Exercises 5

Jan Hohenheim

Setup

```
rm(list = ls())
library(ggplot2)
library(ggthemes)
library(dplyr)
```

```
##
## Attaching package: 'dplyr'

## The following objects are masked from 'package:stats':
##
##     filter, lag

## The following objects are masked from 'package:base':
##
##     intersect, setdiff, setequal, union
```

```
library(purrr)
library(cowplot)
```

```
##
## Attaching package: 'cowplot'

## The following object is masked from 'package:ggthemes':
##
##     theme_map
```

```
library(latex2exp)
library(glue)

theme <-
  theme_solarized() #+
  #theme(text = element_text(family = "Helvetica Neue"))
```

```
HbSS <- c(7.2, 7.7, 8, 8.1, 8.3, 8.4, 8.4, 8.5, 8.6, 8.7, 9.1, 9.1, 9.1, 9.8, 10.1,
10.3)
HbSb <- c(8.1, 9.2, 10, 10.4, 10.6, 10.9, 11.1, 11.9, 12, 12.1)
Hb <- data.frame(level = c(HbSS, HbSb),
                 category = c(rep("HbSS", length(HbSS)), rep("HbSb", length(HbSb))))
print("HbSS")
```

```
## [1] "HbSS"
```

```
Hb |>
  filter(category == "HbSS") |>
  select(level) |>
  glimpse()
```

```
## Rows: 16
## Columns: 1
## $ level <dbl> 7.2, 7.7, 8.0, 8.1, 8.3, 8.4, 8.4, 8.5, 8.6, 8.7, 9.1, 9.1, 9.1,~
```

```
print("HbSb")
```

```
## [1] "HbSb"
```

```
Hb |>
  filter(category == "HbSb") |>
  select(level) |>
  glimpse()
```

```
## Rows: 10
## Columns: 1
## $ level <dbl> 8.1, 9.2, 10.0, 10.4, 10.6, 10.9, 11.1, 11.9, 12.0, 12.1
```

Problem 17

Use again the sickle-cell disease data introduced in Problem 13 (Worksheet 4). For the cases listed below, specify the null and alternative hypothesis. Then use R to perform the tests and give a careful interpretation.

a) $\mu_{\text{HbS}\beta} = 10$ ($\alpha = .05$, two-sided)

- $H_0 : \mu_{\text{HbS}\beta} = 10$
- $H_1 : \mu_{\text{HbS}\beta} \neq 10$

```
HbSb |>
  t.test(mu = 10, conf.level = 0.95)
```

```
##
##  One Sample t-test
##
## data:  HbSb
## t = 1.5514, df = 9, p-value = 0.1552
## alternative hypothesis: true mean is not equal to 10
## 95 percent confidence interval:
##   9.711386 11.548614
## sample estimates:
## mean of x
##     10.63
```

10 is in the CI, accept H_0

b) $\mu_{\text{HbS}\beta} = \mu_{\text{HbSS}}$ ($\alpha = .001$, two-sided)

- $H_0 : \mu_{\text{HbS}\beta} = \mu_{\text{HbS}}$
- $H_1 : \mu_{\text{HbS}\beta} \neq \mu_{\text{HbS}}$

```
HbSb |>
  t.test(HbSS, conf.level = 0.999)
```

```
##
##  Welch Two Sample t-test
##
## data:  HbSb and HbSS
## t = 4.1896, df = 13.913, p-value = 0.0009205
```

```
## alternative hypothesis: true difference in means is not equal to 0
## 99.9 percent confidence interval:
##  0.01954312 3.81545688
## sample estimates:
## mean of x mean of y
##   10.6300    8.7125
```

0 is not in CI, reject H_0

(c) What changes, if one-sided tests are performed instead?

The H_0 doesn't change, but the H_1 can be either: - left-tailed: $H_1^{'} : \mu < n$ or - right-tailed: $H_1^{''} : \mu > n$

```r
HbSb |> t.test(mu = 10, conf.level = 0.95, alternative = "less")
```

```
##
##  One Sample t-test
##
## data:  HbSb
## t = 1.5514, df = 9, p-value = 0.9224
## alternative hypothesis: true mean is less than 10
## 95 percent confidence interval:
##      -Inf 11.37439
## sample estimates:
## mean of x
##     10.63
```

```r
HbSb |> t.test(mu = 10, conf.level = 0.95, alternative = "greater")
```

```
##
##  One Sample t-test
##
## data:  HbSb
## t = 1.5514, df = 9, p-value = 0.07761
## alternative hypothesis: true mean is greater than 10
## 95 percent confidence interval:
##  9.885612      Inf
## sample estimates:
## mean of x
##     10.63
```

```r
HbSb |> t.test(HbSS, conf.level = 0.999, alternative = "less")
```

```
##
##  Welch Two Sample t-test
##
## data:  HbSb and HbSS
## t = 4.1896, df = 13.913, p-value = 0.9995
## alternative hypothesis: true difference in means is less than 0
## 99.9 percent confidence interval:
##      -Inf 3.653291
## sample estimates:
## mean of x mean of y
##   10.6300    8.7125
```

```r
HbSb |> t.test(HbSS, conf.level = 0.999, alternative = "greater")
```

```
##
```

```
##  Welch Two Sample t-test
##
## data:  HbSb and HbSS
## t = 4.1896, df = 13.913, p-value = 0.0004602
## alternative hypothesis: true difference in means is greater than 0
## 99.9 percent confidence interval:
##  0.1817091       Inf
## sample estimates:
## mean of x mean of y
##   10.6300    8.7125
```

All one-sided test variations accept H_0, except for the last, which leads us to accept the alternative $H_1 : \mu_{\text{HbS}\beta} > \mu_{\text{HbSS}}$. ## Problem 18 Anorexia is an eating disorder that is characterized by low weight, food restriction, fear of gaining weight and a strong desire to be thin. The dataset *anorexia* in the package *MASS* gives the weight of 29 females before and after a cognitive behavioral treatment (in pounds). Test whether the treatment was effective.

- $H_0 : \mu_{\text{control}} = \mu_{\text{cbt}}$
- $H_1 : \mu_{\text{control}} < \mu_{\text{cbt}}$
  using $\alpha = 0.05$

```
library(MASS)
```

```
##
## Attaching package: 'MASS'

## The following object is masked from 'package:dplyr':
##
##     select
```

```
?anorexia

anorexia.delta <- anorexia |>
  mutate(Deltawt = Postwt - Prewt)

control <- anorexia.delta |>
  filter(Treat == "Cont")
cbt <- anorexia.delta |>
  filter(Treat == "CBT")

alpha <- 0.05
t.test(control$Deltawt, cbt$Deltawt, alternative = "less", conf.level = 1 - alpha)
```

```
##
##  Welch Two Sample t-test
##
## data:  control$Deltawt and cbt$Deltawt
## t = -1.6677, df = 50.971, p-value = 0.05075
## alternative hypothesis: true difference in means is less than 0
## 95 percent confidence interval:
##       -Inf 0.015656
## sample estimates:
## mean of x mean of y
## -0.450000  3.006897
```

$p > 0.05$, so accept H_0: The CBT treatment had no significant effect.