Empirical Bayes Estimation in Small Area

Ferry Butar Butar

Sam Houston State University
Huntsville, Texas, USA

Yogyakarta 2018, May 7

- The term small area or domain typically refers to a population for which reliable statistics of interest cannot be produced due to certain limitations of the available data.
- A small geographical area such as a county, municipality or a census division.
- Could be a " small domain ", i.e. a subgroup of a population such as a specific age, gender, race group of people within a large geographical area.
- Here in Indonesia could be Kecamatan, Kabupaten

- *Provide reliable estimators of totals and means for large areas or domains.*
- *Direct survey estimators for a small area, based on data only from the sample units in the area, yield large standard errors due to unduly small size of the sample in the area.*
- *The reason behind this is that the original survey was designed to produce accuracy at a much higher level of aggregation than that for local areas.*

Demand for Small Area Statistics: From both Public and Private sectors

Federal Government:

- *need for Small Area Income and Poverty Estimates*
- *need in regional planning and in allocation of government resources*
- *distribution*
- *equity*
- *disparity*

Sam Houston STATE UNIVERSITY A Member of The Texas State University System

- Estimation of per capita income (PCI) for several small places,
- Estimates of median income of four-person families for 50 states and DC,
- EBLUP of areas under corn and soybeans,
- Cressie (1992) proposed EB method in the census undercount,
- Estimation of prob. of false fire alarms reported by many street boxes of NYC,
- Estimation of Cancer Mortality rates.

Sam Houston STATE UNIVERSITY A Member of The Texas State University System

- Let y_{ij} = characteristics of interest for the j th unit in the i th local area $(j = 1, ..., N_i;\ i = 1, ..., m)$
- Let $y_{ij}\ (j = 1, ..., n_i)$ denote the characteristics corresponding to the n_i sample units.
- Population Mean: $\overline{Y}_i = N_i^{-1} \sum_{j=1}^{N_i} y_{ij} \quad (i = 1, ..., m)$

Direct Estimate: (sample mean)

$$\overline{y}_{is} = n_i^{-1} \sum_{j=1}^{n_i} y_{ij} \quad (i = 1, ..., m)$$

- *design unbiased*
- *large variance*

Synthetic Estimates:

$$\overline{y}_s = \frac{\sum_{i=1}^m n_i \overline{y}_{is}}{\sum_{i=1}^m n_i}$$

- *design biased*
- *small variance*

Composite Estimates:

$$w_i \overline{y}_{is} + (1 - w_i)\overline{y}_s$$

how to choose the weights w_i?
both *design* and *model* -based approaches have been proposed
often $w_i = \frac{n_i}{N_i}$ or $\frac{\sum n_i}{\sum N_i}$

Sam Houston STATE UNIVERSITY A Member of The Texas State University System

Suppose the parameter of interest is the small area means θ_i $(i = 1, \cdots, m)$.

- *Direct Sample Survey Estimates:*

$$\hat{\theta}_i = \sum_{j=1}^{n_i} w_{ij} y_{ij}$$

- *Synthetic Estimates:*

$$\hat{\theta}_i^s = x_i' \hat{\beta}$$

- *Composite estimates:*

$$\hat{\theta}_i^c = a_i \hat{\theta}_i + (1 - a_i) \hat{\theta}_i^s$$

Question: *How does one choose the weights a_i?*

- Ad Hoc Methods
- Empirical Best Linear Unbiased Prediction (EBLUP) Method
- Empirical Bayes (EB) Method
- Hierarchical Bayes (HB)

Two Level Model

For $i = 1, \cdots, m$, assume

Level 1: (Sampling Model) $Y_i \mid v_i \overset{ind}{\sim} N[x_i'\beta + v_i, D_i]$;

Level 2: (Linking Model) $v_i \overset{ind}{\sim} N[0, \psi]$.

- The hyper-parameters β and ψ are unknown,
- The sampling variances D_i's are assumed to be known.

