Modelação Ecológica - Tutorial Aula 1 - Final Task

Tiago A. Marques

October 25, 2018

Contents

1 Introduction

This document describes a possible implementation of the final task presented to the students of "Modelação Ecológica" under the tutorial "A hands-on tutorial on R and R Studio".

2 The task

Here we will implement an exercise were we pretend we are sampling an animal population, using some (very basic) simulations to understand the process better. Create plots that represent all the steps of your task, with proper legends, labels, colors, etc, and add all your comments to the dynamic report. 1. Simulate the positions of 10000 animals in a study area, with length 10km and width 1km. Assume that any animal has an equal chance to be at any location in the study area (this corresponds to a uniform density surface). 2. Generate a transect at a random location along the study area. 3. Assume that you can potentially detect at most animals up to 500 meters from your transect. Count all the animals that you would detect if detection was perfect across your transect. 4. Consider that animals far from your transect are harder to detect - yes, you are doing distance sampling! Define a function that represents a distance sampling half-normal detection function (If needed, look back at your slides from the 1st class). Assume that sigma=200m. 5. Simulate the detection process and get a sample of those animals detected. 6. Create a plot that allows you to estimate (at this stage just a visual guess in needed) the detection probability. 7. Repeat the sampling process 500 times, and store the number of animals detected in each one of your simulated surveys. 8. plot the distribution of the number of animals that you would detect each new survey. Take your own conclusions about all that you did.

3 Implementation

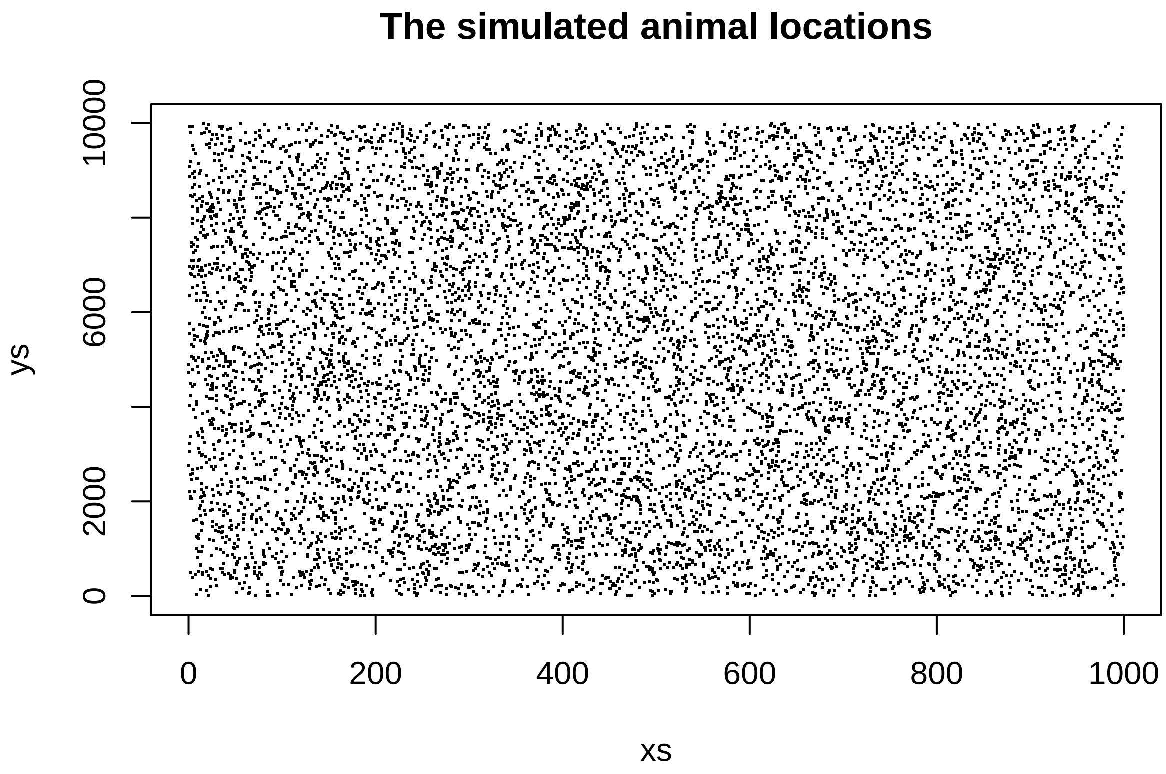
3.1 Simulate animal positions

First we define some state variables

```
#number of animals
N=10000
#units in meters
#length
L=10000
#width
W=1000
```

