

Box-Müller Algorithm

115

to generate numbers that are distributed according to a gaussian.

Based on how the gaussian is integrated.

1 dimension \rightarrow 2 dimensions.

$$p(t) = \frac{1}{\sqrt{2\pi}} e^{-t^2/2}$$

$\pm \infty$

$$\int_{-\infty}^{+\infty} p(t) dt = ?$$

1D \rightarrow 2D:

$$+ \infty \quad + \infty$$

$$\int_{-\infty}^{+\infty} p(t_1) dt_1 \int_{-\infty}^{+\infty} p(t_2) dt_2$$

$$= \iint_{-\infty}^{+\infty} dt_1 dt_2 \frac{e^{-(t_1^2 + t_2^2)/2}}{\sqrt{2\pi}}$$

Change of variable:

$$\underbrace{(t_1, t_2)}_{\alpha} \rightarrow \underbrace{(r, \phi)}_{\alpha} \text{ Polar coordinates.}$$

↳ Cartesian coordinates

$$p(t_1) p(t_2) dt_1 dt_2 = \frac{1}{2\pi} e^{-r^2/2} r dr d\phi$$

$$0 \leq \phi < 2\pi$$

$$0 \leq r < \infty$$

$$\begin{aligned} & \int_0^{2\pi} \int_0^\infty r dr e^{-r^2/2} \\ &= \int_0^\infty \frac{1}{2} dr^2 e^{-r^2/2} \\ &= \int_0^\infty dx e^{-x} = 1. \end{aligned}$$

+ \curvearrowleft

$$\Rightarrow \int_{-\infty}^{\infty} dt p(t) = 1.$$

We use the same idea to transform the gaussian into a distribution whose integral may be diverged.

$$p(t_1) dt_1, p(t_2) dt_2$$

$$= p_\phi(\phi) p_n(n) d\phi dn$$

$$p_\phi(\phi) = \frac{1}{2\pi} ; \quad 0 \leq \phi < 2\pi$$

$$\underbrace{p_n(n)}_{=} = n e^{-n^2/2} \quad 0 \leq n < \infty.$$

Cumulative distribution:

$$P_n(n) = \int_0^n i' di' e^{-i'^2/2} = 1 - e^{-n^2/2}.$$

Inverting this:

$$n(P_n) = \sqrt{-2 \ln(1 - P_n)}$$

Hence,

Generate

$$\left. \begin{array}{l} 0 < z < 1 \\ 0 < n < 1 \end{array} \right\}$$

both from a flat distribution on the unit interval (i.e., use the random number generator).

Then calculate

$$t_1 = \sqrt{-2 \ln \xi} \cos 2\pi\eta \quad \}$$

$$t_2 = \sqrt{-2 \ln \xi} \sin 2\pi\eta \quad \}$$

t_1 and t_2 are gaussian.

