Stochastic Structural Dynamics

Lecture-27

Monte Carlo simulation approach-3

Dr C S Manohar
Department of Civil Engineering
Professor of Structural Engineering
Indian Institute of Science
Bangalore 560 012 India
manohar@civil.iisc.ernet.in

Digital simulation of samples of random variables

Let X be a random variable with PDF $P_X(x)$.

How to generate samples $\{x_i\}_{i=1}^{n}$ of X on a computer so that the these numbers are statistically indistinguishable from realizations of the random variable X ?

Pseudorandom number generators

- Deterministic algorithms which produce outputs which are statistically indistinguishable from realizations of random variables.
- Starting point in digital simulation of random variables
- Random numbers are taken to mean numbers distributed uniformly in $[0,1]$.

Linear congruential generators

$$X_i = (aX_{i-1} + c) \bmod m; \ i = 1, 2, \cdots \quad X_0$$

$$R_i = \frac{X_i}{m} = \text{pseudorandom numbers}$$

a, c, m, X_0 : Integers to be specified by user

a : multiplier (>0)

c : increment (≥ 0)

m : modulus (>0)

X_0 : a seed (≥ 0)

$a, c, m, X_0 \in [0, m-1]$

modulo m : returns the remainder after dividing $(aX_{i-1} + c)$ by m.

Period $k \leq m$

Theorem

A congruential generator has full period m if and only if

(i) $\gcd(c, m) = 1$

(ii) $a \equiv 1 \mod p$ for each prime factor p of m

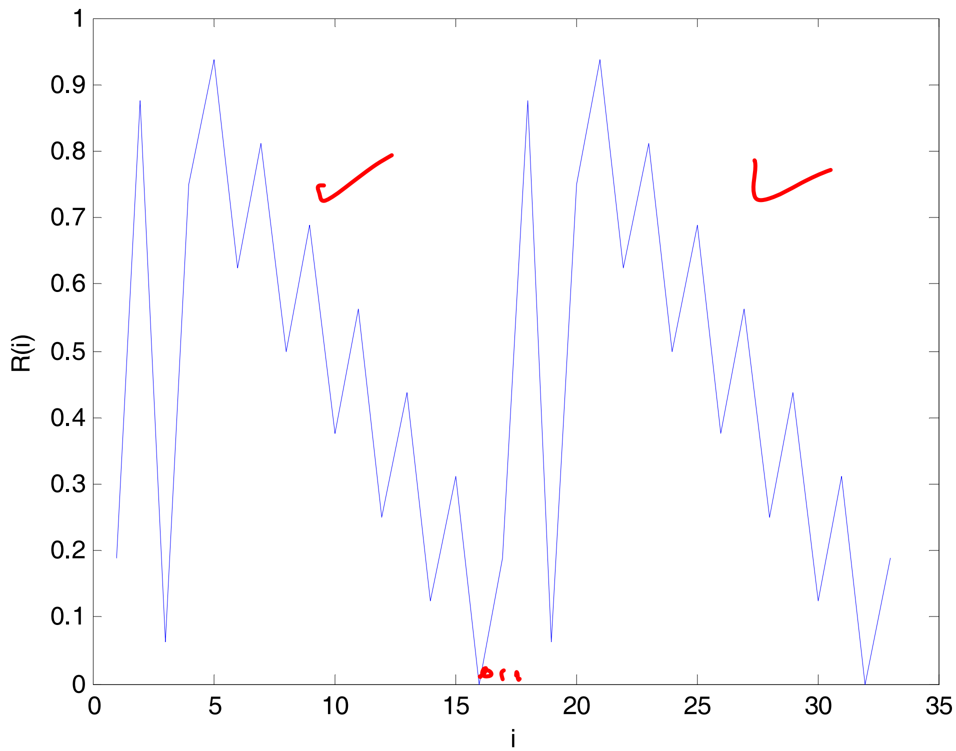
(iii) $a \equiv 1 \mod 4$ if 4 divides m

Example

$$X_i = (9X_i + 3) \bmod 2^4; X_0 = 3$$

$$R_i = \frac{X_i}{m}$$

Period=16



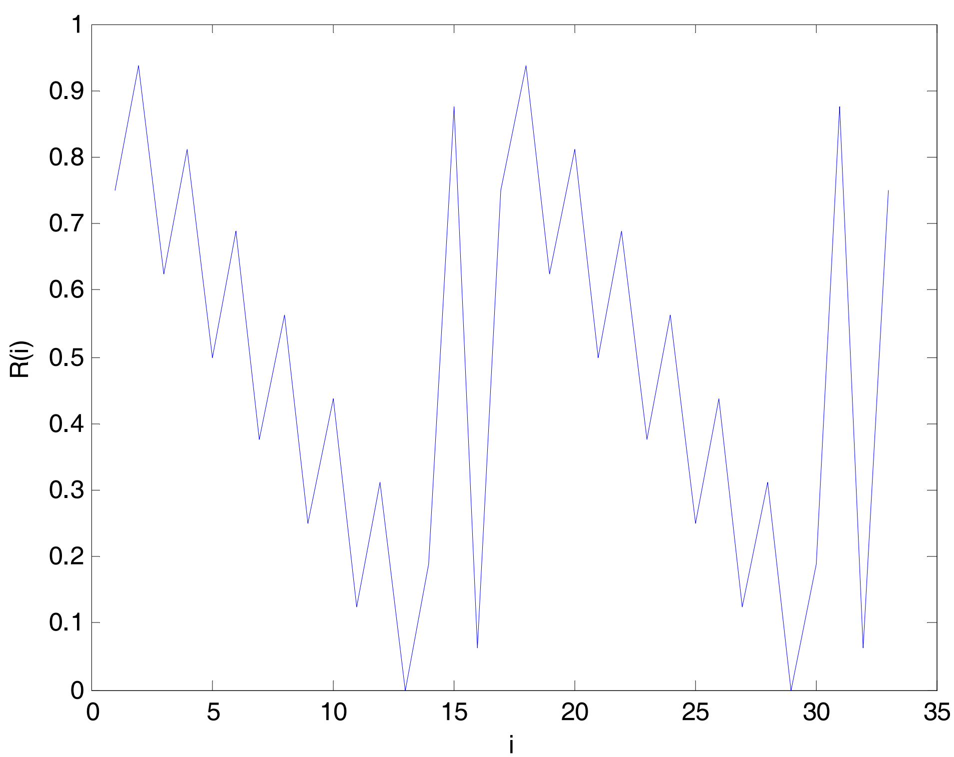
0.1875	0.1875
0.8750	0.8750
0.0625	0.0625
0.7500	0.7500
0.9375	0.9375
0.6250	0.6250
0.8125	0.8125
0.5000	0.5000
0.6875	0.6875
0.3750	0.3750
0.5625	0.5625
0.2500	0.2500
0.4375	0.4375
0.1250	0.1250
0.3125	0.3125
0	0

Example

$$X_i = (9X_i + 3) \bmod 2^4;\ X_0 = 12$$

$$R_i = \frac{X_i}{m}$$

Period=16



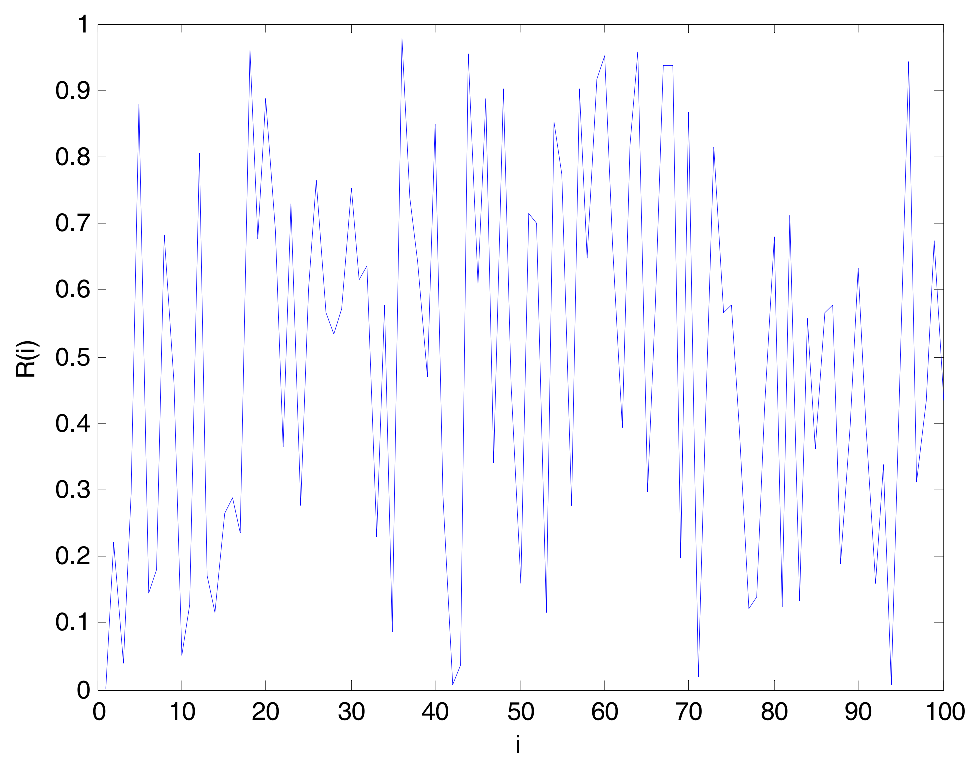
0.7500	0.7500
0.9375	0.9375
0.6250	0.6250
0.8125	0.8125
0.5000	0.5000
0.6875	0.6875
0.3750	0.3750
0.5625	0.5625
0.2500	0.2500
0.4375	0.4375
0.1250	0.1250
0.3125	0.3125
0	0
0.1875	0.1875
0.8750	0.8750
0.0625	0.0625

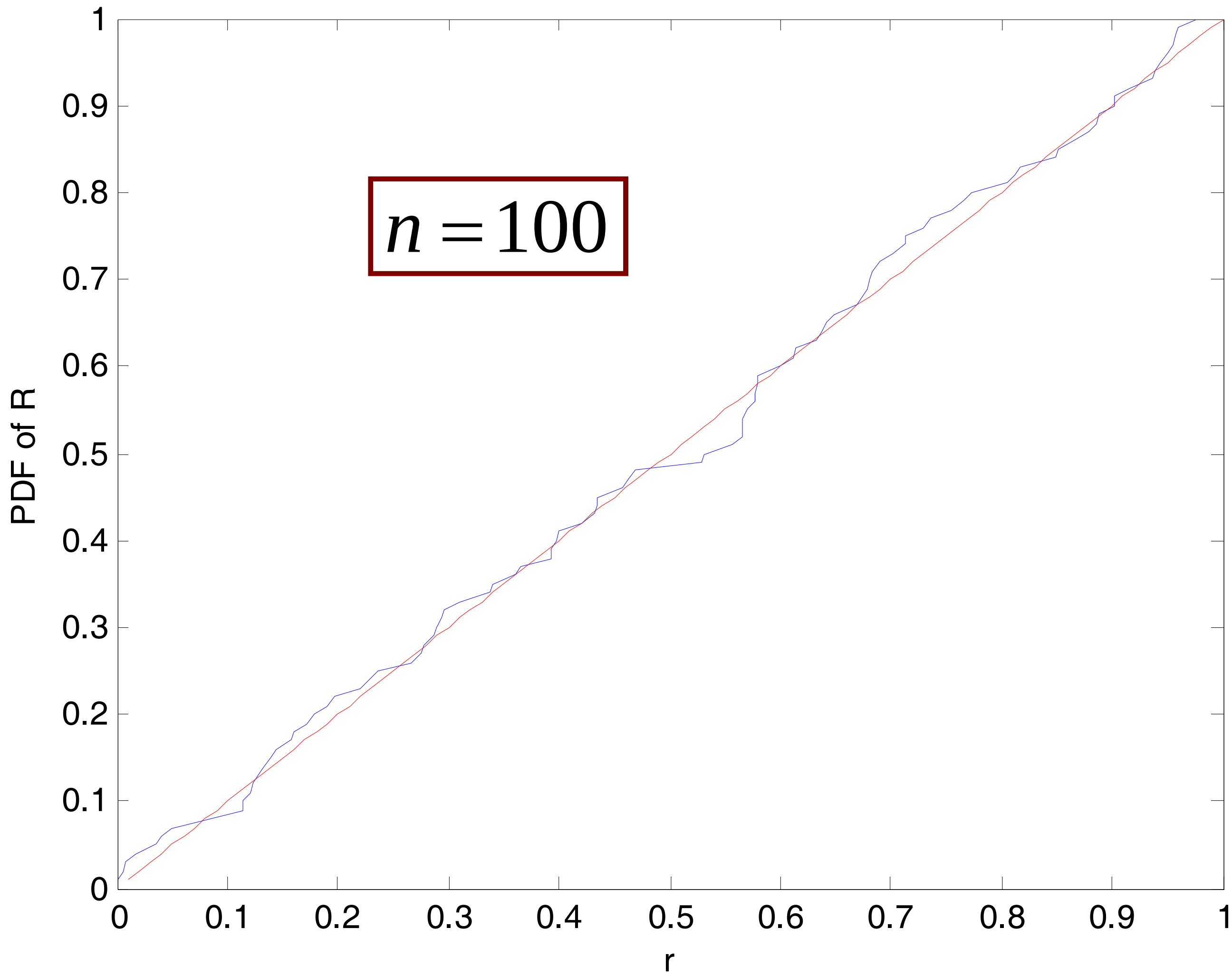
Example

$$X_i = (906185749 X_i + 1) \bmod 2^{31};\ X_0 = 12$$

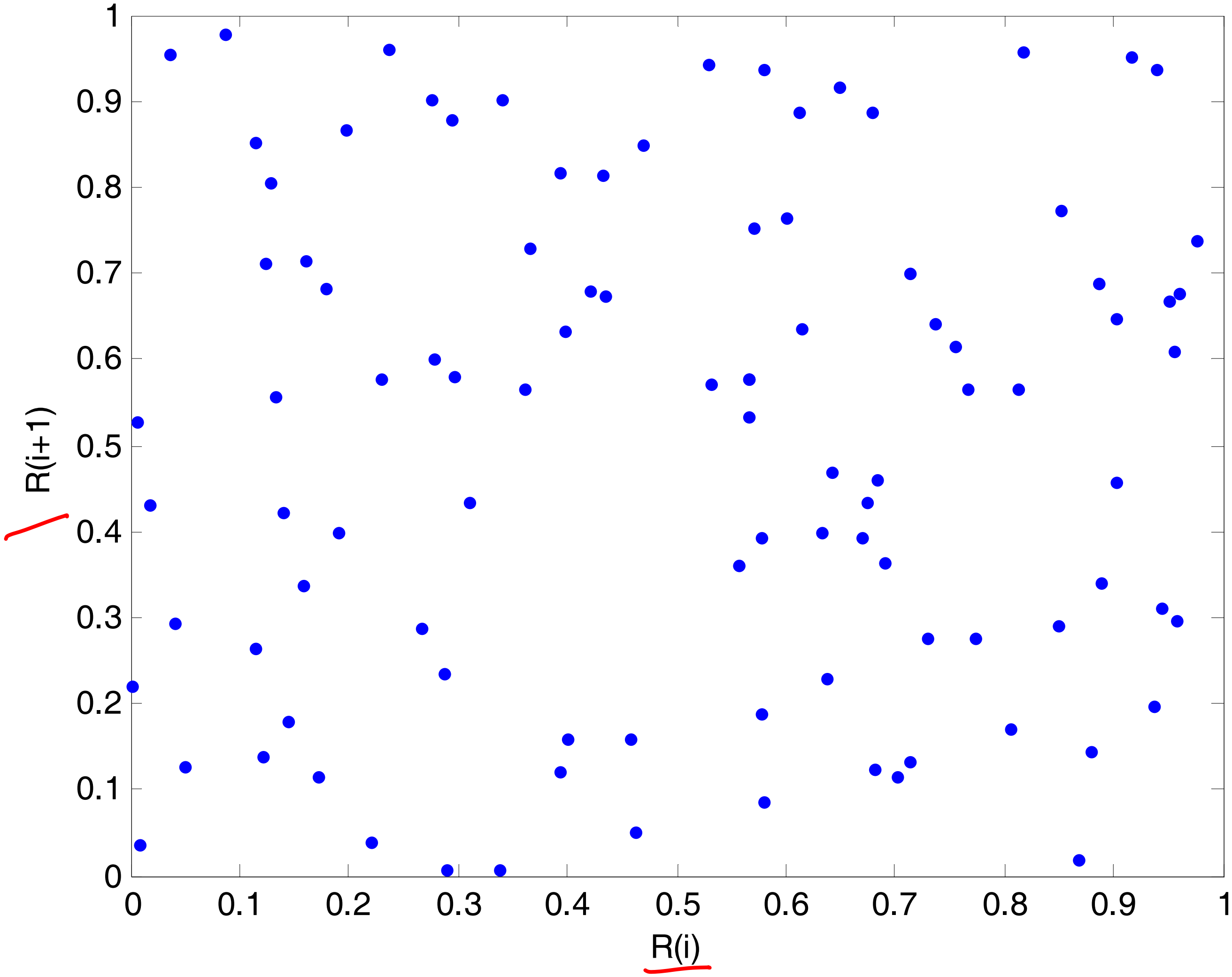
$$R_i = \frac{X_i}{m}$$

Period=2147483648

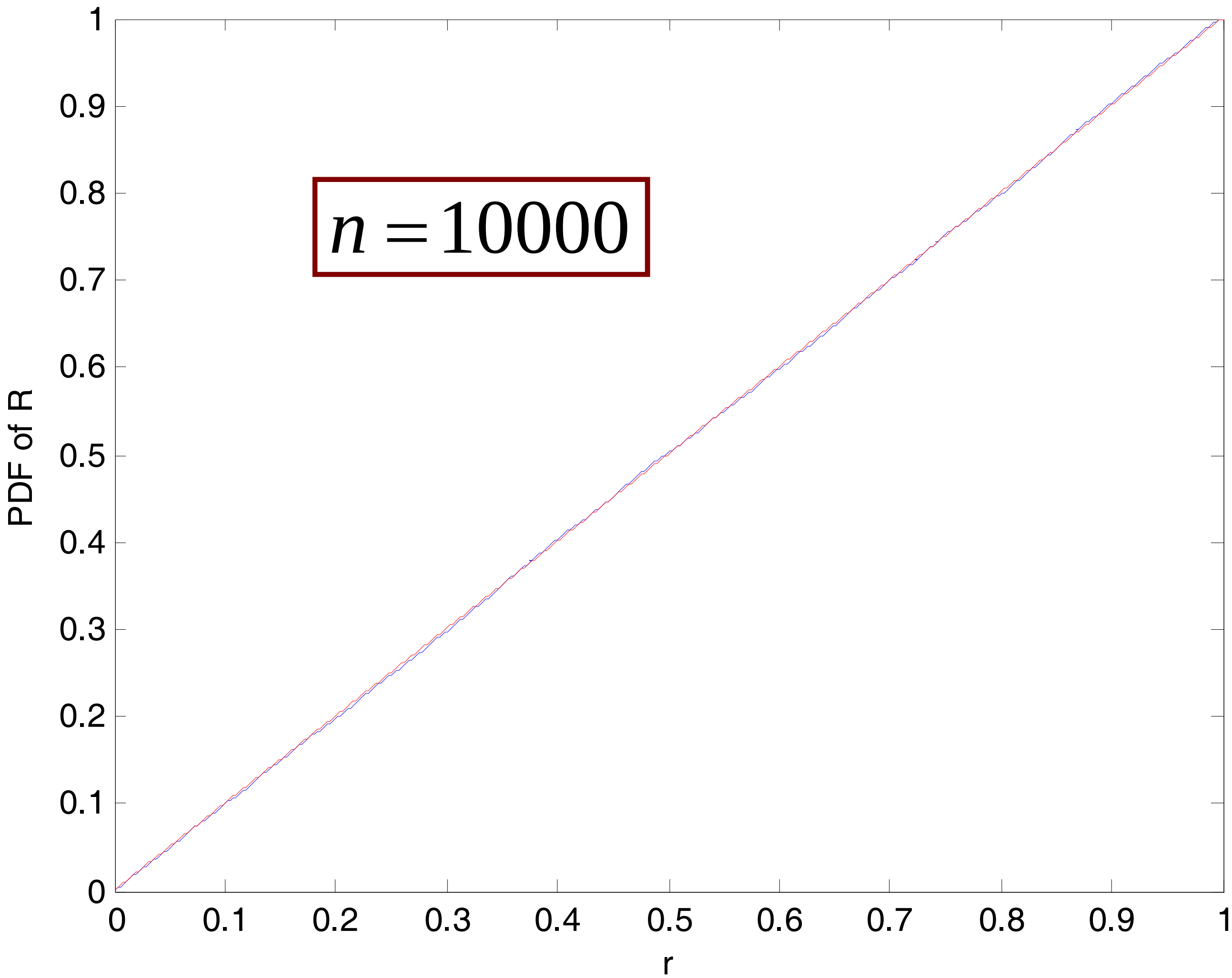




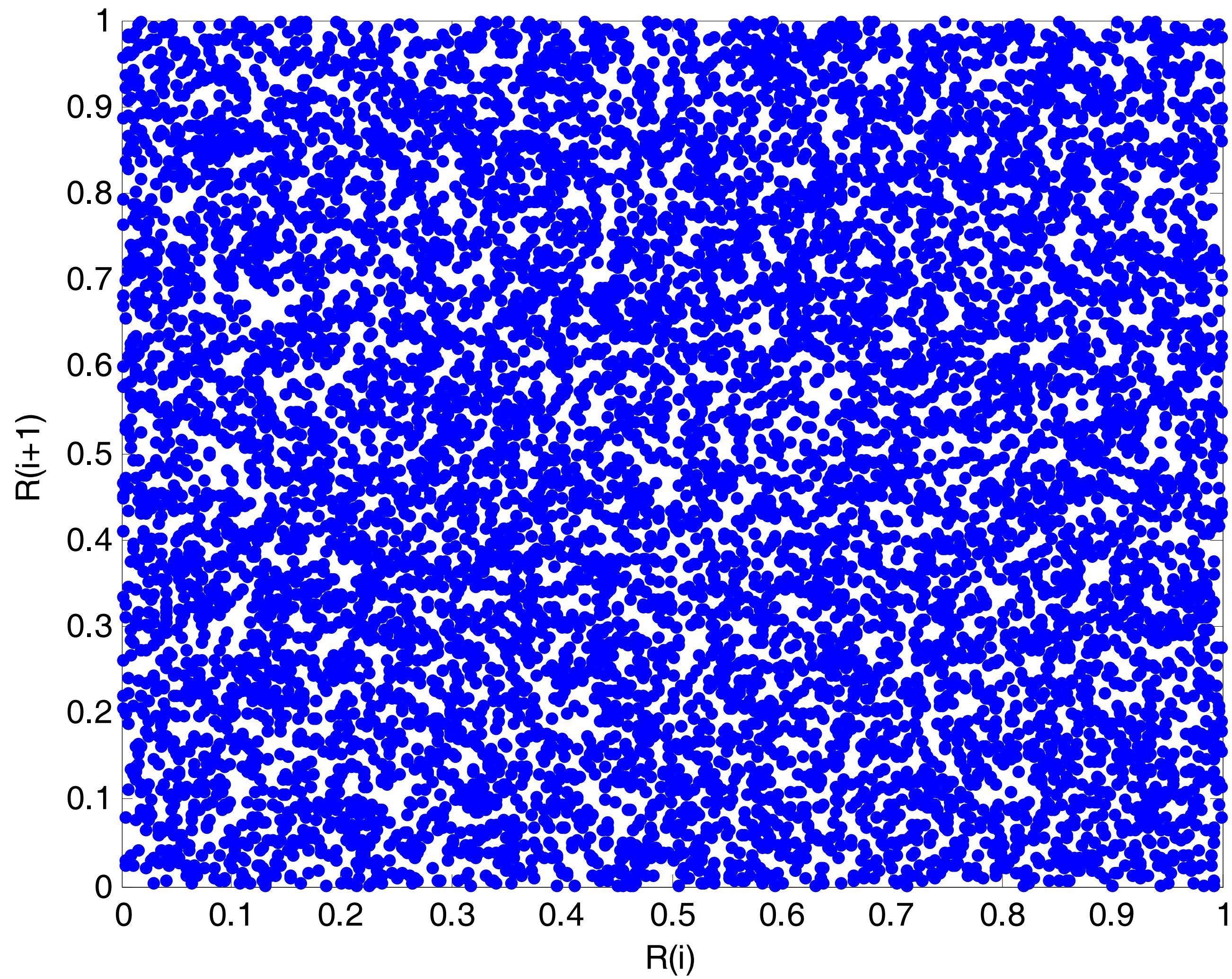
$n = 100$
PDF of R
r
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1



