2012 Joint Statistical Meeting

New Model-Optimized Sampling Techniques

2012. 08. 02

Sung-Joon Hong, SunWoong Kim

Department of Statistics, Dongguk University

Outline

1. **Previous Studies of Model-Optimized Sampling Techniques**
2. **New Approaches**
3. **Illustrations**
4. **Conclusions**

1. Previous Studies of Model-Optimized πPS Sampling

Raj (1956)

- Finite population model

$$y_i = \alpha + \beta x_i, \; i = 1, \cdots, N$$

- The variance form of the H-T estimator

$$Var\left(\hat{Y}_{HT}\right) = \sum_{i=1}^{N} \frac{y_i^2}{\pi_i} + 2\sum_{i=1}^{N}\sum_{j>i}^{N} \frac{\pi_{ij}}{\pi_i \pi_j} y_i y_j - Y^2, \; where Y = \sum_{i=1}^{N} Y_i$$

- Optimization problem
  - Minimizing the variance of the H-T estimator under the finite population model

$$Minimize \sum_{i=1}^{N}\sum_{j>i}^{N} \frac{\pi_{ij}}{\pi_i \pi_j} y_i y_j \;\; subject\ to \sum_{j \neq i}^{N} \pi_{ij} = \pi_i, \; i = 1, \cdots, N$$

Rao and Bayless (1969, 1970)

- Superpopulation model

$$y_i = \beta x_i + \varepsilon_i, \ where\ E_\xi\left(\varepsilon_i\right) = 0,\ \ E_\xi\left(\varepsilon_i^2\right) = a x_i^g \ \ \left(a > 0,\ \ g \ge 0\right),$$

$$and\ \ E_\xi\left(\varepsilon_i \varepsilon_j\right) = 0,\ \ i = 1, \cdots, N$$

- The variance form of the H-T estimator

$$Var\left(\hat{Y}_{HT}\right) = \sum_{i=1}^{N} \sum_{j>i}^{N} \left(\pi_i \pi_j - \pi_{ij}\right)\left(\frac{y_i}{\pi_i} - \frac{y_j}{\pi_j}\right)^2$$

- Optimization problem
  - Minimizing the variance of the H-T estimator under the superpopulation model does not depend on π_{ij}, the selection probability of a sample

Kim, Heeringa, and Solenberger (2006)

- Superpopulation Model

$$y_i = \alpha + \beta x_i + \varepsilon_i,\ where\ E_\xi(\varepsilon_i) = 0,\ V_\xi(\varepsilon_i^2) = \delta x_i^\gamma\ \ (\delta > 0,\ \gamma \geq 0),$$

$$and\ \ E_\xi(\varepsilon_i \varepsilon_j) = 0,\ \ i = 1, \cdots, N$$

- The variance forms of the H-T estimator

$$(1)\ Var\left(\hat{Y}_{HT}\right) = \sum_{i=1}^{N} \frac{y_i^2 (1 - \pi_i)}{\pi_i} + 2 \sum_{i=1}^{N} \sum_{j>i}^{N} \frac{y_i y_j (\pi_{ij} - \pi_i \pi_j)}{\pi_i \pi_j}$$

$$(2)\ Var\left(\hat{Y}_{HT}\right) = \sum_{i=1}^{N} \sum_{j>i}^{N} (\pi_i \pi_j - \pi_{ij}) \left(\frac{y_i}{\pi_i} - \frac{y_j}{\pi_j}\right)^2$$

$$(3)\ var_{SYG}\left(\hat{Y}_{HT}\right) = \sum_{i=1}^{N} \sum_{j>i}^{N} \frac{\pi_i \pi_j - \pi_{ij}}{\pi_{ij}} \left(\frac{y_i}{\pi_i} - \frac{y_j}{\pi_j}\right)^2$$

Kim, Heeringa, and Solenberger (2006) (Cont.)

- Optimization problem
  - Minimizing the variance form (1) of H-T estimator under the superpopulation model (called OP1 (Kim *et al.,* 2008))

$$Minimize\left[\sum_{i=1}^{N}\sum_{j>i}^{N}\frac{\alpha+\beta\left(x_i+x_j\right)}{x_i x_j}\sum_{i,j\in s}p_{\xi}(s)\right]$$

  - Minimizing the variance form (2) of H-T estimator under the superpopulation model (called OP2 (Kim *et al.,* 2008))

$$Minimize\left[\sum_{i=1}^{N}\sum_{j>i}^{N}\left(x_j^{-1}-x_i^{-1}\right)\left(\alpha x_i^{-1}+\beta\right)\sum_{i,j\in s}p_{\xi}(s)\right]$$

Hong, Park, Kim, Ahn, and Heeringa (2009)

- Superpopulation model

$$y_i = \alpha + \beta x_i + \varepsilon_i, \ \ i = 1, \cdots, N,$$

$$\begin{aligned} where \ \ & E_\xi\left(\varepsilon_i \mid x_i\right) = 0, \\ & V_\xi\left(\varepsilon_i \mid x_i\right) = \delta x_i^\gamma \ \ \left(\delta > 0, \ \ \gamma \geq 0\right), \\ & E_\xi\left(\varepsilon_i \varepsilon_j \mid x_i x_j\right) = 0, \ \ i \neq j \end{aligned}$$

Hong, Park, Kim, Ahn, and Heeringa (2009) (Cont.)

