robustness Analysis

1. Dominance Analysis for Changes in the Poverty Cutoff
   – With respect to poverty cutoff (analogous to unidimensional dominance)
   – Multidimensional dominance

2. Rank Robustness Analysis
   – With respect to weights
   – With respect to deprivation cutoffs.

First Order Unidimensional Dominance

Example: Let x=(2,4,6,10) and y=(1,4,5,10) be two income distributions



No part of y lies to the right of x

In this case x dominates y, which means that x has unambiguously less poverty than y according to the headcount ratio.

Dominance for H and M_0

Question: When can we say that a distribution has higher H or M_0 for any poverty cutoff (k), for a given weight vector and a given deprivation cutoff vector?

Hint: The concept can be borrowed from unidimensional stochastic dominance.

Alkire and Foster (2011)

Dominance for H and M_0 in AF

Consider the following deprivation matrix

$$\begin{array}{cccccc} & \textbf{Income} & \textbf{Years of Education} & \textbf{Sanitation (Improved?)} & \textbf{Access to Electricity} \\ g^0 = & \left|\begin{matrix} 0 \\ 1 \\ 1 \\ 0 \end{matrix}\right. & \begin{matrix} 0 \\ 0 \\ 1 \\ 1 \end{matrix} & \begin{matrix} 0 \\ 0 \\ 1 \\ 0 \end{matrix} & \left.\begin{matrix} 0 \\ 1 \\ 1 \\ 0 \end{matrix}\right| \end{array}$$

$$z = \left| \begin{matrix} \mathbf{500} & \mathbf{12} & \mathbf{1} & \mathbf{1} \end{matrix} \right|$$

Dominance for H and M_0 in AF

- For equal weight, the deprivation count vector is c

	Income	Years of Education	Sanitation (Improved?)	Access to Electricity	c
$g^0 =$	0	0	0	0	**0**
	1	0	0	1	**2**
	1	1	1	1	**4**
	0	1	0	0	**1**
$z =$	**500**	**12**	**1**	**1**	

Dominance for H and M_0 in AF

Result (Alkire and Foster 2011)

– If a deprivation score vector c for joint distribution X first order stochastically dominates another deprivation score vector c' of X', then X has *no higher* H and M_0 than X' for all k and X has strictly lower H and M_0 than X' for some k

Note, however, that the distribution functions would be downward sloping instead of upward rising.

Complementary CDF (CCDF)

CDF of a distribution x is denoted by F_x

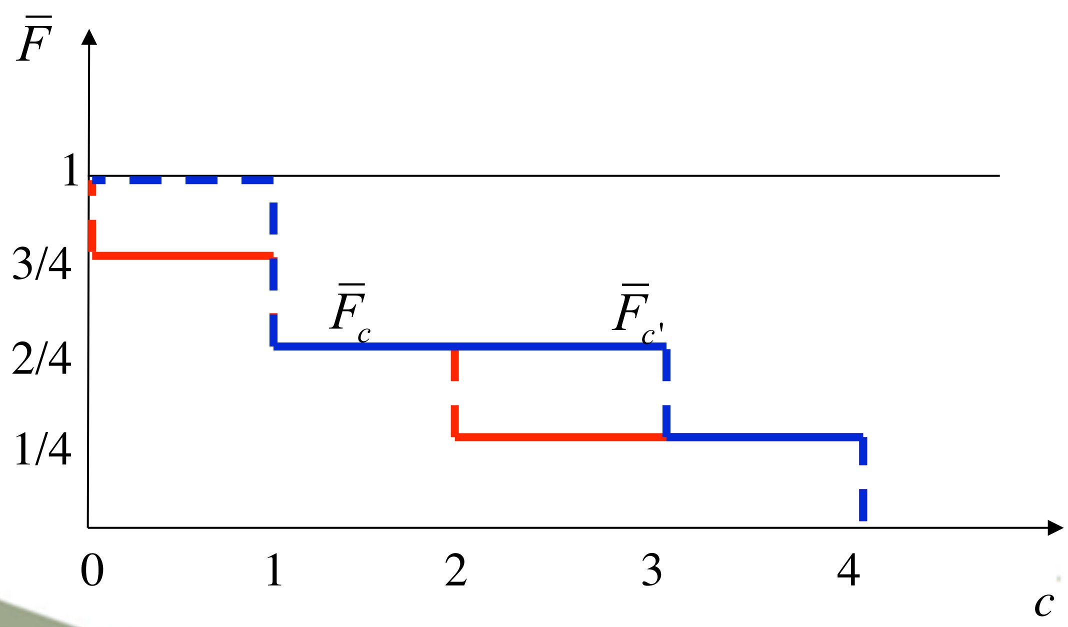
CCDF of a distribution x is $\bar{F}_x = 1 - F_x$

