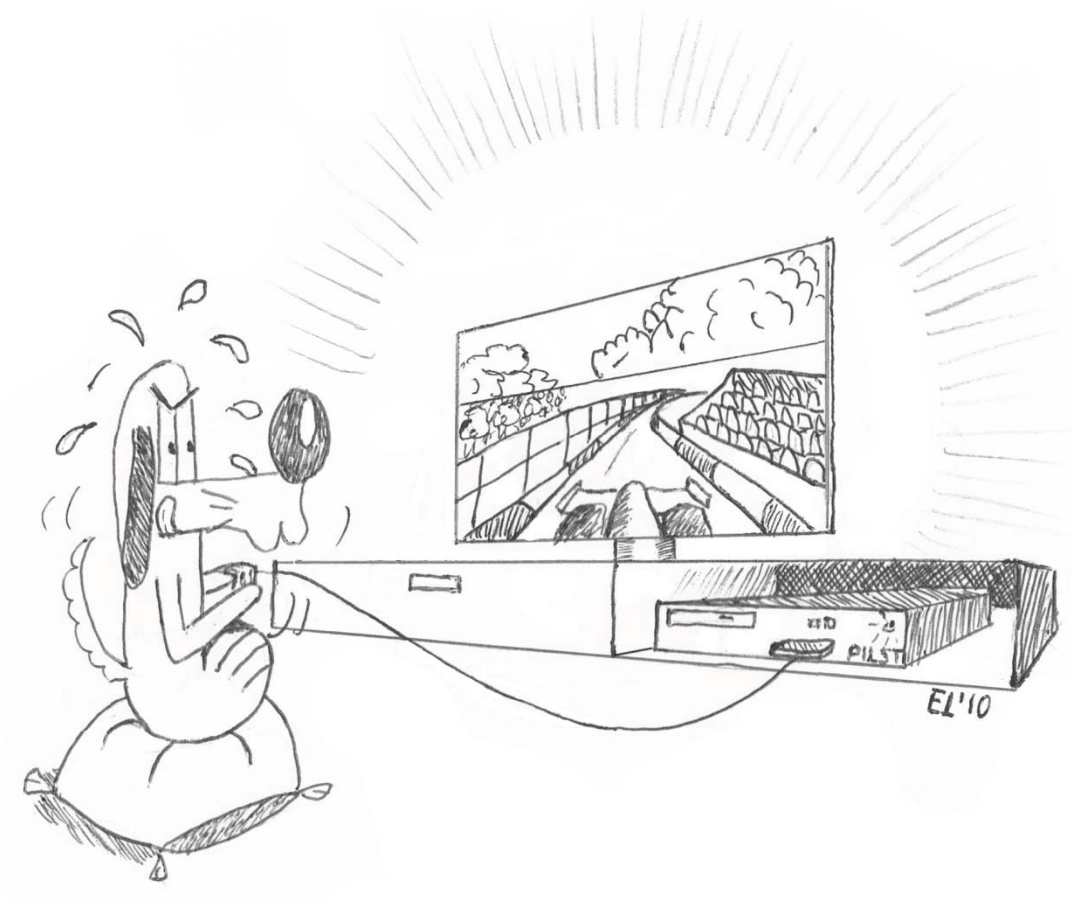
Simulation



Where real stuff
starts

March 2019

ToC

1. What is a simulation ?

Two simulation examples. What is the difference ?

Simulation 1: Time to complete $\boldsymbol{N}$ jobs

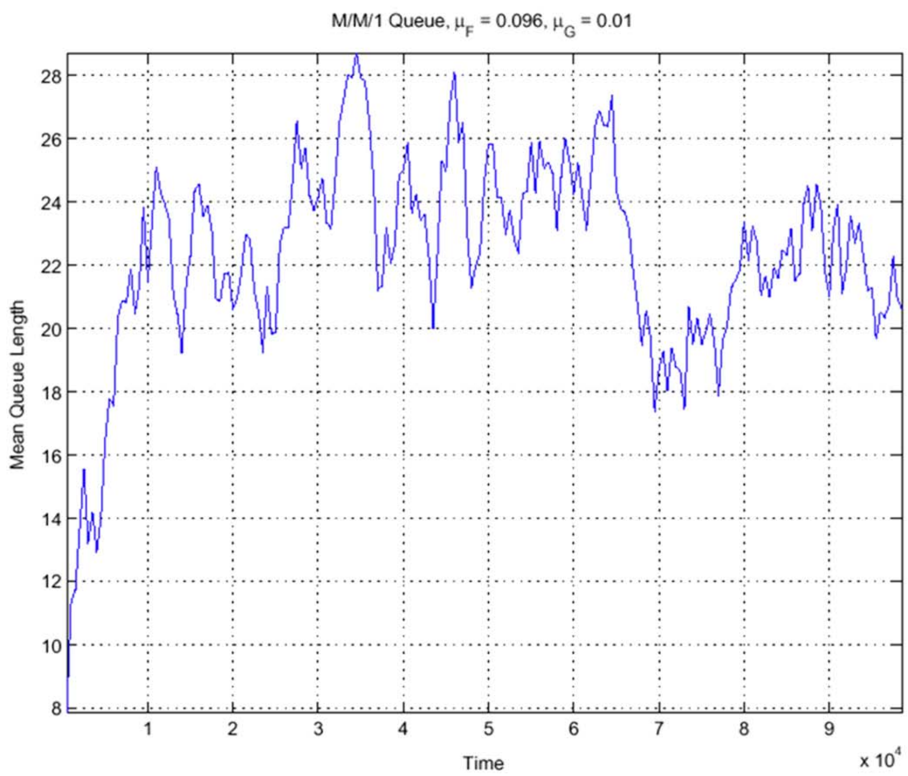
N customers arrive (at the same time) at Joe's shop and download a file. We want to estimate the time T it takes to serve the N customers

Simulation 2: Delay at a server

An information server handles requests according to some scheduling policy. We want to estimate the time it takes to serve one request

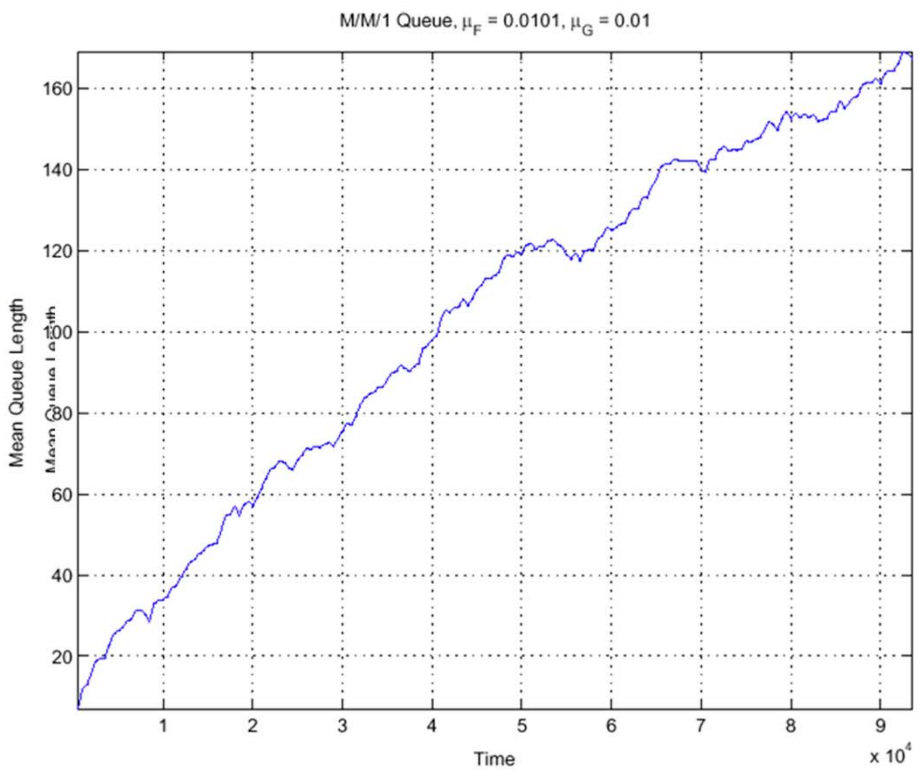
Two Simulation Runns
What is the difference ?

Scenario 1



Average service time = 96 ms

Scenario 2



Average service time = 101 ms

Request arrival rate = 10 /s

Definition of Stationarity

A property of a stochastic model X_t

Let X_t be a stochastic model that represents the state of the simulator at time t. It is stationary iff

for any $s > 0$

$(X_{t1},\ X_{t2}, \dots, X_{tn})$ has same distribution as $(X_{t1+s}, X_{t2+s}, \dots, X_{tn+s})$

i.e. simulation does not get “old”

Classical Cases

Markov models

Definition: State X_t is sufficient to draw the future of the simulation -- Common case for all simulations

For a Markov model, over a discrete state space

If you run the simulation long enough it will either walk to infinity (unstable) or converge to a stationary regime

Ex: queue with $\rho > 1$: unstable

queue with $\rho < 1$: becomes stationary after transient

If the state space is strongly connected (any state can be reached from any state) then there is 0 or 1 stationary regime

Ex: queue

Else, there may be several distinct stationary regimes

Ex: system with failure modes

Stationarity and Transience

Knowing whether a model has a stationary regime is sometimes a hard problem

- We will see important models where this is solved
- Ex: some queuing systems
  Ex: time series models

Reasoning about your system may give you indications

- Do you expect growth ?
- Do you expect seasonality ?

Once you believe your model is stationary, you should handle transients in order to eliminate the impact of initial conditions

- Remove (how ? Look at your output and guess)
- Sometimes it is possible to avoid transients at all (perfect simulation – see later “Importance of the View Point”)

Typical Reasons For Non Stationarity

Obvious dependency on time

- Seasonality, growth
- Can be ignored at small time scale (minute)

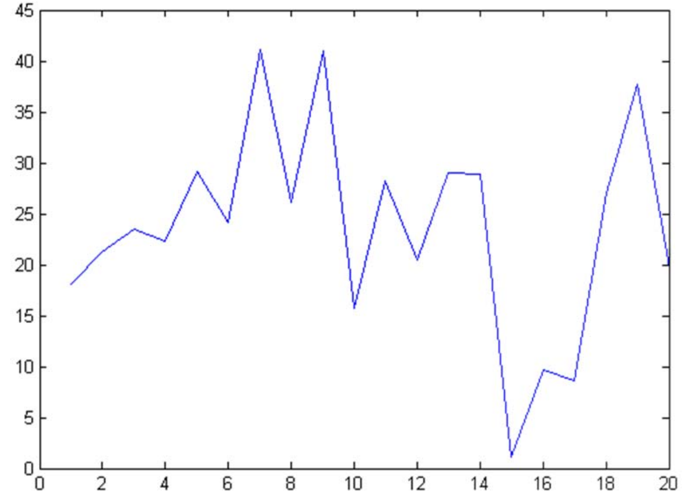
Non Stability: Explosion

- Queue with utilization factor >1

Non Stability: Freezing Simulation

- System becomes slower with time (aging)
- Typically because there are rare events of large impact (« Kings »)
- The longer the simulation, the larger the largest king
- We’ll come back to this in the chapter « Importance of the View Point »

The state of a simulation at time $t = 1, 2, 3, \ldots$ is drawn at random from a distribution F(). Is this a stationary simulation ?