- Optimization problem
  - OP3: Modification of OP1

$$Minimize \left[\sum_{i=1}^{N} \sum_{j>i}^{N} \frac{1}{x_i x_j} \sum_{i,j \in s} p_{\xi}(s) \right]$$

  - OP4: Modification of OP2

$$Minimize \left[\left\{ \alpha \sum_{i=1}^{N} \sum_{j>i}^{N} \left(\frac{1}{x_i x_j} - \frac{1}{x_i^2} \right) - 2\beta \sum_{i=1}^{N} \sum_{j>i}^{N} \frac{1}{x_i} \right\} \sum_{i,j \in s} p_{\xi}(s) \right]$$

$$subject\ to \quad \pi_i = \sum_{i \in s} p_{\xi}(s),\ i = 1, \cdots, N,$$

$$c\pi_i \pi_j \le \sum_{i,j \in s} p_{\xi}(s) \le \pi_i \pi_j,$$

$$where\ c\ is\ a\ real\ number\ between\ 0\ and\ 1$$

2. New Approaches

Superpopulation Model

- General Polynomial Model (Royall and Herson, 1973) :

$$y_i = \sum_{k=0}^{K} \delta_k \beta_k x_i^k + \varepsilon_i \left[v(x_i) \right]^{1/2}, \ i = 1, \cdots, N,$$

$$where \ E_\xi(\varepsilon_i \mid x_i) = 0,$$

$$V_\xi(\varepsilon_i \mid x_i) = \sigma^2,$$

$$E_\xi(\varepsilon_i \varepsilon_j \mid x_i x_j) = 0, \ i \neq j,$$

$$which \ is \ described \ as \ \xi\left(\delta_0, \cdots, \delta_K : v(x)\right),$$

$$where \ \delta_k = 0 \ or \ 1$$

Model Expectation

- Anticipated Variance
  - Under the variance form (1) of H-T estimator

$$E_{\xi}\left[Var\left(\hat{Y}_{HT}\right)\right] = \sum_{i=1}^{N}\left(\frac{X}{nx_i}-1\right)\left\{\sigma^2 v\left(x_i\right)+\mathbf{x}_i'\boldsymbol{\beta}\boldsymbol{\beta}'\mathbf{x}_i\right\}$$

$$+2\frac{X^2}{n}\sum_{i=1}^{N}\sum_{j>i}^{N}\frac{\mathbf{x}_i'\boldsymbol{\beta}\boldsymbol{\beta}'\mathbf{x}_j}{x_i x_j}\pi_{ij} - 2\sum_{i=1}^{N}\sum_{j>i}^{N}\mathbf{x}_i'\boldsymbol{\beta}\boldsymbol{\beta}'\mathbf{x}_j$$

$$, where\ \mathbf{x}_i' = \begin{bmatrix}1 & x_i & x_i^2 & \cdots & x_i^K\end{bmatrix}'$$

$$\boldsymbol{\beta}' = \begin{bmatrix}\delta_0\beta_0 & \delta_1\beta_1 & \delta_2\beta_2 & \cdots & \delta_K\beta_K\end{bmatrix}'$$

Model Expectation (Cont.)

– Under the variance form (2) of H-T estimator

$$E_{\xi}\left[Var\left(\hat{Y}_{HT}\right)\right]=\frac{\sigma^2}{n}\sum_{i=1}^{N}v\left(x_i\right)\frac{\sum_{i=1}^{N}x_i}{x_i}\left(1-\frac{nx_i}{\sum_{i=1}^{N}x_i}\right)$$

$$+2\sum_{i=1}^{N}\sum_{j>i}^{N}x_j\mathbf{x}_i'\boldsymbol{\beta}\boldsymbol{\beta}'\left\{\frac{\mathbf{x}_i}{x_i}-\frac{\mathbf{x}_j}{x_j}\right\}$$

$$+2\frac{X^2}{n^2}\sum_{i=1}^{N}\sum_{j>i}^{N}x_i^{-1}\mathbf{x}_i'\boldsymbol{\beta}\boldsymbol{\beta}'\left\{\frac{\mathbf{x}_j}{x_j}-\frac{\mathbf{x}_i}{x_i}\right\}\pi_{ij}$$

Optimization Problems (OP) under the General Polynomial Model

- New OP_A :

$$Minimize\left[\sum_{i=1}^{N}\sum_{j>i}^{N}\frac{\mathbf{x}_i'\boldsymbol{\beta}\boldsymbol{\beta}'\mathbf{x}_j}{x_i x_j}\sum_{i,j\in s}p_\xi(s)\right]$$

- New OP_B :

$$Minimize\left[\sum_{i=1}^{N}\sum_{j>i}^{N}x_i^{-1}\mathbf{x}_i'\boldsymbol{\beta}\boldsymbol{\beta}'\left\{\frac{\mathbf{x}_j}{x_j}-\frac{\mathbf{x}_i}{x_i}\right\}\sum_{i,j\in s}p_\xi(s)\right]$$

$$subject\ to\quad \pi_i=\sum_{i\in s}p_\xi(s),\ i=1,\cdots,N,$$

$$c\pi_i\pi_j\le\sum_{i,j\in s}p_\xi(s)\le\pi_i\pi_j,$$

$$where\ c\ is\ a\ real\ number\ between\ 0\ and\ 1$$

3. Illustrations

Natural Populations

- Rao and Bayless (1969)