CCDF is also known as survival function or reliability function in other branch of literature

$\bar{F}_x(b)$ denotes the proportion of population with values larger than b.

Example

Let the two deprivation score (count) vectors be
$c = (0, 1, 2, 4)$ and $c' = (1, 1, 3, 4)$



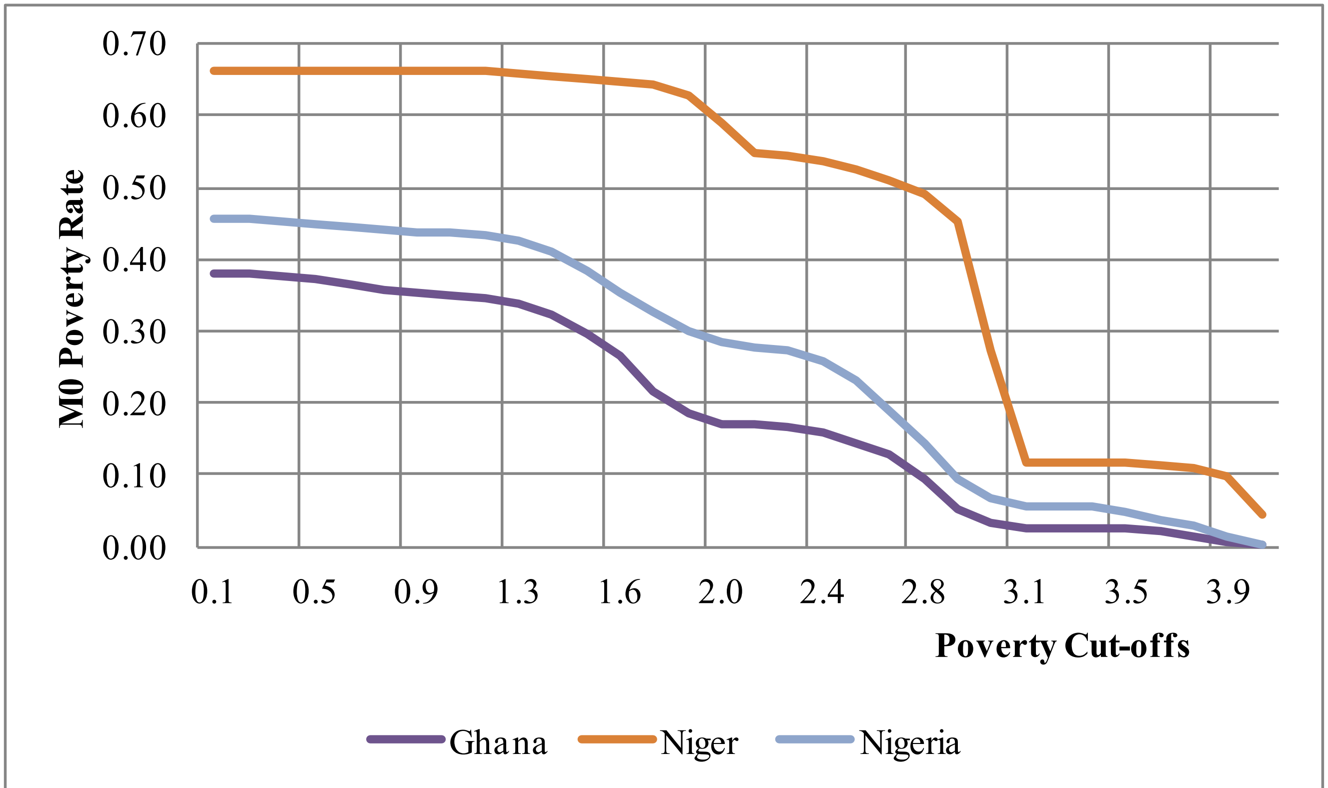
Is there any poverty (k) for which there is more poverty in c than in c'?

