Longitudinal data

Thomas Lumley

13/07/2022

Topics

- differences between years
- repeated measurements on people

Differences over years

Suppose you have 5 two-year waves of NHANES, each estimating the national population, and you want trends over time

- treat each year as a separate set of **strata**
- (therefore, different PSUs)
- maybe rescale weights

You need to rescale weights if you want to estimate *totals*. Divide each by five and maybe rescale so population is the same in each wave or maybe not

Then just do regression.

Alternatively: estimate within each year and combine. This is messy and retro

Repeated measurements on people

Panel surveys: same people get measured multiple times, which increases precision for time trends **a lot**

Measurements within individual are just another nested stage of sampling. If you have the full structure

```
svydesign(id=~cluster+person+time, strata=~whatever, ...)
```

If you don't (public use data)

```
svydesign(id=~cluster, strata=~whatever, ...)
```

Then just `svyglm`. The correlation is accounted for as part of the sampling; you don't need mixed models or anything.