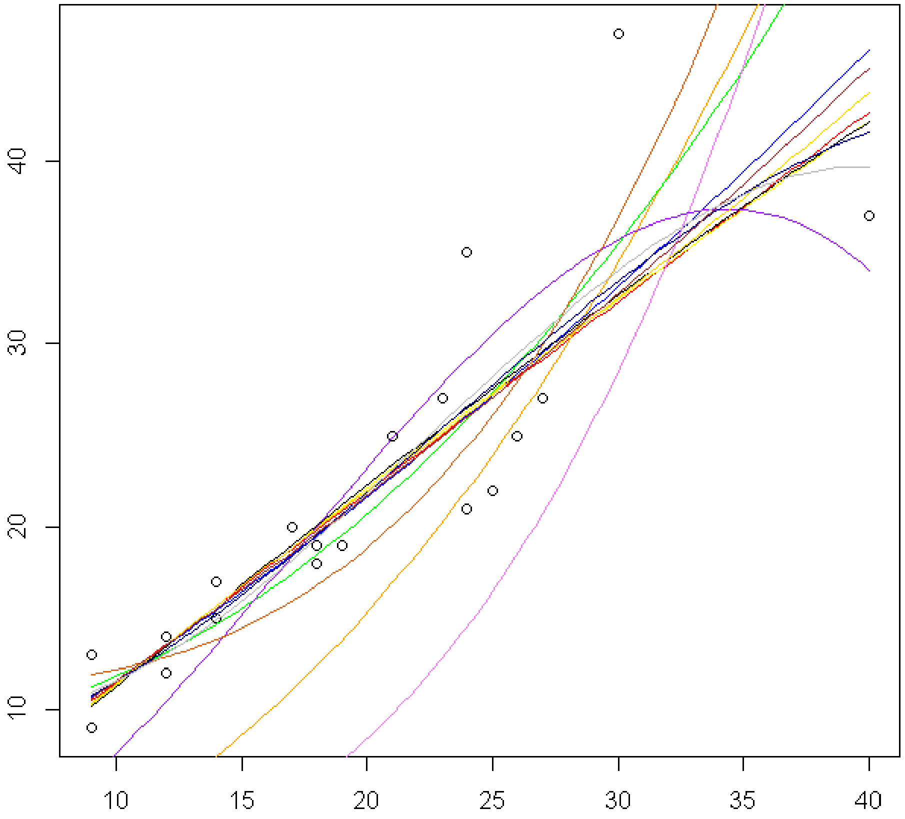
No.	Source	y	x	N
1	Horvitz and Thompson (1952)	No. of Households	Eye-estimated No. of Households	20
2	Sampford (1962)	Oats acreage in 1957	Oats acreage in 1947	35
3	Sukhatme (1954)	Wheat acreage	No. of villages	9

Superpopulation Models

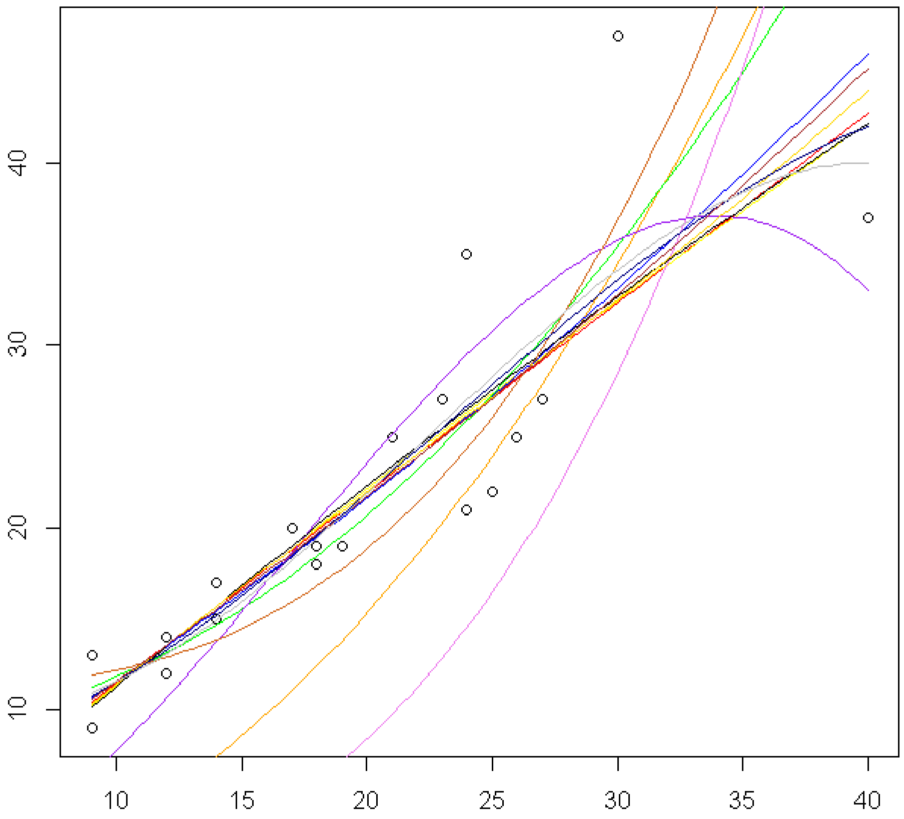
No.	Symbol	Model	No.	Symbol	Model
1	$\xi(1,1:v(x_i))$	$y_i = \beta_0 + \beta_1 x_i$	8	$\xi(0,1,0,1:v(x_i))$	$y_i = \beta_1 x_i + \beta_3 x_i^3$
2	$\xi(0,0,1:v(x_i))$	$y_i = \beta_2 x_i^2$	9	$\xi(0,1,1,1:v(x_i))$	$y_i = \beta_1 x_i + \beta_2 x_i^2 + \beta_3 x_i^3$
3	$\xi(0,1,1:v(x_i))$	$y_i = \beta_1 x_i + \beta_2 x_i^2$	10	$\xi(1,0,0,1:v(x_i))$	$y_i = \beta_0 + \beta_3 x_i^3$
4	$\xi(1,0,1:v(x_i))$	$y_i = \beta_0 + \beta_2 x_i^2$	11	$\xi(1,0,1,1:v(x_i))$	$y_i = \beta_0 + \beta_2 x_i^2 + \beta_3 x_i^3$
5	$\xi(1,1,1:v(x_i))$	$y_i = \beta_0 + \beta_1 x_i + \beta_2 x_i^2$	12	$\xi(1,1,0,1:v(x_i))$	$y_i = \beta_0 + \beta_1 x_i + \beta_3 x_i^3$
6	$\xi(0,0,0,1:v(x_i))$	$y_i = \beta_3 x_i^3$	13	$\xi(1,1,1,1:v(x_i))$	$y_i = \beta_0 + \beta_1 x_i + \beta_2 x_i^2 + \beta_3 x_i^3$
7	$\xi(0,0,1,1:v(x_i))$	$y_i = \beta_2 x_i^2 + \beta_3 x_i^3$			