H dominance implies M_0 dominance.

Similar Concepts: M_0 Curves

Dominance holds in terms of M_0 for all k

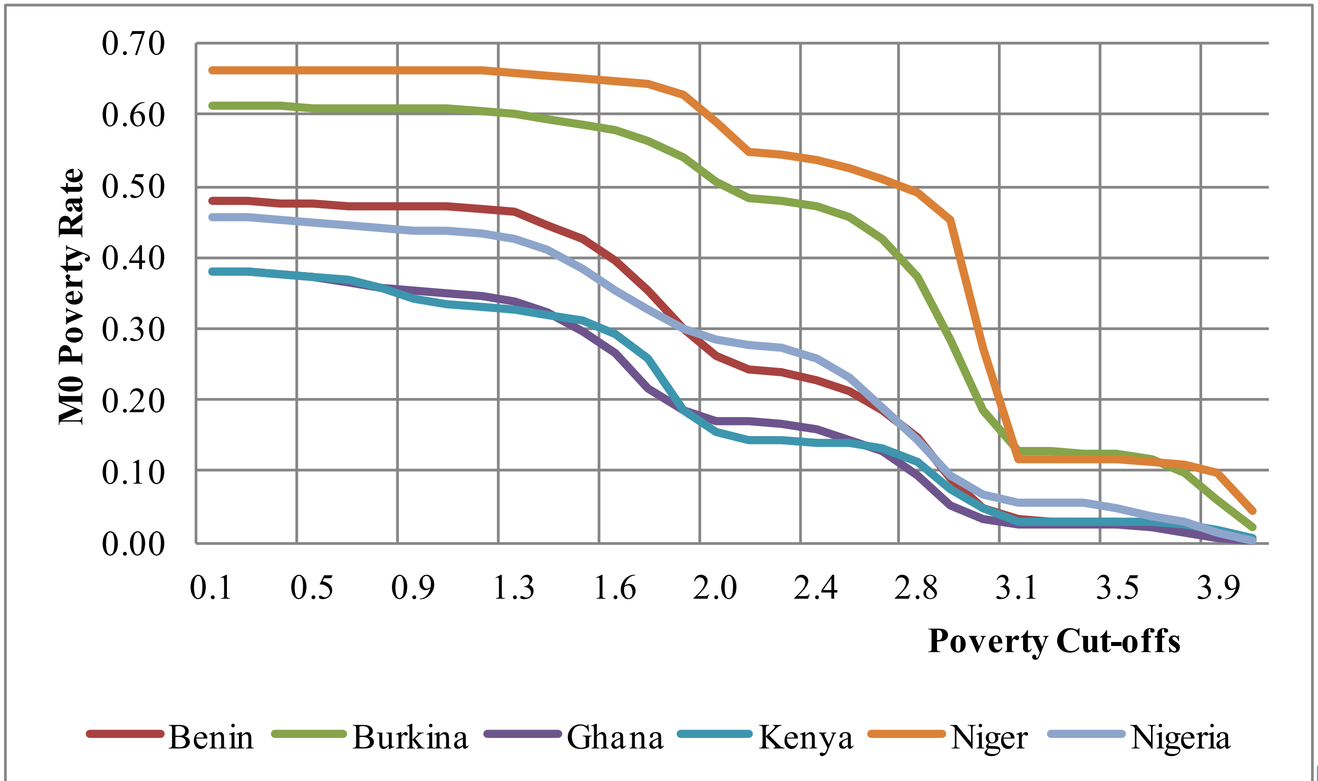
In the case of sample surveys, statistical tests are required to establish dominance



