MATH 60604A Statistical Modelling

Chapitre 1 Solutions

Juliana Schulz

Question 1

We begin by loading the dataset in `R`:

```
# dataset:
ceo.data<-read.csv("ceocompensation.csv")
head(ceo.data)
```

```
  COMP AGE EDUCATN BACKGRD TENURE EXPER SALES VAL PCNTOWN PROF COMPANY  BIRTH
1 1948  55       1       1     23  23.0  1227 7.6    0.55  145    AdvM    chi
2  809  59       1       2     38   0.5 19196 0.4    0.01  505   aetna    chi
3  721  53       2       1     26   0.5   839 1.5    0.10  -60   aller   sanf
4 2027  62       2       2     25   5.0  8379 3.4    0.04  806    amer  vertx
5 2094  63       1       3     41   8.0 10818 5.9    0.04 1166   ameri bigrun
6  570  60       1       4      3   3.0   804 0.2    0.21   49  anchor philie
```

a) Based on the data, provide an estimate along with a corresponding 99% confidence interval for CEOs' mean compensation (the `COMP` variable).

We can obtain the estimated mean along with a 99% C.I. in many ways. Here we will proceed "manually'':

```
# mean:
mu<-mean(ceo.data$COMP)
# standard deviation:
s<-sd(ceo.data$COMP)
# number of obs
n<-nrow(ceo.data)
# conf interval
c(mu-qt(0.995,df=n-1)*s/sqrt(n),mu+qt(0.995,df=n-1)*s/sqrt(n))
```

```
[1]  897.7102 1345.6298
```

We obtain an estimated mean compensation of $\hat{\mu} = 1121.67$ with 99% C.I. $(897.71, 1345.63)$ (measured in thousands of dollars). Note that we can also obtain these values using the `t.test` function in `R` (see part b)).

b) Do the data indicate that, on average, CEOs make \$1 million? (Hint: you many use your answer from part a).)

Here we are interested in testing the hypotheses

$$H_0 : \mu = 1000 \quad \text{vs.} \quad H_1 : \mu \neq 1000$$

where μ represents the average CEO salary (compensation). *Careful: the compensation variable is measured in thousands of dollars, so that one million dollars corresponds to* $\mu = 1000$. We see that the C.I. from part a) contains the value 1000 and thus we fail to reject H_0 at the $\alpha = 0.01$ level. Similarly, from the t-test (`t.test` function) we obtain a p-value of $p = 0.1568$, and since $p > \alpha$ we fail to reject H_0. Thus, the data do not provide sufficient evidence to conclude that the average CEO compensation is significantly different from one million.

```
t.test(ceo.data$COMP,mu=1000,conf.level=0.99)
```

```
    One Sample t-test

data:  ceo.data$COMP
t = 1.4268, df = 99, p-value = 0.1568
alternative hypothesis: true mean is not equal to 1000
99 percent confidence interval:
  897.7102 1345.6298
sample estimates:
mean of x
  1121.67
```

c) A reader of Forbes magazine comments that CEOs tend to hold their position for at least 5 years. Formally test this using the data (in particular, the `EXPER` variable).

Here we are interested in testing the hypotheses

$$H_0 : \mu \leq 5 \quad \text{vs. } H_1 : \mu > 5$$

where μ represents the average number of years a CEO holds their position as CEO. Note that we are interested in a one-sided test. To carry this test out in `R`, we must specify that `alternative="greater"` as per the direction of the alternative hypothesis (see the code below). The t-test yields a test statistic of 4.79 with a small p-value ($p = 5.914e - 06$) and thus for any reasonable level α we can reject H_0. Thus, the data provide sufficient evidence to conclude that CEO's tend to hold their position for longer than 5 years on average. That is, the data provide evidence in support of the readers comment.

```
t.test(ceo.data$EXPER,mu=5,alterative="greater")
```

```
    One Sample t-test

data:  ceo.data$EXPER
t = 4.7877, df = 99, p-value = 5.914e-06
alternative hypothesis: true mean is not equal to 5
95 percent confidence interval:
```

```
  7.318817 10.601183
sample estimates:
mean of x
     8.96
```