

Small group assignment: Working with data in R

Your task on this assignment is to work in a small group (2-3 people) to address the following questions. Only one writeup (as a .Rmd and .pdf) needs to be submitted for each group. The problems below use several different datasets. Your document should include code, figures, and written answers to each question.

1. Fisheries example

The code below includes the Canadian Atlantic cod population (in tonnes) over time. Build a plot of this data and explain in writing what has happened to the population.

```
cod_pop <- c(1450, 1420, 1050, 2643, 1060, 1080, 1410, 1150, 740, 175, 43, 10, 12,
            15, 16, 16, 28, 30, 32, 23, 12, 19, 27)
years <- c(1983, 1984, 1985, 1986, 1987, 1988, 1989, 1990, 1991, 1992, 1993, 1994, 1995, 1996, 1997, 1998, 1999, 2000, 2001, 2002, 2003, 2004, 2005)

fishdata <- data.frame(x = years, y = cod_pop)
ggplot(data = fishdata, aes(x = years, y = cod_pop)) + geom_point() + geom_smooth() +
  xlab("Year") + ylab("Cod Population")
```

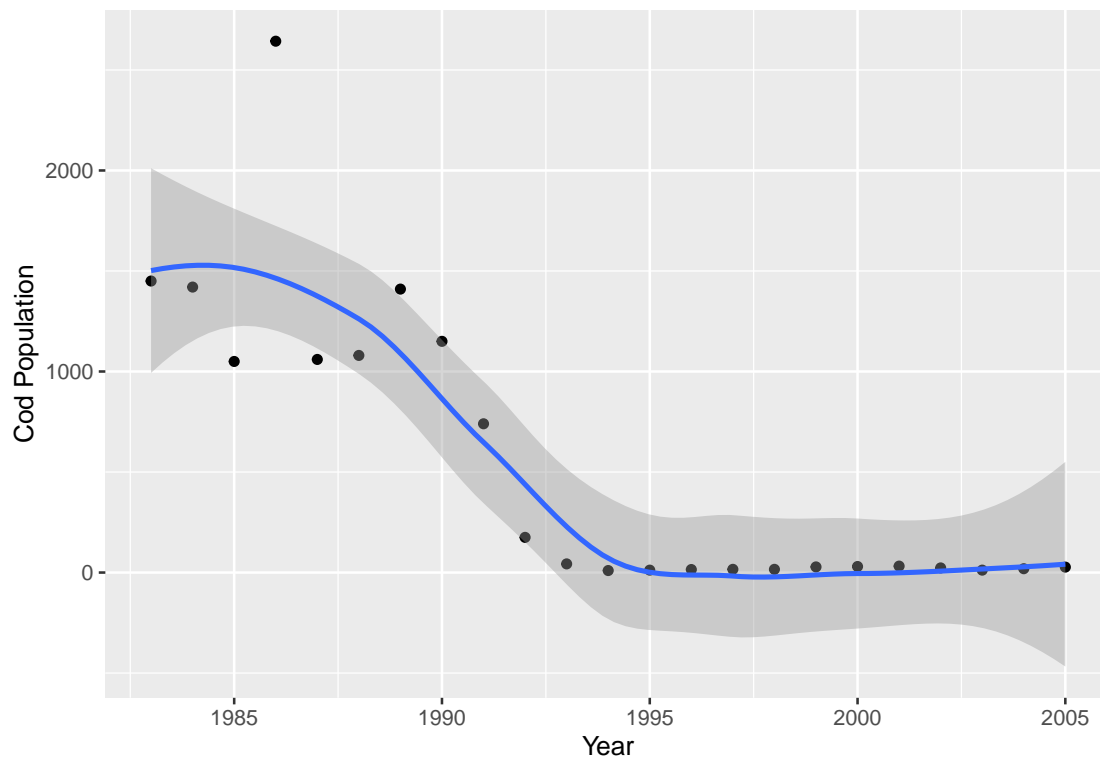


Figure 1: The cod population started declining after the year 1989. Beginning in 1995, the population remained nearly extinct (around 0).

