IMBS 1 **ISCID-CO Dunkerque/ULCO**

Mathematics applied to economics and management

Foundations of Descriptive and Inferential Statistics

October 2015 - Continuous assessment - Semester 1

Time allowed : 1h30 - All documents allowed

Exercice 1 For each question, indicate only one answer (or your answer will be counted as incorrect) and the commands in R you need to use to get the answer. There is no penalty for incorrect answers.

Questions 1. to 3. are based on the following sample of ages (in months) of 18 children at a day care :

36, 42, 18, 32, 22, 22, 25, 29, 30, 31, 19, 24, 35, 29, 26, 36, 24, 28

1. 1pt The median age of the children is 28.5 (answer (d)).

```
> ages<-c(36,42,18,32,22,22,25,29,30,31,19,24,35,29,26,36,24,28)
> median(ages)
[1] 28.5
```

2. 1pt The interquartile range for this data set is given by $Q_3 - Q_1 = 8$ (answer (a)).

```
> quantile(ages,0.75,type=2)-quantile(ages,0.25,type=2)
 75%
   8
```

3. 1pt The standard deviation of the age of children is 6.42 (answer (d)).

```
> sd(ages)
[1] 6.42198
```

Consider the result of a fictional Stat 100 final exam taken by 120 students, as given in the following relative frequency distribution :

Grade	Less than 50	50-59	60-69	70-79	80-89	90-100
Frequency	15%	10%	30%	25%	15%	5%

4. 2pts 45 students received at least a 70 on this exam (answer (b)).

```
> Freq<-c(15,10,30,25,15,5)
> Cum.Freq<-cumsum(Freq)
> Rev.Cum.Freq<-100-cumsum(Freq)+Freq
> class=c("[0,50)","[50,59]","[60,69]","[70,79]","[80,89]","[90,100]")
> tab<-data.frame(Freq,Cum.Freq,Rev.Cum.Freq,row.names=class)
> tab
          Freq  Cum.Freq  Rev.Cum.Freq
 [0,50)     15        15           100
 [50,59]    10        25            85
 [60,69]    30        55            75
 [70,79]    25        80            45
 [80,89]    15        95            20
 [90,100]    5       100             5
> tab["[70,79]","Rev.Cum.Freq"]
[1] 45
```

5. [2pts] According to the empirical rule, approximately 67% of normally distributed data lies within one standard deviation of the mean (answer(b)).

```
> ecdf(ages)(mean(ages)+sd(ages))-ecdf(ages)(mean(ages)-sd(ages))
[1] 0.6666667
```

Questions 6. to 9. are based on the following grouped frequency table of the income, x, of 30 employees at a local small business (in \$1000s).

Income	$26 < x \leq 28$	$28 < x \leq 30$	$30 < x \leq 32$	$32 < x \leq 34$	$34 < x \leq 36$
Frequency	2	11	8	5	4

6. [2pts] The relative cumulative frequency of the $28 < x \leq 30$ class is 0.43 (answer (b))

```
> Freq<-c(2,11,8,5,4)
> Cum.Freq<-cumsum(Freq)
> class<-c("]26,28]","]28,30]","]30,32]","]32,34]","]34,36]")
> tab<-data.frame(Freq,Cum.Freq,Rel.Cum.Freq=Cum.Freq/sum(Freq),row.names=class)
> tab
          Freq  Cum.Freq  Rel.Cum.Freq
 ]26,28]     2         2    0.06666667
 ]28,30]    11        13    0.43333333
 ]30,32]     8        21    0.70000000
 ]32,34]     5        26    0.86666667
 ]34,36]     4        30    1.00000000
> tab["]28,30]","Rel.Cum.Freq"]
[1] 0.4333333
```

7. [1pt] From above, we deduce that the class that contains the 80-th percentile is]32; 34] (answer (d)) :

8. [2pts] Using class-midpoints as representative values, we can estimate the mean for this data as 30, 870\$ (answer (a)).

```
> class<-c(26,28,30,32,34,36)
> center=numeric(length(class)-1)
> for (i in 1 :length(class)-1){
+ center[i]<-(class[i]+class[i+1])/2}
> center
[1] 27 29 31 33 35
> mean(rep(center,Freq))
[1] 30.86667
```