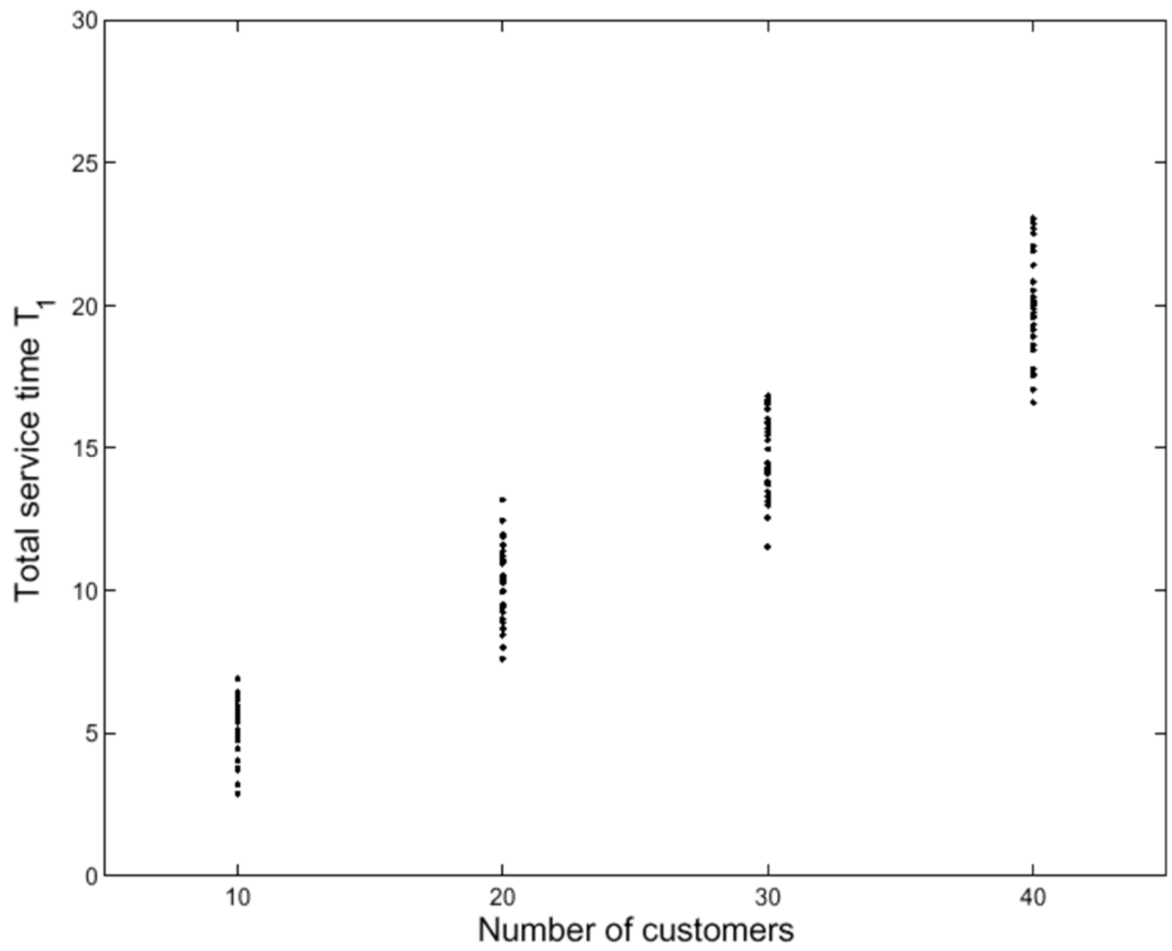
X_t is a sample drawn from the distribution $N(23, 100)$

A. Yes
B. It depends on the distribution $F(\quad)$
C. It depends if the simulation terminates or not
D. No
E. I don’t know

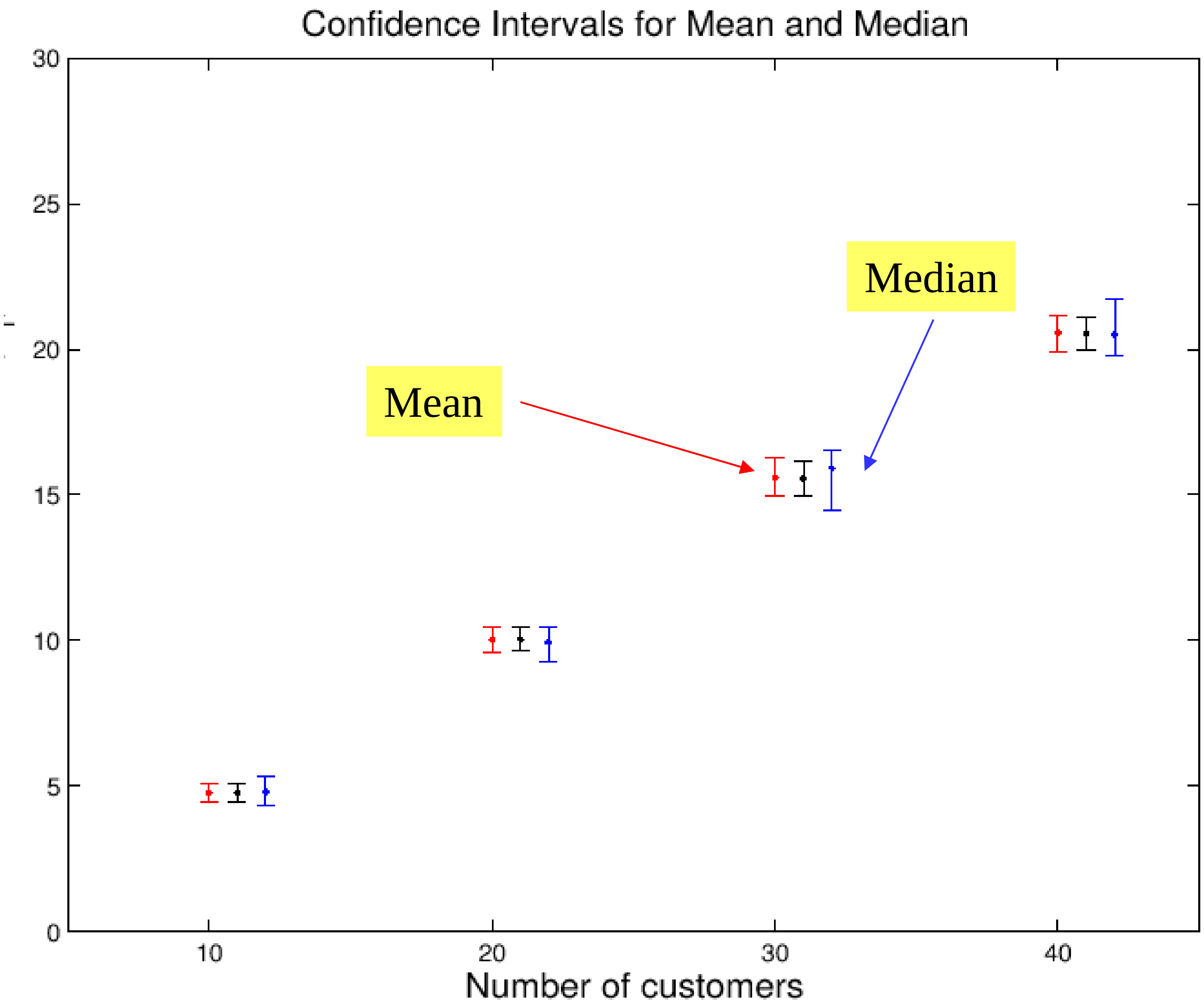
2 Accuracy of Simulation Output

A stochastic simulation produces a random output, we need confidence intervals

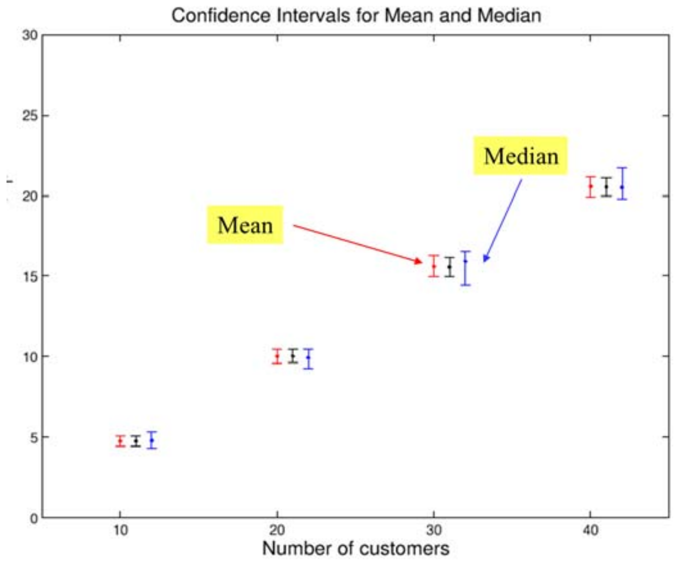
30 runs of Joe's shop: estimate the time to serve a batch of N customers



Confidence Intervals



When computing these confidence intervals we need to make sure that…



A. … the simulation runs are independent
B. … the output is normally distributed
C. Both
D. None
E. I don’t know

3 Random Number Generator

A stochastic simulation does not use truly random numbers but use pseudo-random numbers

The Random Number Generator produces a random number ~ U(0,1) based on a chaotic sequence

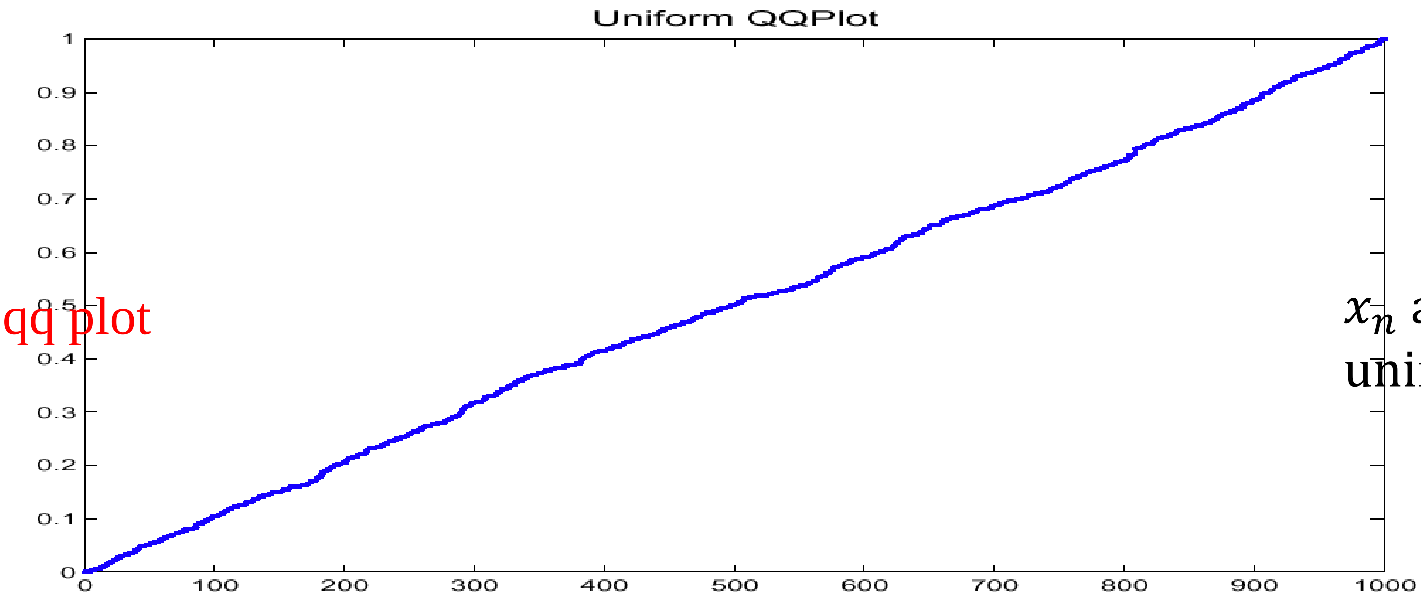
Example (obsolete but commonly used, eg in ns2:)

EXAMPLE 6.8: LINEAR CONGRUENCE. A widespread generator (for example the default in ns2) has $a = 16'807$ and $m = 2^{31} - 1$. The sequence is $x_n = \frac{sa^n \bmod m}{m}$ where s is the seed. m is a prime number, and the smallest exponent h such that $a^h = 1 \bmod m$ is $m - 1$. It follows that for any value of the seed s, the period of x_n is exactly $m - 1$.Figure 6.5 shows that the sequence x_n indeed looks random.

