COGSCI131–Spring 2019 — Homework 5Solutions

Yuchen Zhou, SID 3034489901

1.

By code, we get probability that x=1 is in the concept given that x=0 is is 0.73.

```
import numpy
import matplotlib.pyplot as plt

#x=numpy.arange(-10,10,0.1)
x=1
y=0
#y=numpy.zeros(200)
left=numpy.zeros(1000)
right=numpy.zeros(1000)
length=numpy.zeros(1000)
for i in range(1000):
    left[i] = -numpy.random.uniform()*10
    right[i]= numpy.random.uniform()*10
    length[i]=right[i]-left[i]

sumlength2 = 0
sumlength1=0
for j in range(1000):
    sumlength1+=1/length[j]
    if left[j]<x and right[j]>x:
        sumlength2+=1/length[j]
y=sumlength2/sumlength1
"""
fig=plt.figure()
ax=fig.add_subplot(111)
ax.plot(x,y)
plt.xlabel("x")
plt.ylabel("curve of probability when taking 100 regions")
plt.show()
"""
print(y)
```

2.

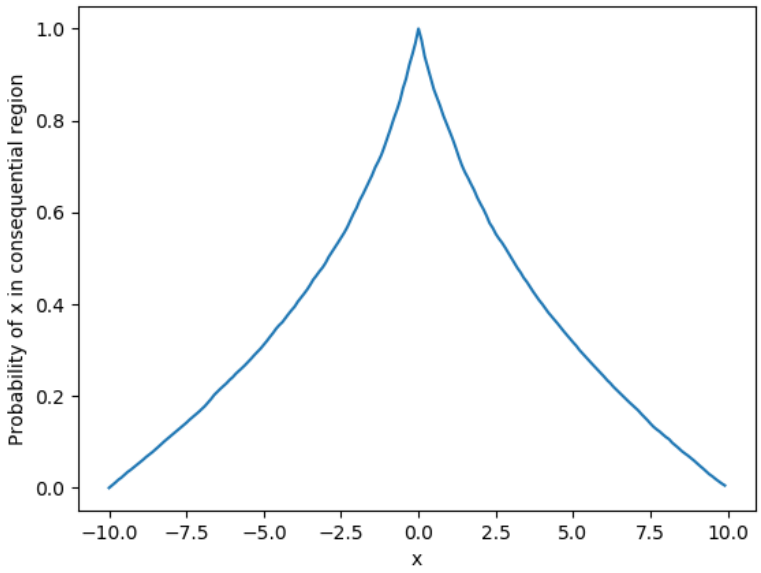

Figure 1: probability of x in consequential region

- The function looks like a negative exponential function. Here is the intuition:
- We know the possibility would decrease if we get further from "0".
- Also, we know that regions that the item loses per step when getting further is proportional to it's distance from "0". So it's an exponential function.

```
import numpy
import matplotlib.pyplot as plt

x=numpy.arange(-10,10,0.1)
y=numpy.zeros(200)
left=numpy.zeros(10000)
right=numpy.zeros(10000)
length=numpy.zeros(10000)
for i in range(10000):
    left[i] = -numpy.random.uniform()*10
    right[i]= numpy.random.uniform()*10
    length[i]=right[i]-left[i]
for i in range(200):
    sumlength2 = 0
    sumlength1=0
    for j in range(10000):
        sumlength1+=1/length[j]
        if left[j]<x[i] and right[j]>x[i]:
            sumlength2+=1/length[j]
    y[i]=sumlength2/sumlength1
fig=plt.figure()
```

```
ax=fig.add_subplot(111)
ax.plot(x,y)
plt.xlabel("x")
plt.ylabel("Probability of x in consequential region")
plt.show()
```

3.

We can not justify an exponential curve by eyes, however, if

$$y = a \times k^x, a, k \in R_+$$

then

$$\log y = \log k \times x + \log a$$

which is linear to x.
In this way,we can draw the curve and by justifying whether it's straight we can check whether y is exponential to x.

4.

