ACE Models with the NLSY

William Howard Beasley (Howard Live Oak LLC, Norman)
Joseph Lee Rodgers (University of Oklahoma, Norman)
Kelly Meredith (University of Oklahoma, Norman)
David Bard (University of Oklahoma Health Sciences Center, OKC)

May 4, 2012

Abstract

We describe how to use the NlsyLinks package to examine various biometric models, using the NLSY79.

Contents

1 Ambiguous twins

What are "ambiguous twins"?

MZ twins share all of the genetic information (*i.e.*, $R = 1$), while DZ twins on average share half (*i.e.*, $R = 0.5$). Sometimes a sibling pair doesn't have enough information for us to classify comfortably as either MZ or DZ. We assign these "ambiguous twins" $R = .75$. Currently there are 13 ambiguous twins in the NLSY79C sample.

Of these 13 pairs, all had close birthdays and were the same gender. 12 pairs are ambiguous because the mother didn't complete an NLSY survey since 1993; the first twin items were presented in 1994 (*e.g.*, R48257.00, and R48260.00). The mother of 13th pair (*i.e.*, subjects 864902 and 864903) simply avoided responding to the twin survey items.

Occasionally they mother of twins provided conflicting evidence. Fortunately, these mother were consistent among their most recent responses. For instance, Subjects 392401 and 392402 were indicated DZ in 1998, but MZ in 2000, 2002, and 2004. This pair was assigned $R = 1$.

Gen2 ambiguous twins can be viewed with:

```
> subset(Links79Pair, RelationshipPath=='Gen2Siblings' & R==.75)

     ExtendedID Subject1Tag Subject2Tag    R RelationshipPath
1309       1460      146001      146002 0.75     Gen2Siblings
5013       5658      565901      565902 0.75     Gen2Siblings
6275       6639      663901      663902 0.75     Gen2Siblings
6665       7111      711101      711102 0.75     Gen2Siblings
7483       7913      791406      791407 0.75     Gen2Siblings
8436       8649      864902      864903 0.75     Gen2Siblings
9338       9596      959601      959602 0.75     Gen2Siblings
```

```
  9725     10012     1001201     1001202 0.75     Gen2Siblings
 10387     11191     1119103     1119104 0.75     Gen2Siblings
 10435     11486     1148601     1148602 0.75     Gen2Siblings
 10461     11733     1173301     1173302 0.75     Gen2Siblings
 10462     11739     1173901     1173902 0.75     Gen2Siblings
 11062     12574     1257402     1257403 0.75     Gen2Siblings
```

The 129 Gen2 siblings who are potentially twins or triplets can be similarly viewed with:

```
subset(Links79PairExpanded, RelationshipPath=='Gen2Siblings' & MultipleBirth != 'No')
```

2   Ambiguous siblings

What are "ambiguous siblings"?

Similar to ambiguous twins, ambiguous siblings are sibling pairs that we cannot comfortably classify as either full-siblings ($R = .5$) or half-siblings ($R = .25$). All siblings in the NLSY79-C/YA dataset share the same biological mother, so for these pairs, the problem is reduced to determining if they share the same biological father. There are two typical reasons for classifying siblings as ambiguous: (a) the relevant items are missing responses, or (b) the existing responses conflict with each other.

For instance, there are at least 173 Gen2 pairs where one sibling explicitly reported they shared a biological father, while the other sibling explicitly reported they did not. These subjects can be viewed with:

```
subset(Links79PairExpanded, RelationshipPath=='Gen2Siblings' &
   ((RExplicitOlderSibVersion==.5 & RExplicitYoungerSibVersion==.25)
      | (RExplicitOlderSibVersion==.25 & RExplicitYoungerSibVersion==.5))).
```

Another example occurs when a subject reports they are unsure or if their own responses are inconsistent over the years. These 69 Gen2 pairs can be viewed with:

```
subset(Links79PairExpanded, RelationshipPath=='Gen2Siblings' &
   (RExplicitOlderSibVersion==.375 | RExplicitYoungerSibVersion==.375))
```

3   Retaining *vs.* dropping the ambiguous twins and siblings

I am running ACE models with sibling pairs. Do you recommend including the pairs who are classified as $R = .375$ or $R = .75$? Or should I exclude them from the analyses?
This issue touches Behavior Genetic concepts and modeling pragmatics.

This is an important question

4   Grant support