Let us check if output looks random

The Linear Congruential Generator of ns2

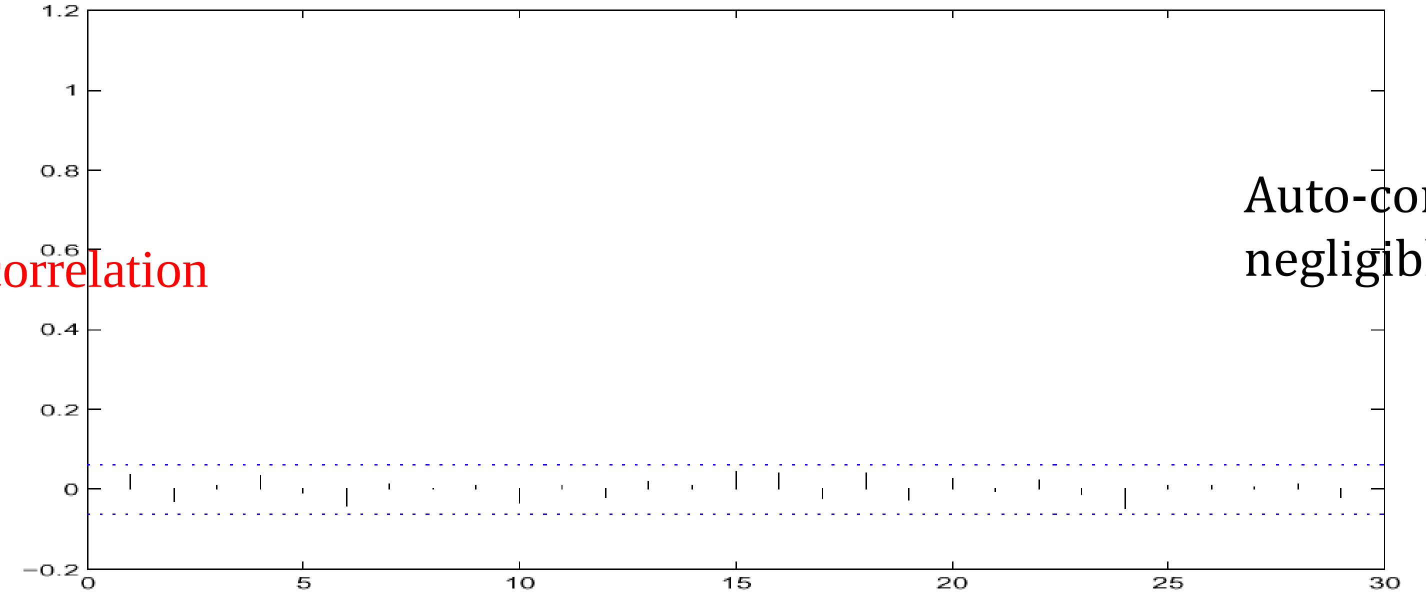
uniform qq plot



x_n appears to be uniform

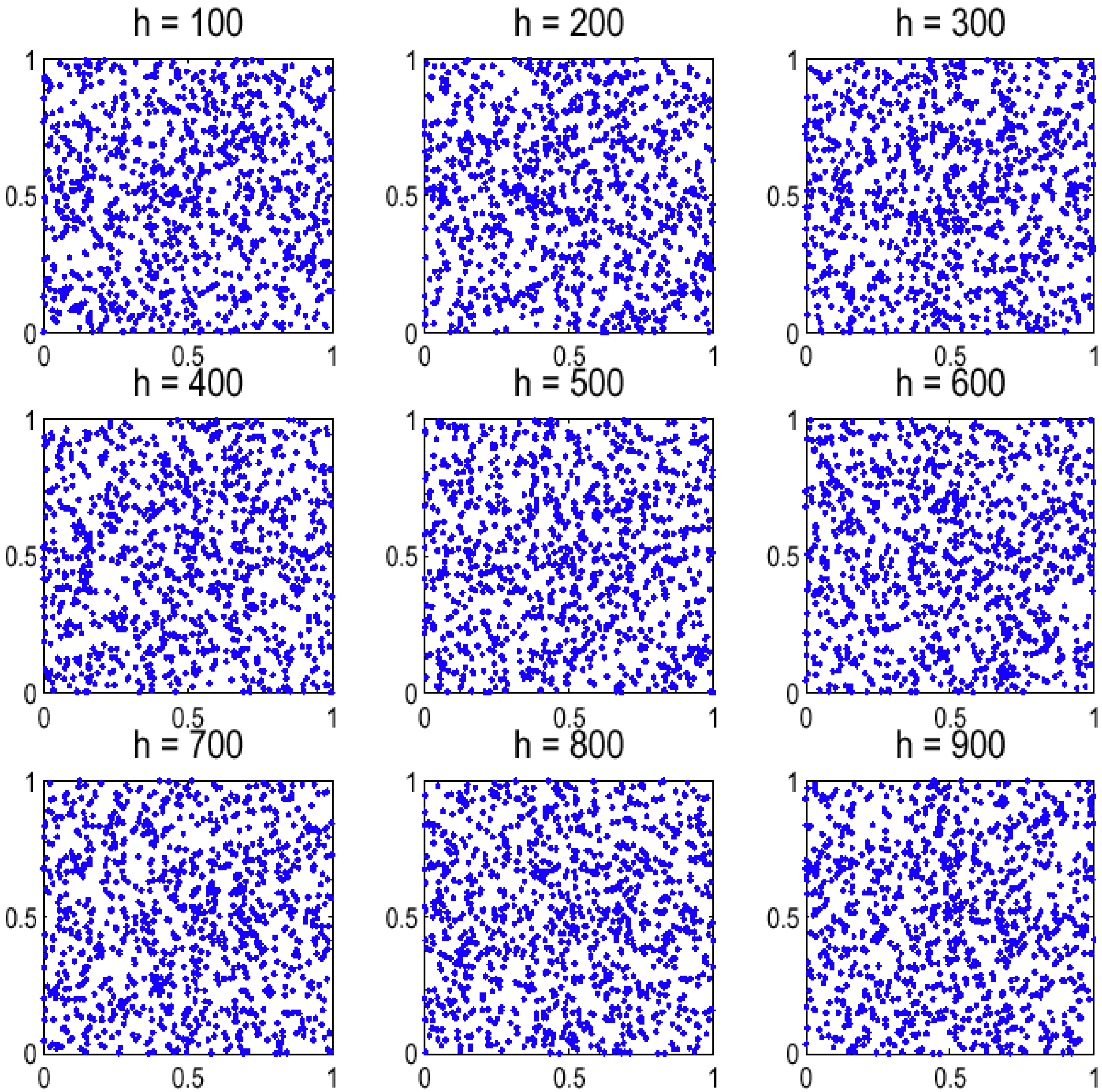
(a)

autocorrelation



Auto-correlation is negligible

Lag Diagram (x_n, x_{n+h}), 1000 points



x_n appears to be independent of x_{n+h}

Period of RNG

This RNG appears to produce an iid, uniform output

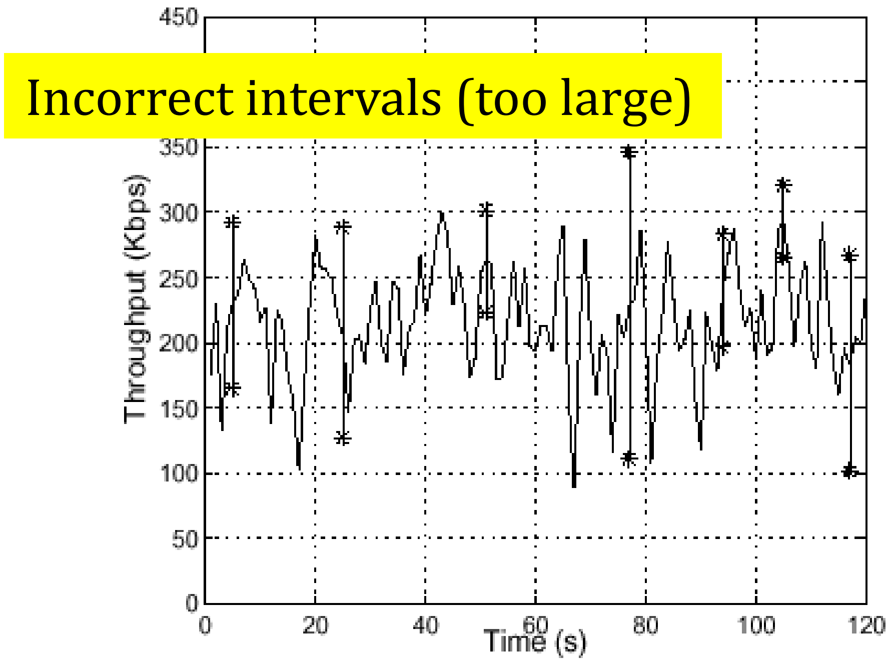
However, it is in fact periodic

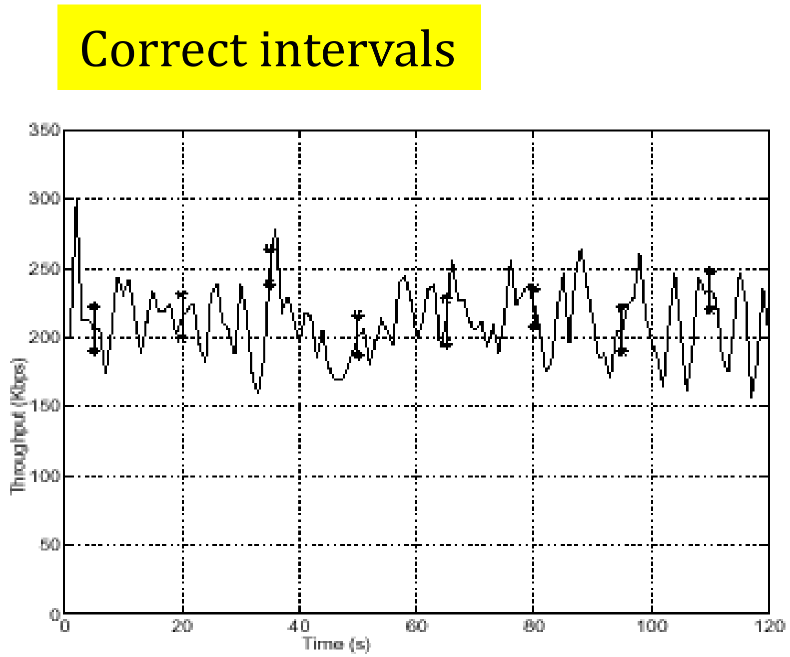
- Any sequence generated from a deterministic algorithm in a given computer must be periodic (because the number of states is finite)

But we can require that the period should be much larger than maximum number of uses

- The ns2 simulator has period $\approx 2 \cdot 10^9$, which is too small
- The “Mersenne twister” (matlab’s default) has period $2^{19937} - 1 \approx 10^{6000}$

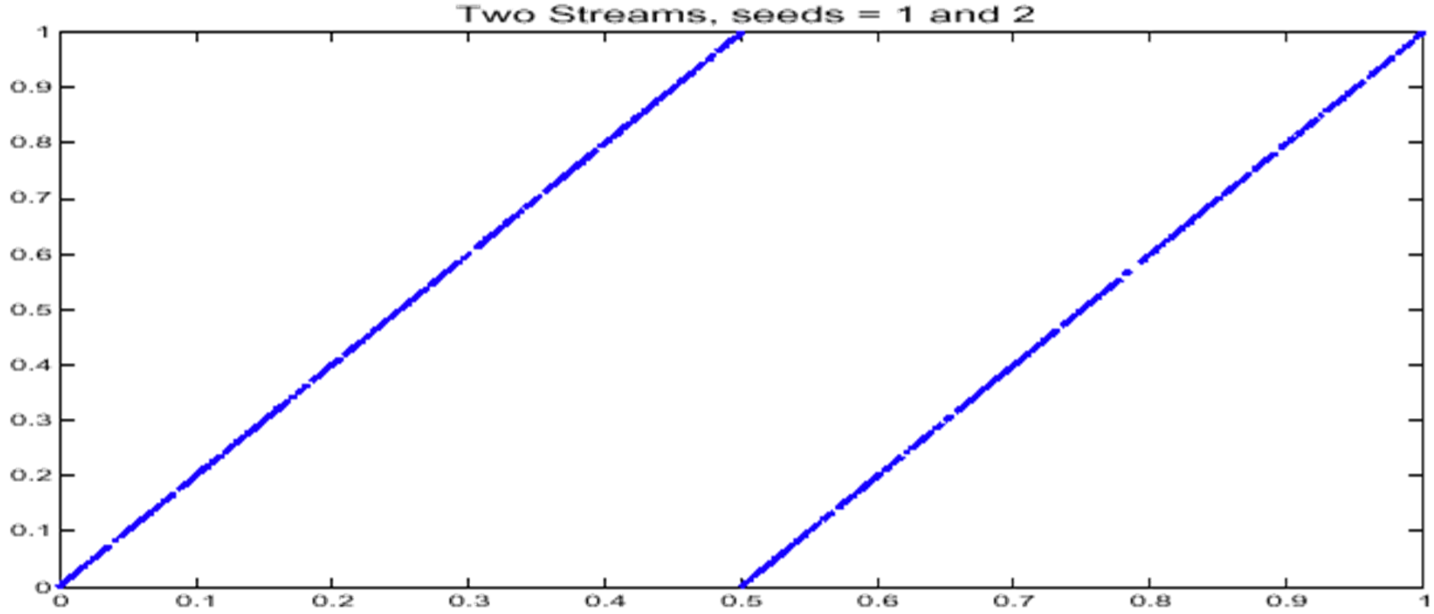
Example when the period is too small



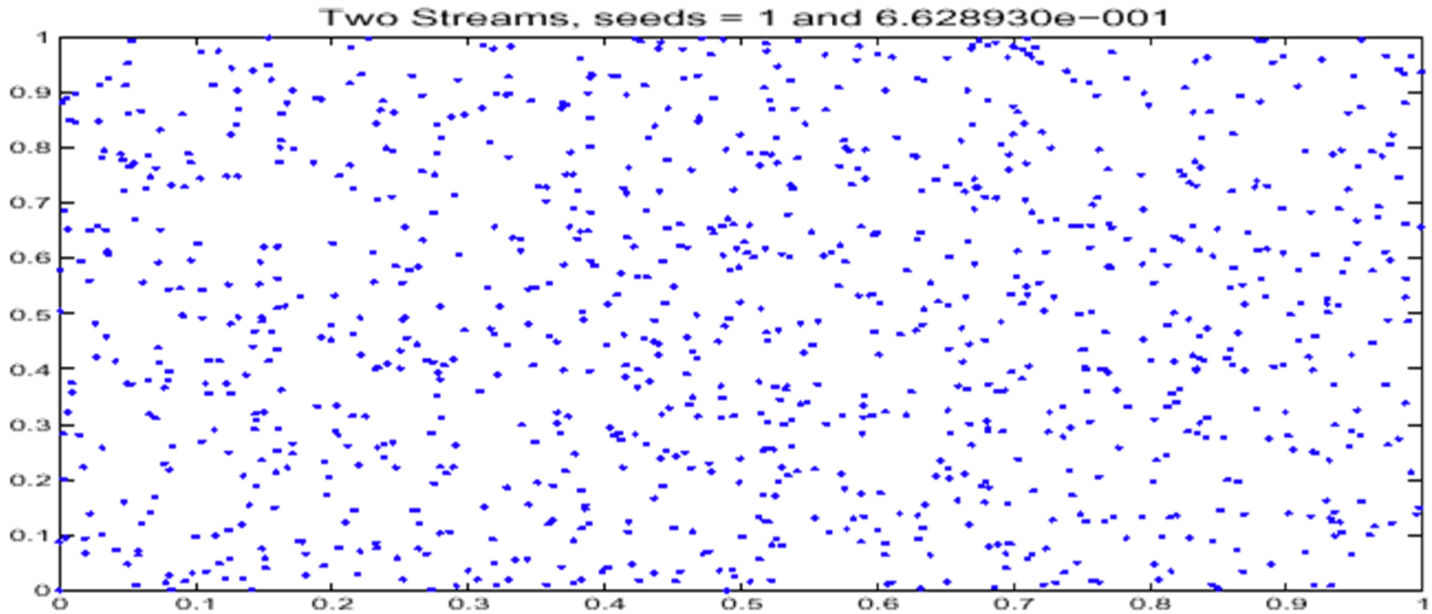


(a) Linear Congruence with $a = 16'807$ and $m = 2^{31} - 1$ (b) L'Ecuyer"'s generator[LecuyerSimConf-01]

