HW7

Ian Driscoll

10/25/2018

7H.

```
library(ggplot2)
```

```
## Warning: package 'ggplot2' was built under R version 3.4.4
```

```
bin_sample = rbinom(1000, 5, 0.4)
table(bin_sample)
```

```
## bin_sample
##   0   1   2   3   4   5
##  69 266 352 229  65  19
```

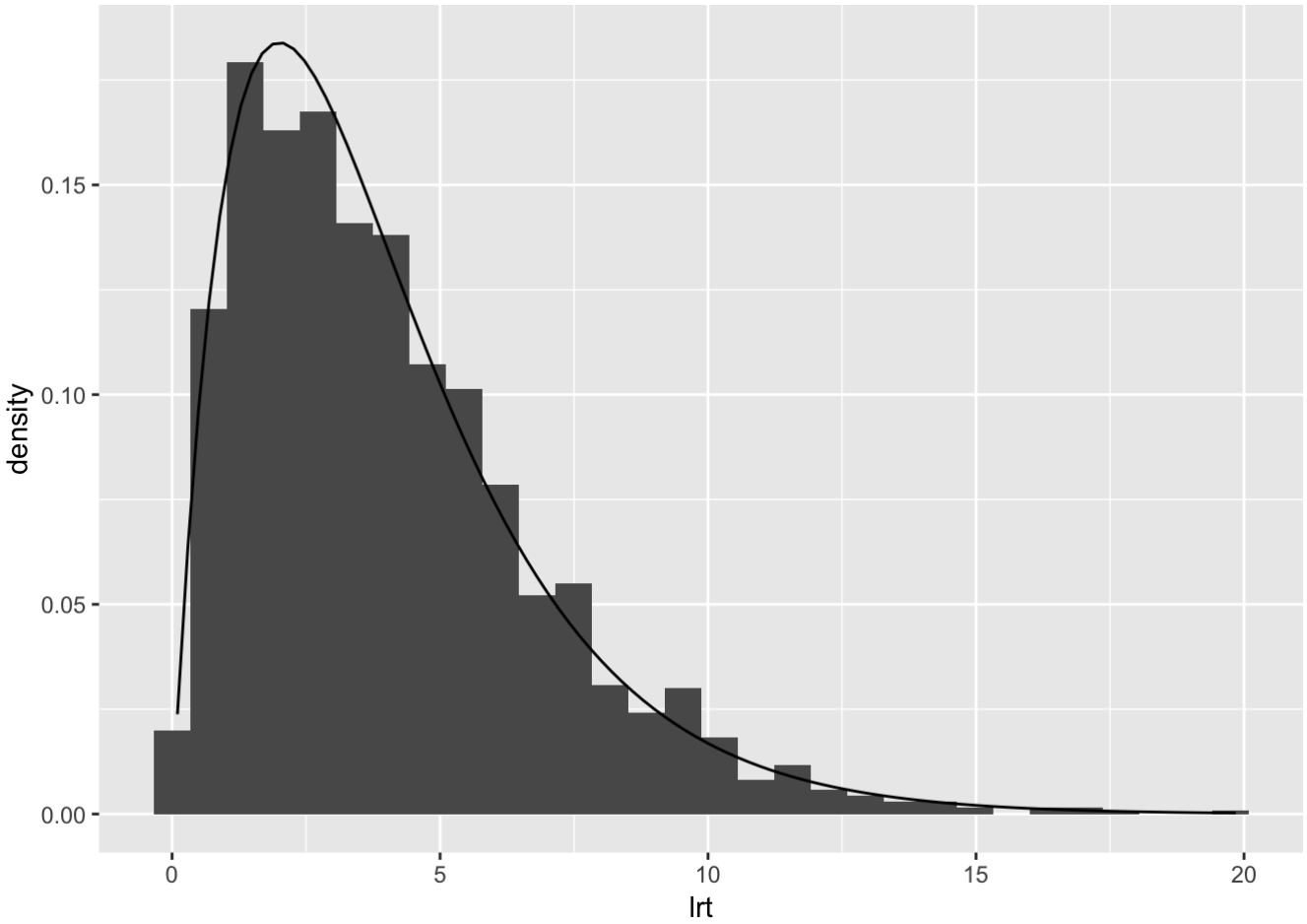
```
pchisq(8.33, 4, lower.tail = FALSE)
```

```
## [1] 0.08021098
```

7I.

