Problem Set 1 - Econometrics I 2016/2017

Vikrant Vig

Due: October 28th, 2016
(until 18h00 at Dimas' pigeon hole)

Analytical Exercises

1. (10pts) Let X and Y be two random variables. Consider the following definitions:

   (I) X and Y are independent: $f_{X,Y}(X, Y) = f_X(X)f_Y(Y)$;

   (II) X and Y are mean independent: $E(Y|X) = E(Y)$;

   (III) X and Y are linear independent: $Cov(X, Y) = E(XY) - E(X)E(Y) = 0$.

   (a) Show that (I) $\Rightarrow$ (II) $\Rightarrow$ (III).
   (b) *Aside (Extra 5pts):* Can you think of some counter examples to explain why the inverse does not always hold?

2. (15pts) Consider the joint pdf for random variables X and Y below:

$$f(x, y) = \begin{cases} \frac{2}{3}(x + 2y) & 0 < x < 1,\ 0 < y < 1 \\ 0 & \text{else} \end{cases}$$

   (a) Find the marginal pdfs.
   (b) Compute $\mathbb{P}(X \leq \frac{1}{2}|Y = \frac{1}{2})$
   (c) Are X and Y independent?

3. (10pts) Let X be a continuous random variable with the pdf below, where a and b are constants.

$$f(x) = \begin{cases} \frac{1}{b-a} & a < x < b \\ 0 & \text{else} \end{cases}$$

   (a) Compute $\mathbb{E}[X]$.
   (b) Let $g(X) = X^2$. Compute $\mathbb{E}[g(X)]$.

4. (20pts) Suppose that random variables y and x only take the values 0 and 1 and have the following joint probability distribution.

	x=0	x=1
y=0	0.1	0.2
y=1	0.4	0.3

   (a) Find $E(y|x)$, $E(y^2|x)$ and $var(y|x)$ for x=0 and x=1

   (b) Now calculate $E(y)$, $E(y^2)$ and $var(y)$ using the Law of Iterated Expectations.

5. (15pts) Let t be a n x 1 vector and A a n x n matrix, show the following results from matrix calculus.

   (a) $\frac{d(At)}{dt} = A$

   (b) $\frac{d(t'A)}{dt} = A'$

   (c) $\frac{d(t'At)}{dt} = t'(A + A')$

   (d) $\frac{d^2t'At}{dtdt'} = (A + A')$

   (Hint: express the matrices multiplications in index notation and then apply the derivative)

Computational Exercises

6. (15pts) (Basic notions of matlab) Present these results using the Matlab's Publish functionality [1]

   (a) Create the vector a = [1 2 3 4]

   (b) Create matrix $A = \begin{pmatrix} 1 & 2 \\ 3 & 4 \end{pmatrix}$ as well as A^T and A^{-1}.

   (c) Compute A^2 and compare the result to that obtained by typing A ^2

   (d) Create matrix $B = \begin{pmatrix} -1 & 4 \\ 2 & 2 \end{pmatrix}$ and compute A·B.

   (e) Create matrix

$$C = \begin{pmatrix} 0.1 & 0.9 & 0 & 0 & 0 \\ 0.1 & 0 & 0.9 & 0 & 0 \\ 0 & 0.1 & 0 & 0.9 & 0 \\ 0 & 0 & 0.1 & 0 & 0.9 \\ 0 & 0 & 0 & 0.1 & 0.9 \end{pmatrix}$$

   using a loop (see function for).

[1] For more information see:
http://uk.mathworks.com/help/matlab/matlab_prog/publishing-matlab-code.html

(f) Given x = [3 1 5 7 9 2 6], explain what the following commands "mean" by by summarizing the net result of the command.

- x(3)
- x(1:7)
- x(1:end)
- x(1:end-1)
- x(6:-2:1)
- x([1 6 2 1 1])
- sum(x)

7. (15pts) (Basic notions of Stata) Answer the following questions, many of which require use of the dataset "PS1_insurance.xlsx" These data contain personal information on people that participated in a lottery to gain access to the publicly provided healthcare (with winners having lottery=1 in the data). These include birth year, household income, sex, and the number of doctors visits in the period after the lottery.

   (a) Import data and compute summary statistics for each variable including sample count, mean, standard deviation, minimum, maximum, 25th 50th, and 75th quantiles. Tip: see `tabstat` in Stata.

   (b) Generate a new variable for each person's age (as of 2010), and a dummy variable that equals one if the person had at least one doctor visit after the lottery. Summarize these variables.

   (c) To visualize the income distribution of the sample, plot a histogram with bin-widths of $2500. Tip: see `histogram` in Stata.