(x_n, x'_n) for two parallel streams with the ns2 RNG



(a)



Using RNGs

Be careful to have a RNG that has a period orders of magnitude larger than what you will ever use in the simulation – choose your RNG carefully

The RNG must have a period several orders of magnitude larger than the maximum number of times you will call it

The RNG uses a seed -- for independent replications you need independent seeds; sequential runs usually are sufficient. If you use parallel runs, you must find an way to obtain truly independent seeds

- Eg based on clock
- Or by using an RNG that supports parallel streams

4 Sampling From A Distribution

Problem is:

Given a distribution $F()$, and a RNG, produce some sample X that is drawn from this distribution

A common task in simulation

Matlab does it for us most of the time, but not always

Two generic methods

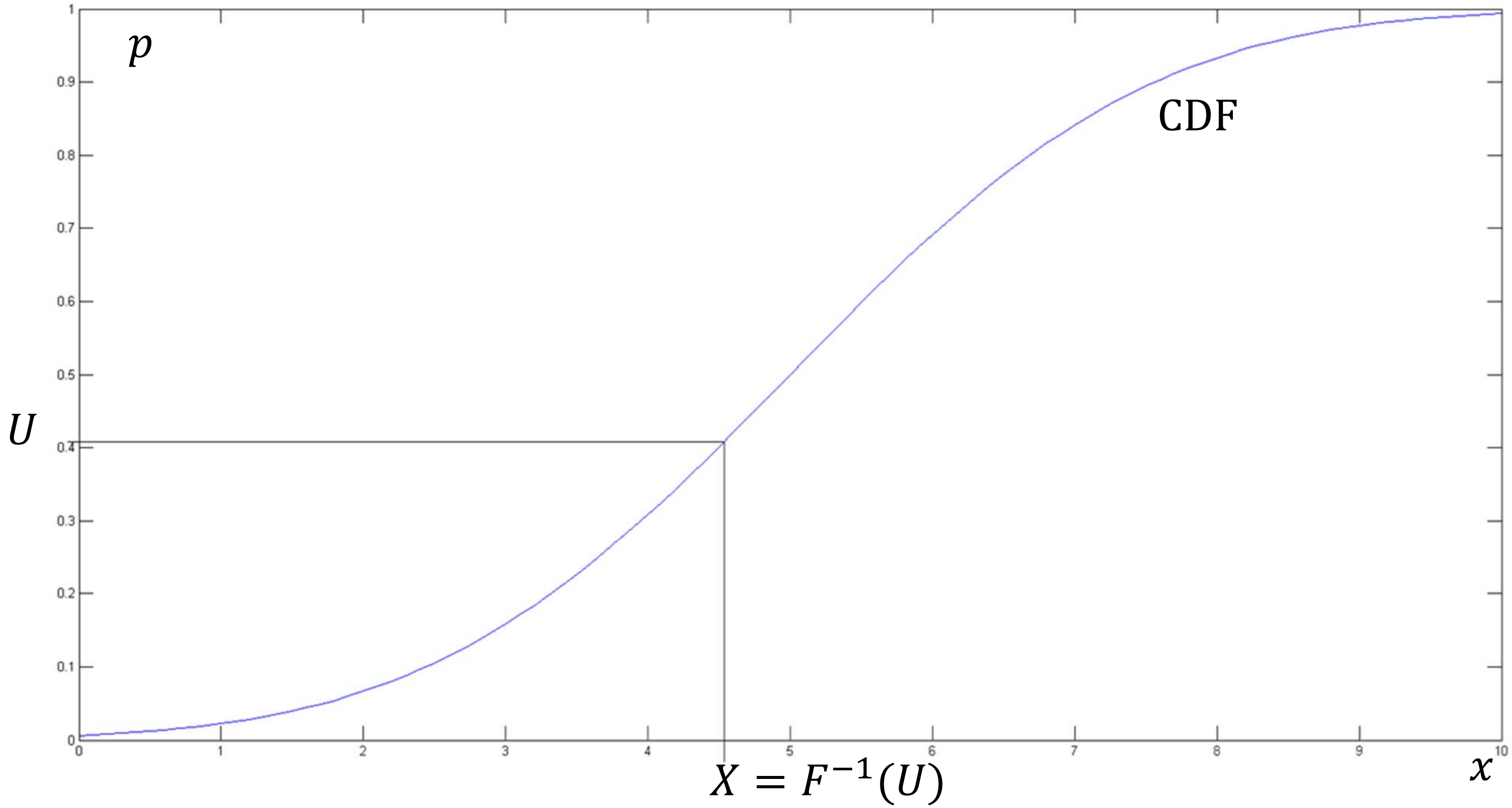
CDF inversion

Rejection sampling

CDF Inversion for distribution with CDF $F()$

Applies to real or integer valued random variable

For a continuous distribution: draw $U \sim U(0,1)$ using RNG; then $X = F^{-1}(U)$ is drawn from the distribution with CDF $F()$



Proof

U is uniformly distributed: $P(U \leq u) = u$ for $u \in [0,1]$

We define X by $X = \mathrm{F}^{-1}(U)$, i.e. $U = F(X)$
Let us compute the CDF of X

$(X \leq x) \Leftrightarrow \big(F(X) \leq F(x)\big)$ because F is monotonic increasing

Therefore

$$P(X \leq x) = P\big(F(X) \leq F(x)\big) = F(x)$$

EXAMPLE 7.10: EXPONENTIAL RANDOM VARIABLE. The CDF of the *exponential distribution* with parameter λ is $F(x) = 1 - e^{-\lambda x}$. The pseudo-inverse is obtained by solving the equation

$$1 - e^{-\lambda x} = p$$

where x is the unknown. The solution is $x = -\frac{\ln(1-p)}{\lambda}$. Thus a sample X of the exponential distribution is obtained by letting $X = -\frac{\ln(1-U)}{\lambda}$, or, since U and $1-U$ have the same distribution:

$$X = -\frac{\ln(U)}{\lambda} \tag{7.9}$$

where U is the output of the random number generator.

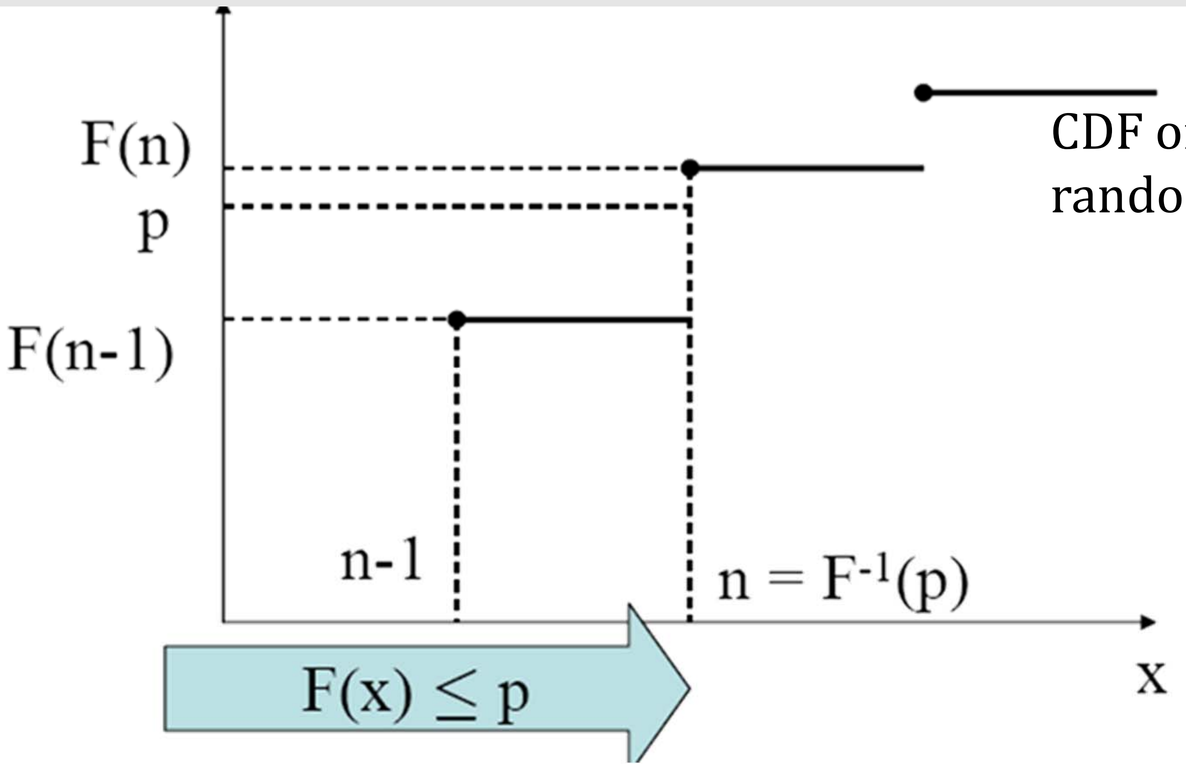
CDF Inversion, when $F()$ has jumps

Replace inverse by pseudo-inverse

THEOREM 6.1. *Let F be the CDF of a random variable X with values in $\mathbb{R}$. Define the pseudo-inverse, F^{-1} of F by*

$$F^{-1}(p) = \sup\{x : F(x) \leq p\}$$

Let U be a sample of a random variable with uniform distribution on $(0, 1)$; $F^{-1}(U)$ is a sample of X.



CDF of an integer-valued random variable

Sampling an integer valued RV

We want to draw a sample of the random integer $N \geq 0$ such that $p_k = P(N = k)$.

Let $F_n = \sum_{i=0}^{n} p_i$ for $n = 0,1,2,\ldots$ and $F_{-1} = 0$

1. Draw U using the RNG
2. Output is n such that $F_{n-1} \leq U < F_n$

EXAMPLE 7.11: GEOMETRIC RANDOM VARIABLE. Here X takes integer values $0, 1, 2, \ldots$. The *geometric distribution* with parameter θ satisfies $\mathbb{P}(X = k) = \theta(1-\theta)^k$, thus for $n \in \mathbb{N}$:

$$F(n) = \sum_{k=0}^{n} \theta(1-\theta)^k = 1 - (1-\theta)^{n+1}$$

by application of Eq.(7.10):

$$F^{-1}(p) = n \Leftrightarrow n \leq \frac{\ln(1-p)}{\ln(1-\theta)} < n+1$$

hence

$$F^{-1}(p) = \left\lfloor \frac{\ln(1-p)}{\ln(1-\theta)} \right\rfloor$$

and, since U and $1-U$ have the same distribution, a sample X of the geometric distribution is

$$X = \left\lfloor \frac{\ln(U)}{\ln(1-\theta)} \right\rfloor \tag{7.11}$$

What does this function compute ?
myfun($\boldsymbol{p}$) $=$ if rand() $< p$ 0 else 1

A. The flip of a coin where 1 is obtained with proba $\boldsymbol{p}$ and 0 with proba 1−$\boldsymbol{p}$
B. The flip of a coin where 1 is obtained with proba $\mathbf{1}-\boldsymbol{p}$ and $\mathbf{0}$ with proba $\boldsymbol{p}$
C. A sample of a geometric random variable with mean $\frac{\mathbf{1}}{\boldsymbol{p}}$
D. A sample of a geometric random variable with mean $\frac{\mathbf{1}}{\mathbf{1}-\boldsymbol{p}}$
E. I don’t know

Which one is a uniform random point ?

