

Lab #3

0. BACKGROUND

This lab is meant to give you practice using G-computation to calculate a relative risk for a binary effect.

1. DOWNLOAD

•Download the model building dataset `lab3.dta`. The data was collected by the United Network on Organ Sharing and contains 1 year survival outcomes on pediatric kidney transplants carried out in the US between 1990 and 2002. The variables are

- death (1=died in first year, 0 = survived first year) — the outcome
- age (age of child at transplant)
- age_don (age of the kidney donor)
- txtype (1=cadaveric donor, 0 = living donor)
- cold_isc (cold ischemia time in minutes)
- hlatmat (number of matching HLA loci)
- year (calendar year of transplant)
- id (participant id)

Here I have completed the cold_ischemia times by using single imputation

Read in the data

```
use "lab3.dta", clear
```

2. REGRESSION

Our objective is to understand the effect of txtype (1=cadaveric, 0=living) on the risk of death. Let's start with a two-way table

```
tabulate death txtype, col chi2
```

Show the proportions. What proportion die in the two txtype groups?

3. REGRESSION

Our objective is to understand the adjusted effect of txtype (1=cadaveric, 0=living) on the risk of death. One way to do that is to fit a relative risk regression — a binomial regression with a log link

```
glm death i.hlamat c.age_don c.age i.txtype c.year,  
family(binomial) link(log) eform
```

What is the relative risk of death associated with the txtype?

Also, fit a logistic regression model.

```
logistic death i.hlamat c.age_don c.age i.txtype c.year
```

What is the odds ratio of death associated with the txtype? Is it similar to the relative risk?

4. G-COMPUTATION (POTENTIAL OUTCOME ESTIMATION)

You can use either model to predict the probability of death if the entire sample received a cadaveric transplant (we will call this “p1”) or if the entire sample received a living transplant (we will call this “p0”). This is done with the margins command. Let's calculate them first based on the logistic regression model.

First, repeat the logistic regression

```
logistic death i.hlamat c.age_don c.age i.txtype c.year
```

You can obtain p1 and p0 with the margins command

```
margins txtype, post
```

What are the estimated probability of death if all patients in the cohort received a living organ ? What are the estimated probability of death if all patients in the cohort received a cadaveric organ ?

It is possible to obtain a relative risk (RR) with confidence interval by typing

```
nlcom log(_b[1.txtype]) - log(_b[0.txtype])
```

You will have to exponentiate the values to get the RR and it's limits. To exponentiate the value , 0.3226191, you type

```
dis exp(.3226191)
```

You may be curious to understand how p0 and p1 are calculated. First, refit the logistic model.

```
logistic death i.hlamat c.age_don c.age i.txtype c.year
```

“p1” is calculated based on the logistic model using the regression equation but with tx-type set to “1” for each person. Before doing that, save the value of txtype in a variable “oldtxtype”

```
gen oldtxtype = txtype  
replace txtype = 1  
predict p1, p
```

This value of p1 is then averaged across the entire sample

```
mean p1
```

Note...this value is identical to the estimated value from the margins command.

If you want, you can attempt the same approach to estimate p0

```
replace txtype = 0  
predict p0, p
```

```
mean p0
```