Michigan SAS Users Group (MSUG):
2016 SAS Conference

Fitting Weighted Multilevel Models to Complex Sample Survey Data in SAS: A Case Study

Brady T. West, PhD

Survey Research Center
Institute for Social Research
University of Michigan-Ann Arbor
Email: bwest@umich.edu
Phone: 734-647-4615

Presentation Overview

- Complex Samples and Design-Based Analysis
- Fitting Multilevel Models to Complex Samples
- Using PROC GLIMMIX to fit these models
- A Case Study: the 2013 MMP Provider Survey

Complex Samples

- Generally, any samples that deviate from simple random samples with replacement
- “Complex” samples are generally characterized by any or all of the following:
  - Unequal probabilities of selection for different population elements
  - Stratification of the target population for sampling
  - **Multi-stage** cluster sampling within strata

Complex Samples

- Variables are sometimes provided in public-use data sets containing weights, replicate weights, stratum codes, and/or cluster codes
- These variables contain information about the sample design that may be relevant / informative for outcome variables of interest!
- **Example:** Systolic blood pressure tends to vary across different U.S. counties
- **Approaches need to be <u>considered</u> when fitting models to survey data that account for these features; don't ignore them entirely!**

Design-based Modeling Approaches

- All based on the “randomization” (or “repeated-sampling”) inferential framework
- The data were collected from a random sample from some well-defined **finite population**; each element in the population had a non-zero probability of being selected
- These probabilities drive all of the inference!

Design-based Modeling Approaches

- Unbiased, weighted estimates of finite population model parameters are computed
- Estimated standard errors of the weighted estimates represent <u>estimates of the standard deviation of the sampling distribution</u> of estimates that would be generated from **hypothetical repeated samples** using the same sample design

Design-based Modeling Approaches

- **Common Estimation Approaches:**
  - Pseudo-Maximum Likelihood Estimation for Model Parameters (Binder, 1981, 1983)
  - Robust, Non-parametric, Design-consistent Variance Estimation Approaches:
    - Taylor Series Linearization (TSL)
    - Jackknife / Balanced Repeated Replication (JRR / BRR)
    - Bootstrapping
  - Variances all estimated with respect to the sample design, rather than the model
- Intervals are then formed using approximate degrees of freedom based on the sample design for robust population inference

Design-based Approaches in SAS

- SAS "survey" procedures for <u>analysis</u>:
  - `proc surveymeans`
  - `proc surveyfreq`
  - `proc surveyreg`
  - `proc surveylogistic`
  - `proc surveyphreg`
- Where is `proc surveymixed`?

Multilevel Models (MLMs) for Complex Samples

- In **model-based approaches**, where we specify the best possible probability model for a given variable, effects of randomly sampled clusters (possibly within strata) are generally treated as **random effects**
  - **This captures within-cluster correlation**
- Effects of strata (fixed by design, and not randomly sampled) are generally treated as **fixed effects**
- A multilevel model therefore seems appropriate, and one needs to have **<u>explicit scientific interest</u>** in between-cluster variance components
- **BUT:** What should be done about the weights?

Weighted Estimation of MLMs

- We use the sampling weights <u>at each stage</u> to estimate the parameters of a multilevel model, and compute appropriate standard errors of the weighted estimates
  - Compute variance estimates with respect to the combined randomization <u>**and**</u> model distributions
- Important paper: Pfeffermann et al. (1998)
- Some more recent important papers: Rabe-Hesketh and Skrondal (2006), Carle (2009)

Important: Data Requirements

- Weighted estimation of MLMs **requires weights at each level of the multilevel data structure**
- **Level-1 weights:** inverses of conditional probabilities of selection, given that a Level-2 cluster was sampled (appropriately scaled)
- **Level-2 weights:** generally inverses of probabilities of selection for sampling clusters; should be conditional weights if considering a three-level model
- **Also need cluster codes and stratum codes!**

Important: Data Requirements

- **Key Point:** Don't use the overall (adjusted) sampling weights that are typically provided in public-use data files; use conditional inverse-probability weights at each level
- Rabe-Hesketh and Skrondal (2006) clearly show how the weighted likelihood function (or **pseudo-likelihood**) for a multilevel model requires the inverse-probability weights at <u>each level</u> of the data hierarchy