Next, we simulate the animals, and show them in a plot. Since the distribution was uniform in space, we use the function `runif`

```
#set the seed to get always the same results
set.seed(123)
xs=runif(N,0,W)
ys=runif(N,0,L)
par(mfrow=c(1,1))
plot(xs,ys,pch=".",main="The simulated animal locations")
```

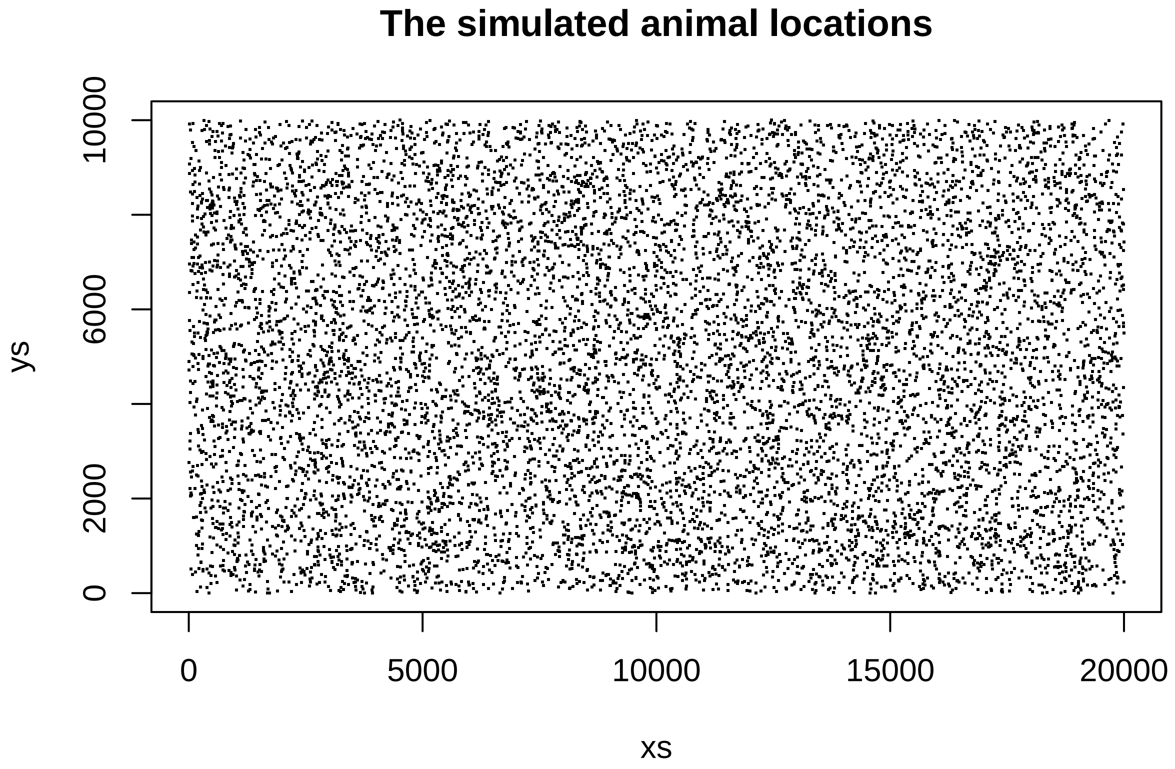


3.2 Generate a transect

Next, we need to generate the transect. The transect goes along the study area. After the fact I realized this is not sensible, since the area is only 1km wide and we will assume that we could see up to 500 meters on each side. Therefore, on the fly, I am going to change the width of the area to be 20 km, so that the exercise is more sensible.

