

## AULA 6

### Análise de Dados Legislativos e Eleitorais Utilizando o Programa Stata

Professor: Ernesto Friedrich de Lima Amaral

Email: [eflamaral@gmail.com](mailto:eflamaral@gmail.com)

Site do curso: [www.ernestoamaral.com/stata20091.html](http://www.ernestoamaral.com/stata20091.html)

Data: 28/05/2009

Horário: 18:00–21:00

### Sites importantes para uso de modelos hierárquicos com Stata

– Stata:

<http://www.stata.com/stata9/mixed.html>

– Portal de Estatística Computacional da Universidade da Califórnia de Los Angeles (UCLA):

<http://statcomp.ats.ucla.edu/mlm/default.htm>

– "Generalized Linear Latent And Mixed Models (GLLAMM)":

<http://www.gllamm.org>

– Instalação do "Generalized Linear Latent And Mixed Models (GLLAMM)":

<http://www.gllamm.org/install.html>

### Vamos rodar os últimos comandos da aula 5

```
clear
set mem 100m
set more off

log using "C:\cursodcp\log\Aula6.log", text replace

*Agrupando os bancos de dados da Argentina, Chile e Guatemala
use "C:\cursodcp\dados\Argentina51.dta", clear
append using "C:\cursodcp\dados\Chile42.dta"
append using "C:\cursodcp\dados\Guatemala52.dta"

*Mantendo somente as variáveis de interesse no banco
keep pais p34 p35 p58 peso
label variable pais "País"

*No Chile, há opção 3 na variável p34,
*que significa "não aumentar impostos".
*Essa opção será colocada como missing
*em nossa nova variável.
tab p34 pais, nolabel missing

*Recodificando opinião dos deputados sobre imposto
tab p34, missing nolabel
gen imposto=p34
replace imposto=. if p34==3 | p34==8 | p34==9
tab p34 imposto, missing nolabel
```

```

*Recodificando variável imposto para 0/1
tab imposto, missing nolabel
gen impostod=.
replace impostod=0 if imposto==1
replace impostod=1 if imposto==2
tab imposto impostod

*Recodificando opinião dos deputados sobre privatização
tab p35, missing nolabel
gen privatizar=p35
replace privatizar=. if p35==8 | p35==9
tab privatizar, missing nolabel

*Criando variáveis dicotômicas de privatizar (uso do tabulate)
tab privatizar, gen(privatizar)

*Recodificando escala esquerda/direita do deputado
tab p58, missing nolabel
gen esqdir=p58
replace esqdir=. if p58==98 | p58==99
tab p58 esqdir, missing nolabel

*Criando variável dicotômica de deputado de esquerda
gen esquerda=.
replace esquerda=1 if esqdir>=1 & esqdir<=4
replace esquerda=0 if esqdir>=5 & esqdir<=10
tab esqdir esquerda, missing nolabel

*Excluindo valores missing de nossas variaveis de interesse
di _N
drop if esquerda==. | impostod==. | pais==.
di _N

```

## Modelos hierárquicos (modelos com intercepto aleatório)

Primeiramente, podemos estimar modelos com intercepto aleatório (*random-intercept models*) com os comandos "xtreg", "xtmixed" e "gllamm". Nos exemplos a seguir vamos estimar o efeito da variável "impostod" (opinião do deputado sobre cobrança de imposto) na variável "esquerda" (indica se o deputado se auto-classificou como esquerdista), levando em consideração que há uma variação aleatória desse efeito nos países de nosso banco de dados. Ou seja, assumimos que a distribuição dos deputados de cada país pela variável "esquerda" não ocorre de maneira uniforme:

```

label define nomes 42 "Chile" 51 "Argentina" 52 "Guatemala"
label values pais nomes
tab esquerda pais, col

```

No comando "xtreg", a opção "i" identifica os diferentes níveis das variáveis independentes. Ou seja, essa opção indica o segundo nível da regressão. Além disso, é utilizada a técnica de estimação estatística de máxima verossimilhança (mle) para estimação dos coeficientes:

```
xtreg esquerda impostod, i(pais) mle
```

Um modelo similar ao anterior pode ser estimado com o comando "xtmixed". A opção "||" separa a parte aleatória da parte fixa do modelo. A parte aleatória é colocada do lado direito dessa opção:

```
xtmixed esquerda impostod || pais:, mle
```