1. Figure of logarithmic probability to x ranging from -10 to 10 as shown:

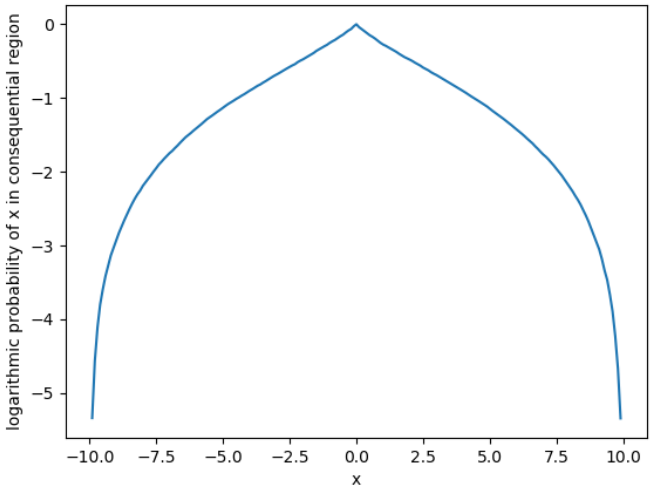

Figure 2: logarithmic probability of x in consequential region

2. Figure of logarithmic probability to x ranging from -5 to 5 as shown:

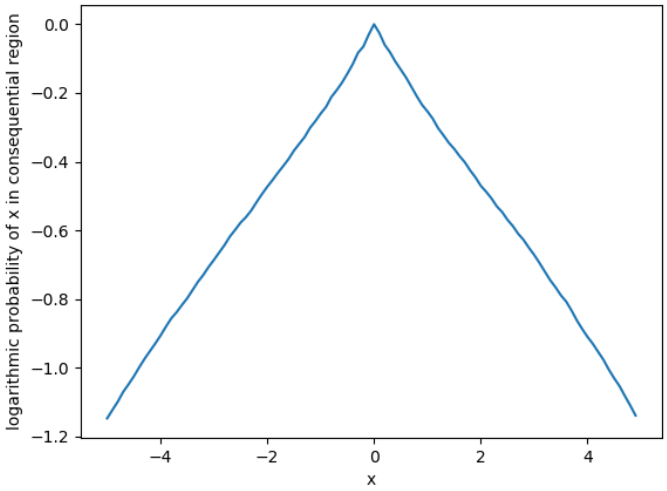

Figure 3: logarithmic probability of x in consequential region

3. From two plots we know that when x is close to 0, the curve perfectly satisfies an exponential function. However, when x is far from 0, it does not observe that law at all.

4. When simulating consequential region, we choose left edges no less than -10 and right edges no more than 10, while in reality, the boundary should extend to negative infinite and infinite. This approximation results in when getting close to the boundary, the probability no longer follow the exponential decreasing law.

5.

1. Taking 10 consequential regions:

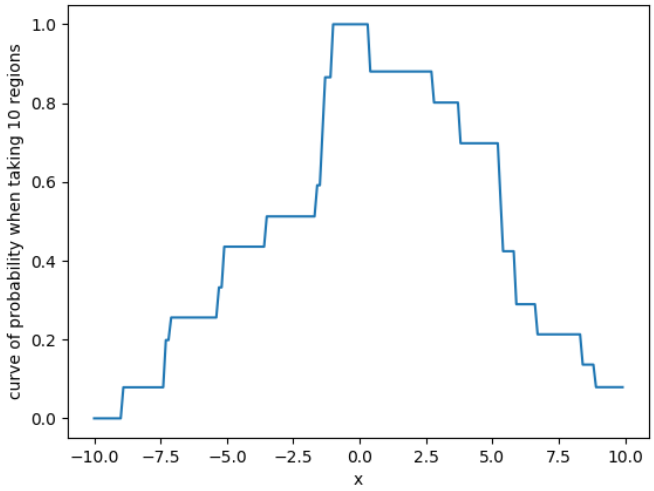

Figure 4: curve of probabilities when taking 10 regions

2. Taking 100 consequential regions:

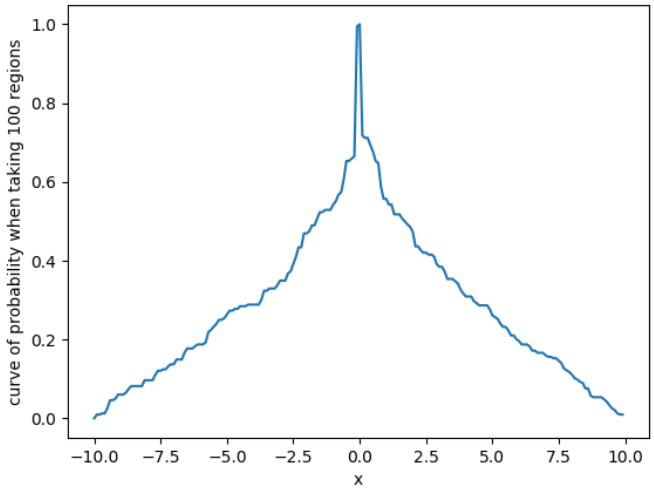

Figure 5: curve of probabilities when taking 100 regions

3. Taking 1000 consequential regions:

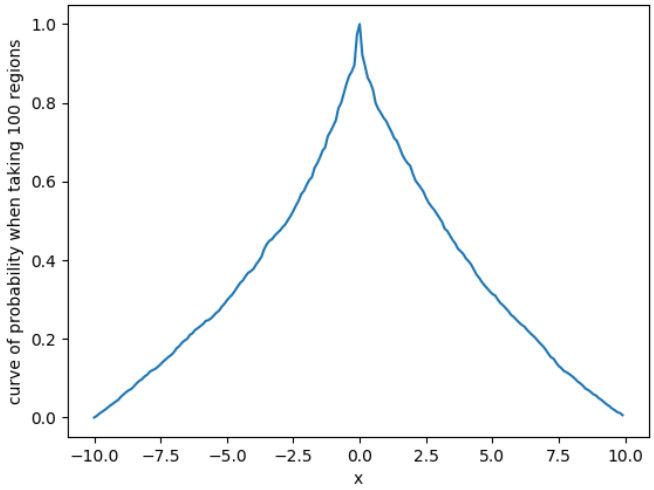

Figure 6: curve of probabilities when taking 1000 regions

4. We can conclude that more regions are taken, smoother the curve will be.
   Also, with less consequential regions, the result will be more noisy in respect to the universal law when judging.

6.

For 10 consequential regions and 10000:

1. Find n people, test their responses to a list of items, like "1", "2", "3", ..., given that they all know "0".
2. Collect their responses of whether item i is "0" by "yes" or "no".
3. Use statistic method: if the standard deviation of data at each item is big we think that people use 10 consequential regions to judge it while if the standard deviation is small we suppose people use 10000 regions.

For 10000 consequential regions and 20000, procedures are all the same.

However, in 10 vs. 10000, we can notice a distinct variance while in 10000 vs. 20000, there is very little difference.