**Lista de Exercícios PNAD 2013 – Inferência**

A base de dados “Mulheres.dta” é uma extração de dados da Pesquisa Nacional por Amostra de Domicílios (PNAD), para o ano de 2013 para o estado de São Paulo. Com essa base, a ideia é estimarmos o modelo a seguir para explicar o número de filhos entre as mulheres.

*FilhosTot = β0 + β1AnosdeEstudo + β2log(Renda) + β3Idade + u*

Onde *FilhosTot* é o número de filhos que a mulher indicou ter, *AnosdeEstudos* é a quantidade de anos de estudos, *Renda* é a renda mensal de todas as fontes e *Idade* é idade calculada baseada no ano de nascimento.

1. Quais são as interpretações dos β’s desse modelo?
2. Usando a base de dados “Mulheres” estime o modelo e reporte seus resultados, interpretando cada coeficiente estimado. (Obs: como a PNAD é uma amostra estratificada da população, a regressão tem que ser ponderada pela variável *Peso*, que representa o quanto cada indivíduo da amostra representa da população. Dica: no R, isso é feito com o comando adicional em negrito: lm(y ~ x, **weights=Peso**)). Mantendo tudo o mais constante, quantos anos de estudo é necessário para reduzir o número de filhos em uma unidade?
3. Calcule os valores do teste de t de cada variável explicativa. Usando o teste t-bicaudal a um nível de significância de 5% quais variáveis foram estatisticamente significativas se a Hipótese Nula é que o coeficiente é igual a zero? Suas conclusões se alteram se o nível de significância for 1%? E a 10%? (H0: *β = 0 vs HA: β ≠ 0*). Suas conclusões se alteram se o teste for monocaudal para os três níveis de significância (1%, 5% e 10%)?
4. No modelo estimado teste a significância conjunta (através do teste de F) das variáveis *AnosdeEstudos e log(renda)*, ou seja, sesão conjuntamente iguais a zero, Use um nível de significância de 5%. Apresente o valor de F calculado e, a partir do F tabelado indique sua conclusão.
5. Construa os intervalos de confiança com 95% de confiança de cada variável. O que você pode concluir?
6. O que é o valor de p (p-value ou p-valor)? Baseado em nossas estimações dos itens 3 e 4, o que o p-valor tem a nos dizer?