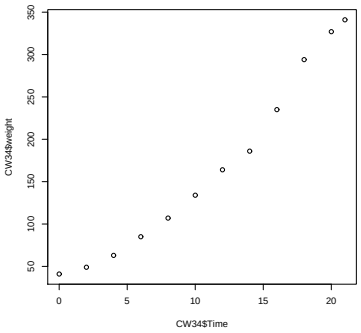
9. [2pts] If the boss' income (the "31-st employee") is \$250,000, the mean income for all 31 workers is approximately equal to 38, 000\$ (answer (c)).

```
> mean(c(rep(center,Freq),250))
[1] 37.93548
```

Exercice 2

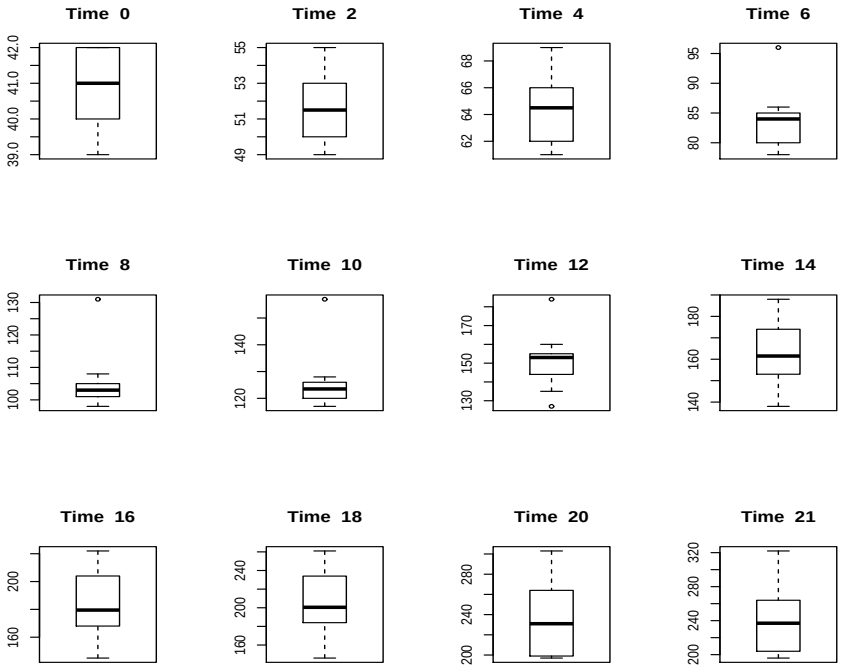
1. [2pts]

```
> CW<-ChickWeight
> subset(CW,CW$Chick==34)->CW34
> plot(CW34$weight~CW34$Time)
```



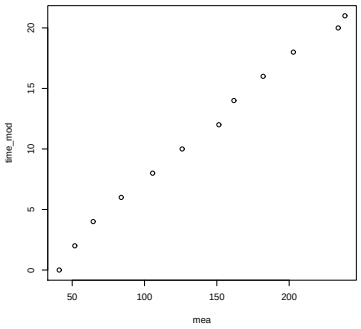
2. 2pts

```
> subset(CW,CW$Diet==4)->CW4
> table(CW4$Time)
  0   2   4   6   8  10  12  14  16  18  20  21
 10  10  10  10  10  10  10  10  10  10   9   9
> time_mod<-c(0,2,4,6,8,10,12,14,16,18,20,21)
> leng<-length(time_mod)
> par(mfrow=c(3,4)) > for (i in 1 :leng){
+ boxplot(CW4$weight[CW4$Time==time_mod[i]],main=paste("Time ",time_mod[i]))}
```



3. 2pts

```
> mea=numeric(leng))
> for (i in 1 :leng)
+ mea[i]<-mean(CW4$weight[CW4$Time==time_mod[i]])
> mea
[1] 41.0000 51.8000 64.5000 83.9000 105.6000 126.0000 151.4000 161.8000 182.0000
[10]  202.9000 233.8889 238.5556
> plot(mea,time_mod)
```


time_mod
0
5
10
15
20
50
100
150
200
mea