R(i+1)
R(i)
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1

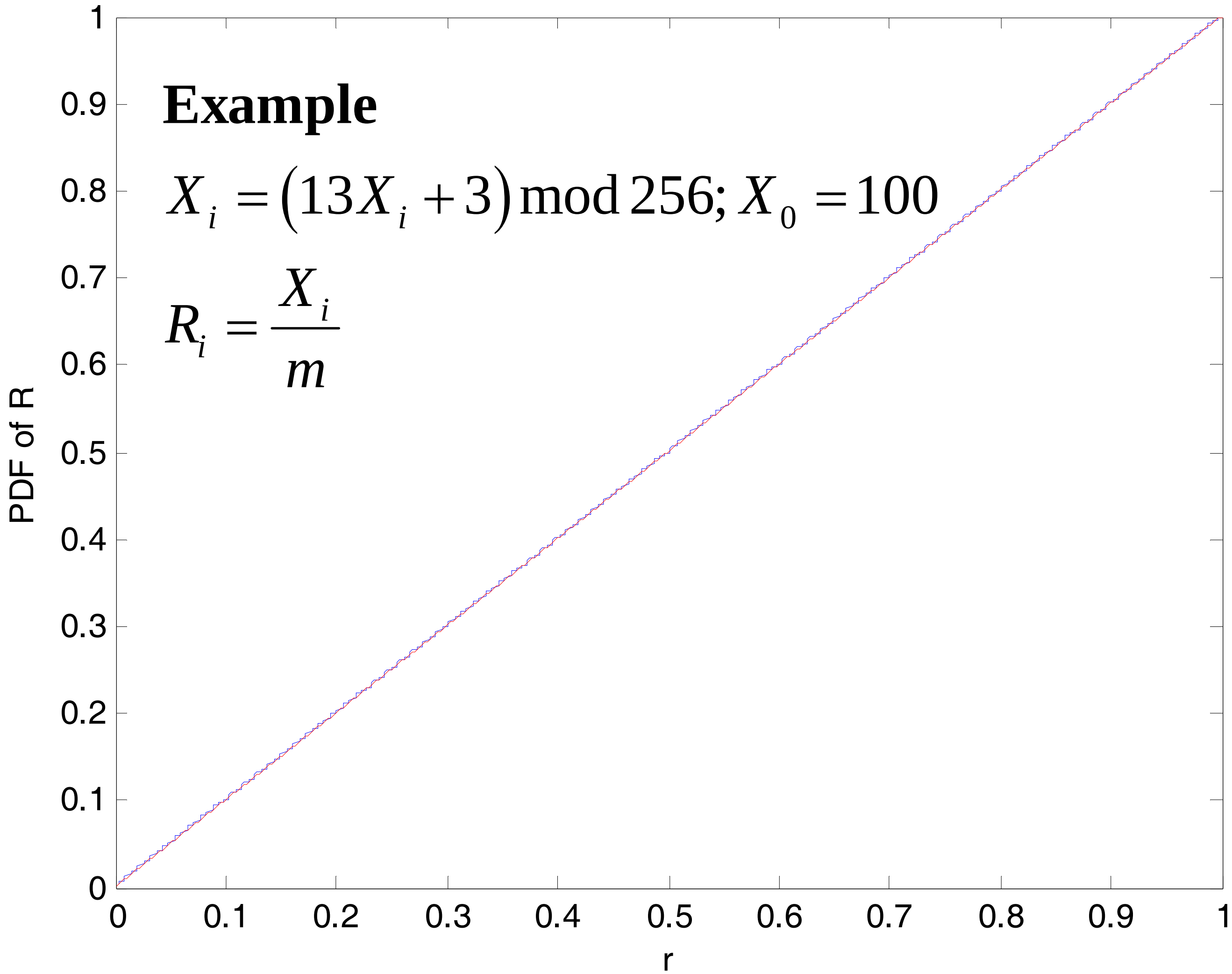


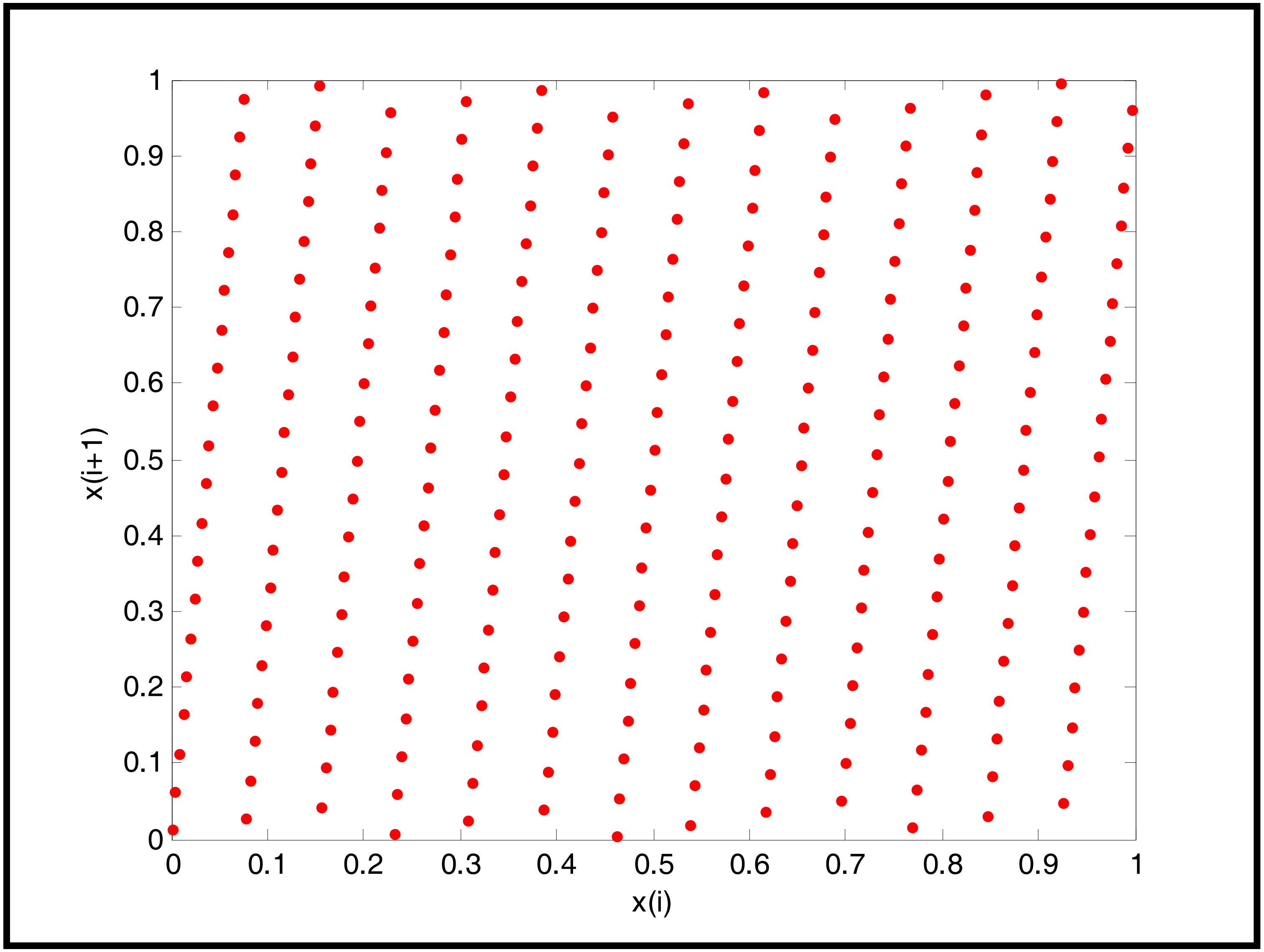
$n = 10000$
PDF of R
r
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1

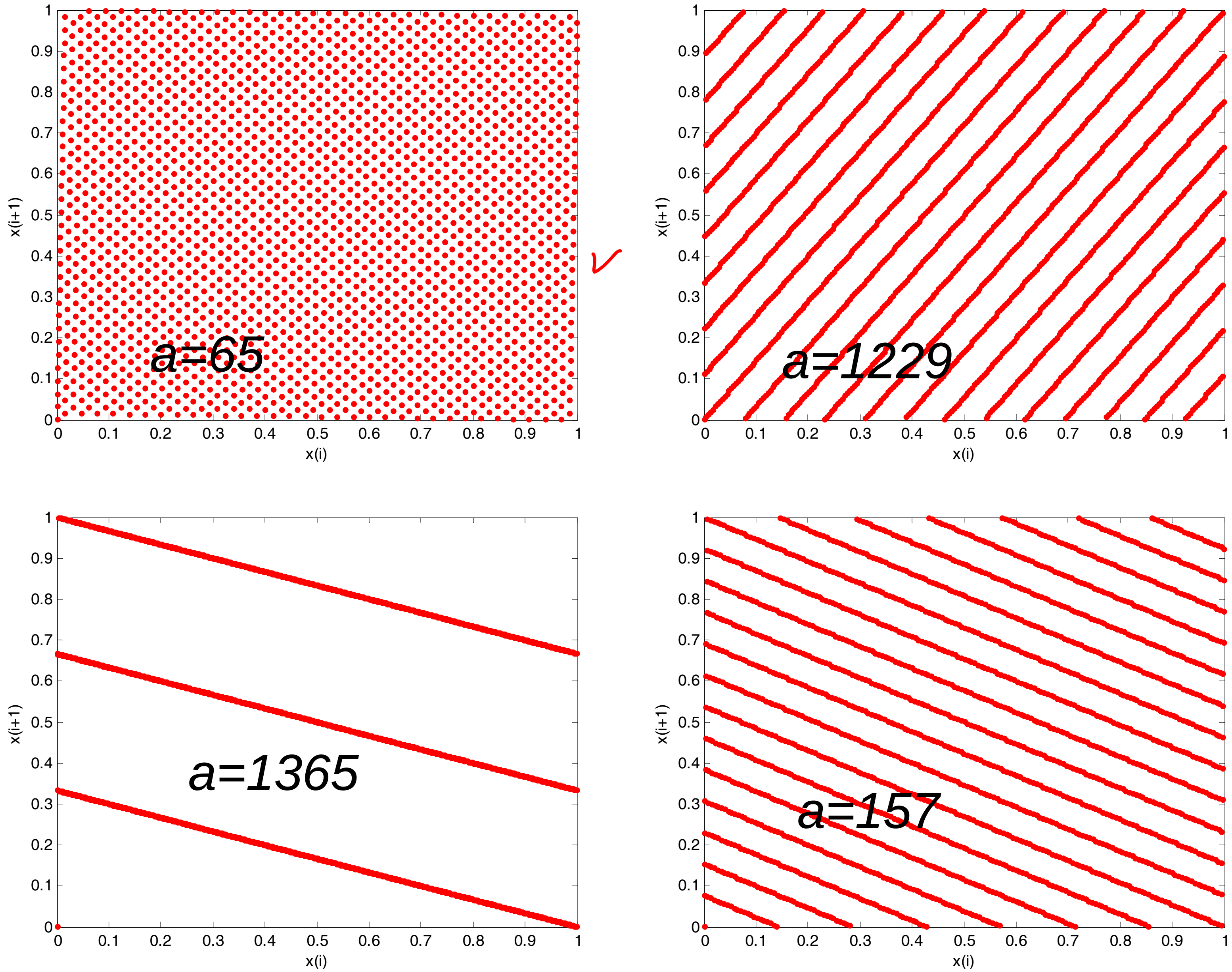


R(i+1)
R(i)
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1

$X_0 = 10$	$X_0 = 11$	$X_0 = 2^{11}$	$X_0 = 2^{12}$
0.0000	0.0000	0.0001	0.0000
0.2198	0.6417	0.3776	0.4121
0.0394	0.5012	0.2258	0.7314
0.2943	0.5644	0.8558	0.8386
0.8787	0.5598	0.8599	0.1702
0.1436	0.0009	0.2307	0.8148
0.1784	0.0543	0.9648	0.6668
0.6822	0.6788	0.2736	0.1618
0.4614	0.0526	0.5027	0.2293
0.0500	0.8025	0.9398	0.7986


Example
$X_i = (13X_i + 3) \bmod 256; X_0 = 100$
$R_i = \frac{X_i}{m}$
PDF of R
r
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1


x(i+1)
x(i)
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1



c=1
m=2048

Caution: hidden orders.

How about order in higher dimensions?
What is acceptable and what is not?
Depends on the application.

Combinations of generators

Motivation: randomness in higher dimensions

$$X_{i+1} = (171 X_i) \bmod 30269$$

$$Y_{i+1} = (172 Y_i) \bmod 30307$$

$$Z_{i+1} = (170 Z_i) \bmod 30323$$

$$R_{i+1} = \left(\frac{X_{i+1}}{30269} + \frac{Y_{i+1}}{30307} + \frac{Z_{i+1}}{30323} \right) \bmod 1$$

Choice of seed

- Store X_0 : helps to reproduce the sequence and avoid storage of random number sequences.
- Use the end of a sequence as a seed in the next sequence
- Start a new sequence with a randomly chosen seed (clock time on the CPU): this approach leads to numbers that are not reproducible.

Selecting pseudorandom number generators

Theoretical studies on random number generators

- Number theory & nonlinear maps

Empirical

- Tests for uniformity
- Tests for independence of pairs and k-tuples

Guidelines

- Period to be at least $2^{27} \approx 10^8$
- k-tuples (k ≤ 10) as uniformly distributed as possible in $[0,1)^k$

Moral

Expert help is needed.

Use algorithms that have been theoretically investigated: naive attempts might be dangerous.

References

•J S Dagpunar, 2007, Simulation and Monte Carlo, Wiley, Chichester.

•R D Ripley, 1987, Stochastic simulaiton, John Wiley, NY

•C P Robert and G Casella, 2004, Monte Carlo statistical methods, Springer, NY.

•J S Liu, 2001, Monte Carlo strategies in scientific computing, Springer, NY.

Generation of Gaussian random numbers

Recall : Box - Muller transformation

Let X and Y be independent and uniformly distributed random variables in 0 to 1. Define

$$U = \left(-2\ln X\right)^{\frac{1}{2}} \cos 2\pi Y$$

$$V = \left(-2\ln X\right)^{\frac{1}{2}} \sin 2\pi Y.$$

Determine jpdf of U and V.

$$u^2 + v^2 = \left(-2\ln x\right) \Rightarrow x = \exp\left[-\frac{u^2+v^2}{2}\right]$$

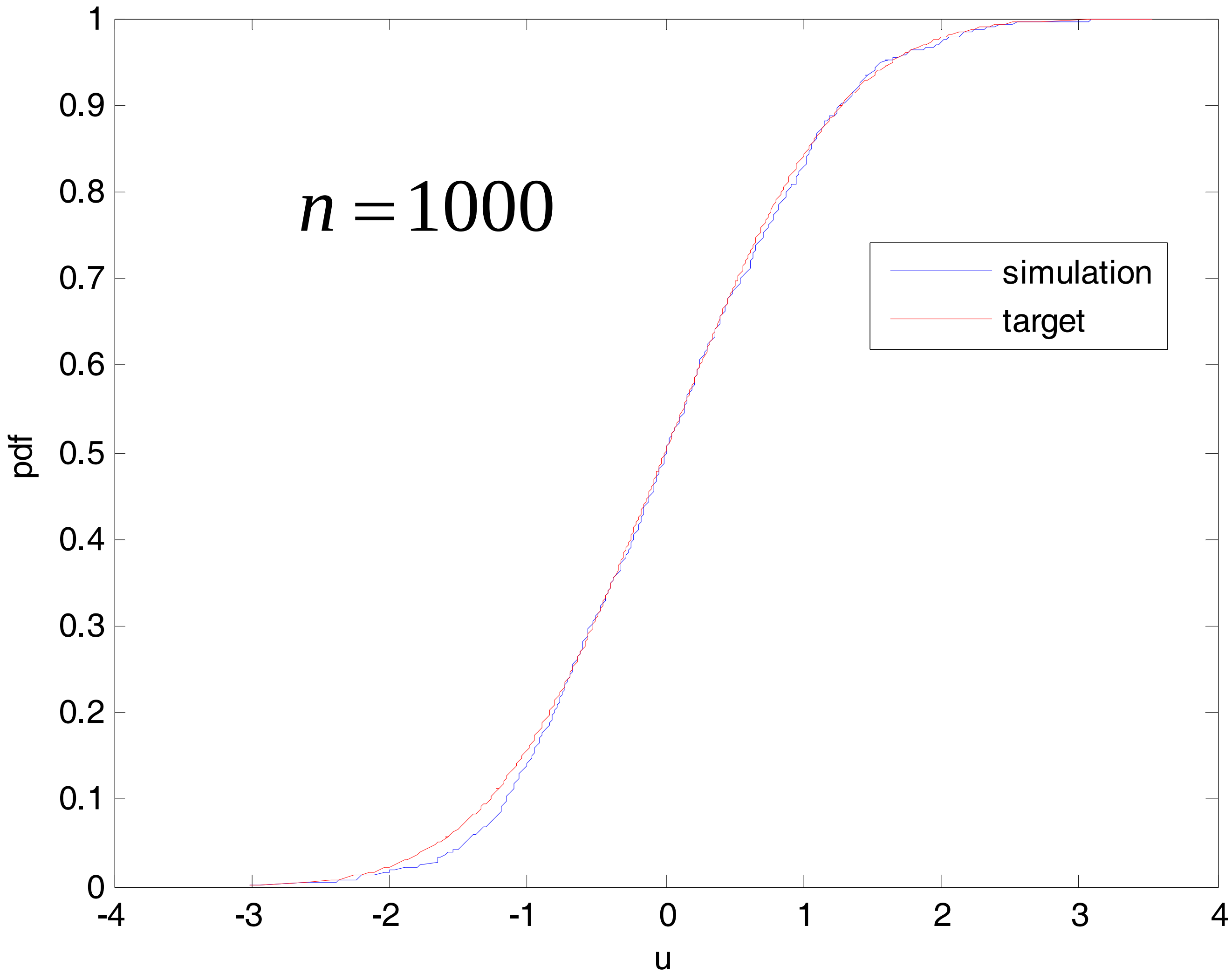
$$\frac{v}{u} = \tan 2\pi y \Rightarrow y = \frac{1}{2\pi}\tan^{-1}\left(\frac{v}{u}\right)$$

$$J^{-1}=\begin{vmatrix} \exp\left[-\dfrac{u^2+v^2}{2}\right](-u) & \exp\left[-\dfrac{u^2+v^2}{2}\right](-v) \\ \dfrac{1}{2\pi}\dfrac{1}{1+\left(\dfrac{v}{u}\right)^2}\left(-\dfrac{v}{u^2}\right) & \dfrac{1}{2\pi}\dfrac{1}{1+\left(\dfrac{v}{u}\right)^2}\left(\dfrac{1}{u}\right) \end{vmatrix}$$

$$=-\frac{1}{2\pi}\exp\left[-\frac{u^2+v^2}{2}\right]$$

$$\Rightarrow p_{UV}(u,v)=\frac{1}{2\pi}\exp\left[-\frac{u^2+v^2}{2}\right];-\infty<u,v<\infty$$

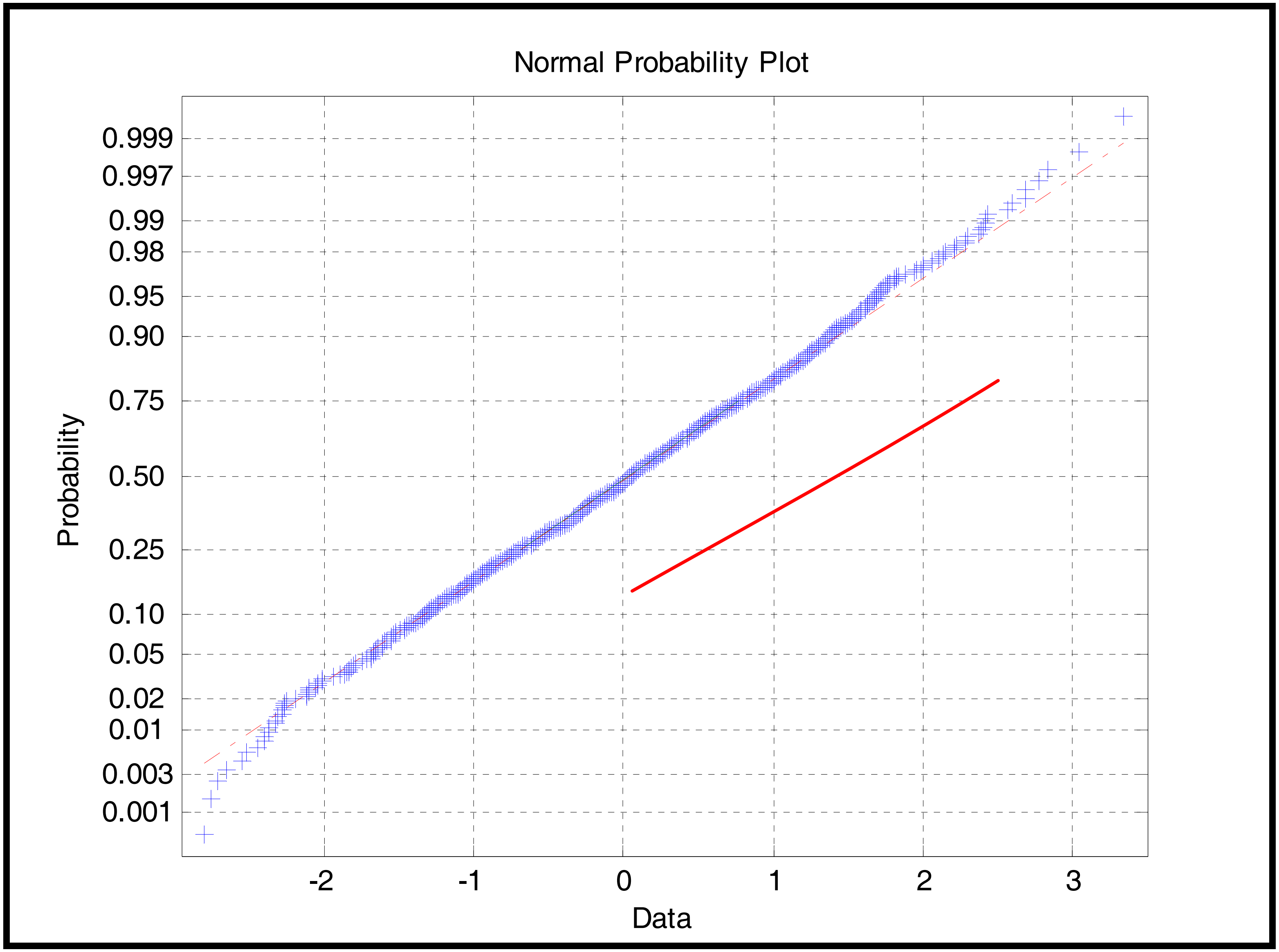
$$\Rightarrow U\perp V\ \&\ U\sim N(0,1),V\sim N(0,1)$$

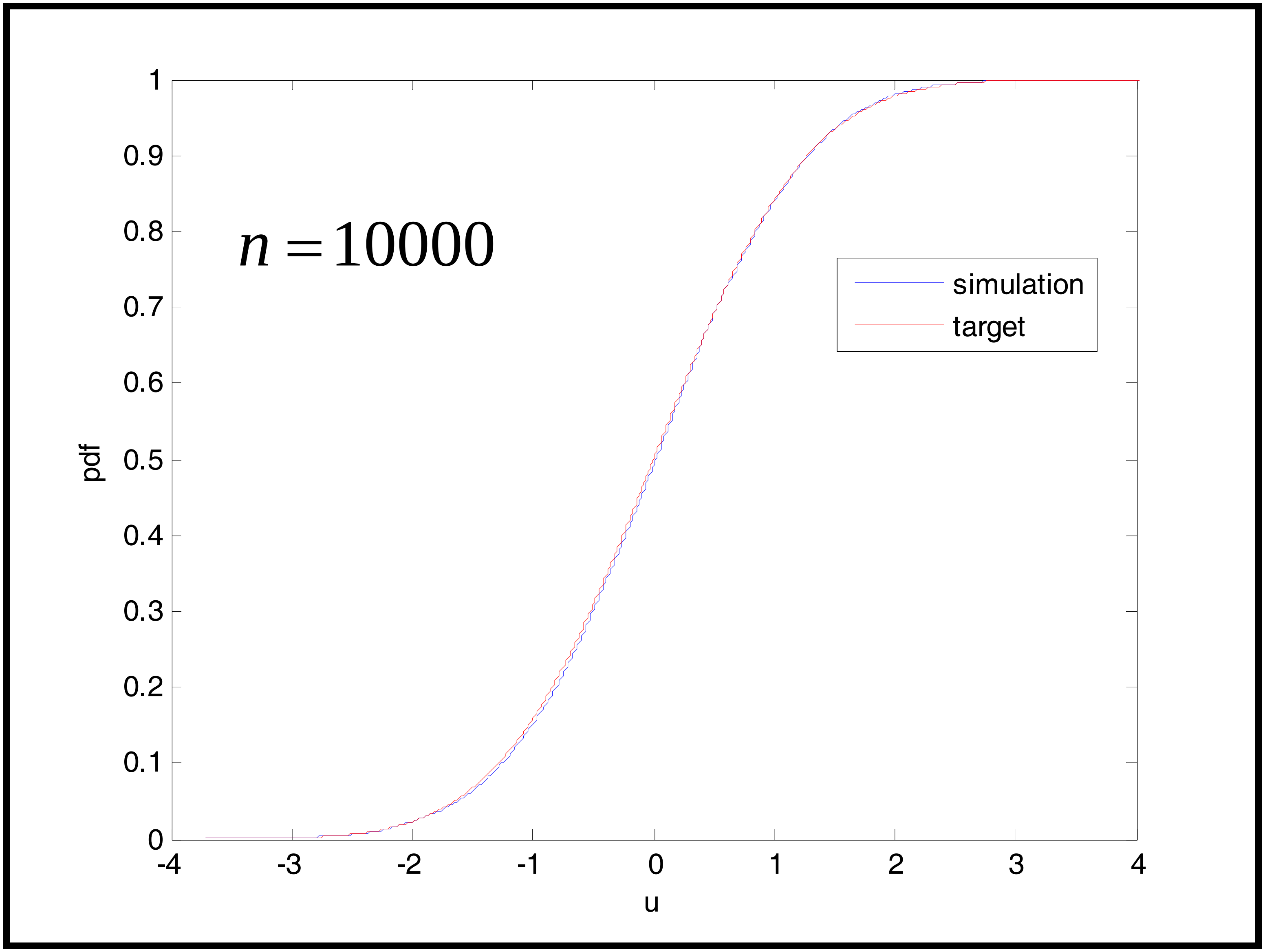


$n = 1000$
simulation
target
pdf
u
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
-4
-3
-2
-1
0
1
2
3
4