É possível ainda usar o comando "gllamm" para estimar modelos com intercepto aleatório. Antes de tudo, é preciso instalar o pacote que foi obtido nesse site (<http://www.gllamm.org/install.html>). Após gravar os arquivos de interesse no diretório "C:\cursodcp\pacotes\gllamm", é usado o comando "adopath" para instalar o pacote. O comando "which" indica a versão do pacote instalado:

```
adopath + "C:\cursodcp\pacotes\gllamm"
which gllamm
```

A estimação do modelo com intercepto aleatório com o comando "gllamm" leva mais tempo para gerar os resultados:

```
gllamm esquerda impostod, i(pais)
```

## Modelos hierárquicos (efeito ENTRE grupos)

O comando "xtreg" tem a opção de estimar modelos que mensuram a regressão sobre médias de grupos (*regression on group means*). Em nosso caso, as médias de países. Dessa forma, o efeito estimado é aquele ENTRE os sujeitos/grupos do segundo nível (*between-subject effect*). O efeito da variável independente (impostod) sobre a variável dependente (esquerda) é estimado para cada grupo (país), e então é realizada uma média entre os efeitos de tais países. No nosso caso, esse modelo não mostra significância estatística:

```
xtreg esquerda impostod, i(pais) be
```

## Modelos hierárquicos (efeito DENTRO de grupos)

Como vimos na aula passada, o comando "xtreg" ainda permite estimar modelos de efeitos fixos (*fixed-effects models*) que mensuram os impactos DENTRO dos sujeitos/grupos do segundo nível (*within-subject effect*). O efeito da variável independente sobre a variável dependente é estimado após controlar a regressão pela variável de segundo nível (pais). Dessa forma, o modelo estima o efeito da opinião acerca de imposto (impostod) sobre posicionamento ideológico (esquerda), após considerar que há um efeito fixo de cada país influenciando na relação entre nossas variáveis de primeiro nível. Em suma, é como se o modelo estimasse coeficientes para cada um dos países, mas que não são exibidos nos resultados da regressão:

```
xtreg esquerda impostod, i(pais) fe
```

## Outra forma de testar o efeito do segundo nível

Os modelos multi-níveis (*multilevel models*) também podem ser estimados somente com o componente de variância do segundo nível. Isso permite verificar se realmente há uma variação significativa nos resultados da variável dependente (esquerda), considerando os valores de nossa variável de segundo nível (pais):

```
xtmixed esquerda || pais:, mle
```

Se houver relação significativa entre as variáveis acima, podemos supor que é importante incluir essa variável de segundo nível (pais), quando modelos de regressão são estimados na explicação de "esquerda". Posteriormente, o modelo com componente de variância em dois níveis (*2-levels variance-component model*) é estimado, como realizado em exemplo anterior:

```
xtmixed esquerda impostod || pais:, mle
```

## Explorando os gráficos de dispersão

Gráfico de dispersão com linha de tendência entre auto-classificação do deputado na escala esquerda/direita (esqdir) e sua opinião sobre privatização (privatizar):

```
twoway (scatter esqdir privatizar) (lfit esqdir privatizar)
```

Gráfico de dispersão com linha de tendência e intervalo de confiança entre "esqdir" e "privatizar":

```
twoway (scatter esqdir privatizar) (lfitci esqdir privatizar)
```

Curiosidade... Gráfico que exhibe a função da curva normal:

```
twoway function y=normden(x), range(-4 4)
```

## Análise de regressão com gráficos

Salvar valores preditos e resíduos:

```
regress esqdir privatizar [pweight=peso]
predict esqpred
gen res=esqdir-esqpred
```

```
label variable esqdir "Escala esquerda/direita"
label variable privatizar "Opinião sobre privatização"
label variable esqpred "Valores preditos de esquerda/direita"
label variable res "Resíduo da regressão entre esq/dir e privatização"
```

Gráfico de dispersão dos valores originais de esquerda/direita (esqdir) e valores preditos de esquerda/direita (esqpred) pela variável independente (privatizar):

```
twoway (scatter esqdir privatizar) ///
      (scatter esqpred privatizar, connect(1) clcolor(red) clpattern(dot))
```

