The stochastic modeling

Karol Dziedziul

February 2011 Gdansk

A schedule of the lecture

- **Generation of randomly arranged numbers**
- **Monte Carlo**
- **Strong Law of Large Numbers SLLN**

We use SAS 9. For convenience programs will be given in an appendix. Programs will be also available on my personal website

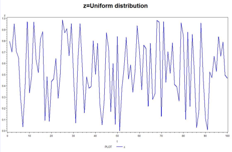
Random number

♣ Create a sequence of random numbers using Program 1. …
0.64828,0.71448, 0.37498, 0.15074, 0.12514, 0.79300, 0.13727, 0.61135, 0.47887, 0.63348, 0.60430, 0.49142, 0.83783, 0.57892, 0.87900, 0.64154, 0.13433, 0.58922, 0.82709, 0.69035, 0.85657, 0.95802, 0.61780, 0.43094, …

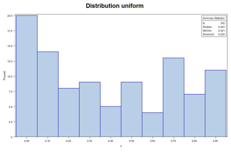
Numbers are chosen completely irregularly from the interval $[0,1]$. In 1927 random number tables was created by Tippett. Here we use a random number generator. Random numbers are drawn from a uniform distribution. It means that percentage of random numbers drawn from any interval $[a,b] \subset [0,1]$ approximately is equal to $(b-a)$.

Uniform Distribution

♣ Arrange a haphazard collection of $n = 100$ random numbers $u_1, u_2, \ldots, u_n$ use Program 1a.



♣ To make sure that random numbers are indeed drawn from a uniform distribution take $n = 1000$ and $n = 10000$ in Program 1a and see the histogram



Application of the random numbers, Monte Carlo

Such sequences of randomly numbers are helpful to calculate value of integrals. Let us consider

$$\int_0^1 sin(x)\, dx \approx 0.459697.$$

We can use the random numbers to do the following approximation

$$\int_0^1 sin(x)\, dx \approx \frac{1}{n}\sum_{j=1}^{n} sin(u_j).$$

♣ Use Program 2 to get an estimated value of the integral. Change n to obtain better accuracy. Here are some results. If $n = 100$

$$\int_0^1 sin(x)\, dx \approx 0.433 \approx 0.460.$$

If $n = 10000$

$$\int_0^1 sin(x)\, dx \approx 0.456 \approx 0.459.$$

Random numbers drawn from Beta distribution

Note that

$$\int_0^1 sin(x)4(1-x)^3dx \approx \frac{1}{n}\sum_{j=1}^{n} sin(u_j)4(1-u_j)^3.$$

♣ Use Program 2a to obtain an estimated value of the above integral.
But

$$\int_0^1 4(1-x)^3dx = 1.$$

This function is an example of a density in a class of Beta distributions, (SAS9 Beta(1,4)).
♣ Create the sequence of randomly arranged numbers $\xi_1, \xi_2, ..., \xi_n$ drawn from Beta distribution. Use Program 3. Observe the histogram. Use Program 4 to calculate the value of the integral by formula

$$\int_0^1 sin(x)4(1-x)^3dx \approx \frac{1}{n}\sum_{j=1}^{n} sin(\xi_j).$$

The probability density function of Beta distribution

The probability density function (density) of beta distribution with parameter $a, b > 0$ is given by

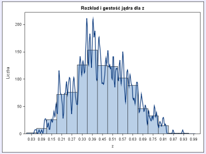
$$\beta_{a,b}(x) = \frac{1}{B(a,b)} x^{a-1}(1-x)^{b-1}, \quad 0 \leq x \leq 1,$$

where $B(a,b)$ is a normalizing constant. Note that the uniform distribution is an example of beta distribution for $a = b = 1$. For almost all parameters $a, b > 0$ the probability density function of beta distribution is a continuous function.

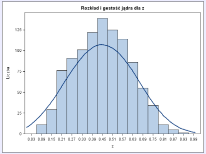
♣ To make a better estimation of the density than given by histogram apply Program 5.

Estimation of density of Beta distribution

♣ Notice the role of parameter bmw (so called bandwidth parameter) in smoothing of histogram. Change bmw in scale from 0.1 to 2.
If bmw=0.1 then



If bmw=2 then



The expected value and variance of Beta distribution

The expected value (expectation, mean) and variance of a Beta distribution radome variable X is given by:

$$EX = \int_0^1 x\beta_{a,b}(x)dx = \frac{a}{a+b}.$$

$$VarX = \int_0^1 x^2\beta_{a,b}(x)dx - (EX)^2 = \frac{ab}{(a+b)^2(a+b+1)}.$$

♣ Choose any $a, b > 0$ in Program 6 to check above formula. Notice the difference between true and simulated mean and variance.