Results of Kolmogorov Smirnov test: $k = 0.0390;\ c = 0.0428; (\alpha = 0.05)$

Accept the hypothesis that sample is drawn from a population of $N(0,1)$

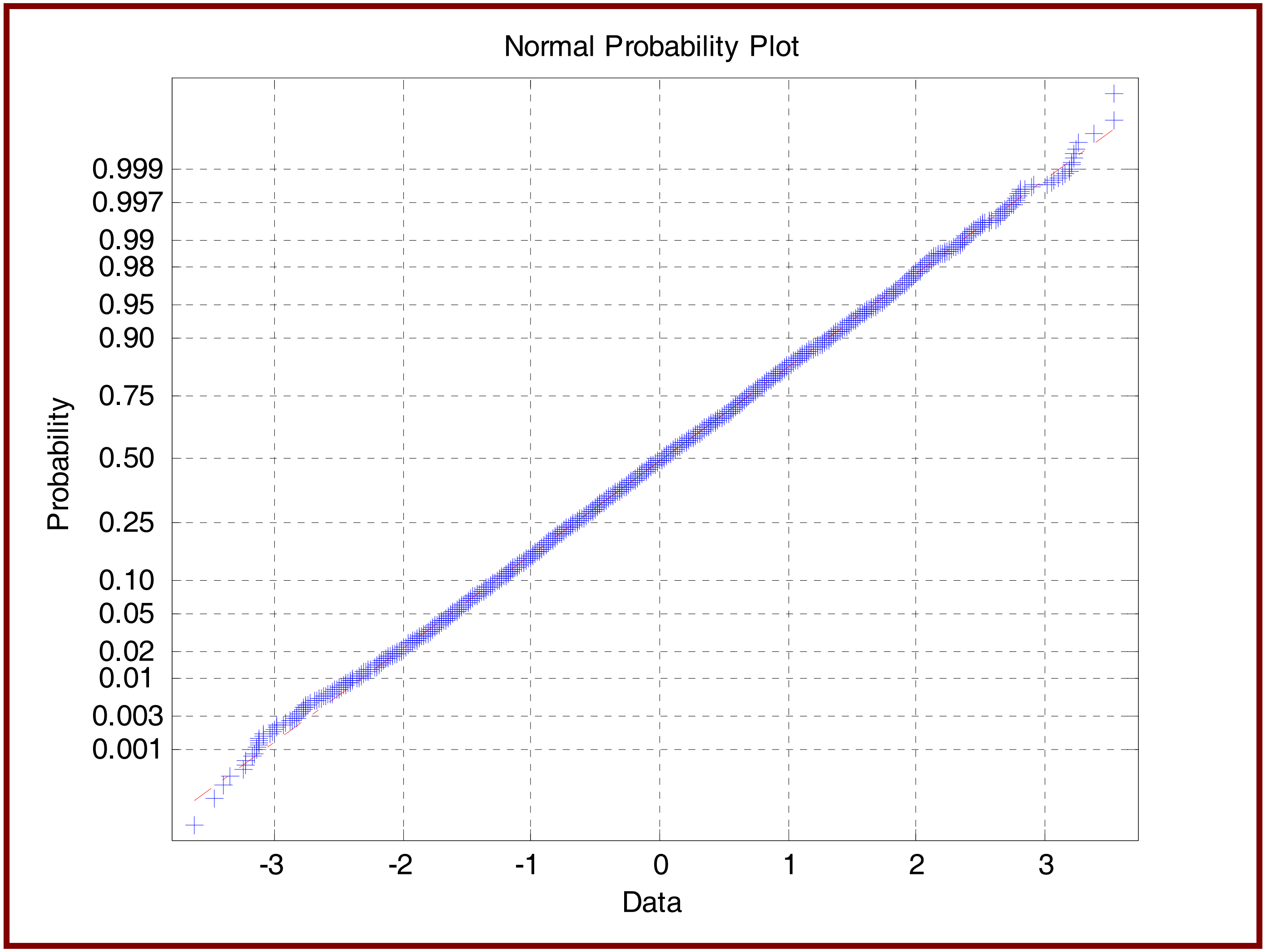


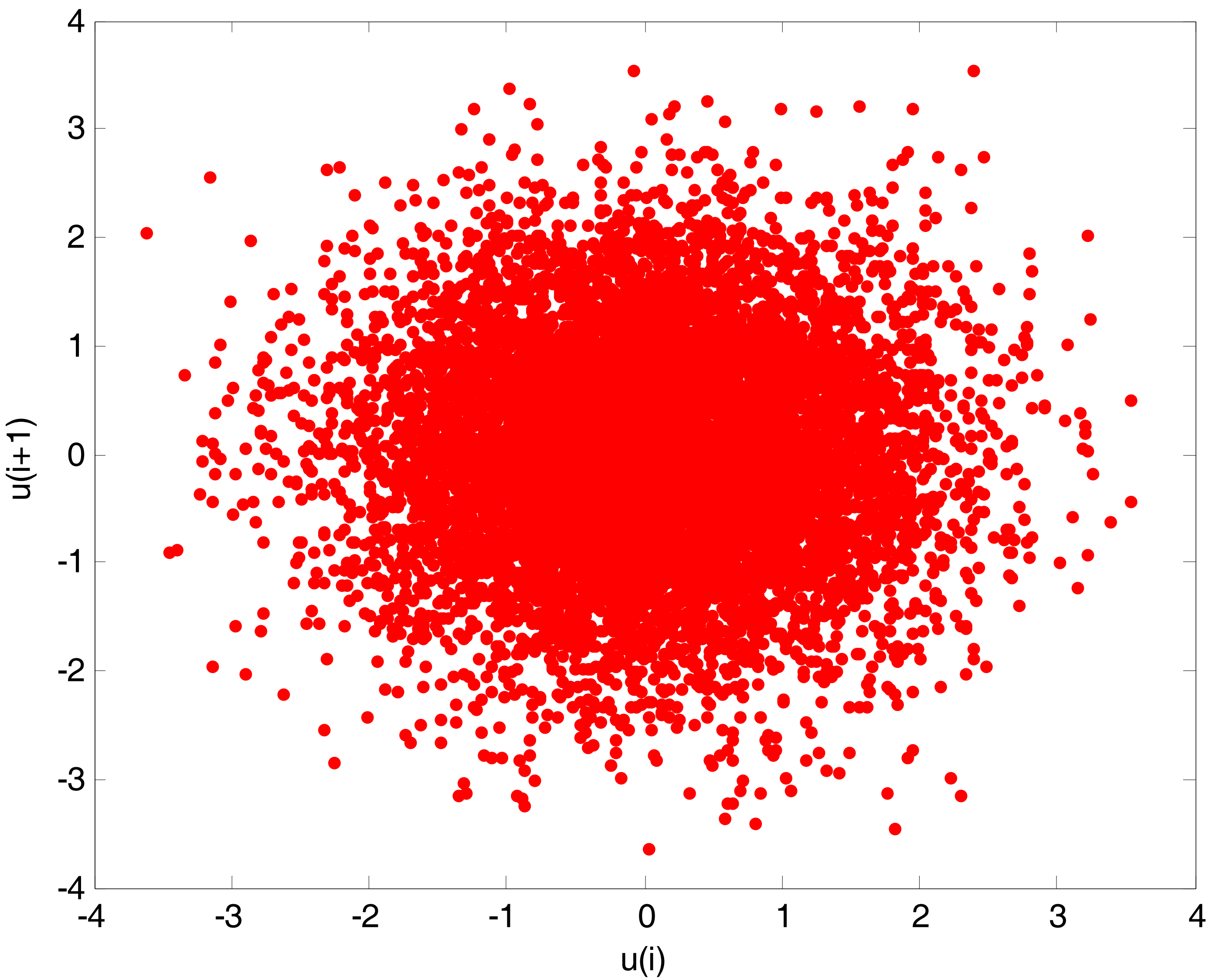


$n = 10000$
simulation
target
pdf
u
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
-4
-3
-2
-1
0
1
2
3
4

Results of Kolmogorov Smirnov test: $k = 0.0122$; $c = 0.0136$; $(\alpha = 0.05)$

Accept the hypothesis that sample is drawn from a population of $N(0,1)$





4
3
2
1
0
-1
-2
-3
-4
u(i+1)
-4
-3
-2
-1
0
1
2
3
4
u(i)

$$X(i) = \left[aX(i-1) + c\right] \bmod m$$

$$Y(i) = \left[aY(i-1) + c\right] \bmod m$$

$$m = 2048$$

$$a = 1229$$

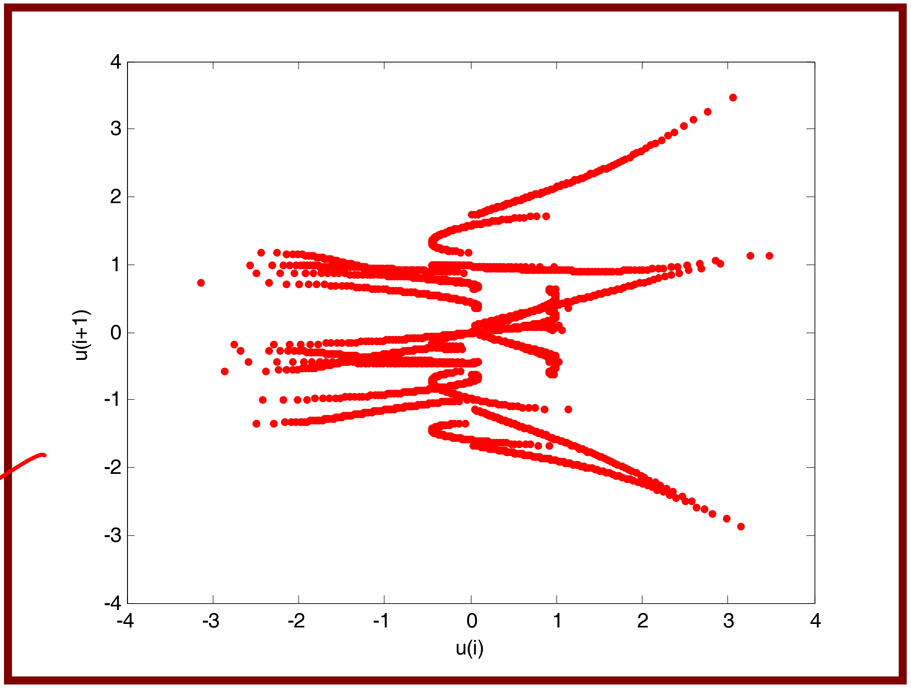
$$c = 1$$

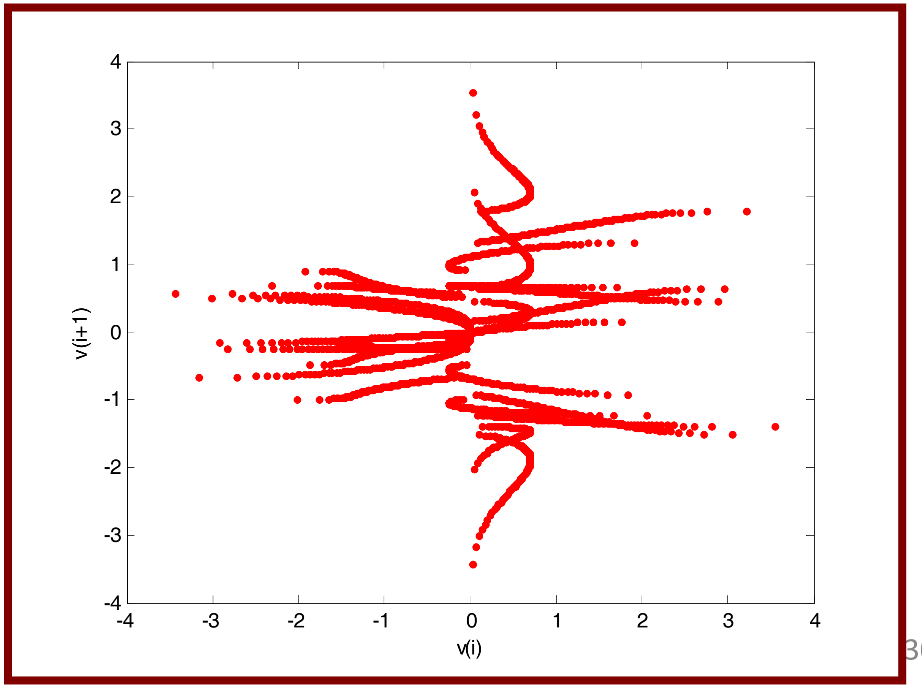
$$X(1) = 0$$

$$Y(1) = 1$$

$$U(i) = \left[-2 \ln X(i)\right]^{\frac{1}{2}} \cos 2\pi Y(i)$$

$$V(i) = \left[-2 \ln X(i)\right]^{\frac{1}{2}} \sin 2\pi Y(i).$$





General procedure for simulation of samples of a random variable X with pdf $p_X(x)$.

Starting point: generator for standard normal rv.

Let $Z \sim N(0,1)$.

Consider

$$P_X(X) = \Phi(Z) \Rightarrow X = P_X^{-1}\left[\Phi(Z)\right]$$

$$\Rightarrow p_X(x)\frac{dx}{dz} = \phi(z)$$

$$\Rightarrow p_X(x) = \frac{\phi(z)}{\left|\frac{dx}{dz}\right|} = p_X(x).$$

Steps

(1) Simulate samples of Z: $\left\{\mathrm{Z_i}\right\}_{i=1}^{N}$.

(2) Simulate samples of X using $X_i = P_X^{-1}\left[\Phi(Z_i)\right]$

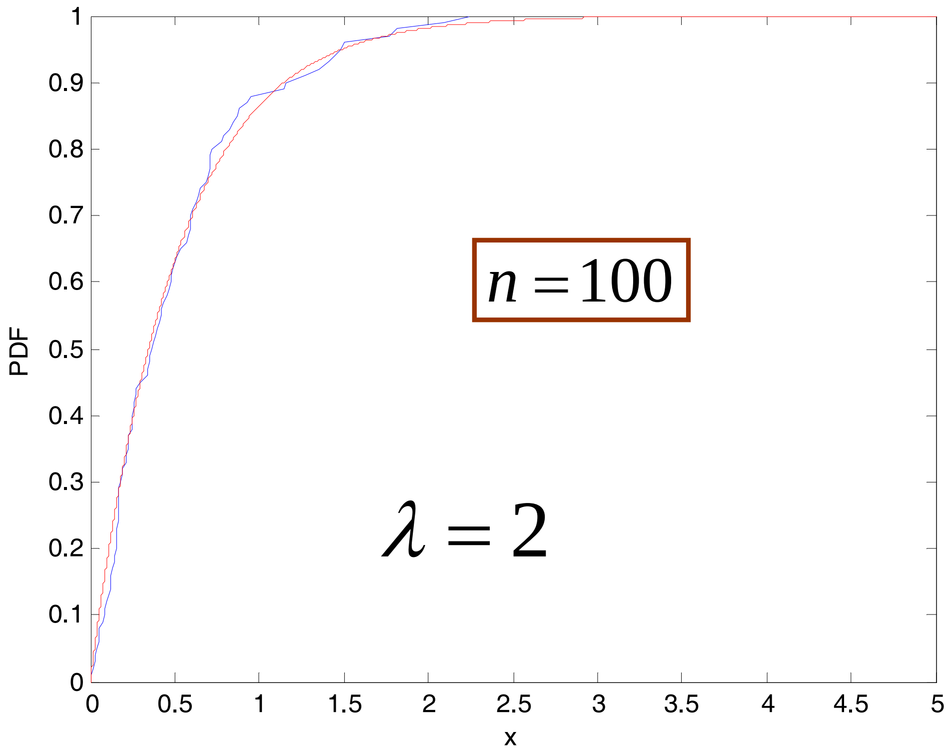
Example : Exponential random variable

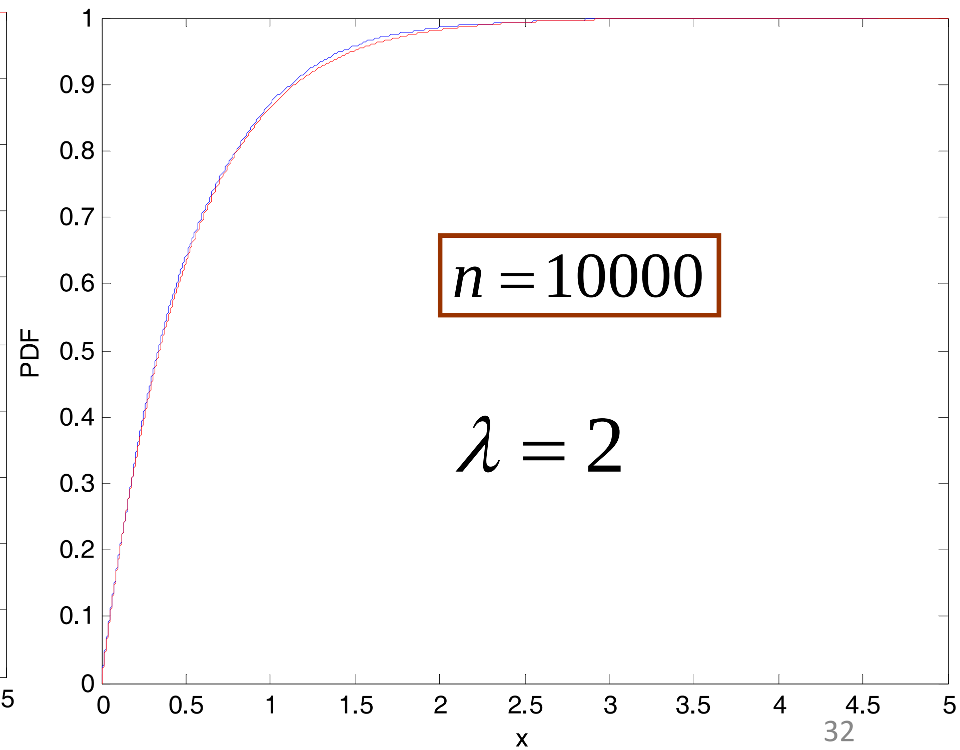
Let $P_X\left(x\right)=1-\exp\left(-\lambda x\right)$;

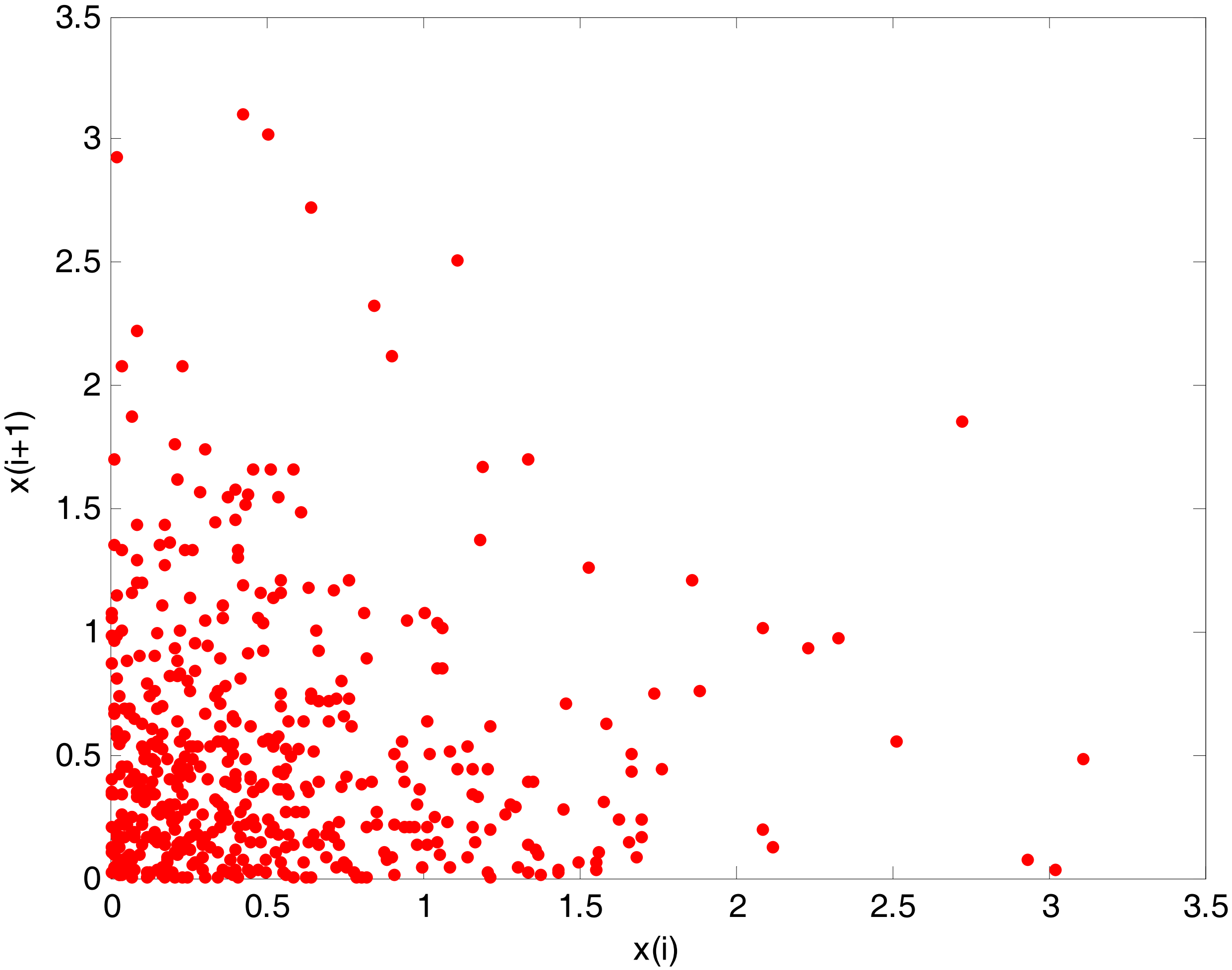
$P_X\left(X\right)=\Phi(Z)$

$\Rightarrow 1-\exp\left(-\lambda X\right)=\Phi\left(Z\right)$

$$X=-\frac{1}{\lambda}\log\left[1-\Phi\left(Z\right)\right]$$







3.5
3
2.5
2
1.5
1
0.5
0
x(i+1)
0
0.5
1
1.5
2
2.5
3
3.5
x(i)

Example : Rayleigh random variable

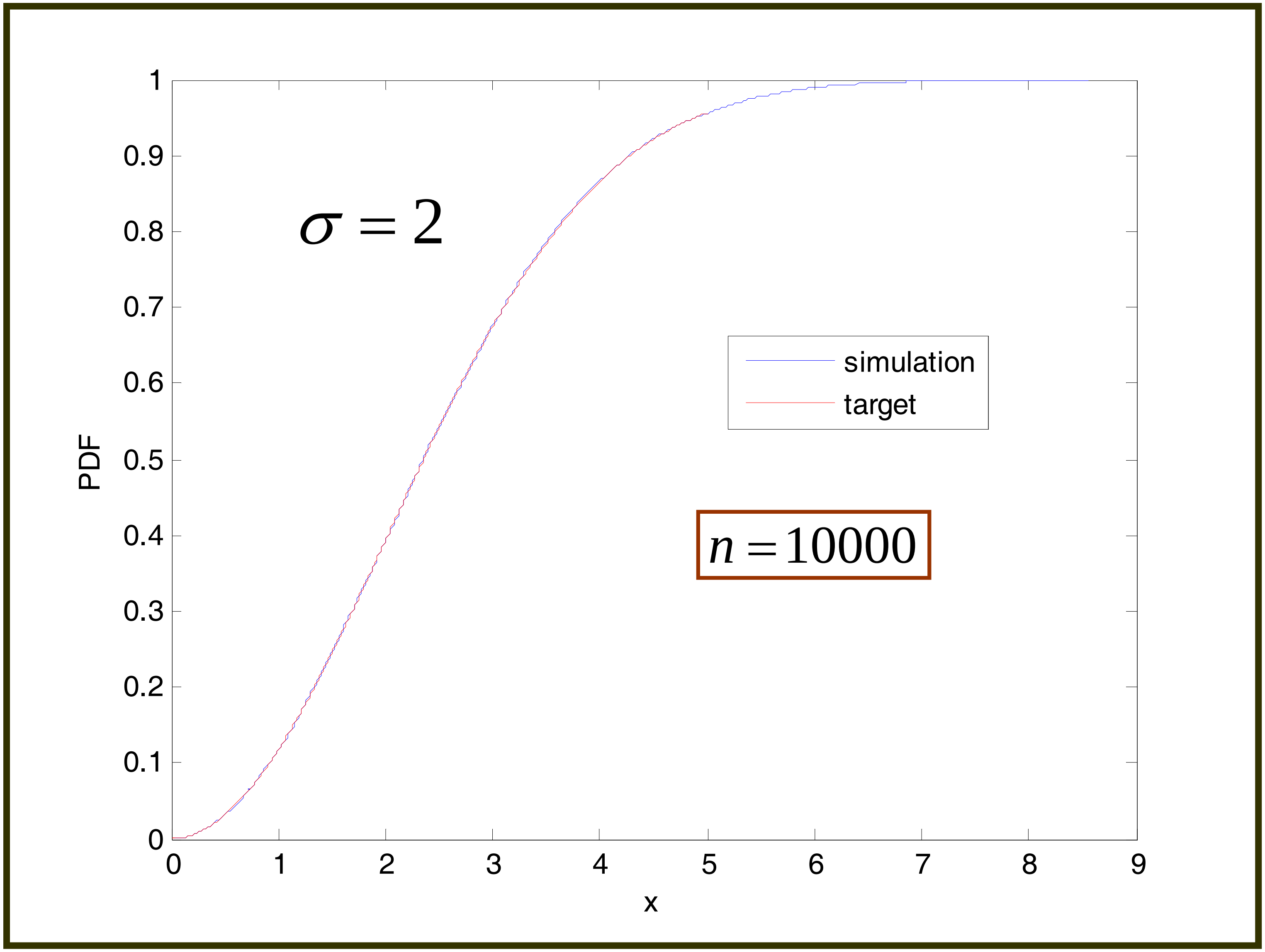
$$\text{Let } p_X\left(x\right)=\frac{x}{\sigma^2}\exp\left(-\frac{x^2}{2\sigma^2}\right);x\geq 0$$

