Documentation For Research Topic 1

Victor Huang

November 8, 2020

Overview:

The purpose of the R program is to use R function to perform Monte-Carlo draw on randomly generated population.
The following input parameters are defined as:
`nSims: number of simulation`
`N: Population size`
`n: Sample size`
`y: population generated from N(theta,sigma2)`
`theta: Mean of y`
`sigma2: Variance of y`
`Y: Sample selected without replacement from y`
`theta_bar: Mean of sample`
`theta_var: Variance of Monte Carolo Draws`

set_para.R

set_para.R defines the function: `set_para`. It takes `theta` and `sigma2` as inputs and outputs a vector representing the parameters of the population. `theta` will be treated as the popluation mean and `sigma2` will be treated as the population variance.

pop_gen.R

pop_gen.R defines the function: `pop_gen`. It takes a vector generated in **set_para.R** and `N`, a numeric input representing the popultion size. The function outputs a vector of length `N` with each element following iid Normal distribution, specified by the input parameter. It also requires a `type` variable which will specify which distribution to draw population from.

pop_draw.R

pop_draw.R defines the function `pop_draw`. It takes a vector of numbers generated in **pop_gen.R** as the population, and a number `n` as the sample size. It will draw `n` elements from the population with replacement and output these draws as a new vector, representing the random samples.

simulation.R

simulation.R defines the function `simulation`. It takes `nSims`, `N`, `n`, `theta`, and `sigma2` as inputs. The function first calls `set_para` to generate a parameter vector based on `theta` and `sigma2`. Then it calss

Table 1: Sample Output

	sampleMean	sampleVar
Observed	0.3889207	0.0116446
Theorethical	0.0000000	0.0500000

	x
nSims	10
theta	0
sigma2	1
N	20
n	10

	x
Expected_Mean	theta
Expected_Variance	(N-n)/(N*n)*sigma2

`pop_gen` to generate a population vector of size `N`, with each element following `N(theta,sigma2)`. Using `pop_draw`, it will generate `nSims` samples of size `n` from the population, and record each sample's mean. The function's output will be a 2 by 2 dataframe showing the observed sample mean, variance of the observed sample mean, the theoretical sample mean and the theoretical variance of the sample mean.

The following is an example output: