# A short introduction to the computation of standard errors for AF measures

Gaston Yalonetzky

Oxford Poverty and Human Development Initiative, University of Oxford

OPHI-HDCA Summer School, Delft, 24 August - 3 September 2011.

We are grateful to the World Bank, two anonymous donors and OPHI for financial support

# Table of contents

# Introduction: Standard errors for the AF measures

- The computation of standard errors is fundamental for the performance of statistical inference and the production of confidence intervals.

# Introduction: Standard errors for the AF measures

- The computation of standard errors is fundamental for the performance of statistical inference and the production of confidence intervals.
- In classical statistics there are two ways of producing standard errors:

# Introduction: Standard errors for the AF measures

- The computation of standard errors is fundamental for the performance of statistical inference and the production of confidence intervals.
- In classical statistics there are two ways of producing standard errors:
    1. Applying the formulas of analytically derived standard errors. Sometimes these can only be asymptotic approximations (we will see a couple of these).

# Introduction: Standard errors for the AF measures

- The computation of standard errors is fundamental for the performance of statistical inference and the production of confidence intervals.
- In classical statistics there are two ways of producing standard errors:
    1. Applying the formulas of analytically derived standard errors. Sometimes these can only be asymptotic approximations (we will see a couple of these).
    2. With resampling methods (not covered in this lecture, but some references given at the end).

# Introduction: Standard errors for the AF measures

In this lecture we will:

- Quickly review how we compute a standard error for $H$ in the case of a simple random sample.

# Introduction: Standard errors for the AF measures

In this lecture we will:

- Quickly review how we compute a standard error for $H$ in the case of a simple random sample.
- Study how we can derive asymptotic standard errors using $A$ as an example.

# Introduction: Standard errors for the AF measures

In this lecture we will:

- Quickly review how we compute a standard error for $H$ in the case of a simple random sample.
- Study how we can derive asymptotic standard errors using $A$ as an example.
- Derive asymptotic standard errors for $\Delta\%$ for cross-sections and panel datasets.

# Introduction: Standard errors for the AF measures

In this lecture we will:

- Quickly review how we compute a standard error for $H$ in the case of a simple random sample.
- Study how we can derive asymptotic standard errors using $A$ as an example.
- Derive asymptotic standard errors for $\Delta\%$ for cross-sections and panel datasets.
- Briefly discuss computation of standard errors when we have more complex surveys.

## Computing the standard error of $H$

We illustrate with the computation of the standard error of $H$ for simple random samples.

## Computing the standard error of $H$

We illustrate with the computation of the standard error of $H$ for simple random samples.

$H = \frac{1}{N}\sum_{n=1}^{N} y_n$ is a sample average of "1"s and "0"s:

$$H(X; w, k, Z) = \frac{1}{N}\sum_{n=1}^{N} \mathbb{I}(\sum_{d=1}^{D} w_d \mathbb{I}(x_{nd} \leq z_d) \geq k) = \frac{1}{N}\sum_{n=1}^{N} y_n$$

# Computing the standard error of $H$

We illustrate with the computation of the standard error of $H$ for simple random samples.

$H = \frac{1}{N}\sum_{n=1}^{N} y_n$ is a sample average of "1"s and "0"s:

$$H(X; w, k, Z) = \frac{1}{N}\sum_{n=1}^{N}\mathbb{I}\Big(\sum_{d=1}^{D} w_d \mathbb{I}(x_{nd} \leq z_d) \geq k\Big) = \frac{1}{N}\sum_{n=1}^{N} y_n$$

Now assume that each of the $N$ observations are *independently* and *identically* distributed.

# Computing the standard error of $H$

We illustrate with the computation of the standard error of $H$ for simple random samples.

