Variate Generator Library

Feng Wang
feng.wang@uni-ulm.de

March 6, 2014

Contents

1 Getting Started

1.1 Welcome

Welcome to Variate Generator library!

By the time you have read through this tutorial, you will be able to play with it.

1.2 Compilation Enviroment Requirement

As some C++11* features are employed when implementing this library, before we get further, please check your compiler for C++11 compatablity.

1.3 A Quick Example

The code listed in Table 1 shows how to generate gaussian random numbers:

*Features such as lambda functions, variadic template and keyword auto, see http://www.open-std.org/jtc1/sc22/wg21/ for more information.

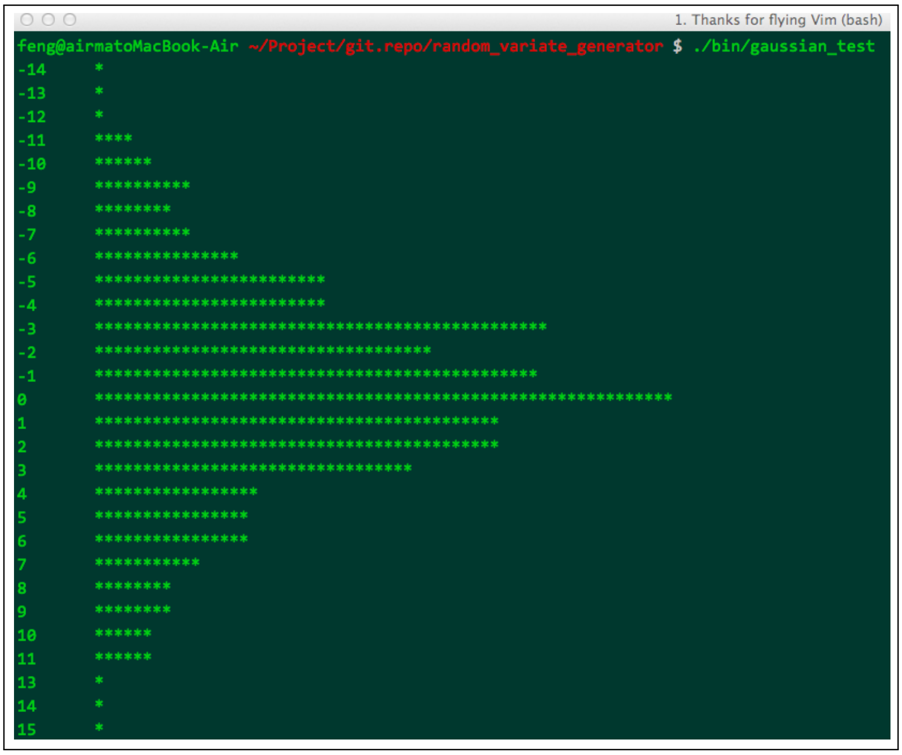

Figure 1: Gaussian Random Number Example

Table 1: Source Code for a Gaussian Variate Example

```cpp
#include <vg.hpp>
#include <cmath>
#include <map>
#include <iostream>
int main()
{
    //generate double precision gaussian random numbers
    //using mt19937 as random generator engine with arguments (0,4)
    vg::variate_generator<double, vg::gaussian, vg::mt19937> vg(0, 4);
    std::map< int, int > sample;
    //generate 500 gaussian numbers and store them in a map
    for ( auto i = vg.begin(); i != vg.begin()+500; ++i )
        sample[std::round(*i)]++;
    //show the number generated
    for ( auto i = sample.begin(); i != sample.end(); ++i )
    {
        std::cout << (*i).first << "\t";
        for ( auto j = 0; j < (*i).second; ++j )
            std::cout << "*";
        std::cout << "\n";
    }
    return 0;
}
```

As this library is header file only, if your c++ compiler is g++, a typical compilation and link command for the example code whose file name is *test_gaussian.cc* can be

g++ -I*PATH_TO_THE_HEADER* -o ./bin/gaussian_test test_gaussian.cc -std=c++11

This command will generate a executable file *gaussian_test* in directory *./bin*, and its executation result is shown in figure 1

2 An Overview of VG

A random variate generator consists of three parts:

- variate type
- distribution type
- engine type

2.1 struct variate_generator

The basic struct used for a generator is variate_generator, which is decleared as:

Table 2: variate_generator struct

```
namespace vg
{
    template
    <
     class T = double,                                   //variate
     template<class, class> class Distribution = uniform,//distribution
     class Engine = mitchell_moore                      //engine
    >
    struct variate_generator
    {
       template< typename ... Tn >
       variate_generator( const Tn ... );
       T operator()() const;
       operator T() const;
       iterator begin() const;
    };
};
```

2.1.1 declaration

So to make a generator to product variates of int type and lagarithmic distribution, with a parameter 0.33, we can simply declare:

```
vg::variate_generator<int, vg::lagarithmic> v( 0.33 );
```

which it is equivalent to

```
vg::variate_generator<int, vg::lagarithmic, vg::mitchell_moore> v( 0.33, 0 );
```

where the last argument 0 is the default engine seed.