Linear Mixed Model

$$Y_i = x_i'\beta + v_i + e_i, \qquad i = 1, \cdots, m,$$

where $e_i \overset{ind}{\sim} N(0, D_i)$ and $v_i \overset{iid}{\sim} N(0, \psi)$.

An Example (Efron and Morris, 1975)

- p_i : batting average of player i for the first $n = 45$ at-bats ($i = 1, \cdots, m$).
- π_i: the true season batting average.
- $np_i \overset{ind}{\sim} Bin(n, \pi_i), i = 1, \cdots, 18.$
- $Y_i = \sqrt{n}\, arcsin(2p_i - 1), \theta_i = \sqrt{n}\, arcsin(2\pi_i - 1),$
- The Efron-Morris model:

$$Y_i|\theta_i \overset{ind}{\sim} N[\theta_i, 1], \theta_i \overset{iid}{\sim} N[\mu, A]$$

- x_i : a covariate obtained from the previous batting average (Jiang and Lahiri, 2006)

Sam Houston STATE UNIVERSITY A Member of The Texas State University System

An Example (Fay-Herriot Model, 1979)

- $y_i = \theta_i + e_i$, $i = 1, \ldots, m$,
- $\theta_i = x_i'\beta + v_i$,
- $v_i \overset{ind}{\sim} N(0, D_i)$
- $e_i \overset{ind}{\sim} N(0, A)$

where y_i is the survey estimator of per-capita income for the ith area, D_i is known the sampling variance of y_i; $x_i = (x_{i1}, x_{i2}, ..., x_{ik})'$ is a vector of known benchmark variables obtainable from the 1969 tax return data and 1970 census data.

The BP, BLUP, EBLUP of θ_i

- The best predictor (BP) of θ_i:

$$\hat{\theta}_i^{BP} = (1 - B_i)y_i + B_i x_i^{'} \beta,$$

where $\xi = (\beta, \psi)^T$ and $B_i = \frac{D_i}{\psi + D_i}, \; i = 1, \cdots, m.$

- The best linear unbiased prediction (BLUP) of θ_i:

$$\hat{\theta}_i^{BLUP} = (1 - B_i)y_i + B_i x_i' \hat{\beta}(\psi),$$

where $\hat{\beta}(\psi) = (X'\Sigma^{-1}X)^{-1}X'\Sigma^{-1}Y$,

- An empirical BLUP (EBLUP) of θ_i:

$$\hat{\theta}_i^{EBLUP} = (1 - \hat{B}_i)y_i + \hat{B}_i x_i^{'} \hat{\beta}(\hat{\psi})$$

where $\hat{B}_i = \frac{D_i}{\hat{\psi} + D_i}$, ψ is a consistent estimator of ψ, and $\hat{\xi} = (\hat{\beta}(\hat{\psi}), \hat{\psi})'$

- Morris (1983)
- Laird and Louis 1987
- Prasad and Rao 1990
- Lahiri and Rao 1995
- among others

$$MSE[\hat{\theta}_i^{EBLUP}] \stackrel{\text{def}}{=} E[\hat{\theta}_i^{EBLUP} - \theta_i]^2$$

$$= E(\theta_i - \hat{\theta}_i^{BP})^2 + E(\hat{\theta}_i^{BP} - \hat{\theta}_i^{BLUP})^2$$

$$+E(\hat{\theta}_i^{BLUP} - \hat{\theta}_i^{EBLUP})^2$$

$$= g_{1i}(\psi) + g_{2i}(\psi) + E[\hat{\theta}_i^{EBLUP} - \hat{\theta}_i^{BLUP}]^2$$

- Let $mse[\hat{\theta}_i(\hat{\psi})]$ be an estimator of $MSE[\hat{\theta}_i(\hat{\psi})]$.
- We shall seek an estimator for which

$$E[mse(\hat{\theta}_i(\hat{\psi})] = MSE[\hat{\theta}_i(\hat{\psi})] + o(m^{-1}),$$