$$\mathrm{P}_X\left(x\right)=\int_0^x\frac{u}{\sigma^2}\exp\left(-\frac{u^2}{2\sigma^2}\right)du=\int_0^{\frac{x^2}{2\sigma^2}}\exp\left(-t\right)dt=1-\exp\left(-\frac{x^2}{2\sigma^2}\right)$$

$$P_X(X)=\Phi(Z)$$

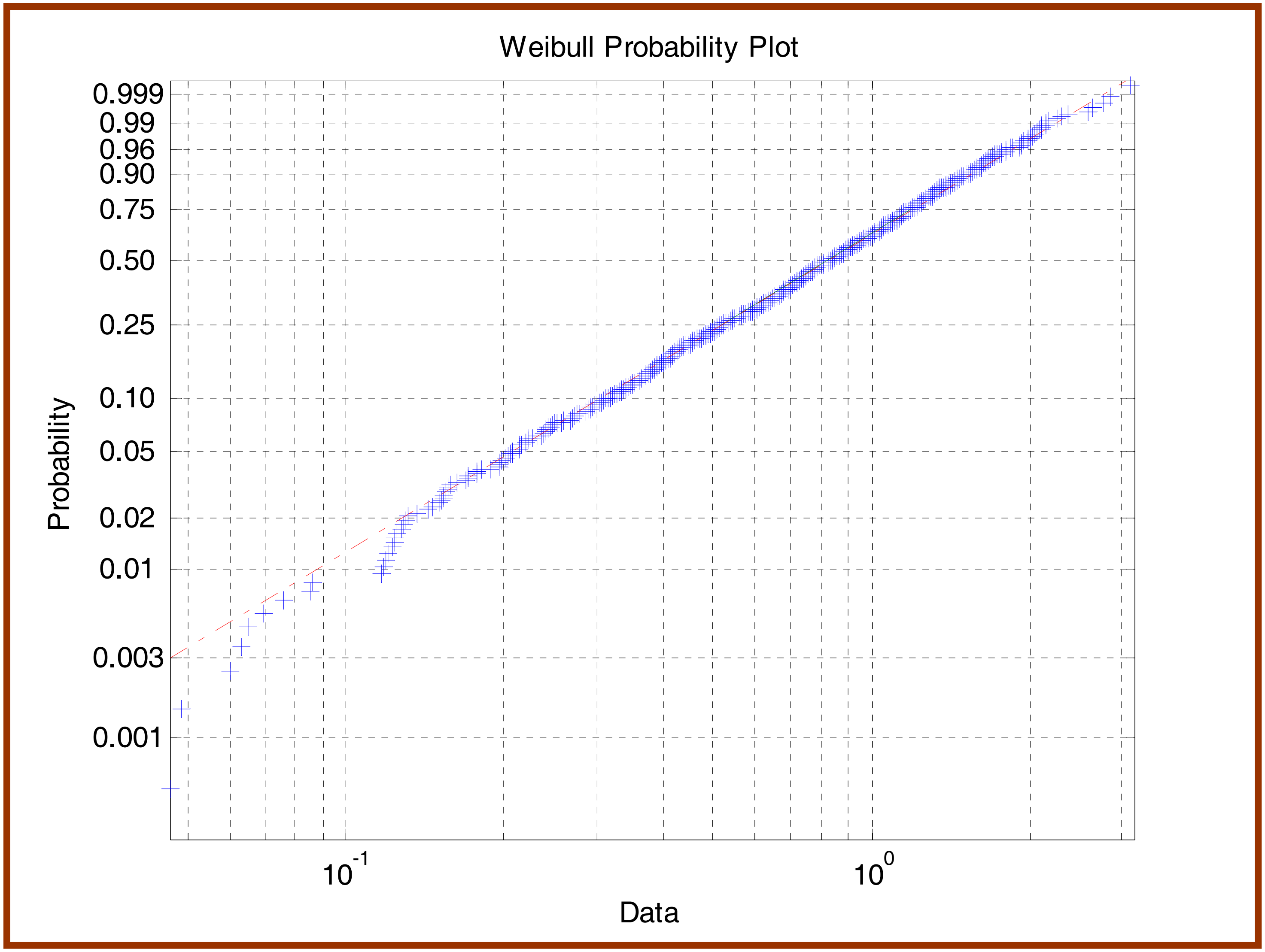
$$\Rightarrow 1-\exp\left(-\frac{X^2}{2\sigma^2}\right)=\Phi\left(Z\right)$$

$$X=\left\{-2\sigma^2\log\left[1-\Phi\left(Z\right)\right]\right\}^{\frac{1}{2}}$$



$\sigma = 2$
simulation
target
$n = 10000$
PDF
x
0
0.1
0.2
0.3
0.4
0.5
0.6
0.7
0.8
0.9
1
0
1
2
3
4
5
6
7
8
9

1000 Weibull random numbers on Weibull probability paper



Question : How to proceed if $P_X(x)$ is not (or not easily) invertible?

Example :

$$p_R(r) = \frac{r}{\sigma_a \sigma_b \sqrt{\left(1-r_{ab}^2\right)}} \exp\left[-r^2\left(\frac{\sigma_a^2+\sigma_b^2}{4\sigma_a^2\sigma_b^2\left(1-r_{ab}^2\right)}\right)\right] I_0\left[r^2\left(\frac{r_{ab}^2+\left\{\frac{\sigma_a^2-\sigma_b^2}{2\sigma_a\sigma_b}\right\}^2}{2\sigma_a\sigma_b\left(1-r_{ab}^2\right)}\right)\right]$$

$0 < r < \infty.$

$I_0(\bullet) =$ Bessel's function of the first kind

$$p_\Phi(\phi) = \frac{\sqrt{\left(1-r_{ab}^2\right)}}{2\pi\sigma_a\sigma_b\left[\frac{\cos^2\phi}{\sigma_b^2}+\frac{\sin^2\phi}{\sigma_a^2}-\frac{r_{ab}\sin\phi\cos\phi}{\sigma_a\sigma_b}\right]}; 0 < \phi < 2\pi$$

Accept - Reject Methods

Let X be a random variable with pdf $p_X(x)$.

Let it be required to simulate samples of X.

Define a random variable U such that it is distributed uniformly in 0 to $p_X(x)$.

$\Rightarrow U$ and X are mutually dependent with

$$p_{UX}(u,x) = 1 \text{ for } 0 < u \leq p_X(x) \ \& -\infty < x < \infty$$

$$= 0 \text{ otherwise.}$$

$$\Rightarrow p_X(x) = \int p_{UX}(u,x)\,du = \int_0^{p_X(x)} du$$

$$p_X(x) = \int p_{UX}(u,x)\,du = \int_0^{p_X(x)} du$$

Thus $p_X(x)$ is obtained as a marginal pdf of $p_{UX}(u,x)$.

Can we simulate samples of (X,U) without inverting $P_X(x)$?

$U =$ auxiliary variable.

Generate (X,U) by generating uniform random variables on the constrained set $\{(x,u): 0 < u < p_X(x)\}$.

Simulating $X \sim p_X(x)$ is equivalent to simulating $(X,U) \sim \mathbf{U}\{(x,u): 0 < u < p_X(x)\}$

$$p_X(x) = \int p_{UX}(u,x)\,du = \int_0^{p_X(x)} du$$

Thus $p_X(x)$ is obtained as a marginal pdf of $p_{UX}(u,x)$.

Can we simulate samples of (X,U) without inverting $P_X(x)$?

$U =$ auxiliary variable.

Generate (X,U) by generating uniform random variables on the constrained set $\{(x,u): 0 < u < p_X(x)\}$.

Simulating $X \sim p_X(x)$ is equivalent to simulating

$(X,U) \sim \mathbf{U}\{(x,u): 0 < u < p_X(x)\}$

Let X be a random variable with pdf $p_X(x), x \in I$.

Represent $p_X(x) = Mg(x)h(x)$

where $M \geq 1,\ 0 < g(x) \leq 1 \,\&\, h(x)$ is a valid pdf.

Let $U \sim \mathbf{U}[0,1] \,\&\, Y \sim h(y)$.

$\Rightarrow p_Y\left[x \mid U \leq g(Y)\right] = p_X(x)$

$$\bullet\, p_Y\left[x \mid U \le g(Y)\right] = \frac{P\left[U \le g(Y) \mid Y = x\right] h(x)}{P\left[U \le g(Y)\right]}$$

$$\left[\text{Bayes' theorem}\right]$$

$$\bullet\, P\left[U \le g(Y) \mid Y = x\right] = P\left[U \le g(x)\right] = g(x)$$

$$\bullet\, P\left[U \le g(Y)\right] = \int P\left[U \le g(Y) \mid Y = x\right] h(x)\, dx$$

$$= \int g(x) h(x)\, dx = \int \frac{p_X(x)}{M}\, dx = \frac{1}{M}$$

$$\bullet\, p_Y\left[x \mid U \le g(Y)\right] = \frac{g(x) h(x)}{1/M} = p_X(x)$$

Algorithm

1. Generate $X \sim g\left(x\right) \& U \sim \mathbf{U}\left[0,1\right]$

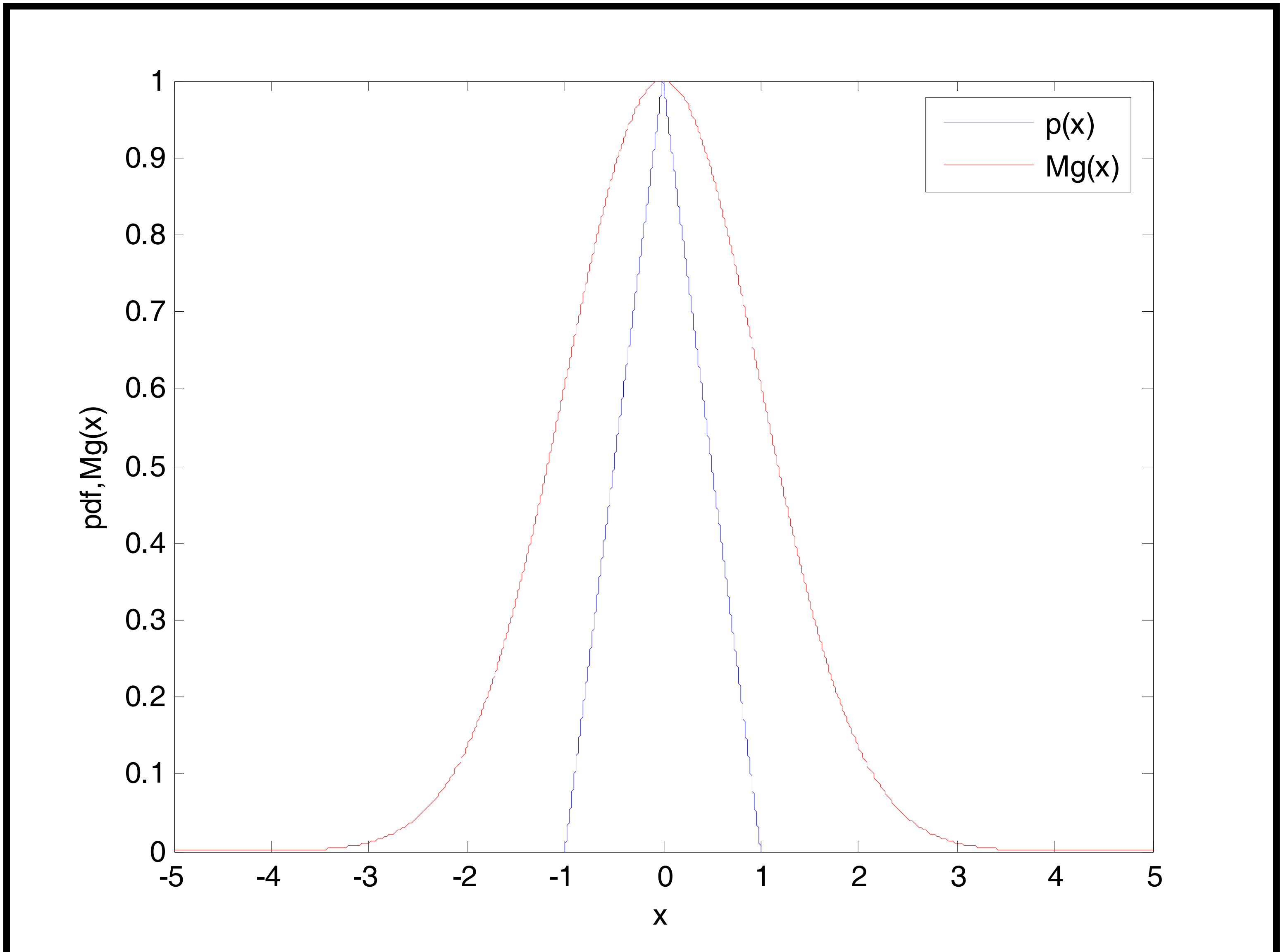
2. Accept $Y = X$ if $U \leq \dfrac{p_X\left(x\right)}{Mg\left(x\right)}$

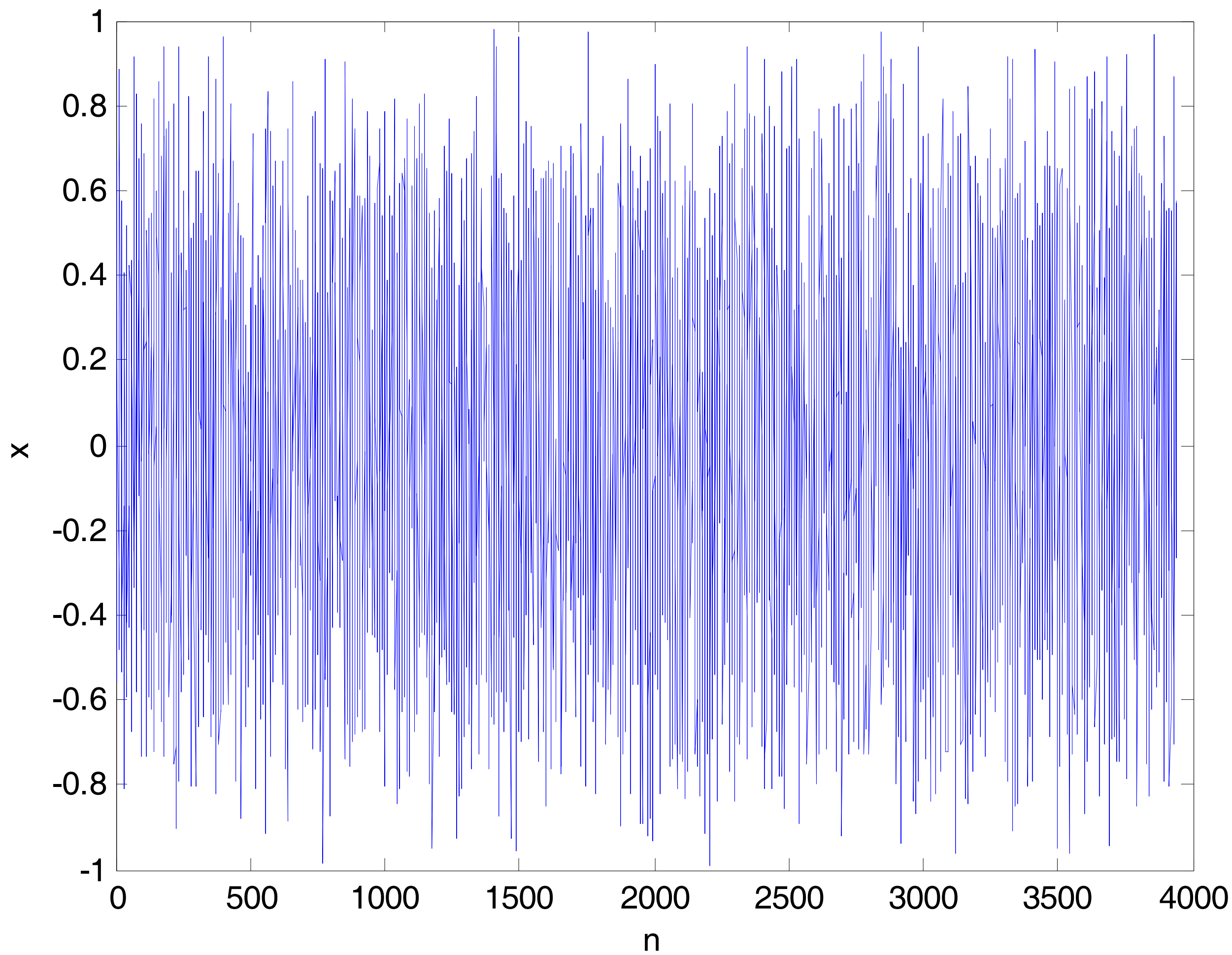
3. Return to 1, otherwise.

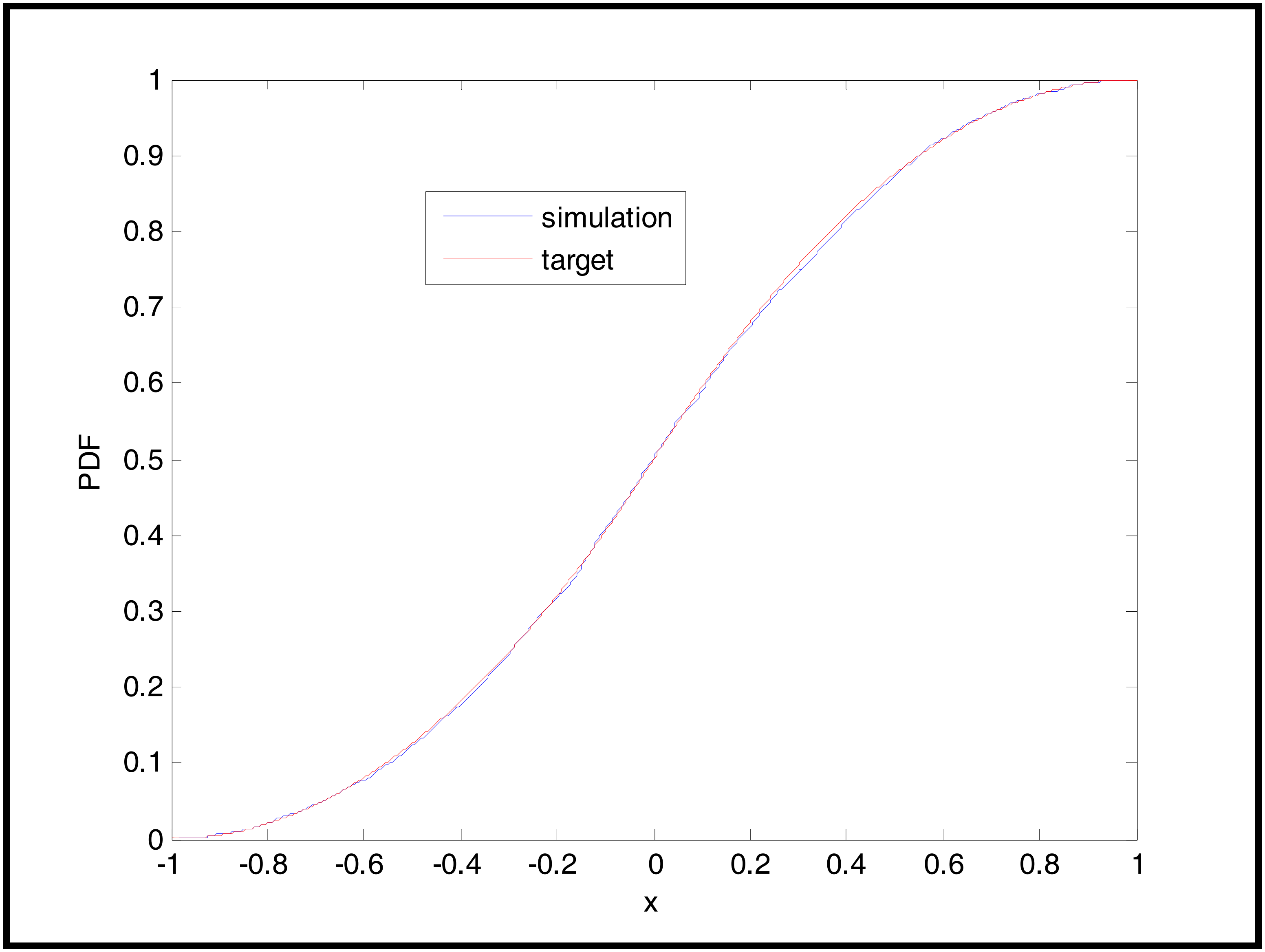
Example

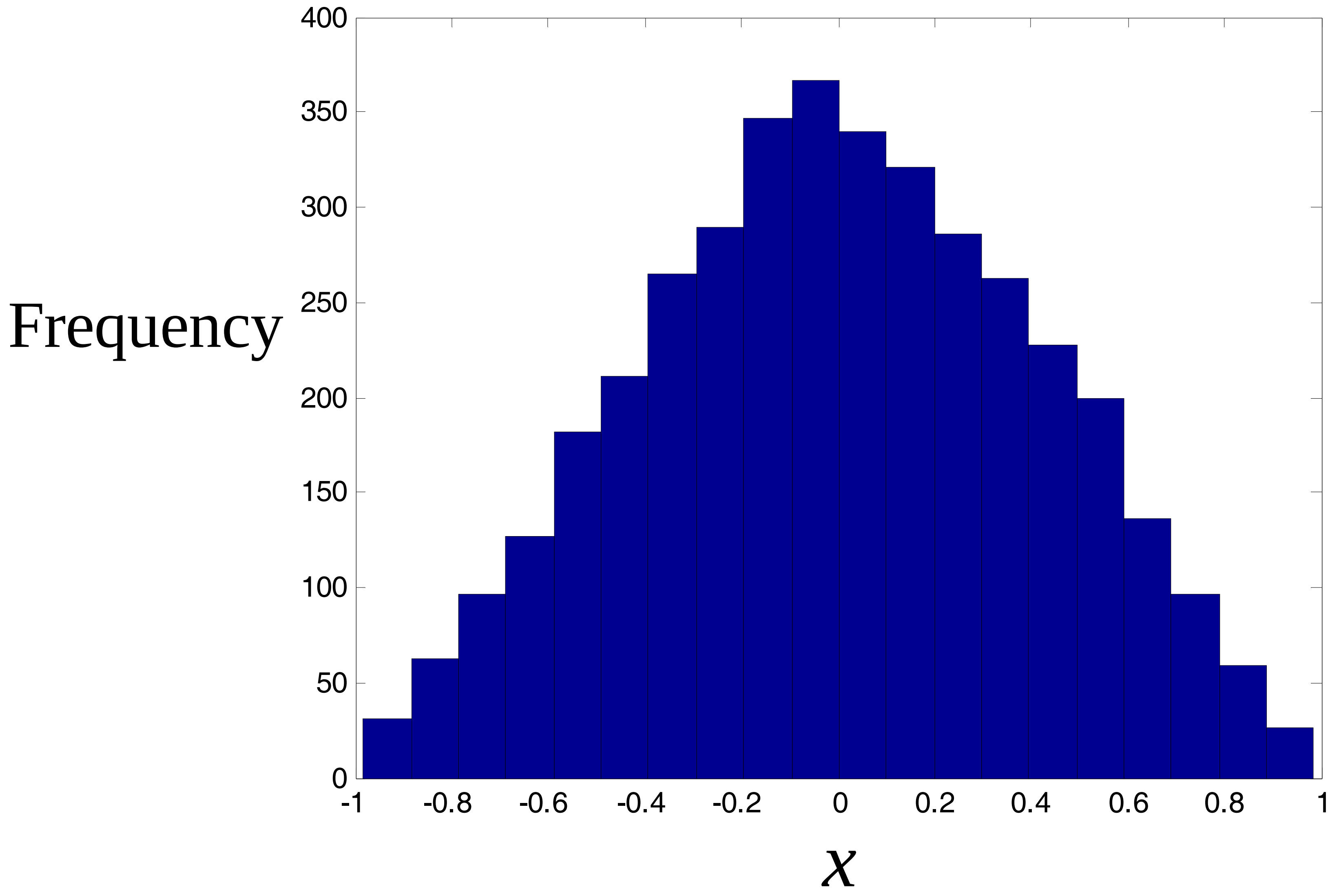
$$p_X(x) = x + 1; -1 < x < 0$$
$$= -x + 1; 0 < x < 1$$

$$P_X(x) = x + \frac{x^2}{2} + \frac{1}{2}; -1 < x < 0$$
$$= x - \frac{x^2}{2} + \frac{1}{2}; 0 < x < 1$$


p(x)
Mg(x)
pdf,Mg(x)
x
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
-5
-4
-3
-2
-1
0
1
2
3
4
5


1
0.8
0.6
0.4
0.2
0
-0.2
-0.4
-0.6
-0.8
-1
x
0
500
1000
1500
2000
2500
3000
3500
4000
n


1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
PDF
simulation
target
-1
-0.8
-0.6
-0.4
-0.2
0
0.2
0.4
0.6
0.8
1
x



Frequency
400
350
300
250
200
150
100
50
0
-1
-0.8
-0.6
-0.4
-0.2
0
0.2
0.4
0.6
0.8
1
x