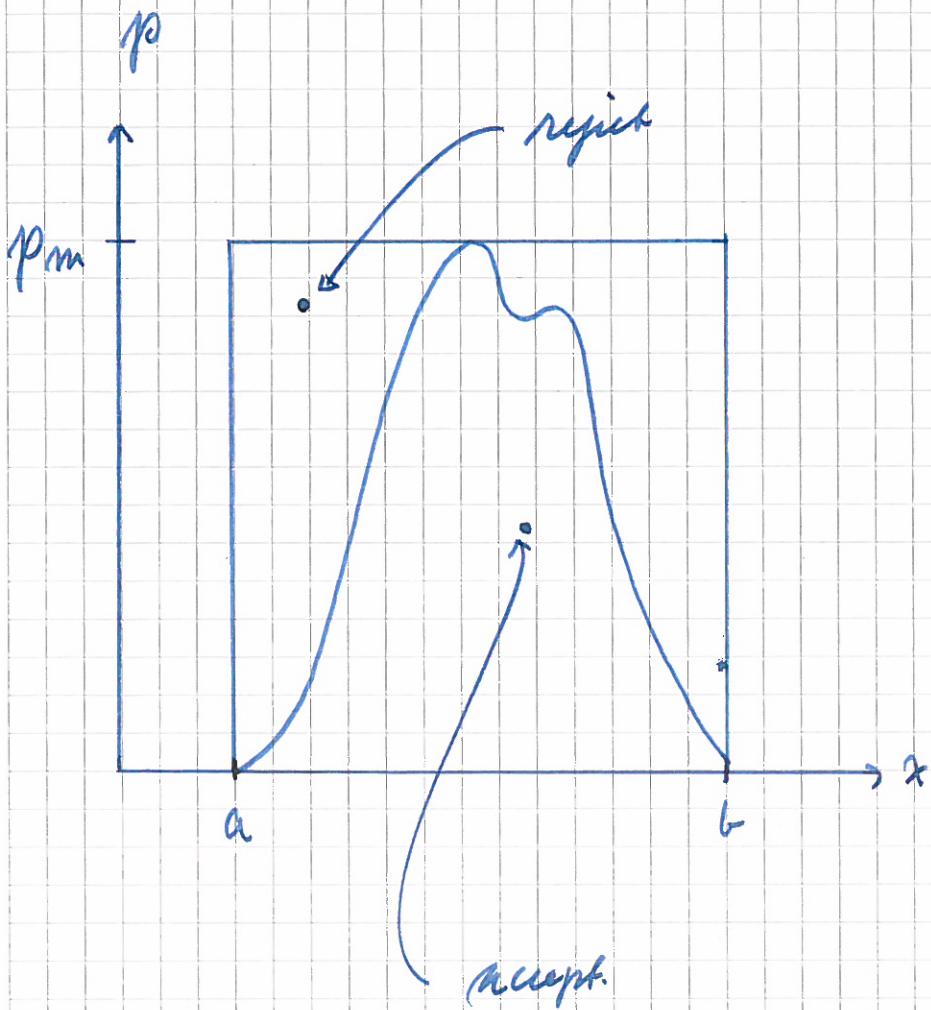
Rejection method for general P.

(from Neumann, 1947)

Probability distribution $p(x)$.

$$a < x < b$$

$$0 < p < p_m$$



Generate (x_i, y_i)

$$a < x_i < b$$

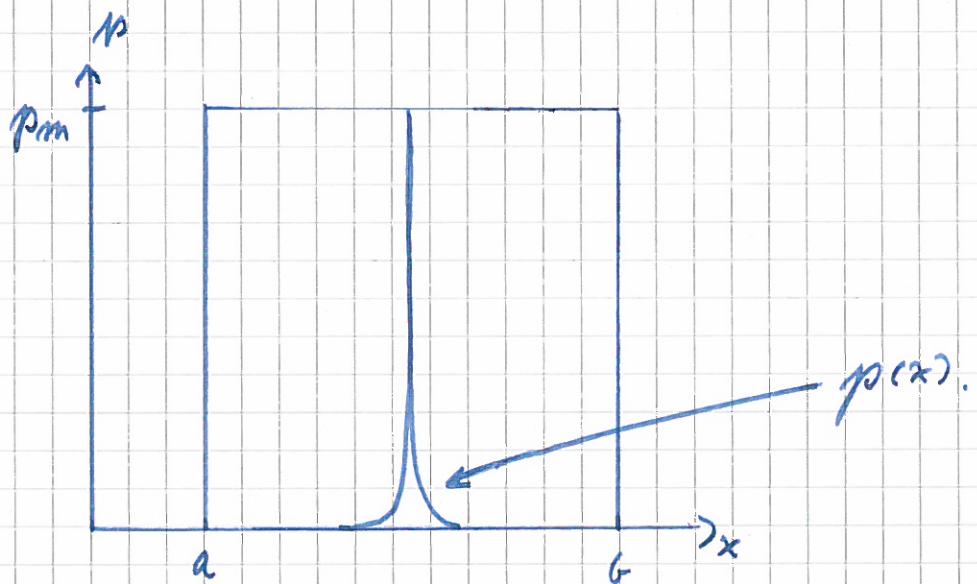
$$0 < y_i < p_m$$

from a flat distribution.