For Fay Herriot Model

- $g_{1i}(\psi) = (1 - B_i)D_i$
- $g_{2i}(\psi) = B_i^2 x_i'(\sum D_i^{-1} B_i x_i x_i')^{-1} x_i$

Sam Houston STATE UNIVERSITY A Member of The Texas State University System

- note: $E[g_{1i}(\hat{A}) + g_{2i}(\hat{A})] = g_{1i}(A) + g_{2i}(A) + 0(m^{-1})$
- note: $E[g_{1i}(\hat{A})] = g_{1i}(A) - g_{3i}(A) + o(m^{-1})$,
- where $g_{3i}(A) = \frac{\hat{B}_i^3}{D_i} \frac{2}{m^2} \Sigma \hat{B}_i^{-2} D_i^2$

We shall propose a parametric bootstrap method to correct for the $0(m^{-1})$ bias of $g_{1i}(\hat{A}) + g_{2i}(\hat{A})$ and to estimate $E[\hat{\theta}_i^{EBLUP}(\hat{A}) - \hat{\theta}_i^{BLUP}(A)]^2$.

Remark:

- $g_{1i}(A)$ is the measure of uncertainty of the Bayes estimator $\hat{\theta}_i^B$,
- $g_{2i}(A)$ is the uncertainty due to estimation of β and
- the third term is due to the estimation of unknown A.
- One may naively approximate $MSE(\hat{\theta}_i^{EBLUP})$ by

$$g_{1i}(A) + g_{2i}(A)$$

which ignores the uncertainty due to estimation of A.

It can be shown that,

$$E(\hat{\theta}_i^{EBLUP} - \theta_i^{BLUP})^2 = g_{3i}(\psi) \; + \; o(m^{-1}),$$

Thus, the naive approximation, *i.e.*,

$$g_{1i}(\psi) \; + \; g_{2i}(\psi),$$

could lead to a serious underestimation since $g_{3i}(\psi)$ is of order $O(m^{-1})$, same as the order of $g_{2i}(\psi)$.

$$y_i^\star = x_i'\hat{\beta}(\hat{A}) + v_i^\star + e_i^\star, \; i = 1, \ldots, m,$$

- Compute $\hat{\beta} = (X'\hat{\Sigma}^{-1}X)^{-1}X'\hat{\Sigma}^{-1}Y$ and $\hat{A}$;
- Draw bootstrap sample from $v_i^\star \overset{ind}{\sim} N(0, \hat{A})$
- and $e_i^\star \overset{ind}{\sim} N(0, D_i)$

$$\widehat{MSE}[\hat{\theta}_i(A)] = g_{1i}(\hat{A}) + g_{2i}(\hat{A}) + g_{3i}(\hat{A})]$$

$$\hat{E}[\hat{\theta}_i(\hat{A}) - \hat{\theta}_i(A)]^2 = E_\star\{\hat{\theta}_i[\hat{\beta}^\star(\hat{A}^\star), \hat{A}^\star; Y_i] - \hat{\theta}_i(\hat{A})\}^2$$

- Compute $\hat{\beta}^*$ and $\hat{\psi}^*$ from Y^*. Then we have $\hat{\theta}_i^{\text{EBLUP}*} = (1 - \hat{B}^*)Y_i^* + \hat{B}^* x_i'\hat{\beta}^*$, and $g_{1i}(\hat{A}^*) = \frac{A^* D_i}{A^* + D_i}$;

Let

$$mse_{boot}[\hat{\theta}_i(\hat{A})] = \widehat{MSE}[\hat{\theta}_i(A)] + \widehat{E}[\hat{\theta}_i(\hat{A}) - \hat{\theta}_i(A)]^2.$$

Then,

$$E\{mse_{boot}[\hat{\theta}_i(\hat{A})]\} = MSE[\hat{\theta}_i(\hat{A})] + o(m^{-1})$$

$$mse_{boot}^{BL}[\hat{\theta}_i(\hat{A})] = g_{1i}(\hat{A}) + g_{2i}(\hat{A}) + g_{3i}(\hat{A}) + g_{4i}(\hat{A}) + o_p(m^{-1})$$