2-3. Population models

The discrete-time logistic growth model has the form $N[t] + r * N[t] * (1 - N[t]/K)$. It is often used to describe population growth, such as in animal and plant populations, cancer tumor growth, and bacterial growth. Your task is to write a for loop that runs this model from time = 1 to time = 50. The starting value of N is 0.1. The default parameter values are $r = 0.4$ and $K = 1$. What is the behavior (e.g., does the population go extinct? are there cycles?) with the default parameter values? What happens if you increase r ? Please provide graphs to support your conclusions.

```
N <- vector("numeric", length = 50)

for (i in 1:50) {
  N[1] <- 0.1
  N[i + 1] <- N[i] + 0.4 * N[i] * (1 - N[i]/1)
}
popdata <- data.frame(1:51, N)

ggplot(popdata, aes(x = X1.51, y = N)) + geom_point() + geom_line(color = "black") +
  ggtitle("R = 0.4") + xlab("Year") + ylab("Population")
```

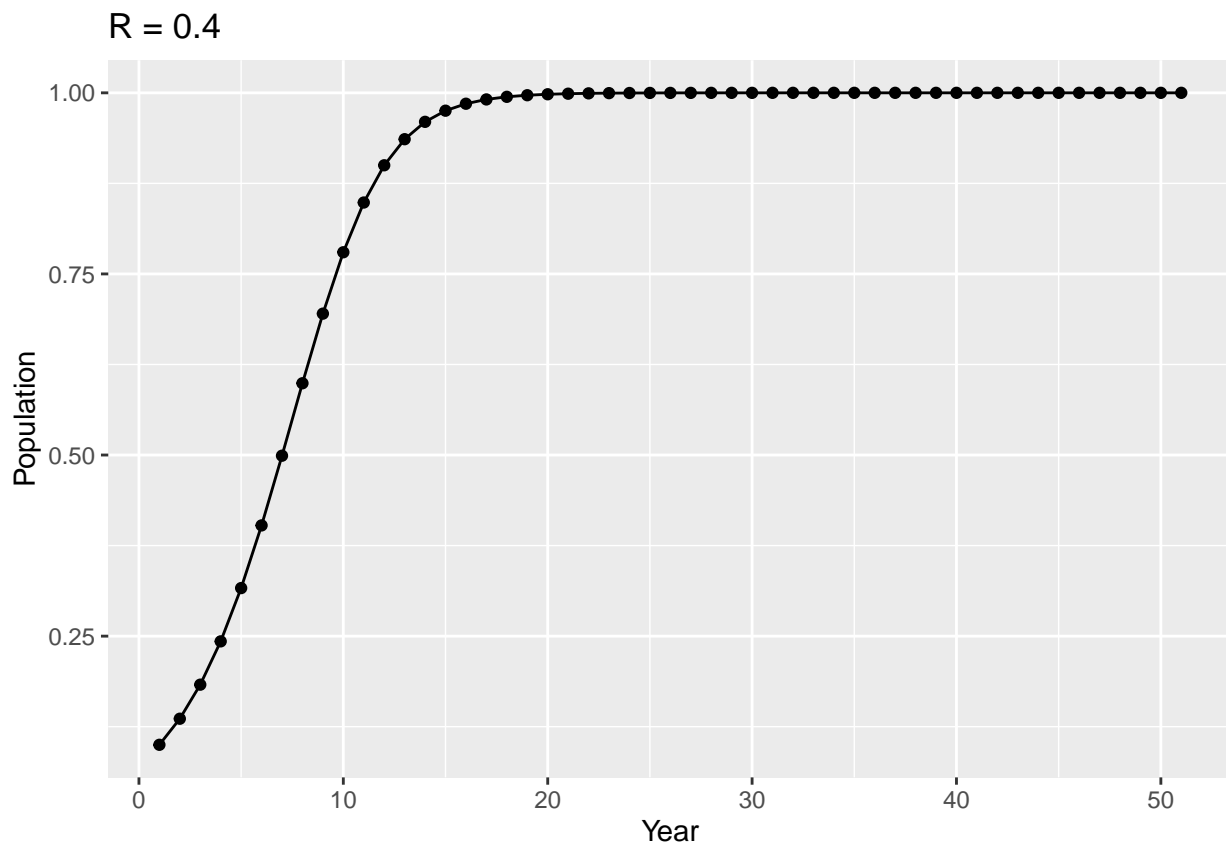


Figure 2: With the default parameter values, the population will grow until it reaches a carrying capacity and then levels off at that capacity.

```

N <- vector("numeric", length = 50)

for (i in 1:50) {
  N[1] <- 0.1
  N[i + 1] <- N[i] + 0.6 * N[i] * (1 - N[i]/1)
}
popdata <- data.frame(1:51, N)

ggplot(popdata, aes(x = X1.51, y = N)) + geom_point() + geom_line(color = "black") +
  ggtitle("R = 0.6") + xlab("Year") + ylab("Population")

for (i in 1:50) {
  N[1] <- 0.1
  N[i + 1] <- N[i] + 0.8 * N[i] * (1 - N[i]/1)
}
popdata <- data.frame(1:51, N)

ggplot(popdata, aes(x = X1.51, y = N)) + geom_point() + geom_line(color = "black") +
  ggtitle("R = 0.8") + xlab("Year") + ylab("Population")

for (i in 1:50) {
  N[1] <- 0.1
  N[i + 1] <- N[i] + 1 * N[i] * (1 - N[i]/1)
}
popdata <- data.frame(1:51, N)

ggplot(popdata, aes(x = X1.51, y = N)) + geom_point() + geom_line(color = "black") +
  ggtitle("R = 1") + xlab("Year") + ylab("Population")

```

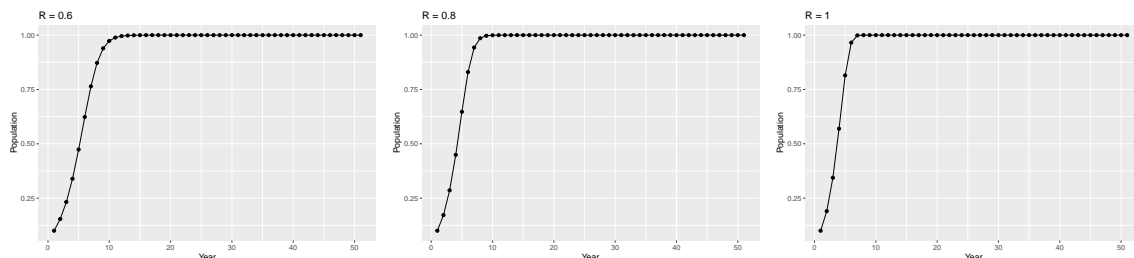


Figure 3: However, if you increase r , the populations will hit this carrying capacity and remain there, but it will hit the carrying capacity in fewer years as r increases, meaning that the r parameter represents the growth rate of the population

4. Measles data in the UK

On Canvas (https://mycourses.unh.edu/courses/88910/files/6917786?module_item_id=2094185), I've uploaded historical data on total measles cases and deaths for Wales and the UK. Load the data into R and build figures. Use the data to explain as much as possible about what is happening with the dynamics of measles.

