

Sol 3063: Ejemplo análisis de sensibilidad

Luis Maldonado

Mayo, 2016

1. El análisis de sensibilidad puede resultar contra-intuitivo. La razón es que estamos hablando de algo (variable omitida) que no observamos. Para ilustrar como funciona este tipo de análisis, veamos el siguiente ejemplo, el cual trabaja con una base de datos con 10 parejas de tratados y controles. Las variables son: id=identificador persona, y1=binary outcome, d1=tratamiento (1 es tratado, 0 es control), y block1=pareja.

```
> rm(list = ls())
> library(foreign)
> # Fijese cual es su directorio
> setwd('/Users/luismaldonado/Dropbox/Sol3026_3063/2016/Clases/Clase 13/Guia Sensibilidad')
> dat <- read.dta("Sensitivity1.dta",convert.factors = F)
> dat
```

	id	y1	d1	block1
1	1	0	0	1
2	2	1	1	1
3	4	1	0	2
4	3	0	1	2
5	6	0	0	3
6	5	0	1	3
7	8	1	0	4
8	7	1	1	4
9	9	0	0	5
10	10	0	1	5
11	12	1	0	6
12	11	1	1	6
13	14	1	0	7
14	13	1	1	7
15	15	0	0	8
16	16	1	1	8
17	17	0	0	9
18	18	0	1	9
19	20	0	0	10
20	19	0	1	10

2. En el caso de outcomes binarios, vimos en clases los Rosenbaum Bounds. Estos se basan en el McNemar T test. Para detalles de este test, vea el cap. 4 del libro de Rosenbaum (2002) *Observational Studies*. He dejado una copia de este capítulo en la web del curso.
3. Tal como vimos en la tabla, lo central, luego de aplicar matching y formar parejas, es completar la siguiente tabla, particularmente las celdas que corresponden a las parejas discordantes.

	(D=1,Y=1)	(D=1,Y=0)
(D=0, Y=1)	a	b
(D=0, Y=0)	c	d

4. Para completar esta tabla con los datos de las 10 parejas arriba mostradas, lo primero es transformar estos datos en una base de parejas. OJO, esto no es por si mismo necesario, es solo un camino de completar la tabla.
5. En el siguiente código, usamos una funcion que se llama `fullmatch2sensmat`, la cual es gentiliza de Jake Bowers y nos facilita mucho el trabajo

```
> # Preparacion de los datos
> fullmatch2sensmat<-function(y,z,fm){
+   ## A function to reformat fullmatches for use with sensmv/mw
+   ## y is the outcome
+   ## z is binary treatment indicator (1=assigned treatment)
+   ## fm is a factor variable indicating matched set membership
+   ## We assume that y,z, and fm have no missing data.
+   dat<-data.frame(y=y,z=z,fm=fm)[order(fm,z,decreasing=TRUE),]
+   numcols<-max(table(fm))
+   resplist<-lapply(split(y,fm),
+     function(x){
+       return(c(x,rep(NA, max(numcols-length(x),0))))
+     })
+   respmat<-t(simplify2array(resplist))
+   return(respmat)
+ }
> respmat<-with(dat,fullmatch2sensmat(y1,d1,block1))
> respmat
```

	[,1]	[,2]
1	0	1
2	1	0
3	0	0
4	1	1
5	0	0
6	1	1
7	1	1
8	0	1
9	0	0
10	0	0

6. Qué es `respmat`? Compare con la base `dat` y encontrará la respuesta.
7. Con los datos expresados en `respmat` ya podemos llenar la tabla, pero voy a dar un paso más, solo para hacer las estimaciones aún más simples. Note lo que hago a continuación.

```
> #Estimares de tipos de parejas
> Y0<-respmat[,1]
> Y1<-respmat[,2]
> dat3 = data.frame(Y0,Y1)
> dat3
```

	Y0	Y1
1	0	1
2	1	0
3	0	0
4	1	1
5	0	0
6	1	1
7	1	1
8	0	1
9	0	0
10	0	0

```
> dat
```

	id	y1	d1	block1
1	1	0	0	1
2	2	1	1	1
3	4	1	0	2
4	3	0	1	2
5	6	0	0	3
6	5	0	1	3
7	8	1	0	4
8	7	1	1	4
9	9	0	0	5
10	10	0	1	5
11	12	1	0	6
12	11	1	1	6
13	14	1	0	7
14	13	1	1	7
15	15	0	0	8
16	16	1	1	8
17	17	0	0	9
18	18	0	1	9
19	20	0	0	10
20	19	0	1	10

8. Note que respmat es una matrix, mientras que dat3 es una data frame. Ahora bien, comparemos con los datos de la base dat.

- Si se fija en las filas de dat3, note que corresponden a las parejas. Por ejemplo la fila 1 corresponde al block1=1 en la base dat. En dat3, Y0 corresponde al valor en el outcome para los controles y Y1 es el valor en el outcome para los tratados. Por ejemplo, para la pareja 8 en dat3, el control tiene un valor de 0 en el outcome, y el tratado un valor de 1 en la variable dependiente.

9. Con esta información estamos listos para completar la tabla.

	(D=1,Y=1)	(D=1,Y=0)
(D=0, Y=1)	a=3 (Blocks 4,6,7)	b=1 (Block 2)
(D=0, Y=0)	c=2 (Blocks 1 y 8)	d=4 (Blocks 3,5,9,10)

10. Como ejemplo, veamos la celda b. Esta corresponde a las parejas en donde i) un miembro de esta si no recibe el tratamiento, tiene el valor de 1 en el outcome (D=0,Y=1) y ii) el otro miembro si recibe el tratamiento, no presenta el valor de 1 en el outcome (D=1,Y=0) (Ver Clases!!). Si observamos la información disponible en dat 3, notamos que solo la pareja/Bloc número 2 cumple esta condición y, por lo tanto, esta celda b tiene un valor de 1, es decir, solo una pareja.

11. Para realizar el McNemar test, necesitamos saber las celdas c (2), b (1) y el c+b (3). En R, podemos usar el siguiente código.

```
> sum(dat3$Y1!=dat3$Y0) #Numero total de pares discordantes=3
[1] 3

> sum(dat3$Y0==1 & dat3$Y1==0) #CELDA B, (D=0,Y=1) y (D=1,Y=0)=1
[1] 1

> sum(dat3$Y0==0 & dat3$Y1==1) #CELDA C, (D=0,Y=0) y (D=1,Y=1)=2
[1] 2
```

12. Con esta información, podemos realizar el test de sensibilidad. Voy a trabajar con un Gamma de 1.10

```

> #A mano con gamma de 1.10
> gamma<-1.10 #Paso 1: especificar el valor de gamma
> pp<-gamma/(1+gamma) #Paso 2: calcular p+ para este gamma
> binom.test(2,3,p=pp,alternative="greater") #Paso 3: test binomial exacto

```

Exact binomial test

```

data: 2 and 3
number of successes = 2, number of trials = 3, p-value = 0.5357
alternative hypothesis: true probability of success is greater than 0.5238095
95 percent confidence interval:
 0.1353504 1.0000000
sample estimates:
probability of success
      0.6666667

```

13. Para interpretación, vea la clase 13b. Inténtelo con el paquete rbounds, debe llegar al mismo resultado.