Compare with:

$$mse^{PR}[\hat{\theta}_i(\hat{A})] = g_{1i}(\hat{A}) + g_{2i}(\hat{A}) + 2g_{3i}(\hat{A})$$

where $g_{4i}(A) = 2\frac{\hat{B}_i^4}{D_i^2}\frac{1}{m^2}\left(\sum D_i^2 \hat{B}_i^{-2}\right)(y_i - x_i'\hat{\beta})^2$

A very special case:

- $y_i = \mu + v_i + e_i, i = 1, ..., m$
- $v_i \overset{iid}{\sim} N(0, A)$
- $e_i \overset{iid}{\sim} N(0, D)$,
- u_i and e_i are independent

Prediction

To predict $\theta_i = \mu + v_i$

$$\hat{\theta}_i^{EBLUP} = \hat{B}_1 \overline{Y} + (1 - \hat{B}_1) Y_i,$$

where

- $\overline{Y} = \frac{1}{m} \sum Y_i$,
- $\hat{B}_1 = min(1.0, \hat{B})$ and
- $\hat{B} = \frac{D (m - 1)}{\sum (Y_i - \overline{Y})^2}$.

- $V_i^{M} = (1 - \hat{B}_1)\, D + \frac{D\ \hat{B}_1}{m} + \frac{2\ \hat{B}_1^2}{m - 3}\,(y_i - \overline{y})^2.$

- $V_i^{PR} = (1 - \hat{B}_1)\, D + \frac{D\ \hat{B}_1}{m} + \frac{4\ D\ \hat{B}_1}{m}.$

- $V_i^{BL} = (1 - \hat{B}_1)D + \frac{2D\hat{B}_1}{m} + \frac{m-3}{m-1}\,\frac{B_1}{m} + \frac{2\ \hat{B}_1^2}{m-5}(y_i - \overline{y})^2,$

- $V_i^{LL} = (1 - \hat{B}_1)\, D + \frac{m-1}{m-5}\,\frac{D\ \hat{B}_1}{m} + \frac{2\ \hat{B}_1^2}{m - 5}\,(y_i - \overline{y})^2.$

Sam Houston
STATE UNIVERSITY
A Member of The Texas State University System

Remark:

- V_i^{BL} and V_i^M are identical upto order $o_p(m^{-1})$.
- the difference between V_i^M and V_i^{LL} is of order $O_p(m^{-1})$.
- The Prasad-Rao measure V_i^{PR} cannot match a hierarchical Bayes solution
- The PR method gives exactly the same measure of uncertainty for all the small-areas
- Since $E\, g_{4i}(\hat{\psi}; Y_i) \;=\; g_{3i}(\psi) \;+\; o(m^{-1})$, it is quite possible that for some i, V_i^{LL} could give us a measure which is less than the naive measure V_i^N (at least for large m).

Sam Houston STATE UNIVERSITY A Member of The Texas State University System

$$\mathbf{Y}_n = \mathbf{X}\beta + \mathbf{Z}\mathbf{v}_q + \mathbf{e}_n$$

where

- $\mathbf{Y}_n \in \mathbb{R}^n$ is the vector of observed data,
- $\mathbf{X}$ is a known $(n \times p)$ matrix,
- $\mathbf{Z}$ is a known $(n \times q)$ matrix,
- $\beta \in \mathbb{R}^p$ is a fixed but unknown parameter vector, and $\mathbf{v}_q \in \mathbb{R}^q$ and $\mathbf{e}_n \in \mathbb{R}^n$ are random variables following the normal distributions $N_q(0, R_n(\psi))$ and $N_n(0, D_q(\psi))$ respectively,
- $d = p + s$, where s is the dimension of ψ.

