Predicting Relative Position of a Continuous Random Variable

Preethi, Atreya, Kailash

March 10, 2023

1 Problem Statement

Person A draws n = 100 independent realizations of a continuous random variable and ranks them in increasing order from 1 to 100. Subsequently, Person B draws a single value from the same population and inserts this value into the rank order created earlier by Person A. For example, if her value is such that 50 of Person A's draws are smaller and 50 are larger, then the rank associated with her draw would be 51 , that is, overall, her value would be the 51st in increasing order. Or, if her value is smaller than all 100 of Person A's, then the rank 1 would be associated with it.

a. Is it more likely that Person B's value will occupy rank 51 than rank 1?

b. Derive for general n the probability that Person B's value will occupy rank k, where $1 \leq k \leq n+1$

2 Motivation

In this problem, we demonstrate how the use of Monte - Carlo simulation can make the problem much simpler to solve.An experiment performed once may not give the expected result and often in such cases we are required to perform it multiple times to obtain the most probable result.For the above problem we are asked what the probable rank of a random variable could be, we solve this by plotting a graph for the frequency of obtaining a particular rank when the experiment is repeated multiple times.

3 Solution

Claim 1 We claim that the probability that b will be between x_1 and x_2 is $x_2 - x_1$. We also claim that the probability that b will occupy the first rank is E_1 and rank 51 is E_{51}-E_{50}.

Claim 2 We claim that the expected position of the particle having the n^{th} rank is

CDF = probability that k particles lie before x = $\int_0^1 {}^nC_k x^k (1-x)^{n-k} dx$.

Now, we have PDF = d/dx(CDF) = ${}^nC_k \cdot k \cdot y^{(k-1)} \cdot (1-y)^{(n-k)} -^n C_{k-1} * (k-1) * y^{(k-2)} * (1-y)^{(n-k)}$

Therefore the expected position of the particle is

$$\int_0^1 x * PDF(x) = \int_0^1 x \cdot ({}^nC_k \cdot k \cdot y^{(k-1)} \cdot (1-y)^{(n-k)} -^n C_{k-1} * (k-1) * y^{(k-2)} * (1-y)^{(n-k)}) \tag{1}$$

$$= \frac{k}{n+1} \tag{2}$$

We thus get that on average the particles are equally spaced over the region 0, 1. Hence, it is equally probable to find the particle having the first rank as it is to find the particle having the 51^{st} rank since the interval from 0 to expected value of the first particle is the same as the interval of the 50th particle to the 51st particle.

3.1 For part 2

We get that since each sub-interval has equal length, the probability that the person B will occupy rank k is 1/n+1

3.2 Results

- It is equally likely for person B to get rank 1 and rank 51.
- The general formula for person B occupying rank k is 1/n+1.

3.3 Code

```
import random

l=100000  #number of Monte-Carlo iterations


cnt=[0]*100 #how many times Person B chose rank "i"


for k in range(l):

    n=[] #contains 100 random variables, later arranged by rank

    for i in range(100):
        n.append(random.random()) #inserting random variables between 0 and 1

    n.sort() #sorting in ascending order
    x=random.random() #Person B chooses from the same random variable pool

    for i in range(100):
        if x<n[i]:
            cnt[i]+=1 #increase the ith position by 1 since that is the rank in this iteration
            break



file=open("expecwaittime.txt", "w")

for m in range(100):
    file.write(str(m+1)+'\t') #rank
    file.write(str(cnt[m])+'\n') #number of times the rank showed up in 100,000 Monte-Carlo experiments
    m+=1


file.close()
```

Listing 1: Code used for question 5.