

STAT 406: HW2

- All computer code should be written using the language R. Type ALL your code into one PLAIN Text format file. Plain text format is available by default in R. Please do not use Microsoft Word .doc format or .rtf format or .pdf format. Inside your plain text file, make sure you identify each problem in a comment placed at the beginning of the problem. The file name should match your name as in 'JohnDoe.R'. Submit your R code file online (under Assignments) at or before the due date, and hand in a hard copy of the code with any additional material to your lab.
- I recommend that before submitting your homework, you also create a new directory and run your R code, to make sure that it is self-contained and runs as you intended.

1. Let U_1, U_2, \dots be independent and identically distributed random variables with uniform distribution $\mathcal{U}(0, 1)$. Define the following random variable

$$N = \min \left\{ n \geq 1 : \sum_{k=1}^n U_k > 1 \right\}.$$

Write an R function that takes an integer argument x and returns a vector (N_1, \dots, N_x) where each N_i is an independent realization of the random variable N defined above.

2. Here is a version of the game of crap. First, you roll two well-balanced, six-sided dice; let x be the sum of the dice of the first roll. If $x = 7$, or $x = 11$ you win, otherwise you keep rolling until either you get x again, in which case you also win, or until you get a sum of 7 or 11 in which case you lose.

Write a function that takes no input, and simulate the game of crap.

3. In the US Congress, a Representative who introduces a bill or other measure on the floor is called its sponsor. Other members can also add

themselves to the bill proposal as co-sponsors if they support the idea. Thus co-sponsorship of a bill is an indication of political proximity. We can also view co-sponsorship as a proxy for political influence in the sense that more influential senators may tend to draw more co-sponsorship. We focus on the senate. The goal of this problem is to quantify the political influence of each senator by computing the number of co-sponsorships they can draw.

The file `senmatrix.txt` contains the sponsorship and co-sponsorship data during the 110-th Senate. Use the following code to load the file in R.

```
senmat<-read.table('senmatrix.txt',header=F,sep=',')
```

Each row of `senmat` represents a senator and each column represents a bill. The entry `senmat[i,j]` can take the values:

- (a) 1 - if senator i sponsored bill j
- (b) 2 - if senator i cosponsored bill j
- (c) there are other possible values, but we will ignore those.

We define the political influence of a given senator i as

$$V_i = \frac{S_i}{b_i},$$

where S_i is the total number of co-sponsorship won by senator i and b_i is the total number of bills sponsored by i . We set $V_i = 0$ if $b_i = 0$. Write code to obtain the vector V , of length N (N is the total number of senators), where $V[i]$ is the political influence of i . The file `senators.txt` contains the names of the senators of the 110-th Senate in the same order as in the file `senmatrix.txt`. Write code to print the 5 most influential senators, and the 5 least influential senators, according to your statistic V .