Bonus assignment 2

Ludwig Tranheden
ludtra@student.chalmers.se
9406238791
MVE155

2016/02/10

Contents

1 Problem 8.44

We are given a datasheet containing body temperature and heart rates of 65 females and 65 males. Assuming that the population distributions are normal the task is to for the males and females estimate and calculate:

(i) Means.

(ii) Standard deviations.

(iii) 95% CI's for the means.

We start by sorting the data into three separate arrays: Male, Female and Both. These arrays contains the data of males, females and both females and males. The code is shown below:

```
temp = readtable('bodytemp.txt');
field1 = 'Gender';
field2 = 'Temperature';
field3 = 'Rate';
s = struct(field1,0,field2,0,field3,0);
l=1;
k=1;
Male(1,65) = struct(s);
Female(1,65) = struct(s);
Both(1,130) = struct(s);

for i=1:130
    Both(i) = struct(field1,temp.gender(i),field2,temp.temperature(i),field3,temp.rate(i));
    if temp.gender(i) == 1
       Male(k) = struct(field1,temp.gender(i),field2,temp.temperature(i),field3,temp.rate(i));
       k = k+1;
    else
       Female(l) = struct(field1,temp.gender(i),field2,temp.temperature(i),field3,temp.rate(i));
       l = l+1;
    end
end
```

Using the maximum likelihood estimates of the mean and standard deviation (unbiased), i.e equation 1 & 2. We can estimate the wanted parameters using the matlab function "normfit".

$$\hat{\mu} = \bar{X} \tag{1}$$

$$\hat{\sigma} = \sqrt{\frac{1}{n-1}\sum_{i}^{n}(X_i - \bar{X})^2} \tag{2}$$

The result are shown in table 1.

Parameter	$\hat{\mu}$	$\hat{\sigma}$
Temperature male	98.1046	0.6988
Temperature female	98.3938	0.7435
Heart rate male	73.3692	5.8752
Heart rate female	74.1538	8.1052

Table 1: Estimates of temperature and heart rate parameters of males and females

Moving on to the confidience intervalls we note that the problem states that the population distributions are normal. Extending this with the assumption that the sample is IID we have that

$$\frac{\bar{X} - \mu}{S_{\bar{X}}} \sim t_{n-1}$$

where n is the sample size, that is 65 each case. This means that we can use equation 3 for the 95% CI's.

$$CI_{95\%}: \quad \bar{X} \pm t_{64}(0.025) * S_{\bar{X}} \tag{3}$$

This gives the followoing 4 CI's for the means seen in table 2:

Parameter	CI
Mean temperature male	[97.9315,98.2778]
Mean temperature female	[98.2096,98.5781]
Mean heart rate male	[71.9134,74.8250]
Mean heart rate female	[72.1455,76.1622]

Table 2: Confidence intervalls for the means

None of the confidence intervalls contains the allegedly average temperature of 98.6. Still, we further investigate if you can draw the conclusion that 98.6 really is the average body temperature. To investigate this we do a student's t-test on the mean of both females and males under the hypothesis that the true mean is $\mu_0 = 98.6$. The significance level i choose is $\alpha = 0.05$. The test statistic is:

$$t = \frac{\bar{X} - \mu_0}{S_{\bar{X}}}$$

Which should follow a t_{n-1} distribution if the hypothesis is true. So if the value of t does not belong to the confidence intervall generated by a $\alpha = 0.05$ we can draw the conclusion that the mean of our sample is not μ_0 (Just a simple hypothesis test with H_1 being $\mu \neq \mu_0$). Using matlab we get $t \approx -5.5$ and we see that t lies outside of the confidence intervall we get with 129 df [-1.98,1.98]. The conclusion is that the average body temperature most probably is not 98.6. The code of the problem is displayed below:

```
load Female
load Male
load Both


[myhattemp_male,sigmahattemp_male] = normfit([Male.Temperature]); %Uses maximum likelihood estimates
[myhatrate_male,sigmahatrate_male] = normfit([Male.Rate]);

[myhattemp_female,sigmahattemp_female] = normfit([Female.Temperature]);
[myhatrate_female,sigmahatrate_female] = normfit([Female.Rate]);

n = 65-1;
apl = 0.05/2;

Ci_Maletemp = [mean([Male.Temperature])+tinv(apl,n)*std([Male.Temperature])/sqrt(65)
                mean([Male.Temperature])-tinv(apl,n)*std([Male.Temperature])/sqrt(65)];
Ci_femaletemp = [mean([Female.Temperature])+tinv(apl,n)*std([Female.Temperature])/sqrt(65)
                mean([Female.Temperature])-tinv(apl,n)*std([Female.Temperature])/sqrt(65)];
Ci_Malerate = [mean([Male.Rate])+tinv(apl,n)*std([Male.Rate])/sqrt(65)
                mean([Male.Rate])-tinv(apl,n)*std([Male.Rate])/sqrt(65)];
Ci_femalerate = [mean([Female.Rate])+tinv(apl,n)*std([Female.Rate])/sqrt(65)
                    mean([Female.Rate])-tinv(apl,n)*std([Female.Rate])/sqrt(65)];

% t-test
mu0 = 98.6;
t = (mean([Both.Temperature]) - mu0)/(std([Both.Temperature])/sqrt(130));
Ci = [tinv(apl,129) -tinv(apl,129)];
[h,p,ci] = ttest([Both.Temperature],mu0);
```

Concluding remark:
I have assumed that we are dealing with IID samples which to me seemed logical because otherwise the calculations would have been almost identical of those in bonus assignment 1. Further i have not defined the basic formulas already defined in the first bonus assignment, e.g $S_{\bar{X}}$.

2 Problem 9.64 (c)

In this problem we are working with the same dataset as the task before but now we are to compare the mean body tempratures of males and females separatly against the null hypothesis that the mean is 98,6. I.e we have the following hypothesis:

$$H_0 : \mu = 98.6$$
$$H_1 : \mu \neq 98.6$$

This is basicly what we did in the previous exercise but for both females and males. So without repeating the theory already used we do the same test using the same significance level, repeating the procedure of the previous problem. The code is the same with minor changes for the different data used.

2.0.1 Males

Using the same formulas as above in the previous problem we get that the test statistic is $t = -5.7158$ and this falls outside the generated confidence intervall with 64 df: [-1.9977,1.9977]. Furthermore the p value is $\approx 3 * 10^{-7}$ and we conclude that the mean body temperature of males is not 98,6; it's probably lower.

2.0.2 Females

Doing the same for females we get $t = -2.2355$ which again falls outside the confidence intervall (same as for males). But now our p-value is 0.0289, so if we where to lower the significance lever to $\alpha = 0.01$ we can not reject the nullhypothesis. Therefore it's not as sure as in the male case that we can reject the nullhypothesis. But standing by the significance level choosen initially we conclude that the mean body temperature of females is not 98,6.