Estimation and Weights

Dr. A.C. Kulshreshtha
U.N. Statistical Institute for Asia and the Pacific (SIAP)

Second RAP Regional Workshop on
Building Training Resources for Improving Agricultural & Rural Statistics
Sampling Methods for Agricultural Statistics-Review of Current Practices
SCI, Tehran, Islamic Republic of Iran
10-17 September 2013

Topics to be covered

- Estimation of population parameter from sample using sample weights
- What are sample weights?
- Why weight?
- How are sample weights computed?

Estimation of population parameter from sample using sample weights - Basic Idea of “Weighting”

- To estimate the total, T_Y of a survey variable Y from sample values y
  1. Multiply each sample value y_i by a *weighting factor*, w_i $$w_i \times y_i$$
  2. Estimate for the total $$\hat{T}_Y = \sum_i w_i \times y_i$$

Derivation of Weighting Factor

- Weighting factor associated with a sample value is based on
  - base or design weight or sampling weight

    = inverse of selection probability of the sample unit
- And is also
  - adjusted for nonresponse
  - calibrated to population control totals
  - {plus other special adjustments}

Illustration

Weighted values

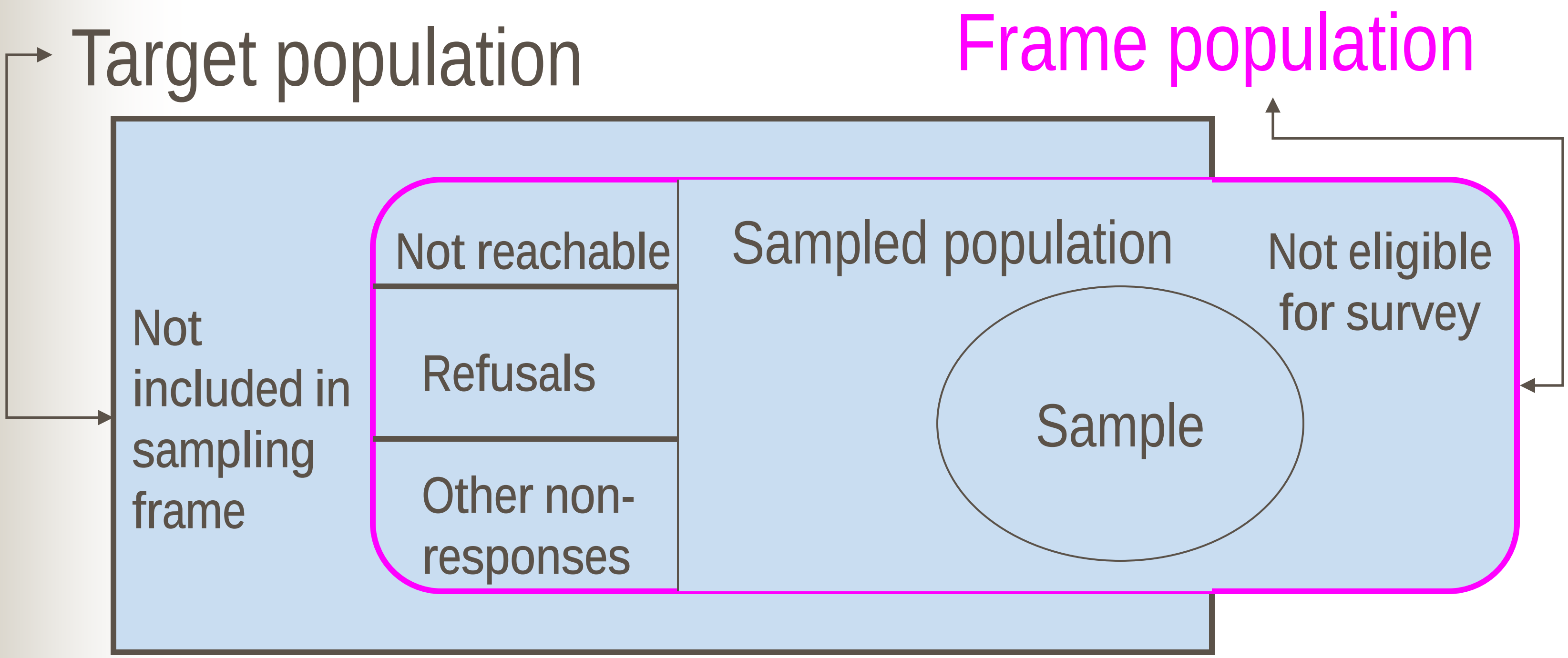
ID	Stratum	Selection probability	Base weight (w_b)	y	Adjustment for non-response (w_r)	Calibration weight (w_c)	Final weight (w_F)	w_F*y
01	1	0.0025	400.0000	8				
02	2	0.0035	285.7143	-				
03	1	0.0018	571.4286	6				
04	2	0.0031	322.5806	10				
05	1	0.0016	625.0000	5				
06	2	0.0035	285.7143	18				
07	2	0.0038	266.6667	18				
08	1	0.0015	666.6667	-				
09	1	0.0028	363.6364	9				
10	1	0.0024	416.6667	8				