If $y_i \leq p(x_i)$, accept x_i .

If $y_i > p(x_i)$, reject x_i .

This method is ineffective if the area under $p(x)$ is much less than that in the bin $(b-a) \times p_m$:



Improved method:

Find a test function $f(x)$ that resembles $p(x)$, and which is such that $f(x) \geq p(x)$ for $a \leq x \leq b$.

Assume that $F(x) = \int_a^x f(x') dx'$

may be inverted.

① Generate a random number $x \in [a, b]$
from the distribution $f(x)$
($x = F^{-1}(z)$).

② Generate a second random number
 $y \in [0, f(x)]$ from a flat
distribution.

If $y \leq p(x)$, accept x

If $y > p(x)$, reject x .

multivariate gaussian distributions.

$$\vec{x} = (x_1, \dots, x_N)$$

$$\langle x_i \rangle = 0$$

Covariance: $\langle x_i x_j \rangle$.

We define the covariance matrix:

$$\overleftrightarrow{S} = \begin{pmatrix} S_{11} & S_{12} & S_{13} & \dots \\ S_{21} & S_{22} & S_{23} & \dots \\ S_{31} & S_{32} & S_{33} & \dots \\ \vdots & & & \end{pmatrix}, \quad S_{ij} = \langle x_i, x_j \rangle$$

The inverse of the covariance matrix
we denote

$$\overleftrightarrow{G} = \overleftrightarrow{S}^{-1} = \begin{pmatrix} G_{11} & G_{12} & G_{13} & \dots \\ G_{21} & G_{22} & G_{23} & \dots \\ G_{31} & G_{32} & G_{33} & \dots \\ \vdots & \vdots & \vdots & \vdots \end{pmatrix}$$

A multi variable gaussian
distribution is then

$$f(\vec{x}) = \frac{1}{\sqrt{(2\pi)^N \det \overleftrightarrow{S}}} e^{-\frac{1}{2} \vec{x}^T \overleftrightarrow{G} \vec{x}}$$

If \hat{S} is diagonal:

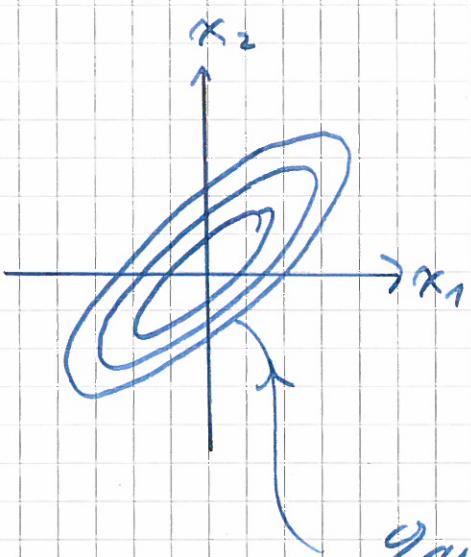
$$p(\vec{x}) = \prod_{i=1}^N \frac{1}{\sqrt{2\pi} s_{ii}} e^{-\frac{1}{2} G_{ii} x_i^2}$$

Product of N
uncorrelated Gaussians.

In general:

Quadratic form

$$Q = \vec{x}^T \tilde{G} \cdot \vec{x}$$



Equipotential curves.

Linear transformation: $\vec{x} = \tilde{T} \cdot \vec{y}$

where \tilde{T} is an orthogonal

matrix (rotation): $\tilde{T} \cdot \tilde{T}^T = \mathbb{I}$.

$$Q = x^T \tilde{G} \cdot \tilde{x} = \tilde{y}^T \{ \tilde{T}^T \tilde{G} \tilde{T} \} \tilde{y}$$

$$= \sum_{i=1}^N G_{ii} y_i^2$$

After transformation, \tilde{G}'
is diagonal.