The new area and animals are shown below:

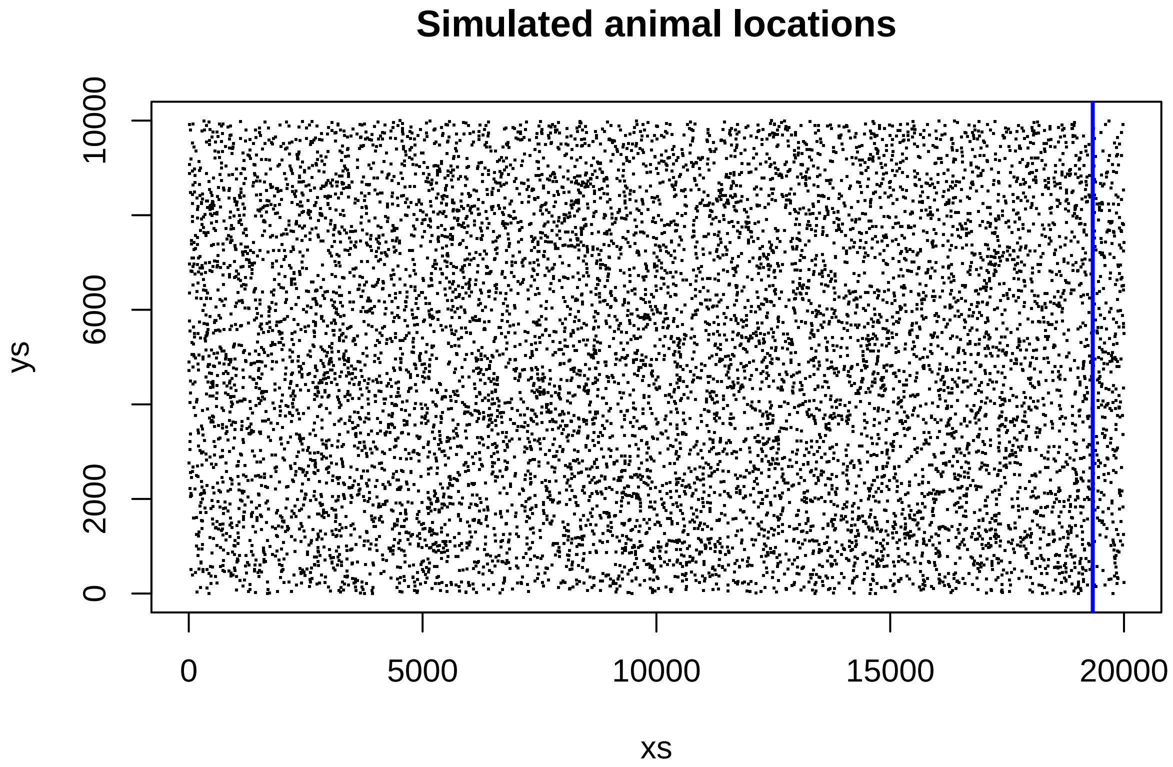
```r
#number of animals
N=10000
#units in meters
#length
L=10000
#width
W=20000
#set the seed to get always the same results
set.seed(123)
xs=runif(N,0,W)
ys=runif(N,0,L)
par(mfrow=c(1,1))
plot(xs,ys,pch=".",main="The simulated animal locations")
```



Then, I generate a transect along the study area, and to avoid edge effects, I assume the transect will not overlap the edges (by defining a buffer where the transect midle point can not be placed).

```r
#define a buffer
buffer=500
```

```
#get transect location
trans=runif(1,0+buffer,W-buffer)
par(mfrow=c(1,1))
plot(xs,ys,pch=".",main="Simulated animal locations")
abline(v=trans,lwd=2,col=4)
```