$$w_F = w_b \cdot w_r \cdot w_c$$

Imperfections in Sample

- Sample is not representative of population
  - Selection of units with unequal probabilities
  - Non-coverage of parts of the population
  - Survey unable to obtain information from some units in the sample (unit non-response)

In Practice …


Target population
Frame population
Not included in sampling frame
Not reachable
Refusals
Other non-responses
Sampled population
Sample
Not eligible for survey

Why Weight?

- To improve "representativeness" of the sample in terms of the size, distribution and characteristics of the survey population
  - To compensate for unequal probabilities of selection
  - To compensate for (unit) nonresponse
  - To adjust weighted sample distribution to make it conform to population distribution

Design (Base) Weight Concept

- Is the inverse of the probability of selection
  - Thus, depends on the sample selection plan
- Number of units in the population being represented by the sample unit
  - In ideal conditions (e.g., negligible nonresponse, good frame), the design weights take care of "representativeness"
  - But, this is not true in less than ideal conditions

Design (Base) Weight - more!

- Self-representing sampling units are those with probability of selection = 1
  - ‘Take-all’ unit
  - ‘Certainty’ unit
- In **self-weighting** samples each sampled unit has the **same design weight**
  - Computation of estimates is further simplified since the weighting factor is a constant number
  - But, such samples may not always be efficient
  - Because of non-sampling errors, a sample designed to be self-weighting may turn out not to be so

Design Weights: SRS & Systematic

$$\hat{T}_Y = \sum_{i}^{n} w_{bi} \times y_i$$

Design	Design Weight
SRSWR SRSWOR Circular Systematic	$w_{bi} = \frac{N}{n} = \frac{1}{f}$
Linear Systematic	$w_{bi} = k = \frac{N}{n}$ rounded to the nearest integer

Example: SRS Design Weights

N=100, n=5→

$$\pi_i = \frac{n}{N} = \frac{5}{100} = 0.05 \Rightarrow w_{b(i)} = \frac{1}{0.05} = 20$$

Sample Farm	π_i	w_b	y_i	w_b*y
A	0.05	20	10	200
B	0.05	20	24	480
C	0.05	20	59	1180
D	0.05	20	72	1440
E	0.05	20	103	2060

Design Weights: PPS

$$\hat{T}_Y = \sum_{i}^{n} w_{bi} \times y_i$$

Design	Design Weight
PPS PPS Systematic	$w_{bi} = \frac{1}{n\pi_i}; \quad \pi_i = \quad selection\ probability$

Example: PPS Design Weight

Sample Farm	Total Area	π_i	w_b	y_i	w_b*y
A	52	0.009	111.1111	10	1111.11
B	110	0.019	52.63158	24	1263.16
C	300	0.052	19.23077	59	1134.62
D	410	0.071	14.08451	72	1014.08
E	430	0.075	13.33333	103	1373.33

Design Weights: Stratified SRSWOR

$$\hat{t}_{strsrs} = \sum_{h=1}^{H} \sum_{i=1}^{n_h} w_{bhi} \cdot y_{hi}$$

Design	Design Weight
Stratified SRSWOR within each stratum	$w_{bhi} = \frac{N_h}{n_h} = \frac{1}{f_h}$
Proportional allocation	$w_{bhi} = \frac{N_h}{n_h} = \frac{N}{n}$

Design weights (Summary)

Single-stage sampling

Sample selection	**Unbiased estimator of $\hat{Y}$**	**Base weights, w_b**	**Remarks**
SRSWOR SRSWR Circular systematic	$\hat{Y} = \sum_{i=1}^{n} \frac{N}{n} y_i$	$w_i = \frac{N}{n} = \frac{1}{f}$	Base weight is same for all sampling units.
Linear systematic	$\hat{Y} = \sum_{i=1}^{n} k y_i$	$w_i = k = \frac{N}{n}$, rounded to nearest integer	Base weight is exactly same, if k is a whole number and approximately same if not.
PPS or PPS systematic	$\hat{Y} = \frac{1}{n} \sum_{i=1}^{n} \frac{y_i}{\pi_i} = \sum_{i=1}^{n} \frac{1}{n\pi_i} y_i$	$w_i = \frac{1}{n\pi_i}$; $\pi_i = \frac{Z}{z_i}$, where z is measure of size	Base weights vary across sampling units.