$H = \frac{1}{N}\sum_{n=1}^{N} y_n$ is a sample average of "1"s and "0"s:

$$H(X; w, k, Z) = \frac{1}{N}\sum_{n=1}^{N} \mathbb{I}(\sum_{d=1}^{D} w_d \mathbb{I}(x_{nd} \le z_d) \ge k) = \frac{1}{N}\sum_{n=1}^{N} y_n$$

Now assume that each of the $N$ observations are *independently* and *identically* distributed. In that case the expected value of $H$ is:

$$E[H(X; w, k, Z)] = \frac{1}{N}\sum_{n=1}^{N} E[y_n] = \frac{N}{N}\mathfrak{H} = \mathfrak{H}$$

# Computing the standard error of $H$

We illustrate with the computation of the standard error of $H$ for simple random samples.

$H = \frac{1}{N}\sum_{n=1}^{N} y_n$ is a sample average of "1"s and "0"s:

$$H(X; w, k, Z) = \frac{1}{N}\sum_{n=1}^{N} \mathbb{I}(\sum_{d=1}^{D} w_d \mathbb{I}(x_{nd} \leq z_d) \geq k) = \frac{1}{N}\sum_{n=1}^{N} y_n$$

Now assume that each of the $N$ observations are *independently* and *identically* distributed. In that case the expected value of $H$ is:

$$E[H(X; w, k, Z)] = \frac{1}{N}\sum_{n=1}^{N} E[y_n] = \frac{N}{N}\mathfrak{H} = \mathfrak{H}$$

# Computing the standard error of $H$

Now the standard error of a random variable is the sample estimation of its (population) standard deviation.

# Computing the standard error of $H$

Now the standard error of a random variable is the sample estimation of its (population) standard deviation. What is the standard deviation of $H$?

# Computing the standard error of $H$

Now the standard error of a random variable is the sample estimation of its (population) standard deviation. What is the standard deviation of $H$? We need its variance first. Let's look at the population variance:

# Computing the standard error of $H$

Now the standard error of a random variable is the sample estimation of its (population) standard deviation. What is the standard deviation of $H$? We need its variance first. Let's look at the population variance:

$$E[H - \mathfrak{H}]^2 = \frac{1}{N^2} N\sigma_I^2 = \frac{1}{N}\sigma_I^2$$

# Computing the standard error of $H$

Now the standard error of a random variable is the sample estimation of its (population) standard deviation. What is the standard deviation of $H$? We need its variance first. Let's look at the population variance:

$$E[H - \mathfrak{H}]^2 = \frac{1}{N^2} N \sigma_I^2 = \frac{1}{N} \sigma_I^2$$

$$\sigma_I^2 = E[y_n - \mathfrak{H}]^2 = E[y_n^2] - \mathfrak{H}^2$$

# Computing the standard error of $H$

Now we compute the sample variance of $H$.

# Computing the standard error of $H$

Now we compute the sample variance of $H$. First we replace $\sigma_I^2$ with its sample estimator $\bar{\sigma}_I^2$:

$$\bar{\sigma}_I^2 = \frac{1}{N-1}\sum_{n=1}^{N}[y_n - H]^2 \simeq H(1-H)$$

# Computing the standard error of $H$

Now we compute the sample variance of $H$. First we replace $\sigma_I^2$ with its sample estimator $\bar{\sigma}_I^2$:

$$\bar{\sigma}_I^2 = \frac{1}{N-1}\sum_{n=1}^{N}[y_n - H]^2 \simeq H(1-H)$$

The sample variance of $H$ is: $\bar{\sigma}_H^2 = \frac{\bar{\sigma}_I^2}{N}$

# Computing the standard error of $H$

Now we compute the sample variance of $H$. First we replace $\sigma_I^2$ with its sample estimator $\bar{\sigma}_I^2$:

$$\bar{\sigma}_I^2 = \frac{1}{N-1} \sum_{n=1}^{N} [y_n - H]^2 \simeq H(1-H)$$

The sample variance of $H$ is: $\bar{\sigma}_H^2 = \frac{\bar{\sigma}_I^2}{N}$

Finally the standard error of $H$ is the square root of $\bar{\sigma}_H^2$:

$$SE(H) = \sqrt{\bar{\sigma}_H^2} = \sqrt{\frac{\bar{\sigma}_I^2}{N}}$$

# Asymptotic standard errors for ratios like $A$

For other averages, like $M^{\alpha}$, the procedure to compute standard errors using simple surveys is the same as outlined before for $H$.