Weight Scaling

- We **scale** the conditional weights at Level 1 (and other lower levels) of the data hierarchy; normalize these conditional weights to sum to within-cluster sample sizes
- Sums of <u>conditional</u> weights at Level 1 over-state the actual sample sizes within clusters!
- Failure to scale leads to bias in parameter estimates, especially for small samples (Pfeffermann et al., 1998)
- This is particularly important for multilevel logistic regression models (Rabe-Hesketh and Skrondal, 2006)

Fitting Weighted MLMs using PROC GLIMMIX

- First possible in SAS/STAT 13.1
- Important Paper: Zhu (2014)
- **Weight scaling needs to be performed <u>PRIOR</u> to fitting the models using GLIMMIX**
- We consider example syntax in the context of a case study: The 2013 Medical Monitoring Project (MMP) HIV Provider Survey (CDC)
- For more details: West et al. (2015, AJPH)

The 2013 MMP HIV Provider Survey

- A probability sample of HIV care providers selected from outpatient HIV care facilities in 16 states and Puerto Rico
- Two-stage PPS sample design: first sampled states and territories, and then sampled facilities within states and territories
- Attempted to survey all HIV providers within selected facilities

The 2013 MMP HIV Provider Survey

- For purposes of this example, we consider facilities at Level 2, and providers at Level 1
- We have <u>explicit scientific interest</u> in estimating between-facility variance in two provider-specific outcomes
- We also want to examine the fixed effects of various predictors from each level on the two provider-specific outcomes

The 2013 MMP HIV Provider Survey

- **Level 2 weights:** reflect probabilities of selection for the facilities, and incorporate adjustments for facility nonresponse
- **Level 1 weights:** reflect probabilities of **response**, given that all providers in a sampled facility were asked to complete the survey
  - Need to be scaled to sum to the sample size within each facility
- **Many surveys do not provide both weights!**

Multilevel Models for the MMP Data

- Two binary DVs (→ multilevel logit models):
  - Does the provider deliver adequate drug use risk reduction guidance to his/her clients? (based on specific, pre-established criteria…)
  - Does the provider deliver adequate sexual risk reduction guidance to his/her clients?
- Random facility intercepts
- Fixed effects of several provider-level and facility-level covariates

Scaling the Level 1 Weights

```
data tempwts3;
   merge tempwts2 (in=ina) facmeans2 (in=inb);
   by facility_id;
   if ina;
   level1wts = level1wt / meanwt;
   * Scaled weight using Method 2 from Pfeffermann
      et al. (1998);
run;
```

- Note that this is done PRIOR to fitting the models using PROC GLIMMIX!

Fitting the Models using GLIMMIX

```
proc glimmix data = tempwts3;
   class facility_id _gend2_ps _prtype_ps _newrace_ps;
   model _sexrr_ps (event = "1") = _gend2_ps _prtype_ps
  _newrace_ps primcare_ps _specialist_ps integteam_ps
  num200 _rw_fund_ps _facilemr_ps
      / solution link=logit dist=bin obsweight=level1wts;
   random int / subject=facility_id weight=level2wt;
   covtest glm;
run;
```

- **Note: this is the basic syntax (many of the options in GLIMMIX are also available when using weights)**

Comparing Selected Estimates

Outcome		Ignoring Sampling Weights	**Weighted Estimation**
Use of Adequate Drug Risk Reduction Practices	*Parameter*	*Estimate (SE)*	*Estimate (SE)*
	Intercept	**-0.826 (0.606)**	**-1.413 (0.571)***
	Male Provider	-0.354 (0.216)	-0.220 (0.272)
	Primary Care Provider	-0.152 (0.482)	0.257 (0.452)
	Delivers care in language other than English	**0.475 (0.223)***	**0.395 (0.307)**
	% patients injecting drug users > 50%	0.728 (0.503)	0.783 (0.448)
	Works on Integrated Team	-0.032 (0.308)	0.248 (0.339)
	# Patients > 200	-0.006 (0.222)	-0.311 (0.327)
	Ryan White HIV/Aids Program Funded Facility	0.483 (0.326)	0.645 (0.510)
	Facility uses Electronic Medical Records	0.040 (0.263)	-0.073 (0.288)
	var(Intercepts)	**0.033 (0.139)**	**<0.001 (<0.001)**
	# providers	405	405
	# facilities	68	68
	-2 (P)LL	525.47	1173.22 (P)

