STATS 205: Homework Assignment 3 (Spring 2019)

4/19/2019

Solve these problems from the textbook *HWC* available here.

Send your Rmd and PDF files to pjeganat [at] stanford [dot] edu.

Due on 4/26/2019 at 1.30 p.m.

Friendly reminder: **Mid-term project proposal due on 5/3/2019 at 1.30 p.m.**

1) In Lecture 6 (Bootstrap II), we saw an example of LSAT and GPA from 15 students. The data is in an R package `bootstrap`.

```
library(bootstrap); data(law)
t(law)
```

```
##           1     2      3      4      5      6   7      8      9     10
## LSAT 576.00 635.0 558.00 578.00 666.00 580.00 555 661.00 651.00 605.00
## GPA    3.39   3.3   2.81   3.03   3.44   3.07   3   3.43   3.36   3.13
##          11     12     13     14     15
## LSAT 653.00 575.00 545.00 572.00 594.00
## GPA    3.12   2.74   2.76   2.88   2.96
```

We are interested in inference on the correlation between GPA and LSAT. The point estimate of correlation coefficient is

```
theta.hat = cor(law$LSAT, law$GPA); theta.hat
```

```
## [1] 0.7763745
```

We know that the all possible distinct bootstrap samples can be defined by the weight vector.

```
library(partitions)
n = 15
allCompositions = compositions(n, n);allCompositions[,1:5]
```

```
##       [,1] [,2] [,3] [,4] [,5]
##  [1,]   15   14   13   12   11
##  [2,]    0    1    2    3    4
##  [3,]    0    0    0    0    0
##  [4,]    0    0    0    0    0
##  [5,]    0    0    0    0    0
##  [6,]    0    0    0    0    0
##  [7,]    0    0    0    0    0
##  [8,]    0    0    0    0    0
##  [9,]    0    0    0    0    0
## [10,]    0    0    0    0    0
## [11,]    0    0    0    0    0
## [12,]    0    0    0    0    0
## [13,]    0    0    0    0    0
## [14,]    0    0    0    0    0
## [15,]    0    0    0    0    0
```

Choose randomly $R = 10,000$ columns of allCompositions.

```
allCompositions.sub = allCompositions[, sample(1:dim(allCompositions)[2], size = 10000, replace = FALSE)
```

(i) Compute bootstrap replicates of correlation (use R codes in Lecture 6).
(ii) Compute standard error of correlation coefficient estimate using these bootstrap replicates.

2) (i) **HWC** Page 21, Problem 9 (discrete data with two categories.)

(ii) In addition, find the power of the test if $p = .7$ and $p = .8$.

3) **HWC** Page 24, Problem 11 (estimate for p and standard error for $\hat{p}$).

4) **HWC** Page 31, Problem 15 (confidence interval for p).

5) **HWC* Page 33, Problem 23 (Chi-squared test for more than two categories).

6) **HWC* Page 513, Problem 13 (Fisher's exact test on difference on probabilities. Use `fisher.test()` function in R).

7) **HWC* Page 509, Problem 3 (approximate confidence interval for difference of proportions)

8) **HWC* Page 510, Problem 6 (Matched Pairs problem. Use `mcnemar.test()` function in R).

9) **HWC* Page 510, Problem 9 (2×2 Chi-squared test for homogeneity).