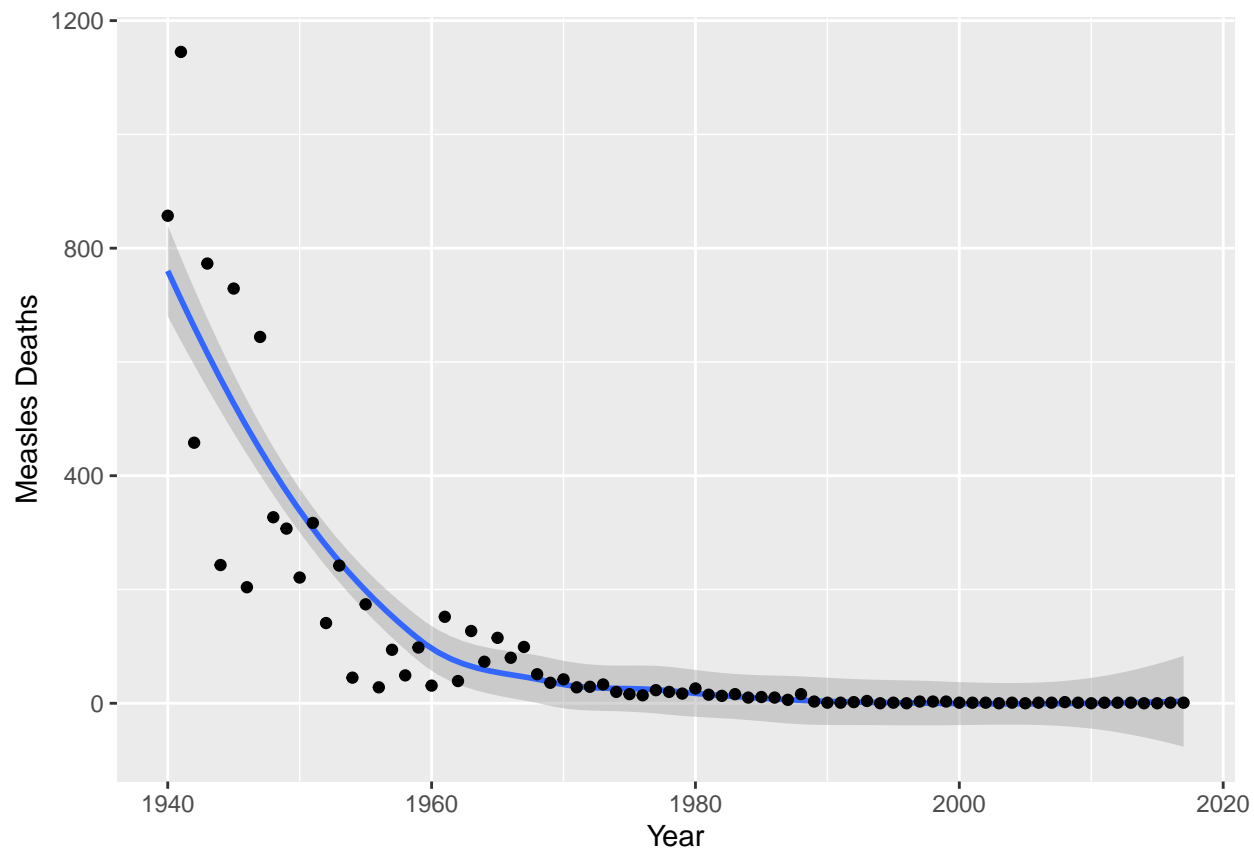


Figure 4: In the total number of deaths per year, we can see that the total number of measles deaths generally decreased since the 40s, probably with the increase in standard public health procedures. Deaths reached close to zero since around 1970 (with the advent of the measles vaccine in 1963)