A. A

B. B

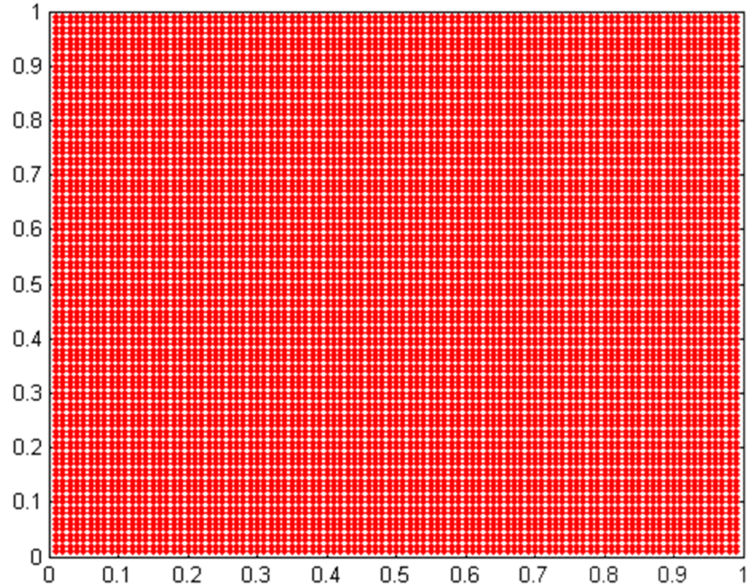
C. Both

D. None

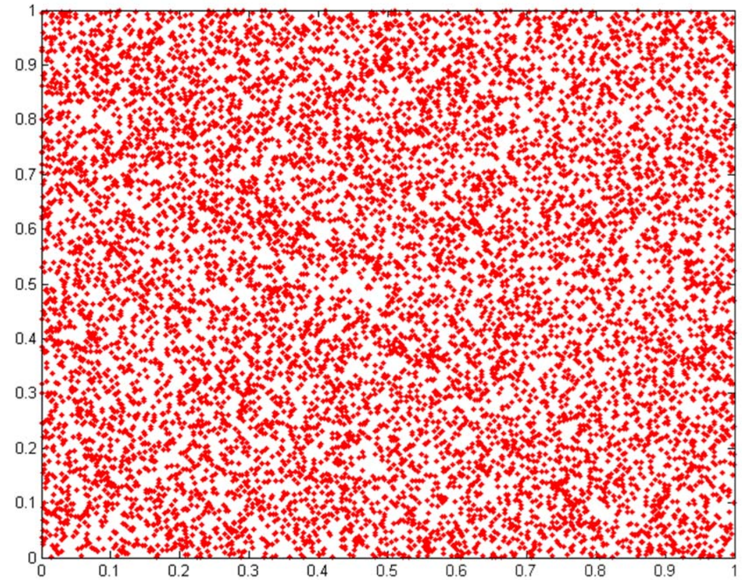
E. I don't know

10'000 samples

A



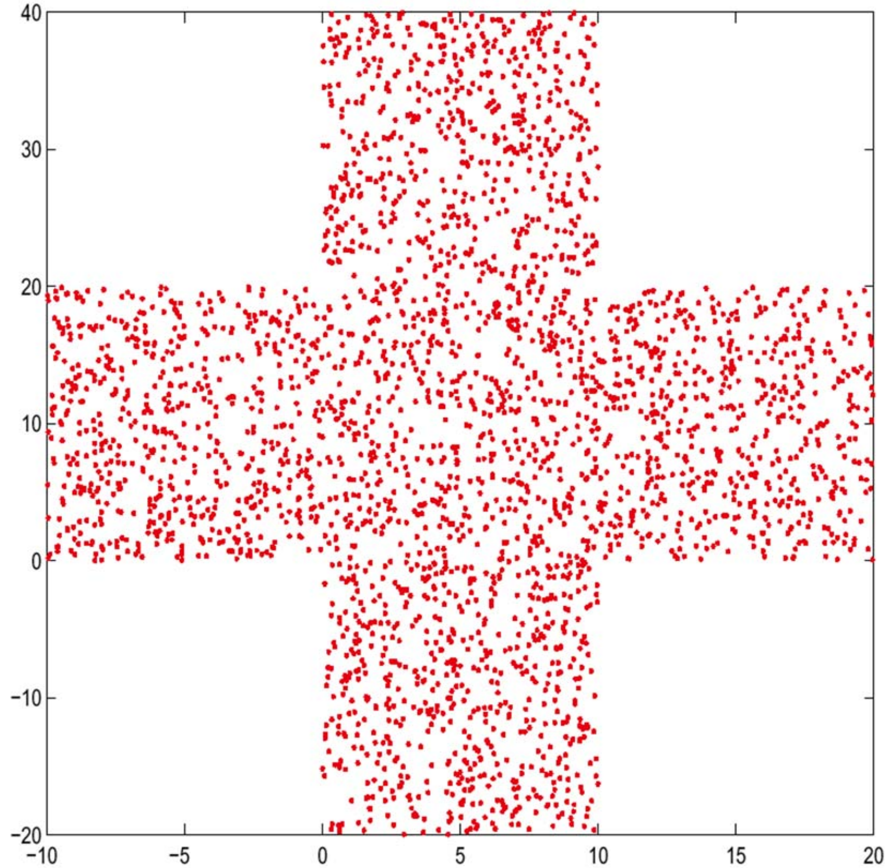
B



How do you sample a uniform point in a rectangle ?

We sample a point $\boldsymbol{M} = (\boldsymbol{X}, \boldsymbol{Y})$ uniformly in some arbitrary area. Are $\boldsymbol{X}$ and $\boldsymbol{Y}$ independent ?

A. Yes
B. It depends on the area
C. No
D. I don’t know



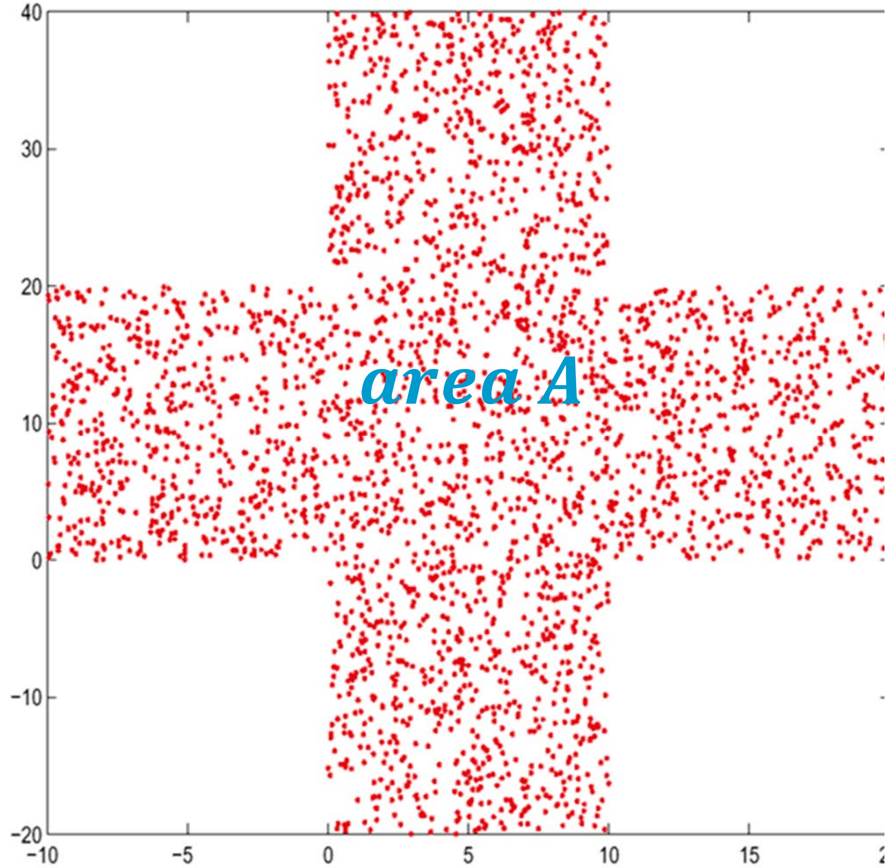
What is the PDF of the uniform distribution in A ?

A. $f_{X,Y}(x,y) = \begin{cases} 1 \text{ if } (x,y) \in A \\ 0 \text{ otherwise} \end{cases}$

B. $f_{X,Y}(x,y) = \begin{cases} xy \ \text{ if } (x,y) \in A \\ 0 \text{ otherwise} \end{cases}$

C. Something else

D. I don't know



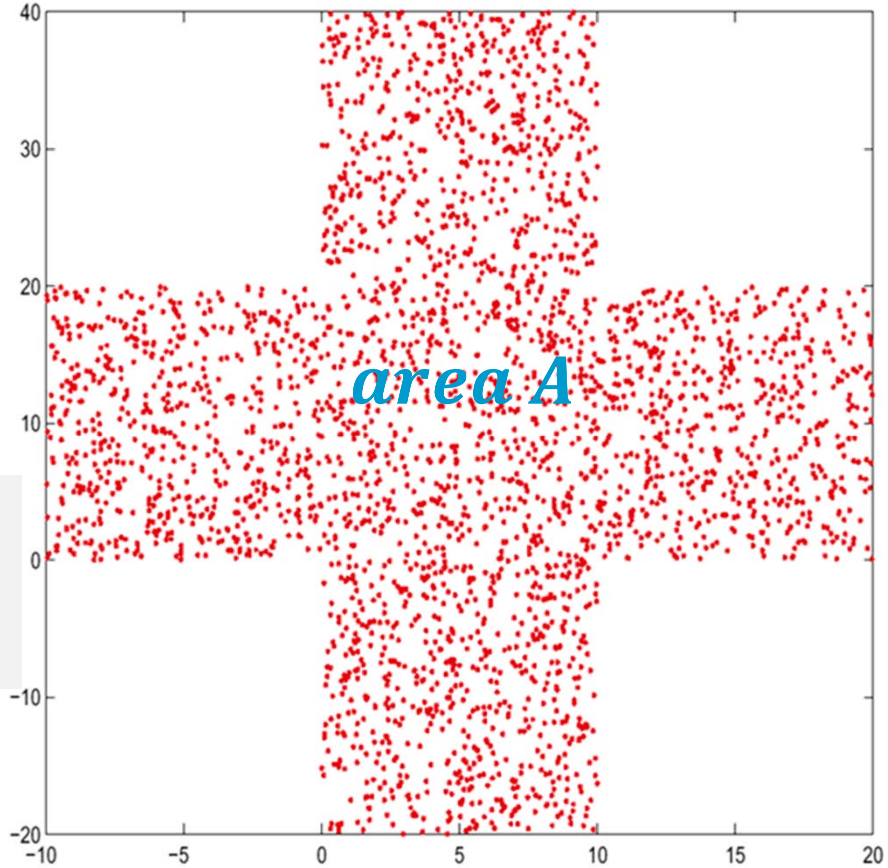
How can we sample a point uniformly distributed in this area $\boldsymbol{A}$?

A

draw $X \sim U(-10,20)$
if $X \in [0;10]$ draw $Y \sim U(-20;40]$
else draw $Y \sim U(0;20]$

B

repeat
draw $X \sim U(-10,20), Y \sim U(-20;40]$
until $(X,Y) \in A$



A. A
B. B
C. Both
D. None
E. I don't know

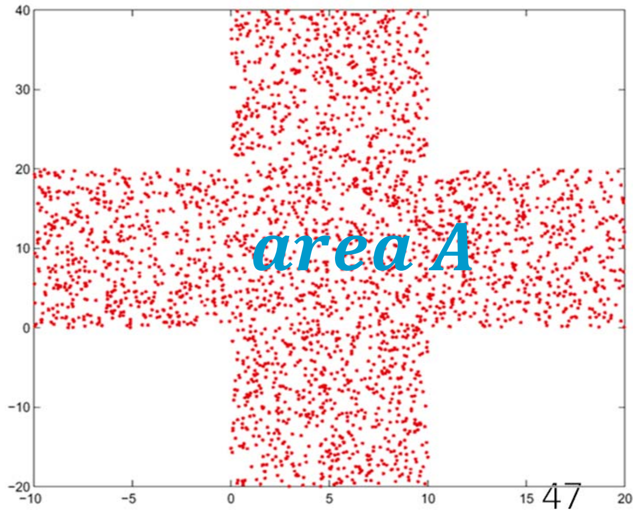
Rejection Sampling for Conditional Distribution

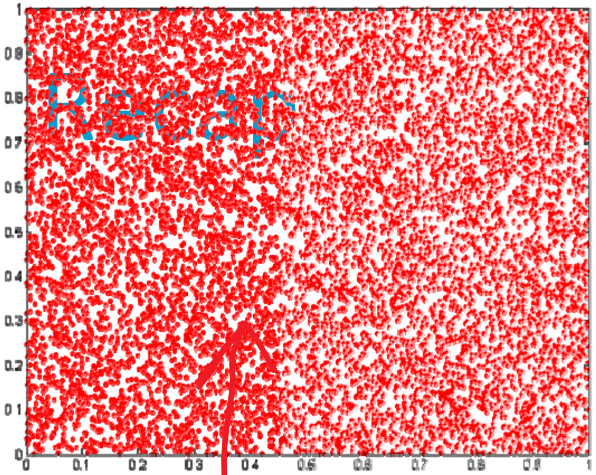
THEOREM 6.2 (Rejection Sampling for a Conditional Distribution). *Let X be a random variable in some space S such that the distribution of X is the conditional distribution of $\tilde{X}$ given that $\tilde{Y} \in \mathcal{A}$, where $(\tilde{X}, \tilde{Y})$ is a random variable in $S \times S'$ and $\mathcal{A}$ is a measurable subset of S.*
A sample of X is obtained by the following algorithm:

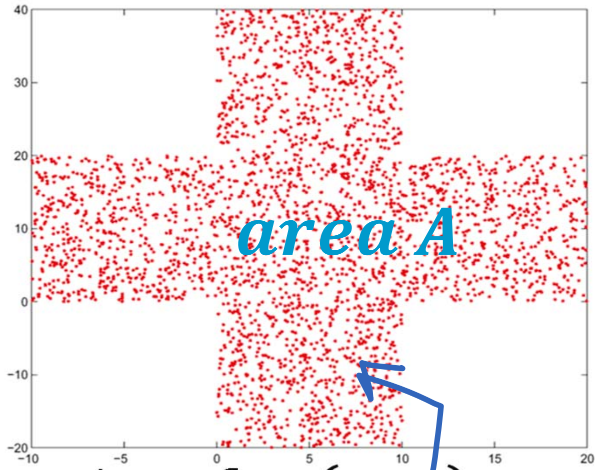
do
  draw a sample of $(\tilde{X}, \tilde{Y})$
until $\tilde{Y} \in \mathcal{A}$
return$(\tilde{X})$

The expected number of iterations of the algorithm is $\frac{1}{\mathbb{P}(\tilde{Y} \in \mathcal{A})}$.

