

IST772 Problem Set 5

Shivani Sanjay Mahaddalkar

Attribution statement: (choose only one) 1. I did this homework by myself, with help from the book and the professor

Chapter 5, Exercise 6

The built-in warpbreaks data set (see "? warpbreaks" for documentation) contains data for the number of warp thread breaks per loom with different amounts of tension (we will not consider the type of wool). The tensions are labelled "L", "M" or "H" for low, medium and high tension. As a reminder, these subsetting statement accesses the breaks data for the low and medium tension:

```
warpbreaks$breaks[warpbreaks$tension=="L"]  
warpbreaks$breaks[warpbreaks$tension=="M"]
```

Run a t-test to compare the means of the breaks for medium tension ("M") and breaks for high tension ("H") in the warpbreaks data. Report the observed value of t, the degrees of freedom, and the p-value associated with the observed value (1 pt). Assuming an alpha threshold of .05, decide whether you should reject the null hypothesis that the means are equal or fail to reject the null hypothesis (1 pt). In addition, report the upper and lower bound of the confidence interval (1 pt).

```
t.test(warpbreaks$breaks[warpbreaks$tension=="M"],  
warpbreaks$breaks[warpbreaks$tension=="H"])  
  
##  
##  Welch Two Sample t-test  
##  
## data:  warpbreaks$breaks[warpbreaks$tension == "M"] and  
warpbreaks$breaks[warpbreaks$tension == "H"]  
## t = 1.6199, df = 33.74, p-value = 0.1146  
## alternative hypothesis: true difference in means is not equal to 0  
## 95 percent confidence interval:  
##  -1.203597 10.648042  
## sample estimates:  
## mean of x mean of y  
##  26.38889  21.66667
```

The observed value of t from the t-test is 1.6199, with 33.74 degrees of freedom and p-value of 0.1146. Assuming an alpha threshold of 0.05, p-value(0.1146) is greater than alpha, which means that we fail to reject the null hypothesis that the means are equal. For this instance, the lower bound and the upper bound of the confidence interval is estimated

to be 0.8976796 and 19.1023204. However, there is no real way of knowing whether the population mean difference lies in this CI.

Chapter 5, Exercise 7

Install and library() the BEST package. Note that you may need to install a program called JAGS onto your computer before you try to install the BEST package inside of R. Use BESTmcmc() to compare the warp breaks for medium tension ("M") to breaks for high tension ("H") (1 pt). Plot the result and document the boundary values that BESTmcmc() calculated for the HDI (1 pt). Write a brief definition of the meaning of the HDI and interpret the results from this comparison (1 pt).

```
#install.packages("BEST") #Commented out because it has been installed already
```

```
library("BEST")
```

```
## Warning: package 'BEST' was built under R version 4.0.4
```

```
## Loading required package: HDInterval
```

```
## Warning: package 'HDInterval' was built under R version 4.0.4
```

```
breaksHM <- BESTmcmc(warpbreaks$breaks[warpbreaks$tension=="M"],  
warpbreaks$breaks[warpbreaks$tension=="H"])
```

```
## Waiting for parallel processing to complete...
```

```
## done.
```

```
breaksHM
```

```
## MCMC fit results for BEST analysis:
```

```
## 100002 simulations saved.
```

```
##           mean      sd median  HDIlo HDIup  Rhat n.eff
```

```
## mu1      26.286  2.423 26.272 21.454 31.07 1.000 59882
```

```
## mu2      21.320  2.135 21.295 17.097 25.55 1.000 56316
```

```
## nu       35.287 30.014 26.630  1.739 94.64 1.002 18784
```

```
## sigma1   9.657  1.919  9.396  6.273 13.47 1.000 40619
```

```
## sigma2   8.498  1.804  8.274  5.325 12.18 1.000 34188
```