3.3 Animals that can be detected

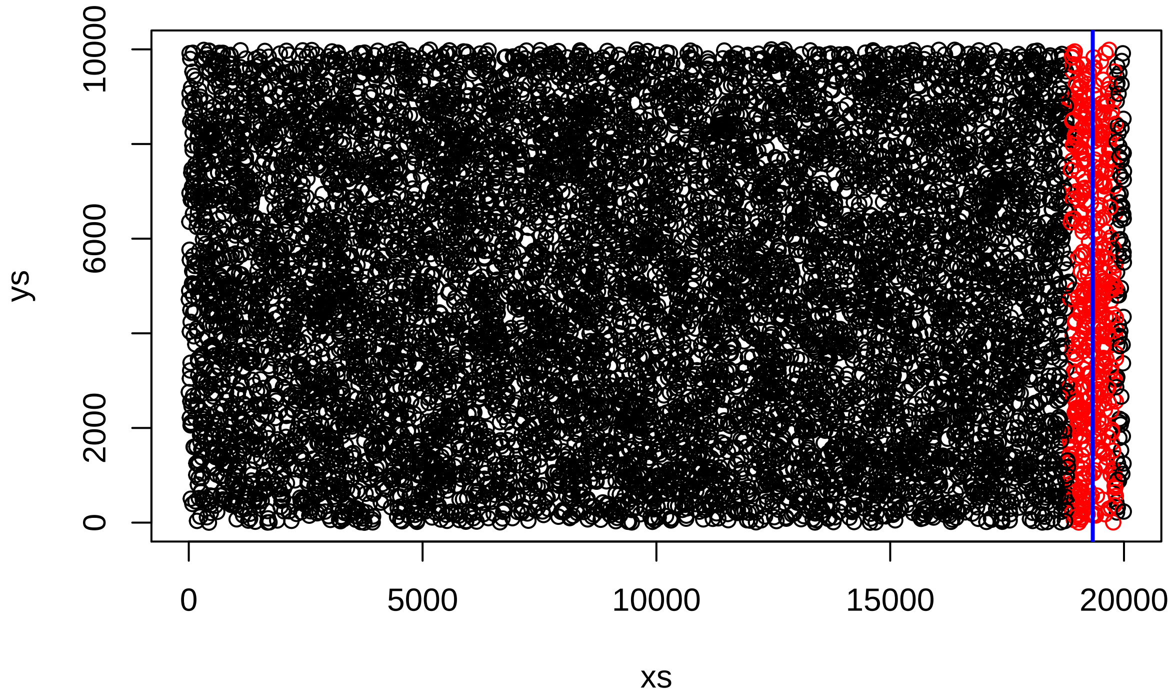
We simply need to see how many animals are within 500 meters of our transect

```
#define what we call in distance sampling the truncation distance
#no animals can be detected beyhond that distance
w=500
#calculate the perpendicular distance to the transect, for each animal
dist=abs(trans-xs)
#how many are at risk
natrisk=sum(dist<w)
```

There are 477 animals within 500

We can plot them with a different color in our study area

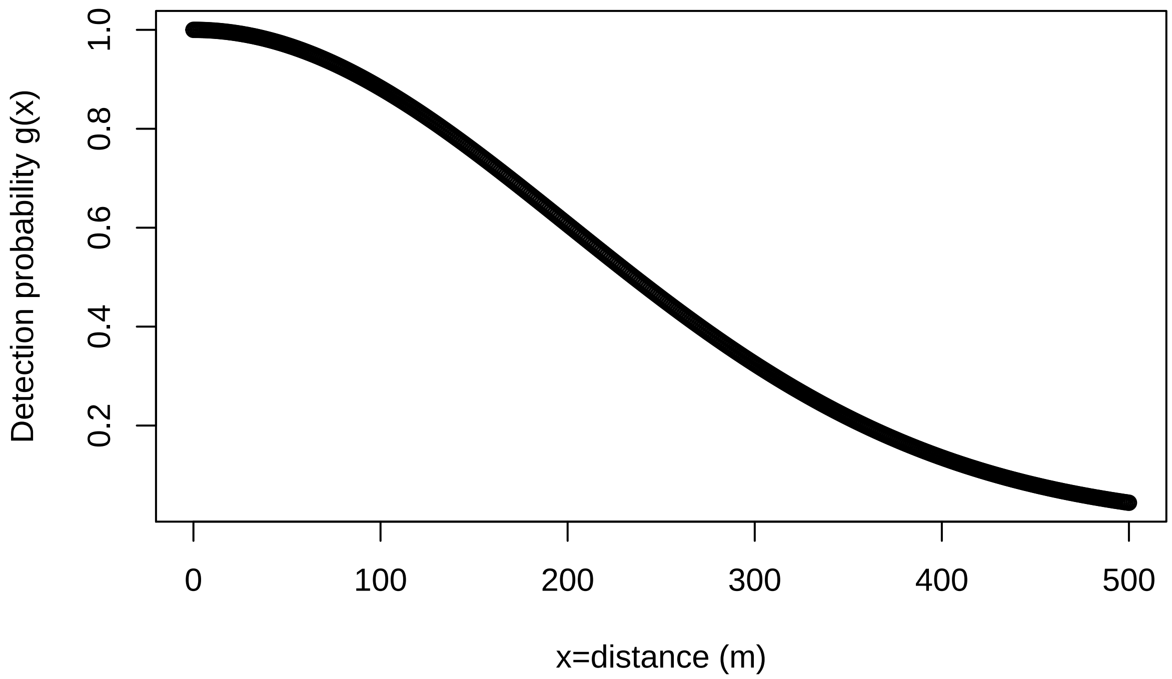
```
plot(xs,ys,col=ifelse(dist<w,2,1))
abline(v=trans,lwd=2,col=4)
```



