Homework 3, Exercise 6(b)

Robert Winter

6(b) Find a 95% HPD credible interval for N (you do not need to match 95% exactly. Get as close as possible).

First, we construct the posterior probability mass function $h(N|x = 203)$, and generate some data $(203, h(203|x = 203)), ... (1000, h(1000|x = 203))$.

```
# Posterior PMF
post = function(n){
  c = 21.253
  if(n>=203){
    h = (c/n)*(0.99)^(n-1)
  }
  else{
    h = 0
  }
  return(h)
}
```

```
# Data for plot
ns = c(203:1000)
hs = c()
for(i in 1:length(ns)){
  hs[i] = post(ns[i])
}
data = cbind(ns, hs) %>% as.data.frame()
```

Since the posterior PMF is monotonically decreasing, it is unimodal, with its mode at the leftmost point of its support, $n = 203$. Since the HPD credible interval for a unimodal distribution contains the mode, it follows that the HPD interval begins at $n = 203$. After some trial and error with the rightmost endpoint of the HPD interval, I found that the interval $(203, 438)$ contains approximately 95.02% of the mass of the distribution, making it the (best possible approximate) 95% HPD credible interval.

```r
# HPD interval
L = 203
R = 438

area = 0
for(n in L:R){
  area = area + post(n)
}
area # 0.9502
```

```
[1] 0.9502448
```

We visualize the 95% HPD credible interval below.

```r
ggplot(data, aes(x=ns, y=hs)) +
  theme_bw() +
  geom_point(size=1, col = "blue") +
  xlab("n") +
  ylab("h(n | x = 203)") +
  geom_vline(xintercept = c(L,R), linetype = "dashed") +
  ggtitle("HPD Credible Interval for N: [203, 438]") +
  theme(plot.title = element_text(hjust = 0.5)) +
  annotate("text", x = 0.5*(L+R), y = 0.0004, label = "Area = 0.950")
```