computer_proj2

Harrison Halesworth

2025-09-10

```r
# Timing how long this takes for the fun of it
start <- Sys.time()

# Set m to be the number of envelopes and letters in the problem
m <- c(10, 1000, 10000, 100000)

# Set n to be number of simulations of the matching problem
n <- 100000

# Store the empirical probabilities for the three trials
probability <- c(0,0,0)

# Simulate for different numbers of letters and envelopes
for (i in 1:length(m)) {
  # Initialize count to 0
  # This will keep track of how many trials exceed more than 1 match and will be updated as such
  count <- 0
  for (j in 1:n) {
    # Set up letters vector of size n
    letters <- c(1:m[i])

    # Set up random permutation of envelopes of size n
    envelopes <- sample(letters, m[i])

    # Sum up the number of matches in the simulation
    matches <- sum(letters==envelopes)

    # If the number of matches is greater than or equal to 1 then increase the count
    if (matches >= 1) {count <- count + 1}
  }
  probability[i] <- count / n
}

# Calculate code duration
end <- Sys.time()
duration <- end - start

print(probability)
```

```
## [1] 0.63404 0.63443 0.63074 0.63047
```

```r
print(duration)
```

```
## Time difference of 31.17328 mins
```

Remarks

As we can see by the results above, via 100000 trial simulations for $n = 10, 1000, 10000, 100000$ envelopes/letters, the probability of at least one match is ≈ 0.632, which is consistent with what we found in class. We can now look to calculate it via the explicit formula we derived in class and see if the values are consistent.

```r
# Calculate probability of at least one match via explicit formula we derived from class

# Initialize empty vector to hold calculated probabilities
probability_calc <- c(0,0,0,0)

# Loop through each value of m to calculate sum
for (k in 1:length(probability_calc)) {
  # Calculate sum for m envelopes/letters
  for (l in 1:m[k]) {
  probability_calc[k] <- probability_calc[k] + ((-1)^(l+1))/factorial(l)
  }
}

print(probability_calc)
```

```
## [1] 0.6321205 0.6321206 0.6321206 0.6321206
```

Remarks

As we can see, the calculated probabilities agree to an extent with the probabilities obtained from our simulation in the previous part. As we expect, the proximity of our simulated probability and calculated probability is the closest when we have $m = 100000$ envelopes, which is the closest to infinity of the values we tried out in this experiment.