

Estimando Medidas de Desigualdade de Renda com a PNADC

Solução

Guilherme Jacob

17/08/2021

Vamos precisar da versão de desenvolvedor do pacote `convey`. Ela pode ser instalada usando o comando a seguir:

```
devtools::install_github( "ajdamico/convey" , ref = "svyby-covmat" )
```

O conjunto de dados que vamos utilizar neste exercício está guardado no arquivo `pnadc-recorte.Rds` na pasta `dados` e pode ser lido usando a função `readRDS`. Para ler estes dados e guardá-los no objeto `pnadc.df`, use o comando abaixo (verifique o endereço do arquivo no seu computador!):

```
pnadc.df <- readRDS( "/home/guilherme/GitHub/cursoIntroRenda/dados/pnadc-recorte.Rds" )
```

Este arquivo contém um recorte dos microdados da 1ª visita da PNAD Contínua. Com ele, realize as tarefas a seguir:

1. Crie um objeto de plano amostral;

```
# carrega libraries
library( survey )
library( convey )

# cria desenho amostral sem pós-estratificação
pnadc.design <-
  svydesign(
    ids = ~ upa + v1008 ,
    strata = ~ estrato ,
    weights = ~ v1032 ,
    data = pnadc.df ,
    nest = TRUE )

# prepara para convey
pnadc.design <- convey_prep( pnadc.design )
```

2. Estime as seguintes medidas de desigualdade:

2.a) Índice de Gini do rendimento domiciliar per capita (`def.rdp`);

```
svygini( ~def.rdp , pnadc.design , na.rm = TRUE )
```

```
##              gini      SE
## def.rdp 0.54297 0.0042
```

2.b) Índices de Gini do rendimento domiciliar per capita (`def.rdp`) nas áreas urbanas usando a função `subset` e a variável `v1022`;

```
svygini( ~def.rdp , subset( pnadc.design , v1022 == "Urbano" ) , na.rm = TRUE )
```

```
##          gini      SE
## def.rdpc 0.53427 0.0044
```

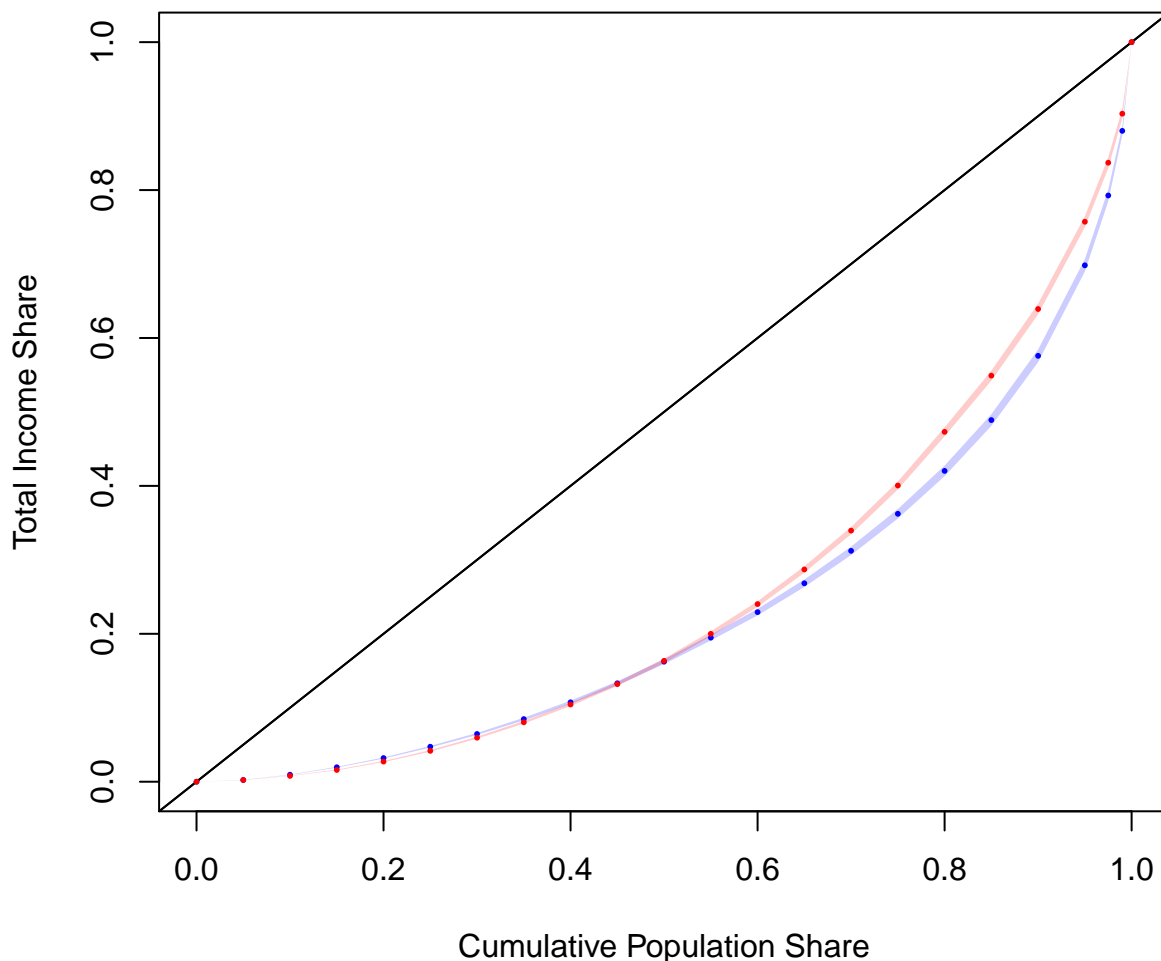
2.c) Índices de Gini do rendimento de todos os trabalhos (def.vd4019) para homens e mulheres usando a função `svyby` e a variável `v2007`;

```
svyby( ~def.vd4019 , ~v2007 , pnadc.design , svygini , na.rm = TRUE )
```

```
##          v2007 def.vd4019      se
## Homem   Homem  0.5145646 0.006202286
## Mulher  Mulher  0.4925970 0.004852430
```