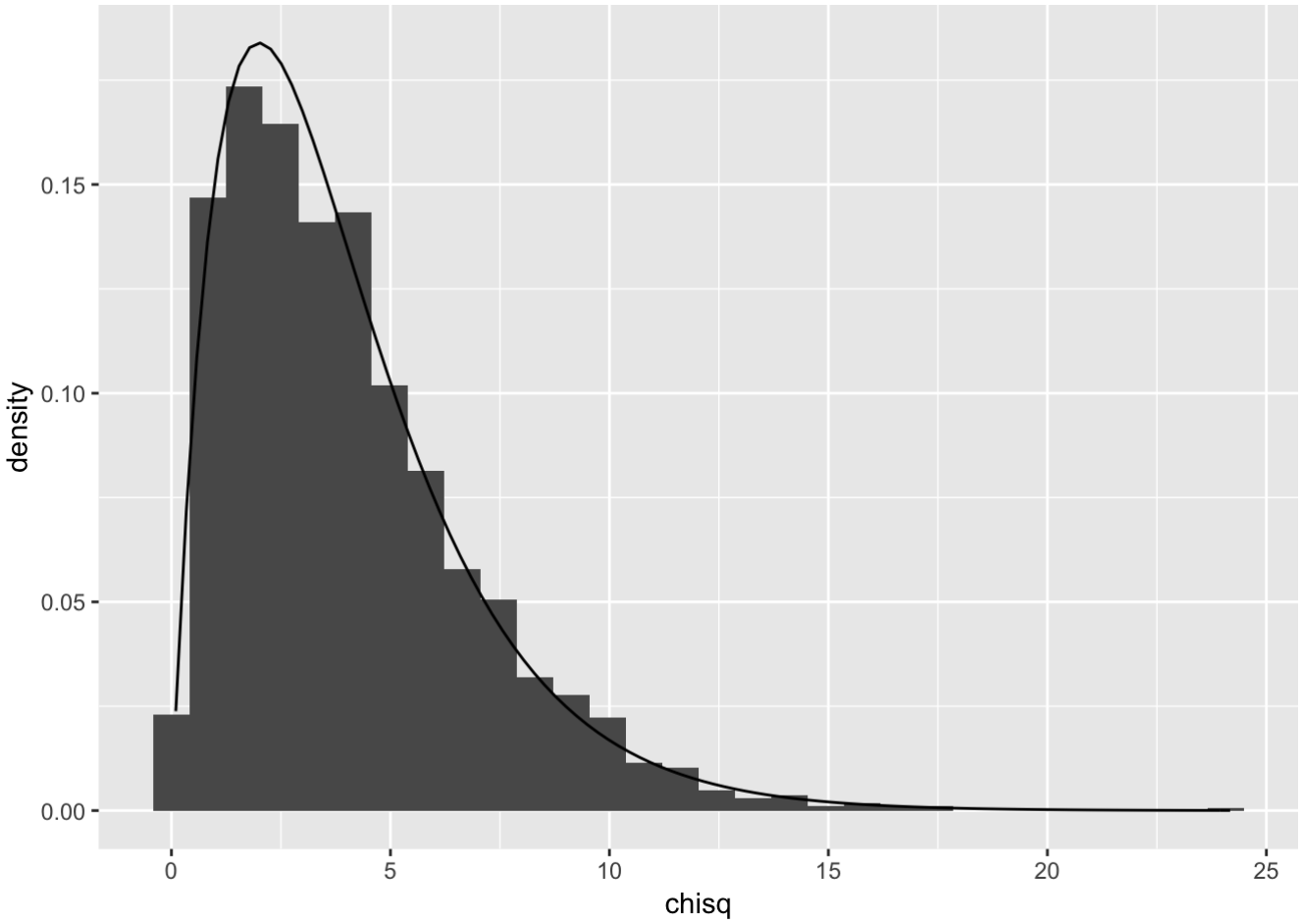
```
vec_lrt = c()
vec_chisq = c()
for (i in 1:2000) {
  bin_sample1 = rbinom(1000, 5, 0.4)
  sample_table1 = table(bin_sample1)
  p = (sample_table1[[2]] + 2 * sample_table1[[3]] + 3 * sample_table1[[4]] + 4 *
    sample_table1[[5]] + 5 * sample_table1[[6]]) / (5 * sample_table1[[1]] + 5 *
  sample_table1[[2]] + 5 * sample_table1[[3]] + 5 * sample_table1[[4]] + 5 *
 sample_table1[[5]] + 5 * sample_table1[[6]])
  expected = c()
  for (k in 0:5) {
    count = 1000 * choose(5, k) * p^k * (1-p)^(5-k)
    expected = c(expected, count)
  }
  total_lrt = 0
  total_chisq = 0
  for (x in 1:6) {
    total_lrt = total_lrt + (sample_table1[[x]] * log(sample_table1[[x]] / expected[x]))
    total_chisq = total_chisq + (((sample_table1[[x]] - expected[x])^2) / expected[x])
  }
  total_lrt = total_lrt * 2
  vec_lrt = c(vec_lrt, total_lrt)
  vec_chisq = c(vec_chisq, total_chisq)
}
df = data.frame(lrt = vec_lrt, chisq = vec_chisq)
lrt_plot = ggplot(df) + geom_histogram(aes(x = lrt, y = ..density..)) + stat_function(fu
n = dchisq, args = list(df = 4))
lrt_plot
```

```
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
```




```
chisq_plot = ggplot(df) + geom_histogram(aes(x = chisq, y = ..density..)) + stat_functio
n(fun = dchisq, args = list(df = 4))
chisq_plot
```

```
## `stat_bin()` using `bins = 30`. Pick better value with `binwidth`.
```


density
0.15
0.10
0.05
0.00
0
5
10
15
20
25
chisq