- Let $\theta = c^T(\mathbf{X}\beta + \mathbf{Z}\mathbf{v})$, where c is a fixed and known $(n \times 1)$ vector. The distribution of θ given $\mathbf{Y}$ is $N(\hat{\theta}^{BP}, g_1(\psi))$, where

$$
\begin{aligned}
\hat{\theta}^{BP} &= c^T\mathbf{X}\beta + c^T\mathbf{Z}R\mathbf{Z}^T\Sigma^{-1}(\mathbf{Y} - \mathbf{X}\beta) \\
&= c^T D\Sigma^{-1}\mathbf{X}\beta + c^T\mathbf{Z}R\mathbf{Z}^T\Sigma^{-1}\mathbf{Y}, \quad \text{and} \\
g_1(\psi) &= c^T\mathbf{Z}\left(R - R\mathbf{Z}^T\Sigma^{-1}\mathbf{Z}R\right)\mathbf{Z}^T c; \\
\Sigma &= R + \mathbf{Z}D\mathbf{Z}^T.
\end{aligned}
$$

- Replace β and ψ by $\hat{\beta}$ and $\hat{\psi}$ to obtain: $\hat{\theta}^{EBLUP}$ and $g_1(\hat{\psi})$;

We shall attempt to estimate

- $MSE[\hat{\theta}(\hat{\psi})] \stackrel{\text{def}}{=} E[\hat{\theta}(\hat{\psi}) - \theta]^2$
- Prasad and Rao and others used the following identity (up to the order $o(m^{-1})$)

$$MSE[\hat{\theta}_i(\hat{\psi})] = MSE[\hat{\theta}_i(\psi)] + E[\hat{\theta}_i(\hat{\psi} - \hat{\theta}_i(\psi)]^2$$

(Kackar and Harville, 1984), The above identity hold under normality of vi and e_i)

- Define

$$\boldsymbol{Y}^* = \boldsymbol{X}\hat{\beta} + \boldsymbol{Z}\boldsymbol{v}^* + \boldsymbol{e}^*$$

where $\boldsymbol{v}^* \sim N_q(0, R(\hat{\psi}))$ and $\boldsymbol{e}^* \sim N_n(0, D(\hat{\psi}))$.
- From $\boldsymbol{Y}^*$, obtain $\hat{\beta}^*$ and $\hat{\psi}^*$
- Obtain $\hat{\theta}^{EBLUP*}$ using $\hat{\beta}^*$ and $\hat{\psi}^*$

- Then

$$\widehat{MSE}[\hat{\theta}_i(\psi)] = g_{1i}(\hat{\psi}) + E_\star[g_{2i}(\hat{\psi}^\star) + g_{3i}(\hat{\psi})]$$

- and

$$\widehat{E}[\hat{\theta}_i(\hat{\psi}) - \hat{\theta}_i(\psi)]^2 = E_\star\{\hat{\theta}_i[\hat{\beta}^\star(\hat{\psi}^\star), \hat{\psi}^\star; Y_i] - \hat{\theta}_i(\hat{\psi})\}^2$$

- Let

$$mse_{boot}[\hat{\theta}_i(\hat{\psi})] = \widehat{MSE}[\hat{\theta}_i(\psi)] + \widehat{E}[\hat{\theta}_i(\hat{\psi}) - \hat{\theta}_i(\psi)]^2$$

- $= g_{1i}(\hat{\psi}) \; + \; m^{-1}B^T(\hat{\psi})\nabla g_{1i}(\hat{\psi}) \; + \; g_{2i}(\hat{\psi})$
  $+g_{3i}(\hat{\psi}) \; + g_{4i}(\hat{\psi}; Y_i)$
- $g_{3i}(\psi) = trace\,[L_i(\psi)\Sigma_i(\psi)L_i'(\psi)\sum(\psi)],$
- $L_i(\psi) = \; col_{1 \leq j \leq s}\, L_{ij}'(\psi),$
- $L_{ij}'(\psi) = \frac{\partial}{\partial \psi_j}\,(\lambda_i' G_i(\psi) Z_i' \Sigma_i^{-1}(\psi))$
- $\sum(\psi) = E(\hat{\psi} \; - \; \psi)(\hat{\psi} \; - \; \psi)'.$

