Example 3.31

This part will provide an `alternative mathematic model` for calculate the probability of each outcome occurs at least one time.

Instead of using venn diagram to solve this problem, example 3.31 using `conditional probability` to calculate P(n,k), which `n` means the execution time, and `k` means the categories of outcome.

How to use

- Requirement
  - Linux - `ubuntu 16.04`
  - gnuplot
- Testbed
  - categories of outcome: `5` (e.g. `range` in *example 2.5*)
  - testcase(for *example 2.5* simulation): `10^6`
  - execution times(for *example 3.31* mathematic model): `30`
    - When execution times increase, the more time will spend

on this program.

- The testbed is built with Makefile, so just need to run `make` and all the program will be compile and run automatically.

```
make
```

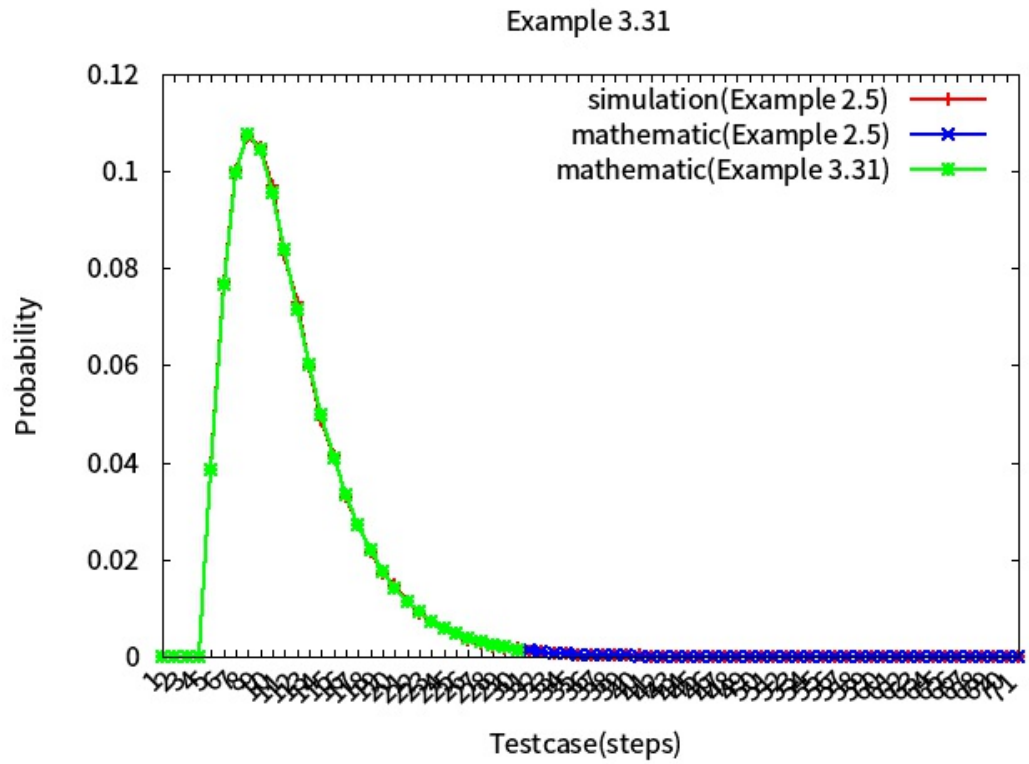
Run the program

- Let's see the content in `Makefile`:
  - **First**:
    - Compile the new mathematic model program (under `example 3.31`) , mathematic model program (under `example 2.5`),and the simulation program (under `example 2.5`)
  - **Second**:
    - Similar to the demo in `example 2.5`, we run the simulation program first, and store the result into `simulation.output`.
    - And then run the mathematic model program (both `2.5` and `3.31`), and directly pipe their output content into `mathematic.output`,`mathematic_3_31.output` respectively.
  - **Last**:
    - After generating all the result into output file with specified format, then we can using the gnuplot script - `result.gp` to plot the result into `PNG` format.

Result

- As you can see the figure down below, the `simulation(red line)` and the `mathematic(green line)` have the similar curve.

- Run with testcase `10^6`:



- That we can say this mathematic model has the correct answer!

Comparison/Time Complexity

> This part will need to install the chrome extension: **`Github with MathJax`**, then can see the correct mathematic formula

Formula

- We can now see the mathematic model between 2 example (e.g. `2.5` & `3.31`)

- *Mathematic model in `Example 2.5`:*

$$P(X=N) = \sum_{i=1}^m P_i (1-P_i)^{n-1} - \sum_{i<j} (P_i+P_j)(1-Pi-Pj)^{n-1} + \sum_{i<j<k} (P_i+P_j+P_k)(1-P_i-P_j-P_k)^{n-1} - \dots $$

- *Mathematic model in `Example 3.31`:*

$$P(m,r) = \sum_{j=1}^{m-r+1} P(m-j,r-1) \cdot C_j^m \cdot (\frac{P_r}

{\sum_{j=1}^{r} P_j})^j \cdot (1 - \frac{P_r}{\sum_{j=1}^{r} P_j})^{m-j}$$

$$Starting \ from: P(m,1) = 1, if \ m>=1 ; P(m,1) = 0, if \ m==0$$

Time Complexity Analysis

Mathematic model in `Example 2.5`

- In the implementation of `example2.5/mathematic.cc`, we can see the formula above. In this term , we have number `r` of combination `C(r,i)`, which `r` represent as the number of categories; And **total number of terms** in formula will be: $$\sum_{i=1}^r C_i^r$$
- And we can see the terms of combinations : $$\sum_{k=0}^n C_k^n = 2^n$$, so in our terms: $$\sum_{k=1}^n C_k^n = 2^n - 1$$
- Then we can get the time complexity will be: $$O(\sum_{i=1}^r C_i^r) = O(2^r-1)=O(2^r)$$
- And why using `combination`? Because as the formula above, we can see each term has: $$\sum_{i<j<k...<r}, r = number\ of\ categories$$, so we can use `combination` rather than `permutation`.

Mathematic model in `Example 3.31`

- In the implementation of `example3.31/mathematic.cc` is using several `recursive` functions to construct the probability result.
- Totally we have number of **(m-r)** terms in **`P(m,r)`**, and the result will be the **combination with repetition**: $$H_r^{m-r}$$, then we can transform into combination :$$C_r^{(m-r)+r-1} = C_r^{m-1} = \frac{(m-1)!}{r! \cdot (m-r-1)!}= \frac{(m-1)\cdot(m-2) ...(m-r)}{r!} $$
- After eliminate `r!`, we can have the polynomial function `m^r`: $$O(\frac{(m-1)\cdot(m-2) ...(m-r)}{r!}) := O((m-1)\cdot(m-2) ...(m-r)) := O(m^r)$$, which `m` represent **`total execution times`**, and `r` represent **`number of categories`**

Comparion with Example `2.5`,`3.31`

	Time Complexity
`Example 2.5`	$$O(2^r)$$
`Example 3.31`	$$O(m^r)$$

As the table showing above, we can see **when the execution time(e.g. `m`) increase, the mathematic model program of `Example 3.31` will more time than `2.5` to complete.**

Author

- Kevin Cyu, kevinbird61@gmail.com