2.d) Curvas de Lorenz do rendimento domiciliar per capita para os domínios urbano e rural.

```
svylorenz( ~def.rdpc , subset( pnadc.design , v1022 == "Urbano" ) ,
  quantile = c( seq( 0 , .95 , .05 ) , .975 , .99 , 1 ) ,
  na.rm = TRUE , curve.col = "blue" )
svylorenz( ~def.rdpc , subset( pnadc.design , v1022 == "Rural" ) ,
  quantile = c( seq( 0 , .95 , .05 ) , .975 , .99 , 1 ) ,
  na.rm = TRUE , add = TRUE , curve.col = "red" )
```



3. Escolha uma região e crie um objeto de plano amostral para ela usando a função `subset`. Com ele:

```
pnadc.regiao <- subset( pnadc.design , regiao == "Norte" )
```

3.a) Índice de Gini do rendimento domiciliar per capita (def.rdpc);

```
svygini( ~def.rdpc , pnadc.regiao , na.rm = TRUE )
```

```
##           gini      SE
## def.rdpc 0.53742 0.0069
```

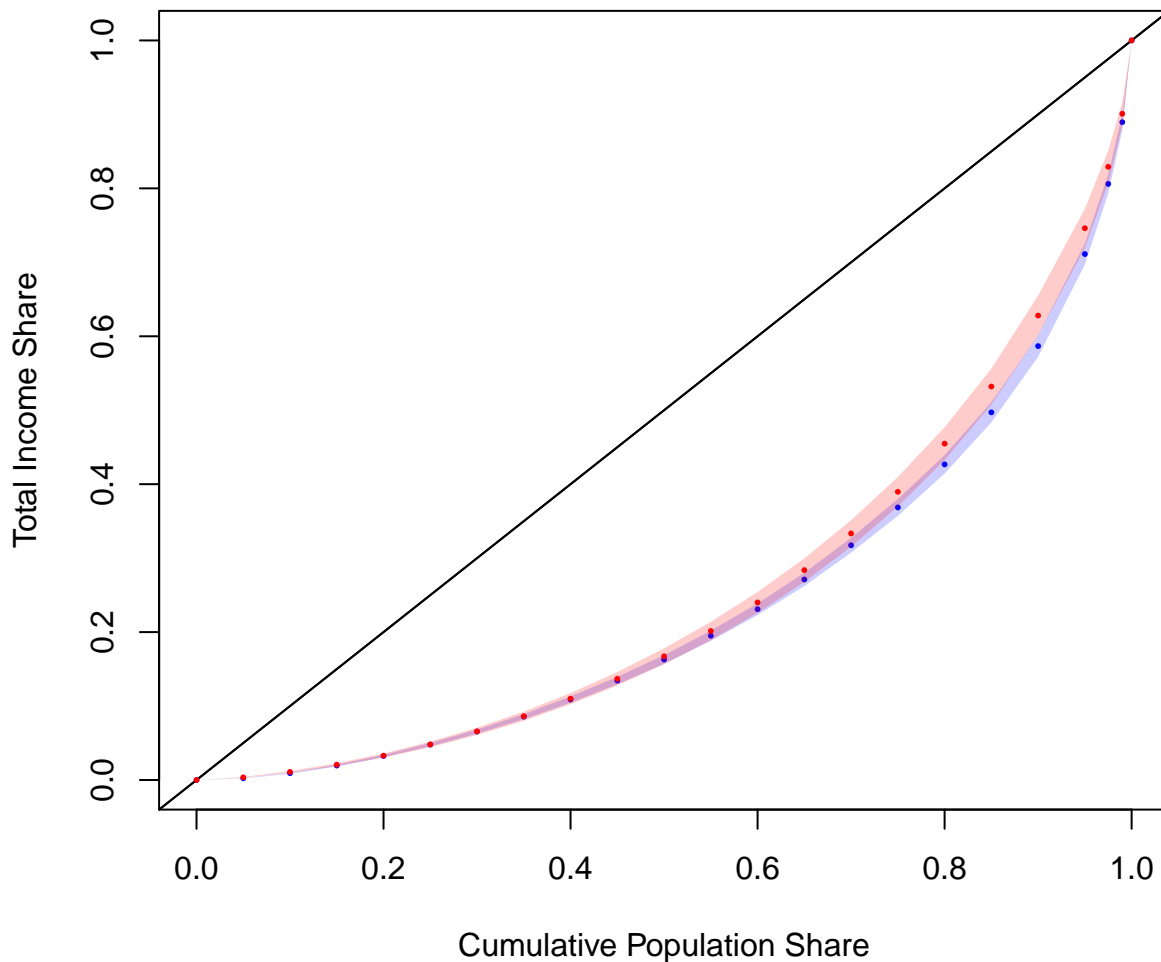
3.b) Índices de Gini do rendimento domiciliar per capita (def.rdpc) na área urbana usando a função `subset` e a variável `v1022`;

```
svygini( ~def.rdpc , subset( pnadc.regiao , v1022 == "Urbano" ) , na.rm = TRUE )
```

```
##           gini      SE
## def.rdpc 0.52781 0.0076
```

3.c) Curvas de Lorenz do rendimento domiciliar per capita para os domínios urbano e rural;

```
svylorenz( ~def.rdpc , subset( pnadc.regiao , v1022 == "Urbano" ) ,
  quantile = c( seq( 0 , .95 , .05 ) , .975 , .99 , 1 ) ,
  na.rm = TRUE , curve.col = "blue" )