Comparing Selected Estimates

Outcome		Ignoring Sampling Weights	**Weighted Estimation**
Use of Adequate Sexual Risk Reduction Practices	**Intercept**	**-1.000 (0.727)**	**-1.278 (0.641)***
	Male Provider	**-0.350 (0.218)**	**-0.390 (0.173)***
	White Provider	-0.169 (0.307)	-0.453 (0.359)
	Black Provider	**0.800 (0.468)**	**1.059 (0.305)****
	Hispanic Provider (Reference: Other)	0.050 (0.402)	-0.259 (0.335)
	Nurse Practitioner (Reference: PA)	**0.817 (0.443)**	**1.175 (0.420)****
	HIV Specialist	-0.011 (0.238)	0.206 (0.373)
	Primary Care Provider	-0.028 (0.454)	-0.110 (0.530)
	Works on Integrated Team	0.552 (0.309)	0.693 (0.416)
	# Patients > 200	0.225 (0.222)	-0.015 (0.245)
	Ryan White HIV/Aids Program Funded Facility	-0.006 (0.310)	0.023 (0.539)
	Facility uses Electronic Medical Records	-0.183 (0.276)	-0.056 (0.269)
	var(Intercepts)	**0.171 (0.150)**	**0.046 (0.161)**
	# providers	514	514
	# facilities	68	68
	-2 (P)LL	636.11	1343.70 (P)

Comparing Selected Estimates

- There were several differences in inference when applying the sampling weights at each level, particularly for the fixed effects
- The weights appeared to be **informative** about these relationships of interest in the target population
- The between-facility variance was being over-stated when ignoring the weights (larger facilities had a higher sampling probability)
- The fixed effects in the weighted models were explaining more of the between-facility variance

Take-Away Points

- Fitting a variety of multilevel models (logistic, normal, etc.) in a way that accounts for sampling weights is now easy in GLIMMIX
- **Best Practice:** consider both weighted and unweighted estimation, and assess the sensitivity of your inferences to the weights
- If models are well-specified or weights are not informative, use of weights may result in inefficient (and similar) estimates

Some Final Points…

- Survey agencies often do not release the weight information necessary to implement these weighted multilevel modeling approaches (mainly the weights associated with sampling clusters…disclosure risk?)
- Analysts thus need to resort to design-based approaches, or including fixed effects of the (centered?) respondent weights in the MLMs…
- Whether using weights or not, **make sure that you have explicit scientific interest in estimating between-cluster variance** (otherwise, why fit MLMs?)

References

- Carle, A.C. (2009). Fitting multilevel models in complex survey data with design weights: Recommendations. *BMC Medical Research Methodology*, 9:49.
- Dumouchel, W.H. and Duncan, G.J. (1983). Using sample survey weights in multiple regression analysis of stratified samples. *Journal of the American Statistical Association*, 78(383), 535-543.
- Fuller, W. (2009). *Sampling Statistics*. Wiley.
- Gelman, A. (2007). Struggles with survey weighting and regression modeling. *Statistical Science*, 22 (2), 153–164.
- Heeringa, S.G., West, B.T., and Berglund, P.A. (2010). *Applied Survey Data Analysis*. Chapman & Hall / CRC Press.

References

- Korn, E.L. and Graubard, B.G. (1999). *Analysis of Health Surveys*. Wiley.
- Little, R.J.A. (1991). Inference with survey weights. *Journal of Official Statistics,* 7, 405-424.
- Pfeffermann, D. (1993). The role of sampling weights when modelling survey data. *International Statistical Review*, 61, 317-337.
- Pfeffermann, D., Skinner, C.J., Holmes, D.J., Goldstein, H., and Rasbash, J. (1998). Weighting for unequal selection probabilities in multilevel models. *Journal of the Royal Statistical Society, Series B,* 60(1), 23-40.
- Rabe-Hesketh, S. and Skrondal, A. (2006). Multilevel modelling of complex survey data. *Journal of the Royal Statistical Society, Series A*, 169(4), 805-827.

References

- West, B.T., Beer, L., Gremel, W., Weiser, J., Johnson, C., Garg, S., and Skarbinski, J. (2015). Weighted Multilevel Models: A Case Study. *American Journal of Public Health*, 105(11), 2214-2215.
- Zhu, M. (2014). Analyzing Multilevel Models With the GLIMMIX Procedure. Cary, NC: SAS Institute Inc. Paper SAS026-2014.

Thanks! Questions?

- Thank you for attending the conference today!
- Thanks to Brandy Sinco for the invitation, and also to my co-authors from the CDC and Altarum Institute for allowing me to work with the MMP data. Views expressed are my own.
- For any follow-up inquiries or additional references, please email **bwest@umich.edu**.