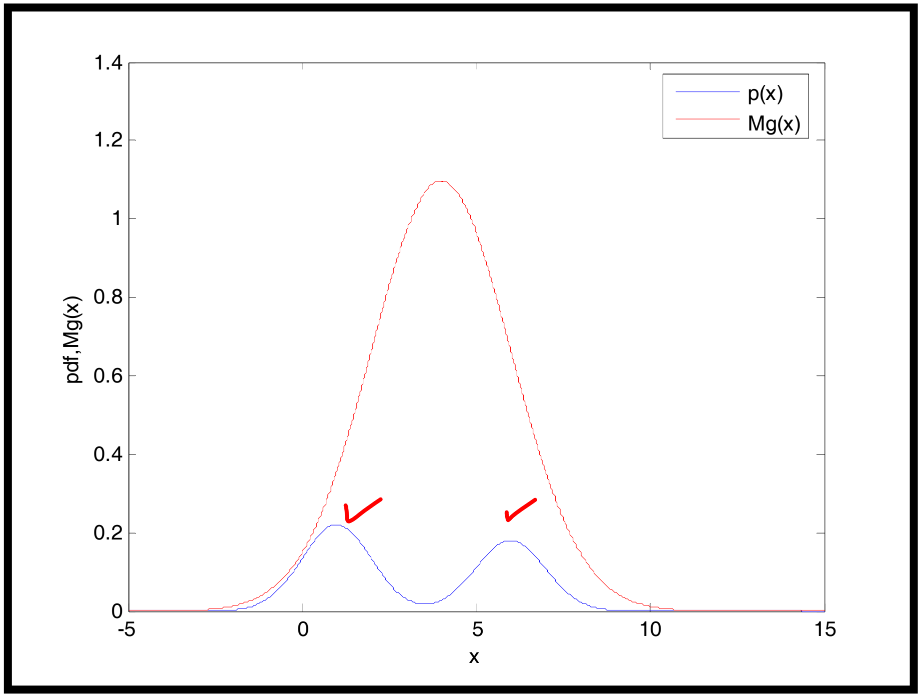
$p_X(x) = a_1 N(x, 2, 1) + a_2 N(x, 6, 1) \quad a_1 + a_2 = 1$

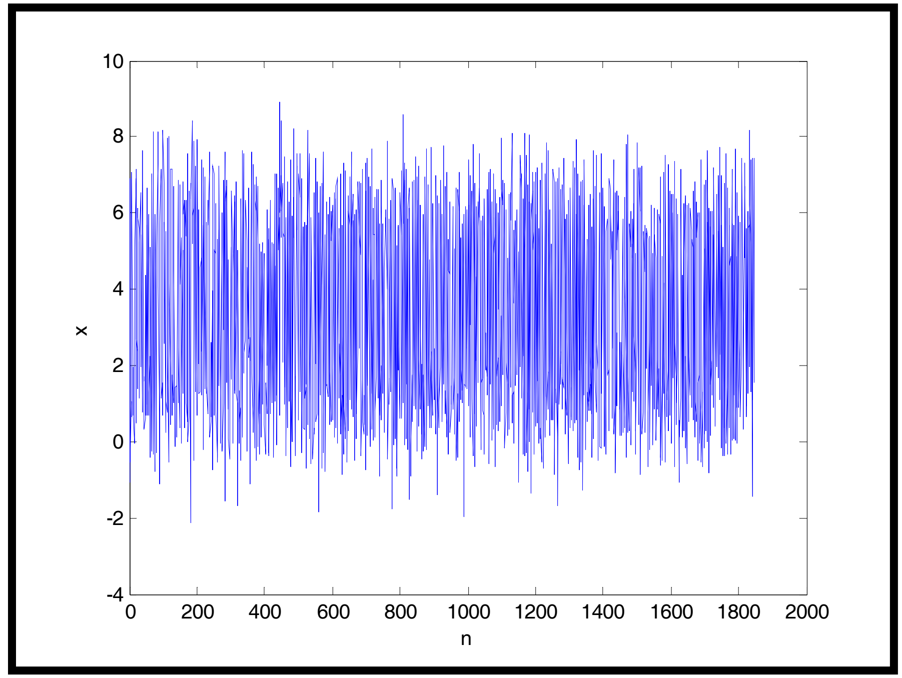
Example

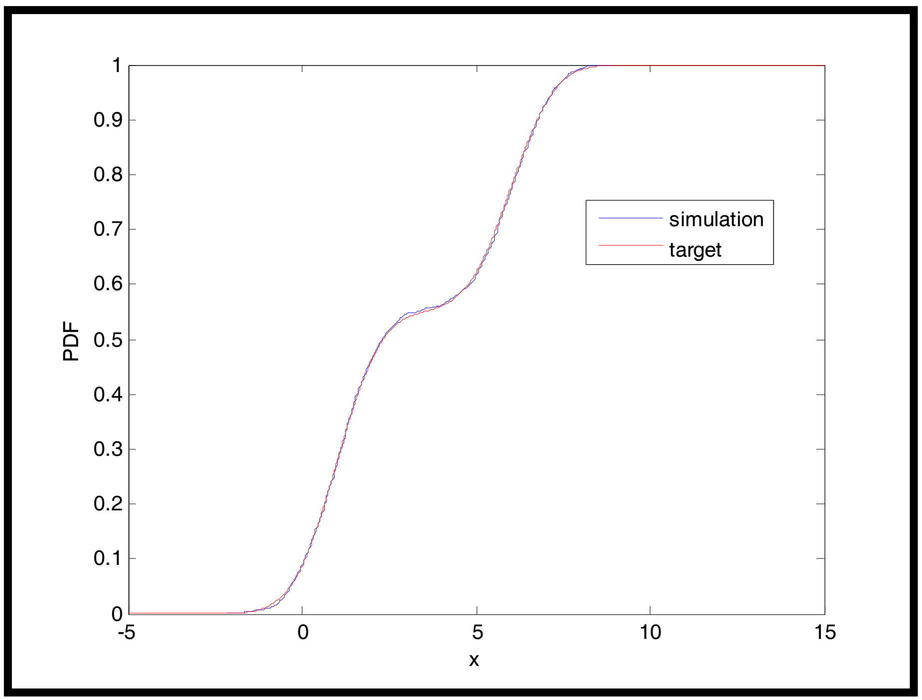
$$p_X(x) = a_1 \left\{ \frac{1}{\sqrt{2\pi}} \exp\left[-\frac{1}{2} \left(\frac{x-2}{1} \right)^2 \right] \right\} +$$

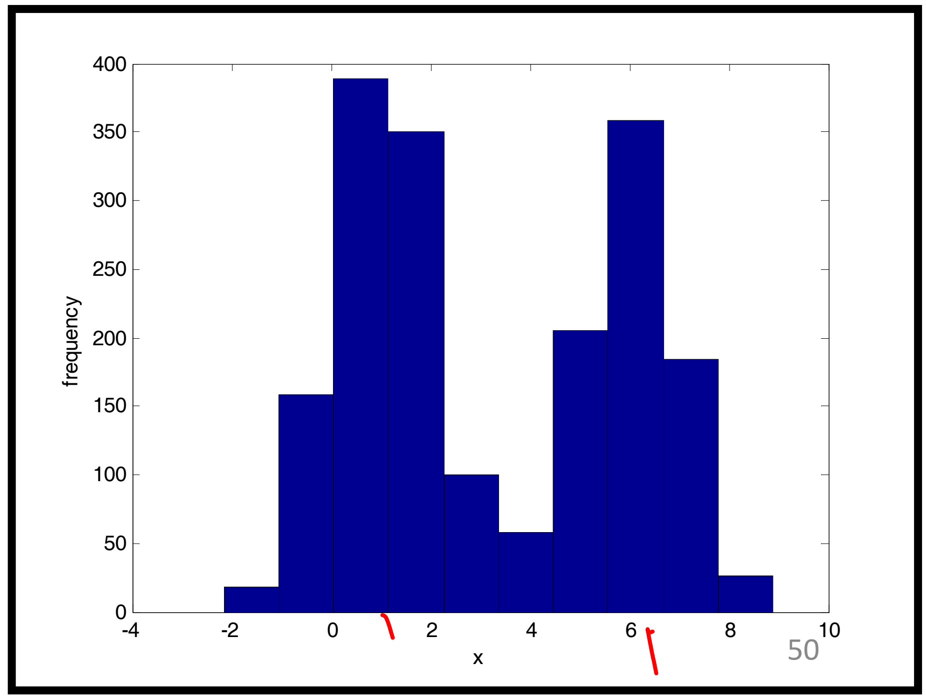
$$a_2 \left\{ \frac{1}{4\sqrt{2\pi}} \exp\left[-\frac{1}{2} \left(\frac{x-6}{1} \right)^2 \right] \right\}; a_1 + a_2 = 1$$

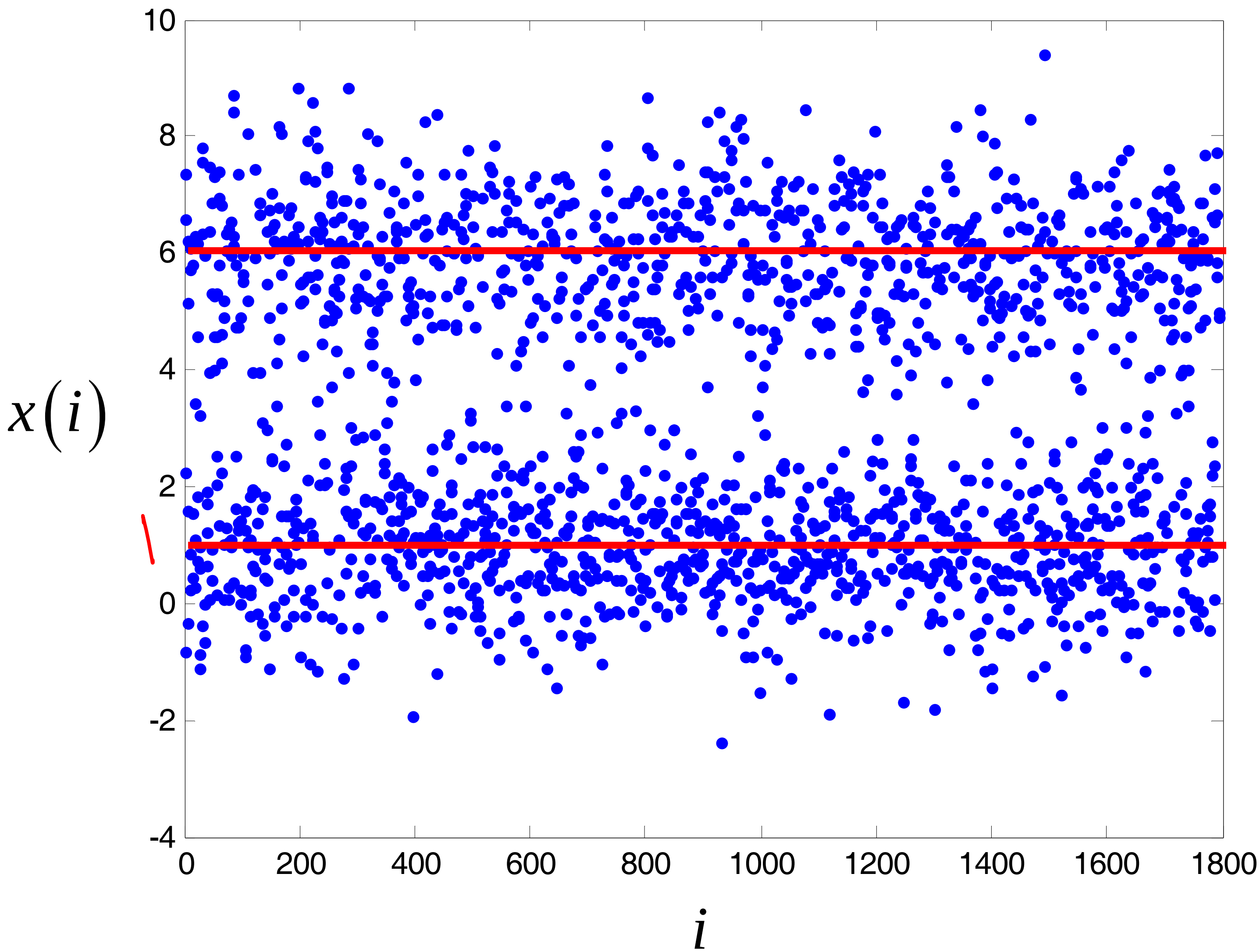
$$-\infty < x < \infty$$


p(x)
Mg(x)
pdf,Mg(x)
x


x
n


simulation
target
PDF
x


frequency
x


10
8
6
4
2
0
-2
-4
0
200
400
600
800
1000
1200
1400
1600
1800
$x(i)$
i

Simulation of vector Gaussian random variables

$\left\{X_i\right\}_{i=1}^{n}$ = correlated Gaussian rvs.

$$\left\langle X_i\right\rangle = \mu_i ; C_{ij} = \left\langle\left(X_i-\mu_i\right)\left(X_j-\mu_j\right)\right\rangle$$

$$X_i' = \frac{X_i-\mu_i}{\sigma_i} \Rightarrow \left\langle X_i'\right\rangle = 0$$

$$C_{ij}' = \left\langle X_i' X_j'\right\rangle = \begin{bmatrix} 1 & C_{12}' & C_{13}' \cdots & C_{1n}' \\ C_{21}' & 1 & C_{23}' \cdots & C_{2n}' \\ \vdots & \vdots & \vdots & \vdots \\ C_{n1}' & C_{n2}' & C_{n3}' \cdots & 1 \end{bmatrix}$$

Transform

$$Y = T^t X' \Rightarrow \left\langle Y\right\rangle = 0; \left\langle YY^t\right\rangle = \left\langle T^t X' X'^t T\right\rangle = T^t C' T$$

Select T such that $T^t C' T = I$

How to select T?

Consider the eigenvalue problem $[C']\{\alpha\} = \lambda\{\alpha\}$

eigenvalues: $|C' - I\lambda| = 0 \Rightarrow n$ eigenvalues $\{\lambda_i\}_{i=1}^{n}$

C' is positive definite $\Rightarrow \lambda_i > 0 \forall i = 1, 2, \cdots, n$

eigenvectors : Φ.

$$C'\phi_i = \lambda_i \phi_i$$

$$C'\phi_j = \lambda_j \phi_j$$

$$\phi_j^t C' \phi_i = \lambda_i \phi_j^t \phi_i$$

$$\phi_i^t C' \phi_j = \lambda_j \phi_i^t \phi_j$$

$$\phi_j^t C' \phi_i = \lambda_j \phi_j^t \phi_i \quad \left(\because C' = C'^t\right)$$

$$\Rightarrow (\lambda_i - \lambda_j)\phi_j^t \phi_i = 0 \Rightarrow \phi_j^t \phi_i = 0 \forall i \neq j \Rightarrow \phi_i^t C' \phi_j = 0 \forall i \neq j$$

Select Φ such that $\Phi^t C' \Phi = I$.

Take $T = \Phi$.

Target

$$\mu = \begin{Bmatrix} 0.1 \\ 2.0 \end{Bmatrix} \; \& \; C = \begin{bmatrix} 1.2 & 0.3 \\ 0.3 & 4.5 \end{bmatrix}$$

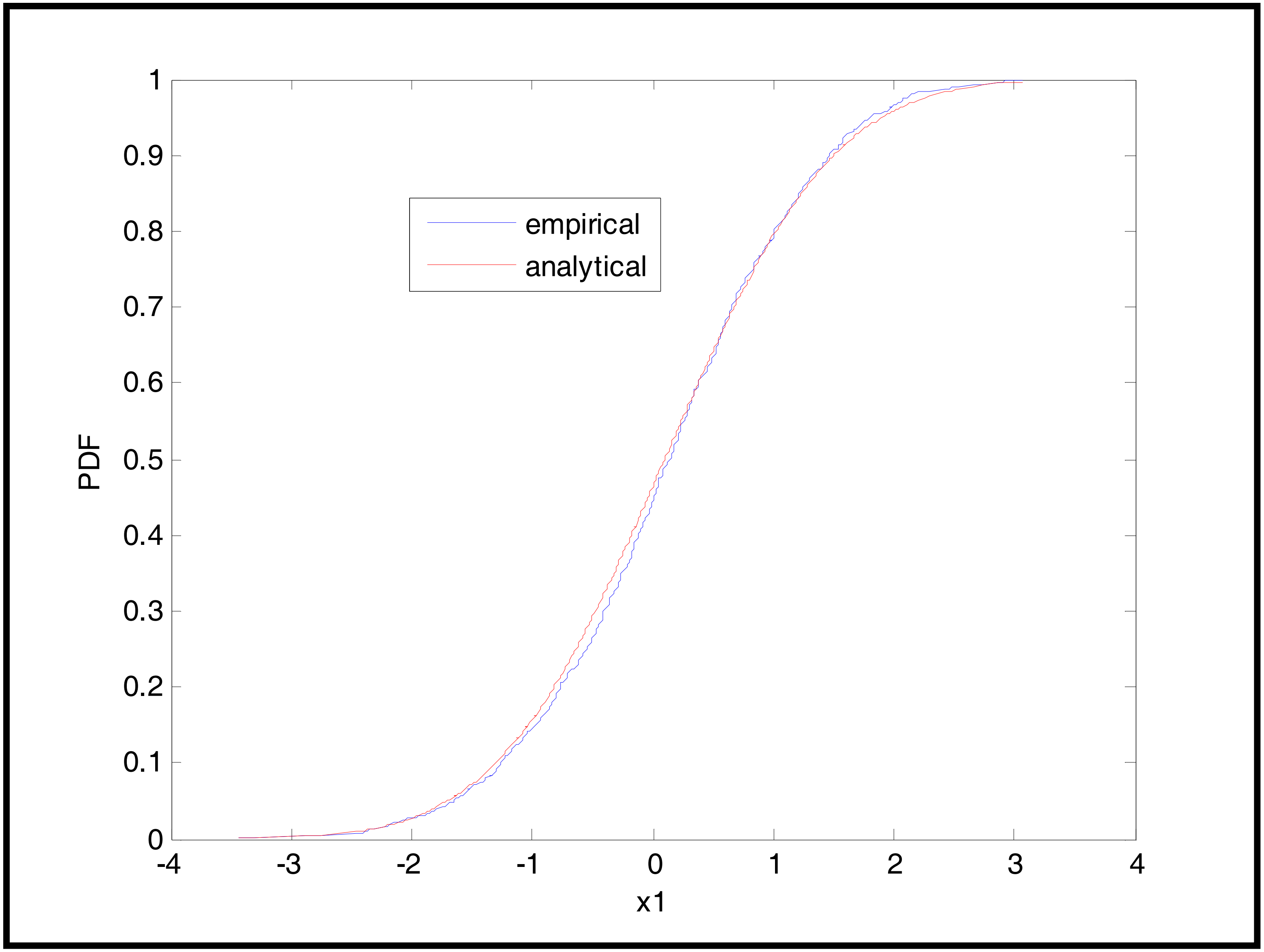
$$C' = \begin{bmatrix} 1.0000 & 0.1291 \\ 0.1291 & 1.0000 \end{bmatrix}$$

$$T = \begin{bmatrix} -0.7071 & 0.7071 \\ 0.7071 & 0.7071 \end{bmatrix}$$

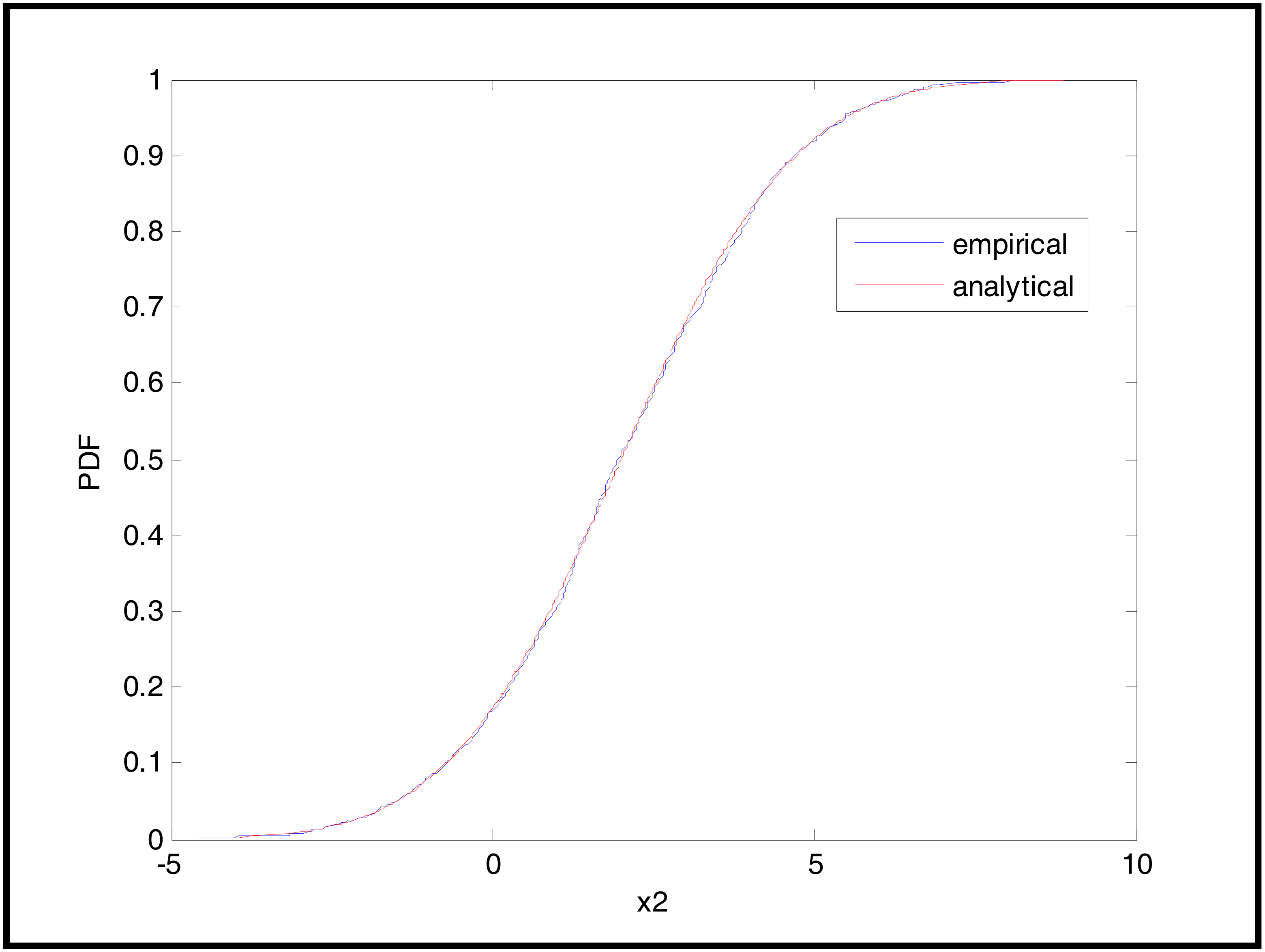
$$\{\lambda\} = \begin{Bmatrix} 0.8709 \\ 1.1291 \end{Bmatrix}$$

Simulated (with 1000 samples)

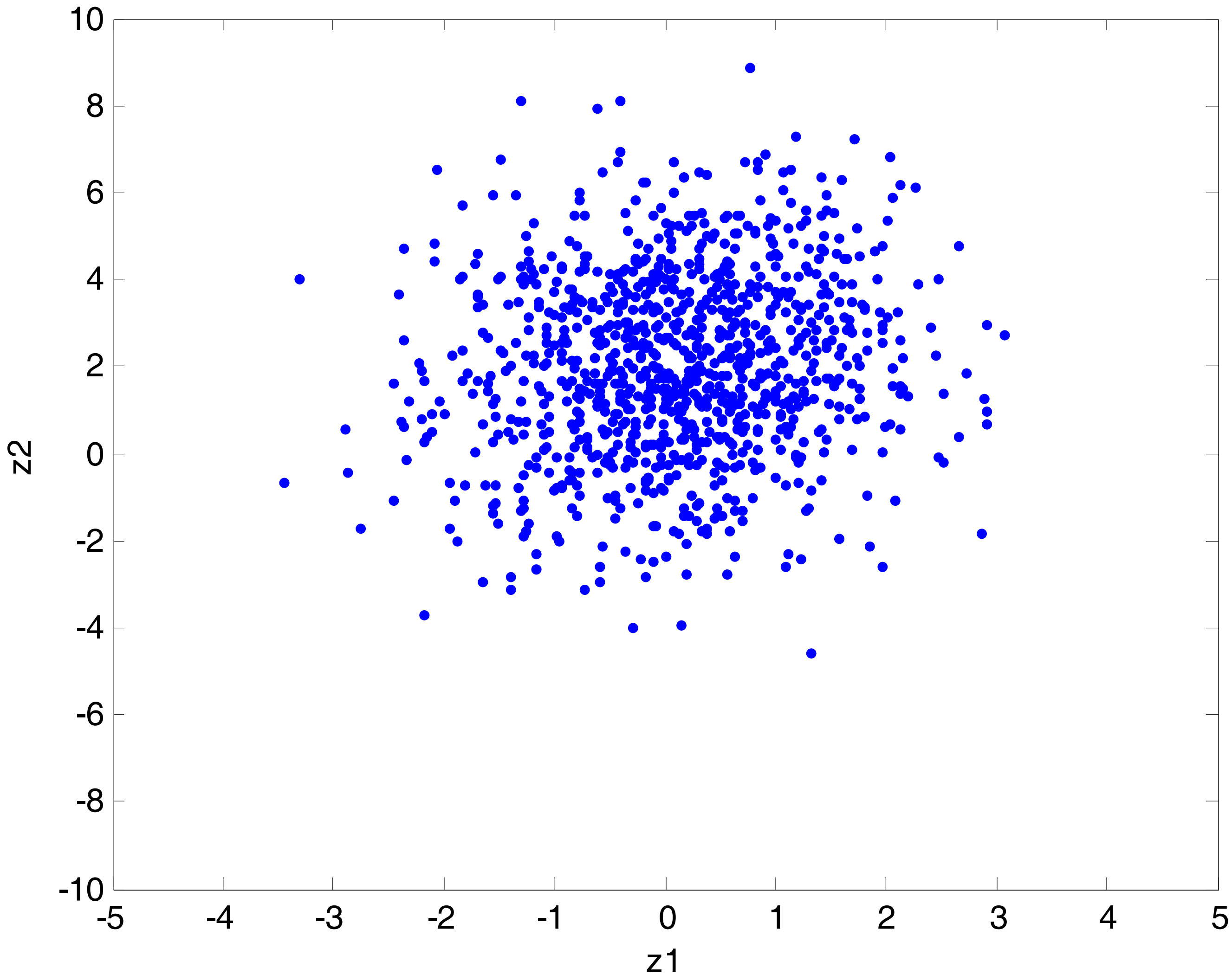
$$\mu = \begin{Bmatrix} 0.1224 \\ 2.0311 \end{Bmatrix} \; \& \; C = \begin{bmatrix} 1.1077 & 0.2826 \\ 0.2826 & 4.4713 \end{bmatrix}$$



empirical
analytical
PDF
x1
1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
-4
-3
-2
-1
0
1
2
3
4



1
0.9
0.8
0.7
0.6
0.5
0.4
0.3
0.2
0.1
0
PDF
-5
0
5
10
x2
empirical
analytical



10
8
6
4
2
0
-2
-4
-6
-8
-10
z2
-5
-4
-3
-2
-1
0
1
2
3
4
5
z1

Simulation of vector non - Gaussian random variables

Rosenblatt transformation

Let X_1 & X_2 be two non-Gaussian RVs.