The same matrix diagonalizes the
covariance matrix:

$$\tilde{S} = \tilde{G}'^{-1}$$

$$\tilde{T}^T \tilde{S} \cdot \tilde{T} = \tilde{T}^{-1} \tilde{S} \tilde{T}$$

$$= (\tilde{T}^{-1} \tilde{S}^{-1} \tilde{T})^{-1} = (\tilde{T}^T \tilde{G} \tilde{T})^{-1}$$

Algorithm

- * Covariance matrix (or its inverse) is given.
- * Use the main axis transformation to construct \tilde{T} .
- * Find the eigenvalues of \tilde{S} :
 σ_i^2 are these eigenvalues.
- * Generate N independent Gaussian random numbers y_i with variance σ_i^2 .
- * Transform $\vec{x} = \tilde{T} \cdot \vec{y}$.

Main axis transformation

Determine eigenvectors ($= \vec{s}_i$)
 and eigenvalues ($= \sigma_i^2$) of
 \hat{S} by using a canned
routine.

Normalize \vec{s}_i : $|\vec{s}_i| = 1$.

Combine N column vectors \vec{s}_i
 to form \hat{T} .

Generating random points on
 a unit sphere in cartesian
 coordinates.

How not to do this:

Generate $\phi \in [0, 2\pi]$.

Generate

$$\begin{cases} x = \cos \phi \\ y = \sin \phi \end{cases}$$

How to do it:

Generate

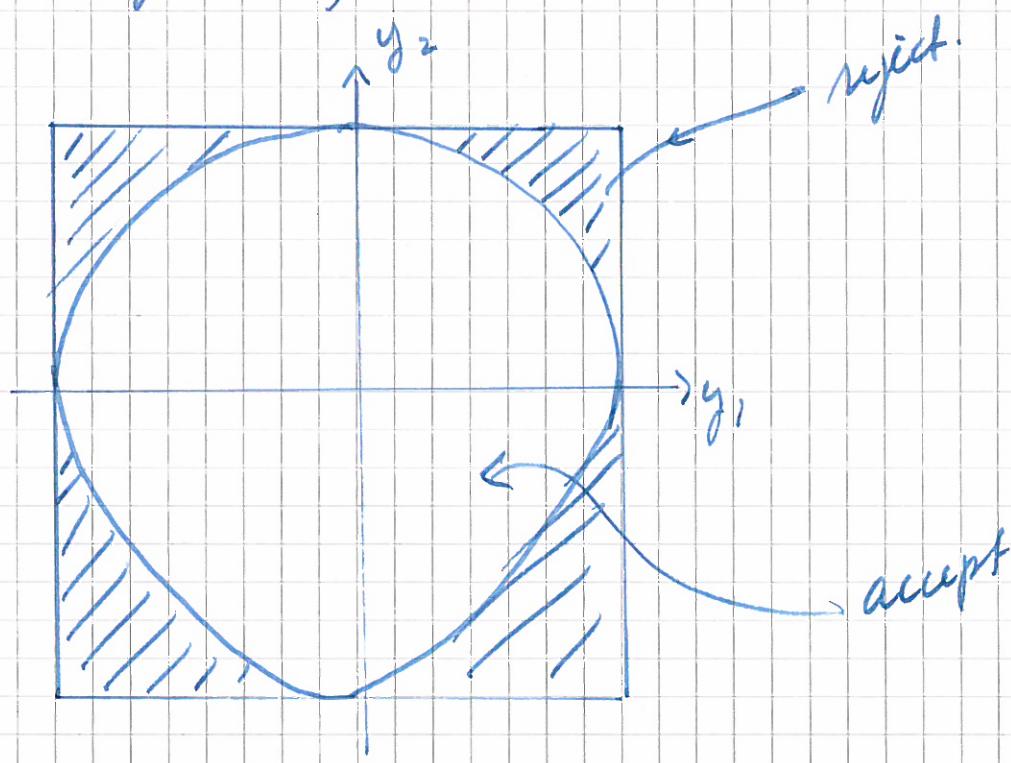
$$y_1 \in [-1, 1]$$

$$y_2 \in [-1, 1]$$

Calculate

$$r^2 = y_1^2 + y_2^2$$

Reject if $r^2 > 1$.