- $g_{4i}(\hat{\psi}) = trace[L_i(\hat{\psi})\ [Y_i - X_i\hat{\beta}(\hat{\psi})][Y_i - X_i\hat{\beta}(\hat{\psi})]'L_i'(\psi)\ \sum(\hat{\psi})]$
- $g_{2i}(\hat{\psi}) = (c_i^T - X_i^T\Sigma_i^{-1}Z_iD_ic_i)^T(\sum_i^m X_i^T\Sigma_i^{-1}X_i)^{-1}(c_i^T - X_i^T\Sigma_i^{-1}Z_iD_ic_i)$

- Then, under certain regularity conditions,

$$E\{mse_{boot}[\hat{\theta}_i(\hat{\psi})]\} = MSE[\hat{\theta}_i(\hat{\psi})] + o(m^{-1})$$

- We generate $N = 10,000$ independent data sets $\{Y_i, i = 1, \cdots, m\}$ using a simplified Fay-Herriot model $Y_i = \mu + v_i + e_i$, where $v_i \sim N(0, A)$, and $e_i \sim N(0, D_i)$.
- For each simulation we found the confidence intervals
- $e_i^{EBLUP} \pm z_{.025}\sqrt{V_i^j}$,
- checked whether θ_i belonged to the confidence interval ($i = 1, ..., m$).

Pattern: $D/A = (< 1, = 1, > 1)$;
Areas: $m = (30, 50, 70, 100)$.

Table : *Average coverage and average length of different intervals (nominal coverage=0.95)*

	$D/A = 0.5$			
	$m = 30$	$m = 50$	$m = 70$	$m = 100$
N	0.937 (2.23)	0.943 (2.25)	0.945 (2.25)	0.947 (2.26)
L-L	0.944 (2.28)	0.946 (2.27)	0.947 (2.27)	0.948 (2.27)
P-R	0.943 (2.28)	0.945 (2.27)	0.947 (2.27)	0.948 (2.27)
BL	0.947 (2.32)	0.948 (2.30)	0.949 (2.29)	0.949 (2.28)

Table : *Average coverage and average length of different intervals (nominal coverage=0.95)*

	$D/A = 1.0$			
	$m = 30$	$m = 50$	$m = 70$	$m = 100$
N	0.911 (2.68)	0.929 (2.72)	0.936 (2.74)	0.941 (2.75)
L-L	0.929 (2.81)	0.937 (2.79)	0.941 (2.78)	0.944 (2.78)
P-R	0.929 (2.80)	0.937 (2.78)	0.941 (2.78)	0.944 (2.78)
BL	0.939 (2.90)	0.943 (2.84)	0.945 (2.82)	0.947 (2.81)

Sam Houston STATE UNIVERSITY A Member of The Texas State University System

Table : *Average coverage and average length of different intervals (nominal coverage=0.95)*

	$D/A = 1.5$			
	$m = 30$	$m = 50$	$m = 70$	$m = 100$
N	0.884 (2.88)	0.909 (2.93)	0.923 (2.96)	0.933 (2.99)
L-L	0.914 (3.10)	0.924 (3.05)	0.932 (3.04)	0.938 (3.04)
P-R	0.921 (3.09)	0.926 (3.05)	0.932 (3.04)	0.938 (3.04)
BL	0.933 (3.25)	0.935 (3.14)	0.938 (3.11)	0.942 (3.09)

Sam Houston
STATE UNIVERSITY
A Member of The Texas State University System

Table : Percent Average Relative Biases of MSE Estimators

	$D/A = 0.5$		$D/A = 1.0$		$D/A = 1.5$	
	$m = 30$	$m = 100$	$m = 30$	$m = 100$	$m = 30$	$m = 100$
N	-7	-2	-12	-4	-16	-6
L-L	-2	-1	-4	-2	-4	-3
P-R	-3	-1	-6	-2	-6	-3
BL	0	0	2	0	4	0

- Our proposed prediction is a bootstrap, which is different from the traditional approaches.
- Our model theory can handle many useful Mixed linear models.
- We need normality assumption. Extension of our theory for non-normal and non-linear model is in progress.
- Zero estimates of variance components pose problems in real life data analysis.

