

**UNIVERSIDADE FEDERAL DE MINAS GERAIS  
FACULDADE DE FILOSOFIA E CIÊNCIAS HUMANAS  
DEPARTAMENTO DE CIÊNCIA POLÍTICA**

**Tópicos Especiais em Teoria e Análise Política:  
Problema de Desenho e Análise Empírica  
(DCP 859B4)**

**EXERCÍCIO 1**

**Informações gerais:**

Carga horária: 45 horas/aula (3 créditos)

Período: 2º semestre de 2010

Localização: FAFICH/UFMG

Horário: quintas-feiras (14:00 às 17:00)

Sala de aula: 3025

**Professores:**

Magna Inácio (magna.inacio@gmail.com)

Ernesto Friedrich de Lima Amaral (eflamaral@gmail.com)

Site do curso: [www.ernestoamaral.com/dcp859b4-102.html](http://www.ernestoamaral.com/dcp859b4-102.html)

**Informação:**

Com o banco de dados de escolha do próprio aluno, o objetivo deste exercício é de aplicar os conhecimentos apreendidos nas aulas de 01 a 05 da disciplina. O estudante deverá desenvolver os seguintes tópicos:

- 1) Elaborar pergunta de pesquisa e construir hipóteses que serão testadas, além de apresentar banco de dados.
- 2) Realizar análise exploratória dos dados para compreender características das variáveis: centro (média, mediana); variação (desvio padrão), distribuição (tabelas de frequência e histogramas) e *outliers* (diagramas de caixa); mudança no tempo.
  - 2.1) Exibição e análise de tabelas e gráficos (histograma, diagrama de caixa, diagrama de dispersão) com estatísticas descritivas, elaborados com pesos populacionais (“fweight”), quando o banco de dados tiver variável de peso. Utilize estatísticas descritivas apropriadas para variáveis contínuas e categóricas.
  - 2.2) Após diagrama de dispersão, calcular coeficiente de correlação linear entre variáveis (“pweight”), quando pertinente.
- 3) Pensar em uma regressão que contenha somente uma variável independente contínua para explicar uma variável dependente binária.
  - 3.1) Estimar regressão simples com modelos de mínimos quadrados ordinários (MQO), logístico e probit.
  - 3.2) Gerar variável do resíduo de cada uma das regressões acima.
  - 3.3) Elaborar três gráficos de dispersão entre resíduos das regressões acima e variável independente.
  - 3.4) Gerar variável predita de cada uma das regressões acima.
  - 3.5) Elaborar único gráfico de dispersão entre variável predita (por MQO, logístico e probit) e variável independente.
  - 3.6) Regressões precisam utilizar pesos amostrais (“pweight”), quando o banco de dados tiver variável de peso.
  - 3.7) Com base no seu exemplo, explicar impropriedade de usar modelos de mínimos quadrados ordinários com variável dependente binária.

4) Elaborar modelos de regressão multivariados com base nas hipóteses levantadas:

4.1) Utilizar modelagem apropriada (MQO, logístico e/ou probit).

4.2) Análise de diferentes estimativas das regressões, tais como: significância econômica, significância estatística, coeficientes padronizados, análise do  $R^2$ . Algumas destas análises são específicas para o modelo que você utilizar.

4.3) Usar do fator de inflação de variância (VIF) para estimar problemas de multicolinearidade em modelos MQO.

4.4) Utilizar comando “test” e cálculos do “log-likelihood” para testar diferenças entre modelos aninhados.

4.5) Testar e interpretar diferentes formas funcionais das variáveis: logaritmo, funções quadráticas, termos de interação.

Escrever relatório em Word contendo as tabelas e gráficos gerados, assim como análises dos mesmos. O exercício deve ser entregue no dia 07/10/2010.

### **Considerações gerais:**

1) As tabelas descritivas devem ser feitas com peso, usando a opção “fweight”.

2) Os modelos devem ser realizados com peso, usando a opção “pweight”.

3) Evitar colocar tabelas e gráficos que não aumentam o poder de explicação de suas análises.

4) Analisar todas tabelas e gráficos incluídos no trabalho.

5) É possível apresentar todos modelos de regressão em uma única tabela. Escolham um modelo inicial (que ficará na primeira coluna), depois apresentem os demais modelos. Lembrem-se de colocar as estimativas dos betas, os erros-padrão (entre parênteses) e a significância estatística (com asteriscos). A última coluna pode ainda apresentar os betas padronizados do melhor modelo. Explore o comando “outreg2”.

6) Se preocupe com a correta formatação de tabelas e gráficos.

7) Há alguns artigos em meu site (<http://www.ernestoamaral.com/papers.html>) que podem servir de base para apresentação de tabelas e gráficos. Por exemplo, o artigo publicado na “Notas de Población” tem tabelas descritivas e de modelos de regressão. Uma grande quantidade de informações pode ser organizada em uma mesma tabela, de forma bem clara e objetiva. Vocês podem olhar outros artigos neste mesmo site.

8) Os trabalhos deverão ser realizados com o uso do programa computacional STATA. A programação realizada neste programa (arquivo “do”) deverá ser anexada no final de cada exercício.

9) O exercício deve ser elaborado em texto contínuo e não em tópicos sem a interconexão das sentenças (lembrem-se do formato de artigo). O trabalho deve ser realizado com configuração de página A4, margens de 2 cm, fonte Times New Roman de tamanho 12, espaço 1,5. As fontes internas das tabelas podem ter tamanho inferior a 12. Geralmente é usado o tamanho 10 para os números e letras dentro das tabelas e tamanho 8 para o rodapé da tabela.