Superpopulation Model Plots

- Population1:



ML



REML

Relative Efficiencies

Population1:

Using ML

ML	New OPA						New OPB					
Model	c						c					
	0	0.1	0.2	0.3	0.4	0.5	0	0.1	0.2	0.3	0.4	0.5
SM1	112.635	123.711	109.571	113.482	107.064	108.649	96.834	97.149	101.175	103.502	107.099	106.747
SM2	103.42	101.483	107.768	107.313	106.577	108.049	113.752	114.5	116.24	109.673	106.891	109.859
SM3	112.635	123.711	109.571	113.482	107.064	108.649	100.937	107.479	98.834	98.744	112.256	104.222
SM4	116.421	123.711	106.607	114.7	100.294	104.616	115.544	116.725	116.518	111.255	108.828	109.742
SM5	112.635	123.711	109.571	113.482	107.064	108.649	113.719	116.665	112.708	113.36	108.675	110.135
SM6	101.186	106.293	111.674	107.227	106.329	108.237	114.008	118.411	115.629	109.673	107.479	109.859
SM7	118.646	108.602	105.333	103.04	102.468	105.849	106.238	102.674	110.965	105.599	104.223	107.54
SM8	112.635	123.711	109.571	113.482	107.064	108.649	96.696	104.96	95.912	104.045	101.479	104.574
SM9	112.635	123.711	109.571	113.482	107.064	108.649	124.648	107.636	109.217	113.262	106.852	106.526
SM10	105.965	100.529	110.08	104.825	104.392	106.096	113.805	113.793	116.171	111.31	108.85	109.652
SM11	112.635	123.711	109.571	113.482	107.064	108.649	115.842	104.184	104.336	107.895	106.205	107.448
SM12	112.635	123.711	109.571	113.482	107.064	108.649	120.823	120.204	116.392	111.813	111.668	110.536
SM13	112.739	123.711	109.571	113.482	107.064	108.649	104.102	100.73	100.205	107.563	106.16	109.227
MEAN	111.294	117.7158	109.0792	111.1508	105.8902	107.8492	110.5345	109.6238	108.7925	108.2842	107.4358	108.159

Midzuno	Sampford	Murthy	PPS_with
50.64523	107.841	107.1007	100

Relative Efficiencies (Cont.)