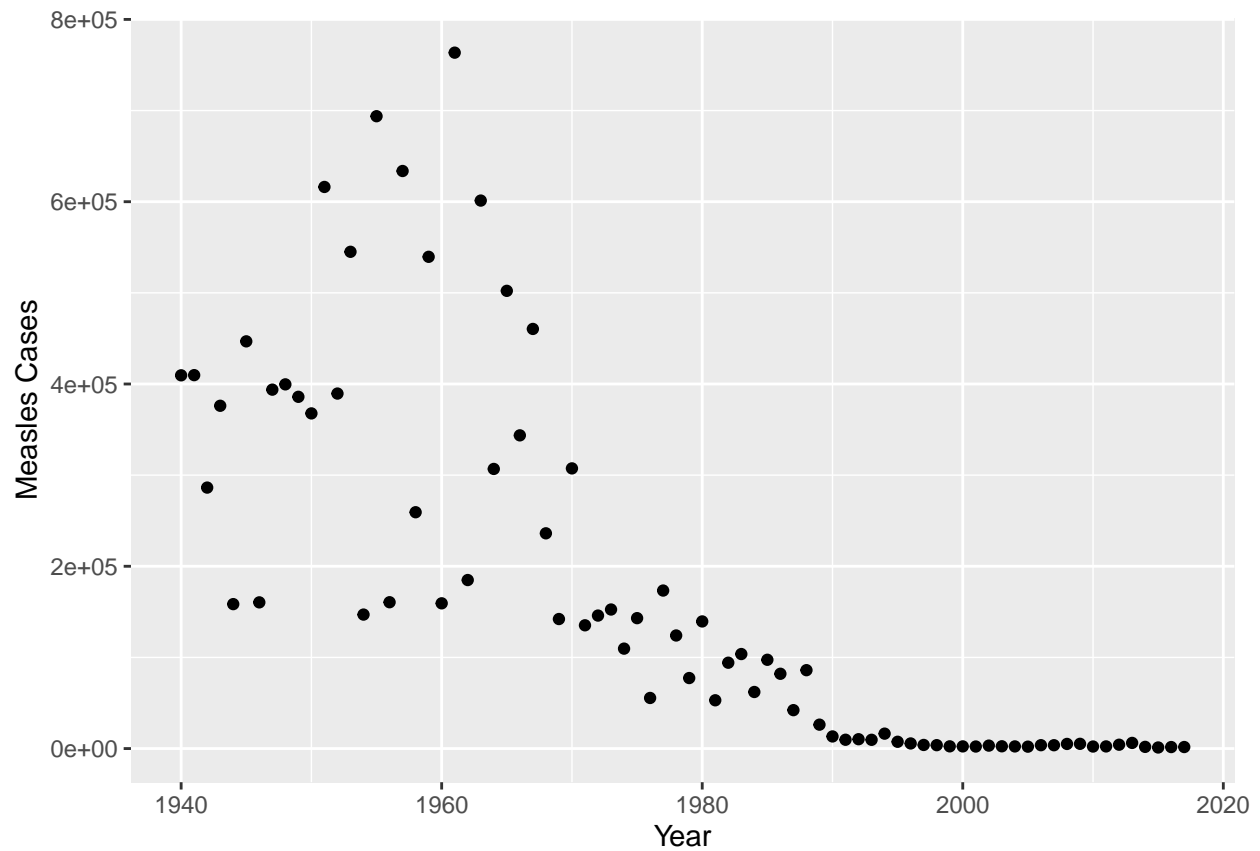


Figure 5: Similar to the Measles Deaths trend, we can see that the total number of measles cases generally decreased since the 40s. However, they didn't seem to almost eradicate cases until the mid 90s, probably representing