3.4 Defining a detection function

Using the material from the slides, we can define the detection function. Below we plot the detection function setting `sigma=200` by default

```
HN=function(x,sigma=200){
  px=exp(-(x^2)/(2*sigma^2))
  return(px)
}
plot(0:500,HN(0:500),xlab="x=distance (m)",ylab="Detection probability g(x)")
```



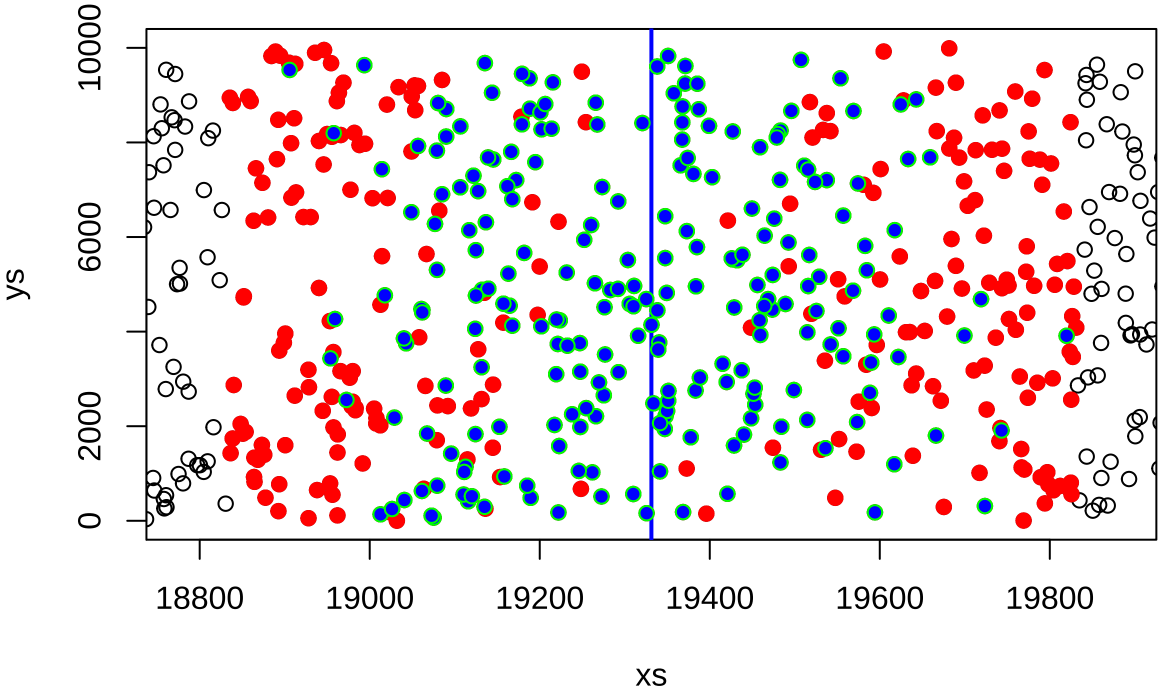
3.5 Simulate the detection process

Now we can simulate the detection process. To do that we first evaluate the detection probability of each animal, and then compare that probability with a random draw from a Uniform in (0,1). If the value is smaller than the probability of detection, the animal is detected, if not, it is missed

```r
#P detectar cada um dos meus individuos
pdet=HN(dist)
#an index for animals potentially detected
inTRANSECT=dist<500
#get coordinates of just those
xs2=xs[inTRANSECT]
ys2=ys[inTRANSECT]
#their distances
dist=dist[inTRANSECT]
#and their pdets
pdet=pdet[inTRANSECT]
#how
n1=sum(inTRANSECT)
#get the uniform
punif=runif(n1,0,1)
#index for those detected detected
inDET=pdet>punif
#how many where they?
n2=sum(inDET)
```

