1 Simulating Dice Rolls

1.1 Simulating 6000 dice rolls

Here’s the R code for the simulation:

```r
trials <- sample(6, 6000, replace = T)
# Note that we have to sample with replacement.
table(trials)
```

The output is:

```
trials
1 2 3 4 5 6
1000 1023 1011 1007 981 978
```

As the output says, we often get numbers that are not very close to 1000 (i.e., in the range 995-1005) even though we are assuming a uniform distribution. This claim is in fact strengthened by the next question which asks to estimate the number of times that one would expect to roll more than 1030 ones.

1.2 Estimating the number of 1’s

```r
{ 
count = 0
for(i in 1:10000)
{
    trials <- sample(6, 6000, replace = T)
    count[i] <- sum(trials == 1)
    # note the usage of sum command
}
sum(count > 1030)
}
```

And the output was: 1470.

So the simulation says that we should expect to roll more than 1030 around 14.7 ≈ 15% of the times.

Let’s try to verify this claim:

Let X denote the total number of 1’s in the 6000 rolls. Then we know that

\[ X \sim \text{Bin}(6000, \frac{1}{6}) \]
Now we need to find out $P(X > 1030)$. Since $X$ is discrete this is by definition is

$$P(X > 1030) = \sum_{i=1031}^{6000} P(X = i)$$

The above sum could be easily evaluated by a small R code, but we also note that we can do it without a computer using the \textbf{Normal} approximation of the binomial distribution.

```r
s = 0
for(i in 1031:6000)
{
  s = s + dbinom(i,6000,1/6)
  # dbinom gives the binomial density taking x,n and p respectively
}
```

And the output not surprisingly is $0.1454537$

Also note that instead of taking the sum, we could have directly used the \textbf{pbinom command} as follows:

$$\text{1 - p}(X<1030) \equiv 1 - \text{pbinom}(1030,6000,1/6)$$

Since the choice of the number 1 was arbitrary, this calculation tells us that we would give a 15% chance to any number showing up more than 1030 when rolled 6000 times.

As we noted, we can use the Normal approximation of binomial distribution. If $X \sim \text{Bin}(n,p)$, we know that $E(X) = np$ and $\text{Var}(X) = np(1-p)$. When $n$ goes large we can approximate the distribution of $X$ to be a normal distribution, with the same mean and variance. So

$$X \sim N(1000, 833.3)$$

Hence we just have to find $1 - \text{pnorm}(1030,1000,\text{sqrt}(833.3))$. (Also this can be very easily read from the table.)