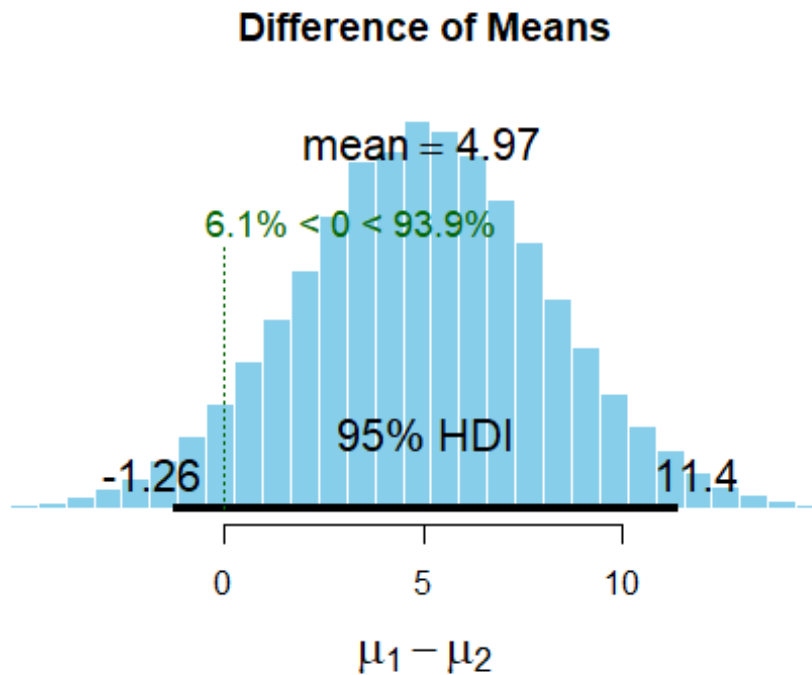
```
##
```

```
## 'HDIlo' and 'HDIup' are the limits of a 95% HDI credible interval.
```

```
## 'Rhat' is the potential scale reduction factor (at convergence, Rhat=1).
```

```
## 'n.eff' is a crude measure of effective sample size.
```

```
plot(breaksHM)
```



The boundary values for the HDI(high density interval) are -1.4 and 11.3. It means that there is a 95% probability that the population mean difference between the two groups falls between -1.4 and 11.3.

Chapter 5, Exercise 8

Compare and contrast the results of Exercise 6 and Exercise 7. You have three types of evidence: the results of the null hypothesis test, the confidence interval, and the HDI from the BESTmcmc() procedure. Each one adds something, in turn, to the understanding of the difference between tensions. Explain what information each test provides about the comparison of the breaks for medium tension and for high tension (1 pt for each report and interpretation).

1. With the t-test we could ascertain that we failed to reject the null hypothesis(that the population mean difference between the two groups is zero). We have not accepted the null hypothesis, we just do not have evidence to either accept or reject it.This means the difference in the tensions could be zero.
2. With the confidence interval from a single t-test we cannot say anything with certainty about the population mean difference or its probable range.
3. The HDI from the BESTmcmc() procedure however gives the lower and upper bounds of the population mean ie -1.35 and 11.4 with 95% probability that the population mean lies between the HDI. From the graph it can be seen that 93.9% of the values are greater than 0, which means that the probability of the mean break for population for

M thread is greater than that of H. Which validates their names of Medium and High tensions effectively.

Chapter 5, Exercise 9

Using the same warpbreaks data set, compare the breaks for medium tension to the breaks for low tension ("L"). Use all of the methods described earlier (t-test, confidence interval, and Bayesian method) and explain all of the results (1 pt for each report and interpretation).

```
t.test(warpbreaks$breaks[warpbreaks$tension=="M"],
warpbreaks$breaks[warpbreaks$tension=="L"])

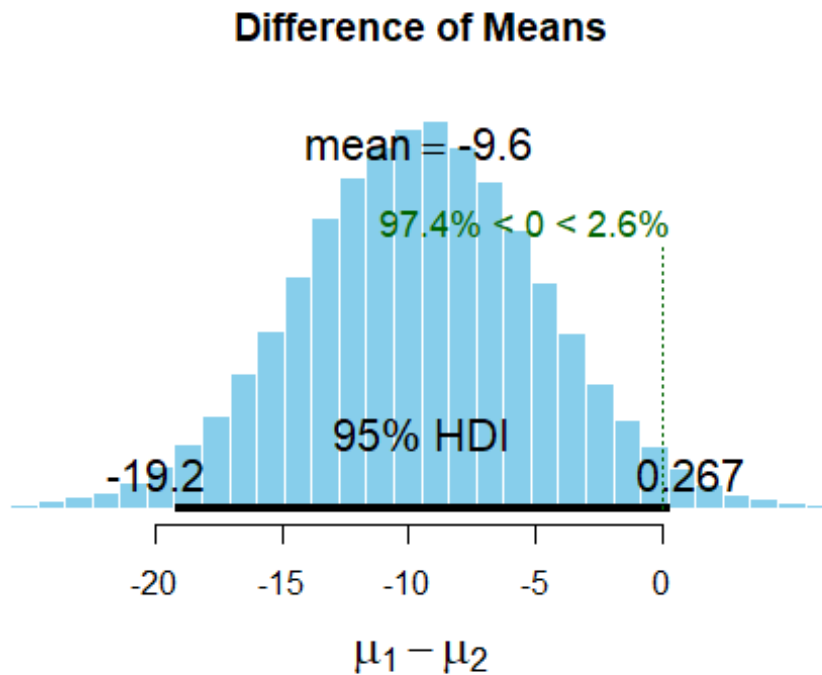
##
##  Welch Two Sample t-test
##
## data:  warpbreaks$breaks[warpbreaks$tension == "M"] and
warpbreaks$breaks[warpbreaks$tension == "L"]
## t = -2.256, df = 26.554, p-value = 0.03252
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -19.1023204  -0.8976796
## sample estimates:
## mean of x mean of y
##  26.38889  36.38889
```

