Assessment 07 - The Big Short

Gabriele Mineo - Harvard Data Science Professional

Bank earnings

Say you manage a bank that gives out 10,000 loans. The default rate is 0.03 and you lose $200,000 in each foreclosure.

Create a random variable S that contains the earnings of your bank. Calculate the total amount of money lost in this scenario.

Instructions

- Using the `sample` function, generate a vector called `defaults` that contains `n` samples from a vector of `c(0,1)`, where `0` indicates a payment and `1` indicates a default
- Multiply the total number of defaults by the loss per foreclosure.

```
# Assign the number of loans to the variable `n`
n <- 10000

# Assign the loss per foreclosure to the variable `loss_per_foreclosure`
loss_per_foreclosure <- -200000

# Assign the probability of default to the variable `p_default`
p_default <- 0.03

# Use the `set.seed` function to make sure your answer matches the expected result after random samplin
set.seed(1)

# Generate a vector called `defaults` that contains the default outcomes of `n` loans
defaults <- sample(c(0,1), n, replace=TRUE, prob=c(1-p_default, p_default))

# Generate `S`, the total amount of money lost across all foreclosures. Print the value to the console.
S <- sum(loss_per_foreclosure * defaults)

print(S)
```

```
## [1] -6.3e+07
```

Bank earnings Monte Carlo

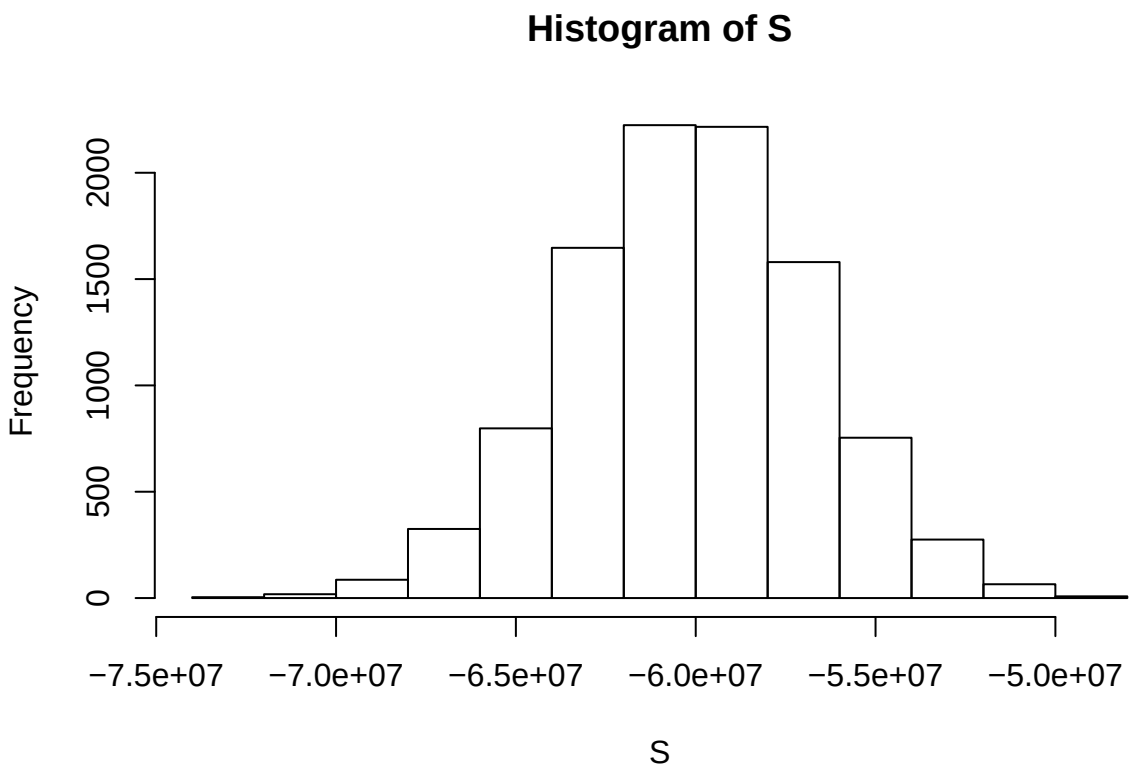
Run a Monte Carlo simulation with 10,000 outcomes for S, the sum of losses over 10,000 loans. Make a histogram of the results.