# Asymptotic standard errors for ratios like $A$

For other averages, like $M^{\alpha}$, the procedure to compute standard errors using simple surveys is the same as outlined before for $H$.

In STATA this is done easily by generating the respective variable $y_n$ and then using the command "summarize".

# Asymptotic standard errors for ratios like $A$

For other averages, like $M^{\alpha}$, the procedure to compute standard errors using simple surveys is the same as outlined before for $H$.

In STATA this is done easily by generating the respective variable $y_n$ and then using the command "summarize".

However, this procedure cannot be implemented when we have ratios of random variables, (as opposed to averages based on sums).

# Asymptotic standard errors for ratios like $A$

For other averages, like $M^{\alpha}$, the procedure to compute standard errors using simple surveys is the same as outlined before for $H$.

In STATA this is done easily by generating the respective variable $y_n$ and then using the command "summarize".

However, this procedure cannot be implemented when we have ratios of random variables, (as opposed to averages based on sums). In the AF family we have ratios like $A$. For these variables we need to compute *asymptotic* standard errors.

# Asymptotic standard errors for ratios like $A$

$A = \frac{M^0}{H}$ is the ratio of two sample averages of "1"s and "0"s:

$$A(X; w, k, Z) = \frac{\frac{1}{ND}\sum_{n=1}^{N} \mathbb{I}(c_n \geq k) \sum_{d=1}^{D} w_d \mathbb{I}(x_{nd} \leq z_d)}{\frac{1}{N}\sum_{n=1}^{N} \mathbb{I}(\sum_{d=1}^{D} w_d \mathbb{I}(x_{nd} \leq z_d) \geq k)}$$

# Asymptotic standard errors for ratios like $A$

$A = \frac{M^0}{H}$ is the ratio of two sample averages of "1"s and "0"s:

$$A(X; w, k, Z) = \frac{\frac{1}{ND}\sum_{n=1}^{N} \mathbb{I}(c_n \geq k) \sum_{d=1}^{D} w_d \mathbb{I}(x_{nd} \leq z_d)}{\frac{1}{N}\sum_{n=1}^{N} \mathbb{I}(\sum_{d=1}^{D} w_d \mathbb{I}(x_{nd} \leq z_d) \geq k)}$$

The first step for the computation is to approximate $A$ with a first-order Taylor expansion around the population ratio $\frac{\mathfrak{M}^0}{\mathfrak{H}}$.

# Asymptotic standard errors for ratios like $A$

$A = \frac{M^0}{H}$ is the ratio of two sample averages of "1"s and "0"s:

$$A(X; w, k, Z) = \frac{\frac{1}{ND}\sum_{n=1}^{N} \mathbb{I}(c_n \geq k) \sum_{d=1}^{D} w_d \mathbb{I}(x_{nd} \leq z_d)}{\frac{1}{N}\sum_{n=1}^{N} \mathbb{I}(\sum_{d=1}^{D} w_d \mathbb{I}(x_{nd} \leq z_d) \geq k)}$$

The first step for the computation is to approximate $A$ with a first-order Taylor expansion around the population ratio $\frac{\mathfrak{M}^0}{\mathfrak{H}}$. This expansion is a linear function of random variables for which we can compute standard errors:

$$A(X; w, k, Z) - \frac{\mathfrak{M}^0}{\mathfrak{H}} \simeq \frac{1}{H}(M^0 - \mathfrak{M}^0) - \frac{M^0}{H^2}(H - \mathfrak{H})$$