Now, we can plot the detected animals on top of the previous plots, zooming in the sampled transect

```
plot(xs,ys,xlim=c(trans-(w+50),trans+(w+50)))
points(xs2,ys2,col="red",bg="red",pch=21)
abline(v=trans,lwd=2,col=4)
points(xs2[inDET],ys2[inDET],col="green",bg="blue",pch=21)
```

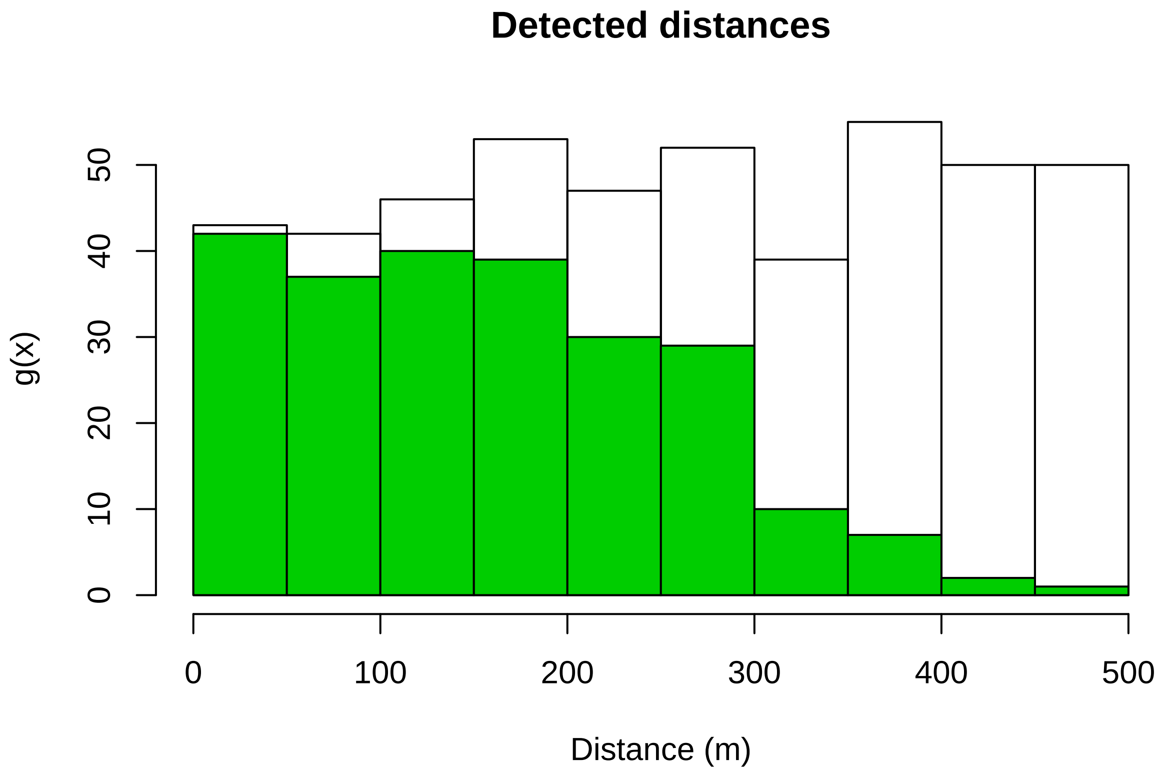


We had 477 animals potentially detected, and we detected 237.

3.6 The detection probability

We simply need an histogram of the distances of animals at risk of being detected, and we can add on top the detected distances

```
hist(dist,main="Detected distances",xlab="Distance (m)",ylab="g(x)")
hist(dist[inDET],add=TRUE,col=3)
```



Roughtly looking at this, we would estimate a probability of detection of about 50 %.

