Week 2 Exercises

Question 2.2

A. The standard deviation of the given proportions is 0.00641 while the estimated standard deviation if the sexes of babies were independently decided with a constant probability over the 24-month period is 0.008003.

B. The difference between these two numbers is not statistically significant. Performing a chi-squared test yields a test statistic of 14.755 which does not exceed the critical value with 23 degrees of freedom, insufficient to demonstrate significance.

Quesion 2.3

```r
set.seed(3)

yuh <- replicate(1000, sum(sample(c(0,1), size = 20, prob = c(0.5,0.5), replace = TRUE)))

yuh <- as.tibble(yuh)
```

```
## Warning: 'as.tibble()' is deprecated as of tibble 2.0.0.
## Please use 'as_tibble()' instead.
## The signature and semantics have changed, see '?as_tibble'.
## This warning is displayed once every 8 hours.
## Call 'lifecycle::last_warnings()' to see where this warning was generated.
```

```r
yuh %>%
  ggplot(aes(x = value)) +
  geom_histogram(binwidth = 1, color = "white", aes(y = ..density..)) +
  stat_function(
        fun = dnorm,
        args = with(yuh, c(mean = mean(value), sd = sd(value))),
        color = "red") +
  theme_light() +
  labs(title = "Density of Sums",
       x = "Sums",
       y = "Density")
```

Density of Sums
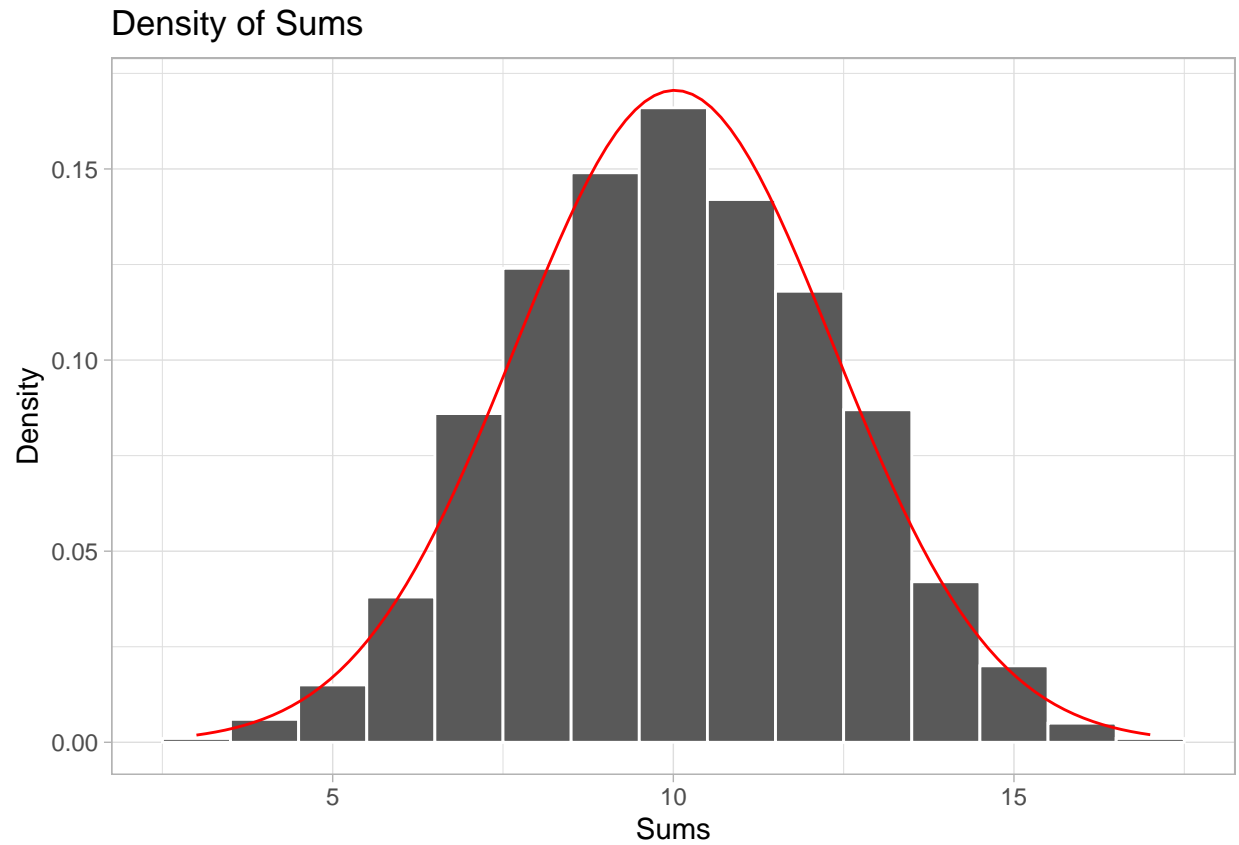
The histogram pretty closely follows the normal distribution curve. It is not as perfectly symmetrical and there is some deviance from the exact predicted value, but this is to be expected in a random sampling of this size.