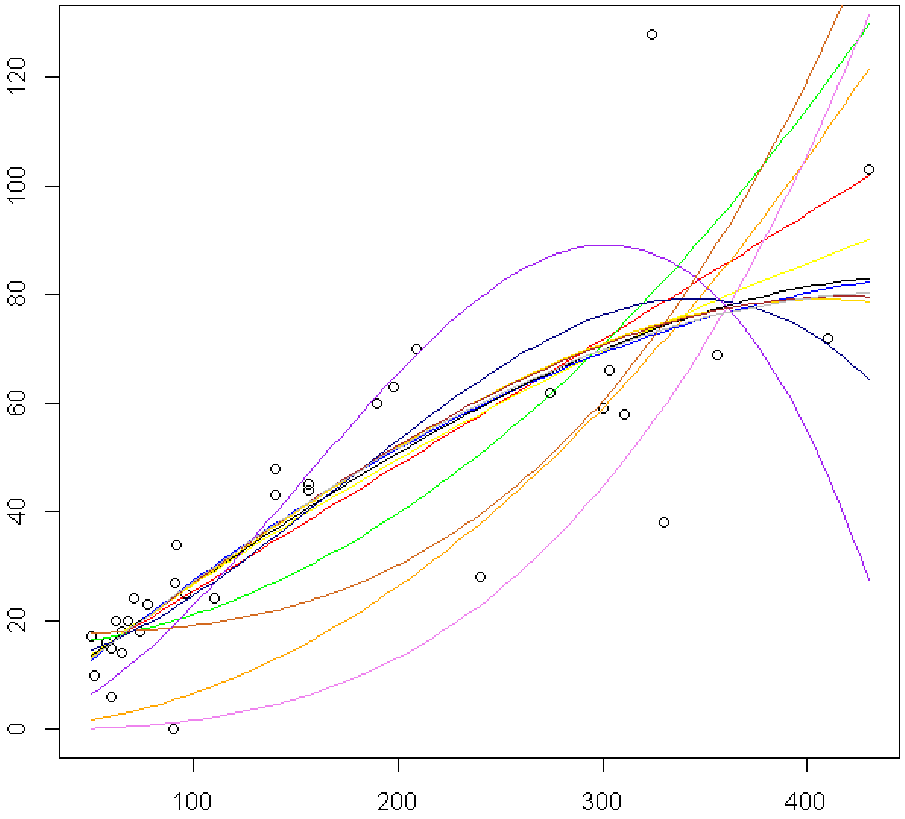
Using REML

REML	New OP_A						New OP_B					
Model	c						c					
	0	0.1	0.2	0.3	0.4	0.5	0	0.1	0.2	0.3	0.4	0.5
SM1	112.635	123.711	109.571	113.482	107.064	108.649	95.855	100.899	94.062	106.264	107.627	105.076
SM2	103.42	101.483	107.768	107.313	106.577	108.049	113.752	114.5	116.24	109.673	106.891	109.859
SM3	112.635	123.711	109.571	113.482	107.064	108.649	100.937	102.734	98.92	98.744	112.213	104.222
SM4	116.421	123.711	107.283	114.7	100.294	104.616	112.286	116.563	116.715	110.705	106.784	109.382
SM5	112.635	123.711	109.571	113.482	107.064	108.649	118.161	114.804	112.805	112.053	108.703	109.393
SM6	101.186	106.293	111.674	107.227	106.329	108.237	114.008	118.411	115.629	109.673	107.479	109.859
SM7	122.074	103.489	103.115	99.795	105.788	104.721	107.919	112.652	111.192	109.498	104.857	107.508
SM8	112.635	123.711	109.571	113.482	107.064	108.649	98.094	100.462	104.498	109.141	106.135	107.889
SM9	112.635	123.711	109.571	113.482	107.064	108.649	112.475	121.045	112.801	112.577	105.818	108.344
SM10	107.479	103.438	105.663	109.54	103.835	104.081	113.713	113.794	116.171	110.358	108.514	109.588
SM11	112.635	123.711	109.571	113.482	107.064	108.649	112.85	101.421	114.827	108.23	107.352	107.633
SM12	112.635	123.711	109.571	113.482	107.064	108.649	111.293	115.378	120.72	110.911	109.164	109.457
SM13	112.739	123.711	109.571	113.482	107.064	108.649	101.154	100.73	100.9	107.611	106.16	109.227
MEAN	111.6742	117.5463	108.6208	111.2639	106.1027	107.6074	108.6536	110.261	110.4215	108.8798	107.5152	108.2644

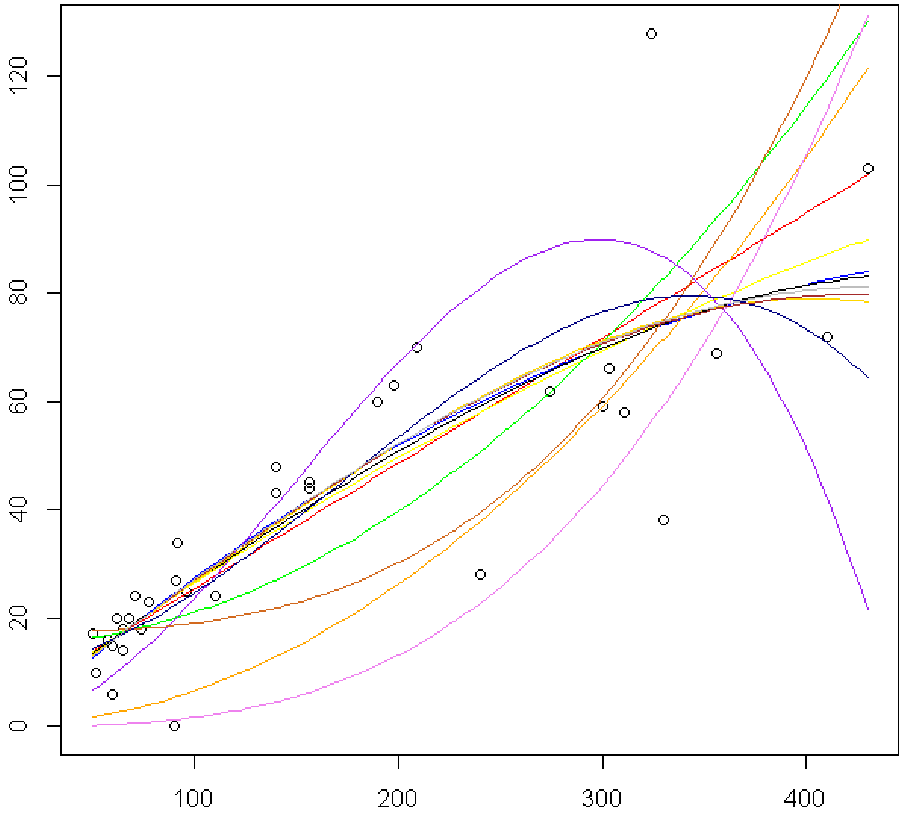
Midzuno	Sampford	Murthy	PPS_with
50.64523	107.841	107.1007	100