The t-test gives a t value of -2.256 with 26.554 degrees of freedom and p-value of 0.03252. The p-value is greater than the threshold alpha therefore we reject the null hypothesis. This means that there is a difference in the number of breaks between the population of M and L threads. However, the t-test does not give us any further direction. The confidence interval in this instance is between -19.1023204 and -0.8976796. Since it is a single run, we cannot say with certainty whether the population mean difference lies in the confidence interval.

```
breaksML <- BESTmcmc(warpbreaks$breaks[warpbreaks$tension=="M"],
warpbreaks$breaks[warpbreaks$tension=="L"])

## Waiting for parallel processing to complete...done.

plot(breaksML)
```



The Bayesian method gives an HDI of (-19.5, 0.0184), which means that there is a 95% probability that the population mean difference lies within the HDI. It also gives another bit of information that the value of the difference is lesser than zero 97.4% of times which means it is very likely that the population mean difference is negative and the M category breaks less number of times than L category.

Chapter 5, Exercise 10

Consider this t-test, which compares two datasets of $n = 100,000$ observations each:

```
t.test(rnorm(100000,mean=17.1,sd=3.8),rnorm(100000,mean=17.2,sd=3.8))

##
##  Welch Two Sample t-test
##
## data:  rnorm(1e+05, mean = 17.1, sd = 3.8) and rnorm(1e+05, mean = 17.2,
## sd = 3.8)
## t = -5.936, df = 2e+05, p-value = 2.925e-09
## alternative hypothesis: true difference in means is not equal to 0
## 95 percent confidence interval:
##  -0.13406769 -0.06751001
## sample estimates:
## mean of x mean of y
## 17.10052 17.20131
```

For rnorm() command was used to generate a random normal distribution of observations. The only difference between the two is that in the first rnorm() call, the mean is set to 17.1 and in the second it is set to 17.2. I think you would agree that this is a negligible difference, if we are discussing breaks. Run this line of code and comment on the results of the t-test (1 pt). What are the implications in terms of using the NHST on very large data sets (e.g., is it a good idea)? (1 pt)

The p-value in this case is much lesser than 0.05 which means we reject the null hypothesis that the population mean difference is zero. However, objectively, this is true, in case of smaller differences, where we can afford to neglect them, it does not help us. It should not be giving us a p-value this small to say that the two samples do not belong to the same group. On running the same distribution for a sample size of 15 each, it gives a p-value of 0.7009 which is greater than 0.05, which would mean that it fails to reject the null hypothesis. This is a more acceptable outcome in our case. This implies that the NHST does not give a clearer picture in case of larger datasets.

As mentioned in the text-book about the number 1 concern with NHST is that it gives a go/no-go decision. In cases like these, when we might need a more probabilistic approach to discard a null hypothesis, NHST shows its limitations.

```
t.test(rnorm(15,mean=17.1,sd=3.8),rnorm(15,mean=17.2,sd=3.8))  
##  
##  Welch Two Sample t-test  
##  
## data:  rnorm(15, mean = 17.1, sd = 3.8) and rnorm(15, mean = 17.2, sd =  
3.8)  
## t = -0.63517, df = 26.534, p-value = 0.5308  
## alternative hypothesis: true difference in means is not equal to 0  
## 95 percent confidence interval:  
##  -3.247739  1.713258  
## sample estimates:  
## mean of x mean of y  
##  15.76379  16.53103
```