JPDF: $P_{12}(x_1, x_2)$

jpdf : $p_{12}(x_1, x_2)$

MPDF: $P_1(x_1)$ & $P_2(x_2)$

mpdf : $p_1(x_1)$ & $p_2(x_2)$

Let $U_1 \to N(0,1)$ & $U_2 \to N(0,1)$ with $U_1 \perp U_2$

Define

$$P_1(X_1) = \Phi(U_1)$$

$$P_2(X_2 \mid X_1) = \Phi(U_2)$$

$$p_1(x_1)\frac{dx_1}{du_1} = \phi(u_1);\ \frac{dx_1}{du_2} = 0$$

$$p_2(x_2 \mid x_1)\frac{dx_2}{du_2} = \phi(u_2)$$

$$\Rightarrow\ p_{12}(x_1, x_2) = \frac{\phi(u_1, u_2)}{\mid J \mid} @ u_1$$

$$= \Phi^{-1}\left\{P_1(x_1)\right\} \& u_2 = \Phi^{-1}\left\{P_2(x_2 \mid x_1)\right\}$$

$$\Rightarrow\ p_{12}(x_1, x_2) =$$

$$\frac{\phi(u_1)\phi(u_2)}{\phi(u_1)\phi(u_2)} p_2(x_2 \mid x_1) p_1(x_1) = p_{12}(x_1, x_2)$$

Generalization

Let X be a non-Gaussian vector of RVs.

Let $\{U_i\}_{i=1}^{n}$ such that $U_i \sim N(0,1)$ & $U_i \perp U_j \forall i \neq j$.

The Rosenblatt transformation is given by

$$\Phi(U_1) = P_1(X_1)$$

$$\Phi(U_2) = P_2(X_2 \mid X_1)$$

$$\Phi(U_3) = P_3(X_3 \mid X_2, X_1)$$

$$\vdots$$

$$\Phi(U_n) = P_n(X_n \mid X_{n-1}, X_{n-2}, \cdots, X_1)$$

Remark

To implement the Rosenblatt transformation we need the complete specification of JPDF of X.

Partially specified non-Gaussian RVs
Nataf's transformation

Let X_1 and X_2 be two random variables such that

- X_1 and X_2 are not completely specified
- Knowledge on X_1 and X_2 is limited to first order pdfs and the covariance matrix.

Question: How to transform X to standard normal space?

JCSS (2002)

Steel as a 5-dimensional random variable

Description	COV
Yield strength	0.07
Ultimate tensile strength	0.04
Young's modulus	0.03
Poisson's ratio	0.03
Ultimate strain	0.06

$$\rho = \begin{bmatrix} 1 & 0.75 & 0 & 0 & -0.45 \\ & 1 & 0 & 0 & -0.60 \\ & & 1 & 0 & 0 \\ & & & 1 & 0 \\ & & & & 1 \end{bmatrix}$$

Distribution: first order pdf-s are lognormal

Let

$$P_{X_1}(X_1) = \Phi(U_1)$$

$$P_{X_2}(X_2) = \Phi\left(U_2\right)$$

with $U_1 \sim N(0,1), U_2 \sim N(0,1)\ \&\ \left\langle U_1 U_2 \right\rangle = \rho_{12}^*$

$$\Rightarrow p_1(x_1)\frac{dx_1}{du_1} = \phi(u_1); \frac{dx_1}{du_2} = 0$$

$$p_2(x_2)\frac{dx_2}{du_2} = \phi(u_2); \frac{dx_2}{du_1} = 0$$

$$J = \begin{vmatrix} \dfrac{\phi(u_1)}{p_1(x_1)} & 0 \\ 0 & \dfrac{\phi(u_2)}{p_2(x_2)} \end{vmatrix} = \frac{\phi(u_1)\phi(u_2)}{p_1(x_1)p_2(x_2)}$$

$$p_{X_1X_2}(x_1,x_2)=\frac{p_{U_1U_2}(u_1,u_2)}{\phi(u_1)\phi(u_2)}p_1(x_1)p_2(x_2)$$

$$@u_1=\Phi^{-1}\left[P_{X_1}(x_1)\right],u_2=\Phi^{-1}\left[P_{X_2}(x_2)\right]$$

$$=\frac{\phi_2\left\{\Phi^{-1}\left[P_{X_1}(x_1)\right],u_2=\Phi^{-1}\left[P_{X_2}(x_2)\right]\right\}}{\phi\left\{\Phi^{-1}\left[P_{X_1}(x_1)\right]\right\}\phi\left\{\Phi^{-1}\left[P_{X_2}(x_2)\right]\right\}}p_1(x_1)p_2(x_2)$$

$$\rho_{12}=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}(x_1-\mu_1)(x_2-\mu_2)p_{X_1X_2}(x_1,x_2)dx_1dx_2$$

$$=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}(x_1-\mu_1)(x_2-\mu_2)\frac{\phi_2\left\{\Phi^{-1}\left[P_{X_1}(x_1)\right],u_2=\Phi^{-1}\left[P_{X_2}(x_2)\right]\right\}}{\phi\left\{\Phi^{-1}\left[P_{X_1}(x_1)\right]\right\}\phi\left\{\Phi^{-1}\left[P_{X_2}(x_2)\right]\right\}}$$

$$p_1(x_1)p_2(x_2)dx_1dx_2$$

Substitute

$$P_{X_1}\left(x_1\right)=\Phi\left(z_1\right)\,\&\,P_{X_2}\left(x_2\right)=\Phi\left(z_1\right)$$

$$\Rightarrow dx_1 dx_2\, p_{\mathrm{X}_1}\left(x_1\right) p_{\mathrm{X}_1}\left(x_1\right)=\phi\left(z_1\right)\phi\left(z_1\right)dz_1 dz_2$$

$$\Rightarrow \rho_{12}=\int_{-\infty}^{\infty}\int_{-\infty}^{\infty}(P_{X_1}^{-1}\left\{\Phi\left(z_1\right)\right\}-\mu_1)(P_{X_1}^{-1}\left\{\Phi\left(z_1\right)\right\}-\mu_2)\phi_2\left(z_1,z_2,\rho_{12}^{*}\right)dz_1 dz_2$$

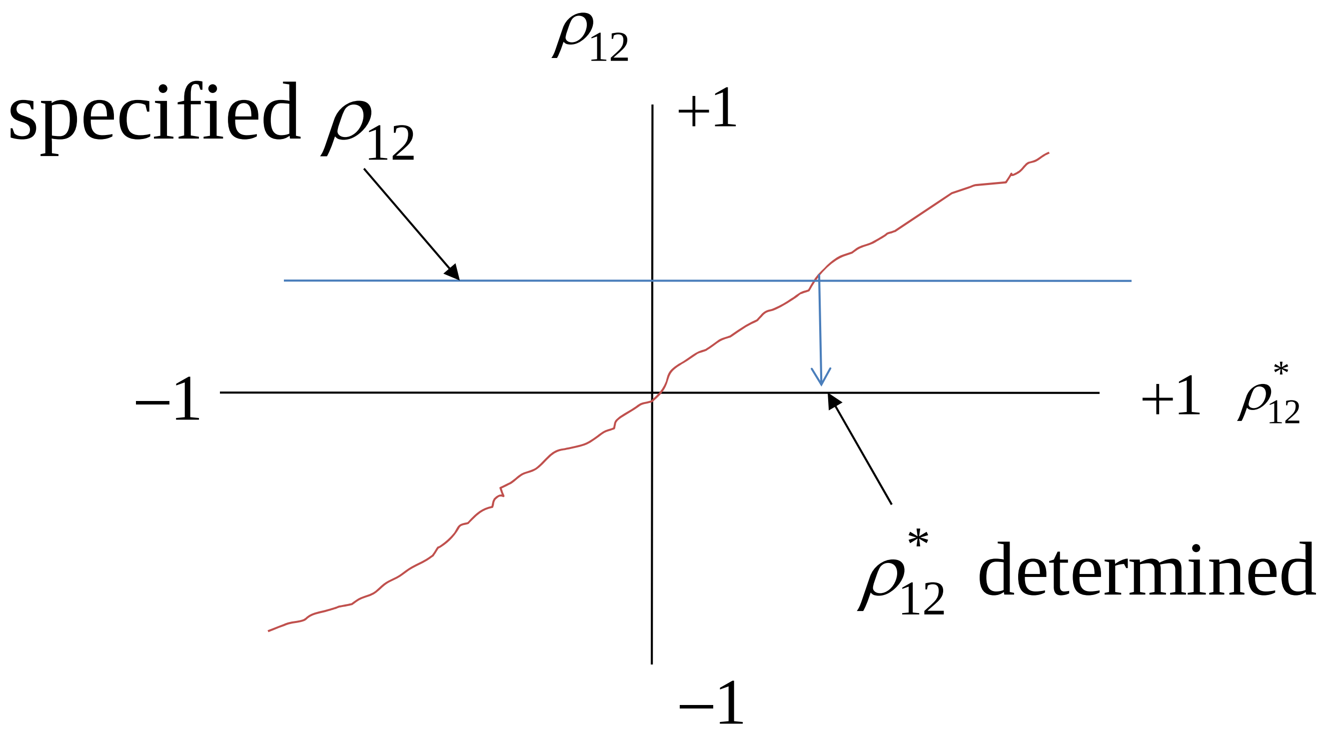
Strategy for the determination of the unknown ρ_{12}^{*}

$$\rho_{12} = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} (P_{X_1}^{-1}\left\{\Phi\left(z_1\right)\right\} - \mu_1)(P_{X_2}^{-1}\left\{\Phi\left(z_2\right)\right\} - \mu_2)\phi_2\left(z_1, z_2, \rho_{12}^{*}\right)dz_1 dz_2$$

(1) Divide the range [-1,1] of ρ_{12}^{*} into L divisions.

(2) For each value of $\left\{\rho_{12}^{*i}\right\}_{i=1}^{L}$ solve the above equation (numerically) and obtain the corresponding values of $\left\{\rho_{12}^{i}\right\}_{i=1}^{L}$. Note that $-1 \le \rho_{12}^{i} \le 1 \forall i = 1, 2, \cdots, L.$

(3) Interpolate $\left\{\rho_{12}^{i}\right\}_{i=1}^{L}$ to obtain the value of ρ_{12}^{*} for which the target value of ρ_{12} is realized.



ρ_{12}
specified ρ_{12}
+1
−1
+1 ρ^*_{12}
ρ^*_{12} determined
−1

Remarks

- Generalization to the *n*-dimensional case is straight forward
- Nataf's transformation leads to correlated Gaussian rvs. A further transformation is needed to reach the standard normal space.
- Requires solution of an integral equation
- This itself can be done numerically.
- For *n*-dimensional case the number of integral equations to be solved becomes $n(n-1)/2$.
- Other forms of partial information can be handled within the framework of the Nataf's model.

Steps for simulation of 2 - dimensional Nataf random variables

Step 1 solve for ρ_{12}^{*} by solving

$$\rho_{12} = \int_{-\infty}^{\infty}\int_{-\infty}^{\infty} (P_{X_1}^{-1}\left\{\Phi\left(z_1\right)\right\} - \mu_1)(P_{X_2}^{-1}\left\{\Phi\left(z_2\right)\right\} - \mu_2)\phi_2\left(z_1, z_2, \rho_{12}^{*}\right)dz_1 dz_2$$

Step 2 Simulate $Z \sim N\left(0, \rho^{*}\right)$.

Step 3 Simulate X_1 and X_2 using

$$X_i = P_X^{-1}\left\{\Phi\left(U_i\right)\right\}; i = 1, 2$$

Example

X is uniformly distributed in 0 to 100.
Y is exponentially distributed with parameter 1/0.08.
Z is Rayleigh distributed with parameter 12.

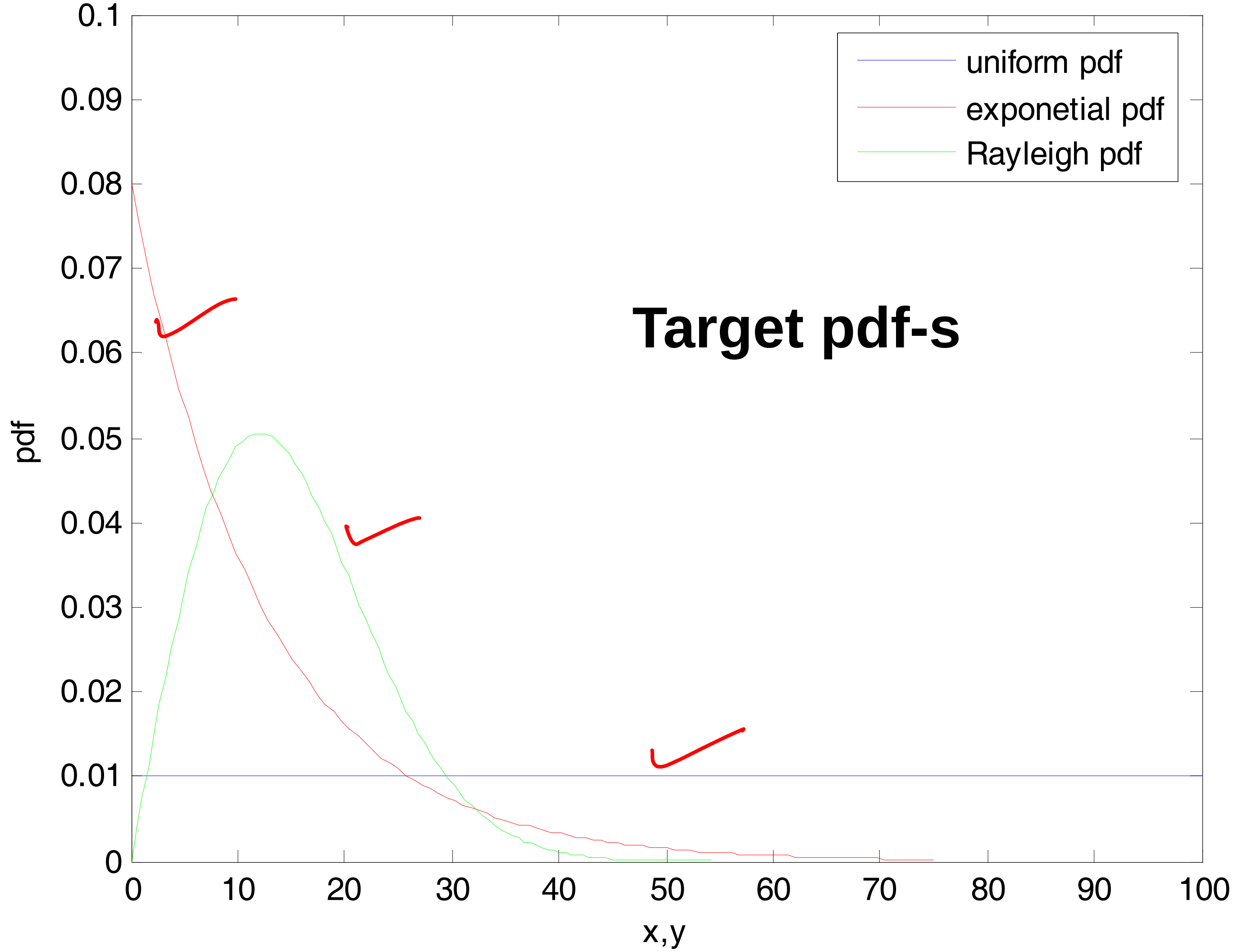
mean_target = 50.0000 12.5000 15.0398
mean_simulated = 50.2530 12.2213 15.2910
std_target = 28.8675 12.5000 7.8616
std_simulated = 29.0108 12.2915 8.0349

rho_target =

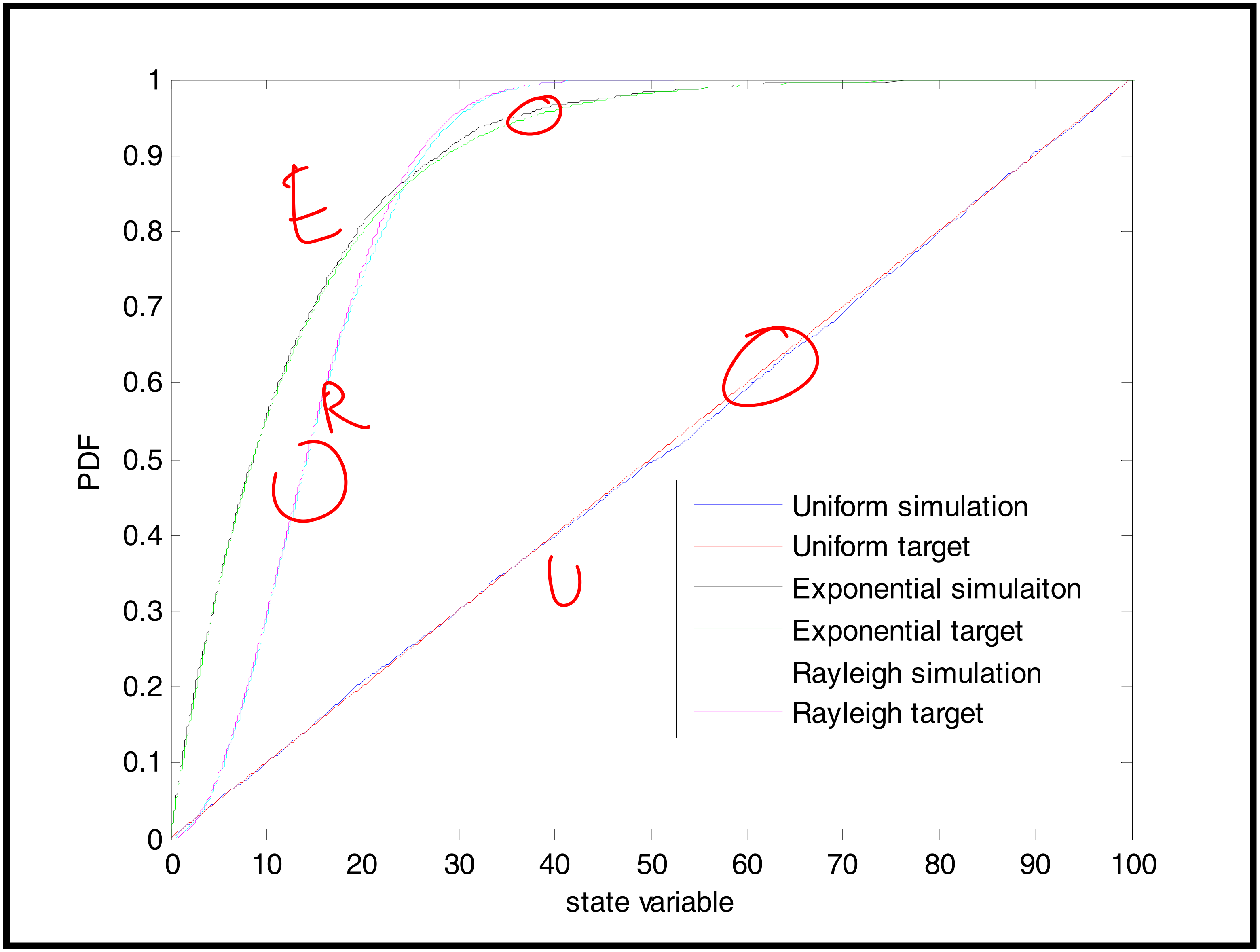
1.0000	0.3400	0.4500
0.3400	1.0000	-0.5000
0.4500	-0.5000	1.0000

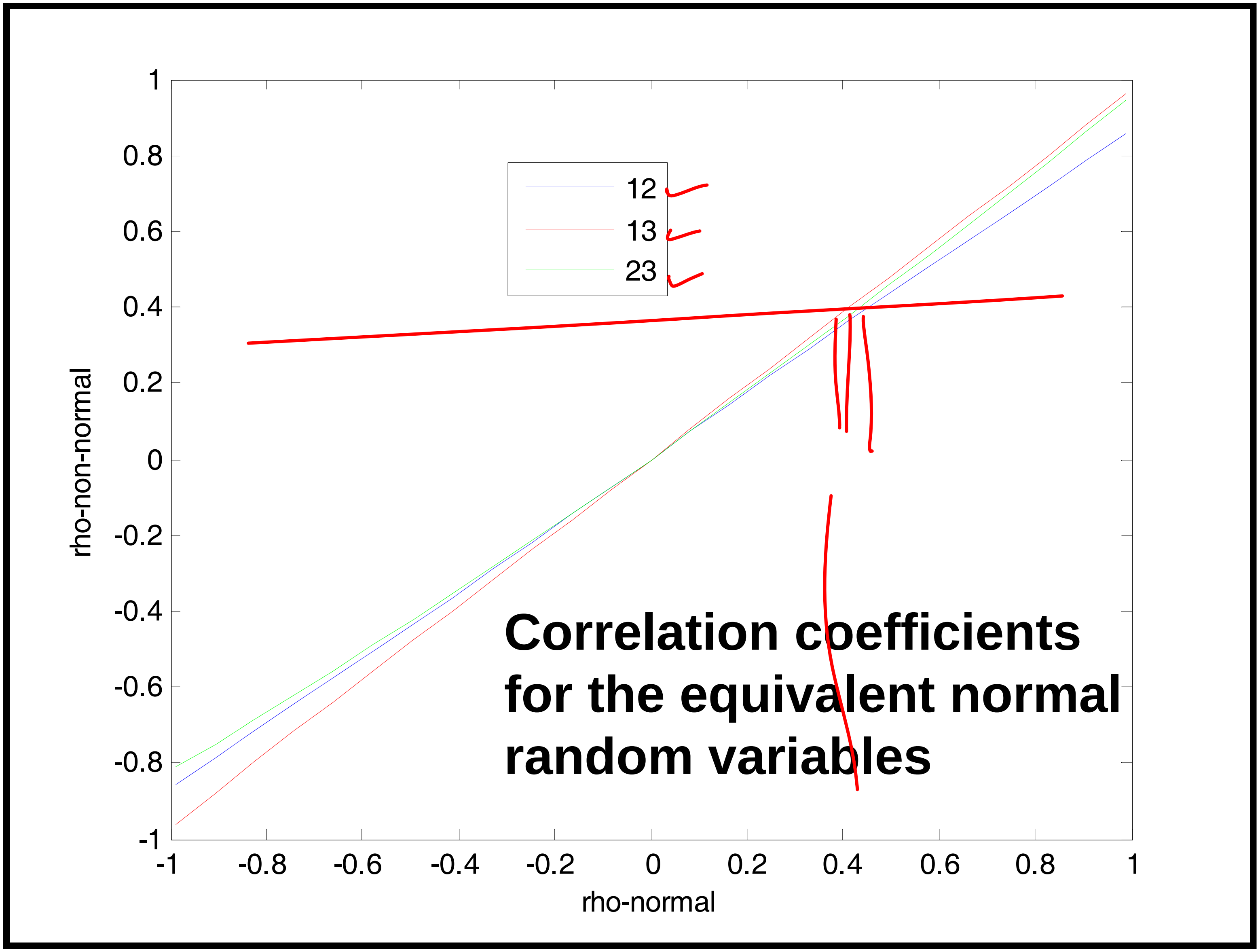
rho_simulated =

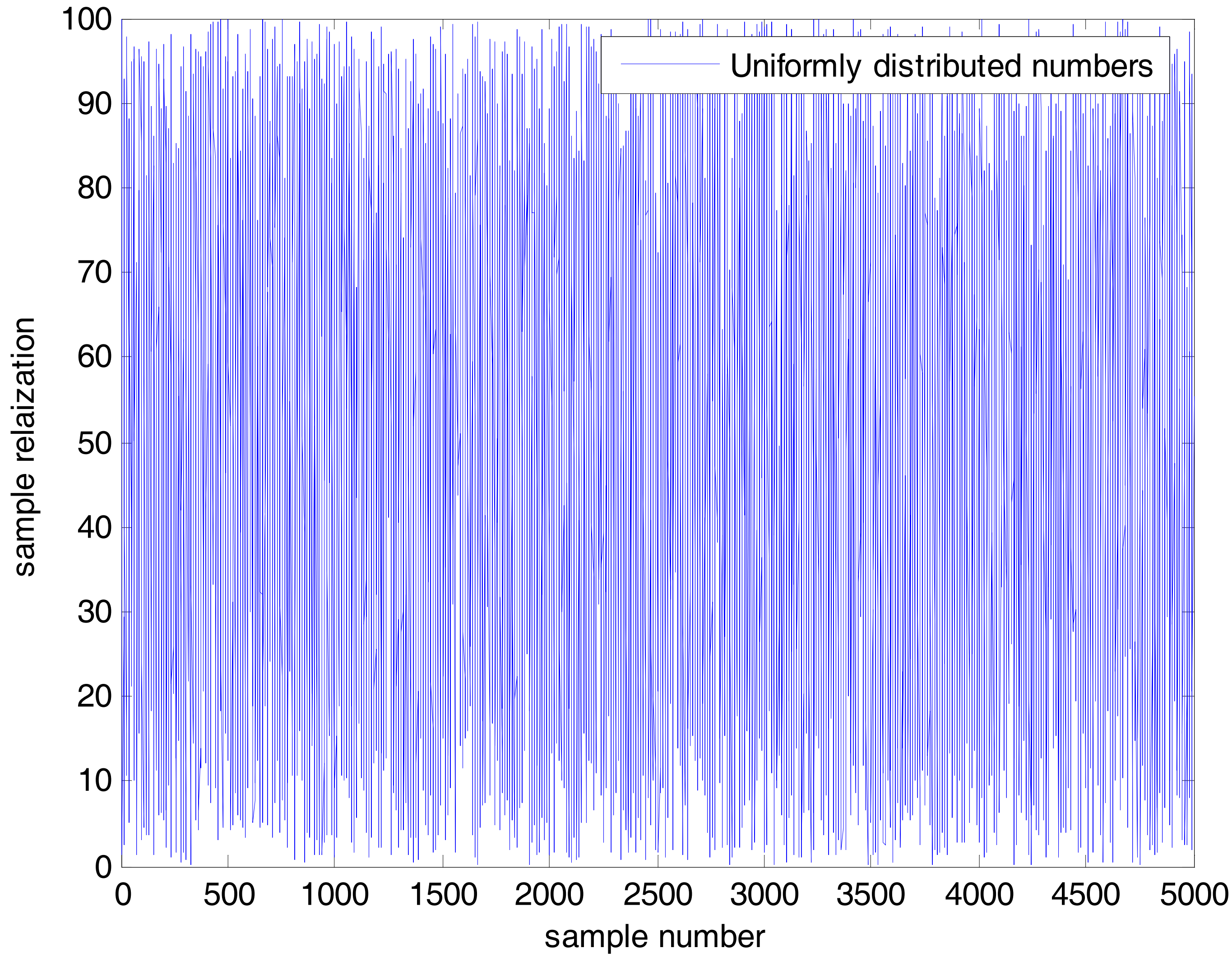
1.0000	0.3146	0.4686
0.3146	1.0000	-0.5014
0.4686	-0.5014	1.0000