Notation in this theorem	Previous example
$\tilde{X}$	$(X, Y) \sim$ Unif(rectangle)
$\tilde{Y}$	$(X, Y) \sim$ Unif(rectangle)
X	$(X, Y) \sim$ Unif(A)







A uniformly distributed point in an area A has pdf equal to $f_{X,Y}(x,y) = 1_{(x,y)\in A} \times \frac{1}{\text{area}(A)}$

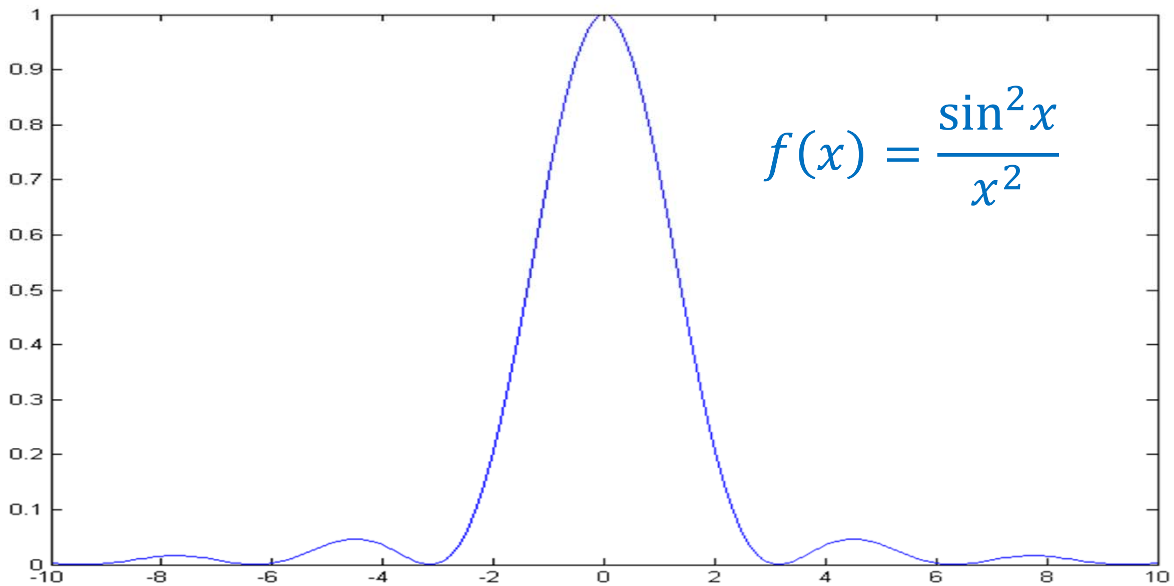
It can be sampled by *rejection sampling*

> sample a point (X,Y) uniform in a bounding box around A
> reject (X,Y) if not in A
> until $(X,Y) \in A$
> return (X,Y)

Note: the distribution *here* is not the same as *there*

What does this algorithm compute ?

1. sample X uniformly in $[-10,10]$;
2. compute $q = f(X)$
3. with probability q
   return (X)
   with probability $1-q$
   goto 1



A. A sample of the random variable with pdf $= Kf(x)$ for some constant K

B. A sample of the random variable with pdf $= K\,|x| f(x)$ for some constant K

C. A sample of the random variable with pdf $= K\,|1-x| f(x)$ for some constant K

D. Nothing, it never terminates

E. None of the above

F. I don't know

Rejection Sampling for General Distributions

THEOREM 6.3 (Rejection Sampling for Distribution with Density). *Consider two random variables X, Y with values in the same space, that both have densities. Assume that:*

- *we know a method to draw a sample of X*
- *the density of Y is known up to a normalization constant K: $f_Y(y) = K f_Y^n(y)$, where f_Y^n is a known function*
- *there exist some $c > 0$ such that*

$$\frac{f_Y^n(x)}{f_X(x)} \leq c$$

A sample of Y is obtained by the following algorithm:

do
draw independent samples of X and U, where $U \sim$Unif$(0, c)$
until $U \leq \frac{f_Y^n(X)}{f_X(X)}$
return(X)

The expected number of iterations of the algorithm is $\frac{c}{K}$.

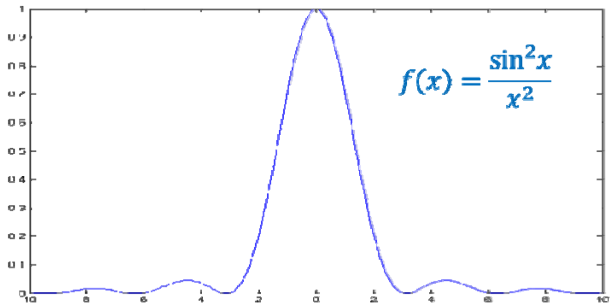
some $c > 0$ such that

$$\frac{f_Y^n(x)}{f_X(x)} \leq c$$

obtained by the following algorithm:

do
draw independent samples of X and U, where $U \sim Unif(0, c)$
until $U \leq \frac{f_Y^n(X)}{f_X(X)}$
return(X)

Example: use rejection sampling to draw a sample of $\boldsymbol{Y}$ with pdf $\boldsymbol{f_Y(y) = K\left(\frac{\sin(y)}{y}\right)^2 1_{\{-a \leq y \leq a\}}}$



$X \sim \text{Unif}(-a, a)$, $f_X(x) = \frac{1}{2a}$ for $-a \leq x \leq a$

$f_Y^n(y) = \left(\frac{\sin(y)}{y}\right)^2 1_{\{-a \leq y \leq a\}}$ $\frac{f^n(x)}{f_X(x)} = \frac{f^n(x)}{\frac{1}{2a}} \leq 2a$ we take $c = 2a$

The rejection sampling algorithm is:

1. Sample $X \sim \text{Unif}(-a, a)$ and $U \sim \text{Unif}(0,2a)$
2. If $U \leq \frac{f_X^n(X)}{1/2a}$ return(X) else goto 1

Example: use rejection sampling to draw a sample of $\boldsymbol{Y}$ with pdf $\boldsymbol{f_Y(y)} = \boldsymbol{K\left(\frac{\sin(y)}{y}\right)^2 1_{\{-a \le y \le a\}}}$

some $c > 0$ such that

$$\frac{f_Y^n(x)}{f_X(x)} \le c$$

obtained by the following algorithm:

do

draw independent samples of X and U, where $U \sim Unif(0, c)$

until $U \le \frac{f_Y^n(X)}{f_X(X)}$

return(X)

The rejection sampling algorithm is:

1. Sample $X \sim \text{Unif}(-a, a)$ and $U \sim \text{Unif}(0,2a)$
2. If $U \le \frac{f_X^n(X)}{1/2a}$ return(X) else goto 1

Replace U by $V = \frac{U}{2a}$, we obtain the equivalent algorithm:

1. Sample $X \sim \text{Unif}(-a, a)$ and $V \sim \text{Unif}(0,1)$
2. If $V \le f_X^n(X)$ return(X) else goto 1

The rejection sampling algorithm is:

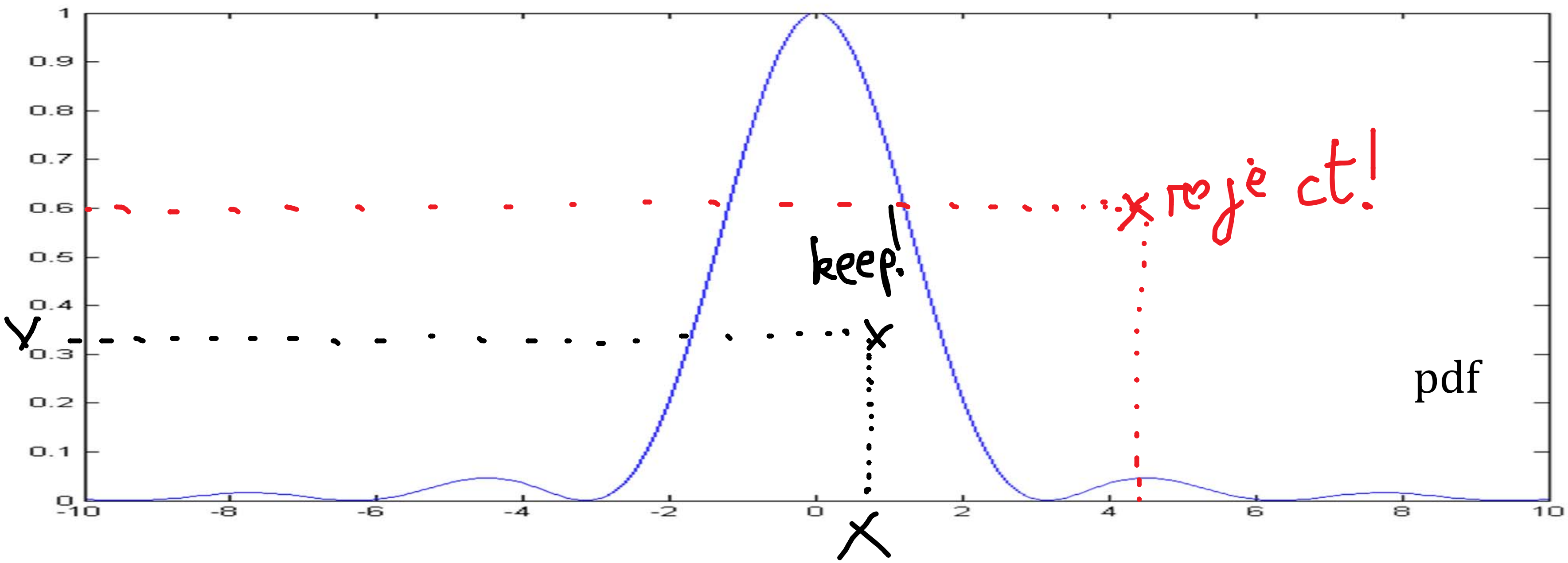
1. Sample $X \sim \text{Unif}(-a, a)$ and $U \sim \text{Unif}(0,2a)$
2. If $U \le \frac{f_X^n(X)}{1/2a}$ return(X) else goto 1

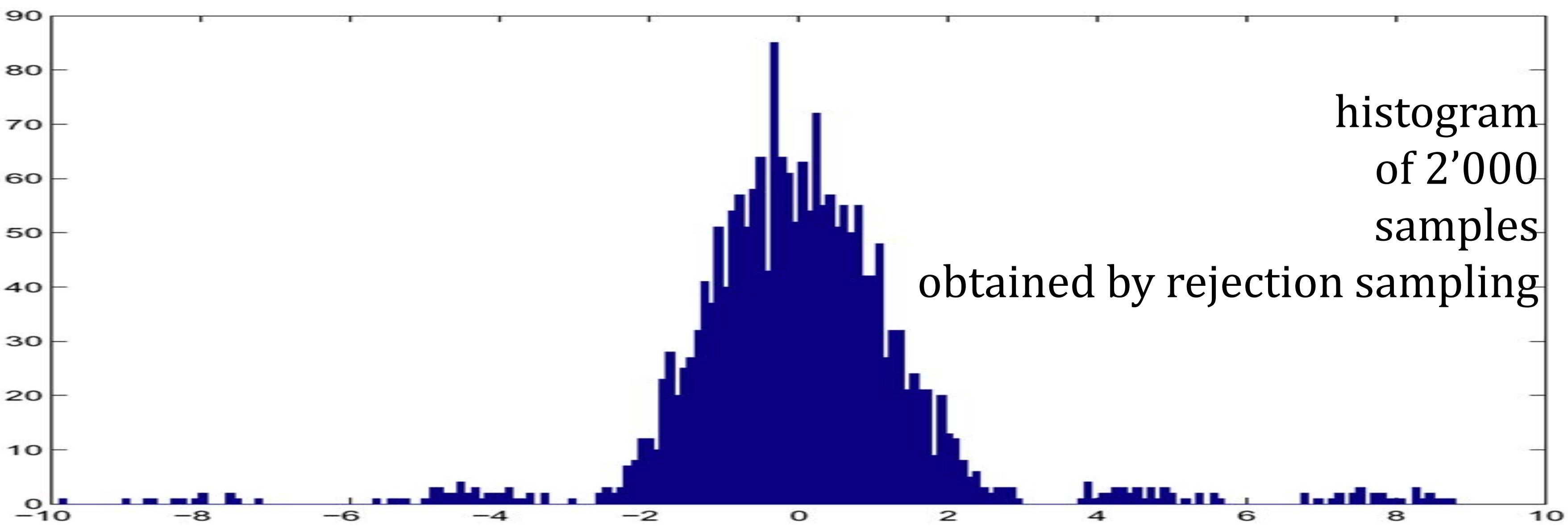
Replace U by $V = \frac{U}{2a}$, we obtain the equivalent algorithm:

1. Sample $X \sim \text{Unif}(-a, a)$ and $V \sim \text{Unif}(0,1)$
2. If $V \le f_X^n(X)$ return(X) else goto 1

Interpretation

1. Sample $X \sim \text{Unif}(-a, a)$ and $V \sim \text{Unif}(0,1)$
2. If $V \leq f_X^n(X)$ return(X) else goto 1





(a)

Arbitrary Distribution with Density Assume that we want a sample of Y, which takes values in the bounded interval $[a, b]$ and has a density $f_Y = K f_Y^n(y)$. Assume that $f_Y^n(y)$ (non normalized density) can easily be computed, but not the normalization constant K which is unknown. Also assume that we know an upper bound M on f_Y^n.