# Asymptotic standard errors for ratios like $A$

The sample covariance of the right-hand side of the expression for $A$ is:

# Asymptotic standard errors for ratios like $A$

The sample covariance of the right-hand side of the expression for $A$ is:

$$\bar{\sigma}_A^2 = \frac{1}{H^2}\frac{\bar{\sigma}_0^2}{N} + (\frac{M^0}{H^2})^2\frac{\bar{\sigma}_H^2}{N} - 2\frac{M^0}{H^3}\frac{\bar{\sigma}_{0,H}}{N}$$

# Asymptotic standard errors for ratios like $A$

The sample covariance of the right-hand side of the expression for $A$ is:

$$\bar{\sigma}_A^2 = \frac{1}{H^2}\frac{\bar{\sigma}_0^2}{N} + (\frac{M^0}{H^2})^2\frac{\bar{\sigma}_H^2}{N} - 2\frac{M^0}{H^3}\frac{\bar{\sigma}_{0,H}}{N}$$

Where:

$$\bar{\sigma}_0^2 = \frac{1}{N-1}\sum_{n=1}^{N}[\mathbb{I}(c_n \geq k)c_n - M^0]^2$$

$$\bar{\sigma}_{0,H} \simeq M^0[1-H]$$

# Asymptotic standard errors for ratios like $A$

The sample covariance of the right-hand side of the expression for $A$ is:

$$\bar{\sigma}_A^2 = \frac{1}{H^2}\frac{\bar{\sigma}_0^2}{N} + (\frac{M^0}{H^2})^2\frac{\bar{\sigma}_H^2}{N} - 2\frac{M^0}{H^3}\frac{\bar{\sigma}_{0,H}}{N}$$

Where:

$$\bar{\sigma}_0^2 = \frac{1}{N-1}\sum_{n=1}^{N}[\mathbb{I}(c_n \geq k)c_n - M^0]^2$$

$$\bar{\sigma}_{0,H} \simeq M^0[1-H]$$

Finally $SE(A) = \sqrt{\bar{\sigma}_A^2}$

# Asymptotic standard errors for $\Delta\%$

Percentage changes are also ratios of random variables: $\Delta\%(H) = \frac{H(t)}{H(t-a)} - 1$.

# Asymptotic standard errors for $\Delta\%$

Percentage changes are also ratios of random variables: $\Delta\%(H) = \frac{H(t)}{H(t-a)} - 1$.Hence a similar approach to the one used for $A$ is required:

# Asymptotic standard errors for $\Delta\%$

Percentage changes are also ratios of random variables: $\Delta\%(H) = \frac{H(t)}{H(t-a)} - 1$.Hence a similar approach to the one used for $A$ is required:

$$\frac{M(t)}{M(t-a)} - \frac{\mathfrak{M}(\mathfrak{t})}{\mathfrak{M}(\mathfrak{t}-\mathfrak{a})} \simeq \frac{1}{M(t-a)}[M(t) - \mathfrak{M}(\mathfrak{t})]$$

$$-\frac{M(t)}{[M(t-a)]^2}[M(t-a) - \mathfrak{M}(\mathfrak{t}-\mathfrak{a})]$$

# Asymptotic standard errors for $\Delta\%$

Percentage changes are also ratios of random variables: $\Delta\%(H) = \frac{H(t)}{H(t-a)} - 1$.Hence a similar approach to the one used for $A$ is required:

$$\frac{M(t)}{M(t-a)} - \frac{\mathfrak{M}(\mathfrak{t})}{\mathfrak{M}(\mathfrak{t}-\mathfrak{a})} \simeq \frac{1}{M(t-a)}[M(t) - \mathfrak{M}(\mathfrak{t})]$$

$$-\frac{M(t)}{[M(t-a)]^2}[M(t-a) - \mathfrak{M}(\mathfrak{t}-\mathfrak{a})]$$

But there is an important difference between the change computed for cross sections and for panel data. The former uses two samples while the latter just one.

