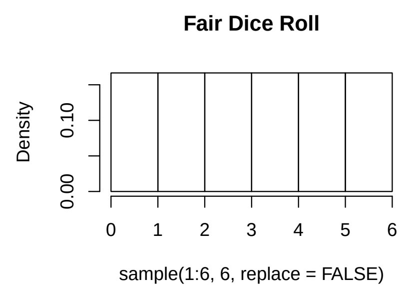
Distribution of Two Dice

October 17, 2014

Consider rolling a dice. It produces a uniform distribution of values between 1 and 6, with each value having a 1/6 likelihood.

```
bins <- seq(from = 0, to = 6, by = 1)
hist(sample(1:6, 6, replace = FALSE), prob = TRUE,
    breaks = bins, main = "Fair Dice Roll")
```



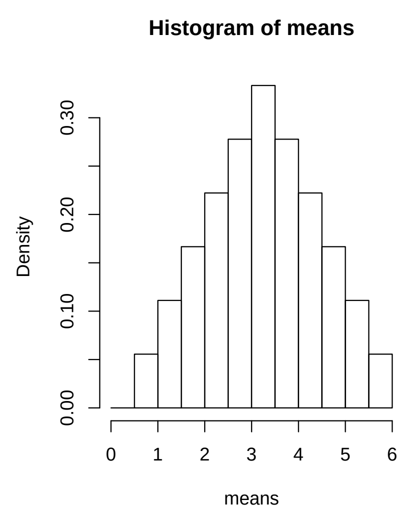
Now consider rolling two dice together and taking the mean of their values. Their mean would be between 1 and 6. But will the mean be uniformly distributed? Let's simulate all possible means.

```
mean_rolls <- matrix(NA, nrow = 6, ncol = 6)
for (i in 1:6) {
    for (j in 1:6) {
        mean_rolls[i, j] = mean(c(i, j))
    }
}
mean_rolls
```

```
##      [,1] [,2] [,3] [,4] [,5] [,6]
## [1,]  1.0  1.5  2.0  2.5  3.0  3.5
## [2,]  1.5  2.0  2.5  3.0  3.5  4.0
## [3,]  2.0  2.5  3.0  3.5  4.0  4.5
## [4,]  2.5  3.0  3.5  4.0  4.5  5.0
## [5,]  3.0  3.5  4.0  4.5  5.0  5.5
## [6,]  3.5  4.0  4.5  5.0  5.5  6.0
```

Let's use a histogram to visualize the distribution of means.

```
bins <- seq(from = 0, to = 6, by = 0.5)
means <- mean_rolls[mean_rolls > 0]
hist(means, breaks = bins, prob = TRUE)
```



We can see that means towards the mean (7) such as 3.0, 3.5, and 4.0 can be produced by many combinations of dice roll. But means far from the center, such as 1 and 6, can only be produced by a single combination each.

In a sense, this illustrates the central limit theorem. Values that are towards the center of possible outcomes are more likely to occur because there are more *random combinations* in with they can be produced. This gives the impression that the means of random dice rolls are converging in a near normal distribution towards central values. If we were to roll more dice, then this distribution would get thinner, and look even more like a classic normal distribution.