If accept, calculate

$$\begin{aligned} x_1 &= y_1/r \\ x_2 &= y_2/r \end{aligned} \quad \left. \right\}$$

We need to calculate a square root, but this is much cheaper than doing two trigonometric functions.

In 3 dimensions (sphere):

Marsaglia algorithm

Generate $y_1 \in [-1, 1]$

$y_2 \in [-1, 1]$

Reject if $y_1^2 + y_2^2 \geq 1$.

If $y_1^2 + y_2^2 < 1$, we have

$$x_1 = 2y_1 \sqrt{1-n^2}$$

$$x_2 = 2y_2 \sqrt{1-n^2}$$

$$x_3 = 1 - 2n^2$$

(Check that

$$x_1^2 + x_2^2 + x_3^2 = 1$$

Sequences of random numbers.

$$x_1, x_2, x_3, \dots, x_n, \dots$$

How to construct correlations
in sequences of random
numbers.

$$\text{So far } \langle x_m x_{m+k} \rangle = \langle x_m \rangle \langle x_{m+k} \rangle.$$

Stochastic processes

So far stochastic numbers.

We generalize to

stochastic functions.

A given function $x(t)$ drawn from the ensemble $\{x(t)\}$ is a realization of the stochastic process.

130

Cumulative probability:

$$P_1(x; t) = \text{Prob}\{x(t) \leq x\}.$$

Probability density:

$$\rho_1(x; t) = \frac{d}{dx} P_1(x; t).$$

We generalize these two concepts:

$$P_1(x; t) \rightarrow P_m(x_1, \dots, x_m; t_1, \dots, t_m).$$

$$= \text{Prob}\{X(t_1) \leq x_1, X(t_2) \leq x_2, \dots, X(t_m) \leq x_m\}.$$

Corresponding probability density:

$$\rho_m(x_1, \dots, x_m; t_1, \dots, t_m) = \frac{\partial^m p_m}{\partial x_1 \dots \partial x_m}.$$

elements

$$\langle x^m(t) \rangle = \int_{-\infty}^{+\infty} dx x^m \rho_1(x; t)$$



$$\langle x^{m_1}(t_1) x^{m_2}(t_2) \dots x^{m_m}(t_m) \rangle$$

$$= \int_{-\infty}^{+\infty} \dots \int_{-\infty}^{+\infty} dx_1 \dots d x_m x_1^{m_1} \dots x_m^{m_m}$$

$$\rho_m(x_1, \dots, x_m; t_1, \dots, t_m)$$

Probability functions can be reconstructed from knowledge of the moments.

Statistical Stationarity:

$$P_m(x_1, \dots, x_n; t_1, \dots, t_n)$$

$$= P_m(x_1, \dots, x_n; t_1 + \epsilon, \dots, t_n + \epsilon)$$

$\uparrow \qquad \qquad \qquad \uparrow$

Shift in time.

A special case: $P_2(x_1, x_2; t_1, t_2)$

Shift time by $-t_1$:

$$P_2(x_1, x_2; t_1, t_2) = P_2(x_1, x_2; 0, \underbrace{t_2 - t_1}_{\uparrow})$$

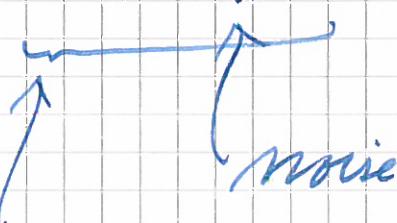
Only dependence
on time
difference.