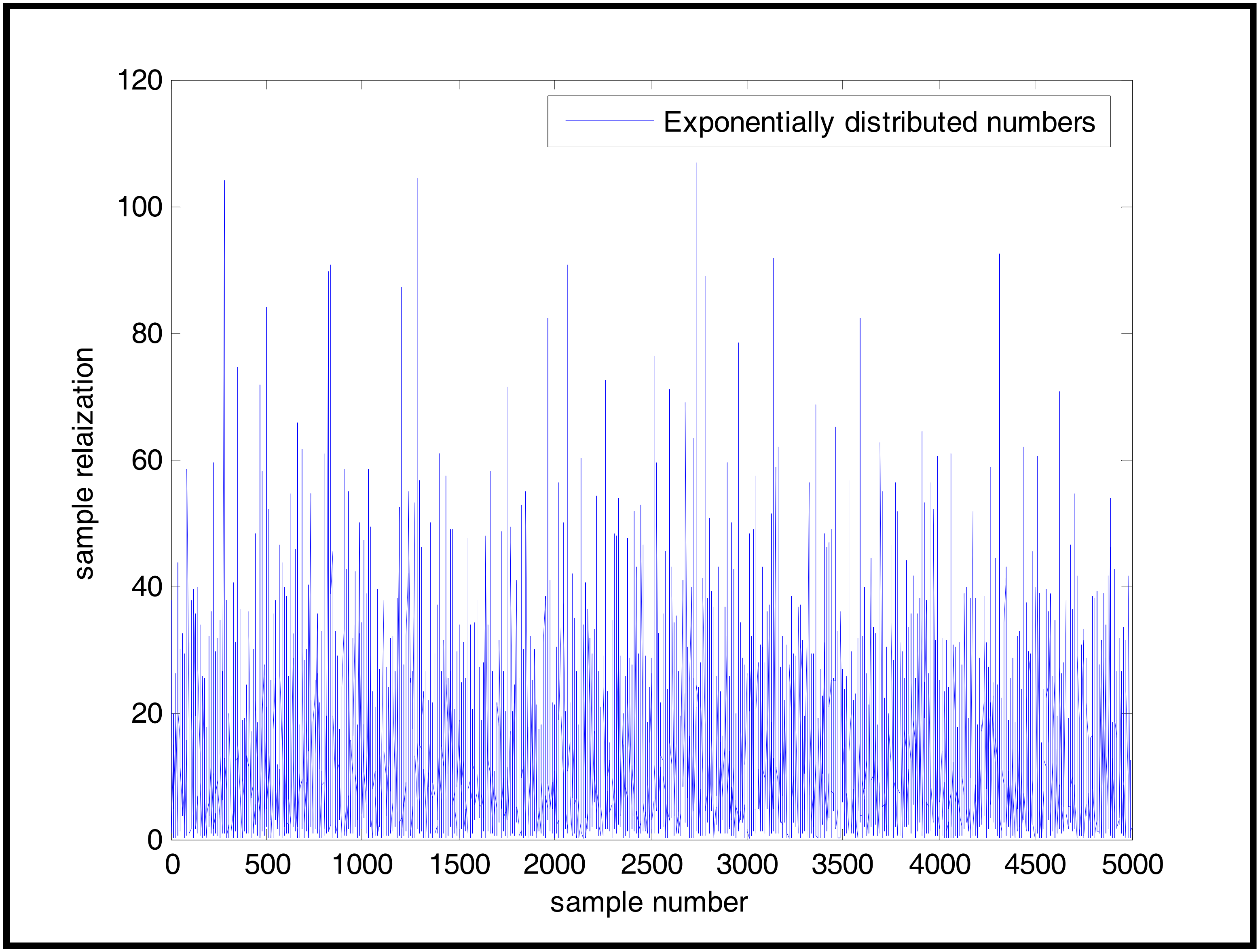

0.1
0.09
0.08
0.07
0.06
0.05
0.04
0.03
0.02
0.01
0
pdf
0
10
20
30
40
50
60
70
80
90
100
x,y
uniform pdf
exponetial pdf
Rayleigh pdf
Target pdf-s

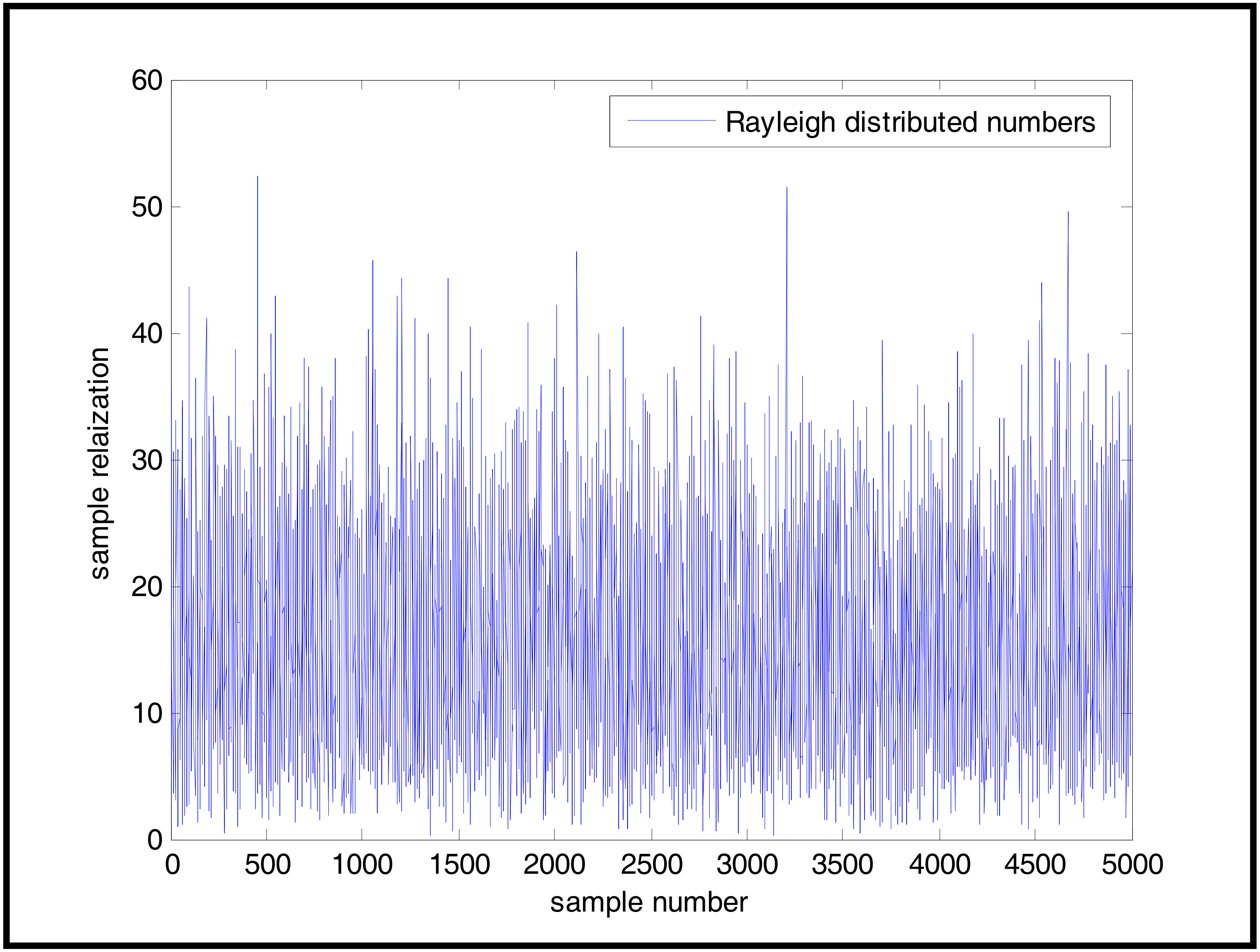
Target and simulated PDF-s




1
0.8
0.6
0.4
0.2
0
-0.2
-0.4
-0.6
-0.8
-1
rho-non-normal
12
13
23
Correlation coefficients
for the equivalent normal
random variables
-1
-0.8
-0.6
-0.4
-0.2
0
0.2
0.4
0.6
0.8
1
rho-normal


Uniformly distributed numbers
sample relaization
sample number
100
90
80
70
60
50
40
30
20
10
0
0
500
1000
1500
2000
2500
3000
3500
4000
4500
5000

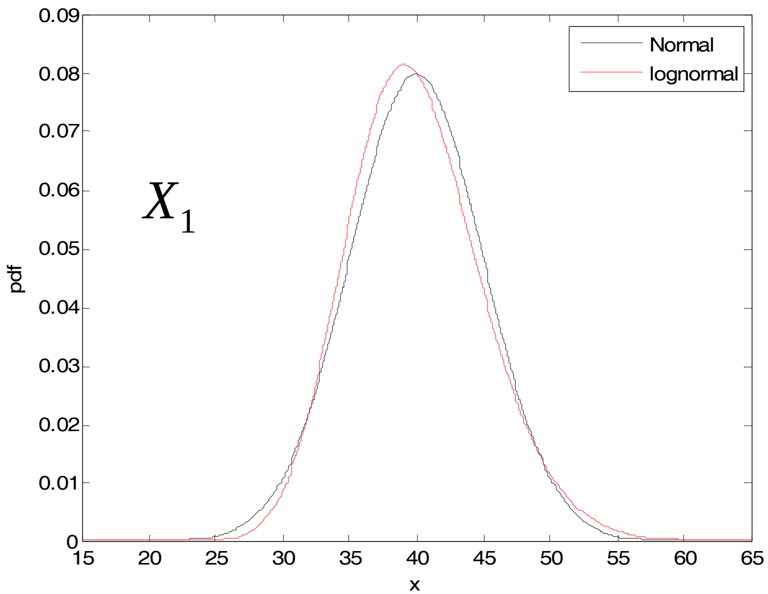

Exponentially distributed numbers
120
100
80
60
40
20
0
0
500
1000
1500
2000
2500
3000
3500
4000
4500
5000
sample relaization
sample number


Rayleigh distributed numbers
sample relaization
sample number
0
10
20
30
40
50
60
0
500
1000
1500
2000
2500
3000
3500
4000
4500
5000

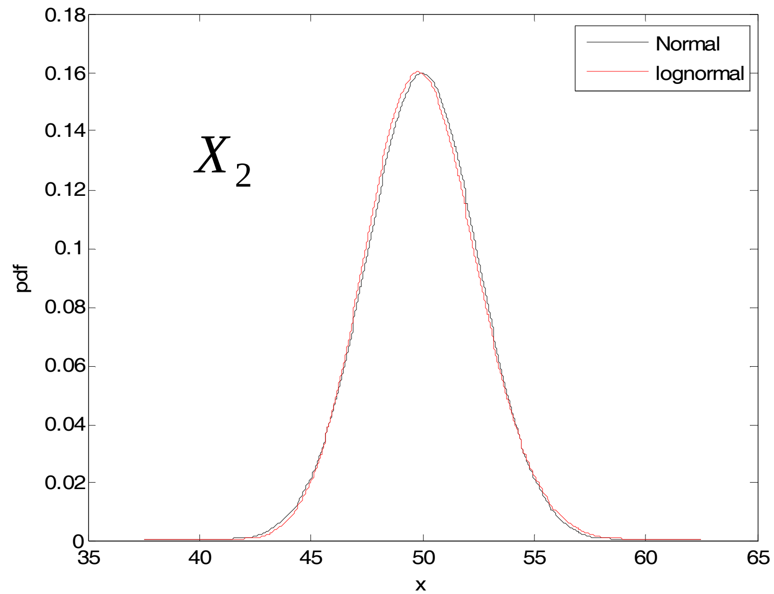
X_1=lognormal
X_2=lognormal
X_3=Type I asymptotic
Mu=40 50 1000
Stdev=5 2.5 200
rho =

1.0	0.4	0.0
0.4	1.0	0.0
0.0	0.0	1.0

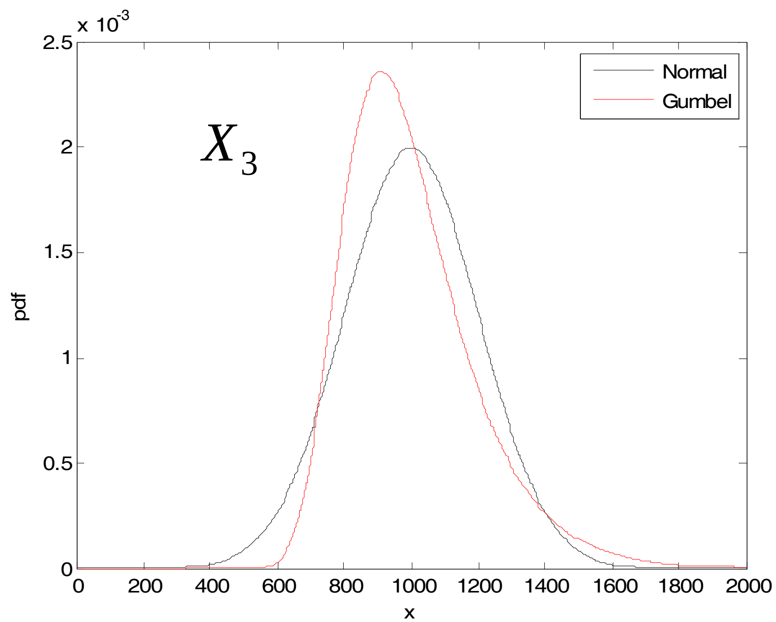
```
Rho_equivalent_gaussian
rho_t =
   1.0000    0.4011    0.0000
   0.4011    1.0000    0.0000
   0.0000    0.0000    1.0000
Check on simulations (with 5000 samples)
Msim=
  40.0210   50.0238  997.3724
Stdsim= 4.9544    2.4913  198.8895
Rhosim=
   1.0000    0.4123   -0.0045
   0.4123    1.0000   -0.0006
  -0.0045   -0.0006    1.0000
```



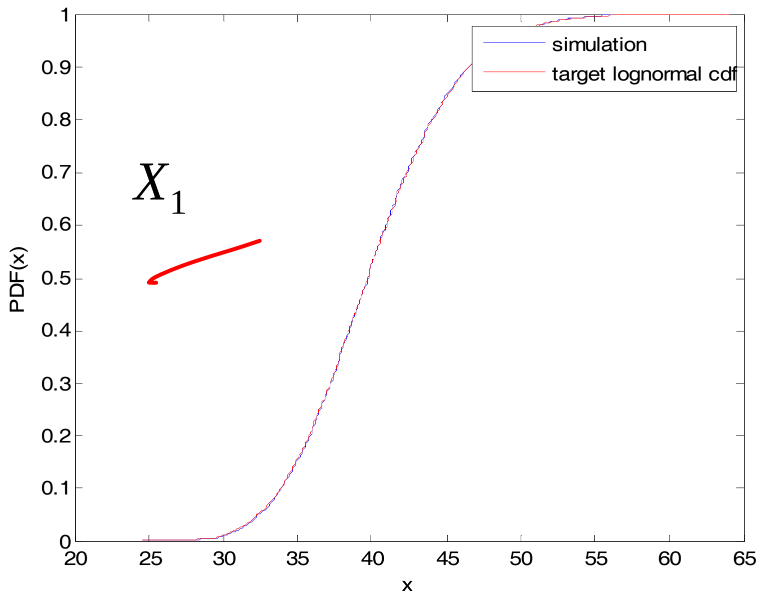
Normal
lognormal
X_1
pdf
x
0.09
0.08
0.07
0.06
0.05
0.04
0.03
0.02
0.01
0
15
20
25
30
35
40
45
50
55
60
65

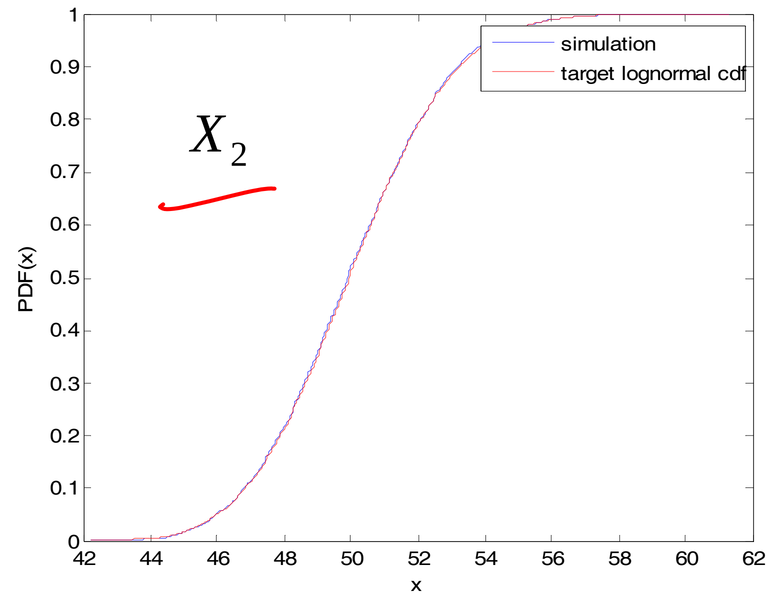


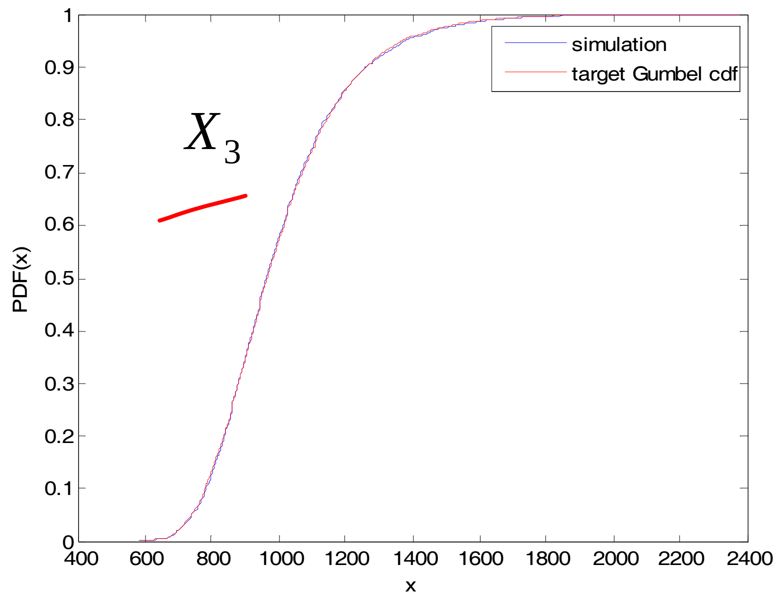
Normal
lognormal
X_2
pdf
x
0.18
0.16
0.14
0.12
0.1
0.08
0.06
0.04
0.02
0
35
40
45
50
55
60
65



x 10^{-3}
Normal
Gumbel
X_3
pdf
x
2.5
2
1.5
1
0.5
0
0
200
400
600
800
1000
1200
1400
1600
1800
2000


simulation
target lognormal cdf
X_1
PDF(x)
x


simulation
target lognormal cdf
X_2
PDF(x)
x


simulation
target Gumbel cdf
X_3
PDF(x)
x

Example

Let $n = 7$.

RV	PDF	P_1	P_2
X_1	Uniform	0.004	0.016
X_2	Lognormal	-0.01205E+02	0.000499E+02
X_3	Lognormal	0.058811E+02	0.000997E+02
X_4	Normal	0.000226	0.0000113
X_5	Evpdf	0.47749	25.65
X_6	Evpdf	0.11729	213.758
X_7	Normal	40.0	6.0

Target mean and standard deviation

Mu=	0.01	0.3	360.0	0.0002	0.50	0.12	40.0
Stdev=	0.0034	0.015	36.0	0.0000113	0.05	0.006	6.0

Target Correlation Coefficient matrix

1.00	0.00	0.10	0.30	0.00	0.40	0.10
0.00	1.00	-0.20	0.40	0.30	0.00	0.00
0.10	-0.20	1.00	-0.20	-0.10	0.00	0.00
0.30	0.40	-0.20	1.00	0.40	0.00	0.00
0.00	0.30	-0.10	0.40	1.00	0.50	0.00
0.40	0.00	0.00	0.00	0.50	1.00	0.00
0.10	0.00	0.00	0.00	0.00	0.00	1.00

Check on simulations (with 5000 samples)

Msim=0.0100 0.3002 360.1200 0.0002 0.4994 0.1200 39.9397

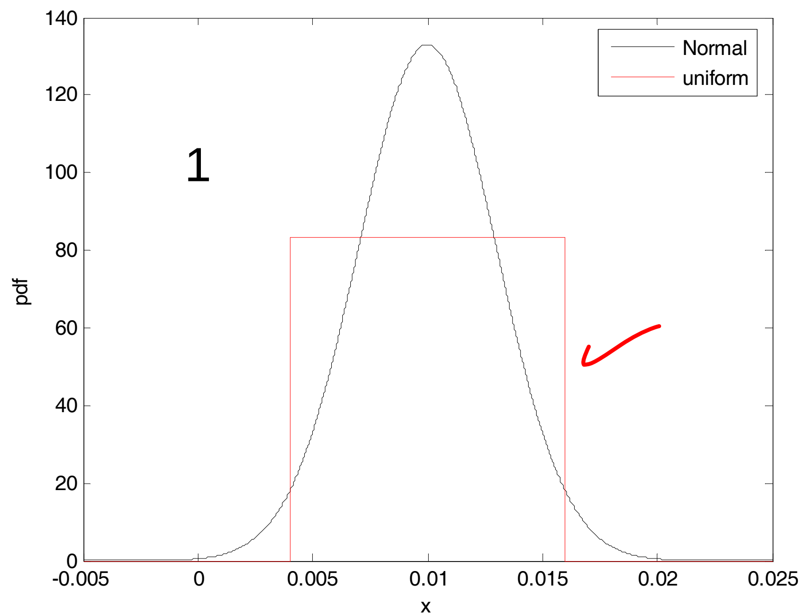
Stdsim=0.00347483519898 0.01492909017911 36.01027683948955 0.00001129812325 0.05008607386972 0.00610687237089 6.03945024214555

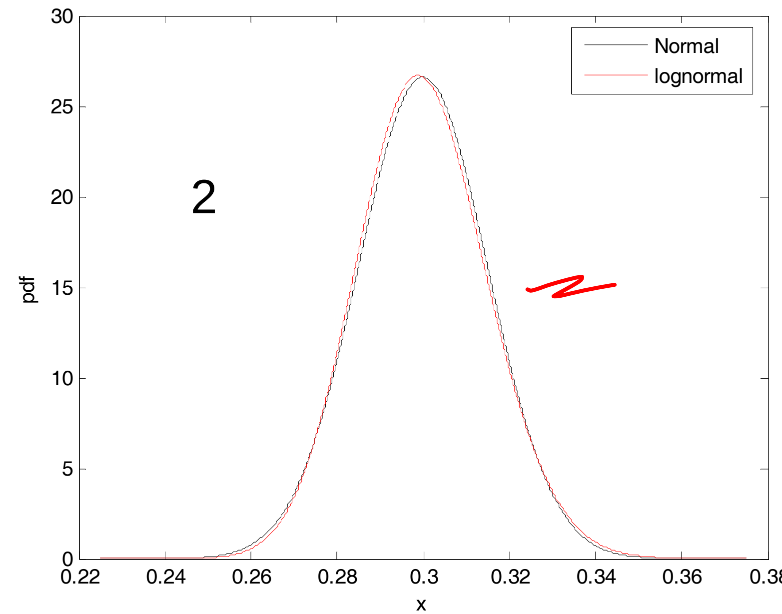
Rhosim=

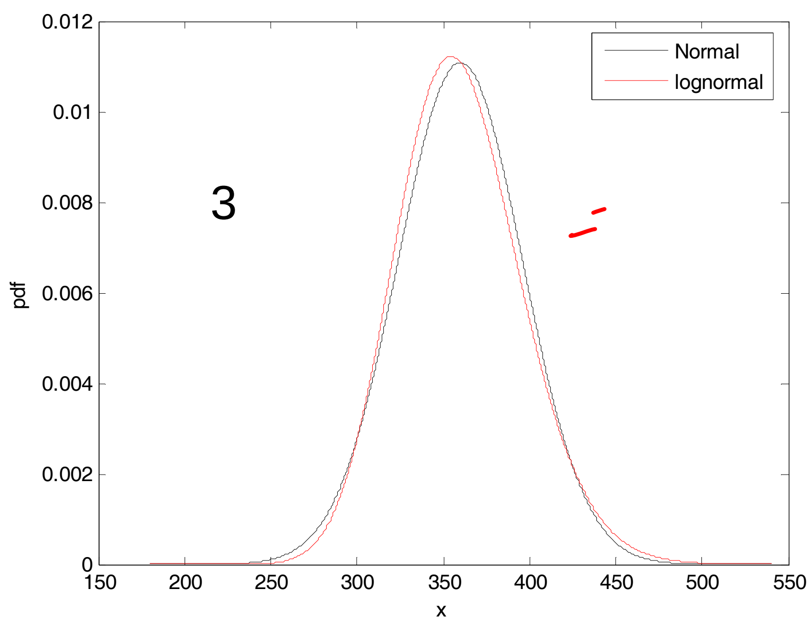
1.0000	0.0117	0.0868	0.3177	0.0072	0.4002	0.0988
0.0117	1.0000	-0.1915	0.4001	0.3020	-0.0030	-0.0024
0.0868	-0.1915	1.0000	-0.2038	-0.0950	0.0061	0.0150
0.3177	0.4001	-0.2038	1.0000	0.4086	0.0151	-0.0043
0.0072	0.3020	-0.0950	0.4086	1.0000	0.4966	-0.0048
0.4002	-0.0030	0.0061	0.0151	0.4966	1.0000	0.0049
0.0988	-0.0024	0.0150	-0.0043	-0.0048	0.0049	1.0000