Superpopulation Model Plots

- Population2:



ML



REML

Relative Efficiencies

Population2:

Using ML

ML	New OP_A						New OP_B					
Model	c						c					
	0	0.1	0.2	0.3	0.4	0.5	0	0.1	0.2	0.3	0.4	0.5
SM1	104.525	99.49	93.926	104.107	103.616	104.241	109.008	110.068	106.622	102.984	100.559	104.595
SM2	106.546	109.913	108.21	99.928	104.154	104.214	98.385	98.25	99.253	102.308	104.26	103.827
SM3	106.385	99.49	93.926	104.107	103.616	104.241	109.276	109.918	110.123	107.469	103.114	104.889
SM4	99.665	104.235	100.071	104.107	103.616	104.241	100.097	98.388	98.872	101.785	101.709	103.898
SM5	106.402	99.49	94.642	104.107	103.616	104.241	110.704	109.314	106.007	106.326	108.501	104.303
SM6	111.556	110.086	110.491	106.894	105.125	104.453	107.724	105.953	98.764	106.064	104.904	104.544
SM7	106.466	106.534	104.636	105.548	108.243	105.152	101.67	100.803	104.103	102.919	105.22	103.929
SM8	106.028	99.49	93.926	104.107	103.616	104.241	110.767	109.1	108.49	106.681	103.536	104.174
SM9	106.84	99.49	93.926	104.107	103.616	104.241	109.915	108.691	105.116	107.672	105.657	104.538
SM10	102.021	106.571	100.626	102.752	103.821	104.254	98.515	97.506	100.062	102.415	103.113	103.816
SM11	107.097	100.464	100.255	101.319	99.987	104.787	103.721	102.246	105.819	102.295	104.547	105.171
SM12	107.432	99.49	93.926	104.107	103.616	104.241	109.56	107.518	105.418	109.739	107.788	104.362
SM13	107.775	99.49	93.926	104.107	103.616	104.241	110.957	109.054	110.701	109.626	105.996	104.418
MEAN	106.0568	102.6333	98.65285	103.7921	103.866	104.3683	106.1768	105.1392	104.5654	105.2525	104.5311	104.3434

Midzuno	Sampford	Murthy	PPS_with
67.63793	104.3026	104.4578	100

Relative Efficiencies (Cont.)