Design weights (Contd.)

Stratified single-stage sampling

Sample selection	Unbiased estimator of $\hat{Y}$	Base weights, w_b	Remarks
Stratified SRSWOR within each stratum	$\hat{Y} = \sum_{h,i} \frac{N_h}{n_h} y_{hi}$ where $N = \sum_h N_h$; $n = \sum_h n_h$	$w_{hi} = \frac{N_h}{n_h} = \frac{1}{f_h}$ for all sample units from the hth stratum	Base weight is same within each stratum but may vary among strata. In proportional allocation, $w_{hi} = \frac{N}{n}$ for all strata.

Design weights (Contd.)

Two-stage sampling

Sample selection	Unbiased estimator of $\hat{Y}$	Base weights, w_b	Remarks
Two-stage sampling with SRS at each stage	$\hat{Y} = \sum_{i,j} \frac{M}{m} \frac{N_i}{n_i} y_{ij}$	$w_{ij} = \frac{M}{m} \frac{N_i}{n_i}$ for all sample ssus from the ith sample psu	Base weight is same for all ssus from same psu but varies across psus. Would be constant if $\frac{N_i}{n_i}$ is same for all sample psus
Two-stage sampling with PPS at first stage and SRS at second stage	$\hat{Y} = \sum_{i,j} \frac{1}{m\pi_i} \frac{N_i}{n_i} y_{ij}$	$w_{ij} = \frac{1}{m\pi_i} \frac{N_i}{n_i}$	Base weight is same for all ssus in same psu but may vary across psus.

Design weights-2 stage

PSU	N_i	Method 1			Method 2			Method 3		
		Prob of selection, stage 1	Prob of selection, stage 2	Prob of selection	Prob of selection, stage 1	Prob of selection, stage 2	Prob of selection	Prob of selection, stage 1	Prob of selection, stage 2	Prob of selection
1	20	3/6	4/20	1/10	3/6	5/20	1/8	3(20/96)	4/20	1/8
2	16	3/6	4/16	1/8	3/6	4/16	1/8	3(16/96)	4/16	1/8
3	12	3/6	4/12	1/6	3/6	3/12	1/8	3(12/96)	4/12	1/8
4	16	3/6	4/16	1/8	3/6	4/16	1/8	3(16/96)	4/16	1/8
5	12	3/6	4/12	1/6	3/6	3/12	1/8	3(12/96)	4/12	1/8
6	20	3/6	4/20	1/10	3/6	4/20	1/8	3(20/96)	4/20	1/8