# Asymptotic standard errors for $\Delta\%$ in cross sections

The asymptotic variance of $M^\alpha$ is:

$$\bar{\sigma}^2_{\Delta\% M^\alpha} = \frac{1}{[M(t-a)]^2}\frac{\bar{\sigma}^2_{\alpha(t)}}{N(t)} + [\frac{M^\alpha(t)}{[M^\alpha(t-a)]^2}]^2 \frac{\bar{\sigma}^2_{\alpha(t-a)}}{N(t-a)}$$

# Asymptotic standard errors for Δ% in cross sections

The asymptotic variance of $M^{\alpha}$ is:

$$\bar{\sigma}^2_{\Delta\% M^{\alpha}} = \frac{1}{[M(t-a)]^2}\frac{\bar{\sigma}^2_{\alpha(t)}}{N(t)} + [\frac{M^{\alpha}(t)}{[M^{\alpha}(t-a)]^2}]^2\frac{\bar{\sigma}^2_{\alpha(t-a)}}{N(t-a)}$$

Then the standard error is:

$$SE(\Delta\% M^{\alpha}) = \sqrt{\bar{\sigma}^2_{\Delta\% M^{\alpha}}}$$

# Asymptotic standard errors for Δ% in panel data

The asymptotic variance of $M^{\alpha}$ is:

$$\bar{\sigma}^2_{\Delta\% M^{\alpha}} = \frac{1}{[M^{\alpha}(t-a)]^2}\frac{\bar{\sigma}^2_{\alpha(t)}}{N} + [\frac{M^{\alpha}(t)}{[M^{\alpha}(t-a)]^2}]^2\frac{\bar{\sigma}^2_{\alpha(t-a)}}{N}$$

$$-2\frac{M^{\alpha}(t)}{[M^{\alpha}(t-a)]^3}\frac{\sigma_{\alpha(t),\alpha(t-a)}}{N}$$

# Asymptotic standard errors for Δ% in panel data

The asymptotic variance of $M^\alpha$ is:

$$\bar{\sigma}^2_{\Delta\% M^\alpha} = \frac{1}{[M^\alpha(t-a)]^2}\frac{\bar{\sigma}^2_{\alpha(t)}}{N} + [\frac{M^\alpha(t)}{[M^\alpha(t-a)]^2}]^2\frac{\bar{\sigma}^2_{\alpha(t-a)}}{N}$$

$$-2\frac{M^\alpha(t)}{[M^\alpha(t-a)]^3}\frac{\sigma_{\alpha(t),\alpha(t-a)}}{N}$$

Where: $\sigma_{\alpha(t),\alpha(t-a)} =$

$$\frac{1}{N-1}\sum_{n=1}^{N-1}[\mathbb{I}(c_n^t \geq k)\mathbb{I}(c_n^{t-a} \geq k)(\sum_{d=1}^{D}\frac{w_d}{D}[\frac{z_d - x_{nd}(t)}{z_d}]_+^\alpha)(\sum_{d=1}^{D}\frac{w_d}{D}[\frac{z_d - x_{nd}(t-a)}{z_d}]_+^\alpha)]$$

$$-M^\alpha(t)M^\alpha(t-a)$$

# Asymptotic standard errors for $\Delta\%$ in panel data

The asymptotic variance of $M^\alpha$ is:

$$\bar{\sigma}^2_{\Delta\% M^\alpha} = \frac{1}{[M^\alpha(t-a)]^2}\frac{\bar{\sigma}^2_{\alpha(t)}}{N} + [\frac{M^\alpha(t)}{[M^\alpha(t-a)]^2}]^2\frac{\bar{\sigma}^2_{\alpha(t-a)}}{N}$$

$$-2\frac{M^\alpha(t)}{[M^\alpha(t-a)]^3}\frac{\sigma_{\alpha(t),\alpha(t-a)}}{N}$$

Where: $\sigma_{\alpha(t),\alpha(t-a)} =$

$$\frac{1}{N-1}\sum_{n=1}^{N-1}[\mathbb{I}(c_n^t \geq k)\mathbb{I}(c_n^{t-a} \geq k)(\sum_{d=1}^{D}\frac{w_d}{D}[\frac{z_d - x_{nd}(t)}{z_d}]_+^\alpha)(\sum_{d=1}^{D}\frac{w_d}{D}[\frac{z_d - x_{nd}(t-a)}{z_d}]_+^\alpha)]$$

$$-M^\alpha(t)M^\alpha(t-a)$$

Then the standard error is:

$$SE(\Delta\% M^\alpha) = \sqrt{\bar{\sigma}^2_{\Delta\% M^\alpha}}$$

# Standard errors with complex surveys

Different types of surveys are available.

# Standard errors with complex surveys

Different types of surveys are available. For instance, some of the country surveys included in the EU-SILC are simple and random; however most household surveys tend to be more complex, especially they tend to be two-stage, stratified.

# Standard errors with complex surveys

Different types of surveys are available. For instance, some of the country surveys included in the EU-SILC are simple and random; however most household surveys tend to be more complex, especially they tend to be two-stage, stratified.

When computing standard errors the survey features need to be taken into account, in particular:

- Strata

# Standard errors with complex surveys

Different types of surveys are available. For instance, some of the country surveys included in the EU-SILC are simple and random; however most household surveys tend to be more complex, especially they tend to be two-stage, stratified.

When computing standard errors the survey features need to be taken into account, in particular:

- Strata
- Clusters

# Standard errors with complex surveys

Different types of surveys are available. For instance, some of the country surveys included in the EU-SILC are simple and random; however most household surveys tend to be more complex, especially they tend to be two-stage, stratified.

When computing standard errors the survey features need to be taken into account, in particular:

- Strata
- Clusters
- Sampling weights

# Standard errors with complex surveys

Different types of surveys are available. For instance, some of the country surveys included in the EU-SILC are simple and random; however most household surveys tend to be more complex, especially they tend to be two-stage, stratified.

When computing standard errors the survey features need to be taken into account, in particular:

- Strata
- Clusters
- Sampling weights

The formulas can be adapted easily (although a bit tediously) following Deaton's book "The analysis of household surveys" (1997)

# Standard errors with complex surveys

Different types of surveys are available. For instance, some of the country surveys included in the EU-SILC are simple and random; however most household surveys tend to be more complex, especially they tend to be two-stage, stratified.

When computing standard errors the survey features need to be taken into account, in particular:

- Strata
- Clusters
- Sampling weights

The formulas can be adapted easily (although a bit tediously) following Deaton's book "The analysis of household surveys" (1997)

In STATA this is done very easily by first indicating the variables that contain the survey information (with svyset) and then using "svy: sum" to produce the required statistics.

# Example of how STATA handles complex surveys

Imagine we want to compute the standard error of $H$ and the variable identifying households as poor is called "mdpoor".

# Example of how STATA handles complex surveys

Imagine we want to compute the standard error of $H$ and the variable identifying households as poor is called "mdpoor".

In the quite typical case of a stratified two-stage survey, first one tells STATA what are the variables that indicate clusters, strata and weights: pause

svyset: namecluster [pweight=nameweight], strata(namestrata) || householdid

# Example of how STATA handles complex surveys

Imagine we want to compute the standard error of $H$ and the variable identifying households as poor is called "mdpoor".