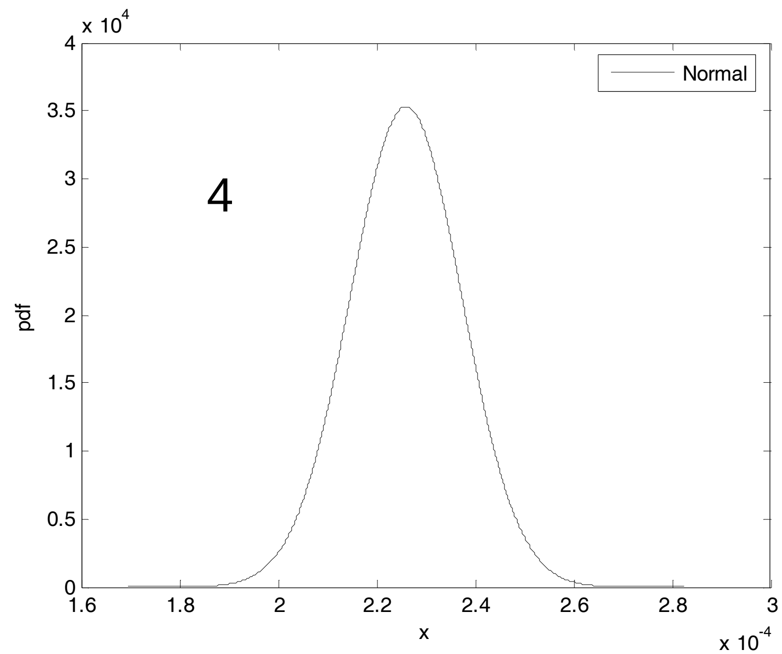
Rho_equivalent_gaussian rho_t =

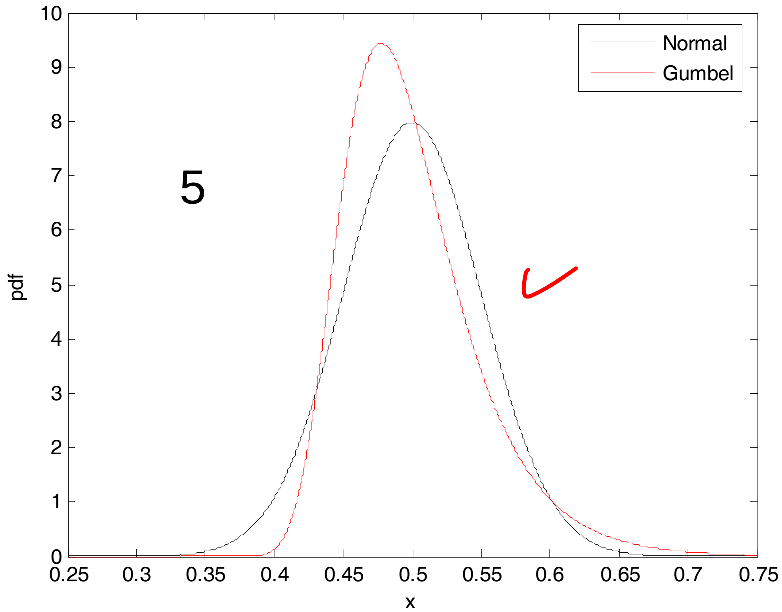
1.0000	0.0000	0.1026	0.3070	0.0000	0.4242	0.1023
0.0000	1.0000	-0.2008	0.4002	0.3085	0.0000	0.0000
0.1026	-0.2008	1.0000	-0.2005	-0.1051	0.0000	0.0000
0.3070	0.4002	-0.2005	1.0000	0.4123	0.0000	0.0000
0.0000	0.3085	-0.1051	0.4123	1.0000	0.5103	0.0000
0.4242	0.0000	0.0000	0.0000	0.5103	1.0000	0.0000
0.1023	0.0000	0.0000	0.0000	0.0000	0.0000	1.0000

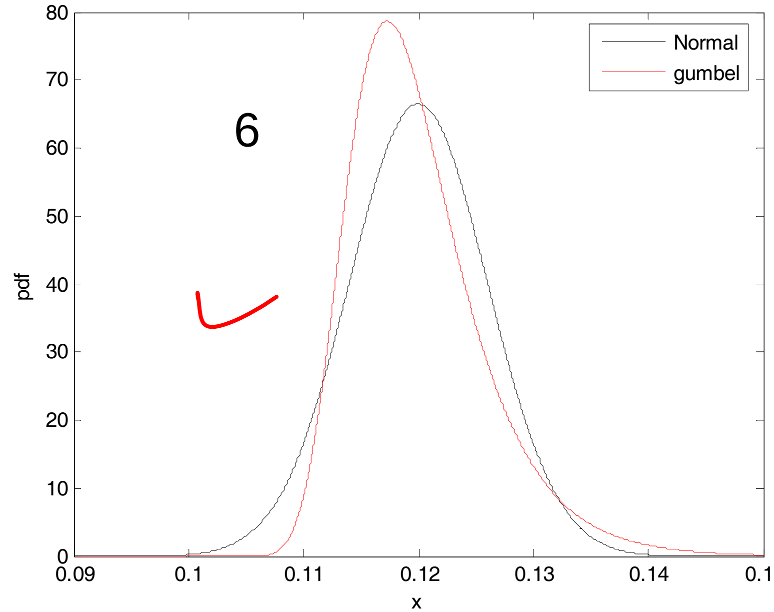

1
Normal
uniform
pdf
x
140
120
100
80
60
40
20
0
-0.005
0
0.005
0.01
0.015
0.02
0.025

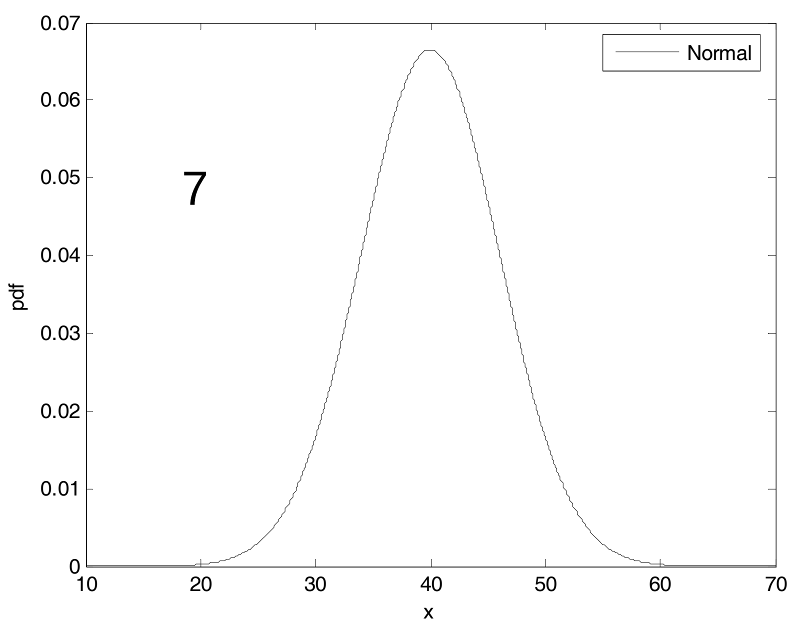

2
Normal
lognormal
pdf
x
30
25
20
15
10
5
0
0.22
0.24
0.26
0.28
0.3
0.32
0.34
0.36
0.38

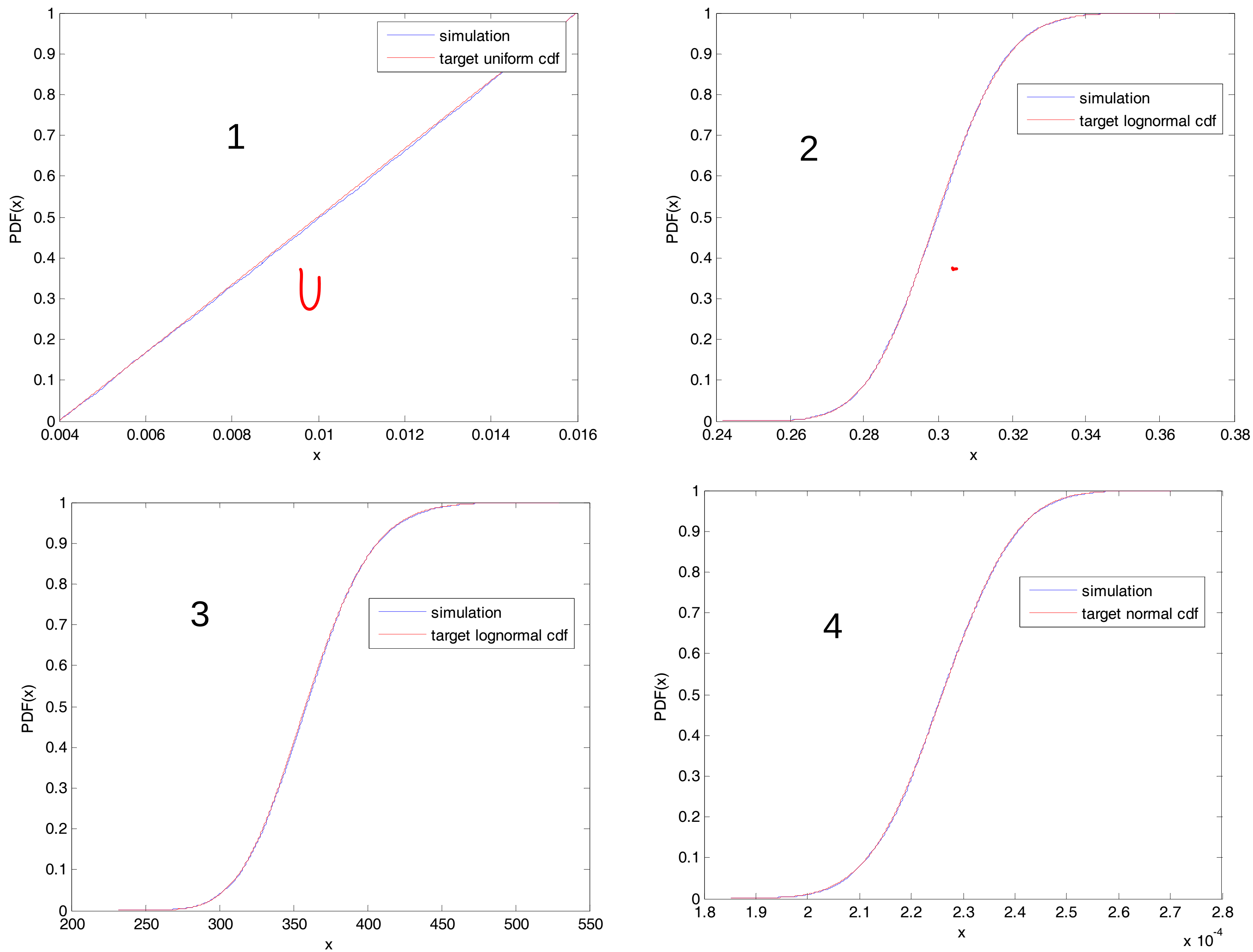

3
Normal
lognormal
pdf
x
0.012
0.01
0.008
0.006
0.004
0.002
0
150
200
250
300
350
400
450
500
550

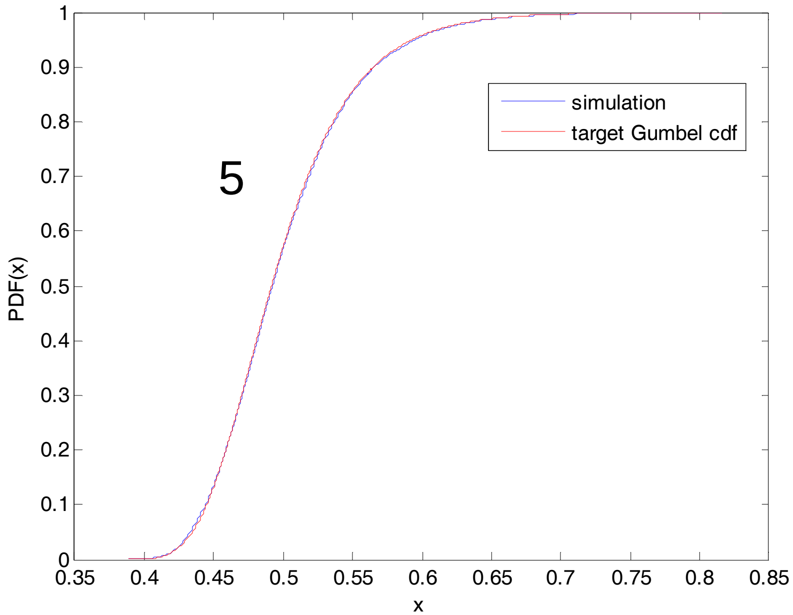

4
Normal
pdf
x
x 10^4
x 10^-4
4
3.5
3
2.5
2
1.5
1
0.5
0
1.6
1.8
2
2.2
2.4
2.6
2.8
3


5
Normal
Gumbel
pdf
x
0.25
0.3
0.35
0.4
0.45
0.5
0.55
0.6
0.65
0.7
0.75
0
1
2
3
4
5
6
7
8
9
10

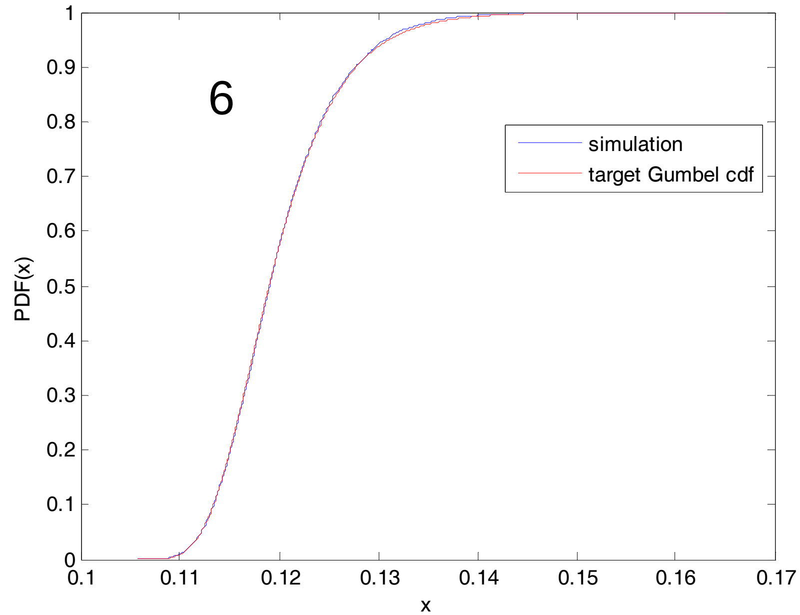

6
Normal
gumbel
pdf
x
0.09
0.1
0.11
0.12
0.13
0.14
0.15
0
10
20
30
40
50
60
70
80


7
Normal
pdf
x
10
20
30
40
50
60
70
0
0.01
0.02
0.03
0.04
0.05
0.06
0.07

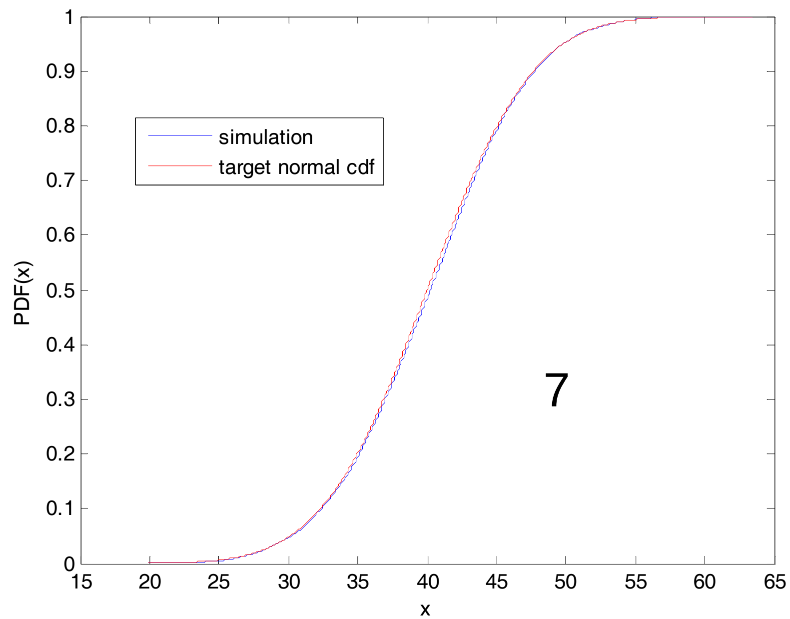

1
simulation
target uniform cdf
PDF(x)
x
2
simulation
target lognormal cdf
PDF(x)
x
3
simulation
target lognormal cdf
PDF(x)
x
4
simulation
target normal cdf
PDF(x)
x
x 10^-4



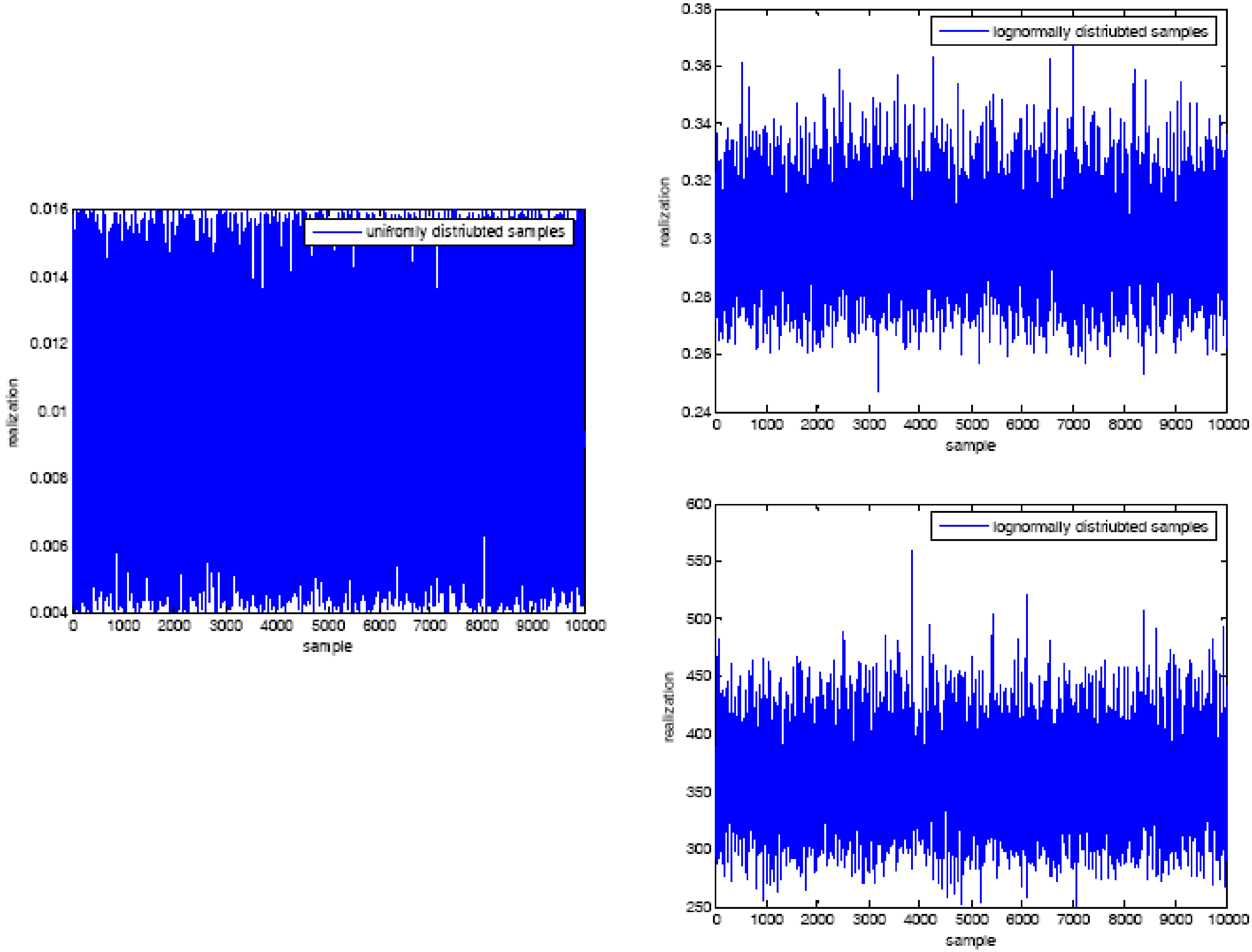
5
simulation
target Gumbel cdf
PDF(x)
x

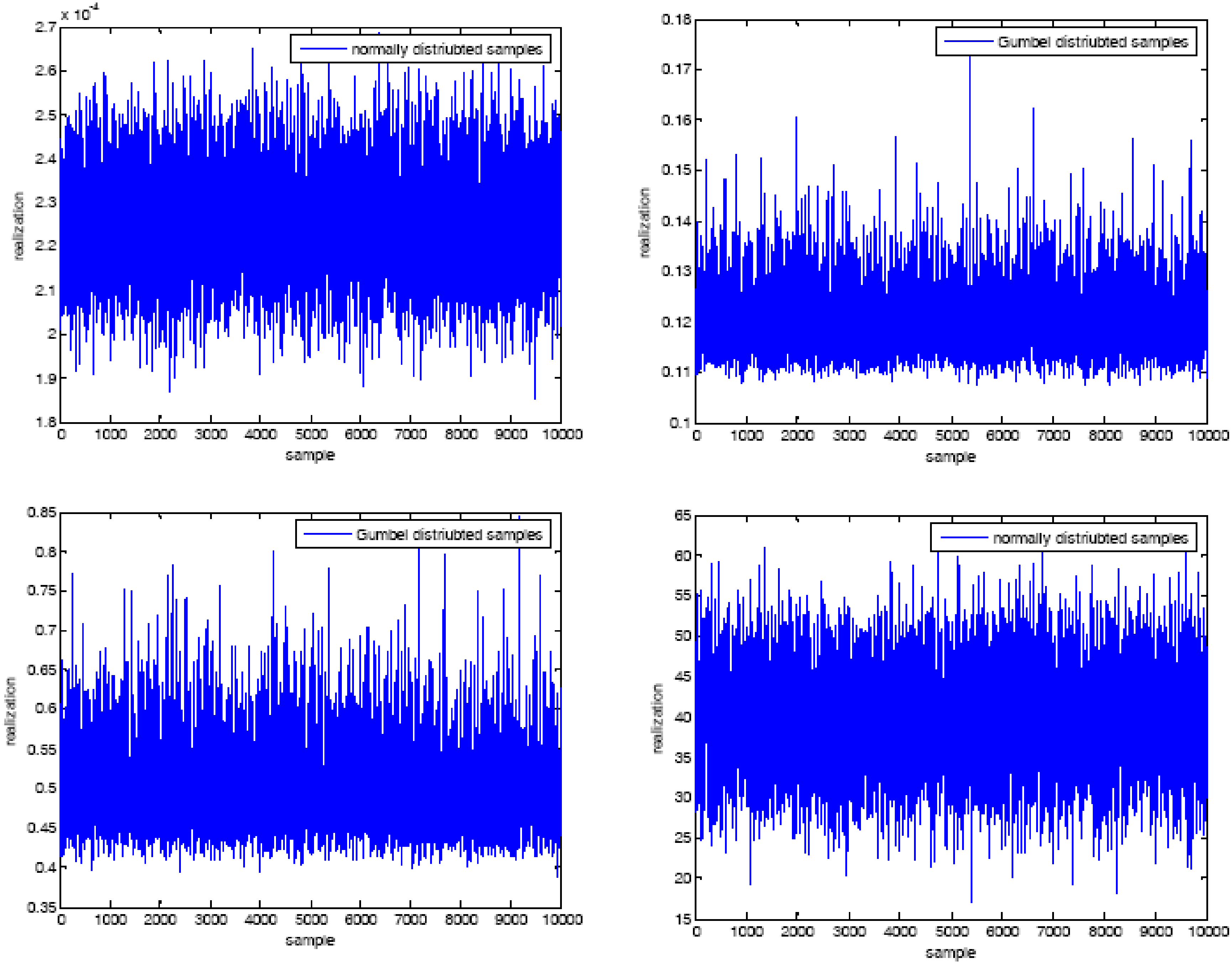


6
simulation
target Gumbel cdf
PDF(x)
x



7
simulation
target normal cdf
PDF(x)
x


unifromly distriubted samples
realization
sample
lognormally distriubted samples
realization
sample
lognormally distriubted samples
realization
sample



x 10^-4
normally distriubted samples
realization
sample
Gumbel distriubted samples
realization
sample
Gumbel distriubted samples
realization
sample
normally distriubted samples
realization
sample