N=96

Self-weighting design

Non-Response Adjustment

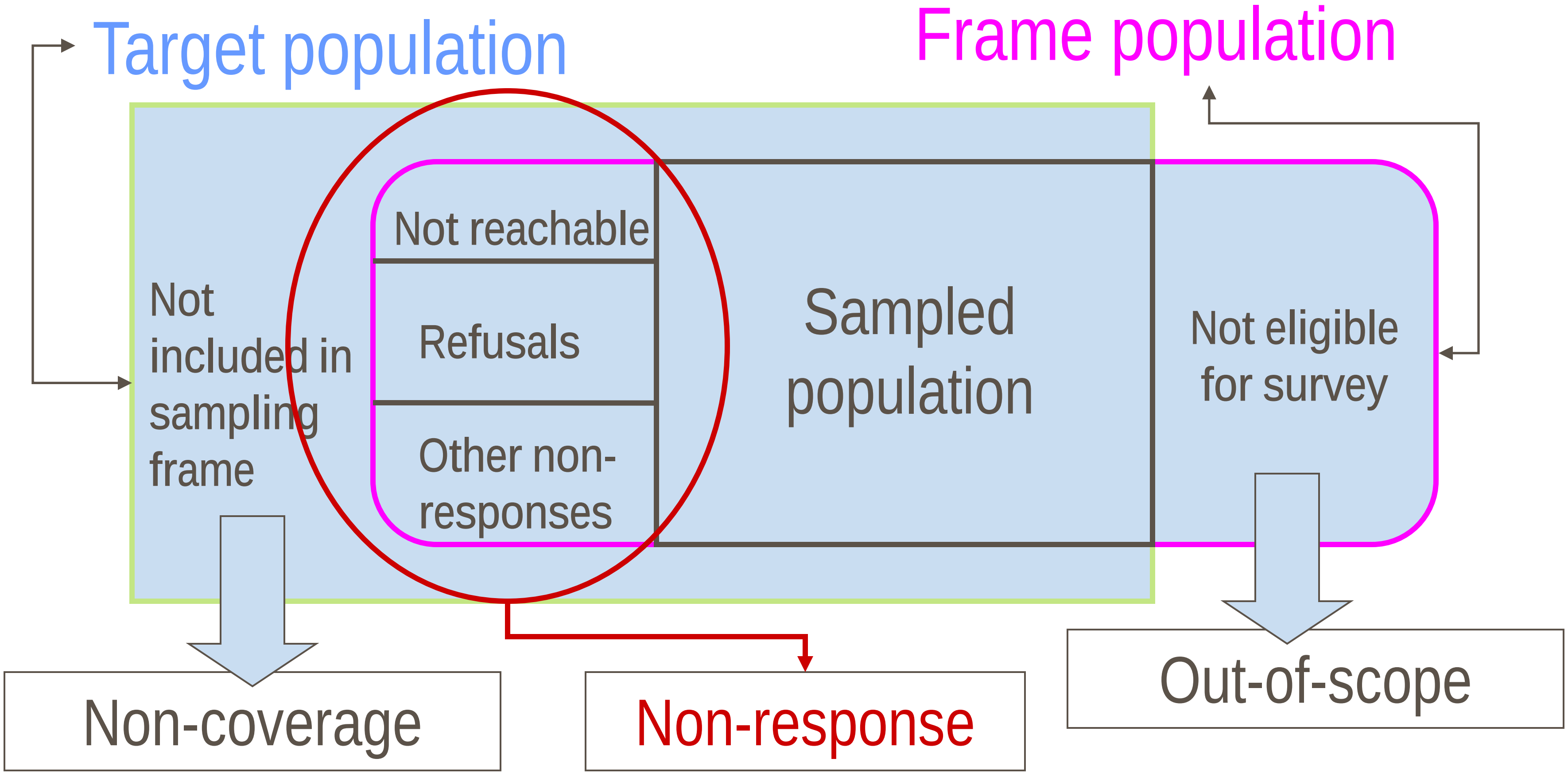


Illustration: Nonresponse Adjustment

ID	Stratum	Selection probability	Base weight (w_b)	y	Adjustment for non-response (w_r)
01	1	0.0025	400.0000	8	
02	2	0.0035	285.7143	-	
03	1	0.0018	571.4286	6	
04	2	0.0031	322.5806	10	
05	1	0.0016	625.0000	5	
06	2	0.0035	285.7143	18	
07	2	0.0038	266.6667	18	
08	1	0.0015	666.6667	-	
09	1	0.0028	363.6364	9	
10	1	0.0024	416.6667	8	

Non-response

- Unit or total nonresponse- no data at all for unit in the sample
- Item nonresponse- no data for some questions in the survey
- *May introduce **non-sampling bias** in the survey estimates by failing to include a portion of the population*

Non-response Bias

- When estimates are based only on values of those who respond …
- Estimates are overestimated or underestimated …
- When characteristics of nonrespondents are different from characteristics of respondents

Illustration of Non-response Bias

Stratum	Size	Total	Mean	Variance
Respondents	N_R	T_R	$\bar{Y}_R$	S_R^2
Nonrespondents	N_{NR}	T_{NR}	$\bar{Y}_{NR}$	S_{NR}^2
Whole Population	N	T	$\bar{Y}$	S^2

Let $\hat{\theta}_R$ be an unbiased estimator for $\bar{Y}_R$ using only respondents

$$\bar{Y} = \frac{N_R}{N} \cdot \bar{Y}_R + \frac{N_{NR}}{N} \cdot \bar{Y}_{NR} \Rightarrow$$

$$Bias(\hat{\theta}_R) = E(\hat{\theta}_R) - \bar{Y} = \frac{N_{NR}}{N} \cdot (\bar{Y}_R - \bar{Y}_{NR})$$

$$Bias(\hat{\theta}_R) = E(\hat{\theta}_R) - \bar{Y} = \frac{N_{NR}}{N} \cdot (\bar{Y}_R - \bar{Y}_{NR})$$