References

- Battese, G.E., Harter, R.M. and Fuller, W.A. (1988), " An error-components Model for Prediction of County Crop Areas using Survey and Satellite Data,", *Journal of the American Statistical Association*, **80**, 28 – 36.
- Butar, B.F. (1997), "Empirical Bayes Methods in Survey Sampling," Unpublished Ph.D. Thesis, Dept. of Mathematics and Statistics, University of Nebraska-Lincoln.
- Butar Butar, F., and Lahiri, P. (2003), "On Measures of Uncertainty of Empirical Bayes Small-Area Estimators,", *Journal of Statistical Planning and Inference*, **112**, 63 – 76
- Butar Butar, F. (2005)," Small Area Estimation, *Encyclopedia of Social Measurement, Academic Press, Vol* **3**, 489 – 494*.*
- *Carter, G.M., and Rolph, J.E.* (1974)*, " Empirical Bayes Methods Applied to Estimating Fire Alarm Probabilities", Journal of the American Statistical Association,* **69**, 269 – 277*.*
- *Fay, R. E., and Herriot, R.A.* (1979)*, "Estimates of Income for Small Places: an Application of James-Stein Procedure to Census Data, " Journal of American Statistical Association,* **74**, 269 – 277*.*
- *Ghosh, M., and Meeden, G.* (1986)*, "Empirical Bayes Estimation in Finite Population Sampling, "Journal of American Statistical Association,* **81**, 1058 – 1062*.*
- *Ghosh, M., and Lahiri, P.* (1987)*, "Robust Empirical Bayes Estimation of Means from Stratified Samples, "Journal of American Statistical Association,* **82**, 1153 – 1162*.*
- *Ghosh, M., Nangia, N., and Kim, D.* (1996)*, "Estimation of Median Income of Four-Person Families: A Bayesian Time Series Approach, " Journal of American Statistical Association,* **91**, 1423 – 1431*.*

- *Ghosh, M., and Rao, J.N.K.* (1994), *"Small Area Estimation: An Appraisal", Statistical Science,* **9**, *No.*1, 55 − 93, *(with discussion).*
- *Jiang, J., and Lahiri, P. (2006). Mixed model prediction and small area estimation (with discussions), Test,* **15**, *1, 1-96.*
- *Kackar, R.N., and Harville, D.A.* (1984), *"Approximations for Standard Errors of Estimators of Fixed and Random Effects in Mixed Linear Models," Journal of the American Statistical Association ,* **79**, 853 − 862*.*
- *Laird, N.M., and Louis, T.A.* (1987), *"Empirical Bayes Confidence Intervals Based on Bootstrap Samples," Journal of the American Statistical Association ,* **82**, 739 − 750*.*
- *Morris, C.N.* (1983), *"Parametric Empirical Bayes Inference: Theory and Applications," Journal of American Statistical Association,* **78**, 47 − 59*.*
- *Prasad, N.G.N., and Rao, J.N.K.* (1990), *"The Estimation of Mean Squared Errors of Small Area Estimators, "Journal of American Statistical Association,* **85**, 163 − 171*.*
- *Rao, J.N.K.* (1999), *"Some recent Advances in Model-Based Small Area Estimation," Survey Methodology,* **25**, *175-186.*
- *Rao, J.N.K. (2001), "EB and EBLUP in Small Area Estimation," in Empirical Bayes and Likelihood Inference (S. E. Ahmed and N. Reid Eds.), 33-44, Springer.*
- *Rao, J. N. K. (2003), Small Area Estimation, Wiley, New York.*
- *Rao, J.N.K., (2005), Inferential issues in small area estimation: some new developments, Statistics in Transition, Vol 7, 2003, 513-526.*

Sam Houston
STATE UNIVERSITY
A Member of The Texas State University System