a vaccine rollout that took a few years.

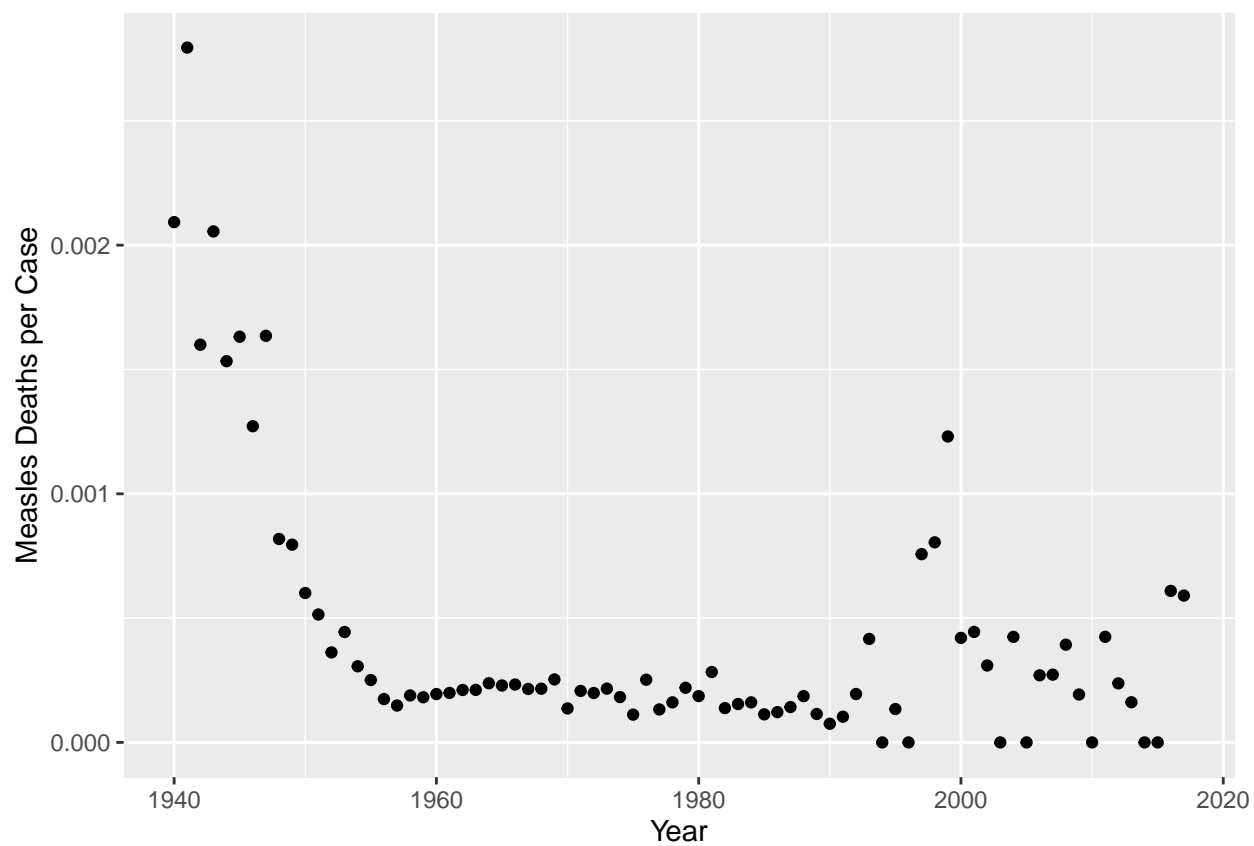


Figure 6: Here, we can see an interesting uptick in Deaths per Case starting in the 2000s, maybe representing an anti-vax movement that resulted in the reemergence of measles cases.