… therefore, bias is small if either

- mean for nonrespondents is close to the mean for the respondents, or
- proportion of nonrespondents in the population is small; that is, there is very little non-response

Reducing Unit Non-response Bias

- Replace by drawing a larger sample than needed and creating a reserve sample from which replacements are selected
  - but, may be useless if replacements are more similar to responding rather than non-responding units
- Replace by substitution with a unit that was not sampled
  - Neighboring household for non-responding household
  - ‘Similar’ establishment for non-responding establishment
- Non-response adjustment of the weights

Non-response Adjustment

- Concept–distribute base weights of all eligible non-responding sampled units among the responding units
- Types of adjustment procedures
  - Weighting class adjustments
  - Regression-analysis based procedures
  - Raking adjustments

Adjustment for nonresponse (example 1)

FSIC	n_i	# responding	Rate-up factor
Complete enumeration			
6111/6112	24	21	1.1429
6121/6122	15	15	1.0000
6130	4	3	1.3333
6140	13	13	1.0000
6151/6152	9	7	1.2857
6171/6172	52	49	1.0612
6180	31	28	1.1071
6160/6191/6192	57	49	1.1633
Sample	14	12	1.1667

$$w_r = \frac{\#\,\text{in sample}}{\#\,\text{responding}}$$

Adjustment for nonresponse (example 2)

- Suppose age of every person selected in sample is known.

	Age					
	15-24	25-34	35-44	45-64	65+	Total
Sample size	202	220	180	195	203	1,000
Respondents	124	187	162	187	203	863

Weighting Classes

Adjustment for nonresponse (example 2)

	Age					
	15-24	25-34	35-44	45-64	65+	Total
Sample size	202	220	180	195	203	1,000
Respondents	124	187	162	187	203	863
Sum of weights of sample	30,322	33,013	27,046	29,272	30,451	150,104
Sum of weights of respondents	18,693	28,143	24,371	28,138	30,451	
Response probability	0.6165	0.8525	0.9011	0.9613	1.0000	
w_r	1.6221	1.1730	1.1098	1.0403	1.0000	

=weights of respondents/weights of sample

Adjustment for nonresponse (example 2)

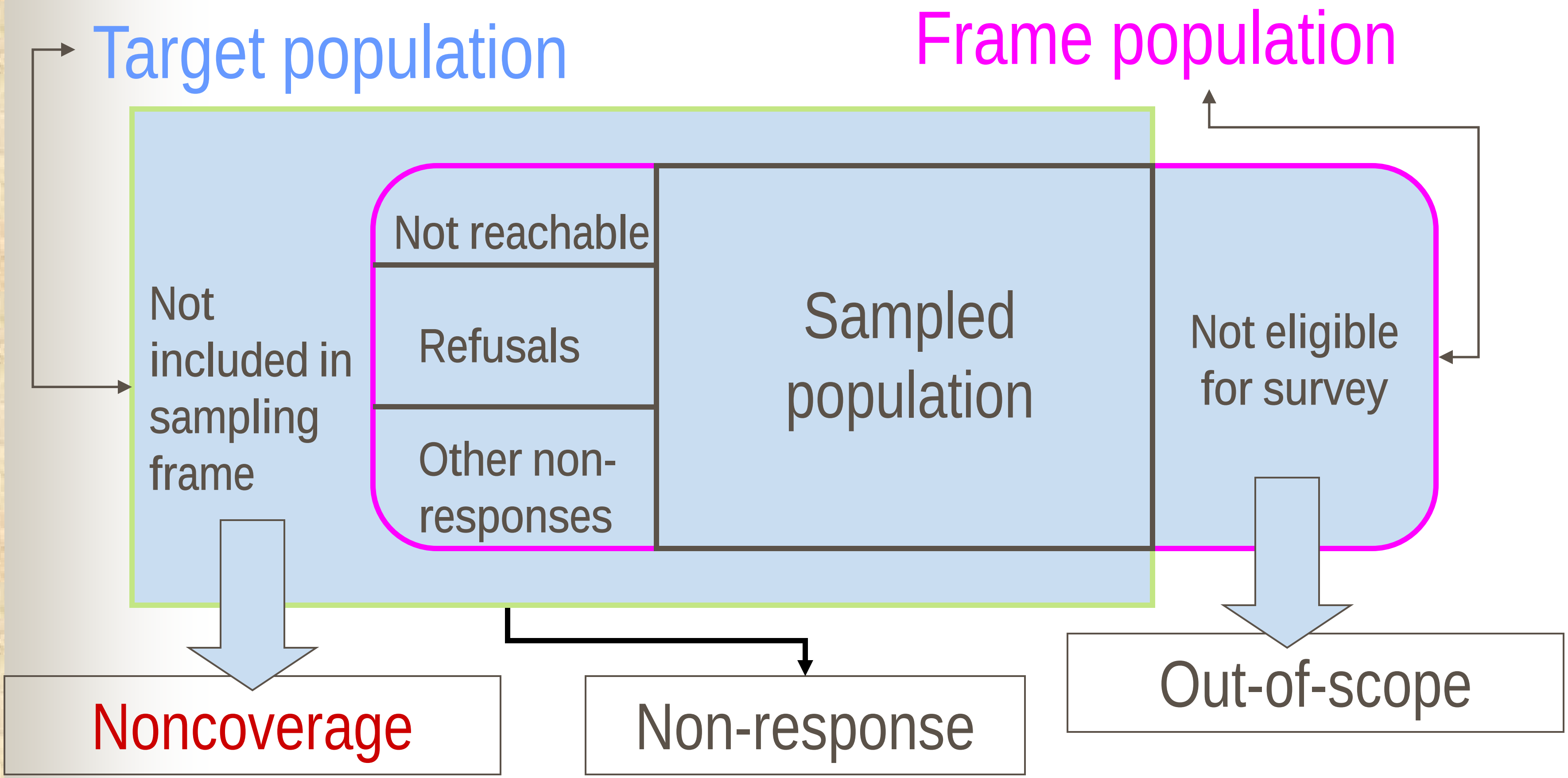
	Age					
	15-24	25-34	35-44	45-64	65+	Total
Sample size	202	220	180	195	203	1,000
Respondents	124	187	162	187	203	863
Sum of weights of sample	30,322	33,013	27,046	29,272	30,451	150,104
Sum of weights of respondents	18,693	28,143	24,371	28,138	30,451	
Response probability	0.6165	0.8525	0.9011	0.9613	1.0000	
w_r	1.6221	1.1730	1.1098	1.0403	1.0000	