Also, to make a generator to product variates of hypergeometric distribution, with int type and paramenters (200, 200, 200), using mt19937 persudo–random engine and engine seed 987654321, we can declare it with one line code like this:

```
vg::variate_generator<int, vg::hypergeometric, vg::mt19937> v( 200, 200, 200, 987654321 );
```

2.1.2 Generation

After the generator v has been declared, we can generate variate in several ways:

Generate only one variate:

```
auto i = v();
int j = v;
auto k = *(v.begin());
```

Generate multiple variates:

```
std::vector<int> array1( v.begin(), v.begin()+100);
std::vector<int> array2( 100 );
std::generate( array2.begin(), array2.end(), v );
std::vector<int> array3;
std::copy( v.begin(), v.begin()+100, std::back_inserter( array3 ) );
```

2.2 Built-in Distributions

Curent we have about fifty distributions implemented:

- arcsine distribution 3.1
- bernoulli distribution
- beta distribution
- beta_binomial distribution
- beta_pascal distribution
- binomial distribution
- burr distribution
- cauchy distribution
- chi_square distribution
- digamma distribution
- erlang distribution
- exponential distribution
- exponential_power distribution
- extreme_value distribution
- f distribution
- factorial distribution
- gamma distribution
- gaussian distribution
- gaussian_tail distribution
- generalized_hypergeometric_b3 distribution
- generalized_waring distribution
- geometric distribution
- grassia distribution
- gumbel_1 distribution
- gumbel_2 distribution
- hyperbolic_secant distribution
- hypergeometric distribution

- inverse_gaussian distribution
- inverse_polya_eggenberger distribution
- lambda distribution
- laplace distribution
- levy distribution
- list distribution
- logarithmic distribution
- logistic distribution
- lognormal distribution
- mizutani distribution
- negative_binomial distribution
- negative_binomial_beta distribution
- normal distribution
- pareto distribution
- pascal distribution
- pearson distribution
- planck distribution
- poisson distribution
- polya distribution
- polya_aeppli distribution
- rayleigh distribution
- rayleigh_tail distribution
- singh_maddala distribution
- t distribution
- teichroew distribution
- triangular distribution
- trigamma distribution
- uniform distribution 3.4
- von_mises distribution
- wald distribution
- waring distribution
- weibull distribution
- yule distribution
- zipf distribution

2.3 Engines

Currently we have 3 engines implemented:

- linear_congruential
- mitchell_moore
- mt19937

3 Distributions

3.1 Arcsine Distribution

Arcsine distribution is a special case of beta distribution when $\alpha = \beta = 0.5$.

3.1.1 Characterization

Probability density function is

$$f(x) = \begin{cases} \frac{1}{\pi\sqrt{x(1-x)}} & \text{if } 0 < x < 1 \\ 0 & \text{otherwise} \end{cases} \tag{1}$$

Cumulative distribution function is

$$F(x) = \begin{cases} 0 & \text{if } x < 0 \\ \frac{2\arcsin(\sqrt{x})}{\pi} & \text{if } 0 \leq x < 1 \\ 1 & \text{otherwise} \end{cases} \tag{2}$$

3.1.2 Usage

Arcsine variate generator is supposed to be initialized with seed s, which is 0 by default.

Table 3: arcsine distribution example code

```
#include <vg.hpp>
#include "test.hpp"
#include <vector>
#include <cstddef>

int main()
{
    vg::variate_generator<double, vg::arcsine, vg::mt19937> vg_;

    std::size_t n = 10000000;
    std::vector<double> x(n);

    std::generate( x.begin(), x.end(), vg_ );

    test( x.begin(), x.end(), "arcsine", 0.5, 0.125, 0);

    return 0;
}
```

Table 4: arcsine distribution example code output

	Mean	Variance	Skewness
Theory	0.500000000000000	0.125000000000000	0.00000000000000
Generated	0.500215862149786	0.125018618223430	-0.000983748428920325

3.2 Bernoulli Distribution

The Bernoulli distribution is a discrete distribution having two possible outcomes labelled by $n=0$ and $n=1$, in which $n=1$("success") occurs with probability p and $n=0$("failure") occurs with probability $q=1-p$, where $0<p<1$.