Using REML

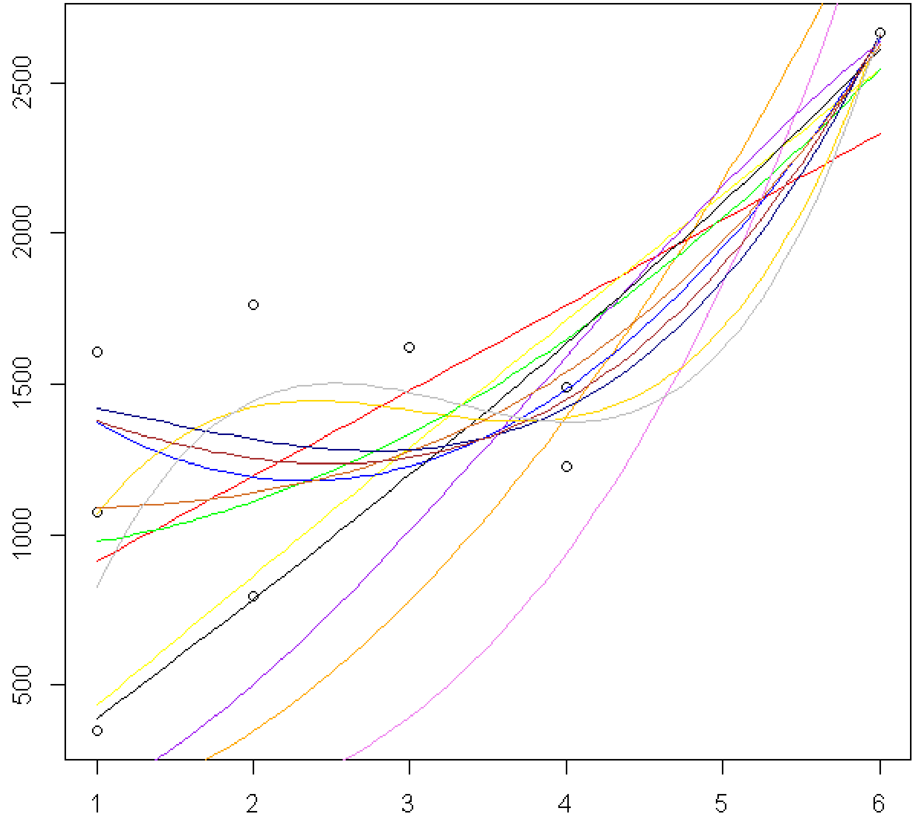
REML	New OP_A						New OP_B					
Model	c						c					
	0	0.1	0.2	0.3	0.4	0.5	0	0.1	0.2	0.3	0.4	0.5
SM1	104.525	99.49	93.926	104.107	103.616	104.241	109.92	109.425	104.016	98.03	101.383	104.205
SM2	106.546	109.913	108.21	99.928	104.154	104.214	98.385	98.25	99.253	102.308	104.26	103.827
SM3	106.385	99.49	93.926	104.107	103.616	104.241	110.282	107.855	109.134	103.099	106.84	104.66
SM4	101.034	95.285	94.075	104.107	103.616	104.241	99.864	98.704	99.161	101.719	101.749	103.904
SM5	106.402	99.49	94.642	104.107	103.616	104.241	110.996	109.546	106.252	106.259	106.447	104.201
SM6	111.556	110.086	110.491	106.894	105.125	104.453	107.724	105.953	98.764	106.064	104.904	104.544
SM7	105.695	102.52	106.83	106.845	107.639	103.062	101.542	102.943	105.053	103.066	104.662	105.388
SM8	106.028	99.49	93.926	104.107	103.616	104.241	110.767	109.934	108.49	106.681	103.712	104.174
SM9	99.688	99.196	101.569	97.454	105.356	104.787	111.259	110.085	105.85	106.353	104.334	104.459
SM10	102.191	111.506	100.243	102.752	103.821	104.254	98.485	97.54	100.147	102.375	103.34	103.856
SM11	104.759	100.464	100.255	101.319	99.987	104.787	107.936	106.039	103.672	103.28	104.315	104.505
SM12	107.432	99.49	93.926	104.107	103.616	104.241	108.542	109.452	108.561	110.401	107.608	104.391
SM13	107.775	99.49	93.926	104.107	103.616	104.241	111.48	109.82	105.811	106.397	108.199	104.162
MEAN	105.3858	101.9931	98.91885	103.3801	103.9534	104.2495	106.7063	105.8112	104.1665	104.3102	104.7502	104.3289

Midzuno	Sampford	Murthy	PPS_with
67.63793	104.3026	104.4578	100

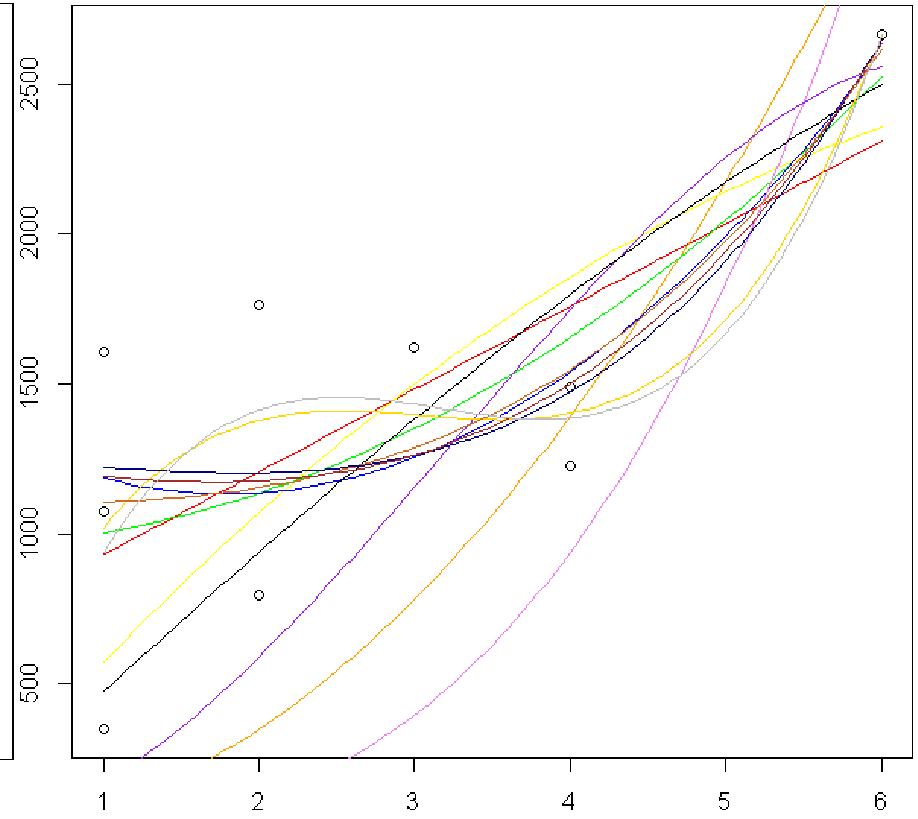
dongguk UNIVERSITY

Superpopulation Model Plots

- Population3:



ML



REML

Relative Efficiencies

Population3:

