

STA5075: Practical 1

Jessica Stow (STWJES003@myuct.ac.za)

2025-02-03

Question 1

Calculate $\log(i + 1)$ for $i = 0$ to 100. The average of these values should = 3.647074.

```
i <- 0:100 # create vector i

# create function
x.log <- function(x) {
  return(log(i+1))
}

x.log(i)

## [1] 0.0000000 0.6931472 1.0986123 1.3862944 1.6094379 1.7917595 1.9459101
## [8] 2.0794415 2.1972246 2.3025851 2.3978953 2.4849066 2.5649494 2.6390573
## [15] 2.7080502 2.7725887 2.8332133 2.8903718 2.9444390 2.9957323 3.0445224
## [22] 3.0910425 3.1354942 3.1780538 3.2188758 3.2580965 3.2958369 3.3322045
## [29] 3.3672958 3.4011974 3.4339872 3.4657359 3.4965076 3.5263605 3.5553481
## [36] 3.5835189 3.6109179 3.6375862 3.6635616 3.6888795 3.7135721 3.7376696
## [43] 3.7612001 3.7841896 3.8066625 3.8286414 3.8501476 3.8712010 3.8918203
## [50] 3.9120230 3.9318256 3.9512437 3.9702919 3.9889840 4.0073332 4.0253517
## [57] 4.0430513 4.0604430 4.0775374 4.0943446 4.1108739 4.1271344 4.1431347
## [64] 4.1588831 4.1743873 4.1896547 4.2046926 4.2195077 4.2341065 4.2484952
## [71] 4.2626799 4.2766661 4.2904594 4.3040651 4.3174881 4.3307333 4.3438054
## [78] 4.3567088 4.3694479 4.3820266 4.3944492 4.4067192 4.4188406 4.4308168
## [85] 4.4426513 4.4543473 4.4659081 4.4773368 4.4886364 4.4998097 4.5108595
## [92] 4.5217886 4.5325995 4.5432948 4.5538769 4.5643482 4.5747110 4.5849675
## [99] 4.5951199 4.6051702 4.6151205

mean(x.log(i), na.rm = TRUE) # average = 3.647074

## [1] 3.647074
```

Question 2

Generate 10000 random values.

```
set.seed(20230130) # set starting point for random number generator
## generate 10000 values from an exponential distribution, mean = 1
y <- rexp(10000)
```

Question 3

Find the largest number and its position. (Answer: 9.131717, 9293)

```
max(y) # find the maximum number in vector y # 9.131717
```

```
## [1] 9.131717
```

```
which(y == max(y)) # find its position # 9293
```

```
## [1] 9293
```

Question 4

How many values are > 2 (absolute and %)? (Answer: 1333, 13.33%)

```
length(which(y > 2)) # 1333
```

```
## [1] 1333
```

```
length(which(y > 2))/length(y)*100 # 13.33
```

```
## [1] 13.33
```

Question 5

y2: select every 2nd element of y, starting from 1st.

```
y2 <- y[seq(from = 1, to = length(y), by = 2)]
```

Question 6

y3: replace values > 3 in y2 with 3.

```
y3 <- y2
y3[y3 > 3] <- 3 # replace values greater than 3 with 3
mean(y3) # average = 0.9362115
```

```
## [1] 0.9362115
```

Question 7

Find the variance of y3 using vector operations. Check with var().

```
sum((y3-mean(y3))^2)/(length(y3)-1) # manual calculation of the variance
```

```
## [1] 0.7019085
```

```
var(y3) == sum((y3-mean(y3))^2)/(length(y3)-1) # TRUE
```

```
## [1] TRUE
```