In the quite typical case of a stratified two-stage survey, first one tells STATA what are the variables that indicate clusters, strata and weights: pause

svyset: namecluster [pweight=nameweight], strata(namestrata) || householdid

Then one can estimate the standard error of $H$ using the command svy: mean applied to "mdpoor":

svy: mean poor

# Illustration: what happens if a complex survey is assumed simple

| | **Simple random sampling** | | **Two-stage stratified design** | |
|---|---|---|---|---|
| K | H | Standard error | H | Standard error |
| 1 | 0.890866 | 0.001469 | 0.861326 | 0.006056 |
| 2 | 0.758537 | 0.001721 | 0.718621 | 0.007356 |
| 3 | 0.589473 | 0.001846 | 0.540775 | 0.008568 |
| 4 | 0.453307 | 0.001857 | 0.409078 | 0.008301 |
| 5 | 0.379032 | 0.001833 | 0.339469 | 0.007826 |
| 6 | 0.268702 | 0.001752 | 0.237406 | 0.006951 |
| 7 | 0.15861 | 0.00159 | 0.14059 | 0.005481 |
| 8 | 0.082467 | 0.00138 | 0.071697 | 0.003729 |
| 9 | 0.03486 | 0.001127 | 0.030272 | 0.00229 |

# Concluding remarks

- The formulas used above are based on a so-called "superpopulation approach".

# Concluding remarks

- The formulas used above are based on a so-called "superpopulation approach". There is an alternative called finite population approach.

# Concluding remarks

- The formulas used above are based on a so-called "superpopulation approach". There is an alternative called finite population approach.
- "A superpopulation approach makes more assumptions, but is in some ways more straightforward".

# Concluding remarks

- The formulas used above are based on a so-called "superpopulation approach". There is an alternative called finite population approach.
- "A superpopulation approach makes more assumptions, but is in some ways more straightforward". "The finite-population approach is more general, in that it makes no assumptions about the homogeneity of the observations in the sample, but it is also more limited, in that it is specifically concerned with one population only , and makes no claim to generality beyond that population". (Deaton, 1997, p. 42)

# Concluding remarks

- The formulas used above are based on a so-called "superpopulation approach". There is an alternative called finite population approach.
- "A superpopulation approach makes more assumptions, but is in some ways more straightforward". "The finite-population approach is more general, in that it makes no assumptions about the homogeneity of the observations in the sample, but it is also more limited, in that it is specifically concerned with one population only , and makes no claim to generality beyond that population". (Deaton, 1997, p. 42)
- In the case of the finite population approach the variance of the sample average has to be multiplied by $(1 - \lambda f)$, where $f = N/\mathcal{N}$, $\mathcal{N}$ is the population and $\lambda$ is equal to 1 if the sampling is without replacement (0 otherwise).

# Concluding remarks

- In practice the differences in estimates using the two approaches are likely to be minimal.

# Concluding remarks

- In practice the differences in estimates using the two approaches are likely to be minimal.
- It is good practice to check for the formulas of different survey designs in survey design textbooks.

# Concluding remarks

- In practice the differences in estimates using the two approaches are likely to be minimal.
- It is good practice to check for the formulas of different survey designs in survey design textbooks. Household surveys come with manuals that (ideally) explain what survey design was implemented.

# Concluding remarks

- In practice the differences in estimates using the two approaches are likely to be minimal.
- It is good practice to check for the formulas of different survey designs in survey design textbooks. Household surveys come with manuals that (ideally) explain what survey design was implemented.
- An alternative to analytically derived standard errors is provided by resampling methods (e.g. bootstraps, jackknifes). There are several of these and they can be quite useful when derivation becomes cumbersome.

# Concluding remarks

- In practice the differences in estimates using the two approaches are likely to be minimal.
- It is good practice to check for the formulas of different survey designs in survey design textbooks. Household surveys come with manuals that (ideally) explain what survey design was implemented.
- An alternative to analytically derived standard errors is provided by resampling methods (e.g. bootstraps, jackknifes). There are several of these and they can be quite useful when derivation becomes cumbersome. For some discussion and examples see Kolenikov ("Resampling inference with complex survey data") and Rao ("Bootstrap methods for analyzing complex sample survey data").