Source: Batana (2013)

M_0 Curves May, However, Cross

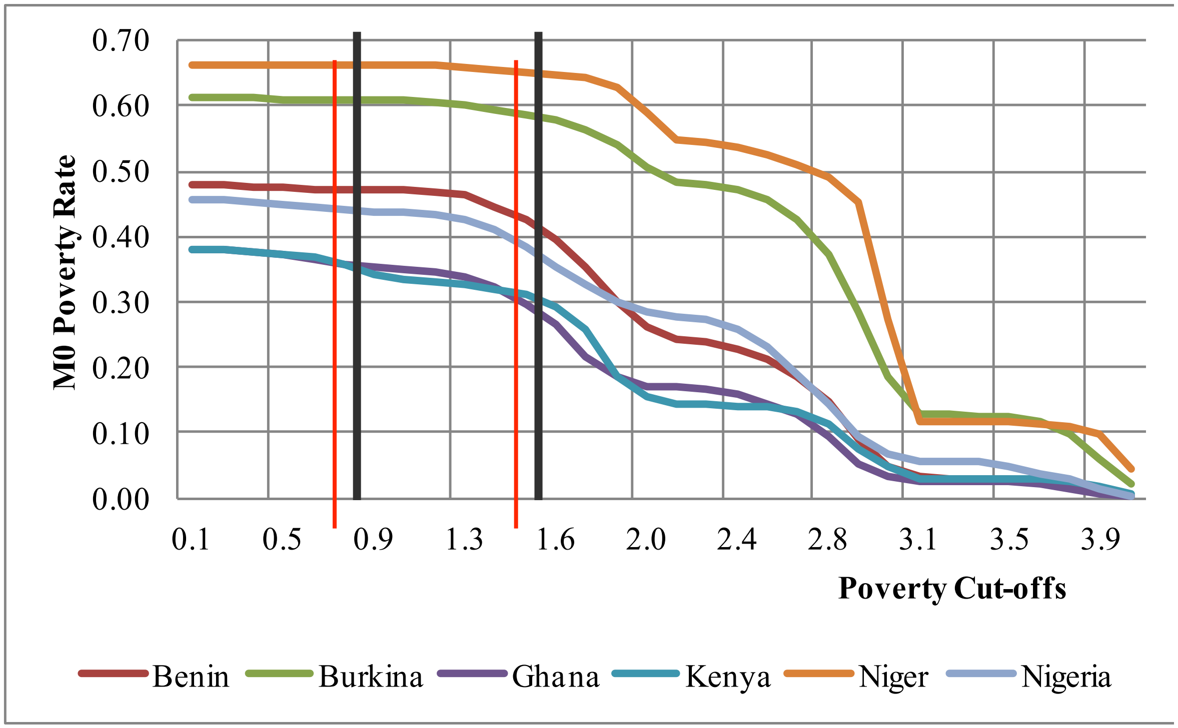
Note below that not all countries stochastically dominate each other (Batana 2013)



M_0 Curves May, However, Cross

Is a particular comparison between two countries or regions robust when *k* varies between an interval?

If *k* is between 0.9 and 1.6, the comparisons between Niger & Burkina Faso and Benin & Nigeria <u>may</u> be robust, but those between Kenya & Ghana are not.



What if ***k*** ranges between 0.9 and 2.4?

OPHI Oxford Poverty & Human Development Initiative

UNIVERSITY OF OXFORD

Rank Robustness Analyses

Dominance across all comparison values is an extreme form of robustness

Stochastic Dominance (SD) conditions are useful for pair by pair analysis and provides the <u>strongest</u> possible comparisons

SD conditions, however, may be too stringent and <u>may not hold</u> for the majority of countries.