We take X uniformly distributed over $[a, b]$ and obtain the sampling method:

> **do**
>
> draw $X \sim$Unif(a, b) and $U \sim$Unif$(0, M)$
>
> **until** $U \leq f_Y^n(X)$
> **return**(X)

Note that we do *not* need to know the multiplicative constant K. For example, consider the distribution with density

$$f_Y(y) = K \frac{\sin^2(y)}{y^2} \mathbf{1}_{\{-a \leq y \leq a\}} \tag{6.13}$$

K is hard to compute, but a bound M on f_Y^n is easy to find ($M = 1$) (Figure 6.10).

We sampled 5'000 points from the distribution in $[\mathbf{0}\,;\mathbf{1}\,]^2$ with pdf $\boldsymbol{A}$, $\boldsymbol{B}$ or $\boldsymbol{C}$. Who's what ?

A. ABC

B. ACB

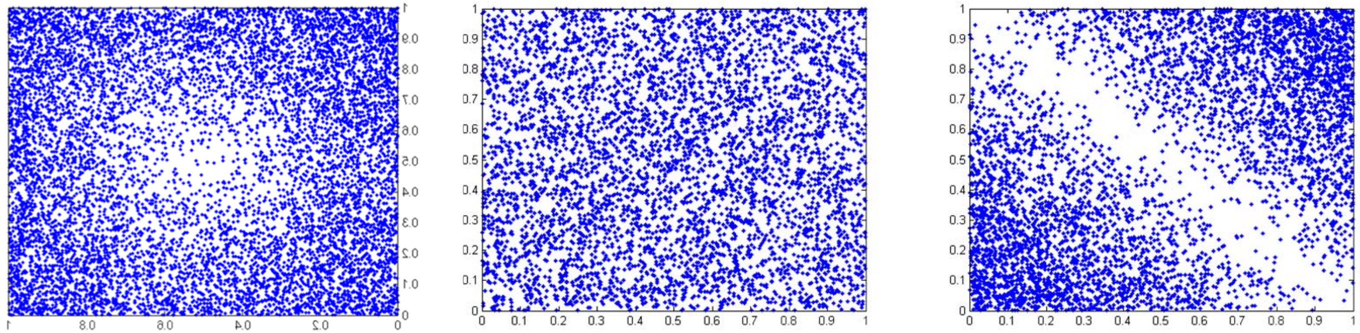
C. BAC

D. CAB

E. BCA

F. CBA

G. I don't know



$$A(x,y) = K_1 \times |x+y-1|$$
$$B(x,y) = K_2$$
$$C(x,y) = K_3 \times \sqrt{(x-0.5)^2 + (y-0.5)^2}$$

Which algorithm is a correct sampling of the distribution in $[\mathbf{0}\,;\mathbf{1}\,]^{\mathbf{2}}$ with pdf proportional to $\boldsymbol{C}$?

A. A

B. B

C. Both

D. None

E. I don't know

$$C(x,y) = \sqrt{(x-0.5)^2 + (y-0.5)^2}$$

A

1. Sample X and Y uniformly in $[0,1]$ and $U \sim \text{Unif}(1, 1/\sqrt{2})$;
2. If $U \leq C(X,Y)$ return(X,Y) else goto 1

B

1. Sample X and Y uniformly in $[0,1]$ and $U \sim \text{Unif}(0,1)$;
2. If $U \leq C(X,Y)$ return(X,Y) else goto 1

EXAMPLE 6.14: A STOCHASTIC GEOMETRY EXAMPLE. We want to sample the random vector (X_1, X_2) that takes values in the rectangle $[0,1] \times [0,1]$ and whose distribution has a density proportional to $|X_1 - X_2|$. We take $f_X =$ the uniform density over $[0,1] \times [0,1]$ and $f_Y^n(x_1, x_2) = |x_1 - x_2|$. An upper bound on the ratio $\frac{f_Y^n(x_1,x_2)}{f_X(x_1,x_2)}$ is 1. The sampling algorithm is thus:

do
  draw X_1, X_2 and U ~Unif$(0,1)$
until $U \leq |X_1 - X_2|$
return(X_1, X_2)

Figure 6.10 shows an example. Note that there is no need to know the normalizing constant to apply the sampling algorithm.

Another Sample from a Weird Distribution

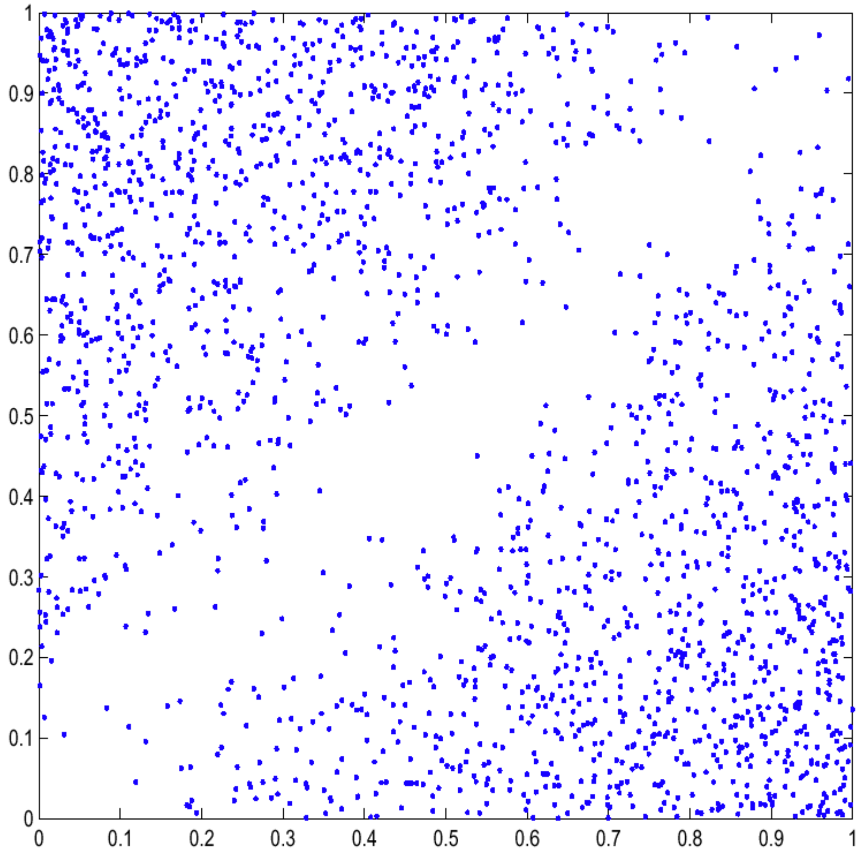


Figure 3.8: (a) Empirical histogram (bin size = 10) of 2000 samples of the distribution with density $f_X(x)$ proportional to $\frac{\sin^2(x)}{x^2} 1_{\{-a \leq y \leq a\}}$ with $a = 10$. (b) 2000 independent samples of the distribution on the rectangle with density $f_{X_1,X_2}(x_1, x_2)$ proportional to $|x_1 - x_2|$.

6.3 Ad-Hoc Methods

Optimized methods exist for some common distributions

Optimization = reduce computing time

If implemented in your tool, use them !

Example: simulating a normal distribution

Inversion method is not simple (no closed form for F^{-1})

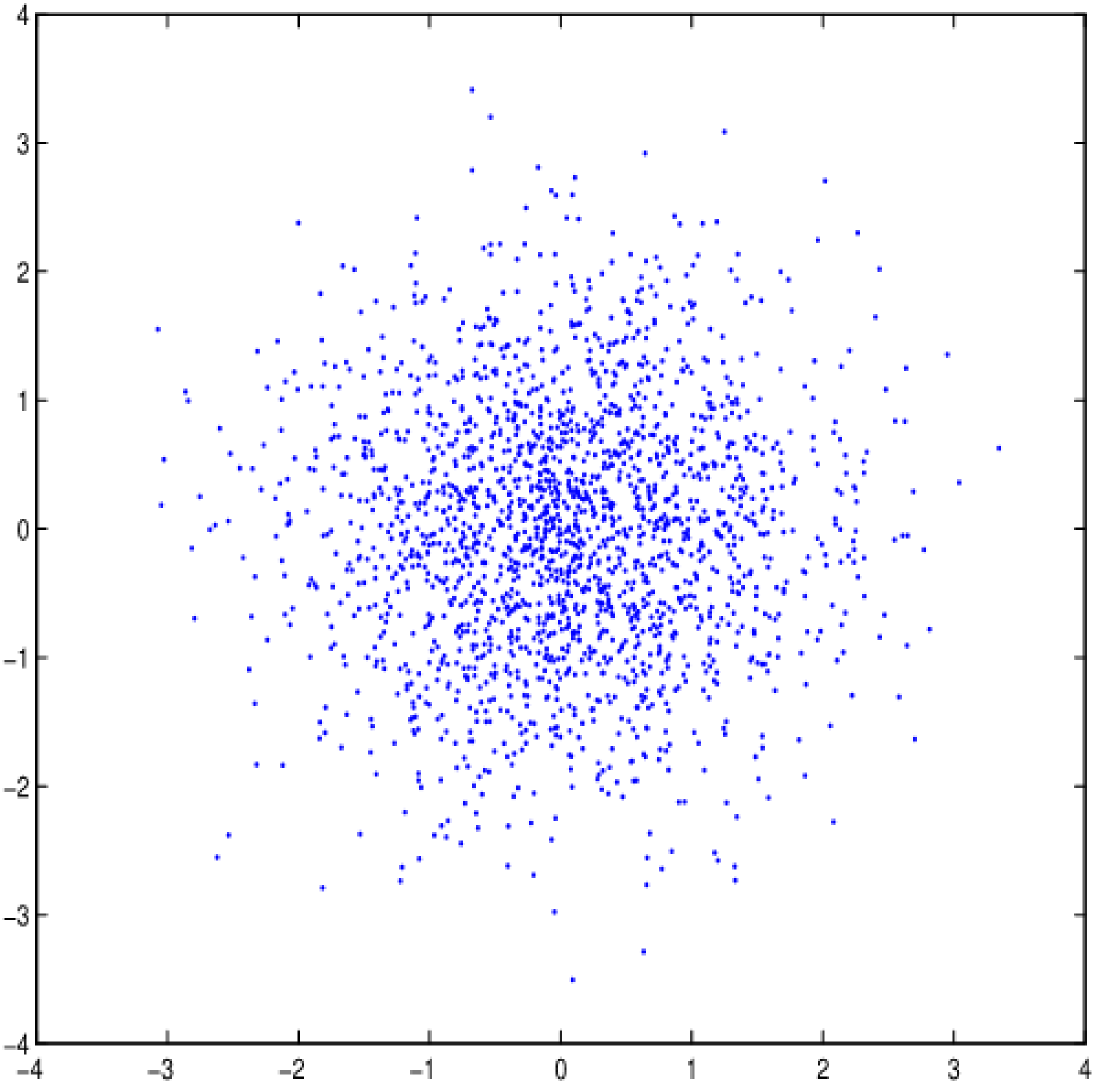
Rejection method is possible

But a more efficient method exists, for drawing jointly *2* independent normal RVs

There are also ad-hoc methods for n-dimensional normal distributions (gaussian vectors)