=1/response probability

Summary of Steps: Weighting Class

- Apply base weights
- Partition sample into subgroups (weighting classes) Construct classes so that units within each class are homogeneous with respect to the study variable.
- Compute weighted response rates for each of these classes
- Use reciprocal of the response rates for each of the weighting classes for non-response adjustment factor, w_r.
- Non-response adjusted weight is w_b*w_r for respondents and 0 for non-respondents

Calibration Factor, w_c



Adjustment of sampling weights for Non-response!

Non-Coverage

- Failure of sampling frame to cover all of the target population
- Thus, some sampling units have zero probability of selection

Compensating for Non-Coverage

- Improved field procedures such as
  - Improved listing procedures
  - Use of multiple frames
- Statistical adjustment of weights (or calibration factors)

Calibration Factor(s)

- Adjust $\{w_{b*}w_r\}$ further to calibrate to known population totals
- Why calibrate?
  - Reduce bias in estimates introduced by non-coverage
  - Compensate partially for nonresponse bias
  - Make estimates consistent with known totals
  - Reduce sampling error for estimates that are highly correlated to known control totals

Calibration factor(s): Illustration

Age Group	Census Distribution Women	Census Distribution Men	Sample Distribution Women	Sample Distribution Men	Calibration Factors Women	Calibration Factors Men
15-24	18	19	18	20	1.000	0.950
25-34	20	22	22	23	0.909	0.957
35-44	25	27	23	25	1.087	1.080
45-64	22	23	21	22	1.048	1.045
65+	15	9	16	10	0.938	0.900
TOTAL	100	100	100	100	1.000	1.000

Calibrate totals of age groups so that survey totals are same as most recent census totals

Post-stratification Weights- Example

- Suppose the gender distribution of responding units in the sample is different from that of the gender distribution in the population

Sex	Number		Distribution	
	in sample	in population	in sample	in population
Women	550	150,000	0.65	0.60
Men	300	100,000	0.35	0.40
TOTAL	850	250,000	1.00	1.00

Post-strata

Post-stratification Weights- Example

Subclasses	w_b	w_r	w_b*w_r	w_c
Majority Group	450	1.25	562.50	
Women	450	1.25	562.50	0.927
Men	450	1.25	562.50	1.133
Minority Group	50	1.11	55.56	
Women	50	1.11	55.56	0.927
Men	50	1.11	55.56	1.133