Instructions

- Given the probability of default, use the function `sample` to generate a list of 10,000 loan outcomes: payment (0) or default (1).
- Given the cost of each default, use the function `sum` to return the sum of all losses across the 10,000 loans.
- Repeat the previous steps a total of 10,000 times.
- Plot the histogram of values using the function `hist`.

```r
# Assign the number of loans to the variable `n`
n <- 10000

# Assign the loss per foreclosure to the variable `loss_per_foreclosure`
loss_per_foreclosure <- -200000

# Assign the probability of default to the variable `p_default`
p_default <- 0.03

# Use the `set.seed` function to make sure your answer matches the expected result after random samplin
set.seed(1)

# The variable `B` specifies the number of times we want the simulation to run
B <- 10000

# Generate a list of summed losses 'S'. Replicate the code from the previous exercise over 'B' iteratio

S <- replicate(B, {
    defaults <- sample( c(0,1), n, prob=c(1-p_default, p_default), replace = TRUE)
  sum(defaults * loss_per_foreclosure)
})

# Plot a histogram of 'S'
hist(S)
```

```
## Warning in n * h: si è prodotto un NA per overflow di interi
```



Bank earnings expected value

What is the expected value of S, the sum of losses over 10,000 loans? For now, assume a bank makes no money if the loan is paid.

Instructions

- Using the chances of default (`p_default`), calculate the expected losses over 10,000 loans.

```
# Assign the number of loans to the variable `n`
n <- 10000

# Assign the loss per foreclosure to the variable `loss_per_foreclosure`
loss_per_foreclosure <- -200000

# Assign the probability of default to the variable `p_default`
p_default <- 0.03

# Calcualte the expected loss due to default out of 10,000 loans
n*(p_default*loss_per_foreclosure + (1-p_default)*0)
```

```
## [1] -6e+07
```

Bank earnings standard error

What is the standard error of S?

Instructions

- Compute the standard error of the random variable `S` you generated in the previous exercise, the summed outcomes of 10,000 loans.

```
# Assign the number of loans to the variable `n`
n <- 10000

# Assign the loss per foreclosure to the variable `loss_per_foreclosure`
loss_per_foreclosure <- -200000

# Assign the probability of default to the variable `p_default`
p_default <- 0.03

# Compute the standard error of the sum of 10,000 loans
sqrt(n) * abs(loss_per_foreclosure) * sqrt(p_default*(1 - p_default))
```

```
## [1] 3411744
```

Bank earnings interest rate - 1

So far, we've been assuming that we make no money when people pay their loans and we lose a lot of money when people default on their loans. Assume we give out loans for $180,000. How much money do we need to make when people pay their loans so that our net loss is $0?

In other words, what interest rate do we need to charge in order to not lose money?

Instructions

- If the amount of money lost or gained equals 0, the probability of default times the total loss per default equals the amount earned per probability of the loan being paid.
- Divide the total amount needed per loan by the loan amount to determine the interest rate.

```
# Assign the loss per foreclosure to the variable `loss_per_foreclosure`
loss_per_foreclosure <- -200000

# Assign the probability of default to the variable `p_default`
p_default <- 0.03

# Assign a variable `x` as the total amount necessary to have an expected outcome of $0
x <- -(loss_per_foreclosure*p_default) / (1 - p_default)

# Convert `x` to an interest rate, given that the loan amount is $180,000. Print this value to the cons
x / 180000
```

```
## [1] 0.03436426
```

Bank earnings interest rate - 2

With the interest rate calculated in the last example, we still lose money 50% of the time. What should the interest rate be so that the chance of losing money is 1 in 20?

In math notation, what should the interest rate be so that $\Pr(S < 0) = 0.05$?

Remember that we can add a constant to both sides of the equation to get:

$$\Pr\left(\frac{S-\mathrm{E}[S]}{\mathrm{SE}[S]} < \frac{-\mathrm{E}[S]}{\mathrm{SE}[S]}\right)$$

which is

$$\Pr\left(Z < \frac{-[lp+x(1-p)]n}{(x-l)\sqrt{np(1-p)}}\right) = 0.05$$

Let `z = qnorm(0.05)` give us the value of `z` for which:

$$\Pr(Z \leq z) = 0.05$$

Instructions

- Use the `qnorm` function to compute a continuous variable at given quantile of the distribution to solve for z.
- In this equation, l, p, and n are known values. Once you've solved for `z`, solve for `x`.
- Divide `x` by the loan amount to calculate the rate.

```
# Assign the number of loans to the variable `n`
n <- 10000

# Assign the loss per foreclosure to the variable `loss_per_foreclosure`
loss_per_foreclosure <- -200000

# Assign the probability of default to the variable `p_default`
p_default <- 0.03

# Generate a variable `z` using the `qnorm` function
z <- qnorm(0.05)

# Generate a variable `x` using `z`, `p_default`, `loss_per_foreclosure`, and `n`
x <- -loss_per_foreclosure*( n*p_default - z*sqrt(n*p_default*(1 - p_default)))/ ( n*(1 - p_default) + z
```

```
# Convert `x` to an interest rate, given that the loan amount is $180,000. Print this value to the cons
x / 180000
```

```
## [1] 0.03768738
```

Bank earnings - minimize money loss

The bank wants to minimize the probability of losing money. Which of the following achieves their goal without making interest rates go up?

Possible Answers

- A smaller pool of loans
- A larger probability of default
- A reduced default rate [X]
- A larger cost per loan default