Using ML

ML	New OP_A						New OP_B					
Model	c						c					
	0	0.1	0.2	0.3	0.4	0.5	0	0.1	0.2	0.3	0.4	0.5
SM1	126.693	120.347	113.556	109.155	106.593	105.937	129.166	124.831	118.496	110.9	107.284	106.197
SM2	125.536	119.128	115.166	108.96	106.593	105.937	95.008	99.255	98.969	99.625	103.604	105.632
SM3	94.975	114.057	112.001	100.987	109.878	106.796	127.591	123.985	115.807	110.226	107.884	106.197
SM4	132.63	125.71	122.7	120.004	114.825	106.921	133.034	126.389	122.091	117.968	110.777	108.435
SM5	121.848	123.083	117.195	115.893	112.914	107.432	126.243	122.188	119.22	118.479	111.02	107.226
SM6	125.517	119.586	113.278	109.837	106.593	105.937	99.145	100.544	97.746	99.433	102.102	105.632
SM7	126.53	118.713	114.433	110.098	106.593	105.937	93.442	98.297	98.702	99.746	103.432	105.855
SM8	124.115	118.814	114.087	109.256	106.32	105.937	95.008	94.168	98.684	99.346	103.432	105.855
SM9	132.095	125.71	123.934	116.497	114.325	108.27	131.769	130.67	123.939	121.321	113.468	107.749
SM10	120.969	118.309	120.132	117.38	112.749	107.27	126.118	125.101	120.217	117.594	111.081	107.065
SM11	120.248	120.911	121.827	114.042	112.072	107.27	126.364	125.29	122.451	117.349	115.077	106.966
SM12	120.969	123.083	119.988	115.893	114.823	107.432	126.243	120.79	119.992	117.741	111.22	107.065
SM13	132.63	130.501	123.934	117.968	114.325	108.27	134.733	128.477	126.245	119.839	112.072	107.929
MEAN	123.4427	121.3809	117.8639	112.7669	110.6618	106.8728	118.7588	116.9219	114.043	111.5052	108.6502	106.7541

Midzuno	Sampford	Murthy	PPS_with
305.457	107.6236	114.3764	100

Relative Efficiencies (Cont.)

Using REML

REML	New OP_A						New OP_B					
Model	c						c					
	0	0.1	0.2	0.3	0.4	0.5	0	0.1	0.2	0.3	0.4	0.5
SM1	126.693	120.347	115.091	109.775	106.593	105.937	129.166	124.2	118.299	111.577	107.284	106.168
SM2	125.536	119.128	115.166	108.96	106.593	105.937	95.008	99.255	98.969	99.625	103.604	105.632
SM3	126.092	120.366	112.977	109.256	106.593	105.937	129.506	124.757	117.144	110.575	107.118	106.384
SM4	132.63	125.71	122.7	117.063	114.193	106.921	131.242	126.357	123.938	118.626	112.774	108.479
SM5	121.848	120.716	120.132	120.993	113.043	107.27	126.243	125.29	122.451	119.435	111.22	107.359
SM6	125.517	119.586	113.278	109.837	106.593	105.937	99.145	100.544	97.746	99.433	102.102	105.632
SM7	132.095	127.738	123.316	121	113.005	107.749	96.134	102.087	98.485	97.802	102.977	105.855
SM8	115.314	118.814	114.087	109.256	106.32	105.937	127.957	122.039	117.336	110.776	106.863	106.197
SM9	132.095	125.71	125.568	116.497	114.325	108.27	134.733	128.444	126.245	119.64	111.445	106.921
SM10	121.848	118.309	120.132	117.38	113.043	107.27	126.489	120.614	120.217	117.261	111.486	107.397
SM11	120.969	118.309	119.988	120.758	113.043	107.27	126.243	125.29	122.451	119.435	111.081	107.226
SM12	121.848	118.309	120.211	113.457	113.043	107.27	126.489	122.665	122.451	119.435	111.22	107.036
SM13	132.095	125.71	125.568	117.968	114.325	108.27	134.733	128.444	123.938	116.494	112.072	107.402
MEAN	125.7369	121.4425	119.0934	114.7846	110.824	106.9212	121.776	119.2297	116.1285	112.3165	108.5574	106.7452

Midzuno	Sampford	Murthy	PPS_with
305.457	107.6236	114.3764	100

4. Conclusions

Conclusions

- The polynomial superpopulation models from 1st to 3rd order (except the simple linear model without intercept) were used to investigate the efficiencies of model-optimized sampling
- There are not much different in relative efficiencies (RE) between the model estimation methods, ML and REML
- It seems that the REs are reduced as the constant for the variance stability, c, increases

Conclusions (Cont.)

- REs are different according to the superpopulation models in new OP_B rather than new OP_A, but the variations of them tend to be stable as c increases
- Because of these fluctuation of REs, we may choose the superpopulation models that give higher REs than Murthy's method, but also be careful of choosing them
- Considering the REs and the stability of them, the appropriate value of c = 0.3 might be good

Further Research

- The empirical studies for more natural population models might be needed
- Studies about below items would be interesting
  - The stability of the variance estimator
  - The relationship between REs and the estimated parameters of superpopulation models
  - The adoption of nonlinear superpopulation model
  - The efficiencies in the larger sample size
  - The comparison of the efficiencies between the H-T estimator and the GREG estimator in the model-optimized sampling

Contact Information

Sung-Joon Hong

hsj8129@dongguk.edu

Sun-Woong Kim

sunwk@dongguk.edu

dongguk UNIVERSITY