svylorenz( ~def.rdpc , subset( pnadc.regiao , v1022 == "Rural" ) ,
  quantile = c( seq( 0 , .95 , .05 ) , .975 , .99 , 1 ) ,
  na.rm = TRUE , add = TRUE , curve.col = "red" )
```



3.d) Índice de Gini, Theil-L e Theil T do rendimento de todos os trabalhos (def.vd4019) para homens e mulheres usando a função `svyby` e a variável `v2007`;

```
# índice de Gini
svyby( ~def.vd4019 , ~v2007 , pnadc.regiao , svygini , na.rm = TRUE )
```

```
##          v2007 def.vd4019          se
## Homem   Homem   0.5070483 0.009660340
## Mulher  Mulher  0.4993774 0.007850074
```

```
# Theil-L
svyby( ~def.vd4019 , ~v2007 , pnadc.regiao , svygei , epsilon = 0 , na.rm = TRUE )
```

```
##          v2007 def.vd4019          se
## Homem   Homem   0.4793381 0.01824961
## Mulher  Mulher  0.4761517 0.01494327
```

```
# Theil-T
svyby( ~def.vd4019 , ~v2007 , pnadc.regiao , svygei , epsilon = 1 , na.rm = TRUE )
```

```
##          v2007 def.vd4019          se
## Homem   Homem   0.5435326 0.03529345
## Mulher  Mulher  0.4826058 0.01838917
```

3.e) Decomponha o Theil-T do rendimento de todos os trabalhos por desigualdade por grupos de anos de estudo (vd3006) usando a função svygeidec.

```
svygeidec( ~def.vd4019 , ~vd3006 , pnadc.regiao , epsilon = 1 , na.rm = TRUE )
```

```
##          gei decomposition          SE
## total                0.52175 0.0256
## within                0.37569 0.0213
## between               0.14606 0.0097
```