=pop'n/sample=0.60/0.65

=0.40/0.35

Post-stratification Using Weights

- Let N_h = number of elements in post-stratum h. Then the post-stratification weight for unit i in stratum h is

$$x_{hi} = \text{indicator for responding unit}$$

$$w_i^c = \sum_{h=1}^{H} w_i^{adj} \cdot x_{hi} \cdot \frac{N_h}{\sum_{j \in sample} w_j^{adj} \cdot x_{hj}}$$

$$w_i^{adj} = w_{bi} \times w_{ri}$$

Raking Adjustments

- Post-strata are formed using more than one variable
- Only the marginal population totals are known
- *Example: (Lohr, S. 1999. Sampling Design & Analysis, pp 270-271)*

	Black	White	Asian	Native	Other	Total
Female						1510
Male						1490
Total	600	2120	150	100	30	3000

Raking Adjustments: Example

- Table of sum of weights

	Black	White	Asian	Native	Other	Total
Female	300	1200	60	30	30	1620
Male	150	1080	90	30	30	1380
Total	450	2280	150	60	60	3000

- Raking adjusts the weights so that the sums of weights in the margins equal the population counts

Example (continued)

- Step 1. Adjust the row sum of weights
  - Multiple female cells weight by 1510/1620
  - Multiply male cells weight by 1490/1380
- Result of raking row totals– now equal to population:

	Black	White	Asian	Native	Other	Sum of Weights	Population
Female	279.6	1118.5	55.9	28.0	28.0	1510	1510
Male	162.0	1166.1	97.2	32.4	32.4	1490	1490
Sum of weights	441.6	2284.6	153.1	60.4	60.4	3000	
Population	600	2120	150	100	30		

Example (continued)

- Step 2. Adjust the column sum of weights obtained after step 1
  - Multiple black cells weight by 600/441.6
  - Multiply white cells weight by 2120/2284.6
  - Multiply Asian cells weight by 150/153.1
  - Multiply native cells weight by 100/60.4
  - Multiply other cells weight by 30/60.4
- Result of raking column totals– now equal to population, but row totals are again different:

	Black	White	Asian	Native	Other	Sum of Weights	
Female	379.9	1037.9	54.8	46.3	13.9	1532.9	1510
Male	220.1	1082.1	95.2	53.7	16.1	1467.1	1490
Sum of weights	600	2120	150	100	30	3000	
Population	600	2120	150	100	30		

Example (continued)

- Repeat steps 1 and 2, using the most recently created tables of adjusted weights until the row sums of weights and column sums of weights converge to the row population totals and column population totals, respectively
- In this example, a total of five steps (iterations) have to be made to get the result:

	Black	White	Asian	Native	Other	Sum of Weights	
Female	375.6	1021.5	53.7	45.6	13.7	1510.0	1510
Male	224.4	1098.5	96.3	54.4	16.3	1490.0	1490
Sum of weights	600.0	2120.0	150.0	100.0	30.0	3000	
Population	600	2120	150	100	30		

Caution About Weights

- Weighting adjustments make use of models or assumptions about the population
- Adjustments may reduce unit nonresponse bias, but not totally eliminate it
- When no adjustments are made, the assumption is that for the entire population, respondents and nonrespondents have similar characteristics
- When post-stratification or weighting class methods are used, the assumption is that in each post-strata or weighting class, respondents and nonrespondents have the same characteristics
- But, what if your model is not correct????

Imputation

- … refers to the procedure of assigning values to a missing data item (item nonresponse).
- Objective is to reduce nonresponse bias
- Also creates a ‘clean’ rectangular data set for analysis
- Caution—
  - Always include information on imputation in data set (e.g., whether imputed or not, donors, etc)
  - Variances with imputed data are underestimated

Methods of Imputation

- Deductive
- Cell Mean
- Hot-Deck
  - Sequential
  - Random
  - Nearest-neighbor
- Regression
- Stochastic Regression
- Cold Deck
- Substitution

- Multiple

Deriving Weights

An Example

Sample Data Records

Identifier	Group	Gender	Income	w_b	w_r	w_c	w_F
1	Majority	Male	200				
2	Majority	Male	-				
3	Minority	Female	130				
4	Majority	Male	500				
.	.	.	.				
.	.	.	.				
.	.	.	.				
999	Majority	Female	-				
1000	Minority	Male	220				

$$w_F = w_b \cdot w_r \cdot w_c$$

Sub-classes

Group/Gender	Women	Men
Majority		
Minority		

Population Sizes			
Group/Gender	Women	Men	Total
Majority			225,000
Minority			25,000
Total	150,000	100,000	250,000

