

Rolling Two Dice Examples Using R
Stat 341 - Fall 2008

This is the R code that looks at two different random variables obtained from the sample space of the experiment of rolling two 6-sided dice: the sum of the values on the two dice and the largest or maximum value of the two dice. Using R, we will study the observed probability distributions of these random variables.

To begin, we need to create a virtual dice in R. The code is

```
dice<- c(1:6)
```

We would then like to roll the dice twice and observe the outcome and then repeat the process 10,000 times. The R code to do this is

```
dice1<- sample(dice,10000,replace = T)
dice2<- sample(dice,10000,replace = T)
```

Now we have 10000 rolls of our two dice in the variables **dice1** and **dice2**. But they are in separate variables and we need to join them together into a matrix. Each dice will become a column in the matrix and each roll of the 2 dice will become a row in the matrix. Here is the R code to do this.

```
dicematrix<- cbind(dice1,dice2)
```

At this point, it may be helpful to look at the matrix **dicematrix**. To look at the first 20 rows, type in the command

```
dicematrix[1:20,]
```

Now we would like to look at the values in each row (the outcomes of the rolls of our 2 dice) and calculate the sum of the two dice and the largest or maximum of the two dice. In R, whenever you have a matrix, you can apply a mathematical function (like **sum**) to each row or column of the matrix with the same command. In this case, we would like to apply two different functions (**sum** and **max**) to each of our 10000 rows of the matrix **dicematrix**. The commands are

```
sumtwodicesim<- apply(dicematrix,1,sum)
maxtwodicesim<- apply(dicematrix,1,max)
```

In the commands above, the first value is the name of the matrix, the second value of 1 specifies we want to apply the function to the rows and the third value is the name of the function we want to apply.

We can then study the 10000 observed values of these two random variables using histograms and summary statistics. For example, to find the smallest and largest values of the observed sums, you can type

```
min(sumtwodicesim)
max(sumtwodicesim)
```

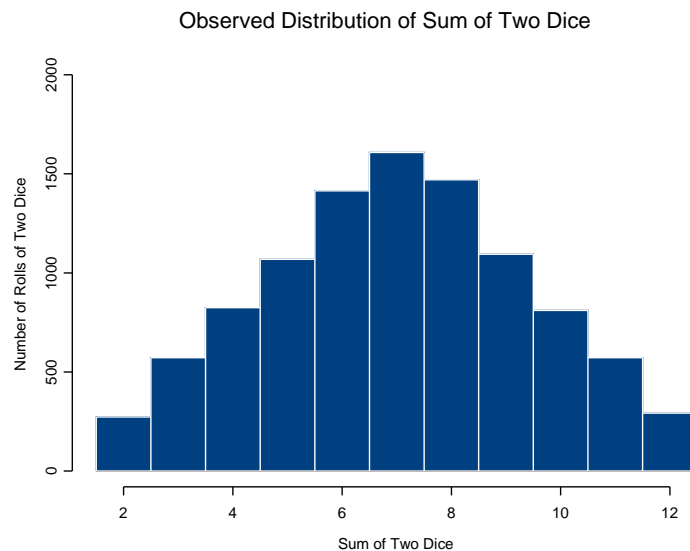
To find other summary statistics of the observed sums, like the mean, median, five number summary, and standard deviation, you can type

```
mean(sumtwodicesim) #mean
sqrt(var(sumtwodicesim)) #std. dev.
fivenum(sumtwodicesim) #five number summary
```

To get a picture of the observed values, you can make a histogram. For these values, you should set up the histogram so that the observed values are centered in the bars of the histogram. For example, for the observed sums, you should set up your histogram as

```
sumdicebreaks<- c(1:12) + 0.5
hist(sumtwodicesim, breaks = sumdicebreaks)
```

Here is a picture of the observed distribution of the sums.



Similar code in R will give you the histogram and summary statistics for the largest or maximum values of the two dice.

```
min(maxtwodicesim) #minimum observed value
max(maxtwodicesim) #maximum observed value
mean(maxtwodicesim) #mean
sqrt(var(maxtwodicesim)) #std. dev.
fivenum(maxtwodicesim) #five number summary
maxdicebreaks<- c(0:6) + 0.5 #set the breaks for the largest value
hist(maxtwodicesim, breaks = maxdicebreaks)
```

Here is a picture of the observed distribution of the maximum values.