3.2.1 Characterization

Probability density function is

$$f(k,p)=\begin{cases} p & \text{if } k=1 \\ 1-p & \text{otherwise} \end{cases} \tag{3}$$

Cumulative distribution function is

$$F(k,p)=\begin{cases} 0 & \text{if } k<0 \\ 1-p & \text{if } 0\le k<1 \\ 1 & \text{otherwise} \end{cases} \tag{4}$$

3.2.2 Usage

Bernoulli variate generator is supposed to be initialized with probability p and seed s, defalut values are $p=0.5$ and $s=0$.

Table 5: bernoulli distribution example code

```
#include <vg.hpp>
#include "test.hpp"
#include <vector>
#include <cstddef>
int main()
{
    vg::variate_generator<double, vg::bernoulli, vg::mt19937> vg_(0.5);
    std::size_t n = 10000000;
    std::vector<double> x(n);
    std::generate( x.begin(), x.end(), vg_ );
    test( x.begin(), x.end(), "bernoulli", 0.5, 0.25, 0 );
    return 0;
}
```

Table 6: bernoulli distribution example code output

	Mean	Variance	Skewness
Theory	0.500000000000000	0.250000000000000	0.00000000000000
Generated	0.500000200000000	0.249999999999885	-7.99999992907318e-07

3.3 Beta-binomial Distribution

Beta-binomial distribution is distributed as a binomial distribution with parameter p, where p is distribution with a beta distribution characterized by parameters α and β.

3.3.1 Characterization

Probability density function is

$$f(x,n,\alpha,\beta)=\frac{B(x+\alpha,n-x+\beta)\binom{n}{x}}{B(\alpha,\beta)} \tag{5}$$

Cumulative distribution function is

$$F(x,n,\alpha,\beta) = 1 - \frac{nB(\beta+n-x-1,\alpha+x+1)\Gamma(n)F_n(\alpha,\beta;x)}{B(\alpha,\beta)B(n-x,x+2)\Gamma(n+2)} \tag{6}$$

3.3.2 Usage

Beta-binomial variate generator is supposed to be initialized with parameters n, α, β and seed s, default values are $n=10$, $\alpha=1$, $\beta=1$ and $s=0$.

Table 7: beta-binomial distribution example code

```
#include <vg.hpp>
#include "test.hpp"
#include <cmath>
#include <vector>
#include <cstddef>
int main()
{
    vg::variate_generator<double, vg::beta_binomial, vg::mt19937> vg_( 100, 0.1, 0.2 );
    std::size_t n = 10000000;
    std::vector<double> x(n);
    std::generate( x.begin(), x.end(), vg_ );
    auto const& mean = []( double n, double a, double b ){ return n * a / ( a+b ); };
    auto const& variance = []( double n, double a, double b ) { return n * a * b * (a+b+n)/((a+b)*(a+b)*(a+b+1)); };
    auto const& skewness = []( double n, double a, double b ) { return ( a+b+n+n)*(b-a)/(a+b+2)*std::sqrt((1+a+b)/(n*a*b*(a+b))); };
    test( x.begin(), x.end(), "beta_binomial with n = 100, alpha = 0.1, beta = 0.2", mean(100, 0.1, 0.2), variance(100, 0.1, 0.2), skewness(100, 0.1, 0.2), -1.5 );
    return 0;
}
```

Table 8: beta-binomial distribution example code output

	Mean	Variance	Skewness
Theory	33.3333333333333	1714.52991452991	12.8188529096048
Generated	33.3402505000000	1714.34470241029	0.700684542622759

3.4 Uniform Distribution

Uniform distribution generates pseudo random variables that uniformly distributed within interval $[a,b]$.

3.4.1 Characterization

Probability density function is

$$f(x) = \begin{cases} \frac{1}{b-a} & \text{if } a \le x \le b \\ 0 & \text{otherwise} \end{cases} \tag{7}$$

Cumulative distribution function is

$$F(x) = \begin{cases} 0 & \text{if } x < a \\ \frac{x-a}{b-a} & \text{if } a \le x < b \\ 1 & \text{otherwise} \end{cases} \tag{8}$$

3.4.2 Usage

Uniform variate generator is supposed to be initialized with parameters a, b and seed s, default values are $a=0.0$, $b=1.0$ and $s=0$.

Table 9: uniform distribution example code

```
#include <vg.hpp>

#include "test.hpp"

```

```
#include <vector>
#include <cstddef>

int main()
{
    vg::variate_generator<double, vg::uniform, vg::mt19937> vg_(-1, 1);

    std::size_t n = 10000000;
    std::vector<double> x(n);

    std::generate( x.begin(), x.end(), vg_ );

    test( x.begin(), x.end(), "unoform", 0, 1.0/12, 0.0, -2 );

    return 0;
}
```

Table 10: uniform distribution example code output

	Mean	Variance	Skewness
Theory	0.00000000000000	0.333333333333333	0.00000000000000
Generated	-3.12801617330395e-06	0.333229960243464	-6.33178467221793e-05