Rank Robustness Analysis

Until now, we have compared the robustness of comparisons across countries or regions to varying poverty cutoffs.

How can we evaluate the ranking of a set of countries or regions, when

- the poverty cut-off varies
- the weights vary
- the deprivation cutoffs vary.

Rank Robustness of Comparisons

A useful method for comparing robustness of ranking is to compute rank correlation coefficients

- Spearman's rank correlation coefficient
- Kendall's rank correlation coefficient
- Percentage of pair-wise comparisons that are robust

First, different rankings of countries or regions are generated for different specifications of parameters

- Different weighting vectors, different poverty/deprivation cutoffs

Next, the pair-wise ranks and rank correlation coefficients are computed.

Kendall's Tau

- For each pair, we find whether the comparison is *concordant* or *discordant*
  - 10 countries means 45 pair-wise comparisons

- The comparison between a pair of countries is *concordant* if one dominates the other for both specifications (C)

- The comparison between a pair of countries is *discordant* if one dominates the other for one specification, but is dominated for the other specification (D).

Kendall's Tau

- The Kendall's Tau rank correlation coefficient (t) is equal to

$$\tau = \frac{C - D}{C + D}$$

- It lies between -1 and +1
- If there are ties, this measure should be adjusted for ties
  - The tie adjusted Tau is known as tau-b

Spearman's Rho

- The Spearman's Rho also measures rank correlation but is slightly different from Tau
  - First, countries are ranked for two specifications
  - Then, for each country the difference in the two ranks are computed (r_i for country i)

- The Spearman's Rho (r) is

$$\rho = 1 - \frac{6\sum_{i=1}^{n} r_i^2}{n(n^2 - 1)}$$

Some Illustrations using the MPI

Robustness to weights

Re-weight each dimension:

– 33%	**50%**	25%	25%
– 33%	25%	**50%**	25%
– 33%	25%	25%	**50%**

Robustness to Weights

			MPI Weights 1	MPI Weights 2	MPI Weights 3
			Equal weights: 33% each **(Selected Measure)**	50% Education 25% Health 25% LS	50% Health 25% Education 25% LS
MPI Weights 2	50% Education	Pearson	0.992		
	25% Health	Spearman	0.979		
	25% LS	Kendall (Taub)	0.893		
MPI Weights 3	50% Health	Pearson	0.995	0.984	
	25% Education	Spearman	0.987	0.954	
	25% LS	Kendall (Taub)	0.918	0.829	
MPI Weights 4	50% LS	Pearson	0.987	0.965	0.975
	25% Education	Spearman	0.985	0.973	0.968
	25% Health	Kendall (Taub)	0.904	0.863	0.854
Number of countries:		**109**			

Alkire and Santos (2010, 2014).

Robustness to Poverty Cutoff (*k*)

Spearman's Rank Correlation Table for different poverty cutoffs out of 10 indicators in India

Cut-off (k)	3	4	5	6	7
4	1.00	-	-	-	-
5	0.99	1.00	-	-	-
6	0.99	1.00	1.00	-	-
7	0.97	0.97	0.98	0.98	-
8	0.96	0.96	0.96	0.97	0.98

Alkire and Seth (2008)

Statistical Inferences

Common Concerns

1. Does the overall poverty measure of a country amount to P?
2. Is the overall poverty larger or smaller in one region than another region?
3. Has the overall poverty increased or decreased over time?

One often needs to infer these conclusions (related to population) from a sample (as collecting data from the population is too expensive).

Some Terminologies

Inferential statistics, such as standard error (SE) and confidence intervals (CI), deal with inferences about populations **based on** the behavior of **samples**.

Both SEs and CIs will help us determine how likely it is that results based on a sample (or samples) are the same results that would have been obtained for the entire population.

Standard Error & Confidence Interval

Standard error of a random variable is the sample estimation of its (population) standard deviation. The standard error gives us an idea of the precision of the sample estimation.

Standard deviation, intuitively, is a notion of uncertainty.