SLLN Kolmogorov, Marcinkiewicz, Zygmund

In the above method (called Monte Carlo) we use the following Theorem.

Theorem

Let $\eta, \eta_1, \eta_2, \eta_3 \ldots$ *be a sequence of (i.i.d.) independent and identically distributed random variables, fix any* $p \in (0, 2)$*. Then*

$$\frac{1}{n^{1/p}} \sum_{j=1}^{n} \eta_j \quad \textit{converges a.s.(almost surely)}$$

if and only if

$$E|\eta|^p < \infty$$

and either $p \leq 1$ *or* $E\eta = 0$*. In that case the limit equals* $E\eta$ *for* $p = 1$ *and is otherwise* 0*.*

Applications Monte Carlo

Theorem

If we take $p = 1$ *and* $\xi, \xi_1, \xi_2, \xi_3 \ldots$ *a sequence of (i.i.d. rvs) such that* $E|f(\xi)| < \infty$ *then puting*

$$\eta_j = f(\xi_j)$$

we obtain for $n \to \infty$ *that*

$$\frac{1}{n} \sum_{j=1}^{n} f(\xi_j) \to Ef(\xi) \quad a.s.$$

Application 2

Theorem

If $X_j \geq 0$, $j = 1, 2...$ *and* $\frac{1}{n}\sum_{j=1}^{n} X_j$ *does not converge then* $EX = \infty$*!!!*

Example. Let W_t be a Wiener process. For W_t we denote by T_a the first hitting time of a level (a barrier) $a > 0$ by W_t, i.e.

$$T_a = \inf\{t > 0 : W_t = a\}.$$

Program 39. We know that for almost all scenarios $\omega \in \Omega$ a trajectory $W_t(\omega)$ reaches the level a. We observe that in given time a trajectory is reluctant to reach level a. We suspect that $ET_a = \infty$. How to verify it? Program 40 shows a distribution of T_a. Change time in which we observe the Wiener process and calculate $\frac{1}{n}\sum_{j=1}^{n} X_j$. Program 40a.Take $p = 1$

Application 3

If we suspect that there is $0 < p^* < 1$ such that for all $p < p^*$ $E|X|^p < \infty$ and $E|X|^p = \infty$ for $p > p^*$ then using SLLN we can estimate p^*. Program 40a. Take $p = 0.4$; $p = 0.5$; $p = 0.6$

SLLN - equivalent formulation

Theorem

Let $\eta, \eta_1, \eta_2, \eta_3 \ldots$ *be a sequence of (i.i.d. rvs) independent and identically distributed random variables, fix any* $p \in (0,2)$*. Then*

$$\frac{1}{n^{1/p}} \left(\sum_{j=1}^{n} \eta_j - an \right) \to 0 \quad a.s.(almost\ surely) \tag{1}$$

for some real a if and only if

$$E|\eta|^p < \infty.$$

In that case we have

$$a = \begin{cases} 0 & p < 1, \\ E\eta & p \in [1,2). \end{cases}$$

Problem

Let

$$E|\eta|^p < \infty$$

Note that

$$\frac{1}{n^{1/p}}\left(\sum_{j=1}^{n}\eta_j - an\right) = n^{1-1/p}\left(\frac{1}{n}\sum_{j=1}^{n}\eta_j - a\right)$$

What happens if $p \nearrow 2$ i.e. $1 - 1/p \searrow 1/2$? Let $\eta, \eta_1, \eta_2, \eta_3 ...$ be a sequence of (i.i.d. rvs) independent and identically distributed random variables, let $p = 2$.

$$\sqrt{n}\left(\frac{1}{n}\sum_{j=1}^{n}\eta_j - a\right) \to ??. \tag{2}$$

SLLN for order statistics

Theorem of this type was proved by van Zwet, Sen... For a random variable X with $E|X| < \infty$ and the distribution function F_X, expected shortfall at confidance level $\alpha \in [0,1)$ is defined as

$$ES_\alpha = \frac{1}{1-\alpha} \int_\alpha^1 q_u(F_X) du,$$

where $Q_u(F_X)$ is the quantile function of F_X.

Theorem

Let $X, X_1, X_2, \dots$ be random variables i.i.d. with above assumptions. We have

$$\lim_{n \to 0} \frac{1}{[n(1-\alpha)]} \sum_{j=1}^{[n(1-\alpha)]} L_{j,n} = ES_\alpha \quad a.e.,$$

where $L_{1,n} \geq L_{2,n} \geq \cdots \geq L_{n,n}$ are the order statistics of $L_1, L_2, \dots, L_n$ and where $[n(1-\alpha)]$ denotes the largest integer not exceeding $n(1-\alpha)$.