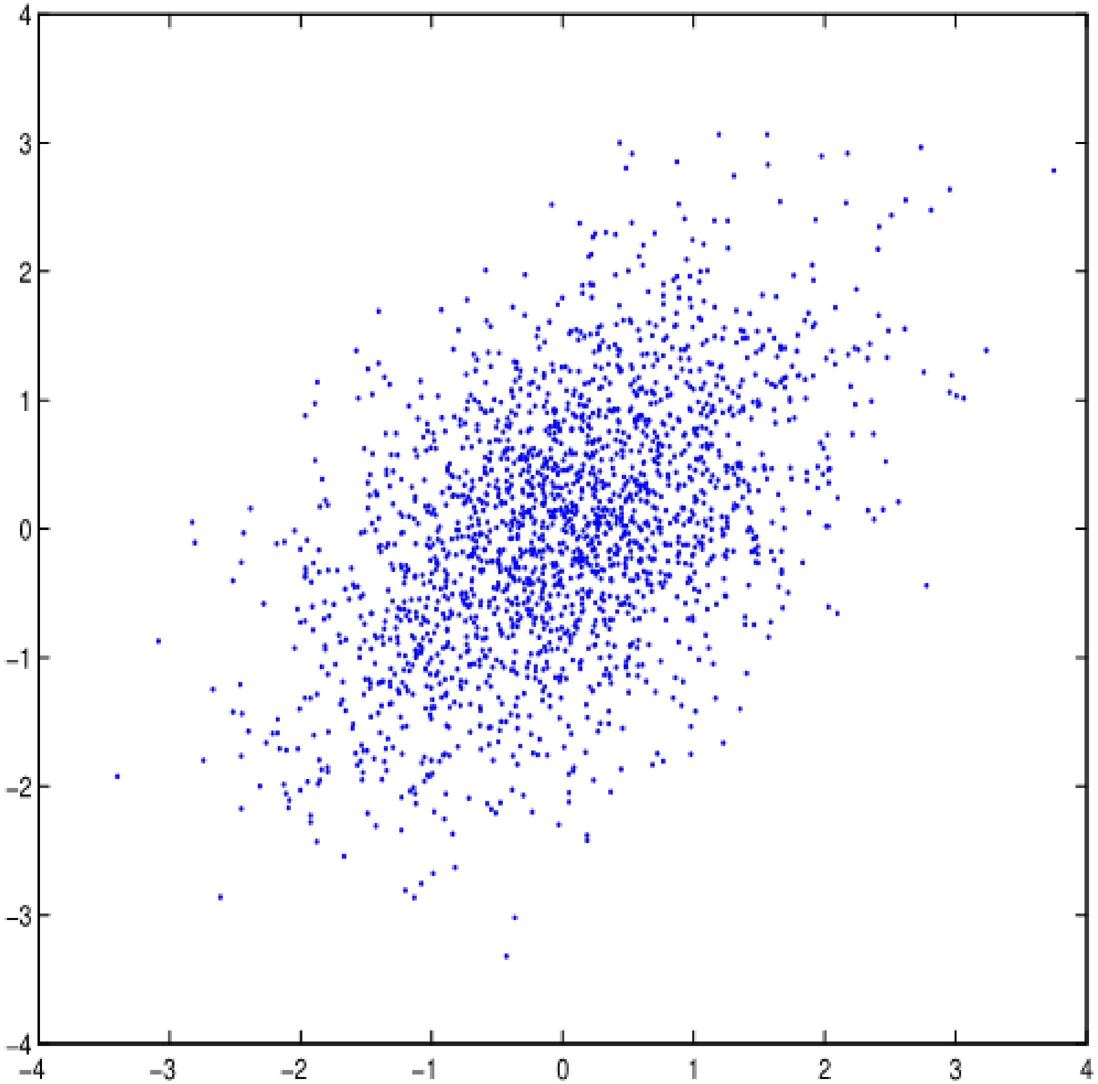
1’000 samples of the random Gaussian vector (X, Y) with zero mean and covariance matrix $\begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix}$

$\boldsymbol{X} \sim \boldsymbol{N}(\boldsymbol{0}, \boldsymbol{1}), \boldsymbol{Y} \sim \boldsymbol{N}(\boldsymbol{0}, \boldsymbol{1})$ and $\boldsymbol{X}, \boldsymbol{Y}$ are independent



1'000 samples of the random Gaussian vector (X, Y) with zero mean and covariance matrix $\begin{pmatrix} 1 & 0.5 \\ 0.5 & 1 \end{pmatrix}$

$\boldsymbol{X}, \boldsymbol{Y}$ are not independent



What is the distribution of $\boldsymbol{Y}$?

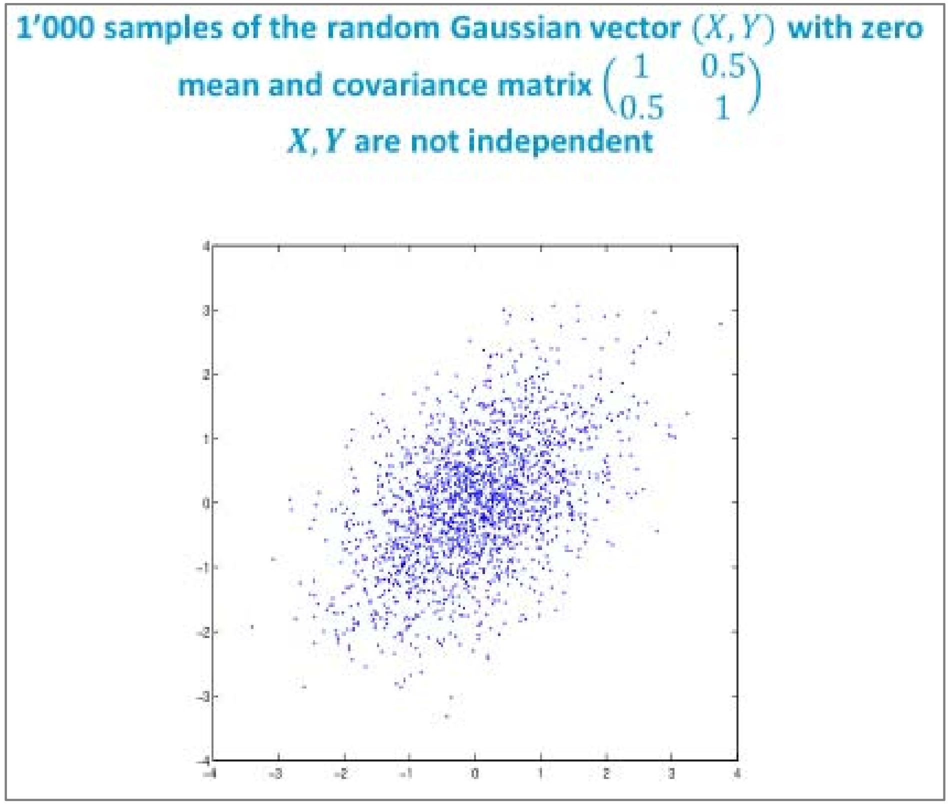
A. Normal with $\mu = 0, \sigma^2 = 1$

B. Normal with $\mu = 0, \sigma^2 = 0.5$

C. Normal but with other parameters

D. Non normal

E. I don’t know



4 Monte Carlo Simulation

A simple method to compute integrals and probabilities

Idea:

interpret the integral or probability β as $\beta = E(\varphi(X))$

simulate as many independent samples of X as you want

estimate the expectation by the mean

1. generate R replicates $X^r, r = 1 \dots R$
2. the Monte-Carlo estimate of β is

$$\hat{\beta} = \frac{1}{R} \sum_{r=1}^{R} \varphi(X_r)$$

3. compute a confidence interval for the mean (since β is the mean of the distribution of $\varphi(X)$)

Example: Compute $\beta = \int_0^1 \int_0^1 \int_0^1 \int_0^1 \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2}\, dx_1 dx_2 dy_1 dy_2$

i.e. the average distance between two points uniformly distributed in the unit square $[0\,;1]^2$. Say which one is a correct implementation of the Monte-Carlo method:

A

1. Draw R points M^r, N^r random uniform in $[0,1]^2$
2. $\hat{\beta} = \text{mean}(d(M^r, N^r), r = 1{:}R)$
3. $\hat{\sigma} = \text{std}(d(M^r, N^r), r = 1{:}R)$
4. $CI = \hat{\beta} \pm 1.96\hat{\sigma}$

B

$b = 0, s = 0$
do R times
  sample (X_1, Y_1, X_2, Y_2) iid ~ Unif (0,1)
  $b = b + \sqrt{(X_1 - X_2)^2 + (Y_1 - Y_2)^2}$
  $s = s + (X_1 - X_2)^2 + (Y_1 - Y_2)^2$
end
$\hat{\beta} = b/R$
$\hat{\sigma} = \sqrt{\frac{s}{R} - \hat{\beta}^2}$
$CI = \hat{\beta} \pm 1.96 \frac{\hat{\sigma}}{\sqrt{R}}$

A. A
B. B
C. Both
D. None
E. I don't know

Example

Compute $\beta =$
$$\int_0^1 \int_0^1 \int_0^1 \int_0^1 \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2}\, dx_1 dx_2 dy_1 dy_2$$
i.e. the average distance between two points uniformly distributed in the unit square $[0\,;1]^2$.

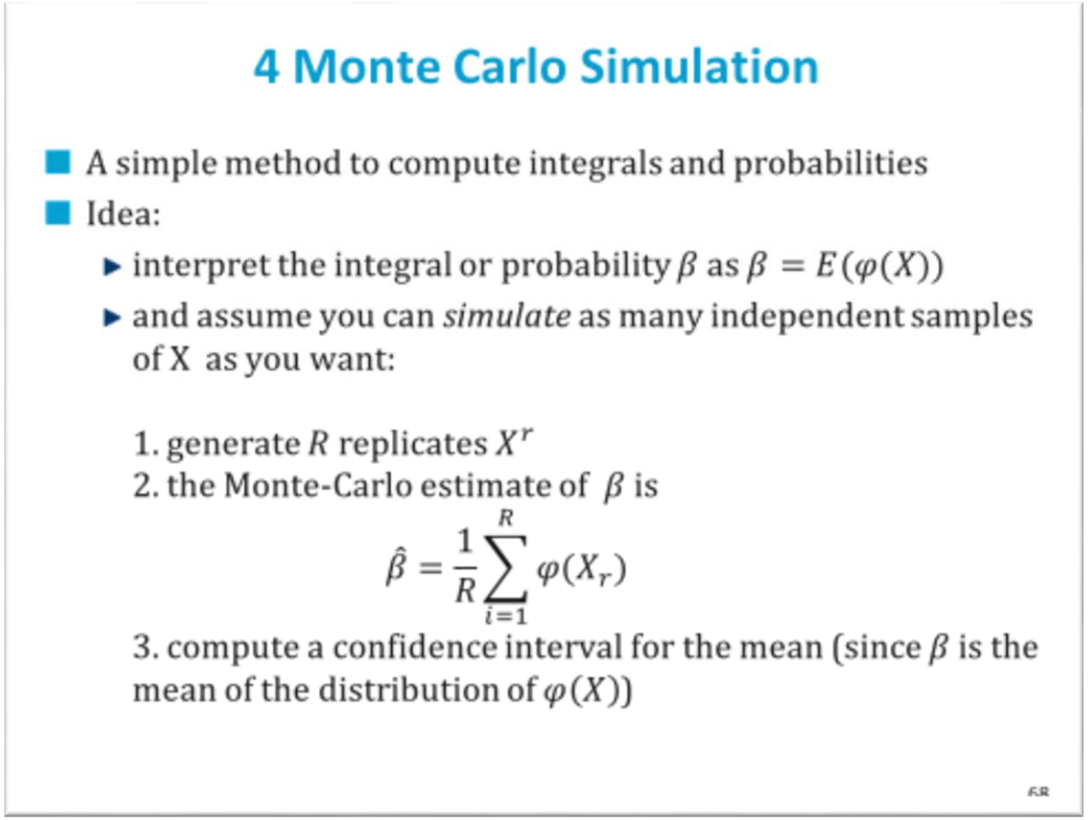
We obtain the following Monte Carlo estimates

R	$\hat{\beta}$	95% confidence interval	
1’000	0.5254	0.5102	0.5405
10’000	0.5227	0.5178	0.5276
100’000	0.5206	0.5190	0.5221
1’000’000	0.5216	0.5211	0.5221

(The answer is known and is equal to 0.5214 [Ghosh B. 1943, On the distribution of random distances in a rectangle])

Is it a good idea to use a confidence interval for the mean ?

A. Yes when R is large
B. Yes but it would be safer to use a CI for median
C. No because we don't know if the output is normally distributed
D. I don't know



4 Monte Carlo Simulation

- A simple method to compute integrals and probabilities
- Idea:
  - interpret the integral or probability β as $\beta = E(\varphi(X))$
  - and assume you can *simulate* as many independent samples of X as you want:

    1. generate R replicates X^r
    2. the Monte-Carlo estimate of β is

    $$\hat{\beta} = \frac{1}{R}\sum_{i=1}^{R} \varphi(X_r)$$

    3. compute a confidence interval for the mean (since β is the mean of the distribution of $\varphi(X)$)

Conclusion

Simulating well requires knowing the concepts of

- Transience
- Confidence intervals
- Sampling methods (Rejection sampling, CDF inversion)

Monte Carlo simulation is an efficient way to compute by simulation quantities that can be expressed as integrals