5-7. Population genetics

The code below implements a Wright-Fisher model with genetic drift and mutation. Use and modify the code to create plots and answer the following questions.

```
# Part 1: Wright-Fisher model with genetic drift and mutation

# Created by Easton R. White

# Last edited: 4-Oct-2021

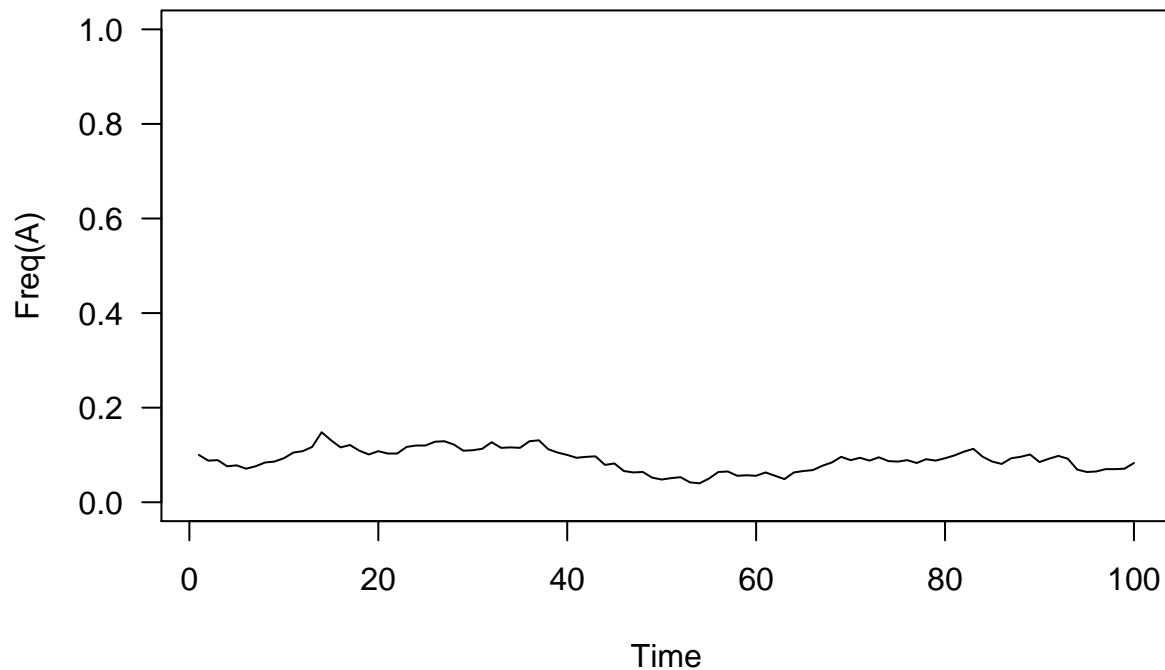
# Setting up parameter values
N = 1000 # population size
p_ini = 0.1 # the initial frequency of the A allele
u = 0.1 # mutation rate from a to A
max.time = 100 # time to run simulations

# Custom built function
sample.population <- function(p, N) {
  next_pop <- vector("character", length = N)
  for (i in 1:N) {
    new_allele <- sample(x = c("A", "a"), size = 1, prob = c(p, 1 - p))
    next_pop[i] <- as.character(new_allele)
    if (x == "a") {
      rbinom("a", N, u)
    }
  }
  return(sum(next_pop == "A")/N) # returns the value for p
}

p <- vector(mode = "numeric", length = max.time)
p[1] <- p_ini

for (t in 1:(max.time - 1)) {
  p[t + 1] = sample.population(p[t], N)
}

plot(1:max.time, p, ylim = c(0, 1), type = "l", las = 1, ylab = "Freq(A)", xlab = "Time")
```



Population genetics questions

1. What is the effect of population size (N) on the probability of extinction of A allele?

The smaller the population size, the greater the probability of the extinction of A allele.

2. How does the initial frequency of the A allele affect the probability it will reach fixation?

The greater the initial frequency of the allele, the more likely it is to reach fixation

3. Let's now study the combined effects of mutation and genetic drift. Modify the above code (specifically the `sample.population()` function) to include a probability, `u`, that if a small “a” allele is chosen, it will mutate to be a big A allele. An if/else statement and or the command `rbinom()` might be helpful. How does this affect our findings in questions 1 and 2?

Including the probability of the mutation, makes it so that the smaller the population size, the lower the probability of extinction of A allele and so that the greater the initial frequency of the allele the more likely it is to reach fixation.