In general, a stochastic process is
stationary to order k if P_1 to
 P_k are stationary.

Example:

(Diffusion process:

$$\dot{x}(t) = \eta(t)$$



Langevin equation \rightarrow

Johker - Planck (or diffusion)
equation:

$$\frac{\partial}{\partial t} p_i(x; t) = \frac{D}{2} \frac{\partial^2}{\partial x^2} p_i(x; t)$$

Initial condition:

$$p_i(x; 0) = \delta(x).$$

Solution:

$$p_i(x; t) = \frac{e^{-x^2/2Dt}}{\sqrt{2\pi Dt}}$$

$x(t)$ is a realization of the stochastic
process described by $p_i(x; t)$.

134

This process is not stationary:

$$p_1(x; t + \tau) \neq p_1(x; t).$$

The auto correlation function:

$$\langle X(t) X(t + \tau) \rangle$$

$$= \int_{-\infty}^{+\infty} dx_1 \int_{-\infty}^{+\infty} dx_2 X_1 X_2$$

$$p_2(x_1, x_2; t_1, t_2)$$

In a stationary process.

$$\langle X(t) X(t + \tau) \rangle = \langle X(0) X(\tau) \rangle.$$

How fast does a process forget a given value x attained at time t ?

Conditional probability:

$$p(x_2/x_1; t_1, t_2)$$

$$= \frac{p_{x_2}(x_2, x_1; t_1, t_2)}{p_{x_1}(x_1; t_1)}$$

Probability that $x = x_2$ at time t_2
given that $x = x_1$ at time t_1 .

Conditional moment:

$$\langle x(t+\tau) | x(t) \rangle$$

$\stackrel{+ \infty}{\rightarrow}$

$$= \int_{-\infty}^{+\infty} dx_2 x_2 p(x_2/x_1; t_1, t+\tau)$$

When $\tau \rightarrow \infty$ and mean is stationary
at first order:

$$\langle x(t+\tau) | x(t) \rangle \rightarrow \langle x \rangle.$$

ExampleStationary gaussian process:

$$p_1(x) = \frac{1}{\sqrt{2\pi\sigma^2}} e^{-x^2/2\sigma^2}$$

$$p_2(x_1, x_2; 0, \tau)$$

$$= \frac{1}{\sqrt{(2\pi)^2 \det(\tilde{S}_2(\tau))}} e^{-\frac{1}{2} \vec{x}^T \tilde{G}_2(\tau) \vec{x}}$$

where $\tilde{G}_2 = \tilde{S}_2^{-1}$ and

$$\tilde{S}_2(\tau) = \begin{pmatrix} \langle x(0)^2 \rangle & \langle x(0)x(\tau) \rangle \\ \langle x(0)x(\tau) \rangle & \langle x(\tau)^2 \rangle \end{pmatrix}$$

$$\langle x(0)^2 \rangle = \langle x(\tau)^2 \rangle = \langle x^2 \rangle.$$

$$\vec{x}^T \tilde{G}_2(\tau) \vec{x} = \frac{1}{\det(\tilde{S}_2(\tau))} (\langle x^2 \rangle x_1^2$$

$$- 2 \langle x(0)x(\tau) \rangle x_1 x_2 + \langle x^2 \rangle x_2^2)$$

In general :

$$P_m(x_1, \dots, x_n; t_1, \dots, t_n)$$

$$= \frac{1}{\sqrt{(2\pi)^n \det(\tilde{S}_m)}} e^{-\frac{1}{2} \tilde{x}^T \tilde{G}_m^{-1} \tilde{x}}$$

where $\tilde{G}_m = \tilde{S}_m^{-1}$ and

$$\tilde{S}_m = \{ \langle x(t_i) x(t_j) \rangle \}$$

Markov processes

Stochastic sequence