We actually detected 49.7 % of the animals in the covered transect. Not bad for a "guessestimate" above!

Note that we can actually calculate the true probability of detection, by integrating the detection function (the area under the curve) and dividing that by `w` (the area of the histogram if we detected all distances).

```
trueP=integrate(f = HN,lower = 0,upper=w)$value/w
```

The true probability of detecting an animal within 500m of the transect is 0.495.

3.7 Repeat the process many times

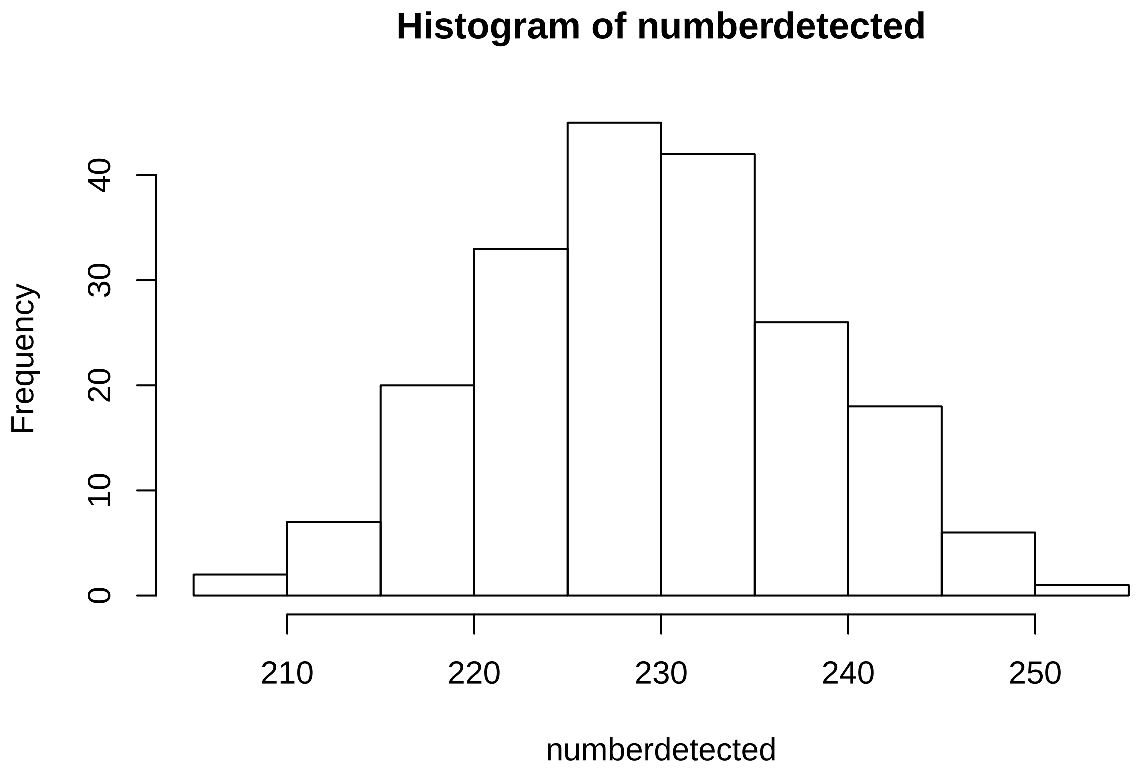
We could now simply repeat all this many times. Note that we could repeat the process of generating the animals' positions, or the positions and the transect, or condition on the observed animals and transect, and then repeat just the detection component. Here we do the latter, but leave the former as an exercise for you to do at home.

```
#how many times do we repeat the process
n3=200
numberdetected=numeric(n3)
for(j in 1:n3){
#get the uniform
punif=runif(n1,0,1)
#index for those detected detected
inDET=pdet>punif
#how many where they?
numberdetected[j]=sum(inDET)
}
```

3.8 Plot the outcome

Finally, we could plot the number of animals detected in each of our 200 simulations

```
hist(numberdetected)
```



4 Conclusions

We were able to simulate an animal population and a simple sampling strategy, to formulate a model for detectability (based on distance sampling) and to simulate the detection process.

We investigated the randomness involved in the number of animals detected conditional on the animals available and a transect position, and noted there was considerable variability across simulations.