Homework 2

Frank Woodling

February 23, 2016

2.01.

```
library('gtools')
```

```
## Warning: package 'gtools' was built under R version 3.2.3
```

```
t1 <- c(1, 3, 5, 7)
t2 <- c(2, 4, 6, 8)

t1 <- c(2, 4, 6, 11)
t2 <- c(3, 5, 7, 12)
t = c(t1, t2)

teststat.obs = sum(t1)
teststat.obs
```

```
## [1] 23
```

```
combo = combinations(8, 4, v = t, set = F, repeats.allowed = F)

W = NULL

for (i in 1:70)
{
  W = c(W, sum(combo[i,]))
}

sum(teststat.obs >= W)/choose(8, 4)
```

```
## [1] 0.3714286
```

Before working out I thought I should get 0.33 or 0.5 for my p-value. After plugging some numbers in I ended up with 0.3429 the first time and 0.3714 the second time.

2.05

```
# library('gtools')
woodland <- read.table('woodland.txt', header=T)
woodland
```

```
##    obs species
## 1  5.1       A
## 2  9.4       A
## 3  7.2       A
## 4  8.1       A
## 5  8.8       A
## 6  2.5       B
## 7  4.2       B
## 8  6.9       B
## 9  5.5       B
## 10 5.3       B
```

```
attach(woodland)
ranks = rank(obs)

w.obs = sum(ranks[species=="A"])
a.ranks = combinations(10, 5, v = ranks, set = F, repeats.allowed = F)
W = rep(NA, choose(10,5))

for(i in 1:choose(10,5))
{
  W[i] = sum(a.ranks[i,])
}

sum(W > w.obs)/choose(10, 5)
```

```
## [1] 0.01587302
```

With a p-value of 0.01587302 we can reject the null hypothesis that there is no difference between the two species' nesting heights.

2.12 complete first part, hodges-lehmann wrong (typo?)

```
# From the table in the book the Hodges-Lehmann estimate should be from pairwise
# difference 5 to 25.

carapace <- read.table('carapace.txt', header=T)
carapace
```

```
##    obs   section
## 1    5 Section_1
## 2   11 Section_1
## 3   16 Section_1
## 4    8 Section_1
## 5   12 Section_1
## 6   17 Section_2
## 7   14 Section_2
## 8   15 Section_2
## 9   21 Section_2
## 10  19 Section_2
## 11  13 Section_2
```

```r
attach(carapace)
```

```
## The following object is masked from woodland:
##
##     obs
```

```r
s1 = c(5, 11, 16, 8, 12)
s2 = c(17, 14, 15, 21, 19, 13)

pwd = NULL

for(i in 1:5)
{
  for(j in 1:6)
  {
    pwd = c(pwd, (i-j))
  }
}

sort(pwd)
```

```
##  [1] -5 -4 -4 -3 -3 -3 -2 -2 -2 -2 -1 -1 -1 -1 -1  0  0  0  0  0  1  1  1
## [24]  1  2  2  2  3  3  4
```

```r
hodges_ci = c(pwd[5], pwd[25])
hodges_ci
```

```
## [1] -4  4
```

```r
m = 5
n = 6
std_err = sqrt(((m-1)*sd(s1)+(n-1)*sd(s2))/(n+m-2))
t = abs(qt(.05, 9, lower.tail = TRUE))
upper = (mean(s1)-mean(s2)) + t*std_err*sqrt((1/m)+(1/n))
lower = (mean(s1)-mean(s2)) - t*std_err*sqrt((1/m)+(1/n))
upper
```

```
## [1] -4.005374
```

```r
lower
```

```
## [1] -8.194626
```

The Hodge-Lehmann confidence interval is (-4, 4). The normal confidence interval is (-8, -4). I cannot figure out why these are not giving me reasonable confidence intervals. I believe if anything it must be a problem with my loop since many of those numbers are negative.

2.19