



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

Trabalho de Análise de Regressão Linear

Informações gerais:

Período: 15 de julho a 02 de agosto de 2013 (segunda-feira a sexta-feira)

Horário: 13:30 às 16:30

Carga horária: 45 horas

Sala de aula: FAFICH 3062

Síte do curso: www.ernestoamaral.com/mq13reg.html

Professor:

Ernesto Friedrich de Lima Amaral

Email: eflamaral@gmail.com

Monitora:

Samantha Haussmann Rodarte Faustino

Email: samantha.haussmann@hotmail.com

Informações gerais:

O trabalho deve ser entregue ao professor e à monitora por email até o dia 05 de agosto de 2013 (segunda-feira). Este trabalho deve ser realizado individualmente.

O trabalho será avaliado com base nos conhecimentos apresentados durante todo o curso. A utilização da maior **quantidade** de técnicas, a preocupação na **qualidade** de suas aplicações e a elaboração de **interpretações coesas** indicarão a nota a ser obtida nesta avaliação. A correta formatação de tabelas e gráficos também será avaliada. A programação em Stata utilizada para este trabalho deve ser incluída no final do documento entregue ao professor.

O trabalho deve ser realizado com configuração de página A4, margens de 2 cm, fonte Times New Roman de tamanho 12, espaço 1,5 e máximo de 10 páginas (excluindo a programação em Stata). 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.

Questões:

1) Com os bancos de dados de pessoas e de domicílios da Pesquisa Nacional por Amostra de Domicílios (PNAD) do Instituto Brasileiro de Geografia e Estatística (IBGE) de 2011 de todo território brasileiro, proponha um conjunto de modelos de regressão linear múltipla com mínimos quadrados ordinários para explicar o rendimento mensal em dinheiro que recebia normalmente, no mês de referência, no trabalho principal da semana de referência (v9532).

1.1) Exibição e análise de tabelas com estatísticas descritivas e resultados da regressão, utilizando os pesos apropriados. Analise diferentes estimativas das regressões, tais como: significância econômica, significância estatística, coeficientes padronizados, análise de multicolinearidade. Teste diferenças entre os modelos estimados com o uso do Teste de F e o R^2 ajustado. Use o fator de inflação de variância (VIF) para estimar problemas de multicolinearidade. Lembre-se de explorar ao máximo as técnicas apreendidas durante o curso.

1.2) Exibição e análise de gráficos que ilustram valores preditos da variável dependente por diferentes variáveis independentes utilizadas.

1.3) Exibição e análise de gráficos de dispersão entre resíduos e demais variáveis (variável dependente original, variável dependente predita e variáveis independentes).

Demais considerações:

- 1) O banco de dados da Pesquisa Nacional por Amostra de Domicílios (PNAD) está disponível nesta página: (<http://www.ibge.gov.br/home/estatistica/populacao/trabalhoerendimento/pnad2011/microdados.shtm>)
- 2) Evitar colocar tabelas e gráficos que não aumentam o poder de explicação de suas análises.
- 3) Analisar todas tabelas e gráficos incluídos no trabalho.
- 4) Se preocupe com a correta formatação de tabelas e gráficos.
- 5) 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 sobre seguro desemprego 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.
- 6) 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.
- 7) As tabelas descritivas devem ser feitas com peso, usando a opção “fweight”.
- 8) Os modelos de regressão devem ser realizados com peso, usando a opção “pweight”.
- 9) O plano amostral complexo da PNAD deve ser levado em consideração, usando o comando “svyset”. Leia o artigo sobre seguro desemprego para aprender como utilizar este comando. A variável de estrato é a v4617 e a variável de unidade primária de amostragem é a v4618, disponíveis no banco de domicílios.
- 10) É possível apresentar vários modelos de regressão em uma única tabela. Escolham um modelo inicial (que ficará na primeira coluna), depois apresentem os demais modelos de maneira progressiva. 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. Utilizem o comando “outreg2” para transferir os modelos de regressão do Stata para o Word (vejam exemplo em outreg2_exemplo.txt, disponível no arquivo Modulos.zip).