Problem Set 1

Siria Angino
Federica Romei

24/09/2013

You can work on these problem in group up to 4 students. However, each student must turn in its own write up of the solution to the problems. You have one week to send them from the day they are online. You should send the solutions (including the .r file) to Federica if your surname starts with the letters from A-F, otherwise to Siria. There is no grade for the problem set. Your solutions can be object of discussions during the T.A. sessions.

Remember that, in order to do the midterm, you should send **ALL** the problem sets.

IMPORTANT: Write in the mail object **ASE2013PS1** for the first problem set, **ASE2013PS2** for the second problem set and so forth.

1. For this exercise we will use the dataset "Birthweight.dta". In this dataset you can find:

```
-------------------------------------------------------------------------------------------------
              storage  display     value
variable name   type   format      label      variable label
-------------------------------------------------------------------------------------------------
y               float   %9.0g                 1988 family income, $1000s
cigtax          float  %9.0g                 cig. tax in home state, 1988
cigprice        float  %9.0g                 cig. price in home state, 1988
bw              int    %8.0g                 birth weight, ounces
fatheduc        byte   %8.0g                 father's yrs of educ
motheduc        byte   %8.0g                 mother's yrs of educ
order           byte   %8.0g                 birth order of child
male            byte   %8.0g                 =1 if male child
white           byte   %8.0g                 =1 if white
cigs            byte   %8.0g                 cigs smked per day while preg
bwlbs           float  %9.0g                 birth weight, pounds
cig             float  %9.0g                 packs smked per day while preg
-------------------------------------------------------------------------------------------------
```

Test the following sentence:

(a) The birth weight of the male babies is higher than the one of the female;

(b) White babies tend to weight more than the non white ones;

(c) The average family income in 1988 was 28 thousands dollars;

(d) Family with an income above the average tends to give birth to more healthy babies (with high weight).

2. Consider a standard normal distribution, i.e. N(0,1). Evaluate:

   (a) $P(x \leq -1.96)$

   (b) $P(x \leq -1.64)$

   (c) $P(x \leq 0)$

   (d) $P(x \leq 1.64)$

   (e) $P(|x| \leq 1.96)$

3. Consider two random variables X and Z, with E[X]=2, E[Z]=1, Var[X]=1, Var[Z]=1, a=0.5 and b=3.

   (a) Calculate E[aX+b, cZ+d].

   (b) Calculate Var[aX+b].

   (c) Assuming that X and Z are independent, calculate Var[X+Z] and SD[X+Z].

   (d) Assuming that Cov(X,Z)=1, calculate Cov(aX+b, cZ+d).

   (e) Generalize previous results for any finite E[X], E[Z], var[X], var[Z], a and b.

   (f) Assuming again Cov(X,Z)=1, what can you say about Corr(X,Z)?

4. Given two variables X and Y:

   (a) Show that Corr(a + bX, c + dY) = Corr(X,Y);

   (b) What is the unit of measure for Corr(X,Y)?

5. Let X be a standard normal random variable and $Y = X^2$.

   (a) Show that $E[Y|X] = X^2$.

   (b) Show that E[Y]=1.

   (c) Show that E[XY]=0 (recall that, for a variable $Z \sim N(0,1)$, $E[X^{2n+1} = 0]$ for all n $\in$ N).

   (d) Show that Cov(X,Y)=0, thus Corr(X,Y)=0.

   As you will find out in this example, Corr(X,Y)=0 does not imply E[Y|X]=0. However the opposite is true: $\text{E}[Y|X] = 0 \Rightarrow Corr(X,Y) = 0$.