Confidence interval contains the true population parameter with some probability that is known as the confidence level. Standard errors are required to compute the confidence interval.

How to Obtain the Standard Error

To compute the standard error, we can use:

1. **Analytical methods**: “Formulas” which either provide the exact or the asymptotic approximation of the standard error (Yalonetzky, 2010).

2. **Resampling methods**: Standard errors and confidence intervals may be computed through bootstrap (Alkire & Santos, 2014).

How Can a Confidence Interval be Used?

- Based on the sample, we can reject any claim that India's M_0 is equal or more than 0.256 or equal or less than 0.245 with 95% probability
- However, we cannot reject with the same probability that India's M_0 lies anywhere strictly between 0.245 and 0.256

	India 2005/6			
Estimate	Value	Standard Error	Confidence Interval (95%)	Confidence Interval (99%)
$\hat{M}_0$	0.251	0.0026	(0.245, 0.256)	(0.244,0.258)
$\hat{H}$	48.5%	0.41%	(47.7%, 49.3%)	(47.4%,49.6%)
$\hat{A}$	51.7%	0.20%	(51.3%, 52.1%)	(51.2%,52.2%)

Source: Alkire and Seth (2013)

Hypothesis Tests

- Confidence intervals when the population parameter is unknown.
- Testing a hypothesis about what the population parameter is; For example, suppose an incumbent government hypothesizes that the Adjusted Headcount Ratio in India in 2006 is 0.26.
- Testing the null hypothesis H_0: $M_0 = 0.26$ VS H_1: $M_0 \neq 0.26$ (or :$M_0 > 0.26$ or M_0: $M0 < 0.26$).

Hypothesis Tests

- Two-tail test; the null hypothesis H_0 can be rejected against the alternative H_1: $M_0 \neq 0.26$ with $(1\text{-}\omega)$ percent confidence if $|(M_0\text{-}0.26)/se_{M0}|>|z_{\omega/2}|$.
- An equivalent procedure entails comparing the significance level against the so-called p-value.
- The conclusions based on the confidence intervals and the one-sample tests are identical.

Statistical Inference in MPI Comparisons

- If the confidence intervals of two point estimates do not overlap, then indeed the comparison is statistically significant
- However, if the confidence intervals do overlap, it does not necessarily mean that the comparison is not statistically significant

Confidence Interval (CI) Overlap | No CI Overlap

Difference is Not Significant | Difference is Significant

Statistical Inference in MPI Comparisons

- With 95% confidence, Punjab's MPI is not larger than 0.103 and no less than 0.073, although the point estimate of MPI is 0.088.

States	MPI	Standard Error	95% Confidence Interval		Difference	
			Lower Bound	Upper Bound	MPI-I	Statistically Significant
Goa	0.057	0.0062	0.045	0.069	0.31	Yes
Punjab	0.088	0.0078	0.073	0.103		
Andhra Pradesh	0.194	0.0093	0.176	0.212	0.32	No
Tripura	0.226	0.0162	0.195	0.258		

Source: Alkire and Seth (2013)

Statistical Inference in MPI Comparisons

- Hypothesis tests on the significance of the comparisons (of A, H or M0) between countries or over time.
- Punjab is significantly poorer than Goa; However, we cannot draw the same kind of conclusion for the comparison between Andhra Pradesh and Tripura.

States	MPI	Standard Error	95% Confidence Interval		Difference	
			Lower Bound	Upper Bound	MPI-I	Statistically Significant
Goa	0.057	0.0062	0.045	0.069	0.31	Yes
Punjab	0.088	0.0078	0.073	0.103		
Andhra Pradesh	0.194	0.0093	0.176	0.212	0.32	No
Tripura	0.226	0.0162	0.195	0.258		

Source: Alkire and Seth (2013)

Conclusions

- It is important to conduct robustness analyses and statistical tests in addition to reporting the estimates
- If estimates and comparisons are not robust to different choices of parameter or not statistically significant, then strong policy conclusions cannot be drawn.