Gráfico de dispersão dos resíduos da regressão entre escala esquerda/direita e opinião sobre privatização (res) pelos valores preditos de esquerda/direita (esqpred):

```
twoway (scatter res esqpred, yline(0))
```

## Adicionando a variável "país" na regressão

Inserindo variáveis categóricas de país no gráfico:

```
tab esquerda pais, col
tab pais, sum(esqdir)
tab pais, gen(pais)
```

```
xi: regress esqdir privatizar i.pais [pweight=peso]
predict esqpred2
gen res2=esqdir-esqpred2
```

```
label variable esqdir "Valores observados de esq/dir"
label variable privatizar "Opinião sobre privatização"
label variable esqpred2 "Valores preditos de esq/dir"
label variable res2 "Resíduo da regressão entre esq/dir e privatização/país"
```

Diagrama de dispersão com uma linha para cada país dos valores originais de esquerda/direita (esqdir) e valores preditos de esquerda/direita (esqpred2) pela variável independente (privatizar):

```
twoway (scatter esqdir privatizar) ///
      (scatter esqpred2 privatizar, connect(1) clcolor(red) clpattern(dot)), ///
///
title("Diagrama de dispersão para legisladores do Chile, Argentina e Guatemala" ///
      "dos valores observados de esquerda/direita e valores preditos de esquerda/direita" ///
      "pela variável de opinião sobre privatização" ///
      , size(medsmall)) ///
///
ytitle("Escala esquerda/direita")
graph export "C:\cursodcp\graf\pred2todos.wmf", replace
```

Diagrama de dispersão para o Chile (pais=42) dos valores originais de esquerda/direita (esqdir) e valores preditos de esquerda/direita (esqpred2) pela variável independente (privatizar):

```
twoway (scatter esqdir privatizar) ///
      (scatter esqpred2 privatizar, connect(1) clcolor(red) clpattern(dot)) ///
      if pais==42, ///
///
title("Diagrama de dispersão para legisladores do Chile" ///
      "dos valores observados de esquerda/direita e valores preditos de esquerda/direita" ///
      "pela variável de opinião sobre privatização" ///
      , size(medsmall)) ///
///
ytitle("Escala esquerda/direita")
graph export "C:\cursodcp\graf\pred2chi.wmf", replace
```

Diagrama de dispersão para a Argentina (pais=51) dos valores originais de esquerda/direita (esqdir) e valores preditos de esquerda/direita (esqpred2) pela variável independente (privatizar):

```
twoway (scatter esqdir privatizar) ///
      (scatter esqpred2 privatizar, connect(1) clcolor(red) clpattern(dot)) ///
      if pais==51, ///
///
title("Diagrama de dispersão para legisladores da Argentina" ///
      "dos valores observados de esquerda/direita e valores preditos de esquerda/direita" ///
      "pela variável de opinião sobre privatização" ///
      , size(medsmall)) ///
///
ytitle("Escala esquerda/direita")
graph export "C:\cursodcp\graf\pred2arg.wmf", replace
```

Diagrama de dispersão para a Guatemala (pais=52) dos valores originais de esquerda/direita (esqdir) e valores preditos de esquerda/direita (esqpred2) pela variável independente (privatizar):

```
twoway (scatter esqdir privatizar) ///
      (scatter esqpred2 privatizar, connect(1) clcolor(red) clpattern(dot)) ///
      if pais==52, ///
///
title("Diagrama de dispersão para legisladores da Guatemala" ///
      "dos valores observados de esquerda/direita e valores preditos de esquerda/direita" ///
      "pela variável de opinião sobre privatização" ///
      , size(medsmall)) ///
///
ytitle("Escala esquerda/direita")
graph export "C:\cursodcp\graf\pred2gua.wmf", replace
```

Diagrama de dispersão com um gráfico para cada país dos valores originais de esquerda/direita (esqdir) e valores preditos de esquerda/direita (esqpred2) pela variável independente (privatizar):

```
twoway (scatter esqdir privatizar) ///
      (scatter esqpred2 privatizar, connect(1) clcolor(red) clpattern(dot)), ///
///
by(pais) ///
///
caption("Opinião sobre privatização") ///
///
xtitle("") ///
///
ytitle("Escala esquerda/direita")
graph export "C:\cursodcp\graf\pred2multi.wmf", replace
```