Sample Selection Plan

- Stratified by GROUP with **equal allocation** of sample size n=1,000; SRSWOR selection within strata

Sample Sizes			
Group/Gender	Women	Men	Total
Majority			500
Minority			500
Total			1,000

Design (Base) Weights

Subclasses	w_b
Majority Group	450
Women	450
Men	450
Minority Group	50
Women	50
Men	50

450 = 225,000/500

50 = 25,000/500

Non-response

- Suppose that there are non-responses:

Group	Total Sample	Responding	Response rate
Majority	500	400	0.80
Minority	500	450	0.90
TOTAL	1,000	850	0.85

= number responding/total sample

Adjustment for nonresponse

Subclasses	w_b	w_r	
Majority Group	450	1.25	
Women	450	1.25	=1/0.80
Men	450	1.25	
Minority Group	50	1.11	
Women	50	1.11	=1/0.90
Men	50	1.11	

Adjustment for non-response (Contd.)

Subclasses	w_b	w_r	w_b*w_r
Majority Group	450	1.25	562.50
Women	450	1.25	562.50
Men	450	1.25	562.50
Minority Group	50	1.11	55.56
Women	50	1.11	55.56
Men	50	1.11	55.56

Calibration to Population Totals

- Suppose the sex distribution in responding sample is different from sex distribution in the population

Gender	Number		Distribution	
	in sample	in population	in sample	in population
Women	550	150,000	0.65	0.60
Men	300	100,000	0.35	0.40
TOTAL	850	250,000	1.00	1.00

Post-stratification Weights

Subclasses	w_b	w_r	$w_b{*}w_r$	w_c
Majority Group	450	1.25	562.50	
Women	450	1.25	562.50	0.927
Men	450	1.25	562.50	1.133
Minority Group	50	1.11	55.56	
Women	50	1.11	55.56	0.927
Men	50	1.11	55.56	1.133

=pop'n/sample=0.60/0.65

=0.40/0.35

Final Weights

Subclasses	w_b	w_r	w_b*w_r	w_c	w_b*w_r*w_c
Majority Group	450	1.25	562.50		
Women	450	1.25	562.50	0.927	521.44
Men	450	1.25	562.50	1.133	637.31
Minority Group	50	1.11	55.56		
Women	50	1.11	55.56	0.927	51.50
Men	50	1.11	55.56	1.133	62.94

Sample Data

$=Income*w_F$

Identifier	Group	Gender	Income	w_F	Weighted Income
1	Majority	Male	200	637.31	127,462.00
2	Majority	Male	-	637.31	-
3	Minority	Female	130	51.50	6,695.00
4	Majority	Male	500	637.31	318,655.00
.	.	.	.	.	.
.	.	.	.	.	.
.	.	.	.	.	.
999	Majority	Female	-	521.44	-
1000	Minority	Male	220	62.94	13,846.80

Estimation

- of average income
  - Obtain sum of *weighted income*
  - Divide by sum of *final sampling weights of responding units*

Others

Adjustments for variability of weights

Why adjust for variability?

- Adjustments to base weights for nonresponse and calibration may result in substantial variations in weights or extremely large values
  - Effect→ larger sampling errors of estimates (but note that adjusting may increase *bias* in estimates)
  - Or, estimates are likely to be highly influenced by a few observations
- Can control for excessive variation in sample design stage but may not be completely successful

For example: Variability in design weights

- Considerations at design stage—
  - Sub-sampling b=1 person leads to large variation in weights
  - Sub-sampling b=2 persons, instead, leads to reduction in variability

	Household size				
Weights for:	1	2	3	4	5
Subsample 1 person per HH	1	2	3	4	5
Subsample 2 persons per HH	1	1	1.5	2	2.5

How to adjust?

- At each stage of adjustment:
  - "Trim" or truncate large weights
  - Redistribute the trimmed weight to the rest of sampled units
- When to trim? Some examples--
  - Weights for a small number of cases are larger than the rest– e.g., 3-4 times larger than the average weights
  - RSE or CV of adjusted weights increases by 20% over unadjusted weights

Summary

- What are sample weights?
  - Design or base weights
  - Non-response adjustment factors
  - Calibration factors
  - Adjustments for variability in weights
- Why weight?
- How are sample weights computed?

THANKS