USP – ICMC – Análise Exploratória de Dados $4^{\underline{a}}$ lista de exercícios – $1^{\underline{o}}/2010$

1. Prove que o valor máximo da área de desigualdade (D) é dado por

$$D_{\max} = \frac{1}{2} \left(1 - \frac{1}{n} \right).$$

2. Desenvolva um código em R para o traçado da curva de Lorenz.

Nos exercícios de 3 a 6 utilize o conjunto de dados Ilocos (data(ilocos)) do pacote ineq (library(ineq)) em R. Os dados foram coletados em domicílios nas Filipinas. O comando ?Ilocos fornece diversas informações sobre as variáveis. Considere as variáveis income, sex, family.size, urbanity e province. É de interesse verificar se a primeira variável (income) está relacionada às demais. Procure levar em conta esta questão de interesse em suas respostas. Comente os aspectos que lhe chamarem atenção.

- 3. Apresente as variáveis utilizando medidas resumo (todas que você considerar pertinentes) e recursos gráficos.
- 4. Procure relações entre a variável family.size e outras que você julgar relevantes.
- 5. Apresente a curva de Lorenz e o índice de Gini da variável renda total do domicílo (income).
- 6. Calcule medidas de associação entre a variável income e as demais variáveis.
- 7. Prove que se y = a + bx e b > 0, então cor(x, y) = 1.
- 8. Prove que se y = a + bx e b < 0, então cor(x, y) = -1.
- 9. Prove que se $a_1 \in \mathbb{R}$, $a_2 \in \mathbb{R}$, $b_1 > 0$ e $b_2 > 0$, então $cor(a_1 + b_1 x, a_2 + b_2 y) = cor(x, y)$.
- 10. Continuando o exercício 9, calcule $cor(a_1 + b_1x, a_2 + b_2y)$ nas seguintes situações: (a) $b_1 > 0$ e $b_2 < 0$, (b) $b_1 < 0$ e $b_2 > 0$ e (c) $b_1 < 0$ e $b_2 < 0$.
- 11. Ilustre graficamente os resultados obtidos nos exercícios 9 e 10.